# U- and V-statistics for Itô Semimartingales

Christian Schmidt

PhD Dissertation
March 17, 2015

Supervisor: Mark Podolskij

Department of Mathematics
Aarhus University
Denmark

# Contents

**Preface** **iii**

**1 Introduction** **1**
- 1.1 Semimartingales . . . 1
- 1.2 U- and V-statistics . . . 3
- 1.3 High Frequency Data and Stable Convergence . . . 4
- Bibliography . . . 6

**A Limit theorems for non-degenerate U-statistics of continuous semimartingales** **9**
- A.1 Introduction . . . 11
- A.2 Preliminaries . . . 12
- A.3 Law of large numbers . . . 13
- A.4 Stable central limit theorem . . . 16
- A.5 Estimation of the conditional variance . . . 27
- A.6 Statistical applications . . . 29
- A.7 Proofs of some technical results . . . 35
- Bibliography . . . 41

**B On U- and V-statistics for discontinuous Itô semimartingales** **45**
- B.1 Introduction . . . 47
- B.2 Preliminaries . . . 49
- B.3 The jump case . . . 50
- B.4 The mixed case . . . 64
- B.5 Appendix . . . 74
- Bibliography . . . 90

# Preface

**Abstract**

Stochastic processes are widely used for modeling and nowadays, in many applications, data is observed very frequently. That makes it necessary to have statistical procedures in order to deal with functionals that are based on these high frequency observations. A particular example comes from finance, where for instance prices of stocks are recorded frequently. Under some type of no-arbitrage conditions it is further known that those price processes must be semimartingales (see [2]).

In this dissertation U- and V-statistics of high frequency observations of such semimartingales are investigated. Those types of statistics are classical objects in mathematical statistics and are widely used in estimation and test theory. In the case that the underlying semimartingale is continuous a functional law of large numbers and central limit theorem are presented. The main tools that are used in the derivation of the theory is the empirical process approach for U-statistics (see [1]) and a general stable central limit theorem for sums of triangular arrays (see [3]). The results are then generalized to the case where the semimartingale is discontinuous. There however, the asymptotic theory is more sophisticated because it heavily depends on the kernel function of the U-statistic. This leads to several different types of limiting processes.

**Resumé**

Stokastiske processer anvendes i mange sammenhænge til modellering af tilfældige fænomener over tid og i mange nutidige anvendelser observeres data ofte med høj frekvens. Dette nødvendiggører statistiske procedurer, som kan behandle funktioner af disse højfrekvente data. Prisen påaktier kan nævnes som et konkret eksempel påsådanne data. Det er yderligere kendt, at under visse antagelser om ingen arbitrage, mådisse priser være semimartingaler (se [2]).

I denne afhandling studeres U- og V-statistik for højfrekvente observation fra disse semimartingaler. Disse statistikker er klassiske elementer i matematisk statistik og anvendes ofte inden for estimations- og testteori. Der vil blive præsenteret en store tals lov og en central grænseværdisætning for funktionaler i tilfældet, hvor den underliggende semimartingal er kontinuert. Hovedredskabet som anvendes til udledningen af teorien er den empiriske process tilgang for U-statistikker (se [1]) og et generel stabilt central grænseværdisætning for summer af uafhængige trekantskemaer (se [3]). Disse resultater generaliseres til situationen, hvor semimartingalen er diskontinuert. I denne situation bliver den asymptotiske teori mere sofistikeret, da den i høj grad afhænger af kernefunktion af U-statistikken. Dette leder til flere forskellige typer af grænseprocesser.

**Acknowledgments**

I wish to thank my supervisor Mark Podolskij for introducing me to the subject and all the input he gave to me. Also I am grateful for all the meetings we had and for many fruitful discussions. I would further like to thank all the people in the departments of mathematics at both Heidelberg and Aarhus University for providing an inspiring atmosphere. Special Thanks also to my current and former office mates Claudio Heinrich and Sophon Tunyavetchakit. Finally I would like to thank my family for the constant support during the time of my PhD studies.

## Bibliography

[1] E. Beutner and H. Zähle. Deriving the asymptotic distribution of U- and V-statistics of dependent data using weighted empirical processes. *Bernoulli*, 18(3):803–822, 2012.

[2] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Math. Ann.*, 300(1):463–520, 1994.

[3] J. Jacod. On continuous conditional gaussian martingales and stable convergence in law. *Sémin. probab. Strasbg*, XXXI:232–246, 1997.

# Introduction

The purpose of this section is to give an introduction to the main objects and problems that appear in this dissertation. In the first section we give a short overview on semimartingales. We further discuss Itô semimartingales, which is a certain subclass of semimartingales, and provide the so-called Grigelionis decomposition for this type of processes. In the second section we introduce the notion of U- and V-statistics and briefly describe the empirical process approach, which is used in this thesis in order to derive an asymptotic theory. In the last section we define what we mean by U- and V- statistics for semimartingales in the high frequency setting. Further we indicate what the main results of this dissertation are and how they are related to the literature.

## 1.1 Semimartingales

In this section we will introduce the class of processes that is studied in this dissertation. For this let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space that fulfills the usual assumptions. A stochastic process $(M_t)_{t\geq 0}$ that is defined on this space is called a local martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if there exists an increasing sequence $(T_n)_{n\in\mathbb{N}}$ of stopping times such that $T_n \to \infty$ almost surely and such that the process $(M_{t\wedge T_n})_{t\geq 0}$ is a martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$ for each $n \in \mathbb{N}$. We say that a function $g : \mathbb{R} \to \mathbb{R}$ is of finite variation on the interval $[0, t]$ for $t > 0$ if

$$\sup_{\mathcal{P}} \sum_{i=1}^{n} |g(t_i) - g(t_{i-1})| < \infty,$$

where $n \in \mathbb{N}$ and the supremum runs over all partitions $\mathcal{P} = \{t_0, \dots, t_n\}$ of $[0, t]$ with $0 = t_0 < t_1 < \cdots < t_n = t$.

A process $(X_t)_{t\geq 0}$ is called semimartingale if it has a decomposition

$$X_t = X_0 + M_t + A_t, \quad t \geq 0,$$

where $X_0$ is $\mathcal{F}_0$-measurable, $(M_t)_{t\geq 0}$ is a local martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$ and $(A_t)_{t\geq 0}$ is adapted, càdlàg and pathwise of finite variation on $[0, t]$ for all $t > 0$. We further require that $M_0 = A_0 = 0$.

We are interested in a certain type of semimartingales, the so-called Itô semimartingales. In order to state their definition we need to introduce the characteristics of a semimartingale. To do this we first define the jump measure $\mu^X$ that is associated to $X$ by

$$\mu^X(\omega, dt, dx) = \sum_s \mathbb{1}_{\{\Delta X_s \neq 0\}} \epsilon_{(s,\Delta X_s(\omega))}(dt, dx),$$

where $\epsilon_u$ is the Dirac measure that puts mass 1 to $u \in \mathbb{R}_+ \times \mathbb{R}$ and $\Delta X_s = X_s - X_{s-}$ stands for the jump size of $X$ at time $s$. We denote the compensator of $\mu^X$ by $\nu^X$. In terms of those random measures the semimartingale $X$ has a decomposition

$$X = X_0 + B + X^c + (x\mathbb{1}_{\{|x|\leq 1\}}) * (\mu^X - \nu^X) + (x\mathbb{1}_{\{|x|>1\}}) * \mu^X,$$

which is known as Lévy-Itô decomposition (see [14]). Here $X^c$ is a continuous local martingale and $B$ a process of locally finite variation. The last two terms in the decomposition stand for stochastic integrals. For a random measure $\gamma$ and an optional function $f$ on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$ the stochastic integral process $f * \gamma$ is defined by

$$\int_{[0,t]\times\mathbb{R}} f(\omega, s, x)\gamma(\omega, ds, dx),$$

whenever the integral exists. We remark that

$$(x\mathbb{1}_{\{|x|>1\}}) * \mu_t^X = \sum_{s\leq t} |\Delta X_s| \mathbb{1}_{\{|\Delta X_s|>1\}}$$

is the sum of the jumps of size bigger than 1, which are only finitely many. Since the jumps of a semimartingale do not need to be absolutely summable, the small jumps are compensated and expressed in the (purely discontinuous) martingale $(x\mathbb{1}_{\{|x|\leq 1\}}) * (\mu^X - \nu^X)$. Given the Lévy-Itô decomposition the characteristics of the semimartingale $X$ are defined to be the triplet $(B, C, \nu^X)$, where $B$ and $\nu^X$ are given above and $C$ is the quadratic variation of $X^c$. When those characteristics are absolutely continuous with respect to the Lebesgue measure, i.e.

$$B_t = \int_0^t b_s ds, \quad C_t = \int_0^t \sigma_s^2 ds, \quad \nu^X(\omega, dt, dx) = dt F_t(\omega, dx),$$

where $(b_s)_{s\geq 0}$ and $(\sigma_s)_{s\geq 0}$ are stochastic processes and $F_t(\omega, dx)$ is a measure on $\mathbb{R}$ for fixed $(\omega, t)$, then $X$ is called Itô semimartingale. In particular Lévy processes are also Itô semimartingales. Because of this special structure Itô semimartingales can be written in the form

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + (x\mathbb{1}_{\{|x|\leq 1\}}) * (\mu^X - \nu^X)_t + (x\mathbb{1}_{\{|x|>1\}}) * \mu_t^X,$$

where $W$ is a standard Brownian motion. For some technical reasons a slightly different decomposition will be used. It is generally possible (see [13]) to replace the jump measure $\mu^X$ by a Poisson random measure $\mathfrak{p}$ and the compensator $\nu^X$ by the compensator $\mathfrak{q}$ of the Poisson random measure, which is given by $\mathfrak{q}(dt, dz) = dt \otimes \lambda(dz)$ for some $\sigma$-finite measure $\lambda$ on $\mathbb{R}$. What changes then is that $x$ has to be replaced by some predictable function $\delta : \Omega \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$. Finally $X$ can be represented as

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + (\delta\mathbb{1}_{\{|\delta|\leq 1\}}) * (\mathfrak{p} - \mathfrak{q})_t + (\delta\mathbb{1}_{\{|\delta|>1\}}) * \mathfrak{p}_t.$$

This decomposition is known as Grigelionis decomposition.

## 1.2 U- and V-statistics

The introduction of U-statistics goes back to Hoeffding in [10], whose work is partly based on a paper by Halmos [8]. The interest in U-statistics is mainly due to its great importance in estimation theory. In the classical setting of i.i.d. data U-statistics form a large class of unbiased estimators for certain parameters of the underlying distribution. More precisely a U-statistic of order $d$ based on a sequence of i.i.d. real-valued random variables $(X_k)_{k\in\mathbb{N}}$ that live on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is defined as

$$U_n = \binom{n}{d}^{-1} \sum_{1\leq i_1<\cdots<i_d\leq n} H(X_{i_1}, \ldots, X_{i_d}),$$

where the kernel function $H : \mathbb{R}^d \to \mathbb{R}$ is generally assumed to be symmetric. Clearly $U_n$ is an unbiased estimator for

$$\theta = \mathbb{E}(H(X_1, \ldots, X_d)),$$

whenever the expectation exists and is finite. The estimator is also optimal in the sense that it has the smallest variance among all unbiased estimators of $\theta$. Simple examples for quantities that can be estimated by U-statistics are of course the sample mean and the sample variance. In the case $d = 2$ we can use $H(x, y) = |x-y|$ as a kernel function, which is known as Gini's mean difference and serves as a measure of concentration of the distribution of $X_1$. Another famous example is the so-called Wilcoxon statistic, which uses $H(x, y) = \mathbb{1}_{\{x\leq y\}}$ as a kernel function (see [17]). In manuscript A we will come back to those examples in our setting.

Closely related to U-statistics are V-statistics, which were first investigated by von Mises [18]. By a V-statistic of order $d$ we mean a statistic of the type

$$V_n = \frac{1}{n^d} \sum_{1\leq i_1,\ldots,i_d\leq n} H(X_{i_1}, \ldots, X_{i_d}),$$

where $H : \mathbb{R}^d \to \mathbb{R}$ is again some kernel function. In some situations (like e.g. in manuscript A) the summands for which not all indices are different, are asymptotically negligible and hence both U- and V-statistics result in the same asymptotic theory. Besides their importance for estimation theory both types of statistics are also used for testing. Since in many situations the exact distributions of $U_n$ or $V_n$ are unknown it is reasonable to develop at least asymptotic tests in such situations which requires to have a law of large numbers and an associated central limit theorem. Even in the given setting of i.i.d. random variables $X_k$ this is not straightforward. A first solution to this problem was given by Hoeffding who used the well-known Hoeffding decomposition. Since then various approaches like Hermite expansion (see [5, 6]) or empirical process theory (see [2]) have been proposed to give an asymptotic theory under weaker assumptions on $(X_k)_{k\in\mathbb{N}}$. For example results are given in the situation where the underlying data inherits some weak dependence ([3, 7]) or long range dependence (see [5, 6]).

For our purpose the empirical process approach seems to be most suitable and we therefore shortly describe what is meant by this. We can always write the a V-statistic as

$$V_n = \int_{\mathbb{R}^d} H(x_1, \ldots, x_d) F_n(dx_1) \ldots F_n(dx_d),$$

where

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{X_i \le x\}}$$

denotes the empirical distribution function of the underlying data. We know that this function converges almost surely in the uniform topology to the distribution function $F$ of $X_1$. Often it possible to deduce

$$V_n \xrightarrow{\mathbb{P}} V = \int_{\mathbb{R}^d} H(x_1, \dots, x_d) F(dx_1) \dots F(dx_d)$$

from this. Assume that we also have a central limit theorem for the empirical distribution function of the form $\sqrt{n}(F_n - F) \xrightarrow{d} G$ for some random variable $G$, which is typically some Brownian bridge in the i.i.d. case. We then use the decomposition

$$\sqrt{n}(V_n - V) = \sum_{k=1}^{d} \int_{\mathbb{R}^d} H(x_1, \dots, x_d) \sqrt{n}(F_n(dx_k) - F(dx_k)) \prod_{i=1}^{k-1} F_n(dx_i) \prod_{i=k+1}^{d} F(dx_i).$$

Knowing the results for the empirical distribution function and assuming the $H$ is symmetric, it is then sometimes possible to deduce

$$\sqrt{n}(V_n - V) \xrightarrow{d} d \int_{\mathbb{R}^d} H(x_1, \dots, x_d) G(dx_1) F(dx_2) \dots F(dx_d)$$

for the V-statistic.

## 1.3 High Frequency Data and Stable Convergence

We generally assume that we observe an Itô semimartingale $X$ on a finite time interval $[0, T]$ at high frequency. This means we observe the data

$$X_{i/n}, \quad i = 0, \dots, \lfloor nT \rfloor$$

and we are interested in the case $n \to \infty$, which is also known as infill asymptotics. In the literature many statistics have been investigated in this setting. Two of the most important ones are given by

$$V(f, X)_t^n = \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} f(\sqrt{n} \Delta_i^n X), \tag{1.1}$$

$$V'(f, X)_t^n = \sum_{i=1}^{\lfloor nt \rfloor} f(\Delta_i^n X), \tag{1.2}$$

where $f : \mathbb{R} \to \mathbb{R}$ is some sufficiently smooth function, $0 \le t \le T$ and $\Delta_i^n X = X_{i/n} - X_{(i-1)/n}$. An asymptotic theory (law of large numbers and central limit theorem) for those statistics under various assumptions is provided by Jacod in [12]. Generally $V(f, X)_t^n$ is used if $X$ is continuous because the order of $\Delta_i^n X$ is $n^{-1/2}$ and hence $\sqrt{n}$ is the right scaling in the function $f$. For discontinuous $X$ the analysis is more complicated. This is mainly due to fact that increments of the jump and continuous part of $X$ are of different order and hence the choice of the statistics depends on the function $f$. For instance, let $f(x) = |x|^p$. If $p > 2$ one uses $V'(f,X)_t^n$ and only the jump part of $X$ appears in the limit. For $p < 2$ one considers $V(f,X)_t^n$ and the limit is dominated by the continuous part of $X$. If $p = 2$ both statistics coincide and both the continuous and the jump part of $X$ appear in the limit.

In the literature many other statistics of similar type have been investigated. See for example [1] for bipower variation or [15] for truncated power variation, which basically deal with the estimation of certain functionals of the volatility $(\sigma_s)_{s\geq 0}$ in the presence of jumps. In [9] the authors considers statistics that are based on multidimensional Itô semimartingales that are observed at high frequency, but with non-synchronous observations times. Many of those cases were by now also extended to the situation that only a noisy version of $X$ is observed (see e.g. [16, 4]).

What essentially all the statistics have in common is that the distribution of the limiting process in the central limit theorem is, at least in all cases that are relevant for applications, mixed normal. This means in general we have limit theorems of the type $L_n \to L$, where the random variable $L$ is mixed normal, that is it can be written as $L = UV$, where $U > 0$ and $V$ are independent random variables and $V \sim \mathcal{N}(0,1)$. In general the law of $U$ is unknown, what makes the central limit theorem infeasible in the given form. Usually it is however possible to consistently estimate $U$ by some estimator $U_n$. If the mode of convergence in the central limit theorem is only weak convergence, then, of course, we cannot automatically deduce the feasible standard central limit theorem $L_n/U_n \xrightarrow{d} \mathcal{N}(0,1)$. This implication is however true if the convergence $L_n \to L$ is stable.

We say that a sequence $(Z_n)_{n\in\mathbb{N}}$ of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in a Polish space $(E, \mathcal{E})$ converges stably in law to a random variable $Z$ that is defined on an extension $(\Omega', \mathcal{F}', \mathbb{P}')$ of $(\Omega, \mathcal{F}, \mathbb{P})$ and takes also values in $(E, \mathcal{E})$, if and only if

$$\mathbb{E}(f(Z_n)Y) \to \mathbb{E}'(f(Z)Y) \quad \text{as} \quad n \to \infty$$

for all bounded and continuous $f$ and bounded, $\mathcal{F}$-measurable $Y$. We write $Z_n \xrightarrow{st} Z$ for stable convergence of $Z_n$ to $Z$. The main theorem that is used in the high frequency setting in order to proof stable central limit theorems is provided by Jacod in [11].

The aim of this thesis is to develop an asymptotic theory for U- and V-statistics when the underlying data comes from an Itô semimartingale $X$ that is observed at high frequency. More precisely we are interested in the asymptotic theory of statistics of the type

$$\frac{1}{n^l} \sum_{1\leq i_1,\ldots,i_d\leq \lfloor nt\rfloor} H(\sqrt{n}\Delta_{i_1}^n X, \ldots \sqrt{n}\Delta_{i_l}^n X, \Delta_{i_{l+1}}^n X, \ldots, \Delta_{i_d}^n X)$$

for some $0 \leq l \leq d$. Clearly this generalizes results from [12] for the statistics defined at (1.1) and (1.2). The main problem here is to deal with the dependency structure of the multiple increments, which makes it impossible to directly apply the techniques that can be found in the literature.

In manuscript A we deal with the case that $l = d$ and $X$ is a continuous Itô semimartingale. The method employed in the proofs is a combination of techniques that were developed in the high frequency setting and the empirical process approach for U- and V- statistics that was discussed in the previous section. In manuscript B the underlying process is a discontinuous Itô semimartingale and the asymptotic theory developed in manuscript A for $l = d$ is generalized to this setting. The case $l = 0$, in which the limiting distribution is dominated by the jump part, is also treated separately. Afterwards both results are combined to obtain a law of large numbers and a central limit theorem for general $l$.

## Bibliography


[1] O. E. Barndorff-Nielsen, O. E. Graversen, J. Jacod, M. Podolskij, and N. Shephard. A central limit theorem for realised power and bipower variations of continuous semimartingales. In Yu. Kabanov, R. Liptser, and J. Stoyanov, editors, *From Stochastic Calculus to Mathematical Finance. Festschrift in Honour of A.N. Shiryaev*, pages 33–68. Springer, Heidelberg, 2006.

[2] E. Beutner and H. Zähle. Deriving the asymptotic distribution of U- and V-statistics of dependent data using weighted empirical processes. *Bernoulli*, 18(3):803–822, 2012.

[3] S. Borovkova, R. Burton, and H. Dehling. Limit theorems for functionals of mixing processes with applications to u-statistics and dimension estimation. *Trans. Amer. Math. Soc.*, 353:4261–4318, 2001.

[4] K. Christensen, M. Podolskij, and M. Vetter. On covariation estimation for multivariate continuous itô semimartingales with noise in non-synchronous observation schemes. *Journal of Multivariate Analysis*, 120:59–84, 2013.

[5] H. Dehling and M.S. Taqqu. The empirical process of some long-range dependent sequences with an application to U-statistics. *Ann. Statist.*, 17(4):1767–1783, 1989.

[6] H. Dehling and M.S. Taqqu. Bivariate symmetric statistics of long-range dependent observations. *J. Statist. Plann. Inference*, 28:153–165, 1991.

[7] M. Denker and G. Keller. On U-statistics and v. miseâĂŹ statistics for weakly dependent processes. *Z. Wahrsch. Verw. Gebiete*, 64(4):505–522, 1983.

[8] P.R. Halmos. The theory of unbiased estimation. *Ann. Math. Statist.*, 17(1):34–43, 1946.

[9] T. Hayashi and N. Yoshida. On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli*, 11(2):359–379, 2005.

[10] W. Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.*, 19:293–325, 1948.

[11] J. Jacod. On continuous conditional gaussian martingales and stable convergence in law. *Sémin. probab. Strasbg*, XXXI:232–246, 1997.

[12] J. Jacod. Asymptotic properties of realized power variations and related functionals of semimartingales. *Stoch. Process. Appl.*, 118:517–559, 2008.

[13] J. Jacod and P. Protter. *Discretization of Processes.* Springer, Berlin, 2012.

[14] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer, Berlin, 2nd edition, 2003.

[15] C. Mancini. Disentangling the jumps of the diffusion in a geometric jumping brownian motion. *Giornale dell'Instituto Italiano degli Attuari*, LXIV:19–47, 2001.

[16] M. Podolskij and M. Vetter. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, 15(3):634–658, 2009.

[17] R.J. Serfling. *Approximation Theorems of Mathematical Statistics.* Wiley, New York, 1980.

[18] R. v. Mises. On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.*, 18(3):309–348, 1947.

Paper A

# Limit theorems for non-degenerate U-statistics of continuous semimartingales

By Mark Podolskij, Christian Schmidt and Johanna F. Ziegel

**Abstract**

This paper presents the asymptotic theory for non-degenerate U-statistics of high frequency observations of continuous Itô semimartingales. We prove uniform convergence in probability and show a functional stable central limit theorem for the standardized version of the U-statistic. The limiting process in the central limit theorem turns out to be conditionally Gaussian with mean zero. Finally, we indicate potential statistical applications of our probabilistic results.

**Keywords:** high frequency data, limit theorems, semimartingales, stable convergence, U-statistics.

**AMS Subject Classification:** Primary 60F05, 60F15, 60F17; secondary 60G48, 60H05.

## A.1 Introduction

Since the seminal work by Hoeffding [15], U-statistics have been widely investigated by probabilists and statisticians. Nowadays, there exists a vast amount of literature on the asymptotic properties of U-statistics in the case of independent and identically distributed (i.i.d.) random variables or in the framework of weak dependence. We refer to [23] for a comprehensive account of the asymptotic theory in the classical setting. The papers [4, 5, 11] treat limit theorems for U-statistics under various mixing conditions, while the corresponding theory for long memory processes has been studied for example in [9, 14]; see [16] for a recent review of the properties of U-statistics in various settings. The most powerful tools for proving asymptotic results for U-statistics include the classical Hoeffding decomposition (see e.g. [15]), Hermite expansions (see e.g. [9, 10]), and the empirical process approach (see e.g. [3]). Despite the activity of this field of research, U-statistics for high frequency observations of a time-continuous process have not been studied in the literature thus far. The notion of high frequency data refers to the sampling scheme in which the time step between two consecutive observations converges to zero while the time span remains fixed. This concept is also known under the name of infill asymptotics. Motivated by the prominent role of semimartingales in mathematical finance, in this paper we present novel asymptotic results for high frequency observations of Itô semimartingales and demonstrate some statistical applications.

The seminal work of Jacod [17] marks the starting point for stable limit theorems for semimartingales. Stimulated by the increasing popularity of semimartingales as natural models for asset pricing, the asymptotic theory for partial sums processes of continuous and discontinuous Itô semimartingales has been developed in [2, 18, 22]; see also the recent book [20]. We refer to [25] for a short survey of limit theorems for semimartingales. More recently, asymptotic theory for Itô semimartingales observed with errors has been investigated in [19].

The methodology we employ to derive a limit theory for U-statistics of continuous Itô semimartingales is an intricate combination and extension of some of the techniques developed in the series of papers mentioned in the previous paragraph, and the empirical process approach to U-statistics.

In this paper we consider a one-dimensional continuous Itô semimartingale of the form

$$X_t = x + \int_0^t a_s ds + \int_0^t \sigma_s dW_s, \qquad t \geq 0,$$

defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ (which satisfies the usual assumptions), where $x \in \mathbb{R}$, $(a_s)_{s\geq 0}$, $(\sigma_s)_{s\geq 0}$ are stochastic processes, and $W$ is a standard Brownian motion. The underlying observations of $X$ are

$$X_{\frac{i}{n}}, \qquad i = 0, \dots, [nt],$$

and we are in the framework of infill asymptotics, i.e. $n \to \infty$. In order to present our main results we introduce some notation. We define

$$\mathcal{A}_t^n(d) := \{\boldsymbol{i} = (i_1, \dots, i_d) \in \mathbb{N}^d \mid 1 \leq i_1 < i_2 < \dots < i_d \leq [nt]\},$$

$$Z_{\boldsymbol{s}} := (Z_{s_1}, \dots, Z_{s_d}), \qquad \boldsymbol{s} \in \mathbb{R}^d,$$

where $Z = (Z_t)_{t\in\mathbb{R}}$ is an arbitrary stochastic process. For any continuous function $H : \mathbb{R}^d \to \mathbb{R}$, we define the U-statistic $U(H)^n_t$ of order $d$ as

$$U(H)^n_t = \binom{n}{d}^{-1} \sum_{\boldsymbol{i}\in\mathcal{A}^n_t(d)} H(\sqrt{n}\Delta^n_{\boldsymbol{i}}X) \tag{A.1}$$

with $\Delta^n_{\boldsymbol{i}}X = X_{\boldsymbol{i}/n} - X_{(\boldsymbol{i}-1)/n}$. For a multi-index $\boldsymbol{i} \in \mathbb{N}^d$, the vector $\boldsymbol{i}-1$ denotes the multi-index obtained by componentwise subtraction of 1 from $\boldsymbol{i}$. In the following we assume that the function $H$ is symmetric, i.e. for all $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ and all permutations $\pi$ of $\{1, \dots, d\}$ it holds that $H(\pi x) = H(x)$, where $\pi x = (x_{\pi(1)}, \dots, x_{\pi(d)})$.

Our first result determines the asymptotic behavior of $U(H)^n_t$:

$$U(H)^n_t \xrightarrow{\text{u.c.p.}} U(H)_t := \int_{[0,t]^d} \rho_{\sigma_{\boldsymbol{s}}}(H) d\boldsymbol{s},$$

where $Z^n \xrightarrow{\text{u.c.p.}} Z$ denotes uniform convergence in probability, that is, for any $T > 0$, $\sup_{t\in[0,T]} |Z^n_t - Z_t| \xrightarrow{\mathbb{P}} 0$, and

$$\rho_{\sigma_{\boldsymbol{s}}}(H) := \int_{\mathbb{R}^d} H(\sigma_{s_1}u_1, \dots, \sigma_{s_d}u_d)\varphi_d(\boldsymbol{u})d\boldsymbol{u} \tag{A.2}$$

with $\varphi_d$ denoting the density of the d-dimensional standard Gaussian law $\mathcal{N}_d(0, \boldsymbol{I}_d)$. The second result of this paper is the stable functional central limit theorem

$$\sqrt{n}(U(H)^n - U(H)) \xrightarrow{st} L,$$

where $\xrightarrow{st}$ denotes stable convergence in law and the function $H$ is assumed to be even in each coordinate. The limiting process $L$ lives on an extension of the original probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and it turns out to be Gaussian with mean zero conditionally on the original $\sigma$-algebra $\mathcal{F}$. The proofs of the asymptotic results rely upon a combination of recent limit theorems for semimartingales (see e.g. [17, 20, 22]) and empirical processes techniques.

The paper is organized as follows. In § A.3 we present the law of large numbers for the U-statistic $U(H)^n_t$. The associated functional stable central limit theorem is provided in §A.4. Furthermore, we derive a standard central limit theorem in § A.5. In § A.6 we demonstrate statistical applications of our limit theory including Gini's mean difference, homoscedasticity testing and Wilcoxon statistics for testing of structural breaks. Some technical parts of the proofs are deferred to § A.7.

## A.2 Preliminaries

We consider the continuous diffusion model

$$X_t = x + \int_0^t a_s ds + \int_0^t \sigma_s dW_s, \qquad t \geq 0, \tag{A.3}$$

where $(a_s)_{s\geq 0}$ is a càglàd process, $(\sigma_s)_{s\geq 0}$ is a càdlàg process, both adapted to the filtration $(\mathcal{F}_s)_{s\geq 0}$. Define the functional class $C^k_p(\mathbb{R}^d)$ via

$$C^k_p(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{R} \mid f \in C^k(\mathbb{R}^d) \text{ and all derivatives up to order } k \text{ are of polynomial growth}\}.$$

Note that $H \in C_p^0(\mathbb{R}^d)$ implies that $\rho_{\sigma_s}(H) < \infty$ almost surely. For any vector $y \in \mathbb{R}^d$ we denote by $\|y\|$ its maximum norm; for any function $f : \mathbb{R}^d \to \mathbb{R}$, $\|f\|_\infty$ denotes its supremum norm. Finally, for any $z \neq 0$, $\Phi_z$ and $\varphi_z$ stand for the distribution function and density of the Gaussian law $\mathcal{N}(0, z^2)$, respectively; $\Phi_0$ denotes the Dirac measure at the origin. The bracket $[M, N]$ denotes the covariation process of two local martingales $M$ and $N$.

## A.3 Law of large numbers

We start with the law of large numbers, which describes the limit of the U-statistic $U(H)_t^n$ defined at (A.1). First of all, we remark that the processes $(a_s)_{s\geq 0}$ and $(\sigma_{s-})_{s\geq 0}$ are locally bounded, because they are both càglàd. Since the main results of this subsection (Proposition A.2 and Theorem A.3) are stable under stopping, we may assume without loss of generality that:

$$\text{The processes } a \text{ and } \sigma \text{ are bounded in } (\omega, t). \tag{A.4}$$

A detailed justification of this statement can be found in [2, § 3].

We start with the representation of the process $U(H)_t^n$ as an integral with respect to a certain empirical random measure. For this purpose let us introduce the quantity

$$\alpha_j^n := \sqrt{n}\sigma_{\frac{j-1}{n}} \Delta_j^n W, \qquad j \in \mathbb{N}, \tag{A.5}$$

which serves as a first order approximation of the increments $\sqrt{n}\Delta_j^n X$. The empirical distribution function associated with the random variables $(\alpha_j^n)_{1\leq j\leq [nt]}$ is defined as

$$F_n(t, x) := \frac{1}{n}\sum_{j=1}^{[nt]} \mathbb{1}_{\{\alpha_j^n \leq x\}}, \qquad x \in \mathbb{R},\ t \geq 0. \tag{A.6}$$

Notice that, for any fixed $t \geq 0$, $F_n(t, \cdot)$ is a finite random measure. Let $\widetilde{U}(H)_t^n$ be the U-statistic based on $\alpha_j^n$'s, i.e.

$$\widetilde{U}(H)_t^n = \binom{n}{d}^{-1} \sum_{\boldsymbol{i}\in\mathcal{A}_t^n(d)} H(\alpha_{\boldsymbol{i}}^n). \tag{A.7}$$

The functional $U_t'^n(H)$ defined as

$$U_t'^n(H) := \int_{\mathbb{R}^d} H(\boldsymbol{x}) F_n^{\otimes d}(t, d\boldsymbol{x}), \tag{A.8}$$

where

$$F_n^{\otimes d}(t, d\boldsymbol{x}) := F_n(t, dx_1)\cdots F_n(t, dx_d),$$

is closely related to the process $\widetilde{U}(H)_t^n$; in fact, if both are written out as multiple sums over nondecreasing multi-indices then their summands coincide on the set $\mathcal{A}_t^n(d)$. They differ for multi-indices that have at least two equal components. However, the number of these diagonal multi-indices is of order $O(n^{d-1})$. We start with a simple lemma, which we will often use throughout the paper. We omit a formal proof since it follows by standard arguments.

**Lemma A.1.** *Let $Z_n, Z : [0,T] \times \mathbb{R}^m \to \mathbb{R}$, $n \geq 1$, be random positive functions such that $Z_n(t,\cdot)$ and $Z(t,\cdot)$ are finite random measures on $\mathbb{R}^m$ for any $t \in [0,T]$. Assume that*

$$Z_n(\cdot, \boldsymbol{x}) \xrightarrow{\text{u.c.p.}} Z(\cdot, \boldsymbol{x}),$$

*for any fixed $\boldsymbol{x} \in \mathbb{R}^m$, and $\sup_{t\in[0,T],\boldsymbol{x}\in\mathbb{R}^m} Z(t,\boldsymbol{x})$, $\sup_{t\in[0,T],\boldsymbol{x}\in\mathbb{R}^m} Z_n(t,\boldsymbol{x})$, $n \geq 1$, are bounded random variables. Then, for any continuous function $Q : \mathbb{R}^m \to \mathbb{R}$ with compact support, we obtain that*

$$\int_{\mathbb{R}^m} Q(\boldsymbol{x}) Z_n(\cdot, d\boldsymbol{x}) \xrightarrow{\text{u.c.p.}} \int_{\mathbb{R}^m} Q(\boldsymbol{x}) Z(\cdot, d\boldsymbol{x}).$$

The next proposition determines the asymptotic behavior of the empirical distribution function $F_n(t,x)$ defined at (A.6), and the U-statistic $U_t'^n(H)$ given at (A.8).

**Proposition A.2.** *Assume that $H \in C_p^0(\mathbb{R}^d)$. Then, for any fixed $x \in \mathbb{R}$, it holds that*

$$F_n(t,x) \xrightarrow{\text{u.c.p.}} F(t,x) := \int_0^t \Phi_{\sigma_s}(x) ds. \tag{A.9}$$

*Furthermore, we obtain that*

$$U_t'^n(H) \xrightarrow{\text{u.c.p.}} U(H)_t := \int_{[0,t]^d} \rho_{\boldsymbol{\sigma_s}}(H) ds, \tag{A.10}$$

*where the quantity $\rho_{\boldsymbol{\sigma_s}}(H)$ is defined at* (A.2).

*Proof.* Recall that we always assume (A.4) without loss of generality. Here and throughout the paper, we denote by $C$ a generic positive constant, which may change from line to line; furthermore, we write $C_p$ if we want to emphasize the dependence of $C$ on an external parameter $p$. We first show the convergence in (A.9). Set $\xi_j^n := n^{-1}\mathbb{1}_{\{\alpha_j^n \leq x\}}$. It obviously holds that

$$\sum_{j=1}^{[nt]} \mathbb{E}[\xi_j^n | \mathcal{F}_{\frac{j-1}{n}}] = \frac{1}{n} \sum_{j=1}^{[nt]} \Phi_{\sigma_{\frac{j-1}{n}}}(x) \xrightarrow{\text{u.c.p.}} F(t,x),$$

for any fixed $x \in \mathbb{R}$, due to Riemann integrability of the process $\Phi_\sigma$. On the other hand, we have for any fixed $x \in \mathbb{R}$

$$\sum_{j=1}^{[nt]} \mathbb{E}[|\xi_j^n|^2 | \mathcal{F}_{\frac{j-1}{n}}] = \frac{1}{n^2} \sum_{j=1}^{[nt]} \Phi_{\sigma_{\frac{j-1}{n}}}(x) \xrightarrow{\mathbb{P}} 0.$$

This immediately implies the convergence (see [20, Lemma 2.2.11, p. 577])

$$F_n(t,x) - \sum_{j=1}^{[nt]} \mathbb{E}[\xi_j^n | \mathcal{F}_{\frac{j-1}{n}}] = \sum_{j=1}^{[nt]} \left( \xi_j^n - \mathbb{E}[\xi_j^n | \mathcal{F}_{\frac{j-1}{n}}] \right) \xrightarrow{\text{u.c.p.}} 0,$$

which completes the proof of (A.9). If $H$ is compactly supported then the convergence in (A.10) follows directly from (A.9) and Lemma A.1.

Now, let $H \in C_p^0(\mathbb{R}^d)$ be arbitrary. For any $k \in \mathbb{N}$, let $H_k \in C_p^0(\mathbb{R}^d)$ be a function with $H_k = H$ on $[-k,k]^d$ and $H_k = 0$ on $([-k-1,k+1]^d)^c$. We already know that

$$U'^n(H_k) \xrightarrow{\text{u.c.p.}} U(H_k),$$

for any fixed $k$, and $U(H_k) \xrightarrow{\text{u.c.p.}} U(H)$ as $k \to \infty$. Since the function $H$ has polynomial growth, i.e. $|H(\boldsymbol{x})| \leq C(1 + \|\boldsymbol{x}\|^q)$ for some $q > 0$, we obtain for any $p > 0$

$$\mathbb{E}[|H(\alpha_{\boldsymbol{i}}^n)|^p] \leq C_p \mathbb{E}[(1 + \|\alpha_{\boldsymbol{i}}^n\|^{qp})] \leq C_p \tag{A.11}$$

uniformly in $\boldsymbol{i}$, because the process $\sigma$ is bounded. Statement (A.11) also holds for $H_k$. Recall that the function $H - H_k$ vanishes on $[-k,k]^d$. Hence, we deduce by (A.11) and Cauchy-Schwarz inequality that

$$\begin{aligned}
&\mathbb{E}[\sup_{t \in [0,T]} |U_t'^n(H - H_k)|] \\
&\quad \leq C \binom{n}{d}^{-1} \sum_{1 \leq i_1, \ldots, i_d \leq [nT]} \left(\mathbb{E}[\mathbb{1}_{\{|\alpha_{i_1}^n| \geq k\}} + \cdots + \mathbb{1}_{\{|\alpha_{i_d}^n| \geq k\}}]\right)^{1/2} \\
&\quad \leq C_T \sup_{s \in [0,T]} \left(\mathbb{E}[1 - \Phi_{\sigma_s}(k)]\right)^{1/2} \to 0
\end{aligned}$$

as $k \to \infty$. This completes the proof of (A.10). □

Proposition A.2 implies the main result of this section.

**Theorem A.3.** *Assume that $H \in C_p^0(\mathbb{R}^d)$. Then it holds that*

$$U(H)_t^n \xrightarrow{\text{u.c.p.}} U(H)_t := \int_{[0,t]^d} \rho_{\sigma_{\boldsymbol{s}}}(H) ds, \tag{A.12}$$

*where the quantity $\rho_{\sigma_{\boldsymbol{s}}}(H)$ is defined at* (A.2).

*Proof.* In § A.7 we will show that

$$U(H)^n - \widetilde{U}(H)^n \xrightarrow{\text{u.c.p.}} 0, \tag{A.13}$$

where the functional $\widetilde{U}(H)_t^n$ is given at (A.7). In view of Proposition A.2, it remains to prove that $\widetilde{U}(H)_t^n - U_t'^n(H) \xrightarrow{\text{u.c.p.}} 0$. But due to the symmetry of $H$ and estimation (A.11), we obviously obtain that

$$\mathbb{E}[\sup_{t \in [0,T]} |\widetilde{U}(H)_t^n - U_t'^n(H)|] \leq \frac{C_T}{n} \to 0,$$

since the summands in $\widetilde{U}(H)_t^n$ and $U_t'^n(H)$ are equal except for diagonal multi-indices. □

**Remark A.1.** The result of Theorem A.3 can be extended to weighted U-statistics of the type

$$U(H;X)_t^n := \binom{n}{d}^{-1} \sum_{\boldsymbol{i}\in\mathcal{A}_t^n(d)} H(X_{\frac{\boldsymbol{i}-1}{n}}; \sqrt{n}\Delta_{\boldsymbol{i}}^n X). \tag{A.14}$$

Here, $H : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is assumed to be continuous and symmetric in the first and last $d$ arguments. Indeed, similar methods of proof imply the u.c.p. convergence

$$U(H;X)_t^n \xrightarrow{\text{u.c.p.}} U(H;X)_t = \int_{[0,t]^d} \rho_{\sigma_{\boldsymbol{s}}}(H;X_{\boldsymbol{s}})d\boldsymbol{s},$$

with

$$\rho_{\sigma_{\boldsymbol{s}}}(H;X_{\boldsymbol{s}}) := \int_{\mathbb{R}^d} H(X_{\boldsymbol{s}};\sigma_{s_1}u_1,\ldots,\sigma_{s_d}u_d)\varphi_d(\boldsymbol{u})d\boldsymbol{u}.$$

It is not essential that the weight process equals the diffusion process $X$. Instead, we may consider any $k$-dimensional $(\mathcal{F}_t)$-adapted Itô semimartingale of the type (A.3). We leave the details to the interested reader. □

## A.4 Stable central limit theorem

In this section we present a functional stable central limit theorem associated with the convergence in (A.12).

### Stable convergence

The concept of stable convergence of random variables has been originally introduced by Renyi [26]. For properties of stable convergence, we refer to [1, 25]. We recall the definition of stable convergence: Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence of random variables defined on $(\Omega,\mathcal{F},\mathbb{P})$ with values in a Polish space $(E,\mathcal{E})$. We say that $Y_n$ converges stably with limit $Y$, written $Y_n \xrightarrow{st} Y$, where $Y$ is defined on an extension $(\Omega',\mathcal{F}',\mathbb{P}')$ of the original probability space $(\Omega,\mathcal{F},\mathbb{P})$, if and only if for any bounded, continuous function $g$ and any bounded $\mathcal{F}$-measurable random variable $Z$ it holds that

$$\mathbb{E}[g(Y_n)Z] \to \mathbb{E}'[g(Y)Z], \quad n\to\infty.$$

Typically, we will deal with $E = \mathbb{D}([0,T],\mathbb{R})$ equipped with the Skorohod topology, or the uniform topology if the process $Y$ is continuous. Notice that stable convergence is a stronger mode of convergence than weak convergence. In fact, the statement $Y_n \xrightarrow{st} Y$ is equivalent to the joint weak convergence $(Y_n,Z) \xrightarrow{d} (Y,Z)$ for any $\mathcal{F}$-measurable random variable $Z$; see e.g. [1].

### Central limit theorem

For the stable central limit theorem we require a further structural assumption on the volatility process $(\sigma_s)_{s\ge 0}$. We assume that $\sigma$ itself is a continuous Itô semimartingale:

$$\sigma_t = \sigma_0 + \int_0^t \tilde{a}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \tilde{v}_s dV_s, \tag{A.15}$$

where the processes $(\tilde{a}_s)_{s\geq 0}$, $(\tilde{\sigma}_s)_{s\geq 0}$, $(\tilde{v}_s)_{s\geq 0}$ are càdlàg, adapted and $V$ is a Brownian motion independent of $W$. This type of condition is motivated by potential applications. For instance, when $\sigma_t = f(X_t)$ for a $C^2$-function $f$, then the Itô formula implies the representation (A.15) with $\tilde{v} \equiv 0$. In fact, a condition of the type (A.15) is nowadays a standard assumption for proving stable central limit theorems for functionals of high frequency data; see e.g. [2, 18]. Moreover, we assume that the process $\sigma$ does not vanish, i.e.

$$\sigma_s \neq 0 \qquad \text{for all } s \in [0,T]. \tag{A.16}$$

We believe that this assumption is not essential, but dropping it would make the following proofs considerably more involved and technical. As in the previous subsection, the central limit theorems presented in this paper are stable under stopping. This means, we may assume, without loss of generality, that

$$\text{The processes } a, \sigma, \sigma^{-1}, \tilde{a}, \tilde{\sigma} \text{ and } \tilde{v} \text{ are bounded in } (\omega, t). \tag{A.17}$$

We refer again to [2, § 3] for a detailed justification of this statement.

We need to introduce some further notation to describe the limiting process. First, we will study the asymptotic properties of the empirical process

$$\mathbb{G}_n(t,x) := \frac{1}{\sqrt{n}} \sum_{j=1}^{[nt]} \Big( \mathbb{1}_{\{\alpha_j^n \leq x\}} - \Phi_{\sigma_{\frac{j-1}{n}}}(x) \Big), \tag{A.18}$$

where $\alpha_j^n$ is defined at (A.5). This process is of crucial importance for proving the stable central limit theorem for the U-statistic $U(H)_t^n$. We start with the derivation of some useful inequalities for the process $\mathbb{G}_n$.

**Lemma A.4.** *For any even number $p \geq 2$ and $x, y \in \mathbb{R}$, we obtain the inequalities*

$$\mathbb{E}[\sup_{t\in[0,T]} |\mathbb{G}_n(t,x)|^p] \leq C_{T,p}\phi(x), \tag{A.19}$$

$$\mathbb{E}[\sup_{t\in[0,T]} |\mathbb{G}_n(t,x) - \mathbb{G}_n(t,y)|^p] \leq C_{T,p}|x-y|, \tag{A.20}$$

*where $\phi : \mathbb{R} \to \mathbb{R}$ is a bounded function (that depends on $p$ and $T$) with exponential decay at $\pm\infty$.*

*Proof.* Recall that the processes $\sigma$ and $\sigma^{-1}$ are assumed to be bounded. We begin with the inequality (A.19). For any given $x \in \mathbb{R}$, $(\mathbb{G}_n(t,x))_{t\in[0,T]}$ is an $(\mathcal{F}_{[nt]/n})$-martingale. Hence, the discrete Burkholder inequality implies that

$$\mathbb{E}[\sup_{t\in[0,T]} |\mathbb{G}_n(t,x)|^p] \leq C_{T,p}\mathbb{E}\Big[\Big|\sum_{j=1}^{[nT]} \zeta_j^n\Big|^{p/2}\Big]$$

with $\zeta_j^n := n^{-1}(\mathbb{1}_{\{\alpha_j^n \leq x\}} - \Phi_{\sigma_{(j-1)/n}}(x))^2$. Recalling that $p \geq 2$ is an even number und applying the Hölder inequality, we deduce that

$$\begin{aligned}
\Big|\sum_{j=1}^{[nT]} \zeta_j^n\Big|^{p/2} &\leq C_T n^{-1} \sum_{j=1}^{[nT]} (\mathbb{1}_{\{\alpha_j^n \leq x\}} - \Phi_{\sigma_{\frac{j-1}{n}}}(x))^p \\
&= C_T n^{-1} \sum_{j=1}^{[nT]} \sum_{k=0}^{p} \binom{p}{k} (-1)^k \Phi^k_{\sigma_{\frac{j-1}{n}}}(x) \mathbb{1}_{\{\alpha_j^n \leq x\}}.
\end{aligned}$$

Thus, we conclude that

$$\mathbb{E}[\sup_{t\in[0,T]} |\mathbb{G}_n(t,x)|^p] \le C_{T,p} \sup_{s\in[0,T]} \mathbb{E}[\Phi_{\sigma_s}(x)(1-\Phi_{\sigma_s}(x))^p] =: C_{T,p}\phi(x),$$

where the function $\phi$ obviously satisfies our requirements. This completes the proof of (A.19). By exactly the same methods we obtain, for any $x \ge y$,

$$\begin{aligned}&\mathbb{E}[\sup_{t\in[0,T]} |\mathbb{G}_n(t,x)-\mathbb{G}_n(t,y)|^p] \\ &\qquad \le C_{T,p} \sup_{s\in[0,T]} \mathbb{E}[(\Phi_{\sigma_s}(x)-\Phi_{\sigma_s}(y))(1-(\Phi_{\sigma_s}(x)-\Phi_{\sigma_s}(y)))^p]\end{aligned}$$

Since $\sigma$ and $\sigma^{-1}$ are both bounded, there exists a constant $M>0$ such that

$$\sup_{s\in[0,T]} |\Phi_{\sigma_s}(x)-\Phi_{\sigma_s}(y)| \le |x-y| \sup_{M^{-1}\le z\le M, y\le r\le x} \varphi_z(r).$$

This immediately gives (A.20). □

Our next result presents a functional stable central limit theorem for the process $\mathbb{G}_n$ defined at (A.18).

**Proposition A.5.** *We obtain the stable convergence*

$$\mathbb{G}_n(t,x) \xrightarrow{st} \mathbb{G}(t,x)$$

*on $\mathbb{D}([0,T])$ equipped with the uniform topology, where the convergence is functional in $t\in[0,T]$ and in finite distribution sense in $x\in\mathbb{R}$. The limiting process $\mathbb{G}$ is defined on an extension $(\Omega',\mathcal{F}',\mathbb{P}')$ of the original probability space $(\Omega,\mathcal{F},\mathbb{P})$ and it is Gaussian conditionally on $\mathcal{F}$. Its conditional drift and covariance kernel are given by*

$$\begin{aligned}\mathbb{E}'[\mathbb{G}(t,x)\mid\mathcal{F}] &= \int_0^t \overline{\Phi}_{\sigma_s}(x)dW_s,\\ \mathbb{E}'[\mathbb{G}(t_1,x_1)\mathbb{G}(t_2,x_2)\mid\mathcal{F}] &- \mathbb{E}'[\mathbb{G}(t_1,x_1)\mid\mathcal{F}]\mathbb{E}'[\mathbb{G}(t_2,x_2)\mid\mathcal{F}]\\ &= \int_0^{t_1\wedge t_2} \Phi_{\sigma_s}(x_1\wedge x_2)-\Phi_{\sigma_s}(x_1)\Phi_{\sigma_s}(x_2)-\overline{\Phi}_{\sigma_s}(x_1)\overline{\Phi}_{\sigma_s}(x_2)ds,\end{aligned}$$

*where $\overline{\Phi}_z(x)=\mathbb{E}[V\mathbb{1}_{\{zV\le x\}}]$ with $V\sim\mathcal{N}(0,1)$.*

*Proof.* Recall that due to (A.17) the process $\sigma$ is bounded in $(\omega,t)$. (However, note that we do not require the condition (A.15) to hold.) For any given $x_1,\dots,x_k\in\mathbb{R}$, we need to prove the functional stable convergence

$$(\mathbb{G}_n(\cdot,x_1),\dots,\mathbb{G}_n(\cdot,x_k)) \xrightarrow{st} (\mathbb{G}(\cdot,x_1),\dots,\mathbb{G}(\cdot,x_k)).$$

We write $\mathbb{G}_n(t,x_l)=\sum_{j=1}^{[nt]}\chi_{j,l}^n$ with

$$\chi_{j,l}^n := \frac{1}{\sqrt{n}}\Big(\mathbb{1}_{\{\alpha_j^n\le x_l\}} - \Phi_{\sigma_{\frac{j-1}{n}}}(x_l)\Big), \qquad 1\le l\le k.$$

According to [21, Theorem IX.7.28] we need to show that

$$\sum_{j=1}^{[nt]} \mathbb{E}[\chi_{j,r}^n \chi_{j,l}^n | \mathcal{F}_{\frac{j-1}{n}}] \xrightarrow{\mathbb{P}} \int_0^t \Big( \Phi_{\sigma_s}(x_r \wedge x_l) - \Phi_{\sigma_s}(x_r)\Phi_{\sigma_s}(x_l) \Big) ds, \tag{A.21}$$

$$\sum_{j=1}^{[nt]} \mathbb{E}[\chi_{j,l}^n \Delta_j^n W | \mathcal{F}_{\frac{j-1}{n}}] \xrightarrow{\mathbb{P}} \int_0^t \overline{\Phi}_{\sigma_s}(x_l) ds, \tag{A.22}$$

$$\sum_{j=1}^{[nt]} \mathbb{E}[|\chi_{j,l}^n|^2 \mathbb{1}_{\{|\chi_{j,l}^n| > \varepsilon\}} | \mathcal{F}_{\frac{j-1}{n}}] \xrightarrow{\mathbb{P}} 0, \qquad \text{for all } \varepsilon > 0, \tag{A.23}$$

$$\sum_{j=1}^{[nt]} \mathbb{E}[\chi_{j,l}^n \Delta_j^n N | \mathcal{F}_{\frac{j-1}{n}}] \xrightarrow{\mathbb{P}} 0, \tag{A.24}$$

where $1 \leq r, l \leq d$ and the last condition must hold for all bounded continuous martingales $N$ with $[W, N] = 0$. The convergence in (A.21) and (A.22) is obvious, since $\Delta_j^n W$ is independent of $\sigma_{(j-1)/n}$. We also have that

$$\sum_{j=1}^{[nt]} \mathbb{E}[|\chi_{j,l}^n|^2 \mathbb{1}_{\{|\chi_{j,l}^n| > \varepsilon\}} | \mathcal{F}_{\frac{j-1}{n}}] \leq \varepsilon^{-2} \sum_{j=1}^{[nt]} \mathbb{E}[|\chi_{j,l}^n|^4 | \mathcal{F}_{\frac{j-1}{n}}] \leq Cn^{-1},$$

which implies (A.23). Finally, let us prove (A.24). We fix $l$ and define $M_u := \mathbb{E}[\chi_{j,l}^n | \mathcal{F}_u]$ for $u \geq (j-1)/n$. By the martingale representation theorem we deduce the identity

$$M_u = M_{\frac{j-1}{n}} + \int_{\frac{j-1}{n}}^u \eta_s dW_s$$

for a suitable predictable process $\eta$. By the Itô isometry we conclude that

$$\mathbb{E}[\chi_{j,l}^n \Delta_j^n N | \mathcal{F}_{\frac{j-1}{n}}] = \mathbb{E}[M_{\frac{j}{n}} \Delta_j^n N | \mathcal{F}_{\frac{j-1}{n}}] = \mathbb{E}[\Delta_j^n M \Delta_j^n N | \mathcal{F}_{\frac{j-1}{n}}] = 0.$$

This completes the proof of Proposition A.5. □

We suspect that the stable convergence in Proposition A.5 also holds in the functional sense in the $x$ variable. However, proving tightness (even on compact sets) turns out to be a difficult task. In particular, inequality (A.20) is not sufficient for showing tightness.

**Remark A.2.** We highlight some probabilistic properties of the limiting process $\mathbb{G}$ defined in Proposition A.5.

(i) Proposition A.5 can be reformulated as follows. Let $x_1, \ldots, x_k \in \mathbb{R}$ be arbitrary real numbers. Then it holds that

$$(\mathbb{G}_n(\cdot, x_1), \ldots, \mathbb{G}_n(\cdot, x_k)) \xrightarrow{st} \int_0^{\cdot} v_s dW_s + \int_0^{\cdot} w_s^{1/2} dW_s',$$

where $W'$ is a $k$-dimensional Brownian motion independent of $\mathcal{F}$, and $v$ and $w$ are $\mathbb{R}^k$-valued and $\mathbb{R}^{k\times k}$-valued processes, respectively, with coordinates

$$\begin{aligned} v_s^r &= \overline{\Phi}_{\sigma_s}(x_r), \\ w_s^{rl} &= \Phi_{\sigma_s}(x_r \wedge x_l) - \Phi_{\sigma_s}(x_r)\Phi_{\sigma_s}(x_l) - \overline{\Phi}_{\sigma_s}(x_r)\overline{\Phi}_{\sigma_s}(x_l), \end{aligned}$$

for $1 \leq r, l \leq k$. This type of formulation appears in [21, Theorem IX.7.28]. In particular, $(\mathbb{G}(\cdot, x_l))_{1\leq l\leq k}$ is a $k$-dimensional martingale. $\square$

(ii) It is obvious from (i) that $\mathbb{G}$ is continuous in $t$. Moreover, $\mathbb{G}$ is also continuous in $x$. This follows from Kolmogorov's criterion and the inequality $(y \leq x)$

$$\begin{aligned} &\mathbb{E}'[|\mathbb{G}(t,x) - \mathbb{G}(t,y)|^p] \\ &\leq C_p \mathbb{E}\Big[\Big(\int_0^t \Big\{\Phi_{\sigma_s}(x) - \Phi_{\sigma_s}(y) - (\Phi_{\sigma_s}(x) - \Phi_{\sigma_s}(y))^2\Big\} ds\Big)^{p/2}\Big] \\ &\leq C_p (x-y)^{p/2}, \end{aligned}$$

for any $p > 0$, which follows by the Burkholder inequality. In particular, $\mathbb{G}(t, \cdot)$ has Hölder continuous paths of order $1/2 - \varepsilon$, for any $\varepsilon \in (0, 1/2)$. $\square$

(iii) A straightforward computation (cf. (A.19)) shows that the function $\mathbb{E}[\sup_{t\in[0,T]} \mathbb{G}(t,x)^2]$ has exponential decay as $x \to \pm\infty$. Hence, for any function $f \in C_p^1(\mathbb{R})$, we have

$$\int_{\mathbb{R}} f(x)\mathbb{G}(t, dx) < \infty, \quad \text{a.s..}$$

If $f$ is an even function, we also have that

$$\int_{\mathbb{R}} f(x)\mathbb{G}(t,dx) = \int_{\mathbb{R}} f(x)(\mathbb{G}(t,dx) - \mathbb{E}'[\mathbb{G}(t,dx)|\mathcal{F}]),$$

since

$$\int_{\mathbb{R}} f(x)\mathbb{E}'[\mathbb{G}(t,dx)|\mathcal{F}] = \int_0^t \left(\int_{\mathbb{R}} f(x)\overline{\Phi}_{\sigma_s}(dx)\right) dW_s,$$

and, for any $z > 0$,

$$\int_{\mathbb{R}} f(x)\overline{\Phi}_z(dx) = \int_{\mathbb{R}} x f(x)\varphi_z(x)dx = 0,$$

because $f\varphi_z$ is an even function. The same argument applies for $z < 0$. Furthermore, the integration by parts formula and the aforementioned argument imply the identity

$$\begin{aligned} &\mathbb{E}'\Big[\Big|\int_{\mathbb{R}} f(x)\mathbb{G}(t,dx)\Big|^2 |\mathcal{F}\Big] \\ &= \int_0^t \left(\int_{\mathbb{R}^2} f'(x)f'(y)\Big(\Phi_{\sigma_s}(x\wedge y) - \Phi_{\sigma_s}(x)\Phi_{\sigma_s}(y)\Big)dxdy\right) ds. \end{aligned}$$

We remark that, for any $z \neq 0$, we have

$$\mathrm{var}[f(V)] = \int_{\mathbb{R}^2} f'(x)f'(y)\Big(\Phi_z(x\wedge y) - \Phi_z(x)\Phi_z(y)\Big)dxdy$$

with $V \sim \mathcal{N}(0, z^2)$. $\square$

Now, we present a functional stable central limit theorem of the U-statistic $U_t'^n(H)$ given at (A.8), which is based on the approximative quantities $(\alpha_j^n)_{1\le j\le [nt]}$ defined at (A.5).

**Proposition A.6.** *Assume that the conditions* (A.15), (A.16), *and* (A.17) *hold. Let $H \in C_p^1(\mathbb{R}^d)$ be a symmetric function that is even in each (or, equivalently, in one) argument. Then we obtain the functional stable convergence*

$$\sqrt{n}(U'^n(H) - U(H)) \xrightarrow{st} L,$$

*where*

$$L_t = d \int_{\mathbb{R}^d} H(x_1,\dots,x_d)\mathbb{G}(t,dx_1)F(t,dx_2)\cdots F(t,dx_d). \tag{A.25}$$

*The convergence takes place in $\mathbb{D}([0,T])$ equipped with the uniform topology. Furthermore, $\mathbb{G}$ can be replaced by $\mathbb{G} - \mathbb{E}'[\mathbb{G}|\mathcal{F}]$ without changing the limit and, consequently, $L$ is a centered Gaussian process, conditionally on $\mathcal{F}$.*

*Proof.* First of all, we remark that

$$\int_{\mathbb{R}} H(x_1,\dots,x_d)\mathbb{E}'[\mathbb{G}(t,dx_1)|\mathcal{F}] = 0$$

follows from Remark A.2(iii). The main part of the proof is divided into five steps:

**(i)** In § A.7 we will show that under condition (A.15) we have

$$\sqrt{n}\Big(U(H)_t - \int_{\mathbb{R}^d} H(\boldsymbol{x})\overline{F}_n^{\otimes d}(t,d\boldsymbol{x})\Big) \xrightarrow{\text{u.c.p.}} 0 \tag{A.26}$$

with

$$\overline{F}_n(t,x) := \frac{1}{n}\sum_{j=1}^{[nt]} \Phi_{\sigma_{\frac{j-1}{n}}}(x).$$

Thus, we need to prove the stable convergence $L^n \xrightarrow{st} L$ for

$$L_t^n := \sqrt{n}\Big(U_t'^n(H) - \int_{\mathbb{R}^d} H(\boldsymbol{x})\overline{F}_n^{\otimes d}(t,d\boldsymbol{x})\Big). \tag{A.27}$$

Assume that the function $H \in C^1(\mathbb{R}^d)$ has compact support. Recalling the definition (A.18) of the empirical process $\mathbb{G}_n$, we obtain the identity

$$L_t^n = \sum_{l=1}^d \int_{\mathbb{R}^d} H(\boldsymbol{x})\mathbb{G}_n(t,dx_l)\prod_{m=1}^{l-1} F_n(t,dx_m)\prod_{m=l+1}^{d} \overline{F}_n(t,dx_m).$$

In step (iv) we will show that both $F_n(t,dx_m)$ and $\overline{F}_n(t,dx_m)$ can be replaced by $F(t,dx_m)$ without affecting the limit. In other words, $L^n - L'^n \xrightarrow{\text{u.c.p.}} 0$ with

$$L_t'^n := \sum_{l=1}^d \int_{\mathbb{R}^d} H(\boldsymbol{x})\mathbb{G}_n(t,dx_l)\prod_{m\neq l} F(t,dx_m).$$

But, since $H$ is symmetric, we readily deduce that

$$L_t'^n = d \int_{\mathbb{R}^d} H(\boldsymbol{x}) \mathbb{G}_n(t, dx_1) \prod_{m=2}^d F(t, dx_m).$$

The random measure $F(t,x)$ has a Lebesgue density in $x$ due to assumption (A.16), which we denote by $F'(t,x)$. The integration by parts formula implies that

$$L_t'^n = -d \int_{\mathbb{R}^d} \partial_1 H(\boldsymbol{x}) \mathbb{G}_n(t, x_1) \prod_{m=2}^d F'(t, x_m) d\boldsymbol{x},$$

where $\partial_l H$ denotes the partial derivative of $H$ with respect to $x_l$. This identity completes step (i).

**(ii)** In this step we will start proving the stable convergence $L'^n \xrightarrow{st} L$ (the function $H \in C^1(\mathbb{R}^d)$ is still assumed to have compact support). Since the stable convergence $\mathbb{G}_n \xrightarrow{st} \mathbb{G}$ does not hold in the functional sense in the $x$ variable, we need to overcome this problem by a Riemann sum approximation. Let the support of $H$ be contained in $[-k,k]^d$. Let $-k = z_0 < \cdots < z_l = k$ be the equidistant partition of the interval $[-k,k]$. We set

$$Q(t, x_1) := \int_{\mathbb{R}^{d-1}} \partial_1 H(x_1, \ldots, x_d) \prod_{m=2}^d F'(t, x_m) dx_2 \ldots dx_d,$$

and define the approximation of $L_t'^n$ via

$$L_t'^n(l) = -\frac{2dk}{l} \sum_{j=0}^{l} Q(t, z_j) \mathbb{G}_n(t, z_j).$$

Proposition A.5 and the properties of stable convergence imply that

$$\Big(Q(\cdot, z_j), \mathbb{G}_n(\cdot, z_j)\Big)_{0 \le j \le l} \xrightarrow{st} \Big(Q(\cdot, z_j), \mathbb{G}(\cdot, z_j)\Big)_{0 \le j \le l}.$$

Hence, we deduce the stable convergence

$$L_\cdot'^n(l) \xrightarrow{st} L_\cdot(l) := -\frac{2dk}{l} \sum_{j=0}^{l} Q(\cdot, z_j) \mathbb{G}(\cdot, z_j).$$

as $n \to \infty$, for any fixed $l$. Furthermore, we obtain the convergence

$$L(l) \xrightarrow{\text{u.c.p.}} L$$

as $l \to \infty$, where we reversed all above transformations. This convergence completes step (ii).

**(iii)** To complete the proof of the stable convergence $L'^n \xrightarrow{st} L$, we need to show that

$$\lim_{l\to\infty} \limsup_{n\to\infty} \sup_{t\in[0,T]} |L'^n_t(l) - L'^n_t| = 0,$$

where the limits are taken in probability. With $h = l/2k$ we obtain that

$$|L'^n_t(l) - L'^n_t| = d\left|\int_{\mathbb{R}} \Big\{Q(t,[xh]/h)\mathbb{G}_n(t,[xh]/h) - Q(t,x)\mathbb{G}_n(t,x)\Big\}dx\right|.$$

Observe that

$$\sup_{t\in[0,T]} |F'(t,x_m)| = \int_0^T \varphi_{\sigma_s}(x_m)ds \le T \sup_{M^{-1}\le z\le M} \varphi_z(x_m), \tag{A.28}$$

where $M$ is a positive constant with $M^{-1} \le |\sigma| \le M$. Recalling the definition of $Q(t,x)$ we obtain that

$$\sup_{t\in[0,T]} |Q(t,x)| \le C_T, \qquad \sup_{t\in[0,T]} |Q(t,x) - Q(t,[xh]/h)| \le C_T\eta(h^{-1}), \tag{A.29}$$

where $\eta(\varepsilon) := \sup\big\{|\partial_1 H(\boldsymbol{y}_1) - \partial_1 H(\boldsymbol{y}_2)| \,\big|\, \|\boldsymbol{y}_1 - \boldsymbol{y}_2\| \le \varepsilon,\ \boldsymbol{y}_1, \boldsymbol{y}_2 \in [-k,k]^d\big\}$ denotes the modulus of continuity of the function $\partial_1 H$. We also deduce by Lemma A.4 that

$$\mathbb{E}[\sup_{t\in[0,T]} |\mathbb{G}_n(t,x)|^p] \le C_T, \tag{A.30}$$

$$\mathbb{E}[\sup_{t\in[0,T]} |\mathbb{G}_n(t,x) - \mathbb{G}_n(t,[xh]/h)|^p] \le C_T h^{-1}, \tag{A.31}$$

for any even number $p \ge 2$. Combining the inequalities (A.29), (A.30) and (A.31), we deduce the convergence

$$\lim_{l\to\infty} \limsup_{n\to\infty} \mathbb{E}[\sup_{t\in[0,T]} |L'^n_t(l) - L'^n_t|] = 0$$

using that $Q(t,\cdot)$ has compact support contained in $[-k,k]$. Hence, $L'^n \xrightarrow{st} L$ and we are done.

**(iv)** In this step we will prove the convergence

$$L^n - L'^n \xrightarrow{\text{u.c.p.}} 0.$$

This difference can be decomposed into several terms; in the following we will treat a typical representative (all other terms are treated in exactly the same manner). For $l < d$ define

$$R^n_t(l) := \int_{\mathbb{R}^d} H(\boldsymbol{x})\mathbb{G}_n(t,dx_l) \prod_{m=1}^{l-1} F_n(t,dx_m) \times \prod_{m=l+1}^{d-1} \overline{F}_n(t,dx_m)[\overline{F}_n(t,dx_d) - F(t,dx_d)].$$

Now, we use the integration by parts formula to obtain that

$$R_t^n(l) = \int_{\mathbb{R}} N_n(t, x_l)\mathbb{G}_n(t, x_l)dx_l,$$

where

$$N_n(t, x_l) = \int_{\mathbb{R}^{d-1}} \partial_l H(\boldsymbol{x}) \prod_{m=1}^{l-1} F_n(t, dx_m) \times \prod_{m=l+1}^{d-1} \overline{F}_n(t, dx_m)[\overline{F}_n(t, dx_d) - F(t, dx_d)].$$

As in step (iii) we deduce for any even $p \geq 2$,

$$\mathbb{E}[\sup_{t\in[0,T]} |\mathbb{G}_n(t, x_l)|^p] \leq C_p, \quad \mathbb{E}[\sup_{t\in[0,T]} |N_n(t, x_l)|^p] \leq C_p.$$

Recalling that the function $H$ has compact support and applying the dominated convergence theorem, it is sufficient to show that

$$N_n(\cdot, x_l) \xrightarrow{\text{u.c.p.}} 0,$$

for any fixed $x_l$. But this follows immediately from Lemma A.1, since

$$F_n(\cdot, x) \xrightarrow{\text{u.c.p.}} F(\cdot, x), \quad \overline{F}_n(\cdot, x) \xrightarrow{\text{u.c.p.}} F(\cdot, x),$$

for any fixed $x \in \mathbb{R}$, and $\partial_l H$ is a continuous function with compact support. This finishes the proof of step (iv).

**(v)** Finally, let $H \in C_p^1(\mathbb{R}^d)$ be arbitrary. For any $k \in \mathbb{N}$, let $H_k \in C_p^1(\mathbb{R}^d)$ be a function with $H_k = H$ on $[-k, k]^d$ and $H_k = 0$ on $([-k-1, k+1]^d)^c$. Let us denote by $L_t^n(H)$ and $L_t(H)$ the processes defined by (A.27) and (A.25), respectively, that are associated with a given function $H$. We know from the previous steps that

$$L^n(H_k) \xrightarrow{st} L(H_k)$$

as $n \to \infty$, and $L(H_k) \xrightarrow{\text{u.c.p.}} L(H)$ as $k \to \infty$. So, we are left to proving that

$$\lim_{k\to\infty} \limsup_{n\to\infty} \sup_{t\in[0,T]} |L_t^n(H_k) - L_t^n(H)| = 0,$$

where the limits are taken in probability. As in steps (ii) and (iii) we obtain the identity

$$\begin{aligned} &L_t^n(H_k) - L_t^n(H) \\ &\quad = \sum_{l=1}^{d} \int_{\mathbb{R}^d} \partial_l(H - H_k)(\boldsymbol{x})\mathbb{G}_n(t, x_l)dx_l \prod_{m=1}^{l-1} F_n(t, dx_m) \prod_{m=l+1}^{d} \overline{F}_n(t, dx_m) \\ &\quad =: \sum_{l=1}^{d} Q^l(k)_t^n. \end{aligned}$$

We deduce the inequality

$$|Q^l(k)^n_t| \le n^{-(l-1)} \sum_{i_1,\dots,i_{l-1}=1}^{[nt]} \int_{\mathbb{R}^{d-l+1}} |\partial_l(H-H_k)(\alpha^n_{i_1},\dots,\alpha^n_{i_{l-1}},x_l,\dots,x_d)| \times |\mathbb{G}_n(t,x_l)| \prod_{m=l+1}^{d} \overline{F}'_n(t,x_m)dx_l\dots dx_d.$$

We remark that $\partial_l(H_k - H)$ vanishes if all arguments lie in the interval $[-k,k]$. Hence,

$$|Q^l(k)^n_t| \le n^{-(l-1)} \sum_{i_1,\dots,i_{l-1}=1}^{[nt]} \int_{\mathbb{R}^{d-l+1}} |\partial_l(H-H_k)(\alpha^n_{i_1},\dots,\alpha^n_{i_{l-1}},x_l,\dots,x_d)| \times \Big(\sum_{m=1}^{l-1} \mathbb{1}_{\{|\alpha^n_{i_m}|>k\}} + \sum_{m=l}^{d} \mathbb{1}_{\{|x_m|>k\}}\Big)|\mathbb{G}_n(t,x_l)| \prod_{m=l+1}^{d} \overline{F}'_n(t,x_m)dx_l\dots dx_d.$$

Now, applying Lemma A.4, (A.11), (A.28) and the Cauchy-Schwarz inequality, we deduce that

$$\mathbb{E}[\sup_{t\in[0,T]} |Q^l(k)^n_t|] \le C_T \int_{\mathbb{R}^{d-l+1}} \Big((l-1)\sup_{M^{-1}\le z\le M}(1-\Phi_z(k)) + \sum_{m=l}^{d} \mathbb{1}_{\{|x_m|>k\}}\Big)^{1/2} \times \psi(x_l,\dots,x_d)\phi(x_l) \prod_{m=l+1}^{d} \sup_{M^{-1}\le z\le M} \varphi_z(x_m)dx_l\dots dx_d,$$

for some bounded function $\phi$ with exponential decay at $\pm\infty$ and a function $\psi \in C^0_p(\mathbb{R}^{d-l+1})$. Hence

$$\int_{\mathbb{R}^{d-l+1}} \psi(x_l,\dots,x_d)\phi(x_l) \prod_{m=l+1}^{d} \sup_{M^{-1}\le z\le M} \varphi_z(x_m)dx_l\dots dx_d < \infty,$$

and we conclude that

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbb{E}[\sup_{t\in[0,T]} |Q^l(k)^n_t|] = 0.$$

This finishes step (v) and we are done with the proof of Proposition A.6. □

Notice that an additional $\mathcal{F}$-conditional bias would appear in the limiting process $L$ if we would drop the assumption that $H$ is even in each coordinate. The corresponding asymptotic theory for the case $d = 1$ has been studied in [22]; see also [17].

**Remark A.3.** Combining limit theorems for semimartingales with the empirical distribution function approach is probably the most efficient way of proving Proposition A.6. Nevertheless, we shortly comment on alternative methods of proof.

Treating the multiple sum in the definition of $U'^n(H)$ directly is relatively complicated, since at a certain stage of the proof one will have to deal with partial sums of functions of $\alpha_j^n$ weighted by an anticipative process. This anticipation of the weight process makes it impossible to apply martingale methods directly.

Another approach to proving Proposition A.6 is a pseudo Hoeffding decomposition. This method relies on the application of the classical Hoeffding decomposition to $U'^n(H)$ by pretending that the scaling components $\sigma_{(\boldsymbol{i}-1)/n}$ are non-random. However, since the random variables $\alpha_j^n$ are not independent when the process $\sigma$ is stochastic, the treatment of the error term connected with the pseudo Hoeffding decomposition will not be easy, because the usual orthogonality arguments of the Hoeffding method do not apply in our setting. □

**Remark A.4.** In the context of Proposition A.6 we would like to mention a very recent work by Beutner and Zähle [3]. They study the empirical distribution function approach to U- and V-statistics for unbounded kernels $H$ in the classical i.i.d. or weakly dependent setting. Their method relies on the application of the functional delta method for quasi-Hadamard differentiable functionals. In our setting it would require the functional convergence

$$\mathbb{G}_n(t,\cdot) \xrightarrow{st} \mathbb{G}(t,\cdot),$$

where the convergence takes place in the space of càdlàg functions equipped with the weighted sup-norm $\|f\|_\lambda := \sup_{x\in\mathbb{R}} |(1+|x|^\lambda)f(x)|$ for some $\lambda > 0$. Although we do not really require such a strong result in our framework (as can be seen from the proof of Proposition A.6), it would be interesting to prove this type of convergence for functionals of high frequency data; cf. the comment before Remark A.2. □

To conclude this section, we finally present the main result: A functional stable central limit theorem for the original U-statistic $U(H)^n$.

**Theorem A.7.** *Assume that the symmetric function $H \in C_p^1(\mathbb{R}^d)$ is even in each (or, equivalently, in one) argument. If $\sigma$ satisfies conditions* (A.15) *and* (A.16)*, we obtain the functional stable central limit theorem*

$$\sqrt{n}\Big(U(H)^n - U(H)\Big) \xrightarrow{st} L, \tag{A.32}$$

*where the convergence takes place in $\mathbb{D}([0,T])$ equipped with the uniform topology and the limiting process $L$ is defined at* (A.25).

*Proof.* In § A.7 we will show the following statement: Under condition (A.15) it holds that

$$\sqrt{n}|U(H)^n - \widetilde{U}(H)^n| \xrightarrow{\text{u.c.p.}} 0. \tag{A.33}$$

In view of Proposition A.6, it remains to prove that $\sqrt{n}|\widetilde{U}(H)_t^n - U_t'^n(H)| \xrightarrow{\text{u.c.p.}} 0$. But due to the symmetry of $H$, we obtain as in the proof of Theorem A.3

$$\mathbb{E}[\sup_{t\in[0,T]} |\widetilde{U}(H)_t^n - U_t'^n(H)|] \le \frac{C_T}{n}.$$

This finishes the proof of Theorem A.7. □

We remark that the stable convergence at (A.32) is not feasible in its present form, since the distribution of the limiting process $L$ is unknown. In the next section we will explain how to obtain a feasible central limit theorem that opens the door to statistical applications.

## A.5 Estimation of the conditional variance

In this section we present a standard central limit theorem for the U-statistic $U(H)_t^n$. We will confine ourselves to the presentation of a result in finite distributional sense. According to Remark A.2 (iii) applied to

$$f_t(x) := d \int_{\mathbb{R}^{d-1}} H(x, x_2 \dots, x_d) F(t, dx_2) \cdots F(t, dx_d),$$

the conditional variance of the limit $L_t$ is given by

$$V_t := \mathbb{E}'[|L_t|^2 | \mathcal{F}] = \int_0^t \left( \int_{\mathbb{R}} f_t^2(x) \varphi_{\sigma_s}(x) dx - \left( \int_{\mathbb{R}} f_t(x) \varphi_{\sigma_s}(x) dx \right)^2 \right) ds.$$

Hence, the random variable $L_t$ is non-degenerate when

$$\mathrm{var}\Big( \mathbb{E}[H(x_1 U_1, \dots, x_d U_d) | U_1] \Big) > 0, \qquad (U_1, \dots, U_d) \sim \mathcal{N}_d(0, \boldsymbol{I}_d),$$

for all $x_1, \dots, x_d \in \{\sigma_s \mid s \in A \subseteq [0,t]\}$ and some set $A$ with positive Lebesgue measure. This essentially coincides with the classical non-degeneracy condition for U-statistics of independent random variables.

We define the functions $G_1 : \mathbb{R}^{2d-1} \to \mathbb{R}$ and $G_2 : \mathbb{R}^2 \times \mathbb{R}^{2d-2} \to \mathbb{R}$ by

$$G_1(\boldsymbol{x}) = H(x_1, x_2 \dots, x_d) H(x_1, x_{d+1}, \dots, x_{2d-1}), \tag{A.34}$$
$$G_2(\boldsymbol{x}; \boldsymbol{y}) = H(x_1, y_1, \dots, y_{d-1}) H(x_2, y_d, \dots, y_{2d-2}), \tag{A.35}$$

respectively. Then $V_t$ can be written as

$$\begin{aligned} V_t = {} & d^2 \int_{[0,t]^{2d-1}} \rho_{\sigma_{\boldsymbol{s}}}(G_1) d\boldsymbol{s} \\ & - d^2 \int_{[0,t]^{2d-2}} \int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}} \rho_{\sigma_{\boldsymbol{s}}}(G_2(x_1, x_2; \cdot)) \varphi_{\sigma_q}(x_1) \varphi_{\sigma_q}(x_2) dx_1\, dx_2\, dq\, d\boldsymbol{s}. \end{aligned}$$

We denote the first and second summand on the right hand side of the preceding equation by $V_{1,t}$ and $V_{2,t}$, respectively. Let $\widetilde{G}_1$ denote the symmetrization of the function $G_1$. By Theorem A.3 it holds that

$$V_{1,t}^n = d^2 U(\widetilde{G}_1)_t^n \xrightarrow{\text{u.c.p.}} d^2 U(\widetilde{G}_1)_t = V_{1,t}.$$

The multiple integral $V_{2,t}$ is almost in the form of the limit in Theorem A.3, and it is indeed possible to estimate it by a slightly modified U-statistic as the following proposition shows. The statistic presented in the following proposition is a generalization of the bipower concept discussed e.g. in [2] in the case $d = 1$.

**Proposition A.8.** *Assume that $H \in C_p^0(\mathbb{R}^d)$. Let*

$$V_{2,t}^n := \frac{d^2}{n} \binom{n}{2d-2}^{-1} \sum_{\boldsymbol{i} \in \mathcal{A}_t^n(2d-2)} \times \sum_{j=1}^{[nt]-1} \widetilde{G}_2(\sqrt{n}\Delta_j^n X, \sqrt{n}\Delta_{j+1}^n X; \sqrt{n}\Delta_{i_1}^n X, \ldots, \sqrt{n}\Delta_{i_{2d-2}}^n X),$$

*where $\widetilde{G}_2$ denotes the symmetrization of $G_2$ with respect to the $\boldsymbol{y}$-values, that is*

$$\widetilde{G}_2(\boldsymbol{x}; \boldsymbol{y}) = \frac{1}{(2d-2)!} \sum_{\pi} G_2(\boldsymbol{x}; \pi\boldsymbol{y}),$$

*for $\boldsymbol{x} \in \mathbb{R}^2$, $\boldsymbol{y} \in \mathbb{R}^{2d-2}$, and where the sum runs over all permutations of $\{1, \ldots, 2d-2\}$. Then*

$$V_2^n \xrightarrow{\text{u.c.p.}} V_2.$$

*Proof.* The result can be shown using essentially the same arguments as in the proofs of Proposition A.2 and Theorem A.3. We provide a sketch of the proof. Similar to (A.7) we define

$$\widetilde{V}_{2,t}^n := \frac{d^2}{n} \binom{n}{2d-2}^{-1} \sum_{\boldsymbol{i} \in \mathcal{A}_t^n(2d-2)} \sum_{j=1}^{[nt]-1} \widetilde{G}_2(\alpha_j^n, \alpha_{j+1}'^n; \alpha_{i_1}^n, \ldots, \alpha_{i_{2d-2}}^n),$$

where $\alpha_{j+1}'^n := \sqrt{n}\sigma_{\frac{j-1}{n}} \Delta_{i+1}^n W$. Analogously to (A.8) we introduce the random process

$$V_{2,t}'^n := d^2 \int_{\mathbb{R}^{2d-2}} \int_{\mathbb{R}^2} \widetilde{G}_2(\boldsymbol{x}; \boldsymbol{y}) \widetilde{F}_n(t, d\boldsymbol{x}) F_n^{\otimes(2d-2)}(t, d\boldsymbol{y}),$$

where

$$\widetilde{F}_n(t, x_1, x_2) = \frac{1}{n} \sum_{j=1}^{[nt]-1} \mathbb{1}_{\{\alpha_j^n \le x_1\}} \mathbb{1}_{\{\alpha_{j+1}'^n \le x_2\}}.$$

Writing out $V_{2,t}'^n$ as a multiple sum over nondecreasing multi-indices in the $\boldsymbol{y}$ arguments, one observes as before that $V_{2,t}'^n$ and $\widetilde{V}_{2,t}^n$ differ in at most $O(n^{2d-3})$ summands. Therefore, using the same argument as in the proof of Theorem A.3

$$\widetilde{V}_{2,t}^n - V_{2,t}'^n \xrightarrow{\text{u.c.p.}} 0.$$

For any fixed $x, y \in \mathbb{R}$ it holds that

$$\widetilde{F}_n(t, x, y) \xrightarrow{\text{u.c.p.}} \widetilde{F}(t, x, y) := \int_0^t \Phi_{\sigma_s}(x) \Phi_{\sigma_s}(y) ds.$$

This can be shown similarly to the proof of Proposition A.2 as follows. Let $\xi_j^n = n^{-1} \mathbb{1}_{\{\alpha_j^n \le x_1\}} \mathbb{1}_{\{\alpha_{j+1}'^n \le x_2\}}$. Then,

$$\sum_{j=1}^{[nt]-1} \mathbb{E}[\xi_j^n | \mathcal{F}_{\frac{j-1}{n}}] = \frac{1}{n} \sum_{j=1}^{[nt]-1} \Phi_{\sigma_{\frac{j-1}{n}}}(x_1) \Phi_{\sigma_{\frac{j-1}{n}}}(x_2) \xrightarrow{\text{u.c.p.}} \widetilde{F}(t, x, y).$$

On the other hand, we trivially have that $\sum_{j=1}^{[nt]-1} \mathbb{E}[|\xi_j^n|^2|\mathcal{F}_{\frac{j-1}{n}}] \xrightarrow{\mathbb{P}} 0$, for any fixed $t > 0$. Hence, the Lenglart's domination property (see [21, p. 35]) implies the convergence

$$\sum_{j=1}^{[nt]-1} \left(\xi_j^n - \mathbb{E}[\xi_j^n|\mathcal{F}_{\frac{j-1}{n}}]\right) \xrightarrow{\text{u.c.p.}} 0,$$

which in turn means that $\widetilde{F}_n(t, x, y) \xrightarrow{\text{u.c.p.}} \widetilde{F}(t, x, y)$.

We know now that $V_{2,t}'^n$ converges to the claimed limit if $G_2$ is compactly supported. For a general $G_2$ with polynomial growth one can proceed exactly as in Proposition A.2. To conclude the proof, one has to show that $V_{2,t}^n - V_{2,t}'^n \xrightarrow{\text{u.c.p.}} 0$. This works exactly as in § A.7. □

The properties of stable convergence immediately imply the following theorem.

**Theorem A.9.** *Let the assumptions of Theorem A.7 be satisfied. Let $t > 0$ be fixed. Then we obtain the standard central limit theorem*

$$\frac{\sqrt{n}\big(U(H)_t^n - U(H)_t\big)}{\sqrt{V_t^n}} \xrightarrow{d} \mathcal{N}(0, 1), \tag{A.36}$$

*where $V_t^n = V_{1,t}^n - V_{2,t}^n$ using the notation defined above.*

The convergence in law in (A.36) is a feasible central limit theorem that can be used in statistical applications. It is possible to obtain similar multivariate central limit theorems for finite dimensional vectors $\sqrt{n}\big(U(H)_{t_j}^n - U(H)_{t_j}\big)_{1\le j\le k}$; we leave the details to the interested reader.

## A.6 Statistical applications

In this section we present some statistical applications of the limit theory for U-statistics of continuous Itô semimartingales.

### Gini's mean difference

Gini's mean difference is a classical measure of statistical dispersion, which serves as robust measure of variability of a probability distribution [7]. Recall that for a given distribution $\mathbb{Q}$, Gini's mean difference is defined as

$$MD := \mathbb{E}[|Y_1 - Y_2|],$$

where $Y_1, Y_2$ are independent random variables with distribution $\mathbb{Q}$. In the framework of i.i.d. observations $(Y_i)_{i\ge 1}$ the measure $MD$ is consistently estimated by the U-statistic $\frac{2}{n(n-1)} \sum_{1\le i<j\le n} |Y_i - Y_j|$. Gini's mean difference is connected to questions of stochasic dominance as shown by [27]. We refer to the recent paper [24] for the estimation theory for Gini's mean difference under long range dependence.

In the setting of continuous Itô semimartingales we conclude by Theorem A.3 that

$$U(H)_t^n \xrightarrow{\text{u.c.p.}} MD_t := m_1 \int_{[0,t]^2} |\sigma_{s_1}^2 + \sigma_{s_2}^2|^{1/2} ds_1 ds_2,$$

where the function $H$ is given by $H(x,y) = |x-y|$ and $m_p$ is the $p$-th absolute moment of $\mathcal{N}(0,1)$. In mathematical finance the quantity $MD_t$ may be viewed as an alternative measure of price variability, which is more robust to outliers than the standard quadratic variation $[X,X]_t$.

Formally, we cannot directly apply Theorem A.7 to obtain a weak limit theory for the statistic $U(H)_t^n$, since the function $H(x,y) = |x-y|$ is not differentiable and $H$ is not even in *each* component. Since $Y_1 - Y_2$ and $Y_1 + Y_2$ have the same distribution for centered independent normally distributed random variables $Y_1, Y_2$, the modification

$$\overline{H}(x,y) := \tfrac{1}{2}(|x-y| + |x+y|),$$

which is even in each component, has the same limit, i.e. $U(\overline{H})_t^n \xrightarrow{\text{u.c.p.}} MD_t$. Moreover, using sub-differential calculus and defining

$$\operatorname{grad} \overline{H}(x,y) := \tfrac{1}{2}\big(\operatorname{sign}(x-y) + \operatorname{sign}(x+y), \operatorname{sign}(x-y) + \operatorname{sign}(x+y)\big),$$

all the proof steps remain valid (we also refer to [2], who prove the central limit theorem for non-differentiable functions). Thus, by the assertion of Theorem A.7, we deduce the stable convergence

$$\sqrt{n}\Big(U(\overline{H})_t^n - MD_t\Big) \xrightarrow{st} L_t = \int_{\mathbb{R}^2} (|x_1 - x_2| + |x_1 + x_2|)\mathbb{G}(t, dx_1)F(t, dx_2),$$

where the stochastic fields $\mathbb{G}(t,x)$ and $F(t,x)$ are defined in Proposition A.5 and (A.9), respectively. Now, we follow the route proposed in § A.5 to obtain a standard central limit theorem. We compute the symmetrization $\widetilde{G}_1, \widetilde{G}_2$ of the functions $G_1, G_2$ defined at (A.34) and (A.35), respectively:

$$\begin{aligned}
\widetilde{G}_1(x_1,x_2,x_3) &= \tfrac{1}{6}\Big((|x_1-x_2| + |x_1+x_2|)(|x_1-x_3| + |x_1+x_3|) \\
&\quad + (|x_2-x_1| + |x_2+x_1|)(|x_2-x_3| + |x_2+x_3|) \\
&\quad + (|x_3-x_1| + |x_3+x_1|)(|x_3-x_2| + |x_3+x_2|)\Big), \\
\widetilde{G}_1(x_1,x_2;y_1,y_2) &= \tfrac{1}{4}\Big((|x_1-y_1| + |x_1+y_1|)(|x_2-y_2| + |x_2+y_2|) \\
&\quad + (|x_1-y_2| + |x_1+y_2|)(|x_2-y_1| + |x_2+y_1|)\Big).
\end{aligned}$$

Using these functions we construct the statistics $V_{1,t}^n$ and $V_{2,t}^n$ (see § A.5). Finally, for any fixed $t > 0$ we obtain a feasible central limit theorem

$$\frac{\sqrt{n}\big(U(\overline{H})_t^n - MD_t\big)}{\sqrt{V_{1,t}^n - V_{2,t}^n}} \xrightarrow{d} \mathcal{N}(0,1).$$

The latter enables us to construct confidence regions for mean difference statistic $MD_t$.

### $\mathbb{L}^p$-type tests for constant volatility

In this subsection we propose a new homoscedasticity test for the volatility process $\sigma^2$. Our main idea relies on a certain distance measure, which is related to $\mathbb{L}^p$-norms;

we refer to [12, 13] for similar testing procedures in the $\mathbb{L}^2$ case. Let us define

$$h(s_1,\ldots,s_d) := \sum_{i=1}^d \sigma_{s_i}^2, \qquad s_1,\ldots,s_d \in [0,1],$$

and consider a real number $p > 1$. Our test relies on the $\mathbb{L}^p$-norms

$$\|h\|_{\mathbb{L}^p} := \left(\int_{[0,1]^d} |h(\boldsymbol{s})|^p d\boldsymbol{s}\right)^{1/p}.$$

Observe the inequality $\|h\|_{\mathbb{L}^p} \geq \|h\|_{\mathbb{L}^1}$ and, when the process $h$ is continuous, equality holds if and only if $h$ is constant. Applying this intuition, we introduce a distance measure $\mathcal{M}^2$ via

$$\mathcal{M}^2 := \frac{\|h\|_{\mathbb{L}^p}^p - \|h\|_{\mathbb{L}^1}^p}{\|h\|_{\mathbb{L}^p}^p} \in [0,1].$$

Notice that a continuous process $\sigma^2$ is constant if and only if $\mathcal{M}^2 = 0$. Furthermore, the measure $\mathcal{M}^2$ provides a quantitative account of the deviation from the homoscedasticity hypothesis, as it takes values in $[0,1]$.

For simplicity of exposition we introduce an empirical analogue of $\mathcal{M}^2$ in the case $d = 2$. We define the functions

$$H_1(x) := \tfrac{1}{2}(|x_1 - x_2|^{2p} + |x_1 + x_2|^{2p}), \qquad H_2(x) := x_1^2 + x_2^2$$

with $x \in \mathbb{R}^2$. Notice that both functions are continuously differentiable and even in each component, hence they satisfy the assumptions of Theorems A.3 and A.7. In particular, Theorem A.3 implies the convergence in probability

$$U(H_1)_1^n \xrightarrow{\mathbb{P}} U(H_1)_1 = m_{2p}\|h\|_{\mathbb{L}^p}^p, \qquad U(H_2)_1^n \xrightarrow{\mathbb{P}} U(H_2)_1 = \|h\|_{\mathbb{L}^1},$$

where the constant $m_{2p}$ has been defined in the previous subsection. The main ingredient for a formal testing procedure is the following result.

**Proposition A.10.** *Assume that conditions of Theorem A.7 hold. Then we obtain the stable convergence*

$$\sqrt{n}\Big(U(H_1)_1^n - m_{2p}\|h\|_{\mathbb{L}^p}^p, U(H_2)_1^n - \|h\|_{\mathbb{L}^1}\Big) \tag{A.37}$$
$$\xrightarrow{st} 2\Big(\int_{\mathbb{R}^2} H_1(x_1,x_2)\mathbb{G}(1,dx_1)F(1,dx_2), \int_{\mathbb{R}^2} H_2(x_1,x_2)\mathbb{G}(1,dx_1)F(1,dx_2)\Big)$$

*Furthermore, the $\mathcal{F}$-conditional covariance matrix $V = (V_{ij})_{1\leq i,j\leq 2}$ of the limiting random variable is given as*

$$V_{ij} = \int_0^1 \bigg(\int_{\mathbb{R}} f_i(x)f_j(x)\varphi_{\sigma_s}(x)dx \tag{A.38}$$
$$- \Big(\int_{\mathbb{R}} f_i(x)\varphi_{\sigma_s}(x)dx\Big)\Big(\int_{\mathbb{R}} f_j(x)\varphi_{\sigma_s}(x)dx\Big)\bigg)ds$$

*with*

$$f_i(x) := 2\int_{\mathbb{R}} H_i(x,y)F(1,dy), \qquad i = 1,2.$$

*Proof.* As in the proof of Theorem A.7 we deduce that

$$\sqrt{n}(U(H_i)_1^n - U(H_i)_1) = L_1'^n(i) + o_{\mathbb{P}}(1), \qquad i = 1, 2,$$

where $L_1'^n(i)$ is defined via

$$L_1'^n(i) = 2 \int_{\mathbb{R}^2} H_i(x_1, x_2)\mathbb{G}_n(1, dx_1)F(1, dx_2).$$

Now, exactly as in steps (ii)–(v) of the proof of Proposition A.6, we conclude the joint stable convergence in (A.37). The $\mathcal{F}$-conditional covariance matrix $V$ is obtained from Remark A.2(iii) as in the beginning of § A.5. □

Let now $\mathcal{M}_n^2$ be the empirical analogue of $\mathcal{M}^2$, i.e.

$$\mathcal{M}_n^2 := \frac{m_{2p}^{-1}U(H_1)_1^n - (U(H_2)_1^n)^p}{m_{2p}^{-1}U(H_1)_1^n} \xrightarrow{\mathbb{P}} \mathcal{M}^2.$$

Observe the identities

$$\mathcal{M}_n^2 = r\Big(U(H_1)_1^n, U(H_2)_1^n\Big), \qquad \mathcal{M}^2 = r\Big(m_{2p}\|h\|_{\mathbb{L}^p}^p, \|h\|_{\mathbb{L}^1}\Big),$$

where $r(x, y) = 1 - m_{2p}\frac{y^p}{x}$. Applying Proposition A.10 and delta method for stable convergence, we conclude that $\sqrt{n}(\mathcal{M}_n^2 - \mathcal{M}^2)$ converges stably in law towards a mixed normal distribution with mean 0 and $\mathcal{F}$-conditional variance given by

$$v^2 := \nabla r\Big(m_{2p}\|h\|_{\mathbb{L}^p}^p, \|h\|_{\mathbb{L}^1}\Big) V \nabla r\Big(m_{2p}\|h\|_{\mathbb{L}^p}^p, \|h\|_{\mathbb{L}^1}\Big)^\star,$$

where the random variable $V \in \mathbb{R}^{2\times 2}$ is defined at (A.38).

For an estimation of $V$ we can proceed as in § A.5. Define the functions $G_1^{ij} : \mathbb{R}^3 \to \mathbb{R}$ and $G_2^{ij} : \mathbb{R}^4 \to \mathbb{R}$ by

$$\begin{aligned} G_1^{ij}(x_1, x_2, x_3) &= H_i(x_1, x_2)H_j(x_1, x_3) \\ G_2^{ij}(x_1, x_2, y_1, y_2) &= H_i(x_1, y_1)H_j(x_2, y_2), \quad i, j = 1, 2. \end{aligned}$$

Let further $\widetilde{G}_1^{ij}$ be the symmetrization of $G_1^{ij}$ and $\widetilde{G}_2^{ij}$ the symmetrization of $G_2^{ij}$ with respect to the $\boldsymbol{y}$-values. With

$$W_{ij} := \frac{4}{n}\binom{n}{2}^{-1} \sum_{i_1=1}^{n-1} \sum_{1\le i_2<i_3\le n} \widetilde{G}_2^{ij}(\sqrt{n}\Delta_{i_1}^n X, \sqrt{n}\Delta_{i_1+1}^n X, \sqrt{n}\Delta_{i_2}^n X, \sqrt{n}\Delta_{i_3}^n X),$$

we can, exactly as in § A.5, deduce that

$$V^n := (4U(\widetilde{G}_1^{ij})_1^n - W_{ij})_{i,j=1,2} \xrightarrow{\mathbb{P}} V.$$

Using the previous results, we directly get

$$v_n^2 := \nabla r\Big(U(H_1)_1^n, U(H_2)_1^n\Big) V^n \nabla r\Big(U(H_1)_1^n, U(H_2)_1^n\Big)^\star \xrightarrow{\mathbb{P}} v^2.$$

Now, the properties of stable convergence yield the following feasible central limit theorem:

$$\frac{\sqrt{n}(\mathcal{M}_n^2 - \mathcal{M}^2)}{\sqrt{v_n^2}} \xrightarrow{d} \mathcal{N}(0,1). \tag{A.39}$$

With these formulas at hand we can derive a formal test procedure for the hypothesis

$$H_0: \ \sigma_s^2 \text{ is constant on } [0,1] \qquad \text{vs.} \qquad H_1: \ \sigma_s^2 \text{ is not constant on } [0,1].$$

These hypotheses are obviously equivalent to

$$H_0: \ \mathcal{M}^2 = 0, \qquad \text{vs.} \qquad H_1: \ \mathcal{M}^2 > 0.$$

Defining the test statistic $S_n$ via

$$S_n := \frac{\sqrt{n}\mathcal{M}_n^2}{\sqrt{v_n^2}},$$

we reject the null hypothesis at level $\gamma \in (0,1)$ whenever $S_t^n > c_{1-\gamma}$, where $c_{1-\gamma}$ denotes the $(1-\gamma)$-quantile of $\mathcal{N}(0,1)$. Now, (A.39) implies that

$$\lim_{n\to\infty} \mathbb{P}_{H_0}(S_n > c_{1-\gamma}) = \gamma, \qquad \lim_{n\to\infty} \mathbb{P}_{H_1}(S_n^n > c_{1-\gamma}) = 1.$$

In other words, our test statistic is consistent and keeps the level $\gamma$ asymptotically.

### Wilcoxon test statistic for structural breaks

Change-point analysis has been an active area of research for many decades (we refer to [6] for a comprehensive overview). The Wilcoxon statistic is a standard statistical procedure for testing structural breaks in location models. Let $(Y_i)_{1\le i\le n}$, $(Z_i)_{1\le i\le m}$ be mutually independent observations with $Y_i \sim \mathbb{Q}_{\theta_1}$, $Z_i \sim \mathbb{Q}_{\theta_2}$, where $\mathbb{Q}_\theta(A) = \mathbb{Q}_0(A-\theta)$ for all $A \in \mathcal{B}(\mathbb{R})$ and $\mathbb{Q}_0$ is a non-atomic probability measure. In this classical framework the Wilcoxon statistic is defined by

$$\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} \mathbb{1}_{\{Y_i \le Z_j\}}.$$

Under the null hypothesis $\theta_1 = \theta_2$, the test statistic is close to $1/2$, while deviations from this value indicate that $\theta_1 \neq \theta_2$. We refer to the recent work [8] for change-point tests for long-range dependent data.

Applying the same intuition we may provide a test statistic for structural breaks in the volatility process $\sigma^2$. Assume that the semimartingale $X$ is observed at high frequency on the interval $[0,1]$ and the volatility is constant on the intervals $[0,t)$ and $(t,1]$ for some $t \in (0,1)$, i.e. $\sigma_s^2 = \sigma_0^2$ on $[0,t)$ and $\sigma_s^2 = \sigma_1^2$ on $(t,1]$. Our aim is to test the null hypothesis $\sigma_0^2 = \sigma_1^2$ or to infer the change-point $t$ when $\sigma_0^2 \neq \sigma_1^2$. In this framework the Wilcoxon type statistic is defined via

$$WL_t^n := \frac{1}{n^2}\sum_{i=1}^{[nt]}\sum_{j=[nt]+1}^{n} \mathbb{1}_{\{|\Delta_i^n X| \le |\Delta_j^n X|\}}.$$

Notice that the kernel is neither symmetric nor continuous. Nevertheless, we deduce the following result.

**Proposition A.11.** *Assume that condition* (A.16) *holds. Then we obtain the convergence.*

$$WL_t^n \xrightarrow{\text{u.c.p.}} WL_t := \int_0^t \int_t^1 \left( \int_{\mathbb{R}^2} \mathbb{1}_{\{|\sigma_{s_1} u_1| \le |\sigma_{s_2} u_2|\}} \varphi_d(\boldsymbol{u}) d\boldsymbol{u} \right) ds_1 ds_2 \qquad \text{(A.40)}$$
$$= \int_0^t \int_t^1 \left( 1 - \frac{2}{\pi} \arctan \left| \frac{\sigma_{s_1}}{\sigma_{s_2}} \right| \right) ds_1 ds_2.$$

*Proof.* As in the proof of Theorem A.3, we first show the convergence (A.40) for the approximations $\alpha_i^n$ of the scaled increments $\sqrt{n}\Delta_i^n X$. We define

$$U_t'^n := \int_{\mathbb{R}^2} \mathbb{1}_{\{|x| \le |y|\}} F_n(t, dx)(F_n(1, dy) - F_n(t, dy)).$$

Since condition (A.16) holds, the measure $F_n(t, dx)$ is non-atomic. Hence, we conclude that

$$U_t'^n \xrightarrow{\text{u.c.p.}} WL_t$$

exactly as in the proof of Proposition A.2. It remains to prove the convergence

$$WL_t^n - U_t'^n \xrightarrow{\text{u.c.p.}} 0.$$

Observe the identity

$$WL_t^n - U_t'^n = \frac{1}{n^2} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} \left( \mathbb{1}_{\{|\Delta_i^n X| \le |\Delta_j^n X|\}} - \mathbb{1}_{\{|\alpha_i^n| \le |\alpha_j^n|\}} \right)$$
$$= \frac{1}{n^2} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} \Big( \mathbb{1}_{\{|\Delta_i^n X| \le |\Delta_j^n X|\}} - \mathbb{1}_{\{|\sqrt{n}\Delta_i^n X| \le |\alpha_j^n|\}} + \mathbb{1}_{\{|\sqrt{n}\Delta_i^n X| \le |\alpha_j^n|\}} - \mathbb{1}_{\{|\alpha_i^n| \le |\alpha_j^n|\}} \Big)$$

In the following we concentrate on proving that

$$\frac{1}{n^2} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} \left( \mathbb{1}_{\{|\sqrt{n}\Delta_i^n X| \le |\alpha_j^n|\}} - \mathbb{1}_{\{|\alpha_i^n| \le |\alpha_j^n|\}} \right) \xrightarrow{\text{u.c.p.}} 0,$$

as the other part is negligible by the same arguments. Using the identity

$$\mathbb{1}_{\{|\sqrt{n}\Delta_i^n X| \le |\alpha_j^n|\}} - \mathbb{1}_{\{|\alpha_i^n| \le |\alpha_j^n|\}} = \mathbb{1}_{\{|\sqrt{n}\Delta_i^n X| \le |\alpha_j^n|, |\alpha_i^n| > |\alpha_j^n|\}} - \mathbb{1}_{\{|\sqrt{n}\Delta_i^n X| > |\alpha_j^n|, |\alpha_i^n| \le |\alpha_j^n|\}}$$

we restrict our attention on proving

$$\frac{1}{n^2} \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} \mathbb{1}_{\{|\sqrt{n}\Delta_i^n X| > |\alpha_j^n|, |\alpha_i^n| \le |\alpha_j^n|\}} \xrightarrow{\text{u.c.p.}} 0.$$

For an arbitrary $q \in (0, 1/2)$, we deduce the inequality

$$\mathbb{E}[\mathbb{1}_{\{|\sqrt{n}\Delta_i^n X| > |\alpha_j^n|, |\alpha_i^n| \le |\alpha_j^n|\}}] \le \mathbb{E}\left[ \frac{|\sqrt{n}\Delta_i^n X - \alpha_i^n|^q}{||\alpha_j^n| - |\alpha_i^n||^q} \right]$$
$$\le \mathbb{E}[|\sqrt{n}\Delta_i^n X - \alpha_i^n|^{2q}]^{1/2} \mathbb{E}[||\alpha_j^n| - |\alpha_i^n||^{-2q}]^{1/2}$$

For a standard normal random variable $U$, and for any $x > 0, y \geq 0$, define

$$g_q(x,y) := \mathbb{E}[|x|U| - y|^{-2q}].$$

Since $2q < 1$, we have

$$\begin{aligned} g_q(x,y) &= \mathbb{E}[|x|U| - y|^{-2q}\mathbb{1}_{\{|x|U|-y|\leq 1\}}] + \mathbb{E}[|x|U| - y|^{-2q}\mathbb{1}_{\{|x|U|-y|>1\}}] \\ &\leq \int_{\mathbb{R}} |x|u| - y|^{-2q}\mathbb{1}_{\{|x|u|-y|\leq 1\}} du + 1 \leq \frac{C_q}{x} + 1 < \infty \end{aligned} \tag{A.41}$$

Due to assumption (A.16) and by a localization argument we can assume that $\sigma_t$ is uniformly bounded away from zero. Therefore and by (A.41) we obtain

$$\begin{aligned} \mathbb{E}[||\alpha_j^n| - |\alpha_i^n||^{-2q}] &= \mathbb{E}[\mathbb{E}[||\alpha_j^n| - |\alpha_i^n||^{-2q}|\mathcal{F}_{\frac{j-1}{n}}]] \\ &= \mathbb{E}[g_q(\sigma_{\frac{j-1}{n}}, \alpha_i^n)] \leq C_q < \infty. \end{aligned}$$

Hence,

$$\begin{aligned} &\frac{1}{n^2}\sum_{i=1}^{[nt]}\sum_{j=[nt]+1}^{n} \mathbb{E}[\mathbb{1}_{\{|\sqrt{n}\Delta_i^n X|>|\alpha_j^n|, |\alpha_i^n|\leq|\alpha_j^n|\}}] \\ &\leq \frac{C}{n^2}\sum_{i=1}^{[nt]}\sum_{j=[nt]+1}^{n} \mathbb{E}[|\sqrt{n}\Delta_i^n X - \alpha_i^n|^{2q}]^{1/2} \xrightarrow{\text{u.c.p.}} 0, \end{aligned}$$

where the last convergence follows as in (A.42). This completes the proof of Proposition A.11. $\square$

Now, observe that when the process $\sigma^2$ has no change-point at time $t \in (0,1)$ (i.e. $\sigma_0^2 = \sigma_1^2$) the limit at (A.40) is given by $WL_t = \frac{1}{2}t(1-t)$. Thus, under the null hypothesis $\sigma_0^2 = \sigma_1^2$, we conclude that $WL_t^n \xrightarrow{\text{u.c.p.}} \frac{1}{2}t(1-t)$. Since the time point $t \in (0,1)$ is unknown in general, we may use the test statistic

$$\sup_{t\in(0,1)} |WL_t^n - \tfrac{1}{2}t(1-t)|$$

to test for a possible change point. Large values of this quantity speak against the null hypothesis. On the other hand, under the alternative $\sigma_0^2 \neq \sigma_1^2$, the statistic $\hat{t}_n := \text{argsup}_{t\in(0,1)}|WL_t^n - \frac{1}{2}t(1-t)|$ provides a consistent estimator of the change-point $t \in (0,1)$. A formal testing procedure would rely on a stable central limit theorem for $WL_t^n$, which is expected to be highly complex, since the applied kernel is not differentiable.

## A.7 Proofs of some technical results

Before we start with the proofs of (A.13) and (A.33) we state the following Lemma, which can be shown exactly as [2, Lemma 5.4].

**Lemma A.12.** *Let $f : \mathbb{R}^d \to \mathbb{R}^q$ be a continuous function of polynomial growth. Let further $\gamma_i^n, \gamma_i'^n$ be real-valued random variables satisfying $\mathbb{E}[(|\gamma_i^n| + |\gamma_i'^n|)^p] \leq C_p$ for all $p \geq 2$ and*

$$\binom{n}{d}^{-1} \sum_{\boldsymbol{i} \in \mathcal{A}_t^n(d)} \mathbb{E}[\|\gamma_{\boldsymbol{i}}^n - \gamma_{\boldsymbol{i}}'^n\|^2] \to 0.$$

*Then we have for all $t > 0$:*

$$\binom{n}{d}^{-1} \sum_{\boldsymbol{i} \in \mathcal{A}_t^n(d)} \mathbb{E}[\|f(\gamma_{\boldsymbol{i}}^n) - f(\gamma_{\boldsymbol{i}}'^n)\|^2] \to 0.$$

Recall that we assume (A.4) without loss of generality; in §A.7 and §A.7 we further assume (A.17), i.e. all the involved processes are bounded.

## Proof of (A.13)

The Burkholder inequality yields that $\mathbb{E}[(|\sqrt{n}\Delta_i^n X| + |\alpha_i^n|)^p] \leq C_p$ for all $p \geq 2$. In view of the previous Lemma $U(H)^n - \widetilde{U}(H)^n \xrightarrow{\text{u.c.p.}} 0$ is a direct consequence of

$$\binom{n}{d}^{-1} \sum_{\boldsymbol{i} \in \mathcal{A}_t^n(d)} \mathbb{E}[\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X - \alpha_{\boldsymbol{i}}^n\|^2] \leq \frac{C}{n} \sum_{j=1}^{[nt]} \mathbb{E}[|\sqrt{n}\Delta_j^n X - \alpha_j^n|^2] \to 0 \tag{A.42}$$

as it is shown in [2, Lemma 5.3]. □

## Proof of (A.33)

We divide the proof into several steps.

**(i)** We claim that

$$\sqrt{n}(U(H)^n - \widetilde{U}(H)^n) - P^n(H) \xrightarrow{\text{u.c.p.}} 0$$

where

$$P_t^n(H) := \sqrt{n}\binom{n}{d}^{-1} \sum_{\boldsymbol{i} \in \mathcal{A}_t^n(d)} \nabla H(\alpha_{\boldsymbol{i}}^n)(\sqrt{n}\Delta_{\boldsymbol{i}}^n X - \alpha_{\boldsymbol{i}}^n).$$

Here, $\nabla H$ denotes the gradient of $H$. This can be seen as follows. Since the process $\sigma$ is itself a continuous Itô semimartingale we have

$$\mathbb{E}[|\sqrt{n}\Delta_i^n X - \alpha_i^n|^p] \leq C_p n^{-p/2} \tag{A.43}$$

for all $p \geq 2$. By the mean value theorem, for any $\boldsymbol{i} \in \mathcal{A}_t^n(d)$, there exists a random variable $\chi_{\boldsymbol{i}}^n \in \mathbb{R}^d$ such that

$$H(\sqrt{n}\Delta_{\boldsymbol{i}}^n X) - H(\alpha_{\boldsymbol{i}}^n) = \nabla H(\chi_{\boldsymbol{i}}^n)(\sqrt{n}\Delta_{\boldsymbol{i}}^n X - \alpha_{\boldsymbol{i}}^n)$$

with $\|\chi_{\boldsymbol{i}}^n - \alpha_{\boldsymbol{i}}^n\| \le \|\sqrt{n}\Delta_{\boldsymbol{i}}^n X - \alpha_{\boldsymbol{i}}^n\|$. Therefore, we have

$$
\begin{aligned}
&\mathbb{E}[\sup_{t\le T} |\sqrt{n}(U(H)_t^n - \widetilde{U}_t(H)^n) - P_t^n(H)|] \\
&\quad \le C\sqrt{n}\binom{n}{d}^{-1} \sum_{\boldsymbol{i}\in\mathcal{A}_T^n(d)} \mathbb{E}[\|(\nabla H(\chi_{\boldsymbol{i}}^n) - \nabla H(\alpha_{\boldsymbol{i}}^n)\|\|(\sqrt{n}\Delta_{\boldsymbol{i}}^n X - \alpha_{\boldsymbol{i}}^n)\|] \\
&\quad \le C\sqrt{n}\binom{n}{d}^{-1} \Big( \sum_{\boldsymbol{i}\in\mathcal{A}_T^n(d)} \mathbb{E}[\|(\nabla H(\chi_{\boldsymbol{i}}^n) - \nabla H(\alpha_{\boldsymbol{i}}^n))\|^2]\Big)^{1/2} \\
&\qquad \times \Big( \sum_{\boldsymbol{i}\in\mathcal{A}_T^n(d)} \mathbb{E}[\|(\sqrt{n}\Delta_{\boldsymbol{i}}^n X - \alpha_{\boldsymbol{i}}^n)\|^2]\Big)^{1/2} \\
&\quad \le C\Big\{\binom{n}{d}^{-1} \sum_{\boldsymbol{i}\in\mathcal{A}_T^n(d)} \mathbb{E}[\|(\nabla H(\chi_{\boldsymbol{i}}^n) - \nabla H(\alpha_{\boldsymbol{i}}^n))\|^2]\Big\}^{1/2} \to 0
\end{aligned}
$$

by (A.42) and Lemma A.12.

**(ii)** In this and the next step we assume that $H$ has compact support. Now we split $P_t^n$ up into two parts:

$$
P_t^n = \sqrt{n}\binom{n}{d}^{-1} \sum_{\boldsymbol{i}\in\mathcal{A}_t^n(d)} \nabla H(\alpha_{\boldsymbol{i}}^n) v_{\boldsymbol{i}}^n(1) + \sqrt{n}\binom{n}{d}^{-1} \sum_{\boldsymbol{i}\in\mathcal{A}_t^n(d)} \nabla H(\alpha_{\boldsymbol{i}}^n) v_{\boldsymbol{i}}^n(2), \quad \text{(A.44)}
$$

where $\sqrt{n}\Delta_{\boldsymbol{i}}^n X - \alpha_{\boldsymbol{i}}^n = v_{\boldsymbol{i}}^n(1) + v_{\boldsymbol{i}}^n(2)$ and $\boldsymbol{i} = (i_1, \dots, i_d)$, with

$$
\begin{aligned}
v_{i_k}^n(1) &= \sqrt{n}\Big(n^{-1} a_{\frac{i_k-1}{n}} \\
&\qquad + \int_{\frac{i_k-1}{n}}^{\frac{i_k}{n}} \{\tilde{\sigma}_{\frac{i_k-1}{n}}(W_s - W_{\frac{i_k-1}{n}}) + \tilde{v}_{\frac{i_k-1}{n}}(V_s - V_{\frac{i_k-1}{n}})\} dW_s\Big) \\
v_{i_k}^n(2) &= \sqrt{n}\Big( \int_{\frac{i_k-1}{n}}^{\frac{i_k}{n}} (a_s - a_{\frac{i_k-1}{n}}) ds + \int_{\frac{i_k-1}{n}}^{\frac{i_k}{n}} \Big\{ \int_{\frac{i_k-1}{n}}^{s} \tilde{a}_u du \\
&\qquad + \int_{\frac{i_k-1}{n}}^{s} (\tilde{\sigma}_{u-} - \tilde{\sigma}_{\frac{i_k-1}{n}}) dW_u + \int_{\frac{i_k-1}{n}}^{s} (\tilde{v}_{u-} - \tilde{v}_{\frac{i_k-1}{n}}) dV_u \Big\} dW_s \Big).
\end{aligned}
$$

We denote the first and the second summand on the right hand side of (A.44) by $S_t^n$ and $\widetilde{S}_t^n$, respectively. First, we show the convergence $\widetilde{S}^n \xrightarrow{\text{u.c.p.}} 0$. Since the first derivative of $H$ is of polynomial growth we have $\mathbb{E}[\|\nabla H(\alpha_{\boldsymbol{i}}^n)\|^2] \le C$ for all $\boldsymbol{i} \in \mathcal{A}_t^n(d)$. Furthermore, we obtain by using the Hölder, Jensen, and Burkholder inequalities

$$
\mathbb{E}[|v_{i_k}^n(2)|^2] \le \frac{C}{n^2} + \int_{\frac{i_k-1}{n}}^{\frac{i_k}{n}} (a_s - a_{\frac{[ns]}{n}})^2 + (\tilde{\sigma}_{s-} - \tilde{\sigma}_{\frac{[ns]}{n}})^2 + (\tilde{v}_{s-} - \tilde{v}_{\frac{[ns]}{n}})^2 ds.
$$

Thus, for all $t > 0$, we have

$$
\begin{aligned}
&\sqrt{n}\binom{n}{d}^{-1}\mathbb{E}\sum_{\boldsymbol{i}\in\mathcal{A}_t^n(d)}|\nabla H(\alpha_{\boldsymbol{i}}^n)v_{\boldsymbol{i}}^n(2)| \\
&\quad\le C\sqrt{n}\binom{n}{d}^{-1}\Big(\mathbb{E}\Big[\sum_{\boldsymbol{i}\in\mathcal{A}_t^n(d)}\|\nabla H(\alpha_{\boldsymbol{i}}^n)\|^2\Big]\Big)^{1/2}\Big(\mathbb{E}\Big[\sum_{\boldsymbol{i}\in\mathcal{A}_t^n(d)}\|v_{\boldsymbol{i}}^n(2)\|^2\Big]\Big)^{1/2} \\
&\quad\le C\Big(n\binom{n}{d}^{-1}\mathbb{E}\Big[\sum_{i_1,\dots,i_d=1}^{[nt]}(|v_{i_1}^n(2)|^2+\cdots+|v_{i_d}^n(2)|^2)\Big]\Big)^{1/2} \\
&\quad\le C\Big(\mathbb{E}\Big[\sum_{j=1}^{[nt]}|v_j^n(2)|^2\Big]\Big)^{1/2} \\
&\quad\le C\Big(n^{-1}+\mathbb{E}\int_0^t(a_s-a_{\frac{[ns]}{n}})^2+(\tilde\sigma_{s-}-\tilde\sigma_{\frac{[ns]}{n}})^2+(\tilde v_{s-}-\tilde v_{\frac{[ns]}{n}})^2ds\Big)^{1/2}\to 0
\end{aligned}
$$

by the dominated convergence theorem and $\widetilde{S}^n \xrightarrow{\text{u.c.p.}} 0$ readily follows.

**(iii)** To show $S^n \xrightarrow{\text{u.c.p.}} 0$ we use

$$
S_t^n=\sum_{k=1}^d\sqrt{n}\binom{n}{d}^{-1}\sum_{\boldsymbol{i}\in\mathcal{A}_t^n(d)}\partial_k H(\alpha_{\boldsymbol{i}}^n)v_{i_k}^n(1)=:\sum_{k=1}^d S_t^n(k)
$$

Before we proceed with proving $S^n(k) \xrightarrow{\text{u.c.p.}} 0$, for $k = 1, \dots, d$, we make two observations: First, by the Burkholder inequality, we deduce

$$
\mathbb{E}[|\sqrt{n}v_{i_k}^n(1)|^p]\le C_p,\quad \text{for all } p\ge 2, \tag{A.45}
$$

and second, for fixed $x \in \mathbb{R}^{d-k}$, and for all $\boldsymbol{i} = (i_1, \dots, i_k) \in \mathcal{A}_t^n(k)$, we have

$$
\mathbb{E}[\partial_k H(\alpha_{\boldsymbol{i}}^n,x)v_{i_k}^n(1)|\mathcal{F}_{\frac{i_k-1}{n}}]=0, \tag{A.46}
$$

since $\partial_k H$ is an odd function in its $k$-th component. Now, we will prove that

$$
\sqrt{n}n^{-k}\sum_{\boldsymbol{i}\in\mathcal{A}_t^n(k)}\partial_k H(\alpha_{\boldsymbol{i}}^n,x)v_{i_k}^n(1)\xrightarrow{\text{u.c.p.}}0, \tag{A.47}
$$

for any fixed $x \in \mathbb{R}^{d-k}$. From (A.46) we know that it suffices to show that

$$
\sum_{i_k=1}^{[nt]}\mathbb{E}\Big[\Big(\sum_{1\le i_1<\cdots<i_{k-1}<i_k}\chi_{i_1,\dots,i_k}\Big)^2\Big|\mathcal{F}_{\frac{i_k-1}{n}}\Big]\xrightarrow{\mathbb{P}}0,
$$

where $\chi_{i_1,\dots,i_k} := \sqrt{n}n^{-k}\partial_k H(\alpha_{\boldsymbol{i}}^n, x)v_{i_k}^n(1)$. (Note that the sum in the expectation only runs over the indices $i_1, \dots, i_{k-1}$.) But this follows from the $L^1$ convergence and (A.45) via

$$\begin{aligned}
&\sum_{i_k=1}^{[nt]} \mathbb{E}\Big[\Big(\sum_{1\le i_1<\cdots<i_{k-1}<i_k} \chi_{i_1,\ldots,i_k}\Big)^2\Big] \\
&\qquad \le \frac{C}{n^k}\sum_{i_k=1}^{[nt]} \sum_{1\le i_1<\cdots<i_{k-1}<i_k} \mathbb{E}\big[(\partial_k H(\alpha_{\boldsymbol{i}}^n, x) v_{i_k}^n(1))^2\big] \\
&\qquad \le \frac{C}{n} \to 0.
\end{aligned}$$

Recall that we still assume that $H$ has compact support. Let the support of $H$ be a subset of $[-K,K]^d$ and further $-K = z_0 < \cdots < z_m = K$ be an equidistant partition of $[-K,K]$. We denote the set $\{z_0,\ldots,z_m\}$ by $Z_m$. Also, let $\eta(\varepsilon) := \sup\big\{\|\nabla H(\boldsymbol{x}) - \nabla H(\boldsymbol{y})\| \,\big|\, \|\boldsymbol{x}-\boldsymbol{y}\| \le \varepsilon\big\}$ be the modulus of continuity of $\nabla H$. Then we have

$$\begin{aligned}
\sup_{t\le T} |S_t^n(k)| &\le C\sqrt{n} n^{-k} \sup_{t\le T} \sup_{x\in[-K,K]^{d-k}} \Big| \sum_{\boldsymbol{i}\in\mathcal{A}_t^n(k)} \partial_k H(\alpha_{\boldsymbol{i}}^n, x) v_{i_k}^n(1)\Big| \\
&\le C\sqrt{n} n^{-k} \sup_{t\le T} \max_{x\in Z_m^{d-k}} \Big| \sum_{\boldsymbol{i}\in\mathcal{A}_t^n(k)} \partial_k H(\alpha_{\boldsymbol{i}}^n, x) v_{i_k}^n(1)\Big| \\
&\quad + C\sqrt{n} n^{-k} \sum_{\boldsymbol{i}\in\mathcal{A}_T^n(k)} \eta\left(\frac{2K}{m}\right) |v_{i_k}^n(1)|.
\end{aligned}$$

Observe that, for fixed $m$, the first summand converges in probability to 0 as $n\to\infty$ by (A.47). The second summand is bounded in expectation by $C\eta(2K/m)$ which converges to 0 as $m\to\infty$. This implies $S_t^n(k) \xrightarrow{\text{u.c.p.}} 0$ which finishes the proof of (A.33) for all $H$ with compact support.

**(iv)** Now, let $H \in C_p^1(\mathbb{R}^d)$ be arbitrary and $H_k$ be a sequence of functions in $C_p^1(\mathbb{R}^d)$ with compact support that converges pointwise to $H$ and fulfills $H = H_k$ on $[-k,k]^d$. In view of step (i) it is enough to show that

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbb{E}\Big[\sup_{t\le T}\Big|\sqrt{n}\binom{n}{d}^{-1} \sum_{\boldsymbol{i}\in\mathcal{A}_t^n(d)} \nabla(H-H_k)(\alpha_{\boldsymbol{i}}^n)(\sqrt{n}\Delta_{\boldsymbol{i}}^n X - \alpha_{\boldsymbol{i}}^n)\Big|\Big] = 0.$$

Since $H - H_k$ is of polynomial growth and by (A.43) we get

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\le T}\Big|\sqrt{n}\binom{n}{d}^{-1} \sum_{\boldsymbol{i}\in\mathcal{A}_t^n(d)} \nabla(H-H_k)(\alpha_{\boldsymbol{i}}^n)(\sqrt{n}\Delta_{\boldsymbol{i}}^n X - \alpha_{\boldsymbol{i}}^n)\Big|\Big] \\
&\qquad \le C\sqrt{n}\binom{n}{d}^{-1} \sum_{\boldsymbol{i}\in\mathcal{A}_T^n(d)} \mathbb{E}[\|\nabla(H-H_k)(\alpha_{\boldsymbol{i}}^n)\|\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X - \alpha_{\boldsymbol{i}}^n\|] \\
&\qquad \le C\binom{n}{d}^{-1} \sum_{\boldsymbol{i}\in\mathcal{A}_T^n(d)} \mathbb{E}\Big[\Big(\sum_{l=1}^d \mathbb{1}_{\{|\alpha_{i_l}^n|>k\}}\Big)^2 \|\nabla(H-H_k)(\alpha_{\boldsymbol{i}}^n)\|^2\Big]^{1/2} \le \frac{C}{k},
\end{aligned}$$

which finishes the proof. □

**Proof of** (A.26)

We can write

$$U(H)_t = \int_{[0,t]^d}\int_{\mathbb{R}^d} H(\boldsymbol{x})\varphi_{\sigma_{s_1}}(x_1)\cdots\varphi_{\sigma_{s_d}}(x_d)d\boldsymbol{x}d\boldsymbol{s}.$$

We also have

$$\overline{F}'_n(t,x) = \int_0^{\frac{[nt]}{n}} \varphi_{\sigma_{\frac{[ns]}{n}}}(x)ds,$$

where $\overline{F}'_n(t,x)$ denotes the Lebesgue density in $x$ of $\overline{F}_n(t,x)$ defined at (A.26). So we need to show that $P^n(H) \xrightarrow{\text{u.c.p.}} 0$, where

$$\begin{aligned}P^n_t(H) := \sqrt{n}\int_{[0,t]^d}\int_{\mathbb{R}^d} & H(\boldsymbol{x}) \\ & \times \big(\varphi_{\sigma_{s_1}}(x_1)\cdots\varphi_{\sigma_{s_d}}(x_d) - \varphi_{\sigma_{\frac{[ns_1]}{n}}}(x_1)\cdots\varphi_{\sigma_{\frac{[ns_d]}{n}}}(x_d)\big)d\boldsymbol{x}d\boldsymbol{s}.\end{aligned}$$

As previously we show the result first for $H$ with compact support.

**(i)** Let the support of $H$ be contained in $[-k,k]^d$. From [2, § 8] we know that, for fixed $x\in\mathbb{R}$, it holds that

$$\sqrt{n}\int_0^t \big(\varphi_{\sigma_s}(x) - \varphi_{\sigma_{\frac{[ns]}{n}}}(x)\big)ds \xrightarrow{\text{u.c.p.}} 0. \tag{A.48}$$

Also, with $\rho(z,x) := \varphi_z(x)$ we obtain, for $x,y\in[-k,k]$,

$$\begin{aligned}&\Big|\int_0^t (\varphi_{\sigma_s}(x) - \varphi_{\sigma_{\frac{[ns]}{n}}}(x)) - (\varphi_{\sigma_s}(y) - \varphi_{\sigma_{\frac{[ns]}{n}}}(y))ds\Big| \\ &\quad\le \int_0^t \big|\partial_1\rho(\xi_s,x)(\sigma_s - \sigma_{\frac{[ns]}{n}}) - \partial_1\rho(\xi'_s,y)(\sigma_s - \sigma_{\frac{[ns]}{n}})\big|ds \\ &\quad\le \int_0^t \big|\partial_{11}\rho(\xi''_s,\eta_s)(\xi_s - \xi'_s) + \partial_{21}\rho(\xi''_s,\eta_s)(x-y)\big|\big|\sigma_s - \sigma_{\frac{[ns]}{n}}\big|ds \\ &\quad\le C\int_0^t |\sigma_s - \sigma_{\frac{[ns]}{n}}|^2 + |\sigma_s - \sigma_{\frac{[ns]}{n}}||y-x|ds,\end{aligned}$$

where $\xi_s, \xi'_s, \xi''_s$ are between $\sigma_s$ and $\sigma_{[ns]/n}$ and $\eta_s$ is between $x$ and $y$. Now, let $Z_m = \{jk/m \mid j = -m,\ldots,m\}$. Then, we get

$$\begin{aligned}\sup_{t\le T}|P^n_t(H)| &\le C_T\sup_{t\le T}\sqrt{n}\int_{[-k,k]}\Big|\int_0^t \varphi_{\sigma_s}(x) - \varphi_{\sigma_{\frac{[ns]}{n}}}(x)ds\Big|dx \\ &\le C_T\sup_{t\le T}\sup_{x\in[-k,k]}\sqrt{n}\Big|\int_0^t \varphi_{\sigma_s}(x) - \varphi_{\sigma_{\frac{[ns]}{n}}}(x)ds\Big| \\ &\le C_T\sup_{t\le T}\max_{x\in Z_m}\sqrt{n}\Big|\int_0^t \varphi_{\sigma_s}(x) - \varphi_{\sigma_{\frac{[ns]}{n}}}(x)ds\Big| \\ &\quad + C_T\sqrt{n}\int_0^T \big(|\sigma_s - \sigma_{\frac{[ns]}{n}}|^2 + \frac{k}{m}|\sigma_s - \sigma_{\frac{[ns]}{n}}|\big)ds\end{aligned}$$

$$
\begin{aligned}
&\leq C_T \sum_{x\in Z_m} \sup_{t\leq T} \sqrt{n} \Big| \int_0^t \varphi_{\sigma_s}(x) - \varphi_{\sigma_{\frac{[ns]}{n}}}(x) ds \Big| \\
&\quad + C_T \sqrt{n} \int_0^T \big( |\sigma_s - \sigma_{\frac{[ns]}{n}}|^2 + \frac{k}{m} |\sigma_s - \sigma_{\frac{[ns]}{n}}| \big) ds
\end{aligned}
$$

Observe that, for fixed $m$, the first summand converges in probability to 0 by (A.48). By the Itô isometry and (A.17) we get for the expectation of the second summand:

$$
\begin{aligned}
&\mathbb{E}\Big[ \sqrt{n} \int_0^T \big( |\sigma_s - \sigma_{\frac{[ns]}{n}}|^2 + \frac{k}{m} |\sigma_s - \sigma_{\frac{[ns]}{n}}| \big) ds \Big] \\
&\quad = \sqrt{n} \int_0^T \mathbb{E}\big[ |\sigma_s - \sigma_{\frac{[ns]}{n}}|^2 + \frac{k}{m} |\sigma_s - \sigma_{\frac{[ns]}{n}}| \big] ds \leq C_T \left( \frac{1}{\sqrt{n}} + \frac{1}{m} \right).
\end{aligned}
$$

Thus, by choosing $m$ large enough and then letting $n$ go to infinity, we obtain the convergence $P_t^n(H) \xrightarrow{\text{u.c.p.}} 0$.

**(ii)** Now let $H \in C_p^1(\mathbb{R}^d)$ and $H_k$ be an approximating sequence of functions in $C_p^1(\mathbb{R}^d)$ with compact support and $H = H_k$ on $[-k, k]^d$. Observe that, for $\boldsymbol{x}, \boldsymbol{s} \in \mathbb{R}^d$, we obtain by the mean value theorem that

$$
\begin{aligned}
&\mathbb{E}\Big[ \big| \varphi_{\sigma_{s_1}}(x_1) \cdots \varphi_{\sigma_{s_d}}(x_d) - \varphi_{\sigma_{\frac{[ns_1]}{n}}}(x_1) \cdots \varphi_{\sigma_{\frac{[ns_d]}{n}}}(x_d) \big| \Big] \\
&\qquad\qquad \leq \psi(\boldsymbol{x}) \sum_{i=1}^d \mathbb{E} |\sigma_{s_i} - \sigma_{\frac{[ns_i]}{n}}| \leq \frac{C}{\sqrt{n}} \psi(\boldsymbol{x}),
\end{aligned}
$$

where the function $\psi$ is exponentially decaying at $\pm\infty$. Thus

$$
\begin{aligned}
&\lim_{k\to\infty} \limsup_{n\to\infty} \mathbb{E}\Big[ \sup_{t\leq T} |P_t^n(H) - P_t^n(H_k)| \Big] \\
&\qquad \leq C_T \lim_{k\to\infty} \limsup_{n\to\infty} \int_{\mathbb{R}^d} |(H - H_k)(\boldsymbol{x})| \psi(\boldsymbol{x}) d\boldsymbol{x} = 0,
\end{aligned}
$$

which finishes the proof of (A.26). □

## Acknowledgements

We would like to thank Herold Dehling for his helpful comments.

## Bibliography


[1] D. J. Aldous and G. K. Eagleson. On mixing and stability of limit theorems. *Ann. Probab.*, 6(2):325, 1978.

[2] O. E. Barndorff-Nielsen, O. E. Graversen, J. Jacod, M. Podolskij, and N. Shephard. A central limit theorem for realised power and bipower variations of continuous semimartingales. In Yu. Kabanov, R. Liptser, and J. Stoyanov, editors, *From Stochastic Calculus to Mathematical Finance. Festschrift in Honour of A.N. Shiryaev*, pages 33–68. Springer, Heidelberg, 2006.

[3] E. Beutner and H. Zähle. Deriving the asymptotic distribution of u- and v-statistics of dependent data using weighted empirical processes. *Bernoulli*, 18(3):803–822, 2012.

[4] S. Borovkova, R. Burton, and H. Dehling. Consistency of the takens estimator for the correlation dimension. *Ann. Appl. Probab.*, 9:376–390, 1999.

[5] S. Borovkova, R. Burton, and H. Dehling. Limit theorems for functionals of mixing processes with applications to u-statistics and dimension estimation. *Trans. Amer. Math. Soc.*, 353:4261–4318, 2001.

[6] M. Csörgó and L. Horváth. *Limit theorems in change-point analysis*. J. Wiley & Sons, 97.

[7] H. A. David. Gini's mean difference rediscovered. *Biometrika*, 55:573–575, 1968.

[8] H. Dehling, A. Rooch, and M. Taqqu. Non-parametric change-point tests for long-range dependent data. *Scandinavian Journal of Statistics*, 40(1):153–173, 2013.

[9] H. Dehling and M. S. Taqqu. The empirical processes of some long-range dependent sequences with an application to u-statistics. *Ann. Statist.*, 17:1767–1783, 1989.

[10] H. Dehling and M. S. Taqqu. Bivariate symmetric statistics of long-range dependent observations. *J. Statist. Plann. Inference*, 28:153–165, 1991.

[11] M. Denker and G. Keller. On u-statistics and von mises' statistic for weakly dependent processes. *Z. Wahrsch. Verw. Gebiete*, 64:505–552, 1983.

[12] H. Dette and M. Podolskij. Testing the parametric form of the volatility in continuous time diffusion models - an empirical process approach. *Journal of Econometrics*, 143:56–73, 2008.

[13] H. Dette, M. Podolskij, and M. Vetter. Estimation of integrated volatility in continuous time financial models with applications to goodness-of-fit testing. *Scandinavian Journal of Statistics*, 33:259–278, 2006.

[14] L. Giraitis and M. S. Taqqu. Limit theorems for bivariate appell polynomials. i. central limit theorems. *Probab. Theory Related Fields*, 107:359–381, 1997.

[15] W. Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.*, 19:293–325, 1948.

[16] T. Hsing and W. B. Wu. On weighted u-statistics for stationary processes. *Ann. Probab.*, 32:1600–1631, 2004.

[17] J. Jacod. On continuous conditional gaussian martingales and stable convergence in law. *Sémin. probab. Strasbg*, XXXI:232–246, 1997.

[18] J. Jacod. Asymptotic properties of realized power variations and related functionals of semimartingales. *Stoch. Process. Appl.*, 118:517–559, 2008.

[19] J. Jacod, M. Podolskij, and M. Vetter. Limit theorems for moving averages of discretized processes plus noise. *Ann. Statist.*, 38(3):1478–1545, 2010.

[20] J. Jacod and P. Protter. *Discretization of Processes.* Springer, Berlin, 2012.

[21] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer, Berlin, 2nd edition, 2003.

[22] S. Kinnebrock and M. Podolskij. A note on the central limit theorem for bipower variation of general functions. *Stoch. Process. Appl.*, 118:1056–1070, 2008.

[23] V. S. Koroljuk and Yu. V. Borovskich. *Theory of $U$-Statistics.* Kluver, Dordrecht, 1994.

[24] C. Lévy-Leduc, H. Boistard, E. Moulines, M. S. Taqqu, and V. A. Reisen. Asymptotic properties of u-processes under long-range dependence. *Annals of Statistics*, 39(3):1399–1426, 2011.

[25] M. Podolskij and M. Vetter. Understanding limit theorems for semimartingales: a short survey. *Stat. Neerl.*, 64(3):329–351, 2010.

[26] A. Renyi. On stable sequences of events. *Sankhya A*, 25:293–302, 1963.

[27] S. Yitzhaki. Stochastic dominance, mean variance, and Gini's mean difference. *Amer. Econ. Rev.*, 71:178–185, 1982.

Paper B

# On U- and V-statistics for discontinuous Itô semimartingales

By Mark Podolskij, Christian Schmidt and Mathias Vetter

**Abstract**

In this paper we examine the asymptotic theory for U- and V-statistics of discontinuous Itô semimartingales that are observed at high frequency. For different types of kernel functions we show laws of large numbers and associated stable central limit theorems. In most of the cases the limiting process will be conditionally centered Gaussian. The structure of the kernel function determines whether the jump and/or the continuous part of the semimartingale contribute to the limit.

**Keywords:** high frequency data, limit theorems, semimartingales, stable convergence, U-statistics, high frequency data, limit theorems, semimartingales, stable convergence, U-statistics, V-statistics.

**AMS Subject Classification:** Primary 60F05, 60F15, 60F17; secondary 60G48, 60H05.

## B.1 Introduction

U- and V-statistics are classical objects in mathematical statistics. They were introduced in the works of Halmos [9], von Mises [22] and Hoeffding [10], who provided (amongst others) the first asymptotic results for the case that the underlying random variables are independent and identically distributed. Since then there was a lot of progress in this field and the results were generalized in various directions. Under weak dependency assumptions asymptotic results are for instance shown in Borovkova, Burton and Dehling [4], in Denker and Keller [8] or more recently in Leucht [16]. The case of long memory processes is treated in Dehling and Taqqu [5, 6] or in Lévy-Leduc et al. [17]. For a general overview we refer to the books of Serfling [21] and Lee [15]. The methods applied in the proofs are quite different. One way are decomposition techniques like the famous Hoeffding decomposition or Hermite expansion as for example in [5, 6, 17]. Another approach is to use empirical process theory (see e.g. [1, 18]). In Beutner and Zähle [2] this method was recently combined with a continuous mapping approach to give a unifying way to treat the asymptotic theory for both U- and V-statistics in the degenerate and non-degenerate case.

In this paper we are concerned with U- and V-statistics where the underlying data comes from a (possibly discontinuous) Itô semimartingale of the form

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t, \quad t \geq 0, \tag{B.1}$$

where $W$ is a standard Brownian motion, $(b_s)_{s\geq 0}$ and $(\sigma_s)_{s\geq 0}$ are stochastic processes and $J_t$ is some jump process that we will specify later. Semimartingales play an important role in stochastic analysis because they form a large class of integrators with respect to which the Itô integral can be defined. This is one reason why they are widely used in applications, for instance in mathematical finance. Since the seminal work of Delbaen and Schachermayer [7] it is known that under certain no arbitrage conditions asset price processes must be semimartingales. Those price processes are nowadays observed very frequently, say for example at equidistant time points $0, 1/n, \dots \lfloor nT \rfloor /n$ for a fixed $T \in \mathbb{R}$ and large $n$. It is therefore of great interest to infer as many properties of the semimartingale from the given data $X_0, X_{1/n}, \dots, X_{\lfloor nT \rfloor/n}$ as possible for large $n$. In particular we are interested in the limiting behavior when $n$ tends to infinity. This setting is known as high frequency or infill asymptotics and is an active field of research since the last decades. For a comprehensive account we refer to the book of Jacod and Protter [12].

In Podolskij, Schmidt and Ziegel [18] an asymptotic theory for U-statistics of continuous Itô semimartingales (i.e. $J_t \equiv 0$ in (B.1)) was developed in the high frequency setting, where a U-statistic of order $d$ is defined by

$$U(X,H)_t^n = \binom{n}{d}^{-1} \sum_{1 \leq i_1 < \dots < i_d \leq \lfloor nt \rfloor} H(\sqrt{n}\Delta_{i_1}^n X, \dots, \sqrt{n}\Delta_{i_d}^n X),$$

for some sufficiently smooth kernel function $H : \mathbb{R}^d \to \mathbb{R}$ and $\Delta_i^n X = X_{i/n} - X_{(i-1)/n}$. In [18] it is shown that $U(X,H)_t^n$ converges in probability to some functional of the volatility $\sigma$. Also an associated functional central limit theorem is given where the limiting process turned out to be conditionally Gaussian.

In this paper we will extend those results to the case of discontinuous Itô semimartingales $X$. A general problem when dealing with discontinuous processes is that, depending on the function $H$, the U-statistic defined above might not converge to any finite limit. Therefore, in this work, we will deal with the slightly different V-statistics, where by V-statistic of order $d$ we mean a statistic of the type

$$Y_t^n(H,X,l) = \frac{1}{n^l} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)} \sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)} H(\sqrt{n}\Delta_{\boldsymbol{i}}^n X, \Delta_{\boldsymbol{j}}^n X),$$

where $0 \leq l \leq d$ and

$$\mathcal{B}_t^n(k) = \big\{ \boldsymbol{i} = (i_1,\dots,i_k) \in \mathbb{N}^k \,\big|\, 1 \leq i_1,\dots,i_k \leq \lfloor nt \rfloor \big\} \qquad (k \in \mathbb{N}).$$

In the definition of $Y_t^n(H,X,l)$ we used a vector notation, that we will employ throughout the paper. For $\boldsymbol{s} = (s_1,\dots,s_d) \in \mathbb{R}^d$ and any stochastic process $(Z_s)_{s\in\mathbb{R}}$, we write

$$Z_{\boldsymbol{s}} = (Z_{s_1},\dots,Z_{s_d}).$$

Comparing the definitions of the U- and V-statistics we can see that they are of very similar type if $l = d$. In fact, for continuous $X$, both statistics will converge to the same limit if $H$ is symmetric. A major difference is the scaling inside the function $H$ whenever $l \neq d$. Already for $d = 1$ we can see why we need different scaling depending on the function $H$. In Jacod [11] the author considers (among others) the statistics

$$Y_t^n(H,X,1) = \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} H(\sqrt{n}\Delta_i^n X) \quad \text{and} \quad Y_t^n(H,X,0) = \sum_{i=1}^{\lfloor nt \rfloor} H(\Delta_i^n X)$$

for some function $H$ and $d = 1$. To give an example we consider the so-called power variation, for which $H_p(x) = |x|^p$. The results in [11] yield that, in the case $0 < p < 2$ and under some additional assumptions, we have

$$Y_t^n(H_p,X,1) \xrightarrow{\mathbb{P}} m_p \int_0^t |\sigma_s|^p ds,$$

where $m_p$ is the $p$-th absolute moment of a standard normal distribution. It follows then that $Y_t^n(H_p,X,0)$ would explode for this specific $H_p$. On the other hand, if $p > 2$, we have

$$Y_t^n(H_p,X,0) \xrightarrow{\mathbb{P}} \sum_{s\leq t} |\Delta X_s|^p, \tag{B.2}$$

where $\Delta X_s = \Delta X_s - \Delta X_{s-}$ stands for the jumps of $X$. Clearly this implies that $Y_t^n(H_p,X,1)$ diverges in this case. For the associated central limit theorems the assumptions need to be stronger. One needs to require $0 < p < 1$ for $Y_t^n(H_p,X,1)$ and $p > 3$ for $Y_t^n(H_p,X,0)$. The limiting processes will in most cases again be conditionally Gaussian (see Theorems B.3 and B.13 for more details). The structure however is quite different. In the $p < 1$ case the conditional variance of the limit will only depend on the continuous part of $X$, whereas in the $p > 3$ case the conditional variance is more complicated and depends on both the jump and the continuous part of $X$.

To accommodate these different behaviors we will in our general setting of V-statistics $Y_t^n(H,X,l)$ of order $d$ consider kernel functions of the form

$$H(x_1,\dots,x_l,y_1,\dots,y_{d-l}) = \prod_{i=1}^{l} |x_i|^{p_i} \prod_{j=1}^{d-l} |y_j|^{q_j} L(x_1,\dots,x_l,y_1,\dots,y_{d-l}),$$

where $L$ has to fulfill some boundedness conditions and needs to be sufficiently smooth. Further we assume $p_1,\dots,p_l < 2$ and $q_1,\dots,q_{d-l} > 2$. Clearly there are two special cases. If $l=0$ we need a generalization of (B.2) to V-statistics of higher order. If $l=d$ the V-statistic is of similar form as the U-statistic $U(X,H)_t^n$ defined above. In particular we need to extend the theory of U-statistics of continuous Itô semimartingales in [18] to the case of discontinuous Itô semimartingales. Finally, in the sophisticated situation of arbitrary $l$, we will combine the two special cases. The limiting processes in the central limit theorem will still be (in most cases) conditionally Gaussian.

The paper is organized as follows. In Section B.2 we give some basic definitions and the notation that we use. In Section B.3 we present a law of large numbers and a central limit theorem in the case $l=0$ for a slightly more general statistics than $Y_t^n(H,X,l)$. In Section B.4 we use those results in order to obtain, under stronger assumptions, again a law of large numbers and an associated central limit theorem for $Y_t^n(H,X,l)$ for general $l$. In the appendix we will prove, besides some technical results, a uniform central limit theorem for U-statistics, which can in some sense be seen as a generalization of the results for U-statistics given in [18] and which is crucial for the proof of the main Theorem B.13 in Section B.4.

## B.2 Preliminaries

Throughout the paper we will assume that for some $T \in \mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$ the process

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + (\delta \mathbb{1}_{\{|\delta|\leq 1\}}) * (\mathfrak{p} - \mathfrak{q})_t + (\delta \mathbb{1}_{\{|\delta|>1\}}) * \mathfrak{p}_t,$$

defined for $t \in [0,T]$, is a 1-dimensional Itô-semimartingale defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ that satisfies the usual assumptions. Further we require that $W$ is a Brownian motion and $\mathfrak{p}$ is a Poisson random measure with compensator $\mathfrak{q}(dt,dz) = dt \otimes \lambda(dz)$ for some $\sigma$-finite measure $\lambda$. Also we assume that $b$ is locally bounded and $\sigma$ is càdlàg and that $\delta$ is some predictable function on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$.

Moreover we will use the following vector notation: If $\boldsymbol{p} = (p_1,\dots,p_n), \boldsymbol{x} = (x_1,\dots,x_n) \in \mathbb{R}^n$, then we let $|\boldsymbol{x}|^{\boldsymbol{p}} := \prod_{k=1}^n |x_k|^{p_k}$. Define further $\boldsymbol{p} \leq \boldsymbol{x} \iff p_i \leq x_i$ for all $1 \leq i \leq n$. If $t \in \mathbb{R}$ we let $\boldsymbol{x} \leq t \iff x_i \leq t$ for all $1 \leq i \leq n$. Also we define $|\boldsymbol{x}|^t = |x_1|^t \cdots |x_n|^t$. By $\|\cdot\|$ we denote the maximum norm for vectors and the supremum norm for functions.

Finally we introduce the notation

$$\mathfrak{P}(l) := \Big\{ p : \mathbb{R}^l \to \mathbb{R} \,\Big|\, p(x_1,\dots,x_l) = \sum_{\boldsymbol{\alpha} \in A} |x_1|^{\alpha_1} \cdots |x_l|^{\alpha_l}, A \subset \mathbb{R}_+^l \text{ finite} \Big\}. \quad \text{(B.3)}$$

We will assume in the whole paper that $K$ is some generic constant that may change from line to line.

## B.3 The jump case

In this section we analyze the asymptotic behavior of the V-statistic $V(H,X,l)_t^n$ defined by

$$V(H,X,l)_t^n := \frac{1}{n^{d-l}} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d)} H(\Delta_{\boldsymbol{i}}^n X) = \frac{1}{n^{d-l}} Y_t^n(H,X,0) \tag{B.4}$$

for different types of continuous functions $H:\mathbb{R}^d\to\mathbb{R}$. As a toy example in the case $d=2$ serve the two kernel functions

$$H_1(x_1,x_2) = |x_1|^p \quad \text{and} \quad H_2(x_1,x_2) = |x_1x_2|^p$$

for some $p>2$. Already for these basic functions it is easy to see why there should be different rates of convergence, i.e. different $l$, in the law of large numbers. Consider

$$V(H_1,X,l)_t^n = \frac{\lfloor nt\rfloor}{n^{2-l}} \sum_{i=1}^{\lfloor nt\rfloor} |\Delta_i^n X|^p \quad \text{and} \quad V(H_2,X,l)_t^n = \frac{1}{n^{2-l}} \left( \sum_{i=1}^{\lfloor nt\rfloor} |\Delta_i^n X|^p \right)^2.$$

In order to get convergence in probability to some non-trivial limit we know from the 1-dimensional theory (see (B.2)) that we have to choose $l=1$ for $H_1$ and $l=2$ for $H_2$.

In the following two subsections we will provide a law of large numbers and an associated central limit theorem for the statistics defined at (B.4).

### Law of large numbers

For the law of large numbers we do not need to impose any additional assumptions on the process $X$. We only need to require that the kernel function $H$ fulfills (B.5), which is in the $d=1$ case the same condition that is given in [11]. We have

**Theorem B.1.** *Let $H:\mathbb{R}^d\to\mathbb{R}$ be continuous and $1\le l\le d$ such that*

$$\lim_{(x_1,\dots,x_l)\to\mathbf{0}} \frac{H(x_1,\dots,x_d)}{|x_1|^2\cdot\ldots\cdot|x_l|^2} = 0. \tag{B.5}$$

*Then, for fixed $t>0$,*

$$V(H,X,l)_t^n \xrightarrow{\mathbb{P}} V(H,X,l)_t := t^{d-l} \sum_{\boldsymbol{s}\in(0,t]^l} H(\Delta X_{\boldsymbol{s}},\mathbf{0}).$$

**Remark B.1.** Note that we can write $H$ in the form

$$H = |x_1\cdot\ldots\cdot x_l|^2 L(x_1,\dots,x_d),$$

where

$$L(x_1,\dots,x_d) = \begin{cases} \frac{H(x_1,\dots,x_d)}{|x_1\cdot\ldots\cdot x_l|^2}, & \text{if } x_1,\dots,x_l\neq 0,\\ 0, & \text{otherwise.}\end{cases}$$

By assumption (B.5), $L$ is continuous and consequently the limit $V(H,X,l)_t$ is well-defined, since the squared jumps of a semimartingale are absolutely summable.

**Remark B.2.** Condition (B.5) is stated in a somewhat asymmetric way because it only concerns the first $l$ arguments of $H$. Generally one should rearrange the arguments of $H$ in a way such that (B.5) is fulfilled for the largest possible $l$. In particular, $H(x_1,\dots,x_l,\mathbf{0})$ is not identically 0 then, which will lead to non-trivial limits.

*Proof of Theorem B.1.* Let $t>0$ be fixed. The proof will be divided into two parts. In the first one we will show that

$$\xi_t^n := \frac{1}{n^{d-l}} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d)} \left(H(\Delta_{\boldsymbol{i}}^n X) - H(\Delta_{i_1}^n X,\dots,\Delta_{i_l}^n X,\mathbf{0})\right) \xrightarrow{\mathbb{P}} 0.$$

Then we are left with proving the theorem in the case $l=d$, which will be done in the second part.

Since the paths of $X$ are càdlàg and therefore bounded on compacts by a constant $A_t(\omega)=\sup_{0\le s\le t}|X_s(\omega)|$, we have the estimate

$$\begin{aligned}|\xi_t^n| &\le \frac{1}{n^{d-l}} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d)} |\Delta_{i_1}^n X\cdot\ldots\cdot\Delta_{i_l}^n X|^2 \delta_{L,A_t}(\max(|\Delta_{i_{l+1}}^n X|,\dots,|\Delta_{i_d}^n X|))\\ &= \left(\sum_{i=1}^{\lfloor nt\rfloor}|\Delta_i^n X|^2\right)^l \frac{1}{n^{d-l}} \sum_{i_{l+1},\dots,i_d=1}^{\lfloor nt\rfloor} \delta_{L,A_t}(\max(|\Delta_{i_{l+1}}^n X|,\dots,|\Delta_{i_d}^n X|)),\end{aligned}$$

where

$$\delta_{L,A_t}(\epsilon) := \sup\left\{|L(\boldsymbol{x})-L(\boldsymbol{y})| \;\middle|\; \boldsymbol{x},\boldsymbol{y}\in[-2A_t,2A_t]^d, \|\boldsymbol{x}-\boldsymbol{y}\|<\epsilon\right\}, \quad \epsilon>0$$

denotes the modulus of continuity of $L$.

We will now use the elementary property of the càdlàg process $X$, that for every $\epsilon>0$, there exists $N\in\mathbb{N}$ such that $|\Delta_i^n X|<2\epsilon$ for all $n\ge N$, if $X$ does not have a jump of size bigger than $\epsilon$ on $\left(\frac{i-1}{n},\frac{i}{n}\right]$. Since the number of those jumps is finite, we obtain for sufficiently large $n$ the estimate

$$\frac{1}{n^{d-l}} \sum_{i_{l+1},\dots,i_d=1}^{\lfloor nt\rfloor} \delta_{L,A_t}(\max(|\Delta_{i_{l+1}}^n X|,\dots,|\Delta_{i_d}^n X|)) \le t^{d-l}\delta_{L,A_t}(2\epsilon) + \frac{K(\epsilon)}{n}.$$

Using the continuity of $L$, the left hand side becomes arbitrarily small, if we first choose $\epsilon$ small and then $n$ large. From [13] we know that

$$[X,X]_t^n := \sum_{i=1}^{\lfloor nt\rfloor}|\Delta_i^n X|^2 \xrightarrow{\mathbb{P}} [X,X]_t = \int_0^t \sigma_s^2\,ds + \sum_{0<s\le t}|\Delta X_s|^2, \tag{B.6}$$

and thus we obtain $\xi_t^n \xrightarrow{\mathbb{P}} 0$.

For the second part of the proof, i.e. the convergence $V(H,X,l)_t^n \xrightarrow{\mathbb{P}} V(H,X,l)_t$ in the case $l=d$, we define the functions $g_k^n:\mathbb{R}^{d-1}\to\mathbb{R}$ by

$$\begin{aligned}g_k^n(\boldsymbol{x}) = &\sum_{i=1}^{\lfloor nt\rfloor}|\Delta_i^n X|^2 L(x_1,\dots,x_{k-1},\Delta_i^n X,x_k,\dots,x_{d-1})\\ &-\sum_{s\le t}|\Delta X_s|^2 L(x_1,\dots,x_{k-1},\Delta X_s,x_k,\dots,x_{d-1})\end{aligned}$$

and deduce

$$|V(H,X,d)^n_t - V(H,X,d)_t| = \Big| \sum_{\boldsymbol{i}\in\mathcal{B}^n_t(d)} H(\Delta^n_{\boldsymbol{i}} X) - \sum_{\boldsymbol{s}\in[0,t]^d} H(\Delta X_{\boldsymbol{s}})\Big|$$

$$= \Big|\sum_{k=1}^d \Big\{ \sum_{\boldsymbol{i}\in\mathcal{B}^n_t(k)} \sum_{\boldsymbol{s}\in[0,t]^{d-k}} H(\Delta^n_{\boldsymbol{i}} X, \Delta X_{\boldsymbol{s}}) - \sum_{\boldsymbol{i}\in\mathcal{B}^n_t(k-1)} \sum_{\boldsymbol{s}\in[0,t]^{d-k+1}} H(\Delta^n_{\boldsymbol{i}} X, \Delta X_{\boldsymbol{s}})\Big\}\Big|$$

$$\le \sum_{k=1}^d ([X,X]^n_t)^{k-1}[X,X]^{d-k}_t \sup_{\|\boldsymbol{x}\|\le 2A_t} |g^n_k(\boldsymbol{x})|.$$

By using (B.6) again we see that it remains to show $\sup_{\|\boldsymbol{x}\|\le 2A_t} |g^n_k(\boldsymbol{x})| \xrightarrow{\mathbb{P}} 0$ for $1 \le k \le d$. In the following we replace the supremum by a maximum over a finite set and give sufficiently good estimates for the error that we make by doing so.

For any $m \in \mathbb{N}$ define the (random) finite set $A^m_t$ by

$$A^m_t := \Big\{ \frac{k}{m} \,\Big|\, k \in \mathbb{Z}, \frac{|k|}{m} \le 2A_t \Big\}.$$

Then we have

$$\sup_{\|\boldsymbol{x}\|\le 2A_t} |g^n_k(\boldsymbol{x})| \le \max_{\boldsymbol{x}\in(A^m_t)^{d-1}} |g^n_k(\boldsymbol{x})| + \sup_{\substack{\|\boldsymbol{x}\|,\|\boldsymbol{y}\|\le 2A_t\\ \|\boldsymbol{x}-\boldsymbol{y}\|\le 1/m}} |g^n_k(\boldsymbol{x}) - g^n_k(\boldsymbol{y})| =: \zeta^n_{k,1}(m) + \zeta^n_{k,2}(m).$$

Since the sets $A^m_t$ are finite, we immediately get $\zeta^n_{k,1}(m) \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$ from Remark 3.3.3 in [12] for any fixed $m$. For the second summand $\zeta^n_{k,2}(m)$ observe that

$$|\zeta^n_{k,2}(m)| \le \Big( \sum_{i=1}^{\lfloor nt \rfloor} |\Delta^n_i X|^2 + \sum_{s\le t} |\Delta X_s|^2 \Big) \delta_{L,A_t}(m^{-1}),$$

which implies

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(|\zeta^n_{k,2}(m)| > \epsilon) = 0 \quad \text{for every} \quad \epsilon > 0.$$

The proof is complete. □

## Central limit theorem

In this section we will show a central limit theorem that is associated to the law of large numbers in Theorem B.1. The mode of convergence will be the so-called stable convergence. This notion was introduced by Renyi [20] and generalized the concept of weak convergence. We say that a sequence $(Z_n)_{n\in\mathbb{N}}$ of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in a Polish space $(E, \mathcal{E})$ converges stably in law to a random variable $Z$ that is defined on an extension $(\Omega', \mathcal{F}', \mathbb{P}')$ of $(\Omega, \mathcal{F}, \mathbb{P})$ and takes also values in $(E, \mathcal{E})$, if and only if

$$\mathbb{E}(f(Z_n)Y) \to \mathbb{E}'(f(Z)Y) \quad \text{as} \quad n \to \infty$$

for all bounded and continuous $f$ and bounded, $\mathcal{F}$-measurable $Y$. We write $Z_n \xrightarrow{st} Z$ for stable convergence of $Z_n$ to $Z$. For a short summary of the properties of stable convergence we refer to [19]. The main property that we will use here is that if we have two sequences $(Y_n)_{n\in\mathbb{N}}, (Z_n)_{n\in\mathbb{N}}$ of real-valued random variables and real-valued random variables $Y, Z$ with $Y_n \xrightarrow{\mathbb{P}} Y$ and $Z_n \xrightarrow{st} Z$, then we automatically obtain joint convergence $(Z_n, Y_n) \xrightarrow{st} (Z, Y)$.

In contrast to the law of large numbers, we need to impose some assumptions on the jumps of the process $X$. We assume that $|\delta(\omega, t, z)| \wedge 1 \le \Gamma_n(z)$ for all $t \le \tau_n(\omega)$, where $\tau_n$ is an increasing sequence of stopping times going to infinity. The functions $\Gamma_n$ are assumed to fulfill

$$\int \Gamma_n(z)^2 \lambda(dz) < \infty.$$

Since the main result of this section, which is Theorem B.5, is stable under stopping, we may assume by a standard localization argument (see [12, Section 4.4.1]) that the following boundedness condition is satisfied:

$$|b_t| \le A, \quad |\sigma_t| \le A, \quad |X_t| \le A, \quad |\delta(t,z)| \le \Gamma(z) \le A$$

holds uniformly in $(\omega, t)$ for some constant $A$ and a function $\Gamma$ with

$$\int \Gamma(z)^2 \lambda(dz) \le A.$$

A common technique for proving central limit theorems for discontinuous semimartingales is to decompose the process $X$ for fixed $m \in \mathbb{N}$ into the sum of two processes $X(m)$ and $X'(m)$, where the part $X'(m)$ basically describes the jumps of $X$, which are of size bigger than $1/m$ and of whom there are only finitely many. Eventually one lets $m$ go to infinity.

So here we define $D_m = \{z \mid \Gamma(z) > 1/m\}$ and $(S(m,j))_{j\ge 1}$ to be the successive jump times of the Poisson process $\mathbb{1}_{\{D_m \setminus D_{m-1}\}} * \mathfrak{p}$. Let $(S_q)_{q\ge 1}$ be a reordering of $(S(m,j))$, and

$$\begin{aligned}
\mathcal{P}_m &= \{p \mid S_p = S(k,j) \text{ for } j \ge 1, k \le m\}, \\
\mathcal{P}_t^n(m) &= \left\{ p \in \mathcal{P}_m \,\middle|\, S_p \le \frac{\lfloor nt \rfloor}{n} \right\}, \\
\mathcal{P}_t(m) &= \{p \in \mathcal{P}_m \mid S_p \le t\}.
\end{aligned}$$

Further let

$$\begin{aligned}
R_-(n,p) &= \sqrt{n}(X_{S_p-} - X_{\frac{i-1}{n}}) \\
R_+(n,p) &= \sqrt{n}(X_{\frac{i}{n}} - X_{S_p}) \\
R(n,p) &= R_-(n,p) + R_+(n,p),
\end{aligned}$$

if $\frac{i-1}{n} < S_p \le \frac{i}{n}$. Now we split $X$ into a sum of $X(m)$ and $X'(m)$, where $X'(m)$ is the "big jump part" and $X(m)$ is the remaining term, by

$$\begin{aligned}
b(m)_t &= b_t - \int_{\{D_m \cap \{z \mid |\delta(t,z)| \le 1\}\}} \delta(t,z)\lambda(dz) \\
X(m)_t &= \int_0^t b(m)_s ds + \int_0^t \sigma_s dW_s + (\delta \mathbb{1}_{D_m^c}) * (\mathfrak{p} - \mathfrak{q})_t \\
X'(m) &= X - X(m) = (\delta \mathbb{1}_{D_m}) * \mathfrak{p}.
\end{aligned}$$

Further let $\Omega_n(m)$ denote the set of all $\omega$ such that the intervals $(\frac{i-1}{n}, \frac{i}{n}]$ $(1 \le i \le n)$ contain at most one jump of $X'(m)(\omega)$, and

$$|X(m)(\omega)_{t+s} - X(m)(\omega)_t| \le \frac{2}{m} \qquad \text{for all } t \in [0,T],\ s \in [0, n^{-1}].$$

Clearly, $\mathbb{P}(\Omega_n(m)) \to 1$, as $n \to \infty$.

Before we state the main result of this section we begin with some important lemmas. The first one gives useful estimates for the size of the increments of the process $X(m)$. For a proof see [12, (2.1.44) and (5.1.24)].

**Lemma B.2.** *For any $p \ge 1$ we have*

$$\mathbb{E}(|X(m)_{t+s} - X(m)_t|^p | \mathcal{F}_t) \le K(s^{(p/2)\wedge 1} + m^p s^p)$$

*for all $t \ge 0, s \in [0,1]$.*

As a simple application of the lemma we obtain for $p \ge 2$ and $\boldsymbol{i} \in \mathcal{B}_t^n(d)$ with $i_1 < \cdots < i_d$

$$\begin{aligned}
&\boldsymbol{E}\big[|\Delta_{i_1}^n X(m)|^p \cdot \ldots \cdot |\Delta_{i_d}^n X(m)|^p\big] \\
&\quad = \boldsymbol{E}\Big[\Delta_{i_1}^n X(m)|^p \cdot \ldots \cdot |\Delta_{i_{d-1}}^n X(m)|^p \boldsymbol{E}\big[|\Delta_{i_d}^n X(m)|^p \big| \mathcal{F}_{\frac{i_d-1}{n}}\big]\Big] \\
&\quad \le K\Big(\frac{1}{n} + \frac{m^p}{n^p}\Big)\boldsymbol{E}\big[|\Delta_{i_1}^n X(m)|^p \cdot \ldots \cdot |\Delta_{i_{d-1}}^n X(m)|^p\big] \le \cdots \le \frac{K(n,m)}{n^d}
\end{aligned}$$

for some positive sequence $K(n,m)$ which satisfies $\limsup_{n\to\infty} K(n,m) \le K$ for any fixed $m$. Consequently, for general $\boldsymbol{i} \in \mathcal{B}_t^n(d)$, we have

$$\boldsymbol{E}\big[|\Delta_{i_1}^n X(m)|^p \cdot \ldots \cdot |\Delta_{i_d}^n X(m)|^p\big] \le K(n,m) n^{-\#\{i_1,\ldots,i_d\}}.$$

Since the number of elements $\boldsymbol{i} = (i_1, \ldots, i_d) \in \mathcal{B}_t^n(d)$ with $\#\{i_1, \ldots, i_d\} = k$ is of order $n^k$, we obtain the useful formula

$$\mathbb{E}\Big[\sum_{\boldsymbol{i} \in \mathcal{B}_t^n(d)} |\Delta_{i_1}^n X(m)|^p \cdot \ldots \cdot |\Delta_{i_d}^n X(m)|^p\Big] \le K(n,m), \tag{B.7}$$

and similarly

$$\frac{1}{\sqrt{n}}\mathbb{E}\Big[\sum_{\boldsymbol{i} \in \mathcal{B}_t^n(d)} |\Delta_{i_1}^n X(m)|^p \cdot \ldots \cdot |\Delta_{i_{d-1}}^n X(m)|^p |\Delta_{i_d}^n X(m)|\Big] \le K(n,m). \tag{B.8}$$

The next lemma again gives some estimate for the process $X(m)$ and is central for the proof of Theorem B.5.

**Lemma B.3.** *Let $C > 0$ be a constant. Assume further that $f : \mathbb{R}\times[-C,C]^{d-1} \to \mathbb{R}$ is defined by $f(\boldsymbol{x}) = |x_1|^p g(\boldsymbol{x})$, where $p > 3$ and $g \in \mathcal{C}(\mathbb{R} \times [-C,C]^{d-1})$ is twice continuously differentiable in the first argument. Then we have*

$$\mathbb{E}\Big(\mathbb{1}_{\Omega_n(m)}\sqrt{n}\Big|\sum_{i=1}^{\lfloor nt \rfloor}\Big(f(\Delta_i^n X(m), x_2, \ldots, x_d) - \sum_{\frac{i-1}{n} < s \le \frac{i}{n}} f(\Delta X(m)_s, x_2, \ldots, x_d)\Big)\Big|\Big) \le \beta_m(t)$$

*for some sequence $(\beta_m(t))$ with $\beta_m(t) \to 0$ as $m \to \infty$, uniformly in $x_2, \ldots, x_d$.*

*Proof.* The main idea is to apply Itô formula to each of the summands and then estimate the expected value. For fixed $x_2, \dots, x_d$ this was done in [12, p. 132]. We remark that the proof there essentially relies on the following inequalities: For fixed $\boldsymbol{z} \in [-C, C]^{d-1}$ and $|x| \leq 1/m$ ($m \in \mathbb{N}$) there exists $\beta_m(\boldsymbol{z})$ such that $\beta_m(\boldsymbol{z}) \to 0$ as $m \to \infty$ and

$$|f(x, \boldsymbol{z})| \leq \beta_m(\boldsymbol{z})|x|^3, \ |\partial_1 f(x, \boldsymbol{z})| \leq \beta_m(\boldsymbol{z})|x|^2, \ |\partial_{11}^2 f(x, \boldsymbol{z})| \leq \beta_m(\boldsymbol{z})|x|. \tag{B.9}$$

Further, for $x, y \in \mathbb{R}$, define the functions

$$k(x, y, \boldsymbol{z}) = f(x + y, \boldsymbol{z}) - f(x, \boldsymbol{z}) - f(y, \boldsymbol{z}), \quad g(x, y, \boldsymbol{z}) = k(x, y, \boldsymbol{z}) - \partial_1 f(x, \boldsymbol{z})y.$$

Following [12] we obtain for $|x| \leq 3/m$ and $|y| \leq 1/m$ that

$$|k(x, y, \boldsymbol{z})| \leq K\beta_m(\boldsymbol{z})|x||y|, \ |g(x, y, \boldsymbol{z})| \leq K\beta_m(\boldsymbol{z})|x||y|^2. \tag{B.10}$$

Since $f$ is twice continuously differentiable in the first argument and $\boldsymbol{z}$ lies in a compact set, the estimates under (B.9) and (B.10) hold uniformly in $\boldsymbol{z}$, i.e. we can assume that the sequence $\beta_m(\boldsymbol{z})$ does not depend on $\boldsymbol{z}$, and hence the proof in [12] in combination with the uniform estimates implies the claim. □

At last we give a lemma that can be seen as a generalization of the fundamental theorem of calculus.

**Lemma B.4.** *Consider a function $f \in \mathcal{C}^d(\mathbb{R}^d)$. Then we have*

$$f(x) = f(0) + \sum_{k=1}^{d} \sum_{1 \leq i_1 < \cdots < i_k \leq d} \int_0^{x_{i_1}} \cdots \int_0^{x_{i_k}} \partial_{i_k} \cdots \partial_{i_1} f(g_{i_1,\dots,i_k}(s_1, \dots, s_k)) ds_k \dots ds_1,$$

*where $g_{i_1,\dots,i_k} : \mathbb{R}^k \to \mathbb{R}^d$ with*

$$(g_{i_1,\dots,i_k}(s_1, \dots, s_k))_j = \begin{cases} 0, & \text{if } j \notin \{i_1, \dots, i_k\} \\ s_l, & \text{if } j = i_l. \end{cases}$$

*Proof.* First write

$$f(x) = f(0) + \sum_{k=1}^{d} \big(f(x_1, \dots, x_k, 0, \dots, 0) - f(x_1, \dots, x_{k-1}, 0, \dots, 0)\big),$$

which yields

$$f(x) = f(0) + \sum_{k=1}^{d} \int_0^{x_k} \partial_k f(x_1, \dots, x_{k-1}, t, 0, \dots, 0) \ dt.$$

Now we can apply the first step to the function $g_t(x_1, \dots, x_{k-1}) := \partial_k f(x_1, \dots, x_{k-1}, t, 0, \dots, 0)$ in the integral and by doing this step iteratively we finally get the result. □

We still need some definitions before we can state the central limit theorem (see for comparison [12, p. 126]). For the definition of the limiting processes we introduce a second probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ equipped with sequences $(\psi_{n+})_{n\geq 1}, (\psi_{n-})_{n\geq 1}$, and $(\kappa_n)_{n\geq 1}$ of random variables, where all variables are independent, $\psi_{n\pm} \sim \mathcal{N}(0,1)$, and $\kappa_n \sim U([0,1])$. Let now $(T_n)_{n\geq 1}$ be a weakly exhausting sequence of stopping times for the jumps of $X$. We then finally define a very good filtered extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t\geq 0}, \tilde{\mathbb{P}})$ of the original space by

$$\tilde{\Omega} = \Omega \times \Omega', \quad \tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}', \quad \tilde{\mathbb{P}} = \mathbb{P} \otimes \mathbb{P}'.$$

The filtration $\tilde{\mathcal{F}}_t$ is chosen in such a way that it is the smallest filtration containing $\mathcal{F}_t$ and that $\kappa_n$ and $\psi_{n\pm}$ are $\tilde{\mathcal{F}}_{T_n}$-measurable. Further let

$$R_n = R_{n-} + R_{n+}, \quad \text{with} \quad R_{n-} = \sqrt{\kappa_n}\sigma_{T_n-}\psi_{n-}, \quad R_{n+} = \sqrt{1-\kappa_n}\sigma_{T_n}\psi_{n+}.$$

Also define the sets

$$\mathcal{A}_l(d) := \big\{ L \in \mathcal{C}^{d+1}(\mathbb{R}^d) \,\big|\, \lim_{\boldsymbol{y}\to 0} \partial_k L(\boldsymbol{x}, \boldsymbol{y}) = 0 \text{ for all } \boldsymbol{x} \in \mathbb{R}^l,\ k = l+1, \dots, d \big\}$$

for $l = 1, \dots, d$.

**Remark B.3.** The following properties hold:

(i) $\mathcal{A}_l(d) = \mathcal{C}^{d+1}(\mathbb{R}^d)$ for $l = d$.

(ii) If $f, g \in \mathcal{A}_l(d)$, then also $f+g, fg \in \mathcal{A}_l(d)$, i.e. $\mathcal{A}_l(d)$ is an algebra.

(iii) Let $f \in \mathcal{C}^{d+1}(\mathbb{R})$ with $f'(0) = 0$, then

$$L(x_1, \dots, x_d) = f(x_1 \cdot \ldots \cdot x_d) \quad \text{and} \quad L(x_1, \dots, x_d) = f(x_1) + \dots + f(x_d)$$

are elements of $\mathcal{A}_l(d)$ for all $1 \leq l \leq d$.

We obtain the following stable limit theorem.

**Theorem B.5.** *Let $1 \leq l \leq d$ and $H : \mathbb{R}^d \to \mathbb{R}$ with $H(\boldsymbol{x}) = |x_1|^{p_1} \cdot \ldots \cdot |x_l|^{p_l} L(\boldsymbol{x})$, where $p_1, \dots, p_l > 3$ and $L \in \mathcal{A}_l(d)$. For $t > 0$ it holds that*

$$\sqrt{n}\Big(V(H,X,l)_t^n - V(H,X,l)_t\Big)$$
$$\xrightarrow{st} U(H,X,l)_t := t^{d-l} \sum_{k_1,\dots,k_l : T_{k_1},\dots,T_{k_l} \leq t} \sum_{j=1}^{l} \partial_j H(\Delta X_{T_{k_1}}, \dots, \Delta X_{T_{k_l}}, \mathbf{0}) R_{n_j}.$$

*The limit is $\mathcal{F}$-conditionally centered with variance*

$$\mathbb{E}(U(H,X,l)_t^2 | \mathcal{F}) = \tfrac{1}{2} t^{2(d-l)} \sum_{s\leq t} \Big( \sum_{k=1}^{l} \bar{V}_k(H,X,l,\Delta X_s) \Big)^2 (\sigma_{s-}^2 + \sigma_s^2),$$

*where*

$$\bar{V}_k(H,X,l,y)$$
$$= \sum_{s_1,\dots,s_{k-1},s_{k+1},\dots,s_l \leq t} \partial_k H(\Delta X_{s_1}, \dots, \Delta X_{s_{k-1}}, y, \Delta X_{s_{k+1}}, \dots, \Delta X_{s_l}, \mathbf{0}).$$

*Furthermore the $\mathcal{F}$-conditional law does not depend on the choice of the sequence $(T_k)_{k\in\mathbb{N}}$ and $U(H,X,l)_t$ is $\mathcal{F}$-conditionally Gaussian, if $X$ and $\sigma$ do not have common jump times.*

**Remark B.4.** In the case $d = 1$ this result can be found in Jacod [11] (see Theorem 2.11 and Remark 2.14 therein). A functional version of the central limit theorem in the given form does not exist even for $d = 1$. For an explanation see Remark 5.1.3 in [12]. In order to obtain functional results one generally needs to consider the discretized sequence

$$\sqrt{n}\Big(V(H,X,l)_t^n - V(H,X,l)_{\lfloor nt\rfloor/n}\Big).$$

In the proof below we would have to show that all approximation steps hold in probability uniformly on compact sets (instead of just in probability), which seems to be out of reach with our methods. What we could show with our approach is that Theorem B.5 holds in the finite distribution sense in $t$.

**Remark B.5.** In the case that the limit is $\mathcal{F}$-conditionally Gaussian we can get a standard central limit theorem by just dividing by the square root of the conditional variance, i.e.

$$\frac{\sqrt{n}\big(V(H,X,l)_t^n - V(H,X,l)_t\big)}{\sqrt{\mathbb{E}(U(H,X,l)_t^2|\mathcal{F})}} \xrightarrow{d} \mathcal{N}(0,1).$$

Since the conditional variance is generally unknown, we might need to consistently estimate it in order to obtain a feasible central limit theorem.

*Proof.* In the appendix we will show that $U(H,X,l)_t$ is in fact well-defined and fulfills the aforementioned conditional properties. To simplify notations we will give a proof only for symmetric $L$ and $p_1 = \cdots = p_l = p$ for some $p > 3$. Note that in this case $H$ is symmetric in the first $l$ components, which implies

$$\partial_j H(x_1,\ldots,x_l,0,\ldots,0) = \partial_1 H(x_j,x_2,\ldots,x_{j-1},x_1,x_{j+1},\ldots,x_l,0,\ldots,0).$$

Therefore, we have for fixed $j$

$$\begin{aligned}
&\sum_{\substack{k_1,\ldots,k_l:\\ T_{k_1},\ldots,T_{k_l}\le t}} \partial_k H(\Delta X_{T_{k_1}},\ldots,\Delta X_{T_{k_l}},\mathbf{0})R_{k_j}\\
&\quad= \sum_{\substack{k_1,\ldots,k_l:\\ T_{k_1},\ldots,T_{k_l}\le t}} \partial_1 H\big(\Delta X_{T_{k_j}},\Delta X_{T_{k_2}},\ldots,\Delta X_{T_{k_{j-1}}},\\
&\qquad\qquad \Delta X_{T_{k_1}},\Delta X_{T_{k_{j+1}}},\ldots,\Delta X_{T_{k_l}},\mathbf{0}\big)R_{k_j}\\
&\quad= \sum_{\substack{k_1,\ldots,k_l:\\ T_{k_1},\ldots,T_{k_l}\le t}} \partial_1 H(\Delta X_{T_{k_1}},\ldots,\Delta X_{T_{k_l}},\mathbf{0})R_{k_1},
\end{aligned}$$

and thus the limit can be written as

$$U(H,X,l)_t = lt^{d-l} \sum_{k_1\ldots,k_l:T_{k_1},\ldots,T_{k_l}\le t} \partial_1 H(\Delta X_{T_{k_1}},\ldots,\Delta X_{T_{k_l}},0\ldots,0)R_{k_1}.$$

Later we will prove $\sqrt{n}(V(H,X,l)_{\frac{\lfloor nt\rfloor}{n}} - V(H,X,l)_t) \xrightarrow{\mathbb{P}} 0$ as $n\to\infty$, so it will be enough to show the discretized version of the central limit theorem, i.e.

$$\xi_t^n := \sqrt{n}(V(H,X,l)_t^n - V(H,X,l)_{\frac{\lfloor nt\rfloor}{n}}) \xrightarrow{st} U(H,X,l)_t. \tag{B.11}$$

For the proof of this result we will use a lot of decompositions and frequently apply the following claim.

**Lemma B.6.** *Let $(Z_n)_{n\in\mathbb{N}}$ be a sequence of random variables, where, for each $m\in\mathbb{N}$, we have a decomposition $Z_n = Z_n(m) + Z'_n(m)$. If there is a sequence $(Z(m))_{m\in\mathbb{N}}$ of random variables and a random variable $Z$ with*

$$Z_n(m) \xrightarrow[n\to\infty]{st} Z(m), \quad Z(m) \xrightarrow[m\to\infty]{\mathbb{P}} Z, \quad \text{and} \quad \lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{P}(|Z'_n(m)|>\eta)=0$$

*for all $\eta>0$, then*

$$Z_n \xrightarrow{st} Z.$$

*Proof.* For a proof of this result see [12, Prop. 2.2.4]. □

For the proof of (B.11) we will successively split $\xi_t^n$ into several terms and then apply Lemma B.6. As a first decomposition we use

$$\xi_t^n = \mathbb{1}_{\Omega_n(m)}\xi_t^n + \mathbb{1}_{\Omega\setminus\Omega_n(m)}\xi_t^n.$$

Since $\mathbb{P}(\Omega_n(m))\to 1$ as $n\to\infty$, the latter term converges to 0 almost surely as $n\to\infty$, so we can focus on the first summand, which we further decompose into

$$\mathbb{1}_{\Omega_n(m)}\xi_t^n = \mathbb{1}_{\Omega_n(m)}\Big(\zeta^n(m) + \sum_{k=0}^{l}\sum_{j=0}^{d-l}\big(\zeta_{k,j}^n(m)-\tilde{\zeta}_{k,j}^n(m)\big) - \sum_{k=1}^{l}\zeta_k^n(m)\Big) \tag{B.12}$$

with

$$\begin{aligned}
\zeta^n(m) &= \frac{\sqrt{n}}{n^{d-l}}\Big(\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d)} H(\Delta_{mathbfi}^n X(m)) \\
&\qquad - \lfloor nt\rfloor^{d-l}\sum_{u_1,\dots,u_l\le\frac{\lfloor nt\rfloor}{n}} H(\Delta X(m)_{u_1},\dots,\Delta X(m)_{u_l},\mathbf{0})\Big) \\
\zeta_{k,j}^n(m) &= \frac{\sqrt{n}}{n^{d-l}} \sum_{\substack{\boldsymbol{p},\in\mathcal{P}_t^n(m)^k\\ \boldsymbol{q}\in\mathcal{P}_t^n(m)^j}} {\sum_{\substack{\boldsymbol{i}\in\mathcal{B}_t^n(l-k)\\ \boldsymbol{r}\in\mathcal{B}_t^n(d-l-j)}}}' \\
&\qquad \beta_{k,j} H\Big(\Delta X_{S_{\boldsymbol{p}}} + \frac{R(n,\boldsymbol{p})}{\sqrt{n}}, \Delta_{\boldsymbol{i}}^n X(m), \Delta X_{S_{\boldsymbol{q}}} + \frac{R(n,\boldsymbol{q})}{\sqrt{n}}, \Delta_{\boldsymbol{r}}^n X(m)\Big) \\
\tilde{\zeta}_{k,j}^n(m) &= \frac{\sqrt{n}}{n^{d-l}} \sum_{\boldsymbol{p},\boldsymbol{q}\in\mathcal{P}_t^n(m)^{k\times j}} {\sum_{\substack{\boldsymbol{i}\in\mathcal{B}_t^n(l-k)\\ \boldsymbol{r}\in\mathcal{B}_t^n(d-l-j)}}}' \\
&\qquad \beta_{k,j} H\Big(\frac{R(n,\boldsymbol{p})}{\sqrt{n}}, \Delta_{\boldsymbol{i}}^n X(m), \frac{R(n,\boldsymbol{q})}{\sqrt{n}}, \Delta_{\boldsymbol{r}}^n X(m)\Big) \\
\zeta_k^n(m) &= \sqrt{n}\frac{\lfloor nt\rfloor}{n^{d-l}}^{d-l} \sum_{\boldsymbol{p}\in\mathcal{P}_t^n(m)^k} \sum_{u_{k+1},\dots,u_l\le\frac{\lfloor nt\rfloor}{n}} \\
&\qquad \binom{l}{k} H\Big(\Delta X_{S_{\boldsymbol{p}}}, \Delta X_{u_{k+1}}(m),\dots,\Delta X_{u_l}(m),\mathbf{0}\Big),
\end{aligned}$$

where

$$\beta_{k,j} = \binom{l}{k}\binom{d-l}{j}.$$

The prime on the sums indicates that we sum only over those indices $\boldsymbol{i}$ and $\boldsymbol{r}$ such that $\Delta_{\boldsymbol{i}}^n X'(m)$ and $\Delta_{\boldsymbol{r}}^n X'(m)$ are vanishing, which in other word means that no big jumps of $X$ occur in the corresponding time intervals.

The basic idea behind the decomposition is that we distinguish between intervals $(\frac{i-1}{n}, \frac{i}{n}]$ where $X$ has a big jump and where not. Essentially we replace the original statistic $\xi_t^n$ by the same statistic $\zeta^n(m)$ for the process $X(m)$ instead of $X$. Using the trivial identity

$$\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d)} H(\Delta_{\boldsymbol{i}}^n X) = \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d)} H(\Delta_{\boldsymbol{i}}^n X(m)) + \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d)} \Big(H(\Delta_{\boldsymbol{i}}^n X) - H(\Delta_{\boldsymbol{i}}^n X(m))\Big)$$

we can see that an error term appears by doing this. Of course, we have $\Delta_{\boldsymbol{i}}^n X(m) = \Delta_{\boldsymbol{i}}^n X$ if no big jump occurs. In the decomposition above, $\zeta_{k,j}^n(m) - \tilde{\zeta}_{k,j}^n(m)$ gives the error term if we have $k$ big jumps in the first $l$ coordinates and $j$ big jumps in the last $d-l$ coordinates. In the same manner the term $\zeta_k^n(m)$ takes into account that we might have big jumps in $k$ arguments of $H(\Delta X_{u_1}, \ldots, \Delta X_{u_l}, \mathbf{0})$. All the binomial coefficients appear because of the symmetry of $H$ in the first $l$ and the last $d-l$ arguments. Note also that this decomposition is not valid without the indicator function $\mathbb{1}_{\Omega_n(m)}$.

In the appendix we will prove the following claim.

**Proposition B.7.** *It holds that*

$$\mathbb{1}_{\Omega_n(m)}\Big|\sum_{k=0}^{l}\sum_{j=0}^{d-l}\big(\zeta_{k,j}^n(m) - \tilde{\zeta}_{k,j}^n(m)\big) - \sum_{k=1}^{l}\zeta_k^n(m) - (\zeta_{l,0}^n(m) - \zeta_l^n(m))\Big| \xrightarrow{\mathbb{P}} 0$$

*if we first let $n \to \infty$ and then $m \to \infty$.*

So in view of Lemma B.6 we are left with considering the terms $\zeta_{l,0}^n(m) - \zeta_l^n(m)$ and $\zeta_n(m)$, where the first one is the only one that contributes to the limiting distribution. We will start with proving the three assertions

$$\lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{P}(\mathbb{1}_{\Omega_n(m)}|\zeta_{l,0}^n(m) - \hat{\zeta}_{l,0}^n(m)| > \eta) = 0 \quad \text{for all} \quad \eta > 0, \tag{B.13}$$

$$\mathbb{1}_{\Omega_n(m)}(\hat{\zeta}_{l,0}^n(m) - \zeta_l^n(m)) \xrightarrow{st} U(H, X'(m), l)_t, \quad \text{as } n \to \infty, \tag{B.14}$$

$$U(H, X'(m))_t \xrightarrow{\tilde{\mathbb{P}}} U(H, X, l)_t, \quad \text{as } m \to \infty, \tag{B.15}$$

where

$$\hat{\zeta}_{l,0}^n(m) := \frac{\sqrt{n}}{n^{d-l}} \sum_{\boldsymbol{p}\in\mathcal{P}_t^n(m)^l} \sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)} H\Big(\Delta X_{S_{\boldsymbol{p}}} + \frac{R(n,\boldsymbol{p})}{\sqrt{n}}, \mathbf{0}\Big).$$

For (B.13) observe that we have

$$\begin{aligned}
&\mathbb{1}_{\Omega_n(m)}|\zeta_{l,0}^n(m) - \hat{\zeta}_{l,0}^n(m)| \\
&\quad\le \mathbb{1}_{\Omega_n(m)} \sum_{\boldsymbol{p}\in\mathcal{P}_t(m)^l} \Big|\Delta X_{S_{\boldsymbol{p}}} + \frac{1}{\sqrt{n}}R(n,\boldsymbol{p})\Big|^p \frac{\sqrt{n}}{n^{d-l}} \\
&\qquad\qquad \times \sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)} \sum_{k=1}^{d-l} \sup_{\substack{\boldsymbol{x}\in[-2A,2A]^l \\ \boldsymbol{y}\in[-2/m,2/m]^{d-l}}} |\partial_k L(\boldsymbol{x},\boldsymbol{y})||\Delta_{j_k}^n X(m)| + O_{\mathbb{P}}(n^{-1/2})
\end{aligned}$$

by the mean value theorem. The error of small order in the estimate above is due to the finitely many large jumps, which are included in the sum over $\boldsymbol{j}$ now, but do not appear in $\zeta^n_{l,0}(m)$ by definition. Clearly,

$$\lim_{M\to\infty}\limsup_{m\to\infty}\limsup_{n\to\infty}\mathbb{P}\Big(\sum_{\boldsymbol{p}\in\mathcal{P}_t(m)^l}\Big|\Delta X_{S_{\boldsymbol{p}}}+\frac{1}{\sqrt{n}}R(n,\boldsymbol{p})\Big|^p > M\Big)=0,$$

and by Lemma B.2 we have

$$\begin{aligned}\mathbb{E}\Big(\frac{\sqrt{n}}{n^{d-l}}&\sum_{\boldsymbol{j}\in\mathcal{B}^n_t(d-l)}\sum_{k=1}^{d-l}\sup_{\boldsymbol{x}\in[-2A,2A]^l}\sup_{\boldsymbol{y}\in[-2/m,2/m]^{d-l}}|\partial_k L(\boldsymbol{x},\boldsymbol{y})||\Delta^n_{j_k}X(m)|\Big)\\ &\le K(1+mn^{-1/2})\sup_{\boldsymbol{x}\in[-2A,2A]^l}\sup_{\boldsymbol{y}\in[-2/m,2/m]^{d-l}}|\partial_k L(\boldsymbol{x},\boldsymbol{y})|,\end{aligned}$$

which converges to 0 if we first let $n\to\infty$ and then $m\to\infty$, since $L\in\mathcal{A}_l(d)$ and $[-2A,2A]^l$ is compact. This immediately implies (B.13).

For the proof of (B.14) we need another Lemma, which can be found in [12, Prop. 4.4.10].

**Lemma B.8.** *For fixed $p\in\mathbb{N}$ the sequence $(R(n,p))_{n\in\mathbb{N}}$ is bounded in probability, and*

$$(R(n,p)_-,R(n,p)_+)_{p\ge1}\xrightarrow{st}(R_{p-},R_{p+})_{p\ge1}$$

*as $n\to\infty$.*

Then we have, by the mean value theorem, Lemma B.8, the properties of stable convergence, and the symmetry of $H$ in the first $l$ components

$$\begin{aligned}&\mathbb{1}_{\Omega_n(m)}(\hat\zeta^n_{l,0}(m)-\zeta^n_l(m))\\ &=\sqrt{n}\mathbb{1}_{\Omega_n(m)}\Bigg(\frac{\lfloor nt\rfloor^{d-l}}{n^{d-l}}\sum_{\boldsymbol{p}\in\mathcal{P}^n_t(m)^l}\Big[H\Big(\Delta X_{S_{\boldsymbol{p}}}+\frac{1}{\sqrt{n}}R(n,\boldsymbol{p}),\boldsymbol{0}\Big)-H\Big(\Delta X_{S_{\boldsymbol{p}}},\boldsymbol{0}\Big)\Big]\Bigg)\\ &\xrightarrow{st}U(H,X'(m),l)_t=lt^{d-l}\sum_{\boldsymbol{p}\in\mathcal{P}_t(m)^l}\partial_1 H\Big(\Delta X_{S_{\boldsymbol{p}}},\boldsymbol{0}\Big)R_{p_1}\quad\text{as}\quad n\to\infty,\end{aligned}$$

i.e. (B.14). For the proof of (B.15) we introduce the notation $\mathcal{P}_t=\{p\in\mathbb{N}\,|\,S_p\le t\}$. We then use the decomposition

$$\begin{aligned}&U(H,X,l)_t-U(H,X'(m),l)_t\\ &\quad=lt^{d-l}\sum_{k=1}^{l}\sum_{\boldsymbol{p}\in\mathcal{P}_t^{k-1}}\sum_{p_k\in\mathcal{P}_t\setminus\mathcal{P}_t(m)}\sum_{\boldsymbol{r}\in\mathcal{P}_t(m)^{l-k}}\partial_1 H(\Delta X_{S_{\boldsymbol{p}}},\Delta X_{S_{p_k}},\Delta X_{S_{\boldsymbol{r}}},\boldsymbol{0})R_{p_1}\\ &\quad=:lt^{d-l}\sum_{k=1}^{l}\psi_k(m).\end{aligned}$$

We have to show that, for each $k$, $\psi_k(m)$ converges in probability to 0 as $m\to\infty$. We will give a proof only for the case $k=1$. Therefore, define the set

$$A(M):=\Big\{\omega\in\Omega\Big|\sum_{s\le t}(|\Delta X_s(\omega)|^p+|\Delta X_s(\omega)|^{2p}+|\Delta X_s(\omega)|^{2p-2})\le M\Big\},\quad M\in\mathbb{R}_+.$$

Then we have

$$\tilde{\mathbb{P}}(|\psi_1(m)| > \eta) \le \tilde{\mathbb{P}}(|\psi_1(m)|\mathbb{1}_{A(M)} > \eta/2) + \mathbb{P}(\Omega\backslash A(M)). \tag{B.16}$$

By the continuity of $L$ and $\partial_1 L$, and since the jumps of $X$ are uniformly bounded in $\omega$, we get

$$\begin{aligned}
\tilde{\mathbb{P}}(|\psi_1(m)|\mathbb{1}_{A(M)} > \eta/2) &\le K\mathbb{E}(\mathbb{1}_{A(M)}\tilde{\mathbb{E}}(\psi_1(m)^2|\mathcal{F})) \\
&\le K\mathbb{E}\Big(\mathbb{1}_{A(M)} \sum_{q\in\mathcal{P}_t\backslash\mathcal{P}_t(m)} \Big(\sum_{\boldsymbol{r}\in\mathcal{P}_t(m)^{l-1}} \partial_1 H(\Delta X_{S_q}, \Delta X_{S_{\boldsymbol{r}}}, 0,\ldots,0)\Big)^2\Big) \\
&\le K\mathbb{E}\Big(\mathbb{1}_{A(M)} \sum_{q\in\mathcal{P}_t\backslash\mathcal{P}_t(m)} (|\Delta X_{S_q}|^p + |\Delta X_{S_q}|^{p-1})^2 \Big(\sum_{r\in\mathcal{P}_t(m)} |\Delta X_{S_r}|^p\Big)^{2(l-1)}\Big) \\
&\le KM^{2(l-1)}\mathbb{E}\Big(\mathbb{1}_{A(M)} \sum_{q\in\mathcal{P}_t\backslash\mathcal{P}_t(m)} (|\Delta X_{S_q}|^{2p} + |\Delta X_{S_q}|^{2p-2})\Big) \to 0
\end{aligned}$$

as $m \to \infty$ by the dominated convergence theorem. Since the second summand in (B.16) is independent of $m$ and converges to 0 as $M \to \infty$, we have

$$\tilde{\mathbb{P}}(|\psi_1(m)| > \eta) \to 0 \quad \text{for all} \quad \eta > 0.$$

The proof for the convergence in probability of $\psi_k(m)$ to 0 for $2 \le k \le l$ is similar.

It remains to show that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(\mathbb{1}_{\Omega_n(m)}|\zeta^n(m)| > \eta) = 0 \tag{B.17}$$

for all $\eta > 0$.

Again, we need several decompositions. We have

$$\begin{aligned}
\zeta^n(m) &= \sqrt{n}\Big(\frac{1}{n^{d-l}} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d)} H(\Delta_{\boldsymbol{i}}^n X(m)) - \frac{\lfloor nt\rfloor^{d-l}}{n^{d-l}} \sum_{\boldsymbol{u}\in[0,n^{-1}\lfloor nt\rfloor]^l} H(\Delta X(m)_{\boldsymbol{u}}, \mathbf{0})\Big) \\
&= \sqrt{n}\Big(\frac{1}{n^{d-l}} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d)} H(\Delta_{\boldsymbol{i}}^n X(m)) - \frac{\lfloor nt\rfloor^{d-l}}{n^{d-l}} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)} H(\Delta_{\boldsymbol{i}}^n X(m), \mathbf{0})\Big) \\
&\quad + \sqrt{n}\Big(\frac{\lfloor nt\rfloor^{d-l}}{n^{d-l}} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)} H(\Delta_{\boldsymbol{i}}^n X(m), \mathbf{0}) - \frac{\lfloor nt\rfloor^{d-l}}{n^{d-l}} \sum_{\boldsymbol{u}\in[0,n^{-1}\lfloor nt\rfloor]^l} H(\Delta X(m)_{\boldsymbol{u}}, \mathbf{0})\Big) \\
&=: \Psi_1^n(m) + \Psi_2^n(m).
\end{aligned}$$

First observe that we obtain by the mean value theorem, and since $X$ is bounded,

$$\begin{aligned}
|\Psi_1^n(m)| &= \frac{\sqrt{n}}{n^{d-l}} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d)} |\Delta_{i_1}^n X(m)\cdots\Delta_{i_l}^n X(m)|^p \\
&\qquad \times |L(\Delta_{\boldsymbol{i}}^n X(m)) - L(\Delta_{i_1}^n X(m),\ldots,\Delta_{i_l}^n X(m), \mathbf{0})| \\
&\le K\frac{\sqrt{n}}{n^{d-l}} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d)} \sum_{k=l+1}^{d} |\Delta_{i_1}^n X(m)\cdots\Delta_{i_l}^n X(m)|^p |\Delta_{i_k}^n X(m)| \\
&= K(d-l)\frac{\sqrt{n}}{n^{d-l}} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d)} |\Delta_{i_1}^n X(m)\cdots\Delta_{i_l}^n X(m)|^p |\Delta_{i_{l+1}}^n X(m)| \\
&\le \frac{K(d-l)}{m^{(p-2)l}} \frac{1}{\sqrt{n}} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l+1)} |\Delta_{i_1}^n X(m)\cdots\Delta_{i_l}^n X(m)|^2 |\Delta_{i_{l+1}}^n X(m)|.
\end{aligned}$$

By (B.8) and $\limsup_{n\to\infty} K(m,n) \le K$ we get

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{E}(\mathbb{1}_{\Omega_n(m)} |\Psi_1^n(m)|) = 0.$$

When showing that $\Psi_2^n(m)$ converges to 0 we can obviously restrict ourselves to the case $l = d$. We need further decompositions:

$$\begin{aligned}\Psi_2^n(m) = \sqrt{n} \sum_{k=1}^d \Big( &\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(k)} \sum_{\boldsymbol{s}\in(0,\frac{\lfloor nt\rfloor}{n}]^{d-k}} H(\Delta_{\boldsymbol{i}}^n X(m), \Delta X(m)_{\boldsymbol{s}}) \\ &- \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(k-1)} \sum_{\boldsymbol{s}\in(0,\frac{\lfloor nt\rfloor}{n}]^{d-k+1}} H(\Delta_{\boldsymbol{i}}^n X(m), \Delta X(m)_{\boldsymbol{s}}) \Big) =: \sum_{k=1}^d \Psi_2^n(m,k).\end{aligned}$$

For a fixed $k$ we have

$$\begin{aligned}\Psi_2^n(m,k) = &\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(k-1)} |\Delta_{i_1}^n X(m) \cdots \Delta_{i_{k-1}}^n X(m)|^p \\ &\times \sum_{\boldsymbol{s}\in(0,\frac{\lfloor nt\rfloor}{n}]^{d-k}} |\Delta X(m)_{s_1} \cdots \Delta X(m)_{s_{d-k}}|^p \\ &\times \sqrt{n}\Big( \sum_{j=1}^{\lfloor nt\rfloor} |\Delta_j^n X(m)|^p L(\Delta_{\boldsymbol{i}}^n X(m), \Delta_j^n X(m), \Delta X(m)_{\boldsymbol{s}}) \\ &\qquad - \sum_{u\le\frac{\lfloor nt\rfloor}{n}} |\Delta X(m)_u|^p L(\Delta_{\boldsymbol{i}}^n X(m), \Delta X(m)_u, \Delta X(m)_{\boldsymbol{s}}) \Big),\end{aligned}$$

where we denote the latter factor by $\Theta_k^n(m,\boldsymbol{i},\boldsymbol{s})$. What causes problems here is that $\Theta_k^n(m,\boldsymbol{i},\boldsymbol{s})$ depends on the random variables $\Delta_{\boldsymbol{i}}^n X(m)$ and $\Delta X(m)_{\boldsymbol{s}}$ and we therefore cannot directly apply Lemma B.3. To overcome this problem we introduce the function $f_y \in \mathcal{C}^{d+1}(\mathbb{R}^{d-1})$ defined by

$$f_y(\boldsymbol{x}) = |y|^p L(x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_d).$$

Then we have

$$\begin{aligned}\Theta_k^n(m,\boldsymbol{i},\boldsymbol{s}) = \sqrt{n}\Big( &\sum_{j=1}^{\lfloor nt\rfloor} f_{\Delta_j^n X(m)}(\Delta_{\boldsymbol{i}}^n X(m), \Delta X(m)_{\boldsymbol{s}}) \\ &- \sum_{u\le\frac{\lfloor nt\rfloor}{n}} f_{\Delta X(m)_u}(\Delta_{\boldsymbol{i}}^n X(m), \Delta X(m)_{\boldsymbol{s}}) \Big).\end{aligned}$$

Now we replace the function $f_y$ according to Lemma B.4 by

$$\begin{aligned}f_y(\boldsymbol{x}) = f_y(\boldsymbol{0}) + \sum_{k=1}^d \sum_{1\le i_1<\cdots<i_k\le d} &\int_0^{x_{i_1}} \cdots \int_0^{x_{i_k}} \\ &\partial_{i_k} \cdots \partial_{i_1} f_y(g_{i_1,\dots,i_k}(s_1, \dots, s_k)) ds_k \dots ds_1.\end{aligned}$$

Since all of the appearing terms have the same structure we will exemplarily treat one of them:

$$\begin{aligned}
&\sqrt{n}\Big|\sum_{j=1}^{\lfloor nt\rfloor}\int_0^{\Delta X^n_{i_1}(m)}|\Delta_j^n X(m)|^p\partial_1 L(s_1,0,\dots,0,\Delta_j^n X(m),0,\dots,0)ds_1\\
&\quad-\sum_{u\leq\frac{\lfloor nt\rfloor}{n}}\int_0^{\Delta X^n_{i_1}(m)}|\Delta X(m)_u|^p\partial_1 L(s_1,0,\dots,0,\Delta X(m)_u,0,\dots,0)ds_1\Big|\\
&\quad\leq\int_{-\frac{2}{m}}^{\frac{2}{m}}\sqrt{n}\Big|\sum_{j=1}^{\lfloor nt\rfloor}|\Delta_j^n X(m)|^p\partial_1 L(s_1,0,\dots,0,\Delta_j^n X(m),0,\dots,0)\\
&\qquad\quad-\sum_{u\leq\frac{\lfloor nt\rfloor}{n}}|\Delta X(m)_u|^p\partial_1 L(s_1,0,\dots,0,\Delta X(m)_u,0,\dots,0)\Big|ds_1.
\end{aligned}$$

This means that we can bound $|\Theta_k^n(m,\boldsymbol{i},\boldsymbol{s})|$ from above by some random variable $\tilde{\Theta}_k^n(m)$ which is independent of $\boldsymbol{i}$ and $\boldsymbol{s}$ and which fulfills

$$\lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{E}\big[\mathbb{1}_{\Omega_n(m)}\tilde{\Theta}_k^n(m)\big]=0 \tag{B.18}$$

by Lemma B.3. Using the previous estimates we have

$$|\Psi_2^n(m,k)|\leq\tilde{\Theta}_k^n(m)\Bigg(\sum_{j=1}^{\lfloor nt\rfloor}|\Delta_j^n X(m)|^p\Bigg)^{k-1}\Bigg(\sum_{u\leq\frac{\lfloor nt\rfloor}{n}}|\Delta X(m)_u|^p\Bigg)^{d-k}.$$

Clearly the latter two terms are bounded in probability and therefore (B.18) yields

$$\lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{P}(\mathbb{1}_{\Omega_n(m)}|\Psi_2^n(m)|>\eta)=0,$$

which proves (B.17).

The last thing we have to show is

$$\sqrt{n}\Big(V(H,X,l)_t-V(H,X,l)_{\frac{\lfloor nt\rfloor}{n}}\Big)\xrightarrow{\mathbb{P}}0,$$

e.g. in the case $l=d$. From [12, p. 133] we know that in the case $d=1$ we have

$$\sqrt{n}\sum_{\frac{\lfloor nt\rfloor}{n}<s_k\leq t}|\Delta X_{s_k}|^p\xrightarrow{\mathbb{P}}0. \tag{B.19}$$

The general case follows by using the decomposition

$$\begin{aligned}
&\Big|\sqrt{n}\Big(\sum_{s_1,\dots,s_d\leq t}H(\Delta X_{s_1},\dots,\Delta X_{s_d})-\sum_{s_1,\dots,s_d\leq\frac{\lfloor nt\rfloor}{n}}H(\Delta X_{s_1},\dots,\Delta X_{s_d})\Big)\Big|\\
&\quad=\Big|\sqrt{n}\sum_{k=1}^{d}\Big(\sum_{s_1,\dots,s_{k-1}\leq t}\ \sum_{s_{k+1},\dots,s_d\leq\frac{\lfloor nt\rfloor}{n}}\ \sum_{\frac{\lfloor nt\rfloor}{n}<s_k\leq t}H(\Delta X_{s_1},\dots,\Delta X_{s_d})\Big)\Big|\\
&\quad\leq\sum_{k=1}^{d}\ \sum_{s_1,\dots,s_{k-1}\leq t}\ \sum_{s_{k+1},\dots,s_d\leq\frac{\lfloor nt\rfloor}{n}}\\
&\qquad\qquad|\Delta X_{s_1}\cdots\Delta X_{s_{k-1}}\Delta X_{s_{k+1}}\cdots\Delta X_{s_d}|^p\Big(\sqrt{n}\sum_{\frac{\lfloor nt\rfloor}{n}<s_k\leq t}|\Delta X_{s_k}|^p\Big),
\end{aligned}$$

which converges to 0 in probability. Hence the proof of Theorem B.5 is complete. □

As a possible application of the theory we want to indicate how one could obtain information about the jump sizes of the process X. For instance we will show how one could test if all the jump sizes lie on a grid $\alpha + \beta\mathbb{Z}$ for a given $\beta$, but unknown $\alpha$.

We start with a slightly more general situation and consider sets $M \subset \mathbb{R}$, for which we can find a non-negative function $g_M : \mathbb{R} \to \mathbb{R}$ that fulfills $g_M(x) = 0$ if and only if $x \in M$, and such that the function $L_M : \mathbb{R}^2 \to \mathbb{R}$ defined by $L_M(x,y) = g_M(x-y)$ lies in $\mathcal{A}_2(2)$. Then our theory gives for $H_M = |x|^{p_1}|y|^{p_2}L_M(x,y)$ that the limit $V(H_M, X, 2)$ in the law of large numbers vanishes if and only if there is $\alpha \in \mathbb{R}$ such that all jump sizes (which are not zero) lie in the set $\alpha + M$. In other words our theory enables us to give testing procedures to test if such an $\alpha$ exists. As a more explicit example we consider the following one.

**Example B.9.** For given $\beta \in \mathbb{R}$ consider the function

$$H(x,y) = |x|^4|y|^4 \sin^2\Big(\frac{\pi(x-y)}{\beta}\Big).$$

Then we have

$$\sum_{i,j=1}^{\lfloor nt \rfloor} H(\Delta_i^n X, \Delta_j^n X)$$
$$\xrightarrow{\mathbb{P}} L(\beta) := \sum_{s_1,s_2 \le t} |\Delta X_{s_1}|^4 |\Delta X_{s_2}|^4 \sin^2\Big(\frac{\pi(\Delta X_{s_1} - \Delta X_{s_2})}{\beta}\Big).$$

It holds that $L(\beta) = 0$ if and only if there exists an $\alpha \in \mathbb{R}$ such that

$$\Delta X_s \in \alpha + \beta\mathbb{Z} \quad \text{for all} \quad s \le t \quad \text{with} \quad \Delta X_s \neq 0.$$

To formally test whether there exists an $\alpha \in \mathbb{R}$ such that the jump sizes lie in the set $\alpha + \beta\mathbb{Z}$ one would of course need to derive estimators for the conditional variance of the limit in Theorem B.5.

## B.4 The mixed case

In this section we will present an asymptotic theory for statistics of the type

$$Y_t^n(H,X,l) = \frac{1}{n^l} \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l)} \sum_{\boldsymbol{j} \in \mathcal{B}_t^n(d-l)} H(\sqrt{n}\Delta_{\boldsymbol{i}}^n X, \Delta_{\boldsymbol{j}}^n X), \tag{B.20}$$

where $H$ behaves as $|x_1|^p \cdots |x_l|^p$ for $p < 2$ in the first $l$ arguments and in the last $d-l$ arguments as $|x_{l+1}|^q \cdots |x_d|^q$ for some $q > 2$ . As already mentioned in the introduction, powers $< 2$ and powers $> 2$ lead to completely different limits. This makes the treatment of $Y_t^n(H,X,l)$ for general $l$ way more complicated than in the case in which only large powers $> 2$ appear as in Section B.3. In fact we use the results from Section B.3 and combine them with quite general results concerning the case $l = d$, that we derive in the appendix. The limits turn out to be a mixture of what one obtains in the $p > 2$ and $p < 2$ setting. In the central limit theorem we get a conditionally Gaussian limit, where the conditional variance is a complicated functional of both the volatility $\sigma$ and the jumps of $X$.

## Law of large numbers

We will prove a law of large numbers for the quantity given in (B.20). As already mentioned we will need a combination of the methods from Section B.3 and methods for U-statistics of continuous Itô-semimartingales that were developed in [18]. We obtain the following result.

**Theorem B.10.** *Let $H(\boldsymbol{x},\boldsymbol{y}) = |x_1|^{p_1}\cdots|x_l|^{p_l}|y_1|^{q_1}\ \cdots|y_{d-l}|^{q_{d-l}}L(\boldsymbol{x},\boldsymbol{y})$ for some continuous function $L:\mathbb{R}^d\to\mathbb{R}$ with $p_1,\dots,p_l<2$ and $q_1,\dots,q_{d-l}>2$ for some $0\le l\le d$. The function $L$ is assumed to fulfill $|L(\boldsymbol{x},\boldsymbol{y})|\le u(\boldsymbol{y})$ for some $u\in\mathcal{C}(\mathbb{R}^{d-l})$. Then, for fixed $t>0$*

$$Y_t^n(H,X,l)\xrightarrow{\mathbb{P}}Y_t(H,X,l)=\sum_{\boldsymbol{s}\in[0,t]^{d-l}}\int_{[0,t]^l}\rho_H(\sigma_{\boldsymbol{u}},\Delta X_{\boldsymbol{s}})d\boldsymbol{u},$$

*where*

$$\rho_H(\boldsymbol{x},\boldsymbol{y})=\mathbb{E}[H(x_1U_1,\dots,x_lU_l,\boldsymbol{y})]$$

*for $\boldsymbol{x}\in\mathbb{R}^l,\boldsymbol{y}\in\mathbb{R}^{d-l}$, and $(U_1,\dots,U_l)\sim\mathcal{N}(\boldsymbol{0},\mathrm{id}_l)$.*

**Remark B.6.** We can see that in the special case $l=0$ we again obtain the result from Theorem B.1. For $l=d$ we basically get the same limit as the case of U-statistics for continuous semimartingales $X$ (see Theorem 3.3 in [18]).

*Proof.* By the standard localization procedure we may assume that $X$ and $\sigma$ are bounded by a constant $A$. We will start by proving the following two assertions:

$$\sup_{\boldsymbol{y}\in[-2A,2A]^{d-l}}\left|\frac{1}{n^l}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)}g(\sqrt{n}\Delta_{\boldsymbol{i}}^nX,\boldsymbol{y})-\int_{[0,t]^l}\rho_g(\sigma_{\boldsymbol{u}},\boldsymbol{y})d\boldsymbol{u}\right|\xrightarrow{\mathbb{P}}0,\tag{B.21}$$

$$\sup_{\boldsymbol{x}\in[-A,A]^l}\left|\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)}\rho_H(\boldsymbol{x},\Delta_{\boldsymbol{j}}^nX)-\sum_{\boldsymbol{s}\in[0,t]^{d-l}}\rho_H(\boldsymbol{x},\Delta X_{\boldsymbol{s}})\right|\xrightarrow{\mathbb{P}}0,\tag{B.22}$$

where $g(\boldsymbol{x},\boldsymbol{y})=|x_1|^{p_1}\cdots|x_l|^{p_l}L(\boldsymbol{x},\boldsymbol{y})$. The proofs mainly rely on the following decomposition for any real-valued function $f$ defined on some compact set $C\subset\mathbb{R}^n$: If $C'\subset C$ is finite and for any $\boldsymbol{x}\in C$ there exists $\boldsymbol{y}\in C'$ such that $\|\boldsymbol{x}-\boldsymbol{y}\|\le\delta$ for some $\delta>0$, then

$$\sup_{\boldsymbol{x}\in C}|f(\boldsymbol{x})|\le\max_{\boldsymbol{x}\in C'}|f(\boldsymbol{x})|+\sup_{\substack{\boldsymbol{x},\boldsymbol{y}\in C\\ \|\boldsymbol{x}-\boldsymbol{y}\|\le\delta}}|f(\boldsymbol{x})-f(\boldsymbol{y})|.$$

Now denote the continuous part of the semimartingale $X$ by $X^c$. For the proof of (B.21) we first observe that for fixed $\boldsymbol{y}\in\mathbb{R}^{d-l}$ we have

$$\frac{1}{n^l}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)}\left(g(\sqrt{n}\Delta_{\boldsymbol{i}}^nX,\boldsymbol{y})-g(\sqrt{n}\Delta_{\boldsymbol{i}}^nX^c,\boldsymbol{y})\right)\xrightarrow{\mathbb{P}}0.$$

We will not give a detailed proof of this "elimination of jumps" step since it follows essentially from the corresponding known case $l=1$ (see [12, Section 3.4.3]) in combination with the methods we use in the proof of (B.23). Using the results of the asymptotic theory for U-statistics of continuous Itô semimartingales given in [18, Prop. 3.2] we further obtain (still for fixed $\boldsymbol{y}$)

$$\frac{1}{n^l} \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l)} g(\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c, \boldsymbol{y}) \xrightarrow{\mathbb{P}} \int_{[0,t]^l} \rho_g(\sigma_{\boldsymbol{u}}, \boldsymbol{y}) d\boldsymbol{u}.$$

To complete the proof of (B.21) we will show

$$\xi^n(m) := \sup_{\substack{\boldsymbol{x},\boldsymbol{y} \in [-2A,2A]^{d-l} \\ \|\boldsymbol{x}-\boldsymbol{y}\| \le \frac{1}{m}}} \frac{1}{n^l} \Big| \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l)} \Big( g(\sqrt{n}\Delta_{\boldsymbol{i}}^n X, \boldsymbol{x}) - g(\sqrt{n}\Delta_{\boldsymbol{i}}^n X, \boldsymbol{y}) \Big) \Big| \xrightarrow{\mathbb{P}} 0 \quad \text{(B.23)}$$

if we first let $n$ and then $m$ go to infinity. The corresponding convergence of the integral term in (B.21) is easy and will therefore be omitted.

Let $\epsilon > 0$ be fixed such that $\max(p_1, \dots, p_l) + \epsilon < 2$, and for all $\alpha > 0$ and $k \in \mathbb{N}$ define the modulus of continuity

$$\delta_k(\alpha) := \sup\Big\{ |g(\boldsymbol{u}, \boldsymbol{x}) - g(\boldsymbol{u}, \boldsymbol{y})| \ \Big| \ \|\boldsymbol{u}\| \le k, \|(\boldsymbol{x}, \boldsymbol{y})\| \le 2A, \|\boldsymbol{x} - \boldsymbol{y}\| \le \alpha \Big\}.$$

Then we have

$$\begin{aligned} \xi^n(m) &\le K\Big( \delta_k(m^{-1}) + \sup_{\boldsymbol{x} \in [-2A,2A]^{d-l}} \frac{1}{n^l} \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l)} \mathbb{1}_{\{\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X\| \ge k\}} |g(\sqrt{n}\Delta_{\boldsymbol{i}}^n X, \boldsymbol{x})| \Big) \\ &\le K\Big( \delta_k(m^{-1}) + \frac{1}{n^l} \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l)} |\sqrt{n}\Delta_{i_1}^n X|^{p_1} \cdots |\sqrt{n}\Delta_{i_l}^n X|^{p_l} \sum_{j=1}^{l} \frac{|\sqrt{n}\Delta_{i_j}^n X|^\epsilon}{k^\epsilon} \Big) \\ &\xrightarrow{\mathbb{P}} K\Big( \delta_k(m^{-1}) + \frac{1}{k^\epsilon} \sum_{j=1}^{l} \prod_{i=1}^{l} \int_0^t m_{p_i + \delta_{ij}\epsilon} |\sigma_s|^{p_i + \delta_{ij}\epsilon} ds \Big) \quad \text{as} \quad n \to \infty, \end{aligned}$$

where $m_p$ is the $p$-th absolute moment of the standard normal distribution and $\delta_{ij}$ is the Kronecker delta (for a proof of the last convergence see [11, Theorem 2.4]) . The latter expression obviously converges to 0 if we let $m \to \infty$ and then $k \to \infty$, which completes the proof of (B.21).

We will prove (B.22) in a similar way. Since $\rho_H(\boldsymbol{x}, \boldsymbol{y})/|y_1 \cdot \ldots \cdot y_{d-l}|^2 \to 0$ as $\boldsymbol{y} \to 0$ , Theorem B.1 implies

$$\sum_{\boldsymbol{j} \in \mathcal{B}_t^n(d-l)} \rho_H(\boldsymbol{x}, \Delta_{\boldsymbol{j}}^n X) \xrightarrow{\mathbb{P}} \sum_{\boldsymbol{s} \in [0,t]^{d-l}} \rho_H(\boldsymbol{x}, \Delta X_{\boldsymbol{s}}),$$

i.e. pointwise convergence for fixed $\boldsymbol{x} \in [-A, A]^l$. Moreover,

$$\begin{aligned} &\sup_{\substack{\boldsymbol{x},\boldsymbol{y} \in [-A,A]^l \\ \|\boldsymbol{x}-\boldsymbol{y}\| \le \frac{1}{m}}} \sum_{\boldsymbol{j} \in \mathcal{B}_t^n(d-l)} \Big| \rho_H(\boldsymbol{x}, \Delta_{\boldsymbol{j}}^n X) - \rho_H(\boldsymbol{y}, \Delta_{\boldsymbol{j}}^n X) \Big| \\ &\qquad \le \Big( \prod_{i=1}^{d-l} \sum_{j=1}^{\lfloor nt \rfloor} |\Delta_j^n X|^{q_i} \Big) \sup_{\substack{\boldsymbol{x},\boldsymbol{y} \in [-A,A]^l \\ \|\boldsymbol{x}-\boldsymbol{y}\| \le \frac{1}{m}}} \ \sup_{\|\boldsymbol{z}\| \le 2A} \Big| \rho_g(\boldsymbol{x}, \boldsymbol{z}) - \rho_g(\boldsymbol{y}, \boldsymbol{z}) \Big|. \end{aligned}$$

The term in brackets converges in probability to some finite limit by Theorem B.1 as $n \to \infty$, and the supremum goes to 0 as $m \to \infty$ because $\rho_g$ is continuous. By similar arguments it follows that

$$\sup_{\substack{\boldsymbol{x},\boldsymbol{y} \in [-A,A]^l \\ \|\boldsymbol{x}-\boldsymbol{y}\| \le \frac{1}{m}}} \sum_{\boldsymbol{s} \in [0,t]^{d-l}} \Big| \rho_H(\boldsymbol{x}, \Delta X_{\boldsymbol{s}}) - \rho_H(\boldsymbol{y}, \Delta X_{\boldsymbol{s}}) \Big| \xrightarrow{\mathbb{P}} 0,$$

if we let $m$ go to infinity. Therefore (B.22) holds.

We will now finish the proof of Theorem B.10 in two steps. First we have

$$\begin{aligned} &\Big| \frac{1}{n^l} \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l)} \sum_{\boldsymbol{j} \in \mathcal{B}_t^n(d-l)} H(\sqrt{n}\Delta_{\boldsymbol{i}}^n X, \Delta_{\boldsymbol{j}}^n X) - \sum_{\boldsymbol{j} \in \mathcal{B}_t^n(d-l)} \int_{[0,t]^l} \rho_H(\sigma_{\boldsymbol{u}}, \Delta_{\boldsymbol{j}}^n X) d\boldsymbol{u} \Big| \\ &\quad \le \Big( \prod_{i=1}^{d-l} \sum_{j=1}^{\lfloor nt \rfloor} |\Delta_j^n X|^{q_i} \Big) \\ &\quad \times \sup_{\boldsymbol{y} \in [-2A,2A]^{d-l}} \Big| \frac{1}{n^l} \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l)} g(\sqrt{n}\Delta_{\boldsymbol{i}}^n X, \boldsymbol{y}) - \int_{[0,t]^l} \rho_g(\sigma_{\boldsymbol{u}}, \boldsymbol{y}) d\boldsymbol{u} \Big| \end{aligned}$$

and the latter expression converges in probability to 0 by (B.21). From (B.22) we obtain the functional convergence

$$\begin{pmatrix} (\sigma_s)_{0 \le s \le t} \\ \sum_{\boldsymbol{j} \in \mathcal{B}_t^n(d-l)} \rho_H(\cdot, \Delta_{\boldsymbol{j}}^n X) \end{pmatrix} \xrightarrow{\mathbb{P}} \begin{pmatrix} (\sigma_s)_{0 \le s \le t} \\ \sum_{\boldsymbol{s} \in [0,t]^{d-l}} \rho_H(\cdot, \Delta X_{\boldsymbol{s}}) \end{pmatrix}$$

in the space $\mathcal{D}([0,t]) \times \mathcal{C}([-A,A]^l)$. Define the mapping

$$\Phi : \mathcal{D}([0,t]) \times \mathcal{C}(\mathbb{R}^l) \to \mathbb{R}, \quad (f,g) \longmapsto \int_{[0,t]^l} g(f(u_1), \dots, f(u_l)) d\boldsymbol{u}.$$

This mapping is continuous and therefore we obtain by the continuous mapping theorem

$$\sum_{\boldsymbol{j} \in \mathcal{B}_t^n(d-l)} \int_{[0,t]^l} \rho_H(\sigma_{\boldsymbol{u}}, \Delta_{\boldsymbol{j}}^n X) d\boldsymbol{u} \xrightarrow{\mathbb{P}} \sum_{\boldsymbol{s} \in [0,t]^{d-l}} \int_{[0,t]^l} \rho_H(\sigma_{\boldsymbol{u}}, \Delta X_{\boldsymbol{s}}) d\boldsymbol{u},$$

which ends the proof. □

### Central limit theorem

In the mixed case we need some additional assumptions on the process $X$. First we assume that the volatility process $\sigma_t$ is not vanishing, i.e. $\sigma_t \neq 0$ for all $t \in [0,T]$, and that $\sigma$ is itself a continuous Itô-semimartingale of the form

$$\sigma_t = \sigma_0 + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \tilde{v}_s dV_s,$$

where $\tilde{b}_s, \tilde{\sigma}_s$, and $\tilde{v}_s$ are càdlàg processes and $V_t$ is a Brownian motion independent of $W$. As a boundedness condition on the jumps we further require that there is a sequence $\Gamma_n : \mathbb{R} \to \mathbb{R}$ of functions and a localizing sequence $(\tau_n)_{n\in\mathbb{N}}$ of stopping times such that $|\delta(\omega,t,z)| \wedge 1 \le \Gamma_n(z)$ for all $\omega, t$ with $t \le \tau_n(\omega)$ and

$$\int \Gamma_n(z)^r \lambda(dz) < \infty$$

for some $0 < r < 1$. In particular, the jumps of the process $X$ are then absolutely summable.

The central limit theorem will be stable under stopping, so we can assume without loss of generality that there is a function $\Gamma : \mathbb{R} \to \mathbb{R}$ and a constant $A$ such that $\delta(\omega,t,z) \le \Gamma(z)$ and

$$|X_t(\omega)|, \quad |b_t(\omega)|, \quad |\sigma_t(\omega)|, \quad |\sigma_t(\omega)^{-1}|, \quad |\tilde{b}_t(\omega)|, \quad |\tilde{\sigma}_t(\omega)|, \quad |\tilde{v}_t(\omega)| \le A,$$

uniformly in $(\omega, t)$. We may further assume $\Gamma(z) \le A$ for all $z \in \mathbb{R}$ and

$$\int \Gamma(z)^r \lambda(dz) < \infty.$$

Before we state the central limit theorem we give a few auxiliary results. At some stages in the proof of Theorem B.13 we will replace the scaled increments of $X$ in the first $l$ arguments by the first order approximation $\alpha_i^n := \sqrt{n}\sigma_{\frac{i-1}{n}}\Delta_i^n W$ of the continuous part of $X$. After several more approximations we will later see that $\sqrt{n}(Y_t^n(H,X,l) - Y_t(H,X,l))$ behaves asymptotically like

$$\sum_{\boldsymbol{q}:S_{\boldsymbol{q}}\le t} \Bigg( \frac{1}{n^l} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)} \sum_{k=l+1}^{d} \partial_k H\big(\alpha_{\boldsymbol{i}}^n, \Delta X_{S_{\boldsymbol{q}}}\big) R(n,q_k) \\ + \sqrt{n}\Big( \frac{1}{n^l} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)} H\big(\alpha_{\boldsymbol{i}}^n, \Delta X_{S_{\boldsymbol{q}}}\big) - \int_{[0,t]^l} \rho_H(\sigma_{\boldsymbol{s}}, \Delta X_{S_{\boldsymbol{q}}}) d\boldsymbol{s}\Big)\Bigg).$$

For now, consider only the term in brackets without $R(n,q_k)$. We can see that if $\Delta X_{S_{\boldsymbol{q}}}$ was just a deterministic number, we could derive the limit by using the asymptotic theory for U-statistics developed in [18]. For the first summand we would need a law of large numbers and for the second one a central limit theorem. Since $\Delta X_{S_{\boldsymbol{q}}}$ is of course in general not deterministic, the above decomposition indicates that it might be useful to have the theorems for U-statistics uniformly in some additional variables. As a first result in that direction we have the following claim.

**Proposition B.11.** *Let $0 \le l \le d$ and $G : \mathbb{R}^l \times [-A,A]^{d-l} \to \mathbb{R}$ be a continuous function that is of polynomial growth in the first $l$ arguments, i.e. $|G(\boldsymbol{x},\boldsymbol{y})| \le (1 + \|\boldsymbol{x}\|^p)w(\boldsymbol{y})$ for some $p \ge 0$ and $w \in \mathcal{C}([-A,A]^{d-l})$. Then*

$$\mathbb{B}_t^n(G,\boldsymbol{x}) := \frac{1}{n^l} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)} G\big(\alpha_{\boldsymbol{i}}^n, \boldsymbol{y}\big) \xrightarrow{\mathbb{P}} \mathbb{B}_t(G,\boldsymbol{y}) := \int_{[0,t]^l} \rho_G(\sigma_{\boldsymbol{s}}, \boldsymbol{y}) d\boldsymbol{s}$$

*in the space $\mathcal{C}([-A,A]^{d-l})$, where*

$$\rho_G(\boldsymbol{x},\boldsymbol{y}) := \mathbb{E}[G(x_1U_1, \ldots, x_lU_l, \boldsymbol{y})]$$

*for a standard normal variable $U = (U_1, \ldots, U_l)$.*

*Proof.* This result follows exactly in the same way as (B.21) without the elimination of jumps step in the beginning. $\square$

In addition to this functional law of large numbers we further need the associated functional central limit theorem for

$$\mathbb{U}_t^n(G,\boldsymbol{y}) = \sqrt{n}\Big(\frac{1}{n^l}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)} G\big(\alpha_{\boldsymbol{i}}^n,\boldsymbol{y}\big) - \int_{[0,t]^l}\rho_G(\sigma_{\boldsymbol{s}},\boldsymbol{y})d\boldsymbol{s}\Big), \tag{B.24}$$

In order to obtain a limit theorem we will need to show tightness and the convergence of the finite dimensional distributions. We will use that, for fixed $\boldsymbol{y}$, an asymptotic theory for (B.24) is given in [18, Prop. 4.3], but under too strong assumptions on the function $G$ for our purpose. In particular, we weaken the assumption of differentiability of $G$ in the following proposition whose proof can be found in the appendix.

**Proposition B.12.** *Let $0\le l\le d$ and let $G:\mathbb{R}^d\to\mathbb{R}$ be a function that is even in the first $l$ arguments and can be written in the form $G(\boldsymbol{x},\boldsymbol{y}) = |x_1|^{p_1}\cdots|x_l|^{p_l}L(\boldsymbol{x},\boldsymbol{y})$ for some function $L\in\mathcal{C}^{d+1}(\mathbb{R}^d)$ and constants $p_1,\dots,p_l\in\mathbb{R}$ with $0<p_1,\dots,p_l<1$. We further impose the following growth conditions:*

$$|L(\boldsymbol{x},\boldsymbol{y})|\le u(\boldsymbol{y}),\quad \Big|\partial_{ii}^2 L(\boldsymbol{x},\boldsymbol{y})\Big| \le (1+\|\boldsymbol{x}\|^{\beta_i})u(\boldsymbol{y})\quad (1\le i\le d),$$
$$\Big|\partial_{j_1}\cdots\partial_{j_k}L(\boldsymbol{x},\boldsymbol{y})\Big| \le (1+\|\boldsymbol{x}\|^{\gamma_{j_1\dots j_k}})u(\boldsymbol{y}),\quad (1\le k\le d;\ 1\le j_1<\dots<j_k\le d)$$

*for some constants $\beta_i,\gamma_{j_1\dots j_k}\ge 0$, and a function $u\in\mathcal{C}(\mathbb{R}^{d-l})$. The constants are assumed to fulfill $\gamma_j+p_i<1$ for $i\ne j$ and $i=1,\dots,l$, $j=1,\dots,d$. Then we have, for a fixed $t>0$*

$$(\mathbb{U}_t^n(G,\cdot),(R_-(n,p),R_+(n,p))_{p\ge1}) \xrightarrow{st} (\mathbb{U}_t(G,\cdot),(R_{p-},R_{p+})_{p\ge1})$$

*in the space $\mathcal{C}([-A,A]^{d-l})\times\mathbb{R}^{\mathbb{N}}\times\mathbb{R}^{\mathbb{N}}$, where $(\mathbb{U}_t(G,\cdot),(R_{p-},R_{p+})_{p\ge1})$ is defined on an extension $(\Omega',\mathcal{F}',\mathcal{P}')$ of the original probability space, $\mathbb{U}_t(G,\cdot)$ is $\mathcal{F}$-conditionally independent of $(\kappa_n,\psi_{n\pm})_{n\ge1}$ and $\mathcal{F}$-conditionally centered Gaussian with covariance structure*

$$\begin{aligned} C(\boldsymbol{y}_1,\boldsymbol{y}_2) &:= \mathbb{E}[\mathbb{U}_t(G,\boldsymbol{y}_1)\mathbb{U}_t(G,\boldsymbol{y}_2)|\mathcal{F}] \qquad\qquad (B.25)\\ &= \sum_{i,j=1}^{l}\int_0^t\Big(\int_{\mathbb{R}} f_i(u,\boldsymbol{y}_1)f_j(u,\boldsymbol{y}_2)\phi_{\sigma_s}(u)du - \Big(\prod_{k=1}^{2}\int_{\mathbb{R}} f_i(u,\boldsymbol{y}_k)\phi_{\sigma_s}(u)du\Big)ds\Big)\end{aligned}$$

*where*

$$f_i(u,\boldsymbol{y}) = \int_{[0,t]^{l-1}}\int_{\mathbb{R}^{l-1}} G(\sigma_{s_1}v_1,\dots,\sigma_{s_{i-1}}v_{i-1},u,\sigma_{s_{i+1}}v_{i+1},\dots,\sigma_{s_l}v_l,\boldsymbol{y})\phi(\boldsymbol{v})d\boldsymbol{v}d\boldsymbol{s}.$$

**Remark B.7.** The proposition is formulated for the approximations $\alpha_i^n$ of the increments of $X$. We remark that the result is still true in the finite dimensional distribution sense if we replace $\alpha_i^n$ by the increments $\Delta_i^n X$. This follows by the same arguments as the elimination of jumps step in Theorem B.13 and Proposition B.14.

We will now state the main theorem of this section. After some approximation steps the proof will mainly consist of an application of the methods from the previous section in combination with the continuous mapping theorem.

**Theorem B.13.** *Let $0 \leq l \leq d$ and $H : \mathbb{R}^d \to \mathbb{R}$ be a function that is even in the first $l$ arguments and can be written in the form $H(\boldsymbol{x}, \boldsymbol{y}) = |x_1|^{p_1} \cdots |x_l|^{p_l} \times |y_1|^{q_1} \cdots |y_{d-l}|^{q_{d-l}} L(\boldsymbol{x}, \boldsymbol{y})$ for some function $L \in \mathcal{C}^{d+1}(\mathbb{R}^d)$ and constants $p_1, \ldots, p_l$, $q_1, \ldots, q_{d-l} \in \mathbb{R}$ with $0 < p_1, \ldots, p_l < 1$ and $q_1, \ldots, q_{d-l} > 3$. We further assume that $L$ fulfills the same assumptions as in Proposition B.12. Then we have, for a fixed $t > 0$*

$$\sqrt{n}\Big(Y_t^n(H,X,l) - Y_t(H,X,l)\Big)$$
$$\xrightarrow{st} V'(H,X,l)_t = \sum_{\boldsymbol{k}: T_{\boldsymbol{k}} \leq t} \Big( \sum_{j=l+1}^{d} \int_{[0,t]^l} \rho_{\partial_j H}(\sigma_{\boldsymbol{u}}, \Delta X_{T_{\boldsymbol{k}}}) d\boldsymbol{u} R_{k_j} + \mathbb{U}_t(H, \Delta X_{T_{\boldsymbol{k}}}) \Big).$$

*The limiting process is $\mathcal{F}$-conditionally centered Gaussian with variance*

$$\mathbb{E}[(V'(H,X,l)_t)^2 | \mathcal{F}]$$
$$= \sum_{s \leq t} \Big( \sum_{k=l+1}^{d} \tilde{V}_k(H,X,l,\Delta X_s) \Big)^2 \sigma_s^2 + \sum_{\boldsymbol{s_1}, \boldsymbol{s_2} \in [0,t]^{d-l}} C(\Delta X_{\boldsymbol{s_1}}, \Delta X_{\boldsymbol{s_2}}),$$

*where the function $C$ is given in* (B.25) *and*

$$\tilde{V}_k(H,X,l,y)$$
$$= \sum_{\substack{s_{l+1},\ldots,s_{k-1} \leq t \\ s_{k+1},\ldots,s_d \leq t}} \int_{[0,t]^l} \rho_{\partial_k H}(\sigma_{\boldsymbol{u}}, \Delta X_{s_{l+1}}, \ldots, \Delta X_{s_{k-1}}, y, \Delta X_{s_{k+1}}, \ldots, \Delta X_{s_d}) d\boldsymbol{u}.$$

*Furthermore the $\mathcal{F}$-conditional law of the limit does not depend on the choice of the sequence $(T_k)_{k \in \mathbb{N}}$.*

**Remark B.8.** The result coincides with the central limit theorem in Section B.3 if $l = 0$, but under stronger assumptions. In particular the assumed continuity of $\sigma$ yields that the limit is always conditionally Gaussian. We further remark that the theorem also holds in the finite distribution sense in $t$.

*Proof.* In the first part of the proof we will eliminate the jumps in the first argument. We split $X$ into its continuous part $X^c$ and the jump part $X^d = \delta * \mathfrak{p}$ via $X = X_0 + X^c + X^d$. Note that $X^d$ exists since the jumps are absolutely summable under our assumptions. We will now show that

$$\xi_n = \frac{\sqrt{n}}{n^l} \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l)} \sum_{\boldsymbol{j} \in \mathcal{B}_t^n(d-l)} \big( H(\sqrt{n}\Delta_{\boldsymbol{i}}^n X, \Delta_{\boldsymbol{j}}^n X) - H(\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c, \Delta_{\boldsymbol{j}}^n X) \big) \xrightarrow{\mathbb{P}} 0.$$

Observe that under our growth assumptions on $L$ we can deduce

$$|L(\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}) - L(\boldsymbol{x}, \boldsymbol{y})| \leq K u(\boldsymbol{y})(1 + \sum_{i=1}^{l} \|\boldsymbol{x}\|^{\gamma_i}) \sum_{j=1}^{l} |z_j|^{p_j} \tag{B.26}$$

This inequality trivially holds if $\|\boldsymbol{z}\| > 1$ because $\|L(\boldsymbol{x},\boldsymbol{y})\| \le u(\boldsymbol{y})$. In the case $\|\boldsymbol{z}\| \le 1$ we can use the mean value theorem in combination with $|z| \le |z|^p$ for $|z| \le 1$ and $0 < p < 1$. Since we also have $\big||x_i + z_i|^{p_i} - |x_i|^{p_i}\big| \le |z_i|^{p_i}$ for $1 \le i \le l$, we have, with $\boldsymbol{q} = (q_1, \dots, q_{d-l})$, the estimate

$$|H(\boldsymbol{x}+\boldsymbol{z},\boldsymbol{y}) - H(\boldsymbol{x},\boldsymbol{y})| \le K u(\boldsymbol{y}) |\boldsymbol{y}|^{\boldsymbol{q}} \sum_{\boldsymbol{m}} P_{\boldsymbol{m}}(\boldsymbol{x}) |\boldsymbol{z}|^{\boldsymbol{m}}$$

where $P_{\boldsymbol{m}} \in \mathfrak{P}(l)$ (see (B.3) for a definition) and the sum runs over all $\boldsymbol{m} = (m_1, \dots, m_l) \neq (0, \dots, 0)$ with $m_j$ either $p_j$ or 0. We do not give an explicit formula here since the only important property is $\mathbb{E}[P_{\boldsymbol{m}}(\sqrt{n}\Delta_{\boldsymbol{i}}^n X')^q] \le K$ for all $q \ge 0$, which directly follows from the Burkholder inequality. Because of the boundedness of $X$ and the continuity of $u$ this leads to the following bound on $\xi_n$:

$$|\xi_n| \le \Big(K \sum_{\boldsymbol{j} \in \mathcal{B}_t^n(d-l)} |\Delta_{\boldsymbol{j}}^n X|^{\boldsymbol{q}}\Big)\Big(\frac{\sqrt{n}}{n^l} \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l)} \sum_{\boldsymbol{m}} P_{\boldsymbol{m}}(\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c) |\sqrt{n}\Delta_{\boldsymbol{i}}^n X^d|^{\boldsymbol{m}}\Big).$$

The first factor converges in probability to some finite limit, and hence it is enough to show that the second factor converges in $L^1$ to 0. Without loss of generality we restrict ourselves to the summand with $\boldsymbol{m} = (p_1, \dots, p_k, 0, \dots, 0)$ for some $1 \le k \le l$. From [12, Lemma 2.1.7] it follows that

$$\mathbb{E}[|\Delta_i^n X^d|^q | \mathcal{F}_{\frac{i-1}{n}}] \le \frac{K}{n} \qquad \text{for all } q > 0. \tag{B.27}$$

Let $r := \max_{1 \le i \le l} p_i$ and $b_k(\boldsymbol{i}) := \#\{i_1, \dots, i_k\}$ for $\boldsymbol{i} = (i_1, \dots, i_l)$. Note that the number of $\boldsymbol{i} \in \mathcal{B}_t^n(l)$ with $b_k(\boldsymbol{i}) = m$ is of order $n^{m+l-k}$ for $1 \le m \le k$. An application of Hölder inequality, successive use of (B.27) and the boundedness of $X$ gives

$$\begin{aligned}
&\mathbb{E}\Big(\frac{\sqrt{n}}{n^l} \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l)} P_{\boldsymbol{m}}(\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c) |\sqrt{n}\Delta_{i_1}^n X^d|^{p_1} \dots |\sqrt{n}\Delta_{i_k}^n X^d|^{p_k}\Big) \\
&\quad \le \frac{n^{1/2+kr/2}}{n^l} \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l)} \Big(\mathbb{E}[P_{\boldsymbol{m}}(\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c)^{\frac{4}{1-r}}]\Big)^{\frac{1-r}{4}} \\
&\qquad\qquad \times \Big(\mathbb{E}\Big[\Big(|\Delta_{i_1}^n X^d|^{p_1} \dots |\Delta_{i_k}^n X^d|^{p_k}\Big)^{\frac{4r}{3+r}}\Big]\Big)^{\frac{3+r}{4}} \\
&\quad \le K \frac{n^{1/2+kr/2}}{n^l} \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l)} n^{-b_k(\boldsymbol{i})(3+r)/4} \le K \frac{n^{1/2+kr/2}}{n^l} \sum_{j=1}^k n^{-j(3+r)/4} n^{j+l-k} \\
&\quad = K \sum_{j=1}^k n^{(2-2k+(2k-j)(r-1))/4}.
\end{aligned}$$

The latter expression converges to 0 since $r < 1$.

In the next step we will show that we can replace the increments $\Delta_i^n X^c$ of the continuous part of $X$ by their first order approximation $\alpha_i^n = \sqrt{n}\sigma_{\frac{i-1}{n}} \Delta_i^n W$.

**Proposition B.14.** *It holds that*

$$\xi_n' = \sqrt{n}\Big(\frac{1}{n^l}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)}\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)} H(\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c, \Delta_{\boldsymbol{j}}^n X) \\ - \frac{1}{n^l}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)}\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)} H(\alpha_{\boldsymbol{i}}^n, \Delta_{\boldsymbol{j}}^n X)\Big) \xrightarrow{\mathbb{P}} 0$$

*as* $n \to \infty$.

We shift the proof of this result to the appendix. Having simplified the statistics in the first argument, we now focus on the second one, more precisely on the process

$$\theta_n(H) = \sqrt{n}\Big(\frac{1}{n^l}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)}\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)} H(\alpha_{\boldsymbol{i}}^n, \Delta_{\boldsymbol{j}}^n X) - \frac{1}{n^l}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)}\sum_{\boldsymbol{s}\in[0,t]^{d-l}} H(\alpha_{\boldsymbol{i}}^n, \Delta X_{\boldsymbol{s}})\Big).$$

In the following we will use the notation from Section B.3. We split $\theta_n(H)$ into

$$\theta_n(H) = \mathbb{1}_{\Omega_n(m)}\theta_n(H) + \mathbb{1}_{\Omega\setminus\Omega_n(m)}\theta_n(H).$$

Since $\Omega_n(m) \xrightarrow{\mathbb{P}} \Omega$ as $n \to \infty$, the latter term converges in probability to 0 as $n \to \infty$. The following result will be shown in the appendix as well.

**Proposition B.15.** *We have the convergence*

$$\mathbb{1}_{\Omega_n(m)}\theta_n(H) - \frac{1}{n^l}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)}\sum_{\boldsymbol{q}\in\mathcal{P}_t^n(m)^{d-l}}\sum_{k=l+1}^{d} \partial_k H(\alpha_{\boldsymbol{i}}^n, \Delta X_{S_{\boldsymbol{q}}})R(n, q_k) \xrightarrow{\mathbb{P}} 0$$

*if we first let* $n \to \infty$ *and then* $m \to \infty$.

Using all the approximations, in view of Lemma B.6 we are left with

$$\Phi_t^n(m) := \frac{1}{n^l}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)}\sum_{\boldsymbol{q}\in\mathcal{P}_t^n(m)^{d-l}}\sum_{k=l+1}^{d} \partial_k H(\alpha_{\boldsymbol{i}}^n, \Delta X_{S_{\boldsymbol{q}}})R(n, q_k) + \sum_{\boldsymbol{s}\in[0,t]^{d-l}} \mathbb{U}_t^n(H, \Delta X_{\boldsymbol{s}}) \\ = \sum_{\boldsymbol{q}\in\mathbb{N}^{d-l}}\Big(\mathbb{1}_{\mathcal{P}_t^n(m)^{d-l}}(\boldsymbol{q})\sum_{k=l+1}^{d} \mathbb{B}_t^n(\partial_k H, \Delta X_{S_{\boldsymbol{q}}})R(n, q_k) + \mathbb{U}_t^n(H, \Delta X_{S_{\boldsymbol{q}}})\Big).$$

The remainder of the proof will consist of four steps. First we use for all $k \in \mathbb{N}$ the decomposition $\Phi_t^n(m) = \Phi_t^n(m,k) + \tilde{\Phi}_t^n(m,k)$, where

$$\Phi_t^n(m,k) := \sum_{q_1,\dots,q_{d-l}\le k} \mathbb{1}_{\mathcal{P}_t^n(m)^{d-l}}(\boldsymbol{q})\sum_{k=l+1}^{d} \mathbb{B}_t^n(\partial_k H, \Delta X_{S_{\boldsymbol{q}}})R(n, q_k) + \sum_{\boldsymbol{q}\in\mathbb{N}^{d-l}} \mathbb{U}_t^n(H, \Delta X_{S_{\boldsymbol{q}}}),$$

i.e. we consider only finitely many jumps in the first summand. We will successively show

$$\lim_{k\to\infty}\limsup_{n\to\infty}\mathbb{P}(|\tilde{\Phi}_t^n(m,k)| > \eta) = 0 \qquad \text{for all } \eta > 0,$$
$$\Phi_t^n(m,k) \xrightarrow{st} \Phi_t(m,k) \qquad \text{as } n \to \infty,$$

for a process $\Phi_t(m,k)$ that will be defined in (B.30). Finally, with $\Phi_t(m)$ defined in (B.33) we will show

$$\Phi_t(m,k) \xrightarrow{\mathbb{P}} \Phi_t(m) \qquad \text{as } k \to \infty, \tag{B.28}$$
$$\Phi_t(m) \xrightarrow{\mathbb{P}} V'(H,X,l)_t. \tag{B.29}$$

For (B.4) observe that we have $\mathcal{P}_t^n(m) \subset \mathcal{P}_t(m)$ and therefore

$$\mathbb{P}\Big(|\tilde{\Phi}_t^n(m,k)| > \eta\Big) \le \mathbb{P}\big(\{\omega : \mathcal{P}_t(m,\omega) \not\subset \{1,\dots,k\}\}\big) \to 0 \quad \text{as} \quad k \to \infty,$$

since the sets $\mathcal{P}_t(m,\omega)$ are finite for fixed $\omega$ and $m$. For (B.4) recall that $g$ was defined by $g(\boldsymbol{x},\boldsymbol{y}) = |x_1|^{p_1} \cdots |x_l|^{p_l} L(\boldsymbol{x},\boldsymbol{y})$. By Propositions B.11 and B.12 and from the properties of stable convergence (in particular, we need joint stable convergence with sequences converging in probability, which is useful for the indicators below) we have

$$\begin{aligned}
&(\mathbb{U}_t^n(g,\cdot), (\mathbb{B}_t^n(\partial_j H,\cdot))_{j=l+1}^d, (\Delta X_{S_p})_{p\in\mathbb{N}}, (R(n,p))_{p\in\mathbb{N}}, (\mathbb{1}_{\mathcal{P}_t^n(m)}(p))_{p\in\mathbb{N}}) \\
&\quad \xrightarrow{st} (\mathbb{U}_t(g,\cdot), (\mathbb{B}_t(\partial_j H,\cdot))_{j=l+1}^d, (\Delta X_{S_p})_{p\in\mathbb{N}}, (R_p)_{p\in\mathbb{N}}, (\mathbb{1}_{\mathcal{P}_t(m)}(p))_{p\in\mathbb{N}})
\end{aligned}$$

as $n \to \infty$ in the space $\mathcal{C}[-A,A]^{(d-l)} \times (\mathcal{C}[-A,A]^{(d-l)})^{d-l} \times \ell_A^2 \times \mathbb{R}^{\mathbb{N}} \times \mathbb{R}^{\mathbb{N}}$, where we denote by $\ell_A^2$ the metric space

$$\ell_A^2 := \{(x_k)_{k\in\mathbb{N}} \in \ell^2 \mid |x_k| \le A \text{ for all } k \in \mathbb{N}\}.$$

For $k \in \mathbb{N}$ we now define a continuous mapping

$$\phi_k : \mathcal{C}[-A,A]^{(d-l)} \times (\mathcal{C}[-A,A]^{(d-l)})^{d-l} \times \ell_A^2 \times \mathbb{R}^{\mathbb{N}} \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$$

by

$$\begin{aligned}
&\phi_k(f, (g_r)_{r=1}^{d-l}, (x_j)_{j\in\mathbb{N}}, (y_j)_{j\in\mathbb{N}}, (z_j)_{j\in\mathbb{N}}) \\
&\quad = \sum_{j_1,\dots,j_{d-l}=1}^{k} z_{j_1} \cdots z_{j_{d-l}} \sum_{r=l+1}^{d} g_r(x_{j_1},\dots,x_{j_{d-l}}) y_{j_r} \\
&\qquad + \sum_{j_1,\dots,j_{d-l}=1}^{\infty} |x_{j_1}|^{q_1} \cdots |x_{j_{d-l}}|^{q_{d-l}} f(x_{j_1},\dots,x_{j_{d-l}}).
\end{aligned}$$

The continuous mapping theorem then yields

$$\begin{aligned}
&\Phi_t^n(m,k) && \text{(B.30)}\\
&\quad = \phi_k(\mathbb{U}_t^n(g,\cdot), (\mathbb{B}_t^n(\partial_r H,\cdot))_{r=l+1}^d, (\Delta X_{S_p})_{p\in\mathbb{N}}, (R(n,p))_{p\in\mathbb{N}}, (\mathbb{1}_{\mathcal{P}_t^n(m)}(p))_{p\in\mathbb{N}}) \\
&\quad \xrightarrow{st} \phi_k(\mathbb{U}_t(g,\cdot), (\mathbb{B}_t(\partial_r H,\cdot))_{r=l+1}^d, (\Delta X_{S_p})_{p\in\mathbb{N}}, (R_p)_{p\in\mathbb{N}}, (\mathbb{1}_{\mathcal{P}_t(m)}(p))_{p\in\mathbb{N}}) \\
&\quad = \sum_{q_1,\dots,q_{d-l}\le k} \mathbb{1}_{\mathcal{P}_t(m)^{d-l}}(\boldsymbol{q}) \sum_{r=l+1}^{d} \mathbb{B}_t(\partial_r H, \Delta X_{S_{\boldsymbol{q}}}) R(n,q_r) && \text{(B.31)}\\
&\qquad + \sum_{\boldsymbol{q}\in\mathbb{N}^{d-l}} \mathbb{U}_t(H, \Delta X_{S_{\boldsymbol{q}}}) \\
&\quad =: \Phi_t(m,k). && \text{(B.32)}
\end{aligned}$$

For $k \to \infty$ we have $\Phi_t(m,k) \xrightarrow{\text{a.s.}} \Phi_t(m)$ with

$$\Phi_t(m) := \sum_{\boldsymbol{q} \in \mathbb{N}^{d-l}} \Big( \mathbb{1}_{\mathcal{P}_t(m)^{d-l}}(\boldsymbol{q}) \sum_{r=l+1}^{d} \mathbb{B}_t(\partial_r H, \Delta X_{S_{\boldsymbol{q}}}) R(n, q_r) + \sum_{\boldsymbol{q} \in \mathbb{N}^{d-l}} \mathbb{U}_t(H, \Delta X_{S_{\boldsymbol{q}}}) \Big), \tag{B.33}$$

i.e. (B.28). For the last assertion (B.29) we have

$$\begin{aligned} &\mathbb{P}(|\Phi_t(m) - V'_t(H,X,l)| > \eta) \\ &\quad \le K\mathbb{E}[(\Phi_t(m) - V'_t(H,X,l))^2] = K\mathbb{E}[\mathbb{E}[(\Phi_t(m) - V'_t(H,X,l))^2 | \mathcal{F}]] \\ &\quad \le K\mathbb{E}[\sum_{\boldsymbol{k} \in \mathbb{N}^{d-l}} \sum_{r=l+1}^{d} (1 - \mathbb{1}_{\mathcal{P}_t(m)^{d-l}}(\boldsymbol{k})) |\mathbb{B}_t(\partial_r H, \Delta X_{S_{\boldsymbol{k}}})|^2] \\ &\quad \le K\mathbb{E}[\sum_{\boldsymbol{k} \in \mathbb{N}^{d-l}} \sum_{r=l+1}^{d} (1 - \mathbb{1}_{\mathcal{P}_t(m)^{d-l}}(\boldsymbol{k})) \prod_{i=1}^{d-l} \big( |\Delta X_{S_{k_i}}|^{q_i} + |\Delta X_{S_{k_i}}|^{q_i - 1} \big)^2]. \end{aligned}$$

Since the jumps are absolutely summable and bounded the latter expression converges to 0 as $m \to \infty$. □

## B.5 Appendix

### Existence of the limiting processes

We give a proof that the limiting processes in Theorem B.5 and Theorem B.13 are well-defined. The proof will be similar to the proof of [12, Prop. 4.1.4]. We restrict ourselves to proving that

$$\sum_{\boldsymbol{k}: T_{\boldsymbol{k}} \le t} \int_{[0,t]^l} \rho_{\partial_{l+1} H}(\sigma_{\boldsymbol{u}}, \Delta X_{T_{\boldsymbol{k}}}) d\boldsymbol{u} R_{k_1} \tag{B.34}$$

is defined in a proper way, corresponding to Theorem B.13. For $l = 0$ we basically get the result for Theorem B.5, but under slightly stronger assumptions. The proof, however, remains the same.

We show that the sum in (B.34) converges in probability for all $t$ and that the conditional properties mentioned in the theorems are fulfilled. Let $I_m(t) = \{n \mid 1 \le n \le m, T_n \le t\}$. Define

$$Z(m)_t := \sum_{\boldsymbol{k} \in I_m(t)^{d-l}} \int_{[0,t]^l} \rho_{\partial_{l+1} H}(\sigma_{\boldsymbol{u}}, \Delta X_{T_{\boldsymbol{k}}}) d\boldsymbol{u} R_{k_1}.$$

By fixing $\omega \in \Omega$, we further define the process $Z^\omega(m)_t$ on $(\Omega', \mathcal{F}', \mathbb{P}')$ by putting $Z^\omega(m)_t(\omega') = Z(m)_t(\omega, \omega')$. The process is obviously centered, and we can immediately deduce

$$\mathbb{E}'(Z^\omega(m)_t^2) = \sum_{k_1 \in I_m(t)} \Big( \sum_{\boldsymbol{k} \in I_m(t)^{d-l-1}} \int_{[0,t]^l} \rho_{\partial_{l+1} H}(\sigma_{\boldsymbol{u}}, \Delta X_{T_{k_1}}, \Delta X_{T_{\boldsymbol{k}}}) d\boldsymbol{u} \Big)^2 \sigma_{T_{k_1}}^2, \tag{B.35}$$

$$\mathbb{E}'(e^{iuZ^\omega(m)_t}) = \prod_{k_1 \in I_m(t)} \int e^{iu \sum_{\boldsymbol{k} \in I_m(t)^{d-l-1}} \int_{[0,t]^l} \rho \partial_{l+1} H(\sigma_{\boldsymbol{u}}, \Delta X_{T_{k_1}}, \Delta X_{T_{\boldsymbol{k}}}) d\boldsymbol{u} R_{k_1}} d\mathbb{P}'. \tag{B.36}$$

The processes $X$ and $\sigma$ are both càdlàg and hence bounded on $[0,T]$ for a fixed $\omega \in \Omega$. Let now $m, m' \in \mathbb{N}$ with $m' \leq m$ and observe that $I_m(t)^q \backslash I_{m'}(t)^q \subset I_m(T)^q \backslash I_{m'}(T)^q$ for all $q \in \mathbb{N}$ and $t \leq T$. Since $L$ and $\partial_1 L$ are bounded on compact sets, we obtain

$$\begin{aligned}
&\mathbb{E}'\Big[\Big(\sup_{t\in[0,T]} |Z^\omega(m)_t - Z^\omega(m')_t|\Big)^2\Big] \\
&\quad = \mathbb{E}'\Big[\Big(\sup_{t\in[0,T]} \Big| \sum_{\boldsymbol{k}\in I_m(t)^{d-l}\backslash I_{m'}(t)^{d-l}} \int_{[0,t]^l} \rho \partial_{l+1} H(\sigma_{\boldsymbol{u}}, \Delta X_{T_{\boldsymbol{k}}}) d\boldsymbol{u} R_{k_1}\Big|\Big)^2\Big] \\
&\quad \leq \mathbb{E}'\Big[\Big(\sum_{\boldsymbol{k}\in I_m(T)^{d-l}\backslash I_{m'}(T)^{d-l}} \int_{[0,T]^l} \big|\rho \partial_{l+1} H(\sigma_{\boldsymbol{u}}, \Delta X_{T_{\boldsymbol{k}}})\big| d\boldsymbol{u} |R_{k_1}|\Big)^2\Big] \\
&\quad \leq K(\omega)\Big( \sum_{\boldsymbol{k}\in I_m(T)^{d-l}\backslash I_{m'}(T)^{d-l}} \\
&\qquad\qquad \big(|\Delta X_{T_{k_1}}|^{q_1-1} + |\Delta X_{T_{k_1}}|^{q_1}\big)|\Delta X_{T_{k_2}}|^{q_2} \cdots |\Delta X_{T_{k_{d-l}}}|^{q_{d-l}}\Big)^2 \\
&\quad \to 0 \qquad \text{as } m, m' \to \infty
\end{aligned}$$

for $\mathbb{P}$-almost all $\omega$, since $\sum_{s\leq t} |\Delta X_s|^p$ is almost surely finite for any $p \geq 2$. Therefore we obtain, as $m, m' \to \infty$,

$$\begin{aligned}
&\tilde{\mathbb{P}}\Big(\sup_{t\in[0,t]} |Z(m)_t - Z(m')_t| > \epsilon\Big) \\
&\qquad = \int \mathbb{P}'\Big(\sup_{t\in[0,T]} |Z^\omega(m)_t - Z^\omega(m')_t| > \epsilon\Big) d\mathbb{P}(\omega) \to 0
\end{aligned}$$

by the dominated convergence theorem. The processes $Z(m)$ are càdlàg and contitute a Cauchy sequence in probability in the supremum norm. Hence they converge in probability to some $\tilde{\mathcal{F}}_t$-adapted càdlàg process $Z_t$. By the previous estimates we also obtain directly that

$$Z^\omega(m)_t \to Z_t(\omega, \cdot) \quad \text{in} \quad L^2(\Omega', \mathcal{F}', \mathbb{P}'). \tag{B.37}$$

As a consequence it follows from (B.35) that

$$\begin{aligned}
&\int Z_t(\omega, \omega')^2 d\mathbb{P}'(\omega') \\
&\quad = \sum_{s_1 \leq t} \Big( \sum_{s_2, \ldots, s_{d-l} \leq t} \int_{[0,t]^l} \rho \partial_{l+1} H(\sigma_{\boldsymbol{u}}, \Delta X_{s_1}, \Delta X_{s_2}, \ldots, \Delta X_{s_{d-l}}) d\boldsymbol{u}\Big)^2 \sigma_{s_1}^2.
\end{aligned}$$

Note that the multiple sum on the right hand side of the equation converges absolutely and hence does not depend on the choice of $(T_n)$. By (B.37) we obtain

$$\mathbb{E}'(e^{iuZ^\omega(m)_t}) \to \mathbb{E}'(e^{iuZ_t(\omega,\cdot)}).$$

Observe that for any centered square integrable random variable $U$ we have

$$\Big|\int e^{iyU} - 1 d\mathbb{P}\Big| \leq \mathbb{E}U^2|y|^2 \qquad \text{for all } y \in \mathbb{R}.$$

Therefore the product in (B.36) converges absolutely as $m \to \infty$, and hence the characteristic function and thus the law of $Z_t(\omega,\cdot)$ do not depend on the choice of the sequence $(T_n)$. Lastly, observe that $R_n$ is $\mathcal{F}$-conditionally Gaussian. (In the case of a possibly discontinuous $\sigma$ as in Theorem B.5 we need to require that $X$ and $\sigma$ do not jump at the same time to obtain such a property.) So we can conclude that $Z^\omega(m)_t$ is Gaussian, and $Z_t(\omega,\cdot)$ as a stochastic limit of Gaussian random variables is so as well.

### Uniform limit theory for continuous U-statistics

In this chapter we will give a proof of Proposition B.12. Mainly we have to show that the sequence in (B.24) is tight and that the finite dimensional distributions converge to the finite dimensional distributions of $\mathbb{U}_t$. For the convergence of the finite dimensional distributions we will generalize Proposition 4.3 in [18]. The basic idea in that work is to write the U-statistic as an integral with respect to the empirical distribution function

$$F_n(t,x) = \frac{1}{n}\sum_{j=1}^{\lfloor nt\rfloor} \mathbb{1}_{\{\alpha_j^n \leq x\}}.$$

In our setting we have

$$\frac{1}{n^l}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l)} G(\alpha_{\boldsymbol{i}}^n,\boldsymbol{y}) = \int_{\mathbb{R}^l} G(\boldsymbol{x},\boldsymbol{y})F_n(t,dx_1)\cdots F_n(t,dx_l).$$

Of particular importance in [18] is the limit theory for the empirical process connected with $F_n$, which is given by

$$\mathbb{G}_n(t,x) = \frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nt\rfloor}\Big(\mathbb{1}_{\{\alpha_j^n\leq x\}} - \Phi_{\sigma_{\frac{j-1}{n}}}(x)\Big),$$

where $\Phi_z$ is the cumulative distribution function of a standard normal random variable with variance $z^2$. As a slight generalization of [18, Prop. 4.2] and by the same arguments as in [12, Prop. 4.4.10] we obtain the joint convergence

$$(\mathbb{G}_n(t,x),(R_-(n,p),R_+(n,p))_{p\geq1}) \xrightarrow{st} (\mathbb{G}(t,x),(R_{p-},R_{p+})_{p\geq1}).$$

The stable convergence in law is to be understood as a process in $t$ and in the finite distribution sense in $x\in\mathbb{R}$. The limit is defined on an extension $(\Omega',\mathcal{F}',\mathcal{P}')$ of the original probability space. $\mathbb{G}$ is $\mathcal{F}$-conditionally independent of $(\kappa_n,\psi_{n\pm})_{n\geq1}$ and $\mathcal{F}$-conditionally Gaussian and satisfies

$$\mathbb{E}'[\mathbb{G}(t,x)|\mathcal{F}] = \int_0^t \overline{\Phi}_{\sigma_s}(x)dW_s,$$

$$\mathbb{E}'[\mathbb{G}(t_1,x_1)\mathbb{G}(t_2,x_2)|\mathcal{F}] - \mathbb{E}'[\mathbb{G}(t_1,x_1)|\mathcal{F}]\mathbb{E}'[\mathbb{G}(t_2,x_2)|\mathcal{F}]$$
$$= \int_0^{t_1\wedge t_2} \Phi_{\sigma_s}(x_1\wedge x_2) - \Phi_{\sigma_s}(x_1)\Phi_{\sigma_s}(x_2) - \overline{\Phi}_{\sigma_s}(x_1)\overline{\Phi}_{\sigma_s}(x_2)ds,$$

where $\overline{\Phi}_z(x) = \mathbb{E}[V \mathbb{1}_{\{zV \le x\}}]$ with $V \sim \mathcal{N}(0,1)$. As in the proof of Prop. 4.3 in [18] we will use the decomposition

$$\begin{aligned}\mathbb{U}_t^n(G,\boldsymbol{y}) &= \sum_{k=1}^l \int_{\mathbb{R}^l} G(\boldsymbol{x},\boldsymbol{y})\mathbb{G}_n(t,dx_k) \prod_{m=1}^{k-1} F_n(t,dx_m) \prod_{m=k+1}^{l} \bar{F}_n(t,dx_m) \\ &\qquad + \sqrt{n}\Big(\frac{1}{n^l}\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(l)} \rho_G(\sigma_{(\boldsymbol{j}-1)/n},\boldsymbol{y}) - \int_{[0,t]^l} \rho_G(\sigma_{\boldsymbol{s}},\boldsymbol{y})d\boldsymbol{s}\Big) \\ &=: \sum_{k=1}^l Z_k^n(G,\boldsymbol{y}) + R^n(\boldsymbol{y}),\end{aligned}$$

where

$$\bar{F}_n(t,x) = \frac{1}{n}\sum_{j=1}^{\lfloor nt \rfloor} \Phi_{\sigma_{(j-1)/n}}(x).$$

From [18, Prop. 3.2] we know that both $F_n$ and $\bar{F}_n$ converge in probability to $F(t,x) = \int_0^t \Phi_{\sigma_s}(x)ds$ for fixed $t$ and $x$. If $G$ is symmetric and continuously differentiable in $\boldsymbol{x}$ with derivative of polynomial growth, [18, Prop. 4.3] gives for fixed $\boldsymbol{y}$

$$\sum_{k=1}^l Z_k^n(G,\boldsymbol{y}) \xrightarrow{st} \sum_{k=1}^l \int_{\mathbb{R}^l} G(\boldsymbol{x},\boldsymbol{y})\mathbb{G}(t,dx_k) \prod_{m\neq k} F(t,dx_m) =: \sum_{k=1}^l Z_k(G,\boldsymbol{y}). \tag{B.38}$$

We remark that the proof of this result mainly relies on the following steps: First, use the convergence of $F_n$ and $\bar{F}_n$ and replace both by their limit $F$, which is differentiable in $x$. Then use the integration by parts formula for the Riemann-Stieltjes integral with respect to $\mathbb{G}_n(t,dx_k)$ plus the differentiability of $G$ in the $k$-th argument to obtain that $Z_k^n(G,\boldsymbol{y})$ behaves asymptotically exactly the same as $-\int_{\mathbb{R}^l} \partial_k G(\boldsymbol{x},\boldsymbol{y})\mathbb{G}_n(t,x_k)\prod_{m\neq k} F'(t,x_m)d\boldsymbol{x}$. Since one now only has convergence in finite dimensional distribution of $\mathbb{G}_n(t,\cdot)$ to $\mathbb{G}(t,\cdot)$, one uses a Riemann approximation of the integral with respect to $dx_k$ and takes limits afterwards. In the end do all the steps backwards.

From the proof and the aforementioned joint convergence of $\mathbb{G}_n$ and the sequence $(R_\pm(n,p))_{p\ge1}$ it is clear that we can slightly generalize (B.38) to

$$\begin{aligned}&\Big((Z_k^n(G,\boldsymbol{y}))_{1\le k\le l}, (R_-(n,p), R_+(n,p))_{p\ge1}\Big) \\ &\qquad \xrightarrow{st} \Big((Z_k(G,\boldsymbol{y}))_{1\le k\le l}, (R_{p-}, R_{p+})_{p\ge1}\Big),\end{aligned} \tag{B.39}$$

where the latter convergence holds in the finite distribution sense in $\boldsymbol{y}$ and also for non-symmetric, but still continuously differentiable functions $G$. A second consequence of the proof of (B.38) is that the mere convergence $Z_k^n(G,\boldsymbol{y}) \xrightarrow{st} Z_k(G,\boldsymbol{y})$ only requires $G$ to be continuously differentiable in the $k$-th argument if $k$ is fixed.

To show that (B.39) holds in general under our assumptions let $\psi_\epsilon \in \mathcal{C}^\infty(\mathbb{R})$ $(\epsilon > 0)$ be functions with $0 \le \psi_\epsilon \le 1$, $\psi_\epsilon(x) \equiv 1$ on $[-\epsilon/2,\epsilon/2]$, $\psi_\epsilon(x) \equiv 0$ outside of $(-\epsilon,\epsilon)$, and $\|\psi_\epsilon'\| \le K\epsilon^{-1}$ for some constant $K$, which is independent of $\epsilon$. Then the function $G(\boldsymbol{x})(1-\psi_\epsilon(x_k))$ is continuously differentiable in the $k$-th argument and hence it is enough to prove

$$\lim_{\epsilon\to 0}\limsup_{n\to\infty}\mathbb{P}(\sup_{\boldsymbol{y}\in[-A,A]^{d-l}}|Z_k^n(G\psi_\epsilon,\boldsymbol{y})|>\eta)=0 \tag{B.40}$$

$$\lim_{\epsilon\to 0}\mathbb{P}(\sup_{\boldsymbol{y}\in[-A,A]^{d-l}}|Z_k(G\psi_\epsilon,\boldsymbol{y})|>\eta)=0 \tag{B.41}$$

for all $\eta>0$ and $1\le k\le l$. For given $k$ the functions $\psi_\epsilon$ are to be evaluated at $x_k$. We show (B.40) only for $k=l$. The other cases are easier since $\bar{F_n}$ is continuously differentiable in $x$ and the derivative is bounded by a continuous function with exponential decay at $\pm\infty$ since $\sigma$ is bounded away from 0.

For $k=l$, some $P\in\mathfrak{P}(1), Q\in\mathfrak{P}(l-1)$ and $x_l\neq 0$, we have

$$|\partial_l(G(\boldsymbol{x},\boldsymbol{y})\psi_\epsilon(x_l))|\le K(1+|x_l|^{p_1-1})P(x_l)Q(x_1,\dots,x_{l-1})+K|x_1|^{p_1}\cdots|x_l|^{p_l}\epsilon^{-1}.$$

Since $p_1-1>-1$ the latter expression is integrable with respect to $x_l$ on compact intervals. Therefore the standard rules for the Riemann-Stieltjes integral and the monotonicity of $F_n$ in $x$ yield

$$\begin{aligned}
&\sup_{\boldsymbol{y}\in[-A,A]^{d-l}}|Z_l^n(G\psi_\epsilon,\boldsymbol{y})|\\
&=\sup_{\boldsymbol{y}\in[-A,A]^{d-l}}\Big|\int_{\mathbb{R}^l}G(\boldsymbol{x},\boldsymbol{y})\psi_\epsilon(x_l)\mathbb{G}_n(t,dx_l)\prod_{m=1}^{l-1}F_n(t,dx_m)\Big|\\
&=\sup_{\boldsymbol{y}\in[-A,A]^{d-l}}\Big|\int_{\mathbb{R}^l}-\mathbb{G}_n(t,x_l)\partial_l(G(\boldsymbol{x},\boldsymbol{y})\psi_\epsilon(x_l))dx_l\prod_{m=1}^{l-1}F_n(t,dx_m)\Big|\\
&\le\int_{\mathbb{R}^{l-1}}\int_{-\epsilon}^{\epsilon}K|\mathbb{G}_n(t,x_l)|(1+|x_l|^{p_1-1})P(x_l)Q(x_1,\dots,x_{l-1})dx_l\prod_{m=1}^{l-1}F_n(t,dx_m)\\
&\quad+\int_{\mathbb{R}^{l-1}}\int_{-\epsilon}^{\epsilon}K|\mathbb{G}_n(t,x_l)||x_1|^{p_1}\cdots|x_l|^{p_l}\epsilon^{-1}dx_l\prod_{m=1}^{l-1}F_n(t,dx_m)\\
&=\int_{-\epsilon}^{\epsilon}K\Big(\frac{1}{n^{l-1}}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)}Q(\alpha_{\boldsymbol{i}}^n)\Big)|\mathbb{G}_n(t,x_l)|(1+|x_l|^{p_1-1})P(x_l)dx_l\\
&\quad+\int_{-\epsilon}^{\epsilon}K\Big(\frac{1}{n^{l-1}}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)}|\alpha_{i_1}^n|^{p_1}\cdots|\alpha_{i_{l-1}}^n|^{p_{l-1}}\Big)|\mathbb{G}_n(t,x_l)||x_l|^{p_l}\epsilon^{-1}dx_l.
\end{aligned}$$

We have $\mathbb{E}|\alpha_i^n|^q\le K$ uniformly in $i$ for every $q\ge 0$. From [18, Lemma 4.1] it further follows that $\mathbb{E}|\mathbb{G}_n(t,x)|^q\le K$ for all $q\ge 2$. Then we deduce from Hölder inequality

$$\mathbb{E}\Big(\sup_{\boldsymbol{y}\in[-A,A]^{d-l}}|Z_l^n(G\psi_\epsilon,\boldsymbol{y})|\Big)\le K\int_{-\epsilon}^{\epsilon}(1+|x_l|^{p_1-1})P(x_l)+|x_l|^{p_l}\epsilon^{-1}dx_l,$$

which converges to 0 if we let $\epsilon\to 0$. We omit the proof of (B.41) since it follows by the same arguments.

So far we have proven that (B.39) holds under our assumptions on $G$. Furthermore, we can easily calculate the conditional covariance structure of the conditionally centered Gaussian process $\sum_{k=1}^l Z_k(G,\boldsymbol{y})$ by simply using that we know the covariance structure of $\mathbb{G}(t,x)$. We obtain the form in (B.25); for more details see [18, sect. 5].

Next we will show that

$$\sup_{\boldsymbol{y}\in[-A,A]^{d-l}} |R^n(\boldsymbol{y})| \xrightarrow{\mathbb{P}} 0 \tag{B.42}$$

as $n \to \infty$. Observe that $\rho_G(\boldsymbol{x},\boldsymbol{y})$ is $\mathcal{C}^{d+1}$ in the $\boldsymbol{x}$ argument. Therefore we get $R^n(\boldsymbol{y}) \xrightarrow{\mathbb{P}} 0$ for any fixed $\boldsymbol{y}$ from [18, sect. 7.3]. Further we can write

$$\begin{aligned} R^n(\boldsymbol{y}) =& \sqrt{n}\int_{[0,\lfloor nt\rfloor/n]^l} (\rho_G(\sigma_{\lfloor n\boldsymbol{s}\rfloor/n},\boldsymbol{y}) - \rho_G(\sigma_{\boldsymbol{s}},\boldsymbol{y}))d\boldsymbol{s} \\ &+ \sqrt{n}\Big(\int_{[0,t]^l} \rho_G(\sigma_{\boldsymbol{s}},\boldsymbol{y})d\boldsymbol{s} - \int_{[0,\lfloor nt\rfloor/n]^l} \rho_G(\sigma_{\boldsymbol{s}},\boldsymbol{y})d\boldsymbol{s}\Big). \end{aligned}$$

The latter term converges almost surely to 0 and hence we can deduce (B.42) from the fact that $\mathbb{E}|R^n(\boldsymbol{y}) - R^n(\boldsymbol{y'})| \le K\,\|\boldsymbol{y}-\boldsymbol{y'}\|$, which follows because $\rho_G(\boldsymbol{x},\boldsymbol{y})$ is continuously differentiable in $\boldsymbol{y}$ and $\mathbb{E}(\sqrt{n}|\sigma_{\lfloor nu\rfloor/n} - \sigma_u|) \le K$ for all $u \in [0,t]$.

Therefore we have proven the convergence of the finite dimensional distributions

$$((\mathbb{U}_t^n(G,\boldsymbol{y}_i))_{i=1}^m, (R_-(n,p), R_+(n,p))_{p\ge1})) \xrightarrow{st} ((\mathbb{U}_t(G,\boldsymbol{y}_i))_{i=1}^m, (R_{p-}, R_{p+})_{p\ge1}).$$

What remains to be shown in order to deduce Proposition B.12 is that the limiting process is indeed continuous and that the sequences $Z_k^n(G,\cdot)$ $(1 \le k \le l)$ are tight. For the continuity of the limit observe that

$$\begin{aligned} &\mathbb{E}[|\mathbb{U}_t(G,\boldsymbol{y}) - \mathbb{U}_t(G,\boldsymbol{y}')|^2|\mathcal{F}] \\ &= \int_0^t \int_{\mathbb{R}} \Big(\sum_{i=1}^l (f_i(u,\boldsymbol{y}) - f_i(u,\boldsymbol{y'}))\Big)^2 \phi_{\sigma_s}(u)du \\ &\qquad - \Big(\sum_{i=1}^l \int_{\mathbb{R}} (f_i(u,\boldsymbol{y}) - f_i(u,\boldsymbol{y'}))\phi_{\sigma_s}(u)du\Big)^2 ds. \end{aligned}$$

Here we can use the differentiability assumptions and the boundedness of $\sigma$ and $\sigma^{-1}$ to obtain

$$\mathbb{E}[|\mathbb{U}_t(G,\boldsymbol{y}) - \mathbb{U}_t(G,\boldsymbol{y}')|^2] = \mathbb{E}[\mathbb{E}[|\mathbb{U}_t(G,\boldsymbol{y}) - \mathbb{U}_t(G,\boldsymbol{y}')|^2|\mathcal{F}]] \le K\,\|\boldsymbol{y}-\boldsymbol{y}'\|^2.$$

Since $\mathbb{U}_t(G,\cdot)$ is $\mathcal{F}$-conditionally Gaussian we immediately get

$$\mathbb{E}[|\mathbb{U}_t(G,\boldsymbol{y}) - \mathbb{U}_t(G,\boldsymbol{y}')|^p] \le K_p\,\|\boldsymbol{y}-\boldsymbol{y}'\|^p$$

for any even $p \ge 2$. In particular, this implies that there exists a continuous version of the multiparameter process $\mathbb{U}_t(G,\cdot)$ (see [14, Theorem 2.5.1]).

The last thing we need to show is tightness. A tightness criterion for multiparameter processes can be found in [3]. Basically we have to control the size of the increments of the process on blocks (and on lower boundaries of blocks, which works in the same way). By a block we mean a set $B \subset [-A,A]^{d-l}$ of the form

$B = (y_1, y_1'] \times \cdots \times (y_{d-l}, y_{d-l}']$, where $y_i < y_i'$. An increment of a process $Z$ defined on $[-A, A]^{d-l}$ on such a block is defined by

$$\Delta_B(Z) := \sum_{i_1,\ldots,i_{d-l}=0}^{1} (-1)^{d-l-\sum_j i_j} Z(y_1 + i_1(y_1' - y_1), \ldots, y_{d-l} + i_{d-l}(y_{d-l}' - y_{d-l})).$$

We remark that if $Z$ is sufficiently differentiable, then

$$\Delta_B(Z) = \partial_1 \cdots \partial_{d-l} Z(\boldsymbol{\xi})(y_1' - y_1) \cdot \ldots \cdot (y_{d-l}' - y_{d-l})$$

for some $\boldsymbol{\xi} \in B$. We will now show tightness for the process $Z_l^n(G, \boldsymbol{y})$. According to [3] it is enough to show

$$\mathbb{E}[|\Delta_B(Z_l^n(G, \cdot))|^2] \le K(y_1' - y_1)^2 \cdot \ldots \cdot (y_{d-l}' - y_{d-l})^2$$

in order to obtain tightness. As before we use the standard properties of the Riemann-Stieltjes integral to deduce

$$\begin{aligned}
\mathbb{E}[|\Delta_B(Z_l^n(G, \cdot))|^2] &= \mathbb{E}\Big[\Big(\int_{\mathbb{R}^l} \Delta_B(G(\boldsymbol{x}, \cdot))\mathbb{G}_n(t, dx_l) \prod_{k=1}^{l-1} F_n(t, dx_k)\Big)^2\Big] \\
&= \mathbb{E}\Big[\Big(\int_{\mathbb{R}^l} \Delta_B(\partial_l G(\boldsymbol{x}, \cdot))\mathbb{G}_n(t, x_l)dx_l \prod_{k=1}^{l-1} F_n(t, dx_k)\Big)^2\Big] \\
&= \mathbb{E}\Big[\Big(\int_{\mathbb{R}^l} \partial_l \partial_{l+1} \cdots \partial_d G(\boldsymbol{x}, \boldsymbol{\xi})\mathbb{G}_n(t, x_l)dx_l \prod_{k=1}^{l-1} F_n(t, dx_k)\Big)^2\Big] \prod_{i=1}^{l} (y_i - y_i')^2
\end{aligned}$$

for some $\boldsymbol{\xi} \in B$. As it is shown in [18] there exists a continuous function $\gamma : \mathbb{R} \to \mathbb{R}$ with exponential decay at $\pm\infty$ such that $\mathbb{E}[\mathbb{G}_n(t, x)^4] \le \gamma(x)$. Using the growth assumptions on $L$ we further know that there exist $P \in \mathfrak{P}(1)$ and $Q \in \mathfrak{P}(l-1)$ such that

$$|\partial_l \partial_{l+1} \cdots \partial_d G(\boldsymbol{x}, \boldsymbol{\xi})| \le K(1 + |x_l|^{p_l - 1})P(x_l)Q(x_1, \ldots, x_{l-1})$$

and hence

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_{\mathbb{R}^l} \partial_l \partial_{l+1} \cdots \partial_d G(\boldsymbol{x}, \boldsymbol{\xi})\mathbb{G}_n(t, x_l)dx_l \prod_{k=1}^{l-1} F_n(t, dx_k)\Big)^2\Big] \\
&\quad \le K\mathbb{E}\Big[\int_{\mathbb{R}^2} \Big(\frac{1}{n^{l-1}} \sum_{\boldsymbol{i} \in \mathcal{B}_t^n(l-1)} Q(\alpha_{\boldsymbol{i}}^n)\Big)^2 \Big(\prod_{j=1}^{2} (1 + |u_j|^{p_l - 1})P(u_j)|\mathbb{G}_n(t, u_j)|\Big) du_1 du_2\Big] \\
&\quad \le K
\end{aligned}$$

by Fubini, the Cauchy-Schwarz inequality, and the aforementioned properties of $\mathbb{G}_n(t, x)$. The proof for the tightness of $Z_k^n(G, \boldsymbol{y})$ $(1 \le k \le l-1)$ is similar and therefore omitted.

### Proofs of some technical results

*Proof of Proposition B.7.* We will prove this in two steps.

**(i)** For $j > 0$ consider the terms $\zeta^n_{k,j}(m)$ and $\tilde{\zeta}^n_{k,j}(m)$, which appear in decomposition (B.12). Since $X$ is bounded and $\mathcal{P}^n_t(m)$ a finite set, we have the estimate

$$\max(|\zeta^n_{k,j}(m)|, |\tilde{\zeta}^n_{k,j}(m)|) \le K(m)\sqrt{n}n^{-j} \sum_{\boldsymbol{i}\in\mathcal{B}^n_t(l-k)} |\Delta^n_{i_1}X(m)|^p \cdots |\Delta^n_{i_{l-k}}X(m)|^p.$$

By (B.7) we therefore obtain

$$\mathbb{E}(\mathbb{1}_{\Omega_n(m)}(|\zeta^n_{k,j}(m)| + |\tilde{\zeta}^n_{k,j}(m)|)) \to 0 \quad \text{as} \quad n \to \infty.$$

In the case $k > 0$ we have

$$|\tilde{\zeta}^n_{k,j}(m)| \le K(m)\sqrt{n}n^{-j-k\frac{p}{2}} \sum_{\boldsymbol{p}\in\mathcal{P}^n_t(m)^k} |R(n,\boldsymbol{p})|^p \sum_{\boldsymbol{i}\in\mathcal{B}^n_t(l-k)} |\Delta^n_{\boldsymbol{i}}X(m)|^p.$$

Since $(R(n,p))$ is bounded in probability as a sequence in $n$, we can deduce

$$\mathbb{1}_{\Omega_n(m)}|\tilde{\zeta}^n_{k,j}(m)| \xrightarrow{\mathbb{P}} 0 \quad \text{as} \quad n \to \infty.$$

Furthermore, in the case $j = k = 0$, we have $\zeta^n_{0,0}(m) = \tilde{\zeta}^n_{0,0}(m)$.

**(ii)** At last we have to show the convergence

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(\mathbb{1}_{\Omega_n(m)}|\sum_{k=1}^{l-1} \big(\zeta^n_{k,0}(m) - \zeta^n_k(m)\big)| > \eta) = 0 \quad \text{for all} \quad \eta > 0.$$

First we will show in a number of steps that we can replace $\Delta X_{S_{\boldsymbol{p}}} + \frac{1}{\sqrt{n}}R(n,\boldsymbol{p})$ by $\Delta X_{S_{\boldsymbol{p}}}$ in $\zeta^n_{k,0}(m)$ without changing the asymptotic behaviour. Fix $k \in \{1,\ldots,l-1\}$. We start with

$$\begin{aligned}
&\Bigg|\binom{l}{k}^{-1}\zeta^n_{k,0}(m) - \frac{\sqrt{n}}{n^{d-l}} \sum_{\substack{\boldsymbol{p}\in\mathcal{P}^n_t(m)^{k-1}\\ p_k\in\mathcal{P}^n_t(m)}} \sideset{}{'}\sum_{\boldsymbol{i}\in\mathcal{B}^n_t(d-k)} \\
&\qquad\qquad H\Big(\Delta X_{S_{\boldsymbol{p}}} + \frac{1}{\sqrt{n}}R(n,\boldsymbol{p}), \Delta X_{S_{p_k}}, \Delta^n_{\boldsymbol{i}}X(m)\Big)\Bigg| \\
&= \Bigg|\frac{\sqrt{n}}{n^{d-l}} \sum_{\substack{\boldsymbol{p}\in\mathcal{P}^n_t(m)^{k-1}\\ p_k\in\mathcal{P}^n_t(m)}} \sideset{}{'}\sum_{\boldsymbol{i}\in\mathcal{B}^n_t(d-k)} \int_0^{\frac{R(n,p_k)}{\sqrt{n}}} \\
&\qquad \partial_k H\Big(\Delta X_{S_{\boldsymbol{p}}} + \frac{1}{\sqrt{n}}R(n,\boldsymbol{p}), \Delta X_{S_{p_k}} + u, \Delta^n_{\boldsymbol{i}}X(m)\Big)du\Bigg| \\
&\le |K \sum_{\substack{\boldsymbol{p}\in\mathcal{P}_t(m)^{k-1}\\ p_k\in\mathcal{P}_t(m)}} |R(n,p_k)| \sup_{|u|,|v|\le\frac{|R(n,p_k)|}{\sqrt{n}}} (|\Delta X_{S_{p_k}} + u|^p + |\Delta X_{S_{p_k}} + v|^{p-1}) \\
&\qquad \times \prod_{r=1}^{k-1} \Big|\Delta X_{S_{p_r}} + \frac{R(n,p_r)}{\sqrt{n}}\Big|^p \sum_{\boldsymbol{i}\in\mathcal{B}^n_t(l-k)} \prod_{j=1}^{l-k} |\Delta^n_{i_j}X(m)|^p \\
&=: K\Phi^n_1(m) \times \Phi^n_2(m).
\end{aligned}$$

The first factor $\Phi_1^n(m)$ converges, as $n \to \infty$, stably in law towards

$$\Phi_1(m) = \sum_{\substack{\boldsymbol{p}\in\mathcal{P}_t(m)^{k-1} \\ p_k\in\mathcal{P}_t(m)}} |R_{p_k}|(|\Delta X_{S_{p_k}}|^p + |\Delta X_{S_{p_k}}|^{p-1}) \prod_{r=1}^{k-1} |\Delta X_{S_{p_r}}|^p.$$

By the Portmanteau theorem we obtain

$$\limsup_{n\to\infty} \mathbb{P}(|\Phi_1^n(m)| \geq M) \leq \tilde{\mathbb{P}}(|\Phi_1(m)| \geq M) \quad \text{for all} \quad M \in \mathbb{R}_+,$$

whereas, as $m \to \infty$,

$$\Phi_1(m) \xrightarrow{\tilde{\mathbb{P}}} \Big(\sum_{s\leq t} |\Delta X_s|^p\Big)^{k-1} \sum_{p_k\in\mathcal{P}_t} R_{p_k}(|\Delta X_{S_{p_k}}|^p + |\Delta X_{S_{p_k}}|^{p-1}).$$

So it follows that

$$\lim_{M\to\infty} \limsup_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(|\Phi_1^n(m)| \geq M) = 0.$$

Furthermore

$$\begin{aligned}
&\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{E}(\mathbb{1}_{\Omega_n(m)} \Phi_2^n(m)) \\
&\qquad \leq \lim_{m\to\infty} \limsup_{n\to\infty} \frac{K}{m^{(l-k)(p-2)}} \mathbb{E}\Big( \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-k)} \prod_{j=1}^{l-k} |\Delta_{i_j}^n X(m)|^2 \Big) = 0
\end{aligned}$$

by Lemma B.2. We finally obtain

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}\big(\mathbb{1}_{\Omega_n(m)} |\Phi_1^n(m)\Phi_2^n(m)| > \eta\big) = 0 \qquad \text{for all } \eta > 0.$$

Doing these steps successively in the first $k-1$ components as well, we obtain

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}\Big(\mathbb{1}_{\Omega_n(m)} \Big| \binom{l}{k}^{-1} \zeta_{k,0}^n(m) - \theta_k^n(m) \Big| > \eta\Big) = 0 \qquad \text{for all } \eta > 0$$

with

$$\theta_k^n(m) := \frac{\sqrt{n}}{n^{d-l}} \sum_{\boldsymbol{p}\in\mathcal{P}_t(m)^k} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d-k)} H\Big(\Delta X_{S_{\boldsymbol{p}}}, \Delta_{\boldsymbol{i}}^n X(m)\Big).$$

By the same arguments as in the proof of the convergence $\mathbb{1}_{\Omega_n(m)} \Psi_1^n(m) \xrightarrow{\mathbb{P}} 0$ in Section B.3 we see that we can replace the last $d-l$ variables of $H$ in $\mathbb{1}_{\Omega_n(m)} \theta_k^n(m)$ by 0 without changing the limit. So we can restrict ourselves without loss of generality to the case $l = d$ now and have to prove

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(\mathbb{1}_{\Omega_n(m)} |\Theta_k^n(m)| > \eta) = 0 \tag{B.43}$$

with

$$\begin{aligned}
\Theta_k^n(m) := \sqrt{n} \sum_{\boldsymbol{p}\in\mathcal{P}_t(m)^k} \Big( &\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(d-k)} H\Big(\Delta X_{S_{\boldsymbol{p}}}, \Delta_{\boldsymbol{i}}^n X(m)\Big) \\
&- \sum_{\boldsymbol{s}\in(0,\frac{\lfloor nt\rfloor}{n}]^{d-k}} H\Big(\Delta X_{S_{\boldsymbol{p}}}, \Delta X(m)_{\boldsymbol{s}}\Big)\Big).
\end{aligned}$$

Since
$$\sum_{q\in\mathcal{P}_t(m)} |\Delta X_{S_q}|^p \le \sum_{s\le t} |\Delta X_s|^p$$
is bounded in probability, we can adopt exactly the same method as in the proof of $\mathbb{1}_{\Omega_n(m)}\Psi_2^n(m) \xrightarrow{\mathbb{P}} 0$ to show (B.43), which finishes the proof of Proposition B.7. $\square$

*Proof of Proposition B.14.* We will only show that we can replace $\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c$ by $\alpha_{\boldsymbol{i}}^n$ in the first argument, i.e. the convergence of

$$\zeta_n := \frac{\sqrt{n}}{n^l} \sum_{k=1}^{\lfloor nt\rfloor} \sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)} \sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)} \Big( H(\sqrt{n}\Delta_k^n X^c, \sqrt{n}\Delta_{\boldsymbol{i}}^n X^c, \Delta_{\boldsymbol{j}}^n X) - H(\alpha_k^n, \sqrt{n}\Delta_{\boldsymbol{i}}^n X^c, \Delta_{\boldsymbol{j}}^n X)\Big) \tag{B.44}$$

to 0 in probability. All the other replacements follow in the same manner. Define the function $g:\mathbb{R}^d\to\mathbb{R}$ by $g(w,\boldsymbol{x},\boldsymbol{y}) = |w|^{p_1}L(w,\boldsymbol{x},\boldsymbol{y})$. In a first step we will show that, for fixed $M>0$, we have

$$\frac{1}{\sqrt{n}} \sup_{\|\boldsymbol{z}\|\le M} \sum_{k=1}^{\lfloor nt\rfloor} \big(g(\sqrt{n}\Delta_k^n X^c, \boldsymbol{z}) - g(\alpha_k^n, \boldsymbol{z})\big) \xrightarrow{\mathbb{P}} 0, \tag{B.45}$$

where $\boldsymbol{z} = (\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}^{l-1}\times\mathbb{R}^{d-l}$. Note that our growth assumptions on $L$ imply the existence of constants $h,h',h''\ge 0$ such that

$$w\neq 0 \implies |\partial_1 g(w,\boldsymbol{x},\boldsymbol{y})| \le Ku(\boldsymbol{y})(1+\|(w,\boldsymbol{x})\|^h)\big(1+|w|^{p_1-1}\big) \tag{B.46}$$

$$\begin{aligned} w\neq 0, |z|\le |w|/2 \implies\ & |\partial_1 g(w+z,\boldsymbol{x},\boldsymbol{y}) - \partial_1 g(w,\boldsymbol{x},\boldsymbol{y})| \\ & \le Ku(\boldsymbol{y})|z|(1+\|(w,\boldsymbol{x})\|^{h'} + |z|^{h'})\big(1+|w|^{p_1-2}\big) \end{aligned}$$

$$|g(w+z,\boldsymbol{x},\boldsymbol{y}) - g(w,\boldsymbol{x},\boldsymbol{y})| \le Ku(\boldsymbol{y})(1+\|(w,\boldsymbol{x})\|^{h''})|z|^{p_1} \tag{B.47}$$

The first inequality is trivial, the second one follows by using the mean value theorem, and the last one can be deduced by the same arguments as in the derivation of (B.26). In particular, for fixed $\boldsymbol{x},\boldsymbol{y}$ all assumptions of [12, Theorem 5.3.6] are fulfilled and hence

$$\frac{1}{\sqrt{n}} \max_{\boldsymbol{z}\in K_m(M)} \sum_{k=1}^{\lfloor nt\rfloor} \big(g(\sqrt{n}\Delta_k^n X^c, \boldsymbol{z}) - g(\alpha_k^n, \boldsymbol{z})\big) \xrightarrow{\mathbb{P}} 0,$$

where $K_m(M)$ is defined to be a finite subset of $[-M,M]^{d-1}$ such that for each $\boldsymbol{z}\in[-M,M]^{d-l}$ there exists $\boldsymbol{z}'\in K_m(M)$ with $\|\boldsymbol{z}-\boldsymbol{z}'\|\le 1/m$. In order to show (B.45) it is therefore enough to prove

$$\frac{1}{\sqrt{n}} \sup_{\substack{\|(\boldsymbol{z_1},\boldsymbol{z_2})\|\le M \\ \|\boldsymbol{z}_1-\boldsymbol{z}_2\|\le 1/m}} \Big| \sum_{k=1}^{\lfloor nt\rfloor} \big(g(\sqrt{n}\Delta_k^n X^c, \boldsymbol{z}_1) - g(\alpha_k^n, \boldsymbol{z}_1) - (g(\sqrt{n}\Delta_k^n X^c, \boldsymbol{z}_2) - g(\alpha_k^n, \boldsymbol{z}_2))\big)\Big| \xrightarrow{\mathbb{P}} 0$$

if we first let $n$ and then $m$ go to infinity.

Now, let $\theta_k^n = \sqrt{n}\Delta_k^n X^c - \alpha_k^n$ and $B_k^n = \{|\theta_k^n| \leq |\alpha_k^n|/2\}$. Clearly, $g$ is differentiable in the last $d-1$ arguments and on $B_k^n$ we can also apply the mean value theorem in the first argument. We therefore get

$$\begin{aligned}
&\mathbb{1}_{B_k^n}\big(g(\sqrt{n}\Delta_k^n X^c, \boldsymbol{z}_1) - g(\alpha_k^n, \boldsymbol{z}_1) - (g(\sqrt{n}\Delta_k^n X^c, \boldsymbol{z}_2) - g(\alpha_k^n, \boldsymbol{z}_2))\big) \\
&\quad = \sum_{j=2}^{d} \mathbb{1}_{B_k^n} \partial_1 \partial_j g(\chi_{j,k}^n, \boldsymbol{\xi}_{j,k}^n)(z_2^{(j)} - z_1^{(j)})\theta_k^n,
\end{aligned}$$

where $\chi_{j,k}^n$ is between $\sqrt{n}\Delta_k^n X^c$ and $\alpha_k^n$ and $\xi_{j,k}^n$ is between $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$. $z_i^{(j)}$ stands for the $j$-th component of $\boldsymbol{z}_i$. We have $|\partial_1 \partial_j g(w, \boldsymbol{z})| \leq p_1 |w|^{p_1 - 1} |\partial_j L(w, \boldsymbol{z})| + |w|^{p_1} |\partial_1 \partial_j L(w, \boldsymbol{z})|$ and therefore the growth conditions on $L$ imply that there exists $q \geq 0$ such that

$$|\partial_1 \partial_j g(w, \boldsymbol{z})| \leq K u(\boldsymbol{y})(1 + |w|^{p_1 - 1})(1 + \|(w, \boldsymbol{x})\|^q).$$

On $B_k^n$ we have $|\chi_{j,k}^n| \leq \frac{3}{2}|\alpha_k^n|$. From $\|\boldsymbol{z}\| \leq M$ we find

$$\begin{aligned}
&\frac{1}{\sqrt{n}} \mathbb{E}\Big( \sup_{\substack{\|(\boldsymbol{z_1}, \boldsymbol{z_2})\| \leq M \\ \|\boldsymbol{z}_1 - \boldsymbol{z}_2\| \leq 1/m}} \Big| \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{1}_{B_k^n} \sum_{j=1}^{2} (-1)^{j-1} \big( g(\sqrt{n}\Delta_k^n X^c, \boldsymbol{z}_j) - g(\alpha_k^n, \boldsymbol{z}_j) \big) \Big| \Big) \\
&\qquad \leq \frac{K(M)}{\sqrt{n} m} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E}\Big( (1 + |\alpha_k^n|^{p_1 - 1})(1 + |\alpha_k^n|^q + |\sqrt{n}\Delta_k^n X^c|^q)|\theta_k^n| \Big).
\end{aligned}$$

By Burkholder inequality we know that $\mathbb{E}\big((1 + |\alpha_k^n|^q + |\sqrt{n}\Delta_k^n X^c|^q)^u\big) \leq K$ for all $u \geq 0$. Since $\sigma$ is a continuous semimartingale we further have $\mathbb{E}(|\theta_k^n|^u) \leq K n^{-u/2}$ for $u \geq 1$. Finally, because $\sigma$ is bounded away from 0, we also have $\mathbb{E}\big((|\alpha_k^n|^{p_1 - 1})^u\big) \leq K$ for all $u \geq 0$ with $u(1 - p_1) < 1$. Using this results in combination with Hölder inequality we obtain

$$\frac{K(M)}{\sqrt{n} m} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E}\Big( (1 + |\alpha_k^n|^{p_1 - 1})(1 + |\alpha_k^n|^q + |\sqrt{n}\Delta_k^n X^c|^q)|\theta_k^n| \Big) \leq \frac{K(M)}{m},$$

which converges to 0 as $m \to \infty$.

Now we focus on $(B_k^n)^C$. Let $2 \leq j \leq d$. Observe that, similarly to (B.26), by distinguishing the cases $|z| \leq 1$ and $|z| > 1$, we find that

$$|\partial_j L(w + z, \boldsymbol{x}, \boldsymbol{y}) - \partial_j L(w, \boldsymbol{x}, \boldsymbol{y})| \leq K(1 + |w|^{\gamma_j + \gamma_{1j}})|z|^{\gamma_j}.$$

We used here that $\|(\boldsymbol{x}, \boldsymbol{y})\|$ is bounded and the simple inequality $1 + a + b \leq 2(1 + a)b$ for all $a \geq 0, b \geq 1$. From this we get

$$\begin{aligned}
&|\partial_j g(w + z, \boldsymbol{x}, \boldsymbol{y}) - \partial_j g(w, \boldsymbol{x}, \boldsymbol{y})| \\
&\quad \leq \big| |w + z|^{p_1} - |w|^{p_1} \big| |\partial_j L(w + z, \boldsymbol{x}, \boldsymbol{y})| \\
&\qquad + |w|^{p_1} |\partial_j L(w + z, \boldsymbol{x}, \boldsymbol{y}) - \partial_j L(w, \boldsymbol{x}, \boldsymbol{y})| \\
&\quad \leq K(1 + |w|^q)(|z|^{\gamma_j + p_1} + |z|^{\gamma_j})
\end{aligned}$$

for some $q \geq 0$. Recall that $\gamma_j < 1$ and $\gamma_j + p_1 < 1$ by assumption. For some $\xi_j^n$ between $z_1^{(j)}$ and $z_2^{(j)}$ we therefore have

$$\frac{1}{\sqrt{n}}\mathbb{E}\Big( \sup_{\substack{\|(\boldsymbol{z_1},\boldsymbol{z_2})\|\leq M \\ \|\boldsymbol{z}_1-\boldsymbol{z}_2\|\leq 1/m}} \Big|\sum_{k=1}^{\lfloor nt\rfloor} \mathbb{1}_{(B_k^n)^C} \sum_{j=1}^{2}(-1)^{j-1}\big(g(\sqrt{n}\Delta_k^n X^c, \boldsymbol{z}_j) - g(\alpha_k^n, \boldsymbol{z}_j)\big)\Big|\Big)$$
$$= \frac{1}{\sqrt{n}}\mathbb{E}\Big( \sup_{\substack{\|(\boldsymbol{z_1},\boldsymbol{z_2})\|\leq M \\ \|\boldsymbol{z}_1-\boldsymbol{z}_2\|\leq 1/m}} \Big|\sum_{k=1}^{\lfloor nt\rfloor}\sum_{j=2}^{d}$$
$$\mathbb{1}_{(B_k^n)^C}\big(\partial_j g(\sqrt{n}\Delta_k^n X^c, \xi_j^n) - \partial_j g(\alpha_k^n, \xi_j^n)\big)(z_2^{(j)} - z_1^{(j)})\Big|\Big)$$
$$\leq \frac{K(M)}{\sqrt{n}m}\sum_{k=1}^{\lfloor nt\rfloor}\mathbb{E}\Big(\mathbb{1}_{(B_k^n)^C}(1+|\alpha_k^n|^q + |\sqrt{n}\Delta_k^n X^c|^q)(|\theta_k^n|^{\gamma_1} + |\theta_k^n|^{\gamma_j+p_1})\Big)$$
$$\leq \frac{K(M)}{\sqrt{n}m}\sum_{k=1}^{\lfloor nt\rfloor}\mathbb{E}\Big(\mathbb{1}_{(B_k^n)^C}(1+|\alpha_k^n|^q + |\sqrt{n}\Delta_k^n X^c|^q)\Big(\frac{|\theta_k^n|}{|\alpha_k^n|^{1-\gamma_1}} + \frac{|\theta_k^n|}{|\alpha_k^n|^{1-(\gamma_j+p_1)}}\Big)\Big)$$
$$\leq \frac{K(M)}{m}$$

by the same arguments as before, and hence (B.45) holds. For any $M > 2A$ we therefore have (with $\boldsymbol{q} = (q_1, \dots, q_{d-l})$)

$$\Big|\frac{\sqrt{n}}{n^l}\sum_{k=1}^{\lfloor nt\rfloor}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)}\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)}\mathbb{1}_{\{\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\|\leq M\}}$$
$$\times\Big(H(\sqrt{n}\Delta_k^n X^c, \sqrt{n}\Delta_{\boldsymbol{i}}^n X^c, \Delta_{\boldsymbol{j}}^n X) - H(\alpha_k^n, \sqrt{n}\Delta_{\boldsymbol{i}}^n X^c, \Delta_{\boldsymbol{j}}^n X)\Big)\Big|$$
$$\leq \Big(\frac{1}{n^{l-1}}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)}\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)}|\sqrt{n}\Delta_{i_1}X^c|^{p_2}\cdots|\sqrt{n}\Delta_{i_{l-1}}X^c|^{p_l}|\Delta_{\boldsymbol{j}}^n X|^{\boldsymbol{q}}\Big)$$
$$\times\Big|\frac{1}{\sqrt{n}}\sup_{\|\boldsymbol{z}\|\leq M}\sum_{k=1}^{\lfloor nt\rfloor}\big(g(\sqrt{n}\Delta_k^n X^c, \boldsymbol{z}) - g(\alpha_k^n, \boldsymbol{z})\big)\Big|$$

The first factor converges in probability to some finite limit, and hence the whole expression converges to 0 by (B.45). In order to show (B.44) we are therefore left with proving that

$$\frac{\sqrt{n}}{n^l}\sum_{k=1}^{\lfloor nt\rfloor}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)}\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)}\mathbb{1}_{\{\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\|>M\}}$$
$$\times\Big(H(\sqrt{n}\Delta_k^n X^c, \sqrt{n}\Delta_{\boldsymbol{i}}^n X^c, \Delta_{\boldsymbol{j}}^n X) - H(\alpha_k^n, \sqrt{n}\Delta_{\boldsymbol{i}}^n X^c, \Delta_{\boldsymbol{j}}^n X)\Big)$$

converges in probability to 0 if we first let $n$ and then $M$ go to infinity. As before we will distinguish between the cases that we are on the set $B_k^n$ and on $(B_k^n)^C$. Let $\tilde{\boldsymbol{p}} = (p_2, \dots, p_l)$. With the mean value theorem and the growth properties of $\partial_1 g$

from (B.46) we obtain for all $M \geq 1$:

$$
\begin{aligned}
&\Big|\frac{\sqrt{n}}{n^l}\sum_{k=1}^{\lfloor nt\rfloor}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)}\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)}\mathbb{1}_{\{\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\|>M\}}\mathbb{1}_{B_k^n}\\
&\quad\times\Big(H(\sqrt{n}\Delta_k^n X^c,\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c,\Delta_{\boldsymbol{j}}^n X)-H(\alpha_k^n,\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c,\Delta_{\boldsymbol{j}}^n X)\Big)\Big|\\
&\quad\le K\Big(\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)}|\Delta_{\boldsymbol{j}}^n X|^{\boldsymbol{q}}\Big)\frac{\sqrt{n}}{n^l}\sum_{k=1}^{\lfloor nt\rfloor}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)}|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c|^{\tilde{\boldsymbol{p}}}\mathbb{1}_{\{\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\|>M\}}\\
&\qquad\times(1+|\alpha_k^n|^h+|\sqrt{n}\Delta_k^n X^c|^h+\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\|^h)\big(1+|\alpha_k^n|^{p_1-1}\big)|\theta_k^n|\\
&\quad\le\Big(K\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)}|\Delta_{\boldsymbol{j}}^n X|^{\boldsymbol{q}}\Big)\\
&\qquad\times\Big(\frac{1}{n^{l-1}}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)}\mathbb{1}_{\{\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\|>M\}}|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c|^{\tilde{\boldsymbol{p}}}\,\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\|^h\Big)\\
&\qquad\times\Big(\frac{1}{\sqrt{n}}\sum_{k=1}^{\lfloor nt\rfloor}(1+|\alpha_k^n|^h+|\sqrt{n}\Delta_k^n X^c|^h)\big(1+|\alpha_k^n|^{p_1-1}\big)|\theta_k^n|\Big)\\
&\quad=:A_nB_n(M)C_n,
\end{aligned}
$$

where we used $M \geq 1$ and $1+a+b \leq 2(1+a)b$ for the final inequality again. As before, we deduce that $A_n$ is bounded in probability and $\mathbb{E}(C_n) \leq K$. We also have $\mathbb{E}(B_n(M)) \leq K/M$ and hence

$$
\lim_{M\to\infty}\limsup_{n\to\infty}\mathbb{P}(A_nB_n(M)C_n>\eta)=0
$$

for all $\eta > 0$.
Again, with (B.47), we derive for $M \geq 1$:

$$
\begin{aligned}
&\Big|\frac{\sqrt{n}}{n^l}\sum_{k=1}^{\lfloor nt\rfloor}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)}\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)}\mathbb{1}_{\{\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\|>M\}}\mathbb{1}_{(B_k^n)^C}\\
&\quad\times\Big(H(\sqrt{n}\Delta_k^n X^c,\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c,\Delta_{\boldsymbol{j}}^n X)-H(\alpha_k^n,\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c,\Delta_{\boldsymbol{j}}^n X)\Big)\Big|\\
&\quad=\frac{\sqrt{n}}{n^l}\sum_{k=1}^{\lfloor nt\rfloor}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)}\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)}\mathbb{1}_{\{\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\|>M\}}\mathbb{1}_{(B_k^n)^C}|\Delta_{\boldsymbol{j}}^n X|^{\boldsymbol{q}}|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c|^{\tilde{\boldsymbol{p}}}\\
&\qquad\times\Big|g(\alpha_k^n+\theta_k^n,\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c,\Delta_{\boldsymbol{j}}^n X)-g(\alpha_k^n,\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c,\Delta_{\boldsymbol{j}}^n X)\Big|\\
&\quad\le K\Big(\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)}|\Delta_{\boldsymbol{j}}^n X|^{\boldsymbol{q}}\Big)\\
&\qquad\times\Big(\frac{1}{n^{l-1}}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)}\mathbb{1}_{\{\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\|>M\}}|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c|^{\tilde{\boldsymbol{p}}}\,\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\|^{h''}\Big)\\
&\qquad\times\Big(\frac{1}{\sqrt{n}}\sum_{k=1}^{\lfloor nt\rfloor}\mathbb{1}_{(B_k^n)^C}(1+|\alpha_k^n|^{h''})|\theta_k^n|^{p_1}\Big)
\end{aligned}
$$

$$\leq K\Big(\sum_{\boldsymbol{j}\in\mathcal{B}_t^n(d-l)}|\Delta_{\boldsymbol{j}}^n X|^{\boldsymbol{q}}\Big)$$
$$\times\Big(\frac{1}{n^{l-1}}\sum_{\boldsymbol{i}\in\mathcal{B}_t^n(l-1)}\mathbb{1}_{\{\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\|>M\}}|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c|^{\tilde{\boldsymbol{p}}}\left\|\sqrt{n}\Delta_{\boldsymbol{i}}^n X^c\right\|^{h''}\Big)$$
$$\times\Big(\frac{1}{\sqrt{n}}\sum_{k=1}^{\lfloor nt\rfloor}(1+|\alpha_k^n|^{h''})|\alpha_k^n|^{p_1-1}|\theta_k^n|\Big).$$

For the last step, recall that we have the estimate $|\theta_k^n|^{1-p_1}\leq K|\alpha_k^n|^{1-p_1}$ on the set $(B_k^n)^C$. Once again, the final random variable converges to 0 if we first let $n$ and then $M$ to infinity. □

*Proof of Proposition B.15:.* We will give a proof only in the case $d=2$ and $l=1$. We use the decomposition

$$\begin{aligned}
&\mathbb{1}_{\Omega_n(m)}\theta_n(H)\\
&\quad=\frac{\mathbb{1}_{\Omega_n(m)}}{\sqrt{n}}\Big(\sum_{i,j=1}^{\lfloor nt\rfloor}H(\alpha_i^n,\Delta_j^n X(m))-\sum_{i=1}^{\lfloor nt\rfloor}\sum_{s\leq\frac{\lfloor nt\rfloor}{n}}H(\alpha_i^n,\Delta X(m)_s)\Big)\\
&\quad\quad-\frac{\mathbb{1}_{\Omega_n(m)}}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}\sum_{\frac{\lfloor nt\rfloor}{n}<s\leq t}H(\alpha_i^n,\Delta X_s)\\
&\quad\quad+\frac{\mathbb{1}_{\Omega_n(m)}}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}\sum_{p\in\mathcal{P}_t^n(m)}\\
&\quad\quad\quad\quad\Big\{H\big(\alpha_i^n,\Delta X_{S_p}+n^{-1/2}R(n,p)\big)-H\big(\alpha_i^n,n^{-1/2}R(n,p)\big)\Big\}\\
&\quad\quad-\frac{\mathbb{1}_{\Omega_n(m)}}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}\sum_{p\in\mathcal{P}_t^n(m)}H(\alpha_i^n,\Delta X_{S_p})\\
&\quad=:\theta_n^{(1)}(H)-\theta_n^{(2)}(H)+\theta_n^{(3)}(H)-\theta_n^{(4)}(H).
\end{aligned}$$

In the general case we would have to use the decomposition given in (B.12) for the last $d-l$ arguments. We first show that we have

$$\lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{P}(|\theta_n^{(1)}(H)|>\eta)=0\quad\text{for all}\quad\eta>0.$$

We do this in two steps.

**(i)** Let $\phi_k$ be a function in $\mathcal{C}^\infty(\mathbb{R}^2)$ with $0\leq\phi_k\leq 1$, $\phi_k\equiv 1$ on $[-k,k]^2$, and $\phi_k\equiv 0$ outside of $[-2k,2k]^2$. Also, let $\tilde{g}:\mathbb{R}^2\to\mathbb{R}$ be defined by $\tilde{g}(x,y)=|y|^{q_1}L(x,y)$

and set $H_k = \phi_k H$ and $\tilde{g}_k = \phi_k \tilde{g}$. Then we have

$$
\begin{aligned}
&|\theta_n^{(1)}(H_k)| \\
&\quad= \Big| \frac{\mathbb{1}_{\Omega_n(m)}}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} |\alpha_i^n|^{p_1} \Big( \sum_{j=1}^{\lfloor nt \rfloor} \tilde{g}_k(\alpha_i^n, \Delta_j^n X(m)) - \sum_{s \le \frac{\lfloor nt \rfloor}{n}} \tilde{g}_k(\alpha_i^n, \Delta X(m)_s) \Big) \Big| \\
&\quad\le \Big| \frac{\mathbb{1}_{\Omega_n(m)}}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} |\alpha_i^n|^{p_1} \Big( \sum_{j=1}^{\lfloor nt \rfloor} \tilde{g}_k(0, \Delta_j^n X(m)) - \sum_{s \le \frac{\lfloor nt \rfloor}{n}} \tilde{g}_k(0, \Delta X(m)_s) \Big) \Big| \\
&\qquad + \Big| \frac{\mathbb{1}_{\Omega_n(m)}}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} |\alpha_i^n|^{p_1} \Big( \sum_{j=1}^{\lfloor nt \rfloor} \int_0^{\alpha_i^n} \partial_1 \tilde{g}_k(u, \Delta_j^n X(m)) du \\
&\qquad\qquad - \sum_{s \le \frac{\lfloor nt \rfloor}{n}} \int_0^{\alpha_i^n} \partial_1 \tilde{g}_k(u, \Delta X(m)_s) du \Big) \Big| \\
&\quad\le \Big( \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} |\alpha_i^n|^{p_1} \Big) \Big( \sqrt{n} \mathbb{1}_{\Omega_n(m)} \Big| \sum_{j=1}^{\lfloor nt \rfloor} \tilde{g}_k(0, \Delta_j^n X(m)) - \sum_{s \le \frac{\lfloor nt \rfloor}{n}} \tilde{g}_k(0, \Delta X(m)_s) \Big| \Big) \\
&\qquad + \Big( \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} |\alpha_i^n|^{p_1} \Big) \Big( \mathbb{1}_{\Omega_n(m)} \int_{-k}^{k} \sqrt{n} \\
&\qquad\qquad \times \Big| \sum_{j=1}^{\lfloor nt \rfloor} \partial_1 \tilde{g}_k(u, \Delta_j^n X(m)) - \sum_{s \le \frac{\lfloor nt \rfloor}{n}} \partial_1 \tilde{g}_k(u, \Delta X(m)_s) \Big| du \Big),
\end{aligned}
$$

which converges to zero in probability by Lemma B.3, if we first let $n \to \infty$ and then $m \to \infty$, since

$$
\frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} |\alpha_i^n|^{p_1}
$$

is bounded in probability by Burkholder inequality.

**(ii)** In this part we show

$$
\lim_{k \to \infty} \lim_{m \to \infty} \limsup_{n \to \infty} \mathbb{P}(|\theta_n^{(1)}(H) - \theta_n^{(1)}(H_k)| > \eta) = 0 \quad \text{for all} \quad \eta > 0.
$$

Observe that we automatically have $|\Delta_i^n X(m)| \le k$ for some $k$ large enough. Therefore,

$$
\begin{aligned}
&|\theta_n^{(1)}(H) - \theta_n^{(1)}(H_k)| \\
&\quad= \Big| \frac{\mathbb{1}_{\Omega_n(m)}}{\sqrt{n}} \sum_{i,j=1}^{\lfloor nt \rfloor} \Big( H(\alpha_i^n, \Delta_j^n X(m)) - H_k(\alpha_i^n, \Delta_j^n X(m)) \Big) \Big| \\
&\quad\le \frac{\mathbb{1}_{\Omega_n(m)}}{\sqrt{n}} \sum_{i,j=1}^{\lfloor nt \rfloor} \mathbb{1}_{\{|\alpha_i^n| > k\}} \Big| H(\alpha_i^n, \Delta_j^n X(m)) - H_k(\alpha_i^n, \Delta_j^n X(m)) \Big|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{\mathbb{1}_{\Omega_n(m)}}{\sqrt{n}} \sum_{i,j=1}^{\lfloor nt \rfloor} \mathbb{1}_{\{|\alpha_i^n|>k\}} \Big| H(\alpha_i^n, \Delta_j^n X(m)) \Big| \\
&\leq \frac{K \mathbb{1}_{\Omega_n(m)}}{\sqrt{n}} \sum_{i,j=1}^{\lfloor nt \rfloor} \mathbb{1}_{\{|\alpha_i^n|>k\}} \big| (1+|\alpha_i^n|^{p_1})(\Delta_j^n X(m))^{q_1} \big| \\
&\leq \frac{K}{\sqrt{n}} \Big( \sum_{i=1}^{\lfloor nt \rfloor} \mathbb{1}_{\{|\alpha_i^n|>k\}} (1+|\alpha_i^n|^{p_1}) \Big) \Big( \mathbb{1}_{\Omega_n(m)} \sum_{j=1}^{\lfloor nt \rfloor} \big| \Delta_j^n X(m) \big|^{q_1} \Big) \\
&\leq K \Big( \sum_{i=1}^{\lfloor nt \rfloor} \mathbb{1}_{\{|\alpha_i^n|>k\}} \Big)^{\frac{1}{2}} \Big( \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} (1+|\alpha_i^n|^{p_1})^2 \Big)^{\frac{1}{2}} \Big( \mathbb{1}_{\Omega_n(m)} \sum_{j=1}^{\lfloor nt \rfloor} \big| \Delta_j^n X(m) \big|^{q_1} \Big)
\end{aligned}
$$

Now observe that we have

$$
\Big( \mathbb{1}_{\Omega_n(m)} \sum_{j=1}^{\lfloor nt \rfloor} \big| \Delta_j^n X(m) \big|^{q_1} \Big) \xrightarrow{\mathbb{P}} \sum_{s \leq t} |\Delta X_s|^{q_1},
$$

if we first let $n \to \infty$ and then $m \to \infty$. Further we have

$$
\mathbb{E}\Big[ \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} (1+|\alpha_i^n|^{p_1})^2 \Big] \leq K
$$

by Burkholder inequality and finally

$$
\mathbb{P}\Big( \Big| \sum_{i=1}^{\lfloor nt \rfloor} \mathbb{1}_{\{|\alpha_i^n|>k\}} \Big| > \eta \Big) \leq \frac{1}{\eta} \mathbb{E}\Big( \sum_{i=1}^{\lfloor nt \rfloor} \mathbb{1}_{\{|\alpha_i^n|>k\}} \Big) \leq \sum_{i=1}^{\lfloor nt \rfloor} \frac{\mathbb{E}[|\alpha_i^n|^2]}{\eta k^2} \leq \frac{K}{\eta k^2} \to 0,
$$

as $k \to \infty$. For $\theta_n^{(2)}(H)$ we have

$$
\begin{aligned}
|\theta_n^{(2)}(H)| &\leq \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{\frac{\lfloor nt \rfloor}{n} < s \leq t} (1+|\alpha_i^n|^{p_1}) |\Delta X_s|^{q_1} u(\Delta X_s) \\
&\leq \Big( \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} (1+|\alpha_i^n|^{p_1}) \Big) \Big( \sqrt{n} \sum_{\frac{\lfloor nt \rfloor}{n} < s \leq t} |\Delta X_s|^{q_1} \Big) \xrightarrow{\mathbb{P}} 0,
\end{aligned}
$$

since the first factor is bounded in expectation and the second one converges in probability to 0 (see (B.19)). For the second summand of $\theta_n^{(3)}(H)$ we get

$$
\begin{aligned}
&\Big| \frac{\mathbb{1}_{\Omega_n(m)}}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{p \in \mathcal{P}_t^n(m)} H\big( \alpha_i^n, n^{-1/2} R(n,p) \big) \Big| \\
&\quad \leq \Big( \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} (1+|\alpha_i^n|^{p_1}) \Big) \Big( \mathbb{1}_{\Omega_n(m)} \sum_{p \in \mathcal{P}_t^n(m)} \Big| \frac{R(n,p)^{q_1}}{n^{\frac{1}{2}(q_1-1)}} \Big| \Big) \xrightarrow{\mathbb{P}} 0
\end{aligned}
$$

as $n \to \infty$ because the first factor is again bounded in expectation and since $(R(n,p))_{n\in\mathbb{N}}$ is bounded in probability and $\mathcal{P}_t^n(m)$ finite almost surely. The remaining terms are $\theta_n^{(4)}(H)$ and the first summand of $\theta_n^{(3)}(H)$, for which we find by the mean value theorem

$$\begin{aligned}
&\frac{\mathbb{1}_{\Omega_n(m)}}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{p\in\mathcal{P}_t^n(m)} \Big\{ H\big(\alpha_i^n, \Delta X_{S_p} + n^{-1/2} R(n,p)\big) - H(\alpha_i^n, \Delta X_{S_p}) \Big\} \\
&\quad = \frac{\mathbb{1}_{\Omega_n(m)}}{n} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{p\in\mathcal{P}_t^n(m)} \partial_2 H\big(\alpha_i^n, \Delta X_{S_p}\big) R(n,p) \\
&\qquad + \Big( \frac{\mathbb{1}_{\Omega_n(m)}}{n} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{p\in\mathcal{P}_t^n(m)} \big(\partial_2 H\big(\alpha_i^n, \Delta X_{S_p} + \xi_i^n(p)\big) - \partial_2 H\big(\alpha_i^n, \Delta X_{S_p}\big)\big) R(n,p) \Big)
\end{aligned}$$

for some $\xi_i^n(p)$ between 0 and $R(n,p)/\sqrt{n}$. The latter term converges to 0 in probability since we have $|\partial_{22} H(x,y)| \le (1+|x|^q)(|y|^{q_1} + |y|^{q_1-1} + |y|^{q_1-2}) u(y)$ for some $q \ge 0$ by the growth assumptions on $L$. Therefore,

$$\begin{aligned}
&\Big| \frac{\mathbb{1}_{\Omega_n(m)}}{n} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{p\in\mathcal{P}_t^n(m)} \big(\partial_2 H\big(\alpha_i^n, \Delta X_{S_p} + \xi_i^n(p)\big) - \partial_2 H\big(\alpha_i^n, \Delta X_{S_p}\big)\big) R(n,p) \Big| \\
&\quad = \Big| \frac{\mathbb{1}_{\Omega_n(m)}}{n} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{p\in\mathcal{P}_t^n(m)} \partial_{22} H\big(\alpha_i^n, \Delta X_{S_p} + \tilde{\xi}_i^n(p)\big) \xi_i^n(p) R(n,p) \Big| \\
&\quad \le \Big( \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} (1 + |\alpha_i^n|^{p_1}) \Big) \sum_{p\in\mathcal{P}_t^n(m)} K \frac{|R(n,p)|^2}{\sqrt{n}} \xrightarrow{\mathbb{P}} 0,
\end{aligned}$$

where $\tilde{\xi}_i^n(p)$ is between 0 and $R(n,p)/\sqrt{n}$. The last inequality holds since the jumps of $X$ are bounded and $|\tilde{\xi}_i^n(p)| \le |R(n,p)|/\sqrt{n} \le 2A$. The convergence holds because $R(n,p)$ is bounded in probability and $\mathcal{P}_t^n(m)$ is finite almost surely. □

# Bibliography


[1] E. Beutner and H. Zähle. Deriving the asymptotic distribution of U- and V-statistics of dependent data using weighted empirical processes. *Bernoulli*, 18(3):803–822, 2012.

[2] E. Beutner and H. Zähle. Continuous mapping approach to the asymptotics of U- and V -statistics. *Bernoulli*, 20(2):846–877, 2014.

[3] P. J. Bickel and M. J. Wichura. Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.*, 42(5):1656–1670, 1971.

[4] S. Borovkova, R. Burton, and H. Dehling. Limit theorems for functionals of mixing processes with applications to u-statistics and dimension estimation. *Trans. Amer. Math. Soc.*, 353:4261–4318, 2001.

[5] H. Dehling and M.S. Taqqu. The empirical process of some long-range dependent sequences with an application to U-statistics. *Ann. Statist.*, 17(4):1767–1783, 1989.

[6] H. Dehling and M.S. Taqqu. Bivariate symmetric statistics of long-range dependent observations. *J. Statist. Plann. Inference*, 28:153–165, 1991.

[7] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Math. Ann.*, 300(1):463–520, 1994.

[8] M. Denker and G. Keller. On U-statistics and v. miseâĂŹ statistics for weakly dependent processes. *Z. Wahrsch. Verw. Gebiete*, 64(4):505–522, 1983.

[9] P.R. Halmos. The theory of unbiased estimation. *Ann. Math. Statist.*, 17(1):34–43, 1946.

[10] W. Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.*, 19:293–325, 1948.

[11] J. Jacod. Asymptotic properties of realized power variations and related functionals of semimartingales. *Stoch. Process. Appl.*, 118:517–559, 2008.

[12] J. Jacod and P. Protter. *Discretization of Processes.* Springer, Berlin, 2012.

[13] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer, Berlin, 2nd edition, 2003.

[14] D. Khoshnevisan. *Multiparameter Processes.* Springer, New York, 2002.

[15] A.J. Lee. *U-Statistics, Theory and Practice.* Dekker, New York, 1990.

[16] A. Leucht. Degenerate U- and V-statistics under weak dependence: Asymptotic theory and bootstrap consistency. *Bernoulli*, 18(2):552–585, 2012.

[17] C. Lévy-Leduc, H. Boistard, E. Moulines, M. S. Taqqu, and V. A. Reisen. Asymptotic properties of U-processes under long-range dependence. *Ann. Statist.*, 39(3):1399–1426, 2011.

[18] M. Podolskij, C. Schmidt, and J. F. Ziegel. Limit theorems for nondegenerate U-statistics of continuous semimartingales. *Ann. Appl. Probab.*, 24(6):2491–2526, 2014.

[19] M. Podolskij and M. Vetter. Understanding limit theorems for semimartingales: a short survey. *Stat. Neerl.*, 64(3):329–351, 2010.

[20] A. Renyi. On stable sequences of events. *Sankhya A*, 25:293–302, 1963.

[21] R.J. Serfling. *Approximation Theorems of Mathematical Statistics.* Wiley, New York, 1980.

[22] R. v. Mises. On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.*, 18(3):309–348, 1947.