

Finite-Sample-Optimized Nonparametric Asymptotic Confidence Intervals for the Mean and Tests for Change in the Mean and Regression Coefficients

by

Jervis Gallanosa

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Statistics)

UNIVERSITY OF MANITOBA

(Winnipeg)

© Jervis Gallanosa 2026

Abstract

Classical asymptotic procedures, such as confidence intervals for the mean and change-point tests, are widely used because they are applicable under weak assumptions. Their finite-sample performance, however, can fall short of what their asymptotic behavior suggests. Confidence intervals may undercover for skewed or heavy-tailed data, while change-point tests may have inflated type I errors or low power for detecting a change near the beginning or end of a sample.

Motivated by these issues, this thesis develops nonparametric asymptotic confidence intervals for the mean, tests for a change in the mean, and tests for changes in regression coefficients. The proposed procedures are based on convergence in distribution of functionals of appropriately weighted and normalized partial-sum-type processes. The central theme is that suitable choices of functionals and weight functions can improve or better balance finite-sample performance. Numerical studies show that the proposed confidence intervals attain higher finite-sample coverage probabilities than the classical CLT-based interval for a range of right-skewed distributions, and that the proposed mean-change tests achieve strong power across all change-point locations while maintaining well-controlled type I errors. For testing for a change in the slope or in the intercept in simple linear regression, the thesis establishes the asymptotic theory needed to pursue similar finite-sample improvements in the regression setting.

Table of Contents

Abstract	iii
Table of Contents	iv
List of Tables	viii
List of Figures	x
Acknowledgements	xi
1 Introduction	1
1.1 Background and Motivation	1
1.2 Main Contributions and Thesis Organization	3
1.3 Contributions of Authors	5
2 Asymptotic Confidence Intervals for the Mean with Increased Finite-Sample Coverage Probabilities	7
Abstract	9
1 Introduction	9
2 Main Results	14
2.1 Preliminaries	14
2.2 FACIs Based on h_2 and h_3	15
3 Conclusions	24
Appendix to Chapter 2	25
References	29

3	Powerful Nonparametric Asymptotic Tests for Change in the Mean with Reduced Type I Errors	31
	Abstract	33
1	Introduction	33
1.1	Nonparametric Asymptotic Tests for H_0 vs. H_A Based on Sup-Functionals of an Appropriately Weighted Tied-Down Partial Sums Process	34
1.2	Nonparametric Asymptotic Tests for H_0 vs. H_A Based on Integral Functionals of an Appropriately Weighted Tied-Down Partial Sums Process	38
1.3	Overview of the Present Work	39
2	Main Results	41
2.1	Preliminaries	41
2.2	Finite-Sample Performance of the M_n^q - and L_n^q -Based Tests	45
2.2.1	Empirical Type I Errors of the M_n^q - and L_n^q -Based Tests	46
2.2.2	Empirical Powers of the M_n^q - and L_n^q -Based Tests	50
3	Conclusions	59
	Appendix to Chapter 3	61
	References	63
4	Nonparametric Asymptotic Tests for a Change in the Slope and Intercept in Simple Linear Regression	65
	Abstract	67
1	Introduction	67
2	Limit Theorems for Change-Point Analysis	70
3	Testing for a Change in the Slope	81
3.1	Model, residual quantities, and assumptions	81
3.2	Main results	84
3.2.1	Supremum statistic with Brownian-bridge-based limit	84
3.2.2	Supremum Darling–Erdős-type statistic	85
3.2.3	Integral statistic with Brownian-bridge-based limit	86

3.2.4	Integral Darling–Erdős-type statistic	86
3.2.5	Joint limit theorem	87
3.2.6	Rényi-type weighted supremum statistic	88
3.3	Proofs of the Slope-Change Results	89
	Auxiliary results for slope-change statistics	89
3.3.1	Proof of Theorem 3.1	98
3.3.2	Proof of Theorem 3.2	104
3.3.3	Proof of Theorem 3.3	117
3.3.4	Proof of Theorem 3.4	128
3.3.5	Proof of Theorem 3.5	138
3.3.6	Proof of Theorem 3.6	147
4	Testing for a Change in the Intercept	153
4.1	Model, residual quantities, and assumptions	153
4.2	Main results	154
4.2.1	Supremum statistic with Brownian-bridge-based limit	154
4.2.2	Supremum Darling–Erdős-type statistic	155
4.2.3	Integral statistic with Brownian-bridge-based limit	156
4.2.4	Integral Darling–Erdős-type statistic	156
4.2.5	Joint limit theorem	156
4.2.6	Rényi-type weighted supremum statistic	158
4.3	Proofs of the Intercept-Change Results	158
	Auxiliary results for intercept-change statistics	158
4.3.1	Proof of Theorem 4.1	160
4.3.2	Proof of Theorem 4.2	161
4.3.3	Proof of Theorem 4.3	163
4.3.4	Proof of Theorem 4.4	165
4.3.5	Proof of Theorem 4.5	168
4.3.6	Proof of Theorem 4.6	171
5	Concluding Remarks	173
	References	176

5 Discussion 178

 1 Summary of Contributions and Novel Developments 178

 2 Overall Impact and Future Directions 179

Full Bibliography 181

List of Tables

Chapter 2

1	Probability density functions for some distributions.	15
2	The $1 - \alpha$ quantile of $\int_0^1 \frac{W^2(t)}{(t(1-t))^{2a}} dt$	15
3	The $1 - \alpha$ quantile of $h_3\left(\frac{W(t)}{q_3(t)}\right) = d_1 \frac{W(0.99)}{q_3(0.99)} + d_2 \int_0^1 \frac{W^2(t)}{q_3^2(t)} dt$	17
4	Comparison of I_1 , $I_{2,q}$ with $q(t) = (t(1-t))^a$, I_3 , and I_4 at 95% confidence level.	18
5	δ_{100} , $\tilde{\delta}_{100}$, and coverages $\widehat{CP}_1$, $\widehat{CP}_{2,q}$, $\widehat{CP}_3$, $\widehat{CP}_4$ ($\cdot 100\%$) for $n = 100$	21
6	Comparison of coverages $\widehat{CP}_3 \cdot 100\%$ of I_3 and $\widehat{CP}_4 \cdot 100\%$ of I_4 (both at 95% confidence level) with coverages $\widehat{CP}_{1,\text{adj}}$ of equal-average-length adjusted I_1 counterparts (at higher nominal levels) for right-skewed distributions. . . .	23
A1	Comparison of I_1 , $I_{2,q}$ with $q(t) = (t(1-t))^a$, I_3 , and I_4 at the 90% confidence level.	25
A2	Comparison of I_1 , $I_{2,q}$ with $q(t) = (t(1-t))^a$, I_3 , and I_4 at the 98% confidence level.	26
A3	Comparison of coverages $\widehat{CP}_3 \cdot 100\%$ of I_3 and $\widehat{CP}_4 \cdot 100\%$ of I_4 (both at 90% confidence level) with coverages $\widehat{CP}_{1,\text{adj}}$ of equal-average-length adjusted I_1 counterparts (at higher nominal levels) for right-skewed distributions. . . .	27
A4	Comparison of coverages $\widehat{CP}_3 \cdot 100\%$ of I_3 and $\widehat{CP}_4 \cdot 100\%$ of I_4 (both at 98% confidence level) with coverages $\widehat{CP}_{1,\text{adj}}$ of equal-average-length adjusted I_1 counterparts (at higher nominal levels) for right-skewed distributions. . . .	28

Chapter 3

1	Distributions with their skewness (if exists), δ_{100} , and probability density functions.	44
2	λ quantiles of $\sup_{0 < t < 1} \frac{ B(t) }{q(t)}$ and $\int_0^1 \frac{B^2(t)}{q(t)} dt$	46
3	The empirical type I errors (%) of the M_n^q and L_n^q tests at the $\alpha = 5\%$ asymptotic significance level.	47
4	The empirical powers for Normal(0, 1), SymPareto(2.1, 1), Weibull(0.4, 1), and Lognormal(0, 1.5) for the change sizes $\Delta = 0.50\sigma$ and 0.75σ	53
A1	The empirical powers for Exponential(1), $t(3)$, and Half- $t(3)$ for the change sizes $\Delta = 0.50\sigma$ and 0.75σ	61
A2	The empirical powers for Frechet(2.5), Inverse-Normal(7, 1), and Pareto(2.1, 1) for the change sizes $\Delta = 0.50\sigma$ and 0.75σ	62

List of Figures

Chapter 3

1	The empirical type I errors of the M_n^q and L_n^q tests for n up to 10,000 (increments of 200).	49
2	The empirical powers of the M_n^q and L_n^q tests as functions of k^*/n at $n = 200$ and the change size $\Delta = 0.50\sigma$	51
3	The empirical powers (averaged over 10 distributions) of the M_n^q and L_n^q tests as functions of the change size Δ	54
4	The empirical power vs. the empirical type I error of the $M_n^{q(\beta)}$ and $L_n^{q(\beta)}$ tests with different β for Lognormal(0, 1.5) and the change size $\Delta = 0.50\sigma$. $M_n^{q(\beta)}$ tests: β increases from -1.2 to -0.6 in steps of 0.05 (right to left). $L_n^{q(\beta)}$ tests: β increases from -5 to -2 in steps of 0.25 (right to left). Highlighted disk and square: $M_n^{q(-0.9)}$ and $L_n^{q(-4)}$ tests. Triangle: M_n^{q2} test.	56

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Dr. Yuliya V. Martsynyuk, for her guidance, support, patience, and encouragement throughout my graduate studies. I am especially grateful for her belief in me from the beginning, for the many discussions we shared over the years, and for her understanding during difficult periods. Her continued support has meant a great deal to me, and I am truly thankful for the opportunity to have been her student.

I would also like to thank the members of my advisory committee, Dr. Saman Muthukumarana and Dr. Brad Johnson, for their time, helpful comments, and support throughout my program. I am especially grateful to Dr. Muthukumarana, in his role as Head of the Department of Statistics, for his guidance on academic and departmental matters and for his support of graduate students' academic and professional development.

I would also like to thank my external examiner, Dr. Andrei Volodin, for his careful reading of my thesis and for his thoughtful comments and suggestions.

I gratefully acknowledge the financial support I received from my supervisor, the Department of Statistics, the Faculty of Science, and the Faculty of Graduate Studies. This support allowed me to continue my studies and focus on my research.

I would also like to thank the faculty, staff, and fellow graduate students in the Department of Statistics for their support and for making my time in the department meaningful. I am grateful for the kindness, assistance, encouragement, and positive experiences I shared with them throughout my graduate studies.

Finally, I am deeply grateful to my family, Edna, Mico, and Pam, for their love, support, and encouragement throughout this journey. I am thankful for the many ways they have helped me along the way.

Chapter 1

Introduction

1.1 Background and Motivation

The themes of this thesis lie within two closely related areas of statistics and probability theory: invariance principles for Student processes and nonparametric change-point analysis. Although the manuscript chapters address different inferential problems, they are connected by a common methodological idea: using functionals of appropriately normalized and weighted partial-sum-type processes to construct nonparametric asymptotic procedures that aim to improve finite-sample performance relative to more classical choices of statistics.

Classical asymptotic procedures are widely used for confidence intervals and hypothesis testing because they are broadly applicable and often require only weak distributional assumptions. For example, the usual asymptotic confidence interval for the mean of independent and identically distributed (i.i.d.) random variables (r.v.'s) is based on the Central Limit Theorem and its Studentized form. Although this interval is asymptotically valid under broad conditions, its finite-sample coverage probability may be lower than the nominal confidence level, especially for skewed or heavy-tailed distributions.

The Functional Central Limit Theorems (FCLTs) established by Csörgő et al. (2003) provide weak convergence results for the Student process based on i.i.d. r.v.'s in the domain of attraction of the normal law. When suitable functionals are applied to this process, the resulting convergence can be used to construct functional asymptotic confidence intervals (FACIs) for the mean. Earlier works by Tuzov (2014) and Martynyuk and Tuzov (2016) used this framework to derive FACIs based on several choices of functionals, including supremum and integral-type functionals. Their numerical studies showed that some of these intervals can have better finite-sample coverage probabilities than the classical

CLT-based interval, especially in settings where the latter may undercover. However, such improvements may come at the cost of longer average interval lengths, reflecting a trade-off between coverage accuracy and interval efficiency.

Weighted versions of the Student-process FCLT were developed by Csörgő et al. (2008) and provide a further source of flexibility. The use of weighted Student processes for a confidence interval construction was explored by Bhardwaj (2022), but the finite-sample trade-off between coverage improvement and interval length remained a central issue. This motivated the study in Chapter 2 of how the choice of a functional and weight function affects the finite-sample performance of FACIs.

Similar finite-sample concerns arise in change-point analysis. In tests for a change in the mean of chronologically ordered observations, a common statistic is based on the largest absolute difference between the sample mean before and after each possible change point. The power of such a test depends not only on the size of the change, but also on where the change occurs in the sample. Changes near the beginning or end of the observation period are typically harder to detect than changes near the middle.

Tests for at most one change in the mean can be formulated using functionals of normalized tied-down partial-sum processes, whose weak limits under the null hypothesis are functionals of a Brownian bridge. Weighted supremum-type tests have been studied in the change-point literature, for example by Csörgő and Horváth (1988), Csörgő and Horváth (1997), and Orasch and Pouliot (2004). In this setting, the weight function affects the sensitivity of the test to changes occurring at different locations in the sample. However, a three-way trade-off is usually observed among a test's type I errors, power for detecting the change near the beginning or end of the sample, and power for detecting the change in the middle of the sample. It appears that previous works did not specifically aim to obtain tests with better-balanced finite-sample behavior through the choice of a functional, weight function, or both. This motivates the study in Chapter 3.

Chapter 4 extends the same nonparametric process-based perspective to regression change-point testing. In simple linear regression, residual-based tied-down partial-sum processes are constructed for testing changes in the slope and intercept separately. Although related component-specific tests for at most one change in the slope or intercept have been

studied previously, the chapter develops additional asymptotic results that provide a basis for future finite-sample comparisons in the regression setting.

1.2 Main Contributions and Thesis Organization

The overall contribution of this thesis is the development and study of nonparametric asymptotic confidence intervals and change-point detection tests based on functionals of weighted normalized partial-sum-type processes. Chapters 2 and 3 include numerical investigations of finite-sample performance, while Chapter 4 develops asymptotic theory for regression change-point testing.

Chapter 2 develops FACIs for the mean based on weighted Student processes. Starting from existing FCLTs, the chapter considers interval constructions based on an integral functional and a mixed functional involving both projection and integral components. It also introduces special choices of weight functions and derives the corresponding confidence intervals by solving the resulting inequalities in μ . The finite-sample coverage probabilities and expected lengths of these intervals are then compared with those of the classical Studentized-CLT-based interval and related FACIs from earlier work.

The chapter shows that the proposed weighted-process-based intervals can have substantially improved finite-sample coverages for right-skewed distributions in the domain of attraction of the normal law, including cases where the classical CLT-based interval significantly undercovers. It also shows that the improved coverage is not explained only by an increased interval length: for many right-skewed distributions, the best proposed intervals have higher empirical coverage probabilities than equal-expected-length versions of the classical interval. These results identify settings in which FACIs may be preferable to the classical Studentized-CLT-based interval.

Chapter 3 studies nonparametric tests for an at-most-one change in the mean. The chapter investigates test statistics based on supremum and integral functionals of weighted tied-down partial-sum processes whose null limits are Brownian-bridge-based. While weighted supremum tests of this type have been considered previously, weighted integral tests for this mean-change problem appear not to have been systematically studied in the litera-

ture. Through numerical comparisons, the chapter examines the effect of weight functions on type I errors and powers across distributions with different degrees of asymmetry and heavy-tailedness, as well as across different sample sizes, change sizes, and change-point locations. The chapter emphasizes the finite-sample trade-off among type I error control, power for changes near the endpoints of the sample, and power for changes near the middle of the sample.

By choosing suitable weight functions, Chapter 3 proposes weighted supremum and integral tests that are nearly as powerful as the strongest previously studied weighted supremum test, while having type I errors closer to the nominal level. The weighted integral test is a particularly important contribution. It emerges as one of the most powerful and well-balanced tests considered in the study, exhibiting strong power across all change-point locations while maintaining the best controlled type I errors among the weighted supremum and integral tests examined.

Chapter 4 extends the weighted tied-down partial-sum process framework from the mean-change problem to simple linear regression with random design. The problems of testing for at most one change in the slope and for at most one change in the intercept are treated separately. For each problem, residual-based quantities are used to construct appropriate tied-down partial-sum processes. This leads to asymptotic null distributions for several slope-specific and intercept-specific statistics, including supremum and integral statistics with Brownian-bridge-based limits, supremum and integral Darling–Erdős-type statistics, selected joint limits, and Rényi-type supremum statistics with Wiener-process limits.

The main contribution of Chapter 4 is the development of an asymptotic framework for these regression change-point tests. This provides the foundation needed to extend the finite-sample ideas of Chapter 3 to tests for changes in regression coefficients. Since Chapter 3 shows that the choice of a functional and weight function can strongly affect the balance between power and type I error control, the asymptotic results developed in Chapter 4 provide a solid basis for future numerical studies aimed at identifying regression change-point tests with better-balanced finite-sample performance.

Finally, Chapter 5 summarizes the main conclusions of the thesis and discusses possible

directions for future work. These include finite-sample studies for the regression tests, further investigation of weight selection, and extensions to more general regression models.

The three manuscript chapters are connected by a common process-based approach and a common finite-sample motivation. Each chapter uses functionals of weighted and normalized partial-sum-type processes to construct nonparametric asymptotic procedures. Chapters 2 and 3 directly investigate finite-sample performance in the confidence-interval and mean-change settings. Chapter 4 builds on the weighted change-point framework of Chapter 3 by developing corresponding residual-based tied-down partial-sum methods for regression coefficient changes and establishing the asymptotic foundation for future finite-sample studies.

1.3 Contributions of Authors

This sandwich thesis was made possible through the contributions of the authors of the manuscripts included in Chapters 2, 3, and 4. Jervis Gallanosa contributed to the development of the research, implementation of the numerical studies, formal analysis, interpretation of results, and preparation of the manuscripts. He also drafted and integrated the thesis manuscript, including the introductory and concluding chapters, and was responsible for maintaining coherence and consistency across the thesis.

Dr. Yuliya V. Martsynyuk provided supervision, conceptual and methodological guidance, and theoretical direction throughout the research. Her contributions were central to the development of the asymptotic frameworks, the formulation of the research problems, the writing and revising of the manuscripts of Chapters 2 and 3, and the refinement of the manuscript of Chapter 4.

For Chapter 2, conceptualization, methodology, and supervision were provided by Y. V. Martsynyuk. Simulation studies were conducted by S. Bhardwaj and J. Gallanosa. Formal analysis was carried out by S. Bhardwaj, J. Gallanosa, and Y. V. Martsynyuk. Writing was completed by J. Gallanosa and Y. V. Martsynyuk.

In Chapter 3, conceptualization and supervision were provided by Y. V. Martsynyuk. Numerical studies, formal analysis, interpretation of results, and writing were carried out

by J. Gallanosa and Y. V. Martsynyuk.

For Chapter 4, theoretical development, formal analysis, and writing were primarily carried out by J. Gallanosa. Y. V. Martsynyuk provided supervision, methodology, and conceptualization.

Chapter 2

Asymptotic Confidence Intervals for the Mean with Increased Finite-Sample Coverage Probabilities

Asymptotic Confidence Intervals for the Mean with Increased Finite-Sample Coverage Probabilities

Shivani Bhardwaj¹, Jervis Gallanosa², and Yuliya V. Martsynyuk²

¹Canada Revenue Agency,
360 Main St.,
Winnipeg R3C 3Z3.

²Department of Statistics,
University of Manitoba,
Winnipeg MB R3T 2N2

The content of this chapter is based on the following publication:

Bhardwaj, S.; Gallanosa, J.; Martsynyuk, Y. V. (2025). Asymptotic Confidence Intervals for the Mean with Increased Finite-Sample Coverage Probabilities. *Mathematics* **13**, 3931.

<https://doi.org/10.3390/math13243931>

Abstract

We consider a Student process based on independent copies of a random variable X . If X is in the domain of attraction of the normal law (DAN), a weighted version of the Student process is known to follow a functional Central Limit Theorem (FCLT). Accordingly, appropriate functionals of such a process converge in distribution to the same functionals of the similarly weighted standard Wiener process. We use such a convergence for an integral functional and derive asymptotic confidence intervals (CIs) for the mean of X . For right-skewed distributions of X in DAN, we show that the obtained CIs have higher finite-sample coverage probabilities than, and may be preferred over, a CI I_1 of the same asymptotic confidence level $1 - \alpha$ that is based on the CLT for the Student t -statistic, since the finite-sample coverage probabilities of the latter CI may be lower than $1 - \alpha$. Moreover, for such distributions, the finite-sample coverage probabilities of our best two CIs are also higher than those of their respective equal-expected-length I_1 counterparts.

Keywords: weighted Student process; functional Central Limit Theorem; integral functional; asymptotic confidence interval; domain of attraction of the normal law; right-skewed distribution; Berry–Esseen bound

MSC: 60E07; 60F17; 62G15

1 Introduction

Let $\{X, X_i, i \geq 1\}$ be a sequence of independent and identically distributed (i.i.d.) random variables (r.v.) with an unknown population mean $\mathbb{E}X = \mu$ throughout this paper. Consider the Student t -statistic

$$T_n(X_1, \dots, X_n) := \frac{\sum_{i=1}^n X_i}{s_n \sqrt{n}}, \quad (1)$$

where $s_n := \sqrt{\sum_{i=1}^n (X_i - \bar{X}_n)^2 / (n-1)}$ and $\bar{X}_n := \sum_{i=1}^n X_i / n$.

When $0 < \text{Var}X < \infty$, it is well-known that

$$T_n(X_1 - \mu, \dots, X_n - \mu) \xrightarrow{\mathcal{D}} N(0, 1), \quad n \rightarrow \infty, \quad (2)$$

which leads to the classical $100(1 - \alpha)\%$ asymptotic confidence interval (CI) for the mean μ :

$$I_1 := \left[\bar{X}_n - \frac{z_{\alpha/2} s_n}{\sqrt{n}}, \bar{X}_n + \frac{z_{\alpha/2} s_n}{\sqrt{n}} \right], \quad (3)$$

where $\alpha \in (0, 1)$ and $z_{\alpha/2}$ are the $1 - \alpha/2$ quantile of the standard normal distribution $N(0, 1)$. In fact, Logan et al. (1973) extended (2) and hence validated (3) for the distributions in the domain of attraction of the normal law (DAN), where $X \in \text{DAN}$ means that there exist constants a_n and $b_n > 0$ such that

$$\frac{\sum_{i=1}^n X_i - a_n}{b_n} \xrightarrow{\mathcal{D}} N(0, 1), \quad n \rightarrow \infty. \quad (4)$$

Remark 1. *The distributions in DAN have finite moments of order $\nu \in (0, 2)$, while their variances are positive, but may be infinite. Clearly, if $0 < \text{Var}X < \infty$, then (4) is satisfied on account of the central limit theorem (CLT), with $a_n = n\mathbb{E}X$ and $b_n = \sqrt{n\text{Var}X}$. If $\text{Var}X = \infty$, then a_n can also be taken as $n\mathbb{E}X$, while $b_n = \sqrt{n}l_X(n)$, where $l_X(n)$ is a slowly varying function at infinity, meaning that $l_X(az)/l_X(z) \rightarrow 1$, as $z \rightarrow \infty$, for all $a > 0$. Moreover, in this case $l_X(n)$ is also nondecreasing and converges to ∞ , as $n \rightarrow \infty$. Consider a Pareto distribution of the first kind with the tail index α and the scale parameter β , denoted by $\text{Pareto}(\alpha, \beta)$ hereafter. $\text{Pareto}(2, 1)$ with the probability density function $2x^{-3}\mathbb{1}_{\{x \geq 1\}}$ is a well-known example of a distribution in DAN with an infinite variance. Moreover, b_n in (4) can be taken as $\sqrt{n \log n}$ in this case. Clearly, if $\alpha > 2$, then $\text{Pareto}(\alpha, 1)$ has a finite variance and hence is also in DAN.*

A Student process in $D[0, 1]$, the space of real-valued functions on $[0, 1]$ that are right-continuous and have left-hand limits, is defined as follows:

$$T_n^t(X_1, \dots, X_n) := \frac{\sum_{i=1}^{\lfloor nt \rfloor} X_i}{s_n \sqrt{n}}, \quad 0 \leq t \leq 1, \quad (5)$$

where $\sum_{i=1}^0 X_i := 0$. Thus, $T_n^t(X_1, \dots, X_n)$ is a random step function on $[0, 1]$:

$$T_n^t(X_1, \dots, X_n) = \begin{cases} 0, & \text{if } 0 \leq t < \frac{1}{n}, \\ \frac{X_1}{s_n \sqrt{n}}, & \text{if } \frac{1}{n} \leq t < \frac{2}{n}, \\ \vdots & \vdots \\ \frac{X_1 + X_2 + \dots + X_{n-1}}{s_n \sqrt{n}}, & \text{if } \frac{n-1}{n} \leq t < 1, \\ \frac{X_1 + X_2 + \dots + X_n}{s_n \sqrt{n}}, & \text{if } t = 1. \end{cases} \quad (6)$$

When $t = 1$, the Student process in (6) becomes the Student t -statistic of (1).

Let $\{W(t), 0 \leq t \leq 1\}$ denote a standard Wiener process. Theorem 1 of Csörgő et al. (2003) extended (2) for $X \in \text{DAN}$ as follows:

$$X \in \text{DAN} \quad \text{and} \quad \mathbb{E}X = \mu \quad \implies \quad h(T_n^t(X_1 - \mu, \dots, X_n - \mu)) \xrightarrow{\mathcal{D}} h(W(t)), \quad n \rightarrow \infty, \quad (7)$$

for all functionals $h: D[0, 1] \rightarrow \mathbb{R}$ that are $\mathcal{D}$ -measurable and ρ -continuous,

where $\mathcal{D}$ is the σ -field of subsets of $D[0, 1]$ that is generated by the finite-dimensional subsets of $D[0, 1]$ and ρ is the sup-norm metric in $D[0, 1]$. The convergence in distribution in (7) is known as a functional Central Limit Theorem (FCLT). We also note that the functionals to be considered in this paper are $\mathcal{D}$ -measurable and ρ -continuous.

In order to see that the CLT of (2) for the Student t -statistic is a special case of the FCLT of (7) for the Student process, one reads (7) with the projection functional $h_{1,t_0}(\cdot)$ with $t_0 = 1$, where

$$h_{1,t_0}(f(t)) = f(t_0), \quad \text{for any } f(t) \in D[0, 1]. \quad (8)$$

Tuzov (2014) and Martsynyuk and Tuzov (2016) considered several concrete functionals as in (7) and derived respective asymptotic CIs, abbreviated as FACIs, for the mean of X based on convergence in distribution of these functionals. The focus of these works was to explore finite-sample properties of the obtained asymptotic CIs and compare them to those of I_1 of (3). Accordingly, from the comparison of the expected lengths and finite-

sample coverage probabilities of these CIs, which was mostly performed numerically, it was concluded that some of the FACIs were shorter than I_1 , had at least as high coverage probabilities as those of I_1 , or both, and thus could potentially serve as alternatives to I_1 . Situations in which these FACIs could be preferred over I_1 were not discussed in these works.

Now, we present an FCLT for a weighted version of the Student process of (5) needed for an introduction of the current work.

Let Q be the class of positive functions $q(t): (0, 1) \rightarrow (0, \infty)$, i.e., $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$, which are non-decreasing near 0 and non-increasing near 1. For $q(t) \in Q$, let

$$\mathcal{I}(q, c) := \int_0^1 \frac{e^{-cq^2(t)/(t(1-t))}}{t(1-t)} dt, \quad c > 0. \quad (9)$$

Assuming that

$$\mathcal{I}(q, c) < \infty \quad \text{for all } c > 0, \quad (10)$$

it follows from Corollary 5 of Csörgő et al. (2008) that (7) for the weighted Student process $T_n^t(X_1 - \mu, \dots, X_n - \mu)/q([nt]/n)$ holds true:

$$X \in \text{DAN} \quad \text{and} \quad \mathbb{E}X = \mu \quad \implies \quad h \left(\frac{T_n^t(X_1 - \mu, \dots, X_n - \mu)}{q \left(\frac{[nt]}{n} \right)} \right) \xrightarrow{\mathcal{D}} h \left(\frac{W(t)}{q(t)} \right), \quad n \rightarrow \infty, \quad (11)$$

for all functionals $h: D[0, 1] \rightarrow \mathbb{R}$ that are $\mathcal{D}$ -measurable and ρ -continuous.

Remark 2. For $q(t) \in Q$,

$$(10) \text{ holds true if and only if } \lim_{t \downarrow 0} \frac{|W(t)|}{q(t)} = \lim_{t \uparrow 1} \frac{|W(t)|}{q(t)} = 0 \text{ a.s.} \quad (12)$$

(see Theorem 3.4 of Csörgő et al. (1986)). Using (12) and Khinchine's local law of the iterated logarithm for the standard Wiener process, one can obtain examples of weight functions in Q satisfying (10), also known as Chibisov–O'Reilly functions, by simply multiplying the function

$$q_1(t) := \sqrt{t(1-t) \log \log(t(1-t))^{-1}}, \quad t \in (0, 1), \quad (13)$$

by an appropriate function on $(0, 1)$ that converges to infinity as $t \downarrow 0$ and $t \uparrow 1$.

Among other functionals, the convergence in distribution of (11) can be read with the integral functional $h(\cdot) = h_2(\cdot)$, where

$$h_2(f(t)) = \int_0^1 f^2(t) dt, \quad \text{for any } f(t) \in D[0, 1]. \quad (14)$$

Moreover, such a convergence can also be concluded by using an L_2 -approximation in probability (see Theorem 3 of Csörgő et al. (2008)), bypassing the FCLT and using a subclass of weight functions $q(t)$ from Q that satisfy the condition

$$\int_0^1 \frac{t(1-t)}{q^2(t)} dt < \infty. \quad (15)$$

Thus, in particular, (11) with $h_2(\cdot)$ holds true with the weight function

$$q_{2,a}(t) := (t(1-t))^a, \quad \text{where } 0 \leq a < 0.5. \quad (16)$$

The FCLT in (11) also holds true with a linear combination of the projection functional h_{1,t_0} of (8) and the integral functional h_2 of (14), in particular, with the following functional:

$$h_3(f(t)) = d_1 h_{1,0.99}(f(t)) + d_2 h_2(f(t)), \quad \text{for any } f(t) \in D[0, 1], \quad (17)$$

where $d_1, d_2 \in \mathbb{R}$.

In the present work, we derive FACIs for the mean μ by using convergence in distribution of the integral functional h_2 of (14) and its modification h_3 of (17) applied to the Student process $T_n^t(X_1 - \mu, \dots, X_n - \mu)/q([nt]/n)$ weighted with some appropriate functions $q(t)$.

We study, numerically, the finite-sample performance of the obtained FACIs and compare it to that of the CLT-based CI I_1 of (3). For asymmetric distributions in DAN, the finite-sample coverage probabilities of I_1 may be lower than its asymptotic confidence level $1 - \alpha$. Moreover, I_1 may suffer from significant undercoverage for smaller sample sizes and more skewed distributions, especially the ones with fewer moments, while the rates of convergence of the coverage probabilities of I_1 to $1 - \alpha$ may be slow. We show that our

$1 - \alpha$ FACIs have higher finite-sample coverage probabilities than, and may be reasonable alternatives to, the CI I_1 for right-skewed distributions in DAN, although the FACIs are longer than I_1 on average.

2 Main Results

2.1 Preliminaries

In order to derive our FACIs for μ from (11), the distribution function of the limiting r.v. $h(W(t)/q(t))$ must either have a closed form expression or be tabulated. This is only the case for $h_2(W(t))$. The rest of the limiting distributions have to be tabulated by us, similarly to the way of Tuzov (2014) and Martsynyuk and Tuzov (2016) for their functionals, which is based on the invariance principle approach of Erdős and Kac (1946). Accordingly, since the limiting distribution in (11) does not depend on the underlying distribution structure of X , we estimate its quantiles by those of the respective empirical distribution of the r.v. $h(T_n^t(X_1 - \mu, X_2 - \mu, \dots, X_n - \mu)/q([nt]/n))$, where the latter distribution is based on 100,000 simulated values that are each computed from an $N(0, 1)$ random sample of size 100,000.

After deriving our FACIs, we conduct a detailed numerical study of their finite-sample coverage probabilities and expected lengths in comparison to those of the CLT-based CI I_1 of (3). The exact finite-sample coverage probabilities of the FACIs and I_1 are approximated by their empirical counterparts, which are based on 10,000 random samples of size n from a distribution of X .

In order to compare the expected length of a FACI (which is also not feasible to obtain in general in a closed form expression) to the expected length of I_1 , we use the ratios of the empirical expected lengths of a FACI to that of I_1 from 10,000 random samples of size n .

In our simulation study, we use seven right-skewed and two symmetric distributions from DAN with different numbers of moments: *Pareto*(2, 1), *Weibull*(0.4, 1), *InverseNormal*(7, 1), *Log-logistic*(2.1, 1), *Log-normal*(0, 1.3), *Frechet*(2.5), *Half-t*(3), *SymPareto*(2, 1), and *t*(3). The probability density functions for some of these distributions are provided in Table 1.

Table 1: Probability density functions for some distributions.

Distribution	Probability Density Function
<i>Weibull</i> (0.4, 1)	$0.4 x^{-0.6} e^{-x^{0.4}}, \quad x > 0$
<i>InverseNormal</i> (7, 1)	$(2\pi x^3)^{-0.5} e^{-(x-7)^2/(98x)}, \quad x > 0$
<i>Log-logistic</i> (2.1, 1)	$2.1 x^{1.1} (1 + x^{2.1})^{-2}, \quad x > 0$
<i>Log-normal</i> (0, 1.3)	$(1.3x\sqrt{2\pi})^{-1} e^{-\ln^2(x)/(2(1.3)^2)}, \quad x > 0$
<i>Frechet</i> (2.5)	$2.5 x^{-3.5} e^{-x^{-2.5}}, \quad x > 0$
<i>Half-t</i> (3)	$\frac{4}{\sqrt{3}\pi} (1 + x^2/3)^{-2}, \quad x > 0$
<i>SymPareto</i> (2, 1)	$ x ^{-3}, \quad x \geq 1$

2.2 FACIs Based on h_2 and h_3

Substituting $h(\cdot) = h_2(\cdot)$ of (14) into (11), we have the following, as $n \rightarrow \infty$:

$$\int_0^1 \frac{(T_n^t(X_1 - \mu, \dots, X_n - \mu))^2}{q^2\left(\frac{[nt]}{n}\right)} dt \xrightarrow{\mathcal{D}} \int_0^1 \frac{W^2(t)}{q^2(t)} dt. \quad (18)$$

Since $T_n^t(X_1, \dots, X_n)$ is a random step function on $[0, 1]$ as in (6), the convergence in (18) amounts to

$$\frac{1}{s_n^2 n^2} \sum_{k=1}^{n-1} \frac{\left(\sum_{i=1}^k (X_i - \mu)\right)^2}{q^2\left(\frac{k}{n}\right)} \xrightarrow{\mathcal{D}} \int_0^1 \frac{W^2(t)}{q^2(t)} dt. \quad (19)$$

We consider (19) with two weight functions $q(t) = q_{2,a}(t) = (t(1-t))^a$ from Q with $a = 0$ and 0.4 , where both $q_{2,0}(t)$ and $q_{2,0.4}(t)$ are Chibisov–O’Reilly functions satisfying (10) as well as (15). The cumulative distribution function of $\int_0^1 (W^2(t)/q^2(t)) dt$ has only been available for $q(t) \equiv 1$ (see Csörgő and Horváth (1981)). We retabulated the quantiles of $\int_0^1 W^2(t) dt$ and they closely match the ones in the latter paper. We also obtained the required quantiles of $\int_0^1 (W^2(t)/q_{2,0.4}^2(t)) dt$. All these quantiles are listed in Table 2.

 Table 2: The $1 - \alpha$ quantile of $\int_0^1 \frac{W^2(t)}{(t(1-t))^{2a}} dt$.

	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.02$
$a = 0$	1.198	1.668	2.292
$a = 0.4$	10.734	14.888	20.690

Let $\alpha \in (0, 1)$ and define $b > 0$ such that $P\left(\int_0^1 \frac{W^2(t)}{q^2(t)} dt \leq b\right) = 1 - \alpha$. It follows from (19) that

$$P\left(\frac{1}{s_n^2 n^2} \sum_{k=1}^{n-1} \frac{\left(\sum_{i=1}^k (X_i - \mu)\right)^2}{q^2\left(\frac{k}{n}\right)} \leq b\right) \xrightarrow{\mathcal{D}} 1 - \alpha, \quad n \rightarrow \infty. \quad (20)$$

The quadratic inequality in μ under the probability sign in (20) reads

$$\mu^2 \sum_{k=1}^{n-1} \frac{k^2}{q^2\left(\frac{k}{n}\right)} - 2\mu \sum_{k=1}^{n-1} \frac{k \sum_{i=1}^k X_i}{q^2\left(\frac{k}{n}\right)} + \sum_{k=1}^{n-1} \frac{\left(\sum_{i=1}^k X_i\right)^2}{q^2\left(\frac{k}{n}\right)} - bn^2 s_n^2 \leq 0. \quad (21)$$

This leads to the following $1 - \alpha$ FACI for μ :

$$I_{2,q} := \left[\frac{2 \sum_{k=1}^{n-1} \frac{k \sum_{i=1}^k X_i}{q^2\left(\frac{k}{n}\right)} - \sqrt{D_n}}{2 \sum_{k=1}^{n-1} \frac{k^2}{q^2\left(\frac{k}{n}\right)}}, \frac{2 \sum_{k=1}^{n-1} \frac{k \sum_{i=1}^k X_i}{q^2\left(\frac{k}{n}\right)} + \sqrt{D_n}}{2 \sum_{k=1}^{n-1} \frac{k^2}{q^2\left(\frac{k}{n}\right)}} \right], \quad (22)$$

where D_n is the discriminant of the quadratic function of (21):

$$D_n := 4 \left(\sum_{k=1}^{n-1} \frac{k \sum_{i=1}^k X_i}{q^2\left(\frac{k}{n}\right)} \right)^2 - 4 \sum_{k=1}^{n-1} \frac{k^2}{q^2\left(\frac{k}{n}\right)} \left(\sum_{k=1}^{n-1} \frac{\left(\sum_{i=1}^k X_i\right)^2}{q^2\left(\frac{k}{n}\right)} - bn^2 s_n^2 \right), \quad (23)$$

where $D_n > 0$ in all the Monte Carlo simulations of this paper. The FACI $I_{2,q}$ with $q(t) \equiv 1$ reduces to one of the FACIs obtained by Tuzov (2014) and Martsynyuk and Tuzov (2016) from (7).

Next, we similarly derive the FACIs based on the convergence in distribution in (11) with the Chibisov–O’Reilly weight function

$$q_3(t) = [t(1-t)]^{0.25} [\log \log(t(1-t))^{-1}]^{-0.5} \quad (24)$$

satisfying (10) and with $h(\cdot) = h_3(\cdot)$ of (17), with two pairs of values for (d_1, d_2) ,

$$(0.3 \cdot q_3(0.99), 0.7) = (0.07671255, 0.7) \quad \text{and} \quad (0.5 \cdot q_3(0.99), 0.5) = (0.1278543, 0.5). \quad (25)$$

Accordingly, for $\alpha \in (0, 1)$ and e defined as $P(h_3(W(t)/q_3(t)) \leq e) = 1 - \alpha$, we have

$$P \left(d_1 \frac{\sum_{i=1}^{\lfloor 0.99n \rfloor} (X_i - \mu)}{s_n \sqrt{n} q_3 \left(\frac{\lfloor 0.99n \rfloor}{n} \right)} + \frac{d_2}{s_n^2 n^2} \sum_{k=1}^{n-1} \frac{\left(\sum_{i=1}^k (X_i - \mu) \right)^2}{q_3^2 \left(\frac{k}{n} \right)} \leq e \right) \xrightarrow{\mathcal{D}} 1 - \alpha, \quad n \rightarrow \infty. \quad (26)$$

Solving the quadratic inequality in μ under the probability sign in (26) yields the following

$1 - \alpha$ FACI for μ :

$$I_{3,q} := \left[\frac{\frac{d_1 s_n n \sqrt{n} \lfloor 0.99n \rfloor}{q_3 \left(\frac{\lfloor 0.99n \rfloor}{n} \right)} + 2d_2 \sum_{k=1}^{n-1} \frac{k \sum_{i=1}^k X_i}{q_3^2 \left(\frac{k}{n} \right)} - \sqrt{\Delta_n}}{2d_2 \sum_{k=1}^{n-1} \frac{k^2}{q_3^2 \left(\frac{k}{n} \right)}}, \frac{\frac{d_1 s_n n \sqrt{n} \lfloor 0.99n \rfloor}{q_3 \left(\frac{\lfloor 0.99n \rfloor}{n} \right)} + 2d_2 \sum_{k=1}^{n-1} \frac{k \sum_{i=1}^k X_i}{q_3^2 \left(\frac{k}{n} \right)} + \sqrt{\Delta_n}}{2d_2 \sum_{k=1}^{n-1} \frac{k^2}{q_3^2 \left(\frac{k}{n} \right)}} \right], \quad (27)$$

where

$$\begin{aligned} \Delta_n := & \left(\frac{d_1 s_n n \sqrt{n} \lfloor 0.99n \rfloor}{q_3 \left(\frac{\lfloor 0.99n \rfloor}{n} \right)} + 2d_2 \sum_{k=1}^{n-1} \frac{k \sum_{i=1}^k X_i}{q_3^2 \left(\frac{k}{n} \right)} \right)^2 \\ & - 4d_2 \sum_{k=1}^{n-1} \frac{k^2}{q_3^2 \left(\frac{k}{n} \right)} \left(\frac{d_1 s_n n \sqrt{n} \sum_{i=1}^{\lfloor 0.99n \rfloor} X_i}{q_3 \left(\frac{\lfloor 0.99n \rfloor}{n} \right)} + d_2 \sum_{k=1}^{n-1} \frac{\left(\sum_{i=1}^k X_i \right)^2}{q_3^2 \left(\frac{k}{n} \right)} - e s_n^2 n^2 \right), \quad (28) \end{aligned}$$

where, for the discriminant Δ_n , we have $\Delta_n > 0$ in all the Monte Carlo simulations in the upcoming tables, for both pairs of values for (d_1, d_2) of (25). Our simulated values of $1 - \alpha$ quantiles e of $h_3(W(t)/q_3(t))$ can be found in Table 3.

Table 3: The $1 - \alpha$ quantile of $h_3 \left(\frac{W(t)}{q_3(t)} \right) = d_1 \frac{W(0.99)}{q_3(0.99)} + d_2 \int_0^1 \frac{W^2(t)}{q_3^2(t)} dt$.

	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.02$
$(d_1, d_2) = (0.07671255, 0.7)$	2.414540	3.360469	4.661709
$(d_1, d_2) = (0.12785430, 0.5)$	1.875837	2.652915	3.697833

Hereafter, we denote the FACIs $I_{3,q}$ of (27) with $(d_1, d_2) = (0.07671255, 0.7)$ and $(0.1278543, 0.5)$ by I_3 and I_4 , respectively.

In Table 4, we numerically compare the finite-sample properties of I_1 , $I_{2,q}$ with the weight functions $q(t) \equiv 1$ and $q(t) = (t(1-t))^{0.4}$, I_3 , and I_4 at the $1 - \alpha = 0.95$ asymptotic confidence level. Table 4 lists the values of the empirical coverage probabilities $\widehat{CP}_i$ of I_i ,

2. Main Results

$i = 1, 3$, and 4, and $\widehat{CP}_{2,q}$ of $I_{2,q}$, together with the corresponding Monte Carlo standard errors in parentheses, all multiplied by 100%, as well as the values of $\widehat{r}_{2,q}$ and $\widehat{r}_i$, the ratios of the empirical expected lengths of $I_{2,q}$ and I_1 , and I_i and I_1 , $i = 3$ and 4. The Monte Carlo standard errors for $\widehat{r}_{2,q}$ (with both $q(t) \equiv 1$ and $q(t) = (t(1-t))^{0.4}$), $\widehat{r}_3$, and $\widehat{r}_4$ are in the range of $[0.000154, 0.000504]$ across all nine distributions from DAN and the sample sizes $n = 50, 100, 300, 500$, and 1000.

Table 4: Comparison of I_1 , $I_{2,q}$ with $q(t) = (t(1-t))^a$, I_3 , and I_4 at 95% confidence level.

Distribution	n	FACI Coverages					Ratios of Average Lengths			
		$\widehat{CP}_1 \cdot 100\%$	$\widehat{CP}_{2,q} \cdot 100\% \ (a = 0)$	$\widehat{CP}_{2,q} \cdot 100\% \ (a = 0.4)$	$\widehat{CP}_3 \cdot 100\%$	$\widehat{CP}_4 \cdot 100\%$	$\widehat{r}_{2,q} \ (a = 0)$	$\widehat{r}_{2,q} \ (a = 0.4)$	$\widehat{r}_3$	$\widehat{r}_4$
Pareto(2, 1)	50	81.41 (0.39)	83.40 (0.37)	88.49 (0.32)	88.65 (0.32)	92.05 (0.27)	1.124	1.354	1.253	1.339
	100	83.74 (0.37)	85.66 (0.35)	89.51 (0.31)	89.94 (0.30)	93.16 (0.25)	1.115	1.274	1.183	1.272
	300	87.05 (0.34)	88.75 (0.32)	91.10 (0.28)	91.83 (0.27)	94.64 (0.23)	1.109	1.187	1.115	1.196
	500	88.06 (0.32)	89.18 (0.31)	91.14 (0.28)	91.81 (0.27)	94.77 (0.22)	1.108	1.157	1.094	1.174
	1000	89.82 (0.30)	90.81 (0.29)	92.12 (0.27)	92.74 (0.26)	95.34 (0.21)	1.107	1.124	1.074	1.151
Weibull(0.4, 1)	50	81.51 (0.39)	83.04 (0.38)	87.92 (0.33)	87.99 (0.33)	90.89 (0.29)	1.124	1.354	1.252	1.338
	100	86.04 (0.35)	87.77 (0.33)	90.98 (0.29)	91.27 (0.28)	93.83 (0.24)	1.115	1.274	1.183	1.273
	300	90.19 (0.30)	90.90 (0.29)	93.21 (0.25)	93.37 (0.25)	95.48 (0.21)	1.109	1.187	1.115	1.196
	500	91.17 (0.28)	92.03 (0.27)	93.84 (0.24)	93.82 (0.24)	95.90 (0.20)	1.108	1.157	1.095	1.174
	1000	92.96 (0.26)	93.31 (0.25)	94.90 (0.22)	94.70 (0.22)	95.92 (0.20)	1.107	1.124	1.074	1.151
InverseNormal(7, 1)	50	82.52 (0.38)	83.97 (0.37)	88.21 (0.32)	88.38 (0.32)	90.84 (0.29)	1.124	1.354	1.253	1.338
	100	87.13 (0.33)	88.26 (0.32)	91.52 (0.28)	91.88 (0.27)	94.20 (0.23)	1.115	1.274	1.183	1.271
	300	91.80 (0.27)	92.10 (0.27)	94.58 (0.23)	94.62 (0.23)	96.02 (0.20)	1.109	1.187	1.114	1.195
	500	92.93 (0.26)	93.52 (0.25)	95.19 (0.21)	95.04 (0.22)	96.25 (0.19)	1.108	1.157	1.094	1.174
	1000	93.78 (0.24)	94.12 (0.24)	95.60 (0.21)	95.21 (0.21)	96.15 (0.19)	1.107	1.124	1.074	1.151
Log-logistic(2.1, 1)	50	84.44 (0.36)	86.48 (0.34)	91.34 (0.28)	91.51 (0.28)	94.40 (0.23)	1.124	1.354	1.253	1.339
	100	87.46 (0.33)	88.93 (0.31)	92.76 (0.26)	93.11 (0.25)	95.79 (0.20)	1.115	1.274	1.183	1.271
	300	89.87 (0.30)	90.96 (0.29)	93.25 (0.25)	93.67 (0.24)	96.11 (0.19)	1.109	1.187	1.115	1.196
	500	90.51 (0.29)	91.44 (0.28)	93.56 (0.25)	94.01 (0.24)	96.03 (0.20)	1.108	1.157	1.094	1.173
	1000	91.07 (0.29)	92.13 (0.27)	93.50 (0.25)	93.63 (0.24)	95.73 (0.20)	1.107	1.124	1.074	1.151
Log-normal(0, 1.3)	50	85.31 (0.35)	86.86 (0.34)	91.08 (0.29)	91.30 (0.28)	93.71 (0.24)	1.124	1.354	1.253	1.339
	100	87.54 (0.33)	88.64 (0.32)	92.31 (0.27)	92.41 (0.26)	94.78 (0.22)	1.115	1.274	1.183	1.272
	300	91.70 (0.28)	92.39 (0.27)	94.64 (0.23)	94.70 (0.22)	96.19 (0.19)	1.109	1.187	1.115	1.196
	500	92.17 (0.27)	92.90 (0.26)	94.89 (0.22)	94.73 (0.22)	96.24 (0.19)	1.108	1.157	1.094	1.173
	1000	93.25 (0.25)	93.82 (0.24)	95.35 (0.21)	95.17 (0.21)	96.43 (0.19)	1.107	1.124	1.074	1.151
Frechet(2.5)	50	87.93 (0.33)	89.24 (0.31)	93.45 (0.25)	93.52 (0.25)	95.64 (0.20)	1.124	1.354	1.253	1.339
	100	89.61 (0.31)	90.57 (0.29)	94.10 (0.24)	94.27 (0.23)	96.36 (0.19)	1.115	1.274	1.183	1.272
	300	91.90 (0.27)	92.13 (0.27)	94.75 (0.22)	94.82 (0.22)	96.45 (0.19)	1.109	1.187	1.114	1.196
	500	92.22 (0.27)	92.91 (0.26)	94.85 (0.22)	94.94 (0.22)	96.76 (0.18)	1.108	1.157	1.094	1.173
	1000	93.54 (0.25)	93.85 (0.24)	95.41 (0.21)	95.36 (0.21)	96.80 (0.18)	1.107	1.124	1.074	1.151
Half-t(3)	50	90.98 (0.29)	91.77 (0.27)	95.75 (0.20)	95.62 (0.20)	97.30 (0.16)	1.124	1.354	1.253	1.339
	100	91.99 (0.27)	92.86 (0.26)	96.11 (0.19)	96.03 (0.20)	97.57 (0.15)	1.115	1.274	1.183	1.272
	300	93.66 (0.24)	93.92 (0.24)	96.33 (0.19)	96.09 (0.19)	97.06 (0.17)	1.109	1.187	1.115	1.196
	500	93.89 (0.24)	94.13 (0.24)	96.08 (0.19)	95.77 (0.20)	96.78 (0.18)	1.108	1.157	1.095	1.174
	1000	94.11 (0.24)	94.51 (0.23)	96.06 (0.19)	95.49 (0.21)	96.32 (0.19)	1.107	1.124	1.074	1.151
SymPareto(2, 1)	50	95.12 (0.22)	95.45 (0.21)	99.10 (0.09)	98.17 (0.13)	97.86 (0.14)	1.124	1.354	1.253	1.339
	100	95.51 (0.21)	95.34 (0.21)	98.56 (0.12)	97.41 (0.16)	97.10 (0.17)	1.115	1.274	1.183	1.271
	300	95.33 (0.21)	95.26 (0.21)	97.96 (0.14)	96.69 (0.18)	96.33 (0.19)	1.109	1.187	1.115	1.196
	500	95.62 (0.20)	95.67 (0.20)	97.89 (0.14)	96.62 (0.18)	96.37 (0.19)	1.108	1.157	1.094	1.173
	1000	95.35 (0.21)	95.20 (0.21)	97.12 (0.17)	96.04 (0.20)	95.96 (0.20)	1.106	1.123	1.074	1.151
t(3)	50	94.80 (0.22)	95.29 (0.21)	98.99 (0.10)	98.09 (0.14)	97.84 (0.15)	1.123	1.354	1.253	1.339
	100	95.02 (0.22)	95.02 (0.22)	98.40 (0.13)	97.37 (0.16)	97.03 (0.17)	1.115	1.274	1.183	1.272
	300	95.11 (0.22)	95.22 (0.21)	97.77 (0.15)	96.49 (0.18)	96.14 (0.19)	1.109	1.187	1.115	1.196
	500	94.85 (0.22)	95.17 (0.21)	97.25 (0.16)	95.98 (0.20)	95.96 (0.20)	1.108	1.157	1.094	1.174
	1000	94.87 (0.22)	94.86 (0.22)	96.60 (0.18)	95.36 (0.21)	95.40 (0.21)	1.107	1.124	1.074	1.151

Based on Table 4, we make the following observations in Remarks 3–6.

Remark 3. *All four FACIs exhibit trade-offs between their respective empirical coverage probabilities and expected lengths. Indeed, they have mostly higher coverage probabilities than those of I_1 , while being longer than I_1 on average. More precisely, the ratios $\hat{r}_{2,q}$, $\hat{r}_3$, and $\hat{r}_4$ are greater than one, being at most 1.354 for $n = 50$ and at most 1.151 for $n = 1000$, and it appears these ratios do not depend on the distribution and decrease as n increases. As for the empirical finite-sample coverage probabilities, $\widehat{CP}_{2,q}$, $\widehat{CP}_3$, and $\widehat{CP}_4$ are higher than $\widehat{CP}_1$, except for a small number of larger samples for $\widehat{CP}_{2,q}$ when $q(t) \equiv 1$. For all the asymmetric distributions, $\widehat{CP}_1 \cdot 100\%$ is lower than $(1 - \alpha)100\% = 95\%$ (more so for smaller n) and converges to 95% slowly. For the Pareto(2, 1) samples, I_1 has the most severe undercoverage, where the rate of convergence of $\widehat{CP}_1 \cdot 100\%$ is so slow that for $n = 1000$, $\widehat{CP}_1 \cdot 100\%$ is still below 95%, at 89.82%. Moreover, as seen from Table 4 (for the 95% confidence level) and Tables A1 and A2 from Appendix to Chapter 2, with similar comparisons of I_1 , $I_{2,q}$, I_3 , and I_4 at the 90% and 98% confidence levels, the four FACIs provide coverage improvements of various degree, with the most significant one being the case of I_4 , and thus may be reasonable alternatives to I_1 for right-skewed distributions. On the other hand, for the symmetric SymPareto(2, 1) and $t(3)$ distributions, the coverage $\widehat{CP}_1 \cdot 100\%$ is very close to the asymptotic confidence level for all our sample sizes and does not require an improvement. We note in passing that the ratios of the empirical expected lengths $\hat{r}_{2,q}$, $\hat{r}_3$, and $\hat{r}_4$ obtained in Tables A1 and A2 are very similar to the ones in Table 4, and that the corresponding Monte Carlo standard errors are in the range of [0.000219, 0.000715] in Table A1 and [0.000109, 0.000359] in Table A2.*

Remark 4. *Various Berry–Esseen upper bounds for*

$$E_n := \sup_x \left| P \left(\sum_{i=1}^n X_i \left(\sum_{i=1}^n X_i^2 \right)^{-1/2} \leq x \right) - P(N(0, 1) \leq x) \right| \quad (29)$$

were established by a number of authors under different conditions on the sequence $\{X_i, i \geq 1\}$ (see Section 3 of Shao and Wang (2013)). For example, assuming that $X_1, \dots, X_n$ are independent but not necessarily identically distributed random variables with zero means and

finite variances, Shao (2005) proved that

$$E_n \leq \delta_n, \quad (30)$$

where

$$\delta_n := \frac{10.2}{B_n^2} \sum_{i=1}^n E(X_i^2 \mathbb{1}_{\{|X_i| \geq B_n/2\}}) + \frac{25}{B_n^3} \sum_{i=1}^n E(|X_i|^3 \mathbb{1}_{\{|X_i| \leq B_n/2\}}), \quad \text{with } B_n^2 := \sum_{i=1}^n EX_i^2. \quad (31)$$

When $\{X_i, i \geq 1\}$ are i.i.d. random variables, the expression δ_n of (31) converges to zero and becomes a uniform upper bound of the rate of convergence in the self-normalized CLT.

Assuming that $\{X, X_i, i \geq 1\}$ are i.i.d. random variables with $X \in \text{DAN}$ and $EX = 0$, and defining

$$\tilde{b}_n := \sup\{x : nx^{-2}E(X^2 \mathbb{1}_{\{|X| \leq x\}}) \geq 1\}, \quad (32)$$

and

$$\tilde{\delta}_n := nP(|X| > \tilde{b}_n) + n\tilde{b}_n^{-1} \left| E(X \mathbb{1}_{\{|X| \leq \tilde{b}_n\}}) \right| + n\tilde{b}_n^{-3} E(|X|^3 \mathbb{1}_{\{|X| \leq \tilde{b}_n\}}), \quad (33)$$

Bentkus and Götze (1996) showed that

$$E_n \leq A\tilde{\delta}_n, \quad (34)$$

where A is an absolute constant and $\tilde{\delta}_n \rightarrow 0$ as $n \rightarrow \infty$.

The bounds δ_n and $A\tilde{\delta}_n$ can also be used to estimate the rate of convergence in the Studentized CLT (2) and how fast the finite-sample coverage probability $\widehat{CP}_1$ of the corresponding CI I_1 of (3) converges to the asymptotic level $1 - \alpha$.

The rate of convergence in the self-normalized and Studentized CLTs, and hence that of $\widehat{CP}_1$ to $1 - \alpha$, is known to be, in general, faster for symmetric distributions and slower for asymmetric distributions, especially when the latter have fewer moments, larger skewness, or both (see, for example, Bentkus and Götze (1996) and references therein). This is clearly observed for $\widehat{CP}_1$ for the symmetric $SymPareto(2, 1)$ and $t(3)$ distributions from Tables 4, A1, and A2. On the other hand, the asymmetric version of $SymPareto(2, 1)$, namely

the ordinary $Pareto(2, 1)$ distribution with an infinite variance, has the lowest $\widehat{CP}_1$ in these tables and converges slowly. Similarly, for the asymmetric $Half-t(3)$, its $\widehat{CP}_1$ is lower compared to that of $t(3)$. For the asymmetric $Log-logistic(2.1, 1)$, $Frechet(2.5)$, and $Half-t(3)$ distributions, we can see from Tables 4, A1, and A2 that their respective $\widehat{CP}_1$ values decrease as their numbers of finite moments decrease.

Interestingly, the behavior of the finite-sample coverage probabilities of our FACIs $I_{2,q}$ closely resembles that of $\widehat{CP}_1$, as evident from Tables 4, A1, and A2. Moreover, in Table 5, not only $\widehat{CP}_1$ but also the four FACI coverages exhibit nearly the same ordering as the values of δ_n and $\tilde{\delta}_n$ computed for our distributions for $n = 100$. The orderings of the values of δ_n , $\tilde{\delta}_n$, $\widehat{CP}_1$, $\widehat{CP}_{2,q}$, $\widehat{CP}_3$, and $\widehat{CP}_4$ vary only slightly for other sample sizes, while remaining in close agreement. Thus, for right-skewed distributions with larger values of δ_n (if defined) or $\tilde{\delta}_n$, one may prefer to rely on one of the FACIs to estimate the mean μ , as I_1 may exhibit undesirably low coverage probabilities.

Table 5: δ_{100} , $\tilde{\delta}_{100}$, and coverages $\widehat{CP}_1$, $\widehat{CP}_{2,q}$, $\widehat{CP}_3$, $\widehat{CP}_4$ ($\cdot 100\%$) for $n = 100$

Distribution	δ_{100}	$\tilde{\delta}_{100}$	$\widehat{CP}_1 \cdot 100\%$	$\widehat{CP}_{2,q} \cdot 100\%$		$\widehat{CP}_3 \cdot 100\%$	$\widehat{CP}_4 \cdot 100\%$
				$q(t) \equiv 1$	$q(t) = (t(1-t))^{0.4}$		
<i>Pareto</i> (2, 1)	NA	1.590	83.74	85.66	89.51	89.94	93.16
<i>Weibull</i> (0.4, 1)	8.700	1.429	86.04	87.77	90.98	91.27	93.83
<i>InverseNormal</i> (7, 1)	8.574	1.247	87.13	88.26	91.52	91.88	94.20
<i>Log-logistic</i> (2.1, 1)	9.365	1.232	87.46	88.93	92.76	93.11	95.79
<i>Log-normal</i> (0, 1.3)	8.310	1.138	87.54	88.64	92.31	92.41	94.78
<i>Frechet</i> (2.5)	8.215	0.974	89.61	90.57	94.10	94.27	96.36
<i>Half-t</i> (3)	7.922	0.603	91.99	92.86	96.11	96.03	97.57
<i>SymPareto</i> (2, 1)	NA	0.451	95.51	95.34	98.56	97.41	97.10
<i>t</i> (3)	6.186	0.328	95.02	95.02	98.40	97.37	97.03

Remark 5. Upon further inspection of Tables 4, A1, and A2, for the asymmetric distributions, we conclude that the I_3 FACI has the finite-sample coverages that are similar to those of $I_{2,q}$ with $q(t) = (t(1-t))^{0.4}$, being shorter than the latter FACI on average. On the other hand, I_4 has a similar average length to that of $I_{2,q}$ with $q(t) = (t(1-t))^{0.4}$, but has higher finite-sample coverages than the latter FACI. Thus, I_3 and I_4 improve on $I_{2,q}$ with $q(t) = (t(1-t))^{0.4}$ in terms of the expected length and finite-sample coverage probability,

respectively. In addition to Table 4, for our right-skewed distributions, we compare the 95% confidence level I_3 and I_4 FACIs to their respective I_1 counterparts of the same average lengths (and thus of higher nominal levels, since the ratios of the average lengths $\hat{r}_3$ and $\hat{r}_4$ reported in Table 4 are greater than one). Thus, in Table 6, for each sample size used in Table 4, we find the respective I_1 matches for I_3 and I_4 (so that the absolute difference between the average lengths of the FACI and I_1 divided by the average length of the FACI is less than or equal to 0.00001) and report the finite-sample coverages $\widehat{CP}_{1,\text{adj}} \times 100\%$ of the thus adjusted I_1 together with their nominal asymptotic coverages and the difference in the finite-sample coverage between the 95% confidence level FACI and the adjusted I_1 CI. Both I_3 and I_4 have higher coverages than their adjusted CLT-based I_1 counterparts, especially for smaller sample sizes and Pareto(2, 1), with a few exceptions for large sample sizes where I_3 and I_4 do not provide more overcoverage than the respective adjusted I_1 CIs. Similar comparisons with analogous conclusions are performed for the 90% and 98% confidence level I_3 and I_4 FACIs with their respective equal-average-length I_1 CIs in Tables A3 and A4 from Appendix to Chapter 2.

2. Main Results

Table 6: Comparison of coverages $\widehat{CP}_3 \cdot 100\%$ of I_3 and $\widehat{CP}_4 \cdot 100\%$ of I_4 (both at 95% confidence level) with coverages $\widehat{CP}_{1,\text{adj}}$ of equal-average-length adjusted I_1 counterparts (at higher nominal levels) for right-skewed distributions.

Distribution	n	I_3				I_4			
		I_3	$\widehat{CP}_3$	$\widehat{CP}_{1,\text{adj}}$ (Nominal)	$(\widehat{CP}_3 - \widehat{CP}_{1,\text{adj}})$	I_4	$\widehat{CP}_4$	$\widehat{CP}_{1,\text{adj}}$ (Nominal)	$(\widehat{CP}_4 - \widehat{CP}_{1,\text{adj}})$
		avg. len.	$\times 100\%$	$\times 100\%$	$\times 100\%$	avg. len.	$\times 100\%$	$\times 100\%$	$\times 100\%$
<i>Pareto</i> (2, 1)	50	1.334	88.65	87.08 (98.595)	1.57	1.426	92.05	88.59 (99.133)	3.46
	100	0.966	89.94	87.82 (97.959)	2.12	1.038	93.16	89.71 (98.734)	3.45
	300	0.614	91.83	90.20 (97.108)	1.63	0.658	94.64	91.74 (98.092)	2.90
	500	0.481	91.81	90.13 (96.803)	1.68	0.516	94.77	91.84 (97.857)	2.93
	1000	0.357	92.74	91.54 (96.472)	1.20	0.382	95.34	92.93 (97.591)	2.41
<i>Weibull</i> (0.4, 1)	50	6.001	87.99	86.62 (98.590)	1.37	6.412	90.89	87.84 (99.128)	3.05
	100	4.249	91.27	89.88 (97.962)	1.39	4.570	93.83	91.14 (98.738)	2.69
	300	2.479	93.37	92.50 (97.109)	0.87	2.660	95.48	93.66 (98.092)	1.82
	500	1.917	93.82	93.10 (96.806)	0.72	2.056	95.90	94.42 (97.856)	1.48
	1000	1.355	94.70	94.42 (96.476)	0.28	1.452	95.92	95.77 (97.596)	0.15
<i>InverseNormal</i> (7, 1)	50	11.086	88.38	87.00 (98.591)	1.38	11.845	90.84	88.02 (99.129)	2.82
	100	7.857	91.88	90.55 (97.955)	1.33	8.446	94.20	91.68 (98.730)	2.52
	300	4.534	94.62	93.78 (97.102)	0.84	4.864	96.02	94.89 (98.086)	1.13
	500	3.469	95.04	94.52 (96.804)	0.52	3.721	96.25	95.53 (97.856)	0.72
	1000	2.434	95.21	95.10 (96.472)	0.11	2.609	96.15	96.19 (97.593)	−0.04
<i>Log-logistic</i> (2.1, 1)	50	1.252	91.51	90.09 (98.594)	1.42	1.338	94.40	91.41 (99.132)	2.99
	100	0.918	93.11	91.53 (97.955)	1.58	0.987	95.79	92.96 (98.727)	2.83
	300	0.548	93.67	92.38 (97.109)	1.29	0.588	96.11	93.82 (98.093)	2.29
	500	0.439	94.01	92.63 (96.804)	1.38	0.471	96.03	94.02 (97.855)	2.01
	1000	0.319	93.63	92.84 (96.475)	0.79	0.342	95.73	94.34 (97.596)	1.39
<i>Log-normal</i> (0, 1.3)	50	2.808	91.30	89.88 (98.596)	1.42	3.001	93.71	91.06 (99.134)	2.65
	100	1.962	92.41	91.31 (97.959)	1.10	2.109	94.78	92.69 (98.735)	2.09
	300	1.153	94.70	93.92 (97.107)	0.78	1.237	96.19	95.02 (98.090)	1.17
	500	0.880	94.73	94.03 (96.803)	0.70	0.944	96.24	95.21 (97.854)	1.03
	1000	0.626	95.17	94.74 (96.476)	0.43	0.671	96.43	96.03 (97.596)	0.40
<i>Frechet</i> (2.5)	50	0.813	93.52	92.37 (98.593)	1.15	0.869	95.64	93.50 (99.132)	2.14
	100	0.564	94.27	93.29 (97.961)	0.98	0.607	96.36	94.45 (98.736)	1.91
	300	0.330	94.82	94.09 (97.105)	0.73	0.354	96.45	95.27 (98.090)	1.18
	500	0.255	94.94	94.22 (96.801)	0.72	0.274	96.76	95.35 (97.850)	1.41
	1000	0.182	95.36	94.87 (96.472)	0.49	0.195	96.80	95.98 (97.592)	0.82
<i>Half-t</i> (3)	50	0.834	95.62	94.98 (98.596)	0.64	0.891	97.30	95.80 (99.134)	1.50
	100	0.571	96.03	95.42 (97.959)	0.61	0.614	97.57	96.52 (98.735)	1.05
	300	0.321	96.09	95.80 (97.109)	0.29	0.344	97.06	96.85 (98.093)	0.21
	500	0.246	95.77	95.66 (96.807)	0.11	0.264	96.78	96.80 (97.858)	−0.02
	1000	0.172	95.49	95.65 (96.474)	−0.16	0.185	96.32	96.91 (97.596)	−0.59

Remark 6. As per Tables 4, A1, and A2, the FACIs $I_{2,q}$ with $q(t) = (t(1-t))^{0.4}$, I_3 ,

and I_4 exhibit a more expensive, yet, in the case of right-skewed distributions with higher δ_{100} (if defined) or $\tilde{\delta}_{100}$, potentially desirable trade-off between their finite-sample coverage probabilities and expected lengths as compared to the FACI $I_{2,q}$ with $q(t) \equiv 1$, which was obtained from the FCLT in (7) for the unweighted Student process by Tuzov (2014) and Martsynyuk and Tuzov (2016). Moreover, the finite-sample coverage probabilities of other FACIs considered in these works can be seen to be significantly exceeded by those of $I_{2,q}$ with $q(t) = (t(1-t))^{0.4}$, I_3 , and I_4 for our right-skewed distributions (we did not include our computations in this regard here).

3 Conclusions

In this paper, we derived FACIs $I_{2,q}$, I_3 , and I_4 for a population mean μ of a distribution in DAN by using convergence in distribution of the integral functional h_2 of (14) and its modification h_3 of (17), applied to the Student process $T_n^t(X_1 - \mu, \dots, X_n - \mu)/q([nt]/n)$ weighted, respectively, with the function $q(t) = (t(1-t))^a$ ($a = 0$ and 0.4) and the function $q(t) = [t(1-t)]^{0.25}[\log \log(t(1-t))^{-1}]^{-0.5}$. We showed, numerically, that our $1 - \alpha$ FACIs have higher finite-sample coverage probabilities than, and may be reasonable alternatives to, the CLT-based $1 - \alpha$ CI I_1 for right-skewed distributions in DAN, although the FACIs are longer than I_1 on average by at most 1.361 for $n = 50$ and only at most 1.156 for $n = 1000$. The obtained FACIs may be especially appealing for right-skewed distributions in DAN with larger values of the Berry–Esseen bounds δ_n of (31) (if defined) or $\tilde{\delta}_n$ of (33), since for such distributions the finite-sample coverage probabilities of I_1 may be lower, sometimes significantly, than the asymptotic confidence level $1 - \alpha$, especially for smaller samples, and may have slow convergence rates. Moreover, $I_{2,q}$ with $q(t) = (t(1-t))^{0.4}$, I_3 , and I_4 can be seen to be better potential alternatives to the CLT-based CI I_1 than the FACIs from the FCLT in (7) for the unweighted Student process, in particular, the $I_{2,q}$ FACI with $q(t) \equiv 1$ obtained in Tuzov (2014) and Martsynyuk and Tuzov (2016). Furthermore, our h_3 -based FACIs, namely I_3 and I_4 , improve on the h_2 -based $I_{2,q}$ FACI with $q(t) = (t(1-t))^{0.4}$ in terms of the expected length and finite-sample coverage probability, respectively.

Appendix to Chapter 2

 Table A1: Comparison of I_1 , $I_{2,q}$ with $q(t) = (t(1-t))^a$, I_3 , and I_4 at the 90% confidence level.

Distribution	n	$\widehat{CP}_1 \cdot 100\%$	FACI Coverages				Ratios of Average Lengths			
			$\widehat{CP}_{2,q} \cdot 100\%$	$\widehat{CP}_{2,q} \cdot 100\%$	$\widehat{CP}_3 \cdot 100\%$	$\widehat{CP}_4 \cdot 100\%$	$\widehat{r}_{2,q}$	$\widehat{r}_{2,q}$	$\widehat{r}_3$	$\widehat{r}_4$
			a = 0	a = 0.4					a = 0	a = 0.4
Pareto(2,1)	50	76.62 (0.42)	77.91 (0.41)	84.46 (0.36)	84.76 (0.36)	88.47 (0.32)	1.121	1.361	1.259	1.344
	100	78.31 (0.41)	79.90 (0.40)	85.04 (0.36)	85.80 (0.35)	89.53 (0.31)	1.111	1.280	1.189	1.280
	300	82.15 (0.38)	83.56 (0.37)	86.88 (0.34)	87.17 (0.33)	90.71 (0.29)	1.106	1.192	1.120	1.202
	500	83.14 (0.37)	83.47 (0.37)	86.88 (0.34)	87.26 (0.33)	90.28 (0.30)	1.104	1.162	1.100	1.179
	1000	84.67 (0.36)	85.93 (0.35)	87.85 (0.33)	88.24 (0.32)	91.03 (0.29)	1.104	1.129	1.079	1.156
Weibull(0.4,1)	50	77.18 (0.42)	78.20 (0.41)	84.06 (0.37)	84.33 (0.36)	87.44 (0.33)	1.120	1.361	1.259	1.343
	100	81.37 (0.39)	82.90 (0.38)	87.86 (0.33)	88.13 (0.32)	90.54 (0.29)	1.111	1.280	1.190	1.281
	300	85.61 (0.35)	85.75 (0.35)	89.68 (0.30)	89.18 (0.31)	90.82 (0.29)	1.106	1.192	1.120	1.202
	500	86.56 (0.34)	87.30 (0.33)	90.22 (0.30)	89.78 (0.30)	91.30 (0.28)	1.105	1.162	1.100	1.179
	1000	87.82 (0.33)	88.35 (0.32)	91.11 (0.28)	90.09 (0.30)	91.17 (0.28)	1.103	1.129	1.079	1.156
InverseNormal(7,1)	50	78.44 (0.41)	79.27 (0.41)	84.90 (0.36)	84.91 (0.36)	87.81 (0.33)	1.121	1.361	1.259	1.343
	100	82.47 (0.38)	83.35 (0.37)	88.35 (0.32)	88.20 (0.32)	90.42 (0.29)	1.111	1.280	1.189	1.279
	300	86.65 (0.34)	87.16 (0.33)	91.11 (0.28)	90.28 (0.30)	91.74 (0.28)	1.105	1.192	1.119	1.201
	500	88.08 (0.32)	87.99 (0.33)	91.94 (0.27)	90.71 (0.29)	91.64 (0.28)	1.104	1.162	1.100	1.179
	1000	89.29 (0.31)	88.79 (0.32)	91.92 (0.27)	90.65 (0.29)	91.10 (0.28)	1.103	1.128	1.079	1.156
Log-logistic(2.1,1)	50	79.73 (0.40)	80.88 (0.39)	87.50 (0.33)	87.82 (0.33)	91.03 (0.29)	1.121	1.361	1.259	1.344
	100	82.65 (0.38)	83.18 (0.37)	88.82 (0.32)	89.19 (0.31)	92.49 (0.26)	1.112	1.280	1.189	1.279
	300	85.16 (0.36)	85.46 (0.35)	89.35 (0.31)	89.49 (0.31)	91.94 (0.27)	1.106	1.192	1.120	1.202
	500	85.18 (0.36)	85.99 (0.35)	89.30 (0.31)	89.28 (0.31)	92.16 (0.27)	1.104	1.162	1.100	1.179
	1000	86.29 (0.34)	86.97 (0.34)	89.34 (0.31)	89.46 (0.31)	91.76 (0.27)	1.104	1.129	1.079	1.156
Log-normal(0,1.3)	50	80.89 (0.39)	82.02 (0.38)	87.69 (0.33)	87.51 (0.33)	90.64 (0.29)	1.121	1.361	1.260	1.344
	100	82.77 (0.38)	83.35 (0.37)	88.79 (0.32)	88.64 (0.32)	91.34 (0.28)	1.111	1.280	1.189	1.280
	300	87.03 (0.34)	87.18 (0.33)	91.27 (0.28)	90.75 (0.29)	92.09 (0.27)	1.106	1.192	1.120	1.202
	500	87.21 (0.33)	87.66 (0.33)	90.93 (0.29)	90.26 (0.30)	91.97 (0.27)	1.104	1.162	1.099	1.179
	1000	88.24 (0.32)	88.52 (0.32)	91.31 (0.28)	90.54 (0.29)	91.96 (0.27)	1.104	1.129	1.079	1.156
Frechet(2.5)	50	82.70 (0.38)	84.05 (0.37)	90.40 (0.29)	90.38 (0.29)	92.77 (0.26)	1.120	1.361	1.259	1.344
	100	84.96 (0.36)	85.53 (0.35)	90.81 (0.29)	90.82 (0.29)	92.89 (0.26)	1.111	1.280	1.190	1.281
	300	87.06 (0.34)	87.11 (0.34)	91.16 (0.28)	90.53 (0.29)	92.24 (0.27)	1.105	1.192	1.120	1.202
	500	87.58 (0.33)	87.65 (0.33)	91.18 (0.28)	90.52 (0.29)	92.41 (0.26)	1.104	1.162	1.099	1.178
	1000	89.03 (0.31)	88.62 (0.32)	91.45 (0.28)	91.06 (0.29)	92.28 (0.27)	1.103	1.129	1.079	1.156
Half-t(3)	50	86.12 (0.35)	86.82 (0.34)	92.92 (0.26)	92.50 (0.26)	94.25 (0.23)	1.121	1.361	1.260	1.344
	100	87.11 (0.34)	87.43 (0.33)	92.98 (0.26)	92.42 (0.26)	93.88 (0.24)	1.111	1.280	1.189	1.280
	300	88.42 (0.32)	88.40 (0.32)	93.07 (0.25)	91.43 (0.28)	92.17 (0.27)	1.106	1.192	1.120	1.202
	500	89.02 (0.31)	88.75 (0.32)	92.55 (0.26)	91.36 (0.28)	91.93 (0.27)	1.105	1.162	1.100	1.179
	1000	89.35 (0.31)	89.29 (0.31)	92.25 (0.27)	91.03 (0.29)	91.65 (0.28)	1.103	1.128	1.079	1.156
SymPareto(2,1)	50	89.94 (0.30)	90.37 (0.30)	96.95 (0.17)	95.04 (0.22)	94.57 (0.23)	1.120	1.361	1.259	1.344
	100	90.29 (0.30)	90.58 (0.29)	95.91 (0.20)	94.14 (0.23)	93.21 (0.25)	1.111	1.280	1.189	1.279
	300	90.44 (0.29)	90.32 (0.30)	94.65 (0.23)	92.50 (0.26)	91.75 (0.28)	1.106	1.192	1.121	1.203
	500	90.53 (0.29)	90.36 (0.30)	94.14 (0.23)	92.37 (0.27)	91.77 (0.27)	1.104	1.162	1.100	1.179
	1000	90.11 (0.30)	90.05 (0.30)	93.35 (0.25)	91.66 (0.28)	91.15 (0.28)	1.102	1.128	1.079	1.156
t(3)	50	89.63 (0.30)	89.55 (0.31)	96.62 (0.18)	94.86 (0.22)	94.56 (0.23)	1.120	1.360	1.259	1.343
	100	89.71 (0.30)	89.65 (0.30)	95.58 (0.21)	93.57 (0.25)	92.91 (0.26)	1.111	1.280	1.189	1.280
	300	90.08 (0.30)	90.38 (0.29)	94.38 (0.23)	92.19 (0.27)	91.68 (0.28)	1.105	1.192	1.120	1.202
	500	90.13 (0.30)	90.30 (0.30)	93.62 (0.24)	91.75 (0.28)	91.28 (0.28)	1.104	1.162	1.100	1.179
	1000	89.38 (0.31)	89.53 (0.31)	92.57 (0.26)	90.81 (0.29)	90.50 (0.29)	1.103	1.129	1.079	1.156

Table A2: Comparison of I_1 , $I_{2,q}$ with $q(t) = (t(1-t))^a$, I_3 , and I_4 at the 98% confidence level.

Distribution	n	$\widehat{CP}_1 \cdot 100\%$	FACI Coverages				Ratios of Average Lengths			
			$\widehat{CP}_{2,q} \cdot 100\%$	$\widehat{CP}_{2,q} \cdot 100\%$	$\widehat{CP}_3 \cdot 100\%$	$\widehat{CP}_4 \cdot 100\%$	$\hat{r}_{2,q}$	$\hat{r}_{2,q}$	$\hat{r}_3$	$\hat{r}_4$
			$a = 0$	$a = 0.4$			$a = 0$	$a = 0.4$		
Pareto(2,1)	50	85.64 (0.35)	87.64 (0.33)	92.21 (0.27)	92.09 (0.27)	94.26 (0.23)	1.119	1.351	1.248	1.330
	100	87.94 (0.33)	89.66 (0.30)	92.90 (0.26)	93.08 (0.25)	95.66 (0.20)	1.111	1.272	1.178	1.261
	300	91.62 (0.28)	92.82 (0.26)	94.49 (0.23)	94.79 (0.22)	96.84 (0.17)	1.105	1.185	1.110	1.187
	500	92.16 (0.27)	93.10 (0.25)	94.64 (0.23)	95.05 (0.22)	97.23 (0.16)	1.103	1.155	1.090	1.165
	1000	93.73 (0.24)	94.56 (0.23)	95.56 (0.21)	95.97 (0.20)	97.77 (0.15)	1.103	1.122	1.070	1.143
Weibull(0.4,1)	50	85.42 (0.35)	87.06 (0.34)	91.02 (0.29)	90.86 (0.29)	93.03 (0.25)	1.119	1.351	1.247	1.330
	100	89.94 (0.30)	91.29 (0.28)	93.73 (0.24)	93.80 (0.24)	95.87 (0.20)	1.110	1.272	1.178	1.262
	300	93.58 (0.25)	94.56 (0.23)	96.00 (0.20)	96.04 (0.20)	97.65 (0.15)	1.105	1.185	1.110	1.187
	500	94.61 (0.23)	95.21 (0.21)	96.57 (0.18)	96.73 (0.18)	97.94 (0.14)	1.104	1.155	1.090	1.165
	1000	96.14 (0.19)	96.44 (0.19)	97.35 (0.16)	97.40 (0.16)	98.39 (0.13)	1.103	1.122	1.070	1.143
InverseNormal(7,1)	50	86.05 (0.35)	87.77 (0.33)	91.01 (0.29)	90.93 (0.29)	92.95 (0.26)	1.119	1.351	1.247	1.330
	100	90.58 (0.29)	91.78 (0.27)	94.19 (0.23)	94.36 (0.23)	96.14 (0.19)	1.111	1.272	1.178	1.261
	300	94.75 (0.22)	95.23 (0.21)	96.65 (0.18)	96.74 (0.18)	98.04 (0.14)	1.105	1.185	1.110	1.187
	500	95.70 (0.20)	96.40 (0.19)	97.31 (0.16)	97.38 (0.16)	98.51 (0.12)	1.104	1.155	1.090	1.165
	1000	96.64 (0.18)	97.03 (0.17)	97.79 (0.15)	97.84 (0.15)	98.62 (0.12)	1.103	1.122	1.070	1.143
Log-logistic(2,1,1)	50	88.88 (0.31)	90.57 (0.29)	94.60 (0.23)	94.47 (0.23)	96.55 (0.18)	1.120	1.351	1.248	1.330
	100	91.63 (0.28)	92.95 (0.26)	95.66 (0.20)	95.75 (0.20)	97.65 (0.15)	1.111	1.272	1.177	1.261
	300	93.68 (0.24)	94.36 (0.23)	96.25 (0.19)	96.43 (0.19)	98.02 (0.14)	1.105	1.185	1.110	1.187
	500	94.16 (0.23)	95.17 (0.21)	96.44 (0.19)	96.73 (0.18)	98.25 (0.13)	1.104	1.155	1.090	1.165
	1000	94.80 (0.22)	95.40 (0.21)	96.34 (0.19)	96.58 (0.18)	97.98 (0.14)	1.103	1.122	1.070	1.143
Log-normal(0,1,3)	50	88.70 (0.32)	90.59 (0.29)	93.95 (0.24)	93.94 (0.24)	95.79 (0.20)	1.119	1.351	1.248	1.331
	100	91.38 (0.28)	92.40 (0.26)	94.83 (0.22)	94.97 (0.22)	96.91 (0.17)	1.111	1.272	1.178	1.262
	300	94.91 (0.22)	95.46 (0.21)	96.79 (0.18)	96.92 (0.17)	98.30 (0.13)	1.105	1.185	1.110	1.187
	500	95.42 (0.21)	95.88 (0.20)	97.15 (0.17)	97.29 (0.16)	98.37 (0.13)	1.104	1.155	1.090	1.165
	1000	96.40 (0.19)	96.83 (0.18)	97.68 (0.15)	97.83 (0.15)	98.78 (0.11)	1.103	1.122	1.070	1.143
Frechet(2.5)	50	91.41 (0.28)	92.66 (0.26)	95.83 (0.20)	95.75 (0.20)	97.38 (0.16)	1.119	1.351	1.248	1.330
	100	93.37 (0.25)	94.13 (0.24)	96.52 (0.18)	96.62 (0.18)	98.17 (0.13)	1.111	1.272	1.178	1.262
	300	95.15 (0.21)	95.69 (0.20)	97.29 (0.16)	97.39 (0.16)	98.48 (0.12)	1.105	1.185	1.110	1.187
	500	95.55 (0.21)	96.15 (0.19)	97.32 (0.16)	97.52 (0.16)	98.65 (0.12)	1.104	1.155	1.090	1.165
	1000	96.44 (0.19)	97.04 (0.17)	97.88 (0.14)	97.86 (0.14)	98.77 (0.11)	1.103	1.122	1.070	1.143
Half-t(3)	50	94.04 (0.24)	95.09 (0.22)	97.68 (0.15)	97.62 (0.15)	98.73 (0.11)	1.120	1.351	1.248	1.331
	100	95.50 (0.21)	96.21 (0.19)	98.23 (0.13)	98.24 (0.13)	99.04 (0.10)	1.111	1.272	1.178	1.262
	300	96.76 (0.18)	96.97 (0.17)	98.35 (0.13)	98.27 (0.13)	98.94 (0.10)	1.105	1.185	1.110	1.187
	500	97.00 (0.17)	97.09 (0.17)	98.43 (0.12)	98.26 (0.13)	98.85 (0.11)	1.104	1.155	1.090	1.165
	1000	97.27 (0.16)	97.44 (0.16)	98.31 (0.13)	98.06 (0.14)	98.71 (0.11)	1.103	1.122	1.070	1.143
SymPareto(2,1)	50	98.22 (0.13)	98.42 (0.12)	99.72 (0.05)	99.49 (0.07)	99.47 (0.07)	1.119	1.351	1.247	1.330
	100	98.36 (0.13)	98.23 (0.13)	99.63 (0.06)	99.16 (0.09)	99.00 (0.10)	1.110	1.272	1.177	1.261
	300	98.32 (0.13)	98.24 (0.13)	99.39 (0.08)	98.84 (0.11)	98.81 (0.11)	1.105	1.185	1.110	1.187
	500	98.57 (0.12)	98.43 (0.12)	99.40 (0.08)	98.95 (0.10)	98.92 (0.10)	1.104	1.155	1.090	1.165
	1000	98.17 (0.13)	98.06 (0.14)	99.02 (0.10)	98.43 (0.12)	98.49 (0.12)	1.102	1.122	1.070	1.143
$t(3)$	50	97.92 (0.14)	98.15 (0.13)	99.75 (0.05)	99.45 (0.07)	99.38 (0.08)	1.119	1.351	1.247	1.330
	100	98.00 (0.14)	98.03 (0.14)	99.59 (0.06)	99.10 (0.09)	99.06 (0.10)	1.110	1.272	1.178	1.261
	300	98.28 (0.13)	98.00 (0.14)	99.32 (0.08)	98.79 (0.11)	98.66 (0.11)	1.105	1.185	1.110	1.187
	500	97.92 (0.14)	98.04 (0.14)	99.06 (0.10)	98.53 (0.12)	98.46 (0.12)	1.104	1.155	1.090	1.165
	1000	97.95 (0.14)	98.05 (0.14)	98.88 (0.11)	98.23 (0.13)	98.24 (0.13)	1.103	1.122	1.070	1.143

Table A3: Comparison of coverages $\widehat{CP}_3 \cdot 100\%$ of I_3 and $\widehat{CP}_4 \cdot 100\%$ of I_4 (both at 90% confidence level) with coverages $\widehat{CP}_{1,\text{adj}}$ of equal-average-length adjusted I_1 counterparts (at higher nominal levels) for right-skewed distributions.

Distribution	n	I_3				I_4			
		avg. len.	$\widehat{CP}_3 \cdot 100\%$	$\widehat{CP}_{1,\text{adj}}$ (Nominal)	$(\widehat{CP}_3 - \widehat{CP}_{1,\text{adj}})$ ·100%	avg. len.	$\widehat{CP}_4 \cdot 100\%$	$\widehat{CP}_{1,\text{adj}}$ (Nominal)	$(\widehat{CP}_4 - \widehat{CP}_{1,\text{adj}})$ ·100%
<i>Pareto</i> (2,1)	50	1.125	84.76	82.80 (96.169)	1.96	1.201	88.47	84.51 (97.294)	3.96
	100	0.815	85.80	83.66 (94.957)	2.14	0.877	89.53	85.52 (96.477)	4.01
	300	0.517	87.17	85.59 (93.457)	1.58	0.555	90.71	87.24 (95.194)	3.47
	500	0.406	87.26	85.97 (92.947)	1.29	0.435	90.28	87.83 (94.756)	2.45
	1000	0.301	88.24	87.08 (92.407)	1.16	0.322	91.03	89.14 (94.270)	1.89
<i>Weibull</i> (0.4,1)	50	5.061	84.33	82.92 (96.157)	1.41	5.399	87.44	84.25 (97.281)	3.19
	100	3.586	88.13	86.00 (94.964)	2.13	3.860	90.54	87.94 (96.487)	2.60
	300	2.090	89.18	88.70 (93.459)	0.48	2.243	90.82	90.45 (95.197)	0.37
	500	1.617	89.78	89.16 (92.954)	0.62	1.734	91.30	90.94 (94.755)	0.36
	1000	1.142	90.09	90.17 (92.414)	−0.08	1.224	91.17	92.09 (94.283)	−0.92
<i>InverseNormal</i> (7,1)	50	9.350	84.91	83.73 (96.160)	1.18	9.975	87.81	85.04 (97.282)	2.77
	100	6.628	88.20	87.08 (94.948)	1.12	7.132	90.42	88.67 (96.464)	1.75
	300	3.823	90.28	90.26 (93.443)	0.02	4.101	91.74	91.95 (95.180)	−0.21
	500	2.925	90.71	90.99 (92.949)	−0.28	3.137	91.64	92.70 (94.755)	−1.06
	1000	2.052	90.65	91.45 (92.406)	−0.80	2.198	91.10	93.11 (94.274)	−2.01
<i>Log-logistic</i> (2.1,1)	50	1.056	87.82	85.91 (96.168)	1.91	1.127	91.03	87.52 (97.291)	3.51
	100	0.774	89.19	87.39 (94.947)	1.80	0.833	92.49	89.20 (96.458)	3.29
	300	0.462	89.49	88.43 (93.460)	1.06	0.496	91.94	90.05 (95.199)	1.89
	500	0.370	89.28	88.27 (92.949)	1.01	0.397	92.16	90.20 (94.752)	1.96
	1000	0.269	89.46	88.53 (92.413)	0.93	0.289	91.76	90.45 (94.281)	1.31
<i>Log-normal</i> (0,1.3)	50	2.368	87.51	86.49 (96.174)	1.02	2.528	90.64	87.61 (97.300)	3.03
	100	1.655	88.64	87.52 (94.958)	1.12	1.782	91.34	89.33 (96.480)	2.01
	300	0.972	90.75	90.18 (93.455)	0.57	1.043	92.09	91.96 (95.190)	0.13
	500	0.742	90.26	90.07 (92.947)	0.19	0.796	91.97	91.89 (94.750)	0.08
	1000	0.528	90.54	90.65 (92.415)	−0.11	0.565	91.96	92.54 (94.281)	−0.58
<i>Frechet</i> (2.5)	50	0.686	90.38	89.06 (96.164)	1.32	0.732	92.77	90.38 (97.292)	2.39
	100	0.476	90.82	89.59 (94.963)	1.23	0.513	92.89	91.21 (96.483)	1.68
	300	0.278	90.53	90.37 (93.449)	0.16	0.298	92.24	92.16 (95.190)	0.08
	500	0.215	90.52	90.46 (92.941)	0.06	0.231	92.41	92.07 (94.738)	0.34
	1000	0.154	91.06	91.22 (92.405)	−0.16	0.164	92.28	92.84 (94.270)	−0.56
<i>Half-t</i> (3)	50	0.703	92.50	92.17 (96.174)	0.33	0.751	94.25	93.25 (97.298)	1.00
	100	0.482	92.42	91.92 (94.959)	0.50	0.518	93.88	93.64 (96.480)	0.24
	300	0.271	91.43	92.09 (93.458)	−0.66	0.290	92.17	93.86 (95.198)	−1.69
	500	0.208	91.36	91.74 (92.955)	−0.38	0.223	91.93	93.67 (94.758)	−1.74
	1000	0.145	91.03	91.68 (92.411)	−0.65	0.156	91.65	93.41 (94.280)	−1.76

Table A4: Comparison of coverages $\widehat{CP}_3 \cdot 100\%$ of I_3 and $\widehat{CP}_4 \cdot 100\%$ of I_4 (both at 98% confidence level) with coverages $\widehat{CP}_{1,\text{adj}}$ of equal-average-length adjusted I_1 counterparts (at higher nominal levels) for right-skewed distributions.

Distribution	n	I_3				I_4			
		avg. len.	$\widehat{CP}_3 \cdot 100\%$	$\widehat{CP}_{1,\text{adj}}$ (Nominal)	$(\widehat{CP}_3 - \widehat{CP}_{1,\text{adj}})$ $\cdot 100\%$	avg. len.	$\widehat{CP}_4 \cdot 100\%$	$\widehat{CP}_{1,\text{adj}}$ (Nominal)	$(\widehat{CP}_4 - \widehat{CP}_{1,\text{adj}})$ $\cdot 100\%$
<i>Pareto</i> (2,1)	50	1.576	92.09	90.68 (99.630)	1.41	1.681	94.26	91.98 (99.803)	2.28
	100	1.141	93.08	91.73 (99.385)	1.35	1.222	95.66	93.02 (99.666)	2.64
	300	0.725	94.79	93.82 (99.018)	0.97	0.776	96.84	94.84 (99.424)	2.00
	500	0.569	95.05	93.76 (98.877)	1.29	0.608	97.23	95.14 (99.328)	2.09
	1000	0.422	95.97	94.99 (98.719)	0.98	0.451	97.77	96.09 (99.216)	1.68
<i>Weibull</i> (0.4,1)	50	7.093	90.86	89.86 (99.629)	1.00	7.563	93.03	90.89 (99.802)	2.14
	100	5.021	93.80	92.63 (99.386)	1.17	5.379	95.87	93.63 (99.667)	2.24
	300	2.930	96.04	95.10 (99.019)	0.94	3.133	97.65	96.08 (99.425)	1.57
	500	2.267	96.73	95.97 (98.878)	0.76	2.423	97.94	96.80 (99.328)	1.14
	1000	1.601	97.40	96.98 (98.720)	0.42	1.711	98.39	97.68 (99.218)	0.71
<i>InverseNormal</i> (7,1)	50	13.104	90.93	89.83 (99.629)	1.10	13.971	92.95	90.90 (99.802)	2.05
	100	9.285	94.36	93.29 (99.384)	1.07	9.943	96.14	94.30 (99.665)	1.84
	300	5.360	96.74	96.17 (99.017)	0.57	5.731	98.04	96.84 (99.423)	1.20
	500	4.101	97.38	96.76 (98.877)	0.62	4.384	98.51	97.47 (99.328)	1.04
	1000	2.878	97.84	97.53 (98.719)	0.31	3.075	98.62	98.18 (99.217)	0.44
<i>Log-logistic</i> (2.1,1)	50	1.480	94.47	93.43 (99.630)	1.04	1.578	96.55	94.46 (99.803)	2.09
	100	1.085	95.75	94.57 (99.384)	1.18	1.162	97.65	95.69 (99.664)	1.96
	300	0.648	96.43	95.42 (99.019)	1.01	0.693	98.02	96.38 (99.425)	1.64
	500	0.519	96.73	95.85 (98.878)	0.88	0.555	98.25	96.84 (99.328)	1.41
	1000	0.378	96.58	95.80 (98.720)	0.78	0.404	97.98	96.61 (99.218)	1.37
<i>Log-normal</i> (0,1.3)	50	3.318	93.94	92.96 (99.630)	0.98	3.538	95.79	93.91 (99.804)	1.88
	100	2.318	94.97	94.16 (99.385)	0.81	2.483	96.91	95.00 (99.666)	1.91
	300	1.363	96.92	96.39 (99.018)	0.53	1.457	98.30	97.13 (99.424)	1.17
	500	1.040	97.29	96.65 (98.877)	0.64	1.112	98.37	97.43 (99.328)	0.94
	1000	0.740	97.83	97.37 (98.720)	0.46	0.791	98.78	97.99 (99.218)	0.79
<i>Frechet</i> (2.5)	50	0.961	95.75	95.07 (99.629)	0.68	1.025	97.38	95.82 (99.803)	1.56
	100	0.667	96.62	95.73 (99.386)	0.89	0.714	98.17	96.49 (99.667)	1.68
	300	0.390	97.39	96.67 (99.017)	0.72	0.417	98.48	97.35 (99.424)	1.13
	500	0.302	97.52	96.77 (98.876)	0.75	0.322	98.65	97.62 (99.326)	1.03
	1000	0.215	97.86	97.49 (98.718)	0.37	0.230	98.77	98.03 (99.217)	0.74
<i>Half-t</i> (3)	50	0.985	97.62	97.05 (99.630)	0.57	1.051	98.73	97.66 (99.804)	1.07
	100	0.674	98.24	97.51 (99.386)	0.73	0.722	99.04	98.26 (99.666)	0.78
	300	0.379	98.27	97.96 (99.019)	0.31	0.406	98.94	98.51 (99.425)	0.43
	500	0.291	98.26	98.09 (98.878)	0.17	0.311	98.85	98.69 (99.329)	0.16
	1000	0.204	98.06	98.05 (98.719)	0.01	0.218	98.71	98.71 (99.218)	0.00

References

- Bentkus, V., & Götze, F. (1996). The berry–esseen bound for student’s statistic. *The Annals of Probability*, 24(1), 491–503. <https://doi.org/10.1214/aop/1042644728>
- Csörgő, M., Csörgő, S., Horváth, L., & Mason, D. M. (1986). Weighted empirical and quantile processes. *The Annals of Probability*, 14(1), 31–85. <https://doi.org/10.1214/aop/1176992617>
- Csörgő, M., Szyszkowicz, B., & Wang, Q. (2003). Donsker’s theorem for self-normalized partial sums processes. *The Annals of Probability*, 31(3), 1228–1240. <https://doi.org/10.1214/aop/1055425777>
- Csörgő, M., Szyszkowicz, B., & Wang, Q. (2008). On weighted approximations in $d[0,1]$ with applications to self-normalized partial sum processes. *Acta Mathematica Hungarica*, 121(4), 307–332. <https://doi.org/10.1007/s10474-008-7216-5>
- Csörgő, S., & Horváth, L. (1981). On the koziol–green model for random censorship. *Biometrika*, 68, 391–401. <https://doi.org/10.1093/biomet/68.2.391>
- Erdős, P., & Kac, M. (1946). On certain limit theorems of the theory of probability. *Bulletin of the American Mathematical Society*, 52, 292–302. <https://doi.org/10.1090/S0002-9904-1946-08560-2>
- Logan, B. F., Mallows, C. L., Rice, S. O., & Shepp, L. A. (1973). Limit distributions of self-normalized sums. *Annals of Probability*, 1(5), 788–809. <https://doi.org/10.1214/aop/1176996846>
- Martsynyuk, Y. V., & Tuzov, E. (2016). Exploring functional clt confidence intervals for a population mean in the domain of attraction of the normal law. *Acta Mathematica Hungarica*, 148(2), 493–508. <https://doi.org/10.1007/s10474-016-0595-0>
- Shao, Q.-M. (2005). An explicit berry–esseen bound for student’s t-statistic via stein’s method. In *Stein’s method and applications* (pp. 143–155, Vol. 5). Singapore University Press. https://doi.org/10.1142/9789812567673_0009
- Shao, Q.-M., & Wang, Q. (2013). Self-normalized limit theorems: A survey. *Probability Surveys*, 10, 69–93. <https://doi.org/10.1214/13-PS216>

Tuzov, E. (2014). *Exploring functional asymptotic confidence intervals for a population mean*
[M.Sc. thesis]. University of Manitoba. <http://hdl.handle.net/1993/23426>

Chapter 3

Powerful Nonparametric

Asymptotic Tests for Change in the Mean with Reduced Type I Errors

Powerful Nonparametric Asymptotic Tests for Change in the Mean with Reduced Type I Errors

Jervis Gallanosa and Yuliya V. Martsynyuk

Department of Statistics,
University of Manitoba,
Winnipeg MB R3T 2N2

The content of this chapter is based on the following publication:

Gallanosa, J.; Martsynyuk, Y. V. (2026). *Powerful Nonparametric Asymptotic Tests for Change in the Mean with Reduced Type I Errors*. **Mathematics** **14**, 78.

<https://doi.org/10.3390/math14010078>

Abstract

We study numerically finite-sample power functions of nonparametric asymptotic tests for at most one change in the mean that are based on convergence in the distribution of sup- and integral functionals of an appropriately weighted and normalized tied-down partial sums process. For each test, a three-way trade-off is observed among its type I errors, power for detecting the change near the beginning or end of the sample, and power for detecting the change in the middle of the sample. By choosing suitable weight functions of a special form, we propose new sup- and integral tests that are shown to be nearly as powerful as the overall most powerful sup-test in the literature, regardless of where the change occurs in the sample. Moreover, the type I errors of the new tests are closer to the asymptotic significance level across various distributions and are lower and converge faster for distributions that are more asymmetric, heavy-tailed, or both.

Keywords: nonparametric asymptotic test for at most one change in the mean; weighted tied-down partial sums process; supremum functional; integral functional; weight function

MSC: 60F17; 60F25; 62G10

1 Introduction

Let $X_1, X_2, \dots, X_n$, $n > 1$, be nondegenerate independent chronologically ordered real-valued observables with finite means. We are to nonparametrically test the “no change in the mean” null hypothesis

$$H_0 : X_1, X_2, \dots, X_n \text{ is a random sample on } X \text{ with } \mu := \mathbb{E}X \text{ and } \sigma^2 := \text{Var}X < \infty. \quad (1)$$

against the “at most one change (AMOC) in the mean” alternative hypothesis

$$H_A : \text{ there is an integer } k^*, 1 \leq k^* < n, \text{ such that} \\ \mathbb{E}X_1 = \dots = \mathbb{E}X_{k^*} \neq \mathbb{E}X_{k^*+1} = \dots = \mathbb{E}X_n, \quad (2)$$

where the time k^* of AMOC in the mean is usually unknown.

1.1 Nonparametric Asymptotic Tests for H_0 vs. H_A Based on Sup-Functionals of an Appropriately Weighted Tied-Down Partial Sums Process

Let $S_k := \sum_{i=1}^k X_i$, $k \geq 1$. To test H_0 versus H_A *nonparametrically*, one can compare the sample mean S_k/k of the first k observables with the sample mean $(S_n - S_k)/(n - k)$ of the last $n - k$ observables for all possible times k of AMOC in the mean, $1 \leq k < n$. Then, when the variance σ^2 is known, using the standardized statistics of the differences in the sample means,

$$\begin{aligned} T_n(k) &:= \left(\frac{1}{k} + \frac{1}{n-k} \right)^{-1/2} \left(\frac{S_k}{k} - \frac{S_n - S_k}{n-k} \right) \\ &= \left(\frac{k}{n} \left(1 - \frac{k}{n} \right) \right)^{-1/2} \frac{S_k - (k/n)S_n}{\sqrt{n}}, \quad 1 \leq k < n, \end{aligned} \quad (3)$$

one may like to reject H_0 in favor of H_A for large observed values of

$$M_n := \frac{1}{\sigma} \max_{1 \leq k < n} |T_n(k)|. \quad (4)$$

Remark 1. If $X_1, \dots, X_n$ are independent normal random variables (r.v.) with a common known variance σ^2 , then for the likelihood ratio statistic Λ_k obtained when the change in the mean occurs at $k^* = k$, we have that

$$-2 \log \Lambda_k = \frac{1}{\sigma^2} (T_n(k))^2 \quad \text{and} \quad \max_{1 \leq k < n} (-2 \log \Lambda_k) = M_n^2. \quad (5)$$

Thus, one may like to reject H_0 in favor of H_A for large observed values of $\max_{1 \leq k < n} (-2 \log \Lambda_k)$ or, equivalently, of M_n . If σ^2 is unknown, then $\max_{1 \leq k < n} (-2 \log \Lambda_k)$ is large whenever

$$M_n^* := \frac{1}{\hat{\sigma}_n} \max_{1 \leq k < n} |T_n(k)| \quad (6)$$

is large, where $\hat{\sigma}_n := \sqrt{\sum_{i=1}^n (X_i - S_n/n)^2 / n}$ (see Section 1.4 of Csörgő and Horváth (1997) and Section 2 of Csörgő and Horváth (1988) for this and a number of other parametric and nonparametric modelings of AMOC in the mean problems that result in the test statistics

M_n and M_n^* .)

In order to proceed with finding nonparametric asymptotic tests for H_0 vs. H_A that would be based on M_n^* , one needs to find the asymptotic distribution of M_n^* under H_0 .

Let $S_0 := 0$, and for $n \geq 1$, consider a tied-down partial sums process on $[0, 1]$

$$Z_n(t) := \begin{cases} \left(S_{[(n+1)t]} - \frac{[(n+1)t]}{n} S_n \right) / \sqrt{n}, & 0 \leq t < 1, \\ 0, & t = 1. \end{cases} \quad (7)$$

Clearly,

$$M_n^* = \frac{1}{\widehat{\sigma}_n} \max_{1 \leq k \leq n} |T_n(k)| = \frac{1}{\widehat{\sigma}_n} \sup_{0 < t < 1} \frac{|Z_n(t)|}{\left(\frac{[(n+1)t]}{n} \left(1 - \frac{[(n+1)t]}{n} \right) \right)^{1/2}}, \quad (8)$$

where $\frac{|Z_n(t)|}{\left(\frac{[(n+1)t]}{n} \left(1 - \frac{[(n+1)t]}{n} \right) \right)^{1/2}} := 0$ for $t \in (0, \frac{1}{n+1}) \cup [\frac{n}{n+1}, 1)$. Thus, the asymptotic behavior of M_n^* under H_0 can be investigated via that of the sup-functional of the standardized process $|Z_n(t)|$ as in (8).

For $0 < a < b < 1$, let $D([a, b], \mathbb{R})$ be the space of real-valued functions on $[a, b]$ that are right continuous and have left-hand limits, $\mathcal{D}$ be the σ -field of subsets of $D([a, b], \mathbb{R})$ that is generated by the finite-dimensional subsets of $D([a, b], \mathbb{R})$, and ρ be the sup-norm metric in $D([a, b], \mathbb{R})$. Donsker's theorem on $(D([a, b], \mathbb{R}), \mathcal{D}, \rho)$ given H_0 yields that, as $n \rightarrow \infty$,

$$\frac{1}{\widehat{\sigma}_n} \sup_{a \leq t \leq b} \frac{|Z_n(t)|}{(t(1-t))^{1/2}} \xrightarrow{\mathcal{D}} \sup_{a \leq t \leq b} \frac{|B(t)|}{(t(1-t))^{1/2}}, \quad (9)$$

where $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge (cf. Sections 3.3 and 3.4 of Csörgő (2002) and Section 2 of Csörgő et al. (2004) for more details on Donsker's theorem). However, (9) is not a satisfactory solution for testing H_0 vs. H_A since the arbitrariness of choosing the interval (a, b) over which the supremum is taken may exclude observing possible changes in the mean that occur “on the tails”, i.e., with k^* near 1, corresponding to $t \in (0, a)$, or k^* near $n - 1$, corresponding to $t \in (b, 1)$. Moreover, one can show that, in contrast to (9), the r.v.

$$\frac{1}{\widehat{\sigma}_n} \sup_{0 < t < 1} \frac{|Z_n(t)|}{(t(1-t))^{1/2}}$$

and hence also M_n^* in (6) converge in distribution to ∞ as $n \rightarrow \infty$, under H_0 . One solution

to secure a nondegenerate limiting distribution while keeping $\sup_{0 < t < 1}$ is by reweighing $|Z_n(t)|$ with an appropriate nonrandom function $q(t)$ as in (12) below.

Let

$$Q \text{ be the class of positive functions } q(t) : (0, 1) \rightarrow (0, \infty), \text{ i.e., } \inf_{\delta \leq t \leq 1-\delta} q(t) > 0 \quad (10)$$

for all $\delta \in (0, 1/2)$, which are nondecreasing near 0 and nonincreasing near 1,

and put

$$I_0^1(q, c) := \int_0^1 \frac{e^{-cq^2(t)/(t(1-t))}}{t(1-t)} dt, \quad c > 0. \quad (11)$$

From Theorem 2.1.1 of Csörgő and Horváth (1997), which is a special case of Theorem 2.1 of Szyszkowicz (1991), on assuming that $\{X, X_i, i \geq 1\}$ are i.i.d. r.v.s with $0 < \text{Var}X = \sigma^2 < \infty$ and that $I_0^1(q, c) < \infty$ for $q(t)$ from (10) and some, or possibly all, $c > 0$, as $n \rightarrow \infty$,

$$M_n^q := \frac{1}{\widehat{\sigma}_n} \sup_{0 < t < 1} \frac{|Z_n(t)|}{q\left(\frac{[(n+1)t]}{n}\right)} = \frac{1}{\widehat{\sigma}_n} \max_{1 \leq k \leq n} \frac{\left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)^{1/2}}{q\left(\frac{k}{n}\right)} |T_n(k)| \xrightarrow{\mathcal{D}} \sup_{0 < t < 1} \frac{|B(t)|}{q(t)}. \quad (12)$$

Remark 2. For $q(t) \in Q$,

$$I_0^1(q, c) < \infty \text{ for } \underline{\text{some}} \ c > 0 \text{ if and only if } \limsup_{t \downarrow 0} \frac{|B(t)|}{q(t)} < \infty \text{ and } \limsup_{t \uparrow 1} \frac{|B(t)|}{q(t)} < \infty \text{ a.s.}, \quad (13)$$

while

$$I_0^1(q, c) < \infty \text{ for } \underline{\text{all}} \ c > 0 \text{ if and only if } \lim_{t \downarrow 0} \frac{|B(t)|}{q(t)} = \lim_{t \uparrow 1} \frac{|B(t)|}{q(t)} = 0 \text{ a.s.} \quad (14)$$

Equivalences of (13) and (14) are contained in Theorems 3.3 and 3.4 of Csörgő et al. (1986).

The functions $q(t) \in Q$ that satisfy (13) are called Erdős–Feller–Kolmogorov–Petrovski (EFKP) upper-class functions of a Brownian bridge. A prototypical EFKP function is

$$q_1(t) := \sqrt{t(1-t) \log \log(t(1-t))}^{-1}, \quad (15)$$

where, hereafter, $\log(x)$ is defined as the natural logarithm of $\max\{e, x\}$. On the other hand, those functions $q(t) \in Q$ that satisfy (14) are the so-called Chibisov–O’Reilly functions.

One can obtain examples of Chibisov–O’Reilly functions in Q by simply multiplying $q_1(t)$ in (15) with an appropriate positive function $m(t)$ on $(0, 1)$, so that $\lim_{t \downarrow 0, t \uparrow 1} m(t) = \infty$ and $q(t) = q_1(t)m(t)$ remain nondecreasing near 0 and nonincreasing near 1. In this regard, in our study, we consider weight functions $q(t)$ of the form

$$q_{(\beta)}(t) := \left[\log(t(1-t))^{-1} \right]^\beta, \quad \beta \leq 0, \quad (16)$$

where $q_{(\beta)}(t) \in Q$ and satisfies (14).

When testing H_0 vs. H_A via (12), one rejects H_0 for large observed values of the statistic M_n^q , where the critical values are obtained from the tabulated distribution function of the r.v. $\sup_{0 < t < 1} |B(t)|/q(t)$.

Remark 3. The distribution of the r.v. $\sup_{0 < t < 1} |B(t)|/q(t)$ with $q(t) = q_1(t)$ was first tabulated in Eastwood and Eastwood (1998). The distribution of $\sup_{0 < t < 1} |B(t)|/q(t)$ with other $q(t)$ can be tabulated by approximating it with the empirical distribution of M_n^q that is based on random samples of a large size n drawn from any distribution with a finite positive variance.

The convergence in distribution in (12) also holds true with a trivial Chibisov–O’Reilly function $q(t) \equiv 1$ on $(0, 1)$, where the distribution of the limiting r.v. $\sup_{0 \leq t \leq 1} |B(t)|$ was first tabulated in Smirnov (1948). The test based on the corresponding *unweighted* test statistic M_n^1 is known to be powerful in detecting changes that occur in the middle of the sample, i.e., with k^* near $[n/2]$, but insensitive to detecting changes that occur on the tails of the sample with k^* close to 1 or $n - 1$ (cf. Csörgő and Horváth (1988)). On the other hand, using the tests based on the *weighted* M_n^q statistics may overcome this problem (cf. Csörgő and Horváth (1988, 1997)). However, little research has been conducted numerically to investigate the extent to which these weighted sup-tests may have better power than their unweighted counterpart.

The work of Orasch and Pouliot (2004) was one of the first to conduct a detailed numerical comparison of the empirical power of the tests based on the M_n^q statistics, weighted with some EFKP upper-class functions of a Brownian bridge, versus the M_n^1 test. Among the

M_n^q tests the authors of Orasch and Pouliot (2004) studied, they found that the overall most powerful one in detecting changes both on the tails and in the middle of the sample, which also outperforms the M_n^1 and $M_n^{q_1}$ tests, is the $M_n^{q_2}$ test with one of their proposed weight functions, a modification of the prototypical EFKP upper-class function of a Brownian bridge $q_1(t)$ of (15), as follows:

$$q_2(t) := \begin{cases} (t(1-t))^{1/2}, & \text{if } t \in (a, b) = (0.071, 0.929), \\ q_1(t), & \text{if } t \in (0, a] \cup [b, 1). \end{cases} \quad (17)$$

The constants a and b in (17) were found so that the quadratic inequality $q_1(t) < (t(1-t))^{1/2}$ holds true for $a < t < b$.

1.2 Nonparametric Asymptotic Tests for H_0 vs. H_A Based on Integral Functionals of an Appropriately Weighted Tied-Down Partial Sums Process

In Section 1.1, we discussed sup-functionals of an appropriately weighted $|Z_n(t)|$ as test statistics for H_0 vs. H_1 or, equivalently, appropriately weighted *maximums* in k of the absolute values of the $T_n(k)$ statistics in (3), which originate from an intuitive idea of comparing the sample mean S_k/k before a possible time k of AMOC in the mean to the sample mean $(S_n - S_k)/(n - k)$ after time k , for all $1 \leq k < n$.

Another natural test statistic for H_0 vs. H_1 can be obtained from employing an appropriately weighted *average* of $T_n^2(k)$ in k or, equivalently, the integral of an appropriately weighted $Z_n^2(t)$. From (ii) of Proposition 2.1.1 of Csörgő and Horváth (1997) proved via an L_2 -approximation in probability, on assuming that $\{X, X_i, i \geq 1\}$ are i.i.d. r.v.s with a finite positive variance σ^2 , as $n \rightarrow \infty$,

$$\begin{aligned} L_n^q &:= \frac{1}{\widehat{\sigma}_n^2} \int_0^1 \frac{Z_n^2(t)}{q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt = \frac{1}{n(n+1)\widehat{\sigma}_n^2} \sum_{k=1}^{n-1} \frac{(S_k - \frac{k}{n}S_n)^2}{q\left(\frac{k}{n}\right)} \\ &= \frac{1}{(n+1)\widehat{\sigma}_n^2} \sum_{k=1}^{n-1} \frac{\frac{k}{n}(1 - \frac{k}{n})}{q\left(\frac{k}{n}\right)} T_n^2(k) \xrightarrow{\mathcal{D}} \int_0^1 \frac{B^2(t)}{q(t)} dt, \end{aligned} \quad (18)$$

provided that $q(t) \in \mathcal{Q}$ satisfies

$$\int_0^1 \frac{t(1-t)}{q(t)} dt < \infty. \quad (19)$$

When testing H_0 vs. H_A via (18), one rejects H_0 for large observed values of the left-hand side r.v. L_n^q . Unlike for the test based on M_n^q , the weight function $q(t) = (t(1-t))^{1/2}$ is an immediate candidate for the L_n^q test due to the finiteness of the integral in (19). The distribution function of the limiting r.v. $\int_0^1 B^2(t) dt$ was tabulated in Anderson and Darling (1952), but tables for the distribution function of $\int_0^1 B^2(t)/q(t) dt$ with a weight function $q(t)$ of (16) used in this work seem to be unavailable. However, one can tabulate the distribution function of the latter r.v. the same way as that of $\sup_{0 < t < 1} |B(t)|/q(t)$ described in Remark 3, via the empirical distribution of the r.v. L_n^q on the left-hand side of (18). It appears that the finite-sample properties of the L_n^q -based tests have not been studied in the literature.

1.3 Overview of the Present Work

In Section 2 of the present work, we conduct an extensive numerical study of the finite-sample performance of the nonparametric tests for H_0 vs. H_A for AMOC in the mean at the 5% asymptotic significance level that are based on the sup- and integral functionals of the appropriately weighted and normalized tied-down partial sums process $Z_n(t)$ of (7), i.e., the tests based on M_n^q in (12) and L_n^q in (18), with some newly proposed weight functions $q(t) = q_{(\beta)}(t)$ of the form in (16) and some previously used $q(t)$ satisfying the integral conditions in (13)/(14) and (19), respectively. For a variety of distributions with finite variances and different degrees of asymmetry and heavy-tailedness, we investigate the empirical type I errors and powers of detecting the change in the mean of the M_n^q and L_n^q tests, where the change occurs due to a shift in the sample, considering various sizes and times of the change and different sample sizes, all specified in Section 2.1 with preliminaries.

In Section 2.2, we propose new sup- and integral tests: the $M_n^{q(-0.9)}$ test with the Chibisov–O’Reilly weight function $q_{(\beta)}(t) = q_{(-0.9)}(t)$ of (16) and the $L_n^{q(-4)}$ test, where $q_{(-4)}(t)$ can be seen to satisfy the integral condition in (19). We also consider the un-

weighted M_n^1 and L_n^1 tests (with $q(t) \equiv 1$), as well as the $M_n^{q_2}$ test with $q_2(t)$ of (17). We compare the sup- and integral tests among each other in order to obtain one of the most powerful and well-balanced tests, regardless of where the change occurs in the sample, which also has well-controlled type I errors across a wide range of distributions.

In Section 2.2.1, our detailed analysis of the empirical type I errors of the sup-tests reveals that they seem to converge overall slowly to the nominal significance level of 5% and that the type I errors of the *weighted* M_n^q tests (with $q(t) \neq 1$) are generally larger for the samples from distributions that are more asymmetric, heavy-tailed, or both, while the trend is generally the opposite for the *unweighted* M_n^1 and L_n^1 tests. Thus, the type I errors of the $M_n^{q_2}$ and $M_n^{q(-0.9)}$ tests may be undesirably larger than 5%, even for rather large sample sizes, but the type I errors of our proposed $M_n^{q(-0.9)}$ and $L_n^{q(-4)}$ tests are overall closer to 5% than those of the $M_n^{q_2}$ test. Notably, for more asymmetric and/or heavy-tailed distributions, the type I errors of the $L_n^{q(-4)}$ test are lower, and approach 5% faster, than the type I errors of the $M_n^{q(-0.9)}$ test.

In Section 2.2.2, we perform a comparative analysis of the empirical powers of the sup- and integral tests in hand, prior to observing that all the tests become more powerful with an increase in the sample size and change size, and with the time of change k^* getting closer towards the middle of the sample, where all the powers for $k^* = 0.50n$ are close to 100% already for $n = 500$. Thus, the sup- and integral tests that are more sensitive to the changes close to the beginning or end of the sample, namely, the *weighted* M_n^q and L_n^q tests, are more desirable. However, for each of the tests, we observe a three-way trade-off among its type I errors, power for detecting the change on the tails of the sample, and power for detecting the change in the middle of the sample. We find that this trade-off may be significant for the M_n^q tests, especially for the samples from the distributions that are more asymmetric, heavy-tailed, or both. Powerful M_n^q tests for k^* near 1 or $n - 1$ may have undesirably large, slowly convergent type I errors and compromised power for k^* close to $[n/2]$. However, due to the suitable selections of β for the weight function $q(t) = q_{(\beta)}(t)$ explained in Remark 6, our proposed $M_n^{q(-0.9)}$ and $L_n^{q(-4)}$ tests have the powers comparable to that of the overall most powerful $M_n^{q_2}$ sup-test previously studied in the literature, both for changes on the tails and in the middle of the sample, while the type I errors of the new tests are overall

closer to 5%. Moreover, for the distributions that are more asymmetric, heavy-tailed, or both, the type I errors of the $M_n^{q(-0.9)}$ and $L_n^{q(-4)}$ tests are lower, and approach 5% faster, than those of the $M_n^{q_2}$ test, more so for the $L_n^{q(-4)}$ test. All in all, the $L_n^{q(-4)}$ integral test emerges as one of the most powerful and well-balanced nonparametric asymptotic tests for H_0 vs. H_A examined in this work, exhibiting consistently high power regardless of where the change occurs in the sample, while also maintaining the best controlled type I errors among the weighted sup- and integral tests across a wide range of distributions.

Other variants of the M_n^q and L_n^q tests and some related tests are considered at the end of Section 2.2, followed by Section 3 with the paper conclusions.

2 Main Results

2.1 Preliminaries

In order to compute the empirical power functions of the M_n^q sup-tests and L_n^q integral tests, we will first need to obtain the quantiles of $\sup_{0 < t < 1} (|B(t)|/q(t))$ and $\int_0^1 (B^2(t)/q(t)) dt$, the respective limiting r.v.s in (12) and (18). The distributions of these r.v.s have only been tabulated or available in closed-form expressions for $\sup_{0 < t < 1} (|B(t)|/q(t))$ with $q(t) \equiv 1$ and $q(t) = q_2(t)$ of (17), and for $\int_0^1 B^2(t)dt$. We will tabulate the quantiles of the distributions of $\sup_{0 < t < 1} (|B(t)|/q(t))$ and $\int_0^1 (B^2(t)/q(t)) dt$, as described in Remark 3, by approximating them with the quantiles of the respective empirical distribution functions of the r.v.s on the left-hand sides of (12) and (18), where the latter functions are computed using $m = 100,000$ simulated values that are each computed from an independent Normal(0, 1) random sample of size $n = 100,000$. The procedure is as follows:

1. For each replication $j = 1, \dots, m$, generate i.i.d. $N(0, 1)$ r.v.s $X_1^{(j)}, \dots, X_n^{(j)}$ and define the partial sums

$$S_k^{(j)} = \sum_{i=1}^k X_i^{(j)}, \quad k = 1, \dots, n.$$

2. For each j , compute the sup-type statistic

$$(M_n^q)^{(j)} = \frac{1}{\sqrt{n} \hat{\sigma}_n^{(j)}} \max_{1 \leq k < n} \frac{S_k^{(j)} - (k/n) S_n^{(j)}}{q(\frac{k}{n})},$$

where $\hat{\sigma}_n^{(j)} = \sqrt{\sum_{i=1}^n \left(X_i^{(j)} - S_n^{(j)}/n\right)^2 / n}$.

3. To obtain the quantiles of $\sup_{0 < t < 1} |B(t)|/q(t)$, use the empirical distribution function of M_n^q given by

$$G_{n,m}(x) = \frac{1}{m} \sum_{j=1}^m \mathbb{1}_{\{(M_n^q)^{(j)} \leq x\}},$$

which approximates the limiting distribution function

$$G(x) = P\left(\sup_{0 < t < 1} \frac{|B(t)|}{q(t)} \leq x\right).$$

4. To obtain the quantiles of $\int_0^1 (B^2(t)/q(t)) dt$, apply steps 1–3 with $(M_n^q)^{(j)}$ and $G_{n,m}(x)$ replaced respectively with

$$(L_n^q)^{(j)} = \frac{1}{n(n+1) (\hat{\sigma}_n^{(j)})^2} \sum_{k=1}^{n-1} \frac{\left(S_k^{(j)} - \frac{k}{n} S_n^{(j)}\right)^2}{q\left(\frac{k}{n}\right)}$$

and

$$H_{n,m}(x) = \frac{1}{m} \sum_{j=1}^m \mathbb{1}_{\{(L_n^q)^{(j)} \leq x\}},$$

where the latter approximates

$$H(x) = P\left(\int_0^1 \frac{B^2(t)}{q(t)} dt \leq x\right).$$

For sufficiently large m and n , the closeness of $G_{n,m}(x)$ and $H_{n,m}(x)$ to $G(x)$ and $H(x)$ can be seen via Proposition 3.1 of Orasch and Pouliot (2004).

Hereafter, we assume that a possible change in the mean in a sample is due to a shift, thus replacing the hypotheses in (1) and (2) with the following ones:

$$H'_0 : X_1, \dots, X_n \text{ are i.i.d. r.v.s with a distr. func. } F(x), \text{ mean } \mu, \text{ and variance } \sigma^2 < \infty, \quad (20)$$

and

H'_A : there exists an integer k^* , $1 \leq k^* < n$, such that $X_1, \dots, X_{k^*}$ have a distribution function $F(x)$ with a mean μ , while independent $X_{k^*+1}, \dots, X_n$ (21) have a distribution function $F(x - \Delta)$ with a mean $\mu + \Delta$, for some $\Delta > 0$,

where Δ is regarded throughout this paper as the size of the change in the mean.

In our simulation study, we use ten distributions with finite variances and different numbers of moments and skewness (if applicable) listed in Table 1. We note that SymPareto(2.1,1) is a symmetric counterpart of Pareto(2.1,1) with the probability density function (PDF) of $\frac{1}{2}f(-x)\mathbb{1}_{(-\infty,-1)} + \frac{1}{2}f(x)\mathbb{1}_{(1,\infty)}$, where $f(x)$ is the PDF of Pareto(2.1,1), while Half- $t(3)$ is a one-sided version of the $t(3)$ distribution with two times the PDF of $t(3)$ over just $x \geq 0$.

In Table 1, the distributions for an r.v. X with a mean μ and finite positive variance are listed according to their respective values of δ_{100} , where

$$\begin{aligned} \delta_n &:= nP(|Y| > b_n) + nb_n^{-1} |\mathbb{E}(Y \mathbb{1}_{\{|Y| \leq b_n\}})| + nb_n^{-3} \mathbb{E}(|Y|^3 \mathbb{1}_{\{|Y| \leq b_n\}}), \\ \text{with } Y &:= X - \mu \quad \text{and} \quad b_n := \sup\{c : nc^{-2} \mathbb{E}(Y^2 \mathbb{1}_{\{|Y| \leq c\}}) \geq 1\}, \end{aligned} \quad (22)$$

where b_{100} used for δ_{100} has been computed numerically. Since $0 < \text{Var}X < \infty$, for i.i.d. r.v.s $\{X, X_i, i \geq 1\}$, the following holds true:

$$\sup_x \left| P \left(\frac{\sum_{i=1}^n (X_i - \mathbb{E}X)}{(\sum_{i=1}^n (X_i - \mathbb{E}X)^2)^{1/2}} \leq x \right) - P(N(0,1) \leq x) \right| \leq A\delta_n, \quad (23)$$

where A is an absolute constant and $\delta_n \rightarrow 0$, as $n \rightarrow \infty$ (see Bentkus and Götze (1996)). Thus, $A\delta_n$ is both a Berry–Esseen upper bound and a uniform upper bound for the rate of convergence in the self-normalized CLT. We note that the ordering of the distributions according to δ_n computed based on the other considered sample sizes, i.e., when $n = 200$ and $n = 500$, varies insignificantly from that based on δ_{100} . A special role of δ_{100} values in our study of the sup- and integral tests will soon be revealed in Section 2.2.1.

Table 1: Distributions with their skewness (if exists), δ_{100} , and probability density functions.

Distribution	Skewness	δ_{100}	Probability Density Function
Normal(0, 1)	0	0.16	$\frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \quad x \in \mathbb{R}$
Exponential(1)	2	0.24	$e^{-x}, \quad x > 0$
$t(3)$	NA	0.33	$\frac{2}{\pi\sqrt{3}} \left(1 + \frac{x^2}{3}\right)^{-2}, \quad x \in \mathbb{R}$
SymPareto(2.1, 1)	NA	0.41	$\frac{2.1}{2 x ^{3.1}}, \quad x \geq 1$
Half- $t(3)$	∞	0.60	$\frac{1}{\pi\sqrt{3}} \left(1 + \frac{x^2}{3}\right)^{-2}, \quad x \geq 0$
Frechet(2.5)	∞	0.97	$2.5x^{-3.5}e^{-x^{-2.5}}, \quad x > 0$
Inverse-Normal(7, 1)	7.94	1.25	$(2\pi x^3)^{-0.5}e^{-(x-7)^2/(98x)}, \quad x > 0$
Pareto(2.1, 1)	∞	1.46	$\frac{2.1}{x^{3.1}}, \quad x \geq 1$
Weibull(0.4, 1)	11.35	1.47	$0.4x^{-0.6}e^{-x^{0.4}}, \quad x > 0$
Lognormal(0, 1.5)	33.47	1.56	$(1.5x\sqrt{2\pi})^{-1}e^{-\ln^2(x)/4.5}, \quad x > 0$

Our investigation of the finite-sample performance of the M_n^q sup-tests and L_n^q integral tests at the $\alpha = 5\%$ significance level in Section 2.2 consists of studying first their empirical type I errors in Section 2.2.1 and then their empirical powers of detecting a change in the mean in Section 2.2.2.

For each of the ten distributions in Table 1, the empirical type I errors of the M_n^q and L_n^q tests are computed as the proportions of $m = 5000$ independent random samples of size n that lead to rejection of H'_0 , where the rejection regions are based on the M_n^q and L_n^q statistics exceeding the corresponding approximate $1 - \alpha$ quantiles of $\sup_{0 < t < 1} (|B(t)|/q(t))$ and $\int_0^1 (B^2(t)/q(t)) dt$. These type I errors are later reported for sample sizes of $n = 100$, 200, and 500. Due to the slow convergence rates in (12) and (18), we also examine type I errors for much larger sample sizes, up to $n = 10,000$, with one of our figures.

To obtain the empirical powers of the M_n^q - and L_n^q -based tests for a given distribution function $F(x)$ as in H'_0 , sample size n , change point k^* , and change size Δ , we estimate the probabilities of rejecting H'_0 using the following procedure with $m = 5000$.

1. For each replication $j = 1, \dots, m$, generate i.i.d. r.v.s $X_1^{(j)}, \dots, X_n^{(j)}$ from $F(x)$ and introduce the change by replacing $X_{k^*+1}^{(j)}, \dots, X_n^{(j)}$ with $X_{k^*+1}^{(j)} + \Delta, \dots, X_n^{(j)} + \Delta$.

2. Compute the statistics $(M_n^q)^{(j)}$ and $(L_n^q)^{(j)}$ defined in the quantile computing procedure for all $j = 1, \dots, m$.
3. Among the m replications, compute the proportions of times when $(M_n^q)^{(j)}$ and $(L_n^q)^{(j)}$ exceed the corresponding $1 - \alpha$ critical values.

We note that the above procedure can also be used for computing the empirical type I errors of the M_n^q and L_n^q tests when setting $\Delta = 0$.

When the empirical powers are viewed as functions of the relative change-point location k^*/n , they are typically nearly symmetric about $k^*/n = 0.50$, as will be illustrated in the results to follow. Consequently, it is sufficient to focus on change points in the first half of the sample. In our study, we therefore consider the four locations $k^* = 0.05n, 0.15n, 0.25n$, and $0.50n$ for sample sizes $n = 100, 200$, and 500 . From this point forward, we say that a change "on the tails" occurs if $k^* = 0.05n$ or $0.15n$, while $k^* = 0.25n$ and $0.50n$ represent changes "in the middle".

We consider two representative shift magnitudes, $\Delta = 0.50\sigma$ and 0.75σ in our tables, where σ is the true standard deviation of the underlying distribution. In addition, with one of our figures, we examine how empirical powers of the M_n^q and L_n^q tests behave as functions of Δ over a wider range of Δ , $[0.50\sigma, 1\sigma]$.

Due to similar behavior of the empirical powers for several of the distributions, it is sufficient to compare the tests using a few representative distributions, such as Normal(0, 1), SymPareto(2.1, 1), Weibull(0.4, 1), and Lognormal(0, 1.5), with additional comparison tables for the remaining distributions included in Appendix 3.

2.2 Finite-Sample Performance of the M_n^q - and L_n^q -Based Tests

In this section, we investigate the finite-sample performance of our newly proposed $M_n^{q(-0.9)}$ and $L_n^{q(-4)}$ tests, as well as the previously considered M_n^1 , M_n^{q2} , and L_n^1 tests.

Table 2 contains the tabulated 0.90, 0.95, and 0.98 quantiles of $\sup_{0 < t < 1} (|B(t)|/q(t))$ and $\int_0^1 (B^2(t)/q(t))dt$. The quantiles of $\sup_{0 < t < 1} (|B(t)|/q(t))$ with $q(t) \equiv 1$ and $q(t) = q_2(t)$ and the ones for $\int_0^1 B^2(t)dt$ were retabulated and correspond closely to those produced by Smirnov (1948), Orasch and Pouliot (2004), and Anderson and Darling (1952), respectively.

The quantiles of $\sup_{0 < t < 1} (|B(t)|/q_{(-0.9)}(t))$ and $\int_0^1 (B^2(t)/q_{(-4)}(t))dt$ were not available before.

Table 2: λ quantiles of $\sup_{0 < t < 1} \frac{|B(t)|}{q(t)}$ and $\int_0^1 \frac{B^2(t)}{q(t)} dt$.

λ quantile of $\sup_{0 < t < 1} \frac{ B(t) }{q(t)}$	$\lambda = 0.90$	$\lambda = 0.95$	$\lambda = 0.98$
$q(t) \equiv 1$	1.2209	1.3586	1.5171
$q(t) = q_{(-0.9)}(t)$	1.9277	2.1044	2.3119
$q(t) = q_2(t)$	2.9186	3.1702	3.4629
λ quantile of $\int_0^1 \frac{B^2(t)}{q(t)} dt$	$\lambda = 0.90$	$\lambda = 0.95$	$\lambda = 0.98$
$q(t) \equiv 1$	0.3478	0.4624	0.6248
$q(t) = q_{(-4)}(t)$	3.5776	4.4467	5.6269

2.2.1 Empirical Type I Errors of the M_n^q - and L_n^q -Based Tests

Table 3 contains the empirical type I errors of the M_n^q and L_n^q tests with the weight functions as in Table 2 for the sample sizes of $n = 100, 200$, and 500 and the 5% asymptotic significance level. The type I errors are tabulated for all the ten distributions listed in Table 1, where, throughout all tables, the distributions are always ordered by their corresponding values of δ_{100} as in Table 1.

Table 3 shows that the empirical type I errors of the *weighted* M_n^q and L_n^q tests (with $q(t) \neq 1$) are generally larger for the samples from the distributions with larger corresponding values of δ_{100} . However, the trend is generally the opposite for the test based on the *unweighted* M_n^1 and L_n^1 statistics, where the type I errors are close to, or lower than, 5% and generally decrease as δ_{100} increases.

Table 3: The empirical type I errors (%) of the M_n^q and L_n^q tests at the $\alpha = 5\%$ asymptotic significance level.

Distribution	Sup-Tests			Integral Tests	
	M_n^1	$M_n^{q(-0.9)}$	M_n^{q2}	L_n^1	$L_n^{q(-4)}$
$n = 100$					
Normal(0, 1)	3.9	2.7	2.3	4.4	3.9
Exponential(1)	3.3	5.0	6.0	5.1	5.7
$t(3)$	2.9	4.3	5.1	4.6	5.6
SymPareto(2.1, 1)	2.8	5.1	6.4	4.5	6.2
Half- $t(3)$	3.0	6.7	8.4	4.7	6.9
Frechet(2.5)	2.6	8.3	10.0	4.6	7.6
Inverse-Normal(7, 1)	1.7	9.8	11.9	4.1	7.5
Pareto(2.1, 1)	1.7	8.6	10.6	3.8	7.1
Weibull(0.4, 1)	1.8	9.7	12.1	3.5	7.5
Lognormal(0, 1.5)	1.6	9.3	10.8	3.8	7.0
$n = 200$					
Normal(0, 1)	3.7	3.4	3.2	5.1	5.0
Exponential(1)	3.6	4.2	5.9	4.7	4.9
$t(3)$	3.1	4.3	6.3	4.7	5.5
SymPareto(2.1, 1)	2.8	5.7	7.4	4.4	6.2
Half- $t(3)$	3.0	6.5	8.5	4.4	6.1
Frechet(2.5)	2.8	8.4	10.8	4.3	7.1
Inverse-Normal(7, 1)	2.8	9.2	12.6	4.7	7.6
Pareto(2.1, 1)	2.1	8.7	11.4	3.9	7.0
Weibull(0.4, 1)	2.2	9.0	11.8	4.0	6.8
Lognormal(0, 1.5)	2.1	10.0	13.0	3.9	7.6
$n = 500$					
Normal(0, 1)	4.3	3.8	3.8	4.7	4.7
Exponential(1)	4.2	4.3	6.4	4.5	4.7
$t(3)$	3.7	4.7	6.5	4.9	5.4
SymPareto(2.1, 1)	3.3	6.5	8.7	5.4	6.4
Half- $t(3)$	3.9	6.8	10.1	4.6	6.1
Frechet(2.5)	3.1	7.4	10.7	4.4	5.9
Inverse-Normal(7, 1)	3.1	7.6	12.3	4.7	6.5
Pareto(2.1, 1)	3.0	9.2	13.1	3.9	7.1
Weibull(0.4, 1)	2.7	8.5	12.7	4.2	6.8
Lognormal(0, 1.5)	2.9	9.1	13.7	4.4	7.4

The rate of convergence in the self-normalized CLT is known to be generally faster for symmetric distributions and slower for asymmetric distributions, especially for those with fewer moments, larger skewness, or both (see, for example, Bentkus and Götze (1996) and references therein). More precisely, as seen from (23), a uniform upper bound for this rate is a multiple of δ_n , which can be viewed as a single general estimator of asymmetry and heavy-tailedness of the underlying distribution. Interestingly, the rates of convergence of the empirical type I errors of the *weighted* tests based on $M_n^{q(-0.9)}$, M_n^{q2} , and $L_n^{q(-4)}$ mirror this behavior. Thus, as seen from Table 3, the type I errors for the symmetric distributions

$t(3)$ and $\text{SymPareto}(2.1,1)$ are closer to 5% and lower compared with their asymmetric counterparts $\text{Half-}t(3)$ and $\text{Pareto}(2.1,1)$. Moreover, the distributions with fewer moments often exhibit larger empirical type I errors, both among the symmetric ($t(3)$ and $\text{SymPareto}(2.1,1)$) and asymmetric distributions ($\text{Half-}t(3)$, $\text{Frechet}(2.5)$, and $\text{Pareto}(2.1,1)$). Additionally, for the distributions with a finite skewness, such as $\text{Normal}(0,1)$, $\text{Exponential}(1)$, $\text{Inverse-Normal}(7,1)$, $\text{Weibull}(0.4,1)$, and $\text{Lognormal}(0,1.5)$, the type I errors of the weighted tests at $n = 500$ are ordered according to their skewness values. Furthermore, we find that the type I errors of the *weighted* tests based on $M_n^{q(-0.9)}$, $M_n^{q_2}$, and $L_n^{q(-4)}$ follow generally similar orderings to that of the values of δ_{100} in Table 1. We note that the orderings of δ_n and the empirical type I errors of these weighted tests vary somewhat for the other sample sizes ($n = 200$ and 500) while still maintaining general resemblance to each other.

The type I errors of the $M_n^{q(-0.9)}$ and $M_n^{q_2}$ tests may be undesirably larger than 5% as δ_{100} increases, but the type I errors of the $M_n^{q(-0.9)}$ test are overall closer to 5%. Notably, as seen from Table 3, for the distributions with larger δ_{100} values, the type I errors of the weighted $L_n^{q(-4)}$ test are lower and more controlled than the type I errors of the $M_n^{q(-0.9)}$ test. For example, when $n = 500$, the type I error of the $L_n^{q(-4)}$ test is 7.4% for the $\text{Lognormal}(0,1.5)$ distribution, which has the largest value of δ_{100} , compared with 9.1% and 13.7% for the $M_n^{q(-0.9)}$ and $M_n^{q_2}$ tests.

We complement Table 3 with Figure 1 below and thus illustrate the behavior of the empirical type I errors as n increases up to 10,000.

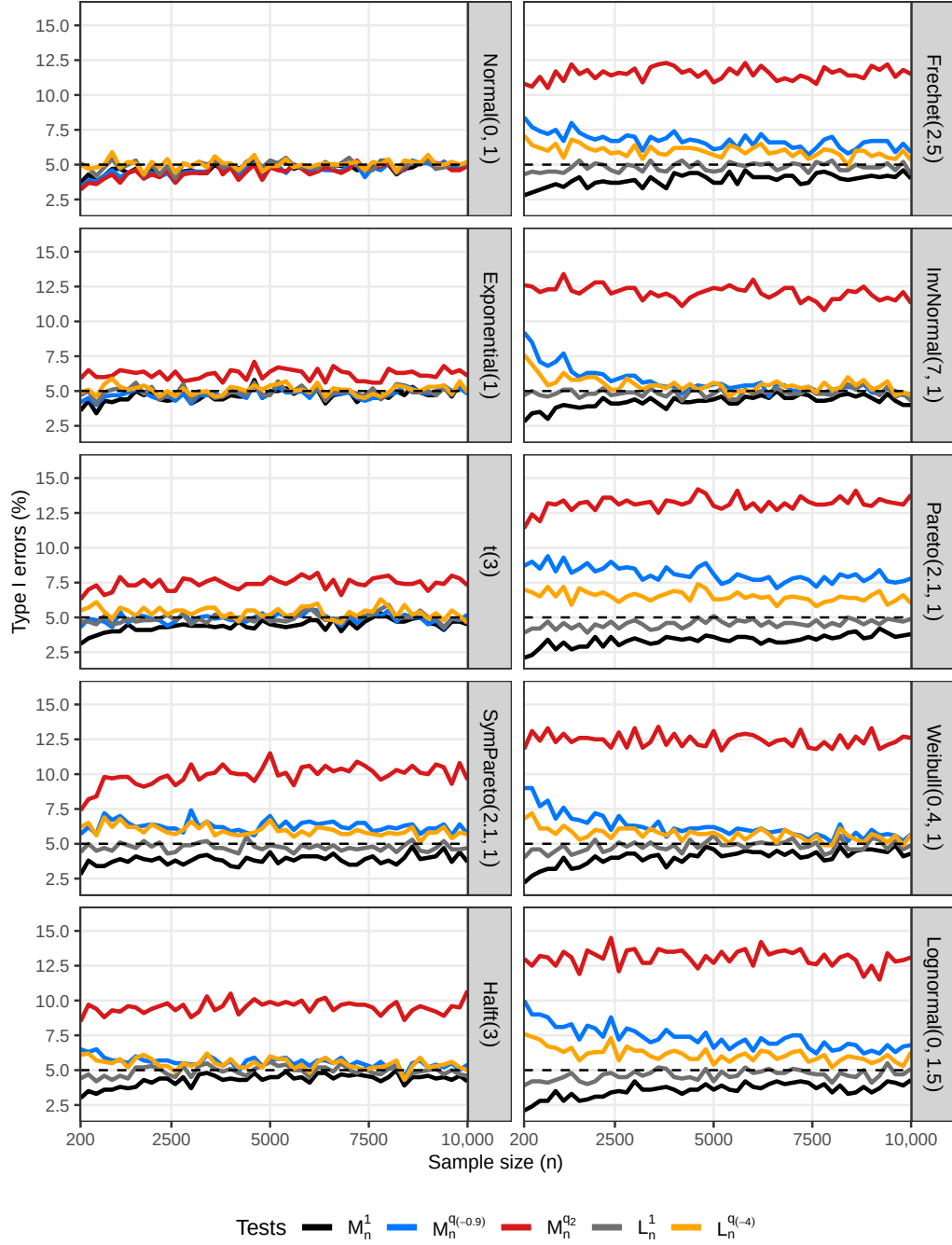


Figure 1: The empirical type I errors of the M_n^q and L_n^q tests for n up to 10,000 (increments of 200).

Figure 1 shows that the empirical type I errors of the weighted $M_n^{q(-0.9)}$ and M_n^{q2} tests are reasonably close to 5% for the samples from the Normal(0,1), Exponential(1), and $t(3)$ distributions. However, for the remaining distributions, the slow convergence of the type I

errors of the $M_n^{q_2}$ test requires a sample size larger than 10,000 for the errors becoming sufficiently close to 5% (at or below 7.5%). On the other hand, the type I errors of the $M_n^{q(-0.9)}$ test become very close to or under 7.5% for $n < 10,000$. Moreover, for the distributions with δ_{100} values larger than that for Exponential(1), the type I errors of the $M_n^{q(-0.9)}$ test are significantly lower and closer to 5% than those of the $M_n^{q_2}$ test. The type I errors of the *unweighted* M_n^1 and L_n^1 tests are under or very close to 5% for all the distributions and mostly all sample sizes smaller than 10,000. The type I errors of the L_n^1 test approach 5% faster than those of the M_n^1 test and *weighted* $L_n^{q(-4)}$ test. Importantly, for the distributions with larger δ_{100} values, the type I errors of the $L_n^{q(-4)}$ integral test approach 5% quicker than those of the $M_n^{q(-0.9)}$ sup-test.

2.2.2 Empirical Powers of the M_n^q - and L_n^q -Based Tests

General Trends of the Empirical Powers of the M_n^q and L_n^q Tests

Before we present our tables with the empirical power values of the M_n^q and L_n^q tests, based on Figure 2, we have the following Remark 4 on an approximate symmetry of the powers as functions of k^*/n about $k^*/n = 0.5$ and on a similarity of the behavior of the powers across the distributions in hand.

Remark 4. *Figure 2 displays the empirical powers of the sup- and integral tests as functions of the change time k^* divided by n for three of our distributions and the change size $\Delta = 0.50\sigma$ at the sample size $n = 200$. The power of each of these tests is nearly symmetric about $k^*/n = 0.50$. This approximate symmetry is analogous for the rest of the distributions considered in this paper, other n , and $\Delta = 0.75\sigma$.*

This allows us to focus our comparative power study only on the performance of the tests when the change occurs in the first half of the sample ($k^/n \leq 0.50$), and only for a few representative distributions such as the two symmetric distributions, Normal(0,1) and SymPareto(2.1,1), and the two skewed and heavy-tailed ones, Weibull(0.4,1) and Lognormal(0,1.5).*

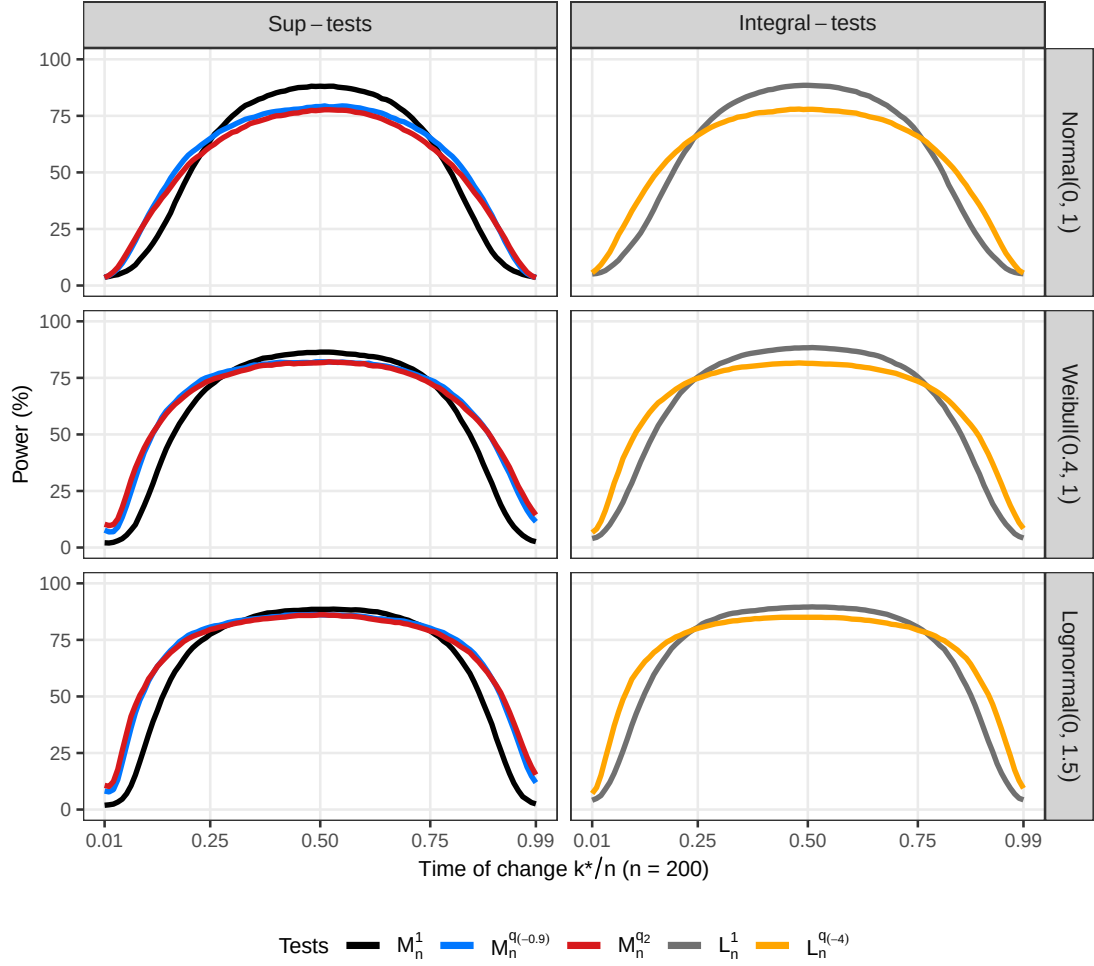


Figure 2: The empirical powers of the M_n^q and L_n^q tests as functions of k^*/n at $n = 200$ and the change size $\Delta = 0.50\sigma$.

Some other general trends of the powers of the M_n^q and L_n^q tests are summarized in Remark 5.

Remark 5. The bell shape of the empirical powers shown in Figure 2 for $\text{Normal}(0,1)$, $\text{Weibull}(0.4,1)$, and $\text{Lognormal}(0,1.5)$ indicates that the closer the change-point time is to the beginning or the end of the sample, the more difficult it is to detect the change. The same holds true for the rest of our ten distributions.

Later, we will show that the power increases with the change size Δ as well as with the sample size, and that it approaches 100% fastest for the changes occurring in the middle of the sample. Thus, for smaller sample sizes, a test may already be very sensitive in detecting

changes in the middle, but not sensitive enough to changes occurring at $k^* < 0.25n$ or $k^* > 0.75n$.

For instance, as shown in Figure 2, the unweighted M_n^1 and L_n^1 tests at $n = 200$ and $\Delta = 0.50\sigma$ are very powerful for changes near $k^* = 0.50n$, but not as powerful as the weighted $M_n^{q(-0.9)}$, M_n^{q2} , and $L_n^{q(-4)}$ tests for changes on the tails.

Comparative Analysis of the Empirical Powers of the M_n^q and L_n^q Tests

Table 4 is a comparison table of the empirical powers at the $\alpha = 5\%$ asymptotic significance level of the tests based on M_n^1 , $M_n^{q(-0.9)}$, M_n^{q2} , L_n^1 , and $L_n^{q(-4)}$ for the change sizes $\Delta = 0.50\sigma$ and 0.75σ ; for changes that occur at $k^* = 0.05n, 0.15n, 0.25n$, and $0.50n$; and for the sample sizes $n = 100, 200$, and 500 . The empirical powers are tabulated for the samples from Normal(0,1), SymPareto(2.1,1), Weibull(0.4,1), and Lognormal(0,1.5) (Tables A1 and A2 from Appendix 3 are the corresponding tables for the remaining six distributions of Table 1). The empirical power tables are complemented with Figure 3 with the average empirical powers over our ten distributions of all the sup- and integral tests as functions of the change size Δ , as Δ varies from 0.5σ to 1σ , for $k^* = 0.05n, 0.15n, 0.25n$, and $0.50n$ and $n = 100, 200$, and 500 .

2. Main Results

Table 4: The empirical powers for Normal(0, 1), SymPareto(2.1, 1), Weibull(0.4, 1), and Lognormal(0, 1.5) for the change sizes $\Delta = 0.50\sigma$ and 0.75σ .

k^*/n	Distribution	$\Delta = 0.50\sigma$					$\Delta = 0.75\sigma$				
		M_n^1	$M_n^{q(-0.9)}$	M_n^{q2}	L_n^1	$L_n^{q(-4)}$	M_n^1	$M_n^{q(-0.9)}$	M_n^{q2}	L_n^1	$L_n^{q(-4)}$
$n = 100$											
0.05	Normal(0, 1)	3.9	4.7	4.8	5.6	8.1	4.6	8.3	9.4	7.4	14.4
	SymPareto(2.1, 1)	5.4	16.7	19.4	10.2	23.9	9.5	40.9	46.9	18.1	48.8
	Weibull(0.4, 1)	2.1	9.6	12.4	7.0	15.6	3.6	20.1	25.6	12.5	30.3
	Lognormal(0, 1.5)	2.5	11.4	14.7	8.6	20.0	4.5	28.4	33.7	14.9	38.9
0.15	Normal(0, 1)	12.5	18.0	16.2	17.6	23.5	28.8	45.4	42.1	36.7	50.3
	SymPareto(2.1, 1)	49.2	66.3	64.3	54.5	68.1	83.0	91.4	90.4	83.9	90.8
	Weibull(0.4, 1)	24.3	39.5	39.1	34.8	44.4	50.4	67.0	65.9	58.0	68.3
	Lognormal(0, 1.5)	34.0	51.2	49.9	43.1	53.2	61.3	76.3	75.1	68.0	77.0
0.25	Normal(0, 1)	30.4	30.4	26.7	36.6	34.9	66.7	67.9	63.6	70.1	68.5
	SymPareto(2.1, 1)	81.1	82.4	80.6	83.2	82.5	94.8	95.7	95.3	95.4	95.4
	Weibull(0.4, 1)	50.4	55.2	53.1	57.4	56.6	75.2	79.0	77.6	79.2	78.5
	Lognormal(0, 1.5)	61.4	65.5	63.4	67.1	65.7	83.0	85.6	84.5	86.0	85.4
0.50	Normal(0, 1)	54.7	40.5	37.6	59.6	44.4	90.0	81.1	79.8	91.2	79.7
	SymPareto(2.1, 1)	92.2	89.3	88.9	93.2	88.6	97.5	97.2	97.1	97.9	97.0
	Weibull(0.4, 1)	69.2	63.5	63.0	73.4	63.0	88.2	84.5	84.2	90.1	84.1
	Lognormal(0, 1.5)	77.4	72.7	72.3	80.7	71.8	91.5	88.9	88.6	92.6	88.8
$n = 200$											
0.05	Normal(0, 1)	5.6	9.8	11.0	8.0	14.7	8.2	23.1	27.1	12.5	30.5
	SymPareto(2.1, 1)	9.9	36.5	42.0	16.4	43.7	24.0	74.2	78.6	33.4	75.7
	Weibull(0.4, 1)	3.9	13.3	17.7	10.1	22.2	8.2	35.4	42.8	18.3	46.1
	Lognormal(0, 1.5)	4.6	20.7	26.0	12.0	29.7	11.9	48.1	54.5	23.4	55.3
0.15	Normal(0, 1)	28.5	44.1	40.9	33.6	48.2	67.9	84.8	82.7	69.0	84.2
	SymPareto(2.1, 1)	79.6	89.4	88.6	80.7	88.7	94.7	97.1	96.9	94.9	97.0
	Weibull(0.4, 1)	43.6	60.3	59.1	50.1	61.9	74.2	85.4	85.0	76.6	85.5
	Lognormal(0, 1.5)	53.3	68.1	67.4	59.6	69.4	79.2	88.1	87.6	81.8	87.6
0.25	Normal(0, 1)	64.2	65.0	61.3	67.1	66.4	96.0	96.3	95.3	95.5	95.3
	SymPareto(2.1, 1)	94.0	94.5	94.1	94.6	94.2	98.1	98.4	98.3	98.2	98.2
	Weibull(0.4, 1)	71.6	75.6	73.9	75.4	74.8	90.8	92.8	92.1	92.2	92.1
	Lognormal(0, 1.5)	77.5	80.7	79.6	80.6	80.1	91.4	93.2	92.9	92.4	92.7
0.50	Normal(0, 1)	88.1	79.1	77.5	88.5	77.9	99.8	99.3	99.3	99.7	98.7
	SymPareto(2.1, 1)	97.4	96.7	96.5	97.6	96.2	99.0	98.8	98.8	99.1	98.8
	Weibull(0.4, 1)	86.4	82.1	81.7	88.3	81.4	96.6	95.5	95.3	97.2	94.8
	Lognormal(0, 1.5)	88.5	86.1	86.1	89.5	85.0	95.8	95.3	95.2	96.4	95.0
$n = 500$											
0.05	Normal(0, 1)	10.2	29.8	34.0	13.8	35.6	21.7	69.7	74.8	28.1	71.0
	SymPareto(2.1, 1)	29.4	78.6	82.5	36.2	79.4	71.1	96.1	96.9	69.3	96.2
	Weibull(0.4, 1)	8.4	34.8	42.6	16.9	43.9	23.9	73.8	79.8	34.8	76.8
	Lognormal(0, 1.5)	12.4	47.4	55.3	21.9	55.1	36.2	80.5	84.6	45.6	82.2
0.15	Normal(0, 1)	78.1	90.7	88.9	75.3	88.3	99.6	100.0	99.9	98.7	99.9
	SymPareto(2.1, 1)	96.5	98.2	98.1	96.6	98.1	98.8	99.4	99.4	99.1	99.3
	Weibull(0.4, 1)	78.6	89.2	89.2	79.3	87.9	95.8	98.2	98.2	95.5	98.0
	Lognormal(0, 1.5)	82.7	91.4	91.1	83.7	90.5	95.6	98.2	98.1	95.8	97.9
0.25	Normal(0, 1)	98.3	98.5	98.1	97.6	97.3	100.0	100.0	100.0	100.0	100.0
	SymPareto(2.1, 1)	98.7	99.0	98.9	98.9	99.0	99.5	99.7	99.7	99.6	99.6
	Weibull(0.4, 1)	94.3	95.7	95.1	94.6	94.6	99.1	99.5	99.4	99.4	99.4
	Lognormal(0, 1.5)	94.5	96.0	95.5	95.2	95.3	98.6	99.1	99.0	99.0	99.0
0.50	Normal(0, 1)	99.9	99.7	99.7	99.9	99.4	100.0	100.0	100.0	100.0	100.0
	SymPareto(2.1, 1)	99.4	99.4	99.3	99.5	99.2	99.7	99.7	99.7	99.7	99.6
	Weibull(0.4, 1)	98.7	97.4	97.3	98.9	97.3	99.9	99.8	99.8	99.9	99.8
	Lognormal(0, 1.5)	98.0	97.6	97.6	98.4	97.2	99.4	99.2	99.2	99.5	99.3

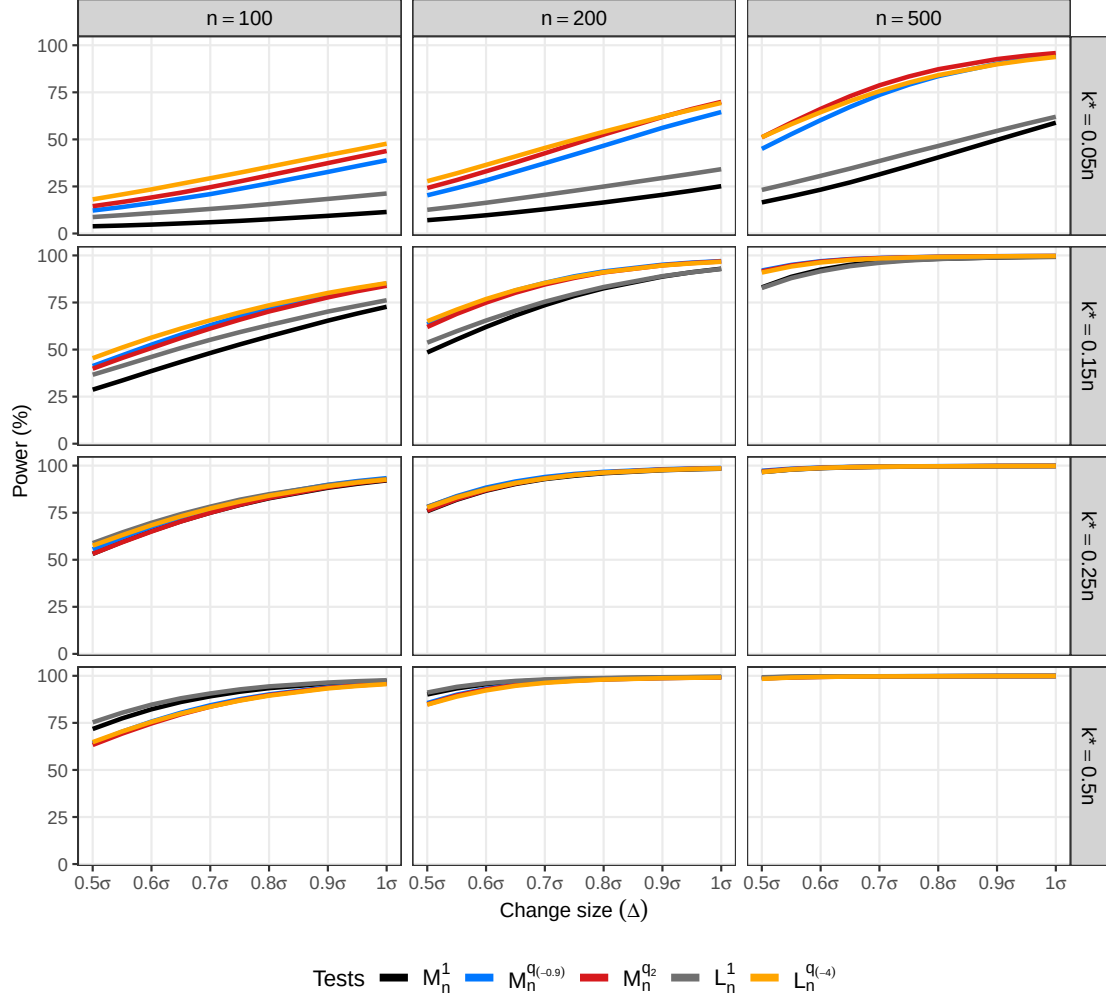


Figure 3: The empirical powers (averaged over 10 distributions) of the M_n^q and L_n^q tests as functions of the change size Δ .

As seen from Figure 3, there is a trade-off between the empirical power for changes on the tails and that for changes in the middle for the M_n^q and L_n^q tests. Thus, all the sup- and integral tests have similar powers at $k^* = 0.25n$, but the most powerful tests (the weighted ones) in detecting changes that occur on the tails (at $k^* = 0.05n$ and $k^* = 0.15n$) are also the least powerful ones for detecting changes in the middle at $k^* = 0.50n$, though when $n = 500$, the powers of all the tests at $k^* = 0.50n$ are close to 100%. Thus, the *weighted* sup- and integral tests are overall more preferable over the *unweighted* M_n^1 and L_n^1 tests, where the L_n^1 test is overall more powerful than the M_n^1 test.

Our proposed $M_n^{q(-0.9)}$ and $L_n^{q(-4)}$ tests are overall nearly as powerful for changes both on the tails and in the middle as the most powerful test obtained in Orasch and Pouliot (2004), the $M_n^{q_2}$ test, as evident from Tables 4, A1, and A2.

However, as concluded in Section 2.2.1, the slowly convergent type I errors of the $M_n^{q_2}$ test may require a sample size larger than 10,000 for becoming reasonably close to 5% (at or below 7.5%), while the type I errors of the $M_n^{q(-0.9)}$ test become very close to or under 7.5% for $n < 10,000$. Moreover, for the distributions with δ_{100} values larger than that for Exponential(1), the type I errors of the $M_n^{q(-0.9)}$ test are significantly lower and closer to 5% than those of the $M_n^{q_2}$ test. Notably, the $L_n^{q(-4)}$ test has better controlled type I errors that, for the distributions with larger δ_{100} values, are lower and approach 5% faster than those of the $M_n^{q(-0.9)}$ test.

In sum, the $L_n^{q(-4)}$ integral test is one of the most powerful and well-balanced nonparametric asymptotic tests for H_0 vs. H_A examined in this work, exhibiting consistently high power regardless of where the change occurs in the sample, while also maintaining the best controlled type I errors among the weighted sup- and integral tests.

On the Selection of the Sup- and Integral-Based Tests

Remark 6. Figure 4 helps to explain the heuristic rationale behind selecting the weight functions $q_{(\beta)}(t)$ for our proposed $M_n^{q(-0.9)}$ and $L_n^{q(-4)}$ tests. It illustrates how the empirical powers and type I errors of the $M_n^{q(\beta)}$ and $L_n^{q(\beta)}$ tests vary across different values of β on the example of the Lognormal(0, 1.5) samples and the change size $\Delta = 0.50\sigma$.

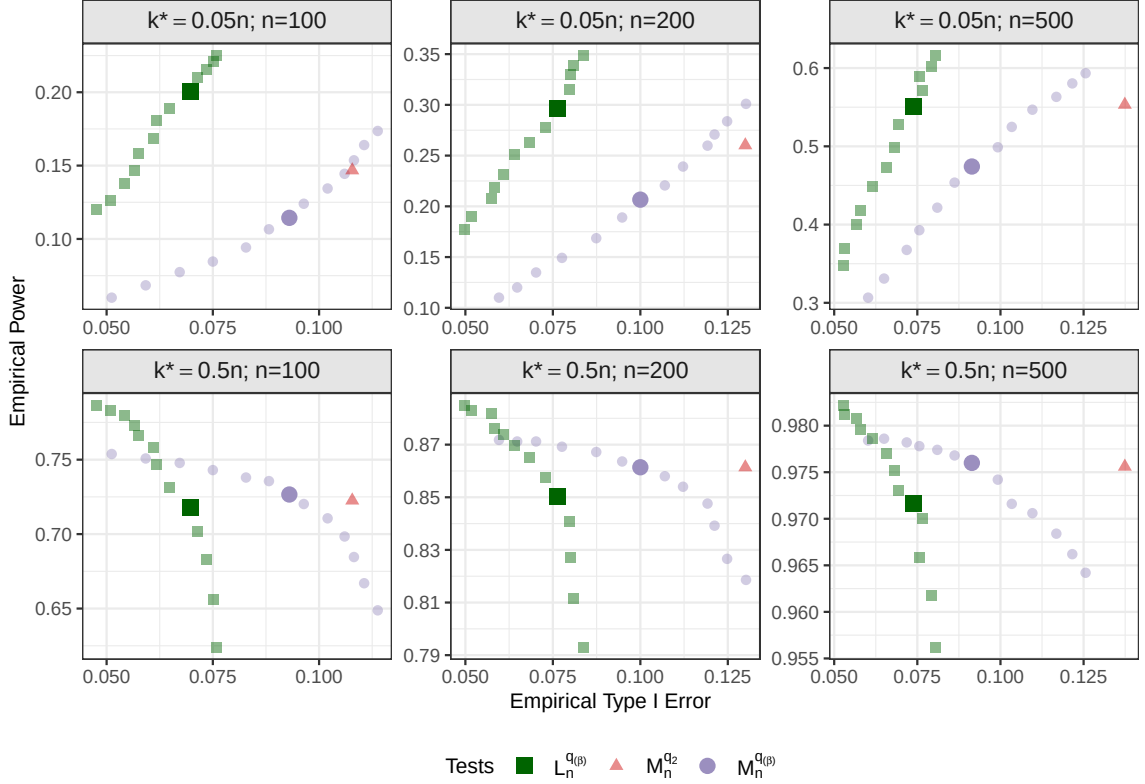


Figure 4: The empirical power vs. the empirical type I error of the $M_n^{q(\beta)}$ and $L_n^{q(\beta)}$ tests with different β for Lognormal(0, 1.5) and the change size $\Delta = 0.50\sigma$. $M_n^{q(\beta)}$ tests: β increases from -1.2 to -0.6 in steps of 0.05 (**right to left**). $L_n^{q(\beta)}$ tests: β increases from -5 to -2 in steps of 0.25 (**right to left**). Highlighted disk and square: $M_n^{q(-0.9)}$ and $L_n^{q(-4)}$ tests. Triangle: $M_n^{q_2}$ test.

For the $M_n^{q(\beta)}$ tests, β increases from -1.2 to -0.6 in steps of 0.05 as we move from right to left in the figure. When the change occurs near the beginning of the sample ($k^*/n = 0.05$), making β smaller increases the power, but this improvement is accompanied by larger type I errors. When the change occurs in the middle ($k^*/n = 0.50$), the pattern reverses: as β becomes smaller, the power decreases while the type I error increases. The choice of $\beta = -0.9$ for our $M_n^{q(-0.9)}$ test is simply the smallest value of β for which the type I errors stay below 10% across all distributions for $n = 100, 200$, and 500 , while still keeping the power reasonably close to that of the $M_n^{q_2}$ test (represented by the triangle in the figure), especially on the tails.

For the integral-based $L_n^{q(\beta)}$ tests, where β ranges from -5 to -2 in steps of 0.25 when reading Figure 4 from right to left, the behavior is different. The curve with the $L_n^{q(\beta)}$ tests lies to the left of that with the $M_n^{q(\beta)}$ tests, reflecting that the integral tests attain similar power with much smaller type I errors. Thus, for the integral tests, making β smaller still improves tail sensitivity, but the resulting type I error inflation is much smaller, leading to a more favorable trade-off between the power and type I error. We selected $\beta = -4$ for our $L_n^{q(-4)}$ test because it yields strong performance for changes on the tails. Indeed, for Lognormal(0, 1.5) at $k^*/n = 0.05$ with $\Delta = 0.5\sigma$, the $L_n^{q(-4)}$ test is more powerful than the M_n^{q2} test at $n = 100$ and 200 , and nearly matches the performance of the M_n^{q2} test at $n = 500$, while maintaining substantially lower type I errors and very similar power in the middle.

On Other Sup- and Integral-Based Tests

Remark 7. A different way of renormalizing $|Z_n(t)|$ while keeping $\sup_{0 < t < 1}$ is via a Darling–Erdős type of statistic

$$M_n^{DE} := \frac{A(\log n)}{\hat{\sigma}_n} \sup_{0 < t < 1} \frac{|Z_n(t)|}{\left(\frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right) \right)^{1/2}} - D(\log n), \quad (24)$$

where $A(x) = (2 \log x)^{1/2}$ and $D(x) = 2 \log x + 0.5 \log \log x - 0.5 \log \pi$. From Theorem 2.1.2 of Csörgő and Horváth (1997), on assuming that H_0 holds and $E(X^2 \log \log(|X| + 1)) < \infty$, as $n \rightarrow \infty$,

$$P(M_n^{DE} \leq y) \rightarrow e^{-2e^{-y}}, \quad \text{for all } y \in \mathbb{R}. \quad (25)$$

Similar to the M_n^q tests, one also rejects H_0 for large values of M_n^{DE} when testing H_0 vs. H_A via (25). The rate of convergence in (25) is slow (cf., e.g., Berkes et al. (2004)). Indeed, upon our own investigation of the power function of the M_n^{DE} test, its type I errors are marginally larger than those of the $M_n^{q(-0.9)}$ test except for Normal(0, 1) at $n = 200$ and 500 , where they are much lower (around 0.9% and 1.1%), and for Frechet(2.5) at $n = 500$. As for the empirical power of the M_n^{DE} test, it is everywhere less powerful than the M_n^{q2} , $M_n^{q(-0.9)}$, and $L_n^{q(-4)}$ tests.

Remark 8. From (ii) of Proposition 2.1.1 of Csörgő and Horváth (1997), on assuming that H_0 holds and $q(t) \in Q$, for $0 < p < \infty$,

$$L_{p,n}^q := \frac{1}{\widehat{\sigma}_n^p} \int_0^1 \frac{|Z_n(t)|^p}{q\left(\frac{[(n+1)t]}{n}\right)} dt \xrightarrow{\mathcal{D}} \int_0^1 \frac{|B(t)|^p}{q(t)} dt, \quad n \rightarrow \infty, \quad (26)$$

provided that

$$\int_0^1 \frac{(t(1-t))^{p/2}}{q(t)} dt < \infty. \quad (27)$$

While in this work we study the $L_{2,n}^{q(\beta)}$ - or simply the $L_n^{q(\beta)}$ -based tests, we also investigated the finite-sample performance of the tests for change in the mean that are based on the $L_{1,n}^{q(\beta)}$ and $L_{3,n}^{q(\beta)}$ statistics, as per the convergence in distribution in (26) and the respective condition on the weight function $q(\beta)$ in (27). We observed that the $L_{1,n}^{q(\beta)}$ tests have type I errors that are already at 5% for smaller samples ($n \leq 500$), even for the Lognormal(0,1.5) distribution. However, there is no possible β such that the test based on $L_{1,n}^{q(\beta)}$ can achieve the same high power as the $L_n^{q(-4)}$ test for detecting changes both on the tails and in the middle. As for the test based on $L_{3,n}^{q(\beta)}$, one can take $\beta = -4.5$ such that this test is as powerful as the $L_n^{q(-4)}$ test for changes at $k^* = 0.05n$, while also gaining some more power at $k^* = 0.15n, 0.25n$, and $0.50n$. However, the type I errors of the $L_{3,n}^{q(-4.5)}$ test are larger than those of the $L_n^{q(-4)}$ test, especially for distributions with larger δ_{100} values. Thus, among the $L_{p,n}^{q(\beta)}$ tests with $p = 1, 2$, and 3 , the ones based on $L_{2,n}^{q(\beta)}$ appeared to have the most favorable and well-balanced trade-offs among their type I errors, power to detect changes on the tails, and power to detect changes in the middle.

Remark 9. We note in passing that one can also define the sup- and integral tests for change in the mean that are based on (12) and (18) by replacing $\widehat{\sigma}_n$ with $\widetilde{\sigma}_{[nt+1]}$ put respectively under the supremum and integral, where

$$\widetilde{\sigma}_k^2 := \begin{cases} \frac{1}{n} \left(\sum_{1 \leq i \leq k} \left(X_i - \frac{S_k}{k} \right)^2 + \sum_{k < i \leq n} \left(X_i - \frac{S_n - S_k}{n - k} \right)^2 \right), & 1 \leq k < n, \\ \frac{1}{n} \sum_{1 \leq i \leq k} \left(X_i - \frac{S_k}{k} \right)^2, & k = n \end{cases} \quad (28)$$

(cf. Corollaries 2.1.1 and 2.1.3 in Csörgő and Horváth (1997)). This is due to the pooled

variances $\{\tilde{\sigma}_k^2, 1 \leq k \leq n\}_{n \geq 1}$ constituting a weakly uniformly consistent sequence of estimators for σ^2 , i.e., as $n \rightarrow \infty$, $\max_{1 \leq k \leq n} |\tilde{\sigma}_k^2 - \sigma^2| = o_P(1)$ (cf. Lemma 2.1.1 in Csörgő and Horváth (1997)). Large values of the corresponding test statistics, denoted by $\widetilde{M}_n^q$ and $\widetilde{L}_n^q$ here, also lead to the rejection of H'_0 . Csörgő et al. (2004) suggests, based on the references therein, that the weighted $\widetilde{M}_n^q$ tests (with $q(t) \not\equiv 1$) should be more powerful than the respective M_n^q tests. We compared the powers of our proposed $M_n^{q(-0.9)}$ sup-test and overall most well-balanced and powerful $L_n^{q(-4)}$ integral test with those of their $\widetilde{M}_n^q$ and $\widetilde{L}_n^q$ versions. We found that the powers of the $\widetilde{M}_n^{q(-0.9)}$ and $\widetilde{L}_n^{q(-4)}$ tests are somewhat higher across all the change times $k^* = 0.05n, 0.15n, 0.25n$, and $0.50n$ and for both change sizes $\Delta = 0.50\sigma$ and 0.75σ , but less significantly so as the sample size increases to $n = 500$. However, such gain in power comes at the cost of the larger type I errors, especially when $n = 100$. As the sample size increases, the difference between the type I errors of the respective tests can be seen to be getting smaller. Therefore, when dealing with distributions that are more asymmetric, heavy-tailed, or both, or, more generally, with the ones having larger δ_{100} values, we would not recommend using the more powerful $\widetilde{M}_n^{q(-0.9)}$ -based version of the $M_n^{q(-0.9)}$ test. For larger sample sizes ($n \geq 500$), the $\widetilde{L}_n^{q(-4)}$ test provides a viable alternative to the $L_n^{q(-4)}$ test, since their type I error levels are then close, but the improvement in power is relatively small.

3 Conclusions

In this paper, we investigate and numerically compare the finite-sample power functions of the nonparametric asymptotic tests for H_0 vs. H_A in (1) and (2) for AMOC change in the mean at the 5% asymptotic significance level that are based on the sup- and integral functionals of the appropriately weighted and normalized tied-down partial sums process $Z_n(t)$ of (7), leading to the test statistics M_n^q of (12) and L_n^q in (18), respectively. In M_n^q and L_n^q , we use some newly proposed weight functions of the form $q(t) = q_{(\beta)}(t)$ of (16) and some previously used $q(t)$ that satisfy conditions (13)/(14) and (19), respectively. For a variety of distributions with finite variances (σ^2) as in Table 1, with different degrees of asymmetry and heavy-tailedness that can be generally estimated with a multiple of the Berry–Esseen

upper bound δ_{100} as in (22), we study the empirical type I errors and powers of detecting the change in the mean of the M_n^q - and L_n^q -based tests, where the change occurs due to a shift in the sample as per (21), considering various change sizes Δ and change times k^* , and different sample sizes.

Both the M_n^q - and L_n^q -based tests become more powerful with an increase in the sample size and size Δ of the change in the mean. They are also more sensitive in detecting changes in the middle of the sample (compared with the ones on the tails of the sample), reaching the power close to 100% for $k^* = 0.50n$ already at $n = 500$. Thus, the M_n^q - and L_n^q -based tests with more power for change-point times close to the beginning and end of the sample, namely, the *weighted* M_n^q and L_n^q tests (with $q(t) \neq 1$), are more desirable. However, for each test in hand, we observe a three-way trade-off among its type I errors, power for detecting the change on the tails of the sample, and power for detecting the change in the middle of the sample. Nevertheless, by suitable choices of β for the weight function $q(t) = q_{(\beta)}(t)$, we find the $M_n^{q(-0.9)}$ sup-test and $L_n^{q(-4)}$ integral test that are nearly as sensitive in detecting changes both on the tails and in the middle of the sample as the overall most powerful $M_n^{q_2}$ test previously studied in the literature, with $q_2(t)$ of (17), while the type I errors of the new tests are overall closer to the 5% nominal level. Moreover, for the distributions with larger δ_{100} values, the type I errors of the $M_n^{q(-0.9)}$ and $L_n^{q(-4)}$ tests are lower, and approach 5% faster, than those of the $M_n^{q_2}$ test, more so for the $L_n^{q(-4)}$ test. Thus, the $L_n^{q(-4)}$ integral test emerges as one of the most powerful and well-balanced nonparametric asymptotic tests for H_0 vs. H_A examined in this work, exhibiting consistently high power regardless of where the change occurs in the sample, while also maintaining the best controlled type I errors among the weighted sup- and integral tests across a wide range of distributions.

Appendix to Chapter 3

Table A1: The empirical powers for Exponential(1), $t(3)$, and Half- $t(3)$ for the change sizes $\Delta = 0.50\sigma$ and 0.75σ .

k^*/n	Distribution	$\Delta = 0.50\sigma$					$\Delta = 0.75\sigma$				
		M_n^1	$M_n^{q(-0.9)}$	M_n^{q2}	L_n^1	$L_n^{q(-4)}$	M_n^1	$M_n^{q(-0.9)}$	M_n^{q2}	L_n^1	$L_n^{q(-4)}$
$n = 100$											
0.05	Exponential(1)	3.7	4.6	5.2	6.6	9.5	4.3	6.6	8.0	8.3	15.4
	$t(3)$	3.8	7.2	8.4	6.2	10.6	4.6	12.0	14.3	8.1	18.0
	Half- $t(3)$	3.2	6.5	8.1	6.4	10.6	4.1	10.5	14.0	9.0	20.6
0.15	Exponential(1)	12.2	18.7	17.1	19.0	25.4	30.7	48.1	44.9	39.5	53.1
	$t(3)$	14.1	23.8	22.5	21.2	28.9	36.2	53.8	51.3	44.6	57.4
	Half- $t(3)$	16.4	28.5	26.7	25.5	34.0	41.4	61.1	58.8	50.0	63.2
0.25	Exponential(1)	32.7	32.7	29.9	39.4	37.8	67.8	69.7	66.1	71.6	70.0
	$t(3)$	37.0	38.0	34.8	44.4	41.7	73.7	74.5	71.8	76.0	74.1
	Half- $t(3)$	42.9	45.7	42.6	49.7	48.4	76.3	78.7	76.3	79.1	77.9
0.50	Exponential(1)	56.1	43.6	42.0	61.3	46.8	89.2	81.1	79.7	90.7	80.2
	$t(3)$	61.8	50.1	48.6	66.6	52.0	90.6	84.8	84.0	91.9	83.8
	Half- $t(3)$	67.2	57.0	56.0	70.8	57.0	90.4	85.8	85.2	91.5	84.3
$n = 200$											
0.05	Exponential(1)	5.2	7.1	9.2	7.8	14.8	7.3	21.4	26.6	12.4	32.1
	$t(3)$	5.0	10.8	13.6	7.9	15.8	7.8	25.1	30.6	12.9	32.9
	Half- $t(3)$	4.9	10.2	13.1	8.9	19.3	7.8	29.0	35.7	16.0	40.6
0.15	Exponential(1)	30.0	47.0	44.6	36.4	50.8	69.7	85.9	84.1	69.3	84.3
	$t(3)$	31.6	48.8	46.0	38.1	50.4	72.5	86.4	84.8	71.0	84.9
	Half- $t(3)$	38.6	57.1	55.5	45.4	59.3	76.8	87.7	87.1	77.3	87.7
0.25	Exponential(1)	65.8	67.8	64.4	67.4	67.4	95.3	95.9	95.3	95.0	95.2
	$t(3)$	66.9	69.7	66.5	69.0	68.8	95.3	95.7	95.2	95.2	94.8
	Half- $t(3)$	72.8	75.2	72.9	75.4	75.0	94.7	95.5	94.9	94.9	94.8
0.50	Exponential(1)	87.3	78.9	77.4	88.4	77.9	99.5	99.1	98.9	99.6	98.6
	$t(3)$	87.5	80.4	79.6	88.4	79.1	99.0	97.9	97.8	99.0	97.7
	Half- $t(3)$	90.1	84.9	83.9	90.8	83.5	98.4	97.4	97.3	98.5	97.0
$n = 500$											
0.05	Exponential(1)	10.0	28.5	34.6	13.8	36.2	20.0	71.0	76.7	27.3	71.1
	$t(3)$	10.1	30.7	36.9	15.3	36.6	22.3	71.4	76.6	29.1	72.4
	Half- $t(3)$	9.7	33.8	40.9	16.6	41.5	23.5	76.1	81.4	33.1	77.0
0.15	Exponential(1)	77.1	89.9	88.5	75.6	88.0	99.2	99.8	99.8	98.3	99.7
	$t(3)$	78.0	89.7	89.2	76.9	88.9	98.4	99.4	99.3	97.5	99.4
	Half- $t(3)$	81.2	91.4	91.0	80.2	90.1	97.6	98.8	98.9	97.1	98.7
0.25	Exponential(1)	97.7	98.0	97.6	97.3	96.8	100.0	100.0	100.0	100.0	100.0
	$t(3)$	97.0	97.4	97.1	96.7	96.7	99.8	99.9	99.9	99.8	99.8
	Half- $t(3)$	96.1	97.0	96.8	96.3	96.3	99.6	99.6	99.6	99.7	99.6
0.50	Exponential(1)	99.9	99.6	99.7	99.9	99.4	100.0	100.0	100.0	100.0	100.0
	$t(3)$	99.4	99.1	99.1	99.5	98.6	99.9	99.9	99.9	99.9	99.9
	Half- $t(3)$	99.3	98.8	98.7	99.4	98.5	99.9	99.8	99.8	99.9	99.8

Table A2: The empirical powers for Frechet(2.5), Inverse-Normal(7, 1), and Pareto(2.1, 1) for the change sizes $\Delta = 0.50\sigma$ and 0.75σ .

k^*/n	Distribution	$\Delta = 0.50\sigma$					$\Delta = 0.75\sigma$				
		M_n^1	$M_n^{q(-0.9)}$	M_n^{q2}	L_n^1	$L_n^{q(-4)}$	M_n^1	$M_n^{q(-0.9)}$	M_n^{q2}	L_n^1	$L_n^{q(-4)}$
$n = 100$											
0.05	Frechet(2.5)	3.8	8.6	10.6	8.5	16.2	5.1	21.6	27.2	13.3	33.6
	Inverse-Normal(7, 1)	2.0	9.3	11.9	7.1	15.1	3.2	17.2	21.7	11.5	27.9
	Pareto(2.1, 1)	8.0	43.5	49.6	21.2	51.7	23.6	71.8	75.9	39.4	75.3
0.15	Frechet(2.5)	28.5	46.3	43.8	39.0	50.5	60.7	76.7	74.9	66.1	77.8
	Inverse-Normal(7, 1)	21.4	34.7	34.2	32.9	40.8	45.1	62.6	60.9	54.8	65.3
	Pareto(2.1, 1)	74.4	85.5	84.5	78.6	85.3	89.5	94.2	94.2	90.9	94.0
0.25	Frechet(2.5)	60.4	63.9	61.6	65.5	64.5	85.2	87.5	86.1	86.8	86.3
	Inverse-Normal(7, 1)	46.1	50.5	48.3	54.5	52.7	72.9	76.4	74.5	78.4	76.9
	Pareto(2.1, 1)	88.8	91.0	90.5	90.6	90.3	95.1	96.2	95.9	96.3	96.0
0.50	Frechet(2.5)	78.0	72.8	72.2	81.2	71.7	93.1	91.4	91.0	94.3	90.5
	Inverse-Normal(7, 1)	66.8	59.7	59.4	71.1	59.6	87.3	83.6	83.4	89.4	82.8
	Pareto(2.1, 1)	94.1	92.8	92.7	94.9	92.7	97.7	97.3	97.2	98.1	97.0
$n = 200$											
0.05	Frechet(2.5)	5.3	16.9	21.5	11.0	27.4	10.6	47.1	53.7	21.1	54.2
	Inverse-Normal(7, 1)	3.8	10.6	15.1	9.5	19.3	6.7	28.4	36.5	16.6	41.8
	Pareto(2.1, 1)	22.6	67.1	72.6	34.3	71.1	54.3	87.8	89.7	60.4	88.5
0.15	Frechet(2.5)	54.0	70.8	69.7	58.2	71.1	83.6	91.6	91.1	85.0	91.8
	Inverse-Normal(7, 1)	37.2	55.5	54.7	45.3	57.8	71.0	84.4	83.8	74.1	84.7
	Pareto(2.1, 1)	88.3	93.3	93.3	89.4	93.0	95.1	97.6	97.5	95.9	97.4
0.25	Frechet(2.5)	81.8	84.3	82.7	83.5	83.1	95.0	96.0	95.7	95.8	95.7
	Inverse-Normal(7, 1)	68.3	71.3	69.9	72.5	71.7	91.7	93.5	92.7	92.8	92.7
	Pareto(2.1, 1)	94.9	96.1	95.8	95.7	95.6	97.9	98.5	98.4	98.4	98.4
0.50	Frechet(2.5)	92.2	89.2	88.6	93.4	89.0	98.0	97.3	97.2	98.2	97.1
	Inverse-Normal(7, 1)	86.3	80.9	80.4	88.3	80.0	97.6	96.2	96.1	98.1	96.0
	Pareto(2.1, 1)	97.2	96.9	96.8	97.8	96.9	99.1	99.0	98.9	99.3	98.8
$n = 500$											
0.05	Frechet(2.5)	11.8	49.0	56.2	20.9	54.7	35.3	84.2	87.9	43.5	83.9
	Inverse-Normal(7, 1)	8.8	29.3	37.7	16.0	39.9	21.4	71.6	78.3	31.3	74.7
	Pareto(2.1, 1)	54.5	87.7	90.2	60.2	88.7	83.2	96.2	97.0	83.8	96.2
0.15	Frechet(2.5)	85.4	93.0	92.9	86.0	92.4	96.3	98.3	98.2	96.5	98.2
	Inverse-Normal(7, 1)	76.6	88.7	88.6	77.0	87.2	96.4	98.7	98.7	95.9	98.5
	Pareto(2.1, 1)	96.1	97.9	97.9	96.6	97.8	98.3	99.1	99.1	98.9	99.2
0.25	Frechet(2.5)	95.7	96.6	96.5	96.1	96.1	98.8	99.0	99.0	99.0	99.0
	Inverse-Normal(7, 1)	95.1	95.8	95.4	94.8	95.0	99.8	99.9	99.8	99.9	99.8
	Pareto(2.1, 1)	98.3	98.7	98.6	98.8	98.7	99.3	99.5	99.5	99.4	99.5
0.50	Frechet(2.5)	98.0	97.6	97.7	98.3	97.6	99.3	99.2	99.3	99.4	99.2
	Inverse-Normal(7, 1)	99.2	98.0	98.0	99.2	97.7	100.0	100.0	100.0	100.0	100.0
	Pareto(2.1, 1)	99.2	99.0	99.0	99.2	99.0	99.7	99.6	99.6	99.7	99.6

References

- Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain “goodness of fit” criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 23, 193–212. <https://doi.org/10.1214/aoms/1177729437>
- Bentkus, V., & Götze, F. (1996). The berry–esseen bound for student’s statistic. *The Annals of Probability*, 24(1), 491–503. <https://doi.org/10.1214/aop/1042644728>
- Berkes, I., Horváth, L., Hušková, M., & Steinebach, J. (2004). Applications of permutations to the simulations of critical values. *Journal of Nonparametric Statistics*, 16, 197–216. <https://doi.org/10.1080/10485250310001624747>
- Csörgő, M. (2002). A glimpse of the impact of pá l erdős on probability and statistics. *Canadian Journal of Statistics*, 30, 493–556. <https://doi.org/10.2307/3316095>
- Csörgő, M., Csörgő, S., Horváth, L., & Mason, D. M. (1986). Weighted empirical and quantile processes. *The Annals of Probability*, 14(1), 31–85. <https://doi.org/10.1214/aop/1176992617>
- Csörgő, M., & Horváth, L. (1988). Nonparametric methods for changepoint problems. In *Quality control and reliability* (pp. 403–425, Vol. 7). Elsevier. [https://doi.org/10.1016/S0169-7161\(88\)07022-1](https://doi.org/10.1016/S0169-7161(88)07022-1)
- Csörgő, M., & Horváth, L. (1997). *Limit theorems in change-point analysis*. Wiley.
- Csörgő, M., Szyszkowicz, B., & Wang, Q. (2004). On weighted approximations and strong limit theorems for self-normalized partial sums processes. In *Asymptotic methods in stochastics* (pp. 489–521, Vol. 44). American Mathematical Society.
- Eastwood, B. J., & Eastwood, V. R. (1998). Tabulating weighted sup-norm functionals of brownian bridges via monte carlo simulation. In *Asymptotic methods in probability and statistics* (pp. 707–719). North-Holland. <https://doi.org/10.1016/B978-044450083-0/50046-0>
- Orasch, M., & Pouliot, W. (2004). Tabulating weighted sup-norm functionals used in change-point analysis. *Journal of Statistical Computation and Simulation*, 74, 249–276. <https://doi.org/10.1080/0094965031000148703>

- Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, 19, 279–281. <https://doi.org/10.1214/aoms/1177730256>
- Szyszkowicz, B. (1991). Weighted stochastic processes under contiguous alternatives. *C. R. Math. Rep. Acad. Sci. Canada*, 13, 211–216.

Chapter 4

Nonparametric Asymptotic Tests for a Change in the Slope and Intercept in Simple Linear Regression

Nonparametric Asymptotic Tests for a Change in the Slope and Intercept in Simple Linear Regression

Jervis Gallanosa and Yuliya V. Martsynyuk

Department of Statistics,
University of Manitoba,
Winnipeg MB R3T 2N2

The content of this chapter is based on the following manuscript to be submitted for
publication:

Gallanosa, J.; Martsynyuk, Y. V. (2026). *Nonparametric Asymptotic Tests for a Change
in the Slope and Intercept in Simple Linear Regression*. Manuscript prepared for
submission.

Abstract

This paper develops nonparametric asymptotic tests for changes in regression coefficients in a simple linear regression model with random design. The slope-change and intercept-change problems are treated separately, leading to scalar residual-based tied-down partial-sum processes tailored to each component. Under the corresponding null hypotheses, these residual processes are shown to have the same limiting behavior as the tied-down partial-sum processes used in nonparametric mean-change testing.

The main contribution is a broad asymptotic framework for several classes of component-specific residual-process statistics. For both the slope-change and intercept-change problems, limiting distributions are established for supremum and integral statistics with Brownian-bridge-based limits, supremum and integral Darling–Erdős-type statistics, selected joint limits, and Rényi-type supremum statistics with Wiener-process limits. The supremum results connect with existing residual-based and tied-down partial-sum approaches in regression change-point analysis, while the integral, joint, and Rényi-type results extend the range of available asymptotic tools for component-specific testing.

The proposed statistics summarize the residual partial-sum path in different ways. Supremum statistics measure the largest suitably normalized difference over possible change-point locations, while integral statistics average deviations across the tied-down partial-sum path. The resulting theory provides an asymptotic basis for testing slope and intercept changes in simple linear regression and motivates future work on finite-sample performance and applications of the proposed procedures.

1 Introduction

Detecting structural change in regression models is a central problem in change-point analysis. In many applications, the relationship between a response variable and its covariates may not remain stable over time. This paper considers the retrospective problem of testing whether a regression coefficient changed at an unknown time within the observed sample. Even in simple linear regression, instability can arise through a change in the intercept, a change in the slope, or both. These two forms of change have different interpretations. An

intercept shift corresponds to a change in the baseline level of the response, while a slope change reflects a change in the strength or direction of the covariate–response association. This distinction motivates procedures that can target the two components separately, rather than relying only on omnibus tests for instability of the full regression relationship.

A substantial literature addresses structural change in linear regression models. Early contributions include Chow (1960), who considered breaks at known times, and Brown et al. (1975), who developed tests based on cumulative sums of recursive residuals. Subsequent work expanded the scope to unknown break points, multiple changes, and changes in subsets or full vectors of regression parameters. In many procedures, the main objective is to test whether the regression relationship, or the full parameter vector, is stable. A rejection then signals that at least one coefficient has changed, but does not by itself identify which component is responsible.

The distinction between intercept and slope changes is important in applications where the regression coefficients have different substantive meanings. For example, Olmo et al. (2011) detect potential insider trading through a structural change in the intercept of an extended capital asset pricing model. In that setting, an intercept shift represents a sudden mean shift in abnormal returns not explained by the model’s market-risk covariates. Their test is an intercept-change procedure based on weighted tied-down partial sums of residual-type quantities, and is designed both to detect and to estimate the timing of an intercept break. Similarly, Pouliot (2016) studies mutual fund performance in a regression framework where the intercept represents abnormal returns, or alpha, and the slope coefficients measure exposures to risk factors. In that setting, a change in alpha and a change in risk-factor exposure have different financial interpretations. These examples illustrate why component-specific information is useful for interpreting regression instability.

Residual-process methods form an important part of the linear regression change-point literature. Gombay and Horváth (1994) studied tests for changes based on maxima of least-squares residual processes, where the regressors may be fixed or random and the errors are independent and identically distributed. Related ideas were extended to weakly dependent errors with non-stochastic regressors by Gombay (2010), while Horváth and Rice (2014) developed likelihood-based maxima-type tests. More recently, Horváth et al. (2023)

studied tests using supremum-type functionals of partial sums of weighted residuals, where the weighted residuals are formed by multiplying fitted residuals by the regressors. These works are generally formulated in terms of changes in the regression parameter vector, or in a specified subset of that vector. In the one-change formulations of Gombay and Horváth (1994), Gombay (2010), and Horváth and Rice (2014), the parameters under consideration change at a common unknown time, while Horváth et al. (2023) allows a more general multiple-change formulation. Nevertheless, when these approaches are viewed in simple linear regression with an unknown intercept and slope, the associated statistics assess stability of the intercept–slope pair rather than separating the two components through scalar slope- and intercept-specific processes.

Several works are more directly related to the component-specific statistics considered in this work. In fixed-design simple regression, Sen (1980) introduced a least-squares change-point process for testing a possible slope change, and Abd-Rabou and Aly (2000) compared Kolmogorov–Smirnov, Cramér–von Mises, Anderson–Darling, and Erdős–Darling-type tests based on this process. For structural errors-in-variables models, Martsynyuk (2021) developed nonparametric weighted tied-down partial-sum tests for separate changes in the slope and intercept. The problem of separating intercept changes from slope changes was also studied by Pouliot (2016), who developed marginal slope- and intercept-related supremum statistics and combined them through a bivariate reformulation for joint and simultaneous testing.

The present paper instead focuses on simple linear regression with random design and on scalar residual-based tied-down partial-sum processes tailored separately to the slope-change and intercept-change problems. For the slope-change problem, the basic quantities are covariate-weighted residuals, or centered-covariate-weighted residuals when the intercept is unknown. For the intercept-change problem, adjusted residuals are used instead. In both cases, the processes have the same tied-down partial-sum structure as those used in nonparametric mean-change testing. This connection allows several classes of mean-change statistics to be adapted to the regression setting, including Brownian-bridge-based supremum and integral statistics, Darling–Erdős-type supremum and integral statistics, selected joint limit results, and Rényi-type supremum statistics with Wiener-process limits.

Together, these results provide a broad asymptotic basis for several statistics derived from the slope- and intercept-specific residual processes.

The inclusion of integral statistics is an important part of the paper. Supremum statistics are common in change-point analysis because they measure the largest suitably normalized difference between the observations before and after each possible change-point location and naturally lead to an estimator of the change location. Integral statistics instead average the accumulated deviations over the full tied-down partial-sum path. The mean-change results of Gallanosa and Martsynyuk (2026) show that suitable weighted integral tests can have competitive finite-sample performance and may give a favorable balance between power and type I error. This motivates the development of regression analogues of the Brownian-bridge-based integral statistics and the integral Darling–Erdős-type statistics derived in this paper.

The remainder of the paper is organized as follows. Section 2 recalls the limit theorems for weighted partial-sum processes used in the subsequent development. Section 3 develops the residual-based statistics for testing a change in the slope and establishes their asymptotic null distributions. Section 4 gives the corresponding statistics and results for testing a change in the intercept. Section 5 concludes with a discussion of possible directions for future work.

2 Limit Theorems for Change-Point Analysis

The regression statistics developed in this paper are motivated by classical nonparametric tests for a change in the mean; see, for example, Chapter 2 of Csörgő and Horváth (1997). Let $Z_1, \dots, Z_n$ be chronologically ordered independent observations with finite means. In the at-most-one-change (AMOC) problem, one tests

$$H_0 : \mathbb{E}Z_1 = \dots = \mathbb{E}Z_n \quad \text{against} \quad H_A : \mathbb{E}Z_1 = \dots = \mathbb{E}Z_{k^*} \neq \mathbb{E}Z_{k^*+1} = \dots = \mathbb{E}Z_n \quad (1)$$

for some unknown k^* , $1 \leq k^* < n$.

For a fixed k , $1 \leq k < n$, write

$$\bar{Z}_{1:k} := \frac{1}{k} \sum_{i=1}^k Z_i, \quad \bar{Z}_{k+1:n} := \frac{1}{n-k} \sum_{i=k+1}^n Z_i.$$

The difference between the two sample means can then be written as

$$\bar{Z}_{1:k} - \bar{Z}_{k+1:n} = \frac{n}{k(n-k)} \left(\sum_{i=1}^k Z_i - \frac{k}{n} \sum_{i=1}^n Z_i \right), \quad (2)$$

so the mean-change problem is naturally associated with the centered partial sums

$$\sum_{i=1}^k Z_i - \frac{k}{n} \sum_{i=1}^n Z_i, \quad 1 \leq k < n.$$

This leads to the following tied-down partial-sum process $Z_n(t)$, which is the main building block for the mean-change statistics recalled in this section and motivates the residual-based processes used later for regression.

Definition 2.1. For random variables $\{Z, Z_i\}_{i \geq 1}$, define the tied-down partial-sum process $\{Z_n(t), 0 \leq t \leq 1\}_{n \geq 1}$ by

$$Z_n(t) := \begin{cases} \frac{1}{\sqrt{n}} \left(\sum_{i=1}^{\lfloor (n+1)t \rfloor} Z_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n Z_i \right), & 0 \leq t < 1, \\ 0, & t = 1, \end{cases} \quad (3)$$

where $\sum_{i=1}^0 Z_i = 0$.

A common approach to testing H_0 against the AMOC alternative is to reject H_0 for large values of statistics based on supremum-type functionals of $Z_n(t)$, where the supremum is taken over the possible change-point locations. Such statistics correspond to the largest in absolute value before-and-after sample mean difference, and the location where the supremum is attained can be used to estimate the time of change. Integral-type functionals provide an alternative by averaging values derived from the process over the range of possible change-point locations.

We now recall the mean-change limit theorems that will be used later. These results

are based on weighted versions of the supremum and integral functionals of $Z_n(t)$. The first two theorems, [Theorems 2.1](#) and [2.2](#), involve a class of nonrandom weight functions $q(t)$ that, under appropriate conditions, allow the corresponding supremum and integral to be taken over the whole interval $(0, 1)$ while securing nondegenerate limiting distributions. The choice of the weight $q(t)$ also plays an important role in change-point testing, since suitable choices can improve the power of the corresponding test, especially for changes that occur early or late in the sample. Numerical comparisons in the mean-change setting by Orasch and Pouliot ([2004](#)) and Gallanosa and Martsynyuk ([2026](#)) illustrate the finite-sample effects of different choices of these weight functions.

Definition 2.2. Let Q be the class of positive functions $q(t) : (0, 1) \rightarrow (0, \infty)$, satisfying

$$\inf_{\delta \leq t \leq 1-\delta} q(t) > 0 \quad \text{for all } \delta \in (0, 1/2),$$

and such that $q(t)$ is non-decreasing near 0 and non-increasing near 1. For $q(t) \in Q$, define

$$I_0^1(q, c) := \int_0^1 \frac{e^{-cq^2(t)/(t(1-t))}}{t(1-t)} dt, \quad c > 0. \quad (4)$$

The integral $I_0^1(q, c)$ appears in the characterization of upper class functions of a Wiener process (see, for example, Csörgő et al. ([1986](#))).

The first mean-change result concerns the supremum functional of $Z_n(t)$ with appropriate weights $q(t) \in Q$. Its limiting distribution is the corresponding weighted supremum functional of a Brownian bridge.

Theorem 2.1. (Csörgő and Horváth, [1997](#); Theorem 2.1.1). Assume that $\{Z, Z_i\}_{i \geq 1}$ are i.i.d. random variables with $0 < \text{Var}(Z) < \infty$, and let $q(t) \in Q$. Suppose that $I_0^1(q, c) < \infty$ for some or possibly all $c > 0$. Then, as $n \rightarrow \infty$,

$$\sup_{0 < t < 1} \frac{|Z_n(t)|}{\hat{\sigma}_n q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} \xrightarrow{d} \sup_{0 < t < 1} \frac{|B(t)|}{q(t)}, \quad (5)$$

where $\{B(t), 0 \leq t \leq 1\}$ is a standard Brownian bridge and

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n \left(Z_i - \frac{1}{n} \sum_{j=1}^n Z_j \right)^2. \quad (6)$$

Remark 1. Since $Z_n(t)$ is constant on the intervals

$$\left[\frac{k}{n+1}, \frac{k+1}{n+1} \right), \quad k = 0, 1, \dots, n,$$

the supremum statistic in [Theorem 2.1](#) can be written as a maximum over the possible change-point locations. In particular,

$$\sup_{0 < t < 1} \frac{|Z_n(t)|}{\hat{\sigma}_n q(\lfloor (n+1)t \rfloor / n)} = \max_{1 \leq k \leq n} \frac{\left| \sum_{i=1}^k Z_i - \frac{k}{n} \sum_{i=1}^n Z_i \right|}{\sqrt{n} \hat{\sigma}_n q(k/n)}.$$

The next mean-change result concerns integral-type functionals of $Z_n(t)$, which can be written as finite averages of weighted powers of the normalized tied-down partial sums over the possible change-point locations, as shown in [Remark 2](#).

Theorem 2.2. (Csörgő and Horváth, [1997](#); Proposition 2.1.1(ii)). Assume that $\{Z, Z_i\}_{i \geq 1}$ are i.i.d. random variables with $0 < \text{Var}(Z) < \infty$, let $0 < p < \infty$, and let $q(t) \in \mathcal{Q}$. Suppose that

$$\int_0^1 \frac{(t(1-t))^{p/2}}{q(t)} dt < \infty. \quad (7)$$

Then, as $n \rightarrow \infty$,

$$\int_0^1 \left(\frac{|Z_n(t)|}{\hat{\sigma}_n} \right)^p \frac{1}{q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt \xrightarrow{d} \int_0^1 \frac{|B(t)|^p}{q(t)} dt, \quad (8)$$

where $\{B(t), 0 \leq t \leq 1\}$ is a standard Brownian bridge.

Remark 2. Using the same step-function representation of $Z_n(t)$, the integral statistic in

Theorem 2.2 can be written as

$$\int_0^1 \left(\frac{|Z_n(t)|}{\widehat{\sigma}_n} \right)^p \frac{1}{q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt = \frac{1}{n+1} \sum_{k=1}^{n-1} \left(\frac{\left| \sum_{i=1}^k Z_i - \frac{k}{n} \sum_{i=1}^n Z_i \right|}{\sqrt{n} \widehat{\sigma}_n} \right)^p \frac{1}{q(k/n)}. \quad (9)$$

Remark 3. *Weighted supremum tests for a change in the mean were studied numerically by Orasch and Pouliot (2004), who compared several weighted sup-norm functionals and tabulated the corresponding limiting distributions. More recently, Gallanosa and Martsynyuk (2026) studied the finite-sample performance of weighted supremum tests more extensively across several distributions, sample sizes, change sizes, and change-point locations. Their results showed that weighted supremum tests can improve power for early and late changes, but that this improvement may come with larger type I errors, especially for asymmetric or heavy-tailed distributions.*

Gallanosa and Martsynyuk (2026) also studied weighted integral tests for the mean-change problem. Although *Theorem 2.2* is stated for any $0 < p < \infty$, the main finite-sample comparison in Gallanosa and Martsynyuk (2026) focused on the case $p = 2$. The weighted integral test with $p = 2$ had power that was competitive with the strongest weighted supremum tests across change-point locations, while often giving better controlled type I errors. The cases $p = 1$ and $p = 3$ were also considered, but among $p = 1, 2, 3$, the case $p = 2$ gave the most favorable balance between power and type I error control in that study.

The weight

$$q^\star(t) := \sqrt{t(1-t)}, \quad 0 < t < 1,$$

is an important special case because it provides the natural pointwise standardization of the tied-down partial-sum process. This is consistent with the Brownian-bridge-based limit in *Theorem 2.1*, since a standard Brownian bridge $B(t)$ satisfies

$$\text{Var}(B(t)) = t(1-t).$$

Thus $q^\star(t)$ corresponds to the standard deviation of the limiting process at each fixed

$t \in (0, 1)$. However, for $c > 0$,

$$I_0^1(q^*, c) = e^{-c} \int_0^1 \frac{dt}{t(1-t)} = \infty.$$

Thus, this weight does not satisfy the integral condition in [Theorem 2.1](#), and the corresponding supremum statistic does not have a nondegenerate limit. Instead, the standardized maximum over all possible change-point locations requires Darling–Erdős-type centering and normalization, which yields a Gumbel-type limiting distribution.

For any result involving the supremum Darling–Erdős-type normalization, logarithms appearing in iterated logarithms are understood under the convention

$$\log(c) := \log\{\max(e, c)\}, \quad c > 0.$$

Theorem 2.3. (*Csörgő and Horváth, 1997; Theorem A.4.2*). Assume that $\{Z, Z_i\}_{i \geq 1}$ are i.i.d. random variables with $0 < \text{Var}(Z) < \infty$. Then, as $n \rightarrow \infty$,

$$\frac{(2 \log \log n)^{-1/2}}{\sqrt{\text{Var}(Z)}} \max_{1 \leq k < n} \left(\frac{n}{k(n-k)} \right)^{1/2} \left| \sum_{i=1}^k Z_i - \frac{k}{n} \sum_{i=1}^n Z_i \right| \xrightarrow{p} 1. \quad (10)$$

Moreover, if

$$\mathbb{E}[Z^2 \log \log(|Z| + 1)] < \infty,$$

then

$$A(n) \frac{1}{\sqrt{\text{Var}(Z)}} \max_{1 \leq k < n} \left(\frac{n}{k(n-k)} \right)^{1/2} \left| \sum_{i=1}^k Z_i - \frac{k}{n} \sum_{i=1}^n Z_i \right| - D(n) \xrightarrow{d} G, \quad (11)$$

where

$$A(n) = \sqrt{2 \log \log n}, \quad D(n) = 2 \log \log n + \frac{1}{2} \log \log \log n - \frac{1}{2} \log \pi,$$

and G has distribution function

$$\mathbb{P}(G \leq x) = \exp(-2e^{-x}), \quad x \in \mathbb{R}.$$

The convergences in (10) and (11) remain valid when $\text{Var}(Z)$ is replaced by $\hat{\sigma}_n^2$.

Remark 4. *The convergence in Darling–Erdős-type results is known to be slow (see Berkes et al. (2004) for related discussion on the simulation of critical values). In the mean-change simulations of Gallanosa and Martsynyuk (2026), the Darling–Erdős-type statistic was found to be less powerful than the tests studied there that were constructed from the weighted supremum and weighted integral statistics in Theorems 2.1 and 2.2 with specially chosen weight functions.*

An integral analogue of the Darling–Erdős-type supremum statistic is obtained by considering an L^p -type functional of the tied-down partial-sum process with the endpoint weight $(t(1-t))^{1+p/2}$. This weight plays, for integral-type functionals, a role analogous to the point-wise standardizing weight $q^*(t) = \sqrt{t(1-t)}$ in the supremum case. The following result gives the corresponding asymptotic normal limit after logarithmic centering and normalization. It appears that the finite-sample properties of this test have not been studied in the literature.

Theorem 2.4. *(Csörgő and Horváth, 1997; Theorem 2.1.5). Assume that $\{Z, Z_i\}_{i \geq 1}$ are i.i.d. random variables with $0 < \text{Var}(Z) < \infty$. Let $1 \leq p < \infty$, and suppose that*

$$\mathbb{E}|Z|^\nu < \infty \quad \text{for some } \nu > 2.$$

Then, as $n \rightarrow \infty$,

$$\frac{1}{\sqrt{2C_p \log n}} \left[\int_0^1 \frac{|Z_n(t)|^p}{\hat{\sigma}_n^p \left(\frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right) \right)^{1+p/2}} dt - 2m_p \log n \right] \xrightarrow{d} N(0, 1), \quad (12)$$

where

$$m_p = \int_{-\infty}^{\infty} |x|^p \phi(x) dx,$$

ϕ is the standard normal density, and

$$C_p = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |xy|^p [f_{\rho(u)}(x, y) - \phi(x)\phi(y)] dx dy du. \quad (13)$$

Here $\rho(u) = \exp(-|u|/2)$, and $f_{\rho}(x, y)$ denotes the bivariate normal density with mean zero,

unit variances, and correlation coefficient ρ .

The preceding marginal limits can also be combined into two joint limits: one for the supremum-based statistics in [Theorems 2.1](#) and [2.3](#), and one for the integral-based statistics in [Theorems 2.2](#) and [2.4](#). In each case, a statistic with a Brownian-bridge-based limit is paired with its corresponding Darling–Erdős-type statistic, and the factorization of the limiting distribution shows that the two component statistics are asymptotically independent.

Theorem 2.5. (*Csörgő and Horváth, 1997; Proposition 2.1.1*). *Let $\{Z, Z_i\}_{i \geq 1}$ be i.i.d. random variables with $0 < \text{Var}(Z) < \infty$.*

(i) *Let $q(t) \in Q$, and suppose that $I_0^1(q, c) < \infty$ for some or possibly all $c > 0$. Assume also that*

$$\mathbb{E}[Z^2 \log \log(|Z| + 1)] < \infty,$$

Then, for all $x, y \in \mathbb{R}$, as $n \rightarrow \infty$,

$$\begin{aligned} & \mathbb{P} \left\{ \sup_{0 < t < 1} \frac{|Z_n(t)|}{\hat{\sigma}_n q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} \leq x, \right. \\ & \quad \left. A(n) \sup_{0 < t < 1} \frac{|Z_n(t)|}{\hat{\sigma}_n \left[\frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n}\right)\right]^{1/2}} - D(n) \leq y \right\} \\ & \rightarrow \mathbb{P} \left\{ \sup_{0 < t < 1} \frac{|B(t)|}{q(t)} \leq x \right\} \exp(-2e^{-y}). \end{aligned} \tag{14}$$

(ii) *Let $0 < p < \infty$, $1 \leq r < \infty$, and let $q(t) \in Q$ satisfy*

$$\int_0^1 \frac{(t(1-t))^{p/2}}{q(t)} dt < \infty. \tag{15}$$

Assume also that

$$\mathbb{E}|Z|^\nu < \infty \quad \text{for some } \nu > 2.$$

Then, for all $x, y \in \mathbb{R}$, as $n \rightarrow \infty$,

$$\begin{aligned} & \mathbb{P} \left\{ \int_0^1 \left(\frac{|Z_n(t)|}{\widehat{\sigma}_n} \right)^p \frac{1}{q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt \leq x, \right. \\ & \quad \left. \frac{1}{\sqrt{2C_r \log n}} \left[\int_0^1 \frac{|Z_n(t)|^r}{\widehat{\sigma}_n^r \left(\frac{\lfloor (n+1)t \rfloor}{n}\right) \left(1 - \frac{\lfloor (n+1)t \rfloor}{n}\right)^{1+r/2}} dt - 2m_r \log n \right] \leq y \right\} \\ & \longrightarrow \mathbb{P} \left\{ \int_0^1 \frac{|B(t)|^p}{q(t)} dt \leq x \right\} \Phi(y), \end{aligned} \quad (16)$$

where Φ is the standard normal distribution function, and m_r and C_r are defined as in [Theorem 2.4](#).

Remark 5. The joint limits in [Theorem 2.5](#) can be used to construct combined test statistics from the two component statistics appearing in each part of the theorem. For example, in the supremum case, if S_n^q denotes the weighted supremum statistic in [Theorem 2.1](#) and D_n denotes the Darling–Erdős-type statistic in [Theorem 2.3](#), then, for fixed constants a, b , the joint limit implies

$$aS_n^q + bD_n \xrightarrow{d} aS^q + bG,$$

where

$$S^q := \sup_{0 < t < 1} \frac{|B(t)|}{q(t)}$$

and S^q and G are independent. Similar combined statistics can be formed by taking linear combinations, maxima, or other continuous functions. The critical values for any such combined statistic should therefore be obtained from the distribution of the corresponding function applied to the limiting variables appearing in the joint limit, rather than by using the marginal critical values separately. Such critical values can be approximated numerically; see Gallanosa and Martsynyuk ([2026](#)) for an example of this type of simulation-based approximation for the critical values of supremum and integral statistics of the forms in [Theorems 2.1](#) and [2.2](#).

Finally, we recall a Rényi-type result for a trimmed supremum statistic. Whereas the Darling–Erdős statistic divides by the pointwise standardizing factor $\{t(1-t)\}^{1/2}$ over the

whole interval, the Rényi-type statistic below divides by $t(1-t)$, giving stronger emphasis to values near the endpoints. The trimming excludes small neighborhoods near 0 and 1, where this stronger endpoint weighting would otherwise dominate the statistic, but these neighborhoods shrink as $n \rightarrow \infty$. Thus, asymptotically, the statistic still covers possible change-point locations throughout $(0, 1)$. After the corresponding normalization, the limit is the supremum functional of a Wiener process rather than a weighted Brownian bridge.

Theorem 2.6. (Martsynyuk, 2016; Theorem 1.1). Let $\{Z, Z_i\}_{i \geq 1}$ be i.i.d. random variables such that $0 < \text{Var}(Z) < \infty$ and

$$\mathbb{E}|Z|^r < \infty \quad \text{for some } r > 2.$$

Let $\{d_n\}_{n \geq 1}$ and $\{e_n\}_{n \geq 1}$ be sequences of positive constants satisfying

$$\left\{ \begin{array}{ll} d_n \left(1 + \frac{1}{e_n}\right) < 1 & \text{for all } n \geq 1, \\ \frac{d_n}{e_n} \leq \text{const} < 1 & \text{for all sufficiently large } n, \\ d_n \rightarrow 0, \quad e_n \rightarrow 0, & \\ nd_n \rightarrow \infty & \text{as } n \rightarrow \infty. \end{array} \right. \quad (17)$$

Then, as $n \rightarrow \infty$,

$$d_n^{1/2} \sup_{d_n/e_n \leq t \leq 1-d_n} \frac{|Z_n(t)|}{\hat{\sigma}_n \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n}\right)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|, \quad (18)$$

and

$$d_n^{1/2} \sup_{d_n \leq t \leq 1-d_n/e_n} \frac{|Z_n(t)|}{\hat{\sigma}_n \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n}\right)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|, \quad (19)$$

where $\{W(t), 0 \leq t \leq 1\}$ is a standard Wiener process.

Remark 6. Let

$$J_n^{(1)} := \left[\frac{d_n}{e_n}, 1 - d_n \right], \quad J_n^{(2)} := \left[d_n, 1 - \frac{d_n}{e_n} \right].$$

Using the step-function representation in Remark 1, the suprema in (18) and (19) can be written as maxima over the indices whose intervals intersect the corresponding trimmed

range. Namely, for $j = 1, 2$,

$$\sup_{t \in J_n^{(j)}} \frac{|Z_n(t)|}{\hat{\sigma}_n \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n}\right)} = \max_{k \in \mathcal{K}_n^{(j)}} \frac{\left| \sum_{i=1}^k Z_i - \frac{k}{n} \sum_{i=1}^n Z_i \right|}{\sqrt{n} \hat{\sigma}_n \frac{k}{n} \left(1 - \frac{k}{n}\right)},$$

where

$$\mathcal{K}_n^{(j)} := \left\{ k \in \{1, \dots, n-1\} : \left[\frac{k}{n+1}, \frac{k+1}{n+1} \right) \cap J_n^{(j)} \neq \emptyset \right\}.$$

Remark 7. The Rényi-type limit in [Theorem 2.6](#) provides another asymptotic framework for supremum-type statistics. For finite-sample use, the performance of the corresponding test depends on the choice of the trimming sequences d_n and e_n , and suitable choices of these sequences require separate numerical investigation. Horváth et al. (2020) studied a related class of Rényi-type statistics and found strong performance for changes near the beginning or end of the sample when compared with unweighted supremum-type statistics of the form in [Theorem 2.1](#) with $q(t) \equiv 1$, Darling–Erdős-type statistics of the form in [Theorem 2.3](#), and related procedures. On the other hand, the mean-change simulations of Gallanosa and Martynyuk (2026) indicate that tests constructed from the integral statistic in [Theorem 2.2](#), especially for $p = 2$, can have competitive power while maintaining good type I error control. Thus, the finite-sample performance of Rényi-type tests, especially relative to the other supremum- and integral-type procedures recalled above, remains an open question for numerical investigation.

Remark 8. The statistics in the theorems above, stated with the full-sample variance estimator $\hat{\sigma}_n$, can also be formulated with pooled sample variances. For $1 \leq k \leq n$, define

$$\tilde{\sigma}_k^2 := \begin{cases} \frac{1}{n} \left[\sum_{i=1}^k \left(Z_i - \frac{1}{k} \sum_{j=1}^k Z_j \right)^2 + \sum_{i=k+1}^n \left(Z_i - \frac{1}{n-k} \sum_{j=k+1}^n Z_j \right)^2 \right], & 1 \leq k < n, \\ \frac{1}{n} \sum_{i=1}^n \left(Z_i - \frac{1}{n} \sum_{j=1}^n Z_j \right)^2, & k = n. \end{cases} \quad (20)$$

Under the null hypothesis of no change in the mean, the sequence $\{\tilde{\sigma}_k^2, 1 \leq k \leq n\}$ is weakly

uniformly consistent for the common variance, namely

$$\max_{1 \leq k \leq n} |\tilde{\sigma}_k^2 - \sigma^2| = o_P(1),$$

see Lemma 2.1.1 of Csörgő and Horváth (1997). Hence the corresponding statistics, with $\hat{\sigma}_n$ replaced by the appropriate $\tilde{\sigma}_k$ term inside the maximum or integral, have the same asymptotic null limits; see also Corollaries 2.1.1 and 2.1.3 of Csörgő and Horváth (1997).

In the finite-sample study of Gallanosa and Martsynyuk (2026), the $\tilde{\sigma}_k$ -standardized versions of the weighted supremum and integral mean-change tests had somewhat higher power than the corresponding $\hat{\sigma}_n$ -standardized versions. However, this improvement was accompanied by larger type I errors, especially for smaller samples. For this reason, the regression results that follow use full-sample residual variance estimators.

The results in this section are stated for generic i.i.d. variables Z_i . In the regression setting, the role of Z_i will be played by suitable residual-based quantities. The main task in the sections that follow is to show that the corresponding residual-based tied-down partial sums have the same limiting behavior as the generic processes recalled above, with any additional remainder terms being asymptotically negligible.

3 Testing for a Change in the Slope

3.1 Model, residual quantities, and assumptions

We first consider the problem of testing for a possible change in the slope parameter in a simple linear regression model with random design. Under the null hypothesis, the observations satisfy

$$y_i = \alpha + \beta x_i + \varepsilon_i, \quad 1 \leq i \leq n, \quad (21)$$

where $(y_i, x_i) \in \mathbb{R}^2$, α and β are real-valued parameters, and ε_i are error variables.

A possible at-most-one change in the slope is described by

$$y_i = \alpha + \begin{cases} \beta_1 x_i + \varepsilon_i, & 1 \leq i \leq k^*, \\ \beta_2 x_i + \varepsilon_i, & k^* < i \leq n, \end{cases} \quad 1 \leq k^* < n, \quad (22)$$

where the change point k^* is unknown. The corresponding hypotheses are

$$H_0^\beta : \beta_1 = \beta_2 \quad \text{against} \quad H_1^\beta : \beta_1 \neq \beta_2. \quad (23)$$

Let

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y}_n = \frac{1}{n} \sum_{i=1}^n y_i.$$

The least squares estimator of the slope in (21) is

$$\hat{\beta} = \frac{\sum_{i=1}^n (x_i - a\bar{x}_n)(y_i - a\bar{y}_n)}{\sum_{i=1}^n (x_i - a\bar{x}_n)^2}, \quad (24)$$

where

$$a = \begin{cases} 0, & \text{if } \alpha \text{ is known to be } 0, \\ 1, & \text{if } \alpha \text{ is unknown.} \end{cases} \quad (25)$$

For the slope-change problem, define the residual-type quantities

$$\hat{v}_i = (x_i - a\bar{x}_n) \left[(y_i - a\bar{y}_n) - \hat{\beta}(x_i - a\bar{x}_n) \right], \quad 1 \leq i \leq n. \quad (26)$$

The corresponding residual partial sums are

$$\frac{1}{\sqrt{n}} \sum_{i=1}^k \hat{v}_i, \quad 1 \leq k < n,$$

where $\frac{1}{n} \sum_{i=1}^n \hat{v}_i = 0$. These normalized partial sums are the slope-change analogues of the tied-down partial-sum process $Z_n(t)$ in (3). The statistics below are constructed from these residual partial sums in forms parallel to the mean-change statistics recalled in Section 2.

Under H_0^β , the residual partial sums can be decomposed into a leading tied-down partial-

sum term plus additional terms. For $a = 0$, the leading variables are $U_i^{(0)} = x_i \varepsilon_i$, while for $a = 1$, they are $U_i^{(1)} = (x_i - \mathbb{E}x) \varepsilon_i$. The corresponding leading tied-down sums are

$$M_{n,k}^{(a)} = \sum_{i=1}^k U_i^{(a)} - \frac{k}{n} \sum_{i=1}^n U_i^{(a)}, \quad 0 \leq k \leq n.$$

Thus $M_{n,k}^{(a)}$ plays the same role as the centered partial sums that lead to the mean-change process $Z_n(t)$ in (3), and the corresponding normalized slope process is obtained by considering $n^{-1/2} M_{n, \lfloor (n+1)t \rfloor}^{(a)}$. The auxiliary results in Section 3.3 give the exact decompositions of $\sum_{i=1}^k \widehat{v}_i$ and show that the residual-based statistics in the next subsection have the same limiting distributions as the corresponding statistics formed from $M_{n,k}^{(a)}$.

The following assumptions specify the random-design framework used for the asymptotic results. The covariates and errors are modeled through the random vectors (x_i, ε_i) , with independence between the design variable and the regression error.

Assumption 3.1. *The random vectors $\{(x_i, \varepsilon_i)\}_{i \geq 1}$ are i.i.d., and x and ε are independent. Moreover,*

$$\mathbb{E}\varepsilon = 0, \quad 0 < \text{Var}(x) < \infty, \quad 0 < \text{Var}(\varepsilon) < \infty, \quad \mathbb{E}x^4 < \infty.$$

For the integral Darling–Erdős-type, and Rényi-type limits below, we use the following stronger moment assumption.

Assumption 3.2. *The random vectors $\{(x_i, \varepsilon_i)\}_{i \geq 1}$ are i.i.d., and x and ε are independent. Moreover,*

$$0 < \text{Var}(x) < \infty, \quad 0 < \text{Var}(\varepsilon) < \infty, \quad \mathbb{E}\varepsilon = 0,$$

and there exists $r > 2$ such that

$$\mathbb{E}|x|^{2r} < \infty, \quad \mathbb{E}|\varepsilon|^r < \infty.$$

3.2 Main results

We now give the slope-change analogues of the mean-change limit theorems recalled in [Section 2](#). All statistics in this subsection are based on the same scalar residual process formed from the quantities $\hat{v}_i$, but different choices of functional, weighting, and normalization lead to different limiting distributions.

3.2.1 Supremum statistic with Brownian-bridge-based limit

Theorem 3.1. *Let $\hat{v}_i$ be defined by (26), where $\hat{\beta}$ and a are defined in (24) and (25), respectively. Assume that [Assumption 3.1](#) holds and that H_0^β holds. Let $q(t) \in Q$, and suppose that $I_0^1(q, c) < \infty$ for some or possibly all $c > 0$. Then, as $n \rightarrow \infty$,*

$$\frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k \leq n} \frac{\left| \sum_{i=1}^k \hat{v}_i \right|}{\sqrt{n} q(k/n)} \xrightarrow{d} \sup_{0 < t < 1} \frac{|B(t)|}{q(t)}, \quad (27)$$

where $\{B(t), 0 \leq t \leq 1\}$ is a standard Brownian bridge and

$$\hat{\sigma}_{n,\hat{v}}^2 = \frac{1}{n} \sum_{i=1}^n \left(\hat{v}_i - \frac{1}{n} \sum_{j=1}^n \hat{v}_j \right)^2 = \frac{1}{n} \sum_{i=1}^n \hat{v}_i^2, \quad (28)$$

with $\frac{1}{n} \sum_{i=1}^n \hat{v}_i = 0$.

Remark 9. *The statistic in [Theorem 3.1](#) has a natural estimator-comparison interpretation. For the case when $\alpha = 0$, after some straightforward algebra, the partial sums $\sum_{i=1}^k \hat{v}_i$ can be written as*

$$\sum_{i=1}^k \hat{v}_i = \frac{\left(\sum_{i=1}^k x_i^2 \right) \left(\sum_{i=k+1}^n x_i^2 \right)}{\sum_{i=1}^n x_i^2} \left(\hat{\beta}_1^k - \hat{\beta}_2^k \right), \quad (29)$$

where

$$\hat{\beta}_1^k = \frac{\sum_{i=1}^k x_i y_i}{\sum_{i=1}^k x_i^2}, \quad \hat{\beta}_2^k = \frac{\sum_{i=k+1}^n x_i y_i}{\sum_{i=k+1}^n x_i^2} \quad (30)$$

are the least squares estimators of the slope before and after a possible change point at time k . Thus, the statistic

$$\frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k \leq n} \frac{\left| \sum_{i=1}^k \hat{v}_i \right|}{\sqrt{n} q(k/n)} \quad (31)$$

can be interpreted as the maximum, over $1 \leq k < n$, of appropriately normalized and weighted absolute differences between the slope estimators computed before and after k .

This estimator comparison interpretation parallels the nonparametric mean-change methods reviewed in [Section 2](#). In the mean-change setting, large values of the maximum of appropriately normalized and weighted differences between sample means before and after k indicate a change in the mean. Analogously, in the regression setting, large observed values of (31) indicate a possible change in the slope parameter.

3.2.2 Supremum Darling–Erdős-type statistic

Theorem 3.2. Assume that [Assumption 3.1](#) holds, that H_0^β holds, and that

$$\mathbb{E}[\varepsilon^2 \log \log(|\varepsilon| + 1)] < \infty.$$

Let

$$q^*(t) = \sqrt{t(1-t)}, \quad 0 < t < 1.$$

Then, as $n \rightarrow \infty$,

$$A(n) \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \hat{v}_i \right|}{\sqrt{n} q^*(k/n)} - D(n) \xrightarrow{d} G, \quad (32)$$

where $A(n)$, $D(n)$, and G are as defined in [Theorem 2.3](#), and $\hat{\sigma}_{n,\hat{v}}^2$ is defined in (28).

Remark 10. The supremum statistics in [Theorems 3.1](#) and [3.2](#) are related to weighted residual methods for regression change-point testing. In particular, Horváth et al. (2023) studied cumulative sums of weighted residuals in linear regression models, including Brownian-bridge-based supremum statistics and Darling–Erdős-type statistics. In the simple through-origin case $a = 0$, the present slope quantity reduces to $\hat{v}_i = x_i(y_i - \hat{\beta}x_i)$, which is exactly the scalar covariate-weighted fitted residual.

When the intercept is unknown, the relationship is no longer identical. The present statistic uses the centered scalar quantity

$$\hat{v}_i = (x_i - \bar{x}_n) \{ (y_i - \bar{y}_n) - \hat{\beta}(x_i - \bar{x}_n) \},$$

and the corresponding test statistics have scalar limits involving a single Brownian bridge. In contrast, the weighted-residual formulation of Horváth et al. (2023) for simple linear regression with an unknown intercept and slope is based on the two-dimensional weighted-residual process associated with both regression coefficients. Therefore, its limit involves the Euclidean norm of the vector of two independent Brownian bridges. Thus, the through-origin case coincides with the scalar covariate-weighted residual form, whereas the unknown-intercept case gives a centered slope-specific construction whose limit involves a single Brownian bridge rather than a norm of multiple Brownian bridges.

The tied-down partial-sum structure is also connected to the slope-change supremum statistic of Martynyuk (2021) for structural errors-in-variables models. The integral and Rényi-type statistics developed below extend the present slope-change framework beyond the supremum cases.

3.2.3 Integral statistic with Brownian-bridge-based limit

Theorem 3.3. Assume that [Assumption 3.1](#) holds and that H_0^β holds. Let $0 < p < \infty$ and $q(t) \in Q$. Suppose that

$$\int_0^1 \frac{(t(1-t))^{p/2}}{q(t)} dt < \infty. \quad (33)$$

Then, as $n \rightarrow \infty$,

$$\frac{1}{\hat{\sigma}_{n,\hat{v}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{v}_i \right|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt \xrightarrow{d} \int_0^1 \frac{|B(t)|^p}{q(t)} dt, \quad (34)$$

where $\{B(t), 0 \leq t \leq 1\}$ is a standard Brownian bridge and $\hat{\sigma}_{n,\hat{v}}^2$ is defined in (28).

3.2.4 Integral Darling–Erdős-type statistic

Theorem 3.4. Assume that [Assumption 3.2](#) holds and that H_0^β holds. Let $1 \leq p < \infty$.

Then, as $n \rightarrow \infty$,

$$\frac{1}{\sqrt{2C_p \log n}} \left[\frac{1}{\hat{\sigma}_{n,\hat{v}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{v}_i \right|^p}{n^{p/2} \left(\frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right) \right)^{1+p/2}} dt - 2m_p \log n \right] \xrightarrow{d} N(0, 1), \quad (35)$$

where

$$m_p = \int_{-\infty}^{\infty} |x|^p \phi(x) dx,$$

ϕ is the standard normal density, and C_p is defined as in [Theorem 2.4](#).

3.2.5 Joint limit theorem

The marginal limits above can also be combined into joint limits. Here “joint” refers to paired limits of different statistics constructed from the same slope-specific residual process, not to a joint slope–intercept testing procedure. The following result is the slope-change analogue of [Theorem 2.5](#). One joint limit combines the supremum-based statistics in [Theorems 3.1](#) and [3.2](#), and the other combines the integral-based statistics in [Theorems 3.3](#) and [3.4](#).

Theorem 3.5. *Let $\widehat{v}_i$ be defined by (26), where $\widehat{\beta}$ and a are defined in (24) and (25), respectively.*

- (i) *Assume that [Assumption 3.1](#) holds and that H_0^β holds. Let $q(t) \in Q$, and suppose that $I_0^1(q, c) < \infty$ for some or possibly all $c > 0$. Assume also that*

$$\mathbb{E}[\varepsilon^2 \log \log(|\varepsilon| + 1)] < \infty.$$

Then, for all $x, y \in \mathbb{R}$, as $n \rightarrow \infty$,

$$\begin{aligned} & \mathbb{P} \left\{ \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k \leq n} \frac{\left| \sum_{i=1}^k \widehat{v}_i \right|}{\sqrt{n} q(k/n)} \leq x, \right. \\ & \quad \left. A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k \leq n} \frac{\left| \sum_{i=1}^k \widehat{v}_i \right|}{\sqrt{n} q^*(k/n)} - D(n) \leq y \right\} \\ & \longrightarrow \mathbb{P} \left\{ \sup_{0 < t < 1} \frac{|B(t)|}{q(t)} \leq x \right\} \exp(-2e^{-y}), \end{aligned} \tag{36}$$

where $q^(t) = \sqrt{t(1-t)}$, and $A(n)$, $D(n)$, and G are as defined in [Theorem 2.3](#).*

- (ii) *Assume that [Assumption 3.2](#) holds and that H_0^β holds. Let $0 < p < \infty$, let $1 \leq \ell < \infty$,*

and let $q(t) \in Q$. Suppose that

$$\int_0^1 \frac{(t(1-t))^{p/2}}{q(t)} dt < \infty. \quad (37)$$

Then, for all $x, y \in \mathbb{R}$, as $n \rightarrow \infty$,

$$\begin{aligned} & \mathbb{P} \left\{ \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{v}_i \right|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt \leq x, \right. \\ & \quad \left. \frac{1}{\sqrt{2C_\ell \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^\ell} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{v}_i \right|^\ell}{n^{\ell/2} \left(\frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right) \right)^{1+\ell/2}} dt - 2m_\ell \log n \right] \leq y \right\} \\ & \rightarrow \mathbb{P} \left\{ \int_0^1 \frac{|B(t)|^p}{q(t)} dt \leq x \right\} \Phi(y), \end{aligned} \quad (38)$$

where Φ is the standard normal distribution function, and m_ℓ and C_ℓ are defined as in [Theorem 2.4](#).

3.2.6 Rényi-type weighted supremum statistic

Theorem 3.6. Assume that [Assumption 3.2](#) holds and that H_0^β holds. Let $\{d_n\}_{n \geq 1}$ and $\{e_n\}_{n \geq 1}$ be sequences of positive constants satisfying

$$\left\{ \begin{array}{ll} d_n \left(1 + \frac{1}{e_n} \right) < 1 & \text{for all } n \geq 1, \\ \frac{d_n}{e_n} \leq \text{const} < 1 & \text{for all sufficiently large } n, \\ d_n \rightarrow 0, \quad e_n \rightarrow 0, & \\ nd_n \rightarrow \infty & \text{as } n \rightarrow \infty. \end{array} \right. \quad (39)$$

Let

$$J_n^{(1)} := \left[\frac{d_n}{e_n}, 1 - d_n \right], \quad J_n^{(2)} := \left[d_n, 1 - \frac{d_n}{e_n} \right].$$

Then, as $n \rightarrow \infty$,

$$d_n^{1/2} \sup_{t \in J_n^{(1)}} \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{v}_i \right|}{\widehat{\sigma}_{n,\widehat{v}} \sqrt{n} \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|, \quad (40)$$

and

$$d_n^{1/2} \sup_{t \in J_n^{(2)}} \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{v}_i \right|}{\widehat{\sigma}_{n,\widehat{v}} \sqrt{n} \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n}\right)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|, \quad (41)$$

where $\widehat{\sigma}_{n,\widehat{v}}^2$ is given by (28), and $\{W(t), 0 \leq t \leq 1\}$ is a standard Wiener process.

3.3 Proofs of the Slope-Change Results

Auxiliary results for slope-change statistics

Throughout this subsection, let

$$\widetilde{x}_{i,n} := x_i - a\bar{x}_n, \quad 1 \leq i \leq n, \quad (42)$$

where $a \in \{0, 1\}$ is defined in (25). Then

$$\widehat{v}_i = \widetilde{x}_{i,n} \left[(y_i - a\bar{y}_n) - \widehat{\beta} \widetilde{x}_{i,n} \right], \quad 1 \leq i \leq n.$$

Lemma 3.1. *We have*

$$\sum_{i=1}^n \widehat{v}_i = 0. \quad (43)$$

Consequently,

$$\widehat{\sigma}_{n,\widehat{v}}^2 = \frac{1}{n} \sum_{i=1}^n \left(\widehat{v}_i - \frac{1}{n} \sum_{j=1}^n \widehat{v}_j \right)^2 = \frac{1}{n} \sum_{i=1}^n \widehat{v}_i^2. \quad (44)$$

Proof. By the definition of $\widehat{v}_i$,

$$\begin{aligned} \sum_{i=1}^n \widehat{v}_i &= \sum_{i=1}^n \widetilde{x}_{i,n} \left[(y_i - a\bar{y}_n) - \widehat{\beta} \widetilde{x}_{i,n} \right] \\ &= \sum_{i=1}^n \widetilde{x}_{i,n} (y_i - a\bar{y}_n) - \widehat{\beta} \sum_{i=1}^n \widetilde{x}_{i,n}^2 \\ &= \sum_{i=1}^n \widetilde{x}_{i,n} (y_i - a\bar{y}_n) - \frac{\sum_{j=1}^n \widetilde{x}_{j,n} (y_j - a\bar{y}_n)}{\sum_{j=1}^n \widetilde{x}_{j,n}^2} \sum_{i=1}^n \widetilde{x}_{i,n}^2 \\ &= 0. \end{aligned}$$

Hence (43) follows, and (44) is immediate. $\square$

For later use, define the leading i.i.d. term

$$U_i^{(a)} := \begin{cases} x_i \varepsilon_i, & a = 0, \\ (x_i - \mathbb{E}x) \varepsilon_i, & a = 1, \end{cases} \quad i \geq 1. \quad (45)$$

Let

$$M_{n,k}^{(a)} := \sum_{i=1}^k U_i^{(a)} - \frac{k}{n} \sum_{i=1}^n U_i^{(a)}, \quad 0 \leq k \leq n, \quad (46)$$

where $\sum_{i=1}^0 U_i^{(a)} := 0$. In particular,

$$M_{n,0}^{(a)} = M_{n,n}^{(a)} = 0.$$

Define the corresponding normalized tied-down partial-sum process by

$$\mathcal{M}_n^{(a)}(t) := \begin{cases} \frac{1}{\sqrt{n}} M_{n, \lfloor (n+1)t \rfloor}^{(a)}, & 0 \leq t < 1, \\ 0, & t = 1. \end{cases} \quad (47)$$

Finally, define

$$\sigma_{v,a}^2 := \text{Var}(U_1^{(a)}) = \begin{cases} \mathbb{E}(x^2) \text{Var}(\varepsilon), & a = 0, \\ \text{Var}(x) \text{Var}(\varepsilon), & a = 1. \end{cases} \quad (48)$$

Lemma 3.2. Under *Assumption 3.1*, for each $a \in \{0, 1\}$, the random variables $\{U_i^{(a)}\}_{i \geq 1}$ are i.i.d. with

$$\mathbb{E}U_1^{(a)} = 0, \quad 0 < \sigma_{v,a}^2 < \infty.$$

Proof. **CASE $a = 0$.** If $a = 0$, then $U_i^{(0)} = x_i \varepsilon_i$. Since x_i and ε_i are independent and $\mathbb{E}\varepsilon = 0$,

$$\mathbb{E}U_1^{(0)} = \mathbb{E}x \mathbb{E}\varepsilon = 0,$$

and

$$\text{Var}(U_1^{(0)}) = \mathbb{E}(x^2 \varepsilon^2) = \mathbb{E}(x^2) \mathbb{E}(\varepsilon^2) = \mathbb{E}(x^2) \text{Var}(\varepsilon).$$

This quantity is finite and positive under [Assumption 3.1](#).

CASE $a = 1$. If $a = 1$, then $U_i^{(1)} = (x_i - \mathbb{E}x)\varepsilon_i$. Similarly,

$$\mathbb{E}U_1^{(1)} = \mathbb{E}(x - \mathbb{E}x)\mathbb{E}\varepsilon = 0,$$

and

$$\text{Var}(U_1^{(1)}) = \mathbb{E}[(x - \mathbb{E}x)^2]\mathbb{E}(\varepsilon^2) = \text{Var}(x)\text{Var}(\varepsilon),$$

which is finite and positive under [Assumption 3.1](#). □

Lemma 3.3. Under [Assumption 3.1](#) and H_0^β ,

$$\sqrt{n}(\hat{\beta} - \beta) = O_p(1). \quad (49)$$

In particular,

$$\hat{\beta} - \beta = o_p(1).$$

Proof. Under H_0^β ,

$$y_i - a\bar{y}_n = \beta(x_i - a\bar{x}_n) + (\varepsilon_i - a\bar{\varepsilon}_n),$$

and therefore

$$\hat{\beta} - \beta = \frac{\sum_{i=1}^n \tilde{x}_{i,n}(\varepsilon_i - a\bar{\varepsilon}_n)}{\sum_{i=1}^n \tilde{x}_{i,n}^2}. \quad (50)$$

CASE $a = 0$. If $a = 0$, then

$$\hat{\beta} - \beta = \frac{\sum_{i=1}^n x_i \varepsilon_i}{\sum_{i=1}^n x_i^2}.$$

The numerator is $O_p(\sqrt{n})$ by the Central Limit Theorem, while

$$\frac{1}{n} \sum_{i=1}^n x_i^2 \xrightarrow{p} \mathbb{E}x^2 > 0.$$

Hence $\sqrt{n}(\hat{\beta} - \beta) = O_p(1)$.

CASE $a = 1$. If $a = 1$, then

$$\hat{\beta} - \beta = \frac{\sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n)}{\sum_{i=1}^n (x_i - \bar{x}_n)^2}.$$

The denominator satisfies

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \xrightarrow{p} \text{Var}(x) > 0.$$

For the numerator, write

$$\sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n) = \sum_{i=1}^n (x_i - \mathbb{E}x)\varepsilon_i - (\bar{x}_n - \mathbb{E}x) \sum_{i=1}^n \varepsilon_i.$$

The first term is $O_p(\sqrt{n})$ by the central limit theorem. Also,

$$\bar{x}_n - \mathbb{E}x = O_p(n^{-1/2}), \quad \sum_{i=1}^n \varepsilon_i = O_p(\sqrt{n}),$$

so the second term is $O_p(1)$. Thus the numerator is $O_p(\sqrt{n})$, and hence $\sqrt{n}(\hat{\beta} - \beta) = O_p(1)$ also when $a = 1$. $\square$

Lemma 3.4. Assume that H_0^β holds. Then the following decompositions hold.

(i) If $a = 0$, then

$$\sum_{i=1}^k \hat{v}_i = M_{n,k}^{(0)} + R_{n,k}^{(0)}, \quad 1 \leq k < n, \quad (51)$$

where

$$R_{n,k}^{(0)} := \left(\frac{k}{n} - \frac{\sum_{i=1}^k x_i^2}{\sum_{i=1}^n x_i^2} \right) \sum_{i=1}^n x_i \varepsilon_i. \quad (52)$$

(ii) If $a = 1$, then

$$\sum_{i=1}^k \hat{v}_i = M_{n,k}^{(1)} - E_{n,k} - F_{n,k} + R_{n,k}^{(1)}, \quad 1 \leq k < n, \quad (53)$$

where

$$E_{n,k} := \bar{\varepsilon}_n \left[\sum_{i=1}^k (x_i - \mathbb{E}x) - \frac{k}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right], \quad (54)$$

$$F_{n,k} := (\bar{x}_n - \mathbb{E}x) \left[\sum_{i=1}^k \varepsilon_i - \frac{k}{n} \sum_{i=1}^n \varepsilon_i \right], \quad (55)$$

and

$$R_{n,k}^{(1)} := \left(\frac{k}{n} - \frac{\sum_{i=1}^k (x_i - \bar{x}_n)^2}{\sum_{i=1}^n (x_i - \bar{x}_n)^2} \right) \sum_{i=1}^n (x_i - \bar{x}_n) (\varepsilon_i - \bar{\varepsilon}_n). \quad (56)$$

Proof. Under H_0^β ,

$$y_i - a\bar{y}_n = \beta \tilde{x}_{i,n} + (\varepsilon_i - a\bar{\varepsilon}_n).$$

Thus

$$\hat{v}_i = (\beta - \hat{\beta}) \tilde{x}_{i,n}^2 + \tilde{x}_{i,n} (\varepsilon_i - a\bar{\varepsilon}_n).$$

Using $\sum_{i=1}^n \hat{v}_i = 0$ from [Lemma 3.1](#), we obtain

$$\beta - \hat{\beta} = - \frac{\sum_{i=1}^n \tilde{x}_{i,n} (\varepsilon_i - a\bar{\varepsilon}_n)}{\sum_{i=1}^n \tilde{x}_{i,n}^2}.$$

Substituting this into the partial sum of $\hat{v}_i$ gives

$$\begin{aligned} \sum_{i=1}^k \hat{v}_i &= \sum_{i=1}^k \tilde{x}_{i,n} (\varepsilon_i - a\bar{\varepsilon}_n) - \frac{\sum_{i=1}^k \tilde{x}_{i,n}^2}{\sum_{i=1}^n \tilde{x}_{i,n}^2} \sum_{i=1}^n \tilde{x}_{i,n} (\varepsilon_i - a\bar{\varepsilon}_n) \\ &= \sum_{i=1}^k \tilde{x}_{i,n} (\varepsilon_i - a\bar{\varepsilon}_n) - \frac{k}{n} \sum_{i=1}^n \tilde{x}_{i,n} (\varepsilon_i - a\bar{\varepsilon}_n) + R_{n,k}, \end{aligned} \quad (57)$$

where

$$R_{n,k} = \left(\frac{k}{n} - \frac{\sum_{i=1}^k \tilde{x}_{i,n}^2}{\sum_{i=1}^n \tilde{x}_{i,n}^2} \right) \sum_{i=1}^n \tilde{x}_{i,n} (\varepsilon_i - a\bar{\varepsilon}_n).$$

CASE $a = 0$. If $a = 0$, then $\tilde{x}_{i,n} = x_i$, and (57) gives (51) with (52).

CASE $a = 1$. If $a = 1$, then $\tilde{x}_{i,n} = x_i - \bar{x}_n$. Writing $m = \mathbb{E}x$, we have

$$(x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n) = (x_i - m)\varepsilon_i - \bar{\varepsilon}_n(x_i - m) - (\bar{x}_n - m)\varepsilon_i + (\bar{x}_n - m)\bar{\varepsilon}_n.$$

The last term is constant in i , so its centered partial-sum contribution is zero. Therefore (57) gives (53) with (54), (55), and (56). $\square$

Lemma 3.5. Under *Assumption 3.1* and H_0^β ,

$$\hat{\sigma}_{n,\hat{v}}^2 \xrightarrow{p} \sigma_{v,a}^2, \quad (58)$$

where $\sigma_{v,a}^2$ is defined in (48).

Proof. By *Lemma 3.1*,

$$\hat{\sigma}_{n,\hat{v}}^2 = \frac{1}{n} \sum_{i=1}^n \hat{v}_i^2.$$

Under H_0^β ,

$$\hat{v}_i = (\beta - \hat{\beta})\tilde{x}_{i,n}^2 + \tilde{x}_{i,n}(\varepsilon_i - a\bar{\varepsilon}_n).$$

CASE $a = 0$. Then $\tilde{x}_{i,n} = x_i$, and

$$\hat{v}_i = (\beta - \hat{\beta})x_i^2 + x_i\varepsilon_i.$$

Hence

$$\hat{\sigma}_{n,\hat{v}}^2 = (\beta - \hat{\beta})^2 \frac{1}{n} \sum_{i=1}^n x_i^4 + 2(\beta - \hat{\beta}) \frac{1}{n} \sum_{i=1}^n x_i^3 \varepsilon_i + \frac{1}{n} \sum_{i=1}^n x_i^2 \varepsilon_i^2. \quad (59)$$

By *Lemma 3.3*, $\hat{\beta} - \beta = o_p(1)$. Also,

$$\frac{1}{n} \sum_{i=1}^n x_i^4 \xrightarrow{p} \mathbb{E}x^4 < \infty, \quad \frac{1}{n} \sum_{i=1}^n x_i^3 \varepsilon_i \xrightarrow{p} \mathbb{E}(x^3 \varepsilon) = 0,$$

and

$$\frac{1}{n} \sum_{i=1}^n x_i^2 \varepsilon_i^2 \xrightarrow{p} \mathbb{E}(x^2) \mathbb{E}(\varepsilon^2) = \mathbb{E}(x^2) \text{Var}(\varepsilon).$$

Therefore,

$$\hat{\sigma}_{n,\hat{v}}^2 \xrightarrow{p} \sigma_{v,0}^2.$$

CASE $a = 1$. Then $\tilde{x}_{i,n} = x_i - \bar{x}_n$, and

$$\hat{v}_i = (\beta - \hat{\beta})(x_i - \bar{x}_n)^2 + (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n).$$

Hence

$$\begin{aligned} \hat{\sigma}_{n,\hat{v}}^2 &= (\beta - \hat{\beta})^2 \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^4 + \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 (\varepsilon_i - \bar{\varepsilon}_n)^2 \\ &\quad + 2(\beta - \hat{\beta}) \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^3 (\varepsilon_i - \bar{\varepsilon}_n). \end{aligned} \quad (60)$$

First term in (60). We first note that

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^4 = O_p(1). \quad (61)$$

Indeed, since

$$x_i - \bar{x}_n = (x_i - \mathbb{E}x) - (\bar{x}_n - \mathbb{E}x),$$

the inequality $|u - v|^4 \leq 8(|u|^4 + |v|^4)$ gives

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^4 \leq \frac{8}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^4 + 8(\bar{x}_n - \mathbb{E}x)^4.$$

By the WLLN,

$$\frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^4 \xrightarrow{p} \mathbb{E}[(x - \mathbb{E}x)^4] < \infty,$$

and $\bar{x}_n - \mathbb{E}x = o_p(1)$. Hence (61) holds. Since $\hat{\beta} - \beta = o_p(1)$, (61) gives

$$(\beta - \hat{\beta})^2 \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^4 = o_p(1). \quad (62)$$

Second term in (60). We show that

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 (\varepsilon_i - \bar{\varepsilon}_n)^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \varepsilon_i^2 + o_p(1). \quad (63)$$

Since

$$(\varepsilon_i - \bar{\varepsilon}_n)^2 = \varepsilon_i^2 + \bar{\varepsilon}_n^2 - 2\varepsilon_i\bar{\varepsilon}_n,$$

we have

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 (\varepsilon_i - \bar{\varepsilon}_n)^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i^2 \\ &\quad + \bar{\varepsilon}_n^2 \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \\ &\quad - 2\bar{\varepsilon}_n \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i. \end{aligned} \tag{64}$$

For the first term on the right-hand side of (64), write

$$x_i - \bar{x}_n = (x_i - \mathbb{E}x) - (\bar{x}_n - \mathbb{E}x).$$

Then

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \varepsilon_i^2 + (\bar{x}_n - \mathbb{E}x)^2 \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 \\ &\quad - 2(\bar{x}_n - \mathbb{E}x) \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \varepsilon_i^2. \end{aligned} \tag{65}$$

By the WLLN and independence of x and ε ,

$$\frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \varepsilon_i^2 \xrightarrow{p} \mathbb{E}[(x - \mathbb{E}x)^2] \mathbb{E}(\varepsilon^2) = \text{Var}(x) \text{Var}(\varepsilon).$$

Also,

$$\frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 \xrightarrow{p} \mathbb{E}\varepsilon^2 < \infty,$$

and

$$\frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \varepsilon_i^2 \xrightarrow{p} \mathbb{E}(x - \mathbb{E}x) \mathbb{E}(\varepsilon^2) = 0.$$

Since $\bar{x}_n - \mathbb{E}x = o_p(1)$, the last two terms in (65) are $o_p(1)$. Hence

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \varepsilon_i^2 + o_p(1). \quad (66)$$

For the second term on the right-hand side of (64),

$$\bar{\varepsilon}_n^2 \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 = o_p(1), \quad (67)$$

because $\bar{\varepsilon}_n = o_p(1)$ and

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \xrightarrow{p} \text{Var}(x).$$

For the third term on the right-hand side of (64), it is enough to show that

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i = O_p(1), \quad (68)$$

since $\bar{\varepsilon}_n = o_p(1)$. Using $x_i - \bar{x}_n = (x_i - \mathbb{E}x) - (\bar{x}_n - \mathbb{E}x)$, we have

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i &= \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \varepsilon_i + (\bar{x}_n - \mathbb{E}x)^2 \bar{\varepsilon}_n \\ &\quad - 2(\bar{x}_n - \mathbb{E}x) \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \varepsilon_i. \end{aligned} \quad (69)$$

By the WLLN and independence,

$$\frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \varepsilon_i \xrightarrow{p} \mathbb{E}[(x - \mathbb{E}x)^2] \mathbb{E}(\varepsilon) = 0,$$

and

$$\frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \varepsilon_i \xrightarrow{p} \mathbb{E}(x - \mathbb{E}x) \mathbb{E}(\varepsilon) = 0.$$

Therefore,

$$\frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \varepsilon_i = O_p(1), \quad \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \varepsilon_i = O_p(1).$$

Since

$$\bar{x}_n - \mathbb{E}x = o_p(1), \quad \bar{\varepsilon}_n = o_p(1),$$

the right-hand side of (69) is $O_p(1)$. Hence (68) holds. Consequently,

$$-2\bar{\varepsilon}_n \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i = o_p(1). \quad (70)$$

Substituting (66), (67), and (70) into (64) proves (63).

Moreover, by the WLLN and independence of x and ε ,

$$\frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \varepsilon_i^2 \xrightarrow{p} \mathbb{E}[(x - \mathbb{E}x)^2] \mathbb{E}[\varepsilon^2] = \text{Var}(x) \text{Var}(\varepsilon).$$

Therefore,

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 (\varepsilon_i - \bar{\varepsilon}_n)^2 \xrightarrow{p} \text{Var}(x) \text{Var}(\varepsilon). \quad (71)$$

Third term in (60). By Cauchy–Schwarz, (61), and (71),

$$\begin{aligned} \left| 2(\beta - \hat{\beta}) \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^3 (\varepsilon_i - \bar{\varepsilon}_n) \right| &\leq 2|\beta - \hat{\beta}| \left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^4 \right)^{1/2} \\ &\quad \times \left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 (\varepsilon_i - \bar{\varepsilon}_n)^2 \right)^{1/2} = o_p(1). \end{aligned} \quad (72)$$

Combining (60), (62), (71), and (72), we obtain

$$\hat{\sigma}_{n,\hat{v}}^2 \xrightarrow{p} \text{Var}(x) \text{Var}(\varepsilon) = \sigma_{v,1}^2.$$

This proves (58). □

3.3.1 Proof of Theorem 3.1

Proof. Assume that Assumption 3.1 holds, that H_0^β holds, and that $I_0^1(q, c) < \infty$ for some or possibly all $c > 0$.

By Lemma 3.2, the random variables $\{U_i^{(a)}\}_{i \geq 1}$ are i.i.d. with

$$\mathbb{E}U_1^{(a)} = 0, \quad 0 < \sigma_{v,a}^2 < \infty.$$

Hence, by [Theorem 2.1](#),

$$\frac{1}{\widehat{\sigma}_{n,U^{(a)}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(a)}|}{\sqrt{n} q(k/n)} \xrightarrow{d} \sup_{0 < t < 1} \frac{|B(t)|}{q(t)}, \quad (73)$$

where

$$\widehat{\sigma}_{n,U^{(a)}}^2 = \frac{1}{n} \sum_{i=1}^n \left(U_i^{(a)} - \frac{1}{n} \sum_{j=1}^n U_j^{(a)} \right)^2 \xrightarrow{p} \sigma_{v,a}^2.$$

Also, by [Lemma 3.5](#),

$$\widehat{\sigma}_{n,\widehat{v}}^2 \xrightarrow{p} \sigma_{v,a}^2.$$

Therefore,

$$\frac{\widehat{\sigma}_{n,U^{(a)}}}{\widehat{\sigma}_{n,\widehat{v}}} \xrightarrow{p} 1. \quad (74)$$

It remains to show that the remainder terms in [Lemma 3.4](#) are negligible after division by $\sqrt{n} q(k/n)$, uniformly over $1 \leq k < n$.

CASE $a = 0$. By [Lemma 3.4\(i\)](#),

$$\sum_{i=1}^k \widehat{v}_i = M_{n,k}^{(0)} + R_{n,k}^{(0)}.$$

We show that

$$\max_{1 \leq k < n} \frac{|R_{n,k}^{(0)}|}{\sqrt{n} q(k/n)} = o_p(1). \quad (75)$$

By [\(52\)](#),

$$\max_{1 \leq k < n} \frac{|R_{n,k}^{(0)}|}{\sqrt{n} q(k/n)} = \frac{|\sum_{i=1}^n x_i \varepsilon_i|}{\sum_{i=1}^n x_i^2} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k x_i^2 - \frac{k}{n} \sum_{i=1}^n x_i^2 \right|}{\sqrt{n} q(k/n)}. \quad (76)$$

The first factor satisfies

$$\frac{|\sum_{i=1}^n x_i \varepsilon_i|}{\sum_{i=1}^n x_i^2} = \frac{\left| \frac{1}{n} \sum_{i=1}^n x_i \varepsilon_i \right|}{\frac{1}{n} \sum_{i=1}^n x_i^2} = o_p(1), \quad (77)$$

because

$$\frac{1}{n} \sum_{i=1}^n x_i \varepsilon_i \xrightarrow{p} \mathbb{E}(x\varepsilon) = \mathbb{E}x \mathbb{E}\varepsilon = 0, \quad \frac{1}{n} \sum_{i=1}^n x_i^2 \xrightarrow{p} \mathbb{E}x^2 > 0.$$

For the maximum in (76), [Theorem 2.1](#) applied to $\{x_i^2\}_{i \geq 1}$ gives

$$\max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k x_i^2 - \frac{k}{n} \sum_{i=1}^n x_i^2 \right|}{\sqrt{n} q(k/n)} = O_p(1), \quad (78)$$

provided $\text{Var}(x^2) > 0$. If $\text{Var}(x^2) = 0$, then x_i^2 is constant a.s., and the maximum in (78) is identically zero. Combining (76), (77), and (78) proves (75).

By [Lemma 3.4\(i\)](#),

$$\max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \widehat{v}_i \right|}{\sqrt{n} q(k/n)} = \max_{1 \leq k < n} \frac{|M_{n,k}^{(0)} + R_{n,k}^{(0)}|}{\sqrt{n} q(k/n)}.$$

Since (75) holds, $R_{n,k}^{(0)}$ is negligible after division by $\sqrt{n} q(k/n)$, uniformly over $1 \leq k < n$.

Combining this with (73) and (74), Slutsky's theorem gives

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \widehat{v}_i \right|}{\sqrt{n} q(k/n)} \xrightarrow{d} \sup_{0 < t < 1} \frac{|B(t)|}{q(t)}$$

when $a = 0$.

CASE $a = 1$. By [Lemma 3.4\(ii\)](#),

$$\sum_{i=1}^k \widehat{v}_i = M_{n,k}^{(1)} - E_{n,k} - F_{n,k} + R_{n,k}^{(1)}.$$

We show that

$$\max_{1 \leq k < n} \frac{|E_{n,k}|}{\sqrt{n} q(k/n)} = o_p(1), \quad (79)$$

$$\max_{1 \leq k < n} \frac{|F_{n,k}|}{\sqrt{n} q(k/n)} = o_p(1), \quad (80)$$

$$\max_{1 \leq k < n} \frac{|R_{n,k}^{(1)}|}{\sqrt{n} q(k/n)} = o_p(1). \quad (81)$$

The term $E_{n,k}$. By (54),

$$\max_{1 \leq k < n} \frac{|E_{n,k}|}{\sqrt{n} q(k/n)} = |\bar{\varepsilon}_n| \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k (x_i - \mathbb{E}x) - \frac{k}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right|}{\sqrt{n} q(k/n)}.$$

The maximum is $O_p(1)$ by Theorem 2.1 applied to $\{x_i - \mathbb{E}x\}_{i \geq 1}$, and

$$\bar{\varepsilon}_n = o_p(1).$$

Therefore (79) holds.

The term $F_{n,k}$. By (55),

$$\max_{1 \leq k < n} \frac{|F_{n,k}|}{\sqrt{n} q(k/n)} = |\bar{x}_n - \mathbb{E}x| \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \varepsilon_i - \frac{k}{n} \sum_{i=1}^n \varepsilon_i \right|}{\sqrt{n} q(k/n)}.$$

The maximum is $O_p(1)$ by Theorem 2.1 applied to $\{\varepsilon_i\}_{i \geq 1}$, and

$$\bar{x}_n - \mathbb{E}x = o_p(1).$$

Therefore (80) holds.

The term $R_{n,k}^{(1)}$. By (56),

$$\max_{1 \leq k < n} \frac{|R_{n,k}^{(1)}|}{\sqrt{n} q(k/n)} = A_n N_n,$$

where

$$A_n := \frac{\left| \sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n) \right|}{\sum_{i=1}^n (x_i - \bar{x}_n)^2}$$

and

$$N_n := \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k (x_i - \bar{x}_n)^2 - \frac{k}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \right|}{\sqrt{n} q(k/n)}.$$

We prove separately that

$$A_n = o_p(1) \tag{82}$$

and

$$N_n = O_p(1). \quad (83)$$

First,

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n) = \frac{1}{n} \sum_{i=1}^n x_i \varepsilon_i - \bar{x}_n \bar{\varepsilon}_n.$$

By the WLLN and independence,

$$\frac{1}{n} \sum_{i=1}^n x_i \varepsilon_i \xrightarrow{p} \mathbb{E}(x\varepsilon) = \mathbb{E}x \mathbb{E}\varepsilon = 0,$$

and

$$\bar{x}_n \bar{\varepsilon}_n \xrightarrow{p} \mathbb{E}x \mathbb{E}\varepsilon = 0.$$

Therefore,

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n) = o_p(1).$$

Also,

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \xrightarrow{p} \text{Var}(x) > 0.$$

Hence

$$A_n = \frac{\left| \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n) \right|}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2} = o_p(1),$$

which proves (82).

Next, we prove (83). Since

$$(x_i - \bar{x}_n)^2 = (x_i - \mathbb{E}x)^2 - 2(\bar{x}_n - \mathbb{E}x)(x_i - \mathbb{E}x) + (\bar{x}_n - \mathbb{E}x)^2,$$

the third term, $(\bar{x}_n - \mathbb{E}x)^2$, cancels exactly in the centered partial-sum expression:

$$\sum_{i=1}^k (\bar{x}_n - \mathbb{E}x)^2 - \frac{k}{n} \sum_{i=1}^n (\bar{x}_n - \mathbb{E}x)^2 = 0.$$

Therefore,

$$\begin{aligned}
 N_n \leq & \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k (x_i - \mathbb{E}x)^2 - \frac{k}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \right|}{\sqrt{n} q(k/n)} \\
 & + 2|\bar{x}_n - \mathbb{E}x| \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k (x_i - \mathbb{E}x) - \frac{k}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right|}{\sqrt{n} q(k/n)}. \tag{84}
 \end{aligned}$$

The first maximum on the right-hand side of (84) is $O_p(1)$ by [Theorem 2.1](#) applied to $\{(x_i - \mathbb{E}x)^2\}_{i \geq 1}$, provided $\text{Var}((x - \mathbb{E}x)^2) > 0$. If $\text{Var}((x - \mathbb{E}x)^2) = 0$, then $(x_i - \mathbb{E}x)^2$ is constant a.s., and the first maximum is identically zero. The second maximum is $O_p(1)$ by [Theorem 2.1](#) applied to $\{x_i - \mathbb{E}x\}_{i \geq 1}$, and

$$\bar{x}_n - \mathbb{E}x = o_p(1).$$

Thus the second term on the right-hand side of (84) is $o_p(1)$, and (83) follows.

Combining (82) and (83), we obtain

$$\max_{1 \leq k < n} \frac{|R_{n,k}^{(1)}|}{\sqrt{n} q(k/n)} = A_n N_n = o_p(1),$$

which proves (81).

By [Lemma 3.4\(ii\)](#),

$$\sum_{i=1}^k \hat{v}_i = M_{n,k}^{(1)} - E_{n,k} - F_{n,k} + R_{n,k}^{(1)}.$$

The bounds (79), (80), and (81) show that $E_{n,k}$, $F_{n,k}$, and $R_{n,k}^{(1)}$ are negligible after division by $\sqrt{n} q(k/n)$, uniformly over $1 \leq k < n$. Combining this with (73) and (74), Slutsky's theorem gives

$$\frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \hat{v}_i \right|}{\sqrt{n} q(k/n)} \xrightarrow{d} \sup_{0 < t < 1} \frac{|B(t)|}{q(t)}$$

when $a = 1$. This completes the proof. $\square$

3.3.2 Proof of Theorem 3.2

Proof. Assume that the conditions of Theorem 3.2 hold and that H_0^β holds. Recall that

$$q^*(t) = \sqrt{t(1-t)}, \quad 0 < t < 1.$$

For $a \in \{0, 1\}$, let $U_i^{(a)}$, $M_{n,k}^{(a)}$, and $\sigma_{v,a}^2$ be defined in (45), (46), and (48). Also define

$$\hat{\sigma}_{n,U^{(a)}}^2 := \frac{1}{n} \sum_{i=1}^n \left(U_i^{(a)} - \frac{1}{n} \sum_{j=1}^n U_j^{(a)} \right)^2.$$

By Lemma 3.2,

$$\mathbb{E}U_1^{(a)} = 0, \quad 0 < \sigma_{v,a}^2 < \infty,$$

and

$$\hat{\sigma}_{n,U^{(a)}}^2 \xrightarrow{p} \sigma_{v,a}^2.$$

Also, by Lemma 3.5,

$$\hat{\sigma}_{n,\hat{v}}^2 \xrightarrow{p} \sigma_{v,a}^2.$$

We first verify the Darling–Erdős moment condition for the leading variables $U_i^{(a)}$. Throughout this verification, logarithms in iterated logarithms are interpreted according to the convention stated before Theorem 2.3. Since $\mathbb{E}x^4 < \infty$ under Assumption 3.1, we have

$$\mathbb{E}[x^2 \log \log(|x| + 1)] < \infty. \tag{85}$$

Indeed, under the logarithm convention, $\log \log(|x| + 1) = o(x^2)$ as $|x| \rightarrow \infty$. Hence there exist constants $R < \infty$ and $C < \infty$ such that

$$x^2 \log \log(|x| + 1) \leq Cx^4, \quad |x| \geq R.$$

On the other hand, $x^2 \log \log(|x| + 1)$ is bounded on $\{|x| \leq R\}$. Therefore, by increasing C

if necessary,

$$x^2 \log \log(|x| + 1) \leq C(1 + x^4), \quad x \in \mathbb{R},$$

and the conclusion follows.

CASE $a = 0$. If $a = 0$, then $U_i^{(0)} = x_i \varepsilon_i$. Since

$$|x\varepsilon| + 1 \leq (|x| + 1)(|\varepsilon| + 1),$$

we have

$$\log(|x\varepsilon| + 1) \leq \log(|x| + 1) + \log(|\varepsilon| + 1).$$

Since $\log(|x| + 1) \geq 1$ and $\log(|\varepsilon| + 1) \geq 1$ under the log convention,

$$\log(|x| + 1) + \log(|\varepsilon| + 1) \leq 2 \log(|x| + 1) \log(|\varepsilon| + 1).$$

Therefore,

$$\begin{aligned} \log \log(|x\varepsilon| + 1) &\leq \log\{2 \log(|x| + 1) \log(|\varepsilon| + 1)\} \\ &\leq \log 2 + \log \log(|x| + 1) + \log \log(|\varepsilon| + 1). \end{aligned}$$

Multiplying by $x^2 \varepsilon^2$ and using the independence of x and ε , we obtain

$$\begin{aligned} &\mathbb{E}[x^2 \varepsilon^2 \log \log(|x\varepsilon| + 1)] \\ &\leq \log 2 \mathbb{E}(x^2) \mathbb{E}(\varepsilon^2) + \mathbb{E}[x^2 \log \log(|x| + 1)] \mathbb{E}(\varepsilon^2) \\ &\quad + \mathbb{E}(x^2) \mathbb{E}[\varepsilon^2 \log \log(|\varepsilon| + 1)] < \infty, \end{aligned}$$

where the second term is finite by (85). Thus

$$\mathbb{E}\left[\{U_1^{(0)}\}^2 \log \log(|U_1^{(0)}| + 1)\right] < \infty. \tag{86}$$

CASE $a = 1$. If $a = 1$, then $U_i^{(1)} = (x_i - \mathbb{E}x)\varepsilon_i$. Let $m = \mathbb{E}x$. Since

$$|x - m| + 1 \leq |x| + |m| + 1 \leq (|m| + 1)(|x| + 1),$$

we have

$$\log(|x - m| + 1) \leq \log(|m| + 1) + \log(|x| + 1).$$

Since $\log(|x| + 1) \geq 1$ under the log convention,

$$\log(|m| + 1) + \log(|x| + 1) \leq \{\log(|m| + 1) + 1\} \log(|x| + 1).$$

Therefore,

$$\begin{aligned} \log \log(|x - m| + 1) &\leq \log(\{\log(|m| + 1) + 1\} \log(|x| + 1)) \\ &\leq \log\{\log(|m| + 1) + 1\} + \log \log(|x| + 1). \end{aligned} \tag{87}$$

Also,

$$(x - m)^2 \leq 2x^2 + 2m^2.$$

By (87),

$$\begin{aligned} (x - m)^2 \log \log(|x - m| + 1) \\ \leq (2x^2 + 2m^2) [\log\{\log(|m| + 1) + 1\} + \log \log(|x| + 1)]. \end{aligned}$$

The condition $\mathbb{E}x^4 < \infty$ and (85) imply

$$\mathbb{E}[(x - m)^2 \log \log(|x - m| + 1)] < \infty.$$

Applying the product argument from the case $a = 0$, with x replaced by $x - m$, yields

$$\mathbb{E}\left[\{U_1^{(1)}\}^2 \log \log(|U_1^{(1)}| + 1)\right] < \infty. \tag{88}$$

Hence, in both cases $a = 0$ and $a = 1$, [Theorem 2.3](#) applies to the i.i.d. sequence $\{U_i^{(a)}\}_{i \geq 1}$. Recalling the definition of $M_{n,k}^{(a)}$ in [\(46\)](#), and noting that

$$\frac{1}{\sqrt{n} q^*(k/n)} = \left(\frac{n}{k(n-k)} \right)^{1/2},$$

we obtain

$$A(n) \frac{1}{\hat{\sigma}_{n,U^{(a)}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(a)}|}{\sqrt{n} q^*(k/n)} - D(n) \xrightarrow{d} G. \quad (89)$$

Equivalently,

$$A(n) \frac{1}{\hat{\sigma}_{n,U^{(a)}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(a)}|}{\sqrt{n} q^*(k/n)} = D(n) + O_p(1). \quad (90)$$

To replace $\hat{\sigma}_{n,U^{(a)}}$ by $\hat{\sigma}_{n,\hat{v}}$, observe that

$$\begin{aligned} & A(n) \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(a)}|}{\sqrt{n} q^*(k/n)} - A(n) \frac{1}{\hat{\sigma}_{n,U^{(a)}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(a)}|}{\sqrt{n} q^*(k/n)} \\ &= A(n) \frac{1}{\hat{\sigma}_{n,U^{(a)}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(a)}|}{\sqrt{n} q^*(k/n)} \left(\frac{\hat{\sigma}_{n,U^{(a)}}}{\hat{\sigma}_{n,\hat{v}}} - 1 \right). \end{aligned} \quad (91)$$

By [\(90\)](#), the first factor on the right-hand side of [\(91\)](#) is $D(n) + O_p(1)$. Thus it is enough to prove

$$D(n) \left(\frac{\hat{\sigma}_{n,U^{(a)}}}{\hat{\sigma}_{n,\hat{v}}} - 1 \right) = o_p(1). \quad (92)$$

It remains to show that the remainder terms in [Lemma 3.4](#) are negligible on the Darling–Erdős scale and that [\(92\)](#) holds.

CASE $a = 0$.

By [Lemma 3.4\(i\)](#),

$$\sum_{i=1}^k \hat{v}_i = M_{n,k}^{(0)} + R_{n,k}^{(0)}, \quad 1 \leq k < n.$$

The remainder term $R_{n,k}^{(0)}$. We show that

$$A(n) \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k < n} \frac{|R_{n,k}^{(0)}|}{\sqrt{n} q^*(k/n)} = o_p(1). \quad (93)$$

By (52), write

$$R_{n,k}^{(0)} = A_{n,0} V_{n,k}^{(0)},$$

where

$$A_{n,0} := \frac{\sum_{i=1}^n x_i \varepsilon_i}{\sum_{i=1}^n x_i^2}, \quad V_{n,k}^{(0)} := \frac{k}{n} \sum_{i=1}^n x_i^2 - \sum_{i=1}^k x_i^2.$$

Then

$$\begin{aligned} A(n) & \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k < n} \frac{|R_{n,k}^{(0)}|}{\sqrt{n} q^*(k/n)} \\ &= \frac{2 \log \log n}{\sqrt{n}} \cdot \frac{1}{\hat{\sigma}_{n,\hat{v}}} \cdot \frac{1}{\frac{1}{n} \sum_{i=1}^n x_i^2} \cdot \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n x_i \varepsilon_i \right| \\ & \quad \times \frac{1}{\sqrt{2 \log \log n}} \max_{1 \leq k < n} \frac{|V_{n,k}^{(0)}|}{\sqrt{n} q^*(k/n)}. \end{aligned} \tag{94}$$

Now

$$\frac{2 \log \log n}{\sqrt{n}} = o(1), \quad \frac{1}{\hat{\sigma}_{n,\hat{v}}} = O_p(1), \quad \frac{1}{\frac{1}{n} \sum_{i=1}^n x_i^2} = O_p(1),$$

by Lemma 3.5 and the WLLN. Also,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n x_i \varepsilon_i = O_p(1)$$

by the central limit theorem. Finally, by (10) in Theorem 2.3, applied to $\{x_i^2\}_{i \geq 1}$,

$$\frac{1}{\sqrt{2 \log \log n}} \max_{1 \leq k < n} \frac{|V_{n,k}^{(0)}|}{\sqrt{n} q^*(k/n)} = O_p(1),$$

provided $\text{Var}(x^2) > 0$. If $\text{Var}(x^2) = 0$, then x_i^2 is constant a.s., and the maximum is identically zero. Thus the last factor in (94) is $O_p(1)$ in either case. Therefore (93) holds.

Variance comparison. We show that

$$D(n) \left(\frac{\hat{\sigma}_{n,U^{(0)}}}{\hat{\sigma}_{n,\hat{v}}} - 1 \right) = o_p(1). \tag{95}$$

Since $D(n) \asymp \log \log n$, it is enough to show that

$$(\log \log n) \left(\hat{\sigma}_{n,U^{(0)}}^2 - \hat{\sigma}_{n,\hat{v}}^2 \right) = o_p(1). \quad (96)$$

Indeed, both $\hat{\sigma}_{n,U^{(0)}}^2$ and $\hat{\sigma}_{n,\hat{v}}^2$ converge in probability to $\sigma_{v,0}^2 > 0$, so

$$\hat{\sigma}_{n,U^{(0)}} - \hat{\sigma}_{n,\hat{v}} = \frac{\hat{\sigma}_{n,U^{(0)}}^2 - \hat{\sigma}_{n,\hat{v}}^2}{\hat{\sigma}_{n,U^{(0)}} + \hat{\sigma}_{n,\hat{v}}},$$

and the denominator is bounded away from zero in probability.

By definition,

$$\hat{\sigma}_{n,U^{(0)}}^2 = \frac{1}{n} \sum_{i=1}^n (U_i^{(0)})^2 - \left(\frac{1}{n} \sum_{i=1}^n U_i^{(0)} \right)^2.$$

Also, by [Lemma 3.1](#) and the $a = 0$ expansion in the proof of [Lemma 3.5](#),

$$\hat{\sigma}_{n,\hat{v}}^2 = \frac{1}{n} \sum_{i=1}^n \left[(\beta - \hat{\beta})x_i^2 + U_i^{(0)} \right]^2.$$

Therefore,

$$\begin{aligned} \hat{\sigma}_{n,U^{(0)}}^2 - \hat{\sigma}_{n,\hat{v}}^2 &= - \left(\frac{1}{n} \sum_{i=1}^n U_i^{(0)} \right)^2 - (\beta - \hat{\beta})^2 \frac{1}{n} \sum_{i=1}^n x_i^4 \\ &\quad - 2(\beta - \hat{\beta}) \frac{1}{n} \sum_{i=1}^n x_i^3 \varepsilon_i. \end{aligned} \quad (97)$$

For the first term in (97),

$$(\log \log n) \left(\frac{1}{n} \sum_{i=1}^n U_i^{(0)} \right)^2 = \frac{\log \log n}{n} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n U_i^{(0)} \right)^2 = o_p(1),$$

because $n^{-1/2} \sum_{i=1}^n U_i^{(0)} = O_p(1)$. For the second term,

$$(\log \log n) (\beta - \hat{\beta})^2 \frac{1}{n} \sum_{i=1}^n x_i^4 = \frac{\log \log n}{n} (\sqrt{n}(\hat{\beta} - \beta))^2 \frac{1}{n} \sum_{i=1}^n x_i^4 = o_p(1),$$

by Lemma 3.3 and the WLLN. For the third term,

$$(\log \log n)(\beta - \hat{\beta}) \frac{1}{n} \sum_{i=1}^n x_i^3 \varepsilon_i = \frac{\log \log n}{\sqrt{n}} (\sqrt{n}(\beta - \hat{\beta})) \left(\frac{1}{n} \sum_{i=1}^n x_i^3 \varepsilon_i \right).$$

Here $\sqrt{n}(\beta - \hat{\beta}) = O_p(1)$, and

$$\frac{1}{n} \sum_{i=1}^n x_i^3 \varepsilon_i \xrightarrow{p} \mathbb{E}(x^3 \varepsilon) = \mathbb{E}(x^3) \mathbb{E}(\varepsilon) = 0.$$

In particular, the last sample average is $O_p(1)$. Hence the third term is $o_p(1)$. Therefore (96) holds, and hence (95) follows.

By Lemma 3.4(i),

$$\max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \hat{v}_i \right|}{\sqrt{n} q^*(k/n)} = \max_{1 \leq k < n} \frac{|M_{n,k}^{(0)} + R_{n,k}^{(0)}|}{\sqrt{n} q^*(k/n)}.$$

The bound (93) shows that $R_{n,k}^{(0)}$ is negligible on the Darling–Erdős scale, and (95) gives the variance replacement on the $D(n)$ -scale. Combining these facts with (89) and applying Slutsky's theorem gives

$$A(n) \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \hat{v}_i \right|}{\sqrt{n} q^*(k/n)} - D(n) \xrightarrow{d} G$$

when $a = 0$.

CASE $a = 1$.

By Lemma 3.4(ii),

$$\sum_{i=1}^k \hat{v}_i = M_{n,k}^{(1)} - E_{n,k} - F_{n,k} + R_{n,k}^{(1)}, \quad 1 \leq k < n.$$

We first show that

$$A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k < n} \frac{|E_{n,k}|}{\sqrt{n} q^*(k/n)} = o_p(1), \quad (98)$$

$$A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k < n} \frac{|F_{n,k}|}{\sqrt{n} q^*(k/n)} = o_p(1), \quad (99)$$

$$A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k < n} \frac{|R_{n,k}^{(1)}|}{\sqrt{n} q^*(k/n)} = o_p(1). \quad (100)$$

The term $E_{n,k}$. By (54),

$$\begin{aligned} & A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k < n} \frac{|E_{n,k}|}{\sqrt{n} q^*(k/n)} \\ &= \frac{2 \log \log n}{\sqrt{n}} \cdot \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \cdot |\sqrt{n} \bar{\varepsilon}_n| \cdot \frac{1}{\sqrt{2 \log \log n}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k (x_i - \mathbb{E}x) - \frac{k}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right|}{\sqrt{n} q^*(k/n)}. \end{aligned}$$

Here

$$\sqrt{n} \bar{\varepsilon}_n = O_p(1), \quad \widehat{\sigma}_{n,\widehat{v}}^{-1} = O_p(1),$$

and the last factor is $O_p(1)$ by (10) in Theorem 2.3. Since

$$\frac{2 \log \log n}{\sqrt{n}} \rightarrow 0,$$

we obtain (98).

The term $F_{n,k}$. Similarly, by (55),

$$\begin{aligned} & A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k < n} \frac{|F_{n,k}|}{\sqrt{n} q^*(k/n)} \\ &= \frac{2 \log \log n}{\sqrt{n}} \cdot \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \cdot |\sqrt{n}(\bar{x}_n - \mathbb{E}x)| \cdot \frac{1}{\sqrt{2 \log \log n}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \varepsilon_i - \frac{k}{n} \sum_{i=1}^n \varepsilon_i \right|}{\sqrt{n} q^*(k/n)}. \end{aligned}$$

Here

$$\sqrt{n}(\bar{x}_n - \mathbb{E}x) = O_p(1), \quad \widehat{\sigma}_{n,\widehat{v}}^{-1} = O_p(1),$$

and the last factor is $O_p(1)$ by (10) in Theorem 2.3. Therefore (99) holds.

The term $R_{n,k}^{(1)}$. By (56), write

$$R_{n,k}^{(1)} = A_{n,1} W_{n,k}^{(1)},$$

where

$$A_{n,1} := \frac{\sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n)}{\sum_{i=1}^n (x_i - \bar{x}_n)^2},$$

and

$$W_{n,k}^{(1)} := \frac{k}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 - \sum_{i=1}^k (x_i - \bar{x}_n)^2.$$

Then

$$\max_{1 \leq k < n} \frac{|R_{n,k}^{(1)}|}{\sqrt{n} q^*(k/n)} = |A_{n,1}| \max_{1 \leq k < n} \frac{|W_{n,k}^{(1)}|}{\sqrt{n} q^*(k/n)}.$$

We first show that

$$A_{n,1} = O_p(n^{-1/2}). \quad (101)$$

Since

$$A_{n,1} = \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n)}{\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \right)},$$

and

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \xrightarrow{p} \text{Var}(x) > 0,$$

it is enough to show that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n) = O_p(1).$$

Indeed,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n) = \frac{1}{\sqrt{n}} \sum_{i=1}^n x_i \varepsilon_i - \sqrt{n} \bar{x}_n \bar{\varepsilon}_n.$$

By the Central Limit Theorem,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n x_i \varepsilon_i = O_p(1).$$

Also,

$$\bar{x}_n \xrightarrow{p} \mathbb{E}x, \quad \sqrt{n} \bar{\varepsilon}_n = O_p(1),$$

so

$$\sqrt{n} \bar{x}_n \bar{\varepsilon}_n = O_p(1).$$

Thus (101) holds.

Next, we show that

$$\max_{1 \leq k < n} \frac{|W_{n,k}^{(1)}|}{\sqrt{n} q^*(k/n)} = O_p(\sqrt{\log \log n}). \quad (102)$$

Using

$$(x_i - \bar{x}_n)^2 = (x_i - \mathbb{E}x)^2 - 2(\bar{x}_n - \mathbb{E}x)(x_i - \mathbb{E}x) + (\bar{x}_n - \mathbb{E}x)^2,$$

the third term, $(\bar{x}_n - \mathbb{E}x)^2$, cancels exactly in the centered partial-sum expression. Hence

$$\begin{aligned} \max_{1 \leq k < n} \frac{|W_{n,k}^{(1)}|}{\sqrt{n} q^*(k/n)} &\leq \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k (x_i - \mathbb{E}x)^2 - \frac{k}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \right|}{\sqrt{n} q^*(k/n)} \\ &\quad + 2|\bar{x}_n - \mathbb{E}x| \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k (x_i - \mathbb{E}x) - \frac{k}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right|}{\sqrt{n} q^*(k/n)}. \end{aligned} \quad (103)$$

By (10) in Theorem 2.3, applied to $\{(x_i - \mathbb{E}x)^2\}_{i \geq 1}$,

$$\max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k (x_i - \mathbb{E}x)^2 - \frac{k}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \right|}{\sqrt{n} q^*(k/n)} = O_p(\sqrt{\log \log n}),$$

if $\text{Var}((x - \mathbb{E}x)^2) > 0$; otherwise the maximum is identically zero. Also, by (10) applied to $\{x_i - \mathbb{E}x\}_{i \geq 1}$,

$$\max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k (x_i - \mathbb{E}x) - \frac{k}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right|}{\sqrt{n} q^*(k/n)} = O_p(\sqrt{\log \log n}),$$

and

$$\bar{x}_n - \mathbb{E}x = O_p(n^{-1/2}).$$

Therefore the second term in (103) is $o_p(1)$, and (102) follows.

Combining (101) and (102), and using $\hat{\sigma}_{n,\hat{v}}^{-1} = O_p(1)$, gives

$$\begin{aligned} A(n) \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k \leq n} \frac{|R_{n,k}^{(1)}|}{\sqrt{n} q^*(k/n)} \\ = O_p(\sqrt{\log \log n}) \cdot O_p(n^{-1/2}) \cdot O_p(\sqrt{\log \log n}) = O_p\left(\frac{\log \log n}{\sqrt{n}}\right) = o_p(1). \end{aligned}$$

Thus (100) follows.

Variance comparison. It remains to prove

$$D(n) \left(\frac{\hat{\sigma}_{n,U^{(1)}}}{\hat{\sigma}_{n,\hat{v}}} - 1 \right) = o_p(1). \quad (104)$$

Since $D(n) \asymp \log \log n$, it is enough to show that

$$(\log \log n) \left(\hat{\sigma}_{n,U^{(1)}}^2 - \hat{\sigma}_{n,\hat{v}}^2 \right) = o_p(1). \quad (105)$$

As before, this implication follows because both variances converge in probability to $\sigma_{v,1}^2 > 0$.

By definition,

$$\hat{\sigma}_{n,U^{(1)}}^2 = \frac{1}{n} \sum_{i=1}^n (U_i^{(1)})^2 - \left(\frac{1}{n} \sum_{i=1}^n U_i^{(1)} \right)^2. \quad (106)$$

Also, from (60),

$$\begin{aligned} \hat{\sigma}_{n,\hat{v}}^2 &= (\beta - \hat{\beta})^2 \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^4 + \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 (\varepsilon_i - \bar{\varepsilon}_n)^2 \\ &\quad + 2(\beta - \hat{\beta}) \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^3 (\varepsilon_i - \bar{\varepsilon}_n). \end{aligned} \quad (107)$$

First, by Lemma 3.3 and (61),

$$(\beta - \hat{\beta})^2 \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^4 = O_p(n^{-1}). \quad (108)$$

Second, by Cauchy–Schwarz, (61), and (71),

$$\begin{aligned} \left| \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^3 (\varepsilon_i - \bar{\varepsilon}_n) \right| &\leq \left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^4 \right)^{1/2} \\ &\quad \times \left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 (\varepsilon_i - \bar{\varepsilon}_n)^2 \right)^{1/2} = O_p(1). \end{aligned} \quad (109)$$

Together with Lemma 3.3, this gives

$$2(\beta - \hat{\beta}) \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^3 (\varepsilon_i - \bar{\varepsilon}_n) = O_p(n^{-1/2}). \quad (110)$$

It remains to compare

$$\frac{1}{n} \sum_{i=1}^n (U_i^{(1)})^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \varepsilon_i^2$$

with

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 (\varepsilon_i - \bar{\varepsilon}_n)^2.$$

We prove the sharper comparison

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 (\varepsilon_i - \bar{\varepsilon}_n)^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \varepsilon_i^2 + O_p(n^{-1/2}). \quad (111)$$

As in the proof of (63), write

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 (\varepsilon_i - \bar{\varepsilon}_n)^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i^2 \\ &\quad + \bar{\varepsilon}_n^2 \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \\ &\quad - 2\bar{\varepsilon}_n \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i. \end{aligned} \quad (112)$$

For the first term on the right-hand side of (112),

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \varepsilon_i^2 + (\bar{x}_n - \mathbb{E}x)^2 \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 \\ &\quad - 2(\bar{x}_n - \mathbb{E}x) \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \varepsilon_i^2. \end{aligned} \quad (113)$$

Now

$$\bar{x}_n - \mathbb{E}x = O_p(n^{-1/2}), \quad \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 = O_p(1), \quad \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \varepsilon_i^2 = O_p(1).$$

Therefore,

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \varepsilon_i^2 + O_p(n^{-1/2}). \quad (114)$$

For the second term on the right-hand side of (112),

$$\bar{\varepsilon}_n^2 \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 = O_p(n^{-1}), \quad (115)$$

because

$$\bar{\varepsilon}_n = O_p(n^{-1/2}), \quad \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 = O_p(1).$$

For the third term on the right-hand side of (112), we use (68), namely

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i = O_p(1).$$

Since $\bar{\varepsilon}_n = O_p(n^{-1/2})$, we get

$$-2\bar{\varepsilon}_n \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \varepsilon_i = O_p(n^{-1/2}). \quad (116)$$

Substituting (114), (115), and (116) into (112) proves (111).

Finally,

$$\left(\frac{1}{n} \sum_{i=1}^n U_i^{(1)} \right)^2 = O_p(n^{-1}), \quad (117)$$

because $n^{-1/2} \sum_{i=1}^n U_i^{(1)} = O_p(1)$. Combining (106), (107), (108), (110), (111), and (117), we obtain

$$\hat{\sigma}_{n,U^{(1)}}^2 - \hat{\sigma}_{n,\hat{v}}^2 = O_p(n^{-1/2}).$$

Therefore,

$$(\log \log n) \left(\hat{\sigma}_{n,U^{(1)}}^2 - \hat{\sigma}_{n,\hat{v}}^2 \right) = \frac{\log \log n}{\sqrt{n}} O_p(1) = o_p(1),$$

which proves (105) and hence (104).

By Lemma 3.4(ii), the bounds (98), (99), and (100) show that $E_{n,k}$, $F_{n,k}$, and $R_{n,k}^{(1)}$ are negligible on the Darling–Erdős scale. The variance comparison (104) gives the variance replacement on the $D(n)$ -scale. Combining these facts with (89) and applying Slutsky's theorem gives

$$A(n) \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \hat{v}_i \right|}{\sqrt{n} q^*(k/n)} - D(n) \xrightarrow{d} G$$

when $a = 1$.

The proof is complete. □

3.3.3 Proof of Theorem 3.3

We first record several elementary inequalities that will be used in the integral proofs below.

For $0 < p \leq 1$,

$$||u + v|^p - |u|^p| \leq |v|^p. \tag{118}$$

Indeed, the function

$$f : [0, \infty) \rightarrow [0, \infty), \quad f(s) = s^p,$$

is subadditive; that is,

$$(s + r)^p \leq s^p + r^p, \quad s, r \geq 0.$$

Taking $s = |u|$ and $r = |v|$, and using the triangle inequality,

$$|u + v|^p \leq (|u| + |v|)^p \leq |u|^p + |v|^p.$$

Also, since $u = (u + v) - v$, taking $s = |u + v|$ and $r = |v|$ gives

$$|u|^p \leq (|u + v| + |v|)^p \leq |u + v|^p + |v|^p.$$

Hence (118) follows.

For $p > 1$, we will use the following consequence of the mean value theorem:

$$||u + v|^p - |u|^p| \leq p|v|(|u| + |v|)^{p-1}. \quad (119)$$

To see this, consider the function

$$f : [0, \infty) \rightarrow [0, \infty), \quad f(s) = s^p,$$

whose derivative is

$$f'(s) = ps^{p-1}.$$

Applying the mean value theorem to f between the nonnegative real numbers $|u|$ and $|u + v|$, there exists some ξ between $|u|$ and $|u + v|$ such that

$$f(|u + v|) - f(|u|) = f'(\xi)(|u + v| - |u|).$$

Taking absolute values and substituting $f(s) = s^p$ and $f'(\xi) = p\xi^{p-1}$ gives

$$||u + v|^p - |u|^p| = p\xi^{p-1} ||u + v| - |u||.$$

By the reverse triangle inequality,

$$||u + v| - |u|| \leq |v|,$$

and, since ξ lies between $|u|$ and $|u + v|$,

$$\xi \leq \max\{|u|, |u + v|\} \leq |u| + |v|.$$

Therefore,

$$||u + v|^p - |u|^p| \leq p|v|(|u| + |v|)^{p-1},$$

which proves (119).

Finally, for $m \geq 2$ and $0 < p < \infty$,

$$\left| \sum_{j=1}^m u_j \right|^p \leq C_{p,m} \sum_{j=1}^m |u_j|^p, \quad C_{p,m} := \begin{cases} 1, & 0 < p \leq 1, \\ m^{p-1}, & p > 1. \end{cases} \quad (120)$$

Indeed, for $0 < p \leq 1$, the triangle inequality and the subadditivity of $s \mapsto s^p$ on $[0, \infty)$ give

$$\left| \sum_{j=1}^m u_j \right|^p \leq \left(\sum_{j=1}^m |u_j| \right)^p \leq \sum_{j=1}^m |u_j|^p.$$

For $p > 1$, the triangle inequality and Hölder's inequality give

$$\left| \sum_{j=1}^m u_j \right|^p \leq \left(\sum_{j=1}^m |u_j| \right)^p \leq m^{p-1} \sum_{j=1}^m |u_j|^p.$$

Proof. Suppose that H_0^β holds and that $q(t) \in Q$ satisfies

$$\int_0^1 \frac{(t(1-t))^{p/2}}{q(t)} dt < \infty.$$

By Lemma 3.2 and Lemma 3.5,

$$\widehat{\sigma}_{n,U^{(a)}}^2 \xrightarrow{p} \sigma_{v,a}^2, \quad \widehat{\sigma}_{n,\widehat{v}}^2 \xrightarrow{p} \sigma_{v,a}^2,$$

where $0 < \sigma_{v,a}^2 < \infty$ is defined in (48). Hence

$$\frac{\widehat{\sigma}_{n,U^{(a)}}}{\widehat{\sigma}_{n,\widehat{v}}} \xrightarrow{p} 1. \quad (121)$$

Since Theorem 2.2 applies to the i.i.d. sequence $\{U_i^{(a)}\}_{i \geq 1}$, recalling the definition of $M_{n,k}^{(a)}$

in (46), we have

$$\frac{1}{\widehat{\sigma}_{n,U^{(a)}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(a)}|^p}{n^{p/2}q\left(\frac{\lfloor(n+1)t\rfloor}{n}\right)} dt \xrightarrow{d} \int_0^1 \frac{|B(t)|^p}{q(t)} dt. \quad (122)$$

By (121),

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(a)}|^p}{n^{p/2}q\left(\frac{\lfloor(n+1)t\rfloor}{n}\right)} dt \xrightarrow{d} \int_0^1 \frac{|B(t)|^p}{q(t)} dt. \quad (123)$$

CASE $a = 0$.

By Lemma 3.4(i),

$$\sum_{i=1}^k \widehat{v}_i = M_{n,k}^{(0)} + R_{n,k}^{(0)}, \quad 1 \leq k < n.$$

The remainder term $R_{n,k}^{(0)}$. We show that

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|R_{n,\lfloor(n+1)t\rfloor}^{(0)}|^p}{n^{p/2}q\left(\frac{\lfloor(n+1)t\rfloor}{n}\right)} dt = o_p(1). \quad (124)$$

By (52), write

$$R_{n,k}^{(0)} = A_{n,0} V_{n,k}^{(0)},$$

where

$$A_{n,0} := \frac{\sum_{i=1}^n x_i \varepsilon_i}{\sum_{i=1}^n x_i^2}, \quad V_{n,k}^{(0)} := \frac{k}{n} \sum_{i=1}^n x_i^2 - \sum_{i=1}^k x_i^2.$$

First,

$$A_{n,0} = \frac{\frac{1}{n} \sum_{i=1}^n x_i \varepsilon_i}{\frac{1}{n} \sum_{i=1}^n x_i^2} = o_p(1), \quad (125)$$

because

$$\frac{1}{n} \sum_{i=1}^n x_i \varepsilon_i \xrightarrow{p} \mathbb{E}(x\varepsilon) = \mathbb{E}x \mathbb{E}\varepsilon = 0, \quad \frac{1}{n} \sum_{i=1}^n x_i^2 \xrightarrow{p} \mathbb{E}x^2 > 0.$$

Next, by Theorem 2.2 applied to the sequence $\{x_i^2\}_{i \geq 1}$,

$$\int_0^1 \frac{|V_{n,\lfloor(n+1)t\rfloor}^{(0)}|^p}{n^{p/2}q\left(\frac{\lfloor(n+1)t\rfloor}{n}\right)} dt = O_p(1), \quad (126)$$

provided $\text{Var}(x^2) > 0$. If $\text{Var}(x^2) = 0$, then x_i^2 is constant a.s., and the integral in (126) is identically zero. Since $\widehat{\sigma}_{n,\widehat{v}}^{-p} = O_p(1)$, combining (125) and (126) proves (124).

Completion of the case $a = 0$.

The bound (124) gives the required integral-scale negligibility of $R_{n,k}^{(0)}$. It remains only to verify that inserting this negligible remainder inside the power does not affect the limiting integral functional. Let

$$\begin{aligned} I_{n,M}^{(0)} &:= \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(0)}|^p}{n^{p/2} q\left(\frac{\lfloor(n+1)t\rfloor}{n}\right)} dt, \\ I_{n,R}^{(0)} &:= \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|R_{n,\lfloor(n+1)t\rfloor}^{(0)}|^p}{n^{p/2} q\left(\frac{\lfloor(n+1)t\rfloor}{n}\right)} dt, \\ I_{n,M+R}^{(0)} &:= \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(0)} + R_{n,\lfloor(n+1)t\rfloor}^{(0)}|^p}{n^{p/2} q\left(\frac{\lfloor(n+1)t\rfloor}{n}\right)} dt. \end{aligned}$$

By (123), $I_{n,M}^{(0)} = O_p(1)$, and by (124), $I_{n,R}^{(0)} = o_p(1)$.

It remains to show that

$$I_{n,M+R}^{(0)} - I_{n,M}^{(0)} = o_p(1).$$

If $0 < p \leq 1$, applying (118) with

$$u = M_{n,\lfloor(n+1)t\rfloor}^{(0)}, \quad v = R_{n,\lfloor(n+1)t\rfloor}^{(0)},$$

gives

$$\left| I_{n,M+R}^{(0)} - I_{n,M}^{(0)} \right| \leq I_{n,R}^{(0)} = o_p(1).$$

If $p > 1$, the function $s \mapsto s^p$ is convex on $[0, \infty)$, and the subadditivity property used for $0 < p \leq 1$ fails. We therefore compare the corresponding L^p norms instead. Define

$$N_{n,M}^{(0)} := \left(I_{n,M}^{(0)} \right)^{1/p}, \quad N_{n,R}^{(0)} := \left(I_{n,R}^{(0)} \right)^{1/p}, \quad N_{n,M+R}^{(0)} := \left(I_{n,M+R}^{(0)} \right)^{1/p}.$$

Minkowski's inequality gives

$$\left| N_{n,M+R}^{(0)} - N_{n,M}^{(0)} \right| \leq N_{n,R}^{(0)} = o_p(1).$$

Since $N_{n,M}^{(0)} = O_p(1)$, it follows that

$$N_{n,M+R}^{(0)} = O_p(1).$$

Applying (119) with

$$u = N_{n,M}^{(0)}, \quad v = N_{n,M+R}^{(0)} - N_{n,M}^{(0)},$$

we obtain

$$\begin{aligned} \left| I_{n,M+R}^{(0)} - I_{n,M}^{(0)} \right| &= \left| \left(N_{n,M+R}^{(0)} \right)^p - \left(N_{n,M}^{(0)} \right)^p \right| \\ &\leq p \left| N_{n,M+R}^{(0)} - N_{n,M}^{(0)} \right| \left(N_{n,M}^{(0)} + \left| N_{n,M+R}^{(0)} - N_{n,M}^{(0)} \right| \right)^{p-1} \\ &= o_p(1), \end{aligned}$$

since $N_{n,M}^{(0)} = O_p(1)$ and $N_{n,M+R}^{(0)} - N_{n,M}^{(0)} = o_p(1)$.

Thus, for all $0 < p < \infty$,

$$I_{n,M+R}^{(0)} - I_{n,M}^{(0)} = o_p(1). \tag{127}$$

Combining (123) with (127), and using Lemma 3.4(i), gives

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{v}_i \right|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt \xrightarrow{d} \int_0^1 \frac{|B(t)|^p}{q(t)} dt$$

when $a = 0$.

CASE $a = 1$.

By Lemma 3.4(ii),

$$\sum_{i=1}^k \widehat{v}_i = M_{n,k}^{(1)} - E_{n,k} - F_{n,k} + R_{n,k}^{(1)}, \quad 1 \leq k < n.$$

Since the integral functional based on $M_{n,k}^{(1)}$ already has the desired limit by (123), it remains to show that the three additional terms are asymptotically negligible. We first establish their individual integral-scale negligibility:

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|E_{n,\lfloor(n+1)t\rfloor}|^p}{n^{p/2}q\left(\frac{\lfloor(n+1)t\rfloor}{n}\right)} dt = o_p(1), \quad (128)$$

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|F_{n,\lfloor(n+1)t\rfloor}|^p}{n^{p/2}q\left(\frac{\lfloor(n+1)t\rfloor}{n}\right)} dt = o_p(1), \quad (129)$$

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|R_{n,\lfloor(n+1)t\rfloor}^{(1)}|^p}{n^{p/2}q\left(\frac{\lfloor(n+1)t\rfloor}{n}\right)} dt = o_p(1). \quad (130)$$

The term $E_{n,k}$.

By (54),

$$E_{n,k} = \bar{\varepsilon}_n \left[\sum_{i=1}^k (x_i - \mathbb{E}x) - \frac{k}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right].$$

Therefore,

$$\begin{aligned} & \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|E_{n,\lfloor(n+1)t\rfloor}|^p}{n^{p/2}q\left(\frac{\lfloor(n+1)t\rfloor}{n}\right)} dt \\ &= |\bar{\varepsilon}_n|^p \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor(n+1)t\rfloor} (x_i - \mathbb{E}x) - \frac{\lfloor(n+1)t\rfloor}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right|^p}{n^{p/2}q\left(\frac{\lfloor(n+1)t\rfloor}{n}\right)} dt. \end{aligned}$$

Here $\bar{\varepsilon}_n = o_p(1)$, $\widehat{\sigma}_{n,\widehat{v}}^{-p} = O_p(1)$, and the integral factor is $O_p(1)$ by Theorem 2.2 applied to $\{x_i - \mathbb{E}x\}_{i \geq 1}$. Hence (128) holds.

The term $F_{n,k}$.

By (55),

$$F_{n,k} = (\bar{x}_n - \mathbb{E}x) \left[\sum_{i=1}^k \varepsilon_i - \frac{k}{n} \sum_{i=1}^n \varepsilon_i \right].$$

Therefore,

$$\begin{aligned} & \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|F_{n, \lfloor (n+1)t \rfloor}|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt \\ &= |\bar{x}_n - \mathbb{E}x|^p \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \varepsilon_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \varepsilon_i \right|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt. \end{aligned}$$

Here $\bar{x}_n - \mathbb{E}x = o_p(1)$, $\widehat{\sigma}_{n,\widehat{v}}^{-p} = O_p(1)$, and the integral factor is $O_p(1)$ by [Theorem 2.2](#) applied to $\{\varepsilon_i\}_{i \geq 1}$. Hence (129) holds.

The term $R_{n,k}^{(1)}$.

By (56), write

$$R_{n,k}^{(1)} = A_{n,1} W_{n,k}^{(1)},$$

where

$$A_{n,1} := \frac{\sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n)}{\sum_{i=1}^n (x_i - \bar{x}_n)^2},$$

and

$$W_{n,k}^{(1)} := \frac{k}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 - \sum_{i=1}^k (x_i - \bar{x}_n)^2.$$

First,

$$A_{n,1} = o_p(1). \tag{131}$$

Indeed,

$$A_{n,1} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n)}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2}.$$

The numerator satisfies

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n) = \frac{1}{n} \sum_{i=1}^n x_i \varepsilon_i - \bar{x}_n \bar{\varepsilon}_n = o_p(1),$$

while

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \xrightarrow{p} \text{Var}(x) > 0.$$

This proves (131).

Next, we show that

$$\int_0^1 \frac{|W_{n, \lfloor (n+1)t \rfloor}^{(1)}|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt = O_p(1). \quad (132)$$

Using

$$(x_i - \bar{x}_n)^2 = (x_i - \mathbb{E}x)^2 - 2(\bar{x}_n - \mathbb{E}x)(x_i - \mathbb{E}x) + (\bar{x}_n - \mathbb{E}x)^2,$$

the third term, $(\bar{x}_n - \mathbb{E}x)^2$, cancels exactly in the centered partial-sum expression. Thus $W_{n,k}^{(1)}$ is the sum of the two corresponding centered partial-sum terms. Hence, by (120) with $m = 2$,

$$\begin{aligned} & \int_0^1 \frac{|W_{n, \lfloor (n+1)t \rfloor}^{(1)}|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt \\ & \leq C_{p,2} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} (x_i - \mathbb{E}x)^2 - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \right|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt \\ & \quad + C_{p,2} 2^p |\bar{x}_n - \mathbb{E}x|^p \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} (x_i - \mathbb{E}x) - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt. \end{aligned} \quad (133)$$

The first integral on the right-hand side of (133) is $O_p(1)$ by Theorem 2.2 applied to $\{(x_i - \mathbb{E}x)^2\}_{i \geq 1}$, provided $\text{Var}((x - \mathbb{E}x)^2) > 0$; otherwise it is identically zero. The second integral is $O_p(1)$ by Theorem 2.2 applied to $\{x_i - \mathbb{E}x\}_{i \geq 1}$, and

$$\bar{x}_n - \mathbb{E}x = o_p(1).$$

Thus (132) holds.

Combining (131) and (132), and using $\hat{\sigma}_{n,\hat{v}}^{-p} = O_p(1)$, proves (130).

Completion of the case $a = 1$.

The bounds (128), (129), and (130) show that each of the three additional terms in Lemma 3.4(ii) is negligible on the integral scale. We first combine these terms into a single remainder.

Define

$$H_{n,k}^{(1)} := -E_{n,k} - F_{n,k} + R_{n,k}^{(1)}.$$

Then

$$\sum_{i=1}^k \widehat{v}_i = M_{n,k}^{(1)} + H_{n,k}^{(1)}.$$

By (120) with $m = 3$,

$$|H_{n,k}^{(1)}|^p \leq C_{p,3} \left(|E_{n,k}|^p + |F_{n,k}|^p + |R_{n,k}^{(1)}|^p \right).$$

Consequently,

$$\begin{aligned} & \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|H_{n,\lfloor (n+1)t \rfloor}^{(1)}|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt \\ & \leq C_{p,3} \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|E_{n,\lfloor (n+1)t \rfloor}|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt \\ & \quad + C_{p,3} \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|F_{n,\lfloor (n+1)t \rfloor}|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt \\ & \quad + C_{p,3} \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|R_{n,\lfloor (n+1)t \rfloor}^{(1)}|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt = o_p(1), \end{aligned}$$

where the last step follows from (128), (129), and (130). Thus

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|H_{n,\lfloor (n+1)t \rfloor}^{(1)}|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt = o_p(1). \quad (134)$$

It remains to verify that inserting this combined remainder inside the p th power does not affect the limiting integral functional. Let

$$\begin{aligned} I_{n,M}^{(1)} &:= \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(1)}|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt, \\ I_{n,H}^{(1)} &:= \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|H_{n,\lfloor (n+1)t \rfloor}^{(1)}|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt, \\ I_{n,M+H}^{(1)} &:= \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(1)} + H_{n,\lfloor (n+1)t \rfloor}^{(1)}|^p}{n^{p/2} q \left(\frac{\lfloor (n+1)t \rfloor}{n} \right)} dt. \end{aligned}$$

By (123), $I_{n,M}^{(1)} = O_p(1)$, and by (134), $I_{n,H}^{(1)} = o_p(1)$.

It remains to show that

$$I_{n,M+H}^{(1)} - I_{n,M}^{(1)} = o_p(1).$$

If $0 < p \leq 1$, applying (118) with

$$u = M_{n,[(n+1)t]}^{(1)}, \quad v = H_{n,[(n+1)t]}^{(1)},$$

gives

$$\left| I_{n,M+H}^{(1)} - I_{n,M}^{(1)} \right| \leq I_{n,H}^{(1)} = o_p(1).$$

If $p > 1$, we compare the corresponding L^p norms. Define

$$N_{n,M}^{(1)} := \left(I_{n,M}^{(1)} \right)^{1/p}, \quad N_{n,H}^{(1)} := \left(I_{n,H}^{(1)} \right)^{1/p}, \quad N_{n,M+H}^{(1)} := \left(I_{n,M+H}^{(1)} \right)^{1/p}.$$

Minkowski's inequality gives

$$\left| N_{n,M+H}^{(1)} - N_{n,M}^{(1)} \right| \leq N_{n,H}^{(1)} = o_p(1).$$

Since $N_{n,M}^{(1)} = O_p(1)$, it follows that

$$N_{n,M+H}^{(1)} = O_p(1).$$

Applying (119) with

$$u = N_{n,M}^{(1)}, \quad v = N_{n,M+H}^{(1)} - N_{n,M}^{(1)},$$

we obtain

$$\begin{aligned} \left| I_{n,M+H}^{(1)} - I_{n,M}^{(1)} \right| &= \left| \left(N_{n,M+H}^{(1)} \right)^p - \left(N_{n,M}^{(1)} \right)^p \right| \\ &\leq p \left| N_{n,M+H}^{(1)} - N_{n,M}^{(1)} \right| \left(N_{n,M}^{(1)} + \left| N_{n,M+H}^{(1)} - N_{n,M}^{(1)} \right| \right)^{p-1} \\ &= o_p(1), \end{aligned}$$

since $N_{n,M}^{(1)} = O_p(1)$ and $N_{n,M+H}^{(1)} - N_{n,M}^{(1)} = o_p(1)$.

Thus, for all $0 < p < \infty$,

$$I_{n,M+H}^{(1)} - I_{n,M}^{(1)} = o_p(1). \quad (135)$$

Combining (123) with (135), and using Lemma 3.4(ii), gives

$$\frac{1}{\hat{\sigma}_{n,\hat{v}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{v}_i \right|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt \xrightarrow{d} \int_0^1 \frac{|B(t)|^p}{q(t)} dt$$

when $a = 1$. This completes the proof. $\square$

3.3.4 Proof of Theorem 3.4

Proof. Assume that Assumption 3.2 holds and that H_0^β holds. Let

$$\Delta_n(t) := \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right).$$

By Lemma 3.2 and Lemma 3.5,

$$\hat{\sigma}_{n,U^{(a)}}^2 \xrightarrow{p} \sigma_{v,a}^2, \quad \hat{\sigma}_{n,\hat{v}}^2 \xrightarrow{p} \sigma_{v,a}^2,$$

where $0 < \sigma_{v,a}^2 < \infty$ is defined in (48).

We first verify that Theorem 2.4 applies to the leading variables $\{U_i^{(a)}\}_{i \geq 1}$. By Assumption 3.2, there exists $r > 2$ such that

$$\mathbb{E}|x|^{2r} < \infty, \quad \mathbb{E}|\varepsilon|^r < \infty.$$

CASE $a = 0$. If $a = 0$, then $U_i^{(0)} = x_i \varepsilon_i$. By independence,

$$\mathbb{E}|U_1^{(0)}|^r = \mathbb{E}|x\varepsilon|^r = \mathbb{E}|x|^r \mathbb{E}|\varepsilon|^r < \infty.$$

CASE $a = 1$. If $a = 1$, then $U_i^{(1)} = (x_i - \mathbb{E}x)\varepsilon_i$. Since $\mathbb{E}|x|^{2r} < \infty$, we also have

$$\mathbb{E}|x - \mathbb{E}x|^r < \infty.$$

Therefore, by independence,

$$\mathbb{E}|U_1^{(1)}|^r = \mathbb{E}|x - \mathbb{E}x|^r \mathbb{E}|\varepsilon|^r < \infty.$$

Thus, in both cases, $\mathbb{E}|U_1^{(a)}|^\nu < \infty$ for some $\nu > 2$.

Hence, [Theorem 2.4](#) applies to the i.i.d. sequence $\{U_i^{(a)}\}_{i \geq 1}$. Recalling the definition of $M_{n,k}^{(a)}$ in [\(46\)](#), we obtain

$$\frac{1}{\sqrt{2C_p \log n}} \left[\frac{1}{\widehat{\sigma}_{n,U^{(a)}}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(a)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt - 2m_p \log n \right] \xrightarrow{d} N(0, 1). \quad (136)$$

Equivalently,

$$\frac{1}{\widehat{\sigma}_{n,U^{(a)}}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(a)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt = 2m_p \log n + O_p(\sqrt{\log n}). \quad (137)$$

We next replace $\widehat{\sigma}_{n,U^{(a)}}$ by $\widehat{\sigma}_{n,\widehat{v}}$. We show that

$$\sqrt{\log n} \left(\frac{\widehat{\sigma}_{n,U^{(a)}}}{\widehat{\sigma}_{n,\widehat{v}}} - 1 \right) = o_p(1). \quad (138)$$

For $a = 0$, [\(97\)](#) gives

$$\widehat{\sigma}_{n,U^{(0)}}^2 - \widehat{\sigma}_{n,\widehat{v}}^2 = O_p(n^{-1/2}).$$

For $a = 1$, combining [\(106\)](#), [\(107\)](#), [\(108\)](#), [\(110\)](#), [\(111\)](#), and [\(117\)](#) gives

$$\widehat{\sigma}_{n,U^{(1)}}^2 - \widehat{\sigma}_{n,\widehat{v}}^2 = O_p(n^{-1/2}).$$

Thus, for both $a = 0$ and $a = 1$,

$$\hat{\sigma}_{n,U^{(a)}}^2 - \hat{\sigma}_{n,\hat{v}}^2 = O_p(n^{-1/2}). \quad (139)$$

Since both $\hat{\sigma}_{n,U^{(a)}}^2$ and $\hat{\sigma}_{n,\hat{v}}^2$ converge in probability to $\sigma_{v,a}^2 > 0$,

$$\hat{\sigma}_{n,U^{(a)}} - \hat{\sigma}_{n,\hat{v}} = \frac{\hat{\sigma}_{n,U^{(a)}}^2 - \hat{\sigma}_{n,\hat{v}}^2}{\hat{\sigma}_{n,U^{(a)}} + \hat{\sigma}_{n,\hat{v}}} = O_p(n^{-1/2}).$$

Therefore,

$$\frac{\hat{\sigma}_{n,U^{(a)}}}{\hat{\sigma}_{n,\hat{v}}} - 1 = O_p(n^{-1/2}),$$

and (138) follows.

Since $p \geq 1$ is fixed and

$$\frac{\hat{\sigma}_{n,U^{(a)}}}{\hat{\sigma}_{n,\hat{v}}} \xrightarrow{p} 1,$$

(138) also gives

$$\sqrt{\log n} \left[\left(\frac{\hat{\sigma}_{n,U^{(a)}}}{\hat{\sigma}_{n,\hat{v}}} \right)^p - 1 \right] = o_p(1). \quad (140)$$

Moreover,

$$\begin{aligned} & \frac{1}{\hat{\sigma}_{n,\hat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(a)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt - \frac{1}{\hat{\sigma}_{n,U^{(a)}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(a)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt \\ &= \left[\left(\frac{\hat{\sigma}_{n,U^{(a)}}}{\hat{\sigma}_{n,\hat{v}}} \right)^p - 1 \right] \frac{1}{\hat{\sigma}_{n,U^{(a)}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(a)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt. \end{aligned} \quad (141)$$

By (137), the second factor on the right-hand side of (141) is $O_p(\log n)$. Together with (140), this gives

$$\frac{1}{\sqrt{\log n}} \left| \frac{1}{\hat{\sigma}_{n,\hat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(a)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt - \frac{1}{\hat{\sigma}_{n,U^{(a)}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(a)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt \right| = o_p(1). \quad (142)$$

Combining (136) and (142), we obtain

$$\frac{1}{\sqrt{2C_p \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(a)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt - 2m_p \log n \right] \xrightarrow{d} N(0, 1). \quad (143)$$

Equivalently,

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(a)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt = 2m_p \log n + O_p(\sqrt{\log n}). \quad (144)$$

CASE $a = 0$.

By Lemma 3.4(i),

$$\sum_{i=1}^k \widehat{v}_i = M_{n,k}^{(0)} + R_{n,k}^{(0)}, \quad 1 \leq k < n.$$

The remainder term $R_{n,k}^{(0)}$. We show that

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|R_{n,\lfloor (n+1)t \rfloor}^{(0)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt = O_p(n^{-p/2} \log n). \quad (145)$$

By (52), write

$$R_{n,k}^{(0)} = A_{n,0} V_{n,k}^{(0)},$$

where

$$A_{n,0} := \frac{\sum_{i=1}^n x_i \varepsilon_i}{\sum_{i=1}^n x_i^2}, \quad V_{n,k}^{(0)} := \frac{k}{n} \sum_{i=1}^n x_i^2 - \sum_{i=1}^k x_i^2.$$

First,

$$A_{n,0} = O_p(n^{-1/2}). \quad (146)$$

Indeed,

$$A_{n,0} = \frac{n^{-1/2} \sum_{i=1}^n x_i \varepsilon_i}{\sqrt{n} n^{-1} \sum_{i=1}^n x_i^2},$$

where

$$n^{-1/2} \sum_{i=1}^n x_i \varepsilon_i = O_p(1), \quad n^{-1} \sum_{i=1}^n x_i^2 \xrightarrow{p} \mathbb{E} x^2 > 0.$$

Next, by [Theorem 2.4](#) applied to the sequence $\{x_i^2\}_{i \geq 1}$,

$$\int_0^1 \frac{|V_{n, \lfloor (n+1)t \rfloor}^{(0)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt = O_p(\log n), \quad (147)$$

provided $\text{Var}(x^2) > 0$. If $\text{Var}(x^2) = 0$, then x_i^2 is constant a.s., and the integral in (147) is identically zero. Since $\widehat{\sigma}_{n, \widehat{v}}^{-p} = O_p(1)$, combining (146) and (147) proves (145).

Completion of the case $a = 0$.

The bound (145) gives the required size of the remainder contribution on the integral Darling–Erdős scale. It remains only to verify that inserting this remainder inside the power changes the statistic by $o_p(\sqrt{\log n})$. Let

$$\begin{aligned} I_{n,M}^{(0)} &:= \frac{1}{\widehat{\sigma}_{n, \widehat{v}}^p} \int_0^1 \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(0)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt, \\ I_{n,R}^{(0)} &:= \frac{1}{\widehat{\sigma}_{n, \widehat{v}}^p} \int_0^1 \frac{|R_{n, \lfloor (n+1)t \rfloor}^{(0)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt, \\ I_{n,M+R}^{(0)} &:= \frac{1}{\widehat{\sigma}_{n, \widehat{v}}^p} \int_0^1 \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(0)} + R_{n, \lfloor (n+1)t \rfloor}^{(0)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt. \end{aligned}$$

By (144),

$$I_{n,M}^{(0)} = O_p(\log n),$$

and by (145),

$$I_{n,R}^{(0)} = O_p(n^{-p/2} \log n).$$

For $p > 1$, (119) gives

$$||u + v|^p - |u|^p| \leq p|v|(|u| + |v|)^{p-1}.$$

Since

$$(s + r)^{p-1} \leq 2^{p-1} (s^{p-1} + r^{p-1}), \quad s, r \geq 0,$$

we obtain

$$||u + v|^p - |u|^p| \leq p2^{p-1} (|u|^{p-1}|v| + |v|^p).$$

Hence, with $K_p := p2^p$,

$$||u + v|^p - |u|^p| \leq K_p (|u|^{p-1}|v| + |v|^p). \quad (148)$$

If $p = 1$, (118) gives

$$||u + v| - |u|| \leq |v|,$$

and therefore

$$|I_{n,M+R}^{(0)} - I_{n,M}^{(0)}| \leq I_{n,R}^{(0)} = O_p(n^{-1/2} \log n).$$

If $p > 1$, then by (148) and Hölder's inequality,

$$\begin{aligned} |I_{n,M+R}^{(0)} - I_{n,M}^{(0)}| &\leq K_p \left(I_{n,M}^{(0)}\right)^{(p-1)/p} \left(I_{n,R}^{(0)}\right)^{1/p} + K_p I_{n,R}^{(0)} \\ &= O_p\left((\log n)^{(p-1)/p}\right) O_p\left(n^{-1/2}(\log n)^{1/p}\right) + O_p(n^{-p/2} \log n) \\ &= O_p\left(\frac{\log n}{\sqrt{n}}\right) + O_p(n^{-p/2} \log n). \end{aligned} \quad (149)$$

Thus, for all $1 \leq p < \infty$,

$$\frac{I_{n,M+R}^{(0)} - I_{n,M}^{(0)}}{\sqrt{\log n}} = o_p(1). \quad (150)$$

Combining (143) with (150), and using Lemma 3.4(i), gives

$$\frac{1}{\sqrt{2C_p \log n}} \left[\frac{1}{\hat{\sigma}_{n,\hat{v}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{v}_i \right|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt - 2m_p \log n \right] \xrightarrow{d} N(0, 1)$$

when $a = 0$.

CASE $a = 1$.

By Lemma 3.4(ii),

$$\sum_{i=1}^k \hat{v}_i = M_{n,k}^{(1)} - E_{n,k} - F_{n,k} + R_{n,k}^{(1)}, \quad 1 \leq k < n.$$

We show that

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|E_{n,\lfloor(n+1)t\rfloor}|^p}{n^{p/2}\Delta_n(t)^{1+p/2}} dt = O_p(n^{-p/2} \log n), \quad (151)$$

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|F_{n,\lfloor(n+1)t\rfloor}|^p}{n^{p/2}\Delta_n(t)^{1+p/2}} dt = O_p(n^{-p/2} \log n), \quad (152)$$

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|R_{n,\lfloor(n+1)t\rfloor}^{(1)}|^p}{n^{p/2}\Delta_n(t)^{1+p/2}} dt = O_p(n^{-p/2} \log n). \quad (153)$$

The term $E_{n,k}$. By (54),

$$E_{n,k} = \bar{\varepsilon}_n \left[\sum_{i=1}^k (x_i - \mathbb{E}x) - \frac{k}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right].$$

Therefore,

$$\begin{aligned} & \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|E_{n,\lfloor(n+1)t\rfloor}|^p}{n^{p/2}\Delta_n(t)^{1+p/2}} dt \\ &= |\bar{\varepsilon}_n|^p \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor(n+1)t\rfloor} (x_i - \mathbb{E}x) - \frac{\lfloor(n+1)t\rfloor}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right|^p}{n^{p/2}\Delta_n(t)^{1+p/2}} dt. \end{aligned}$$

Here

$$\bar{\varepsilon}_n = O_p(n^{-1/2}), \quad \widehat{\sigma}_{n,\widehat{v}}^{-p} = O_p(1),$$

and the integral factor is $O_p(\log n)$ by Theorem 2.4 applied to $\{x_i - \mathbb{E}x\}_{i \geq 1}$. Hence (151) holds.

The term $F_{n,k}$. By (55),

$$F_{n,k} = (\bar{x}_n - \mathbb{E}x) \left[\sum_{i=1}^k \varepsilon_i - \frac{k}{n} \sum_{i=1}^n \varepsilon_i \right].$$

Therefore,

$$\begin{aligned} & \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|F_{n,\lfloor (n+1)t \rfloor}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt \\ &= |\bar{x}_n - \mathbb{E}x|^p \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \varepsilon_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \varepsilon_i \right|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt. \end{aligned}$$

Here

$$\bar{x}_n - \mathbb{E}x = O_p(n^{-1/2}), \quad \widehat{\sigma}_{n,\widehat{v}}^{-p} = O_p(1),$$

and the integral factor is $O_p(\log n)$ by [Theorem 2.4](#) applied to $\{\varepsilon_i\}_{i \geq 1}$. Hence (152) holds.

The term $R_{n,k}^{(1)}$. By (56), write

$$R_{n,k}^{(1)} = A_{n,1} W_{n,k}^{(1)},$$

where

$$A_{n,1} := \frac{\sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n)}{\sum_{i=1}^n (x_i - \bar{x}_n)^2},$$

and

$$W_{n,k}^{(1)} := \frac{k}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 - \sum_{i=1}^k (x_i - \bar{x}_n)^2.$$

By (101),

$$A_{n,1} = O_p(n^{-1/2}). \tag{154}$$

Recall from (120) that, for $m \geq 2$,

$$\left| \sum_{j=1}^m u_j \right|^p \leq C_{p,m} \sum_{j=1}^m |u_j|^p, \quad C_{p,m} := \begin{cases} 1, & 0 < p \leq 1, \\ m^{p-1}, & p > 1. \end{cases}$$

Next, we show that

$$\int_0^1 \frac{|W_{n,\lfloor (n+1)t \rfloor}^{(1)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt = O_p(\log n). \tag{155}$$

Using

$$(x_i - \bar{x}_n)^2 = (x_i - \mathbb{E}x)^2 - 2(\bar{x}_n - \mathbb{E}x)(x_i - \mathbb{E}x) + (\bar{x}_n - \mathbb{E}x)^2,$$

the last term cancels exactly in the centered partial-sum expression. Hence, by (120) with $m = 2$,

$$\begin{aligned} & \int_0^1 \frac{|W_{n, \lfloor (n+1)t \rfloor}^{(1)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt \\ & \leq C_{p,2} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} (x_i - \mathbb{E}x)^2 - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \right|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt \\ & \quad + C_{p,2} 2^p |\bar{x}_n - \mathbb{E}x|^p \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} (x_i - \mathbb{E}x) - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt. \end{aligned} \quad (156)$$

The first integral on the right-hand side of (156) is $O_p(\log n)$ by Theorem 2.4 applied to $\{(x_i - \mathbb{E}x)^2\}_{i \geq 1}$, provided $\text{Var}((x - \mathbb{E}x)^2) > 0$; otherwise it is identically zero. The second integral is $O_p(\log n)$ by Theorem 2.4 applied to $\{x_i - \mathbb{E}x\}_{i \geq 1}$, and

$$\bar{x}_n - \mathbb{E}x = O_p(n^{-1/2}).$$

Thus (155) holds. Combining (154) and (155), and using $\hat{\sigma}_{n,\hat{v}}^{-p} = O_p(1)$, proves (153).

Completion of the case $a = 1$.

The bounds (151), (152), and (153) give the required size of the additional terms on the integral Darling–Erdős scale. It remains only to verify that inserting their combined contribution inside the power changes the statistic by $o_p(\sqrt{\log n})$. Define

$$H_{n,k}^{(1)} := -E_{n,k} - F_{n,k} + R_{n,k}^{(1)}.$$

Then, by Lemma 3.4(ii),

$$\sum_{i=1}^k \hat{v}_i = M_{n,k}^{(1)} + H_{n,k}^{(1)}.$$

By (120) with $m = 3$,

$$|H_{n,k}^{(1)}|^p \leq C_{p,3} \left(|E_{n,k}|^p + |F_{n,k}|^p + |R_{n,k}^{(1)}|^p \right).$$

Together with (151), (152), and (153), this gives

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|H_{n,[(n+1)t]}^{(1)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt = O_p(n^{-p/2} \log n). \quad (157)$$

Let

$$\begin{aligned} I_{n,M}^{(1)} &:= \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,[(n+1)t]}^{(1)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt, \\ I_{n,H}^{(1)} &:= \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|H_{n,[(n+1)t]}^{(1)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt, \\ I_{n,M+H}^{(1)} &:= \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,[(n+1)t]}^{(1)} + H_{n,[(n+1)t]}^{(1)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt. \end{aligned}$$

By (144),

$$I_{n,M}^{(1)} = O_p(\log n),$$

and by (157),

$$I_{n,H}^{(1)} = O_p(n^{-p/2} \log n).$$

If $p = 1$, applying (118) pointwise with

$$u = M_{n,[(n+1)t]}^{(1)}, \quad v = H_{n,[(n+1)t]}^{(1)},$$

gives

$$\left| I_{n,M+H}^{(1)} - I_{n,M}^{(1)} \right| \leq I_{n,H}^{(1)} = O_p(n^{-1/2} \log n).$$

If $p > 1$, then by (148) and Hölder's inequality,

$$\begin{aligned} \left| I_{n,M+H}^{(1)} - I_{n,M}^{(1)} \right| &\leq K_p \left(I_{n,M}^{(1)} \right)^{(p-1)/p} \left(I_{n,H}^{(1)} \right)^{1/p} + K_p I_{n,H}^{(1)} \\ &= O_p\left(\frac{\log n}{\sqrt{n}}\right) + O_p(n^{-p/2} \log n). \end{aligned} \quad (158)$$

Thus, for all $1 \leq p < \infty$,

$$\frac{I_{n,M+H}^{(1)} - I_{n,M}^{(1)}}{\sqrt{\log n}} = o_p(1). \quad (159)$$

Combining (143) with (159), and using Lemma 3.4(ii), gives

$$\frac{1}{\sqrt{2C_p \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{v}_i \right|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt - 2m_p \log n \right] \xrightarrow{d} N(0, 1)$$

when $a = 1$. This completes the proof. $\square$

3.3.5 Proof of Theorem 3.5

Proof. We prove the two parts separately. In both parts, we first apply the corresponding joint limit theorem to the leading i.i.d. process, then show explicitly that the residual-based statistics are asymptotically equivalent to the corresponding leading statistics. The final transfer to the residual-based joint limit follows from Slutsky's theorem for random vectors.

PART (i).

By Lemma 3.2, $\{U_i^{(a)}\}_{i \geq 1}$ are i.i.d. with

$$\mathbb{E}U_1^{(a)} = 0, \quad 0 < \sigma_{v,a}^2 < \infty.$$

Moreover, the log-log moment verification in the proof of Theorem 3.2 shows that

$$\mathbb{E} \left[\{U_1^{(a)}\}^2 \log \log(|U_1^{(a)}| + 1) \right] < \infty;$$

see (85), (86), and (88).

Hence Theorem 2.5(i) applies to the i.i.d. sequence $\{U_i^{(a)}\}_{i \geq 1}$. Recalling the definition of

$M_{n,k}^{(a)}$ in (46), define the leading statistics

$$S_{n,q}^{(a)} := \frac{1}{\hat{\sigma}_{n,U^{(a)}}} \max_{1 \leq k \leq n} \frac{|M_{n,k}^{(a)}|}{\sqrt{n} q(k/n)},$$

$$D_n^{(a)} := A(n) \frac{1}{\hat{\sigma}_{n,U^{(a)}}} \max_{1 \leq k \leq n} \frac{|M_{n,k}^{(a)}|}{\sqrt{n} q^*(k/n)} - D(n).$$

Then Theorem 2.5(i) gives, for all $x, y \in \mathbb{R}$,

$$\mathbb{P} \left\{ S_{n,q}^{(a)} \leq x, D_n^{(a)} \leq y \right\} \longrightarrow \mathbb{P} \left\{ \sup_{0 < t < 1} \frac{|B(t)|}{q(t)} \leq x \right\} \exp(-2e^{-y}). \quad (160)$$

We now transfer (160) to the residual-based statistics. Define

$$\hat{S}_{n,q} := \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k \leq n} \frac{\left| \sum_{i=1}^k \hat{v}_i \right|}{\sqrt{n} q(k/n)},$$

$$\hat{D}_n := A(n) \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k \leq n} \frac{\left| \sum_{i=1}^k \hat{v}_i \right|}{\sqrt{n} q^*(k/n)} - D(n).$$

The Brownian-bridge-based supremum statistic.

First consider $a = 0$. By Lemma 3.4(i),

$$\sum_{i=1}^k \hat{v}_i = M_{n,k}^{(0)} + R_{n,k}^{(0)}.$$

Therefore,

$$\left| \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k \leq n} \frac{\left| \sum_{i=1}^k \hat{v}_i \right|}{\sqrt{n} q(k/n)} - \frac{1}{\hat{\sigma}_{n,U^{(0)}}} \max_{1 \leq k \leq n} \frac{|M_{n,k}^{(0)}|}{\sqrt{n} q(k/n)} \right|$$

$$\leq \left| \frac{\hat{\sigma}_{n,U^{(0)}}}{\hat{\sigma}_{n,\hat{v}}} - 1 \right| \frac{1}{\hat{\sigma}_{n,U^{(0)}}} \max_{1 \leq k \leq n} \frac{|M_{n,k}^{(0)}|}{\sqrt{n} q(k/n)} + \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k \leq n} \frac{|R_{n,k}^{(0)}|}{\sqrt{n} q(k/n)}.$$

The first term is $o_p(1)$ by (73) and (74). The second term is $o_p(1)$ by (75) and $\hat{\sigma}_{n,\hat{v}}^{-1} = O_p(1)$.

Hence

$$\hat{S}_{n,q} = S_{n,q}^{(0)} + o_p(1). \quad (161)$$

Now consider $a = 1$. By Lemma 3.4(ii),

$$\sum_{i=1}^k \hat{v}_i = M_{n,k}^{(1)} - E_{n,k} - F_{n,k} + R_{n,k}^{(1)}.$$

Thus,

$$\begin{aligned} & \left| \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \hat{v}_i \right|}{\sqrt{n} q(k/n)} - \frac{1}{\hat{\sigma}_{n,U^{(1)}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(1)}|}{\sqrt{n} q(k/n)} \right| \\ & \leq \left| \frac{\hat{\sigma}_{n,U^{(1)}}}{\hat{\sigma}_{n,\hat{v}}} - 1 \right| \frac{1}{\hat{\sigma}_{n,U^{(1)}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(1)}|}{\sqrt{n} q(k/n)} \\ & \quad + \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k < n} \frac{|E_{n,k}|}{\sqrt{n} q(k/n)} + \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k < n} \frac{|F_{n,k}|}{\sqrt{n} q(k/n)} \\ & \quad + \frac{1}{\hat{\sigma}_{n,\hat{v}}} \max_{1 \leq k < n} \frac{|R_{n,k}^{(1)}|}{\sqrt{n} q(k/n)}. \end{aligned}$$

The first term is $o_p(1)$ by (73) and (74). The remaining three terms are $o_p(1)$ by (79), (80), and (81), together with $\hat{\sigma}_{n,\hat{v}}^{-1} = O_p(1)$. Hence

$$\hat{S}_{n,q} = S_{n,q}^{(1)} + o_p(1). \quad (162)$$

Combining (161) and (162), we have, for both $a = 0$ and $a = 1$,

$$\hat{S}_{n,q} = S_{n,q}^{(a)} + o_p(1). \quad (163)$$

The supremum Darling–Erdős-type statistic.

First consider $a = 0$. By Lemma 3.4(i),

$$\sum_{i=1}^k \hat{v}_i = M_{n,k}^{(0)} + R_{n,k}^{(0)}.$$

Using the triangle inequality,

$$\begin{aligned} & \left| \widehat{D}_n - D_n^{(0)} \right| \\ & \leq A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k < n} \frac{|R_{n,k}^{(0)}|}{\sqrt{n} q^*(k/n)} \\ & \quad + \left| A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(0)}|}{\sqrt{n} q^*(k/n)} - A(n) \frac{1}{\widehat{\sigma}_{n,U^{(0)}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(0)}|}{\sqrt{n} q^*(k/n)} \right|. \end{aligned}$$

The first term is $o_p(1)$ by (93). The second term can be written as

$$A(n) \frac{1}{\widehat{\sigma}_{n,U^{(0)}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(0)}|}{\sqrt{n} q^*(k/n)} \left| \frac{\widehat{\sigma}_{n,U^{(0)}}}{\widehat{\sigma}_{n,\widehat{v}}} - 1 \right|,$$

which is $o_p(1)$ by (90) and (95). Hence

$$\widehat{D}_n = D_n^{(0)} + o_p(1). \quad (164)$$

Now consider $a = 1$. By Lemma 3.4(ii),

$$\sum_{i=1}^k \widehat{v}_i = M_{n,k}^{(1)} - E_{n,k} - F_{n,k} + R_{n,k}^{(1)}.$$

Therefore,

$$\begin{aligned} & \left| \widehat{D}_n - D_n^{(1)} \right| \\ & \leq A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k < n} \frac{|E_{n,k}|}{\sqrt{n} q^*(k/n)} + A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k < n} \frac{|F_{n,k}|}{\sqrt{n} q^*(k/n)} \\ & \quad + A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k < n} \frac{|R_{n,k}^{(1)}|}{\sqrt{n} q^*(k/n)} \\ & \quad + \left| A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(1)}|}{\sqrt{n} q^*(k/n)} - A(n) \frac{1}{\widehat{\sigma}_{n,U^{(1)}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(1)}|}{\sqrt{n} q^*(k/n)} \right|. \end{aligned}$$

The first three terms are $o_p(1)$ by (98), (99), and (100). The last term is $o_p(1)$ by (90) and (104). Hence

$$\widehat{D}_n = D_n^{(1)} + o_p(1). \quad (165)$$

Combining (164) and (165), we have, for both $a = 0$ and $a = 1$,

$$\widehat{D}_n = D_n^{(a)} + o_p(1). \quad (166)$$

By (163) and (166),

$$(\widehat{S}_{n,q}, \widehat{D}_n) = (S_{n,q}^{(a)}, D_n^{(a)}) + o_p(1)$$

componentwise. Together with the leading joint convergence in (160), Slutsky's theorem for random vectors gives

$$(\widehat{S}_{n,q}, \widehat{D}_n) \xrightarrow{d} \left(\sup_{0 < t < 1} \frac{|B(t)|}{q(t)}, G \right),$$

with the product-form limiting distribution in (160). Hence, for all $x, y \in \mathbb{R}$,

$$\mathbb{P} \left\{ \widehat{S}_{n,q} \leq x, \widehat{D}_n \leq y \right\} \longrightarrow \mathbb{P} \left\{ \sup_{0 < t < 1} \frac{|B(t)|}{q(t)} \leq x \right\} \exp(-2e^{-y}),$$

which proves (36).

PART (ii).

Let

$$\Delta_n(t) := \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right).$$

By Assumption 3.2, there exists $r > 2$ such that

$$\mathbb{E}|x|^{2r} < \infty, \quad \mathbb{E}|\varepsilon|^r < \infty.$$

As in the proof of Theorem 3.4, this implies that, for each $a \in \{0, 1\}$,

$$\mathbb{E}|U_1^{(a)}|^\nu < \infty \quad \text{for some } \nu > 2.$$

Therefore, Theorem 2.5(ii) applies to the i.i.d. sequence $\{U_i^{(a)}\}_{i \geq 1}$, with the first integral power p and the integral Darling–Erdős power ℓ .

Define the leading integral statistics

$$L_{n,p}^{(a)} := \frac{1}{\widehat{\sigma}_{n,U^{(a)}}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(a)}|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt,$$

$$J_{n,\ell}^{(a)} := \frac{1}{\sqrt{2C_\ell \log n}} \left[\frac{1}{\widehat{\sigma}_{n,U^{(a)}}^\ell} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(a)}|^\ell}{n^{\ell/2} \Delta_n(t)^{1+\ell/2}} dt - 2m_\ell \log n \right].$$

Then [Theorem 2.5](#)(ii) gives, for all $x, y \in \mathbb{R}$,

$$\mathbb{P} \left\{ L_{n,p}^{(a)} \leq x, J_{n,\ell}^{(a)} \leq y \right\} \longrightarrow \mathbb{P} \left\{ \int_0^1 \frac{|B(t)|^p}{q(t)} dt \leq x \right\} \Phi(y). \quad (167)$$

We now transfer [\(167\)](#) to the residual-based statistics. Define

$$\widehat{L}_{n,p} := \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{v}_i \right|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt,$$

$$\widehat{J}_{n,\ell} := \frac{1}{\sqrt{2C_\ell \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^\ell} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{v}_i \right|^\ell}{n^{\ell/2} \Delta_n(t)^{1+\ell/2}} dt - 2m_\ell \log n \right].$$

The Brownian-bridge-based integral statistic.

First consider $a = 0$. By [Lemma 3.4](#)(i),

$$\sum_{i=1}^k \widehat{v}_i = M_{n,k}^{(0)} + R_{n,k}^{(0)}.$$

The completion of the case $a = 0$ in the proof of [Theorem 3.3](#) shows that inserting $R_{n,k}^{(0)}$ inside the power does not change the integral functional asymptotically; see [\(127\)](#). Hence,

$$\widehat{L}_{n,p} = \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(0)}|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt + o_p(1).$$

By (121) and (122),

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(0)}|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt = L_{n,p}^{(0)} + o_p(1).$$

Therefore,

$$\widehat{L}_{n,p} = L_{n,p}^{(0)} + o_p(1). \quad (168)$$

Now consider $a = 1$. Define

$$H_{n,k}^{(1)} := -E_{n,k} - F_{n,k} + R_{n,k}^{(1)}.$$

Then, by Lemma 3.4(ii),

$$\sum_{i=1}^k \widehat{v}_i = M_{n,k}^{(1)} + H_{n,k}^{(1)}.$$

The completion of the case $a = 1$ in the proof of Theorem 3.3 shows that inserting $H_{n,k}^{(1)}$ inside the power does not change the integral functional asymptotically; see (135). Hence,

$$\widehat{L}_{n,p} = \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(1)}|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt + o_p(1).$$

Again, by (121) and (122),

$$\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(1)}|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt = L_{n,p}^{(1)} + o_p(1).$$

Thus,

$$\widehat{L}_{n,p} = L_{n,p}^{(1)} + o_p(1). \quad (169)$$

Combining (168) and (169), we have, for both $a = 0$ and $a = 1$,

$$\widehat{L}_{n,p} = L_{n,p}^{(a)} + o_p(1). \quad (170)$$

The integral Darling–Erdős-type statistic.

First consider $a = 0$. By Lemma 3.4(i),

$$\sum_{i=1}^k \widehat{v}_i = M_{n,k}^{(0)} + R_{n,k}^{(0)}.$$

The completion of the case $a = 0$ in the proof of Theorem 3.4, with p there replaced by ℓ , gives

$$\frac{1}{\sqrt{\log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^\ell} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(0)} + R_{n,\lfloor (n+1)t \rfloor}^{(0)}|^\ell}{n^{\ell/2} \Delta_n(t)^{1+\ell/2}} dt - \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^\ell} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(0)}|^\ell}{n^{\ell/2} \Delta_n(t)^{1+\ell/2}} dt \right] = o_p(1);$$

see (150). Hence,

$$\widehat{J}_{n,\ell} = \frac{1}{\sqrt{2C_\ell \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^\ell} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(0)}|^\ell}{n^{\ell/2} \Delta_n(t)^{1+\ell/2}} dt - 2m_\ell \log n \right] + o_p(1).$$

By (142), with the power p there replaced by ℓ ,

$$\frac{1}{\sqrt{2C_\ell \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^\ell} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(0)}|^\ell}{n^{\ell/2} \Delta_n(t)^{1+\ell/2}} dt - 2m_\ell \log n \right] = J_{n,\ell}^{(0)} + o_p(1).$$

Therefore,

$$\widehat{J}_{n,\ell} = J_{n,\ell}^{(0)} + o_p(1). \quad (171)$$

Now consider $a = 1$. With

$$H_{n,k}^{(1)} = -E_{n,k} - F_{n,k} + R_{n,k}^{(1)},$$

Lemma 3.4(ii) gives

$$\sum_{i=1}^k \widehat{v}_i = M_{n,k}^{(1)} + H_{n,k}^{(1)}.$$

The completion of the case $a = 1$ in the proof of Theorem 3.4, with p there replaced by ℓ ,

gives

$$\frac{1}{\sqrt{\log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^\ell} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(1)} + H_{n,\lfloor(n+1)t\rfloor}^{(1)}|^\ell}{n^{\ell/2} \Delta_n(t)^{1+\ell/2}} dt - \frac{1}{\widehat{\sigma}_{n,\widehat{v}}^\ell} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(1)}|^\ell}{n^{\ell/2} \Delta_n(t)^{1+\ell/2}} dt \right] = o_p(1);$$

see (159). Hence,

$$\widehat{J}_{n,\ell} = \frac{1}{\sqrt{2C_\ell \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^\ell} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(1)}|^\ell}{n^{\ell/2} \Delta_n(t)^{1+\ell/2}} dt - 2m_\ell \log n \right] + o_p(1).$$

By (142), with the power p there replaced by ℓ ,

$$\frac{1}{\sqrt{2C_\ell \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{v}}^\ell} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(1)}|^\ell}{n^{\ell/2} \Delta_n(t)^{1+\ell/2}} dt - 2m_\ell \log n \right] = J_{n,\ell}^{(1)} + o_p(1).$$

Thus,

$$\widehat{J}_{n,\ell} = J_{n,\ell}^{(1)} + o_p(1). \quad (172)$$

Combining (171) and (172), we have, for both $a = 0$ and $a = 1$,

$$\widehat{J}_{n,\ell} = J_{n,\ell}^{(a)} + o_p(1). \quad (173)$$

By (170) and (173),

$$(\widehat{L}_{n,p}, \widehat{J}_{n,\ell}) = (L_{n,p}^{(a)}, J_{n,\ell}^{(a)}) + o_p(1)$$

componentwise. Together with the leading joint convergence in (167), Slutsky's theorem for random vectors gives

$$(\widehat{L}_{n,p}, \widehat{J}_{n,\ell}) \xrightarrow{d} \left(\int_0^1 \frac{|B(t)|^p}{q(t)} dt, N \right),$$

with the product-form limiting distribution in (167). Consequently, for all $x, y \in \mathbb{R}$,

$$\mathbb{P} \left\{ \widehat{L}_{n,p} \leq x, \widehat{J}_{n,\ell} \leq y \right\} \longrightarrow \mathbb{P} \left\{ \int_0^1 \frac{|B(t)|^p}{q(t)} dt \leq x \right\} \Phi(y),$$

which proves (38). The proof is complete. $\square$

3.3.6 Proof of Theorem 3.6

Proof. Assume that Assumption 3.2 holds and that H_0^β holds. Let

$$\Delta_n(t) := \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right).$$

We prove the two convergences simultaneously by working with $j \in \{1, 2\}$, where

$$J_n^{(1)} = \left[\frac{d_n}{e_n}, 1 - d_n \right], \quad J_n^{(2)} = \left[d_n, 1 - \frac{d_n}{e_n} \right].$$

We first apply Theorem 2.6 to the leading i.i.d. sequence $\{U_i^{(a)}\}_{i \geq 1}$. By Lemma 3.2,

$$\mathbb{E}U_1^{(a)} = 0, \quad 0 < \sigma_{v,a}^2 < \infty,$$

where $\sigma_{v,a}^2$ is defined in (48). It remains only to verify the moment condition required in Theorem 2.6.

If $a = 0$, then $U_i^{(0)} = x_i \varepsilon_i$. By independence and Assumption 3.2,

$$\mathbb{E}|U_1^{(0)}|^r = \mathbb{E}|x\varepsilon|^r = \mathbb{E}|x|^r \mathbb{E}|\varepsilon|^r < \infty.$$

If $a = 1$, then $U_i^{(1)} = (x_i - \mathbb{E}x)\varepsilon_i$. Since $\mathbb{E}|x|^{2r} < \infty$, we also have $\mathbb{E}|x - \mathbb{E}x|^r < \infty$, and hence

$$\mathbb{E}|U_1^{(1)}|^r = \mathbb{E}|x - \mathbb{E}x|^r \mathbb{E}|\varepsilon|^r < \infty.$$

Thus, in both cases, $\{U_i^{(a)}\}_{i \geq 1}$ satisfies the assumptions of Theorem 2.6. Recalling the definition of $M_{n,k}^{(a)}$ in (46), Theorem 2.6 gives, for $j = 1, 2$,

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(a)}|}{\widehat{\sigma}_{n, U^{(a)}} \sqrt{n} \Delta_n(t)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|. \quad (174)$$

In particular,

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(a)}|}{\widehat{\sigma}_{n, U^{(a)}} \sqrt{n} \Delta_n(t)} = O_p(1). \quad (175)$$

By [Lemma 3.2](#) and [Lemma 3.5](#),

$$\widehat{\sigma}_{n, U^{(a)}}^2 \xrightarrow{p} \sigma_{v, a}^2, \quad \widehat{\sigma}_{n, \widehat{v}}^2 \xrightarrow{p} \sigma_{v, a}^2.$$

Therefore,

$$\frac{\widehat{\sigma}_{n, U^{(a)}}}{\widehat{\sigma}_{n, \widehat{v}}} \xrightarrow{p} 1. \quad (176)$$

Combining [\(174\)](#) with [\(176\)](#), we obtain, for $j = 1, 2$,

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(a)}|}{\widehat{\sigma}_{n, \widehat{v}} \sqrt{n} \Delta_n(t)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|. \quad (177)$$

CASE $a = 0$.

By [Lemma 3.4\(i\)](#),

$$\sum_{i=1}^k \widehat{v}_i = M_{n, k}^{(0)} + R_{n, k}^{(0)}, \quad 1 \leq k < n.$$

The remainder term $R_{n, k}^{(0)}$. We show that, for $j = 1, 2$,

$$d_n^{1/2} \frac{1}{\widehat{\sigma}_{n, \widehat{v}}} \sup_{t \in J_n^{(j)}} \frac{|R_{n, \lfloor (n+1)t \rfloor}^{(0)}|}{\sqrt{n} \Delta_n(t)} = o_p(1). \quad (178)$$

By [\(52\)](#), write

$$R_{n, k}^{(0)} = A_{n, 0} V_{n, k}^{(0)},$$

where

$$A_{n, 0} := \frac{\sum_{i=1}^n x_i \varepsilon_i}{\sum_{i=1}^n x_i^2}, \quad V_{n, k}^{(0)} := \frac{k}{n} \sum_{i=1}^n x_i^2 - \sum_{i=1}^k x_i^2.$$

First,

$$A_{n, 0} = \frac{\frac{1}{n} \sum_{i=1}^n x_i \varepsilon_i}{\frac{1}{n} \sum_{i=1}^n x_i^2} = o_p(1), \quad (179)$$

because

$$\frac{1}{n} \sum_{i=1}^n x_i \varepsilon_i \xrightarrow{p} \mathbb{E}(x\varepsilon) = \mathbb{E}x \mathbb{E}\varepsilon = 0, \quad \frac{1}{n} \sum_{i=1}^n x_i^2 \xrightarrow{p} \mathbb{E}x^2 > 0.$$

Next, by [Theorem 2.6](#) applied to $\{x_i^2\}_{i \geq 1}$,

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|V_{n, \lfloor (n+1)t \rfloor}^{(0)}|}{\sqrt{n} \Delta_n(t)} = O_p(1), \quad (180)$$

provided $\text{Var}(x^2) > 0$. If $\text{Var}(x^2) = 0$, then x_i^2 is constant a.s., and the supremum in (180) is identically zero. Since $\hat{\sigma}_{n,\hat{v}}^{-1} = O_p(1)$, combining (179) and (180) proves (178).

Completion of the case $a = 0$. Using the decomposition above and the triangle inequality, for $j = 1, 2$,

$$\begin{aligned} & \left| d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{v}_i \right|}{\hat{\sigma}_{n,\hat{v}} \sqrt{n} \Delta_n(t)} - d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(0)}|}{\hat{\sigma}_{n,\hat{v}} \sqrt{n} \Delta_n(t)} \right| \\ & \leq d_n^{1/2} \frac{1}{\hat{\sigma}_{n,\hat{v}}} \sup_{t \in J_n^{(j)}} \frac{|R_{n, \lfloor (n+1)t \rfloor}^{(0)}|}{\sqrt{n} \Delta_n(t)} = o_p(1), \end{aligned}$$

by (178). Therefore, by (177),

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{v}_i \right|}{\hat{\sigma}_{n,\hat{v}} \sqrt{n} \Delta_n(t)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|$$

when $a = 0$, for $j = 1, 2$.

CASE $a = 1$.

By [Lemma 3.4\(ii\)](#),

$$\sum_{i=1}^k \hat{v}_i = M_{n,k}^{(1)} - E_{n,k} - F_{n,k} + R_{n,k}^{(1)}, \quad 1 \leq k < n.$$

We show that, for $j = 1, 2$,

$$d_n^{1/2} \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \sup_{t \in J_n^{(j)}} \frac{|E_{n,\lfloor(n+1)t\rfloor}|}{\sqrt{n} \Delta_n(t)} = o_p(1), \quad (181)$$

$$d_n^{1/2} \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \sup_{t \in J_n^{(j)}} \frac{|F_{n,\lfloor(n+1)t\rfloor}|}{\sqrt{n} \Delta_n(t)} = o_p(1), \quad (182)$$

$$d_n^{1/2} \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \sup_{t \in J_n^{(j)}} \frac{|R_{n,\lfloor(n+1)t\rfloor}^{(1)}|}{\sqrt{n} \Delta_n(t)} = o_p(1). \quad (183)$$

The term $E_{n,k}$. By (54),

$$E_{n,k} = \bar{\varepsilon}_n \left[\sum_{i=1}^k (x_i - \mathbb{E}x) - \frac{k}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right].$$

Therefore,

$$\begin{aligned} & d_n^{1/2} \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \sup_{t \in J_n^{(j)}} \frac{|E_{n,\lfloor(n+1)t\rfloor}|}{\sqrt{n} \Delta_n(t)} \\ &= |\bar{\varepsilon}_n| \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{\left| \sum_{i=1}^{\lfloor(n+1)t\rfloor} (x_i - \mathbb{E}x) - \frac{\lfloor(n+1)t\rfloor}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right|}{\sqrt{n} \Delta_n(t)}. \end{aligned}$$

Here $\bar{\varepsilon}_n = o_p(1)$ and $\widehat{\sigma}_{n,\widehat{v}}^{-1} = O_p(1)$. The supremum factor is $O_p(1)$ by Theorem 2.6 applied to $\{x_i - \mathbb{E}x\}_{i \geq 1}$. Hence (181) holds.

The term $F_{n,k}$. By (55),

$$F_{n,k} = (\bar{x}_n - \mathbb{E}x) \left[\sum_{i=1}^k \varepsilon_i - \frac{k}{n} \sum_{i=1}^n \varepsilon_i \right].$$

Therefore,

$$\begin{aligned} & d_n^{1/2} \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} \sup_{t \in J_n^{(j)}} \frac{|F_{n,\lfloor(n+1)t\rfloor}|}{\sqrt{n} \Delta_n(t)} \\ &= |\bar{x}_n - \mathbb{E}x| \frac{1}{\widehat{\sigma}_{n,\widehat{v}}} d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{\left| \sum_{i=1}^{\lfloor(n+1)t\rfloor} \varepsilon_i - \frac{\lfloor(n+1)t\rfloor}{n} \sum_{i=1}^n \varepsilon_i \right|}{\sqrt{n} \Delta_n(t)}. \end{aligned}$$

Here $\bar{x}_n - \mathbb{E}x = o_p(1)$ and $\hat{\sigma}_{n,\hat{v}}^{-1} = O_p(1)$. The supremum factor is $O_p(1)$ by [Theorem 2.6](#) applied to $\{\varepsilon_i\}_{i \geq 1}$. Hence [\(182\)](#) holds.

The term $R_{n,k}^{(1)}$. By [\(56\)](#), write

$$R_{n,k}^{(1)} = A_{n,1} W_{n,k}^{(1)},$$

where

$$A_{n,1} := \frac{\sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n)}{\sum_{i=1}^n (x_i - \bar{x}_n)^2},$$

and

$$W_{n,k}^{(1)} := \frac{k}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 - \sum_{i=1}^k (x_i - \bar{x}_n)^2.$$

First,

$$A_{n,1} = o_p(1). \tag{184}$$

Indeed,

$$A_{n,1} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)(\varepsilon_i - \bar{\varepsilon}_n)}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2},$$

where the numerator converges in probability to

$$\mathbb{E}[(x - \mathbb{E}x)\varepsilon] = \mathbb{E}(x - \mathbb{E}x) \mathbb{E}\varepsilon = 0,$$

and the denominator converges in probability to $\text{Var}(x) > 0$.

Next, we show that

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|W_{n, \lfloor (n+1)t \rfloor}^{(1)}|}{\sqrt{n} \Delta_n(t)} = O_p(1). \tag{185}$$

Using

$$(x_i - \bar{x}_n)^2 = (x_i - \mathbb{E}x)^2 - 2(\bar{x}_n - \mathbb{E}x)(x_i - \mathbb{E}x) + (\bar{x}_n - \mathbb{E}x)^2,$$

the last term cancels in the tied-down partial-sum expression. Therefore,

$$\begin{aligned}
 & d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|W_{n, \lfloor (n+1)t \rfloor}^{(1)}|}{\sqrt{n} \Delta_n(t)} \\
 & \leq d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} (x_i - \mathbb{E}x)^2 - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n (x_i - \mathbb{E}x)^2 \right|}{\sqrt{n} \Delta_n(t)} \\
 & \quad + 2|\bar{x}_n - \mathbb{E}x| d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} (x_i - \mathbb{E}x) - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n (x_i - \mathbb{E}x) \right|}{\sqrt{n} \Delta_n(t)}.
 \end{aligned}$$

The first supremum is $O_p(1)$ by [Theorem 2.6](#) applied to $\{(x_i - \mathbb{E}x)^2\}_{i \geq 1}$, provided $\text{Var}((x - \mathbb{E}x)^2) > 0$. If $\text{Var}((x - \mathbb{E}x)^2) = 0$, then this first supremum is identically zero. The second supremum is $O_p(1)$ by [Theorem 2.6](#) applied to $\{x_i - \mathbb{E}x\}_{i \geq 1}$, and $\bar{x}_n - \mathbb{E}x = o_p(1)$. This proves [\(185\)](#).

Combining [\(184\)](#) and [\(185\)](#), and using $\hat{\sigma}_{n,\hat{v}}^{-1} = O_p(1)$, gives [\(183\)](#).

Completion of the case $a = 1$. Using the decomposition above and the triangle inequality, for $j = 1, 2$,

$$\begin{aligned}
 & \left| d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{v}_i \right|}{\hat{\sigma}_{n,\hat{v}} \sqrt{n} \Delta_n(t)} - d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(1)}|}{\hat{\sigma}_{n,\hat{v}} \sqrt{n} \Delta_n(t)} \right| \\
 & \leq d_n^{1/2} \frac{1}{\hat{\sigma}_{n,\hat{v}}} \sup_{t \in J_n^{(j)}} \frac{|E_{n, \lfloor (n+1)t \rfloor}|}{\sqrt{n} \Delta_n(t)} + d_n^{1/2} \frac{1}{\hat{\sigma}_{n,\hat{v}}} \sup_{t \in J_n^{(j)}} \frac{|F_{n, \lfloor (n+1)t \rfloor}|}{\sqrt{n} \Delta_n(t)} \\
 & \quad + d_n^{1/2} \frac{1}{\hat{\sigma}_{n,\hat{v}}} \sup_{t \in J_n^{(j)}} \frac{|R_{n, \lfloor (n+1)t \rfloor}^{(1)}|}{\sqrt{n} \Delta_n(t)} = o_p(1),
 \end{aligned}$$

by [\(181\)](#), [\(182\)](#), and [\(183\)](#). Therefore, by [\(177\)](#),

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{v}_i \right|}{\hat{\sigma}_{n,\hat{v}} \sqrt{n} \Delta_n(t)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|$$

when $a = 1$, for $j = 1, 2$. Since $j = 1$ and $j = 2$ correspond respectively to $J_n^{(1)}$ and $J_n^{(2)}$, [\(40\)](#) and [\(41\)](#) follow. This completes the proof. $\square$

4 Testing for a Change in the Intercept

4.1 Model, residual quantities, and assumptions

We next consider the problem of testing for a possible change in the intercept. Under the null hypothesis, the observations satisfy

$$y_i = \alpha + \beta x_i + \varepsilon_i, \quad 1 \leq i \leq n. \quad (186)$$

A possible intercept change is described by

$$y_i = \begin{cases} \alpha_1 + \beta x_i + \varepsilon_i, & 1 \leq i \leq k^*, \\ \alpha_2 + \beta x_i + \varepsilon_i, & k^* < i \leq n, \end{cases} \quad 1 \leq k^* < n, \quad (187)$$

where k^* is unknown. We test

$$H_0^\alpha : \alpha_1 = \alpha_2 \quad \text{against} \quad H_1^\alpha : \alpha_1 \neq \alpha_2. \quad (188)$$

For the intercept-change problem, define

$$\hat{w}_i = y_i - \hat{\beta} x_i, \quad 1 \leq i \leq n, \quad (189)$$

where $\hat{\beta}$ is defined in (24). Let

$$\hat{\sigma}_{n,\hat{w}}^2 = \frac{1}{n} \sum_{i=1}^n \left(\hat{w}_i - \frac{1}{n} \sum_{j=1}^n \hat{w}_j \right)^2. \quad (190)$$

Under H_0^α , the tied-down partial sums of $\hat{w}_i$ decompose into the tied-down partial sums of the regression errors plus a slope-estimation remainder term. This decomposition is made precise in the auxiliary results, and the remainder term is shown in the proofs to be asymptotically negligible under the normalizations and functionals considered below. Thus, the intercept-change statistics are regression analogues of the mean-change statistics in [Section 2](#), with the residual-based quantities $\hat{w}_i$ playing the role of the observations.

The results below use the same random-design assumptions introduced in [Section 3.1](#), with additional moment conditions stated when needed.

4.2 Main results

4.2.1 Supremum statistic with Brownian-bridge-based limit

Theorem 4.1. *Assume that [Assumption 3.1](#) holds and that H_0^α holds. Let $q(t) \in Q$, and suppose that $I_0^1(q, c) < \infty$ for some or possibly all $c > 0$. Then, as $n \rightarrow \infty$,*

$$\frac{1}{\hat{\sigma}_{n,\hat{w}}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \hat{w}_i - \frac{k}{n} \sum_{i=1}^n \hat{w}_i \right|}{\sqrt{n} q(k/n)} \xrightarrow{d} \sup_{0 < t < 1} \frac{|B(t)|}{q(t)}, \quad (191)$$

where $\{B(t), 0 \leq t \leq 1\}$ is a standard Brownian bridge and $\hat{\sigma}_{n,\hat{w}}^2$ is defined in [\(190\)](#).

Remark 11. *For each k , $1 \leq k < n$, let*

$$\bar{x}_k := \frac{1}{k} \sum_{i=1}^k x_i, \quad \bar{y}_k := \frac{1}{k} \sum_{i=1}^k y_i,$$

and

$$\tilde{x}_k := \frac{1}{n-k} \sum_{i=k+1}^n x_i, \quad \tilde{y}_k := \frac{1}{n-k} \sum_{i=k+1}^n y_i.$$

Define

$$\hat{\alpha}_1^k := \bar{y}_k - \hat{\beta} \bar{x}_k, \quad \hat{\alpha}_2^k := \tilde{y}_k - \hat{\beta} \tilde{x}_k.$$

Then

$$\sum_{i=1}^k \hat{w}_i - \frac{k}{n} \sum_{i=1}^n \hat{w}_i = \frac{k(n-k)}{n} (\hat{\alpha}_1^k - \hat{\alpha}_2^k). \quad (192)$$

Thus, the statistic

$$\frac{1}{\hat{\sigma}_{n,\hat{w}}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \hat{w}_i - \frac{k}{n} \sum_{i=1}^n \hat{w}_i \right|}{\sqrt{n} q(k/n)} \quad (193)$$

can be interpreted as the maximum, over all possible change points k , of appropriately normalized and weighted differences between the estimated intercepts computed from the observations before and after k , using the common full-sample slope estimate $\hat{\beta}$. Consequently, large observed values of [\(193\)](#) provide evidence against H_0^α in favor of H_1^α .

Remark 12. *The statistic in [Theorem 4.1](#), which has a weighted Brownian-bridge supremum limit, is connected to residual-type methods for detecting intercept changes. In particular, the intercept-change supremum statistic of Martsynyuk ([2021](#)) for structural errors-in-variables models has a similar tied-down partial-sum structure; when the measurement-error component in the covariate is removed, that construction reduces to a residual-based statistic of the same general form.*

The statistic is also conceptually related to the intercept-change methods of Pouliot ([2016](#)), which were developed to distinguish changes in intercept from changes in slope. The connection is mainly through the use of weighted tied-down partial sums of residual-type quantities, but the statistic considered here is constructed directly from the full-sample residual quantity $\widehat{w}_i$ introduced above.

The Brownian-bridge-based integral statistic, the supremum and integral Darling–Erdős-type statistics, and the Rényi-type statistics developed below extend the present intercept-change framework beyond the Brownian-bridge-based supremum case.

4.2.2 Supremum Darling–Erdős-type statistic

Theorem 4.2. *Assume that [Assumption 3.1](#) holds, that H_0^α holds, and that*

$$\mathbb{E}[\varepsilon^2 \log \log(|\varepsilon| + 1)] < \infty.$$

Let

$$q^*(t) = \sqrt{t(1-t)}, \quad 0 < t < 1.$$

Then, as $n \rightarrow \infty$,

$$A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{w}}} \max_{1 \leq k \leq n} \frac{\left| \sum_{i=1}^k \widehat{w}_i - \frac{k}{n} \sum_{i=1}^n \widehat{w}_i \right|}{\sqrt{n} q^*(k/n)} - D(n) \xrightarrow{d} G, \quad (194)$$

where $A(n)$, $D(n)$, and G are as defined in [Theorem 2.3](#), and $\widehat{\sigma}_{n,\widehat{w}}^2$ is defined in [\(190\)](#).

4.2.3 Integral statistic with Brownian-bridge-based limit

Theorem 4.3. Assume that [Assumption 3.1](#) holds and that H_0^α holds. Let $0 < p < \infty$ and $q(t) \in Q$. Suppose that

$$\int_0^1 \frac{(t(1-t))^{p/2}}{q(t)} dt < \infty. \quad (195)$$

Then, as $n \rightarrow \infty$,

$$\frac{1}{\hat{\sigma}_{n,\hat{w}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \hat{w}_i \right|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt \xrightarrow{d} \int_0^1 \frac{|B(t)|^p}{q(t)} dt, \quad (196)$$

where $\{B(t), 0 \leq t \leq 1\}$ is a standard Brownian bridge and $\hat{\sigma}_{n,\hat{w}}^2$ is defined in [\(190\)](#).

4.2.4 Integral Darling–Erdős-type statistic

Theorem 4.4. Assume that [Assumption 3.2](#) holds and that H_0^α holds. Let $1 \leq p < \infty$.

Then, as $n \rightarrow \infty$,

$$\frac{1}{\sqrt{2C_p \log n}} \left[\frac{1}{\hat{\sigma}_{n,\hat{w}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \hat{w}_i \right|^p}{n^{p/2} \left(\frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right) \right)^{1+p/2}} dt - 2m_p \log n \right] \xrightarrow{d} N(0, 1), \quad (197)$$

where

$$m_p = \int_{-\infty}^{\infty} |x|^p \phi(x) dx,$$

ϕ is the standard normal density, C_p is defined as in [Theorem 2.4](#), and $\hat{\sigma}_{n,\hat{w}}^2$ is defined in [\(190\)](#).

The marginal limits above can also be combined into joint limits. The following result is the intercept-change analogue of [Theorem 2.5](#). One joint limit combines the supremum-based statistics in [Theorems 4.1](#) and [4.2](#), and the other combines the integral-based statistics in [Theorems 4.3](#) and [4.4](#).

4.2.5 Joint limit theorem

Theorem 4.5. Let $\hat{w}_i$ be defined by [\(189\)](#), where $\hat{\beta}$ is defined in [\(24\)](#).

- (i) Assume that [Assumption 3.1](#) holds and that H_0^α holds. Let $q(t) \in Q$, and suppose that $I_0^1(q, c) < \infty$ for some or possibly all $c > 0$. Assume also that

$$\mathbb{E}[\varepsilon^2 \log \log(|\varepsilon| + 1)] < \infty.$$

Then, for all $x, y \in \mathbb{R}$, as $n \rightarrow \infty$,

$$\begin{aligned} & \mathbb{P} \left\{ \frac{1}{\widehat{\sigma}_{n,\widehat{w}}} \max_{1 \leq k \leq n} \frac{\left| \sum_{i=1}^k \widehat{w}_i - \frac{k}{n} \sum_{i=1}^n \widehat{w}_i \right|}{\sqrt{n} q(k/n)} \leq x, \right. \\ & \quad \left. A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{w}}} \max_{1 \leq k \leq n} \frac{\left| \sum_{i=1}^k \widehat{w}_i - \frac{k}{n} \sum_{i=1}^n \widehat{w}_i \right|}{\sqrt{n} q^*(k/n)} - D(n) \leq y \right\} \\ & \longrightarrow \mathbb{P} \left\{ \sup_{0 < t < 1} \frac{|B(t)|}{q(t)} \leq x \right\} \exp(-2e^{-y}), \end{aligned} \quad (198)$$

where $q^*(t) = \sqrt{t(1-t)}$, and $A(n)$, $D(n)$, and G are as defined in [Theorem 2.3](#).

- (ii) Assume that [Assumption 3.2](#) holds and that H_0^α holds. Let $0 < p < \infty$, let $1 \leq \ell < \infty$, and let $q(t) \in Q$. Suppose that

$$\int_0^1 \frac{(t(1-t))^{p/2}}{q(t)} dt < \infty. \quad (199)$$

Then, for all $x, y \in \mathbb{R}$, as $n \rightarrow \infty$,

$$\begin{aligned} & \mathbb{P} \left\{ \frac{1}{\widehat{\sigma}_{n,\widehat{w}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \widehat{w}_i \right|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt \leq x, \right. \\ & \quad \left. \frac{1}{\sqrt{2C_\ell \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{w}}^\ell} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \widehat{w}_i \right|^\ell}{n^{\ell/2} \left(\frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n}\right)\right)^{1+\ell/2}} dt - 2m_\ell \log n \right] \leq y \right\} \\ & \longrightarrow \mathbb{P} \left\{ \int_0^1 \frac{|B(t)|^p}{q(t)} dt \leq x \right\} \Phi(y), \end{aligned} \quad (200)$$

where Φ is the standard normal distribution function, and m_ℓ and C_ℓ are defined as in [Theorem 2.4](#).

4.2.6 Rényi-type weighted supremum statistic

Theorem 4.6. Assume that [Assumption 3.2](#) holds and that H_0^α holds. Let $\{d_n\}_{n \geq 1}$ and $\{e_n\}_{n \geq 1}$ be sequences of positive constants satisfying (17). Let

$$J_n^{(1)} := \left[\frac{d_n}{e_n}, 1 - d_n \right], \quad J_n^{(2)} := \left[d_n, 1 - \frac{d_n}{e_n} \right].$$

Then, as $n \rightarrow \infty$,

$$d_n^{1/2} \sup_{t \in J_n^{(1)}} \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \hat{w}_i \right|}{\hat{\sigma}_{n,\hat{w}} \sqrt{n} \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|, \quad (201)$$

and

$$d_n^{1/2} \sup_{t \in J_n^{(2)}} \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \hat{w}_i \right|}{\hat{\sigma}_{n,\hat{w}} \sqrt{n} \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|, \quad (202)$$

where $\{W(t), 0 \leq t \leq 1\}$ is a standard Wiener process and $\hat{\sigma}_{n,\hat{w}}^2$ is defined in (190).

4.3 Proofs of the Intercept-Change Results

Auxiliary results for intercept-change statistics

For $1 \leq k < n$, define

$$M_{n,k}^{(\alpha)} := \sum_{i=1}^k \varepsilon_i - \frac{k}{n} \sum_{i=1}^n \varepsilon_i, \quad (203)$$

and

$$X_{n,k} := \sum_{i=1}^k x_i - \frac{k}{n} \sum_{i=1}^n x_i. \quad (204)$$

Also define

$$\hat{\sigma}_{n,\varepsilon}^2 := \frac{1}{n} \sum_{i=1}^n \left(\varepsilon_i - \frac{1}{n} \sum_{j=1}^n \varepsilon_j \right)^2. \quad (205)$$

Lemma 4.1. Assume that [Assumption 3.1](#) holds and that H_0^α holds. Then, for $1 \leq k < n$,

$$\sum_{i=1}^k \hat{w}_i - \frac{k}{n} \sum_{i=1}^n \hat{w}_i = M_{n,k}^{(\alpha)} - (\hat{\beta} - \beta) X_{n,k}. \quad (206)$$

Proof. Under H_0^α ,

$$y_i = \alpha + \beta x_i + \varepsilon_i, \quad 1 \leq i \leq n.$$

Therefore,

$$\hat{w}_i = y_i - \hat{\beta}x_i = \alpha + \varepsilon_i - (\hat{\beta} - \beta)x_i.$$

Hence

$$\begin{aligned} \sum_{i=1}^k \hat{w}_i - \frac{k}{n} \sum_{i=1}^n \hat{w}_i &= \left[k\alpha + \sum_{i=1}^k \varepsilon_i - (\hat{\beta} - \beta) \sum_{i=1}^k x_i \right] \\ &\quad - \frac{k}{n} \left[n\alpha + \sum_{i=1}^n \varepsilon_i - (\hat{\beta} - \beta) \sum_{i=1}^n x_i \right]. \end{aligned}$$

The constant intercept terms cancel, and the result follows from (203) and (204). $\square$

Lemma 4.2. Assume that *Assumption 3.1* holds and that H_0^α holds. Then

$$\hat{\beta} - \beta = O_p(n^{-1/2}), \quad (207)$$

$$\hat{\sigma}_{n,\varepsilon}^2 \xrightarrow{p} \sigma_\varepsilon^2, \quad (208)$$

$$\hat{\sigma}_{n,\hat{w}}^2 \xrightarrow{p} \sigma_\varepsilon^2, \quad (209)$$

where $\sigma_\varepsilon^2 := \text{Var}(\varepsilon) > 0$. Moreover,

$$\hat{\sigma}_{n,\hat{w}}^2 - \hat{\sigma}_{n,\varepsilon}^2 = O_p(n^{-1/2}). \quad (210)$$

Consequently,

$$\frac{\hat{\sigma}_{n,\varepsilon}}{\hat{\sigma}_{n,\hat{w}}} \xrightarrow{p} 1. \quad (211)$$

Proof. Under H_0^α , the least-squares slope estimator satisfies

$$\hat{\beta} - \beta = \frac{\sum_{i=1}^n (x_i - \bar{x}_n) \varepsilon_i}{\sum_{i=1}^n (x_i - \bar{x}_n)^2}. \quad (212)$$

By *Assumption 3.1*,

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \xrightarrow{p} \text{Var}(x) > 0.$$

Also,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (x_i - \bar{x}_n) \varepsilon_i = \frac{1}{\sqrt{n}} \sum_{i=1}^n (x_i - \mathbb{E}x) \varepsilon_i - \sqrt{n}(\bar{x}_n - \mathbb{E}x) \bar{\varepsilon}_n = O_p(1),$$

because $\mathbb{E}[(x - \mathbb{E}x)\varepsilon] = 0$ and $\sqrt{n}(\bar{x}_n - \mathbb{E}x) \bar{\varepsilon}_n = O_p(n^{-1/2})$. Thus (207) follows from (212).

The consistency (208) follows from the law of large numbers. Next,

$$\hat{w}_i - \widehat{w}_n = (\varepsilon_i - \bar{\varepsilon}_n) - (\hat{\beta} - \beta)(x_i - \bar{x}_n).$$

Therefore,

$$\begin{aligned} \hat{\sigma}_{n,\widehat{w}}^2 &= \frac{1}{n} \sum_{i=1}^n \left[(\varepsilon_i - \bar{\varepsilon}_n) - (\hat{\beta} - \beta)(x_i - \bar{x}_n) \right]^2 \\ &= \hat{\sigma}_{n,\varepsilon}^2 - 2(\hat{\beta} - \beta) \frac{1}{n} \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon}_n)(x_i - \bar{x}_n) + (\hat{\beta} - \beta)^2 \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2. \end{aligned}$$

The sample covariance term is $O_p(1)$, while $\hat{\beta} - \beta = O_p(n^{-1/2})$ and $n^{-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2 = O_p(1)$. Hence

$$\hat{\sigma}_{n,\widehat{w}}^2 - \hat{\sigma}_{n,\varepsilon}^2 = O_p(n^{-1/2}),$$

which proves (210). Combining this with (208) gives (209). Finally, (211) follows from (208) and (209). $\square$

4.3.1 Proof of Theorem 4.1

Proof. Assume that Assumption 3.1 holds and that H_0^α holds. Let $q(t) \in \mathcal{Q}$, and suppose that $I_0^1(q, c) < \infty$ for some or possibly all $c > 0$.

By Theorem 2.1 applied to the i.i.d. sequence $\{\varepsilon_i\}_{i \geq 1}$, we have

$$\frac{1}{\hat{\sigma}_{n,\varepsilon}} \max_{1 \leq k \leq n} \frac{|M_{n,k}^{(\alpha)}|}{\sqrt{n} q(k/n)} \xrightarrow{d} \sup_{0 < t \leq 1} \frac{|B(t)|}{q(t)}. \quad (213)$$

Using (211), (213) also gives

$$\frac{1}{\hat{\sigma}_{n,\widehat{w}}} \max_{1 \leq k \leq n} \frac{|M_{n,k}^{(\alpha)}|}{\sqrt{n} q(k/n)} \xrightarrow{d} \sup_{0 < t \leq 1} \frac{|B(t)|}{q(t)}. \quad (214)$$

The slope-estimation remainder. By [Lemma 4.1](#),

$$\sum_{i=1}^k \hat{w}_i - \frac{k}{n} \sum_{i=1}^n \hat{w}_i = M_{n,k}^{(\alpha)} - (\hat{\beta} - \beta) X_{n,k}.$$

We show that

$$\frac{|\hat{\beta} - \beta|}{\hat{\sigma}_{n,\hat{w}}} \max_{1 \leq k < n} \frac{|X_{n,k}|}{\sqrt{n} q(k/n)} = o_p(1). \quad (215)$$

By [Theorem 2.1](#) applied to the i.i.d. sequence $\{x_i\}_{i \geq 1}$,

$$\frac{1}{\hat{\sigma}_{n,x}} \max_{1 \leq k < n} \frac{|X_{n,k}|}{\sqrt{n} q(k/n)} = O_p(1),$$

where

$$\hat{\sigma}_{n,x}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \xrightarrow{p} \text{Var}(x) > 0.$$

Thus

$$\max_{1 \leq k < n} \frac{|X_{n,k}|}{\sqrt{n} q(k/n)} = O_p(1).$$

Together with $\hat{\beta} - \beta = o_p(1)$ and $\hat{\sigma}_{n,\hat{w}}^{-1} = O_p(1)$, this proves (215).

Completion of the proof. By the triangle inequality and [Lemma 4.1](#),

$$\begin{aligned} & \left| \frac{1}{\hat{\sigma}_{n,\hat{w}}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \hat{w}_i - \frac{k}{n} \sum_{i=1}^n \hat{w}_i \right|}{\sqrt{n} q(k/n)} - \frac{1}{\hat{\sigma}_{n,\hat{w}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(\alpha)}|}{\sqrt{n} q(k/n)} \right| \\ & \leq \frac{|\hat{\beta} - \beta|}{\hat{\sigma}_{n,\hat{w}}} \max_{1 \leq k < n} \frac{|X_{n,k}|}{\sqrt{n} q(k/n)} = o_p(1), \end{aligned}$$

by (215). Combining this with (214) gives (191). □

4.3.2 Proof of [Theorem 4.2](#)

Proof. Assume that [Assumption 3.1](#) holds, that H_0^α holds, and that

$$\mathbb{E}[\varepsilon^2 \log \log(|\varepsilon| + 1)] < \infty.$$

Let $q^*(t) = \sqrt{t(1-t)}$. By [Theorem 2.3](#) applied to the i.i.d. sequence $\{\varepsilon_i\}_{i \geq 1}$,

$$A(n) \frac{1}{\widehat{\sigma}_{n,\varepsilon}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(\alpha)}|}{\sqrt{n} q^*(k/n)} - D(n) \xrightarrow{d} G. \quad (216)$$

Consequently,

$$A(n) \frac{1}{\widehat{\sigma}_{n,\varepsilon}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(\alpha)}|}{\sqrt{n} q^*(k/n)} = D(n) + O_p(1). \quad (217)$$

Replacing $\widehat{\sigma}_{n,\varepsilon}$ by $\widehat{\sigma}_{n,\widehat{w}}$. We show that

$$A(n) \left| \frac{1}{\widehat{\sigma}_{n,\widehat{w}}} - \frac{1}{\widehat{\sigma}_{n,\varepsilon}} \right| \max_{1 \leq k < n} \frac{|M_{n,k}^{(\alpha)}|}{\sqrt{n} q^*(k/n)} = o_p(1). \quad (218)$$

Indeed,

$$\left| \frac{\widehat{\sigma}_{n,\varepsilon}}{\widehat{\sigma}_{n,\widehat{w}}} - 1 \right| = O_p(n^{-1/2})$$

by [\(210\)](#), [\(208\)](#), and [\(209\)](#). Hence the left-hand side of [\(218\)](#) equals

$$\left| \frac{\widehat{\sigma}_{n,\varepsilon}}{\widehat{\sigma}_{n,\widehat{w}}} - 1 \right| A(n) \frac{1}{\widehat{\sigma}_{n,\varepsilon}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(\alpha)}|}{\sqrt{n} q^*(k/n)} = O_p(n^{-1/2}) \{D(n) + O_p(1)\} = o_p(1),$$

because $D(n) = O(\log \log n)$. Therefore,

$$A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{w}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(\alpha)}|}{\sqrt{n} q^*(k/n)} - D(n) \xrightarrow{d} G. \quad (219)$$

The slope-estimation remainder. By [Lemma 4.1](#), the residual partial sum differs from $M_{n,k}^{(\alpha)}$ by

$$-(\widehat{\beta} - \beta)X_{n,k}.$$

We show that

$$A(n) \frac{|\widehat{\beta} - \beta|}{\widehat{\sigma}_{n,\widehat{w}}} \max_{1 \leq k < n} \frac{|X_{n,k}|}{\sqrt{n} q^*(k/n)} = o_p(1). \quad (220)$$

Since [Assumption 3.1](#) implies

$$\mathbb{E}[x^2 \log \log(|x| + 1)] < \infty,$$

Theorem 2.3 applied to $\{x_i\}_{i \geq 1}$ gives

$$A(n) \frac{1}{\widehat{\sigma}_{n,x}} \max_{1 \leq k < n} \frac{|X_{n,k}|}{\sqrt{n} q^*(k/n)} - D(n) = O_p(1).$$

Thus

$$A(n) \max_{1 \leq k < n} \frac{|X_{n,k}|}{\sqrt{n} q^*(k/n)} = O_p(D(n)).$$

Together with $\widehat{\beta} - \beta = O_p(n^{-1/2})$, $\widehat{\sigma}_{n,\widehat{w}}^{-1} = O_p(1)$, and $D(n) = O(\log \log n)$, this proves (220).

Completion of the proof. Using the triangle inequality and **Lemma 4.1**,

$$\begin{aligned} & \left| A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{w}}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \widehat{w}_i - \frac{k}{n} \sum_{i=1}^n \widehat{w}_i \right|}{\sqrt{n} q^*(k/n)} - A(n) \frac{1}{\widehat{\sigma}_{n,\widehat{w}}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(\alpha)}|}{\sqrt{n} q^*(k/n)} \right| \\ & \leq A(n) \frac{|\widehat{\beta} - \beta|}{\widehat{\sigma}_{n,\widehat{w}}} \max_{1 \leq k < n} \frac{|X_{n,k}|}{\sqrt{n} q^*(k/n)} = o_p(1), \end{aligned}$$

by (220). Combining this with (219) gives (194). $\square$

4.3.3 Proof of **Theorem 4.3**

Proof. Assume that **Assumption 3.1** holds and that H_0^α holds. Let $0 < p < \infty$, let $q(t) \in \mathcal{Q}$, and suppose that

$$\int_0^1 \frac{(t(1-t))^{p/2}}{q(t)} dt < \infty.$$

By **Theorem 2.2** applied to the i.i.d. sequence $\{\varepsilon_i\}_{i \geq 1}$,

$$\frac{1}{\widehat{\sigma}_{n,\varepsilon}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(\alpha)}|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt \xrightarrow{d} \int_0^1 \frac{|B(t)|^p}{q(t)} dt. \quad (221)$$

In particular, the left-hand side of (221) is $O_p(1)$. By (211), this also gives

$$\frac{1}{\widehat{\sigma}_{n,\widehat{w}}^p} \int_0^1 \frac{|M_{n,\lfloor (n+1)t \rfloor}^{(\alpha)}|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt \xrightarrow{d} \int_0^1 \frac{|B(t)|^p}{q(t)} dt. \quad (222)$$

The slope-estimation remainder. By [Lemma 4.1](#), with $k = \lfloor (n+1)t \rfloor$,

$$\sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \hat{w}_i = M_{n, \lfloor (n+1)t \rfloor}^{(\alpha)} - (\hat{\beta} - \beta) X_{n, \lfloor (n+1)t \rfloor}.$$

Define

$$\begin{aligned} I_{n,M}^{(\alpha)} &:= \frac{1}{\hat{\sigma}_{n,\hat{w}}^p} \int_0^1 \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(\alpha)}|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt, \\ I_{n,R}^{(\alpha)} &:= \frac{|\hat{\beta} - \beta|^p}{\hat{\sigma}_{n,\hat{w}}^p} \int_0^1 \frac{|X_{n, \lfloor (n+1)t \rfloor}|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt, \\ I_{n,M+R}^{(\alpha)} &:= \frac{1}{\hat{\sigma}_{n,\hat{w}}^p} \int_0^1 \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(\alpha)} - (\hat{\beta} - \beta) X_{n, \lfloor (n+1)t \rfloor}|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt. \end{aligned}$$

By [Theorem 2.2](#) applied to $\{x_i\}_{i \geq 1}$,

$$\int_0^1 \frac{|X_{n, \lfloor (n+1)t \rfloor}|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt = O_p(1).$$

Since $\hat{\beta} - \beta = o_p(1)$ and $\hat{\sigma}_{n,\hat{w}}^{-1} = O_p(1)$, it follows that

$$I_{n,R}^{(\alpha)} = o_p(1). \quad (223)$$

Also, by [\(222\)](#),

$$I_{n,M}^{(\alpha)} = O_p(1). \quad (224)$$

Inserting the remainder inside the power. If $0 < p \leq 1$, then

$$||u + v|^p - |u|^p| \leq |v|^p, \quad u, v \in \mathbb{R}.$$

Thus, using $u = M_{n, \lfloor (n+1)t \rfloor}^{(\alpha)}$ and $v = -(\hat{\beta} - \beta) X_{n, \lfloor (n+1)t \rfloor}$,

$$|I_{n,M+R}^{(\alpha)} - I_{n,M}^{(\alpha)}| \leq I_{n,R}^{(\alpha)} = o_p(1).$$

If $p > 1$, define

$$N_{n,M}^{(\alpha)} := \left(I_{n,M}^{(\alpha)} \right)^{1/p}, \quad N_{n,R}^{(\alpha)} := \left(I_{n,R}^{(\alpha)} \right)^{1/p}, \quad N_{n,M+R}^{(\alpha)} := \left(I_{n,M+R}^{(\alpha)} \right)^{1/p}.$$

Minkowski's inequality gives

$$\left| N_{n,M+R}^{(\alpha)} - N_{n,M}^{(\alpha)} \right| \leq N_{n,R}^{(\alpha)} = o_p(1).$$

Moreover, $N_{n,M}^{(\alpha)} = O_p(1)$ by (224), and hence $N_{n,M+R}^{(\alpha)} = O_p(1)$. Therefore,

$$\left| I_{n,M+R}^{(\alpha)} - I_{n,M}^{(\alpha)} \right| = \left| \left(N_{n,M+R}^{(\alpha)} \right)^p - \left(N_{n,M}^{(\alpha)} \right)^p \right| = o_p(1).$$

Thus, for all $0 < p < \infty$,

$$I_{n,M+R}^{(\alpha)} - I_{n,M}^{(\alpha)} = o_p(1). \quad (225)$$

Combining (222) with (225) gives

$$\frac{1}{\widehat{\sigma}_{n,\widehat{w}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \widehat{w}_i \right|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt \xrightarrow{d} \int_0^1 \frac{|B(t)|^p}{q(t)} dt.$$

This proves (196). □

4.3.4 Proof of Theorem 4.4

Proof. Assume that Assumption 3.2 holds and that H_0^α holds. Let $1 \leq p < \infty$, and put

$$\Delta_n(t) := \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right).$$

By Theorem 2.4 applied to the i.i.d. sequence $\{\varepsilon_i\}_{i \geq 1}$,

$$\frac{1}{\sqrt{2C_p \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\varepsilon}^p} \int_0^1 \frac{\left| M_{n, \lfloor (n+1)t \rfloor}^{(\alpha)} \right|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt - 2m_p \log n \right] \xrightarrow{d} N(0, 1). \quad (226)$$

Consequently,

$$\frac{1}{\widehat{\sigma}_{n,\varepsilon}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(\alpha)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt = 2m_p \log n + O_p(\sqrt{\log n}). \quad (227)$$

Replacing $\widehat{\sigma}_{n,\varepsilon}$ by $\widehat{\sigma}_{n,\widehat{w}}$. We show that

$$\begin{aligned} & \frac{1}{\sqrt{2C_p \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{w}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(\alpha)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt - 2m_p \log n \right] \\ & - \frac{1}{\sqrt{2C_p \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\varepsilon}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(\alpha)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt - 2m_p \log n \right] = o_p(1). \end{aligned} \quad (228)$$

Indeed, the left-hand side of (228) is bounded in absolute value by

$$\frac{1}{\sqrt{2C_p \log n}} \left| \frac{\widehat{\sigma}_{n,\varepsilon}^p}{\widehat{\sigma}_{n,\widehat{w}}^p} - 1 \right| \frac{1}{\widehat{\sigma}_{n,\varepsilon}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(\alpha)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt.$$

By (210), (208), and (209),

$$\left| \frac{\widehat{\sigma}_{n,\varepsilon}^p}{\widehat{\sigma}_{n,\widehat{w}}^p} - 1 \right| = O_p(n^{-1/2}).$$

Together with (227), this gives

$$\frac{1}{\sqrt{\log n}} O_p(n^{-1/2}) \{2m_p \log n + O_p(\sqrt{\log n})\} = o_p(1),$$

which proves (228). Hence, by (226),

$$\frac{1}{\sqrt{2C_p \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{w}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(\alpha)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt - 2m_p \log n \right] \xrightarrow{d} N(0, 1). \quad (229)$$

The slope-estimation remainder. By Lemma 4.1, with $k = \lfloor(n+1)t\rfloor$,

$$\sum_{i=1}^{\lfloor(n+1)t\rfloor} \widehat{w}_i - \frac{\lfloor(n+1)t\rfloor}{n} \sum_{i=1}^n \widehat{w}_i = M_{n,\lfloor(n+1)t\rfloor}^{(\alpha)} - (\widehat{\beta} - \beta) X_{n,\lfloor(n+1)t\rfloor}.$$

Define

$$\begin{aligned} I_{n,M}^{(\alpha,p)} &:= \frac{1}{\widehat{\sigma}_{n,\widehat{w}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(\alpha)}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt, \\ I_{n,R}^{(\alpha,p)} &:= \frac{|\widehat{\beta} - \beta|^p}{\widehat{\sigma}_{n,\widehat{w}}^p} \int_0^1 \frac{|X_{n,\lfloor(n+1)t\rfloor}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt, \\ I_{n,M+R}^{(\alpha,p)} &:= \frac{1}{\widehat{\sigma}_{n,\widehat{w}}^p} \int_0^1 \frac{|M_{n,\lfloor(n+1)t\rfloor}^{(\alpha)} - (\widehat{\beta} - \beta) X_{n,\lfloor(n+1)t\rfloor}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt. \end{aligned}$$

By [Theorem 2.4](#) applied to the i.i.d. sequence $\{x_i\}_{i \geq 1}$,

$$\int_0^1 \frac{|X_{n,\lfloor(n+1)t\rfloor}|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt = O_p(\log n).$$

Since $\widehat{\beta} - \beta = O_p(n^{-1/2})$ and $\widehat{\sigma}_{n,\widehat{w}}^{-1} = O_p(1)$, it follows that

$$I_{n,R}^{(\alpha,p)} = O_p(n^{-p/2} \log n). \quad (230)$$

In particular,

$$\frac{I_{n,R}^{(\alpha,p)}}{\sqrt{\log n}} = o_p(1). \quad (231)$$

Also, by [\(229\)](#),

$$I_{n,M}^{(\alpha,p)} = O_p(\log n). \quad (232)$$

Inserting the remainder inside the power. Since $p \geq 1$, define

$$N_{n,M}^{(\alpha,p)} := \left(I_{n,M}^{(\alpha,p)} \right)^{1/p}, \quad N_{n,R}^{(\alpha,p)} := \left(I_{n,R}^{(\alpha,p)} \right)^{1/p}, \quad N_{n,M+R}^{(\alpha,p)} := \left(I_{n,M+R}^{(\alpha,p)} \right)^{1/p}.$$

Minkowski's inequality gives

$$\left| N_{n,M+R}^{(\alpha,p)} - N_{n,M}^{(\alpha,p)} \right| \leq N_{n,R}^{(\alpha,p)}.$$

By [\(230\)](#),

$$N_{n,R}^{(\alpha,p)} = O_p(n^{-1/2} (\log n)^{1/p}),$$

and by (232),

$$N_{n,M}^{(\alpha,p)} = O_p((\log n)^{1/p}).$$

Thus

$$N_{n,M+R}^{(\alpha,p)} = O_p((\log n)^{1/p}).$$

For nonnegative a and b , the mean value theorem gives

$$|a^p - b^p| = p\xi^{p-1}|a - b| \leq p(a + b)^{p-1}|a - b|,$$

for some ξ between a and b . Therefore,

$$\begin{aligned} |I_{n,M+R}^{(\alpha,p)} - I_{n,M}^{(\alpha,p)}| &= \left| \left(N_{n,M+R}^{(\alpha,p)} \right)^p - \left(N_{n,M}^{(\alpha,p)} \right)^p \right| \\ &\leq p \left(N_{n,M+R}^{(\alpha,p)} + N_{n,M}^{(\alpha,p)} \right)^{p-1} N_{n,R}^{(\alpha,p)} \\ &= O_p((\log n)^{(p-1)/p}) O_p(n^{-1/2}(\log n)^{1/p}) \\ &= O_p(n^{-1/2} \log n). \end{aligned}$$

Consequently,

$$\frac{I_{n,M+R}^{(\alpha,p)} - I_{n,M}^{(\alpha,p)}}{\sqrt{\log n}} = o_p(1). \quad (233)$$

Combining (229) with (233) gives

$$\frac{1}{\sqrt{2C_p \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{w}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \widehat{w}_i \right|^p}{n^{p/2} \Delta_n(t)^{1+p/2}} dt - 2m_p \log n \right] \xrightarrow{d} N(0, 1).$$

This proves (197). □

4.3.5 Proof of Theorem 4.5

Proof. We prove the two parts separately. In both parts, the corresponding joint limit theorem is first applied to the leading i.i.d. process $\{\varepsilon_i\}_{i \geq 1}$. The equivalences established in the preceding proofs then transfer the leading joint limit to the residual-based statistics by Slutsky's theorem for random vectors.

PART (i).

Assume that [Assumption 3.1](#) holds and that H_0^α holds. Let $q(t) \in \mathcal{Q}$, suppose that $I_0^1(q, c) < \infty$ for some or possibly all $c > 0$, and assume that

$$\mathbb{E}[\varepsilon^2 \log \log(|\varepsilon| + 1)] < \infty.$$

Define the leading statistics

$$S_{n,q}^{(\alpha)} := \frac{1}{\hat{\sigma}_{n,\varepsilon}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(\alpha)}|}{\sqrt{n} q(k/n)},$$

$$D_n^{(\alpha)} := A(n) \frac{1}{\hat{\sigma}_{n,\varepsilon}} \max_{1 \leq k < n} \frac{|M_{n,k}^{(\alpha)}|}{\sqrt{n} q^*(k/n)} - D(n).$$

By [Theorem 2.5](#)(i) applied to $\{\varepsilon_i\}_{i \geq 1}$, for all $x, y \in \mathbb{R}$,

$$\mathbb{P} \left\{ S_{n,q}^{(\alpha)} \leq x, D_n^{(\alpha)} \leq y \right\} \longrightarrow \mathbb{P} \left\{ \sup_{0 < t < 1} \frac{|B(t)|}{q(t)} \leq x \right\} \exp(-2e^{-y}). \quad (234)$$

Now define the residual-based statistics

$$\hat{S}_{n,q}^{(\alpha)} := \frac{1}{\hat{\sigma}_{n,\hat{w}}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \hat{w}_i - \frac{k}{n} \sum_{i=1}^n \hat{w}_i \right|}{\sqrt{n} q(k/n)},$$

$$\hat{D}_n^{(\alpha)} := A(n) \frac{1}{\hat{\sigma}_{n,\hat{w}}} \max_{1 \leq k < n} \frac{\left| \sum_{i=1}^k \hat{w}_i - \frac{k}{n} \sum_{i=1}^n \hat{w}_i \right|}{\sqrt{n} q^*(k/n)} - D(n).$$

The proof of [Theorem 4.1](#) showed, using [Lemma 4.1](#), (211), and (215), that

$$\hat{S}_{n,q}^{(\alpha)} = S_{n,q}^{(\alpha)} + o_p(1). \quad (235)$$

Similarly, the proof of [Theorem 4.2](#) showed, using [Lemma 4.1](#), (218), and (220), that

$$\hat{D}_n^{(\alpha)} = D_n^{(\alpha)} + o_p(1). \quad (236)$$

Thus,

$$\left(\hat{S}_{n,q}^{(\alpha)}, \hat{D}_n^{(\alpha)} \right) = \left(S_{n,q}^{(\alpha)}, D_n^{(\alpha)} \right) + o_p(1)$$

componentwise. Together with (234), Slutsky's theorem for random vectors gives

$$\left(\widehat{S}_{n,q}^{(\alpha)}, \widehat{D}_n^{(\alpha)}\right) \xrightarrow{d} \left(\sup_{0 < t < 1} \frac{|B(t)|}{q(t)}, G\right),$$

with the product-form limiting distribution in (234). Therefore, for all $x, y \in \mathbb{R}$,

$$\mathbb{P}\left\{\widehat{S}_{n,q}^{(\alpha)} \leq x, \widehat{D}_n^{(\alpha)} \leq y\right\} \longrightarrow \mathbb{P}\left\{\sup_{0 < t < 1} \frac{|B(t)|}{q(t)} \leq x\right\} \exp(-2e^{-y}),$$

which proves (198).

PART (ii).

Assume that **Assumption 3.2** holds and that H_0^α holds. Let $0 < p < \infty$, let $1 \leq \ell < \infty$, let $q(t) \in \mathcal{Q}$, and suppose that

$$\int_0^1 \frac{(t(1-t))^{p/2}}{q(t)} dt < \infty.$$

Let

$$\Delta_n(t) := \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n}\right).$$

Define the leading integral statistics

$$\begin{aligned} L_{n,p}^{(\alpha)} &:= \frac{1}{\widehat{\sigma}_{n,\varepsilon}^p} \int_0^1 \frac{\left|M_{n,\lfloor (n+1)t \rfloor}^{(\alpha)}\right|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt, \\ J_{n,\ell}^{(\alpha)} &:= \frac{1}{\sqrt{2C_\ell \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\varepsilon}^\ell} \int_0^1 \frac{\left|M_{n,\lfloor (n+1)t \rfloor}^{(\alpha)}\right|^\ell}{n^{\ell/2} \Delta_n(t)^{1+\ell/2}} dt - 2m_\ell \log n \right]. \end{aligned}$$

By **Theorem 2.5(ii)** applied to $\{\varepsilon_i\}_{i \geq 1}$, for all $x, y \in \mathbb{R}$,

$$\mathbb{P}\left\{L_{n,p}^{(\alpha)} \leq x, J_{n,\ell}^{(\alpha)} \leq y\right\} \longrightarrow \mathbb{P}\left\{\int_0^1 \frac{|B(t)|^p}{q(t)} dt \leq x\right\} \Phi(y). \quad (237)$$

Define the residual-based integral statistics

$$\begin{aligned}\widehat{L}_{n,p}^{(\alpha)} &:= \frac{1}{\widehat{\sigma}_{n,\widehat{w}}^p} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \widehat{w}_i \right|^p}{n^{p/2} q\left(\frac{\lfloor (n+1)t \rfloor}{n}\right)} dt, \\ \widehat{J}_{n,\ell}^{(\alpha)} &:= \frac{1}{\sqrt{2C_\ell \log n}} \left[\frac{1}{\widehat{\sigma}_{n,\widehat{w}}^\ell} \int_0^1 \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \widehat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \widehat{w}_i \right|^\ell}{n^{\ell/2} \Delta_n(t)^{1+\ell/2}} dt - 2m_\ell \log n \right].\end{aligned}$$

The proof of [Theorem 4.3](#) showed, using [Lemma 4.1](#), [\(211\)](#), and [\(225\)](#), that

$$\widehat{L}_{n,p}^{(\alpha)} = L_{n,p}^{(\alpha)} + o_p(1). \quad (238)$$

The proof of [Theorem 4.4](#), with the power p there replaced by ℓ , showed, using [\(228\)](#) and [\(233\)](#), that

$$\widehat{J}_{n,\ell}^{(\alpha)} = J_{n,\ell}^{(\alpha)} + o_p(1). \quad (239)$$

Therefore,

$$\left(\widehat{L}_{n,p}^{(\alpha)}, \widehat{J}_{n,\ell}^{(\alpha)} \right) = \left(L_{n,p}^{(\alpha)}, J_{n,\ell}^{(\alpha)} \right) + o_p(1)$$

componentwise. Together with [\(237\)](#), Slutsky's theorem for random vectors gives

$$\left(\widehat{L}_{n,p}^{(\alpha)}, \widehat{J}_{n,\ell}^{(\alpha)} \right) \xrightarrow{d} \left(\int_0^1 \frac{|B(t)|^p}{q(t)} dt, N \right),$$

with the product-form limiting distribution in [\(237\)](#). Consequently, for all $x, y \in \mathbb{R}$,

$$\mathbb{P} \left\{ \widehat{L}_{n,p}^{(\alpha)} \leq x, \widehat{J}_{n,\ell}^{(\alpha)} \leq y \right\} \longrightarrow \mathbb{P} \left\{ \int_0^1 \frac{|B(t)|^p}{q(t)} dt \leq x \right\} \Phi(y),$$

which proves [\(200\)](#). The proof is complete. $\square$

4.3.6 Proof of [Theorem 4.6](#)

Proof. Assume that [Assumption 3.2](#) holds and that H_0^α holds. Let

$$\Delta_n(t) := \frac{\lfloor (n+1)t \rfloor}{n} \left(1 - \frac{\lfloor (n+1)t \rfloor}{n} \right).$$

We prove the two convergences simultaneously by working with $j \in \{1, 2\}$, where

$$J_n^{(1)} = \left[\frac{d_n}{e_n}, 1 - d_n \right], \quad J_n^{(2)} = \left[d_n, 1 - \frac{d_n}{e_n} \right].$$

By [Theorem 2.6](#) applied to the i.i.d. sequence $\{\varepsilon_i\}_{i \geq 1}$,

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(\alpha)}|}{\hat{\sigma}_{n, \varepsilon} \sqrt{n} \Delta_n(t)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|. \quad (240)$$

In particular,

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(\alpha)}|}{\hat{\sigma}_{n, \varepsilon} \sqrt{n} \Delta_n(t)} = O_p(1). \quad (241)$$

Using [\(211\)](#), [\(240\)](#) also gives

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(\alpha)}|}{\hat{\sigma}_{n, \hat{w}} \sqrt{n} \Delta_n(t)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|. \quad (242)$$

The slope-estimation remainder. By [Lemma 4.1](#), with $k = \lfloor (n+1)t \rfloor$,

$$\sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \hat{w}_i = M_{n, \lfloor (n+1)t \rfloor}^{(\alpha)} - (\hat{\beta} - \beta) X_{n, \lfloor (n+1)t \rfloor}.$$

We show that, for $j = 1, 2$,

$$d_n^{1/2} \frac{|\hat{\beta} - \beta|}{\hat{\sigma}_{n, \hat{w}}} \sup_{t \in J_n^{(j)}} \frac{|X_{n, \lfloor (n+1)t \rfloor}|}{\sqrt{n} \Delta_n(t)} = o_p(1). \quad (243)$$

By [Theorem 2.6](#) applied to the i.i.d. sequence $\{x_i\}_{i \geq 1}$,

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|X_{n, \lfloor (n+1)t \rfloor}|}{\hat{\sigma}_{n, x} \sqrt{n} \Delta_n(t)} = O_p(1),$$

where

$$\hat{\sigma}_{n, x}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \xrightarrow{p} \text{Var}(x) > 0.$$

Thus,

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|X_{n, \lfloor (n+1)t \rfloor}|}{\sqrt{n} \Delta_n(t)} = O_p(1).$$

Together with $\hat{\beta} - \beta = O_p(n^{-1/2})$ and $\hat{\sigma}_{n, \hat{w}}^{-1} = O_p(1)$, this proves (243).

Completion of the proof. Using the triangle inequality and Lemma 4.1, for $j = 1, 2$,

$$\begin{aligned} & \left| d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \hat{w}_i \right|}{\hat{\sigma}_{n, \hat{w}} \sqrt{n} \Delta_n(t)} - d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{|M_{n, \lfloor (n+1)t \rfloor}^{(\alpha)}|}{\hat{\sigma}_{n, \hat{w}} \sqrt{n} \Delta_n(t)} \right| \\ & \leq d_n^{1/2} \frac{|\hat{\beta} - \beta|}{\hat{\sigma}_{n, \hat{w}}} \sup_{t \in J_n^{(j)}} \frac{|X_{n, \lfloor (n+1)t \rfloor}|}{\sqrt{n} \Delta_n(t)} = o_p(1), \end{aligned}$$

by (243). Combining this with (242) gives

$$d_n^{1/2} \sup_{t \in J_n^{(j)}} \frac{\left| \sum_{i=1}^{\lfloor (n+1)t \rfloor} \hat{w}_i - \frac{\lfloor (n+1)t \rfloor}{n} \sum_{i=1}^n \hat{w}_i \right|}{\hat{\sigma}_{n, \hat{w}} \sqrt{n} \Delta_n(t)} \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|.$$

Taking $j = 1$ gives (201), and taking $j = 2$ gives (202). This completes the proof. $\square$

5 Concluding Remarks

This paper developed nonparametric asymptotic tests for changes in regression coefficients in a simple linear regression model with random design. The slope-change and intercept-change problems were treated separately. For each problem, a scalar residual-based tied-down partial-sum process was constructed so that, under the corresponding null hypothesis, the resulting statistics have the same limiting behavior as suitable functionals of the partial-sum processes used in nonparametric mean-change testing.

The main contribution is a broad collection of asymptotic results for component-specific residual-process statistics. For both the slope-change and intercept-change problems, we established limiting distributions for supremum and integral statistics with Brownian-bridge-based limits, supremum and integral Darling–Erdős-type statistics, selected joint limits, and Rényi-type supremum statistics with Wiener-process limits. The supremum results connect with existing residual-based and tied-down partial-sum approaches in regression

change-point analysis, while the integral, joint, and Rényi-type results extend the range of available component-specific asymptotic tools.

Although the slope-change and intercept-change statistics are constructed from different residual quantities, the same general strategy is used in both cases. The residual-based process is shown to be asymptotically equivalent to a tied-down partial-sum process to which the mean-change limit theorems can be applied.

The results also emphasize that different choices of a functional, weight, and normalization lead to different asymptotic procedures. Supremum statistics measure the largest in absolute value suitably normalized mean difference over possible change-point locations, while integral statistics average deviations across the tied-down partial-sum path. Darling–Erdős-type and Rényi-type statistics use different normalizations and have different limiting distributions. Thus, the collection of statistics developed here gives several ways to summarize the same underlying residual partial-sum behavior.

The focus of this paper was asymptotic theory. A natural next step is a detailed finite-sample study of the proposed regression tests. Such a study should compare type I error control and power across the different statistics, sample sizes, change sizes, covariate distributions, error distributions, and change-point locations. The mean-change results discussed earlier suggest that the choice of a functional and weight can substantially affect the balance between power and type I error, but these comparisons need to be carried out separately for the regression residual processes.

Several extensions are possible. The joint limit results obtained here may be used to construct combined statistics from different functionals or normalizations, although their critical values and finite-sample behavior would require separate numerical investigation. Another direction is to extend the present collection of limit theorems to broader regression change-point frameworks in which the intercept and slope may both change, possibly at different times. Existing work has developed procedures for distinguishing changes in regression parameters, but the integral, integral Darling–Erdős-type, joint, and Rényi-type directions considered here suggest further asymptotic and finite-sample questions in such settings. Extensions to multiple regression, dependent observations, heteroscedastic errors, and other structural change settings would further broaden the applicability of the residual-

based weighted partial-sum approach.

References

- Abd-Rabou, A.-E. S., & Aly, E.-E. A. A. (2000). A comparison of weighted tests for a slope change in simple regression. *Biometrical Journal*, 42(8), 1033–1038. [https://doi.org/10.1002/1521-4036\(200012\)42:8<1033::AID-BIMJ1033>3.0.CO;2-0](https://doi.org/10.1002/1521-4036(200012)42:8<1033::AID-BIMJ1033>3.0.CO;2-0)
- Berkes, I., Horváth, L., Hušková, M., & Steinebach, J. (2004). Applications of permutations to the simulations of critical values. *Journal of Nonparametric Statistics*, 16, 197–216. <https://doi.org/10.1080/10485250310001624747>
- Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society. Series B*, 37, 149–192.
- Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3), 591–605. <https://doi.org/10.2307/1910133>
- Csörgő, M., Csörgő, S., Horváth, L., & Mason, D. M. (1986). Weighted empirical and quantile processes. *The Annals of Probability*, 14(1), 31–85. <https://doi.org/10.1214/aop/1176992617>
- Csörgő, M., & Horváth, L. (1997). *Limit theorems in change-point analysis*. Wiley.
- Gallanosa, J., & Martsynyuk, Y. V. (2026). Powerful nonparametric asymptotic tests for change in the mean with reduced type i errors. *Mathematics*, 14, 78. <https://doi.org/10.3390/math14010078>
- Gombay, E. (2010). Change detection in linear regression with time series errors. *Canadian Journal of Statistics*, 38(1), 65–79. <https://doi.org/10.1002/cjs.10043>
- Gombay, E., & Horváth, L. (1994). Limit theorems for change in linear regression. *Journal of Multivariate Analysis*, 48(1), 43–69. [https://doi.org/10.1016/0047-259X\(94\)80004-F](https://doi.org/10.1016/0047-259X(94)80004-F)
- Horváth, L., Miller, C., & Rice, G. (2020). A new class of change point test statistics of rényi type. *Journal of Business & Economic Statistics*, 38(3), 570–579. <https://doi.org/10.1080/07350015.2018.1537923>
- Horváth, L., & Rice, G. (2014). Extensions of some classical methods in change point analysis. *TEST*, 23(2), 219–255. <https://doi.org/10.1007/s11749-014-0368-4>

- Horváth, L., Rice, G., & Zhao, Y. (2023). Testing for changes in linear models using weighted residuals. *Journal of Multivariate Analysis*, 198, 105210. <https://doi.org/10.1016/j.jmva.2023.105210>
- Martsynyuk, Y. V. (2016). Testing for change in the mean via convergence in distribution of sup-functionals of weighted tied-down partial sums processes. *Mathematical Methods of Statistics*, 25(3), 219–232. <https://doi.org/10.3103/S1066530716030054>
- Martsynyuk, Y. V. (2021). Nonparametric detection of change in the slope and intercept in linear structural errors-in-variables models. *Statistics & Probability Letters*, 169, 108957. <https://doi.org/10.1016/j.spl.2020.108957>
- Olmo, J., Pilbeam, K., & Pouliot, W. (2011). Detecting the presence of insider trading via structural break tests. *Journal of Banking & Finance*, 35(11), 2820–2828. <https://doi.org/10.1016/j.jbankfin.2011.03.013>
- Orasch, M., & Pouliot, W. (2004). Tabulating weighted sup-norm functionals used in change-point analysis. *Journal of Statistical Computation and Simulation*, 74, 249–276. <https://doi.org/10.1080/0094965031000148703>
- Pouliot, W. (2016). Robust tests for change in intercept and slope in linear regression models with application to manager performance in the mutual fund industry. *Economic Modelling*, 58, 523–534. <https://doi.org/10.1016/j.econmod.2016.03.011>
- Sen, P. K. (1980). Asymptotic theory of some tests for a possible change in the regression slope occurring at an unknown time point. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 52(2), 203–218. <https://doi.org/10.1007/BF00531606>

Chapter 5

Discussion

1 Summary of Contributions and Novel Developments

This thesis studied nonparametric asymptotic procedures for confidence interval estimation and change-point testing. The common methodological theme was the use of functionals of appropriately normalized and weighted partial-sum-type processes. Although the manuscript chapters considered different inferential problems, they were connected by the same general idea: suitable choices of functionals and weight functions can be used to construct asymptotic procedures with the aim of improving, or better balancing, finite-sample performance.

Chapter 2 contributed to confidence interval construction for the mean. The chapter developed functional asymptotic confidence intervals based on integral functionals of weighted Student processes and compared their finite-sample coverage probabilities and expected lengths with those of the classical CLT-based interval. The proposed intervals had substantially higher finite-sample coverages for right-skewed distributions in the domain of attraction of the normal law, including cases where the classical interval significantly undercovers. The results also showed that the improved coverage was not explained only by longer intervals, since the best proposed intervals often had higher empirical coverage probabilities than equal-expected-length versions of the classical interval.

Chapter 3 contributed to nonparametric testing for an at-most-one change in the mean. The chapter studied weighted supremum and weighted integral statistics based on a tied-down partial-sum process. Weighted supremum tests had been considered previously, but weighted integral tests for this problem had not been studied to the same extent. The numerical results showed a finite-sample trade-off among type I error control, power for

changes near the endpoints of the sample, and power for changes near the middle of the sample. By choosing suitable weight functions, the chapter identified weighted tests with favorable finite-sample behavior. In particular, the weighted integral test emerged as one of the most powerful and well-balanced tests considered, with strong power across all change-point locations and well-controlled type I errors.

Chapter 4 contributed to regression change-point analysis by extending weighted partial-sum methods from the mean-change problem of Chapter 3 to simple linear regression with random design. The slope-change and intercept-change problems were treated separately. For each problem, residual-based tied-down partial-sum processes were constructed, and asymptotic null theory was developed for several classes of test statistics. These include supremum and integral statistics with Brownian-bridge-based limits, supremum and integral Darling–Erdős-type statistics, selected joint limits, and Rényi-type supremum statistics with Wiener-process limits. The chapter provides a theoretical foundation for future finite-sample studies of tests for changes in regression coefficients.

Taken together, these contributions show that the thesis provides more than three separate manuscript chapters. The distinct contribution of the thesis is the development and study of a common weighted-process approach for nonparametric asymptotic inference. In the confidence-interval problem, weighted Student processes and their functionals are used to improve interval coverages while keeping interval lengths under consideration. In the mean-change problem, functionals of weighted tied-down partial-sum processes are used to balance power across change-point locations and type I error control. In the regression setting, residual-based tied-down partial sums extend the mean-change approach of Chapter 3 to tests for changes in regression coefficients and provide the asymptotic foundation needed for future numerical comparisons.

2 Overall Impact and Future Directions

The broader impact of the thesis is that it demonstrates how weight functions and functionals can be used as practical tools for modifying the behavior of nonparametric asymptotic procedures. In the confidence-interval setting, different weights and functionals can af-

fect coverage probability and expected length. In the change-point testing setting, they can affect type I error control, overall power, and how power is distributed across possible change-point locations.

The findings also clarify important trade-offs. In confidence interval estimation, improved coverage may come with longer expected length. In change-point testing, improved sensitivity to changes near the beginning or end of the sample may affect type I error control or power for changes near the middle of the sample. The thesis therefore suggests that there may not be a single universally best functional or weight function. Instead, suitable choices depend on the inferential goal and on the finite-sample properties one wishes to prioritize.

Several directions remain for future work. First, the regression change-point procedures developed in Chapter 4 should be studied numerically. The results of Chapter 3 suggest that the choice of a functional and weight function can play an important role in balancing type I error control and power. Accordingly, future simulations should compare the proposed regression tests with existing procedures and investigate whether similar finite-sample improvements can be obtained for changes in regression coefficients.

Second, further work is needed on weight function selection. This thesis studies special forms of weight functions that perform well in the settings considered, but more systematic or data-adaptive approaches may be possible. Such methods would need to preserve the relevant asymptotic theory while improving finite-sample performance for particular classes of distributions or change-point alternatives.

Third, the regression framework of Chapter 4 can be extended beyond simple linear regression with random design. Natural directions include multiple regression, dependent observations, heteroscedastic errors, and broader structural change settings in which more than one regression coefficient may change. These extensions would broaden the applicability of the weighted partial-sum methodology developed in this thesis.

Overall, the thesis shows that suitable choices of weight functions and functionals can be used to shape nonparametric asymptotic procedures toward more favorable finite-sample performance. This provides a unified contribution to confidence interval estimation, mean-change testing, and regression change-point analysis.

Full Bibliography

- Abd-Rabou, A.-E. S., & Aly, E.-E. A. A. (2000). A comparison of weighted tests for a slope change in simple regression. *Biometrical Journal*, 42(8), 1033–1038. [https://doi.org/10.1002/1521-4036\(200012\)42:8<1033::AID-BIMJ1033>3.0.CO;2-0](https://doi.org/10.1002/1521-4036(200012)42:8<1033::AID-BIMJ1033>3.0.CO;2-0)
- Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain “goodness of fit” criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 23, 193–212. <https://doi.org/10.1214/aoms/1177729437>
- Bentkus, V., & Götze, F. (1996). The berry–esseen bound for student’s statistic. *The Annals of Probability*, 24(1), 491–503. <https://doi.org/10.1214/aop/1042644728>
- Berkes, I., Horváth, L., Hušková, M., & Steinebach, J. (2004). Applications of permutations to the simulations of critical values. *Journal of Nonparametric Statistics*, 16, 197–216. <https://doi.org/10.1080/10485250310001624747>
- Bhardwaj, S. (2022). *Finite-sample properties and applicability of functional clt based confidence intervals for a population mean* [M.Sc. thesis]. University of Manitoba. <http://hdl.handle.net/1993/36759>
- Bhardwaj, S., Gallanosa, J., & Martsynyuk, Y. V. (2025). Asymptotic confidence intervals for the mean with increased finite-sample coverage probabilities. *Mathematics*, 13, 3931. <https://doi.org/10.3390/math13243931>
- Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society. Series B*, 37, 149–192.
- Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3), 591–605. <https://doi.org/10.2307/1910133>
- Csörgő, M. (2002). A glimpse of the impact of páł erdős on probability and statistics. *Canadian Journal of Statistics*, 30, 493–556. <https://doi.org/10.2307/3316095>

- Csörgő, M., Csörgő, S., Horváth, L., & Mason, D. M. (1986). Weighted empirical and quantile processes. *The Annals of Probability*, 14(1), 31–85. <https://doi.org/10.1214/aop/1176992617>
- Csörgő, M., & Horváth, L. (1988). Nonparametric methods for changepoint problems. In *Quality control and reliability* (pp. 403–425, Vol. 7). Elsevier. [https://doi.org/10.1016/S0169-7161\(88\)07022-1](https://doi.org/10.1016/S0169-7161(88)07022-1)
- Csörgő, M., & Horváth, L. (1997). *Limit theorems in change-point analysis*. Wiley.
- Csörgő, M., Szyszkowicz, B., & Wang, Q. (2003). Donsker's theorem for self-normalized partial sums processes. *The Annals of Probability*, 31(3), 1228–1240. <https://doi.org/10.1214/aop/1055425777>
- Csörgő, M., Szyszkowicz, B., & Wang, Q. (2004). On weighted approximations and strong limit theorems for self-normalized partial sums processes. In *Asymptotic methods in stochastics* (pp. 489–521, Vol. 44). American Mathematical Society.
- Csörgő, M., Szyszkowicz, B., & Wang, Q. (2008). On weighted approximations in $d[0,1]$ with applications to self-normalized partial sum processes. *Acta Mathematica Hungarica*, 121(4), 307–332. <https://doi.org/10.1007/s10474-008-7216-5>
- Csörgő, S., & Horváth, L. (1981). On the koziol–green model for random censorship. *Biometrika*, 68, 391–401. <https://doi.org/10.1093/biomet/68.2.391>
- Eastwood, B. J., & Eastwood, V. R. (1998). Tabulating weighted sup-norm functionals of brownian bridges via monte carlo simulation. In *Asymptotic methods in probability and statistics* (pp. 707–719). North-Holland. <https://doi.org/10.1016/B978-044450083-0/50046-0>
- Erdős, P., & Kac, M. (1946). On certain limit theorems of the theory of probability. *Bulletin of the American Mathematical Society*, 52, 292–302. <https://doi.org/10.1090/S0002-9904-1946-08560-2>
- Gallanosa, J., & Martsynyuk, Y. V. (2026). Powerful nonparametric asymptotic tests for change in the mean with reduced type i errors. *Mathematics*, 14, 78. <https://doi.org/10.3390/math14010078>
- Gombay, E. (2010). Change detection in linear regression with time series errors. *Canadian Journal of Statistics*, 38(1), 65–79. <https://doi.org/10.1002/cjs.10043>

- Gombay, E., & Horváth, L. (1994). Limit theorems for change in linear regression. *Journal of Multivariate Analysis*, 48(1), 43–69. [https://doi.org/10.1016/0047-259X\(94\)80004-F](https://doi.org/10.1016/0047-259X(94)80004-F)
- Horváth, L., Miller, C., & Rice, G. (2020). A new class of change point test statistics of rényi type. *Journal of Business & Economic Statistics*, 38(3), 570–579. <https://doi.org/10.1080/07350015.2018.1537923>
- Horváth, L., & Rice, G. (2014). Extensions of some classical methods in change point analysis. *TEST*, 23(2), 219–255. <https://doi.org/10.1007/s11749-014-0368-4>
- Horváth, L., Rice, G., & Zhao, Y. (2023). Testing for changes in linear models using weighted residuals. *Journal of Multivariate Analysis*, 198, 105210. <https://doi.org/10.1016/j.jmva.2023.105210>
- Logan, B. F., Mallows, C. L., Rice, S. O., & Shepp, L. A. (1973). Limit distributions of self-normalized sums. *Annals of Probability*, 1(5), 788–809. <https://doi.org/10.1214/aop/1176996846>
- Martsynyuk, Y. V. (2016). Testing for change in the mean via convergence in distribution of sup-functionals of weighted tied-down partial sums processes. *Mathematical Methods of Statistics*, 25(3), 219–232. <https://doi.org/10.3103/S1066530716030054>
- Martsynyuk, Y. V. (2021). Nonparametric detection of change in the slope and intercept in linear structural errors-in-variables models. *Statistics & Probability Letters*, 169, 108957. <https://doi.org/10.1016/j.spl.2020.108957>
- Martsynyuk, Y. V., & Tuzov, E. (2016). Exploring functional clt confidence intervals for a population mean in the domain of attraction of the normal law. *Acta Mathematica Hungarica*, 148(2), 493–508. <https://doi.org/10.1007/s10474-016-0595-0>
- Olmo, J., Pilbeam, K., & Pouliot, W. (2011). Detecting the presence of insider trading via structural break tests. *Journal of Banking & Finance*, 35(11), 2820–2828. <https://doi.org/10.1016/j.jbankfin.2011.03.013>
- Orasch, M., & Pouliot, W. (2004). Tabulating weighted sup-norm functionals used in change-point analysis. *Journal of Statistical Computation and Simulation*, 74, 249–276. <https://doi.org/10.1080/0094965031000148703>

- Pouliot, W. (2016). Robust tests for change in intercept and slope in linear regression models with application to manager performance in the mutual fund industry. *Economic Modelling*, 58, 523–534. <https://doi.org/10.1016/j.econmod.2016.03.011>
- Sen, P. K. (1980). Asymptotic theory of some tests for a possible change in the regression slope occurring at an unknown time point. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 52(2), 203–218. <https://doi.org/10.1007/BF00531606>
- Shao, Q.-M. (2005). An explicit berry–esseen bound for student’s t-statistic via stein’s method. In *Stein’s method and applications* (pp. 143–155, Vol. 5). Singapore University Press. https://doi.org/10.1142/9789812567673_0009
- Shao, Q.-M., & Wang, Q. (2013). Self-normalized limit theorems: A survey. *Probability Surveys*, 10, 69–93. <https://doi.org/10.1214/13-PS216>
- Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, 19, 279–281. <https://doi.org/10.1214/aoms/1177730256>
- Szyszkowicz, B. (1991). Weighted stochastic processes under contiguous alternatives. *C. R. Math. Rep. Acad. Sci. Canada*, 13, 211–216.
- Tuzov, E. (2014). *Exploring functional asymptotic confidence intervals for a population mean* [M.Sc. thesis]. University of Manitoba. <http://hdl.handle.net/1993/23426>