

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN MATHÉMATIQUES ET INFORMATIQUE APPLIQUÉES

PAR
FRÉDÉRIC BOUCHARD

UNE CLASSE DE TESTS BASÉE SUR LA FONCTION CARACTÉRISTIQUE
POUR LA DÉTECTION DE CHANGEMENT GRADUEL DANS UNE SÉRIE
CHRONOLOGIQUE

MAI 2025

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

AVANT-PROPOS

Il est impensable de débiter ce travail sans remercier sincèrement Monsieur Jean-François Quessy, mon directeur de recherche, qui a su me transmettre sa passion pour les mathématiques. Il a été un point central de mon parcours universitaire par son implication dans ma réussite et en m'inculquant sa méthode de travail hors pair. Ses précieux conseils ont fait de cette expérience un bagage d'une valeur inestimable, qui me servira certainement lors de mes futurs projets.

Je souhaite également mentionner ma compagne, ma famille et mes amis, envers qui je suis infiniment redevable. Sans ces personnes, la réalisation de ce mémoire aurait été impossible.

Enfin, je veux remercier les professeurs Mhamed Mesfioui et Jean-Sébastien Dessureault, du Département de mathématiques et d'informatique de l'UQTR, pour avoir gracieusement accepté de lire et de commenter ce travail de recherche. Leurs remarques judicieuses ont permis d'améliorer la version finale de mon travail.

Le financement durant mes études a été fourni par un octroi individuel à M. Jean-François Quessy dans le cadre du programme des Subventions à la Découverte du Conseil de Recherche en Sciences Naturelles et en Génie (CRSNG) du Canada, ainsi que par l'Institut des Sciences Mathématiques du Québec.

Table des matières

Avant-propos	ii
Table des matières	iii
Liste des tableaux	vi
Table des figures	viii
1 Préliminaires	1
1.1 Organisation du mémoire	1
1.2 Détecter un changement de régime dans une série chronologique	2
1.2.1 Énoncé de la problématique	2
1.2.2 Un aperçu historique	3
1.2.3 Limites des approches actuelles	6
1.2.4 Voici comment on répondra aux constats #1 et #2	7
1.3 Les « pièces maitresses » de l'article	7
1.3.1 Le modèle à changement graduel	7
1.3.2 Approche de Quessy (2019)	8
1.3.3 Utilisation de la fonction caractéristique	10
1.3.4 Estimation d'une fonction caractéristique	11
1.3.5 Processus de type « α -mixing »	13
1.3.6 Méthode Monte Carlo pour approximer la puissance d'un test	14
1.4 Un condensé sur les processus stochastiques	15
1.4.1 Processus gaussiens	16

1.4.2	Mouvement brownien	17
1.4.3	Processus de Lévy	19
1.4.4	Fonction de répartition empirique	21
1.4.5	Processus empirique	23
1.4.6	Processus séquentiel	25
1.5	Un condensé sur les méthodes de rééchantillonnage	26
1.5.1	La méthode du <i>jackknife</i>	26
1.5.2	La technique du <i>bootstrap</i>	28
1.5.3	La méthode du multiplicateur	29
2	Article	32
2.1	Introduction	33
2.2	Basic tools	35
2.2.1	A criterion for testing the null hypothesis of no change	35
2.2.2	Unbiased estimation of $Q_{\mathbf{K}}$	37
2.3	A general class of test statistics	38
2.3.1	From Székely and Rizzo (2007) to Lévy measures	38
2.3.2	Test statistics based on Lévy measures	39
2.3.3	Some specific Lévy measures	40
2.3.4	Relationship with a test by [24]	41
2.4	Large-sample results	43
2.4.1	Relationship with the sequential empirical process	43
2.4.2	Weak convergence of the test statistics	44
2.4.3	A resampling method for the computation of p-values	45
2.4.4	Consistency of the tests and change-point estimators	47
2.5	Size and power of the proposed tests	48
2.5.1	A warp-speed bootstrap method	48
2.5.2	Ability of the tests to keep their nominal level	49
2.5.3	Investigation of the power of the tests	50
2.6	Illustrations	55

2.6.1	Historic temperature data in the Canadian Far North	55
2.6.2	Standard and Poor's 500	57
2.7	Conclusion	59
2.8	Proofs	61
2.8.1	Proof of Lemma 1	61
2.8.2	Proof of Lemma 2	61
2.8.3	Proof of Proposition 1	62
2.8.4	Proof of Proposition 2	63
2.8.5	Proof of Proposition 3	65
2.8.6	Proof of Proposition 4	68
2.8.7	Proof of Lemma 3	69
Conclusion		72
Bibliographie		73
A Analyse des données du Klementinum		79
B Code Matlab		82
B.1	Simulations de séries chronologiques	82
B.2	Calcul des statistiques	83
B.3	Puissance des tests	88
B.4	Analyse des données du Klementinum	91

Liste des tableaux

2.1	Percentages of rejection of the null hypothesis under $\mathbb{H}_0$, as estimated from $B = 1000$ warp-speed bootstrap replicates and $N = 40$, under Normal and Gamma marginal distributions when $n = 100$ (upper panel) and $n = 200$ (bottom panel)	50
2.2	Percentages of rejection of the null hypothesis, as estimated from $B = 1000$ warp-speed bootstrap replicates and $N = 40$, under independent Normal random variables with a unit standard deviation and a mean that gradually shifts from μ_1 to μ_2 when $n = 100$ (upper panel) and $n = 200$ (bottom panel)	52
2.3	Percentages of rejection of the null hypothesis, as estimated from $B = 1000$ warp-speed bootstrap replicates and $N = 40$, under independent Normal random variables with a zero mean and a standard deviation that gradually shifts from σ_1 to σ_2 when $n = 100$ (upper panel) and $n = 200$ (bottom panel)	53
2.4	Percentages of rejection of the null hypothesis under series of size $n = 100$ with Gamma marginal distributions whose mean and/or standard deviation gradually shifts, as estimated from $B = 1000$ warp-speed bootstrap replicates and $N = 40$	54
2.5	Estimated p-values (in %) of the six newly proposed tests of change-point detection in the monthly maxima collected at the Iqaluit weather station from January 1948 to July 2008	57

2.6	Estimated p-values (in %) of the tests of change-point detection in the daily closing time of the Standard and Poor 500 financial index based on $T_{n,\Sigma}^\nu$ and $T_{n,\vee}^\nu$ for three Levy measures when $\lambda = 3/2$	59
A.1	P-valeurs estimées (en %) de deux tests de détection de temps de changement appliqués à la médiane mensuelle des maxima journaliers observés à l'observatoire de Klementinum à Prague de janvier 1890 à décembre 2004	81

Table des figures

1.1	Données simulées pour illustrer quelques types de changements qui peuvent survenir dans une série chronologique. Figures du haut : changements graduel (à gauche) et abrupt (à droite) dans la moyenne ; figures du bas : changements graduel (à gauche) et abrupt (à droite) dans la variance	9
1.2	Illustration du processus $\mathbb{B}_n$ lorsque $n = 1000$ (en haut) et $n = 10000$ (en bas).	25
2.1	Time series of the monthly maximum temperature (in degrees Celcius) measured at the Iqaluit weather station (Nunavut territory, Canada) from January 1946 to December 2008	56
2.2	Lag-1 differentiated series based on the Standard and Poor's 500 daily closing prices	58
A.1	Séries chronologiques des médianes mensuelles de la température (en degrés Celsius) extraite des données de l'observatoire de Klementinum couvrant la période de janvier 1890 à décembre 2004	80

Chapitre 1

Préliminaires

1.1 Organisation du mémoire

Le cœur de ce mémoire est un article scientifique en cours d'évaluation par les pairs pour la revue *Econometrics and Statistics*. Ce manuscrit constitue l'entièreté du Chapitre 2 du mémoire. Dans cet article, on élabore des tests statistiques qui permettent la détection de changements stochastiques dans une série chronologique. Une particularité de ces tests est qu'ils sont en mesure de détecter des changements qui se font graduellement, en opposition aux changements abrupts. Des estimateurs pour les temps de changement sont également proposés.

Le chapitre 1 permet d'avoir une vision d'ensemble de la problématique abordée et de comprendre les motivations derrière ce projet. On fournit également des détails concernant certaines notions et certains outils qui sont employés dans l'article. Spécifiquement, la sous-section 1.2.1 énonce formellement la problématique qui est traitée dans ce mémoire, alors que la sous-section 1.2.2 situe historiquement le problème. Les sous-sections 1.2.3 et 1.2.4 relèvent quelques limites et faiblesses des approches

actuelles et, surtout, décrivent comment s'y prendre pour les résoudre.

La Section 1.3 donne des détails sur les fondements de l'article. On y décrit formellement le modèle à changements graduels, ainsi que la méthodologie de Quessy [44] basée autour de ce modèle. On définit aussi la fonction caractéristique d'une variable aléatoire et la façon de l'estimer à partir d'un jeu de données. Des détails sur les processus stochastiques « α -mélangeant » (en anglais, α -*mixing*) terminent cette sous-section. La Section 1.4 est consacrée à l'introduction des processus stochastiques. Ces connaissances sont fondamentales pour la construction des statistiques de test dans l'article et pour l'obtention de leurs propriétés asymptotiques. Le concept de mesure de Lévy, qui est étroitement lié aux processus stochastiques, y est aussi abordé. On conclut le Chapitre 1 avec une description de trois méthodes de ré-échantillonnage non paramétriques, à savoir le *jackknife*, le *bootstrap* et la méthode du multiplicateur ; ceci est l'objet de la Section 1.5.

1.2 Détecter un changement de régime dans une série chronologique

1.2.1 Énoncé de la problématique

Soit une réalisation $Y_1, \dots, Y_n$ d'un processus stochastique univarié $(Y_t)_{t \in \mathbb{Z}}$ à valeurs dans $\mathcal{Y} \subseteq \mathbb{R}$. Ici, l'indice $t \in \mathbb{Z}$ représente le temps. Le problème de point de changement consiste à vérifier si le comportement stochastique de $(Y_t)_{t \in \mathbb{Z}}$ est demeuré stable au fil du temps, ou si, à l'inverse, il s'est modifié à partir d'un certain temps $K \in \{1, \dots, n-1\}$. Plus précisément, on cherche à tester l'hypothèse nulle $\mathbb{H}_0$ qui

stipule que, pour tout $y \in \mathcal{Y}$,

$$\mathbb{P}(Y_1 \leq y) = \cdots = \mathbb{P}(Y_n \leq y)$$

en la confrontant à l'hypothèse alternative $\mathbb{H}_1$ qui stipule qu'il existe un temps $K \in \{1, \dots, n-1\}$ tel que

$$\mathbb{P}(Y_1 \leq y) = \cdots = \mathbb{P}(Y_K \leq y) \neq \mathbb{P}(Y_{K+1} \leq y) = \cdots = \mathbb{P}(Y_n \leq y).$$

Ce problème a attiré l'attention de quelques chercheurs dès le milieu du vingtième siècle. On verra qu'en dépit de nombreuses contributions qui ont été faites depuis ce temps, il reste encore des avancées significatives à effectuer dans ce domaine.

La prochaine sous-section fait un survol historique des principales contributions à la pratique et à la théorie concernant la détection de changements. La sous-section qui la suit dresse ensuite un portrait des limites des approches disponibles, ce qui permet de mettre la table sur les contributions effectuées dans ce mémoire de maîtrise.

1.2.2 Un aperçu historique

Les premières contributions concernant la détection et l'estimation de points de changement sont dues à Page [37] et à Page [38]. Dans ces travaux, la confrontation des hypothèses $\mathbb{H}_0$ et $\mathbb{H}_1$ est réduite à la recherche d'un changement potentiel concernant un certain paramètre θ . Ce paramètre peut être, par exemple, la moyenne ou la variance d'une distribution. Les recherches de Page [37, 38] se fondent sur les sommes cumulatives $S_1, \dots, S_n$, où

$$S_r = \sum_{i=1}^r (x_i - \theta).$$

Une statistique globale qui combine toutes ces sommes cumulatives est alors

$$\mathcal{S}_n = \max_{0 \leq i \leq n} \left(S_r - \min_{0 \leq i < r} S_i \right).$$

On rejette alors $\mathbb{H}_0$ lorsque $\mathcal{S}_n > q_\alpha$, où la valeur critique q_α est telle que sous $\mathbb{H}_0$,

$$\mathbb{P}(\mathcal{S}_n > q_\alpha) = \alpha.$$

Ces recherches établissent un cadre qui permettra l'exploration de méthodes de détection de changement plus élaborées. De façon générale, les tests retrouvés dans la littérature se classent selon deux catégories :

- 1) Les tests paramétriques, pour lesquels certaines hypothèses concernant la forme de la distribution sous-jacente sont émises ;
- 2) Les tests non paramétriques, où aucune hypothèse (ou très peu ...) n'est faite sur la loi des observations.

Dans le cas d'un test paramétrique, le but est de tester l'hypothèse d'un changement par rapport à un certain paramètre θ en supposant que les fonctions de répartition $F_1, \dots, F_n$ associées aux observations appartiennent à une famille paramétrique. Par exemple, on pourrait supposer que $F_1, \dots, F_n$ appartiennent à la famille des lois normales, de sorte que

$$F_j(y) = \Phi \left(\frac{y - \mu_j}{\sigma_j} \right),$$

où Φ est la fonction de répartition de la loi $\mathcal{N}(0, 1)$. Il s'agit justement du cadre utilisé par Chernoff and Zacks [11], qui proposent une méthode de détection de changements dans la moyenne de données distribuées selon la loi normale. Par ailleurs, Hinkley [21] élargit cette approche en s'intéressant à un contexte plus général concernant des distributions autres que la normale. Hinkley [21] propose aussi un cadre novateur qui permet la détection de ruptures multiples.

Un grand nombre de méthodologies ont été développées depuis, si bien qu'en dresser une liste exhaustive n'est pas réaliste. On peut néanmoins référer le lecteur à l'ouvrage de Chen and Gupta [10], qui propose un recensement des méthodes disponible. De plus, pour montrer que cette branche de la recherche en statistique est toujours active, on peut citer les travaux plus récents de Perreault, Luc and Bernier, J. and Bobée, Bernard and Parent, Éric [39], Zou et al. [61] et Fotopoulos et al. [16].

Les tests non paramétriques permettent de s'affranchir de la nécessité d'émettre une hypothèse sur la forme de la loi sous-jacente aux données. Les premiers tests non paramétriques de détection de changement ont été proposés par Bhattacharyya and Johnson [4] et Pettitt [40]. Ces tests statistiques sont basés sur les rangs des observations. D'autres travaux dans cette lignée ont suivi, notamment ceux de Lombard [33] et Lombard [34]. Ces contributions [4, 33, 34, 40] s'attardent à certains moments d'une distribution. Les travaux de Gombay and Horvath [20] permettent (du moins en principe) la détection de tous les types de changement dans une distribution. Des contributions plus récentes sont attribuables notamment à Zou et al. [62], Dubey and Müller [13] et Mohr and Neumeier [36].

Plusieurs auteurs ont utilisé les tests disponibles pour analyser des phénomènes physiques. Par exemple, Ali and Giaccotto [1] ont étudié les changements de moyenne et de variance dans ces séries boursières afin de mieux comprendre leur volatilité et, par conséquent, mieux évaluer les risques d'investissement. Un autre domaine où l'on s'intéresse beaucoup à la recherche de points de changement est l'hydrologie. Par exemple, Ryberg et al. [48] appliquent plusieurs méthodes de détection de points de changement afin d'identifier des modifications dans les crues annuelles.

1.2.3 Limites des approches actuelles

La motivation de ce mémoire repose en grande partie sur deux constats concernant les tests de détection de changement dans une série chronologique.

Constat #1

La plupart des travaux sur la détection de points de changement suggèrent que, si jamais une rupture s'est produite, elle est nécessairement abrupte.

Constat #2

Les tests non paramétriques sont habituellement basés sur la fonction de répartition de la loi sous-jacente aux observations.

Le Constat #1 à l'effet que les travaux sur la détection de changements assument une rupture abrupte n'est pas problématique quand cette hypothèse peut être raisonnablement soutenue. Toutefois, il y a plusieurs contextes où cette hypothèse n'est pas réaliste. On peut penser immédiatement aux données climatiques, qui sont réputées être sujettes à des changements qui s'étalent sur une longue période, parfois plusieurs années. Pour ce type de contexte, supposer une rupture abrupte peut mener à des conclusions erronées.

Le Constat #2 souligne que les méthodes sont généralement basées sur la fonction de répartition. Cependant, la fonction de répartition n'est pas la seule manière de caractériser le comportement d'une variable aléatoire. En effet, une variable aléatoire peut aussi se caractériser par sa fonction caractéristique. Cette approche dans le cas de la détection de changements a été relativement peu explorée. À ce titre, deux contributions retiennent particulièrement notre attention, à savoir Hušková and Meintanis [24] et Hušková and Meintanis [25]. Ces auteurs proposent des tests de détection basés justement sur des fonctions caractéristiques empiriques. Il est crucial de souligner ici que ces tests reposent sur l'hypothèse de changements abrupts.

1.2.4 Voici comment on répondra aux constats #1 et #2

Dans ce mémoire, des tests non paramétriques qui répondent en quelque sorte au Constat #1 et au Constat #2 sont développés. Pour répondre au Constat #1, on se basera sur un modèle à changements graduels initialement proposé par Quessy [44]. Alors que ce modèle comprend comme cas particuliers les changements abrupts, il permet la prise en compte d'un plus large éventail de ruptures possibles. Ce modèle est formellement décrit à la Section 1.3.1.

Pour répondre au Constat #2, on adopte une approche basée sur la fonction caractéristique. Cette façon de procéder permet de généraliser les tests de Hušková and Meintanis [24] au cas de ruptures graduelles. Pour ce faire, une attention particulière sera portée aux mesures de Lévy. Ces mesures permettent de définir une riche famille de statistiques de test. L'usage des mesures de Lévy en statistique est relativement récent. On peut retrouver cette idée dans Böttcher et al. [6], où un lien est établi entre les mesures de Lévy et la distance de covariance proposée dans le désormais célèbre article de Székely et al. [53].

1.3 Les « pièces maitresses » de l'article

1.3.1 Le modèle à changement graduel

Afin de s'affranchir de l'hypothèse restrictive d'un changement abrupt, Lombard [34] a proposé un modèle à changement graduel. Ce modèle suppose que la loi sous-jacente à une suite de variables aléatoires $Y_1, \dots, Y_n$ se modifie graduellement entre les temps

K_1 et K_2 . Spécifiquement, pour certaines fonctions de répartition F_1 et F_2 , on a

$$\mathbb{P}(Y_j \leq y) = \omega_j^{\mathbf{K}} F_1(y) + (1 - \omega_j^{\mathbf{K}}) F_2(y), \quad (1.1)$$

où pour $\mathbf{K} = (K_1, K_2)$, la suite de poids $\omega_1^{\mathbf{K}}, \dots, \omega_n^{\mathbf{K}}$ est définie par

$$\omega_j^{\mathbf{K}} = \begin{cases} 1, & j \leq K_1; \\ \frac{K_2 - j}{K_2 - K_1}, & K_1 < j < K_2; \\ 0, & j \geq K_2. \end{cases} \quad (1.2)$$

Sur la base de ce modèle, Lombard [34] a ensuite développé des tests de détection de ruptures basés sur les rangs des observations. Ces tests sont non paramétriques au sens où aucune hypothèse sur la loi des observations n'est émise.

1.3.2 Approche de Quessy (2019)

La détection de changement graduel dans le comportement stochastique d'une série d'observations multivariées $\mathbf{Y}_1, \dots, \mathbf{Y}_n \in \mathbb{R}^d$ est exploré dans l'article de Quessy [44]. L'idée élaborée par l'auteur s'appuie sur une mesure de distance entre les distributions $F_1, \dots, F_n$ et le modèle (2.1). Conséquemment, il s'agit de considérer le produit scalaire centré défini entre deux vecteurs $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ par

$$\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b}).$$

Si on pose $\mathbf{F} = (F_1, \dots, F_n)$ et $\boldsymbol{\omega}^{\mathbf{K}} = (\omega_1^{\mathbf{K}}, \dots, \omega_n^{\mathbf{K}})$, alors

$$L_{\mathbf{K}}(\mathbf{y}) = \frac{\langle \boldsymbol{\omega}^{\mathbf{K}}, \mathbf{F}(\mathbf{y}) \rangle}{n}. \quad (1.3)$$

Supposons que la suite des distributions sous-jacentes aux données $F_1, \dots, F_n$ suivent le modèle (2.1) pour un temps de changement $\mathbf{K}_0 \in \mathcal{C}_n = \{(c_1, c_2) \in \mathbb{N}^2 : 1 \leq c_1 <$

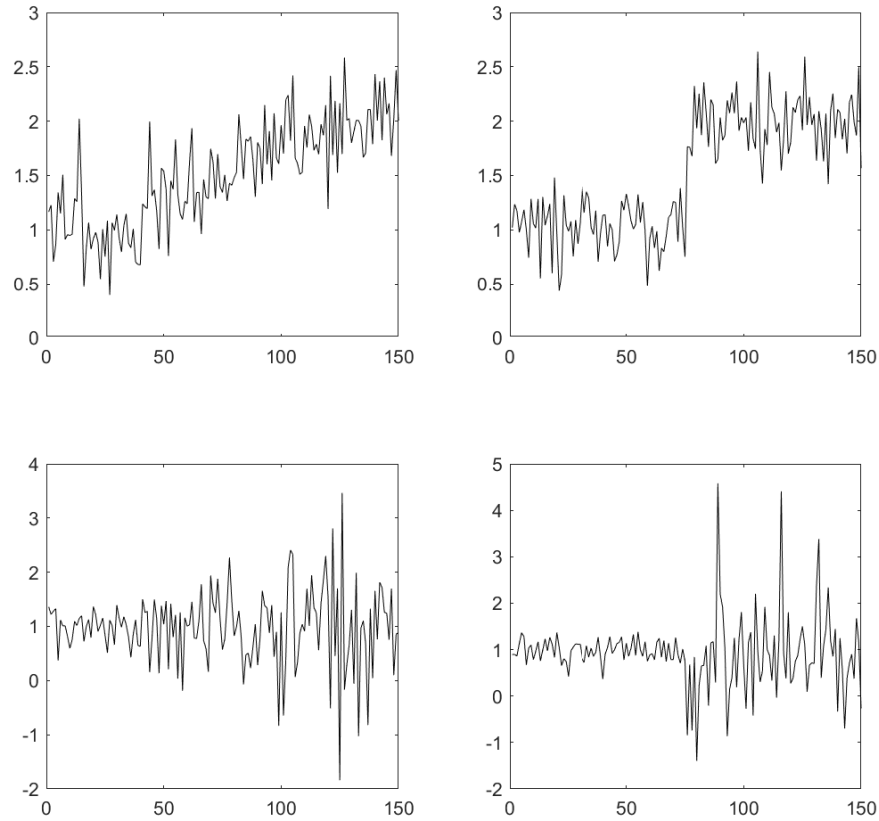


FIGURE 1.1 – Données simulées pour illustrer quelques types de changements qui peuvent survenir dans une série chronologique. Figures du haut : changements graduel (à gauche) et abrupt (à droite) dans la moyenne ; figures du bas : changements graduel (à gauche) et abrupt (à droite) dans la variance

$c_2 \leq n\}$. Il suffit d'appliquer l'expression à l'Équation (1.3) pour obtenir

$$L_{\mathbf{K}}(\mathbf{y}) = \{D_1(y) - D_2(y)\} \frac{\langle \omega^{\mathbf{K}}, \omega^{\mathbf{K}_0} \rangle}{n}, \quad (1.4)$$

où D_1 et D_2 sont respectivement les distributions avant et après le changement. L'Équation (1.4) favorise la construction d'un critère en vue de vérifier la présence d'une variation dans le comportement de D_1 et D_2 . Ainsi, sous le modèle à changement graduel, l'hypothèse nulle ne tient que si $D_1 = D_2$. Cela suggère de confronter

les hypothèses

$$\mathbb{H}_0 : \max_{\mathbf{K} \in \Delta_n} |L_{\mathbf{K}}| \equiv 0 \quad vs \quad \mathbb{H}_1 : \max_{\mathbf{K} \in \Delta_n} |L_{\mathbf{K}}| \neq 0. \quad (1.5)$$

Dans cette optique, on cherche à mesurer l'écart entre l'origine et l'Équation (1.4). Pour ce faire, l'auteur propose de se baser sur les statistiques globales

$$S_n^\zeta = \frac{1}{n^2} \sum_{\mathbf{K} \in \Delta} \eta\left(\frac{\mathbf{K}}{n}\right) \zeta(L_{\mathbf{K},n}) \quad \text{et} \quad T_n^\zeta = \max_{\mathbf{K} \in \Delta} \eta\left(\frac{\mathbf{K}}{n}\right) \zeta(L_{\mathbf{K},n}),$$

où ζ est une fonctionnelle de distance. Comme cas particuliers, on a

$$\begin{aligned} \zeta^{\text{CvM}}(g) &= \sqrt{\int_{\mathbb{R}}^d \{g(\mathbf{y})\}^2 dF_{1:n}(\mathbf{y})}, \\ \zeta^{\text{KS}}(g) &= \max_{\mathbf{y} \in V_n} |g(\mathbf{y})| \\ \text{et } \zeta^{\text{Kui}}(g) &= \left| \max_{\mathbf{y} \in V_n} g(\mathbf{y}) - \min_{\mathbf{y} \in V_n} g(\mathbf{y}) \right|, \end{aligned}$$

qui sont respectivement les fonctionnelles de Cramér–von Mises, de Kuiper et de Kolmogorov–Smirnov.

1.3.3 Utilisation de la fonction caractéristique

La fonction caractéristique d'une variable aléatoire Y est définie par

$$\psi_Y(t) = \mathbb{E}(e^{itY}), \quad t \in \mathbb{R}. \quad (1.6)$$

où $i^2 = -1$. En invoquant la formule d'Euler, c'est-à-dire $e^{ix} = \cos x + i \sin x$, on obtient la représentation équivalente $\psi_Y(t) = \mathbb{E}\{\cos(tY)\} + i \mathbb{E}\{\sin(tY)\}$. Si Y possède une densité f_Y , alors

$$\psi_Y(t) = \int_{\mathbb{R}} e^{itY} f_Y(y) dy.$$

Cette définition peut s'étendre à n'importe quel type de variable aléatoire en invoquant l'intégrale de Riemann–Stieltjes. Ainsi, si $F_Y = \mathbb{P}(Y \leq y)$ est la fonction de répartition de Y , alors

$$\psi_Y(t) = \int_{\mathbb{R}} e^{ity} dF_Y(y), \quad (1.7)$$

où l'intégrale ci-dessus est l'intégrale prise au sens de Riemann–Stieltjes.

Supposons maintenant une suite de variables aléatoires $Y_1, \dots, Y_n$ soumises au modèle à changement graduel de l'Équation (2.1), c'est-à-dire que $F_j(y) = \mathbb{P}(Y_j \leq y) = \omega_j^{\mathbf{K}} F_1(y) + (1 - \omega_j^{\mathbf{K}}) F_2(y)$. À la lumière de l'Équation (1.7), la fonction caractéristique de Y_j peut s'exprimer par

$$\begin{aligned} \psi_j(t) &= \int_{\mathbb{R}} e^{ity} dF_j(y) \\ &= \int_{\mathbb{R}} e^{ity} \{ \omega_j^{\mathbf{K}} dF_1(y) + (1 - \omega_j^{\mathbf{K}}) dF_2(y) \} \\ &= \omega_j^{\mathbf{K}} \int_{\mathbb{R}} e^{ity} dF_1(y) + (1 - \omega_j^{\mathbf{K}}) \int_{\mathbb{R}} e^{ity} dF_2(y) \\ &= \omega_j^{\mathbf{K}} \psi_1(t) + (1 - \omega_j^{\mathbf{K}}) \psi_2(t), \end{aligned}$$

où ψ_1 et ψ_2 sont les fonctions caractéristiques de F_1 et F_2 , respectivement. La suite de fonctions caractéristiques $\Psi_1, \dots, \Psi_n$ associées à une série chronologique $Y_1, \dots, Y_n$ est à la base des méthodologies statistiques qui sont développées au Chapitre 2.

1.3.4 Estimation d'une fonction caractéristique

Supposons un échantillon $Y_1, \dots, Y_n$ i.i.d. d'une loi dont la fonction de répartition est F_Y . On souhaite estimer la fonction caractéristique ψ_Y associée à F_Y . Puisque

$\psi_Y(t) = E(e^{itY})$, un estimateur naturel est donné par

$$\psi_n(t) = \frac{1}{n} \sum_{j=1}^n e^{itY_j}.$$

Définition 1.1. Soit $Z = Z_1 + i Z_2$, une variable aléatoire à valeur dans l'espace $\mathbb{C}$ des nombres complexes. Alors l'espérance de Z est $E(Z) = E(Z_1) + i E(Z_2)$ et sa variance est $\text{var}(Z) = E(|Z|^2) - |E(Z)|^2$, où $|z| = \sqrt{z_1^2 + z_2^2}$ est le complexe conjugué de $z = z_1 + i z_2 \in \mathbb{C}$.

À la lumière de la Définition 1.1, l'estimateur ψ_n est sans biais pour ψ_Y car

$$E\{\psi_n(t)\} = \frac{1}{n} \sum_{j=1}^n E(e^{itY_j}) = \frac{1}{n} \sum_{j=1}^n \psi_Y(t) = \psi_Y(t).$$

De plus, en invoquant à nouveau la Définition 1.1,

$$\begin{aligned} \text{var}\{\psi_n(t)\} &= E\left(\left|\frac{1}{n} \sum_{j=1}^n e^{itY_j}\right|^2\right) - |\psi_Y(t)|^2 \\ &= \frac{1}{n^2} E\left\{\sum_{j,j'=1}^n e^{it(Y_j - Y_{j'})}\right\} - |\psi_Y(t)|^2 \\ &= \frac{1}{n^2} \sum_{j,j'=1}^n E\{e^{it(Y_j - Y_{j'})}\} - |\psi_Y(t)|^2. \end{aligned} \quad (1.8)$$

On remarque ensuite que

$$\begin{aligned} \sum_{j,j'=1}^n E\{e^{it(Y_j - Y_{j'})}\} &= \sum_{j=1}^n E\{e^{it(Y_j - Y_j)}\} + \sum_{j \neq j'}^n E\{e^{it(Y_j - Y_{j'})}\} \\ &= n + \sum_{j \neq j'=1}^n E(e^{itY_j}) E(e^{-itY_{j'}}) \\ &= n + \sum_{j \neq j'=1}^n \psi_Y(t) \psi_Y(-t) \\ &= n + n(n-1) |\psi_Y(t)|^2, \end{aligned}$$

où la dernière ligne découle de $|\psi_Y(t)|^2 = \psi_Y(t) \psi_Y(-t)$. En injectant cette expression dans l'Équation (1.8), on obtient

$$\begin{aligned} \text{var} \{\psi_n(t)\} &= \frac{n + n(n-1) |\psi_Y(t)|^2}{n^2} - |\psi_Y(t)|^2 \\ &= \frac{1 + (n-1) |\psi_Y(t)|^2}{n} - |\psi_Y(t)|^2 \\ &= \frac{1 - |\psi_Y(t)|^2}{n}. \end{aligned}$$

1.3.5 Processus de type « α -mixing »

Un processus stochastique est considéré α -mixing si les observations sont asymptotiquement indépendantes lorsque le temps qui les sépare est grand. En d'autres mots, plus les observations sont éloignées, moins elles sont dépendantes. De ce fait, soit la séquence $(X_t)_{t \in \mathbb{Z}}$ définie sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, alors la dépendance entre X_{t_1} et X_{t_2} s'approche de zéro plus $|t_2 - t_1|$ est grand. Pour la suite, on supposera que les processus satisfont l'hypothèse de strong-mixing. Plus spécifiquement, soit $A, B \subset \mathcal{F}$, on définit les coefficients α de $(X_t)_{t \in \mathbb{Z}}$ tel que $\alpha_X(0) = 1/2$ et $\forall r \geq 1$,

$$\alpha_X(r) = \sup_{t \in \mathbb{Z}} \sup_{(A, B) \in \mathcal{F}_{-\infty}^t \times \mathcal{F}_{t+r}^\infty} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|$$

Ici, $\mathcal{F}^a$ est la σ -algèbre générée par la séquence $(X_t)_{a \leq t \leq b}$ où $a, b \in \mathbb{Z} \cup \{-\infty, \infty\}$. Ainsi, le processus est α -mixing si $\alpha(r) \rightarrow 0$ lorsque $r \rightarrow \infty$. Cette condition permet de couvrir un large éventail de modèles, tels que les processus autorégressifs (AR) et les modèles autorégressifs à moyenne mobile (ARMA).

1.3.6 Méthode Monte Carlo pour approximer la puissance d'un test

La découverte de la version moderne de la méthode Monte Carlo est attribuée à Stanislaw Ulam lors de sa collaboration avec le célèbre John Von Neumann au laboratoire national de Los Alamos dans les années 1940. Avec les avancées technologiques des ordinateurs, elle intervient maintenant de façon systématique dans le calcul de puissance pour des tests statistiques. Cette méthode est utilisée abondamment à la Section 2.5 de l'article présenté intégralement au Chapitre 2.

Tout d'abord, la puissance $\mathcal{P}$ d'un test statistique est définie comme la probabilité de rejeter l'hypothèse nulle $\mathbb{H}_0$ lorsque les données sont supposées provenir d'une loi sous une hypothèse alternative $\mathbb{H}_1$. Ainsi,

$$\mathcal{P} = \mathbb{P}(\text{Rejeter } \mathbb{H}_0 \mid \mathbb{H}_1).$$

Sauf dans de rares cas, comme par exemple pour des tests basés sur des statistiques dont la loi est Normale ou Student, le calcul de $\mathcal{P}$ est difficile, voire impossible. Une manière simple et efficace pour approximer la valeur de $\mathcal{P}$ consiste à procéder par simulations Monte Carlo. Supposons donc un test basé sur une statistique S_n , où n est la taille d'échantillon, dont la région de rejet au seuil α est R_α . L'approximation de $\mathcal{P}$ procède de la manière suivante :

- (1) On simule M échantillons indépendants de taille n provenant de la loi sous l'hypothèse alternative $\mathbb{H}_1$;
- (2) On calcule les statistiques de test pour chacun des échantillons de l'étape (1), ce qui fournit $S_n^{(1)}, \dots, S_n^{(M)}$.

(3) On approxime la puissance du test avec

$$\hat{\mathcal{P}} = \frac{1}{M} \sum_{m=1}^M \mathbb{I} \{S_n^{(m)} \in R_\alpha\}.$$

Plus la valeur de M est élevée, plus l'approximation de la puissance sera précise.

1.4 Un condensé sur les processus stochastiques

L'objectif de cette section est d'introduire la notion de processus séquentiel et de mesure de Lévy. Pour ce faire, on débute en abordant les processus gaussiens, en particulier le mouvement brownien et le pont brownien, Quelques propriétés de ces processus seront également explorées.

D'abord, un processus stochastique est une collection de variables aléatoires définie sur un espace de probabilité et indexée par un certain ensemble T . Formellement, un processus stochastique défini sur $(\Omega, \mathcal{F}, \mathbb{P})$ et indexé par un ensemble T est tel que $X_t : \Omega \rightarrow \mathbb{R}$ pour chaque $t \in T$. De façon générale, un processus dépend des paramètres ω et t . En fixant, $\omega \in \Omega$, on obtient la trajectoire du processus donnée par la fonction $t \in T \rightarrow X_t(\omega)$. De même, en fixant $t \in T$ on trouve $\omega \in \Omega \rightarrow X_t(\omega)$, une variable aléatoire sur l'espace de probabilité. Finalement, quelques propriétés de continuité qui seront exploitées ultérieurement sont données.

Définition 1.2. Soit l'espace métrique $\mathcal{C} = (T, d)$, un processus $(X_t)_{t \in T}$ est continu en trajectoires si pour tout $\omega \in \Omega$, la fonction $t \in T \rightarrow X_t(\omega)$ est continue sur $\mathcal{C}$.

Définition 1.3. Un processus $(X_t)_{t \in \mathcal{T}}$ est continue en probabilité si $X_s \xrightarrow{\mathbb{P}} X_t$ lorsque $s \rightarrow t$.

1.4.1 Processus gaussiens

Définition 1.4. *Un processus stochastique $(X_t)_{t \in T}$ sur $(\Omega, \mathcal{F}, \mathbb{P})$ est dit «gaussien» lorsque toute combinaison finie de $(X_{t_1}, \dots, X_{t_n})$ suit une distribution normale.*

La définition 1.4 confirme l'intuition selon laquelle un processus gaussien est une collection de variables aléatoires normales définies sur un même espace de probabiliste. En fait, il est entièrement défini par sa fonction de covariance et son espérance. De plus, l'utilisation des propriétés de la loi normale implique qu'un processus stochastique $(X_t)_{t \in T}$ est gaussien si et seulement si pour tout $t_1, \dots, t_K \in T$, la combinaison linéaire

$$Y = a_1 X_{t_1} + \dots + a_n X_{t_K}$$

est distribuée selon une loi normale unidimensionnelle. En d'autres mots, un processus stochastique est gaussien si et seulement si toutes combinaisons linéaires finies de ses lois marginales sont aussi de loi gaussienne.

En s'inspirant de la Définition 1.10 de Breton [7], voici quelques propriétés fondamentales de processus stochastiques plus complexes.

Définition 1.5. *Soit un processus stochastique $(X_t)_{t \in T}$. Alors*

- (i) *$(X_t)_{t \in T}$ est strictement stationnaire si $X_{t+s} \stackrel{d}{=} X_t$ pour tout $s, t \geq 0$;*
- (ii) *$(X_t)_{t \in T}$ est à accroissements stationnaires si la loi des accroissements $X_{t+s} - X_t$ ne dépend pas de $s > 0$ au sens où $X_{t+s} - X_t \stackrel{d}{=} X_s$;*
- (iii) *$(X_t)_{t \in T}$ est à accroissements indépendants si pour tout $0 < t_1 < t_2 < \dots < t_K$, les variables aléatoires $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_K} - X_{t_{K-1}}$ sont indépendantes.*

Le processus gaussien, notamment ses propriétés d'accroissements stationnaires et indépendants, sont à la base de la définition du mouvement brownien. Ce processus fait l'objet de la prochaine sous-section.

1.4.2 Mouvement brownien

Le mouvement brownien fut décrit pour la première fois par l'écossais Robert Brown, en 1827, lorsqu'il observa des particules de pollen de la plante de *Clarkia pulchella* en suspension dans l'eau. Le processus fut modélisé par Louis Bachelier en 1900 dans le cadre de sa thèse doctorale sous la supervision de Henri Poincaré. Cette théorie sera explorée par le célèbre physicien Albert Einstein [15] où il réalise la première analyse statistique de la physique du mouvement. Par la suite, Norbert Wiener [59] définit de façon rigoureuse le mouvement brownien et sa théorie sera enrichie par les travaux de Kolmogorov sur la théorie des probabilités dans les années 1930. Cela permettra de formaliser la notion de processus stochastique qui sera ensuite utilisée dans de multiples disciplines, comme la finance, la biologie et même la physique quantique.

Définition 1.6. *Un mouvement brownien (processus de Wiener) $(W_t)_{t \geq 0}$ est un processus gaussien centré à trajectoires continues dont la covariance est donnée par*

$$\text{cov}(W_s, W_t) = \min(s, t).$$

Trois propriétés du mouvement brownien se déduisent de la Définition 1.6.

Lemme 1.1. *Soit $(W_t)_{t \geq 0}$, un mouvement brownien. Alors*

- (i) $W_0 = 0$;
- (ii) *Les incréments de W_t sont stationnaires, car $W_s - W_t \stackrel{d}{=} W_{s-t} \sim N(0, s - t)$ pour tout $t < s$;*
- (iii) *Pour $t_1 < \dots < t_K$ arbitraires, les incréments $W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_K} - W_{t_{K-1}}$ sont indépendants.*

Le Lemme 1.1 offre les outils nécessaires pour montrer que $\text{cov}(W_s, W_t) = \min(s, t)$.

Sans perte de généralité, on suppose que $s \leq t$. Par la définition de la covariance,

$$\begin{aligned}
 \text{cov}(W_s, W_t) &= E(W_s W_t) - E(W_s) E(W_t) \\
 &= E\{W_s(W_s + W_t - W_s)\} \\
 &= E(W_s^2) + E\{W_s(W_t - W_s)\} \\
 &= E(W_s^2) \\
 &= s.
 \end{aligned}$$

On montre de la même manière que $\text{cov}(W_s, W_t) = t$ lorsque $t \leq s$ respectivement.

On a donc $\text{cov}(W_s, W_t) = \min(s, t)$.

On conclut cette section en introduisant le pont brownien.

Définition 1.7. Soit le mouvement brownien $(W_t)_{t \geq 0}$. Alors le pont brownien est le processus $(B_t)_{t \in [0,1]}$ tel que

$$B_t = W_t - t W_1.$$

Par conséquent, $B_0 = 0$ et $B_1 = 0$. Le processus B_t est donc “attaché” aux extrémités de l’intervalle $[0, 1]$, d’où le vocable de «pont brownien».

La fonction de covariance de B_t est $\text{cov}(B_s, B_t) = \min(s, t) - st$. En effet, par les propriétés du Lemme 1.1 concernant le mouvement brownien, on obtient

$$\begin{aligned}
 \text{cov}(B_s, B_t) &= \text{cov}(W_s - sW_1, W_t - tW_1) \\
 &= E\{(W_s - sW_1)(W_t - tW_1)\} - E(W_s - sW_1) E(W_t - tW_1).
 \end{aligned}$$

Le terme de droite est égal à zéro, puisque W est un mouvement brownien, donc

$E(W_s - sW_1) = 0$. Ainsi,

$$\begin{aligned} \text{cov}(B_s, B_t) &= E(W_s W_t) - t E(W_s W_1) - s E(W_1 W_t) + st E(W_1 W_1) \\ &= \min(s, t) - ts - st + st \\ &= \min(s, t) - st, \end{aligned}$$

où la dernière égalité découle du fait que $\text{cov}(W_s, W_t) = E(W_s W_t) = \min(s, t)$.

1.4.3 Processus de Lévy

Afin d'introduire la notion de mesure de Lévy, on considère une classe de processus stochastiques plus générale : les processus de Lévy.

Définition 1.8. $(X_t)_{t \in T}$ est un processus de Lévy s'il possède des accroissements indépendants et stationnaires.

Cette classe est caractérisée par les processus aux trajectoires $t \in T \rightarrow X_t$ continues à droite presque partout ayant une limite à gauche (Càdlàg). Cela rend possible l'étude d'un large éventail de modèles possédant des discontinuités appelées « sauts ». Par exemple, une telle particularité peut se retrouver dans les processus de Poisson. En revanche, le mouvement brownien se distingue comme étant le seul processus de Lévy à trajectoires continues. On décrit ci-dessous une propriété clé de ces processus.

Définition 1.9. Si une variable aléatoire X possède une distribution infiniment divisible, alors pour tout entier $n \geq 1$, il existe des variables aléatoires indépendantes $X_1, \dots, X_n$ identiquement distribuées telles que $X \stackrel{d}{=} X_1 + \dots + X_n$.

Si $(X_t)_{t \in T}$ est un processus de Lévy, alors $\forall t \in T$, X_t dispose d'une distribution infiniment divisible. La loi Gamma, la loi Normale ainsi que la loi de Poisson sont des distributions connues possédant cette particularité. Effectivement, soit $X \sim \mathcal{N}(\mu, \sigma^2)$, alors

la loi de X est la même que la somme de n variables aléatoires de loi $\mathcal{N}(\mu/n, \sigma^2/n)$. On peut ainsi conclure que la distribution gaussienne est pourvue de la propriété de divisibilité infinie.

Théorème 1.1 (Décomposition de Lévy-Itô). *Soit X_t , un processus de Lévy, alors on peut l'exprimer par*

$$X_t = mt + \sigma W_t + \mathcal{J}_t.$$

Avant d'introduire le Théorème de Lévy-Khintchine, il est important de noter que la décomposition de Lévy-Itô décrite au Théorème 1.1 indique qu'un processus de Lévy est représentable de façon unique par trois composantes, à savoir

- (1) Le paramètre de dérive m qui représente le changement moyen ou attendu du processus ;
- (2) Un mouvement brownien W_t de variance σ^2 ;
- (3) Un processus de sauts $\mathcal{J}_t$, prenant en compte les discontinuités, définit par une mesure de Lévy ν .

Il est à noter que le processus $\mathcal{J}_t$ peut s'écrire comme la somme de deux processus de Poisson. Soit donc P_t , un processus de Poisson composé auquel on associe la mesure de Lévy ν_P et qui satisfait les conditions

$$\nu_P(|x| \leq 1) = 0 \quad \text{et} \quad \nu_P(|x| > 1) < \infty.$$

Soit aussi un processus de Poisson généralisé compensé G_t tel que

$$\int_{-\infty}^{\infty} x^2 \nu_G(dx) < \infty, \quad \nu_G(\{0\}) = 0 \quad \text{et} \quad \nu_G(|x| > 1) = 0.$$

Alors $\mathcal{J}_t$ admet la représentation $\mathcal{J}_t = P_t + G_t$ définie par la mesure de Lévy $\nu = \nu_P + \nu_G$. Comme la mesure ν décrit l'intensité des sauts d'amplitude $|x|$, on peut interpréter P_t et G_t comme les processus modélisant les sauts de valeur $|x| > 1$ et $|x| \leq 1$ respectivement. En résumé, cet important résultat affirme que X_t est

entièrement déterminé par son triplet (m, σ^2, ν) . La preuve de son unicité n'est pas présentée ici, car elle dépasse le cadre de ce travail. Ainsi, on remarque qu'en posant $m = 0$ et $\nu = 0$, on retrouve le mouvement brownien.

On clôt cette section avec un résultat tissant le lien avec les mesures de Lévy.

Théorème 1.2 (Lévy–Khintchine). *Soit X , une variable aléatoire infiniment divisible de fonction caractéristique $\psi(t)$, alors son exposant caractéristique ϕ est donné par*

$$\phi_X(t) = imt - \frac{\sigma^2}{2}t^2 + \int_{\mathbb{R}} \{e^{itx} - 1 - itx \mathbb{I}(|x| < 1)\} \nu(dx)$$

où $m \in \mathbb{R}, \sigma \geq 0$ sont la dérive et la variance du mouvement brownien associées au processus, alors que ν est une mesure de Lévy sur $\mathbb{R}$ satisfaisant les conditions

$$\int_{\mathbb{R}} \min(|t|^2, 1) \nu(dt) < \infty \quad \text{et} \quad \nu(\{0\}) = 0.$$

1.4.4 Fonction de répartition empirique

Soit une série d'observations $X_1, \dots, X_n$ indépendantes et identiquement distribuées caractérisée par une certaine fonction de répartition $F_X(x) = \mathbb{P}(X \leq x)$. En général, il est bénéfique d'estimer F_X et une telle estimation est donnée par la fonction de répartition empirique, à savoir

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(X_i \leq x).$$

Lemme 1.2. *Soit une valeur fixe $x_0 \in \mathbb{R}$.*

1. $\mathbb{E}\{F_n(x_0)\} = F_X(x_0)$, donc F_n est un estimateur sans biais de F_X ;
2. $\text{var}\{F_n(x_0)\} = \sigma_{x_0}^2/n$, où $\sigma_{x_0}^2 = F_X(x_0)\{1 - F_X(x_0)\}$.

Il suffit d'invoquer les propriétés de base de l'espérance pour écrire

$$\begin{aligned}
 \mathbb{E}\{F_n(x_0)\} &= \mathbb{E}\left\{\frac{1}{n} \sum_{i=1}^n \mathbb{I}(X_i \leq x_0)\right\} \\
 &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}\{\mathbb{I}(X_i \leq x_0)\} \\
 &= \frac{1}{n} \sum_{i=1}^n \mathbb{P}(X_i \leq x_0) \\
 &= F_X(x_0).
 \end{aligned}$$

On constate que $F_n(x_0)$ est un estimateur sans biais de $F_X(x_0)$. Dans un même ordre d'idées, la variance de la fonction de répartition est

$$\begin{aligned}
 \text{var}\{F_n(x_0)\} &= \text{var}\left\{\frac{1}{n} \sum_{i=1}^n \mathbb{I}(X_i \leq x_0)\right\} \\
 &= \frac{1}{n^2} \sum_{i=1}^n \text{var}\{\mathbb{I}(X_i \leq x_0)\} \\
 &= \frac{\text{var}\{\mathbb{I}(X_1 \leq x_0)\}}{n} \\
 &= \frac{\mathbb{E}\{\mathbb{I}(X_1 \leq x_0)\} - \mathbb{E}^2\{\mathbb{I}(X_1 \leq x_0)\}}{n} \\
 &= \frac{F_X(x_0) - \{F_X(x_0)\}^2}{n} \\
 &= \frac{F_X(x_0) \{1 - F_X(x_0)\}}{n}.
 \end{aligned}$$

Plus généralement, on peut obtenir la covariance de F_n .

Lemme 1.3. *Soit $X_1, \dots, X_n$ i.i.d F_X . Alors pour $x_1, x_2 \in \mathbb{R}$ fixés,*

$$\text{cov}\{F_n(x_1), F_n(x_2)\} = \frac{F_X(x_1 \wedge x_2) - F_X(x_1) F_X(x_2)}{n}.$$

Pour démontrer le lemme précédent, il s'agit d'abord de noter que

$$\begin{aligned} \text{cov} \{F_n(x_1), F_n(x_2)\} &= \text{cov} \left\{ \frac{1}{n} \sum_{i=1}^n \mathbb{I}(X_i \leq x_1), \frac{1}{n} \sum_{j=1}^n \mathbb{I}(X_j \leq x_2) \right\} \\ &= \frac{1}{n^2} \sum_{i,j=1}^n \text{cov} \{ \mathbb{I}(X_i \leq x_1), \mathbb{I}(X_j \leq x_2) \}. \end{aligned}$$

Puisque les variables $X_1, \dots, X_n$ sont i.i.d.,

$$\begin{aligned} \text{cov} \{F_n(x_1), F_n(x_2)\} &= \frac{1}{n^2} \sum_{i=1}^n \text{cov} \{ \mathbb{I}(X_i \leq x_1), \mathbb{I}(X_i \leq x_2) \} \\ &= \frac{\text{cov} \{ \mathbb{I}(X_1 \leq x_1), \mathbb{I}(X_1 \leq x_2) \}}{n}. \end{aligned}$$

Le résultat annoncé s'obtient en calculant

$$\begin{aligned} \text{cov} \{ \mathbb{I}(X_1 \leq x_1), \mathbb{I}(X_1 \leq x_2) \} &= \mathbb{E} \{ \mathbb{I}(X_1 \leq x_1) \mathbb{I}(X_1 \leq x_2) \} \\ &\quad - \mathbb{E} \{ \mathbb{I}(X_1 \leq x_1) \} \mathbb{E} \{ \mathbb{I}(X_1 \leq x_2) \} \\ &= \mathbb{E} \{ \mathbb{I}(X_1 \leq x_1 \wedge x_2) \} - F_X(x_1) F_X(x_2) \\ &= F_X(x_1 \wedge x_2) - F_X(x_1) F_X(x_2). \end{aligned}$$

1.4.5 Processus empirique

Dans l'optique d'introduire les processus empiriques, il est nécessaire d'invoquer la loi des grands nombres stipulant que $F_n(x_0) \rightarrow F_X(x_0)$. Ce résultat complète les propriétés de la fonction de répartition empirique et assure sa convergence ponctuelle lorsque $n \rightarrow \infty$. Toutefois, il est pertinent d'étudier le comportement de F_n sans fixer une valeur de x parmi les nombres réels. Spécifiquement, le Théorème de Glivenko-Cantelli assure que F_n est un bon estimateur de F_X pour tout $x \in \mathbb{R}$. Au préalable,

on définit la distance uniforme par

$$D_n = \sup_{x \in \mathcal{X}} |F_n(x) - F_X(x)|.$$

Cette distance fut introduite par Kolmogorov [28] dans un article célèbre qui constitue la base sur laquelle le théorème suivant s'appuie.

Théorème 1.3 (Glivenko–Cantelli). *Soit $X_1, \dots, X_n$ une série de variables aléatoires indépendantes et identiquement distribuées ayant une distribution F_X sur les réels. Alors lorsque $n \rightarrow \infty$,*

$$D_n = \sup_{x \in \mathcal{X}} |F_n(x) - F_X(x)| \rightarrow 0.$$

Ce résultat établit la convergence uniforme presque sûre de F_n vers F_X . On termine en montrant le lien entre les processus empiriques et le pont brownien. Il s'agit de faire appel au Théorème de la limite centrale pour définir le processus empirique par

$$\mathbb{F}_n(x) = \sqrt{n} \{F_n(x) - F_X(x)\} \sim \mathcal{N}(0, \sigma_x^2). \quad (1.9)$$

Théorème 1.4 (Donsker). *Le processus $\mathbb{F}_n$ converge en distribution vers un processus gaussien $\mathbb{G}$ de moyenne 0 et de covariance donnée par*

$$\text{cov}\{\mathbb{F}(s), \mathbb{F}(t)\} = F_X\{\min(s, t)\} - F_X(s)F_X(t).$$

dans l'espace des fonctions continues à droite ayant une limite à gauche (Càdlàg).

Advenant le cas où $X_1, \dots, X_n$ sont i.i.d. de loi Uniforme sur $(0, 1)$, alors le Théorème de Donsker établit la convergence du processus empirique $\mathbb{F}_n$ vers un pont brownien $\mathbb{B}$. En effet, puisque $F_X(x) = x$, on trouve immédiatement

$$\mathbb{E}\{\mathbb{F}(t)\} = 0 \quad \text{et} \quad \text{cov}\{\mathbb{F}(s), \mathbb{F}(t)\} = \min(s, t) - st.$$

L'utilisation directe de l'Équation (1.9) amène $\mathbb{B}_n = \sqrt{n}\{F_n(s) - s\}$. Une simulation

de ce processus est illustrée à la Figure 1.2.

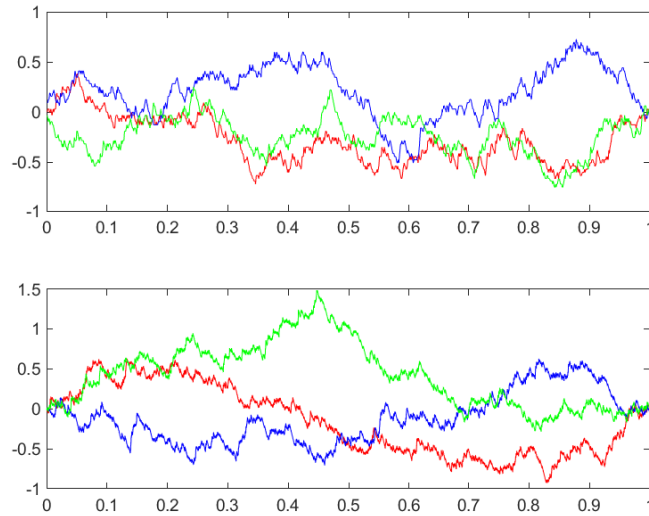


FIGURE 1.2 – Illustration du processus $\mathbb{B}_n$ lorsque $n = 1000$ (en haut) et $n = 10000$ (en bas).

1.4.6 Processus séquentiel

On termine cette courte mise en contexte des processus stochastique avec une notion importante retrouvée dans l'article présenté à la section 2. Soit une série d'observations univariées $X_1, \dots, X_n$ défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. Le processus séquentiel est fréquemment utilisé avec pour objectif de mesurer la distance entre la fonction de répartition $F_{1:K}$, avant le changement, et $F_{K+1:n}$, après le changement. Cela permet notamment de vérifier si une rupture s'est produite dans une série d'observations. En appliquant cette définition dans un cas à changement graduel $1 \leq K_1 \leq K_2 \leq n$, on obtient

$$F_{K_1:K_2}(x) = \frac{1}{K_2 - K_1 + 1} \sum_{i=K_1}^{K_2} \mathbb{I}(X_i \leq x). \quad (1.10)$$

À la lumière de la Définition 1.9, on peut définir rigoureusement le processus séquentiel. Spécifiquement, pour $(s, x) \in [0, 1] \times \mathbb{R}$ alors

$$\mathbb{F}_n(s, x) = \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{n^{3/2}} \{F_{1:\lfloor ns \rfloor}(x) - F_{\lfloor ns \rfloor+1:n}(x)\} \quad (1.11)$$

Lorsque $X_1, \dots, X_n$ sont i.i.d, un important théorème de [41] assure la convergence faible de $\mathbb{F}_n$ vers un processus de Kiefer. Une extension de ce théorème permet d'appliquer ce résultat aux processus α -mélangeants.

1.5 Un condensé sur les méthodes de rééchantillonnage

Dans cette section, plusieurs méthodes de rééchantillonnages sont présentées. Celles-ci consistent à réutiliser l'information contenue dans un échantillon de manière astucieuse. L'objectif est souvent d'estimer un paramètre, une statistique et même d'étudier les propriétés d'un test d'hypothèse. Pour ce faire, quelques approches comme le *jackknife*, le *bootstrap* et la méthode du multiplicateur sont introduites.

1.5.1 La méthode du *jackknife*

La méthode dite du jackknife, introduite par Quenouille [42], est la première technique de rééchantillonnage qui a été mise en avant. Dans un article subséquent, [43] raffine cette approche avec comme objectif de corriger le biais d'un estimateur. Tukey [55] investigue la possibilité d'utiliser le *jackknife* afin de construire des intervalles de confiance et apporte une contribution aux travaux précédents. L'idée principale est d'analyser le comportement d'une statistique lorsque l'on retire une donnée. Spécifiquement, soit une série d'observations $\mathbf{X} = (X_1, \dots, X_n)$, on définit l'échantillon

jackknife « retire-1 » de taille $n - 1$ comme suit

$$\mathbf{X}_{(i)} = (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n).$$

Conséquemment, pour $\mathcal{T}_n = \mathcal{T}(X_1, \dots, X_n)$ on considère $\mathcal{T}_{(i)} = \mathcal{T}(\mathbf{X}_{(i)})$ comme la statistique $\mathcal{T}_n$ calculée en retirant l'observation X_i . Par exemple, si l'on considère la moyenne empirique $\bar{X}_n$, on peut écrire

$$\mathcal{T}_{(i)} = \bar{X}_{(i)} = \frac{1}{n-1} \sum_{j=1, j \neq i}^n X_j.$$

Avant d'énoncer le prochain résultat, soit la moyenne des répliques *jackknife* de $\mathcal{T}_n$ définie par

$$\bar{\mathcal{T}}_{(\cdot)} = \frac{1}{n} \sum_{j=1}^n \mathcal{T}_{(j)}.$$

Lemme 1.4. *Si $\mathcal{T}_n$ est la moyenne empirique, alors*

$$\mathcal{T}_{(i)} - \bar{\mathcal{T}}_{(\cdot)} = \frac{\bar{X}_n - X_i}{n-1}.$$

Afin d'alléger les calculs, on pose $\bar{X}_n = \bar{X}$

$$\begin{aligned} \mathcal{T}_{(i)} - \bar{\mathcal{T}}_{(\cdot)} &= \frac{1}{n-1} \left\{ \left(\sum_{j=1}^n X_j \right) - X_i \right\} - \frac{1}{n} \sum_{i=1}^n \bar{X}_{(i)} \\ &= \frac{n\bar{X} - X_i}{n-1} - \frac{1}{n} \sum_{i=1}^n \frac{1}{n-1} \left\{ \left(\sum_{j=1}^n X_j \right) - X_i \right\} \\ &= \frac{n\bar{X} - X_i}{n-1} - \frac{1}{n(n-1)} \sum_{i=1}^n (n\bar{X} - X_i) \\ &= \frac{n\bar{X} - X_i - (n\bar{X} - \bar{X})}{n-1} \\ &= \frac{\bar{X}_n - X_i}{n-1} \end{aligned}$$

Lemme 1.5. *La variance jackknife $\hat{\sigma}_{\mathcal{T}}^2$ est un estimateur sans biais de $\text{var}(\bar{X}_n) = \sigma^2/n$ et son expression est donnée par*

$$\hat{\sigma}_{\mathcal{T}}^2 = \frac{n-1}{n} \sum_{i=1}^n (\mathcal{T}_{(i)} - \bar{\mathcal{T}}_{(\cdot)})^2.$$

Il s'agit d'invoquer le résultat du Lemme 1.4 et d'exploiter le fait que S_n^2 est un estimateur sans biais de σ^2 pour trouver

$$\mathbb{E}(\hat{\sigma}_{\mathcal{T}}^2) = \mathbb{E}\left\{\sum_{i=1}^n \frac{1}{n(n-1)} (X_i - \bar{X}_n)^2\right\} = \frac{\sigma^2}{n}.$$

1.5.2 La technique du *bootstrap*

La méthode du *bootstrap*, telle que suggérée par Efron [14], est une technique de ré-échantillonnage devenue populaire avec les avancées technologiques des ordinateurs. Cette approche est non paramétrique dans le sens où l'on n'émet aucune hypothèse quant à la loi sous-jacente aux observations. Ainsi, l'objectif est de maximiser l'utilisation de l'information retrouvée dans l'échantillon. Pour ce faire, on effectue un tirage sans remise dans la série d'observations originale afin de construire des ensembles appelés échantillons *bootstrap*. En répétant ce procédé un grand nombre de fois, il devient possible d'estimer des paramètres et de calculer des p-valeurs, ce qui permet d'établir un critère de rejet pour les tests d'hypothèses.

Plus formellement, soit l'événement A : *Piger X_i dans l'échantillon $\mathbf{X}$* . Alors

$$\mathbb{P}(A) = \frac{1}{n}, \quad \forall i \in \{1, \dots, n\}.$$

On définit alors l'échantillon *bootstrap* $\mathbf{X}^B = X_1^*, \dots, X_n^*$, où les X_i^* sont les observations originales pigées sans remise avec une probabilité $\mathbb{P}(A) = 1/n$. Ainsi, il s'agit de

calculer la valeur du paramètre pour chaque échantillon $\mathbf{X}^B$. Préalablement, on note

$$\bar{\theta}^B = \frac{1}{n} \sum_{i=1}^n \theta_i^B. \quad (1.12)$$

comme la moyenne du paramètre θ^B calculé à partir des échantillons *bootstrap*. De cette manière, on estime la variance de l'estimateur en faisant appel à l'Équation (1.12)

$$\hat{\sigma}_\theta^2 = \frac{1}{M} \sum_{i=1}^M (\theta_i^B - \bar{\theta}^B)^2, \quad (1.13)$$

où M est le nombre d'échantillons *bootstrap*. En pratique, ce nombre doit être élevé afin d'obtenir une bonne estimation. Dans le cas d'un calcul de p-valeur, la méthode se résume en un algorithme facile à implémenter dans la mesure où l'on recherche une valeur de rejet de $\mathbb{H}_0$. Donc, soit S_n et $\hat{S}_n^B$,

1. Calculer S_n à partir de $X_1, \dots, X_n$;
2. Piger n fois avec remise dans $\mathbf{X} = X_1, \dots, X_n$ en accordant une probabilité de $\frac{1}{n}$ à chaque observation pour obtenir un échantillon *bootstrap* $\mathbf{X}^B = (X_1^*, \dots, X_n^*)$;
3. Calculer $\hat{S}_n^B$ à partir de $\mathbf{X}^B$;
4. Répéter M fois les étapes 2 et 3 ;
5. On obtient alors les répliques $\hat{S}_n^{B,(1)}, \dots, \hat{S}_n^{B,(M)}$ ce qui permet d'évaluer

$$\mathcal{P} = \frac{1}{n} \sum_{i=1}^M \mathbb{I} \left\{ \hat{S}_n^{B,(i)} > S_n \right\}.$$

On rejette l'hypothèse $\mathbb{H}_0$ si $\mathcal{P} < \alpha$.

1.5.3 La méthode du multiplicateur

La méthode du multiplicateur est une autre technique de rééchantillonnage non paramétrique. L'idée est d'utiliser des variables « multiplicateurs » de manière à attribuer

un poids aux données originales. Cela rend possible la création de plusieurs répliques d'échantillons se comportant de façon analogue aux données de départ. En ce sens, l'objectif sera d'effectuer un calcul de p-valeur dans le cadre de tests d'hypothèses. Soit $\xi_1, \dots, \xi_n$ i.i.d, des variables indépendantes de $X_1, \dots, X_n$, telles que

$$\mathbb{E}(\xi_i) = 0 \quad \text{et} \quad \text{var}(\xi_i) = 1.$$

Comme le Théorème de Donsker assure la convergence de $\mathbb{F}_n$ vers un pont brownien, on définit le processus multiplicateur par

$$\mathbb{M}_n(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_i \{ \mathbb{I}(X_i \leq x) - F_X(x) \}. \quad (1.14)$$

On retrouve l'expression du processus empirique à l'intérieur de l'Équation (1.14). Conséquemment, le Théorème de la limite central des multiplicateurs présenté dans les travaux de Kosorok [29] affirme que $\mathbb{M}_n$ converge faiblement vers un processus gaussien centré $\mathbb{G}(x)$. Sa structure de covariance est donnée par

$$\text{cov}\{\mathbb{G}(x), \mathbb{G}(x')\} = F_X\{\min(x, x')\} - F_X(x)F_X(x').$$

On peut vérifier que $\mathbb{G}$ est en fait un pont brownien $\mathbb{B}\{F_X(x)\}$. Ce résultat est d'une grande importance, puisqu'il assure que $\mathbb{M}_n$ se comporte de façon analogue à $\mathbb{F}_n$. Donc, il sera possible de reproduire son comportement à l'aide de simulations. Soit $\xi^{(1)}, \dots, \xi^{(m)}$, des échantillons de variables multiplicateurs générés de façon indépendante alors, la version multiplicateur du processus empirique $\mathbb{F}_n$ est donnée par

$$\mathbb{M}_n^{(h)} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_i^{(h)} \{ \mathbb{I}(X_i \leq x) - F_X(x) \} \sim \mathbb{B}\{F_X(x)\}.$$

On conclut en exprimant les étapes d'un algorithme facile à mettre en œuvre ayant une méthodologie similaire à celui présenté dans la Section 1.5.2. Soit une statistique S_n , ainsi que sa version multiplicateur $\widehat{S}_n^{(h)}$,

- (1) Calculer S_n à partir des données $X_1, \dots, X_n$;

- (2) Générer un échantillon $\xi_1, \dots, \xi_n$ indépendant de l'échantillon original ;
- (3) Calculer $\widehat{S}_n^{(h)}$ à partir de $\xi_1, \dots, \xi_n$;
- (4) Répéter les étapes 2 et 3 un grand nombre de fois M ;
- (5) La p-valeur est alors estimée par

$$\mathcal{P} = \frac{1}{n} \sum_{h=1}^M \mathbb{I} \left\{ \widehat{S}_n^{(h)} > S_n \right\}.$$

On rejette alors l'hypothèse $\mathbb{H}_0$ lorsque $\mathcal{P} < \alpha$.

Chapitre 2

Article

Mots-clés : Fonction caractéristique, modèle à changement graduel, multiplicateur sériel, alpha-mixing, warp-speed bootstrap

Résumé : Cet article propose des tests destinés à la détection de points de changement graduels dans des séries temporelles univariées sous une condition générale de « α -mixing ». Les statistiques de test proposées reposent sur la fonction caractéristique, avec des fonctions de poids appartenant à une classe générale de mesures de Lévy. Les distributions asymptotiques des statistiques de test sous les hypothèses nulle et alternative sont établies à l'aide des résultats de [44]. De plus, la validité d'une procédure de « bootstrap » par la méthode du multiplicateur est démontrée pour le calcul des p-valeurs. Des expériences de Monte Carlo sont menées, intégrant une méthode de Warp-speed bootstrap afin d'accélérer considérablement le temps de calcul. Les résultats indiquent que les tests proposés conservent correctement leur niveau nominal et affichent une bonne puissance contre un large éventail d'alternatives.

2.1 Introduction

Let $Y_1, \dots, Y_n$ be realizations of a univariate process $(Y_t)_{t \in \mathbb{Z}}$ taking values in $\mathcal{Y} \subseteq \mathbb{R}$. The so-called change-point problem consists in checking whether or not the stochastic behavior of $(Y_t)_{t \in \mathbb{Z}}$ changed at some unknown time $K \in \{1, \dots, n-1\}$. Specifically, the interest is to test the null hypothesis $\mathbb{H}_0 : \mathbb{P}(Y_1 \leq y) = \dots = \mathbb{P}(Y_n \leq y)$ for all $y \in \mathcal{Y}$ against the alternative hypothesis $\mathbb{H}_1$ that there exists $K \in \{1, \dots, n-1\}$ such that for all y in $\mathcal{Y}' \subset \mathcal{Y}$ of non-null Lebesgue measure,

$$\mathbb{P}(Y_1 \leq y) = \dots = \mathbb{P}(Y_K \leq y) \neq \mathbb{P}(Y_{K+1} \leq y) = \dots = \mathbb{P}(Y_n \leq y).$$

The first contributions on change-point analysis were concerned with the detection of a change in a given parameter, often the mean or the variance, at an unknown time point. Such a procedure can be traced back to [38], where a test for a change in a parameter is proposed; see also [11] in the case of normally distributed observations. Nonparametric approaches have also been central in the development of tests of change-point detection. For example, rank tests for a change in location were considered by [4], [40] and [49], whereas the procedure in [1] allows to detect shifts in both location and scale. An overview of nonparametric methods can be found in the excellent monograph of [8]. Alongside, research has developed around procedures that focus on U-statistics. Some examples include [19] in the case of independent observations, and [12] for time series subject to short-range dependence.

The assumption that a change would necessarily occur abruptly, as is implicit in the above-mentioned works, may be unrealistic in many contexts of application. For example, changes in temperature data subject to climatic perturbations are more likely to occur gradually than abruptly. Some authors addressed this issue. For example, [5] and [17] assumed a location model with a linear drift in CUSUM charts. An estimator of the time-of-change was proposed by [23], where it was assumed that the sequence

of means $\mu_1, \dots, \mu_n$ associated to $Y_1, \dots, Y_n$ are such that for a fixed $\alpha \in [0, 1]$,

$$\mu_j = \mu + \delta \left\{ \max \left(\frac{j - K}{n}, 0 \right) \right\}^\alpha,$$

where δ is a scale parameter and μ is the mean before a shift at time $K \in \{1, \dots, n-1\}$ happens. A generalisation of this model is considered in [2]. Nonparametric methods are developed in [58] to estimate the change in locally stationary processes.

An alternative way to deal with gradual changes has been proposed in [44] in order to overcome the limitations of the abrupt-change approaches. This method led to fruitful applications, see for instance [46] and [30] for the detection of gradual changes at the level of dependence measures. In the case of $\mathbb{R}$ -valued processes, the model in [44] assumes that for some univariate distribution functions F_1 and F_2 , the sequence of distribution functions associated to $Y_1, \dots, Y_n$ is such that for some $\mathbf{K} = (K_1, K_2) \in \mathcal{C}_n = \{(c_1, c_2) \in \mathcal{N}^2 : 1 \leq c_1 < c_2 \leq n\}$,

$$\mathbb{P}(Y_j \leq y) = \omega_j^{\mathbf{K}} F_1(y) + (1 - \omega_j^{\mathbf{K}}) F_2(y). \quad (2.1)$$

In the expression above, the sequence of weights $\omega_1^{\mathbf{K}}, \dots, \omega_n^{\mathbf{K}}$ is defined by $\omega_j^{\mathbf{K}} = 1$ if $j \leq K_1$, $\omega_j^{\mathbf{K}} = 0$ if $j \geq K_2$, whereas $\omega_j^{\mathbf{K}} = (K_2 - j)/(K_2 - K_1)$ otherwise. This model is a distributional extension of the smooth-change model proposed by [34], where rank tests and estimators for changes in a real-valued parameter are proposed; see also [47] for the exploration of inference procedures in Lombard's model. The model in Equation (2.1) implies that $\mathbb{E}(Y_j) = \omega_j^{\mathbf{K}} \mu_1 + (1 - \omega_j^{\mathbf{K}}) \mu_2$ and $\text{var}(Y_j) = \omega_j^{\mathbf{K}} \sigma_1^2 + (1 - \omega_j^{\mathbf{K}}) \sigma_2^2 + \omega_j^{\mathbf{K}} (1 - \omega_j^{\mathbf{K}}) (\mu_1 - \mu_2)^2$, where μ_1 and σ_1^2 (resp. μ_2 and σ_2^2) are the mean and variance of the law induced by F_1 (resp. F_2).

This paper introduces a general class of test statistics for the detection of gradual change-point in a series of univariate observations based on the sequence $\psi_1, \dots, \psi_n$ of characteristic functions stemming from $Y_1, \dots, Y_n$, where for $i^2 = -1$ and $t \in \mathbb{R}$, $\psi_j(t) = \mathbb{E}(e^{itY_j})$. Because of the one-to-one relationship between a distribution

function and its associated characteristic function (Cf), these tests can be seen as Cf counterparts of those in [44]. Such a Cf-based approach has been adopted in [24] and [25] for the development of tests for detecting abrupt changes. In addition to providing extensions to the procedure in [24] that accommodate for gradual changes, a large class of Cf-type statistics based on Lévy measures is introduced, relying on the generalization in [6] of the distance covariance introduced in the seminal paper by [53]. However, instead of using it to test for stochastic independence, the idea is adapted here to provide statistics for the testing and estimation of gradual changes.

The remainder of the paper is organised as follows. In Section 2, the tools that will serve as the basis of the proposed methodologies are described. A general class of test statistics based on Lévy measures is introduced in Section 3, and many specific examples are given. In Section 4, the weak convergence of the test statistics under the null hypothesis is established, as well as their consistency against any kinds of gradual-change alternatives; a method to compute asymptotically valid p-values is also described, as well as a consistent estimator of the time of change. The results of a simulation study are presented in Section 5. In Section 6, the methodologies are illustrated on climatic and financial data. Some future perspectives are discussed in Section 7. All the technical details are postponed to an appendix.

2.2 Basic tools

2.2.1 A criterion for testing the null hypothesis of no change

As is well known, there is a one-to-one correspondence between the distribution function $F_Y(y) = \mathbb{P}(Y \leq y)$ of a random variable Y and its characteristic function

$\psi_Y(t) = E(e^{itY})$, where $i^2 = -1$ and $t \in \mathbb{R}$. This relationship can be stated as

$$\psi_Y(t) = \int_{\mathbb{R}} e^{ity} dF_Y(y),$$

where the integral above is understood in the Riemann–Stieltjes sense.

Now suppose that a random sequence $Y_1, \dots, Y_n$ follows the gradual-change model in Equation (2.1) with time of change $\mathbf{K}_0 \in \mathcal{C}_n$, where $\mathcal{C}_n = \{(c_1, c_2) \in \mathcal{N}^2 : 1 \leq c_1 < c_2 \leq n\}$. Letting $\psi_1, \dots, \psi_n$ be the sequence of characteristic functions associated to $Y_1, \dots, Y_n$, one easily establishes that

$$\psi_j(t) = \omega_j^{\mathbf{K}_0} \Psi_1(t) + (1 - \omega_j^{\mathbf{K}_0}) \Psi_2(t), \quad (2.2)$$

where Ψ_1 and Ψ_2 are the characteristic functions associated to F_1 and F_2 , respectively. Proceeding as in [44], one defines the rescaled scalar product between $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$ and $\mathbf{b} = (b_1, \dots, b_n) \in \mathbb{R}^n$ by

$$\langle \mathbf{a}, \mathbf{b} \rangle = \frac{1}{n} \sum_{j=1}^n (a_j - \bar{a}) (b_j - \bar{b}) = \frac{1}{n} \sum_{j=1}^n (a_j - \bar{a}) b_j,$$

where $\bar{a} = (a_1 + \dots + a_n)/n$ and $\bar{b} = (b_1 + \dots + b_n)/n$. Letting $\boldsymbol{\omega}^{\mathbf{K}} = (\omega_1^{\mathbf{K}}, \dots, \omega_n^{\mathbf{K}})$ for $\mathbf{K} \in \mathcal{C}_n$ and $\boldsymbol{\psi} = (\psi_1, \dots, \psi_n)$, define the complex-valued function

$$Q_{\mathbf{K}}(t) = \langle \boldsymbol{\omega}^{\mathbf{K}}, \boldsymbol{\psi}(t) \rangle. \quad (2.3)$$

Lemma 1. When $Y_1, \dots, Y_n$ follows the gradual-change model in Eq. (2.1) with time of change $\mathbf{K}_0 \in \mathcal{C}_n$, then

$$Q_{\mathbf{K}}(t) = \{\Psi_1(t) - \Psi_2(t)\} \langle \boldsymbol{\omega}^{\mathbf{K}}, \boldsymbol{\omega}^{\mathbf{K}_0} \rangle.$$

Lemma 1 indicates that $Q_{\mathbf{K}}$ can be used as a criterion for testing the null hypothesis of

no change in a univariate series against the alternative of a gradual-change at unknown time points. Specifically, since the null hypothesis of a stable series is equivalent to $\Psi_1 = \Psi_2$ under the gradual-change model, the null and alternative hypotheses of interest can be written as

$$\mathbb{H}_0 : Q_{\mathbf{K}} \equiv 0, \quad \forall \mathbf{K} \in \mathcal{C}_n \quad \text{and} \quad \mathbb{H}_1 : Q_{\mathbf{K}} \neq 0, \quad \exists \mathbf{K} \in \mathcal{C}_n.$$

In the above formulation, $Q_{\mathbf{K}} \neq 0$ is understood in the sense that there exists some subset $\mathcal{R} \subset \mathbb{R}$ of non-null Lebesgue measure such that $Q_{\mathbf{K}}(t) \neq 0$ for all $t \in \mathcal{R}$.

2.2.2 Unbiased estimation of $Q_{\mathbf{K}}$

Based on realizations $Y_1, \dots, Y_n$ of a univariate process $(Y_t)_{t \in \mathbb{Z}} \subseteq \mathbb{R}$, an unbiased estimator of the vector $\boldsymbol{\psi} = (\psi_1, \dots, \psi_n)$ of characteristic functions is given by $\widehat{\boldsymbol{\psi}}(t) = (e^{itY_1}, \dots, e^{itY_n})$. In view of Equation (2.3), an unbiased estimator of $Q_{\mathbf{K}}$ is

$$Q_{\mathbf{K},n}(t) = \langle \omega^{\mathbf{K}}, \widehat{\boldsymbol{\psi}}(t) \rangle = \frac{1}{n} \sum_{j=1}^n (\omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) e^{itY_j}. \quad (2.4)$$

When $Y_1, \dots, Y_n$ follows the gradual-change model with time of change $\mathbf{K}_0$, then

$$\mathbb{E} \{Q_{\mathbf{K},n}(t)\} = \{\Psi_1(t) - \Psi_2(t)\} \langle \omega^{\mathbf{K}}, \omega^{\mathbf{K}_0} \rangle.$$

As a consequence, $\mathbb{E}\{Q_{\mathbf{K},n}(t)\} \equiv 0$ for all $\mathbf{K} \in \mathcal{C}_n$ under the null hypothesis of no change-point, since then, $\Psi_1 \equiv \Psi_2$.

2.3 A general class of test statistics

2.3.1 From Székely and Rizzo (2007) to Lévy measures

Test statistics based on characteristic functions are traditionally based on integrable weights that ensure their existence. A first step to overcome the need for integrable weights was taken in [53] upon defining⁷ the distance covariance. Specifically, the distance covariance between the random vectors $\mathbf{W}_1 \in \mathbb{R}^{d_1}$ and $\mathbf{W}_2 \in \mathbb{R}^{d_2}$ is defined for some $\lambda \in (0, 2)$ by

$$\text{dCov}_\lambda(\mathbf{W}_1, \mathbf{W}_2) \propto \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \frac{|\psi_{\mathbf{W}_1, \mathbf{W}_2}(\mathbf{t}_1, \mathbf{t}_2) - \psi_{\mathbf{W}_1}(\mathbf{t}_1) \psi_{\mathbf{W}_2}(\mathbf{t}_2)|^2}{|\mathbf{t}_1|^{d_1+\lambda} |\mathbf{t}_2|^{d_2+\lambda}} d\mathbf{t}_2 d\mathbf{t}_1, \quad (2.5)$$

where here and in the sequel, $|z| = \sqrt{a^2 + b^2}$ is the modulus of the complex number $z = a + bi$. Distance covariance is therefore a weighted distance between the joint characteristic function $\psi_{\mathbf{W}_1, \mathbf{W}_2}$ of $(\mathbf{W}_1, \mathbf{W}_2)$ and the product of the marginal characteristic functions. A key feature of distance covariance is that $\text{dCor}_\lambda(\mathbf{W}_1, \mathbf{W}_2) = 0$ if and only if $\mathbf{W}_1$ and $\mathbf{W}_2$ are independent, whereas $\text{dCor}_\lambda(\mathbf{W}_1, \mathbf{W}_2) > 0$ otherwise.

An interesting generalization of the distance covariance has been proposed in [6], based on the fact that the non-integrable weights $|\mathbf{t}_1|^{-(d_1+\lambda)}$ and $|\mathbf{t}_2|^{-(d_2+\lambda)}$ considered by [53] are particular cases of Lévy measures. This suggests substituting these weights with arbitrary Lévy measures ν_1 and ν_2 , *i.e.* Borel measures on $\mathbb{R}^{d_1} \setminus \{0\}$ and $\mathbb{R}^{d_2} \setminus \{0\}$, respectively, that satisfy

$$\int_{\mathbb{R}^{d_1} \setminus \{0\}} \min(|\mathbf{t}|, 1) \nu_1(d\mathbf{t}) < \infty \quad \text{and} \quad \int_{\mathbb{R}^{d_2} \setminus \{0\}} \min(|\mathbf{t}|, 1) \nu_2(d\mathbf{t}) < \infty.$$

By substituting $|\mathbf{t}_1|^{-(d_1+\lambda)}$ and $|\mathbf{t}_2|^{-(d_2+\lambda)}$ with ν_1 and ν_2 in Equation (2.5) yields the

generalized distance covariance as defined in [6], namely

$$\text{dCov}_{\nu_1, \nu_2}(\mathbf{W}_1, \mathbf{W}_2) = \int |\psi_{\mathbf{W}_1, \mathbf{W}_2}(\mathbf{t}_1, \mathbf{t}_2) - \psi_{\mathbf{W}_1}(\mathbf{t}_1) \psi_{\mathbf{W}_2}(\mathbf{t}_2)|^2 \nu_2(d\mathbf{t}_2) \nu_1(d\mathbf{t}_1),$$

where the integral above is taken on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \setminus \{0\}$.

2.3.2 Test statistics based on Lévy measures

The use of the distance covariance beyond the purpose of testing for independence is not very widespread, although it has already been explored by some authors. For example, [9] observed that their tests of homogeneity and symmetry are closely related to the distance covariance and to the related Székely–Rizzo inequality that first appeared in [54]. The distance covariance has also been used in [45] in the context of testing homogeneity hypotheses for copulas and in [3] for testing exchangeability of multivariate copulas. Let also mention the work by [60] for testing the equality of two survival functions based on their distance covariance.

In the sequel, a class of test statistics based on functionals of $Q_{\mathbf{K},n}$ is introduced. Specifically, for a given symmetric Lévy measure ν on $\mathbb{R} \setminus \{0\}$, define

$$S_{\mathbf{K},n}^\nu = n \int_{\mathbb{R} \setminus \{0\}} |Q_{\mathbf{K},n}(t)|^2 \nu(dt). \quad (2.6)$$

Lemma 2 below provides an explicit expression for $S_{\mathbf{K},n}^\nu$. Before stating it, note that Corollary 3.7.9 in [26] ensures that a Lévy measure ν on $\mathbb{R} \setminus \{0\}$ determines a unique negative definite function $\beta_\nu : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$\beta_\nu(y) = \int_{\mathbb{R} \setminus \{0\}} (1 - e^{ity}) \nu(dt). \quad (2.7)$$

Lemma 2. If ν is a Lévy measure characterized by the negative definite function

β_ν defined in Equation (2.7), then $S_{\mathbf{K},n}^\nu$ is a weighted U-statistic of degree two with kernel $-2\beta_\nu(y_1 - y_2)$ and sequence of weights $(\omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}})(\omega_{j'}^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}})$, *i.e.*

$$S_{\mathbf{K},n}^\nu = \frac{1}{n} \sum_{1 \leq j < j' \leq n} (\omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}})(\omega_{j'}^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) \{-2\beta_\nu(Y_j - Y_{j'})\}.$$

The expression of $S_{\mathbf{K},n}^\nu$ in Lemma 2 is helpful for its implementation, as well as for the derivation of its asymptotic distribution.

In order to be able to capture gradual changes that may occur at any possible times, statistics that combine $S_{\mathbf{K},n}^\nu$ for each $\mathbf{K} \in \mathcal{C}_n$ are considered. Specifically, consider the global test statistics

$$T_{n,\Sigma}^\nu = \frac{1}{n^2} \sum_{\mathbf{K} \in \mathcal{C}_n} \eta\left(\frac{\mathbf{K}}{n}\right) S_{\mathbf{K},n}^\nu \quad \text{and} \quad T_{n,\vee}^\nu = \max_{\mathbf{K} \in \mathcal{C}_n} \left(\frac{\mathbf{K}}{n}\right) S_{\mathbf{K},n}^\nu. \quad (2.8)$$

In the above expression $\eta : \mathcal{C} \rightarrow \mathbb{R}$ is a continuous weight function, where $\mathcal{C} = \{\kappa = (\kappa_1, \kappa_2) \in [0, 1]^2 : \kappa_1 < \kappa_2\}$.

2.3.3 Some specific Lévy measures

Any density f defined on $\mathbb{R}$ is a Lévy measure, since

$$\int_{\mathbb{R} \setminus \{0\}} \min(|t|, 1) f(t) \leq \int_{\mathbb{R}} f(t) < \infty.$$

In that case, the associated negative definite function is $\beta_\nu(y) = 1 - \Psi_f(y)$, where Ψ_f is the characteristic function of f . For a symmetric spherical law with characteristic exponent $\lambda \in (0, 2]$, one has $\Psi_f(t) = e^{-|t|^\lambda}$ and then $\beta_\nu(y) = 1 - e^{-|y|^\lambda}$. The Cauchy and Normal distributions are members of the spherical family when, respectively, $\lambda = 1$

and $\lambda = 2$. The formula in Lemma 2 then becomes

$$S_{\mathbf{K},n}^\nu = \frac{2}{n} \sum_{1 \leq j < j' \leq n} (\omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) (\omega_{j'}^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) \exp(|Y_j - Y_{j'}|^\lambda).$$

It has already been seen that $\nu(d\mathbf{t}) \propto |\mathbf{t}|^{-(d+\lambda)}$ is a d -variate Lévy measure for any $\lambda \in (0, 2)$. When $d = 1$, one has $\nu(dt) = |t|^{-(1+\lambda)}$. As already computed by [52], $\beta_\nu(y) \propto |y|^\lambda$ in that case. Therefore, the associated test statistic is proportional to

$$S_{\mathbf{K},n}^\nu = \frac{2}{n} \sum_{1 \leq j < j' \leq n} (\omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) (\omega_{j'}^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) |Y_j - Y_{j'}|^\lambda.$$

Another example of a d -variate Lévy measure can be found in [50]. When $d = 1$, the latter expresses as

$$\nu(dt) = \frac{\lambda^2}{2\sqrt{\pi}} \int_0^\infty u^{-1/2} \exp\left(\lambda^2 u - \frac{|t|}{4u}\right) du dt,$$

where $\lambda > 0$. In that case, $\beta_\nu(y) = \{1 + (\lambda/y)^2\}^{-1}$ and then

$$S_{\mathbf{K},n}^\nu = -\frac{2}{n} \sum_{1 \leq j < j' \leq n} (\omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) (\omega_{j'}^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) \left(1 + \frac{\lambda^2}{|Y_j - Y_{j'}|^2}\right)^{-1}.$$

2.3.4 Relationship with a test by [24]

The abrupt-change model is a special case of the gradual-change model in Equation (2.1) when $K_1 = K$ and $K_2 = K + 1$ for some $K \in \{1, \dots, n - 1\}$. Under this

model, $\omega_j^{\mathbf{K}} = \mathbb{I}(j \leq K)$ and $\bar{\omega}^{\mathbf{K}} = K/n$. Since $\beta_\nu(0) = 0$ and $\beta_\nu(Y_j - Y_{j'}) = \beta_\nu(Y_{j'} - Y_j)$,

$$\begin{aligned} S_{K,n}^\nu &= -\frac{2}{n} \sum_{1 \leq j < j' \leq n} \left\{ \mathbb{I}(j \leq K) - \frac{K}{n} \right\} \left\{ \mathbb{I}(j' \leq K) - \frac{K}{n} \right\} \beta_\nu(Y_j - Y_{j'}) \\ &= -\frac{1}{n} \sum_{j,j'=1}^n \left\{ \mathbb{I}(j \leq K) - \frac{K}{n} \right\} \left\{ \mathbb{I}(j' \leq K) - \frac{K}{n} \right\} \beta_\nu(Y_j - Y_{j'}) \\ &= -\frac{1}{n} \sum_{j,j'=1}^K \beta_\nu(Y_j - Y_{j'}) + \frac{2K}{n^2} \sum_{j=1}^K \sum_{j'=1}^n \beta_\nu(Y_j - Y_{j'}) - \frac{K^2}{n^3} \sum_{j,j'=1}^n \beta_\nu(Y_j - Y_{j'}). \end{aligned}$$

If the Lévy measure ν is a density f , it has already been shown that $\beta_\nu(y) = 1 - \Psi_f(y)$.

The above expression then becomes

$$S_{K,n}^f = \frac{1}{n} \sum_{j,j'=1}^K \Psi_f(Y_j - Y_{j'}) - \frac{2K}{n^2} \sum_{j=1}^K \sum_{j'=1}^n \Psi_f(Y_j - Y_{j'}) + \frac{K^2}{n^3} \sum_{j,j'=1}^n \Psi_f(Y_j - Y_{j'}).$$

Long but straightforward calculations yield $S_{K,n}^f = \{K^2(n-K)^2/n^3\} V_{K,n}$, where

$$\begin{aligned} V_{K,n} &= \frac{1}{K^2} \sum_{j,j'=1}^K \Psi_f(Y_j - Y_{j'}) - \frac{2}{K(n-K)} \sum_{j=1}^K \sum_{j'=K+1}^n \Psi_f(Y_j - Y_{j'}) \\ &\quad + \frac{1}{(n-K)^2} \sum_{j,j'=K+1}^n \Psi_f(Y_j - Y_{j'}). \end{aligned}$$

The expression above for $V_{K,n}$ is the same as that of Equation (8) in [24]. Hence, the test statistic $T_{n,\nu}^\nu$ in Equation (2.8) when $\nu := f$ and in the special case of an abrupt change corresponds to the test statistic in Equation (7) in [24] when $\gamma = 1$, namely

$$T_{n,\nu}^{\text{HM}} = \max_{1 \leq K < n} S_{K,n}^f.$$

The test statistic $T_{n,\nu}^\nu$ can then be seen as a generalization of $T_{n,\nu}^{\text{HM}}$ to the case of gradual changes and to the general family of Lévy measures.

2.4 Large-sample results

2.4.1 Relationship with the sequential empirical process

The sequential empirical process appears naturally when testing for an abrupt change point. In broad terms, it is defined as a weighted distance between the “after the change” and “before the change” distribution functions. In the case of a univariate series $Y_1, \dots, Y_n$, the sequential empirical process is defined for $(s, y) \in [0, 1] \times \mathbb{R}$ by

$$\mathbb{F}_n(s, y) = \frac{\lfloor ns \rfloor (n - \lfloor ns \rfloor)}{n^{3/2}} \{F_{1:\lfloor ns \rfloor}(y) - F_{\lfloor ns \rfloor+1:n}(y)\}, \quad (2.9)$$

where for $1 \leq L_1 \leq L_2 \leq n$,

$$F_{L_1:L_2}(y) = \frac{1}{L_2 - L_1 + 1} \sum_{\ell=L_1}^{L_2} \mathbb{I}(Y_\ell \leq y).$$

The next result shows that $Q_{\mathbf{K},n}$, and consequently $S_{\mathbf{K},n}^\nu$, can be expressed as a functional of $\mathbb{F}_n$ via a generalisation of the sequential empirical process, namely

$$\mathbb{L}_{\mathbf{K},n}(y) = \frac{1}{K_2 - K_1} \sum_{j=K_1}^{K_2-1} \mathbb{F}_n\left(\frac{j}{n}, y\right). \quad (2.10)$$

Proposition 1. The empirical process $Q_{\mathbf{K},n}$ can be expressed as

$$Q_{\mathbf{K},n}(t) = \frac{1}{\sqrt{n}} \int_{\mathbb{R}} e^{ity} \mathbb{L}_{\mathbf{K},n}(dy),$$

where the integral above is understood in the Riemann–Stieltjes sense. Moreover,

$$S_{\mathbf{K},n}^\nu = - \int_{\mathbb{R}^2} \beta_\nu(y - y') \mathbb{L}_{\mathbf{K},n}(dy) \mathbb{L}_{\mathbf{K},n}(dy').$$

Proposition 1 shows that $S_{\mathbf{K},n}^\nu$ expresses as a quadratic functional of the empirical process $\mathbb{L}_{\mathbf{K},n}$. Even if this functional representation will not be exploited for the derivation of the asymptotic behavior of $S_{\mathbf{K},n}^\nu$, it will prove useful for the derivation of a suitably adapted resampling method.

2.4.2 Weak convergence of the test statistics

The convergence in law of $S_{\mathbf{K},n}^\nu$ for a fixed time of change $\mathbf{K} \in \mathcal{C}_n$ is a special case of a result due to [35] for weighted U-statistics. Before stating it, let $\mathcal{C} = \{\boldsymbol{\kappa} = (\kappa_1, \kappa_2) \in [0, 1]^2 : \kappa_1 < \kappa_2\}$ and recall that a F -Wiener process W_F is a centered Gaussian process on $\mathbb{R}$ such that $\mathbb{E}\{W_F(y) W_F(y')\} = \min\{F(y), F(y')\}$. Let also $\Lambda(\boldsymbol{\kappa}, \boldsymbol{\kappa}_0)$ be the limit of $\langle \omega(\lfloor n\boldsymbol{\kappa} \rfloor), \omega(\lfloor n\boldsymbol{\kappa}_0 \rfloor) \rangle$ as $n \rightarrow \infty$, which according to Eq. (7) of [44] expresses as

$$\Lambda(\boldsymbol{\kappa}, \boldsymbol{\kappa}_0) = \int_{\kappa_1}^{\kappa_2} \int_{\kappa_{01}}^{\kappa_{02}} \frac{\{\min(s, s') - ss'\}}{(\kappa_2 - \kappa_1)(\gamma_2 - \gamma_1)} ds' ds. \quad (2.11)$$

In particular,

$$\Lambda(\boldsymbol{\kappa}, \boldsymbol{\kappa}) = \frac{2\kappa_1 + \kappa_2}{3} - \left(\frac{\kappa_1 + \kappa_2}{2} \right)^2.$$

Proposition 2. Under $Y_1, \dots, Y_n$ i.i.d. F and as long as $\mathbb{E}\{\beta_\nu^2(Y_1, Y_2)\} < \infty$, the statistic $S_{\lfloor n\boldsymbol{\kappa} \rfloor, n}^\nu$ converges in distribution to the random variable

$$\mathbb{S}_{\boldsymbol{\kappa}}^\nu = \Lambda(\boldsymbol{\kappa}, \boldsymbol{\kappa}) \int_{\mathbb{R}^2} \beta(y_1 - y_2) \{F(dy_1) F(dy_2) - W_F(dy_1) W_F(dy_2)\}.$$

Proposition 3. Under the conditions of Proposition 2, the statistic $S_{\mathbf{K},n}^\nu$ converges weakly in the space $\ell^\infty(\mathcal{C})$ of bounded functions on $\mathcal{C}$ to a tight limit $\mathbb{S}_{\boldsymbol{\kappa}}^\nu$.

Upon noting that

$$T_{n,\Sigma}^\nu = \frac{1}{n^2} \sum_{\mathbf{K} \in \mathcal{C}_n} \eta\left(\frac{\mathbf{K}}{n}\right) S_{\mathbf{K},n}^\nu = \int_{[0,1]^2} \mathbb{I}(\lfloor n\kappa_1 \rfloor < \lfloor n\kappa_2 \rfloor) \eta\left(\frac{\lfloor n\boldsymbol{\kappa} \rfloor}{n}\right) S_{\lfloor n\boldsymbol{\kappa} \rfloor, n}^\nu d\boldsymbol{\kappa},$$

a consequence of Proposition 3 is that

$$T_{n,\Sigma}^\nu \rightsquigarrow \mathbb{T}_\Sigma^\mu = \int_{\mathcal{C}} \eta(\boldsymbol{\kappa}) S_\kappa^\nu d\boldsymbol{\kappa}.$$

Similarly,

$$T_{n,\vee}^\nu = \max_{\mathbf{K} \in \mathcal{C}_n} \eta\left(\frac{\mathbf{K}}{n}\right) S_{\mathbf{K},n}^\nu = \sup_{\boldsymbol{\kappa} \in \mathcal{C}} \eta\left(\frac{\lfloor n\boldsymbol{\kappa} \rfloor}{n}\right) S_{\lfloor n\boldsymbol{\kappa} \rfloor, n}^\nu \rightsquigarrow \mathbb{T}_\vee^\mu = \sup_{\boldsymbol{\kappa} \in \mathcal{C}} \eta(\boldsymbol{\kappa}) S_\kappa^\nu.$$

2.4.3 A resampling method for the computation of p-values

The problem of computing p-values for non-parametrically specified models under the null hypothesis can often be tackled by using the multiplier bootstrap. The method is described for instance in [29] in the i.i.d. case. There do not appear to be any results concerning multiplier versions of weighted U-statistics whose kernel is degenerate, as is the case for the introduced test statistics.

Let $\delta_1, \dots, \delta_n$ be i.i.d. Exponential with mean one, independent of the observations $Y_1, \dots, Y_n$. These former are called the “multiplier variables”. Letting $\bar{\delta}_{1:j} = (\delta_1 + \dots + \delta_j)/j$, define the multiplier version of the sequential empirical process $\mathbb{F}_n$ in Equation (2.9) as

$$\widehat{\mathbb{F}}_n\left(\frac{j}{n}, y\right) = \frac{1}{\sqrt{n}} \left\{ \sum_{\ell=1}^j \left(\frac{\delta_\ell}{\bar{\delta}_{1:j}} - 1 \right) \mathbb{I}(Y_\ell \leq y) - \frac{j}{n} \sum_{\ell=1}^n \left(\frac{\delta_\ell}{\bar{\delta}_{1:n}} - 1 \right) \mathbb{I}(Y_\ell \leq y) \right\}. \quad (2.12)$$

The use of stochastic weights of the form $(\delta_\ell/\bar{\delta}_{1:j}) - 1$ is referred to the bayesian bootstrap in the literature, see for instance [57] and [29]. A multiplier version of $Q_{\mathbf{K},n}$

is then

$$\widehat{Q}_{\mathbf{K},n}(t) = \frac{-it}{\sqrt{n}} \int_{\mathbb{R}} e^{ity} \widehat{\mathbb{L}}_{\mathbf{K},n}(y) dy,$$

where in view of Equation (2.10), the multiplier version of $\mathbb{L}_{\mathbf{K},n}$ is given by

$$\widehat{\mathbb{L}}_{\mathbf{K},n}(y) = \frac{1}{K_2 - K_1} \sum_{j=K_1}^{K_2-1} \widehat{\mathbb{F}}_n \left(\frac{j}{n}, y \right). \quad (2.13)$$

In view of Proposition 1 and based on the form of the test statistic in Equation (2.6), the proposed multiplier version of $S_{\mathbf{K},n}^\nu$ is

$$\widehat{S}_{\mathbf{K},n}^\nu = \int_{\mathbb{R}} \left| \widehat{Q}_{\mathbf{K},n}(t) \right|^2 \nu(dt).$$

Lemma 3. Define the matrix $\mathfrak{B} \in \mathbb{R}^{n \times n}$ whose entries are $\mathfrak{B}_{\ell\ell'} = \beta_\nu(Y_\ell - Y_{\ell'})$ and let $\mathfrak{B}_{1:j,1:k} \in \mathbb{R}^{j \times k}$ be the block matrix of the first j lines and first k columns of $\mathfrak{B}$, $j, k \in \{1, \dots, n\}$. The multiplier version of $S_{\mathbf{K},n}^\nu$ can then be expressed as

$$\begin{aligned} \widehat{S}_{\mathbf{K},n}^\nu = \frac{1}{n^2(K_2 - K_1)^2} \sum_{j,j'=K_1+1}^{K_2} \left\{ \boldsymbol{\delta}_{1:j} \mathfrak{B}_{1:j,1:j'} \boldsymbol{\delta}_{1:j'}^\top - \frac{j'}{n} \boldsymbol{\delta}_{1:j} \mathfrak{B}_{1:j,1:n} \boldsymbol{\delta}_{1:n}^\top \right. \\ \left. - \frac{j}{n} \boldsymbol{\delta}_{1:n} \mathfrak{B}_{1:n,1:j'} \boldsymbol{\delta}_{1:j'}^\top + \frac{jj'}{n^2} \boldsymbol{\delta}_{1:n} \mathfrak{B}_{1:n,1:n} \boldsymbol{\delta}_{1:n}^\top \right\}, \end{aligned}$$

where $\boldsymbol{\delta}_{1:j} = (\delta_1/\bar{\delta}_{1:j} - 1, \dots, \delta_j/\bar{\delta}_{1:j} - 1)$.

The suggested multiplier versions of the global test statistics $T_{n,\Sigma}^\nu$ and $T_{n,\vee}^\nu$ are given respectively by

$$\widehat{T}_{n,\Sigma}^\nu = \frac{1}{n^2} \sum_{\mathbf{K} \in \mathcal{C}_n} \eta \left(\frac{\mathbf{K}}{n} \right) \widehat{S}_{\mathbf{K},n}^\nu \quad \text{and} \quad \widehat{T}_{n,\vee}^\nu = \max_{\mathbf{K} \in \mathcal{C}_n} \left(\frac{\mathbf{K}}{n} \right) \widehat{S}_{\mathbf{K},n}^\nu.$$

2.4.4 Consistency of the tests and change-point estimators

It will first be established that the tests based on $T_{n,\Sigma}^\nu$ and $T_{n,\nu}^\nu$ are consistent under all kinds of gradual-change alternatives.

Proposition 4. Suppose that $Y_1, \dots, Y_n$ follows the gradual-change model in Eq. (2.1) with time of change $\mathbf{K}_0 \in \mathcal{C}_n$ and characteristic functions such that $\Psi_1 \neq \Psi_2$. Then, for $\boldsymbol{\kappa}, \boldsymbol{\kappa}_0 \in \mathcal{C}$ such that $\mathbf{K} = \lfloor n\boldsymbol{\kappa} \rfloor$ and $\mathbf{K}_0 = \lfloor n\boldsymbol{\kappa}_0 \rfloor$,

$$\frac{S_{\mathbf{K},n}^\nu}{n} \xrightarrow{\mathbb{P}} \{\Lambda(\boldsymbol{\kappa}, \boldsymbol{\kappa}_0)\}^2 \int_{\mathbb{R} \setminus \{0\}} |\Psi_1(t) - \Psi_2(t)|^2 \nu(dt) > 0.$$

As a consequence, the tests based on $T_{n,\Sigma}^\nu$ and $T_{n,\nu}^\nu$ are consistent since $S_{\mathbf{K},n}^\nu \xrightarrow{\mathbb{P}} \infty$.

Proposition 4 suggests a way to estimate the unknown location $\mathbf{K}_0$ of a gradual change, based on the fact that

$$\mathbb{E}(S_{\mathbf{K},n}^\nu) = \{\langle \omega^{\mathbf{K}}, \omega^{\mathbf{K}_0} \rangle\}^2 \int_{\mathbb{R} \setminus \{0\}} |\Psi_1(t) - \Psi_2(t)|^2 \nu(dt).$$

Invoking the Cauchy–Schwarz inequality,

$$\begin{aligned} \frac{\mathbb{E}(S_{\mathbf{K},n}^\nu)}{\langle \omega^{\mathbf{K}}, \omega^{\mathbf{K}} \rangle} &= \frac{\{\langle \omega^{\mathbf{K}}, \omega^{\mathbf{K}_0} \rangle\}^2}{\langle \omega^{\mathbf{K}}, \omega^{\mathbf{K}} \rangle} \int_{\mathbb{R} \setminus \{0\}} |\Psi_1(t) - \Psi_2(t)|^2 \nu(dt) \\ &\leq \frac{\langle \omega^{\mathbf{K}}, \omega^{\mathbf{K}} \rangle \langle \omega^{\mathbf{K}_0}, \omega^{\mathbf{K}_0} \rangle}{\langle \omega^{\mathbf{K}}, \omega^{\mathbf{K}} \rangle} \int_{\mathbb{R} \setminus \{0\}} |\Psi_1(t) - \Psi_2(t)|^2 \nu(dt) \\ &= \langle \omega^{\mathbf{K}_0}, \omega^{\mathbf{K}_0} \rangle \int_{\mathbb{R} \setminus \{0\}} |\Psi_1(t) - \Psi_2(t)|^2 \nu(dt), \end{aligned}$$

the equality occurring if and only if $\mathbf{K} = \mathbf{K}_0$. This suggests estimating the time of change $\mathbf{K}_0$ with the estimator $\widehat{\mathbf{K}} = \lfloor n\widehat{\boldsymbol{\kappa}} \rfloor$, where

$$\widehat{\boldsymbol{\kappa}} = \operatorname{argsup}_{\boldsymbol{\kappa} \in \mathcal{C}} \frac{S_{\lfloor n\boldsymbol{\kappa} \rfloor, n}^\nu}{\Lambda(\boldsymbol{\kappa}, \boldsymbol{\kappa})}.$$

The estimator $\hat{\boldsymbol{\kappa}}$ is consistent for $\boldsymbol{\kappa}_0$, since in view of Proposition 4,

$$\hat{\boldsymbol{\kappa}} \xrightarrow{\mathbb{P}} \operatorname{argsup}_{\boldsymbol{\kappa} \in \mathcal{C}} \frac{\{\Lambda(\boldsymbol{\kappa}, \boldsymbol{\kappa}_0)\}^2}{\Lambda(\boldsymbol{\kappa}, \boldsymbol{\kappa})} = \boldsymbol{\kappa}_0.$$

2.5 Size and power of the proposed tests

2.5.1 A warp-speed bootstrap method

Suppose that the goal is to evaluate the power of a statistical test whose implementation is based on a certain bootstrap method in order to approximate its p-value. In a typical investigation of the power of such a test, a large number M of data sets are simulated from a given distributional model under the alternative hypothesis, and then for each of these simulated data sets, a large number B of parametric bootstrap samples are generated. This results in a total of $M \times B$ data sets of size n that must be simulated and for which the statistical methodology under investigation must be performed; this task may be very time consuming. The Warp-speed bootstrap introduced in [18] allows to considerably reduce the calculation time in these kinds of Monte Carlo power studies.

A version of the warp-speed bootstrap is introduced here when the re-sampling method is based on the multiplier bootstrap. Specifically, for each $b \in \{1, \dots, B\}$, a sample of size n from the alternative hypothesis is generated. Then, the test statistic $S_n^{(b)}$ and one multiplier version $\hat{S}_n^{(b)}$ are computed. This results in B replicates $S_n^{(1)}, \dots, S_n^{(B)}$ of the test statistic under the alternative hypothesis of interest, as well as B replicates $\hat{S}_n^{(1)}, \dots, \hat{S}_n^{(B)}$ of the test statistic under the null hypothesis. Letting $\hat{q}_\alpha$ be the α -th empirical upper percentile of $\hat{S}_n^{(1)}, \dots, \hat{S}_n^{(B)}$, the estimated probability

of rejection of the null hypothesis is given by

$$\widehat{\mathbb{P}} = \frac{1}{B} \sum_{b=1}^B \mathbb{I}(S_n^{(b)} > \widehat{q}_\alpha).$$

In total, only $2B$ samples need to be simulated, compared to the MB needed in a standard Monte Carlo methodology.

2.5.2 Ability of the tests to keep their nominal level

The first aspect to investigate is to ensure that the tests manage to hold their nominal level under a null hypothesis. Since no assumption is made about the underlying marginal distribution, apart from being continuous, the proposed tests are non-parametric in nature. Nevertheless, the behavior of the test statistics depends on the unknown marginal distribution. In order to take this into account, two distributions have been considered. The first consists of a Normal distribution with constant mean $\mu = 0$ and standard deviation $\sigma = 1$, whereas the second is the Gamma with parameters $\alpha = 9$ and $\beta = 1/3$ ($\mu = 3$ and $\sigma = 1$). The tests under investigation are those based on the statistics $T_{n,\Sigma}^\nu$ and $T_{n,\vee}^\nu$ in the case of the *Spherical* weight based on $\beta_\nu(y) = 1 - e^{|y|^\lambda}$, $\lambda \in (0, 2]$, the *Székelly-Rizzo* weight based on $\beta_\nu(y) = |y|^\lambda$, $\lambda \in (0, 2)$, and the *Schilling-Schnurr* weight, for which, $\beta_\nu(y) = 1/\{1 + (\lambda/y)^2\}^2$, $\lambda > 0$.

The results of this investigation are found in Table 2.1 in the case when $\lambda \in \{1, 3/2, 2\}$. As can be seen, the ability of the tests to keep their 5% nominal level is good, regardless of the type of marginal distributions. As expected, the results are a little better, *i.e.*, closer to 5%, when $n = 200$ compared to $n = 100$. The only slightly problematic case is when the *Schilling-Schnurr* weight is used with $\lambda = 2$ when $n = 100$, but the problem resolves when $n = 200$.

TABLE 2.1 – Percentages of rejection of the null hypothesis under $\mathbb{H}_0$, as estimated from $B = 1000$ warp-speed bootstrap replicates and $N = 40$, under Normal and Gamma marginal distributions when $n = 100$ (upper panel) and $n = 200$ (bottom panel)

Model	λ	Spherical		Székely–R.		Schilling–S.	
		$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$
Normal $\mu = 0, \sigma = 1$	1	2.4	2.3	3.1	2.5	3.4	2.8
	3/2	5.8	4.5	6.0	5.5	2.9	2.4
	2	4.5	4.6	4.4	4.3	7.0	6.9
Gamma $\mu = 3, \sigma = 1$	1	4.7	4.5	4.3	4.2	4.9	4.3
	3/2	4.6	4.2	4.2	4.0	5.7	5.4
	2	4.8	4.7	4.3	3.6	5.8	5.9
Normal $\mu = 0, \sigma = 1$	1	4.5	3.8	3.6	3.7	4.1	3.9
	3/2	4.5	4.6	5.2	4.8	4.6	4.0
	2	5.1	4.7	3.6	2.9	4.9	4.6
Gamma $\mu = 3, \sigma = 1$	1	4.8	4.2	5.3	4.7	4.7	4.5
	3/2	3.9	3.9	6.2	5.4	4.6	4.6
	2	5.7	5.9	5.2	4.8	4.8	4.5

2.5.3 Investigation of the power of the tests

Now that the tests' ability to keep their nominal level is confirmed, the next step consists in checking at which point they are sensitive against change-point alternatives. In what follows, the approximate power of the test statistics $T_{n,\Sigma}^\nu$ and $T_{n,\vee}^\nu$ when ν is either the *Spherical*, *Székely–Rizzo* and *Schilling–Schnurr* weights when $\lambda = 3/2$ are obtained and compared to that of a test in [44]. The latter is based on the Cramér–von Mises type statistic

$$T_{n,\Sigma}^Q = \frac{1}{n^{3/2}} \sum_{\mathbf{K} \in \mathcal{C}_n} \left(\frac{K_2 - K_1}{n} \right) \sqrt{\frac{1}{n} \sum_{i=1}^n \mathbb{L}_{\mathbf{K},n}^2(Y_i)},$$

where $\mathbb{L}_{\mathbf{K},n}$ is defined in Equation (2.10). Among the tests considered in [44], the one based on $T_{n,\Sigma}^Q$ was identified as being, overall, the most powerful.

The results in Table 2.2 concern the power of the seven tests when the underlying distribution is the Normal with a constant standard deviation and a mean that gradually shifts between two times that are more or less distant. Table 2.3 presents similar results in the case where the mean is stable in time, but the standard deviation gradually shifts between two change-points. The results both in Table 2.2 and Table 2.3 indicate, unsurprisingly, that the power of a given test increases when the sample size goes from $n = 100$ to $n = 200$. As expected, the tests are more sensitive as the difference between the before-the-change and after-the-change values of the parameter increases. Overall, the ability of a given test to detect a change is similar when $\boldsymbol{\kappa} \in \{(.5, .5), (.4, .5), (.4, .6)\}$, and decreases significantly when $\boldsymbol{\kappa} = (.2, .8)$, and even more when $\boldsymbol{\kappa} = (.1, .5)$.

Among the newly proposed tests, those based on the *Schilling–Schnurr* weight are systematically more powerful for detecting changes in the mean; the latter perform slightly better than the Cramér–von Mises statistic $T_{n,\Sigma}^Q$ of [44]. However, the performance of the tests based on the *Schilling–Schnurr* weight is much less convincing for detecting changes in the standard deviation. In that case, it is the tests based on the *Spherical* and *Székel–Rizzo* weights that stand out as being the most performing, being much more powerful than the test based on $T_{n,\Sigma}^Q$.

One closes this subsection by considering the power of the seven tests under five scenarios involving the Gamma distribution whose density is given by

$$f_{\alpha,\beta}(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha}, \quad x > 0.$$

In order that the mean and standard deviation be respectively μ and σ , the parameters have been set to $\alpha = \mu^2/\sigma^2$ and $\beta = \sigma^2/\mu$. The results are in Table 2.4.

TABLE 2.2 – Percentages of rejection of the null hypothesis, as estimated from $B = 1000$ warp-speed bootstrap replicates and $N = 40$, under independent Normal random variables with a unit standard deviation and a mean that gradually shifts from μ_1 to μ_2 when $n = 100$ (upper panel) and $n = 200$ (bottom panel)

(κ_1, κ_2)	(μ_1, μ_2)	Spherical		Székely–R.		Schilling–S.		$T_{n,\Sigma}^Q$
		$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	
(.5,.5)	(0,1/3)	19.4	17.8	23.2	21.6	29.9	28.9	28.0
	(0,2/3)	65.3	63.0	77.1	73.4	81.5	79.6	79.1
(.4,.5)	(0,1/3)	19.0	18.0	26.8	26.2	32.8	31.4	32.0
	(0,2/3)	63.8	60.1	74.1	72.7	80.6	79.3	77.2
(.4,.6)	(0,1/3)	18.7	18.4	26.0	24.2	30.0	28.1	26.4
	(0,2/3)	63.2	61.8	77.3	76.8	82.4	81.4	80.9
(.2,.8)	(0,1/3)	18.4	16.6	21.4	20.6	21.7	21.0	17.5
	(0,2/3)	53.2	51.8	64.7	62.0	71.6	70.1	65.9
(.1,.5)	(0,1/3)	14.8	14.7	13.8	13.4	17.2	16.5	18.9
	(0,2/3)	44.9	42.4	49.9	49.9	61.1	61.9	53.0
(.5,.5)	(0,1/3)	37.4	35.2	44.3	41.5	60.7	59.7	51.4
	(0,2/3)	94.3	93.7	97.0	96.5	98.7	98.7	98.0
(.4,.5)	(0,1/3)	33.2	32.7	42.6	43.9	52.3	49.7	49.8
	(0,2/3)	93.4	93.1	97.1	96.9	98.8	98.6	97.6
(.4,.6)	(0,1/3)	34.7	34.0	46.7	46.4	52.1	50.8	46.2
	(0,2/3)	91.9	91.3	97.5	97.3	98.0	97.6	98.4
(.2,.8)	(0,1/3)	27.6	28.2	36.1	33.9	43.8	42.3	40.3
	(0,2/3)	84.2	83.6	89.7	89.4	92.5	92.4	93.0
(.1,.5)	(0,1/3)	25.7	25.2	31.5	30.6	35.9	35.3	34.5
	(0,2/3)	70.1	71.0	78.4	80.8	87.1	88.3	84.1

TABLE 2.3 – Percentages of rejection of the null hypothesis, as estimated from $B = 1000$ warp-speed bootstrap replicates and $N = 40$, under independent Normal random variables with a zero mean and a standard deviation that gradually shifts from σ_1 to σ_2 when $n = 100$ (upper panel) and $n = 200$ (bottom panel)

(κ_1, κ_2)	(σ_1, σ_2)	Spherical		Székely–R.		Schilling–S.		$T_{n,\Sigma}^Q$
		$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	
(.5,.5)	(1,4/3)	14.0	12.4	14.9	12.7	5.9	5.3	9.4
	(1,5/3)	48.4	44.9	48.3	41.9	13.1	10.9	15.9
(.4,.5)	(1,4/3)	12.2	11.0	14.7	12.3	7.3	6.0	5.6
	(1,5/3)	44.9	41.8	48.2	44.6	11.7	10.5	19.7
(.4,.6)	(1,4/3)	19.0	16.8	16.7	12.1	4.1	3.8	7.1
	(1,5/3)	52.7	48.5	47.9	44.4	11.9	11.1	11.4
(.2,.8)	(1,4/3)	13.7	10.7	10.8	9.8	7.6	6.5	7.8
	(1,5/3)	33.2	30.5	31.1	28.1	10.1	8.6	16.7
(.1,.5)	(1,4/3)	8.7	7.4	8.1	7.7	5.3	4.5	5.6
	(1,5/3)	22.4	22.7	24.3	21.7	9.0	7.3	10.7
(.5,.5)	(1,4/3)	43.1	38.2	29.7	27.5	9.4	8.7	15.6
	(1,5/3)	79.6	78.3	87.4	86.2	31.2	24.7	44.3
(.4,.5)	(1,4/3)	32.4	29.3	29.9	26.0	7.6	7.4	9.5
	(1,5/3)	81.5	79.0	83.6	81.7	22.5	17.6	45.0
(.4,.6)	(1,4/3)	32.4	28.6	31.5	29.0	6.6	6.2	13.0
	(1,5/3)	83.7	81.0	88.4	86.4	21.2	16.5	46.0
(.2,.8)	(1,4/3)	28.2	27.8	25.7	24.8	9.3	9.1	8.0
	(1,5/3)	69.7	67.4	70.1	66.3	17.5	14.4	29.9
(.1,.5)	(1,4/3)	16.2	16.9	14.8	13.5	7.0	5.8	7.5
	(1,5/3)	52.8	52.4	50.6	51.8	17.5	13.6	26.1

TABLE 2.4 – Percentages of rejection of the null hypothesis under series of size $n = 100$ with Gamma marginal distributions whose mean and/or standard deviation gradually shifts, as estimated from $B = 1000$ warp-speed bootstrap replicates and $N = 40$

(κ_1, κ_2)	(μ_1, μ_2)	(σ_1, σ_2)	Spherical		Székely–R.		Schilling–S.		$T_{n,\Sigma}^Q$
			$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	
(.5,.5)	(3,3.5)	(1,1)	46.5	45.3	41.3	39.8	60.2	57.7	54.6
	(3,4)	(1,1)	98.0	97.3	96.0	95.8	99.6	99.2	99.5
	(3,3)	(1,1.5)	29.6	27.3	28.4	24.3	20.8	16.6	12.3
	(3,3)	(1,2)	81.8	76.4	82.4	80.2	61.4	53.5	50.1
	(3,3.5)	(1,1.5)	29.4	26.8	30.9	26.1	44.8	42.5	31.7
(.4,.6)	(3,3.5)	(1,1)	51.3	49.7	40.7	38.5	59.4	56.6	58.3
	(3,4)	(1,1)	96.8	95.9	96.5	96.1	98.4	98.4	98.9
	(3,3)	(1,1.5)	30.7	29.3	31.3	25.8	20.4	17.0	11.7
	(3,3)	(1,2)	83.1	81.2	81.0	79.2	51.1	46.2	46.4
	(3,3.5)	(1,1.5)	31.9	29.9	29.3	27.4	42.0	38.8	35.6
(.2,.8)	(3,3.5)	(1,1)	37.2	35.3	32.4	30.2	49.5	48.4	42.0
	(3,4)	(1,1)	90.7	90.1	90.1	89.2	96.2	96.2	97.7
	(3,3)	(1,1.5)	22.5	21.0	22.7	23.0	17.4	15.4	14.1
	(3,3)	(1,2)	64.2	63.7	67.5	65.8	40.0	31.9	31.5
	(3,3.5)	(1,1.5)	25.6	25.4	25.7	23.9	34.9	32.9	30.7

As was previously observed in Tables 2.2–2.3 under the Normal distribution, the tests are more sensitive when a change takes place over a short period of time. For the results in Tables 2.2–2.3, it means that the tests are more powerful when $\kappa \in \{(.5, .5), (.4, .6)\}$ compared to $\kappa = (.2, .8)$. The tests' behavior with respect to the type of distributional change is also consistent with what has been found under the Normal. Namely, the tests based on the *Schilling–Schnurr* weight are better at detecting a change in the mean compared to the other tests, whereas those based on the *Spherical* and *Székely–Rizzo* weights are more powerful when a change occurs in the standard deviation. When both the mean and standard deviation shift, the *Schilling–Schnurr* tests compare favorably with its competitors. Overall, the six newly proposed tests outperform that based on $T_{n,\Sigma}^Q$.

2.6 Illustrations

2.6.1 Historic temperature data in the Canadian Far North

There is growing recognition in the scientific community that the effects of climate change are more pronounced in northern regions. For example, according to Ouranos, an organization located in the province of Quebec (Canada), “In northern regions, warming is more pronounced than elsewhere in the world. While the Earth has warmed by 0.8°C over the last 70 years, Quebec’s average annual temperature has increased by 1.1°C.” In order to investigate this phenomenon, a change-point analysis is conducted on a time series of temperatures at Iqaluit, which is the capital of the Nunavut territory in Canada. The data comprises the monthly maxima recorded at the Iqaluit weather station from January 1946 and July 2008, for a total of 746 observations. Since there is an obvious strong seasonality effect in these data, the analysis is performed separately on the twelve time series corresponding to the different months. These series are shown

in Figure 2.1, where the few missing values (one or two per series) were simply ignored.

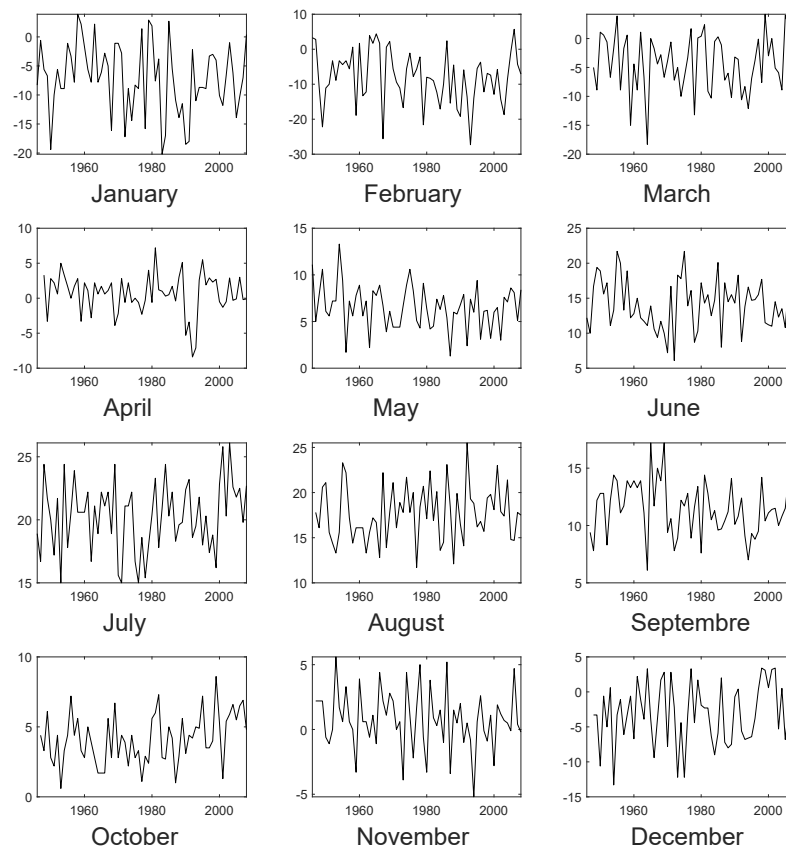


FIGURE 2.1 – Time series of the monthly maximum temperature (in degrees Celcius) measured at the Iqaluit weather station (Nunavut territory, Canada) from January 1946 to December 2008

A look at the sample autocorrelations and partial autocorrelations convince that the serial independence hypothesis is reasonable. The p-values estimated from $B = 500$ multipliers for the six newly proposed tests with weights $\lambda = 3/2$ are in Table 2.5. At the 10% level, the two tests based on the *Schilling–Schnurr* weight detect a gradual change for the month of May; the estimated years between which this change took place are 1948 and 1998. It looks like the variability changes during this period. Another significant change is detected in the October time series between 1968 and 1996. It seems that the average is drifting upwards during this period. To conclude, one may state that these results align with the expectation that the shifts from winter to spring and from summer to fall are particularly sensitive to temperature changes.

Of course, more research will be needed to rule out these effects, for example by looking at data series over longer periods and at other Nordic stations.

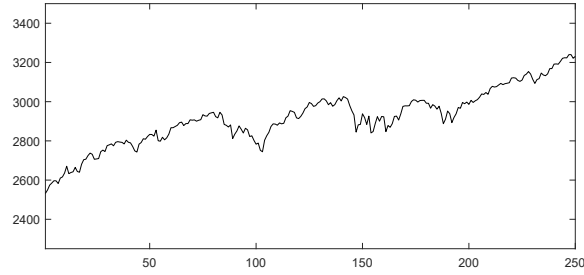
TABLE 2.5 – Estimated p-values (in %) of the six newly proposed tests of change-point detection in the monthly maxima collected at the Iqaluit weather station from January 1948 to July 2008

Period	Spherical		Székely–R.		Schilling–S.		Change-point estimation	
	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$\hat{K}_1$	$\hat{K}_2$
January	27.9	35.0	25.8	33.0	12.4	11.8	—	—
February	13.3	15.3	11.8	14.5	9.1	10.4	—	—
March	25.1	29.1	37.3	41.7	68.9	74.4	—	—
April	13.4	16.0	17.7	21.8	38.6	39.3	—	—
May	32.3	39.6	31.6	36.6	7.3	8.5	1948	1998
June	21.5	25.7	28.5	32.6	47.9	54.4	—	—
July	14.7	16.9	29.6	33.4	30.6	25.9	—	—
August	14.6	17.0	12.3	15.2	19.3	22.4	—	—
September	11.5	14.1	10.0	11.4	12.1	12.4	—	—
October	13.1	13.2	10.6	9.8	5.2	5.0	1968	1996
November	40.6	42.2	30.8	35.1	10.8	12.6	—	—
December	11.2	14.1	12.8	15.2	38.7	41.0	—	—

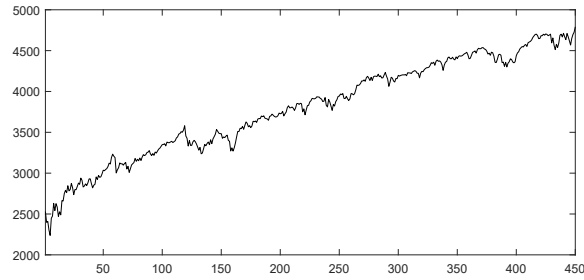
2.6.2 Standard and Poor’s 500

A second illustration of the methodologies is provided on the daily closing prices of the Standard and Poor’s 500 financial index (SP500). The SP500 comprises five hundred United States companies listed on the stock markets, including most of the leading organizations in sectors such as technology, financial services, healthcare, and energy. In the forthcoming analysis, three data sets have been investigated in the light of the newly developed tools of change-point detection, namely a pre-Covid-19, during Covid-19 and post-Covid-19 time series. The first time series is made of $n = 250$ daily measures from January 4, 2019 to December 31, 2019; the second series consists of the $n = 450$ measures from March 17, 2020 to December 27, 2021, *i.e.*, at the height of the pandemic, whereas the third is made of $n = 250$ post-pandemic daily measures from January 3, 2022 to December 29, 2022. This time, the sample autocorrelations and partial autocorrelations indicate that the three series behave as a moving average

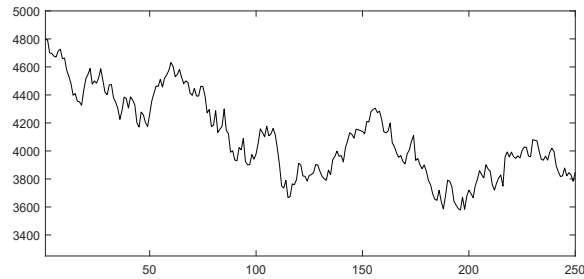
process of order one. In order to remove this serial structure, the lag-1 differentiated series have been considered instead. The three resulting series are shown in Figure 2.2.



(a) Pre-Covid-19



(b) During Covid-19



(c) Post-Covid-19

FIGURE 2.2 – Lag-1 differentiated series based on the Standard and Poor's 500 daily closing prices

The estimated p-values obtained from $B = 500$ multipliers for the tests based on $T_{n,\Sigma}^\nu$ and $T_{n,\vee}^\nu$ when ν is the *Székely-Rizzo* weight with $\lambda = 3/2$ are given in Table 2.6. For each of the three series, no evidence of a stochastic regime change is detected, and

this is true for each of the tests that was performed.

TABLE 2.6 – Estimated p-values (in %) of the tests of change-point detection in the daily closing time of the Standard and Poor 500 financial index based on $T_{n,\Sigma}^\nu$ and $T_{n,\vee}^\nu$ for three Levy measures when $\lambda = 3/2$

Period	Spherical		Székely–R.		Schilling–S.	
	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$
Pre-Covid-19	67.4	72.2	82.0	88.0	78.8	67.2
During Covid-19	34.8	39.2	16.0	16.2	18.8	15.8
Post-Covid-19	70.0	73.0	75.2	76.2	51.8	62.4

2.7 Conclusion

This paper introduces new test statistics for the change-point problem in the case of univariate data sets. The approach that is developed is based on characterizing the stochasticity of these observations via their characteristic function. The two main contributions of the paper are the following :

- (1) Consideration of the gradual change model, as opposed to the usual abrupt-change model, which allows the detection, and by extension the estimation, of changes that occur gradually ;
- (2) The use of Lévy measures in a setting other than to test independence.

The proposed test statistics are weighted U-statistics. In order to use them to test the presence or absence of changes in a time $\mathbf{K} = (K_1, K_2)$ that is undetermined in advance, a uniform weak convergence result was obtained. It was thus established that these statistics converge towards non-degenerate distributions under the null hypothesis. Since these limiting distributions depend on the unknown distribution of the observations, a resampling method based on multiplier bootstrap was proposed for the computation of p-values.

Overall, the proposed tests perform well for relatively small sample sizes. On one side, they hold their nominal level reasonably well under the null hypothesis. On the other side, they show appreciable power under a wide range of gradual change alternatives. However, several challenges related to the introduced methodology merit further investigation.

A significant step forward will be achieved if one can derive the asymptotic distribution of the statistic $S_{\mathbf{K},n}^\nu$ under serial dependence. There is work on weighted U-statistics in the case of time series, for example [22], but the results apply to non-degenerate kernels. On the other hand, there are results for degenerate U-statistics, but not in the weighted case, see for example the work of [32] and [31]. The serial extension in a framework of weighted U-statistics is therefore far from straightforward. Moreover, these results must be obtained in the context of the gradual change model, thus requiring a uniform asymptotic treatment over all possible times of change.

Acknowledgements. The first author acknowledges financial support of the Institut des Sciences Mathématiques de Montréal (ISM). The second author acknowledges financial support by an individual grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

2.8 Proofs

2.8.1 Proof of Lemma 1

From the form of ψ_j in Eq (2.2), one has

$$\begin{aligned}\psi_j - \bar{\psi} &= \omega_j^{\mathbf{K}_0} \Psi_1(t) + (1 - \omega_j^{\mathbf{K}_0}) \Psi_2(t) - \bar{\omega}^{\mathbf{K}_0} \Psi_1(t) - (1 - \bar{\omega}^{\mathbf{K}_0}) \Psi_2(t) \\ &= \Psi_1(t)(\omega_j^{\mathbf{K}_0} - \bar{\omega}^{\mathbf{K}_0}) + \Psi_2(t)(\bar{\omega}^{\mathbf{K}_0} - \omega_j^{\mathbf{K}_0}) \\ &= (\omega_j^{\mathbf{K}_0} - \bar{\omega}^{\mathbf{K}_0}) \{\Psi_1(t) - \Psi_2(t)\}.\end{aligned}$$

Substituting the last expression into the definition of $Q_{\mathbf{K}}$ in Eq (2.3) yields

$$\begin{aligned}Q_{\mathbf{K}}(t) &= \frac{1}{n} \sum_{j=1}^n (\omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) (\omega_j^{\mathbf{K}_0} - \bar{\omega}^{\mathbf{K}_0}) \{\Psi_1(t) - \Psi_2(t)\} \\ &= \{\Psi_1(t) - \Psi_2(t)\} \langle \omega^{\mathbf{K}}, \omega^{\mathbf{K}_0} \rangle.\end{aligned}$$

2.8.2 Proof of Lemma 2

Since the complex conjugate of $Q_{\mathbf{K},n}(t)$ is $Q_{\mathbf{K},n}(-t)$, one has

$$\begin{aligned}S_{\mathbf{K},n}^\nu &= n \int_{\mathbb{R}} Q_{\mathbf{K},n}(t) Q_{\mathbf{K},n}(-t) \nu(dt) \\ &= \frac{1}{n} \sum_{j,j'=1}^n \int_{\mathbb{R}} (\omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) (\omega_{j'}^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) e^{it(Y_j - Y_{j'})} \nu(dt) \\ &= -\frac{1}{n} \sum_{j,j'=1}^n \int_{\mathbb{R}} (\omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) (\omega_{j'}^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) (1 - e^{it(Y_j - Y_{j'})}) \nu(dt) \\ &= -\frac{1}{n} \sum_{j,j'=1}^n (\omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) (\omega_{j'}^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) \beta_\nu(Y_j - Y_{j'}).\end{aligned}$$

Since $\beta_\nu(0) = 0$ and β_ν is symmetric, one can write

$$\begin{aligned} S_{\mathbf{K},n}^\nu &= -\frac{2}{n} \sum_{1 \leq j < j' \leq n} (\omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) (\omega_{j'}^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) \beta_\nu(Y_j - Y_{j'}) \\ &= \frac{1}{n} \sum_{1 \leq j < j' \leq n} (\omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) (\omega_{j'}^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}) \{-2 \beta_\nu(Y_j - Y_{j'})\}. \end{aligned}$$

2.8.3 Proof of Proposition 1

One deduces from Lemma 1 in [44] that for any $\mathbf{b} = (b_1, \dots, b_n) \in \mathbb{R}^n$,

$$n \langle \omega^{\mathbf{K}}, \mathbf{b} \rangle = \frac{1}{K_2 - K_1} \sum_{j=K_1}^{K_2-1} j (\bar{b}_{1:j} - \bar{b}_{1:n}),$$

where $\bar{b}_{1:j}$ is the mean of $b_1, \dots, b_j$. Applying this identity to the definition of $Q_{\mathbf{K},n}$ in Equation (2.4) yields

$$n Q_{\mathbf{K},n}(t) = n \langle \omega^{\mathbf{K}}, \hat{\psi}(t) \rangle = \frac{1}{K_2 - K_1} \sum_{j=K_1}^{K_2-1} j \{ \bar{\psi}_{1:j}(t) - \bar{\psi}_{1:n}(t) \},$$

where $\bar{\psi}_{1:j}(t) = \{\psi_1(t) + \dots + \psi_j(t)\}/j$. Since $\bar{\psi}_{1:n} = \{j \bar{\psi}_{1:j} + (n-j) \bar{\psi}_{j+1:n}\}/n$, and using the fact that $\bar{\psi}_{L_1:L_2}(t) = \int_{\mathbb{R}} e^{ity} F_{L_1:L_2}(\mathrm{d}y)$ for any $1 \leq L_1 \leq L_2 \leq n$,

$$\begin{aligned} n Q_{\mathbf{K},n}(t) &= \frac{1}{K_2 - K_1} \sum_{j=K_1}^{K_2-1} \frac{j(n-j)}{n} \{ \bar{\psi}_{1:j}(t) - \bar{\psi}_{j+1:n}(t) \} \\ &= \frac{1}{K_2 - K_1} \sum_{j=K_1}^{K_2-1} \frac{j(n-j)}{n} \int_{\mathbb{R}} e^{ity} \{ F_{1:j}(\mathrm{d}y) - F_{j+1:n}(\mathrm{d}y) \} \\ &= \sqrt{n} \times \frac{1}{K_2 - K_1} \sum_{j=K_1}^{K_2-1} \int_{\mathbb{R}} e^{ity} \left\{ \frac{j(n-j)}{n^{3/2}} (F_{1:j}(\mathrm{d}y) - F_{j+1:n}(\mathrm{d}y)) \right\} \\ &= \sqrt{n} \int_{\mathbb{R}} e^{ity} \frac{1}{K_2 - K_1} \sum_{j=K_1}^{K_2-1} \mathbb{F}_n \left(\frac{j}{n}, \mathrm{d}y \right) \\ &= \sqrt{n} \int_{\mathbb{R}} e^{ity} \mathbb{L}_{\mathbf{K},n}(\mathrm{d}y). \end{aligned}$$

Hence, $Q_{\mathbf{K},n}(t) = \int_{\mathbb{R}} e^{ity} \mathbb{L}_{\mathbf{K},n}(dy) / \sqrt{n}$. Now, from the definition of $S_{\mathbf{K},n}^\nu$ and using the fact that the complex conjugate of $Q_{\mathbf{K},n}(t)$ is $Q_{\mathbf{K},n}(-t)$,

$$\begin{aligned} S_{\mathbf{K},n}^\nu &= n \int_{\mathbb{R} \setminus \{0\}} |Q_{\mathbf{K},n}(t)|^2 \nu(dt) \\ &= n \int_{\mathbb{R} \setminus \{0\}} Q_{\mathbf{K},n}(t) Q_{\mathbf{K},n}(-t) \nu(dt) \\ &= \int_{\mathbb{R} \setminus \{0\}} \left\{ \int_{\mathbb{R}} e^{ity} \mathbb{L}_{\mathbf{K},n}(dy) \right\} \left\{ \int_{\mathbb{R}} e^{-ity'} \mathbb{L}_{\mathbf{K},n}(dy') \right\} \nu(dt). \end{aligned}$$

Since $\mathbb{L}_{\mathbf{K},n}(y) \rightarrow 0$ as $y \rightarrow \pm\infty$, one can write

$$\begin{aligned} S_{\mathbf{K},n}^\nu &= \int_{\mathbb{R} \setminus \{0\}} \left\{ \int_{\mathbb{R}} (1 - e^{ity}) \mathbb{L}_{\mathbf{K},n}(dy) \right\} \left\{ \int_{\mathbb{R}} (1 - e^{-ity'}) \mathbb{L}_{\mathbf{K},n}(dy') \right\} \nu(dt) \\ &= \int_{\mathbb{R}^2} \left\{ \int_{\mathbb{R} \setminus \{0\}} (1 - e^{ity} - e^{-ity'} + e^{it(y-y')}) \nu(dt) \right\} \mathbb{L}_{\mathbf{K},n}(dy) \mathbb{L}_{\mathbf{K},n}(dy') \\ &= \int_{\mathbb{R}^2} \{\beta_\nu(y) + \beta_\nu(-y') - \beta_\nu(y - y')\} \mathbb{L}_{\mathbf{K},n}(dy) \mathbb{L}_{\mathbf{K},n}(dy'). \end{aligned}$$

Since $\int_{\mathbb{R}} \mathbb{L}_{\mathbf{K},n}(dy) = 0$, one has

$$\int_{\mathbb{R}^2} \beta_\nu(y) \mathbb{L}_{\mathbf{K},n}(dy) \mathbb{L}_{\mathbf{K},n}(dy') = \int_{\mathbb{R}} \beta_\nu(y) \mathbb{L}_{\mathbf{K},n}(dy) \int_{\mathbb{R}} \mathbb{L}_{\mathbf{K},n}(dy') = 0,$$

and similarly, $\int_{\mathbb{R}^2} \beta_\nu(-y') \mathbb{L}_{\mathbf{K},n}(dy) \mathbb{L}_{\mathbf{K},n}(dy') = 0$. One can then conclude that

$$S_{\mathbf{K},n}^\nu = - \int_{\mathbb{R}^2} \beta_\nu(y - y') \mathbb{L}_{\mathbf{K},n}(dy) \mathbb{L}_{\mathbf{K},n}(dy').$$

2.8.4 Proof of Proposition 2

The proof is based on the following result in [35] for the convergence in distribution of weighted U-statistics. The result is stated next in the special case of a kernel of degree two.

Theorem 1 (Theorem 4, [35]). For a symmetric kernel $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ and a sequence

$\{a_j\}_{j \in \mathcal{N}}$, consider the weighted U-statistic

$$U_n = \frac{1}{n} \sum_{1 \leq j < j' \leq n} a_j a_{j'} h(Y_j, Y_{j'}),$$

where $Y_1, \dots, Y_n$ are i.i.d. F . Suppose that $E\{h^2(Y_1, Y_2)\} < \infty$ and that the sequence of weights $\{a_j\}_{j \in \mathcal{N}}$ is bounded and such that the following limit exists :

$$A_1 = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n a_j^2 > 0.$$

Then, U_n converges in distribution to the random variable

$$\begin{aligned} U &= \left(\frac{A_2^2 - A_1}{2} \right) E\{h(Y_1, Y_2)\} + A_2 \sqrt{A_1} \int_{\mathbb{R}} E\{h(y_1, Y_2)\} W_F(dy_1) \\ &\quad + \frac{A_1}{2} \int_{\mathbb{R}^2} h(y_1, y_2) W_F(dy_1) W_F(dy_2), \end{aligned}$$

where W_F is a F -Wiener process and

$$A_2 = \lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} \sum_{j=1}^n a_j.$$

Here, $h(y_1, y_2) = -2\beta_\nu(y_1 - y_2)$ and $a_j = \omega_j^{\lfloor n\kappa \rfloor} - \bar{\omega}^{\lfloor n\kappa \rfloor}$, so that $A_2 = 0$ and

$$A_1 = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \left(\omega_j^{\lfloor n\kappa \rfloor} - \bar{\omega}^{\lfloor n\kappa \rfloor} \right)^2 = \lim_{n \rightarrow \infty} \langle \omega(\lfloor n\kappa \rfloor), \omega(\lfloor n\kappa \rfloor) \rangle = \Lambda(\kappa, \kappa).$$

It follows that the asymptotic distribution of $S_{\lfloor n\kappa \rfloor, n}^\nu$ is that of the random variable

$$\begin{aligned} \mathbb{S}_\kappa^\nu &= -\frac{\Lambda(\kappa, \kappa)}{2} (-2E\{\beta_\nu(Y_1 - Y_2)\}) + \frac{\Lambda(\kappa, \kappa)}{2} \int_{\mathbb{R}^2} -2\beta_\nu(y_1, y_2) W_F(dy_1) W_F(dy_2) \\ &= \Lambda(\kappa, \kappa) \int_{\mathbb{R}^2} \beta(y_1 - y_2) F(dy_1) F(dy_2) - \Lambda(\kappa, \kappa) \int_{\mathbb{R}^2} \beta(y_1 - y_2) W_F(dy_1) W_F(dy_2) \\ &= \Lambda(\kappa, \kappa) \int_{\mathbb{R}^2} \beta(y_1 - y_2) \{F(dy_1) F(dy_2) - W_F(dy_1) W_F(dy_2)\}. \end{aligned}$$

2.8.5 Proof of Proposition 3

Theorem 1.5.4 of [57] states that a net $X_\alpha : \Omega_\alpha \rightarrow \ell^\infty(T)$ converges weakly to a tight limit X if and only if (1) there is a stochastic process X such that $(X_\alpha(t_1), \dots, X_\alpha(t_N))$ converges weakly to $(X(t_1), \dots, X(t_N))$ for every finite set of points $t_1, \dots, t_N \in T$ and (2) X_α is asymptotically tight. When (1) and (2) are satisfied, one then writes $X_\alpha \rightsquigarrow X$.

In the current context, $X_\alpha := S_{[n\kappa],n}^\nu$ and $T := \mathcal{C}$. For arbitrary $\kappa_1, \dots, \kappa_N \in \mathcal{C}$, the convergence of the finite-dimensional distribution

$$(S_{[n\kappa_1],n}^\nu, \dots, S_{[n\kappa_N],n}^\nu) \rightsquigarrow (S_{\kappa_1}^\nu, \dots, S_{\kappa_N}^\nu)$$

is a consequence of Proposition 2. Now according to Theorem 1.5.7 of [57], $S_{[n\kappa],n}^\nu$ is asymptotically tight in $\mathcal{C}$ if and only if :

(B_1) $S_{[n\kappa],n}^\nu$ is asymptotically tight in $\mathbb{R}$ for each $\kappa \in \mathcal{C}$, which means that for every $\epsilon > 0$, one can find $M_\epsilon \in (0, \infty)$ such that

$$\liminf_{n \rightarrow \infty} \mathbb{P}(S_{[n\kappa],n}^\nu \leq M_\epsilon) \geq 1 - \epsilon.$$

(B_2) For all $\eta, \epsilon > 0$, one can find $\delta > 0$ such that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sup_{(\kappa, \kappa') \in \mathcal{C}_\delta} |S_{[n\kappa],n}^\nu - S_{[n\kappa'],n}^\nu| > \epsilon \right) \leq \eta,$$

where $\mathcal{C}_\delta = \{(\kappa, \kappa') : \kappa, \kappa' \in \mathcal{C}, |\kappa - \kappa'| \leq \delta\}$.

Proposition 2 ensures that (B_1) holds. To show that (B_2) holds as well, let $\alpha_j^\kappa =$

$\omega_j^{\lfloor n\kappa \rfloor} - \bar{\omega}^{\lfloor n\kappa \rfloor}$, so that one can write

$$S_{\lfloor n\kappa \rfloor, n}^\nu = -\frac{2}{n} \sum_{j < j'} \alpha_j^\kappa \alpha_{j'}^\kappa \beta_\nu (Y_j - Y_{j'}).$$

One then has

$$\begin{aligned} |S_{\lfloor n\kappa \rfloor, n}^\nu - S_{\lfloor n\kappa' \rfloor, n}^\nu| &= \frac{2}{n} \left| \sum_{j < j'} \beta_\nu (Y_j - Y_{j'}) \left(\alpha_j^\kappa \alpha_{j'}^\kappa - \alpha_j^{\kappa'} \alpha_{j'}^{\kappa'} \right) \right| \\ &\leq \frac{2}{n} \sum_{j < j'} \beta_\nu (Y_j - Y_{j'}) \left| \alpha_j^\kappa \alpha_{j'}^\kappa - \alpha_j^{\kappa'} \alpha_{j'}^{\kappa'} \right|. \end{aligned}$$

Upon noting that $\alpha_j^\kappa \alpha_{j'}^\kappa - \alpha_j^{\kappa'} \alpha_{j'}^{\kappa'} = \alpha_j^\kappa (\alpha_{j'}^\kappa - \alpha_{j'}^{\kappa'}) + \alpha_{j'}^{\kappa'} (\alpha_j^\kappa - \alpha_j^{\kappa'})$ and since the sequence $\{\alpha_j^\kappa\}$ is bounded,

$$\sup_{(\kappa, \kappa') \in \mathcal{C}_\delta} \left| \alpha_j^\kappa \alpha_{j'}^\kappa - \alpha_j^{\kappa'} \alpha_{j'}^{\kappa'} \right| \leq L \sup_{(\kappa, \kappa') \in \mathcal{C}_\delta} \left| \alpha_j^\kappa - \alpha_j^{\kappa'} \right| + L \sup_{(\kappa, \kappa') \in \mathcal{C}_\delta} \left| \alpha_{j'}^{\kappa'} - \alpha_{j'}^\kappa \right|.$$

Note that the weights can be written as

$$\omega_j^{\lfloor n\kappa \rfloor} = \mathbb{I}(j \leq \lfloor n\kappa_2 \rfloor) \min \left(1, \frac{\lfloor n\kappa_2 \rfloor - j}{\lfloor n\kappa_2 \rfloor - \lfloor n\kappa_1 \rfloor} \right),$$

whereas it can be shown easily that $\bar{\omega}^{\lfloor n\kappa \rfloor} = \{\lfloor n\kappa_1 \rfloor + \lfloor n\kappa_2 \rfloor - 1\} / (2n)$. Since $|\lfloor nr \rfloor / n - r| < 1/n$, $r \in \mathbb{R}$,

$$\begin{aligned} \left| \bar{\omega}^{\lfloor n\kappa \rfloor} - \bar{\omega}^{\lfloor n\kappa' \rfloor} \right| &= \left| \frac{\lfloor n\kappa_1 \rfloor - \lfloor n\kappa'_1 \rfloor}{2n} + \frac{\lfloor n\kappa_2 \rfloor - \lfloor n\kappa'_2 \rfloor}{2n} \right| \\ &\leq \frac{|\kappa_1 - \kappa'_1|}{2} + \frac{|\kappa_2 - \kappa'_2|}{2} + o(1/n) \\ &\leq \delta + o(1/n). \end{aligned}$$

Now, $\omega_j^{\lfloor n\kappa \rfloor} - \omega_j^{\lfloor n\kappa' \rfloor} = 0$ if $j > \max\{\lfloor n\kappa_2 \rfloor, \lfloor n\kappa'_2 \rfloor\}$.

If $j \leq \min\{\lfloor n\kappa_2 \rfloor, \lfloor n\kappa'_2 \rfloor\}$, there is three cases to consider :

$$1) j \leq \min \{ \lfloor n\kappa_1 \rfloor, \lfloor n\kappa'_1 \rfloor \}$$

$$2) j \in [\lfloor n\kappa_1 \rfloor, \lfloor n\kappa'_1 \rfloor]$$

$$3) j \geq \max \{ \lfloor n\kappa_1 \rfloor, \lfloor n\kappa'_1 \rfloor \}$$

In the first case, the definition of ω_j yields $\omega_j^{\lfloor n\kappa \rfloor} - \omega_j^{\lfloor n\kappa' \rfloor} = 0$. For 2), without loss of generality, let's assume $\kappa'_1 \geq \kappa_1$ which implies

$$\omega_j^{\lfloor n\kappa \rfloor} = \frac{\lfloor n\kappa_2 \rfloor - j}{\lfloor n\kappa_2 \rfloor - \lfloor n\kappa_1 \rfloor} \quad \text{and} \quad \omega_j^{\lfloor n\kappa' \rfloor} = 1$$

$$\begin{aligned} \left| \omega_j^{\lfloor n\kappa \rfloor} - 1 \right| &= \left| \frac{\lfloor n\kappa_1 \rfloor - j}{\lfloor n\kappa_2 \rfloor - \lfloor n\kappa_1 \rfloor} \right| \\ &\leq \left| \frac{\lfloor n\kappa_1 \rfloor - \lfloor n\kappa'_1 \rfloor}{\lfloor n\kappa_2 \rfloor - \lfloor n\kappa_1 \rfloor} \right| \\ &\leq n |\kappa_1 - \kappa'_1| \\ &= n \delta \end{aligned}$$

In the third case, let's assume $\min \{ \lfloor n\kappa_2 \rfloor, \lfloor n\kappa'_2 \rfloor \} = \lfloor n\kappa'_2 \rfloor$

$$\begin{aligned} \left| \omega_j^{\lfloor n\kappa \rfloor} - \omega_j^{\lfloor n\kappa' \rfloor} \right| &= \left| \frac{\lfloor n\kappa_2 \rfloor - j}{\lfloor n\kappa_2 \rfloor - \lfloor n\kappa_1 \rfloor} - \frac{\lfloor n\kappa'_2 \rfloor - j}{\lfloor n\kappa'_2 \rfloor - \lfloor n\kappa'_1 \rfloor} \right| \\ &\leq \left| \frac{\lfloor n\kappa_2 \rfloor - \lfloor n\kappa'_2 \rfloor}{\lfloor n\kappa_2 \rfloor - \lfloor n\kappa_1 \rfloor} \right| \\ &\leq n |\kappa_2 - \kappa'_2| \\ &= n \delta \end{aligned}$$

It follows that,

$$\sup_{(\kappa, \kappa') \in \mathcal{C}_\delta} \left| \alpha_j^\kappa \alpha_{j'}^\kappa - \alpha_j^{\kappa'} \alpha_{j'}^{\kappa'} \right| \leq \frac{2LL'\delta}{n}.$$

As a consequence,

$$\begin{aligned}
\mathbb{P} \left(\sup_{(\kappa, \kappa') \in \mathcal{C}_\delta} |S_{[n\kappa], n}^\nu - S_{[n\kappa'], n}^\nu| > \epsilon \right) &\leq \mathbb{P} \left(\frac{2}{n} \sum_{j < j'} \beta_\nu(Y_j - Y_{j'}) \times \frac{2LL'\delta}{n} > \epsilon \right) \\
&= \mathbb{P} \left(\frac{4}{n^2} \sum_{j < j'} \beta_\nu(Y_j - Y_{j'}) > \frac{\epsilon}{LL'\delta} \right) \\
&= \mathbb{P} \left(U_n > \frac{\epsilon}{2LL'\delta} \times \frac{n}{n-1} \right),
\end{aligned}$$

where U_n^ν is a (non-weighted) U-statistic with kernel β_ν , namely

$$U_n = \binom{n}{2}^{-1} \sum_{j < j'} \beta_\nu(Y_j - Y_{j'}).$$

2.8.6 Proof of Proposition 4

For simplicity, let $W_j^{\mathbf{K}} = \omega_j^{\mathbf{K}} - \bar{\omega}^{\mathbf{K}}$, and observe that

$$\begin{aligned}
S_{\mathbf{K}, n}^\nu &= \frac{1}{n} \sum_{j, j'=1}^n W_j^{\mathbf{K}} W_{j'}^{\mathbf{K}} \int_{\mathbb{R} \setminus \{0\}} e^{it(Y_j - Y_{j'})} \nu(dt) \\
&= \frac{1}{n} \sum_{j, j'=1}^n W_j^{\mathbf{K}} W_{j'}^{\mathbf{K}} \int_{\mathbb{R} \setminus \{0\}} \{e^{it(Y_j - Y_{j'})} - \psi_j(t) \psi_{j'}(-t)\} \nu(dt) \\
&\quad + n \int_{\mathbb{R} \setminus \{0\}} \left\{ \frac{1}{n} \sum_{j=1}^n W_j^{\mathbf{K}} \psi_j(t) \right\} \left\{ \frac{1}{n} \sum_{j'=1}^n W_{j'}^{\mathbf{K}} \psi_{j'}(t) \right\} \nu(dt) \\
&= \tilde{S}_{\mathbf{K}, n}^\nu + n \int_{\mathbb{R} \setminus \{0\}} \langle \omega^{\mathbf{K}}, \psi(t) \rangle \langle \omega^{\mathbf{K}}, \psi(-t) \rangle \nu(dt),
\end{aligned}$$

where $\psi = (\psi_1, \dots, \psi_n)$ and

$$\tilde{S}_{\mathbf{K}, n}^\nu = \frac{1}{n} \sum_{j, j'=1}^n W_j^{\mathbf{K}} W_{j'}^{\mathbf{K}} \int_{\mathbb{R} \setminus \{0\}} \{e^{it(Y_j - Y_{j'})} - \psi_j(t) \psi_{j'}(-t)\} \nu(dt).$$

Following for instance [56], the projection of the statistic $\tilde{S}_{\mathbf{K},n}^\nu$ is

$$\mathbb{E}(\tilde{S}_{\mathbf{K},n}^\nu | Y_{j'}) = \frac{1}{n} \sum_{j,j'=1}^n W_j^{\mathbf{K}} W_{j'}^{\mathbf{K}} \int_{\mathbb{R} \setminus \{0\}} \psi_j(t) \{e^{-itY_{j'}} - \psi_{j'}(-t)\} \nu(dt).$$

Since the latter is non-degenerate, it implies that $S_{\mathbf{K},n}^\nu / \sqrt{n}$ converges in distribution to a non-degenerate normal limit S_κ^ν . Now invoking Lemma 1,

$$\begin{aligned} \langle \omega^{\mathbf{K}}, \psi(t) \rangle \langle \omega^{\mathbf{K}}, \psi(-t) \rangle &= \{ \langle \omega^{\mathbf{K}}, \omega^{\mathbf{K}_0} \rangle \}^2 \{ \Psi_1(t) - \Psi_2(t) \} \{ \Psi_1(-t) - \Psi_2(-t) \} \\ &= \{ \langle \omega^{\mathbf{K}}, \omega^{\mathbf{K}_0} \rangle \}^2 | \Psi_1(t) - \Psi_2(t) |^2. \end{aligned}$$

Consequently, as $n \rightarrow \infty$,

$$\begin{aligned} \frac{S_{\mathbf{K},n}^\nu}{n} &= \frac{1}{\sqrt{n}} \left(\frac{\tilde{S}_{\mathbf{K},n}^\nu}{\sqrt{n}} \right) + \{ \langle \omega^{\mathbf{K}}, \omega^{\mathbf{K}_0} \rangle \}^2 \int_{\mathbb{R} \setminus \{0\}} | \Psi_1(t) - \Psi_2(t) |^2 \nu(dt) \\ &\xrightarrow{\mathbb{P}} \{ \Lambda(\kappa, \kappa) \}^2 \int_{\mathbb{R} \setminus \{0\}} | \Psi_1(t) - \Psi_2(t) |^2 \nu(dt) > 0. \end{aligned}$$

It follows that $S_{\mathbf{K},n}^\nu \rightarrow \infty$ in probability, and it may be conclude that the tests based on $T_{n,\Sigma}^\nu$ and $T_{n,\vee}^\nu$ are consistent.

2.8.7 Proof of Lemma 3

From the definition of the multiplier version of the sequential empirical process in Equation (2.12) and the definition of $\delta_{1:j}$, one can write

$$\sqrt{n} \widehat{\mathbb{F}}_n \left(\frac{j}{n}, y \right) = \delta_{1:j} \mathcal{I}_{1:j}^\top(y) - \frac{j}{n} \delta_{1:n} \mathcal{I}_{1:n}^\top(y),$$

where $\mathcal{I}_{1:j}^\top(y) = (\mathbb{I}(Y_1 \leq y), \dots, \mathbb{I}(Y_j \leq y))$. With this notation and using Equation (2.13), one can write

$$\widehat{\mathbb{L}}_{\mathbf{K},n}(y) = \frac{1}{\sqrt{n}(K_2 - K_1)} \sum_{j=K_1}^{K_2-1} \left\{ \boldsymbol{\delta}_{1:j} \mathcal{I}_{1:j}^\top(y) - \frac{j}{n} \boldsymbol{\delta}_{1:n} \mathcal{I}_{1:n}^\top(y) \right\}.$$

The multiplier version of $\widehat{Q}_{\mathbf{K},n}(t)$ can then be expressed as

$$\begin{aligned} \widehat{Q}_{\mathbf{K},n}(t) &= \frac{-it}{\sqrt{n}} \int_{\mathbb{R}} e^{ity} \widehat{\mathbb{L}}_{\mathbf{K},n}(y) dy \\ &= \frac{-it}{n} \int_{\mathbb{R}} e^{ity} \left[\frac{1}{K_2 - K_1} \sum_{j=K_1}^{K_2-1} \left\{ \boldsymbol{\delta}_{1:j} \mathcal{I}_{1:j}^\top(y) - \frac{j}{n} \boldsymbol{\delta}_{1:n} \mathcal{I}_{1:n}^\top(y) \right\} \right] dy \\ &= \frac{1}{n(K_2 - K_1)} \sum_{j=K_1}^{K_2-1} \int_{\mathbb{R}} (-it) e^{ity} \left\{ \boldsymbol{\delta}_{1:j} \mathcal{I}_{1:j}^\top(y) - \frac{j}{n} \boldsymbol{\delta}_{1:n} \mathcal{I}_{1:n}^\top(y) \right\} dy. \end{aligned}$$

Since

$$\int_{\mathbb{R}} (-it) e^{ity} \mathbb{I}(Y_j \leq y) dy = \lim_{N \rightarrow \infty} \int_{Y_j}^N (-it) e^{ity} dy = e^{itY_j} - \lim_{N \rightarrow \infty} e^{itN},$$

one can deduce that

$$\int_{\mathbb{R}} (-it) e^{ity} \boldsymbol{\delta}_{1:j} \mathcal{I}_{1:j}^\top(y) dy = \boldsymbol{\delta}_{1:j} \Psi_{1:j}(t),$$

where $\Psi_{1:j}(t) = (e^{itY_1}, \dots, e^{itY_j})$. Hence,

$$\widehat{Q}_{\mathbf{K},n}(t) = \frac{1}{n(K_2 - K_1)} \sum_{j=K_1}^{K_2-1} \left\{ \boldsymbol{\delta}_{1:j} \Psi_{1:j}^\top(t) - \frac{j}{n} \boldsymbol{\delta}_{1:n} \Psi_{1:n}^\top(t) \right\}.$$

Now since the complex conjugate of $\widehat{Q}_{\mathbf{K},n}(t)$ is $\widehat{Q}_{\mathbf{K},n}(-t)$,

$$\begin{aligned}
n^2(K_2 - K_1)^2 \left| \widehat{Q}_{\mathbf{K},n}(t) \right|^2 &= n^2(K_2 - K_1)^2 \widehat{Q}_{\mathbf{K},n}(t) \widehat{Q}_{\mathbf{K},n}(-t) \\
&= \sum_{j,k=K_1}^{K_2-1} \left\{ \boldsymbol{\delta}_{1:j} \Psi_{1:j}^\top(t) - \frac{j}{n} \boldsymbol{\delta}_{1:n} \Psi_{1:n}^\top(t) \right\} \\
&\quad \times \left\{ \boldsymbol{\delta}_{1:k} \Psi_{1:k}^\top(t) - \frac{k}{n} \boldsymbol{\delta}_{1:n} \Psi_{1:n}^\top(t) \right\} \\
&= \sum_{j,k=K_1}^{K_2-1} \left\{ \boldsymbol{\delta}_{1:j} \Psi_{1:j}^\top(t) \Psi_{1:k}(-t) \boldsymbol{\delta}_{1:k}^\top \right. \\
&\quad - \frac{k}{n} \boldsymbol{\delta}_{1:j} \Psi_{1:j}^\top(t) \Psi_{1:n}(-t) \boldsymbol{\delta}_{1:n}^\top \\
&\quad - \frac{j}{n} \boldsymbol{\delta}_{1:n} \Psi_{1:n}^\top(t) \Psi_{1:k}(-t) \boldsymbol{\delta}_{1:k}^\top \\
&\quad \left. + \frac{jk}{n^2} \boldsymbol{\delta}_{1:n} \Psi_{1:n}^\top(t) \Psi_{1:n}(-t) \boldsymbol{\delta}_{1:n}^\top \right\}.
\end{aligned}$$

Upon noting that $\mathfrak{B}_{1:j,1:k}^\nu = \int_{\mathbb{R}} \Psi_{1:j}^\top(t) \Psi_{1:k}(-t) \nu(dt)$, one concludes that

$$\begin{aligned}
\widehat{S}_{\mathbf{K},\nu} &= \int_{\mathbb{R}} \left| \widehat{Q}_{\mathbf{K},n}(t) \right|^2 \nu(dt) \\
&= \frac{1}{n^2(K_2 - K_1)^2} \sum_{j,k=K_1}^{K_2-1} \left\{ \boldsymbol{\delta}_{1:j} \mathfrak{B}_{1:j,1:k} \boldsymbol{\delta}_{1:k}^\top - \frac{k}{n} \mathfrak{B}_{1:j,1:n} \boldsymbol{\delta}_{1:n}^\top \right. \\
&\quad \left. - \frac{j}{n} \mathfrak{B}_{1:n,1:k} \boldsymbol{\delta}_{1:k}^\top + \frac{jk}{n^2} \mathfrak{B}_{1:n,1:n} \boldsymbol{\delta}_{1:n}^\top \right\}.
\end{aligned}$$

Conclusion

Plusieurs thèmes ont été abordés dans ce mémoire. Tout d'abord, une revue des principales avancées dans la théorie de la détection de points de changement a été présentée. Afin de mettre en lumière les motivations derrière l'article introduit au chapitre 2, un résumé concis des idées proposées dans l'article de Quessy [44] a été exposé. Par la suite, quelques notions préalables à la Section 2, telles que la fonction caractéristique, les processus stochastiques et les techniques de rééchantillonnage, ont été illustrées. Enfin, au deuxième chapitre, un article soumis à la revue *Econometrics and Statistics* est inséré, détaillant une méthodologie innovante pour la détection de ruptures graduelles dans le cas de processus stochastiques univariés.

Ce travail a étendu le champ de recherche actuel sous plusieurs aspects. D'abord, une famille de tests statistiques inédite basée sur la fonction caractéristique et les mesures de Lévy a été introduite. En outre, l'utilisation de ces outils a permis l'identification de point de changement sous le modèle graduel de Lombard [34]. Finalement, ce contexte plus général fut favorable au développement de tests statistiques non paramétriques par la mise en œuvre de la méthode du multiplicateur.

Dans l'optique de poursuivre ce travail, une approche plus large aurait pu être considérée. En effet, une extension de cette méthodologie aux séries chronologiques multivariées pourrait être envisagée dans de futurs travaux. De plus, une élaboration de tests pour la détection de multiples ruptures permettrait de généraliser la théorie suggérée dans ce mémoire et rendrait possible l'étude de jeux de données plus complexes.

Bibliographie

- [1] M. M. Ali and C. Giaccotto. The identical distribution hypothesis for stock market prices. location-and scale-shift alternatives. *Journal of the American Statistical Association*, 77(377) :19–28, 1982.
- [2] A. Aue and J. Steinebach. A note on estimating the change-point of a gradually changing stochastic process. *Statist. Probab. Lett.*, 56(2) :177–191, 2002. ISSN 0167-7152.
- [3] T. Bahraoui and J.-F. Quessy. Tests of multivariate copula exchangeability based on Lévy measures. *Scand. J. Stat.*, 49(3) :1215–1243, 2022. ISSN 0303-6898.
- [4] G. K. Bhattacharyya and R. A. Johnson. Nonparametric tests for shift at an unknown time point. *Ann. Math. Statist.*, 39 :1731–1743, 1968. ISSN 0003-4851.
- [5] A. F. Bissell. The performance of control charts and cusums under linear trend. *Applied Statistics*, pages 145–151, 1984.
- [6] B. Böttcher, M. Keller-Ressel, and R. L. Schilling. Detecting independence of random vectors- generalized distance covariance and Gaussian covariance. *Mod. Stoch. Theory Appl.*, 5(3) :353–383, 2018. ISSN 2351-6046.
- [7] J.-C. Breton. Processus stochastique. *Université de Rennes1*, 2013.
- [8] B. E. Brodsky and B. S. Darkhovsky. *Nonparametric methods in change-point problems*, volume 243 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1993. ISBN 0-7923-2122-7.

- [9] F. Chen, S. G. Meintanis, and L. Zhu. On some characterizations and multi-dimensional criteria for testing homogeneity, symmetry and independence. *J. Multivariate Anal.*, 173 :125–144, 2019. ISSN 0047-259X.
- [10] J. Chen and A. K. Gupta. *Parametric statistical change point analysis*. Birkhäuser Boston, Inc., Boston, MA, 2000. ISBN 0-8176-4169-6.
- [11] H. Chernoff and S. Zacks. Estimating the current mean of a normal distribution which is subjected to changes in time. *Ann. Math. Statist.*, 35 :999–1018, 1964.
- [12] H. Dehling, K. Vuk, and M. Wendler. Change-point detection based on weighted two-sample U-statistics. *Electron. J. Stat.*, 16(1) :862–891, 2022.
- [13] P. Dubey and H.-G. Müller. Fréchet change-point detection. *Ann. Statist.*, 48 (6) :3312–3335, 2020.
- [14] B. Efron. Bootstrap methods : another look at the jackknife. *Ann. Statist.*, 7 (1) :1–26, 1979.
- [15] A. Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 322(8) :549–560, 1905.
- [16] S. B. Fotopoulos, V. K. Jandhyala, and L. Tan. Asymptotic study of the change-point mle in multivariate Gaussian families under contiguous alternatives. *J. Statist. Plann. Inference*, 139(3) :1190–1202, 2009.
- [17] F. Gan. Cusum control charts under linear drift. *The Statistician*, pages 71–84, 1992.
- [18] R. Giacomini, D. N. Politis, and H. White. A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Econometric Theory*, 29(3) :567–589, 2013. ISSN 0266-4666.
- [19] E. Gombay and L. Horváth. An application of U -statistics to change-point analysis. *Acta Sci. Math. (Szeged)*, 60(1-2) :345–357, 1995. ISSN 0001-6969.

- [20] E. Gombay and L. Horv  th. Change-points and bootstrap. *Environmetrics*, 10 (6) :725–736, 1999.
- [21] D. V. Hinkley. Inference about the change-point in a sequence of random variables. *Biometrika*, 57 :1–17, 1970.
- [22] T. Hsing and W. B. Wu. On weighted U -statistics for stationary processes. *Ann. Probab.*, 32(2) :1600–1631, 2004. ISSN 0091-1798.
- [23] M. Hu  skov  . Gradual changes versus abrupt changes. *J. Statist. Plann. Inference*, 76(1-2) :109–125, 1999. ISSN 0378-3758.
- [24] M. Hu  skov   and S. G. Meintanis. Change point analysis based on empirical characteristic functions. *Metrika*, 63(2) :145–168, 2006. ISSN 0026-1335.
- [25] M. Hu  skov   and S. G. Meintanis. Change-point analysis based on empirical characteristic functions of ranks. *Sequential Anal.*, 25(4) :421–436, 2006. ISSN 0747-4946.
- [26] N. Jacob. *Pseudo differential operators and Markov processes. Vol. I*. Imperial College Press, London, 2001. ISBN 1-86094-293-8. Fourier analysis and semi-groups.
- [27] D. Jaru  skov   and J. Antoch. Changepoint analysis of Klementinum temperature series. *Environmetrics*, 31(1) :e2570, 13, 2020. ISSN 1180-4009.
- [28] A. Kolmogorov. Sulla determinazione empirica di una legge di distribuzione. *Giornale dell’Istituto Italiano degli Attuari*, 4 :83–91, 1933.
- [29] M. Kosorok. *Introduction to empirical processes and semiparametric inference*. Springer, New York, 2008.
- [30] F. C. Lemyre and J.-F. Quessy. Kendall’s tau-based inference for gradually changing dependence structures. *Statist. Papers*, 65(4) :2033–2075, 2024. ISSN 0932-5026.

- [31] A. Leucht. Degenerate U - and V -statistics under weak dependence- asymptotic theory and bootstrap consistency. *Bernoulli*, 18(2) :552–585, 2012. ISSN 1350-7265.
- [32] A. Leucht and M. H. Neumann. Consistency of general bootstrap methods for degenerate U -type and V -type statistics. *J. Multivariate Anal.*, 100(8) :1622–1633, 2009. ISSN 0047-259X.
- [33] F. Lombard. Asymptotic distributions of rank statistics in the change-point problem. *South African Statist. J.*, 17(1) :83–105, 1983. ISSN 0038-271X.
- [34] F. Lombard. Rank tests for changepoint problems. *Biometrika*, 74(3) :615–624, 1987.
- [35] P. Major. Asymptotic distributions for weighted U -statistics. *Ann. Probab.*, 22(3) :1514–1535, 1994. ISSN 0091-1798.
- [36] M. Mohr and N. Neumeyer. Nonparametric volatility change detection. *Scand. J. Stat.*, 48(2) :529–548, 2021.
- [37] E. S. Page. Continuous inspection schemes. *Biometrika*, 41 :100–115, 1954.
- [38] E. S. Page. A test for a change in a parameter occurring at an unknown point. *Biometrika*, 42 :523–527, 1955. ISSN 0006-3444.
- [39] Perreault, Luc and Bernier, J. and Bobée, Bernard and Parent, Éric. Bayesian change-point analysis in hydrometeorological time series. Part 1. The normal model revisited. *Journal of Hydrology*, 235(3–4) :221–241, 2000.
- [40] A. Pettitt. A non-parametric approach to the change point problem. *Applied Statistics*, 28(2) :126–135, 1979.
- [41] W. Philipp and L. Pinzur. Almost sure approximation theorems for the multivariate empirical process. *Z. Wahrsch. Verw. Gebiete*, 54(1) :1–13, 1980. ISSN 0044-3719. doi : 10.1007/BF00535346.

- [42] M. H. Quenouille. Approximate tests of correlation in time-series. *Proc. Cambridge Philos. Soc.*, 45 :483–484, 1949.
- [43] M. H. Quenouille. Notes on bias in estimation. *Biometrika*, 43 :353–360, 1956.
- [44] J.-F. Quessy. Consistent nonparametric tests for detecting gradual changes in the marginals and the copula of multivariate time series. *Statist. Papers*, 60(3) : 367–396, 2019. ISSN 0932-5026.
- [45] J.-F. Quessy. A Székely-Rizzo inequality for testing general copula homogeneity hypotheses. *J. Multivariate Anal.*, 186 :Paper No. 104815, 13, 2021. ISSN 0047-259X.
- [46] J.-F. Quessy. Gradual change-point analysis based on Spearman matrices for multivariate time series. *Ann. Inst. Statist. Math.*, 76(3) :423–446, 2024. ISSN 0020-3157.
- [47] J.-F. Quessy, A.-C. Favre, M. Saïd, and M. Champagne. Statistical inference in lombard’s smooth-change model. *Environmetrics*, 22(7) :882–893, 2011.
- [48] K. R. Ryberg, G. A. Hodgkins, and R. W. Dudley. Change points in annual peak streamflows : Method comparisons and historical change points in the united states. *Journal of Hydrology*, 583 :124307, 2020.
- [49] E. Schechtman. A nonparametric test for detecting changes in location. *Communications in Statistics-Theory and Methods*, 11(13) :1475–1482, 1982.
- [50] R. L. Schilling and A. Schnurr. The symbol associated with the solution of a stochastic differential equation. *Electron. J. Probab.*, 15 :1369–1393, 2010.
- [51] J. Střeščík. Long-term climatic changes observed in prague-klementinum. *Contributions to geophysics and geodesy*, 30 :169–196, 2000.
- [52] G. J. Székely and M. L. Rizzo. Energy statistics- a class of statistics based on distances. *J. Statist. Plann. Inference*, 143(8) :1249–1272, 2013. ISSN 0378-3758.

- [53] G. J. Székely, M. L. Rizzo, and N. K. Bakirov. Measuring and testing dependence by correlation of distances. *Ann. Statist.*, 35(6) :2769–2794, 2007. ISSN 0090-5364.
- [54] G. J. Székely and T. F. Mori. A characteristic measure of asymmetry and its application for testing diagonal symmetry. *Comm. Statist. Theory Methods*, 30 (8-9) :1633–1639, 2001. ISSN 0361-0926. Statistics- reflections on the past and visions for the future (San Antonio, TX, 2000).
- [55] J. W. Tukey. Bias and confidence in not quite large samples. *Annals of Mathematical Statistics*, 29, 1958.
- [56] A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. ISBN 0-521-49603-9 ; 0-521-78450-6.
- [57] A. W. van der Vaart and J. A. Wellner. *Weak convergence and empirical processes*. Springer Series in Statistics. Springer-Verlag, New York, 1996. ISBN 0-387-94640-3. With applications to statistics.
- [58] M. Vogt and H. Dette. Detecting gradual changes in locally stationary processes. *Ann. Statist.*, 43(2) :713–740, 2015. ISSN 0090-5364.
- [59] N. Wiener. Differential-space. *J. Math. and Phys.*, 2 :131–174, 1923.
- [60] Q. Zhang. A nonparametric test for comparing survival functions based on restricted distance correlation. *Depend. Model.*, 11(1) :Paper No. 20230108, 15, 2023.
- [61] C. Zou, Y. Liu, P. Qin, and Z. Wang. Empirical likelihood ratio test for the change-point problem. *Statistics and Probability Letters*, 77(4) :374–382, 2007.
- [62] C. Zou, G. Yin, L. Feng, and Z. Wang. Nonparametric maximum likelihood approach to multiple change-point problems. *Ann. Statist.*, 42(3) :970–1002, 2014.

Annexe A

Analyse des données du Klementinum

La série de températures du Klementinum est une compilation de mesures de la température de l'air effectuées à la plus ancienne station météorologique d'Europe centrale, à savoir l'observatoire du Klementinum situé à Prague, en République tchèque. Le point de départ de cette série est considéré comme l'année 1775. Les observations météorologiques à l'observatoire du Klementinum se poursuivent jusqu'à aujourd'hui et constituent une source d'information unique et précieuse sur les conditions météorologiques et climatiques au cours de l'histoire moderne. Voir [51] pour une analyse qualitative de cette série et [27] pour une analyse formelle des points de rupture.

La série des maxima journaliers mesurés à l'observatoire du Klementinum (en degrés Celsius) est disponible du 1er janvier 1775 au 30 avril 2005, soit un total de 84 126 observations. Dans la suite, la méthodologie introduite dans la section 2 sera illustrée sur la série des médianes mensuelles, en prenant chaque mois séparément. Cela permet de minimiser les effets des valeurs extrêmes et d'éliminer l'effet saisonnier. Les douze séries qui seront analysées couvrent la période de janvier 1890 à décembre 2004, soit un total de $n = 115$ observations, et peuvent être visualisées sur la Figure A.1.

Les p-valeurs, estimées à partir de $B = 500$ échantillons bootstrap multiplicateurs en série, des tests basés sur T_n^Σ et $T_n^\vee$ dans le cas du poids de Schilling–Schurr lorsque $\beta = 1$, sont présentées dans le Tableau A.1. L’estimation des temps de changement est également fournie lorsque l’hypothèse nulle est rejetée au seuil de $\alpha = 0.10$. À cet effet, le paramètre de discrétisation a été fixé à $N = 100$.

Les résultats du Tableau A.1 mettent en lumière la présence de changements abruptes dans toutes les séries mensuelles excluant le mois de Janvier. En effet, l’estimation des temps de changement $\hat{K}_1$ et $\hat{K}_2$ semblent être exclusivement abrupte. De façon générale, les tests détectent une rupture se produisant au milieu de la série d’observations, c’est-à-dire à l’intérieur de l’intervalle $[n/4, 3n/4]$.

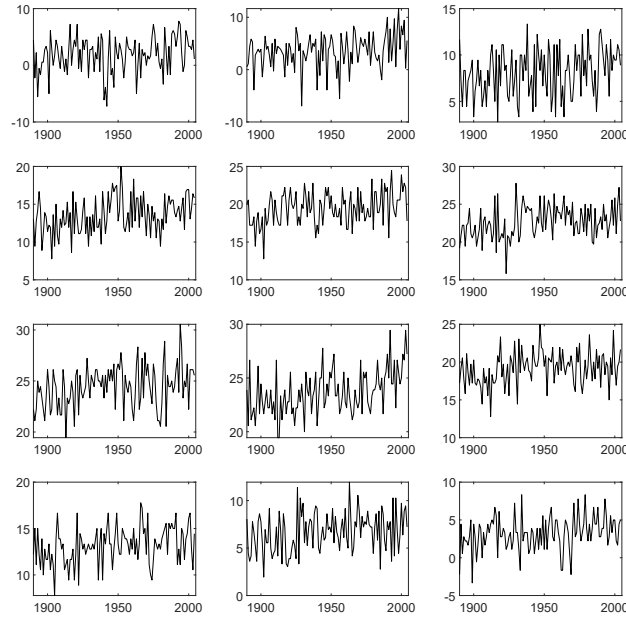


FIGURE A.1 – Séries chronologiques des médianes mensuelles de la température (en degrés Celsius) extraite des données de l’observatoire de Klementinum couvrant la période de janvier 1890 à décembre 2004

TABLE A.1 – P-valeurs estimées (en %) de deux tests de détection de temps de changement appliqués à la médiane mensuelle des maxima journaliers observés à l’observatoire de Klementinum à Prague de janvier 1890 à décembre 2004

Period	$\nu_{3/2}^{\text{Int}}$		$\nu_{3/2}^{\text{Sz}}$		$\nu_{3/2}^{\text{Sch}}$		Change-point estimation	
	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$T_{n,\Sigma}^\nu$	$T_{n,\vee}^\nu$	$\hat{K}_1$	$\hat{K}_2$
January	36.8	39.6	26.2	26.6	8.6	8.6	—	—
February	6.2	7.0	6.2	6.4	5.0	5.4	1965	1966
March	12.4	12.0	5.4	5.6	4.0	4.2	1971	1972
April	7.4	9.0	4.4	4.4	1.0	1.4	1942	1943
May	9.0	9.4	7.8	7.2	2.8	2.8	1914	1915
June	2.2	3.2	0.8	1.6	0.4	0.6	1929	1930
July	11.2	13.4	4.8	6.4	1.8	1.6	1919	1920
August	2.4	3.0	1.4	1.6	1.6	1.6	1941	1942
September	3.2	3.0	1.2	1.2	1.2	1.6	1925	1926
October	7.8	9.4	5.2	5.4	1.4	2.0	1936	1937
November	6.8	8.0	3.6	4.2	0.6	1.0	1925	1926
December	22.4	26.6	9.6	13.4	3.6	4.2	1970	1971

Annexe B

Code Matlab

B.1 Simulations de séries chronologiques

```
1 function Y = GradualUnivCopula_Simulation(T,C,tauC,kap,mu,s,Margin)
2
3 % "Simulation of univariate observations whose lag-1 dependence is
4 % managed by a given copula and a marginal distribution such that
5 % the mean and/or the variance gradually shift between two times"
6 % Input -> T: Sample size
7 %           C: Lag-1 copula (either 1, [3 nu], 11)
8 %           tauC: Kendall's tau of C
9 %           kap=[K1,K2]: Times of change
10 %           mu=[mu1,mu2]: means of the first K1 and last n-K2
    observations
11 %           s=[s1,s2]: std dev of the first K1 and last n-K2
    observations
12 %           Margin: either 'N' (Normal), 'G' (Gamma) or 'B' (Beta)
13 % Output -> Y: Simulated time series
14 % Necessary procedures: TimeSerieCopula_Simulation
15
16 U = TimeSerieCopula_Simulation(T,C,tauC);
```

```

17
18 % Setting up the parameters
19 K = floor(T*kap); K1 = K(1); K2 = K(2);
20 mu1 = mu(1); mu2 = mu(2);
21 s1 = s(1); s2 = s(2);
22
23 % Setting up the marginal distribution
24 if strcmp(Margin, 'N')
25     Q = @(u,mu,s)(norminv(u,mu,s));
26 elseif strcmp(Margin, 'G')
27     Q = @(u,mu,s)(gaminv(u,mu^2/s^2,s^2/mu));
28 elseif strcmp(Margin, 'B')
29     Q = @(u,mu,s)(betainv(u,mu*(mu*(1-mu)-s^2)/s^2,(1-mu)*(mu*(1-mu)-
        s^2)/s^2));
30 end
31
32 % Starting the simulation
33 Y = zeros(T,1);
34 Y(1:K1) = Q(U(1:K1),mu1,s1);
35 Y(K2+1:T) = Q(U(K2+1:T),mu2,s2);
36 for i=K1+1:K2
37     mu = mu1 + (i-K1)*(mu2-mu1)/(K2-K1);
38     s = s1 + (i-K1)*(s2-s1)/(K2-K1);
39     Y(i) = Q(U(i),mu,s);
40 end

```

B.2 Calcul des statistiques

```

1 function [Stats,PVs,Mult] = GradCf_TestsIID(Y,Levy,M,N,Weights)
2
3 % "Tests of gradual change-points in a univariate time series based
4 % on the characteristic function and Levy measures
5 % Input -> Y: Vector of size n of the entries of the time series
6 %           Levy: [nu,kappa], where kappa is the parameter value and

```

```

7 %           - nu=1 -> Integrable weight
8 %           - nu=2 -> Energy weight
9 %           - nu=3 -> Schilling-Schnurr weight
10 %           M: Number of multiplier bootstrap samples
11 %           N: Approximation parameter
12 %           Weights: w1, ..., wn
13 % Output -> Stats: 2-dim vector of the Sum-type and Max-type
    statistics
14 %           PVs: 2-dim vector of the estimated p-values
15 %           Mult: Two-column matrix of the multiplier bootstrap
    samples
16 %           of the two test statistics
17
18 fct = Levy(1); kappa = Levy(2);
19 if fct==1
20     nu = @(t)(1 - exp(-(abs(t))^kappa));
21 elseif fct==2
22     nu = @(t)(t^2/(t^2+kappa^2));
23 elseif fct==3
24     nu = @(t)(abs(t)^kappa);
25 end
26
27 % Entries of matrix B
28 n = length(Y);
29 B = zeros(n,n);
30 for j=1:n
31     for j1=1:n
32         B(j,j1) = nu(Y(j) - Y(j1));
33     end
34 end
35
36 % Computation of the test statistics
37 S = zeros(N,N);
38 for i=1:N-1
39     for j=i+1:N
40         S(i,j) = - Weights(:,i,j).'*B*Weights(:,i,j) / n^2;

```

```

41     S(i,j) = S(i,j)*(j-i)/N;
42     end
43 end
44 T_sum = sum(sum(S)) / n^(3/2);
45 T_max = sqrt(n)*max(max(S));
46 Stats = [T_sum T_max];
47
48 % Computation of the multiplier versions of the statistics
49 Mult = zeros(M,2);
50 parfor h=1:M
51     delta = exprnd(1,1,n);
52
53     Q1 = double.empty;
54     Q = zeros(n,n);
55     An = (delta(1:n) / mean(delta(1:n))) - 1;
56     x = An*B(1:n,1:n)*An.';
57     for j=1:n
58         Aj = (delta(1:j) / mean(delta(1:j))) - 1;
59         Q2 = Aj(1:j)*B(1:j,1:n)*An.';
60         for k=1:n
61             Ak = (delta(1:k) / mean(delta(1:k))) - 1;
62             Q1 = cat(2,Q1,Aj(1:j)*B(1:j,k));
63             Q3 = (j/n)*(k/n)*x;
64             Q(j,k) = Q1*Ak(1:k)' - 2*(k/n)*Q2 + Q3;
65         end
66         Q1 = double.empty;
67     end
68
69     S_Mult = zeros(N,N);
70     for ell1=1:N-1
71         for ell2=ell1+1:N
72             K1 = int16(n*ell1/N); K2 = int16(n*ell2/N);
73             k1 = double(K1); k2 = double(K2);
74             tempo = ones(1,k2-k1)*Q(k1+1:k2,k1+1:k2)*ones(k2-k1,1);
75             S_Mult(ell1,ell2) = - tempo / (n*(k2-k1))^2;
76             S_Mult(ell1,ell2) = S_Mult(ell1,ell2)*(ell2-ell1)/N;

```

```

77         end
78     end
79     T_sum_mult = sum(sum(S_Mult)) / (n^(3/2));
80     T_max_mult = sqrt(n) * max(max(S_Mult));
81     Mult(h,:) = [T_sum_mult T_max_mult];
82 end
83 PVs = sum(Mult > Stats) / M;

```

```

1 function [Stats,PVs,Mult] = GradCf_TestsSerial(Y,Levy,M,N,Weights)
2
3 % "Tests of gradual change-points in a univariate time series based
4 % on the characteristic function and Levy measures
5 % Input -> Y: Vector of size n of the entries of the time series
6 %         Levy: [nu,kappa], where kappa is the parameter value and
7 %         - nu=1 -> Integrable weight
8 %         - nu=2 -> Schilling-Schnurr weight
9 %         - nu=3 -> Energy weight
10 %         M: Number of multiplier bootstrap samples
11 %         N: Approximation parameter
12 %         Weights: w1, ..., wn
13 % Output -> Stats: 2-dim vector of the Sum-type and Max-type
14 %             statistics
15 %             PVs: 2-dim vector of the estimated p-values
16 %             Mult: Two-column matrix of the multiplier bootstrap
17 %                 samples
18 %                 of the two test statistics
19
20 fct = Levy(1); kappa = Levy(2);
21 if fct==1
22     nu = @(t)(1 - exp(-(abs(t))^kappa));
23 elseif fct==2
24     nu = @(t)(t^2/(t^2+kappa^2));
25 elseif fct==3
26     nu = @(t)(abs(t)^kappa);
27 end

```

```

27 % Computation of the test statistics
28 n = length(Y); B = zeros(n,n);
29 for j=1:n
30     for j1=1:n
31         t = Y(j) - Y(j1);
32         B(j,j1) = nu(t);
33     end
34 end
35
36 S = zeros(N,N);
37 for i=1:N-1
38     for j=i+1:N
39         S(i,j)= - Weights(:,i,j).'*B*Weights(:,i,j) / n^2;
40         S(i,j) = S(i,j)*(j-i)/N;
41     end
42 end
43 T_sum = sum(sum(S)) / n^(3/2);
44 T_max = sqrt(n)*max(max(S));
45 Stats = [T_sum T_max];
46
47 % Computation of the multiplier versions of the statistics
48 Mult = zeros(M,2);
49 ell_n = int16(1.1*n^(1/4));
50 parfor h=1:M
51     q = 1/(2*double(ell_n)-1);
52     delta = zeros(1,n);
53     w = gamrnd(q,q,n+2*ell_n-1,1);
54     for j=1:n
55         delta(j) = mean(w(j:j+2*ell_n-2));
56     end
57
58     Q1 = double.empty;
59     Q = zeros(n,n);
60     An = (delta(1:n) / mean(delta(1:n))) - 1;
61     x = An*B(1:n,1:n)*An.';
62     for j=1:n

```



```

63     Aj = (delta(1:j) / mean(delta(1:j))) - 1;
64     Q2 = Aj(1:j)*B(1:j,1:n)*An.';
65     for k=1:n
66         Ak = (delta(1:k) / mean(delta(1:k))) - 1;
67         Q1 = cat(2,Q1,Aj(1:j)*B(1:j,k));
68         Q3 = (j/n)*(k/n)*x;
69         Q(j,k) = Q1*Ak(1:k)' - 2*(k/n)*Q2 + Q3;
70     end
71     Q1 = double.empty;
72 end
73
74 S_Mult = zeros(N,N);
75 for ell1=1:N-1
76     for ell2=ell1+1:N
77         K1 = int16(n*ell1/N); K2 = int16(n*ell2/N);
78         k1 = double(K1); k2 = double(K2);
79         tempo = ones(1,k2-k1)*Q(k1+1:k2,k1+1:k2)*ones(k2-k1,1);
80         S_Mult(ell1,ell2) = - tempo / (n*(k2-k1))^2;
81         S_Mult(ell1,ell2) = S_Mult(ell1,ell2)*(ell2-ell1)/N;
82     end
83 end
84 T_sum_mult = sum(sum(S_Mult)) / (n^(3/2));
85 T_max_mult = sqrt(n) * max(max(S_Mult));
86 Mult(h,:) = [T_sum_mult T_max_mult];
87 end
88 PVs = sum(Mult > Stats) / M;

```

B.3 Puissance des tests

```

1 function Power = GradCf_PowerIID(kap,mu,s,Margin,Levy,n,N,B)
2
3 % "Estimation of the power of the test of Bouchard & Quessy (2025)
   for

```

```

4 % the detection of a gradual change-point in a univariate time
   series
5 % based on the characteristic function"
6 % Input -> kap=[K1,K2]: Times of change
7 %           mu=[mu1,mu2]: means of the first K1 and last n-K2
   observations
8 %           s=[s1,s2]: std dev of the first K1 and last n-K2
   observations
9 %           Margin: either 'N' (Normal distr) or 'G' (Gamma distr)
10 %           Levy: [nu,alpha], where alpha is the parameter value and
11 %           - nu=1 -> Integrable weight
12 %           - nu=2 -> Energy weight
13 %           - nu=3 -> Schilling-Schnurr weight
14 %           n: Sample size
15 %           N: Approximation parameter
16 %           B: Number of Warp-speed bootstrap samples
17 % Output -> Power: Estimated power of the test
18 % Necessary procedures: GradualUnivCopula_Simulation, GradCf_TestsIID
19
20 Weights = zeros(n,N,N);
21 for i=1:N-1
22     for j=i+1:N
23         K1 = int16(n*i/N); k1 = double(K1);
24         K2 = int16(n*j/N); k2 = double(K2);
25
26         w_bar = (k1+k2-1)/(2*n);
27         for ell=1:n
28             if ell<=k1
29                 W = 1;
30             elseif ell>=k2
31                 W = 0;
32             else
33                 W = (k2-ell)/(k2-k1);
34             end
35             Weights(ell,i,j) = W - w_bar;
36         end

```

```

37     end
38 end
39
40 Stats = zeros(B,2); Mult = zeros(B,2);
41 parfor b=1:B
42     Y = GradualUnivCopula_Simulation(n,1,0,kap,mu,s,Margin);
43     [Stats(b,:),~,Mult(b,:)] = GradCf_TestsIID(Y,Levy,1,N,Weights);
44 end
45 VC = prctile(Mult,95);
46 Power = sum(Stats > VC) / B;

```

```

1 function Power = GradCf_PowerSerial(kap,mu,s,Margin,Levy,C,tauC,n,N,B
   )
2
3 % "Estimation of the power of the test of Bouchard & Quessy (2025)
   for
4 % the detection of a gradual change-point in a univariate time
   series
5 % based on the characteristic function"
6 % Input -> kap=[K1,K2]: Times of change
7 %           mu=[mu1,mu2]: means of the first K1 and last n-K2
   observations
8 %           s=[s1,s2]: std dev of the first K1 and last n-K2
   observations
9 %           Margin: either 'N' (Normal distr) or 'G' (Gamma distr)
10 %           Levy: [nu,alpha], where alpha is the parameter value and
11 %                - nu=1 -> Integrable weight
12 %                - nu=2 -> Energy weight
13 %                - nu=3 -> Schilling-Schnurr weight
14 %           C: Lag-1 copula (either 1, [3 nu], 11)
15 %           tauC: Kendall's tau of C
16 %           n: Sample size
17 %           N: Approximation parameter
18 %           B: Number of Warp-speed bootstrap samples
19 % Output -> Power: Estimated power of the test
20 % Necessary procedures: GradualUnivCopula_Simulation,

```

```

GradCf_TestsSerial
21
22 Weights = zeros(n,N,N);
23 for i=1:N-1
24     for j=i+1:N
25         K1 = int16(n*i/N); k1 = double(K1);
26         K2 = int16(n*j/N); k2 = double(K2);
27
28         w_bar = (k1+k2-1)/(2*n);
29         for ell=1:n
30             if ell<=k1
31                 W = 1;
32             elseif ell>=k2
33                 W = 0;
34             else
35                 W = (k2-ell)/(k2-k1);
36             end
37             Weights(ell,i,j) = W - w_bar;
38         end
39     end
40 end
41
42 Stats = zeros(B,2); Mult = zeros(B,2);
43 parfor b=1:B
44     Y = GradualUnivCopula_Simulation(n,C,tauC,kap,mu,s,Margin);
45     [Stats(b,:),~,Mult(b,:)] = GradCf_TestsSerial(Y,Levy,1,N,Weights)
46     ;
47 end
48 VC = prctile(Mult,95);
Power = sum(Stats > VC) / B;

```

B.4 Analyse des données du Klementinum

```

1 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2 % Construction of the time series of the
3 % median of monthly maxima (in Celsius)
4 load('KlementinumFULL.mat');
5 A = KlementinumFULL;
6 B = varfun(@median,A,'GroupingVariables', ...
7           {'Year','Month'},'InputVariable','TMAX');
8 B = B(1:2760,:);
9 KlementinumByMonths = zeros(230,12);
10 for j=1:12
11     KlementinumByMonths(:,j) = B.median_TMAX(B.Month == j);
12 end
13 KlementinumByMonths = (5/9)*(KlementinumByMonths-32);
14
15 % FIGURES
16 load('KlementinumByMonths.mat');
17 Z = KlementinumByMonths(116:230,:);
18 v = 1890:2004;
19 for j=1:12
20     subplot(4,3,j); plot(v,Z(:,j),'color','k'); xlim([1890 2005]);
21 end
22
23 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
24 % Table 7
25 % Tests of change-point detection
26 % and change-point estimation
27 load('KlementinumByMonths.mat'); Z = KlementinumByMonths(116:230,:);
28 N = 40; n = length(Z); M = 500; W = GradCf_Weights(n,N);
29 PV = zeros(12,6); Khat = zeros(12,2); t=clock;
30 for j=1:12
31     Y = Z(:,j);
32     [~,a1,~] = GradCf_TestsSerial(Y,[1 3/2],M,N,W);
33     [~,a2,~] = GradCf_TestsSerial(Y,[2 3/2],M,N,W);
34     [~,a3,~] = GradCf_TestsSerial(Y,[3 3/2],M,N,W);
35     PV(j,:) = 100 * [a1 a2 a3];
36     Khat(j,:) = int16(GradualCf_CPEstimators(Y,[3 3/2],100)+1889);

```

```

37 end
38 etime(clock,t)
39 LatexTable(PV,1,'PVs-Klementinum.txt');

```

```

1 function Weights = GradCf_Weights(n,N)
2
3 % "Computation of the weights w1,...,wn for each K1<K2 in {1,...,n}
4 % in the gradual-change model described by Bouchard & Quessy (2025),
5 % where wj=1 if j<=K1, wj=0 if j>=K2, and wj=(K2-j)/(K2-K1)
6 % otherwise"
7 % Input -> n: Sample size
8 %          N: Approximation parameter
9 % Output -> Weights: n x N x N matrix of weights
10
11 Weights = zeros(n,N,N);
12 for i=1:N-1
13     for j=i+1:N
14         K1 = int16(n*i/N); k1 = double(K1);
15         K2 = int16(n*j/N); k2 = double(K2);
16
17         w_bar = (k1+k2-1)/(2*n);
18         for ell=1:n
19             if ell<=k1
20                 W = 1;
21             elseif ell>=k2
22                 W = 0;
23             else
24                 W = (k2-ell)/(k2-k1);
25             end
26             Weights(ell,i,j) = W - w_bar;
27         end
28     end
29 end

```

```

1 function Khat = GradualCf_CPEstimators(Y,Levy,N)
2

```

```

3 % "Estimation of the two times of change in a univariate time series
4 %   subject to a gradual change based on the characteristic function
5 %   as described by Bouchard & Quessy (2025)"
6 %   Input  -> Y: Vector of size n of the entries of the time series
7 %           Levy: [nu,alpha], where alpha is the parameter value and
8 %           - nu=1 -> Integrable weight
9 %           - nu=2 -> Energy weight
10 %           - nu=3 -> Schilling-Schnurr weight
11 %           N: Approximation parameter
12 % Output -> Khat: Estimator of (K1,K2)
13 % Necessary procedures: GradCf_Weights
14
15 fct = Levy(1); kappa = Levy(2);
16 if fct==1
17     nu = @(t)(1 - exp(-(abs(t))^kappa));
18 elseif fct==2
19     nu = @(t)(t^2/(t^2+kappa^2));
20 elseif fct==3
21     nu = @(t)(abs(t)^kappa);
22 end
23
24 % Entries of matrix B
25 n = length(Y);
26 B = zeros(n,n);
27 for j=1:n
28     for j1=1:n
29         B(j,j1) = nu(Y(j) - Y(j1));
30     end
31 end
32
33 % Computation of the estimator
34 % of the two times of change
35 Weights = GradCf_Weights(n,N);
36 S = zeros(N,N);
37 for j=1:N-1
38     for j1=j+1:N

```

```
39         u1 = j/N; u2 = j1/n;  
40         F = (2*u1+u2)/3 - (u1+u2)^2/4;  
41         TEMPO = -Weights(:,j,j1) .* B * Weights(:,j,j1);  
42         S(j,j1) = TEMPO / sqrt(F);  
43     end  
44 end  
45 [maximum,id] = max(S);  
46 [~,K2hat] = max(maximum);  
47 K1hat = id(K2hat);  
48 Khat = (n/N)*[K1hat K2hat];
```