# TWO APPROACHES TO THE ISOTONIC CHANGE-POINT PROBLEM:

# NONPARAMETRIC AND MINIMAX

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Karl D'Silva

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Department:
Statistics

April 2014

Fargo, North Dakota

# North Dakota State University
Graduate School

**Title**

TWO APPROACHES TO THE ISOTONIC CHANGE-POINT PROBLEM: NONPARAMETRIC AND MINIMAX

**By**

KARL D'SILVA

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

**DOCTOR OF PHILOSOPHY**

SUPERVISORY COMMITTEE:

GANG SHEN
Chair

RHONDA MAGEL

MEGAN ORR

JESSICA STRIKER

Approved:

APRIL 14TH, 2014 — Date

RHONDA MAGEL — Department Chair

## ABSTRACT

A change in model parameters over time often characterizes major events. Situations in which this may arise include observing increasing temperatures, intense rainfall, and the valuation of a stock. The question is whether these observations are simply the result of natural variation, or rather are indicative of an underlying monotonic trend. This is known as the isotonic change-point problem. Two approaches to this problem are considered: Firstly, for correlated data with short-range dependence, we prove that a particular U-statistic based on a modified version of the Jonckheere-Terpstra test statistic is asymptotically equivalent to a more complex U-statistic discussed by Shen and Xu (2013); one that has been shown to outperform other existing tests in a variety of situations. Secondly, we shall justify and utilize the minimax criterion in order to identify the optimal test statistic within a specified class. We shall see that, as motivated by the projection method, the aforementioned class is the class of contrasts. It shall be proven that the set of coefficients originally proposed by Abelson and Tukey (1963), and utilized by Brillinger (1989) in the isotonic change-point setting, are in fact minimax in the independent data case. For correlated data with short-range dependence, we shall demonstrate a sufficient condition for minimaxity to hold.

## DEDICATION

To my parents Milbhor and Pearl.

**TABLE OF CONTENTS**

# LIST OF FIGURES

# 1. INTRODUCTION

A change in model parameters over time often characterizes major events. Situations in which this may arise include observing increasing temperatures, intense rainfall, and the valuation of a stock. The question is whether these observations are simply the result of natural variation, or rather are indicative of an underlying monotonic trend. This is known as the isotonic change-point problem.

We shall formulate the problem as follows: Unless noted, we shall assume throughout that the random process has the form

$$X_i \triangleq \mu_i + Z_i,\ i = 1, 2, \ldots, n \tag{1}$$

where the $\mu_i$ are non-decreasing deterministic signals, and the $Z_i$ are zero-mean strictly stationary noise. We shall also require that the $Z_i$ satisfy the following mixing condition:

Let $\mathscr{F}_a^b$ be the $\sigma$-algebra generated by $\{Z_i\}_a^b$ and

$\beta_n \triangleq E \sup_{A \in \mathscr{F}_n^{+\infty}} |P(A|\mathscr{F}_{-\infty}^0) - P(A)|$; it is assumed that

$$\text{there exists a } \delta > 0 \text{ such that } \lim_{n \to +\infty} n^{1+\delta}\beta_n = 0 \tag{2}$$

This condition is fulfilled by a wide range of stochastic processes with short-range dependence. These include, but are not limited to, m-dependent processes and invertible ARMA processes (Shen and Xu 2013). In addition, all observations are assumed to have been obtained prior to analysis- this is known as an offline or nonsequential analysis.

Throughout, we shall consider the following hypothesis test regarding the $\mu_i$ :

$$H_0 : \mu_i = \mu_{i+1} \text{ for all } i = 1, 2, \ldots, n-1 \tag{3}$$

$$H_a : \mu_i < \mu_{i+1} \text{ for some } i = 1, 2, \ldots, n-1 \tag{4}$$

Our work will be motivated by the latter two approaches mentioned in the next chapter. More specifically, in Chapter 3, we shall consider a nonparametric approach and in Chapter 4, a criterion-based approach to the isotonic change-point problem. In particular, we will justify and utilize the minimax criterion in order to identify the optimal test statistic within a specified class. We shall see that, as motivated by the projection method, the aforementioned class is the class of contrasts.

# 2. LITERATURE OVERVIEW

There have been four main approaches to the isotonic change-point problem. Alvarez and Dey (2009) consider Bayesian estimation of $\mu_1, \mu_2, \ldots, \mu_n$, which is outside the scope of our research. However, credible sets can be used to perform hypothesis tests on desired parameters. A Bayesian treatment allows for more general hypothesis tests concerning the specific form of the trend, as opposed to simply testing for the existence of an isotonic trend. The main drawback of this work is that the Bayesian isotonic method approach is only applied and discussed within the i.i.d. setting, which limits its practical usefulness.

Wu, Woodroofe, and Mentz (2001) consider a test for the existence of a monotonic trend in short-range dependent sequences. The main drawback is that this test statistic has no explicit distribution and depends on a tuning parameter. Furthermore, it has been demonstrated via simulation by Shen and Xu (2013), that the aforementioned test statistic stabilizes slowly to said distribution, and in certain settings, has an inflated type I error rate.

The test statistic proposed by Brillinger (1989) is based on a contrast of the data, with the coefficients specified in Abelson and Tukey (1963). These specific coefficients, however, were derived under the assumption that the random variables that form the stochastic process are independent. This test statistic has an asymptotically normal distribution.

Shen and Xu (2013) propose a nonparametric test statistic that is a modified version of the traditional Jonckheere test statistic of homogeneity versus monotonicity. It is a U-statistic of degree 2, that is shown to be asymptotically normal.

# 3. NONPARAMETRIC APPROACH

Recall that the Jonckheere-Terpstra test statistic for trend is $\sum_{i<j} \mathbb{1}_{\{X_i<X_j\}}$, which is across all the groups. In addition, observe that $\mathbb{1}_{\{X_i<X_j\}}$ is simply the Mann-Whitney statistic, in the case where the sample size is equal to 1 for both groups $i$ and $j$. As an aside, note that the Jonckheere-Terpstra test is performed under the assumption that the samples are independent (between groups). Unless stated, all the following derivations are performed under $H_0$.

## 3.1. Defining $U_n^*$

Consider the following modification of the Jonckheere-Terpstra test statistic:

$$U_n^* \triangleq \tfrac{n-1}{n+1}\binom{n}{2}^{-1}\sum_{i<j}(\mathbb{1}_{\{X_i<X_j\}} - \tfrac{1}{2}) \tag{5}$$

Note that there are $\binom{n}{2}$ summands in $U_n^*$, and that we have centered each one by its unconditional expectation.

We wish to find the asymptotic distribution of $U_n^*$. However, the exact distribution of a U-statistic can often be difficult to derive. To that end, observe that $U_n^*$ is a statistic based on $X_1, \ldots, X_n$. We shall approximate $U_n^*$ by its projection on $X_1, ..., X_n$, which we denote by $T_n^*$ :

$$T_n^* \triangleq EU_n^* + \sum_{k=1}^{n}[E(U_n^*|X_k) - EU_n^*] \tag{6}$$

This is known as the Hájek projection; which projects a random variable onto the class consisting of sums of measurable mappings that have a finite second moment. The projection is the closest element in the aforementioned class to the original random variable, in the sense of squared expectation (van de Vaart 1998, p. 153).

We shall now derive an explicit form for $T_n^*$, using the results given below:

### 3.2. Derivation of $T_n^*$

**Definition 1.** *Let $F(x) \triangleq P(X_1 \leq x)$, where $P$ is the appropriate probability measure. This is the cumulative distribution function of $X_i$, $i = 1, 2, \ldots, n$.*

**Proposition 1.** $E(\mathbb{1}_{\{X_i < X_j\}} | X_i) = 1 - F(X_i)$.

**Proposition 2.** $E(\mathbb{1}_{\{X_i < X_j\}} | X_j) = F(X_j)$.

**Proposition 3.** $EU_n^* = 0$.

So by the definition of $T_n^*$ and Proposition 3,

$$T_n^* = \sum_{k=1}^{n} E(U_n | X_k) \tag{7}$$

$$= \frac{n-1}{n+1} \binom{n}{2}^{-1} \sum_{k=1}^{n} E\left[ \sum_{i<j} (\mathbb{1}_{\{X_i < X_j\}} - \tfrac{1}{2}) \,|X_k \right]. \tag{8}$$

**Proposition 4.** *If $k \neq i, j$, then* $E\left[ \sum_{i<j} (\mathbb{1}_{\{X_i < X_j\}} - \frac{1}{2}) \,|X_k \right] = 0$.

Proposition 5 follows after some algebraic manipulation involving rearrangement and rexpression of summands:

**Proposition 5.** $T_n^* = 4n^{-1} \sum_{k=1}^{n} [(\frac{k}{n+1} - \frac{1}{2})(F(X_k) - \frac{1}{2})]$.

### 3.3. Asymptotic distribution of $U_n^*$- independent case

For this subsection, we assume that the $Z_i$ are independent. Noting that under $H_0$, it follows that $F(X_i) \sim \mathscr{U}(0,1),\ i = 1, 2, \ldots, n$ , where $\mathscr{U}(0,1)$ is the uniform distribution with support on $[0, 1]$. Thus, $ET_n^* = 0$. Next, regarding the variance of $T_n^*$:

**Proposition 6.** $\text{var}(T_n^*) = \frac{n-1}{9n(n+1)}$.

Notice that we have a weighted sum of independent and identically distributed random variables. Define

$$\sigma_k^2 \triangleq \text{var}((\tfrac{k}{n+1} - \tfrac{1}{2})(F(X_k) - \tfrac{1}{2})), \tag{9}$$

$$s_n^2 \triangleq \sum_{k=1}^{n} \sigma_k^2. \tag{10}$$

Noting that $\sigma_k^2 \leq \frac{1}{24}$ for all $k$, and $s_n^2 \propto \frac{n(n-1)}{(n+1)}$,
it follows that

$$\max_{k=1,\dots,n} \frac{\sigma_k^2}{s_n^2} \to 0 \text{ as } n \to +\infty, \tag{11}$$

which is a sufficient condition for Lindeberg's central limit theorem (Billingsley 1995, p.369).

Thus, by the Lindeberg central limit theorem,

$$\frac{\sqrt{n}(T_n^* - 0)}{(\frac{n-1}{n+1})^{1/2}} \xrightarrow{d} N(0, \tfrac{1}{9}) \quad \text{as} \quad n \to +\infty. \tag{12}$$

Furthermore, since $\frac{n-1}{n+1} \xrightarrow{p} 1$, it follows by Slutsky's theorem and the continuous mapping theorem that

$$\sqrt{n}(\tfrac{\sqrt{3}}{2}T_n^* - 0) \xrightarrow{d} N(0, \tfrac{1}{12}) \tag{13}$$

Hence, we have shown that the test statistic $T_n^*$ is asymptotically normal. With that said, $U_n^*$ can also be shown to be asymptotically normal if it can be demonstrated that the term $\sqrt{n}(U_n^* - T_n^*)$ goes to 0 in probability. It turns out that this is in fact true even in the general case. This is discussed in the next section.

### 3.4. Asymptotic distribution of $U_n^*$- dependent case

In this section, we assume the $Z_i$ satisfy the mixing condition specified in (2). Define

$$\gamma_k \triangleq \operatorname{cov}(F(X_1), F(X_{k+1})), \quad k = 0, 1, \ldots \tag{14}$$

$$\sigma^2 \triangleq \gamma_0 + 2\sum_{k=1}^{+\infty} \gamma_k \neq 0, \tag{15}$$

$$Y_i \triangleq (\tfrac{i}{n+1} - \tfrac{1}{2})(F(X_i) - \tfrac{1}{2}), \quad i = 1, 2, \ldots, n \tag{16}$$

$$T_n^* \triangleq 4n^{-1}\sum_{i=1}^{n} Y_i \tag{17}$$

In order to demonstrate the asymptotic normality of $T_n^*$, we shall utilize the following central limit theorem for weakly dependent sequences owed to Herrndoff (1984, p.142):

**Corollary 1.** *Let $q > 2$ and $p = 2/q$. If $\operatorname{var}[\sqrt{n}(T_n^* - 0)] \to \sigma^2 > 0$, $EY_j = 0, EY_j^2 < +\infty$, and there exists $q > 2$ such that $\limsup_{n\to\infty} \|Y_n\|_q < +\infty$ where $\|Y_n\|_q \triangleq [E|Y_n|^q]^{1/q}$, and $\sum_{k=1}^{+\infty} \beta_k^{1-p} < +\infty$, then $\sqrt{n}(\frac{\sqrt{3}}{2}T_n^* - 0) \xrightarrow{d} N(0, \sigma^2)$.*

Therefore, to demonstrate the asymptotic normality of $T_n^*$, we will need to verify the preceding conditions stated above.

**Lemma 1.** $\operatorname{var}[\sqrt{n}(\frac{\sqrt{3}}{2}T_n^* - 0)] \to \sigma^2 > 0.$

**Lemma 2.** $EY_j = 0, EY_j^2 < +\infty$

**Lemma 3.** *There exists $q > 2$ such that* $\limsup_{n\to\infty}[E|Y_n|^q]^{1/q} < +\infty$, *and* $\sum_{k=1}^{+\infty}\beta_k^{1-p} < +\infty$, *where* $p \triangleq 2/q$.

Since the aforementioned conditions are met, then the asymptotic normality of $T_n^*$ is established, via Corollary 1. We state this as a theorem:

**Theorem 1.** *Assume the random process in* (1) *satisfies the mixing condition specified in* (2)*. Then, under* $H_0$, $\sqrt{n}(\frac{\sqrt{3}}{2}T_n^* - 0) \xrightarrow{d} N(0, \sigma^2)$.

Define $R_n^* \triangleq \frac{\sqrt{3}}{2}(U_n^* - T_n^*)$, or equivalently

$$R_n^* \triangleq \left[\tfrac{\sqrt{3}}{2}(\tfrac{n-1}{n+1})\binom{n}{2}^{-1}\sum_{i<j}(\mathbb{1}_{\{X_i<X_j\}} - \tfrac{1}{2})\right] - \left[2\sqrt{3}n^{-1}\sum_{k=1}^{n}[(\tfrac{k}{n+1} - \tfrac{1}{2})(F(X_k) - \tfrac{1}{2})]\right] \tag{18}$$

If it can be shown that $\sqrt{n}(R_n^* - 0) \xrightarrow{p} 0$, then the asymptotic normality of $U_n^*$ is easily verified. This can be proven by utilizing a result due to Yoshihara (1976). In order to facilitate usage of this lemma, we reexpress $T_n^*$ as follows:

**Proposition 7.** $T_n^* = \frac{n-1}{n+1}\binom{n}{2}^{-1}\sum_{i<j}(F(X_j) - F(X_i))$

Now we can express $R_n^*$ in a simpler form:

$$R_n^* = \frac{\sqrt{3}(n-1)}{2(n+1)}\binom{n}{2}^{-1}\sum_{i<j}(\mathbb{1}_{\{X_i<X_j\}} - F(X_j) + F(X_i) - \tfrac{1}{2}). \tag{19}$$

Next, in preparation for the following exposition, we define the following quantities:

Let $\{X_i, -\infty < i < +\infty\}$ be a strictly stationary sequence of random variables defined on a probability space $(\Omega, \mathscr{A}, P)$. Let $i_1 < i_2 < \cdots < i_k$ be otherwise arbitrary integers. Corresponding to these integers, let $F$ denote the cumulative distribution function of $(X_{i_1}, \ldots, X_{i_k})$. Then, there exists a Lebesgue-Stieltjes probability measure $\mu_F$ on $(\mathbf{R}^k, \mathscr{B}^k)$ such that $\mu_F(-\infty, \mathbf{x}) = F(\mathbf{x})$, where $\mathbf{x} \in \mathbf{R}^k$ (Athreya and Lahiri 2006, p.46).

Next, for any $j$ such that $1 \leq j < k$, define

$$F_j^{(k)}(B^{(j)} \times B^{(k-j)}) \triangleq \mu_F((X_{i_1}, \ldots, X_{i_j}) \in B^{(j)})\mu_F((X_{i_{j+1}}, \ldots, X_{i_k}) \in B^{(k-j)}) \tag{20}$$

and

$$F_0^{(k)}(B^{(k)}) \triangleq \mu_F((X_{i_1}, \ldots, X_{i_k}) \in B^{(k)}), \tag{21}$$

where $B^{(j)}$ denotes a Borel set in $\mathbf{R}^j$.

Now we will state the lemma due to Yoshihara (1976):

**Lemma 4.** *For any $j$ such that $0 \leq j < k$, let $h(x_1, x_2, \ldots, x_k)$ be a Borel function such that $\int_{\mathbf{R}^k} |h(x_1, x_2, \ldots, x_k)|^{1+\delta} dF_j^{(k)} \leq M$ for some $\delta > 0, M \geq 0$. Then $|\int_{\mathbf{R}^k} h(x_1, x_2, \ldots, x_k) dF_0^{(k)} - \int_{\mathbf{R}^k} h(x_1, x_2, \ldots, x_k) dF_j^{(k)}| \leq 4M^{1/(1+\delta)} \beta_{i_{j+1}-i_j}^{\delta/(1+\delta)}$.*

By using this lemma, it can be shown that $E(\sqrt{n}R_n^*)^2 \to 0$, and thus $\sqrt{n}R_n^* \xrightarrow{p} 0$. We state this as a theorem.

**Theorem 2.** *Assume the random process in* (1) *satisfies the mixing condition specified in* (2)*. Then, under $H_0$, $\sqrt{n}(R_n^* - 0) \xrightarrow{p} 0$.*

Finally, we can state the main result:

**Theorem 3.** *Assume the random process in* (1) *satisfies the mixing condition specified in* (2)*. Then, under $H_0$, $\sqrt{n}(\frac{\sqrt{3}}{2}U_n^* - 0) \xrightarrow{d} N(0, \sigma^2)$.*

This is easily shown via Slutsky's theorem, and by using the results from Theorems 1 and 2.

### 3.5. Comment

From Shen, Xu (2013), we have:

$$U_n \triangleq \sqrt{3}\binom{n}{2}^{-1}\sum_{i,j=1}^{n}(\tfrac{j-i}{n+1})^{+}(\mathbb{1}_{\{X_i<X_j\}}-\tfrac{1}{2}) \tag{22}$$

$$T_n \triangleq (2\sqrt{3})(n-1)^{-1}\sum_{k=1}^{n}[(\tfrac{k}{n+1}-\tfrac{1}{2})(F(X_k)-\tfrac{1}{2})] \tag{23}$$

Of theoretical interest is the fact that it has been shown, via the projection method, that the "linearization" of $U_n$, denoted by $T_n$ is asymptotically equivalent to the "linearization" of $U_n^*$, denoted by $T_n^*$. In addition, we have demonstrated that the remainder term $R_n^*$ goes to 0; thus implying that $U_n$ and $U_n^*$ are asymptotically equivalent in distribution. Thus, for large $n$, $U_n^*$ would be preferred to $U_n$ owing to its simpler form, and thus its greater computational efficiency.

### 3.6. Numerical study

In this subsection, we shall consider various configurations of the $\mu_i$ under $H_a$, and examine via a simulation study, the power of the four tests discussed up to this point for said configurations. In addition, we shall consider four random error structures detailed below.

Recall the form of the random process: $X_i = \mu_i + Z_i$. Consider $\mu_i = \delta^*\phi(i/n),\ i = 1, 2, \ldots, n$, where either $\phi(x) = \mathbb{1}_{\{x>0.5\}} - 0.5$ or $\phi(x) = x - 0.5$ respectively, where $\delta^*$ is a constant taking the value of $j/100$ with $j = 0, 1, \ldots, 99$.

Here, the value of $\delta^*$ controls the magnitude that the $\mu_i$ are deviating from $H_0$. On the other hand, $\phi(\cdot)$ controls the form of how the $\delta_i$ are deviating from the null hypothesis. The forms of $\phi$ given above correspond to a sudden jump or an increasing linear trend respectively.

The following four different zero-mean standardized stationary processes are taken into consideration:

1) $Z_i \overset{iid}{\sim} N(0,1)$

2) $Z_i \overset{iid}{\sim} t_3/\sqrt{3}$, where $t_3$ is the $t$ distribution with 3 degrees of freedom

3) $Z_i = (\sqrt{1-0.3^2})\nu_i$, where $\nu_i = 0.3\nu_{i-1} + \epsilon_i$ and $\epsilon_i \overset{iid}{\sim} N(0,1)$

4) $Z_i = (\sqrt{1-0.7^2})\nu_i$,where $\nu_i = 0.7\nu_{i-1} + \epsilon_i$ and $\epsilon_i \overset{iid}{\sim} N(0,1)$

(We note that 3) and 4) are standardized Gaussian AR(1) processes)

When applicable, for all the forthcoming simulations and tests, we shall use $m_n = n^{1/3}$, where $m_n$ is a deterministic sequence that plays a role in the lag-window estimator of the variance (refer to Shen and Xu 2013 for further details).

For each of the four tests, we perform 10,000 simulations with a sample size equal to 200, at the different combinations of the $\mu_i$ described above, all at the $\alpha = 0.05$ significance level.

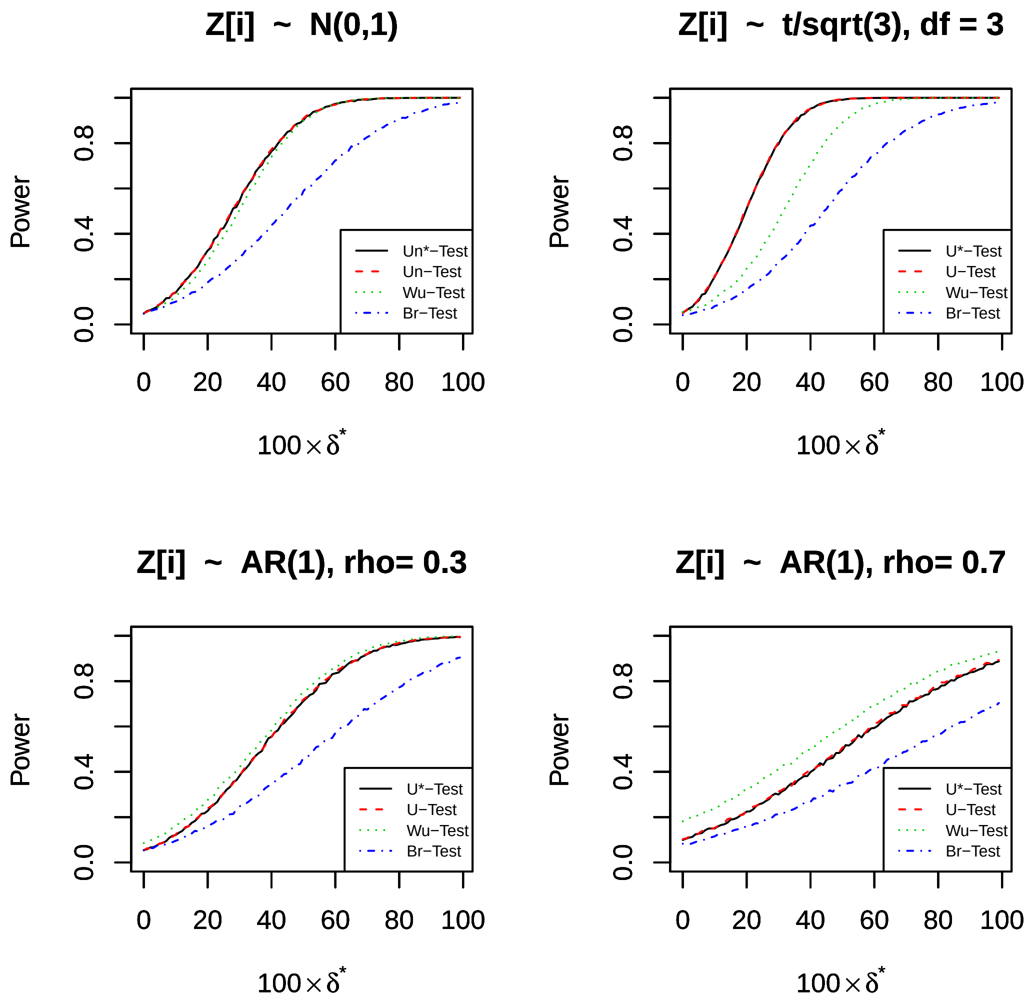


Figure 1. Power curves for $U_n^*$, $U_n$, Wu's test, and Brillinger's test where $\phi(x) = \mathbb{1}_{\{x>0.5\}} - 0.5$

As expected due to Theorem 3, the power curves of $U_n^*$ and $U_n$ are equivalent. Note that Wu's test does not actually have higher power in the latter two cases above, due to its simulated type I error rate being greater than the nominal type I error rate of 0.05 (Shen and Xu, 2013). The same remarks apply to Figure 2.

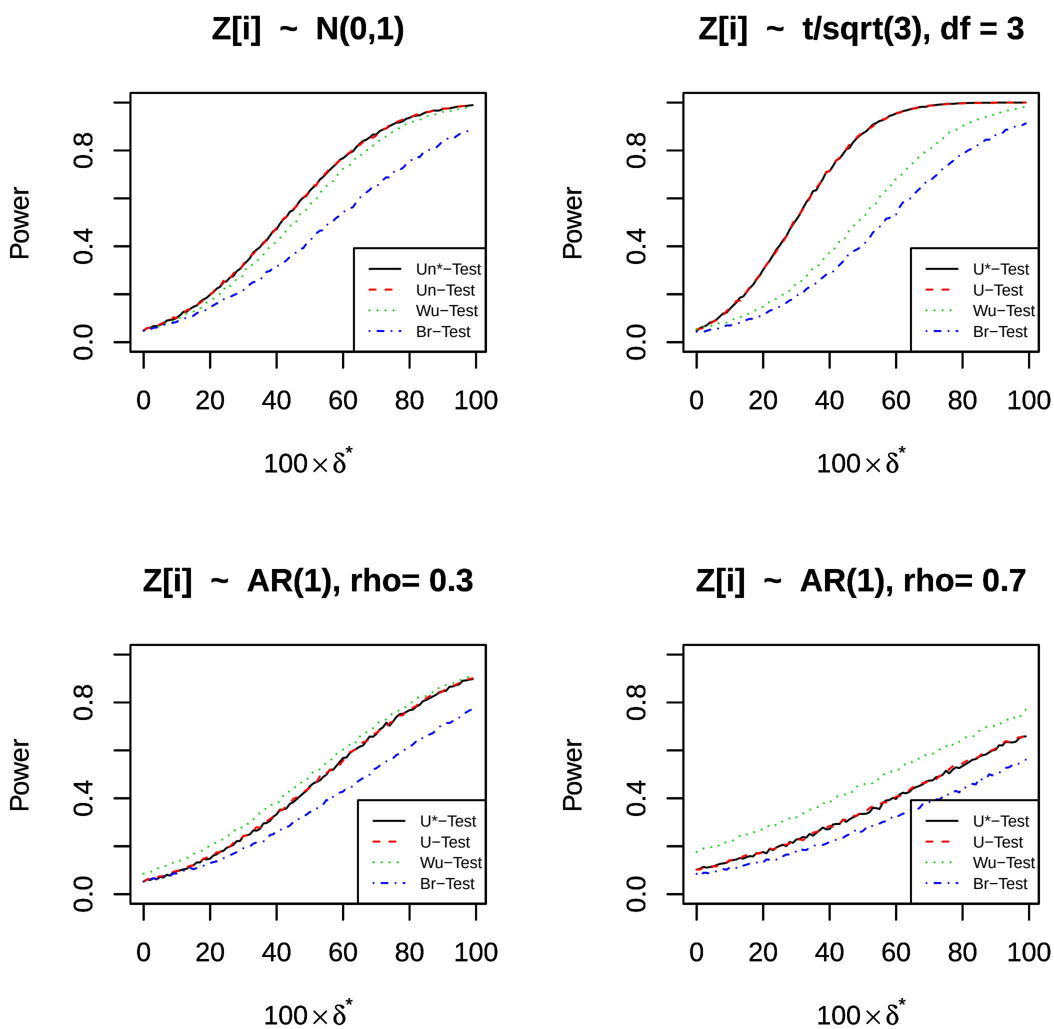


Figure 2. Power curves for $U_n^*$, $U_n$, Wu's test, and Brillinger's test where $\phi(x) = x - 0.5$

### 3.7. Example: Argentina rainfall data

Rainfall in Tucumán, Argentina, directly affects the economic well-being of this predominantly agricultural region. Annual rainfall data was collected over 113 consecutive years (1884-1996) for this region, and these records indicate a large jump in annual rainfall totals around the years 1955-1956. The construction of a dam is believed to have been the cause for this apparent irreversible change in annual rainfall amounts.

Wu, Woodroofe and Mentz (2001) examined the rainfall data and found that their test was strongly significant at the 0.01 level, as opposed to Brillinger's test which had a p-value of 0.0796. However, as discussed before, Wu's test stabilizes very slowly, which can result in invalid critical values for relatively small sample sizes.

We apply the $U_n^*$ test and the $U_n$ test to this data. For the $U_n^*$ test we obtain a p-value of 0.053, whilst for the $U_n$ test, we obtain a p-value of 0.048. Although the $U_n^*$ test narrowly fails to reject $H_0$ at the $\alpha = 0.05$ level, the p-values for these two tests are quite comparable, especially given the relatively small data set for which these tests were applied to.

## 4. MINIMAX APPROACH

Consider the functional form of $T_n^* = 4n^{-1} \sum_{k=1}^{n} [(\frac{k}{n+1} - \frac{1}{2})(F(X_k) - \frac{1}{2})]$. Defining $Y_i \triangleq F(X_i) - \frac{1}{2}$, we observe that $T_n^*$ is a linear combination of the $Y_i$. More specifically, $\{\frac{k}{n+1} - \frac{1}{2}\}_{k=1}^{n}$ is a contrast. With the preceding as motivation, we shall restrict our search for the "best" test to the class of test statistics possessing the following form:

$$T_n^*(c_1, ..., c_n) \triangleq \sum_{i=1}^{n} c_i X_i = \mathbf{c}^T \mathbf{X}, \tag{24}$$

subject to the restriction $\sum_{i=1}^{n} c_i = 0$, where $\mathbf{c}^T \triangleq (c_1, ..., c_n)$, $\mathbf{X}^T \triangleq (X_1, ..., X_n)$.

Also, let $\Gamma \triangleq \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{n-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{n-3} \\ \vdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{n-1} & \gamma_{n-2} & \gamma_{n-3} & \cdots & \gamma_0 \end{bmatrix}$,

which is the autocovariance matrix of $\mathbf{X}$, where $\gamma_k \triangleq \mathrm{cov}(X_1, X_{k+1})$. Finally, define

$$\Delta_1 \triangleq \{\boldsymbol{\delta} : \delta_1 \leq \delta_2 \leq \cdots \leq \delta_n\}, \tag{25}$$

$$\mathscr{C} \triangleq \{\mathbf{c} : \mathbf{c}^T \mathbb{1}_n = 0\}, \tag{26}$$

$$A \triangleq \{1, 2, \ldots, n-1\}, \tag{27}$$

whilst noting that $\boldsymbol{\delta}^T = (\delta_1, \ldots, \delta_n) \triangleq E_{H_a} \mathbf{X}^T$.

### 4.1. Derivation of a minimax criterion

Operating within the decision-theoretic framework, we shall use the 0-1 loss function. The corresponding risk function is

$$R(\boldsymbol{\delta}, T_n^*) = P_{H_0}(T_n^* \geq t) + P_{H_a}(T_n^* < t). \tag{28}$$

We shall use minimax as our optimality criteria; the $T_n^*$ within the class of contrasts that minimizes the maximum risk will be considered the "best", or equivalently, the minimax rule. So we seek to identify the contrast $\mathbf{c}$ that minimizes the maximum risk (0-1 loss):

$$\arg\min_{\{T_n^*\}} \max_{\{\boldsymbol{\delta}\in\Delta_1\}} R(\boldsymbol{\delta}, T_n^*) = \arg\min_{\{\mathbf{c}\in\mathscr{C}\}} \max_{\{\boldsymbol{\delta}\in\Delta_1\}} [P_{H_0}(T_n^* \geq t) + P_{H_a}(T_n^* < t)] \tag{29}$$

Observe that $E_{H_0}T_n^* = 0,\ \text{var}_{H_0}\, T_n^* = \mathbf{c}^T\Gamma\mathbf{c}$.

As in Chapter 3, by utilizing Corollary 1, we see that

$$\frac{\sqrt{n}(T_n^* - E_{H_0}T_n^*)}{\sqrt{\mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c}}} \xrightarrow{d} N(0,1). \tag{30}$$

Operating within the Neyman-Pearson framework of testing, we fix $t$ to achieve the desired Type I error rate. Asymptotically, the following can be shown:

**Proposition 8.**

$$\arg\min_{\{\mathbf{c}\in\mathscr{C}\}} \max_{\{\boldsymbol{\delta}\in\Delta_1\}} [P_{H_0}(T_n^* \geq t) + P_{H_a}(T_n^* < t)] = \arg\max_{\{\mathbf{c}\in\mathscr{C}\}} \min_{\{\boldsymbol{\delta}\in\Delta_1\}} \frac{\mathbf{c}^T\boldsymbol{\delta}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}}$$

Thus, the optimality criterion that any minimax rule $\mathbf{c}$ must satisfy is:

$$\arg\max_{\{\mathbf{c}\in\mathscr{C}\}} \min_{\{\boldsymbol{\delta}\in\Delta_1\}} \frac{\mathbf{c}^T\boldsymbol{\delta}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}} \tag{31}$$

Note that we are now dealing with a purely deterministic quantity.

### 4.2. Admissible $\boldsymbol{\delta}$: A simplification of the parameter space $\Delta_1$

Deriving $\arg\max_{\{\mathbf{c}\in\mathscr{C}\}} \min_{\{\boldsymbol{\delta}\in\Delta_1\}} \frac{\mathbf{c}^T\boldsymbol{\delta}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}}$ directly is not a trivial matter. However, the paper by Abelson and Tukey (1963, p. 1353) discusses a geometric characterization that greatly simplifies the space of admissible $\boldsymbol{\delta}$; i.e. all $\boldsymbol{\delta}$ that satisfy $\delta_1 \leq \delta_2 \leq \cdots \leq \delta_n$. In our setting, it turns out that we only need to consider a specific set of $n-1$ vectors when identifying where the minimum of $\frac{\mathbf{c}^T\boldsymbol{\delta}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}}$ with respect to $\boldsymbol{\delta}$ is potentially achieved. The characterization of these vectors is provided below within the change-point problem setting:

**Definition 2.** *If the solitary change-point occurs at $i = k+1$ (i.e. when $\delta_1 = \delta_2 = \cdots = \delta_k < \delta_{k+1} = \delta_{k+2} = \cdots$), where $k = 1, 2, \ldots, n-1$, then the corresponding standard corner vector is* $\boldsymbol{\delta}^{(k)} \triangleq (\underbrace{0, 0, \ldots,}_{k} \underbrace{1,}_{(k+1)th\ element} \underbrace{1, \ldots}_{n-k-1})$.

So for any candidate $\mathbf{c} \in \mathscr{C}$, the maximum risk will occur in the case when there is only a single change-point. This result can also be found within the linear programming framework, for example in Feiring (1986, p. 31). We state this as a theorem for our setting.

**Theorem 4.** *For any* $\mathbf{c} \in \mathscr{C}$, $\min_{\{\boldsymbol{\delta}\in\boldsymbol{\Delta}_1\}} \frac{\mathbf{c}^T\boldsymbol{\delta}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}} = \min_{\{i\in A\}} \frac{\mathbf{c}^T\boldsymbol{\delta}^{(i)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}}$.

Hence, the optimality criterion can be expressed as:

$$\arg\max_{\{\mathbf{c}\in\mathscr{C}\}} \min_{\{i\in A\}} \frac{\mathbf{c}^T\boldsymbol{\delta}^{(i)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}} \tag{32}$$

Next, we shall standardize the $\boldsymbol{\delta}^{(k)}$ in order to facilitate the derivation of future results. We end up with the following. Details are provided in the Appendix.

$$\acute{\boldsymbol{\delta}}^{(k)} \triangleq (\underbrace{-\{\tfrac{(n-k)}{kn}\}^{1/2}, -\{\tfrac{(n-k)}{kn}\}^{1/2}, \ldots,}_{k} \underbrace{\{\tfrac{k}{(n-k)n}\}^{1/2}}_{\text{(k+1)th element}}, \underbrace{\{\tfrac{k}{(n-k)n}\}^{1/2}, \ldots}_{n-k-1}) \tag{33}$$

## 4.3. Nature of the independent case minimax rule and derivation of the Brillinger coefficients

A further result in Abelson and Tukey (1963, p. 1354) states that any $\mathbf{c}$ that achieves $\max_{\{\mathbf{c}\in\mathscr{C}\}} \min_{\{i\in A\}}(\mathbf{c}^T\boldsymbol{\delta}^{(i)})$ must have the same dot product with *every* corner vector $\boldsymbol{\delta}^{(k)}$. Keep in mind that this corresponds to the independent random variable case. So we end up with the following system of equations that will determine such a $\mathbf{c}$ :

$$\begin{bmatrix} (\acute{\boldsymbol{\delta}}^{(1)})^T \\ (\acute{\boldsymbol{\delta}}^{(2)})^T \\ \vdots \\ (\acute{\boldsymbol{\delta}}^{(n-1)})^T \\ (\mathbb{1}_n)^T \end{bmatrix} \mathbf{c} = \begin{bmatrix} \zeta \\ \zeta \\ \vdots \\ \zeta \\ 0 \end{bmatrix}, \text{ where } \zeta \text{ is some constant.}$$

It can be shown that the coefficients derived in Abelson and Tukey (1963) and used by Brillinger, can be derived in an algebraic, as opposed to a geometric, manner. The aforementioned coefficients are simply the solution to the system of equations mentioned above.

**Proposition 9.** $\tilde{c}_j = \zeta \left\{ \left(\frac{(j-1)(n-(j-1))}{n}\right)^{1/2} - \left(\frac{j(n-j)}{n}\right)^{1/2} \right\}$, *for* $j = 1, 2, \ldots, n$.

**Definition 3.** *Let* $\mathbf{c}_0 \triangleq (\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_n)$. *These are the coefficients used by Brillinger.*

### 4.4. The structure of $\boldsymbol{\Delta}$

Now consider the subset of (the vector space) $\mathbf{R}^n$ defined as

$$\boldsymbol{\Delta} \triangleq \{\mathbf{c} \in \mathbf{R}^n : c_1 \leq c_2 \leq \cdots \leq c_n, \mathbf{c}^T \mathbb{1}_n = 0\}. \tag{34}$$

With a view to proving minimaxity, we will consider the structure of this subset. Throughout, it shall be tacitly assumed that any set or space considered is nonempty.

Notice that $\boldsymbol{\Delta}$ is a convex cone, since for any $a, b \in \mathbf{R}_+, \mathbf{x}, \mathbf{y} \in \boldsymbol{\Delta}$, it follows that $a\mathbf{x} + b\mathbf{y} \in \boldsymbol{\Delta}$. By Guler (2010, p. 94), this implies that $\boldsymbol{\Delta}$ itself is the convex conical hull of $\boldsymbol{\Delta}$. We state this result below:

**Lemma 5.** $\boldsymbol{\Delta}$ *is a convex cone. Moreover,* $\boldsymbol{\Delta}$ *itself is the convex conical hull of* $\boldsymbol{\Delta}$.

By a result owing to Carathéodory (1911), it follows that every element of the convex conical hull $\boldsymbol{\Delta}$ can be expressed as a conical combination of $n-1$ linearly independent elements of $\boldsymbol{\Delta}$. Again, we state this as a lemma for our situation:

**Lemma 6.** *Every element of the convex conical hull* $\boldsymbol{\Delta}$ *can be expressed as a conical combination of* $\dim(\mathrm{span}(\boldsymbol{\Delta})) = n-1$ *linearly independent elements of* $\boldsymbol{\Delta}$.

Observe that $\left\{\acute{\boldsymbol{\delta}}^{(1)}, \acute{\boldsymbol{\delta}}^{(2)}, \ldots, \acute{\boldsymbol{\delta}}^{(n-1)}\right\}$ are linearly independent. Hence, for any $\mathbf{c} \in \boldsymbol{\Delta}$, it follows that

$$\mathbf{c} = \sum_{i=1}^{n-1} a_i \acute{\boldsymbol{\delta}}^{(i)}, a_i \geq 0 \text{ for all } i \in A. \tag{35}$$

Now, clearly $\mathbf{c}_0 \in \boldsymbol{\Delta}$ so

$$\mathbf{c}_0 = \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}^{(i)}, w_i \geq 0 \text{ for all } i \in A. \tag{36}$$

### 4.5. A sufficient condition for minimaxity of $\mathbf{c}_0$

Utilizing the above representation of $\mathbf{c}_0$, we shall now demonstrate a sufficient condition for the minimaxity of $\mathbf{c}_0$ to hold.

First, we construct an inner product space as follows:

Define the inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle \triangleq \mathbf{x}^T \Gamma \mathbf{y},\ \mathbf{x}, \mathbf{y} \in \mathbf{R}^{n \times 1}. \tag{37}$$

The associated norm is then

$$\|\mathbf{x}\|_\Gamma \triangleq \langle \mathbf{x}, \mathbf{x} \rangle^{1/2}, \mathbf{x} \in \mathbf{R}^{n \times 1}. \tag{38}$$

**Proposition 10.** *For* $\mathbf{c} \in \mathscr{C}$, $\Gamma$ *symmetric positive definite, if* $\Gamma \mathbf{c}_0 \in \boldsymbol{\Delta}$, *and there exist* $j, l \in A$ *such that* $\min_{k \in A} \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(k)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} = \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} < \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}}$, *then* $\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} < \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}}$ .

**Proposition 11.** *For* $\mathbf{c} \in \mathscr{C}$, $\Gamma$ *symmetric positive definite, if* $\Gamma \mathbf{c}_0 \in \boldsymbol{\Delta}$, *and*

$\frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}} = \frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}}$ *for all* $i, l \in A$*, then* $\frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}} = \frac{\mathbf{c}_0^T\acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}_0^T\Gamma\mathbf{c}_0)^{1/2}}$ *for all* $i, l \in A$*.*

Define

$$\mathscr{A} \triangleq \left\{\mathbf{c} \in \mathrm{R}^{n\times 1} : \frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}} = \frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}} \text{ for all } i, l \in A\right\}. \tag{39}$$

**Theorem 5.** *Let the random process be defined as in Equation* (1)*. If* $\Gamma\mathbf{c}_0 \in \boldsymbol{\Delta}$*, then* $\mathbf{c}_0$ *is minimax. Moreover, membership in* $\mathscr{A}$ *is a necessary and sufficient condition for* $\mathbf{c} \in \mathscr{C}$ *to be minimax.*

**Corollary 2.** $\mathbf{c}_0$ *is the unique minimax, up to a positive scaling.*

All the following statements are contingent on the condition $\Gamma\mathbf{c}_0 \in \Delta$: Proposition 10 states that any candidate $\mathbf{c}$ that does not have an equal dot product with all of the corner vectors $\acute{\boldsymbol{\delta}}^{(i)}$ is not admissible. Proposition 11 states that any candidate $\mathbf{c}$ that does have an equal dot product with all of the corner vectors $\acute{\boldsymbol{\delta}}^{(i)}$ matches, but does not beat $\mathbf{c}_0$. Theorem 6 claims that $\mathbf{c}_0$ is indeed minimax. Corollary 2 says that any contrast that is a positive multiple of $\mathbf{c}_0$ is also minimax.

Of course, if $\Gamma = \sigma^2 I_n$, then the condition $\Gamma\mathbf{c}_0 \in \Delta$ is satisfied, since by the construction of $\acute{\boldsymbol{\delta}}^{(j)}$ in (33), without loss of generality we can set $\sigma^2 = 1$, and by the monotonicity of $\mathbf{c}_0$. Since this is an important special case, we restate the main result in terms of this situation:

**Theorem 6.** *Let the random process be defined as in Equation* (1)*, with autocovariance matrix* $\Gamma = \sigma^2 I_n$*. Then* $\mathbf{c}_0$ *is minimax. Moreover, membership in* $\mathscr{A}$ *is a necessary and sufficient condition for* $\mathbf{c} \in \mathscr{C}$ *to be minimax.*

Hence, we have proven that Brillinger's test for monotonic trend is actually minimax in the independent case, given that it is a member of the class of test statistics outlined in (24).

### 4.6. Simulation: example where Brillinger's test is minimax

All the simulation details discussed in the previous chapter carry over here. Notice in Figure 3 that Brillinger's test is the most powerful for this particular configuration of $\delta_i$; one that corresponds to a single change-point occurring at the sixth observation. In Figure 4, we increase the sample size to 1,000 observations in the case where the change point occurs at the third observation. We examine, in addition to Brillinger's test and Wu's test, the performance of $T_n^*$. Recall that finding the minimax rule within the class of contrasts was originally motivated by the functional form of the projection $T_n^*$. Again, Brillinger's test performs the best for this early-change-point setting. Noting that Wu's test has a simulated type I error rate of 0.0618, both the aforementioned test and the $T_n^*$ test perform poorly.

Also recall the discussion regarding the fact that the maximum risk will occur at one of the corner vectors. As mentioned before, any single change-point case corresponds to one of the corner vectors. Intuitively, any test will have difficulty detecting a monotonic trend if the singular change-point occurs near the beginning and end time points (finite sample case).

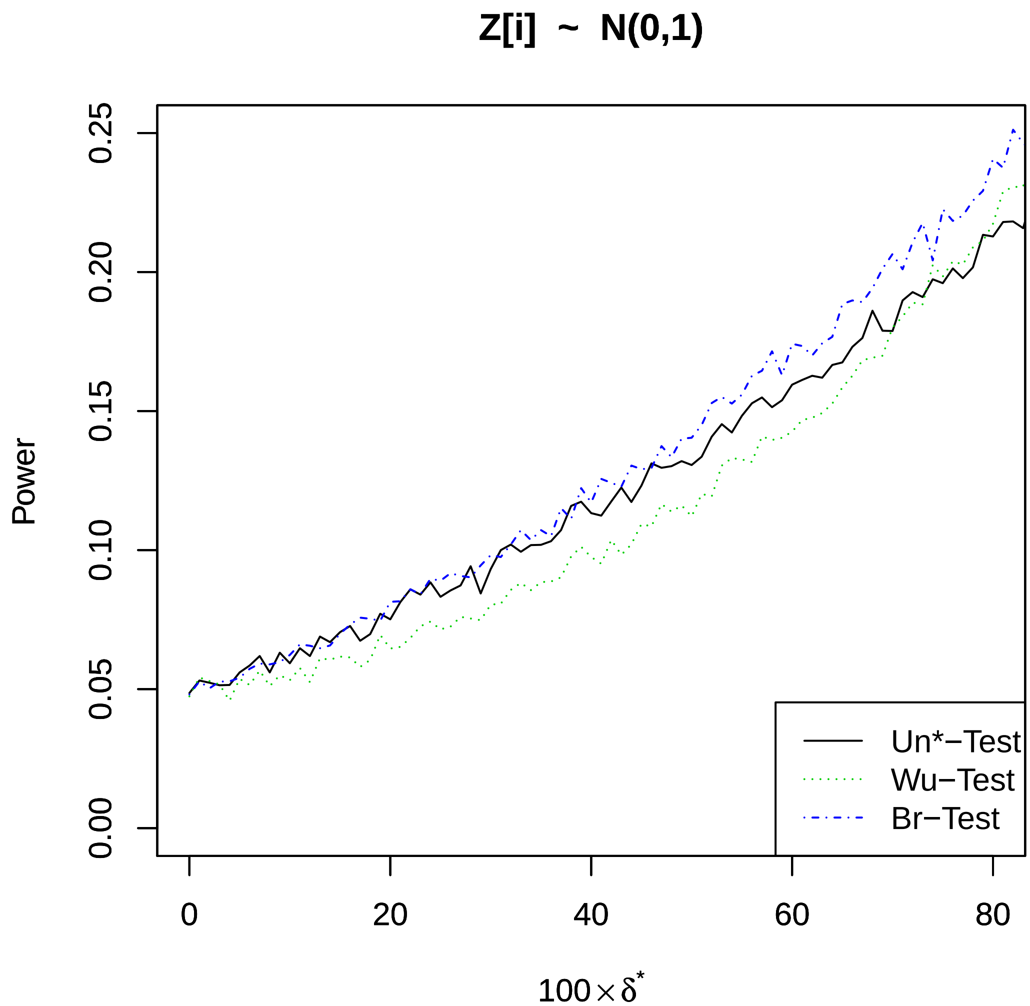


Figure 3. Power curves for $U_n^*$, Wu's test, and Brillinger's test where $\phi(x) = \mathbb{1}_{\{x>0.025\}} - 0.5$

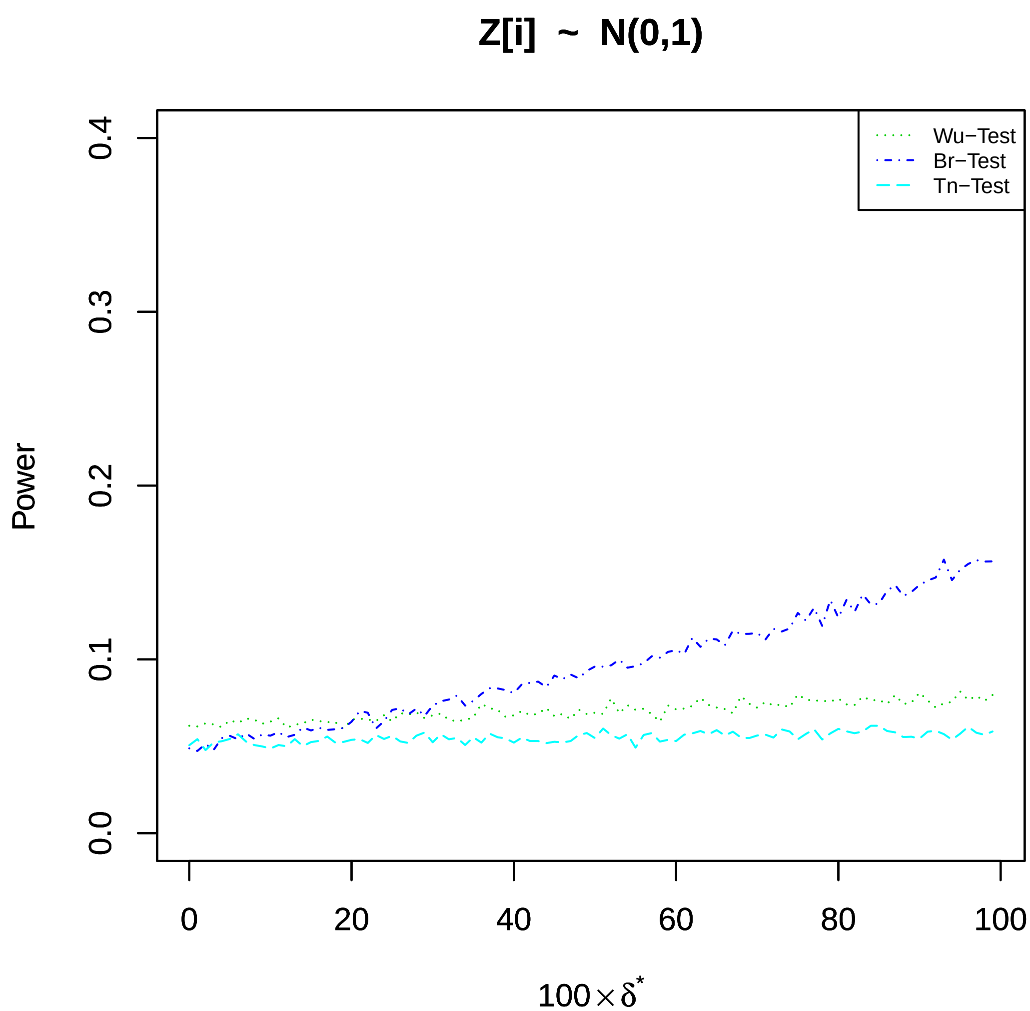


Figure 4. Power curves for $T_n^*$, Wu's test, and Brillinger's test where $\phi(x) = \mathbb{1}_{\{x>0.001\}} - 0.5$

### 4.7. An alternate characterization of the minimax rule

We have already exhibited the minimax rule in the preceding section. However, the following discussion and characterization frames the problem in a slightly different manner. Hence, for theoretical and pedagogical reasons we reproduce it here.

From Lemma 6, it follows that

$$\Gamma \mathbf{c}_0 = \sum_{i=1}^{n-1} w_i \Gamma \acute{\boldsymbol{\delta}}^{(i)} \tag{40}$$

Define $\Gamma\boldsymbol{\Delta} \triangleq \{\Gamma\mathbf{c} : \mathbf{c} \in \boldsymbol{\Delta}\}$ Using a result given in Rockafellar (1970, p. 19), it follows that $\Gamma\boldsymbol{\Delta}$ is also a convex cone, and thus is a convex conical hull. We state this as a lemma:

**Lemma 7.** $\Gamma\boldsymbol{\Delta}$ *is a convex conical hull.*

Now define $Y \triangleq \left\{ \acute{\boldsymbol{\delta}}^{(1)} \; \acute{\boldsymbol{\delta}}^{(2)} \; \cdots \; \acute{\boldsymbol{\delta}}^{(n-1)} \right\} \in \mathbf{R}^{n\times(n-1)}$
Consider the mapping $T(\Gamma) : \mathbf{R}^{n\times n} \to \mathbf{R}^{n\times(n-1)} \triangleq \Gamma Y$, whose domain is the set of all correlation matrices $\Gamma \in \mathbf{R}^{n\times n}$.
Recall that $\Gamma$ is the correlation matrix of the random process. It is positive-definite and thus of full rank. We now demonstrate that $\Gamma\boldsymbol{\Delta}$ is full rank (Dattoro 2013, p. 37).

**Lemma 8.** *The columns of* $\Gamma Y$ *are linearly independent.*

As a side note, observe that by Lemma A9, $\text{rank}(\Gamma Y) = n - 1$.

**Lemma 9.** *Any* $\mathbf{c} \in \Gamma\boldsymbol{\Delta}$ *can be expressed as a conical combination of the elements of* $\left\{ \Gamma\acute{\boldsymbol{\delta}}^{(i)}, i \in A \right\}$

**Definition 4.** *The generators for a closed convex cone* $C$ *are any collection of directions whose convex conical hull constructs* $C$*.*

**Lemma 10.** $\left\{ \Gamma\acute{\boldsymbol{\delta}}^{(i)}, i \in A \right\}$ *comprise the generators for* $\Gamma\boldsymbol{\Delta}$*. Moreover, this is a minimal set of generators for* $\Gamma\boldsymbol{\Delta}$*.*

**Definition 5.** *An extreme direction for a pointed closed convex cone $C$ is any nonnegative scaling of a vector that cannot be expressed as a conic combination of any other vectors in $C$.*

**Definition 6.** *A set of directions is said to be conically independent if no direction from the set can be expressed as a conic combination of the remaining directions.*

Clearly, linear independence is a stronger property.
The following fact is given in Dattoro (2013, p. 145):

**Lemma 11.** *When a set of conically independent directions from a pointed closed convex cone $C$ is comprised of generators, then all of these directions must be extreme directions of $C$.*

**Lemma 12.** $\left\{\Gamma\acute{\boldsymbol{\delta}}^{(i)}, i \in A\right\}$ *are the extreme directions of* $\Gamma\boldsymbol{\Delta}$.

Define

$$\acute{\boldsymbol{\delta}}_{\Gamma}^{(i)} \triangleq \Gamma\acute{\boldsymbol{\delta}}^{(i)} \text{ for all } i \in A \tag{41}$$

Since

$$\mathbf{c}_0 = \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}^{(i)}, w_i \geq 0 \text{ for all } i \in A, \tag{42}$$

it follows that

$$\Gamma\mathbf{c}_0 = \sum_{i=1}^{n-1} w_i \Gamma\acute{\boldsymbol{\delta}}^{(i)} = \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}_{\Gamma}^{(i)} \tag{43}$$

Notice that the problem now has the same structure as the one discussed in Abelson and Tukey (1963). In other words, for a contrast $\mathbf{c} \in \mathbf{R}^n$ and a vector $\mathbf{a}$ in $\mathbf{R}^n$ belonging to the convex cone $C$, the minimum value of $\mathbf{c}^T\mathbf{a}$ for a fixed $\mathbf{c}$ must be attained at one of the extreme directions (Abelson and Tukey, 1963, p. 1353).

The extreme directions are referred to as corner vectors in the aforementioned paper. We state this as a theorem for our generalized setting.

**Theorem 7.** *For any* $\mathbf{c} \in \mathscr{C}$, $\min_{\{\boldsymbol{\delta} \in \Gamma \boldsymbol{\Delta}\}} \mathbf{c}^T \boldsymbol{\delta} = \min_{\{i \in A\}} \mathbf{c}^T \acute{\boldsymbol{\delta}}_{\Gamma}^{(i)}$

**Proposition 12.** *For* $\mathbf{c} \in \mathscr{C}$, $\Gamma$ *symmetric positive definite,* $\Gamma \mathbf{c}_0 \in \Gamma \boldsymbol{\Delta}$, *and* $\mathbf{c}_0^T \boldsymbol{\delta}_{\Gamma}^{(i)} = \mathbf{c}_0^T \boldsymbol{\delta}_{\Gamma}^{(l)}$ *for all* $i, l \in A$, *if there exist* $j, l \in A$ *such that* $\min_{k \in A} \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(k)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} = \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} < \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}}$, *then* $\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} < \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}}$ .

**Proposition 13.** *For* $\mathbf{c} \in \mathscr{C}$, $\Gamma$ *symmetric positive definite,* $\Gamma \mathbf{c}_0 \in \Gamma \boldsymbol{\Delta}$, *and* $\mathbf{c}_0^T \boldsymbol{\delta}_{\Gamma}^{(i)} = \mathbf{c}_0^T \boldsymbol{\delta}_{\Gamma}^{(l)}$ *for all* $i, l \in A$, , *if* $\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} = \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}}$ *for all* $i, l \in A$, *then* $\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} = \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}}$ *for all* $i, l \in A$.

**Theorem 8.** *Let the random process be defined as in Equation* (1). *If* $\mathbf{c}_0^T \boldsymbol{\delta}_{\Gamma}^{(i)} = \mathbf{c}_0^T \boldsymbol{\delta}_{\Gamma}^{(l)}$ *for all* $i, l \in A$, *then* $\mathbf{c}_0$ *is minimax. Moreover, membership in* $\mathscr{A}$ *is a necessary and sufficient condition for* $\mathbf{c} \in \mathscr{C}$ *to be minimax.*

**Corollary 3.** $\mathbf{c}_0$ *is the unique minimax, up to a positive scaling.*

# REFERENCES


Abelson, R. P., and Tukey, J. W. (1963), "Efficient utilization of non-numerical information in quantitative analysis: general theory and the case of simple order," *The Annals of Mathematical Statistics*, 34, 4, 1347-1369.

Alvarez, E. E., and Dey, D. K. (2009), "Bayesian isotonic changepoint analysis," *Ann Inst Stat Math*, **61**, 355-370.

Athreya, K. B., and Lahiri, S. N. (2010), *Measure Theory and Probability Theory*, Springer Science+Business Media, LLC.

Billingsley, P. (1995), *Probability and Measure* (3rd ed.), New York: John Wiley & Sons.

Brillinger, D. R. (1989), "Consistent detection of a monotonic trend superimposed on a stationary time series," *Biometrika*, 76, 1, 23-30.

Carathéodory, C. (1911), "Uber den Variabilitätsbereich der Fourier'schen Konstanten von positiven harmonischen Funktionen," *Rend. Circ. Mat. Palermo* **32**, 193-217.

Dattoro, J. (2013), *Convex Optimization & Euclidean Distance Geometry*, Meboo Publishing.

Doukhan, P. (1985), *Mixing: Properties and Examples*, Lecture Notes in Statistics 85, New York: Springer-Verlag.

Feiring, B. (1986), *Linear Programming: An Introduction*, Thousand Oaks, CA: SAGE Publications, Inc.

Güler, O. (2010), *Foundations of Optimization*, New York: Springer.

Herrndorf, N. (1984), "A Functional Central Limit Theorem for Weakly Dependent Sequences on Random Variables," *The Annals of Probability*, 12, 1, 141-153.

Jonckheere, A. R. (1954). "A distribution-free k-sample test against ordered alternatives," *Biometrika*, **41**, 133-145.

Mann, H. B., and Whitney, D. R. (1947), "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other," *Annals of Mathematical Statistics*, **18**, 1, 50-60.

Rockafellar, R. T. (1970), *Convex Analysis*, Princeton University Press.

Shen, G., and Xu, H. (2013), "On the isotonic change-point problem," *Journal of Nonparametric Statistics*, DOI: 10.1080/10485252.2013.821472. Available at http://dx.doi.org/10.1080/10485252.2013.821472.

van de Vaart, A. W. (1998), *Asymptotic Statistics* (1st ed.), Cambridge: Cambridge University Press.

Wu, W. B., Woodroofe, M., and Mentz, G. (2001), "Isotonic Regression: Another look at the change-point problem," *Biometrika*, 88, 3, 793-804.

Yoshihara, K. (1976), "Limiting Behavior of U-Statistics for Stationary Absolutely Regular Processes", *Z. Wahrscheinlichkeitstheorieverw. Gebiete*, 35, 237-252.

## APPENDIX. PROOFS

Let P be the appropriate probability measure defined on $(\mathbb{R}, \sigma(X_i))$, and let $\lambda$ be Lebesgue measure defined on $(\mathbb{R}, \mathscr{B})$. By the Radon-Nikodym Theorem, since $P \ll \lambda$ ($P$ absolutely continuous with respect to $\lambda$; in other words, $\forall\ A \in \sigma(X_i), \lambda(A) = 0$ implies $P(A) = 0$), it follows that there exists an unique Borel-measurable function $f \geq 0$ ($P$ $a.s.$) such that $\forall A \in \sigma(X_i), P(A) = \int_A f d\lambda$. Here $f$ is the Radon-Nikodym derivative of P with respect to $\lambda$ ($f \triangleq \frac{dP}{d\lambda}$).

**Proposition A1.** *$E(\mathbb{1}_{\{X_i<X_j\}}|X_i) = 1 - F(X_i)$.*

*Proof.*

$$
\begin{aligned}
E(\mathbb{1}_{\{X_i<X_j\}}|X_i) &= \int \mathbb{1}_{\{X_i<X_j\}} dP && (44)\\
&= \int_{\mathbb{1}_{\{X_i<X_j\}}} dP && (45)\\
&= \int_{\mathbb{1}_{\{X_i<t\}}} f(t|x_i) d\lambda && (46)\\
&= \int_{X_i}^{+\infty} f(t|x_i) d\lambda && (47)\\
&= \int_{X_i}^{+\infty} f(t) d\lambda && (48)\\
&= P\{X_i < X < +\infty\} && (49)\\
&= 1 - P\{-\infty < X < X_i\} && (50)\\
&= 1 - F(X_i) && (51)
\end{aligned}
$$

□

Again let P be the appropriate probability measure.

**Proposition A2.** $E(\mathbb{1}_{\{X_i<X_j\}}|X_j) = F(X_j)$.

*Proof.*

$$\begin{aligned}
E(\mathbb{1}_{\{X_i<X_j\}}|X_j) &= \int \mathbb{1}_{\{X_i<X_j\}} dP && (52)\\
&= \int_{\mathbb{1}_{\{X_i<X_j\}}} dP && (53)\\
&= \int_{\mathbb{1}_{\{t<X_j\}}} f(t|x_j)d\lambda && (54)\\
&= \int_{-\infty}^{X_j} f(t|x_j)d\lambda && (55)\\
&= \int_{-\infty}^{X_j} f(t)d\lambda && (56)\\
&= P\{-\infty < X < X_j\} && (57)\\
&= F(X_j) && (58)
\end{aligned}$$

□

**Proposition A3.** $EU_n^* = 0$

*Proof.*

$$\begin{aligned}
EU_n^* &= \frac{n-1}{n+1}\binom{n}{2}^{-1} E\left[\sum_{i<j}(\mathbb{1}_{\{X_i<X_j\}} - \tfrac{1}{2})\right] && (59)\\
&= \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{i<j} E(\mathbb{1}_{\{X_i<X_j\}} - \tfrac{1}{2}) && (60)\\
&= \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{i<j} (P\{X_i < X_j\} - \tfrac{1}{2}) && (61)\\
&= 0 && (62)
\end{aligned}$$

□

So by Proposition A3,

$$T_n^* = \sum_{k=1}^{n} E(U_n|X_k) \tag{63}$$

$$= \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{k=1}^{n} E\left[\sum_{i<j} (\mathbb{1}_{\{X_i<X_j\}} - \tfrac{1}{2})\,|X_k\right] \tag{64}$$

**Proposition A4.** *If $k \neq i, j$, then* $E\left[\sum_{i<j} (\mathbb{1}_{\{X_i<X_j\}} - \frac{1}{2})\,|X_k\right] = 0$.

*Proof.* Note if $k \neq i, j$,

$$E(\sum_{i<j} (\mathbb{1}_{\{X_i<X_j\}} - \tfrac{1}{2})\,|X_k) = \sum_{i<j} E(\mathbb{1}_{\{X_i<X_j\}} - \tfrac{1}{2}|X_k) \tag{65}$$

$$= \sum_{i<j} E(\mathbb{1}_{\{X_i<X_j\}} - \tfrac{1}{2}) \tag{66}$$

$$= 0, \tag{67}$$

where (66) is due to the independence of the $\sigma$-fields in question. □

**Proposition A5.** $T_n^* = 4n^{-1} \sum_{k=1}^{n} [(\frac{k}{n+1} - \frac{1}{2})(F(X_k) - \frac{1}{2})]$.

*Proof.*

$$T_n^* = \sum_{k=1}^{n} E(U_n|X_k) \tag{68}$$

$$= \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{k=1}^{n} E\left[\sum_{i<j} (\mathbb{1}_{\{X_i<X_j\}} - \tfrac{1}{2})\,|X_k\right] \tag{69}$$

$$\tag{70}$$

Consider the outer summand for each value of $k$.

Utilizing Propositions A1 and A2, for:

$$k = 1 : n - 1 \text{ terms of } (1 - F(X_1)) - \tfrac{1}{2} = \tfrac{1}{2} - F(X_1) \tag{71}$$

$$k = 2 : n - 2 \text{ terms of } \tfrac{1}{2} - F(X_2), \text{ and 1 term of } F(X_2) - \tfrac{1}{2} \tag{72}$$

In general, for:

$$0 < q < n + 1 : n - q \text{ terms of } \tfrac{1}{2} - F(X_q), \text{ and } q - 1 \text{ terms of } F(X_q) - \tfrac{1}{2} \tag{73}$$

Therefore,

$$T_n^* = \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{k=1}^{n} [(n-k)(\tfrac{1}{2} - F(X_k)) + (k-1)(F(X_k) - \tfrac{1}{2})] \tag{74}$$

$$= \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{k=1}^{n} [((n-k) - (k-1))(\tfrac{1}{2} - F(X_k))] \tag{75}$$

$$= \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{k=1}^{n} [(n-2k+1)(\tfrac{1}{2} - F(X_k))] \tag{76}$$

$$= \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{k=1}^{n} [(2k-n-1)(F(X_k) - \tfrac{1}{2})] \tag{77}$$

$$= 4n^{-1} \sum_{k=1}^{n} [(\tfrac{k}{n+1} - \tfrac{1}{2})(F(X_k) - \tfrac{1}{2})] \tag{78}$$

□

**Proposition A6.** $\text{var}(T_n^*) = \frac{n-1}{9n(n+1)}$.

*Proof.* Using the expression for $T_n^*$ given in Proposition A5,

$$\text{var}(T_n^*) = 16n^{-2} \,\text{var}\left[\sum_{k=1}^{n} [(\tfrac{k}{n+1} - \tfrac{1}{2})(F(X_k) - \tfrac{1}{2})]\right] \tag{79}$$

$$= 16n^{-2}\left[\sum_{k=1}^{n} \text{var}[(\tfrac{k}{n+1} - \tfrac{1}{2})(F(X_k) - \tfrac{1}{2})]\right] \tag{80}$$

$$= 16n^{-2}\left[\sum_{k=1}^{n} \text{var}[(\tfrac{k}{n+1} - \tfrac{1}{2})F(X_k)]\right] \tag{81}$$

$$= \tfrac{4}{3}n^{-2}\left[\sum_{k=1}^{n} (\tfrac{k}{n+1} - \tfrac{1}{2})^2\right] \tag{82}$$

$$= \tfrac{n-1}{9n(n+1)} \tag{83}$$

□

Define

$$\gamma_k \triangleq \text{cov}(F(X_1), F(X_{k+1})), \quad k = 0, 1, \ldots \tag{84}$$

$$\sigma^2 \triangleq \gamma_0 + 2\sum_{k=1}^{+\infty} \gamma_k \neq 0, \tag{85}$$

$$Y_i \triangleq (\tfrac{i}{n+1} - \tfrac{1}{2})(F(X_i) - \tfrac{1}{2}) \tag{86}$$

$$T_n^* \triangleq 4n^{-1}\sum_{i=1}^{n} Y_i \tag{87}$$

In order to demonstrate the asymptotic normality of $T_n^*$, we shall utilize the following central limit theorem for weakly dependent sequences owed to Herrndoff (1984, p.142):

**Corollary A1.** *Let $q > 2$ and $p = 2/q$. If $\text{var}[\sqrt{n}(T_n^* - 0)] \to \sigma^2 > 0$, $EY_j = 0, EY_j^2 < +\infty$, and there exists $q > 2$ such that $\limsup_{n\to\infty} \|Y_n\|_q < +\infty$ where $\|Y_n\|_q \triangleq [E|Y_n|^q]^{1/q}$, and $\sum_{k=1}^{+\infty} \beta_k^{1-p} < +\infty$, then $\sqrt{n}(\frac{\sqrt{3}}{2}T_n^* - 0) \xrightarrow{d} N(0, \sigma^2)$.*

These conditions for Corollary A1 to hold for $T_n^*$ will now be verified.

**Lemma A1.** $\text{var}[\sqrt{n}(\frac{\sqrt{3}}{2}T_n^* - 0)] \to \sigma^2 > 0$.

*Proof.*

$$\mathrm{var}[\sqrt{n}(T_n - 0)] \tag{88}$$

$$= \mathrm{var}[\sqrt{n}T_n] \tag{89}$$

$$= n\,\mathrm{var}(T_n) \tag{90}$$

$$= 16n^{-1}\,\mathrm{var}\left[\sum_{i=1}^{n} Y_i\right] \tag{91}$$

$$= 16n^{-1}\,\mathrm{var}\left[\sum_{i=1}^{n}(\tfrac{i}{n+1} - \tfrac{1}{2})(F(X_i) - \tfrac{1}{2})\right] \tag{92}$$

$$= 16n^{-1}\,\mathrm{var}\left[\sum_{i=1}^{n}(\tfrac{i}{n+1} - \tfrac{1}{2})F(X_i)\right] \tag{93}$$

$$= 16n^{-1}\left[\gamma_0 \sum_{i=1}^{n}(\tfrac{i}{n+1} - \tfrac{1}{2})^2 + 2\sum_{i<j}(\tfrac{i}{n+1} - \tfrac{1}{2})(\tfrac{j}{n+1} - \tfrac{1}{2})\,\mathrm{cov}(F(X_i), F(X_j))\right] \tag{94}$$

$$= 16n^{-1}\left[\gamma_0 \sum_{i=1}^{n}(\tfrac{i}{n+1} - \tfrac{1}{2})^2 + 2\sum_{i<j}(\tfrac{i}{n+1} - \tfrac{1}{2})(\tfrac{j}{n+1} - \tfrac{1}{2})\gamma_{j-i}\right] \tag{95}$$

$$= \left[\tfrac{8(2n+1)}{3(n+1)} - 4\right]\gamma_0 + 32n^{-1}\sum_{i<j}(\tfrac{i}{n+1} - \tfrac{1}{2})(\tfrac{j}{n+1} - \tfrac{1}{2})\gamma_{j-i} \tag{96}$$

$$= \left[\tfrac{4(n-1)}{3(n+1)}\right]\gamma_0 + 32n^{-1}\sum_{i<j}(\tfrac{i}{n+1} - \tfrac{1}{2})(\tfrac{j}{n+1} - \tfrac{1}{2})\gamma_{j-i} \tag{97}$$

$$= \left[\tfrac{4(n-1)}{3(n+1)}\right]\gamma_0 + 32n^{-1}\sum_{i<j}(\tfrac{i}{n+1} - \tfrac{1}{2})[(\tfrac{i}{n+1} - \tfrac{1}{2}) + (\tfrac{j-i}{n+1})]\gamma_{j-i} \tag{98}$$

$$= \left[\tfrac{4(n-1)}{3(n+1)}\right]\gamma_0 + 32n^{-1}\sum_{i<j}[(\tfrac{i}{n+1} - \tfrac{1}{2})^2 + (\tfrac{i}{n+1} - \tfrac{1}{2})(\tfrac{j-i}{n+1})]\gamma_{j-i} \tag{99}$$

$$= \left[\tfrac{4(n-1)}{3(n+1)}\right]\gamma_0 + 32n^{-1}\sum_{k>0}[(\tfrac{i}{n+1} - \tfrac{1}{2})^2 + (\tfrac{i}{n+1} - \tfrac{1}{2})(\tfrac{k}{n+1})]\gamma_k \tag{100}$$

$$= \left[\tfrac{4(n-1)}{3(n+1)}\right]\gamma_0 + 32n^{-1}\sum_{k=1}^{n-1}\sum_{i=1}^{n-k}[(\tfrac{i}{n+1} - \tfrac{1}{2})^2 + (\tfrac{i}{n+1} - \tfrac{1}{2})(\tfrac{k}{n+1})]\gamma_k \tag{101}$$

$$= \left[\tfrac{4(n-1)}{3(n+1)}\right]\gamma_0 + 32n^{-1}\left[\sum_{k=1}^{n-1}\sum_{i=1}^{n-k}(\tfrac{i}{n+1}-\tfrac{1}{2})^2\gamma_k + \sum_{k=1}^{n-1}\sum_{i=1}^{n-k}(\tfrac{i}{n+1}-\tfrac{1}{2})(\tfrac{k}{n+1})\gamma_k\right] \quad (102)$$

$$= \left[\tfrac{4(n-1)}{3(n+1)}\right]\gamma_0 + 32n^{-1}\left[\sum_{k=1}^{n-1}\gamma_k\sum_{i=1}^{n-k}(\tfrac{i}{n+1}-\tfrac{1}{2})^2 + \sum_{k=1}^{n-1}(\tfrac{k}{n+1})\gamma_k\sum_{i=1}^{n-k}(\tfrac{i}{n+1}-\tfrac{1}{2})\right] \quad (103)$$

$$= \left[\tfrac{4(n-1)}{3(n+1)}\right]\gamma_0 + 32n^{-1}\left[\sum_{k=1}^{n-1}\gamma_k\tfrac{-4k^3+6nk^2+(1-3n^2)k+n^3-n}{12(n+1)^2} + \sum_{k=1}^{n-1}(\tfrac{k}{n+1})\gamma_k\tfrac{k^2-nk}{2(n+1)}\right] \quad (104)$$

$$= \left[\tfrac{4(n-1)}{3(n+1)}\right]\gamma_0 + 8n^{-1}\left[\sum_{k=1}^{n-1}\gamma_k\tfrac{-4k^3+6nk^2+(1-3n^2)k+n^3-n}{3(n+1)^2} + \sum_{k=1}^{n-1}\gamma_k\tfrac{6k^3-6nk^2}{3(n+1)^2}\right] \quad (105)$$

$$= \left[\tfrac{4(n-1)}{3(n+1)}\right]\gamma_0 + 8n^{-1}\left[\sum_{k=1}^{n-1}\gamma_k\tfrac{2k^3+(1-3n^2)k+n^3-n}{3(n+1)^2}\right] \quad (106)$$

$$= \left[\tfrac{4(n-1)}{3(n+1)}\right]\gamma_0 + \tfrac{8(n^2-1)}{3(n+1)^2}\left[\sum_{k=1}^{n-1}\gamma_k\right] + \left[\sum_{k=1}^{n-1}\gamma_k\tfrac{16k^3+(8-24n^2)k}{3n(n+1)^2}\right] \quad (107)$$

$$= \left[\tfrac{4(n-1)}{3(n+1)}\right]\left[\gamma_0 + 2\sum_{k=1}^{n-1}\gamma_k\right] + \left[\sum_{k=1}^{n-1}k\gamma_k\tfrac{16k^2+(8-24n^2)}{3n(n+1)^2}\right]. \quad (108)$$

Note that $|\gamma_k| \leq 4\beta_k$ (Doukhan 1985, sec.1.2.2, Lemma 3). By the preceding and (2), since $\beta_n = o(n^{-1-\delta})$, $o(|\gamma_n|) = o(n^{-1-\delta})$ also.

Now,

$$\left|\sum_{k=1}^{n-1}k\gamma_k\tfrac{16k^2+8-24n^2}{3n(n+1)^2}\right| \leq \sum_{k=1}^{n-1}k|\gamma_k|\left|\tfrac{16n^2+8-24n^2}{3n(n+1)^2}\right| \quad (109)$$

$$\leq \sum_{k=1}^{n-1}k|\gamma_k|\tfrac{8n^2}{3n(n+1)^2} \quad (110)$$

$$\leq \tfrac{8n^2}{3n(n+1)^2}\sum_{k=1}^{n-1}k|\gamma_k| \quad (111)$$

$$\leq \tfrac{32n^2}{3n(n+1)^2}\sum_{k=1}^{n-1}k\beta_k \quad (112)$$

$$= O\left[\tfrac{1}{n}\sum_{k=1}^{n-1}k\beta_k\right] \quad (113)$$

$$= o(1) \tag{114}$$

Equation (114) follows from Equations (165)-(170).

Noting that

$$\lim_{n\to+\infty}\left[\gamma_0 + 2\sum_{k=1}^{n-1}\gamma_k\right] = \gamma_0 + 2\sum_{k=1}^{+\infty}\gamma_k = \sigma^2, \tag{115}$$

it follows that

$$\text{var}[\sqrt{n}(T_n - 0)] \to \tfrac{4}{3}\sigma^2 > 0 \tag{116}$$

Of course, this implies that

$$\text{var}[\sqrt{n}(\tfrac{\sqrt{3}}{2}T_n^* - 0)] \to \sigma^2 > 0, \tag{117}$$

as required. □

**Lemma A2.** $EY_j = 0, EY_j^2 < +\infty$

*Proof.* Under $H_0$,

$$EY_j = E((\tfrac{j}{n+1} - \tfrac{1}{2})(F(X_j) - \tfrac{1}{2})) \tag{118}$$

$$= (\tfrac{j}{n+1} - \tfrac{1}{2})E(F(X_j) - \tfrac{1}{2}) \tag{119}$$

$$= 0 \tag{120}$$

Trivially, $EY_j^2 = \text{var}(Y_j) < +\infty$. □

**Lemma A3.** *There exists* $q > 2$ *such that* $\limsup_{n\to\infty}[E|Y_n|^q]^{1/q} < +\infty$, *and* $\sum_{k=1}^{+\infty}\beta_k^{1-p} < +\infty$, *where* $p \triangleq 2/q$.

*Proof.* As motivated by the formulation of Lemma A5, consider $0 < \delta_1 < \delta$. Then, $0 < \frac{1+\delta}{\delta} < \frac{1+\delta_1}{\delta_1}$, which implies $0 < \frac{\delta_1}{1+\delta_1} < \frac{\delta}{1+\delta}$. Therefore, from Lemma A5, it is known that $\sum_{k=1}^{+\infty} \beta_k^{1-(\delta_1/1+\delta_1)} < +\infty$. So let $p = \frac{\delta_1}{1+\delta_1}$. Since $p = 2/q$ by definition, this implies $q = \frac{2(1+\delta_1)}{\delta_1}$, which satisfies the first claim, by Lemma A2. □

**Theorem A1.** $\sqrt{n}(\frac{\sqrt{3}}{2}T_n^* - 0) \xrightarrow{d} N(0, \sigma^2)$.

*Proof.* This immediately follows by Lemmas A1-A3 and Corollary A1. □

**Proposition A7.** $T_n^* = \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{i<j}(F(X_j) - F(X_i))$

*Proof.*

$$T_n^* = \sum_{k=1}^{n} E(U_n^*|X_k) \tag{121}$$

$$= \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{k=1}^{n} E(\sum_{i<j} (\mathbb{1}_{\{X_i<X_j\}} - \tfrac{1}{2})\,|X_k) \tag{122}$$

$$= \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{i<j} E(\sum_{k=1}^{n} (\mathbb{1}_{\{X_i<X_j\}} - \tfrac{1}{2})\,|X_k) \tag{123}$$

$$= \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{i<j}(\sum_{k=1}^{n} E(\mathbb{1}_{\{X_i<X_j\}} - \tfrac{1}{2}\,|X_k)) \tag{124}$$

$$= \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{i<j}(1 - F(X_i) - \tfrac{1}{2} + F(X_j) - \tfrac{1}{2}) \tag{125}$$

$$= \frac{n-1}{n+1}\binom{n}{2}^{-1} \sum_{i<j}(F(X_j) - F(X_i)) \tag{126}$$

□

**Theorem A2.** *Under* $H_0$, $\sqrt{n}(R_n^* - 0) \xrightarrow{p} 0$

*Proof.* We shall demonstrate that $E(nR_n^{*2}) = o(1)$. To that end, we will utilize the following result, due to Yoshihara (1976):

**Lemma A4.** *For any $j$ such that $0 \leq j < k$, let $h(x_1, x_2, \ldots, x_k)$ be a Borel function such that $\int_{\mathbf{R}^k} |h(x_1, x_2, \ldots, x_k)|^{1+\delta} dF_j^{(k)} \leq M$ for some $\delta > 0, M \geq 0$. Then $|\int_{\mathbf{R}^k} h(x_1, x_2, \ldots, x_k) dF_0^{(k)} - \int_{\mathbf{R}^k} h(x_1, x_2, \ldots, x_k) dF_j^{(k)}| \leq 4M^{1/(1+\delta)} \beta_{i_{j+1}-i_j}^{\delta/(1+\delta)}$.*

Note that Lemma A4 still holds when the quantity $\beta_{i_{j+1}-i_j}^{\delta/(1+\delta)}$ is replaced by $\beta_{i_{j+1}-i_j}^{1-\nu}$, where $0 < \nu < \frac{\delta}{1+\delta}$, and where $\delta$ is the quantity defined in (2). This is because $|\beta_n| \leq 1$ for all $n$. We now restate Lemma A4 in terms of $\nu$:

**Lemma A5.** *For any $j$ such that $0 \leq j < k$, let $h(x_1, x_2, \ldots, x_k)$ be a Borel function such that $\int_{\mathbf{R}^k} |h(x_1, x_2, \ldots, x_k)|^{1+\delta} dF_j^{(k)} \leq M$ for some $\delta > 0, M \geq 0$. Then $|\int_{\mathbf{R}^k} h(x_1, x_2, \ldots, x_k) dF_0^{(k)} - \int_{\mathbf{R}^k} h(x_1, x_2, \ldots, x_k) dF_j^{(k)}| \leq 4M^{\nu} \beta_{i_{j+1}-i_j}^{1-\nu}$, where $0 < \nu < \frac{\delta}{1+\delta}$.*

Recall that

$$R_n^* = \frac{\sqrt{3}(n-1)}{2(n+1)} \binom{n}{2}^{-1} \sum_{i<j} (\mathbb{1}_{\{X_i < X_j\}} - F(X_j) + F(X_i) - \tfrac{1}{2}). \tag{127}$$

With that in mind, define

$$h(x, y) \triangleq \mathbb{1}_{\{x<y\}} - F(y) + F(x) - \tfrac{1}{2}. \tag{128}$$

Recalling that $E(F(X_i)) = \frac{1}{2}$, note that

$$\int_{\mathbf{R}} h(x, y) dF(x) = 0, \int_{\mathbf{R}} h(x, y) dF(y) = 0. \tag{129}$$

This implies that, for all $g < i < l < m$,

$$\int_{\mathbf{R}^4} h(x_g, x_i) h(x_l, x_m) dF(x_g) dF(x_i, x_l, x_m) = 0, \tag{130}$$

and

$$\int_{\mathbf{R}^4} h(x_g, x_i) h(x_l, x_m) dF(x_g, x_i, x_l) dF(x_m) = 0. \tag{131}$$

Also observe that $|h(x, y)| < 4$. It follows that for any $j = 1, 2, 3$,

$$\int_{\mathbf{R}^4} |h(x_g, x_i) h(x_l, x_m)|^{1+\delta} dF_j^{(4)} \leq 16^{1+\delta}. \tag{132}$$

Therefore, by Lemma A5,

$$\left| \int_{\mathbf{R}^4} h(x_g, x_i) h(x_l, x_m) dF(x_g, x_i, x_l, x_m) - \int_{\mathbf{R}^4} h(x_g, x_i) h(x_l, x_m) dF_1^{(4)} \right| \leq 64 \beta_{i-g}^{1-\nu}, \tag{133}$$

and

$$\left| \int_{\mathbf{R}^4} h(x_g, x_i) h(x_l, x_m) dF(x_g, x_i, x_l, x_m) - \int_{\mathbf{R}^4} h(x_g, x_i) h(x_l, x_m) dF_3^{(4)} \right| \leq 64 \beta_{m-l}^{1-\nu}. \tag{134}$$

Using (130) and (131) respectively, the above can be simplified:

$$\left| \int_{\mathbf{R}^4} h(x_g, x_i) h(x_l, x_m) dF(x_g, x_i, x_l, x_m) \right| \leq 64 \beta_{i-g}^{1-\nu}, \tag{135}$$

and

$$\left| \int_{\mathbf{R}^4} h(x_g, x_i) h(x_l, x_m) dF(x_g, x_i, x_l, x_m) \right| \leq 64 \beta_{m-l}^{1-\nu}. \tag{136}$$

The above two inequalities can be "combined" to obtain a sharper upper bound:

$$\left| \int_{\mathbf{R}^4} h(x_g, x_i) h(x_l, x_m) dF(x_g, x_i, x_l, x_m) \right| \leq 64 \beta_{(i-g)\vee(m-l)}^{1-\nu}. \tag{137}$$

Next, for all $g < i < l$,

$$\left| \int_{\mathbf{R}^3} h(x_g, x_i) h(x_i, x_l) dF(x_g, x_i, x_l) - \int_{\mathbf{R}^3} h(x_g, x_i) h(x_i, x_l) dF_1^{(3)} \right| \leq 64 \beta_{i-g}^{1-\nu}, \tag{138}$$

and

$$\left| \int_{\mathbf{R}^3} h(x_g, x_i) h(x_i, x_l) dF(x_g, x_i, x_l) - \int_{\mathbf{R}^3} h(x_g, x_i) h(x_i, x_l) dF_2^{(3)} \right| \leq 64 \beta_{l-i}^{1-\nu}. \tag{139}$$

Using the same reasoning which resulted in equations (130) and (131), we see that

$$\int_{\mathbf{R}^3} h(x_g, x_i) h(x_i, x_l) dF(x_g) dF(x_i, x_l) = 0 \tag{140}$$

and

$$\int_{\mathbf{R}^3} h(x_g, x_i) h(x_i, x_l) dF(x_g, x_i) dF(x_l) = 0. \tag{141}$$

So we see that

$$\left| \int_{\mathbf{R}^3} h(x_g, x_i) h(x_i, x_l) dF(x_g, x_i, x_l) \right| \leq 64 \beta_{(i-g) \vee (l-i)}^{1-\nu}. \tag{142}$$

Similarly,

$$\left| \int_{\mathbf{R}^3} h(x_g, x_i) h(x_g, x_l) dF(x_g, x_i, x_l) \right| \leq 64 \beta_{l-i}^{1-\nu}, \tag{143}$$

and

$$\left| \int_{\mathbf{R}^3} h(x_g, x_i) h(x_l, x_i) dF(x_g, x_i, x_l) \right| \leq 64 \beta_{i-g}^{1-\nu} \tag{144}$$

Next, for all $g < i, l < m$, define

$$J((g, i), (l, m)) \triangleq h(X_g, X_i) h(X_l, X_m). \tag{145}$$

It follows from the definition of expectation, (137), (142), (143) and (144) that:

$$|E(J((g,i),(l,m)))| \leq 64\beta_{(i-g)\vee(m-l)}^{1-\nu}, \qquad |E(J((g,i),(g,l)))| \leq 64\beta_{l-i}^{1-\nu}$$

$$|E(J((g,i),(i,l)))| \leq 64\beta_{(i-g)\vee(l-i)}^{1-\nu}, \qquad |E(J((g,i),(l,i)))| \leq 64\beta_{i-g}^{1-\nu}$$

We define the following index sets:

$$I_1 \triangleq \{(g,i,l,m) : 1 \leq g < i \leq l < m \leq n\}, \tag{146}$$

$$I_2 \triangleq \{(g,i,l,m) : 1 \leq g < l \leq i < m \leq n\}, \tag{147}$$

$$I_3 \triangleq \{(g,i,l,m) : 1 \leq g \leq l < m \leq i \leq n\}. \tag{148}$$

Note that

$$\{I_1 \cup I_2 \cup I_3\} = \{(g,i,l,m) : 1 \leq g < i \leq n, 1 \leq l < m \leq n\}. \tag{149}$$

Next, we follow the general approach taken by Yoshihara (1976) (cf. Equations (2.15)-(2.21) therein):

$$\sum_{I_1} |E(J((g,i),(l,m)))| \leq 64n^2 \sum_{k=1}^{n} (k+1)\beta_k^{1-\nu}, \tag{150}$$

$$\sum_{I_2} |E(J((g,i),(l,m)))| \leq 64n^2 \sum_{k=1}^{n} (k+1)\beta_k^{1-\nu}, \tag{151}$$

$$\sum_{I_3} |E(J((g,i),(l,m)))| \leq 64n^2 \left(1 + 2\sum_{k=1}^{n} \beta_k^{1-\nu}\right). \tag{152}$$

Observe that

$$R^{*2} = \sum_{g<i} \sum_{l<m} J((g,i),(l,m)) = \sum_{I_1 \cup I_2 \cup I_3} J((g,i),(l,m)); \tag{153}$$

so

$$E(R_n^{*2}) = E\left[\frac{3(n-1)^2}{4(n+1)^2}\binom{n}{2}^{-2}\sum_{I_1\cup I_2\cup I_3} J((g,i),(l,m))\right] \tag{154}$$

$$\leq \frac{3(n-1)^2}{4(n+1)^2}\binom{n}{2}^{-2} E\left[\sum_{I_1\cup I_2\cup I_3} |J((g,i),(l,m))|\right] \tag{155}$$

$$\leq \frac{3(n-1)^2}{4(n+1)^2}\binom{n}{2}^{-2} \sum_{I_1\cup I_2\cup I_3} E|J((g,i),(l,m))| \tag{156}$$

Noting that

$$\sum_{I_1\cup I_2\cup I_3} E|J((g,i),(l,m))| \tag{157}$$

$$\leq \sum_{I_1} E|J((g,i),(l,m))| + \sum_{I_2} E|J((g,i),(l,m))| + \sum_{I_3} E|J((g,i),(l,m))| \tag{158}$$

$$= 128n^2\sum_{k=1}^{n}(k+1)\beta_k^{1-\nu} + 64n^2(1+2\sum_{k=1}^{n}\beta_k^{1-\nu}) \tag{159}$$

$$\leq 192n^2\left[\sum_{k=1}^{n}(k+1)\beta_k^{1-\nu} + 1 + 2\sum_{k=1}^{n}\beta_k^{1-\nu}\right] \tag{160}$$

$$= 192n^2\left[\sum_{k=1}^{n}(k+3)\beta_k^{1-\nu} + 1\right], \tag{161}$$

it follows that

$$E(R_n^{*2}) \leq \frac{3(n-1)^2}{4(n+1)^2}\binom{n}{2}^{-2} 192n^2\left[\sum_{k=1}^{n}(k+3)\beta_k^{1-\nu} + 1\right]. \tag{162}$$

Since $\beta_n = o(n^{-1-\delta})$, this implies that

$$\beta_n^{1-\nu} = \beta_n^{1/(1+\delta)} = o(n^{-1}). \tag{163}$$

Thus, $\sum_{k=1}^{+\infty} \beta_k^{1-\nu} < +\infty$, and so for all $\epsilon > 0$, there exists a $n_0$ large enough such that $\sum_{k=n_0+1}^{+\infty} \beta_k^{1-\nu} < \frac{\epsilon}{2}$. In addition,

$$0 \le \sum_{k=n_0+1}^{n} (\tfrac{k+3}{n})\beta_k^{1-\nu} \le \sum_{k=n_0+1}^{n} \beta_k^{1-\nu}. \tag{164}$$

So,

$$\frac{1}{n}\sum_{k=1}^{n}(k+3)\beta_k^{1-\nu} = \sum_{k=1}^{n}(\tfrac{k+3}{n})\beta_k^{1-\nu} \tag{165}$$

$$\le \sum_{k=1}^{n_0}(\tfrac{n_0+3}{n})\beta_k^{1-\nu} + \sum_{k=n_0+1}^{n}(\tfrac{k+3}{n})\beta_k^{1-\nu} \tag{166}$$

$$= \tfrac{n_0+3}{n}\sum_{k=1}^{n_0}\beta_k^{1-\nu} + \sum_{k=n_0+1}^{n}(\tfrac{k+3}{n})\beta_k^{1-\nu} \tag{167}$$

$$< \tfrac{n_0+3}{n}\sum_{k=1}^{n_0}\beta_k^{1-\nu} + \sum_{k=n_0+1}^{+\infty}\beta_k^{1-\nu} \tag{168}$$

$$< \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} \text{ for } n > \max\left\{\tfrac{2(n_0+3)}{\epsilon}\sum_{k=1}^{n_0}\beta_k^{1-\nu}, n_0\right\} \tag{169}$$

$$= \epsilon \tag{170}$$

Therefore,

$$E(nR_n^{*2}) \le \frac{3(n-1)^2}{4(n+1)^2}\binom{n}{2}^{-2} 192n^3\left[\sum_{k=1}^{n}(k+3)\beta_k^{1-\nu} + 1\right] \tag{171}$$

$$= O\left(\tfrac{1}{n}\sum_{k=1}^{n}(k+3)\beta_k^{1-\nu}\right) \tag{172}$$

$$= o(1) \tag{173}$$

Hence, $E(\sqrt{n}R_n^*)^2 \to 0$, and thus $\sqrt{n}R_n^* \xrightarrow{p} 0$.

□

**Theorem A3.** *Assume the random process in* (1) *satisfies the mixing condition specified in* (2)*. Then, under* $H_0$, $\sqrt{n}(\frac{\sqrt{3}}{2}U_n^* - 0) \xrightarrow{d} N(0, \sigma^2)$.

*Proof.* By Theorem 2,

$$\sqrt{n}(\tfrac{\sqrt{3}}{2}(U_n^* - T_n^*)) \xrightarrow{p} 0. \tag{174}$$

By Theorem 1,

$$\sqrt{n}(\tfrac{\sqrt{3}}{2}T_n^* - 0) \xrightarrow{d} N(0, \sigma^2) \tag{175}$$

Thus, by Slutsky's theorem,

$$\sqrt{n}(\tfrac{\sqrt{3}}{2}(U_n^* - T_n^*)) + \sqrt{n}(\tfrac{\sqrt{3}}{2}T_n^* - 0) \xrightarrow{d} N(0, \sigma^2), \tag{176}$$

or equivalently,

$$\sqrt{n}(\tfrac{\sqrt{3}}{2}U_n^*) \xrightarrow{d} N(0, \sigma^2). \tag{177}$$

□

**Proposition A8.**

$$\arg\min_{\{\mathbf{c}\in\mathscr{C}\}} \max_{\{\boldsymbol{\delta}\in\Delta_1\}} [P_{H_0}(T_n^* \geq t) + P_{H_a}(T_n^* < t)] = \arg\max_{\{\mathbf{c}\in\mathscr{C}\}} \min_{\{\boldsymbol{\delta}\in\Delta_1\}} \frac{\mathbf{c}^T\boldsymbol{\delta}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}}$$

*Proof.* Define

$$\boldsymbol{\delta}^T \triangleq (\delta_1, \ldots, \delta_n) = E_{H_a}\mathbf{X}^T, \tag{178}$$

where $\{\delta_i\}_{i=1}^n$ is a monotonically increasing sequence. Then,

$$\arg\min_{\{\mathbf{c}\in\mathscr{C}\}} \max_{\{\boldsymbol{\delta}\in\Delta_1\}} [P_{H_0}(T_n^* \geq t) + P_{H_a}(T_n^* < t)] \tag{179}$$

$$= \arg\min_{\{\mathbf{c}\in\mathscr{C}\}} \max_{\{\boldsymbol{\delta}\in\Delta_1\}} P_{H_a}\left(\frac{\sqrt{n}\,\mathbf{c}^T\mathbf{X}}{\sqrt{\mathbf{c}^T\Gamma\mathbf{c}}} < k\right) \tag{180}$$

$$= \arg\min_{\{\mathbf{c}\in\mathscr{C}\}} \max_{\{\boldsymbol{\delta}\in\Delta_1\}} P_{H_a}\left(\tfrac{\sqrt{n}\,\mathbf{c}^T(\mathbf{X}-\boldsymbol{\delta}+\boldsymbol{\delta})}{\sqrt{\mathbf{c}^T\Gamma\mathbf{c}}} < k\right) \tag{181}$$

$$= \arg\min_{\{\mathbf{c}\in\mathscr{C}\}} \max_{\{\boldsymbol{\delta}\in\Delta_1\}} P_{H_a}\left(\tfrac{\sqrt{n}\,\mathbf{c}^T(\mathbf{X}-\boldsymbol{\delta})}{\sqrt{\mathbf{c}^T\Gamma\mathbf{c}}} + \tfrac{\sqrt{n}\,\mathbf{c}^T\boldsymbol{\delta}}{\sqrt{\mathbf{c}^T\Gamma\mathbf{c}}} < k\right) \tag{182}$$

$$= \arg\min_{\{\mathbf{c}\in\mathscr{C}\}} \max_{\{\boldsymbol{\delta}\in\Delta_1\}} P_{H_a}\left(\sqrt{n}Z + \tfrac{\sqrt{n}\,\mathbf{c}^T\boldsymbol{\delta}}{\sqrt{\mathbf{c}^T\Gamma\mathbf{c}}} < k\right), \tag{183}$$

where Equation (183) follows since under $H_a$,

$$\tfrac{\sqrt{n}\,\mathbf{c}^T(\mathbf{X}-\boldsymbol{\delta})}{\sqrt{\mathbf{c}^T\Gamma\mathbf{c}}} \xrightarrow{d} N(0,1). \tag{184}$$

Note that:

$$\max_{\{\boldsymbol{\delta}\in\Delta_1\}} P_{H_a}\left(\sqrt{n}Z + \tfrac{\sqrt{n}\,\mathbf{c}^T\boldsymbol{\delta}}{\sqrt{\mathbf{c}^T\Gamma\mathbf{c}}} < k\right) \Longrightarrow \min_{\{\boldsymbol{\delta}\in\Delta_1\}} \left(\tfrac{\sqrt{n}\,\mathbf{c}^T\boldsymbol{\delta}}{\sqrt{\mathbf{c}^T\Gamma\mathbf{c}}}\right), \tag{185}$$

and

$$\min_{\{\mathbf{c}\in\mathscr{C}\}} P_{H_a}\left(\sqrt{n}Z + \tfrac{\sqrt{n}\,\mathbf{c}^T\boldsymbol{\delta}}{\sqrt{\mathbf{c}^T\Gamma\mathbf{c}}} < k\right) \Longrightarrow \max_{\{\mathbf{c}\in\mathscr{C}\}} \left(\tfrac{\sqrt{n}\,\mathbf{c}^T\boldsymbol{\delta}}{\sqrt{\mathbf{c}^T\Gamma\mathbf{c}}}\right). \tag{186}$$

So the optimality criterion that the minimax rule $\mathbf{c}$ must satisfy can be expressed as follows:

$$\arg\max_{\{\mathbf{c}\in\mathscr{C}\}} \min_{\{\boldsymbol{\delta}\in\Delta_1\}} \left(\tfrac{\mathbf{c}^T\boldsymbol{\delta}}{\sqrt{\mathbf{c}^T\Gamma\mathbf{c}}}\right). \tag{187}$$

□

The following calculations are done in preparation for the upcoming proof: Standardize each $\boldsymbol{\delta}^{(k)}$ such that:

$$\tilde{\delta}_j^{(k)} \triangleq \frac{\delta_j^{(k)} - \bar{\delta}^{(k)}}{\sigma^{-1}\|\boldsymbol{\delta}^{(k)} - \bar{\boldsymbol{\delta}}^{(k)}\|}, \text{ where } \bar{\delta}^{(k)} \triangleq n^{-1}\sum_{i=1}^{n} \delta_i^{(k)} \tag{188}$$

It follows that

$$(\tilde{\boldsymbol{\delta}}^{(k)})^T \mathbb{1}_n = 0, \text{ and } \|\tilde{\boldsymbol{\delta}}^{(k)}\| = \sigma \tag{189}$$

Now,

$$\boldsymbol{\delta}^{(k)} - \bar{\boldsymbol{\delta}}^{(k)} = (\underbrace{\tfrac{-(n-k)}{n}, \ldots,}_{k} \underbrace{\tfrac{k}{n},}_{\text{(k+1)th element}} \underbrace{\tfrac{k}{n}, \ldots}_{n-k-1})^T, \tag{190}$$

and,

$$\|\boldsymbol{\delta}^{(k)} - \bar{\boldsymbol{\delta}}^{(k)}\| = \left(\tfrac{k(n-k)}{n}\right)^{1/2} \tag{191}$$

so,

$$\tilde{\boldsymbol{\delta}}^{(k)} \triangleq \frac{\boldsymbol{\delta}^{(k)} - \bar{\boldsymbol{\delta}}^{(k)}}{\sigma^{-1}\|\boldsymbol{\delta}^{(k)} - \bar{\boldsymbol{\delta}}^{(k)}\|} \tag{192}$$

$$= \sigma(\underbrace{-\{\tfrac{(n-k)}{kn}\}^{1/2}, -\{\tfrac{(n-k)}{kn}\}^{1/2}, \ldots,}_{k} \underbrace{\{\tfrac{k}{(n-k)n}\}^{1/2},}_{\text{(k+1)th element}} \underbrace{\{\tfrac{k}{(n-k)n}\}^{1/2}, \ldots}_{n-k-1}) \tag{193}$$

Define

$$\acute{\boldsymbol{\delta}}^{(k)} \triangleq \sigma^{-1}\tilde{\boldsymbol{\delta}}^{(k)} \text{ for all } k = 1, 2, \ldots, n-1, \tag{194}$$

Then

$$\frac{\mathbf{c}^T \tilde{\boldsymbol{\delta}}^{(k)}}{\sigma(\mathbf{c}^T\mathbf{c})^{1/2}} = \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(k)}}{(\mathbf{c}^T\mathbf{c})^{1/2}}, \quad \|\acute{\boldsymbol{\delta}}^{(k)}\| = 1 \text{ for all } k = 1, 2, \ldots, n-1 \tag{195}$$

**Proposition A9.** $c_j = \zeta\left\{ \left(\frac{(j-1)(n-(j-1))}{n}\right)^{1/2} - \left(\frac{j(n-j)}{n}\right)^{1/2} \right\}$, *for* $j = 1, 2, \ldots, n$.

*Proof.* We use (strong) induction.

Base case:

$$\tfrac{1}{(n-1)}\left[\sum_{i=2}^{n} c_i\right] - \zeta(\tfrac{n}{n-1})^{1/2} = c_1 \tag{196}$$

$$\Longrightarrow \frac{-1}{(n-1)}c_1 - \zeta(\tfrac{n}{n-1})^{1/2} = c_1 \text{ (since } \sum_{i=1}^{n} c_i = 0) \tag{197}$$

$$\Longrightarrow -\zeta(\tfrac{n}{n-1})^{1/2} = (\tfrac{n}{n-1})c_1 \tag{198}$$

$$\Longrightarrow c_1 = -\zeta(\tfrac{n-1}{n})^{1/2} \tag{199}$$

□

Now assume that $c_k = \zeta\left\{ \left(\frac{(k-1)(n-(k-1))}{n}\right)^{1/2} - \left(\frac{k(n-k)}{n}\right)^{1/2} \right\}$ holds, for a fixed $k < n$, and for all $j \leq k$.

Then,

$$\frac{k+1}{(n-(k+1))} \sum_{i=k+2}^{n} c_i - \zeta\left(\frac{(k+1)n}{n-(k+1)}\right)^{1/2} = \left(\sum_{i=1}^{k+1} c_i\right) \tag{200}$$

$$\Longrightarrow -\frac{k+1}{(n-(k+1))}\left(\sum_{i=1}^{k+1} c_i\right) - \zeta\left(\frac{(k+1)n}{n-(k+1)}\right)^{1/2} = \left(\sum_{i=1}^{k+1} c_i\right) \tag{201}$$

$$\Longrightarrow -\zeta\left(\frac{(k+1)n}{n-(k+1)}\right)^{1/2} = \frac{n}{(n-(k+1))}\left(\sum_{i=1}^{k+1} c_i\right) \tag{202}$$

$$\Longrightarrow -\zeta\left(\frac{(k+1)(n-(k+1))}{n}\right)^{1/2} = \left(\sum_{i=1}^{k+1} c_i\right) \tag{203}$$

$$\Longrightarrow -\zeta\left(\frac{(k+1)(n-(k+1))}{n}\right)^{1/2} - \left(\sum_{i=1}^{k} c_i\right) = c_{k+1} \tag{204}$$

$$\Longrightarrow -\zeta\left(\frac{(k+1)(n-(k+1))}{n}\right)^{1/2} = \left(\sum_{i=1}^{k+1} c_i\right) \tag{205}$$

$$\Longrightarrow -\zeta\left(\frac{(k+1)(n-(k+1))}{n}\right)^{1/2} - \sum_{i=1}^{k} \zeta\left[\left(\frac{(i-1)(n-(i-1))}{n}\right)^{1/2} - \left(\frac{i(n-i)}{n}\right)^{1/2}\right] = c_{k+1} \tag{206}$$

Equation (206) follows by the inductive step.

Observing that we have a telescoping sum,

$$-\zeta\left(\tfrac{(k+1)(n-(k+1))}{n}\right)^{1/2}+\zeta\left(\tfrac{k(n-k)}{n}\right)^{1/2}=c_{k+1} \tag{207}$$

Now, for $k=n$, it follows that

$$c_n=-\sum_{i=1}^{n-1}c_i \tag{208}$$

$$=-\zeta\sum_{i=1}^{n-1}\left[\left(\tfrac{(i-1)(n-(i-1))}{n}\right)^{1/2}-\left(\tfrac{i(n-i)}{n}\right)^{1/2}\right] \tag{209}$$

$$=\zeta(\tfrac{n-1}{n})^{1/2} \tag{210}$$

□

**Lemma A6.** *Every convex cone is equal to its convex conical hull.*

**Lemma A7.** *Every element of a convex conical hull* $\boldsymbol{\Upsilon}$ *can be expressed as a conical combination of* $\dim(\mathrm{span}(\boldsymbol{\Upsilon}))$ *linearly independent elements of* $\boldsymbol{\Upsilon}$.

**Proposition A10.** *For* $\mathbf{c}\in\mathscr{C}$, $\Gamma$ *symmetric positive definite, if* $\Gamma\mathbf{c}_0\in\boldsymbol{\Delta}$, *and there exist* $j,l\in A$ *such that* $\min_{k\in A}\frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(k)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}}=\frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}}<\frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}}$, *then* $\frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}}<\frac{\mathbf{c}_0^T\acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T\Gamma\mathbf{c}_0)^{1/2}}$ .

*Proof.* Suppose for such a $\mathbf{c}\in\mathscr{C}$, there exists $j\in A$ such that

$$\frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}}\geq\frac{\mathbf{c}_0^T\acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T\Gamma\mathbf{c}_0)^{1/2}}. \tag{211}$$

By construction,

$$\frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \equiv \zeta_2 > 0 \text{ for all } i \in A, \text{ where } \zeta_2 \text{ is a constant;} \tag{212}$$

this implies that

$$\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} > \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} \geq \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \text{ for all } i \in A \tag{213}$$

Since $\Gamma \mathbf{c}_0 \in \boldsymbol{\Delta}$, it follows that

$$\Gamma \mathbf{c}_0 = \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}^{(i)}, w_i \geq 0 \text{ for all } i \in A. \tag{214}$$

So,

$$\begin{aligned}
\frac{\mathbf{c}_0^T \Gamma \mathbf{c}_0}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} &= \frac{\mathbf{c}_0^T \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} && (215)\\
&< \frac{\mathbf{c}^T \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} && (216)\\
&= \frac{\mathbf{c}^T \Gamma \mathbf{c}_0}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}}, && (217)
\end{aligned}$$

or equivalently,

$$(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2} (\mathbf{c}^T \Gamma \mathbf{c})^{1/2} < (\mathbf{c}^T \Gamma \mathbf{c}_0), \tag{218}$$

which can also be expressed as

$$\|\mathbf{c}\|_\Gamma \|\mathbf{c}_0\|_\Gamma < \langle \mathbf{c}, \mathbf{c}_0 \rangle \tag{219}$$

However, by the Schwarz inequality,

$$\langle \mathbf{c}, \mathbf{c}_0 \rangle \leq \|\mathbf{c}\|_\Gamma \|\mathbf{c}_0\|_\Gamma \tag{220}$$

A contradiction is reached. Therefore,

$$\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} < \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \tag{221}$$

and hence such a $\mathbf{c}$ cannot be minimax. □

**Proposition A11.** *For* $\mathbf{c} \in \mathscr{C}$, $\Gamma$ *symmetric positive definite, if* $\Gamma \mathbf{c}_0 \in \boldsymbol{\Delta}$, *and* $\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} = \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}}$ *for all* $i, l \in A$, *then* $\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} = \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}}$ *for all* $i, l \in A$.

*Proof.* Suppose for such a $\mathbf{c} \in \mathscr{C}$,

$$\text{there exists } j \in A \text{ such that } \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} > \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}}. \tag{222}$$

By construction,

$$\frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \equiv \zeta_2 > 0 \text{ for all } i \in A, \tag{223}$$

which implies that

$$\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} > \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \text{ for all } i, l \in A. \tag{224}$$

Since $\Gamma \mathbf{c}_0 \in \boldsymbol{\Delta}$, it follows that

$$\Gamma \tilde{\mathbf{c}} = \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}^{(i)}, w_i \geq 0 \text{ for all } i \in A. \tag{225}$$

So,

$$\frac{\mathbf{c}_0^T \Gamma \mathbf{c}_0}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} = \frac{\mathbf{c}_0^T \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \tag{226}$$

$$< \frac{\mathbf{c}^T \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} \tag{227}$$

$$= \frac{\mathbf{c}^T \Gamma \mathbf{c}_0}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} \tag{228}$$

or equivalently,

$$(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2} (\mathbf{c}^T \Gamma \mathbf{c})^{1/2} < (\mathbf{c}^T \Gamma \mathbf{c}_0), \tag{229}$$

which can also be expressed as

$$\|\mathbf{c}\|_\Gamma \|\mathbf{c}_0\|_\Gamma < \langle \mathbf{c}, \mathbf{c}_0 \rangle \tag{230}$$

However, by the Schwarz inequality,

$$\langle \mathbf{c}, \mathbf{c}_0 \rangle \leq \|\mathbf{c}\|_\Gamma \|\mathbf{c}_0\|_\Gamma \tag{231}$$

A contradiction is reached. Therefore,

$$\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} \leq \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \text{ for all } j \in A. \tag{232}$$

In the same manner, by instead initially assuming there exists a $j \in A$ such that $\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} < \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}}$, it can also be shown that $\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} \geq \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}}$ for all $j \in A$. Hence, we conclude that

$$\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} = \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \text{ for all } i, l \in A. \tag{233}$$

In other words, $\mathbf{c}$ is minimax if and only if $\mathbf{c}_0$ is minimax. □

Define

$$\mathscr{A} \triangleq \left\{ \mathbf{c} \in \mathrm{R}^{n \times 1} : \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} = \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} \text{ for all } i, l \in A \right\}. \tag{234}$$

**Theorem A4.** *Let the random process be defined as in Equation* (1)*. Then* $\mathbf{c}_0$ *is minimax. Moreover, membership in* $\mathscr{A}$ *is a necessary and sufficient condition for* $\mathbf{c} \in \mathscr{C}$ *to be minimax.*

*Proof.* Observe that

$$\mathscr{A}^c = \left\{ \mathbf{c} \in \mathrm{R}^{n \times 1} : \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} < \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} \text{ for some } i, l \in A \right\}. \tag{235}$$

Proposition A10 implies that any $\mathbf{c} \in (\mathscr{C} \cap \mathscr{A}^c)$ cannot be minimax as it is beaten by $\mathbf{c}_0$. Proposition A11 demonstrates that any $\mathbf{c} \in (\mathscr{C} \cap \mathscr{A})$ matches, but does not beat $\mathbf{c}_0$. But since $\mathbf{c} \in (\mathscr{C} \cap \mathscr{A}) \cup (\mathscr{C} \cap \mathscr{A}^c) = \mathscr{C}$, the claim immediately follows; in other words,

$$\mathbf{c}_0 = \arg\max_{\{\mathbf{c} \in \mathscr{C}\}} \min_{\{k \in A\}} \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(k)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}}. \tag{236}$$

□

**Corollary A2.** $\mathbf{c}_0$ *is the unique minimax, up to a positive scaling.*

*Proof.* If $\hat{\mathbf{c}}$ is minimax, then the following must hold, by Theorem A4:

$$\frac{\hat{\mathbf{c}}^T \acute{\boldsymbol{\delta}}^{(k)}}{(\hat{\mathbf{c}}^T \Gamma \hat{\mathbf{c}})^{1/2}} = \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(k)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \equiv \zeta_2 > 0 \text{ for all } k \in A. \tag{237}$$

Re-expressing the above equality as

$$\frac{\hat{\mathbf{c}}^T \acute{\boldsymbol{\delta}}^{(k)}}{\|\hat{\mathbf{c}}\|_\Gamma} = \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(k)}}{\|\mathbf{c}_0\|_\Gamma} \text{ for all } k \in A, \tag{238}$$

after algebraic manipulation, we see that

$$(\|\mathbf{c}_0\|_\Gamma \hat{\mathbf{c}}^T - \|\hat{\mathbf{c}}\|_\Gamma \mathbf{c}_0^T) \acute{\boldsymbol{\delta}}^{(k)} = 0. \tag{239}$$

There are three cases to consider:

Clearly, $\acute{\boldsymbol{\delta}}^{(k)}$ is not the $(n-1)$-vector of 0's.

Next, if $\|\mathbf{c}_0\|_\Gamma \hat{\mathbf{c}}^T - \|\hat{\mathbf{c}}\|_\Gamma \mathbf{c}_0^T$ equals the $(n-1)$-vector of 0's, then it follows that

$$\frac{\hat{\mathbf{c}}^T}{\|\hat{\mathbf{c}}\|_\Gamma} = \frac{\mathbf{c}_0^T}{\|\mathbf{c}_0\|_\Gamma}. \tag{240}$$

This shows that $\hat{\mathbf{c}} \in \text{span}\{\mathbf{c}_0\}$.

The last case assumes that neither $\|\mathbf{c}_0\|_\Gamma \hat{\mathbf{c}}^T - \|\hat{\mathbf{c}}\|_\Gamma \mathbf{c}_0^T$ nor $\acute{\boldsymbol{\delta}}^{(k)}$ equal the $(n-1)$-vector of 0's. Then

$$(\|\mathbf{c}_0\|_\Gamma \hat{\mathbf{c}}^T - \|\hat{\mathbf{c}}\|_\Gamma \mathbf{c}_0^T) \acute{\boldsymbol{\delta}}^{(k)} = 0. \tag{241}$$

But this implies that

$$(\|\mathbf{c}_0\|_\Gamma \hat{\mathbf{c}}^T - \|\hat{\mathbf{c}}\|_\Gamma \mathbf{c}_0^T) \in N\left\{\acute{\boldsymbol{\delta}}^{(1)}\ \acute{\boldsymbol{\delta}}^{(2)} \cdots\ \acute{\boldsymbol{\delta}}^{(n-1)}\right\} \tag{242}$$

and thus

$$\mathbf{c}_0 \in N\left\{\hat{\boldsymbol{\delta}}^{(1)}\ \hat{\boldsymbol{\delta}}^{(2)} \cdots \hat{\boldsymbol{\delta}}^{(n-1)}\right\}, \tag{243}$$

which cannot be true as

$$\mathbf{c}_0 \in C\left\{\hat{\boldsymbol{\delta}}^{(1)}\ \hat{\boldsymbol{\delta}}^{(2)} \cdots \hat{\boldsymbol{\delta}}^{(n-1)}\right\},\ \text{unless}\ \hat{\mathbf{c}} \in \operatorname{span}\{\mathbf{c}_0\}. \tag{244}$$

Finally, note that $\hat{\mathbf{c}} = -\mathbf{c}_0$ cannot be minimax; if it were, then

$$\frac{\hat{\mathbf{c}}^T \hat{\boldsymbol{\delta}}^{(k)}}{(\hat{\mathbf{c}}^T \Gamma \hat{\mathbf{c}})^{1/2}} = \frac{-\mathbf{c}_0^T \hat{\boldsymbol{\delta}}^{(k)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} = \zeta_2 > 0 \text{ for all } k \in A. \tag{245}$$

But by construction it follows that $\dfrac{-\mathbf{c}_0^T \hat{\boldsymbol{\delta}}^{(k)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} = -\zeta_2 < 0$. Thus, only positive scalings of $\mathbf{c}_0$ are admitted. □

From Lemma A7, it follows that

$$\Gamma \mathbf{c}_0 = \sum_{i=1}^{n-1} w_i \Gamma \hat{\boldsymbol{\delta}}^{(i)} \tag{246}$$

Define $\Gamma\boldsymbol{\Delta} \triangleq \{\Gamma\mathbf{c} : \mathbf{c} \in \boldsymbol{\Delta}\}$ Using a result given in Rockafellar (1970, p. 19), it follows that $\Gamma\boldsymbol{\Delta}$ is also a convex cone, and thus is a convex conical hull. We state this as a lemma:

**Lemma A8.** $\Gamma\boldsymbol{\Delta}$ *is a convex conical hull.*

Now define $Y \triangleq \left\{\hat{\boldsymbol{\delta}}^{(1)}\ \hat{\boldsymbol{\delta}}^{(2)}\ \cdots\ \hat{\boldsymbol{\delta}}^{(n-1)}\right\} \in \mathbf{R}^{n\times(n-1)}$
Consider the mapping $T(\Gamma) : \mathbf{R}^{n\times n} \to \mathbf{R}^{n\times(n-1)} \triangleq \Gamma Y$, whose domain is the set of all correlation matrices $\Gamma \in \mathbf{R}^{n\times n}$.
Recall that $\Gamma$ is the correlation matrix of the random process. It is positive-

definite and thus of full rank. We now demonstrate that $\Gamma\boldsymbol{\Delta}$ is full rank (Dattoro 2013, p. 37).

**Lemma A9.** *The columns of* $\Gamma Y$ *are linearly independent.*

*Proof.* $\left\{\Gamma\acute{\boldsymbol{\delta}}^{(i)}, i \in A\right\}$ are linearly independent by definition if

$$\sum_{i=1}^{n-1} a_i \Gamma \acute{\boldsymbol{\delta}}^{(i)} = \mathbf{0} \Longrightarrow a_i \equiv 0 \text{ for all } i \in A. \tag{247}$$

Since $\Gamma$ is invertible, it has a left inverse, so it can be seen that

$$\sum_{i=1}^{n-1} a_i \Gamma \acute{\boldsymbol{\delta}}^{(i)} = \Gamma \left( \sum_{i=1}^{n-1} a_i \acute{\boldsymbol{\delta}}^{(i)} \right) = \mathbf{0} \Longleftrightarrow \sum_{i=1}^{n-1} a_i \acute{\boldsymbol{\delta}}^{(i)} = \mathbf{0}. \tag{248}$$

But, by the linear independence of $\left\{\acute{\boldsymbol{\delta}}^{(1)}, \acute{\boldsymbol{\delta}}^{(2)}, \ldots, \acute{\boldsymbol{\delta}}^{(n-1)}\right\}$, this can only be achieved when $a_i \equiv 0$ for all $i \in A$, as claimed.

□

As a side note, observe that by Lemma A9, $\text{rank}(\Gamma Y) = n - 1$.

**Lemma A10.** *Any* $\mathbf{c} \in \Gamma\boldsymbol{\Delta}$ *can be expressed as a conical combination of the elements of* $\left\{\Gamma\acute{\boldsymbol{\delta}}^{(i)}, i \in A\right\}$

*Proof.* This immediately follows by Lemmas A8, A9 and A7 respectively. □

**Definition A1.** *The generators for a closed convex cone* $C$ *are any collection of directions whose convex conical hull constructs* $C$*.*

**Lemma A11.** $\left\{\Gamma\acute{\boldsymbol{\delta}}^{(i)}, i \in A\right\}$ *comprise the generators for* $\Gamma\boldsymbol{\Delta}$*. Moreover, this is a minimal set of generators for* $\Gamma\boldsymbol{\Delta}$*.*

*Proof.* The first claim immediately follows by Lemmas A8 and A10 respectively. Next, suppose that $\left\{\Gamma\acute{\boldsymbol{\delta}}^{(i)}, i \in A\right\}$ is not a minimal set of generators; so there exists a $i \in A$, say $k$, such that $\Gamma\acute{\boldsymbol{\delta}}^{(k)}$ is superfluous. But this is not true as $\Gamma\acute{\boldsymbol{\delta}}^{(k)}$ is independent of $\left\{\Gamma\acute{\boldsymbol{\delta}}^{(i)}, i \in A\backslash\{k\}\right\}$ by Lemma A9, so it cannot be expressed as a conical combination of the aforementioned collection of generators. But this is a contradiction, as clearly $\Gamma\acute{\boldsymbol{\delta}}^{(k)} \in \Gamma\boldsymbol{\Delta}$, by definition of $\Gamma\boldsymbol{\Delta}$. □

**Definition A2.** *An extreme direction for a pointed closed convex cone $C$ is any nonnegative scaling of a vector that cannot be expressed as a conic combination of any other vectors in $C$.*

**Definition A3.** *A set of directions is said to be conically independent if no direction from the set can be expressed as a conic combination of the remaining directions.*

Clearly, linear independence is a stronger property.

The following fact is given in Dattoro (2013, p. 145):

**Lemma A12.** *When a set of conically independent directions from a pointed closed convex cone $C$ is comprised of generators, then all of these directions must be extreme directions of $C$.*

**Lemma A13.** $\left\{\Gamma\acute{\boldsymbol{\delta}}^{(i)}, i \in A\right\}$ *are the extreme directions of* $\Gamma\boldsymbol{\Delta}$*.*

*Proof.* This follows from Lemmas A9, A11 and A12 respectively. □

Define

$$\acute{\boldsymbol{\delta}}_{\Gamma}^{(i)} \triangleq \Gamma\acute{\boldsymbol{\delta}}^{(i)} \text{ for all } i \in A \tag{249}$$

Since

$$\mathbf{c}_0 = \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}^{(i)}, w_i \geq 0 \text{ for all } i \in A, \tag{250}$$

it follows that

$$\Gamma \mathbf{c}_0 = \sum_{i=1}^{n-1} w_i \Gamma \acute{\boldsymbol{\delta}}^{(i)} = \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}_{\Gamma}^{(i)} \tag{251}$$

**Proposition A12.** *For* $\mathbf{c} \in \mathscr{C}$, $\Gamma$ *symmetric positive definite,* $\Gamma \mathbf{c}_0 \in \Gamma \boldsymbol{\Delta}$, *and* $\mathbf{c}_0^T \boldsymbol{\delta}_{\Gamma}^{(i)} = \mathbf{c}_0^T \boldsymbol{\delta}_{\Gamma}^{(l)}$ *for all* $i, l \in A$, *if there exist* $j, l \in A$ *such that* $\min_{k \in A} \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(k)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} = \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} < \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}}$, *then* $\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} < \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}}$ .

*Proof.* Suppose for such a $\mathbf{c} \in \mathscr{C}$, there exists $j \in A$ such that

$$\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} \geq \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}}. \tag{252}$$

By construction,

$$\frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \equiv \zeta_2 > 0 \text{ for all } i \in A, \text{ where } \zeta_2 \text{ is a constant;} \tag{253}$$

this implies that

$$\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} > \frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} \geq \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \text{ for all } i \in A \tag{254}$$

Since $\Gamma \mathbf{c}_0 \in \Gamma \boldsymbol{\Delta}$, it follows that

$$\Gamma \mathbf{c}_0 = \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}_{\Gamma}^{(i)}, w_i \geq 0 \text{ for all } i \in A. \tag{255}$$

So,

$$\frac{\mathbf{c}_0^T \Gamma \mathbf{c}_0}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} = \frac{\mathbf{c}_0^T \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}_\Gamma^{(i)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \tag{256}$$

$$< \frac{\mathbf{c}^T \sum_{i=1}^{n-1} w_i \acute{\boldsymbol{\delta}}_\Gamma^{(i)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} \tag{257}$$

$$= \frac{\mathbf{c}^T \Gamma \mathbf{c}_0}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}}, \tag{258}$$

or equivalently,

$$(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2} (\mathbf{c}^T \Gamma \mathbf{c})^{1/2} < (\mathbf{c}^T \Gamma \mathbf{c}_0), \tag{259}$$

which can also be expressed as

$$\|\mathbf{c}\|_\Gamma \|\mathbf{c}_0\|_\Gamma < \langle \mathbf{c}, \mathbf{c}_0 \rangle \tag{260}$$

However, by the Schwarz inequality,

$$\langle \mathbf{c}, \mathbf{c}_0 \rangle \leq \|\mathbf{c}\|_\Gamma \|\mathbf{c}_0\|_\Gamma \tag{261}$$

A contradiction is reached. Therefore,

$$\frac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} < \frac{\mathbf{c}_0^T \acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \tag{262}$$

and hence such a $\mathbf{c}$ cannot be minimax. □

**Proposition A13.** *For* $\mathbf{c} \in \mathscr{C}$, $\Gamma$ *symmetric positive definite, and* $\mathbf{c}_0^T \boldsymbol{\delta}_\Gamma^{(i)} = \mathbf{c}_0^T \boldsymbol{\delta}_\Gamma^{(l)}$ *for all* $i, l \in A$, $\Gamma \mathbf{c}_0 \in \Gamma \boldsymbol{\Delta}$, *if* $\dfrac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} = \dfrac{\mathbf{c}^T \acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}}$ *for all* $i, l \in A$, *then*

$$\frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}} = \frac{\mathbf{c}_0^T\acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}_0^T\Gamma\mathbf{c}_0)^{1/2}} \textit{ for all } i, l \in A.$$

*Proof.* Suppose for such a $\mathbf{c} \in \mathscr{C}$,

$$\text{there exists } j \in A \text{ such that } \frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}} > \frac{\mathbf{c}_0^T\acute{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T\Gamma\mathbf{c}_0)^{1/2}}. \tag{263}$$

By construction,

$$\frac{\mathbf{c}_0^T\acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}_0^T\Gamma\mathbf{c}_0)^{1/2}} \equiv \zeta_2 > 0 \text{ for all } i \in A, \tag{264}$$

which implies that

$$\frac{\mathbf{c}^T\acute{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}} > \frac{\mathbf{c}_0^T\acute{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}_0^T\Gamma\mathbf{c}_0)^{1/2}} \text{ for all } i, l \in A. \tag{265}$$

Since $\Gamma\mathbf{c}_0 \in \Gamma\boldsymbol{\Delta}$, it follows that

$$\Gamma\mathbf{c}_0 = \sum_{i=1}^{n-1} w_i\acute{\boldsymbol{\delta}}_\Gamma^{(i)}, w_i \geq 0 \text{ for all } i \in A. \tag{266}$$

So,

$$\begin{aligned}
\frac{\mathbf{c}_0^T\Gamma\mathbf{c}_0}{(\mathbf{c}_0^T\Gamma\mathbf{c}_0)^{1/2}} &= \frac{\mathbf{c}_0^T\sum_{i=1}^{n-1} w_i\acute{\boldsymbol{\delta}}_\Gamma^{(i)}}{(\mathbf{c}_0^T\Gamma\mathbf{c}_0)^{1/2}} &(267)\\
&< \frac{\mathbf{c}^T\sum_{i=1}^{n-1} w_i\acute{\boldsymbol{\delta}}_\Gamma^{(i)}}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}} &(268)\\
&= \frac{\mathbf{c}^T\Gamma\mathbf{c}_0}{(\mathbf{c}^T\Gamma\mathbf{c})^{1/2}} &(269)
\end{aligned}$$

or equivalently,

$$(\mathbf{c}_0^T\Gamma\mathbf{c}_0)^{1/2}(\mathbf{c}^T\Gamma\mathbf{c})^{1/2} < (\mathbf{c}^T\Gamma\mathbf{c}_0), \tag{270}$$

which can also be expressed as

$$\|\mathbf{c}\|_\Gamma \|\mathbf{c}_0\|_\Gamma < \langle \mathbf{c}, \mathbf{c}_0 \rangle \tag{271}$$

However, by the Schwarz inequality,

$$\langle \mathbf{c}, \mathbf{c}_0 \rangle \leq \|\mathbf{c}\|_\Gamma \|\mathbf{c}_0\|_\Gamma \tag{272}$$

A contradiction is reached. Therefore,

$$\frac{\mathbf{c}^T \hat{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} \leq \frac{\mathbf{c}_0^T \hat{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \text{ for all } j \in A. \tag{273}$$

In the same manner, by instead initially assuming there exists a $j \in A$ such that $\frac{\mathbf{c}^T \hat{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} < \frac{\mathbf{c}_0^T \hat{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}}$, it can also be shown that $\frac{\mathbf{c}^T \hat{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} \geq \frac{\mathbf{c}_0^T \hat{\boldsymbol{\delta}}^{(j)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}}$ for all $j \in A$. Hence, we conclude that

$$\frac{\mathbf{c}^T \hat{\boldsymbol{\delta}}^{(i)}}{(\mathbf{c}^T \Gamma \mathbf{c})^{1/2}} = \frac{\mathbf{c}_0^T \hat{\boldsymbol{\delta}}^{(l)}}{(\mathbf{c}_0^T \Gamma \mathbf{c}_0)^{1/2}} \text{ for all } i, l \in A. \tag{274}$$

In other words, $\mathbf{c}$ is minimax if and only if $\mathbf{c}_0$ is minimax. □

**Theorem A5.** *Let the random process be defined as in Equation* (1). *If* $\mathbf{c}_0^T \boldsymbol{\delta}_\Gamma^{(i)} = \mathbf{c}_0^T \boldsymbol{\delta}_\Gamma^{(l)}$ *for all* $i, l \in A$, *then* $\mathbf{c}_0$ *is minimax. Moreover, membership in* $\mathscr{A}$ *is a necessary and sufficient condition for* $\mathbf{c} \in \mathscr{C}$ *to be minimax.*

*Proof.* The proof is identical to the proof of Theorem A4. □

**Corollary A3.** $\mathbf{c}_0$ *is the unique minimax, up to a positive scaling.*

*Proof.* The proof is identical to the proof of Corollary A2. □