

ORIGINAL ARTICLE

# PERMUTATION TESTING FOR DEPENDENCE IN TIME SERIES

JOSEPH P. ROMANO<sup>a,b\*</sup>  AND MARIUS A. TIRLEA<sup>a</sup>

<sup>a</sup>*Department of Statistics, Stanford University, Stanford, CA, USA*

<sup>b</sup>*Department of Economics, Stanford University, Stanford, CA, USA*

Given observations from a stationary time series, permutation tests allow one to construct exactly level  $\alpha$  tests under the null hypothesis of an i.i.d. (or, more generally, exchangeable) distribution. On the other hand, when the null hypothesis of interest is that the underlying process is an uncorrelated sequence, permutation tests are not necessarily level  $\alpha$ , nor are they approximately level  $\alpha$  in large samples. In addition, permutation tests may have large Type 3, or directional, errors, in which a two-sided test rejects the null hypothesis and concludes that the first-order autocorrelation is larger than 0, when in fact it is less than 0. In this article, under weak assumptions on the mixing coefficients and moments of the sequence, we provide a test procedure for which the asymptotic validity of the permutation test holds, while retaining the exact rejection probability  $\alpha$  in finite samples when the observations are independent and identically distributed. A Monte Carlo simulation study, comparing the permutation test to other tests of autocorrelation, is also performed, along with an empirical example of application to financial data.

*Received 3 March 2021; Revised 25 September 2021; Accepted 7 December 2021*

**Keywords:** Autocorrelation; hypothesis testing; stationary process.

**MOS subject classification:** 62G10, 62M10.

## 1. INTRODUCTION

In this article, we investigate the use of permutation tests for detecting dependence in a time series. When testing the null hypothesis that the underlying time series consists of i.i.d. observations, permutation tests can be constructed that control the probability of a Type 1 error exactly, for any choice of test statistic. Typically, the choice of test statistic is the first-order sample autocorrelation, or some function of many of the sample autocorrelations. However, significant problems of error control arise, stemming from the fact that zero autocorrelation and independence are actually quite different. It is crucial to carefully specify the null hypothesis of interest, whether it is the case that the observations are i.i.d. or that the observations have zero autocorrelation. For example, if the null hypothesis specifies that the autocorrelation is zero and one uses the sample first-order autocorrelation as a test statistic when applying a permutation test, then the Type 1 error can be shockingly different from the nominal level, even asymptotically. Nevertheless, one might think it reasonable to reject based on such a permutation test, since the test statistic appears ‘large,’ relative to the null reference permutation distribution. However, even if one views the null hypothesis as specifying that the time series is i.i.d., a rejection of the null hypothesis based on the sample autocorrelation is inevitably accompanied by the interpretation then that the true underlying autocorrelation is non-zero. Indeed, one typically makes the further claim that this correlation is positive (negative), when the sample autocorrelation can be large and positive (large and negative). When the true autocorrelation is zero and there is a large probability of a Type 1 error, then lack of Type 1 error control also implies lack of Type 3, or directional, error control. That is, there can be a large probability that one declares the underlying correlation to be positive when it is in fact negative.

\* Correspondence to: Joseph Romano, Departments of Statistics and Economics, Stanford University, Stanford, CA, USA.  
 Email: romano@stanford.edu

Assume  $X_1, \dots, X_n$  are jointly distributed according to some strictly stationary, infinite dimensional distribution  $P$ , where the distribution  $P$  belongs to some family  $\Omega$ . Consider the problem of testing  $H : P \in \Omega_0$ , where  $\Omega_0$  is some subset of stationary processes. For example, we might be interested in testing that the underlying  $P$  is a product of its marginals, that is, the underlying process is i.i.d.

The problem of testing independence in time series and time series residuals is fundamental to understanding the stochastic process under study. A frequently used analog for testing independence is that of testing the hypothesis

$$H_r : \rho(1) = \dots = \rho(r) = 0, \quad (1.1)$$

for some fixed  $r$ , where  $\rho(k)$  is the  $k$ th-order autocorrelation. Examples of such tests include those proposed by Box and Pierce (1970) and Ljung and Box (1978). However, such tests assume that the data-generating model is parametric or semi-parametric and, in particular, follows an ARIMA model. This assumption is, in general, violated for arbitrary  $P$ , and so these tests will not be exact for finite samples or asymptotically valid, as will be shown later.

We propose a non-parametric testing procedure for the hypothesis

$$H^{(k)} : \rho(k) = 0, \quad (1.2)$$

based on permutation testing, whence we may construct a testing procedure for the hypothesis  $H_m$  using multiple testing procedures. Later, we will also consider this joint testing of many autocorrelations simultaneously in a multiple testing framework.

To review the testing procedure in application to this problem: let  $S_n$  be the symmetric, or permutation, group of order  $n$ . Then, given any test statistic  $T_n(X) = T_n(X_1, \dots, X_n)$ , for each element  $\pi_n \in S_n$ , let  $\hat{T}_{\pi_n} = T_n(X_{\pi_n(1)}, \dots, X_{\pi_n(n)})$ . Let the ordered values of the  $\hat{T}_{\pi_n}$  be

$$\hat{T}_n^{(1)} \leq \dots \leq \hat{T}_n^{(n!)} . \quad (1.3)$$

Fix a nominal level  $\alpha \in (0, 1)$ , and let  $m = n! - [\alpha n!]$ , where  $[x]$  denotes the largest integer less than or equal to  $x$ . Let  $M^+(x)$  and  $M^0(x)$  be the number of values  $\hat{T}_n^{(j)}(x)$  which are greater than and equal to  $\hat{T}_n^{(m)}(x)$  respectively. Let

$$a(x) = \frac{\alpha n! - M^+(x)}{M^0(x)} . \quad (1.4)$$

Define the permutation test  $\phi(X)$  to be equal to 1,  $a(X)$ , or 0, according to whether  $T_n(X)$  is greater than, equal to, or less than  $\hat{T}_n^{(m)}(X)$  respectively. Additionally, define the permutation distribution

$$\hat{R}_n^{T_n}(t) := \frac{1}{n!} \sum_{\pi_n \in S_n} I \{ \hat{T}_{\pi_n} \leq t \} . \quad (1.5)$$

Let  $[n] = \{1, \dots, n\}$ . We observe that, for  $\Pi_n \sim \text{Unif}(S_n)$ , independent of the sequence  $\{X_i, i \in [n]\}$ , and  $X_{\Pi_n} = (X_{\Pi_n(1)}, \dots, X_{\Pi_n(n)})$ , the permutation distribution is the distribution of  $T_n(X_{\Pi_n})$  conditional on the sequence  $\{X_i, i \in [n]\}$ . Also, accounting for discreteness, the permutation test rejects if the observed test statistic  $T_n$  exceeds the  $1 - \alpha$  quantile of the permutation distribution  $\hat{R}_n^{T_n}$ .

Under the randomization hypothesis that the joint distribution of the  $X_i$  is invariant under permutation, the permutation test  $\phi$  is exact level  $\alpha$  (see Lehmann and Romano, 2005, Theorem 15.2.1), but problems may arise when the null hypothesis  $H^{(k)} : \rho(k) = 0$  holds true, but the sequence  $X$  is not independent and identically distributed. Indeed, the distribution of an uncorrelated sequence is not invariant under permutations, and the randomization hypothesis does not hold (the randomization hypothesis guarantees finite-sample validity of

the permutation test; see Lehmann and Romano, 2005, Section 15.2). Such issues may hinder the use of permutation testing for valid inference, but we will show how to restore asymptotic validity to the permutation test.

For instance, consider the problem of testing  $H^{(1)} : \rho(1) = 0$ , for some sequence  $\{X_i, i \in [n]\} \sim P$ , where  $\rho(1) = 0$ . If the sequence is not i.i.d., the permutation test may have rejection probability significantly different from the nominal level, which leads to several issues. If the rejection probability is greater than the nominal level, we may reject the null hypothesis, and conclude that there is non-zero first order autocorrelation, whereas in fact we have autocorrelation of some higher order, or some other unobserved dependence structure. A further issue is that of Type 3, or directional, error, in a two-sided test of  $H^{(1)}$ . In this situation, one runs the risk of rejecting the null and concluding, for instance, that the first-order autocorrelation is larger than 0, when in fact it is less than 0. To illustrate this, if there exists some distribution  $P_n$  of the sequence  $\{X_i, i \in [n]\}$  with first-order autocorrelation  $\rho(1) = 0$  but rejection probability equal to  $\gamma \gg \alpha$ , by continuity it follows that there exists some distribution  $Q_n$  of the sequence with first-order autocorrelation  $\rho(1) < 0$ , but two-sided rejection probability almost as large as  $\gamma$ . Under such a distribution, with probability almost  $\gamma/2$ , not only would we reject the null, but we would also falsely conclude that the first-order sample autocorrelation is greater than 0, when in fact the opposite holds. We will later show that  $\gamma$  may be arbitrarily close to 1; see Example 3.3 and Remark 3.3. There are also issues if the rejection probability under the null is much smaller than the nominal level. In this case, again by continuity, we would have power significantly less than the nominal level even if the alternative is true, that is, the test would be biased. The strategy to overcome these issues is essentially as follows: assuming stationarity of the sequence  $\{X_i, i \geq 1\}$ , we wish to show that the permutation distribution based on some test statistic is asymptotically pivotal, that is, does not depend on the distribution of the  $X_i$ , in order for the critical region of the associated hypothesis test to not depend on parameters of the distribution of the  $X_i$ . We then wish to show that the limiting distribution of the test statistic under  $H^{(1)}$  is the same as the permutation distribution, so that we may perform (asymptotically) valid inference. Without this matching condition, we may not claim that a permutation test is asymptotically valid, despite being exact under the additional assumption of independence of the sequence. The main technical challenge in analyzing the ‘random’ permutation distribution stems from the fact that, under weak dependence assumptions, the rationale for permuting no longer applies as it did in the case of the sequence being independent and identically distributed.

Significant work has been done on these issues in the context of other problems. Neuhaus (1993) discovers the idea of studentizing test statistics to allow for asymptotically valid inference in the permutation testing setting, Janssen (1997) compares means by appropriate studentization in a permutation test, Janssen and Pauls (2003) give general results about permutation testing, Chung and Romano (2013) consider studentizing linear statistics in a two-sample setting, Omelka and Pauly (2012) compare correlations by permutation testing, and DiCiccio and Romano (2017) consider testing correlation structure and regression coefficients. In the context of time series data, Nichols and Holmes (2002) discuss the application of permutation testing to neuroimaging data, and Ptitsyn *et al.* (2006) consider the application of permutation testing as a method of testing for periodicity in biological data. In a more theoretical setting, Jentsch and Pauly (2015) use randomization methods to test equality of spectral densities, and Ritzwoller and Romano (2020) consider permutation testing in the setting of dependent Bernoulli sequences.

In the broader context of testing for serial correlation in time series, Francq *et al.* (2005) consider testing for autocorrelation of errors in ARMA models, Lobato *et al.* (2002) and Escanciano and Lobato (2009) provide improvements to the Box–Pierce test, Shao (2011) constructs a test using the bootstrap of the null hypothesis of white noise, and Zhu (2016) improves the Ljung–Box test in application to ARMA models by bootstrapping.

The goal of this article is to provide a framework for the use of permutation testing as a valid method for testing the hypothesis  $H^{(k)} : \rho(k) = 0$ , which retains the exactness property under the assumption of independence of the  $X_i$ , but is also asymptotically valid for a large class of weakly dependent stationary sequences. In particular, throughout this article, we consider the problem of testing  $H^{(1)} : \rho(1) = 0$ , for  $\{X_i, i \in [n]\}$  a weakly dependent sequence, and with test statistic a possibly studentized version of the sample autocorrelation  $\hat{\rho}_n = \hat{\rho}_n(1)$ , where

$$\hat{\rho}_n(k) \equiv \hat{\rho}_n(X_1, \dots, X_n; k) = \frac{\frac{1}{n-k} \sum_{i=1}^{n-k} (X_i - \bar{X}_n)(X_{i+k} - \bar{X}_n)}{\hat{\sigma}_n^2}. \quad (1.6)$$

$\hat{\sigma}_n^2$  is the sample variance, given by

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2, \quad (1.7)$$

and  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ . We note that  $\bar{X}_n$  and  $\hat{\sigma}_n^2$  are permutation invariant. Unless otherwise stated, we consider the problem of testing  $H^{(1)} : \rho_1 = 0$ , where  $\rho_1$  is the first-order autocorrelation, and  $\hat{\rho}_n$  refers to the first-order sample autocorrelation.

There are several different notions of weak dependence (see Bradley, 2005 for a discussion thereof). Throughout this article, we focus on the notions of  $m$ -dependence and  $\alpha$ -mixing.

Section 2 provides several useful preliminary definitions and results. The main results are given in Section 3, in which we give conditions for the asymptotic validity of the permutation test when  $\{X_i, i \in [n]\}$  is a stationary,  $\alpha$ -mixing sequence, satisfying fairly standard technical moment and mixing conditions. For example, the results in this section are applicable to a large class of stationary ARMA processes. These results should be compared with those of Section S.2 of the Supporting Information, which provides analogous results for  $m$ -dependent sequences under weaker moment conditions. The technical arguments for Sections S.2 and 3 are rather distinct, though the results in both sections allow one to construct valid permutation tests of correlations by appropriate studentization. Section 4 provides a framework for using individual permutation tests for different order autocorrelations in a multiple testing setting. Section 5 provides simulations illustrating the results. Section 6 gives an application of the testing procedure to financial data. The proofs are quite lengthy due to the technical requirements needed to prove the results; consequently, all proofs are deferred to the Supporting Information.

## 2. PRELIMINARIES

We establish notation and definitions which will be used in Section 3, in addition to providing a useful central limit theorem. We begin with a brief discussion of  $\alpha$ -mixing. Let  $\{X_n, n \in \mathbb{Z}\}$  be a stationary sequence of random variables, adapted to the filtration  $\{\mathcal{F}_n\}$ . Let  $\mathcal{G}_n = \sigma(X_r : r \geq n)$ . For  $n \in \mathbb{N}$ , let  $\alpha_X(n)$  be Rosenblatt's  $\alpha$ -mixing coefficient, defined as

$$\alpha_X(n) = \sup_{A \in \mathcal{F}_0, B \in \mathcal{G}_n} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|. \quad (2.1)$$

We say that  $\{X_n\}$  is  $\alpha$ -mixing if  $\alpha_X(n) \rightarrow 0$  as  $n \rightarrow \infty$ .

In Section 3, to obtain an asymptotically valid permutation testing procedure, it is necessary to compare the true limiting sample distribution to the permutation distribution. To this end, we provide the following result, which follows from a straightforward application of Ibragimov's central limit theorem (Ibragimov, 1962) and the delta method.

**Theorem 2.1.** Let  $X_1, \dots, X_n$  be a strictly stationary sequence, with variance  $\sigma^2 > 0$ , and first-order autocorrelation  $\rho_1$ , such that one of the following two conditions holds.

- (i)  $\{X_i, i \in [n]\}$  is  $m$ -dependent, for some  $m \in \mathbb{N}$ , and  $\mathbb{E}[X_1^4] < \infty$ .
- (ii)  $\{X_i, i \in [n]\}$  is  $\alpha$ -mixing, and, for some  $\delta > 0$ , we have that

$$\mathbb{E}[|X_1|^{4+2\delta}] < \infty, \quad (2.2)$$

and the  $\alpha$ -mixing coefficients  $\alpha_X(\cdot)$  satisfy

$$\sum_{n \geq 1} \alpha_X(n)^{\frac{\delta}{2+\delta}} < \infty. \quad (2.3)$$

Let  $\hat{\rho}_n$  be the sample first-order autocorrelation. Let

$$\begin{aligned} \kappa^2 &= \text{Var}(X_1^2) + 2 \sum_{k \geq 2} \text{Cov}(X_1^2, X_k^2) \\ \tau_1^2 &= \text{Var}(X_1 X_2) + 2 \sum_{k \geq 2} \text{Cov}(X_1 X_2, X_k X_{k+1}) \\ \nu_1 &= \text{Cov}(X_1 X_2, X_1^2) + \sum_{k \geq 2} \text{Cov}(X_1^2, X_k X_{k+1}) + \sum_{k \geq 2} \text{Cov}(X_1 X_2, X_k^2). \end{aligned} \quad (2.4)$$

Let

$$\gamma_1^2 = \frac{1}{\sigma^4} (\tau_1^2 - 2\rho_1 \nu_1 + \rho_1^2 \kappa^2). \quad (2.5)$$

Suppose that  $\kappa^2, \tau_1^2, \gamma_1^2 \in (0, \infty)$ . Then, as  $n \rightarrow \infty$ ,

$$\sqrt{n}(\hat{\rho}_n - \rho_1) \xrightarrow{d} N(0, \gamma_1^2). \quad (2.6)$$

A similar result to Theorem 2.1 has been shown in Francq *et al.* (2005). It is provided here to make clear the limiting distribution of the sample first-order autocorrelation. With this result in hand, we have an explicit basis with which to compare the limiting sample distribution and the permutation distribution. It will be useful to consider estimators of the asymptotic variance  $\gamma_1^2$  in Theorem 2.1. To this end, we provide the following definition.

**Definition 2.1.** Let  $\{X_i, i \in \mathbb{N}\}$  be a stationary sequence of random variables. For  $i \in \mathbb{N}$ , let  $Y_i = (X_i - \bar{X}_n)(X_{i+1} - \bar{X}_n)$ , and let  $Z_i = (X_i - \bar{X}_n)^2$ . Let  $\{b_n, n \in \mathbb{N}\}$  be a non-decreasing sequence of natural numbers. Let

$$\begin{aligned} \hat{K}_n^2 &= \frac{1}{n} \sum_{i=1}^n (Z_i - \bar{Z}_n)^2 + \frac{2}{n} \sum_{j=1}^{b_n} \sum_{i=1}^{n-j} (Z_i - \bar{Z}_n)(Z_{i+j} - \bar{Z}_n) \\ \hat{T}_n^2 &= \frac{1}{n} \sum_{i=1}^{n-1} (Y_i - \bar{Y}_n)^2 + \frac{2}{n} \sum_{j=1}^{b_n} \sum_{i=1}^{n-j-1} (Y_i - \bar{Y}_n)(Y_{i+j} - \bar{Y}_n) \\ \hat{\nu}_n &= \frac{1}{n} \sum_{i=1}^{n-1} (Y_i - \bar{Y}_n)(Z_i - \bar{Z}_n) + \frac{1}{n} \sum_{j=1}^{b_n} \sum_{i=1}^{n-j-1} (Z_i - \bar{Z}_n)(Y_{i+j} - \bar{Y}_n) + \\ &\quad + \frac{1}{n} \sum_{j=1}^{b_n} \sum_{i=1}^{n-j} (Y_i - \bar{Y}_n)(Z_{i+j} - \bar{Z}_n). \end{aligned} \quad (2.7)$$

Let

$$\hat{\gamma}_n^2 = \frac{1}{\hat{\sigma}_n^4} \left[ \hat{T}_n^2 - \hat{\rho}_n \hat{\nu}_n + \hat{\rho}_n^2 \hat{K}_n^2 \right]. \quad (2.8)$$

In Section 3, we will show consistency of the estimator  $\hat{\gamma}_n$  under standard moment and mixing conditions on the sequence  $\{X_i, i \in \mathbb{N}\}$ . Having established this notation, we may begin to consider the problem of testing for serial correlation for a broad class of  $\alpha$ -mixing sequences.

### 3. PERMUTATION DISTRIBUTION FOR $\alpha$ -MIXING SEQUENCES

We consider the problem of testing the null hypothesis

$$H^{(1)} : \rho_1 = 0, \quad (3.1)$$

where  $\rho_1 = \rho(\mathbf{X}; 1)$  is the first-order autocorrelation, in the setting where the sequence  $\{X_i, i \in [n]\}$  is stationary and  $\alpha$ -mixing. When the distribution of  $(X_1, \dots, X_n)$  is invariant under permutation, that is, the sequence is exchangeable, the randomization hypothesis holds, and so one may construct permutation tests of the hypothesis  $H_0$  with exact level  $\alpha$ . Note that in the case of  $\alpha$ -mixing, exchangeability and independence are equivalent conditions.<sup>1</sup> An implication of this result is that any permutation testing procedure will retain the exactness property under the additional assumption of independence of the  $X_i$ , and this is the only condition under which the randomization hypothesis holds.

However, if the realizations of the sequence are not independent, the test may not be valid even asymptotically, that is, the rejection probability of such a test need not be  $\alpha$  for finite samples or even near  $\alpha$  in the limit as  $n \rightarrow \infty$ . Hence the goal is to construct a testing procedure, based on some appropriately chosen test statistic, which has asymptotic rejection probability equal to  $\alpha$ , but which also retains the finite sample exactness property under the assumption of independence of the  $X_i$ . It is therefore important to analyze the asymptotic properties of the permutation distribution.

We wish to consider a permutation test based on the first-order sample autocorrelation,  $\hat{\rho}_n$ . Our strategy is as follows: to determine the limiting behavior of the permutation distribution,  $\hat{R}_n$ , we will show that an appropriately studentized version of the first-order sample autocorrelation has permutation distribution asymptotically not depending on the underlying process  $\{X_n, n \in \mathbb{N}\}$ , and that, under the null hypothesis  $H^{(1)} : \rho_1 = 0$ , the test statistic has asymptotic distribution equal to that of the permutation distribution.

We proceed as follows, in the spirit of Noether (1950). Suppose, for now, that the sequence  $\{X_n, n \in \mathbb{N}\}$  is uniformly bounded. For  $\Pi_n \sim \text{Unif}(S_n)$ , observing that the permutation distribution based on some test statistic  $T_n(X_1, \dots, X_n)$  is the empirical distribution of  $T_n(X_{\Pi_n(1)}, \dots, X_{\Pi_n(n)})$  conditional on the data  $\{X_i, i \in [n]\}$ , we may condition on the data and apply the central limit theorem of Wald and Wolfowitz (1943), checking that appropriate conditions on the sample variance are satisfied. This allows us to obtain a convergence result for a distribution very closely related to that of the permutation distribution, but with additional centering and scaling factors.

We are now in a position to use a double application of Slutsky's theorem for randomization distributions (Chung and Romano, 2013, Theorem 5.2), thus obtaining the following result.

**Theorem 3.1.** Let  $\{X_n, n \in \mathbb{N}\}$  be a stationary, bounded,  $\alpha$ -mixing sequence. Suppose that

$$\sum_{n \geq 1} \alpha_X(n) < \infty. \quad (3.2)$$

The permutation distribution of  $\sqrt{n}\hat{\rho}_n$  based on the test statistic  $\hat{\rho}_n = \hat{\rho}(X_1, \dots, X_n)$ , with associated group of transformations  $S_n$ , the symmetric group of order  $n$ , satisfies

$$\sup_{t \in \mathbb{R}} |\hat{R}_n(t) - \Phi(t)| \xrightarrow{P} 0, \quad (3.3)$$

as  $n \rightarrow \infty$ , where  $\Phi(t)$  is the distribution of a standard Gaussian random variable.

<sup>1</sup> A proof of this statement is given in Lemma S.3.1.

We now wish to remove the boundedness constraint of Theorem 3.1 and extend its result to the setting of stationary,  $\alpha$ -mixing sequences with uniformly bounded moments of some order. To do this, we require the following lemma.

**Lemma 3.1.** For each  $N, n \in \mathbb{N}$ , let  $G_{N, n} : \mathbb{R} \rightarrow \mathbb{R}$  be a non-decreasing random function such that, for each  $N$ , and for all  $t \in \mathbb{R}$ , as  $n \rightarrow \infty$ ,

$$G_{N, n}(t) \xrightarrow{p} g_N(t), \quad (3.4)$$

where, for each  $N$ ,  $g_N : \mathbb{R} \rightarrow \mathbb{R}$  is a function. Suppose further that, as  $N \rightarrow \infty$ , for each  $t \in \mathbb{R}$ ,

$$g_N(t) \rightarrow g(t), \quad (3.5)$$

where  $g : \mathbb{R} \rightarrow \mathbb{R}$  is continuous. Then, there exists a sequence  $N_n \rightarrow \infty$  such that, for all  $t \in \mathbb{R}$ , as  $n \rightarrow \infty$ ,

$$G_{N_n, n}(t) \xrightarrow{p} g(t). \quad (3.6)$$

We may proceed to extend Theorem 3.1 as follows. Let  $\{X_i, i \geq 1\}$  be an  $\alpha$ -mixing sequence, with summable  $\alpha$ -mixing coefficients, and let  $G_{N, n}(t)$  be the permutation distribution, evaluated at  $t$ , of the truncated sequence  $Y_i = (X_i \wedge N) \vee (-N)$ , based on the test statistic  $\sqrt{n}\hat{\rho}_n$ . Let  $g_N = g = \Phi$ , where  $\Phi$  is the distribution of a standard Gaussian random variable. By Theorem 3.1, the conditions of Lemma 3.1 are satisfied, so we apply Lemma 3.1 to find an appropriate sequence of truncation parameters  $N_n$ .

Then, for  $\Pi_n \sim \text{Unif}(S_n)$ , and  $Y_i = (X_i \wedge N_n) \vee (-N_n)$ , we relate the first-order sample autocorrelation  $\hat{\rho}_n(X_{\Pi_n(1)}, \dots, X_{\Pi_n(n)})$  to the truncated first-order sample autocorrelation  $\hat{\rho}_n(Y_{\Pi_n(1)}, \dots, Y_{\Pi_n(n)})$ . Bounding the difference of these two autocorrelations in probability using Doukhan (1994), Section 1.2.2, Theorem 3, and applying Slutsky's theorem for randomization distributions once more, we obtain the following result.

**Theorem 3.2.** Let  $\{X_n, n \in \mathbb{N}\}$  be a stationary,  $\alpha$ -mixing sequence, with mean 0 and variance 1. Suppose that, for some  $\delta > 0$ ,

$$\sum_{n \geq 1} \alpha_X(n)^{\frac{\delta}{2+\delta}} < \infty, \quad (3.7)$$

and

$$\mathbb{E} \left[ |X_1|^{8+4\delta} \right] < \infty. \quad (3.8)$$

Then, the permutation distribution of  $\sqrt{n}\hat{\rho}_n$  based on the test statistic  $\hat{\rho}_n = \hat{\rho}_n(X_1, \dots, X_n)$ , with associated group of transformations  $S_n$ , the symmetric group of order  $n$ , satisfies

$$\sup_{t \in \mathbb{R}} \left| \hat{R}_n(t) - \Phi(t) \right| \xrightarrow{p} 0, \quad (3.9)$$

where  $\Phi(t)$  is the distribution of a standard Gaussian random variable.

We have shown that the permutation distribution is asymptotically Gaussian, with mean and variance not depending on the underlying process, and that the result holds irrespective of whether or not the null hypothesis  $H^{(1)}$  holds. This result may be interpreted as follows. Under the action of a random permutation, for large values of  $n$ , one would expect that the first-order sample autocorrelation of the sequence  $\{X_n, n \in \mathbb{N}\}$  behaves similarly to the case of  $X_i \stackrel{\text{i.i.d.}}{\sim} F$ , where  $F$  is the marginal distribution of the  $X_i$ , since the dependence between consecutive

terms in the permuted sequence will be very weak, on account of the large sample size and the localized dependence structure of the original sequence. However, the same is not true of the asymptotic distribution of the test statistic. Indeed, under the null hypothesis, the asymptotic distribution of  $\sqrt{n}\hat{\rho}_n$  is also Gaussian with mean 0, but with variance not necessarily equal to 1. Therefore it is not possible to claim asymptotic validity of the permutation test based on this test statistic.

Since, clearly, in the context of Theorem 2.1,  $\gamma_1^2 = 1$  does not hold in general, a permutation test based on the test statistic  $\sqrt{n}\hat{\rho}_n$  will not be asymptotically valid. However, note that, under the additional restriction of independence,  $\gamma_1^2 = 1$  always, hence, as is consistent with the test being exact under independence, the permutation test will also be asymptotically valid in this case. One could conclude that the permutation test based on the above test statistic is not asymptotically valid in general.

Alternatively, one could adapt the test statistic above in some fashion, to resolve the issue of (asymptotically) mismatched variances in the permutation distribution and distribution of the test statistic. In particular, a natural way to adapt the test statistic is to studentize it by some estimator of  $\gamma_1^2$ , motivated by the heuristics that, under permutations, all dependence structure in the sequence will be broken, and the estimator will behave as if it were computed from an i.i.d. sample. Therefore, despite the limiting distribution of  $\sqrt{n}\hat{\rho}_n$  being different, in general, from the case when  $X_i \stackrel{\text{i.i.d.}}{\sim} F$ , where  $F$  is the marginal distribution of the  $X_i$ , under appropriate studentization, the limiting behaviors will be the same.

To this end, we now consider a permutation test based on some studentized version of the test statistic  $\sqrt{n}\hat{\rho}_n$ . Provided we can find a weakly consistent estimator  $\hat{\gamma}_n^2 = \hat{\gamma}_n^2(X_1, \dots, X_n)$  of  $\gamma_1^2$ , such that, for  $\Pi_n$  a random permutation independent of the sequence  $\{X_i, i \in [n]\}$ , we also have that  $\hat{\gamma}_n^2(X_{\Pi_n(1)}, \dots, X_{\Pi_n(n)}) = \text{Var}(X_1)^2 + o_p(1)$ , we may apply Slutsky's theorem for randomization distributions to studentize the test statistic and construct an asymptotically valid permutation test.

**Lemma 3.2.** Let  $\{X_n, n \in \mathbb{N}\}$  be a stationary,  $\alpha$ -mixing sequence, with variance  $\sigma^2 > 0$  such that, for some  $\delta > 0$ ,

$$\mathbb{E} \left[ |X_1|^{8+4\delta} \right] < \infty, \quad (3.10)$$

and

$$\sum_{n \geq 1} \alpha_X(n)^{\frac{\delta}{2+\delta}} < \infty. \quad (3.11)$$

Let  $\gamma_1^2$  be as defined in (2.5). Let  $b_n = o(\sqrt{n})$  be such that  $b_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Let  $\hat{\gamma}_n^2$  be as in (2.8). Then, as  $n \rightarrow \infty$ ,

$$\hat{\gamma}_n^2 \xrightarrow{P} \gamma_1^2. \quad (3.12)$$

**Lemma 3.3.** In the setting of Lemma 3.2, let  $\Pi_n \sim \text{Unif}(S_n)$ , independent of the sequence  $\{X_n, n \in \mathbb{N}\}$ . For  $i \in \mathbb{N}$ , let

$$\begin{aligned} \tilde{Y}_i &= (X_{\Pi_n(i)} - \bar{X}_n) (X_{\Pi_n(i+1)} - \bar{X}_n) \\ \tilde{Z}_i &= (X_{\Pi_n(i)} - \bar{X}_n)^2. \end{aligned} \quad (3.13)$$

Let

$$\hat{K}_n^2 = \frac{1}{n} \sum_{i=1}^n (\tilde{Z}_i - \bar{\tilde{Z}}_n)^2 + \frac{2}{n} \sum_{j=1}^{b_n} \sum_{i=1}^{n-j} (\tilde{Z}_i - \bar{\tilde{Z}}_n) (\tilde{Z}_{i+j} - \bar{\tilde{Z}}_n)$$



$$\begin{aligned}
\hat{T}_n^2 &= \frac{1}{n} \sum_{i=1}^{n-1} \left( \tilde{Y}_i - \bar{\tilde{Y}}_n \right)^2 + \frac{2}{n} \sum_{j=1}^{b_n} \sum_{i=1}^{n-j-1} \left( \tilde{Y}_i - \bar{\tilde{Y}}_n \right) \left( \tilde{Y}_{i+j} - \bar{\tilde{Y}}_n \right) \\
\hat{v}_n &= \frac{1}{n} \sum_{i=1}^{n-1} \left( \tilde{Y}_i - \bar{\tilde{Y}}_n \right) \left( \tilde{Z}_i - \bar{\tilde{Z}}_n \right) + \frac{1}{n} \sum_{j=1}^{b_n} \sum_{i=1}^{n-j-1} \left( \tilde{Z}_i - \bar{\tilde{Z}}_n \right) \left( \tilde{Y}_{i+j} - \bar{\tilde{Y}}_n \right) + \\
&\quad + \frac{1}{n} \sum_{j=1}^{b_n} \sum_{i=1}^{n-j} \left( \tilde{Y}_i - \bar{\tilde{Y}}_n \right) \left( \tilde{Z}_{i+j} - \bar{\tilde{Z}}_n \right).
\end{aligned} \tag{3.14}$$

Let

$$\hat{\gamma}_n^2 = \frac{1}{\hat{\sigma}_n^4} \left[ \hat{T}_n^2 - \hat{\rho}_n \hat{v}_n + \hat{\rho}_n^2 \hat{K}_n^2 \right]. \tag{3.15}$$

We have that, as  $n \rightarrow \infty$ ,

$$\hat{\gamma}_n^2 \xrightarrow{p} 1. \tag{3.16}$$

Combining the results of Lemmas 3.2 and 3.3 with Slutsky's theorem for randomization distributions, we may conclude the following.

**Theorem 3.3.** Let  $\{X_n, n \in \mathbb{N}\}$  be a strictly stationary,  $\alpha$ -mixing sequence, with variance  $\sigma^2$  and first-order autocorrelation  $\rho_1$ , such that, for some  $\delta > 0$ ,

$$\mathbb{E} \left[ |X_1|^{8+4\delta} \right] < \infty, \tag{3.17}$$

and

$$\sum_{n \geq 1} \alpha_X(n)^{\frac{\delta}{2+\delta}} < \infty. \tag{3.18}$$

Let  $\kappa^2$ ,  $\tau_1^2$ ,  $v_1$  and  $\gamma_1^2$  be as in Theorem 2.1. Suppose that  $\kappa^2$ ,  $\tau_1^2$ ,  $\gamma_1^2 \in (0, \infty)$ . Let  $b_n = o(\sqrt{n})$  be such that  $b_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Let  $\hat{\gamma}_n^2$  be as in (2.8).

(i) We have that, as  $n \rightarrow \infty$ ,

$$\frac{\sqrt{n}(\hat{\rho}_n - \rho_1)}{\hat{\gamma}_n} \xrightarrow{d} N(0, 1). \tag{3.19}$$

(ii) Let  $\hat{R}_n$  be the permutation distribution, with associated group of transformations  $S_n$ , the symmetric group of order  $n$ , based on the test statistic  $\sqrt{n}\hat{\rho}_n/\hat{\gamma}_n$ . Then, as  $n \rightarrow \infty$ ,

$$\sup_{t \in \mathbb{R}} \left| \hat{R}_n(t) - \Phi(t) \right| \xrightarrow{p} 0, \tag{3.20}$$

where  $\Phi$  is the standard Gaussian c.d.f.

**Remark 3.1.** Under the assumptions set out in Theorem 3.3, in particular as a result of (3.20), the level  $\alpha$  permutation test of the null  $H^{(1)} : \rho_1 = 0$  based on the test statistic  $\sqrt{n}\hat{\rho}_n/\hat{\gamma}_n$  is asymptotically valid.

**Remark 3.2.** Note that, under the stronger assumption of  $m$ -dependence of the sequence  $\{X_i, i \in \mathbb{N}\}$ , we have that the summability condition (3.18) on the  $\alpha$ -mixing coefficients holds for any  $\delta > 0$ , and so the result of Theorem 3.3 also holds in this case. However, one may obtain an analogous result to that of Theorem 3.3 with weaker moment conditions; see Section S.2 of the Supporting Information for further details.

We now illustrate the application of Theorem 3.3 to the class of stationary ARMA processes, as well as to a class of stationary  $m$ -dependent sequences.

**Example 3.1.** (ARMA process). Let  $\{X_i, i \in \mathbb{Z}\}$  satisfy the equation

$$\sum_{i=0}^p B_i X_{t-i} = \sum_{k=0}^q A_k \epsilon_k, \quad (3.21)$$

where the  $\epsilon_k$  are independent and identically distributed, and  $\mathbb{E}\epsilon_k = 0$ , that is,  $X$  is an ARMA( $p, q$ ) process. Let  $X$  have first-order autocorrelation  $\rho = 0$ . Let

$$P(z) := \sum_{i=0}^p B_i z^i. \quad (3.22)$$

If the equation  $P(z) = 0$  has no solutions inside the unit circle  $\{z \in \mathbb{C} : |z| \leq 1\}$ , there exists a unique stationary solution to (3.21). By Mokkadem (1988), Theorem 1, if the distribution of the  $\epsilon_k$  is absolutely continuous with respect to Lebesgue measure on  $\mathbb{R}$ , and also that, for some  $\delta > 0$ ,

$$\mathbb{E} \left[ |\epsilon_1|^{8+4\delta} \right] > 0, \quad (3.23)$$

we have that the sequence  $\{X_i, i \in \mathbb{N}\}$  satisfies the conditions of Theorem 3.3, as long as  $\gamma_1^2$ , as defined in (2.5), is finite and positive. Therefore, asymptotically, the rejection probability of the permutation test applied to such a sequence will be equal to the nominal level  $\alpha$ .

**Example 3.2.** (AR(2) process). We specialize Example 3.1 to the case of an AR(2) process with first-order autocorrelation equal to 0. Suppose that the strictly stationary sequence  $\{X_i, i \geq 1\}$  satisfies, for all  $t > 2$ , the equation

$$X_t = \phi X_{t-1} + \rho X_{t-2} + \epsilon_t, \quad (3.24)$$

where the  $\epsilon_t$  are as in Example 3.1. The first-order autocorrelation of  $X$  is given by

$$\rho(1) = \frac{\phi}{1 - \rho}. \quad (3.25)$$

Hence, for  $X$  to be such that  $\rho(1) = 0$ , we must have that  $\phi = 0$ . In particular, it follows that  $\{X_{2i}, i \geq 1\}$  and  $\{X_{2i-1}, i \geq 1\}$  are independent and identically distributed stationary AR(1) processes with parameter  $\rho$ . By the same argument as in Example 3.1, the requisite  $\alpha$ -mixing condition is satisfied, and so, in order for the result of Theorem 3.3 to apply, it suffices to show that  $\tau_1^2$ ,  $\kappa^2$ , and  $\nu_1$ , as defined in (2.7), are finite and non-zero. If so, since  $\rho(1) = 0$ , we have that  $\gamma_1^2 = \tau_1^2 / \sigma^4$ , and so the variance condition on  $\gamma_1^2$  is automatically satisfied. We begin by noting that, for  $i$  odd,

$$\text{Cov}(X_1, X_i) = \rho^{\frac{i-1}{2}} \text{Var}(X_1), \quad (3.26)$$

and similarly for the covariance between  $X_2$  and  $X_i$ , for  $i$  even. Also, note that  $\mathbb{E}[X_1] = 0$ . Simple calculations show that

$$\begin{aligned}\text{Var}(X_1 X_2) &= \text{Var}(X_1)^2 \\ \text{Cov}(X_1 X_2, X_k X_{k+1}) &= \rho^{k-1} \text{Var}(X_1)^2, \quad k \geq 2.\end{aligned}\quad (3.27)$$

It follows that  $\tau_1^2 \in (0, \infty)$ . Similarly, we have that  $\nu_1 = 0$ , and

$$\text{Cov}(X_1^2, X_k^2) = \begin{cases} \rho^{k-1} \text{Var}(X_1^4), & \text{if } k \in 2\mathbb{N}, \\ 0, & \text{otherwise.} \end{cases}\quad (3.28)$$

It follows that  $\gamma_1^2 \in (0, \infty)$ , and so the result of Theorem 3.3 holds in this case.

**Example 3.3.** (Products of i.i.d. random variables). Let  $\{Z_n, n \in \mathbb{N}\}$  be mean zero, i.i.d., non-constant random variables, such that

$$\mathbb{E}[|Z_1|^{8+4\delta}] < \infty. \quad (3.29)$$

Fix  $m \in \mathbb{N}$ , and, for each  $i$ , let

$$X_i = \prod_{j=i}^{i+m-1} Z_j. \quad (3.30)$$

We observe that the sequence  $\{X_i, i \geq 1\}$  is stationary and  $m$ -dependent, and that, by Fubini's theorem, the  $X_i$  have uniformly bounded  $(8 + 4\delta)$  moments. It now suffices to show that  $\kappa^2$  and  $\tau_1^2$  are finite and strictly greater than 0, and  $\gamma_1^2$ , as defined in (2.5), is finite and strictly greater than zero. Let  $M_k$  be the  $k$ th moment of  $Z_1$ ,  $k \geq 1$ . Simple calculations show that

$$\begin{aligned}\text{Var}(X_1 X_2) &= M_2^2 M_4^{m-1} \\ \text{Cov}(X_1 X_2, X_k X_{k+1}) &= 0, \quad k \geq 2.\end{aligned}\quad (3.31)$$

Hence

$$\tau_1^2 = M_2^2 M_4^{m-1} \in (0, \infty).$$

Additionally, we have that

$$\begin{aligned}\kappa^2 &= M_4^m - M_2^{2m} \in (0, \infty) \\ \nu_1 &= 0.\end{aligned}$$

Hence we have that  $\gamma_1^2 = M_4^{m-1} / M_2^{2(m-1)} \in (0, \infty)$ . It follows that we may apply the result of Theorem 3.2, and conclude that the rejection probability of the permutation test based on the test statistic  $\sqrt{n} \hat{\rho}_n / \hat{\gamma}_n$  converges to  $\alpha$  as  $n \rightarrow \infty$ .

**Remark 3.3.** Example 3.3 also provides an illustrative example of the need for studentization in the permutation test. Indeed, in the setting of Example 3.3, for  $r \in \mathbb{N}$  odd, let

$$Z_i = G_i^r, \quad (3.32)$$

where  $\{G_i, i \in \mathbb{Z}\}$  are independent standard Gaussian random variables. Hence

$$\gamma_1^2 = \left( \frac{\mathbb{E}[G_i^{4r}]}{\mathbb{E}[G_i^{2r}]^2} \right)^{m-1} = \left( \frac{(4r-1)!!}{((2r-1)!!)^2} \right)^{m-1}.$$

By Theorems 2.1 and 3.2, it follows that the asymptotic rejection probability of the level  $\alpha$  two-sided permutation test, based on the test statistic  $\sqrt{n}\hat{\rho}_n$ , converges to

$$2 \left( 1 - \Phi \left( \left( \frac{((2r-1)!!)^2}{(4r-1)!!} \right)^{\frac{m-1}{2}} z_{1-\alpha/2} \right) \right),$$

as  $n \rightarrow \infty$ , where  $z_{1-\alpha/2}$  is the  $\alpha/2$  quantile of the standard normal distribution. It follows that this rejection probability can be arbitrarily close to 1 for large values of  $n$ ,  $m$  and  $r$ . Therefore, by continuity, there exists a distribution  $Q_n$  of  $(X_1, \dots, X_n)$  such that  $\rho(1) < 0$ , but the two-sided permutation test based on  $\sqrt{n}\hat{\rho}_n$  would reject  $H^{(1)}$ , with probability arbitrarily close to  $1/2$ , and conclude that the first-order sample autocorrelation is greater than zero, when, in fact, the opposite is true. \*\*\*

**Remark 3.4.** In this section, we have only considered a permutation test of the hypothesis  $H^{(1)} : \rho(1) = 0$ . However, analogously, one may prove a similar result for a permutation testing procedure for the hypothesis  $H^{(k)} : \rho_k = \rho(k) = 0$ , where  $k \in \mathbb{N}$  is fixed. Indeed, note that the sequence  $\{\tilde{Y}_i : i \geq 1\}$ , where  $\tilde{Y}_i = X_i X_{i+k}$  is  $\alpha$ -mixing, with  $\alpha$ -mixing coefficients given by  $\alpha_\varepsilon(n) = \alpha_X(n-k)$ .

Furthermore, for  $\Pi_n$  a random permutation independent of the  $X_i$ , under appropriate moment conditions for the  $X_i$ , we also have that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n-1} X_{\Pi_n(i)} X_{\Pi_n(i+1)} \stackrel{d}{=} \frac{1}{\sqrt{n}} \sum_{i=1}^{n-k} X_{\Pi_n(i)} X_{\Pi_n(i+k)} + o_p(1), \quad (3.33)$$

since, for any fixed element  $\sigma \in S_n$ ,  $\Pi_n \sigma \stackrel{d}{=} \Pi_n$ . Hence, defining an appropriate estimator of the variance of  $\hat{\rho}_n(k)$ , we may similarly construct an asymptotically valid permutation test under the hypothesis  $H^{(k)}$ . To be precise, let  $Y_i = (X_i - \bar{X}_n)(X_{i+k} - \bar{X}_n)$ , and let  $Z_i = (X_i - \bar{X}_n)^2$ . Let

$$\begin{aligned} \hat{T}_{n,k}^2 &= \frac{1}{n} \sum_{i=1}^{n-1} (Y_i - \bar{Y}_n)^2 + \frac{2}{n} \sum_{j=1}^{b_n} \sum_{i=1}^{n-j-k} (Y_i - \bar{Y}_n)(Y_{i+j} - \bar{Y}_n) \\ \hat{v}_{n,k} &= \frac{1}{n} \sum_{i=1}^{n-1} (Y_i - \bar{Y}_n)(Z_i - \bar{Z}_n) + \frac{1}{n} \sum_{j=1}^{b_n} \sum_{i=1}^{n-j-k} (Z_i - \bar{Z}_n)(Y_{i+j} - \bar{Y}_n) + \\ &\quad + \frac{1}{n} \sum_{j=1}^{b_n} \sum_{i=1}^{\min\{n-j, n-k\}} (Y_i - \bar{Y}_n)(Z_{i+j} - \bar{Z}_n). \end{aligned} \quad (3.34)$$

Let  $\hat{K}_n^2$  and  $\kappa^2$  be defined as in Definition 2.1 and Theorem 2.1 respectively. Let

$$\hat{\gamma}_{n,k}^2 = \frac{1}{\hat{\sigma}_n^4} \left( \hat{T}_{n,k}^2 - 2\hat{\rho}_n(k)\hat{v}_{n,k} + \hat{\rho}_n(k)^2 \hat{K}_n^2 \right), \quad (3.35)$$

where  $\hat{\rho}_n(k)$  is the sample  $k$  th order autocorrelation. Let

$$\begin{aligned}\tau_k^2 &= \text{Var}(X_1 X_{k+1}) + 2 \sum_{j \geq 2} \text{Cov}(X_1 X_{k+1}, X_j X_{j+k}) \\ \nu_k &:= \text{Cov}(X_1 X_{k+1}, X_1^2) + \sum_{j \geq 2} \text{Cov}(X_1^2, X_j X_{j+k}) + \sum_{j \geq 2} \text{Cov}(X_1 X_{k+1}, X_j^2),\end{aligned}\quad (3.36)$$

and let

$$\gamma_k^2 = \frac{1}{\sigma^4} (\tau_k^2 - 2\rho_k \nu_k + \rho_k^2 \kappa^2). \quad (3.37)$$

Assume that  $\tau_k^2, \kappa^2 \in \mathbb{R}_+$  and  $\gamma_k^2 \in \mathbb{R}_+$ . Under the same conditions on the sequence  $\{X_n, n \in \mathbb{N}\}$  as in Theorem 3.3, by an identical argument to the one given in the case  $k = 1$ , we will have that the permutation distribution based on the test statistic  $\sqrt{n}\hat{\rho}_n(k)/\hat{\gamma}_n$ , with associated group of transformations  $S_n$ , will satisfy (3.20), and the test statistic will satisfy a central limit theorem analogous to (3.19).

Having developed a permutation testing framework, we further derive an array version of Theorem 3.3, to provide a procedure under which one may compute the limiting power of the permutation test under local alternatives.

**Theorem 3.4.** For each  $n \in \mathbb{N}$ , let  $\{X_i^{(n)}, i \in [n]\}$ , be stationary sequences of random variables. Suppose that the  $X_i^{(n)}$  are bounded, uniformly in  $i$  and  $n$ , and that

$$\sum_{n \geq 1} \sup_{r \geq n+1} \alpha_{X^{(r)}}(n) < \infty. \quad (3.38)$$

The permutation distribution of  $\sqrt{n}\hat{\rho}_n$ , based on the test statistic  $\hat{\rho}_n = \hat{\rho}_n(X_1^{(n)}, \dots, X_n^{(n)})$ , with associated group of transformations  $S_n$ , the symmetric group of order  $n$ , satisfies, as  $n \rightarrow \infty$ ,

$$\sup_{t \in \mathbb{R}} |\hat{R}_n(t) - \Phi(t)| \xrightarrow{P} 0. \quad (3.39)$$

We may view Theorem 3.4 as an extension of Theorem 3.1. Analogously, we may extend Theorem 3.2, and Lemmas 3.2 and 3.3, and hence the result of Theorem 3.3 holds for triangular arrays of stationary,  $\alpha$ -mixing sequences, replacing the condition (3.17) with

$$\sup_{n \geq 1} \mathbb{E} \left[ |X_1^{(n)}|^{8+4\delta} \right] < C, \quad (3.40)$$

and the condition (3.18) with

$$\sum_{n \geq 1} \max_{r \geq n+1} \alpha_{X^{(r)}}(n)^{\frac{\delta}{2+\delta}} < \infty. \quad (3.41)$$

In particular, it follows that we may apply the result of Theorem 3.4 to triangular arrays of stationary sequences, where instead of (3.19), we have the result

$$\frac{\sqrt{n}(\hat{\rho}_n - \rho_n)}{\hat{\gamma}_n} \xrightarrow{d} N(0, 1), \quad (3.42)$$

where  $\rho_n$ , and  $\hat{\gamma}_n$  are defined analogously to Theorem 3.3.

It follows that we may compute the power function of the permutation test under appropriate sequences of limiting local alternatives.

**Example 3.4.** (AR(1) process) Consider a triangular array of AR(1) processes, given by

$$X_i^{(n)} = \rho_n X_{i-1}^{(n)} + \epsilon_i^{(n)}, \quad i \in \{2, \dots, n\}, \quad (3.43)$$

where  $\rho_n = h/\sqrt{n}$ , for some fixed constant  $h \in (0, 1)$ , and the  $\epsilon_i^{(n)}$  form a triangular array of independent standard Gaussian random variables. For each  $n$ , the autoregressive process defined in (3.43) has a unique stationary solution, in which  $(X_1^{(n)}, \dots, X_n^{(n)})$  follows a multivariate Gaussian distribution, with mean 0 and covariance matrix given by

$$\text{Cov}(X_i^{(n)}, X_j^{(n)}) = \frac{\rho_n^{|i-j|}}{1 - \rho_n^2}. \quad (3.44)$$

Consider the problem of testing  $H^{(1)} : \rho_1 = 0$  against the alternative  $\rho_1 > 0$ , using the permutation test described in Theorem 3.4.

By Theorem 1 of Mokkadem (1988), we have that condition (3.41) is satisfied for example,  $\delta = 1/2$ , and, since the  $X_i^{(n)}$  have uniformly bounded second moment and are normally distributed, we also have that condition (3.40) is satisfied.

Hence, letting  $\phi_n$  denote the permutation test conducted on the sequence  $\{X_i^{(n)}, i \in [n]\}$ , we may apply the analogous result of Theorem 3.4 to the triangular array of AR processes, whence, by an application of Slutsky's theorem, we obtain that

$$\frac{\sqrt{n}\hat{\rho}_n}{\hat{\gamma}_n} - \frac{h}{\gamma^{(n)}} \xrightarrow{d} N(0, 1), \quad (3.45)$$

and that the local limiting power function satisfies, for  $z_{1-\alpha}$  the upper  $\alpha$  quantile of the standard Gaussian distribution,

$$\mathbb{E}_{\rho_n} \phi_n \rightarrow 1 - \Phi\left(z_{1-\alpha} - \lim_{n \rightarrow \infty} \frac{h}{\gamma^{(n)}}\right), \quad (3.46)$$

and

$$(\gamma^{(n)})^2 = \frac{1}{\sigma_n^4} \left[ (\tau^{(n)})^2 - 2\rho_n v_1^{(n)} + \rho_n^2 (\kappa^{(n)})^2 \right], \quad (3.47)$$

where

$$\begin{aligned} (\kappa^{(n)})^2 &= \text{Var}\left(\left(X_1^{(n)}\right)^2\right) + 2 \sum_{k \geq 2} \text{Cov}\left(\left(X_1^{(n)}\right)^2, \left(X_k^{(n)}\right)^2\right) \\ (\tau^{(n)})^2 &= \text{Var}\left(X_1^{(n)} X_2^{(n)}\right) + 2 \sum_{k \geq 2} \text{Cov}\left(X_1^{(n)} X_2^{(n)}, X_k^{(n)} X_{k+1}^{(n)}\right) \\ v_1^{(n)} &:= \text{Cov}\left(X_1^{(n)} X_2^{(n)}, \left(X_1^{(n)}\right)^2\right) + \sum_{k \geq 2} \text{Cov}\left(\left(X_1^{(n)}\right)^2, X_k^{(n)} X_{k+1}^{(n)}\right) + \\ &\quad + \sum_{k \geq 2} \text{Cov}\left(X_1^{(n)} X_2^{(n)}, \left(X_k^{(n)}\right)^2\right). \end{aligned} \quad (3.48)$$

Example 7.16 of van der Vaart (1998) establishes local asymptotic normality of the local alternative sequence to the null model corresponding to  $h = 0$ . Hence, by contiguity of the sequence of alternatives, it follows that, as  $n \rightarrow \infty$ ,  $\gamma^{(n)} \rightarrow 1$ , and so, as  $n \rightarrow \infty$ ,

$$\mathbb{E}_{\rho_n} \phi_n \rightarrow 1 - \Phi(z_{1-\alpha} - h). \quad (3.50)$$

**Remark 3.5.** We observe that, in the setting of Example 3.4, the one-sided studentized permutation test is locally asymptotically uniformly most powerful (LAUMP; see Lehmann and Romano, 2005, Definition 13.3.3). Indeed, by Example 7.16 and Theorem 15.4 in van der Vaart (1998), coupled with the result of Lemma 13.3.2 in Lehmann and Romano (2005), we observe that the optimal local power of a one-sided test, against the alternatives  $\rho_n = h/\sqrt{n}$ , in the setting of Example 3.4, is

$$\beta^* = 1 - \Phi(z_{1-\alpha} - h).$$

Since this is exactly the power in (3.50), it follows that the studentized permutation test is LAUMP.

**Remark 3.6.** Note that, more generally, the same argument applies when computing the limiting local power of the studentized permutation test with respect to contiguous alternatives. Indeed, if contiguity can be established for some sequence of alternatives  $\{P_n, n \in \mathbb{N}\}$  with first-order autocorrelations  $\rho_{1,n} = h/\sqrt{n}$ , by a similar argument to the one presented in Example 3.4, we will have that the convergence of  $\hat{\gamma}_n^2$  to  $\gamma_1^2$  (in probability) also holds under the contiguous sequence of alternatives. Hence the limiting power of the one-sided level  $\alpha$  studentized permutation test, under the contiguous sequence of alternatives, will also be given by

$$\mathbb{E}_{P_n} \phi_n \rightarrow 1 - \Phi\left(z_{1-\alpha} - \frac{h}{\gamma_1}\right),$$

as  $n \rightarrow \infty$ .

#### 4. MULTIPLE AND JOINT HYPOTHESIS TESTING

We outline multiple testing procedures which may be applied to test the hypotheses  $H^{(k)}$ , as defined in (1.2), simultaneously. While we make use of the standard Bonferroni method of combining  $p$ -values, we argue such an approach is not overly conservative. We develop a method for testing joint null hypotheses of the form

$$H_r : \rho(1) = \rho(2) = \cdots = \rho(r) = 0.$$

It is desirable to perform such a test in a multiple testing framework, that is, in the case of rejection of the null hypothesis, we often wish to accompany this rejection with inference on which of the individual hypotheses  $H^{(k)}$  do not hold. To this end, it is necessary to construct a procedure allowing for valid inference, in the sense that the family-wise error rate (FWER) is controlled at the nominal level  $\alpha$ . In general, we may apply the canonical Bonferroni correction; that is, given marginal  $p$ -values  $\hat{p}_1, \dots, \hat{p}_r$  and a nominal level  $\alpha$ , we reject the null hypothesis  $H_r$  if  $\min_i \hat{p}_i \leq \alpha/r$ , and assert that the hypothesis  $H^{(k)}$  does not hold for any  $k$  such that  $\hat{p}_k \leq \alpha/r$ . Under the further assumption of independence of  $p$ -values, we may use multiple testing procedures such as the Šidák correction, which rejects any  $H^{(k)}$  for which  $\hat{p}_k \leq 1 - (1 - \alpha)^{1/r}$ .

This procedure is marginally more powerful than the canonical Bonferroni procedure, but may not control the FWER at the nominal level  $\alpha$  if there is negative dependence between the  $\hat{p}_k$ . To understand the dependence structure between sample autocorrelations, and their corresponding permutation  $p$ -values, we have the following result.

**Theorem 4.1.** In the setting of Theorem 3.3, let  $r \in \mathbb{N}$ ,  $r > 1$ . For  $k \in [r]$ , let  $\rho_k$  be the  $k$ th-order autocorrelation, and let  $\hat{\rho}_k$  be the  $k$ th-order sample autocorrelation. Let  $\Sigma \in \mathbb{R}^{(r+1) \times (r+1)} = (\sigma_{ij})_{i,j=0}^r$  be such that

$$\sigma_{ij} = \begin{cases} \text{Var}(X_1 X_{1+i}) + 2 \sum_{l>1} \text{Cov}(X_1 X_{1+i}, X_l X_{l+i}), & i = j \\ \text{Cov}(X_1 X_{1+i}, X_1 X_{1+j}) + \sum_{l>1} [\text{Cov}(X_1 X_{1+i}, X_l X_{l+j}) + \text{Cov}(X_1 X_{1+j}, X_l X_{l+i})], & i \neq j. \end{cases}$$

Let  $A \in \mathbb{R}^{(r+1) \times r}$  be given by

$$A = \begin{pmatrix} -\frac{\rho_1}{\sigma^4} & \dots & \dots & \dots & -\frac{\rho_r}{\sigma^4} \\ \frac{1}{\sigma^2} & 0 & \dots & \dots & 0 \\ 0 & \frac{1}{\sigma^2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & \dots & \frac{1}{\sigma^2} \end{pmatrix}.$$

Then, as  $n \rightarrow \infty$ ,

$$\sqrt{n} \left[ \begin{pmatrix} \hat{\rho}_1 \\ \vdots \\ \hat{\rho}_r \end{pmatrix} - \begin{pmatrix} \rho_1 \\ \vdots \\ \rho_r \end{pmatrix} \right] \xrightarrow{d} N(0, A^T \Sigma A).$$

**Remark 4.1.** We observe that, in the i.i.d. setting, the sample autocorrelations are asymptotically independent. Indeed, in this case, we have that  $\Sigma$ , as defined in (4.1), is diagonal, and, for  $i \neq j$ , for all  $l \in \{0, \dots, r\}$ ,  $A_{li} A_{lj} = 0$ . Therefore, for  $i, j \in [r]$ ,  $i \neq j$ ,

$$(A^T \Sigma A)_{ij} = \sum_{l,s=0}^r A_{li} A_{sj} \sigma_{ls} = \sum_{l=0}^r A_{li} A_{lj} \sigma_{ll} = 0.$$

By Remark 4.1 and the uniform convergence of the permutation distribution  $\hat{R}_n$  in Theorem 3.3, we have that, for  $1 \leq k \leq r$ , leaving the dependence of  $\hat{\rho}_k$  on  $n$  implicit,

$$\hat{\rho}_k = 1 - \Phi \left( \frac{\hat{\rho}_n(k)}{\hat{\gamma}_{n,k}} \right) + o_p(1),$$

where  $\hat{\gamma}_{n,k}$  is as defined in (3.35). It follows that, in some settings, such as the i.i.d. setting, the marginal  $p$ -values are asymptotically independent. Therefore we may use the Šidák correction if, for instance, we use the null hypothesis  $H_r$  as a portmanteau test of independence of realizations. However, more generally, using the Bonferroni cutoff of  $\alpha/r$  is only marginally larger than the Bonferroni-Šidák correction, and it applies irrespective of the asymptotic dependence structure of the marginal  $p$ -values. Indeed, since, by Theorem 4.1 and Remark 4.1, any method must at least account for possibility of asymptotic independence, it follows that the cutoff should be at least as large as the Šidák correction. But since the Bonferroni correction is not much larger than the Šidák correction, there does not appear to be much gain, in terms of power, in devising a method that precisely accounts for the joint dependence among the marginal  $p$ -values. Despite this, we may use a step-down procedure, such as that of Holm (1979), to obtain a larger power.

We illustrate the application of the canonical Bonferroni procedure in Section 6, in application to historical log-return data.



Table I. Monte Carlo simulation results for null rejection probabilities for tests of  $\rho(1) = 0$ , in an  $m$ -dependent Gaussian product setting

| $m$ | $n$         | 10     | 20     | 50     | 80     | 100    | 500    | 1000   |
|-----|-------------|--------|--------|--------|--------|--------|--------|--------|
| 0   | Stud. Perm. | 0.0511 | 0.0489 | 0.0465 | 0.0452 | 0.0500 | 0.0488 | 0.0525 |
|     | Unst. Perm. | 0.0503 | 0.0511 | 0.0493 | 0.0470 | 0.0494 | 0.0480 | 0.0527 |
|     | Ljung–Box   | 0.0534 | 0.0544 | 0.0474 | 0.0488 | 0.0516 | 0.0488 | 0.0482 |
|     | Box–Pierce  | 0.0198 | 0.0365 | 0.0407 | 0.0448 | 0.0484 | 0.0480 | 0.0478 |
| 1   | Stud. Perm. | 0.0654 | 0.0578 | 0.0611 | 0.0586 | 0.0582 | 0.0532 | 0.0534 |
|     | Unst. Perm. | 0.1010 | 0.1249 | 0.1388 | 0.1512 | 0.1553 | 0.1692 | 0.1749 |
|     | Ljung–Box   | 0.0888 | 0.1390 | 0.1873 | 0.2084 | 0.2102 | 0.2359 | 0.2651 |
|     | Box–Pierce  | 0.0342 | 0.1057 | 0.1737 | 0.2001 | 0.2039 | 0.2341 | 0.2645 |
| 2   | Stud. Perm. | 0.0718 | 0.0638 | 0.0661 | 0.0615 | 0.0683 | 0.0608 | 0.0582 |
|     | Unst. Perm. | 0.1288 | 0.1588 | 0.2041 | 0.2189 | 0.2327 | 0.2580 | 0.2721 |
|     | Ljung–Box   | 0.0999 | 0.1912 | 0.2975 | 0.3420 | 0.3494 | 0.4425 | 0.4645 |
|     | Box–Pierce  | 0.0455 | 0.1555 | 0.2841 | 0.3319 | 0.3410 | 0.4414 | 0.4638 |
| 3   | Stud. Perm. | 0.0714 | 0.0708 | 0.0638 | 0.0713 | 0.0706 | 0.0647 | 0.0566 |
|     | Unst. Perm. | 0.1411 | 0.1716 | 0.2332 | 0.2582 | 0.2748 | 0.3269 | 0.3364 |
|     | Ljung–Box   | 0.1000 | 0.2056 | 0.3404 | 0.4026 | 0.4310 | 0.5644 | 0.6034 |
|     | Box–Pierce  | 0.0451 | 0.1693 | 0.3252 | 0.3946 | 0.4233 | 0.5634 | 0.6033 |

Table II. Monte Carlo simulation results for null rejection probabilities for tests of  $\rho(1) = 0$ , in an  $\alpha$ -mixing setting

| Distribution                | $n$         | 10     | 20     | 50     | 80     | 100    | 500    | 1000   |
|-----------------------------|-------------|--------|--------|--------|--------|--------|--------|--------|
| AR(2), N(0, 1) innov.       | Stud. Perm. | 0.0418 | 0.0215 | 0.0399 | 0.0420 | 0.0448 | 0.0464 | 0.0480 |
|                             | Unst. Perm. | 0.0594 | 0.0901 | 0.1212 | 0.1403 | 0.1372 | 0.1541 | 0.1658 |
|                             | Ljung–Box   | 0.2101 | 0.2360 | 0.2570 | 0.2492 | 0.2537 | 0.2527 | 0.2607 |
|                             | Box–Pierce  | 0.1283 | 0.1972 | 0.2411 | 0.2399 | 0.2468 | 0.2516 | 0.2603 |
| AR(2) Prod., N(0, 1) innov. | Stud. Perm. | 0.0385 | 0.0370 | 0.0332 | 0.0375 | 0.0382 | 0.0350 | 0.0366 |
|                             | Unst. Perm. | 0.0555 | 0.0648 | 0.0836 | 0.0923 | 0.0939 | 0.1136 | 0.1181 |
|                             | Ljung–Box   | 0.0938 | 0.0878 | 0.1012 | 0.1151 | 0.1189 | 0.1674 | 0.1708 |
|                             | Box–Pierce  | 0.0493 | 0.0658 | 0.0913 | 0.1081 | 0.1138 | 0.1653 | 0.1705 |
| AR(2), $U[-1, 1]$ innov.    | Stud. Perm. | 0.0496 | 0.0243 | 0.0390 | 0.0433 | 0.0444 | 0.0470 | 0.0464 |
|                             | Unst. Perm. | 0.0628 | 0.0940 | 0.1276 | 0.1412 | 0.1395 | 0.1569 | 0.1572 |
|                             | Ljung–Box   | 0.2272 | 0.2464 | 0.2545 | 0.2539 | 0.2522 | 0.2530 | 0.2522 |
|                             | Box–Pierce  | 0.1403 | 0.2051 | 0.2394 | 0.2434 | 0.2445 | 0.2517 | 0.2517 |
| AR(2), $t_{9.5}$ innov.     | Stud. Perm. | 0.0432 | 0.0218 | 0.0385 | 0.0423 | 0.0437 | 0.0531 | 0.0459 |
|                             | Unst. Perm. | 0.0582 | 0.0902 | 0.1182 | 0.1346 | 0.1362 | 0.1581 | 0.1634 |
|                             | Ljung–Box   | 0.2050 | 0.2316 | 0.2567 | 0.2522 | 0.2524 | 0.2654 | 0.2590 |
|                             | Box–Pierce  | 0.1206 | 0.1941 | 0.2416 | 0.2436 | 0.2455 | 0.2638 | 0.2579 |

## 5. SIMULATION RESULTS

Monte Carlo simulations illustrating our results are given in this section. Tables I and II tabulate the rejection probabilities of one-sided tests for the permutation tests, in addition to those of the Ljung–Box and Box–Pierce tests. The nominal level considered is  $\alpha = 0.05$ . The simulation results confirm that the permutation test is valid, in that, in large samples, it approximately attains level  $\alpha$ . The simulation results also confirm that, by contrast, both the Ljung–Box and Box–Pierce tests perform extremely poorly in non-i.i.d. settings.

As a review, the Ljung–Box and Box–Pierce tests are used to test for independence of residuals in fitting ARMA models. This is done by a portmanteau test of the null hypothesis  $H_r$ , as defined in (1.1), which is tested under the assumption that the residuals follow a Gaussian white noise process. For each  $k \in \mathbb{N}$ , let  $\hat{\rho}_k$  be the sample  $k$ th order autocorrelation. The one-sided Ljung–Box test compares the test statistic

$$\hat{Q}_{LB, n} = n(n+2) \sum_{k=1}^r \frac{\hat{\rho}_k^2}{n-k} \quad (5.1)$$

to the quantiles of a  $\chi_r^2$  distribution, with rejection occurring for large values of  $\hat{Q}_{LB, n}$ . Similarly, the one-sided Box–Pierce test compares the test statistic

$$\hat{Q}_{BP, n} = n \sum_{k=1}^r \hat{\rho}_k^2 \quad (5.2)$$

to the quantiles of a  $\chi_r^2$  distribution, with rejection occurring for large values of  $\hat{Q}_{BP, n}$ . In the case of  $r = 1$ , both the one-sided Ljung–Box and Box–Pierce tests compare the test statistic

$$\hat{Q}_n = C_n \hat{\rho}_1^2 \quad (5.3)$$

to the quantiles of a  $\chi_1^2$  distribution, with rejection in both tests occurring for large values of  $\hat{Q}_n$ . The Ljung–Box and Box–Pierce tests primarily differ in their scaling in this case; namely, the Ljung–Box test takes  $C_n = n(n+2)/(n-1)$  in (5.3), while the Box–Pierce test uses  $C_n = n$ .

In this simulation, we consider both  $m$ -dependent (in Table I) and  $\alpha$ -mixing (in Table II) processes. Table I gives the null rejection probabilities for sampling distributions of the form described in Example 3.3, in the case of Gaussian products, where the values of  $m$  are listed in the first column. Note that  $m = 0$  corresponds to the setting where the  $X_i$  are independent standard Gaussian random variables.

Table II gives the null rejection probabilities for processes of the form described in Example 3.2, with  $\rho = 0.5$ . We include one additional example in the second row of Table II. The sample distribution in this row is as follows.  $\{X_{2i}, i \in [n/2]\}$  and  $\{X_{2i-1}, i \in [n/2]\}$  are independent and identically distributed sequences, with

$$X_{2i} = Y_{2i} Y_{2(i+1)}, \quad (5.4)$$

for  $Y$  as in Example 3.2, with  $\phi = 0$  and  $\rho = 0.5$ , and standard Gaussian innovations.

For each situation, 10,000 simulations were performed. Within each simulation, the permutation test was calculated by randomly sampling 2000 permutations.

The results of the simulation are further illustrated in Figures 1 and 2. Figure 1 shows kernel density estimates<sup>2</sup> of the distributions of the test statistic and the permutation distribution in the  $m$ -dependent setting described above. Figure 2 provides QQ plots of the simulated  $p$ -values against the theoretical quantiles of a  $U[0, 1]$  distribution, also in the  $m$ -dependent setting. These figures further confirm the asymptotic validity of the permutation test procedure in the  $m$ -dependent setting.

We observe several computational choices to be made when applying the permutation testing framework in practice. By the results of Lemmas 3.2 and 3.3, for large values of  $n$ , the estimate  $\hat{\gamma}_n^2$  will be strictly positive with high probability. However, for smaller values of  $n$ , it may be the case that a numerically negative value of  $\hat{\gamma}_n^2$  is observed, either when computing the test statistic or the permutation distribution. A trivial solution to this issue is the truncate the estimate at some sufficiently small fixed lower bound  $\epsilon > 0$ . Note that, for appropriately small choices of  $\epsilon$ , that is,  $\epsilon < \gamma_1^2$ , the results of Lemmas 3.2 and 3.3 still hold, that is, inference based on this

<sup>2</sup> These were obtained using the `density` function in R, using the default Gaussian kernel.

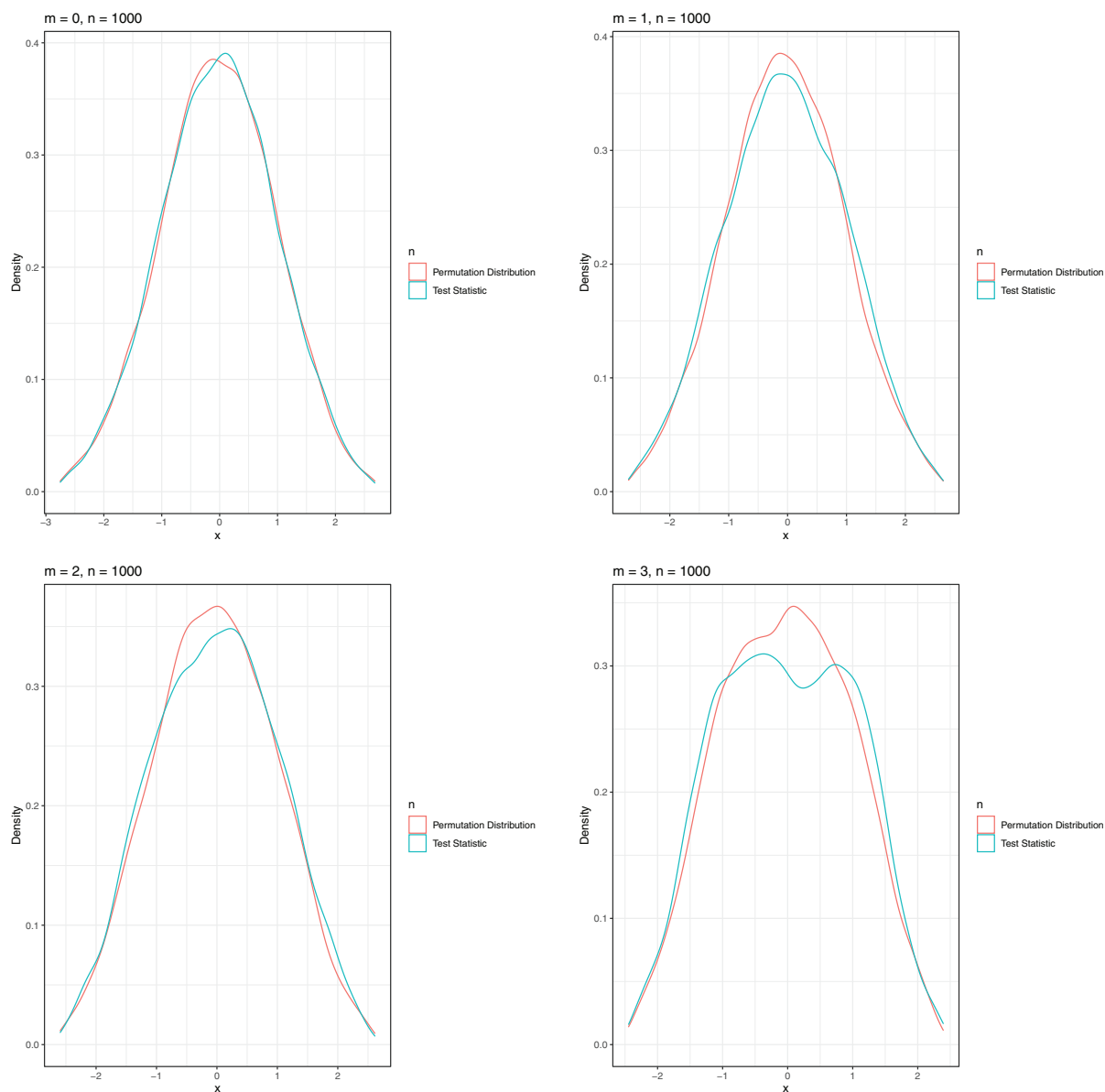


Figure 1. Kernel density estimates of the densities of the test statistic and permutation distribution in the  $m$ -dependent case. In the case of the permutation distribution, the KDE of the permutation distribution, pooled across simulations, is provided. The kernel used for the KDE is a Gaussian kernel

choice of studentization is still asymptotically valid. In practice, however, the suitability of a choice of  $\epsilon$  for a particular numerical application is affected by the distribution of the  $X_i$ . For the above simulation, a constant value of  $\epsilon = 10^{-6}$  was used.

A further choice is that of the truncation sequence  $\{b_n, n \in \mathbb{N}\}$  used in the definition of  $\hat{\gamma}_n^2$ . Any sequence  $\{b_n\}$  such that, as  $n \rightarrow \infty$ ,  $b_n \rightarrow \infty$  and  $b_n = o(\sqrt{n})$  is theoretically justified by Theorem 3.3, although, in a specific setting, some choices of  $\{b_n\}$  will lead to more numerical stability than others. In practice, several choices

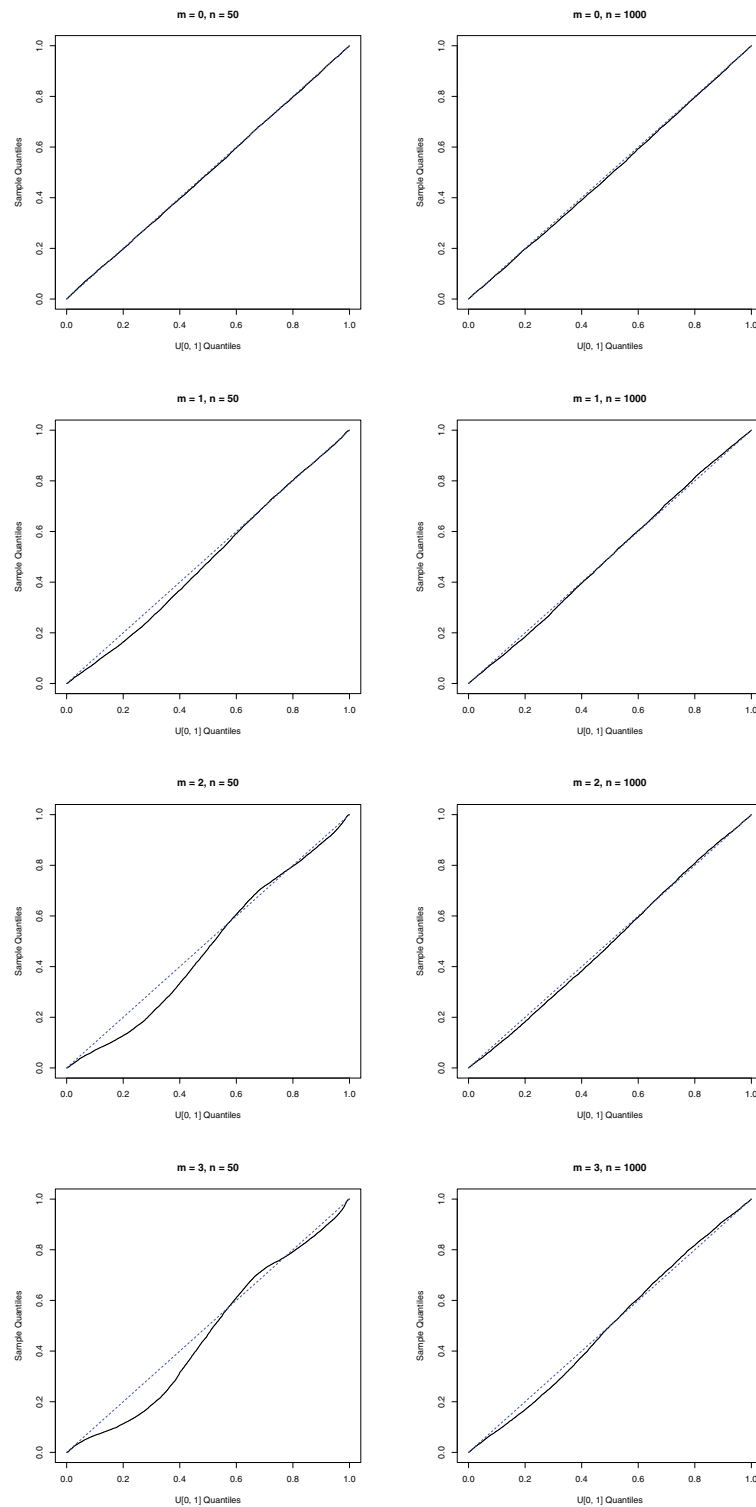


Figure 2. QQ plots of the sample  $p$ -values obtained from one-sided permutation tests, in the  $m$ -dependent Gaussian product setting

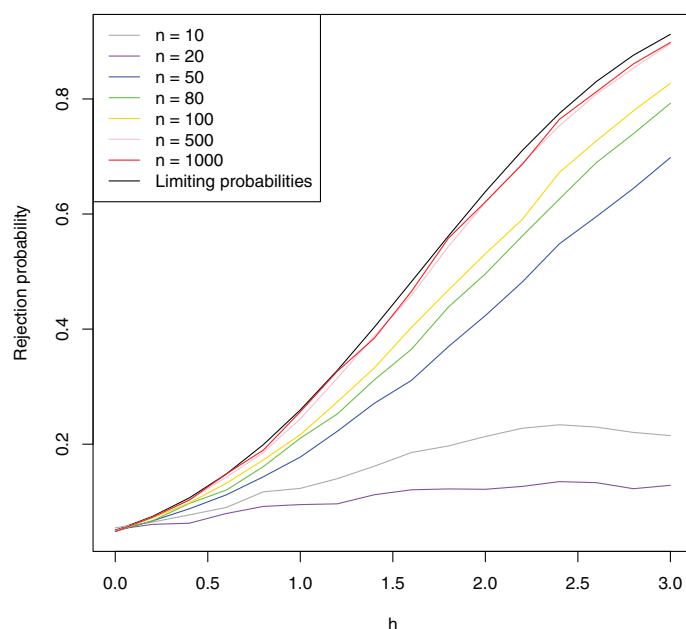


Figure 3. Monte Carlo simulation results for rejection probabilities of the studentized permutation tests of  $\rho(1) = 0$ , in the setting of Example 3.4, for different values of  $h$  and  $n$

Table III. Monte Carlo simulation results for null rejection probabilities for tests of  $\rho(1) = 0$ , using the test of Shao (2011), in an  $m$ -dependent Gaussian product setting

| $m$ | $n$ | 10     | 20     | 50     | 80     | 100    | 500    | 1000   |
|-----|-----|--------|--------|--------|--------|--------|--------|--------|
| 0   |     | 0.2256 | 0.1664 | 0.1142 | 0.0932 | 0.0932 | 0.0714 | 0.0512 |
| 1   |     | 0.1930 | 0.1240 | 0.1000 | 0.0726 | 0.0816 | 0.0678 | 0.0668 |
| 2   |     | 0.1888 | 0.1090 | 0.0658 | 0.0516 | 0.0646 | 0.0516 | 0.0642 |
| 3   |     | 0.1792 | 0.0996 | 0.0514 | 0.0400 | 0.0466 | 0.0386 | 0.0436 |

of  $\{b_n\}$  were implemented, but were found to make little difference to the rejection probabilities observed. In the simulations above,  $\{b_n\}$  was taken to be  $b_n = \lfloor n^{1/3} \rfloor + 1$ , where  $\lfloor x \rfloor$  denotes the integer part of  $x$ .

We also provide Monte Carlo simulation results for the local limiting power of the one-sided studentized permutation test with local alternatives of the form described in Example 3.4. The nominal level considered is  $\alpha = 0.05$ . For each situation, 10,000 simulations were performed. Within each simulation, the permutation test was calculated by randomly sampling 2000 permutations. Figure 3 shows the null rejection probabilities for different values of  $h$ .

We observe that, for large values of  $n$ , the sample rejection probabilities are very close to the theoretical rejection probabilities computed in Example 3.4.

We also provide Monte Carlo simulation results illustrating a comparison between the studentized permutation test and the bootstrap-assisted test of Shao (2011). Table III tabulates the rejection probabilities of this bootstrap-assisted test in the  $m$ -dependent Gaussian products case, presented in Section 5 of the main article. The nominal level considered is  $\alpha = 0.05$ . For each situation, 5000 simulations were performed. Within each simulation, the bootstrap distribution was calculated by using  $M = 500$  bootstrap samples, as in the original article, and the block size used was  $b_n = \lfloor \sqrt{n} \rfloor$ . Note that the sequences in the  $m$ -dependent Gaussian products case satisfy  $\rho(k) = 0$  for all  $k$ , and so a direct comparison of the rejection probabilities of both tests is suitable.

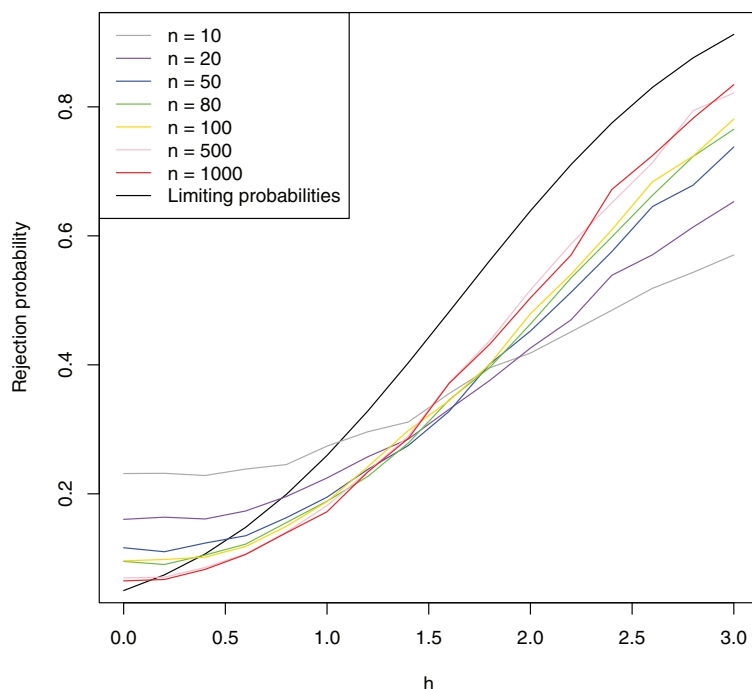


Figure 4. Monte Carlo simulation results for rejection probabilities of the tests  $\rho(1) = 0$  using the bootstrap-assisted test of Shao (2011), in the setting of Example 3.4, for different values of  $h$  and  $n$

We observe that the rejection probabilities obtained in Table III are comparable to those obtained for the permutation test based on the studentized sample autocorrelation only for large values of  $n$ , with the permutation test having simulated rejection probabilities slightly closer to the nominal level  $\alpha = 0.05$ . For smaller values of  $n$ , the bootstrap-assisted test does not reject the null with probability close to the nominal level  $\alpha = 0.05$ . For example, for  $m = 0$  and  $n = 10$ , Table III shows a rejection probability of 0.2256, while Table I in the main article shows a rejection probability of 0.0511 for the permutation test in this case.

We additionally provide Monte Carlo simulation results for the local limiting power of the one-sided bootstrap-assisted test with local alternatives of the form described in Example 3.4. The nominal level considered is  $\alpha = 0.05$ . For each situation, 5000 simulations were performed. Within each simulation, the bootstrap-assisted test was calculated using  $M = 500$  bootstrap samples, and the block size used was  $b_n = \lfloor \sqrt{n} \rfloor$ . Figure 4 shows the null rejection probabilities for different values of  $n$  and  $h$ . We observe that the power of the bootstrap-assisted test against these local alternatives is significantly lower than that of the permutation test; indeed, it appears that the rejection probabilities of the bootstrap-assisted test converge to values significantly below the theoretical maximal limiting rejection probabilities, for moderate to large values of  $n$ . Since, as illustrated in Table III, the Type 1 error is not controlled by the bootstrap-assisted test for smaller values of  $n$ , and this thereby artificially increases the power since the test is rejecting too frequently.

## 6. APPLICATION TO FINANCIAL DATA

We describe an application of the permutation test to financial stock data.

Under the assumption that a certain version of the Efficient Market Hypothesis holds true (see Fama, 1970; Malkiel, 2003 for details), we have that the daily log-returns of a stock  $R_t$ , that is, for  $S_t$  the stock price at time  $t$ ,

$$R_t = \log(S_t) - \log(S_{t-1}),$$

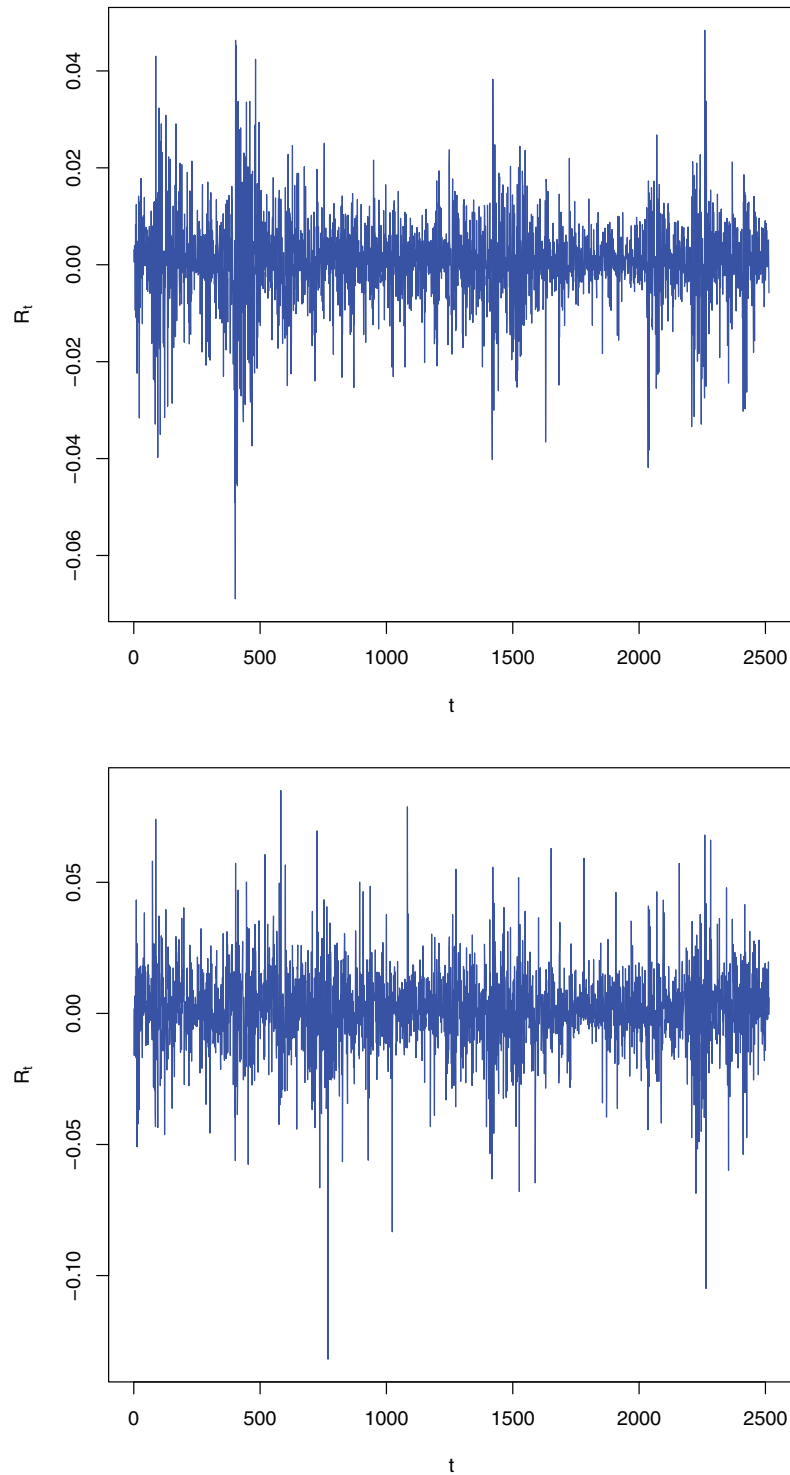


Figure 5. Plots of the daily log-returns  $R_t$  against time. The top plot shows log-returns for the S&P 500 index, and the bottom plot shows log-returns for Apple stock. In both plots,  $t = 0$  corresponds to the first day of trading after 1 January 2010, that is, 1 April 2010

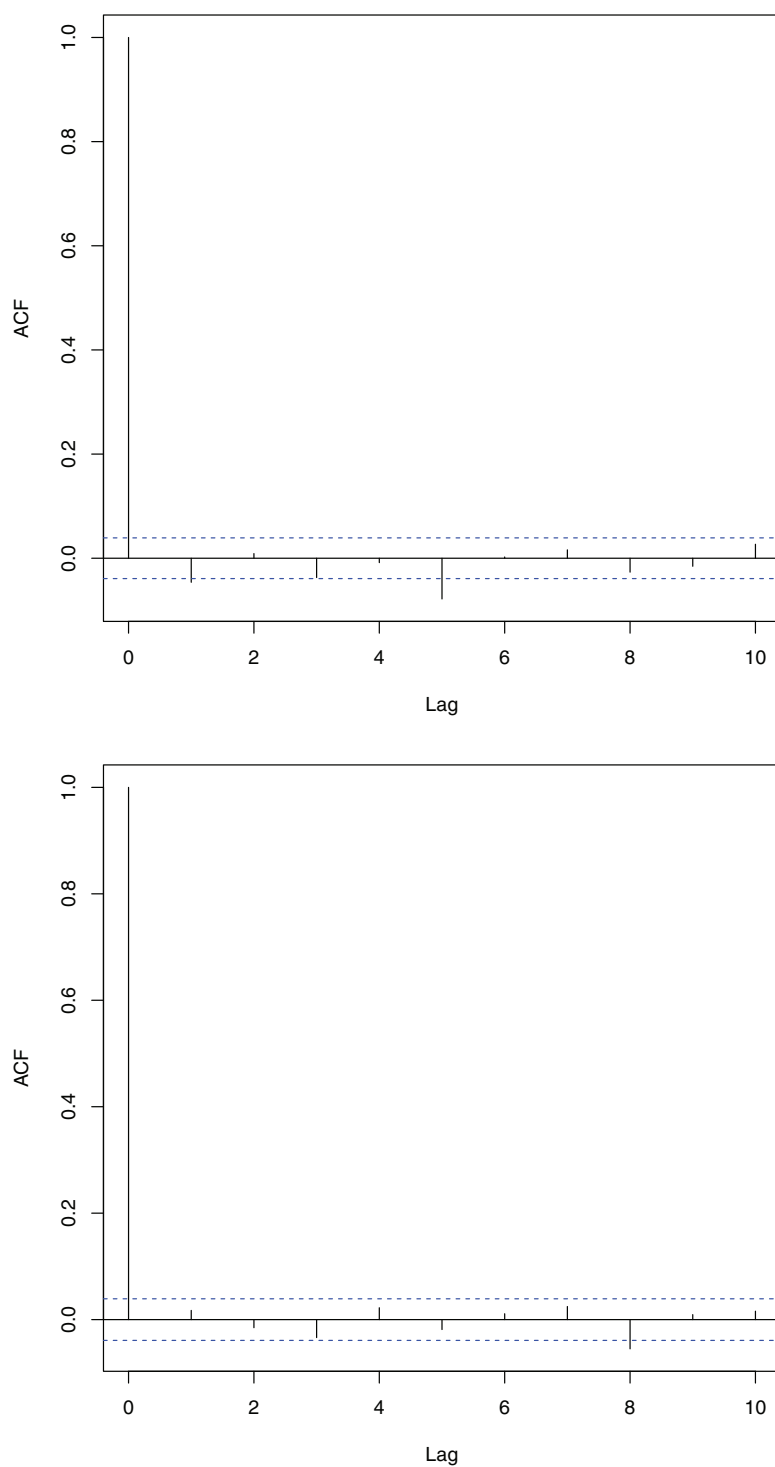


Figure 6. Plots of the sample autocorrelations  $\hat{\rho}_k$  against time. The top plot shows sample autocorrelations for the S&P 500 index, and the bottom plot shows sample autocorrelations for Apple stock. The dotted blue lines show 95% confidence intervals under the assumption that the sequence is a Gaussian white noise process



Table IV. The marginal  $p$ -values obtained using the permutation test for the S&P 500 index and Apple stock data

| $k$  | 1      | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      | 10     |
|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| SPX  | 0.1675 | 0.8665 | 0.4035 | 0.8305 | 0.0340 | 0.9340 | 0.6025 | 0.4100 | 0.4975 | 0.3610 |
| AAPL | 0.3375 | 0.5145 | 0.1255 | 0.3435 | 0.4715 | 0.6310 | 0.3845 | 0.0120 | 0.6470 | 0.4895 |

are serially uncorrelated. Stronger versions of the Efficient Market Hypothesis assert that the daily log-returns either form a martingale sequence, or are independent. It follows that a test of the lack of correlation of observed daily log-returns can provide evidence for, or against, these versions of the Efficient Market Hypothesis.

We illustrate such portmanteau tests performed on daily closing prices for the S&P 500 index (SPX) and Apple (AAPL) stock. In particular, the hypothesis  $H_r$ , for  $r = 10$ , where  $H_r$  is as in (1.1), was tested using the studentized permutation tests described in Section 3, with a Bonferroni correction. The results of these tests were compared to the corresponding results of the Ljung–Box test when applied to the data. The test was performed using closing price data, obtained from Yahoo! Finance, between the dates of 1 January 2010 and 31 December 2019. In both cases, days for which data were unavailable, such as weekends and holidays, were omitted. Plots of the log-returns are shown in Figure 5, and plots of the sample autocorrelations are shown in Figure 6. In both cases, the permutation distribution was approximated using 2000 random permutations. In addition, the summation parameter used in the studentization term  $\hat{\gamma}_n$  used was  $b_n = \lfloor n^{1/3} \rfloor + 1$ . The  $p$ -values obtained are shown in Table IV.

We observe that, marginally, in the case of the S&P 500 index, the  $p$ -value for  $k = 5$  was significant, and, in the case of Apple stock, the marginal  $p$ -value for  $k = 8$  was significant. However, in both cases, none of the  $p$ -values is significant at the  $\alpha = 5\%$  level when adjusted using a Bonferroni correction, and so we may conclude that there is no significant evidence in the data for the daily log-returns to indicate deviation from the Efficient Market Hypothesis.

By contrast, the  $p$ -values obtained using the Ljung–Box test, for all 10 lags simultaneously, were 0.0010 (in the case of the S&P 500 index), and 0.0827, in the case of Apple stock. While the results in the case of Apple stock are consistent with those of the permutation test, in the case of the S&P 500 index, we observe that the Ljung–Box test rejects the null hypothesis at the  $\alpha = 5\%$  level. However, in light of the results of the permutation test and the simulation results in Section 5, this should not cast doubt on our conclusion that no significant deviation from the Efficient Market Hypothesis is observed.

## 7. CONCLUSIONS

When the fundamental assumption of exchangeability does not necessarily hold, permutation tests are invalid unless strict conditions on underlying parameters of the problem are satisfied. For instance, the permutation test of  $\rho(1) = 0$  based on the sample first-order autocorrelation is asymptotically valid only when  $\sigma^2$ , the marginal variance of the distribution, and  $\gamma_1^2$ , where  $\gamma_1^2$  is as defined in (2.5), are equal. Hence rejecting the null must be interpreted correctly, since rejection of the null with this permutation test does not necessarily imply that the true first-order autocorrelation of the sequence is non-zero. We provide a testing procedure that allows one to obtain asymptotic rejection probability  $\alpha$  in a permutation test setting. A significant advantage of this test is that it has the exactness property, absent from the Ljung–Box and Box–Pierce tests, under the assumption of i.i.d., as well as achieving asymptotic level  $\alpha$  in a much wider range of settings than the aforementioned tests. An analogous testing procedure permits for asymptotically valid inference in a test of the  $k$ th order autocorrelation.

Correct implementation of a permutation test is crucial if one is interested in confirmatory inference via hypothesis testing; indeed, proper error control of Type 1, 2 and 3 errors can be obtained for tests of autocorrelations by basing inference on test statistics which are asymptotically pivotal. A framework has been provided for a test of serial lack of correlation in time series data, where tests for  $\rho(j) = 0$  are conducted simultaneously for a large number of values of  $j$ , while maintaining error control with respect to the family-wise error rate.

In this article, a test of lack of serial autocorrelation of data was presented. In the context of other work on testing autocorrelation in time series, such as that of Francq *et al.* (2005), future work will expand on the ideas

presented, to provide an analogous permutation testing procedure to test the lack of serial correlation in the errors of a large class of models.

## ACKNOWLEDGEMENTS

Supported by NSF Grant MMS-1949845 (J.P.R. and M.A.T.). We thank Kevin Guo and Benjamin Seiler for helpful comments and conversations.

## DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed during this study. Historical stock data was obtained from Yahoo! Finance (finance.yahoo.com).

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

## REFERENCES

- Box GEP, Pierce DA. 1970. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association* **65**(332):1509–1526.
- Bradley RC. 2005. Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys* **2**:107–144.
- Chung E, Romano JP. 2013. Exact and asymptotically robust permutation tests. *Annals of Statistics* **41**(2):484–507.
- DiCiccio CJ, Romano JP. 2017. Robust permutation tests for correlation and regression coefficients. *Journal of the American Statistical Association* **112**(519):1211–1220.
- Doukhan P. 1994. *Mixing: Properties and Examples (Lecture Notes in Statistics)* Springer, New York.
- Escanciano JC, Lobato IN. 2009. An automatic portmanteau test for serial correlation. *Journal of Econometrics* **151**(2):140–149.
- Fama EF. 1970. Efficient capital markets: a review of theory and empirical work. *The Journal of Finance* **25**(2):383–417.
- Franco C, Roy R, Zakoian J-M. 2005. Diagnostic checking in ARMA models with uncorrelated errors. *Journal of the American Statistical Association* **100**(470):532–544.
- Holm S. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* **6**(2):65–70.
- Ibragimov IA. 1962. Some limit theorems for stationary processes. *Theory of Probability & Its Applications* **7**(4):349–382.
- Janssen A. 1997. Studentized permutation tests for non-i.i.d. hypotheses and the generalized Behrens–Fisher problem. *Statistics & Probability Letters* **36**(1):9–21.
- Janssen A, Pauls T. 2003. How do bootstrap and permutation tests work? *Annals of Statistics* **31**(3):768–806.
- Jentsch C, Pauly M. 2015. Testing equality of spectral densities using randomization techniques. *Bernoulli* **21**(2):697–739.
- Lehmann EL, Romano JP. 2005. *Testing Statistical Hypotheses (Springer Texts in Statistics)*, 3rd ed. Springer, New York.
- Ljung GM, Box GEP. 1978. On a measure of lack of fit in time series models. *Biometrika* **65**(2):297–303.
- Lobato I, Nankervis JC, Savin N. 2002. Testing for zero autocorrelation in the presence of statistical dependence. *Econometric Theory* **18**(3):730–743.
- Malkiel BG. 2003. The efficient market hypothesis and its critics. *Journal of Economic Perspectives* **17**(1):59–82.
- Mokkadem A. 1988. Mixing properties of ARMA processes. *Stochastic Processes and their Applications* **29**(2):309–315.
- Neuhäus G. 1993. Conditional rank tests for the two-sample problem under random censorship. *Annals of Statistics* **21**(4):1760–1779.
- Nichols TE, Holmes AP. 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human Brain Mapping* **15**(1):1–25.
- Noether GE. 1950. Asymptotic properties of the Wald–Wolfowitz test of randomness. *Annals of Mathematical Statistics* **21**(2):231–246.
- Omelka M, Pauly M. 2012. Testing equality of correlation coefficients in two populations via permutation methods. *Journal of Statistical Planning and Inference* **142**:1396–1406.
- Pitsyn A, Zvonik S, Gimble J. 2006. Permutation test for periodicity in short time series data. *BMC Bioinformatics* **7**(Suppl. 2):S10.

- Ritzwoller DM, Romano JP. 2020. *Uncertainty in the hot hand fallacy: detecting streaky alternatives to random Bernoulli sequences*. Technical report 2020-02. Department of Statistics, Stanford University.
- Shao X. 2011. A bootstrap-assisted spectral test of white noise under unknown dependence. *Journal of Econometrics* **162**:213–224.
- van der Vaart AW. 1998. *Asymptotic statistics*. In *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, Cambridge.
- Wald A, Wolfowitz J. 1943. An exact test for randomness in the non-parametric case based on serial correlation. *Annals of Mathematical Statistics* **14**(4):378–388.
- Zhu K. 2016. Bootstrapping the portmanteau tests in weak auto-regressive moving average models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **78**(2):463–485.