A NONPARAMETRIC TEST FOR A CONSTANT CORRELATION MATRIX

DOMINIK WIED

TU Dortmund

October 29, 2014

Abstract

We propose a nonparametric procedure to test for changes in correlation matrices at an unknown point in time. The new test requires constant expectations and variances, but only mild assumptions on the serial dependence structure and has considerable power in finite samples. We derive the asymptotic distribution under the null hypothesis of no change as well as local power results and apply the test to stock returns.

Keywords: Fluctuation Test, Functional Delta Method, Gaussian Process, Local Power.

JEL codes: C12, C14, C32, C58

*TU Dortmund, Fakultät Statistik, 44221 Dortmund, Germany. Email: wied@statistik.tu-dortmund.de, Phone: +49 231 755 5419. Financial support by Deutsche Forschungsgemeinschaft (SFB 823, project A1) and useful comments by the editor Esfandiar Maasoumi, an associate editor, a referee, Matthias Arnold, Axel Bücher, Walter Krämer and Martin Wendler are gratefully acknowledged.
1. Introduction

The Bravais-Pearson correlation coefficient is arguably the most widely used measure of dependence between random variables. For financial time series, correlations among returns are for instance widely used in risk management. However, there is compelling empirical evidence that the correlation structure of financial returns cannot be assumed to be constant over time, see e.g. Krishan et al. (2009). In particular, in periods of financial crisis, correlations often increase, a phenomenon which is sometimes referred to as “Diversification Meltdown”. As most often potential change points are not known a priori, practitioners are interested in testing correlation constancy in financial time series at an unknown point in time.

Wied et al. (2012) propose a nonparametric retrospective kernel-based correlation constancy test (referred to as KB-test in what follows) and Wied and Galeano (2012) propose a sequential monitoring procedure. These papers complement other approaches for related measures of dependence, e.g. for the whole covariance matrix (Aue et al., 2009, Galeano and Peña, 2007), the copula (Na et al., 2012, Krämer and van Kampen, 2011), Spearman’s rho (Gaißler and Schmid, 2010), Kendall’s tau (Dehling et al., 2014), autocovariances in a linear process (Lee et al., 2003) and covariance operators in the context of functional data analysis (Fremdt et al., 2012).

In what follows, we stick to correlation. Wied et al. (2012) show that a correlation test can be more powerful than a covariance test when we have more than one change point in the covariance structure.

However, the KB-test only considers bivariate correlations, whereas in portfolio management, where we typically have more than two assets, constancy of the whole correlation matrix is of interest. In this context, it would be possible to perform several pairwise tests and to use a level correction like Bonferroni-Holm. However, in this paper, we consider the correlation matrix. In a simulation study, we see that the matrix-based test outperforms the Bonferroni-Holm approach in some (although not in all) situations. We extend the methodology from the KB-test to higher dimensions, but on the other hand
keep its nonparametric and model-free approach.

We consider the \( \frac{p(p-1)}{2} \)-vector of successively calculated pairwise correlation coefficients and derive its limiting distribution with the functional delta method approach and some proof ideas from Wied et al. (2012). We use a bootstrap approximation for a normalizing constant in order to approximate the asymptotic limit distribution of the test statistic. This may be an alternative for the bivariate case as well.

In an application of this test to Value-at-Risk forecasts (Berens et al., 2013) it is seen that this proposed test might indeed be useful in practical situations. That is, it might be a promising approach to combine the well-known CCC (constant conditional correlation) and DCC (dynamic conditional correlation) model with this test for structural breaks in correlations.

The paper is organized as follows: In Section 2, we present the test statistic and derive the asymptotic null distribution, Section 3 deals with local power, Section 4 presents simulation evidence, Section 5 provides an empirical illustration and Section 6 a conclusion. All proofs are deferred to the appendix.

2. The Fluctuation Test

Let \( X_t = (X_{1,t}, X_{2,t}, \ldots, X_{p,t}) \), \( t \in \mathbb{Z} \), be a sequence of \( p \)-variate random vectors on a probability space \( (\Omega, \mathcal{A}, \mathbb{P}) \) with finite 4-th moments and (unconditional) correlation matrix \( R_t = (\rho_t^{ij})_{1 \leq i,j \leq p} \), where

\[
\rho_t^{ij} = \frac{\text{Cov}(X_{i,t}, X_{j,t})}{\sqrt{\text{Var}(X_{i,t})\text{Var}(X_{j,t})}}.
\]

Furthermore, we call \( || \cdot ||_r \) the \( L_r \)-norm, \( r > 0 \), and \( D(I, \mathbb{R}^d), d \in \mathbb{N} \), the space of \( d \)-dimensional càdlàg functions on an interval \( I \subseteq [0,1] \) (compare Billingsley, 1968 and related literature for details). We write \( A \sim (m,n) \) for a matrix \( A \) with \( m \) rows and \( n \) columns. Throughout the paper, we denote by \( \rightarrow_d \) and \( \rightarrow_p \) convergence in distribution and probability, respectively, of random variables or vectors. By \( \Rightarrow_d \), we denote conver-
gence in distribution of stochastic processes on a function space, which will be specified depending upon the situation, and with respect to the corresponding supremum norm.

For $T \in \mathbb{N}$, the hypothesis pair is given by $H_0 : R_1 = \ldots = R_T$ vs. $H_1 : \neg H_0$. Under $H_0$, we denote $\rho_{ij}^T =: \rho_{ij}$.

The “preliminary version” of the test statistic is given by

$$Q_T := \max_{2 \leq k \leq T} \sum_{1 \leq i < j \leq p} \frac{k}{\sqrt{T}} |\hat{\rho}_{ij}^k - \hat{\rho}_{ij}^T| =: \max_{2 \leq k \leq T} \frac{k}{\sqrt{T}} \|P_{k,T}\|_1,$$

where

$$\hat{\rho}_{ij}^k = \frac{\sum_{t=1}^k (X_{i,t} - \bar{X}_{i,k})(X_{j,t} - \bar{X}_{j,k})}{\sqrt{\sum_{t=1}^k (X_{i,t} - \bar{X}_{i,k})^2 \sqrt{\sum_{t=1}^k (X_{j,t} - \bar{X}_{j,k})^2}},$$

$\bar{X}_{i,k} = \frac{1}{k} \sum_{t=1}^k X_{i,t}$, $\bar{X}_{j,k} = \frac{1}{k} \sum_{t=1}^k X_{j,t}$ and $P_{k,T} = (\hat{\rho}_{ij}^k - \hat{\rho}_{ij}^T)_{1 \leq i < j \leq p} \in \mathbb{R}^{p(p-1)/2}$.\(^1\) The value $\hat{\rho}_{ij}^k$ is the empirical pairwise correlation coefficient for the random variables $X_i$ and $X_j$, calculated from the first $k$ observations. Thus, the test statistic compares the pairwise successively calculated correlation coefficients with the corresponding correlation coefficients calculated from the whole sample. The null hypothesis is rejected whenever $Q_T$ becomes too large, i.e., whenever at least one of these differences become too large over time or, equivalently, whenever the successively calculated correlation coefficients of at least one pair fluctuate too much over time. The weighting factor $\frac{k}{\sqrt{T}}$ serves for compensating the fact that the correlations are typically estimated better in the middle or in the end of sample compared to the beginning of the sample. We will see later on in the context of discussing the bootstrap approximation that it might be more convenient to use a slightly modified version of $Q_T$.

For deriving the limiting null distribution and local power results, some additional assumptions are necessary. The following assumptions concern moments and serial dependencies of the random variables and correspond to (A1), (A2) and (A3) in Wied et al.\(^2\)

\(^1\)Here and analogously in the following, the expression $1 \leq i < j \leq p$ for a vector means that the first entry or entries consist of the expressions for $i = 1$, followed by the one(s) for $i = 2$ and so on.
Assumption 1. For

\[ U_t := \begin{pmatrix}
X_{1,t}^2 - \mathbb{E}(X_{1,t}^2) \\
\vdots \\
X_{p,t}^2 - \mathbb{E}(X_{p,t}^2) \\
X_{1,t} - \mathbb{E}(X_{1,t}) \\
\vdots \\
X_{p,t} - \mathbb{E}(X_{p,t}) \\
X_{1,t}X_{2,t} - \mathbb{E}(X_{1,t}X_{2,t}) \\
X_{1,t}X_{3,t} - \mathbb{E}(X_{1,t}X_{3,t}) \\
\vdots \\
X_{p-1,t}X_{p,t} - \mathbb{E}(X_{p-1,t}X_{p,t})
\end{pmatrix} \]

and \( S_j := \sum_{t=1}^{j} U_t \), we have

\[
\lim_{m \to \infty} \mathbb{E} \left( \frac{1}{m} S_m S_m' \right) =: D_1,
\]

where \( D_1 \) is a finite and positive definite matrix with \( 2p + \frac{p(p-1)}{2} \) rows and \( 2p + \frac{p(p-1)}{2} \) columns.

Assumption 2. For some \( r > 2 \), the \( r \)-th absolute moments of the components of \( U_t \) are uniformly bounded, that means, \( \sup_{t \in \mathbb{Z}} \mathbb{E} \| U_t \|_r < \infty \).

Assumption 3. For \( r \) from Assumption 2, the vector \( (X_{1,t}, \ldots, X_{p,t}) \) is \( L_2 \)-NED (near-epoch dependent) with size \( -\frac{r-1}{r-2} \) and constants \( (c_t), t \in \mathbb{Z} \), on a sequence \( (V_t), t \in \mathbb{Z} \), which is \( \alpha \)-mixing of size \( \phi^* := -\frac{r-1}{r-2} \), i.e.,

\[
\| (X_{1,t}, \ldots, X_{p,t}) - \mathbb{E} ((X_{1,t}, \ldots, X_{p,t})|\sigma(V_{t-l}, \ldots, V_{t+i})) \|_2 \leq c_l v_l
\]

with \( \lim_{t \to \infty} v_l = 0 \). The constants \( (c_t), t \in \mathbb{Z} \) fulfill \( c_l \leq 2 \| U_t \|_2 \) with \( U_t \) from Assumption
Assumption 1 is a regularity condition which rules out trending random variables. As we have financial returns in mind, this is no issue.

Assumption 2 is more critical because it requires finite $|4 + \gamma|$-th moments of $X_t$ with an arbitrary $\gamma > 0$ (note that the components of $X_t$ enter $U_t$ quadratically). In fact, there is evidence that even variances might not exist for some financial series, cf. Mandelbrot (1963). However, simulation evidence below shows that the test still works under the $t_3$-distribution.

Assumption 3 is a very general serial dependence assumption which holds in relevant econometric models, e.g. in GARCH-models under certain conditions (cf. Carrasco and Chen, 2002). It guarantees that the vector

$$(X^2_{1,t}, \ldots, X^2_{p,t}, X_{1,t}, \ldots, X_{p,t}, X_{1,t}X_{2,t}, X_{1,t}X_{3,t}, \ldots, X_{p-1,t}X_{p,t})$$

is $L_2$-NED (near-epoch dependent) with size $-\frac{1}{2}$, cf. Davidson (1994), p. 273. This allows for applying a functional central limit theorem later on.

Next, we impose a stationarity condition which is in line with Aue et al. (2009).

**Assumption 4.** $(X_{1,t}, \ldots, X_{p,t}), t \in \mathbb{Z}$, has constant expectation and variances, that means, $E(X_{i,t}), i = 1, \ldots, p$, and $0 < E(X^2_{i,t}), 1 \leq i \leq p$, do not depend on $t$.

This condition might be slightly relaxed to allow for some fluctuations in the first and second moments (see (A4) and (A5) in Wied et al., 2012), but for ease of exposition and because the procedure would remain exactly the same, we stick to this assumption. Note that most financial time series processes as for example GARCH are (unconditionally) stationary under certain conditions. Clearly, the original test problem is invariant under heteroscedasticity. But we believe that it is at least extremely difficult if not impossible to design a fluctuation test for correlations in which arbitrary variance changes are allowed under the null hypothesis.

Our main result is:
Theorem 1. Under $H_0$ and Assumptions 1,2,3,4, for $T \to \infty$,

$$
\frac{\tau(s)}{\sqrt{T}}(\hat{\rho}^{ij}_{T(s)} - \rho^{ij}_{T})_{1 \leq i < j \leq p} \Rightarrow_d E^{1/2} B^{\frac{p(p-1)}{2}}(s),
$$
on $D\left([0, 1], \mathbb{R}^{\frac{p(p-1)}{2}}\right)$, where $\tau(s) = \lfloor 2 + s(T - 2) \rfloor$.

$$
E = \lim_{T \to \infty} \text{Cov}\left(\sqrt{T}(\hat{\rho}^{ij}_{T})_{1 \leq i < j \leq p}\right) \sim \left(\frac{p(p-1)}{2} \times \frac{p(p-1)}{2}\right)
$$

and $B^{\frac{p(p-1)}{2}}(s)$ is a vector of $\frac{p(p-1)}{2}$ independent standard Brownian Bridges.

The proof of the theorem can be found in the appendix. It relies on the application of an adapted functional delta method. We want to stress that simply applying a functional central limit theorem is not enough here due to the cumbersome, non-linear structure of the correlation coefficient.

From the previous theorem, we directly obtain with the Continuous Mapping Theorem

Corollary 1. Under $H_0$ and Assumptions 1,2,3,4, for $T \to \infty$,

$$
Q_T \to_d \sup_{0 \leq s \leq 1} \left\| E^{1/2} B^{\frac{p(p-1)}{2}}(s) \right\|_1.
$$

In order to obtain critical values, we need information about $E$. There are several possibilities for estimating $E$; one possibility is the estimator $\hat{E}$, given by a bootstrap approximation. For this estimation, one can for example use the moving block bootstrap from Künsch (1989) and Liu and Singh (1992), cf. also Lahiri (1999), Conçalves and White (2002), Conçalves and White (2003), Calhoun (2013), Radulović (2012) and Sharipov and Wendler (2012).

Defining a block length $l_T$, we divide the time series into $T - l_T - 1$ overlapping blocks $B_i, i = 1, \ldots, T - l_T - 1$, with length $l_T$ such that $B_1 = (X_1, \ldots, X_{l_T}), B_2 = (X_2, \ldots, X_{l_T+1}), \ldots$.

Then, in each bootstrap repetition $b, b = 1, \ldots, B$ for some large $B$, we sample $\left\lfloor \frac{T}{l_T} \right\rfloor$ times with replacement one of the $T - l_T - 1$ blocks and concatenate the blocks. So, we obtain
\( B \) \( p \)-dimensional time series with length \( \left\lfloor \frac{T}{l_T} \right\rfloor \cdot l_T \). For each bootstrapped time series we calculate the vector \( v_b := \sqrt{T} (\hat{\rho}_{ij}^{b,T})_{1 \leq i < j \leq p} \). The estimator \( \hat{E} \) is then the empirical covariance matrix of these \( B \) vectors, i.e.,

\[
\hat{E} = \frac{1}{B} \sum_{b=1}^{B} (v_b - \bar{v})(v_b - \bar{v})'
\]

with \( \bar{v} = \frac{1}{B} \sum_{b=1}^{B} v_b \). The bootstrap estimator “replaces” the rather complicated kernel estimator \( \tilde{E} \) from the KB-test (Appendix A.1 in Wied et al., 2012). The advantage of the bootstrap estimator is the fact that it can be derived easily even in higher dimensions. It would be possible to obtain a kernel estimator also in higher (> 2) dimensions. However, its structure would then depend on the structure of derivatives of certain non-linear, higher-dimensional functions which transform a high-dimensional vector of moments to the vector of correlation coefficients. (More information is given in the proof of Theorem 1). The arguably complicated transformation makes the calculation of a kernel estimator very cumbersome and much harder to implement. Moreover, a kernel estimator depends on the choice of the bandwidth and the kernel. The disadvantage of the bootstrap is that it is computationally more intensive. In addition, the choice of the block length is required.

The matrix \( \hat{E} \) is an estimator for \( \text{Cov}^*(v_b) \) which is the (theoretical) covariance matrix of \( v_b \) with respect to the bootstrap sample conditionally on the original data \( X_1, \ldots, X_T \).

In order to validate the bootstrap, the key point is the proof that, for \( T \to \infty \), \( \text{Cov}^*(v_b) \) converges in probability to \( E \). In order to obtain such an asymptotic result, we need an assumption on the block length.

**Assumption 5.** For \( T \to \infty \), \( l_T \to \infty \) and \( l_T \sim T^\alpha \) for \( \alpha \in (0, 1) \).

The assumption is similar to the one for the moving block bootstrap in Theorem 1 (Condition 4) of Calhoun (2013). It guarantees that the block length becomes large but not too large compared to \( T \).

Moreover, we need an assumption which ensures that the bootstrap correlation coefficients
are sufficiently close to the correlation coefficients obtained from the data.

**Assumption 6.** For \(1 \leq i < j \leq p\), some \(\delta > 0\) and \(b = 1, \ldots, B\), the random variable

\[
C_T := \mathbb{E} \left( \left| \sqrt{T} \left( \hat{\rho}_{b,T}^{ij} - \hat{\rho}_{T}^{ij} \right) \right|^{2+\delta} \right|_{X_1, \ldots, X_T}
\]

is stochastically bounded \((C_T = O_p(1))\).

The next theorem gives the theoretical validation for the bootstrap.

**Theorem 2.** Under \(H_0\) and Assumptions 1, 2, 3, 4, 5, 6, for \(T \to \infty\),

\[
\text{Cov}^* \left( \sqrt{T} \left( \hat{\rho}_{b,T}^{ij} \right)_{1 \leq i < j \leq p} \right) \to_p \mathcal{E}.
\]

Given the theoretical results, it is reasonable to consider the “test statistic”

\[
A_T := \max_{2 \leq k \leq T} \frac{k}{\sqrt{T}} \left\| \hat{E}^{-1/2}P_{k,T} \right\|_1
\]

in applications. Then, the null hypothesis is rejected whenever \(A_T\) is larger than the \((1 - \alpha)\)-quantile of the random variable \(A := \sup_{0 \leq s \leq 1} \left\| B^{\frac{p(p-1)}{2}}(s) \right\|_1\). The quantiles of \(A\), which serve as an approximation for the quantiles of the finite sample distribution, can easily be obtained by Monte Carlo simulations, i.e., by approximating the paths of the Brownian Bridge on fine grids.

There might be situations in practice in which \(\hat{E}^{1/2}\) is not positive definite so that \(\hat{E}^{-1/2}\) would not be defined. However, due to Assumption 1, at least for larger \(T\) and \(B\), we can virtually assume positive definiteness.\(^2\)

\(^2\)To circumvent the problem of impossible or numerically unstable inversion of \(\hat{E}^{1/2}\), one could calculate the statistic \(Q_T\) and simulate critical values from the limit random variable in Corollary 1 in which \(E\) is replaced by \(\hat{E}\).
3. Local Power

Econometricians are often not only interested in the behavior of a test under the null hypothesis, but would like to get information about the behavior under some local alternatives. For simplicity, we consider a setting in which the expectations and variances remain constant such that a covariance change is equal to a change in correlations. To be more precise, under $H_1$, in at least one of the components of $X_t$, there is a correlation change of order $\frac{M}{\sqrt{T}}$ ($M > 0$ arbitrary) with constant expectations and variances and

$$(E(X_{i,t}X_{j,t}))_{1 \leq i < j \leq p} = v + \frac{M}{\sqrt{T}}g\left(\frac{t}{T}\right).$$

Here, $v \in \mathbb{R}^{(p-1)^2}$ is a constant vector and $g(s) = (g_1(s), \ldots, g_{\binom{p-1}{2}}(s))$ is a bounded $(p-1)^2$-dimensional function that is not constant and that can be approximated by step functions such that the function

$$\int_0^s g(u)du - s \int_0^1 g(u)du$$

is different from $0 \in \mathbb{R}^{\frac{p(p-1)}{2}}$ for at least one $s \in [0, 1]$. The integral is defined component by component.

Note that we now deal with triangular arrays because the distribution of the $X_t$ changes with $T$, but, for simplicity, we do not change our notation.

A typical example for the function $g$ would be a step function with a jump from 0 to $g_0$ in a given point $z_0$ in one of the components. This implies that the correlation of one pair jumps at time $[T \cdot z_0]$. A step function with several jumps would correspond to multiple change points. With a continuous function $g$, one would obtain continuously changing correlations.

The following Theorem 3 is an analogue to Theorem 1 and yields the distribution under the sequence of local alternatives.

Theorem 3. Under the sequence of local alternatives and Assumptions 1,2,3,4, for $T \to$
\[
\frac{\tau(s)}{\sqrt{T}}(\hat{\rho}^ij - \hat{\rho}^ij)_{1 \leq i < j \leq p} \Rightarrow_d E^{1/2}B^{p(p-1)/2}(s) + E^{1/2}C(s),
\]
on \(D([0, 1], \mathbb{R}^{p(p-1)/2})\), where

\[
C(s) = M \left( \begin{array}{c} 
\frac{1}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} \left( \int_0^s g_1(u)du - s \int_0^1 g_1(u)du \right) \\
\frac{1}{\sqrt{\text{Var}(X_1)\text{Var}(X_3)}} \left( \int_0^s g_2(u)du - s \int_0^1 g_2(u)du \right) \\
\vdots \\
\frac{1}{\sqrt{\text{Var}(X_{p-1})\text{Var}(X_p)}} \left( \int_0^s g_{p(p-1)/2}(u)du - s \int_0^1 g_{p(p-1)/2}(u)du \right) 
\end{array} \right)
\]
is a deterministic function that depends on the specific form of the local alternative under consideration, characterized by \(g\).

In Theorem 3, the supremum is taken over the absolute value of a Brownian Bridge plus a deterministic function \(C(s)\). As the main characteristic of the function \(C(s)\) from Theorem 3, we have the factor \(M\) times the expression \(\int_0^s g_i(u)du - s \int_0^1 g_i(u)du\) in each component \(i = 1, \ldots, \frac{p(p-1)}{2}\). This follows from the structure of a Brownian Bridge.

The previous theorem directly yields with the Continuous Mapping Theorem

**Corollary 2.** Under the sequence of local alternatives and Assumptions 1,2,3,4, for \(T \to \infty\),

\[
Q_T \Rightarrow \sup_{0 \leq s \leq 1} \left\| E^{1/2}B^{p(p-1)/2}(s) + E^{1/2}C(s) \right\|_1
\]

Also under local alternatives we want to estimate \(E\) with the bootstrap. It turns out that the estimator presented in Section 2 has the same limit distribution as under the null hypothesis. Thus, the bootstrap approach is valid both under the null and under the alternative.

**Theorem 4.** Under the sequence of local alternatives and Assumptions 1,2,3,4,5,6, for \(T \to \infty\),

\[
\text{Cov}^*(\sqrt{T}(\hat{\rho}^ij)_{1 \leq i < j \leq p}) \to_p E.
\]
The theoretical results in this section imply that the quantity $A_T$ is close to $A_L := \sup_{0 \leq s \leq 1} \left\| B^{s(p-1)}(s) + C(s) \right\|_1$ for large $T$ and $B$. Moreover, for every $B \geq 1$, the test statistic becomes arbitrarily large for large $M$ and $T$.

4. Finite Sample Evidence

We illustrate the finite sample properties of our multivariate test with Monte Carlo simulations in different settings: We consider a series of four-variate random vectors which are, on the one hand, serially independent and, on the other hand, fulfill a four-variate MA(1)-structure with MA-parameters 0.5. This means that, for $t = 0, \ldots, T$, there are serially independent vectors $u_t := (u_t,1, u_t,2, u_t,3, u_t,4)$ such that the data generating process is defined by

$$X_t = u_t + A u_{t-1}, t = 1, \ldots, T$$

with $X_t = (X_t,1, X_t,2, X_t,3, X_t,4)$, $A = \text{diag}(\theta, \theta, \theta, \theta)$ and $\theta \in \{0, 0.5\}$. The lengths of the series are chosen as $T \in \{200, 500\}$, the block lengths are $l_T = \lfloor T^{1/4} \rfloor$, respectively$^3$, the number of bootstrap replications is 999 and the number of Monte Carlo replications is 10000. We consider, on the one hand, a four-variate normal distribution (ND) and, on the other hand, a four-variate $t_3$-distribution. The $t_3$-distribution is not covered by our assumptions, but we analyze it to get a picture of the behavior of the test in settings which are realistic in financial applications.

For simulating the behavior under the null, we set the variances of the $u_{i,t}, i = 1, 2, 3, 4,$ to 1 and the correlations of the $u_{i,t}$ to $\rho_{12} = \ldots = \rho_{34} =: \rho_0 \in \{0, 0.5\}$. Under the alternative, the $u_{i,t}$ have correlation $\rho_0$ in the first half of the sample. Moreover, we have a change in all six pairwise correlations of the $u_{i,t}$ with shifts $\Delta \rho = -0.2, -0.4, 0.2, 0.4$ in the middle of the sample.

The results (empirical rejection probabilities, not-size-adjusted, nominal level 0.05 which corresponds to a simulated critical value of 4.47) are given in Table 1.

---

$^3$For, $T = 200$, $\left\lfloor \frac{T}{l_T} \right\rfloor \neq \frac{T}{l_T}$, so that the length of the bootstrapped time series is not exactly equal to $T$. However, we consider the difference as negligible.
It is seen that there are some size distortions for the heavy-tailed distribution and/or serial dependence although the level seems to converge to 0.05 for higher $T$ in all cases. The power of the test increases in $T$ and in absolute values of the correlation changes. For the $t_3$-distribution, the power is in general considerably lower. Further simulations show that the size properties become worse for even higher $\rho_0$ and even higher serial dependence.

Moreover, we compare the test for constant correlation matrix with a multivariate procedure based on the pairwise correlation test from Wied et al. (2012) (with bandwidth $\lfloor \log(T) \rfloor$) and the Bonferroni-Holm correction. That means that we perform $m = 6$ pairwise tests and denote by $p_{(1)}, \ldots, p_{(m)}$ the corresponding p-values in increasing order. We declare the null hypothesis of constant correlation matrix to be invalid if there is at least one $j = 1, \ldots, m$ such that

$$p_{(j)} < \frac{0.05}{m + 1 - j}.$$ 

The results are also presented in Table 1. Depending on the situation, sometimes the one and sometimes the other procedure performs better. While the Bonferroni-Holm procedure has in general slightly better size and power properties for $\rho_0 = 0.5$, the matrix-based test performs better with $\rho_0 = 0$, especially with the normal distribution and decreasing correlation. There is even one case in which the Bonferroni-Holm procedure is not unbiased (the power is smaller than the size) which does not occur with the matrix-based test.

In another setting, we have compared the bivariate bootstrap with the bivariate kernel-based and have seen that both tests behave more or less similarly.

5. APPLICATION TO STOCK RETURNS

Next, we show how the proposed test can be applied in financial time series. For this, we consider the correlation of four stocks. In order to avoid issues due to market trading in
different time zones, we just consider the European market. We look at the four companies of Euro Stoxx 50 with the highest weights in the index in the end of May 2012, that means Total, Sanofi, Siemens and BASF, and consider the time span 01.01.2007 - 01.06.2012 such that $T = 1414$. The data was obtained from the database Datastream. Figures 1, 2, 3 plot rolling windows of the six pairwise correlations of the continuous daily returns from each asset with the window length 120. This corresponds to the trading time of about half a year. The days on the x-axis show the first day of the windows, respectively.

We identify time-varying correlations. It is for example interesting to see that the correlation between Total and Sanofi is close to 0 in the beginning of February 2008 and much higher after this. The correlation between Sanofi and BASF is interestingly low in the middle of 2009.

The test statistic $Q_T$ applied on the four-variate return vector is equal to 10.49. With $B = 10999$ bootstrap replications, we obtain $A_T = 6.55$. With this value, we cannot yet determine if the null hypothesis of constant correlation is rejected. So we calculate the statistic $A_T$ with $B = 10999$. The 0.95-quantile of $\sup_{0 \leq s \leq 1} ||B^6(t)||_1$ is equal to 4.47 and so, the null hypothesis is rejected on the significance level $\alpha = 0.05$. The approximate p-value is smaller than 0.001.

Figure 4 shows the process

\[
\left( \sum_{1 \leq i < j \leq p} \frac{k}{\sqrt{T}} |\hat{\rho}_{ij}^k - \hat{\rho}_{ij}^T| \right)_{2 \leq k \leq T},
\]

that means the evolution of the successively calculated correlations over time. In the context of CUSUM tests, the point of the maximum is often considered as a reasonable estimator for the (most important) change point if the test decides that such a point actually exists, see Vostrikova (1981) and the related literature. In this case, the maximum
is obtained at the 11th of September 2008 which corresponds quite well to the insolvency of Lehman Brothers. A discussion on dating multiple change points in the correlation matrix can be found in Galeano and Wied (2014).

- Figure 4 here -

6. Conclusion

We have presented a new fluctuation test for constant correlations in the multivariate setting for which the location of potential change points need not be specified a priori. The new test is based on a bootstrap approximation, works under mild assumptions regarding the dependence structure, has appealing properties in simulations and seems to be useful in empirical applications. Potential drawbacks of the test are the requirement of finite fourth moments and the assumption of constant expectations and variances. It might be an interesting question for the future to thoroughly investigate to which extent these drawbacks could be overcome by some kind of prefiltering and/or other transformations. Moreover, it could be worthwhile to extend the present approach to the problem of monitoring correlation changes or to other, perhaps more robust measures of dependence.
Aue, A., S. H"ormann, L. Horvath, and M. Reimherr (2009): “Break Detection in the Covariance Structure of Multivariate Time Series Models,” Annals of Statistics, 37(6B), 4046–4087.

Berens, T., G. Weiss, and D. Wied (2013): “Testing for Structural Breaks in Correlations: Does it Improve Value-at-Risk Forecasting?” SSRN Working Paper (http://ssrn.com/abstract=2265488).

Billingsley, P. (1968): Convergence of Probability Measures, Wiley, New York.

Calhoun, G. (2013): “Block Bootstrap Consistency Under Weak Assumptions,” Iowa State Working Paper, version 25.03.2013, http://www.econ.iastate.edu/sites/default/files/publications/papers/p14313-2011-09-23.pdf.

Carrasco, M. and X. Chen (2002): “Mixing and Moment Properties of Various GARCH and Stochastic Volatility Models,” Econometric Theory, 18(1), 17–39.

Cheng, G. (2011): “A Note on Bootstrap Moment Consistency for Semiparametric M-Estimations,” Working Paper, arXiv: 1109.4204v1.

Concalves, S. and H. White (2002): “The Bootstrap of the Mean for Dependent Heterogeneous Arrays,” Econometric Theory, 18(6), 1367–1384.

——— (2003): “Consistency of the Stationary Bootstrap Under Weak Moment Assumptions,” Economics Letters, 81(2), 273–278.

Davidson, J. (1994): Stochastic Limit Theory, Oxford University Press.

Dehling, H., D. Vogel, M. Wendler, and D. Wied (2014): “Testing for Changes in the Rank Correlation of Time Series,” Working Paper, arXiv:1203.4871v3.
Fremdt, S., J. Steinebach, L. Horváth, and P. Kokoszka (2012): “Testing the Equality of Covariance Operators in Functional Samples,” Scandinavian Journal of Statistics, 40(1), 138–152.

Gaissler, S. and F. Schmid (2010): “On Testing Equality of Pairwise Rank Correlations in a Multivariate Random Vector,” Journal of Multivariate Analysis, 101(10), 2598–2615.

Galeano, P. and D. Wied (2014): “Dating Multiple Changes in the Correlation Matrix,” SSRN Working Paper (http://ssrn.com/abstract=2431468).

Galeano, P. and Peña, D. (2007): “Covariance Changes Detection in Multivariate Time Series,” Journal of Statistical Planning and Inference, 137(1), 194–211.

Künsch, H. (1989): “The Jackknife and the Bootstrap for General Stationary Observations,” Annals of Statistics, 17(3), 1217–1241.

Krämer, W. and M. van Kampen (2011): “A Simple Nonparametric Test for Structural Change in Joint Tail Probabilities,” Economics Letters, 110(3), 245–247.

Krishan, C., R. Petkova, and P. Ritchken (2009): “Correlation Risk,” Journal of Empirical Finance, 16(3), 353–367.

Lahiri, S. (1999): “Theoretical Comparisons of Block Bootstrap Methods,” Annals of Statistics, 27(1), 386–404.

Lee, S., J. Ha, O. Na, and S. Na (2003): “The CUSUM Test for Parameter Change in Time Series Models,” Scandinavian Journal of Statistics, 30(4), 781–796.

Liu, R. and K. Singh (1992): “Moving Blocks Jackknife and Bootstrap Capture Weak Dependence,” in Exploring the Limits of the Bootstrap, Wiley, New York, 224–248.

Mandelbrot, B. (1963): “The Variation of Certain Speculative Prices,” Journal of Business, 36(4), 394–419.
NA, O., J. LEE, AND S. LEE (2012): “Change Point Detection in Copula ARMA–GARCH Models,” *Journal of Time Series Analysis*, 33(4), 554–569.

PAULY, M. (2009): *Eine Analyse bedingter Tests mit bedingten Zentralen Grenzwertsätzen für Resampling-Statistiken*, Dissertation, Heinrich-Heine-Universität Düsseldorf.

RADULOVIĆ, D. (2012): “Necessary and Sufficient Conditions for the Moving Blocks Bootstrap Central Limit Theorem of the Mean,” *Journal of Nonparametric Statistics*, 24(2), 343–357.

SHARIPOV, O. AND M. WENDLER (2012): “Bootstrap for the Sample Mean and for U-statistics of Mixing and Near-Epoch Dependent Processes,” *Journal of Nonparametric Statistics*, 24(2), 317–342.

VAN DER VAART, A. AND J. WELLNER (1996): *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer Verlag.

VOSTRIKOVA, L. (1981): “Detecting ‘Disorder’ in Multidimensional Random Processes,” *Soviet Mathematics Doklady*, 24, 55–59.

WIED, D. AND P. GALEANO (2012): “Monitoring Correlation Change in a Sequence of Random Variables,” *Journal of Statistical Planning and Inference*, 143(1), 186–196.

WIED, D., W. KRÄMER, AND H. DEHLING (2012): “Testing for a Change in Correlation at an Unknown Point in Time Using an Extended Functional Delta Method,” *Econometric Theory*, 28(3), 570–589.
A. Appendix

Proof of Theorem 1

Note that the null hypothesis and Assumption 4 imply that \( E(X_{i,t}X_{j,t}), 1 \leq i, j \leq p \), do not depend on \( t \).

At first, we need an invariance principle for the vector

\[
V_T(s) := \frac{1}{\sqrt{T}} \sum_{t=1}^{\tau(s)} \begin{pmatrix}
X_{1,t}^2 & - & E(X_{1,t}^2) \\
\vdots & & \vdots \\
X_{p,t}^2 & - & E(X_{p,t}^2) \\
X_{1,t} & - & E(X_{1,t}) \\
\vdots & & \vdots \\
X_{p,t} & - & E(X_{p,t}) \\
X_{1,t}X_{2,t} & - & E(X_{1,t}X_{2,t}) \\
X_{1,t}X_{3,t} & - & E(X_{1,t}X_{3,t}) \\
\vdots & & \vdots \\
X_{p-1,t}X_{p,t} & - & E(X_{p-1,t}X_{p,t})
\end{pmatrix},
\]

which is provided by Davidson (1994), p. 492. Thus, it holds, for \( T \to \infty \), \( V_T(s) \Rightarrow_d D_1^{1/2} W_{2p + \frac{2(p-1)}{3}} (s) \) on \( D \left( [0, 1], \mathbb{R}^{2p + \frac{p(p-1)}{2}} \right) \), where \( W_{2p + \frac{2(p-1)}{3}} (s) \) is a \( \left( \frac{p(p-1)}{2} + 2p \right) \)-dimensional Brownian Motion with independent components and \( D_1 \) is given in Assumption 1.
Now, one makes the observation that

\[
V_T(s) = \frac{\tau(s)}{\sqrt{T}} \begin{pmatrix}
\bar{X}_1^2(s) - \mathbb{E}(X_{1,t}^2) \\
\vdots \\
\bar{X}_p^2(s) - \mathbb{E}(X_{p,t}^2) \\
\bar{X}_1(s) - \mathbb{E}(X_{1,t}) \\
\vdots \\
\bar{X}_p(s) - \mathbb{E}(X_{p,t}) \\
\bar{X}_1\bar{X}_2(s) - \mathbb{E}(X_{1,t}X_{2,t}) \\
\bar{X}_1\bar{X}_3(s) - \mathbb{E}(X_{1,t}X_{3,t}) \\
\vdots \\
\bar{X}_{p-1}\bar{X}_p(s) - \mathbb{E}(X_{p-1,t}X_{p,t})
\end{pmatrix},
\]

where, for \(i = 1, \ldots, p\), \(\bar{X}_i^2(s) = \frac{1}{\tau(s)} \sum_{t=1}^{\tau(s)} X_{i,t} X_{i,t}\), \(\bar{X}_i^2(s) = \frac{1}{\tau(s)} \sum_{t=1}^{\tau(s)} X_{i,t}^2\) and, for \(1 \leq i < j \leq p\), \(\bar{X}_i\bar{X}_j(s) = \frac{1}{\tau(s)} \sum_{t=1}^{\tau(s)} X_{i,t} X_{j,t}\). The goal is to transform this vector of simple first and second order moments into the vector with the successively calculated correlation coefficients and then to apply the adapted functional delta method, Theorem A.1 in Wied et al. (2012). The transforming functions are

\[
f_1 : \mathbb{R}^{2p + \frac{p(p-1)}{2}} \to \mathbb{R}^{p + \frac{p(p-1)}{2}}
\]

\[
(x_1, \ldots, x_{2p + \frac{p(p-1)}{2}}) \to \begin{pmatrix}
x_1 - (x_{p+1}^2) \\
\vdots \\
x_p - (x_{2p}^2) \\
x_{2p+1} - x_{p}x_{p+1} \\
x_{2p+2} - x_{p}x_{p+2} \\
\vdots \\
x_{\frac{p(p+1)}{2}} - x_{p-1}x_{2p}
\end{pmatrix}
\]
for the transformation on the vector of variances and covariances and
\[
f_2 : \mathbb{R}^{p + \frac{p(p-1)}{2}} \rightarrow \mathbb{R}^{\frac{p(p-1)}{2}}
\]
\[
(x_1, \ldots, x_{p + \frac{p(p-1)}{2}}) \rightarrow \left( \begin{array}{c}
\frac{x_{p+1}}{\sqrt{x_1x_2}} \\
\frac{x_{p+2}}{\sqrt{x_1x_3}} \\
\vdots \\
\frac{x_{p+\frac{p(p-1)}{2}}}{\sqrt{x_{p-1}x_p}} \\
\end{array} \right)
\]

for the transformation on the vector of correlations.

We obtain, for \( T \rightarrow \infty \) and for arbitrary \( \epsilon > 0 \),

\[
W_T(s) := \frac{\tau(s)}{\sqrt{T}}(\hat{p}^{ij}_T - \rho^{ij})_{1 \leq i < j \leq p} \Rightarrow_d D_3D_2D_1^{1/2}W^{2p+\frac{p(p-1)}{2}}(s) \quad (1)
\]
on \( D([\epsilon, 1], \mathbb{R}^{2p+\frac{p(p-1)}{2}}) \) for matrices \( D_2 \sim \left( \left( p + \frac{p(p+1)}{2} \right) \times \left( 2p + \frac{p(p+1)}{2} \right) \right) \) and \( D_3 \sim \left( \frac{p(p+1)}{2} \times \left( p + \frac{p(p+1)}{2} \right) \right) \). Here, \( D_2 \) is the Jacobian matrix of \( f_1 \) and \( D_3 \) is the Jacobian matrix of \( f_2 \), evaluated at certain moments.

We are not interested in the exact (and cumbersome) structure of these matrices. But we observe that \( D_2 \) contains all \( \left( p + \frac{p(p+1)}{2} \right) \)-dimensional unit vectors and \( D_3 \) contains all \( \left( \frac{p(p+1)}{2} \right) \)-dimensional unit vectors (weighted with some constants) in its columns. Thus, \( D_2 \) and \( D_3 \) have full column rank. Together with Assumption 1, this implies that \( D_3D_2D_1^{1/2} \) has full column rank. Consequently, \( D_3D_2D_1D_2D_3' \) is invertible and positive definite.

Now, with an application of Theorem 4.2 in Billingsley (1968), we obtain, for \( T \rightarrow \infty \),

\[
W_T(s) \Rightarrow_d D_3D_2D_1^{1/2}W^{2p+\frac{p(p-1)}{2}}(s) \quad \text{on } D \left( [0, 1], \mathbb{R}^{2p+\frac{p(p-1)}{2}} \right). \quad \text{Moreover, it holds}
\]

\[
D_3D_2D_1^{1/2}W^{2p+\frac{p(p-1)}{2}}(s) \overset{d}{=} (D_3D_2D_2'D_3')^{1/2}W^{\frac{p(p-1)}{2}}(s)
\]

and from (1) it is easy to see (with \( s = 1 \)) that the asymptotic covariance matrix of

\[
\sqrt{T} (\hat{p}^{ij}_T)_{1 \leq i < j \leq p}
\]
is equal to \( D_3D_2D_1D_2'D_3' =: E \).
Proof of Theorem 2

We use a bootstrap theorem for near epoch dependent data for $V_T(1)$. Note that a univariate bootstrap central limit theorem conditionally on the original data (see for example Pauly, 2009, Lemma and Definition 2.7, for a precise definition of this type of convergence) is obtained by Calhoun (2013), Corollary 2.

For the multivariate generalization, we use an argument based on the Cramér-Wold device. Since we consider convergence of conditional distributions which are random variables and since an uncountable union of null sets is not necessary a null set again, we cannot directly apply the Cramér-Wold device. However, we can use an argument based on the Cramér-Wold device and Assumption 1 for the multivariate generalization (see Pauly, 2009, Theorem 3.19, Theorem 3.20 and the related material in this reference; the main argument is that we just consider rational $\lambda$ when applying the Cramér-Wold device).

Then, Condition 1 of Calhoun (2013), Corollary 2, is fulfilled with our Assumption 3, Condition 2 as well as the condition $\sum_{t=1}^{n}(\mu_{nt} - \bar{\mu})^2 = o(n^{1/2})$ with our Assumption 1 and our Assumption 4, Condition 3 with our Assumption 2 and Condition 4 with our Assumption 5.

Summing up the previous discussion, the block bootstrap consistently estimates the distribution law of $V_T(1)$. But then, with the standard (functional) delta method for the bootstrap (van der Vaart and Wellner, 1996, Theorem 3.9.11.) transforming $V_T(1)$ to $W_T(1)$, also the law of $W_T(1)$ is consistently estimated. That means that, for $T \to \infty$,

$$d \left( \mathcal{L} \left( \sqrt{T} (\hat{\rho}^{ij}_{b,T} - \hat{\rho}^{ij}_{T})_{1 \leq i < j \leq p} | X_1, \ldots, X_T \right), \mathcal{L} \left( D_3 D_2 D_1^{1/2} W^{2p+\frac{p(p+1)}{4}}(1) \right) \right) \to_p 0,$$

where $d$ is a metric of weak convergence (see Pauly, 2009, p. 36) and $\mathcal{L}(\cdot)$ denotes the distribution of a random vector.
Now consider, for $1 \leq i < j \leq p$ and the $\delta$ from Assumption 6, the conditional expectation
\[
E \left( \left| \sqrt{T} (\hat{\rho}^{ij}_{b,T} - \hat{\rho}^{ij}_T) \right|^{2+\delta} \mid X_1, \ldots, X_T \right) =: C_T.
\]
By Assumption 6, $C_T$ is stochastically bounded. Then, with Lemma 1 in Cheng, 2011, we can consistently estimate the asymptotic covariance matrix of $W_T(1)$.

**Proof of Theorem 3**

Transferring the proof of Theorem 1, we obtain, under $H_1$, for $T \to \infty$, $V_T(s) \Rightarrow_d D_1^{1/2} W_{2p+\frac{p(p-1)}{2}}(s) + A(s)$ on $D \left([0,1], \mathbb{R}^{2p+\frac{p(p-1)}{2}} \right)$. Here,
\[
A = \left(0, \ldots, 0, \int_0^s g(u) \, du \right)'
\]
(note that $g(u)'$ is the transpose of the function $g$).

So,
\[
W_T(s) := \frac{\tau(s)}{\sqrt{T}} (\hat{\rho}^{ij}_{T(s)} - \rho^{ij})_{1 \leq i < j \leq p} \Rightarrow_d D_1^{1/2} W_{2p+\frac{p(p-1)}{2}}(s) + D_3 D_2 A(s),
\]
where $D_3$ and $D_2$ are the matrices mentioned in the proof of Theorem 1. Due to the structure of $D_3$ and $D_2$, we have
\[
D_3 D_2 A(s) = M \begin{pmatrix}
\frac{1}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} \int_0^s g_1(u) \, du \\
\frac{1}{\sqrt{\text{Var}(X_1)\text{Var}(X_3)}} \int_0^s g_2(u) \, du \\
\vdots \\
\frac{1}{\sqrt{\text{Var}(X_{p-1})\text{Var}(X_p)}} \int_0^s g_{p(p-1)}(u) \, du
\end{pmatrix}
\]
This completes the proof.

**Proof of Theorem 4**

Under local alternatives, for $i = 1, \ldots, p$, $E(X_{i,t})$ and $E(X_{i,t}^2)$ are constant, respectively.
Moreover, for $1 \leq i < j \leq p$, it holds

$$\sum_{t=1}^{T} (E(X_{i,t}X_{j,t}) - \overline{X_iX_j}(1))^2 = o(T^{1/2}).$$

Therefore, the condition “$\sum_{t=1}^{n} (\mu_{nt} - \bar{\mu})^2 = o(n^{1/2})$” of Corollary 2 in Calhoun (2013) is fulfilled. The other conditions are fulfilled with the same arguments as in Theorem 2. Then, by this corollary and the Cramér-Wold Theorem, we estimate $E$ consistently with the bootstrap estimator as described in the proof of Theorem 2. ■
Table 1: Empirical size and empirical power (times 100, respectively) of the multivariate correlation test; columns 5,6 give empirical rejection probabilities for the matrix-based test, columns 7,8 give rejection probabilities for the Bonferroni-Holm procedure

| $MA$ dist. | $\rho_0$ | rej.prob. | rej.prob. |
|------------|----------|-----------|-----------|
|            | $T = 200$ | $T = 500$ | $T = 200$ | $T = 500$ |
| $\Delta \rho = 0$ |          |           |           |
| 0          | N 0      | 2.8       | 3.8       | 2.9       | 3.5       |
| 0          | N 0.5    | 4.4       | 4.4       | 5.5       | 4.3       |
| 0          | t 0      | 8.7       | 6.7       | 7.8       | 4.9       |
| 0          | t 0.5    | 11.5      | 8.1       | 10.1      | 6.3       |
| 0.5        | N 0      | 4.8       | 5.3       | 4.0       | 4.0       |
| 0.5        | N 0.5    | 7.4       | 6.2       | 7.5       | 5.1       |
| 0.5        | t 0      | 13.1      | 9.4       | 9.7       | 5.5       |
| 0.5        | t 0.5    | 17.1      | 12.3      | 13.2      | 8.1       |
| $\Delta \rho = 0.2$ |          |           |           |
| 0          | N 0      | 32.2      | 89.1      | 26.3      | 73.5      |
| 0          | N 0.5    | 43.6      | 90.5      | 55.6      | 93.7      |
| 0          | t 0      | 19.0      | 32.3      | 15.8      | 20.2      |
| 0          | t 0.5    | 30.1      | 41.3      | 35.4      | 44.9      |
| 0.5        | N 0      | 30.3      | 80.8      | 24.7      | 63.9      |
| 0.5        | N 0.5    | 42.3      | 82.9      | 52.3      | 87.9      |
| 0.5        | t 0      | 24.5      | 34.7      | 18.4      | 22.2      |
| 0.5        | t 0.5    | 36.6      | 44.8      | 38.9      | 46.7      |
| $\Delta \rho = -0.2$ |          |           |           |
| 0          | N 0      | 74.4      | 100.0     | 29.2      | 80.7      |
| 0          | N 0.5    | 16.8      | 79.5      | 20.3      | 80.7      |
| 0          | t 0      | 36.4      | 64.0      | 16.4      | 21.2      |
| 0          | t 0.5    | 13.6      | 22.5      | 10.1      | 15.6      |
| 0.5        | N 0      | 65.8      | 99.6      | 25.9      | 70.1      |
| 0.5        | N 0.5    | 15.4      | 66.6      | 15.9      | 67.9      |
| 0.5        | t 0      | 39.9      | 64.9      | 20.1      | 23.4      |
| 0.5        | t 0.5    | 18.1      | 24.7      | 12.4      | 16.8      |
(a) Rolling correlations between Total and Sanofi

(b) Rolling correlations between Total and Siemens

Figure 1: Rolling correlations
(a) Rolling correlations between Total and BASF

(b) Rolling correlations between Sanofi and Siemens

Figure 2: Rolling correlations
(a) Rolling correlations between Sanofi and BASF

(b) Rolling correlations between Siemens and BASF

Figure 3: Rolling correlations
Figure 4: Evolution of successively calculated correlations