Generalized Spectral Tests for High Dimensional Multivariate Martingale Difference Hypotheses

Xuexin Wang

School of Economics and WISE

Xiamen University, China

Abstract

In this study, we propose new generalized spectral tests for high dimensional multivariate martingale difference hypotheses. Essentially, we introduce a nonintegrable weighting function, which is pivotal in generalized spectral tests. We demonstrate rigorously that this weighting function enjoys attractive advantages in multivariate situations, which is strikingly different from the conventional weighting function employed in previous works. First, the weighting function tends to deliver infinite weighting values in a neighborhood of the origin, which has very important power implications for generalized spectral tests. Second, its weighting values increase with the dimension of the conditioning set, which lends our proposed statistics excellent theoretical and empirical properties in high-dimensionality situations. Third, under this weighting function, a bias-reduced statistic can be constructed to improve finite sample properties. Last, under this weighting function, the proposed statistics enjoy analytical forms and are, therefore, quite easy to compute in practice. Our new generalized spectral tests are robust to high dimensionality of the conditioning set, which can be larger than the sample size. They are consistent against a wide class of nonlinear-dependent multivariate processes. Monte Carlo simulations reveal that the bias-reduced statistic generally performs better than its competitors substantially. Moreover, it is robust to the heteroskedasticity of unknown forms and heavy-tails in the data generating processes. We apply our approach to test the

Email: xuexinwang@xmu.edu.cn. Address correspondence to Xuexin WANG, WISE,Xiamen University, China.

efficient market hypothesis on the US stock market, using data sets on the monthly and daily data of portfolios sorted by industry. Our test results provide strong evidence against the efficient market hypothesis with respect to the US stock market at monthly frequency.

Keywords: Efficient Market Hypothesis; Generalized Spectral Tests; Nonintegrable Weighting Function; High-dimensionality; Bias Reduction

JEL Classification Numbers: C12, C22

1 Introduction

In dynamic contexts, many economic and financial theories indicate martingale or martingale difference sequence (MDS) for some economic variables. For example, Hall (1978) demonstrates that consumption follows a random walk under a dynamic utility optimization situation, which is a special case of univariate martingale in the sense that the information set only contains historical information of consumption. The efficient market hypothesis (EMH), developed by Fama (1970), states that, in an efficient asset market, asset prices always "fully reflect" available information. Its statistical implications are associated with martingale or MDS directly. To be specific, suppose there are q assets in a market. Following Fama (1970), an efficient market implies that, for any asset i,

$$E\left(R_{i,t+1}|\Phi_t\right) = 0,$$

where $R_{i,t+1} = r_{i,t+1} - E\left(r_{i,t+1}|\Phi_t\right)$, the return at $t+1$ in excess of the equilibrium expected return projected at t, Φ_t contains fully available information on the market. In aggregate, we have

$$E\left(R_{t+1}|\Phi_t\right) = 0, \tag{1}$$

where $R_{t+1} = \left(R_{1,t+1}, ..., R_{q,t+1}\right)'$. Without further restrictions on the information set Φ_t, (1) is not statistically testable. Fama (1970) introduced a weak form of EMH by setting $\Phi_t = I_t = \left(R_t', R_{t-1}', ...\right)'$,[1] i.e.,

$$E\left(R_{t+1}|I_t\right) = 0, \tag{2}$$

[1] The definition of the information set avoids the measure-theoretic terminology, and emphasizes the dimensionality involved. More formally, the information set is the σ-field generated by the history information on R_t.

which is a multivariate MDS in the sense that the elements in I_t, that is, $R_t, R_{t-1}, ...$, are q-dimensional vectors, when $q > 1$.

Naturally, it is of great interest to check martingale hypotheses or martingale difference hypotheses (MDHs) implied by theories. In econometrics, there is a large body of literature on testing (directly or indirectly) a univariate MDH or a univariate martingale hypothesis. Notably, the variance ratio test, proposed by Lo and MacKinlay (1988), is popular in finance literature. Further development along this line includes Chow and Denning (1993), Choi (1999), Wright (2000), Chen and Deo (2006), and Kim (2006), among others. An alternative approach is the Box–Pierce-Ljung Portmanteau test proposed by Box and Pierce (1970) and Ljung and Box (1978), which tests the lack of autocorrelations up to some finite lags for a time series sequence. Lobato et al. (2002), Lobato et al. (2001), Lobato (2001), and Wang and Sun (2020), among others, propose modified Portmanteau tests, allowing for weaker assumptions. Test statistics allowing for an infinite number of autocorrelations are proposed by Durlauf (1991), Deo (2000) and Shao (2011), based on spectral transformation. The tests mentioned above only check the MDH indirectly. Moreover, they cannot detect nonlinear dependence with zero correlations, which seems to be a common feature in asset returns. Tests which can detect nonlinear dependence are proposed by Hong (1999), Domínguez and Lobato (2003), Hong and Lee (2003), Hong and Lee (2005), Kuan and Lee (2004), and Escanciano and Velasco (2006) (hereafter, EV (2006)), among others. See Escanciano and Lobato (2009) for a comprehensive review of univariate MDH testing. Literature on testing martingale hypotheses includes Park and Whang (2005), Escanciano (2007) and Phillips and Jin (2014).

On the other hand, literature on testing high dimensional multivariate MDHs is scarce. In practice, testing multivariate MDHs, such as (2), reduces to checking whether returns on an asset index or an asset portfolio follow a univariate MDS, e.g., Lo and MacKinlay (1988), among others. That is, one tests the MDH

$$E\left(T'R_{t+1}|I_t\right) = 0, \tag{3}$$

where T is a $q \times 1$ weighting vector, $I_t = (T'R_t, T'R_{t-1}...)'$. This practice is a convenient dimension reduction. However, given a weighting vector T, (3) does not imply (2). Only in recent years have tests for high dimensional multivariate MDHs been proposed. Chang et al. (2017) propose tests for high dimensional white noise with maximum cross-correlations, while Tsay (2020) employs Spearman's rank correlation and the theory of extreme values to test high dimensional white noise. As in the univariate situation, these tests cannot detect nonlinear dependence. For multivariate tests that can detect nonlinear dependence, EV (2006) argue that their generalized spectral tests can be extended to multivariate situations with ease. However, theoretical analysis and Monte Carlo simulations in our study reveal that the multivariate versions of these tests suffer from "curse of dimensionality" substantially: their empirical sizes and powers are as low as zero, even when the dimension is moderate.

In this study, we propose new generalized spectral tests for multivariate MDHs, especially suitable for high-dimensionality situations. On the surface, our test statistics seem to bear some resemblance to the one proposed by EV (2006); however, the representation of the statistic, the asymptotic theory, and the theoretical and empirical properties are different. Essentially, we introduce a nonintegrable weighting function that is pivotal in generalized spectral tests and deviates from EV (2006) substantially. We demonstrate rigorously that this weighting function enjoys attractive advantages in multivariate situations, which is strikingly different from the conventional weighting function employed in previous works. First, the weighting function tends to deliver infinite weighting values in a neighborhood of the origin, which has very important power implications. Second, its weighting values increase with the dimension of the conditioning set, which lends our proposed statistics excellent theoretical and empirical properties in high-dimensionality situations. Third, under this weighting function, a bias-reduced statistic can be constructed to improve finite sample properties. Finally, under this weighting function, the proposed statistics enjoy analytical forms, and are, therefore, quite easy to compute in practice.

Our new generalized spectral tests are robust to high dimensionality of the conditioning set, which can be larger than the sample size. They are consistent against a wide class of nonlinear-dependent multivariate processes. Monte Carlo simulations reveal that the bias-reduced statistic generally performs better than its competitors substantially. Moreover, it is robust to the heteroskedasticity of unknown forms and heavy-tails in the data generating processes (DGPs).

Using the new generalized spectral test, we test the EMH on the US stock market. Our data sets are the monthly and daily data on portfolios sorted by industry. We find strong evidence against the EMH on the US stock market at monthly frequency.

The remainder of the paper is organized as follows. Section 2 presents the multivariate MDHs that we focus on and the generalized spectral testing framework. Section 3 first introduces the nonintegrable weighting function and discusses its theoretical properties, and then presents two versions of our proposed statistics. Section 4 is devoted to establishing the asymptotic theory for these statistics. Section 5 proposes a bootstrapping procedure to approximate the asymptotic distribution of our proposed statistics. Section 6 reports simulation evidence, while Section 7 presents an empirical application. The last section concludes. All proofs are presented in the Appendix.

We introduce some notations. For a vector X, $X^{'}$ is its transpose vector. The imaginary unit $i = \sqrt{-1}$. For a complex-valued function $f(\cdot)$, its complex conjugate is denoted by $f^c(\cdot)$ and $|f(\cdot)|^2 = f(\cdot) f^c(\cdot)$. In a Euclidean space, the scalar product of the vectors τ and ς is denoted by $\langle \tau, \varsigma \rangle$. The Euclidean norm of $X = (X_1, ..., X_q)$ in $\mathbb{C}^q$ is $\|X\|$, where $\|X\|^2 = X'X^c$. X^+ and X^{++} are independent copies of X, that is, X^+, X^{++}, and X are independent and identically distributed (i.i.d). For elements f and g in a Hilbert space, their inner product is denoted as $\langle f, g \rangle_H$, and the norm of f is denoted as $\|f\|_H := \sqrt{\langle f, f \rangle_H}$.

2 Basic setting and testing framework

Let $Y_t = (Y_{1,t}, ..., Y_{q,t})' \in \mathbb{R}^q$, $t \in \mathbb{Z}$ be a q-variate real-valued stationary and ergodic time series, the information set $I_t = \left\{Y_t', Y_{t-1}', ...\right\}'$. We tackle the problem of testing that, almost surely (a.s),

$$H_0^* : E\left[Y_t | I_{t-1}\right] = E\left(Y_t\right).$$

H_0^* is a natural generalization of the univariate MDH. The methodology developed here can be easily extended to test other types of multivariate MDHs without many modifications. Notably, we can test an individual element or a subvector of Y_t following a martingale difference:

$$H_0^{sub*} : E\left[Y_t^{sub}|I_{t-1}\right] = E\left(Y_t^{sub}\right), \text{a.s.}, E\left(Y_t^{sub}\right) \in \mathbb{R}^p$$

where Y_t^{sub} is a subsector of Y_t, $p < q$.

The high dimensionality of H_0^* or H_0^{sub*} comes from two aspects: one is that I_t contains all the historical information on Y_t; the other is the dimension of Y_t. This makes the problem of testing H_0^* or H_0^{sub*} extremely challenging. To circumvent the difficulty, we focus on testing the weaker null hypotheses

$$H_0 : E\left[Y_t|Y_{t-j}\right] = E\left(Y_t\right), \text{a.s. for all } j = 1, 2, ... \tag{4}$$

$$H_A : P\left(E\left[Y_t|Y_{t-j}\right] \neq E\left(Y_t\right)\right) > 0, \text{ for some } j,$$

and

$$H_0^{sub} : E\left[Y_t^{sub}|Y_{t-j}\right] = E\left(Y_t^{sub}\right), \text{a.s. for all } j = 1, 2, ... \tag{5}$$

$$H_A^{sub} : P\left(E\left[Y_t^{sub}|Y_{t-j}\right] \neq E\left(Y_t^{sub}\right)\right) > 0, \text{ for some } j.$$

In other words, in conditioning sets, it is equally important to consider contemporaneous variables, as these variables are potentially contemporaneously and serially dependent. The widely accepted argument regarding the necessity of multivariate methods holds in constructing H_0. Clearly, H_0^* implies H_0; thus, a rejection of H_0 implies a rejection of H_0^*. Note that Y_{t-j} still has a q dimension, which can be quite large.

To test H_0, we rely on the equivalence between conditional moment restriction and an infinite number of unconditional moment conditions, i.e., for $j = 1, 2, ...,$

$$E\left[Y_t|Y_{t-j}\right] = E\left(Y_t\right), \text{a.s.} \Longleftrightarrow E\left[\left(Y_t - E\left(Y_t\right)\right)\exp\left(i\left\langle\tau, Y_{t-j}\right\rangle\right)\right] = 0, \text{ for almost all } \tau \in \mathbb{R}^q.$$

Denote

$$\gamma_j\left(\tau\right) = E\left[\left(Y_t - E\left(Y_t\right)\right)\exp\left(i\left\langle\tau, Y_{t-j}\right\rangle\right)\right].$$

Define $\gamma_{-j}(\tau) = \gamma_j(\tau)$ for $j = 1, 2, ...$ The Fourier transform of the functions $\gamma_j(\tau)$ is

$$f(\kappa, \tau) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j(\tau) \exp(-ij\kappa), \ \forall \kappa \in [-\pi, \pi], \tau \in \mathbb{R}^q.$$

Note that $f(\kappa, \tau)$ exists if

$$\sup_{\tau \in \mathbb{R}^q} \left\| \gamma_j(\tau) \right\| < \infty,$$

which holds under some proper conditions.

We use a generalized spectral distribution function as in EV(2006), i.e.,

$$H(\lambda, \tau) = 2 \int_0^{\lambda\pi} f(\kappa, \tau) \, d\kappa = \gamma_0(\tau) \lambda + 2 \sum_{j=1}^{\infty} \gamma_j(\tau) \frac{\sin(j\pi\lambda)}{j\pi}.$$

Then, testing H_0 is equivalent to testing

$$H(\lambda, \tau) = \gamma_0(\tau) \lambda, \text{ for all } \lambda \in [0, 1] \text{ and } \tau \in \mathbb{R}^q.$$

It is pivotal to construct a distance measure between $H(\lambda, \tau)$ and $\gamma_0(\tau)\lambda$ such that

$$\int_{\mathbb{R}^q} \int_0^1 \|H(\lambda, \tau) - \gamma_0(\tau) \lambda\|^2 \, d\lambda W(d\tau),$$

where $W(d\tau)$ is a proper weighting function. Under some regularity conditions, it is not difficult to obtain

$$\int_{\mathbb{R}^q} \int_0^1 \|H(\lambda, \tau) - \gamma_0(\tau) \lambda\|^2 \, d\lambda W(d\tau) = 2 \sum_{j=1}^{\infty} \frac{\int_{\mathbb{R}^q} \left\| \gamma_j(\tau) \right\|^2 W(d\tau)}{(j\pi)^2}. \tag{6}$$

It is noted that this measure is a sum of individual distance measures between $\gamma_j(\tau)$ and 0_q, which involve the weighting function $W(d\tau)$.

A test statistic for testing H_0 can then be constructed based on a sample analog of (6). Specifically, given a sample $\{Y_t\}_1^n$, define

$$\hat{\gamma}_j(\tau) = \frac{1}{n-j} \sum_{t=j+1}^{n} \left(Y_t - \bar{Y}_{n-j} \right) \exp(i \langle \tau, Y_{t-j} \rangle),$$

with

$$\bar{Y}_{n-j} = \frac{1}{n-j} \sum_{t=j+1}^{n} Y_t.$$

A generalized spectral test has the representation

$$\begin{aligned} D_n^2 &= 2\sum_{j=1}^{n-1} \frac{(n-j)/2}{(j\pi)^2} \int_{\mathbb{R}^q} \left\|\hat{\gamma}_j(\tau)\right\|^2 W(d\tau) \\ &= \sum_{j=1}^{n-1} \frac{(n-j)}{(j\pi)^2} \int_{\mathbb{R}^q} \left\|\hat{\gamma}_j(\tau)\right\|^2 W(d\tau), \end{aligned}$$

where $(n-j)/2$ is a rescaling term for $\int_{\mathbb{R}^q} \left\|\hat{\gamma}_j(\tau)\right\|^2 W(d\tau)$, which considers the sample size in computing $\hat{\gamma}_j(\tau)$ to improve the finite sample properties of D_n^2. This representation is convenient in calculating D_n^2 in practice.

Alternatively, define

$$\hat{H}_n(\lambda,\tau) = \hat{\gamma}_0(\tau)\lambda + 2\sum_{j=1}^{n-1} \left(1-\frac{j}{n}\right)^{1/2} \hat{\gamma}_j(\tau) \frac{\sin j\pi\lambda}{j\pi}$$

with $\left(1-\frac{j}{n}\right)^{1/2}$ a finite sample correction term, as in Hong (1999a), and define

$$\begin{aligned} \hat{S}_n(\lambda,\tau) &= \left(\frac{n}{2}\right)^{1/2} \left(\hat{H}_n(\lambda,\tau) - \hat{\gamma}_0(\tau)\lambda\right) \\ &= \sum_{j=1}^{n-1} (n-j)^{1/2} \hat{\gamma}_j(\tau) \frac{\sqrt{2}\sin j\pi\lambda}{j\pi}. \end{aligned}$$

It can then be shown, after some calculations, that

$$D_n^2 = \int_{\mathbb{R}^q} \int_0^1 \left\|\hat{S}_n(\lambda,\tau)\right\|^2 d\lambda W(d\tau).$$

This representation of D_n^2 is useful in deriving asymptotic theory.

3 A Unique Weighting Function and New Test Statistics

It is obvious that the choice of $W(\tau)$ has crucial implications on the power performance of D_n^2. The optimal choice of $W(\cdot)$ depends on the true alternative at hand, as pointed out by Epps and Pulley (1983) in a univariate situation involving empirical characteristic functions. Additionally, they pointed out three considerations on the choice of $W(\tau)$. In our scenario, first, $W(\tau)$ should assign high weight values where $\hat{\gamma}_j(\tau)$ is large under alternatives. Second, $W(\tau)$ should give high weighting values around the neighborhood of the origin of τ , as $\hat{\gamma}_j(\tau)$ can be shown to be more accurately an estimate of $\gamma_j(\tau)$ in this neighborhood. Third, from a practical perspective, it is important to choose a $W(\tau)$ to give D_n^2 an analytical form, otherwise, it would be a formidable task to solve integrals numerically in multivariate situations.

Some weighting functions, which are integrable, have been chosen in previous literature in the case of $q = 1$, for example, Kuan and Lee (2004) and EV(2006). However, theoretical analysis and Monte Carlo simulations reveal that generalized spectral tests based on these integrable weighting functions are not good choices in multivariate situations, as they have extremely poor finite sample properties in this case: their empirical sizes and powers can be as low as zero for a moderate q, not to mention a higher dimension.

To address this issue, we choose an alternative weighting function, which is nonintegrable, such that

$$W(\tau) = \frac{1}{c_q \, ||\tau||^{q+1}} \tag{7}$$

with c_q defined in Lemma 3.1. In the following, we write $\int_{\mathbb{R}^q} \frac{\|\gamma_j(\tau)\|^2}{c_q\|\tau\|^{q+1}} d\tau = \int_{\mathbb{R}^q} \left\|\gamma_j(\tau)\right\|^2 \omega(d\tau)$, where $\omega(d\tau) = \left(c_q \|\tau\|^{q+1}\right)^{-1} d\tau$ for notational simplicity.

By employing $\omega(d\tau)$, $\int_{\mathbb{R}^q} \left\|\gamma_j(\tau)\right\|^2 \omega(d\tau)$ has a very convenient analytical form. To derive this property, we rely on the following Lemma.

Lemma 3.1 *For all $X \in \mathbb{R}^q$*

$$\int_{R^q} \frac{1 - \cos\langle\tau, X\rangle}{\|\tau\|^{q+1}} d\tau = c_q \|X\|, \tag{8}$$

where

$$c_q = \frac{\pi^{(q+1)/2}}{\Gamma\left(\left(q+1\right)/2\right)},$$

in which $\Gamma(r) = \int_0^\infty \tau^{r-1} e^{-\tau} d\tau, r \neq 0, -1, -2....$ *The integrals at* 0 *and* ∞ *are meant in the principal value sense:* $lim_{\varepsilon \to 0} \int_{\mathbb{R}^q \backslash \{\varepsilon B + \varepsilon^{-1} B^c\}}$*, where* B *is the unit ball centered at 0 and* B^c *is the complement of* B*, and*

$$\int_{\mathbb{R}^q} \frac{\sin\left(\langle \tau, X \rangle\right)}{\|\tau\|^{q+1}} d\tau = 0.$$

Proof. See the Appendix. ■

Remark 3.1 *A more general result, allowing for* $1/\|\tau\|^{q+\alpha}$*, where* $0 < \alpha < 2$*, was established first in Székely et al. (2007).*

Based on Lemma 3.1, it is easy to show that

$$\begin{aligned} \int_{\mathbb{R}^q} \left\|\gamma_j(\tau)\right\|^2 \omega(d\tau) &= \int_{\mathbb{R}^q} E\left[\left(Y_t - E\left(Y_t\right)\right)' \left(Y_t^+ - E\left(Y_t\right)\right) \exp\left(i\left\langle \tau, Y_{t-j} - Y_{t-j}^+ \right\rangle\right)\right] \omega(d\tau) \\ &= -E\left[\left(Y_t - E\left(Y_t\right)\right)' \left(Y_t^+ - E\left(Y_t\right)\right) \left\|Y_{t-j} - Y_{t-j}^+\right\|\right], \end{aligned}$$

where (Y_t, Y_{t-j}) and $\left(Y_t^+, Y_{t-j}^+\right)$ are i.i.d. A natural sample analog for $\int_{\mathbb{R}^q} \left\|\gamma_j(\tau)\right\|^2 \omega(d\tau)$ in finite samples is then

$$-\frac{1}{(n-j)^2} \sum_{k=j+1}^{n} \sum_{l=j+1}^{n} \left[\left(Y_k - \bar{Y}_{n-j}\right)' \left(Y_l - \bar{Y}_{n-j}\right) \|Y_{k-j} - Y_{l-j}\|\right]. \tag{9}$$

Accordingly, one version of our proposed statistic, which is denoted as MD_n^2, is

$$MD_n^2 = -\sum_{j=1}^{n-1} \frac{1}{(n-j)(j\pi)^2} \sum_{k=j+1}^{n} \sum_{l=j+1}^{n} \left[\left(Y_k - \bar{Y}_{n-j}\right)' \left(Y_l - \bar{Y}_{n-j}\right) \|Y_{k-j} - Y_{l-j}\|\right].$$

Similarly, the statistic for testing H_0^{sub} is

$$MD_{sub,n}^2 = -\sum_{j=1}^{n-1} \frac{1}{(n-j)(j\pi)^2} \sum_{k=j+1}^{n} \sum_{l=j+1}^{n} \left[\left(Y_k^{sub} - \bar{Y}_{n-j}^{sub}\right)' \left(Y_l^{sub} - \bar{Y}_{n-j}^{sub}\right) \|Y_{k-j} - Y_{l-j}\|\right].$$

It is noted that the test statistics proposed by EV(2006) in the univariate situation also enjoy an-

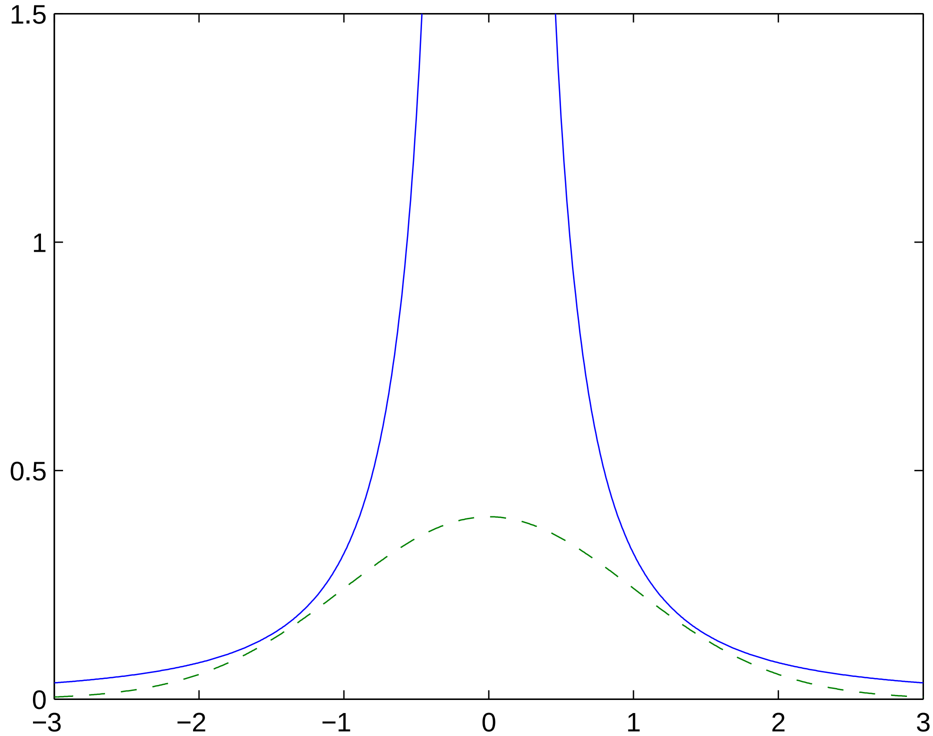


Figure 1: Standard normal density function (dashed curve) vs. $1/(\pi\tau^2)$ (solid curve)

alytical forms in the multivariate situation, when proper weighting functions are chosen. Notably, when a multivariate standard normal density function is chosen, their statistic is

$$EV_n^2 = \sum_{j=1}^{n-1} \frac{1}{(n-j)\,(j\pi)^2} \sum_{k=j+1}^{n} \sum_{l=j+1}^{n} \left[\left(Y_k - \bar{Y}_{n-j}\right)' \left(Y_l - \bar{Y}_{n-j}\right) \exp\left(-0.5\,\|Y_{k-j} - Y_{l-j}\|^2\right) \right].$$

However, additional features of (7) stand out, which are strikingly different from the conventional integrable weighting functions and lend our generalized spectral tests excellent finite sample properties for testing high dimensional multivariate MDHs. One outstanding feature is that its weighting values go to infinity as $\|\tau\| \to 0$, which accords perfectly with the second principle of Epps and Pulley (1983). By contrast, an integrable weighting function only delivers finite values in this neighborhood. Figure 1 demonstrates the striking difference between a standard normal density function and (7).

Furthermore, (7) is an increasing function of q, which renders our generalized spectral tests robust to high dimensionality of Y_{t-j}. Note that

$$\begin{aligned}\left(c_q \|\tau\|^{q+1}\right)^{-1} &= \frac{\Gamma\left(\left(q+1\right)/2\right)}{\pi^{(q+1)/2}}\left(\|\tau\|^{q+1}\right)^{-1} \\ &= \frac{(q-1)!!}{(2\pi)^{q/2}\|\tau\|^{q+1}},\end{aligned}$$

where $q!!$ is a double factorial such that

$$q!! = \begin{cases} q\cdot(q-2)\ldots 5\cdot 3\cdot 1 & q>0 \text{ odd} \\ q\cdot(q-2)\ldots 6\cdot 4\cdot 2 & q>0 \text{ even} \\ 1 & q=-1,0. \end{cases}$$

By applying an approximation to $(q-1)!!$, we have

$$\begin{aligned}\left(c_q \|\tau\|^{q+1}\right)^{-1} &\approx c\,(q-1)^{q/2}\, e^{-(q-1)/2}\frac{1}{(2\pi)^{q/2}\|\tau\|^{q+1}} \\ &\approx \frac{c\sqrt{e}}{\|\tau\|}\left(\frac{q-1}{2\pi e\,\|\tau\|^2}\right)^{q/2},\end{aligned}$$

where $c=\sqrt{\pi}$ for $q-1$ is even, $\sqrt{2}$ for $q-1$ is odd. It is obvious, for a fixed value $\|\tau\|$ in a neighborhood of the origin, that (7) is an increasing function of q. On the other hand, for a multivariate standard normal density function, it is a decreasing function of q. Notably, its weighting value equals $(2\pi)^{-q/2}$ at the origin, being the maximum. It shrinks to zero quickly when q increases. Consequently, EV_n^2 can be very close to zero under the null and alternatives, even when q is moderate.

The nonintegrable weighting function (7) has been employed in Székely et al. (2007), Székely and Rizzo (2009), Székely and Rizzo (2014), Shao and Zhang (2014), Davis et al. (2018), Zhang et al. (2018), and Yao et al. (2018), among others, in the statistics literature. Both Yao et al. (2018) and Davis et al. (2018) apply the distance correlation (variance) proposed by Székely et al. (2007) to test independence, not MDH, for high dimensional data or time series data. Zhang et al. (2018) propose a test for a MDH where the dimension of information set increases with the sample size. Both the testing framework and the test statistic are different from ours.

Our contribution here is that we sort out the attractive advantages of (7) in relation to high-dimensionality situations. To the best of our knowledge, there are no other classes of weighting functions that share the same features as (7). In this sense, it is unique.

3.1 A Bias-reduced Estimator for $\int_{\mathbb{R}^q} \left\|\gamma_j(\tau)\right\|^2 \omega(d\tau)$ and A Bias-reduced Statistic

It is noted that MD_n^2 avoids lag term truncations, as we consider $\int_{\mathbb{R}^q} \left\|\hat{\gamma}_j(\tau)\right\|^2 \omega(d\tau)$ up to $j = n-1$. However, (9) is a biased estimator for $\int_{\mathbb{R}^q} \left\|\gamma_j(\tau)\right\|^2 \omega(d\tau)$. As j increases, this finite-sample bias becomes more severe, since sample size becomes smaller. These biases, in aggregate, can result in poor finite sample properties for MD_n^2 in practice. To address this shortcoming, we employ an alternative estimator for $\int_{\mathbb{R}^q} \left\|\gamma_j(\tau)\right\|^2 \omega(d\tau)$, which will be shown to be unbiased or bias reduced.

Denote

$$a_{jkl} = \|Y_{k-j} - Y_{l-j}\|, b_{kl} = \frac{1}{2}\|Y_k - Y_l\|^2, k, l = j+1, ..., T.$$

It can be shown that (see the proof of Lemma 3.2 for some clues)

$$-\sum_{k=j+1}^{n}\sum_{l=j+1}^{n}\left[\left(Y_k - \bar{Y}_{n-j}\right)'\left(Y_l - \bar{Y}_{n-j}\right)\|Y_{k-j} - Y_{l-j}\|\right] = \sum_{k=j+1}^{n}\sum_{l=j+1}^{n} A_{jkl}B_{kl},$$

where matrices A_j and B have (k,l)th entry

$$B_{kl} = b_{kl} - \frac{1}{n-j}\sum_{l=j+1}^{n} b_{kl} - \frac{1}{n-j}\sum_{k=j+1}^{n} b_{kl} + \frac{1}{(n-j)^2}\sum_{k=j+1}^{n}\sum_{l=j+1}^{n} b_{kl},$$

$$A_{jkl} = a_{jkl} - \frac{1}{n-j}\sum_{l=j+1}^{n} a_{jkl} - \frac{1}{n-j}\sum_{k=j+1}^{n} a_{jkl} + \frac{1}{(n-j)^2}\sum_{k=j+1}^{n}\sum_{l=j+1}^{n} a_{jkl}.$$

We define new matrices $\widetilde{A}_j$ and $\tilde{B}$ as having (k,l)th entry

$$\widetilde{B}_{kl} = \begin{cases} b_{kl} - \frac{1}{n-j-2}\sum_{l=j+1}^{n} b_{kl} - \frac{1}{n-j-2}\sum_{k=j+1}^{n} b_{kl} + \frac{1}{(n-j-1)(n-j-2)}\sum_{k=j+1}^{n}\sum_{l=j+1}^{n} b_{kl} & k \neq l \\ 0 & k = l \end{cases}$$

$$\widetilde{A}_{jkl} = \begin{cases} a_{jkl} - \frac{1}{n-j-2}\sum_{l=j+1}^{n} a_{jkl} - \frac{1}{n-j-2}\sum_{k=j+1}^{n} a_{jkl} + \frac{1}{(n-j-1)(n-j-2)}\sum_{k=j+1}^{n}\sum_{l=j+1}^{n} a_{jkl} & k \neq l \\ 0 & k = l \end{cases}$$

The alternative estimator for $\int_{\mathbb{R}^q} \left\|\gamma_j(\tau)\right\|^2 \omega(d\tau)$ is

$$\frac{1}{(n-j)(n-j-3)} \sum_{k=j+1}^{n} \sum_{l=j+1}^{n} \tilde{A}_{jkl} \tilde{B}_{kl}. \tag{10}$$

The following Lemma shows that (10) is an unbiased (bias reduced) estimator for $\int_{\mathbb{R}^q} \left\|\gamma_j(\tau)\right\|^2 \omega(d\tau)$ under some conditions.

Lemma 3.2 *When $\{Y_t\}$ is i.i.d, and $E\left\|Y_t\right\|^3 < \infty$, for $j = 1, ..., n-4$,*

$$E\left(\frac{1}{(n-j)(n-j-3)} \sum_{k=j+1}^{n} \sum_{l=j+1}^{n} \tilde{A}_{jkl} \tilde{B}_{kl}\right) = -E\left[(Y_t - E(Y_t))' \left(Y_t^+ - E(Y_t)\right) \left\|Y_{t-j} - Y_{t-j}^+\right\|\right].$$

Under Assumption 4.1, (10) is less biased than (9).

Proof. *See the Appendix.* ■

Lemma 3.2 demonstrates that (10) is an unbiased estimator for $\int_{\mathbb{R}^q} \left\|\gamma_j(\tau)\right\|^2 \omega(d\tau)$ under the i.i.d condition, which was derived by Park et al. (2015). In relation to time series data, (10) is a bias-reduced estimator, compared to (9).

By employing this estimator, our statistic becomes

$$\widetilde{MD}_n^2 = \sum_{j=1}^{n-4} \frac{1}{(n-j-3)(j\pi)^2} \sum_{k=j+1}^{n} \sum_{l=j+1}^{n} \tilde{A}_{jkl} \tilde{B}_{kl},$$

for $n > 4$. it is expected that $\widetilde{MD}_n^2$ performs better than MD_n^2 in finite samples.

Remark 3.2 *The bias-reduced statistic for testing H_0^{sub} can be constructed analogously.*

4 Asymptotic Theory

In this section, we establish the asymptotic theory for MD_n^2 and $\widetilde{MD}_n^2$. Given that $MD_n^2 = \int_{\mathbb{R}^q} \int_0^1 \left\| \hat{S}_n(\lambda, \tau) \right\|^2 d\lambda \omega(d\tau)$, it is more convenient to establish the null limit distribution of the process $\hat{S}_n(\lambda, \tau)$ in a Hilbert space, as in EV (2006). Let $\eta = (\lambda, \tau) \in \Pi = [0,1] \times \mathbb{R}^q$ and υ the product measure of $\omega(\cdot)$ and the lebesgue measure on $[0,1]$, i.e., $d\upsilon(\eta) = \omega(d\tau) d\lambda$. Then we consider $\hat{S}_n(\lambda, \tau)$ as a random element in the Hilbert space $L_2(\Pi, \upsilon)$ of all square integrable functions (with respect to the measure υ) with the inner product

$$\langle f, g \rangle_H = \int_\Pi f(\eta)' g^c(\eta) d\upsilon(\eta) = \int_\Pi f(\eta)' g^c(\eta) \omega(d\tau) d\lambda.$$

$L_2(\Pi, \upsilon)$ is endowed with the natural Borel σ-field induced by the norm $\|f\|_H = \langle f, f \rangle_H^{1/2}$. If Z is a $L_2(\Pi, \upsilon)$-valued random element and has a probability μ_Z, we say Z has mean m and $E(\langle Z, h \rangle_H) = \langle m, h \rangle_H$ for any $h \in L_2(\Pi, \upsilon)$. If $E\|Z\|_H^2 < \infty$ and Z has zero mean, then the covariance operator of Z (or μ_Z), $C_Z(\cdot)$ say, is a continuous, linear, symmetric positive definite operator from $L_2(\Pi, \upsilon)$ to $L_2(\Pi, \upsilon)$, defined by

$$C_Z(h) = E[\langle Z, h \rangle_H Z].$$

An operator s on a Hilbert space is called nuclear if it can be represented as $s(h) = \sum_{j=1}^\infty l_j \langle h, f_j \rangle_H f_j$, where $\{f_j\}$ is an orthonormal basis of the Hilbert space and $\{l_j\}$ is a real sequence, such that $\sum_{j=1}^\infty |l_j| < \infty$. It is easy to show, see, e.g., Bosq (2000), that the covariance operator $C_Z(\cdot)$ is a nuclear operator, provided that $E\|Z\|_H^2 < \infty$. By introducing these notations, we can write $MD_n^2 = \left\| \hat{S}_n(\lambda, \tau) \right\|_H^2$.

To derive the asymptotic theory, we consider the following assumptions.

Assumption 4.1 $\{Y_t\}$ *is a strictly stationary and ergodic process.* $E\|Y_1\|^3 < \infty$.

Assumption 4.1 is slightly stronger than EV (2006), where they assume a finite second moment for $\{Y_t\}$. In our study, the existence of the third moment for $\{Y_t\}$ is required in constructing the bias-reduced statistic. Since we employ $\omega(\cdot)$, the nonintegrable weighting function in a q-

dimensional Hilbert space, the proofs of the theorems are fundamentally different from EV(2006), although the frameworks seem similar.

Theorem 4.1 *Under Assumption 4.1 and* H_0

$$MD_n^2, \widetilde{MD}_n^2 \xrightarrow{d} MD_\infty^2 := \int_{\mathbb{R}^q} \int_0^1 \|Z(\lambda, \tau)\|^2 \, d\lambda \omega(d\tau),$$

where $Z(\lambda, \tau)$ *is a Gaussian process with zero mean and covariance operator* $\sigma_h^2 = \langle C_S(h), h \rangle_H$*, where*

$$\sigma_h^2 = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} E\left[\int_{\Pi \times \Pi} (Y_t - \mu)' \phi_{t-j}(\tau) \Psi_j(\lambda) h^c(\eta) (Y_t - \mu)' \phi_{t-k}(\tau^*) \Psi_k(\lambda^*) h^c(\eta^*) \, dv(\eta) dv(\eta^*)\right],$$

where $\phi_t(\tau) := \exp(i \langle \tau, Y_t \rangle) - \varphi(\tau)$ *with* $\varphi(\tau) = E[\exp(i \langle \tau, Y_t \rangle)]$*,* $\mu = E(Y_t)$*,* $\Psi_j(\lambda) := \sqrt{2} \sin(j\pi\lambda) / j\pi, \eta = (\lambda, \tau)$ *and* $\eta^* = (\lambda^*, \tau^*)$*.*

Proof. See the Appendix. ■

By analyzing the principal components of MD_∞^2, it is possible to demonstrate that the asymptotic distribution of MD_∞^2 can be expressed as a weighted sum of independent χ_1^2 distributed random variables with weights implicitly determined by the DGP. Therefore, the asymptotic distribution of MD_∞^2 is not standard.

4.1 Consistency

The consistency properties of MD_n^2 and $\widetilde{MD}_n^2$ under alternative H_A are stated in the following theorem.

Theorem 4.2 *Under Assumption 4.1 and* H_A*,*

$$\frac{1}{n} MD_n^2, \frac{1}{n} \widetilde{MD}_n^2 \xrightarrow{p} \sum_{j=1}^{\infty} \frac{\int_{\mathbb{R}^q} \left\|\gamma_j(\tau)\right\|^2 \omega(d\tau)}{(j\pi)^2}.$$

Proof. *See the Appendix.* ■

$$\sum_{j=1}^{\infty} \frac{\int_{\mathbb{R}^q} \left\| \gamma_j(\tau) \right\|^2 \omega(d\tau)}{(j\pi)^2} > 0$$

under H_A, since there exists at least $j \geq 1$ such that $\int_{\mathbb{R}^q} \left\| \gamma_j(\tau) \right\|^2 \omega(d\tau) \neq 0$. In other words, the test is consistent against all block-wise alternatives H_A. Of course, the common criticism on generalized spectral tests also applies here, as they can only detect alternatives which are only a subset of the negation of H_0^*.

To further appreciate the consistency properties of our tests, we consider the following local nonparametric alternatives:

$$H_{A,n} : E(Y_t - \mu | \mathcal{F}_{t-1}) = \frac{g_t}{\sqrt{n}}, \text{a.s.}, \tag{11}$$

where $\mathcal{F}_{t-1}$ is the σ-field generated by I_{t-1}, and the sequence $\{g_t\}$ satisfies the following assumption:

Assumption 4.2 *$\{g_t\}$ is measurable with respect to I_{t-1} ($\mathcal{F}_{t-1}$-measurable), zero mean, strictly stationary, ergodic, and a square integrable sequence, such that there exists a $j \geq 1$ with $E[g_t|Y_{t-j}] \neq 0$ with positive probability.*

Theorem 4.3 *Under the local alternatives (11) and Assumptions 4.1 and 4.2*

$$MD_n^2, \widetilde{MD}_n^2 \xrightarrow{d} MD_{local,\infty}^2 := \int_{\mathbb{R}^q} \int_0^1 \left\| Z(\lambda, \tau) + G((\lambda, \tau)) \right\|^2 d\lambda \omega(d\tau),$$

where

$$G((\lambda, \tau)) = \sum_{j=1}^{\infty} E\left[g_t \exp\left(i \langle \tau, Y_{t-j} \rangle\right)\right] \frac{\sqrt{2} \sin j\pi\lambda}{j\pi}.$$

Proof. See the Appendix. ■

Therefore, $MD_n^2, \widetilde{MD}_n^2$ can detect all block-wise alternatives (11) at the parametric rate $n^{-1/2}$.

5 The Bootstrap Test

As demonstrated in the previous section, the asymptotic distribution of our tests is nonstandard. We can approximate the asymptotic distribution of $MD_n^2, \widetilde{MD}_n^2$ by a bootstrapping method. In particular, we can simulate their critical values by the following algorithm:

1. Calculate the test statistic $MD_n^2, \widetilde{MD}_n^2$ with the original sample $\{Y_t\}_{t=1}^n$.

2. Generate $\{w_t\}_{t=1}^n$, a sequence of independent random variables with zero mean, unit variance, and bounded support, and independent of the sample $\{Y_t\}_{t=1}^n$. Then generate $\{Y_t^*\}_{t=1}^n$, where $Y_t^* = w_t Y_t$.

3. Then, compute

$$MD_n^{*2} = -\sum_{j=1}^{n-1} \frac{1}{(n-j)(j\pi)^2} \sum_{k=j+1}^{n} \sum_{l=j+1}^{n} \left[\left(Y_k^* - \bar{Y}_{n-j}^*\right)^T \left(Y_l^* - \bar{Y}_{n-j}^*\right) \|Y_{k-j} - Y_{l-j}\| \right],$$

where $\bar{Y}_{n-j}^* = \frac{1}{n-j} \sum_{k=j+1}^n Y_k^*$.

$$\widetilde{MD}_n^{*2} = \sum_{j=1}^{n-4} \frac{1}{(n-j-3)(j\pi)^2} \sum_{k=j+1}^{n} \sum_{l=j+1}^{n} \tilde{A}_{jkl} \tilde{B}_{kl}^*,$$

where

$$\widetilde{B}_{kl}^* = \begin{cases} b_{kl}^* - \frac{1}{n-j-2} \sum_{l=j+1}^n b_{kl}^* - \frac{1}{n-j-2} \sum_{k=j+1}^n b_{kl}^* + \frac{1}{(n-j-1)(n-j-2)} \sum_{k=j+1}^n \sum_{l=j+1}^n b_{kl}^* & k \neq l \\ 0 & k = l \end{cases}$$

with $b_{kl}^* = \frac{1}{2} \|Y_k^* - Y_l^*\|^2, k, l = j+1, ..., n$.

4. Repeat steps 2 and 3 B times and compute the empirical $(1-\alpha)$th sample quantile of $MD_n^{*2}, \widetilde{MD}_n^{*2}$ with the B values, $MD_{n,\alpha}^{*2}, \widetilde{MD}_{n,\alpha}^{*2}$, say. The proposed test rejects the null hypothesis at the significance level α if $MD_n^2 > MD_{n,\alpha}^{*2}$, $\widetilde{MD}_n^2 > \widetilde{MD}_{n,\alpha}^{*2}$.

To derive the consistency of the bootstrapping procedure, it is necessary to consider the alternative representation of MD_n^{*2}:

$$MD_n^{*2} = \int_{\mathbb{R}^q} \int_0^1 \left\| \hat{S}_n^*(\eta) \right\|^2 d\lambda \omega\left(d\tau\right),$$

where

$$\hat{S}_n^*(\eta) = \sum_{j=1}^{n-1} (n-j)^{1/2}\, \hat{\gamma}_j^*\left(\tau\right) \frac{\sqrt{2}\sin j\pi\lambda}{j\pi},$$

in which

$$\hat{\gamma}_j^*\left(\tau\right) = \frac{1}{n-j} \sum_{t=j+1}^{n} \left(Y_t w_t - \bar{Y}_{n-j}^*\right) \exp\left(i\left\langle \tau, Y_{t-j}\right\rangle\right).$$

For MD_n^{*2} and $\widetilde{MD}_n^{*2}$, let their distribution, conditional on a sample $\{Y_t\}_1^n$, be $L\left(MD_n^{*2} |\{Y_t\}_1^n\right)$ and $L\left(\widetilde{MD}_n^{*2} |\{Y_t\}_1^n\right)$, then the consistency of the bootstrapping procedure is established in the following theorem:

Theorem 5.1 *Under Assumption 4.1, then under the null hypothesis (4),*

$$\begin{aligned}
\rho_w\left(L\left(MD_n^{*2} |\{Y_t\}_1^n\right), MD_\infty^2\right) &\rightarrow 0 \ a.s. \ as\ n - \infty, \\
\rho_w\left(L\left(\widetilde{MD}_n^{*2} |\{Y_t\}_1^n\right), MD_\infty^2\right) &\rightarrow 0 \ a.s. \ as\ n - \infty.
\end{aligned}$$

Under the local alternatives (11), and Assumptions 4.1, 4.2,

$$\begin{aligned}
\rho_w\left(L\left(MD_n^{*2} |\{Y_t\}_1^n\right), MD_{local,\infty}^2\right) &\rightarrow 0 \ a.s. \ as\ n - \infty, \\
\rho_w\left(L\left(\widetilde{MD}_n^{*2} |\{Y_t\}_1^n\right), MD_{local,\infty}^2\right) &\rightarrow 0 \ a.s. \ as\ n - \infty.
\end{aligned}$$

where ρ_w is any metric that metricizes weak convergence in $L_2(\Pi, v)$, see Politis and Romano (1994).

Proof. *See the Appendix.* ■

Note that, given the result obtained in Theorem 5.1, the proposed bootstrap test has a correct asymptotic level, is consistent, and is able to detect alternatives tending to the null at the parametric rate $n^{-1/2}$.

This procedure is similar to the wild bootstrap used in EV (2006), which is in the spirit of wild bootstrap procedures proposed by Wu (1986), Liu (1988) or Mammen (1993). As in EV (2006), in our Monte Carlo simulations, we employ the $\{w_t\}_{t=1}^n$ sequence of i.i.d Bernoulli variates with

$$\begin{aligned} P\left(w_t = 0.5\left(1-\sqrt{5}\right)\right) &= \left(1+\sqrt{5}\right)/2\sqrt{5} \\ P\left(w_t = 0.5\left(1+\sqrt{5}\right)\right) &= 1-\left(1+\sqrt{5}\right)/2\sqrt{5}. \end{aligned}$$

6 Monte Carlo Evidence

In this section, we examine the finite sample properties of our test statistics, comparing with EV_n^2, the multivariate version of EV (2006). Sample sizes are 100 and 300. The nominal size is 5%. The number of Monte Carlo experiments is $1,000$, while the number of bootstrap replications is $B = 300$. The random sequence in the bootstrap is generated by the Bernoulli variates. In all the replications, 200 pre-sample data values were generated and discarded. We consider two groups of qs: a moderate-value group in which $q = 1,2,3,4,5,6,7,8$, and a high-value group in which $q = 20,40,60,80,100,120,140,160$.

In the following, $a \circ b$ denotes element-wise multiplication of two vectors a, b. We consider the following 6 null hypotheses:

N1. Uncross-correlated i.i.d process: an i.i.d $N(0, I_q)$ distributed q-variate data sequence, where I_q is a $q \times q$ identity matrix.

N2. Cross-correlated i.i.d process: an i.i.d $N(0, \Omega)$ distributed q-variate data sequence, where Ω is a $q \times q$ positive definite matrix.[2]

N3. Cross-correlated i.i.d $t_3(\Omega)$ process: a student's t distributed q-variate data sequence with 3 degrees of freedom, where Ω is a $q \times q$ positive definite matrix.

N4. Stochastic volatility process: $Y_t = \varepsilon_t \circ \exp(\sigma_t)$, with $\sigma_t = 0.936\sigma_{t-1} + 0.32\upsilon_t$, where $\{\varepsilon_t\}$ follows an i.i.d $N(0, \Omega_1)$ and $\{\upsilon_t\}$ follows an i.i.d $N(0, \Omega_2)$ in which Ω_1 and Ω_2 are $q \times q$ positive definite matrices.

[2] To generate a positive definite square matrix, we first generate a square matrix Υ. The positive definite square matrix Ω is then constructed via $\Omega = \Upsilon'\Upsilon$. In the Monte Carlo simulations, values of Ω are predetermined by specifying the seed for the Matlab random number generator. In all the simulations, we set the seed equal to 1, and generate Υ based on a uniformly distributed random numbers generator in the interval $(0,1)$.

N5. Stochastic volatility $+t_3$ process: $Y_t = \varepsilon_t \circ \exp(\sigma_t)$, with $\sigma_t = 0.936\sigma_{t-1} + 0.32\upsilon_t$, where $\{\varepsilon_t\}$ follows an i.i.d $t_3(\Omega_1)$ and $\{\upsilon_t\}$ follows an i.i.d $N(0, \Omega_2)$ in which Ω_1 and Ω_2 are $q \times q$ positive definite matrices.

N6. GARCH process: $Y_t = \varepsilon_t \circ \sigma_t$, where $\sigma_t \circ \sigma_t = 0.1 + 0.05Y_{t-1} \circ Y_{t-1} + 0.9\ \sigma_{t-1} \circ \sigma_{t-1}$ and $\{\varepsilon_t\}$ follows an i.i.d $N(0, \Omega)$ in which Ω is a $q \times q$ positive definite matrix.

N1 is the benchmark. In N2-N6, we allow for cross-correlations between contemporaneous variables. In N3 and N5, there exist heavy tails explicitly, and in N4-N6, there exists heteroskedasticity. The empirical sizes of $\widetilde{MD}_n^2$, MD_n^2, and EV_n^2 for sample sizes 100 and 300 are depicted in Figures 2 and 3, respectively. In each figure, the subplots in the left column report the results of moderate q values; those in the right column report the results of high q values.

It is observed that $\widetilde{MD}_n^2$ has very accurate size control for both moderate and high values of q. When $q = 1$, the empirical sizes of $\widetilde{MD}_n^2$, MD_n^2, and EV_n^2 are comparable; however, when q values increase, the empirical sizes of both MD_n^2 and EV_n^2 deteriorate to zero quickly, especially in the cases involving heavy tails and heteroskedasticity. As sample size increases, the finite-sample properties of MD_n^2 improve, although they remain much worse than those of $\widetilde{MD}_n^2$.

For the power check, we consider the following 6 non-MDSs:

A7. VAR process: $Y_t = 0.2Y_{t-1} + \varepsilon_t$, where $\{\varepsilon_t\}$ follows an i.i.d $N(0, I_q)$.

A8. First order exponential autoregressive process (EXP(1)): $Y_t = 0.6Y_{t-1} + \exp(-0.5Y_{t-1} \circ Y_{t-1}) + \varepsilon_t$, where $\{\varepsilon_t\}$ follows an i.i.d $N(0, I_q)$.

A9. Multiple non-linear moving average process (NLMA): $Y_t = \varepsilon_{t-1} \circ \varepsilon_{t-2} \circ (\varepsilon_{t-2} + \varepsilon_t + 1)$,where $\{\varepsilon_t\}$ follows an i.i.d $N(0, \Omega)$.

A10. Bilinear process: $Y_t = \varepsilon_t + b_1\varepsilon_{t-1} \circ Y_{t-1} + b_2\varepsilon_{t-1} \circ Y_{t-2}$, where $\{\varepsilon_t\}$ follows an i.i.d $N(0, \Omega)$.

A11. The sum of a white noise and the first difference of a stationary autoregressive process of order one (NDAR):

$$\begin{aligned} Y_t &= \varepsilon_t + X_t - X_{t-1} \\ X_t &= 0.85X_{t-1} + \upsilon_t, \end{aligned}$$

where $\{\varepsilon_t\}$ follows an i.i.d $N(0, \Omega_1)$ and $\{\upsilon_t\}$ follows an i.i.d $N(0, \Omega_2)$ in which Ω_1 and Ω_2 are $q \times q$ positive definite matrices.

A12. Threshold autoregressive process of order one (TAR(1)):

$$Y_t = \begin{cases} -0.5Y_{t-1} + \varepsilon_t, & \text{if } \|Y_{t-1}\| > 1 \\ 0.4Y_{t-1} + \varepsilon_t, & \text{otherwise,} \end{cases}$$

where $\{\varepsilon_t\}$ follows an i.i.d $N(0, \Omega)$.

The alternatives A8-A12 in the case of $q = 1$ were considered by EV (2006). The empirical powers of $\widetilde{MD}_n^2$, MD_n^2, and EV_n^2 for sample sizes 100 and 300 are presented in Figures 4 and 5, respectively. In each figure, the subplots in the left column report the results of moderate q values; those in the right column report the results of high q values.

As expected, the testing power of $\widetilde{MD}_n^2$ increases when sample size increases. When q value is large, $\widetilde{MD}_n^2$ maintains its testing power for all the alternatives. On the other hand, EV_n^2 has empirical powers as low as zero in many cases, especially in relation to high q values. For alternatives A7 and A8, where the multivariate time series sequences are uncross-correlated, the testing power of $\widetilde{MD}_n^2$ increases as q value increases. This demonstrates empirically the attractive advantages of the nonintegrable weighting function in high-dimensionality situations.

Overall, these Monte Carlo simulation results reveal that our bias-reduced test maintains excellent finite-sample properties, specially in relation to high-dimensionality situations.

7 Application

In this section, we apply our newly proposed generalized spectral tests to test the EMH on the US stock market. We first consider data sets on monthly average value weighted returns of industry stock portfolios from July 1926 to October 2020. In these data sets, each NYSE, AMEX, and NASDAQ stock is assigned to an industry portfolio at the end of June of year t based on its four-digit SIC code at that time. We consider 5, 10, 17 and 30 industry stock portfolios. The data resource is from the data library maintained by Prof. French.[3]

[3]https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/.

In each of these data sets, we first apply Chang et al.'s (2017) omnibus test to check whether these multivariate return sequences are white noise up to the first 10 lags. We use the pre-transformed data sequence with cross validation on the tuning parameter for estimating the contemporaneous correlations. The number of bootstrap replications is $1,000$.[4] We report its p-values of 5, 10, 17 and 30 industry stock portfolios in Table 1. The results reveal that, for all the monthly returns of industry portfolios considered here, there is no strong evidence that there exists linear dependence.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
5 portfolios	0.052	0.052	0.114	0.107	0.121	0.136	0.128	0.162	0.218	0.193
10 portfolios	0.042	0.045	0.058	0.078	0.100	0.074	0.087	0.138	0.145	0.168
17 portfolios	0.074	0.060	0.067	0.078	0.063	0.074	0.075	0.081	0.148	0.144
30 portfolios	0.068	0.058	0.084	0.072	0.073	0.078	0.074	0.065	0.083	0.123

Table 1: P-values of the omnibus test of Chang et al. (2017) for monthly returns of industry stock portfolios.

We now apply the generalized spectral tests for these industry portfolios to test the weak form of the EMH. For illustrations, denote the null for q industry portfolios as

$$H_0^{(q)} : E\left[Y_t^{(q)} | Y_{t-j}^{(q)}\right] = E\left(Y_t^{(q)}\right), \text{a.s. for all } j = 1, 2, ...$$

where $Y_t^{(q)}$ is a q-dimensional industry portfolio return vector. The EMH corresponds to the case when q equals the number of stocks in the US stock market. By considering the industry portfolios here, we test the dimension-reduced null hypotheses. By the structure of these portfolios, when $p < q$, $H_0^{(q)}$ implies $H_0^{(p)}$. In other words, a rejection of $H_0^{(p)}$ implies a rejection of $H_0^{(q)}$.

The number of bootstrap replications is also $1,000$. The p-values of the generalized spectral tests are reported in Table 2. Some interesting patterns appear. For the monthly returns of the 5 and 10 industry portfolios, our tests can not reject the nulls at the 5% significance level. However, for the monthly returns of the 17 and 30 industry portfolios, the nulls are rejected by our bias-reduced statistic at the 5% significance level. On the other hand, the EV test cannot reject the nulls at the 5% significance level in all the cases.

[4]We use the function **wntest** in the R package HDtest with opt=3.

	5 portfolios	10 portfolios	17 portfolios	30 portfolios
BMD	0.12	0.09	0.04	0.01
MD	0.16	0.11	0.09	0.04
EV	0.60	0.60	0.58	0.58

Table 2: P-values of generalized spectral tests. BMD represents the bias-reduced statistic $\widetilde{MD}_n^2$; MD represents the statistic MD_n^2; EV represents the statistic EV_n^2.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
5 portfolios	0.107	0.131	0.132	0.135	0.163	0.155	0.085	0.087	0.102	0.113
10 portfolios	0.117	0.148	0.197	0.210	0.201	0.203	0.115	0.116	0.135	0.127
17 portfolios	0.158	0.181	0.243	0.260	0.305	0.299	0.136	0.118	0.133	0.117
30 portfolios	0.214	0.242	0.327	0.365	0.371	0.380	0.128	0.168	0.189	0.166
49 portfolios	0.278	0.270	0.332	0.363	0.375	0.400	0.238	0.232	0.228	0.242

Table 3: P-values of the omnibus test of Chang et al. (2017) for daily returns of industry stock portfolios.

We then consider data sets on daily average value weighted returns of stock portfolios from Dec 1, 2016 to Nov 30, 2020. We consider 5, 10, 17, 30 and 49 industry stock portfolios. The p-values from Chang et al.'s (2017) omnibus test and the generalized spectral tests are reported in Tables 3 and 4, respectively. It is observed that, up to the 49 industry portfolios, there is no strong evidence against the EMH.

At a monthly frequency, and not at a daily frequency, the EMH of the US stock market is fully rejected by our test. These results echo a large body of literature on stock predictability, wherein it is argued that the US stock returns are predictable for longer horizons. Especially, Hong et al. (2007) found that the returns on industry portfolios predicted stock market movements at a monthly frequency. Another implication from the results is that nonlinear modeling of the stock returns may offer better predictions.

	5 portfolios	10 portfolios	17 portfolios	30 portfolios	49 portfolios
BMD	0.052	0.100	0.130	0.169	0.213
MD	0.064	0.122	0.160	0.194	0.236
EV	0.652	0.575	0.482	0.485	0.481

Table 4: P-values of generalized spectral tests. BMD represents the bias-reduced statistic $\widetilde{MD}_n^2$; MD represents the statistic MD_n^2; EV represents the statistic EV_n^2.

8 Conclusions

In this study, we propose new generalized spectral tests for multivariate MDHs, which especially suit high-dimensionality situations. Essentially, we introduce a unique, nonintegrable weighting function that is pivotal in generalized spectral tests. We demonstrate rigorously that this weighting function enjoys attractive advantages in high dimensionality situations. Simulation results reveal that the newly proposed bias-reduced generalized spectral test statistic outperforms its competitors substantially. The application to testing the EMH on the US stock market demonstrates the usefulness of our proposal. It is of interest to extend our approach to test the correct specification for multivariate parametric time series models, which in many cases implies a multivariate MDS on the unobservable model errors. In this scenario, however, the parameter uncertainty must be considered, which will result in more complicated asymptotic theory for the test statistic and require a new bootstrapping procedure. We will report this important extension in a separate work.

9 Appendix of Proofs

Proof of Lemma 3.1. For any $x \in \mathbb{R}$,

$$\begin{aligned} \frac{d \int_0^\infty \frac{1-\cos(xs)}{s^2} ds}{dx} &= \int_0^\infty \frac{\sin(xs)}{s} ds \\ &= \frac{\pi}{2} sgn(x), \end{aligned}$$

where $sgn(x)$ denotes the sign of x. So

$$\int_0^\infty \frac{1-\cos(xs)}{s^2} ds = \frac{\pi}{2} |x|. \tag{12}$$

By 3.3.2.3, P586, Prudnikov et al. (1986) and applying (12), we have

$$\begin{aligned}
\int_{\mathbb{R}^q} \frac{1-\cos\left(\langle\tau, X\rangle\right)}{\|\tau\|^{q+1}} d\tau &= \frac{2\pi^{(q-1)/2}}{\Gamma\left(\frac{q-1}{2}\right)} \int_0^{\pi} \int_0^{\infty} \frac{1-\cos\left(\|X\| s\cos u\right)}{s^2} ds \sin^{q-2}(u)\, du \\
&= \|X\| \frac{2\pi^{(q+1)/2}}{\Gamma\left(\frac{q-1}{2}\right)} \int_0^{\pi} |\cos u| \sin^{q-2}(u)\, du \\
&= \|X\| \frac{\pi^{(q+1)/2}}{\Gamma\left(\frac{q+1}{2}\right)}.
\end{aligned}$$

$$\begin{aligned}
\int_{\mathbb{R}^q} \frac{\sin\left(\langle\tau, X\rangle\right)}{\|\tau\|^{q+1}} d\tau &= \frac{2\pi^{(q-1)/2}}{\Gamma\left(\frac{q-1}{2}\right)} \int_0^{\pi} \int_0^{\infty} \frac{\sin\left(\|X\| s\cos u\right)}{s^2} ds \sin^{q-2}(u)\, du \\
&= \|X\| \frac{2\pi^{(q-1)/2}}{\Gamma\left(\frac{q-1}{2}\right)} \int_0^{\infty} \frac{\sin(s)}{s^2} \int_0^{\pi} \cos u \sin^{q-2}(u)\, du ds \\
&= 0.
\end{aligned}$$

So the proof is complete. ■

Proof of Lemma 3.2. Firstly, let E_{Y_t} denote the expectation with respect to Y_t, $E_{Y_t^+}$ follows similarly. We have

$$\begin{aligned}
& \frac{1}{2}\left\|Y_t - Y_t^+\right\|^2 - E_{Y_t}\left(\frac{1}{2}\left\|Y_t - Y_t^+\right\|^2\right) - E_{Y_t^+}\left(\frac{1}{2}\left\|Y_t - Y_t^+\right\|^2\right) + E\left(\frac{1}{2}\left\|Y_t - Y_t^+\right\|^2\right) \\
= \; & \frac{1}{2}\left(Y_t - Y_t^+\right)'\left(Y_t - Y_t^+\right) - \frac{1}{2}\int_{\mathbb{R}^q} \left(y_t - Y_t^+\right)'\left(y_t - Y_t^+\right) dF_{Y_t}(y_t) \\
& -\frac{1}{2}\int_{\mathbb{R}^q} \left(Y_t - y_t^+\right)'\left(Y_t - y_t^+\right) dF_{Y_t^+}\left(y_t^+\right) \\
& + \int_{\mathbb{R}^q}\int_{\mathbb{R}^q} \frac{1}{2}\left(y_t - y_t^+\right)'\left(y_t - y_t^+\right) dF_{Y_t^+}\left(y_t^+\right) dF_{Y_t}(y_t) \\
= \; & \frac{1}{2}\left[-2Y_t'Y_t^+ + 2E\left(Y_t\right)'Y_t^+ + 2Y_t'E\left(Y_t\right) - 2E\left(Y_t\right)'EY_t\right] \\
= \; & -\left(Y_t - E\left(Y_t\right)\right)'\left(Y_t^+ - E\left(Y_t\right)\right).
\end{aligned}$$

Then

$$
\begin{aligned}
\int_{\mathbb{R}^q} \left\|\gamma_j(\tau)\right\|^2 \omega(d\tau) &= -E\left[\left(Y_t - E\left(Y_t\right)\right)'\left(Y_t^+ - E\left(Y_t\right)\right)\left\|Y_{t-j} - Y_{t-j}^+\right\|\right] \\
&= E\left[\left[\begin{array}{c} \frac{1}{2}\left\|Y_t - Y_t^+\right\|^2 - E_{Y_t}\left(\frac{1}{2}\left\|Y_t - Y_t^+\right\|^2\right) \\ -E_{Y_t^+}\left(\frac{1}{2}\left\|Y_t - Y_t^+\right\|^2\right) + E\left(\frac{1}{2}\left\|Y_t - Y_t^+\right\|^2\right) \end{array}\right]\left\|Y_{t-j} - Y_{t-j}^+\right\|\right] \\
&= \gamma_j + \alpha_j\beta - 2\delta_j,
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha_j &: = E\left\|Y_{t-j} - Y_{t-j}^+\right\|, \beta := E\left(\frac{1}{2}\left\|Y_t - Y_t^+\right\|^2\right), k \neq l, \\
\delta_j &: = E\left(\left\|Y_{t-j} - Y_{t-j}^+\right\|\left(\frac{1}{2}\left\|Y_t - Y_t^{++}\right\|^2\right)\right), k, l, m \text{ distinct}, \\
\gamma_j &: = E\left(\left\|Y_{t-j} - Y_{t-j}^+\right\|\left(\frac{1}{2}\left\|Y_t - Y_t^+\right\|^2\right)\right), k \neq l
\end{aligned}
$$

with $(Y_t, Y_{t-j}), \left(Y_t^+, Y_{t-j}^+\right)$ and $\left(Y_t^{++}, Y_{t-j}^{++}\right)$ are i.i.d. Note that the existence of δ_j and γ_j requires the existence of the third moment for $\{Y_t\}$, since by Hölder's inequality and Minkowski's Inequality

$$
\begin{aligned}
\delta_j &\leq \frac{1}{2}\left[E\left(\left\|Y_t - Y_t^{++}\right\|^3\right)\right]^{2/3}\left[E\left(\left\|Y_{t-j} - Y_{t-j}^+\right\|^3\right)\right]^{1/3} \\
&\leq \frac{1}{2}4\left(E\left\|Y_t\right\|^3\right)^{2/3} 2\left(E\left\|Y_t\right\|^3\right)^{1/3} \\
&\leq 4E\left\|Y_t\right\|^3.
\end{aligned}
$$

Define

$$
T_{j1} = \sum_{k,l=j+1} a_{jkl} b_{kl}, T_{j2} = a_{j..} b_{..}, \; T_{j3} = \sum_{k=j+1}^{n} a_{jk.} b_{k.},
$$

where

$$
a_{j..} = \sum_{k,l=j+1}^{n} a_{jkl}, b_{..} = \sum_{k,l=j+1}^{n} b_{kl}, a_{jk.} = \sum_{l=j+1}^{n} a_{jkl}, b_{k.} = \sum_{l=j+1}^{n} b_{kl}.
$$

It is not difficult to obtain

$$\sum_{k=j+1}^{n}\sum_{l=j+1}^{n}\tilde{A}_{jkl}\tilde{B}_{kl}=T_{j1}+\frac{T_{j2}}{(n-j-1)(n-j-2)}-\frac{2T_{j3}}{n-j-2}.$$

When $\{Y_t\}$ is i.i.d, $E\|Y_t\|^3<\infty$,

$$E(T_{j1})=(n-j)(n-j-1)\gamma_j,$$

$$E(T_{j2})=(n-j)(n-j-1)\left[(n-j-2)(n-j-3)\alpha_j\beta+2\gamma_j+4(n-j-2)\delta_j\right],$$

$$E(T_{j3})=(n-j)(n-j-1)\left[(n-j-2)\delta_j+\gamma_j\right].$$

To get some idea about the results above, note, for example, in T_{j1}, for $k,l=j+1,...,n$,

$$E(a_{jkl}b_{kl}):=\begin{cases} E\left(\frac{1}{2}\|Y_k-Y_l\|^2\|Y_{k-j}-Y_{l-j}\|\right), & k\neq l\\ 0, & k=l.\end{cases}$$

Since $(Y_k,\, Y_{k-j})$ and $(Y_l,\, Y_{l-j})$ are i.i.d for $k\neq l$, so $E(a_{jkl}b_{kl})=\gamma_j$, for $k\neq l$. The arguments for T_{j2} and T_{j3} are similar. Therefore, after some calculations,

$$E\left(\sum_{k=j+1}^{n}\sum_{l=j+1}^{n}A_{jkl}B_{kl}\right)=(n-j)(n-j-3)\left(\gamma_j+\alpha_j\beta-2\delta_j\right).$$

If $\{Y_t\}$ is a strictly stationary and ergodic, $E\|Y_1\|^3<\infty$,

$$E(T_{j1})=(n-j)(n-j-1)\gamma_j+o\left((n-j)^2\right).$$

To get some idea about this result, for T_{j1}, denote $\gamma_j(s)=E\left(\frac{1}{2}\|Y_{j+1}-Y_{j+1+s}\|^2\|Y_1-Y_{1+s}\|\right)$, for $s=1,...,n-j-1$.

$$\begin{aligned} E(T_{j1}) &= 2\sum_{m=1}^{n-j}\sum_{s=1}^{m}\gamma_j(s)=(n-j)(n-j-1)\gamma_j+2\sum_{m=1}^{n-j-1}\sum_{s=1}^{m}\left[\gamma_j(s)-\gamma_j\right]\\ &= (n-j)(n-j-1)\gamma_j\\ &\quad +(n-j)(n-j-1)\sum_{m=1}^{n-j-1}\frac{m}{(n-j)(n-j-1)}\frac{1}{m}\sum_{s=1}^{m}\left[\gamma_j(s)-\gamma_j\right].\end{aligned}$$

Note that, given Y_{j+1},Y_1, $\left\{\frac{1}{2}\left\|Y_{j+1}-Y_{j+1+s}\right\|^2\left\|Y_1-Y_{1+s}\right\|\right\}$ is ergodic, so by the ergodic theorem,

$$\frac{1}{m}\sum_{s=1}^{m}\frac{1}{2}\left\|Y_{j+1}-Y_{j+1+s}\right\|^2\left\|Y_1-Y_{1+s}\right\|\overset{a.s}{\to}E\left(\frac{1}{2}\left\|Y_{j+1}-Y_{j+1}^{+}\right\|^2\left\|Y_1-Y_1^{+}\right\||Y_{j+1},Y_1\right), \text{ as } m\to\infty.$$

Then by monotone convergence theorem,

$$\frac{1}{m}\sum_{s=1}^{m}E\left(\frac{1}{2}\left\|Y_{j+1}-Y_{j+1+s}\right\|^2\left\|Y_1-Y_{1+s}\right\|\right)\to E\left(\frac{1}{2}\left\|Y_{j+1}-Y_{j+1}^{+}\right\|^2\left\|Y_1-Y_1^{+}\right\|\right), \text{ as } m\to\infty.$$

By the Toeplitz Lemma,

$$\sum_{m=1}^{n-j-1}\frac{m}{(n-j)(n-j-1)}\frac{1}{m}\sum_{s=1}^{m}\left[\gamma_j(s)-\gamma_j\right]=o(1).$$

Similarly,

$$E(T_{j2})=(n-j)(n-j-1)\left[(n-j-2)(n-j-3)\alpha_j\beta+2\gamma_j+4(n-j-2)\delta_j\right]+o\left((n-j)^4\right), \tag{13}$$

$$E(T_{j3})=(n-j)(n-j-1)\left[(n-j-2)\delta_j+\gamma_j\right]+o\left((n-j)^3\right). \tag{14}$$

Now

$$
\begin{aligned}
\sum_{k=j+1}^{n}\sum_{l=j+1}^{n} A_{jkl}B_{kl} &= \left\{\begin{array}{cccc}
T_{j1} & -\frac{T_{j3}}{n} & -\frac{T_{j3}}{n} & +\frac{T_{j2}}{n^2} \\
-\frac{T_{j3}}{n} & +\frac{T_{j3}}{n} & +\frac{T_{j2}-T_{j3}}{n^2} & -\frac{T_{j2}}{n^2} \\
-\frac{T_{j3}}{n} & +\frac{T_{j2}-T_{j3}}{n^2} & +\frac{T_{j3}}{n} & -\frac{T_{j2}}{n^2} \\
+\frac{T_{j2}}{n^2} & -\frac{T_{j2}}{n^2} & -\frac{T_{j2}}{n^2} & +\frac{T_{j2}}{n^2}
\end{array}\right\} \\
&= T_{j1} + \frac{T_{j2}}{(n-j)^2} - \frac{2(n-j+1)}{(n-j)^2}T_{j3} \\
&= \sum_{k=j+1}^{n}\sum_{l=j+1}^{n} \tilde{A}_{jkl}\tilde{B}_{kl} \\
&\quad + \frac{T_{j2}}{(n-j)^2} - \frac{T_{j2}}{(n-j-1)(n-j-2)} \\
&\quad - \left(\frac{2(n-j+1)}{(n-j)^2}T_{j3} - \frac{2T_{j3}}{n-j-2}\right) \\
&= \sum_{k=j+1}^{n}\sum_{l=j+1}^{n} \tilde{A}_{jkl}\tilde{B}_{kl} \\
&\quad + \frac{-3(n-j)+2}{(n-j)^2(n-j-1)(n-j-2)}T_{j2} \\
&\quad -2\frac{-(n-j)-2}{(n-j)^2(n-j-2)}T_{j3}
\end{aligned}
$$

$$
\begin{aligned}
E\left(\frac{1}{(n-j)^2}\sum_{k=j+1}^{n}\sum_{l=j+1}^{n} A_{jkl}B_{kl}\right) &= E\left(\frac{1}{(n-j)(n-j-3)}\sum_{k=j+1}^{n}\sum_{l=j+1}^{n} \tilde{A}_{jkl}\tilde{B}_{kl}\right) \\
&\quad - \frac{3\alpha_j\beta - 2\delta_j}{n-j} + o\left((n-j)^{-1}\right).
\end{aligned}
$$

So (9) is more biased than (10). ■

In order to prove Theorem 4.1, we introduce several Lemmas firstly.

Lemma 9.1 *Under Assumption 4.1*

$$
E\int_{\mathbb{R}^q} |\phi_t(Y_t,\tau)|^2 \omega(d\tau) = E\left\|Y_t - Y_t^+\right\|, \tag{15}
$$

$$
E\int_{\mathbb{R}^q} \phi_t(Y_t,\tau)\phi_s(Y_s,\tau)^c \omega(d\tau) = E\left\|Y_t - Y_s\right\| - E\left\|Y_t - Y_t^+\right\|, \tag{16}
$$

for any $t \neq s,\ t, s = 1, ..., n.$

Proof.

$$\begin{aligned}
E \int_{\mathbb{R}^q} |\phi_t(Y_t, \tau)|^2 \omega(d\tau) &= E \int_{\mathbb{R}^q} \exp(i \langle \tau, Y_t \rangle) - \varphi(Y_t, \tau) (\exp(-i \langle \tau, Y_t \rangle) - \varphi(Y_{t-j}, \tau)^c) \omega(d\tau) \\
&= E \int_{\mathbb{R}^q} 1 - E[\exp(i \langle \tau, Y_t \rangle)] \exp(-i \langle \tau, Y_t \rangle) + \\
&\quad 1 - E[\exp(-i \langle \tau, Y_t \rangle)] \exp(i \langle \tau, Y_t \rangle) \\
&\quad - (1 - E[\exp(i \langle \tau, Y_t \rangle)] E[\exp(-i \langle \tau, Y_t \rangle)]) \omega(d\tau) \\
&= E\left[2 E_{Y_t} \left\| Y_t^+ - Y_t \right\| - E \left\| Y_t - Y_t^+ \right\|\right] \\
&= E \left\| Y_t - Y_t^+ \right\|.
\end{aligned}$$

$$\begin{aligned}
E \int_{\mathbb{R}^q} \phi_t(Y_t, \tau) \phi_s(Y_s, \tau)^c \omega(d\tau) &= E \int_{\mathbb{R}^q} \exp(i \langle \tau, Y_t \rangle) - \varphi(Y_t, \tau) (\exp(-i \langle \tau, Y_s \rangle) - \varphi(Y_s, \tau)^c) \omega(d\tau) \\
&= E \int_{\mathbb{R}^q} \exp(i \langle \tau, Y_t - Y_s \rangle) - E[\exp(i \langle \tau, Y_t \rangle)] \exp(-i \langle \tau, Y_s \rangle) \\
&\quad - E[\exp(-i \langle \tau, Y_s \rangle)] \exp(i \langle \tau, Y_t \rangle) \\
&\quad + E[\exp(i \langle \tau, Y_t \rangle)] E[\exp(-i \langle \tau, Y_t \rangle)] \omega(d\tau) \\
&= E \int_{\mathbb{R}^q} - (1 - \exp(i \langle \tau, Y_t - Y_s \rangle)) + 1 - E_{Y_t}[\exp(i \langle \tau, Y_t - Y_s \rangle)] \\
&\quad + 1 - E_{Y_s}[\exp(i \langle \tau, Y_t - Y_s \rangle)] \\
&\quad - \left(1 - E\left[\exp\left(i \left\langle \tau, Y_t - Y_t^+ \right\rangle\right)\right]\right) \omega(d\tau) \\
&= E\left[- \|Y_t - Y_s\| + E_{Y_t} \|Y_t - Y_s\| + E_{Y_s} \|Y_t - Y_s\| - E \left\| Y_t - Y_t^+ \right\|\right] \\
&= E \|Y_t - Y_s\| - E \left\| Y_t - Y_t^+ \right\|.
\end{aligned}$$

In above derivations, we repeatedly apply the result in Lemma 3.1. ■

Define

$$\hat{Z}_n(\lambda, \tau) = \sum_{j=1}^{n-1} (n-j)^{1/2} \hat{r}_j(\tau) \frac{\sqrt{2} \sin j\pi\lambda}{j\pi},$$

where

$$\hat{r}_j(\tau) = \frac{1}{n-j} \sum_{t=j+1}^{n} (Y_t - \mu) \phi_{t-j}(Y_{t-j}, \tau).$$

Lemma 9.2 *Under Assumption 4.1 and* H_0^*

$$\left\|\hat{S}_n(\lambda,\tau)\right\|_H^2 = \left\|\hat{Z}_n(\lambda,\tau)\right\|_H^2 + o_p(1).$$

Proof. Note

$$\begin{aligned}
\hat{\gamma}_j(\tau) &= \frac{1}{n-j}\sum_{t=j+1}^{n}\left(Y_t-\bar{Y}_{n-j}\right)\exp\left(i\langle\tau,Y_{t-j}\rangle\right) \\
&= \frac{1}{n-j}\sum_{t=j+1}^{n}\left(Y_t-\mu+\mu-\bar{Y}_{n-j}\right)\left[\phi_{t-j}(Y_{t-j},\tau)+E\left[\exp\left(i\langle\tau,Y_{t-j}\rangle\right)\right]\right] \\
&= \frac{1}{n-j}\sum_{t=j+1}^{n}\left(Y_t-\mu\right)\phi_{t-j}(Y_{t-j},\tau) \\
&\quad -\frac{1}{n-j}\sum_{t=j+1}^{n}\left(Y_t-\mu\right)\frac{1}{n-j}\sum_{t=j+1}^{n}\phi_{t-j}(Y_{t-j},\tau)
\end{aligned}$$

Then

$$\hat{S}_n(\lambda,\tau) = \hat{Z}_n(\lambda,\tau) - \hat{R}_n(\lambda,\tau),$$

where

$$\hat{R}_n(\lambda,\tau) = \sum_{j=1}^{n-1}(n-j)^{1/2}\frac{1}{n-j}\sum_{t=j+1}^{n}\left(Y_t-\mu\right)\frac{1}{n-j}\sum_{t=j+1}^{n}\phi_{t-j}(Y_{t-j},\tau)\frac{\sqrt{2}\sin j\pi\lambda}{j\pi}.$$

Then

$$\left\|\hat{S}_n(\lambda,\tau)\right\|_H^2 = \left\|\hat{Z}_n(\lambda,\tau)\right\|_H^2 + \left\|\hat{R}_n(\lambda,\tau)\right\|_H^2 - 2\,\mathrm{Re}\left(\left\langle\hat{S}_n(\lambda,\tau),\hat{R}_n(\lambda,\tau)\right\rangle_H\right),$$

where $\mathrm{Re}(A)$ denotes the real part of the complex number A and

$$\begin{aligned}
\left\|\hat{R}_n(\lambda,\tau)\right\|_H^2 &= \sum_{j=1}^{n-1}\frac{(n-j)}{(\pi j)^2}\left(\frac{1}{n-j}\sum_{t=j+1}^{n}\left(Y_t-\mu\right)\right)' \\
&\quad \times\frac{1}{n-j}\sum_{t=j+1}^{n}\left(Y_t-\mu\right)\int_{\mathbb{R}^q}\left|\frac{1}{n-j}\sum_{t=j+1}^{n}\phi_{t-j}(Y_{t-j},\tau)\right|^2\omega(d\tau).
\end{aligned}$$

By a strictly stationary, ergodic MDS central limit theorem, it is easy to get

$$(n-j)^{-1/2}\sum_{t=j+1}^{n}(Y_t-\mu)=O_p(1).$$

On the other hand,

$$\begin{aligned}
E\int_{\mathbb{R}^q}\left|\frac{1}{n-j}\sum_{t=j+1}^{n}\phi_{t-j}(Y_{t-j},\tau)\right|^2\omega(d\tau) &= \int_{\mathbb{R}^q}\frac{1}{(n-j)^2}\sum_{t=j+1}^{n}\sum_{s=j+1}^{n}E\left[\phi_{t-j}(Y_{t-j},\tau)\phi_{s-j}(Y_{s-j},\tau)^c\right]\omega(d\tau)\\
&= \frac{1}{(n-j)^2}\sum_{t=j+1}^{n}E\int_{\mathbb{R}^q}|\phi_t(Y_t,\tau)|^2\omega(d\tau)\\
&\quad+\frac{2}{(n-j)^2}\sum_{t,s=j+1,t<s}^{n}E\int_{\mathbb{R}^q}\left[\phi_{t-j}(Y_{t-j},\tau)\phi_{s-j}(Y_{s-j},\tau)^c\right]\omega(d\tau)\\
&= \frac{1}{(n-j)}E\left\|Y_t-Y_t^+\right\|\\
&\quad+\frac{2}{(n-j)}\sum_{m=1}^{n-j-1}\left(1-\frac{m}{n-j}\right)\left(E\left\|Y_t-Y_{t+m}\right\|-E\left\|Y_t-Y_t^+\right\|\right)
\end{aligned}$$

By the ergodicity condition and the Toeplitz Lemma, for a fixed j,

$$\frac{2}{(n-j)}\sum_{m=1}^{n-j-1}\left(1-\frac{m}{n-j}\right)\left(E\left\|Y_t-Y_{t+m}\right\|-E\left\|Y_t-Y_t^+\right\|\right)\to 0,$$

as $n\to\infty$. Then we conclude that

$$E\left(\int_{\mathbb{R}^q}\left|\frac{1}{n-j}\sum_{t=j+1}^{n}\phi_{t-j}(Y_{t-j},\tau)\right|^2\omega(d\tau)\right)=o(1).$$

Therefore

$$\int\left|\frac{1}{n-j}\sum_{t=j+1}^{n}\phi_{t-j}(Y_{t-j},\tau)\right|^2\omega(d\tau)=o_p(1).$$

So

$$\left\|\hat{R}_n(\lambda,\tau)\right\|_H^2=o_p(1)\sum_{j=1}^{n-1}\frac{1}{(\pi j)^2}=o_p(1).$$

$$\left\|\hat{R}_n(\lambda,\tau)\right\|_H=o_p(1).$$

On the other hand,

$$E\left\|\hat{Z}_n\left(\lambda,\tau\right)\right\|_H^2=\sum_{j=1}^{n-1}\frac{(n-j)}{(j\pi)^2}E\int_{\mathbb{R}^q}\left\|\hat{r}_j\left(\tau\right)\right\|^2\omega\left(d\tau\right).$$

$$\begin{aligned}
E\int_{\mathbb{R}^q}\left\|\hat{r}_j\left(\tau\right)\right\|^2\omega\left(d\tau\right) &= E\left(\int_{\mathbb{R}^q}\left|\frac{1}{n-j}\sum_{t=j+1}^{n}\left(Y_t-\mu\right)\phi_{t-j}(Y_{t-j},\tau)\right|^2\omega\left(d\tau\right)\right)\\
&= E\int_{\mathbb{R}^q}\frac{1}{(n-j)^2}\sum_{t=j+1}^{n}\left(Y_t-\mu\right)'\phi_{t-j}(Y_{t-j},\tau)\sum_{t=j+1}^{n}\left(Y_t-\mu\right)\phi_{t-j}(Y_{t-j},\tau)^c\omega\left(d\tau\right)\\
&= \int_{\mathbb{R}^q}\frac{1}{(n-j)^2}\sum_{t=j+1}^{n}E\left|\left(Y_t-\mu\right)\phi_{t-j}(Y_{t-j},\tau)\right|^2\omega\left(d\tau\right)\\
&\quad+\int_{\mathbb{R}^q}\frac{1}{(n-j)^2}\sum_{t,s=j+1,t\neq s,}^{n}E\left(\left(Y_t-\mu\right)'\left(Y_s-\mu\right)\phi_{t-j}(Y_{t-j},\tau)\phi_{s-j}(Y_{s-j},\tau)^c\right)\omega\left(d\tau\right)\\
&= \frac{1}{n-j}E\int_{\mathbb{R}^q}\left|\left(Y_t-\mu\right)\phi_{t-j}(Y_{t-j},\tau)\right|^2\omega\left(d\tau\right)\\
&= \frac{1}{n-j}E\left[\left(Y_t-\mu\right)'\left(Y_t-\mu\right)\left(2E_{Y_{t-j}}\left\|Y_{t-j}-Y_{t-j}^+\right\|-E\left\|Y_{t-j}-Y_{t-j}^+\right\|\right)\right]\\
&= \frac{1}{n-j}E\left[\left(Y_t-\mu\right)'\left(Y_t-\mu\right)\right]\left(E\left\|Y_{t-j}-Y_{t-j}^+\right\|\right).
\end{aligned}$$

Since, under Assumption (4.1) and H_0^*, by the the law of iterated expectation

$$E\left(\left(Y_t-\mu\right)'\left(Y_s-\mu\right)\phi_{t-j}(Y_{t-j},\tau)\phi_{s-j}(Y_{s-j},\tau)^c\right)=0,$$

for any $t\neq s$. Then

$$E\int_{\mathbb{R}^q}\left\|\hat{r}_j\left(\tau\right)\right\|^2\omega\left(d\tau\right)=O\left((n-j)^{-1}\right).$$

So

$$\begin{aligned}
E\left\|\hat{Z}_n\left(\lambda,\tau\right)\right\|_H^2 &= \sum_{j=1}^{n-1}\frac{(n-j)}{(j\pi)^2}O\left((n-j)^{-1}\right)\\
&= O\left(1\right).
\end{aligned}$$

So

$$\left\|\hat{Z}_n(\lambda,\tau)\right\|_H = O_p(1).$$

By Cauchy-Schwartz inequality

$$\begin{aligned}\left\langle \hat{Z}_n(\lambda,\tau), \hat{R}_n(\lambda,\tau)\right\rangle_H &\leq \left\|\hat{R}_n(\lambda,\tau)\right\|_H \left\|\hat{Z}_n(\lambda,\tau)\right\|_H \\ &= o_p(1)\,O_p(1) \\ &= o_p(1).\end{aligned}$$

Then

$$\left\|\hat{S}_n(\lambda,\tau)\right\|_H^2 = \left\|\hat{Z}_n(\lambda,\tau)\right\|_H^2 + o_p(1).$$

■

Lemma 9.3 *Let $\Longrightarrow$ denote weak convergence in the Hilbert space $L_2(\Pi, \upsilon)$ endowed with the norm metric. Under Assumption 4.1 and H_0,*

$$\hat{Z}_n(\lambda,\tau) \Rightarrow Z(\lambda,\tau)$$

on the Hilbert space $L_2(\Pi,\upsilon)$, where $Z(\lambda,\tau)$ is a Gaussian process with zero mean and covariance operator $\sigma_h^2 = \langle C_S(h), h\rangle_H$, where

$$\sigma_h^2 = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} E\left[\int_{\Pi\times\Pi} (Y_t-\mu)'\,\phi_{t-j}(\tau)\Psi_j(\lambda)h^c(\eta)\,(Y_t-\mu)'\,\phi_{t-k}(\tau^*)\,\Psi_k(\lambda^*)\,h^c(\eta^*)\,dv(\eta)dv(\eta^*)\right].$$

Proof. The proof is similar to the proof of Theorem 1 in EV (2006). We only give some sketches here. By a similar decomposition as (28) in EV (2006),

$$\hat{Z}_n(\eta) = \hat{Z}_n^K(\eta) + \hat{R}_n^K(\eta).$$

It can be shown that, for any $h \in L_2(\Pi,\upsilon)$, in $\left\langle \hat{Z}_n^K(\eta), h\right\rangle$ converges to $\left\langle Z^K(\eta), h\right\rangle$, where

$Z^K(\eta)$ is a Gaussian process with zero mean and asymptotic variances

$$\sigma_{h,K}^2 = \sum_{j=1}^{K}\sum_{k=1}^{K} E\left[\int_{\Pi\times\Pi} (Y_t-\mu)'\,\phi_{t-j}(\tau)\Psi_j(\lambda)h^c(\eta)\,(Y_t-\mu)'\,\phi_{t-k}(\tau^*)\,\Psi_k(\lambda^*)\,h^c(\eta^*)\,dv(\eta)dv(\eta^*)\right].$$

Moreover, the sequence $\left\{\hat{Z}_n^K(\eta)\right\}$ is tight. Finally, we explicitly use the result that $\int_{\mathbb{R}^q} |\phi_t(Y_t,\tau)|^2\,\omega(d\tau) = O_p(1)$ to show that, for any $\varepsilon > 0$,

$$\lim_{K\to\infty}\lim_{n\to\infty} P\left(\left\|\hat{R}_n^K(\eta)\right\| > \varepsilon\right) = 0.$$

Then the conclusion follows. ■

Lemma 9.4 *Under Assumption 4.1 and* H_0,

$$MD_n^2 = \widetilde{MD}_n^2 + o_p(1).$$

Proof.

$$\begin{aligned}
MD_n^2 - \widetilde{MD}_n^2 &= \sum_{j=1}^{n-1}\frac{1}{(n-j)(j\pi)^2}\sum_{k=j+1}^{n}\sum_{l=j+1}^{n}A_{jkl}B_{kl} - \sum_{j=1}^{n-4}\frac{1}{(n-j-3)(j\pi)^2}\sum_{k=j+1}^{n}\sum_{l=j+1}^{n}\tilde{A}_{jkl}\tilde{B}_{kl} \\
&= \sum_{j=1}^{n-4}\frac{3}{(n-j-3)(n-j)(j\pi)^2}\left(-\frac{(3(n-j)+2)\,T_{j2}}{(n-j)^2(n-j-1)(n-j-2)} + \frac{(2(n-j)+4)\,T_{j3}}{(n-j)^2(n-j-2)}\right) \\
&\quad + \sum_{j=n-3}^{n-1}\frac{1}{(n-j)(j\pi)^2}\sum_{k=j+1}^{n}\sum_{l=j+1}^{n}A_{jkl}B_{kl}
\end{aligned}$$

By (13) and (14),

$$T_{j2} = O_p\left((n-j)^4\right),$$

$$T_{j3} = O_p\left((n-j)^3\right).$$

On the other hand,

$$\sum_{j=n-3}^{n-1} \frac{1}{(n-j)(j\pi)^2} \sum_{k=j+1}^{n} \sum_{l=j+1}^{n} A_{jkl} B_{kl} = \sum_{j=n-3}^{n-1} \frac{(n-j)}{(j\pi)^2} \int_{\mathbb{R}^q} \left\| \hat{\gamma}_j(\tau) \right\|^2 \omega(d\tau)$$
$$\leq C \sum_{j=n-3}^{n-1} \frac{1}{(j\pi)^2} \to 0\,, \text{as } n \to \infty.$$

We conclude that $MD_n^2 = \widetilde{MD}_n^2 + o_p(1)$. ■

Proof of Theorem 4.1. Define the region $D(\delta) = \left\{ \tau : \delta \leq \|\tau\| \leq \frac{1}{\delta} \right\}$ for each $0 < \delta < 1$.

For a δ, based on Lemma 9.3 and the Continuous Mapping Theorem

$$Q_n(\delta) := \int_{D(\delta)} \int_0^1 \|Z_n(\lambda, \tau)\|^2 \omega(d\tau)\, d\lambda \xrightarrow{d} Q(\delta) := \int_{D(\delta)} \int_0^1 \|Z(\lambda, \tau)\|^2 \omega(d\tau)\, d\lambda.$$

Then we need to show

$$\limsup_{n\to\infty} E \left| Q_n(\delta) - \left\| \hat{Z}_n(\lambda, \tau) \right\|_H^2 \right| = 0, \text{ as } \delta \to 0.$$

$$\limsup_{n\to\infty} E \left| Q(\delta) - \|Z(\lambda, \tau)\|_H^2 \right| = 0, \text{ as } \delta \to 0.$$

$$E \left| Q_n(\delta) - \left\| \hat{Z}_n(\lambda, \tau) \right\|_H^2 \right| = E \left(\int_{\|\tau\|<\delta} \int_0^1 \left\| \hat{Z}_n(\lambda, \tau) \right\|^2 \omega(d\tau)\, d\lambda \right)$$
$$+ E \left(\int_{\|\tau\|>1/\delta} \int_0^1 \left\| \hat{Z}_n(\lambda, \tau) \right\|^2 \omega(d\tau)\, d\lambda \right)$$
$$= C_1 + C_2.$$

$$
\begin{aligned}
C_1 &= \sum_{j=1}^{n-1} \frac{(n-j)}{(j\pi)^2} \int_{\|\tau\|<\delta} E \left\| \hat{r}_j(\tau) \right\|^2 \omega(d\tau) \\
&= \sum_{j=1}^{n-1} \frac{(n-j)}{(j\pi)^2} \int_{\|\tau\|<\delta} E \left(\left\| \frac{1}{n-j} \sum_{t=j+1}^{n} (Y_t - \mu) \phi_{t-j}(Y_{t-j}, \tau) \right\|^2 \right) \omega(d\tau) \\
&= \sum_{j=1}^{n-1} \frac{(n-j)}{(j\pi)^2} \int_{\|\tau\|<\delta} E \left(\frac{1}{(n-j)^2} \sum_{t=j+1}^{n} (Y_t - \mu)' \phi_{t-j}(Y_{t-j}, \tau) \sum_{t=j+1}^{n} (Y_t - \mu) \phi_{t-j}(Y_{t-j}, \tau)^c \right) \omega(d\tau) \\
&= \sum_{j=1}^{n-1} \frac{1}{(j\pi)^2} \int_{\|\tau\|<\delta} E \left(\left\| (Y_t - \mu) \phi_{t-j}(Y_{t-j}, \tau) \right\|^2 \right) \omega(d\tau).
\end{aligned}
$$

Similarly

$$
C_2 = \sum_{j=1}^{n-1} \frac{1}{(j\pi)^2} \int_{\|\tau\|>1/\delta} E \left(\left\| (Y_t - \mu) \phi_{t-j}(Y_{t-j}, \tau) \right\|^2 \right) \omega(d\tau).
$$

As $\int_{\mathbb{R}^q} E \left(\left\| (Y_t - \mu) \phi_{t-j}(Y_{t-j}, \tau) \right\|^2 \right) \omega(d\tau)$ is finite, so

$$
\begin{aligned}
\int_{\|\tau\|>1/\delta} E \left(\left\| (Y_t - \mu) \phi_{t-j}(Y_{t-j}, \tau) \right\|^2 \right) \omega(d\tau) &\to 0, \\
\int_{\|\tau\|<\delta} E \left(\left\| (Y_t - \mu) \phi_{t-j}(Y_{t-j}, \tau) \right\|^2 \right) \omega(d\tau) &\to 0
\end{aligned}
$$

as $\delta \to 0$. So

$$
\begin{aligned}
\limsup_{n\to\infty} E \int_{\|\tau\|<\delta} \int_0^1 \left\| \hat{Z}_n(\lambda, \tau) \right\|^2 \omega(d\tau)\, d\lambda &= 0, \\
\limsup_{n\to\infty} E \int_{\|\tau\|>1/\delta} \int_0^1 \left\| \hat{Z}_n(\lambda, \tau) \right\|^2 \omega(d\tau)\, d\lambda &= 0
\end{aligned}
$$

as $\delta \to 0$. Therefore,

$$
\limsup_{n\to\infty} E \left| Q_n(\delta) - \left\| \hat{Z}_n(\lambda, \tau) \right\|_H^2 \right| = 0, \text{ as } \delta \to 0
$$

Similar arguments apply to prove that

$$
\limsup_{n\to\infty} E \left| Q(\delta) - \left\| Z(\lambda, \tau) \right\|_H^2 \right| = 0.
$$

Then, we conclude, by Theorem 8.6.2 of Resnick (1999),

$$MD_n^2 \xrightarrow{d} MD_\infty^2 := \int_{\mathbb{R}^q} \int_0^1 \|Z(\lambda, \tau)\|^2 \omega(d\tau) d\lambda.$$

Lastly, by Lemma 9.4, and the Slutsky's Thereom, we obtain

$$\widetilde{MD}_n^2 \xrightarrow{d} MD_\infty^2.$$

■

Proof of Theorem 4.2. Note

$$n^{-1/2}\hat{S}_n(\lambda, \tau) = n^{-1/2}\hat{Z}_n(\lambda, \tau) - n^{-1/2}\hat{R}_n(\lambda, \tau),$$

$$\frac{1}{n} MD_n^2 = \left\| n^{-1/2}\hat{S}_n(\lambda, \tau) \right\|_H^2.$$

Under H_A,

$$\begin{aligned} E\int_{\mathbb{R}^q} \|\hat{r}_j(\tau)\|^2 \omega(d\tau) &= E\left(\int_{\mathbb{R}^q} \left| \frac{1}{n-j} \sum_{t=j+1}^n (Y_t - \mu)\, \phi_{t-j}(Y_{t-j}, \tau) \right|^2 \omega(d\tau)\right) \\ &= E\int_{\mathbb{R}^q} \frac{1}{(n-j)^2} \sum_{t=j+1}^n (Y_t - \mu)' \phi_{t-j}(Y_{t-j}, \tau) \sum_{t=j+1}^n (Y_t - \mu)\, \phi_{t-j}(Y_{t-j}, \tau)^c \omega(d\tau) \\ &= \int_{\mathbb{R}^q} \frac{1}{(n-j)^2} \sum_{t=j+1}^n E\left|(Y_t - \mu)\, \phi_{t-j}(Y_{t-j}, \tau)\right|^2 \omega(d\tau) \\ &\quad + \int_{\mathbb{R}^q} \frac{1}{(n-j)^2} \sum_{t,s=j+1, t\neq s,}^n E\left((Y_t - \mu)'(Y_s - \mu)\, \phi_{t-j}(Y_{t-j}, \tau)\phi_{s-j}(Y_{s-j}, \tau)^c\right) \omega(d\tau) \\ &= \frac{1}{n-j} E\left[(Y_t - \mu)'(Y_t - \mu)\right] \left(E\left\|Y_{t-j} - Y_{t-j}^+\right\|\right) \\ &\quad + \frac{2}{(n-j)} \sum_{m=1}^{n-j-1} \left(1 - \frac{m}{n-j}\right) E\left((Y_t - \mu)'(Y_{t+m} - \mu)\left(\|Y_t - Y_{t+m}\| - \left\|Y_t - Y_t^+\right\|\right)\right). \end{aligned}$$

Again by the ergodicity condition and the Toeplitz Lemma, for a fixed j,

$$\frac{2}{(n-j)} \sum_{m=1}^{n-j-1} \left(1 - \frac{m}{n-j}\right) E\left((Y_t - \mu)'(Y_{t+m} - \mu)\left(\|Y_t - Y_{t+m}\| - \left\|Y_t - Y_t^+\right\|\right)\right) \to 0,$$

as $n \to \infty$. So

$$E \int_{\mathbb{R}^q} \|\hat{r}_j(\tau)\|^2 \omega(d\tau) = o(1).$$

Then

$$\left\| n^{-1/2} \hat{Z}_n(\lambda, \tau) \right\|_H = o_p(1).$$

Moreover, it is easy to show

$$\left\| n^{-1/2} \hat{R}_n(\lambda, \tau) \right\|_H = o_p(1),$$

Then

$$\left\| n^{-1/2} \hat{S}_n(\lambda, \tau) \right\|_H^2 = \left\| n^{-1/2} \hat{Z}_n(\lambda, \tau) \right\|_H^2 + o_p(1).$$

Furthermore, similar to the proof of Lemma 9.4,

$$\frac{1}{n} \widetilde{MD}_n^2 = \frac{1}{n} MD_n^2 + o_p(1).$$

Given the region $D(\delta) = \left\{ \tau : \delta \leq \|\tau\| \leq \frac{1}{\delta} \right\}$ for each $0 < \delta < 1$, by the weak law of large number under Assumption 4.1 and the Continuous Mapping theorem,

$$\frac{1}{n} Q_n(\delta) := \int_{D(\delta)} \int_0^1 \left\| n^{-1/2} Z_n(\lambda, \tau) \right\|^2 \omega(d\tau)\, d\lambda \xrightarrow{p} \frac{1}{n} Q(\delta) := \sum_{j=1}^{\infty} \frac{\int_{D(\delta)} \left\| \gamma_j(\tau) \right\|^2 \omega(d\tau)}{(j\pi)^2}.$$

Following similar arguments in the proof of Theorem 4.1, we can get

$$\limsup_{n \to \infty} E \left| \frac{1}{n} Q_n(\delta) - \left\| n^{-1/2} \hat{Z}_n(\lambda, \tau) \right\|_H^2 \right| = 0, \text{ as } \delta \to 0.$$

$$\limsup_{n \to \infty} E \left| \frac{1}{n} Q(\delta) - \sum_{j=1}^{\infty} \frac{\int_{\mathbb{R}^q} \left\| \gamma_j(\tau) \right\|^2 \omega(d\tau)}{(j\pi)^2} \right| = 0, \text{ as } \delta \to 0.$$

Then we have

$$\frac{1}{n} MD_n^2, \frac{1}{n} \widetilde{MD}_n^2 \xrightarrow{p} \sum_{j=1}^{\infty} \frac{\int_{\mathbb{R}^q} \left\| \gamma_j(\tau) \right\|^2 \omega(d\tau)}{(j\pi)^2}.$$

■

Proof of Theorem 4.3. Under Assumptions 4.1 and 4.2, denote

$$\hat{S}_{1n}(\lambda,\tau)=\sum_{j=1}^{n-1}(n-j)^{1/2}\,\hat{\gamma}_{1j}(\tau)\,\frac{\sqrt{2}\sin j\pi\lambda}{j\pi},$$

where

$$\hat{\gamma}_{1j}(\tau)=\frac{1}{n-j}\sum_{t=j+1}^{n}\left(Y_t-\bar{Y}_{n-j}-\frac{g_t}{\sqrt{n}}\right)\exp\left(i\left\langle\tau,Y_{t-j}\right\rangle\right).$$

Then

$$\hat{S}_n(\lambda,\tau)=\hat{S}_{1n}(\lambda,\tau)+\hat{G}_n(\lambda,\tau),$$

where

$$\hat{G}_n(\lambda,\tau)=\sum_{j=1}^{n-1}\sqrt{\frac{n-j}{n}}\frac{1}{(n-j)}\sum_{t=j+1}^{n}g_t\exp\left(i\left\langle\tau,Y_{t-j}\right\rangle\right)\frac{\sqrt{2}\sin j\pi\lambda}{j\pi}.$$

Denote

$$\hat{Z}_{1n}(\lambda,\tau)=\sum_{j=1}^{n-1}(n-j)^{1/2}\,\hat{r}_{1j}(\tau)\,\frac{\sqrt{2}\sin j\pi\lambda}{j\pi},$$

where

$$\hat{r}_{1j}(\tau)=\frac{1}{n-j}\sum_{t=j+1}^{n}\left(Y_t-\mu-\frac{g_t}{\sqrt{n}}\right)\phi_{t-j}(Y_{t-j},\tau).$$

Then

$$\hat{S}_{1n}(\lambda,\tau)=\hat{Z}_{1n}(\lambda,\tau)-\hat{R}_{1n}(\lambda,\tau),$$

where

$$\begin{aligned}\hat{R}_{1n}(\lambda,\tau) &= \sum_{j=1}^{n-1}(n-j)^{1/2}\frac{1}{n-j}\sum_{t=j+1}^{n}\left(Y_t-\mu-\frac{g_t}{\sqrt{n}}\right)\\ &\quad\times\frac{1}{n-j}\sum_{t=j+1}^{n}\phi_{t-j}(Y_{t-j},\tau)\frac{\sqrt{2}\sin j\pi\lambda}{j\pi}.\end{aligned}$$

Note that, under $H_{A,n}$, $Y_t-\mu-\frac{g_t}{\sqrt{n}}$ is a MDS with respect to I_{t-1}, then similar to Lemma 9.2,

$$\left\|\hat{S}_{1n}(\lambda,\tau)\right\|_H^2=\left\|\hat{Z}_{1n}(\lambda,\tau)\right\|_H^2+o_p(1).$$

Similar to Lemma 9.3

$$\hat{Z}_{1n}(\lambda,\tau) \Rightarrow Z(\lambda,\tau)$$

on the Hilbert space $L_2(\Pi,\upsilon)$. On the other hand, $\hat{G}_n(\lambda,\tau) = G(\lambda,\tau) + o_p(1)$. So

$$\hat{S}_n(\lambda,\tau) \Rightarrow Z(\lambda,\tau) + G(\lambda,\tau)$$

on the Hilbert space $L_2(\Pi,\upsilon)$. Finally, following similar arguments as in the proof of Theorem 4.1, we arrive at the conclusion. ■

Proof of Theorem 5.1. Note

$$\begin{aligned}
\hat{\gamma}_j^*(\tau) &= \frac{1}{n-j}\sum_{t=j+1}^{n}\left(Y_t w_t - \bar{Y}_{n-j}^*\right)\exp\left(i\langle\tau, Y_{t-j}\rangle\right) \\
&= \frac{1}{n-j}\sum_{t=j+1}^{n}\left((Y_t-\mu)w_t + \mu w_t - \bar{Y}_{n-j}^*\right)\left[\phi_{t-j}(Y_{t-j},\tau) + E\left[\exp\left(i\langle\tau, Y_{t-j}\rangle\right)\right]\right] \\
&= \frac{1}{n-j}\sum_{t=j+1}^{n}(Y_t-\mu)w_t\phi_{t-j}(Y_{t-j},\tau) - \frac{1}{n-j}\sum_{t=j+1}^{n}(Y_t-\mu)w_t\frac{1}{n-j}\sum_{t=j+1}^{n}\phi_{t-j}(Y_{t-j},\tau).
\end{aligned}$$

Then

$$\hat{S}_n^*(\lambda,\tau) = \hat{Z}_n^*(\lambda,\tau) - \hat{R}_n^*(\lambda,\tau),$$

where

$$\hat{Z}_n^*(\lambda,\tau) = \sum_{j=1}^{n-1}(n-j)^{1/2}\,\hat{r}_j^*(\tau)\,\frac{\sqrt{2}\sin j\pi\lambda}{j\pi},$$

with

$$\hat{r}_j^*(\tau) = \frac{1}{n-j}\sum_{t=j+1}^{n}(Y_t-\mu)w_t\phi_{t-j}(Y_{t-j},\tau),$$

$$\hat{R}_n^*(\lambda,\tau) = \sum_{j=1}^{n-1}(n-j)^{1/2}\frac{1}{n-j}\sum_{t=j+1}^{n}(Y_t-\mu)w_t\frac{1}{n-j}\sum_{t=j+1}^{n}\phi_{t-j}(Y_{t-j},\tau)\frac{\sqrt{2}\sin j\pi\lambda}{j\pi}.$$

Then

$$\left\|\hat{S}_n^*(\lambda,\tau)\right\|_H^2 = \left\|\hat{Z}_n^*(\lambda,\tau)\right\|_H^2 + \left\|\hat{R}_n^*(\lambda,\tau)\right\|_H^2 - 2\,\mathrm{Re}\left\langle\hat{S}_n^*(\lambda,\tau), \hat{R}_n^*(\lambda,\tau)\right\rangle_H.$$

Following similar arguments as in the proof of Theorem 4.1, and 4.3, we arrive at the conclusion, with the understanding that the convergences are conditional on the sample $\{Y_t\}_1^n$. ■

References

Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and Applications*. Springer-Verlag, New York.

Box, G. E. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332):1509–1526.

Chang, J., Yao, Q., and Zhou, W. (2017). Testing for high-dimensional white noise using maximum cross-correlations. *Biometrika*, 104(1):111–127.

Chen, W. W. and Deo, R. S. (2006). The variance ratio statistic at large horizons. *Econometric Theory*, 22(2):206–234.

Choi, I. (1999). Testing the random walk hypothesis for real exchange rates. *Journal of Applied Econometrics*, 14(3):293–308.

Chow, K. and Denning, K. C. (1993). A simple multiple variance ratio test. *Journal of Econometrics*, 58(3):385–401.

Davis, R., Matsui, M., Mikosch, T. V., and Wan, P. (2018). Applications of distance correlation to time series. *Bernoulli*, 24:3087–3116.

Deo, R. S. (2000). Spectral tests of the martingale hypothesis under conditional heteroscedasticity. *Journal of Econometrics*, 99(2):291–315.

Domínguez, M. A. and Lobato, I. N. (2003). Testing the martingale difference hypothesis. *Econometric Reviews*, 22(4):351–377.

Durlauf, S. N. (1991). Spectral based testing of the martingale hypothesis. *Journal of Econometrics*, 50(3):355–376.

Epps, T. W. and Pulley, L. B. (1983). A test for normality based on the empirical characteristic function. *Biometrika*, 70(3):723–726.

Escanciano, J. C. (2007). Weak convergence of non-stationary multivariate marked processes with applications to martingale testing. *Journal of Multivariate Analysis*, 98(7):1321–1336.

Escanciano, J. C. and Lobato, I. N. (2009). *Testing the Martingale Hypothesis*, pages 972–1003. Palgrave Macmillan UK, London.

Escanciano, J. C. and Velasco, C. (2006). Generalized spectral tests for the martingale difference hypothesis. *Journal of Econometrics*, 134(1):151–185.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2):383–417.

Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy*, 86(6):971–987.

Hong, H., Torous, W. N., and Valkanov, R. (2007). Do industries lead stock markets. *Journal of Financial Economics*, 83(2):367–396.

Hong, Y. (1999). Hypothesis testing in time series via the empirical characteristic function: A generalized spectral density approach. *Journal of the American Statistical Association*, 94(448):1201–1220.

Hong, Y. and Lee, T. H. (2003). Inference on predictability of foreign exchange rates via generalized spectrum and nonlinear time series models. *The Review of Economics and Statistics*, 85(4):1048–1062.

Hong, Y. and Lee, Y.-J. (2005). Generalized spectral tests for conditional mean models in time series with conditional heteroscedasticity of unknown form. *The Review of Economic Studies*, 72(2):499–541.

Kim, J. H. (2006). Wild bootstrapping variance ratio tests. *Economics Letters*, 92(1):38–43.

Kuan, C.-M. and Lee, W.-M. (2004). A new test of the martingale difference hypothesis. *Studies in Nonlinear Dynamics and Econometrics*, 8(4):1–26.

Liu, R. Y. (1988). Bootstrap procedures under some non-i.i.d. models. *Annals of Statistics*, 16(4):1696–1708.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.

Lo, A. W. and MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1(1):41–66.

Lobato, I., Nankervis, J. C., and Savin, N. (2001). Testing for autocorrelation using a modified box-pierce q test. *International Economic Review*, 42(1):187–205.

Lobato, I., Nankervis, J. C., and Savin, N. (2002). Testing for zero autocorrelation in the presence of statistical dependence. *Econometric Theory*, 18(3):730–743.

Lobato, I. N. (2001). Testing that a dependent process is uncorrelated. *Journal of the American Statistical Association*, 96(455):1066–1076.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics*, 21(1):255–285.

Park, J. Y. and Whang, Y.-J. (2005). A test of the martingale hypothesis. *Studies in Nonlinear Dynamics and Econometrics*, 9(2):1–32.

Park, T., Shao, X., and Yao, S. (2015). Partial martingale difference correlation. *Electronic Journal of Statistics*, 9(1):1492–1517.

Phillips, P. C. B. and Jin, S. (2014). Testing the martingale hypothesis. *Journal of Business & Economic Statistics*, 32(4):537–554.

Politis, D. N. and Romano, J. P. (1994). Limit theorems for weakly dependent hilbert space valued random variables with application to the stationary bootstrap. *Statistica Sinica*, 4(2):461–476.

Prudnikov, A. P., Brychkov, Y. A., and Marichev, O. I. (1986). *Integrals and series*. Gordon and Breach, New York.

Resnick, S. (1999). *A Probability Path*. Birkhäuser, Boston, MA.

Shao, X. (2011). A bootstrap-assisted spectral test of white noise under unknown dependence. *Journal of Econometrics*, 162(2):213–224.

Shao, X. and Zhang, J. (2014). Martingale difference correlation and its use in high-dimensional variable screening. *Journal of the American Statistical Association*, 109(507):1302–1318.

Székely, G. J. and Rizzo, M. L. (2009). Brownian distance covariance. *The Annals of Applied Statistics*, 3(4):1236–1265.

Székely, G. J. and Rizzo, M. L. (2014). Partial distance correlation with methods for dissimilarities. *Ann. Statist.*, 42(6):2382–2412.

Székely, G. J., Rizzo, M. L., and Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *Annals of Statistics*, 35(6):2769–2794.

Tsay, R. S. (2020). Testing serial correlations in high-dimensional time series via extreme value theory. *Journal of Econometrics*, 216(1):106–117.

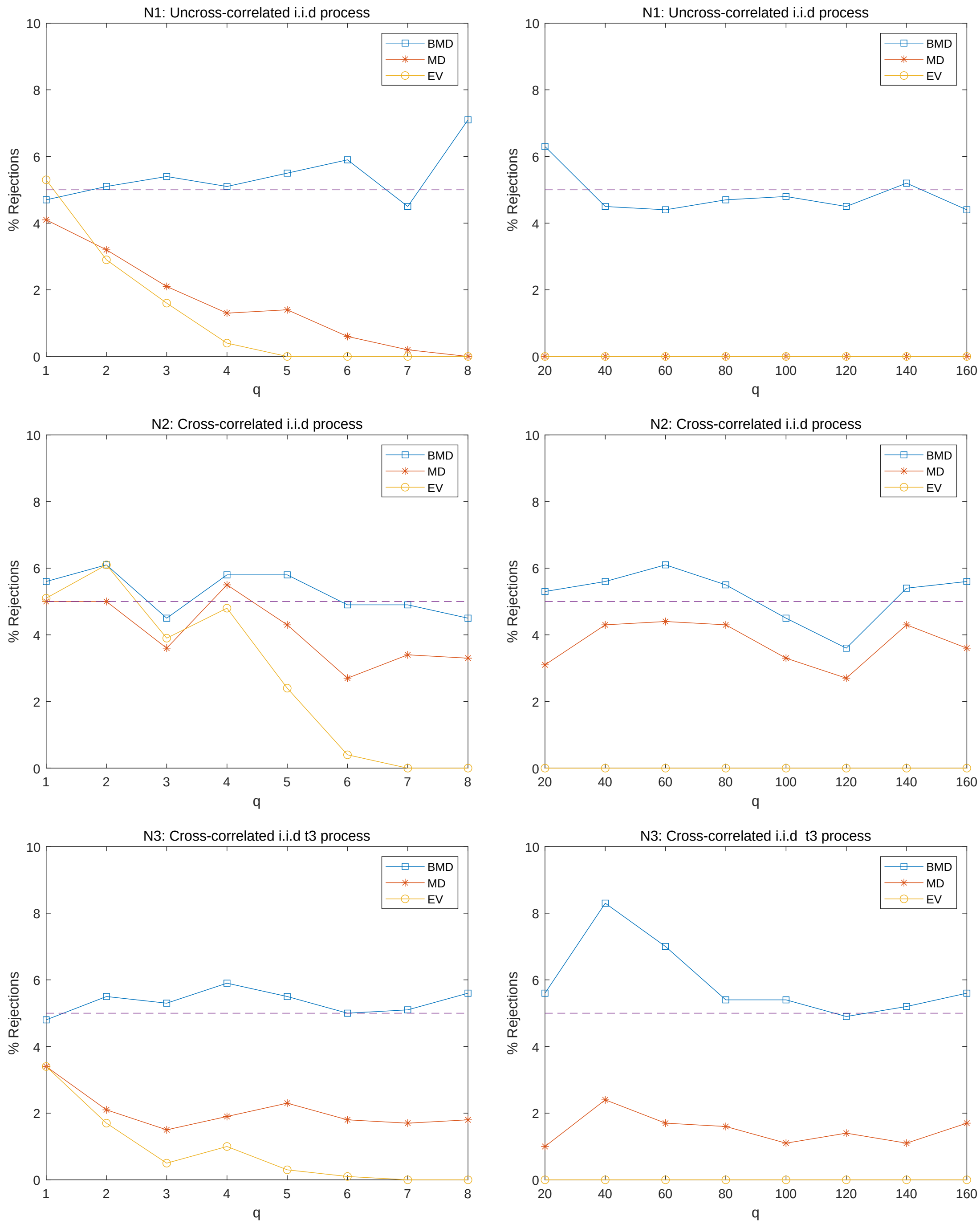
Wang, X. and Sun, Y. (2020). An asymptotic f test for uncorrelatedness in the presence of time series dependence. *Journal of Time Series Analysis*, 41(4):536–550.

Wright, J. H. (2000). Alternative variance-ratio tests using ranks and signs. *Journal of Business & Economic Statistics*, 18(1):1–9.

Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Annals of Statistics*, 14(4):1261–1295.

Yao, S., Zhang, X., and Shao, X. (2018). Testing mutual independence in high dimension via distance covariance. *Journal of The Royal Statistical Society Series B-statistical Methodology*, 80(3):455–480.

Zhang, X., Yao, S., and Shao, X. (2018). Conditional mean and quantile dependence testing in high dimension. *Annals of Statistics*, 46(1):219–246.


N1: Uncross-correlated i.i.d process
N1: Uncross-correlated i.i.d process
N2: Cross-correlated i.i.d process
N2: Cross-correlated i.i.d process
N3: Cross-correlated i.i.d t3 process
N3: Cross-correlated i.i.d t3 process
% Rejections
q
BMD
MD
EV

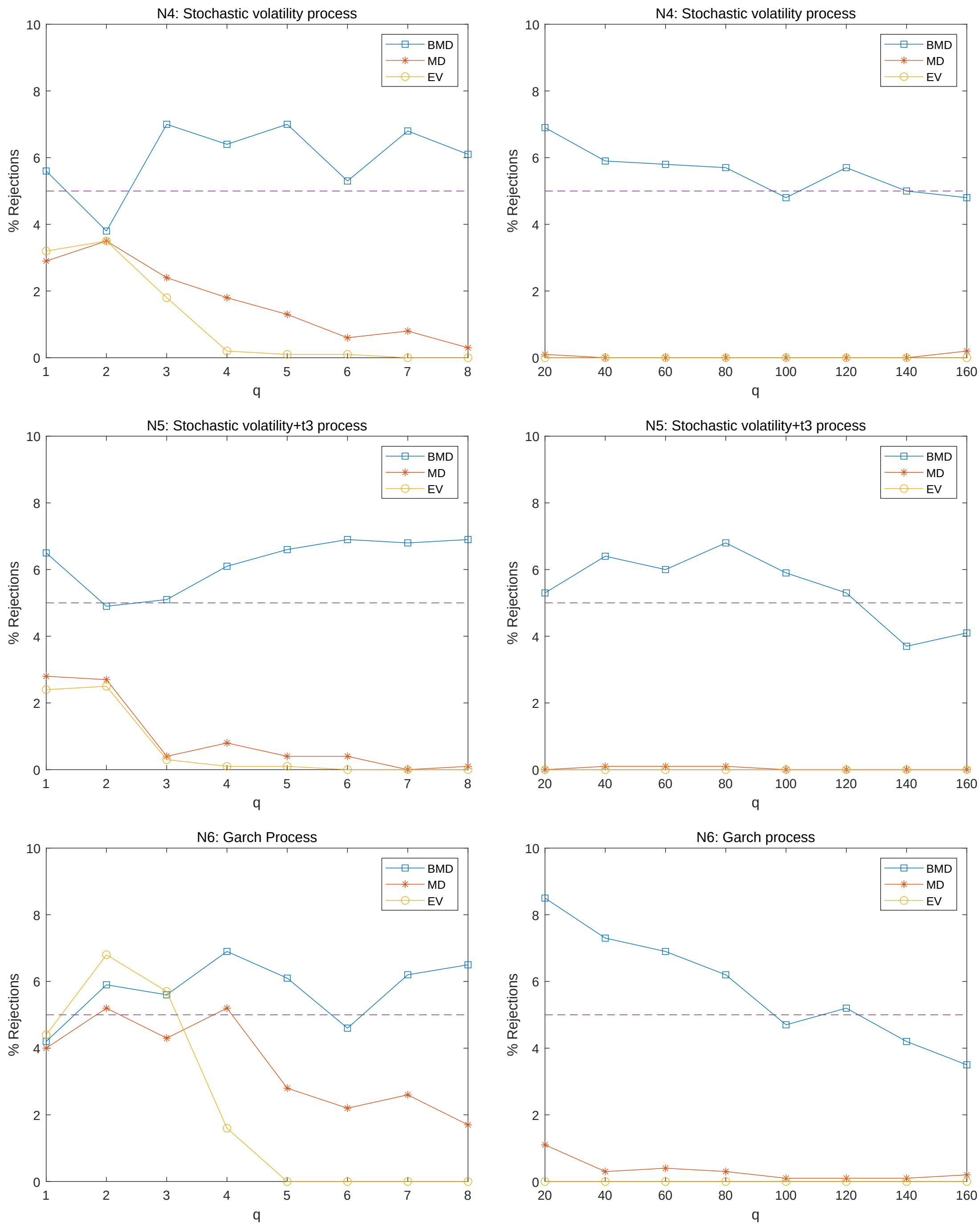
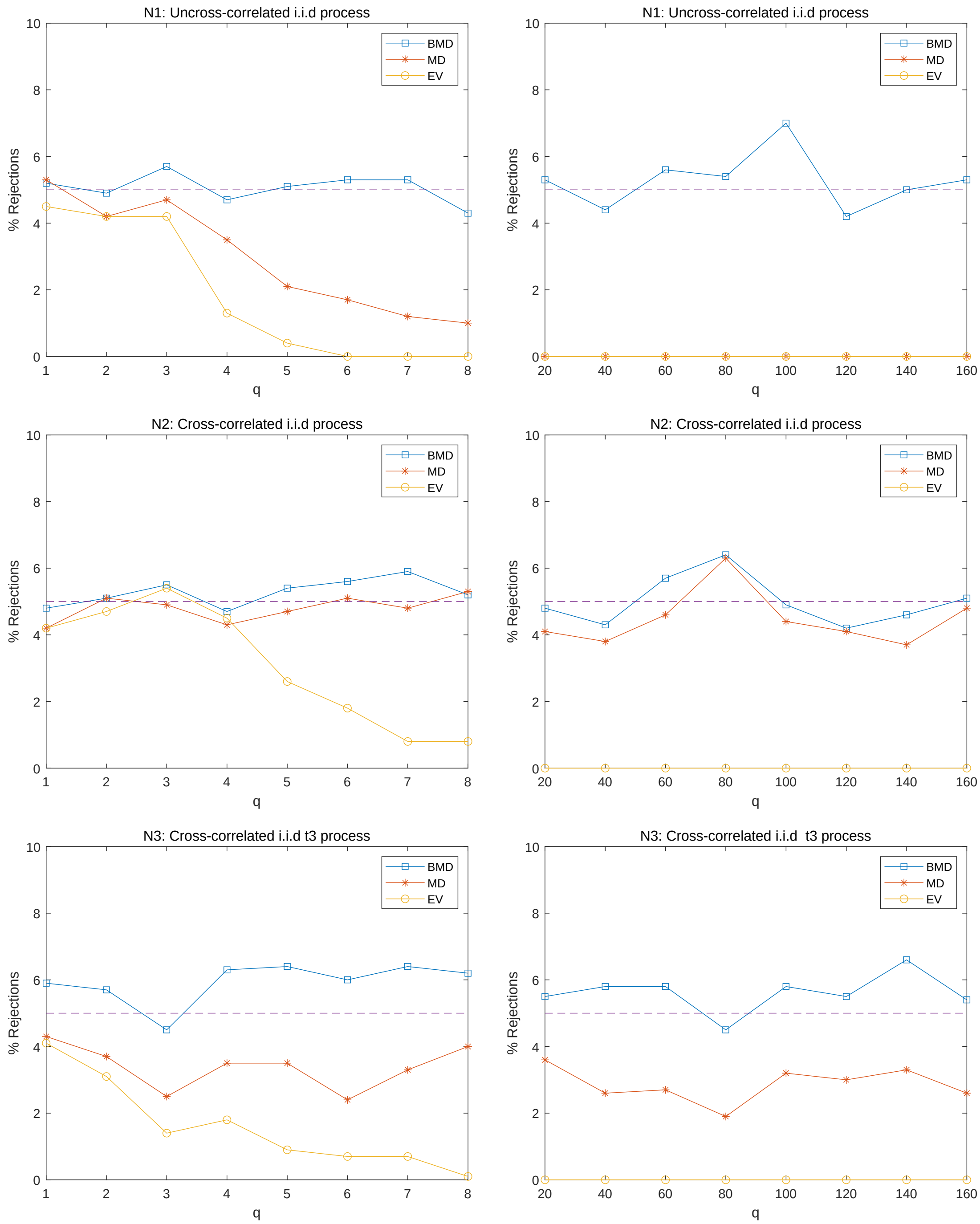


Figure 2: Empirical sizes of generalized spectral tests under the null hypotheses N1-N6. Sample size $n = 100$. BMD represents the bias-corrected statistic $\widetilde{MD}_n^2$; MD represents the statistic MD_n^2; EV represents the statistic EV_n^2.


N1: Uncross-correlated i.i.d process
N1: Uncross-correlated i.i.d process
N2: Cross-correlated i.i.d process
N2: Cross-correlated i.i.d process
N3: Cross-correlated i.i.d t3 process
N3: Cross-correlated i.i.d t3 process
% Rejections
q
BMD
MD
EV

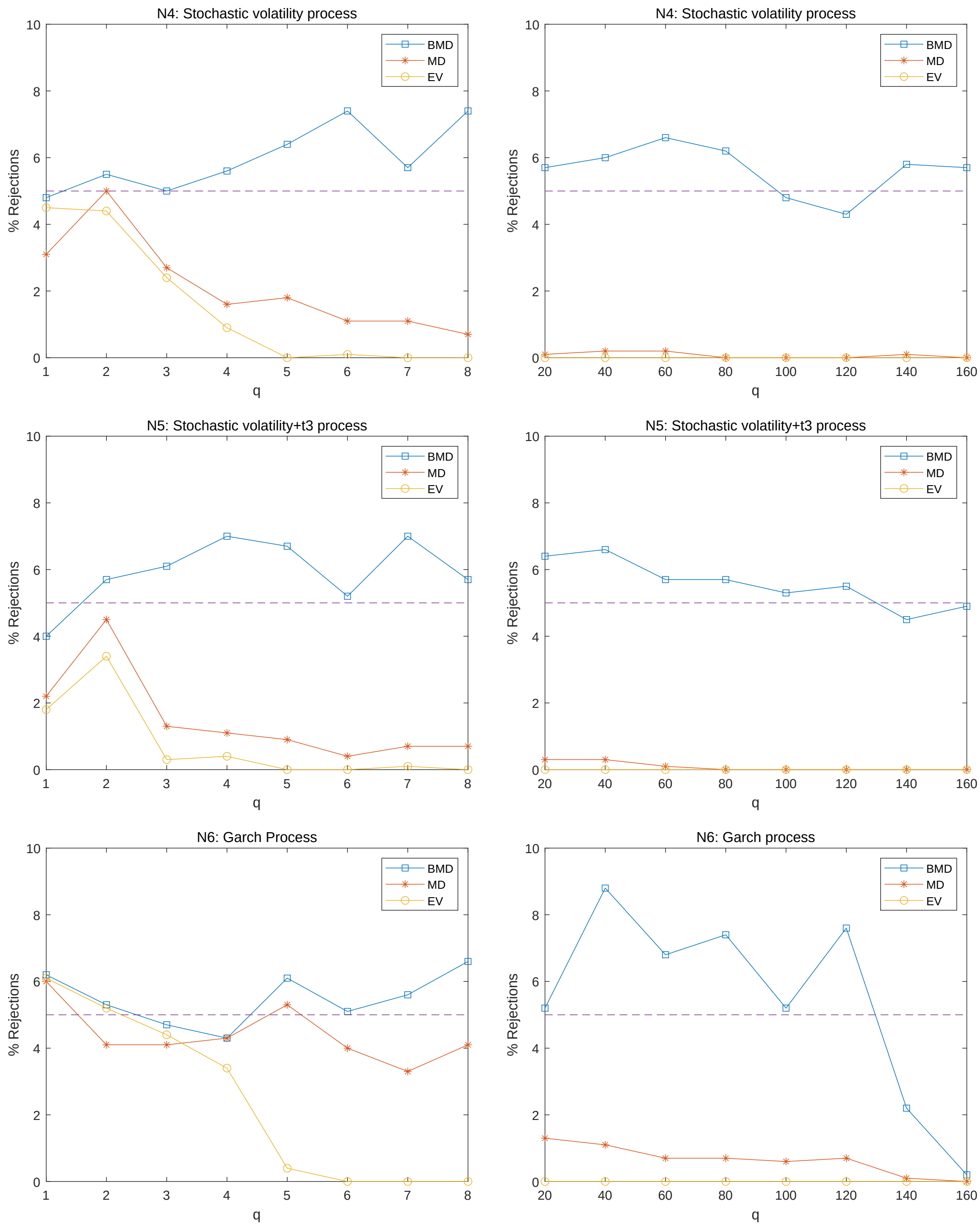
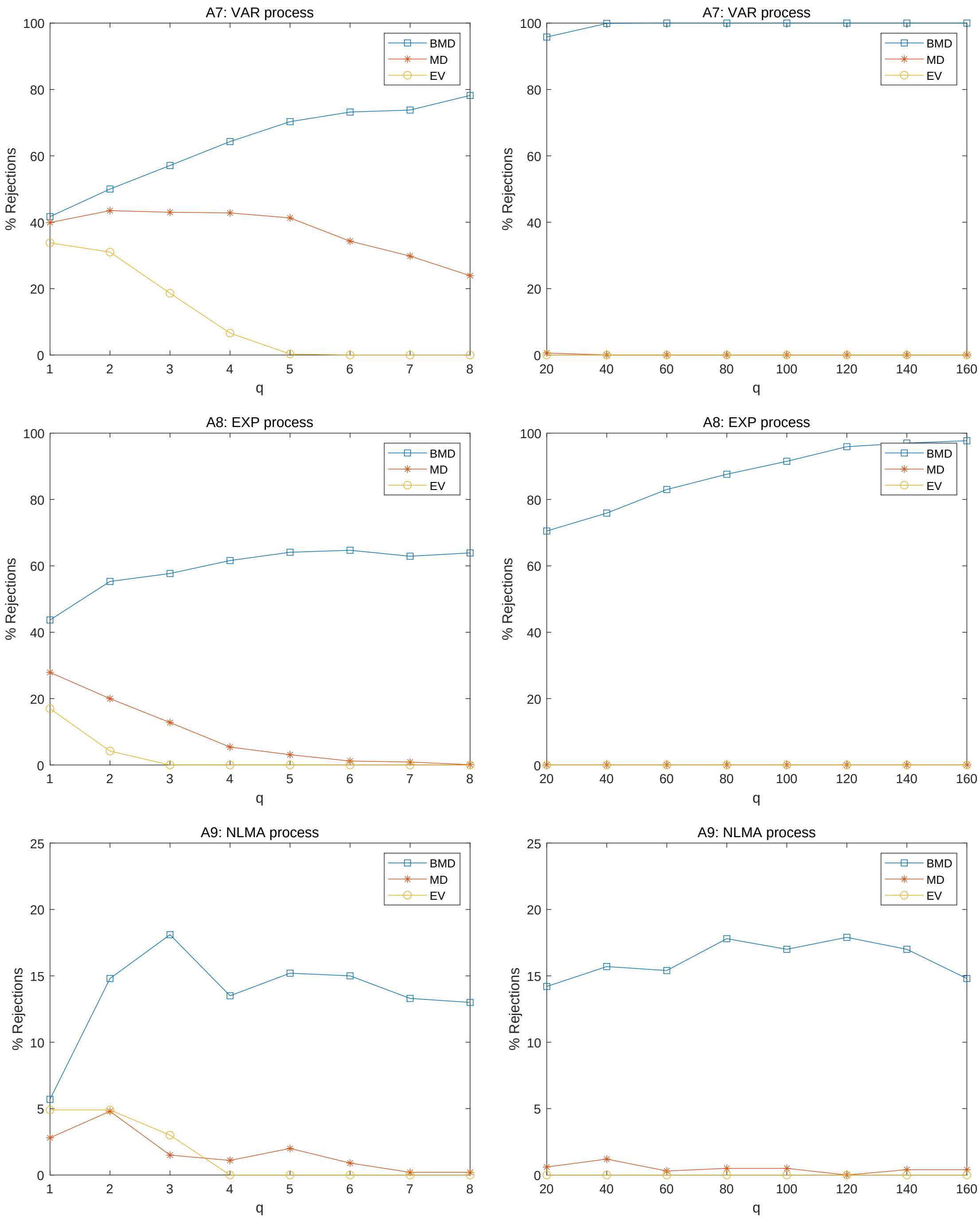


Figure 3: Empirical sizes of generalized spectral tests under the null hypotheses N1-N6. Sample size $n = 300$. BMD represents the bias-corrected statistic $\widetilde{MD}_n^2$; MD represents the statistic MD_n^2; EV represents the statistic EV_n^2.


A7: VAR process
A7: VAR process
A8: EXP process
A8: EXP process
A9: NLMA process
A9: NLMA process
BMD
MD
EV
% Rejections
q

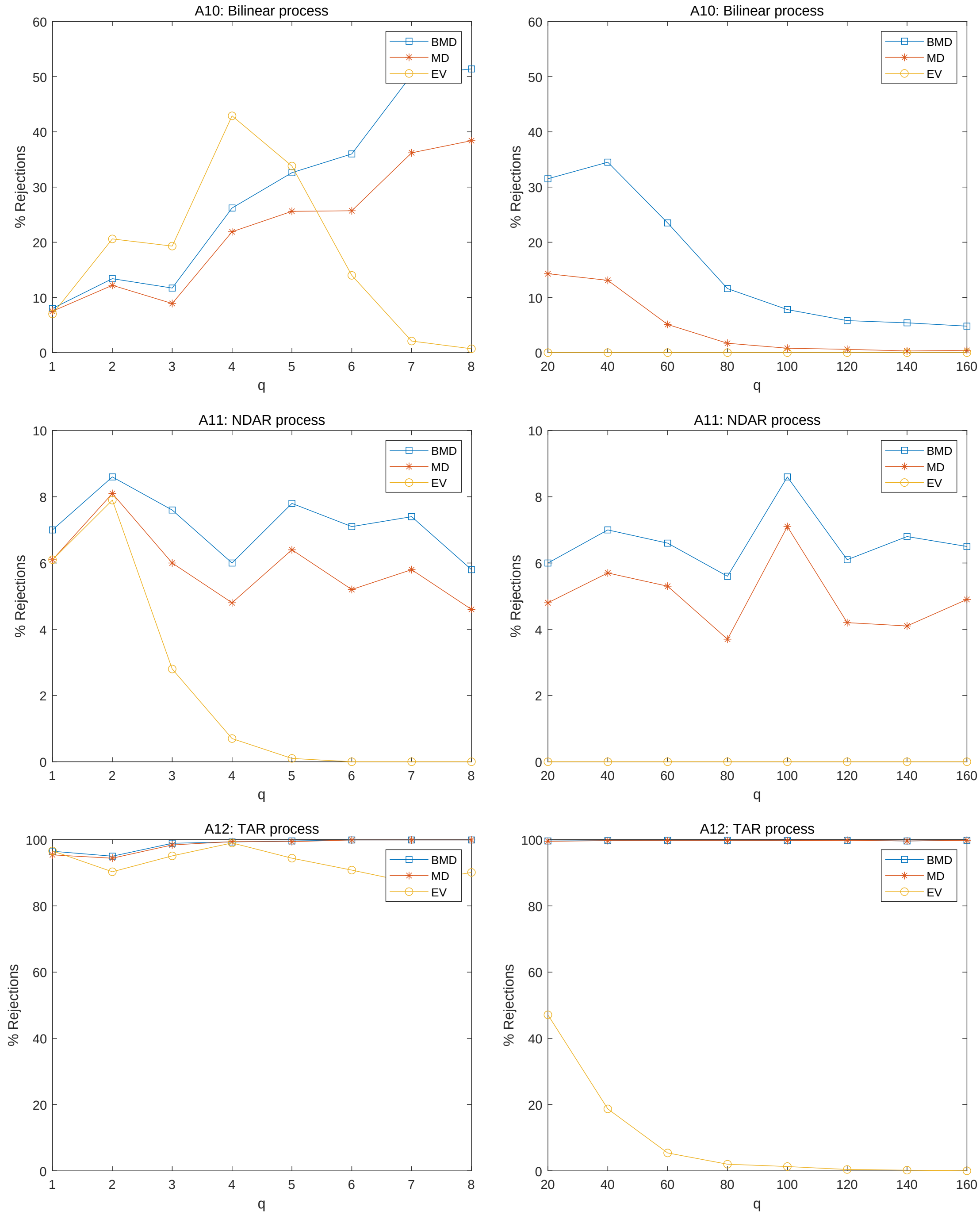
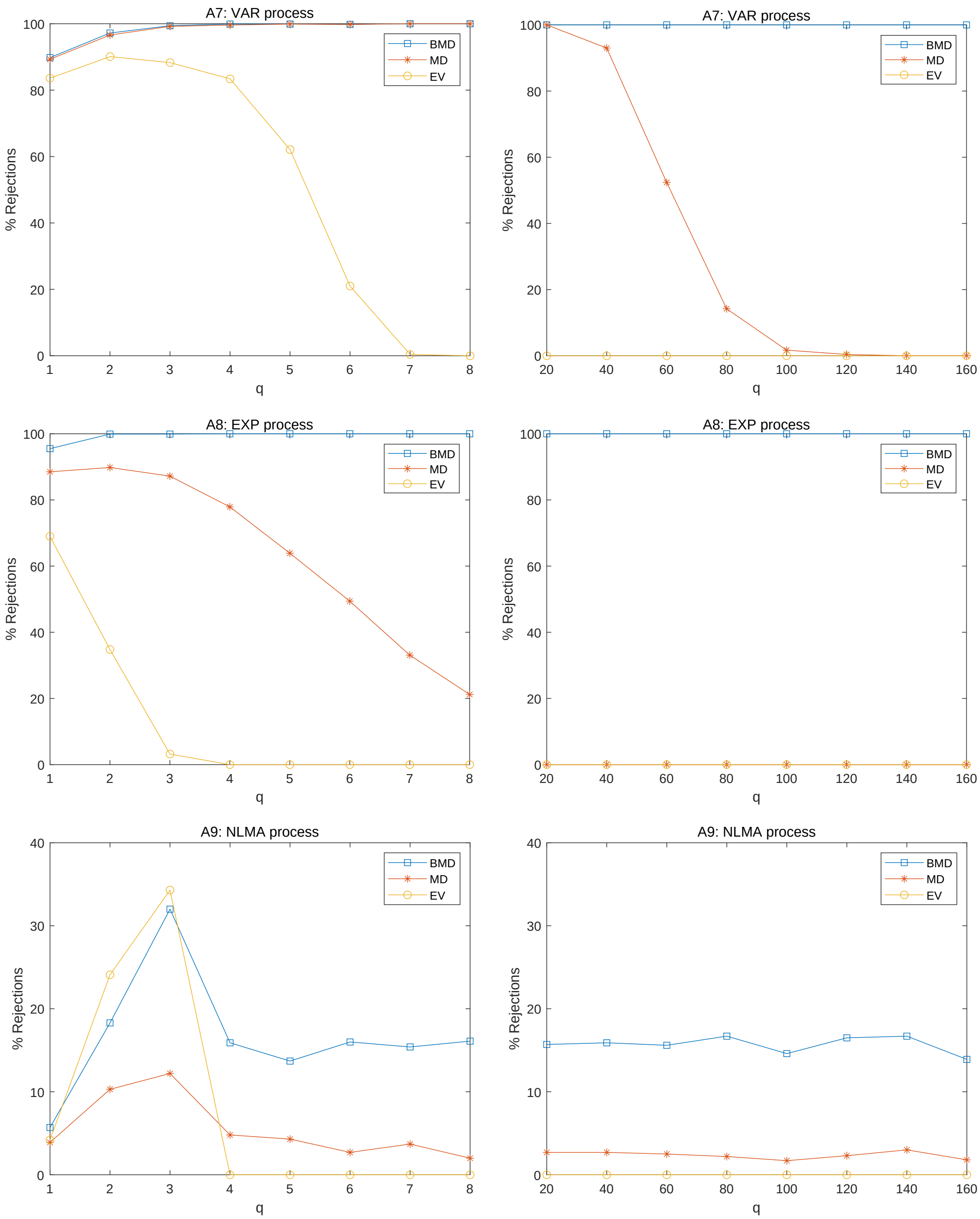


Figure 4: Empirical powers of generalized spectral tests under alternative hypotheses A7-A12. Sample size $n = 100$. BMD represents the bias-corrected statistic $\widetilde{MD}_n^2$; MD represents the statistic MD_n^2; EV represents the statistic EV_n^2.


A7: VAR process
A7: VAR process
A8: EXP process
A8: EXP process
A9: NLMA process
A9: NLMA process
% Rejections
q
BMD
MD
EV

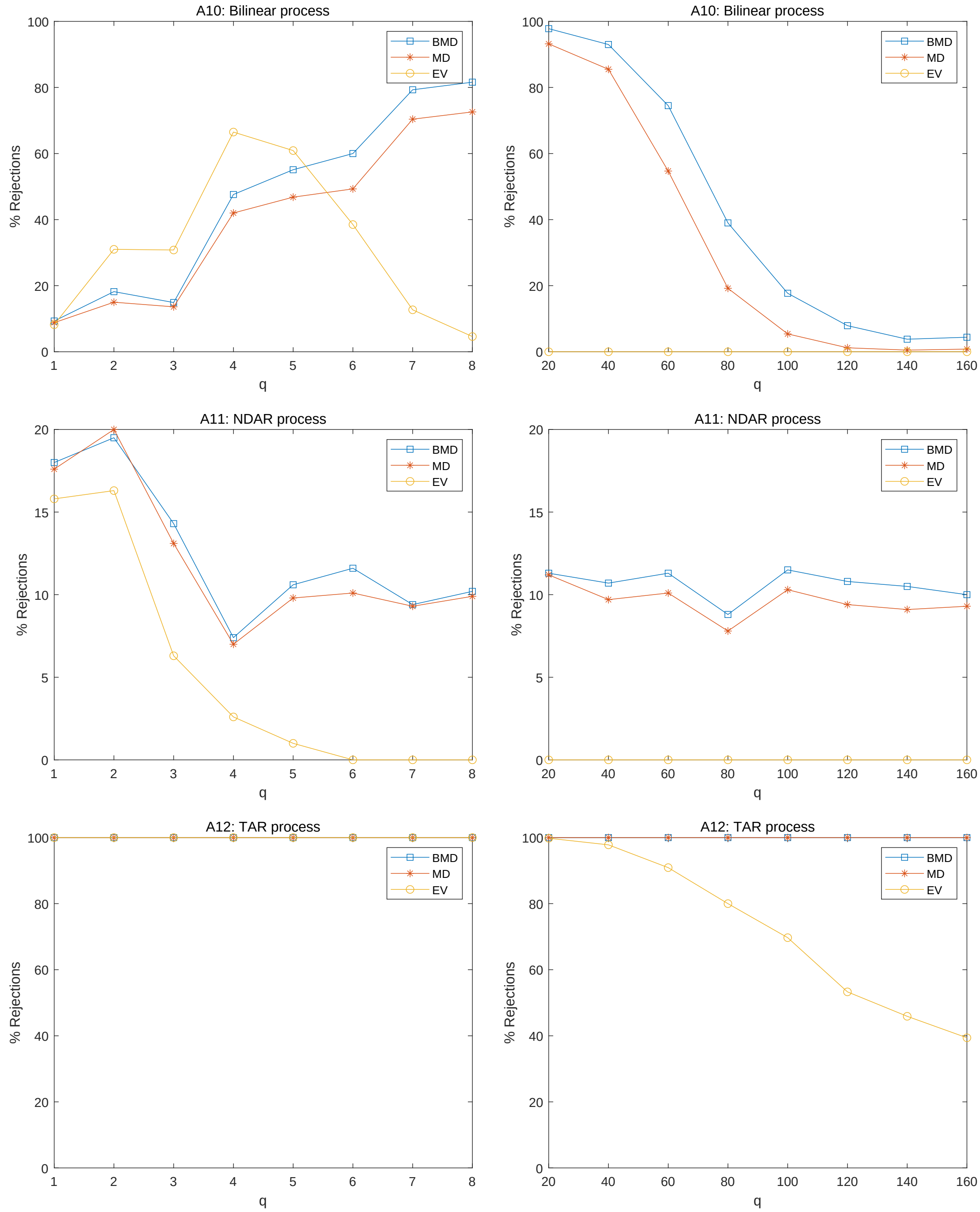


Figure 5: Empirical powers of generalized spectral tests under alternative hypotheses A7-A12. Sample size $n = 300$. BMD represents the bias-corrected statistic $\widetilde{MD}_n^2$; MD represents the statistic MD_n^2; EV represents the statistic EV_n^2.