

# Detecting serial dependence in tail events: A test dual to the BDS test

Cees Diks<sup>1</sup>

CeNDEF, Faculty of Economics and Econometrics, University of Amsterdam, and  
Tinbergen Institute, Roetersstraat 11, 1018 WB, Amsterdam, The Netherlands

July 4, 2002

<sup>1</sup>The author's research is supported by the Netherlands Organization for Scientific Research (NWO) under a MaGW-Pionier grant and a MaGW-Vernieuwingsimpuls grant.

## Abstract

A test for serial independence is proposed which is related to the BDS test but focuses on tail event probabilities rather than probabilities near the center of the distribution. The motivation behind this approach is to make the test more suitable for detecting structure in the tails, such as remaining ARCH- or GARCH-type structure in standardized residuals of financial time series. The new test can be implemented easily by slight modification of the standard BDS test, and is also suitable for model identification. The BDS test and the modified version are compared numerically. To enable fair power comparisons, both tests are implemented as exact level Monte Carlo tests, enabling power calculations of the tests at identical actual sizes. The Monte Carlo implementation allows the use of test statistics which are considerably simpler than for the standard BDS test. For all nonlinear stochastic models examined the power of the new test is found to be uniformly larger over all practically reasonable values of the bandwidth parameter. The test is illustrated with an empirical application.

*Keywords:* Nonparametric tests; Serial dependence; Correlation integral; Monte Carlo tests; Volatility clustering;

*JEL classification:* C12, C14, C52

# 1 Introduction

Recent developments in nonlinear time series analysis have triggered an increasing interest in correlation integral based methods for detecting nonlinear structure in financial time series. Comparisons of various correlation integral based tests concerning the dependence structure in time series, such as that performed by Liu *et al.* (1998), have shown that the BDS test introduced by Brock *et al.* (1987) is highly suitable for economic time series analysis. It enables testing the null hypothesis that a time series consists of IID (independent, identically distributed) observations, against a wide range of alternatives. Indeed, the BDS test by now is an integral part of most standard econometric data analysis software packages. A generalization which can be used to test for independence of a variable and a vector was introduced by Johnson and McLelland (1998).

Although the BDS test is known to have very high power against deterministic structure, simulation studies (Brock *et al.*, 1991; Brock *et al.*, 1996) indicate that the power against conditional heteroskedasticity of the ARCH/GARCH type as specified by Engle (1982) and Bollerslev (1986) is much less impressive. In this paper it is argued that this lack of power may be ascribed, at least partially, to the fact that the BDS test statistic is dominated by contributions from periods in the data with small volatility rather than large volatility. Because empirical evidence suggests that dependence in financial time series is more strong in volatile periods, a modified version of the BDS test which focuses on high volatility regimes can be expected to have larger power against the type of structure present in financial time series.

The aim of this paper is to introduce such a modified test, which will be referred to as the *dual BDS test*, and to perform power comparisons by numerical simulation. For the power comparisons Monte Carlo implementations of both tests are used. The Monte Carlo approach consists of generating random data under the null, conditionally on an observed set of minimal sufficient statistics. For the IID null hypothesis this implies one should condition on the order statistic which gives rise to a permutation test. The Monte Carlo approach has the advantage that tests are obtained with a size exactly equal to a nominal size chosen, so that fair power comparisons can be performed at identical nominal size. For the BDS test as well as the dual BDS test the permutation tests have the additional advantage that simpler test statistics can be used, and that the tests do not break down when the test statistics are degenerate  $U$ -statistics.

This paper is organized as follows. In section 2 the BDS test is described and its dual version introduced. Section 3 describes the implementation of both tests as exact level Monte Carlo tests. The numerical results of the power calculations are presented in section 4, while section 4.4 shows some applications to empirical time series. Section 5 summarizes and discusses the results.

## 2 A dual to the BDS test

Let  $\{X_t\}_{t=1}^N$  denote a sequence of length  $N$  of scalar valued observations from a strictly stationary time series process. The  $m$ -histories  $\mathbf{X}_t^m$  are defined as  $(X_t, \dots, X_{t+m-1})^T$ , where the superscript  $T$  denotes the transposed. The corresponding correlation integrals are then defined as

$$C_m(\epsilon) = E \left[ I_{[0, \epsilon]}(\|\mathbf{X} - \mathbf{Y}\|) \right] = P[\|\mathbf{X} - \mathbf{Y}\| \leq \epsilon] \quad (1)$$

where  $\mathbf{X}$  and  $\mathbf{Y}$  are independent and distributed according to the stationary distribution of the  $m$ -histories,  $I(\cdot)$  denotes the indicator function, which takes either the value 0 or 1 according to

$$I_{[0,\epsilon]}(s) = \begin{cases} 1 & \text{if } s \in [0, \epsilon] \\ 0 & \text{if } s \notin [0, \epsilon], \end{cases}$$

and  $\|\cdot\|$  denotes the supremum norm, given by

$$\|\mathbf{u}\| = \sup_{i=1,\dots,m} |u_i|.$$

The key observation which leads to the BDS test is that correlation integrals factorize for IID time series processes. For the IID case one has

$$C_m(\epsilon) = [C_1(\epsilon)]^m,$$

which follows straightforwardly from considering

$$\begin{aligned} C_m(\epsilon) &= P[\|\mathbf{X} - \mathbf{Y}\| \leq \epsilon] \\ &= P[|X_1 - Y_1| \leq \epsilon, \dots, |X_m - Y_m| \leq \epsilon] \\ &= P[|X_1 - Y_1| \leq \epsilon] \times \dots \times P[|X_m - Y_m| \leq \epsilon] \\ &= [C_1(\epsilon)]^m. \end{aligned}$$

The BDS test consists of comparing estimates of  $C_m(\epsilon)$  and  $[C_1(\epsilon)]^m$ , obtained using  $U$ -statistics estimators of the form

$$\hat{C}_m(\epsilon) = \frac{2}{n(n-1)} \sum_t \sum_{s,s < t} I_{[0,\epsilon]}(\|\mathbf{X}_t^m - \mathbf{X}_s^m\|).$$

it can be shown that the standardized test statistic

$$T = \sqrt{n} \frac{\hat{C}_m(\epsilon) - [\hat{C}_1(\epsilon)]^m}{\hat{\sigma}_m(\epsilon)} \tag{2}$$

with  $\hat{\sigma}_m(\epsilon)$  as in Brock *et al.* (1987) is asymptotically standard normally distributed. The fact that both  $\hat{C}_1(\epsilon)$  and  $\hat{\sigma}_m(\epsilon)$  are functions of the order statistic allows for considerable simplifications of the test statistic when the BDS test is implemented as a permutation test. The statistical motivation for using permutation tests will be discussed in detail in section 3.

Although, strictly speaking,  $C_m(\epsilon) - [C_1(\epsilon)]^m$  is not semi-positive definite, a one-sided implementation of the test, rejecting the IID null hypothesis if  $T$  is too large, usually has larger power than a two-sided test of the same nominal level. This can be understood from the fact that the correlation integral may be interpreted as a measure for randomness. The more dependence structure is present in a stationary time series process, given the marginal distribution of the process, the larger the correlation integral tends to be. For processes which are deterministic or nearly deterministic in that randomness is small, this can be seen from the following argument. For deterministic processes two initially close time paths (represented as  $m$ -histories) tend to stay close in the short run. This leads to a larger conditional probability  $P[\|\mathbf{X}_{s+1}^m - \mathbf{X}_{t+1}^m\| \leq \epsilon \mid \|\mathbf{X}_s^m - \mathbf{X}_t^m\| \leq \epsilon]$  than under the IID null, and hence to larger correlation integrals than expected under the null hypothesis.

A similar argument also provides insights into the power (albeit small) of the BDS test against dependence in higher order moments, such as conditional heteroskedasticity. Consider a process with constant conditional mean and time-varying, auto-correlated conditional variance (i.e. a process with volatility clustering, such as a GARCH process or a stochastic volatility process). Two  $m$ -histories of returns  $\mathbf{X}_s^m$  and  $\mathbf{X}_t^m$  that are  $\epsilon$ -close are likely to be both realisations from periods in which the volatility was relatively small compared to the mean volatility of the process, since delay vectors from low volatility periods have a smaller variance and hence a higher probability of being  $\epsilon$ -close. In the presence of volatility clustering it is then more likely than under the IID null that the short term future volatility is also smaller than average for the time periods following these two  $m$ -histories. This gives a larger conditional probability for the distance  $\|\mathbf{X}_s^m - \mathbf{X}_t^m\|$  to be smaller than  $\epsilon$  than for an IID process.

The previous arguments indicate that the power of the BDS test against volatility clustering is obtained through contributions of  $m$ -histories from time periods with relatively small volatility (tranquil periods). However, typical models with volatility clustering explicitly account for the empirically observed fact that large current volatility implies above average short-term future volatility. The fact that this also implies that small returns tend to follow periods in which returns were small is a side-effect rather than a main characteristic of these models. These considerations suggest that a modified version of the BDS test which is sensitive to persistence of high volatility instead of persistence of low volatility might have more power against the structure represented by standard econometric time series models with conditional dependence in higher moments.

The main idea behind the alternative test proposed here is the following: instead of estimating the joint probability of consecutive *small* differences between two time-paths, estimate the joint probability of consecutive *large* differences between two time paths. To this end we introduce a modified version of the correlation integral, referred to as the *dual correlation integral*:

$$C'_m(\epsilon) = E \left[ I_{[\epsilon, \infty)}(\|\mathbf{X} - \mathbf{Y}\|') \right] = P[\|\mathbf{X} - \mathbf{Y}\|' \geq \epsilon],$$

where the indicator function equals one on  $[\epsilon, \infty)$ , and is zero otherwise, and the supremum norm is replaced by the infimum norm:

$$\|\mathbf{u}\|' = \inf_{i=1, \dots, m} |u_i|.$$

For the usual correlation integral, given in Eq. (1), nonzero contributions to the estimated correlation integral are obtained from pairs of time indices for which the differences between trajectories are smaller than  $\epsilon$  for each of  $m$  consecutive time steps, that is, when of the  $m$  elements of  $\mathbf{X}_s^m - \mathbf{X}_t^m$  are smaller than  $\epsilon$  in absolute value. In contrast, for the dual correlation integral the requirement is exactly opposite in that each of the  $m$  elements of the vector should be larger than  $\epsilon$  in absolute value. The situation is illustrated in Fig. 1 for the case  $m = 2$ .

It can be easily verified that, like the correlation integral, the dual correlation integral factorizes for IID processes:

$$\begin{aligned} C'_m(\epsilon) &= P[\|\mathbf{X} - \mathbf{Y}\|' \geq \epsilon] \\ &= P[|X_1 - Y_1| \geq \epsilon, \dots, |X_m - Y_m| \geq \epsilon] \\ &= P[|X_1 - Y_1| \geq \epsilon] \times \dots \times P[|X_m - Y_m| \geq \epsilon] \\ &= [C'_1(\epsilon)]^m. \end{aligned}$$

The dual correlation integral is estimated in a similar way as the usual correlation integral

$$\hat{C}'_m(\epsilon) = \frac{2}{n(n-1)} \sum_t \sum_{s,s<t} I_{[\epsilon,\infty)}(\|\mathbf{X}_t^m - \mathbf{X}_s^m\|').$$

Existing software for calculating the sample correlation integral can be easily adapted for the dual correlation integral. This merely involves replacing the appropriate smaller than or equal signs by larger than or equal signs in the source code.

It is possible to develop asymptotic theory for the dual correlation integral in a way completely analogous to the asymptotic theory for the usual correlation integral. Upon replacing the kernel function of the usual correlation integral by the kernel function by which the dual correlation integral is defined, the results from Brock *et al.* (1996) carry over directly to the dual correlation integral, giving the asymptotic standard normality of  $T' = \sqrt{N}(\hat{C}'_m(\epsilon) - [\hat{C}'_1(\epsilon)]^m)/\hat{\sigma}'_m(\epsilon)$  under similar conditions as for the BDS test statistic  $T$  given in Eq. (2). However, for reasons that will be discussed in detail in the next section, it is more convenient at this point to proceed by implementing both tests as Monte Carlo tests. Not only does this facilitate power comparisons at the same nominal size, but it also allows the use of simplified test statistics, and does not break down in cases where the test statistics are degenerate  $U$ -statistics.

### 3 Monte Carlo tests

When possible, using a Monte Carlo test for testing a given null hypothesis has several advantages over using alternative methods, arguably the most important one being the fact that Monte Carlo tests have exact level for finite sample sizes. Another advantage is that Monte Carlo tests can be used in conjunction with arbitrary test statistics, thus giving the econometrician freedom of choosing any test statistic of choice. For the BDS test and the dual test proposed here, there is the additional advantage that the Monte Carlo approach permits the use of simplified test statistics; the test statistics neither need to be centered around zero, nor scaled to unit variance, as the  $p$ -values obtained are invariant under such transformations.

The Monte Carlo approach was first suggested by Barnard (1963) in the context of testing a simple null hypothesis, that is, the case where the null model is fully specified. The main idea is to explicitly construct the null distribution of the test statistic by generating a large number of independent realizations using the null model, and calculating the test statistic for each of the realizations. The significance of an empirically observed value of the test statistic is then determined given the fact that, under the null hypothesis, the test statistic for the original data and the artificial data are independent draws from the same distribution. In most situations however, the null hypothesis of interest is not of the simple type but composite, and the procedure cannot be applied directly, since the distribution of the test statistic may depend on unknown model parameters. Besag and Diggle (1977) and also Engen and Lillegård (1997) have shown that the Monte Carlo approach can be used for testing composite null hypotheses, provided that the data generating process is conditioned on a minimal set of sufficient statistics under the null, thus removing the dependence on any unknown nuisance parameters.

At first sight it might seem problematic that the null hypothesis considered here (IID with an unspecified marginal distribution) cannot be represented as a finite dimensional parametric

family of distributions. However, Diks and Tong (1999) show that the Monte Carlo approach is flexible enough to enable testing for a given symmetry of an otherwise entirely unspecified distribution. The situation considered here is similar, with the unknown marginal distribution being the nuisance parameter. It is well-known that under the IID null the order statistics provide a minimal and sufficient statistic. Because, given the order statistic, each permutation of the observed data is equally probable under the null, conditioning on the order statistic automatically leads to a permutation test. Since permutation tests can be implemented straightforwardly, this offers a convenient way of constructing an exact test for independence based on any test statistic of choice.

The  $p$ -values, in case a real-valued test statistic  $T$  is used, can be obtained as follows. Let  $T_0$  denote the value of the test statistic for the original time series and let  $T_1, \dots, T_{M-1}$  denote the values of the test statistic calculated from  $M-1$  randomly permuted replications of the original data. Let us for the moment assume that large values of  $T$  indicate dependence, so that one-sided  $p$ -values are desired, the modifications for two-sided tests being obvious. The  $p$ -value is defined as

$$\hat{p} = \frac{\#\{i = 0, \dots, M-1; T_i \geq T_0\}}{M},$$

where the nominator represents the number of values in  $T_0, \dots, T_{M-1}$  that are at least as large as  $T_0$ . Under the null hypothesis,  $T_1, \dots, T_{M-1}$  obey permutation symmetry, so that  $\hat{p}$  is distributed uniformly on  $1/M, \dots, M/M = 1$ , under the assumption that the conditional distribution of  $T$  is continuous. The probability that  $T_0$  is among the largest  $k$  values is then equal to  $k/M$ , and rejecting the null hypothesis whenever  $\hat{p} \leq \alpha$ , where  $\alpha = l/M$  for integer  $l$ , yields an exact level- $\alpha$  test. More generally, an exact level- $\alpha$  test is obtained also in cases where the distribution of  $T$  is not continuous, provided that  $T_0$  is given a rank uniformly distributed among the observed test statistics with identical value (i.e. ties are ranked randomly).

Apart from having the desired exact size property, Monte Carlo tests have the apparent drawback of being random in the sense that they do not accept or reject the null with certainty for a given dataset. In the Monte Carlo literature it was observed that this uncertainty may lead to a small loss of power for values of  $T$  near the critical value. The fact that the uncertainty decreases with  $M$ , then naturally leads to the question of which value to take for  $M$  given that one wishes to test at nominal size  $\alpha$ . Hope (1968) and Marriott (1979) showed that the power indeed decreases when smaller values of  $M$  are used, but it does so only slightly so that there is no need to choose an excessively large value of  $M$ . (Marriott, 1979) suggests that taking  $l$  as small as 5 is already reasonable from which it follows that  $M = 5/\alpha$  is a sufficiently large number of simulations. Here  $M = 200$  (199 random permutations) are used throughout which leads to  $l = 10$  for a test of size  $\alpha = 0.05$ .

As mentioned above, the statistics used by the BDS test and the dual test simplify considerably when the tests are implemented as Monte Carlo tests. This is shown here only for the BDS test, the arguments being similar for the dual BDS test. Consider implementing a permutation test based on the BDS test statistic  $T$  given in Eq. (2). Both the estimators  $\hat{C}_1(\epsilon)$  and  $\hat{\sigma}_m(\epsilon)$  are independent of the ordering of the data, and hence are functions of the order statistic only. Therefore, their values will be identical for the original time series and all possible Monte Carlo replications. Since  $\hat{\sigma}_m(\epsilon)/\sqrt{N}$ , conditionally on the order statistic, is merely a constant scale parameter and  $\hat{C}_1(\epsilon)$  a location parameter, they have no effect on the  $p$ -values obtained eventually. Therefore, the Monte Carlo test can be implemented without scaling of the test statistic, that is, simply using a test statistic equal to the estimated

correlation integral,  $\tilde{T} = \hat{C}_m(\epsilon)$ , suffices. This observation is in exact accordance with the remark made previously, that conditionally on the order statistic, for  $m > 1$ , larger correlation integrals typically are to be expected for dependent processes than for IID processes. For the dual BDS test similar arguments imply that the estimated dual correlation integral,  $\tilde{T}' = \hat{C}'_m(\epsilon)$  can be used as a test statistic in the permutation test.

Notice that since Monte Carlo tests can be based on any test statistic of choice, they remain valid even when the test statistics happen to be degenerate  $U$ -statistics, in which case inference by using an asymptotic normal approximation would fail.

Finally it should be noted that a Monte Carlo test differs from a bootstrap test in that it is a resampling scheme which by construction preserves the sufficient statistic exactly. For the IID null this leads to a permutation test, that is, resampling without replacement. Resampling with replacement would not preserve the order statistic exactly and is thus not guaranteed to have exact size properties. Although resampling with replacement could provide an exact size test asymptotically, for finite sample sizes it will certainly not provide a test with exact nominal size for arbitrary test statistics.

## 4 Power comparisons

Since Monte Carlo testing amounts to performing exact level- $\alpha$  tests, the size of both tests equals the nominal level  $\alpha$  chosen. Therefore, the size of the tests will not be examined numerically here and focus will be on the power of the tests for various alternatives to the null hypothesis. Chaotic as well as stochastic time series are considered. All time series are normalized to unit sample standard deviation prior to analysis, and  $p$ -values are estimated using  $M = 200$  (199 random permutations).

### 4.1 Chaotic time series

The power is studied for three chaotic series models, the logistic map,

$$X_{t+1} = 4X_t(1 - X_t),$$

the time series consisting of the  $X$ -values generated by the two-dimensional Hénon map

$$\begin{aligned} X_{t+1} &= 1 - 1.4X_t^2 + Y_t \\ Y_{t+1} &= 0.3X_t, \end{aligned}$$

and the  $X$ -values of the Lozi map

$$\begin{aligned} X_{t+1} &= 1 - 1.7|X_t| + Y_t \\ Y_{t+1} &= 0.5X_t. \end{aligned}$$

The power of the tests was determined at a nominal level of 5% ( $\alpha = 0.05$ ) for a range of  $\epsilon$  values equal to  $2^{(i-12)/4}$  for  $i = 0, \dots, 20$ , which includes the range of values used most often in practice:  $\epsilon \in [0.5, 1.5]$ . Because the tests have very high power against determinism, time series as short as  $N = 15$  were used for the chaotic maps. The results based on 1000 replications are given in the first three panels of Fig. 2. The results indicate that the power for the BDS test and the modified test are similar for the logistic map, whereas the power for the other two chaotic models is larger for the dual test, in particular for larger values of  $\epsilon$ .



## 4.2 Stochastic time series

Next the power of the two tests will be examined using time series simulated from the models considered by Brock *et al.* (1996). These are frequently used as models for logarithmic return time series in economics and finance. For these stochastic models much longer time series are required than for the chaotic models discussed above to obtain powers which are considerably larger than the level  $\alpha = 0.05$ . To facilitate comparison with the results of Brock *et al.* (1996) time series of the same length ( $N = 250$ ) are used.

The first two models considered are ARCH (Engle, 1982) and GARCH (Bollerslev, 1986) models. Both the ARCH and GARCH time series were generated according to

$$X_t \sim N(0, \sigma_t^2)$$

with

$$\sigma_t^2 = 1 + 0.5X_{t-1}^2$$

for the ARCH model, and

$$\sigma_t^2 = 1 + 0.1X_{t-1}^2 + 0.8\sigma_{t-1}^2$$

for the GARCH model.

The third model is a NLMA (nonlinear moving average) model, defined as

$$X_t = 0.5\epsilon_{t-2}\epsilon_{t-1} + \epsilon_t$$

with  $\epsilon_t \sim N(0, 1)$ .

The fourth model is an ENLMA (extended nonlinear moving average) model, with

$$X_t = 0.8\epsilon_{t-1} \sum_{j=2}^{20} (0.8)^{j-2} \epsilon_{t-j} + \epsilon_t.$$

The last model considered is the TAR (threshold autoregressive) model:

$$X_t = \begin{cases} -0.5X_{t-1} + \epsilon_t & \text{if } X_{t-1} \leq 1 \\ 0.4X_{t-1} + \epsilon_t & \text{otherwise.} \end{cases}$$

The last five panels in Fig. 2 shows the obtained average power for 1000 realizations. The tests were performed at nominal (and actual) size  $\alpha = 0.05$  using  $M = 200$  Monte Carlo replications. The embedding dimension  $m$  was set equal to 3. The results clearly indicate that the modified test outperforms the traditional BDS test almost uniformly. Only for very small values of  $\epsilon$  the BDS test in some cases rejected more often than the dual test. These values of  $\epsilon$  lie well outside the practically used, and, as the results indicate, relevant, range of  $\epsilon$ -values. Notice that for all models, except possibly the TAR model, the optimal bandwidth for the dual test is larger than for the BDS test. A possible explanation for this phenomenon, at least for processes with volatility clustering in the case  $m = 2$  is the following. If large shocks are likely to be followed by relatively large shocks, situations where  $|X_t - X_s|$  is large and  $|X_{t+1} - X_{s+1}|$  is small will be relatively low compared to that expected under the null. Upon increasing  $\epsilon$  these low probability regions are included in the estimates of the correlation integral, as a result of which the power of the usual BDS test decreases.

Table 1 shows a comparison with two parametric tests for independence, the Tsay (1986) test and the test proposed by Engle (1982). The powers quoted for the Tsay and Engle tests

Power	T	E	BDS	dual BDS
ARCH	0.44	0.98	0.93	0.98
GARCH	0.18	0.49	0.44	0.50
NLMA	0.99	0.72	0.62	0.65
ENLMA	0.64	0.89	0.93	0.96
TAR	0.99	0.63	0.41	0.63

**Table 1:** Powers of the Tsay test, the Engle test, and of the BDS test and the dual BDS test for  $\epsilon = 1$ . The first two tests are based on models of order 2, which corresponds with dependence in 3-histories,  $m = 3$ .

are those published by Brock *et al.* (1996). The powers for the Monte Carlo BDS and dual BDS test are quoted for  $m = 3$  and  $\epsilon = 1$ . The orders for the parametric tests are such that they correspond with tests for dependence of the elements of the three-histories  $\mathbf{X}_t^3$ . Notice that the power of the dual BDS test is comparable to that of the Engle test, a parametric test which is explicitly designed to have power against ARCH type structure.

### 4.3 Linear combinations of the test statistics

Because the Monte Carlo approach can be used with any test statistic it is straightforward to modify the test statistic. This offers the possibility of combining the BDS test and the dual test, for example by using a weighted average

$$T_\lambda = (1 - \lambda)\widehat{C}_m(\epsilon) + \lambda\widehat{C}'_m(\epsilon)$$

with  $\lambda \in (0, 1)$ , or using a weighted average with different bandwidth parameters  $(1 - \lambda)\widehat{C}_m(\epsilon_1) + \lambda\widehat{C}'_m(\epsilon_2)$ . Figure 3 shows the observed powers for the chaotic as well as the dstochastic models described in the previous section, for the test statistic  $T_\lambda$ . The contour lines on the base of these plots correspond to power values  $0, 0.1, \dots, 1$ . For the logistic map small values of  $\lambda$  can be seen to be optimal. Except for the NLMA model, for the nonlinear stochastic models the optimal values of  $\lambda$  are found to be equal to or very close to 1, which indicates that putting all weight on the dual BDS test is optimal for those models.

### 4.4 Application

We examine the behaviour of the test for 1013 daily IBM returns from January 2, 1990 until December 31, 1993. The returns time series as well as the standardized residuals for a GARCH(1,1) model are shown in Fig. 4. We determined  $p$ -values for the standardized residuals on the basis of  $M = 1000$  monte carlo simulations, for various values of the weight  $\lambda$ . Figure 5 shows the results. It can be observed that the most significant results (small  $p$ -values) are obtained for  $\epsilon$ -values around one standard deviation. Also, it can be seen that the  $p$ -values increase for small  $\lambda$ , that is, when the test is close to the standard BDS test. This is in accordance with the results found before for the nonlinear stochastic models, which indicated that a value of  $\lambda$  close to one is optimal.

## 5 Summary and Discussion

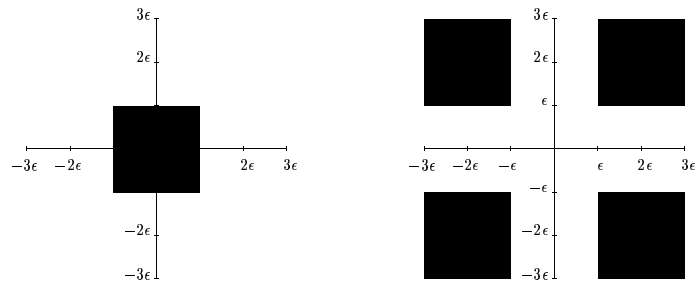
In this paper a modification of the BDS test is proposed which is more suitable for detecting conditional heteroskedasticity. It is designed to be sensitive to persistence in volatility rather than persistence in tranquility, as the BDS test. The power of the tests is then determined numerically for several chaotic and stochastic processes. The results indicate that the test has indeed larger power against ARCH structure, while an increase in power is also observed for some deterministic time series.

In the bootstrap literature it is often emphasized that one should consider test statistics which are, at least asymptotically, pivotal under the null hypothesis, that is, their distribution should not depend on any unknown parameters under the null, see e.g. Davidson and MacKinnon (1999). The Monte Carlo approach, by conditioning on a minimal sufficient statistic, automatically satisfies this requirement, even for finite sample sizes. After conditioning on a minimal and sufficient statistic under the null, the null distribution of any statistic, by construction, does not depend on any unknown parameters.

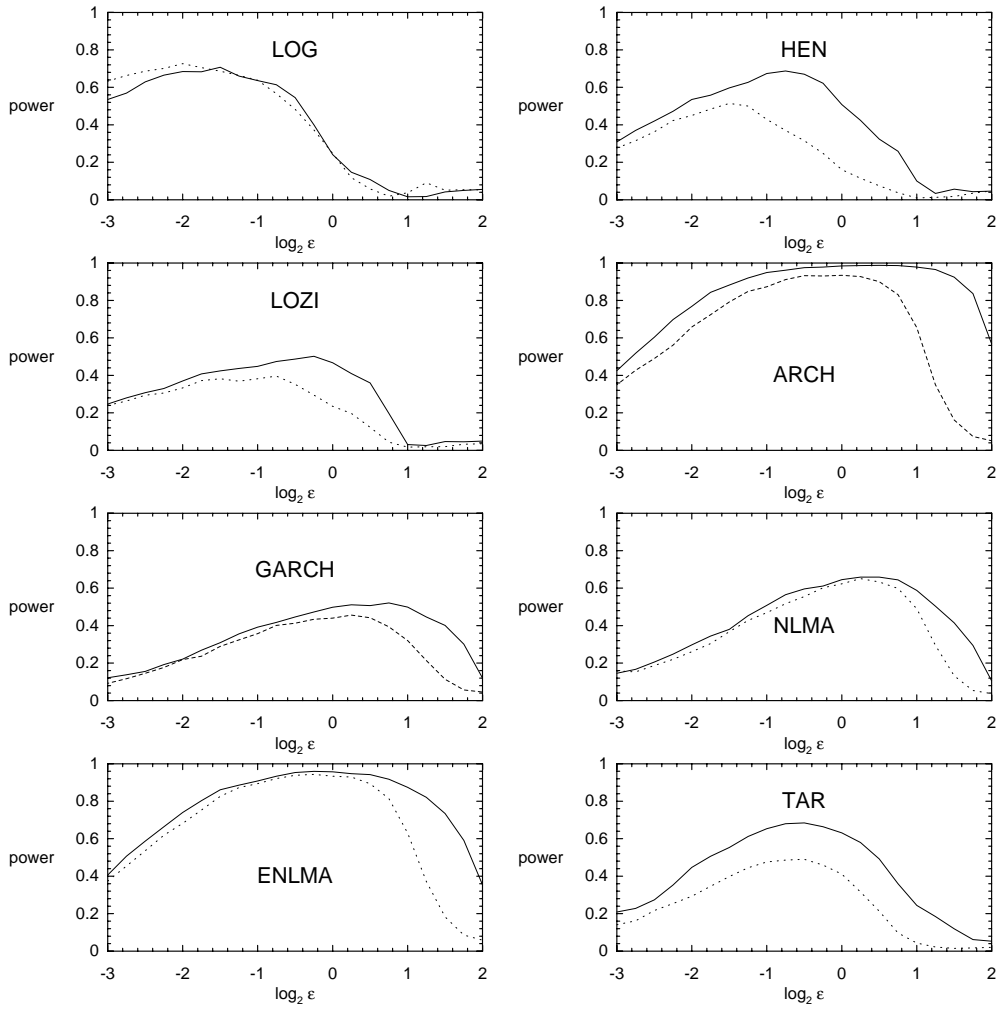
Finally, it is worth noting that the dual BDS test can be used as a diagnostic model verification tool (by applying it to the estimated residuals of a fitted parametric model) under the same conditions as given by Brock *et al.* (1996) for the original BDS test, such as the root- $N$  consistency of the model parameter estimators used. The conditions and proofs given by Brock *et al.* (1996) apply to the general case of a  $U$ -statistic based on any factorizing kernel. Since this includes the dual correlation integral, no adaptations are needed for the dual BDS test. De Lima (1996) presents some additional conditions under which correlation integral based test statistics for estimated residuals from dynamic models are free of nuisance parameters.

## References

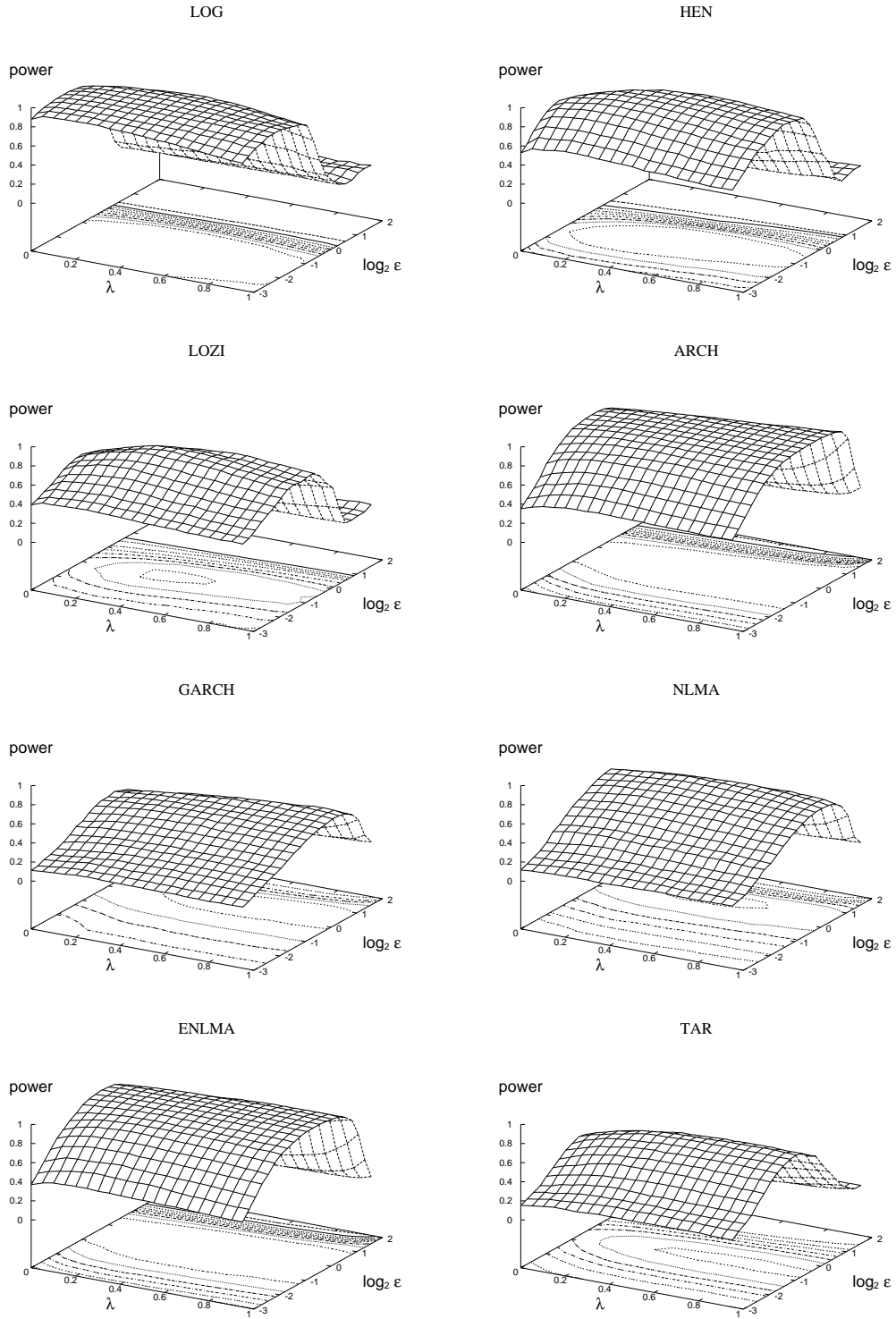
- Barnard, G. A. (1963). Discussion of Professor Bartlett's paper. *J. R. Statist. Soc. B*, **25**, 294.
- Besag, J. and Diggle, P. J. (1977). Simple Monte Carlo tests for spatial pattern. *Appl. Statist.*, **26**, number 3, 327–333.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, **31**, 307–327.
- Brock, W. A., Dechert, W. D. and Scheinkman, J. (1987). A test for independence based on the correlation dimension. Technical report 8702. Social Systems Research Institute, Univ. of Wisconsin, Madison.
- Brock, W. A., Dechert, W. D., Scheinkman, J. A. and B, LeBaron (1996). A test for independence based on the correlation dimension. *Econometrics Review*, **15**, number 3, 197–235.
- Brock, W. A., Hsieh, D. and LeBaron, B. (1991). *Nonlinear Dynamics, Chaos and Instability: Statistical Theory and Economic Evidence*. MIT: Cambridge MA.
- Davidson, R. and MacKinnon, J. G. (1999). The size distortion of bootstrap tests. *International Economic Review*, **40**, 487–508.
- de Lima, P. (1996). Nuisance parameter free properties of correlation integral based statistics. *Econometric Reviews*, **15**, 237–259.
- Diks, C. and Tong, H. (1999). A test for symmetries of multivariate probability distributions. *Biometrika*, **86**, 605–614.
- Engen, S. and Lillegård, M. (1997). Stochastic simulations conditioned on sufficient statistics. *Biometrika*, **84**, number 1, 235–240.
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, **50**, 987–1007.
- Grassberger, P. and Procaccia, I. (1983). Measuring the strangeness of strange attractors. *Physica D*, **9**, 189–208.
- Hope, A. C. A. (1968). A simplified Monte Carlo significance test procedure. *J. R. Statist. Soc. B*, **30**, 582–598.
- Johnson, D. and McLelland, R. (1998). A general dependence test and applications. *Journal of Applied Econometrics*, **13**, 627–644.
- Liu, T., Granger, C. W. J. and Heller, W. P. (1998). Using the correlation exponent to decide whether an economic time series is chaotic. *Journal of Applied Econometrics*, **7**, S25–S39.
- Marriott, F. H. C. (1979). Barnard's Monte Carlo tests: How many simulations? *Appl. Statist.*, **28**, number 1, 75–77.
- Tsay, R. (1986). Nonlinearity tests for time series. *Biometrika*, **72**, 461–466.



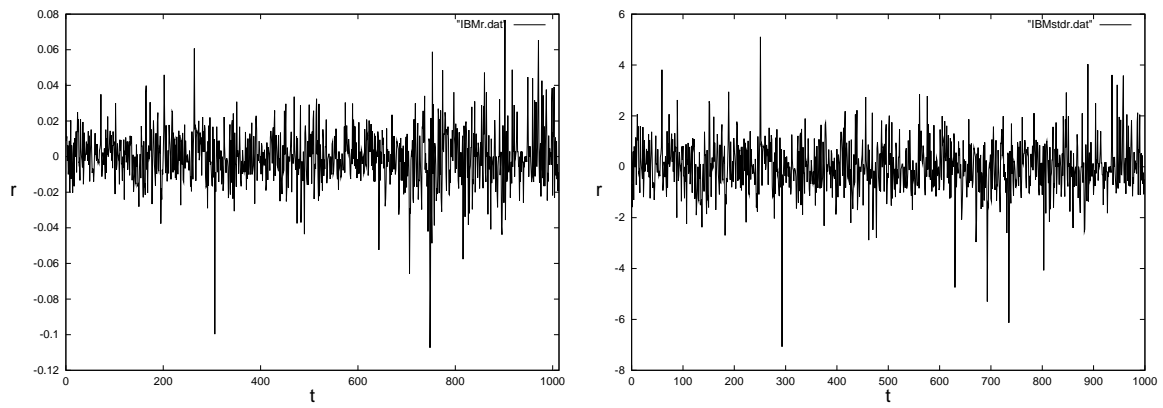
**Figure 1:** Regions in the plane in which  $(|X_t - X_s|, |X_{t+1} - X_{s+1}|)$  should fall to give a nonzero contribution to the correlation integral. The usual correlation integral counts the number of points falling within a distance  $\epsilon$  from the origin (dark area on the left). The dual correlation integral counts the number of points in the regions for which all coordinates are bigger than  $\epsilon$  in absolute value (dark areas on the right).



**Figure 2:** Estimated power at nominal level  $\alpha = 0.05$  as a function of  $\epsilon$  for the dual correlation integral (solid lines) and the usual correlation integral (dotted lines). Time series length  $N = 15$  for the chaotic models, and  $N = 250$  for the stochastic models, number of replications  $M = 200$ , embedding dimension  $m = 3$ . Rejection rates were determined over 1000 independent realisations.

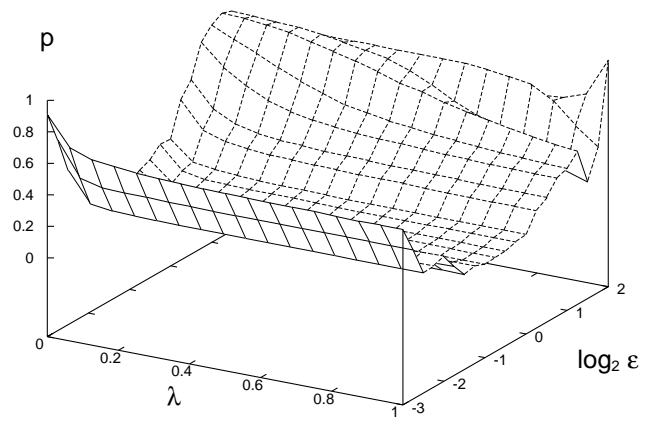


**Figure 3:** Power of the combined test statistic  $\lambda \widehat{C}_m(\epsilon) + (1 - \lambda) \widehat{C}'_m(\epsilon)$  as a function of  $\lambda$  and  $\epsilon$  for the logistic map, the Hénon map, and the ARCH and GARCH process. Other parameters are the same as for Fig. 2



**Figure 4:** *IBM returns and standardized GARCH(1,1) residuals.*





**Figure 5:** *Obtained p-values for the standardized IBM residuals, as a function of  $\epsilon$  and the weight  $\lambda$ .*