

Baruník, Jozef; Kraicová, Lucie

Working Paper

Estimation of long memory in volatility using wavelets

IES Working Paper, No. 33/2014

Provided in Cooperation with:

Charles University, Institute of Economic Studies (IES)

Suggested Citation: Baruník, Jozef; Kraicová, Lucie (2014) : Estimation of long memory in volatility using wavelets, IES Working Paper, No. 33/2014, Charles University in Prague, Institute of Economic Studies (IES), Prague

This Version is available at:

<https://hdl.handle.net/10419/120443>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Institute of Economic Studies, Faculty of Social Sciences
Charles University in Prague

Estimation of Long Memory in Volatility Using Wavelets

Jozef Baruník
Lucie Kraicová

IES Working Paper: 33/2014



Institute of Economic Studies,
Faculty of Social Sciences,
Charles University in Prague

[UK FSV – IES]

Opletalova 26
CZ-110 00, Prague
E-mail : ies@fsv.cuni.cz
<http://ies.fsv.cuni.cz>

Institut ekonomických studií
Fakulta sociálních věd
Univerzita Karlova v Praze

Opletalova 26
110 00 Praha 1

E-mail : ies@fsv.cuni.cz
<http://ies.fsv.cuni.cz>

Disclaimer: The IES Working Papers is an online paper series for works by the faculty and students of the Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Czech Republic. The papers are peer reviewed, but they are *not* edited or formatted by the editors. The views expressed in documents served by this site do not reflect the views of the IES or any other Charles University Department. They are the sole property of the respective authors. Additional info at: [ies@fsv.cuni.cz](http://ies.fsv.cuni.cz)

Copyright Notice: Although all documents published by the IES are provided without charge, they are licensed for personal, academic or educational use. All rights are reserved by the authors.

Citations: All references to documents served by this site must be appropriately cited.

Bibliographic information:

Baruník J., Kraicová L. (2014). “Estimation of Long Memory in Volatility Using Wavelets” IES Working Paper 33/2014. IES FSV. Charles University.

This paper can be downloaded at: <http://ies.fsv.cuni.cz>

Estimation of Long Memory in Volatility Using Wavelets

Jozef Baruník^a
Lucie Kraicová^b

^aInstitute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Smetanovo nábřeží 6, 111 01 Prague 1, Czech Republic
email: kraicova.l@seznam.cz

^bInstitute of Information Theory and Automation, Academy of Sciences of the Czech Republic: Pod Vodarenskou Veží 4, 182 00, Prague, Czech Republic
email: barunik@fsv.cuni.cz

September 2014

Abstract:

In this work we focus on the application of wavelet-based methods in volatility modeling. We introduce a new, wavelet-based estimator (wavelet Whittle estimator) of a FIEGARCH model, ARCH-family model capturing long-memory and asymmetry in volatility, and study its properties. Based on an extensive Monte Carlo experiment, both the behavior of the new estimator in various situations and its relative performance with respect to two more traditional estimators (maximum likelihood estimator and Fourier-based Whittle estimator) are assessed, along with practical aspects of its application. Possible solutions are proposed for most of the issues detected, including suggestion of a new specification of the estimator. This uses maximal overlap discrete wavelet transform, which improves the estimator performance, as we show in the experiment extension. Next, we study all the estimators in case of a FIEGARCH-Jump model, which brings interesting insights to their mechanism. We conclude that, after optimization of the estimation setup, the wavelet-based estimator may become an attractive robust alternative to the traditional methods.

Keywords: volatility, long memory, FIEGARCH, wavelets, Whittle, Monte Carlo
JEL: C13, C18, C51, G17

Acknowledgements: We would like to express our gratitude to Ana Perez, who provided us with her code for the Monte Carlo experiments using FIEGARCH MLE and FWE, and we gratefully acknowledge the support from the Czech Science Foundation under the 13-32263S project.

1 Introduction

Volatility is one of the most important concepts in finance. As a natural measure of risk of financial assets, it is frequently utilized by both academics and practitioners as a core variable in a wide range of applications. Hence, there is a great motivation to study patterns in its behavior and to improve our ability to forecast its future development. Over time, various approaches to reach these goals have been developed, all of them having some strengths and weaknesses.

In this paper we focus on one of these approaches that is widely used in practice, the parametric volatility modeling. Generally, the concept is based on approximating the true data generating processes in terms of reasonably complex structures (volatility models), which enable data interpretation and forecasting. Compared to non-parametric or semi-parametric methods, the models are fully specified by a finite set of parameters, which are estimated using empirical data and/or theory. Due to the wide range of their practical applications, the importance of the models (not only) in finance is substantial, and thus, requirements on their quality as well. Although the meaning of “quality” may differ in various contexts, for a good performance of any parametric model two things are always essential: first, it has to be well specified (no relevant patterns, such as long memory or asymmetries, should be omitted); second, an appropriate estimator of its parameters has to be applied. As they are complementary, simultaneous research in both these areas is relevant.

In our work, we address the latter characteristics, the optimal method of estimation. We introduce a new, wavelet-based estimator (wavelet Whittle estimator, WWE) of an econometric model that is frequently used for volatility analysis and forecasting (FIEGARCH model), and assess its properties. Although the presented method is not entirely new in econometrics, see e.g. Jensen (1999), to our knowledge, this is the first time its FIEGARCH specification is provided and studied. As will be explained in more detail in the next section, the estimator is based on maximizing an approximated likelihood function derived for the FIEGARCH process, where the approximation is done using wavelet transform; in our case using discrete wavelet transform (DWT) and, as an alternative, maximal overlap discrete wavelet transform (MODWT). The two transforms differ in some technical details (see e.g. Percival & Walden (2000)) but the main idea behind them is the same - representation of data in a way that enables their analysis in time and scale (frequencies) at the same time. Properties of these transforms can help us to get an estimator that is more robust to unpredictable time-localized irregularities in the data than its more traditional counterpart that is based on Fourier transform.

The ultimate goal of our work is to show the performance of the new estimator; discuss its advantages, disadvantages and applicability in practice; and to contribute to the literature on wavelet-based methods in economics. To enable a comprehensive study of the new method suggested, an extensive Monte Carlo experiment is designed. Based on its results both accuracy and empirical convergence of the new estimator under changing optional settings can be examined, as well as its relative performance with respect to two more traditional estimators: Fourier-based Whittle estimator (FWE, approximation of the likelihood function using Fourier transform) and maximum likelihood estimator (MLE, no approximation).

Although many questions remain open, our results are believed to contribute to the knowledge of the WWE behavior and applicability; especially due to the detailed analysis we provide and number of questions we raise. The work shows that even in the case of simulated data, which follow a pure FIEGARCH process, and thus do not allow to fully utilize the advantages of WWE over its traditional counterparts, the estimator can work reasonably well. When we focus on the individual parameters estimation, in terms of bias the performance is comparable, in some cases even better than that of the FWE, while in terms of RMSE the latter is usually better. As expected, the exact MLE mostly outperforms both of the Whittle estimators, with

rare exceptions in situations, which are of small importance for most practical applications in finance. Though, due to the computational complexity of the MLE, for large data sets FWE and WWE represent attractive alternatives. Concerning the experiment setup, we show that the presence and strength of long memory in the data affects both absolute and relative performance of the estimators, but the effect is mostly not strong enough to change their relative positions.

Next to the main results, significant part of the discussion focuses on the practical aspects of the WWE application. Based on observation of its behavior under changing experiment setup, several issues are detected. All of them are properly discussed and possible solutions are proposed. The most important outcome of the discrete wavelet transform-based WWE analysis is a detailed study of the empirical convergence of the estimator (with sample size and with level of wavelet transform applied - analogy to the local whittle estimation studied in Moulines et al. (2008)) and suggestion of an alternative WWE specification (based on maximal overlap discrete wavelet transform, MODWT) that should significantly improve its performance in all applications, not only in the context of FIEGARCH model estimation. In turn, this is shown in a subsequent Monte Carlo experiment. In this experiment we not only add the new WWE estimator, but we also extend it to enable forecasting performance analysis. The results show both the superiority of MODWT-based WWE over the DWT-based one and the overall similarity of the WWE and FWE performance that looks significantly better in terms of forecasting performance than in the case of individual parameters estimation. As a next step, we show the results of a Monte Carlo experiment, where the pure FIEGARCH model is enriched by jumps. It is shown that due to the nature of the FIEGARCH model and the implicit forecasting algorithms, the WWE cannot be as fully utilized to improve the estimates, as has been expected.

Focusing on more technical aspects, we also discuss the possibility that the relative performance of estimators is affected by optimization issues. It is shown that the identification of the parameters in the Whittle objective functions is problematic and may significantly increase the importance of the optimization algorithm choice for estimation accuracy. This choice is considered to be a non-trivial task, but when dealt with correctly, the estimates may be significantly improved.

Due to the pioneering nature of this work and the complexity of the problem concerned, the results presented are not intended to be directly projected to changes in estimation methods used in practice. For practitioners, the presented conclusions should be interesting as a message that given sufficient demand for further research in this area, new, possibly highly efficient methods based on wavelet transform could be available in the future. Though, the target group are the academics. It is believed that the results provided are a good basis for future research.

The paper is structured as follows: section 2 provides a literature review; section 3 introduces the FIEGARCH model and the individual estimators; in section 4 the setup of the Monte Carlo experiment is described and results are discussed (all tables and figures referred to are available in the appendices, either at the end of this paper or in the online supplement that is available at [<http://ies.fsv.cuni.cz/en/staff/kraicova>] or on request from the authors); next, in section 5 we present the extended experiment; in section 6 we compare our results with related literature; section 7 concludes.

2 Literature Review

Compared to the wide range of studies on semi-parametric Wavelet Whittle estimators (for relative performance of local FWE and WWE of ARFIMA model see e.g. Fay et al. (2009) or Frederiksen & Nielsen (2005) and related works), literature assessing performance of their parametric counterparts is not extensive. Though, results of the studies completed so far suggest

that the performance of WWE in parametric setting is an interesting and important research topic.

One of the important papers on parametric (DWT-based) WWE is that of Jensen (1999). The author introduces the WWE of ARFIMA process and compares its performance with that of FWE under various settings. In most cases, both estimators perform similarly, but when the true MA coefficients are close to the boundary of invertibility, the Fourier-based estimator deteriorates whereas the wavelet-based Whittle retains its accuracy. Speed of the two estimators is similar.

Percival & Walden (2000) describe a wavelet-based approximate MLE for both stationary and non-stationary fractionally differenced processes, discuss its properties and show its performance on a simulated (Fractionally differenced process) as well as on real data. Relatively good performance is demonstrated even on very short samples (128 observations).

Whitcher (2004) applies WWE based on a discrete wavelet packet transform (DWPT) to a seasonal persistent process. Good performance of the estimator is shown in both Monte Carlo experiment and real data application.

Heni & Mohamed (2011) use the approach introduced in Whitcher (2004) and apply it on a FIGARCH-GARMA model. Using this specification, nominal daily exchange rates of the Tunisian dinar relative to the US dollar, the Euro and the Japanese yen are modeled. Various filters are applied, all with support larger than 8, which is argued to be appropriate given the process characteristics. Good forecasting performance of the model-estimator set is shown. Further application can be seen e.g. in Gonzaga & Hauser (2011).

All these works estimate parameters in various models using the WWE. Although none of them focuses on a FIEGARCH model, they can be used to compare the performance of the WWE as applied in different situations, as well as they provide an inspiration for the WWE setup. Yet, our work focuses on FIEGARCH modeling, and thus it should extend results of a study, where this particular model is estimated using estimators different from the WWE, such as the paper by Perez & Zaffaroni (2008). In their work the authors compare finite sample properties of the MLE and two FWEs (univariate and bivariate) for (FI)EGARCH models on simulated series of length $T = 2^9, 2^{10}$ and 2^{11} ; the MLE represents the traditional approach, while the two FWEs are introduced as possible alternatives, performance of which is to be tested. It is concluded that “MLE is more efficient, but bivariate FWE sometimes performs comparably”. FWE becomes dominant in cases of processes close to nonstationarity. For EGARCH(1,1)¹ FWE is, in most cases, more biased than the other two estimators, but in some cases it even outperforms the MLE. For EGARCH(1,2) Whittle estimators perform worse than MLE in terms of RMSE, but outperform the former method in terms of bias. For FIEGARCH(1,d,2) Whittle estimators are always underestimating the long memory parameter, whereas FWE becomes dominant at high levels of persistence in the time series. Compared to MLE, Whittle methods seem to offer lower bias at the cost of higher RMSE.

Our work represents a direct extension of this paper. In an experiment setup mirroring that of Perez & Zaffaroni (2008), the WWE is introduced as an alternative to the methods above. The main focus is on its performance compared to that of MLE and univariate FWE, which are explicitly included in the Monte Carlo experiment, while comparison with results for the bivariate FWE in Perez & Zaffaroni (2008) is also possible due to the same estimation setup.

¹using definition of EGARCH used in this paper, which is equivalent to EGARCH(1,0) in the original paper

3 FIEGARCH Model and Parameter Estimators

3.1 FIEGARCH Model

Despite the extremely wide spectrum of processes generating financial returns time series, there are some stylized features which many of them have in common. They have been detected over years of financial market analysis and have shaped the means of financial time series modeling. One of these features, which initiated a real breakthrough in financial modeling, is their time-variant, past-dependent volatility. A well known response to this observed pattern was the introduction of conditional volatility models by Engle (1982). Over time, performance of these models (ARCH family models) in practical applications has demonstrated the importance of conditional volatility in time series analysis and feasibility of direct volatility estimation and forecasting. Although several alternative concepts based on explicitly modeled volatility have been developed since Engle (1982), generalized ARCH models are still among those best performing in practical applications. This makes it relevant to study the performance of new parameter estimators in their context. In our study we focus on one of the generalizations of the basic ARCH model, the FIEGARCH model, where the log-returns $\{\epsilon_t\}$ are modeled conditionally on their past realizations as:

$$\epsilon_t = z_t h_t^{1/2} \quad (1)$$

$$\ln(h_t) = \omega + \Phi(L)g(z_{t-1}) \quad (2)$$

$$g(z_t) = \theta z_t + \gamma[|z_t| - E(|z_t|)] \quad (3)$$

where:

$$\Phi(L) = (1 - L)^{-d}[1 + \alpha(L)][\beta(L)]^{-1} \quad (4)$$

$$\alpha(L) = 1 + \alpha_{[2]}(L) = 1 + \sum_{i=2}^p \alpha_i L^{i-1} \quad (5)$$

$$\beta(L) = 1 - \sum_{i=1}^q \beta_i L^i \quad (6)$$

$$(1 - L)^d = 1 - d \sum_{k=1}^{\infty} \Gamma(k - d) \Gamma(1 - d)^{-1} \Gamma(k + 1)^{-1} L^k \quad (7)$$

and where for FIEGARCH(1,d,2), i.e. q=1, p=2, we have:

$$\alpha(L) = 1 + \alpha L \quad (8)$$

$$\beta(L) = 1 - \beta L. \quad (9)$$

The $\{z_t\}$ (iid. unobservable standardized innovations) and $\{\epsilon_t\}$ (observable log-returns series) are discrete-time real valued processes satisfying $E_{t-1}(z_t) = 0$, $E_{t-1}(z_t^2) = 1$, $E_{t-1}(\epsilon_t) = 0$, $E_{t-1}(\epsilon_t^2) = h_t$, L is a lag operator $L^i g_t = g_{t-i}$ and $\Gamma(\cdot)$ is the gamma function. The polynomials $\alpha(L)$ and $\beta(L)$ have no zeros in common, their roots are outside the unit circle, $\theta\gamma \neq 0$ and $d < 0.5$. The $\{h_t\}_{t=1}^T$ is the conditional log-variance process dependent on the past innovations. See Perez & Zaffaroni (2008). Any FIEGARCH(q,d,p) process is then fully determined by the number of parameters, their values and distribution of the standardized innovations $z(t)$. Concerning the last factor, the three most frequent assumptions in the literature are the Standard normal distribution $N(0,1)$ providing a convenient estimation environment, Student-t distribution assuming thicker tails, and Generalized Error Distribution (GED) with parameter v determining the tail thickness. Normal distribution is nested as a special case of GED for $v=2$.

By its construction the model captures the following stylized features of the real financial time series data: short-term temporal variation in financial returns volatility (Volatility clustering), long-term temporal variation in financial returns volatility (Long memory), negative relationship between past returns and volatility (Leverage effect) and fat-tailed sample distribution of returns. For plots of a FIEGARCH process with three different levels of long memory see Appendix H (online).

While correct model specification is one thing, feasibility of estimation of its parameters is another. In general, estimation of the FIEGARCH model can be carried out by various methods. Below, those considered in this work (the benchmark estimators MLE and FWE, and the newly introduced WWE) are described together with practical aspects of their application.

3.2 Maximum Likelihood Estimator

Maximum likelihood estimation is one of the two benchmark methods we use. It is a traditional, in many cases highly efficient method applicable to a wide range of models, including FIEGARCH model. For a general Gaussian process $\{x_t\}_{t=1}^T$ (usual basic setting for ARCH family models that we apply), the maximum likelihood estimator (MLE) is defined as

$$\hat{\zeta}^{MLE} = \operatorname{argmin}_{\zeta \in \Theta} \ln L(x, \zeta), \quad (10)$$

where $\ln L(x, \zeta)$ is the negative log-likelihood function

$$\ln L(x, \zeta) = \frac{T}{2} \ln(2\pi) + \frac{1}{2} \ln \left| \sum_T \right| + \frac{1}{2} \left[(x^*)^{trans} \left(\sum_T \right)^{-1} x^* \right], \quad (11)$$

where $x^* = x - E(x)$, \sum_T is the covariance matrix of $\{x_t\}_{t=1}^T$, $|\sum_T|$ is its determinant and ζ is the vector of parameters determining $E(x)$ and \sum_T .

Despite the favorable properties of the MLE, there are some issues limiting its practical applicability. The usual problem is that we have to deal with the inversion of the covariance matrix of the process and with its determinant. Although it may not be a big problem when the matrix is diagonal or sufficiently sparse, in cases of dense covariance matrices (characteristic for long memory processes) it may be extremely time demanding, or even unfeasible in case of large datasets. Moreover, as discussed in (Beran (1994), chapter 5), solution may be even unstable in the presence of long memory, when the covariance matrix is close to singularity. Next, the deduction of the mean in x^* may be problematic in practical applications, since whenever the mean is not known, it has to be estimated. The efficiency and bias of the estimator of the mean contributes to the efficiency and bias of the MLE. In case of long-memory processes it can cause significant deterioration of the MLE. This has been shown e.g. in Cheung & Diebold (1994). Both these issues have motivated construction of alternative estimators, usually formulated as approximate MLE and defined by an approximated log-likelihood function. See for example Beran (1994) or Nielsen & Frederiksen (2005).

Next, since the assumption of a specific distribution (in our case Gaussian) is usually too restrictive for practical applications, it is important to study the estimator in situations when it is constructed for some (Gaussian) process but applied to a different (non-Gaussian) process. The resulting estimator is called Quasi-Maximum Likelihood Estimator (QMLE). In the context of GARCH processes with non-normal error distribution this has been studied by Bollerslev & Wooldridge (1992), who show that the estimator remains consistent, but loses efficiency. The efficiency loss, as argued in Engle & Gonzalez-Rivera (1991), is rather small for symmetric t-distributed processes, but can be significant under asymmetric distribution. As discussed in

Bollerslev & Wooldridge (1992), standard test statistics become biased and to ensure valid inference, their robustified counterparts, such as those proposed by the authors, should be used. Unlike the case of GARCH model, the asymptotic theory for EGARCH and FIEGARCH QMLE is not yet available. Though, the (Q)MLE is the basic tool for their estimation suggested by their authors in Nelson (1991) and Bollerslev & Mikkelsen (1996) respectively and the properties of the estimator are assessed via empirical studies; see e.g. Baillie et al. (1996), Bollerslev & Mikkelsen (1996) or Perez & Zaffaroni (2008).

To derive the FIEGARCH (Q)MLE, we just rewrite the general formula for the negative log-likelihood function (11). Since we assume a Gaussian, zero-mean process of independent variables (log-returns) $\{\epsilon_t\}_{t=1}^T$, i.e. assuming $E(\epsilon_t) = 0$ and a diagonal \sum_T with elements σ_t^2 , which are the unconditional variances of ϵ_t . Then without affecting the optimization results, the function (11) can be reduced to

$$L^{MLE}(x, \zeta) = \sum_{t=1}^T \left\{ \ln [\sigma_t^2] + \frac{x_t^2}{\sigma_t^2} \right\}. \quad (12)$$

Next, conditionally on the past realizations of $\{\epsilon_t\}$, the $\{\epsilon_t, \epsilon_j; i, j = 1, \dots, T; i \neq j\}$ remain uncorrelated; the respective conditional covariance matrix is diagonal with elements h_t ; and the determinant again reduces to sum of the diagonal terms. Then, under Gaussianity assumption on z_t in (1), we can use the reduced $L^{MLE}(x)$ from above and define the reduced negative conditional log-likelihood function for the process $\{\epsilon\}_{t=1}^T$ as

$$L^{MLE}(\epsilon, \zeta | \psi_{t-1}) = \sum_{t=1}^T \left\{ \ln[h_t(\zeta)] + \frac{\epsilon_t^2}{h_t(\zeta)} \right\}, \quad (13)$$

where ζ is a vector of parameters in the equation for conditional log-variance in (1), ψ_{t-1} is the information set at time t , containing all past observations available at time t and $h_t(\zeta)$ is the conditional variance defined by the FIEGARCH model. Then the MLE estimator is defined as $\hat{\zeta}^{MLE} = \underset{\zeta \in \Theta}{\operatorname{argmin}} L^{MLE}(\epsilon, \zeta | \psi_{t-1})$, where Θ is the parameter space.

Despite the simple formula, some practical issues arise because of the need to model the conditional volatility process during the optimization. As discussed in Bollerslev & Mikkelsen (1996), since in the context of GARCH models the forecasted volatility depends on the past forecasts, the log-likelihood function becomes dependent on initial conditions I_0 , the basis for conditional variance recursion. So the MLE based on sample $\{\epsilon_t, t = 1, 2, \dots, T\}$ relies on the maximization of conditional log-likelihood function $\log L(\zeta, \epsilon | I_0)$, where $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_T)$ is a T -dimensional vector of observations of the process ϵ_t and I_0 stands for initial conditions. Since, under the general specification of the model, h_t depends on $\{h_{t-i}, i = 1, 2, \dots, \infty\}$, a truncation is needed. In case of a short memory process, the dependence on the past declines rapidly, which enables to use a relatively small number of initializing values. In contrast, when long memory is present, large number of pre-sample “observations” is necessary to prevent a significant loss of information about long-run dependencies in the process. Following the approach of Bollerslev & Mikkelsen (1996), in our Monte Carlo experiment, the pre-sample conditional volatilities are set to the sample volatility, conditional mean is assumed to be known and equal to zero (later on also the non-zero case is assumed) and for both the simulation and estimation the truncation is done at lag 1000. Using these initializing values, for each set of parameters considered by the optimization algorithm the whole vector of conditional volatilities is estimated, together with the implied squared returns needed for further iterations. Given the complexity of this procedure, the method is significantly time consuming.

3.3 Fourier-based Whittle Estimator

Fourier-based Whittle Estimation is our second benchmark method. In general, the respective estimator (FWE) is a spectral-based counterpart of the MLE, where the problematic terms in the log-likelihood function (with the possibly dense covariance matrix) the $[det(\sum_T)]$ and the $[x *^{trans} (\sum_T)^{-1} x^*]$, are replaced by their asymptotic frequency domain representations. The link between time domain and frequency domain is defined by means of Fourier transform. The approximation is based on a study of eigenvectors and respective eigenvalues of the covariance matrix leading to a conclusion that the matrix can be diagonalized by means of Fourier transform. Orthogonality of the Fourier transform projection matrix then allows to achieve the approximation by means of multiplications by identity matrices, simple rearrangements and approximation of integrals by Riemann sums, see Beran (1994). The approximated Whittle negative log-likelihood function for estimation of parameters in the distribution function of the process $\{x\}_{t=1}^T$ under Gaussianity assumption can be written as:

$$\tilde{L}_W(x, \zeta) = \frac{2}{(2\pi)} \sum_{j=1}^{m^*} \ln [f(\lambda_j, \zeta)] \frac{2\pi}{T} + \frac{2}{(2\pi)^2} \sum_{j=1}^{m^*} \frac{I(\lambda_j, x)}{f(\lambda_j, \zeta)} \frac{2\pi}{T}, \quad (14)$$

which can be reduced to

$$\tilde{L}_W(x, \zeta) = \frac{1}{T} \sum_{j=1}^{m^*} \left[\ln [f(\lambda_j, \zeta)] + \frac{I(\lambda_j, x)}{f(\lambda_j, \zeta)} \right]; \quad \lambda_j = 2\pi j/T \quad (15)$$

$$j = 1, 2, \dots, m^*; \quad m^* = \max \{m \in \mathbb{Z}; m \leq (T-1)/2\}, \quad i.e. \lambda_j < \pi, \quad (16)$$

where $f(\lambda, \zeta)$ is the spectral density of process x decomposing the process variance with respect to frequency:

$$cov(x_t, x_s) = \frac{\sigma^2}{2\pi} \int_0^{2\pi} f(\lambda, \zeta) e^{i(s-t)\lambda} d\lambda; \quad (17)$$

see Percival & Walden (2000); and $I(\lambda_j, x)$ is the value of periodogram of x at j th Fourier frequency, whereas there are at most $m^* = (T-1)/2$ frequencies available, with the highest frequency $\lambda_1 = 2\pi/T$ and the smallest frequency detectable $\lambda_{m^*} = 2\pi(T-1)/2T = \pi - (\pi/T)$:

$$I(\lambda_j, x) = (2\pi T)^{-1} \left| \sum_{t=1}^T x_t \exp(i\lambda_j t) \right|^2; \quad \lambda_j = 2\pi j/T; \quad j = 1, 2, \dots, m^*. \quad (18)$$

The respective Fourier-based Whittle estimator is defined as

$$\hat{\zeta}^{FWE} = \underset{\zeta \in \Theta}{\operatorname{argmin}} \tilde{L}_W(x, \zeta). \quad (19)$$

For a detailed FWE derivation see e.g. Beran (1994).

It can be shown that the FWE has the same asymptotic distribution as the exact MLE, hence is asymptotically efficient for Gaussian processes. See Fox & Taquq (1986), Dahlhaus (1989) and Dahlhaus (2006). In the literature, FWE is frequently applied to both Gaussian and non-Gaussian processes (equivalent to QMLE), whereas even in the later case, both finite sample and asymptotic properties of the estimator are often shown to be very favorable and the complexity of the computation depends on the form of the spectral density of the process. Next to a significant reduction in estimation time, the FWE also offers an efficient solution for

long-memory processes with an unknown mean, which can impair efficiency of the MLE. By elimination of the zero frequency coefficient FWE becomes robust to addition of constant terms to the series, and thus in case, when no efficient estimator of the mean is available, FWE can become an appropriate choice even for time series where the MLE is still computable within reasonable time.

Concerning the FIEGARCH estimation, the FIEGARCH-FWE is, to the authors' knowledge, the only one out of the three estimators considered in this work, for which an asymptotic theory is currently available. The theory is derived in Zaffaroni (2009) for a whole class of exponential volatility processes; both strong consistency and asymptotic normality are established, even though the estimator works as an approximate QMLE of a process with an asymmetric distribution, rather than an approximate MLE. This is due to the need to adjust the model to enable derivation of the spectral density of the estimated process. More specifically, as discussed and derived in Perez & Zaffaroni (2008) and Zaffaroni (2009), it is necessary to rewrite the model in a signal plus noise form

$$x_t = \ln(\epsilon_t^2) = \ln(z_t^2) + \omega + \sum_{s=0}^{\infty} \Phi_s g(z_{t-s-1}) \quad (20)$$

$$g(z_t) = \theta z_t + \gamma[|z_t| - E(|z_t|)] \quad (21)$$

$$\Phi(L) = (1 - L)^{-d} [1 + \alpha_{[2]}(L)][\beta(L)]^{-1} \quad (22)$$

$$\text{where for FIEGARCH}(1, d, 2): \quad (23)$$

$$\alpha(L) = 1 + \alpha L \quad (24)$$

$$\beta(L) = 1 - \beta L \quad (25)$$

$$(1 - L)^d = 1 - d \sum_{k=1}^{\infty} \Gamma(k - d) \Gamma(1 - d)^{-1} \Gamma(k + 1)^{-1} L^k, \quad (26)$$

for which the spectral density can be derived. The $\{z_t\}$ and $\{\epsilon_t\}$ denote discrete-time real valued processes satisfying $E_{t-1}(z_t) = 0$, $E_{t-1}(z_t^2) = 1$, $E_{t-1}(\epsilon_t) = 0$, $E_{t-1}(\epsilon_t^2) = h_t$ and Φ_s is the coefficient s of the polynomial $\Phi(L)$. The process $\{x_t\}_{t=1}^T$ then enters the FWE objective function instead of the process $\{\epsilon_t\}_{t=1}^T$. For the detailed derivation of the transformed process and for the formula for its spectral density see Appendix A.

Having the formula for the spectral density, the only term remaining to be specified is the periodogram of $x_t = \ln(\epsilon_t^2)$ at Fourier frequencies. In MATLAB the respective coefficients can be calculated using Fast Fourier Transform (FFT), that gives us coefficients $I_T(j)$, and appropriate rescaling reflecting switch from evaluation at simple frequencies to evaluation at Fourier frequencies. The periodogram takes the form

$$I_T(\lambda_j, x) = \frac{I_T(j)}{2\pi} = \frac{\left| \sum_{t=1}^T (x_t - \hat{\mu}_x) e^{-i2\pi j t / T} \right|^2}{2\pi T} \quad (27)$$

$$\lambda_j = 2\pi j / T; \quad j = 1, 2, \dots, m^*; \quad m^* = \max\{m \in \mathbb{Z}; \quad m \leq (T - 1)/2\}, \quad (28)$$

where $\hat{\mu}_x$ is the estimate of the mean of the process x . Now, the objective function (Equation 16) can be minimized with respect to all its parameters and the vector of Fourier-based Whittle estimates can be obtained.

3.4 Wavelet-Whittle Estimator

Although, based on Perez & Zaffaroni (2008), the FWE seems to be a good alternative to MLE for the FIEGARCH model estimation in the case of application on pure FIEGARCH processes, its use on real data may be, in some cases, problematic. This is because the FWE performance depends on the accuracy of the spectral density estimation using periodogram, which may be impaired by various time-localized patterns in the data diverging from the underlying FIEGARCH process. Motivated by the advances in the spectral density estimation using wavelets, we propose a wavelet-based estimator, the Wavelet Whittle Estimator (WWE), as an alternative. As in the case of Fourier-based Whittle, the Wavelet Whittle effectively overcomes the problem with the $[det(\sum_T)]$ and the $[x *^{trans} (\sum_T)^{-1} x *]$ by means of transformation. The difference is that instead of using discrete Fourier transform (DFT, projection of the time series on periodic functions with infinite support, complex exponentials, at Fourier frequencies $\lambda_j = (2\pi j)/T$; $j = 1, \dots, m^* \leq (T-1)/2$), the data are subject to discrete wavelet transform (DWT, projection on a finite-support function, wavelet, at various levels of translation (k) and dilation (j), $\{k, j \in Z\}$).

3.4.1 Discrete Wavelet Transform

To provide a good introduction to the WWE, it seems to be reasonable to start with a brief description of the wavelet transform that determines its properties and makes it different from the FWE. The core of any wavelet transform is a wavelet system, whose construction, together with means of the projection applied, determine the characteristics of the transformed data. A basic wavelet system can be defined as a set $\{\{\varphi_{j_0,k}(x)\}, \{\psi_{j,k}(x)\}; k \in Z, j = j_0, j_0 - 1, j_0 - 2, \dots\}$ creating an orthonormal basis (ONB) in $L_2(R)$; which means that any function $f \in L_2(R)$ can be expressed as

$$f(x) = \sum_k \alpha_{j_0,k} \varphi_{j_0,k}(x) + \sum_{j=-\infty}^{j_0} \sum_k \beta_{j,k} \psi_{j,k}(x) \quad (29)$$

$\alpha_k = \int f(x) \overline{\varphi_{j_0,k}(x)} dx$, $\beta_{j,k} = \int f(x) \overline{\psi_{j,k}(x)} dx$, where the elements α_k , $\beta_{j,k}$, $\varphi(x)$ and $\psi(x)$ are called scaling coefficients, detail (wavelet) coefficients, scaling function (father wavelet) and wavelet function (mother wavelet) respectively, and the translated and dilated transformations of the mother wavelet $\psi_{j,k}(x) = s(j)^{(-0.5)} \psi(s(j)x - k)$ are called daughter wavelets.

The basic conditions for $\psi(x)$ to be a valid mother wavelet are that $\int \psi(x) dx = 0$ and $\int \psi^2(x) dx = 1$, while the usual requirement is also the “admissibility” condition $\int \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty$, where $\hat{\psi}$ is the Fourier transform of ψ . This condition ensures that we can reconstruct the original time series from its transform. For complete conditions on $\varphi(x)$, $\psi(x)$ to be valid father and mother wavelets in the context various subsets of $L_2(R)$ and for other details concerning construction of wavelet systems see Hardle et al. (1998), for examples of wavelets see Figure 17 in Appendix J.

Next, any method that decomposes original data using the wavelet system and expresses them in terms of coefficients $\{\alpha_k, \beta_{j,k}\}$ and functions $\{\varphi(x), \psi(x)\}$ defined above, is a wavelet transform. In case of $j \in Z$, as applied in this thesis, we speak about a discrete wavelet transform (DWT), while for $j \in R$ the transform is continuous (CWT). By tradition, the default choice of scales is $\{2^{-j}; j \in Z\}$, thus the standard (traditional, not modified) DWT can be defined in terms of the wavelet expansion (29) with scaling defined as $s(j) = 2^{-j}$ (i.e. “scale j ” refers to the scaling 2^{-j}). The DWT coefficients are obtained using two-channel filter banks and down-sampling, so that at each level of decomposition j of a series of length M we get $M/2^j$ DWT

coefficients, see e.g. Jensen (2000). These coefficients can be in turn used for decomposition of the variance σ^2 of the process x :

$$\sigma^2 = E(x^2) - [E(x)]^2 = \frac{\|x\|^2}{T} - [E(x)]^2 = \frac{\sum_{j=1}^J \|W_j\|^2 + \|V_J\|^2}{T} - [E(x)]^2, \quad (30)$$

where W_j ; $j = 1, \dots, J$ and V_J are vectors of wavelet and scaling coefficients respectively and the $[E(x)]^2$ can be estimated using the squared sample mean \bar{x}^2 , or using the true squared mean whenever known. Alternatively, we can use the coefficients for estimating the spectral density $f(\lambda, \zeta)$ of x using the relationship:

$$\frac{\|W_j\|^2}{T} = \frac{\sigma_{W,j}^2}{2^j} \approx 2 \int_{1/2^{j+1}}^{1/2^j} f(\lambda, \zeta); \quad j = 1, 2, \dots, J \quad (31)$$

$$\frac{\|V_J\|^2}{T} - \bar{x}^2 = \frac{\sigma_{V,J}^2}{2^J} - \left(1 - \frac{1}{2^J}\right) \bar{x}^2 \approx 2 \int_0^{1/2^{J+1}} f(\lambda, \zeta), \quad (32)$$

where $\sigma_{W,j}^2$ and $\sigma_{V,J}^2$ are the sample variances of the wavelet and scaling coefficients respectively.

3.4.2 Wavelet Whittle Estimator

Analogically to the FWE, we use the relationship between wavelet coefficients and the spectral density of x to approximate the problematic likelihood function of the MLE. The main advantage is, compared to the FWE, that the wavelets have limited support, and thus, the coefficients are not determined by the whole time series, but by a limited number of observations only. This increases the robustness of the resulting estimator to irregularities in the data well localized in time, such as jumps. These may be poorly detectable in the data, especially in the case of strong long memory that itself creates jump-like patterns, but at the same time, their presence can significantly impair the FWE performance. On the other hand, the main disadvantages of using the DWT are the restriction to sample lengths 2^j and the low number of coefficients at the highest levels of decomposition j .

Skiping the details of wavelet-based approximation of the covariance matrix and the detailed WWE derivation, which can be found e.g. in Percival & Walden (2000), the Wavelet-Whittle objective function can be defined as

$$\ln WWL(x, \zeta) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \ln |O^T \Lambda_T O| - \frac{1}{2} [x^T (O^T \Lambda_T O) x] \quad (33)$$

$$\begin{aligned} \ln WWL(W_{j,k}, \zeta) = & -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \ln (|O^T| |\Lambda_T| |O|) - \\ & - \frac{1}{2} [(Ox)^T (\Lambda_T)^{-1} (Ox)] \end{aligned} \quad (34)$$

$$\begin{aligned} \ln WWL(W_{j,k}, \zeta) = & -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \ln (|\Lambda_T|) - \\ & - \frac{1}{2} [W_{j,k}^t (\Lambda_T)^{-1} W_{j,k}] \end{aligned} \quad (35)$$

$$\begin{aligned} \ln WWL(W_{j,k}, \zeta) = & -\frac{T}{2} \ln(2\pi) - \\ & - \frac{1}{2} \sum_{j=1}^J \left[N_j \ln \int_{1/2^{j+1}}^{1/2^j} 2 * 2^j f(\lambda, \zeta) d\lambda + \sum_{k=1}^{N_j} \frac{W_{j,k}^2}{\int_{1/2^{j+1}}^{1/2^j} 2 * 2^j f(\lambda, \zeta) d\lambda} \right] \end{aligned} \quad (36)$$

Where $W_{j,k}$ are the wavelet (detail) coefficients, ($O = O_{orth}$) is an orthogonal matrix representing the wavelet transform $x = O_{orth}^T x_{transf} = O_{orth}^T O_{orth} x = I_N x = x$ and Λ_T is a diagonal matrix with elements $\{C_1, C_1, \dots, C_1, C_2, \dots, C_J\}$, where for each level j we have N_j elements $(C_j = \int_{1/2^{j+1}}^{1/2^j} 2 * 2^j f(\lambda, \zeta) d\lambda)$, where N_j is the number of DWT coefficients at level j . The Wavelet Whittle Estimator is then defined as

$$\hat{\zeta}_j^{WW} = \operatorname{argmax}_{\zeta \in \Theta} \ln WWL(W_{j,k}, \zeta) \quad (37)$$

and as in the case of Fourier-based Whittle, the problem can be simplified to

$$\hat{\zeta}_j^{WW} = \operatorname{argmax}_{\zeta \in \Theta} \ln WWL(W_{j,k}, \zeta) = \operatorname{argmin}_{\zeta \in \Theta} L_{WW}(W_{j,k}, \zeta) \quad (38)$$

$$L_{WW}(W_{j,k}, \zeta) = \sum_{j=1}^J \left[N_j \ln \int_{1/2^{j+1}}^{1/2^j} 2 * 2^j f(\lambda, \zeta) d\lambda + \sum_{k=1}^{N_j} \frac{W_{j,k}^2}{\int_{1/2^{j+1}}^{1/2^j} 2 * 2^j f(\lambda, \zeta) d\lambda} \right]. \quad (39)$$

Similarly to the Fourier-based Whittle, the estimator is equivalent to a (Q)MLE of parameters in the probability density function of wavelet coefficients under normality assumption. At this time, the negative log-likelihood function can be rewritten as a sum of partial negative log-likelihood functions respective to individual levels of decomposition, whereas at each level, the coefficients are assumed to be homoskedastic, while across levels the variances differ. All wavelet coefficients are assumed to be (approximately) uncorrelated (the DWT approximately diagonalizes the covariance matrix), which requires an appropriate filter choice. Next, in our work the variance of scaling coefficients is excluded. This is possible due to the WWE construction, the only result is that the part of the spectrum respective to this variance is neglected in the estimation. This is optimal especially in cases of long-memory processes, where the spectral density goes to infinity at zero frequency, and where the sample variance of scaling coefficients may be significantly inaccurate estimate of its true counterpart due to the embedded estimation of the process mean.

3.4.3 Full vs. Partial Decomposition

Similarly to the omitted scaling coefficients, we can exclude any number of the sets of wavelet coefficients at the highest and/or lowest levels of decomposition. What we get is a parametric analogy to the Local Wavelet Whittle Estimator (LWWE) developed in Wornell & Oppenheim (1992) and studied by Moulines et al. (2008), who derive the asymptotic theory for LWWE with general upper and lower bound for levels of decomposition $\{j \in \langle L, U \rangle; 0 \leq L < U \leq J_{max}\}$.

Although, in the parametric context, it seems to be natural to use the full decomposition, there are several features of the WWE causing that it may not be optimal. To make the point, let's rewrite the WWE objective function as:

$$L_{WW}(W_{j,DWT}, \zeta) = \sum_{j=1}^J N_j \left[\ln \tilde{\sigma}_{W,j,DWT}^2(\zeta) + \frac{\hat{\sigma}_{W,j,DWT}^2}{\tilde{\sigma}_{W,j,DWT}^2(\zeta)} \right], \quad (40)$$

where $\tilde{\sigma}_{W,j,DWT}^2(\zeta)$ is the theoretical variance of j^{th} level DWT coefficients and $\hat{\sigma}_{W,j,DWT}^2$ is its sample counterpart, ζ is the vector of parameters in $\tilde{\sigma}_{W,j,DWT}^2(\zeta)$ and $\{W_{j,DWT}; j = 1, \dots, J_{max}\}$ are vectors of DWT coefficients used to calculate $\hat{\sigma}_{W,j,DWT}^2$. Using the definition of wavelet variance $v_j^2 = \int_{1/2^{j+1}}^{1/2^j} 2 * f(\lambda, \zeta) d\lambda = \frac{\sigma_{W,j,DWT}^2}{2^j}$; $j = 1, 2, \dots, J$ and using the fact that the optimization

problem does not change by multiplying the right-hand side term by $1/N^*$, the total number of coefficients used in the estimation, the $L_{WW}(W_{j,DWT}, \zeta)$ above is equivalent to

$$L_{WW}(W_{j,DWT}, \zeta) = \sum_{j=1}^J \frac{N_j}{N^*} \left[\ln \tilde{\sigma}_{W,j,DWT}^2(\zeta) + \frac{\hat{v}_{W,j,DWT}^2}{\tilde{v}_{W,j,DWT}^2(\zeta)} \right], \quad (41)$$

where $\tilde{v}_{W,j,DWT}^2(\zeta)$ is the theoretical j^{th} level wavelet variance and $\hat{v}_{W,j,DWT}^2$ is its estimate using DWT coefficients; ζ is the vector of parameters in $\tilde{v}_{W,j,DWT}^2(\zeta)$ and $\{W_{j,DWT}; j = 1, \dots, J_{max}\}$ are vectors of DWT coefficients used to calculate $\hat{v}_{W,j,DWT}^2$.

So in fact, the quality of our estimate of ζ depends on the the quality of our estimates of $\tilde{\sigma}_{W,j,DWT}^2(\zeta)$ using sample variance of DWT coefficients, or equivalently, on the quality of our estimates of $\tilde{v}_{W,j,DWT}^2(\zeta)$ using the rescaled sample variance of DWT coefficients, whereas each level of decomposition has a different weight (N_j/N^*) in the objective function. The weights reflect the number of DWT coefficients at individual levels of decomposition and, asymptotically, the width of the intervals of frequencies (scales) which they represent (i.e. the intervals $(2^{-(j+1)}, 2^{-j})$).

The problem, and one of the motivations for the partial decomposition, is the deteriorating quality of information provided by coefficients at subsequent levels of decomposition. With declining number of coefficients, the averages of their squares are becoming poor estimates of their variances. Consequently, at these levels, the estimator is trying to match inaccurate approximations of the spectral density, and the quality of estimates is impaired. Then the full decomposition, that uses even the highest levels with just a few coefficients, may not be optimal. The importance of this effect should increase with the total energy concentrated at the lowest frequencies used for the estimation and with the level of inaccuracy of the variance estimates. To get a preliminary notion of the magnitude of the problem in the case of FIEGARCH model, see Table 1, Figure 3 and Figure 4 in Appendix C, where integrals of the spectral density $\int_{1/2^{j+1}}^{1/2^j} f(\lambda, \zeta)$ (for several sets of coefficients) over intervals respective to individual levels are presented, together with the implied theoretical variances of the DWT coefficients $\sigma_{W,j,DWT}^2 = \int_{1/2^{j+1}}^{1/2^j} 2 * 2^j f(\lambda, \zeta)$. By their nature, the variances of the DWT coefficients reflect not only the shape of the spectral density (the integral of the spectral density multiplied by two), but also the decline in their number at subsequent levels (the 2^j term). This results in the interesting patterns observable in Figure 3, which suggest to think about both the direct effect of the decreasing number of coefficients on the variance estimates and about the indirect effect that changes their theoretical magnitudes. This indirect effect can be especially important in case of long-memory processes, where a significant portion of energy is located at low frequencies, the respective wavelet coefficients variances to be estimated become very high, while the accuracy of their estimates is poor. In general, dealing with this problem can be very important in case of small samples, where the share of the coefficients at “biased levels” is significant, but the effect should die out with increasing sample size.

One of the possible means of dealing with the latter problem is to use a partial decomposition. The idea is to set a minimal required number of coefficients at the highest level of decomposition considered in the estimation and discard all levels with lower number of coefficients. Under such a setting, the number of levels is increasing with the sample size, as in the case of full decomposition, but levels with small number of coefficients are cut off. According to Percival & Walden (2000), the convergence of the wavelet variance estimator is relatively fast, so that 128 (2^7) coefficients should already ensure a reasonable accuracy². Though, for small samples

²Accuracy of the wavelet variance estimate, not the parameters in approximate MLE.

(such as 2^9) this means a significant cut leading to estimation based on high frequencies only, which may cause even larger problems than the inaccuracy of wavelet variances estimates itself. The point is that every truncation implies a loss of information about the shape of the spectral density, whose quality depends on the accuracy of the estimates of wavelet variances. Especially for small samples, this means a tradeoff between inaccuracy due to poor variance estimation and inaccuracy due to insufficient level of decomposition. As far as our results for FIEGARCH model, based on partial decomposition suggest, somewhat inaccurate information may be still better than no information at all, and consequently, the use of truncation of 6 lags ensuring 128 coefficients at the highest level of decomposition may not be optimal. The optimal level, will be discussed together with the experiment results.

Next possible solution to the problem can be based on a direct improvement of the variances estimates at the high levels of decomposition (low frequencies). Based on the theoretical results on wavelet variance estimation provided in Percival (1995) and summarized in Percival & Walden (2000), where the wavelet variance is defined as

$$v_j^2 = \frac{\sigma_{W,j,DWT}^2}{2^j}; \quad j = 1, 2, \dots, J, \quad (42)$$

where $\sigma_{W,j,DWT}^2$ is the sample variance of the DWT wavelet coefficients, this should be possible by applying maximal overlap discrete wavelet transform (MODWT) instead of DWT. The main difference between the two transforms is that there is no sub-sampling in the case of MODWT. The number of coefficients at each level of decomposition is equal to the sample size, which can improve our estimates of the coefficients' variance. Generally, it is a highly redundant non-orthogonal transform, but in our case this is not an issue. Since the MODWT can be used for wavelet variance estimation, it can be used also for the estimation of the variances of DWT coefficients, and thus, it can be used as a substitute for the DWT in the WWE. Let's recall the WWE definition:

$$\hat{\zeta}^{WW} = \underset{\zeta \in \Theta}{\operatorname{argmin}} \sum_{j=1}^J \left[\frac{N_j}{N^*} \left(\ln \tilde{\sigma}_{W,j,DWT}^2 + \sum_{k=1}^{N_j} \frac{\hat{\sigma}_{W,j,DWT}^2}{\tilde{\sigma}_{W,j,DWT}^2} \right) \right]. \quad (43)$$

Using the definitions of variances of DWT and MODWT coefficients at level j and their relation to the original data spectral density $f(\lambda, \zeta)$ described in Percival & Walden (2000)

$$\hat{\sigma}_{W,j,DWT}^2 = \frac{\sum_{k=1}^{N_j} W_{(DWT)j,k}^2}{N_j} = 2^{j+1} \int_{1/2^{j+1}}^{1/2^j} f(\lambda, \zeta); \quad N_j = \frac{T}{2^j} \quad (44)$$

$$\hat{\sigma}_{W,j,MODWT}^2 = \frac{\sum_{k=1}^T W_{(MODWT)j,k}^2}{T} = 2 \int_{1/2^{j+1}}^{1/2^j} f(\lambda, \zeta), \quad (45)$$

it follows that

$$\hat{\sigma}_{W,j,DWT}^2 = 2^j \hat{\sigma}_{W,j,MODWT}^2. \quad (46)$$

Then the MODWT-based WWE can be defined as

$$\hat{\zeta}_{MODWT}^{WW} = \underset{\zeta \in \Theta}{\operatorname{argmin}} \sum_{j=1}^J \left[\frac{N_j}{N^*} \left(\ln \tilde{\sigma}_{W,j}^2(\zeta) + \frac{2^j \hat{\sigma}_{W,j,MODWT}^2}{\tilde{\sigma}_{W,j}^2(\zeta)} \right) \right]; \quad N_j = \frac{T}{2^j}. \quad (47)$$

According to Percival (1995), in theory, the estimates of wavelet variance using MODWT can never be less efficient than those provided by the DWT, and thus the approach described above

should improve the estimates. Results for this alternative estimator are presented later in the text.

Next interesting question related to the optimal level of decomposition concerns the possibility to make the estimation faster by using a part of the spectrum only. The idea is based on the shape of the spectral density determining the energy at every single interval of frequencies. As can be seen in table Table 1 and Figure 3 in Appendix C, for FIEGARCH model, under a wide range of parameter sets most of the energy is concentrated at the upper intervals. Therefore, whenever it is reasonable to assume that the data-generating process is not an extreme case with parameters implying extremely strong long memory, estimation using a part of the spectrum only may be reasonable. In general, this method should be both better applicable and more useful in case of very long time-series compared to the short ones, especially when fast real-time estimation is required. In case of small samples the partial decomposition can be used as a simple solution to the inaccurate variance estimates at the highest levels of decomposition, but in most cases it is not reasonable to apply it just to speed up the estimation.

At this point the questions raised above represent just preliminary notions based mostly on common sense and the results of Moulines et al. (2008) in the semi-parametric setup. To treat them properly, an asymptotic theory, in our case for the FIEGARCH-WWE, needs to be derived. This should enable to study all the various patterns in detail, decompose the overall convergence of the estimates into convergence with increasing sample size and convergence with increasing level of decomposition and to optimize the estimation setup respectively. Yet, this would be beyond the scope of our current research. Therefore, the analysis we present reduces to an extension of the set of Monte Carlo experiments to cover both the full and the partial decomposition, to demonstrate the relevancy of the problems mentioned above and to provide a motivation for further research in this area.

3.4.4 FIEGARCH WWE

After defining the general form of the estimator and discussing its properties, let's focus on the FIEGARCH WWE application. First, using WWE, the same transformation of the data as in the case of the FWE is necessary. Second, due to the flexibility of the DWT, important choices have to be made before the WWE can be applied. Percival & Walden (2000) in chapter 4 discuss some general practical considerations, including the wavelet choice, handling of boundary coefficients, choice of the decomposition level and application of the DWT on series whose length is $N \neq 2^y$. Focusing on the choices we have applied: The filters chosen for the Monte Carlo experiment are the same as those chosen in Percival & Walden (2000), i.e. Haar wavelet, D4 (Daubechies) wavelet and LA8 (Least asymmetric) wavelet, but the need of a detailed study focusing on the optimal wavelet choice for FIEGARCH WWE is apparent. The only property of the filters that was tested before the estimation was their ability to decorrelate the FIEGARCH process, that is important for the WWE derivation and its performance (see Percival & Walden (2000), Jensen (1999), Jensen (2000) or Johnstone & Silverman (1997)). In Appendix J (online) the quality of the DWT-based decorrelation is assessed based on the dependencies among the resulting wavelet coefficients. We provide estimates of autocorrelation functions (ACFs) of wavelet coefficients respective to FIEGARCH processes for ($T = 2^{11}$; $d = 0.25$, $d = 0.45$, $d = -0.25$) and filters Haar, D4 and LA8. Both sample mean and 95% confidence intervals based on 500 FIEGARCH simulations are provided for each lag available. Based on the results, the approximation of the spectral density can be applied. Next, to avoid the problem with boundary coefficients, they are excluded from the analysis; sample sizes considered are: 2^k ; $k = 9, 10, 11, 12, 13, 14$ and concerning the level of decomposition, both full and partial decomposition are used, the respective results are compared. Making all these choices, the WWE is fully specified and the

objective function is ready for parameters estimation.

3.4.5 Preliminary Results: FWE vs. WWE

Since the relative accuracy of the Fourier- and wavelet-based spectral density estimates determine the relative performance of the parameters estimators, it is interesting to see how the sample Fourier- and wavelet-based approximations of the spectral density match its true shape. For this purpose, a set of figures in Appendix B is provided, showing the rationale for WWE application. Figure 1 shows the true shape of a FIEGARCH spectral density under three different parameter sets, demonstrating the smoothness of this function and the importance of the long memory. Figure 2, then provides the wavelet-based approximations based on the simple assumption that the spectral density is constant over the whole intervals, equal to the estimated averages. Using this specification is relevant given the definition of the WWE. Wavelet-based approximations are compared with the respective true spectral densities, true averages of these spectral densities over intervals of frequencies, as well as with two Fourier-based approximations, one providing point estimates and the second estimating the averages over whole intervals. The figures show a good fit of both Fourier-based and wavelet-based approximations at most of the intervals, some problems can be seen at the lowest frequencies, which supports the idea of partial decomposition. In general, the wavelet-based approximation works well especially for processes with well behaved spectral densities without significant patterns well localized in the frequency domain, when the average energy over the whole intervals of frequencies represents a sufficient information about the shape of the true spectral density. For these processes, the wavelet transform can be effectively used for visual data analysis and both parametric and semi-parametric estimation of parameters in the spectral density function. More figures for the spectral density approximation are available in Appendix I (online).

4 Monte Carlo Experiments

In order to study how the WWE performs compared to the two benchmark estimators (MLE and FWE), an extensive Monte Carlo experiment has been carried out. Each round consisted of 1000 simulations of a FIEGARCH process at a fixed set of parameters, and estimation of these parameters by all methods of interest. The experiment setup mirrors that of Perez & Zaffaroni (2008), which ensures consistency with that work and enables interpretation of the new results as an extension to those already published. Since in this benchmark study no wavelet-based methods are used, choices concerning the WWE application and extension of simulations to longer data sets have been made with respect to other relevant literature (Jensen (1999), Percival & Walden (2000)), as discussed earlier. Technical details of the experiment and tables with results are provided in Appendix K (online).

4.1 Results I: Maximal Decomposition

At first, the experiment has been performed using MLE, FWE and DWT-based WWE with maximal level of decomposition. The maximal level of decomposition means that for sample of length 2^j , using Haar, D4 and LA8 wavelets, we have levels j , $j-2$ and $j-3$ respectively. This is due to the truncation of boundary coefficients explained in the previous section.

In general, the WWE works fairly well in all setups. Especially biases are low, and in most cases decline with increasing sample size. The exception is parameter α for all filters and parameter θ for Haar filter, where the convergence is problematic. Yet, even in these situations the bias remains low for all filters and samples up to 2^{11} , and RMSE (although relatively high) is

declining with sample size as for all the other parameter estimates across all setups. Focusing on the differences between individual filters, the strength of long memory, sample size and parameter concerned seem to be important. In all setups, Haar performs the best in estimating the long-memory parameter d . Other parameters (α , β , γ and in case of small samples also the θ), under $d = 0.25$, are better estimated using filters with larger support. In case of $d = 0.45$ and $d = -0.25$ the relative performance slightly improves for filters with smaller support. The overall performance of the wavelet-based estimators (WWE using various filters) in the experiment suggests using D4 for 2^{10} and 2^{11} and switching to LA8 for 2^9 and $\{2^j; j > 11\}$ in case of long memory in the data (a simple ACF analysis before estimation should reveal this pattern). For negative dependence the optimal choice seems to be Haar for 2^9 and D4 otherwise (with possible shift to LA8 for samples longer than 2^{14}).

Concerning the relative performance of the WWE, FWE and MLE, the WWE works in general comparably to the FWE. In many cases it outperforms the FWE in terms of bias, while in terms of RMSE the FWE is better. Yet, the absolute differences are usually small. As expected, estimates using MLE are in most cases the best. This remains true even in cases with strong long memory, since the long memory is in the variances, not in the (log-return) process itself and the problem with mean estimation under long memory does not apply. The Whittle estimators outperform the exact MLE in some cases, but usually it is in situations with negative memory in the data, which is, based on the current literature on financial returns analysis, not of a great interest for most practical applications.

4.2 Results II: Partial Decomposition

The additional Monte Carlo experiments have been designed to mirror the setup used in the case of full decomposition, with the only difference in the number of levels used for the estimation. For all sample lengths (2^M , $M = 9, 10, \dots, 14$) experiments for levels J , $J = 4, 5, \dots, M$ have been carried out. Results are available for both processes with long memory ($d = 0.25$ and $d = 0.45$), which are of the most interest for practical applications, the case of $d = -0.25$ is omitted to keep the extent of simulations reasonable. For results including mean estimates, respective levels of bias and RMSE see tables in Appendix L(online).

As the results suggest, for small samples, estimation under the restriction to first four levels of decomposition leads to better estimates of d and worse estimates of α in terms of both bias and RMSE, while for longer samples the opposite holds. Other coefficients are estimated sometimes with lower bias, sometimes with lower RMSE than in the case of full decomposition depending on the sample size, strength of the long memory and also on the filter applied. With increasing sample size the performance of the estimator under partial decomposition deteriorates relatively to that using full decomposition. Comparing the performance of individual filters, in most cases LA8 provides the best sets of estimates for both $d = 0.25$ and $d = 0.45$, except for the case of small samples with $d = 0.25$, where Haar seems to be better. In general, it can be said that this partial decomposition setup offers just a different set of bias and RMSE for all parameters than the full decomposition. The choice would depend on the weights assigned to the bias and RMSE and the importance we attach to individual parameters, which could be based on a bias and RMSE of one day forecasts added to the Monte Carlo experiment. While there is a relatively high probability that based on the more detailed analysis the level 4 setup may be preferred in case of short samples, for long samples the full decomposition is likely to be more appropriate.

Moving on to the truncation at level 5, significant overall improvement in the short-sample estimates is apparent for both $d = 0.25$ and $d = 0.45$. Not only are they better compared to the level 4 setup, but also compared to the full decomposition. Relative performance with respect to FWE and MLE also changes, WWE works in most cases better than the FWE for

all filter specifications. Focusing on the relative performance of the filters considered, the results suggest to use D4 for $2^{10} - 2^{13}$ and switching to LA8 for 2^9 and 2^j ; $j > 13$ in case of $d = 0.25$; under $d = 0.45$ LA8 performs the best for all sample sizes as in the case of preceding partial decomposition setup.

Next, under truncation at level 6 the estimator seems to work comparably to the case of truncation at level 5. In most cases it offers an alternative of somewhat lower RMSE at the cost of slightly higher bias, for some parameters even the bias improves. Though, due to the significantly worse estimates of long memory, even to some extent counterbalanced by better estimates of other parameters, the truncation at level 5 may be preferred. The relative performance could be assessed based on the bias and RMSE of one day forecasts added to the Monte Carlo experiment as already proposed in the case of level 4 truncation. Regarding the relative performance of the filters considered, in case of $d = 0.25$ D4 performs the best for almost all sample sizes, while it is outperformed by LA8 when the parameter d becomes larger. Compared to the full decomposition, in case of small samples the estimator works better in most cases in terms of both bias and RMSE. In case of longer samples, the estimates of the long memory parameter deteriorate relatively to their full decomposition counterparts, while short-term dynamics parameters are still estimated in most cases with lower bias and in case of $d = 0.45$ also with lower RMSE under the truncation.

We conclude that the results well demonstrate the effects mentioned when discussing the partial decomposition in 3.4.3. We can see how the partial decomposition helps in the case of short samples and how the benefits from truncation (no use of inaccurate information) decrease relative to the costs (more weight on the high-frequency part of the spectra and no information at all about the spectral density shape at lower frequencies) as the sample size increases, as the long-memory strengthens and as the truncation becomes excessive. Moreover, the effect becomes negligible with longer samples, as the share of problematic coefficients goes to zero. Yet, the convergence with sample size and with the level of decomposition is not easy to interpret. To see the interesting convergence patterns determined by the synergy of various effects of the truncation on the estimates see 3D plots in Appendix D providing a graphical decomposition of the convergence into the convergence with sample size and convergence with increasing level of decomposition; graphs for the estimates of d and α under $d = 0.25$; D4, LA8 and $d = 0.45$; LA8 are available. To make the figures comprehensive and well interpretable, additional Monte Carlo experiments have been performed enabling to present the whole spectrum of possible truncations from that leading to estimation at level 4 to full decomposition. As can be seen, the optimal setup choice for small samples is a non-trivial problem that cannot be reduced to a simple method of cutting a fixed number of highest levels of decomposition to ensure some minimal number of coefficients at each level. Although in case of long samples a nice convergence with both sample size and level of decomposition can be seen for all specifications, the results for small samples are mixed. In this latter case the convergence with sample size still works relatively well, but the increase in level of decomposition does not always improve the estimates. To understand the specific patterns, next to the derivation of the asymptotic theory, it would be interesting to compare the results with their MODWT-based counterparts, which would enable to separate the effect of deteriorating DWT variance estimates and would potentially lead to better interpretable convergence patterns.

5 Monte Carlo Extension: Jumps and Forecasting

As has been concluded in the previous section, on simulated pure FIEGARCH processes the best estimator in terms of both bias and RMSE (in case of individual parameters estimation) seems to be the MLE, followed by FWE and somewhat less “accurate” WWE. But, as discussed in

the sequel, deprecating WWE based on these results only might be premature. Let's look what happens when we assume a more realistic scenario, where the simulated process is augmented by specific time-localized irregularities - jumps - in the log-return process. Next, since the evaluation based on individual parameters estimation only may not be the best practice when forecasting is the main concern, let's analyze also the relative forecasting performance. As a motivation for this step additional plots have been prepared, which can be found in Appendix E (online) and some of them also in Appendix N. They show the bias of the mean estimated spectral densities using FWE and DWT-based WWE under various setups. Except for small samples, where the performance of the FWE is significantly worse than that of WWEs in terms of bias, both the estimators perform very well and in most cases differences are almost negligible. This suggests that the forecasting RMSE should play the major role. Then, based on the results for individual coefficients, FWE can be expected to dominate the DWT-based WWE, at least in case of larger samples and, of course, data generated by a pure FIEGARCH process. But this is just an ex ante guess, the need for a Monte Carlo experiment extension is apparent. Practical issues of this kind of evaluation are discussed later in this section.

5.1 FIEGARCH-Jump Model

Jumps are one of the several well known stylized features of log-returns and/or realized volatility time series and there is a lot of studies on incorporating this pattern in volatility models. For a summary see e.g. Mancini & Calvori (2012). So even if the FIEGARCH process could well approximate the true underlying volatility, it is important to study the jump process and use the additional information for forecasts improvement.

To test the performance of the individual estimators in the case of FIEGARCH-Jump processes, an additional Monte Carlo experiment has been conducted. The simulations are augmented by additional jumps, which do not enter the conditional volatility process, but the log-returns process only. This represents the situation, when the jumps are not resulting from the long memory in the volatility process, which can produce patterns similar to jumps in some cases, as well as they do not determine the volatility process in any way. The log-return process is then specified as:

$$\epsilon_t = z_t h_t^{1/2} + J_t(\lambda), \quad (48)$$

where the process h_t remains the same as in the original FIEGARCH model (Equation 1) and J_t ; $t = 1, 2, \dots, T$ is a Jump process. The jumps are modeled as a sum of intraday jumps, whereas the number of intraday jumps in one day follows a Poisson process with parameter $\lambda = 0.028$ and their size is drawn from a normal distribution $N(0, 0.2)$. The Jump process is based on Mancini & Calvori (2012), with parameters slightly adjusted (originally $\lambda = 0.014$ and sizes follow $N(0, 0.25)$) based on analysis of resulting simulations and comparison with real data. Moreover, unlike in the previous Monte Carlo experiment, a non-zero constant is assumed. Since we would like to keep consistency in the research (keep the parameters the same throughout this paper) and at the same time to simulate time series as close to the real ones as possible, we have compared our simulated time series with real data from "yahoo.finance.com" (for facebook and zynga) and found that the optimal mean for series like these could be around -7 . It is a rule of thumb decision, but we believe that it is a reasonable value at least as a starting point for the future research.

5.2 Forecasting

Next extension, as mentioned above, is the evaluation of the in-sample and out-of-sample forecasting performance. For each simulation the fitted values and a one day ahead forecast per each estimator are calculated. The out-of-sample forecasts are directly stored for further analysis, the in-sample forecasts are transformed to mean error, mean absolute deviation and mean squared error statistics. These statistics are stored and used for overall statistics calculation. When we get the data from all 1000 simulations, we compute the mean error, mean absolute deviation and root mean squared error for both the in-sample and out-of-sample forecasts.

Although the idea of forecasting evaluation seems to be simple, there are some issues we had to deal with. The most important one is the dependency of the forecasting results on the fitting algorithm. This algorithm is technically the same as the one used in MLE and it is in fact possible to manage the maximal error of the forecasts in case of divergence of the fitted time series as well as to ensure robustness to extreme log-returns observations in the data input. We have chosen a basic algorithm that only ensures that the operation does not break and that in case of non-positive and/or infinite fitted conditional variance the algorithm returns sample variance instead and continues computing. But then we have situations, especially in case of jumps, where a few forecasts can have a huge finite error, but other are quite accurate. Then, average error measures are not the best practice, since then the estimator with slightly lower maximal error would be considered better even though in most cases it could be much worse than the alternatives. In case of out-of-sample forecasts (one forecast per simulation) we solve this problem by using error and absolute deviation quantiles. In case of in-sample forecast we set a rule of thumb level of allowed excessive errors (1%) and we analyze further only the 99% of better forecasts. Then, if any NaN is present, we count this observation but exclude it from the further analysis (for all estimators). Next, extremely high in-sample mean absolute errors are considered to be equivalent to infinite ones and treated accordingly, in case of out-of-sample forecasts they are not excluded and their magnitude can be studied - we study the real out-of sample forecasting performance in case of non-extreme in-sample fit. Therefore it is essential to look at the quantile results rather than on the summary measures, where in case of one excessive error out of the 1000 realizations the mean, mean absolute deviation and RMSE are huge even though 999 estimates may be pretty accurate.

5.3 Practical Aspects

Although we expected that the WWE could be easily adjusted to be robust to jumps (compared to FWE and MLE) and thus become a good alternative in case of FIEGARCH estimation on the real data, there are two technical details which make this hardly possible and the WWE theoretically rather than empirically evincible. First, the transformation needed for FWE and WWE derivation hides the jumps in the process (they add volatility, but are not detectable in the transformed data), as can be seen in Appendix F; second, for forecasting we need jump-free data as the input - else we get inaccurate estimates even in case of perfect coefficient estimates. Thus, the jump detection and data adjustment has to be done before the actual parameters estimation takes place. To deal with the jumps we apply one of the well performing wavelet-based jump estimators that is based on a universal threshold of Donoho & Johnstone (1994) and that is described in detail and successfully applied in Barunik & Vacha (2012). When detected, the jumps are replaced by average of the two adjacent values. This, of course, is not the best practice in case of large double-jumps, where this transformation leads to two smaller jumps instead of getting rid of them. Yet, in case of long memory that can produce jump-like patterns, which are usually clustered in high volatility intervals, getting rid of the multiple jumps may not be the best thing to do. So we use this simple transform for our data with moderate jumps, but

in case of data with extreme jumps, such as those in Appendix F, we propose to use a different method that would get rid of all the jumps, otherwise the estimation results would be poor. Anyway, the important thing here is that the jump detection and model estimation are two separable tasks. This holds even in cases of large jumps which are detectable in the transformed data, since in real applications we do not know what kind of jump process are we dealing with and also even in this case it is easier to found the jumps in the data before the transformation. Then we can only study how are the individual estimators able to deal with the residual jumps, which are not detected and subtracted from the time series. And of course, the better the jump estimation method, the lesser the residuals impact.

5.4 Results III: FIEGARCH-Jump

The main results of the Monte Carlo experiment are summarized in tables in Appendix G. In the first two tables MLE, FWE and MODWT-based WWE are compared in terms of individual parameters estimation performance, results for DWT-based WWE are not included due to the limited space. Concerning the comparison of these two estimators, the overall performance of the MODWT-WWE is better than that of the DWT-WWE both in terms of bias and RMSE and considering also the loss of sample size limitation, the MODWT-WWE is strictly preferred. This is only supported by the forecasting results presented in the next tables. Next, focusing on the MLE, FWE and MODWT-WWE relative performance in terms of RMSE for jumps and $d = 0.25$, the MLE, despite being affected by the residual jump effects, it remains the best followed by the two Whittles, which perform comparably, FWE in most cases works slightly better. Yet, the bias of the MLE is significant and we would prefer the use of FWE considering both the bias and the RMSE and in case of longer time series, WWE seems to be the best option due to the faster bias decay. Next, for $d = 0.45$, the MLE performance is very poor and the use of FWE is preferable. As expected, the bias and RMSE in case of individual parameters estimates as well as the mean absolute deviation and RMSE of the out-of-sample forecasts decline and the overall in-sample fit improves with sample size increase and long memory weakening. Next, the constant term estimation performance is worth mentioning, since it is very poor in the case of MLE and strong long memory, and therefore an ex ante estimation as in the case of FWE and WWE is appropriate.

On the other hand, when we look at the forecasting performance, the results are much more clear. The best in all scenarios and by all indicators is the MLE, followed by the FWE and a little less accurate WWE. The impact of jumps depends, of course, on the jump estimator performance and in our case, for forecasting, it is very limited, although the same cannot be said about the impact on individual parameters estimates. Unfortunately, as discussed earlier, WWE does not provide any significant estimation improvement or time savings. Moreover, the use of WWE causes about twice as many cases with extremely poor in-sample fit than the FWE. By its nature, MLE does not cause these poor fit situations at all. But of course, in practice, adjustment of the optimization algorithms as well as the forecasting algorithm could prevent these cases for all estimators. We did not apply any special adjustments just to keep the estimators comparison as “fair” as possible and we propose the question of the algorithms optimization as a topic for future research.

6 Comparison with Literature

As already emphasized, the basic Monte Carlo setup has been chosen to mirror that of Perez & Zaffaroni (2008) to keep consistency in the research and enable direct comparison of the results. Since the benchmark paper focuses on the relative performance of FWE and MLE under the

same conditions as applied in our paper, it is interesting to check whether the results for these two estimators are in both works the same. In the case of MLE, the answer is yes, up to small differences caused by the uniqueness of every simulated time series. For FWE, the differences are somewhat larger (although not extreme; the maximal deviation for both bias and RMSE is less than 0.1, mostly amounting to about 0.06 or less). Based on an analysis of the code underlying the benchmark paper, this may be caused by a different optimization setup, that utilizes an explicitly formulated analytical gradient of the objective function. This helps to estimate the individual coefficients more accurately. Then, it seems natural to run the Monte Carlo experiments for our research using this, virtually more efficient, algorithm (or search for even better one). Yet, this is a non-trivial task. This follows from the form of the WWE objective function and the respective derivations. To enable the re-estimation, more efficiently written code or an adjusted means of gradient calculation would be needed, which is beyond the scope of our current work. In general, the comparison above shows the importance of taking the estimation as a complex problem including many, both theoretical and practical, issues. Most importantly, when comparing several estimators, it is usually not feasible to separate the performance of the estimator from the performance of the optimization algorithm applied. Even using the same algorithm, as applied in this work, cannot generally solve this problem, since each of the estimators may be affected differently. Then, in case of empirical analysis, it is reasonable to speak about comparison of methods of estimation instead of comparison of the individual estimators. The focus is then on the whole sets of estimators and respective means of optimization. This highlights the importance to analyze all details of the methods and optimize the estimation setup as a whole, before any definite conclusions can be made.

Next, comparing the results in this paper with some other works on wavelet-based maximum likelihood estimation, no strange patterns that would contradict the earlier conclusions are found. A relatively good performance of the WWE comparable to that of FWE is observed, which is in compliance with studies using simulated smooth processes. The absolute performance of the WWE is somewhat worse than in the benchmark papers, which is expectable given the complexity of the FIEGARCH model implying more difficult parameters identification compared to the other models estimated in the related works, as well as given the asymmetry of the FIEGARCH process that makes the QMLE less accurate (other works focus on symmetric processes). As mentioned above, optimization of the estimation setup should improve the overall performance and lead to absolute results closer to those in the benchmark studies. Next, focusing on the filter choice, the relative performance of the Haar, D4 and LA8 filters seem to be in compliance with that in Percival & Walden (2000), as well as it supports the conclusion in Jensen (1999) that Haar can be dominated by longer filters. To sum it up, the current work seems to extend the current literature without any contradiction with earlier works. Given the lack of related studies, this seems to be good news. Though, to make any strong conclusions about the WWE performance in various applications, a lot of work has to be done in the future.

7 Conclusion

In this paper, we introduce a new, wavelet-based estimator (wavelet Whittle estimator, WWE) of a FIEGARCH model, ARCH-family model allowing for long-memory and asymmetry in volatility, and study its properties. Based on several Monte Carlo experiments its accuracy and empirical convergence are examined, as well as its relative performance with respect to two more traditional estimators: Fourier-based Whittle estimator (FWE) and maximum likelihood estimator (MLE). It is shown that even in the case of simulated pure FIEGARCH processes, which do not allow to fully utilize the advantages of the WWE, the estimator can work reasonably well. In terms

of bias, it often outperforms the FWE, while in terms of RMSE the FWE is better. Yet, the absolute differences are usually small. As expected, MLE is in most cases the best, even under strong long memory. The Whittle estimators outperform the MLE in some cases, but usually it is in situations with negative memory, which is not of a great interest for most practical applications. The forecasting performance analysis has a similar conclusion, just the differences across estimators are even smaller. Yet, since the Whittle estimators are significantly faster and the differences in the performance are small, they are a good alternative to the MLE for long samples. Concerning the optimal WWE settings studied, the strength of long memory, sample size and parameter concerned seem to be important for the optimal filter (wavelet) choice, but further research in this area is needed.

Next, some practical aspects of the WWE application are discussed. The main focus is on the problem of declining number of wavelet coefficients at subsequent levels of decomposition, which impairs the estimates accuracy. Two solutions to this problem are suggested. One is based on a partial decomposition (local WWE) that ensures some minimal number of coefficients at the highest level of decomposition, the other applies an alternative specification of the WWE (using maximal overlap discrete wavelet transform, MODWT). First, we show that the partial decomposition can improve the estimates in case of short samples, even make the WWE superior to the FWE (and to the MLE for negative memory), while in case of long samples, full decomposition is more appropriate. Yet, the second solution (MODWT-WWE) is argued to be better. Compared to the former method, it ensures the number of coefficients at every level equal to the sample size and does not lead to any decline in the share of spectrum used in the estimation (information loss). The only cost to bear is a somewhat longer estimation time. As our results suggest, using the MODWT instead of the DWT improves the WWE performance in all scenarios.

Next, we study the estimators in case of jumps. The accuracy of individual parameters estimates using MLE is significantly impaired, even if we apply a simple data correction; the FWE and the WWE are superior. Yet, based on the forecasting performance, MLE should be preferred in all scenarios at least in case of small samples, where it can be computed in reasonable time; FWE and WWE can be recommended only as a faster, but slightly less accurate alternative. From these two FWE performs slightly better. Yet, we believe that after optimization of the estimation and forecasting algorithms, the differences between the FWE and WWE disappear, or even the WWE becomes superior.

Last but not least, we discuss the effects of optimization algorithm choice on the experiment results. It is argued that in cases, when the identification of individual parameters in the objective function is problematic, as is the case of the Whittle estimators applied in this work, the performance of the estimator and of the optimization algorithm cannot be well separated. Based on a comparison of our results with those of Perez & Zaffaroni (2008), it is argued that application of more sophisticated optimization algorithms to both FWE and WWE should improve their absolute performance, and potentially also change the conclusions about their relative performance. Therefore, the question of algorithm choice is an important topic to address in the future.

It can be concluded that after optimization of the estimation setup, the WWE may become a very attractive alternative to the traditional estimation methods. Although it is not as useful in case of jumps in the data as we expected, the statement that, compared to FWE, it is more robust to time-localized irregularities is still valid. The only additional requirement is that the irregularities remain detectable even after the data transformation that is necessary for the FWE and WWE application. Although a lot of work has to be done before the WWE applicability and performance will be properly assessed, importance of the research results for volatility modeling should be a sufficient motivation.

References

- Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1), 3–30.
- Barunik, J. & Vacha, L. (2012). Realized wavelet-based estimation of integrated variance and jumps in the presence of noise. Papers 1202.1854, arXiv.org.
- Beran, J. (1994). *Statistics for long-memory processes*. Monographs on statistics and applied probability, 61. Chapman & Hall.
- Bollerslev, T. & Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, 73(1), 151 – 184.
- Bollerslev, T. & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, 11, 143–172.
- Cheung, Y.-W. & Diebold, F. X. (1994). On maximum likelihood estimation of the differencing parameter of fractionally-integrated noise with unknown mean. *Journal of Econometrics*, 62(2), 301 – 316.
- Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *The Annals of Statistics*, 17(4), 1749–1766.
- Dahlhaus, R. (2006). Correction: Efficient parameter estimation for self-similar processes. *The Annals of Statistics*, 34(2), pp. 1045–1047.
- Donoho, D. L. & Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3), 425–455.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4), 987–1007.
- Engle, R. F. & Gonzalez-Rivera, G. (1991). Semiparametric arch models. *Journal of Business & Economic Statistics*, 9(4), 345–359.
- Fay, G., Moulines, E., Roueff, F. R., & Taqqu, M. S. (2009). Estimators of long-memory: Fourier versus wavelets. *Journal of Econometrics*, 151(2), 159 – 177.
- Fox, R. & Taqqu, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary gaussian time series. *The Annals of Statistics*, 14(2), 517–532.
- Frederiksen, P. H. & Nielsen, M. O. (2005). Finite sample comparison of parametric, semiparametric, and wavelet estimators of fractional integration. *Econometric Reviews*, 24(4), 405–443.
- Gonzaga, A. & Hauser, M. (2011). A wavelet whittle estimator of generalized long-memory stochastic volatility. *Statistical Methods & Applications*, 20(1), 23–48.
- Hardle, W., Kerkycharian, G., Tsybakov, A. B., & Picard, D. (1998). *Wavelets, Approximation and Statistical Applications*. John Wiley & Sons, Incorporated.
- Heni, B. & Mohamed, B. (2011). A wavelet-based approach for modelling exchange rates. *Statistical Methods & Applications*, 20(2), 201–220.
- Jensen, M. J. (1999). An approximate wavelet mle of short- and long-memory parameters. *Studies in Nonlinear Dynamics Econometrics*, 3(4), 5.

- Jensen, M. J. (2000). An alternative maximum likelihood estimator of long-memory processes using compactly supported wavelets. *Journal of Economic Dynamics and Control*, 24(3), 361–387.
- Johnstone, I. M. & Silverman, B. W. (1997). Wavelet threshold estimators for data with correlated noise. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 59(2), 319–351.
- Mancini, C. & Calvori, F. (2012). *Jumps*, chapter 17, (pp. 403–445). John Wiley & Sons, Inc.
- Moulines, E., Roueff, F., & Taqqu, M. S. (2008). A wavelet whittle estimator of the memory parameter of a nonstationary gaussian time series. *The Annals of Statistics*, 36(4), 1925–1956.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2), 347–370.
- Nielsen, M. O. & Frederiksen, P. H. (2005). Finite sample comparison of parametric, semiparametric, and wavelet estimators of fractional integration. *Econometric Reviews*, 24(4), 405–443.
- Percival, D. B. & Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press.
- Percival, D. P. (1995). On estimation of the wavelet variance. *Biometrika*, 82(3), 619–631.
- Perez, A. & Zaffaroni, P. (2008). Finite-sample properties of maximum likelihood and whittle estimators in egarch and fiegarch models. *Quantitative and Qualitative Analysis in Social Sciences*, 2, 78–97.
- Whitcher, B. (2004). Wavelet-based estimation for seasonal long-memory processes. *Technometrics*, 46(2), 225–238.
- Wornell, G. W. & Oppenheim, A. (1992). Estimation of fractal signals from noisy measurements using wavelets. *Signal Processing, IEEE Transactions on*, 40(3), 611–623.
- Zaffaroni, P. (2009). Whittle estimation of egarch and other exponential volatility models. *Journal of Econometrics*, 151(2), 190–200.

A FIEGARCH Transformation

As discussed and derived in Perez & Zaffaroni (2008) and Zaffaroni (2009), to estimate the FIEGARCH model using FWE, it is necessary to rewrite the model in a signal plus noise form, for which the spectral density can be derived. Let's begin with the original FIEGARCH(1,d,2) model:

$$\epsilon_t = z_t h_t^{1/2} \quad (49)$$

$$\ln(h_t) = \omega + \Phi(L)g(z_{t-1}) \quad (50)$$

$$g(z_t) = \theta z_t + \gamma[|z_t| - E(|z_t|)] \quad (51)$$

$$\Phi(L) = (1 - L)^{-d}[1 + \alpha_{[2]}(L)][\beta(L)]^{-1} \quad (52)$$

$$\text{where for FIEGARCH}(1,d,2): \alpha(L) = 1 + \alpha L \quad (53)$$

$$\beta(L) = 1 - \beta L \quad (54)$$

$$(1 - L)^d = 1 - d \sum_{k=1}^{\infty} \Gamma(k - d) \Gamma(1 - d)^{-1} \Gamma(k + 1)^{-1} L^k, \quad (55)$$

where $\{z_t\}$ and $\{\epsilon_t\}$ denote discrete-time real valued processes satisfying $E_{t-1}(z_t) = 0$, $E_{t-1}(z_t^2) = 1$, $E_{t-1}(\epsilon_t) = 0$, $E_{t-1}(\epsilon_t^2) = h_t$. Following Zaffaroni (2009), this can be rewritten as

$$\epsilon_t = z_t h_t^{1/2} \quad (56)$$

$$\epsilon_t^2 = z_t^2 h_t \quad (57)$$

$$\ln(\epsilon_t^2) = \ln(z_t^2) + \ln(h_t) \quad (58)$$

$$\ln(h_t) = \omega + \Phi(L)g(z_{t-1}) \quad (59)$$

$$\Phi(L)g(z_{t-1}) = \sum_{s=0}^{\infty} \Phi_s g(z_{t-s-1}) \quad (60)$$

$$g(z_t) = \theta z_t + \gamma[|z_t| - E(|z_t|)], \quad (61)$$

which leads to

$$\ln(\epsilon_t^2) = \ln(z_t^2) + \omega + \sum_{s=0}^{\infty} \Phi_s g(z_{t-s-1}) \quad (62)$$

$$g(z_t) = \theta z_t + \gamma[|z_t| - E(|z_t|)] \quad (63)$$

$$\Phi(L) = (1 - L)^{-d}[1 + \alpha L][1 - \beta L]^{-1}. \quad (64)$$

From Perez & Zaffaroni (2008), the spectral density of this process for $\{z_t\} \sim N(0, 1)$, $\omega = 0$

and Fourier (angular) frequencies $\lambda_j = 2\pi j/T$; $\lambda \in \langle -\pi, \pi \rangle$ is

$$f(\lambda, v) = \frac{A(\xi)}{2\pi} + \frac{B(v)}{2\pi} \left| \sum_{s=0}^{\infty} \Phi_s(\zeta) \exp(is\lambda) \right|^2 + \quad (65)$$

$$+ \frac{C(v)}{2\pi} \left(\exp(i\lambda) \sum_{s=0}^{\infty} \Phi_s(\zeta) \exp(is\lambda) \right) +$$

$$+ \frac{C(v)}{2\pi} \left(\exp(-i\lambda) \sum_{s=0}^{\infty} \Phi_s(\zeta) \exp(-is\lambda) \right)$$

$$A(\xi) = \text{var}(\ln(z_0^2)) \quad (66)$$

$$B(v) = \text{var}(g(z_0)) \quad (67)$$

$$C(v) = \text{cov}(\ln(z_0^2), g(z_0)) \quad (68)$$

$$v = (\xi', \zeta')', \quad (69)$$

where ξ is the vector of parameters in the distribution function of variable z . For $\{z_t\} \sim N(0, 1)$:

$$A(\xi) = \Psi\left(\frac{1}{2}\right) \quad (70)$$

$$B(v) = \theta_2 + \delta^2 * (1 - \mu_{|z|}^2) \quad (71)$$

$$C(v) = \delta * \mu_z * (\psi(1) - \psi(\frac{1}{2})) \quad (72)$$

$$\mu_{|z|} = E(|z|) = \sqrt{\frac{2}{\pi}}, \quad (73)$$

where $\psi(x)$ and $\Psi(x)$ are digamma and trigamma functions respectively. Evaluated at Fourier frequencies, this spectral density occurs in both terms of the FWE objective function. For a generalization to $\{z_t\}$ following GED or Student-t distribution, see Perez & Zaffaroni (2008).

B Spectral Density Estimation

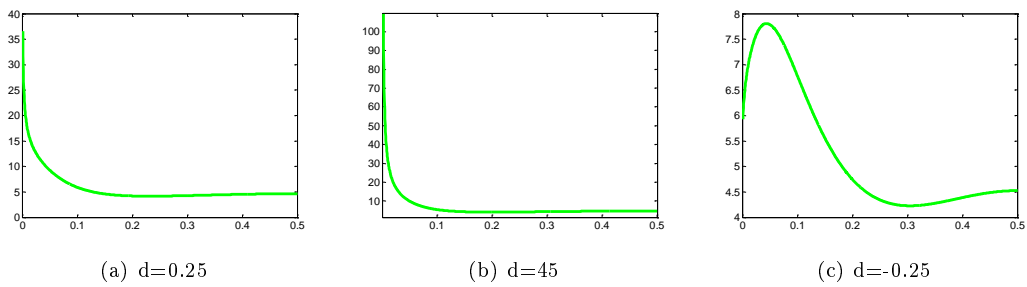
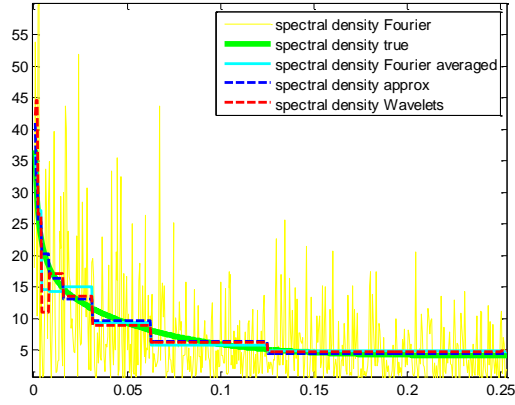
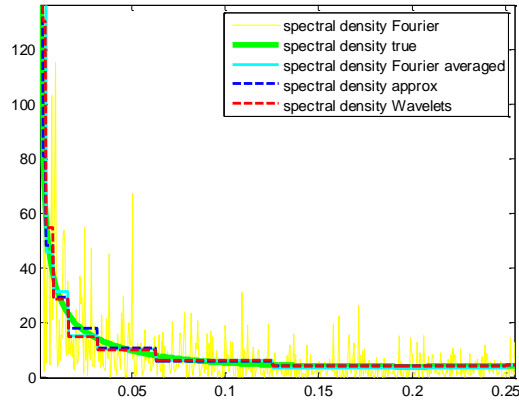


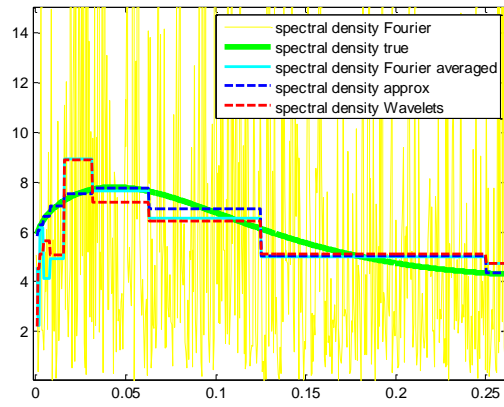
Figure 1: True spectral density



(a) $d=0.25$



(b) $d=0.45$



(c) $d=-0.25$

Figure 2: Spectral density estimation ($d=0.25/0.45/-0.25$), $T=2048$ (2^{11}), level=10,
zoom

C Partial Decomposition

(a) Coefficient Sets

Coefficients:	d	ω	α	β	θ	γ
A:	0.25	0	0.5	0.5	-0.3	0.5
B:	0.45	0	0.5	0.5	-0.3	0.5
C:	-0.25	0	0.5	0.5	-0.3	0.5
D:	0.25	0	0.9	0.9	-0.3	0.5
E:	0.45	0	0.9	0.9	-0.3	0.5
F:	-0.25	0	0.9	0.9	-0.3	0.5
G:	0.25	0	0.9	0.9	-0.9	0.9
H:	0.45	0	0.9	0.9	-0.9	0.9

(b) Integrals over frequencies respective to levels for the coefficient sets from Tab.E.1
(a)

	A	B	C	D	E	F	G	H
Level 1	1.1117	1.1220	1.0897	1.1505	1.1622	1.1207	1.1261	1.1399
Level 2	0.5473	0.5219	0.6274	0.4776	0.4691	0.5306	0.6187	0.6058
Level 3	0.3956	0.3693	0.4330	0.3246	0.3056	0.3959	1.1354	1.3453
Level 4	0.3029	0.3341	0.2425	0.5559	0.7712	0.3528	2.9558	4.8197
Level 5	0.2035	0.2828	0.1175	1.0905	2.1758	0.3003	6.0839	13.2127
Level 6	0.1279	0.2297	0.0550	1.4685	3.9342	0.1965	8.2136	23.4144
Level 7	0.0793	0.1883	0.0259	1.3523	4.7975	0.0961	7.6026	28.4723
Level 8	0.0495	0.1584	0.0123	1.0274	4.8302	0.0408	5.8268	28.7771
Level 9	0.0313	0.1368	0.0059	0.7327	4.5720	0.0169	4.1967	27.3822
Level 10	0.0201	0.1206	0.0029	0.5141	4.2610	0.0071	2.9728	25.6404
Level 11	0.0130	0.1080	0.0014	0.3597	3.9600	0.0030	2.0977	23.9192
Level 12	0.0086	0.0979	0.0007	0.2518	3.6811	0.0013	1.4793	22.2986

(c) Sample Variances of DWT Wavelet Coefficients for the coefficient sets from Table E.1 (a)

	A	B	C	D	E	F	G	H
Level 1	4.4468	4.4880	4.3588	4.6020	4.6488	4.4828	4.5044	4.5596
Level 2	4.3784	4.1752	5.0192	3.8208	3.7528	4.2448	4.9496	4.8464
Level 3	6.3296	5.9088	6.9280	5.1936	4.8896	6.3344	18.1664	21.5248
Level 4	9.6928	10.6912	7.7600	17.7888	24.6784	11.2896	94.5856	154.2304
Level 5	13.0240	18.0992	7.5200	69.7920	139.2512	19.2192	389.3696	845.6128
Level 6	16.3712	29.4016	7.0400	187.9680	503.5776	25.1520	1051.3408	2997.0432
Level 7	20.3008	48.2048	6.6304	346.1888	1228.1600	24.6016	1946.2656	7288.9088
Level 8	25.3440	81.1008	6.2976	526.0288	2473.0624	20.8896	2983.3216	14733.8752
Level 9	32.0512	140.0832	6.0416	750.2848	4681.7280	17.3056	4297.4208	28039.3728
Level 10	41.1648	246.9888	5.9392	1052.8768	8726.5280	14.5408	6088.2944	52511.5392
Level 11	53.2480	442.3680	5.7344	1473.3312	16220.1600	12.2880	8592.1792	97973.0432
Level 12	70.4512	801.9968	5.7344	2062.7456	30155.5712	10.6496	12118.4256	182670.1312

Table 1: Energy decomposition

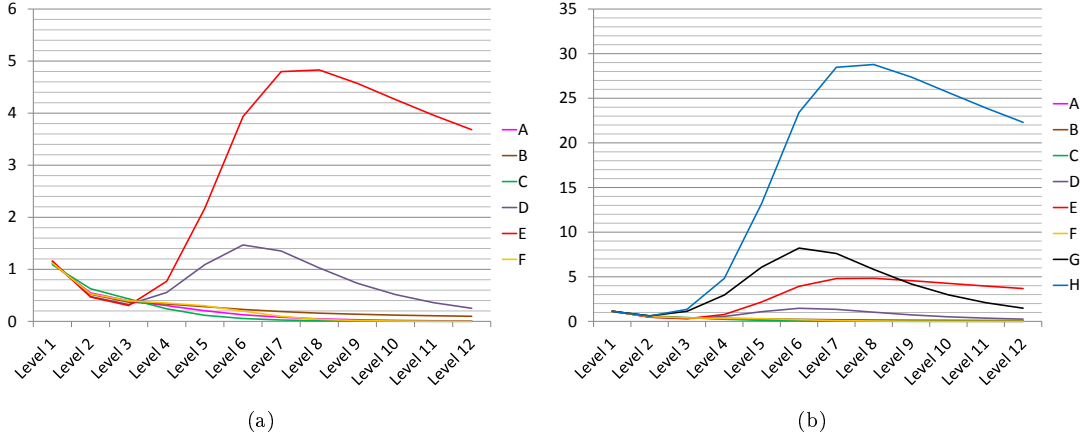


Figure 3: Energy decomposition: Integrals of FIEGARCH spectral density over frequency intervals respective to individual levels of decomposition, assuming various levels of long memory ($d=0.25$, $d=0.45$, $d=-0.25$) and the coefficient sets from Table E.1 (a)

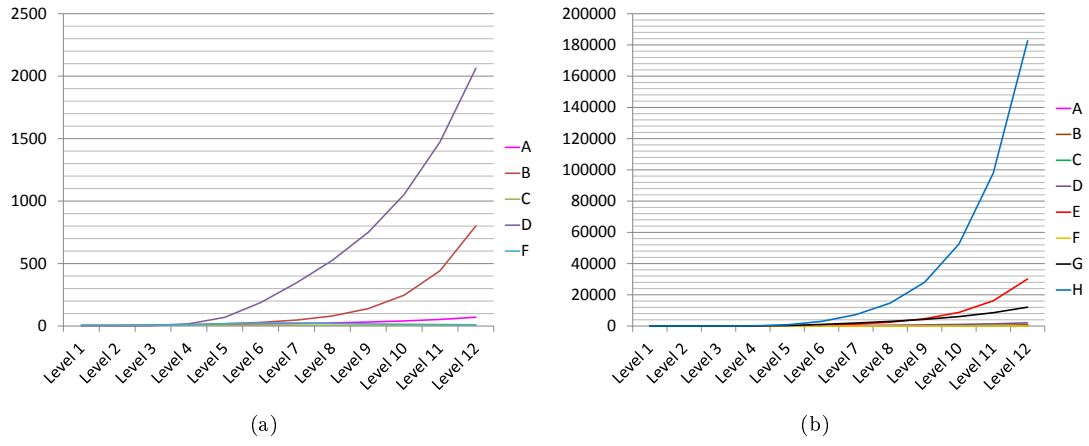


Figure 4: Energy decomposition: True variances of wavelet coefficients respective to individual levels of decomposition and various levels of long memory ($d=0.25$, $d=0.45$, $d=-0.25$) and the coefficient sets from Table E.1 (a)

D Convergence analysis

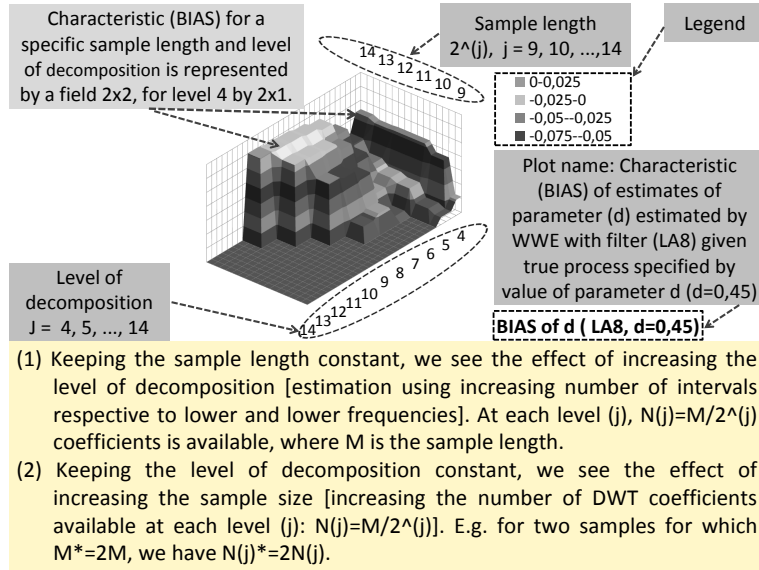


Figure 5: 3D Plots Guide

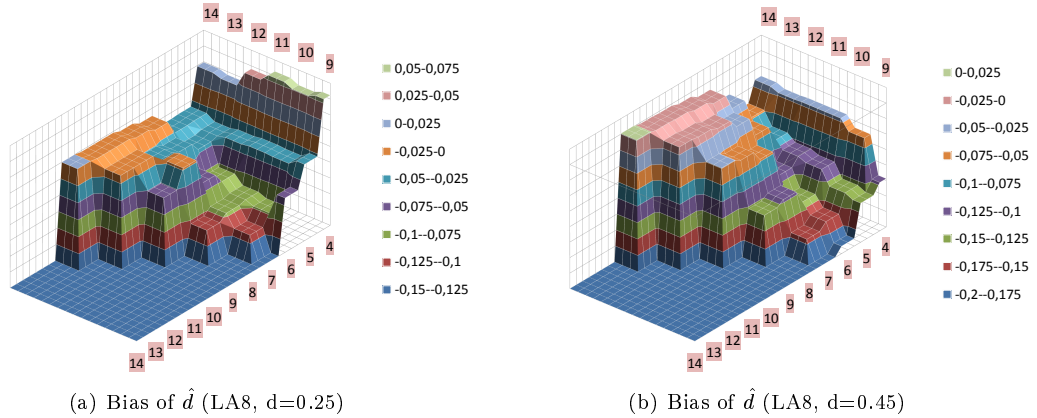
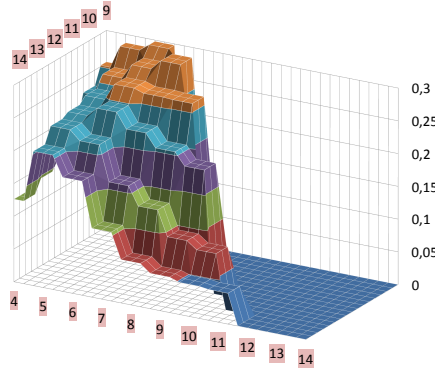
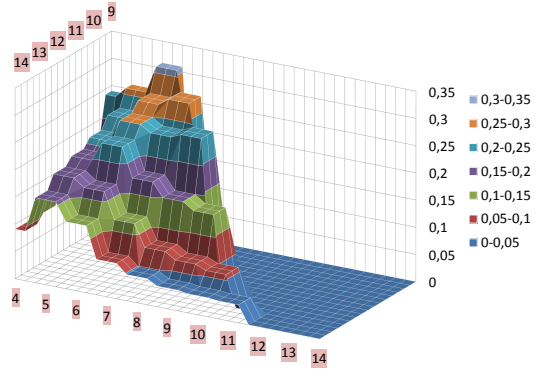


Figure 6: 3D Plots: Partial decomposition: \hat{d} : Bias

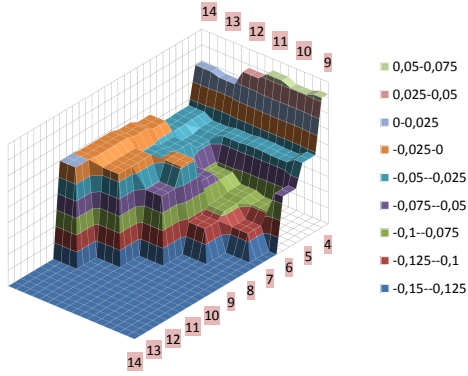


(a) RMSE of \hat{d} (LA8, $d=0.25$)

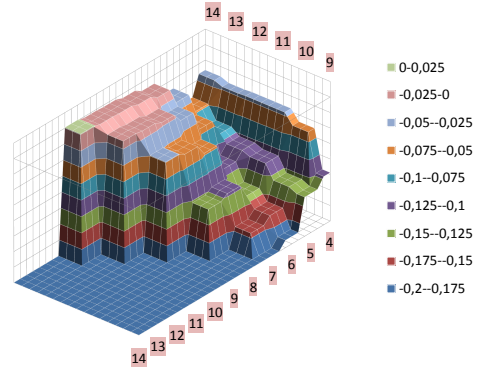


(b) RMSE of \hat{d} (LA8, $d=0.45$)

Figure 7: 3D Plots: Partial decomposition: \hat{d} : RMSE

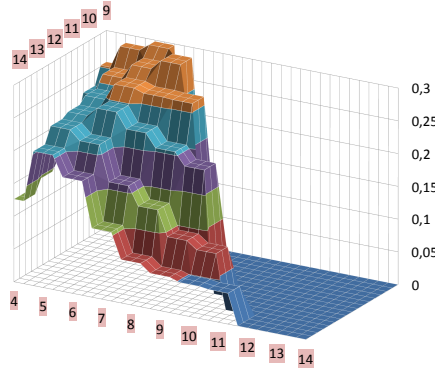


(a) Bias of $\hat{\alpha}$ (LA8, $d=0.25$)

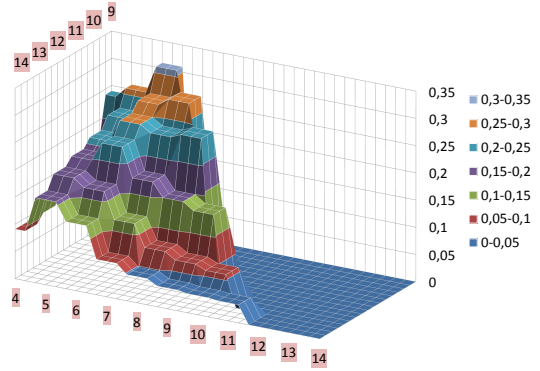


(b) Bias of $\hat{\alpha}$ (LA8, $d=0.45$)

Figure 8: 3D Plots: Partial decomposition: $\hat{\alpha}$: Bias



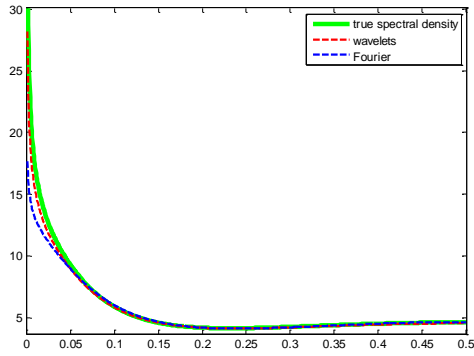
(a) RMSE of $\hat{\alpha}$ (LA8, $d=0.25$)



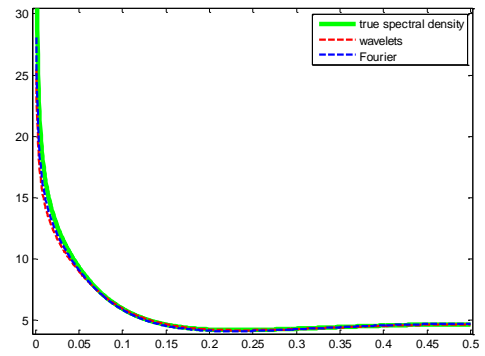
(b) RMSE of $\hat{\alpha}$ (LA8, $d=0.45$)

Figure 9: 3D Plots: Partial decomposition: $\hat{\alpha}$: RMSE

E Relative Performance



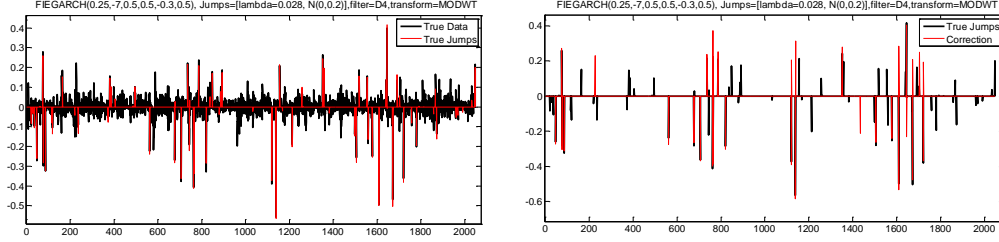
(a) $T=512$ (2^9), level=5(D4)



(b) $T=2048$ (2^{11}), level=5(D4)

Figure 10: Spectral density estimation: Wavelets (Level 5) vs Fourier

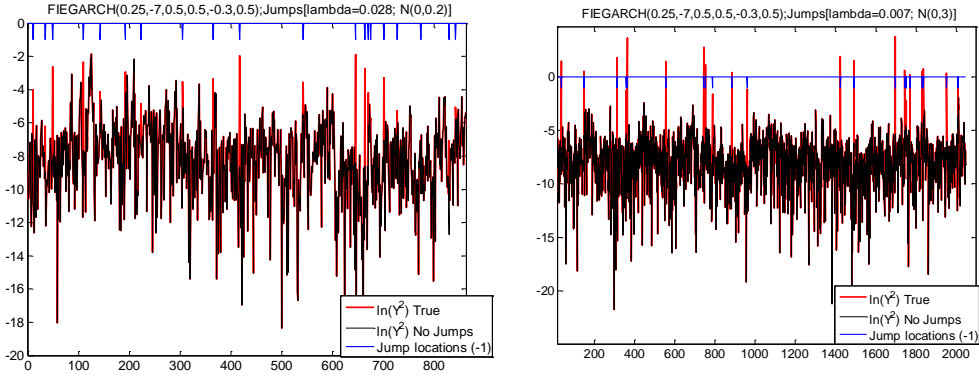
F Jumps: Features



(a) Effects of jumps in the time series

(b) Jump detection

Figure 11: Jumps in the data



(a) $\lambda=0.028$; $N(0,0.2)$

(b) Stronger jumps: $\lambda=0.007$; $N(0,3)$

Figure 12: Jumps after transformation $\ln(y^2)$

G Jumps: Results

PAR	TRUE	METHOD	No Jumps;N=2048				No Jumps;N=16384				No Jumps;N=2048			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	PAR	MEAN	BIAS	RMSE	
\hat{d}	0.250	WWE MODWT	0.165	-0.085	0.225		0.253	0.003	0.042	0.450	0.362	-0.088	0.170	
		WWE MODWT 2	0.168	-0.082	0.227		-	-	-	-	-	-	-	
		FWE	0.212	-0.038	0.147		0.251	0.001	0.036	-	0.415	-0.035	0.087	
		FWE 2	0.213	-0.037	0.146		-	-	-	-	-	-	-	
$\hat{\omega}$	-7.000	MLE	0.220	-0.030	0.085		-	-	-	-	0.433	-0.017	0.043	
		MLE 2	0.228	-0.022	0.086		-	-	-	-	-	-	-	
		MLE	-7.076	-0.076	0.174		-	-	-	-	-7.458	-0.458	0.739	
		MLE 2	-7.083	-0.083	0.182		-	-	-	-	-	-	-	
$\hat{\alpha}_2$	0.500	OTHER	-7.002	-0.002	0.197		-7.003	-0.003	0.074	-	-6.999	0.001	0.696	
		OTHER 2	-7.015	-0.015	0.198		-	-	-	-	-	-	-	
		WWE MODWT	0.434	-0.066	0.349		0.328	-0.172	0.229	-	0.324	-0.176	0.395	
		WWE MODWT 2	0.426	-0.074	0.358		-	-	-	-	-	-	-	
$\hat{\beta}_1$	0.500	FWE	0.527	0.027	0.343		0.512	0.012	0.168	-	0.475	-0.025	0.348	
		FWE 2	0.521	0.021	0.333		-	-	-	-	-	-	-	
		MLE	0.503	0.003	0.121		-	-	-	-	0.487	-0.013	0.128	
		MLE 2	0.464	-0.036	0.136		-	-	-	-	-	-	-	
$\hat{\theta}$	-0.300	WWE MODWT	0.559	0.059	0.249		0.523	0.023	0.078	-	0.610	0.110	0.178	
		WWE MODWT 2	0.561	0.061	0.253		-	-	-	-	-	-	-	
		FWE	0.520	0.020	0.199		0.499	-0.001	0.065	-	0.554	0.054	0.135	
		FWE 2	0.517	0.017	0.214		-	-	-	-	-	-	-	
$\hat{\gamma}$	0.500	MLE	0.529	0.029	0.101		-	-	-	-	0.527	0.027	0.063	
		MLE 2	0.537	0.037	0.109		-	-	-	-	-	-	-	
		WWE MODWT	-0.283	0.017	0.180		-0.337	-0.037	0.078	-	-0.314	-0.014	0.146	
		WWE MODWT 2	-0.261	0.039	0.182		-	-	-	-	-	-	-	
$\hat{\gamma}$	0.500	FWE	-0.244	0.056	0.182		-0.279	0.021	0.077	-	-0.242	0.058	0.158	
		FWE 2	-0.222	0.078	0.189		-	-	-	-	-	-	-	
		MLE	-0.301	-0.001	0.026		-	-	-	-	-0.301	-0.001	0.024	
		MLE 2	-0.282	0.018	0.031		-	-	-	-	-	-	-	
$\hat{\gamma}$	0.500	WWE MODWT	0.481	-0.019	0.196		0.489	-0.011	0.085	-	0.504	0.004	0.218	
		WWE MODWT 2	0.472	-0.028	0.193		-	-	-	-	-	-	-	
		FWE	0.509	0.009	0.175		0.504	0.004	0.083	-	0.526	0.026	0.202	
		FWE 2	0.497	-0.003	0.174		-	-	-	-	-	-	-	
$\hat{\gamma}$	0.500	MLE	0.499	-0.001	0.045		-	-	-	-	0.507	0.007	0.044	
		MLE 2	0.491	-0.009	0.048		-	-	-	-	-	-	-	

Table 2: Monte Carlo No Jumps: d=0.25/0.45; MLE, FWE,MODWT(D4), 2: Correct-
ed using Donoho and Johnstone threshold

PAR	TRUE	METHOD	Jump[0.028,N(0,0.2)];N=2048				Jump[0.028,N(0,0.2)];N=16384				Jump[0.028,N(0,0.2)];N=2048			
			MEAN	BIAS	RMSE	PAR	MEAN	BIAS	RMSE	PAR	MEAN	BIAS	RMSE	PAR
\hat{d}	0.250	WWE MODWT	0.145	-0.105	0.214	-	-	-	-	0.450	-	-	-	-
		WWE MODWT 2	0.154	-0.096	0.225	0.235	-0.015	0.042	-	-	0.347	-0.103	0.179	-
		FWE	0.195	-0.055	0.142	-	-	-	-	-	-	-	-	-
		FWE 2	0.206	-0.044	0.143	0.231	-0.019	0.038	-	-	0.403	-0.047	0.091	-
		MLE	0.018	-0.232	0.353	-	-	-	-	-	-	-	-	-
$\hat{\omega}$	-7.000	MLE 2	0.099	-0.151	0.251	-	-	-	-	-	0.187	-0.263	0.314	-
		MLE	-5.662	1.338	1.450	-	-	-	-	-	-	-	-	-
		MLE 2	-6.282	0.718	0.801	-	-	-	-	-	-5.529	1.471	1.662	-
		OTHER	-6.887	0.113	0.221	-	-	-	-	-	-	-	-	-
		OTHER 2	-6.942	0.058	0.203	-6.941	0.059	0.096	-	-	-6.946	0.054	0.677	-
$\hat{\alpha}_2$	0.500	WWE MODWT	0.475	-0.025	0.437	-	-	-	-	-	-	-	-	-
		WWE MODWT 2	0.492	-0.008	0.402	0.557	0.057	0.243	-	-	0.390	-0.110	0.447	-
		FWE	0.561	0.061	0.454	-	-	-	-	-	-	-	-	-
		FWE 2	0.582	0.082	0.400	0.731	0.231	0.297	-	-	0.535	0.035	0.428	-
		MLE	0.667	0.167	0.385	-	-	-	-	-	-	-	-	-
$\hat{\beta}_1$	0.500	MLE 2	0.652	0.152	0.287	-	-	-	-	-	0.605	0.105	0.308	-
		WWE MODWT	0.592	0.092	0.290	-	-	-	-	-	-	-	-	-
		WWE MODWT 2	0.578	0.078	0.264	0.535	0.035	0.087	-	-	0.636	0.136	0.203	-
		FWE	0.546	0.046	0.266	-	-	-	-	-	-	-	-	-
		FWE 2	0.529	0.029	0.231	0.519	0.019	0.068	-	-	0.579	0.079	0.164	-
$\hat{\theta}$	-0.300	MLE	0.406	-0.094	0.452	-	-	-	-	-	-	-	-	-
		MLE 2	0.503	0.003	0.250	-	-	-	-	-	0.619	0.119	0.240	-
		WWE MODWT	-0.491	-0.191	0.272	-	-	-	-	-	-	-	-	-
		WWE MODWT 2	-0.385	-0.085	0.189	-0.398	-0.098	0.108	-	-	-0.384	-0.084	0.174	-
		FWE	-0.455	-0.155	0.246	-	-	-	-	-	-	-	-	-
$\hat{\gamma}$	0.500	FWE 2	-0.348	-0.048	0.175	-0.356	-0.056	0.065	-	-	-0.324	-0.024	0.153	-
		MLE	-0.214	0.086	0.130	-	-	-	-	-	-	-	-	-
		MLE 2	-0.211	0.089	0.104	-	-	-	-	-	-0.176	0.124	0.137	-
		WWE MODWT	0.203	-0.297	0.365	-	-	-	-	-	-	-	-	-
		WWE MODWT 2	0.322	-0.178	0.271	0.287	-0.213	0.231	-	-	0.276	-0.224	0.322	-
$\hat{\gamma}$	0.500	FWE	0.257	-0.243	0.315	-	-	-	-	-	-	-	-	-
		FWE 2	0.365	-0.135	0.231	0.313	-0.187	0.202	-	-	0.317	-0.183	0.291	-
		MLE	0.287	-0.213	0.256	-	-	-	-	-	-	-	-	-
		MLE 2	0.340	-0.160	0.180	-	-	-	-	-	0.347	-0.153	0.179	-

Table 3: Monte Carlo Jumps: $d = 0.25/0.45$; Poisson lambda = 0.028; $N(0, 0.2)$; FWE, MODWT (D4), 2: Corrected using Donoho and Johnstone threshold

Cat	METHOD	Main Stats				MAD quantiles				
		MEAN ERR	MAD	RMSE		0.50	0.90	0.95	0.99	
in	WWE MODWT	6.1308e-05	0.00039032	0.0052969		-	-	-	-	
	WWE MODWT 2	2.2383e-05	0.00040778	0.0034541		-	-	-	-	
	WWE DWT	8.3135e-05	0.00044932	0.011577		-	-	-	-	
	WWE DWT 2	2.363e-05	0.00043981	0.0050438		-	-	-	-	
	FWE	5.9078e-05	0.00037064	0.00087854		-	-	-	-	
	FWE 2	1.4242e-05	0.00038604	0.0011961		-	-	-	-	
	MLE	6.0381e-06	9.3694e-05	0.00019804		-	-	-	-	
	MLE 2	-3.3242e-05	0.00011734	0.00028776		-	-	-	-	
	WWE MODWT	105.3851	105.3856	3277.1207		0.00015361	0.0010525	0.0019431	0.0060853	
	WWE MODWT 2	-7.6112e-05	0.00058276	0.0020436		0.00016482	0.0010512	0.0020531	0.0072711	
out	WWE DWT	0.00013817	0.00066928	0.0031579		0.00017219	0.0012156	0.0020541	0.0065611	
	WWE DWT 2	0.00087082	0.0015663	0.027558		0.00017181	0.0012497	0.0022271	0.0072191	
	FWE	2.9498e-05	0.00050763	0.0015745		0.00014566	0.0010839	0.0018926	0.005531	
	FWE 2	0.00038499	0.0010395	0.014191		0.00015425	0.0011351	0.001989	0.0081017	
	MLE	-1.5041e-06	0.00012211	0.00035147		4.0442e-05	0.00024483	0.00043587	0.001595	
	MLE 2	-0.00010455	0.00024125	0.0015605		4.3599e-05	0.00030553	0.00063378	0.0038043	
					Error quantiles A					Error quantiles B
		0.01	0.05	0.10			0.90	0.95		0.99
	WWE MODWT	-0.0033827	-0.0010722	-0.0004865			0.00063385	0.0010345		0.0048495
	WWE MODWT 2	-0.0045312	-0.0013419	-0.00062698			0.00053531	0.00089684		0.0051791
out	WWE DWT	-0.0040691	-0.001223	-0.00059588			0.00065501	0.0012118		0.004838
	WWE DWT 2	-0.003994	-0.0013952	-0.00071166			0.00059679	0.0010616		0.0050457
	FWE	-0.0035752	-0.0010712	-0.00053169			0.00061822	0.001086		0.004636
	FWE 2	-0.0042148	-0.00129	-0.00063194			0.0005657	0.0010577		0.0048622
	MLE	-0.00079412	-0.00024382	-0.00013312			0.00010297	0.00026481		0.0013195
	MLE 2	-0.0019587	-0.00046351	-0.00021734			7.5622e-05	0.00018216		0.00077541

Not Corrected: Mean of N-next true values to be forecasted: 0.0016845; Total valid=967, i.e. 96.7 % of M
fails-MLE= 0 %, fails-FWE= 0.8 %, fails-MODWT= 2.1 %, fails-DWT= 1.5 %
Corrected: Mean of N-next true values to be forecasted: 0.0016852 ; Total valid= 959, i.e. 95.9 % of M
fails-MLE= 0 %, fails-FWE= 1.3 %, fails-MODWT= 1.7 %, fails-DWT= 2.7 %

Table 4: Forecasting: N=2048; No Jumps; MLE, FWE, MODWT(D4), DWT(D4), 2:
Corrected using Donoho and Johnstone threshold

Cat	METHOD	Main Stats				MAD quantiles				
		MEAN ERR	MAD	RMSE		0.50	0.90	0.95	0.99	
in	WWE MODWT	7.522e-06	0.00015889	0.00032423		-	-	-	-	-
	WWE DWT	8.5378e-06	0.00017736	0.00038026		-	-	-	-	-
	FWE	6.3006e-06	0.00014367	0.00032474		-	-	-	-	-
	MLE	-	-	-		-	-	-	-	-
out	WWE MODWT	9.3951e-06	0.00018394	0.00054618		7.1556e-05	0.00040438	0.0007369	0.0015565	
	WWE DWT	2.1579e-05	0.0002107	0.00084705		7.1483e-05	0.00041876	0.00074376	0.0018066	
	FWE	-3.3569e-06	0.00017794	0.00067805		5.6684e-05	0.00035132	0.0005922	0.001811	
	MLE	-	-	-		-	-	-	-	-
Error quantiles A										
Error quantiles B										
out	WWE MODWT	0.01	0.05	0.10		0.90	0.95	0.99		
	WWE DWT	-0.0011566	-0.00040454	-0.00021569		0.00021003	0.0004025	0.0013515		
	FWE	-0.0011033	-0.00038209	-0.00018247		0.00023587	0.00049025	0.0016387		
	MLE	-0.00087002	-0.00034408	-0.00018571		0.00018593	0.00035741	0.0010787		
Not Corrected: Mean of N-next true values to be forecasted: 0.0015516; Total valid=1000, i.e. 100 % of M										
fails-MLE= 0 %, fails-FWE= 0 %, fails-MODWT= 0 %, fails-DWT= 0 %										

Table 5: Forecasting: N=16384; No Jumps; MLE, FWE, MODWT(D4), DWT(D4), 2:
Corrected using Donoho and Johnstone threshold

Cat	METHOD	Main Stats				MAD quantiles				
		MEAN ERR	MAD	RMSE		0.50	0.90	0.95	0.99	
in	WWE MODWT	0.00026281	0.0010841	0.02023		-	-	-	-	
	WWE DWT	0.00022639	0.0011206	0.013151		-	-	-	-	
	FWE	0.00027127	0.0010458	0.005243		-	-	-	-	
	MLE	5.7279e-05	0.00026995	0.0012167		-	-	-	-	
out	WWE MODWT	Inf	Inf	Inf		0.00016653	0.0024597	0.005308	0.040031	
	WWE DWT	924.8354	924.8375	28648.7373		0.00017788	0.0024428	0.0049679	0.040403	
	FWE	-0.00010684	0.0015807	0.0078118		0.00016022	0.0025471	0.0057388	0.031548	
	MLE	0.0002289	0.0004843	0.0039972		4.2589e-05	0.00052307	0.0010187	0.0078509	
Error quantiles A										
		0.01	0.05	0.10	Error quantiles B					
out	WWE MODWT	-0.013427	-0.0024269	-0.00087296			0.0010521	0.0026019	0.013128	
	WWE DWT	-0.013075	-0.0025811	-0.0010018			0.00089545	0.0023095	0.0165	
	FWE	-0.012356	-0.002209	-0.00081063			0.0010042	0.002777	0.014773	
	MLE	-0.0016025	-0.00044789	-0.0002568			0.00017179	0.00056713	0.0051968	
Not Corrected: Mean of N-next true values to be forecasted: 0.0064152; Total valid=962, i.e. 96.2 % of M fails-MLE= 0 %, fails-FWE= 1%, fails-MODWT= 1.9 %, fails-DWT= 2.1 %										

Table 6: Forecasting: N=2048; d=0.45; No Jumps; MLE, FWE, MODWT(D4), DWT(D4), 2: Corrected using Donoho and Johnstone threshold

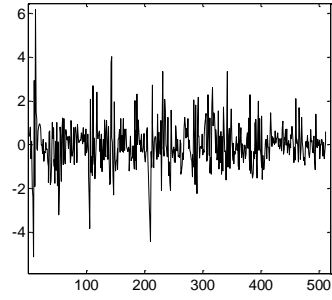
Cat	METHOD	Main Stats				MAD quantiles			
		MEAN ERR	MAD	RMSE		0.50	0.90	0.95	0.99
in	WWE MODWT	0.0024292	0.0027027	0.031109		-	-	-	-
	WWE MODWT 2	0.00049847	0.0008238	0.010266					
	WWE DWT	0.0022873	0.0025833	0.029094		-	-	-	-
	WWE DWT 2	0.00051788	0.00092081	0.013946					
	FWE	0.0024241	0.0026398	0.030474		-	-	-	-
	FWE 2	0.00046896	0.00080786	0.01562					
	MLE	0.00099962	0.0013136	0.0022127		-	-	-	-
	MLE 2	0.00021708	0.0005767	0.0011708					
	WWE MODWT	Inf	Inf	Inf		0.00027911	0.0019584	0.0043937	0.074726
	WWE MODWT 2	12837.4644	12837.4647	361761.8976		0.00020755	0.0012984	0.0021954	0.064956
out	WWE DWT	0.010776	0.010967	0.14233		0.00032328	0.0020108	0.0048951	0.086811
	WWE DWT 2	Inf	Inf	Inf		0.00019516	0.0013594	0.002609	0.08235
	FWE	0.0098899	0.010026	0.15737		0.00025416	0.0019525	0.0047286	0.073048
	FWE 2	1.4046	1.4048	36.7106		0.00017622	0.0012063	0.0024169	0.057823
	MLE	0.0014788	0.0017573	0.0066777		0.00099906	0.001951	0.0027302	0.019721
	MLE 2	0.0022114	0.0025386	0.053463		0.00034636	0.00094187	0.0016393	0.0082677
		Error quantiles A				Error quantiles B			
		0.01	0.05	0.10		0.90	0.95		0.99
	WWE MODWT	-0.0014262	-0.00061569	-0.00023517		0.0018119	0.0043937	0.074726	
	WWE MODWT 2	-0.002004	-0.0007648	-0.00033609		0.0010397	0.0018978	0.064956	
out	WWE DWT	-0.0014902	-0.00065937	-0.00028371		0.0018904	0.0048951	0.086811	
	WWE DWT 2	-0.002635	-0.00076962	-0.00031466		0.0009973	0.0020434	0.08235	
	FWE	-0.00097176	-0.00039409	-0.00021056		0.0019269	0.0047286	0.073048	
	FWE 2	-0.0019851	-0.00057545	-0.00025878		0.00087435	0.0020268	0.057823	
	MLE	-0.0033888	-0.0004285	0.00016064		0.0018323	0.0023536	0.019688	
	MLE 2	-0.0030739	-0.00075814	-0.00029836		0.00066214	0.00099507	0.0059062	
	Corrected: Mean of N-next true values to be forecasted: 0.001324; Total valid=801, i.e. 80.1 % of M								
	fails-MLE= 0 %, fails-FWE= 7 %, fails-MODWT= 13.4 %, fails-DWT= 12.5 %								
	Not Corrected: Mean of N-next true values to be forecasted: 0.001324; Total valid=45.1 , i.e. % of M								
	fails-MLE= 0 %, fails-FWE= 35.2 %, fails-MODWT= 43.4 %, fails-DWT= 41.6 %								

Table 7: Forecasting: N=2048; Jumps Lambda=0.028, N(0, 0.2); MLE, FWE,MODWT(D4),DWT(D4), 2: Corrected using Donoho and Johnstone threshold

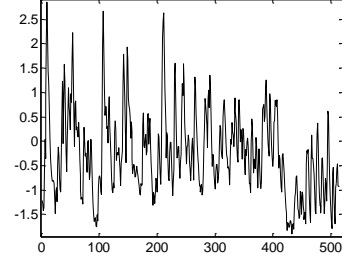
Cat	METHOD	Main Stats				MAD quantiles			
		MEAN ERR	MAD	RMSE		0.50	0.90	0.95	0.99
in	WWE MODWT	0.00079621	0.0010536	0.019562		-	-	-	-
	WWE DWT	0.00074677	0.0010165	0.018346		-	-	-	-
	FWWE	0.00052705	0.00074759	0.0094745		-	-	-	-
	MLE	-	-	-		-	-	-	-
out	WWE MODWT	Inf	Inf	Inf		0.0002394	0.0015	0.0037696	0.088609
	WWE DWT	Inf	Inf	Inf		0.00022172	0.0016482	0.0039952	0.043464
	FWWE	0.010034	0.010249	0.23919		0.00017951	0.0014685	0.0037604	0.04528
	MLE	-	-	-		-	-	-	-
Error quantiles A									
Error quantiles B									
out	WWE MODWT	0.01	0.05	0.10			0.90	0.95	0.99
	WWE DWT	-0.0025009	-0.0006746	-0.00028513			0.0011776	0.0033089	0.088609
	FWWE	-0.0025573	-0.00057909	-0.00028814			0.0012299	0.0033547	0.043464
	MLE	-0.0018013	-0.00048645	-0.00022585			0.0011465	0.0030639	0.04528
Corrected:Mean of N-next true values to be forecasted: 0.0016747; Total valid=948, i.e. 94.8 % of M fails-MLE= - %, fails-FWE= 0.4 %, fails-MODWT= 3.1 %, fails-DWT= 2.5 %									

Table 8: Forecasting: N=16384; Jumps Lambda=0.028, N(0, 0.2); MLE, FWE,MODWT(D4),DWT(D4), 2: Corrected using Donoho and Johnstone threshold

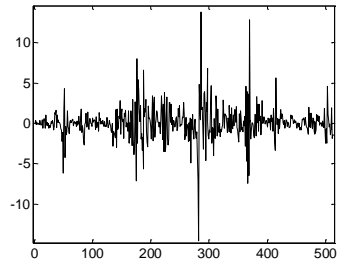
H FIEGARCH Process



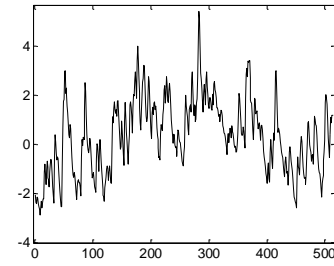
(a) FIEGARCH Process, $d=0.25$



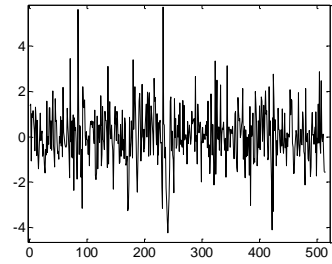
(b) Conditional Volatility, $d=0.25$



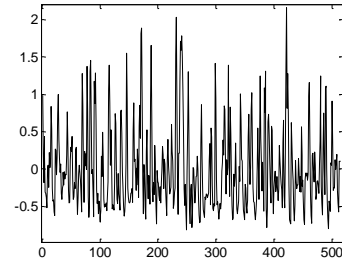
(c) FIEGARCH Process, $d=0.45$



(d) Conditional Volatility, $d=0.45$



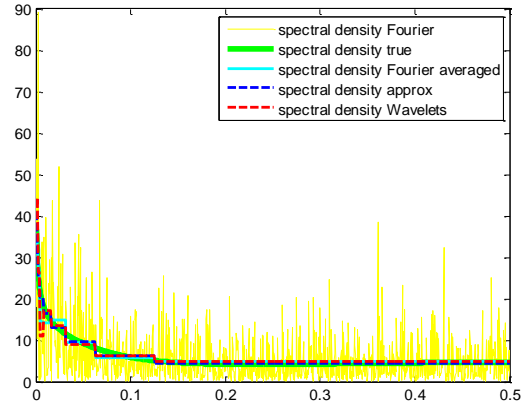
(e) FIEGARCH Process, $d=-0.25$



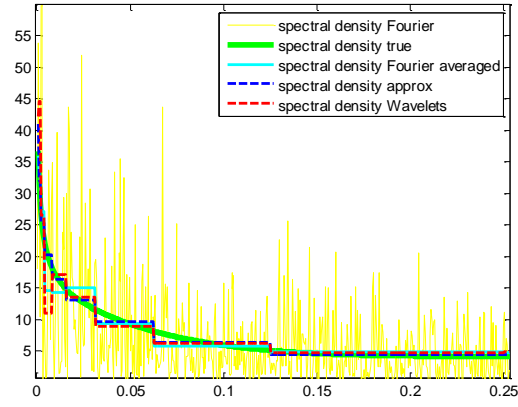
(f) Conditional Volatility, $d=-0.25$

Figure 13: FIEGARCH: Process and respective conditional volatilities for $d=0.25$, $d=0.45$, $d=-0.25$

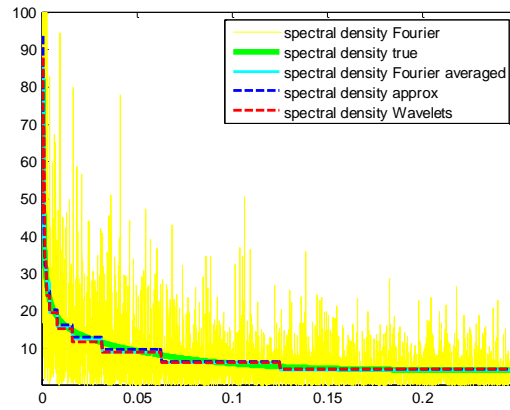
I Spectral density estimation II



(a) $T=2048$ (2^{11}), level=10

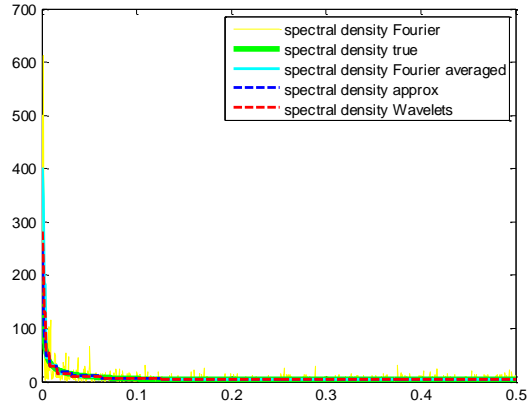


(b) $T=2048$ (2^{11}), level=10, zoom

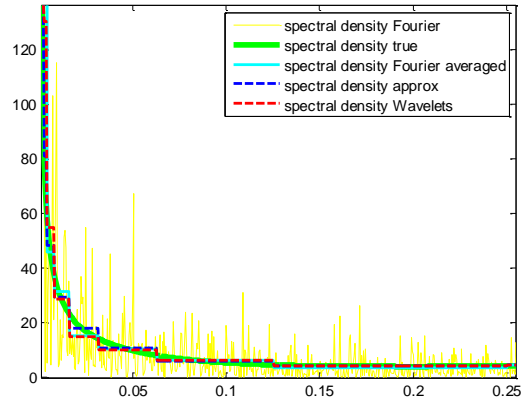


(c) $T=16384$ (2^{14}), level=13, zoom

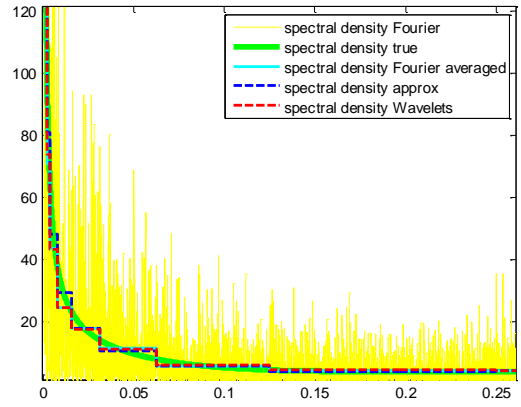
Figure 14: Spectral density estimation ($d=0.25$)



(a) $T=2048$ (2^{11}), level=10

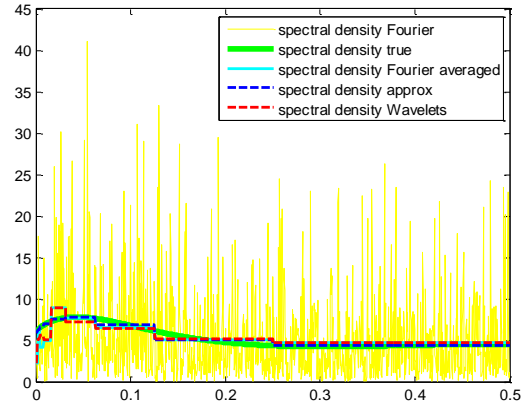


(b) $T=2048$ (2^{11}), level=10, zoom

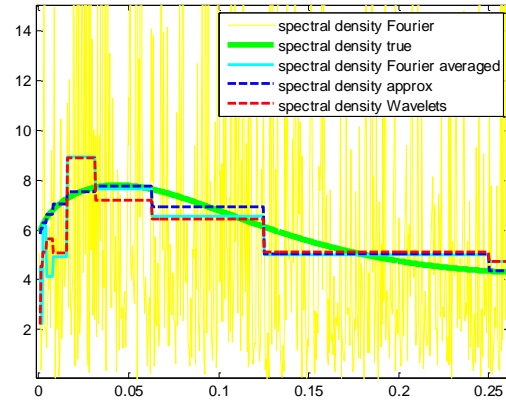


(c) $T=16384$ (2^{14}), level=13, zoom

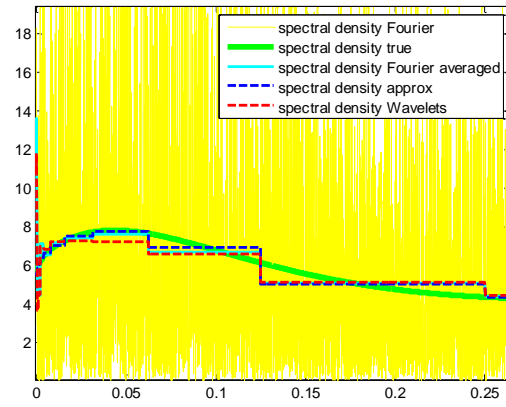
Figure 15: Spectral density estimation ($d=0.45$)



(a) $T=2048$ (2^{11}), level=10



(b) $T=2048$ (2^{11}), level=10, zoom



(c) $T=16384$ (2^{14}), level=13, zoom

Figure 16: Spectral density estimation ($d=-0.25$)

J Wavelets and Wavelet coefficients analysis

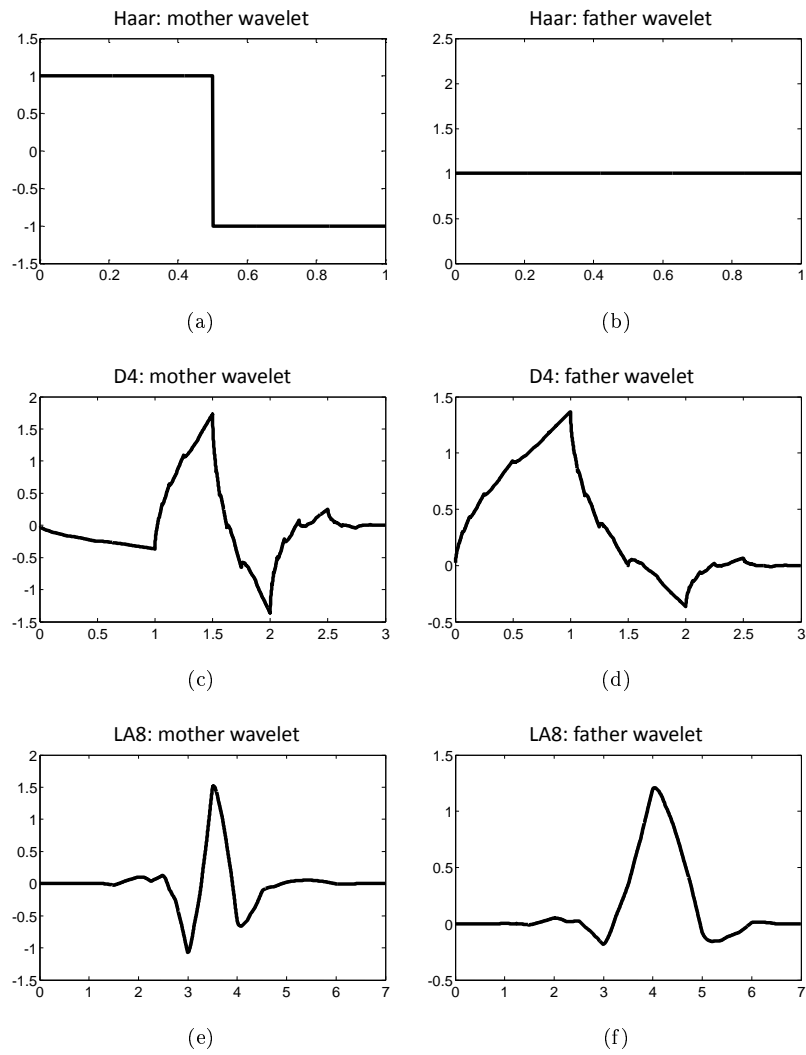


Figure 17: Wavelets: Haar, D4, LA8

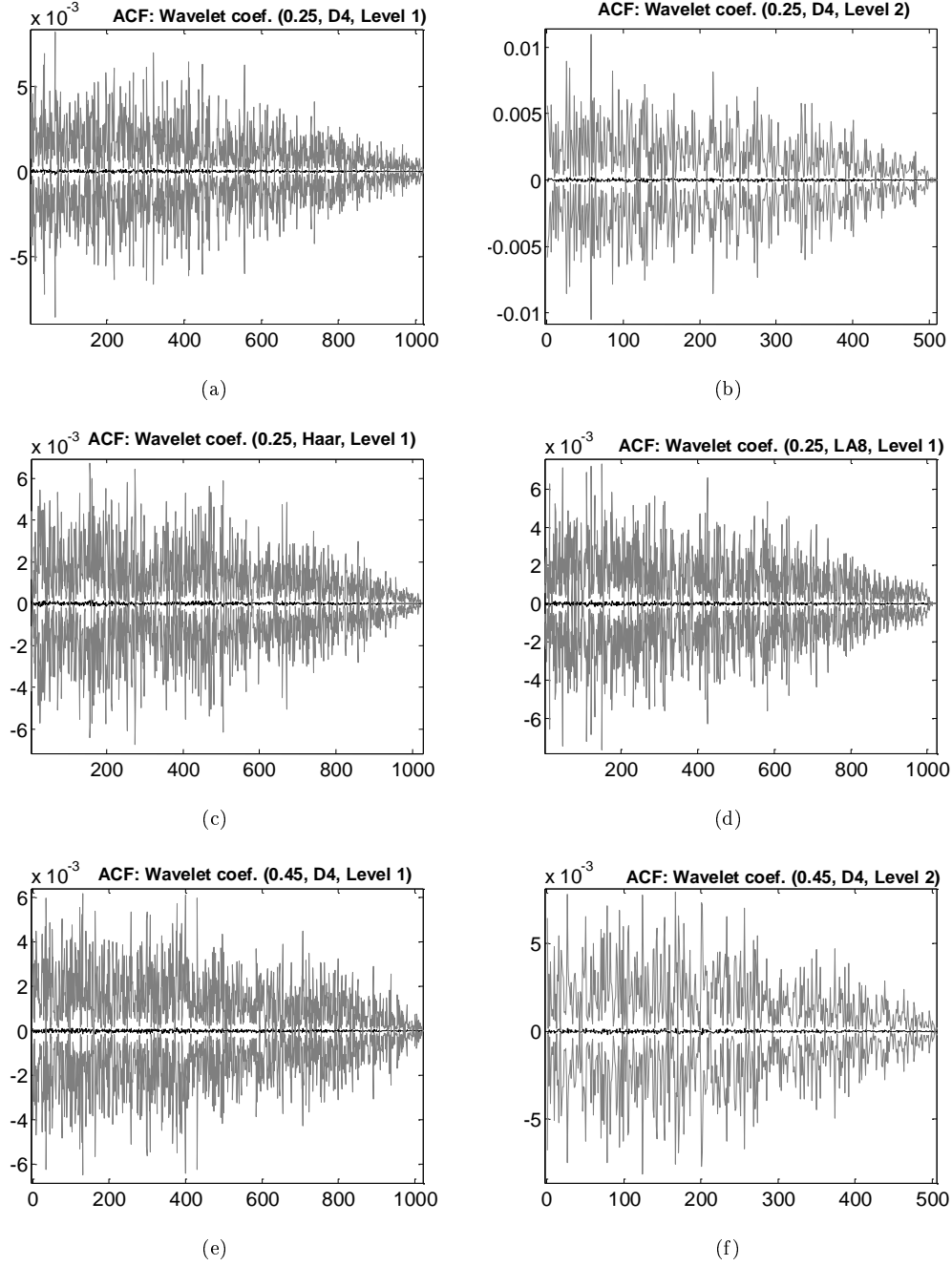


Figure 18: Estimates of autocorrelation functions (ACFs) of wavelet coefficients respective to FIEGARCH processes for $T = 2^{11}$; $d = 0.25$, $d = 0.45$ and filters Haar, D4 and LA8; both sample mean and 95% confidence intervals based on 500 FIEGARCH simulations are provided for each lag available.

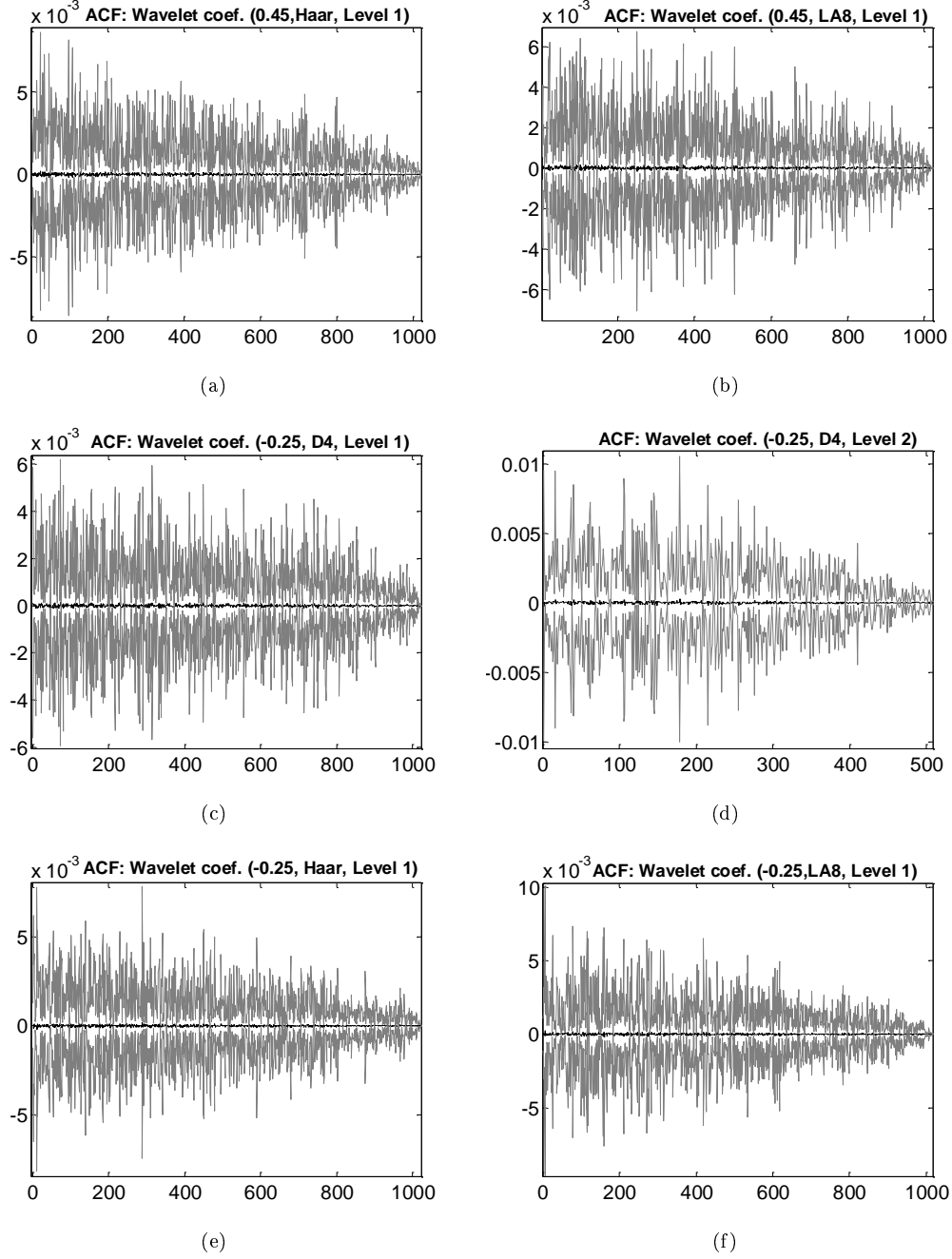


Figure 19: Estimates of autocorrelation functions (ACFs) of wavelet coefficients respective to FIEGARCH processes for $T = 2^{11}$; $d = 0.45$, $d = -0.25$ and filters Haar, D4 and LA8; both sample mean and 95% confidence intervals based on 500 FIEGARCH simulations are provided for each lag available.

K Results: Maximal Decomposition

Simulation:

- FIEGARCH(1,d,2)
- Truncation of infinite ARCH representation at lag = 1000
- Discarding first 1000 simulated numbers
- Number of simulations after truncation = 1000
- Sample lengths: $2^9, 2^{10}, 2^{11}, 2^{12}, 2^{13}, 2^{14}$
- Parameters: $(\omega_1, \alpha_1, \alpha_2, \beta_1, \theta, \gamma) = (0, 1, 0.5, 0.5, -0.3, 0.5)$, $d = 0.25, 0.45, -0.25$
- Estimators: MLE, FWE, WWE(Haar), WWE(D4), WWE(LA8)
- Levels of decomposition: Maximal (depending on filter support)

PAR	TRUE	METHOD	N=2 ⁹ (512)				N=2 ¹⁰ (1024)				N=2 ¹¹ (2048)			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	0.250	WWE1	0.109	-0.141	0.309		0.146	-0.104	0.270		0.202	-0.048	0.183	
		WWE2	0.132	-0.118	0.309		0.138	-0.112	0.287		0.171	-0.079	0.232	
		WWE3	0.176	-0.074	0.306		0.135	-0.115	0.298		0.145	-0.105	0.263	
		FWE	0.127	-0.123	0.284		0.172	-0.078	0.224		0.215	-0.035	0.147	
		MLE	0.174	-0.076	0.183		0.221	-0.029	0.111		0.232	-0.018	0.081	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.393	-0.107	0.557		0.419	-0.081	0.522		0.406	-0.094	0.504	
		WWE2	0.434	-0.066	0.580		0.502	0.002	0.499		0.500	-0.000	0.475	
		WWE3	0.439	-0.061	0.584		0.531	0.031	0.508		0.572	0.072	0.426	
		FWE	0.570	0.070	0.460		0.611	0.111	0.402		0.592	0.092	0.348	
		MLE	0.536	0.036	0.244		0.529	0.029	0.174		0.516	0.016	0.123	
$\hat{\beta}_1$	0.500	WWE1	0.542	0.042	0.392		0.529	0.029	0.347		0.509	0.009	0.285	
		WWE2	0.548	0.048	0.371		0.532	0.032	0.329		0.530	0.030	0.288	
		WWE3	0.526	0.026	0.377		0.553	0.053	0.313		0.539	0.039	0.278	
		FWE	0.523	0.023	0.347		0.512	0.012	0.277		0.501	0.001	0.203	
		MLE	0.549	0.049	0.205		0.516	0.016	0.139		0.515	0.015	0.097	
$\hat{\theta}$	-0.300	WWE1	-0.319	-0.019	0.318		-0.310	-0.010	0.250		-0.332	-0.032	0.197	
		WWE2	-0.272	0.028	0.305		-0.260	0.040	0.245		-0.278	0.022	0.195	
		WWE3	-0.256	0.044	0.307		-0.240	0.060	0.236		-0.252	0.048	0.194	
		FWE	-0.258	0.042	0.289		-0.224	0.076	0.221		-0.237	0.063	0.178	
		MLE	-0.308	-0.008	0.054		-0.302	-0.002	0.037		-0.300	-0.000	0.026	
$\hat{\gamma}$	0.500	WWE1	0.463	-0.037	0.354		0.467	-0.033	0.259		0.458	-0.042	0.209	
		WWE2	0.521	0.021	0.357		0.515	0.015	0.246		0.496	-0.004	0.191	
		WWE3	0.526	0.026	0.363		0.512	0.012	0.240		0.499	-0.001	0.180	
		FWE	0.490	-0.010	0.301		0.502	0.002	0.214		0.498	-0.002	0.172	
		MLE	0.499	-0.001	0.088		0.500	0.000	0.063		0.501	0.001	0.044	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 9: Monte Carlo 1a: d=0.25; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ¹² (4096)				N=2 ¹³ (8192)				N=2 ¹⁴ (16384)			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	0.250	WWE1	0.242	-0.008	0.103		0.252	0.002	0.058		0.257	0.007	0.041	
		WWE2	0.216	-0.034	0.148		0.244	-0.006	0.070		0.255	0.005	0.044	
		WWE3	0.199	-0.051	0.186		0.237	-0.013	0.083		0.250	-0.000	0.047	
		FWE	0.240	-0.010	0.086		0.248	-0.002	0.049		0.251	0.001	0.035	
		MLE	0.244	-0.006	0.041		-	-	-		-	-	-	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.350	-0.150	0.466		0.278	-0.222	0.411		0.240	-0.260	0.358	
		WWE2	0.444	-0.056	0.411		0.392	-0.108	0.356		0.358	-0.142	0.291	
		WWE3	0.534	0.034	0.366		0.462	-0.038	0.296		0.441	-0.059	0.241	
		FWE	0.559	0.059	0.288		0.533	0.033	0.230		0.519	0.019	0.165	
		MLE	0.503	0.003	0.086		-	-	-		-	-	-	
$\hat{\beta}_1$	0.500	WWE1	0.495	-0.005	0.226		0.512	0.012	0.161		0.518	0.018	0.110	
		WWE2	0.522	0.022	0.212		0.517	0.017	0.147		0.516	0.016	0.103	
		WWE3	0.509	0.009	0.234		0.513	0.013	0.146		0.507	0.007	0.104	
		FWE	0.495	-0.005	0.136		0.497	-0.003	0.094		0.496	-0.004	0.066	
		MLE	0.505	0.005	0.058		-	-	-		-	-	-	
$\hat{\theta}$	-0.300	WWE1	-0.348	-0.048	0.165		-0.384	-0.084	0.130		-0.396	-0.096	0.118	
		WWE2	-0.286	0.014	0.158		-0.321	-0.021	0.108		-0.334	-0.034	0.079	
		WWE3	-0.260	0.040	0.157		-0.292	0.008	0.107		-0.305	-0.005	0.073	
		FWE	-0.242	0.058	0.152		-0.268	0.032	0.105		-0.279	0.021	0.074	
		MLE	-0.301	-0.001	0.018		-	-	-		-	-	-	
$\hat{\gamma}$	0.500	WWE1	0.460	-0.040	0.160		0.455	-0.045	0.123		0.454	-0.046	0.098	
		WWE2	0.496	-0.004	0.158		0.491	-0.009	0.118		0.490	-0.010	0.091	
		WWE3	0.502	0.002	0.148		0.500	-0.000	0.112		0.500	0.000	0.086	
		FWE	0.503	0.003	0.139		0.502	0.002	0.107		0.504	0.004	0.082	
		MLE	0.501	0.001	0.029		-	-	-		-	-	-	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 10: Monte Carlo 1b: d=0.25; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ⁹ (512)				N=2 ¹⁰ (1024)				N=2 ¹¹ (2048)			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	0.450	WWE1	0.272	-0.178	0.301		0.343	-0.107	0.225		0.411	-0.039	0.120	
		WWE2	0.243	-0.207	0.322		0.294	-0.156	0.271		0.367	-0.083	0.191	
		WWE3	0.262	-0.188	0.310		0.281	-0.169	0.280		0.322	-0.128	0.246	
		FWE	0.316	-0.134	0.254		0.360	-0.090	0.201		0.425	-0.025	0.103	
		MLE	0.389	-0.061	0.152		0.424	-0.026	0.085		0.446	-0.004	0.042	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.341	-0.159	0.593		0.341	-0.159	0.573		0.287	-0.213	0.537	
		WWE2	0.414	-0.086	0.583		0.431	-0.069	0.528		0.413	-0.087	0.484	
		WWE3	0.434	-0.066	0.571		0.478	-0.022	0.513		0.494	-0.006	0.419	
		FWE	0.510	0.010	0.492		0.525	0.025	0.438		0.523	0.023	0.366	
		MLE	0.521	0.021	0.257		0.500	0.000	0.185		0.496	-0.004	0.124	
$\hat{\beta}_1$	0.500	WWE1	0.618	0.118	0.330		0.579	0.079	0.277		0.539	0.039	0.211	
		WWE2	0.628	0.128	0.323		0.622	0.122	0.267		0.574	0.074	0.214	
		WWE3	0.631	0.131	0.311		0.631	0.131	0.262		0.599	0.099	0.223	
		FWE	0.595	0.095	0.281		0.582	0.082	0.219		0.529	0.029	0.147	
		MLE	0.552	0.052	0.161		0.529	0.029	0.104		0.508	0.008	0.066	
$\hat{\theta}$	-0.300	WWE1	-0.342	-0.042	0.280		-0.351	-0.051	0.231		-0.386	-0.086	0.183	
		WWE2	-0.277	0.023	0.271		-0.276	0.024	0.220		-0.298	0.002	0.170	
		WWE3	-0.243	0.057	0.266		-0.245	0.055	0.219		-0.260	0.040	0.161	
		FWE	-0.247	0.053	0.245		-0.233	0.067	0.194		-0.245	0.055	0.147	
		MLE	-0.304	-0.004	0.050		-0.304	-0.004	0.036		-0.302	-0.002	0.024	
$\hat{\gamma}$	0.500	WWE1	0.498	-0.002	0.367		0.499	-0.001	0.291		0.501	0.001	0.226	
		WWE2	0.538	0.038	0.355		0.525	0.025	0.279		0.531	0.031	0.218	
		WWE3	0.535	0.035	0.340		0.524	0.024	0.259		0.524	0.024	0.203	
		FWE	0.511	0.011	0.315		0.515	0.015	0.245		0.521	0.021	0.190	
		MLE	0.514	0.014	0.090		0.513	0.013	0.064		0.510	0.010	0.042	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 11: Monte Carlo 2a: d=0.45; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ¹² (4096)				N=2 ¹³ (8192)				N=2 ¹⁴ (16384)			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	0.450	WWE1	0.439	-0.011	0.064		0.450	0.000	0.039		0.456	0.006	0.028	
		WWE2	0.422	-0.028	0.097		0.449	-0.001	0.043		0.458	0.008	0.031	
		WWE3	0.401	-0.049	0.137		0.437	-0.013	0.058		0.451	0.001	0.033	
		FWE	0.445	-0.005	0.050		0.454	0.004	0.034		0.456	0.006	0.026	
		MLE	0.450	0.000	0.031		-	-	-		-	-	-	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.274	-0.226	0.476		0.214	-0.286	0.423		0.211	-0.289	0.366	
		WWE2	0.374	-0.126	0.425		0.363	-0.137	0.352		0.340	-0.160	0.297	
		WWE3	0.452	-0.048	0.373		0.431	-0.069	0.304		0.421	-0.079	0.246	
		FWE	0.543	0.043	0.309		0.537	0.037	0.244		0.525	0.025	0.182	
		MLE	0.498	-0.002	0.091		-	-	-		-	-	-	
$\hat{\beta}_1$	0.500	WWE1	0.515	0.015	0.158		0.513	0.013	0.117		0.502	0.002	0.087	
		WWE2	0.539	0.039	0.153		0.513	0.013	0.108		0.504	0.004	0.081	
		WWE3	0.549	0.049	0.161		0.524	0.024	0.106		0.508	0.008	0.078	
		FWE	0.508	0.008	0.098		0.495	-0.005	0.075		0.492	-0.008	0.056	
		MLE	0.502	0.002	0.047		-	-	-		-	-	-	
$\hat{\theta}$	-0.300	WWE1	-0.410	-0.110	0.154		-0.432	-0.132	0.150		-0.437	-0.137	0.146	
		WWE2	-0.327	-0.027	0.120		-0.346	-0.046	0.086		-0.355	-0.055	0.076	
		WWE3	-0.290	0.010	0.117		-0.309	-0.009	0.072		-0.319	-0.019	0.051	
		FWE	-0.261	0.039	0.109		-0.278	0.022	0.071		-0.288	0.012	0.043	
		MLE	-0.301	-0.001	0.017		-	-	-		-	-	-	
$\hat{\gamma}$	0.500	WWE1	0.484	-0.016	0.178		0.479	-0.021	0.129		0.476	-0.024	0.096	
		WWE2	0.513	0.013	0.171		0.506	0.006	0.128		0.504	0.004	0.096	
		WWE3	0.512	0.012	0.164		0.508	0.008	0.124		0.507	0.007	0.092	
		FWE	0.508	0.008	0.154		0.505	0.005	0.116		0.504	0.004	0.088	
		MLE	0.506	0.006	0.030		-	-	-		-	-	-	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 12: Monte Carlo 2b: d=0.45; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ⁹ (512)				N=2 ¹⁰ (1024)				N=2 ¹¹ (2048)			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	-0.250	WWE1	-0.213	0.037	0.321		-0.219	0.031	0.285		-0.243	0.007	0.246	
		WWE2	-0.153	0.097	0.366		-0.222	0.028	0.301		-0.255	-0.005	0.247	
		WWE3	-0.077	0.173	0.422		-0.170	0.080	0.345		-0.275	-0.025	0.250	
		FWE	-0.249	0.001	0.263		-0.259	-0.009	0.233		-0.275	-0.025	0.202	
		MLE	-0.115	0.135	0.200		-0.141	0.109	0.158		-0.165	0.085	0.124	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.423	-0.077	0.595		0.432	-0.068	0.588		0.449	-0.051	0.541	
		WWE2	0.426	-0.074	0.607		0.521	0.021	0.540		0.523	0.023	0.503	
		WWE3	0.396	-0.104	0.622		0.474	-0.026	0.574		0.548	0.048	0.492	
		FWE	0.563	0.063	0.474		0.597	0.097	0.403		0.609	0.109	0.329	
		MLE	0.550	0.050	0.248		0.539	0.039	0.173		0.520	0.020	0.118	
$\hat{\beta}_1$	0.500	WWE1	0.472	-0.028	0.425		0.492	-0.008	0.394		0.513	0.013	0.363	
		WWE2	0.424	-0.076	0.436		0.455	-0.045	0.407		0.519	0.019	0.322	
		WWE3	0.396	-0.104	0.438		0.437	-0.063	0.406		0.528	0.028	0.313	
		FWE	0.430	-0.070	0.404		0.435	-0.065	0.365		0.469	-0.031	0.297	
		MLE	0.327	-0.173	0.291		0.356	-0.144	0.224		0.400	-0.100	0.158	
$\hat{\theta}$	-0.300	WWE1	-0.267	0.033	0.320		-0.268	0.032	0.270		-0.272	0.028	0.236	
		WWE2	-0.241	0.059	0.315		-0.233	0.067	0.265		-0.240	0.060	0.237	
		WWE3	-0.236	0.064	0.315		-0.231	0.069	0.264		-0.231	0.069	0.231	
		FWE	-0.255	0.045	0.308		-0.247	0.053	0.259		-0.228	0.072	0.224	
		MLE	-0.297	0.003	0.059		-0.298	0.002	0.040		-0.299	0.001	0.029	
$\hat{\gamma}$	0.500	WWE1	0.418	-0.082	0.295		0.398	-0.102	0.236		0.399	-0.101	0.184	
		WWE2	0.496	-0.004	0.317		0.466	-0.034	0.241		0.449	-0.051	0.178	
		WWE3	0.540	0.040	0.335		0.504	0.004	0.248		0.471	-0.029	0.163	
		FWE	0.501	0.001	0.244		0.489	-0.011	0.172		0.485	-0.015	0.133	
		MLE	0.488	-0.012	0.098		0.487	-0.013	0.068		0.493	-0.007	0.049	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 13: Monte Carlo 3a: d=-0.25; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ¹² (4096)				N=2 ¹³ (8192)				N=2 ¹⁴ (16384)			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	-0.250	WWE1	-0.246	0.004	0.212		-0.232	0.018	0.176		-0.202	0.048	0.147	
		WWE2	-0.274	-0.024	0.193		-0.269	-0.019	0.164		-0.253	-0.003	0.137	
		WWE3	-0.284	-0.034	0.197		-0.280	-0.030	0.166		-0.267	-0.017	0.133	
		FWE	-0.272	-0.022	0.168		-0.277	-0.027	0.144		-0.269	-0.019	0.119	
		MLE	-0.183	0.067	0.095		-	-	-		-	-	-	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.465	-0.035	0.481		0.441	-0.059	0.419		0.430	-0.070	0.370	
		WWE2	0.527	0.027	0.440		0.528	0.028	0.368		0.478	-0.022	0.311	
		WWE3	0.576	0.076	0.410		0.565	0.065	0.359		0.546	0.046	0.292	
		FWE	0.586	0.086	0.256		0.559	0.059	0.189		0.530	0.030	0.134	
		MLE	0.512	0.012	0.077		-	-	-		-	-	-	
$\hat{\beta}_1$	0.500	WWE1	0.522	0.022	0.303		0.507	0.007	0.268		0.486	-0.014	0.224	
		WWE2	0.528	0.028	0.284		0.521	0.021	0.236		0.526	0.026	0.195	
		WWE3	0.514	0.014	0.268		0.515	0.015	0.223		0.508	0.008	0.187	
		FWE	0.479	-0.021	0.240		0.499	-0.001	0.168		0.504	0.004	0.133	
		MLE	0.427	-0.073	0.115		-	-	-		-	-	-	
$\hat{\theta}$	-0.300	WWE1	-0.302	-0.002	0.196		-0.325	-0.025	0.161		-0.345	-0.045	0.124	
		WWE2	-0.269	0.031	0.199		-0.283	0.017	0.162		-0.307	-0.007	0.123	
		WWE3	-0.254	0.046	0.201		-0.263	0.037	0.170		-0.282	0.018	0.126	
		FWE	-0.252	0.048	0.193		-0.257	0.043	0.164		-0.274	0.026	0.125	
		MLE	-0.299	0.001	0.021		-	-	-		-	-	-	
$\hat{\gamma}$	0.500	WWE1	0.393	-0.107	0.161		0.395	-0.105	0.137		0.392	-0.108	0.126	
		WWE2	0.449	-0.051	0.134		0.446	-0.054	0.106		0.446	-0.054	0.089	
		WWE3	0.472	-0.028	0.117		0.472	-0.028	0.090		0.472	-0.028	0.071	
		FWE	0.487	-0.013	0.101		0.493	-0.007	0.078		0.496	-0.004	0.057	
		MLE	0.499	-0.001	0.033		-	-	-		-	-	-	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 14: Monte Carlo 3b: d=-0.25; MLE, FWE, Haar, D4, LA8

L Partial Decomposition II

PAR	TRUE	METHOD	N=2 ⁹ (512); level 4				N=2 ¹⁰ (1024); level 4				N=2 ¹¹ (2048); level 4			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	0.250	WWE1	0.268	0.018	0.279		0.270	0.020	0.245		0.275	0.025	0.210	
		WWE2	0.293	0.043	0.267		0.304	0.054	0.228		0.309	0.059	0.196	
		WWE3	0.306	0.056	0.252		0.301	0.051	0.226		0.307	0.057	0.189	
		FWE	0.131	-0.119	0.287		0.163	-0.087	0.236		0.214	-0.036	0.145	
		MLE	0.177	-0.073	0.176		0.212	-0.038	0.124		0.236	-0.014	0.065	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.289	-0.211	0.649		0.334	-0.166	0.610		0.302	-0.198	0.603	
		WWE2	0.328	-0.172	0.637		0.351	-0.149	0.620		0.386	-0.114	0.568	
		WWE3	0.360	-0.140	0.636		0.409	-0.091	0.592		0.431	-0.069	0.523	
		FWE	0.572	0.072	0.471		0.564	0.064	0.401		0.571	0.071	0.348	
		MLE	0.552	0.052	0.251		0.509	0.009	0.187		0.507	0.007	0.127	
$\hat{\beta}_1$	0.500	WWE1	0.525	0.025	0.420		0.504	0.004	0.365		0.528	0.028	0.279	
		WWE2	0.521	0.021	0.417		0.522	0.022	0.344		0.504	0.004	0.266	
		WWE3	0.530	0.030	0.411		0.524	0.024	0.320		0.505	0.005	0.252	
		FWE	0.525	0.025	0.350		0.522	0.022	0.292		0.512	0.012	0.199	
		MLE	0.547	0.047	0.198		0.530	0.030	0.141		0.511	0.011	0.089	
$\hat{\theta}$	-0.300	WWE1	-0.307	-0.007	0.299		-0.326	-0.026	0.273		-0.337	-0.037	0.233	
		WWE2	-0.270	0.030	0.293		-0.279	0.021	0.260		-0.285	0.015	0.217	
		WWE3	-0.252	0.048	0.284		-0.261	0.039	0.260		-0.263	0.037	0.210	
		FWE	-0.236	0.064	0.272		-0.229	0.071	0.226		-0.225	0.075	0.179	
		MLE	-0.302	-0.002	0.053		-0.305	-0.005	0.038		-0.302	-0.002	0.025	
$\hat{\gamma}$	0.500	WWE1	0.539	0.039	0.420		0.515	0.015	0.314		0.487	-0.013	0.229	
		WWE2	0.612	0.112	0.434		0.565	0.065	0.340		0.527	0.027	0.231	
		WWE3	0.613	0.113	0.444		0.567	0.067	0.327		0.533	0.033	0.228	
		FWE	0.514	0.014	0.269		0.513	0.013	0.214		0.505	0.005	0.170	
		MLE	0.496	-0.004	0.089		0.503	0.003	0.061		0.502	0.002	0.042	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 15: Monte Carlo 2a (Level 4): d=0.25; MLE, FWE, Haar, D4, LA8

PAR	TRUE	METHOD	N=2 ¹² (4096); level 4				N=2 ¹³ (8192); level 4				N=2 ¹⁴ (16384); level 4			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	0.250	WWE1	0.268	0.018	0.182		0.275	0.025	0.159		0.296	0.046	0.136	
		WWE2	0.290	0.040	0.163		0.281	0.031	0.142		0.276	0.026	0.132	
		WWE3	0.295	0.045	0.160		0.270	0.020	0.145		0.274	0.024	0.126	
		FWE	0.235	-0.015	0.091		0.251	0.001	0.050		0.250	0.000	0.035	
		MLE	0.241	-0.009	0.046		-	-	-		-	-	-	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.348	-0.152	0.495		0.310	-0.190	0.420		0.290	-0.210	0.351	
		WWE2	0.397	-0.103	0.482		0.393	-0.107	0.375		0.361	-0.139	0.313	
		WWE3	0.463	-0.037	0.429		0.469	-0.031	0.342		0.431	-0.069	0.260	
		FWE	0.555	0.055	0.287		0.547	0.047	0.228		0.516	0.016	0.168	
		MLE	0.502	0.002	0.087		-	-	-		-	-	-	
$\hat{\beta}_1$	0.500	WWE1	0.489	-0.011	0.221		0.480	-0.020	0.165		0.463	-0.037	0.135	
		WWE2	0.501	0.001	0.202		0.485	-0.015	0.152		0.492	-0.008	0.127	
		WWE3	0.484	-0.016	0.195		0.483	-0.017	0.144		0.487	-0.013	0.111	
		FWE	0.505	0.005	0.130		0.488	-0.012	0.097		0.498	-0.002	0.064	
		MLE	0.509	0.009	0.063		-	-	-		-	-	-	
$\hat{\theta}$	-0.300	WWE1	-0.357	-0.057	0.181		-0.383	-0.083	0.142		-0.393	-0.093	0.125	
		WWE2	-0.311	-0.011	0.173		-0.327	-0.027	0.121		-0.337	-0.037	0.099	
		WWE3	-0.285	0.015	0.165		-0.297	0.003	0.114		-0.309	-0.009	0.089	
		FWE	-0.246	0.054	0.147		-0.269	0.031	0.101		-0.279	0.021	0.076	
		MLE	-0.301	-0.001	0.018		-	-	-		-	-	-	
$\hat{\gamma}$	0.500	WWE1	0.461	-0.039	0.171		0.453	-0.047	0.128		0.454	-0.046	0.101	
		WWE2	0.505	0.005	0.168		0.492	-0.008	0.122		0.494	-0.006	0.092	
		WWE3	0.511	0.011	0.163		0.503	0.003	0.118		0.504	0.004	0.089	
		FWE	0.500	0.000	0.139		0.501	0.001	0.107		0.504	0.004	0.083	
		MLE	0.501	0.001	0.030		-	-	-		-	-	-	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 16: Monte Carlo 2b (Level 4): d=0.25; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ⁹ (512); level 5			N=2 ¹⁰ (1024); level 5			N=2 ¹¹ (2048); level 5		
			MEAN	BIAS	RMSE	MEAN	BIAS	RMSE	MEAN	BIAS	RMSE
\hat{d}	0.250	WWE1	0.184	-0.066	0.299	0.206	-0.044	0.276	0.190	-0.060	0.258
		WWE2	0.215	-0.035	0.289	0.214	-0.036	0.261	0.203	-0.047	0.245
		WWE3	0.221	-0.029	0.287	0.212	-0.038	0.260	0.213	-0.037	0.243
		FWE	0.127	-0.123	0.284	0.172	-0.078	0.224	0.212	-0.038	0.156
		MLE	0.174	-0.076	0.183	0.221	-0.029	0.111	0.232	-0.018	0.072
$\hat{\omega}$	0.000	ALL	-	-	-	-	-	-	-	-	-
$\hat{\alpha}_2$	0.500	WWE1	0.337	-0.163	0.627	0.399	-0.101	0.597	0.455	-0.045	0.529
		WWE2	0.375	-0.125	0.629	0.450	-0.050	0.563	0.495	-0.005	0.494
		WWE3	0.422	-0.078	0.590	0.484	-0.016	0.532	0.557	0.057	0.459
		FWE	0.570	0.070	0.460	0.611	0.111	0.402	0.583	0.083	0.341
		MLE	0.536	0.036	0.244	0.529	0.029	0.174	0.507	0.007	0.123
$\hat{\beta}_1$	0.500	WWE1	0.528	0.028	0.396	0.498	-0.002	0.365	0.494	-0.006	0.323
		WWE2	0.516	0.016	0.374	0.508	0.008	0.327	0.500	-0.000	0.298
		WWE3	0.506	0.006	0.385	0.514	0.014	0.313	0.487	-0.013	0.282
		FWE	0.523	0.023	0.347	0.512	0.012	0.277	0.502	0.002	0.202
		MLE	0.549	0.049	0.205	0.516	0.016	0.139	0.515	0.015	0.092
$\hat{\theta}$	-0.300	WWE1	-0.308	-0.008	0.316	-0.306	-0.006	0.253	-0.313	-0.013	0.198
		WWE2	-0.269	0.031	0.308	-0.257	0.043	0.244	-0.265	0.035	0.199
		WWE3	-0.251	0.049	0.301	-0.238	0.062	0.238	-0.238	0.062	0.192
		FWE	-0.258	0.042	0.289	-0.224	0.076	0.221	-0.230	0.070	0.180
		MLE	-0.308	-0.008	0.054	-0.302	-0.002	0.037	-0.301	-0.001	0.025
$\hat{\gamma}$	0.500	WWE1	0.487	-0.013	0.379	0.480	-0.020	0.275	0.465	-0.035	0.208
		WWE2	0.548	0.048	0.386	0.526	0.026	0.271	0.508	0.008	0.200
		WWE3	0.540	0.040	0.380	0.530	0.030	0.260	0.516	0.016	0.194
		FWE	0.490	-0.010	0.301	0.502	0.002	0.214	0.505	0.005	0.167
		MLE	0.499	-0.001	0.088	0.500	0.000	0.063	0.502	0.002	0.042
WWE1(Haar)	WWE2(D4)	WWE3(LA8)									

Table 17: Monte Carlo 2c (Level 5): d=0.25; MLE, FWE, Haar, D4, LA8

PAR	TRUE	METHOD	N=2 ¹² (4096); level 5			N=2 ¹³ (8192); level 5			N=2 ¹⁴ (16384); level 5		
			MEAN	BIAS	RMSE	MEAN	BIAS	RMSE	MEAN	BIAS	RMSE
\hat{d}	0.250	WWE1	0.210	-0.040	0.240	0.219	-0.031	0.229	0.236	-0.014	0.204
		WWE2	0.213	-0.037	0.229	0.208	-0.042	0.229	0.224	-0.026	0.212
		WWE3	0.211	-0.039	0.234	0.208	-0.042	0.228	0.216	-0.034	0.205
		FWE	0.240	-0.010	0.086	0.250	0.000	0.049	0.251	0.001	0.035
		MLE	0.244	-0.006	0.041	-	-	-	-	-	-
$\hat{\omega}$	0.000	ALL	-	-	-	-	-	-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.472	-0.028	0.451	0.453	-0.047	0.377	0.399	-0.101	0.337
		WWE2	0.510	0.010	0.430	0.509	0.009	0.365	0.483	-0.017	0.295
		WWE3	0.574	0.074	0.387	0.584	0.084	0.329	0.547	0.047	0.268
		FWE	0.559	0.059	0.288	0.539	0.039	0.223	0.519	0.019	0.165
		MLE	0.503	0.003	0.086	-	-	-	-	-	-
$\hat{\beta}_1$	0.500	WWE1	0.450	-0.050	0.324	0.444	-0.056	0.304	0.442	-0.058	0.279
		WWE2	0.486	-0.014	0.274	0.477	-0.023	0.277	0.469	-0.031	0.259
		WWE3	0.475	-0.025	0.271	0.468	-0.032	0.265	0.474	-0.026	0.244
		FWE	0.495	-0.005	0.136	0.493	-0.007	0.091	0.496	-0.004	0.066
		MLE	0.505	0.005	0.058	-	-	-	-	-	-
$\hat{\theta}$	-0.300	WWE1	-0.335	-0.035	0.162	-0.365	-0.065	0.122	-0.383	-0.083	0.110
		WWE2	-0.279	0.021	0.160	-0.308	-0.008	0.118	-0.325	-0.025	0.080
		WWE3	-0.255	0.045	0.158	-0.279	0.021	0.115	-0.295	0.005	0.081
		FWE	-0.242	0.058	0.152	-0.264	0.036	0.112	-0.279	0.021	0.074
		MLE	-0.301	-0.001	0.018	-	-	-	-	-	-
$\hat{\gamma}$	0.500	WWE1	0.457	-0.043	0.161	0.448	-0.052	0.126	0.452	-0.048	0.097
		WWE2	0.498	-0.002	0.160	0.489	-0.011	0.118	0.488	-0.012	0.090
		WWE3	0.505	0.005	0.152	0.499	-0.001	0.112	0.499	-0.001	0.086
		FWE	0.503	0.003	0.139	0.501	0.001	0.107	0.504	0.004	0.082
		MLE	0.501	0.001	0.029	-	-	-	-	-	-
WWE1(Haar)	WWE2(D4)	WWE3(LA8)									

Table 18: Monte Carlo 2d (Level 5): d=0.25; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ⁹ (512); level 6				N=2 ¹⁰ (1024); level 6				N=2 ¹¹ (2048); level 6			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	0.250	WWE1	0.150	-0.100	0.314		0.143	-0.107	0.307		0.148	-0.102	0.279	
		WWE2	0.144	-0.106	0.302		0.167	-0.083	0.279		0.162	-0.088	0.268	
		WWE3	0.186	-0.064	0.298		0.152	-0.098	0.283		0.157	-0.093	0.269	
		FWE	0.114	-0.136	0.297		0.174	-0.076	0.226		0.211	-0.039	0.157	
		MLE	0.170	-0.080	0.188		0.208	-0.042	0.123		0.231	-0.019	0.075	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.413	-0.087	0.567		0.449	-0.051	0.550		0.451	-0.049	0.485	
		WWE2	0.458	-0.042	0.566		0.498	-0.002	0.532		0.523	0.023	0.466	
		WWE3	0.457	-0.043	0.570		0.531	0.031	0.509		0.559	0.059	0.440	
		FWE	0.574	0.074	0.453		0.611	0.111	0.396		0.576	0.076	0.351	
		MLE	0.555	0.055	0.256		0.523	0.023	0.179		0.508	0.008	0.126	
$\hat{\beta}_1$	0.500	WWE1	0.528	0.028	0.375		0.520	0.020	0.369		0.526	0.026	0.328	
		WWE2	0.550	0.050	0.350		0.523	0.023	0.345		0.523	0.023	0.311	
		WWE3	0.541	0.041	0.324		0.548	0.048	0.301		0.534	0.034	0.286	
		FWE	0.543	0.043	0.340		0.506	0.006	0.290		0.510	0.010	0.211	
		MLE	0.554	0.054	0.197		0.529	0.029	0.144		0.517	0.017	0.096	
$\hat{\theta}$	-0.300	WWE1	-0.265	0.035	0.297		-0.307	-0.007	0.243		-0.321	-0.021	0.198	
		WWE2	-0.225	0.075	0.287		-0.258	0.042	0.239		-0.270	0.030	0.195	
		WWE3	-0.221	0.079	0.289		-0.243	0.057	0.233		-0.250	0.050	0.193	
		FWE	-0.215	0.085	0.272		-0.230	0.070	0.216		-0.236	0.064	0.178	
		MLE	-0.302	-0.002	0.051		-0.303	-0.003	0.037		-0.302	-0.002	0.026	
$\hat{\gamma}$	0.500	WWE1	0.474	-0.026	0.329		0.473	-0.027	0.254		0.459	-0.041	0.204	
		WWE2	0.520	0.020	0.327		0.516	0.016	0.245		0.496	-0.004	0.197	
		WWE3	0.528	0.028	0.318		0.511	0.011	0.232		0.503	0.003	0.189	
		FWE	0.519	0.019	0.275		0.504	0.004	0.209		0.499	-0.001	0.177	
		MLE	0.491	-0.009	0.091		0.500	0.000	0.064		0.501	0.001	0.044	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 19: Monte Carlo 2e (Level 6): d=0.25; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ¹² (4096); level 6				N=2 ¹³ (8192); level 6				N=2 ¹⁴ (16384); level 6			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	0.250	WWE1	0.194	-0.056	0.251		0.208	-0.042	0.230		0.243	-0.007	0.170	
		WWE2	0.171	-0.079	0.255		0.211	-0.039	0.210		0.223	-0.027	0.177	
		WWE3	0.170	-0.080	0.248		0.205	-0.045	0.214		0.215	-0.035	0.174	
		FWE	0.240	-0.010	0.079		0.250	0.000	0.049		0.250	0.000	0.035	
		MLE	0.243	-0.007	0.041		-	-	-		-	-	-	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.459	-0.041	0.431		0.448	-0.052	0.373		0.358	-0.142	0.330	
		WWE2	0.539	0.039	0.395		0.489	-0.011	0.349		0.424	-0.076	0.286	
		WWE3	0.583	0.083	0.364		0.568	0.068	0.319		0.488	-0.012	0.247	
		FWE	0.556	0.056	0.284		0.539	0.039	0.223		0.516	0.016	0.168	
		MLE	0.502	0.002	0.090		-	-	-		-	-	-	
$\hat{\beta}_1$	0.500	WWE1	0.472	-0.028	0.317		0.457	-0.043	0.301		0.465	-0.035	0.233	
		WWE2	0.507	0.007	0.292		0.487	-0.013	0.257		0.499	-0.001	0.209	
		WWE3	0.512	0.012	0.269		0.482	-0.018	0.247		0.504	0.004	0.195	
		FWE	0.500	-0.000	0.132		0.493	-0.007	0.091		0.498	-0.002	0.064	
		MLE	0.506	0.006	0.061		-	-	-		-	-	-	
$\hat{\theta}$	-0.300	WWE1	-0.347	-0.047	0.150		-0.367	-0.067	0.120		-0.385	-0.085	0.110	
		WWE2	-0.285	0.015	0.146		-0.311	-0.011	0.113		-0.329	-0.029	0.084	
		WWE3	-0.260	0.040	0.151		-0.281	0.019	0.114		-0.300	-0.000	0.080	
		FWE	-0.248	0.052	0.140		-0.264	0.036	0.112		-0.279	0.021	0.076	
		MLE	-0.302	-0.002	0.018		-	-	-		-	-	-	
$\hat{\gamma}$	0.500	WWE1	0.455	-0.045	0.156		0.449	-0.051	0.125		0.451	-0.049	0.099	
		WWE2	0.491	-0.009	0.149		0.489	-0.011	0.117		0.492	-0.008	0.090	
		WWE3	0.499	-0.001	0.141		0.499	-0.001	0.111		0.502	0.002	0.086	
		FWE	0.501	0.001	0.135		0.501	0.001	0.107		0.504	0.004	0.083	
		MLE	0.501	0.001	0.031		-	-	-		-	-	-	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 20: Monte Carlo 2f (Level 6): d=0.25; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ⁹ (512); level 4				N=2 ¹⁰ (1024); level 4				N=2 ¹¹ (2048); level 4			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	0.450	WWE1	0.294	-0.156	0.272		0.359	-0.091	0.201		0.368	-0.082	0.167	
		WWE2	0.309	-0.141	0.257		0.384	-0.066	0.168		0.394	-0.056	0.142	
		WWE3	0.328	-0.122	0.240		0.389	-0.061	0.167		0.405	-0.045	0.122	
		FWE	0.320	-0.130	0.250		0.371	-0.079	0.189		0.428	-0.022	0.093	
		MLE	0.388	-0.062	0.144		0.431	-0.019	0.074		0.449	-0.001	0.041	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.289	-0.211	0.643		0.296	-0.204	0.614		0.317	-0.183	0.550	
		WWE2	0.355	-0.145	0.620		0.344	-0.156	0.600		0.386	-0.114	0.515	
		WWE3	0.406	-0.094	0.601		0.409	-0.091	0.568		0.453	-0.047	0.488	
		FWE	0.507	0.007	0.499		0.502	0.002	0.424		0.542	0.042	0.368	
		MLE	0.523	0.023	0.258		0.492	-0.008	0.191		0.496	-0.004	0.129	
$\hat{\beta}_1$	0.500	WWE1	0.608	0.108	0.350		0.579	0.079	0.308		0.566	0.066	0.236	
		WWE2	0.616	0.116	0.327		0.597	0.097	0.306		0.571	0.071	0.228	
		WWE3	0.598	0.098	0.324		0.588	0.088	0.292		0.563	0.063	0.218	
		FWE	0.592	0.092	0.278		0.566	0.066	0.221		0.526	0.026	0.142	
		MLE	0.559	0.059	0.151		0.526	0.026	0.097		0.503	0.003	0.063	
$\hat{\theta}$	-0.300	WWE1	-0.321	-0.021	0.288		-0.362	-0.062	0.270		-0.376	-0.076	0.211	
		WWE2	-0.259	0.041	0.274		-0.288	0.012	0.246		-0.305	-0.005	0.202	
		WWE3	-0.232	0.068	0.261		-0.259	0.041	0.244		-0.265	0.035	0.192	
		FWE	-0.232	0.068	0.247		-0.235	0.065	0.196		-0.244	0.056	0.147	
		MLE	-0.303	-0.003	0.049		-0.306	-0.006	0.036		-0.303	-0.003	0.024	
$\hat{\gamma}$	0.500	WWE1	0.535	0.035	0.378		0.539	0.039	0.327		0.493	-0.007	0.236	
		WWE2	0.577	0.077	0.371		0.580	0.080	0.348		0.532	0.032	0.231	
		WWE3	0.566	0.066	0.371		0.568	0.068	0.319		0.527	0.027	0.224	
		FWE	0.534	0.034	0.297		0.529	0.029	0.240		0.511	0.011	0.194	
		MLE	0.514	0.014	0.088		0.515	0.015	0.062		0.511	0.011	0.043	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 21: Monte Carlo 3a (Level 4): d=0.45; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ¹² (4096); level 4				N=2 ¹³ (8192); level 4				N=2 ¹⁴ (16384); level 4			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	0.450	WWE1	0.372	-0.078	0.153		0.380	-0.070	0.126		0.388	-0.062	0.110	
		WWE2	0.396	-0.054	0.128		0.400	-0.050	0.106		0.406	-0.044	0.088	
		WWE3	0.405	-0.045	0.116		0.404	-0.046	0.105		0.409	-0.041	0.091	
		FWE	0.447	-0.003	0.053		0.454	0.004	0.034		0.456	0.006	0.026	
		MLE	0.453	0.003	0.031		-	-	-		-	-	-	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.320	-0.180	0.468		0.268	-0.232	0.416		0.243	-0.257	0.355	
		WWE2	0.374	-0.126	0.452		0.375	-0.125	0.372		0.350	-0.150	0.300	
		WWE3	0.437	-0.063	0.401		0.428	-0.072	0.334		0.433	-0.067	0.246	
		FWE	0.531	0.031	0.297		0.523	0.023	0.249		0.523	0.023	0.183	
		MLE	0.503	0.003	0.090		-	-	-		-	-	-	
$\hat{\beta}_1$	0.500	WWE1	0.536	0.036	0.189		0.543	0.043	0.153		0.536	0.036	0.116	
		WWE2	0.554	0.054	0.183		0.542	0.042	0.137		0.540	0.040	0.108	
		WWE3	0.545	0.045	0.157		0.546	0.046	0.121		0.534	0.034	0.094	
		FWE	0.503	0.003	0.102		0.496	-0.004	0.074		0.491	-0.009	0.056	
		MLE	0.498	-0.002	0.048		-	-	-		-	-	-	
$\hat{\theta}$	-0.300	WWE1	-0.392	-0.092	0.170		-0.412	-0.112	0.144		-0.424	-0.124	0.137	
		WWE2	-0.321	-0.021	0.149		-0.331	-0.031	0.098		-0.345	-0.045	0.076	
		WWE3	-0.285	0.015	0.136		-0.297	0.003	0.090		-0.309	-0.009	0.061	
		FWE	-0.259	0.041	0.112		-0.275	0.025	0.075		-0.288	0.012	0.044	
		MLE	-0.301	-0.001	0.017		-	-	-		-	-	-	
$\hat{\gamma}$	0.500	WWE1	0.494	-0.006	0.173		0.483	-0.017	0.133		0.474	-0.026	0.098	
		WWE2	0.530	0.030	0.176		0.514	0.014	0.134		0.501	0.001	0.096	
		WWE3	0.528	0.028	0.165		0.515	0.015	0.131		0.505	0.005	0.092	
		FWE	0.517	0.017	0.150		0.513	0.013	0.122		0.505	0.005	0.088	
		MLE	0.506	0.006	0.030		-	-	-		-	-	-	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 22: Monte Carlo 3b (Level 4): d=0.45; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ⁹ (512); level 5			N=2 ¹⁰ (1024); level 5			N=2 ¹¹ (2048); level 5		
			MEAN	BIAS	RMSE	MEAN	BIAS	RMSE	MEAN	BIAS	RMSE
\hat{d}	0.450	WWE1	0.272	-0.178	0.287	0.295	-0.155	0.258	0.312	-0.138	0.240
		WWE2	0.301	-0.149	0.268	0.309	-0.141	0.243	0.328	-0.122	0.216
		WWE3	0.313	-0.137	0.259	0.319	-0.131	0.229	0.341	-0.109	0.208
		FWE	0.310	-0.140	0.255	0.382	-0.068	0.174	0.426	-0.024	0.105
		MLE	0.384	-0.066	0.155	0.433	-0.017	0.073	0.445	-0.005	0.047
$\hat{\omega}$	0.000	ALL	-	-	-	-	-	-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.284	-0.216	0.642	0.334	-0.166	0.579	0.355	-0.145	0.525
		WWE2	0.361	-0.139	0.614	0.393	-0.107	0.565	0.421	-0.079	0.487
		WWE3	0.389	-0.111	0.599	0.440	-0.060	0.529	0.491	-0.009	0.449
		FWE	0.506	0.006	0.496	0.559	0.059	0.429	0.548	0.048	0.367
		MLE	0.511	0.011	0.252	0.511	0.011	0.179	0.500	-0.000	0.128
$\hat{\beta}_1$	0.500	WWE1	0.625	0.125	0.339	0.603	0.103	0.282	0.584	0.084	0.246
		WWE2	0.606	0.106	0.334	0.611	0.111	0.264	0.593	0.093	0.224
		WWE3	0.613	0.113	0.320	0.612	0.112	0.238	0.579	0.079	0.204
		FWE	0.593	0.093	0.286	0.552	0.052	0.204	0.520	0.020	0.147
		MLE	0.561	0.061	0.164	0.518	0.018	0.096	0.508	0.008	0.068
$\hat{\theta}$	-0.300	WWE1	-0.336	-0.036	0.304	-0.340	-0.040	0.241	-0.361	-0.061	0.185
		WWE2	-0.270	0.030	0.288	-0.272	0.028	0.228	-0.285	0.015	0.175
		WWE3	-0.250	0.050	0.281	-0.245	0.055	0.214	-0.252	0.048	0.165
		FWE	-0.253	0.047	0.266	-0.229	0.071	0.189	-0.242	0.058	0.149
		MLE	-0.308	-0.008	0.051	-0.302	-0.002	0.035	-0.302	-0.002	0.024
$\hat{\gamma}$	0.500	WWE1	0.518	0.018	0.412	0.513	0.013	0.303	0.497	-0.003	0.228
		WWE2	0.561	0.061	0.403	0.547	0.047	0.289	0.527	0.027	0.218
		WWE3	0.540	0.040	0.396	0.536	0.036	0.278	0.527	0.027	0.208
		FWE	0.507	0.007	0.325	0.518	0.018	0.242	0.517	0.017	0.190
		MLE	0.517	0.017	0.089	0.513	0.013	0.063	0.510	0.010	0.043
WWE1(Haar)	WWE2(D4)	WWE3(LA8)									

Table 23: Monte Carlo 3c (Level 5): d=0.45; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ¹² (4096); level 5				N=2 ¹³ (8192); level 5				N=2 ¹⁴ (16384); level 5			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	0.450	WWE1	0.336	-0.114	0.209		0.354	-0.096	0.191		0.388	-0.062	0.154	
		WWE2	0.337	-0.113	0.204		0.350	-0.100	0.186		0.376	-0.074	0.163	
		WWE3	0.346	-0.104	0.193		0.356	-0.094	0.180		0.377	-0.073	0.156	
		FWE	0.447	-0.003	0.048		0.454	0.004	0.034		0.456	0.006	0.026	
		MLE	0.453	0.003	0.030		-	-	-		-	-	-	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.331	-0.169	0.443		0.291	-0.209	0.385		0.241	-0.259	0.349	
		WWE2	0.421	-0.079	0.410		0.390	-0.110	0.344		0.359	-0.141	0.281	
		WWE3	0.482	-0.018	0.363		0.474	-0.026	0.295		0.445	-0.055	0.230	
		FWE	0.547	0.047	0.307		0.537	0.037	0.243		0.523	0.023	0.183	
		MLE	0.500	0.000	0.089		-	-	-		-	-	-	
$\hat{\beta}_1$	0.500	WWE1	0.565	0.065	0.210		0.559	0.059	0.177		0.523	0.023	0.183	
		WWE2	0.582	0.082	0.193		0.577	0.077	0.168		0.562	0.062	0.144	
		WWE3	0.575	0.075	0.175		0.568	0.068	0.155		0.556	0.056	0.136	
		FWE	0.501	0.001	0.100		0.494	-0.006	0.072		0.491	-0.009	0.056	
		MLE	0.499	-0.001	0.046		-	-	-		-	-	-	
$\hat{\theta}$	-0.300	WWE1	-0.389	-0.089	0.153		-0.414	-0.114	0.139		-0.429	-0.129	0.140	
		WWE2	-0.305	-0.005	0.130		-0.333	-0.033	0.094		-0.346	-0.046	0.072	
		WWE3	-0.269	0.031	0.128		-0.295	0.005	0.084		-0.309	-0.009	0.056	
		FWE	-0.257	0.043	0.115		-0.279	0.021	0.071		-0.288	0.012	0.044	
		MLE	-0.301	-0.001	0.017		-	-	-		-	-	-	
$\hat{\gamma}$	0.500	WWE1	0.483	-0.017	0.176		0.473	-0.027	0.130		0.476	-0.024	0.097	
		WWE2	0.515	0.015	0.175		0.501	0.001	0.127		0.501	0.001	0.096	
		WWE3	0.513	0.013	0.166		0.505	0.005	0.121		0.504	0.004	0.092	
		FWE	0.511	0.011	0.156		0.504	0.004	0.116		0.505	0.005	0.088	
		MLE	0.506	0.006	0.028		-	-	-		-	-	-	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 24: Monte Carlo 3d (Level 5): d=0.45; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ⁹ (512); level 6				N=2 ¹⁰ (1024); level 6				N=2 ¹¹ (2048); level 6			
			MEAN	BIAS	RMSE		MEAN	BIAS	RMSE		MEAN	BIAS	RMSE	
\hat{d}	0.450	WWE1	0.259	-0.191	0.314		0.278	-0.172	0.288		0.314	-0.136	0.251	
		WWE2	0.271	-0.179	0.297		0.300	-0.150	0.257		0.307	-0.143	0.249	
		WWE3	0.291	-0.159	0.282		0.291	-0.159	0.267		0.307	-0.143	0.250	
		FWE	0.302	-0.148	0.269		0.381	-0.069	0.172		0.421	-0.029	0.107	
		MLE	0.377	-0.073	0.164		0.426	-0.024	0.083		0.445	-0.005	0.043	
$\hat{\omega}$	0.000	ALL	-	-	-		-	-	-		-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.341	-0.159	0.599		0.358	-0.142	0.562		0.330	-0.170	0.507	
		WWE2	0.412	-0.088	0.581		0.409	-0.091	0.557		0.474	-0.026	0.468	
		WWE3	0.433	-0.067	0.574		0.486	-0.014	0.495		0.506	0.006	0.442	
		FWE	0.510	0.010	0.498		0.532	0.032	0.425		0.551	0.051	0.380	
		MLE	0.533	0.033	0.259		0.508	0.008	0.182		0.505	0.005	0.128	
$\hat{\beta}_1$	0.500	WWE1	0.632	0.132	0.318		0.608	0.108	0.288		0.590	0.090	0.244	
		WWE2	0.624	0.124	0.306		0.623	0.123	0.261		0.601	0.101	0.233	
		WWE3	0.625	0.125	0.289		0.617	0.117	0.254		0.609	0.109	0.223	
		FWE	0.612	0.112	0.275		0.561	0.061	0.203		0.528	0.028	0.148	
		MLE	0.564	0.064	0.166		0.525	0.025	0.102		0.507	0.007	0.064	
$\hat{\theta}$	-0.300	WWE1	-0.297	0.003	0.277		-0.352	-0.052	0.233		-0.382	-0.082	0.192	
		WWE2	-0.238	0.062	0.277		-0.278	0.022	0.224		-0.288	0.012	0.166	
		WWE3	-0.217	0.083	0.268		-0.248	0.052	0.208		-0.260	0.040	0.161	
		FWE	-0.212	0.088	0.247		-0.238	0.062	0.186		-0.248	0.052	0.145	
		MLE	-0.303	-0.003	0.049		-0.304	-0.004	0.035		-0.301	-0.001	0.024	
$\hat{\gamma}$	0.500	WWE1	0.518	0.018	0.370		0.510	0.010	0.279		0.490	-0.010	0.226	
		WWE2	0.561	0.061	0.337		0.541	0.041	0.265		0.514	0.014	0.217	
		WWE3	0.549	0.049	0.337		0.526	0.026	0.252		0.514	0.014	0.203	
		FWE	0.539	0.039	0.309		0.518	0.018	0.236		0.511	0.011	0.190	
		MLE	0.509	0.009	0.091		0.513	0.013	0.064		0.508	0.008	0.042	
WWE1(Haar)	WWE2(D4)	WWE3(LA8)												

Table 25: Monte Carlo \mathfrak{E} (Level 6): d=0.45; MLE, FWE, Haar, D4, LA8

PARAM	TRUE	METHOD	N=2 ¹² (4096); level 6			N=2 ¹³ (8192); level 6			N=2 ¹⁴ (16384); level 6		
			MEAN	BIAS	RMSE	MEAN	BIAS	RMSE	MEAN	BIAS	RMSE
\hat{d}	0.450	WWE1	0.348	-0.102	0.222	0.373	-0.077	0.191	0.422	-0.028	0.119
		WWE2	0.341	-0.109	0.219	0.378	-0.072	0.178	0.417	-0.033	0.124
		WWE3	0.334	-0.116	0.228	0.383	-0.067	0.166	0.410	-0.040	0.123
		FWE	0.446	-0.004	0.048	0.453	0.003	0.035	0.456	0.006	0.027
		MLE	0.452	0.002	0.029	-	-	-	-	-	-
$\hat{\omega}$	0.000	ALL	-	-	-	-	-	-	-	-	
$\hat{\alpha}_2$	0.500	WWE1	0.311	-0.189	0.444	0.268	-0.232	0.394	0.224	-0.276	0.361
		WWE2	0.422	-0.078	0.398	0.379	-0.121	0.351	0.354	-0.146	0.288
		WWE3	0.491	-0.009	0.355	0.465	-0.035	0.301	0.425	-0.075	0.238
		FWE	0.544	0.044	0.301	0.530	0.030	0.249	0.526	0.026	0.184
		MLE	0.498	-0.002	0.093	-	-	-	-	-	-
$\hat{\beta}_1$	0.500	WWE1	0.569	0.069	0.207	0.553	0.053	0.172	0.522	0.022	0.125
		WWE2	0.582	0.082	0.201	0.561	0.061	0.164	0.532	0.032	0.121
		WWE3	0.587	0.087	0.196	0.553	0.053	0.152	0.540	0.040	0.120
		FWE	0.505	0.005	0.100	0.497	-0.003	0.076	0.492	-0.008	0.058
		MLE	0.500	-0.000	0.048	-	-	-	-	-	-
$\hat{\theta}$	-0.300	WWE1	-0.402	-0.102	0.152	-0.420	-0.120	0.142	-0.435	-0.135	0.144
		WWE2	-0.319	-0.019	0.119	-0.338	-0.038	0.089	-0.353	-0.053	0.075
		WWE3	-0.282	0.018	0.118	-0.300	0.000	0.079	-0.318	-0.018	0.053
		FWE	-0.264	0.036	0.104	-0.276	0.024	0.076	-0.289	0.011	0.045
		MLE	-0.302	-0.002	0.017	-	-	-	-	-	-
$\hat{\gamma}$	0.500	WWE1	0.482	-0.018	0.169	0.483	-0.017	0.131	0.475	-0.025	0.095
		WWE2	0.510	0.010	0.164	0.507	0.007	0.133	0.501	0.001	0.094
		WWE3	0.511	0.011	0.155	0.511	0.011	0.127	0.503	0.003	0.090
		FWE	0.507	0.007	0.148	0.509	0.009	0.119	0.502	0.002	0.085
		MLE	0.506	0.006	0.031	-	-	-	-	-	-
WWE1(Haar)	WWE2(D4)	WWE3(LA8)									

Table 26: Monte Carlo 3f (Level 6): d=0.45; MLE, FWE, Haar, D4, LA8

M Convergence analysis II

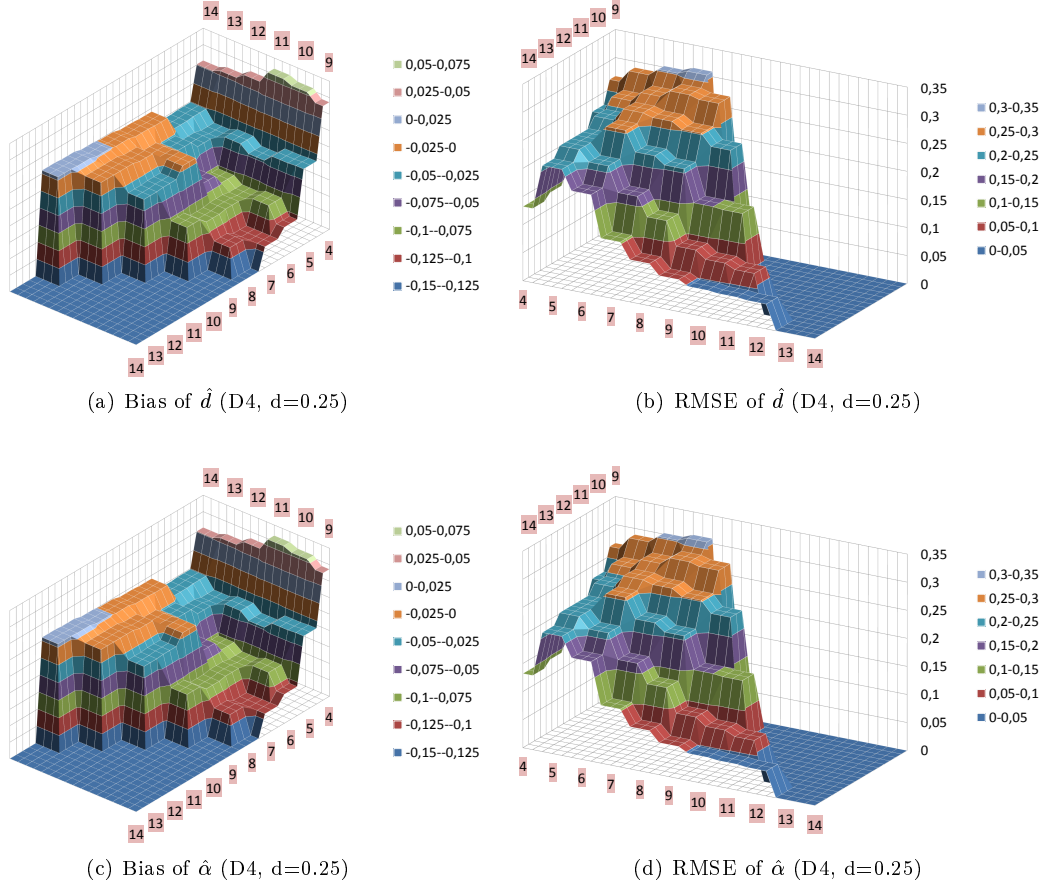
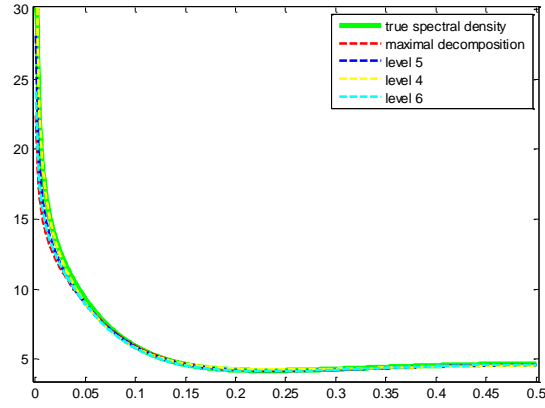
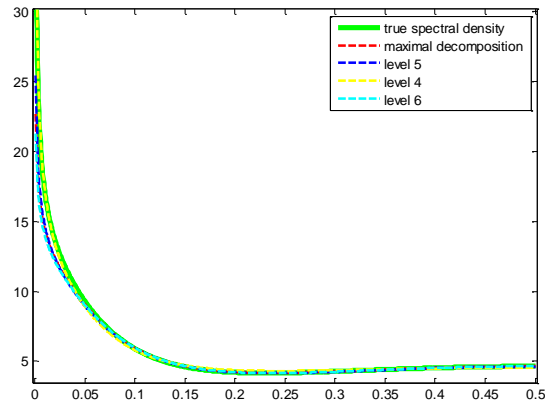


Figure 20: 3D Plots: Partial decomposition: \hat{d} , $\hat{\alpha}$: Bias, RMSE

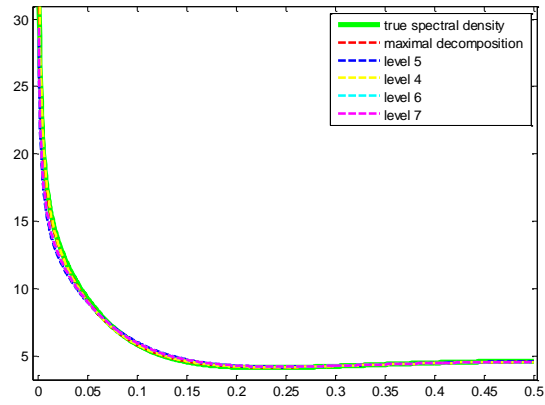
N Relative Performance



(a) $T=512$ (2^9), level= max(9), 4(Haar),5(D4),6(LA8)

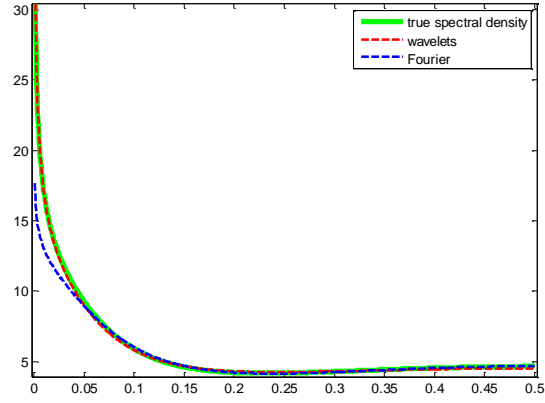


(b) $T=2048$ (2^{11}), level= max(11)(LA8), 4(Haar), 5(D4), 6(D4)

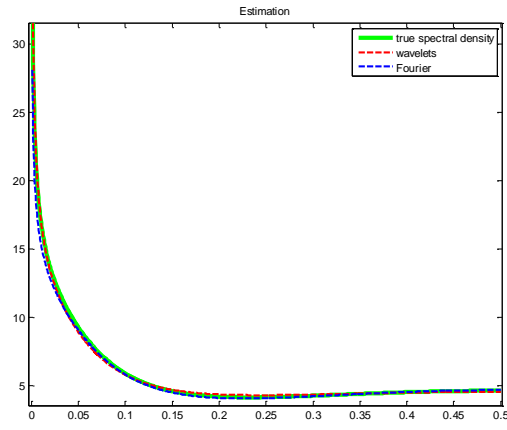


(c) $T=16384$ (2^{14}), level= max(14)(LA8), 4(LA8), 5(D4), 6(LA8), 7(LA8)

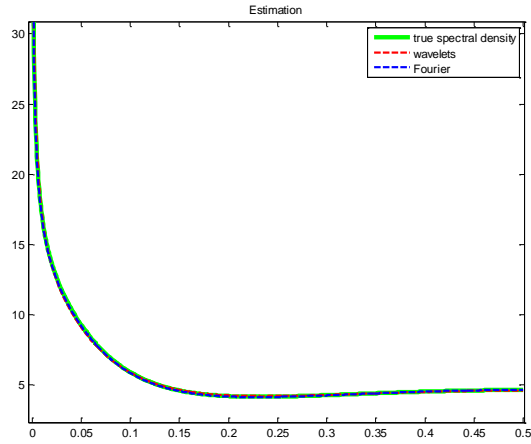
Figure 21: Spectral density estimation: Partial decomposition 0.25



(a) $T=512$ (2^9), level= 4(Haar)

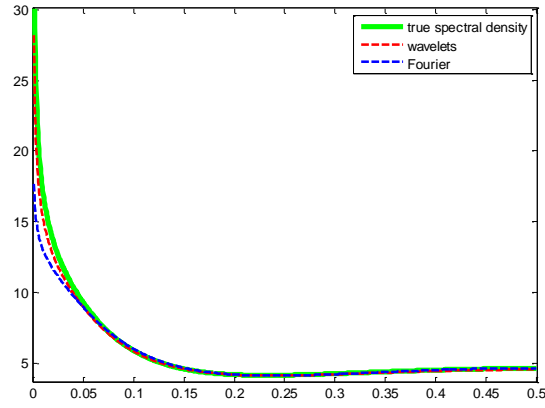


(b) $T=2048$ (2^{11}), level= 4(Haar)

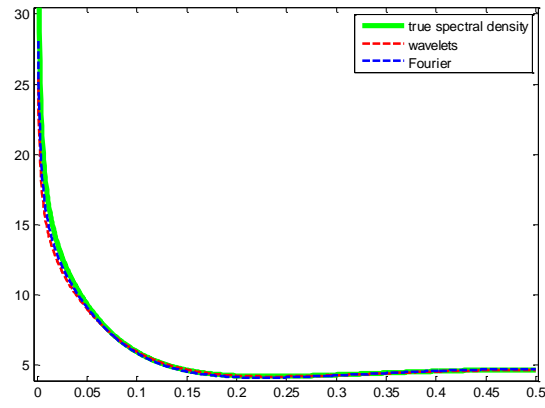


(c) $T=16384$ (2^{14}), level=4(Haar)

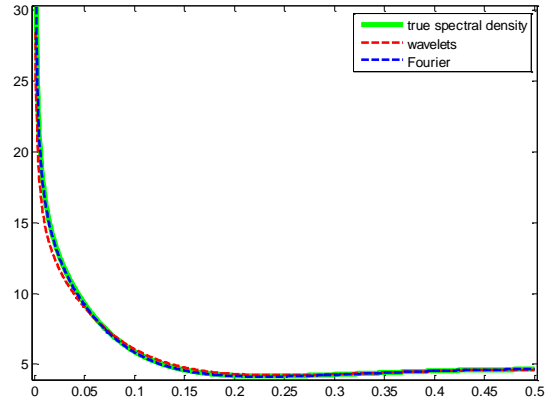
Figure 22: Spectral density estimation: Wavelets (Level 4) vs Fourier



(a) $T=512$ (2^9), level=5(D4)

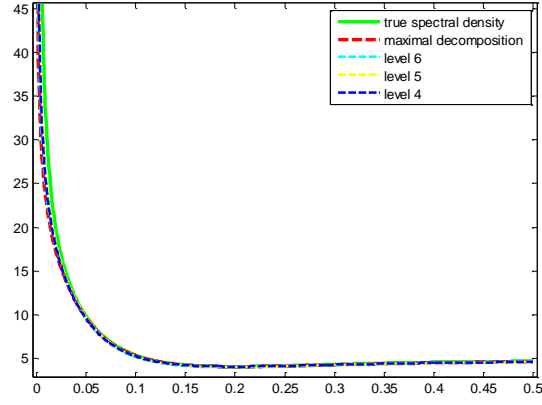


(b) $T=2048$ (2^{11}), level=5(D4)

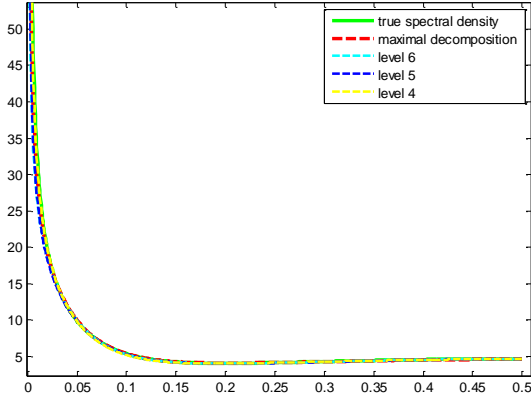


(c) $T=16384$ (2^{14}), level=5(D4)

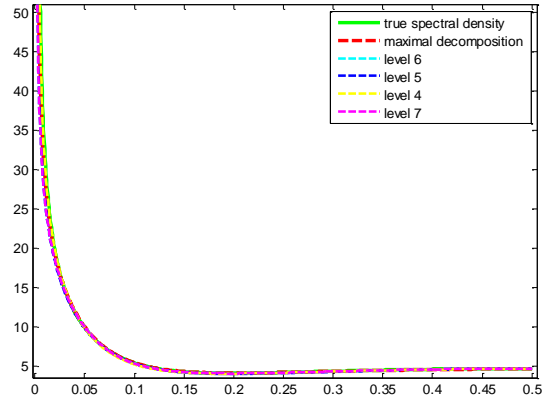
Figure 23: Spectral density estimation: Wavelets (Level 5) vs Fourier



(a) $T=512$ (2^9), level= max(9)(LA8), 4(Haar), 5(D4), 6(LA8)

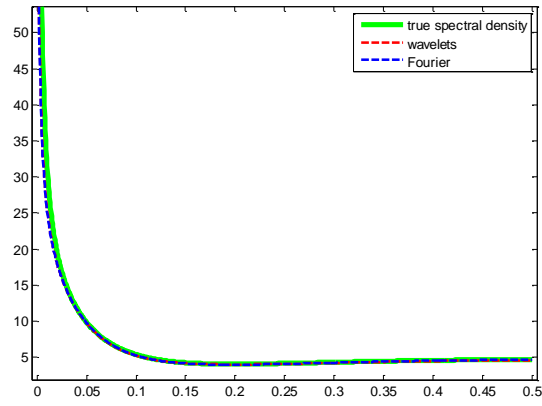


(b) $T=2048$ (2^{11}), level= max(11)(D4), 4(Haar), 5(D4), 6(D4)

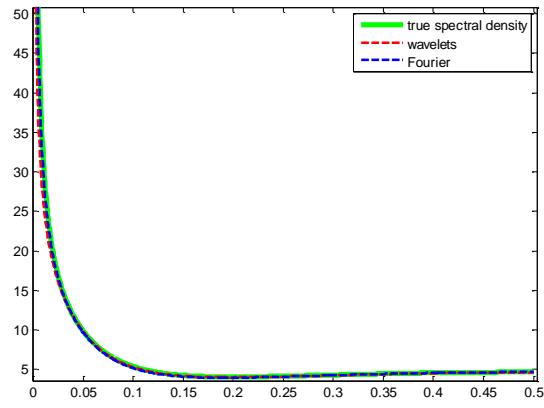


(c) $T=16384$ (2^{14}), level= max(14)(LA8), 4(LA8), 5(D4), 6(LA8), 7(LA8)

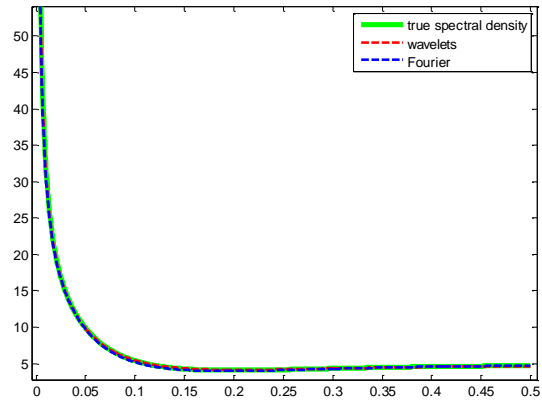
Figure 24: Spectral density estimation: Partial decomposition 0.45



(a) $T=512$ (2^9), level= 4(Haar)



(b) $T=2048$ (2^{11}), level= 5(D4)



(c) $T=16384$ (2^{14}), level= 7(LA8)

Figure 25: Spectral density estimation: Wavelets vs Fourier

IES Working Paper Series

2014

1. Diana Zigraiova, Petr Jakubik: *Systemic Event Prediction by Early Warning System*
2. Pavla Blahova, Karel Janda, Ladislav Kristoufek: *The Perspectives for Genetically Modified Cellulosic Ethanol in the Czech Republic*
3. Oxana Babecká Kucharčuková: *The impact of exchange rate volatility on trade: Evidence for the Czech Republic*
4. Oliver Polyák: *The Impact of Euro Adoption on Export Performance: Comparison of the Czech Republic and Slovakia*
5. Lucie Davidová, Vladimír Benáček: *Determinants of Austrian International Trade: Analysis Based on the Gravity Model*
6. Julie Chytilová, Václav Korbel: *Individual and Group Cheating Behavior: A Field Experiment with Adolescents*
7. Petr Janský: *Consumer Demand System Estimation and Value Added Tax Reforms in the Czech Republic*
8. Pavel Hait, Petr Janský: *Inflation Differentials among Czech Households*
9. Hana Džmuráňová, Petr Teplý: *Risk management of savings accounts*
10. Tomáš Fiala, Tomáš Havránek: *Ailing Mothers, Healthy Daughters? Contagion in the Central European Banking Sector*
11. Petra Andrlíková: *Is Barrier version of Merton model more realistic? Evidence from Europe*
12. Martin Pospíšil, Jiří Schwarz: *Bankruptcy, Investment, and Financial Constraints: Evidence from a Post-Transition Economy*
13. Matt Hunt, Goran Serdarević, Tom Ovington, Clive Kenny: *Evidence for a ladder of investment in Central and Eastern European countries*
14. Petra Andrlíková: *Bayesian default probability models*
15. Petra Buzková: *Has the Relationship Between Market and Model CDS Price Changed during the EMU Debt Crisis?*
16. Jakub Černý, Jiří Witzany: *Interest Rate Swap Credit Valuation Adjustment*
17. Martin Gregor: *Receiver's access fee for a single sender*
18. Pavel Ryska: *Deflation and Economic Growth in Long-Term Perspective*
19. Michal Paulus, Eva Michalíková, Vladimír Benáček: *German International Trade: Interpreting Export Flows According to the Gravity Model*
20. Michal Bauer, Nathan Fiala, Ian Levely: *Trusting Former Rebels: An Experimental Approach to Understanding Reintegration after Civil War*
21. Michal Skořepa: *Family Matters: Concurrent Capital Buffers in a Banking Group*
22. Martin Gregor: *Access Fees for Competing Lobbies*
23. Jozef Baruník, František Čech: *On the Modelling and Forecasting Multivariate Realized Volatility: Generalized Heterogeneous Autoregressive (GHAR) Model*
24. Roman Horváth, Ayaz Zeynalov: *The Natural Resource Curse and Institutions in Post-Soviet Countries*
25. Petr Janský: *Impact of Excise Duties on Households in the Czech Republic*

26. Petr Jakubík, Thomas Reininger: *What Are the Key Determinants of Nonperforming Loans in CESEE?*
27. Jaromír Baxa, Tomáš Olešňaník: *Current Accounts in the Eurozone Countries: The Role of Euro, Fiscal Policies and Financial Developments*
28. Petr Polák: *The Productivity Paradox: A Meta-analysis*
29. Tomáš Havránek: *Selective Reporting and the Social Cost of Carbon*
30. Jozef Baruník, Tomáš Křehlík: *Coupling High-Frequency Data with Nonlinear Models in Multiple-Step-ahead Forecasting of Energy Markets' Volatility*
31. Lenka Šťastná, Jana Votápková: *Efficiency of Hospitals in the Czech Republic: Conditional Efficiency Approach*
32. Eva Michalíková, Michal Paulus: *Gravity Model Analysis: Robust Evidence from the Czech Republic and Corruption Matching*
33. Jozef Baruník, Lucie Kraicová: *Estimation of Long Memory in Volatility Using Wavelets*

All papers can be downloaded at: <http://ies.fsv.cuni.cz>



Univerzita Karlova v Praze, Fakulta sociálních věd

Institut ekonomických studií [UK FSV – IES] Praha 1, Opletalova 26

E-mail : ies@fsv.cuni.cz

<http://ies.fsv.cuni.cz>