

Working Paper No. E2026/4

Adaptive LASSO-MGARCH for Multivariate Volatility Forecasting

Yongdeng Xu, Juyi Lyu and Wenna Lu

March 2026

ISSN 1749-6010

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
business.cardiff.ac.uk

This working paper is produced for discussion purpose. These working papers are expected to be published in due course, in revised form.

Cardiff Economics Working Papers are available online from:

http://econpapers.repec.org/paper/cdfwpaper/ and
https://carbsecon.com/wp/index_abstracts.html

Enquiries: EconWP@cardiff.ac.uk

Adaptive LASSO-MGARCH for Multivariate Volatility Forecasting

Yongdeng Xu * Juyi Lyu † Wenna Lu ‡

March 13, 2026

Abstract

This paper evaluates an Adaptive LASSO-MGARCH model for multivariate volatility forecasting, with an application to green and conventional bonds, equities, energy commodities, and EU carbon allowances. By introducing coefficient-specific adaptive penalisation directly into the multivariate GARCH variance equations, the model delivers a sparse and data-driven volatility spillover structure while preserving positive definiteness of the conditional covariance matrix. Using daily data on green and conventional bonds, equities, energy commodities, and carbon allowances, we show that adaptive regularisation substantially reduces model complexity and improves economic interpretability relative to an unpenalised MGARCH benchmark. Out-of-sample forecasting experiments at multiple horizons demonstrate that the Adaptive LASSO-MGARCH model consistently achieves lower covariance forecast losses, and statistical tests based on the White reality check confirm that these improvements are significant across alternative loss functions.

Keywords: Adaptive LASSO; Multivariate GARCH; Volatility Forecasting; High-Dimensional; Green Finance.

JEL classification: C32, C58, G17

*Corresponding Author, Cardiff University, Cardiff, UK; Email: xuy16@cardiff.ac.uk
†Loughborough University, UK; Email: julielyu2014@gmail.com
‡Cardiff Metropolitan University, Cardiff, UK; Email: wlu@cardiffmet.ac.uk

1 Introduction

Understanding how volatility is transmitted across financial markets is central to risk management, asset allocation, and financial stability analysis. Multivariate GARCH (MGARCH) models provide a natural framework for this task because they model time-varying variances and covariances jointly, allowing volatility shocks in one market to propagate to others (Hamao et al., 1990; Engle and Kroner, 1995; Bauwens et al., 2006). In applications where portfolios span heterogeneous assets, such as bonds, equities, energy commodities, and carbon permits, accurate covariance forecasts are particularly valuable because they directly determine hedge ratios, portfolio risk, and stress test outcomes.

A persistent challenge is that standard MGARCH specifications become difficult to estimate and to forecast as the cross-sectional dimension grows. Unrestricted multivariate variance dynamics typically involve $\mathcal{O}(N^2)$ parameters[1] , which can lead to overfitting and unstable forecasts in realistic sample sizes (Bauwens et al., 2006; Sun and Lin, 2012; Chuliá et al., 2019; Poignard and Fermanian, 2021; Dhaene et al., 2022a). At the same time, the conditional covariance matrix must remain positive definite at all dates, otherwise likelihood evaluation and inference break down (Engle and Kroner, 1995). The literature has proposed several solutions, but each comes with trade-offs. The BEKK family ensures positive definiteness by construction (Engle and Kroner, 1995), yet it remains parameter-heavy and quickly becomes computationally burdensome in systems beyond a handful of assets (Bauwens et al., 2006). Constant conditional correlation (CCC) models reduce dimensionality by separating univariate variances from correlation dynamics (Tse and Tsui, 2002; Engle, 2002), but flexible variants either impose strong homogeneity restrictions or reintroduce high-dimensional parameterisation when correlations are allowed to interact richly (Cappiello et al., 2006; Hafner and Franses, 2009;

[1]Although our empirical application considers $N = 8$ assets, the unrestricted MGARCH variance equations are still parameter-rich due to numerous cross-market terms; thus "high-dimensional" here refers to the dimensionality of the *parameter space* rather than N alone.

Poignard and Fermanian, 2021; Dhaene et al., 2022a). Factor-based MGARCH models achieve parsimony by attributing co-movement to a small set of latent drivers (Engle et al., 1990; Bauwens et al., 2006). Still, they rely on strong structural assumptions that can be restrictive when dependence patterns are heterogeneous across markets.

Related work has therefore explored dimension-reduction and regularisation strategies for multivariate volatility models, aiming to control parameter proliferation and improve out-of-sample covariance forecasting. In particular, L_1-type penalisation has been used to promote sparse spillover structures in multivariate GARCH settings, reducing estimation noise and enhancing interpretability (Sun and Lin, 2012; Wu and Dhaene, 2016a; Dhaene et al., 2022b). Forecast evaluation of multivariate volatility models is commonly conducted using loss functions designed for covariance matrices, and systematic comparisons highlight that gains in parsimony can translate into forecast improvements when models are assessed out of sample (Laurent et al., 2012, 2013).

This paper adopts a complementary approach: rather than imposing strong *ex ante* restrictions on the transmission structure, we allow the data to determine which volatility spillover channels are economically and statistically relevant. Specifically, we propose an *Adaptive-Lasso MGARCH* model that introduces coefficient-specific L_1 penalisation into the multivariate variance equations. Regularisation methods are well established in statistics and machine learning as tools for controlling complexity and improving out-of-sample performance (Tibshirani, 1996; Hastie et al., 2009). In econometrics, uniform LASSO penalties have been applied to multivariate volatility models to promote sparsity (Sun and Lin, 2012; Wu and Dhaene, 2016b; Dhaene et al., 2022a; Yao et al., 2024), but uniform penalisation can overshrink economically important coefficients and may induce non-negligible bias. The adaptive LASSO of Zou (2006) addresses this issue by weighting the penalty term using preliminary estimates, penalising small coefficients strongly while shrinking large coefficients less. This preserves the computational advantages of convex L_1 regularisation and, under suitable conditions, is known to deliver selection consistency

and oracle-type behaviour in the adaptive LASSO literature under suitable regularity conditions (Zou, 2006; Huang et al., 2008; Medeiros and Mendes, 2016). These properties are attractive in MGARCH settings where many cross-market links are plausibly weak, while a small subset may be persistent and forecast-relevant.

Our modelling choice combines adaptive regularisation with a CCC decomposition of the conditional covariance matrix. The CCC structure preserves positive definiteness by construction and isolates the high-dimensional component to the variance equations, where sparsity is imposed directly. The resulting framework produces (i) a parsimonious and interpretable volatility-spillover structure and (ii) covariance forecasts that are less sensitive to estimation noise than dense MGARCH benchmarks.

We apply the proposed AL-MGARCH to an eight-asset system spanning green and conventional bonds, global equities and green equities, energy commodities (oil, gas, coal), and carbon permits (EUAs). This system is economically relevant for studying transition-related risk because it links financial markets that respond to climate policy, energy-price shocks, and global risk appetite. The empirical analysis is designed to address the paper's central question: *does adaptive regularisation improve multivariate covariance forecasting relative to standard MGARCH?*

Related to the present study, Xu et al. (2025) propose an Adaptive LASSO-MGARCH framework for high-dimensional multivariate volatility modelling and demonstrate its ability to recover sparse volatility spillover structures across financial markets. Their analysis primarily focuses on structural interpretation and in-sample transmission mechanisms, while the out-of-sample predictive performance of the Adaptive-Lasso MGARCH model is not formally evaluated. The current paper complements and extends this line of research by shifting the emphasis toward forecasting performance. In particular, we conduct a systematic out-of-sample comparison between AL-MGARCH and a standard MGARCH benchmark across multiple forecast horizons using statistically validated loss-based criteria. This allows us to assess whether the theoretical advantages of adaptive

regularisation, namely sparsity, reduced estimation noise, and oracle-type behaviour, translate into economically meaningful gains in covariance forecasting accuracy. Hence, while Xu et al. (2025) establish the modelling framework and its spillover implications, this paper provides the first dedicated evidence on its forecasting effectiveness. A concise comparison between the two papers is provided in Appendix A.

The main findings are threefold. First, the Adaptive-Lasso selects a sparse set of volatility-transmission channels, shrinking many cross-effects to exactly zero while retaining economically meaningful persistence and spillovers. Second, model selection through out-of-sample tuning yields a substantially lower information criterion relative to the unpenalised benchmark, consistent with improved fit after accounting for complexity. Third, and most importantly, AL-MGARCH delivers more accurate out-of-sample covariance forecasts than MGARCH at daily, weekly, and monthly horizons under both statistical and economic loss functions; the improvements are statistically supported using the White (2000) reality check.

The remainder of the paper is organised as follows. Section 2 introduces the MGARCH benchmark and the Adaptive-Lasso extension, and describes estimation and multi-step forecasting. Section 3 summarises the data. Section 4 reports the estimation results and the out-of-sample forecasting comparison. Section 5 concludes.

2 Adaptive-Lasso MGARCH Model

2.1 Baseline specification

Let $\mathbf{r}_t$ be the $N \times 1$ vector of demeaned daily asset returns and let $\mathcal{F}_{t-1}$ denote the information set available at time $t-1$. We consider a constant conditional correlation (CCC) MGARCH model in which

$$\mathbf{r}_t = \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t = \mathbf{D}_t \mathbf{e}_t, \tag{1}$$

where $\mathbf{D}_t = \operatorname{diag}(h_{1t}^{1/2}, \ldots, h_{Nt}^{1/2})$ and $\mathbf{e}_t \mid \mathcal{F}_{t-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{P})$. The constant correlation matrix is $\mathbf{P} = \{\rho_{ij}\}$ with $\rho_{ii} = 1$. Hence, $\mathbb{E}(\boldsymbol{\varepsilon}_t \mid \mathcal{F}_{t-1}) = \mathbf{0}$ and

$$\mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \mid \mathcal{F}_{t-1}) = \mathbf{H}_t = \mathbf{D}_t \mathbf{P} \mathbf{D}_t.$$

Under CCC, $\mathbf{H}_t$ is positive definite if $\mathbf{P}$ is positive definite and diagonal elements of $\mathbf{D}_t$ are strictly positive. In our implementation, $\mathbf{P}$ is estimated as a correlation matrix of standardized residuals (hence positive semidefinite by construction) and that the variance recursion is implemented with positivity constraints on h_{it}. We adopt CCC to guarantee positive definiteness and to keep the focus on improving high-dimensional variance estimation; the same variance-regularisation approach can be combined with time-varying correlation models (e.g., DCC) in a straightforward extension.

Conditional variances follow a first-order multivariate GARCH recursion:

$$\mathbf{h}_t = \boldsymbol{\omega} + \mathbf{A}\, \boldsymbol{\varepsilon}_{t-1}^2 + \mathbf{B}\, \mathbf{h}_{t-1}, \tag{2}$$

where $\mathbf{h}_t = (h_{1t}, \ldots, h_{Nt})'$, $\boldsymbol{\omega} > 0$, and $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{N \times N}$. Stationarity requires $\rho(\mathbf{A} + \mathbf{B}) < 1$, where $\rho(\cdot)$ denotes the *spectral radius*, i.e., the largest modulus of the eigenvalues.

2.2 Quasi-maximum likelihood estimation

Under the Gaussian quasi-likelihood, the time-t log-density of $\boldsymbol{\varepsilon}_t$ is

$$\log f(\boldsymbol{\varepsilon}_t \mid \boldsymbol{\theta}) = -\frac{N}{2} \log(2\pi) - \frac{1}{2} \log |\mathbf{H}_t| - \frac{1}{2} \boldsymbol{\varepsilon}_t' \mathbf{H}_t^{-1} \boldsymbol{\varepsilon}_t, \tag{3}$$

where $\boldsymbol{\theta}$ collects the parameters in $(\boldsymbol{\omega}, \mathbf{A}, \mathbf{B}, \mathbf{P})$. The (negative) average log-likelihood is

$$\ell_T(\boldsymbol{\theta}) = -\frac{1}{T} \sum_{t=1}^{T} \log f(\boldsymbol{\varepsilon}_t \mid \boldsymbol{\theta}). \tag{4}$$

The QMLE is $\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\Theta} \ell_T(\boldsymbol{\theta})$. In practice, $\mathbf{P}$ is estimated from standardized residuals $\hat{\mathbf{e}}_t = \hat{\mathbf{D}}_t^{-1}\hat{\boldsymbol{\varepsilon}}_t$ as $\hat{\mathbf{P}} = \text{corr}(\hat{\mathbf{e}}_t)$.

2.3 Adaptive-Lasso regularisation

With unrestricted $\mathbf{A}$ and $\mathbf{B}$, the number of parameters grows on the order of N^2, and many cross-effects may be weak. To obtain a parsimonious variance transmission structure, we estimate a penalised version of the Gaussian QML objective:

$$\widehat{\boldsymbol{\theta}}(\lambda_T) = \arg\min_{\boldsymbol{\theta}\in\Theta} \left\{ \ell_T(\boldsymbol{\theta}) + \lambda_T \sum_{j=1}^{d} w_j |\theta_j| \right\}, \tag{5}$$

where $\lambda_T \geq 0$ is the shrinkage parameter and w_j are adaptive weights. We set $w_j = 1/\max\{|\tilde{\theta}_j|, \epsilon_w\}$, where θ_j is the $j$the element of the preliminary estimate $\tilde{\theta}$. We use a small constant ϵ_w to avoid division by zero, and set $\epsilon_w = 0.005$ following Cattivelli and Gallo (2019).

The AL-MGARCH estimator is obtained in three sequential steps.

1. **Initial QML.** Estimate an *unrestricted* CCC-MGARCH(1,1) on the full sample and collect $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\omega}}, \tilde{\mathbf{A}}, \tilde{\mathbf{B}})$. Adaptive weights are then defined as

$$w_j = 1/\max\{|\tilde{\theta}_j|, \epsilon_w\}, \qquad \epsilon_w = 0.005,$$

   The values of ϵ_w are chosen following Audrino and Knaus (2016) and Cattivelli and Gallo (2019).

2. **Selection of the shrinkage parameter $\hat{\lambda}_T$.** Using only the first $2T/3$ observations, solve the penalised problem (5) over a grid of $n_\lambda = 30$ equally spaced values $\{\lambda_k\}_{k=1}^{n_\lambda} \subset [\lambda_{\max}/n_\lambda, \lambda_{\max}]$, where

$$\lambda_{\max} = \frac{|\ell_T(\tilde{\boldsymbol{\theta}})|}{Tdv}, \qquad d = 2N^2 + N, \ .$$

where $v = 10$ following Cattivelli and Gallo (2019) For each λ_k, produce one-step forecasts on the hold-out sample ($T/3$ observations) and compute the QLIK loss (Bauwens and Xu, 2023)

$$\text{QLIK}_k = \sum_{t=T-T/3+1}^{T} \left\{ \text{tr}\left[(\hat{\mathbf{H}}_{t|t-1}^{(k)})^{-1} \mathbf{H}_t \right] + \log |\hat{\mathbf{H}}_{t|t-1}^{(k)}| \right\}. \tag{6}$$

Choose $\hat{\lambda}_T = \arg\min_k \text{QLIK}_k$.[2]

3. **Adaptive-Lasso estimation.** Re-estimate (5) on the full sample with $\lambda_T = \hat{\lambda}_T$, yielding $\hat{\boldsymbol{\theta}}_{\hat{\lambda}_T}$.

Under standard regularity conditions (e.g., strict stationarity and finite fourth moments) and an appropriate penalty sequence, the AL-MGARCH estimator is selection-consistent and asymptotically normal. See the Appendix B for details on the asymptotic properties.

2.4 Multi-step-ahead covariance forecasts

Under CCC, the s-step-ahead covariance forecast is

$$\widehat{\mathbf{H}}_{t+s|t} = \widehat{\mathbf{D}}_{t+s|t} \, \mathbf{P} \, \widehat{\mathbf{D}}_{t+s|t}, \qquad \widehat{\mathbf{D}}_{t+s|t} = \text{diag}\left(\widehat{\mathbf{h}}_{t+s|t}^{1/2} \right). \tag{7}$$

From (2), the one-step variance forecast is

$$\widehat{\mathbf{h}}_{t+1|t} = \boldsymbol{\omega} + \mathbf{A} \, \boldsymbol{\varepsilon}_t^2 + \mathbf{B} \, \mathbf{h}_t. \tag{8}$$

For $s \geq 2$, $\boldsymbol{\varepsilon}_{t+s-1}^2$ is unobserved at time t and is replaced by its conditional expectation, and

$$\mathbf{C} = \mathbf{A} + \mathbf{B}, \tag{9}$$

[2] A BIC or economic-loss criterion is admissible; see Bauwens and Xu, 2025.

The variance forecasts satisfy

$$\widehat{\mathbf{h}}_{t+s|t} = \boldsymbol{\omega} + \mathbf{C}\,\widehat{\mathbf{h}}_{t+s-1|t}, \qquad s \geq 2. \tag{10}$$

Solving recursively gives the closed-form expression

$$\widehat{\mathbf{h}}_{t+s|t} = \widetilde{\boldsymbol{\omega}}_s + \mathbf{C}^{\,s-1}\widehat{\mathbf{h}}_{t+1|t}, \qquad \widetilde{\boldsymbol{\omega}}_s = (\mathbf{I} - \mathbf{C})^{-1}(\mathbf{I} - \mathbf{C}^{\,s-1})\boldsymbol{\omega}, \tag{11}$$

for $\rho(\mathbf{C}) < 1$. Equations (7)–(11) define the s-step-ahead covariance forecasts used below.

3 Data

We study an eight-asset system designed to capture volatility transmission across (i) fixed income (green and conventional bonds), (ii) energy and commodity markets (Brent crude, natural gas, coal), (iii) carbon pricing (EU Emissions Allowances), and (iv) global equities (broad market and green/ESG equities). All series are daily close-to-close prices in USD (or USD-hedged for AGGB). Returns are log-differences multiplied by 100.

Green bonds are proxied by the Solactive Green Bond Index (GBOND). Conventional bonds are proxied by the USD-hedged Bloomberg Global Aggregate Bond Index (AGGB). Energy commodities include Brent crude futures (BRENT), thermal coal (COAL), and natural gas (GAS). Carbon risk is captured by EU Emissions Allowances (EUA). Global equity risk is represented by MSCI ACWI (STOCK) and green equities by the MSCI Global Environment Index (ESG).

The sample spans January 2018 to June 2025 with $T = 1509$ observations, which is the longest common window across all series.

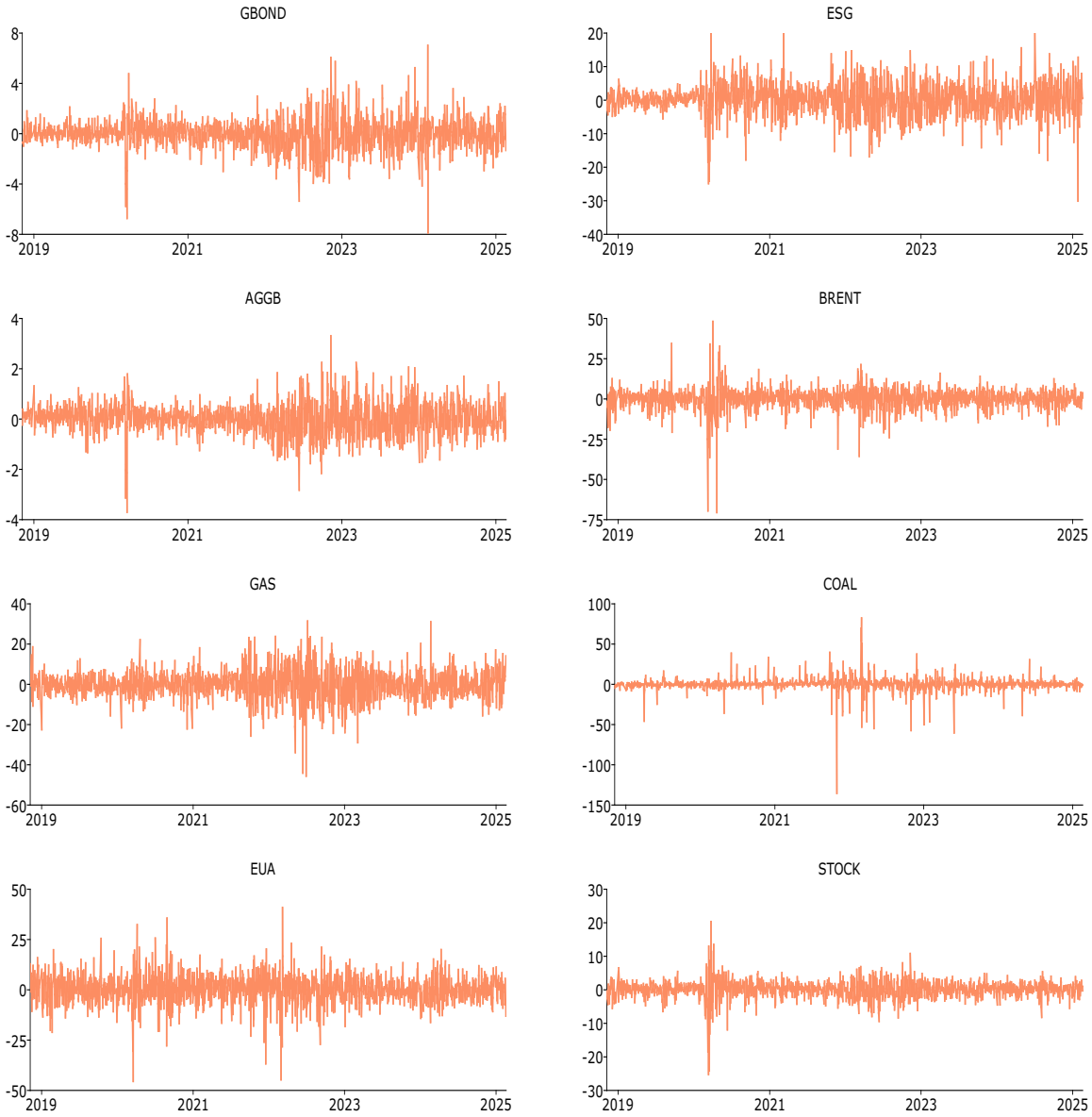


Figure 1 – Time evolution of returns for the eight assets

Figure 1 plots daily log-returns for the eight assets over the common sample (January 2018–June 2025). The sample is constructed as the longest overlapping window across all series; if any asset is missing on a given trading day, that date is removed for all assets to maintain a balanced panel. The figure also highlights several notable volatility

episodes. Volatility rises sharply during the March 2020 COVID-19 sell-off across bonds, equities, EUA and Brent, while GAS and COAL are comparatively less affected. From mid-2021, EUA enters a persistent high-volatility phase, and from February 2022 GAS and COAL display pronounced swings around the Russia–Ukraine energy shock. We retain these observations unchanged (i.e., no trimming or winsorisation), as they reflect economically meaningful shocks that the model is designed to capture.

Table 1 reports descriptive statistics. It reveals three stylised facts. First, at the mean level, most series are mean-reverting around zero: EU Emission Allowances (EUA) are the only asset with a significant positive drift (0.333), while the green-equity index (ESG, 0.095) and the conventional equity index (STOCK, 0.086) show merely modest upticks. By contrast, natural gas records a small negative drift (GAS, -0.09), reflecting the prolonged weakness in global gas prices. The aggregate bond index (AGGB, -0.03) also trends slightly downward, consistent with the post-2017 'repricing of duration' triggered by rising long-term yields. Second, volatility varies widely across markets. The standard deviations of crude oil (6.613) and natural gas (7.116) are more than six times that of conventional bonds (AGGB, 0.9), with green bonds (GBOND, 1.193) sitting in between. This pattern aligns with the summary statistics reported in Reboredo et al. (2020) and Mensi et al. (2024). Third, all distributions deviate from normality. Every asset shows negative skewness, reflecting frequent downside moves, and kurtosis is far above the Gaussian benchmark of three, reaching 42.6 for conventional bonds and 69.3 for coal. These results indicate fat tails and sharp peaks, characteristic of heavy-tailed behaviour. The Ljung-Box statistics for squared returns (lag 20) are highly significant in every column, confirming pronounced ARCH effects and justifying the use of an MGARCH specification in the subsequent analysis.

Table 1 – Descriptive statistics of daily log-returns

	GBOND	ESG	AGGB	BRENT	GAS	COAL	EUA	STOCK
Min	-7.874	-24.843	-13.95	-70.50	-51.553	-135.294	-49.022	-25.192
Max	7.010	37.313	6.03	48.075	31.369	82.206	40.802	20.308
Mean	0.001	0.095	-0.03	0.036	-0.090	0.019	0.333	0.086
Std	1.193	4.764	0.90	6.613	7.116	8.046	7.464	2.643
Skew	-0.261	-0.001	-2.76	-1.344	-0.353	-2.606	-0.333	-1.059
Kurt	8.507	7.720	42.60	21.771	6.920	69.350	7.372	17.253
LB^2	312.148	251.603	10.501	435.717	398.796	59.899	213.785	1581.352

Note. GBOND: Solactive Green Bond Index; ESG: MSCI Global Environment index; AGGB: USD-hedged Bloomberg Global Aggregate; BRENT: ICE Brent Crude futures contract front-month; GAS: S&P GSCI 3-Month Forward Natural Gas Index (ER); COAL: ICE Rotterdam coal nearby; EUA: EU Emissions Allowances (EEX EUA); STOCK: MSCI All-Country World Index. LB^2 is the Ljung-Box statistic for squared returns at lag 20; the 5% critical value is 31.41..

4 Empirical Results

This section evaluates the proposed Adaptive-Lasso MGARCH (AL-MGARCH) model. We first report the benchmark MGARCH(1,1) estimates, then show how the Adaptive-Lasso delivers a more parsimonious volatility-transmission structure, and finally assess whether this parsimony translates into better out-of-sample covariance forecasts.

4.1 MGARCH benchmark

We use an unrestricted CCC–MGARCH(1,1) as the benchmark because it is the standard baseline specification and provides a direct dense-versus-sparse comparison under identical variance dynamics.

Table 2 reports the unrestricted MGARCH(1,1) estimates[3]

The diagonal elements of **A** and **B** are sizeable across assets, confirming strong own-market volatility feedback and persistence. A number of off-diagonal coefficients are also statistically significant, indicating that volatility shocks can propagate across markets (bonds, energy commodities, carbon permits, and equities). The unrestricted MGARCH benchmark produces a dense volatility-transmission structure. In our eight-asset system,

[3]All estimations are conducted in MATLAB using the `fmincon` optimisation routine.

approximately 37% of the off-diagonal elements in the **A** and **B** matrices are statistically significant at the 5% level, implying widespread cross-market volatility interactions. While this richness may appear desirable from a structural perspective, it comes at a cost for forecasting. A dense parameterisation forces the model to estimate many small but persistent cross-effects, which are often weakly identified and sensitive to sampling variation. As a result, these nonzero coefficients may primarily reflect estimation noise rather than stable economic linkages, leading to overfitting and reduced out-of-sample performance.

Table 2 – MGARCH estimates of the coefficient matrices **A** and **B**

	GBOND	ESG	AGGB	BRENT	GAS	COAL	EUA	STOCK
				Matrix **A**				
GBOND	**0.1128**	0.0315	**0.0850**	-0.0053	0.0430	-0.0041	0.0136	**0.0152**
ESG	0.0207	**0.1128**	0.0184	0.0003	-0.0161	**-0.0059**	-0.0063	**0.0042**
AGGB	**0.0574**	-0.0005	**0.1598**	-0.0006	**0.0603**	-0.0029	**-0.0119**	0.0343
BRENT	-0.0021	**-0.0018**	0.0173	**0.1109**	0.0123	0.0113	0.0062	**0.0426**
GAS	**0.0261**	-0.0203	0.0078	-0.0004	**0.1298**	-0.0018	-0.0045	**0.0102**
COAL	**0.1015**	**-0.0238**	**-0.0686**	-0.0021	**0.1373**	**0.0774**	-0.0061	-0.0007
EUA	0.0202	0.0083	**-0.0199**	0.0077	0.0110	0.0085	**0.1161**	**0.0430**
STOCK	0.0200	-0.0063	**-0.0232**	0.0038	0.0169	**-0.0015**	0.0166	**0.1723**
				Matrix **B**				
GBOND	**0.5193**	**0.0837**	**0.0413**	**-0.0091**	**0.0885**	**0.0509**	0.0337	-0.0150
ESG	**-0.0110**	**0.8306**	-0.0356	-0.0211	**0.0597**	**0.0221**	-0.0022	0.0215
AGGB	0.0532	**0.0193**	**0.4764**	-0.0141	**0.0929**	**0.0437**	**-0.0306**	-0.0170
BRENT	**-0.0180**	-0.0098	0.0114	**0.7644**	**-0.0410**	0.0012	**-0.0185**	0.0491
GAS	-0.0166	0.0656	**-0.0186**	0.0104	**0.8156**	0.0002	0.0143	-0.0375
COAL	-0.0001	0.0234	0.0047	0.0203	**0.1282**	**0.5480**	0.0319	**-0.0311**
EUA	-0.0030	-0.0056	-0.0307	-0.0140	**0.0225**	**-0.0129**	**0.7412**	0.0155
STOCK	0.0018	**0.0574**	-0.0041	**0.0418**	-0.0098	-0.0024	**-0.0168**	**0.7133**
				Log-likelihood: 10.2094				
				BIC: 916.4390				

Notes: Bold entries indicate statistical significance at the 5% level. Log-likelihood and Bayesian Information Criterion (BIC) are reported at the bottom of the table. BIC uses active (nonzero) coefficients as effective degrees of freedom.

4.2 Adaptive-Lasso MGARCH: sparse transmission and improved fit

Our main contribution is to regularise the MGARCH variance dynamics using the Adaptive Lasso, which selects a sparse set of economically relevant volatility spillover

channels while shrinking weak cross-effects exactly to zero. The shrinkage parameter λ_T is selected using a rolling validation scheme: the estimation window is split into a training subsample used for tuning λ_T and a validation subsample used to minimise one-step-ahead forecast loss. Specifically, the first 1000 observations are used as the training subsample to construct the adaptive weights and estimate the model for a given λ_T, while the last 509 observations form the validation subsample used to select $\hat{\lambda}_T$ by minimising one-step-ahead QLIK. The penalised model is then re-estimated on the full sample at the optimal $\hat{\lambda}_T$.

Table 3 – Adaptive-Lasso estimates of the coefficient matrices **A** and **B**

	GBOND	ESG	AGGB	BRENT	GAS	COAL	EUA	STOCK
				Matrix **A**				
GBOND	0.1518	0.0104	0.0858	0	0.0123	-0.0028	-0.0011	0.0045
ESG	0.0012	0.1677	0.0041	0	0.0019	0	0	0.0015
AGGB	0.0181	0	0.2377	0	-0.0011	0	-0.0031	0.0125
BRENT	0.0013	0	0	0.1691	0	0	0	-0.0011
GAS	0.0069	-0.0029	0	0	0.2326	0.0019	0	0.0014
COAL	0.0508	-0.0130	-0.0672	0.0017	0.1900	0.1293	0	0
EUA	0.0026	0.0026	0.0059	-0.0045	0.0011	0.0025	0.1623	0
STOCK	-0.0056	0	0	-0.0111	-0.0022	-0.0012	0.0036	0.2370
				Matrix **B**				
GBOND	0.5964	0.0145	0.0562	0.0072	0.0256	0.0401	-0.0042	0.0081
ESG	0	0.7897	-0.0020	0.0030	0.0408	0.0020	0	-0.0056
AGGB	0.0135	-0.0042	0.5944	0.0022	0.0641	-0.0117	-0.0099	-0.0023
BRENT	0	-0.0023	-0.0014	0.7497	-0.0046	0	0.0178	0.0289
GAS	0	-0.0104	0	0.0040	0.6914	0	-0.0064	0
COAL	0	0	0	0	0.1430	0.4911	-0.0015	0
EUA	0	0	-0.0047	0.0032	-0.0021	-0.0041	0.6680	0
STOCK	0.0016	0.0034	0	0.0080	0.0016	0	-0.0015	0.7290
				Log-likelihood: 11.2304				
				BIC: 636.2674				

Note: Elements set exactly to zero by the penalty appear as "0". Standard errors (or t-ratios) are available upon request. BIC uses active (nonzero) coefficients as effective degrees of freedom.

Table 3 reports the resulting Adaptive-Lasso estimates. Relative to the dense MGARCH benchmark, the coefficient matrices become markedly sparse, particularly in off-diagonal elements. In the short-run shock matrix **A**, the largest surviving coefficients are concentrated on the diagonal, while only a limited number of cross-market effects remain,

indicating that immediate volatility transmission is confined to a small subset of markets. By contrast, in the persistence matrix $\mathbf{B}$, diagonal elements remain large, and several cross-market channels survive, suggesting that volatility spillovers are more pronounced at longer horizons than at the daily impact horizon.

This sparsity yields a parsimonious and interpretable volatility-transmission structure and substantially reduces effective model complexity. Consistent with this interpretation, the AL-MGARCH attains a higher log-likelihood and a substantially lower BIC than the unpenalised MGARCH. Here BIC is computed using the number of active (nonzero) penalised coefficients as effective degrees of freedom, so it should be interpreted as a fit–parsimony trade-off under regularisation. Together with its superior out-of-sample forecasting performance, these results indicate that the reduction in estimation variance achieved through adaptive regularisation dominates any bias introduced by shrinkage, leading to improved generalisation.

4.3 Out-of-sample forecasting comparison

The central empirical question is whether the proposed AL-MGARCH improves predictive performance relative to the unrestricted MGARCH benchmark. We therefore compare out-of-sample forecasts of the conditional covariance matrix at horizons $h \in \{1, 5, 22\}$ trading days (daily, weekly, and monthly horizons). Forecast accuracy is evaluated using two standard statistical loss functions for covariance forecasts: QLIK and FN (Frobenius norm). Realised covariance is proxied by $\mathbf{RC}_t = \mathbf{r}_t\mathbf{r}_t'$. See Bauwens and Xu (2023) and Bauwens and Xu (2025) for the details of the two loss functions.

Forecasts are generated in a rolling-window design: each model is re-estimated every fifth observation using a rolling window of 1,000 observations, and h-step-ahead covariance forecasts are produced recursively for $h > 1$. To test whether forecast improvements are statistically meaningful, we apply the White (2000) reality check based on the stationary bootstrap of Politis and Romano (1994).

Table 4 reports average out-of-sample losses. Across all horizons and both loss functions, AL-MGARCH yields lower losses than MGARCH. The gains are present at the one-step horizon and persist at longer horizons, which is consistent with the motivation for regularisation: by eliminating weak and unstable spillover coefficients, AL-MGARCH improves generalisation and delivers more accurate covariance forecasts. In sum, the results provide empirical support for the proposed AL-MGARCH framework as a practical alternative to dense MGARCH specifications when the cross-sectional dimension is moderate and forecasting performance is the primary objective.

Table 4 – Out-of-sample forecasting performance: Statistical loss functions

Horizon	QLIK		FN	
	AL-MGARCH	MGARCH	AL-MGARCH	MGARCH
$h = 1$	**4.4026**	4.5574	**7.9173**	7.9591
$h = 5$	**4.9232**	5.1342	**7.9569**	7.9954
$h = 22$	**5.9333**	6.1867	**8.0863**	8.1516

Note: Bold indicates significance at the 5% level based on the White (2000) reality check.

The economic loss functions are relevant for evaluating covariance matrix forecasts, as they are based on forecasted portfolio performance. The same economic loss functions as in Engle and Kelly (2012) are used, namely the global minimum variance portfolio (GMV) and the minimum variance portfolio (MV); see also Bauwens and Xu (2023) and Bauwens and Xu (2025). These loss functions are the variances of the forecasts of portfolio returns. Given a covariance matrix forecast $H^a_{t+s|t}$, the GMV portfolio weight vector w^a_{t+s} is computed as the minimizer of the portfolio variance $(w^a_{t+s})' H^a_{t+s|t} w^a_{t+s}$ subject to the constraint that the weights add to unity. Next the portfolio return $w^a_{t+s} r_{t+s}$ is computed. Once this procedure is carried out for each forecast date, the variance of the resulting time series of optimal portfolio returns is used as the loss function in the MCS procedure, since a superior model produces optimal portfolios with lower forecast variance.

The MV portfolio is obtained by minimising the portfolio variance subject to the additional constraint that the expected portfolio return be larger than a chosen value.

Following Engle and Kelly (2012), this value is fixed at 10% and the expected portfolio return at the mean of the data.

Table 5 reports the out-of-sample economic loss functions based on portfolio variance for the global minimum variance (GMV) and mean-variance (MV) portfolios. In both cases, the loss corresponds to the realised variance of the optimal portfolio returns constructed using the forecasted covariance matrix. Lower values therefore indicate superior economic performance, as they imply more stable portfolios and improved risk control.

Across all forecast horizons ($h = 1, 5, 22$), the AL-MGARCH model consistently delivers lower portfolio variance than the benchmark MGARCH model for both portfolio strategies. For the GMV portfolio, the variance reductions are economically meaningful and statistically significant at the 5% level according to the White (2000) reality check. For instance, at the one-day horizon, the GMV variance decreases from 1.382 under MGARCH to 1.261 under AL-MGARCH, corresponding to an approximate 9% reduction in realised portfolio risk. Similar improvements persist at the weekly ($h = 5$) and monthly ($h = 22$) horizons, indicating that the gains from sparse covariance estimation remain robust over longer investment horizons.

The economic gains are even more pronounced for the constrained MV portfolio, which incorporates a minimum expected return requirement. At all horizons, AL-MGARCH achieves substantially lower portfolio variance than MGARCH, with statistically significant improvements. This suggests that adaptive regularisation improves not only statistical forecast accuracy but also portfolio risk performance under constraints—an implication that is broadly relevant for global institutional investors and risk managers, with particular salience for those exposed to Europe's carbon and energy markets (e.g., EUAs).

Overall, these results confirm that the forecasting gains of AL-MGARCH are economically meaningful. By eliminating noisy and weak spillover channels, the adaptive penalisation delivers more stable covariance forecasts, leading to improved portfolio al-

location and superior out-of-sample risk performance. The consistency of the results across portfolio objectives and forecast horizons strengthens the practical relevance of the proposed AL-MGARCH framework.

Table 5 – Out-of-sample forecasting performance: Economic loss functions

Horizon	GMV			MV		
	AL-MGARCH	MGARCH	Impr. (%)	AL-MGARCH	MGARCH	Impr. (%)
$h = 1$	**1.261**	1.382	8.76	**2.241**	2.452	8.61
$h = 5$	**1.334**	1.456	8.38	**2.352**	2.567	8.38
$h = 22$	**1.402**	1.524	8.01	**2.437**	2.583	5.65

Note: Impr. (%) is the percentage reduction in loss relative to MGARCH, (MGARCH − AL-MGARCH)/MGARCH × 100. Bold indicates significance at the 5% level based on the White (2000) reality check.

4.4 Robustness

A natural extension would be to assess robustness through dedicated subsample analysis and formal structural-stability tests. In this paper, however, the empirical design already relies on an out-of-sample rolling-window framework with a 1,000-observation estimation window, with parameters re-estimated repeatedly over time. As such, forecasting performance is evaluated across many overlapping subsamples and changing market conditions, providing an informal assessment of stability.

The results show that AL-MGARCH consistently outperforms the unpenalised MGARCH benchmark across forecast horizons and evaluation criteria. This suggests that the improvement is not driven by a particular sample period. Since the main objective of the paper is to compare the out-of-sample forecasting performance of AL-MGARCH and MGARCH, we consider the rolling-window evidence sufficient for the present analysis. Nevertheless, explicit subsample splits and formal break or stability tests would be useful extensions for future research.

5 Conclusion

This paper proposes and evaluates an Adaptive LASSO-MGARCH framework for modelling and forecasting high-dimensional conditional covariance dynamics. By introducing coefficient-specific penalisation directly into the multivariate GARCH variance equations, the model delivers a sparse yet flexible representation of volatility spillovers while mitigating over-parameterisation and estimation noise.

An empirical application to a system of bond, equity, energy, and carbon markets shows that adaptive regularisation substantially simplifies the volatility transmission structure without sacrificing in-sample fit. More importantly, out-of-sample forecasting results demonstrate that AL-MGARCH consistently outperforms an unpenalised MGARCH benchmark across short-, medium-, and long-horizon covariance forecasts. These gains are statistically significant under multiple loss functions and remain robust to rolling-window re-estimation.

The AL-MGARCH framework therefore offers a practical and scalable tool for high-dimensional risk forecasting. Improved covariance forecasts translate into more stable risk estimates, hedge ratios, and portfolio weights. This is relevant for institutional investors (e.g., constrained minimum-variance allocation and risk management) and also for individual investors using risk-based allocation rules. The main practical benefit of AL-MGARCH is that it reduces estimation noise by shrinking weak spillover channels, leading to more robust covariance forecasts.

Our framework adopts a CCC decomposition, so correlation dynamics are not modelled explicitly; extending the approach to time-varying correlations (e.g., DCC-type models) is a natural next step. Future work could also incorporate asymmetric volatility effects and heavier-tailed innovations, and consider alternative realised covariance measures when higher-frequency data are available.

References

Audrino, F., Knaus, S.D., 2016. Lassoing the har model: A model selection perspective on realized volatility dynamics. Econometric Reviews 35, 1485–1521. doi:10.1080/07474938.2015.1092801.

Bauwens, L., Laurent, S., Rombouts, J.V., 2006. Multivariate garch models: a survey. Journal of applied econometrics 21, 79–109.

Bauwens, L., Xu, Y., 2023. Dcc- and deco-heavy: Multivariate garch models based on realized variances and correlations. International Journal of Forecasting 39, 938–955.

Bauwens, L., Xu, Y., 2025. The contribution of realized variance–covariance models to the economic value of volatility timing. International Journal of Forecasting 41, 1165–1183.

Cappiello, L., Engle, R.F., Sheppard, K., 2006. Asymmetric dynamics in the correlations of global equity and bond returns. Journal of Financial econometrics 4, 537–572.

Cattivelli, L., Gallo, G.M., 2019. Adaptive lasso for vector multiplicative error models. Quantitative Finance , 1–20.

Chuliá, H., Furió, D., Uribe, J.M., 2019. Volatility spillovers in energy markets. The Energy Journal 40, 173–198.

Dhaene, G., Sercu, P., Wu, J., 2022a. Volatility spillovers: A sparse multivariate garch approach with an application to commodity markets. Journal of Futures Markets 42, 868–887.

Dhaene, G., Sercu, P., Wu, J., 2022b. Volatility spillovers: A sparse multivariate GARCH approach with an application to commodity markets. Journal of Futures Markets 42, 868–887. doi:10.1002/fut.22312.

Engle, R., 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. Journal of business & economic statistics 20, 339–350.

Engle, R.F., Kelly, B., 2012. Dynamic equicorrelation. Journal of Business & Economic Statistics 30, 212–228. doi:10.1080/07350015.2011.652048.

Engle, R.F., Kroner, K.F., 1995. Multivariate simultaneous generalized arch. Econometric theory 11, 122–150.

Engle, R.F., Ng, V.K., Rothschild, M., 1990. Asset pricing with a factor-arch covariance structure: Empirical estimates for treasury bills. Journal of econometrics 45, 213–237.

Hafner, C.M., Franses, P.H., 2009. A generalized dynamic conditional correlation model: simulation and application to many assets. Econometric Reviews 28, 612–631.

Hamao, Y., Masulis, R.W., Ng, V., 1990. Correlations in price changes and volatility across international stock markets. The review of financial studies 3, 281–307.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed., Springer.

Huang, J., Ma, S., Zhang, C.H., 2008. Adaptive lasso for sparse high-dimensional regression models. statistica sinica , 1603–1618.

Laurent, S., Rombouts, J.V.K., Violante, F., 2012. On the forecasting accuracy of multivariate GARCH models. Journal of Applied Econometrics 27, 934–955.

Laurent, S., Rombouts, J.V.K., Violante, F., 2013. On loss functions and ranking forecasting performances of multivariate volatility models. Journal of Econometrics 173, 1–10. doi:10.1016/j.jeconom.2012.08.004.

Medeiros, M.C., Mendes, E.F., 2016. 1-regularization of high-dimensional time-series models with non-gaussian and heteroskedastic errors. Journal of Econometrics 191, 255–271.

Mensi, W., Selmi, R., Al-Kharusi, S., Belghouthi, H.E., Kang, S.H., 2024. Connectedness between green bonds, conventional bonds, oil, heating oil, natural gas, and petrol: new evidence during bear and bull market scenarios. Resources Policy 91, 104888.

Poignard, B., Fermanian, J.D., 2021. High-dimensional penalized arch processes. Econometric Reviews 40, 86–107.

Politis, D.N., Romano, J.P., 1994. The stationary bootstrap. Journal of the American Statistical Association 89, 1303–1313. doi:10.1080/01621459.1994.10476870.

Reboredo, J.C., Ugolini, A., Aiube, F.A.L., 2020. Network connectedness of green bonds and asset classes. Energy Economics 86, 104629.

Sun, Y., Lin, X., 2012. Regularization for stationary multivariate time series. Quantitative Finance 12, 573–586.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B: Statistical Methodology 58, 267–288.

Tse, Y.K., Tsui, A.K.C., 2002. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. Journal of Business & Economic Statistics 20, 351–362.

White, H., 2000. A reality check for data snooping. Econometrica 68, 1097–1126. doi:10.1111/1468-0262.00152.

Wu, J., Dhaene, G., 2016a. Sparse multivariate GARCH. Working Paper. KU Leuven, Department of Economics.

Wu, J., Dhaene, G., 2016b. Sparse multivariate garch. Discussion paper series, DPS16. 11 , 1–17.

Xu, Y., Lyu, J., Mazouz, K., 2025. Adaptive-lasso mgarch for the volatility spillover of transition finance. URL: https://ssrn.com/abstract=5660580. available at SSRN 5660580.

Yao, S., Zou, H., Xing, H., 2024. L1 regularization for high-dimensional multivariate garch models. Risks 12, 34.

Zou, H., 2006. The adaptive lasso and its oracle properties. Journal of the American statistical association 101, 1418–1429.

Appendix A: Comparison with Xu et al. (2025)

Table 6 highlights the key differences between Xu et al. (2025) and this paper. In brief, Xu et al. (2025) focus on modelling and in-sample spillover interpretation, whereas this paper serves as a forecasting-evaluation companion and adds theoretical and forecasting details.

Appendix B: Asymptotic properties of the Adaptive-LASSO estimator

This appendix states a sufficient-condition result for the Adaptive-LASSO estimator in the CCC–MGARCH setting. The statement follows the logic of adaptive LASSO

Table 6 – Key differences between Xu et al. (2025) and this paper

	Xu et al. (2025)	This paper
Main focus	Model development and in-sample spillover interpretation	Out-of-sample covariance forecasting evaluation
Added contributions	Not the central focus	Asymptotic properties of the AL-MGARCH estimator; explicit multi-step-ahead forecasting formulas
Forecasting design	Not the central focus	Rolling-window forecasts; multi-horizon evaluation ($h \in \{1, 5, 22\}$)
Evaluation and tests	In-sample emphasis	Statistical and economic loss functions; White (2000) reality check

oracle results for likelihood-based multivariate volatility dynamics (e.g., Theorem 1 in Cattivelli and Gallo, 2019) and standard QMLE regularity conditions for GARCH-type models.

Setup

Let $\theta \in \Theta \subset \mathbb{R}^p$ collect the parameters of the variance dynamics (e.g., the elements of ω, A, and B). Define the (average) Gaussian QML criterion

$$\ell_T(\theta) = -\frac{1}{T}\sum_{t=1}^{T} \log f(\varepsilon_t \mid \theta),$$

where $f(\cdot \mid \theta)$ is the Gaussian density implied by the CCC covariance $\mathbf{H}_t(\theta) = \mathbf{D}_t(\theta)\mathbf{P}\mathbf{D}_t(\theta)$. Let the Adaptive-LASSO estimator be

$$\widehat{\theta}(\lambda_T) = \arg\min_{\theta\in\Theta} \left\{ \ell_T(\theta) + \lambda_T \sum_{j=1}^{p} w_j |\theta_j| \right\},$$

with weights $w_j = 1/\max\{|\tilde{\theta}_j|, \epsilon_w\}^\gamma$, where $\tilde{\theta}$ is an initial (unpenalised) QMLE, $\epsilon_w > 0$, and $\gamma > 0$.

Let the true parameter be θ_0 and define the active set $S = \{j : \theta_{0j} \neq 0\}$ with $s = |S|$ and complement S^c.

Assumptions

Assumption A1 (stationarity and ergodicity). Under θ_0, the CCC–MGARCH recursion admits a unique strictly stationary and ergodic solution with $h_{it}(\theta_0) > 0$ a.s., and $\mathbb{E}\|\varepsilon_t\|^{2+\delta} < \infty$ for some $\delta > 0$.

Assumption A2 (smoothness and identification). The map $\theta \mapsto \ell_T(\theta)$ is three-times continuously differentiable in a neighbourhood of θ_0 a.s.; θ_0 is identifiable; and the Fisher information matrix

$$\mathcal{I}(\theta_0) = \lim_{T\to\infty} \mathbb{E}\big[\nabla^2 \ell_T(\theta)\big]\,\big|_{\theta=\theta_0}$$

exists and is positive definite on the active coordinates S.

Assumption A3 (initial estimator). The unpenalised QMLE $\tilde{\theta}$ satisfies $\sqrt{T}(\tilde{\theta} - \theta_0) = O_p(1)$.

Assumption A4 (penalty rate). The tuning sequence satisfies $\lambda_T\sqrt{T} \to 0$ and $\lambda_T T^{(\gamma+1)/2} \to \infty$ as $T \to \infty$.

Oracle-type result

Theorem 5.1 (Selection consistency and asymptotic normality)**.** *Under Assumptions A1–A4, the Adaptive-LASSO estimator $\widehat{\theta}(\lambda_T)$ satisfies:*

1. ***Selection consistency:*** $\Pr\Big(\widehat{S}(\lambda_T) = S\Big) \to 1$ *as* $T \to \infty$, *where* $\widehat{S}(\lambda_T) = \{j : \widehat{\theta}_j(\lambda_T) \neq 0\}$.
2. ***Asymptotic normality on the active set:***

$$\sqrt{T}\left(\widehat{\theta}_S(\lambda_T) - \theta_{0S}\right) \xrightarrow{d} \mathcal{N}\big(0,\ \mathcal{I}_{SS}(\theta_0)^{-1}\big)\,,$$

   where $\mathcal{I}_{SS}(\theta_0)$ denotes the active-set block of the information matrix.

Hence, on the active coordinates, $\widehat{\theta}_S(\lambda_T)$ is asymptotically equivalent to the QMLE computed under the oracle model that knows S in advance.

Proof sketch. The argument follows standard adaptive-LASSO proofs for likelihood-based M-estimators: (i) a local quadratic expansion of $\ell_T(\theta)$ around θ_0 using Assumption A2 and a CLT under Assumption A1; (ii) $\sqrt{T}$-consistency of $\tilde{\theta}$ (Assumption A3) implies the adaptive weights diverge on S^c and remain bounded on S; (iii) the tuning rate in Assumption A4 guarantees that the penalty is asymptotically negligible on S but dominates on S^c, yielding selection consistency; (iv) conditional on correct selection, the penalised estimator coincides asymptotically with the unpenalised QMLE on the active set, implying the stated limit distribution. □