

Statistica Sinica Preprint No: SS-2025-0271

Title	Asymmetric Garch Modelling Without Moment Conditions
Manuscript ID	SS-2025-0271
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202025.0271
Complete List of Authors	Yuxin Tao, Huan Gong and Dong Li
Corresponding Authors	Huan Gong
E-mails	gonghuan@nudt.edu.cn
Notice: Accepted author version.	

ASYMMETRIC GARCH MODELLING WITHOUT MOMENT CONDITIONS

Yuxin Tao, Huan Gong* , Dong Li

*Southern University of Science and Technology, National University of
Defense Technology, Tsinghua University*

Abstract: Heavy tails and stability are two persistent challenges in modelling financial time series, yet most existing approaches rely on finite-moment assumptions and pay insufficient attention to stability issues. To bridge this gap, we propose an asymmetric GARCH model with standardized non-Gaussian stable innovations (sAGARCH), which accommodates infinite variance and even infinite mean. We establish a comprehensive inference framework for both stationary and explosive cases, proving the strong consistency and asymptotic normality of the maximum likelihood estimator, including the tail index parameter. We also discuss multiple estimators for the asymptotic variance. Additionally, we propose a modified Kolmogorov-type test statistic for diagnostic checking, along with tests for strict stationarity and asymmetry. Through Monte Carlo simulations with heavy-tailed innovations, we provide further insight into the finite-sample performance of the intercept estimator. Empirical applications to stock returns further highlight the usefulness and merits of the proposed sAGARCH model.

Key words and phrases: Heavy tails, Kolmogorov-type test, Maximum likelihood estimation, Nonstationarity, Stable distribution.

*Corresponding author. *E-mail:* gonghuan@nudt.edu.cn

1. Introduction

Heavy-tailed phenomena are ubiquitous across scientific fields (Harvey, 2013; Nolan, 2020), and have received sustained attention. Financial time series are a typical example, as they often exhibit heavy tails, volatility clustering, and asymmetry. Since the seminal work of Engle (1982) and Bollerslev (1986), the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model has become a benchmark for econometricians and financial practitioners (Fan and Yao, 2017), with numerous variants proposed for practical needs (Francq and Zakoïan, 2019).

Although GARCH-type models are designed to capture several stylized features, they often struggle with tail thickness, and excess kurtosis may remain in the standardized residuals of fitted models (Bai, 2003). Thus, the commonly imposed finite fourth- or second-moment assumptions on the innovations can be too restrictive. Student's t and generalized Gaussian innovations are useful alternatives, but they lack stability under addition; see, e.g., Fama (1965), Samuelson (1967), Nolan (2020), and discussions in Li, et.al. (2023).

To address this issue, we propose the use of stable distributions for innovations in volatility models. Their application in finance dates back to Mandelbrot (1963) and Fama (1965), who advocated stable laws for

modelling unconditional asset returns, although early use was limited by the lack of closed-form densities. With modern numerical methods, stable models have become computationally feasible. A common concern is that non-normal stable distributions may have infinite mean or variance, seemingly contradicting the finite sample moments observed in practice. However, as noted by Mandelbrot (1997), sample moments need not converge asymptotically, so empirical finiteness does not imply finite theoretical moments. Indeed, power-law models with divergent moments (tail indices below 2 or even 1) are well-documented in finance and physics (Uchaikin and Zolotarev, 1999), and infinite-mean models may provide useful statistical fits and financial risk insights (Chen and Wang, 2025).

The use of stable innovations in GARCH models was first proposed by McCulloch (1985) and later extended by Liu and Brorsen (1995), mainly from an applied perspective. Stationarity conditions were studied by Panorska et al. (1995) and Mittnik et al. (2002). For estimation, Liu and Brorsen (1995) proposed the maximum likelihood estimation (MLE) approach, while Calzolari et al. (2014) developed an indirect inference method using Student's t innovations as auxiliary models. On the other hand, Hall and Yao (2003) analyzed heavy-tailed GARCH models with innovations having infinite fourth moments but finite variance. Zhang et al. (2022) explored the

least absolute deviation estimator for autoregressive models with heavy-tailed G-GARCH noise. However, to the best of our knowledge, the asymptotic properties of MLE and statistical inference for GARCH-type models with infinite-moment innovations remain unexplored in the literature.

Conditional asymmetry, or leverage effects, is another important feature of financial time series. Recent works on asymmetric GARCH-type models include Li and Zhu (2020), Wang et al. (2022), and Zhu et al. (2023). For stable innovations, asymmetry could be introduced through a nonzero skewness parameter β . However, skew stable densities are computationally more complicated and may cause numerical instability. Moreover, as the stability exponent α approaches 2, stable distributions tend to symmetry, making β meaningless and harder to estimate (especially for $\alpha = 2$, which is the Gaussian case); see Nolan (2020). We therefore model asymmetry through the volatility structure rather than through skewed stable innovations.

Our major contributions are threefold. First, we propose an asymmetric GARCH model with stable innovations, which accommodates excess kurtosis, volatility clustering, and leverage effects without imposing moment conditions. We establish the strong consistency and asymptotic normality of the MLE under a unified framework covering both stationary and explosive cases, with $n^{-1/2}$ convergence rate except for the intercept in the

explosive case. It differs from the findings of Hall and Yao (2003). We also provide two estimators of the Fisher information matrix and a universal estimator of the asymptotic variance that does not require stationarity.

Second, we develop several technical tools for handling infinite-moment stable innovations in GARCH-type models. These include refined asymptotic bounds for derivatives of log-stable densities in Appendix S3, an identifiability argument for the stability parameter α in Lemma 1, and an improved exponential convergence rate for the volatility recursion in Lemma 4. These results are essential for removing moment restrictions.

Third, we propose inference procedures for diagnostic checking, strict stationarity, and asymmetry testing. In particular, we develop a modified Kolmogorov-type diagnostic statistic using the transformation method of Bai (2003). Monte Carlo studies verify the finite-sample performance of the MLE and the proposed tests, and empirical applications to stock returns illustrate the practical advantages of the proposed model.

The remainder of the paper is organized as follows. Section 2 proposes the sAGARCH model and derives its stationarity conditions. Section 3 considers the asymptotic properties of the MLE, a universal estimator of its asymptotic variance, and computational issues. Section 4 develops tests for diagnostic checking, strict stationarity, and asymmetry. Section 5 reports

the finite-sample performance of MLE and test statistics. Section 6 analyzes individual stock returns. Section 7 concludes the paper. All technical proofs are postponed to the Supplementary Material.

2. The sAGARCH model: Stability and Explosivity

2.1 Stable distribution

Stable distributions are defined by their characteristic functions, since their densities have no closed forms except for three special cases (i.e., Gaussian, Cauchy, and Lévy distributions). A general stable distribution $\mathbf{S}(\alpha, \beta, \gamma, \delta)$ contains four parameters: α, β, γ and δ , representing stability, skewness, scale, and location, respectively. In this paper, we consider the standardized symmetric stable distribution $\mathbf{S}(\alpha, 0, 1, 0)$, of which the characteristic function is $\phi(s) = \exp(-|s|^\alpha)$, $s \in \mathbb{R}$, where $\alpha \in (0, 2]$. Its density is

$$f_\alpha(x) = \frac{1}{\pi} \int_0^\infty \exp(-s^\alpha) \cos(sx) ds, \quad x \in \mathbb{R}. \quad (2.1)$$

It has no variance when $\alpha < 2$ and no expectation when $\alpha \leq 1$. Particularly, when $\alpha = 1$, it is the standard Cauchy distribution, and when $\alpha = 2$, it reduces to $\mathcal{N}(0, 2)$. More useful properties on stable distribution can be found in Propositions 1-3 in the Appendix S3.

2.2 The sAGARCH model

To better capture excess kurtosis, asymmetry and volatility clustering jointly, we define the first-order asymmetric GARCH model with standardized non-Gaussian symmetric stable innovation (hereafter sAGARCH(1,1)) as

$$\begin{cases} y_t = \sigma_t \eta_t, & t \in \mathbb{Z}_+ := \{1, 2, \dots\}, \\ \sigma_t^2 = \omega + \phi_+ (y_{t-1}^+)^2 + \phi_- (y_{t-1}^-)^2 + \psi \sigma_{t-1}^2, \end{cases} \quad (2.2)$$

with initial values y_0 and $\sigma_0 \geq 0$, where $\omega > 0$, $\phi_+ > 0$, $\phi_- > 0$, $\psi > 0$, $x^+ = \max\{x, 0\}$, $x^- = -\min\{x, 0\}$. $\{\eta_t\}$ is a sequence of i.i.d. standardized non-Gaussian symmetric stable random variables with the stable exponent $\alpha \in (0, 2)$. Heavy-tailedness is modeled through stable innovation $\{\eta_t\}$, and asymmetry is captured by different parameters ϕ_+ and ϕ_- .

By Bougerol and Picard (1992a,b), model (2.2) is strictly stationary *if and only if* the top Lyapunov exponent

$$\gamma_\alpha := E \log a(\eta_1) < 0, \quad \text{where } a(x) = \phi_+(x^+)^2 + \phi_-(x^-)^2 + \psi. \quad (2.3)$$

The plots of the stable densities and the strict stationarity regions under different scenarios are provided in the Appendix S1.1.

3. Maximum Likelihood Estimation

Let $\vartheta = (\phi_+, \phi_-, \psi, \alpha)'$, and $\theta = (\omega, \vartheta)' = (\omega, \phi_+, \phi_-, \psi, \alpha)' \in \mathbb{R}_+^4 \times (0, 2)$, where $\mathbb{R}_+ = (0, \infty)$. Suppose that the observations $\{y_0, y_1, \dots, y_n\}$ are from model (2.2) with true parameter $\theta_0 = (\omega_0, \vartheta_0)' = (\omega_0, \phi_{0+}, \phi_{0-}, \psi_0, \alpha_0)'$. Then the (conditional) log-likelihood function is defined as

$$\tilde{L}_n(\theta) = \sum_{t=1}^n \tilde{\ell}_t(\theta), \quad \tilde{\ell}_t(\theta) = -\log \tilde{\sigma}_t(\theta) + \log f_\alpha\left(\frac{y_t}{\tilde{\sigma}_t(\theta)}\right), \quad (3.4)$$

where $f_\alpha(x)$ is defined in (2.1), and $\tilde{\sigma}_t^2(\theta) = \omega + \phi_+(y_{t-1}^+)^2 + \phi_-(y_{t-1}^-)^2 + \psi \tilde{\sigma}_{t-1}^2(\theta)$, $t \geq 1$, with initial values $(y_0, \tilde{\sigma}_0(\theta)) \equiv (0, 0)$. The MLE of θ_0 is

$$\hat{\theta}_n := (\hat{\omega}_n, \hat{\vartheta}_n) := (\hat{\omega}_n, \hat{\phi}_{n+}, \hat{\phi}_{n-}, \hat{\psi}_n, \hat{\alpha}_n)' = \arg \max_{\theta \in \Theta} \tilde{L}_n(\theta),$$

where the parameter space Θ is a subset of $\mathbb{R}_+^4 \times (0, 2)$.

3.1 Consistency and asymptotic normality

To study the asymptotics of $\hat{\theta}_n$, we require the following assumptions.

Assumption 1. $\{\eta_t\}$ is a sequence of i.i.d. standardized non-Gaussian symmetric stable random variables with the density function $f_{\alpha_0}(x)$.

Assumption 2. The parameter space Θ is compact and $\theta_0 \in \Theta$.

Assumption 3. The true parameter θ_0 is an interior point of Θ .

We are now ready to present our main results. Theorem 1 establishes the results for the stationary case, and Theorem 2 addresses the explosive case. To study the asymptotic properties of $\hat{\theta}_n$ for the stationary case, we define the theoretical (conditional) log-likelihood function as $L_n(\theta) = \sum_{t=1}^n \ell_t(\theta)$ with $\ell_t(\theta) = -\log \sigma_t(\theta) + \log f_\alpha(y_t/\sigma_t(\theta))$, $\sigma_t^2(\theta) = \omega + \phi_+(y_{t-1}^+)^2 + \phi_-(y_{t-1}^-)^2 + \psi\sigma_{t-1}^2(\theta)$, $t \in \mathbb{Z}$. Then when $0 < \psi < 1$,

$$\sigma_t^2(\theta) = \sum_{j=0}^{\infty} \psi^j \{\omega + \phi_+(y_{t-1-j}^+)^2 + \phi_-(y_{t-1-j}^-)^2\}, \quad t \in \mathbb{Z}. \quad (3.5)$$

Since $\{y_t : t \in \mathbb{Z}\}$ is strictly stationary and ergodic, so is $\{\sigma_t^2(\theta) : t \in \mathbb{Z}\}$ by the measurability of $\sigma_t^2(\theta)$ with respect to $\{y_t\}$.

Theorem 1. Suppose that Assumptions 1–2 hold and $\gamma_{\alpha_0} < 0$.

- (i). For Θ such that $\forall \theta \in \Theta$, $\psi < 1$, then $\hat{\theta}_n \xrightarrow{a.s.} \theta_0$, as $n \rightarrow \infty$.
- (ii). Further, if Assumption 3 holds, then $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, \Sigma^{-1})$, where \xrightarrow{d} stands for convergence in distribution and

$$\Sigma = E \left\{ \frac{\partial \ell_t(\theta_0)}{\partial \theta} \frac{\partial \ell_t(\theta_0)}{\partial \theta'} \right\} = \begin{pmatrix} \Sigma_{\hat{\theta}\hat{\theta}'} & \Sigma'_{\hat{\theta}\alpha} \\ \Sigma_{\hat{\theta}\alpha} & \Sigma_{\alpha\alpha} \end{pmatrix}$$

with $\tilde{\theta} = (\omega, \phi_+, \phi_-, \psi)'$ and

$$\begin{aligned}\Sigma_{\tilde{\theta}\tilde{\theta}'} &= \frac{1}{4}E\left\{\frac{1}{\sigma_t^4(\theta_0)}\frac{\partial\sigma_t^2(\theta_0)}{\partial\tilde{\theta}}\frac{\partial\sigma_t^2(\theta_0)}{\partial\tilde{\theta}'}\right\}E\left[\left\{1+\frac{\partial\log f_{\alpha_0}(\eta_t)}{\partial x}\eta_t\right\}^2\right], \\ \Sigma_{\tilde{\theta}\alpha} &= -\frac{1}{2}E\left\{\frac{1}{\sigma_t^2(\theta_0)}\frac{\partial\sigma_t^2(\theta_0)}{\partial\tilde{\theta}}\right\}E\left\{\frac{\partial\log f_{\alpha_0}(\eta_t)}{\partial x}\frac{\partial\log f_{\alpha_0}(\eta_t)}{\partial\alpha}\eta_t\right\}, \\ \Sigma_{\alpha\alpha} &= E\left[\left\{\frac{\partial\log f_{\alpha_0}(\eta_t)}{\partial\alpha}\right\}^2\right].\end{aligned}$$

Theorem 2. *Suppose that Assumptions 1–2 hold and $\gamma_{\alpha_0} > 0$, then*

(i). $\hat{\vartheta}_n \xrightarrow{a.s.} \vartheta_0$, as $n \rightarrow \infty$.

(ii). *Further, if Assumption 3 holds, then $\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \xrightarrow{d} \mathcal{N}(0, \Upsilon^{-1})$, where*

$$\Upsilon = \begin{pmatrix} \Upsilon_{\tilde{\vartheta}\tilde{\vartheta}'} & \Upsilon_{\tilde{\vartheta}\alpha}' \\ \Upsilon_{\tilde{\vartheta}\alpha} & \Upsilon_{\alpha\alpha} \end{pmatrix} \text{ with } \tilde{\vartheta} = (\phi_+, \phi_-, \psi)'$$

$$\Upsilon_{\tilde{\vartheta}\tilde{\vartheta}'} = \frac{1}{4}E(d_t d_t')E\left[\left\{1+\frac{\partial\log f_{\alpha_0}(\eta_t)}{\partial x}\eta_t\right\}^2\right],$$

$$\Upsilon_{\tilde{\vartheta}\alpha} = -\frac{1}{2}E(d_t)E\left\{\frac{\partial\log f_{\alpha_0}(\eta_t)}{\partial x}\frac{\partial\log f_{\alpha_0}(\eta_t)}{\partial\alpha}\eta_t\right\}, \quad \Upsilon_{\alpha\alpha} = E\left[\left\{\frac{\partial\log f_{\alpha_0}(\eta_t)}{\partial\alpha}\right\}^2\right],$$

$$E(d_t) = E(d_t^{\phi_+}, d_t^{\phi_-}, d_t^\psi)' = \left(\frac{1-\nu_{1+}}{\phi_{0+}(1-\nu_1)}, \frac{1-\nu_{1-}}{\phi_{0-}(1-\nu_1)}, \frac{\nu_1}{\psi_0(1-\nu_1)}\right)',$$

$$E\{(d_t^{\phi_+})^2\} = \frac{(1-2\nu_{1+}+\nu_{2+})(1-\nu_1)+2(\nu_{1+}-\nu_{2+})(1-\nu_{1+})}{\phi_{0+}^2(1-\nu_1)(1-\nu_2)},$$

$$E(d_t^{\phi_+} d_t^{\phi_-}) = \frac{(\nu_{1+}-\nu_{2+})(1-\nu_{1-})+(\nu_{1-}-\nu_{2-})(1-\nu_{1+})}{\phi_{0+}\phi_{0-}(1-\nu_1)(1-\nu_2)},$$

$$E(d_t^{\phi_+} d_t^\psi) = \frac{\nu_2(1-\nu_{1+})+\nu_{1+}-\nu_{2+}}{\psi_0\phi_{0+}(1-\nu_2)(1-\nu_1)}, \quad E\{(d_t^\psi)^2\} = \frac{\nu_2(1+\nu_1)}{\psi_0^2(1-\nu_2)(1-\nu_1)},$$

and $E\{(d_t^{\phi^-})^2\}$ (or $E(d_t^{\phi^-} d_t^{\psi})$) is obtained by replacing ϕ_{0+} by ϕ_{0-} and ν_{i+} by ν_{i-} in $E\{(d_t^{\phi^+})^2\}$ (or $E(d_t^{\phi^+} d_t^{\psi})$), with $\nu_i = E\{\{\psi_0/a_0(\eta_t)\}^i\}$, $\nu_{i+} = E\{\{\psi_0/(\phi_{0+}(\eta_t^+)^2 + \psi_0)\}^i\}$, $\nu_{i-} = E\{\{\psi_0/(\phi_{0-}(\eta_t^-)^2 + \psi_0)\}^i\}$, for $i = 1, 2$.

Remark 1. When $\alpha_0 = 1$, similar to Li, et.al. (2023), the exact value of the Fisher matrix Υ in the explosive case can be calculated. The derivation is presented in the Appendix S4 in the Supplementary Material.

Remark 2. Theorem 2 shows that the intercept ω_0 is not estimable in the explosive case, due to the non-identifiability of ω in the limit of $L_n(\theta)/n$. This partial parameter unidentifiability was first observed in QMLE of non-stationary ARCH(1) (Jensen and Rahbek, 2004a,b) and later in nonstationary GARCH models (Francq and Zakoian, 2012, 2013).

Remark 3. Theorems 1-2 exclude the critical case $\gamma_{\alpha_0} = 0$. To the best of our knowledge, the behavior of y_t in this critical case remains largely unresolved in the literature. For example, although Francq and Zakoian (2012) claimed to address this issue, their additional Assumption A is difficult to verify under stable innovations. Our extensive simulations suggest that their assumption does not hold. While we could impose simpler assumptions on y_t to obtain the consistency and asymptotic normality of $\hat{\vartheta}_n$, such as $|y_t|/\rho^t \xrightarrow{a.s.} \infty$, $\rho > 1$, as $t \rightarrow \infty$, such conditions are typically unverifiable in practice. Thus, we leave this problem for future research.

Remark 4. As discussed in the Introduction, another way to model asymmetry is to use $\eta_t \sim \mathbf{S}(\alpha, \beta, 1, 0)$ with skewness parameter $\beta \in (-1, 1)$. However, this poses substantial numerical challenges, as the density $f_{\alpha, \beta}$ has no closed form, exhibits unreliable behavior near boundary cases, and is computationally expensive to evaluate repeatedly in optimization. In addition, as $\alpha \rightarrow 2$, stable laws approach the Gaussian law, and β becomes less informative and harder to estimate reliably. Theoretically, extending to skew-stable GARCH(1,1) setting mainly requires: (i) the asymptotic behavior of partial derivatives of $\log f_{\alpha, \beta}$ involving β , and (ii) the identifiability of the skewness parameter β . We leave this extension to future work.

3.2 Estimation of Fisher matrices Σ and Υ

To conduct statistical inference on θ_0 , we provide two estimation methods for the Fisher information matrix Σ or Υ in Theorems 1 – 2.

First, for Σ in the stationary case, note that each entry of Σ contains two components. For example, the entry $\Sigma_{\phi_+ \phi_+} := \Sigma_\sigma \Sigma_\eta$, where $\Sigma_\sigma = \frac{1}{4} E \left\{ \frac{1}{\sigma_t^4(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \phi_+} \frac{\partial \sigma_t^2(\theta_0)}{\partial \phi_+} \right\}$, and $\Sigma_\eta = E \left[\left\{ 1 + \frac{\partial \log f_{\alpha_0}(\eta)}{\partial x} \eta \right\}^2 \right]$. Given observations $\{y_0, y_1, \dots, y_n\}$, it is straightforward that $\widehat{\Sigma}_\sigma := \frac{1}{4n} \sum_{t=1}^n \frac{1}{\widehat{\sigma}_t^4(\widehat{\theta}_n)} \frac{\partial \widehat{\sigma}_t^2(\widehat{\theta}_n)}{\partial \phi_+} \frac{\partial \widehat{\sigma}_t^2(\widehat{\theta}_n)}{\partial \phi_+}$. As for Σ_η , we define two estimators, where $\widehat{\Sigma}_\eta^{\text{int}}$ is based on the integral ex-

pression of Σ_η , and $\widehat{\Sigma}_\eta^{\text{res}}$ is based on the residuals $\widehat{\eta}_t = y_t/\tilde{\sigma}_t(\widehat{\theta}_n)$:

$$\widehat{\Sigma}_\eta^{\text{int}} := \int_{\mathbb{R}} \left\{ 1 + \frac{\partial \log f_{\widehat{\alpha}_n}(u)}{\partial x} u \right\}^2 f_{\widehat{\alpha}_n}(u) du, \quad \widehat{\Sigma}_\eta^{\text{res}} := \frac{1}{n} \sum_{t=1}^n \left\{ 1 + \frac{\partial \log f_{\widehat{\alpha}_n}(\widehat{\eta}_t)}{\partial x} \widehat{\eta}_t \right\}^2.$$

By combining $\widehat{\Sigma}_\sigma$ with $\widehat{\Sigma}_\eta^{\text{int}}$ and $\widehat{\Sigma}_\eta^{\text{res}}$, we obtain two estimators of $\Sigma_{\psi\psi}$, say, $\widehat{\Sigma}_{\psi\psi}^{\text{int}}$ and $\widehat{\Sigma}_{\psi\psi}^{\text{res}}$. The remaining entries can be dealt with analogously, yielding two final estimators $\widehat{\Sigma}^{\text{int}}$ and $\widehat{\Sigma}^{\text{res}}$.

As for Υ in the explosive case, the terms $E(d_t d_t')$ and $E(d_t)$ depend only on the innovation η . For example, $E\{(d_t^\psi)^2\} = \nu_2(1 + \nu_1)/\{\psi_0^2(1 - \nu_2)(1 - \nu_1)\}$, where $\nu_i = E\{\{\psi_0/a_0(\eta_t)\}^i\}$ for $i = 1, 2$. Thus, we similarly define

$$\begin{aligned} \widehat{\mathcal{V}}_i^{\text{int}} &= \int_{\mathbb{R}} \left\{ \frac{\widehat{\psi}_n}{\widehat{\phi}_{n+}(u^+)^2 + \widehat{\phi}_{n-}(u^-)^2 + \widehat{\psi}_n} \right\}^i f_{\widehat{\alpha}_n}(u) du, \\ \widehat{\mathcal{V}}_i^{\text{res}} &= \frac{1}{n} \sum_{t=1}^n \left\{ \frac{\widehat{\psi}_n}{\widehat{\phi}_{n+}(\widehat{\eta}_t^+)^2 + \widehat{\phi}_{n-}(\widehat{\eta}_t^-)^2 + \widehat{\psi}_n} \right\}^i, \quad i = 1, 2. \end{aligned}$$

The remaining part of Υ can be handled similarly. Consequently, we obtain two corresponding estimators, denoted as $\widehat{\Upsilon}^{\text{int}}$ and $\widehat{\Upsilon}^{\text{res}}$. The following theorem shows that these estimators are all strongly consistent.

Theorem 3. *Suppose that Assumptions 1–2 hold.*

- (i). *If $\gamma_{\alpha_0} < 0$, then $\widehat{\Sigma}^{\text{int}} \xrightarrow{a.s.} \Sigma$, and $\widehat{\Sigma}^{\text{res}} \xrightarrow{a.s.} \Sigma$, as $n \rightarrow \infty$.*
- (ii). *If $\gamma_{\alpha_0} > 0$, then $\widehat{\Upsilon}^{\text{int}} \xrightarrow{a.s.} \Upsilon$, and $\widehat{\Upsilon}^{\text{res}} \xrightarrow{a.s.} \Upsilon$, as $n \rightarrow \infty$.*

Numerical studies in Section 5 show that both types of estimators (i.e., $\widehat{\Sigma}^{\text{int}}$ and $\widehat{\Sigma}^{\text{res}}$, $\widehat{\Upsilon}^{\text{int}}$ and $\widehat{\Upsilon}^{\text{res}}$) perform well, each with its own pros and cons. Based on our findings, we provide a practical guidance: for $\widehat{\alpha}_n \in (0, 1]$, the integration-based estimators ($\widehat{\Sigma}^{\text{int}}$ or $\widehat{\Upsilon}^{\text{int}}$) are recommended; and for $\widehat{\alpha}_n \in (1, 2)$, the residual-based estimators ($\widehat{\Sigma}^{\text{res}}$ or $\widehat{\Upsilon}^{\text{res}}$) are preferable.

Regarding computation, since the density $f_\alpha(x)$ involves improper integrals of oscillating functions and lacks an explicit formula, numerical integration techniques are essential. For all numerical calculations involving $f_\alpha(x)$, we recommend applying transformation techniques in Nolan (1997) and Matsui and Takemura (2006). The R code for the numerical calculation of stable densities is available at the GitHub link in Li, et.al. (2023).

3.3 A universal estimator of the asymptotic variance

In Theorems 1–2, we establish the asymptotic variance of $\widehat{\vartheta}_n$ separately for the stationary and explosive cases. However, in practice, the stationarity of the underlying process is typically unknown. Importantly, valid inference for ϑ_0 can still be conducted without imposing any stationarity condition.

Specifically, from Theorem 1(ii), in the stationary case when $\gamma_{\alpha_0} < 0$,

the asymptotic distribution of $\hat{\vartheta}_n$ is $\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \xrightarrow{d} \mathcal{N}(0, \Upsilon_*^{-1})$, where

$$\Upsilon_* = \Sigma_{\vartheta\vartheta} - \Sigma_{\vartheta\omega} \Sigma_{\omega\omega}^{-1} \Sigma'_{\vartheta\omega} = \begin{pmatrix} \Sigma_{\tilde{\vartheta}\tilde{\vartheta}} & \Sigma'_{\tilde{\vartheta}\alpha} \\ \Sigma_{\tilde{\vartheta}\alpha} & \Sigma_{\alpha\alpha} \end{pmatrix} - \Sigma_{\omega\omega}^{-1} \begin{pmatrix} \Sigma_{\tilde{\vartheta}\omega} \\ \Sigma_{\alpha\omega} \end{pmatrix} (\Sigma'_{\tilde{\vartheta}\omega} \quad \Sigma_{\alpha\omega}).$$

Define the residual-based estimator of $\Sigma_{\tilde{\vartheta}\tilde{\vartheta}}$ as follows:

$$\hat{\Sigma}_{\tilde{\vartheta}\tilde{\vartheta}} = \frac{1}{4} \left\{ \frac{1}{n} \sum_{t=1}^n \frac{1}{\tilde{\sigma}_t^4(\hat{\theta}_n)} \frac{\partial \tilde{\sigma}_t^2(\hat{\theta}_n)}{\partial \tilde{\vartheta}} \frac{\partial \tilde{\sigma}_t^2(\hat{\theta}_n)}{\partial \tilde{\vartheta}'} \right\} \frac{1}{n} \sum_{t=1}^n \left\{ 1 + \frac{\partial \log f_{\hat{\alpha}_n}(\hat{\eta}_t)}{\partial x} \hat{\eta}_t \right\}^2,$$

and other submatrices in Υ_* can be estimated accordingly. Let $\hat{\Upsilon}_* = \hat{\Sigma}_{\vartheta\vartheta} - \hat{\Sigma}_{\vartheta\omega} \hat{\Sigma}_{\omega\omega}^{-1} \hat{\Sigma}'_{\vartheta\omega}$. It can be shown that $\hat{\Upsilon}_*$ is a consistent estimator of Υ_* in the stationary case. The following theorem shows that $\hat{\Upsilon}_*$ also consistently estimates the inverse asymptotic variance of $\hat{\vartheta}_n$ in the explosive case.

Theorem 4. *Suppose that Assumptions 1–2 hold.*

- (i). *If $\gamma_{\alpha_0} < 0$, then $\hat{\Upsilon}_* \xrightarrow{a.s.} \Upsilon_*$, as $n \rightarrow \infty$.*
- (ii). *If $\gamma_{\alpha_0} > 0$, then $\hat{\Upsilon}_* \xrightarrow{a.s.} \Upsilon$, as $n \rightarrow \infty$.*

In both cases, $\hat{\Upsilon}_*^{-1}$ is a strongly consistent estimator of the asymptotic variance of $\hat{\vartheta}_n$. Thus, within this unified framework, we can conduct asymptotically valid inference for ϑ_0 without requiring a prior stationarity test.

The asymmetry test below is a direct application of this result.

4. Testing

In this section, we discuss three types of hypothesis tests: stationarity testing, asymmetry testing, and Kolmogorov-type test for diagnostic checking.

4.1 Strict stationarity testing

To ensure valid inference for all estimators including ω_0 , we must study the stationarity of the underlying process. Consider the following two tests:

$$(i) \text{ Strict stationarity test: } H_0 : \gamma_{\alpha_0} < 0 \text{ vs } H_1 : \gamma_{\alpha_0} \geq 0; \quad (4.6)$$

$$(ii) \text{ Explosivity test: } H_0 : \gamma_{\alpha_0} > 0 \text{ vs } H_1 : \gamma_{\alpha_0} \leq 0. \quad (4.7)$$

To estimate γ_{α_0} , we use the residual-based estimator: $\hat{\gamma}_n := \frac{1}{n} \sum_{t=1}^n \log [\hat{\phi}_{n+}(\hat{\eta}_t^+)^2 + \hat{\phi}_{n-}(\hat{\eta}_t^-)^2 + \hat{\psi}_n]$. Theorem 5 gives the asymptotic distribution of $\hat{\gamma}_n$.

Theorem 5. *Let $u_t = \log a_0(\eta_t) - \gamma_{\alpha_0}$, and $\sigma_u^2 = Eu_t^2 < \infty$. Suppose that Assumptions 1–3 hold, it follows that*

$$\sqrt{n}(\hat{\gamma}_n - \gamma_{\alpha_0}) \xrightarrow{d} \mathcal{N}(0, \sigma_\gamma^2) \text{ as } n \rightarrow \infty,$$

$$\text{where } \sigma_\gamma^2 = \begin{cases} \sigma_u^2 + \{a_1' \Sigma^{-1} a_2 - 4(1 - \nu_1)^2 / c_1\}, & \text{if } \gamma_{\alpha_0} < 0, \\ \sigma_u^2, & \text{if } \gamma_{\alpha_0} > 0, \end{cases} \text{ and the ex-}$$

plicit forms of a_1, a_2, ν_1, c_1 are given in the Appendix S2.5.

Remark 5. Compared with Theorem 4.1 in Francq and Zakoïan (2013), the asymptotic variance of $\hat{\gamma}_n$ takes a more complicated form. This difference arises from our distinct innovation assumptions, so the asymptotic behavior of $\hat{\gamma}_n$ also depends on the estimation effect of stability parameter α . The condition $\sigma_u^2 = Eu_t^2 < \infty$ is ensured by the fact that $E(|\eta_t|^{\alpha_0/2}) < \infty$.

Denote $\hat{\sigma}_u^2 = \frac{1}{n} \sum_{t=1}^n \{ \log [\hat{\phi}_{n+}(\hat{\eta}_t^+)^2 + \hat{\phi}_{n-}(\hat{\eta}_t^-)^2 + \hat{\psi}_n] \}^2 - \{ \frac{1}{n} \sum_{t=1}^n \log [\hat{\phi}_{n+}(\hat{\eta}_t^+)^2 + \hat{\phi}_{n-}(\hat{\eta}_t^-)^2 + \hat{\psi}_n] \}^2$. Under the assumptions of Theorem 5, it can be shown that $\hat{\sigma}_u^2$ converges in probability to σ_u^2 . Construct the test statistic $T_n := \sqrt{n}\hat{\gamma}_n/\hat{\sigma}_u$. Then, for the testing problem (4.6) (resp., (4.7)), the test defined by the stationary (resp., explosive) critical region $C^{ST} = \{T_n > \Phi^{-1}(1 - \underline{\alpha})\}$ (resp., $C^{EP} = \{T_n < \Phi^{-1}(\underline{\alpha})\}$), has its asymptotic significance level bounded by $\underline{\alpha} \in (0, 1)$ and is consistent.

4.2 Asymmetry testing

It is particularly interesting to test for the existence of a leverage effect in financial asset returns. Benefiting from the framework of our sAGARCH model (2.2), the asymmetry testing problem takes a simple form of

$$H_0 : \phi_{0+} = \phi_{0-} \quad \text{vs} \quad H_1 : \phi_{0+} \neq \phi_{0-}. \quad (4.8)$$

Consider the following test statistic for symmetry:

$$T_n^S = \frac{\sqrt{n}(\hat{\phi}_{n+} - \hat{\phi}_{n-})}{\hat{\sigma}_s}, \quad \text{with } \hat{\sigma}_s = \sqrt{\mathbf{e}'\hat{\Upsilon}_*^{-1}\mathbf{e}} \quad \text{and } \mathbf{e} = (1, -1, 0, 0)'$$

Note that this symmetry test does not require any stationarity assumption, as it leverages the universal estimator $\hat{\Upsilon}_*$ of the asymptotic variance of $\hat{\vartheta}_n$.

The following corollary holds directly from Theorem 4.

Corollary 1. *Suppose that Assumptions 1–3 hold. For the symmetry test (4.8), the test defined by the critical region $C^S = \{|T_n^S| > \Phi^{-1}(1 - \underline{\alpha}/2)\}$ has the asymptotic significance level $\underline{\alpha} \in (0, 1)$ and is consistent.*

4.3 Diagnostic checking

Diagnostic checking is crucial in time series modeling, while the standard portmanteau test based on autocorrelations of residuals or squared residuals is unsuitable for assessing the adequacy of model (2.2), where the innovation is assumed to follow stable distribution. Here, we propose a Kolmogorov-type test for diagnostic checking. Consider the following null hypothesis:

$$H_0 : \eta_t \sim f_{\alpha_*}(x) \quad (\text{i.e., } \alpha_0 = \alpha_*), \quad (4.9)$$

where $\alpha_* \in (0, 2)$ is a fixed constant. In practice, α_* can be chosen to be $\hat{\alpha}_n$ or an approximate value of $\hat{\alpha}_n$.

Let $F_\alpha(x)$ be the CDF of $\mathbf{S}(\alpha, 0, 1, 0)$, $U_t = F_{\alpha_*}(\eta_t)$, and $V_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^n [\mathbb{1}(U_t \leq r) - r]$ for $r \in [0, 1]$. By the Donsker Theorem and continuous mapping theorem, under H_0 , $\{U_t\}$ are i.i.d. $U[0, 1]$ r.v., and

$$\sup_{0 \leq r \leq 1} |V_n(r)| \xrightarrow{d} \sup_{0 \leq r \leq 1} |\mathbb{B}(r) - r\mathbb{B}(1)| \quad \text{as } n \rightarrow \infty, \quad (4.10)$$

where $\mathbb{B}(\cdot)$ is a standard Brownian motion. Since η_t is not observable in practice, we replace it with the residual $\tilde{\eta}_t$, where $\tilde{\eta}_t = y_t/\sigma_t(\tilde{\theta}_n)$, $t = 1, \dots, n$, and $\tilde{\theta}_n = (\hat{\omega}_n, \hat{\phi}_{n+}, \hat{\phi}_{n-}, \hat{\psi}_n)'$ is the restricted MLE of $\tilde{\theta}_0$ under H_0 . Accordingly, we replace U_t and $V_n(r)$ by \hat{U}_t and $\hat{V}_n(r)$ respectively, where $\hat{U}_t = F_{\alpha_*}(\tilde{\eta}_t)$ and $\hat{V}_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^n [\mathbb{1}(\hat{U}_t \leq r) - r]$, for $r \in [0, 1]$.

Unfortunately as opposed to (4.10), $\sup_{0 \leq r \leq 1} |\hat{V}_n(r)|$ is no longer asymptotically distribution-free, since it involves the effect of parameter estimation. To obtain an asymptotically distribution-free test statistic, we adopt the martingale transformation of Khmaladze (1981) inspired by Bai (2003). Specifically, let $g(r) = (g_1(r), g_2(r))' = (r, f_{\alpha_*}(F_{\alpha_*}^{-1}(r))F_{\alpha_*}^{-1}(r))'$ with its derivative $\dot{g}(r) = (1, \dot{g}_2(r))'$, where $\dot{g}_2(r) = 1 + \frac{\dot{f}_{\alpha_*}(F_{\alpha_*}^{-1}(r))}{f_{\alpha_*}(F_{\alpha_*}^{-1}(r))} F_{\alpha_*}^{-1}(r)$ and $\dot{f}_{\alpha_*}(x) = \partial f_{\alpha_*}(x)/\partial x$. Define Khmaladze's transformation $\widehat{W}_n(r) = \widehat{V}_n(r) -$

$\int_0^r (\dot{g}(s)'C^{-1}(s) \int_s^1 \dot{g}(u)d\widehat{V}_n(u))ds$, where $C(r) = \int_r^1 \dot{g}(s)\dot{g}(s)'ds$. Then, the Kolmogorov-type test statistic is defined as $T_n^D = \sup_{0 \leq r \leq 1} |\widehat{W}_n(r)|$.

Theorem 6. *Suppose model (2.2) is well specified with $\alpha_0 = \alpha_*$ and the Assumptions 1–3 hold. Then for both $\gamma_{\alpha_*} < 0$ and $\gamma_{\alpha_*} > 0$ cases,*

$$T_n^D \xrightarrow{d} \sup_{0 \leq r \leq 1} |\mathbb{B}(r)| \quad \text{as } n \rightarrow \infty.$$

Thus, for the testing problem (4.9), the critical region is defined as $C^D = \{T_n^D > cv_{\underline{\alpha}}^D\}$ at the asymptotic significance level $\underline{\alpha} \in (0, 1)$, where $cv_{\underline{\alpha}}^D$ is the $(1 - \underline{\alpha})$ quantile of the limiting distribution of T_n^D . The critical values at the significance levels of 10%, 5%, and 1% are 1.9571, 2.2340, and 2.8080, respectively, via simulation. Numerical simulations in Section 5 show that T_n^D has satisfactory power in all cases, even for small samples.

Remark 6. In application, T_n^D can be approximately computed with

$$\max_{1 \leq j \leq n} \sqrt{n} \left| \frac{j}{n} - \frac{1}{n} \sum_{k=1}^j \dot{g}(v_k)'C_k^{-1}D_k(v_k - v_{k-1}) \right|,$$

where $D_k = \sum_{i=k}^n \dot{g}(v_i)$, $C_k = \sum_{i=k}^n \dot{g}(v_i)\dot{g}(v_i)'(v_{i+1} - v_i)$, and v_1, \dots, v_n are ordered values of $\widehat{U}_1, \dots, \widehat{U}_n$ with the convention $v_0 = 0$ and $v_{n+1} = 1$.

5. Simulation Studies

5.1 Performance of the MLE

To assess the finite-sample performance of the MLE of θ_0 in the sAGARCH model (2.2), we choose $n = 200, 500, 1000, 2000$ with 1000 replications. First, we consider three stationary cases corresponding to $\alpha_0 = 1.5, 1.0$, and 0.5 . To streamline the presentation, Table 1 reports only the case $\alpha_0 = 1.5$, which is close to the empirical tail index in Section 6. The complete results for all three values of α_0 are provided in the Appendix S1.2.

Table 1 reports the empirical biases (Bias), empirical standard deviations (ESD), asymptotic standard deviations (ASD) of the MLE $\hat{\theta}_n$, together with two estimators $\widehat{\text{ASD}}^{\text{int}}$ and $\widehat{\text{ASD}}^{\text{res}}$ of ASDs from $\hat{\Sigma}^{\text{int}}$ and $\hat{\Sigma}^{\text{res}}$ in Section 3.2. The ASDs are calculated from Theorem 1 (ii). Table 1 shows that the biases are small and decrease with n , while the ESDs, ASDs, and estimated ASDs become close as n increases. The two estimators of the ASD perform similarly well. For heavier-tailed cases, the results are similar for most parameters, but $\hat{\omega}_n$ may be less stable when $\alpha_0 \leq 1$. According to our model, it is likely related to the extreme values of the observations. Since the intercept ω can be viewed as a *scale parameter* for y_t , the estimator of ω is likely influenced by the scale of y_t and may be overestimated in

Table 1: Simulation results for the MLE $\hat{\theta}_n$ for sAGARCH(1,1) under the stationary case with $\alpha_0 = 1.5$.

n		$\omega_0 = 0.2$	$\phi_{0+} = 0.1$	$\phi_{0-} = 0.2$	$\psi_0 = 0.5$	$\alpha_0 = 1.5$
		$\hat{\omega}_n$	$\hat{\phi}_{n+}$	$\hat{\phi}_{n-}$	$\hat{\psi}_n$	$\hat{\alpha}_n$
200	Bias	0.0360	0.0038	0.0008	-0.0007	0.0253
	ESD	0.1213	0.0610	0.0883	0.1086	0.1131
	ASD	0.0890	0.0504	0.0813	0.0878	0.1085
	$\widehat{\text{ASD}}^{\text{int}}$	0.1095	0.0519	0.0818	0.0948	0.1070
	$\widehat{\text{ASD}}^{\text{res}}$	0.1099	0.0521	0.0821	0.0951	0.1076
500	Bias	0.0181	0.0001	-0.0026	0.0016	0.0113
	ESD	0.0653	0.0330	0.0518	0.0585	0.0712
	ASD	0.0563	0.0319	0.0514	0.0555	0.0687
	$\widehat{\text{ASD}}^{\text{int}}$	0.0627	0.0320	0.0511	0.0576	0.0683
	$\widehat{\text{ASD}}^{\text{res}}$	0.0628	0.0320	0.0512	0.0576	0.0684
1000	Bias	0.0080	0.0008	0.0005	-0.0002	0.0052
	ESD	0.0435	0.0239	0.0361	0.0409	0.0474
	ASD	0.0398	0.0225	0.0364	0.0392	0.0485
	$\widehat{\text{ASD}}^{\text{int}}$	0.0422	0.0226	0.0366	0.0397	0.0485
	$\widehat{\text{ASD}}^{\text{res}}$	0.0423	0.0227	0.0367	0.0397	0.0485
2000	Bias	0.0014	-0.0002	-0.0005	0.0005	0.0025
	ESD	0.0284	0.0164	0.0253	0.0276	0.0328
	ASD	0.0282	0.0159	0.0257	0.0278	0.0343
	$\widehat{\text{ASD}}^{\text{int}}$	0.0288	0.0159	0.0256	0.0279	0.0343
	$\widehat{\text{ASD}}^{\text{res}}$	0.0288	0.0159	0.0256	0.0279	0.0343

some cases; see Appendix S1.2 for more discussions.

Fig. 1 plots the histograms of $\sqrt{n}(\hat{\theta}_n - \theta_0)$ with $n = 1000$ and $\theta_0 = (0.2, 0.1, 0.2, 0.5, 1.5)'$. The results demonstrate that $\hat{\theta}_n$ performs well overall and conforms to the asymptotic normal distribution. Further results on the finite-sample performance of MLE in the explosive cases and the estimation of the Lyapunov exponents are also presented in the Appendix S1.3.

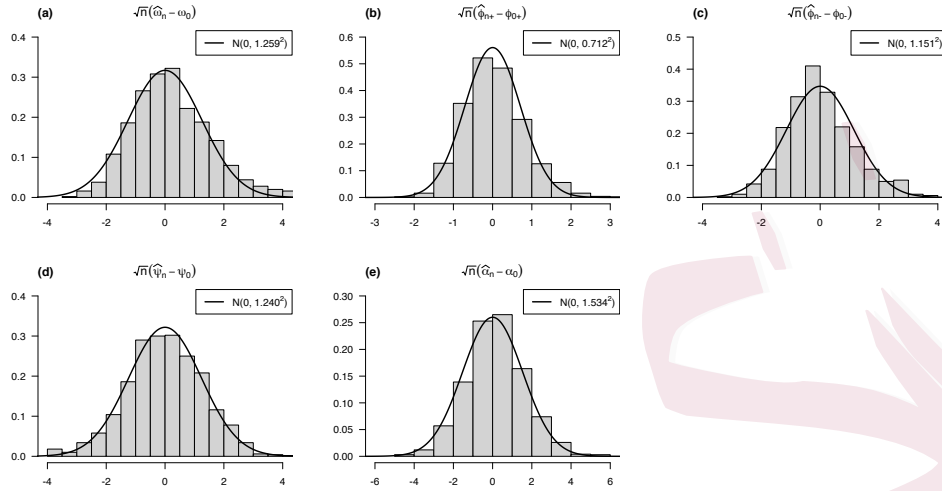


Figure 1: The histograms of $\sqrt{n}(\hat{\theta}_n - \theta_0)$ with $n = 1000$ and $\theta_0 = (0.2, 0.1, 0.2, 0.5, 1.5)'$.

5.2 Performance of the test statistics

Next, we examine the performance of diagnostic checking. The results for the stationarity and asymmetry tests are presented in Appendix S1.4.

For the stationary cases, consider two scenarios: (I) sGARCH(1,1) model with $\theta_0 = (0.2, 0.1, 0.2, 0.5, 1.5)'$. The null hypothesis is $H_0^D : \eta_t \sim f_{1.5}$, motivated by the empirical example in Section 6, and the alternative $H_1 : \eta_t \sim f_{\alpha_*}$ with $\alpha_* \neq 1.5$. (II) sGARCH(1,1) model with $\theta_0 = (0.1, 0.1, 0.2, 0.3, 1)'$. The null $H_0^D : \eta_t \sim f_1$ (i.e., standard Cauchy), with alternative $\eta_t \sim$ Student's t_ν -distribution with $\nu \in [0.5, 5]$. Fig. 2 reports the size and power of T_n^D at the significance level 5%. Here n is 500,

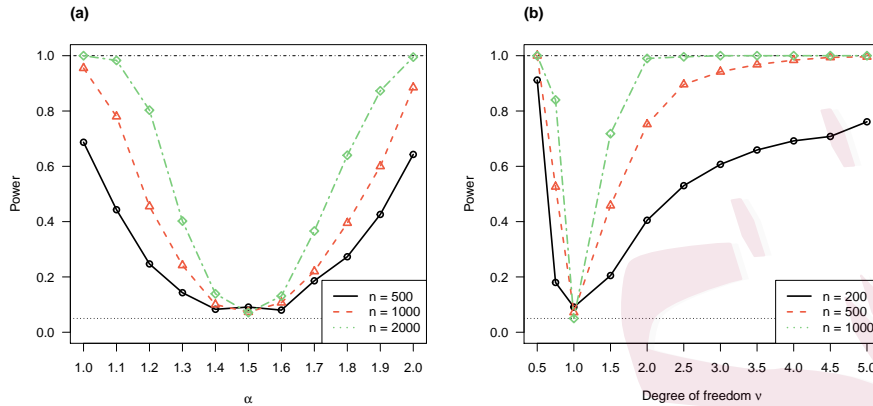


Figure 2: The size and power of T_n^D for stationary cases. (a) $H_1: \eta_t \sim f_\alpha(x)$ with $\alpha \in [1, 2]$. The size corresponds to $\alpha = 1.5$. (b) $H_1: \eta_t \sim \text{Student's } t_\nu$ -distribution with $\nu \in [0.5, 5]$. The size corresponds to $\nu = 1$ (i.e., $\eta_t \sim f_1$). The horizontal dotted line at the bottom denotes the 5% significance level.

1000, 2000 for (I), and 200, 500,1000 for (II), each with 1000 replications. The size is generally close to the nominal level, except for $n = 500$ in (I), and the power increases as the stable exponent α or the degree of freedom ν deviates from the null. The size and power improve with increasing n .

We further examine the performance of T_n^D under explosive cases. For sAGARCH(1,1) model with $\theta_0 = (0.1, 0.1, 0.2, 0.5, 1)'$, consider two alternatives to the null $H_0^D: \eta_t \sim f_1$: (I) $\eta_t \sim f_{\alpha_*}$ with $\alpha_* \neq 1$, and (II) $\eta_t \sim t_\nu$ with $\nu \in [0.5, 5]$. Fig. 3 reports the corresponding results for explosive cases, again based on 1000 replications and the same sample sizes. The test shows satisfactory size and increasing power as the innovation distribution moves

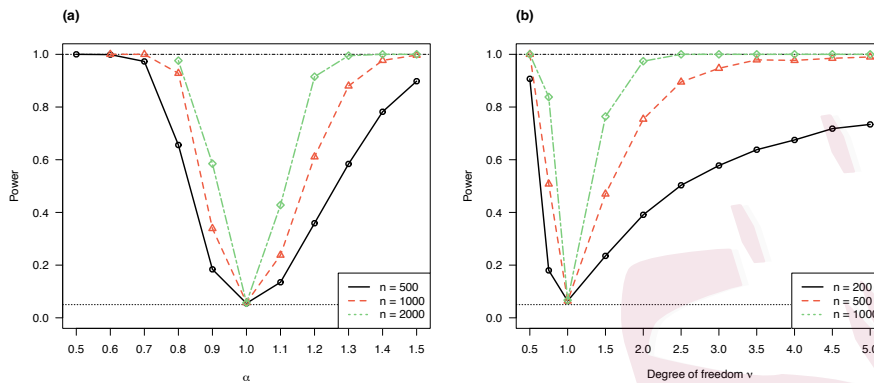


Figure 3: The size and power of T_n^D for explosive cases. (a) $H_1 : \eta_t \sim f_\alpha(x)$, with $\alpha \in [0.5, 1.5]$; (b) $H_1 : \eta_t \sim t_\nu$ -distribution, with $\nu \in [0.5, 5]$.

away from the null specification. In particular, (II) demonstrates that T_n^D can effectively distinguish Student's t innovations from misspecified stable innovations, confirming its usefulness for finite-sample diagnostic checking.

6. Empirical Examples

In this section, we analyze four individual stock return series, which are the same as those in Francq and Zakoian (2012, 2013) for ease of comparison. They are the daily series of Icagen (NasdaqGM: ICGN, May 31, 2007–Feb. 7, 2011), Monarch Community Bancorp (NasdaqCM: MCBF, Aug. 28, 2007 – Feb. 7, 2011), KV Pharmaceutical (NYSE: KV-A, Mar. 31, 2006 – Feb. 7, 2011), and China MediaExpress (NasdaqGS: CCME, Mar. 31, 2009 – Feb. 7, 2011). Fig. 4 displays the four stock return series.

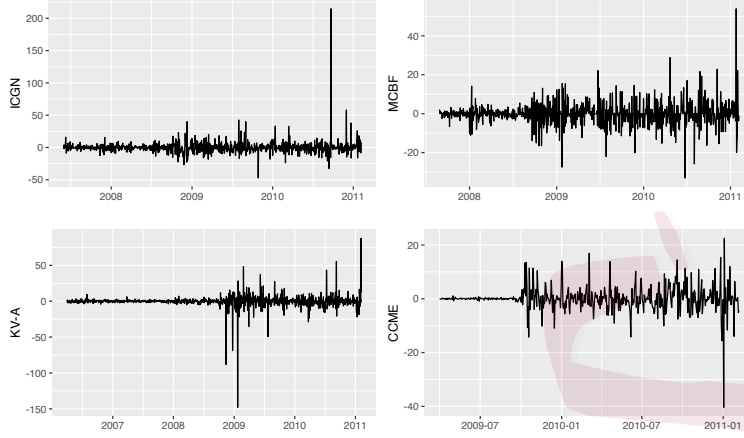


Figure 4: The graphs of four individual stock return series (%).

We fit an sAGARCH(1,1) model to each series and report the results in Table 2, including parameter estimates, the explosivity test T_n , asymmetry test T_n^S , and diagnostic statistics T_n^D , all with corresponding p -values. The estimated ASDs in parentheses are computed using the universal estimator $\hat{\Upsilon}_*$ in Section 3.3. We further conduct out-of-sample VaR/ES backtesting. Define $\widehat{\text{VaR}}_{t+1}(\tau) = -\sigma_{t+1}(\hat{\theta}_n)q_{\hat{\alpha}_n}(\tau)$, $\widehat{\text{ES}}_{t+1}(\tau) = -\sigma_{t+1}(\hat{\theta}_n)E[\eta|\eta \leq q_{\hat{\alpha}_n}(\tau)]$, where $q_\alpha(\tau)$ is the τ -quantile of $\eta \sim f_\alpha(\cdot)$. We use the first 80% of the sample for estimation, the remaining 20% for one-step-ahead VaR/ES forecasting, and set $\tau = 10\%$. We compute three backtesting statistics: Kupiec's unconditional coverage test applied to the exceedance indicators $I_t = \mathbb{1}\{-y_t > \widehat{\text{VaR}}_{t+1}(\tau)\}$, Christoffersen's independence test to examine clustering in $\{I_t\}$, and McNeil-Frey ES test of whether the excess condi-

tional shortfall is i.i.d. with zero mean. Rejection of any of the three null hypotheses indicates that the risk model is misspecified. For comparison, we also report the QMLE results for the asymmetric GARCH(1,1) model in Francq and Zakoïan (2013), and the MLE results for two benchmark models that incorporate asymmetry in the innovation: a skew- t GARCH(1,1) model and a skew-stable GARCH(1,1) model that allows skew-stable innovation ($\beta \neq 0$). To save space, we report the AIC, computational time (s), and the minimum p -value among three VaR/ES backtests (“VaR/ES”). For AGARCH(1,1), p -values of nonstationarity test (“ p -value- T_n ”) and asymmetry test (“ p -value- T_n^S ”) are also reported. Bold entries highlight notable differences in testing results or models with better performance.

Table 2 reveals several notable findings:

(i) First, the sAGARCH(1,1) model provides a competitive fit. All estimated parameters are significant, and its AIC values are lower than those of the AGARCH(1,1) and skew- t GARCH(1,1) models, except for MCBF. The skew-stable GARCH(1,1) model performs comparably in terms of AIC, but requires substantially more computational time due to the optimization of an additional skewness parameter and the complexity of stable densities. It demonstrates that incorporating asymmetry in the volatility process σ_t achieves similar fitting performance while reducing computational cost.

Table 2: Empirical comparison of sAGARCH(1,1) and benchmark GARCH-type models. Notable differences in test results and the best-fitting models are highlighted in bold.

Model		ICGN	MCBF	KV-A	CCME
	n	928	868	1221	488
sAGARCH (1,1)	$\hat{\phi}_{n+}$	0.098 (0.026)	0.022 (0.008)	0.033 (0.008)	0.084 (0.022)
	$\hat{\phi}_{n-}$	0.164 (0.035)	0.029 (0.008)	0.039 (0.008)	0.090 (0.027)
	$\hat{\psi}_n$	0.419 (0.056)	0.884 (0.022)	0.835 (0.018)	0.766 (0.034)
	$\hat{\alpha}_n$	1.556 (0.057)	1.369 (0.065)	1.587 (0.045)	1.527 (0.084)
	p -value- T_n	0.000	0.056	0.000	0.285
	p -value- T_n^S	0.108	0.532	0.527	0.849
	p -value- T_n^D	0.014	0.000	0.588	0.045
	AIC	6011.4	4987.2	6278.1	2185.4
	comp-time	56.3	39.8	83.4	36.7
	VaR/ES	0.106	0.149	0.122	0.384
AGARCH (1,1) with QMLE	p -value- T_n	0.008	0.515	0.708	0.611
	p -value- T_n^S	0.037	0.850	0.052	0.503
	AIC	6430.5	5275.5	7395.7	2272.1
	comp-time	0.2	0.3	0.03	0.2
skew- t GARCH(1,1)	VaR/ES	0.041	0.009	0.001	0.028
	AIC	6056.0	4966.3	6431.2	2248.2
	comp-time	4.6	5.6	7.9	3.1
skew-stable GARCH(1,1)	VaR/ES	0.012	0.098	0.040	0.020
	AIC	6013.1	4987.2	6277.5	2183.1
	comp-time	109.4	217.9	234.1	70.8
	VaR/ES	0.012	0.151	0.811	0.533

(ii) Second, the testing results highlight the role of heavy-tailed modeling.

For stationarity testing, the results for ICGN and CCME are consistent with those in Francq and Zakoïan (2013), whereas KV-A and MCBF lead to different conclusions under the stable-innovation framework. For ICGN, the null hypothesis of symmetry is not rejected at the 5% level using our model, whereas the AGARCH(1,1) model in Francq and Zakoïan (2013) detects asymmetry. This indicates that apparent explosivity or leverage

effects may be sensitive to the assumed innovation distribution and may partly reflect the influence of extreme observations. In fact, after removing an obvious outlier around September 22, 2010 from ICGN, the series is tested to be symmetric even under the AGARCH(1,1) model.

(iii) Third, the values of T_n^D for the ICGN, KV-A, and CCME are below the critical values of $\sup_{0 \leq r \leq 1} |\mathbb{B}(r)|$ at the significance levels of either 1% or 5%. In contrast, the values for the MCBF are relatively large. This discrepancy is likely due to the apparent presence of change points in the MCBF time series, as seen in Fig. 4, which also contributes to the series' nonstationarity. Specifically, volatility increases sharply around mid-2008, possibly related to the financial crisis during that period. Testing for change points is thus an interesting topic for future research.

(iv) Fourth, the VaR/ES backtesting results support the adequacy of the proposed model for tail-risk evaluation. For sAGARCH(1,1), all reported VaR/ES p -values exceed 0.1, suggesting no evidence of misspecification in the tail-risk forecasts. In contrast, the AGARCH(1,1) and skew- t GARCH(1,1) models reject at least one VaR/ES backtest in most cases, and the skew-stable GARCH(1,1) performs worse than sAGARCH(1,1) for ICGN.

We further examine the model-fitting performance of sAGARCH(1,1) by analyzing residuals in Fig. 5. It compares the CDF of the stable dis-

tribution (with estimated $\hat{\alpha}_n$) against the empirical CDF (ECDF) of the residual. For ICGN, KV-A and CCME, the empirical CDF closely aligns with the true one, illustrating the goodness-of-fit of sAGARCH(1,1) model for these datasets. As for MCBF, there exists a certain difference between two CDFs, which suggests room for improvement—for instance, by incorporating change points into the model. Despite this, our sAGARCH(1,1) model still outperforms the original AGARCH(1,1) model.

We also evaluate the performance of model (2.2) against the AGARCH model with t -innovations on several well-known portfolio returns in Appendix S1.5. The results illustrate the merits of our model in capturing aggregate market behavior and more effectively characterizing portfolio risk.

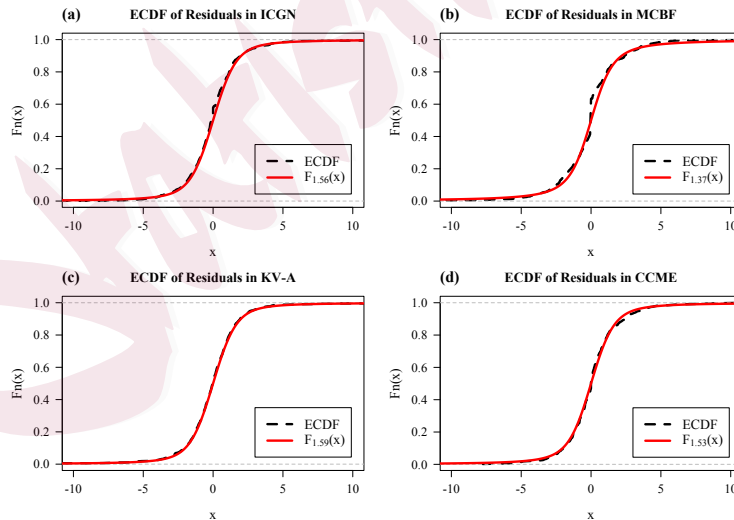


Figure 5: The ECDFs of residuals of four stock return series (%).

7. Concluding Remarks

In this paper, we propose a novel sAGARCH model that operates without moment conditions. The model effectively captures excess kurtosis, volatility clustering, and leverage effects of financial return series. We develop a comprehensive statistical inference framework for the sAGARCH model, unifying both stationary and explosive cases. This addresses a critical gap in the literature on inference of heavy-tailed GARCH-type models. Through extensive simulations, we uncover that the intercept estimator in heavy-tailed GARCH-type models can exhibit poor finite-sample performance. This finding highlights potential risks in financial return forecasting.

To conclude, we outline several promising extensions. First, generalizing the sAGARCH(1,1) model to higher-order specifications (sAGARCH(p, q)) is natural. As addressed in Chan and Ng (2009), such an extension would involve technical challenges in deriving the statistical inference theory. Second, over-parametrization arises in estimating sAGARCH(p, q) models. When true parameters lie on the boundary of parameter space, inference will become nonstandard and more complicated. Third, return series often exhibit structural breaks, making change point detection important for sAGARCH(p, q) modeling. Finally, the asymptotic behavior of sAGARCH(1,1) model in the critical case ($\gamma_{\alpha_0} = 0$) remains an open question.

Supplementary Materials

The Supplementary Material contains additional simulation results, empirical analyses on portfolio returns, and proofs of all theorems.

Acknowledgements

The authors are very grateful to the co-editor, the associate editor, and two anonymous referees for their constructive suggestions and comments, which led to significant improvement. We would like to thank Professor Howell Tong for his insightful suggestions. Tao's research is supported by the NSFC (No.72503086) and the Natural Science Foundation of Guangdong Province (No.2025A1515010697). Gong's research is supported by the NSFC (No.72501294) and the Innovation Research Foundation of NUDT (No.ZK25-61). Li's research is supported by the NSFC (No.72471127).

References

- Andrews, B., Calder, M. and Davis, R.A. (2009). Maximum likelihood estimation for α -stable autoregressive processes. *Ann. Statist.* **37**(4), 1946–1982.
- Bai, J. (2003). Testing parametric conditional distributions of dynamic models. *Review of Economics and Statistics* **85**(3), 531–549

REFERENCES

- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31**(3), 307–327.
- Bougerol, P. and Picard, N. (1992a) Stationarity of GARCH processes and of some nonnegative time series. *J. Econometrics* **52**(1-2), 115–127.
- Bougerol, P. and Picard, N. (1992b) Strict Stationarity of generalized autoregressive processes. *Ann. Probability* **20**(4), 1714–1730.
- Chan, N.H. and Ng, C.T. (2009). Statistical inference for non-stationary GARCH(p, q) models. *Electron. J. Stat.* **3**, 956–992.
- Calzolari, G., Halbleib, R. and Parrini, A. (2014). Estimating GARCH-type models with symmetric stable innovations: Indirect inference versus maximum likelihood. *Comput. Statist. and Data Analysis* **76**, 158–171.
- Chen, Y. and Wang, R., (2025). Infinite-mean models in risk management: Discussions and recent advances. *Risk Sciences*, **1**(100003).
- Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**(4), 987–1007.
- Fama, E. (1965). The behavior of stock market prices. *J. Bus.* **38**(1), 34–105.
- Fan, J. and Yao, Q. (2017). *The Elements of Financial Econometrics*. Cambridge University Press, Cambridge.
- Franco, C. and Zakoian, J.-M. (2004). Maximum likelihood estimation of pure GARCH and

- ARMA-GARCH processes. *Bernoulli* **10**(4), 605–637.
- Franqc, C. and Zakoian, J.-M. (2012). Strict stationarity testing and estimation of explosive and stationary generalized autoregressive conditional heteroscedasticity models. *Econometrica* **80**(2), 821–861.
- Franqc, C. and Zakoian, J.-M. (2013). Inference in nonstationary asymmetric GARCH models. *Ann. Stat.* **41**(4), 1970–1998.
- Franqc, C. and Zakoian, J.-M. (2019). *GARCH Models: Structure, Statistical Inference and Financial Applications*, 2nd Edn. John Wiley.
- Hall, P. and Yao, Q. (2003). Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* **71**(1), 285–317.
- Harvey, A.C. (2013). *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Cambridge University Press, Cambridge.
- Jensen, S. T. and Rahbek, A. (2004a). Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. *Econometrica* **72**(2), 641–646.
- Jensen, S. T. and Rahbek, A. (2004b). Asymptotic inference for nonstationary GARCH. *Econometric Theory* **20**(6), 1203–1226.
- Khmaladze, E.V. (1981). Martingale approach in the theory of goodness-of-fit tests. *Theory of Probability and its Applications* **26**(2), 240–257.
- Li, D., Tao, Y., Yang, Y. and Zhang, R. (2023). Maximum likelihood estimation for α -stable

REFERENCES

- double autoregressive models. *J. Econometrics* **236**(1), 105471.
- Li, D. and Zhu, K. (2020). Inference for asymmetric exponentially weighted moving average models. *Journal of Time Series Analysis*, **41**(1), 154–162.
- Liu, S.M. and Brorsen, B.W. (1995). Maximum likelihood estimation of a GARCH-stable model. *J. Applied Econometrics* **10**(3), 273–285.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *J. Bus.* **36**(4), 394–419.
- Mandelbrot, B. (1997). *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*. Springer, New York.
- Matsui, M. and Takemura, A. (2006). Some improvements in numerical evaluation of symmetric stable density and its derivatives. *Comm. Statist. Theory Methods* **35**(1), 149–172.
- McCulloch, J. (1985). Interest-risk sensitive deposit insurance premia: Stable ACH estimates. *Journal of Banking and Finance* **9**(1), 137–156.
- Mittnik, S., Paoletta, M. and Rachev, S. (2002). Stationarity of stable power-GARCH processes. *J. Econometrics* **106**(1), 97–107.
- Nolan, J.P. (1997). Numerical calculation of stable densities and distribution functions. *Comm. Statist. Stochastic Models*. **13**(4), 759–774.
- Nolan, J.P. (2020). *Univariate Stable Distributions: Models for Heavy Tailed Data*. Springer, Cham.
- Panorska, A., Mittnik, S. and Rachev, S. (1995). Stable GARCH models for financial time series.

Applied Mathematics Letters **8**(5), 33–37.

Samuelson, P.A. (1967). Efficient portfolio selection for Pareto-Lévy investments. *The Journal of Financial and Quantitative Analysis* **2**(2), 107–122.

Uchaikin, V.V. and Zolotarev, V.M. (1999). *Chance and Stability: Stable Distributions and their Applications*. VSP, Utrecht.

Wang, G., Zhu, K., Li, G. and Li, W.K. (2022). Hybrid quantile estimation for asymmetric power GARCH models. *J. Econometrics*, **227**(1), 264–284.

Zhang, X., Zhang, R., Li, Y. and Ling, S. (2022). LADE-based inferences for autoregressive models with heavy-tailed G-GARCH(1, 1) noise. *J. Econometrics*, **227**(1), 228–240.

Zhu, Q., Tan, S., Zheng, Y. and Li, G. (2023). Quantile autoregressive conditional heteroscedasticity. *J. Royal Statistical Society Series B: Statistical Methodology*, **85**(4), 1099–1127.

Department of Statistics and Data Science, Southern University of Science and Technology, Shenzhen 518055, China

E-mail: taoyx@sustech.edu.cn

College of Systems Engineering, National University of Defense Technology, Changsha, China

E-mail: gonghuan@nudt.edu.cn

Department of Statistics and Data Science, Tsinghua University, Beijing 100084, China

E-mail: malidong@tsinghua.edu.cn