Heavy Tails and Financial Time Series Models

Richard A. Davis
Columbia University
www.stat.columbia.edu/~rdavis

Thomas Mikosch
University of Copenhagen

Outline

☞ Financial time series modeling

- General comments
- Characteristics of financial time series
- Examples (exchange rate, Merck, Amazon)
- Multiplicative models for log-returns (GARCH, SV)

☞ Regular variation

- Multivariate case

☞ Applications of regular variation

- Stochastic recurrence equations (GARCH)
- Stochastic volatility
- Extremes and extremal index
- Limit behavior of sample correlations

☞ Wrap-up

Financial Time Series Modeling

One possible goal: Develop models that capture essential features of financial data.

Strategy: Formulate families of models that at least exhibit these key characteristics. (e.g., GARCH and SV)

Linkage with goal: Do fitted models actually capture the desired characteristics of the real data?

Answer wrt to GARCH and SV models: Yes and no. Answer may depend on the features.

Stărică's paper: "Is GARCH(1,1) Model as Good a Model as the Nobel Accolades Would Imply?"

Stărică's paper discusses inadequacy of GARCH(1,1) model as a "data generating process" for the data.

Financial Time Series Modeling (cont)

Goal of this talk: compare and contrast some of the features of GARCH and SV models.

- Regular-variation of finite dimensional distributions
- Extreme value behavior
- Sample ACF behavior

Characteristics of financial time series

Define $X_t = \ln(P_t) - \ln(P_{t-1})$ (log returns)

- heavy tailed

$$P(|X_1| > x) \sim RV(-\alpha), \quad 0 < \alpha < 4.$$

- uncorrelated

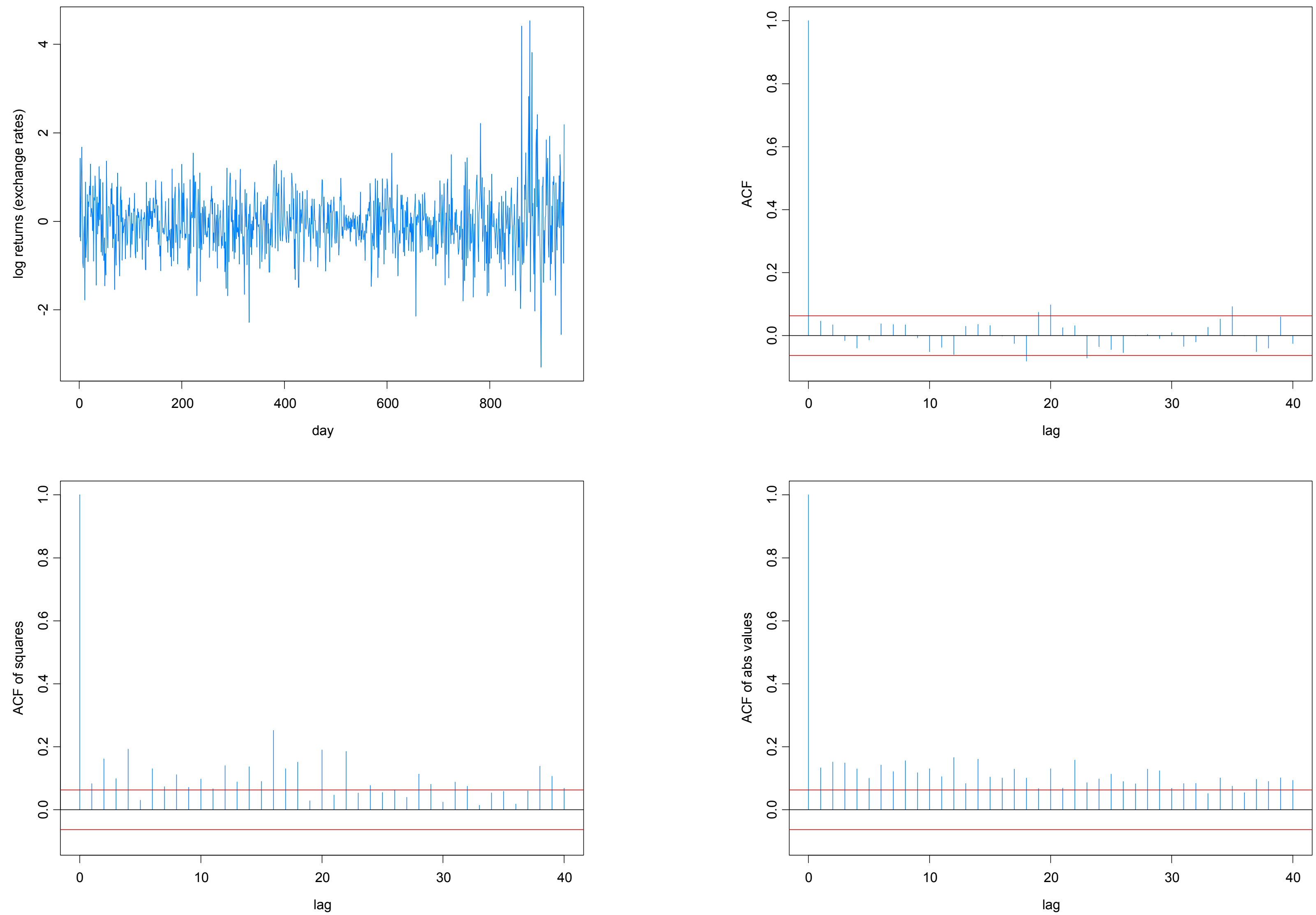
$\hat{\rho}_X(h)$ near 0 for all lags h > 0

- $|X_t|$ and X_t^2 have slowly decaying autocorrelations

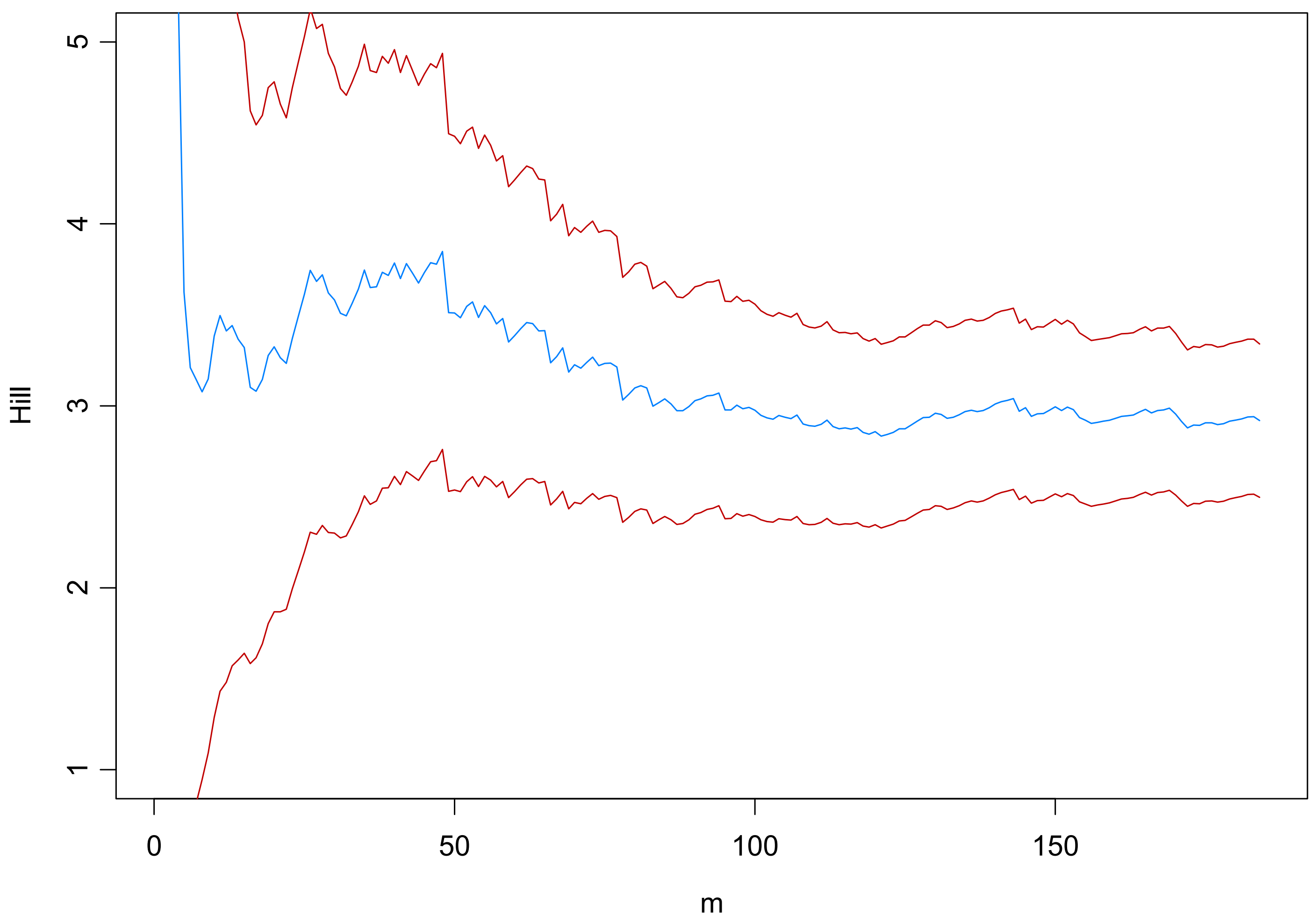
$\hat{\rho}_{|X|}(h)$ and $\hat{\rho}_{X^2}(h)$ converge to 0 *slowly* as h increases.

- process exhibits 'volatility clustering'.

Example: Pound-Dollar Exchange Rates
(Oct 1, 1981 – Jun 28, 1985; Koopman website)

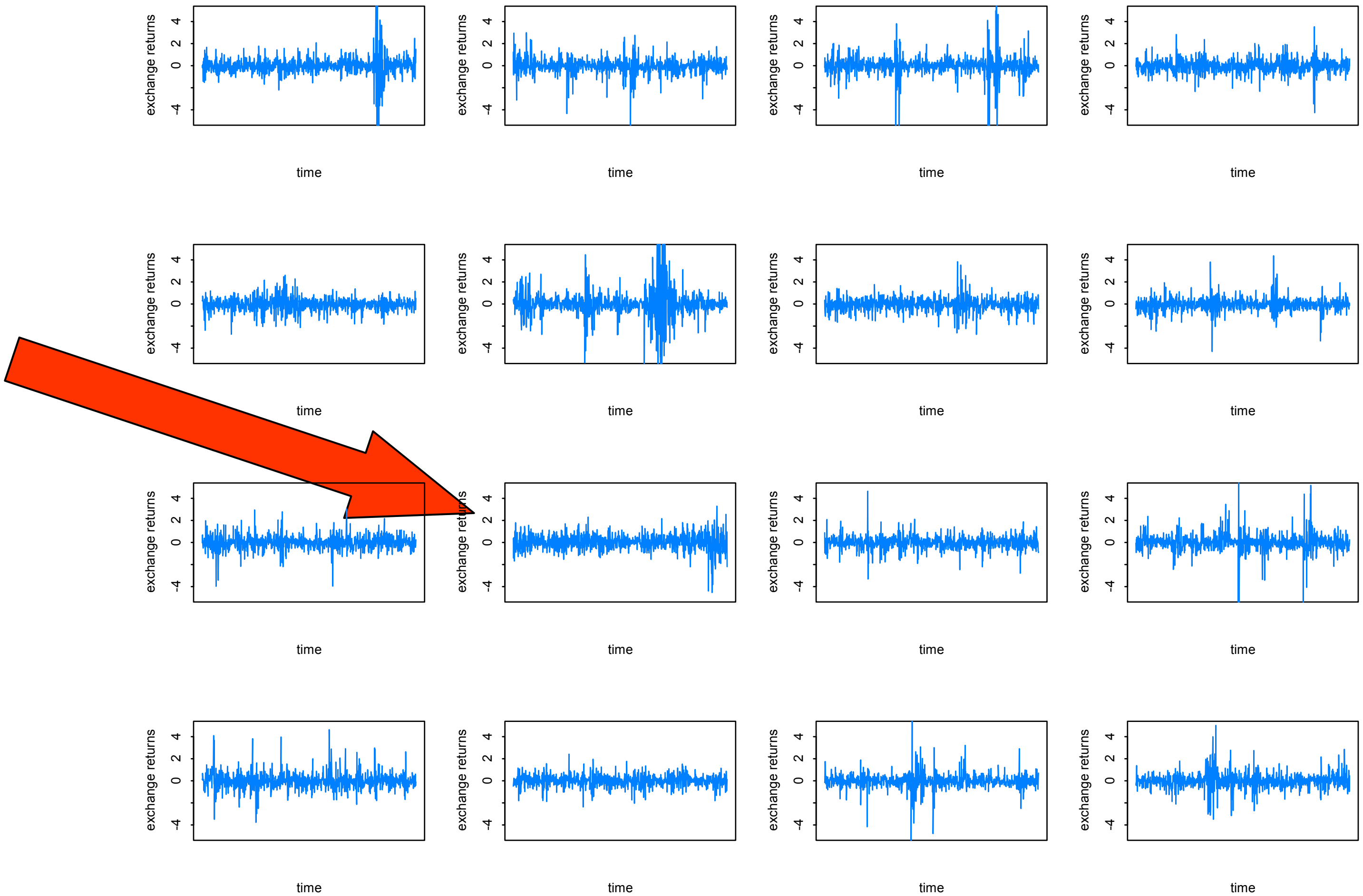


Example: Pound-Dollar Exchange Rates
Hill's estimate of alpha (Hill Horror plots-Resnick)



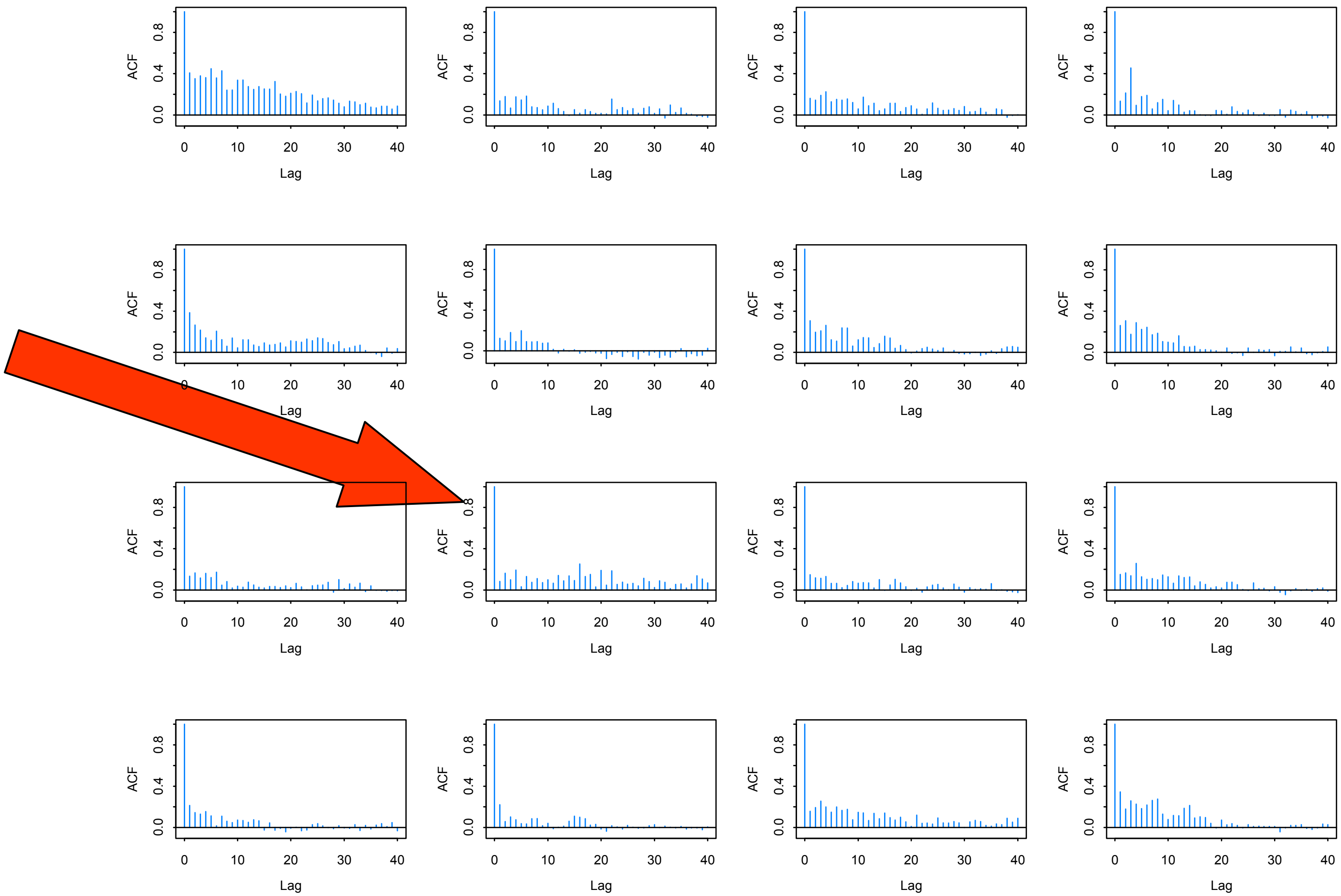
Stărică Plots for Pound-Dollar Exchange Rates

15 realizations from GARCH model fitted to exchange rates + real exchange rate data. Which one is the real data?

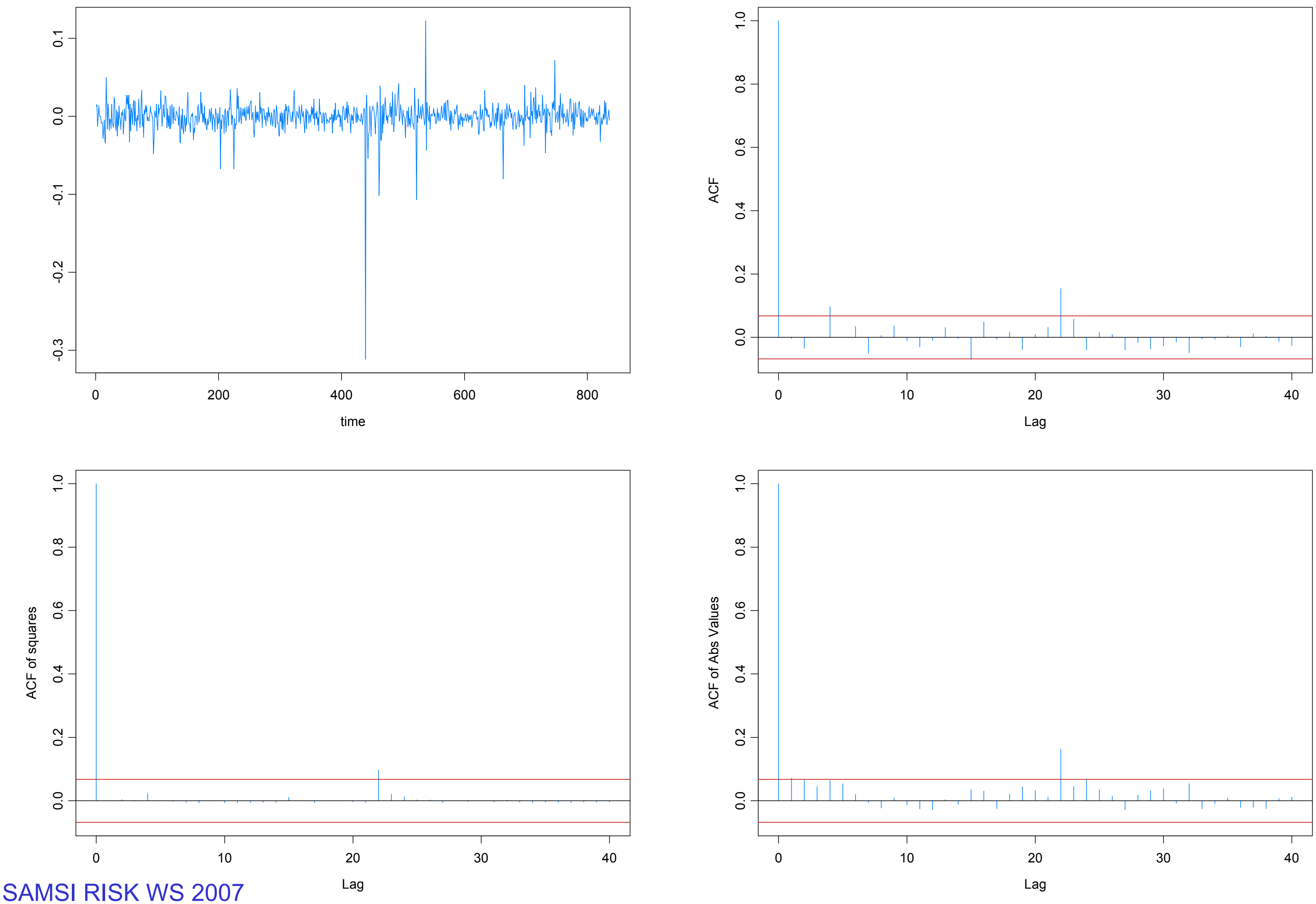


Stărică Plots for Pound-Dollar Exchange Rates

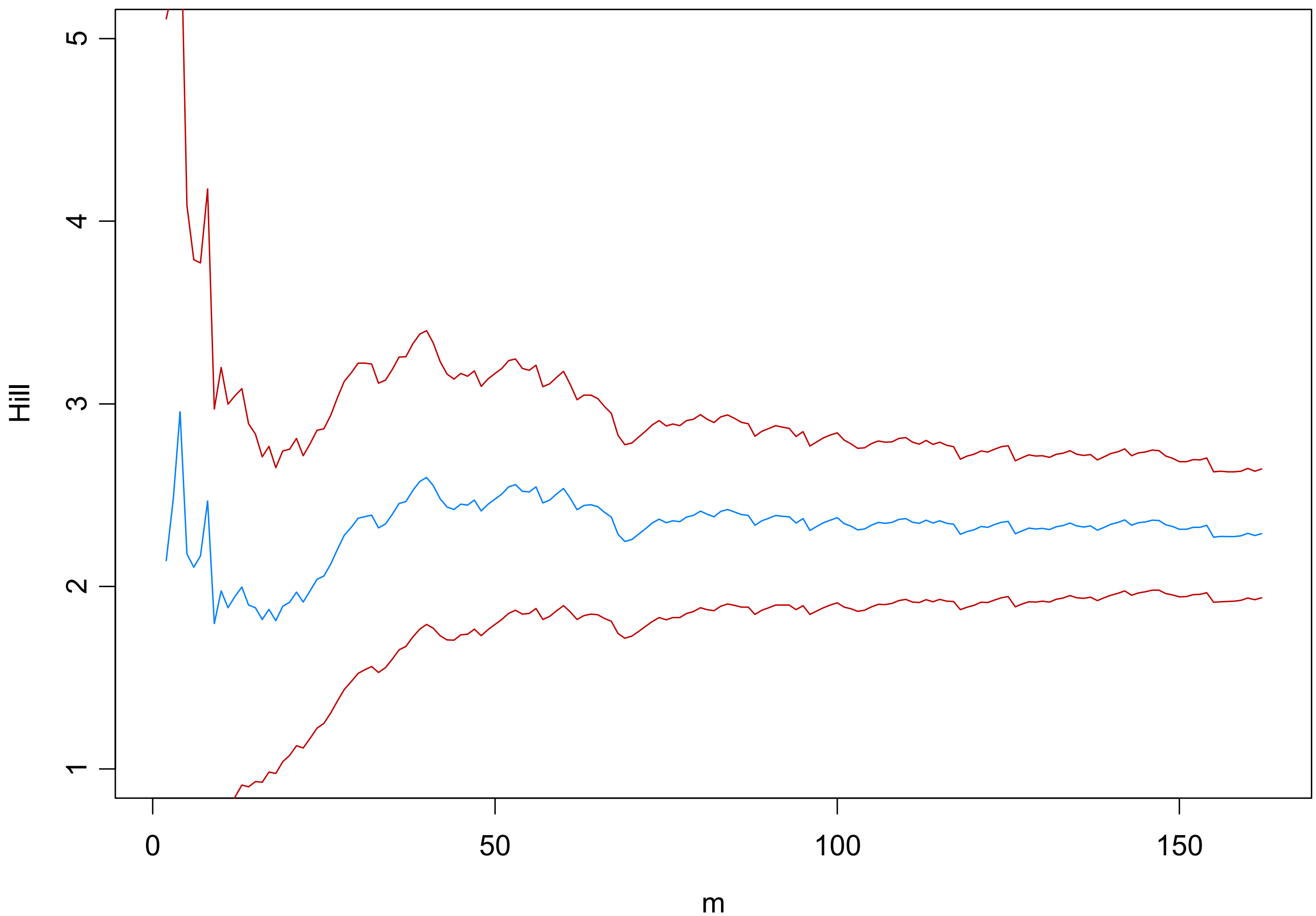
ACF of the squares from the 15 realizations from the GARCH model on previous slide.



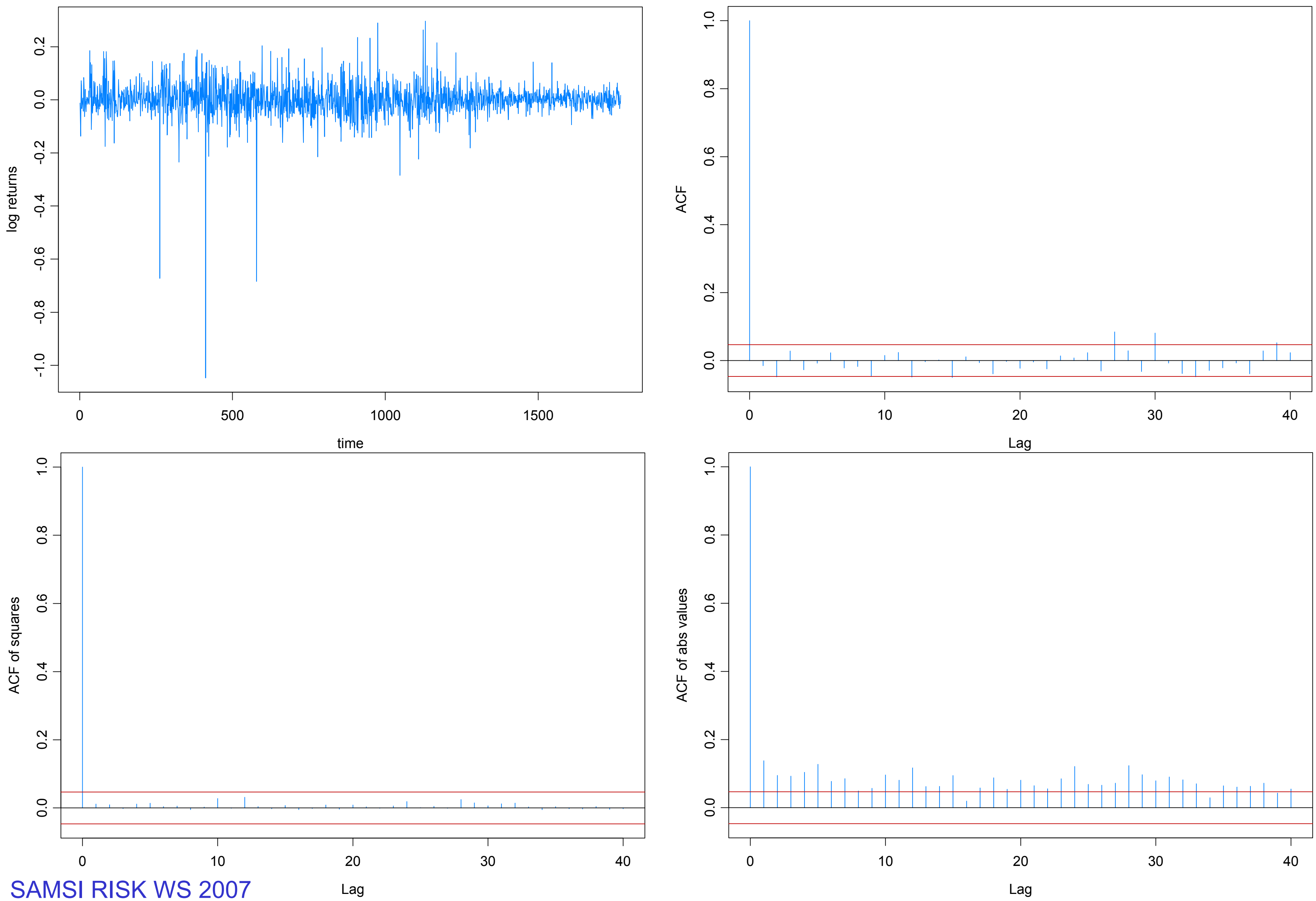
Example: Merck log(returns)
(Jan 2, 2003 – April 28, 2006; 837 observations)



Example: Merck log-returns
Hill's estimate of alpha (Hill Horror plots-Resnick)

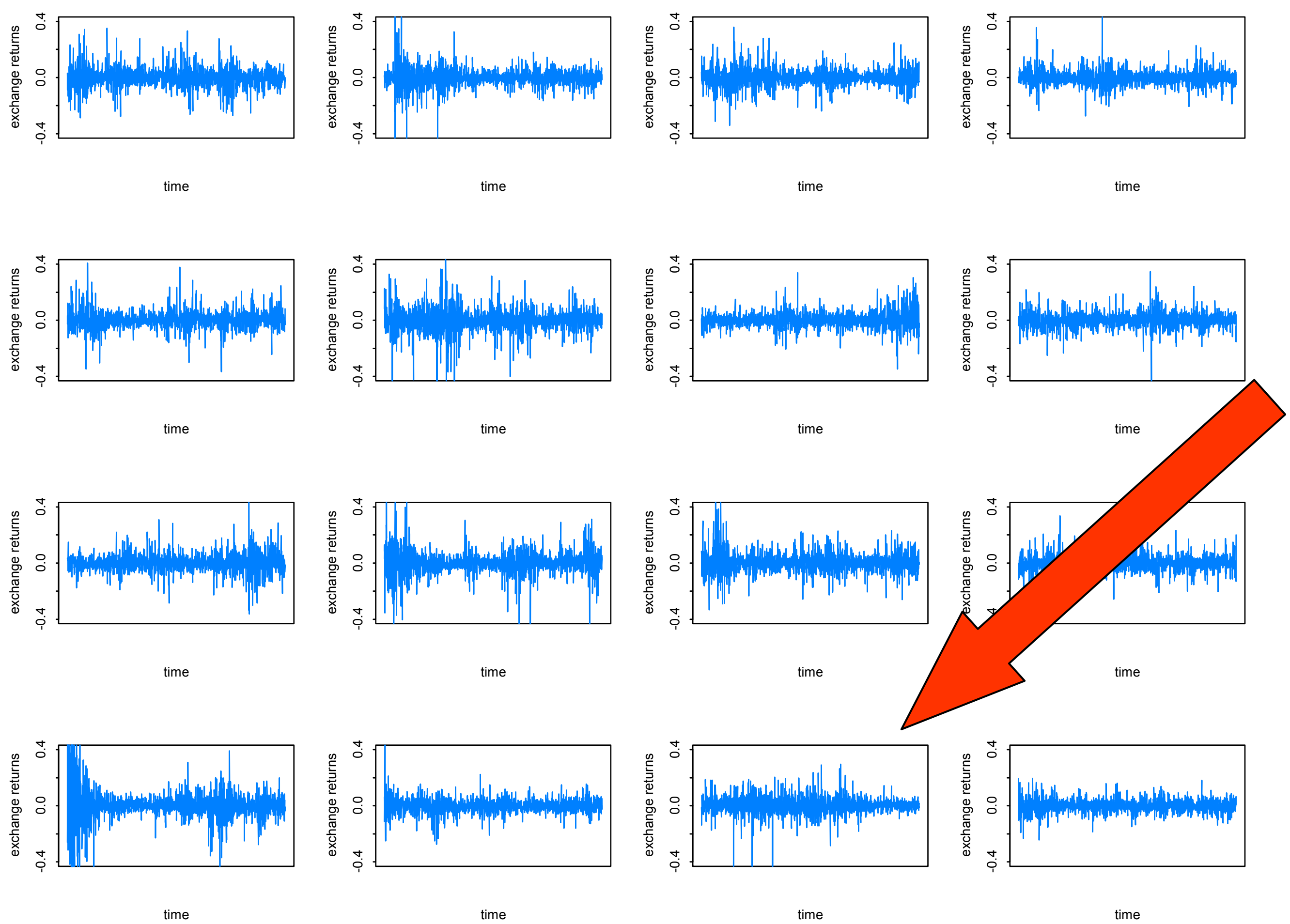


Example: Amazon-returns (May 16, 1997 – June 16, 2004)



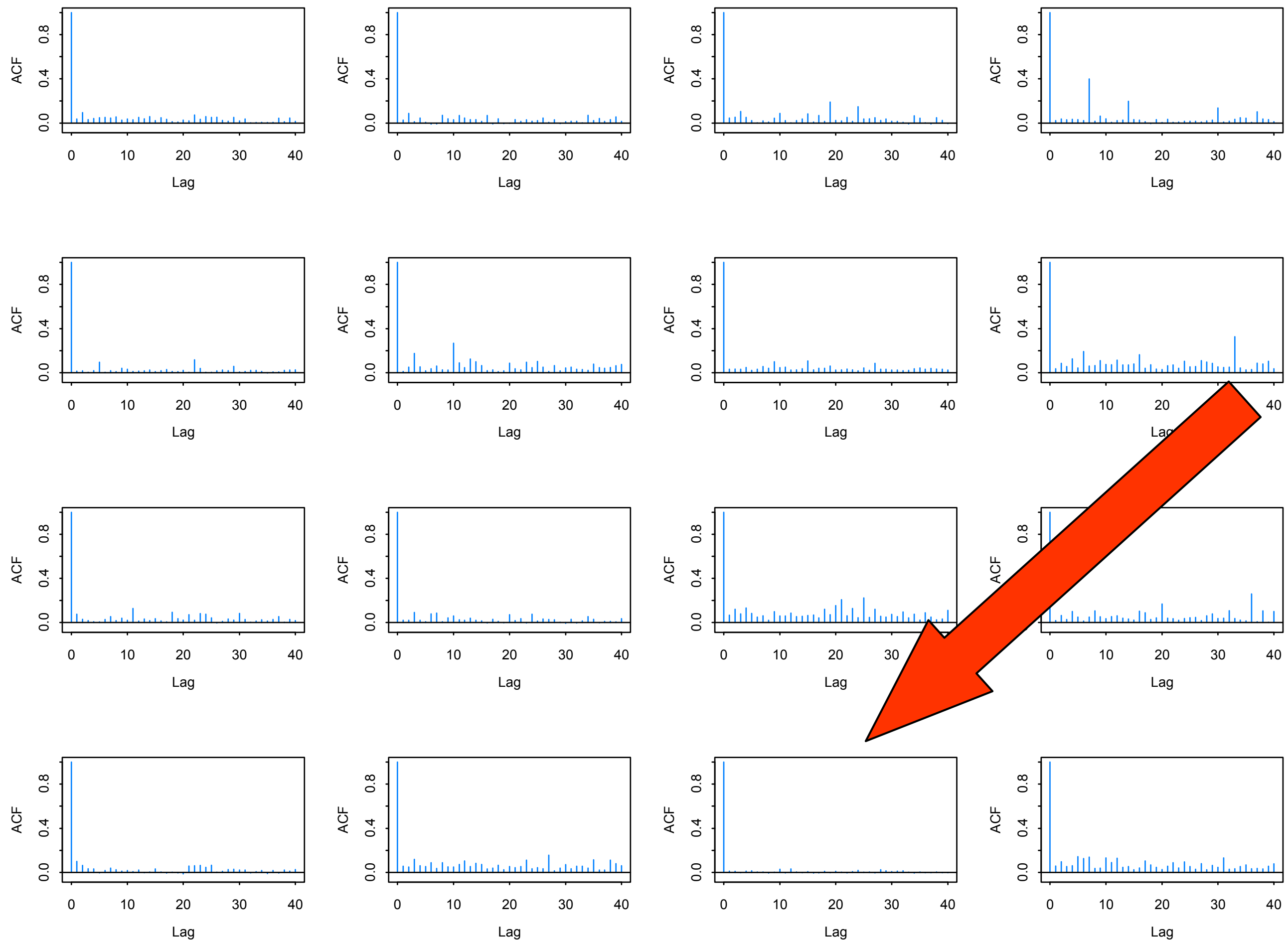
Stărică Plots for the Amazon Data

15 realizations from GARCH model fitted to Amazon + exchange rate data. Which one is the real data?



Stărică Plots for Amazon

ACF of the squares from the 15 realizations from the GARCH model on previous slide.



Multiplicative models for log(returns)

Basic model

$$X_t = \ln(P_t) - \ln(P_{t-1}) \quad \text{(log returns)}$$
$$= \sigma_t Z_t ,$$

where

- $\{Z_t\}$ is IID with mean 0, variance 1 (if exists). (e.g. N(0,1) or a *t*-distribution with ν df.)
- $\{\sigma_t\}$ is the volatility process
- σ_t and Z_t are independent.

Properties:

- $EX_t = 0$, $Cov(X_t, X_{t+h}) = 0$, $h>0$ (uncorrelated if $Var(X_t) < \infty$)
- conditional heteroscedastic (condition on σ_t).

Two models for log(returns)-cont

$X_t = \sigma_t Z_t$ (observation eqn in state-space formulation)

(i) GARCH(1,1) (General AutoRegressive Conditional Heteroscedastic – observation-driven specification):

$$X_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \ \{Z_t\} \sim \mathrm{IID}(0,1)$$

(ii) Stochastic Volatility (parameter-driven specification):

$$X_t = \sigma_t Z_t, \quad \log \sigma_t^2 = \phi_0 + \phi_1 \log \sigma_{t-1}^2 + \varepsilon_t, \ \{\varepsilon_t\} \sim \mathrm{IID\,N}(0,\sigma^2)$$

Main question:

What intrinsic features in the data (*if any*) can be used to discriminate between these two models?

Regular variation — multivariate case

Multivariate regular variation of **X**$=(X_1, \ldots, X_m)$: There exists a random vector $\theta \in S^{m-1}$ such that

$$P(|\mathbf{X}|> t\,x,\ \mathbf{X}/|\mathbf{X}| \in \bullet\ )/P(|\mathbf{X}|>t) \rightarrow_v x^{-\alpha} P(\ \theta \in \bullet\ )$$

($\rightarrow_v$ vague convergence on S^{m-1}, unit sphere in R^m) .

- $P(\ \theta \in \bullet)$ is called the spectral measure
- α is the index of **X**.

Equivalence:

$$\frac{P(\mathbf{X} \in t\bullet)}{P(|\mathbf{X}|>t)} \rightarrow_v \mu(\bullet)$$

μ is a measure on R^m which satisfies for $x > 0$ and A bounded away from 0,

$$\mu(xA) = x^{-\alpha} \mu(A).$$

Regular variation — multivariate case (cont)

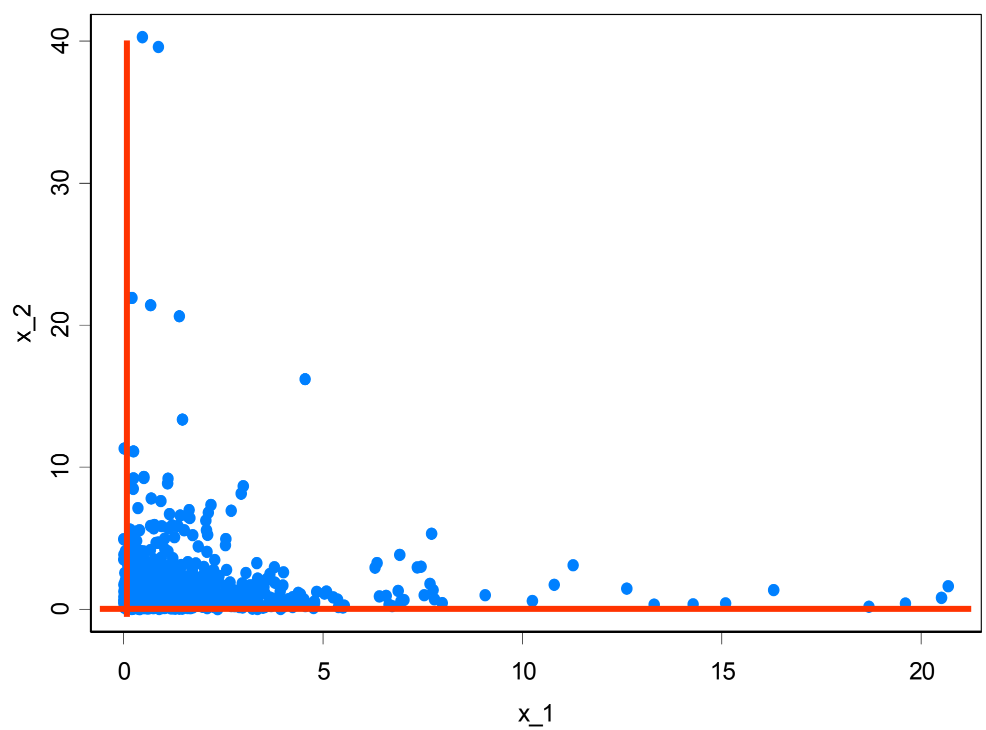
Examples:

1. If $X_1 > 0$ and $X_2 > 0$ are iid RV(α), then **X**= (X_1, X_2) is multivariate regularly varying with index α and *spectral distribution*

$$P(\theta = (0,1)) = P(\theta = (1,0)) = .5$$ (mass on axes).

Interpretation: Unlikely that X_1 and X_2 are very large at the same time.

Figure: plot of (X_{t1}, X_{t2}) for realization of 10,000.



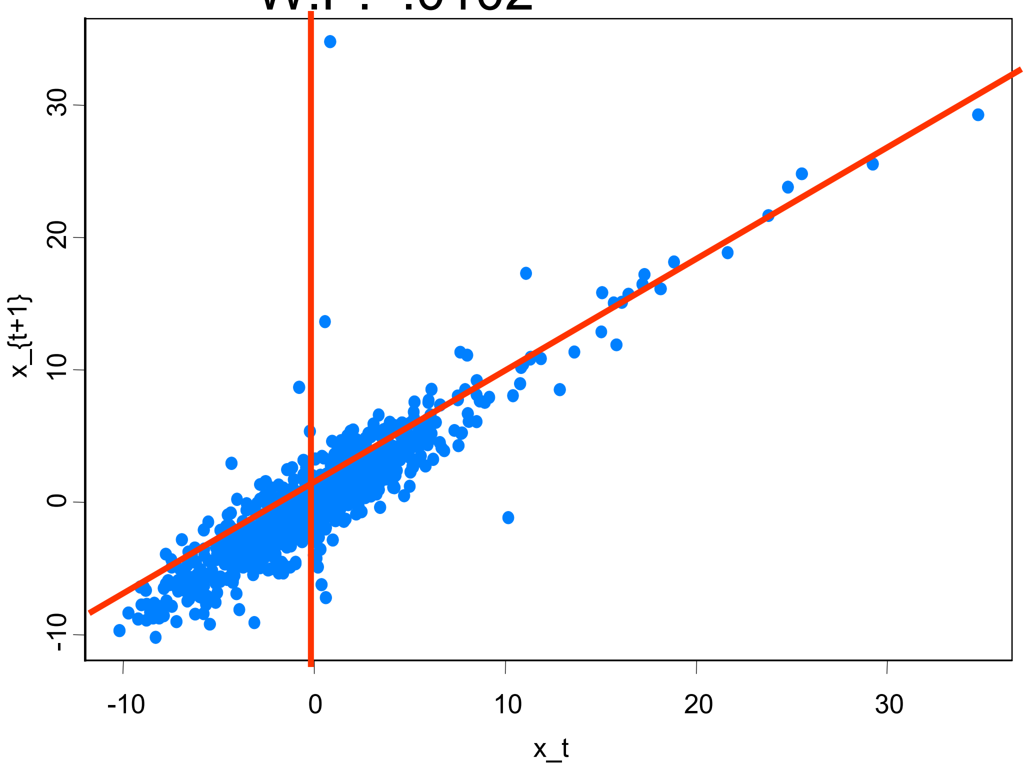
2. If $X_1 = X_2 > 0$, then $\mathbf{X} = (X_1, X_2)$ is multivariate regularly varying with index α and *spectral distribution*

$$P(\theta = (1/\sqrt{2}, 1/\sqrt{2})) = 1.$$

3. AR(1): $X_t = .9\, X_{t-1} + Z_t$, $\{Z_t\}$~IID symmetric stable (1.8)

$$\text{Distr of } \theta: \begin{cases} \pm(1,.9)/\text{sqrt}(1.81), & \text{W.P. } .9898 \\ \pm(0,1), & \text{W.P. } .0102 \end{cases}$$

Figure: scatter plot of (X_t, X_{t+1}) for realization of 10,000



Applications of multivariate regular variation (cont)

Linear combinations:

X ~RV(α) $\Rightarrow$ all linear combinations of **X** are regularly varying

i.e., there exist α and slowly varying fcn L(.), s.t.

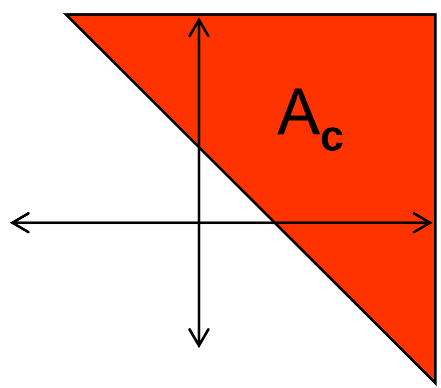
$$P(\mathbf{c}^T\mathbf{X} > t)/(t^{-\alpha}L(t)) \to w(\mathbf{c}),$$ exists for all real-valued **c**,

where

$$w(t\mathbf{c}) = t^{-\alpha}w(\mathbf{c}).$$

Use vague convergence with $A_{\mathbf{c}}=\{\mathbf{y}: \mathbf{c}^T\mathbf{y} > 1\}$, i.e.,

$$\frac{P(\mathbf{X}\in tA_{\mathbf{c}})}{t^{-\alpha}L(t)} = \frac{P(\mathbf{c}^T\mathbf{X} > t)}{P(|\mathbf{X}| > t)} \to \mu(A_{\mathbf{c}}) =: w(\mathbf{c}),$$



where $t^{-\alpha}L(t) = P(|\mathbf{X}| > t)$.

Applications of multivariate regular variation (cont)

Converse?

X ~RV(α) $\Leftarrow$ all linear combinations of **X** are regularly varying?

There exist α and slowly varying fcn $L(.)$, s.t.

(LC) $P(\mathbf{c}^T\mathbf{X} > t)/(t^{-\alpha}L(t)) \to w(\mathbf{c})$, exists for all real-valued **c.**

Theorem (Basrak, Davis, Mikosch, `02). Let **X** be a random vector.

1. If **X** satisfies (LC) with α non-integer, then **X** is RV(α).
2. If **X** > 0 satisfies (LC) for non-negative **c** and α is non-integer, then **X** is RV(α).
3. If **X** > 0 satisfies (LC) with α an odd integer, then **X** is RV(α).

Applications of multivariate regular variation (cont)

There exist α and slowly varying fcn $L(.)$, s.t.

(LC) $P(\mathbf{c}^T\mathbf{X} > t)/(t^{-\alpha}L(t)) \rightarrow w(\mathbf{c})$, exists for all real-valued **c.**

1. If **X** satisfies (LC) with α non-integer, then **X** is RV(α).
2. If **X** > 0 satisfies (LC) for non-negative **c** and α is non-integer, then **X** is RV(α).
3. If **X** > 0 satisfies (LC) with α an odd integer, then **X** is RV(α).

Remarks:

- 1 cannot be extended to integer α (Hult and Lindskog `05)
- 2 cannot be extended to integer α (Hult and Lindskog `05)
- 3 can be extended to even integers (Lindskog et al. `07, under review).

Applications of theorem

1. Kesten (1973). Under general conditions, (LC) holds with L(t)=1 for stochastic recurrence equations of the form

$$\mathbf{Y}_t = \mathbf{A}_t \mathbf{Y}_{t-1} + \mathbf{B}_t, \quad (\mathbf{A}_t, \mathbf{B}_t) \sim \text{IID},$$

$\mathbf{A}_t$ $d \times d$ random matrices, $\mathbf{B}_t$ random d-vectors.

It follows that the distributions of $\mathbf{Y}_t$, and in fact all of the finite dim'l distrs of $\mathbf{Y}_t$ are regularly varying (no longer need α to be non-even).

2. GARCH processes. Since squares of a GARCH process can be embedded in a SRE, the *finite dimensional distributions* of a *GARCH* are regularly varying.

Examples

Example of ARCH(1): $X_t=(\alpha_0+\alpha_1 X^2_{t-1})^{1/2}Z_t$, $\{Z_t\}$~IID.

α found by solving $E|\alpha_1 Z^2|^{\alpha/2} = 1$.

α_1	.312	.577	1.00	1.57
α	8.00	4.00	2.00	1.00

Distr of θ:

$$P(\theta \in \bullet) = E\{\|(B,Z)\|^\alpha I(\arg((B,Z)) \in \bullet)\} / E\|(B,Z)\|^\alpha$$

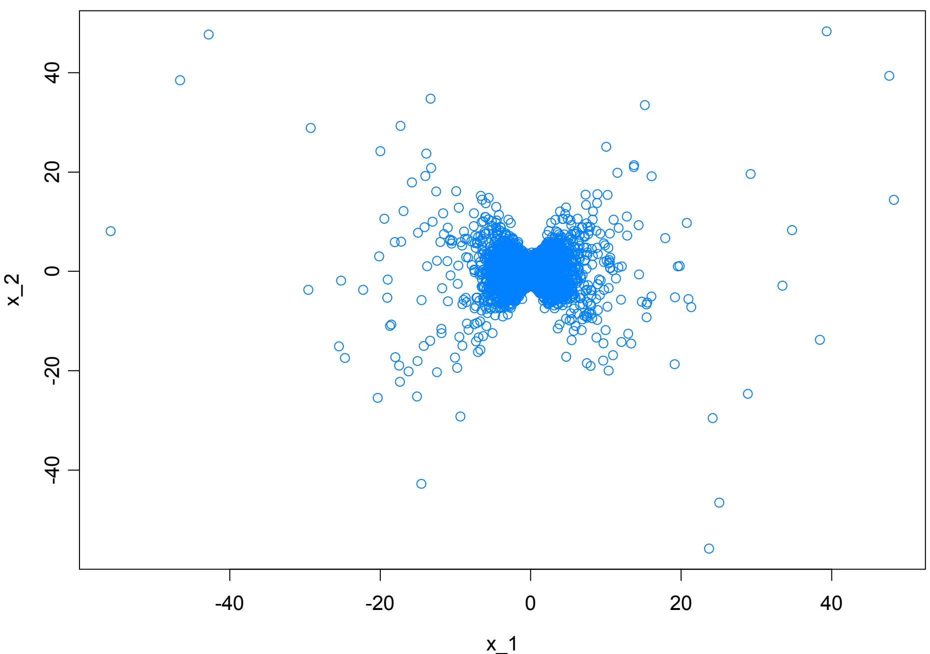
where

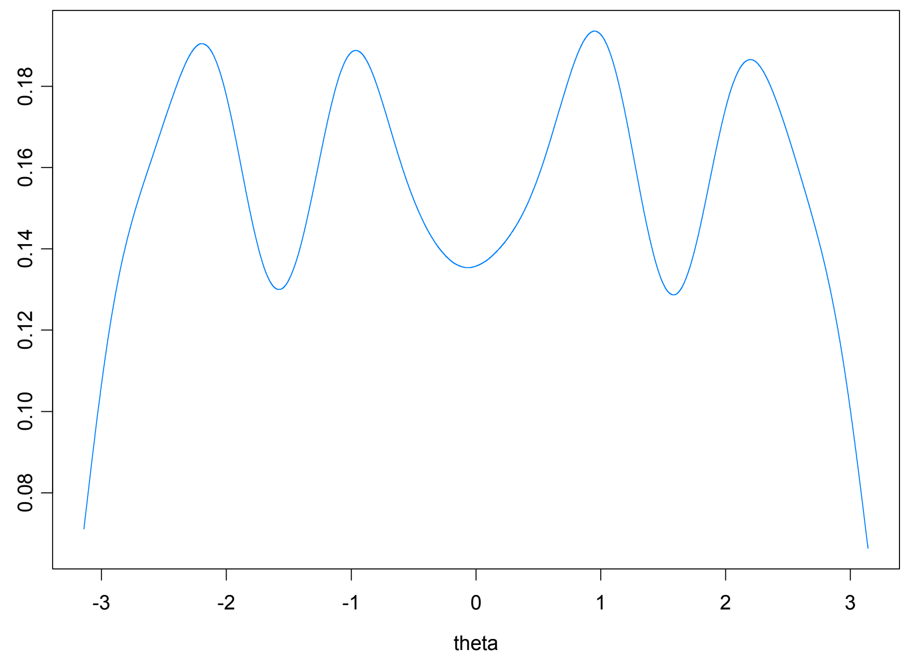
$$P(B = 1) = P(B = -1) = .5$$

Examples (cont)

Example of ARCH(1): $\alpha_0=1$, $\alpha_1=1$, $\alpha=2$, $X_t=(\alpha_0+\alpha_1 X^2_{t-1})^{1/2}Z_t$, $\{Z_t\}$~IID

Figures: plots of (X_t, X_{t+1}) and estimated distribution of θ for realization of 10,000.



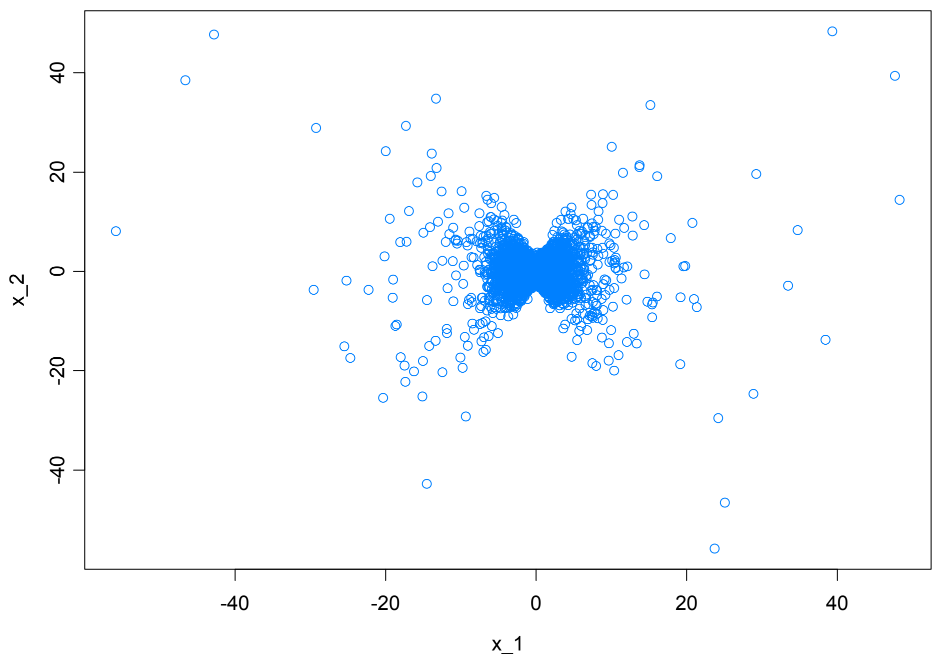


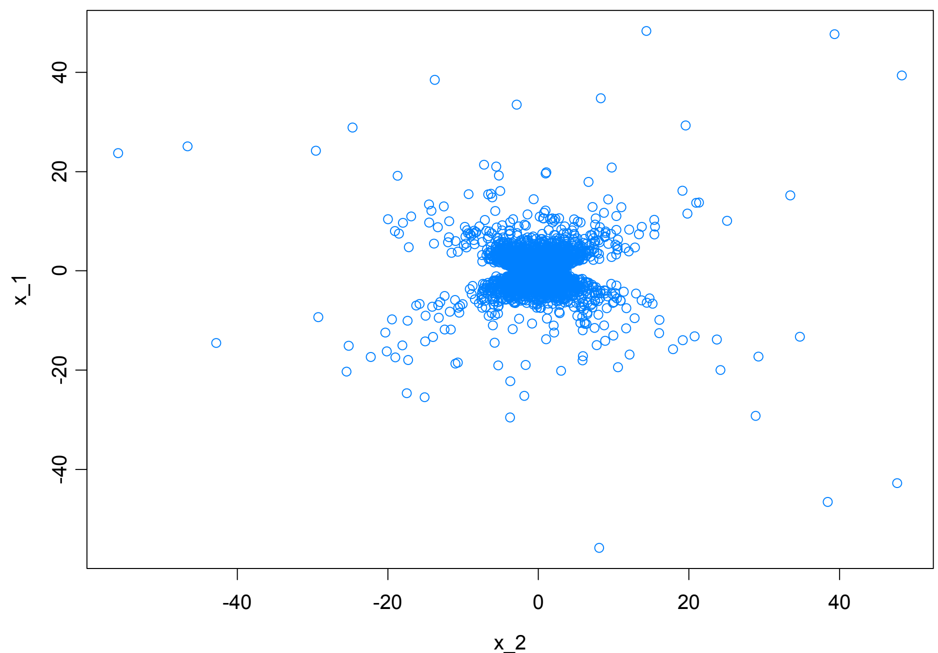
Examples (cont)

Example of ARCH(1): $\alpha_0=1$, $\alpha_1=1$, $\alpha=2$, $X_t=(\alpha_0+\alpha_1 X^2_{t-1})^{1/2}Z_t$, $\{Z_t\}$~IID

Is this process time-reversible?

Figures: plots of (X_t, X_{t+1}) and (X_{t+1}, X_t) imply *non-reversibility.*





Examples (cont)

Example: SV model $X_t = \sigma_t Z_t$

Suppose $Z_t \sim RV(\alpha)$ and

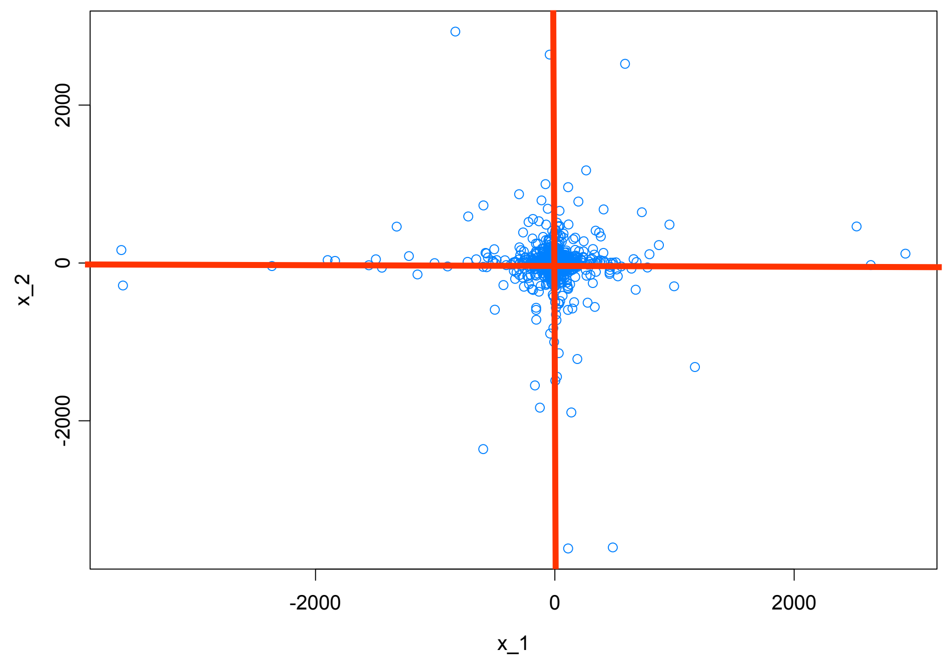
$$X_t = \sigma_t Z_t, \quad \log \sigma_t^2 = \phi_0 + \phi_1 \log \sigma_{t-1}^2 + \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{IID N}(0, \sigma^2)$$

Then $\mathbf{Z}_n = (Z_1, \ldots, Z_n)'$ is regulary varying with index α and so is

$$\mathbf{X}_n = (X_1, \ldots, X_n)' = \text{diag}(\sigma_1, \ldots, \sigma_n)\, \mathbf{Z}_n$$

with spectral distribution concentrated on $(\pm 1, 0)$, $(0, \pm 1)$.

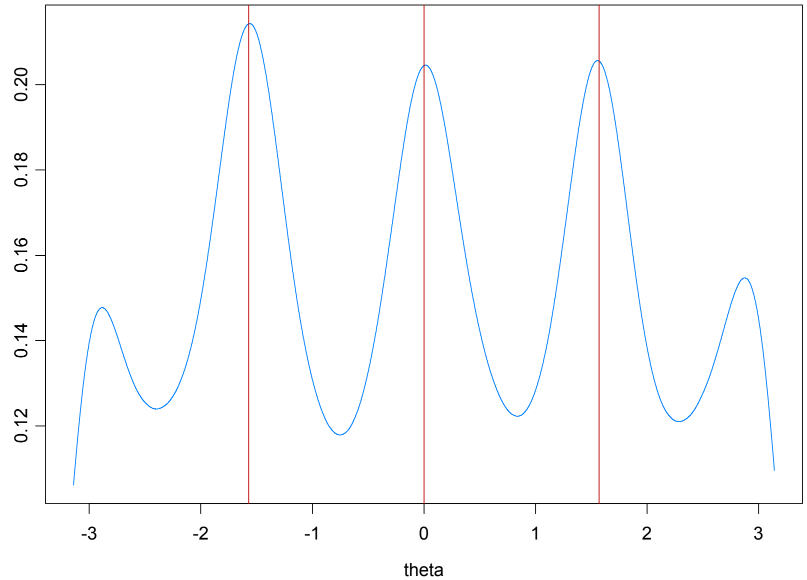
Figure: plot of (X_t, X_{t+1}) for realization of 10,000.

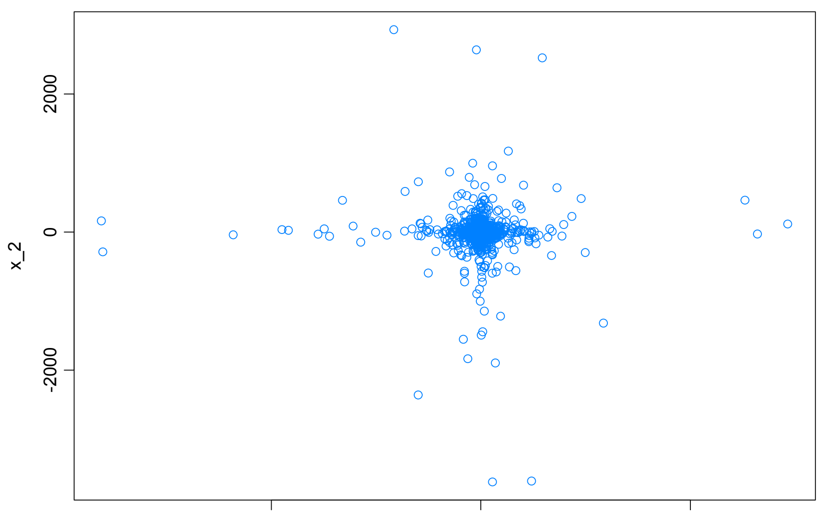


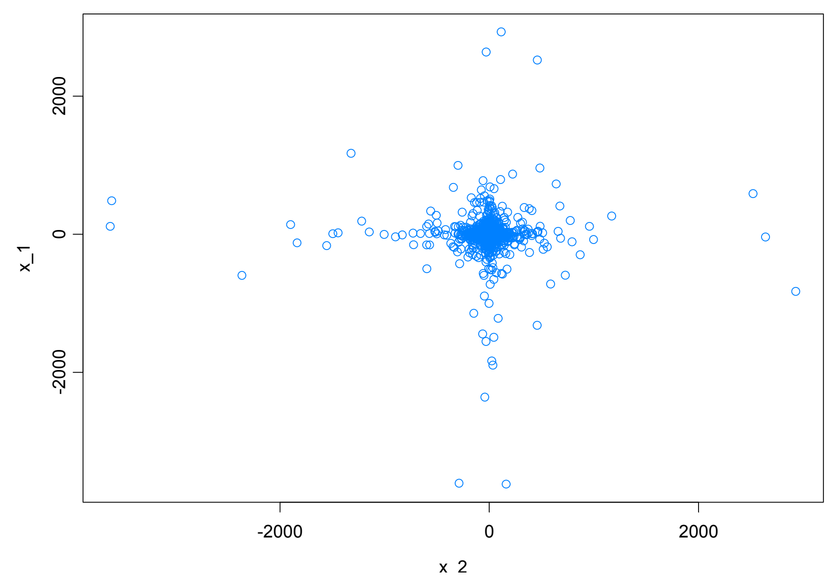
Examples (cont)

Example: SV model $X_t = \sigma_t Z_t$

- SV processes are time-reversible if log-volatility is Gaussian.
- Asymptotically time-reversible if log-volatility is nonGaussian







Extremes for GARCH and SV processes

Setup

- $X_t = \sigma_t Z_t$, $\{Z_t\} \sim$ IID (0,1)
- X_t is RV (α)
- Choose $\{b_n\}$ s.t. $nP(X_t > b_n) \to 1$

Then

$$P^n(b_n^{-1} X_1 \le x) \to \exp\{-x^{-\alpha}\}.$$

Then, with $M_n = \max\{X_1, \ldots, X_n\}$,

(i) GARCH:

$$P(b_n^{-1} M_n \le x) \to \exp\{-\gamma x^{-\alpha}\},$$

γ is extremal index ($0 < \gamma < 1$).

(ii) SV model:

$$P(b_n^{-1} M_n \le x) \to \exp\{-x^{-\alpha}\},$$

extremal index $\gamma = 1$ no clustering.

Extremes for GARCH and SV processes (cont)

(i) GARCH: $P(b_n^{-1}M_n \le x) \to \exp\{-\gamma x^{-\alpha}\}$

(ii) SV model: $P(b_n^{-1}M_n \le x) \to \exp\{-x^{-\alpha}\}$

Remarks about extremal index.

(i) $\gamma < 1$ implies clustering of exceedances

(ii) Numerical example. Suppose c is a threshold such that

$$P^n(b_n^{-1}X_1 \le c) \sim .95$$

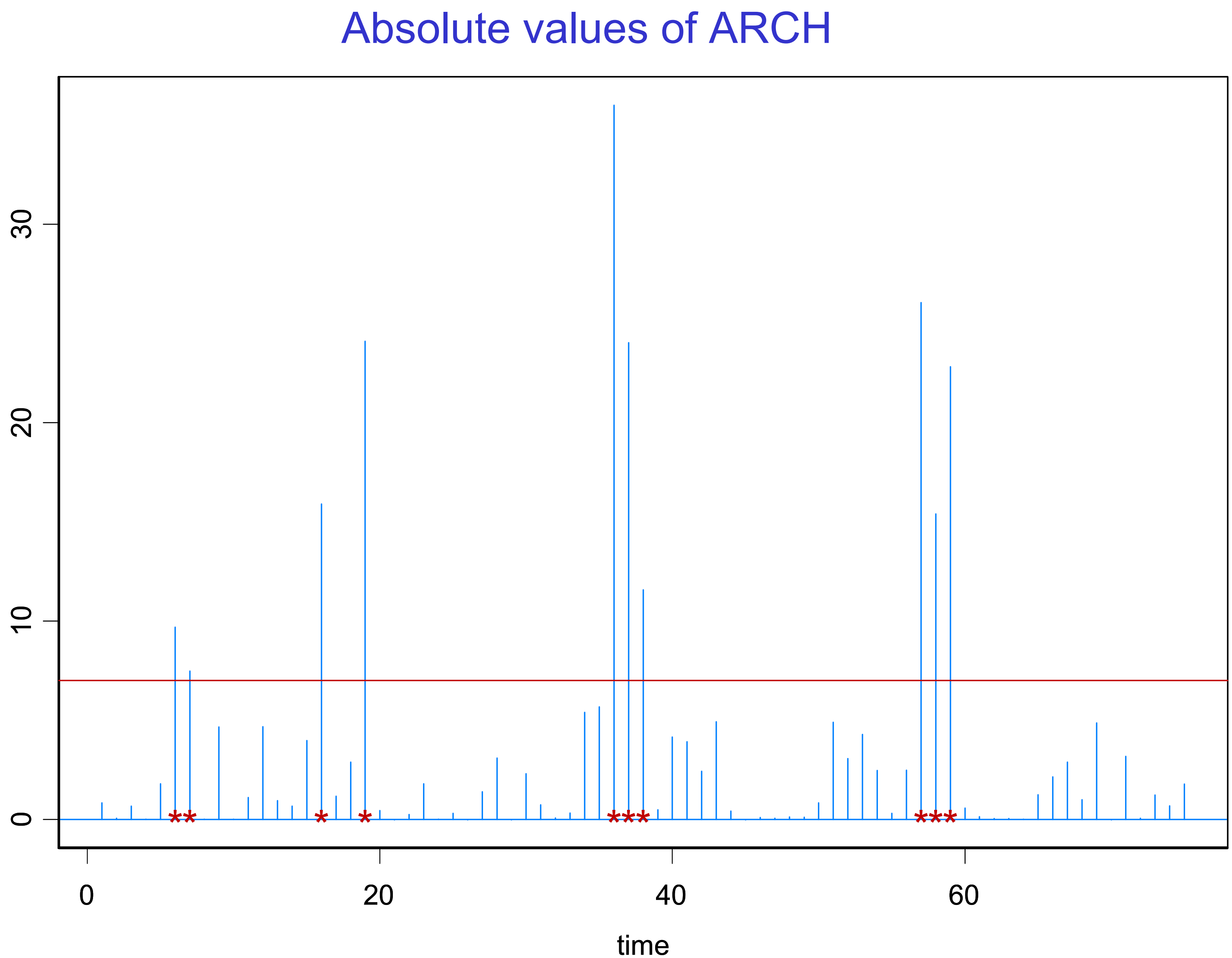
Then, if $\gamma = .5$, $P(b_n^{-1}M_n \le c) \sim (.95)^{.5} = .975$

(iii) $1/\gamma$ is the *mean cluster size* of exceedances.

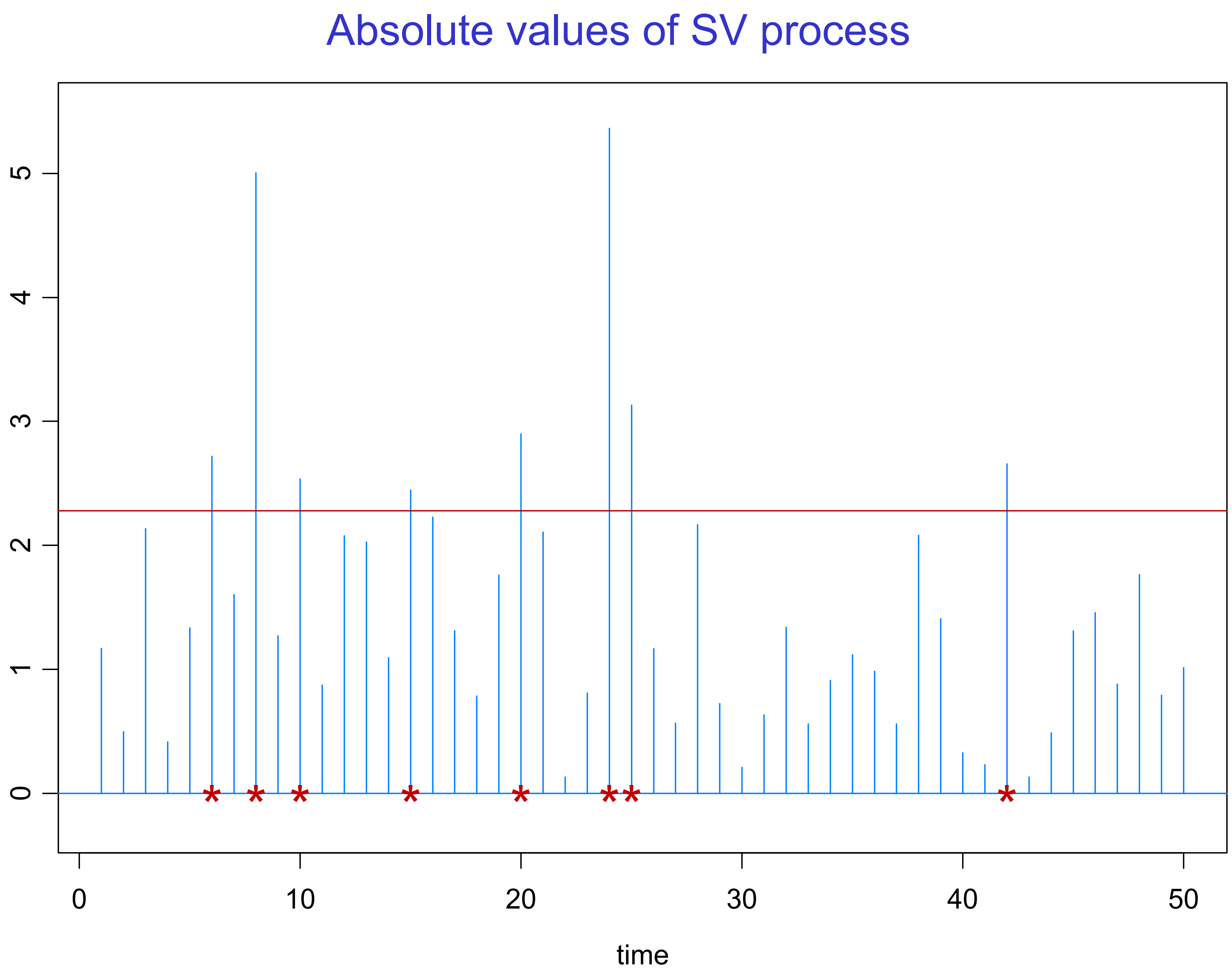
(iv) Use γ to *discriminate* between GARCH and SV models.

(v) Even for the light-tailed SV model (i.e., $\{Z_t\}$ ~IID N(0,1), the *extremal index* is 1 (see Breidt and Davis `98)

Extremes for GARCH and SV processes (cont)



Extremes for GARCH and SV processes (cont)



Summary of results for ACF of GARCH(p,q) and SV models

GARCH(p,q)

$\alpha \in (0,2)$:

$$(\hat{\rho}_X(h))_{h=1,\ldots,m} \xrightarrow{d} (V_h / V_0)_{h=1,\ldots,m},$$

$\alpha \in (2,4)$:

$$\left(n^{1-2/\alpha}\hat{\rho}_X(h)\right)_{h=1,\ldots,m} \xrightarrow{d} \gamma_X^{-1}(0)\left(V_h\right)_{h=1,\ldots,m}.$$

$\alpha \in (4,\infty)$:

$$\left(n^{1/2}\hat{\rho}_X(h)\right)_{h=1,\ldots,m} \xrightarrow{d} \gamma_X^{-1}(0)\left(G_h\right)_{h=1,\ldots,m}.$$

Remark: Similar results hold for the sample ACF based on $|X_t|$ and X_t^2.

Summary of results for ACF of GARCH(p,q) and SV models (cont)

SV Model

$\alpha \in (0,2)$:

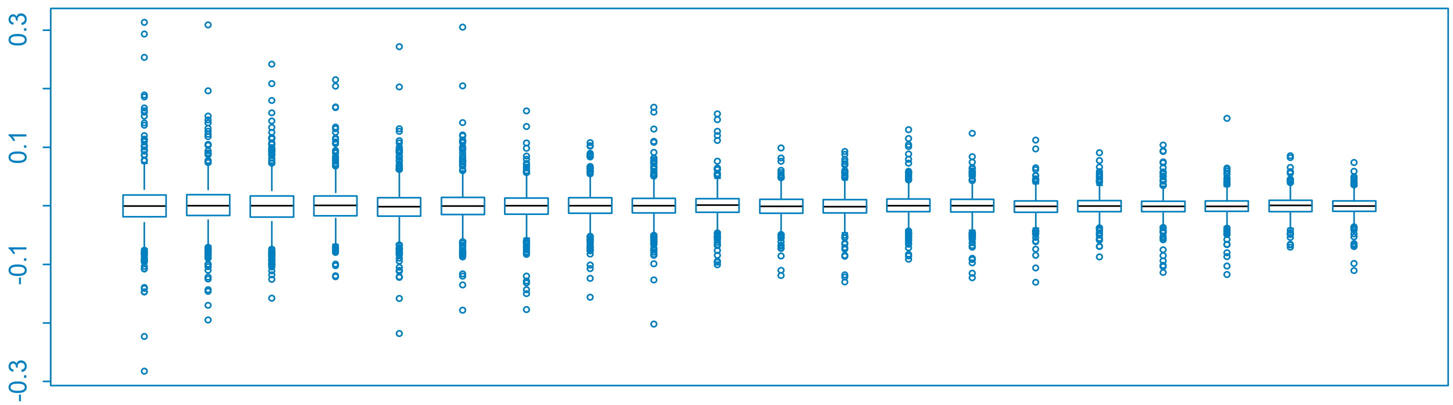
$$(n/\ln n)^{1/\alpha}\hat{\rho}_X(h) \xrightarrow{d} \frac{\|\sigma_1\sigma_{h+1}\|_\alpha}{\|\sigma_1\|_\alpha^2}\frac{S_h}{S_0}.$$

$\alpha \in (2, \infty)$:

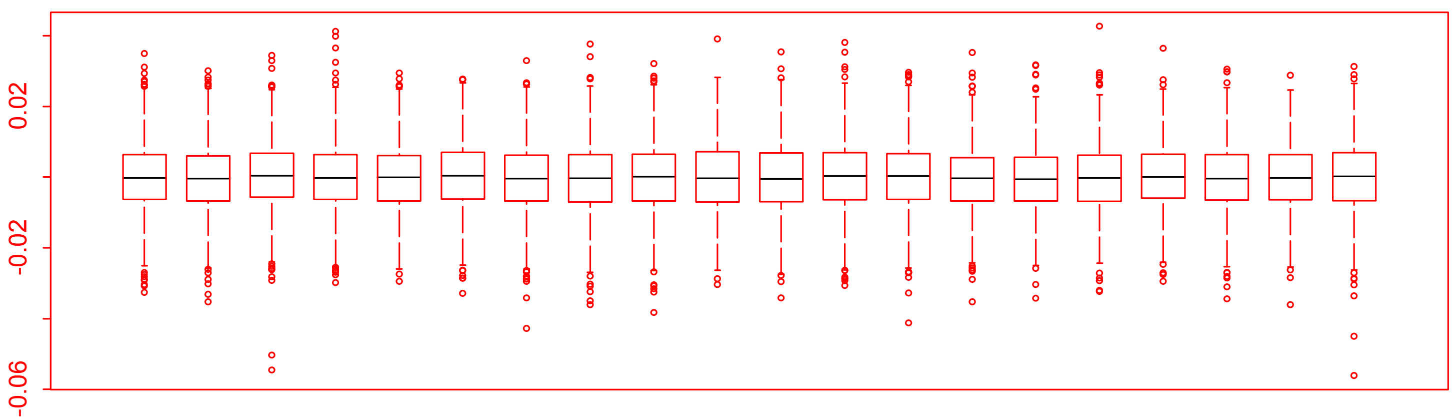
$$\left(n^{1/2}\hat{\rho}_X(h)\right)_{h=1,\ldots,m} \xrightarrow{d} \gamma_X^{-1}(0)\left(G_h\right)_{h=1,\ldots,m}.$$

Sample ACF for GARCH and SV Models (1000 reps)

(a) GARCH(1,1) Model, n=10000

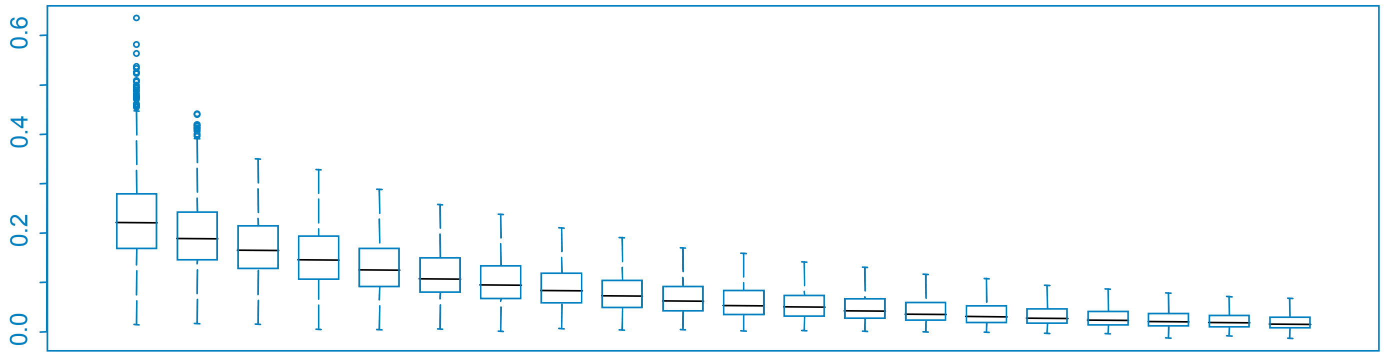


(b) SV Model, n=10000

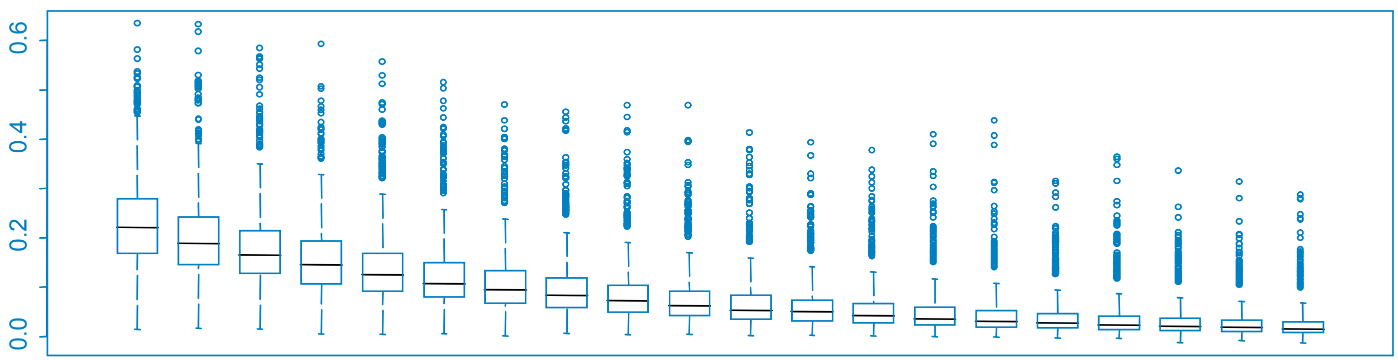


Sample ACF for Squares of GARCH (1000 reps)

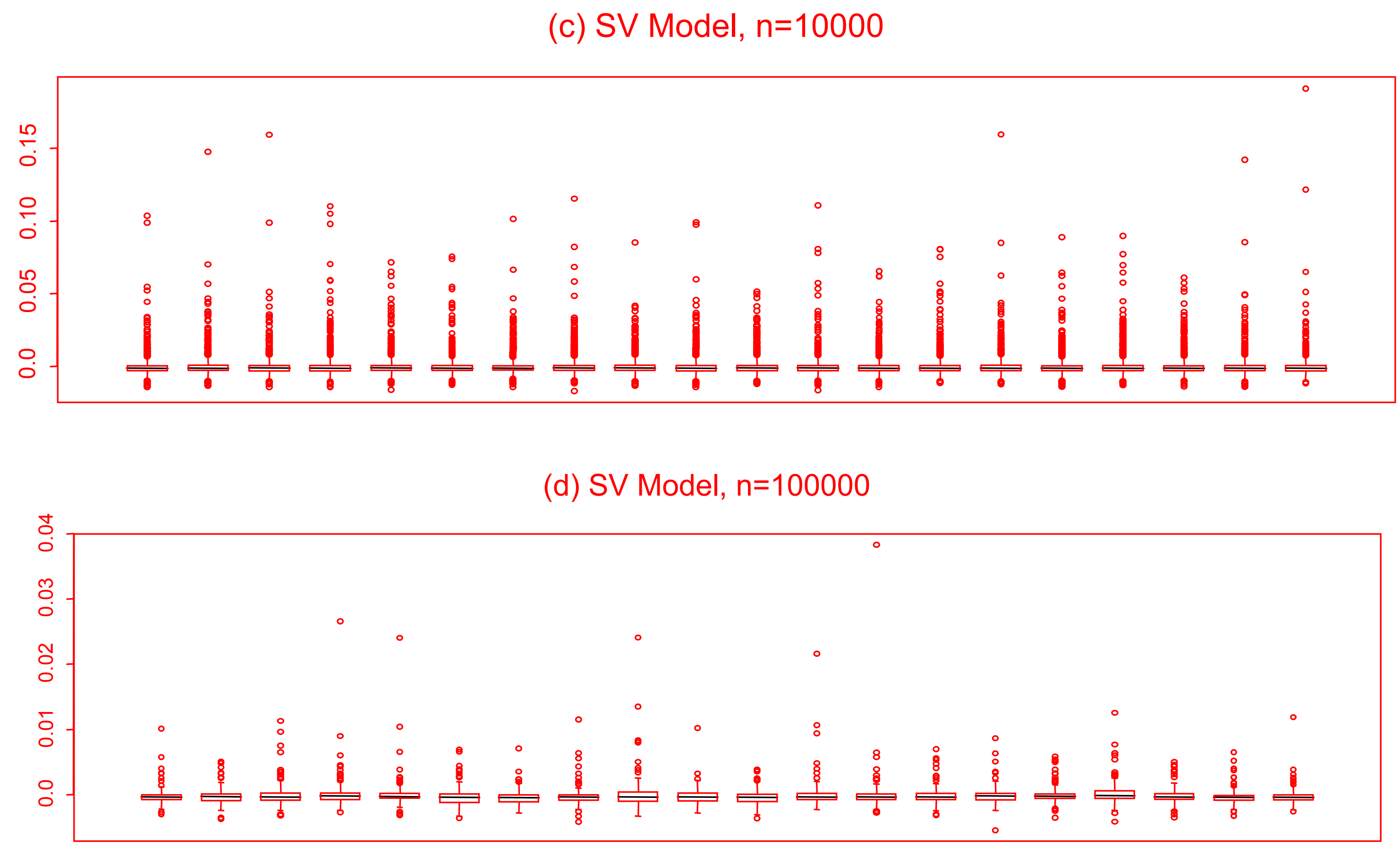
(a) GARCH(1,1) Model, n=10000



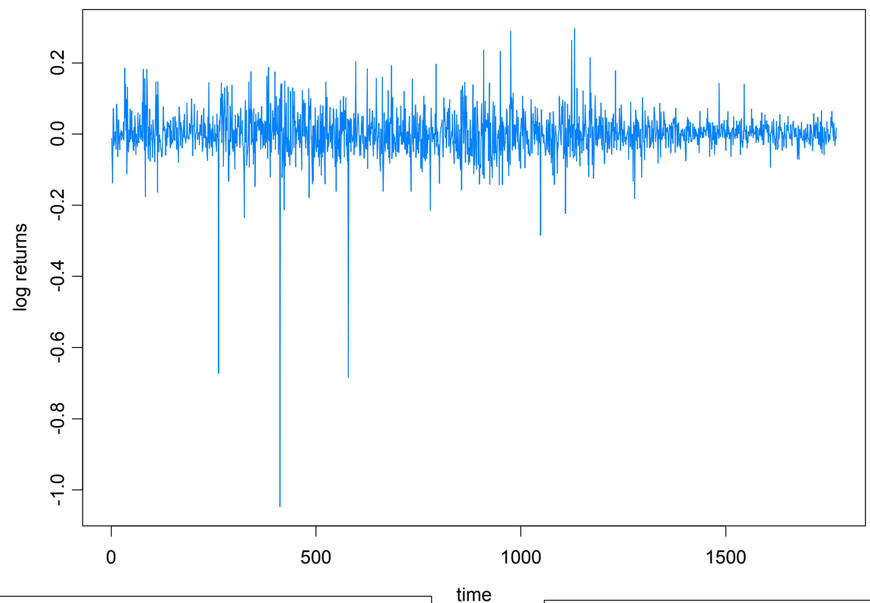
b) GARCH(1,1) Model, n=100000

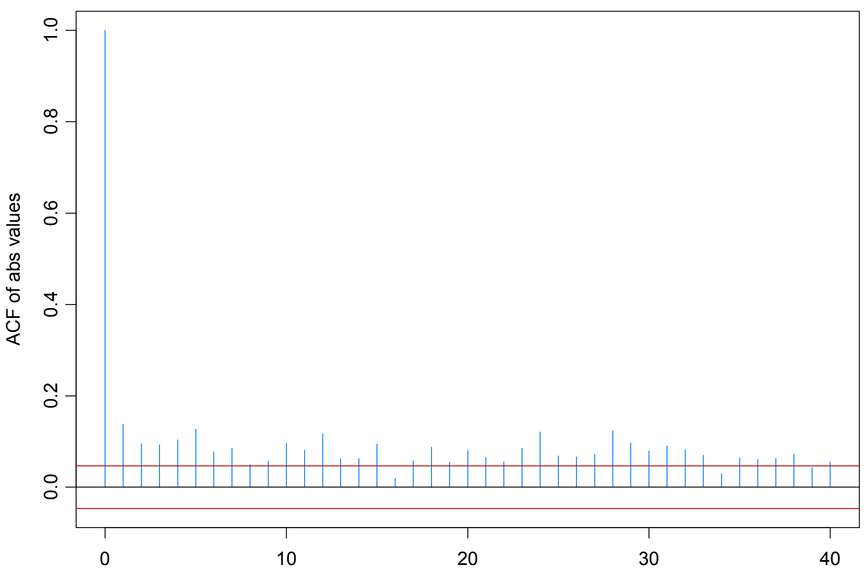


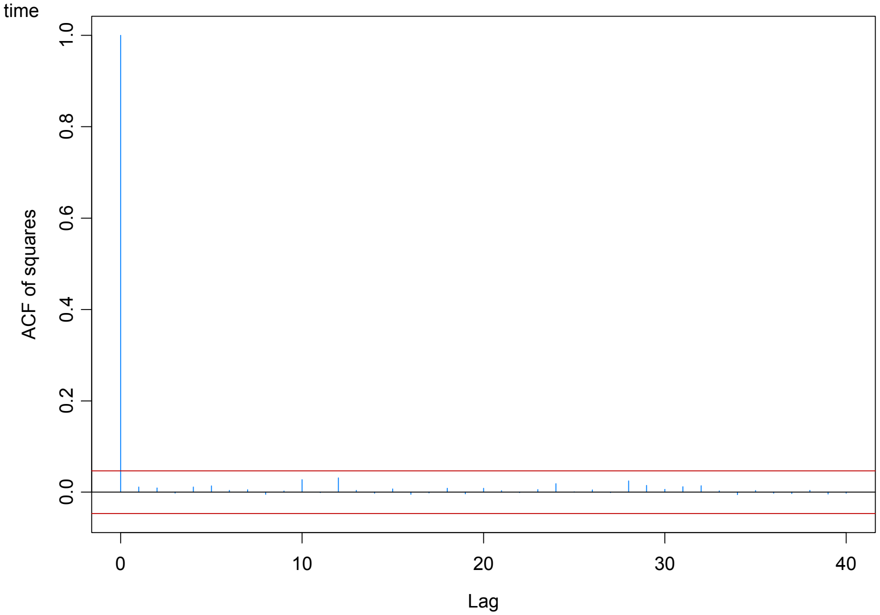
Sample ACF for Squares of SV (1000 reps)



Example: Amazon-returns (May 16, 1997 – June 16, 2004)





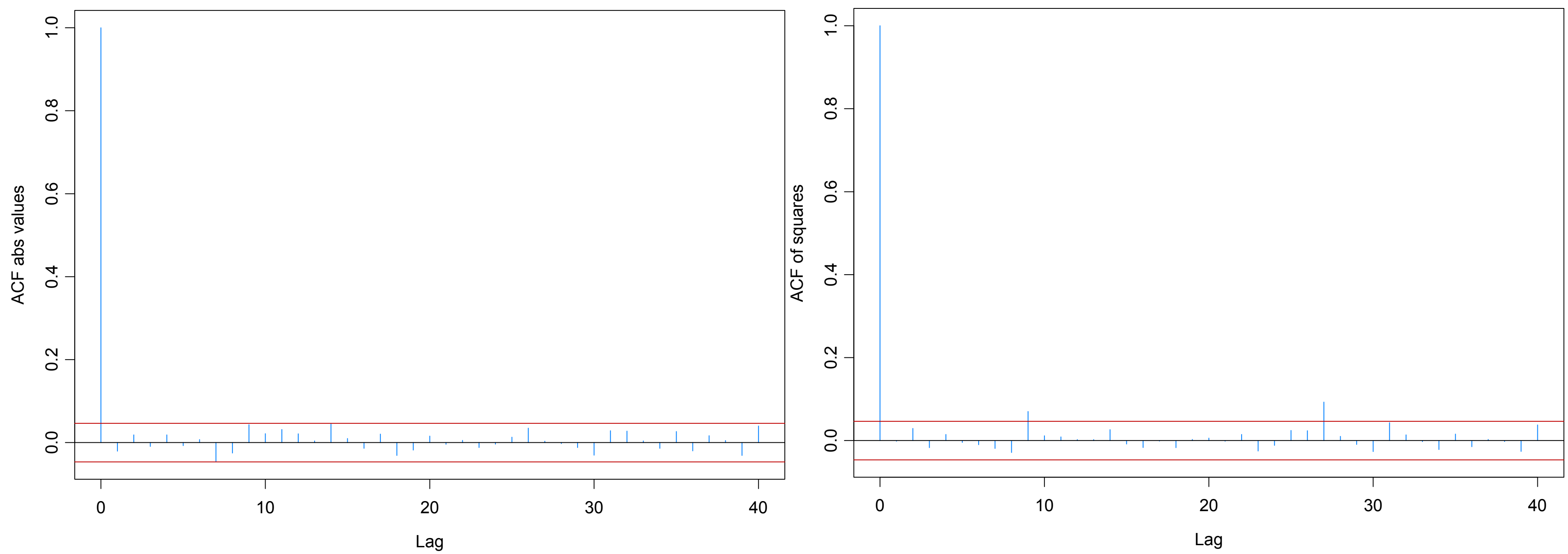


Amazon returns (GARCH model)

GARCH(1,1) model fit to Amazon returns:

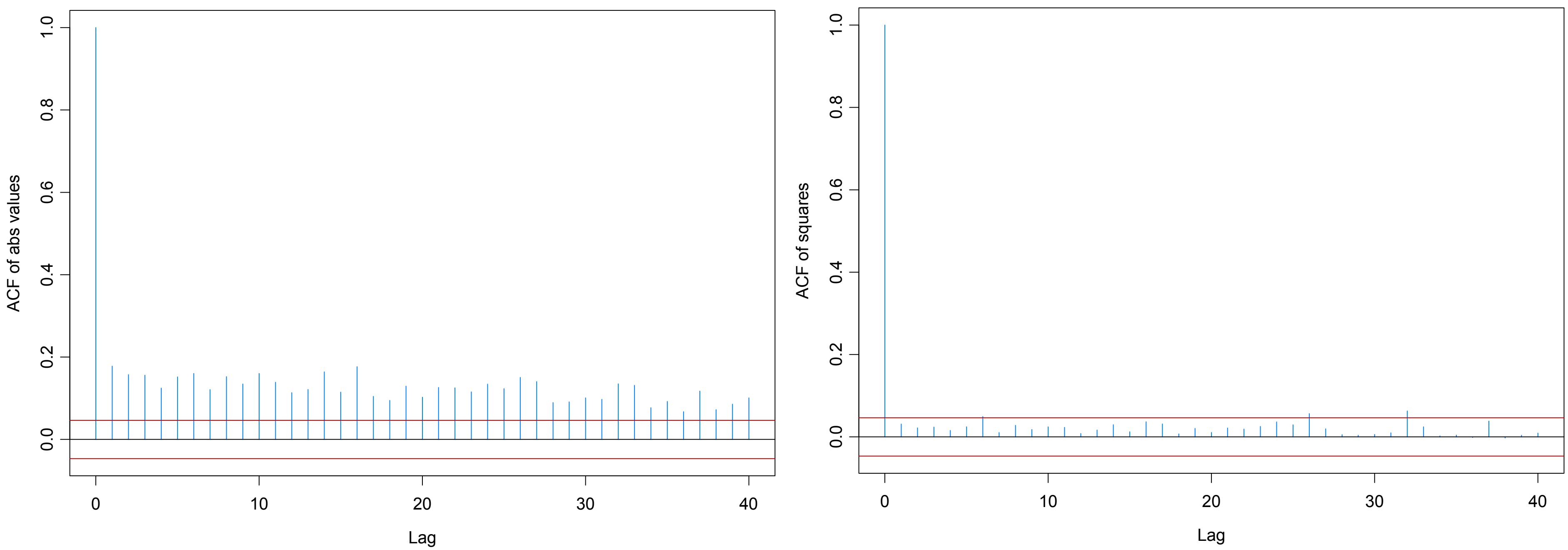
$\alpha_0 = .00002493$, $\alpha_1 = .0385$, $\beta_1 = .957$, $X_t = (\alpha_0 + \alpha_1 X^2_{t-1})^{1/2} Z_t$, $\{Z_t\} \sim \text{IID } t(3.672)$

Simulation from GARCH(1,1) model



Amazon returns (SV model)

Stochastic volatility model fit to Amazon returns:



Wrap-up

- *Regular variation* is a flexible tool for modeling both *dependence* and *tail heaviness.*
- Useful for establishing *point process convergence* of heavy-tailed time series.
- *Extremal index* $\gamma < 1$ for GARCH and $\gamma = 1$ for SV.
- ACF has faster convergence for SV.