Linear Time Series With Nonlinear Behavior

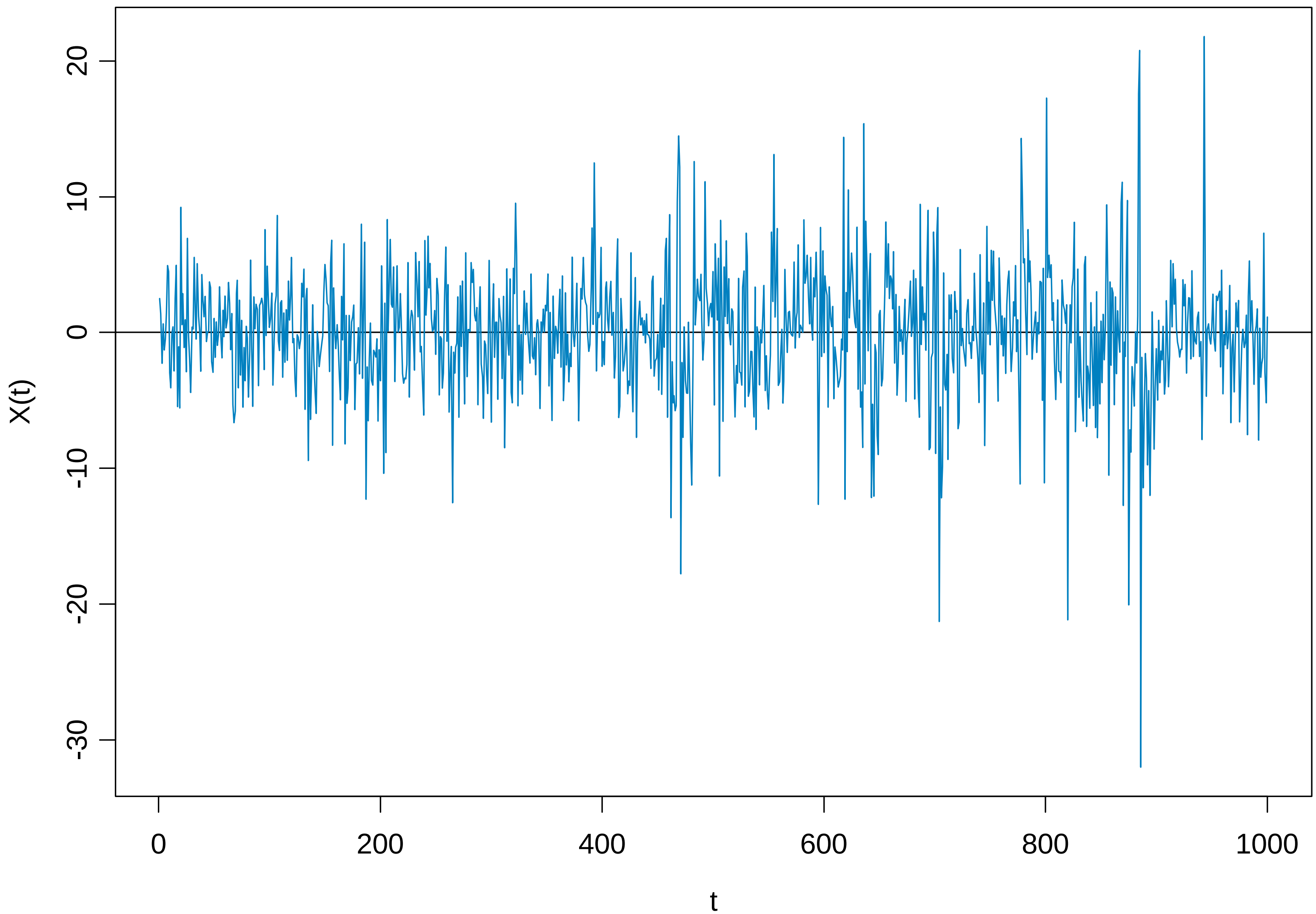
F. Jay Breidt, Iowa State University

Richard A. Davis, Colorado State University

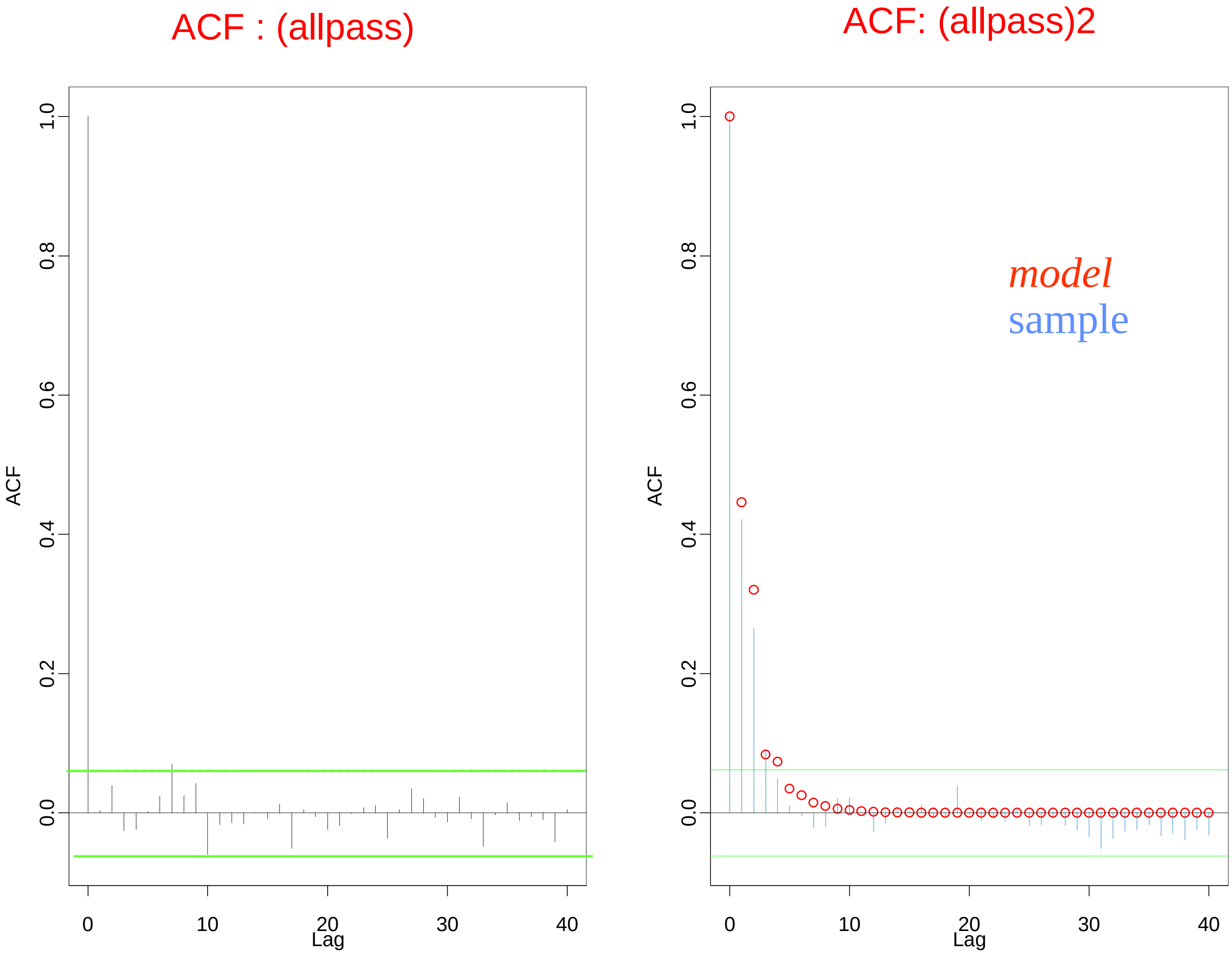
Alex Trindade, Colorado State University

- ☞ All-pass models
  - properties
  - approximation to likelihood
- ☞ Asymptotic theory
  - LAD
- ☞ Order selection
  - PACF
  - AIC
- ☞ Simulation results
- ☞ Application to financial data
- ☞ Application to non-causal AR fitting
  - volumes of MSFT stock

Realization of a Time Series Model



All-pass model of order 2 (t3 noise)



All-pass Models

Causal AR polynomial: $\phi(z)=1-\phi_1 z - \cdots - \phi_p z^p$, $\phi(z) \neq 0$ for $|z| \leq 1$.

Define MA polynomial:

$$\theta(z) = -z^p \phi(z^{-1})/\phi_p = -(z^p - \phi_1 z^{p-1} - \cdots - \phi_p)/\phi_p$$

$\neq 0$ for $|z| \geq 1$ (MA polynomial is non-invertible).

Model for data $\{X_t\}$: $\phi(B)X_t = \theta(B) Z_t$, $\{Z_t\} \sim \text{IID}$ (non-Gaussian)

Examples:

All-pass(1): $X_t - \phi X_{t-1} = Z_t - \phi^{-1} Z_{t-1}$, $|\phi| < 1$.

All-pass(2): $X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = Z_t + \phi_1/\phi_2 Z_{t-1} - 1/\phi_2 Z_{t-2}$

Properties:

- uncorrelated (flat spectrum)

$$f_X(\omega) = \frac{\left|e^{-ip\omega}\right|^2 \left|\phi(e^{i\omega})\right|^2}{\phi_p^2 \left|\phi(e^{-i\omega})\right|^2} \frac{\sigma^2}{2\pi} = \frac{\sigma^2}{\phi_p^2 \, 2\pi}$$

- data are dependent if noise is non-Gaussian (e.g. Breidt & Davis `91).
- squares and absolute values are correlated.
- X_t is heavy-tailed if noise is heavy-tailed.

Approximating the likelihood

Data: $(X_1, \ldots, X_n)$

Model: $X_t = \phi_{01}X_{t-1} + \cdots + \phi_{0p}X_{t-p}$

$$+ (Z_{t-p} - \phi_{01}Z_{t-p+1} - \cdots - \phi_{0p}Z_t)/\phi_{0r}$$

where ϕ_{0r} is the last non-zero coefficient among the ϕ_{0j}'s.

Noise: $z_{t-p} = \phi_{01}z_{t-p+1} + \cdots + \phi_{0p}z_t - (X_t - \phi_{01}X_{t-1} - \cdots - \phi_{0p}X_{t-p})$,

where $z_t = Z_t / \phi_{0r}$.

More generally define,

$$z_{t-p}(\phi) = \begin{cases} 0, & \text{if } t = n+p, \ldots, n+1, \\ \phi_1 z_{t-p+1}(\phi) + \cdots + \phi_p z_t(\phi) - \phi(B)X_t, & \text{if } t = n, \ldots, p+1. \end{cases}$$

Note: $z_t(\phi_0)$ is a close approximation to z_t (initialization error)

Assume that Z_t has density function f_σ and consider the vector

$$\mathbf{z} = (\underbrace{X_{1-p},\ldots, X_0, z_{1-p}(\phi),\ldots, z_0(\phi)}, \underbrace{z_1(\phi),\ldots, z_{n-p+1}(\phi),\ldots, z_n(\phi)})'$$

independent pieces

Joint density of **z**:

$$h(\mathbf{z}) = h_1(X_{1-p},\ldots, X_0, z_{1-p}(\phi),\ldots, z_0(\phi))$$

$$\bullet\left(\prod_{t=1}^{n-p} f_\sigma(\phi_q z_t(\phi))\,|\,\phi_q\,|\right) h_2(z_{n-p+1}(\phi),\ldots, z_n(\phi)),$$

and hence the joint density of the data can be approximated by

$$h(\mathbf{x}) = \left(\prod_{t=1}^{n-p} f_\sigma(\phi_q z_t(\phi))\,|\,\phi_q\,|\right)$$

where $q=\max\{0 \le j \le p: \phi_j \ne 0\}$.

Log-likelihood:

$$L(\phi, \sigma) = -(n-p)\ln(\sigma / |\phi_q|) + \sum_{t=1}^{n-p} \ln f(\sigma^{-1}\phi_q z_t(\phi))$$

where $f_\sigma(z) = \sigma^{-1} f(z/\sigma)$.

Least absolute deviations: choose Laplace density

$$f(z) = \frac{1}{\sqrt{2}} \exp(-\sqrt{2}\,|z|)$$

and log-likelihood becomes

$$\text{constant} - (n-p)\ln\kappa - \sum_{t=1}^{n-p} \sqrt{2}\,|z_t(\phi)| / \kappa$$

Concentrated Laplacian likelihood

$$l(\phi) = \text{constant} - (n-p)\ln \sum_{t=1}^{n-p} |z_t(\phi)|$$

Maximizing $l(\phi)$ is equivalent to minimizing the absolute deviations

$$m_n(\phi) = \sum_{t=1}^{n-p} |z_t(\phi)|.$$

Asymptotic Results:

Theorem 1. Let $\{Y_t\}$ be the linear process

$$Y_t = \sum_{j=-\infty}^{\infty} c_j z_{t-j},$$

where $c_0=0$, $\sum_{j=-\infty}^{\infty} |c_j| < \infty$, $\{z_t\} \sim \text{IID}(0,\sigma^2)$, $\text{median}(z_1)=0$, $g(0)>0$ (g density of z_1). Then

$$S_n = \sum_{t=1}^{n-p} \left(| z_t - n^{-1/2} Y_t | - | z_t | \right)$$

$$\to Var(Y_1) g(0) + N$$

where $N \sim \text{N}(0, \gamma^*(0) + 2\sum_{h\geq 1} \gamma^*(h))$ and $\gamma^*(h)$ is the covariance function for $Y_t \, \text{sgn}(z_t)$

Key idea:

$$S_n = \sum_{t=1}^{n-p} \left(| z_t - n^{-1/2} Y_t | - | z_t | \right)$$

$$= -n^{-1/2} \sum_{t=1}^{n-p} Y_t \,\mathrm{sgn}(z_t)$$

$$+ 2 \sum_{t=1}^{n-p} (n^{-1/2} Y_t - z_t) \left\{ 1_{\{0 < z_t < n^{-1/2} Y_t\}} - 1_{\{n^{-1/2} Y_t < z_t < 0\}} \right\}$$

$$\to N + Var(Y_1) g(0)$$

Theorem 2. On C(R^p),

$$S_n(\mathbf{u}) = \sum_{t=1}^{n-p} \left(| z_t(\phi_0 + n^{-1/2}\mathbf{u}) | - | z_t(\phi_0) | \right)$$

$$\to S(\mathbf{u}),$$

where

$$S(\mathbf{u}) = \frac{f_\sigma(0)}{|\phi_{0r}|} \mathbf{u}'\Gamma_p \mathbf{u} + \mathbf{u}'\mathbf{N},$$

$$\mathbf{N} \sim N(\mathbf{0}, \frac{2Var(|Z_1|)}{\phi_{0r}^2 \sigma^2} \Gamma_p),$$

and Γ_p is the covariance matrix of a causal AR(p).

Limit theory for LAD estimate. Note that

$$\hat{\phi}_{\text{LAD}} = \phi_0 + \hat{\mathbf{u}}_n / \sqrt{n}$$

so that $\hat{\mathbf{u}}_n = \sqrt{n}(\hat{\phi}_{\text{LAD}} - \phi_0) = \arg\min S_n(\mathbf{u})$

$$\to \hat{\mathbf{u}} = \arg\min S(\mathbf{u}).$$

Minimizing S, we find that the minimizer or limit random variable is

$$\hat{\mathbf{u}}_n = \sqrt{n}(\hat{\phi}_{\text{LAD}} - \phi_0) \to -\frac{|\phi_{0r}|\Gamma_p^{-1}}{2 f_\sigma(0)}\mathbf{N}$$

$$-\frac{|\phi_{0r}|\Gamma_p^{-1}}{2 f_\sigma(0)}\mathbf{N} \sim N(\mathbf{0}, \frac{Var(|Z_1|)}{2\sigma^4 f_\sigma^2(0)}\sigma^2\Gamma_p^{-1})$$

Remarks.

1. Need $E|z(\phi)|$ to have a unique minimum at $\phi=\phi_0$. True if Z_1 has heavier tails than Gaussian,

$$E\,|\sum_{j=-\infty}^{\infty} c_j Z_j| > E\,|Z_1|$$

(See Jian and Pawitan (1998) for sufficient conditions.)

2. Asymptotic covariance matrix is scalar multiple of the limiting covariance matrix of AR(p) using Gaussian MLE.

Examples: scalar = .5 Laplace

= .7377 for t-distribution with 3 d.f.

Order Selection:

Partial ACF From the previous result, if true model is of order r and fitted model is of order p > r, then

$$n^{1/2}\hat{\phi}_{p,LAD} \to N(0, \frac{Var(|Z|)}{2\sigma^4 f_\sigma^2(0)})$$

where $\hat{\phi}_{p,LAD}$ is the pth element of $\hat{\phi}_{LAD}$.

Procedure:

1. Fit high order (P-th order) and obtain residuals and estimate scalar,

$$\theta^2 = \frac{Var(|Z_1|)}{2\sigma^4 f_\sigma^2(0)},$$

by empirical moments of residuals and density estimates.

2. Fit all-pass models of order p=1,2, . . . , P via LAD and obtain p-th coefficient $\hat{\phi}_{p,p}$ for each.

3. Choose model order r as the smallest order beyond which the estimated coefficients are statistically insignificant.

AIC: An approximate unbiased estimate of -2 log(like) based on an independent realization is

$$AIC(p) := -2L_X(\hat{\phi}, \hat{\kappa}) + \frac{Var(|Z_1|)}{E|Z_1|\sigma^2 f_\sigma(0)} p$$

Estimate coefficient of p using empirical moments of residuals. (Coefficient is 2 in traditional case.)

Simulation results:

- 1000 replicates of all-pass models
- model order parameter value

model order	parameter value
1	$\phi_1 = .4$
2	$\phi_1 = .3,\ \phi_2 = .4$

- noise distribution is t with 3 d.f.
- sample sizes n=500, 5000
- estimation method is LAD

To guard against being trapped in local minima, we adopted the following strategy.

- 250 random starting values were chosen at *random*. For model of order p, k-th starting value was computed recursively as follows:

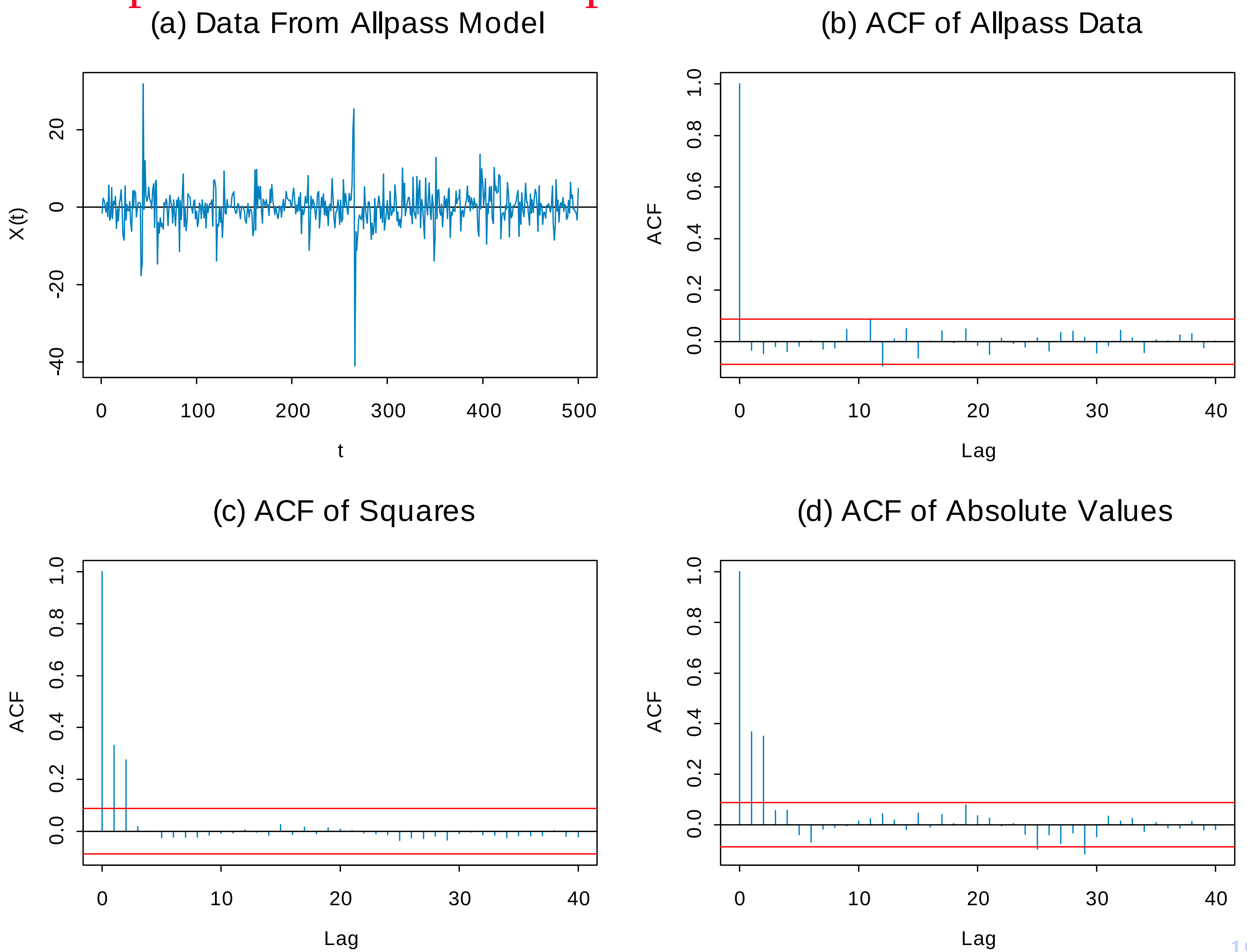
1. Draw $\phi_{11}^{(k)}, \phi_{22}^{(k)}, \ldots, \phi_{pp}^{(k)}$ iid uniform (-1,1).
2. For j=2, …, p, compute

$$\begin{bmatrix} \phi_{j1}^{(k)} \\ \vdots \\ \phi_{j,j-1}^{(k)} \end{bmatrix} = \begin{bmatrix} \phi_{j-1,1}^{(k)} \\ \vdots \\ \phi_{j-1,j-1}^{(k)} \end{bmatrix} - \phi_{jj}^{(k)} \begin{bmatrix} \phi_{j-1,j-1}^{(k)} \\ \vdots \\ \phi_{j-1,1}^{(k)} \end{bmatrix}$$

- Select top 10 based on minimum function evaluation.

- Run Hooke and Jeeves with each of the 10 starting values and choose best optimized value.

Sample realization of all-pass of order 2



Estimates:

$$\hat{\phi}_1 = .297(.0381), \hat{\phi}_2 = .374(.0381)$$

Standard errors computed as $\hat{\theta}\, sqrt\{(1-\hat{\phi}_2^2)/500\}$

where $\hat{\theta} = .919$

Order selection:

- cut-off value for PACF is 1.96*.908/sqrt(500)=.0796
- $AIC(p) := -2L_X(\hat{\phi}, \hat{\kappa}) + 1.896\, p$

	1	2	3	4	5
phi_p	0.289	0.374	0.009	0.011	0.01
AIC(p)	2451	2346	2347	2348	2350

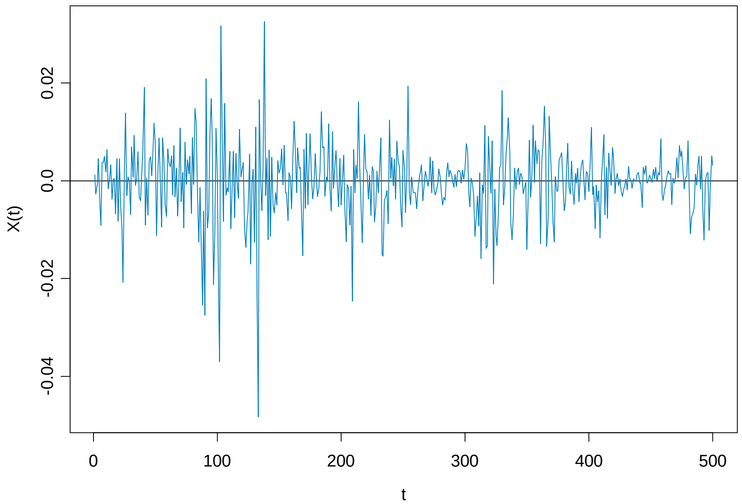
	6	7	8	9	10
phi_p	0.047	0.034	-0.05	0.083	0.021
AIC(p)	2348	2349	2345	2343	2345

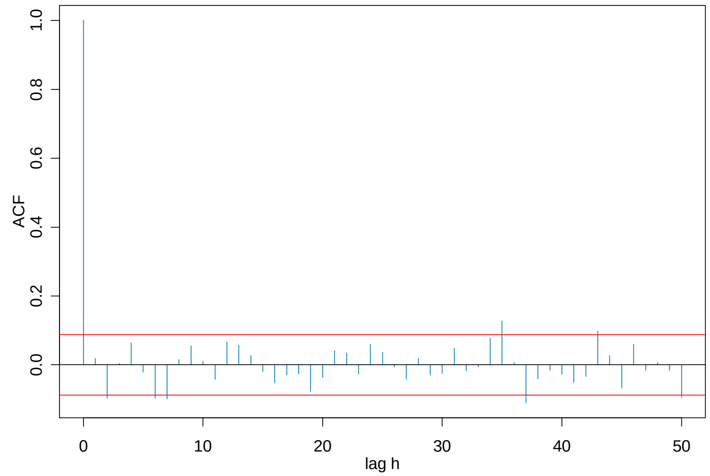
	Asymptotic			Empirical	
N	mean	std dev	mean	std dev	%coverage
500	ϕ_1=.5	.0332	.4979	.0397	94.2
5000	ϕ_1=.5	.0105	.4998	.0109	95.4

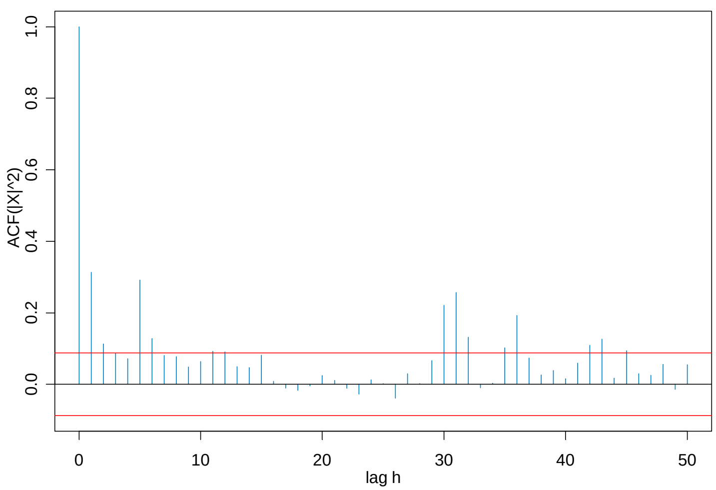
	Asymptotic			Empirical	
N	mean	std dev	mean	std dev	%coverage
500	ϕ_1=.3	.0351	.2990	.0456	92.5
	ϕ_2=.4	.0351	.3965	.0447	92.1
5000	ϕ_1=.3	.0111	.3003	.0118	95.5
	ϕ_2=.4	.0111	.3990	.0117	94.7

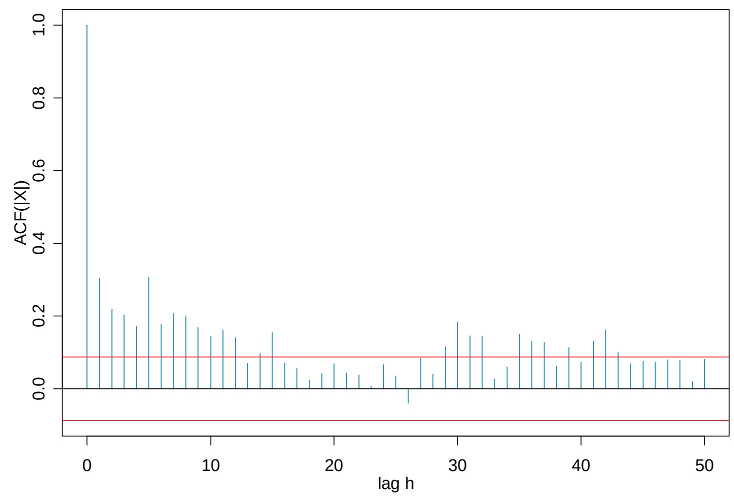
Application to financial data

500-daily log-returns of NZ/US exchange rate





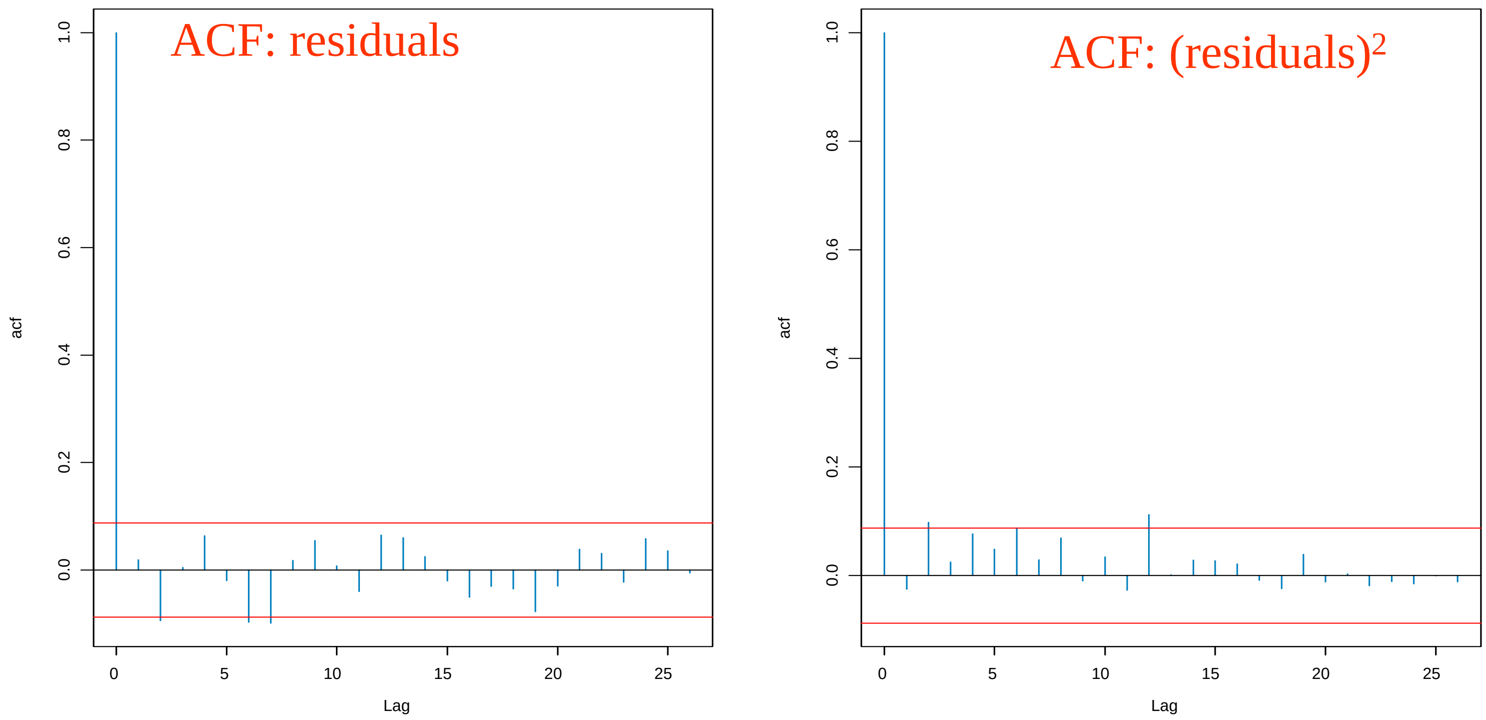




All-pass model fitted to NZ-USA exchange rates :

Order = 6, ϕ_1=-.367, ϕ_2=-.750, ϕ_3=-.391, ϕ_4=.088, ϕ_5= -.193, ϕ_6=-.096

(AIC had local minima at p=6 and 10)



Noncausal AR (p) models (with heavy tailed noise.)

$$X_t - \phi_1 X_{t-1} - \cdot\cdot\cdot - \phi_p X_{t-p} = Z_t,$$

a. $\{Z_t\} \sim \text{IID}(\alpha)$ with Pareto tails

b. $\phi(z) = 1 - \phi_1 z - \cdot\cdot\cdot - \phi_p z^p$

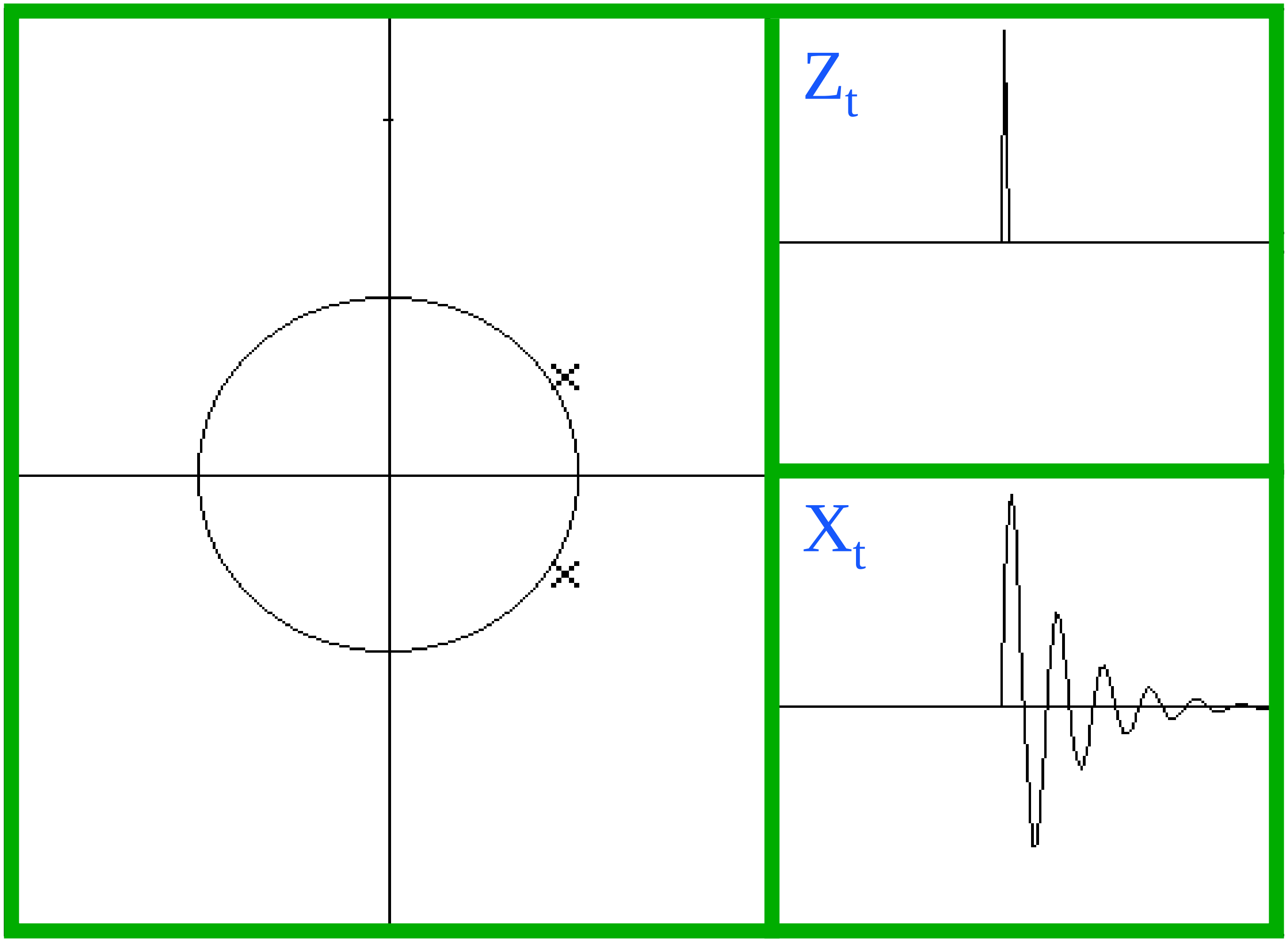
No zeros on the unit circle $\Rightarrow$ stationary.

No zeros inside the unit circle $\Rightarrow$ causal.

Some zero(s) inside the unit circle $\Rightarrow$ non-causal.

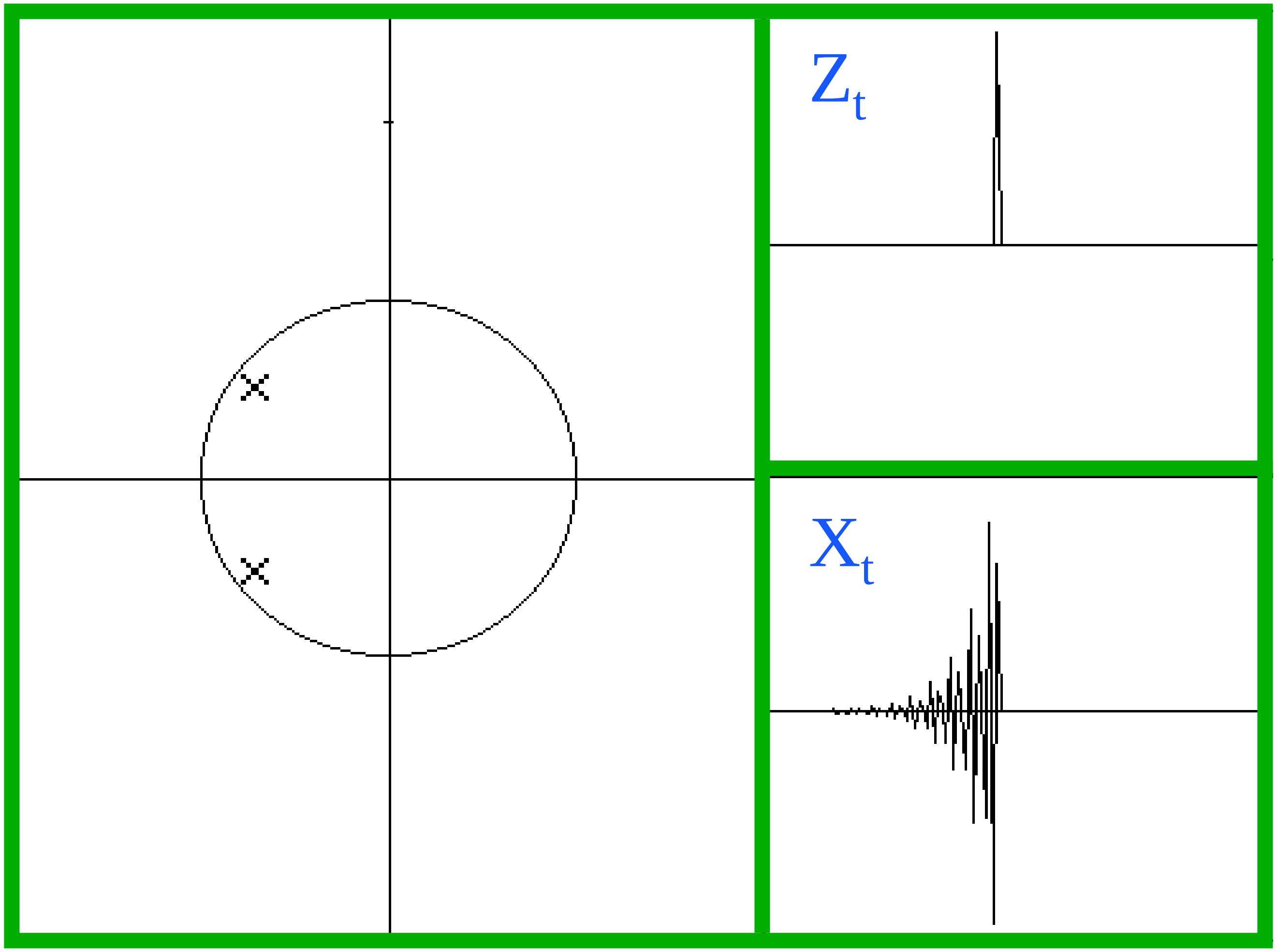
Impulse Response

Causal - Low frequency



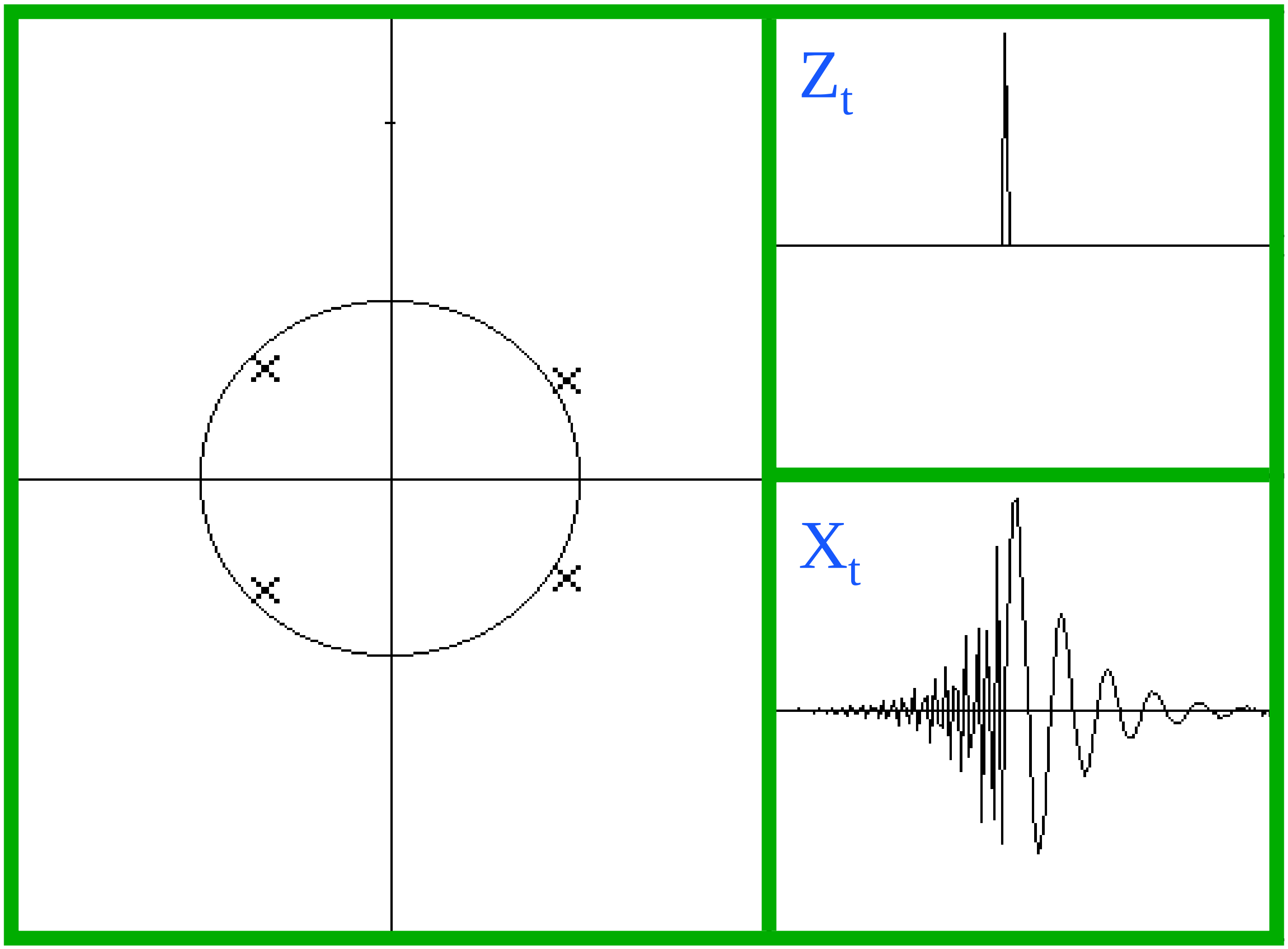
Impulse Response

Noncausal - High frequency



Impulse Response

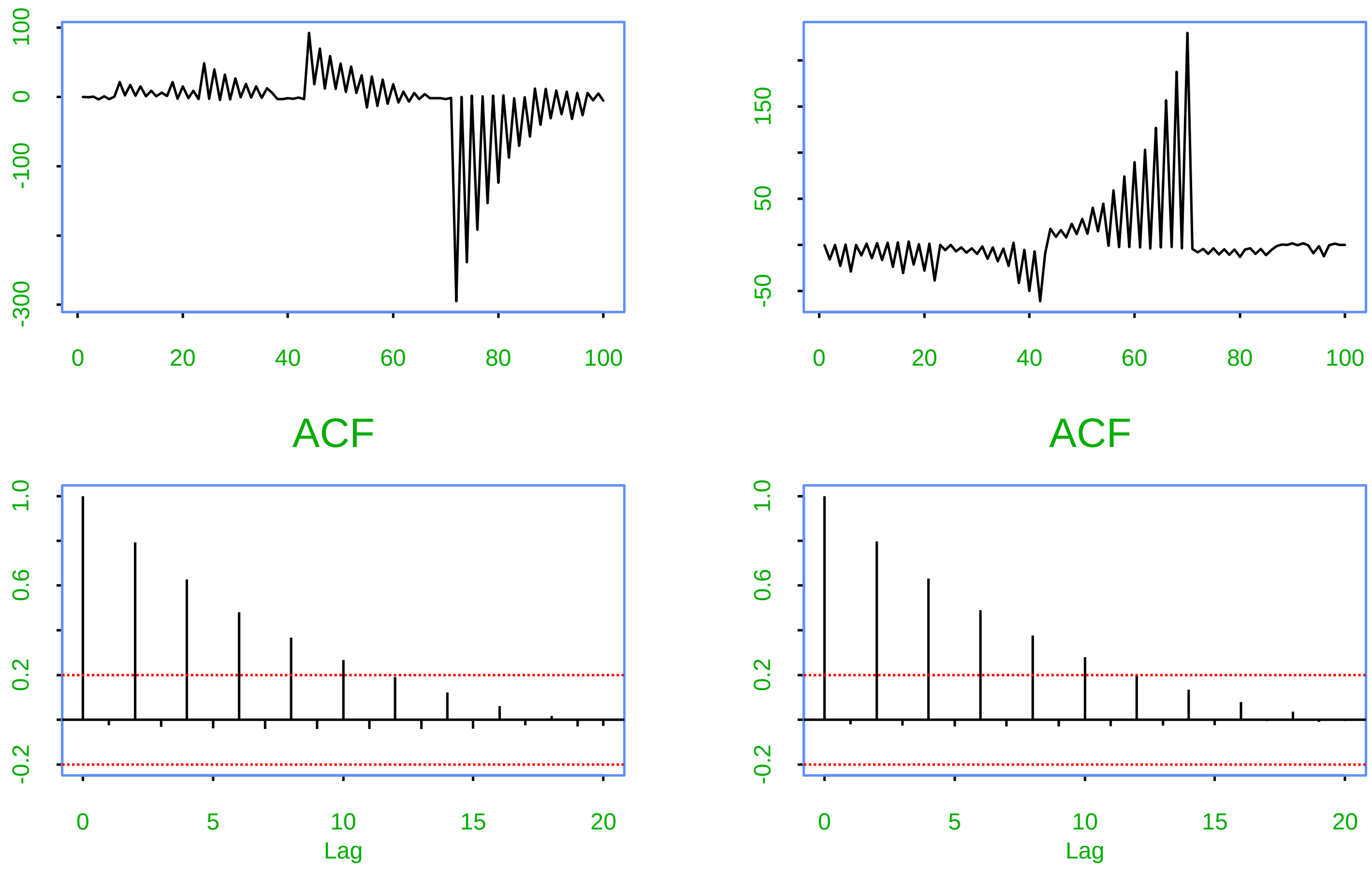
Mixed: High (non-causal) & Low (causal) frequency



Realization of a causal AR(2), and a non-causal AR(2)

Model: $\phi_*(B)X_t = Z_t$, $\{Z_t\} \sim \text{IID}(\alpha = 1)$, where

$\phi_c(B) = (1 - 0.9B)(1 + 0.9B)$ and $\phi_{nc}(B) = (1 - 1.1B)(1 + 1.1B)$



Application of all-pass to non-causal AR model fitting

Suppose $\{X_t\}$ follows the non-causal AR model

$$\phi_c(B)\,\phi_{nc}(B)\,X_t = Z_t\,,\quad \{Z_t\} \sim \text{IID}.$$

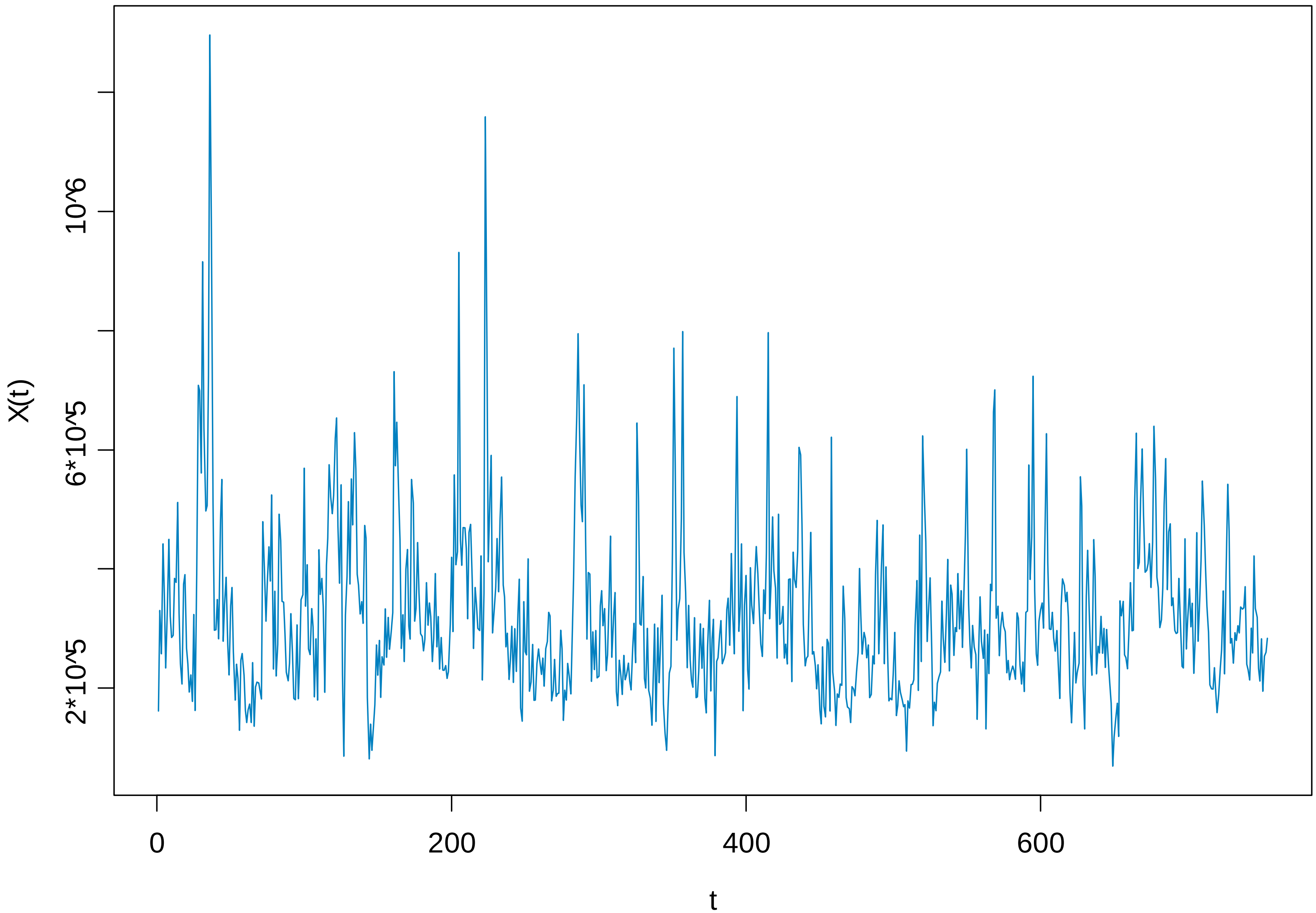
Let $\{U_t\}$ be the residuals obtained by fitting a purely causal AR model, i.e.,

$$\begin{aligned} U_t &= \hat{\phi}(B)X_t \\ &\approx \phi_c(B)\tilde{\phi}_{nc}(B)X_t, \quad (\tilde{\phi}_{nc} \text{ is the causal version of } \phi_{nc}) \\ &= \frac{\tilde{\phi}_{nc}(B)}{\phi_{nc}(B)} Z_t \end{aligned}$$

Thus $\{U_t\}$ follows the purely non-causal all-pass model,

$$\phi_{nc}(B)U_t = \tilde{\phi}_{nc}(B)Z_t.$$

Volumes of Microsoft (MSFT) stock traded over 754 transaction days (6/3/96 to 5/27/99)



Analysis of MSFT:

Log(volume) follows AR(1) or AR(3).

$$U_t = (1-.5834\,B)\,X_t \qquad \text{(causal AR(1))}$$

All-pass model of order 1 fitted to $\{U_t\}$:

$$(1-1.752\,B)U_t = (1-.5708\,B)Z_t.$$

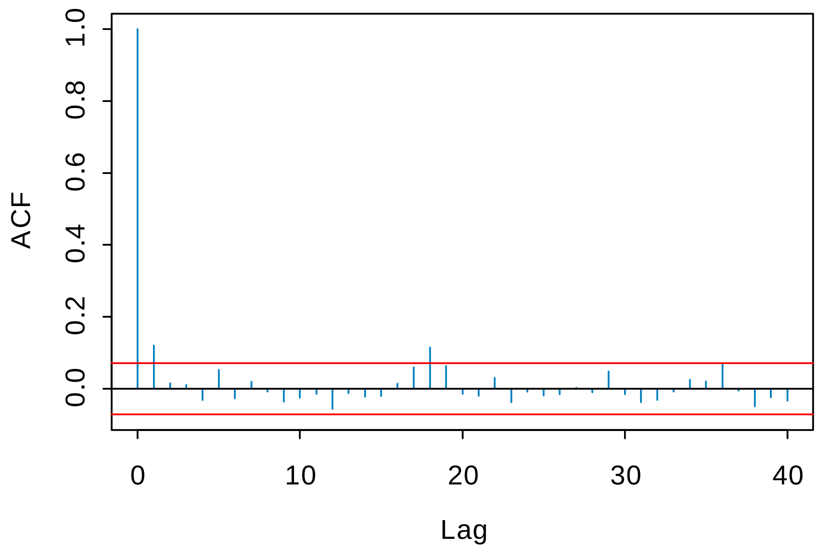
Combining the two models, we obtain the approximate non-causal model for $\{X_t\}$:

$$(1-1.752\,B)X_t = \frac{(1-.5708\,B)}{(1-.5834\,B)}Z_t \approx Z_t.$$

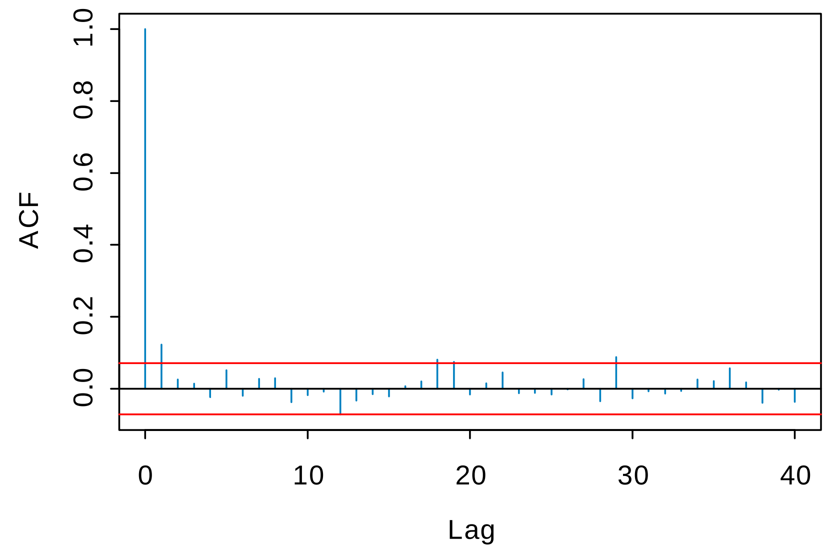
Estimated residuals from all-pass model fit:

$$\tilde{Z}_t = \frac{(1-1.752B)(1-.5834\,B)}{(1-.5708\,B)}X_t$$

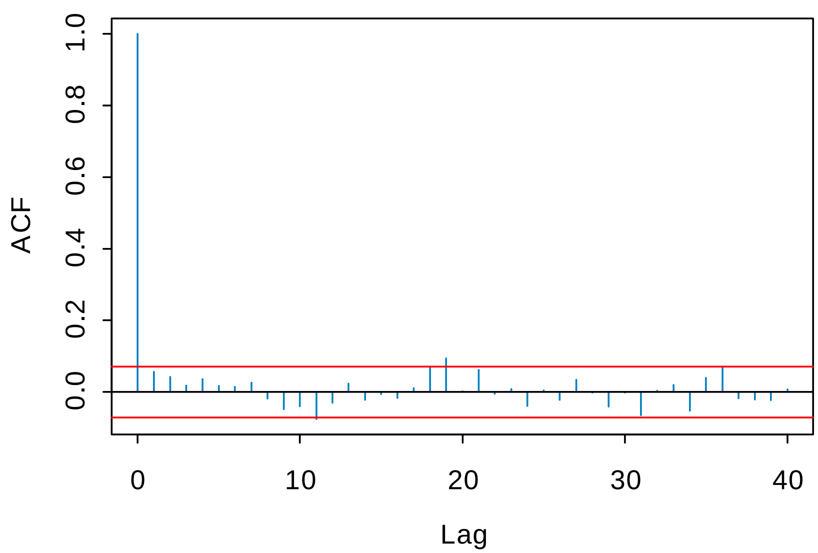
(a) ACF of Squares of Ut



(b) ACF of Absolute Values of Ut



(c) ACF of Squares of Zt



(d) ACF of Absolute Values of Zt