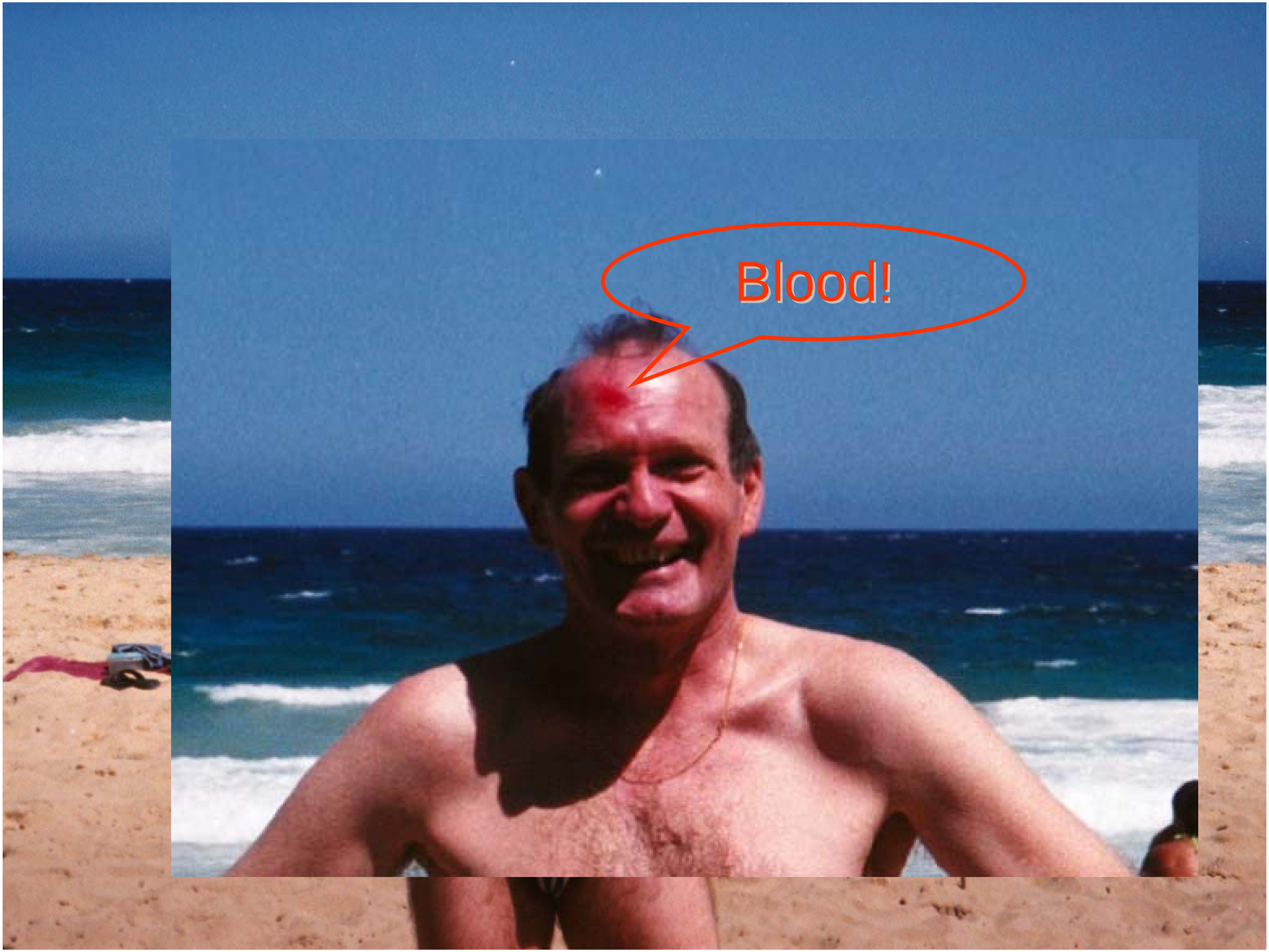
Maximum Likelihood Estimation for α-Stable AR and Allpass Processes

Beth Andrews
Northwestern University

Matt Calder
PHZ Capital Partners

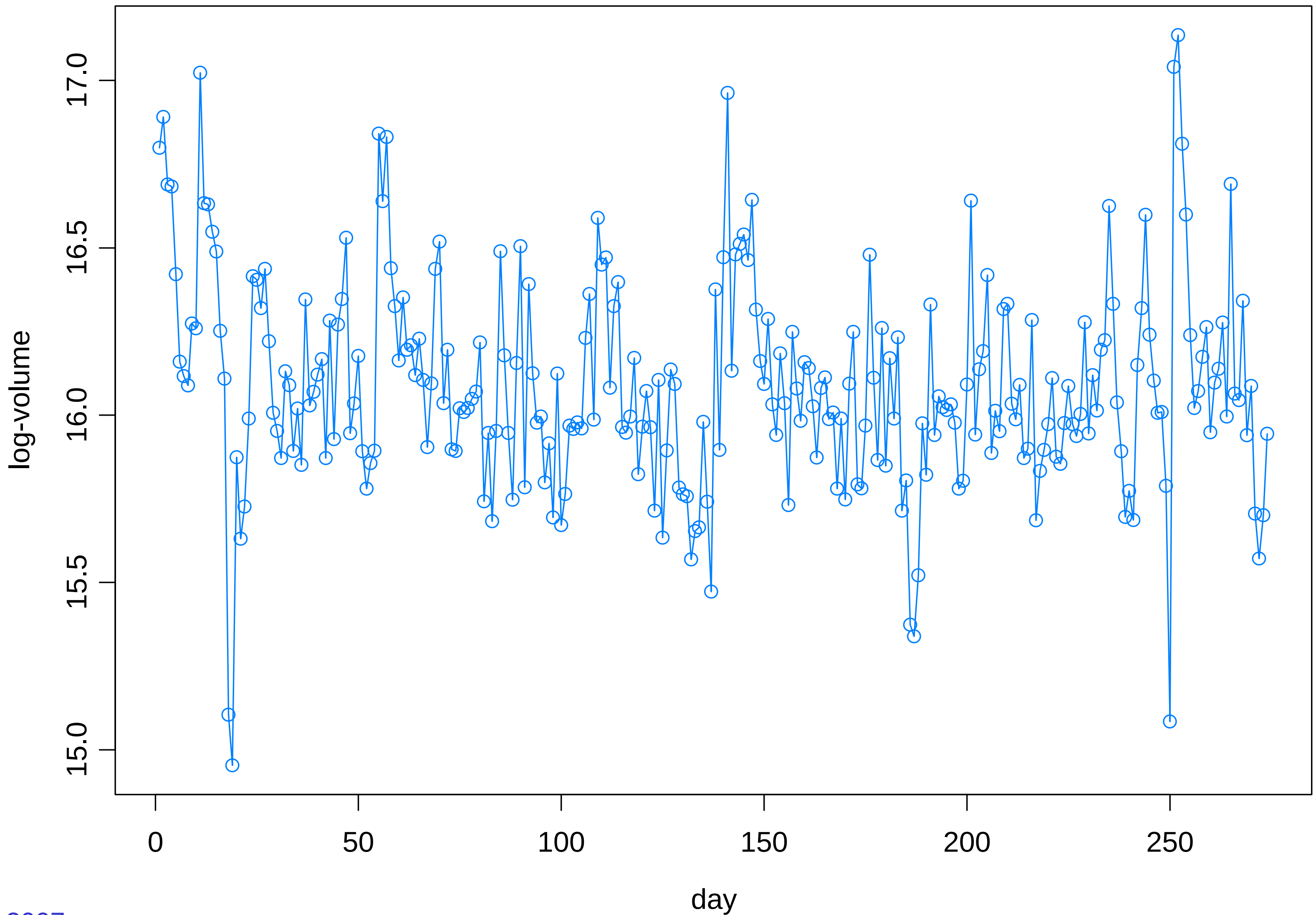
Richard A. Davis
Columbia University

(http://www.stat.columbia.edu/~rdavis)

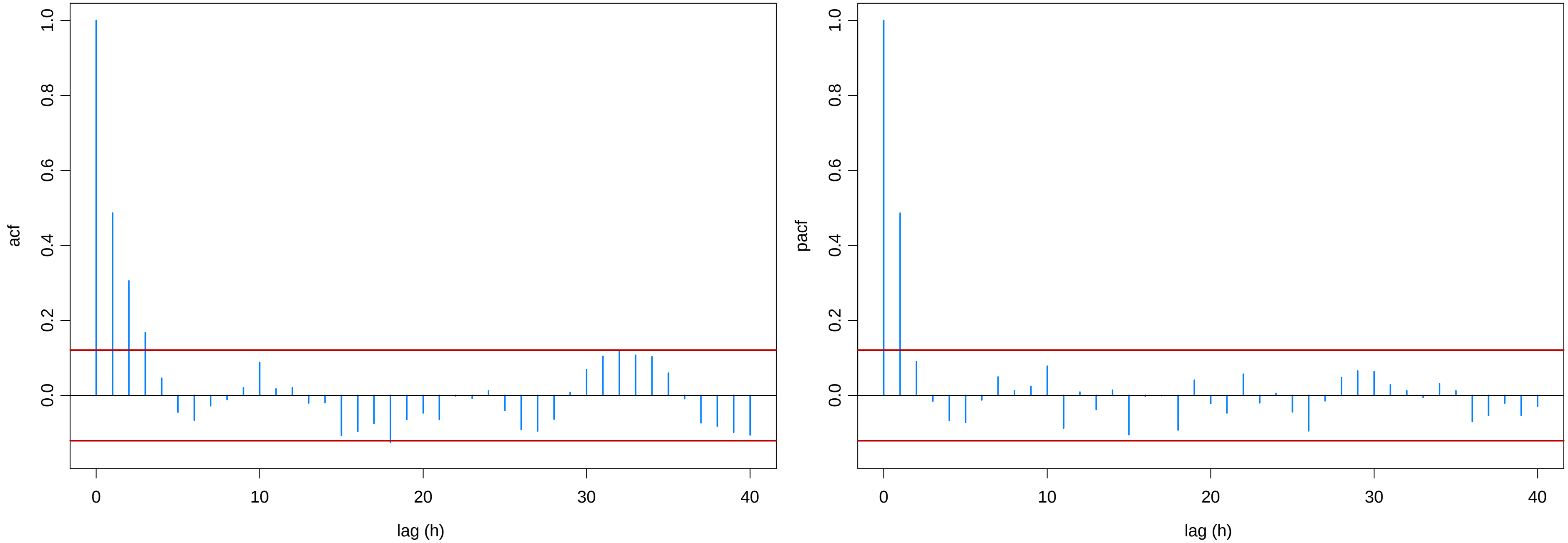

Blood!

1. Motivating Example

Log(volume) of Walmart stock 12/1/03-12/31/04



1. Motivating Example (cont)



Analysis suggests that $\{X_t\}$ follows an AR (1) or AR(2).

A *causal* AR(2) fit is

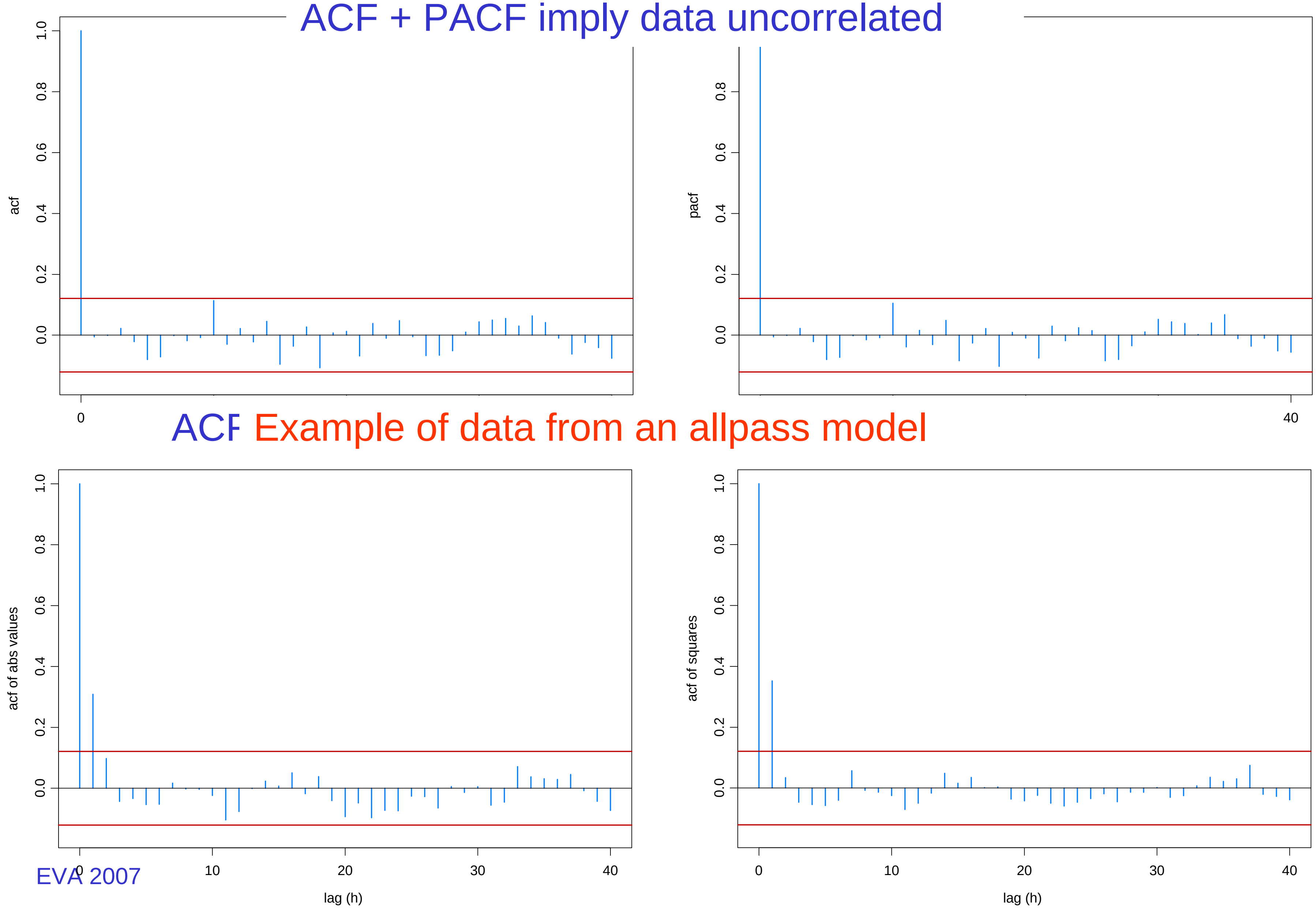
$$X_t = .4455\, X_{t-1} + .1025\, X_{t-2} + Z_t$$

Are the estimated residuals iid?

Analysis of residuals from causal AR(2) fit

ACF + PACF imply data uncorrelated

ACF Example of data from an allpass model



Game Plan

1. Motivating example: Wal-Mart volume
2. Setup
   - AR models
   - Allpass models
   - Stable noise
3. Maximum likelihood estimation
4. Simulation results
5. Wal-Mart revisited

2. Setup—AR models

Assume $\{X_t\}$ follows an autoregressive model or an allpass model.

AR(p) model:

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t, \text{ where } \{Z_t\} \sim \text{IID}$$

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

- ***causal*** if $\phi(z) \neq 0$ for $|z| \leq 1$, i.e., fcn of past values of Z_t

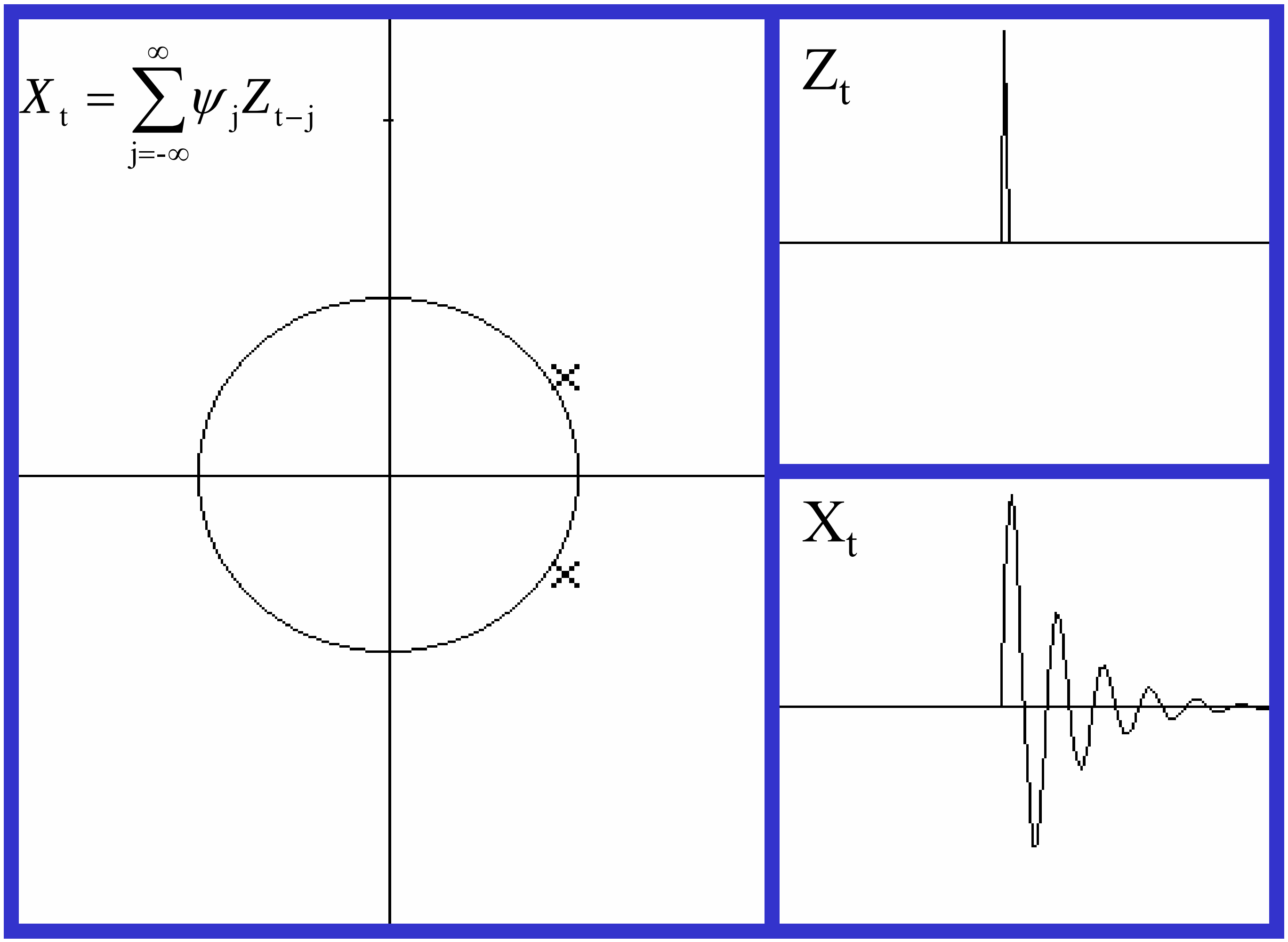
$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

- ***purely noncausal*** if $\phi(z) \neq 0$ for $|z| > 1$, i.e., fcn of future values of Z_t

$$X_t = \sum_{j=p+1}^{\infty} \psi_j Z_{t+j}$$

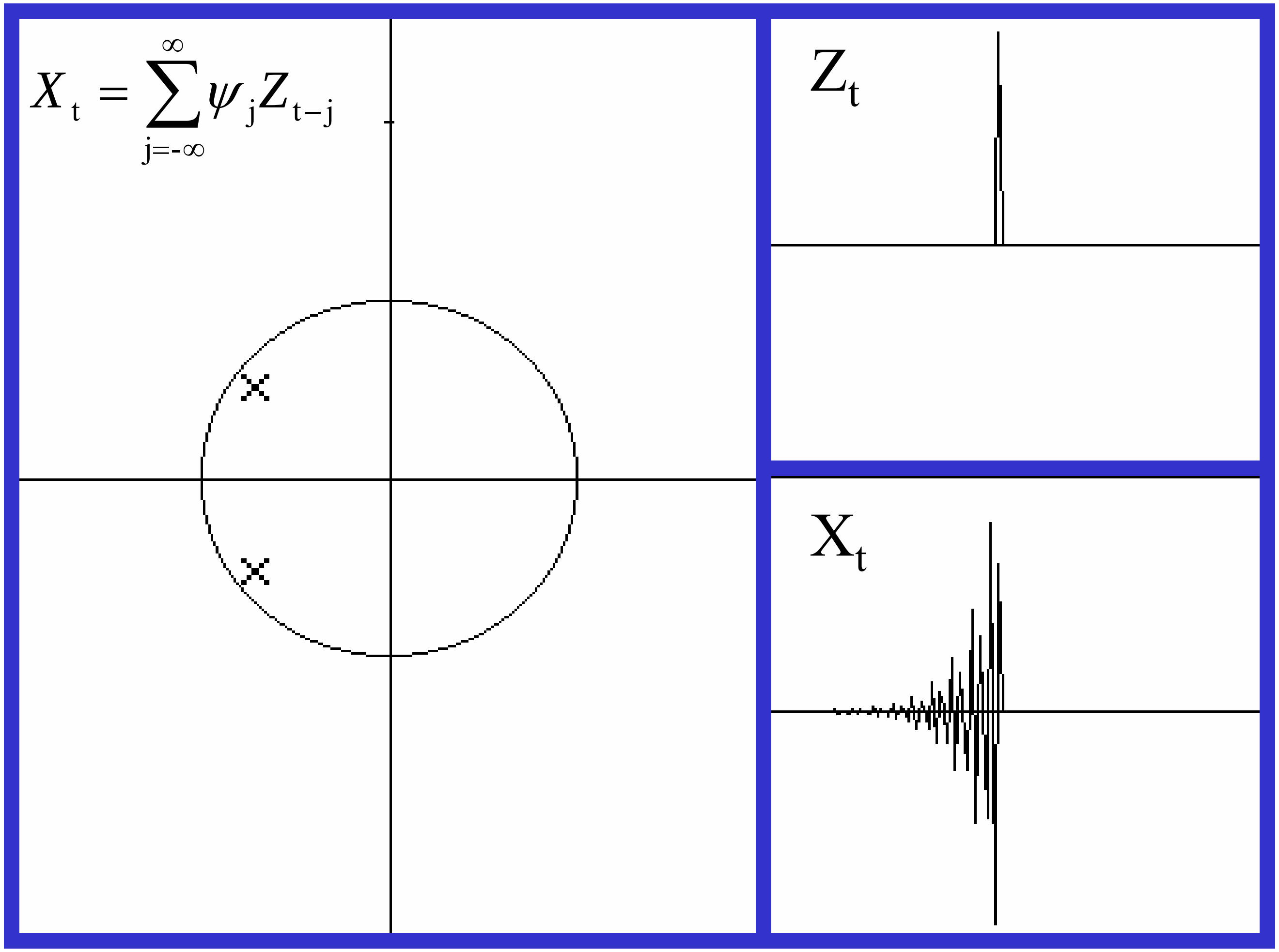
2. Setup—AR Models

Impulse response: causal & low frequency



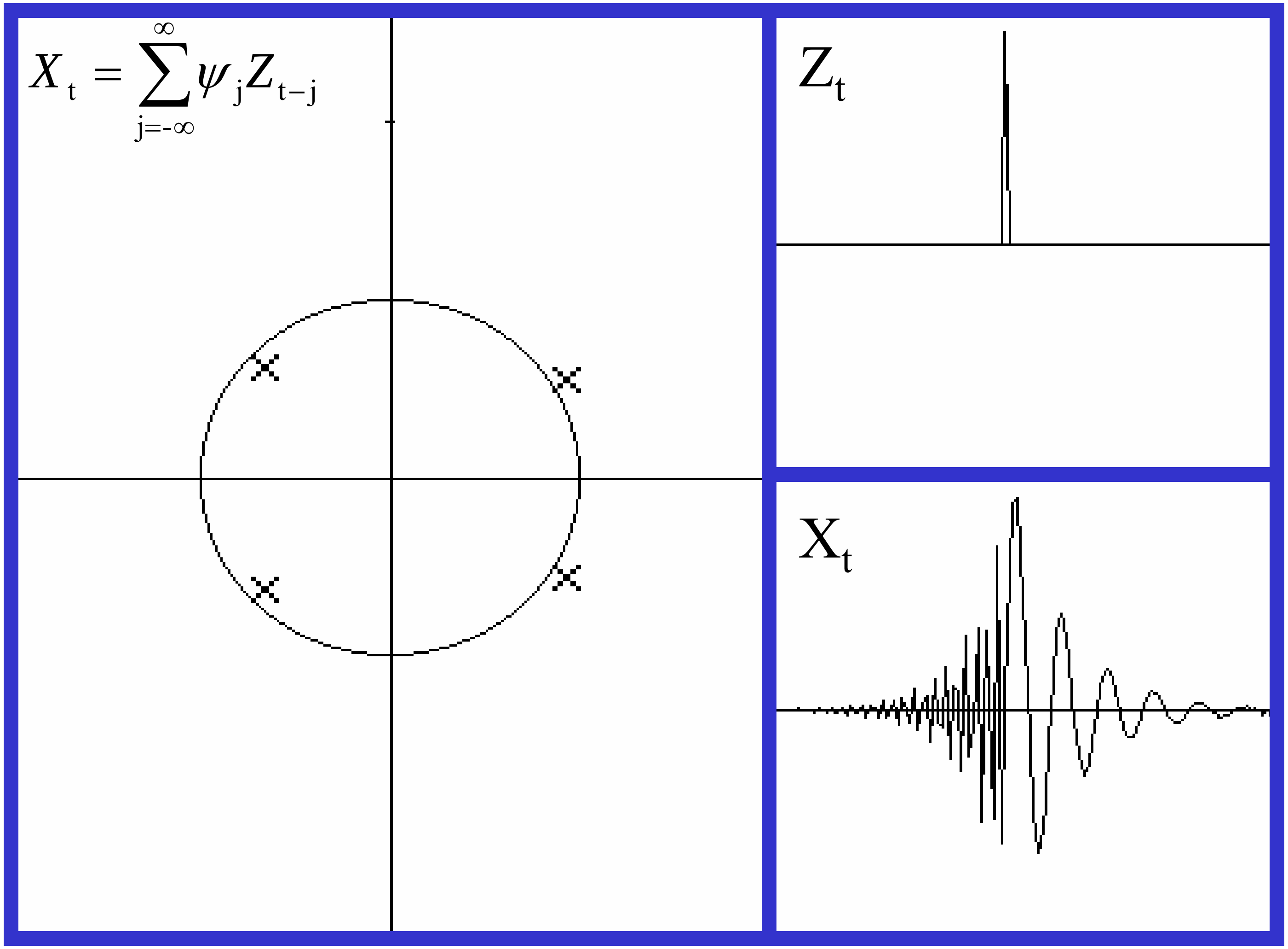
2. Setup—AR Models

Impulse response: noncausal & high frequency



2. Setup—AR Models

Impulse response: mixed causal (low frequency) & noncausal (high frequency)



2. Setup—Allpass 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 $\{X_t\}$: $\phi(B)X_t = \theta(B) Z_t$, $\{Z_t\} \sim$ IID (non-Gaussian)

$$B^k X_t = X_{t-k}$$

Examples:

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

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

2. Setup—Allpass models

Properties: *linear process with nonlinear behavior*

- causal, non-invertible ARMA with MA representation

$$X_t = \frac{B^p \phi(B^{-1})}{-\phi_p \phi(B)} Z_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

- uncorrelated (flat spectrum)

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

- zero mean
- data are dependent if noise is non-Gaussian (e.g. Breidt & Davis 1991).
- squares and absolute values are correlated.
- X_t is heavy-tailed if noise is heavy-tailed.

2. Setup—Allpass models

Properties cont:

- Suppose the time series X_t follows a noncausal AR and/or a non-invertible MA process. For definiteness, suppose $\{X_t\}$ follows the noninvertible MA model

$$X_t = \theta_i(B)\,\theta_{ni}(B)\,Z_t\,, \quad \{Z_t\} \sim \text{IID}.$$

Step 1: Let $\{U_t\}$ be the residuals obtained by fitting a purely invertible MA model, i.e.,

$$\begin{aligned} X_t &= \hat{\theta}(B)U_t \\ &\approx \theta_i(B)\tilde{\theta}_{ni}(B)U_t, \quad (\tilde{\theta}_{ni} \text{ is the invertible version of } \theta_{ni}). \end{aligned}$$

So

$$U_t \approx \frac{\theta_{ni}(B)}{\tilde{\theta}_{ni}(B)} Z_t$$

Step 2: Fit a purely causal AP model to $\{U_t\}$

2. Setup—stable noise

The noise is assumed to have a nonGaussian stable distribution

Noise distribution: $\{Z_t\}$ ~ IID with a stable distribution

That is, Z_t has characteristic function given by

$$\varphi_0(s) = E\{\exp(isZ_t)\} = \exp\{-\sigma^\alpha |s|^\alpha [1 + i\beta \operatorname{sgn}(s)\tan(\pi\alpha/2)(|\sigma s|^{1-\alpha} - 1)] + i\mu s\}$$

where

+ exponent $\alpha \in (0,2)$

+ symmetry parameter $\beta \in (-1,1)$

+ scale parameter $\sigma > 0 \in (0,2)$

+ location parameter μ.

Density function: inverse Fourier transform

$$f(z;(\alpha,\beta,\sigma,\mu)) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-isz} \varphi_0(s) ds$$

can be computed numerically reasonably quickly (Nolan, `97).

2. Setup—stable noise (cont)

Remarks:

1. There are a number of parameterizations of a stable distribution (see Zolotarev `86), but this is the one advocated by Nolan `01.

   - differentiable wrt to z
   - differentiable wrt to the parameter vector $(\alpha,\beta,\sigma,\mu)'$

2. Stable distributions have pareto-like heavy tails

$$x^{\alpha} P(|Z_1| > x) \to \sigma^{\alpha} c(\alpha) \quad \text{and} \quad c(\alpha) = \left(\int_0^{\infty} t^{-\alpha} \sin(t) dt \right)^{-1}.$$

3. Location/scale family

$$f\left(z;(\alpha,\beta,\sigma,\mu)\right) = \sigma^{-1} f\left(\sigma^{-1}(z-\mu);(\alpha,\beta,1,0)\right)$$

4. Unimodal. $f\left(\bullet;(\alpha,\beta,1,0)\right)$ is unimodal

3. Maximum Likelihood Estimation

Estimation for nonCausal and/or All-Pass Models

☞ Second-order moment techniques do not work (work for causal AR!)

- least squares
- Gaussian likelihood

☞ Higher-order cumulant methods

- Giannakis and Swami (1990)
- Chi and Kung (1995)

☞ Non-Gaussian likelihood methods

- likelihood approximation assuming known density
- quasi-likelihood

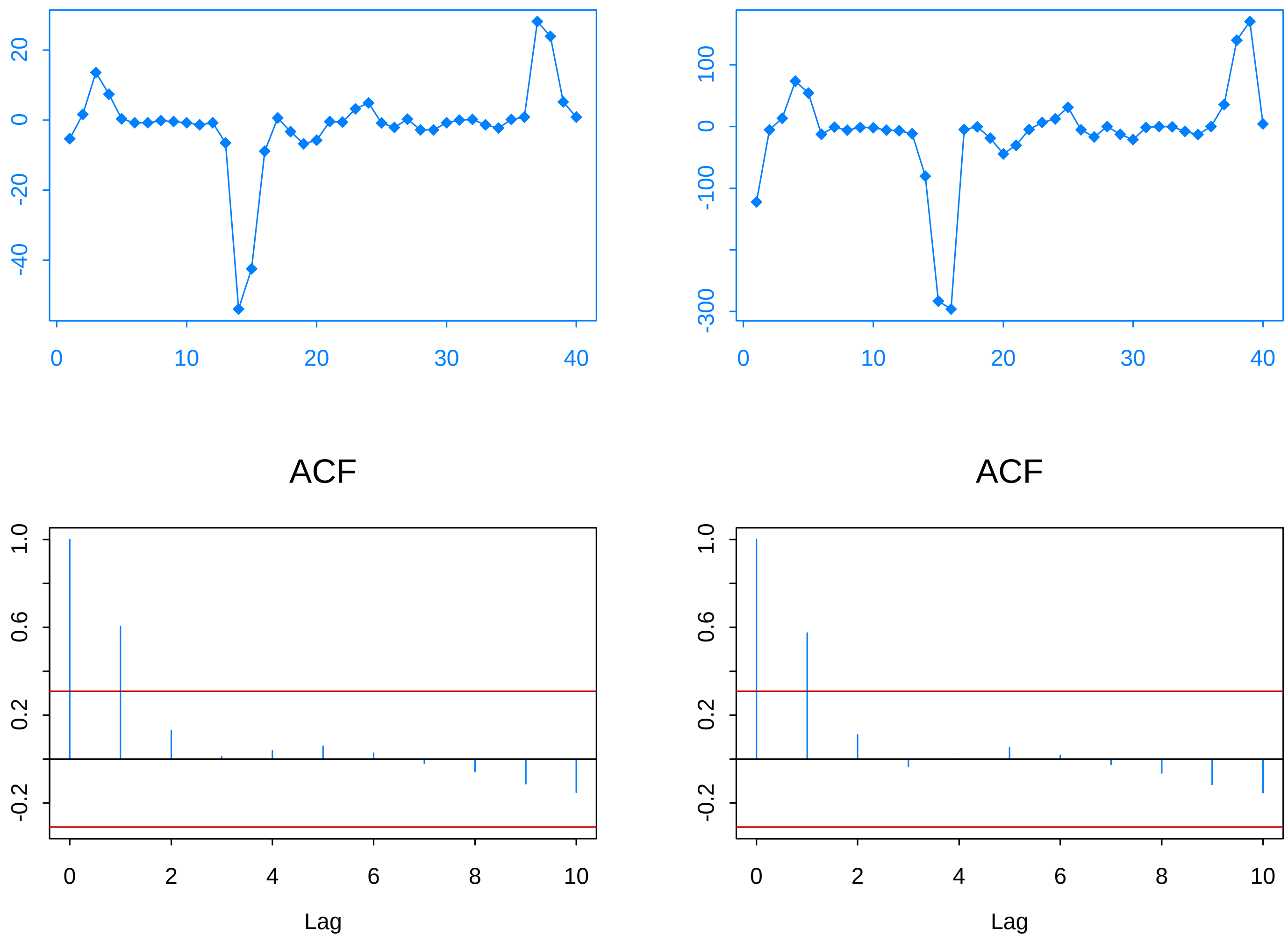
☞ Other

- LAD- least absolute deviation
- Rank-based estimation (minimum dispersion)

Realizations of an invertible and noninvertible MA(2) processes

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

$\theta_i(B) = (1 + 1/2B)(1 + 1/3B)$ and $\theta_{ni}(B) = (1 + 2B)(1 + 3B)$



3. Estimation for AR Models

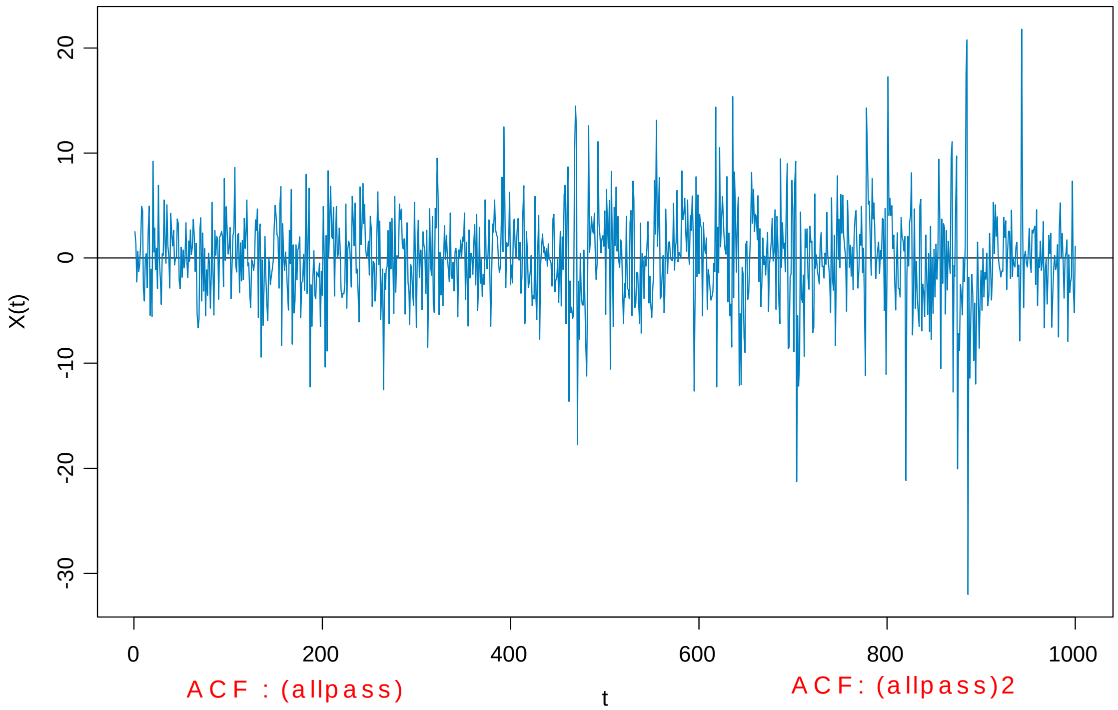
Estimation strategies for *causal* AR models:

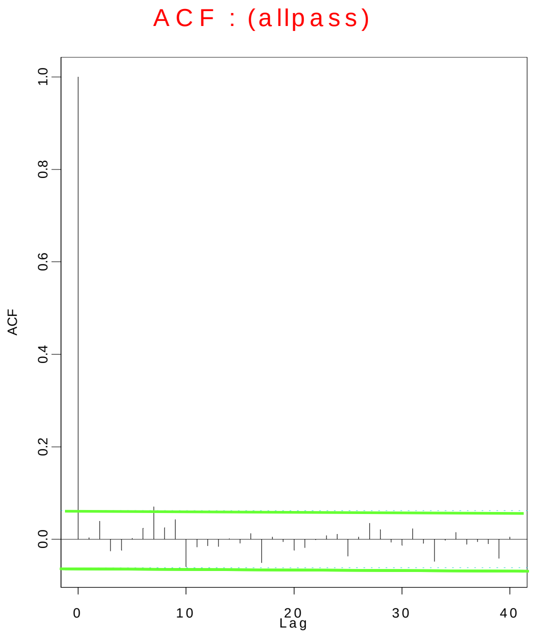
- LS estimation (Davis and Resnick `86)
- Maximum Gaussian likelihood (Mikosch, Gadrich, Kluppelberg, Adler `95)
- LAD- least absolute deviation (special case of M-Estimation) (Davis, Knight, Liu `92; Davis `96)

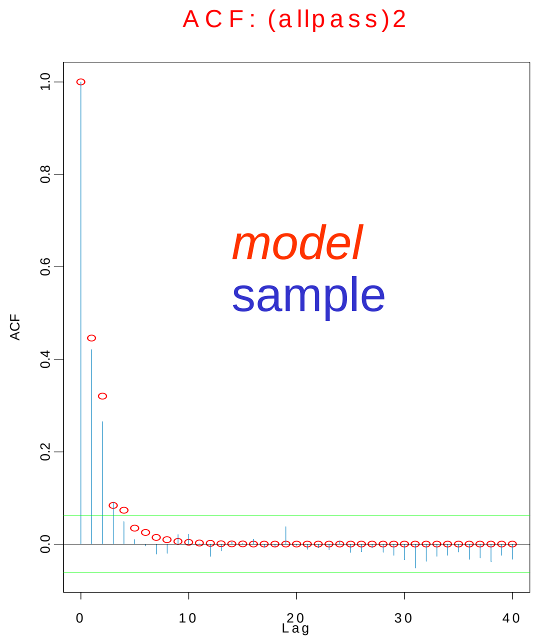
4. Allpass models

Realization from an all-pass model of order 2

(t3 noise)







3. MLE—allpass models: 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 stable pdf f_τ 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**:

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

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

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

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

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

Log-likelihood: Let $\tau = (\alpha, \mathrm{sgn}(\phi_q)\beta, \sigma / |\phi_q|, \mu / \phi_q)'$

$$L(\boldsymbol{\phi}, \tau) = \sum_{t=1}^{n-p} (\ln f(\phi_q z_t(\boldsymbol{\phi}); (\alpha, \beta, \sigma, \mu)) + \ln|\phi_q|)$$

$$= \sum_{t=1}^{n-p} \ln\left[\frac{|\phi_q|}{\sigma} f\left(\frac{z_t(\boldsymbol{\phi}) - \mu/\phi_q}{\sigma/\phi_q}; (\alpha, \beta, 1, 0)\right)\right]$$

$$= \sum_{t=1}^{n-p} \ln\left[f\left(z_t(\boldsymbol{\phi}); (\alpha, \mathrm{sgn}(\phi_q)\beta, \sigma/|\phi_q|, \mu/\phi_q)\right)\right]$$

$$= \sum_{t=1}^{n-p} \ln\left[f\left(z_t(\boldsymbol{\phi}); \tau\right)\right]$$

3. MLE—noncausal AR models

One can go through a similar calculation for noncausal AR models.

Model:

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t$$

$$Z_t = (1-\phi_1 B + \cdots + \phi_p B^p) X_t$$ B=backward shift operator

$$Z_t = (1-\theta_1 B + \cdots + \theta_r B^r)(1-\theta_{r+1} B + \cdots + \theta_{r+s} B^s) X_t$$

$$Z_t = \theta^+(B)\theta^*(B) X_t$$

where $\theta^+(z)$ is the good (***causal***) AR polynomial and $\theta^*(B)$ is the bad (***purely noncausal***) AR polynomial.

Likelihood: $\tau = (\alpha, \operatorname{sgn}(\phi_q)\beta, \sigma/|\phi_q|, \mu/\phi_q)'$

$$L(\theta,\tau) = \sum_{t=p+1}^{n} \left(\ln\left[f\left(Z_t(\theta);\tau\right)\right] + \ln|\theta_p| I(s>0) \right)$$

Reparameterize: Denote the true parameters by ϕ_0 and τ_0 and set

$$u = n^{1/\alpha_0}(\phi - \phi_0) \quad \text{and} \quad v = n^{1/2}(\tau - \tau_0)$$

which are elements of R^p and R^4, respectively. Now the log-likelihood can be re-expressed as a continuous function on $R^p \times R^4$ given by

$$\begin{aligned} W_n(u,v) &= L(\phi_0 + n^{-1/\alpha_0}u, \tau_0 + n^{-1/2}v) - L(\phi_0, \tau_0) \\ &= \sum_{t=1}^{n-p} \ln\left[f\left(z_t(\phi_0 + n^{-1/\alpha_0}u); \tau_0 + n^{-1/2}v\right)\right] - \sum_{t=1}^{n-p} \ln\left[f\left(z_t(\phi_0); \tau_0\right)\right] \end{aligned}$$

Note:

$$\left(\hat{u}_n, \hat{v}_n\right) = \arg\max_{u,v} W_n(u,v) = \left(n^{1/\alpha_0}(\hat{\phi} - \phi_0), n^{1/2}(\hat{\tau} - \tau_0)\right)$$

Result: $W_n(u,v) \to_d W(u) + v'\mathbf{N} - 2^{-1} v'\mathbf{I}(\tau_0)v$ on $C(R^p \times R^4)$, where

- $\mathbf{I}(\tau_0) := -\left[\mathrm{E}\{\partial^2 \ln f(z;\tau)/(\partial\tau_i \partial\tau_j)\} \right]_{i,j=1}^{4}$ is the Fisher information for a stable density.
- $\mathbf{N} \sim N(\mathbf{0}, \mathbf{I}(\tau_0))$ is a normal random vector independent of W.
- W is the process

$$W(u) = \sum_{k=1}^{\infty} \sum_{j \neq 0} \left\{ \ln f\left(z_{k,j} + [\tilde{c}(\alpha_0)]^{1/\alpha_0} \sigma_0 |\phi_{0r}|^{-1} c_j(u) \delta_k \Gamma_k^{-1/\alpha_0}; \tau_0 \right) - \ln f\left(z_{k,j}; \tau_0 \right) \right\}$$

  - $c_j(u)$ are found through the Laurent series identity

$$\sum_{j \neq 0} c_j(u) z^j = u' \left[-(z^{-k}/\phi_0(z)) + (z^k/\phi_0(z^{-1})) \right]_{k=1}^{p}$$

  - $\{z_{k,j}\}$ is iid $z_{1,1} =_d Z_1/\phi_{0r}$
  - $\{\delta_k\}$ is iid $P(\delta_k=1) = p = 1 - P(\delta_k=-1)$.
  - $\Gamma_k = E_1 + \cdots + E_k$, where $\{E_k\}$ is iid unit exponentials
  - $\{z_{k,j}\}$, $\{\delta_k\}$, and $\{E_k\}$ are mutually independent

Deconstructing the limit result and limit process:

The limit process in the following

$$W_n(u,v) \to_d W(u) + v'\mathbf{N} - 2^{-1} v'\mathbf{I}(\tau_0)v$$

consists of two independent pieces:

1. W(u) which governs the limit behavior of the AR or AP model parameters $\phi_1, \ldots, \phi_p$. Set $\xi = \arg\max_u W(u)$
2. $v'\mathbf{N} - 2^{-1}v'\mathbf{I}(\tau_0)v$ which governs the limit behavior of the stable parameters $(\alpha, \beta, \sigma, \mu)$. Set $\eta = \arg\max_v v'\mathbf{N} - 2^{-1}v'\mathbf{I}(\tau_0)v$, is equal to $\eta = \mathbf{I}^{-1}(\tau_0)\mathbf{N} \sim \mathrm{N}(0, \mathbf{I}^{-1}(\tau_0))$.

Theorem: There exists a sequence of maximizers of the likelihood function such that

$$n^{1/\alpha_0}(\hat{\phi}_{ML} - \phi_0) \to_d \xi \quad \text{and} \quad n^{1/2}(\hat{\tau}_{ML} - \tau_0) \to_d \eta.$$

3. MLE—allpass models

$$n^{1/\alpha_0}(\hat{\phi}_{ML} - \phi_0) \to_d \xi, \qquad n^{1/2}(\hat{\tau}_{ML} - \tau_0) \to_d \eta \sim \mathrm{N}(0, \mathrm{I}^{-1}(\tau_0)).$$

Remarks:

- The distribution of ξ is generally intractable. However, one can use bootstrapping techniques (see later example and Davis and Wu `97).
- The scaling for the AR parameters is $n^{1/\alpha}$, which is must faster than the standard $n^{1/2}$ rate.
- The limit behavior of the estimates of the stable parameters is the same as for an iid sequence (see DuMouchel `73).

4. Simulation Results

Simulation setup:

- 300 replicates of a non-causal AR(2) model

  $X_t = -1.2\, X_{t-1} + 1.6\, X_{t-2} + Z_t$ (zeros of AR polyn 1.25, -.5)

- noise distribution is stable with two sets of parameter values:

  - $\alpha = .8 \quad \beta = .5 \quad \sigma = 1.0 \quad \mu = 0.0$ (really heavy!)
  - $\alpha = 1.5 \quad \beta = .5 \quad \sigma = 1.0 \quad \mu = 0.0$

- sample sizes n=500

- estimation is maximum likelihood

4. Simulation Results

	Asymp SD	Empirical Mean	Empirical Std Dev		Asymp SD	Empirical Mean	Empirical Std Dev
$\phi_1 = -1.2$		-1.200	0.004	$\phi_1 = -1.2$		-1.200	0.004
$\phi_2 = 1.6$		1.600	0.004	$\phi_2 = 1.6$		1.600	0.004
$\alpha = 0.8$	0.051	0.798	0.041	$\alpha = 0.8$	0.049	0.800	0.039
$\beta = 0.0$	0.067	-0.001	0.068	$\beta = 0.5$	0.058	0.502	0.056
$\sigma = 1.0$	0.077	0.997	0.073	$\sigma = 1.0$	0.074	0.997	0.071
$\mu = 0.0$	0.054	-0.002	0.057	$\mu = 0.0$	0.062	-0.004	0.064
$\phi_1 = -1.2$		-1.212	0.083	$\phi_1 = -1.2$		-1.204	0.078
$\phi_2 = 1.6$		1.605	0.065	$\phi_2 = 1.6$		1.598	0.062
$\alpha = 1.5$	0.071	1.502	0.069	$\alpha = 1.5$	0.070	1.499	0.071
$\beta = 0.0$	0.137	0.010	0.128	$\beta = 0.5$	0.121	0.509	0.128
$\sigma = 1.0$	0.048	0.999	0.066	$\sigma = 1.0$	0.047	0.997	0.056
$\mu = 0.0$	0.078	-0.006	0.078	$\mu = 0.0$	0.078	0.000	0.083

5. Walmart revisited---residuals from noncausal model

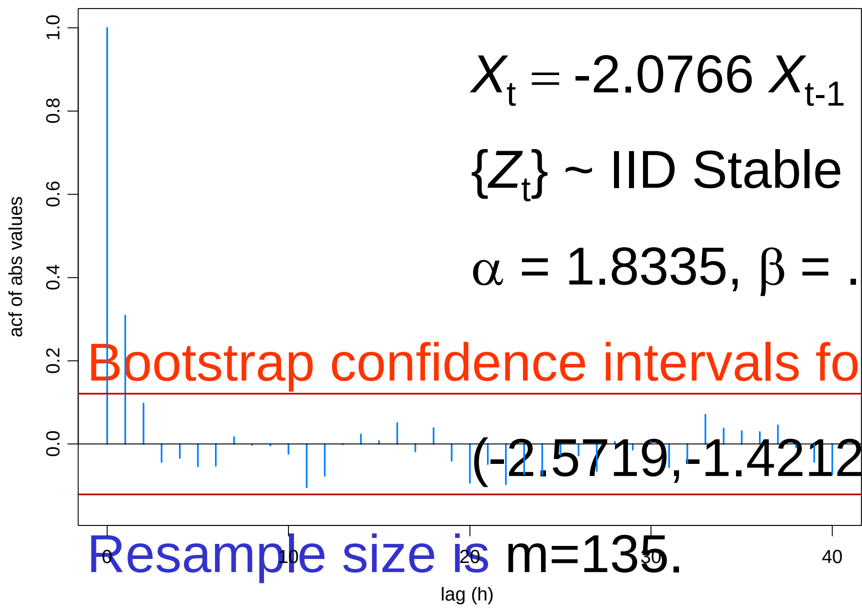
Analysis of the ACF and PACF of the time series (n=274) suggests that $\{X_t\}$ follows an AR (1) or AR(2).

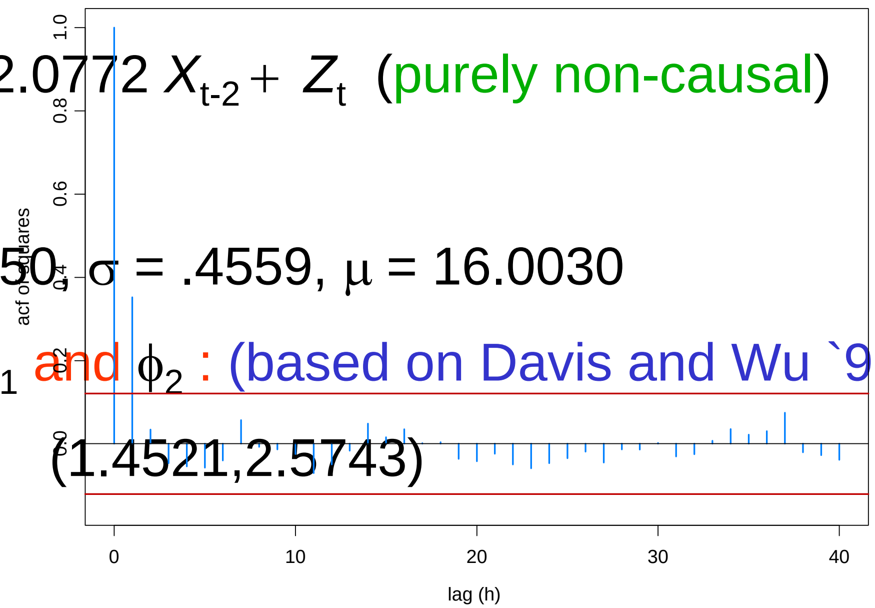
A *causal* AR(2) fit (using Gaussian MLE) is

$$X_t = .4455\, X_{t-1} + .1025\, X_{t-2} + Z_t$$

The estimated residuals were uncorrelated but ***dependent***.

Maximum-likelihood model:





$$X_t = -2.0766\, X_{t-1} + 2.0772\, X_{t-2} + Z_t \quad \text{(purely non-causal)}$$

$\{Z_t\}$ ~ IID Stable

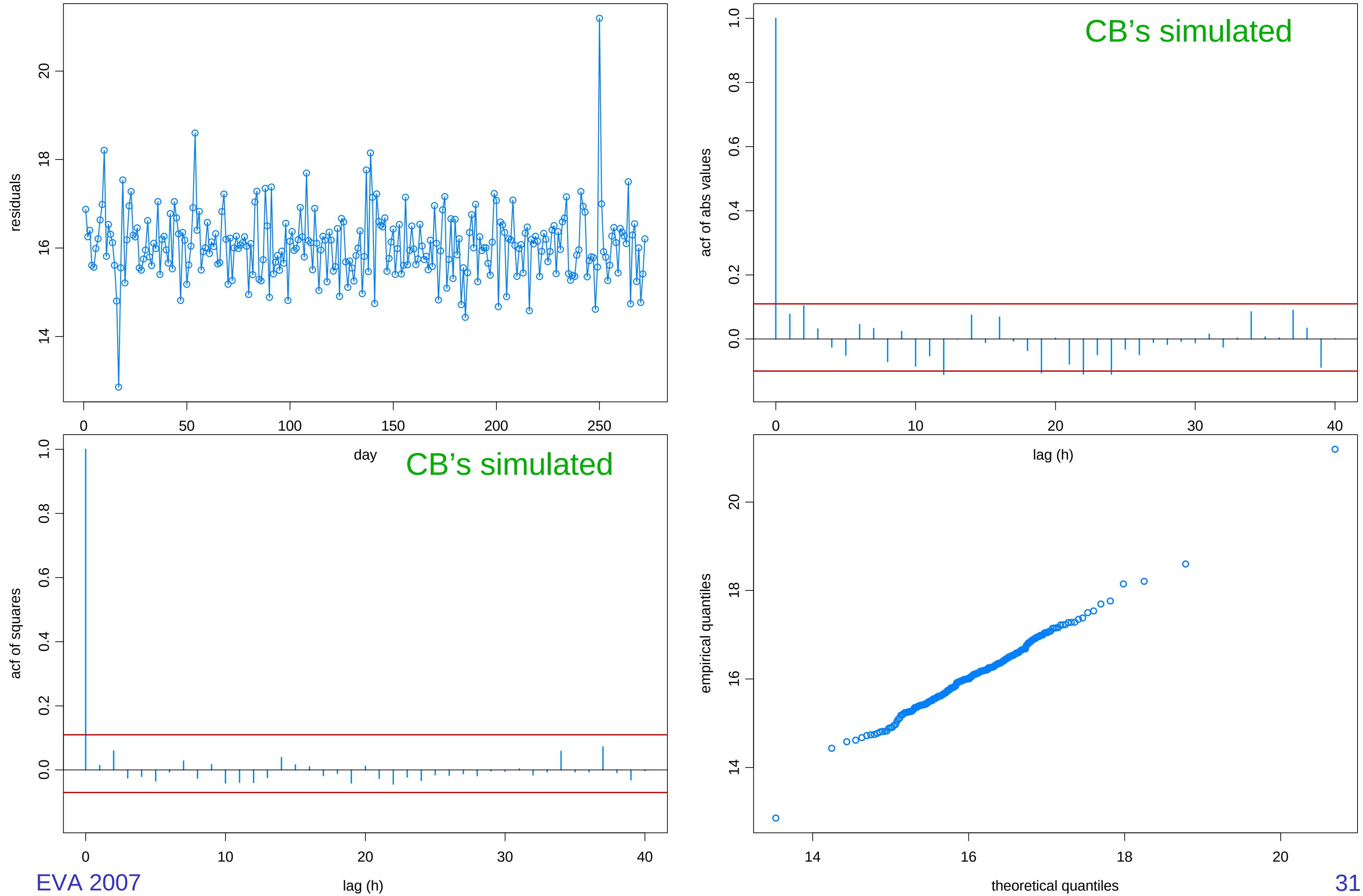
α = 1.8335, β = .5650, σ = .4559, μ = 16.0030

Bootstrap confidence intervals for ϕ_1 and ϕ_2 : (based on Davis and Wu `97)

(-2.5719,-1.4212) (1.4521,2.5743)

Resample size is m=135.

5. Walmart—analysis of residuals from noncausal model



6th Conference on Extremes: Fort Collins, Colorado

June 22-26, 2009


Graybill VIII Conference in 2009
6th Conference on Extremes
Fort Collins, Colorado
June 22- 26, 2009