KING'S College LONDON

University of London

Spectra of Sample Auto-Covariance Matrices Derived from Time Series

Reimer Kühn, Peter Sollich

Disordered System Group, Department of Mathematics, King's College London

Open Statistical Physics, March 7, 2012

arXiv:1112.4877v2 [cond-mat.dis-nn]

Overview

- Sample Auto-Covariance Matrices of Time Series
- Spectral Density and Resolvent (Edwards and Jones, 1976)
  - Annealed Average
  - Exploit Szegö's theorem
  - Scaling
- Numerical Tests
- Summary

Sample Auto-Covariance Matrices of Time Series

- Auto-covariance matrix of stationary stochastic process $(x_n)_{n\in\mathbb{Z}}$:
  From $N \times M$ matrix $X = (x_{it})$ with entries $x_{it} = x_{i+t}$ compute

$$C_{ij} = \frac{1}{M}(XX^T)_{ij} = \frac{1}{M}\sum_{t=0}^{M-1} x_{i+t}x_{j+t} \ .$$

  Expect finite sample fluctuation around mean

$$C_{ij} = \langle x_i x_j \rangle \pm \mathcal{O}(1/\sqrt{M}) = \bar{C}(i-j) \pm \mathcal{O}(1/\sqrt{M})$$

  $\Rightarrow C$ is randomly perturbed Toeplitz matrix.

- Spectrum of C as $N \to \infty$, $M \to \infty$ @ fixed $\alpha = N/M$?
  Known result as $\alpha \to 0$ (Szegö's Theorem):

$$\rho_0(\lambda) = \int_0^{2\pi} \frac{\mathrm{d}q}{2\pi}\, \delta(\lambda - \hat{C}(q))$$

Compare with Wishart–Laguerre Ensemble

- Empirical covariances for N data, evaluated on the basis of M measurements for each variable. Use $N \times M$ matrices $X = (x_{it})$ with i.i.d. entries x_{it} to compute:

$$C_{ij} = \frac{1}{M}(XX^T)_{ij} = \frac{1}{M}\sum_{t=1}^{M} x_{it}x_{jt} \ .$$

  Expect finite sample fluctuation around mean.

$$C_{ij} = \langle x_i x_j \rangle \pm \mathcal{O}(1/\sqrt{M}) = \delta_{ij} \pm \mathcal{O}(1/\sqrt{M})$$

- Spectrum of C as $N \to \infty$, $M \to \infty$ @ fixed $\alpha = N/M$?
  $\Rightarrow$ Marčenko Pastur-Law

$$\rho_\alpha(\lambda) = \frac{1}{2\pi\alpha\lambda}\sqrt{4\alpha - (\lambda - (1+\alpha))^2}$$

Principal Differences

- Rows of X for the auto-covariance problem are sections of a single time series $(x_t)_{t\in\mathbb{Z}}$

$$X = \begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_M \\ x_2 & x_3 & x_4 & \cdots & x_{1+M} \\ \vdots & & & \ddots & \vdots \\ x_N & x_{N+1} & x_{N+2} & \cdots & x_{N+M} \end{pmatrix}$$

- Number of random variables in the problem is $\mathcal{O}(N)$, rather than $\mathcal{O}(N^2)$ as in the Wishart Laguerre ensemble.

- Extensive body of knowledge about the Wishart-Laguerre ensemble and its variants (applications in multivariate statistics, signal-processing, finance, . . .)

- Little is known about the auto-covariance problem

  Basak, Bose & Sen (2011): Existence of Spectral Density for MA-k processes

  Bryc, Dembo & Jiang (2006) Existence of Spectral Density for Random Toeplitz matr.

Spectral Density and Resolvent

- Spectral density of sample covariance matrix from resolvent

$$\rho(\lambda) = \lim_{N\to\infty} \frac{1}{\pi N} \mathrm{Im}\ \mathrm{Tr} \left\langle \left[\lambda_\varepsilon \mathbb{1}\mathrm{I} - C\right]^{-1} \right\rangle , \qquad \lambda_\varepsilon = \lambda - i\varepsilon$$

- express as (S F Edwards & R C Jones, JPA, 1976)

$$\begin{aligned} \rho(\lambda) &= \lim_{N\to\infty} \frac{1}{\pi N} \mathrm{Im}\ \frac{\partial}{\partial\lambda}\ \mathrm{Tr} \left\langle \ln \left[\lambda_\varepsilon \mathbb{1}\mathrm{I} - C\right] \right\rangle \\ &= \lim_{N\to\infty} -\frac{2}{\pi N} \mathrm{Im}\ \frac{\partial}{\partial\lambda} \left\langle \ln Z_N \right\rangle , \end{aligned}$$

  where Z_N is a Gaussian integral:

$$Z_N = \int \prod_{k=1}^{N} \frac{\mathrm{d}u_k}{\sqrt{2\pi/i}}\ \exp\left\{ -\frac{i}{2} \sum_{k,\ell} u_k (\lambda_\varepsilon \delta_{k\ell} - C_{k\ell}) u_\ell \right\}$$

Performing the Average

- Standard Approach – Replica Method

$$\Big\langle \ln Z_N \Big\rangle = \lim_{n\to 0} \frac{1}{n} \ln \Big\langle Z_N^n \Big\rangle$$

- For integer n, Z_N^n is partition function of n identical copies of the system (n-th power of Gaussian integral)

- Experience: final result has structure of replica-symmetric high-temperature solution $\Leftrightarrow$ annealed calculation ($n = 1$).
  $\Rightarrow$ Do annealed calculation from the start ($\langle \ln Z_N \rangle \simeq \ln \langle Z_N \rangle$)

$$\langle Z_N \rangle = \left\langle \int \prod_k \frac{\mathrm{d}u_k}{\sqrt{2\pi/\mathrm{i}}} \, \exp\left\{ -\frac{\mathrm{i}}{2}\lambda_\varepsilon \sum_k u_k^2 + \frac{\mathrm{i}}{2} \sum_{k\ell} C_{k\ell} u_k u_\ell \right\} \right\rangle \ ,$$

- Insert definition of C

$$\langle Z_N \rangle = \left\langle \int \prod_k \frac{\mathrm{d}u_k}{\sqrt{2\pi/\mathrm{i}}} \, \exp\left\{ -\frac{\mathrm{i}}{2}\lambda_\varepsilon \sum_k u_k^2 + \frac{\mathrm{i}}{2}\alpha \sum_{i=1}^{M} \left(\frac{1}{\sqrt{N}} \sum_k x_{k+i} u_k \right)^2 \right\} \right\rangle$$

- Disorder dependence of Z_N only through the variables

$$z_i = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} x_{k+i} u_k \ .$$

- By CLT (for weakly dependent rv's) normally distributed for large M with

$$\langle z_i \rangle = 0 \ , \qquad \langle z_i z_j \rangle = \frac{1}{N} \sum_{k\ell} \langle x_{k+i} x_{\ell+j} \rangle u_k u_\ell \equiv Q_{i,j}$$

and Q given in terms of true process auto-covariance

$$Q_{i,j} = \langle z_i z_j \rangle = \frac{1}{N} \sum_{k\ell} \bar{C}(i-j+k-\ell)\, u_k u_\ell$$

- $\{z_i\}$ average is Gaussian

$$\langle Z_N \rangle = \int \prod_k \frac{\mathrm{d}u_{ka}}{\sqrt{2\pi/\mathrm{i}}} \exp\left\{ -\frac{\mathrm{i}}{2}\lambda_\varepsilon \sum_k u_k^2 - \frac{1}{2}\ln\det(\mathbb{1} - \mathrm{i}\alpha Q) \right\}$$

- Q is a Toeplitz matrix.
  $\Rightarrow$ evaluate spectral sum ln det $(\mathbb{1} - \mathrm{i}\alpha Q)$ using

- **Szegö's theorem**: Given an $N \times N$ Toeplitz matrix A with elements $A_{ik} = a(i-k)$, where $a = a(n) \in \ell_1(\mathbb{Z})$. Then the spectral density has a (weak) limit

$$\rho_N(\lambda) = \frac{1}{N}\sum_{i=1}^{N} \delta(\lambda - \lambda_i) \longrightarrow \int_{-\pi}^{\pi} \frac{\mathrm{d}q}{2\pi} \delta(\lambda - \widehat{a}(q)) \ ,$$

  as $N \to \infty$, where $\widehat{a}(q)$ is called the 'symbol', and is nothing but the Fourier transform of a

$$\widehat{a}(q) = \sum_{n=-\infty}^{\infty} a(n)\mathrm{e}^{\mathrm{i}qn} \ .$$

- Szegö (keeping track of finite-M finite-N expressions)

$$\ln\det(\mathbb{1}\!\mathrm{I} - \mathrm{i}\alpha Q) \sim \sum_{\mu=-(M-1)/2}^{(M-1)/2} \ln\Big(1 - \mathrm{i}\alpha Q_\mu\Big)$$

where

$$Q_\mu = \frac{1}{N}\sum_{k\ell} \widehat{C}(q_\mu)\mathrm{e}^{-\mathrm{i}q_\mu(k-\ell)}\, u_k u_\ell = \widehat{C}(q_\mu)|\widehat{u}(q_\mu)|^2 \equiv Q(q_\mu)$$

with

$$\widehat{u}(q_\mu) = \frac{1}{\sqrt{N}}\sum_{k=1}^{N} \mathrm{e}^{\mathrm{i}q_\mu k}\, u_k \ , \quad q_\mu = \frac{2\pi}{M}\mu$$

- Enforce Q_μ definitions using δ-functions and their Fourier representations. $\Rightarrow$ get Gaussian u_k-integrals.

- Results involve similar spectral sum $\ln\det(\lambda_\varepsilon \mathbb{1}\!\mathrm{I} - R)$, with a matrix R which, too, is Toeplitz.

- Allow closed form expression of $\langle Z_N \rangle$,

$$\langle Z_N \rangle \simeq \prod_{\nu=0}^{(N-1)/2} \left\{ \frac{2\,\mathrm{i}}{\alpha \widehat{C}(p_\nu)} \int_0^\infty \mathrm{d}y \, \frac{\mathrm{e}^{-\mathrm{i}y\lambda_\varepsilon 2/(\alpha \widehat{C}(p_\nu))}}{\left(1-\mathrm{i}y\right)^{2/\alpha}} \right\}$$

- and hence $\rho(\lambda)$:

$$\rho(\lambda) = \int_0^\pi \frac{\mathrm{d}q}{\pi} \frac{1}{\widehat{C}(q)} \rho_\alpha^{(0)}\left(\frac{\lambda}{\widehat{C}(q)}\right)$$

with

$$\rho_\alpha^{(0)}(\lambda) = -\lim_{\varepsilon \to 0} \frac{1}{\pi} \mathrm{Im} \frac{\partial}{\partial \lambda} \ln I_\alpha\left(\frac{2}{\alpha}\lambda_\varepsilon\right)$$

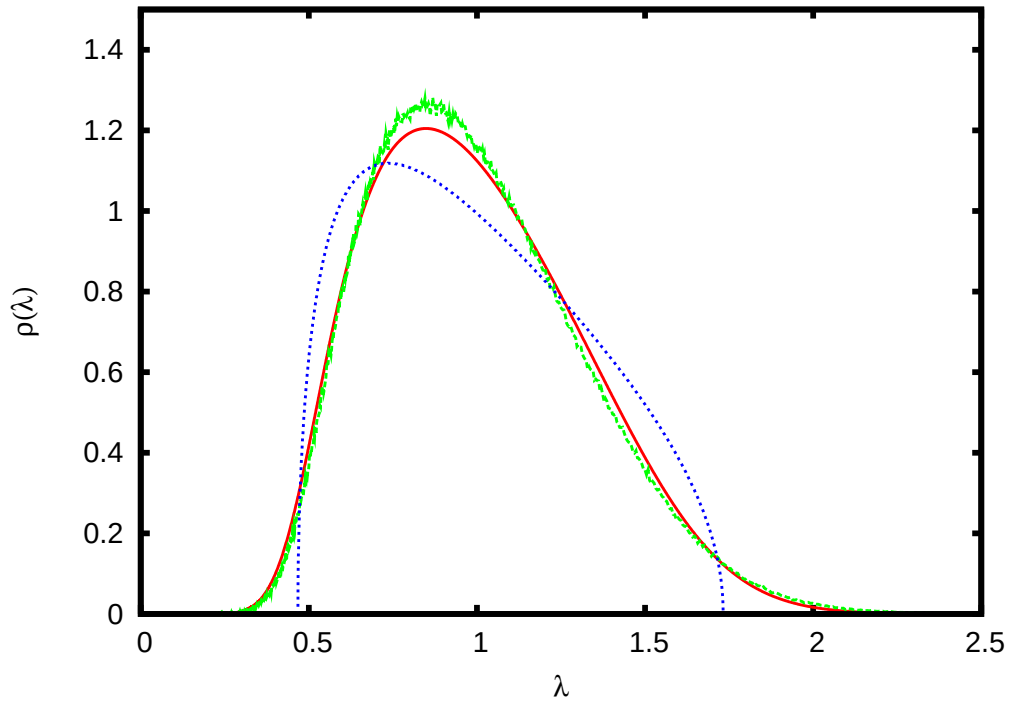
and I_α expressible in terms of an incomplete Γ-function

$$I_\alpha(x) = \mathrm{i}\,(-x)^{-1+2/\alpha}\,\mathrm{e}^{-x}\Gamma(1-2/\alpha,-x)\ , \quad \mathrm{Im}\,x < 0\ .$$

- Have to identify $\rho_\alpha^{(0)}$ with spectral density for auto-covariance matrices of sequences of i.i.d. (uncorrelated) data, for which $\widehat{C}(q) \equiv 1$.

Numerical Tests

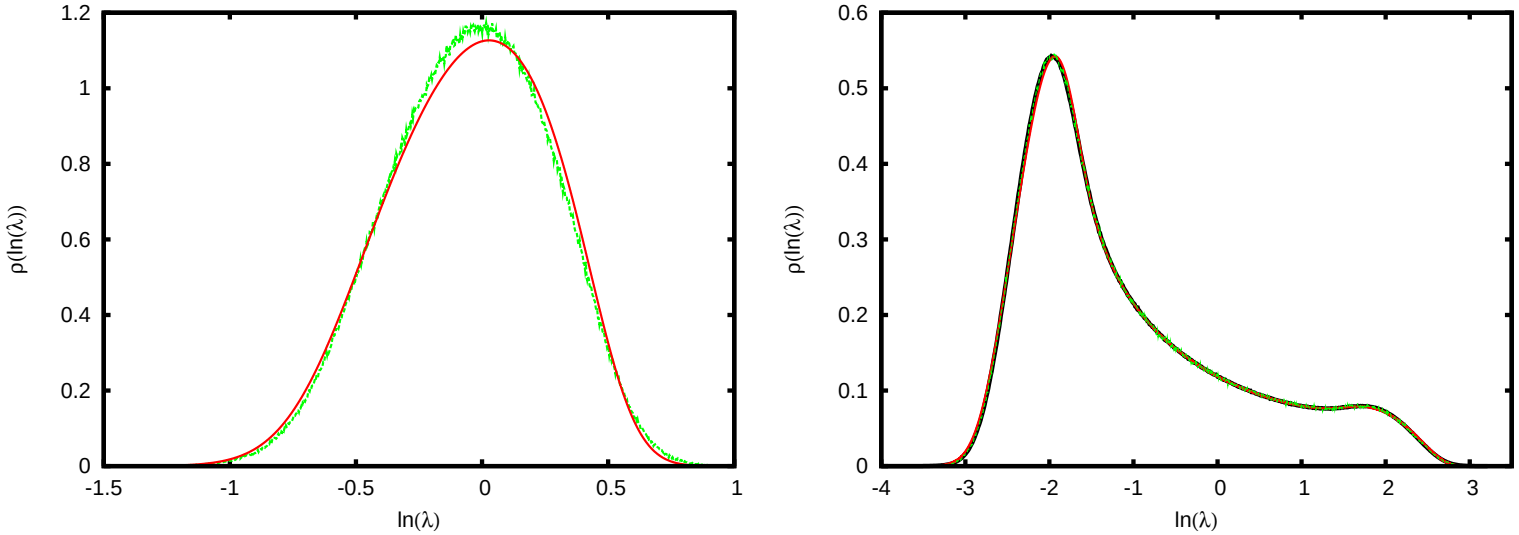
- Spectral density for $x_n \sim \mathcal{N}(0,1)$ i.i.d. @ $\alpha = 0.1$



Simulation results (green); analytic approximation for $\rho_\alpha^{(0)}(\lambda)$ (red), Marčenko-Pastur law (blue-dashed).

- (Logarithmic) Spectral density for AR-1 process @ $\alpha = 0.1$

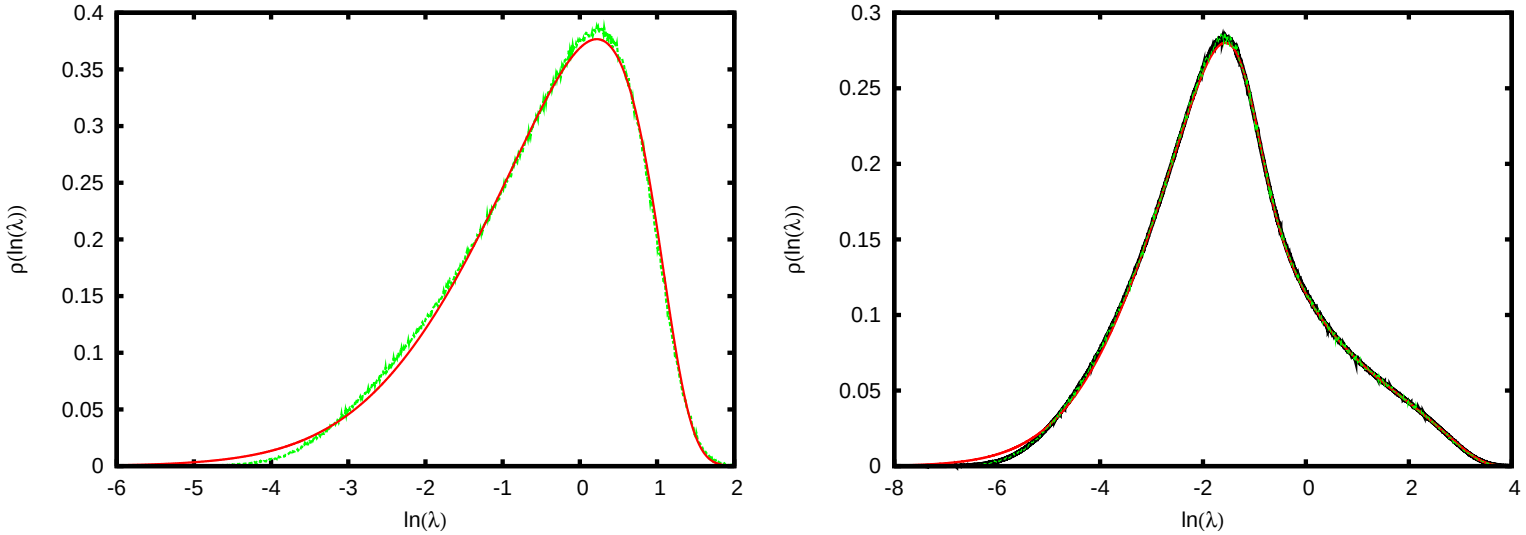
$$x_n = a_1\, x_{n-1} + \sqrt{1 - a_1^2}\, \xi_n$$



Left: $a_1 = 0$ (i.i.d. data), simulation (green) and analytic result (red). **Right**: $a_1 = 0.8$. Comparing scaling based on the empirical scaling function (black) with that based on the analytic result (red) and simulations (green).

- (Logarithmic) Spectral density for AR-1 process @ $\alpha = 0.8$

$$x_n = a_1\, x_{n-1} + \sqrt{1 - a_1^2}\, \xi_n$$

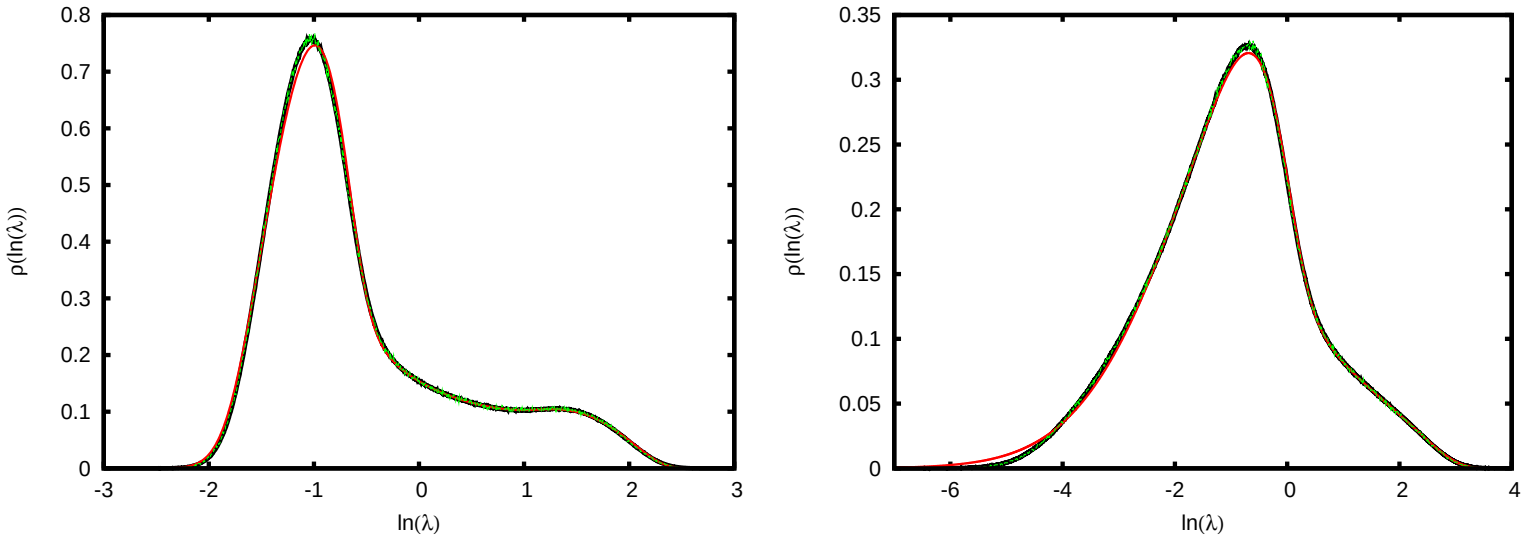


Left: $a_1 = 0$ (i.i.d. data), simulation (green) and analytic result (red). **Right** $a_1 = 0.8$. Comparing scaling based on the empirical scaling function (black) with that based on the analytic result (red) and simulations (green).

- (Logarithmic) Spectral density for AR-2 processes

$$x_n + a_1 x_{n-1} + a_2 x_{n-2} = \sigma \xi_n \ , \quad \sigma \text{ s.t. } \bar{C}(0) = 1 \ ,$$

$$a_1 = 0.5 \ , \quad a_2 = -3/16 \ .$$

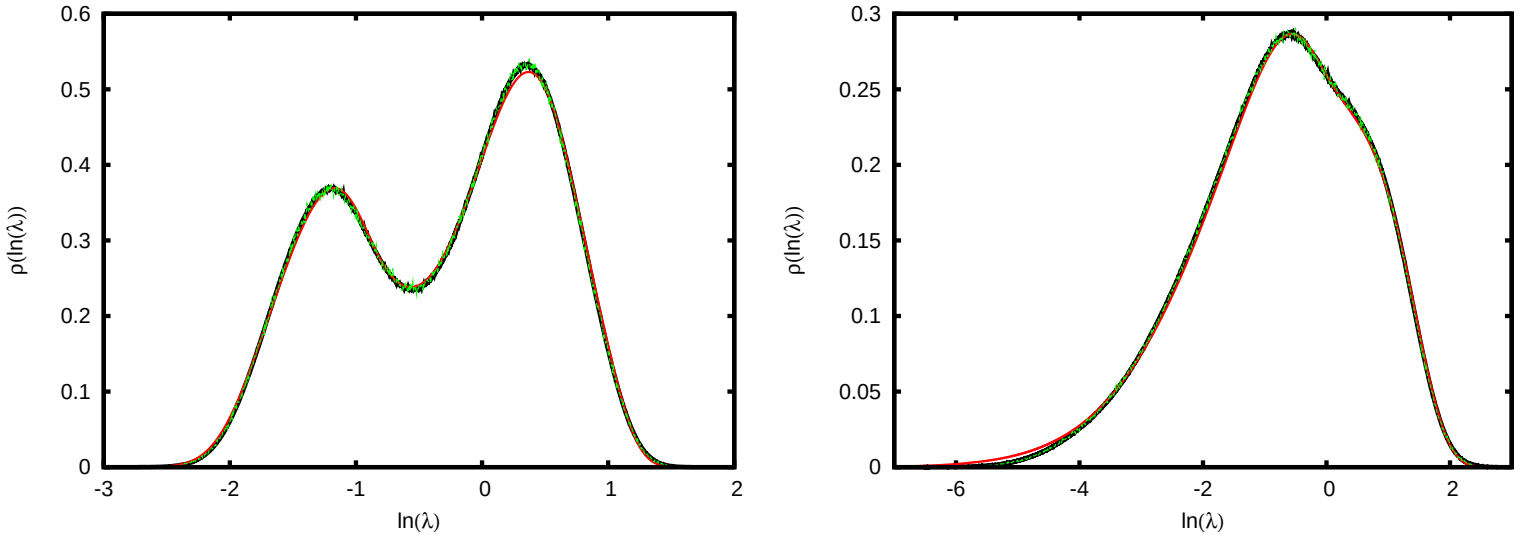


Comparing scaling based on the empirical scaling function (black) with that based on the analytic result (red) and simulations (green). **Left**: $\alpha = 0.1$, **Right**: $\alpha = 0.8$.

- (Logarithmic) Spectral density for AR-2 processes

$$x_n + a_1 x_{n-1} + a_2 x_{n-2} = \sigma \xi_n \ , \quad \sigma \text{ s.t. } \bar{C}(0) = 1 \ ,$$

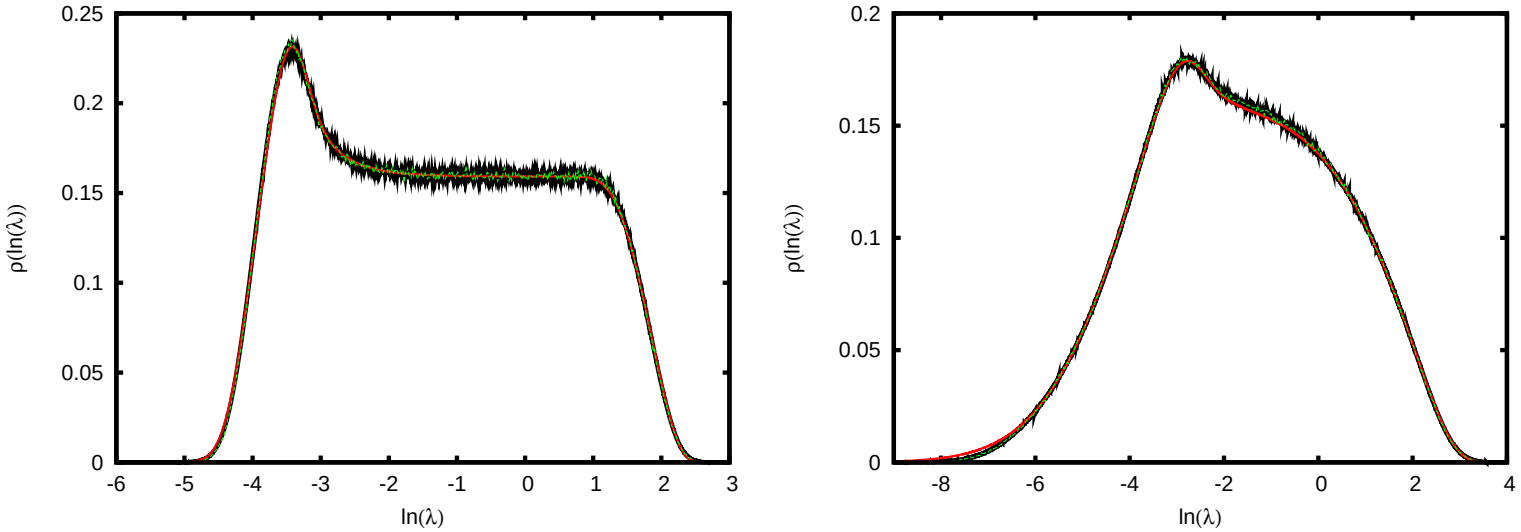
$$a_1 = 0.5 \ , \quad a_2 = 5/16 \ .$$



Comparing scaling based on the empirical scaling function (black) with that based on the analytic result (red) and simulations (green). **Left**: $\alpha = 0.1$, **Right**: $\alpha = 0.8$.

- (Logarithmic) Spectral density for processes with long range auto-correlation

$$\bar{C}(k) = \frac{1}{1 + (k/2)^2} \ ,$$



Comparing scaling based on the empirical scaling function (black) with that based on the analytic result (red) and simulations (green). **Left**: $\alpha = 0.1$, **Right**: $\alpha = 0.8$.

Summary

- Computed DOS of sample auto-covariance matrices using annealed calculation.
- Key ingredient: Szegö's theorem for Toeplitz matrices
- Rectangular window and decorrelation approximation
  $\Rightarrow$ Closed form approximation.
- Use of Szegös theorem suggests a scaling form for DOS.
  – results suggest that scaling is exact
  – ideas for independent proof
- Applications: time-series analysis, signal processing, information theory, finance . . .
- **Thanks!** K. Anand, L. Dall'Asta, P. Vivo