Time-lagged partial correlations of financial time series with high dimensional conditions

Econophysics PY538

Sebastian Gemsheim

April 28, 2015

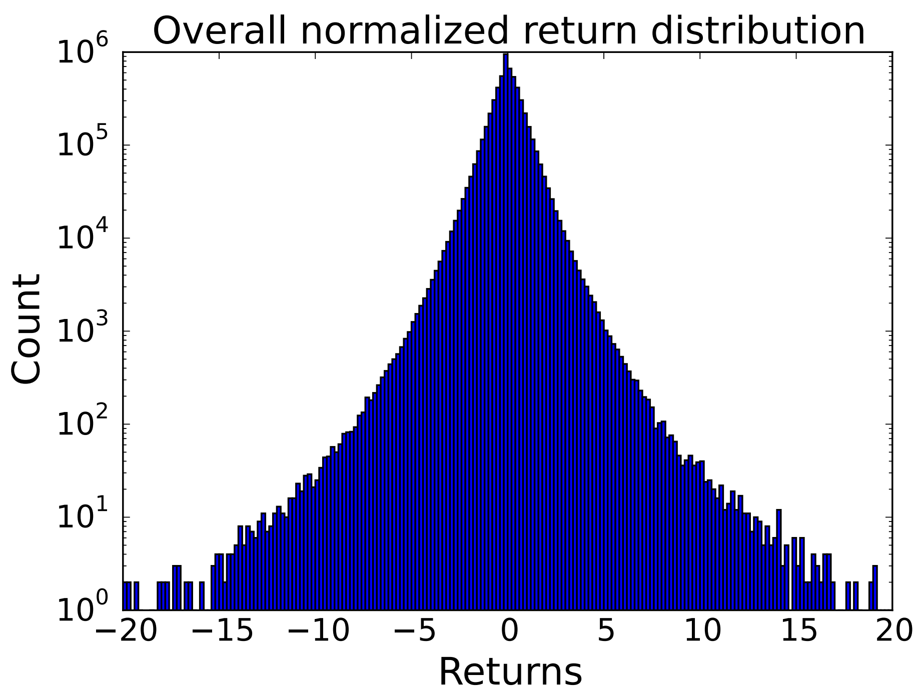
Outline

Data

- New York Stock Exchange 2001-2003
- Returns of the $N = 100$ largest capitalized stocks
- 748 trading days, 78 data points per day, 5 min interval
- Total: $T = 58344$ data points
- Data matrix X with dimension $N \times T$

Return distribution

- Rescaled data: zero mean, unit variance $x_i(t) = \frac{\tilde{x}_i(t) - \mu_{\tilde{x},i}}{\sigma_{\tilde{x},i}}$



Market mode

Covariance & Correlation matrix

$$\Sigma(X,X) = \rho(X,X) = \frac{1}{T} X X^T$$

with Eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots$ and eigenvectors $u_1, u_2, \dots$

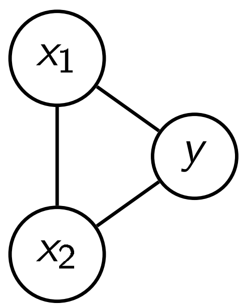
Market mode

$$x_m(t) = \sum_{j=1}^{N} u_{1j} x_j(t) \qquad \Rightarrow \qquad x_i(t) = \underbrace{\alpha_i}_{=0} + \beta_i x_m(t) + \epsilon_i(t)$$

$\rightarrow$ Market mode removed data X_{res} with $\epsilon_i(t)$

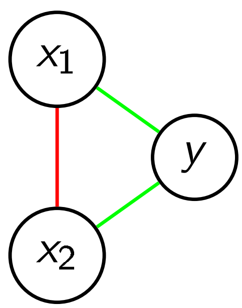
Partial Correlation

- **Question**: What is the correlation between two variables x_1, x_2 given y, a third one?



Partial Correlation

- **Question**: What is the correlation between two variables x_1, x_2 given y, a third one?



- **Answer**: Partial correlation $\rho(x_1, x_2 | y)$

Partial Correlation

Conditional mean

$$\hat{x}_i(y) = \underbrace{\mathbb{E}(x_i)}_{=0} + \frac{\sigma(x,y)}{\sigma(y,y)}\Big(y - \underbrace{\mathbb{E}(y)}_{=0}\Big)$$

Partial covariance

$$\sigma(x_1, x_2|y) = \mathrm{Cov}\left(x_1 - \hat{x}_1(y), x_2 - \hat{x}_2(y)\right)$$

Partial Correlation

Conditional mean for $X = \{x_1, x_2\}$ and $Y = \{y_1, y_2, \ldots, y_m\}$

$$\hat{X}(Y) = \Sigma_{XY}\Sigma_{YY}^{-1}Y$$

Partial covariance

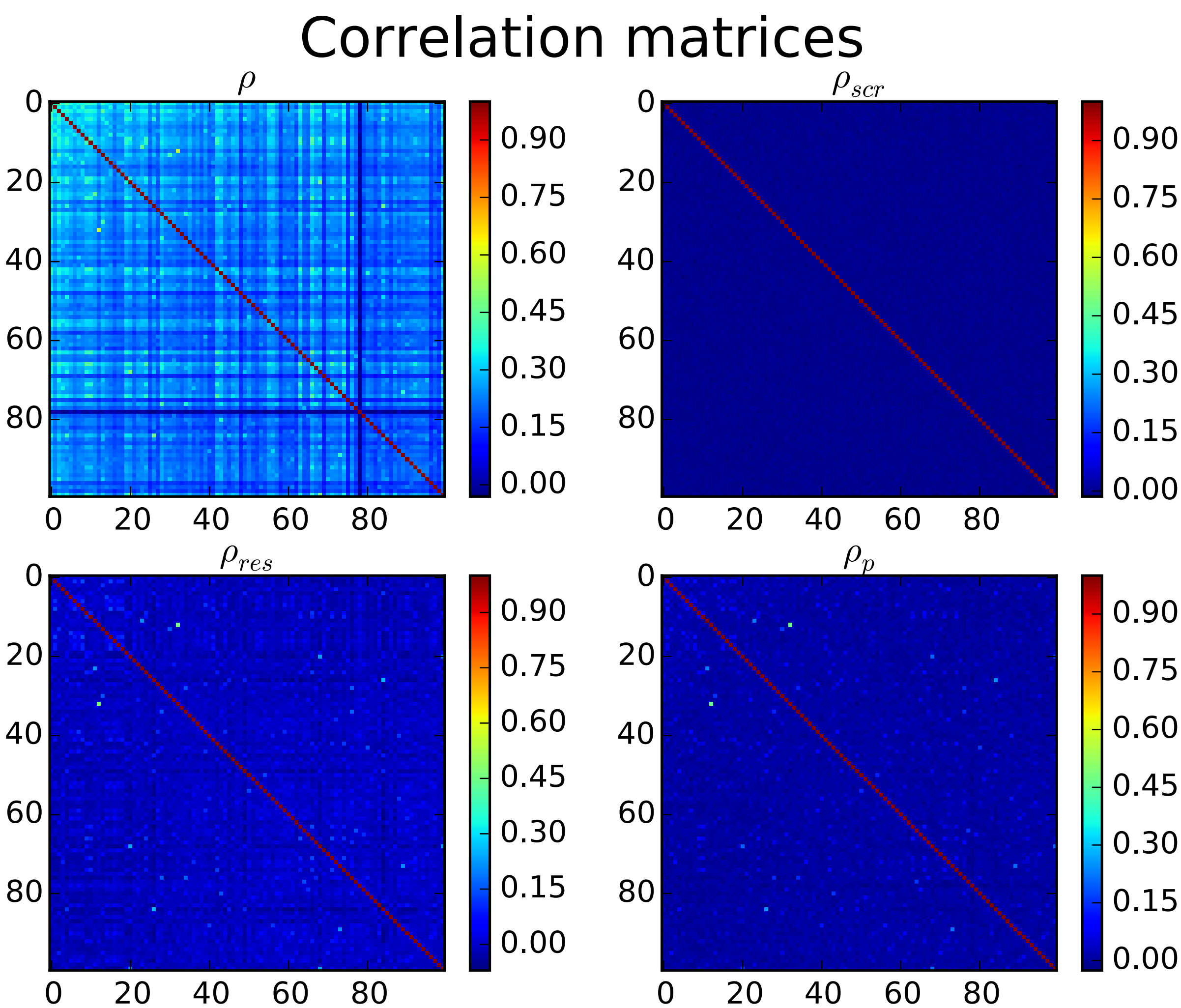
$$\Sigma_{XX|Y} = \mathrm{Cov}\left(X - \hat{X}(Y), X - \hat{X}(Y)\right)$$

$$= \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX} = \begin{pmatrix} \sigma_{11|Y} & \sigma_{12|Y} \\ \sigma_{21|Y} & \sigma_{22|Y} \end{pmatrix}$$

Partial correlation

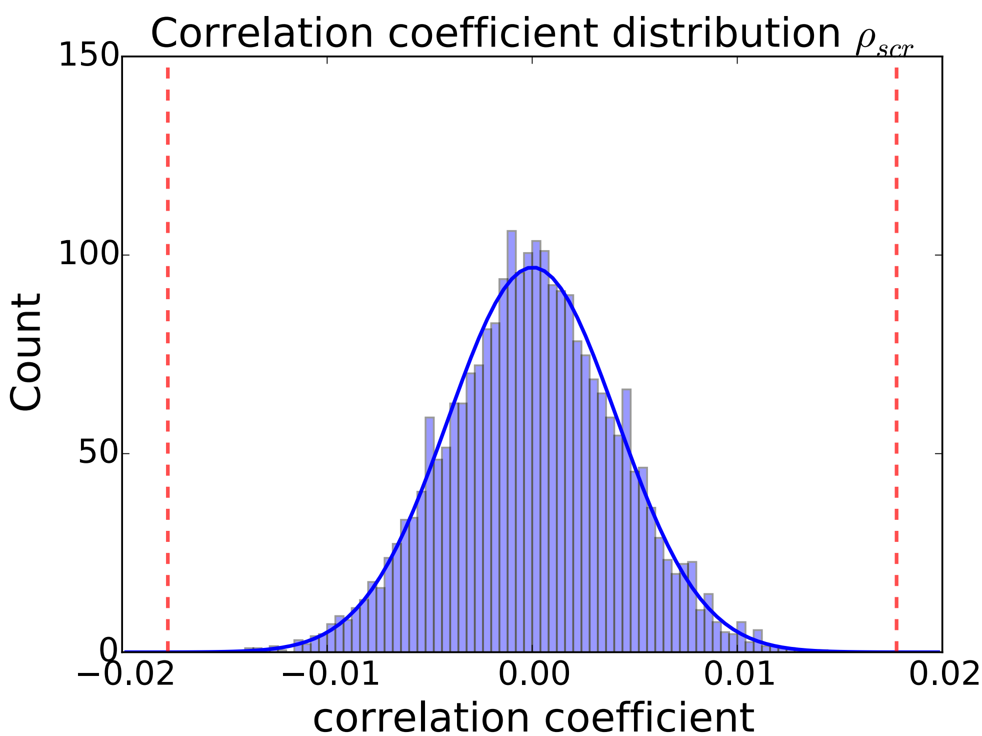
$$\rho_{12|Y} = \frac{\sigma_{12|Y}}{\sqrt{\sigma_{11|Y}\sigma_{22|Y}}}$$

Synchronous Correlation



Synchronous Correlation

- Noise limit: $\rho_{\max} \sim \sqrt{2 \ln \left(N^2\right) / T} = 0.01777$



Time-lagged correlation

Market mode removed data X_{res}

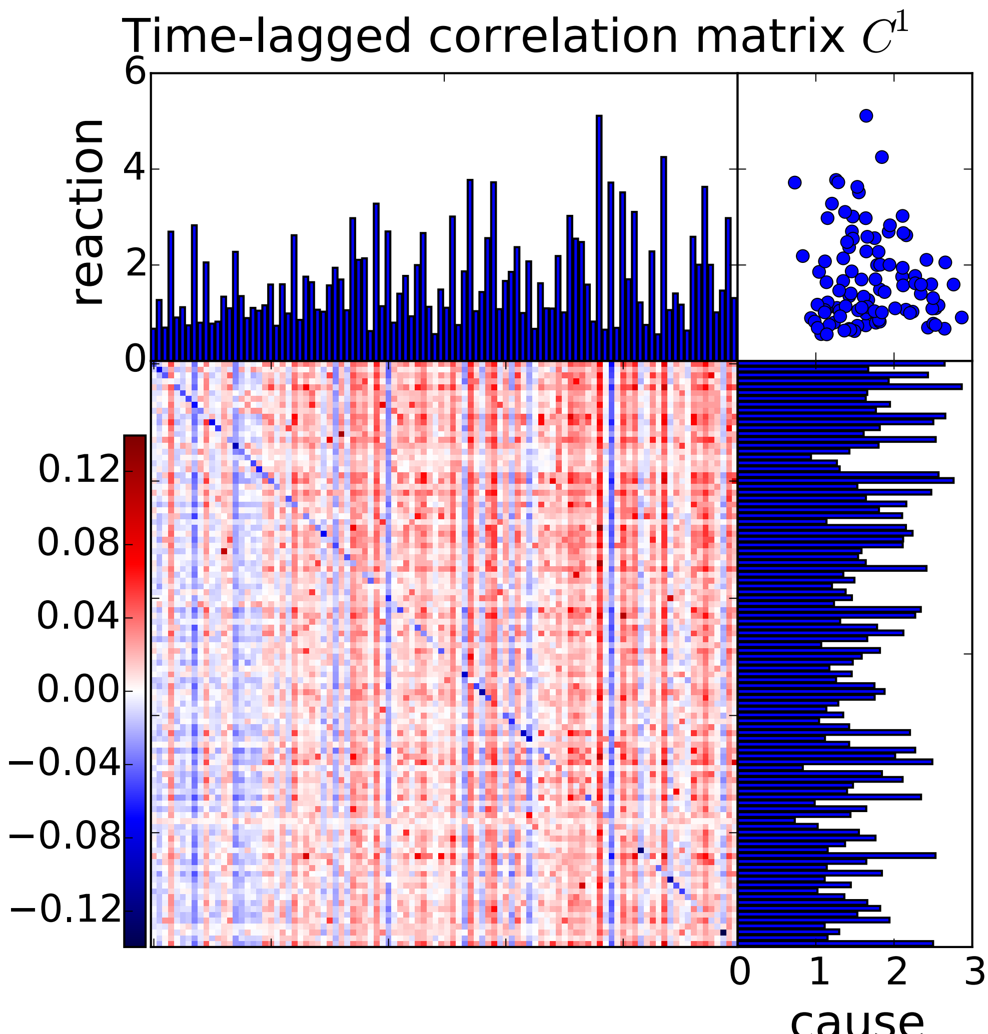
$$C^{\tau}_{res} = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau} X_{res}(t) X^{T}_{res}(t+\tau)$$

Time-lagged partial correlation

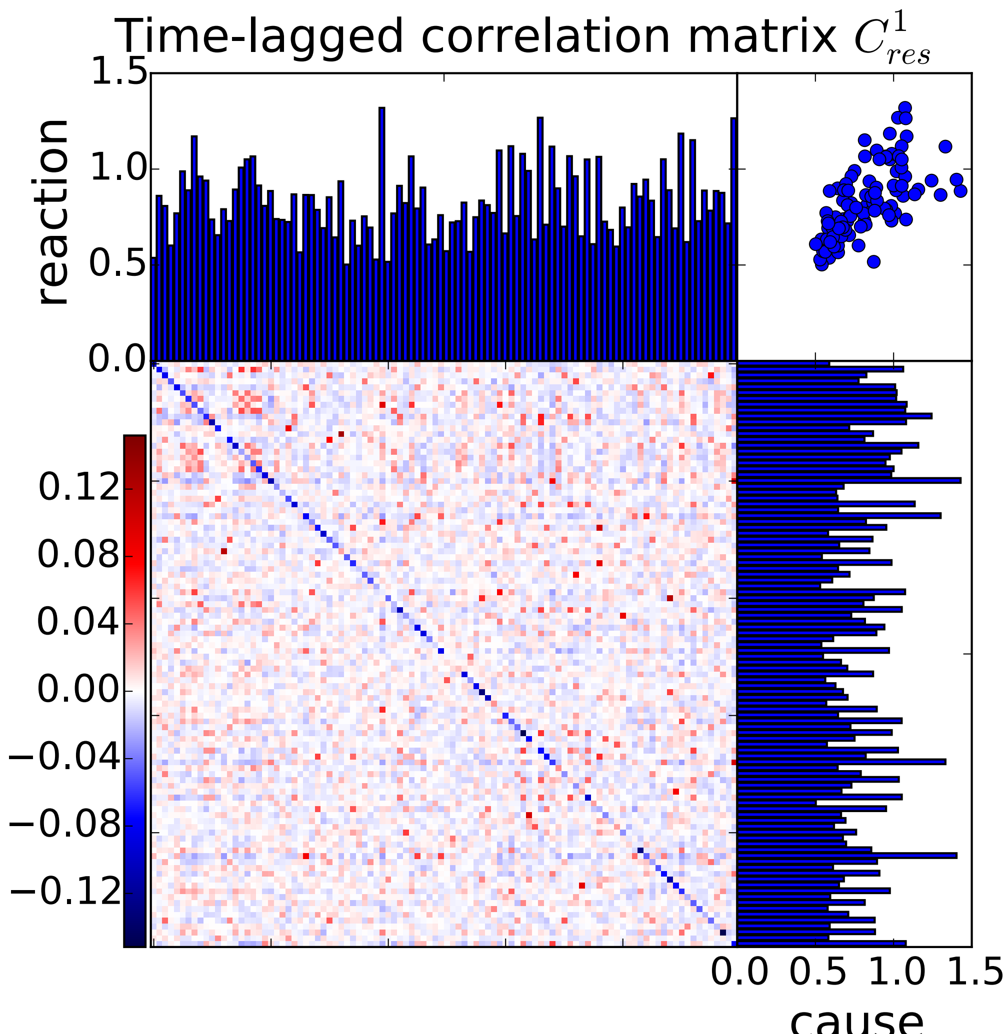
High dimensional condition vector, dim: $(\tau N - 2) \times (T - \tau)$

$$\begin{aligned} Y = \Big\{ & x_1(t), \ldots, x_{i-1}(t), x_{i+1}(t), \ldots, x_N(t), \ldots, \\ & x_1(t+(\tau-k)), \ldots, x_N(t+(\tau-k)), \ldots, \\ & x_1(t+\tau), \ldots, x_{j-1}(t+\tau), x_{j+1}(t+\tau), \ldots, x_N(t+\tau) \Big\} \end{aligned}$$

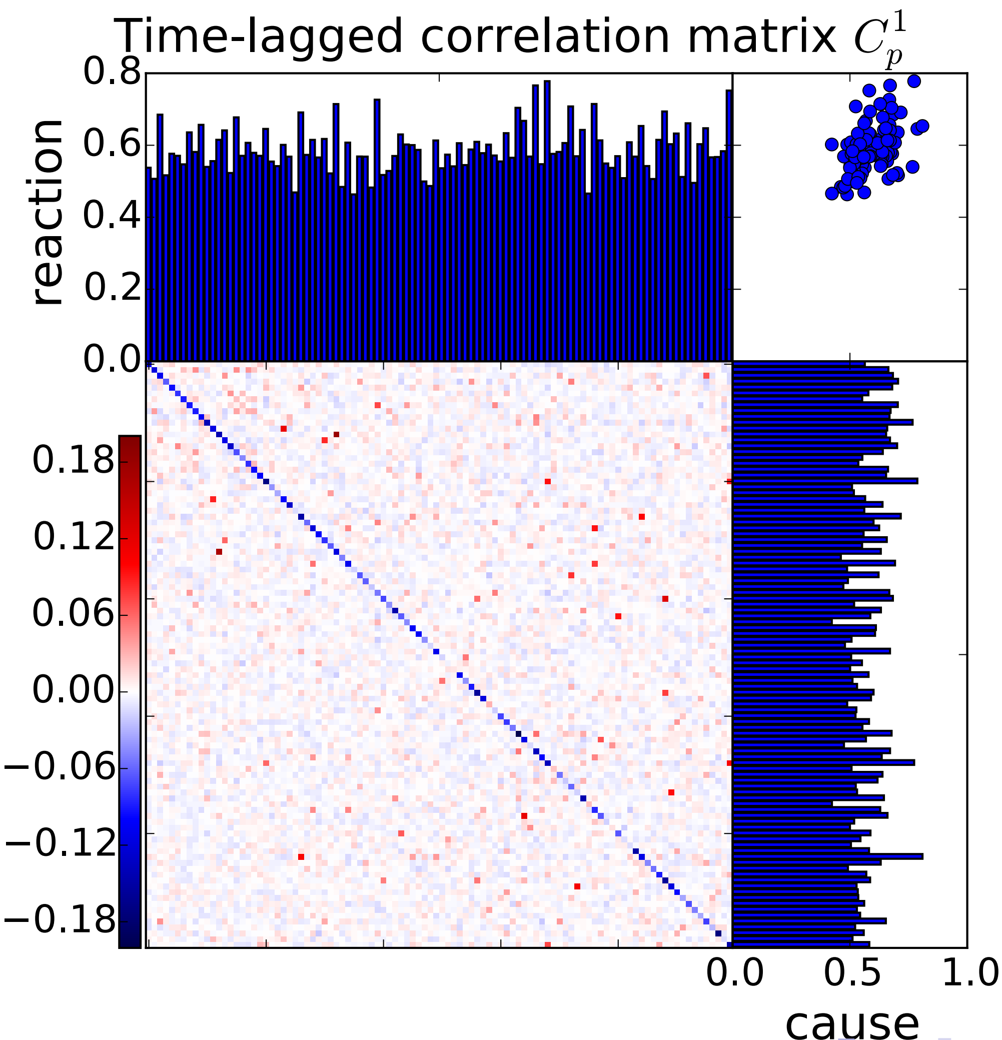
Time-lagged Correlation ($\tau = 1$)



Time-lagged Correlation ($\tau = 1$)

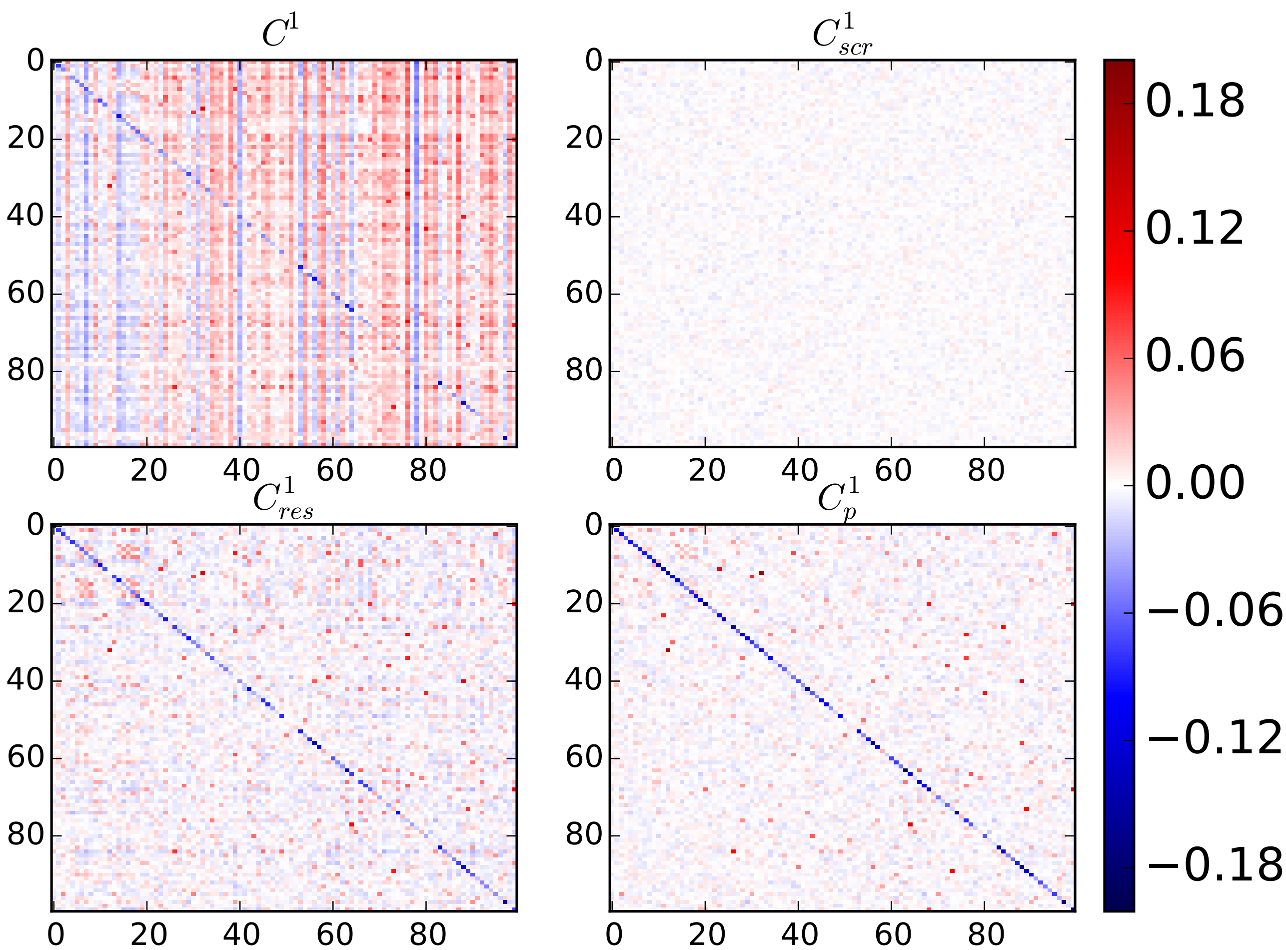


Time-lagged Correlation ($\tau = 1$)



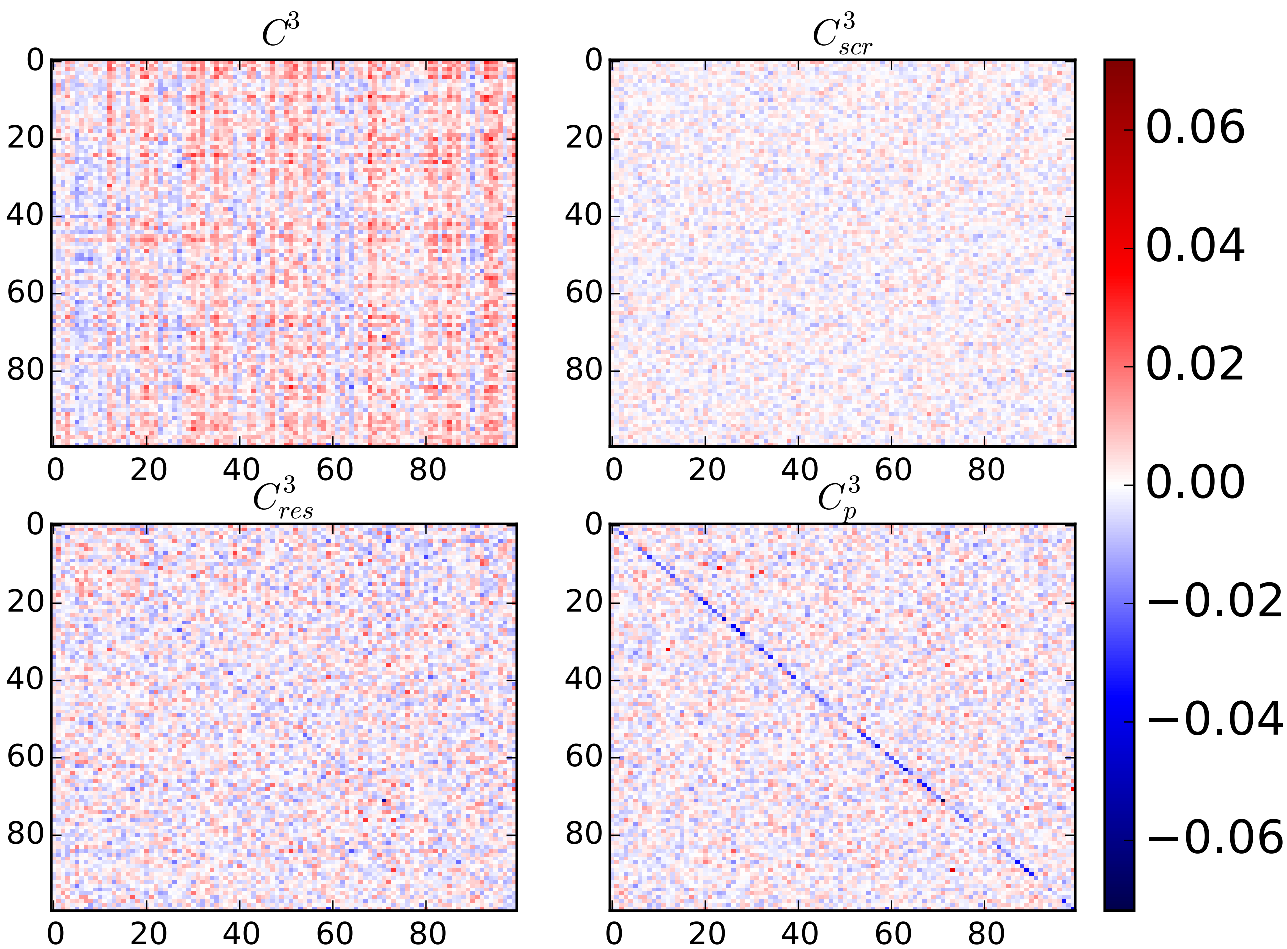
Time-lagged Correlation ($\tau = 1$)

Time-lagged Correlation matrices for lag 1



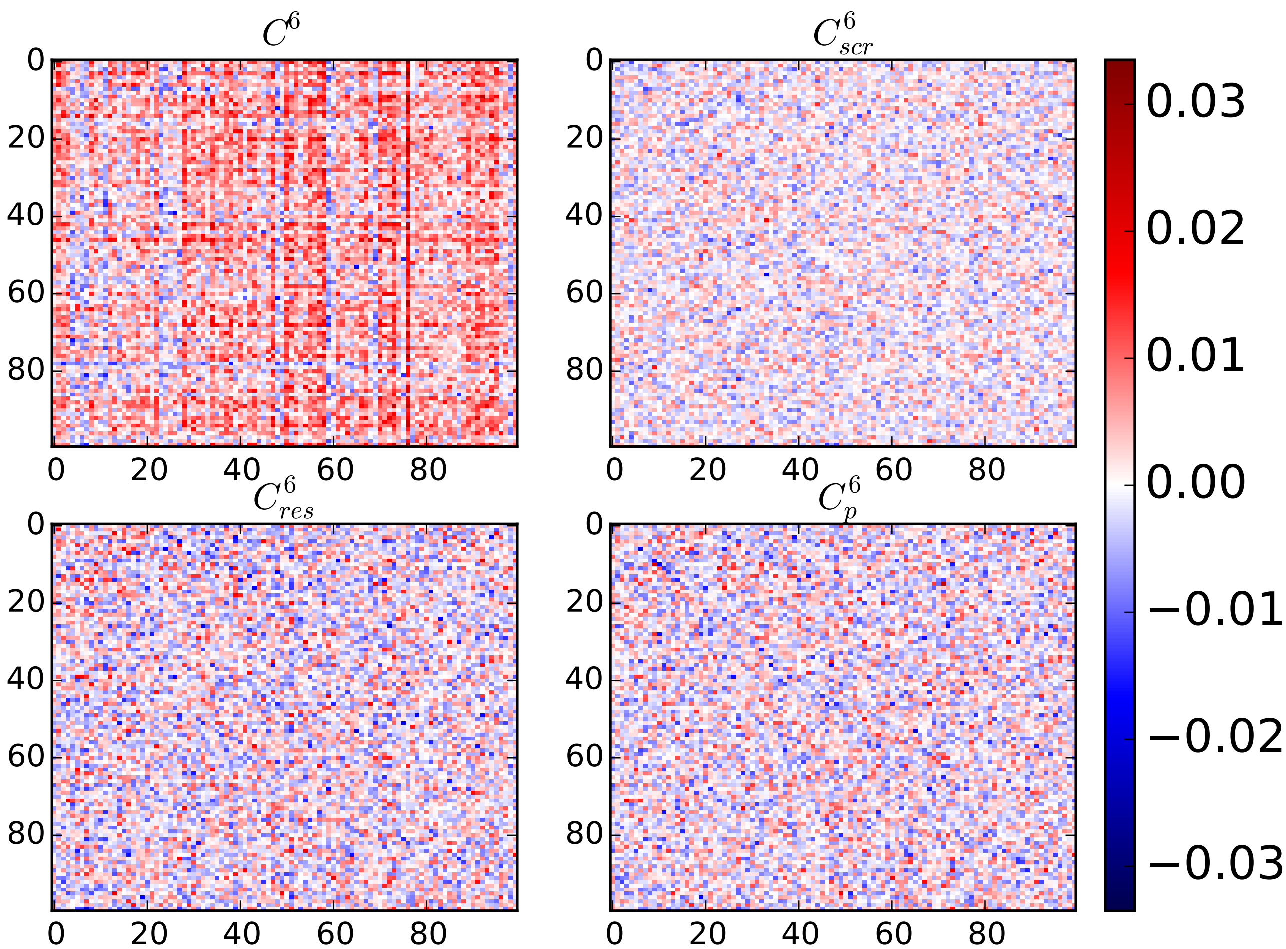
Time-lagged Correlation ($\tau = 3$)

Time-lagged Correlation matrices for lag 3



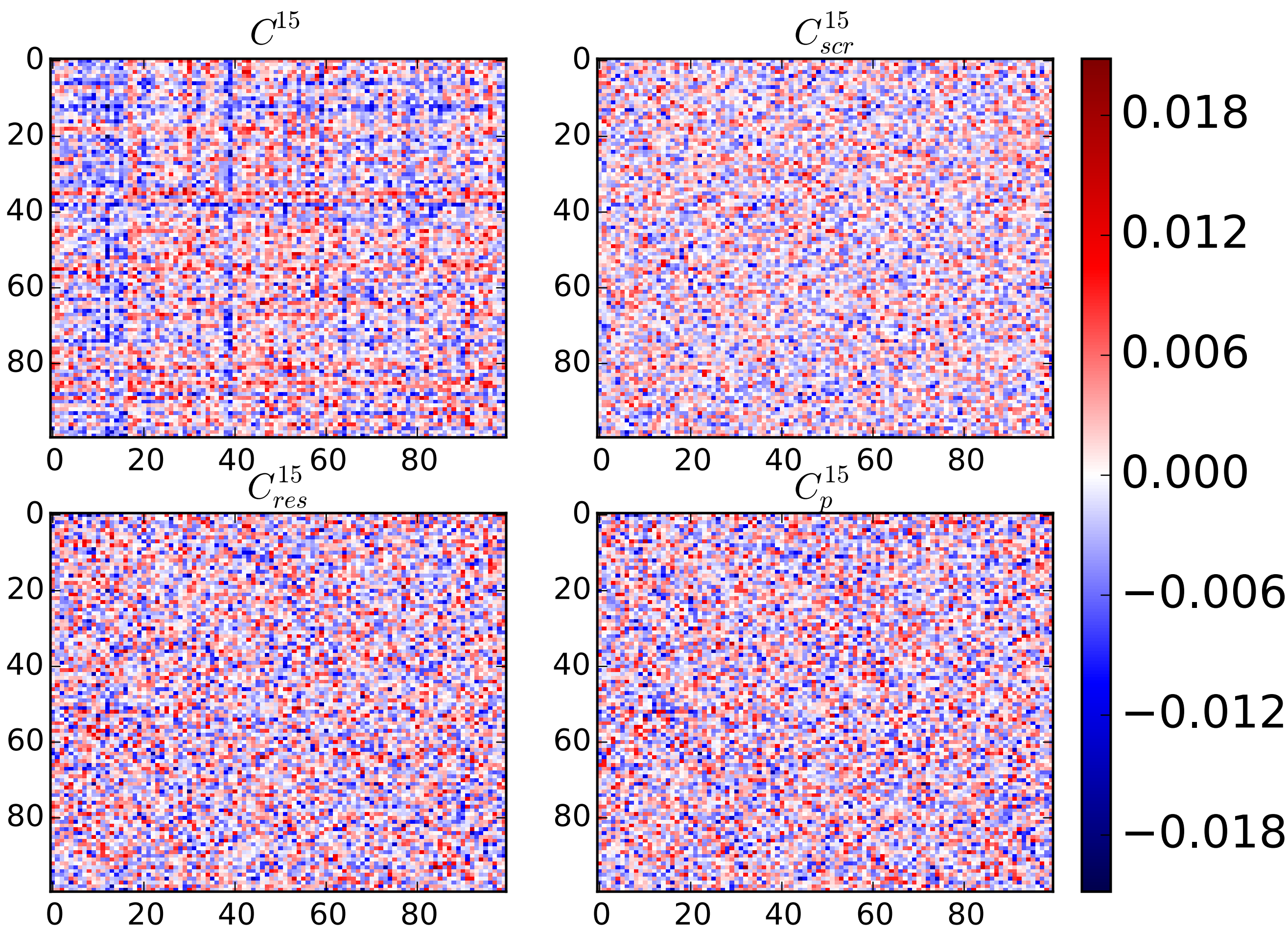
Time-lagged Correlation ($\tau = 6$)

Time-lagged Correlation matrices for lag 6

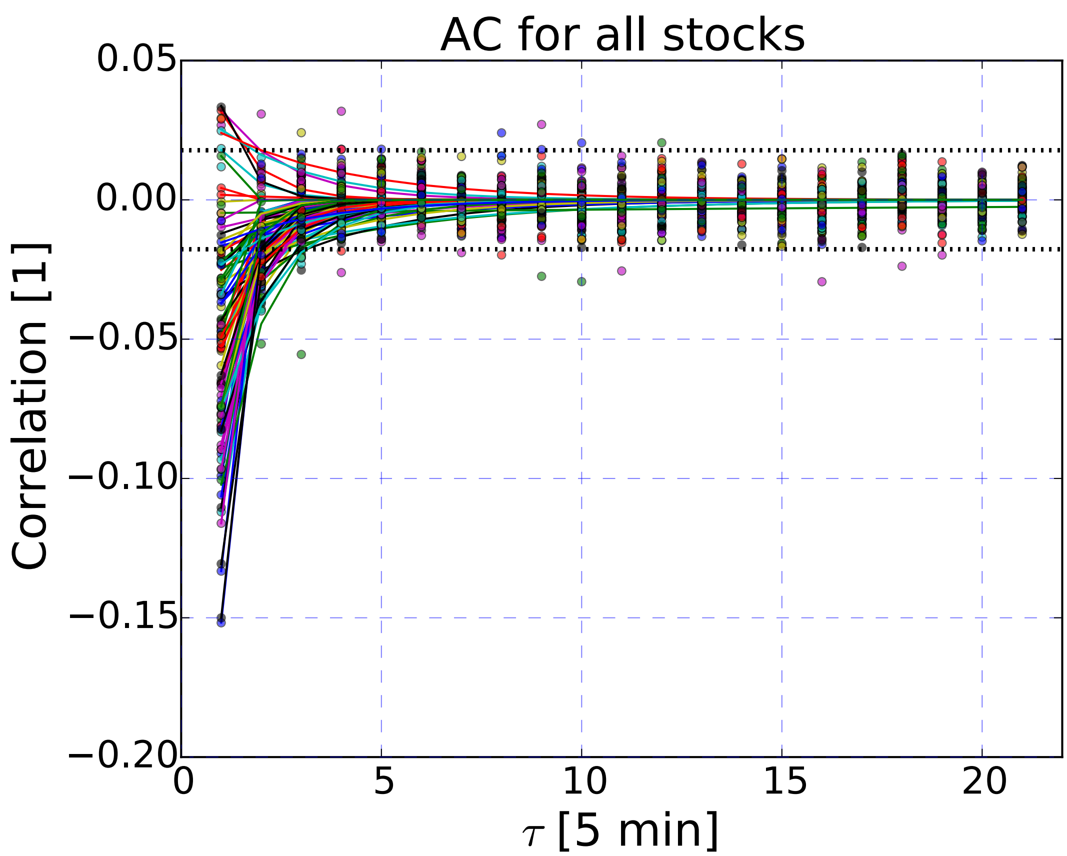


Time-lagged Correlation ($\tau = 15$)

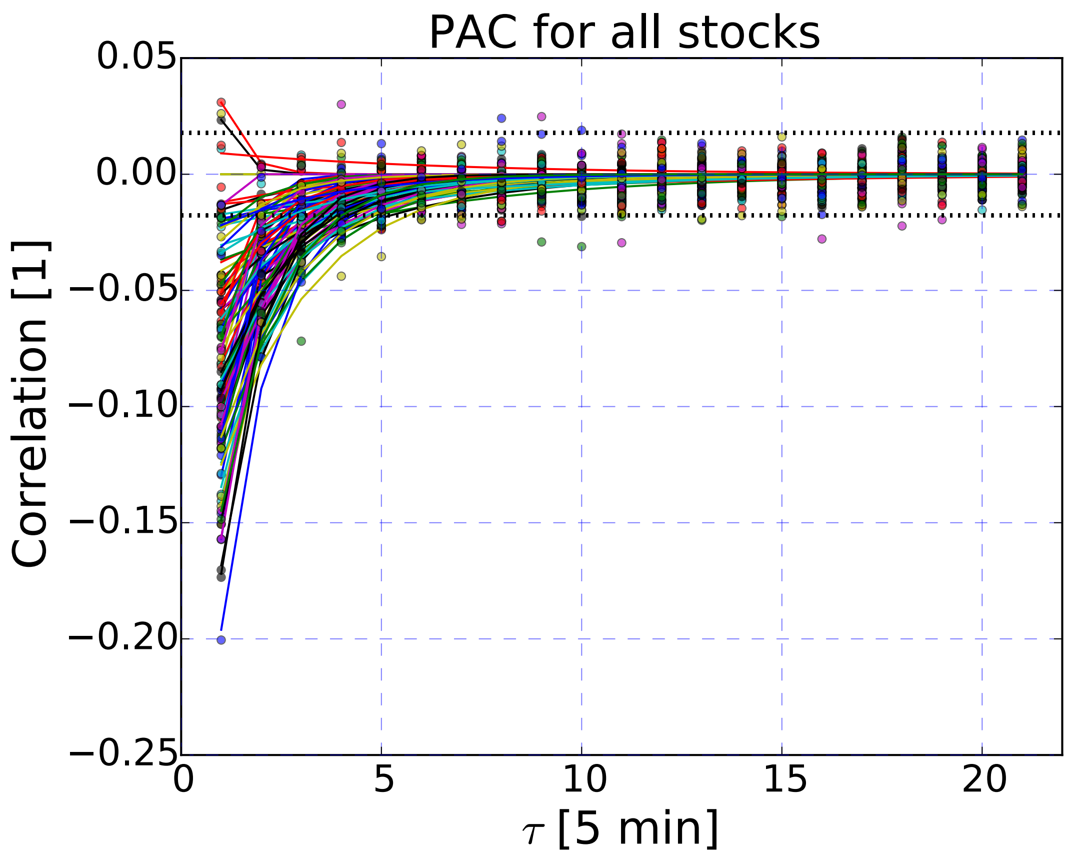
Time-lagged Correlation matrices for lag 15



Autocorrelations - market mode removed

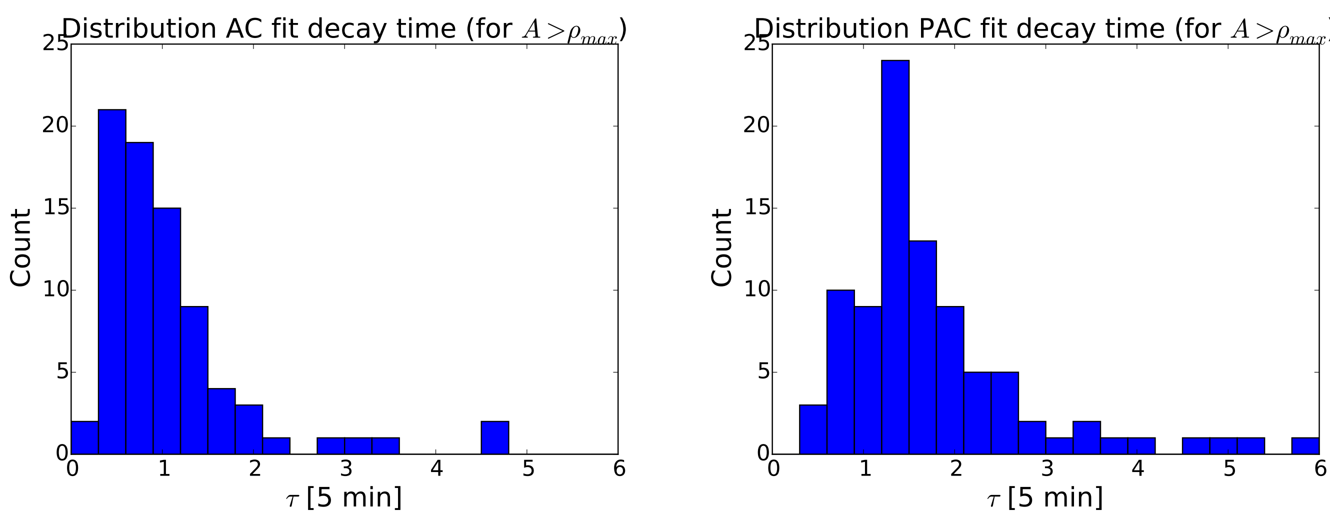


Partial autocorrelations



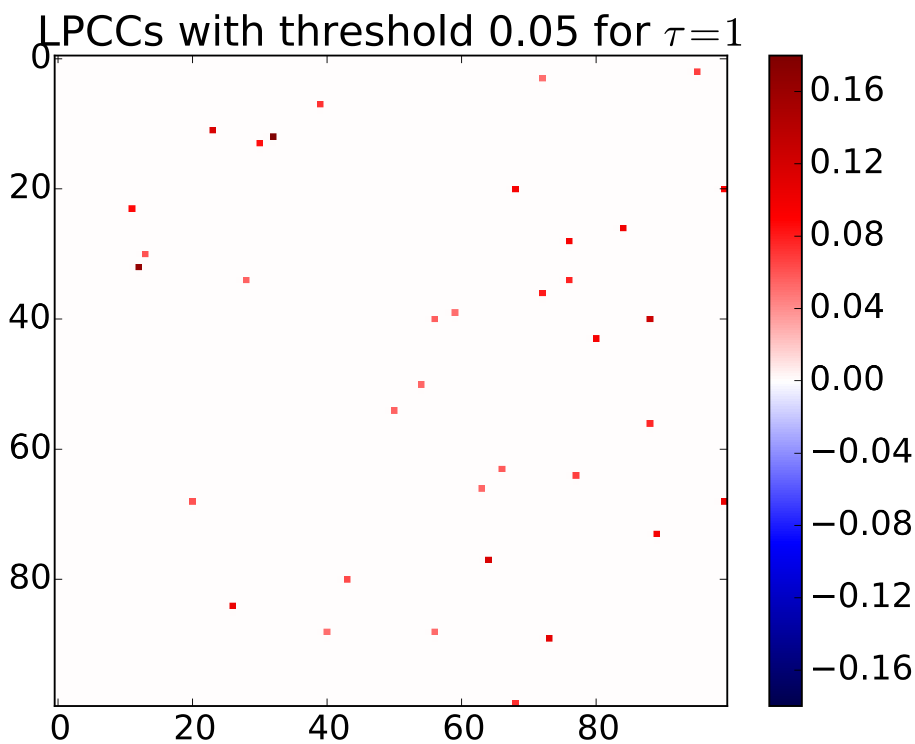
Fit parameter: exponential decay time

- Consider only if fit amplitude A is outside noise region
- AC: decay time $\sim 3 - 5\,\text{min}$
- PAC: decay time $\sim 7\,\text{min}$

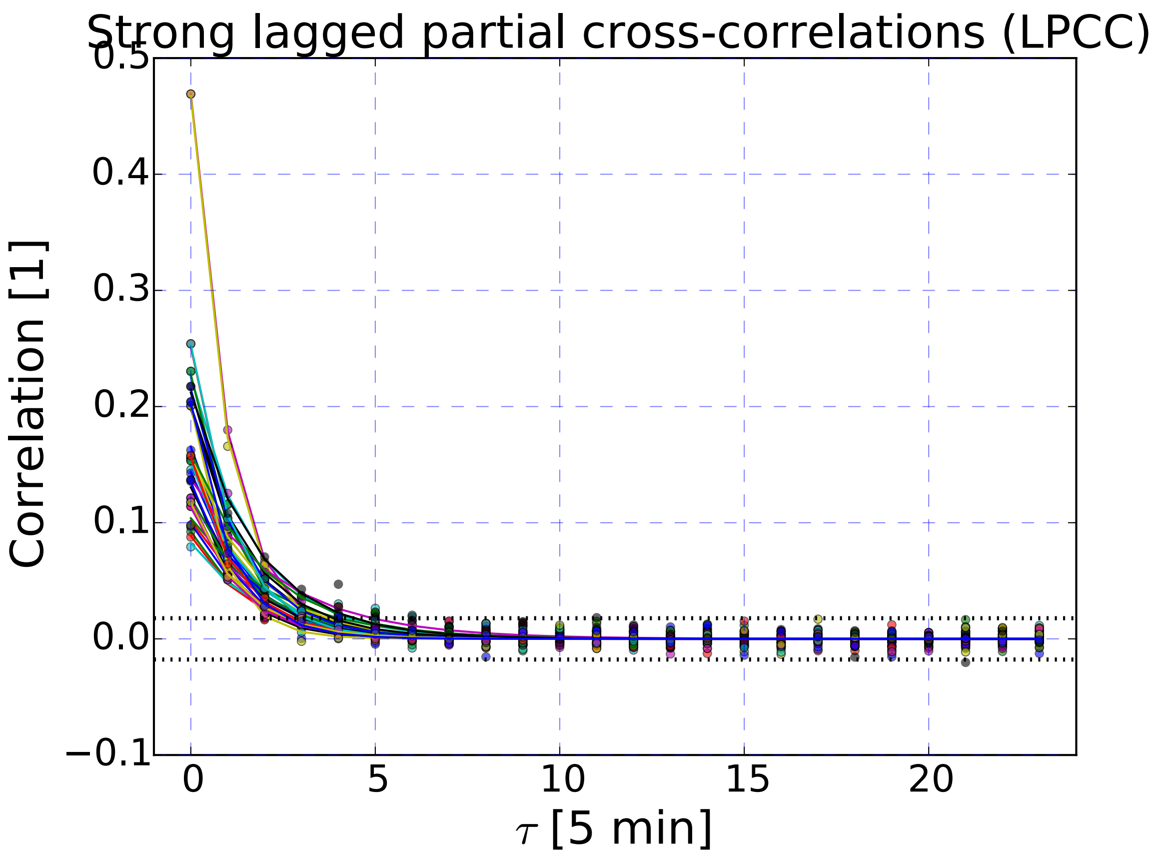


Strong partial cross-correlations

- Filter threshold for lag 1: $0.05 \approx 3\rho_{max}$

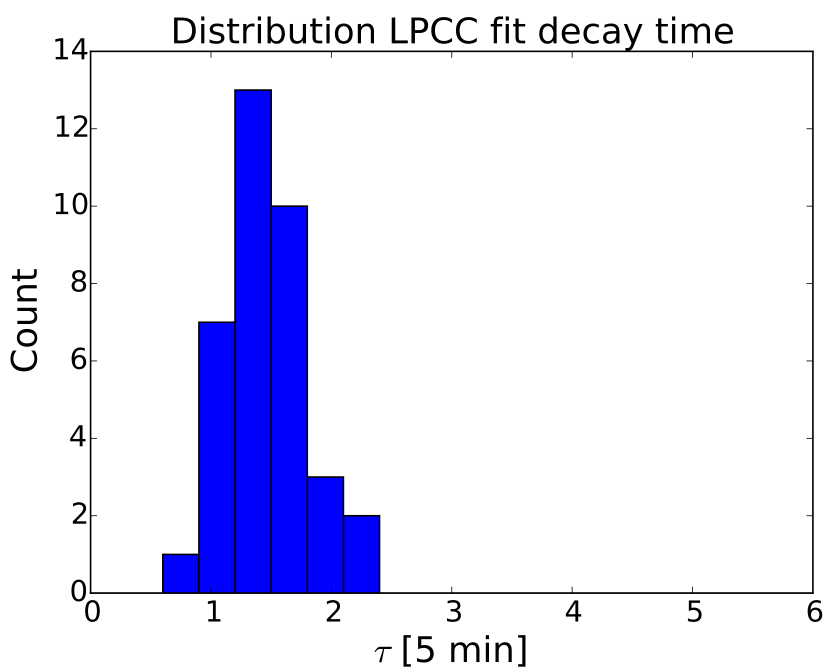


Partial cross-correlations

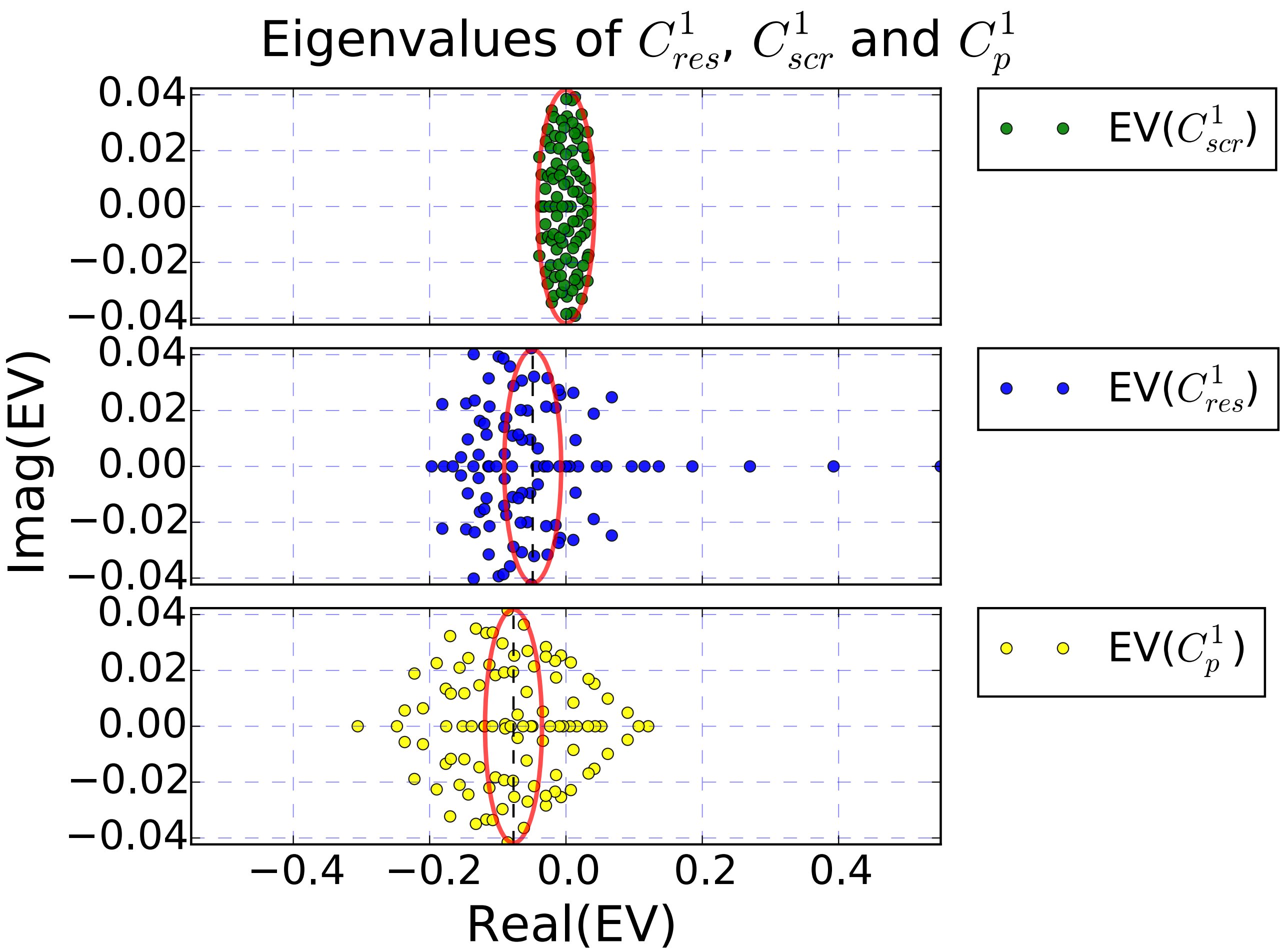


Partial cross-correlations

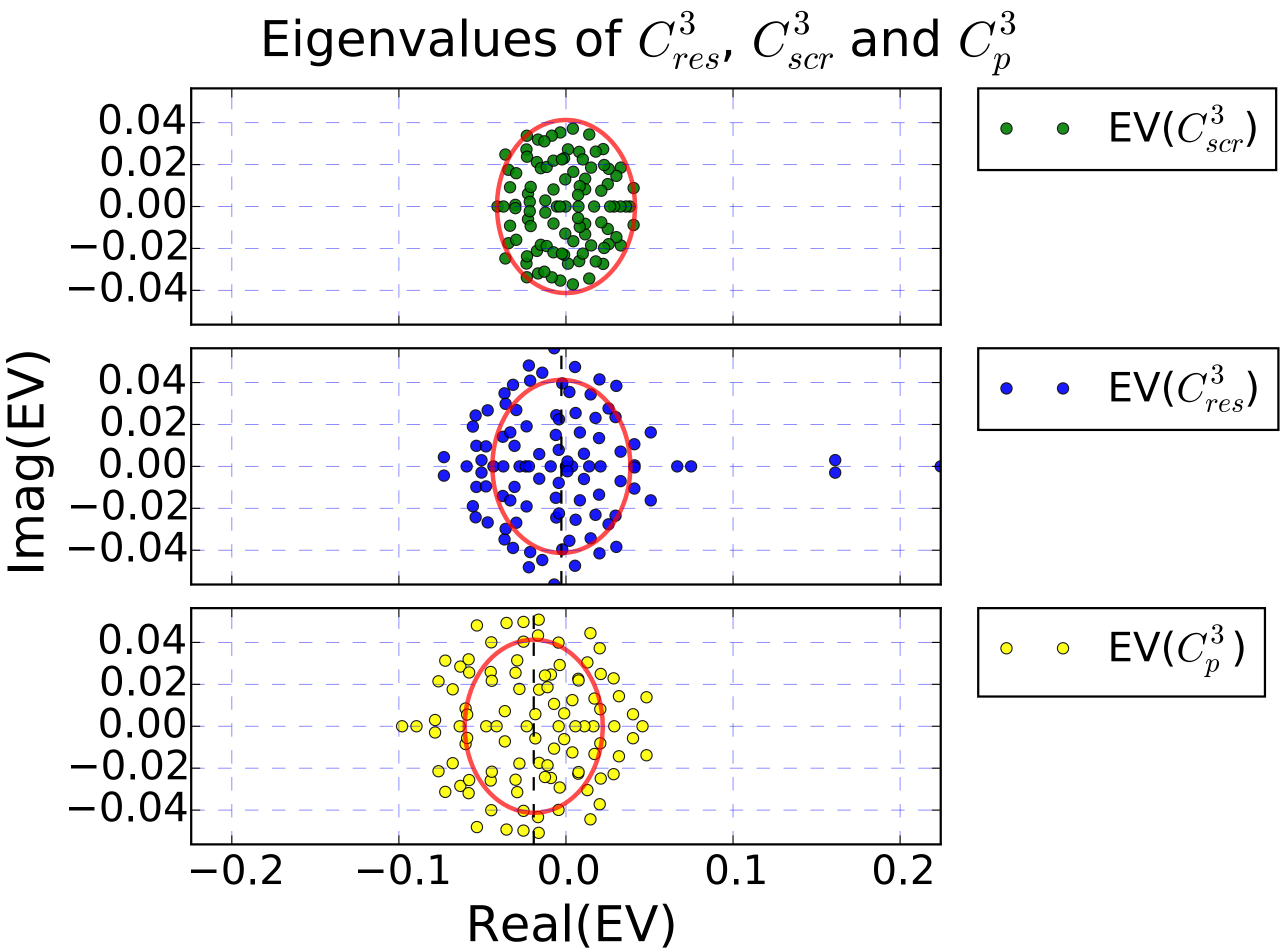
- Same decay time scale as partial autocorrelations, $\tau \approx 7\,\text{min}$



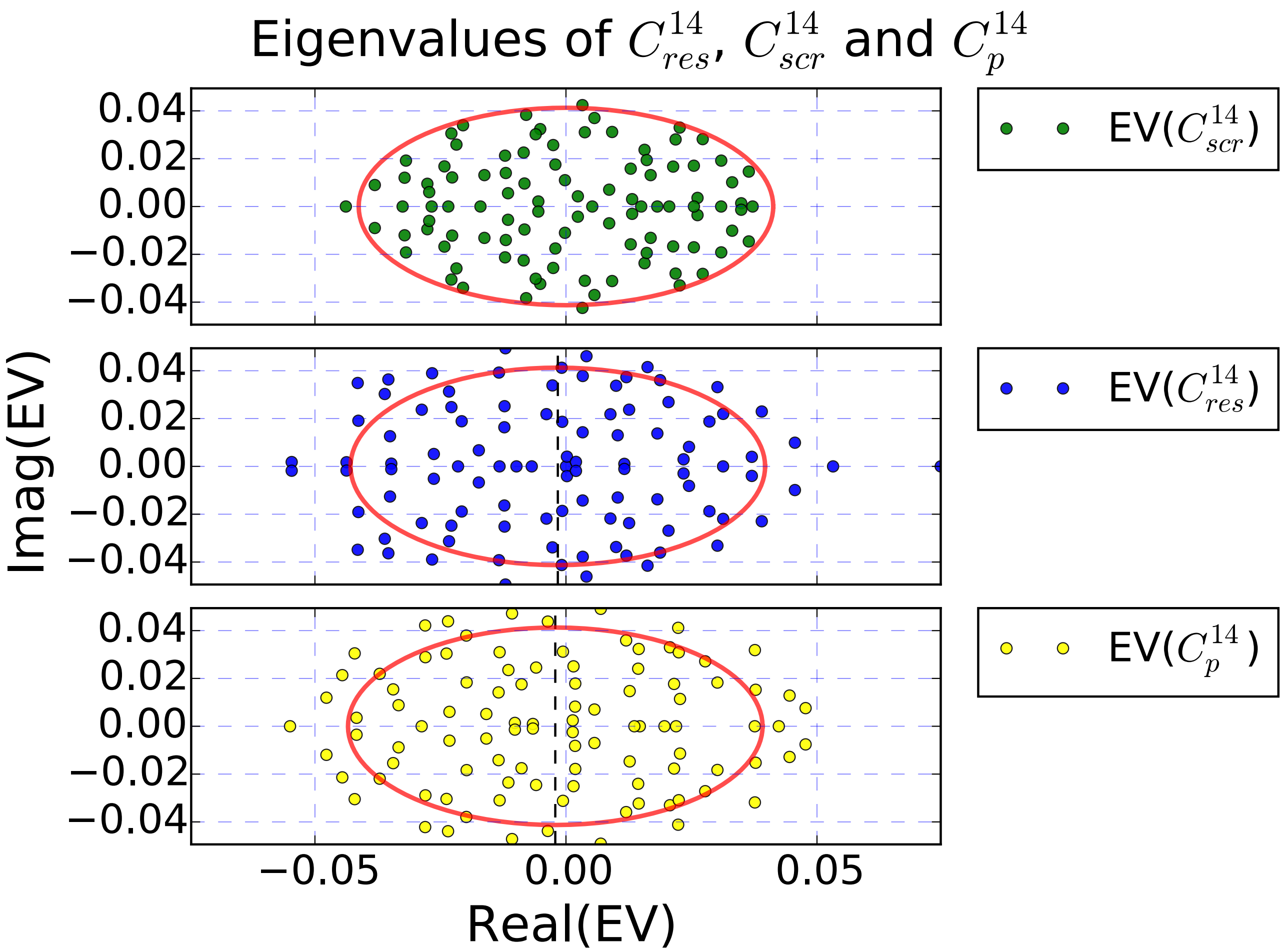
Eigenvalue distribution ($\tau = 1$)



Eigenvalue distribution ($\tau = 3$)



Eigenvalue distribution ($\tau = 14$)



Conclusion

General

Conclusion

General

- good tool to investigate underlying correlation network of a system

Conclusion

General

- good tool to investigate underlying correlation network of a system
- conditions can be extended arbitrarily

Conclusion

General

- good tool to investigate underlying correlation network of a system
- conditions can be extended arbitrarily

Stock market - NYSE

Conclusion

General

- good tool to investigate underlying correlation network of a system
- conditions can be extended arbitrarily

Stock market - NYSE

- typical decay time for correlations: 7 min

Conclusion

General

- good tool to investigate underlying correlation network of a system
- conditions can be extended arbitrarily

Stock market - NYSE

- typical decay time for correlations: 7 min
- raw correlation damped by mutual third party correlations

Conclusion

General

- good tool to investigate underlying correlation network of a system
- conditions can be extended arbitrarily

Stock market - NYSE

- typical decay time for correlations: 7 min
- raw correlation damped by mutual third party correlations
- almost no negative time-lagged cross-correlations

Outlook

General

Outlook

General

- Parallel computing could speed up calculations

Outlook

General

- Parallel computing could speed up calculations

Stock market - NYSE

Outlook

General

- Parallel computing could speed up calculations

Stock market - NYSE

- identify sectors and subsectors with synchronous partial correlations and compare to older results

Outlook

General

- Parallel computing could speed up calculations

Stock market - NYSE

- identify sectors and subsectors with synchronous partial correlations and compare to older results
- include time-lagged partial correlations in cluster identification$\rightarrow$ new dimension

Outlook

General

- Parallel computing could speed up calculations

Stock market - NYSE

- identify sectors and subsectors with synchronous partial correlations and compare to older results
- include time-lagged partial correlations in cluster identification$\rightarrow$ new dimension
- Plot correlation network with time dimension

Outlook

General

- Parallel computing could speed up calculations

Stock market - NYSE

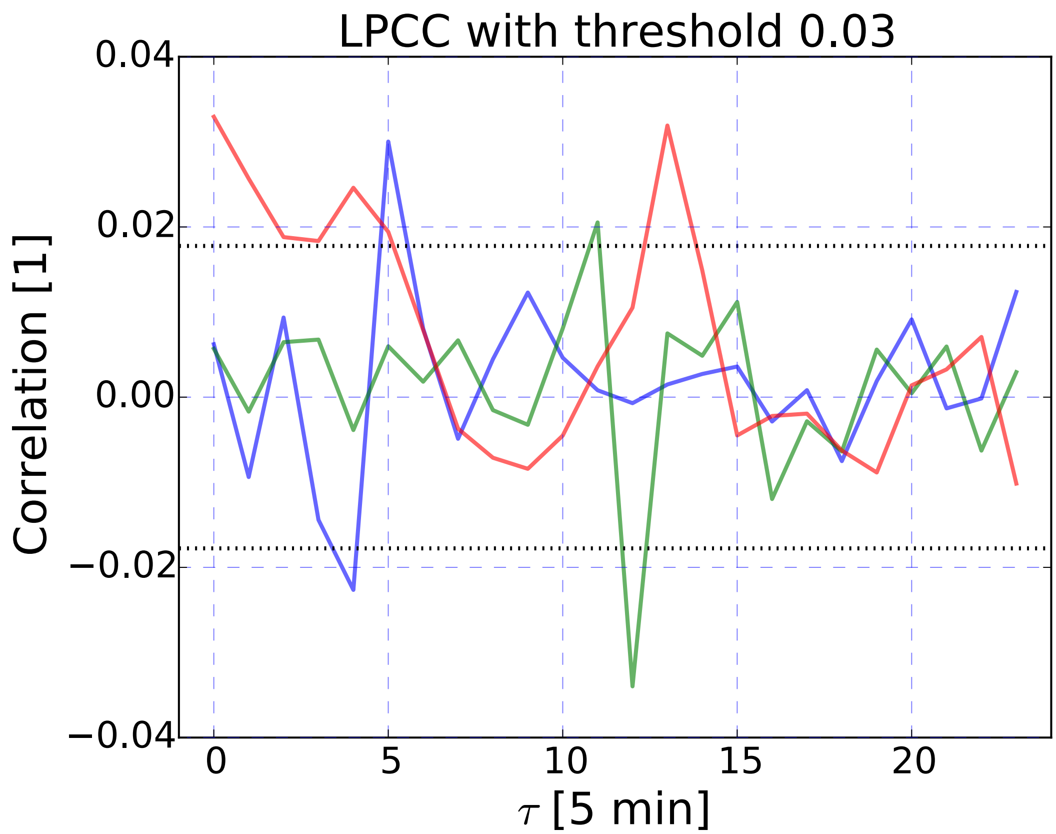
- identify sectors and subsectors with synchronous partial correlations and compare to older results
- include time-lagged partial correlations in cluster identification$\rightarrow$ new dimension
- Plot correlation network with time dimension
- Study SVD decompositions

End

Thank you for your attention!

And thanks to Chester!

Backup slides



Backup slides