

# Genetics and/of basket options

Wolfgang Karl Härdle

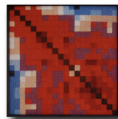
Elena Silyakova

Ladislaus von Bortkiewicz

Chair of Statistics

Humboldt-Universität zu Berlin

<http://lvb.wiwi.hu-berlin.de>



## Basket derivatives



## Basket derivatives



## Basket derivatives



## Basket derivatives

Let us consider a basket of  $N$  assets with value at time  $t$  defined by  $B(t) = \sum_{i=1}^N a_i S_i(t)$ . Then payoffs of some basket options:

□ Basket call:  $\{B(T) - K_B\}^+$

□ Rainbow (best-of-N):  $\left[ \max_{1 \leq i \leq N} \{S_i(T)\} - K_B \right]^+$

□ Atlas (Mountain range):

$$\left\{ \frac{1}{N - (N_1 + N_2)} \sum_{j=1+N_1}^{N-N_2} \frac{S_j(T)}{S_j(0)} - K_B \right\}^+$$

where  $S_i(t)$  - price of the  $i$ -th basket constituent at time  $t$ ,  $a_i$  - quantity of the  $i$ -th asset,  $K_B$  - exercise price (strike) of a basket option,  $T$  - time of the option's expiry,  $N_1, N_2$  - number of best and worst performing stocks.



## Basket derivatives are popular

Cheaper than the corresponding portfolio of plain vanillas:

- portfolio effect
- smaller transaction costs

However create large correlation exposures.

Need to hedge!



## Research questions

1. Which pricing model is suitable for multiasset options?
2. How to estimate dependence (correlation) between assets in the basket?
3. How to estimate correlations in large dimensional baskets?



# Outline

1. Motivation ✓
2. Basket dynamics in the Black-Scholes framework
3. Estimating correlation matrix
  - ▶ Historical (time series) correlation
  - ▶ Implied correlation
4. From equicorrelation to block correlation
5. Improving correlation estimates
6. Conclusion



## Price dynamics of basket constituents

The price dynamic of the  $i$ -th stock in a basket is given by:

$$\frac{dS_i(t)}{S_i(t)} = (r - q_i)dt + \sigma_i dW_i(t) \quad (1)$$

$$\rho_{ij}dt = dW_i(t)dW_j(t) \quad (2)$$

where  $r$  - interest rate,  $q_i$  - dividend yield of a stock  $i$ ,  $\sigma_i$  - constant volatility of the  $i$ -th stock,  $\rho_{ij}$  - constant correlation between the  $i$ -th and the  $j$ -th stock,  $W$  - Brownian motion.



## Dynamics of the basket's value

The dynamics of the basket's value is then given by:

$$\begin{aligned}\frac{dB(t)}{B(t)} &= (r - q_B)dt + \frac{\sum_{i=1}^N w_i S_i(t) \sigma_i dW_i(t)}{\sum_{i=1}^N w_i S_i(t)} = & (3) \\ &= (r - q_B)dt + dZ(t)\end{aligned}$$

where  $q_B$  is the dividend yield of the basket and the relative weight  $w_i$  of the  $i$ -th constituent varies over time and is given by:

$$w_i = \frac{a_i S_i(t)}{\sum_{l=1}^N a_l S_l(t)} \quad (4)$$



## Dynamics of correlated basket constituents

Let

$$\Sigma = \begin{pmatrix} \rho_{11} & \cdots & \rho_{1N} \\ \vdots & \ddots & \vdots \\ \rho_{N1} & \cdots & \rho_{NN} \end{pmatrix}$$

the correlation matrix of a basket.

By Cholesky decomposition  $\Sigma = MM^T$  we obtain

$M = (m_{i,j})_{1 \leq i \leq N, 1 \leq j \leq N}$ , a lower triangular matrix, a "square root" of  $\Sigma$ .

The process for every individual asset  $S_i$  is then defined by:

$$\frac{dS_i(t)}{S_i(t)} = (r - q)dt + \sigma_i \sum_{l=1}^N m_{i,l} dW_l(t) \quad (5)$$



Finally applying Itô's lemma we obtain the closed-form expression for simulation of the  $i$ -th stock process on a time interval  $\Delta t = [t_1, t_2]$ :

$$S_i(t_2) = S_i(t_1) \exp \left\{ \left( r - d - \frac{\sigma_i^2}{2} \right) \Delta t + \sigma_i \sum_{l=1}^N m_{i,l} \sqrt{\Delta t} g_l \right\} \quad (6)$$

where  $g_l \sim N(0, 1)$ , i.i.d.



## Historical correlation

$X_i(t) = \log S_i(t) - \log S_i(t - 1)$ , log returns:

$$\rho_{ij} = \frac{\sum_{k=0}^T \lambda^k \{X_i(t-k) - \bar{X}_i(t)\} \{X_j(t-k) - \bar{X}_j(t)\}}{\sqrt{\sum_{k=0}^T \lambda^k \{X_i(t-k) - \bar{X}_i(t)\}^2 \sum_{k=0}^T \lambda^k \{X_j(t-k) - \bar{X}_j(t)\}^2}}$$

to obtain the historical correlation matrix

$$\begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1N} \\ \rho_{12} & 1 & \cdots & \rho_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1N} & \rho_{N2} & \cdots & 1 \end{bmatrix}$$

Here  $\bar{X}_i(t)$  the arithmetic mean of the  $i$ -th log return calculated at time  $t$ ,  $\lambda$  - decay parameter (RiskMetrics:  $\lambda = 0.94$ ).



## Equicorrelation matrix

Basket variance

$$\sigma_{Basket}^2 = \sum_{i=1}^N w_i^2 \sigma_i^2 + 2 \sum_{i=1}^N \sum_{j=i+1}^N w_i w_j \sigma_i \sigma_j \rho_{ij} \quad (7)$$

replace  $\begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1N} \\ \rho_{21} & 1 & \cdots & \rho_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N1} & \rho_{N2} & \cdots & 1 \end{bmatrix}$  with  $\begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}$ ,

then

$$\rho = \frac{\sigma_{Basket}^2 - \sum_{i=1}^N w_i^2 \sigma_i^2}{2 \sum_{i=1}^N \sum_{j=i+1}^N w_i w_j \sigma_i \sigma_j} \quad (8)$$

is the average basket correlation.

Nice property: for  $-\{1/(N-1)\} < \rho < 1$  - positive definite (

Mardia et al, 1979).



# Outline

1. Motivation
2. Basket dynamics in the Black-Scholes framework
3. Estimating correlation matrix
  - ▶ Historical (time series) correlation
  - ▶ Implied correlation ✓
4. From equicorrelation to block correlation
5. Improving correlation estimates
6. Conclusion

## Implied volatility

- consider an option with Black-Scholes price  $V_{BS}(S(t), \kappa, \sigma, \tau)$
- given the observed market price  $\tilde{V}(t)$ , the implied volatility  $\hat{\sigma}$  can be found by solving:

$$V_{BS}(S(t), \kappa, \hat{\sigma}, \tau) - \tilde{V}(t, \kappa, \tau) = 0$$

where  $\kappa = \frac{K}{S(t) \exp\{(r-d)\tau\}}$  - moneyness metric with  $K$  - strike of the option,  $\tau$  - time to maturity of the option.





## **DAX implied volatility, 20070402-20070502**



## Implied correlation

Using (8) map the implied volatility surfaces of a basket  $\hat{\sigma}_{Basket}(\kappa, \tau)$  and  $N$  constituents  $\hat{\sigma}_i(\kappa, \tau)$  to  $\hat{\rho}(\tau, \kappa)$  the **average implied correlation surface of a basket**:

$$\hat{\rho}(\kappa, \tau) = \frac{\hat{\sigma}_{Basket}^2(\kappa, \tau) - \sum_{i=1}^N w_i^2 \hat{\sigma}_i^2(\kappa, \tau)}{2 \sum_{i=1}^N \sum_{j=i+1}^N w_i w_j \hat{\sigma}_i(\kappa, \tau) \hat{\sigma}_j(\kappa, \tau)} \quad (9)$$

## Dynamic modeling of correlation surfaces

Every  $t$  we observe  $(X_{t,j}, Y_{t,j})$ ,  $1 \leq j \leq J_t$ ,  $1 \leq t \leq T$  where

- ▣  $Y_{t,j}$  - implied correlation
- ▣  $X_{t,j}$  - two-dimensional vector of  $\kappa$  and  $\tau$
- ▣  $T$  - number of observed time periods (days)
- ▣  $J_t$  - number of observations at day  $t$



## Dynamic modeling of correlation surfaces

In an orthogonal  $L$ -factor model, a  $J$ -dimensional random vector  $Y_t = (Y_{t,1}, \dots, Y_{t,J})$  can be represented as

$$Y_{t,j} = Z_{t,1}m_{1,j} + \dots + Z_{t,L}m_{L,j} + \varepsilon_{t,j} \quad (10)$$

where

$Z_{t,j}$  - common factors,  $\varepsilon_{t,j}$  - errors and the coefficients  $m_{l,j}$  - factor loadings.

It such setting the modelling of  $Y_t$  can be simplified to modelling of  $Z_t = (Z_{t,1}, \dots, Z_{t,L})$ , which is more feasible for  $L \ll J$ .



## Dynamic modeling of correlation surfaces

Including explanatory variables  $X_{t,j}$  influencing the factor loadings  $m_{l,j}$  rewrite (10)

$$Y_{t,j} = \sum_{l=1}^L Z_{t,l} m_l(X_{t,j}) + \varepsilon_{t,j} = Z_t^\top m(X_{t,j}) + \varepsilon_{t,j} \quad (11)$$

where

- $Z_t = (Z_{t1}, \dots, Z_{tL})^\top$  - unobservable  $L$ -dimensional process
- $m$  -  $L$ -tuple  $(m_1, \dots, m_L)$  of unknown real-valued functions
- $X_{t,j}, \dots, X_{T,J_T}$  and  $\varepsilon_{t,j}, \dots, \varepsilon_{T,J_T}$  are independent
- $\varepsilon_{t,j}$  are *i.i.d.* with zero mean and finite second moment

## Dynamic modeling of correlation surfaces

$$Y_{t,j} = \sum_{l=1}^L Z_{t,l} \sum_{k=1}^K a_{l,k} \psi_k(X_{t,j}) + \varepsilon_{t,j} = Z_t^\top A \Psi_t + \varepsilon_t \quad (12)$$

where

- $A$  -  $L \times K$  coefficient matrix
- $\Psi_t = \{\psi_1(X_t), \dots, \psi_R(X_t)\}^\top$  - space basis, in Park et al. (2009) a tensor product of one dimensional B-spline basis.

## Estimation

Define the least squares estimators  $\hat{Z}_t$  and  $\hat{A}$ :

$$\sum_{t=1}^T \sum_{j=1}^J \left\{ Y_{t,j} - Z_t^\top A \psi(X_{t,j}) \right\}^2 = \min_{A, Z}!$$

Need to estimate space basis  $\Psi_t$ !



## Choice of space basis

Estimate basis functions in a FPCA framework, motivated by Hall et. al (2006):

Find eigenfunctions corresponding to the  $K$  largest eigenvalues of the smoothed operator

$$\hat{\psi}(u, v) = \hat{\phi}(u, v) - \hat{\mu}(u)\hat{\mu}(v)$$





## Choice of space basis

1. estimate  $\hat{\mu}(u)(\mu(v))$ :

$$\sum_{t=1}^T \sum_{j=1}^J \left\{ Y_{tj} - a - \sum_{c=1}^2 b^c (u^c - X_{tj}^c) \right\}^2 K \left( \frac{X_{tj} - u}{h_\mu} \right)$$

2. estimate  $\hat{\phi}(u, v)$ :

$$\sum_{t=1}^T \sum_{1 \leq j \neq k \leq J_t} \left\{ Y_{tj} Y_{tk} - a_0 - \sum_{c=1}^2 b_1^c (u^c - X_{tj}^c) - \sum_{c=1}^2 b_2^c (v^c - X_{tk}^c) \right\}^2$$

$$\times K \left( \frac{X_{tj} - u}{h_\phi} \right) K \left( \frac{X_{tk} - v}{h_\phi} \right)$$

3. compute  $\hat{\psi}(u, v) = \hat{\phi}(u, v) - \hat{\mu}(u)\hat{\mu}(v)$  and take  $K$  eigenfunctions corresponding to the largest eigenvalues

## Basis functions

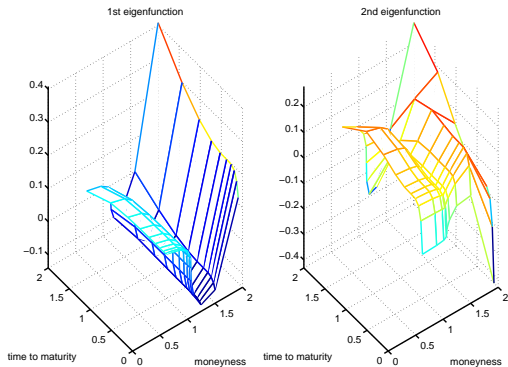


Figure 1: Eigenfunctions as basis functions estimated on 10x10 grid



## Basis functions

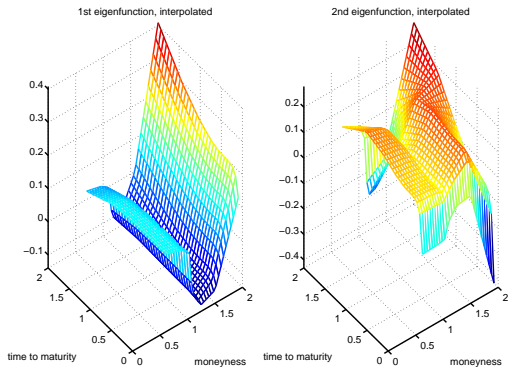
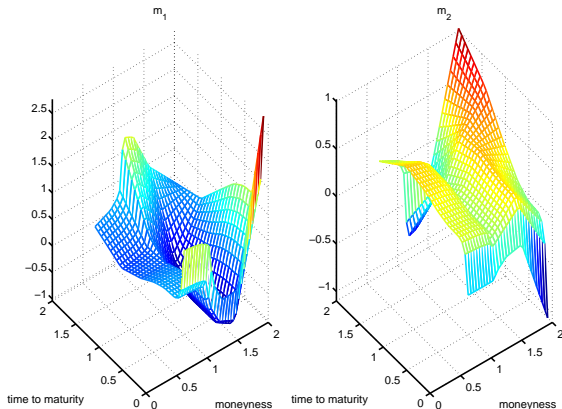


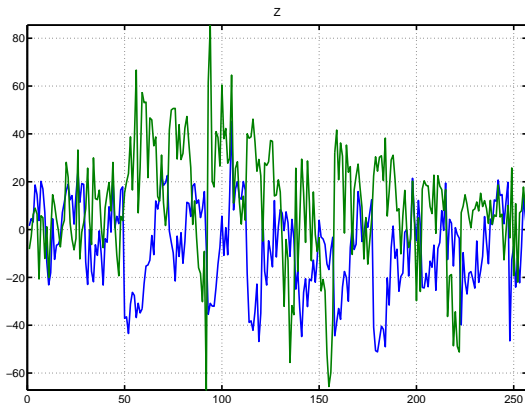
Figure 2: Eigenfunctions as basis functions interpolated



## Estimated functions $\hat{m}_I$



## Estimated time series of factors $\hat{z}_{t1}$ , $\hat{z}_{t2}$



## From equicorrelation to block correlation

Group assets in the basket into  $k$  blocks, then

$$\begin{bmatrix} \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_1 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_1 & \rho_1 & \cdots & 1 \end{bmatrix} & \cdots & \begin{bmatrix} \rho_{k+1} \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \\ \vdots & \ddots & \vdots \\ \begin{bmatrix} \rho_{k+1} \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} & \cdots & \begin{bmatrix} 1 & \rho_k & \cdots & \rho_k \\ \rho_k & 1 & \cdots & \rho_k \\ \vdots & \vdots & \ddots & \vdots \\ \rho_k & \rho_k & \cdots & 1 \end{bmatrix} \end{bmatrix}$$



## Correlation matrix for 2 groups of assets (3 blocks)

$$\begin{bmatrix}
 \begin{bmatrix}
 1 & \rho_1 & \cdots & \rho_1 \\
 \rho_1 & 1 & \cdots & \rho_1 \\
 \vdots & \vdots & \ddots & \vdots \\
 \rho_1 & \rho_1 & \cdots & 1
 \end{bmatrix} & & \\
 & \rho_3 & \\
 & & \begin{bmatrix}
 1 & \rho_2 & \cdots & \rho_2 \\
 \rho_2 & 1 & \cdots & \rho_2 \\
 \vdots & \vdots & \ddots & \vdots \\
 \rho_2 & \rho_2 & \cdots & 1
 \end{bmatrix}
 \end{bmatrix}$$



## Block implied correlation, 3 blocks

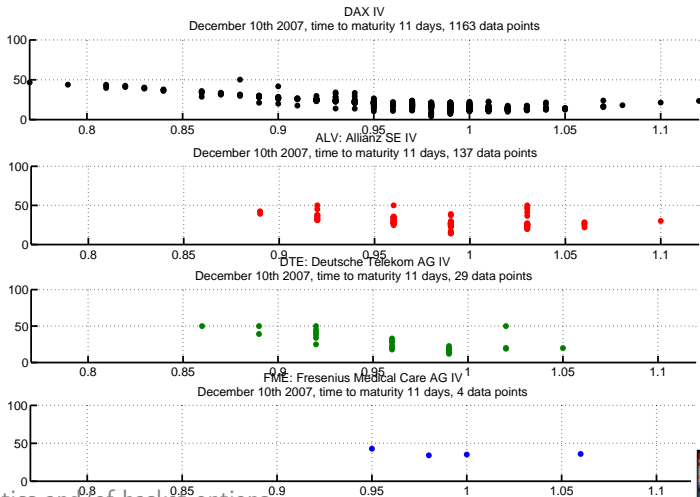
$$\begin{aligned}
 \sigma_{Basket}^2(K, \tau) &= \sum_{i=1}^N w_i^2 \sigma_i^2(K, \tau) + \\
 &+ 2 \sum_{i=1}^M \sum_{j=i+1}^M w_i w_j \sigma_i(K, \tau) \sigma_j(K, \tau) \rho_1(K, \tau) + \\
 &+ 2 \sum_{i=1}^{N-M} \sum_{j=i+1}^{N-M} w_i w_j \sigma_i(K, \tau) \sigma_j(K, \tau) \rho_2(K, \tau) + \\
 &+ 2 \sum_{i=1}^M \sum_{j=M+1}^N w_i w_j \sigma_i(K, \tau) \sigma_j(K, \tau) \rho_3(K, \tau)
 \end{aligned}$$

where  $M$  - number of assets in the 1-st block.

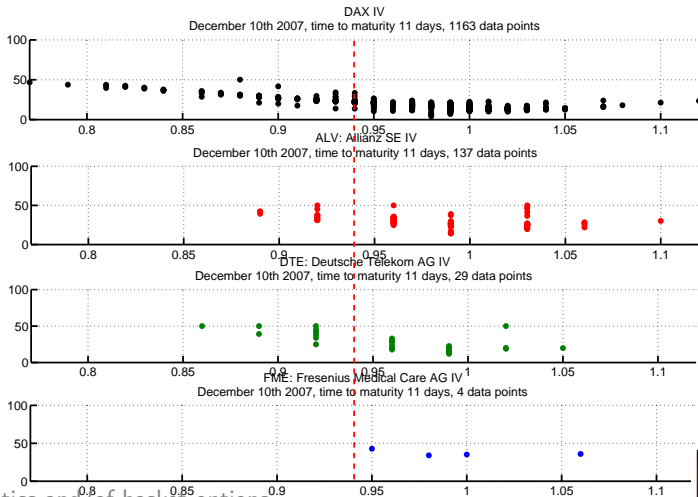




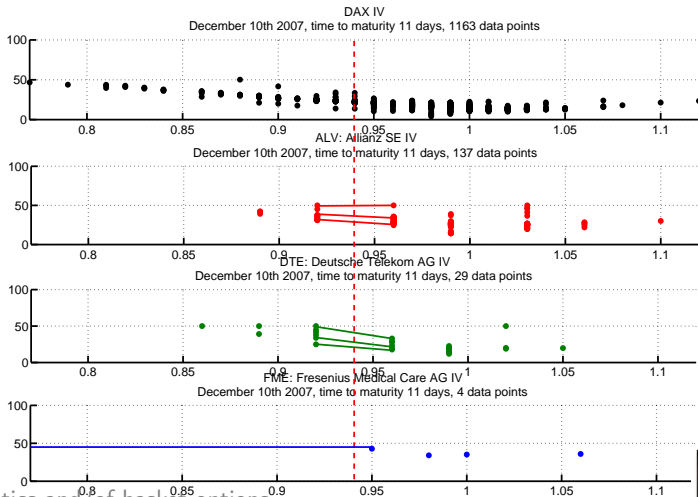
# Synchronization in moneyness dimension



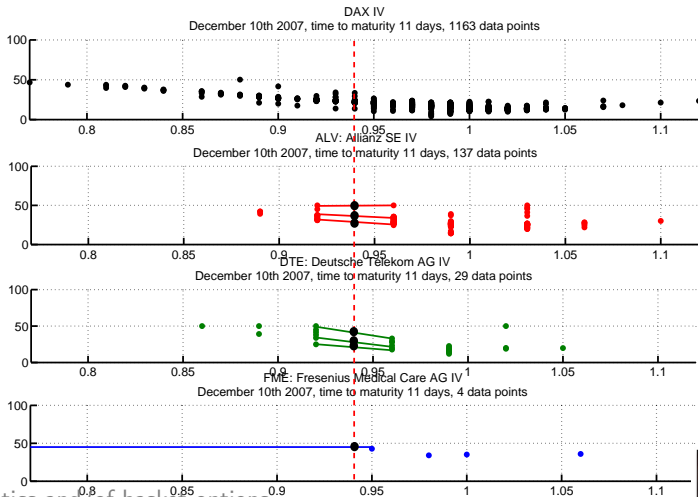
# Synchronization in moneyness dimension



# Synchronization in moneyness dimension



# Synchronization in moneyness dimension



## Challenges

- Moving to high-dimensional portfolios ( $N \nearrow$ ) with block structure of covariance matrix:
  - ▶ need well-conditioned estimate of covariance matrix (Ledoit and Wolf (2003), Bickel and Levina (2008))
  - ▶ need to define the grouping procedure and way of finding the optimal block size (Hautsch, Kyj and Oomen (2009))
- Improving correlation surface modeling:
  - ▶ need to expand the time effect in a series model  $Z_t$  as a sum of basis functions ( Song Härdle and Ritov (2010))



# Genetics and/of basket options

Wolfgang Karl Härdle

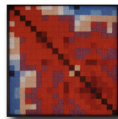
Elena Silyakova

Ladislaus von Bortkiewicz

Chair of Statistics

Humboldt-Universität zu Berlin

<http://lvb.wiwi.hu-berlin.de>



## Bibliography



Alexander, C.

*Market Models, A Guide to Financial Data Analysis*

John Wiley & Sons (2001)



Bai, Z.D.

Methodologies in Spectral Analysis of Large Dimensional  
Random Matrices, A Review

*Statistica Sinica*, (1999)






Efron, B.

Bootstrap Methods: Another Look at the Jackknife

*Annals of Statistics*, (1979)



## Bibliography

-  Fengler, M. R., Pilz K.F. and P. Schwendner  
Basket Volatility and Correlation  
*Volatility As An Asset Class, Risk Publications* (2007)
-  Fengler, M. R. and P. Schwendner  
Quoting multiasset equity options in the presence of errors  
from estimating correlations  
*Journal of Derivatives*, (2004)
-  Hall, P., Müller, H. G. and Wang J.  
Properties of principal component methods for functional and  
longitudinal data analysis  
*Ann. Statist.*, 34(3): 1493-1517, (2006)





## Bibliography



Laloux, L., et al.

Random Matrix Theory and Financial Correlations  
*International Journal of Theoretical and Applied Finance*,  
(2000)



Ledoit, O., and M. Wolf

Improved Estimation of the Covariance Matrix of Stock  
Returns with an Application to Portfolio Selection  
*Journal of Empirical Finance* 105, (2003)



Mardia, K. V., Kent, J. T. and Bibby, J. M.

Multivariate Analysis  
*Academic Press, Duluth, London*, (1979)



Plerou, V., et al.

Random Matrix Approach to Cross Correlations in Financial  
Data

