Cointegration Analysis and Community Detection in Cointegration-based Financial Networks

Yanhua Chen

University of Liverpool

Institute for Risk and Uncertainty

Yanhua.Chen@liverpool.ac.uk

8 September 2015

➢ Contents

➢ Global Stock Markets

Q: Financial Market Integration?

Q: Financial Market Co-Movement?

Wavelet analysis
GARCH model
Correlation analysis
…

➢ Why Choose Cointegration Theory

"Non-stationary"

- The goodness of fit R^2 'too high'
- Durbin-Watson (*DW*) statistic 'too low'
- Obtain apparently significant relationships from unrelated time series
- Spurious regression

"Difference"

To difference the time series successively until stationary is achieved

But the long-run information in the data will be lost

Cointegration Theory

- Avoid spurious results
- Estimate long-run (i.e. long-equilibrium) relationship between non-stationary time series
- If two time series move together in the long-run, it is possible to predict each series' movement .

Granger and Newbold (1974) , Engle and Granger (1987) ;

What is Cointegration?

Two random walks (i.e. non-stationary time series)

Unrelated, no meaningful relationship between their paths

Two random walks (i.e. non-stationary time series)

The random processes describing their paths are cointegration

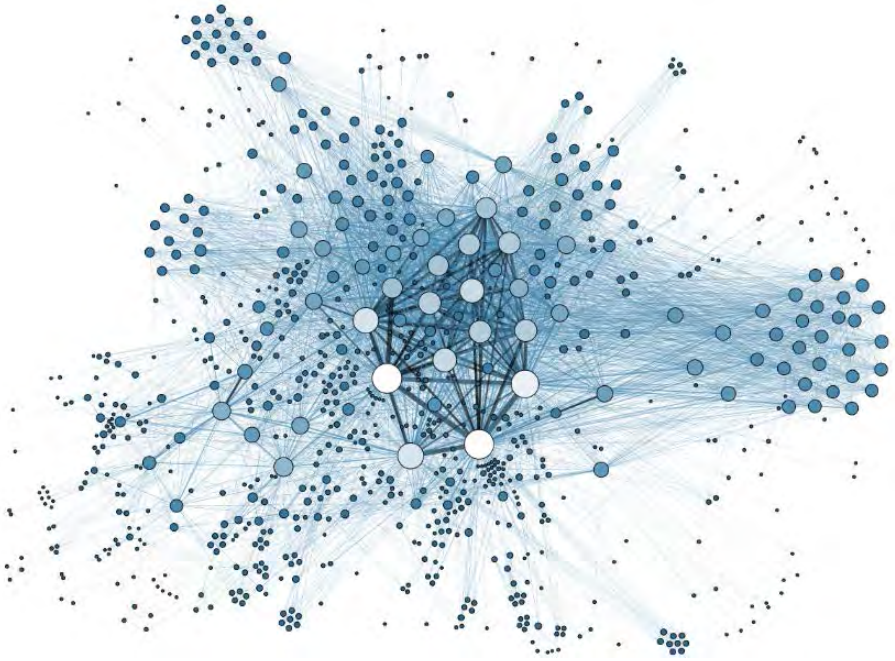
Two non-stationary time series become stationary after differenced, when a regression of one on the other, a linear combination of the two non-stationary time series is stationary, the two non-stationary time series are cointegrated (long-run equilibrium relationship)

Engle and Granger, 1987

Financial Market as a "complex system"

Many, heterogeneous elements

Highly complex interaction



Unveil the complex structure of the Financial Market

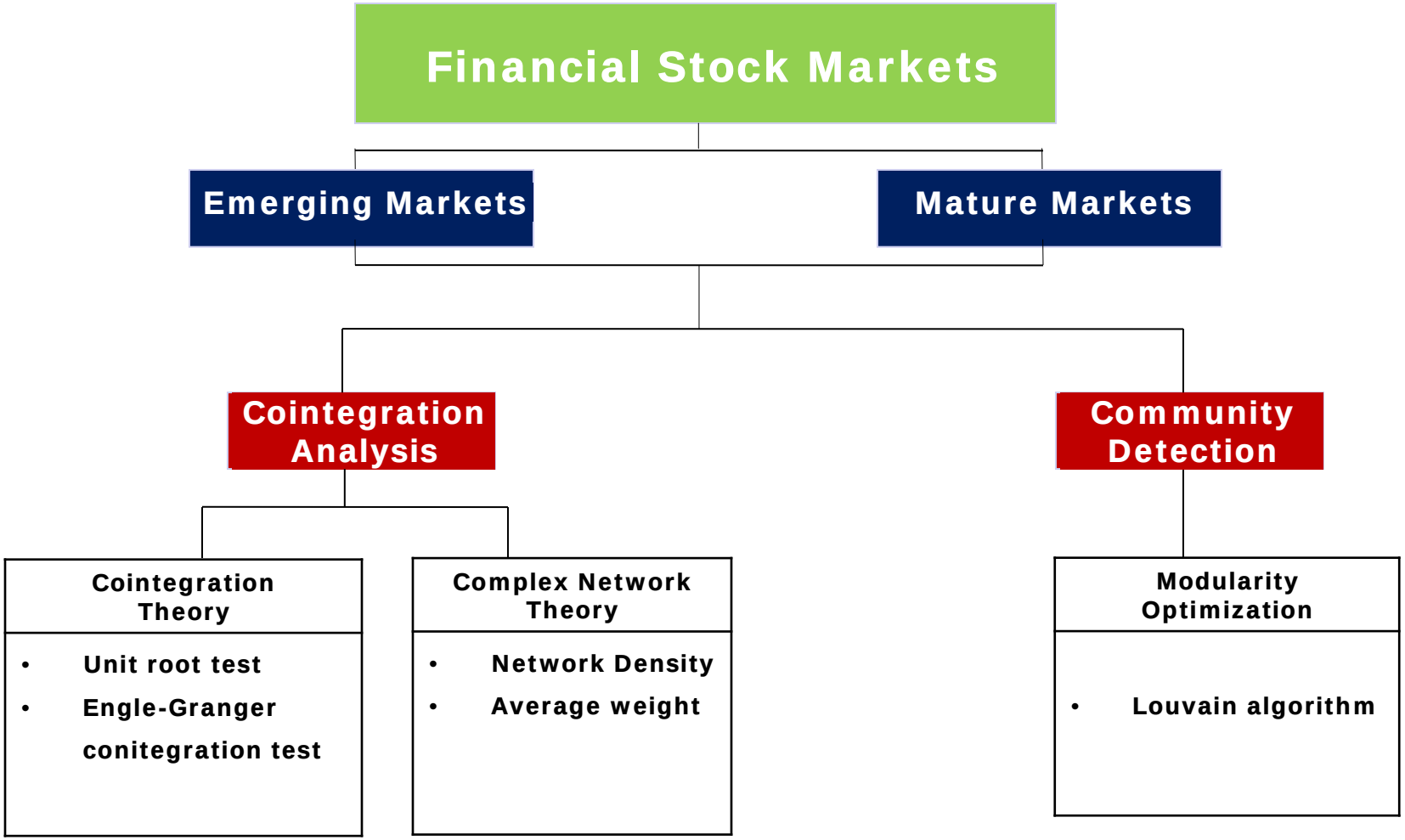
Asset allocation

Portfolio risk estimation

Complex Network Theory

Cointegration Theory

➢ Our Focus and Methodology



- **How the financial crises (U.S. sub-prime crisis and European debts crisis) impact on the global stock markets?**

➢ Financial Market Data Used

26 Global stock market index (Jan.2002—Apr.2012)

Table 1: List of 26 financial markets.

Europe	1	AEX	Netherlan
	2	ATX	Austria
	3	FCHI	France
	4	GDAXI	German
	5	FTSE	U.K
	6	OSEAX	Norway
	7	OMXSPI	Sweden
	8	SSMI	Switzerla
	9	RTS	Russia
Asia&Oceania	10	AORD	Australia
	11	BSESN	India
	12	HSI	Hong Kong
	13	VN	Vietnam
	14	KLSE	Malaysia
	15	KS11	Korea
	16	N225	Japan
	17	SSEC	China
	18	STI	Singapore
	19	TWII	Taiwan
Middle East	20	TA−100	Israel
Americas	21	BVSP	Brazil
	22	MXX	Mexico
	23	MERV	Argentina
	24	GSPTSE	Canada
	25	GSPC	U.S.
Africa	26	JALSH	South Africa

4 sub-periods

Period I: before sub-prime crisis
(Jan. 2002–Dec. 2005)

Period II: early stage of sub-prime crisis and the recession of U.S.
(Jan. 2006–Sep. 2008)

Period III: After the collapse of Lehman Brother
(Sep. 2008-Dec. 2009)

Period IV: during the European debts crisis
(Jan. 2010–Apr. 2012)

C. Y. Chong(2011), J. C. Shambaugh et al. (2012)

➢ One basic requirement: all the time series are non-stationary and must integrated of order 1, *I(1)*

➢ Method: Unit Root Test (Augment Dickey-Fuller Test)

STEP 1 Assume index series y_t is random walk process

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t \qquad \varepsilon_t : \text{is a white noise}$$

STEP 2 Run the ADF test model:

$$\Delta y_t = \rho y_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta y_{t-i} + \varepsilon_t$$

$$\Delta y_t = y_t - y_{t-1}$$

$$\rho = \sum_{i=1}^{p} \phi_i - 1$$

$$\gamma_i = -\sum_{j=1}^{i} \phi_{p-i+j}$$

Dickey and Fuller (1981), Said and Dickey (1984),

STEP 3 The null hypothesis: $H_0: \rho = 0 \quad (\rho = \sum_{i=1}^{p} \phi_i - 1, \quad \sum_{i=1}^{p} \phi_i = 1)$

Then y_t has a unit root

ρ is estimated by OLS(ordinary least squares) method

If ADF t-statistic > the critical value at 1%,5%,10% significant level

Accept the null hypothesis, y_t has a unit root(non-stationary)

If ADF t-statistic < the critical value at 1%,5%,10% significant level

Reject the null hypothesis, y_t is stationary

Asymptotic critical values for the unit root test

No.	Test	Significant level		
Variables	statistic	0.01	0.05	0.1
N=1	τ_μ	-2.58	-1.96	-1.61

MacKinnon (1991) , Davidson and MacKinnon (1993)

➢Cointegration Theory---Engle-Granger Cointegration Test

➢ When all the time series are non-stationary after first differenced are stationary $y_t \sim I(1)$

➢Method: Engle-Granger Cointegration Two–Step Test

Assume if y_t and x_t are cointegrated, a linear combination of them must be stationary, *i.e.*

$$y_t = \alpha x_t + \mu_t \qquad \hat{\mu}_t = y_t - \hat{\alpha} x_t \qquad \hat{\mu}_t \text{ is stationary}$$

STEP 1 Estimate α by using OLS(ordinary least squares)method

STEP 2 Run unit root test(ADF Test) on the estimated $\hat{\mu}_t$ series

If $\hat{\mu}_t$ is a stationary series, then y_t cointegrates with x_t

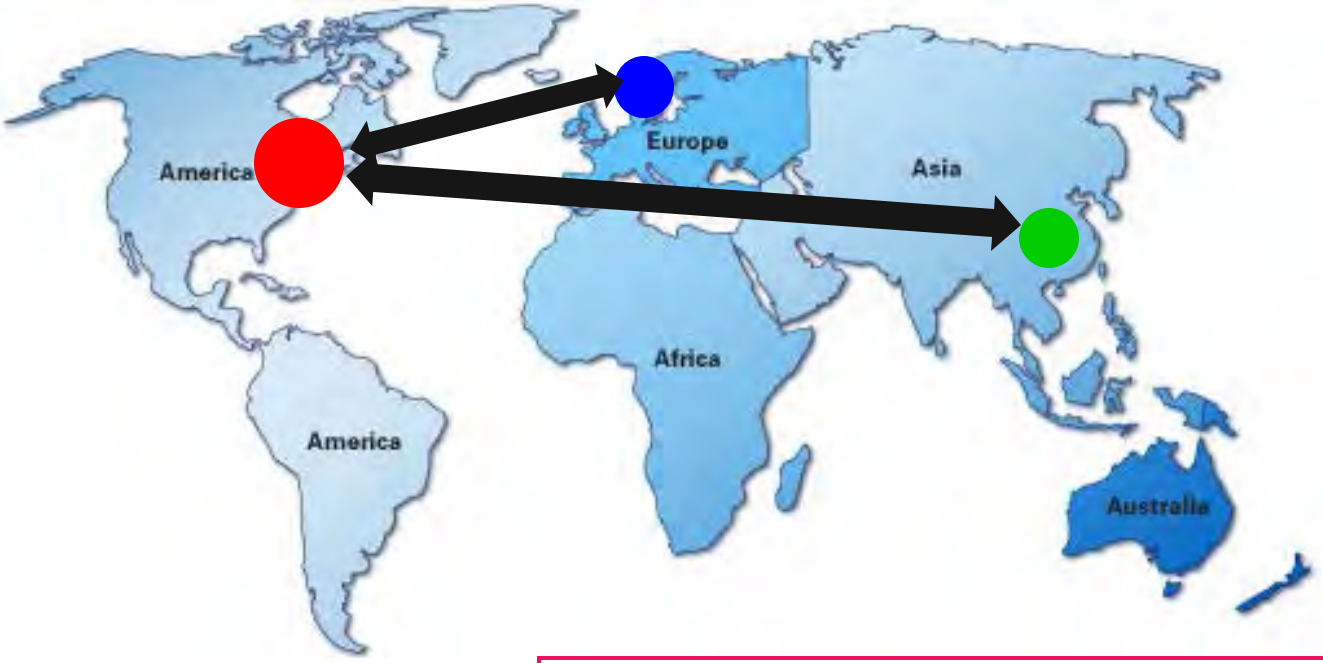
➢ Cointegration Coefficient α

y_t cointegrates with x_t, the regression tells us that on average, a unit increase in x_t is accompanied with a 3.5 % increase in y_t

$$y_t = 3.5x_t + \mu_t \qquad x_t = 1.4y_t + \mu_t'$$

➢What is Cointegration-based Financial Network?

➢ Network Construction



Nodes: stock market indices

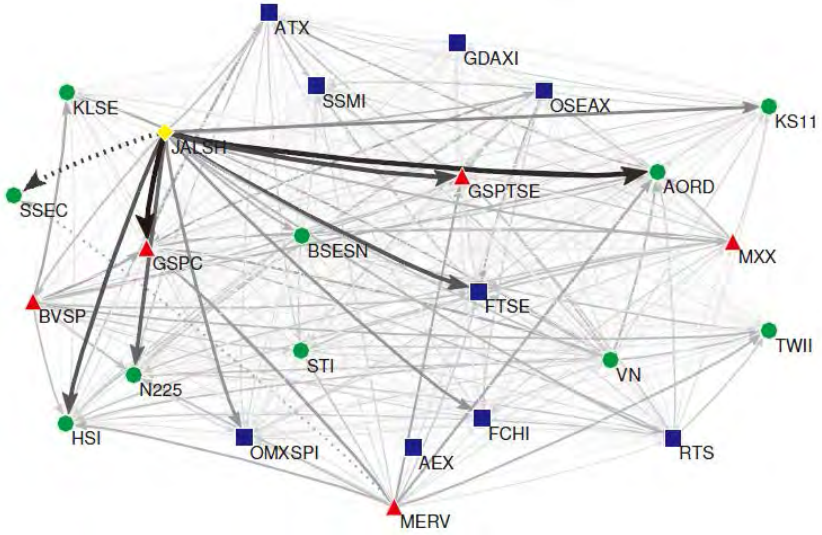
A Weighted and Directed Financial Network

Edges: cointegration relations between different stock markets

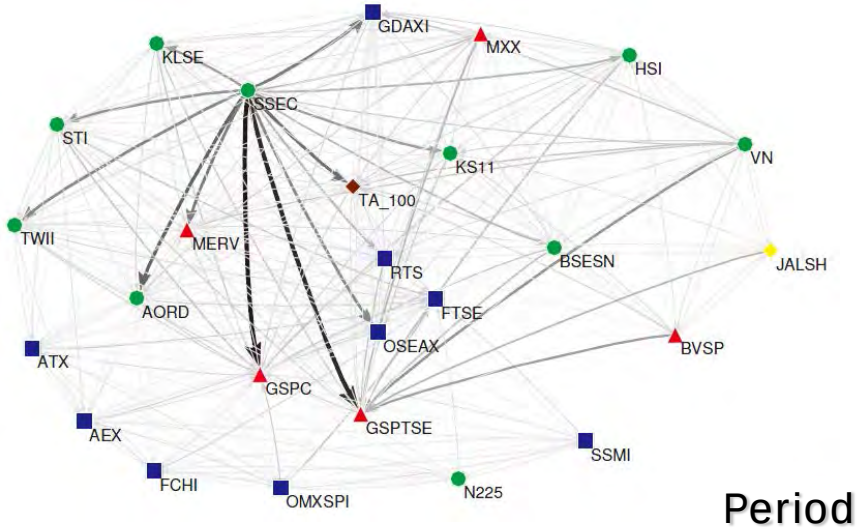
Direction → *i cointegrates with j or j cointegrates with i*

Weight → *cointegration coefficient* $w_{ij}(w_{ji}) = \alpha_{ij}(\alpha_{ji})$

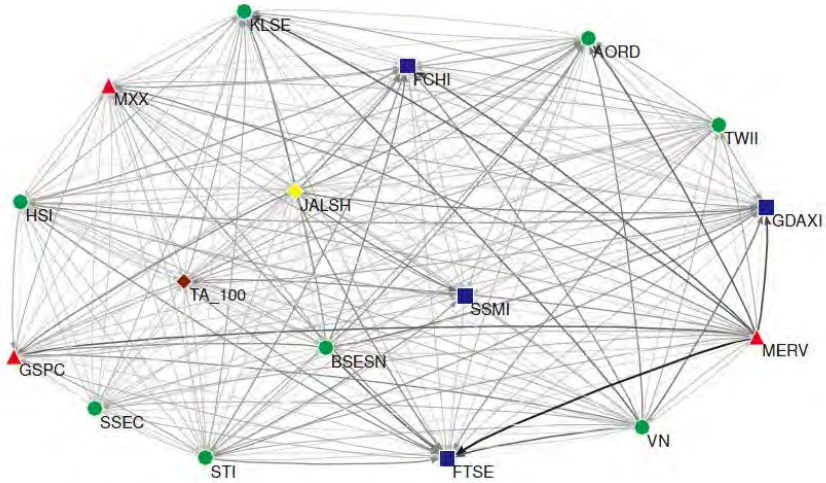
➢Cointegration-based Financial Networks in different periods in our eyes



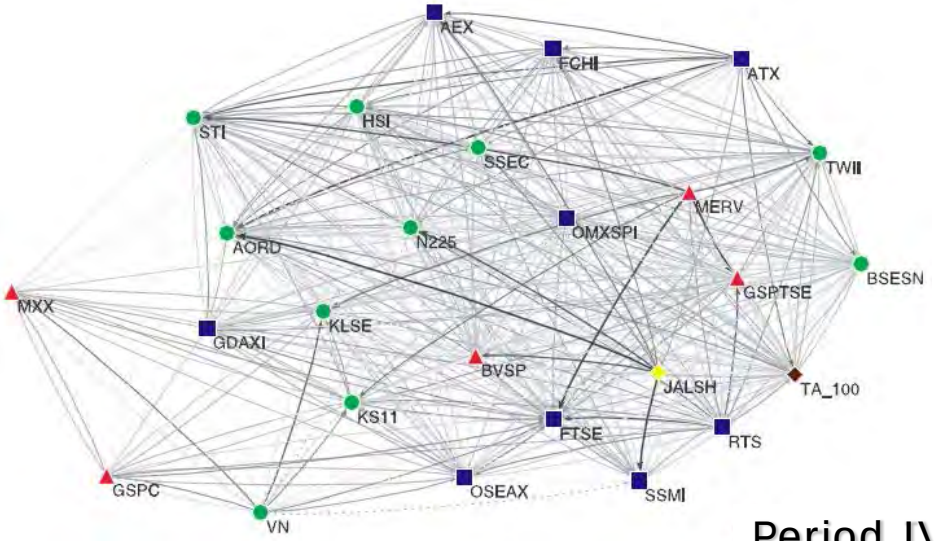
Period I



Period II



Period III



Period IV

- Europe
- Asia & Pacific
- America
- Africa–South Africa
- Middle East-Israel

Cointegration-based financial networks during 4 Periods.

➢ Cointegration-based Financial Network Analysis

Network property of cointegration-based financial networks

The Network Density

$$D = E / N(N-1)$$

Vary from 0 to 1

The higher scores, the higher density, the more cointegration relationships

The Average Weight

$$\langle W \rangle = W / E$$

The degree of the cointegration relationships

	D	<W>
Period I	0.9	1.023
Period II	0.602	0.951
Period III	0.967	0.947
Period IV	0.743	0.667

⇒ Cointegration relationship significantly increased after the Lehman Brothers collapse

⇒ While the degree of cointegration gradually weakened

Financial Crises have changed the cointegration relationships among global stock markets

➢ Community Detection—Modularity

Community is the groups of network nodes. Within these groups there have dense internal links, but between groups there are fewer edges.

The metric for quality of partition is its modularity:

$$Q = \frac{1}{|N|}\sum_{ij}\left[A_{ij} - \frac{k_i^{in}k_j^{out}}{|N|}\right]\delta(c_i, c_j)$$

$\delta(c_i, c_j)$: is the Kronecker delta symbol

A_{ij} : the adjacency matrix of the network

k_i^{in} : is the in-weight of node *i*

k_j^{out} : is the out-weight of node *j*

N : the total number of nodes in the network

c_i : the community to which node *i* belongs

M. E. J. Newman, Eur. Phys. J. B 38, 321-330 (2004); Blondel et al. (2008)

➢ Community Detection—A greedy optimization method

➢ Method: **Louvain algorithm**

Attempt to optimize the modularity of a partition of the network

Initially, every node belongs to the community of itself (|N| communities)

Phase 1 Find 'small' communities by optimizing modularity locally

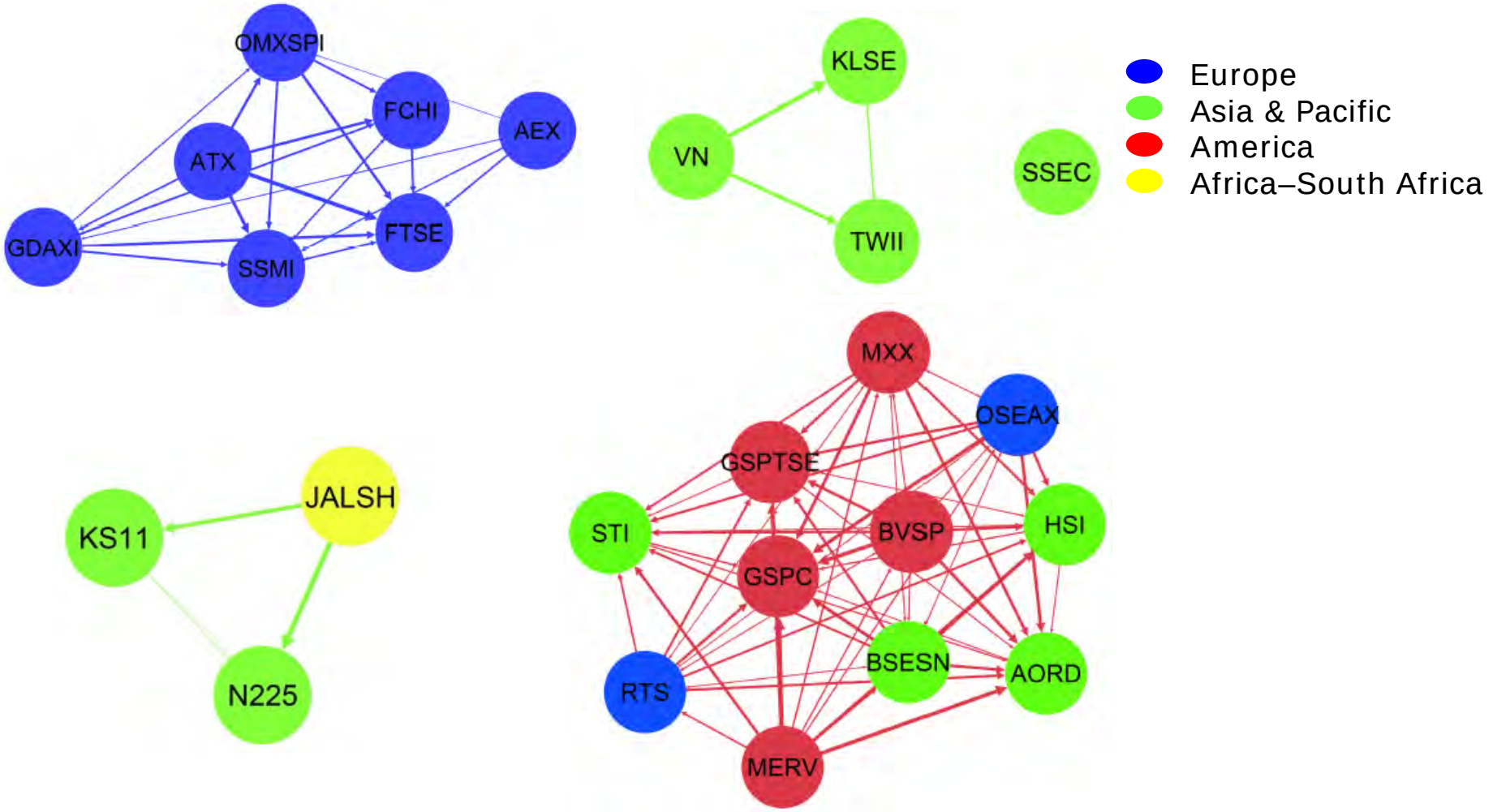
Phase 2 Aggregate nodes belonging to the same community and build a new network whose nodes are the communities

These steps are repeated iteratively until a maximum of modularity is attained and a hierarchy of communities is produced.

Blondel et al. (2008)

➢Community Detection

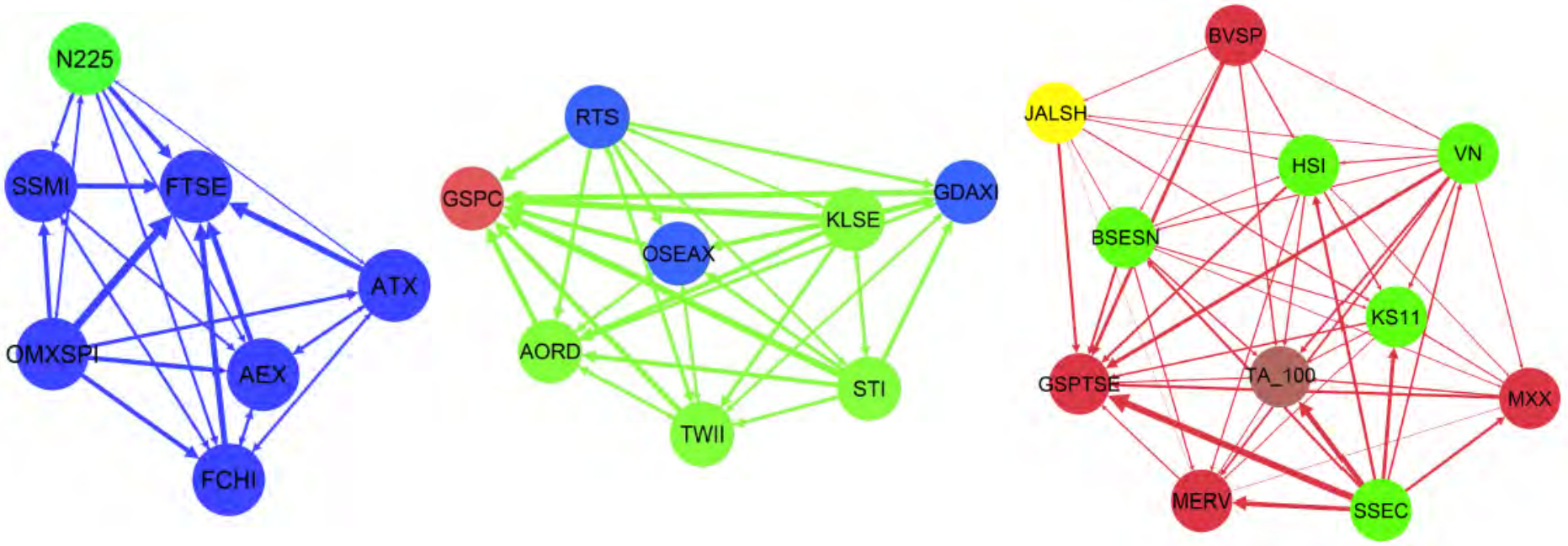
➢ Period I: Before sub-prime crisis(Jan.2002-Dec.2005)



Communities detected in the cointegration-based financial network during Period I.

➢Community Detection

➢ Period II: Early stage of sub-prime crisis and recession of US

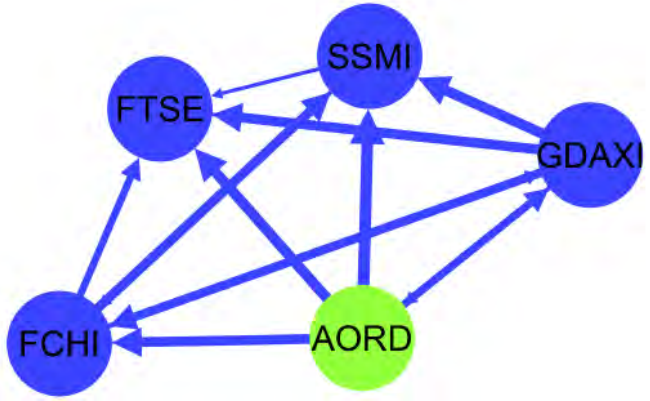


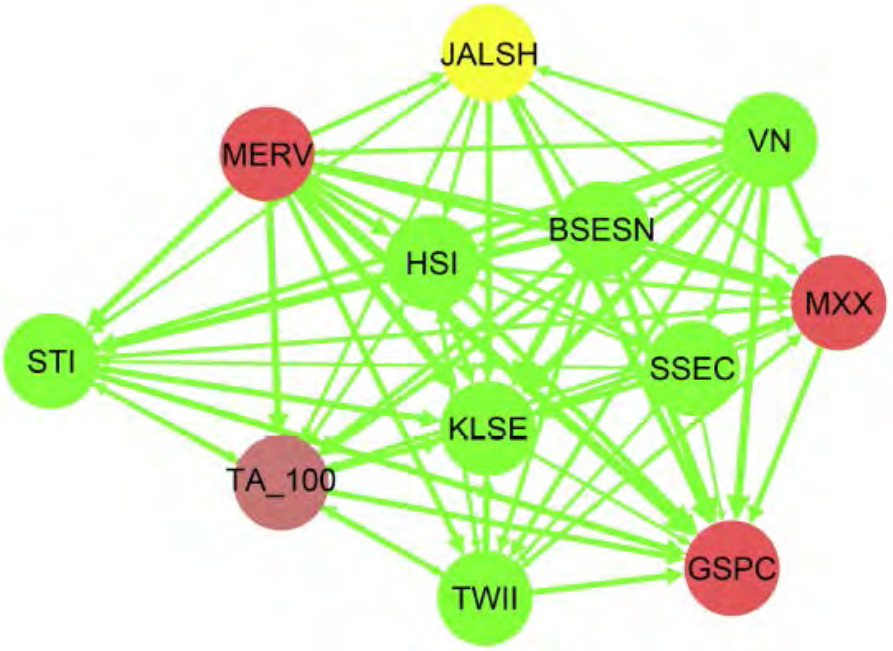
- Europe
- Asia & Pacific
- America
- Africa–South Africa
- Middle East-Israel

Communities detected in the cointegration-based financial network during Period II.

➢Community Detection

➢ Period III: After the collapse of Lehman Brother



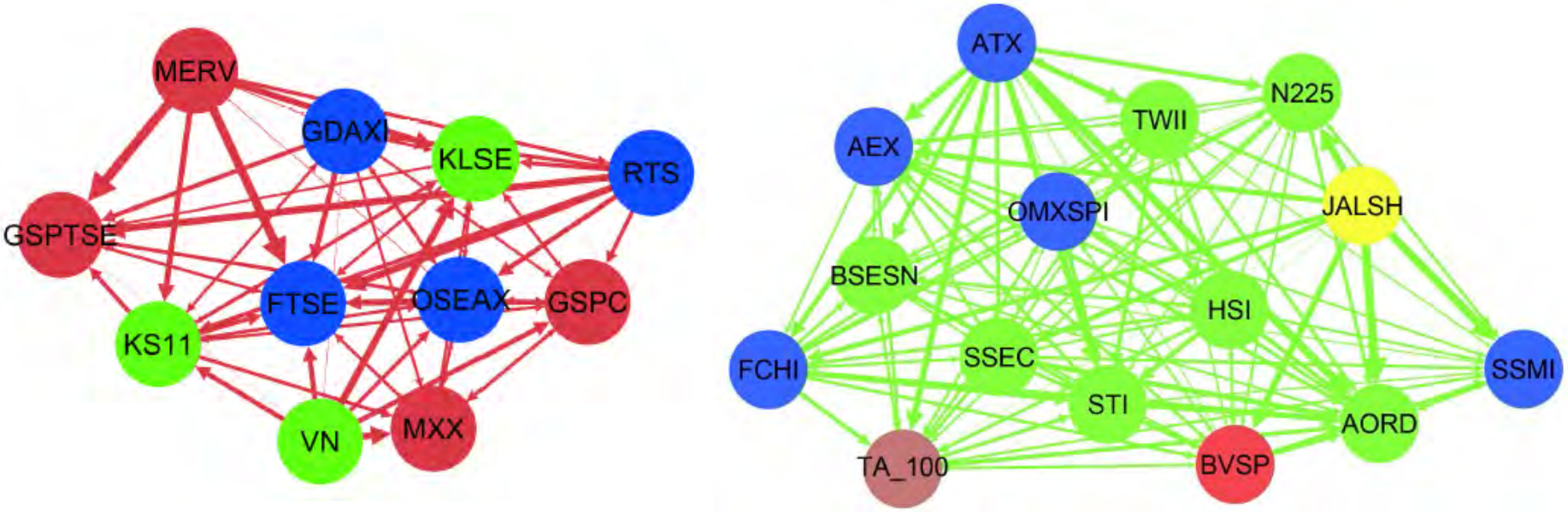


- Europe
- Asia & Pacific
- America
- Africa–South Africa
- Middle East-Israel

Communities detected in the cointegration-based financial network during Period III.

➢Community Detection

➢ Period IV: During the European debts crisis



- Europe
- Asia & Pacific
- America
- Africa–South Africa
- Middle East-Israel

Communities detected in the cointegration-based financial network during Period IV.

➢ Main Results of Community Detection

- The time-varying nature of community during different periods. From Period I to Period IV, the number of community decreased, while the size of these communities significantly increased. **The global stock markets experienced a higher market integration.**

- The financial markets in same geographical regions show highly cointegration relationship with each other.

- The BRICS grouping members(China, Brazil, South Africa, India) are always contained in same community from Period I to Period IV. While another member Russia, is always included in another community with U.S.

Summary

- Financial Crises have changed the cointegration relationships among global stock markets. Cointegration relationship significantly increased after the Lehman Brothers collapse. While the degree of cointegration gradually weakened.

- The time-varying nature of community during different periods. From Period I to Period IV, the number of community decreased, while the size of these communities significantly increased. The global stock markets experienced a higher market integration.

- The financial markets in same geographical regions show highly cointegration relationship with each other.

- The BRICS grouping members(China, Brazil, South Africa, India) are always contained in same community from Period I to Period IV. While another member Russia, is always included in other community with U.S.

Thank you *for your attention!*

Yanhua.Chen@liverpool.ac.uk