The betweenness centrality in random planar graphs: Some results

Marc Barthelemy

CEA, Institut de Physique Théorique, Saclay, France

EHESS, Centre d'Analyse et de Mathématique sociales, Paris, France

marc.barthelemy@cea.fr
www.quanturb.com

2nd Symposium Spatial Networks, Oxford 2017

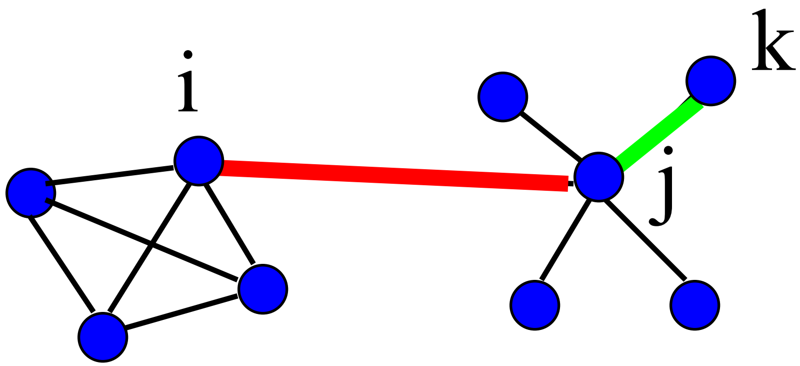
Outline

- Why is the BC interesting ?
- Simple remarks and empirical facts
- Theoretical approaches
  1. Loop versus center: a toy model
  2. The distribution of the BC: structural invariant
  3. Note on the dense limit
- Discussion and perspectives

Characterizing spatial networks

- Richness of spatial networks: topology and space (adjacency matrix + position of the nodes)
  - Usual measures irrelevant (Clustering, assortativity, degree distribution,…). New measures needed !
  - Path-related quantities
  - BC: non trivial spatial distribution. Good candidate for characterizing the structure of a planar network

More interesting: Betweenness Centrality (Freeman '77)



ij: large centrality

jk: small centrality

$$g(ij) = \sum_{s,t} \frac{\sigma_{st}(ij)}{\sigma_{st}}$$

σ_{st} = # of shortest paths from s to t

$\sigma_{st}(ij)$= # of shortest paths from s to t **via** (ij)

Measures the importance of a segment in the shortest paths flow

Betweenness centrality

- Backbone of stable roads

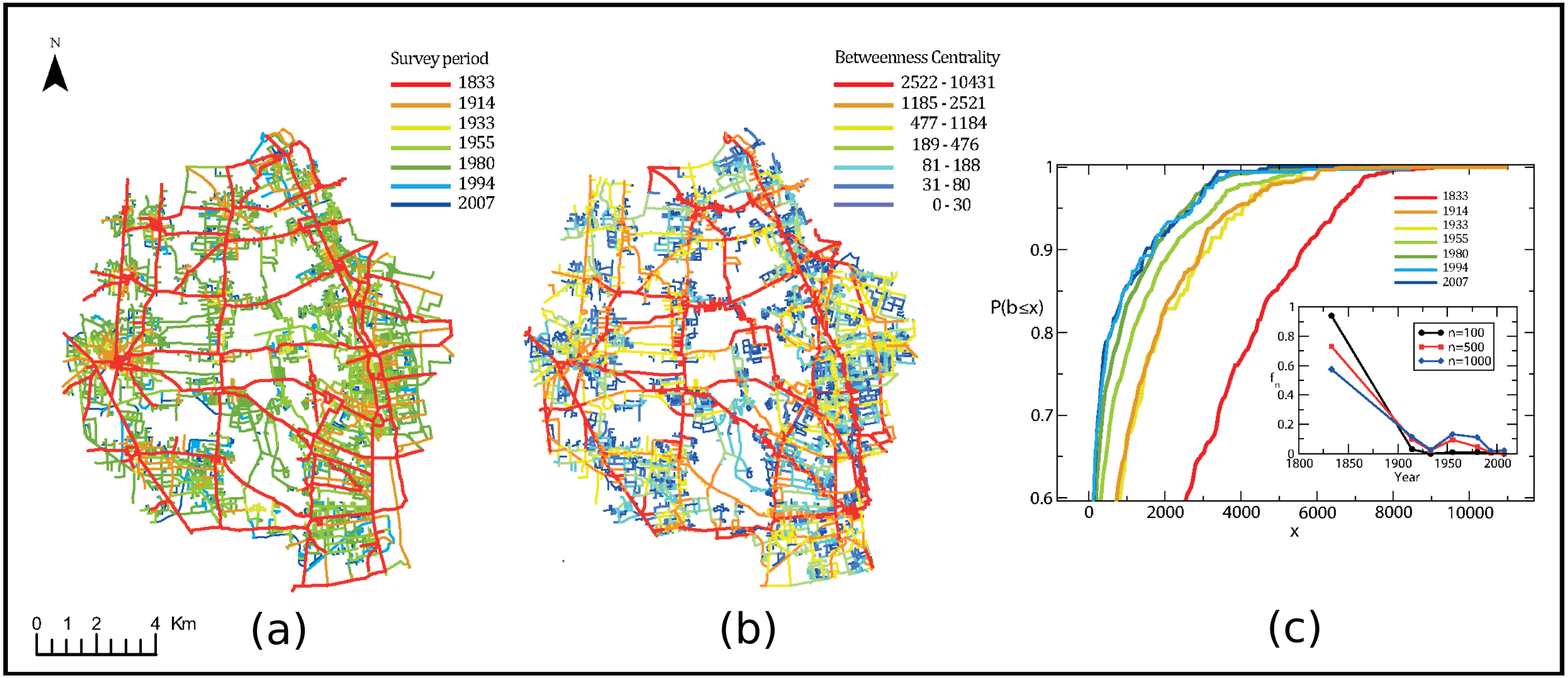


Figure 6 | Color maps indicating (a) the time of creation of each link and (b) its value of betweenness centrality (BC) at year 2007. (c) The cumulative distribution of BC of links added at different times. The inset reports the percentage of edges added at a certain time which are ranked in the top n positions according to the BC. Different curves correspond to $n = 100, 500, 1000$.

Characterization of new links: BC impact

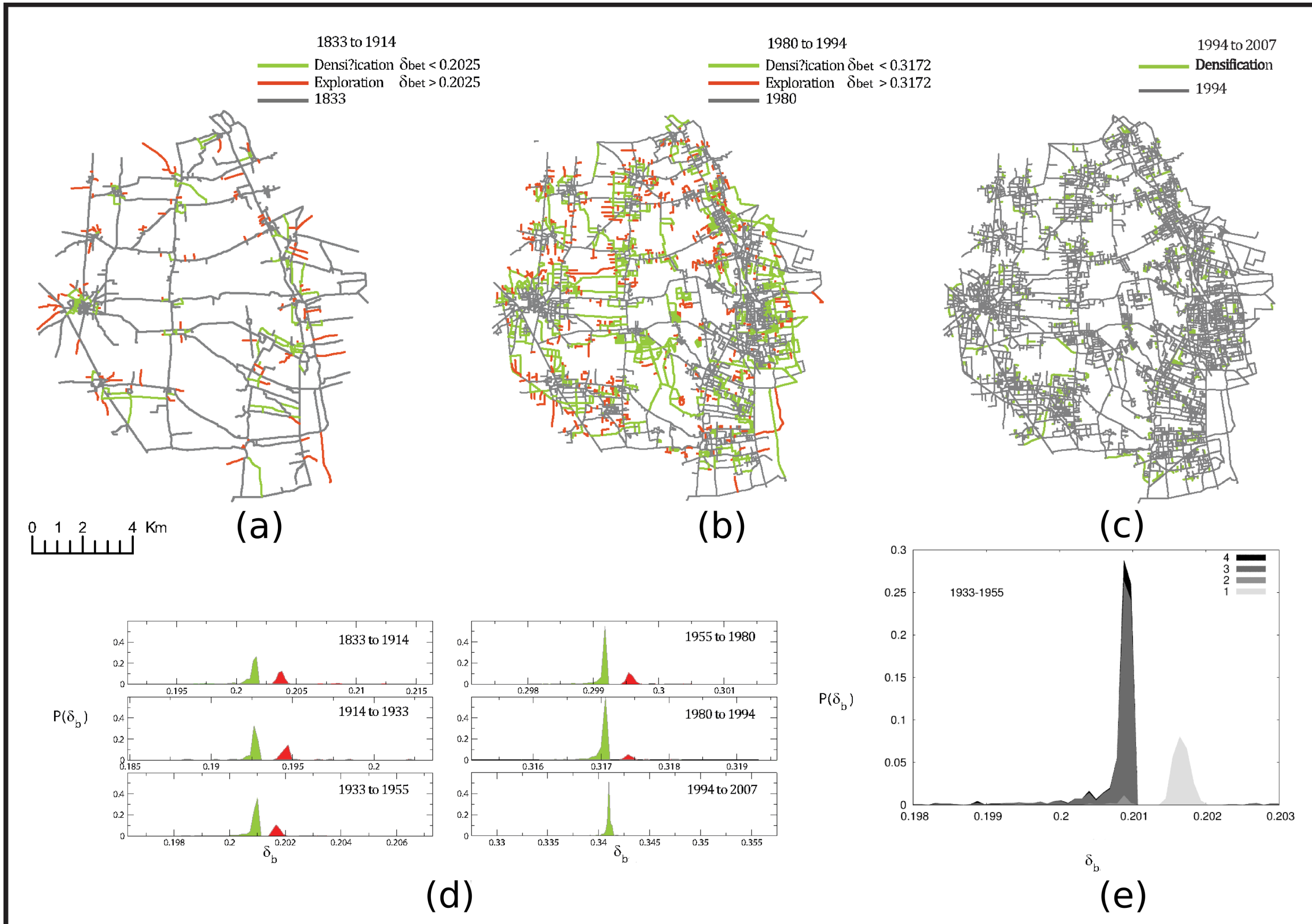
- Average BC of the graph at time t:

$$\bar{b}(G_t) = \frac{1}{(N(t)-1)(N(t)-2)} \sum_{e \in E_t} b(e)$$

- BC impact of new edge e*:

$$\delta b(e*) = \frac{\bar{b}(G_t) - \bar{b}(G_t \setminus \{e*\})}{\bar{b}(G_t)}$$

Evolution: two processes



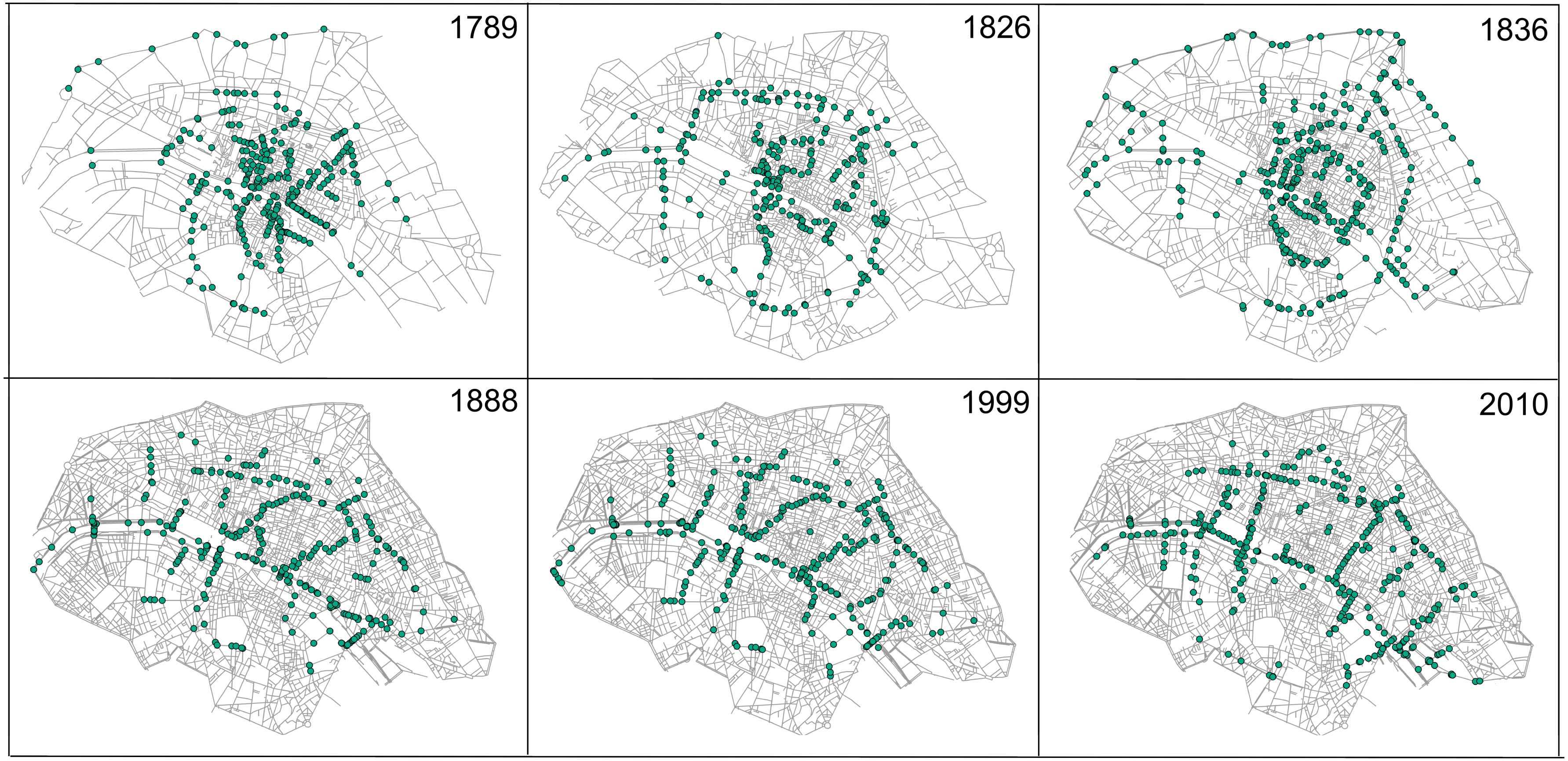
- Two different categories of new links: ‘densification’ and ‘exploration’ clearly identified by the BC impact

Structural effects of changes (the "Haussmann effect")

- Spatial distribution of centrality (most central nodes)

Structural effects of changes (the "Haussmann effect")

- Spatial distribution of centrality (most central nodes)



Scaling of the maximum of the betweenness centrality

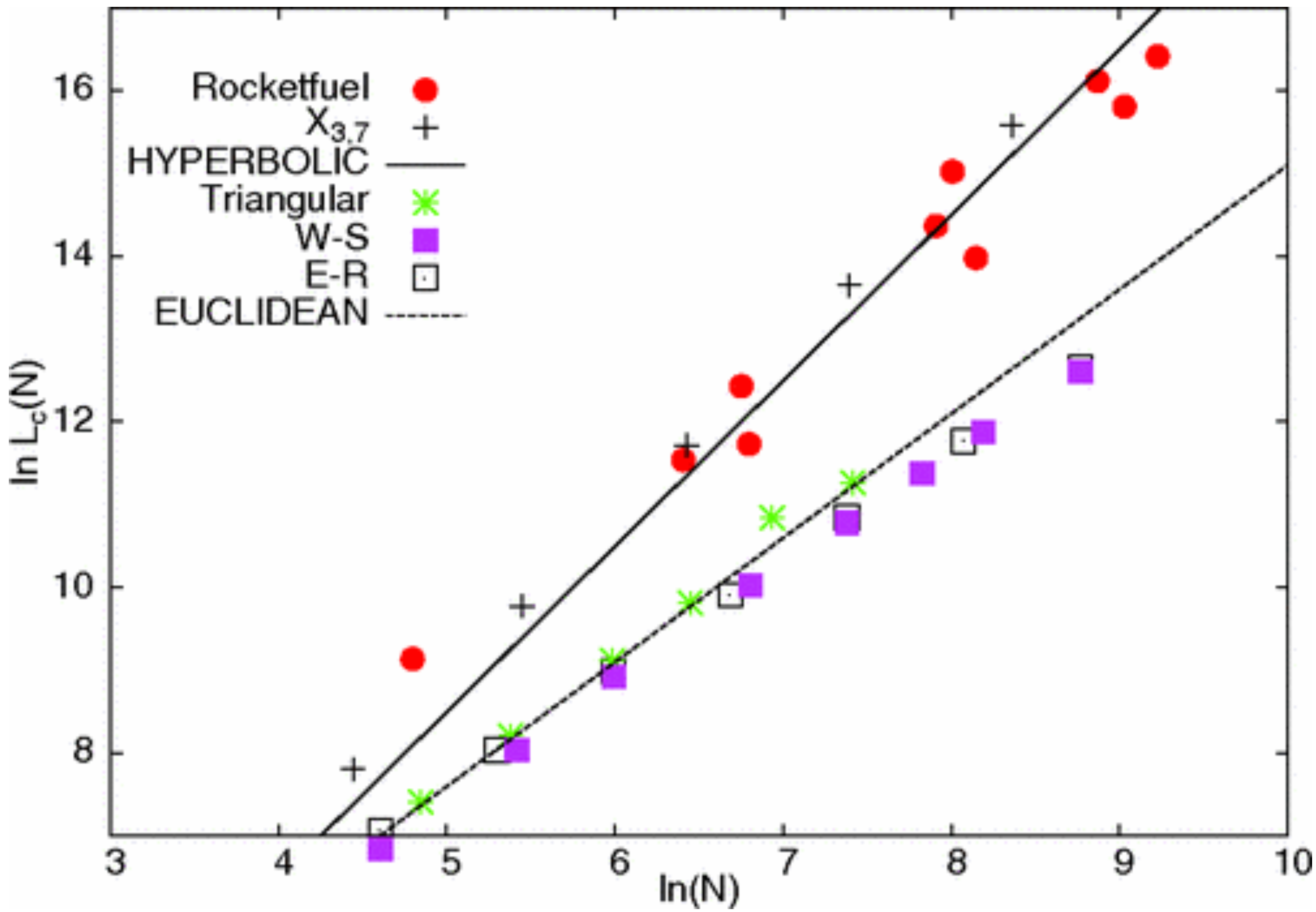
Traffic ≅ BC and maximum gives a bound on the load of nodes

Scaling with N depends on the curvature of the network

$$\max g \sim N^{\beta}$$

Lattice: $\beta = 2$

SW: $\beta = 3/2$

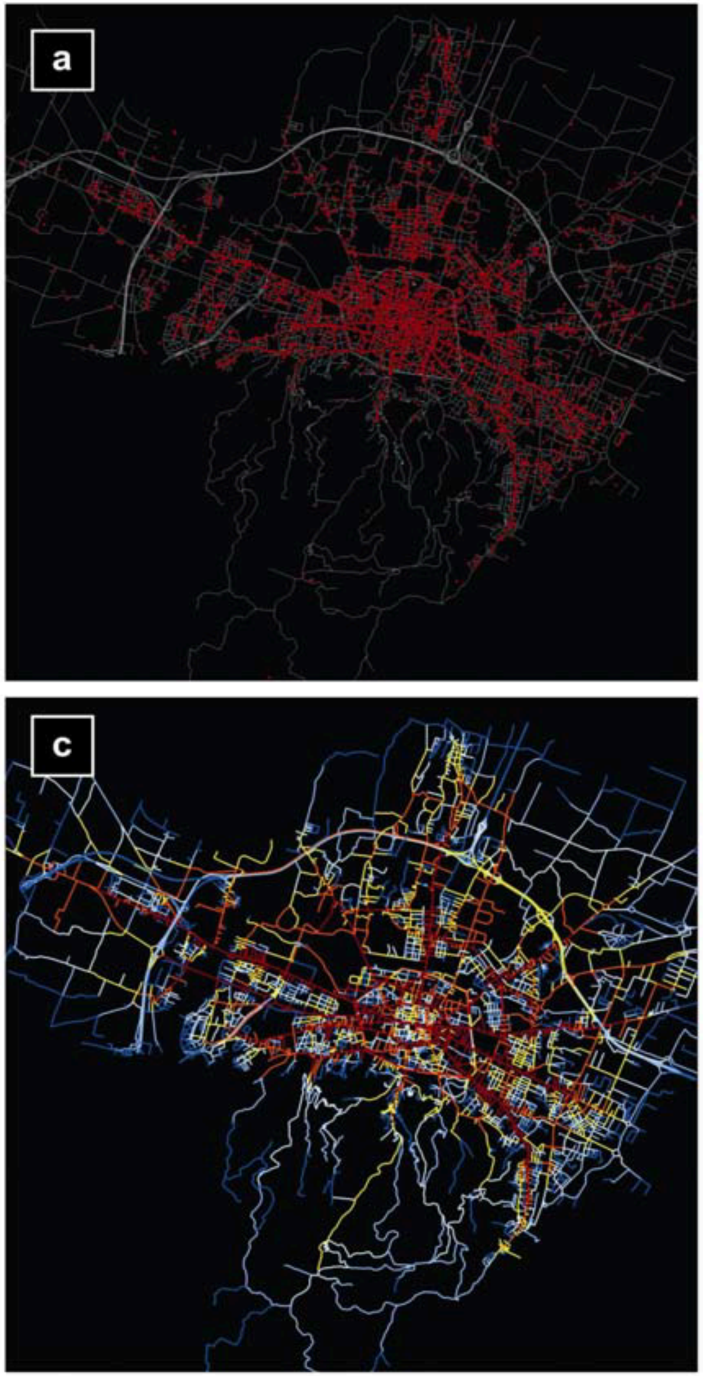


Narayan and Saniee, 2009

Betweenness centrality and activity

(a) location of commercial and service activities (red dots);

(c) street global betweenness BC (blue for lower values and red for higher);



S. Porta et al. Env Plan B, 2009

Summary

- The BC is correlated with many important features (traffic, socio-economic variables, …)

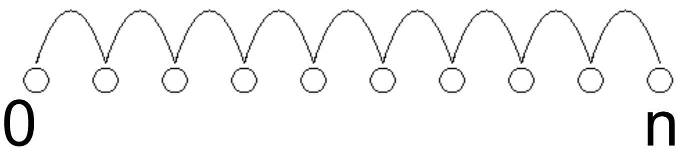
- The BC seems to reveal a lot about the structure of a spatial network: allows to track important structural changes in the evolution of a network

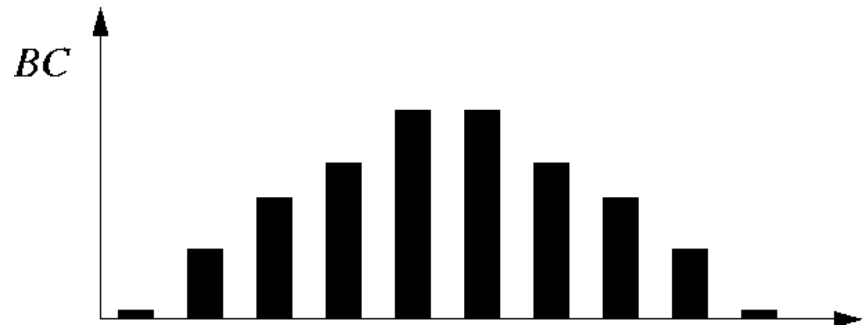
- Spatial distribution of centrality: where are the most central nodes

II. Simple remarks and empirical facts

Betweenness centrality and space

Regular lattice (1d)





$$g(x) = x(n-x)$$

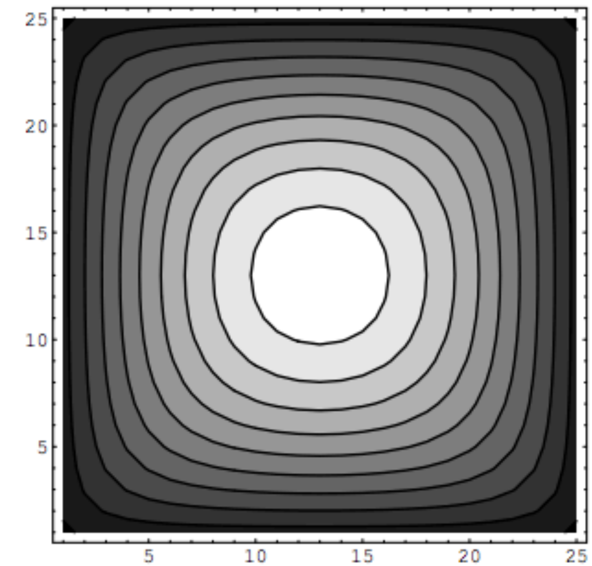
Maximum at the barycenter

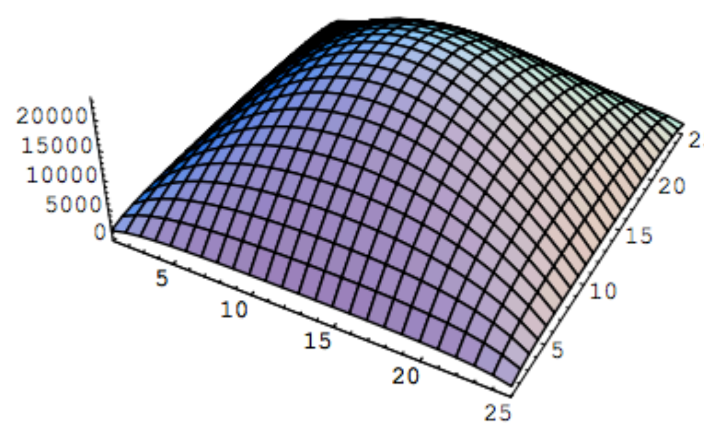
$$x^* = \frac{n}{2}$$

$$g_{max} \sim n^2$$

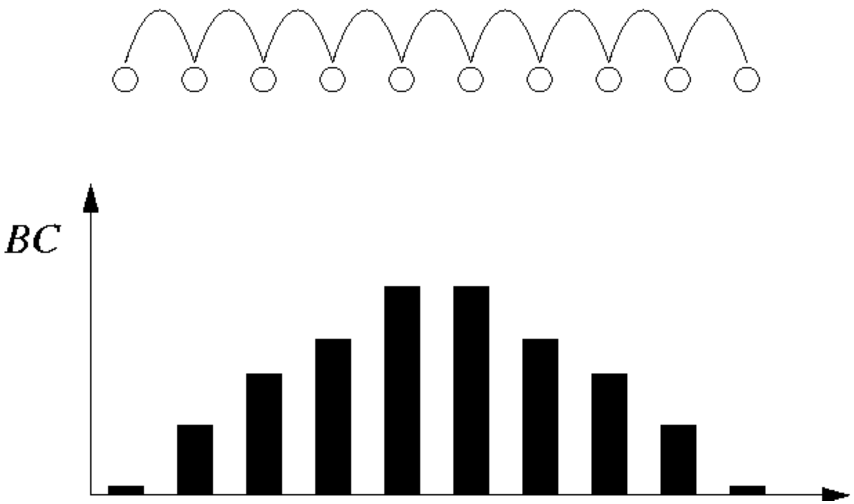
Betweenness centrality and space

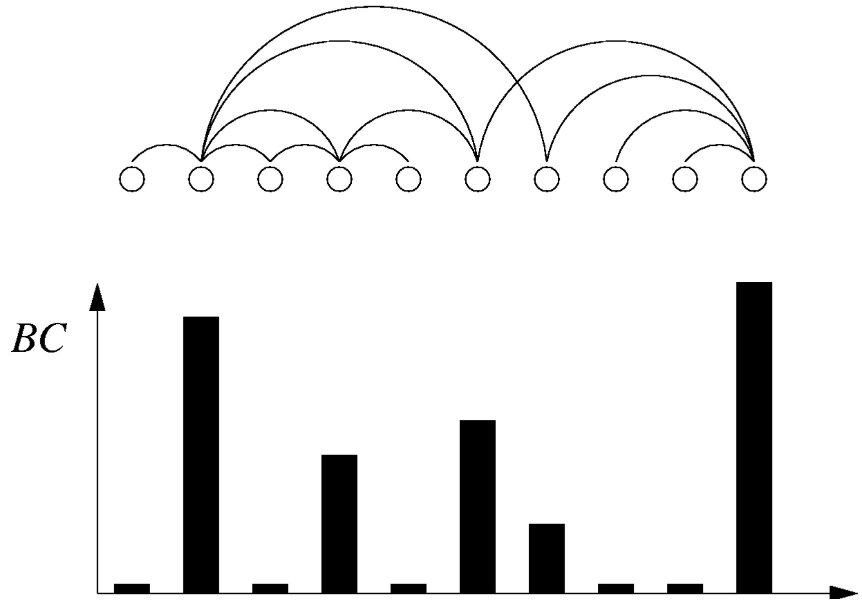
Regular lattice (2d)
Exact calculation ?
Surface with a maximum at (0,0)





Betweenness centrality and space





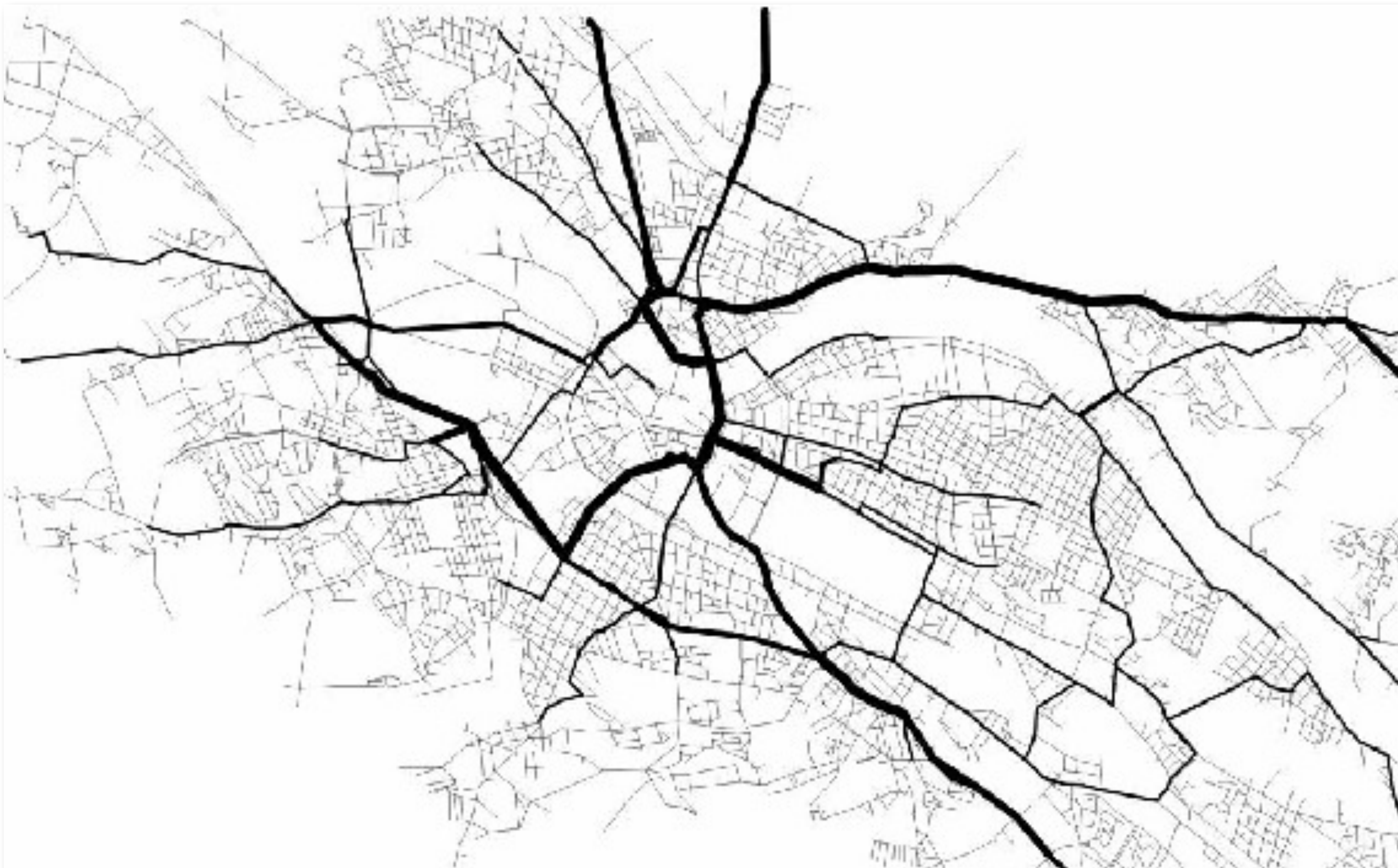
Large BC: small distance to the barycenter

Large BC: large degree

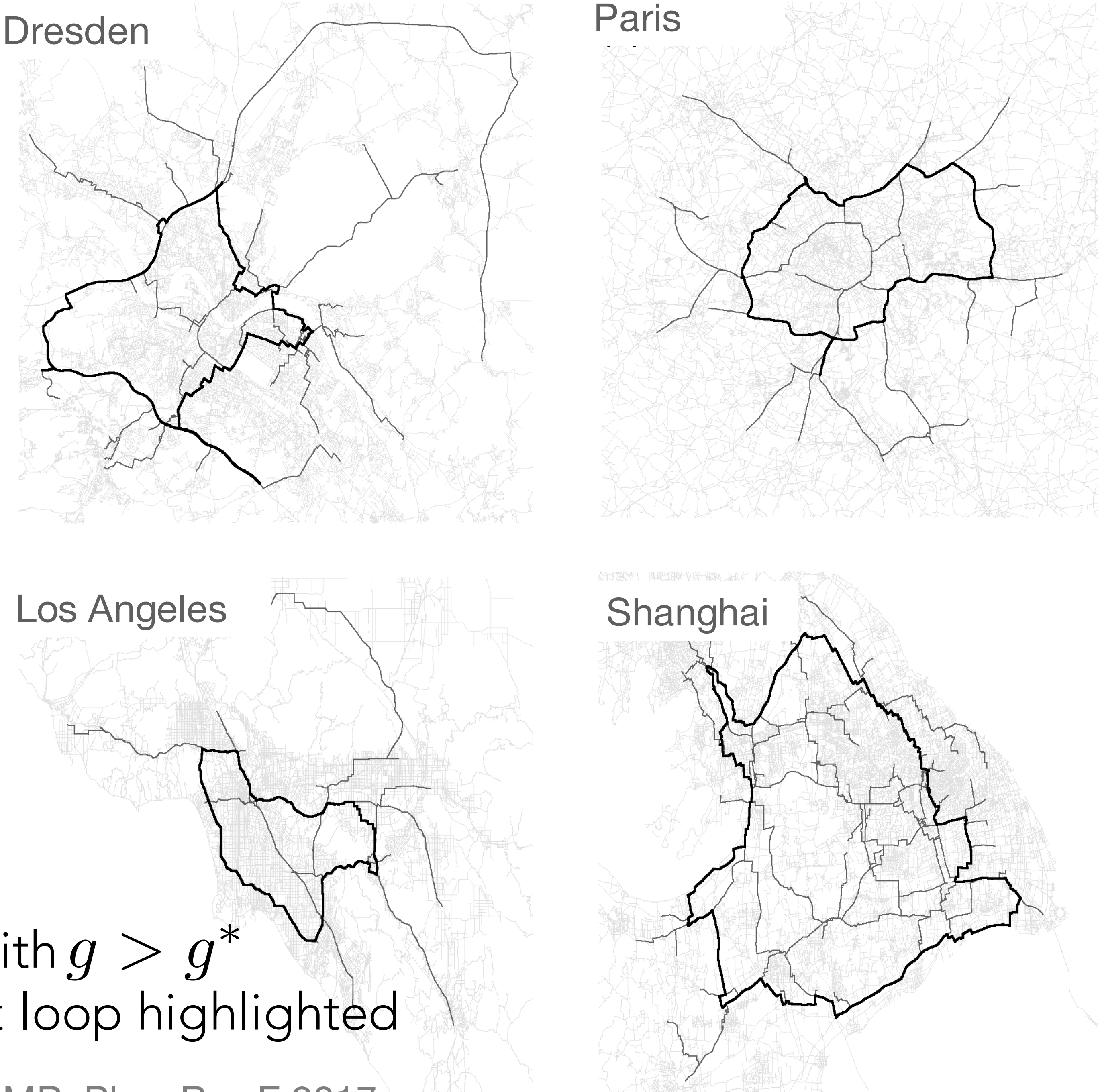
No space : $g(k) \sim k^{\eta}$

Betweenness centrality and space

With disorder: the BC is not a decreasing function of the distance to the barycenter. Patterns of large BC appear



Lammer et al, Physica A 2006



Links with $g > g^*$
Largest loop highlighted

B. Lion & MB, Phys Rev E 2017

Empirical facts: Summary

- The introduction of disorder in planar graphs induce in general the formation of non-trivial structures made of links with a large BC

- In particular, we observe the appearance of loops made of links that can have a BC larger than the barycenter

- In other words, disorder can invert the typical behavior observed for regular lattice where the BC is decreasing monotonously from the barycenter.

- In the following, we propose a toy model which allows to discuss and to understand under which conditions a loop can become more central than the spatial center.

III.1. A toy model

Can a loop be more central than the barycenter of nodes ?

A toy model

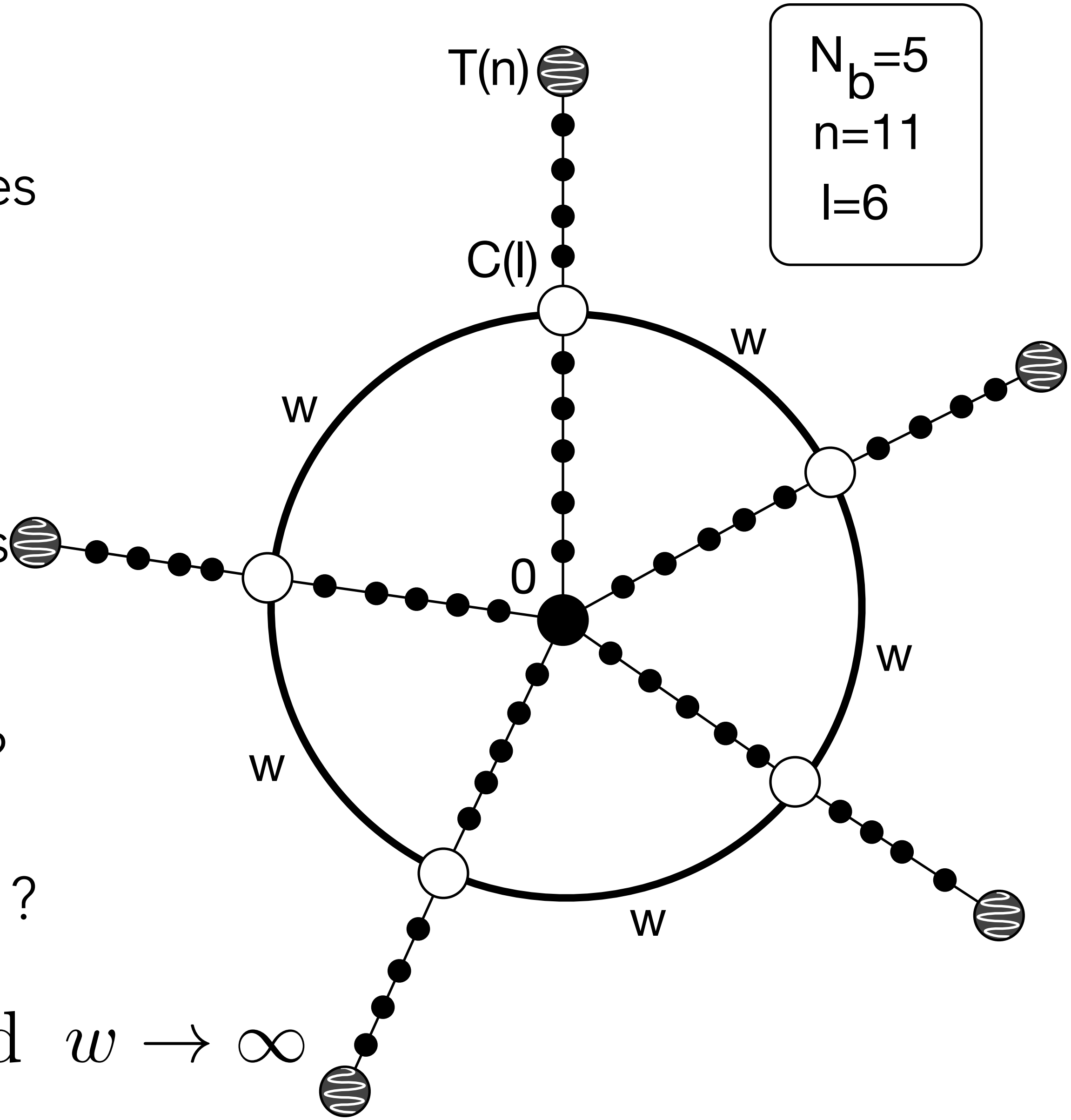
N_b branches with n nodes
+a loop at distance l

Links on branches: w=1
Links of loop: w
Weighted shortest paths

BC of origin O ?
BC of nodes on loop C ?

Condition for g(C)>g(O) ?

Two limits: $w \to 0$ and $w \to \infty$
Non-weighted case w=1



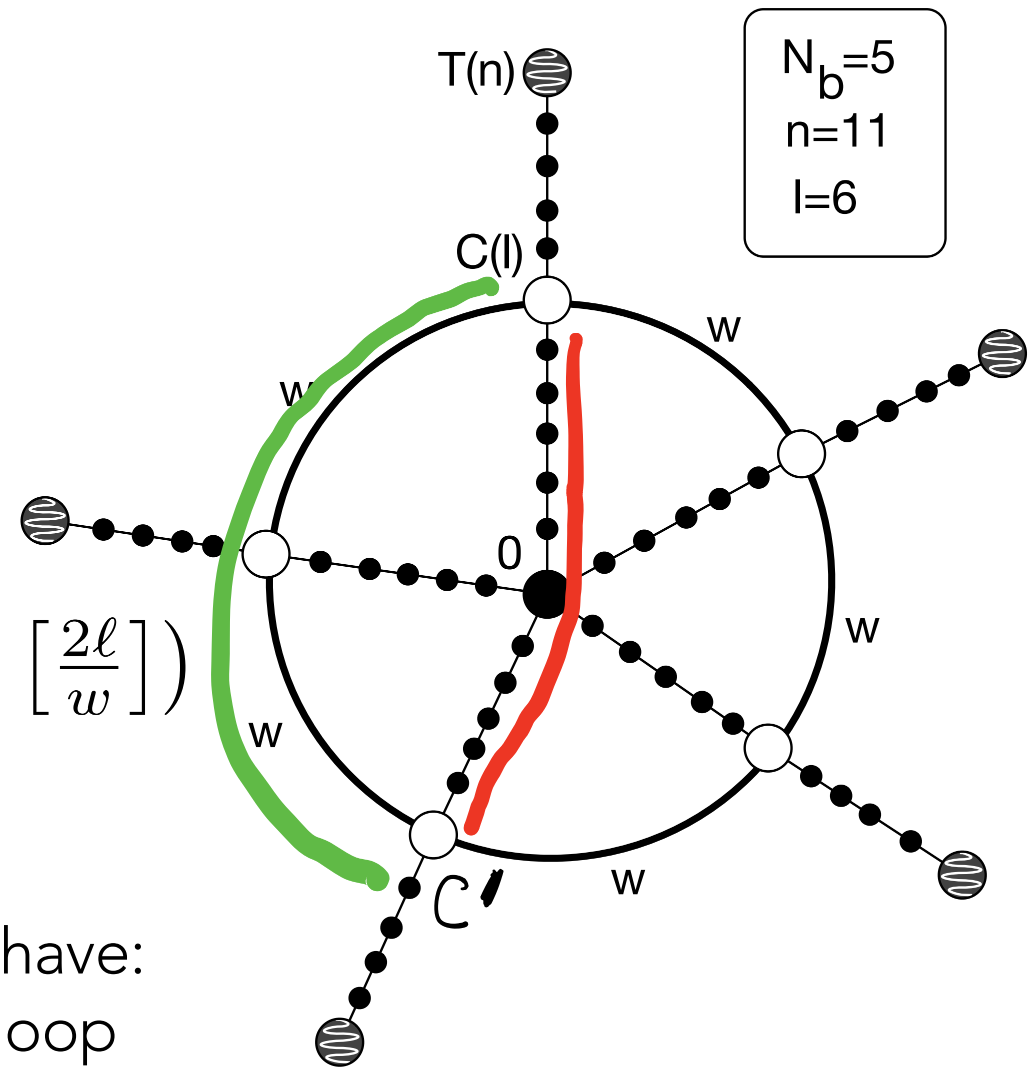
A toy model

Important quantity

$$2\ell \lessgtr wn\ ?$$

$$\chi = \min\left(\frac{N_b-1}{2}, \left[\frac{2\ell}{w}\right]\right)$$

For the branch j we have:
If $j < \chi$ => take the loop
If $j > \chi$ => take the radial path



A toy model

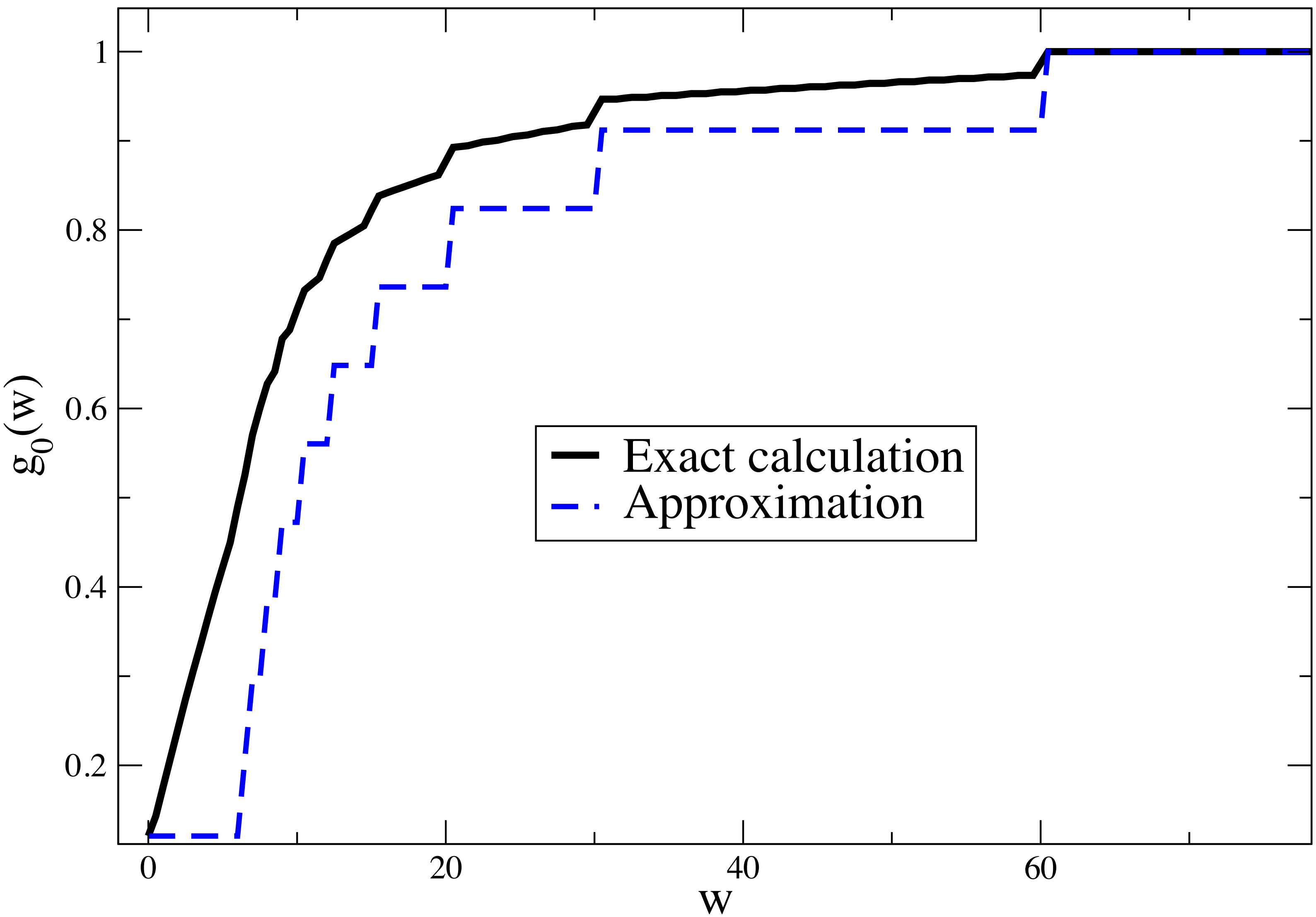
Approximate formulas
(exact closed forms ?)

$$g_0(w) \approx \left\{ \left(\frac{N_b-1}{2} - \chi \right) n^2 + \chi \frac{(\ell-1)(\ell-2)}{2} \right\} N_b$$

Far-away branches Near branches

which recovers both exact limits:

$$g_0(w) \approx \begin{cases} N_b \frac{N_b-1}{2} \frac{(\ell-1)(\ell-2)}{2} & \text{for } w \to 0 \\ n^2 \frac{N_b(N_b-1)}{2} & \text{for } w \to \infty \end{cases}$$



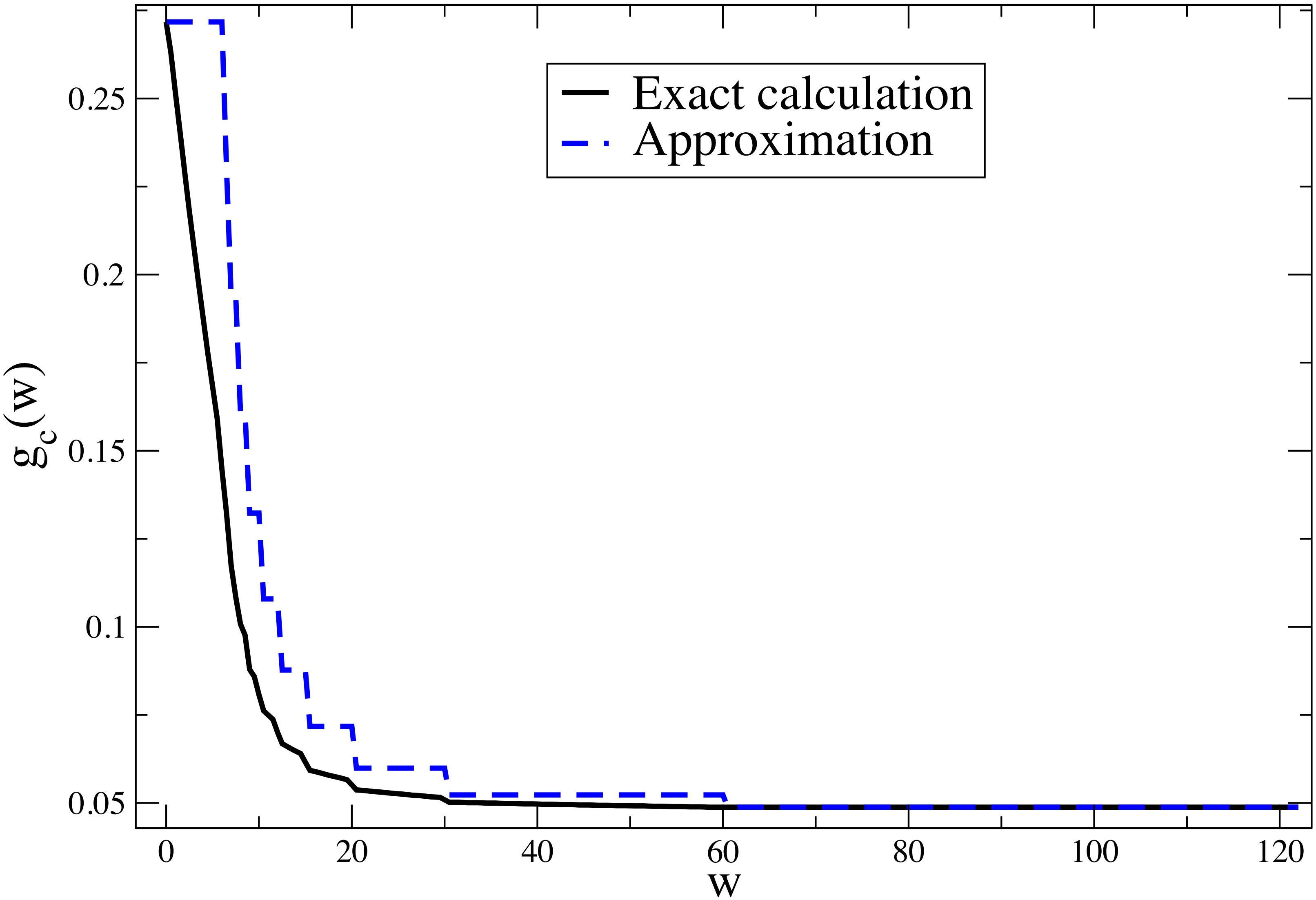
1
0.8
0.6
0.4
0.2
$g_0(w)$
0
20
40
60
w
Exact calculation
Approximation

A toy model

Approximate formulas

BC for C more involved (for l and n large and x=l/n fixed)

$$g_C(x=\ell/n) \approx (1-x)(x+N_b-1)+$$
$$2\chi x(1-\tfrac{x}{2})+\frac{\chi(\chi-1)}{2}\left(1-\frac{x^2}{2}\right)$$



Exact calculation
Approximation
0.25
0.2
0.15
0.1
0.05
$g_c(w)$
0
20
40
60
80
100
120
w

A toy model

- We have expressions for g_0 and g_C
- We can now compare g_0 and g_C but a simple argument allows to understand the `physics'.

A toy model: a simple argument

- Optimal value of g_C versus l

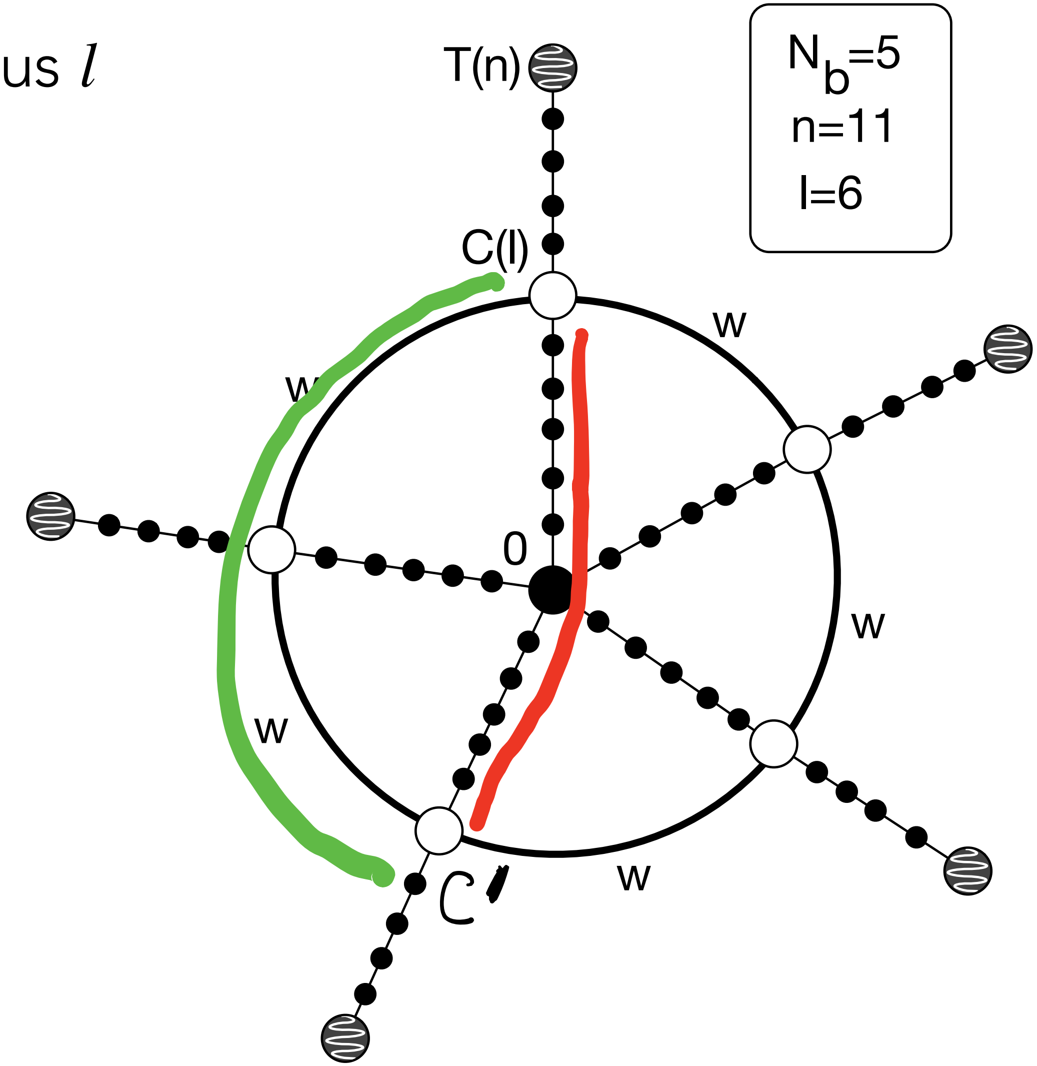
If l too small: most paths will go through O

If l is too large, nodes of lower branches will go through O again

$$2\ell_{opt} = w\frac{N_b - 1}{2}$$

$$\ell_{opt} = \frac{w(N_b - 1)}{4}$$

(is in fact an exact result)



A toy model: a simple argument

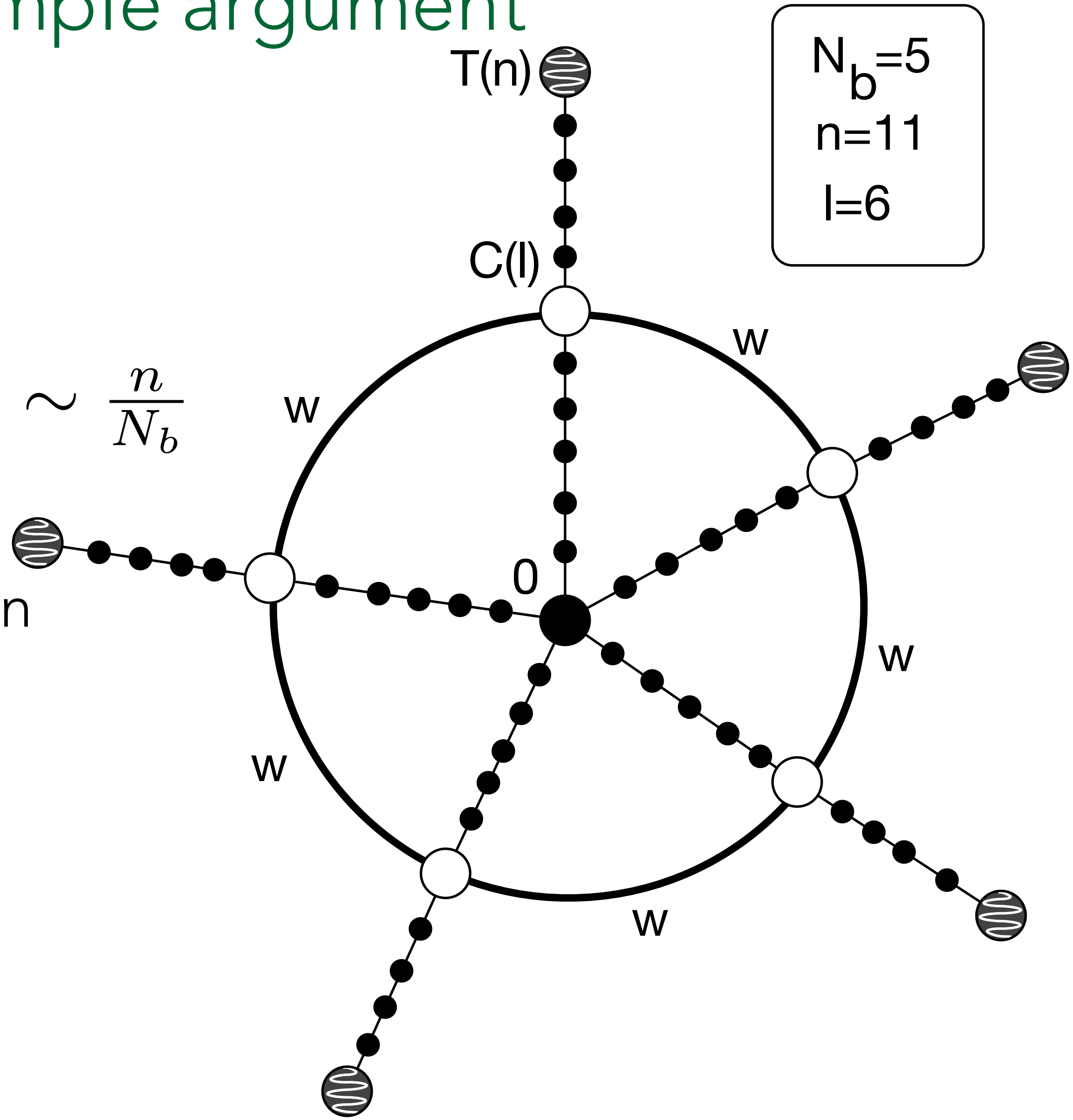
$$\ell_{opt} = \frac{w(N_b - 1)}{4}$$

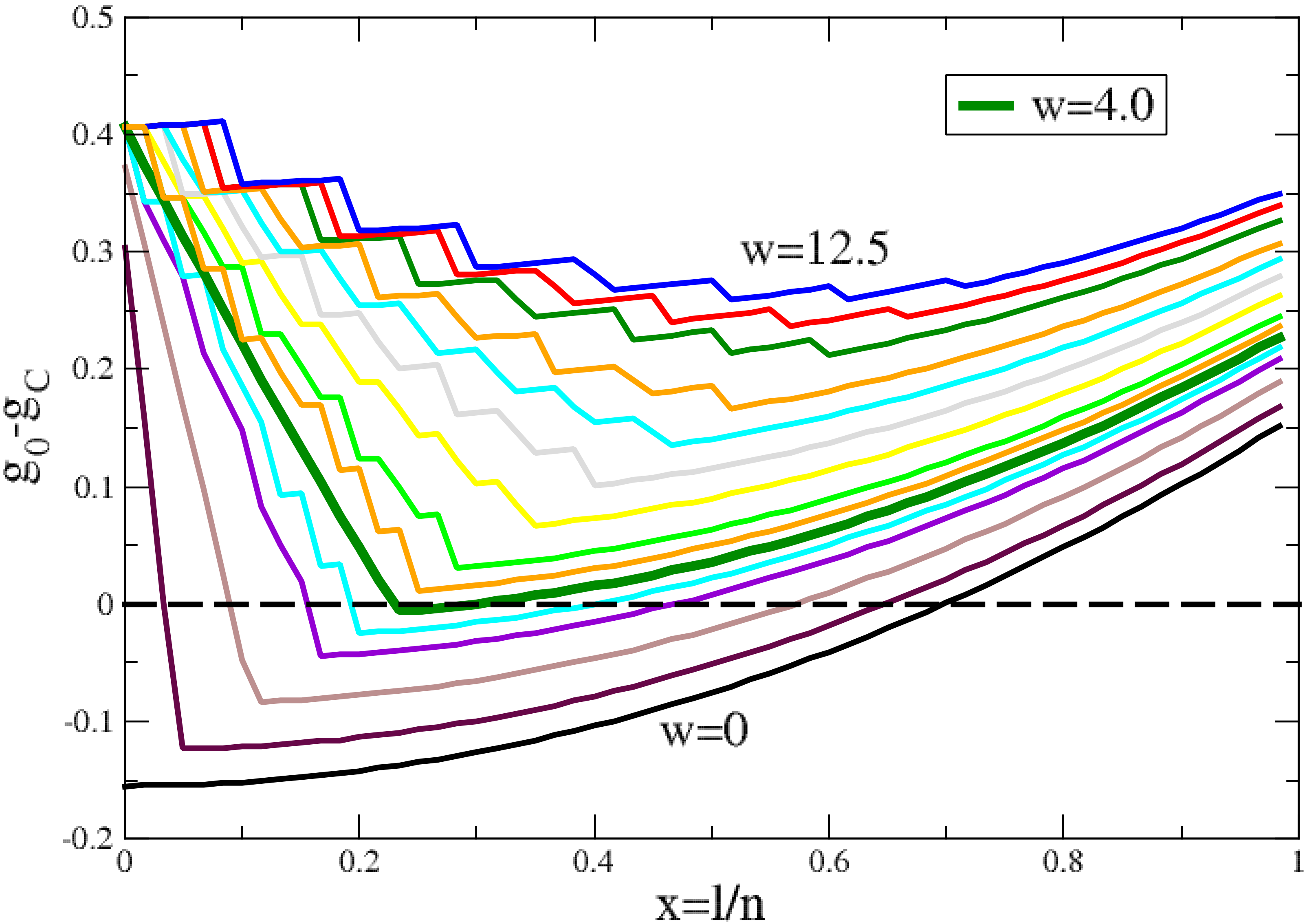
$$\ell_{opt} < n \Rightarrow w < w_c \sim \frac{n}{N_b}$$

Existence of a transition for a value $w = w_c$:

$w > w_c$ $g_0 > g_C$

$w < w_c$ $\exists \ell$ s.t. $g_0 < g_C$




w=4.0
w=12.5
w=0
$g_0 - g_C$
x=l/n
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
0
0.2
0.4
0.6
0.8
1

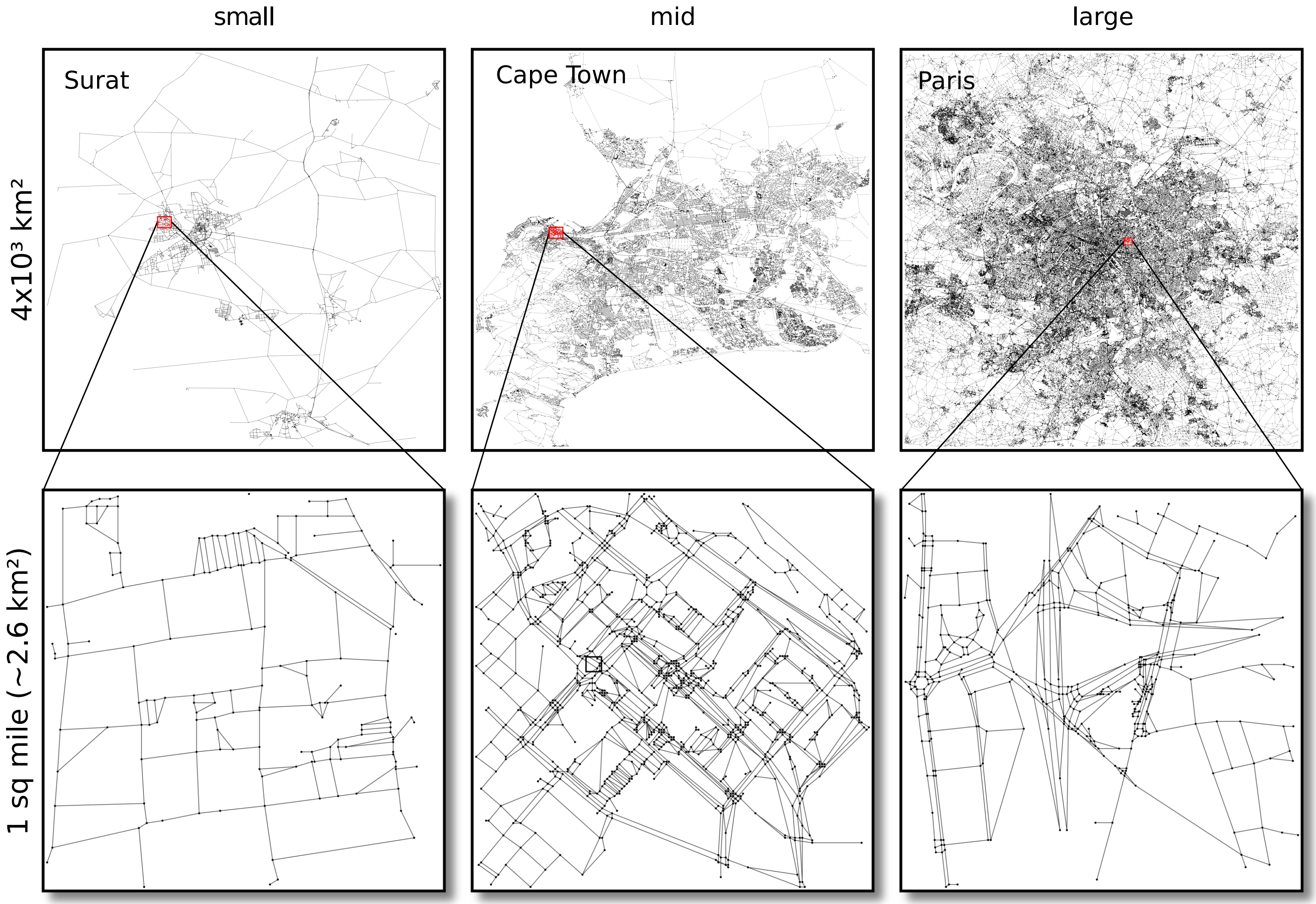
A toy model

- The loop can be more central than the center !
- In particular if w_c>1 then even in the non-weighted case (w=1) this can happen
- $w_c=n/N_b$ which implies that n has to be large enough compared to N_b
- Not the end of the story but we see here that the radial structure impacts the patterns of high BC nodes

III.2. P(g) as a Structural invariant

Determinants of the BC distribution P(g) ?

97 street networks: all continents, all sizes



Kirkley, Barbosa, MB, Ghoshal
in preparation (2017)

The spatial distribution of high BC nodes

- Metrics with threshold on BC: $g>\theta$

- Extent of high BC nodes: $$C_\theta = \frac{\sum_{i=1}^{N_\theta} ||x_i - x_G||}{\sum_{i=1}^{N} ||x_i - x_G||} \frac{N}{N_\theta}$$

- Anisotropy of the set of nodes: $$A_\theta = \frac{\lambda_1}{\lambda_2}$$
(eigenvalues of the inertia tensor)

- Detour factor: $$D = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{d_R(i,j)}{d_E(i,j)}$$

- Control parameter:

$$\rho_e = \frac{E}{E_{DT}} \qquad (\sim \langle k \rangle)$$

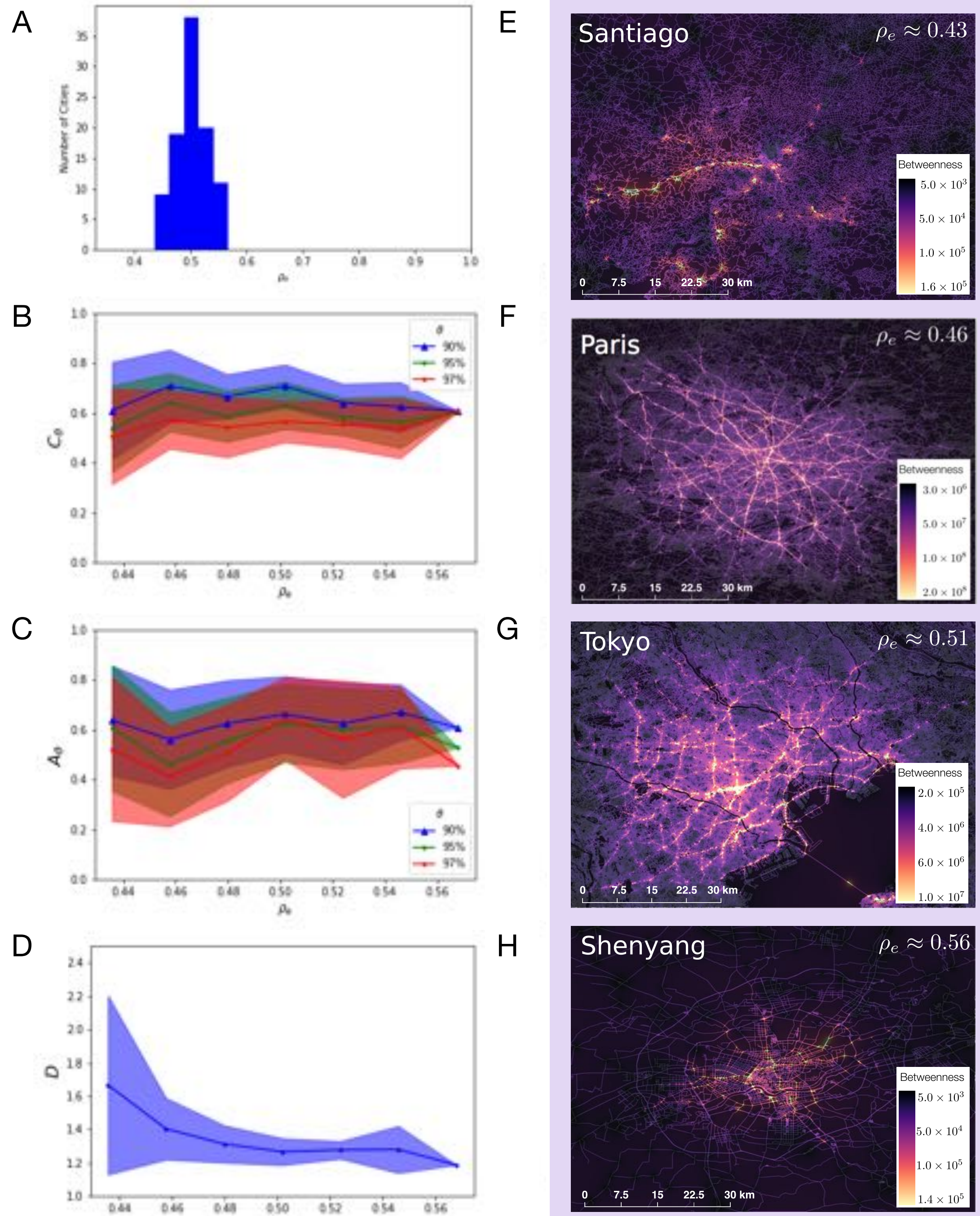
Kirkley, Barbosa, MB, Ghoshal
in preparation (2017)

Metrics for the 97 cities

A: Density distribution
Peaked

B & C: C_θ and A_θ almost constant

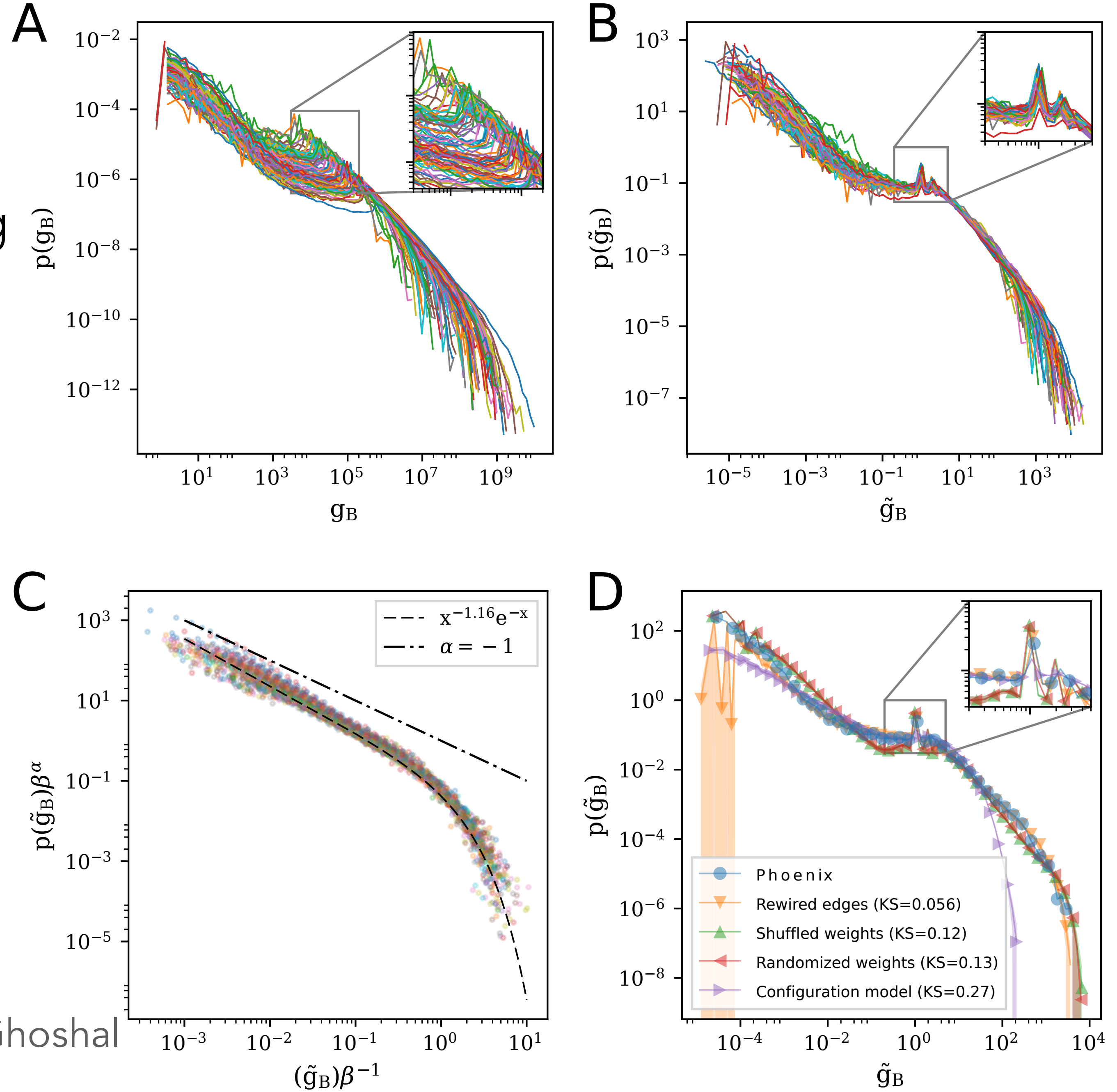
D: decrease ('transition'?) of the detour factor



Kirkley, Barbosa, MB, Ghoshal
in preparation (2017)

Distribution P(g)

- A&B: Rescaling $\tilde{g} = g/N$



- C: Tail $P(\tilde{g}) \sim \frac{1}{\tilde{g}}$

- D: Only planarity seems to matter !

Kirkley, Barbosa, MB, Ghoshal in preparation (2017)

Understanding these results: a toy model

- MST

- Delaunay triangulation & remove randomly links until desired density ρ_e

Kirkley, Barbosa, MB, Ghoshal
in preparation (2017)

Understanding these results: a toy model

A: MST

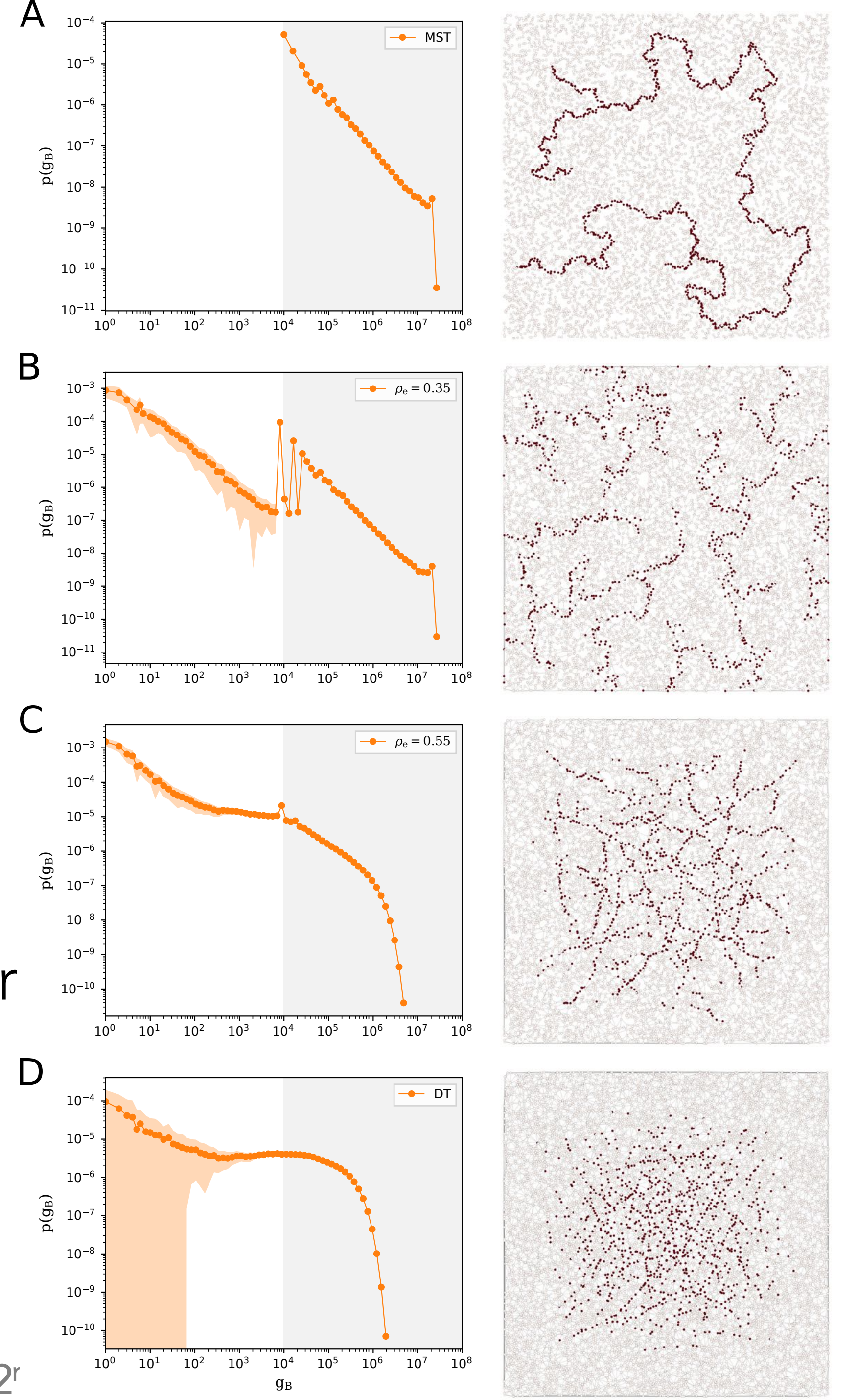
$$g \in [N, N^2]$$

B-D: Delaunay triangulation &
Remove randomly links
until desired density

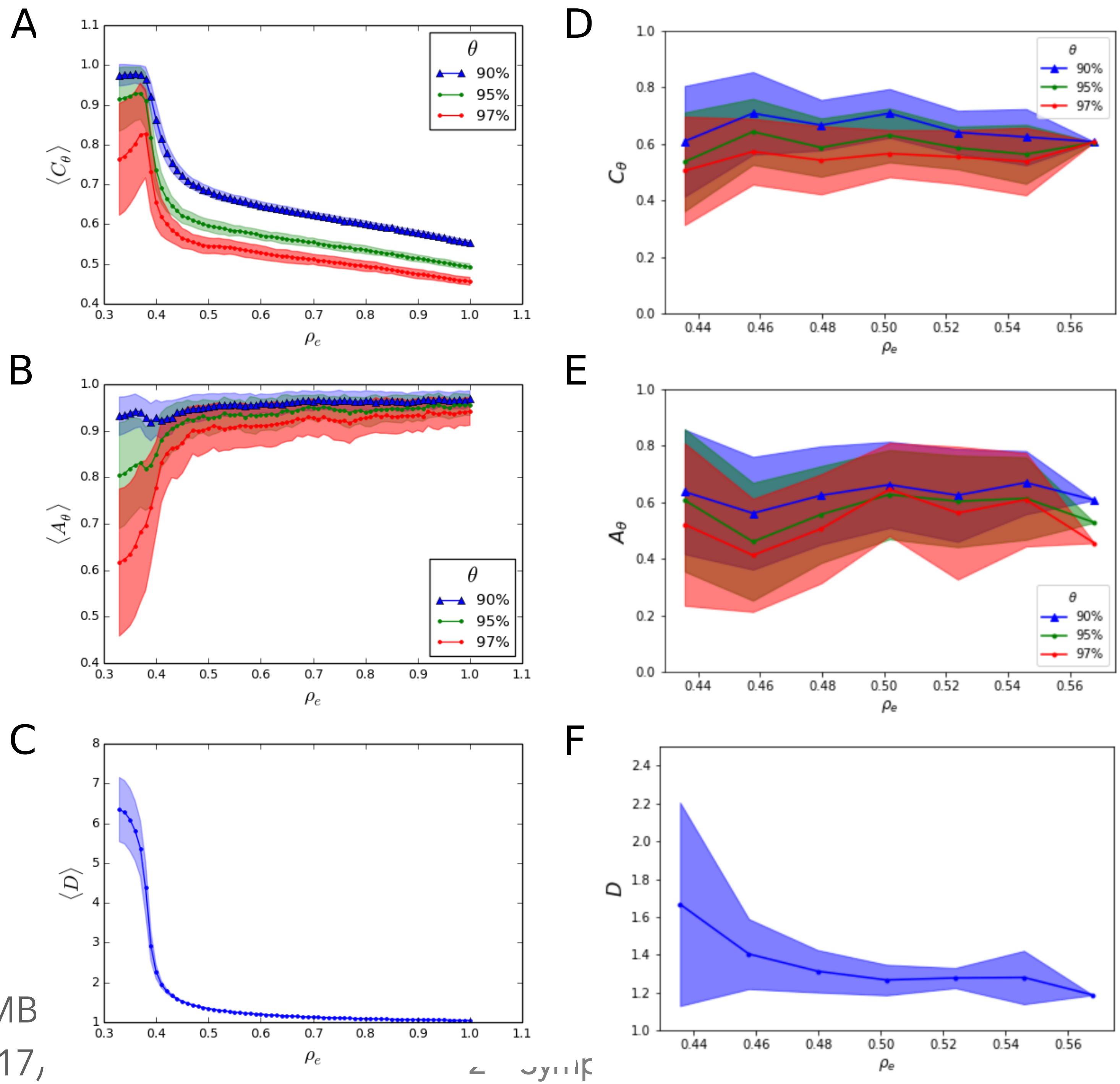
Adding links: increases the number of loops=> smaller values

$$g \in [1, N]$$

Kirkley, Barbosa, MB, Ghoshal
in preparation (2017)



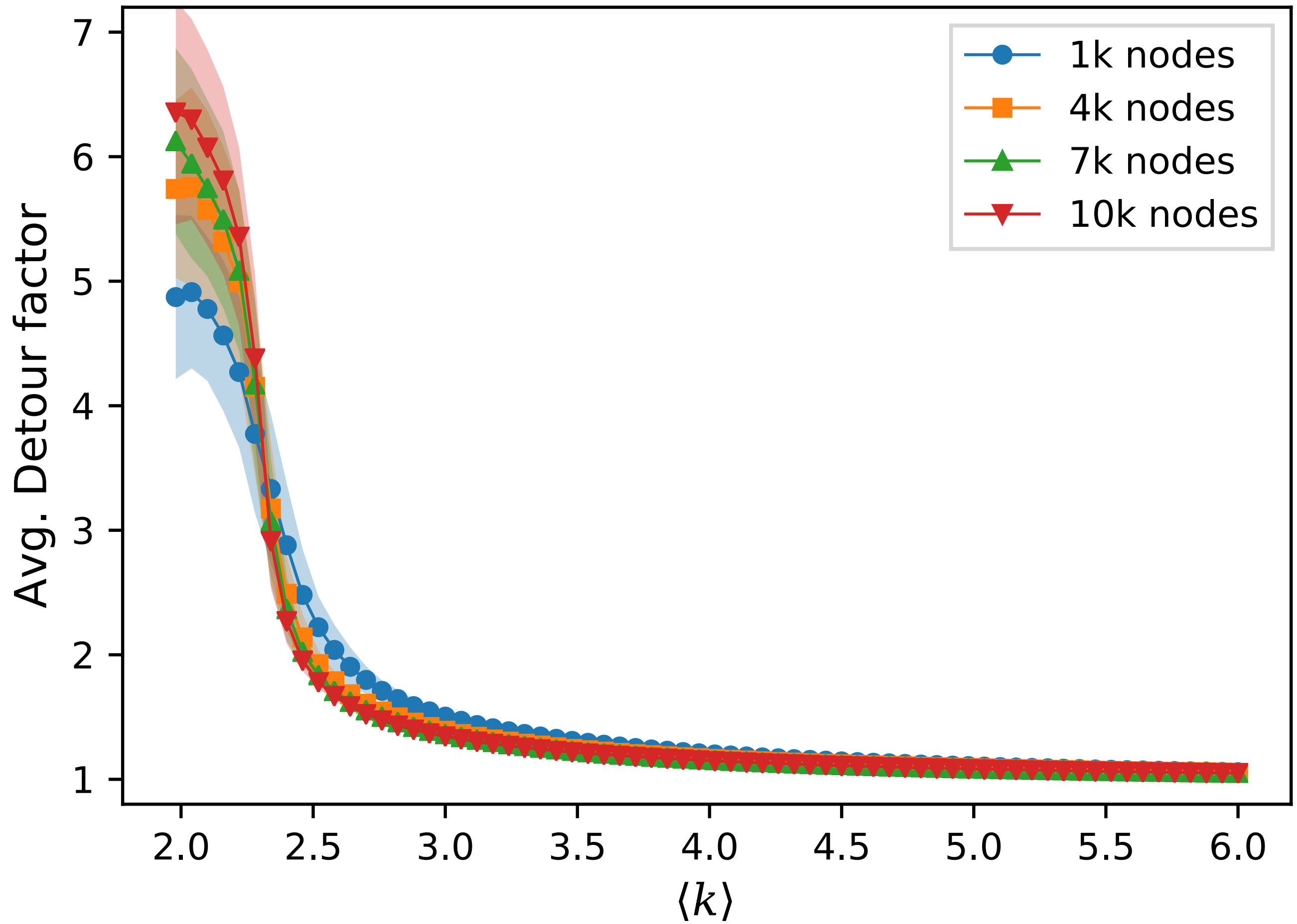
Metrics: model vs. Cities



Kirkley, Barbosa, MB
in preparation (2017,

Detour factor: transition

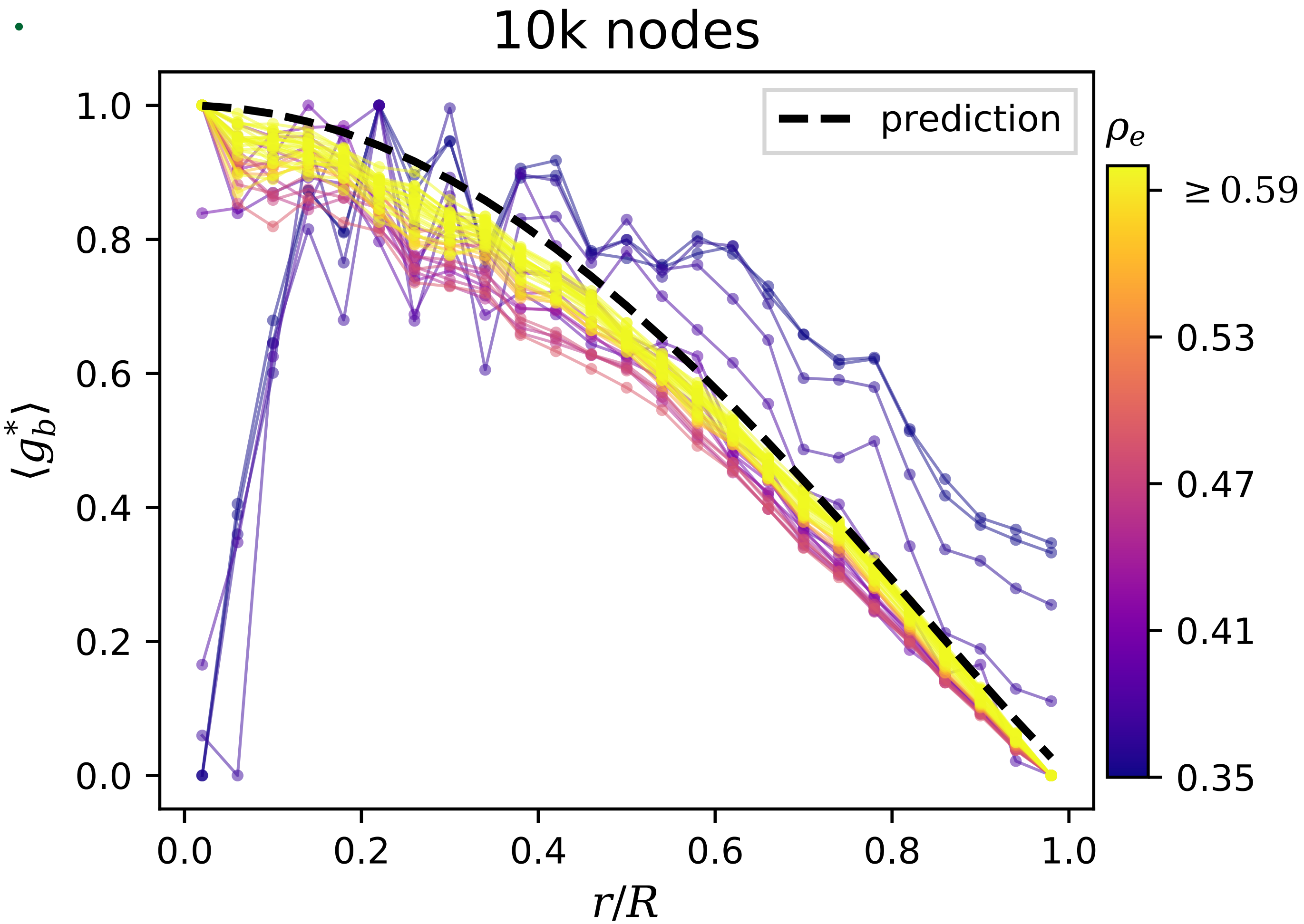
Very quickly, when adding edges a straight shortest path can be found !



Kirkley, Barbosa, MB, Ghoshal in preparation (2017)

Detour factor: transition

If we have straight shortest paths: the calculation by Giles, Georgiou and Dettmann should work



Analytical curve:
Giles, Georgiou, Dettmann, IEEE 2015

Kirkley, Barbosa, MB, Ghoshal
in preparation (2017)

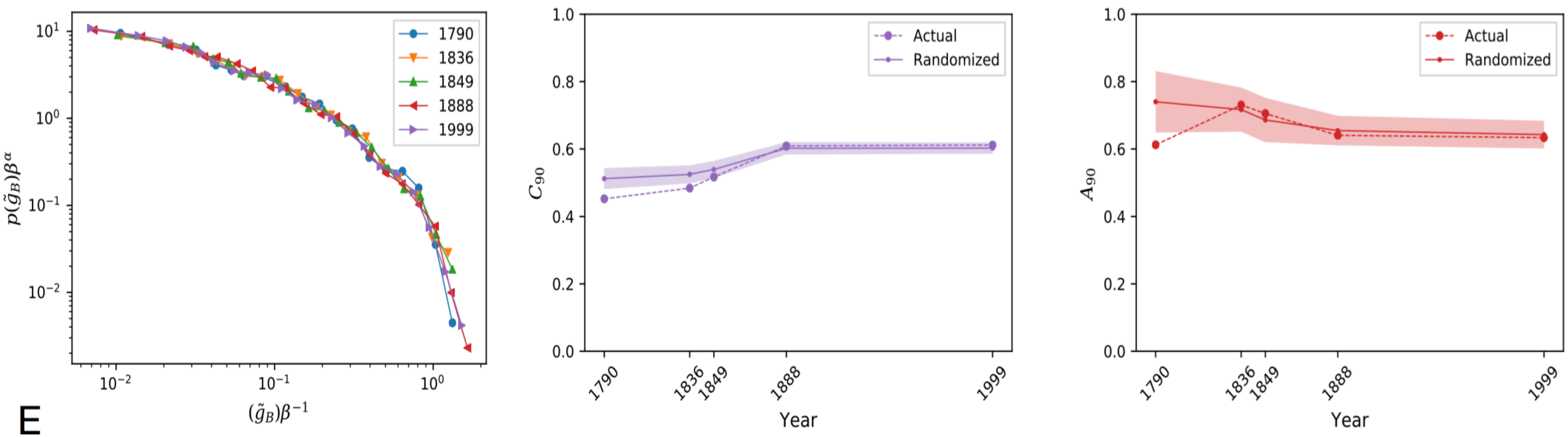
Evolving networks case

1790 ($\rho_e = 0.50$) **1836** ($\rho_e = 0.52$) **1849** ($\rho_e = 0.53$)

1888 ($\rho_e = 0.50$) **1999** ($\rho_e = 0.52$)

Kirkley, Barbosa, MB, Ghoshal
in preparation (2017)

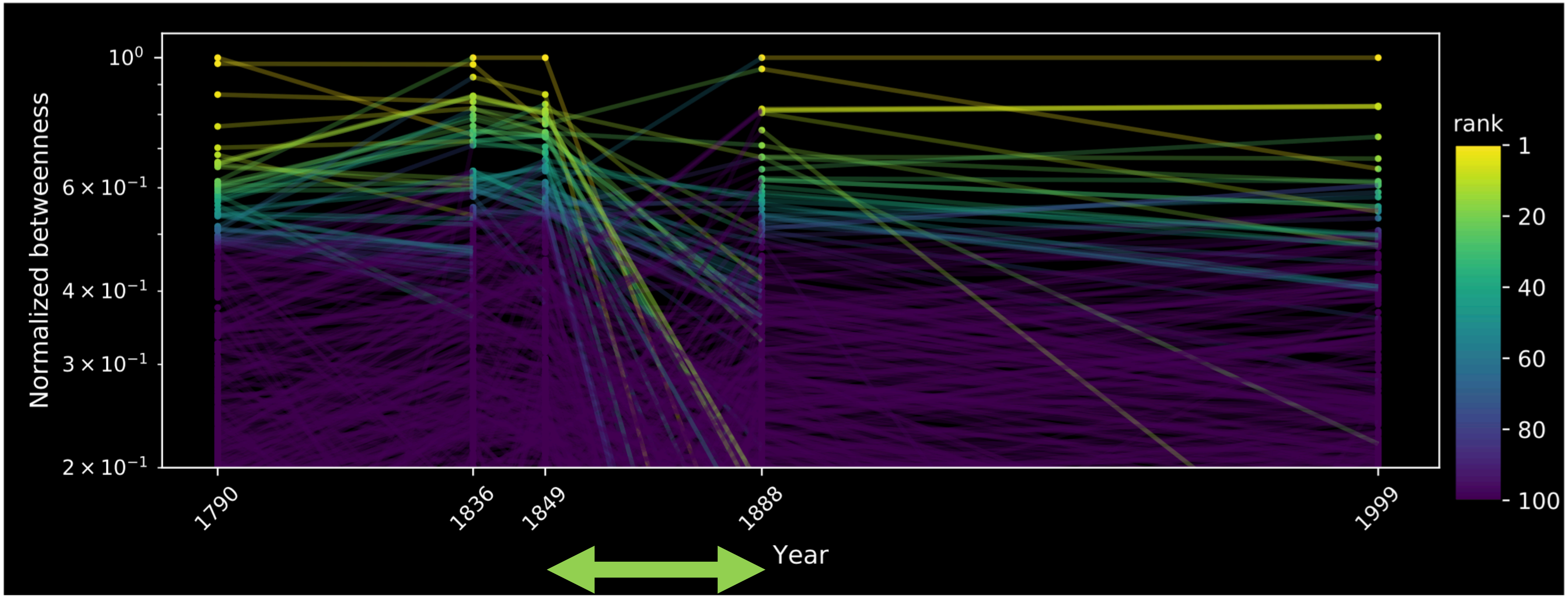
Evolving networks: high BC nodes



- Invariant BC distribution
- Constant C and A (the density of edges is almost constant)
- consistent with the simple model

Kirkley, Barbosa, MB, Ghoshal
in preparation (2017)

Evolving networks: high BC nodes



Haussmann period

Kirkley, Barbosa, MB, Ghoshal
in preparation (2017)

III.3. Some remarks on the dense limit

The BC distribution in the very dense limit of spatial networks

The BC for dense networks

Calculation (in the dense limit of the RGG) proposed by
Giles, Georgiou, Dettmann (IEEE, 2015)

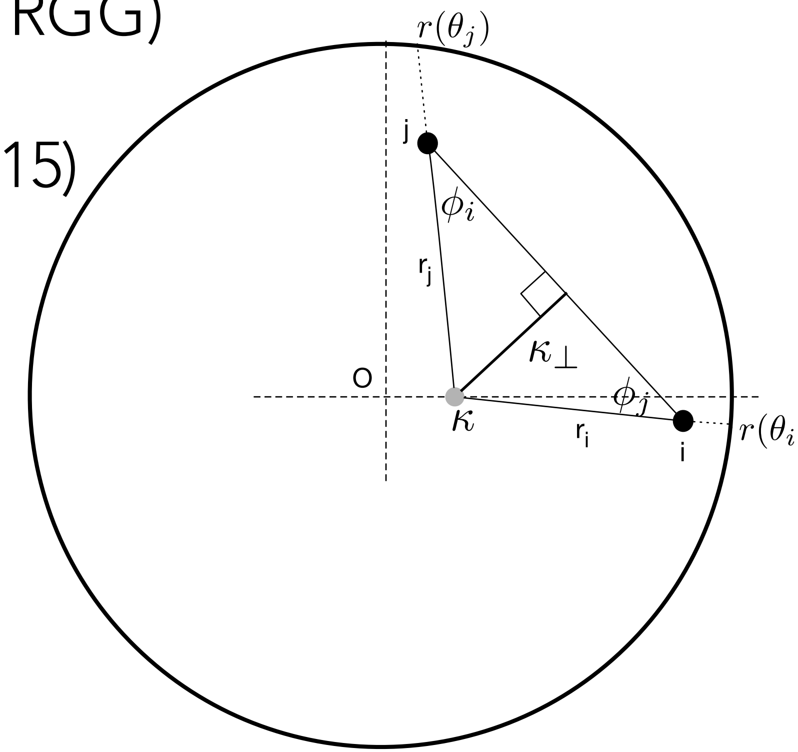
They assume statistical isotropy

$g = g(\kappa)$

where κ is the distance to the center
For a node at distance κ

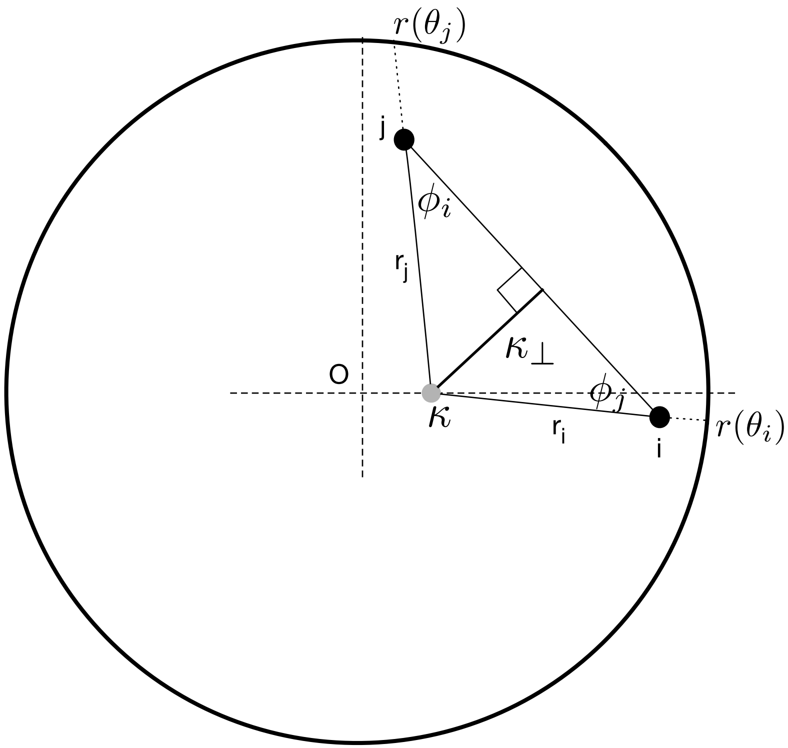
$$g(\kappa) = \int \frac{dr_i}{V}\frac{dr_j}{V}\chi_{ij}(\kappa)$$

$$\chi_{ij}(\kappa) = 1 \ \text{ if } \ \kappa \in \ SP(i,j)$$



The BC for dense networks

Main assumption: the network is dense enough so that shortest paths are straight lines



κ belongs to the shortest path from i to j if $\kappa_\perp = 0$

$$g(\kappa) = \frac{1}{V^2} \int \mathrm{d}r_i \mathrm{d}r_j \delta(\kappa_\perp)$$

$$\Rightarrow \quad \frac{g(\kappa)}{g(0)} = \frac{2}{\pi} \left(1 - \frac{\kappa^2}{R^2}\right) E\left(\frac{\kappa}{R}\right)$$

E(x) Complete elliptic 2[nd] kind -not too different from $\pi/2$

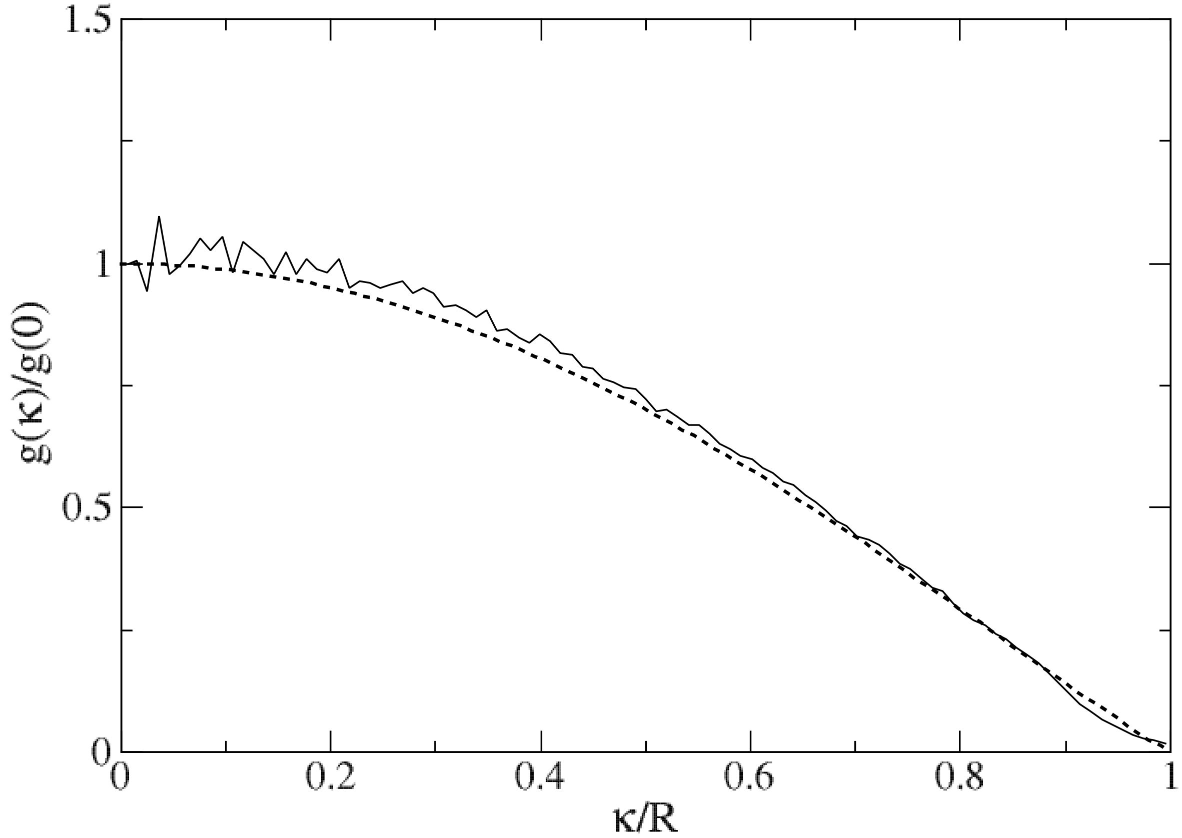
$$g(\kappa)/g(0) \sim 1 - \kappa^2/R^2$$

Giles, Georgiou, Dettmann, IEEE 2015

The BC for dense networks

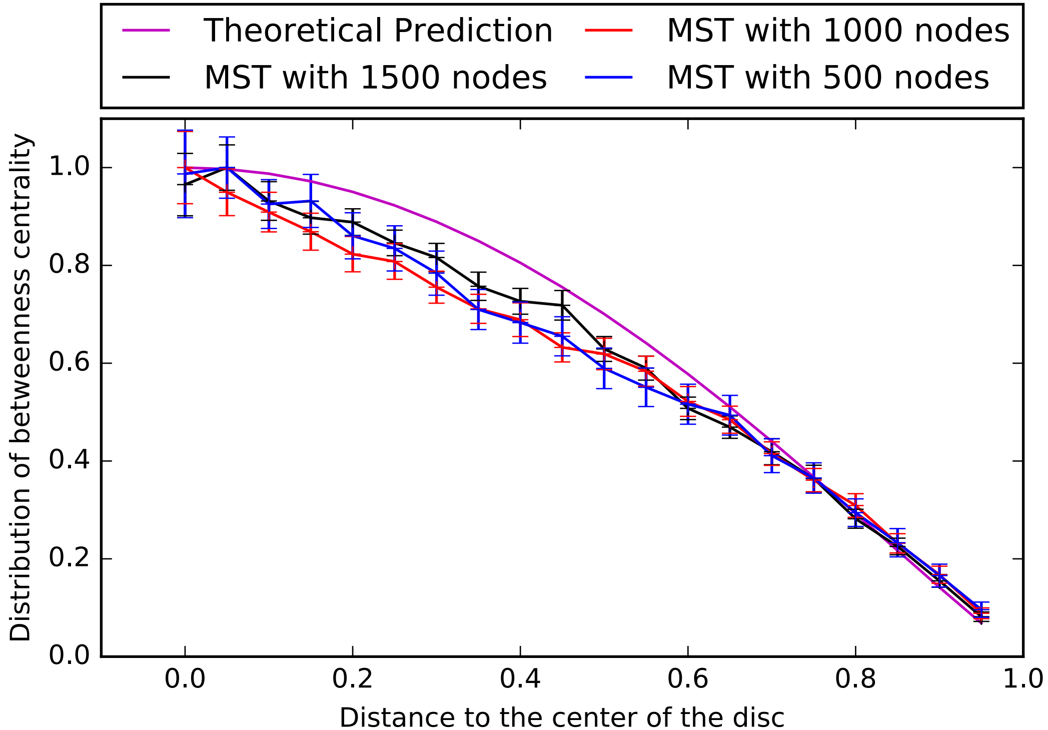
RGG dense limit
Unit disk R=1

$$\rho = N/\pi = 200$$



The BC for dense networks

This result is
- independent from the topology of the graph (!)
- No information about the topology and should be valid for many graphs that are dense enough and s.t. shortest paths are straight lines

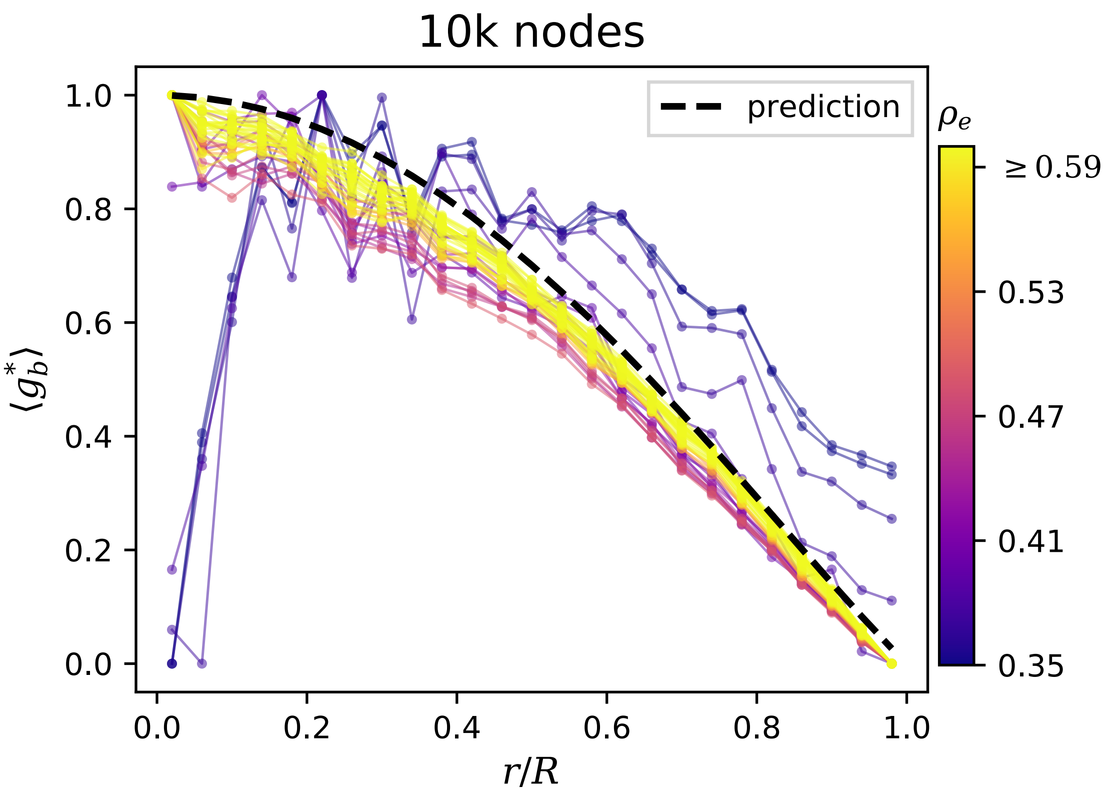


A. Bourges and MB, Internship report (2017)

The BC for dense networks

This result is

- independent from the topology of the graph (!)
- No information about the topology and should be valid for many graphs that are dense enough



A. Bourges and MB, Internship report (2017)

The BC for dense networks

Expansion in 1/density (?)

$$g(r) = \underbrace{g_\infty(r)}_{'universal'} + \frac{A}{\rho} + \mathcal{O}\left(\frac{1}{\rho^2}\right)$$

where A depends on the graph (via some correlation function, typically)

Discussion

- The BC is probably the simplest among interesting quantities that we can compute for spatial networks

- The BC distribution
  - Is invariant and depends only on the planarity constraint
  - The first –low BC – part [1,N] is controlled by loops
  - The tail $[N,N^2]$ is governed by the MST
  - If you stay planar you will always have large BC nodes…you can only play on the spatial pattern (urban planning: become nonplanar !)

- Evolution of planar graphs
  - Spatial structure of high BC nodes: control parameter(s) ?
  - Transition for the detour index
  - Large density: universal BC behavior (?); expansion ?

Thank you for your attention.

http://www.quanturb.com
marc.barthelemy@cea.fr

Students:

A. Bourges (Master Student)
B. Lion (Master student)

Collaborators:

M. Batty	H. Barbosa	H. Berestycki
P. Bordin	S. Dobson	G. Goushal
A. Kirkley	V. Latora	V. Nicosia
S. Shay	E. Strano	S. Porta

Book announcement:

Morphogenesis of Spatial Networks (Springer fall 2017)

Additional slides

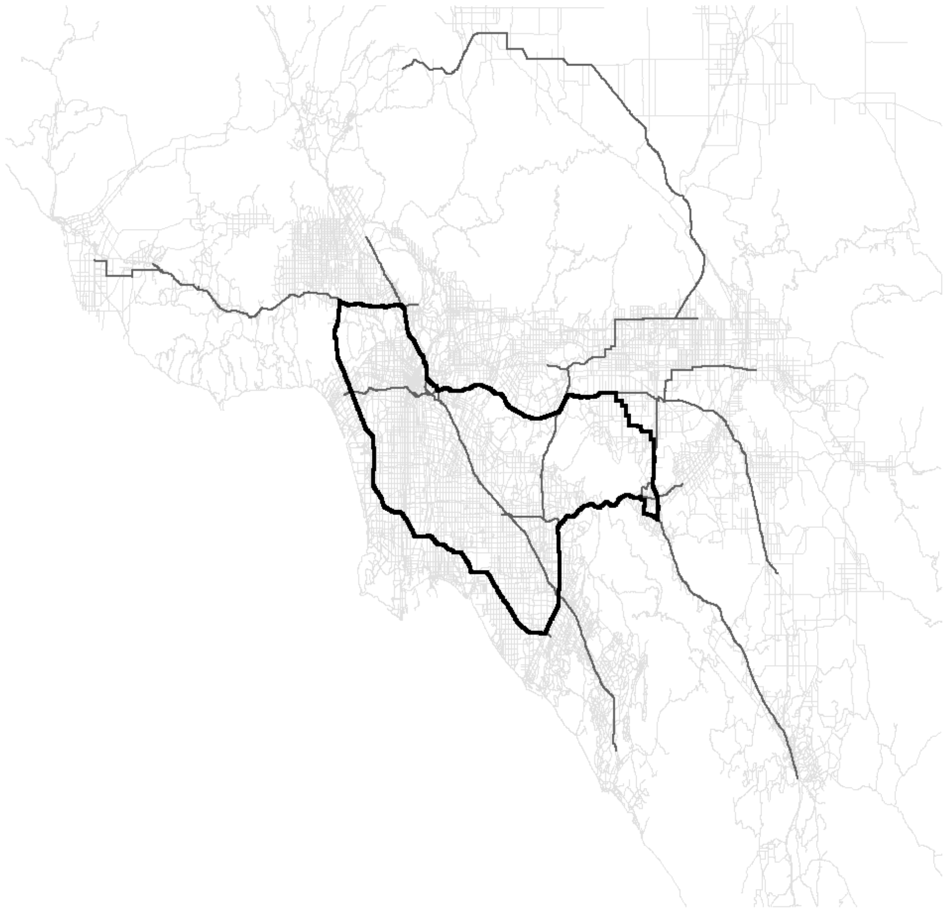
Spatial networks: importance, needs

- Spatial networks are present in many practical applications
  - Transportation networks (roads, subway, …)
  - Infrastructure networks (power grids, water distribution, …)
  - Biological networks (veination pattern, …)
- Need for 'global' characterization of graphs
  - Characterization the structure at a large scale: 'motifs'
  - Compare different graphs
  - 'User-friendly'

Betweenness centrality and space

Paris
Lion & Barthelemy, 2016

Betweenness centrality and space



LA
Lion & Barthelemy, 2016

Be e

Dresden
B. Lion & MB, Phys Rev E 2017

Betweenness centrality and space

Shanghai
Lion & Barthelemy, 2016

Perimeter of the largest loop