Quantum PageRank for Complex Networks

Giuseppe Davide Paparo

UCM, Madrid, Spain

Joint work with M.A. Martin-Delgado, M. Müller and F. Comellas

Sci. Rep. **2**, 444 (2012) (ArXiv 1112.2079)
Sci. Rep. **3**, 2773 (2013) (Arxiv 1303.3891)
Eur. Phys. J. Plus 129: 150 (2014) (Arxiv 1409.3793)

QUITEMAD
Quantum Information Technologies in Madrid

QCC, Nottingham University, 21 / January / 2015

A collaboration of

Quantum

Complex

G. Davide Paparo

Markus Müller

Miguel Angel
Martin-Delgado

Francesc Comellas

QUITEMAD
Quantum Information Technologies in Madrid

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

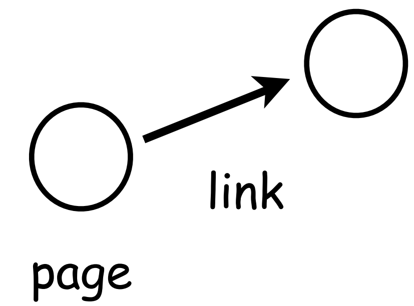
Outline

- Google's PageRank
- Quantum PageRank, a route for Quantum Networks
- Results on tree networks and more general ones
- Quantum PageRank for large scale Complex Networks
- Features on scale-free, hierarchical and Erdös-Renyi
- General properties of the algorithm (localization, scaling behavior, stability)

Classical ranking problem

The internet in 1998: the -mostly static- WWW: how to search?

"Searching" in the WWW:
database search vs objective ranking



database search

link analysis

Problems: Scalability and Objectivity of the search results

Marchiori's idea (among others): the relevance of a page is given by the relation to the web!
Look at the hyperlink structure!

Brin and Page (Google's founders) base PageRank on this idea

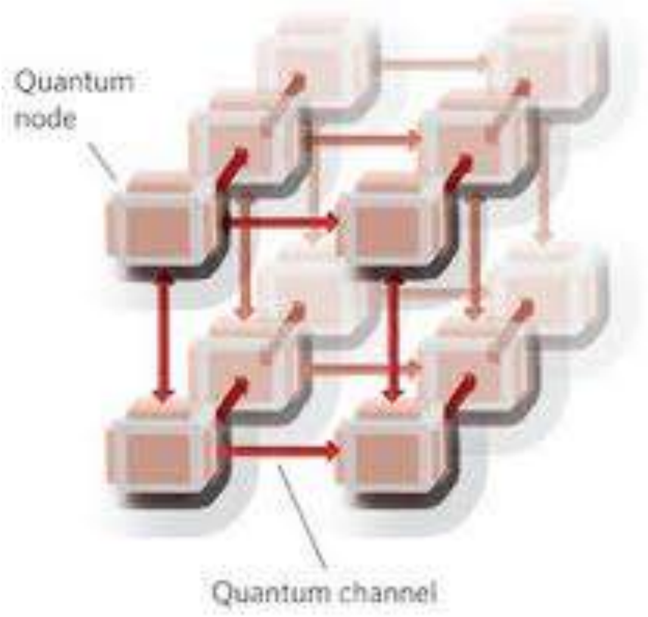
In doing this a webpage is reduced to a number: a ranking!
An audacious idea!

Massimo Marchiori, Proceedings of the Sixth International World Wide Web Conference (WWW6), 1997.

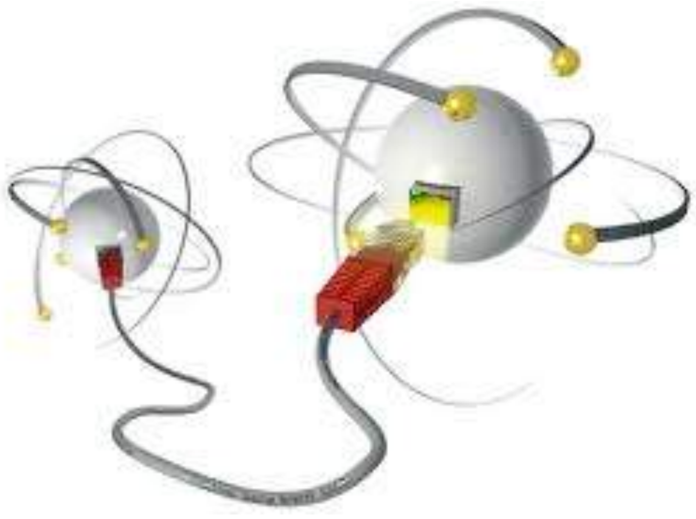
The scenario

New Quantum technologies:

•Quantum Networks: storing information on quantum degrees of freedom.



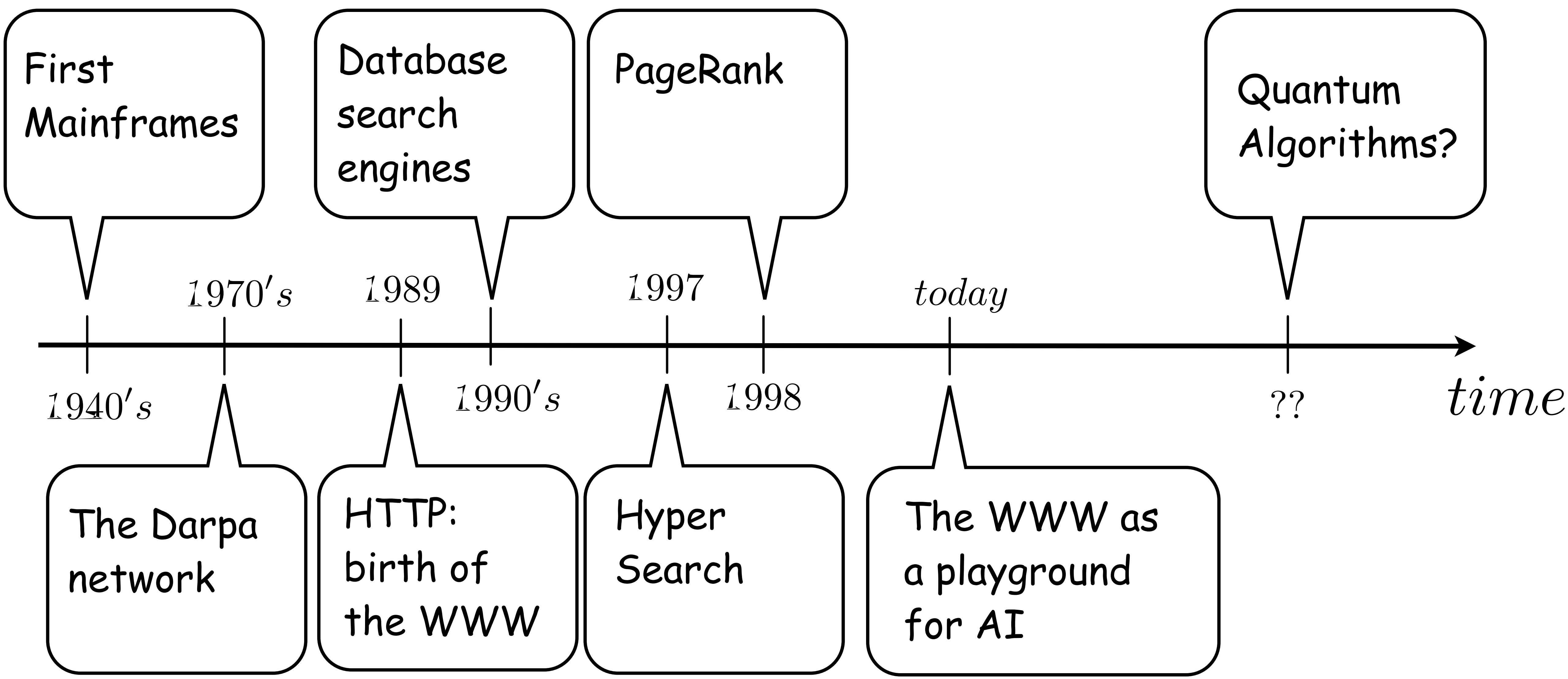
Allows for provably secure quantum communication.



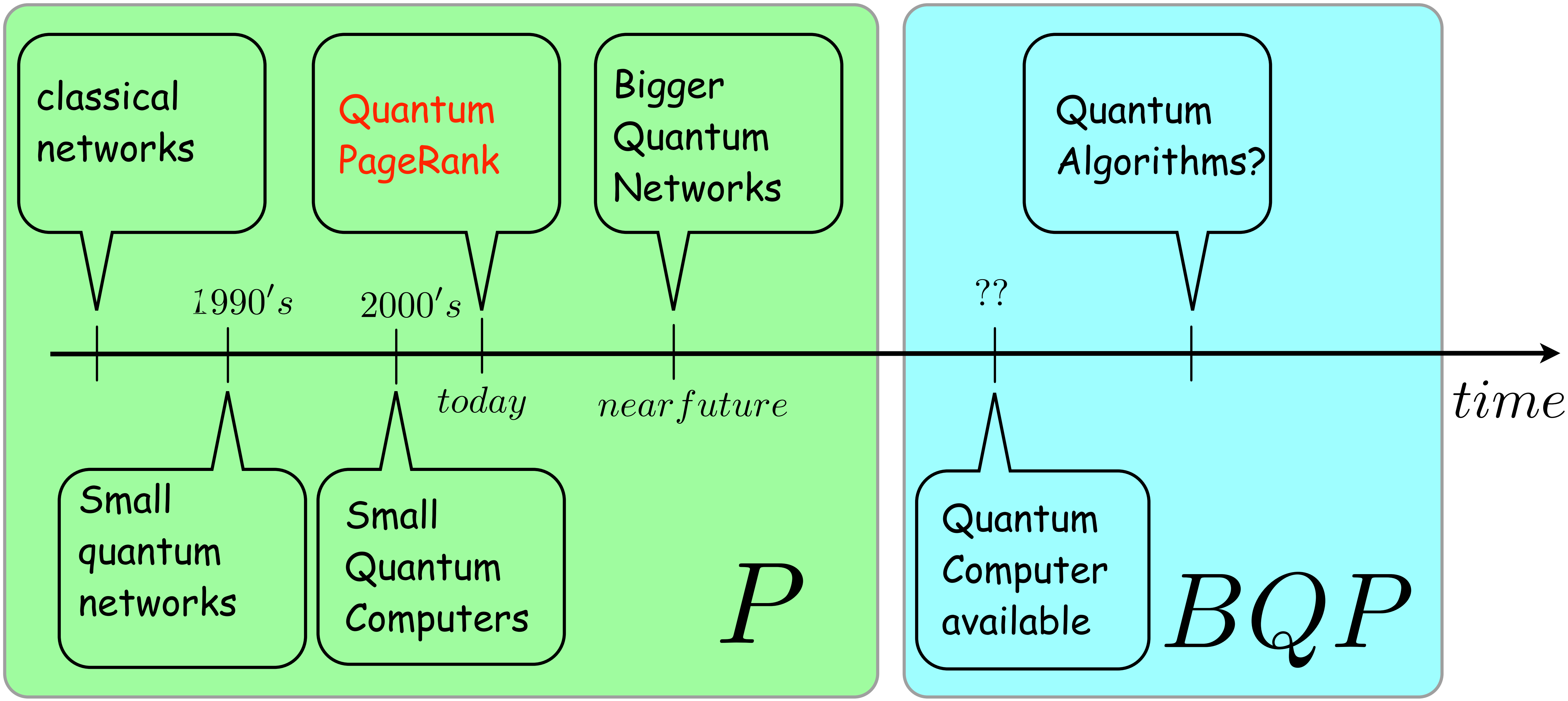
Nowadays: Small (~10 nodes) Quantum Networks have been built.

•Quantum Internet: a large scale quantum network: the quantum follow up of the WWW.

The Timeline



The Timeline



Searching in a Classical Web

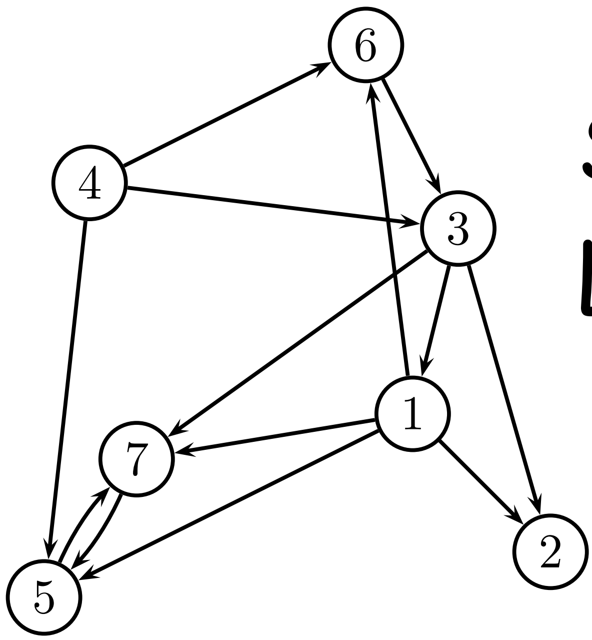
Task: search for "a word" on the internet

what is behind a search engine?

Two step process:

1. Output all pages containing "a word".
2. Rank the most important/relevant first .

Google's PageRank /1



S. Brin and L. Page's idea:
Look into the hyperlink structure!

PageRank's Key Idea:

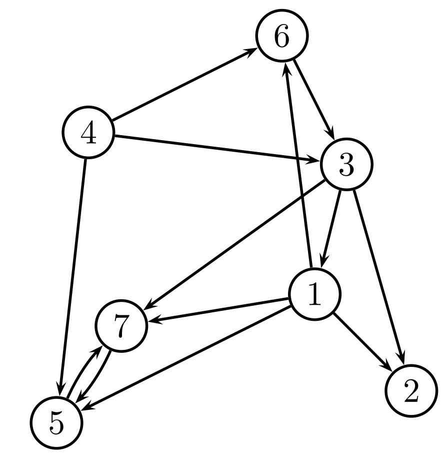
A node's importance is given by the pages that link to it.
The more important these pages are the better.
The fewer the outgoing links they have the better.

$$I(P_i) := \sum_{j \in B_i} \frac{I(P_j)}{\mathrm{outdeg}(\mathrm{P_j})}$$

S. Brin, L. Page: Computer Networks and ISDN Systems 33 (1998) 107-17.

Google's PageRank /2

$$I(P_i) := \sum_{j \in B_i} \frac{I(P_j)}{\mathrm{outdeg(P_j)}}$$



$$H_{ij} := \begin{cases} 1/\mathrm{outdeg(P_j)} & \text{if } P_j \in B_i \\ 0 & \text{otherwise} \end{cases}$$

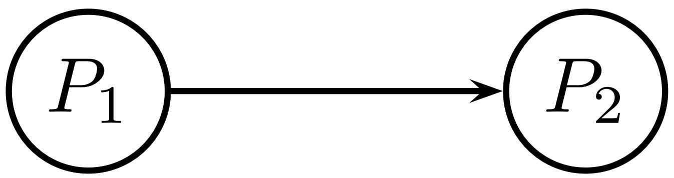
Hyperlink matrix

Computing PageRank is equivalent to: $I = HI$

Solving it iteratively: the "Power Method": $I^{k+1} = HI^k$

But... not that easy... some tinkering is needed...

Problem and Patch/1

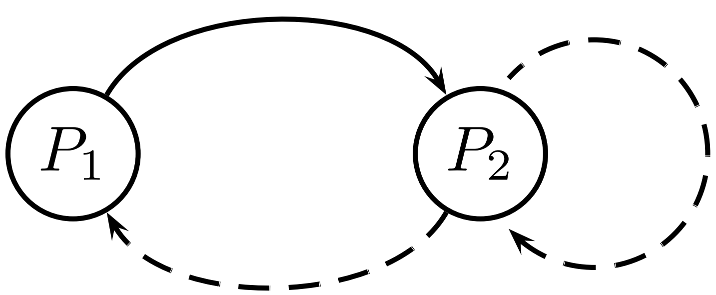


Hyperlink matrix: $H = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$

Easy guess: 2 is more important than 1

feed $I_0 = (1,0)^t \longrightarrow I^{k+1} = H I^k \longrightarrow I = (0,0)^t$?!?!

Patch: add artificial links to "dangling nodes"

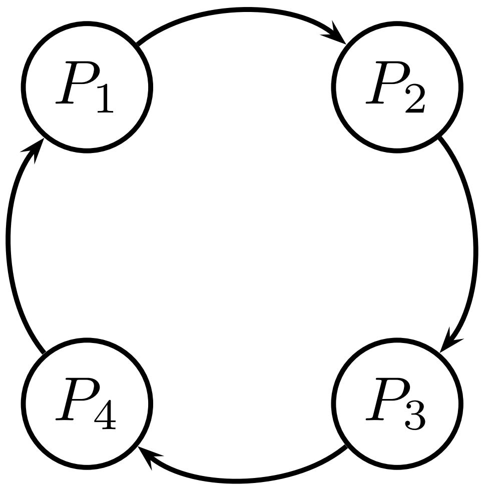


with matrix: $E = \begin{pmatrix} 0 & 1/2 \\ 1 & 1/2 \end{pmatrix}$

Now: $I = (1/3,\ 2/3)^t$ Sound!

We have a stochastic matrix! (columns sum to 1)

Problem and Patch/2



Hyperlink matrix: $E = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

feed $I_0 = (1, 0, 0, 0)^t \longrightarrow I^k = EI^{k-1} \rightarrow$ no convergence ?!?!

Hint from theory: convergence ensured by second eigenvalue smaller than 1

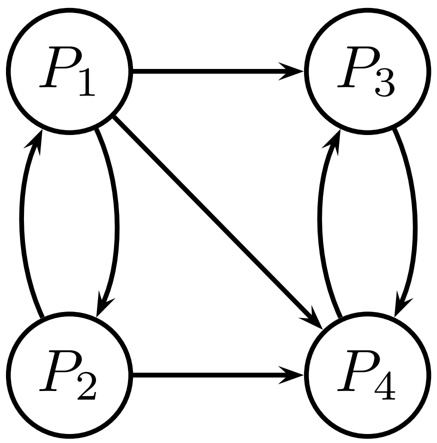
An m exists s.t. E^m with positive entries $\longleftrightarrow$ (Primitivity)

Insight: Importance = probability to find the walker

Interpretation: after m steps any node is reachable, wherever the walker starts

Patch: require primitivity

Problem and Patch/3



Hyperlink matrix: $E = \begin{pmatrix} 0 & 1/2 & \mathbf{0} & \mathbf{0} \\ 1/3 & 0 & \mathbf{0} & \mathbf{0} \\ 1/3 & 1/2 & 0 & 1 \\ 1/3 & 0 & 1 & 0 \end{pmatrix}$

$I_0 = (1, 0, 0, 0)^t \longrightarrow I^k = EI^{k-1} \rightarrow$ page 1 and 2 with ZERO importance ?!?!

No links from subgraph (3,4) to subgraph (1,2)

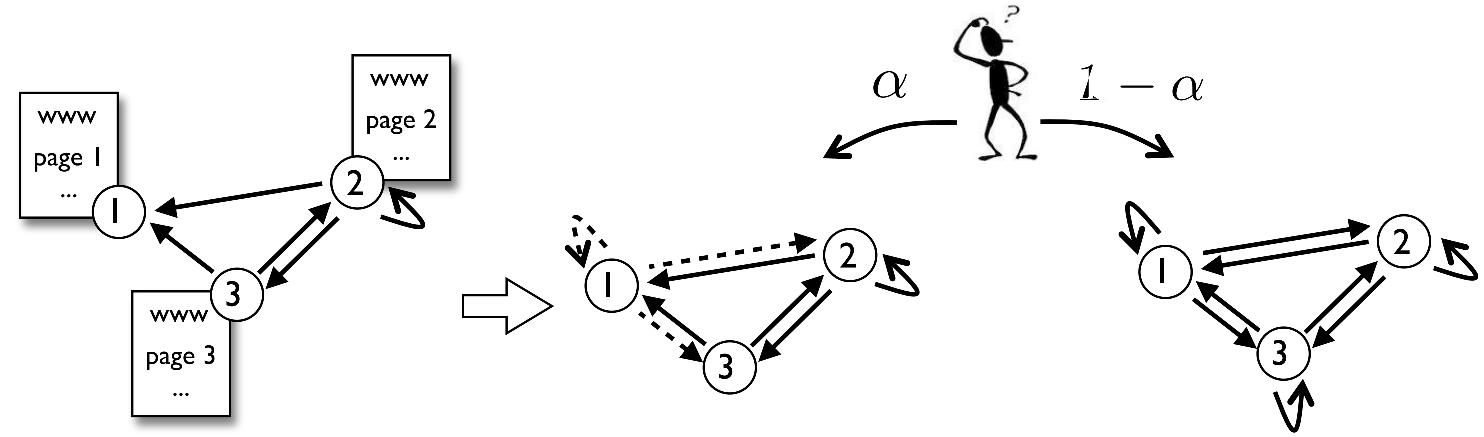
"Drain" of importance from subgraph (1,2)

Reason: E is reducible $\longleftrightarrow$ Graph not "strongly connected"

Patch: require Irreducibility

Google's PageRank

Solution: follow the web for a fraction α of time and jump anywhere else for a fraction $1-\alpha$ of time!



i.e. $$G := \alpha E + \frac{(1-\alpha)}{N} 1$$

- Stochastic
- Primitive
- Irreducible

Now: "Power method":

- Converges
- to the unique stationary vector
- Not dependent on initial condition

Tune the parameter $\alpha \longrightarrow \alpha = 0.85$

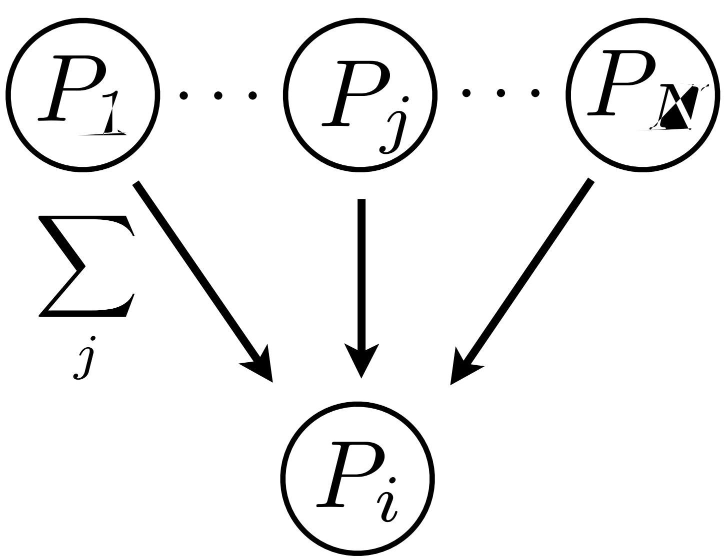
Random walks

PageRank has a random walk at its heart

from $I^k = G I^{k-1}$ **with** $G := \alpha E + \frac{(1-\alpha)}{N}\mathbf{1}$

I^k **is a vector of probabilities of finding the walker**

Google's matrix G is a transition matrix:



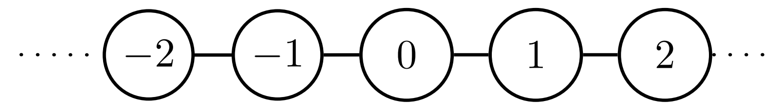
$$\Pr(X^k = P_i) = \sum_j G_{ij} \Pr(X^{(k-1)} = P_j)$$

$$G_{ij} = \Pr(X^k = P_i | X^{(k-1)} = P_j)$$

Markov Chain or Random Walk! How about quantizing?

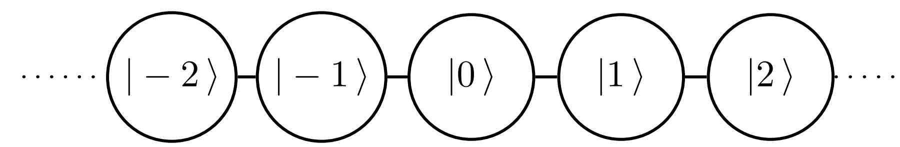
Quantum walks

a walk on a line:



with prob. p go right,
with 1-p go left

Naive quantization:



evolution: $U = \sqrt{p}\,|i+1\rangle\langle i| + \sqrt{1-p}\,|i-1\rangle\langle i|$

But... start from $|0\rangle$ or $|2\rangle$ after a time step BOTH nonzero amplitude on $|1\rangle$

So $\langle 0|2\rangle = 0$ but $\langle 0|U^{\dagger}U|2\rangle \neq 0$ No unitarity!?

Solutions:

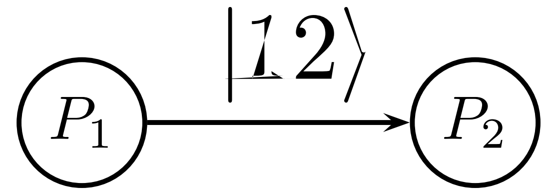
enlarge Hilbert space (coin space)

Scattering Quantum Walk

Szegedy's Quantum Walk

Szegedy's Quantum Walks/I

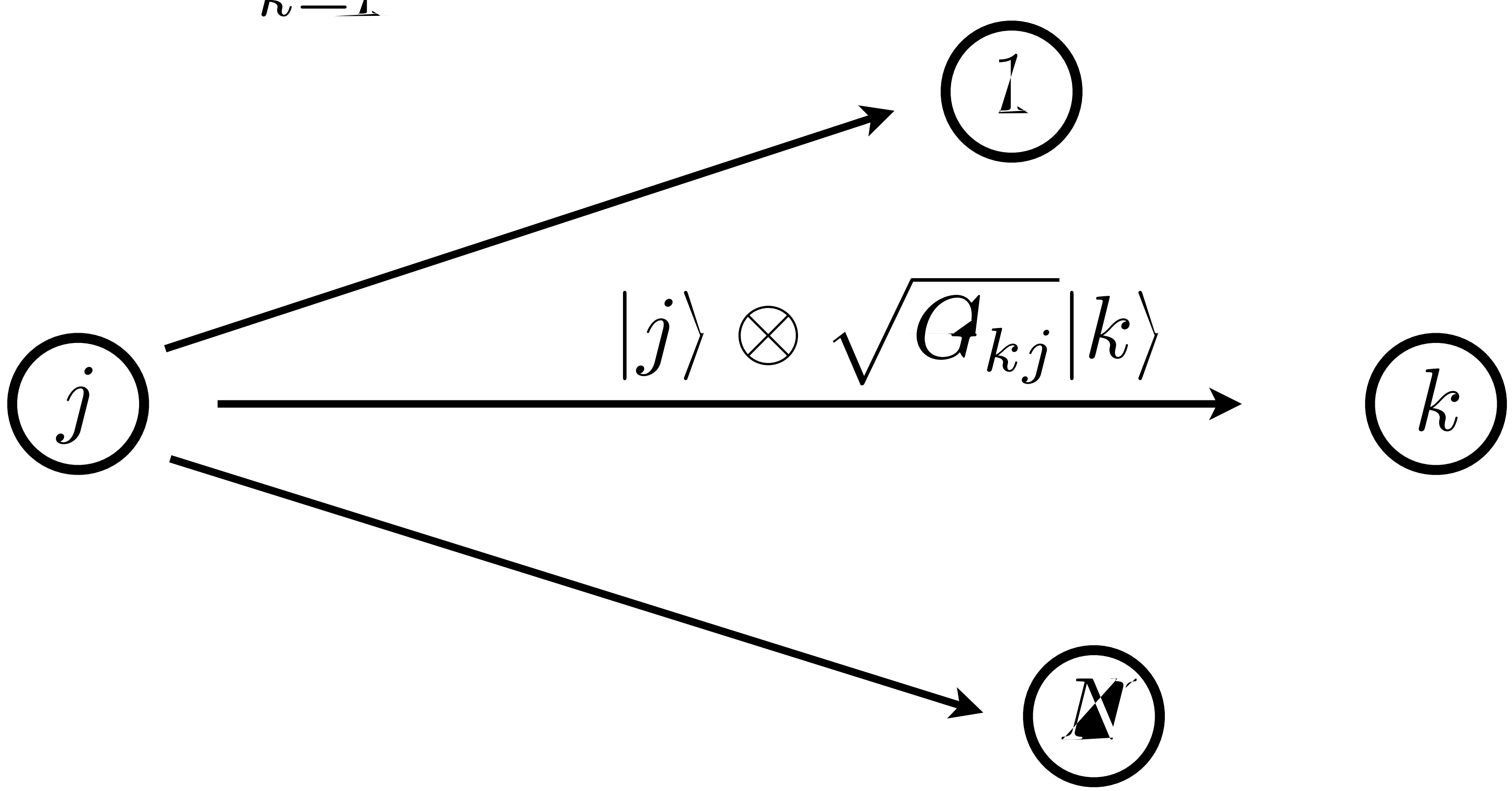
states are links:



States containing info on outgoing links of j:

$$|\psi_j\rangle := |j\rangle_1 \otimes \sum_{k=1}^{N} \sqrt{G_{kj}}\,|k\rangle_2$$

G: Markov Chain

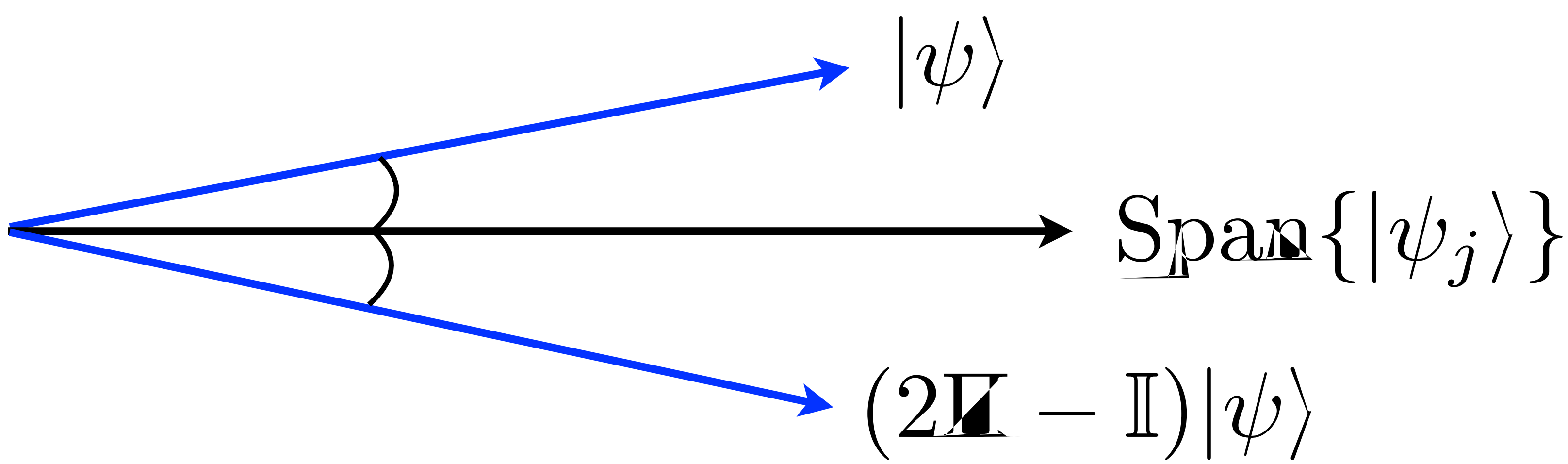


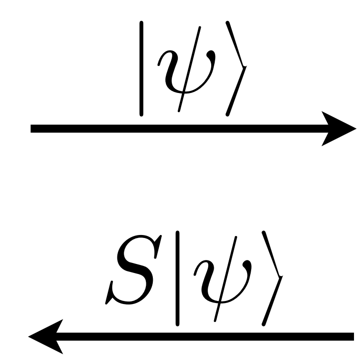
M. Szegedy: Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (2004), pp. 32-41.

Szegedy's Quantum Walks/II

evolution is a reflection around these states and a swap!

$$2\Pi - \mathbb{1} = \sum_{j=1}^{N}\left(2\,|\psi_j\rangle\langle\psi_j| - \frac{1}{N}\mathbb{1}\right) \qquad S = \sum_{j,k=1}^{N} |j,k\rangle\langle k,j|$$





$U = S(2\Pi - \mathbb{1})$ Unitary and directedness preserving

We'll consider two-step evolution operators: U^2

M. Szegedy: Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (2004), pp. 32-41.

Quantum PageRank/1

Key idea: use quantization of Markov Chain to obtain a Quantum PageRank algorithm

Use Google Matrix: $G := \alpha E + \frac{(1-\alpha)}{N}1$

$$|\psi_j\rangle := |j\rangle_1 \otimes \sum_{k=1}^{N} \sqrt{G_{kj}}\,|k\rangle_2$$

Idea: the (instantaneous) Quantum PageRank of a node is the probability to find a quantum Walker that has evolved under a Quantized Markov Chain.

$$I_q(P_i, m) = \langle\psi(0)|\, U^{\dagger\, 2m} |i\rangle_2 \langle i|\, U^{2m} |\psi(0)\rangle.$$

Unitarity suggests it will vary in time

- Average
- Error (variance)

Quantum PageRank/2

Algorithm to calculate the Quantum PageRank:

1. Write down the Google Matrix: G

2. Start from the state: $|\psi_0\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} |\psi_j\rangle$

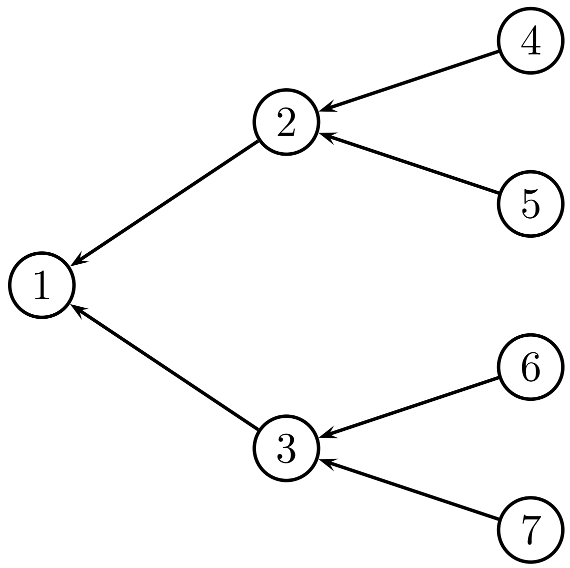
3. Let it evolve according to a Szegedy Walk: $U^2|\psi_0\rangle$

4. Calculate the instantaneous Quantum PageRank

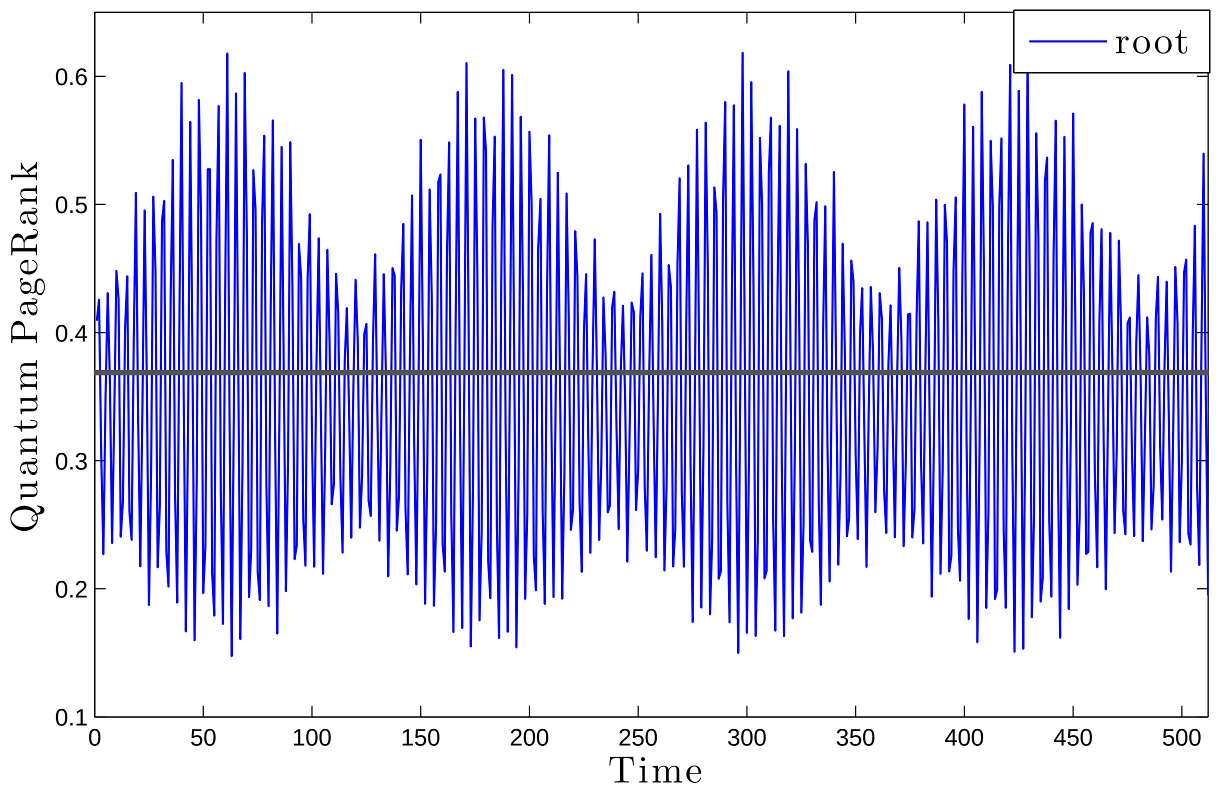
$$I_q(P_i, m) = \langle\psi(0)|U^{\dagger 2m}|i\rangle_2\langle i|U^{2m}|\psi(0)\rangle.$$

5. Calculate the time averaged Quantum PageRank and its variance

Quantum PageRank/Tree

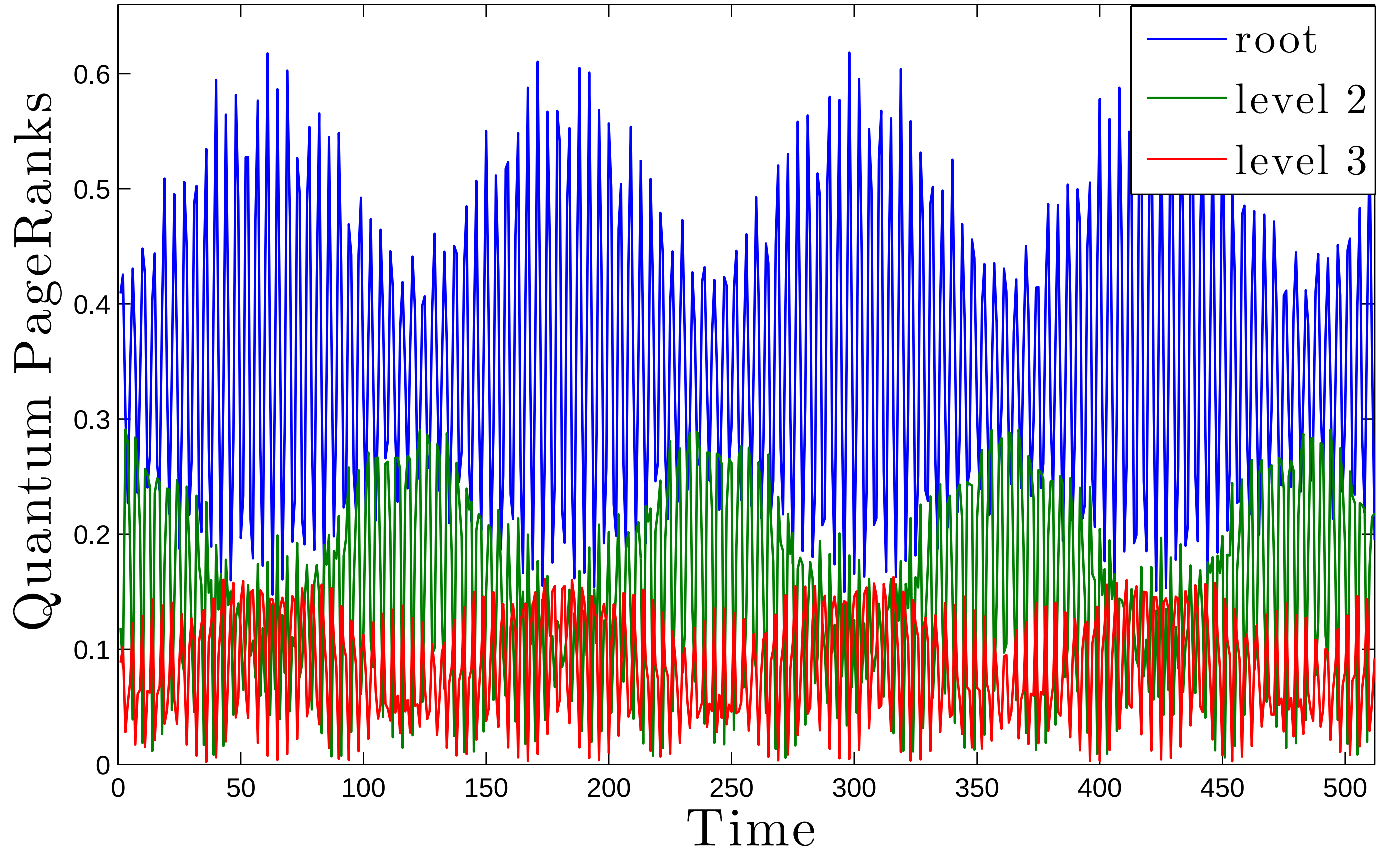


Root:

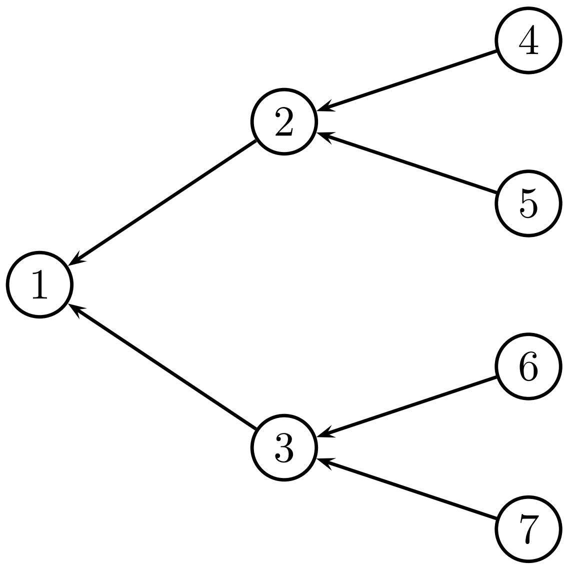


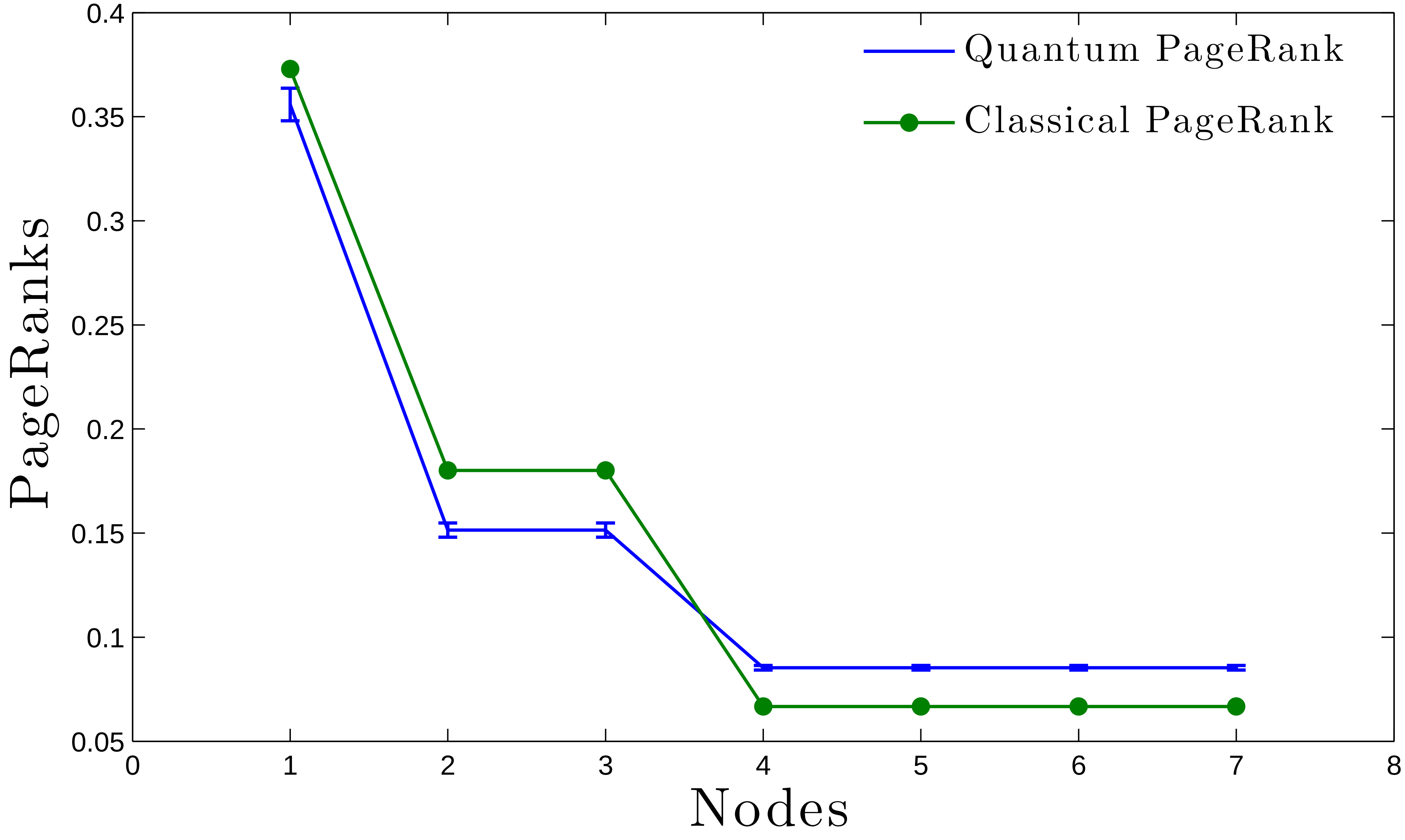
1. The root's Q-PageRank is higher: (Instantaneous) Outperformance!

2. Hierarchy is not preserved at any given time!



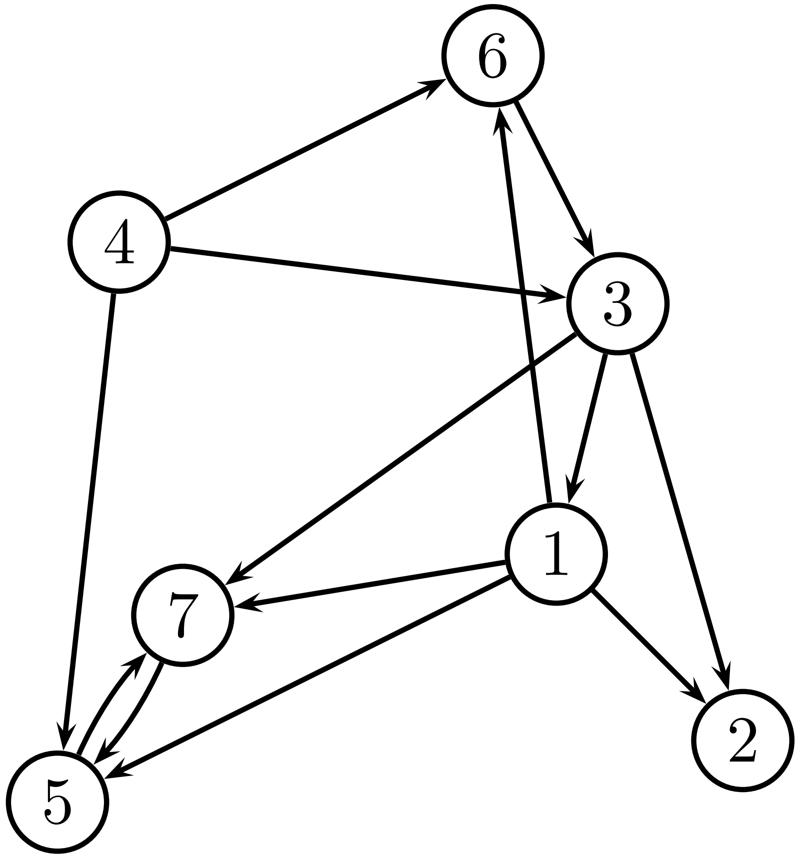
Quantum PageRank/Tree 2

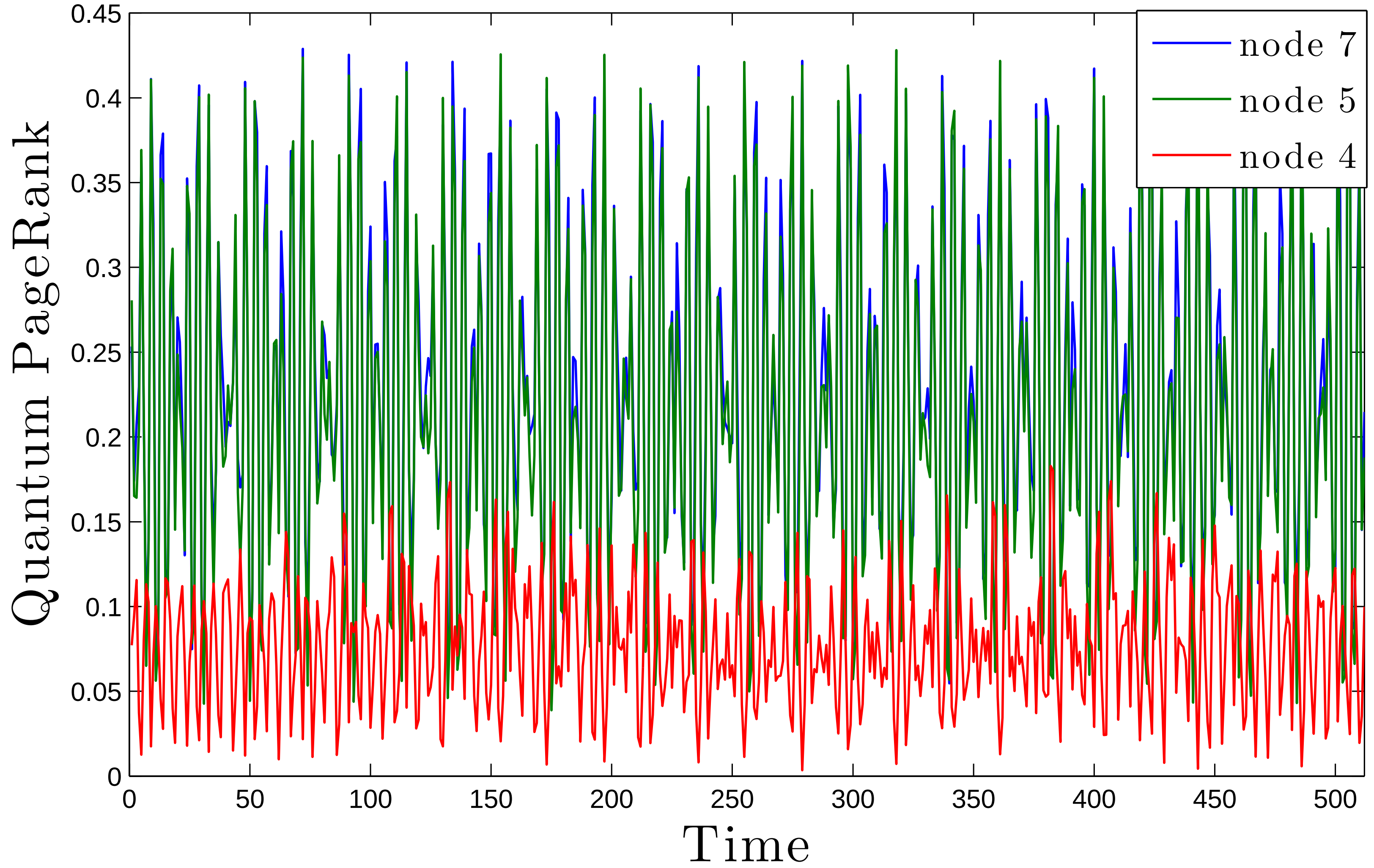




Strong Hierarchical preserving on average!

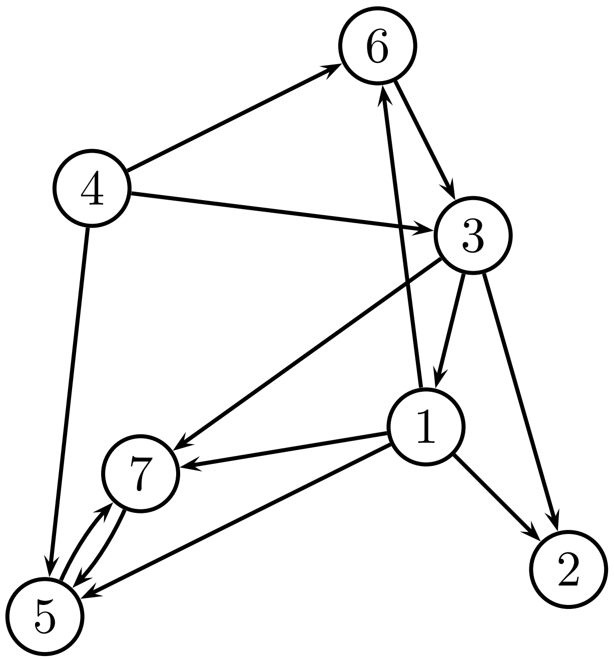
Quantum PageRank/a Graph

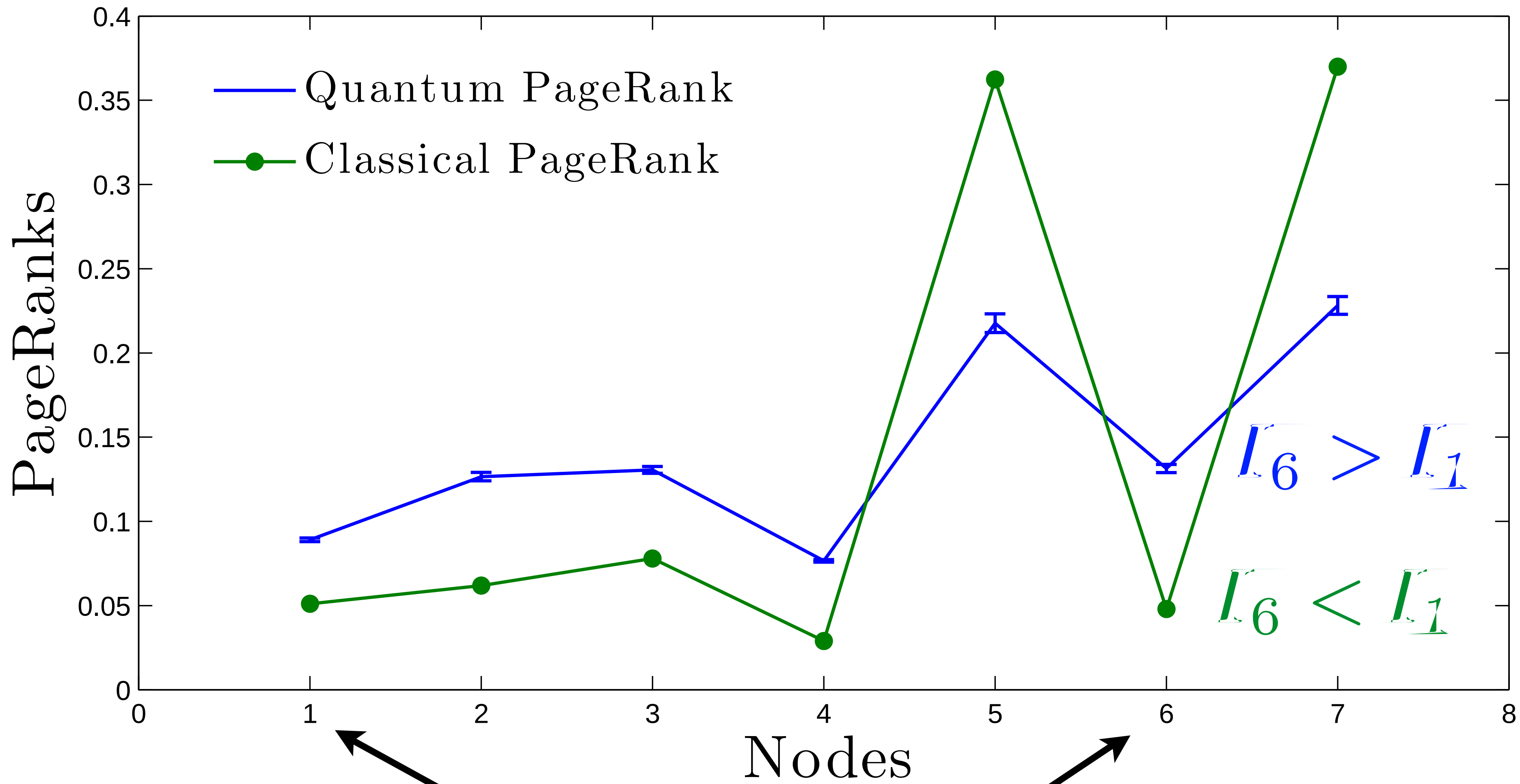




Hierarchy is not preserved at any given time!

Quantum PageRank/a Graph





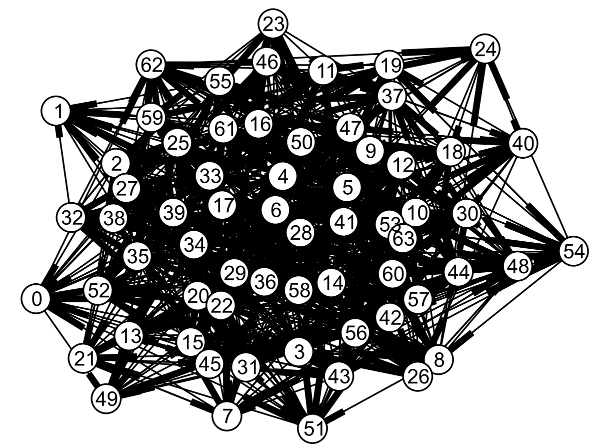
Swap!

Hierarchy is not preserved even on average!

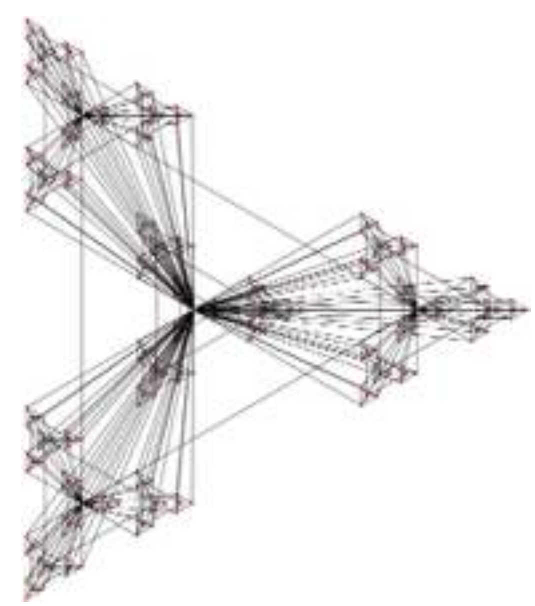
More homogenous importance distribution

Q.PageRank:Bigger Networks

Do these effects persist on bigger networks? ... studies for up to 512 nodes

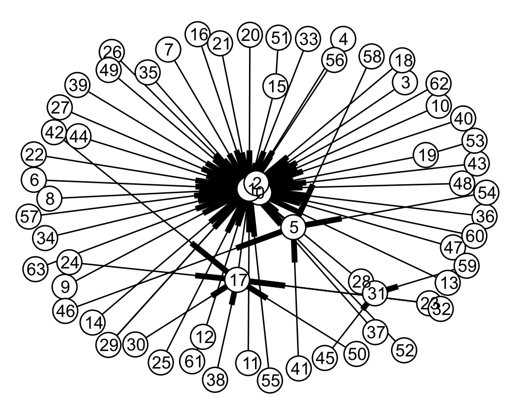


1. Erdös Rényi graphs



2. Hierarchical graphs

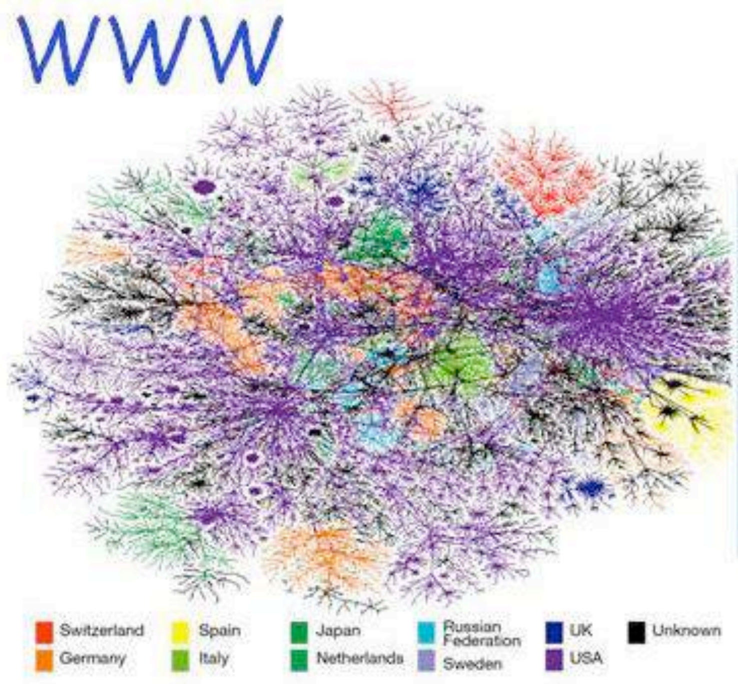
3. Scale free graphs



GDP, M. Müller, F. Comellas, M.Á. Martin-Delgado, Sci. Rep. **3**, 2773 (2013)

Scale-Free Networks

Model of the WWW



Model of the Air routes network

scale-free degree distribution: $P(k) \approx k^{-\gamma}$ $2 \leq \gamma \leq 3$

How to grow a scale-free: Barabási-Albert model with preferential attachment

Quantum networks will likely grow with this topology and will reuse the existing communication networks with the underlying classical communication channels

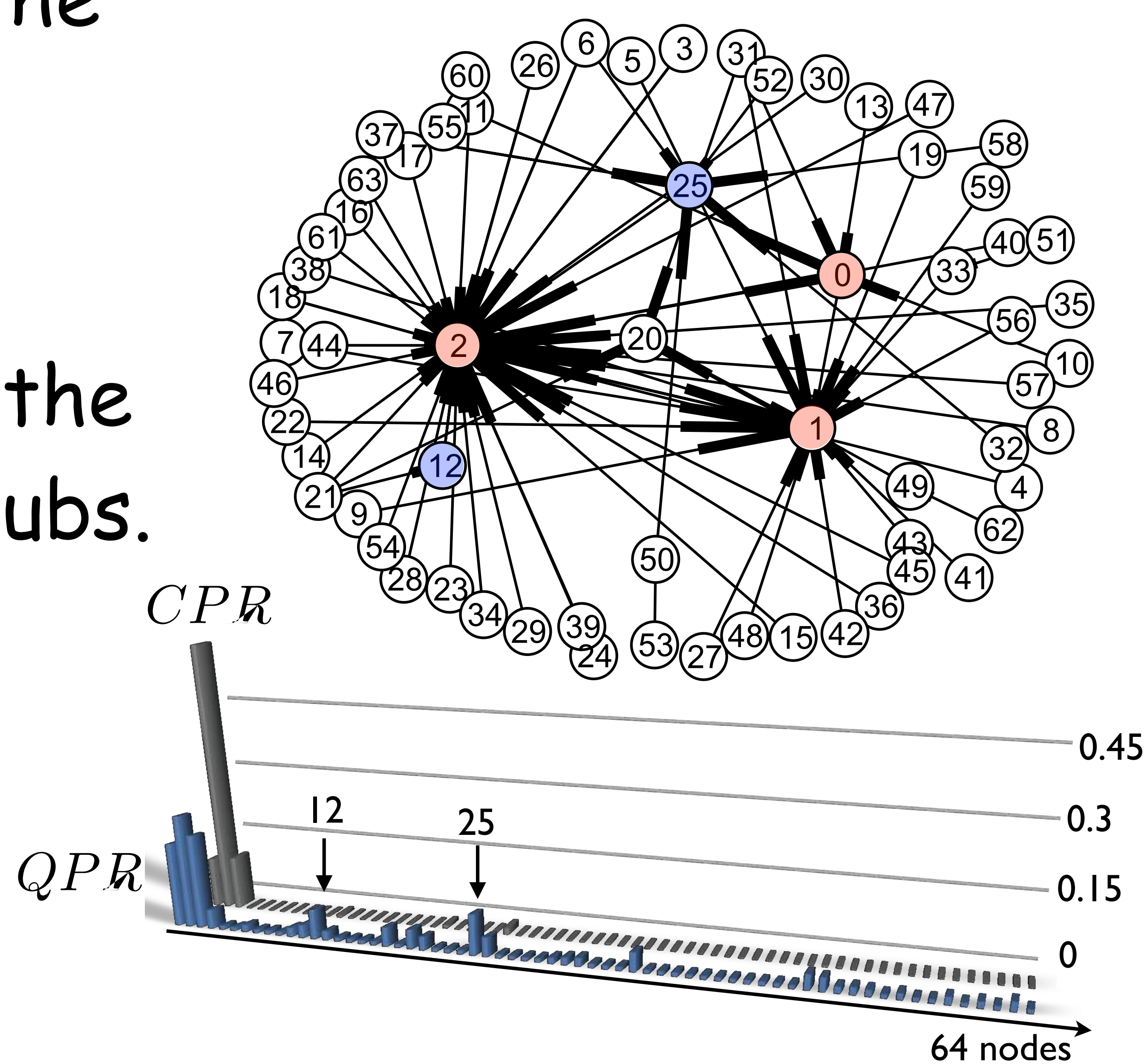
A.-L. Barabási, R. Albert, and H. Jeong, Physica A:Statistical Mechanics and its Applications 281, 69 (2000); A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani, PNAS 101, 3747 (2004); A.-L. Barabási and R. Albert, Science 286, 509 (1999).

Scale-Free Networks

Results of the Quantum PageRank

Highlights the main hubs.

Increased visibility of the secondary hubs.



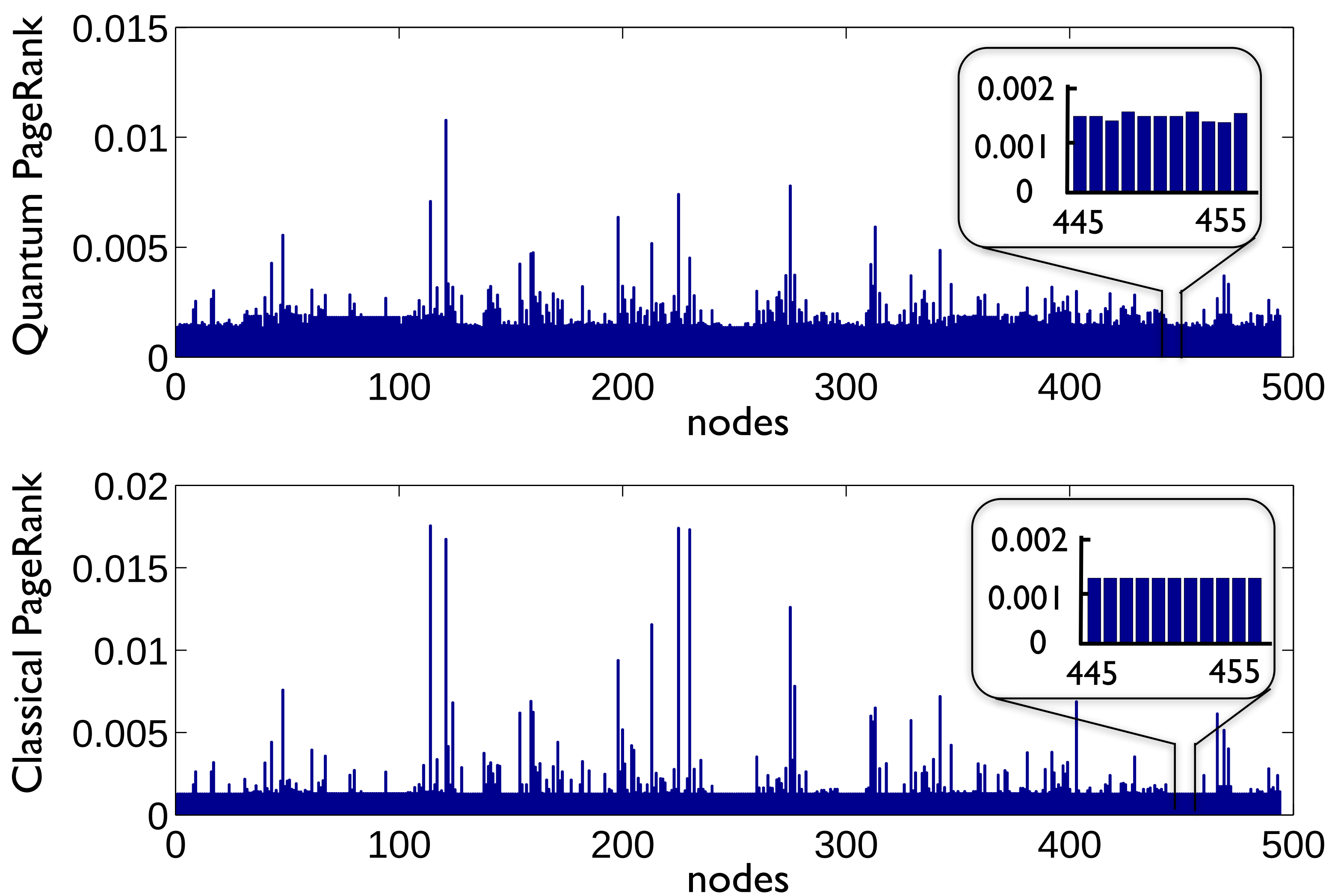
Observations

- Increased resolution

Scale-Free Networks/2

Results of the Quantum PageRank

Lifts the degeneracy of nodes



Observations

- Increased resolution
- Degeneracy partially lifted

Scale-Free Networks only?

What about the other known topologies?

Like Hierarchical graphs?

Observations

- Increased resolution
- Degeneracy partially lifted

Questions

- Dependence on the topology

Observations

- Increased resolution
- Degeneracy partially lifted

Questions

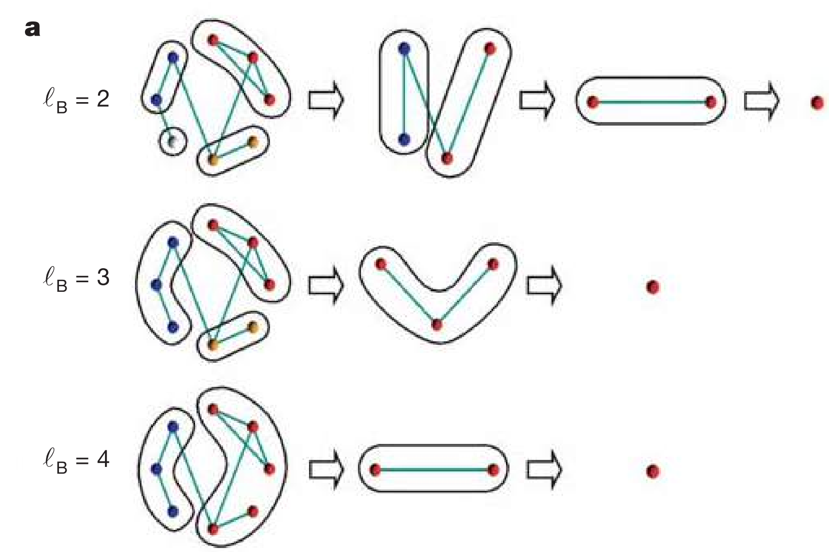
- Dependence on the topology
- Smoothness of ranking
- Stability with respect to the damping parameter
- Localization properties

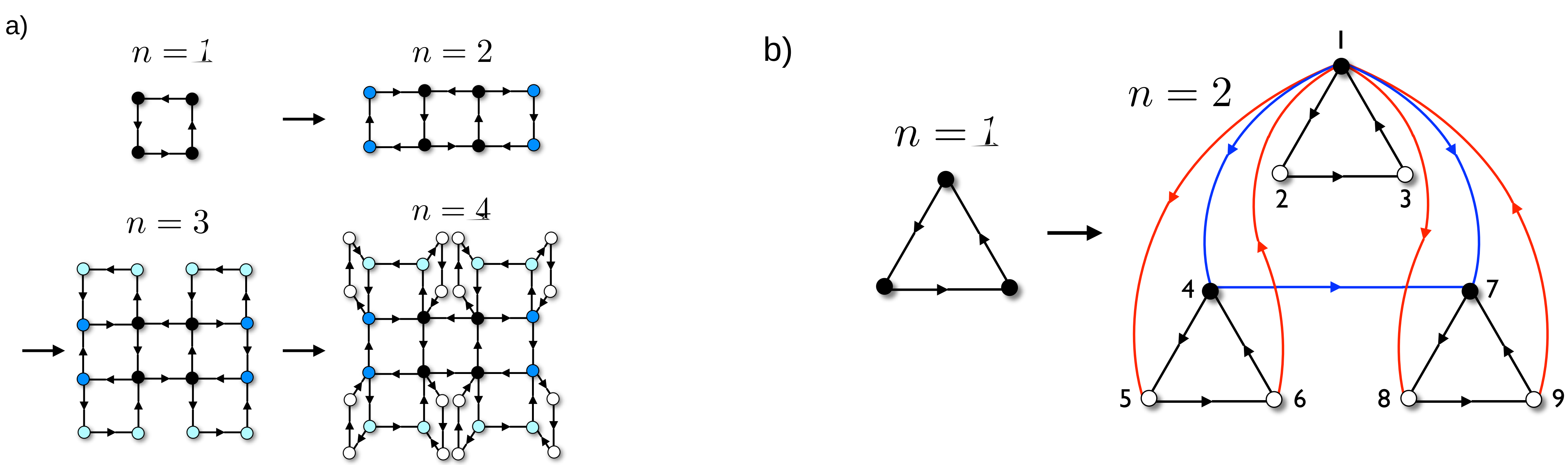
Hierarchical Networks

Networks composed of self repeating moduli

WWW is self-similar

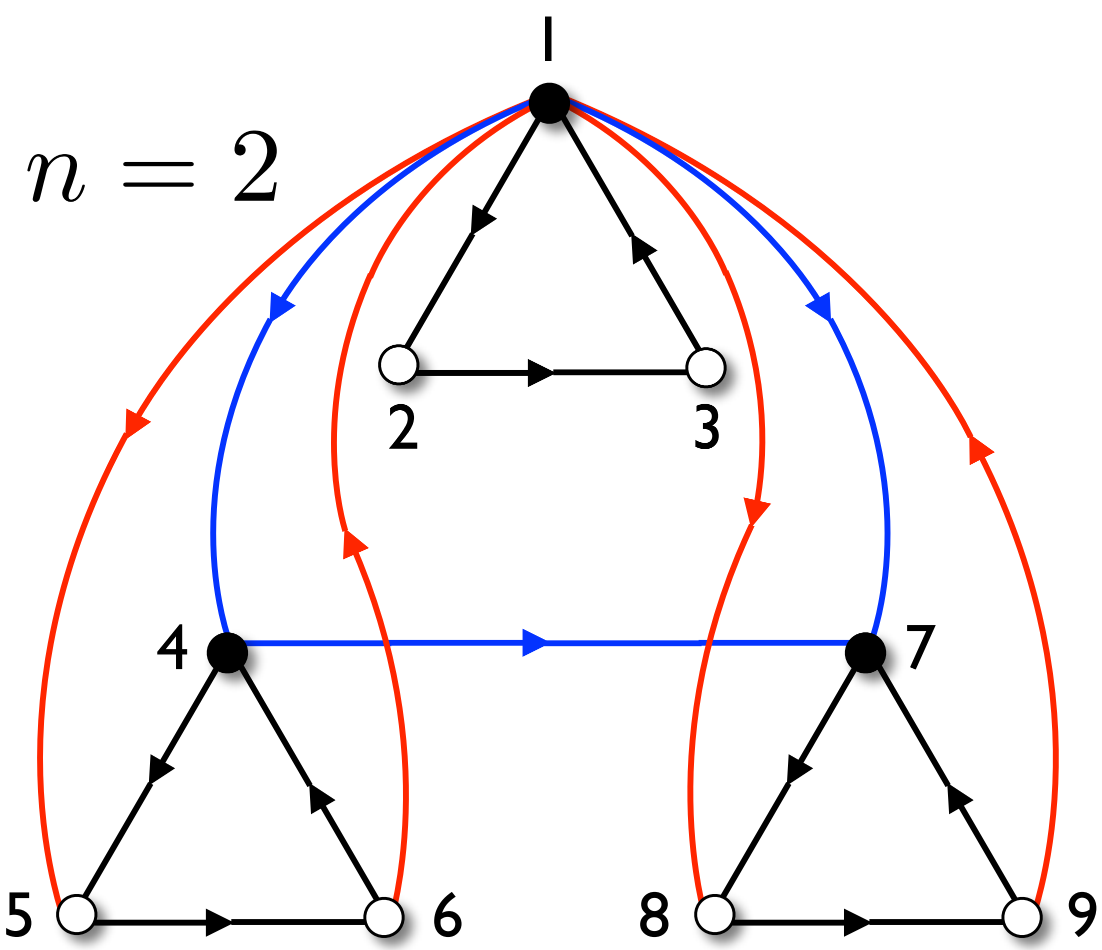
Hierarchical networks are often amenable of exact treatment

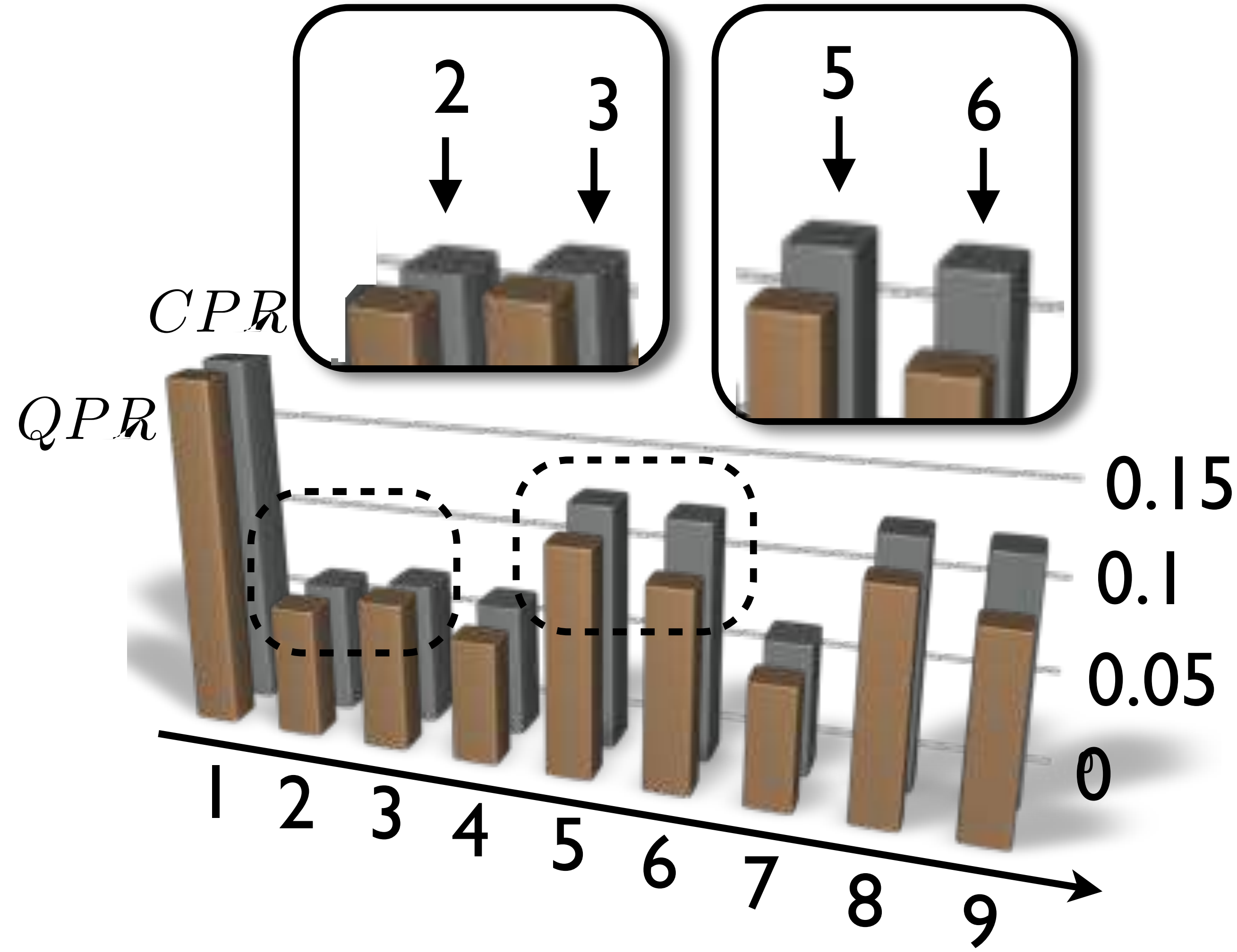




C. Song, S. Havlin, and H. A. Makse, Nature 433, 392 (2005); F. Comellas and A. Miralles, Physica A: Statistical Mechanics and its Applications 388, 2227 (2009); F. Comellas and A. Miralles, Journal of Physics A: Mathematical and Theoretical 42, 425001 (2009).

Hierarchical networks





Global hierarchy is preserved but local connectivity structure is more visible.

Scaling Behavior

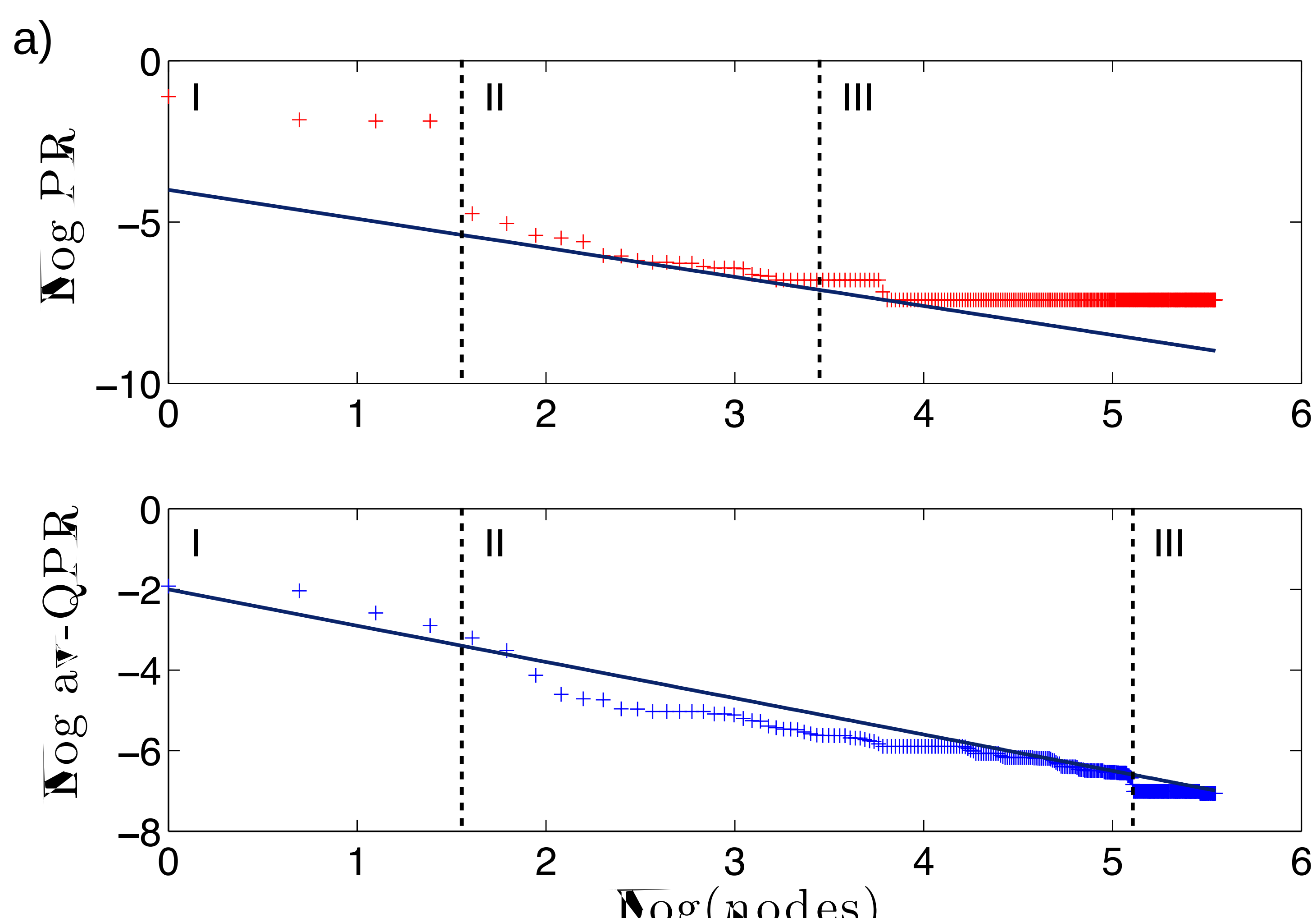
Scaling behavior of the PageRanks: $I_j \sim j^{-\beta}$ $\beta_{cl} \approx 0.9$

How do Quantum PageRanks scale? $\langle I_q(P_j)\rangle \sim j^{-\beta_q}$

Power law! With a smoother scaling $\beta_q = 0.85$

CPR

QPR



3 Zones:

I) Hubs

II) Well fitted

III) Degenerate

Scaling Behavior

A real network from EPA

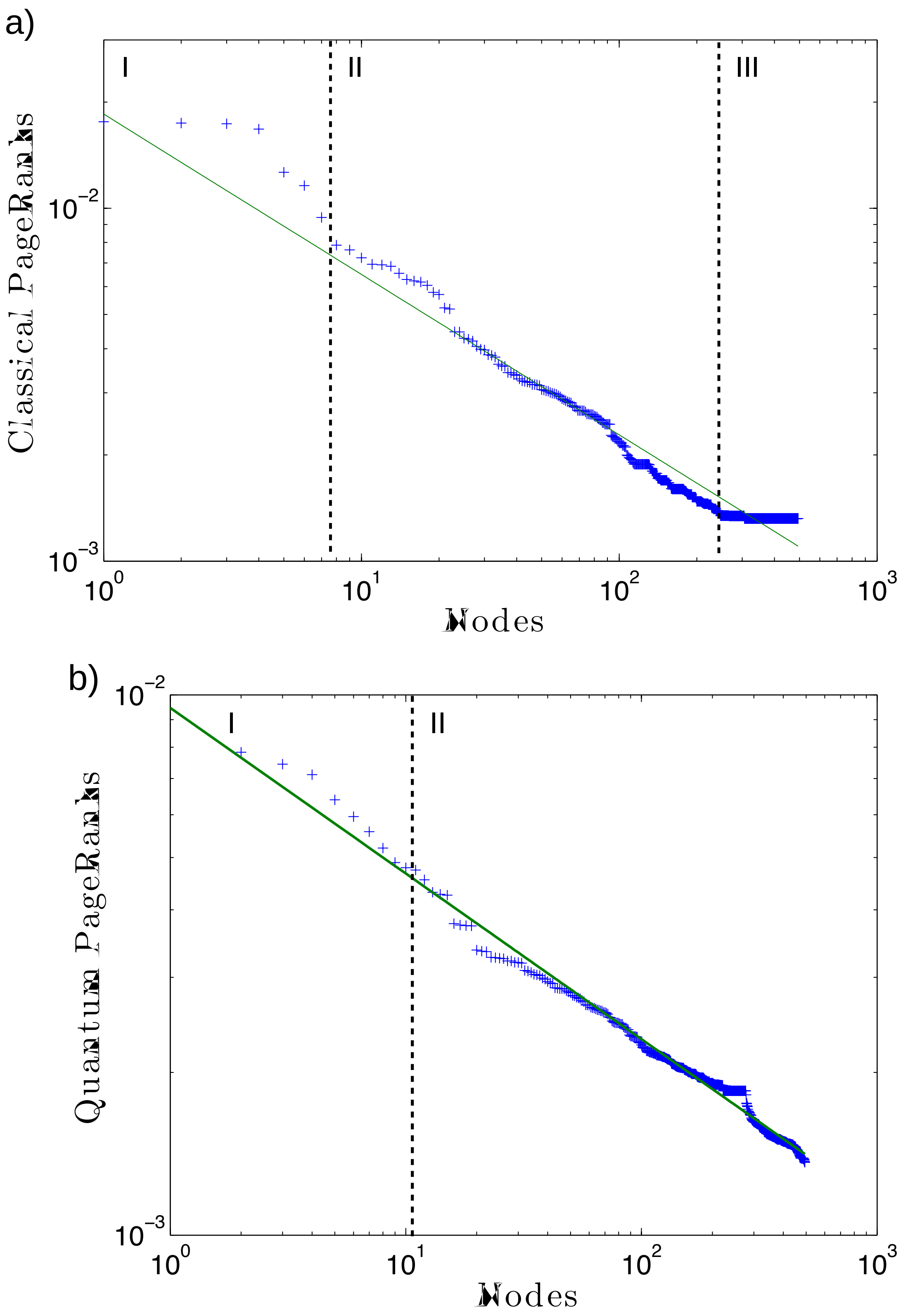
Smoother behavior means better ranking

$$\beta_q = 0.30 \ vs \ \beta_{cl} = 0.45$$

Smoother scaling

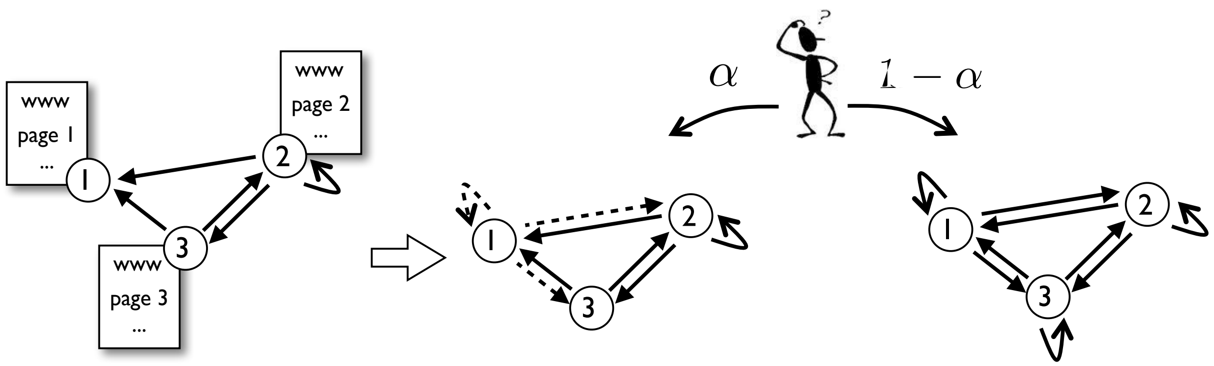
Well fitted in the whole range

Degeneracy resolved



The Problem of Stability

Damping parameter α in the Google matrix is arbitrary



$$G := \alpha E + \frac{(1-\alpha)}{N} 1$$

We need a way to measure the "fidelity"

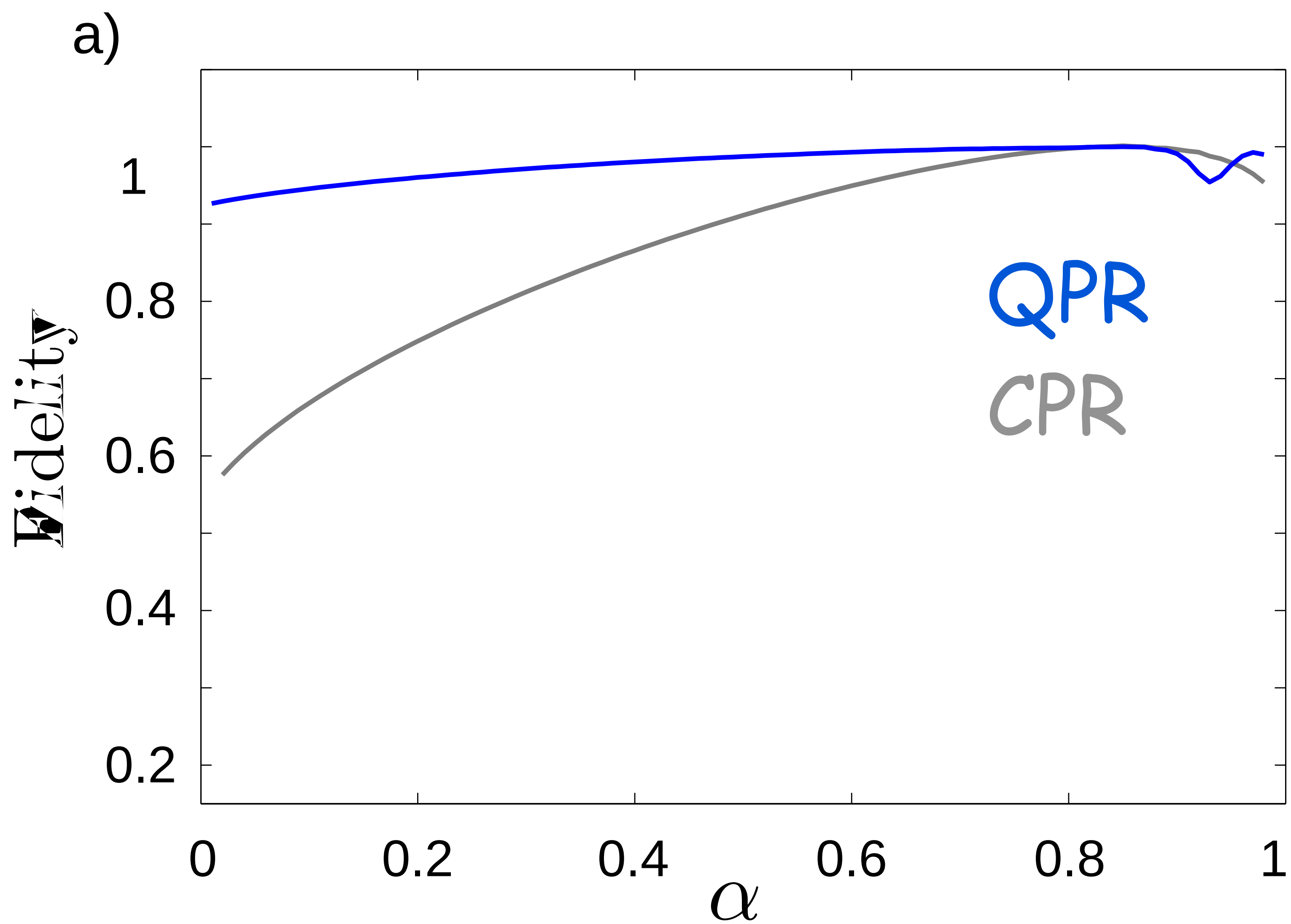
$$f(\alpha, \alpha') = \sum_j \sqrt{\Gamma(P_j, \alpha)\Gamma(P_j, \alpha')}$$

Classically the fidelity can reach zero for values of the damping parameter that are sufficiently far.

B. Georgeot, O. Giraud, and D. L. Shepelyansky, Physical Review E 81, 056109 (2010).

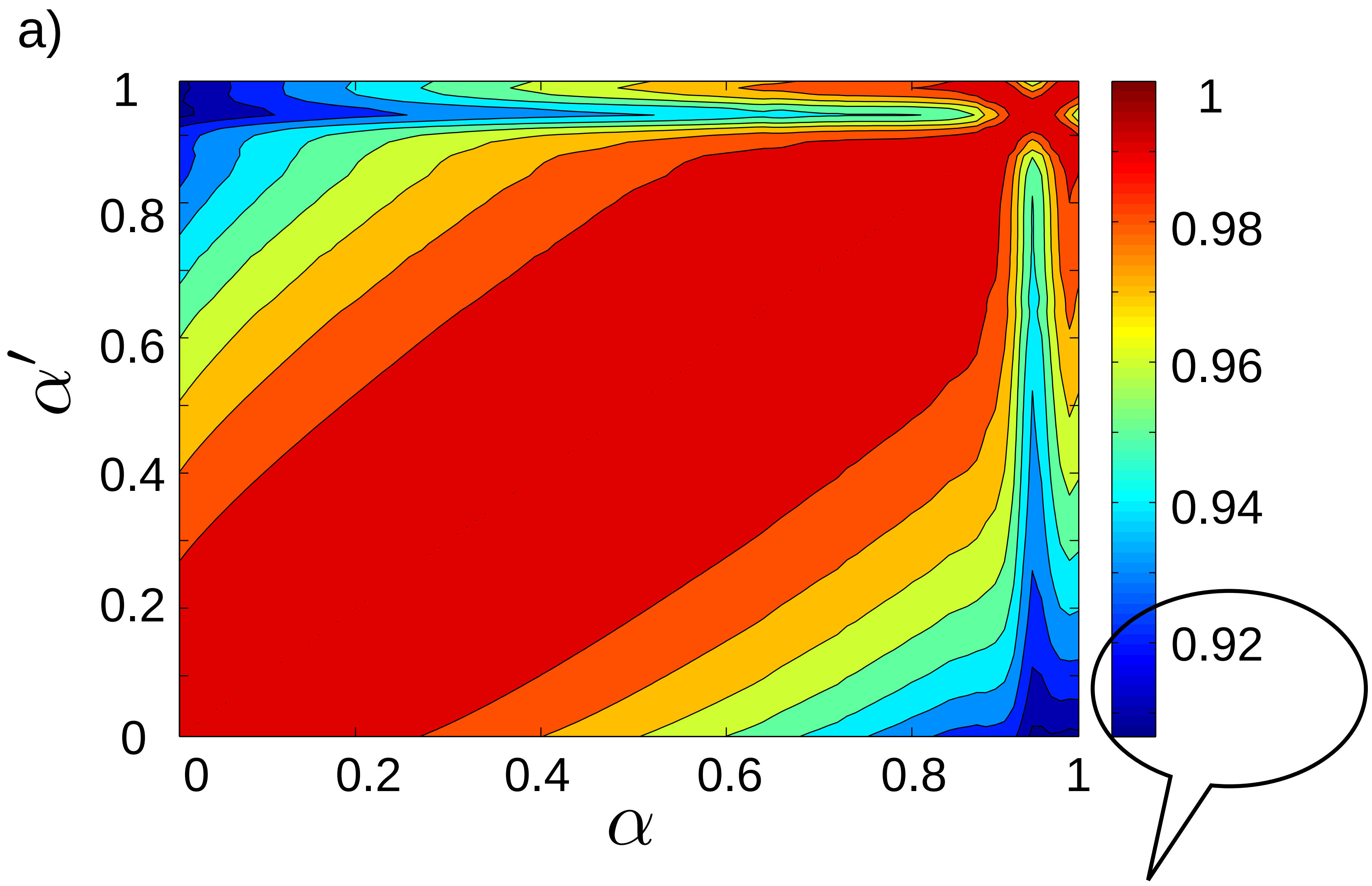
Stability

More robust with respect to the variation of the damping parameter.



$$f(\alpha, 0.85) = \sum_j \sqrt{\Gamma(P_j, \alpha)\Gamma(P_j, 0.85)}$$

Stability



Minimum Fidelity ≥ 0.90

More robust with respect to the variation of the damping parameter.

Conclusions

- Found a valid quantization of Google's PageRank
- Quantum PageRanks show nontrivial features when ranking
- It's able to rank better e.g. in scale-free networks
- When applied on bigger SF Complex Networks the algorithm displays: **localization**, a more **favorable scaling behavior** and it is more **stable**
- All these elements make it a valuable tool to analyze classical networks

Collaborators

GICC (Grupo de Información y Computacíon Cuántica)

Markus Müller
PostDoc

G. Davide Paparo
PhD student

Miguel Angel
Martin-Delgado
(Group leader)

Angel Rivas
PostDoc

Alexandre Dauphin
PhD student

Oscar Viyuela
PhD student

Laura Martin
PhD student

Universidad Complutense de Madrid (UCM)

Francesc Comellas

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

QUITEMAD
Quantum Information Technologies in Madrid

Thank You for your attention!

Sci. Rep. **2**, 444 (2012) (ArXiv: 1112.2079) and
Sci. Rep. **3**, 2773 (2013) (Arxiv 1303.3891)
Eur. Phys. J. Plus 129: 150 (2014) (Arxiv 1409.3793)

PICC: The Physics of Ion Coulomb Crystals

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN

SEVENTH FRAMEWORK PROGRAMME

QUITEMAD

La Suma de Todos
CONSEJERÍA DE EDUCACIÓN
Comunidad de Madrid
www.madrid.org

mi+d www.madrimasd.org

UNION EUROPEA
FONDO SOCIAL EUROPEO

UNIVERSITAS COMPLVTENSIS MATRITENSIS

Santander

Additional Slides

Localization of Walker

Why is studying the localization important?

Localization implies good ranking: scale-free display a little fraction of "hubs"

The method: The Inverse Participation Ratio (IPR)

$$\xi_{cl} := \sum_{i=1}^{N} [\Pr(X = i)]^{2r} \quad vs \quad \xi_q := \sum_{i=1}^{N} \langle I_q(P_i) \rangle^{2r} \qquad \xi := \begin{cases} 1 & \text{if the walker is localized} \\ N^{1-2r} & \text{if the walker is delocalized.} \end{cases}$$

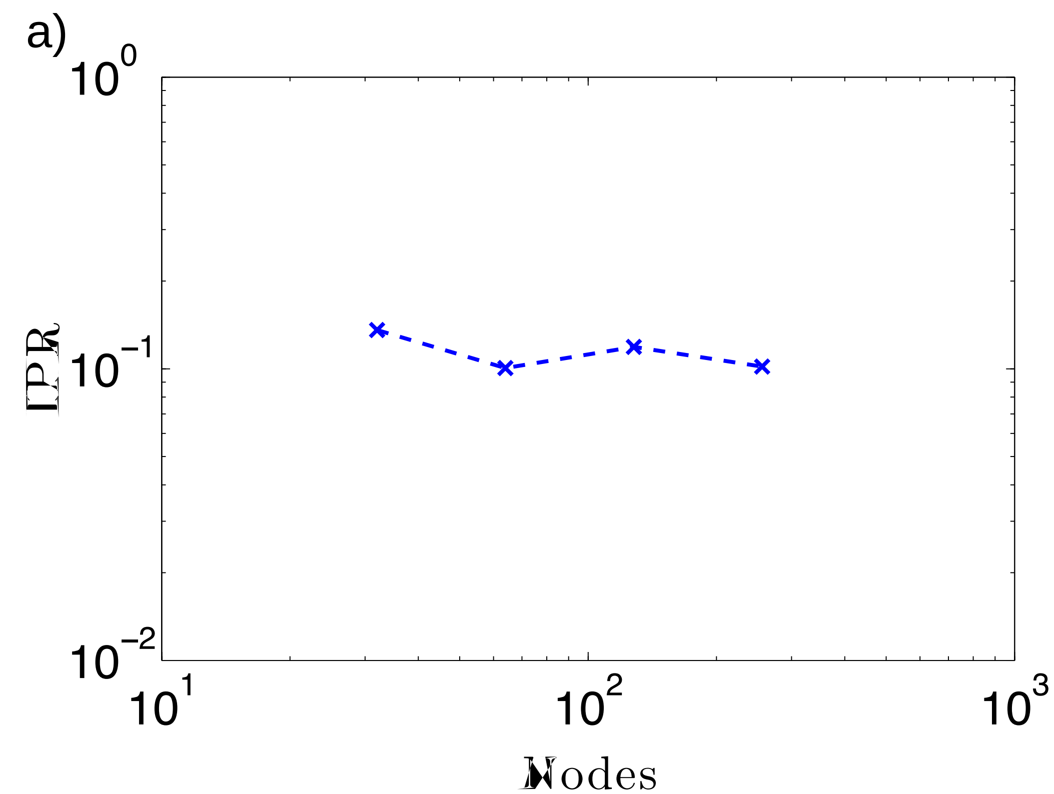
Good witness of localization: we will look at the slope of:

$$\log \xi \ vs \ \log N$$

Localization: Results on SF

Localization found in SF networks classically

and Quantumly!
(slope close to zero)



$$\log \xi \ vs \ \log N$$

A small fraction of nodes will concentrate all the importance: This algorithm can rank well.

Other Approaches / Outlook

- Use Adiabatic Q. Computation to calculate the classical Google PageRank (Silvano Garnerone, P. Zanardi & D. Lidar)
- Use dissipative protocols (Gómez-Gardeñes & Zueco, Silvano Garnerone)
- Multiparticle quantum walks to rank nodes in quantum networks ?

S. Garnerone, P. Zanardi, and D. A. Lidar, Phys. Rev. Lett. 108, 230506 (2012); E. Sánchez-Burillo, J. Duch, J. Gómez-Gardeñes, and D. Zueco, Scientific Reports 2 (2012); S. Garnerone, Phys. Rev. A 86, 032342 (2012).