Modeling the Evolution of Networks as Shrinking Structural Diversity

Jérôme Kunegis

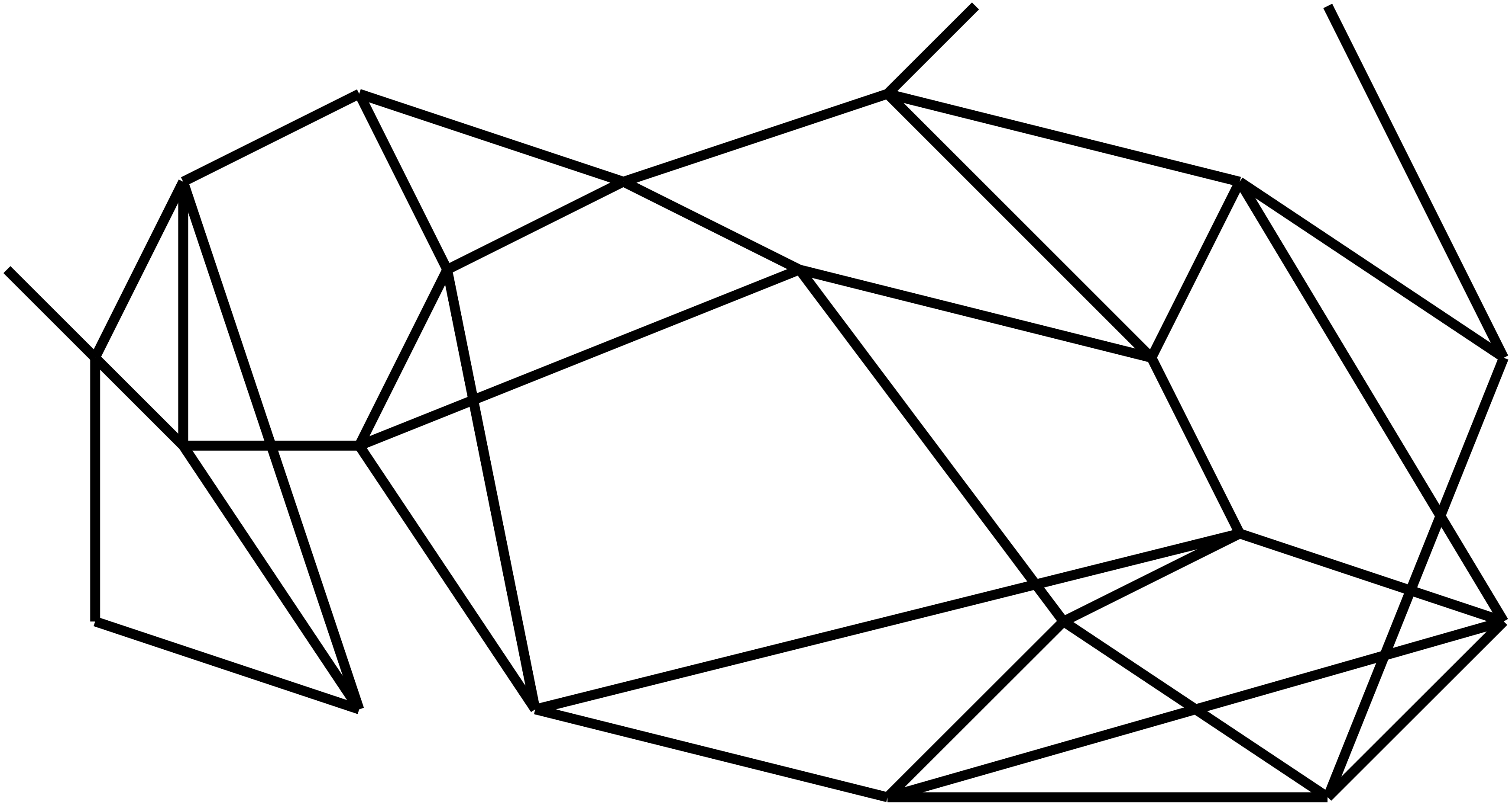
based on work with Damien Fay, Sergej Sizov, Julia Perl, Felix Schwagereit

People and Knowledge Networks

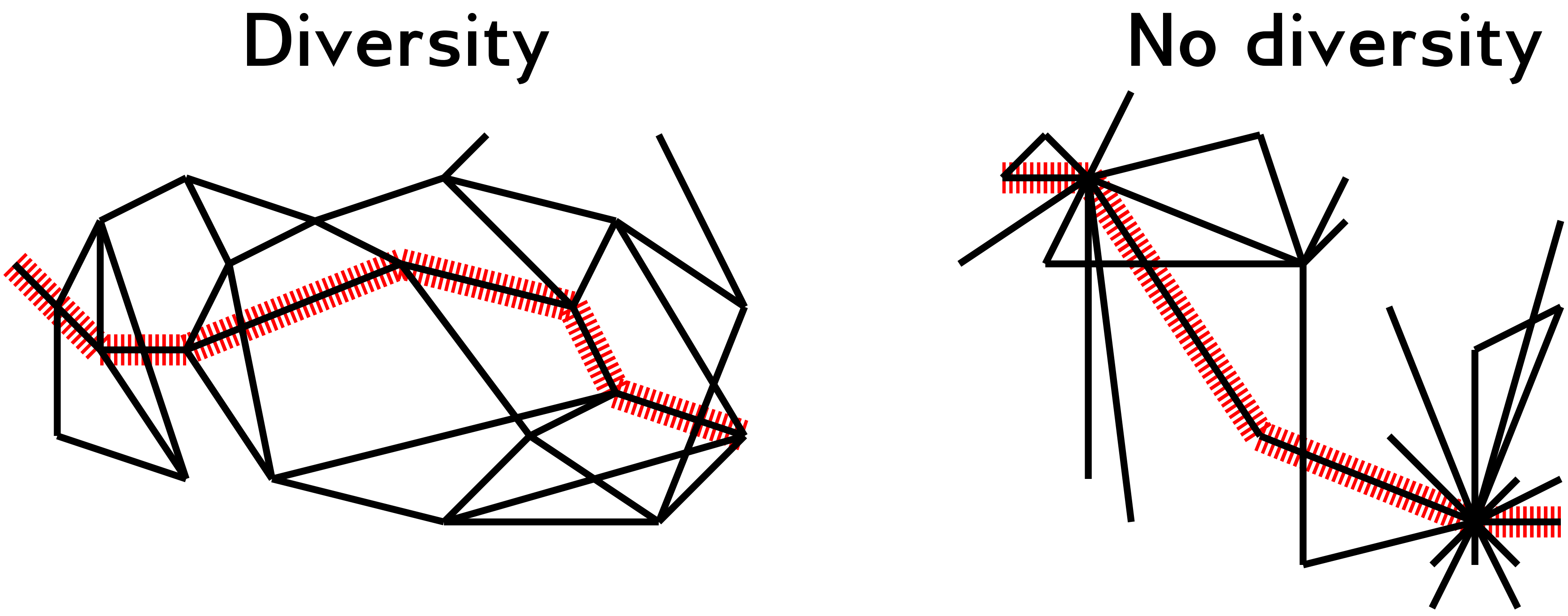
Everyone likes good things:

Or even better: Diversity!

Structural Diversity

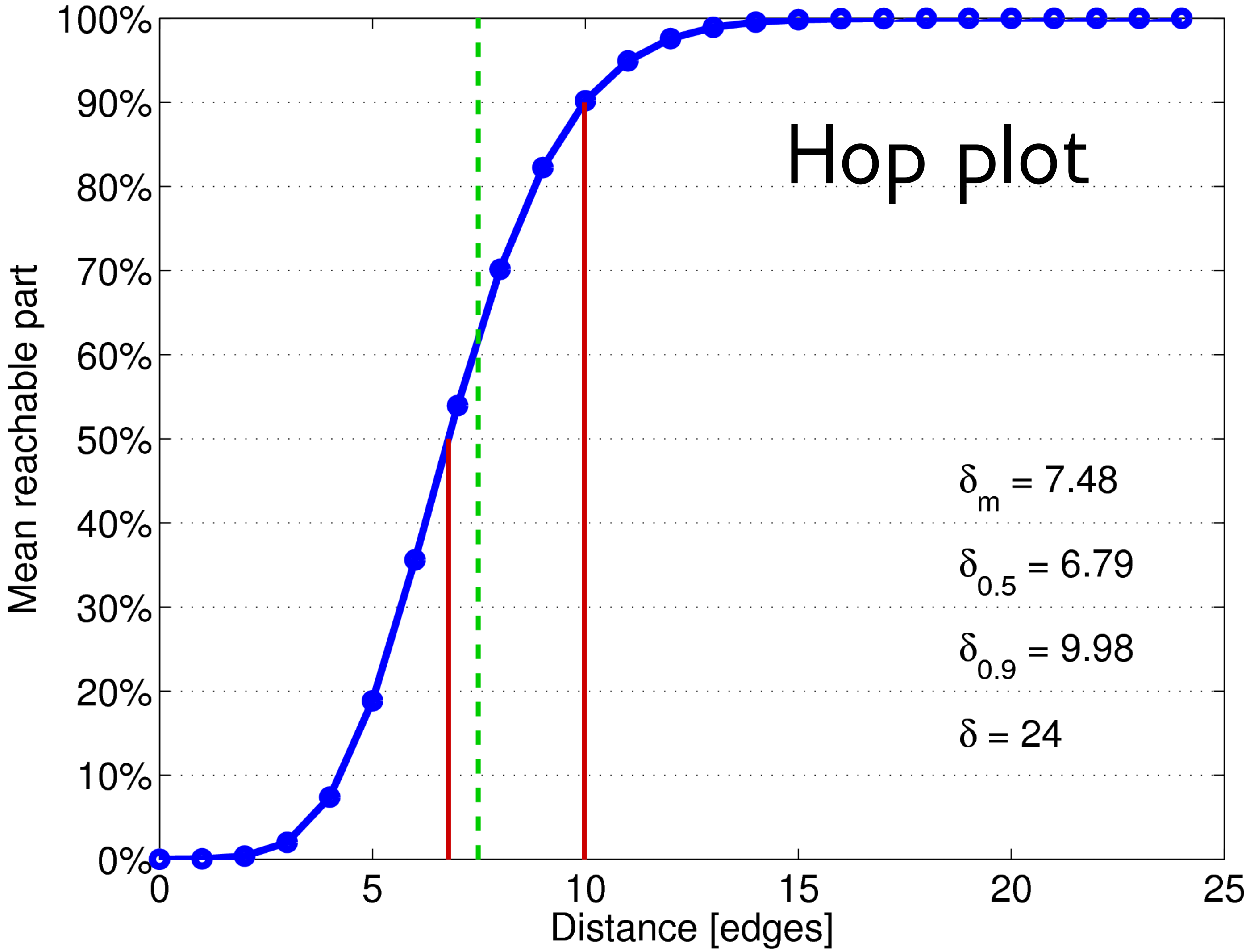


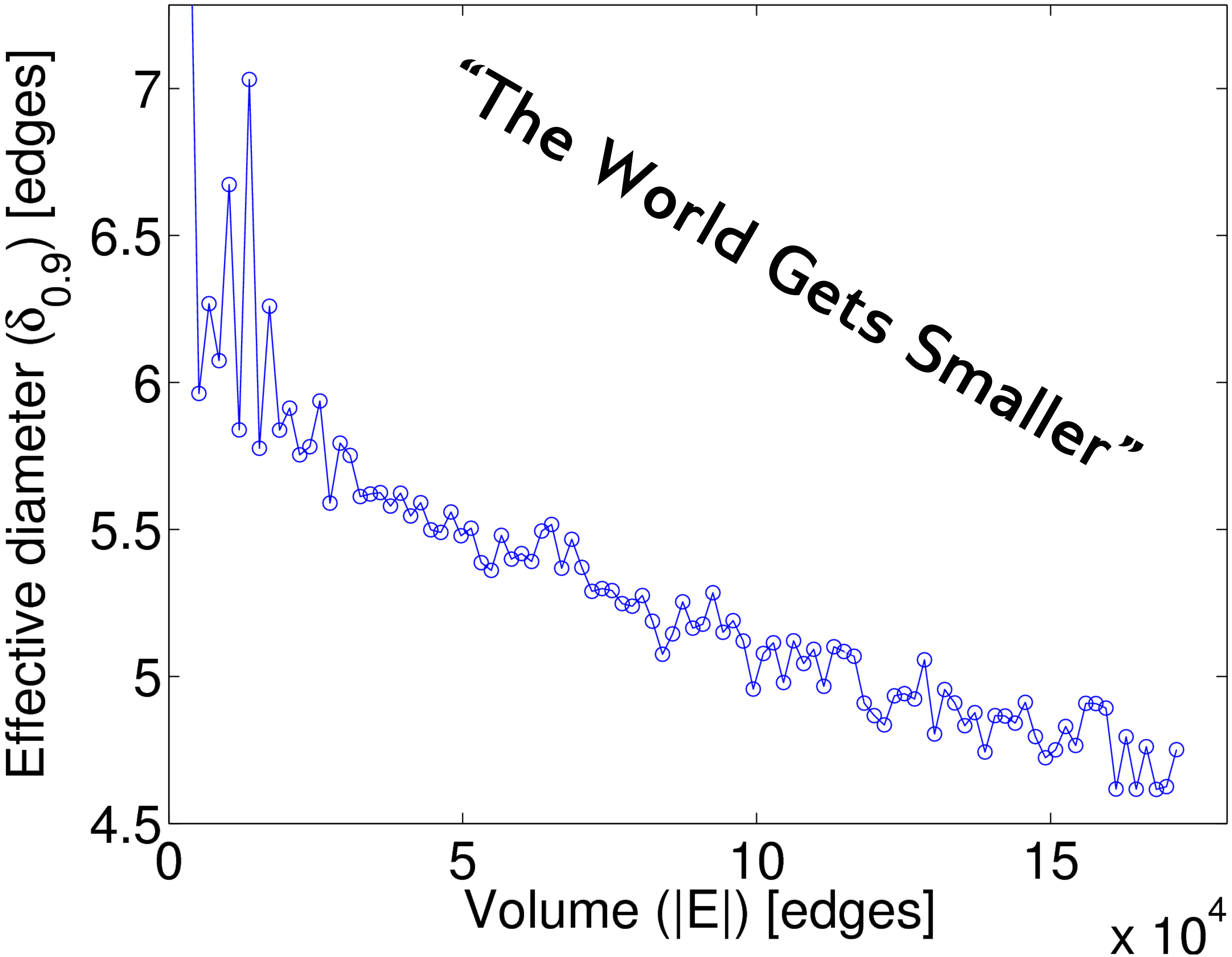
(1) Length of paths



“Large world” “Small world”

90-percentile effective diameter $\delta_{0.9}$





(Leskovec, Kleinberg & Faloutsos 2007)

Outline

(A) How can structural diversity be measured?

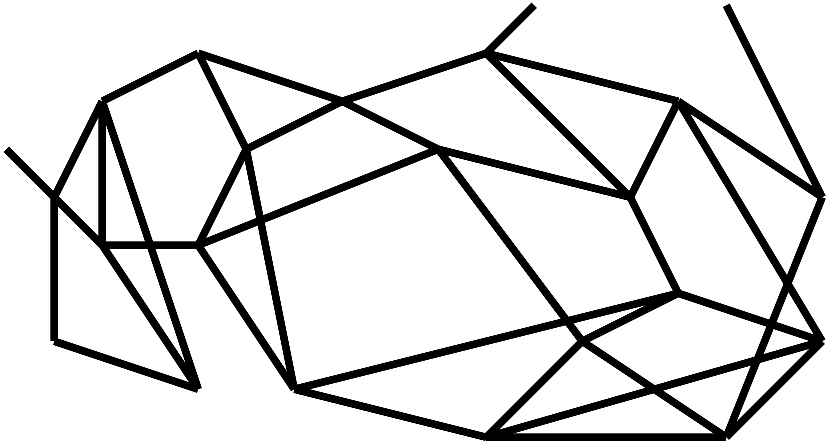
(B) How does diversity change?

(A) How to Measure Diversity in a Network?

(1) Length of paths
(2) Numbers of neighbors
(3) Size of communities
(4) Random walks
(5) Controllability

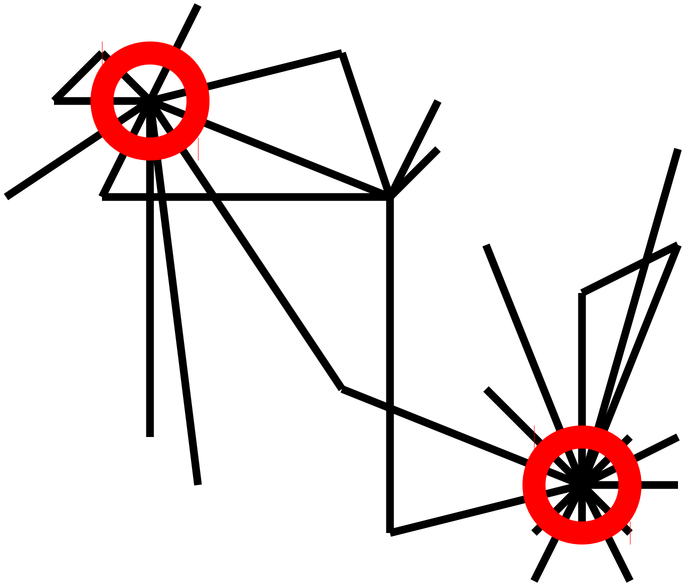
(2) Number of neighbors

Diversity

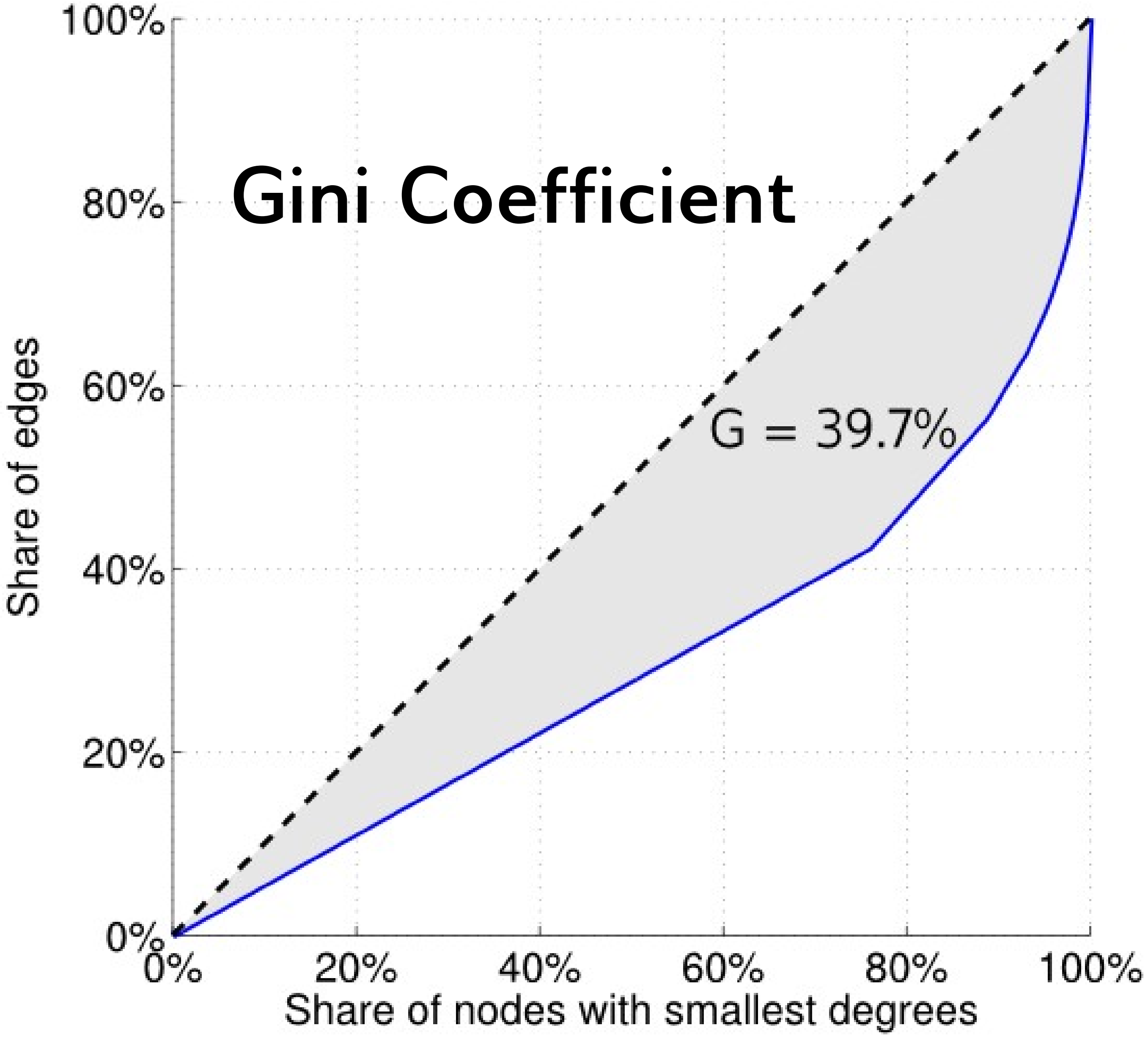


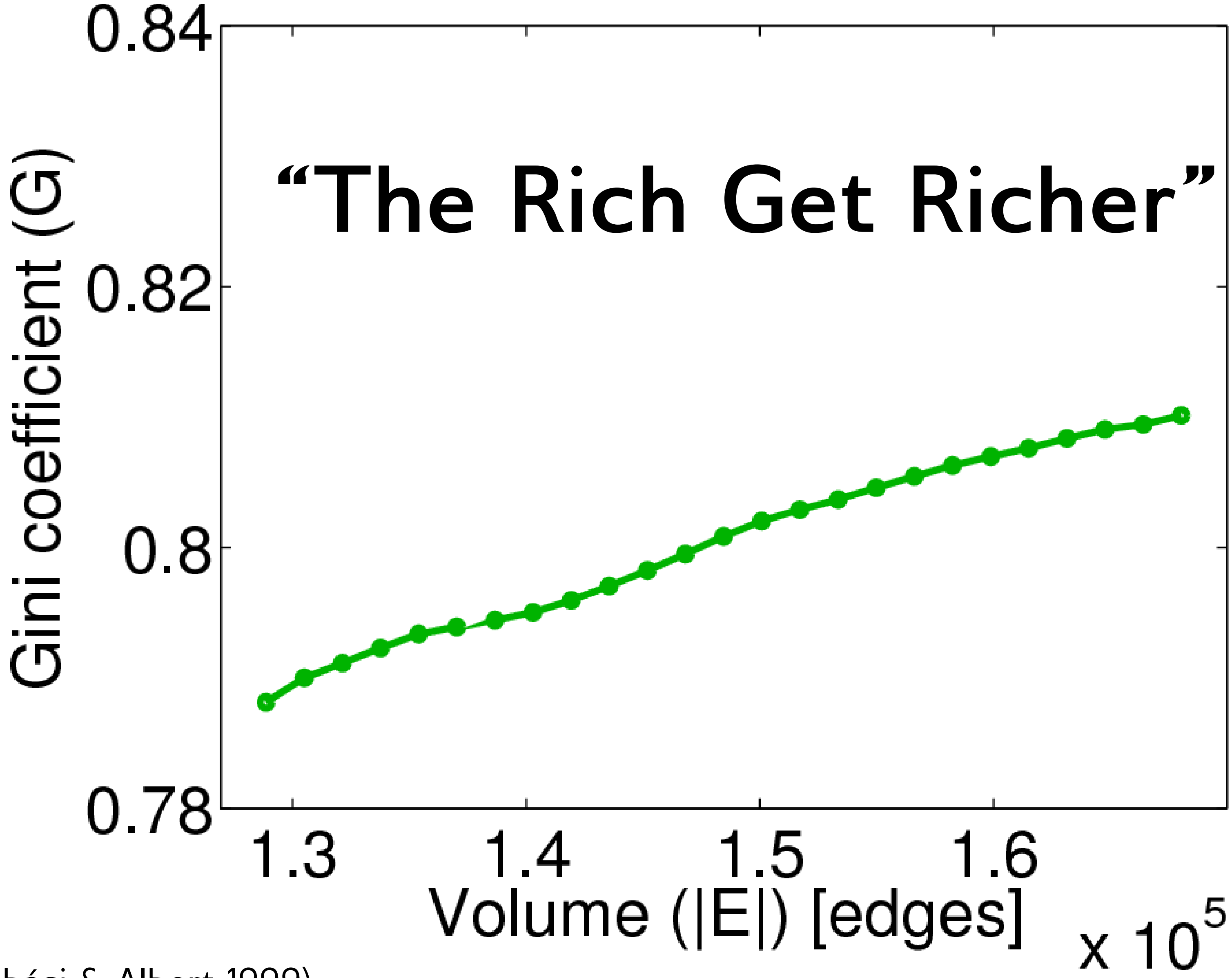
$$d(i) \approx d(j)$$

No diversity



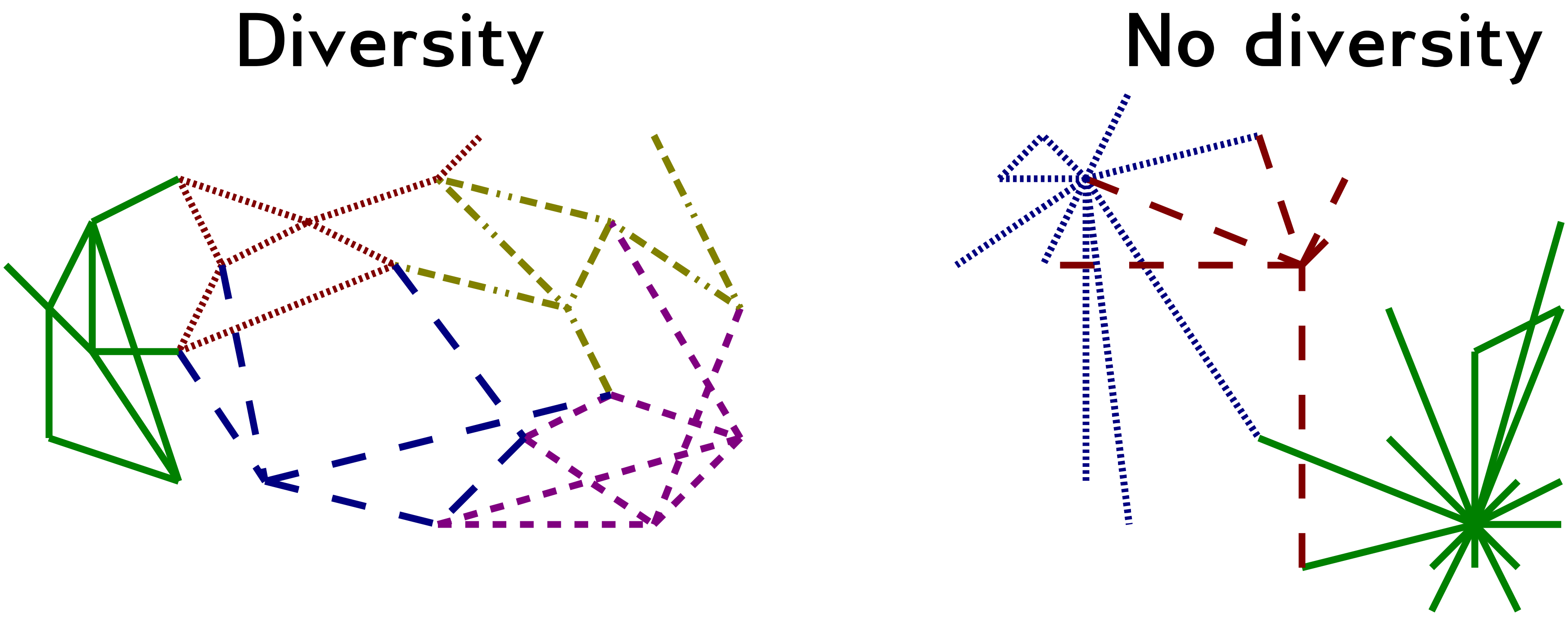
$$d(i) \ll d(j)$$


Gini Coefficient
G = 39.7%
100%
80%
60%
40%
20%
0%
Share of edges
0%
20%
40%
60%
80%
100%
Share of nodes with smallest degrees



(Barabási & Albert 1999)

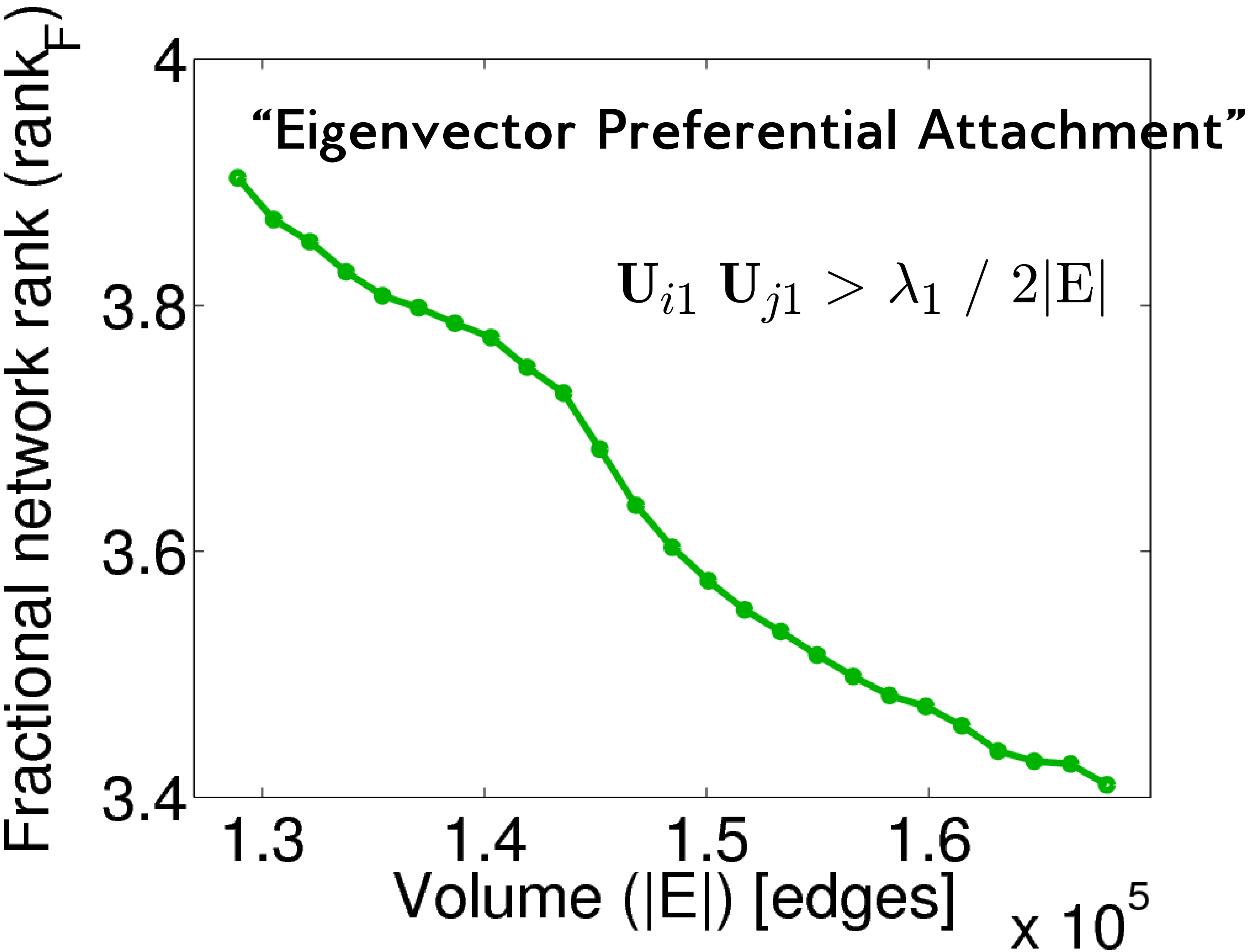
(3) Size of communities



Fractional Rank

Spectrum of the graph $= \{\lambda_1, \lambda_2, \lambda_3, \ldots\}$

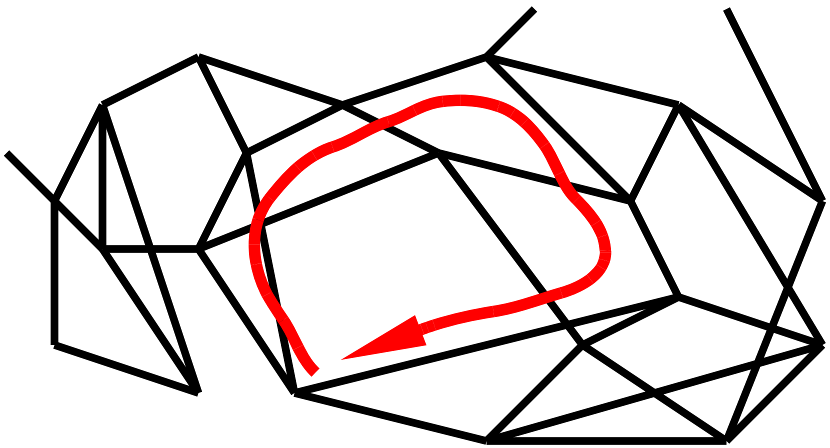
$$\mathrm{rank_F} = \sum\nolimits_{\mathrm{k}} (\lambda_{\mathrm{k}} \;/\; \lambda_1)^2 = (\|\mathbf{A}\|_{\mathrm{F}} \;/\; \|\mathbf{A}\|_2)^2 = 2|\mathrm{E}| \;/\; {\lambda_1}^2$$



“Eigenvector Preferential Attachment”
$\mathbf{U}_{i1}\ \mathbf{U}_{j1} > \lambda_1\ /\ 2|\mathrm{E}|$
Fractional network rank (rank_F)
4
3.8
3.6
3.4
1.3
1.4
1.5
1.6
Volume (|E|) [edges]
x 10^5

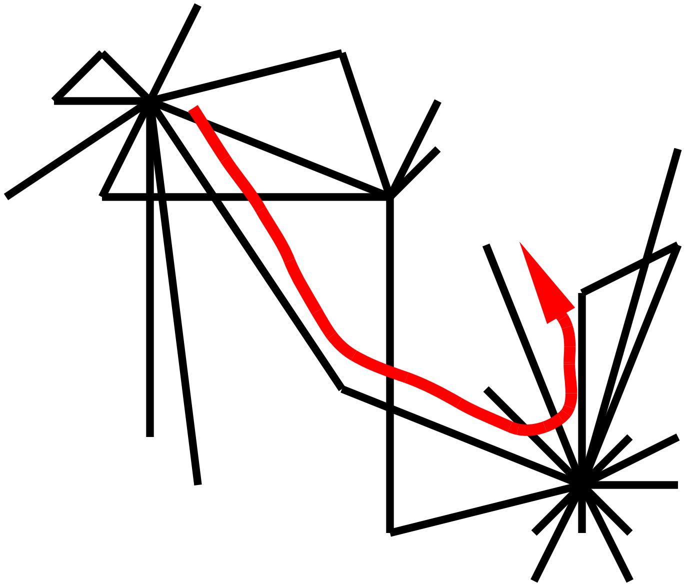
(4) Random walks

Diversity



$P_{\mathrm{ret}}(L)$ large

No diversity



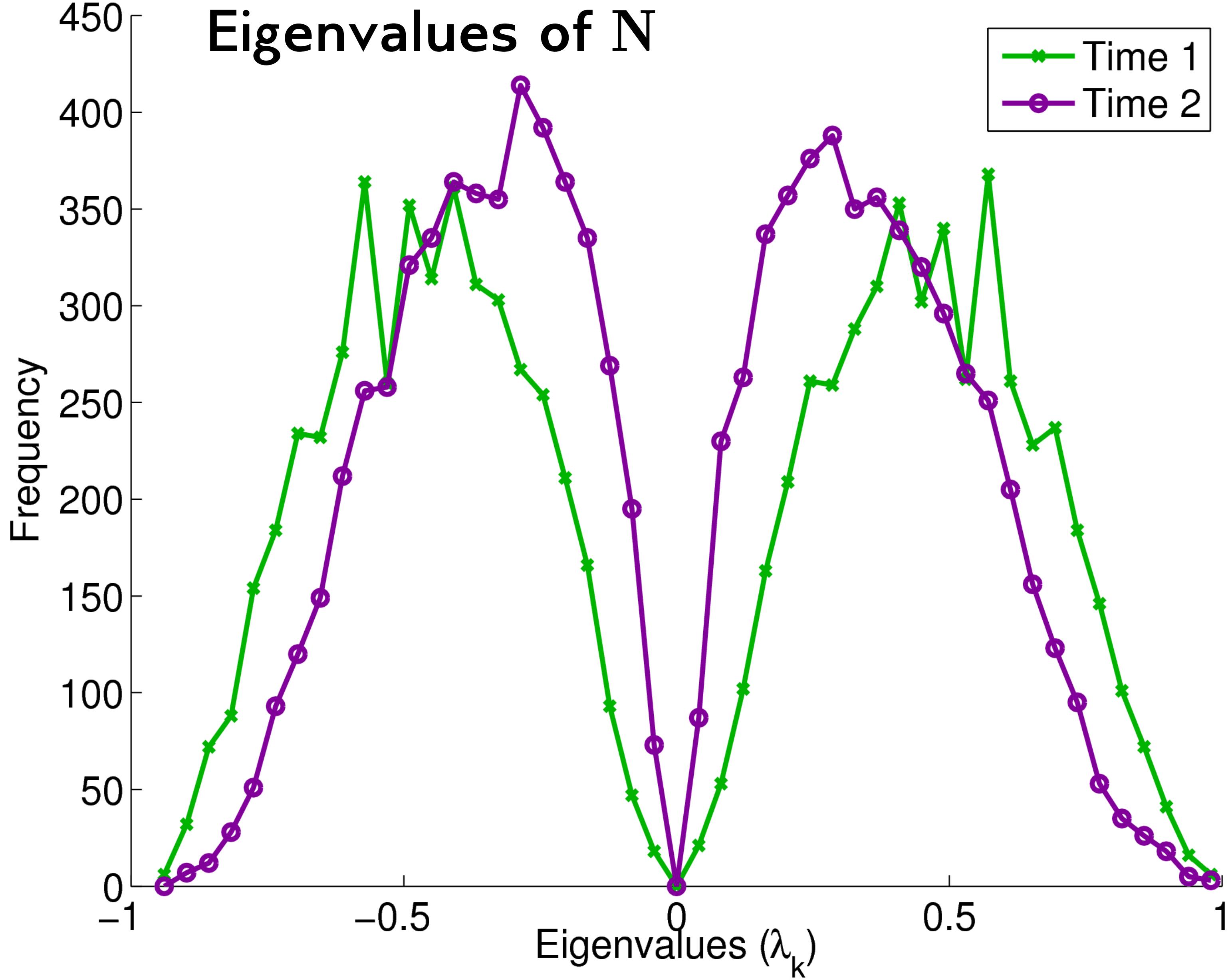
$P_{\mathrm{ret}}(L)$ small

Weighted Spectral Distribution

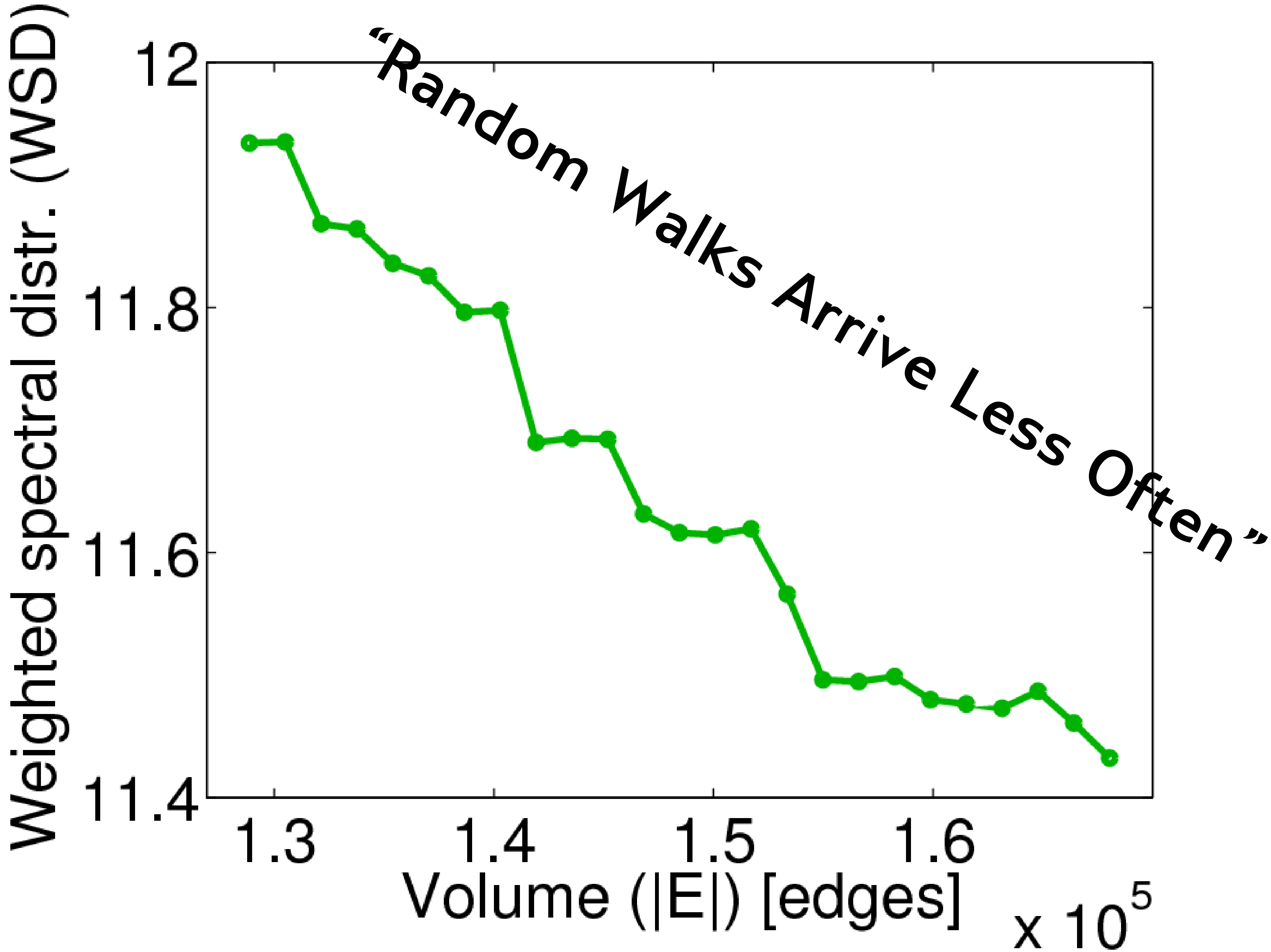
$$P_{\mathrm{ret}}(L) = \Sigma_{(i,\ j,\ .\ .\ .\ k)}\ (d(i)\ d(j)\ .\ .\ .\ d(k))^{-1}$$
$$= \mathrm{tr}(\mathbf{N}^L)$$
$$= \Sigma_k\ {\lambda_k}^L$$

where λ_k are eigenvalues of $\mathbf{N} = \mathbf{D}^{-1/2}\ \mathbf{A}\ \mathbf{D}^{-1/2}$.

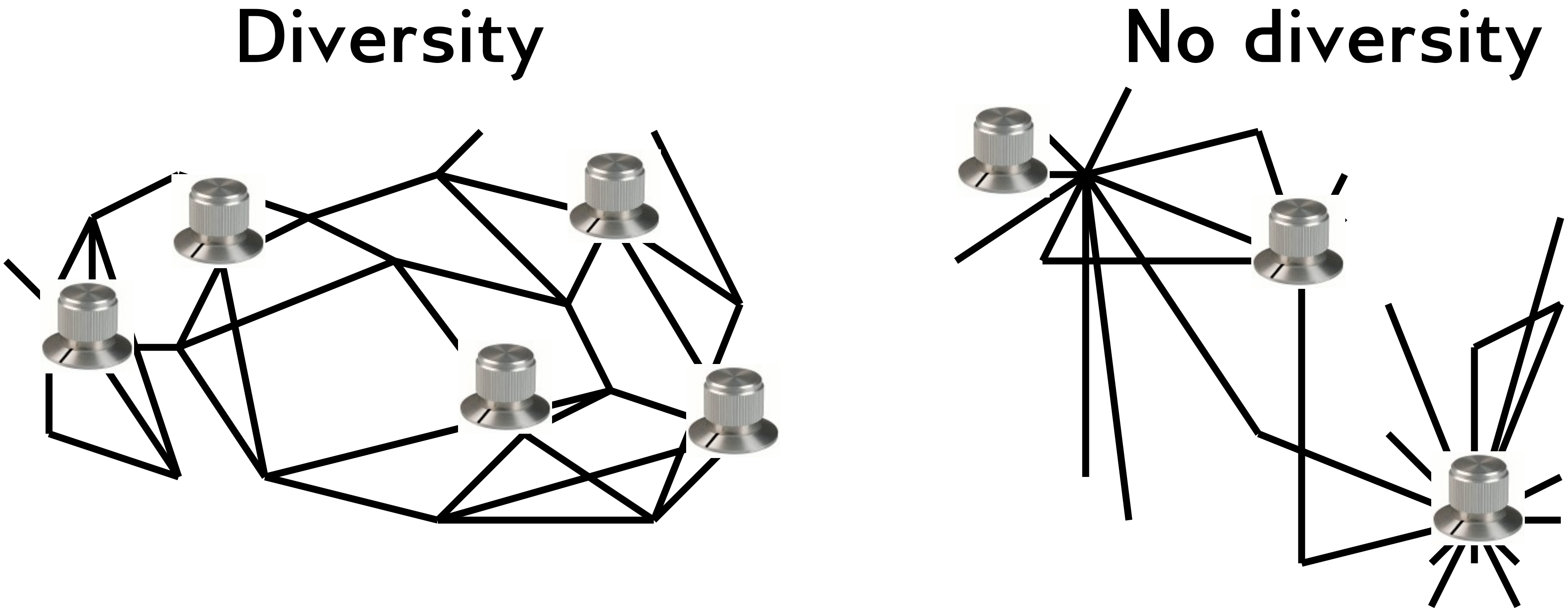
Here: Use $L = 4$ and $k \leq R$



Eigenvalues of N
Time 1
Time 2
Frequency
Eigenvalues (λ_k)
450
400
350
300
250
200
150
100
50
0
−1
−0.5
0
0.5
1

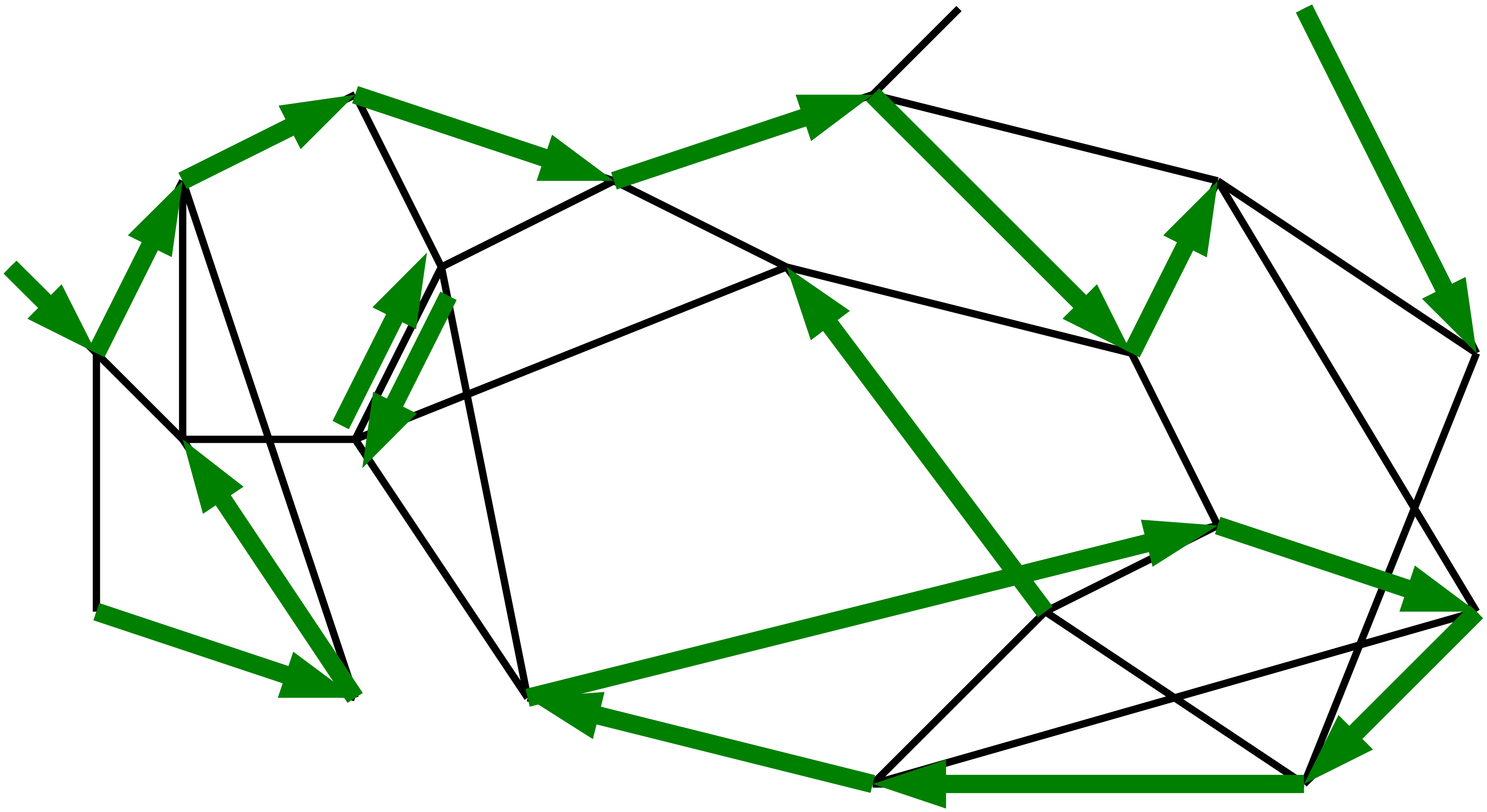

“Random Walks Arrive Less Often”
Weighted spectral distr. (WSD)
12
11.8
11.6
11.4
1.3
1.4
1.5
1.6
Volume (|E|) [edges]
x 10^5

(5) Controllability

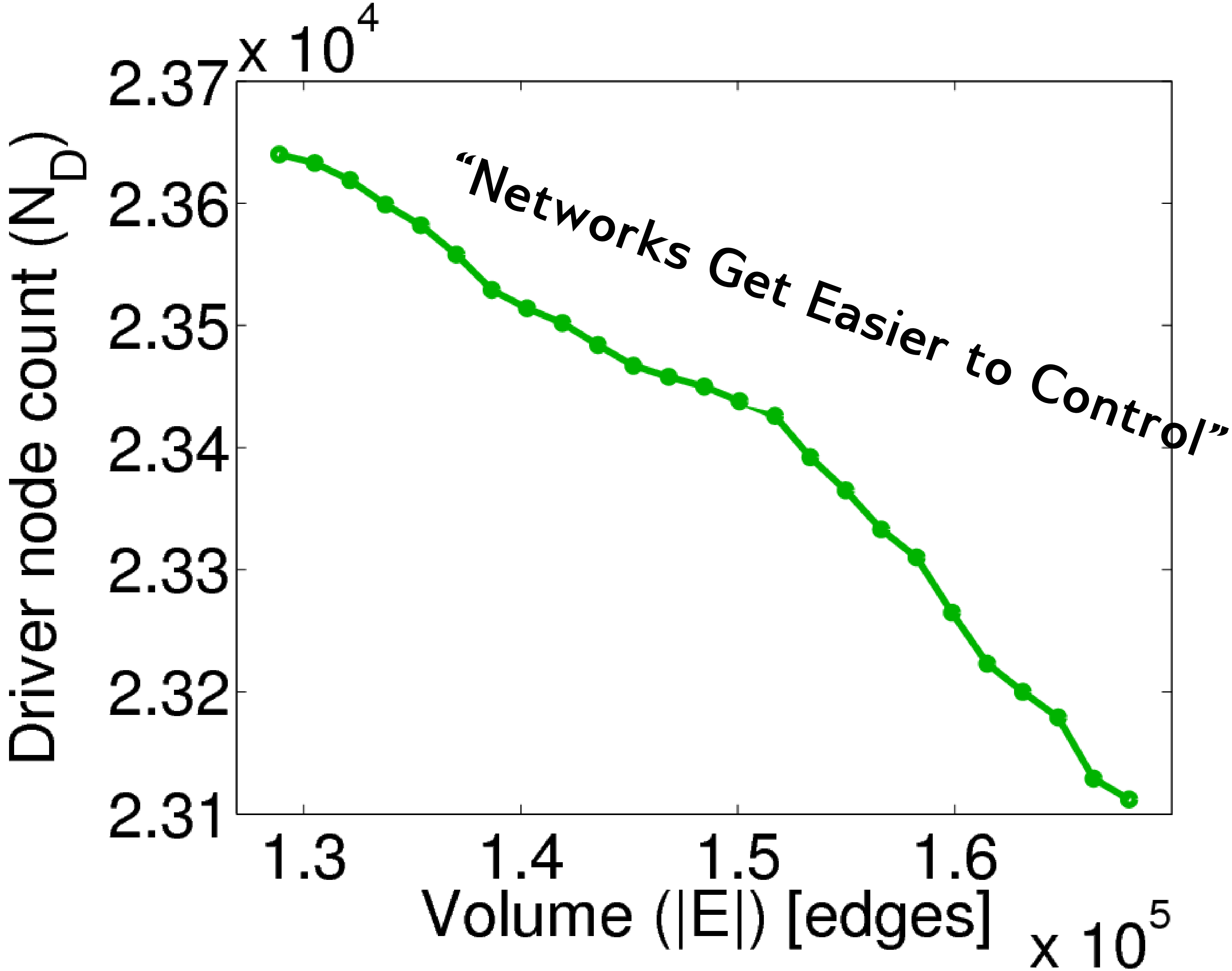


(Liu, Slotine & Barabási 2011)

Find a maximal directed 2-matching



$$\#\text{Knobs needed} = |\mathrm{V}| - \max |\mathrm{M}|$$


x 10^4
2.37
2.36
2.35
2.34
2.33
2.32
2.31
Driver node count (N_D)
"Networks Get Easier to Control"
1.3
1.4
1.5
1.6
Volume (|E|) [edges]
x 10^5

(B) Experiments

27 networks from konect.uni-koblenz.de

	Measure	Observed trends FULL	Observed trends CONNECTED	Predicted trends	Monotonicity CONNECTED
	d	(24) Up	(27) Up		Up
Pref. att.	G	(24) Up	(17) —	Up	
	J	(23) Up	(20) Up	Down	
	γ	(21) Down	(25) Down	Down	
	H_{er}	(19) Down	(12) —	Down	
Connect.	$\delta_{0.9}$	(18) Down	(26) Down	Down	Down
	$\vartheta_r(n)$	(10) —	(22) Down	Down	
	C_{r}	(12) —	(22) Down	Down	Down
	a	(15) —	(27) Up	Up	Up
L. pred.	c	(7) — [a]	(10) Up [a]	Up	
	$\mathrm{rank}_{\mathrm{F}}$	(13) —	(19) Down	Down	
	α	(19) Up	(23) Up	Up	

[a] For the clustering coefficient, the total number of networks is 13, since bipartite networks are excluded.

Thank! You!

Jérôme Kunegis
Sergej Sizov
Julia Perl
Felix Schwagereit
Damien Fay

@kunegis

University of Koblenz–Landau, Germany
Bournemouth University, UK

konect.uni-koblenz.de

Networks Used

		Network	Flags	$|V|$	$|E|$
[64]	ben	Wikibooks, English	B	167,525	1,164,576
[64]	bfr	Wikibooks, French	B	30,997	201,727
[10]	DG	Digg	U M	30,398	87,627
[64]	el	Wikipedia, Greek	B	149,904	1,837,141
[43]	EL	Wikipedia elections	U M	8,297	107,071
[33]	EN	Enron	U M	87,273	1,148,072
[47]	EP	Epinions trust	U M	131,828	841,372
[56]	Fc	Filmtipset	B	75,360	1,266,753
[9]	HA	Haggle	U	274	28,244
[29]	HY	Hypertext 2009	U	113	20,818
[29]	IF	Infectious	U	410	17,298
[19]	M1	MovieLens 100k	B	2,625	100,000
[19]	M2	MovieLens 1M	B	9,746	1,000,209
[19]	Mti	MovieLens tag–movie	B	24,129	95,580
[19]	Mui	MovieLens user–movie	B	11,610	95,580
[19]	Mut	MovieLens user–tag	B	20,537	95,580
[64]	nen	Wikinews, English	B	173,772	901,416
[64]	nfr	Wikinews, French	B	26,546	193,618
[61]	Ol	Facebook friendships	U	63,731	1,545,686
[61]	Ow	Facebook wall posts	U M	63,891	876,993
[64]	qen	Wikiquote, English	B	116,363	549,210
[13]	RM	Reality Mining	U	96	1,086,404
[20]	SD	Slashdot threads	U M	51,083	140,778
[54]	SX	Sexual escorts	B	16,730	50,632
[67]	TO	Internet topology	U	34,761	171,403
[52]	UC	UC Irvine messages	U M	1,899	59,835
[53]	UF	UC Irvine forum	B	1,421	33,720

U Unipartite network
B Bipartite network
M Network with multiple edges

Questions

Did you try the power law exponent instead of the Gini coefficient?

→ Yes, but see (Kunegis & Preusse 2012)

Did you try the absolute value instead of the square in rank_F ?

→ Yes, it leads to the nuclear norm instead of the Frobenius norm, which is harder to compute and highly correlates with it

Isn't it hard to find a maximal directed 2-matching?

→ It takes a runtime of $O(|V|^{1/2} |E|)$, we use Boost Graph Lib

How is the approximation using only *r* eigenvalues for the WSD justified?

→ By observing that all eigenvalues shrink in unison

References

J. Kunegis, S. Sizov, F. Schwagereit, D. Fay. Diversity Dynamics in Online Networks. Proc. Conf. on Hypertext and Social Media, 2012.

J. Kunegis, J. Preusse. Fairness on the Web: Alternatives to the Power Law. Proc. Web Science Conf., 2012.

Y.-Y. Liu, J.-J. Slotine, A.-L. Barabási. Controllability of Complex Networks. Nature, 473:167–173, May 2011.

J. Leskovec, J. Kleinberg, C. Faloutsos. Graph Evolution: Densification and Shrinking Diameters. ACM Trans. Knowledge Discovery from Data, 1(1):1–40, 2007.

A.-L. Barabási, R. Albert. Emergence of Scaling in Random Networks. Science, 286(5439):509–512, 1999.

Credits

http://www.shewearsshortshorts.com/2012/01/downside.html
https://twitter.com/#!/justinbieber
http://www.iconspedia.com/icon/nerd-4255.html
http://hk.digikey.com/1/3/index1227.html