Milgram's experiment in various networks

K.Malarz and K.Kułakowski

Faculty of Physics & Nuclear Techniques

AGH University of Science & Technology

Kraków, Poland

1 Introduction

- sociology (the strength of weak ties, terrorism)
- computer science (Internet, WWW)

1.1 Milgram’s experiment

A group of individuals was asked to send a letter to a target person in Boston via an acquaintance who was supposed to be closer to the target, than the sender.

The mean length of the letter chain was less than seven.

1.2 Small world

- $d \equiv$ the mean length of the shortest path between two vertices
- The growth of d slower than any positive power of N is called a small-world effect.
- By definition, a networks is a small world if it shows the small-world effect.

1.3 Dictionary

- **a graph** $\equiv$ a set of nodes and a set of edges
- **a simple graph** $\equiv$ a graph without loops and without multiple edges
- **a forest** $\equiv$ a graph without cycles
- **a tree** $\equiv$ a compact forest
- **a classical random graph** (CRG) $\equiv$ Erdös–Rényi and/or Gilbert model

By **growing** we mean adding subsequent nodes to an already existing graph with m links:

- $m = 1 \longrightarrow$ trees
- $m = 2 \longrightarrow$ simple graphs

$P(q)$ = probability that a new node will be attached to existing node q

- for **exponential** networks $P(q)$ is uniform
- for **scale-free** (preferential, Albert–Barabási) networks $P(q)$ is proportional to the node degree (i.e. its number of edges)

- a network containing N nodes is fully characterized by its **connectivity matrix** $\mathbf{C}$: $c_N(i, j) = 1$ if the nodes i, j are linked together, and $c_N(i, j) = 0$ elsewhere

- in **a distance matrix** $\mathbf{S}$, the matrix element $s_N(i, j)$ is the number of links along the shortest path from i^{th} to j^{th}

1.4 Search strategy

- In this kind of contact experiments, to find an appropriate next person in the path is a nontrivial task, and several strategies are possible.
- One of the most obvious is to find a person most connected, i.e. a neighbouring node with the highest degree (**MCNS**).

- This strategy has been shown to be effective in networks with power-law degree distribution, but not in random graphs.
- **Here we ask, how the MCNS is effective in the exponential networks**.

2 Numerical approach

For growing networks the starting point is a matrix $\mathbf{S}$ for the tree of linked together two nodes:

$$\mathbf{S} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Selecting a node to link a new node is equivalent to select a number q of column/row of the matrix.

2.1 Trees

$$\forall\, 1 \le i \le N : s_{N+1}(N+1, i) = s_{N+1}(i, N+1) = s_N(q, i) + 1 \quad (1a)$$

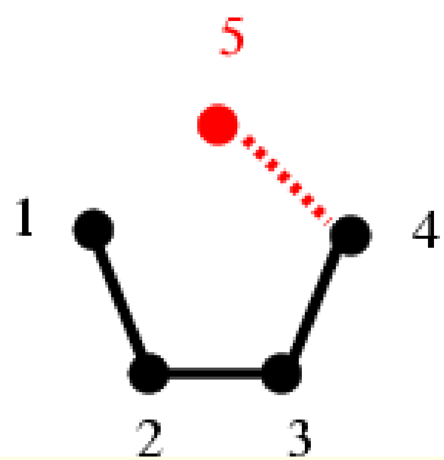
and

$$s_{N+1}(N+1, N+1) = 0 \quad (1b)$$

Distance matrix evolution for trees:

0	1	2	3
1	0	1	2
2	1	0	1
3	2	1	0

0	1	2	3	4
1	0	1	2	3
2	1	0	1	2
3	2	1	0	1
4	3	2	1	0



2.2 Simple graphs

In simple graphs cyclic paths are possible and the distance matrix $\mathbf{S}$ is to be rebuilt when adding each node. The $(N+1)^{\text{th}}$ node is added to existing nodes p and $q \neq p$:

$$\forall\, 1 \leq i, j \leq N : s_{N+1}(i,j)$$

$$= \min\left(s_N(i,j), s_N(i,p) + 2 + s_N(q,j),\right. \qquad \text{(2a)}$$

$$\left. s_N(i,q) + 2 + s_N(p,j)\right).$$

For new, $(N+1)^{\text{th}}$, column/row

$$\forall\, 1 \le i \le N : s_{N+1}(N+1, i) = s_{N+1}(i, N+1) = \min\left(s_N(p,i), s_N(q,i)\right) + 1 \qquad (2b)$$

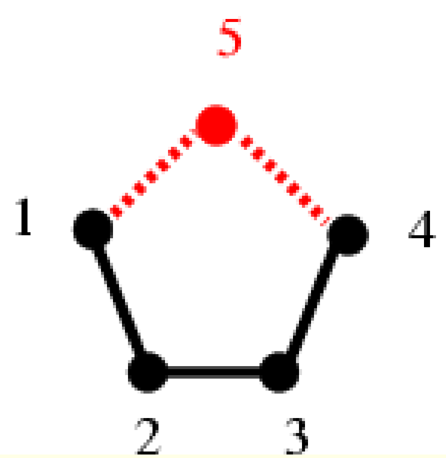
and again for the diagonal element

$$s_{N+1}(N+1, N+1) = 0 \qquad (2c)$$

Distance matrix evolution for simple graphs:

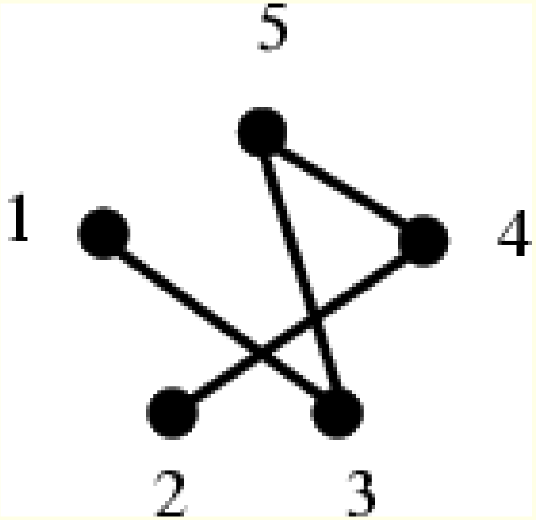
0	1	2	3
1	0	1	2
2	1	0	1
3	2	1	0

0	1	2	2	1
1	0	1	2	2
2	1	0	1	2
2	2	1	0	1
1	2	2	1	0



2.3 CRG

$$N = 5, L = 4, p = \frac{2L}{N(N-1)} = 0.4$$



- We start the simulation with an $N \times N$ matrix with all non-diagonal elements equal to N.

- We go through all non-diagonal elements of $\mathbf{S}$ and set $s(i, j < i)$ equal to one with the probability p.

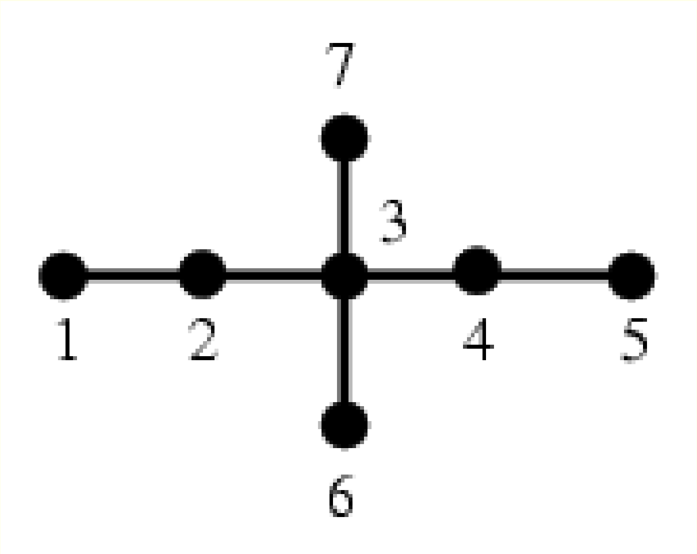
- The matrix $\mathbf{S}$ is kept symmetric.

- Each time, when a new edge is added, we have to rebuild the whole matrix $\mathbf{S}$ due to link between nodes i^{th} and j^{th}:

$$\forall 1 \leq m, n \leq N : s(m,n) = \min\big(s(m,n), s(m,i) + 1 + s(j,n), s(m,j) + 1 + s(i,n)\big). \quad (3)$$

- After such a procedure the matrix $\mathbf{S}_{N \times N}$ contains elements equal to N only if graph is not connected.

2.4 The Kertész list



1	2	2	3	3	3	3	4	4	5	6	7

2.5 Network characteristics

In a connected graph, nodes can be characterised locally (with their degree k) or globally (e.g. with their average length path ξ to other nodes).

Here we investigate how ξ depends on k.

- k_i is # of '1' in i^{th} row/column,
- $\xi_i \equiv [N^{-1} \sum_{j=1}^{N} s_N(i,j)]$,
- $d = [N^{-2} \sum_{i=1}^{N} \sum_{j=1}^{N} s_N(i,j)] = [N^{-1} \sum_{i=1}^{N} \xi_i]$,
- $\xi(k) \equiv$ an average of ξ_i for nodes with given degree $k = k_i$,

where $[\cdots]$ is an average over N_{run} different matrices.

3 Results

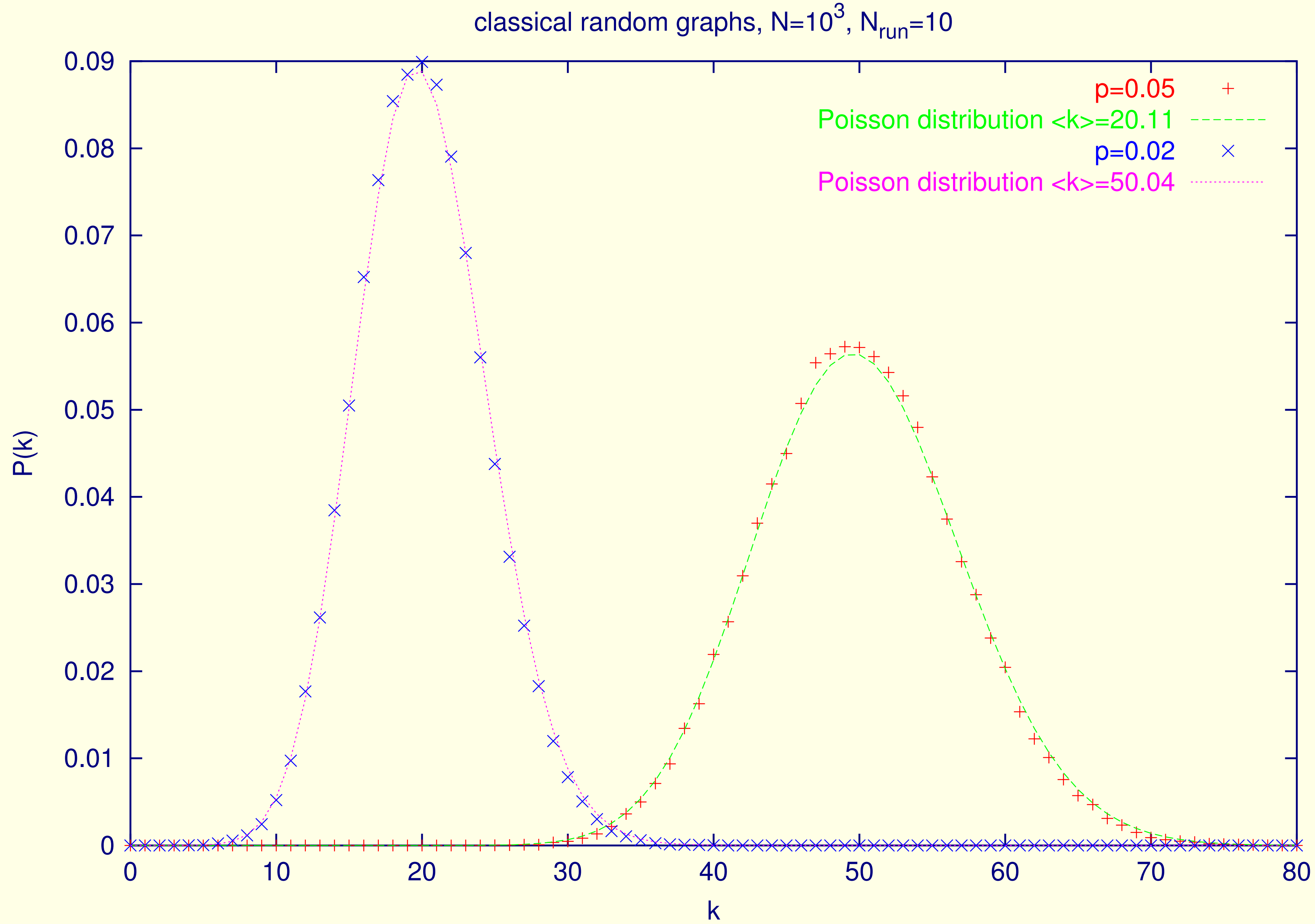
Each network contains at least a thousand of nodes ($N = 10^3$).

The results are typically averaged over $N_{\text{run}} = 10^7$, 10^3 and 100 different networks for trees, simple graphs and CRG, respectively.

3.1 Degree distribution

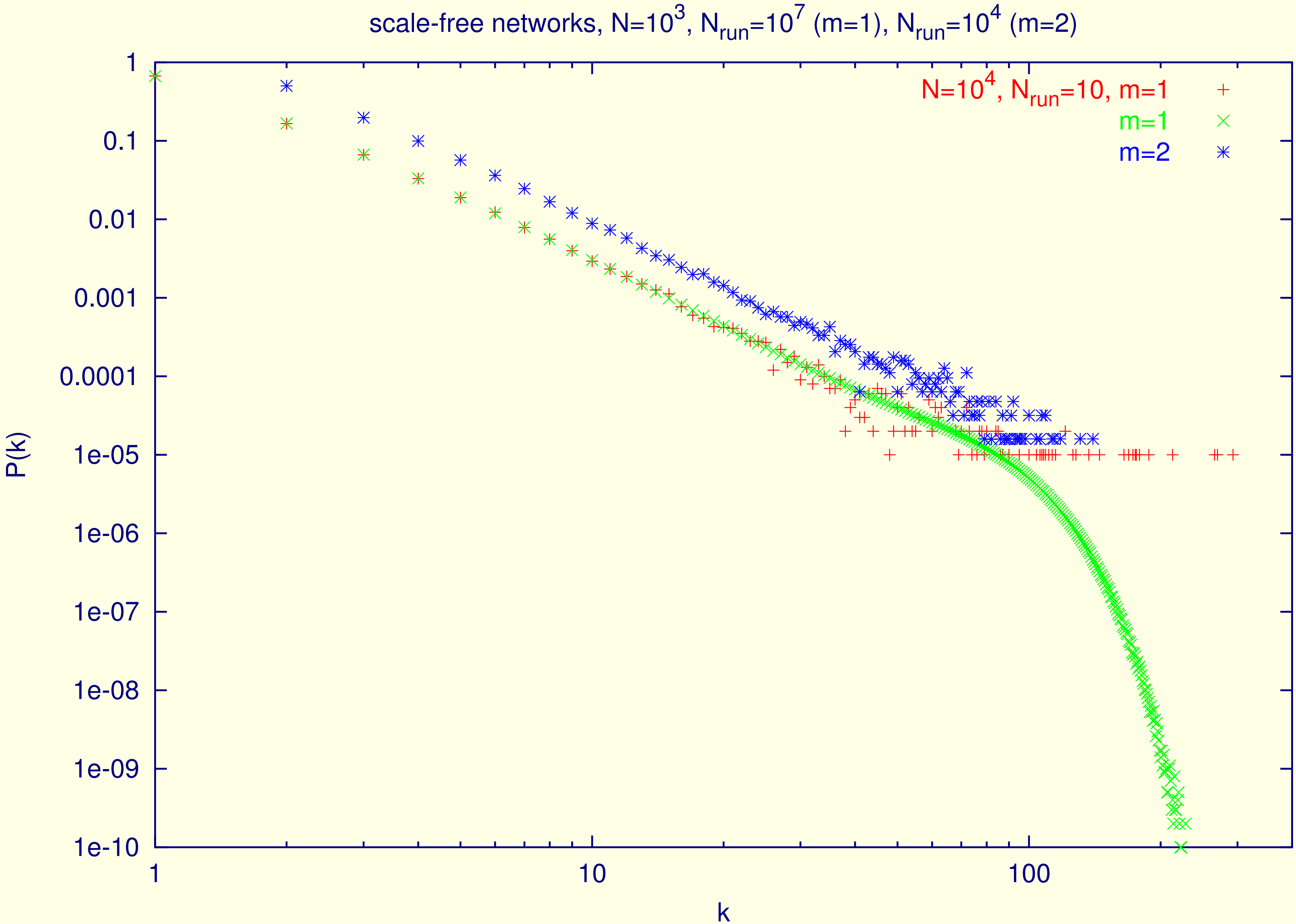
3.1.1 CRG

The degree distribution for CRG follows the Poisson distribution $P(k) = \exp(-\langle k \rangle) \cdot \langle k \rangle^k / k!$, with $\langle k \rangle \approx 20$ and $\langle k \rangle \approx 50$ for $p = 0.02$ and $p = 0.05$, respectively.


classical random graphs, N=10^3, N_run=10
P(k)
k
p=0.05
Poisson distribution <k>=20.11
p=0.02
Poisson distribution <k>=50.04
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
10
20
30
40
50
60
70
80

3.1.2 Scale-free natworks

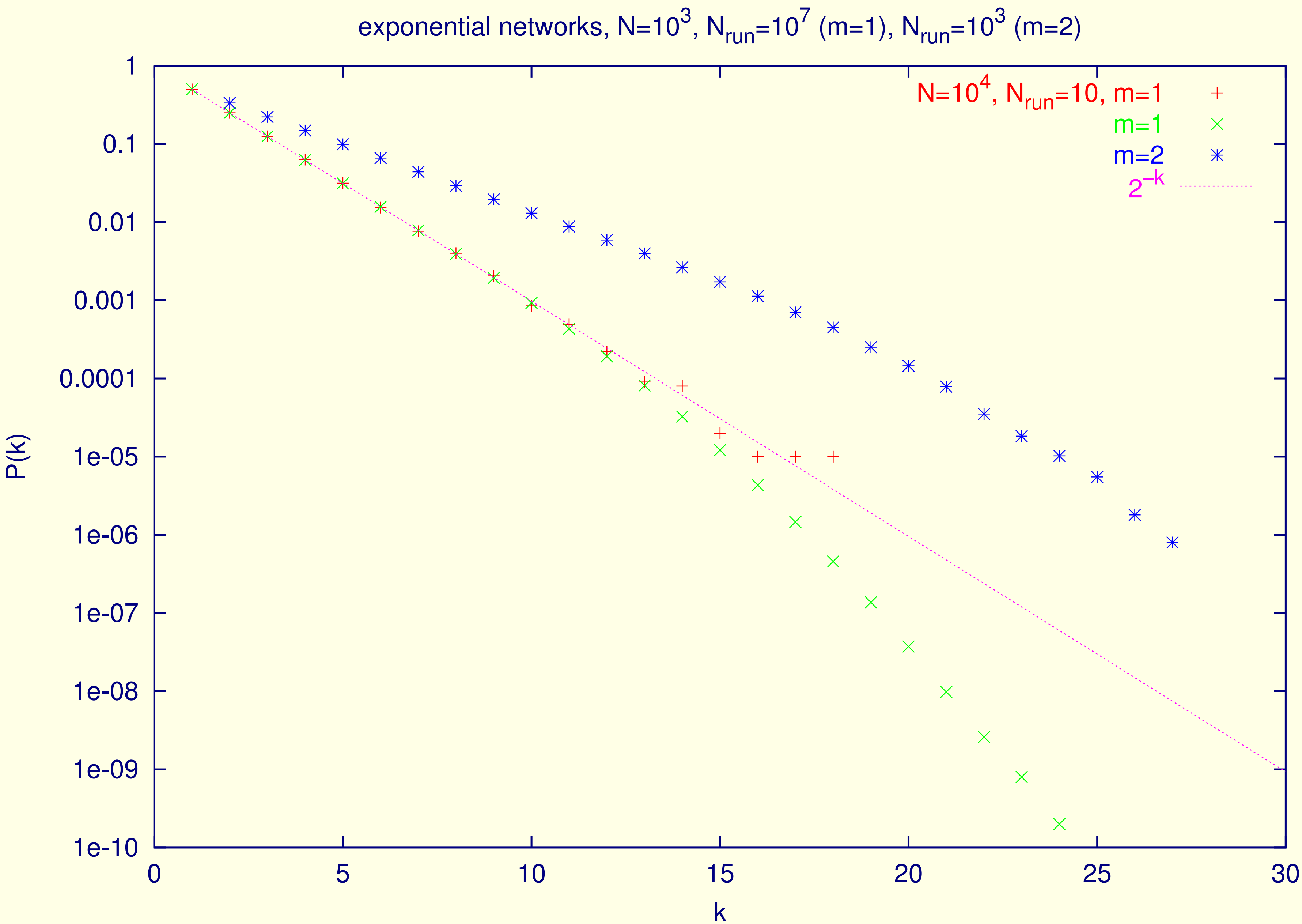
For the scale-free networks we reproduce $P(k) \propto k^{-\gamma}$ with $\gamma \approx 2.7$, while the theoretical value is 3.0. The numerical reduction of γ is known to be caused by the finite-size effect.



scale-free networks, N=10^3, N_run=10^7 (m=1), N_run=10^4 (m=2)
N=10^4, N_run=10, m=1
m=1
m=2
P(k)
k
1
0.1
0.01
0.001
0.0001
1e-05
1e-06
1e-07
1e-08
1e-09
1e-10
1
10
100

3.1.3 Exponential networks

For the exponential trees the node degree distribution is verified to be $P(k) = 2^{-k}$.



exponential networks, N=10^3, N_run=10^7 (m=1), N_run=10^3 (m=2)
N=10^4, N_run=10, m=1
m=1
m=2
2^-k
P(k)
k
1
0.1
0.01
0.001
0.0001
1e-05
1e-06
1e-07
1e-08
1e-09
1e-10
0
5
10
15
20
25
30

3.2 Small-world effect

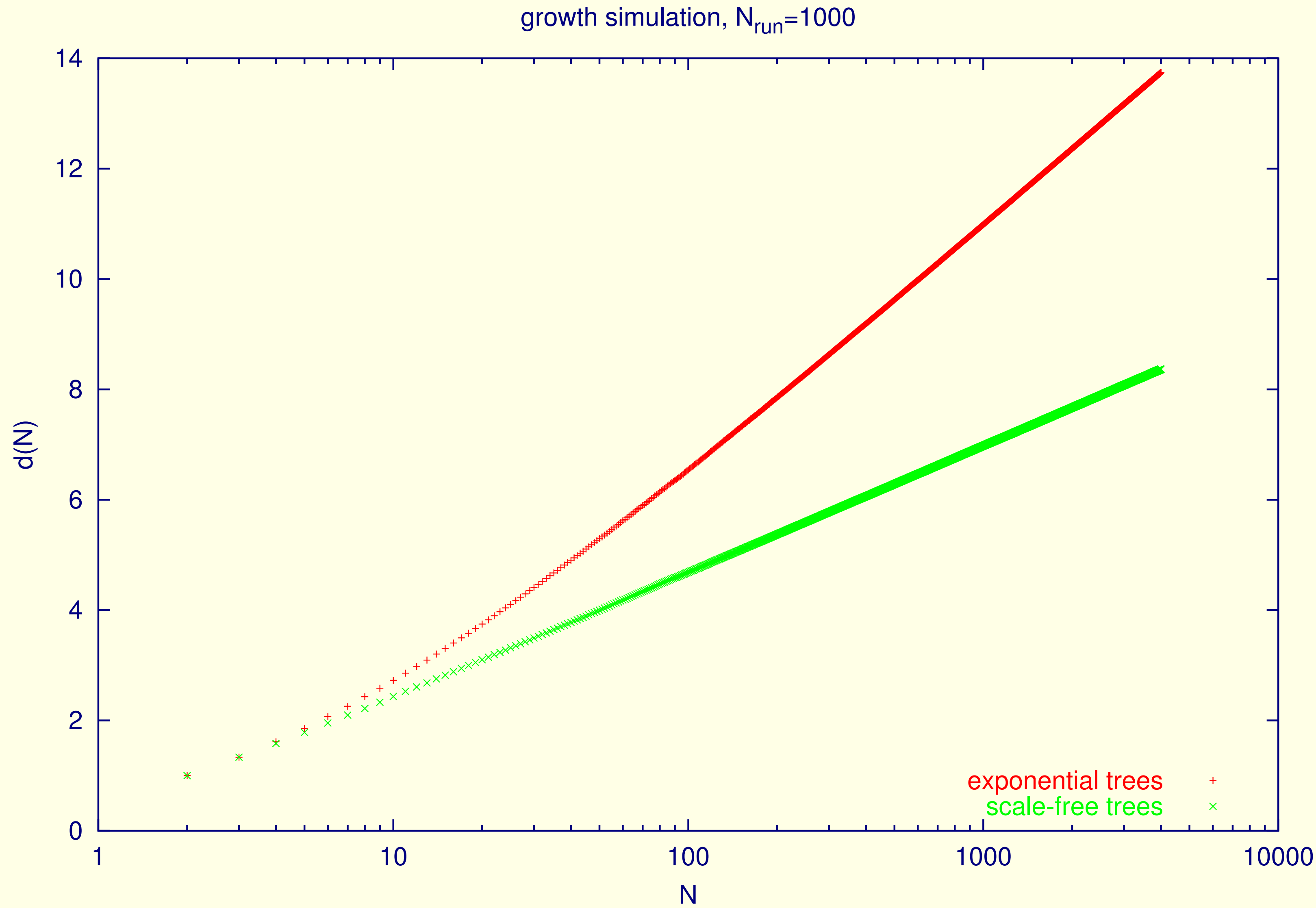


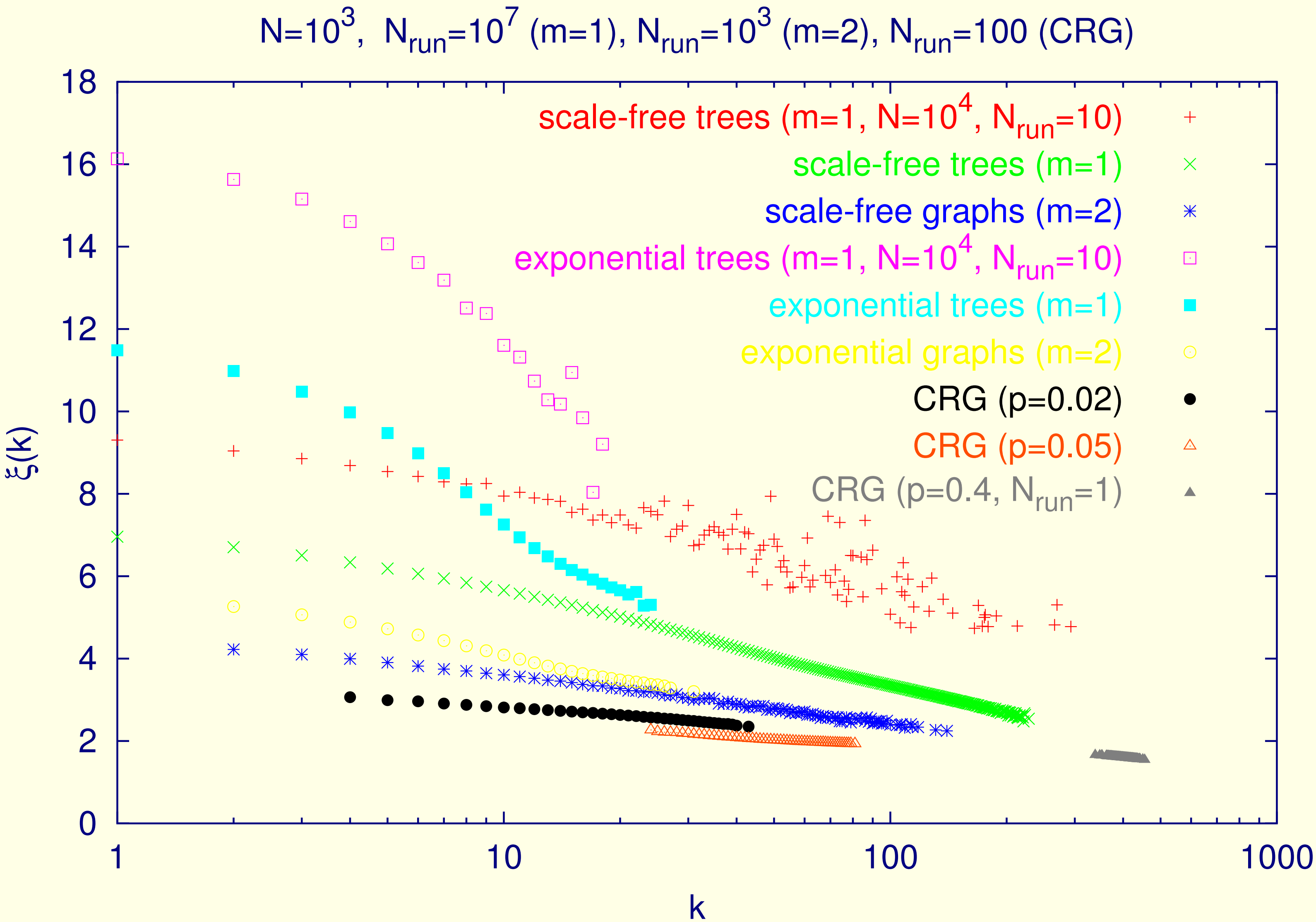
Table 1: The mean distance $d(N) = a \ln N + b$ for different evolving exponential networks.

m	1	2
a	2.00	0.71
b	-2.84	0.16

Table 2: The mean distance $d(N) = a \ln N + b$ for different evolving scale-free networks.

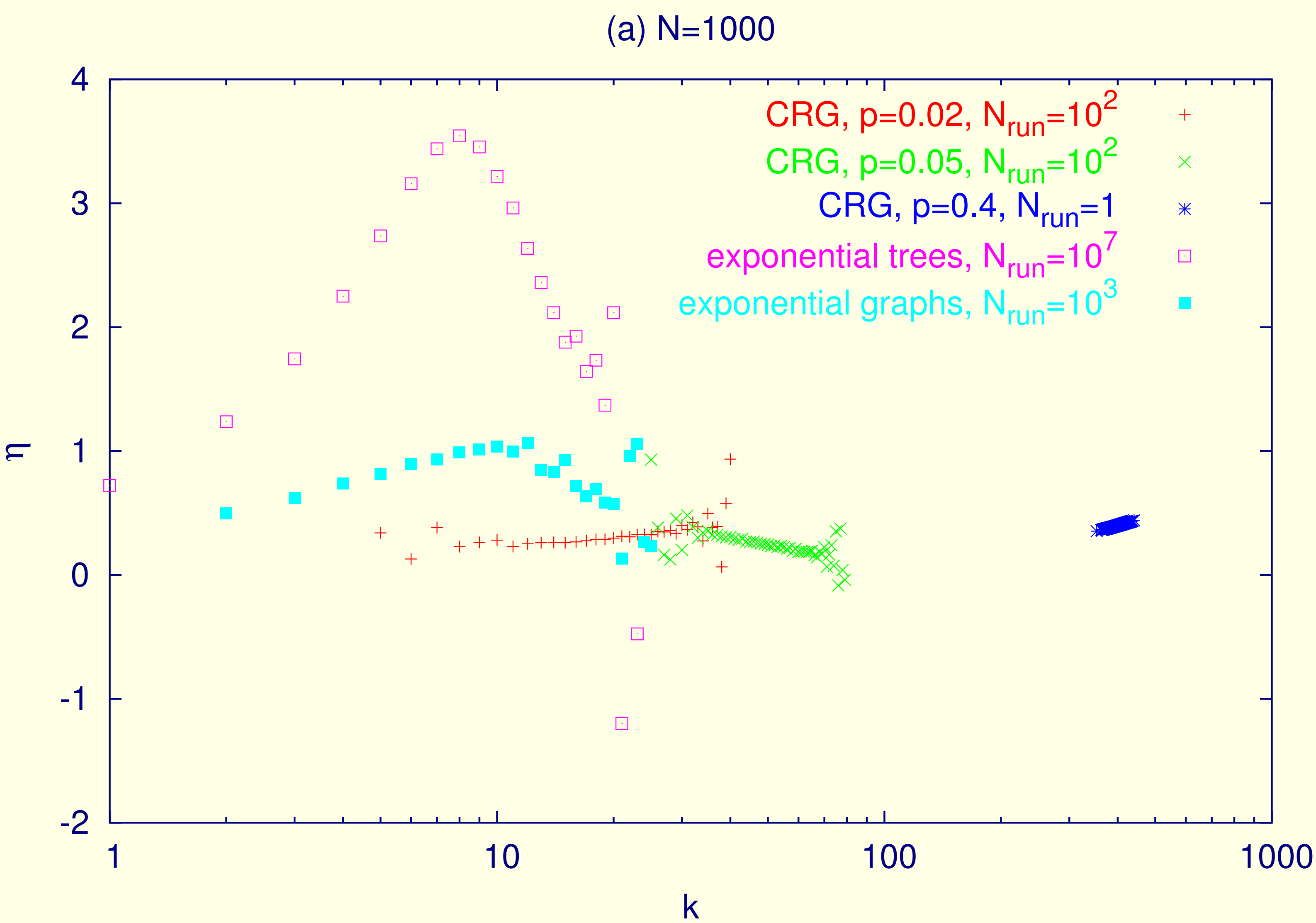
m	1	2
a	1.00	0.48
b	-0.08	0.83

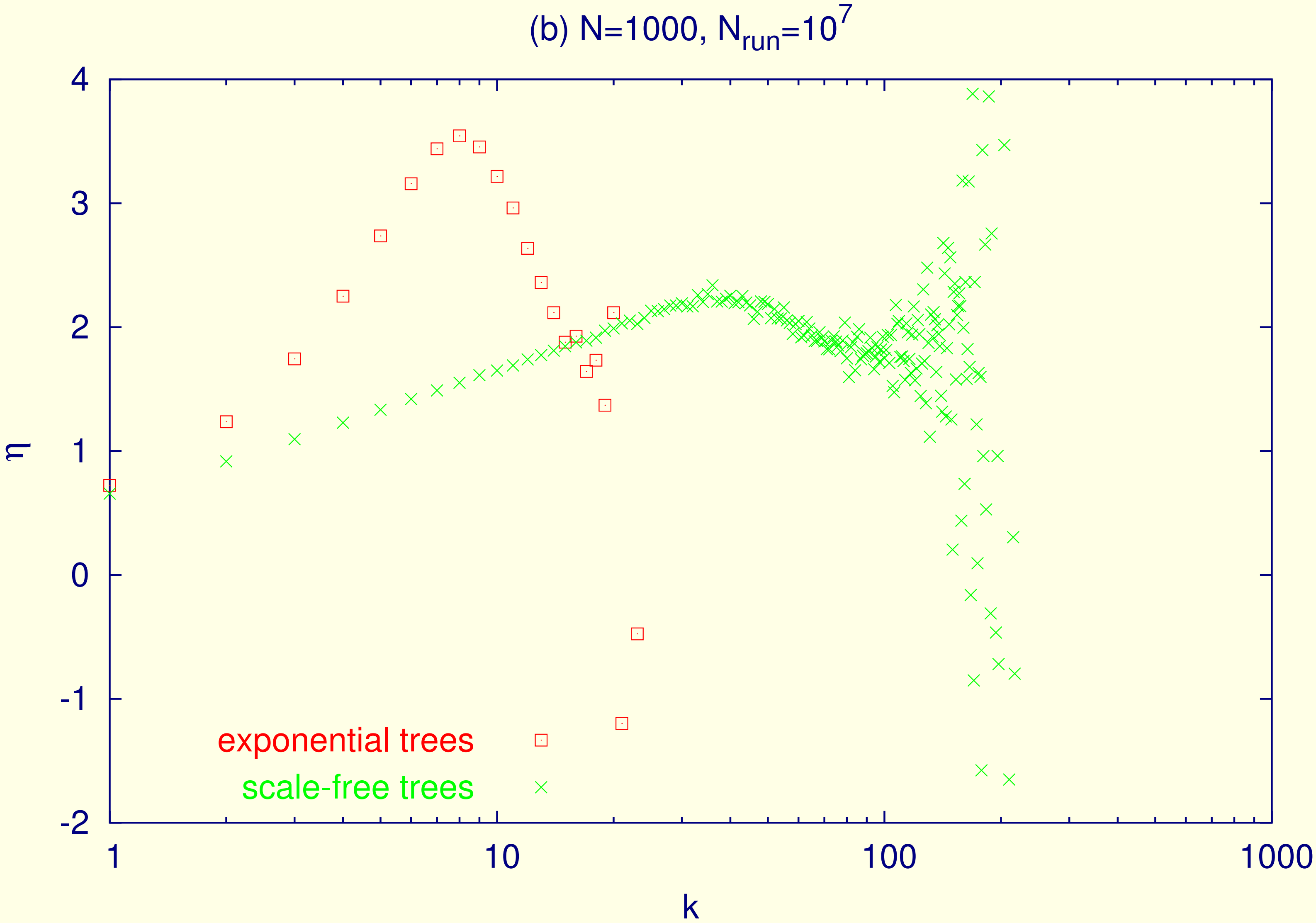
3.3 Effectiveness of MCNS



The slope of the curve $\xi(k)$ brings an information, how this strategy is effective for a given network. The effectiveness of MCNS for nodes of given k can be evaluated by an index

$$\eta = -\frac{\partial \xi}{\partial \ln k}. \tag{4}$$


(a) N=1000
CRG, p=0.02, $N_{run}=10^2$
CRG, p=0.05, $N_{run}=10^2$
CRG, p=0.4, $N_{run}=1$
exponential trees, $N_{run}=10^7$
exponential graphs, $N_{run}=10^3$
η
k
4
3
2
1
0
-1
-2
1
10
100
1000



(b) N=1000, $N_{run}=10^7$
η
k
exponential trees
scale-free trees
4
3
2
1
0
-1
-2
1
10
100
1000

4 Conclusioins

- For the random graphs the mean distance ξ practically does not depend on k, and the index η is close to zero.
- MCNS works better for scale-free networks than for CRG.
- The new result is that MCNS applied in an exponential tree is even more effective, than in the scale-free tree.

- In the scale-free networks, local fluctuations of the degree are enhanced by subsequent linkings.
- The multiple centres of high degree can be created, and the growing concentrates on these centres.
- Then, MCNS can be misleading, as it leads always to a local centre; however, sometimes the target is somewhere else.

- This enhancement is absent in the exponential networks, and that is why MCNS works better there. We note that this argumentation works well for trees.
- For other systems, there is more than one path between each pair of nodes, and any educated but general strategy cannot replace the knowledge where the target is.

Acknowledgements

The machine time on SGI2800 in ACK-CYFRONET-AGH is financed by the Ministry of Science and Information Technology in Poland with grant No. KBN/SGI2800/AGH/018/2003.

References

[1] K.M. et al, IJMPC <u>14</u> (2003) 1201.

[2] K.M. et al, TASK Quarterly <u>8</u> (2004) 115.

[3] K.M. + K.K., `cond-mat/0304693`.

[4] K.M. + K.K., `cond-mat/0402474`.