TASK Quarterly 2 No 1 (1998), 85-95

Brief Reports on Current Research

HOW PHYSICISTS WASTE SUPERCOMPUTERS TIME IN ACADEMIC COMPUTER CENTER CYFRONET-KRAKÓW

KRZYSZTOF MALARZ, ANDRZEJ Z. MAKSYMOWICZ AND KRZYSZTOF KUŁAKOWSKI

Department of Theoretical and Computational Physics
Faculty of Physics and Nuclear Techniques,
University of Mining and Metallurgy,
Mickiewicza 30, PL-30-059 Cracow, Poland
tel. (+48 12) 6173539, fax: (+48 12) 6340010
malarz, amax, kulakowski@ftj.agh.edu.pl

Abstract: In this paper computer facilities for scientific research groups in Cracow are presented. Some problems solved by using supercomputers in Academic Computer Center Cyfronet-Kraków are discussed. The main flowcharts of computer algorithms and programs are included. General results are also presented.

1. Introduction

The Academic Computer Centre Cyfronet-Kraków (ACC CK) was established in 1973, as an independent non-profit organisation under the authority of the Ministry of Science, Higher Education and Technology (today the Ministry of National Education), to provide Cracow academic community and research institutes with powerful facilities for tasks which were beyond the scope of their own computer resources.

ACC CK delivers computer power combined with a wide range of application software to members of the academic and research staff and students of the Jagiellonian University, the University of Mining and Metallurgy, Cracow University of Technology, the Agricultural University, the Pedagogical University, the Economic University, the Collegium Medicum, the Academy of Physical Education, the Institute of Nuclear Physics and the Polish Academy of Science. Representatives of these institutions constitute a body of Scientific Committee of ACC CK and Users Committee for High Performance Computing and Metropolitan Area Network.

Cracow university community performs scientific and educational activities in many directions of pure and applied science. ACC CK serves specialists of practically all

disciplines: mathematicians, physicists, physicians, sociologists, chemists, astronomers, engineers, linguists, biologists, electronic engineers, and many, many others. Originally ACC CK was fitted out with a Cyber computer, which has for many years been the Centre's basic computing facility. Meanwhile needs were rocketing. In 1989, when Cyber was already outdated, the decision was made to buy a Convex computer. After two years of arduous attempts to obtain a US export licence, a C120 single-processor system arrived in ACC CK building. The Convex 120 was the first vector computer in Eastern Europe. It was followed a year later by a C3210, a more advanced model which the scientists of Cracow utilised to 90% of its total capacity. This provided an incentive for further computer purchases, which had to be planned at least a year in advance, as Poland was still subject to COCOM restrictions. In 1993 another Convex model, the C3820, reached Cracow. In 1994 the usage of CPU time on two computers C3220 (upgraded C3210) and C3820 was about 90%. The new need for a powerful computer was apparent. The Cracow academic community has made its choice for the first parallel computer from a number of vendors. The competitive offers were: CONVEX Exemplar SPP/XA-16, IBM 9076 SP2 with 16 processors, Silicon Graphics Power Challenge XL with 14 processors, Hewlett Packard cluster of eight 735CL processors. After due consideration made by a special commission the Exemplar SPP1000/XA-16 with 1.5 GB memory and 32 GB disk space was chosen. At the end of 1995 the Exemplar SPP1000/XA-16 was upgraded to SPP1200/XA-32 with 3.5 GB memory and 40 GB disk space. In November 1996 the Exemplar SPP1200/XA-32 was upgraded to SPP1600/XA-32 and C3820 was upgraded to C3840 by adding two processors. In September 1996 ACC CK obtained 5 processor parallel computer IBM RS/6000 SP.

Besides high performance computers ACC CK have got graphics workstations: SGI Indigo 2, HP VISUALIZE/K460-XP and workstations: HP 712/80, SGI IRIS 4D. Problems with the increasing needs of disk space were solved by installing Automated Tape Libraries ACL 4/52.

ACC CK is not only a high performance computer centre, but also the main node of Cracow Metropolitan Area Network (MAN). It delivers the capability of accessing world-wide internet resources to scientists, students, and many others from the area of Cracow and Southern Poland. In fact, the resources are open for all: high school pupils, commercial institutions, and others. The HP800/H50 is the main net server. Serving nearly 3000 users, it is the key mailbox in Cracow. The Sun Sparc 20 and HP 9000/80 K400 are the servers for many other network applications: Usenet, WWW, gopher, X500 among many others. ACC CK manages the Cracow MAN, which is built on 100Mb/s FDDI ring. More than 40 km of fibreoptic cable have already been laid. ACC CK links all the main educational and scientific institutions mentioned above, offering them its resources for their needs. Being eager to adapt new technologies, it experiments with new solutions, such as ATM (Asynchronous Transfer Mode).

The most part of computer and software resources, available in ACC CK, was financed by the State Committee for Scientific Research (KBN) The role of ACC CK

is not only limited to the above mentioned activities. ACC CK organises seminars, conferences and other activities that popularise new solutions in the world of networking and computing (information science)[1, 2, 3].

2. High performance computers in ACC Cyfronet-Kraków

Our calculations may be carried out on 3 supercomputers in ACC CK:

- vector computer Convex C3840 running under ConvexOS operating system, with four scalar-vector processors (peak performance 960 Mflops), 512 MB of memory and 22 GB of disk space
- massively parallel computer HP Exemplar SPP1600/XA-32 with operating system: SPP-UX based on HP-UX, with 32 processors HP PA-RISC 7200 (peak performance — 7.68 Gflops), 2.5 GB operating memory and 40 GB of disk space
- and parallel computer IBM RS6000/SP with 5 processors Power2 (peak performance 1.3 Gflops), AIX 4.2 operating system, 27 GB of disk space and 2.5 GB of operating memory

Automated tape libraries ACL 2640 together with dedicated server HP D350 (128 MB of operating memory and clock cycle — 100 MHz) have also been installed in ACC CK. This is an essential facility for all users who need large disk space for archiving. Library capacity is 2.64 TB, throughput — 13.5 GB/hr and recording medium is Compac Tape III [2, 3, 4].

3. Computational physics

As it was mentioned in the introduction, scientists in the fields of chemistry, physics and engineering are the main users of supercomputers in ACC CK. Among others, we use the presented above supercomputers and automated tape libraries to investigate some properties of thin and/or ultra-thin films. We are also involved in theoretical problems of classification cellular automata. In this section we present some examples of our activities and main results of numerical investigations

3.1 Computer simulation of thin film growth

3.1.1 Short theoretical background

The thin surface growth is a very common phenomenon in the world of microprocessors and supercomputers. It is particularly interesting if we rememmeber that materials which are used in our electronics are still in US dollars worth. Computer simulations in this field are needed because satisfactory and general enough theory of surface growth does not exist. In the last 20 years many models of film growth were proposed [5], based both on stochastic Langevin differential equations [6] and Monte Carlo

simulations [7]. The first approach was connected with continuous models. On the other hand, the computer simulations are adequate to discrete description of growth.

We consider the anisotropic growth of thin films. In the first approximation we assume the model of random deposition. In this model a particle simply falls down until it reaches the top of column where it is dropped or until it reaches a substrate. Then the particle freezes for the rest of simulation. This trivial case gives good opportunity to test proposed algorithm and computer implementation because the heights of surface on two dimensional square lattice follow the well-known in mathematics and physics Poisson distribution with deviation σ^2 equal average $<h>$.

More realistic model includes particle local migration to the nearest sites. The probability of such movement is determined by Boltzmann factor (which is very common in Monte Carlo simulations in the field of physics) and depends on total particle energy E in unit of $k_B T$. The total energy E has two independent components: binding energy J describing particle-particle forces and energetical diffusion barrier V correlated with a movement resistance. Additionally we assume that values of J and V could be different in different directions, which allows to consider an anisotropic surface growth.

3.1.2 Basic flowchart of the algorithm

The simulation took place on two dimensional square lattice (500 × 500) and 16 layers were deposited. The proposed algorithm is not too much CPU time consuming but for reasonable behaviour of system a large enough lattice is essential. Additionally, we assume periodic boundary conditions in order to minimalise boundary effects. For our simulation carried on HP-Convex supercomputers about 64 MB is necessary for memory and disk, both in order to get only one point in our plots! If we make, for example, about 200 independent runnings, also place for storing compressed binary files starts to be a problem and we have to use tape library archive. In the proposed algorithm:

1. Declare all variables
2. Import initial data from command line and/or from file
2. Allocate memory, initialise tables and variables
3. Prepare substrate of fixed $n \times m$ bottom layer atoms
4. Pick randomly the hit (x, y) site on the $n \times m$ lattice
5. Calculate local energy at the site $[x, y, h(x, y)]$, and its vicinity
6. Draw the coin and decide where the atom freezes
7. Repeat steps 5, 6 and 7 until all atoms have been deposited, then go to step 9
8. Export results to screen and/or output files the step 6 seems to be an excellent point for a parallelization

3.1.3 Main results

We easily find out that for negative values of J — when there are attractive forces between deposited particles — flat surface is produced (see Figure 1); on the contrary, for positive J spiky and roughy surface appears (see Fugure 2).

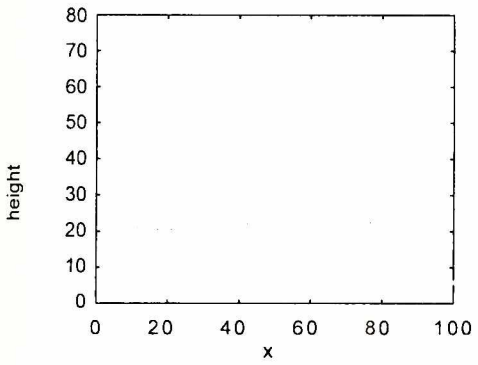

Figure 1. *Film heights profile for large negative value of J*

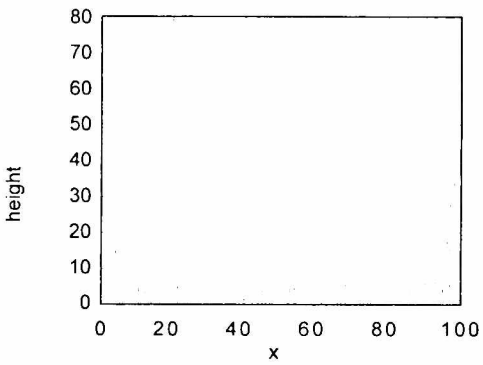

Figure 2. *Film heights profile for big positive value of J*

The spatial correlation function $F(\vec{s})$ of the local film thickness $h(\vec{r})$ is very useful in investigations of film properties [8, 9, 10].

$$F(\vec{s}) \equiv \frac{\langle h(\vec{r}) \cdot h(\vec{r} \cdot \vec{s}) \rangle - \langle h(\vec{r}) \rangle^2}{\langle h^2(\vec{r}) \rangle - \langle h(\vec{r}) \rangle^2},$$

The difference $e = F(\hat{x}) - F(\hat{y})$ is zero for isotropic case, and non zero when anisotropic growth takes place. The sign of ellipticity parameter e shows a tendency of surface to generate terraces-like structures ($e < 0$) or, on the contrary, spiky surfaces for positive value of e. Further investigation of surface morphology is available for vectors $\vec{s}$ other than only for perpendicular axis versors x and y. We check correlation function $F(\vec{s})$ for vectors $\vec{u} = [i, 0]$ and $\vec{v} = [0, i]$ when $i = 1 \ldots 5$. We find $F(\vec{u}) - F(\vec{v}) = 0$ for random deposition, as expected. When diffusion is accounted for, we expect decreasing correlations with increasing distances i. For anisotropic surfaces, we get different distance dependence of $F(\vec{s})$ along x-axis, $F(\vec{u})$ in Figure 3, and y-axis, $F(\vec{v})$ in Figure 4.

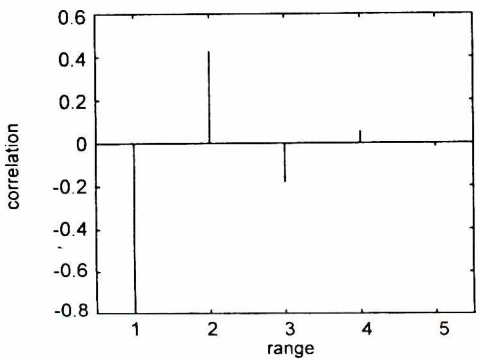

Figure 3. *Structural correlation along x-axis for J^x=-J^y=5 and V=-2*

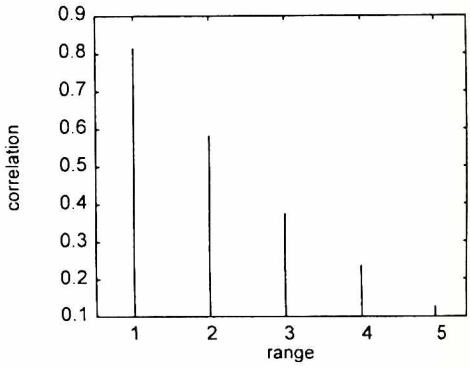

Figure 4. *Structural correlation along y-axis for J^x=-J^y=5 and V=-2*

3.2 Cellular automata classification

3.2.1 Theoretical background

Here, Conway's „*Game of Life*" [11]. For this automaton we prepare a 2-dimensional square lattice with randomly occupied sites. The time evolution of each site the system depends on situation on a 3 × 3 square of cells (the Moore neighbourhood). The rules of game for this automaton could be compared to a time evolution of live populations, where a creature dies when it is alone or has too many neighbours, but is created if 2 or 3 „parents" appear in the neighbourhood. Additionally, we provoke „gliders rain" at the center of the observed system, which makes the system open. One criteria of cellular automata classification arises from the question of *damage_spreading*. The question is how sensitive our system is with respect to small perturbations. In general we can distinguish three cases for the spreading of the damage during later iterations [12]:

1. The damage disappears (stable) — class I and II
2. The damage neither heals nor spreads (marginal stability) — class IV
3. The damage spreads over the whole lattice (unstable) — class III

The „*Game of Life*" is known to belong to the class IV.

3.2.2 Numerical results

To answer our question, two lattices are simulated at the same time, with the same rule and the same initial configuration. However, at one point of the initial lattice a damage appears (by switching free cell to occupied or occupied to free). The Hamming distance is the number of sites which differ the two lattices during time evolution. The main aim of this work is to show that „*Game of Life*" modified by frequently

touching gliders is unstable. The plot in Figure 5 shows growth of the Hamming distance as a function of time. We also find that the stream of gliders does not change the stationary density of „*life*" (for classical Conway's automata it was about 3%, here we are still below 4% — see Figure 6).

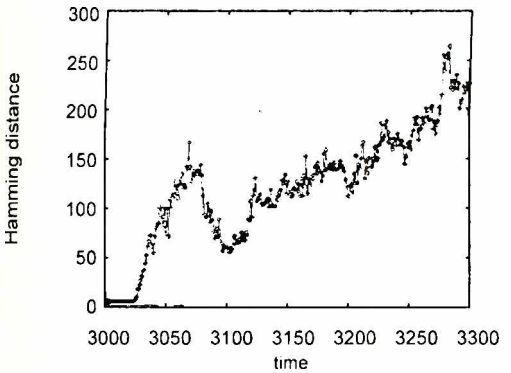

Figure 5. *Hamming distance as function of time*

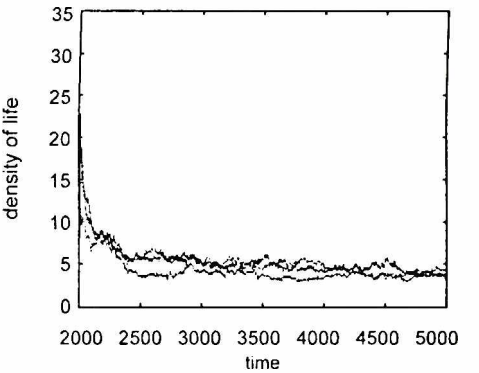

Figure 6. *Density of "living cells" for modified Conway's "Game of Life" cellular automaton and different initial densities*

These investigations are dedicated to parallel computing. They are also memory and time consuming. The reason of CPU time consuming is that at each time step of 3000 iterations for each cell of (1000 × 1000) large lattice we have to check rules of automata:

```
for(i = 1; i < max-1; i++)
for(j = 1; j < max-1; j++)
{
sum = 0;
if(tab_old[i - 1][j - 1]) sum++;
if(tab_old[i - 1][j    ]) sum++;
if(tab_old[i - 1][j + 1]) sum++;
if(tab_old[i    ][j - 1]) sum++;
if(tab_old[i    ][j + 1]) sum++;
if(tab_old[i + 1][j    ]) sum++;
if(tab_old[i + 1][j + 1]) sum++;
if((tab_old[i][j] && (sum==2 || sum==3))||
(!tab_old[i][j] && sum==3))
tab[i][j] = OCCUPIED;
else
tab[i][j] = FREE;
}
```

Our results [13] show that in an open system, the „*Game of Life*" automaton is chaotic — damages spread over the whole lattice. This automaton belongs simultaneously to both III-rd and IV-th class. Stationary state of „*Game of Life*" seems to be a self organized critical state [14].

3.3 Two-dimensional clusters properties investigation near percolation threshold

3.3.1 Short theoretical background

Percolation is simply a model of many and quite different physical phenomena such as: epidemics, forest fires under the influence of strong wind in one direction, kinetic gelation or rock corrosion, *etc* [15, 16].

Consider a square lattice, where each site is occupied randomly with probability p or empty with probability $1 - p$. Occupied and empty sites can — for instance — be adequate to electrical conductors and insulators, respectively, and that electrical current can flow only between nearest neighbours of conductor sites.

You can easily imagine that at low concentration p, the conductor sites are either isolated or form small islands. Two occupied sites belong to the same island (cluster) if they are connected by a solid line of nearest neighbour sites. Then the current can flow between them. At low p values, the mixture is an insulator, since a conducting path connecting opposite edges of our lattice does not exist. On the other hand, at large p values, many connections between opposite edges exist, and the mixture becomes a conductor.

We can find such concentration p_c that we will observe that our lattice is an insulator for p below p_c and it is conducting above p_c. The threshold concentration is called the *percolation threshold*. We can also assume that p_c separates two different phases. In this sense the percolation transition is a simple example of a phase transition phenomenon (such as the solid/liquid transition or conductivity/superconductivity or paramagnetic/ferromagnetic — where in all three examples changes take place at some *critical* temperature T_c).

The percolation transition is characterised by the geometrical properties of the cluster near p_c. For example, the linear size of the *finite* clusters, below and above p_c, is characterised by the *correlation length* ξ. When p approaches p_c, ξ increases as:

$$\xi \propto |p - p_c|^{-\nu},$$

with exponent ν for two dimensional lattice equal 4/3 below and above the threshold [15, 16, 17].

$$\begin{cases} \xi^{<} \propto (p_c - p)^{-\nu} \Leftrightarrow p < p_c, \\ \xi^{>} \propto (p - p_c)^{-\nu} \Leftrightarrow p > p_c. \end{cases}$$

The other important quantity is the probability p_∞ that a site belongs to infinite cluster. For $p < p_c$ only finite clusters exist, and $p_\infty = 0$. In other cases, when $p > p_c$, p_∞ increases with p by a power low:

$$p_\infty \propto (p - p_c)^{\beta},$$

and in our case the exponent $\beta = 5/36$ [15].

The exponents β and ν describe the critical behaviour of typical quantities associated with the clusters properties for $p \approx p_c$. They are called *critical exponents* and are universal (depend only on the dimension d of the lattice).

3.3.2 Numerical algorithm

There are several numerical approaches which allow determining that sites belong to the same cluster. In our case it is the Hoshen-Kopelman method [18]. In this algorithm all sites in the percolation system are labelled in such a way that the sites with the same label belong to the same cluster and different labels are assigned to different clusters. The Hoshen-Kopelman algorithm can be used to determine p_c and to generate the infinite percolation cluster in percolation, but for these purposes nowadays more efficient algorithms are available [19, 20].

3.3.3 Main results

Our main aim in this percolation phenomenon investigation is to find also universal ratio $\chi = \xi^{>} / \xi^{<}$ where $\xi^{>}$ and $\xi^{<}$ are connectivity lengths ξ for p above and below p_c at the same small distance ε. The quotient χ — as it is described by theory — is equal 0.5 [21] but previous Monte Carlo simulation showed it near 0.25 [22]. We consider a connectivity function $g(x)$ which describes probability that a site at distance x to first arbitrarily chosen and fully occupied edge is still connected to this edge. This function is described as:

$$\left(g(x) - p_\infty\right) \propto \begin{cases} \exp(-x/\xi) \Leftrightarrow p > p_c \wedge p_\infty = 0, \\ \exp(-x/\xi) \Leftrightarrow p > p_c \wedge p_\infty \neq 0. \end{cases}$$

and it allows to find χ easily. The $g(x)$ function is presented in Figures 7 and 8.

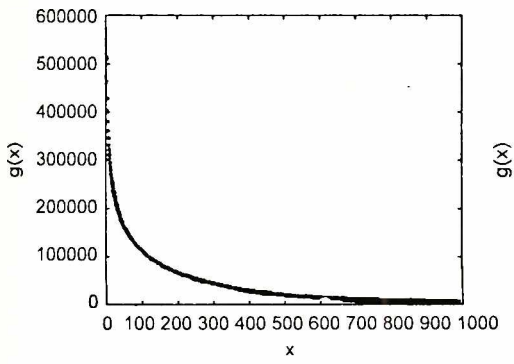

Figure 7. *Function g(x) for $p<p_c$ ($p_\infty=0$)*

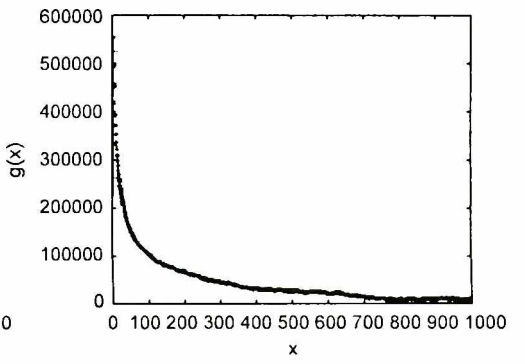

Figure 8. *Function g(x) for $p<p_c$ ($p_\infty\neq0$)*

Our first [23] results show that a definition of the ξ via the exponential decay yields quite different results than the previous one via the integral of the connectivity length. This could be the reason for the earlier discrepancy in determinig χ between mathematical theories and Monte Carlo simulations.

4. Concluding remarks

While writing this papear new HP Exemplar S2000 apears in ACC CK. This massively parallel computer has 16 PA 8000 processors and 4 GB of operating memory. We believe that we are able to find some new exciting problems which allow us to make this supercomputer always busy. We still intend to waste a supercomputer time in Academic Computer Centre Cyfronet-Kraków.

Acknowledgements

We are grateful to Dietrich Stauffer for his guidance to authors and helpful discussions during his one-week-long stay at our Department. KM thanks the staff of ACC CK for their patience and help. The information about ACC CK were browsed from ACC CK www service. The numerical calculations were carried out on computers in Academic Computer Center Cyfronet-Cracow. The time on computers is financed by Polish Committee for Research (KBN) with grants: KBN/C3840/AGH/008/1997, KBN/SPP/AGH/110/1997, KBN/IBMSP/AGH/114/1997.

References

[1] Wielgus M., Kwaśniewski M., http://puma.cyf-kr.edu.pl/BIULETYN/ENGLISH/index.html, 1996

[2] Księżyk M., Mosurska., and Noga M.. In *Proc. Of Conf. On High Performance Computing on Hewlett-Packard Systems*, pages 107-118, Cracow, November 1997

[3] Kwaśniewski N., Wielgus M., *Akademickie Centrum Komputerowe Cyfronet-Kraków 1997/98*, Akademickie Centrum Komputerowe Cyfronet-Kraków, 1997

[4] http://puma.cyf-kr.edu.pl/BIULETYN/ENGLISH/noble2.html, 1996

[5] Levi A. C. and Kotrla M.,*J. Phys. Cond. Matt.*, 9(2):299-344, 1997

[6] Predota M. and Kotrla M.,*Phys. Rev.* E, 54:4, 1-9, 1996

[7] Kotrla M., *Czechoslovak J. Phys.*, 42:449-544 1992

[8] Maksymowicz A. Z., Magdoń M. S., Whiting J. S. S., *Comp. Phys. Comm.*, 97:101-105, 1996

[9] Maksymowicz A. Z., Malarz K., Magdoń M. S.,Thompson S., Whiting J. S. S., In *Proc. of Conf. on High Performance Computing on Hewlett-Packard Systems*, 211-218, Cracow, November 1997

[10] Maksymowicz A. Z., Malarz K., Nagel G., Magdoń M. S., In *Proc. of Conf. on High Performance Computing on Hewlett-Packard Systems*, 227-234, Cracow, November 1997

[11] Berlekamp E. R. and ConwayJ. H, *Winning Ways for Your Mathematical Plays*, vol. 2, Academic Press, London 1982

[12] Wolfram S., *Rev. Mod. Phys*, 55:601, 1983

[13] Malarz K., Kułakowski K., Antoniuk M., Grodecki M., Stauffer P., *Int. J. Mod. Phys. C* – in print

[14] Flyvbjerg H., *Phys. Rev. Lett.*, 76:940, 1996

[15] Bunde A., Havlin S., *Fractals and disordered systems*, Springer-Verlag, 1991

[16] Sahimi M., *Applications of Percolation Theory*, Taylor and Francis, London, 1994

[17] Stauffer D., Aharony A., *Introduction to Percolation Theory*, Taylor and Francis, London, 1994

[18] Hoshen J., Kopelman R., *Phys. Rev.* B, 14:3428, 1976

[19] Leath P. L., *Phys. Rev.* B, 14:5046, 1976

[20] Ziff R. M., *J. Stat. Phys.*, 28:838, 1982

[21] Chayes J. T., Chayes L., Grimmett G. R., Kesten H., Schonmann R. H., *Ann. Prob.*, 17:1277, 1989

[22] Corsten M., Jan N., Jerrard R., *Phys. A*, 156:781-794, 1989

[23] Malarz K., Vidales A. M., Int. J. Mod. Phys. C9 (1998) 147-155