Invasion percolation: new algorithms and universality classes

## LETTER TO THE EDITOR

# Invasion percolation: new algorithms and universality classes

Adrian P Sheppard†‡, Mark A Knackstedt†‡, W V Pinczewski‡ and
Muhammad Sahimi§

† Department of Applied Mathematics, Research School of Physical Sciences and Engineering,
Australian National University, Canberra ACT 0200, Australia
‡ Australian Petroleum Cooperative Research Centre, University of New South Wales, Sydney,
NSW 2052, Australia
§ Department of Chemical Engineering, University of Southern California, Los Angeles, CA
90089-1211, USA

**Abstract.** Employing highly efficient algorithms for simulating invasion percolation (IP), whose execution time scales as $O[M \log(M)]$ or better for a cluster of $M$ sites, and for determining the backbone of the cluster, we obtain precise estimates for the fractal dimensions of the sample-spanning cluster, the backbone, and the minimal path in order to identify the universality classes of four different IP processes (site and bond IP, with and without trapping). In two dimensions IP is characterized by *two* universality classes, one each for IP without trapping, and site and bond IP with trapping. In a three-dimensional site IP with and without trapping is in the universality class of random percolation, while bond IP with trapping is in a distinct universality class, which may be the same as that of optimal paths in strongly disordered media.

Multiphase flow phenomena in porous media are relevant to many problems of great scientific and industrial importance, including extraction of oil, gas and geothermal energy from underground reservoirs, food and soil sciences, powder technology and materials science [1]. Invasion percolation (IP), a model introduced [2] for describing the evolution of the interface between an invading and a defending fluid in a porous medium, has provided deep insight into such phenomena. In addition, IP is relevant to a host of other problems, including characterization of optimal paths and domain walls in strongly disordered media [3, 4], and even simulation of the Ising model at the critical temperature [5]. Moreover, IP is one of the simplest parameter-free models which exhibits self-organized criticality [6], another subject of current interest.

Two main variants of IP have been studied so far. In nontrapping IP (NTIP) the defending fluid is compressible and the invading fluid can, potentially, enter any region on the interface which is occupied by the defending fluid. In the second and more common case, the defending fluid is incompressible and is trapped if a portion of it is surrounded by the invading fluid. In addition to the compressibility, one must also take into account the capacity of the fluids to wet the internal surface of the medium [1]. The process by which a wetting fluid is drawn spontaneously into a porous medium is called imbibition, while the forcing of a nonwetting fluid into the pore space is called drainage. We model the porous medium as a network of pores or sites connected by throats or bonds which have smaller radii than the pores. In IP, the potential displacement events are ranked according to the capillary pressure threshold that must be exceeded before that event takes place. During imbibition, the invading fluid is drawn

first into the smallest constrictions, for which the capillary pressure is large and negative, and it goes last into the widest pores. Displacement events are therefore ranked in terms of the largest opening that the invading fluid must travel through, since it is from these larger capillaries that it is most difficult to displace the defender. Imbibition is therefore a *site* IP and, in contrast, drainage in which the invader has most difficulty with the smallest constrictions is a *bond* IP.

Important differences arise in the structure of the invading fluid paths depending on whether one considers NTIP or TIP [2, 7, 8]. The question of the universality class of IP has not conclusively been established, and the literature contains contradictory claims about it. The scaling properties of NTIP are believed to be consistent with random percolation (RP). For trapping IP (TIP) the fractal dimension $D_f$ of the sample-spanning cluster (SSC) in two dimensions is smaller than that of RP [2]. In three dimensions no significant difference between TIP and RP has been reported. It was originally argued [2] that the fractal properties of IP do not depend on whether one simulates a site or bond IP. Recently, however, it was argued [9] that important differences arise in the structure of the invading fluid's paths when comparing site and bond IP. Porto *et al* [4] used a mapping from the minimal (shortest) paths in TIP to the optimal paths in strongly disordered media, and presented numerical evidence that for TIP the fractal dimension $D_{\min}$ of the minimal path in all dimensions is not the same as that of RP, and hence they argued that TIP and RP do not belong to the same universality class. Barabási [10] argued that loopless bond TIP is in the universality class of RP.

The reason that the universality class of IP has not conclusively been established is the relatively limited accuracy of the numerical results reported so far. This is due to the difficulty of simulating IP models, especially TIP, with very large lattices, since for *all* the present IP algorithms the simulation time grows as O($NM$), where $N$ is the number of sites (or bonds) in the lattice and $M$ the number of sites in the cluster. This has severely hampered studies of IP models. Since the differences between values of various fractal dimensions and scaling exponents of IP and RP appear to be small, it is critical to be able to use very large lattices to establish the universality classes of IP models. The purpose of this letter is twofold. First, we describe a highly efficient algorithm for simulating IP in which the simulation time grows as O[$M \log(M)$], which enables us to use very large lattices. Second, using the new algorithm we measure the fractal properties of various IP models with an accuracy which is at least an order of magnitude better than the best previous estimates. This enables us to test various hypotheses and conclusively establish the true universality classes of the IP models.

In the conventional algorithms [1, 2, 4, 8] the search for the trapped regions is done after every invasion event using a Hoshen–Kopelman algorithm [7, 8], which traverses the whole lattice, labels all the connected regions, and then only those sites (bonds) that are connected to the outlet face are considered as potential invasion sites (bonds). A second sweep of the lattice is then done to determine which of the potential sites is to be invaded in the next time step. Thus, each invasion event demands O($N^2$) calculations. This is highly inefficient for two reasons. First, after each invasion event only a small local change is made to the interface; implementing the global Hoshen–Kopelman search is unnecessary. Secondly, it is wasteful to traverse the entire network at each step to find the most favourable site on the interface since the interface is largely static. We tackle the first problem by searching the neighbours of each newly invaded site to check for trapping. This is ruled out in almost all instances. If trapping is possible, then several simultaneous breadth-first 'forest-fire' searches are used to update the cluster labelling as necessary [11]. This restricts the changes to the most local region possible. Since each site can be invaded or trapped at most once during an invasion, this part of the algorithm scales as O($N$). This cluster searching method has some similarities with the 'perimeter scouting' algorithm for two-dimensional (2D) clusters [12, 13]. In this algorithm one checks whether the most recently invaded sites could have been trapped in the interior of

the cluster. If so, oriented walks are started on the just invaded site, pointing away from it to the neighbouring sites which are those that neither belong to the cluster nor are candidates for invasion. The walks continue until all but one of them have again reached the site of origin. The growth sites visited by these walks are then eliminated from the list of active sites. This method is effective in two dimensions but not as efficient in three dimensions. Our method differs from it by searching cluster volumes rather than perimeters, and incorporating local checking to minimize cluster searching and is thus equally effective in three dimensions.

The second problem is solved by storing the sites (bonds) on the fluid–fluid interface in a list, sorted according to the capillary pressure threshold needed to invade them. This list is implemented via a balanced binary search tree, so that insertion and deletion operations on the list can be performed in $\log(n)$ time, where $n$ is the list size. Sites (bonds) that are designated trapped using the above-described procedures are removed from the invasion list. Each site is added and removed from the interface list at most once, limiting the cost of this part of the algorithm to $O[N \log(n)]$. Thus, the execution time for $N$ sites (bonds) is dominated (for large $N$) by list manipulation and scales at most as $O[N \log(N)]$.

While the execution time is approximately $O[N \log(n)]$, in practice the time and memory requirements depend on the total number of lattice sites and those forming the cluster boundary. For example, we find empirically that for 3D TIP the execution time scales as $M^{1.24}$, and the memory use is 20 bytes for each lattice site plus 64 bytes for each cluster site. On a 500 MHz 21164A Alpha microprocessor, a trapping cluster of $2 \times 10^5$ sites is grown on a $181 \times 181 \times 181$ lattice in 12.0 s, using 120 Mb of memory, while in two dimensions a cluster of $5 \times 10^5$ sites is grown on a $2000 \times 2000$ lattice in 12.0 s, using only 52 Mb. Complete details of the algorithm, which can be used for arbitrary networks, are given elsewhere [14].

We have also used a new optimized algorithm to identify the minimal path length, the sites comprising both the elastic backbone [15], i.e., the set of the sites that lie on the union of all the shortest paths between two widely separated points, and the usual transport backbone, so that the backbone search and computations do not affect the overall execution time of the algorithm. In the past, numerous algorithms have been proposed in the literature [15–18], most of which are too slow or limited to 2D systems. For example, a recent method [19] that uses a matching algorithm takes longer to identify the backbone than the percolation algorithm used here takes to generate it.

An alternative method was recently presented by Babalievski [11], based on depth-first searching out from the elastic backbone [15] to identify loops of occupied sites. This method works well for low connectivity clusters but loses efficiency where the SSC is composed of large well-connected regions. In the latter case, some sites need to be visited numerous times before their status is decided. The method used here is an optimization of this in which the distance on the cluster from the inlet face to each cluster site is used to guide the depth-first search. In this algorithm, there are three major steps which are as follows.

(i) Using a breadth-first search algorithm, we label each site in the cluster with its 'cluster distance' from the inlet face, and then use this information to burn backwards from the outlet face and identify the elastic backbone. At the same time, we construct the 'branch points list'—a list of all the cluster sites that are adjacent to the elastic backbone but are not part of it. The branch points list should be ordered with the sites closest to the inlet face listed first. Note that the elastic backbone sites are part of the backbone.

(ii) We stop if the branch points list is empty. Otherwise, we perform a depth-first search from the last site in the branch points list, flagging all the sites that are visited. During the search, unexplored branch points are added to the branch points list, while another list tracks the sites that have been flagged as visited. We then perform an important optimization during the depth-first search: if there are multiple branches from a single site, the site labelled as

being closest to the inlet face is always the first to be explored.

(iii) The depth-first search terminates when one of two conditions are satisfied: (1) the search contacts the backbone again at a different site from whence it started, in which case the sites in the visited-sites list are flagged as backbone sites; or (2) it retreats back to its starting site, at which point there will be no sites left in the visited-sites list.

(iv) The algorithm continues at step (ii).

In this way the elastic backbone, the transport backbone and the dangling ends of the SSC are all identified.

Examples of execution times for this algorithm running in three dimensions on a 533 MHz 21164A Alpha processor are 0.02 and 0.12 CPU s for $32^3$ and $64^3$ lattices, respectively. The cluster on which these calculations were performed was a SSC generated by a NTIP. When compared with the timings reported in [19] on an equivalent hardware, our algorithm is faster by a factor of 7 for the $32^3$ lattice and by a factor of 12 for a $128^3$ lattice, and thus the larger the lattice size, the more efficient is this algorithm. Comparison with the other algorithms [15–17] is also favourable.

We used $L^{d-1} \times 2L$ lattices in $d$ dimensions with reflecting boundary conditions on the edges. Cluster properties were measured within the central $L^d$ region. We also considered lattices of size $L^{d-1} \times nL$ in $d$ dimensions, where $n = 4, 8, 16$, and measured the cluster properties within the central $L^d$ region to check for end effects. We also studied the effect of periodic versus reflecting boundary conditions. The effect on the results was negligible. The simulations reported in this letter used up to $16\,384^2$ and $512^3$ lattices in two and three dimensions, respectively, larger than the largest previously-used lattices by a factor of 5–10. At least 2000 realizations were used for the largest lattices and over $10^5$ realizations for smaller $L$.

We first consider the fractal dimension $D_f$ of the SSC. If we define a local fractal dimensions $D_f(M) \equiv \mathrm{d} \ln M / \mathrm{d} \ln L$, then according to finite-size scaling (FSS) $D_f(M)$ will converge to the asymptotic value for large $M$ according to, $|D_f - D_f(M)| \sim M^{-\alpha}$, where $\alpha$ is a correction-to-scaling exponent to be estimated from the data. Combining the definition of $D_f(M)$ with the FSS equation gives us a differential equation which has an analytical solution:

$$c_1 + D_f M^\alpha = c_2 L^{\alpha D_f} \tag{1}$$

where $c_1$ and $c_2$ are constants. We then fit the data to equation (1) to estimate *both $D_f$ and $\alpha$ simultaneously*. By so doing we also avoid statistical pitfalls of the two-stage process used in [13] in which the data are first divided into various bins, the local $D_f(M)$ are estimated by numerical differentiation of $\ln M$ with respect to $\ln L$, and then $D_f(M)$ is plotted versus $M^{-\alpha}$ and $\alpha$ is varied until the the 'best' straight line is obtained. In addition, we also obtain reliable estimates for the confidence intervals of $\alpha$ and the fractal dimensions (see below). From this analysis we obtained, $D_f \simeq 1.8949 \pm 0.0009$ ($\alpha \simeq 0.475$), completely consistent with the exact value, $D_f = \frac{91}{48} = 1.895\,83$ for RP. For TIP we obtained, $D_f \simeq 1.825 \pm 0.004$, more than an order of magnitude more precise than the previous estimates [2–4]. In three dimensions we found that, as far as $D_f$ is concerned, there is no significant difference between NTIP and TIP, hence confirming explicitly that trapping occurs rarely in three dimensions; we found $D_f \simeq 2.528 \pm 0.002$, consistent with the most recent estimate of $D_f$ for RP [20], $D_f \simeq 2.523 \pm 0.004$. Moreover, no difference was found between values of $D_f$ for site and bond TIP.

However, the universality classes of IP are not determined solely by $D_f$. The fractal properties of the backbone and the minimal paths, which are particularly important for flow and transport processes, must also be considered. Our simulations indicate that, as a consequence of the trapping and the choice of site or bond IP, strong differences exist between the backbone

**Table 1.** Fractal dimensions for invasion percolation in two dimensions. Numbers in parenthesis are the estimates obtained from local $D_f(M)$ analysis.

| Model | $D_{min}$ | $D_b$ |
|---|---|---|
| NTIP | $1.1293 \pm 0.0010$ | $1.6422 \pm 0.0040$ |
| Site TIP | $1.214 \pm 0.001$ | $1.217 \pm 0.020$ |
| Bond TIP | $1.2170 \pm 0.0007$ | $1.217 \pm 0.001$ |
| RP | $1.1307 \pm 0.0004$ | $1.6432 \pm 0.0008$ |
| Optimal paths | $1.21 \pm 0.01$ | $1.21 \pm 0.01$ |

**Table 2.** Backbone and minimal path fractions in three dimensions.

| Site IP without trapping | | | |
|---|---|---|---|
| $L$ | Runs | $X_{min}$ | $X_b$ |
| 32 | 100 000 | 0.002 909(2) | 0.012 71(2) |
| 45 | 75 000 | 0.001 670(2) | 0.008 62(2) |
| 64 | 50 000 | 0.000 938(2) | 0.005 756(2) |
| 90 | 30 000 | 0.000 537 9(8) | 0.003 91(2) |
| 128 | 10 000 | 0.000 303 8(8) | 0.002 61(2) |
| 256 | 5 000 | 0.000 098 3(3) | 0.001 21(1) |
| 362 | 5 000 | 0.000 055 8(8) | 0.000 79(1) |
| 512 | 2 000 | 0.000 032(1) | 0.000 53(1) |

| Site TIP | | | |
|---|---|---|---|
| $L$ | Runs | $X_{min}$ | $X_b$ |
| 32 | 100 000 | 0.002 913(2) | 0.012 73(3) |
| 45 | 75 000 | 0.001 670(2) | 0.008 61(2) |
| 64 | 50 000 | 0.000 939(1) | 0.005 77(2) |
| 90 | 30 000 | 0.000 537 9(8) | 0.003 87(2) |
| 128 | 10 000 | 0.000 304 2(8) | 0.002 62(2) |
| 256 | 5 000 | 0.000 098 2(3) | 0.001 20(1) |
| 362 | 5 000 | 0.000 055 7(8) | 0.000 79(1) |
| 512 | 2 000 | 0.000 032(1) | 0.000 53(1) |

| Bond TIP | | |
|---|---|---|
| $L$ | Runs | $X_{min} = X_b$ |
| 32 | 100 000 | 0.001 446(1) |
| 45 | 75 000 | 0.000 844(1) |
| 64 | 50 000 | 0.000 483(1) |
| 90 | 30 000 | 0.000 284(1) |
| 128 | 10 000 | 0.000 163(1) |
| 256 | 5 000 | 0.000 056 5(3) |
| 362 | 5 000 | 0.000 032 9 |

and the minimal path structures. For site TIP *local* trapping thwarts extensive growth of the backbone off the minimal path, leading to a greatly diminished backbone. For bond TIP the backbone *coincides* with the minimal path, which has also been reported previously [3, 4, 9]; see table 1 for the 2D results. The estimated fractal dimensions indicate that the scaling properties of NTIP are consistent with those of 2D RP [21]: $D_{min} = 1.1307 \pm 0.0004$ and $D_b = 1.6432 \pm 0.0008$. The estimate for $D_{min}$ for bond TIP is consistent with the fractal dimension $D_{op}$ for the optimal path in the limit of strong disorder [3]. The estimates for site TIP do *not* agree with *any* estimates for RP.
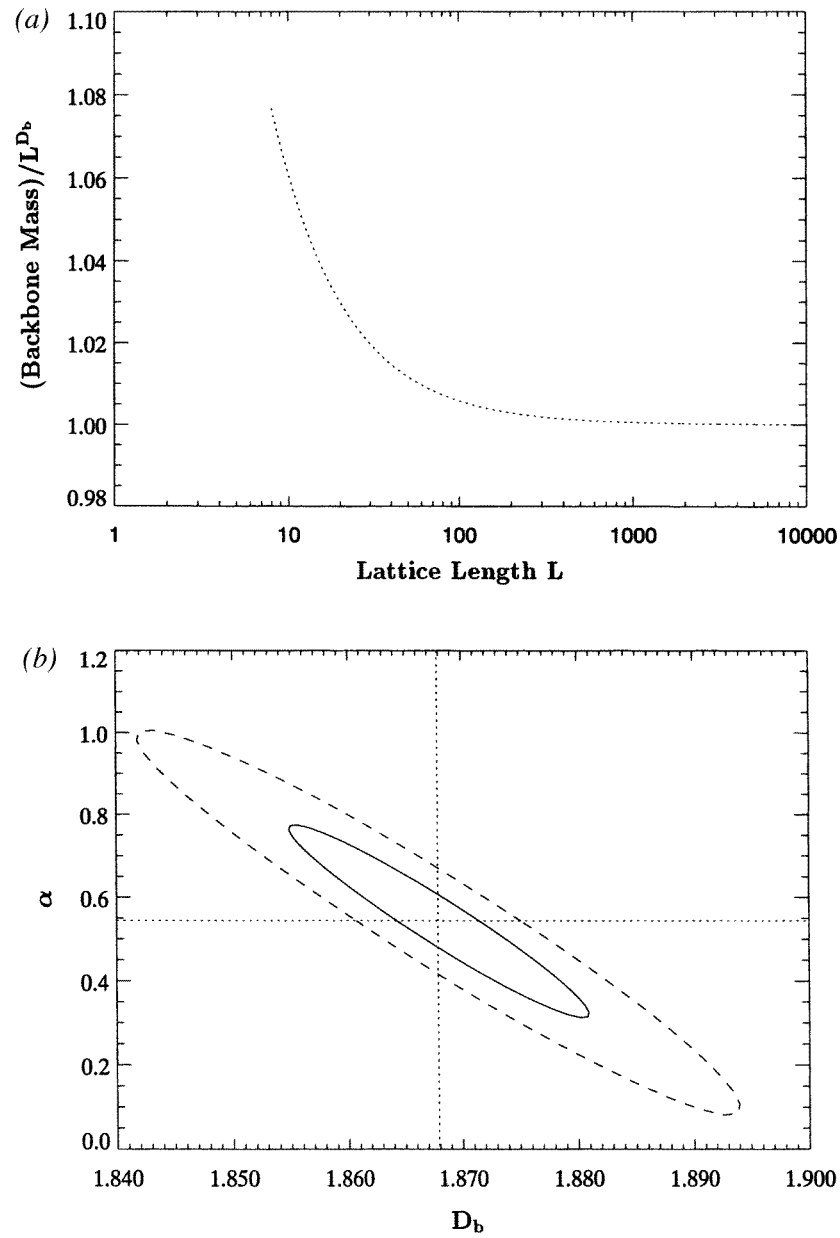
**Figure 1.** (*a*) Scaling of (mass of the backbone)/$L^{D_b}$ versus $L$ for 3D site NTIP, where $D_b$ is estimated from equation (1). (*b*) Variations of $\alpha$ with $D_b$ in the confidence region ellipses corresponding to 90% (solid) and 99% (dashed) confidence intervals.

In three dimensions several thousand realizations for each $L$ were required to obtain estimates of sufficient accuracy. As shown in table 2, the backbone and minimal path fractions are very different for bond TIP. Figure 1 presents the results for the backbone of site NTIP. The upper figure shows how the ratio of the mass of the backbone and $L^{D_b}$ approaches unity as $L$ becomes large, where $D_b$ is the estimate from equation (1). The data and the fit are
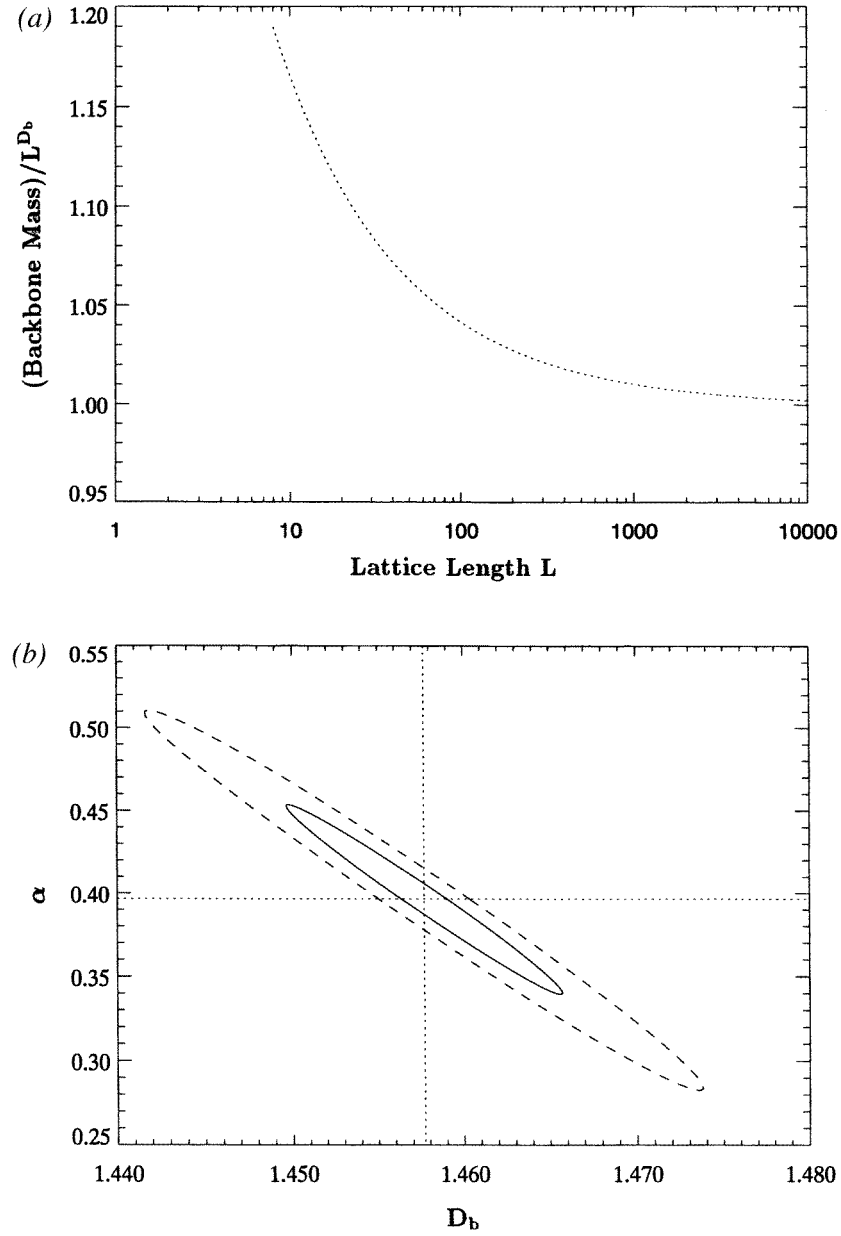
**Figure 2.** Same as in figure 1, but for the backbone of 3D bond TIP.

indistinguishable, confirming the accuracy of the estimated $D_b$. The lower figure shows the variations of $D_b$ with $\alpha_b$ in the confidence region ellipse [22]. As can be seen, requiring 99% confidence limit (the dashed ellipse) also requires larger error bars. Figure 2 presents the same but for the backbone of bond TIP, while table 3 summarizes all the estimates in 3D. For site IP our estimate of $D_{min}$ is in agreement with that of RP [18, 19, 21], while the estimate of $D_b$ is consistent with, but more than an order of magnitude more precise than, the recent estimate [19], $D_b \simeq 1.86 \pm 0.01$, for RP. All the results for site IP are consistent with those of RP,

**Table 3.** Fractal dimensions for the minimal path and the backbone of invasion percolation in three dimensions.

| Model | $D_{min}$ | $D_b$ |
|---|---|---|
| Site NTIP | $1.3697 \pm 0.0005$ | $1.868 \pm 0.010$ |
| Site TIP | $1.3697 \pm 0.0005$ | $1.861 \pm 0.005$ |
| Bond TIP | $1.458 \pm 0.008$ | $1.458 \pm 0.008$ |
| RP | $1.374 \pm 0.004$ | $1.87 \pm 0.03$ |
| Optimal path | $1.43 \pm 0.01$ | $1.42 \pm 0.02$ |

and, therefore, the two models belong to the same universality class. Unlike two dimensions though, the value of $D_{min}$ for bond TIP is not in agreement with the fractal dimension $D_{op}$ of the optimal path in an energy landscape in the limit of strong disorder [3, 4], although this may be due to the limited accuracy of $D_{op}$. This result demonstrates explicitly that the fractal dimensions associated with the transport paths (the backbone and the minimal paths) for bond TIP are distinct from those of RP. Therefore, while the SSC has a fractal dimension $D_f$ consistent with RP, the fractal dimensions associated with the transport paths are close to that of optimal paths in energy landscapes in the limit of strong disorder.

Summarizing, using a highly efficient algorithm, we have shown that 2D IP is characterized by *two* universality classes, one for NTIP which is the same as that of RP, and another for bond and site TIP. In three dimensions, one has *two* distinct universality classes, one for site NTIP and TIP which is the same as that of RP, and one for bond TIP. Thus, while RP, a static process, is described by a unique universality class, IP, a dynamical phenomenon, does not possess a unique universality class. This has important implications for modelling multiphase flow through porous media: while drainage paths are essentially optimal routes in disordered porous media, those for imbibition are very different and *not* optimal. Moreover, since TIP describes waterflooding processes in secondary oil recovery, the difference in the topology of the transport paths for the bond IP (nonwetting fluid invasion) and site IP (wetting fluid invasion) will have a profound effect on the conductivity of the invading phase at the breakthrough point where a SSC forms for the first time. This difference has been noted in core-scale measurements of the conductivity of waterflooded rock under different wetting conditions. For an oil-wet rock the conductivity of the water channels can be several orders of magnitude smaller at breakthrough of the invading water phase than for a comparable water-wet rock [23], consistent with the different topology of the flow paths noted in this letter.

## References

[1] Sahimi M 1993 *Rev. Mod. Phys.* **65** 1393
Sahimi M 1995 *Flow and Transport in Porous Media and Fractured Rock* (Weinheim: VCH)
[2] Chandler R, Koplik J, Lerman K and Willemsen J F 1982 *J. Fluid Mech.* **119** 249
Wilkinson D and Willemsen J 1983 *J. Phys. A: Math. Gen.* **16** 3365
[3] Cieplak M, Maritan A and Banavar J R 1994 *Phys. Rev. Lett.* **72** 2320
Cieplak M, Maritan A and Banavar J R 1996 *Phys. Rev. Lett.* **76** 3754
[4] Porto M, Havlin S, Schwarzer S and Bunde A 1997 *Phys. Rev. Lett.* **79** 4060
Porto M, Bunde A, Havlin S and Roman H E 1997 *Phys. Rev.* E **56** 1667

 [5] Machta J, Choi Y S, Lucke A and Schweizer T 1995 *Phys. Rev. Lett.* **75** 2792
     Franzese G, Cataudella V and Coniglio A 1998 *Phys. Rev.* E **57** 88
 [6] Stark C P 1991 *Nature* **352** 423
     Cafiero R, Gabrielli A, Marsili M and Pietronero L 1996 *Phys. Rev.* E **54** 1406
 [7] Stauffer D and Aharony A 1992 *Introduction to Percolation Theory* 2nd edn (London: Taylor and Francis)
     Bunde A and Havlin S 1996 *Fractals and Disordered Systems* 2nd edn (Berlin: Springer)
 [8] Sahimi M 1994 *Applications of Percolation Theory* (London: Taylor and Francis)
 [9] Sahimi M, Hashemi M and Ghassemzadeh J 1998 *Physica* A **260** 231
[10] Barabási A-L 1996 *Phys. Rev. Lett.* **76** 3750
[11] Babalievski F 1998 *Int. J. Mod. Phys.* C **9** 43
[12] Meakin P 1991 *Physica* A **173** 305
[13] Schwarzer S, Havlin S and Bunde A 1999 *Phys. Rev.* E **59** 3262
     Porto M, Bunde A, Havlin S and Roman H E 1997 *Phys. Rev.* E **56** 1667
     Porto M, Havlin S, Roman H E and Bunde A 1998 *Phys. Rev.* E **58** 5205
[14] Knackstedt M A, Sahimi M and Sheppard A P *Phys. Rev.* E to be published
[15] Herrmann H J, Hong D C and Stanley H E 1984 *J. Phys. A: Math. Gen.* **17** L261
[16] Liem C and Jan N 1988 *J. Phys. A: Math. Gen.* **21** L243
[17] Rintoul M D and Nakanishi H 1994 *J. Phys. A: Math. Gen.* **27** 5445
[18] Grassberger P 1999 *Physica* A **262** 251
[19] Moukarzel C 1998 *Int. J. Mod. Phys.* C **9** 887
[20] Lorenz C D and Ziff R M 1998 *Phys. Rev.* E **57** 230
[21] Grassberger P 1992 *J. Phys. A: Math. Gen.* **25** 5477
     Grassberger P 1992 *J. Phys. A: Math. Gen.* **25** 5867
[22] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes* 2nd edn (Cambridge:
     Cambridge University Press) ch 15
[23] Abdassah D, Permadi P and Sumantri Y 1996 *Society of Petroleum Engineers Paper 36978* (Richardson, TX:
     Society of Petroleum Engineers)