

Empires and percolation: stochastic merging of adjacent regions

D J Aldous[‡], J R Ong and W Zhou

Department of Statistics

367 Evans Hall # 3860

U.C. Berkeley CA 94720

E-mail: aldous@stat.berkeley.edu

www.stat.berkeley.edu/users/aldous

Abstract. We introduce a stochastic model in which adjacent planar regions A, B merge at some stochastic rate $\lambda(A, B)$, and observe analogies with the well-studied topics of mean-field coagulation and of bond percolation. Do infinite regions appear in finite time? We give a simple condition on λ for this *hegemony* property to hold, and another simple condition for it to not hold, but there is a large gap between these conditions, which includes the case $\lambda(A, B) \equiv 1$. For this case, a non-rigorous analytic argument and simulations suggest hegemony.

AMS classification scheme numbers: 82C21, 60K35

1. Introduction

We study random processes defined as follows.

- (i) at each time t the plane is partitioned into polygonal regions
- (ii) as t increases, adjacent regions A, B merge into one region $A \cup B$ at some stochastic rate (probability per unit time) $\lambda(A, B)$.

Here $\lambda(A, B)$ is a specified function of the geometry of regions A and B , typically a simple formula involving quantities such as the areas $\text{area}(A)$, $\text{area}(B)$, the perimeters (boundary lengths) $\text{peri}(A)$, $\text{peri}(B)$ and the length $L(A, B)$ of the boundary between A and B . The *adjacency* condition is enforced by the assumptions

$$\text{if } L(A, B) > 0 \text{ then } \lambda(A, B) > 0; \quad \text{if } L(A, B) = 0 \text{ then } \lambda(A, B) = 0.$$

This is intended as an abstract model for spatial growth via merging – we mentally picture countries merging into empires. Despite its conceptual simplicity, this model has apparently never been studied, so let us name it the *empire process*. The purpose of this article is to initiate its study by setting out some simple analytic and simulation results, describing connections with other models, and providing suggestions for further study and methodology.

1.1. Analogy with mean-field coalescence

Consider moving particles of different masses x (we take the continuous case $0 < x < \infty$), where particles of masses x and y may coalesce when they meet. Ignoring geometry (positions and motion of particles) and simply assuming there is a "relative propensity to meet and coalesce" function $K(x, y)$ which applies to all pairs of particles, leads one to formulate a "mean-field" model. Write $f(x, t)dx$ for the density per unit volume at time t of particles with mass in $[x, x + dx]$. The density function $f(x, t)$ satisfies the well-known *Smoluchowski coagulation equation* [1, 2, 3]

$$\frac{d}{dt}f(x, t) = \frac{1}{2} \int_0^x K(y, x-y)f(y, t)f(x-y, t)dy - f(x, t) \int_0^\infty K(x, y)f(y, t)dy. \quad (1)$$

In the empire process the conserved quantity is area, so there is an analogous density $f(a, t)$ of area- a regions per unit area. But – even when the merger rate function is of the form $\lambda(A, B) = K(\text{area}(A), \text{area}(B))$ – we cannot write down autonomous equations analogous to (1) for $\frac{d}{dt}f(a, t)$, because spatial relationships, and the stochastic nature of mergers, matter in our model.

Versions (now called *stochastic coalescents*) with finite total mass and stochastic merging but no geometry have been studied [2, 4]. And there is recent progress [5] in rigorous verification of the underlying presumption that in models of spatial diffusion and merging, the limiting (low density of massive particles) behaviour is as predicted by the Smoluchowski coagulation equation. Our empire process model is just different, but our general theme of relating qualitative properties of the process to properties of the rate function is obviously parallel to the same theme in mean-field coalescence.

1.2. Analogy with bond percolation

To each edge e of the square lattice assign a random time T_e , with Exponential(1) distribution $P(T_e \leq t) = 1 - e^{-t}$, at which the edge becomes “open”. The configuration of open edges at time t is just the usual bond percolation [6, 7] process with $p = 1 - e^{-t}$. A connected component of m open edges can be identified, as in Figure 1, with a region of area m in the dual square grid.

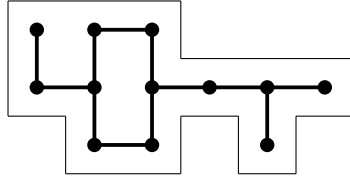


Figure 1. The region associated with a component in bond percolation.

The number of edges between two adjacent components of the bond percolation process equals the boundary length between the two associated dual regions. In the bond percolation process as t increases to $t + dt$, each closed edge has chance dt to become open, so the bond percolation components merge at rate “number of closed edges linking the components”, and so in the dual process two regions A, B merge at rate $L(A, B)$. So we have shown

In the empire process with $\lambda(A, B) = L(A, B)$, started at $t = 0$ with the unit squares of the square grid, the areas of regions at time t are distributed as the numbers of edges of the connected components of bond percolation with $p(t) = 1 - e^{-t}$.

The celebrated result [6, 7] that the critical value in bond percolation equals $1/2$ implies that, in the particular empire process above, infinite regions appear at time $\log 2$.

2. Hegemonic or not?

Whether or not the qualitative property

$$\text{infinite-area regions appear in finite time} \tag{2}$$

holds is perhaps the most interesting question to ask about empire processes. In percolation theory the analogous property is described by phrases like “supercritical” or “percolation occurs”, while in coagulation theory it is called “gelation”, but to maintain the visualization of empires let us call (2) the *hegemonic* case. (See Notes below for details of interpretation of (2)). Can we give conditions, in terms of the merger rate function $\lambda(A, B)$, for whether the empire process is hegemonic or non-hegemonic? Here is one easy result.

Proposition 1 (i) Suppose (for some $c > 0$) we have

$$\lambda(A, B) \geq c L(A, B) \text{ for all } A, B.$$

Then the empire process is hegemonic.

(ii) Suppose (for some $C < \infty$) we have

$$\sum_j \lambda(A, B_j) \text{peri}(B_j) \leq C \text{peri}(A) \text{ for all } A \text{ with adjacent regions } (B_j).$$

Then the empire process is not hegemonic.

See (7) for an explicit example of (ii).

Before giving the simple proofs, we should point out a subtlety whose resolution will be useful. Consider two empire processes, where the second has faster rates than the first:

$$\lambda_1(A, B) \leq \lambda_2(A, B) \text{ for all } A, B. \quad (3)$$

The assertion

$$\begin{aligned} &\text{if the } \lambda_1 \text{ empire process is hegemonic and (3) holds then} \\ &\text{the } \lambda_2 \text{ empire process is hegemonic} \end{aligned} \quad (4)$$

seems plausible at first sight, but we suspect it is not always true (see Notes below). What is true? If we can couple (define simultaneously) the two processes in such a way that the natural refinement-coarsening relationship is maintained (each region of the first is a subset of a region of the second) then the conclusion “if the first process is hegemonic then the second process is also hegemonic” is certainly correct. And the natural condition which ensures the relation can be maintained is the following.

$$\sum_{i,j} \lambda_1(A_i, B_j) \leq \lambda_2(A, B), \quad \text{for all partitions } (A_i) \text{ of } A \text{ and } (B_j) \text{ of } B. \quad (5)$$

This condition includes the case $A_1 = A, B_1 = B$ and so (5) is a stronger assumption than (3).

To investigate when condition (5) might hold, consider a rate function with the *superadditive* property

$$\lambda_1(A_1, B) + \lambda_1(A_2, B) \leq \lambda_1(A_1 \cup A_2, B), \quad \text{for all } B \text{ and disjoint } (A_1, A_2). \quad (6)$$

This condition implies that (5) holds with λ_2 replaced by λ_1 . So if we assume both (3) and superadditivity of λ_1 then we have (5) and the desired implication “if the first process is hegemonic then the second process is also hegemonic” is correct.

Because the rate function $L(A, B)$ is superadditive (in fact, additive) we can immediately deduce part (i) of the Proposition (the constant c is just time-scaling) from the bond-percolation fact.

To prove (ii), fix an arbitrary reference point in the plane and consider the perimeter (X_t , say) of the region \mathcal{R}_t containing the reference point at time t . The definition of the empire process gives growth dynamics as follows, where (B_i) are the regions adjacent to \mathcal{R}_t .

$$E(dX_t | X_s, 0 \leq s \leq t) = \sum_i (\text{peri}(\mathcal{R}_t \cup B_i) - \text{peri}(\mathcal{R}_t)) \lambda(\mathcal{R}_t, B_i) dt$$

$$\begin{aligned}
&\leq \sum_i \text{peri}(B_i) \lambda(\mathcal{R}_t, B_i) dt \\
&\leq C \text{peri}(\mathcal{R}_t) dt \text{ by assumption (ii)} \\
&= C X_t dt.
\end{aligned}$$

So $\frac{d}{dt} EX_t \leq C X_t$ and so $EX_t \leq e^{Ct} EX_0$. So regions have finite mean size at all times.

An explicit example where (ii) applies is

$$\lambda(A, B) = \frac{L(A, B)}{\max(\text{peri}(A), \text{peri}(B))}. \quad (7)$$

Because in this case

$$\sum_j \lambda(A, B_j) \text{peri}(B_j) \leq \sum_j L(A, B_j) = \text{peri}(A).$$

Any merger rate λ' which is slower than this λ in sense (3) will also be non-hegemonic, because case (ii) will apply. But, to repeat the warning at (4), in general if a merger rate is slower in sense (3) than a known non-hegemonic merger rate, then we cannot immediately deduce that it is also non-hegemonic.

2.1. Notes on section 2

Details of the definition of empire process. It is natural to assume the initial configuration is statistically translation-invariant, so that it will remain so at all times. In the section 4 simulations we started with the square grid for simplicity, though for certain theoretical analyses (section 3) it is more convenient to use the hexagonal lattice, and as a continuum model it is perhaps conceptually more natural to start with a non-lattice configuration such as the Voronoi tessellation on random points.

Details of the definition of hegemonic. Property (2) is imprecise in two ways, which we mention here without seeking a more precise analysis. First, we are presuming that the choice of initial configuration does not affect (2). Second, we are presuming that (2) has probability zero or one. Both presumptions are plausible by comparison with standard percolation results.

Limitations of the coupling methodology. One might hope that the coupling methodology used to prove (i) might be useful more widely to prove hegemony under different assumptions. But, by considering small squares on either side of a boundary line, one sees that any superadditive rate function λ must satisfy the hypothesis of (i). Since we need superadditivity to use the coupling, one cannot handle any essentially different cases this way.

Limitations of the methodology for (ii). The result in (ii) remains true if we replace “perimeter” by “area” or indeed any subadditive functional, but we don’t know any interesting cases that can be handled by this variant and not by the stated form of (ii).

Why might the implication (4) fail? This very sketchy outline indicates why (4) is not obvious. Consider a hegemonic empire process with rate λ_1 , starting with the square grid. Write \mathcal{R}_t for the region containing a reference point. Take a sequence m_1, m_2, \dots such that (with probability one) only finitely many values m_j are seen as values of $\text{area}(\mathcal{R}_t)$. Then we should be able to modify λ_1 by redefining $\lambda_1(A, B)$ to be small whenever $\text{area}(A)$ or $\text{area}(B)$ equals some m_j , preserving the hegemonic property and the “only finitely many values m_j are seen” property. Now take λ_2 as the modification of λ_1 in which rates $\lambda_2(A, B)$ are made very large when $\text{area}(A) + \text{area}(B) = \text{some } m_j$. This modification should force the sequence $\text{area}(\mathcal{R}_t)$ (for the λ_2 empire process) to re-enter the sequence (m_j) infinitely often and (having slow merger rates therein) to be non-hegemonic.

3. Analysis of the case $\lambda = 1$

In this section we present a non-rigorous, though convincing, analytic argument that the constant rate case ($\lambda(A, B) = 1$ for adjacent A, B) is hegemonic, and this is supported by simulations – see section 4.

Note that a special feature of this case is that any initial boundary line will still be a piece of some boundary line at time t with probability exactly e^{-t} . However, since these events for different initial lines are dependent, it is not clear how to exploit this formula directly.

It is convenient (the reason is explained below) to take the initial configuration to be the hexagonal lattice, scaled so that boundary edges have length 1. Consider a circuit \mathcal{C} which encloses a reference point b . Write \mathcal{R}_t for the boundary of the region of the empire process containing b at time t . As with the standard Peierls contour method [6, 7] in percolation, to prove hegemonic it is enough to prove that

$$\int_0^\infty \sum_{\mathcal{C}} P(\mathcal{R}_t = \mathcal{C}) dt < \infty$$

where the sum is over all circuits enclosing b . The length of a circuit must be even and ≥ 6 : write $2n$ for the length. The number of possible circuits enclosing b of length $2n$ can be bounded as order 2^{2n} without taking into account their self-avoiding property; but taking this into account reduces the bound to order $2^{(2-\delta)n}$ for some $\delta > 0$ (see [8] for discussion of the value of δ , which is not important for our calculation). So if $p_{2n}(t)$ is an upper bound of the probability that any particular contour of length $2n$ is the boundary of \mathcal{R}_t , then it is enough to show

$$\sum_n 2^{(2-\delta)n} \int_0^\infty p_{2n}(t) dt < \infty. \quad (8)$$

To study $p_{2n}(t)$ consider Figure 2, which shows part of a circuit \mathcal{C} which is present at time t , and shows the regions whose boundaries include part of \mathcal{C} : shown are interior regions A, B, C, D and exterior regions W, X, Y, Z .

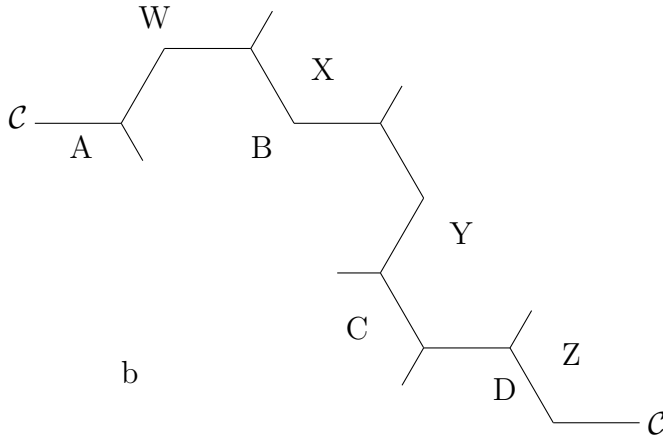


Figure 2. Part of a contour \mathcal{C} enclosing a point b , illustrating regions adjacent to the contour.

This picture can change in two ways at time increases.

- (i) Two regions that are adjacent along the circuit – either exterior regions such as X, Y or interior regions such as B, C – may merge into one. This preserves the circuit and (typically) decreases the number of regions adjacent to the circuit by 1.
- (ii) An interior and exterior region – such as B, X – may merge, destroying the circuit. Note this is where it is convenient to start with the hexagonal lattice – the same position on \mathcal{C} cannot separate two exterior regions and two interior regions, so the number of number of adjacent interior-exterior region pairs equals $i + j$, provided neither i nor j equals 1.

We continue the analysis assuming the “typical” behaviour above always occurs – see Discussion below. Because each possible merger occurs at rate 1, the process (I_t, J_t) where

$$I_t = \text{number of exterior regions adjacent to the circuit}$$

$$J_t = \text{number of interior regions adjacent to the circuit}$$

behaves as the continuous-time Markov chain with transition rates

$$\begin{aligned} (i, j) &\rightarrow (i - 1, j) : \text{rate } i \\ &\rightarrow (i, j - 1) : \text{rate } j \\ &\rightarrow \text{destroyed} : \text{rate } i + j. \end{aligned}$$

started at time 0 with $(I_0, J_0) = (n + 3, n - 3)$. Strictly, these rates only apply when $i, j \geq 2$, but allowing “fictitious” transitions $1 \rightarrow 0$ simplifies the analysis (see Discussion below). We can compare the Markov chain (I_t, J_t) to the chain (I_t^*, J_t^*) defined in the same way but without the “destroyed” possibility. Because in the process (I_t, J_t) the next transition has chance exactly 1/2 to be “destroyed”, we find

$$P(I_t = i, J_t = j) = 2^{i+j-2n} P(I_t^* = i, J_t^* = j). \tag{9}$$

We are interested in $p_{2n}(t) = P(J_t = 1 \text{ or } 0)$ (the 0 captures the fictitious transition), and combining (9) and (8) we see that we need to prove

$$\sum_n 2^{-\delta n} \sum_{i=1}^{n-3} 2^i \int_0^\infty P_{n+3,n-3}(I_t^* = i, J_t^* = j \text{ or } 0) dt < \infty \quad (10)$$

where we wrote $P_{n+3,n-3}(\cdot)$ as a reminder that the chain (I_t^*, J_t^*) starts at $(n+3, n-3)$.

Now if the chain (I_t^*, J_t^*) ever hits state (i, j) then it remains in that state for mean time $1/(i+j)$, so we can write

$$\int_0^\infty P_{n+3,n-3}(I_t^* = i, J_t^* = j \text{ or } 0) dt = (i+1)^{-1} q_{n+3,n-3}(i, 1) + i^{-1} q_{n+3,n-3}(i, 0)$$

where $q_{n+3,n-3}(i, j)$ is the chance that the embedded discrete-time jump chain (\hat{X}_s, \hat{Y}_s) ever hits state (i, j) . But this jump chain is simply the chain obtained by drawing without replacement from a box which initially has $n+3$ balls labeled "exterior" and $n-3$ balls labeled "interior", and writing (\hat{X}_s, \hat{Y}_s) for the number of remaining balls of each type after s draws. In particular, by considering reversed order of draws

$$q_{n+3,n-3}(i, 1) = (i+1) \frac{n-3}{2n} \frac{(n+3)_i}{(2n-1)_i} \leq C(i+1)2^{-i}$$

for some constant C , with a smaller similar bound for $q_{n+3,n-3}(i, 0)$. We now see the sum in (10) is bounded by order n and thus the outer sum is indeed finite.

3.1. Discussion of approximations above.

What is assumed in setting up the rates for the continuous-time Markov chain is that the j successive exterior (and similarly for interior) regions touching the circuit \mathcal{C}

- (a) are distinct
- (b) meet only if they are consecutive pairs along \mathcal{C}
- (c) $j \geq 3$.

Under these assumptions, the combined merger rate is indeed j . If instead of (c) we have $j = 2$ then the merger rate is 1, and if $j = 1$ the rate is 0; however it is not hard to see that these changed rates for $j = 1, 2$ don't affect the "finiteness" result (8).

But the other simplifying assumption are more significant. If in Figure 2 the regions W and Y are adjacent, by meeting behind region X , then the combined merger rate is larger than j ; moreover if they do merge then (a) fails for the merged region. The resulting "combinatorial explosion" of possibilities for successive regions along the boundary being distinct or identical seems very hard to analyze rigorously.

4. Simulation results

We will show simulation results for the following four rate functions.

Model 1: $\lambda(A, B) = 1$.

Model 2: $\lambda(A, B) = \text{area}(A) \times \text{area}(B)$.

Model 3: $\lambda(A, B) = L(A, B)$.

Model 4: $\lambda(A, B) = 1/(\text{area}(A) \times \text{area}(B))$.

The quantities we calculated were the spatial averages

$$S(t) = \text{average (area of empire)}^2 \text{ per unit area}$$

$$D(t) = \text{average number of empires per unit area.}$$

Note the alternative interpretations: the average area of an empire (for uniform random choice of empire) equals $1/D(t)$, while the average area of the empire containing a random point (i.e. the average when empires are chosen with probability proportional to area) equals $S(t)$. We started at $t = 0$ with the square grid, so $S(0) = D(0) = 1$; as t increases, $S(t)$ increases and $D(t)$ decreases. In order to compare different models it is convenient to plot the curve $(D(t), S(t))$, regarded as a function $S(D)$. This is a natural way to summarize the non-hegemonic phase. If the model is hegemonic then, as D decreases from 1, $S(D)$ will increase to infinity at some critical value $D_{\text{crit}} > 0$, whereas for a non-hegemonic model $D_{\text{crit}} = 0$.

There is a different way to summarize the hegemonic phase. Restrict the model to a large finite region of area A . Then as D decreases from 1 to 0, the quantity $S(D)/A$ increases from $1/A$ to 1, and $S(D)/A$ equals the chance that two random points in the region are in the same empire. In the $A \rightarrow \infty$ limit, the function $S(D)/A$ will become a function $f(D)$ analogous to the *percolation function* in percolation theory, with $f(D) = 0$ for $D > D_{\text{crit}}$ and $f(D) > 0$ for $D < D_{\text{crit}}$.

These two ways to think about D_{crit} are shown in Figure 3 for the four models. The upper graph in each pair shows $S(D)$, and the lower graph shows $S(D)/A$, for two models. The points D_{crit} are where (in the $A \rightarrow \infty$ limit) the upper graph goes to infinity and where the lower graph leaves zero. Data from simulations on an 81×81 grid.

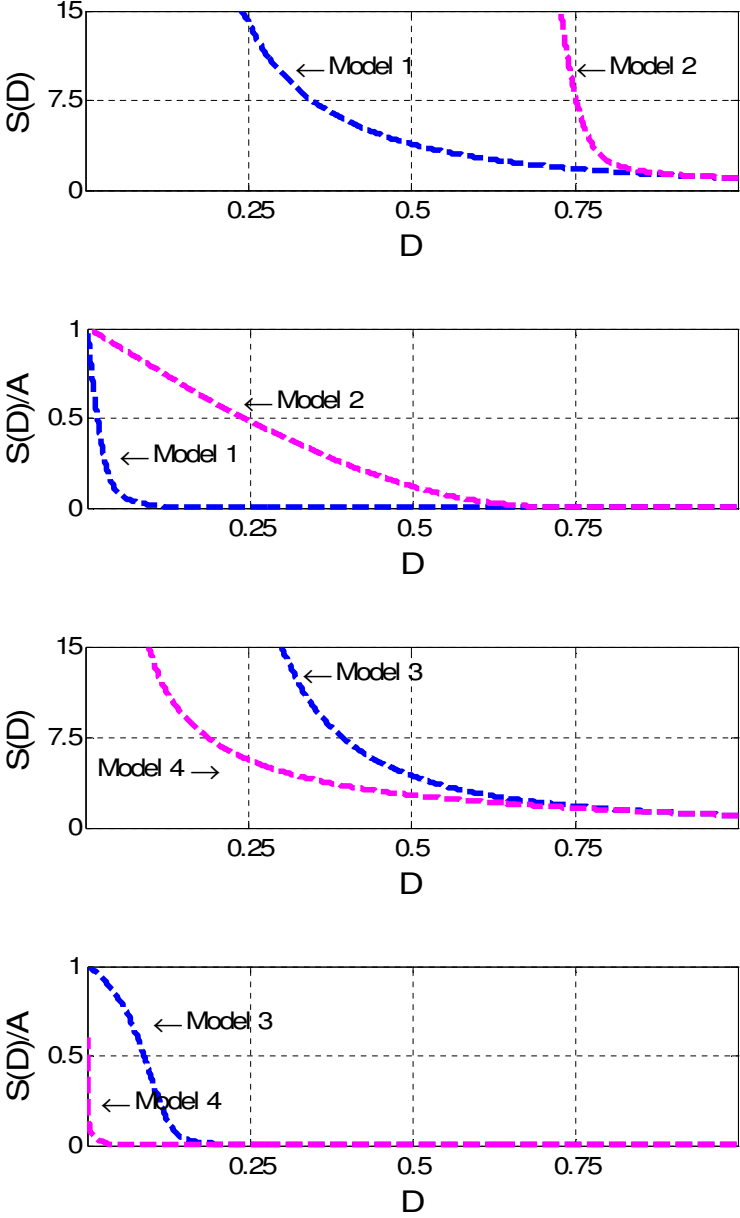


Figure 3. Simulation results for 4 models, explained in text.

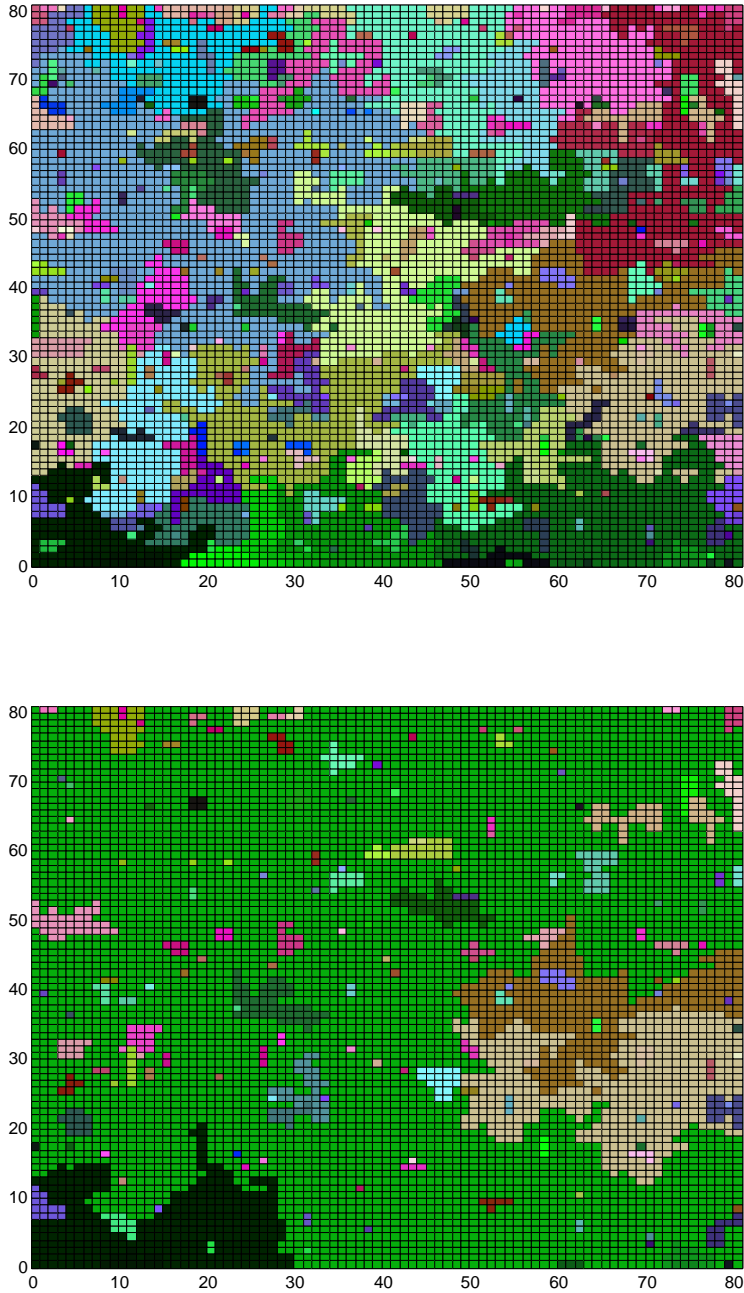


Figure 4. The emergence of a giant component, in the case $\lambda = 1$.

The data from Figure 3 suggests the following rough numerical estimates, which we then compare to theory. The scaling exponents γ will be discussed below.

Model 1: [$\gamma = 0$]: $D_{\text{crit}} \approx 0.15$. This is consistent with the non-rigorous argument for hegemony ($D_{\text{crit}} > 0$).

Model 2: [$\gamma = 2$]: $D_{\text{crit}} \approx 0.6$. In this case Proposition 1(a) applies, so we know $D_{\text{crit}} > 0$.

Model 3: [$\gamma = 1/2$]: $D_{\text{crit}} \approx 0.2$. This is the “bond percolation” case, so we know $D_{\text{crit}} > 0$.

Model 4: [$\gamma = -2$]: $D_{\text{crit}} \approx 0$. In this case Proposition 1(b) applies, so we know $D_{\text{crit}} = 0$.

Figure 4 shows emergence of the giant component in the case $\lambda = 1$. The visual appearance is rather different from that of near-critical bond percolation, in that the large components appear less “fractal” and that distinct moderately large components coexist over longer time periods.

Scaling exponents. In the context of mean-field coalescence, a kernel such that $K(cx, cy) = c^\gamma K(x, y)$ is said to have scaling exponent γ . It has long been understood, mostly non-rigorously (but see [9] for references to recent rigorous work) that the coalescence process should be non-gelling if $\gamma \leq 1$ but gelling if $\gamma > 1$.

We can define a scaling exponent γ analogously for the empire process: $\lambda(cA, cB) = c^{2\gamma}\lambda(A, B)$, where cA denotes linear scaling by a factor c . The data above is consistent with the possibility that there is some critical value γ_{crit} such that empire processes are typically hegemonic for $\gamma > \gamma_{\text{crit}}$ and are typically non-hegemonic for $\gamma < \gamma_{\text{crit}}$. However, the rate function at (7) which is known to be non-hegemonic has $\gamma = 0$, as does the $\lambda = 1$ case presumed to be hegemonic, so it may be that scaling exponents are not so definitive for empire processes.

References

- [1] Drake RL 1972 A general mathematical survey of the coagulation equation *Topics in Current Aerosol Research (Part 2) (International Reviews in Aerosol Physics and Chemistry vol 3)* ed G M Hidy and J R Brock pages (Oxford: Pergamon) pp 201–376
- [2] Aldous DJ 1999 Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists *Bernoulli* **5** pp 3–48
- [3] Norris JR 1999 Smoluchowski’s coagulation equation: uniqueness, non-uniqueness and a hydrodynamic limit for the stochastic coalescent *Ann. Appl. Probab.* **9** pp 78–109
- [4] Bertoin J 2006 *Random Fragmentation and Coagulation Processes (Cambridge Studies in Advanced Mathematics vol. 102)* (Cambridge: Cambridge University Press)
- [5] Hammond A and Rezakhanlou F 2007 Moment bounds for the Smoluchowski equation and their consequences. *Comm. Math. Phys.* **276** pp 645–70
- [6] Grimmett GR 1999 *Percolation (2nd Ed.)* (Berlin: Springer-Verlag)
- [7] Bollobás B and Riordan O 2006 *Percolation* (Cambridge: Cambridge University Press)
- [8] Alm SE and Parviainen R 2004 Bounds for the connective constant of the hexagonal lattice *J. Phys. A* **37** pp 549–60
- [9] Escobedo M and Mischler S 2006 Dust and self-similarity for the Smoluchowski coagulation equation *Ann. Inst. H. Poincaré Anal. Non Linéaire* **23** pp 331–62