Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 24 February 1997 (MN LATEX style file v1.3)

Coagulation and branching process models of gravitational clustering

Ravi K. Sheth[1] and Jim Pitman[2]
[1] *Astronomy Department, University of California, Berkeley, CA 94720*
[2] *Statistics Department, University of California, Berkeley, CA 94720*
Email: sheth@mpa-garching.mpg.de, pitman@stat.berkeley.edu

Submitted to MNRAS May 1996

ABSTRACT
Smoluchowski's coagulation equation describes the growth of clustering in a system in which clusters coalesce irreversibly by binary mergers. If the rate with which two objects merge is proportional to the sum of their masses, and the process starts with all clusters having the same mass initially (monodisperse initial conditions), then the growth of clustering in the system can be derived from a Poisson Galton–Watson branching process. The branching process is consistent with, and provides a more detailed description of gravitational clustering from Poisson initial conditions, than the excursion set, Press–Schechter approach. The correspondence between the branching process and additive coagulation schemes is used to efficiently simulate realizations of the merger history tree.

These results are easily extended to describe clustering from white noise Gaussian initial conditions. Extending this approach to describe clustering from Gaussian fields with arbitrary power spectra is more complicated. To illustrate this, Smoluchowski's additive coagulation equation with arbitrary discrete initial conditions is solved. The result is used to argue that if Smoluchowski's equation is to describe the growth of clustering from arbitrary Gaussian fluctuation fields, and if it is to be consistent with the excursion set results, then the merger rule must depend on the power spectrum.

However, with the exception of the white noise case, the excursion set approach is formally inconsistent with the binary merger assumption, although, for all power spectra of current interest, binary mergers remain a reasonable approximation. This means that, except for the white noise case, an analogue of the Smoluchowski coagulation approach cannot provide a description of clustering that is exactly equivalent to the excursion set description. Nevertheless, since binary mergers remain a good approximation, Smoluchowski's equation, with a merger rate that can be determined from the excursion set approach, so that it is a calculable function of the power spectrum, can provide a useful approximate description of the growth of clustering.

Key words: galaxies: clustering – cosmology: theory – dark matter.

1 INTRODUCTION

This is the third in a series of papers in which models based on the Poisson distribution are used to provide insight into the growth of gravitational clustering in an expanding Universe. This approach is useful, not so much because one thinks that the initial conditions for clustering in our Universe were Poisson, but because the Poisson assumption allows one to solve many problems that, at present, have no solution in the more realistic Gaussian case. Insofar as clustering from an initially Poisson distribution should resemble that from white-noise (Gaussian with power spectrum $P(k) \propto k^n$ and $n = 0$) initial conditions, the results of this approach are useful.

The growth of clustering can be formulated in terms of a Poisson Galton–Watson branching process, or its associated queue (Sheth 1996). This branching process reproduces known excursion set (Epstein 1983; Bond et al. 1991; Lacey & Cole 1993; Sheth 1995) and cloud-in-cloud type (Jedamzik 1995; Sheth 1995) results. However, in addition to these results, the branching process formulation yields an explicit, analytic expression for the reverse transition probability function of a Markovian merger history process, from which most quantities of interest can be obtained. That is to say, for any clump containing N particles, the branching process allows one to enumerate all possible merger histories of the clump, and to compute the probability that each occurs.

Many questions of astrophysical interest can be addressed using Monte–Carlo realizations of such merger histories (e.g. Kauffmann & White 1993). However, the branching process shows that the number of possible merger histories of an N-clump grows rapidly as N increases. Therefore, an efficient method for generating Monte-Carlo realizations of the clustering process is essential for addressing such questions. In this paper we show that the Poisson Galton–Watson branching process is equivalent to a 'sticky-gas' model of the type first formulated by Smoluchowski (1916, 1917). Only a few analytical solutions of Smoluchowski's coagulation equation are known. One of these, in which the merger cross-section is additive in mass, is equivalent to the branching process discussed above (Pitman 1996). A stochastic form of the Smoluchowski model with additive rates allows us to generate Monte-Carlo realizations of the clustering process efficiently.

Section 2 summarizes the relevant branching process results of Sheth (1996). Section 3 develops various explicit descriptions of the merger history process implied by the branching process model, following Pitman (1996). Section 4 uses the results of the previous sections to generate Monte-Carlo realizations of the merger history tree. It compares the ensemble of trees obtained in this way with those obtained using the algorithm developed by Kauffmann & White (1993) in their treatment of merger histories. The closely related Smoluchowski model with additive rates is discussed in Section 5. The Smoluchowski equation with additive rates and arbitrary discrete initial conditions is solved in Appendix B. Section 6 discusses the relation between the coagulation model studied in this paper and the excursion set approach of Bond et al. (1991) and Lacey & Cole (1993). It also discusses the possibility of extending our results to describe the growth of clustering from white-noise as well as more general Gaussian initial conditions. Appendix C shows that, if the coagulation and the excursion set descriptions are to yield the same description of clustering, then the merger rate law must depend on the initial conditions. Section 7 summarizes our results.

2 THE POISSON GALTON–WATSON BRANCHING PROCESS

This section summarizes the relevant results of Sheth (1996), where details of all the arguments are given.

The evolution of clustering from an initially Poisson distribution has been studied using the excursion set (Epstein 1983; Sheth 1995) and cloud-in-cloud formalisms (Sheth 1995). These analyses show that the probability that a clump has k associated particles is

$$\eta(k,b) = \frac{(kb)^{k-1}\mathrm{e}^{-kb}}{k!}, \tag{1}$$

where $k \geq 1$ and $0 \leq b \leq 1$. Equation (1) is known as the Borel distribution (Borel 1942). The probability $f(k,b)$ that a randomly chosen particle belongs to a k-clump is given by multiplying $\eta(k,b)$ by k, and dividing by the mean clump size $\sum_k k\,\eta(k,b) = 1/(1-b)$:

$$f(k,b) = k\,(1-b)\,\eta(k,b). \tag{2}$$

The evolution of clustering is described by the evolution of the b parameter, which is the ratio of the average density in the universe to the critical overdensity that is required for an object to collapse and virialize at cosmological time t. That is,

$$b(t) = \frac{1}{1+\delta_{\mathrm{c}}(t)} \qquad \text{where} \quad \delta_{\mathrm{c}}(t) \propto \frac{1.686}{a(t)}, \tag{3}$$

t denotes cosmological time, and $a(t)$ is the factor by which the universe has expanded since some fiducial initial epoch, say, t_i. Notice that the critical density decreases as the universe expands, so that b increases with expansion. The numerical value of 1.686, and the dependence of the density threshold on expansion factor can be derived by assuming that the collapse is spherical (e.g. Peebles 1980). The relation between δ_{c} and the expansion factor a shown above is correct only in a universe that has critical density (see Lacey & Cole 1993 for the relevant expression when $\Omega \neq 1$). Since b increases monotonically as cosmological time increases, it will be treated as a pseudo-time variable in the remainder of this paper.

The excursion set description also allows one to compute the probability that a randomly chosen particle of a clump having exactly k particles at the time corresponding to b_2, was a member of a j-clump at the earlier epoch $b_1 < b_2$. This probability is

$$\begin{aligned} f(j,b_1|k,b_2) \quad &= \quad k\left(1-\frac{b_1}{b_2}\right)\binom{k}{j}\frac{j^j}{k^k} \\ &\times \quad \left(\frac{b_1}{b_2}\right)^{j-1}\left[k - j\frac{b_1}{b_2}\right]^{k-j-1}, \end{aligned} \tag{4}$$

where $1 \leq j \leq k$ (Sheth 1995). This implies that the average number of progenitor j-subclumps identified at b_1 that are later (at $b_2 > b_1$) in a k-clump is $(k/j)\,f(j,b_1|k,b_2)$. Although the excursion set formalism provides information about the distribution of progenitor subclumps of any given parent clump, it does not provide a complete description of the entire merger history tree (e.g. Lacey & Cole 1993).

This is because one can construct many different algorithms which can be used to generate Monte–Carlo merger histories that are consistent with equation (4) (e.g. Kauffmann & White 1993). For example, each of these different algorithms may yield different predictions for the probability, to be denoted $n(m,b_1|k,b_2)$, that a clump with k particles at b_2 had exactly m progenitor subclumps at the epoch $b_1 < b_2$. The excursion set approach cannot answer the question of which algorithm is correct. Fig. 1 illustrates this problem.

To describe the merger history as a random process it is necessary to make a more elaborate probabilistic model consistent with formulae (1) and (4). The Poisson Galton–Watson branching process described in Sheth (1996) provides one such model. It is motivated by the observation that, since gravity is an attractive pairwise force, to a crude approximation, one may assume that near neighbours in the initial Poisson distribution clump first with their nearest neighbours, and then these clumps merge with their nearest neighbour clumps, and so on. This means that clustering grows by a friends-of-friends, percolation-like process,

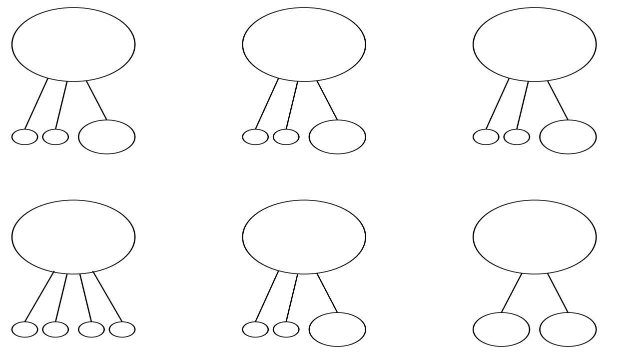

Figure 1. Examples of merger history trees constructed from the same progenitor distribution using two different merger history algorithms (top and bottom). The algorithm for the top level produces trees that are homogeneous; in comparison, the trees on the bottom, produced using a different algorithm, are all different from each other. The excursion set formalism cannot answer the question of which ensemble of trees is more likely, so it cannot discriminate between different algorithms for building merger history trees.

with a friends link-length that increases with time. This percolation-like process is similar to the model of organized spherical collapse envisioned by Press & Schechter (1974) and described most recently by Mo & White (1996). Unfortunately, there are no known analytical descriptions of the percolation merger tree. However, the Poisson Galton–Watson branching process is a simple approximation to the percolation process (see figs. 1 and 2 of Sheth 1996) for which analytical results can be obtained.

For instance, consider $n(m, b_1|k, b_2)$. In the branching process model this probability is given by a binomial distribution:

$$n(m, b_1|k, b_2) = \binom{k-1}{m-1} \left(\frac{b_1}{b_2}\right)^{k-m} \left(1 - \frac{b_1}{b_2}\right)^{m-1}. \qquad (5)$$

Let $p(n_1 \cdots n_k; b_1|k; b_2)$, with $\sum_{j=1}^{k} n_j = m$ and with $\sum_{j=1}^{k} j\, n_j = k$, denote the probability that a k-clump at the epoch b_2 was previously in m subclumps at the epoch b_1, and that there were n_1 single particle subclumps, n_2 subclumps with exactly two particles, n_j subclumps with exactly j particles, and so on. Then the branching process model yields the formula

$$p(n_1 \cdots n_k; b_1|k; b_2) = \frac{(kb_{21})^{m-1} \mathrm{e}^{-kb_{21}}}{\eta(k, b_2)} \prod_{j=1}^{k} \frac{\eta(j, b_1)^{n_j}}{n_j!}, \qquad (6)$$

where $\sum_{j=1}^{k} n_j = m$, $\sum_{j=1}^{k} j\, n_j = k$, and we have defined $b_{21} = (b_2 - b_1)$. We call this function $p(n_1 \cdots n_k; b_1|k; b_2)$ the *partition probability function*.

The partition probability function serves as the transition probability function in reverse time of a Markov process describing how a k-clump existing at time b_2 was broken up into subclumps at time $b_1 < b_2$. In principle, repeated application of this partition probability function determines the joint distribution of clump sizes and the merger history, at any finite number of time epochs prior to b_2, leading up to a k-clump at time b_2. Some equivalent and more intuitive descriptions of this merger history process are presented in the next section.

It is useful to re-write the partition probability function in terms of the distribution of the sum $S_m = X_1 + \cdots + X_m$ of a sequence of independent random variables X_i, each Borel-distributed with parameter b, which is given by

$$P(b, S_m = k) = \frac{m}{k} \frac{(kb)^{k-m} \mathrm{e}^{-kb}}{(k-m)!} \quad \text{where} \quad 1 \le m \le k, \qquad (7)$$

(Tanner 1953, 1961). Then

$$p(n_1, \cdots n_k; b_1|k; b_2) = \frac{n(m, b_1|k, b_2)}{P(b_1, S_m = k)}\, m! \prod_{j=1}^{k} \frac{\eta(j, b_1)^{n_j}}{n_j!}, \qquad (8)$$

where $n(m, b_1|k, b_2)$ is given by equation (5). The first term expresses the probability that there are m subclumps, given that there are k particles in total. The $m!$ and product terms have a simple combinatorial origin (see equation 11 and associated discussion in Spouge 1983b). The form of equation (8) shows clearly that our partition probability function solves a polymer physics problem of the type discussed by Flory (1941a,b,c) and Stockmayer (1943). See Spouge (1983b) for details.

This partition probability function depends only on the ratio b_1/b_2, and not on the actual values of b_1 or b_2. In the limit as $b_1 \to 1$ and $b_2 \to 1$, equation (3) shows that this translates into a dependence on $\delta_{c1} - \delta_{c2}$, rather than the actual values of δ_{c1} or δ_{c2} themselves. This is consistent with the excursion set results for Gaussian random fields (e.g. Lacey & Cole 1993). Moreover, in this limit ($\delta \ll 1$), Stirling's approximation for the factorials reduces all the Poisson excursion set results (e.g. equations 1 and 4) to the corresponding results for $n = 0$, white-noise Gaussian initial conditions. This dependence on the ratio b_1/b_2 implies that the merger histories of k-clumps formed at the epochs b_2 and b_2' are related to each other simply by rescaling the time variable. In other words, except for this rescaling of time, the distribution of merger history trees of k-clumps is independent of their environment.

By taking relevant sums over the partition probability function, one can show explicitly that the branching process expressions for the clump size distribution and the merger probabilities $f(j|k)$ are identical to the excursion set results (equations 1 and 4). However, the partition probability function also allows one to compute quantities that cannot be estimated using the excursion set formalism. For example, the branching process gives an expression for the number m of subclumps of a k-clump (equation 5). The excursion set formalism does not. Furthermore, recall (equation 4) that the mean number of j-subclumps per k-clump, $\langle n_j|k\rangle$, is $(k/j)\, f(j|k)$. Whereas the excursion set formulation only allows one to compute this mean value, the partition probability function allows one to compute the higher order moments of the n_j distribution also. For example, the factorial moments are

$$\left\langle \frac{n_j!}{(n_j - i)!} \,\middle|\, k \right\rangle = \frac{k!}{k^{k-1}} \left[k\left(1 - \frac{b_1}{b_2}\right)\right]^{i} \left[k - ij\frac{b_1}{b_2}\right]^{k-ij-1} \times \frac{1}{(k-ij)!} \left(\frac{j^{j-1}}{j!}\right)^{i} \left(\frac{b_1}{b_2}\right)^{ij-i}. \qquad (9)$$

When $i = 1$, this is equivalent to the relation obtained from equation (4). When $i = 2$, equation (9) gives an estimate of the scatter among different merger history trees. In Section 5 we use this estimate of the scatter to compare different merger tree algorithms.

The partition probability function (equation 6) is a useful quantity for computing some sums of astrophysical interest (e.g. equations 4, 5, and 9). However, many quantities of interest, such as the average of the ratio of the sizes of the two largest progenitor subclumps (e.g. Kauffmann & White 1993), cannot be computed exactly analytically (but see, e.g., Pavlov 1977 and Pitman 1996 for analytical estimates that are useful in certain limiting regimes). In principle, one can compute such sums numerically, since the partition probability function is known, provided one can enumerate all possible partitions of any given integer k efficiently. Unfortunately, the number of partitions of k grows extremely rapidly as k increases, so that it is impractical to compute such sums even for $k \sim 10^2$ or so. Therefore, one would like to have an efficient scheme for generating Monte–Carlo realizations of the merger history tree in such a way that the ensemble of trees so obtained is consistent with what is required by the branching process partition probability function. The next section provides such a prescription.

3 THE MERGER HISTORY PROCESS DERIVED FROM THE BRANCHING PROCESS MODEL

This section offers some representations of time evolution determined by the branching process model of the previous section, following Pitman (1996), where proofs and further details are provided. These representations provide simplified constructions of the Markov process of merger histories determined by the partition probability function (6). The constructions allow efficient simulation of the merger history process (see Section 4), and provide the basis for a more detailed analysis of the branching process model.

A tree labelled by the set $[k]$ of integers $\{1, 2, \cdots, k\}$ will be called here a *k-tree*. Such a tree is a connected graph with vertex set $[k]$ and no cycles, formally identified by its collection of $k-1$ *branches*, where each branch is an unordered pair of vertices chosen from the set $[k]$. According to Cayley's formula, there are k^{k-2} distinct k-trees. A *uniform random k-tree* is a k-tree picked at random in such a way that each of these k^{k-2} trees is equally likely to be chosen. It is known that if a Poisson Galton–Watson family tree is conditioned to have total size k, the k-tree obtained by a uniform random labelling of the k individuals in the family is a uniform random k-tree (Aldous 1991). It follows that the partition probability function $p(n_1 \cdots n_k; b_1|k; b_2)$ displayed in equation (6) describes the distribution of sizes of sub-trees in the random forest obtained from a uniform random k-tree by cutting each of the $k-1$ branches in this tree with probability $1-(b_1/b_2)$, independently of all other branches. This model for generating a random forest was proposed by Moon (1971), who derived the formula (4) for the probability that a randomly chosen vertex is a member of a j-subtree. This representation gives the binomial distribution (equation 5) with no calculation: by construction the number of branches cut has binomial distribution with parameters $k-1$ and $1-(b_1/b_2)$, and the number of subtrees in the partititon is this number of cuts plus one.

The merger history process determined by the branching process model can now be described as follows. To obtain the history over the time interval $[0, b_2]$ of a collection of particles that ends up in a k-clump at time b_2, let the particles in the k-clump be arbitrarily labelled by the set $[k]$ and then connected by a uniform random k-tree. Label the branches of this tree in an arbitrary way by i with $1 \le i \le k-1$, and assign to branch i a random variable U_i, where the U_i are independent with uniform distribution on $[0, b_2]$. For $0 \le b_1 \le b_2$ let the clumps at time b_1 be formed by the subtrees of the k-tree obtained by cutting those branches i with $U_i > b_1$. This construction defines a *merger history process*, say $Y = (Y(b_1), 0 \le b_1 \le b_2)$, which is a Markov process whose time parameter b_1 ranges over $[0, b_2]$, and whose state-space may be identified either as the set of all forests labelled by $[k]$, or as the set of all partitions of the set $[k]$. In the latter case the partition is that generated by the subtrees of the forest. Note that by construction the times U_i are the times at which mergers occur. Since these times are all distinct with probability one, the process evolves by binary mergers. If branch i of the ultimate k-tree at time b_2 joins two particles labelled say p and q, then at time U_i a subtree containing p is joined to a subtree containing q by creation of a branch from p to q. At the level of partitions, at time U_i the clump containing p merges with the clump containing q.

Given that the k-clump at time b_2 has exactly m subclumps at some time $b_1 < b_2$, let $Z_1, \cdots, Z_m$ denote the random sizes of these subclumps, listed in a random order independent of the sizes. It follows from the form of the partition probability function (equation 8) that the random vector $(Z_1, \cdots, Z_m)$ has the same joint distribution as $(X_1, \cdots, X_m)$ given $S_m = k$ for a sequence of independent random variables X_i, each with the Borel(b) distribution (equation 1). This result can also be deduced from the random forest description of the merger history process given above, and Pavlov's (1977) description of the distribution of the sizes of tree components of a random forest. To illustrate, given that at time b_1 there are $m = 2$ clumps of sizes Z_1 and Z_2, where the order of the two clumps is decided by the toss of a fair coin independent of their sizes,

$$\begin{aligned} P(Z_1 = z) &= \frac{\eta(z, b_1)\,\eta(k-z, b_1)}{P(b_1, S_2 = k)} \\ &= \frac{k}{2}\,\frac{(k-2)!}{k^{k-2}}\,\frac{z^{z-1}}{z!}\,\frac{(k-z)^{k-z-1}}{(k-z)!}, \end{aligned} \tag{10}$$

where $1 \le z \le k-1$. The first line follows from conditioning on $m = 2$ in equation (8), and the second from using equations (1) and (7) for the Borel and Borel–Tanner distributions, respectively.

The time reversed merger history process of a k-clump can be simulated by repeated application of this rule to determine how clumps are split, as the time parameter b_1 decreases from b_2 to 0. This *splitting process* starts with the k-clump at time b_2, then splits it into fragments in $k-1$ steps. The splitting times are at $U_{(1)} > \cdots > U_{(k-1)}$ obtained by ranking the values of $k-1$ independent uniform $[0, b_2]$ variables U_i. Independently of these splitting times, the first split at time $U_{(1)}$ is made into components of sizes Z_1 and $k - Z_1$ according to equation (10). After the k-clump has been split into m parts, of sizes say $(z_1 \cdots z_m)$ in some arbitrary order, the split at time $U_{(m)}$ is defined as follows: pick the ith clump with probability $(z_i - 1)/(k-m)$, and given that the ith clump has been picked, split this clump of size z_i into two subclumps of sizes X and $z_i - X$, where X is a random variable with the distribution of Z_1 in equation (10) for z_i instead of k. In terms of the random forest model, this splitting rule corresponds to picking the branches to cut by a process of sampling without replacement. After $m-1$ branches have been cut and there are m subtrees of sizes $z_1, \cdots, z_m$, the ith subtree contains $z_i - 1$ branches, and there are $k-m$ branches left altogether. So the probability that the next cut falls in the ith subtree is $(z_i - 1)/(k-m)$, and given the cut is made in a subtree of size z_i, this cut is made just as if it were the first cut starting from a clump of size z_i.

By application of Bayes' rule, the reversed time splitting process of the last paragraph can be shown to be equivalent to the following construction of the merger history process $Y = (Y(b_1), 0 \le b_1 \le b_2)$ with time running forwards by a *discrete time additive coalescent rule*: Start at time 0 with k separate particles of equal mass, and as before let mergers occur at $k-1$ random times in the interval $[0, b_2]$ defined by ranking $k-1$ independent uniform variables on $[0, b_2]$. Independently of these times, merge clumps by adding a single branch between one randomly chosen particle in each of two clumps, with the probability of choosing a pair of clumps proportional to the sum of the two clump masses. See Appendix A for details.

To summarize, we have provided three different ways of constructing the merger history process derived from the branching process model of gravitational clustering. The merger history process is formalized as a Markov process $Y = (Y(b_1), 0 \le b_1 \le b_2)$, with statespace the set of partitions of $[k]$. The reverse transition probabilities of Y induce the partition probability function displayed in equation (6). The process has the feature that, for any $0 < b < b_2$, and given the partition $Y(b)$, the histories over $[0, b]$ of the different clumps of $Y(b)$ are mutually independent, each governed by a simple rescaling of the original process.

The results of this section concern the merger histories of clumps containing exactly k particles. For results of astrophysical interest, it is interesting to consider the limit in which the mass of any given subclump is a negligible fraction of the total mass of the system. This limit is discussed more fully in Section 5. Before doing so, it is useful to consider the distribution of the partition at time b_1 defined by the particles of a k-clump at time b_2, in an asymptotic regime with k large, and b_1/b_2 bounded away from 0 and 1. Observe first that, by the law of large numbers for the binomial distribution, it is overwhelmingly likely that the number m of clumps will be such that $m/k \approx 1 - b_1/b_2$. Since there are about $(1 - b_1/b_2)k$ clumps with a total of k particles, the expected number of particles in each clump is about $(1 - b_1/b_2)^{-1}$. It follows that the sequence of clump sizes in random order $(Z_1, Z_2, \cdots)$ is a sequence of about $(1 - b_1/b_2)k$ approximately independent random variables, each with approximately Borel(b_1/b_2) distribution (Pavlov 1977). Put another way, among the large number of clumps at time b_1 that are destined to form a particular large clump at time b_2, the relative frequencies of clumps of various sizes will be approximately Borel(b_1/b_2) distributed, with high probability. Call this result the *law of large numbers* for the merger history process. Equation (4) is consistent with this law: taking the limit $k \gg j$ shows that $f(j, b_1|k, b_2) \to (1 - b_1/b_2)\, j\, \eta(j, b_1/b_2)$, which is just like equation (2).

time

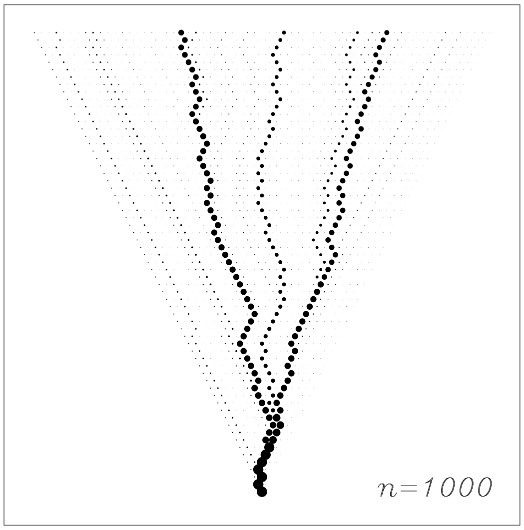

Figure 2. An example of a tree computed using the split rule. The figure shows the evolution of a clump with 10^3 particles; the area of each filled circle indicates the mass of the subclump it represents. There is only one split per time-step, and only the first few splits are shown. To convert the time axis to one that shows b_1/b_2, the split times must be distributed uniformly on the interval $[0, b_2]$. This distribution of waiting times between splits is not shown by the Figure.

Consider the process $Y^* = (Y^*(\tau), 0 \le \tau \le \infty)$ derived from the merger history process $Y = (Y(b_1), 0 \le b_1 \le b_2)$ by the following deterministic change of time parameter: $Y^*(\tau) = Y(b_1)$ where $1 - e^{-\tau} = b_1/b_2$. (Note that the time parameter τ here is different from cosmological time t. We discuss the relation between τ and t in Section 5.) The process Y^* is a partition valued Markov process with stationary forwards transition probabilities determined by the following *additive coalescent rate*: at each given time τ, each pair of components in the partition Y^* is merging at a rate which is the sum of the masses of the two components, where each of the k particles is assigned mass $1/k$. This kind of *stochastic coalescent* process, with statespace the set of partitions of $[k]$, and a Markovian transition mechanism which is determined by a *collision rate* S_{xy} giving the rate at which clumps of mass x and mass y collide, was introduced by Marcus (1968) and studied also by Lushnikov (1978a,b) and others. In the present *additive coalescent* model, $S_{xy} = x + y$. In his study of this model, Lushnikov (1978b) obtained our partition probability function in generating function form (see Hendriks et al. 1985).

Kingman (1982) gave a detailed analysis of the simplest stochastic coalescent process, with kernel S_{xy} identically equal to 1, which has found numerous applications in genetics. As remarks of Ziff (1980) suggest, Kingman's coalescent is related to cutting a random linear graph to form a collection of segments in much the same way as the additive coalescent is related to cutting a random tree to form a random forest.

4 MONTE-CARLO REALIZATIONS OF MERGER HISTORY TREES

In this section, the algorithms suggested by the coalescent and splitting processes described in Section 3 are used to produce an ensemble of realizations of the merger history tree. The trees obtained in this way are compared with the analytic results obtained using the branching process. This is done mainly to show that the numerical algorithms work. Then this ensemble of trees is compared with that which is obtained by using the algorithm developed by Kauffmann & White (1993).

4.1 Tests of the algorithm

Fig. 2 shows an example of a tree computed using the split rule described in the previous section. It shows the evolution of a clump with 10^3 particles; the area of each filled circle indicates the mass of the subclump it represents. There is only one split per time-step, and only the first few splits are shown. The figure is not exactly correct, as it shows the same constant interval between each split, rather than a distribution of waiting times. To convert the time axis to one that shows b_1/b_2, the split times must be distributed uniformly on the interval $[0, b_2]$. This means that, for large k, the distribution of waiting times between splits is approximately exponential. Once the tree is known as a function of b_1/b_2, it is straightforward to convert to cosmological time (Sheth 1995, 1996).

Fig. 3 shows $\langle n_j|k\rangle = (k/j)\, f(j|k)$, the average number of progenitor subclumps having exactly j particles when the parent is a k-clump, as a function of j, for a representative value of $k = 100$. The split-rule algorithm was used to generate 10^4 Monte-Carlo realizations of trees similar to those shown in Fig. 2. The figure shows $(k/j)\, f(j|k)$ given $k = 100$, for a range of output times, corresponding to a range of choices of b_1/b_2. These output times are at values of $b_1/b_2 = 0.22, 0.39, 0.63, 0.86$, and 0.98, which correspond to $\tau = 0.25, 0.5, 1, 2$, and 4, where $b_1/b_2 = 1 - \mathrm{e}^{-\tau}$.

The curves that are higher on the left correspond to smaller values of b_1/b_2. The solid lines show the analytical result obtained from equation (4). The symbols show the result measured from the Monte–Carlo trees. They trace the analytic result well, confirming that the numerical algorithm works. By construction, the split-rule algorithm gives the correct Binomial $n(m|k)$ distribution. The numerical results fit the Binomial distribution very well, although we do not show the corresponding figure here.

Recall that $(k/j)\, f(j|k)$ is the mean number of j-subclumps per parent k-clump (equation 4). It is necessary to check that the numerical algorithm gives the correct higher order moments of this n_j distribution also. To illustrate this, Fig. 4 shows the second factorial moment $\langle n_j(n_j - 1)|k\rangle$ of this distribution. The solid curves show the analytical result ($i = 2$ in equation 9) for the same values of b_1/b_2 as in Fig. 3. Again, curves that are higher on the left correspond to smaller values of b_1/b_2. The symbols show the same quantity measured in the Monte–Carlo ensemble. They trace the analytic result well. We conclude that our scheme is able to generate merger history trees that have the correct ensemble properties.

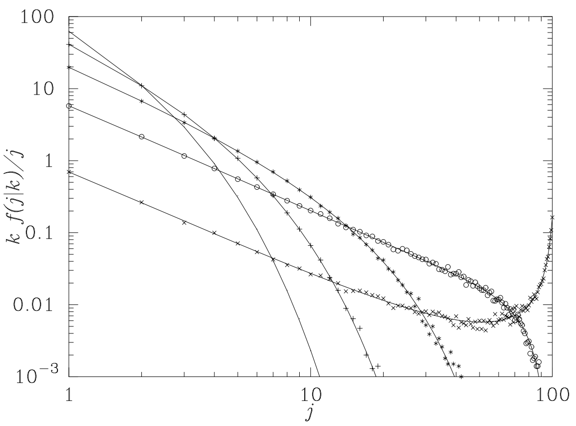

Figure 3. The mean number of j-subclumps per parent k-clump, with $k = 100$, for a range of choices of the ratio b_1/b_2 chosen so that $b_1/b_2 = 1 - \mathrm{e}^{-\tau}$ for $\tau = 1/4, 1/2, 1, 2$ and 4. The curves that are higher on the left correspond to smaller values of b_1/b_2. The solid lines show the analytical result obtained from equation (4). The symbols show the result measured from the Monte–Carlo trees. They trace the analytic result well, confirming that the numerical algorithm works.

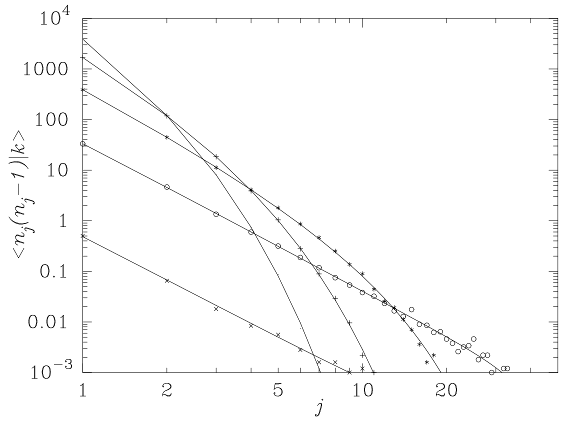

Figure 4. The second factorial moment of the distribution of the number of j-subclumps of a parent k-clump: $\langle n_j(n_j - 1)|k\rangle$. The solid curves show the analytic result ($i = 2$ in equation 9) for the same values of b_1/b_2 as in Fig. 3. Again, curves that are higher on the left correspond to smaller values of b_1/b_2. The symbols show the same quantity measured in the Monte–Carlo ensemble; they trace the analytic result well. We conclude that our scheme generates merger history trees that have the correct ensemble properties.

4.2 Comparison with other algorithms

It is interesting to compare this merger history process with the merger history algorithm developed by Kauffmann & White (1993). We will do this as follows. First, we construct a large number (10^4) of trees using the split rule. Then we choose a set of output times, which means we specify a few values of b_1/b_2. For a given choice of b_1/b_2, the split rule gives a set of progenitor subclumps. We use these progenitor clumps as input to the Kauffmann–White algorithm. This insures that the Kauffmann–White trees are constructed from the same set of progenitors as the branching process, split rule trees.

For each given value of b_1/b_2, we implement their al-

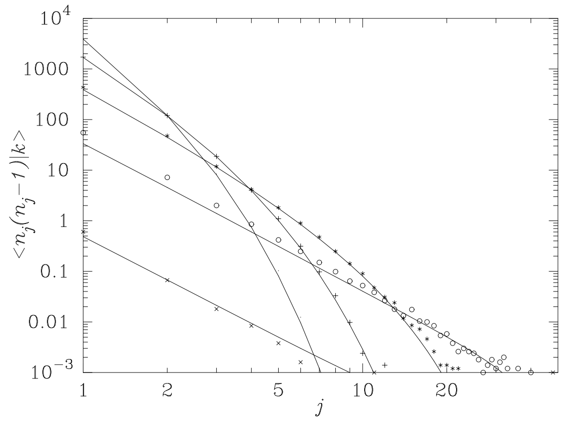

Figure 5. The second factorial moment of the distribution of the number of j-subclumps of a parent k-clump: $\langle n_j(n_j-1)|k\rangle$. The solid curves show the analytic result (equation 9, with $i=2$) for the same values of b_1/b_2 as in Figs. 3 and 4. Curves that are higher on the left correspond to smaller values of b_1/b_2. The symbols show the same quantity measured in the Kauffmann–White ensemble. This figure should be compared with Fig. 4.

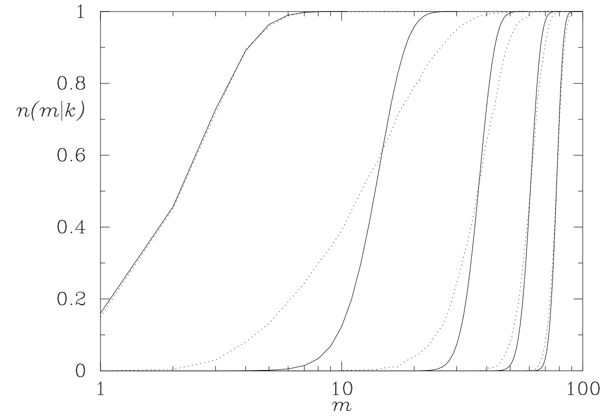

Figure 6. The cumulative distribution of the number of subclumps m for a tree with k-members. The solid lines show the analytic branching process result (the sum of equation 4 from 1 to m) for the values of b_1/b_2 used in Figs. 3 and 4. Curves on the right correspond to smaller values of b_1/b_2 (they have a larger number of subclumps). The dotted lines show the corresponding cumulative distribution for the Kauffmann–White algorithm.

gorithm as follows. Start with the 10^4 parent clumps, each having k empty mass elements. Also, start with the progenitor clumps from the split rule algorithm. Recall that the analytic results of the previous sections, and Fig. 3 of this section, show that this distribution of subclumps is consistent with what is required by the excursion set formalism (equation 4). The Kauffmann–White algorithm assigns the most massive remaining progenitor subclump to one of the (10^4) parent clumps with probability proportional to the amount of mass remaining to be filled in each parent clump, subject to the requirement that no parent become over-filled. This continues until all the progenitors have been assigned to all the parent clumps. In this way, the Kauffmann–White algorithm builds an ensemble of trees that is consistent with the excursion set constraints (equation 4).

How different are the Kauffmann–White trees from our Poisson Galton–Watson branching process trees? By construction, the Kauffmann–White algorithm will have the correct mean number of progenitor subclumps $\langle m\rangle$, and the correct $\langle n_j|k\rangle$ distributions. There is no guarantee that higher order statistics of the Kauffmann–White ensemble will be the same as for the branching process trees. Therefore, we will compare the two tree ensembles using two different higher order statistics: the distribution of the number m of subclumps of a given parent clump, and the second factorial moment distribution $\langle n_j(n_j-1)|k\rangle$, where n_j is the number of j-subclumps within a parent k-clump. For the branching process merger history trees, these distributions are given by equations (5) and (9), respectively.

Fig. 5 is similar to Fig. 4 shown earlier. It shows the second factorial moment of the n_j distribution, $\langle n_j(n_j-1)|k\rangle$, which is a measure of the variance of different merger histories. The solid lines show the analytic branching process result that describes the split-rule trees well (equation 9), for the range of choices of b_1/b_2 considered in the other figures. The symbols show the corresponding quantity measured using the Kauffmann–White trees. They follow the analytic curves reasonably well (but see the discussion below).

Figure 6 shows the cumulative $n(m|k)$ distribution for the Kauffmann–White trees. The solid lines show the analytic branching process result as a function of m (the sum of equation 5 from 1 to m) for the values of b_1/b_2 used in Figs. 3 and 4. Curves on the right correspond to smaller values of b_1/b_2 (they have a larger number of subclumps). By construction, the split-rule algorithm fits the solid lines extremely well. The dotted lines show the corresponding cumulative distribution for the Kauffmann–White algorithm. The distributions are similar at small 'look-back' times, when the parent clumps have not broken up into very many smaller pieces (curve on the left), and also at large look-back times (curve on the right), when a majority of subclumps are single particles. However, in the intermediate regime, the $n(m|k)$ distribution for the Kauffmann–White trees can differ significantly from the Binomial distribution of the branching process (curves in the middle).

Closer inspection of Fig. 5 shows this same discrepancy. Consider the curve that is second from the bottom on the left of that figure. The open circles representing the Kauffmann–White algorithm lie above the analytic curve. Since $\langle n_j(n_j-1)|k\rangle$ is a measure of the variance between different merger histories, this suggests that the Kauffmann–White histories are more varied than the branching process histories. Moreover, this curve corresponds to the penultimate output time, and so to the set of curves that are second from the left in Fig. 6. For this output time, the solid and dotted curves differ significantly: the solid curve rises more steeply than the dotted curve. This implies that the $n(m|k)$ distribution for the branching process (represented by the solid curve) is more like a delta function than is the corresponding Kauffmann–White distribution. This suggests that, at this output time, the branching process trees are more homogeneous than the Kauffmann–White trees, which is consistent with Fig. 5.

Since the branching process description has some physical justification, whereas the Kauffmann–White algorithm does not, it is interesting to see if a simple modification of the Kauffmann–White algorithm can remedy the discrepancy between its merger histories and those of the branching process.

Suppose that the most massive remaining progenitor subclump is assigned to a parent clump that is chosen uniformly at random from among the eligible parents, with eligibility determined solely by requiring that the parent not become over-filled. That is, other than determining eligibility, the mass remaining to be filled in a given parent clump is irrelevant. We used this modified algorithm to build merger history trees—the difference between these trees and those of the branching process was slightly more obvious than the difference between the original Kauffmann–White trees and the branching process trees.

If the largest remaining progenitor is assigned to the emptiest parent with probability one, which we will call a deterministic rule, then the algorithm does not produce trees having, say, only massive progenitors, or only tiny progenitors. Instead, it makes merger history trees that are extremely homogeneous (trees that are more like those at the top of Fig. 1 than those at the bottom). The $n(m|k)$ distributions resulting from this modification become more like delta functions than like the Binomial distribution of the branching process. Moreover, the homogeneity of the trees means that the $\langle n_j(n_j-1)|k\rangle$ distributions differ substantially from those shown in Figs. 4 and 5. By relaxing this deterministic rule, the original Kauffmann–White algorithm is able to produce merger history trees that are more varied. However, Figs. 5 and 6 show that, in the interesting regime, in which subclumps are numerous and many are still a substantial fraction of the mass of the parent clump, the Kauffmann–White merger histories are too varied compared to those of the branching process.

We conclude that, at least for clustering from Poisson initial conditions, properties of ensembles of merger history trees that are generated using the Kauffmann–White algorithm may be quite different from those of an ensemble that was generated using the split-rule that follows from the branching process description. Therefore, quantitative predictions about the evolution of clustering, from Poisson as well as other initial conditions, which depend on the details of the Kauffmann–White algorithm should be interpretted cautiously.

5 RELATION TO SMOLUCHOWSKI'S COAGULATION EQUATION

The previous sections considered the merger histories of clumps containing exactly k particles. For results of astrophysical interest, it is interesting to consider the limit in which the mass of any given subclump is a negligible fraction of the total mass of the system. In the limiting regime where the total number of particles in the ultimate clump tends to infinity, the number densities of clumps of various sizes in a stochastic coalescent are governed by a deterministic equation, known as Smoluchowski's coagulation equation (Marcus 1968; Lushnikov 1978b; Hendriks et al. 1985). This limit is discussed more fully below.

Consider a set of identical particles that all have the same mass initially, and that clump together irreversibly, through binary mergers, as time progresses. Smoluchowski's (1916, 1917) coagulation equation describes the evolution of such a sticky-gas system. Smoluchowski argued that, if $c(k,\tau)$ denotes the number density of clumps having exactly k particles at the time τ [strictly speaking, $c(k,\tau)$ is the number of clumps each having exactly k particles at time τ, divided by the total number of particles, in the limit as the number of particles becomes infinite, where it is assumed that the number of particles is conserved], then the time evolution of $c(k,\tau)$, when irreversible coagulation takes place by binary mergers, is given by

$$\begin{aligned}\frac{\mathrm{d}\,c(k,\tau)}{\mathrm{d}\tau} &= \frac{1}{2}\sum_{i+j=k} S_{ij}\,c(i,\tau)c(j,\tau) - c(k,\tau)\sum_{j=1}^{\infty} S_{jk}\,c(j,\tau)\\ &\equiv R_{\rm form} - R_{\rm dest}, \qquad (11)\end{aligned}$$

where S_{ij} describes the rate with which objects of mass i and j merge with each other. The first term on the right is a formation term; it describes the increase in the number of k-clumps due to the fact that two less massive clumps can merge to make a k-clump. The second term is a loss or destruction term; it allows for the fact that k-clumps may themselves merge with other clumps, and so become more massive than k.

Smoluchowki's formulation of the sticky-gas model is sufficiently general that it has been used to describe a wide variety of physical phenomena (Drake 1972). For example, in this discrete form, it has been used to study the onset of gelation in polymer physics (Ernst 1986 and references therein). In astrophysics, Field & Saslaw (1965) and Silk & Takahashi (1979) used a continuous version of this equation (all sums become integrals, and the mass becomes continuous rather than integer valued) to model the initial stellar mass function, and Silk & White (1978) used it to study the distribution of cluster masses.

Two important ingredients are required to specify solutions to Smoluchowski's equation. The first is the initial particle distribution. In this paper we will mainly consider the case in which all particles have the same (unit) mass at the initial time, say $\tau=0$. This is known as the case of monodisperse initial conditions. The polydisperse case is considered in the Appendix. The second ingredient is the merger cross-section S_{ij}. In general, S_{ij} can be very complicated, so that the most general analytic solution of Smoluchowski's equation has not been written down. For example, one can imagine that S_{ij} could depend on the masses, the relative velocities, the local density, the time τ, and so on. In fact, even for relatively simple cross-sections, few analytical solutions of Smoluchowski's equation are known.

Smoluchowski (1917) solved the case in which S_{ij} is a constant, independent of i, j, time, etc., for monodisperse initial conditions. When $S_{ij}=ij$, the solution is due to McLeod (1962a,b,c). Scott (1968) considered the cases $S_{ij} = 1, (i+j)$, and ij separately, and Spouge (1983) obtained the solution in an implicit form for the general case with $S_{ij}=A+B(i+j)+Cij$, where A, B, and C are arbitrary constants.

The solution to Smoluchowski's equation when the merger cross section is additive, that is, $S_{ij}=i+j$, and all particles are identical at the initial time $\tau=0$ (monodisperse initial conditions), is known (Scott 1968; Lushnikov 1978; van Dongen 1987) to be

$$c_k(\tau) = \mathrm{e}^{-\tau}\eta(k, 1-\mathrm{e}^{-\tau}) \qquad (12)$$

for $\eta(k,b)$ as in equation (1). If we set

$$b = 1 - \mathrm{e}^{-\tau} \tag{13}$$

then we see that Smoluchowski's coagulation process with an additive rate law, when started from monodisperse initial conditions, results in a Borel distribution of clump sizes. The parameter b of the Borel distribution increases as the time τ in the coagulation process increases. Note that the time parameter in the coagulation process is not the same as cosmological time. Cosmological time and coagulation time are related, through b, by equations (3) and (13). As a check, exactly the same distribution (12) of clump sizes is obtained from the additive stochastic coalescent process Y^* of the previous section, in the limit for a large number of particles in the ultimate clump, by application of the law of large numbers for the merger history process.

Appendix B shows that the partition probability function of equation (6) is consistent with the additive coagulation equation. This involves solving Smoluchowski's additive equation for the case of polydisperse initial conditions, and using the Borel distribution with a given value of b_1 as the initial clump-size distribution. In essence, this provides yet another derivation of the scaling solution described in section 3.2 of Sheth (1995).

6 RELATION TO EXCURSION SET RESULTS

This section shows how the results of the previous sections are related to the excursion set approach of Bond et al. (1991) and Lacey & Cole (1993). It also discusses ways in which the branching process and coagulation model results of the previous sections may be used to provide information about the growth of clustering from Gaussian, as opposed to Poisson, initial conditions.

6.1 Poisson initial conditions

To relate the coagulation process description of clustering to the excursion set description, it is useful to express $R_{\rm form}$ and $R_{\rm dest}$ in terms of cosmological, as opposed to coagulation, time.

First, recall that coagulation time τ and cosmological time t are related through the b parameter (see equations 3 and 13). This means that derivatives with respect to cosmological time can be expressed in terms of derivatives with respect to coagulation time:

$$\frac{\mathrm{d}}{\mathrm{d}t} = \left(\frac{\mathrm{d}\tau}{\mathrm{d}b}\frac{\mathrm{d}b}{\mathrm{d}\delta_{\rm c}}\frac{\mathrm{d}\delta_{\rm c}}{\mathrm{d}t}\right)\frac{\mathrm{d}}{\mathrm{d}\tau} = \frac{b^2}{1-b}\left(-\frac{\mathrm{d}\delta_{\rm c}}{\mathrm{d}t}\right)\frac{\mathrm{d}}{\mathrm{d}\tau}. \tag{14}$$

Next, let $N(j,b_1|k,b_2)$ denote the average number of j-sized b_1-progenitor subclumps of a k-clump. That is, $N(j|k)$ is the same as $n_j\,p(n_1\cdots n_k;b_1|k;b_2)$ summed over all partitions of k, so it is equal to $(k/j)\,f(j|k)$ (equation 4). In the limit of small time steps, this can be used to estimate the rate at which j-clumps merge to form k-clumps. In particular, for all $j<k$,

$$\begin{aligned} N(j\to k|\delta_{\rm c})\,\mathrm{d}\delta_{\rm c} &\equiv \lim_{b_1\to b_2}\frac{k}{j}\,f(j,b_1|k,b_2) \\ &= 2\,\frac{\eta(j,b_1)\,\eta(k-j,b_1)}{P(b_1,S_2=k)}\,n(2,b_1|k,b_2) \\ &\to k\,\frac{\eta(j,b)\,\eta(k-j,b)}{\eta(k,b)}\,\frac{\mathrm{d}b}{\mathrm{d}\delta_{\rm c}}\,\mathrm{d}\delta_{\rm c}, \end{aligned} \tag{15}$$

where the second line in the expression above comes from using equations (1), (4) and (7), and setting $m=2$ in equation (5). Notice that the final expression is symmetric in j and $k-j$. Moreover, it is similar to equation (10), which was derived from the binary merger interpretation of the branching process (Section 3).

This, with equations (11) and (14), means that, in the Poisson case, the formation rate in terms of cosmological time is

$$\begin{aligned} R_{\rm form} &= \frac{b^2}{1-b}\left(-\frac{\mathrm{d}\delta_{\rm c}}{\mathrm{d}t}\right)\frac{k}{2}\sum_{j=1}^{k}(1-b)^2\,\eta(j,b)\eta(k-j,b) \\ &= (1-b)\,\eta(k,b)\left(-\frac{\mathrm{d}\delta_{\rm c}}{\mathrm{d}t}\right)\sum_{j=1}^{k}\frac{N(j\to k|\delta_{\rm c})}{2} \\ &= (1-b)\,\eta(k,b)\left(-\frac{\mathrm{d}\delta_{\rm c}}{\mathrm{d}t}\right)(k-1)\,b. \end{aligned} \tag{16}$$

The second equality follows from the definition of $N(j\to k|\delta_{\rm c})$ in equation (15), and the final expression from recognizing that the sum over j is similar to that which defines the Borel–Tanner distribution (equation 7) with $m=2$.

It is easy to see that this expression for the formation rate is sensible. Recall that $N(j|k)$ is the average number of j-clumps per k-clump. This means that $N(j\to k|\delta_{\rm c})$ is the average number of j-clumps that undergo a binary merger to form a k clump. Since mergers are binary, and we are summing over all $j<k$, we must divide this number by two. This, then, gives the fraction of k-clumps that were made by binary mergers involving j-subclumps. This must be multiplied by the total number of k-clumps to give the total number of k-clumps that were formed by binary mergers from j-clumps.

Another way in which to see that equation (15) is sensible is to note that the number density of j-clumps from the epoch b_1 that are later incorporated into k-clumps at the epoch b_2 is the average density times $(1-b_1)\,\eta(j,b_1)$ times $f(k,b_2|j,b_1)$ (see equation 2.24 and associated discussion in Lacey & Cole 1993). Bayes' rule implies that this is equal to the density times $f(j,b_1)\,f(k,b_2|j,b_1)/j = (1-b_2)\,\eta(k,b_2)\,(k/j)\,f(j,b_1|k,b_2)$. Taking the limit $b_1\to b_2$ in this expression yields $(1-b_2)\,\eta(k,b_2)\,N(j\to k|\delta_{\rm c})$, which can be interpretted as the rate with which j-clumps merge to form k-clumps (see the discussion following equation 2.17 of Lacey & Cole 1993). If, in the limit of small time steps considered here, all mergers are binary, then each j-clump that forms a k-clump increases the number density of k-clumps. Thus, this expression provides an estimate of the rate of change (increase) in the number density of k-clumps that is due to binary mergers between j- and $(k-j)$-clumps. In other words, this expression gives the formation rate of k-clumps from j-clumps. Summing this expression over all j and dividing by two (since mergers are binary) gives the formation rate of k-clumps. This is the same as the sum in the penultimate line of equation (16).

Similarly, the destruction rate is

$$R_{\rm dest} = \frac{b^2}{1-b}\left(-\frac{\mathrm{d}\delta_{\rm c}}{\mathrm{d}t}\right)(1-b)^2\,\eta(k,b)\sum_{j=1}^{\infty}(k+j)\,\eta(j,b)$$

$$= (1-b)\,\eta(k,b)\,\left(-\frac{\mathrm{d}\delta_{\rm c}}{\mathrm{d}t}\right) \times \sum_{k'=k+1}^{\infty} \frac{N(k\to k'|\delta_{\rm c})\,\eta(k',b)}{\eta(k,b)}$$

$$= (1-b)\,\eta(k,b)\,\left(-\frac{\mathrm{d}\,\delta_{\rm c}}{\mathrm{d}t}\right)\,\left(kb^2 + \frac{b^2}{(1-b)}\right), \quad (17)$$

where $N(k \to k'|\delta_{\rm c})$ was defined in equation (15). The term in the sum is just $f(k', b_3|k, b_2)$, where $k' > k$ and $b_3 > b_2$, in the limit of small time steps, and the final expression comes from evaluating this sum.

Notice that the first expressions in these two equations for $R_{\rm form}$ and $R_{\rm dest}$ were derived from Smoluchowski's equation, under the assumption that all mergers are binary, and the merger rate is additive. However, the penultimate expressions in equations (16) and (17) are written entirely in terms of quantities that can be determined directly from the excursion set approach. This shows explicitly that, at least for the case of Poisson initial conditions, the Smoluchowski-type description can be obtained directly from the excursion set functions themselves.

Furthermore, to derive the formation rate from the excursion set functions one needed to know that mergers were binary (e.g., this required that we include a factor of 1/2 when computing $R_{\rm form}$). If, for example, ternary mergers were allowed, and the ratio of binary to ternary mergers was unknown, then one would not have been able to write down the formation rate. On the other hand, the destruction rate simply counts the rate at which clumps of mass k become incorporated into clumps with mass $k' > k$, so it does not require knowledge of whether the mergers are binary or not. We will use this fact in the next subsection.

Since the excursion set functions are known for arbitrary Gaussian initial conditions (e.g. Lacey & Cole 1993), we can try to generalize the Poisson treatment above to the Gaussian case. That is, we can try to formulate the excursion set description in terms of a binary merger coagulation equation. We will do this in two steps. The next subsection considers the case of white-noise initial conditions, and section 6.3 considers Gaussian fields with more general power-spectra.

6.2 White-noise initial conditions

In the limit of small δ, which corresponds to the $b \to 1$ limit, Stirling's approximation for the factorials reduces all the Poisson excursion set statements (equations 1 and 4) to the corresponding statements for $n = 0$ white-noise initial conditions (Sheth 1995, 1996). However, it is not straightforward to translate the branching process partition function results into statements about the $n = 0$ Gaussian case (but see Spouge 1984, 1985). Therefore, it is interesting to see if the additive coagulation scheme can at least describe clustering from white-noise, if not all, Gaussian initial conditions.

In the excursion set approach, the number density of clumps that have mass in the range $\mathrm{d}M$ about M at cosmological time t is

$$n(M,t)\,\mathrm{d}M = \frac{\rho_0}{\sqrt{2\pi}}\,\frac{\delta_{\rm c}(t)}{\sigma^3(M)}\,\left|\frac{\mathrm{d}\sigma^2(M)}{\mathrm{d}M}\right| \times \exp\left(-\frac{1}{2}\frac{\delta_{\rm c}^2(t)}{\sigma^2(M)}\right)\,\frac{\mathrm{d}M}{M}, \quad (18)$$

where ρ_0 is the comoving density of the universe, $\delta_{\rm c}(t)$ is the critical density required for a spherical perturbation to collapse by time t (cf. equation 3), and $\sigma^2(M)$ is the variance in the initial density fluctuation field when smoothed over regions that contain mass M on average. The variance decreases monotonically with M in a way that depends on the power spectrum (e.g. Press & Schechter 1974; Bond et al. 1991; Lacey & Cole 1993). This means that the rate of change of the excursion set mass functions is

$$\frac{\mathrm{d}\,n(M,t)}{\mathrm{d}t} = n(M,t)\,\left(-\frac{\mathrm{d}\,\delta_{\rm c}}{\mathrm{d}t}\right)\left(\frac{\delta_{\rm c}(t)}{\sigma^2(M)} - \frac{1}{\delta_{\rm c}(t)}\right). \quad (19)$$

For white noise, $\sigma^2(M) \propto M^{-1}$, so that this expression is similar to the Poisson expression $\mathrm{d}\,(1-b)\eta(k,b)/\mathrm{d}t$, when $b \to 1$.

The excursion set merger statements in the limit of small time steps are

$$N(M_1 \to M_2|\delta_{\rm c})\,\mathrm{d}\delta_{\rm c} = \frac{M_2}{M_1}\,\frac{\mathrm{d}\delta_{\rm c}}{\sqrt{2\pi}}\,\frac{1}{(\sigma_1^2-\sigma_2^2)^{3/2}}\,\left|\frac{\mathrm{d}\sigma_1^2}{\mathrm{d}M_1}\right|, \quad (20)$$

where $M_1 < M_2$ and $\sigma_{1,2} = \sigma(M_{1,2})$ (e.g. Lacey & Cole 1993). For white-noise, $N(M_1 \to M_2|\delta_{\rm c})$ is symmetric in M_1 and $M_2 - M_1$, and the exact expression is similar to the corresponding Poisson statement (when Stirling's approximation for the factorials is used), suggesting that binary mergers dominate. Thus, the formation rate of equation (16) becomes

$$R_{\rm form} = \frac{n(M,t)}{2}\,\left(-\frac{\mathrm{d}\,\delta_{\rm c}}{\mathrm{d}t}\right)\int_0^M N(M'\to M|\delta_{\rm c})\,\mathrm{d}M'$$

$$\propto \frac{M}{2}\int_0^M n(M',t)\,n(M-M',t)\,\mathrm{d}M', \quad (21)$$

where the final expression, which applies only in the white-noise case and is additive as it was for the discrete Poisson case, holds provided $\delta_{\rm c} \ll 1$ (i.e., the late time limit). Similarly, one can show that the destruction rate, computed using the final expression of equation (17), is also additive (in the limit of late times). Although the integrals diverge, this suggests that the Smoluchowski equation with additive rate reproduces the $n = 0$ excursion set results.

This conclusion is consistent with the fact that the late time solutions to the additive coagulation scheme are known to have the same form as the $n = 0$ excursion set mass functions (e.g. Scott 1968; Trubnikov 1971; Silk & White 1978). Thus, the additive coagulation scheme with binary mergers can be used to construct the merger history tree for the $n = 0$ case, just as it was used for the Poisson case in the previous section.

Since the integral for the formation rate in equation (21) diverges, and the corresponding integral for the destruction rate also diverges, it is useful to study the Poisson expressions further. Taking the limit $\delta \to 0$, so that $b \to 1 - \delta_{\rm c}$, and using Stirling's approximation for the factorials in equation (16) yields

$$R_{\rm form} \to n(M,t)\,\left(-\frac{\mathrm{d}\,\delta_{\rm c}}{\mathrm{d}t}\right)\,\left(\frac{1}{\sigma^2(M)} - 1\right)\,(1-\delta_{\rm c}), \quad (22)$$

where we have set $\sigma^2(M) = 1/M$. Similarly, the limit of equation (17) is

$$R_{\rm dest} \to n(M,t)\left(-\frac{{\rm d}\,\delta_{\rm c}}{{\rm d}t}\right)\left(\frac{(1-\delta_{\rm c})^2}{\sigma^2(M)} + \frac{1-\delta_{\rm c}}{\delta_{\rm c}}\right), \qquad (23)$$

so that

$$R_{\rm form} - R_{\rm dest} \to n(M,t)\left(-\frac{{\rm d}\,\delta_{\rm c}}{{\rm d}t}\right)\left(\frac{\delta_{\rm c}}{\sigma^2(M)} + \delta_{\rm c} - \frac{1}{\delta_{\rm c}}\right). \qquad (24)$$

When $\delta_{\rm c} \ll 1$, the second term in the final set of brackets is negligible, and this expression reduces to equation (19). Keeping only the largest terms in the previous expressions for the formation and destruction rates implies that

$$R_{\rm form} \approx n(M,t)\left(-\frac{{\rm d}\,\delta_{\rm c}}{{\rm d}t}\right)\frac{(1-\delta_{\rm c})}{\sigma^2(M)}, \qquad (25)$$

and

$$R_{\rm dest} \approx n(M,t)\left(-\frac{{\rm d}\,\delta_{\rm c}}{{\rm d}t}\right)\left(\frac{1-2\delta_{\rm c}}{\sigma^2(M)} + \frac{1}{\delta_{\rm c}}\right). \qquad (26)$$

These expressions may provide useful approximations to the formation and destruction rates.

6.3 Arbitrary Gaussian initial conditions

Section 6.1 showed that the additive coagulation process is consistent with the excursion set description of clustering from Poisson initial conditions. The previous subsection argued that the two approaches also provide similar descriptions of evolution from $n = 0$ white-noise Gaussian initial conditions. The advantage of the coagulation formulation over the excursion set approach is that it shows clearly how to generate a set of merger history trees, that has the correct ensemble properties, efficiently. Since the excursion set approach is known to be accurate for arbitrary Gaussian initial conditions (Lacey & Cole 1994), it is interesting to see if a coagulation-type approach can be found that is equivalent to the excursion set description for other initial conditions also.

It is not obvious that this can be done. This is because the coagulation scheme that is relevant to the Poisson and $n = 0$ cases depends on three assumptions: binary mergers dominate, the merger rule is additive, and initially, all particles were identical. It may be that the excursion set description is consistent with all three of these assumptions only for Poisson and $n = 0$ initial conditions.

If we assume that changing the initial conditions in the excursion set description corresponds to changing the initial conditions in the coagulation description (but binary mergers still dominate and the merger rule is still additive), then we can use the solution of the polydisperse additive coagulation equation given in the Appendix (equation B24) to see if this assumption is correct. If it is, then, we should be able to find initial conditions $[c_k(\tau = 0)]$ that result in cluster distributions $[c_k(\tau)]$ which, in the appropriate limits (large k and late times), evolve similarly to the excursion set mass functions $n(M,t)$ of equation (18). The late-time behaviour of this solution has been studied by Trubnikov (1971) who argues that, even for arbitrary initial conditions, if the merger rule is additive, then the large mass cutoff scales as $\exp(-aM)$, and not as $\exp(-aM^{\alpha})$ as required by equation (18). Appendix C also shows that self-similar solutions of Smoluchowski's equation that have the form given by equation (18) are inconsistent with the assumption of an additive merger rule.

Moreover, there is at least one reason to suspect that none of the three assumptions above generalize to $n \neq 0$ Gaussian fields. This is because equation (20) is symmetric in M_1 and $M_2 - M_1$ only for a white-noise Gaussian field. To see this, first consider the case of scale-free clustering, for which the initial Gaussian field has power-spectrum $P(k) \propto k^n$ (we discuss the general case briefly later). Then $\sigma^2(M) \propto M^{-\alpha}$, where $\alpha = (3+n)/3$. Let $x = M_1/M_2$. Then equation (20) implies that

$$N(M_1 \to M_2) \propto f(x) \equiv \frac{1}{x^{2+\alpha}}\left(\frac{x^{\alpha}}{1-x^{\alpha}}\right)^{3/2}. \qquad (27)$$

Symmetry would imply that $f(x) = f(1-x)$, which is only satisfied when $\alpha = 1$, i.e., the $n = 0$ white noise case. Therefore, when $n \neq 0$, the excursion set approach is formally inconsistent with the assumption that binary mergers dominate. This means that, when $n \neq 0$, we can no longer use the excursion set statements to estimate $R_{\rm form}$ as we did in (equation 21), nor can we build the merger tree using the coagulation model. For this reason, it appears that our methods, which are consistent with the excursion set description of the growth of clustering from Poisson and $n = 0$ initial conditions, cannot be made consistent with the excursion set description of clustering from more general initial conditions.

Although $N(M_1 \to M_2)$ is not exactly symmetric in M_1 and $M_2 - M_1$ for arbitrary initial conditions, it is interesting to see how much it departs from symmetry. To illustrate this, Fig. 7 shows $[f(x) - f(1-x)]/f(x)$, which is the relative departure from symmetry, as a function of x, for a range of power-spectra. The solid, dashed, dot-dashed and dotted curves correspond to initial spectra with slopes $n = -2, -1, 0$ and 1, respectively. (The dot-dashed curve for the $n = 0$ case is zero for all values of x, as expected.) The curves show that departures from symmetry exceed ten percent in magnitude only when $x < 0.1$ or so. This suggests that a binary merger coagulation model of clustering from these more general initial conditions, may not be a bad approximation to the excursion set description after all.

This observation has an interesting relation to previous work. Figure 7 of Lacey & Cole (1993) shows the distribution of halo formation epochs calculated using the excursion set approach by using an expression like our equation (20), for power-law power spectra. If $n > 0$, they find that their formula predicts probability densities that become slightly negative, while if $n < 0$, they predict non-zero probabilities for formation at arbitrarily early times. Of course, neither of these features are acceptable. Only the $n = 0$ case is well behaved. Our Fig. 7 shows that there is good reason to believe that the $n = 0$ case is special. It is the only case for which the excursion set description is exactly equivalent to a binary merger model.

Since $f(x)$ is so close to symmetric, Fig. 7 suggests the following procedure. In principle, for any given power spectrum, one can combine equation (20) with the expression on the first line of equation (21). This would yield an effective coagulation rule for that power spectrum. Since the excursion set statements are not symmetric, this rule would

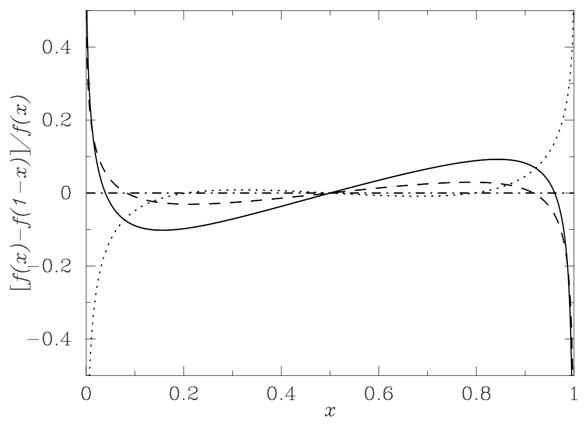

Figure 7. The relative departure from symmetry, $[f(x) - f(1-x)]/f(x)$, where $f(x)$ is given by equation (22), as a function of $x = M_1/M_2$. The solid, dashed, dot-dashed and dotted curves correspond to $n = -2, -1, 0$ and 1, respectively. The dot-dashed curve for the $n = 0$ case is zero for all values of x, as expected. Departures from symmetry exceed ten percent in magnitude only when $x < 0.1$ or so.

not be symmetric either. For example, for initially scale free power spectra, the merger rule has the form

$$S_{M_1,M_2-M_1} = M_2 \times \frac{(1-x)^{2-(\alpha/2)}}{(1-x^\alpha)^{3/2}} \times \mathrm{e}^{-\frac{\delta^2 M_2^\alpha}{2}[1-x^\alpha-(1-x)^\alpha]}, \tag{28}$$

where $x = M_1/M_2$ as before. In the white noise case ($\alpha = 1$), this reduces to M_2 (compare equation 21). So one would then have to find an approximation to this rule that was symmetric. One could then compute Monte–Carlo realizations of the merger history tree using this symmetric merger rule, as we did for the Poisson case in the previous section. However, in contrast to our results for the initially Poisson case, these merger trees would not represent the excursion set results exactly. Nevertheless, Fig. 7 suggests that they should be a reasonable approximation.

The case of initial power spectra that are not scale free is even more complicated, since then the exponent α is, in effect, mass dependent. This means, for example, that the exponents for x and $(1-x)$ would be different. Furthermore, in the scale free case, equation (27) holds for all values of M_2. In the general case, different values of M_2 have different $f(x)$ curves. So, to determine the departures from symmetry, one must first compute expressions like equation (27), and then show figures like Fig. 7, for a range of masses. Therefore, although it appears likely that such a binary merger coagulation approach could provide a good approximation to the excursion set description, we have not pursued any of this further.

Although the binary merger coagulation approach is not exactly equivalent to the excursion set approach for arbitrary power spectra, recall that the equation (23) provides an estimate of the destruction rate, at least for white-noise initial conditions. When written in terms of the variance, excursion set statements have the same form for all initial power spectra. Therefore, it might be that, when written in terms of the variance, equation (23) gives a good approximation to the destruction rate for all initial power spectra also. If so, then, since $\mathrm{d}n(M,t)/\mathrm{d}t$ is known, we can estimate $R_{\rm form}$ by subtracting $R_{\rm dest}$ from equation (19). Thus, although we cannot build a merger history tree in this way, we can at least separate the rate of change of $n(M,t)$ into $R_{\rm form}$ and $R_{\rm dest}$. This is useful, since it allows one to compute some quantities of astrophysical interest.

For example, in the Efstathiou & Rees (1988) model of the quasar luminosity function, one must use $R_{\rm form}$ to compute observable quantities. Previous authors (Haehnelt & Rees 1993) simply set $R_{\rm form} = \mathrm{d}n/\mathrm{d}t$. This means that, in effect, they assume that $R_{\rm form} \gg R_{\rm dest}$. It is not obvious that, for the range of M and t in which they are interested, their approximation is justified. Our formulation may allow one to quantify the accuracy of their assumption. We have not pursued this further.

7 DISCUSSION

We showed that the evolution of clustering in the Poisson Galton–Watson branching process model (section 2) is similar to that which is described by Smoluchowski's binary merger, sticky-gas, coagulation model, if the initial distribution is monodisperse, and the merger cross-section is additive in the mass of the merging clusters (sections 3 and 4). We solved the corresponding polydisperse system, and used the solution to argue that the coagulation equation and branching process descriptions are consistent (Appendix B). We used this correspondence to construct Monte–Carlo realizations of the merger history tree having the correct ensemble properties efficiently (Fig. 2–4). We showed that ensembles of trees constructed in this way are different from ensembles constructed using the Kauffmann–White (1993) algorithm (Figs. 5 and 6).

We also discussed how these results can be extended to describe clustering from Gaussian initial conditions. The additive coagulation process and the well-studied excursion set descriptions of clustering are equivalent for Poisson and $n = 0$ initial conditions (section 6.1). However, the relation between these two approaches in the case of clustering from more general Gaussian initial conditions is more complicated. This is mainly because the excursion set statements and the assumption that, in the limit of small time steps, binary mergers dominate, are formally equivalent only for $n = 0$ initial conditions (section 6.2). This means that the binary merger coagulation equation and the excursion set approach cannot provide equivalent descriptions of clustering from more general initial conditions.

ACKNOWLEDGMENTS

It is a pleasure to thank D. Aldous, D. Earn, I. M. Hook, & G. Lemson for stimulating discussions. RKS thanks Avishai Dekel and the Racah Institute of Physics at the Hebrew University, Jerusalem, for their hospitality during the completion of this work. JP is supported in part by NSF grant DMS-9404345.

REFERENCES

Aldous D.J. 1991, Ann. Probability 19, 1

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Borel E., 1942, Compt. rend., 214, 452
Consul P. C., 1989, Generalized Poisson Distributions: Properties and Applications. M. Dekker, New York
Drake R. L., 1972, in Topics in Current Aerosol Research, eds., G. M. Hiddy and J. R. Brock, 3, 201
Efstathiou G., Rees M. J., 1988, MNRAS, 230, 5P
Epstein R. I., 1983, MNRAS, 205, 207
Ernst M. H., 1986, in Fractals in Physics, eds., L. Pietronero and L. Tosatti, North Holland, Amsterdam, p. 289
Field G. B., Saslaw W. C., 1965, ApJ, 142, 568
Flory P. J., 1941a, J. Amer. Chem. Soc., 63, 3083
Flory P. J., 1941b, J. Amer. Chem. Soc., 63, 3091
Flory P. J., 1941c, J. Amer. Chem. Soc., 63, 3096
Cornell University Press, Ithaca
Haehnelt M., Rees M. J., MNRAS, 263, 168
Hendriks E. M., Spouge J. L., Eibl M., Schreckenberg M., 1985, Z. Phys. B. Cond. Matter 58, 219
Jedamzik K., 1995, ApJ, 448, 1
Kauffmann G., White S. D. M., 1993, MNRAS, 261, 921
Kingman J., 1982, Stoch. Proc. Appl., 13, 235
Kitayama T., Suto Y., 1996, MNRAS, in press
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lacey C., Cole S., 1994, MNRAS, 271, 676
Lu B., 1987, J. Phys. A: Math Gen., 20, 2347
Lushnikov A. A., 1978a, Izv. Atmos. Ocen Phys., 14, 738
Lushnikov A. A., 1978b, J. Colloid Interface. Sci., 65, 276
Marcus A. H., 1968, Technometrics, 10, 133
McLeod J. B., 1962a, Q. J. Math, 13, 119
McLeod J. B., 1962b, Q. J. Math, 13, 192
McLeod J. B., 1962c, Q. J. Math, 13, 283
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
Moon J. W., 1967, in Harary F., ed., Graph Theory and Theoretical Physics, Academic Press, p. 261
Moon J. W., 1971, J. Comb. Th., 10, 201
Pavlov Yu. L., 1977, Th. Prob. Appl., 22, 509
Peebles P. J. E., 1980, The Large Scale Structure of the Universe. Princeton Univ. Press, Princeton
Pitman J., 1996, Berk. Stat. Tech. Rep., 457
Press W., Schechter P., 1974, ApJ, 187, 425
Riordan J., 1979, Combinatorial Identities. Robert E. Krieger Publ. Co., New York
Scott W. T., 1968, J. Atmos. Sci., 25, 54
Sheth R. K., 1995, MNRAS, 276, 796
Sheth R. K., 1996, MNRAS, 281, 1277
Silk J., Takahashi T., 1979, ApJ, 229, 242
Silk J., White S. D. M., 1978, ApJ, 223, L59
Spouge J. L., 1983a, J. Phys. A: Math Gen., 16, 3127
Spouge J. L., 1983b, J. Stat. Phys., 31, 363
Spouge J. L., 1983c, Proc. R. Soc. Lon. A, 387, 351
Spouge J. L., 1984, Adv. Appl. Prob., 16, 56
Spouge J. L., 1985, J. Stat. Phys., 38, 573
Stockmayer W. H., 1943, J. Chem. Phys., 11, 45
Tanner J. C., 1953, Biometrika, 40, 58
Tanner J. C., 1961, Biometrika, 48, 222
Trubnikov B. A., 1971, Soviet Phys.–Doklady, 16, 124
van Dongen P. G. J., 1987, J. Stat. Phys., 49, 927
von Smoluchowski M., 1916, Phys. Z., 17, 585
von Smoluchowski M., 1917, Z. Phys. Chem., 92, 129
Ziff R. M., 1980, J. Stat. Phys., 23, 241

APPENDIX A: EQUIVALENCE BETWEEN THE SPLIT RULE OF SECTION 3 AND THE ADDITIVE COALESCENT

This Appendix shows the equivalence of the split rule (equation 10) and the additive merger rule, as promised in section 3. See Pitman (1996) for a more detailed treatment.

Refer to the second paragraph of section 3 for the definition of an n-tree, and recall that each n-tree has $n-1$ branches. Let $\mathcal{T}_n$ denote the set of n-trees. Cayley showed that the number of distinct trees in this set is n^{n-2}. Let $t \in \mathcal{T}_n$ denote one member of this set. Define a forest over $[n]$ as a graph with vertex set n whose connected components form a collection of trees labeled by the sets of some partition of $[n]$. Consider forests in which there are exactly k trees. Let $\mathcal{F}_{kn}$ denote the set of forests consisting of k trees over $[n]$. Suppose that $f_k \in \mathcal{F}_{kn}$ consists of k trees of sizes $n_1, \cdots, n_k$, where n_i denotes the number of vertices in the ith tree, so that $\sum_i n_i = n$. Moon (1967) showed that the number $N(f_k)$ of trees $t \in \mathcal{T}_n$ which contain the particular forest f_k is

$$N(f_k) = \left(\prod_i n_i\right) n^{k-2}. \tag{A1}$$

A forest F_k is determined by choosing a tree $t \in \mathcal{T}_n$, and then erasing a subset of $k-1$ branches of t. Since an n-tree has $n-1$ branches, there are $n^{n-2}\binom{n-1}{k-1}$ possible choices for F_k. Of these, the number for which $F_k = f_k$ is the number of trees $t \in \mathcal{T}_n$ which contain f_k, so it is given by equation (A1). If all choices of forest are equally likely, then the probability that $F_k = f_k$ is

$$p(F_k = f_k) = \frac{\left(\prod_i n_i\right) n^{k-2}}{n^{n-2}\binom{n-1}{k-1}} = \frac{\left(\prod_i n_i\right)}{n^{n-k}\binom{n-1}{k-1}}. \tag{A2}$$

Suppose that F_k is derived from F_{k-1} by deleting a random branch from the set of branches in F_{k-1}. (So, F_1 is a forest containing just one tree, namely, one of the trees $t \in \mathcal{T}_n$, and F_n is the trivial forest of n disconnected vertices.) This deletion will split one of the trees in the forest F_{k-1} into two smaller trees. Suppose that $F_k = f_k$ was made by taking a tree with $n_i + n_j$ vertices in the forest $F_{k-1} = f_{k-1}$ and splitting it into two trees n_i and n_j. If all remaining branches in the forest F_{k-1} were equally likely to be chosen for deletion, then

$$p(F_k = f_k | F_{k-1} = f_{k-1}) = \frac{1}{n-1-(k-2)}, \tag{A3}$$

since one and only one choice allows one to to go from the forest $F_{k-1} = f_{k-1}$ to the forest $F_k = f_k$. Moreover, notice that equation (A2) implies that

$$\begin{aligned} \frac{p(F_{k-1} = f_{k-1})}{p(F_k = f_k)} &= \frac{(n_i + n_j)\left(\prod_{l\neq i,j} n_l\right)}{n^{n-k+1}\binom{n-1}{k-2}} \bigg/ \frac{\left(\prod_l n_l\right)}{n^{n-k}\binom{n-1}{k-1}} \\ &= \frac{(n_i + n_j)}{n_i\, n_j}\,\frac{(n-k+1)}{n\,(k-1)}. \end{aligned} \tag{A4}$$

Now, Bayes' rule requires that

$$\begin{aligned} p(F_{k-1} = f_{k-1} | F_k = f_k) &= p(F_k = f_k | F_{k-1} = f_{k-1}) \\ &\quad \times \frac{p(F_{k-1} = f_{k-1})}{p(F_k = f_k)} \end{aligned} \tag{A5}$$

Substituting the previous expressions in this one implies that

$$p(F_{k-1}=f_{k-1}|F_k=f_k)=\frac{(n_i+n_j)}{n\,(k-1)}\,\left(\frac{1}{n_i\,n_j}\right). \tag{A6}$$

Suppose we think of this expression as describing the probability that a tree with n_i particles merged with one containing n_j particles. Then this merger is the result of two choices. The first is the choice of the trees n_i and n_j, and the second is the choice of the two vertices, one from each tree, where the trees were actually connected. If any vertex of a tree is as likely to be chosen as any other, then the probability that any one choice occured is $1/(n_i\,n_j)$, the second term on the right of equation (A6). Therefore, the first term on the right of equation (A6) describes the probability that the two trees chosen had n_i and n_j vertices, respectively.

This process is equivalent to the one described in section 3. Thus, this appendix shows that the time reverse process of the split-rule discussed there is equivalent to the process in which the probability that two clumps merge is proportional to the sum of their masses.

APPENDIX B: SMOLUCHOWSKI'S COAGULATION EQUATION WITH ADDITIVE RATE LAW AND ARBITRARY (POLYDISPERSE) INITIAL CONDITIONS

The discrete version of Smoluchowski's coagulation equation, with merger rate $S_{ij}=i+j$ and arbitrary (polydisperse) initial conditions, is solved in this Appendix. This problem was considered by Lu (1987), and the formulation here follows Lu's closely. However, Lu's solution contains an error which we correct here.

Assume that the rate with which two objects merge is proportional to the sum of their masses, and let $c_k(t)$ denote the number density of clusters with mass k at the time t. Then Smoluchowski's equation for the rate of change of this number density is

$$\begin{aligned}\frac{\mathrm{d}\,c_k(t)}{\mathrm{d}t} &= \frac{1}{2}\sum_{i+j=k}(i+j)\,c_i(t)\,c_j(t)-c_k(t)\sum_{i=1}^{\infty}(k+j)\,c_j(t)\\ &= \frac{k}{2}\sum_{i=1}^{k}c_i(t)\,c_{k-i}(t)\\ &\qquad -c_k(t)\Big(k\,M_0(t)+M_1(t)\Big),\end{aligned} \tag{B1}$$

where $M_n(t)$ is the nth moment defined as

$$M_n(t)=\sum_{k=1}^{\infty}k^n\,c_k(t). \tag{B2}$$

Since we are interested in arbitrary initial conditions, at the initial time, set $c_k(t=0)=c_k(0)$. Equation (B1) can be written in terms of the evolution of the moments $M_n(t)$. Namely,

$$\frac{\mathrm{d}\,M_n(t)}{\mathrm{d}t}=\frac{1}{2}\sum_{i,j}^{\infty}\Big[(i+j)^n-i^n-j^n\Big]\,(i+j)\,c_i(t)\,c_j(t), \tag{B3}$$

where the sum is over all values of i and j that are greater than zero. The zeroth moment is the total number of clusters. Its evolution is given by

$$\frac{\mathrm{d}M_0(t)}{\mathrm{d}t}=-M_0(t), \tag{B4}$$

which means that

$$M_0(t)=M_0(0)\,\mathrm{e}^{-t}. \tag{B5}$$

Thus, as small clusters merge to form larger ones, the total number of clusters decreases exponentially as time increases.

Equation (B3) shows that the first moment is a constant. This just expresses mass conservation. We are free to choose units of mass and volume. The choice $M_1(0)=M_1(t)=1$ for all t is natural. Also, changing the unit of time simply multiplies the merger cross-section $S_{ij}=(i+j)$ by a constant. The relation between this and the more general case where $S_{ij}=A(i+j)+B$ is discussed by Lu (1987), so we will not do so here.

The evolution of the second moment is given by

$$\frac{\mathrm{d}M_2(t)}{\mathrm{d}t}=2M_2(t), \tag{B6}$$

so that

$$M_2(t)=\frac{1}{2}\Big[2M_2(0)\,\mathrm{e}^{2t}\Big]. \tag{B7}$$

This shows that $M_2(t)\to\infty$ as $t\to\infty$.

Equations (B1), (B2) and (B5) imply that the evolution of $c_1(t)$ is trivial:

$$\begin{aligned}\frac{\mathrm{d}\,c_1(t)}{\mathrm{d}t} &= -c_1(t)\,M_0(t)-c_1(t)\,M_1(t)\\ &= -c_1(t)\Big(M_0(t)+1\Big),\end{aligned} \tag{B8}$$

so that

$$\begin{aligned}c_1(t) &= \exp\left(-\int_0^t(M_0(t)+1)\,\mathrm{d}t\right)\\ &= \exp\left(-t-\int_0^t M_0(0)\mathrm{e}^{-t}\,\mathrm{d}t\right)\\ &= \exp\Big(-t-M_0(0)+M_0(0)\,\mathrm{e}^{-t}\Big).\end{aligned} \tag{B9}$$

It is convenient to define new variables $x_k(t)$ by

$$c_k(t)=x_k(t)\,\exp\left(-\int_0^t(kM_0(t)+1)\,\mathrm{d}t\right). \tag{B10}$$

If we also define

$$\tau=1-e^{-t}, \tag{B11}$$

then equation (B1) becomes

$$\frac{\mathrm{d}\,x_k(t)}{\mathrm{d}\tau}=\frac{1}{2}\sum_{i+j=k}(i+j)\,x_i(t)\,x_j(t). \tag{B12}$$

This means that equation (B9) implies $x_1(t)=c_1(0)$.

Rather than solving this equation order by order, it is convenient to introduce the generating function

$$g(z,t)=\sum_{k=1}^{\infty}x_k(t)\,\mathrm{e}^{kz}. \tag{B13}$$

Also, define

$$\phi(z) \equiv g(z,0) = \sum_{k=1}^{\infty} x_k(0)\,\mathrm{e}^{kz}. \tag{B14}$$

Equations (B1), (B11) and (B13) imply that

$$\frac{\partial g}{\partial \tau} = \left(\frac{\partial g}{\partial z}\right) g, \tag{B15}$$

which means that g must have the form

$$g = \phi(z + \tau g). \tag{B16}$$

Again, it is convenient to change variables. If we define

$$s \equiv z + \tau\phi(s), \quad \xi \equiv \mathrm{e}^{z}, \quad \text{and} \quad y \equiv \mathrm{e}^{s}, \tag{B17}$$

then

$$g(z,t) = \phi(s) \quad \text{and} \quad \xi = y\,\mathrm{e}^{-\tau\phi(y)}. \tag{B18}$$

This means that

$$y = \xi\,\mathrm{e}^{\tau\phi(y)} = \xi\,\mathrm{e}^{\tau g}. \tag{B19}$$

Now, if

$$y = a + \xi\psi(y), \tag{B20}$$

then Lagrange's formula for the inversion of series (e.g. Whitaker & Watson 1929) shows that

$$f(y) = f(a) + \sum_{n=1}^{\infty} \frac{\xi^n}{n!}\frac{\mathrm{d}^{n-1}}{\mathrm{d}y^{n-1}}\Big(f'(y)\,[\psi(y)]^n\Big)\Big|_{y=a}. \tag{B21}$$

If we set $f(y) = \phi(y) = g(\xi,t)$, $a = 0$, and $\psi(y) = \mathrm{e}^{\tau\phi(y)}$, as suggested by equations (B18) and (B19), then Lagrange's formula allows us to express g in terms of a power series. In particular, this means that $f'(y) = \phi'(y)$, so that

$$\begin{aligned} g(\xi,t) &= \phi(0) + \sum_{k=1}^{\infty}\frac{\xi^k}{k!}\left[\left(\frac{\mathrm{d}}{\mathrm{d}y}\right)^{k-1}\phi'(y)\left(\mathrm{e}^{\tau\phi(y)}\right)^k\right]_{y=0} \\ &= \sum_{k=1}^{\infty}\frac{\xi^k}{k!}\frac{1}{k\tau}\left[\left(\frac{\mathrm{d}}{\mathrm{d}y}\right)^{k-1}\left(k\tau\phi'\mathrm{e}^{k\tau\phi}\right)\right]_{y=0} \\ &= \sum_{k=1}^{\infty}\frac{\xi^k}{k!}\frac{1}{k\tau}\left(\frac{\mathrm{d}}{\mathrm{d}y}\right)^{k}\Big[\exp\ \tau k\,\phi(y)\Big]_{y=0} \\ &= \sum_{k=1}^{\infty}\frac{\xi^k}{k!}\frac{1}{k\tau}\sum_{m=1}^{k}(k\tau)^m \\ &\qquad \times \sum_{m\text{ parts.}}\frac{k!}{n_1!\cdots n_k!}\,c_1(0)^{n_1}\cdots c_k(0)^{n_k}, \end{aligned} \tag{B22}$$

where $\sum_j n_j = m$ and $\sum_j j\,n_j = k$, and the innermost sum in the final expression is over all distinct partitions of k into m parts, as required by the rule of differentiation of composite functions. [This solution is different from that given by Lu (1987), since Lu's equation (56) for the Lagrange formula does not contain the $\phi'(y)$ factor. This error led Lu to conclude that the final sum runs over all partitions of $k-1$ instead of k.]

Equation (B22) expresses $g(z,t)$ in terms of the initial ($t = 0$) clump size distribution. However, equations (B13) and (B17) show that $g(z,t)$ is a power series involving $x_k(t)$, so it is related to the clump size distribution at time t. Thus, we can now express $c_k(t)$ in terms of the distribution at $t = 0$. In particular, if $h(k,t)$ denotes the probability that, at t, a randomly chosen clump has exactly k particles, so that

$$h(k,t) \equiv c_k(t)/M_0(t), \tag{B23}$$

then equation (B22) implies that

$$\begin{aligned} h(k,t) &= \sum_{m=1}^{k}\Big[k\tau M_0(0)\Big]^{m-1}\frac{\mathrm{e}^{-k\tau M_0(0)}}{m!} \\ &\qquad \times \sum_{m\text{ parts.}}\frac{m!}{n_1!\cdots n_k!}\prod_{j=1}^{k}h(j,0)^{n_j}. \end{aligned} \tag{B24}$$

This is the solution to the additive coagulation equation with polydisperse initial conditions.

Let X denote a random variable that has the same distribution as the initial clump size distribution. That is, assume that $P(X = k) = h(k,0)$. Also, let $P_{h_0}(S_m = k)$ denote the distribution of $S_m = X_1 + \cdots + X_m$. Then

$$h(k,t) = \sum_{m=1}^{k}\Big[k\tau M_0(0)\Big]^{m-1}\frac{\mathrm{e}^{-k\tau M_0(0)}}{(m-1)!}\frac{P_{h_0}(S_m = k)}{m}. \tag{B25}$$

If $P_{h_0}(S_m = k)$ can be written explicitly, then it may be useful to re-write equation (B24) in this form instead.

It is easy to verify that this solution is consistent with other known results. If the initial distribution is a delta function, so that all particles initially have the same mass (monodisperse initial conditions), then equation (B24) reduces to the Borel distribution, as it should (e.g. Scott 1968; Lushnikov 1978b; van Dongen 1987). That is, if $c_1(0) = 1$ and $c_k(0) = 0$ otherwise, then setting $b = \tau = 1 - \mathrm{e}^{-t}$ in equation (B24) means that $c_k(t) = (1-b)\eta(k,b)$, where $\eta(k,b)$ is given by equation (1).

If the initial distribution is Borel with parameter b_1 [so $c_k(0) = (1-b_1)\eta(k,b_1)$], then it is instructive to set $\tau = 1 - \mathrm{e}^{-t} = (b_2 - b_1)/(1 - b_1)$, and to define t_1 and t_2 by $b_1 = \tau_1 = 1 - \mathrm{e}^{-t_1}$, and similarly for t_2 and b_2. Then $\mathrm{e}^{-t} = (1-b_2)/(1-b_1) = \mathrm{e}^{-t_2}/\mathrm{e}^{-t_1} = \mathrm{e}^{-(t_2-t_1)}$. This implies that $t = t_2 - t_1$. However, equation (B24) shows that, for this choice of initial distribution, $c_k(t) = (1-b_2)\eta(k,b_2)$. In other words, the additive coagulation process which starts with a Borel distribution with parameter b_1, differs from the initially monodisperse process (a Borel distribution with $b_0 = 0$ say) only by a time shift. This is, of course, sensible.

Finally, notice that when $\tau = (b_2-b_1)/(1-b_1)$ then the terms in the partition sum in equation (B24) look just like those in the partition probability function of equation (6). However, equation (6) was derived from a Poisson Galton–Watson branching process (Sheth 1996). So, the correspondence between equations (B24) and (6) shows explicitly that the additive Smoluchowski coagulation equation and the Poisson Galton–Watson branching process are consistent.

The relation between Galton–Watson branching processes and Smoluchowski's equation has been noted previously (e.g. Ziff 1980). Also, Spouge (1983c) used the fact that branching processes provide an efficient way to enumerate tree-partitions of integers to provide implicit solutions to the coagulation equation (Spouge 1983a, 1984, 1985).

APPENDIX C: SELF-SIMILAR SOLUTIONS TO SMOLUCHOWSKI'S EQUATION

This Appendix considers the self-similar excursion set mass functions described by equation (18); i.e., those for which $\sigma^2(M) \propto 1/M^\alpha$. It shows that only when $\alpha = 1$ are these mass functions solutions to Smoluchowski's equation with additive rate law. In the excursion set description, this corresponds to the case of white-noise initial conditions.

Let x denote the mass of an object, and τ play the role of the time variable. If we set $x_*(\tau) \propto (\mathrm{e}^\tau)^{2/\alpha}$, then, in terms of scaled variables $u = x/x_*(\tau)$, the mass functions have the form

$$n(x,\tau) = \frac{1}{x_*(\tau)^2}\,\phi\left(\frac{x}{x_*(\tau)}\right), \tag{C1}$$

where

$$\phi(u) = \frac{\alpha}{\sqrt{2\pi}}\,\frac{u^{\alpha/2}}{u^2}\,\exp(-u^\alpha/2), \tag{C2}$$

where the relation to the excursion set mass functions is seen by setting $\alpha = (n+3)/3$; we will only consider $4/3 \geq \alpha \geq 1/3$. Notice that $\int u\,\phi(u)\,\mathrm{d}u = 1$.

In these variables $x_*(\tau)^2$ times the left-hand side of Smoluchowski's equation is

$$-\frac{2}{\alpha}\left[2\phi(u) + u\phi'(u)\right] = \phi(u)\,(u^\alpha - 1) \tag{C3}$$

and $x_*(\tau)^2$ times the right-hand side is

$$\frac{1}{2}\int_0^u \phi(v)\,\phi(u-v)\,S(v,u-v)\,\mathrm{d}v - \phi(u)\int_0^\infty \phi(v)\,S(u,v)\,\mathrm{d}v. \tag{C4}$$

In the main text we argued that Smoluchowski's equation with additive rate law has a consistent self-similar solution that corresponds to setting $\alpha = 1$. To see that no other self-similar excursion set mass functions are consistent with an additive rate law, it is useful to multiply both expressions (C3) and (C4) by u^μ and integrate over all $u \geq 0$. Then the left-hand side reads

$$\int_0^\infty \phi(u)\,(u^\alpha - 1)\,u^\mu\,\mathrm{d}u \tag{C5}$$

which can be written as a Gamma function times some normalizing factors. The other side is

$$\int_0^\infty \mathrm{d}v \int_0^\infty \mathrm{d}u\,\left[\frac{(u+v)^\mu}{2} - u^\mu\right]\,\phi(u)\,\phi(v)\,S(u,v). \tag{C6}$$

If we set $S(u,v) = (u+v)$, and expand the term in square brackets using the Binomial theorem, then this second expression can be written as a sum over products of Gamma functions. Thus, it is easy to verify that the two expressions are equal only when $\alpha = 1$. This shows explicitly that self-similar mass functions of the form (C2), with $\alpha \neq 1$ are not solutions of Smoluchowski's equation with additive rate law. In other words, if we seek self-similar solutions to Smoluchowski's equation of the form (C2) then the merger rule $S(u,v)$ must depend on α.