

# One-Dimensional Poisson Growth Models With Random and Asymmetric Growth

Fred W. Huffer  
Department of Statistics  
Florida State University  
Tallahassee, FL 32306  
huffer@stat.fsu.edu

September 4, 2002

## Abstract

Suppose there is a Poisson process of points  $X_i$  on the line. Starting at time zero, a grain begins to grow from each point  $X_i$ , growing at rate  $A_i$  to the left and rate  $B_i$  to the right, with the pairs  $(A_i, B_i)$  being iid. A grain stops growing as soon as it touches another grain. When all growth stops, the line consists of covered intervals (made up of contiguous grains) separated by gaps. We show (i) a fraction  $1/e$  of the line remains uncovered, (ii) the fraction of covered intervals which contain exactly  $k$  grains is  $(k-1)/k!$ , (iii) the length of a covered interval containing  $k$  grains has a gamma( $k-1$ ) distribution, (iv) the distribution of the grain sizes depends only on the distribution of the total growth rate  $A_i + B_i$ , and other results. Similar theorems are obtained for growth processes on a circle; in this case we need only assume the pairs  $(A_i, B_i)$  are exchangeable. These results extend those of Daley, Mallows and Shepp (2000) who studied the case where  $A_i = B_i = 1$ . Simulation results are given to illustrate the various theorems.

*Key words and phrases:* germ-grain model, non-overlapping intervals, Poisson process.

## 1 Introduction and Main Results

Daley, Mallows and Shepp (2000) (referred to as DMS below) discuss a one-dimensional Poisson growth model which is a special case of a  $d$ -dimensional growth model studied by Häggström and Meester (1996) and Daley, Stoyan and Stoyan (1999). DMS demonstrate many interesting properties of this model. The goal of this paper is to show that some of these properties remain true under more general growth conditions.

The model we investigate is now described. Let  $\{X_i : i \in \mathbf{Z}\}$  be the ordered points of a Poisson process on the line with rate 1. (Here  $\mathbf{Z}$  denotes the set of all integers.)

The points  $X_i$  are called ‘germs’. A ‘grain’ grows from each germ  $X_i$ . Starting at time zero, grain  $i$  (located at  $X_i$ ) grows at rate  $A_i$  to the left, and at rate  $B_i$  to the right. The growth rates  $\{(A_i, B_i) : i \in \mathbf{Z}\}$  satisfy  $A_i \geq 0$ ,  $B_i \geq 0$ , and  $A_i + B_i > 0$  for all  $i$ . Grain  $i$  stops growing (in both directions) as soon as it touches another grain. We are interested in the situation on the line after all growth has stopped. The line now consists of a series of covered intervals (consisting of contiguous grains) separated by gaps.

Each covered interval grows in the following fashion. First, two grains meet to form the nucleus (which we call the initial doublet) of the interval. Then subsequently, the interval grows as grains join it one at a time on either the left or right hand side of the interval. [These intuitive remarks follow from Lemma 10 of DMS which continues to hold in our more general situation so long as we restate the lemma in terms of the times at which grains stop growing instead of in terms of their radii.] Some covered intervals will contain only two grains (the initial doublet).

We consider three sets of assumptions A1, A2, A3 on the growth rates  $(A_i, B_i)$  which grow increasingly more restrictive. We assume throughout that the random variables  $\{(A_i, B_i) : i \in \mathbf{Z}\}$  are independent of the Poisson process  $\{X_i : i \in \mathbf{Z}\}$ .

**A1** The random variables  $\{(A_i, B_i) : i \in \mathbf{Z}\}$  have an arbitrary joint distribution.

**A2** The pairs  $(A_i, B_i)$  are iid (The bivariate distribution of  $(A_i, B_i)$  is arbitrary.)

**A3** The pairs  $(A_i, B_i)$  are constructed as follows. Let  $\{R_i : i \in \mathbf{Z}\}$  be iid random variables with  $P(R_i > 0) = 1$ , and  $\{S_i : i \in \mathbf{Z}\}$  be iid random variables with  $P(0 \leq S_i \leq 1) = 1$ . Assume the sequences  $\{R_i\}$  and  $\{S_i\}$  are independent. Take  $A_i = S_i R_i$  and  $B_i = (1 - S_i) R_i$  for all  $i$ .

We shall consider the following list of properties.

**P1** The distribution of the length of a randomly chosen covered interval containing  $k$  contiguous grains is  $\text{gamma}(k-1)$ , the distribution of the sum of  $k-1$  independent exponential random variables with mean 1.

**P2** The fraction of the line which is *not* covered is  $1/e$ .

**P3** The probability that a randomly chosen covered interval contains exactly  $k$  contiguous grains is  $(k-1)/k!$ .

**P4** The distribution of the length of the grain grown from a randomly sampled germ and the distribution of the length of the initial doublet in a randomly chosen covered interval depend only on the distribution of  $A_i + B_i$ .

**P5** Given that a randomly chosen covered interval contains  $k$  grains, the number of grains to the right of the initial doublet has a  $\text{binomial}(k-2, p)$  distribution with  $p = ES_j$ .

In this paper, we shall prove the following results:

## Theorem 1:

- (a) P1 is true under A1.
- (b) P1 to P4 are true under A2.
- (c) P1 to P5 are true under A3.

DMS consider the special case of A3 in which  $R_i$  and  $S_i$  are nonrandom with  $R_i = 2$  and  $S_i = 1/2$ , that is, all the grains grow at rate 1 in both directions. Under these assumptions they prove P1, P2, P3, and much more. The contribution of this paper is to allow the grains to grow asymmetrically and to have random growth rates. As immediate consequences of Theorem 1 we note the following. Theorem 8 of DMS gives the distribution of the length of a randomly chosen covered interval. This theorem remains true under A2 because both P1 and P3 are true under A2. Theorem 3 of DMS gives the length of a randomly chosen grain. This continues to hold under A2 so long as  $P(A_i + B_i = c) = 1$  for some constant  $c$ . This follows from P4.

O’Hely and Sudbury (2001) and Sudbury (2002) study an ‘annihilating process’ which is closely related to the growth model of this paper. (We indicate the connection in Section 4.2.) For this process they prove results corresponding to P2 and P3.

We shall prove Theorem 1 by considering a finite version of the growth process in which there are  $n$  random germs on a circle. In Section 2 we describe this finite growth process and state finite analogs of the various results in Theorem 1. Section 3 presents simulation studies of the finite growth process which illustrate the results of Section 2. Section 4 contains proofs of the results in Section 2.

## 2 Results for the Circle

Our approach in the proofs is to prove finite versions of the results for a situation in which there are  $n$  random germs on a circle with the spacings between the germs given by iid exponential random variables (so that the circle has a random circumference). Letting  $n \rightarrow \infty$  in these results then gives the desired results for the Poisson process of germs on the line. DMS also consider the growth process on the circle and, under their assumptions, prove the property Q2 below. O’Hely and Sudbury (2001) study a version of their annihilating process with a finite number of sites arranged in a circle. For this process, they, too, prove Q2.

Here is our growth model. Let  $Z_1, Z_2, \dots, Z_n$  be iid exponential random variables with mean 1. Let  $X_k = Z_1 + Z_2 + \dots + Z_k$  for  $1 \leq k \leq n$ . Wrap the interval  $[0, X_n]$  into a circle, identifying the points 0 and  $X_n$ . This gives us a circle with  $n$  germs at points  $X_1, X_2, \dots, X_n (= 0)$ . Germs  $i - 1$  and  $i$  are separated by a gap of length  $Z_i$ . From each germ there grows a grain exactly as described in Section 1.

Two notational remarks: When describing directions on the circle we will use the words right and left instead of the more cumbersome clockwise and counterclockwise. Also, the numbering of the germs on the circle will be modulo  $n$ . In particular, germs 0 and 1 are the same as germs  $n$  and  $n + 1$  respectively.

The assumptions we need for the results on the circle are B1, B2, B3 given below. These assumptions are slightly weaker than A1, A2, A3 in that we have replaced independence by exchangeability in two places.

- B1** The random variables  $\{(A_i, B_i) : 1 \leq i \leq n\}$  have an arbitrary joint distribution.
- B2** The pairs  $\{(A_i, B_i) : 1 \leq i \leq n\}$  are exchangeable. (The bivariate distribution of  $(A_i, B_i)$  is arbitrary.)
- B3** The pairs  $(A_i, B_i)$  are constructed as follows. Let  $\{R_i : 1 \leq i \leq n\}$  be exchangeable random variables with  $P(R_i > 0) = 1$ , and  $\{S_i : 1 \leq i \leq n\}$  be iid random variables with  $P(0 \leq S_i \leq 1) = 1$ . Assume the sequences  $\{R_i\}$  and  $\{S_i\}$  are independent. Take  $A_i = S_i R_i$  and  $B_i = (1 - S_i) R_i$  for all  $i$ .

In order to state the properties of the germ-grain process on the circle, we need some notation. Let  $\mathcal{I}$  be the covered interval containing grain 1. Define the events  $F$  and  $F_k$  as follows:  $F$  is the event that  $\mathcal{I}$  begins with grain 1, that is, grain 1 is the leftmost grain in  $\mathcal{I}$ .  $F_k$  is the event that  $\mathcal{I}$  begins with grain 1 and contains exactly  $k$  grains. Finally, define  $\pi_k = \sum_{i=0}^k (-1)^i / i!$ .

For the growth process on the circle, the analogs of properties P1 to P5 are the following:

- Q1** Given  $F_k$ , the length of  $\mathcal{I}$  has a gamma( $k - 1$ ) distribution.
- Q2** The expected fraction of the circle which is *not* covered is  $\pi_n$ .
- Q3**  $P(F_k | F) = \frac{k - 1}{k!} \cdot \frac{\pi_{n-k}}{\pi_n}$  for  $2 \leq k \leq n$ .
- Q4** The distribution of the length of the grain grown from a randomly sampled germ and the distribution of the length of the initial doublet in a randomly chosen covered interval depend only on the joint distribution of  $(A_1 + B_1, \dots, A_n + B_n)$ .
- Q5** Given  $F_k$ , the number of grains of  $\mathcal{I}$  to the right of the initial doublet has a binomial( $k - 2, p$ ) distribution with  $p = ES_j$ .

We will prove:

**Theorem 2:**

- (a) Q1 is true under B1.
- (b) Q1 to Q4 are true under B2.
- (c) Q1 to Q5 are true under B3.

By taking  $n \rightarrow \infty$ , it is clear that Theorem 2 implies Theorem 1.

### 3 Examples

We now present some simulation studies to illustrate the results of the previous section. In these simulations we use  $n = 8$  grains on the circle. We choose a small value of  $n$  for two reasons. First, this makes it possible to distinguish easily between exchangeable and iid random variables, and thus enables us to give a more convincing demonstration that only exchangeability is needed in B2 and the first part of B3. Secondly, with a small value of  $n$  we can see the effect of the factor  $\pi_{n-k}/\pi_n$  in Q3. This factor is essentially 1 for large  $n$ . Some additional simulation results, using a large value of  $n$ , are given in an appendix to this report.

We consider five scenarios denoted (a), (b), (c), (d), (e). In scenarios (a), (b), and (c), the growth rates  $(A_i, B_i)$  of the grains are either  $(1, 0)$  or  $(0, 1)$ , that is, all the grains grow at rate 1, but some grow only to the left, and others only to the right. In scenario (d), the grains still grow only to the right or to the left, but the growth rates vary. In scenario (e), the grains grow at equal rates in both directions.

Here are the five scenarios in detail.

- (a)  $(A_i, B_i) = (0, 1)$  for  $i = 1, 3, 5, 7$  and  $(A_i, B_i) = (1, 0)$  for  $i = 2, 4, 6, 8$ , that is, we alternate between left-growing and right-growing grains.
- (b) Exactly 4 grains grow to the left ( $(A_i, B_i) = (1, 0)$ ) and 4 to the right ( $(A_i, B_i) = (0, 1)$ ), but the ordering of the 8 grains is random.
- (c) The growth rates are iid with  $P\{(A_i, B_i) = (1, 0)\} = P\{(A_i, B_i) = (0, 1)\} = 1/2$ .
- (d) The growth rates  $(A_i, B_i)$  are a random permutation of  $(x, 0)$ ,  $(x, 0)$ ,  $(y, 0)$ ,  $(y, 0)$ ,  $(0, x)$ ,  $(0, x)$ ,  $(0, y)$ ,  $(0, y)$  where  $x = 0.25$  and  $y = 1.75$ .
- (e) The growth rates  $(A_i, B_i)$  are a random permutation of  $(v, v)$ ,  $(v, v)$ ,  $(v, v)$ ,  $(v, v)$ ,  $(w, w)$ ,  $(w, w)$ ,  $(w, w)$ ,  $(w, w)$  where  $v = 0.125$  and  $w = 0.875$ .

It is easily seen that (a) satisfies B1, but not B2; (b) and (d) satisfy B2, but not B3; and (c) and (e) satisfy B3. In (c) we have  $P(R_i = 1) = 1$  and  $P(S_i = 0) = P(S_i = 1) = 1/2$ . In (e),  $P(R_i = 2v) = P(R_i = 2w) = 1/2$  and  $P(S_i = 1/2) = 1$ . In all the scenarios, the average of the total growth rate  $A_i + B_i$  is 1, but note that rescaling the growth rates (that is, replacing  $(A_i, B_i)$  by  $(cA_i, cB_i)$  for all  $i$ ) does not affect the outcome of the growth process.

To compare these scenarios, we carried out 10,000,000 simulations of each. The average fraction of the circle left uncovered in each situation is given in Table 1. Since

(a)	(b)	(c)	(d)	(e)	$\pi_8$
0.50002	0.36786	0.36795	0.36785	0.36792	0.36788

Table 1: Average fraction of circumference not covered.

$\pi_8 \approx 0.36788$ , the simulation results for models (b) to (e) are seen to be in good agreement with Q2. The results for model (a) are easily understood after a little thought. Under this model, there will always be exactly 4 covered intervals, each containing two grains. The lengths of the covered intervals and the lengths of the uncovered gaps between them are iid exponential random variables with mean 1. Thus, the expected fraction of the circle left uncovered is exactly  $1/2$ .

In order to study probabilistic properties of the covered intervals (such as their lengths, the number of grains they contain, etc.), we collect together all the covered intervals created in the 10,000,000 simulations. Properties Q1, Q3, and Q5 provide the relevant distributions for a randomly selected interval from this collection (as in the wording of properties P1, P3, and P5 in Section 1). Table 2 gives the distribution of the number of grains in a randomly selected interval for each of the scenarios. The values predicted by Q3 are given in the last column. As we expect from Theorem 2,

$k$	(a)	(b)	(c)	(d)	(e)	$\frac{k-1}{k!} \cdot \frac{\pi_{8-k}}{\pi_8}$
2	1.0	0.50007	0.50020	0.50014	0.50024	0.49998
3	0.0	0.33250	0.33228	0.33238	0.33223	0.33242
4	0.0	0.12731	0.12741	0.12728	0.12742	0.12747
5	0.0	0.03021	0.03022	0.03026	0.03020	0.03024
6	0.0	0.00945	0.00941	0.00947	0.00944	0.00941
7	0.0	0.00000	0.00000	0.00000	0.00000	0.00000
8	0.0	0.00047	0.00047	0.00048	0.00047	0.00048

Table 2: Distribution of the number  $k$  of grains in a covered interval.

the results for models (b) to (e) agree well with the values from Q3.

According to Theorem 2, property Q1 should hold in all five scenarios; the length of a randomly chosen covered interval containing  $k$  grains should follow a  $\text{gamma}(k-1)$  distribution. As one way of investigating this property, we use our simulated data to compute the first three sample moments of these lengths. For models (b), (c), and (d), these sample moments (denoted  $\mu_1, \mu_2, \mu_3$ ) are displayed in Table 3. In model (a),

$k$	(b)			(c)			(d)		
	$\mu_1$	$\mu_2$	$\mu_3$	$\mu_1$	$\mu_2$	$\mu_3$	$\mu_1$	$\mu_2$	$\mu_3$
2	1.000	2.002	6.008	1.000	2.000	5.998	1.000	2.000	6.002
3	2.000	6.001	24.012	1.999	5.998	23.996	2.000	5.998	23.990
4	3.000	11.998	60.011	2.999	11.993	59.941	3.001	12.008	60.077
5	4.002	20.004	119.880	4.005	20.042	120.286	4.001	20.010	120.061
6	4.992	29.900	208.854	4.997	29.982	210.009	4.998	29.989	210.063
8	7.002	55.914	501.478	6.992	55.824	501.183	7.008	55.995	501.863

Table 3: The moments of the lengths of the covered intervals containing  $k$  grains.

all the covered intervals contain  $k = 2$  grains. The sample moments for these lengths were  $\mu_1 = 1.000, \mu_2 = 2.001, \mu_3 = 6.002$ . (We have omitted model (e) to save space.) All these moments are in good agreement with the moments of the  $\text{gamma}(k-1)$  distribution given by  $\mu_j = (k-2+j)!/(k-2)!$ .

Now we examine the distributions of the lengths of the grains and the initial doublets. Tables 4 and 5 give the first four sample moments of these lengths for models (b)

	(a)	(b)	(c)	(d)	(e)
$\mu_1$	0.5000	0.6321	0.6320	0.6321	0.6321
$\mu_2$	0.5002	0.9698	0.9693	1.0555	1.0554
$\mu_3$	0.7503	2.5307	2.5287	2.8816	2.8823
$\mu_4$	1.5003	9.4344	9.4263	10.8747	10.8878

Table 4: Moments of the distribution of grain sizes.

to (e). All of these models satisfy B2. Models (b) and (c) both have  $A_i + B_i = 1$  with probability 1. Thus, Q4 implies they must have the same distributions for the lengths of the grains and the initial doublets. The same must hold for models (d) and (e) since they, too, have the same joint distribution for the quantities  $A_i + B_i$ , that is,  $A_i + B_i$  equals 0.25 for four values of  $i$ , and 1.75 for the other four values of  $i$ . Our simulation output is consistent with Q4; the columns for (b) and (c) are essentially identical, as are the columns for (d) and (e).

	(a)	(b)	(c)	(d)	(e)
$\mu_1$	1.0001	0.8007	0.8006	0.8452	0.8449
$\mu_2$	2.0006	1.3719	1.3712	1.5242	1.5228
$\mu_3$	6.0023	3.7045	3.7011	4.2653	4.2572
$\mu_4$	24.0047	13.8009	13.7877	16.2266	16.1670

Table 5: Moments of the length of the initial doublet.

Finally, we examine property Q5. Models (c) and (e) satisfy B3 with  $ES_i = 1/2$  so that property Q5 predicts that, for covered intervals with  $k$  grains, the number of grains to the right of the initial doublet should have a binomial( $k - 2, 1/2$ ) distribution. This is in agreement with the mean  $\mu$  and variance  $\sigma^2$  from our simulations given in Table 6. The variances for models (b) and (d) are much smaller than those for (c) and (e), and, for  $k \geq 4$ , the results for (b) and (d) are not consistent with a binomial( $k - 2, p$ ) distribution for any value of  $p$ . All the scenarios (b) to (e) are symmetric, that is, they do not change when the left and right growth rates are interchanged. This symmetry implies that, for covered intervals containing  $k$  grains, the expected number of grains to the right of the initial doublet is  $(k - 2)/2$ . This fact is apparent in Table 6.

## 4 Proofs

### 4.1 Independent Exponentials

Before beginning the proofs we state some properties of independent exponential random variables that we shall need. Let  $U \sim \exp(\theta)$  mean that  $U$  has an exponential distribution with mean  $1/\theta$ .

Let  $U_1, U_2, \dots, U_k$  be independent with  $U_i \sim \exp(\theta_i)$ . Define  $D$  to be the index of the minimum  $U_i$ , that is,  $D = j$  if and only if  $U_j = \min_i U_i$ . The following standard

$k$	(b)		(c)		(d)		(e)	
	$\mu$	$\sigma^2$	$\mu$	$\sigma^2$	$\mu$	$\sigma^2$	$\mu$	$\sigma^2$
3	0.500	0.250	0.500	0.250	0.500	0.250	0.500	0.250
4	1.000	0.429	1.000	0.500	1.000	0.423	1.000	0.500
5	1.499	0.536	1.500	0.752	1.500	0.528	1.500	0.749
6	2.000	0.572	1.996	1.002	2.000	0.564	1.998	0.998
8	3.005	0.423	3.010	1.471	3.003	0.430	3.000	1.515

Table 6: The mean and variance of the number of grains to the right of the initial doublet for covered intervals with  $k$  grains.

facts are used repeatedly throughout the paper:

$$\begin{aligned} \min U_i &\sim \exp(\sum_i \theta_i), \\ P(D = j) &= \theta_j / \sum_i \theta_i, \\ \min U_i &\text{ is independent of } D. \end{aligned}$$

These facts are listed in (or easily derivable from) Lemma 2.1 of Sackrowitz and Samuel-Cahn (1984).

We also need the following. Let  $\mathbf{V} = (V_1, \dots, V_k)$  be any vector of nonnegative random variables independent of  $U_1, \dots, U_k$  given above. Then

$$\mathcal{L}(U_1 - V_1, \dots, U_k - V_k \mid U_1 > V_1, \dots, U_k > V_k, \mathbf{V}) = \mathcal{L}(U_1, \dots, U_k). \quad (1)$$

This is a consequence of the memoryless property (that is,  $\mathcal{L}(U - c \mid U > c) = \mathcal{L}(U)$  for all  $c > 0$ ) which, conditional on  $\mathbf{V}$ , may be applied separately to each of the  $U_i$ ,  $i = 1, \dots, k$ .

Finally, we need that

$$\mathcal{L}(\sum_i \theta_i U_i \mid U_1 < U_2 < \dots < U_k) = \mathcal{L}(\sum_i \theta_i U_i) = \text{gamma}(k). \quad (2)$$

This fact may be derived from Lemma 2.1(ii) of Sackrowitz and Samuel-Cahn which states that, conditional on  $U_1 < U_2 < \dots < U_k$ , the differences  $Y_i = U_i - U_{i-1}$ ,  $i = 1, \dots, k$  (with  $U_0 \equiv 0$ ) are independent exponential random variables with parameters  $\beta_i = \sum_{j=i}^k \theta_j$ ,  $i = 1, \dots, k$ . Using this, fact (2) follows immediately from the algebraic identity  $\sum_{i=1}^k \theta_i U_i = \sum_{i=1}^k \beta_i Y_i$ .

Fact (2) underlies property Q1; it says that the (unconditional) distribution of  $\sum_i \theta_i U_i$  is the same as the conditional distribution given *complete* information about the ordering of the values  $U_i$ . But if this is so, it must also be true that the unconditional distribution coincides with the conditional distribution given any *partial* information about the ordering of the  $U_i$ 's. In particular, if the integers  $j_1, \dots, j_p$  and  $k_1, \dots, k_p$  are such that

$$P(U_{j(i)} < U_{k(i)} \text{ for } i = 1, \dots, p) > 0,$$

then we have

$$\mathcal{L}\left(\sum_i \theta_i U_i \mid U_{j(i)} < U_{k(i)} \text{ for } i = 1, \dots, p\right) = \mathcal{L}(\sum_i \theta_i U_i) = \text{gamma}(k). \quad (3)$$

Here and in the sequel we use parentheses to avoid double subscripts ( $j(i) = j_i$ ).

## 4.2 The “Gap” Description

We shall assume initially that the growth rates  $(A_i, B_i)$  are constants. The results we obtain may be regarded as holding conditionally given the growth rates. Let  $(a_i, b_i)$ ,  $i = 1, \dots, n$  be constants satisfying  $a_i \geq 0$ ,  $b_i \geq 0$ , and  $a_i + b_i > 0$  for all  $i$ . Suppose the grain at germ  $i$  grows at rate  $a_i$  to the left and rate  $b_i$  to the right.

There are two different ways to describe the growth process. One way focuses on the gaps between the grains. The other describes the process more in terms of the grains. The “gap” description is the more straightforward of the two approaches, but the “grain” description is more useful in the proofs. In the following, we present these two approaches and demonstrate their equivalence.

The  $i$ -th gap (between germs  $i - 1$  and  $i$ ) is gradually being covered by the grains growing on either side of it. Let  $Z_i(t)$  denote the length of the remaining uncovered part of the  $i$ -th gap at time  $t$ . Note that  $Z_i(0) = Z_i$ . We shall focus on the lengths  $Z_i(t)$ . We say the  $i$ -th gap “dies” at time  $\tau$  if  $\tau = \inf\{t : Z_i(t) = 0\}$ ; at time  $\tau$  the two grains adjoining this gap touch and stop growing. Let  $M$  be the random number of gaps which die during the entire course of time, and  $\tau_1 < \tau_2 < \dots < \tau_M$  be the random times at which they die. Let  $I_{jk}$  denote the indicator that the grain at germ  $j$  is growing during the period between times  $\tau_{k-1}$  and  $\tau_k$ . Define  $D_k$  to be the gap which dies at time  $\tau_k$ , that is,  $D_k = i$  if  $\tau_k = \inf\{t : Z_i(t) = 0\}$ .

Consider the situation leading up to the first death at time  $\tau_1$ . The  $i$ -th gap is being consumed at rate  $\lambda_{i1} = a_i + b_{i-1}$  by the two adjoining grains so that  $\tau_1 = \min\{Z_i/\lambda_{i1} : 1 \leq i \leq n\}$ . Define  $\lambda_1 = \sum_i \lambda_{i1} = \sum_{i=1}^n (a_i + b_i)$ . From elementary properties of the exponential distribution we know that  $\tau_1 \sim \exp(\lambda_1)$  and  $P(D_1 = i) = \lambda_{i1}/\lambda_1$  for  $i = 1, \dots, n$ . Also, the random variables  $\tau_1$  and  $D_1$  are independent.

After the closing of the first gap at time  $\tau_1$ , the memoryless property of the exponential distribution guarantees that the remaining gaps  $Z_i(\tau_1)$ ,  $i \neq D_1$ , are iid  $\exp(1)$  random variables. We give a brief argument for this. Suppose, for convenience, that  $D_1 = 1$  and thus  $\tau_1 = Z_1/\lambda_{11}$ . Since  $Z_i(\tau_1) = Z_i - \lambda_{i1}\tau_1$ , and the event  $\{D_1 = 1\}$  is equivalent to  $\{Z_1/\lambda_{11} < Z_i/\lambda_{i1} \text{ for } i \neq 1\}$ , we see that

$$\begin{aligned} & \mathcal{L}(Z_2(\tau_1), \dots, Z_n(\tau_1) \mid D_1 = 1, \tau_1 = t) \\ &= \mathcal{L}\left(Z_i - \frac{\lambda_{i1}Z_1}{\lambda_{11}}, i \neq 1 \mid Z_i > \frac{\lambda_{i1}Z_1}{\lambda_{11}}, i \neq 1, Z_1 = \lambda_{11}t\right) \\ &= \mathcal{L}(Z_2, \dots, Z_n) \end{aligned}$$

by property (1). In a similar fashion we can show that, at any time  $\tau_k$ , the remaining gaps  $Z_i(\tau_k)$ ,  $i \notin \{D_1, \dots, D_k\}$  are iid  $\exp(1)$  random variables conditional on  $D_1, \dots, D_k$  and  $\tau_1, \dots, \tau_k$ .

Thus we may repeat the earlier argument at all the times  $\tau_k$ . Consider the situation leading up to the  $k$ -th death at time  $\tau_k$ . At any time  $t$  between  $\tau_{k-1}$  and  $\tau_k$ , the  $i$ -th gap is being consumed at rate  $\lambda_{ik} = a_i I_{ik} + b_{i-1} I_{i-1,k}$  by the two adjoining grains so that  $\tau_k - \tau_{k-1} = \min\{Z_i(\tau_{k-1})/\lambda_{ik} : \lambda_{ik} > 0\}$ . Thus, conditional on  $D_1, D_2, \dots, D_{k-1}$  and  $\tau_1, \dots, \tau_{k-1}$ , we know that  $\tau_k - \tau_{k-1} \sim \exp(\lambda_k)$  with  $\lambda_k = \sum_i \lambda_{ik} = \sum_{i=1}^n (a_i + b_i) I_{ik}$ , and  $P(D_k = i) = \lambda_{ik}/\lambda_k$  for  $i = 1, \dots, n$ . Also, the random variables  $\tau_k$  and  $D_k$  are conditionally independent.

After all growth has stopped (at time  $\tau_M$ ), the remaining  $n - M$  gaps (those values of  $Z_i(\tau_M)$  which are positive) are iid  $\exp(1)$  random variables which are independent of  $\tau_1, \dots, \tau_M$  and  $D_1, \dots, D_M$ .

Define  $J_i(t) = I(Z_i(t) > 0)$  for  $i = 1, \dots, n$ ; we say that  $J_i(t)$  indicates whether the  $k$ -th gap is still “alive” at time  $t$ . The process  $(J_1(t), J_2(t), \dots, J_n(t))$  is an annihilating process of the type studied by O’Hely and Sudbury (2001). They study interacting particle systems in which neighboring particles interact only by trying to annihilate each other. The gaps interact with each other in just this way: a living gap is at risk of death if either of the neighboring gaps is still alive (for this implies that at least one of the grains adjoining it is still growing); if both neighbors are dead, the gap will persist forever. In terms of our growth processes, the results of O’Hely and Sudbury describe the situation where the growth rates satisfy  $a_i = p$  and  $b_i = 1 - p$  for all  $i$ .

### 4.3 The “Grain” Description

A useful equivalent description of the growth process is the following. Define  $r_i = a_i + b_i$  and  $s_i = a_i/r_i$  for  $i = 1, \dots, n$ . Let  $T_1, T_2, \dots, T_n$ , and  $\epsilon_1, \epsilon_2, \dots, \epsilon_n$  be independent with  $T_i \sim \exp(r_i)$  and  $\epsilon_i = \pm 1$  with probabilities  $P(\epsilon_i = -1) = s_i$  and  $P(\epsilon_i = +1) = 1 - s_i$ . Suppose that grains  $1, 2, \dots, n$  are arranged in a circle and that each grain grows according to the following rule. Grain  $i$  grows at rate  $r_i$  and continues to grow until time  $T_i$  or until the first time it is “bumped” by another grain, whichever comes first. If grain  $i$  stops growing at time  $T_i$ , then it bumps against the neighboring grain  $i + \epsilon_i$ . (If grain  $i$  stops growing before time  $T_i$  because it was bumped by another grain, then nothing happens at time  $T_i$  and the value of  $\epsilon_i$  is never used.) When two grains bump, they join together to become part of the same covered interval. After all growth has stopped, the grains have grouped themselves into covered intervals. The final lengths of these grains and intervals are implicitly determined by the above description. The lengths of the gaps between the covered intervals are *not* determined by the above description; we take them to be iid  $\exp(1)$  random variables.

It may help to think of growing the grains as  $n$  disconnected intervals. The rules above determine the final size of each grain, which grains will group together to form covered intervals, and the lengths of the gaps to place between these covered intervals. Once all growth has stopped, we can imagine bending the grains into arcs and assembling them into the circle.

For convenience in our explanations, when two grains bump we say that one of them is doing the bumping (and the other is being bumped). In particular, if grain  $i$  stops growing at time  $T_i$ , we say that grain  $i$  bumps grain  $i + \epsilon_i$  (and not vice versa).

We now show that the “grain description” of this section is equivalent to the earlier one in terms of the gaps. We reuse the variable names from the earlier description. Let  $\tau_1 < \tau_2 < \dots$  be the times at which gaps die (that is, two grains bump together) under this new description. Let  $D_j$  be the gap which dies at time  $\tau_j$ ; gap  $i$  dies at time  $\tau_j$  if grains  $i - 1$  and  $i$  bump at this time. Similarly, define  $I_{jk}$  to be the indicator that grain  $j$  is still growing at times between  $\tau_{k-1}$  and  $\tau_k$  (with  $\tau_0 = 0$ ).

Clearly  $\tau_1 = \min\{T_i : 1 \leq i \leq n\} \sim \exp(\sum_i r_i)$  and  $\sum_i r_i = \lambda_1$  (defined earlier in Section 4.2). Also,  $D_1 = i$  if either  $T_i = \tau_1$  and  $\epsilon_i = -1$ , or  $T_{i-1} = \tau_1$  and  $\epsilon_{i-1} = +1$ .

This event has probability  $(r_i/\lambda_1)s_i + (r_{i-1}/\lambda_1)(1 - s_{i-1}) = (a_i + b_{i-1})/\lambda_1 = \lambda_{i1}/\lambda_1$ . Moreover,  $\tau_1$  and  $D_1$  are independent.

Similarly, conditional on  $D_1, \dots, D_{k-1}$  and  $\tau_1, \dots, \tau_{k-1}$ , the memoryless property implies  $\tau_k - \tau_{k-1} \stackrel{d}{=} \min\{T_i : I_{ik} = 1\} \sim \exp(\sum_i r_i I_{ik})$  and  $\sum_i r_i I_{ik}$  is equal to  $\lambda_k$  defined earlier. Also,  $D_k = i$  if either  $T_i = \tau_k$  and  $\epsilon_i = -1$ , or  $T_{i-1} = \tau_k$  and  $\epsilon_{i-1} = +1$ . This event has probability  $(r_i I_{ik}/\lambda_k)s_i + (r_{i-1} I_{i-1,k}/\lambda_k)(1 - s_{i-1}) = (a_i I_{ik} + b_{i-1} I_{i-1,k})/\lambda_k = \lambda_{ik}/\lambda_k$  as defined earlier. Finally,  $\tau_k - \tau_{k-1}$  and  $D_k$  are conditionally independent.

We see that both descriptions of the growth model lead to the same joint distribution for the random variables  $\tau_1, \tau_2, \dots, D_1, D_2, \dots$ , and hence are equivalent. We will use the the grain description of the growth process to prove various results. But first, we must introduce some notation.

## 4.4 History Vectors

Every covered interval is described by a ‘history vector’  $\mathbf{h}$  which records the time order in which the grains joined to form the interval. For example, suppose a covered interval containing  $k$  grains developed in the following way: first grains  $h_1$  and  $h_2$  (with  $h_2 = h_1 + 1$ ) joined to form the nucleus (or initial doublet), and then were joined by grains  $h_3, \dots, h_k$  in that order. The history vector for this interval is simply  $\mathbf{h} = (h_1, h_2, \dots, h_k)$ .

For  $2 \leq k \leq n$ , let  $\mathcal{H}_k$  denote the set of possible history vectors of length  $k$ . A vector  $\mathbf{h} = (h_1, \dots, h_k) \in \mathbf{Z}^k$  belongs to  $\mathcal{H}_k$  if  $h_2 = h_1 + 1$  and, for all  $j \geq 3$ , the value of  $h_j$  is either  $\min(h_1, \dots, h_{j-1}) - 1$  or  $\max(h_1, \dots, h_{j-1}) + 1$ . This last requirement just says that, after the initial doublet has formed, a covered interval can only grow by adding grains immediately to the left or right of those grains already in the interval. For  $\mathbf{h} = (h_1, \dots, h_k) \in \mathcal{H}_k$ , we define  $|\mathbf{h}| = k$ ,  $\{\mathbf{h}\} = \{h_1, \dots, h_k\}$ ,  $\min(\mathbf{h}) = \min(h_1, \dots, h_k)$ , and  $\max(\mathbf{h}) = \max(h_1, \dots, h_k)$ . For a covered interval with history  $\mathbf{h}$ , the leftmost and rightmost grains in the interval are  $\min(\mathbf{h})$  and  $\max(\mathbf{h})$ , respectively.

For any history vector  $\mathbf{h}$ , we define the event  $A(\mathbf{h})$  by

$$\begin{aligned} A(\mathbf{h}) &= \{T_{h(1)} < T_{h(2)} \text{ and } \epsilon_{h(1)} = +1, \text{ or } T_{h(1)} > T_{h(2)} \text{ and } \epsilon_{h(2)} = -1\} \quad (4) \\ &\cap \{\min(T_{h(1)}, T_{h(2)}) < T_{h(3)} < T_{h(4)} < \dots < T_{h(n)}\} \\ &\cap \{\epsilon_{h(i)} = +1 \text{ for } h_i < h_1\} \\ &\cap \{\epsilon_{h(i)} = -1 \text{ for } h_i > h_2\}. \end{aligned}$$

The conditions in  $A(\mathbf{h})$  are exactly the conditions required for the grains in  $\{\mathbf{h}\}$  to form a covered interval with history vector  $\mathbf{h}$ , assuming there is no interference from grains outside  $\{\mathbf{h}\}$ . The first event in this intersection simply states that in order for the initial doublet (consisting of grains  $h_1$  and  $h_2$ ) to form, either grain  $h_1$  must bump grain  $h_2$  or grain  $h_2$  must bump grain  $h_1$ . The initial doublet forms at time  $\min(T_{h(1)}, T_{h(2)})$ . The second event in the intersection imposes the time ordering required by the history vector  $\mathbf{h}$ . The third and fourth events state that, in order for the interval to form, the grains to the left of grain  $h_1$  must bump the grains to their right, and the grains to the right of grain  $h_2$  must bump the grains to their left.

If  $\mathbf{h} \in \mathcal{H}_n$  and  $\min(\mathbf{h}) = 1$ , then  $A(\mathbf{h})$  is the event that, after all growth has stopped, the circle contains a single covered interval which begins at grain 1 and has history vector  $\mathbf{h}$ . (There are no grains outside  $\{\mathbf{h}\}$  to interfere with the formation of this covered interval.) For any history vectors  $\mathbf{h}^1, \mathbf{h}^2, \dots, \mathbf{h}^p$ , define

$$A(\mathbf{h}^1, \mathbf{h}^2, \dots, \mathbf{h}^p) = \bigcap_{i=1}^p A(\mathbf{h}^i). \quad (5)$$

We say that  $\mathbf{h}^1, \dots, \mathbf{h}^p$  are “compatible” if  $\min(\mathbf{h}^1) = 1$ ,  $\max(\mathbf{h}^p) = n$ , and  $\max(\mathbf{h}^i) + 1 = \min(\mathbf{h}^{i+1})$  for  $i = 1, 2, \dots, p - 1$ . For compatible history vectors, the event  $A(\mathbf{h}^1, \dots, \mathbf{h}^p)$  occurs if and only if, after all growth has stopped, the circle contains  $p$  covered intervals with history vectors  $\mathbf{h}^1, \dots, \mathbf{h}^p$ . (Note that, for any  $i$ , the conditions imposed by  $A(\mathbf{h}^i)$  prevent the grains in  $\{\mathbf{h}^i\}$  from interfering with the growth of the other covered intervals.)

## 4.5 The Case of a Single Interval

The results in Theorem 2 essentially follow from corresponding results for the situation in which the circle contains only a single covered interval. This section deals with this situation. Define  $\mathcal{H}_k^j$  to be the set of history vectors for covered intervals which contain  $k$  grains and begin with grain number  $j$ , that is,  $\mathcal{H}_k^j = \{\mathbf{h} : \mathbf{h} \in \mathcal{H}_k \text{ and } \min(\mathbf{h}) = j\}$ . We assume throughout this section that  $\mathbf{h} \in \mathcal{H}_n^1$ .

Given  $A(\mathbf{h})$ , let  $Y$  denote the length of the single covered interval. This length is the sum of the lengths of the  $n$  grains. Since grains  $h_1$  and  $h_2$  grow at rates  $r_{h(1)}$  and  $r_{h(2)}$  respectively, and their growth stops at time  $\tilde{T} \equiv \min(T_{h(1)}, T_{h(2)})$ , their combined length is  $(r_{h(1)} + r_{h(2)})\tilde{T}$ . For  $i \geq 3$ , grain  $h_i$  grows at rate  $r_{h(i)}$  for a period of time  $T_{h(i)}$ . Thus

$$Y = (r_{h(1)} + r_{h(2)})\tilde{T} + \sum_{i=3}^n r_{h(i)}T_{h(i)}. \quad (6)$$

Since the random variables  $\tilde{T}, T_{h(3)}, \dots, T_{h(n)}$  are independent of the first, third, and fourth events in (4), we see (using (2)) that

$$\mathcal{L}(Y | A(\mathbf{h})) = \mathcal{L}(Y | \tilde{T} < T_{h(3)} < T_{h(4)} < \dots < T_{h(n)}) = \text{gamma}(n - 1). \quad (7)$$

Alternatively, one may argue that  $Y$  has a  $\text{gamma}(n - 1)$  distribution because it is independent of the length of the single uncovered gap (which has an  $\text{exp}(1)$  distribution) and the total circumference of the circle has a  $\text{gamma}(n)$  distribution. However, this simple argument does not suffice when there is more than one covered interval (the situation discussed in the next section).

Now consider the event  $B(k)$  that, after all growth has stopped, the circle contains a single covered interval which begins at grain 1 and has exactly  $k$  grains to the left of the initial doublet. Having  $k$  grains to the left of the initial doublet means that the history vector  $\mathbf{h}$  for this interval has  $h_1 = k + 1$ . The grains  $1, 2, \dots, k$  to the left of the initial doublet must join the interval in the order  $k, k - 1, \dots, 1$  and must each bump the grains to their right. Similarly, the grains  $k + 3, k + 4, \dots, n$  to the right of

the initial doublet must join the interval in the order  $k + 3, k + 4, \dots, n$  and must each bump the grains to their left. Thus, we have

$$\begin{aligned} B(k) &= \{T_{k+1} < T_{k+2} \text{ and } \epsilon_{k+1} = +1, \text{ or } T_{k+1} > T_{k+2} \text{ and } \epsilon_{k+2} = -1\} \\ &\quad \cap \{\tilde{T} < T_k < T_{k-1} < \dots < T_1 \text{ and } \tilde{T} < T_{k+3} < T_{k+4} < \dots < T_n\} \\ &\quad \cap \{\epsilon_i = +1 \text{ for } i \leq k\} \\ &\quad \cap \{\epsilon_i = -1 \text{ for } i \geq k + 3\} \end{aligned} \quad (8)$$

with  $\tilde{T} = T_{k+1} \wedge T_{k+2}$ . Let  $Y$  again denote the length of the covered interval. Using (6), (8) and (3) and repeating the argument which led to (7), we obtain

$$\begin{aligned} \mathcal{L}(Y | B(k)) & \\ &= \mathcal{L}(Y | \tilde{T} < T_k < T_{k-1} < \dots < T_1 \text{ and } \tilde{T} < T_{k+3} < T_{k+4} < \dots < T_n) \\ &= \text{gamma}(n - 1) \end{aligned} \quad (9)$$

Alternatively, we may obtain this result as a direct consequence of (7).

Until now we have viewed the growth rates as fixed constants, that is, we may think of the results proved so far as holding conditionally given  $(A_i, B_i) = (a_i, b_i)$  for  $i = 1, \dots, n$ . We now consider the random growth assumptions B1, B2, B3, and we think of the beginning of Section 4.3 as describing the conditional distribution of  $T_1, \dots, T_n, \epsilon_1, \dots, \epsilon_n$  given the growth rates. The unconditional distributions satisfy:

$$\text{Under B2, the pairs } (T_1, \epsilon_1), \dots, (T_n, \epsilon_n) \text{ are exchangeable.} \quad (10)$$

$$\begin{aligned} \text{Under B3, } T_1, \dots, T_n \text{ are exchangeable, } \epsilon_1, \dots, \epsilon_n \text{ are iid, and} & \quad (11) \\ T_1, \dots, T_n \text{ are independent of } \epsilon_1, \dots, \epsilon_n. & \end{aligned}$$

These intuitive properties are easily shown from the following concrete representation. Let  $W_1, \dots, W_n$  and  $U_1, \dots, U_n$  be random variables which are independent of each other and of the growth rates with  $W_i \sim \exp(1)$  and  $U_i \sim \text{uniform}(0, 1)$  for all  $i$ . For all  $i$ , define

$$\begin{aligned} T_i &= W_i / (A_i + B_i) \quad \text{and} \\ \epsilon_i &= +1 \text{ if } U_i > A_i / (A_i + B_i) \text{ and } \epsilon_i = -1 \text{ otherwise.} \end{aligned} \quad (12)$$

Conditional on  $(A_i, B_i) = (a_i, b_i)$ , these random variables have the properties required of  $T_i$  and  $\epsilon_i$  in Section 4.3. Under B2, the vectors  $(A_i, B_i, W_i, U_i)$ ,  $i = 1, \dots, n$ , are exchangeable so that (10) is true. Under B3, we have  $\epsilon_i = +1$  if  $U_i > S_i$  and  $-1$  otherwise, and  $T_i = W_i / R_i$  and thus (11) holds.

Under assumptions B1, it is clear that (7) and (9) continue to hold; if  $Y$  has a  $\text{gamma}(n - 1)$  distribution conditional on the growth rates, it will have the same distribution unconditionally. This implies a special case of property Q1 for grains on the circle: when all the  $n$  grains form a single interval, the length of this interval has a  $\text{gamma}(n - 1)$  distribution.

Now assume B2. For any history vector  $\mathbf{h}$ , property (10) implies

$$((T_{h(1)}, \epsilon_{h(1)}), \dots, (T_{h(n)}, \epsilon_{h(n)})) \stackrel{d}{=} ((T_1, \epsilon_1), \dots, (T_n, \epsilon_n)).$$

Thus, equation (4) yields

$$\begin{aligned}
P(A(\mathbf{h})) &= P(\{T_1 < T_2 \text{ and } \epsilon_1 = +1, \text{ or } T_1 > T_2 \text{ and } \epsilon_2 = -1\} \\
&\quad \cap \{T_1 \wedge T_2 < T_3 < T_4 < \dots < T_n\} \\
&\quad \cap \{\epsilon_i = +1 \text{ for } h_i < h_1\} \\
&\quad \cap \{\epsilon_i = -1 \text{ for } h_i > h_2\}).
\end{aligned} \tag{13}$$

Let  $C$  be the event that, after all growth has stopped, the circle has a single covered interval that begins with grain 1. Then

$$P(C) = \sum_{\mathbf{h} \in \mathcal{H}_n^1} P(A(\mathbf{h})). \tag{14}$$

There is a one-to-one correspondence between the history vectors in  $\mathcal{H}_n^1$  and the  $2^{n-2}$  binary vectors of length  $n-2$ . This correspondence is given by

$$\mathbf{h} \longleftrightarrow (I(h_3 < h_1), I(h_4 < h_1), \dots, I(h_n < h_1)).$$

(To see that the indicators determine  $\mathbf{h}$  note that  $\sum_{i \geq 3} I(h_i < h_1)$  is the number of grains to the left of the initial doublet. This reveals which two grains are in the doublet (that is,  $h_1$  and  $h_2$ ) since the interval starts with grain 1. Now, since the interval can only grow by adding grains one by one, it is easy to deduce from  $I(h_i < h_1)$ ,  $i = 3, \dots, n$ , the actual values of  $h_3, \dots, h_n$ .) From this correspondence we see that summing  $P(A(\mathbf{h}))$  given in (13) over  $\mathbf{h} \in \mathcal{H}_n^1$  amounts to summing over all the possible values for  $(\epsilon_3, \epsilon_4, \dots, \epsilon_n)$ . This leads to

$$\begin{aligned}
P(C) &= P(\{T_1 < T_2 \text{ and } \epsilon_1 = +1, \text{ or } T_1 > T_2 \text{ and } \epsilon_2 = -1\} \\
&\quad \cap \{T_1 \wedge T_2 < T_3 < T_4 < \dots < T_n\}) \\
&= P(T_1 < T_2, T_1 < T_3 < \dots < T_n, \epsilon_1 = +1) \\
&\quad + P(T_2 < T_1, T_2 < T_3 < \dots < T_n, \epsilon_2 = -1) \\
&= P(T_1 < T_2, T_1 < T_3 < \dots < T_n) \\
&= \frac{1}{n} \cdot \frac{1}{(n-2)!} = \frac{n-1}{n!}.
\end{aligned} \tag{15}$$

We have used the exchangeability in (10), interchanging 1 and 2, to go from the second to the third equality above. The last equality is true because the probability that  $T_1$  is the smallest value among  $T_1, \dots, T_n$  is  $1/n$ , and there is probability  $1/(n-2)!$  that  $T_3, \dots, T_n$  will assume the given ordering. The final result in (15) is the source of the factor  $(k-1)/k!$  in Q3.

Now assume B3. From (8) and (11) we obtain

$$\begin{aligned}
P(B(k)) &= P(\{T_{k+1} < T_{k+2} \text{ and } \epsilon_{k+1} = +1, \text{ or } T_{k+1} > T_{k+2} \text{ and } \epsilon_{k+2} = -1\} \\
&\quad \cap \{\tilde{T} < T_k < T_{k-1} < \dots < T_1 \text{ and } \tilde{T} < T_{k+3} < T_{k+4} < \dots < T_n\}) \\
&\quad \times \xi^k (1 - \xi)^{n-k-2}.
\end{aligned} \tag{16}$$

where  $\xi = P(\epsilon_i = +1) = 1 - ES_i$ . The first term in (16) is

$$\begin{aligned}
& P(\{T_{k+1} < T_{k+2} \text{ and } \epsilon_{k+1} = +1, \text{ or } T_{k+1} > T_{k+2} \text{ and } \epsilon_{k+2} = -1\} \\
& \quad \cap \{\tilde{T} < T_k < T_{k-1} < \dots < T_1 \text{ and } \tilde{T} < T_{k+3} < T_{k+4} < \dots < T_n\}) \\
& = P(T_{k+1} < T_{k+2}, T_{k+1} < T_k < \dots < T_1, T_{k+1} < T_{k+3} < \dots < T_n, \epsilon_{k+1} = +1) \\
& \quad + P(T_{k+2} < T_{k+1}, T_{k+2} < T_k < \dots < T_1, T_{k+2} < T_{k+3} < \dots < T_n, \epsilon_{k+2} = -1) \\
& = P(T_{k+1} < T_{k+2}, T_{k+1} < T_k < T_{k-1} < \dots < T_1, T_{k+1} < T_{k+3} < T_{k+4} < \dots < T_n) \\
& = \frac{1}{n} \cdot \frac{1}{k!} \cdot \frac{1}{(n-k-2)!}.
\end{aligned}$$

The last equality is true because the probability that  $T_{k+1}$  is the smallest value among  $T_1, \dots, T_n$  is  $1/n$ , and there is probability  $1/k!$  that any  $k$  of the  $T_i$  will assume a given ordering. Plugging this into (16) gives us

$$P(B(k)) = \frac{\xi^k (1 - \xi)^{n-k-2}}{n k! (n - k - 2)!}. \quad (17)$$

Now using (15) we obtain

$$P(B(k)|C) = \binom{n-2}{k} \xi^k (1 - \xi)^{n-k-2}. \quad (18)$$

So we conclude that, under B3, the number of grains to the left of the initial doublet (given that the  $n$  grains form a single interval which begins at grain 1) has a binomial distribution.

## 4.6 Multiple Intervals

We now return to thinking of the growth rates as constants  $(a_i, b_i)$  as in Section 4.3. For any history vector  $\mathbf{h}$ , let  $Y(\mathbf{h})$  denote the length that a covered interval with this history would have. From the discussion leading to (6) we see that

$$Y(\mathbf{h}) = (r_{h(1)} + r_{h(2)}) \min(T_{h(1)}, T_{h(2)}) + \sum_{i=3}^{|\mathbf{h}|} r_{h(i)} T_{h(i)}.$$

Let  $\mathbf{h}^1, \mathbf{h}^2, \dots, \mathbf{h}^p$  be any compatible history vectors (see the discussion following (5)). Suppose that  $A(\mathbf{h}^1, \mathbf{h}^2, \dots, \mathbf{h}^p)$  has occurred, and let  $Y_i \equiv Y(\mathbf{h}^i)$  denote the length of the  $i$ -th covered interval. Since  $T_1, \dots, T_n, \epsilon_1, \dots, \epsilon_n$  are mutually independent, and the sets  $\{\mathbf{h}^1\}, \dots, \{\mathbf{h}^p\}$  are disjoint, the same reasoning which led to (7) produces

$$\begin{aligned}
& \text{conditional on } A(\mathbf{h}^1, \mathbf{h}^2, \dots, \mathbf{h}^p), \\
& \text{the random variables } Y_1, \dots, Y_p \text{ are independent and } Y_i \sim \text{gamma}(|\mathbf{h}^i| - 1).
\end{aligned} \quad (19)$$

This continues to hold under model B1 and proves Q1; if  $Y_1$  has the required gamma distribution conditional on all the growth rates and the history vectors for all the covered intervals, it will also have this distribution without the conditioning.

Now assume the model B2. Define  $C(n_1, n_2, \dots, n_p)$  to be the event that, after all growth has stopped, (i) the circle has  $p$  covered intervals, (ii) the first of these intervals begins at grain 1, and (iii) the number of grains in each interval is  $n_1, n_2, \dots, n_p$  respectively where  $n = n_1 + \dots + n_p$ . We may evaluate  $P(C(n_1, n_2, \dots, n_p))$  by repeated application of the argument which led to (15). For  $1 \leq k \leq p$ , define

$$\mathcal{G}(k) = \mathcal{H}_{n_1}^{f_1} \times \mathcal{H}_{n_2}^{f_2} \times \dots \times \mathcal{H}_{n_k}^{f_k}$$

where  $f_1 = 1$  and  $f_i = 1 + \sum_{j < i} n_j$  for  $i > 1$ . Then

$$\begin{aligned} & P(C(n_1, n_2, \dots, n_p)) \\ &= \sum_{(\mathbf{h}^1, \mathbf{h}^2, \dots, \mathbf{h}^p) \in \mathcal{G}(p)} P(A(\mathbf{h}^1, \mathbf{h}^2, \dots, \mathbf{h}^p)) \\ &= \sum_{(\mathbf{h}^1, \dots, \mathbf{h}^{p-1}) \in \mathcal{G}(p-1)} P(A(\mathbf{h}^1, \dots, \mathbf{h}^{p-1})) \sum_{\mathbf{h}^p \in \mathcal{H}_{n_p}^{f_p}} P(A(\mathbf{h}^p) | A(\mathbf{h}^1, \dots, \mathbf{h}^{p-1})). \end{aligned}$$

As a consequence of (10), given the values of any subset of the pairs  $(T_i, \epsilon_i)$ , the remaining pairs  $(T_i, \epsilon_i)$  are conditionally exchangeable. Thus, we may repeat the argument leading to (15) to show that

$$\sum_{\mathbf{h}^p \in \mathcal{H}_{n_p}^{f_p}} P(A(\mathbf{h}^p) | (T_i, \epsilon_i) \text{ for } i \in \{\mathbf{h}^1\} \cup \dots \cup \{\mathbf{h}^{p-1}\}) = \frac{n_p - 1}{n_p!}$$

which implies

$$\sum_{\mathbf{h}^p \in \mathcal{H}_{n_p}^{f_p}} P(A(\mathbf{h}^p) | A(\mathbf{h}^1, \dots, \mathbf{h}^{p-1})) = \frac{n_p - 1}{n_p!}.$$

Continuing in this way leads to

$$P(C(n_1, n_2, \dots, n_p)) = \prod_{i=1}^p \frac{n_i - 1}{n_i!}. \quad (20)$$

We see below that this fact implies Q3.

Now consider the model B3. Define  $B((k_1, k_2, \dots, k_p), (n_1, n_2, \dots, n_p))$  to be the event that, after all growth has stopped, (i) the circle contains  $p$  covered intervals, (ii) the first of these intervals begins at grain 1, (iii) the intervals contain  $n_1, n_2, \dots, n_p$  grains respectively, and (iv) the number of grains to the left of the initial doublet in each interval is  $k_1, k_2, \dots, k_p$  respectively. This event may be written as an intersection of  $p$  events which are defined in a way similar to  $B(k)$  in (8). Using the conditional exchangeability noted prior to (20), we can repeat the arguments involved in (16) and (17) to each of the  $p$  events in this intersection to obtain an expression for  $P(B((k_1, \dots, k_p), (n_1, \dots, n_p)))$  as a product of probabilities like those in (17). Then dividing by  $P(C(n_1, \dots, n_p))$  in (20) leads to

$$P(B((k_1, \dots, k_p), (n_1, \dots, n_p)) | C(n_1, \dots, n_p)) = \prod_{i=1}^p \binom{n_i - 2}{k_i} \xi^{k(i)} (1 - \xi)^{n(i) - k(i) - 2} \quad (21)$$

where  $\xi = 1 - ES_i$ . This fact implies property Q5.

If B2 and B3 are strengthened by replacing ‘exchangeable’ by ‘iid’, then somewhat simpler arguments for (20) and (21) can be given. For example, under the iid assumption, the events  $A(\mathbf{h}^1), \dots, A(\mathbf{h}^p)$  are independent whenever the sets  $\{\mathbf{h}^1\}, \dots, \{\mathbf{h}^p\}$  are disjoint, whereas under our weaker B2 and B3, the events  $A(\mathbf{h}^1), \dots, A(\mathbf{h}^p)$  are *not* independent (but will only be *conditionally* independent given the growth rates).

## 4.7 Properties Q2 and Q3

In the previous section we proved Q1 and Q5. We now show that Q2 and Q3 hold under B2.

With  $n$  grains on the circle, let  $F$  be the event that there is a covered interval which begins with grain 1. Equivalently,  $F$  is the event that there is an uncovered gap between grains 0 and 1. We now show that

$$P(F) = \pi_n \equiv \sum_{i=0}^n \frac{(-1)^k}{k!}. \quad (22)$$

From the definition of  $C(n_1, n_2, \dots, n_p)$ , we see that

$$P(F) = \sum_{p \geq 1} \sum_{(n_1, n_2, \dots, n_p)} P(C(n_1, n_2, \dots, n_p)) \quad (23)$$

with the inner sum over all  $(n_1, n_2, \dots, n_p)$  satisfying  $n_1 + \dots + n_p = n$ . Standard generating function arguments now lead to (22) as follows. Define the generating function

$$\phi(x) = \sum_{k=2}^{\infty} \frac{k-1}{k!} x^k = 1 + (x-1)e^x. \quad (24)$$

From (20), we see that the inner sum in (23) is the coefficient of  $x^n$  in the series expansion of  $\phi(x)^p$ . Summing this over  $p$ , we find that  $P(F)$  is the coefficient of  $x^n$  in the function

$$\sum_{p \geq 1} \phi(x)^p = \frac{\phi(x)}{1 - \phi(x)} = \frac{e^{-x}}{1-x} - 1.$$

This coefficient is  $\pi_n$ .

Define  $F_k$  to be the event that there is a covered interval containing exactly  $k$  grains which begins with grain 1. For  $k = n$ , equation (15) gives us  $P(F_n) = (n-1)/n!$ . In general, we have

$$P(F_k) = \sum_{p \geq 1} \sum_{n_2, \dots, n_p} P(C(k, n_2, \dots, n_p)) \quad (25)$$

where the inner sum is over  $n_2, \dots, n_p$  satisfying  $n_2 + \dots + n_p = n - k$ . Using (20) and repeating the generating function argument which led to (22) we obtain

$$P(F_k) = \frac{k-1}{k!} \pi_{n-k}. \quad (26)$$

Thus

$$P(F_k|F) = \frac{k-1}{k!} \cdot \frac{\pi_{n-k}}{\pi_n}. \quad (27)$$

Conditional on there being a covered interval which begins at grain 1, we may regard this interval as a “randomly chosen” interval. Thus, equation (27) tells us the probability that a randomly chosen interval contains exactly  $k$  grains.

Property Q2 essentially follows from Q1 and (22) plus the fact that the lengths of the uncovered gaps are iid  $\exp(1)$  random variables. Here is the argument. Let  $C$  and  $U$  denote the total lengths of the covered and uncovered parts of the circle after all growth has stopped. Since the lengths of the uncovered gaps are iid  $\exp(1)$ , we conclude from equation (19) that, conditional on  $A(\mathbf{h}^1, \mathbf{h}^2, \dots, \mathbf{h}^p)$ , the random variables  $C$  and  $U$  are independent,  $C \sim \text{gamma}(n-p)$ , and  $U \sim \text{gamma}(p)$ . Thus, conditional on the number of covered intervals being  $p$ , the fraction  $V = U/(U+C)$  of the circle which is not covered has a beta distribution with mean  $p/n$ . Let  $G$  denote the random number of covered intervals on the circle. It is clear by symmetry that  $P(F|G = p) = p/n$ . Therefore,  $EV = P(F) = \pi_n$  by (22).

## 4.8 Property Q4

We prove property Q4 by a coupling argument. A careful statement of this proof would require rather elaborate notation, so we shall be fairly informal.

Let  $T_i$  and  $\epsilon_i$  be as in Section 4.3. Under B2, the vectors  $(A_i, B_i, T_i, \epsilon_i)$ ,  $i = 1, \dots, n$  are exchangeable (see the discussion after (10)). This exchangeability allows us to modify the “grain description” of Section 4.3 in the following way: the  $n$  grains will no longer be in a given fixed ordering around the circle, but will instead be placed in a random ordering, with the position of grain  $i$  being decided at time  $T_i$ .

Here is a detailed description of this modified growth model (referred to as MGM below). We are given  $n$  grains with values  $(A_i, B_i, T_i, \epsilon_i)$ ,  $i = 1, \dots, n$ . Let  $T_{(1)} < T_{(2)} < \dots < T_{(n)}$  be the order statistics of  $T_1, \dots, T_n$ , and  $\sigma_1, \sigma_2, \dots, \sigma_n$  be the associated grain numbers so that  $T_{(i)} = T_{\sigma(i)}$  for all  $i$ . Carry out the following sequential construction. Suppose there are  $n$  positions arranged in a circle, and that initially (at time zero) all these positions are empty. At time  $T_{(1)}$ , place grain  $\sigma_1$  in a randomly chosen position. If  $\epsilon_{\sigma(1)} = -1$ , then grain  $\sigma_1$  bumps the position to the left, otherwise it bumps the position to the right. At time  $T_{(2)}$ , grain  $\sigma_2$  is placed at random in one of the remaining  $n-1$  positions. If this position is not the one bumped by grain  $\sigma_1$ , then grain  $\sigma_2$  bumps the position to its left or right, depending on the value of  $\epsilon_{\sigma(2)}$ . And so on for the other grains. At time  $T_{(i)}$ , grain  $\sigma_i$  is placed at random in one of the remaining  $n-i+1$  positions. If this position has not been bumped by any of the previously placed grains, then grain  $\sigma_i$  bumps the position to its left or right, depending on the value of  $\epsilon_{\sigma(i)}$ . Grains (or positions) which bump each other join together to form covered intervals. At time  $T_{(n)}$  the process is complete. In the final configuration, the length of grain  $i$  is taken to be  $(A_i + B_i)T_i^*$  where  $T_i^*$  is either  $T_i$  or the first time that the position occupied by grain  $i$  was bumped by another grain, whichever is smaller.

It is easily seen that this description is the same as that in Section 4.3 except that now the grains are in a random order. Under the assumption of exchangeability, this reordering has no effect on the distribution of the covered intervals.

At any time in the MGM process, the unoccupied positions are of two types: those which have already been bumped by some grain, and those which have not. The

unoccupied positions of the first type will be said to be “frozen” at the time they were first bumped. The positions of the second type will be called “free”. When a grain is assigned to a free position and bumps another free position, this creates a new covered interval with the grains in these two positions forming the initial doublet. If a grain is assigned to a frozen position or bumps a frozen position, then the grain joins an already existing covered interval.

Now suppose we have two sets of growth rates  $(A_i, B_i)$  and  $(A'_i, B'_i)$ ,  $i = 1, \dots, n$  which both satisfy B2. Assume that the joint distribution of  $(A_1 + B_1, \dots, A_n + B_n)$  is the same as that of  $(A'_1 + B'_1, \dots, A'_n + B'_n)$ . Then without loss of generality we may define all the random variables  $A_i, B_i, A'_i, B'_i$ ,  $i = 1, \dots, n$ , on the same probability space with  $A_i + B_i = A'_i + B'_i$  almost surely for all  $i$ . Let  $(T_i, \epsilon_i)$  and  $(T'_i, \epsilon'_i)$  be the random variables described in Section 4.3 corresponding to these two sets of growth rates. By using the construction in (12) with the same  $W_i$ 's and  $U_i$ 's for both sets of growth rates, we may assume that  $T_i = T'_i$  for all  $i$ .

We simultaneously carry out the MGM construction for both sets of random variables  $(A_i, B_i, T_i, \epsilon_i)$  and  $(A'_i, B'_i, T'_i, \epsilon'_i)$  with the random placement of the grains coupled in a special way. For descriptive purposes, we say the two constructions take place in universe 1 and universe 2 (denoted  $\mathcal{U}_1$  and  $\mathcal{U}_2$ ) respectively. Since  $T_i = T'_i$  for all  $i$ , events in the two universes occur simultaneously. The coupling will be such that, at any time  $t$ , the unoccupied positions in the two universes are in one-to-one correspondence:

- (i) For every frozen position in  $\mathcal{U}_1$  there is a corresponding frozen position in  $\mathcal{U}_2$  which was frozen at the same time.
- (ii) The occupied and frozen positions break up the free positions into blocks such that every block of  $k$  consecutive free positions in  $\mathcal{U}_1$  corresponds to a block of  $k$  consecutive free positions in  $\mathcal{U}_2$ .

The coupling is defined inductively. Suppose that we have carried out the construction up to time  $t$  and are about to place the next grain which is (let us say) grain  $i$ . In  $\mathcal{U}_1$ , choose an unoccupied position (call it  $p$ ) at random and place grain  $i$  in this position. If this position  $p$  happens to be frozen, then in  $\mathcal{U}_2$  place grain  $i$  in the corresponding frozen position. If the position  $p$  is free and  $\epsilon_i = \epsilon'_i$ , then in  $\mathcal{U}_2$  place grain  $i$  in the corresponding free position. To be precise, if position  $p$  is the  $j$ -th in a block of  $k$  free positions, then there will be a corresponding block of  $k$  free positions in  $\mathcal{U}_2$ , and we place grain  $i$  in the  $j$ -th of these. If the position  $p$  is free but  $\epsilon_i \neq \epsilon'_i$ , then in  $\mathcal{U}_2$  we place grain  $i$  in the opposite (mirror image) location of the corresponding block of free positions. That is, if position  $p$  is the  $j$ -th in a block of  $k$  free positions, then in  $\mathcal{U}_2$  we place grain  $i$  in the  $(k - j + 1)$ -th position in the corresponding block of  $k$  free positions.

A little thought shows that properties (i) and (ii) remain true after the placement of grain  $i$  so that the coupling construction can continue until all the  $n$  grains have been placed. In each universe the construction is proceeding according to the MGM. Note that grain  $i$  has the same length in both universes. This holds for all  $i$ . Note also that whenever an initial doublet forms in  $\mathcal{U}_1$  (that is, a grain is assigned to a free position and bumps another free position), the same thing happens in  $\mathcal{U}_2$ , and moreover, the two initial doublets have the same length. Thus, upon completion of

the coupled construction, the grains and initial doublets in both universes have the same length. This implies Q4. (We note that both universes contain the same number of covered intervals, but that the number of grains in these intervals need not be the same.)

In Table 7 we give an example of this coupling construction with  $n = 9$  grains. For convenience in the presentation we take  $T_i = i$  for all  $i$  so that events in the two universes occur at times  $1, 2, \dots, 9$ , with grain  $i$  being placed at time  $i$ . The grains are placed in positions numbered 1 to 9 which the reader must imagine as being wrapped around a circle. At each time, the table describes the configuration of the grains which have been placed so far, and gives the one-to-one correspondence (satisfying (i) and (ii)) between the unoccupied positions in  $\mathcal{U}_1$  and  $\mathcal{U}_2$ . Free positions are indicated by a circle, and frozen positions by a triangle, with the attached subscripts giving the correspondence. For example, at time 3, the two positions marked  $\bigcirc_3$  and  $\triangle_2$  in  $\mathcal{U}_1$  correspond to the positions marked  $\bigcirc_3$  and  $\triangle_2$  in  $\mathcal{U}_2$ . At time  $i$ , grain  $i$  is placed in a randomly chosen unoccupied position; in the table, this position is marked with the value  $i$ . If this position was free prior to the placement of grain  $i$ , we also mark it with a plus or minus sign to indicate the sign of  $\epsilon_i$  (or  $\epsilon'_i$ ). (If the position was frozen, we do not bother to mark the sign since it has no effect on the construction.) For example, position 9 in  $\mathcal{U}_1$  is marked “3-” at time 3 and all subsequent times to indicate that grain 3 was placed in the (formerly free) position 9 and that  $\epsilon_3 = -1$ . In  $\mathcal{U}_2$ , grain 3 is placed in the corresponding position (position 6) which is marked “3+” indicating that  $\epsilon'_3 = +1$ .

Some remarks on the coupling construction: In Table 7, grain 1 is placed in the same position in both universes, but this is not required. The first grain can be placed completely arbitrarily in both universes since no matter where it is placed conditions (i) and (ii) will hold. In  $\mathcal{U}_1$  grain 2 is placed in the third position from the *left* in a block of 7 free positions. In  $\mathcal{U}_2$ , since  $\epsilon_2 \neq \epsilon'_2$ , grain 2 is placed in the third position from the *right* of the corresponding block. In both universes, the placement of grain 2 creates two new blocks of free positions (of sizes 2 and 3) and one new frozen position; these are placed in correspondence with each other in the obvious way. At time 5, there are two isolated free positions remaining in each universe, and the given correspondence associates positions 1 and 5 in  $\mathcal{U}_1$  with positions 5 and 2, respectively, in  $\mathcal{U}_2$ . But we can interchange this correspondence without affecting the validity of the construction. In general, if at any time  $t$  in the construction there happen to be  $r$  blocks of  $k$  free positions in both universes, any of the  $r!$  ways of associating these blocks is equally valid. The only requirement is that the method of associating the blocks at time  $t$  not depend on the random positions chosen for the placement of grains at times later than  $t$ ; peeking into the future would destroy the uniformity (randomness) of the placement of the grains in  $\mathcal{U}_2$ . There is no similar freedom in setting up the correspondence between the frozen positions; every frozen position in  $\mathcal{U}_1$  must be associated with the position in  $\mathcal{U}_2$  which was frozen at the same time. Finally, we note that at the end of the construction both universes contain 3 covered intervals. Listing the grains from left to right, in  $\mathcal{U}_1$  these intervals consist of grains 7, 4, 1, 5, 9 (interval 1), 2, 6 (interval 2), and 8, 3 (interval 3). In  $\mathcal{U}_2$  the intervals consist of grains 1, 4 (interval 1), 6, 2, 5, 9 (interval 2), and 7, 3, 8 (interval 3). The intervals have been numbered so that interval

	Position									
	1	2	3	4	5	6	7	8	9	
Time 1	$\bigcirc_7$	$\triangle_1$	1-	$\bigcirc_1$	$\bigcirc_2$	$\bigcirc_3$	$\bigcirc_4$	$\bigcirc_5$	$\bigcirc_6$	$\mathcal{U}_1$
	$\bigcirc_6$	$\bigcirc_7$	1+	$\triangle_1$	$\bigcirc_1$	$\bigcirc_2$	$\bigcirc_3$	$\bigcirc_4$	$\bigcirc_5$	$\mathcal{U}_2$
Time 2	$\bigcirc_5$	$\triangle_1$	1-	$\bigcirc_1$	$\bigcirc_2$	2+	$\triangle_2$	$\bigcirc_3$	$\bigcirc_4$	$\mathcal{U}_1$
	$\bigcirc_1$	$\bigcirc_2$	1+	$\triangle_1$	$\bigcirc_3$	$\bigcirc_4$	$\bigcirc_5$	$\triangle_2$	2-	$\mathcal{U}_2$
Time 3	$\bigcirc_1$	$\triangle_1$	1-	$\bigcirc_2$	$\bigcirc_3$	2+	$\triangle_2$	$\triangle_3$	3-	$\mathcal{U}_1$
	$\bigcirc_2$	$\bigcirc_3$	1+	$\triangle_1$	$\bigcirc_1$	3+	$\triangle_3$	$\triangle_2$	2-	$\mathcal{U}_2$
Time 4	$\bigcirc_1$	4	1-	$\bigcirc_2$	$\bigcirc_3$	2+	$\triangle_1$	$\triangle_2$	3-	$\mathcal{U}_1$
	$\bigcirc_2$	$\bigcirc_3$	1+	4	$\bigcirc_1$	3+	$\triangle_2$	$\triangle_1$	2-	$\mathcal{U}_2$
Time 5	$\bigcirc_1$	4	1-	5-	$\bigcirc_2$	2+	$\triangle_1$	$\triangle_2$	3-	$\mathcal{U}_1$
	5-	$\bigcirc_2$	1+	4	$\bigcirc_1$	3+	$\triangle_2$	$\triangle_1$	2-	$\mathcal{U}_2$
Time 6	$\bigcirc_1$	4	1-	5-	$\bigcirc_2$	2+	6	$\triangle_1$	3-	$\mathcal{U}_1$
	5-	$\bigcirc_2$	1+	4	$\bigcirc_1$	3+	$\triangle_1$	6	2-	$\mathcal{U}_2$
Time 7	7+	4	1-	5-	$\bigcirc_1$	2+	6	$\triangle_1$	3-	$\mathcal{U}_1$
	5-	$\bigcirc_1$	1+	4	7+	3+	$\triangle_1$	6	2-	$\mathcal{U}_2$
Time 8	7+	4	1-	5-	$\bigcirc_1$	2+	6	8	3-	$\mathcal{U}_1$
	5-	$\bigcirc_1$	1+	4	7+	3+	8	6	2-	$\mathcal{U}_2$
Time 9	7+	4	1-	5-	9-	2+	6	8	3-	$\mathcal{U}_1$
	5-	9-	1+	4	7+	3+	8	6	2-	$\mathcal{U}_2$

Table 7: Example of coupling construction.

$i$  in both universes has the same pair of grains in the initial doublet.

**Acknowledgement:** I thank the referee for many suggestions which have improved the paper.

## References

Daley, D.J., Mallows, C.L., Shepp, L.A. (2000). A one-dimensional Poisson growth model with non-overlapping intervals. *Stochastic Processes and their Applications* **90**, 223-241.

Daley, D.J., Stoyan, D., Stoyan, H. (1999). The volume fraction of a Poisson germ model with maximally non-overlapping spherical grains. *Adv. Appl. Prob.* **31**, 610-624.

Häggström, O., Meester, R. (1996). Nearest neighbor and hard sphere models in continuum percolation. *Random Structures and Algorithms* **9**, 295-315.

Huffer, F.W. (2002). One-dimensional Poisson growth models with random and asymmetric growth. Technical report, Dept. of Statistics, Florida State University.

O'Hely, M., Sudbury, A. (2001). The annihilating process. *Journal of Applied Probability* **38**, 223-231.

Sackrowitz, H., Samuel-Cahn, E. (1984). Estimation of the mean of a selected negative exponential population. *J. R. Statist. Soc. B* **46**, 242-249.

Sudbury, A. (2002). Inclusion-exclusion methods for treating annihilating and deposition processes. *Journal of Applied Probability* (to appear).

## Appendix: Additional Simulation Results

This appendix gives some additional simulations to illustrate the results in Section 2. These examples use  $n = 100$  grains on the circle. There are five different scenarios which are described below.

- (2a)  $(A_i, B_i) = (2, 1)$  for odd  $i$ , and  $(A_i, B_i) = (3, 6)$  for even  $i$ .
- (2b) The growth rates are iid with  $P\{(A_i, B_i)=(2,1)\} = P\{(A_i, B_i)=(3,6)\} = 1/2$ .
- (2c) The growth rates are iid with  $P\{(A_i, B_i)=(1,2)\} = P\{(A_i, B_i)=(3,6)\} = 1/2$ .
- (2d) The growth rates are iid with  $P\{(A_i, B_i)=(1,5)\} = P\{(A_i, B_i)=(3,3)\} = 1/2$
- (2e)  $(A_i, B_i) = (3, 3)$  for all  $i$ .

Note that (2a) satisfies B1, but not B2. (2b) satisfies B2, but not B3. Models (2c), (2d) and (2e) satisfy B3. (2e) is equivalent to the model studied by DMS. In (2c), we have  $P(R_i = 3) = P(R_i = 9) = 1/2$  and  $P(S_i = 1/3) = 1$ . In (2d), we have  $P(R_i = 6) = 1$  and  $P(S_i = 1/6) = P(S_i = 1/2) = 1/2$ . In both (2b) and (2c),  $A_i + B_i$  takes the values 3 and 9 with probability 1/2 each. In both (2d) and (2e),  $A_i + B_i$  takes the value 6 with probability 1.

There were 1,000,000 simulations of each scenario. The results are summarized in Tables 8 to 13 which follow the formats of the tables in Section 3. The last column of Table 9 displays  $(k - 1)/k!$ ; we have omitted the factor of  $\pi_{n-k}/\pi_n$  in Q3 which is essentially 1 for large  $n$ . In Table 10, in order to save space we present results only for models (2a), (2b), and (2d). The results for the other two models are similar.

We briefly comment on the tables. Table 10 suggests that Q1 is true for all the models. Tables 8 and 9 indicate that Q2 and Q3 are true for models (2b), (2c), (2d), and (2e) as expected. In Table 13, the values of  $\mu$  and  $\sigma^2$  reported for models (2c), (2d), and (2e) are in good agreement with those of the binomial( $k - 2, p$ ) distributions (with  $p = ES_i$  equal to 1/3, 1/3, and 1/2, respectively) predicted by Q5. Finally, in Tables 11 and 12, note the agreement between models (2b) and (2c), and between models (2d) and (2e), exactly as predicted by Q4.

(2a)	(2b)	(2c)	(2d)	(2e)	1/e
0.38467	0.36787	0.36781	0.36793	0.36794	0.36788

Table 8: Average fraction of circumference not covered.

$k$	(2a)	(2b)	(2c)	(2d)	(2e)	$(k-1)/k!$
2	0.58886	0.49990	0.49999	0.49987	0.50002	0.50000
3	0.25442	0.33327	0.33339	0.33344	0.33327	0.33333
4	0.13078	0.12511	0.12494	0.12502	0.12502	0.12500
5	0.02032	0.03337	0.03336	0.03333	0.03335	0.03333
6	0.00498	0.00695	0.00693	0.00695	0.00695	0.00694
7	0.00054	0.00119	0.00119	0.00118	0.00119	0.00119
8	0.00008	0.00018	0.00018	0.00018	0.00017	0.00017

Table 9: Distribution of the number  $k$  of grains in a covered interval.

$k$	(2a)			(2b)			(2d)		
	$\mu_1$	$\mu_2$	$\mu_3$	$\mu_1$	$\mu_2$	$\mu_3$	$\mu_1$	$\mu_2$	$\mu_3$
2	1.000	2.000	6.000	1.000	2.000	6.001	1.000	2.001	6.003
3	2.000	6.001	24.007	2.000	6.001	24.005	2.000	5.998	23.996
4	3.000	11.996	59.989	2.998	11.982	59.848	3.000	11.999	60.019
5	3.999	19.989	119.961	4.002	20.019	120.156	3.998	19.969	119.602
6	5.004	30.070	210.946	5.000	29.99	209.865	4.993	29.910	209.069
7	6.005	42.062	336.587	6.009	42.151	338.077	5.985	41.645	329.887
8	7.004	56.073	505.955	7.061	56.863	514.271	6.999	55.808	499.537

Table 10: The moments of the lengths of the covered intervals containing  $k$  grains.

	(2a)	(2b)	(2c)	(2d)	(2e)
$\mu_1$	0.6154	0.6321	0.6323	0.6321	0.6321
$\mu_2$	0.9889	1.0020	1.0026	0.9692	0.9694
$\mu_3$	2.6428	2.6615	2.6645	2.5284	2.5304
$\mu_4$	9.8150	9.9458	9.9663	9.4292	9.4454

Table 11: Moments of the distribution of grain sizes.

	(2a)	(2b)	(2c)	(2d)	(2e)
$\mu_1$	0.8382	0.8165	0.8167	0.8008	0.8007
$\mu_2$	1.4953	1.4266	1.4273	1.3715	1.3712
$\mu_3$	4.1690	3.9040	3.9070	3.7024	3.7009
$\mu_4$	15.9007	14.6454	14.6641	13.7975	13.7971

Table 12: Moments of the length of the initial doublet.

$k$	(2a)		(2b)		(2c)		(2d)		(2e)	
	$\mu$	$\sigma^2$	$\mu$	$\sigma^2$	$\mu$	$\sigma^2$	$\mu$	$\sigma^2$	$\mu$	$\sigma^2$
3	0.500	0.250	0.529	0.249	0.333	0.222	0.333	0.222	0.500	0.250
4	1.184	0.466	1.043	0.500	0.667	0.444	0.667	0.444	1.001	0.500
5	1.499	0.669	1.551	0.751	1.000	0.667	1.001	0.666	1.500	0.749
6	2.185	1.006	2.057	0.996	1.334	0.892	1.334	0.891	1.999	0.999
7	2.491	1.168	2.547	1.260	1.659	1.114	1.675	1.119	2.503	1.253
8	3.191	1.456	3.082	1.469	2.008	1.327	1.992	1.346	3.001	1.484

Table 13: The mean and variance of the number of grains to the right of the initial doublet for covered intervals with  $k$  grains.