Commun. Korean Math. Soc. **20** (2005), No. 4, pp. 813–826

SIZE OF THE CLUSTERS UNDER LOW DENSITY ZERO-RANGE INVARIANT MEASURES

INTAE JEON

ABSTRACT. Regarding all particles at a fixed site as a cluster, the size of the largest cluster under the zero range invariant measures is well studied by Jeon et al.[5] for the case of density one. Here, the density of the finite zero-range process is given by the ratio between the number m of particles and the number n of sites. In this paper, we study the lower density case, i.e., the case $m = o(n)$. Especially, when $m \sim n^{\beta}, 0 < \beta < 1$, we show that there is an interesting cutoff point around $\beta = 1/2$.

0. Introduction

Consider m Markovian particles moving around on finite number n of sites with transition matrix $\{P_{ij}\}_{i,j=1}^{n}$. If there are k particles at one site, say site i, then we regard the set of particles as a cluster of size k, or k-cluster, locates at site i or i-site. The dynamics of the sizes of the clusters is called the zero-range process and is an example of an interacting random system with the following interaction. (See [6]) Let g be a given monotone real valued function from nonnegative integers. Any k-cluster locates at i-site waits an exponential amount of time with parameter $g(k)$, picks j sites with probability P_{ij}, then gives one particle to j site. As a result, the size of the cluster at site i is reduced by one and the size of the cluster at site j is increased by one. If $\{P_{ij}\}_{i,j=1}^{n}$is symmetric and irreducible, then, as Spitzer[7] has shown, there is a unique invariant measure.

To find a new tractable situation similar to the well known Smoluchowski coagulation fragmentation dynamics (see [1, 3]), Jeon et al.[5]

Received February 28, 2005. Revised September 20, 2005.
2000 Mathematics Subject Classification: 60K35, 82C22.
Key words and phrases: zero-range process, low density, invariant measure, cluster size, cutoff point.

and Jeon and March[4] considered the density one case, i.e., $m = n$, and showed that there are two striking transitions of the size of the largest cluster under the invariant measures. For example, if $g(k) = k^{-\alpha}, -\infty < \alpha < \infty$ and let $Z \doteq (Z_1, Z_2, \cdots, Z_n)$ be the equilibrium zero range process with rate function $g(k)$, then we have the following theorem.

THEOREM 0.1. [5] *Let* $Z_n^* = \max_{1 \le i \le n} Z_i$.

(a) *If* $\alpha > 1$, *then* $n - Z_n^*$ *converges to 0 in probability.*
(b) *If* $\alpha = 1$, *then* $n - Z_n^*$ *converges weakly to a Poisson distribution of parameter 1.*
(c) *If* $0 < \alpha < 1$, *then* $(n - Z_n^*)/n^{1+\alpha}$ *converges to 1 in probability.*
(d) *If* $\alpha = 0$, *then* $Z_n^*/\log n$ *converges to 2 in probability.*
(e) *If* $\alpha < 0$, *then* $Z_n^* \log\log n/\log n$ *converges to* α^{-1} *in probability.*

Since any configuration $(Z_1, Z_2, \cdots, Z_n)$ satisfies $Z_1 + Z_2 + \cdots + Z_n = n$, the configuration space consists of integer partitions of n and the equilibrium measure is, then, a random measure of the partitions. The analysis of this random structure can be carried out through the help of Arratia and Tavaré's results [2]. Consider one parameter family of i.i.d. random variables $\{X_i\}_{i=1}^n$ on $\{0, 1, \cdots\}$, where $P\{X_i = k\} = x^k/g!(k), g!(k) = g(k)g(k-1)\cdots g(1)g(0), g(0) = g!(0) = 1$, for $x \in R$. Then one can prove easily that $(Z_1, Z_2, \cdots, Z_n) =_d (X_1, X_2, \cdots, X_n | X_1 + X_2 + \cdots + X_n = n)$, where $=_d$ means that the both terms are equal in distribution. Now we simply choose x which maximizes $P\{X_1 + X_2 + \cdots + X_n = n\}$ and successfully remove the dependence structure.

Our aim is to look at the lower density case, i.e., the case $m = o(n)$. One difficulty is that the above technique does not seem to be applicable in this case, since the probability $P\{X_1 + X_2 + \cdots + X_n = m\}$ is too small to analyze. Instead, we deal with, though less sophisticated, the configuration space and the invariant measures directly.

In this paper, we consider the case $m \sim n^\beta, 0 < \beta < 1$. Suppose β is close to 0, i.e., $\beta << 1$, then, since there are not many particles compared to the sites, it is hard to form 2-clusters. On the other hand, if β is close to 1, then, since there are many particles, with high probability, there exist bigger size clusters. Our natural guess is that, as β increases from 0, the maximum cluster size should increase.

It turns out that there is an interesting cutoff point around $\beta = 1/2$. Indeed, we are able to show that if $g(l-1)/g(l)$ is bounded and

1) if $\beta < 1/2$, then the maximum cluster size is 1. (Theorem 2.6)

2) if $\beta = 1/2$, then the maximum cluster size is ≤ 2. Moreover, the number of 2. clusters converges to a Poisson distribution with parameter $\frac{g(1)^2}{g(2)}$. (Theorem 2.7)

3) if $\beta > 1/2$, then the maximum cluster size is ≥ 2. Moreover, the probability that the number of 2-clusters is finite converges to 0. (Theorem 2.11)

Note that, if $g(l) = l^{-a}$, $-\infty < a < \infty$, as given in Theorem 0.1, then $g(l-1)/g(l)$ is bounded. The condition that $g(l-1)/g(l)$ is bounded is satisfied even $g(l)$ decreases or increases exponentially. For example, let $g(l) = e^{al}$ for some constant a, then $g(l-1)/g(l) = e^{-a}$, which is bounded for l.

1. Zero-range process

Let $N_n = \{1, 2, \cdots, n\}$ and let $\Omega_n^* = \{0, 1, 2, \dots\}^{N_n}$. Fix a stochastic matrix $\{P_{ij}\}_{i \leq i,j \leq n}$ with $P_{ij} = P_{ji}$ and $\sum_{j=1}^n P_{ij} = 1$ for all i, which makes the chain irreducible, and let g be a monotone real valued function from nonnegative integers which is called "rate function". Then Zero-range process is the stochastic process with the following generator. For $\eta = (\eta(1), \eta(2), \cdots, \eta(n))$,

$$(L_n f)(\eta) = \sum_{i=1}^n \sum_{j=1}^n P_{ij} g(\eta(i)) \{f(\eta^{i,j}) - f(\eta)\}, \tag{1.1}$$

where f is any bounded function on Ω_n^*, and $\eta^{i,j}$ is given by

$$\begin{cases} \text{if } \eta(i) = 0, \text{ then } \eta^{i,j} = \eta \\ \text{if } \eta(i) \neq 0, \text{ then } \eta^{i,j}(k) = \begin{pmatrix} \eta(i) - 1 & \text{if } k = i \\ \eta(j) + 1 & \text{if } k = j \\ \eta(k) & \text{otherwise.} \end{pmatrix} \end{cases}$$

Informally, we can describe the dynamics as following: If the process is in state η at certain time, then at any site i, each cluster waits for exponential time with rate $g(\eta(i))$ independently to other sites, picks site j with probability P_{ij} and gives one particle to the site so that $\eta(i)$ decreases to $\eta(i) - 1$, while $\eta(j)$ increases to $\eta(j) + 1$. Let $\eta_t \doteq (\eta_t(1), \eta_t(2), \cdots, \eta_t(n)), 0 \leq t < \infty$, be the Markov process with generator L_n. Since η_t preserves the total number of particles, i.e., $\sum_{i=1}^n \eta_t(i) =$

$\sum_{i=1}^{n} \eta_0(i)$ for all t, if we let $\Omega_n^m = \{\eta \in \Omega_n^* : \sum_{i=1}^{n} \eta(i) = m\}, 1 \leq m < \infty$, then the process restricted to Ω_n^m is ergodic, so there is a unique invariant measure on Ω_n^m, say ν_n^m, since P_{ij} is irreducible. Let $Z \doteq (Z_1, Z_2, \cdots, Z_n)$ be the equilibrium zero range process with rate function g and let $Z_n^* = \max_{1 \leq i \leq n} Z_i$.

The following proposition gives the explicit invariant measure on Ω_n^m. The proof can be found in [5, 7].

LEMMA 1.1. *For any rate function $g(l)$, and for any $\eta \in \Omega_n^m$, let*

$$\mu_n^m(\eta) = \prod_{i=1}^{n} \{g!(\eta(i))\}^{-1}, \tag{1.2}$$

where $g!(l) = g(l)g(l-1)g(l-2)\cdots g(1)g(0)$, with convention $g!(0) = g(0) = 1$. Let

$$\nu_n^m(\eta) = \frac{1}{\Gamma}\mu_n^m(\eta), \text{ where } \Gamma = \mu_n^m(\Omega_n^m) = \sum_{\eta \in \Omega_n^m} \mu_n^m(\eta).$$

Then ν_n^m is the equilibrium measure corresponding to $g(l)$.

Let $S(n)$ be the permutation group of n letters. For any $\sigma_n \in S(n)$, and for any $\eta = (\eta(1), \cdots, \eta(n)) \in \Omega_n^m$, let $\sigma_n \eta = (\eta(\sigma_n(1)), \eta(\sigma_n(2)) \cdots \eta(\sigma_n(n)))$.

LEMMA 1.2. *For any $\eta \in \Omega_n^m, \sigma_n \in S(n)$, let ν_n be the invariant measure corresponding to g. Then $\nu_n^m(\eta) = \nu_n^m(\sigma_n \eta)$.*

PROOF. $\nu_n^m(\sigma_n \eta)$ is just the change of order in multiplication of $\nu_n^m(\eta)$. □

2. The main theorems and proofs

In this section, we will prove three main theorems and several preliminary Lemmas. To simplify, we will drop m, n in the notation, if there is no confusion. For example, $\nu = \nu_n^m$ and $\mu = \mu_n^m$.

For any $\eta = (\eta_1, \eta_2, \cdots, \eta_n) \in \Omega_n^m$, let $l_i(\eta) = \sharp\{l : \eta_l = i\}$, i.e., the number of i-clusters in the configuration. Let

$$\Lambda_n^m = \left\{ \bar{a} = (a_1, a_2, \cdots, a_n) : a_i \text{ is a nonnegative integer and} \sum_{i=1}^{n} i a_i = m \right\}$$

For $\bar{a} = (a_1, a_2, \cdots, a_n) \in \Lambda_n^m$, let $A_{\bar{a}} = \{\eta \in \Omega_n^m : l_i(\eta) = a_i, 1 \le i \le n\}$. Note that $A_{\bar{a}}$ is the set of all configurations such that the number of 1-clusters is a_1, the number of 2-clusters is a_2, etc. Then for any $\eta \in A_{\bar{a}}$, by Lemma 1.2,

$$\nu(\eta) = \frac{1}{\mu(\Omega_n^m)} \cdot \frac{1}{g!(1)^{a_1} g!(2)^{a_2} \cdots g!(n)^{a_n}} \tag{2.1}$$

and, therefore,

$$\nu(A_{\bar{a}}) = \frac{|A_{\bar{a}}|}{\mu(\Omega_n^m)} \cdot \frac{1}{g!(1)^{a_1} g!(2)^{a_2} \cdots g!(n)^{a_n}}, \tag{2.2}$$

$$|A_{\bar{a}}| = \binom{n}{a_1} \binom{n - a_1}{a_2} \cdots \binom{n - \sum_{i=1}^{n-1} a_i}{a_n}. \tag{2.3}$$

DEFINITION 2.1. For $\bar{a}, \bar{b} \in \Lambda_n^m$ such that $\bar{a} = (a_1, a_2, \cdots, a_l, 0, 0, \cdots, 0)$ and $\bar{b} = (b_1, b_2, \cdots, b_l, 0, 0, \cdots, 0)$,

(1) we say $\bar{b}$ is an immediate descendent of $\bar{a}$, if $a_l > 0$ and

$$\begin{aligned}&(b_1, b_2, \cdots, b_{l-1}, b_l, 0, 0, \cdots, 0)\\ &= (a_1 - 1, a_2, \cdots, a_{l-1} - 1, a_l + 1, 0, 0, \cdots, 0),\end{aligned}$$

or, if $a_l = 0, a_{l-1} > 0$, and

$$(b_1, b_2, \cdots, b_l, 0, 0, \cdots, 0) = (a_1 - 1, a_2, \cdots, a_{l-1} - 1, 1, 0, 0, \cdots, 0).$$

If $\bar{b}$ is an immediate descendent of $\bar{a}$, then we write $\bar{a} \succ \bar{b}$, and we call the operation obtaining $\bar{b}$ from $\bar{a}$ a descending operation.

(2) If there is a sequence $\bar{a}_1, \bar{a}_2, \cdots, \bar{a}_l$ such that $\bar{a} \succ \bar{a}_1 \succ \bar{a}_2 \succ \cdots \succ \bar{a}_l \succ \bar{b}$, then we write $\bar{a} \succ\succ \bar{b}$ and we say that $\bar{b}$ is a descendent of $\bar{a}$.

(3) We denote the largest nonzero coordinate, i.e. the maximum size of the clusters, of $\bar{a} = (a_1, a_2, \cdots, a_n)$ by $\max(\bar{a})$. For example, if $\bar{a} = (a_1, a_2, \cdots, a_l, 0, \cdots, 0)$, $a_l > 0$ then $\max(\bar{a}) = l$.

REMARK. There are two different types of descending operations. The first operation raises the number of maximum clusters and the second raises the size of the maximum cluster. Note that the operation can change only the size of the maximum cluster and the number of single particles (1-cluster) or possibly the number of clusters of the second largest size.

The following Lemma shows how the weight is changed by the descending operations.

LEMMA 2.2.

(1) $$\frac{\mu(A_{(a_1-1,a_2,\cdots,a_{l-1}-1,a_l+1,0,0,\cdots,0)})}{\mu(A_{(a_1,a_2,\cdots,a_l,0,0,\cdots,0)})} = \frac{g(1)g(l-1)a_1 a_{l-1}}{g(l)(a_l+1)(n-\sum_{i=1}^{l} a_i + 1)}, \quad l > 2,$$

(2) $$\frac{\mu(A_{(a_1-1,a_2,\cdots,a_{l-1}-1,1,0,0,\cdots,0)})}{\mu(A_{(a_1,a_2,\cdots,a_{l-1},0,0,\cdots,0)})} = \frac{g(1)g(l-1)a_1 a_{l-1}}{g(l)(n-\sum_{i=1}^{l} a_i + 1)}, \quad l > 2,$$

(3) $$\frac{\mu(A_{(a_1-2,1,0,0,\cdots,0)})}{\mu(A_{(a_1,0,0,\cdots,0)})} = \frac{g(1)^2 a_1(a_1-1)(n-1)}{g(2)(n-a_1+1)(n-a_1+2)},$$

(4) $$\frac{\mu(A_{(a_1-2(l+1),l+1,0,0,\cdots,0)})}{\mu(A_{(a_1-2l,l,0,\cdots,0)})} = \frac{g(1)^2(a_1-2l)(a_1-2l-1)}{g(2)(l+1)(n-a_1+l+1)}.$$

PROOF. It is a straightforward application of (2.1), (2.2), and (2.3). For (1), since $l > 2$,

$$\frac{\mu(A_{(a_1-1,a_2,\cdots,a_{l-1}-1,a_l+1,0,0,\cdots,0)})}{\mu(A_{(a_1,a_2,\cdots,a_l,0,0,\cdots,0)})}$$

$$= \frac{\nu(A_{(a_1-1,a_2,\cdots,a_{l-1}-1,a_l+1,0,0,\cdots,0)})}{\nu(A_{(a_1,a_2,\cdots,a_l,0,0,\cdots,0)})}$$

$$= \frac{|A_{(a_1-1,a_2,\cdots,a_{l-1}-1,a_l+1,0,0,\cdots,0)}|g!(1)^{a_1}g!(2)^{a_2}\cdots g!(l-1)^{a_{l-1}}g!(l)^{a_l}}{|A_{(a_1,a_2,\cdots,a_l,0,0,\cdots,0)}|g!(1)^{a_1-1}g!(2)^{a_2}\cdots g!(l-1)^{a_{l-1}-1}g!(l)^{a_l+1}}$$

$$= \frac{\binom{n-\sum_{i=2}^{l-2} a_i}{a_1-1}\binom{n-\sum_{i=1}^{l-2} a_i+1}{a_{l-1}-1}\binom{n-\sum_{i=1}^{l-1} a_i+2}{a_l+1} g(1)g(l-1)}{\binom{n-\sum_{i=2}^{l-2} a_i}{a_1}\binom{n-\sum_{i=1}^{l-2} a_i}{a_{l-1}}\binom{n-\sum_{i=1}^{l-1} a_i}{a_l} g(l)},$$

which becomes, after simple calculation,

$$\frac{g(1)g(l-1)a_1a_{l-1}}{g(l)(a_l+1)(n-\sum_{i=1}^{l}a_i+1)}.$$

Other equations are similar. □

LEMMA 2.3. *For $\bar{a},\bar{b}\in\Lambda_n^m$ such that $\bar{a}\succ\succ\bar{b}$, there exists a unique sequence $\bar{a}_1,\bar{a}_2,\cdots,\bar{a}_l$ satisfying $\bar{a}\succ\bar{a}_1\succ\bar{a}_2\succ\cdots\succ\bar{a}_l\succ\bar{b}$.*

PROOF. Suppose there are two sequences $\{\bar{a}_i\}_{i=1}^{l},\{\bar{b}_i\}_{i=1}^{m}$ such that $\bar{a}\succ\bar{a}_1\succ\bar{a}_2\succ\cdots\succ\bar{a}_l\succ\bar{b}$ and $\bar{a}\succ\bar{b}_1\succ\bar{b}_2\succ\cdots\succ\bar{b}_m\succ\bar{b}$. Let l_0 be the first number satisfying $\bar{a}_{l_0}=\bar{b}_{l_0}$ and $\bar{a}_{l_0+1}\neq\bar{b}_{l_0+1}$. Let $\bar{a}_{l_0}=\bar{b}_{l_0}=(a_1,a_2,\cdots,a_r,0,0,\cdots,0)$ with $a_r>0$. Note that descending operation yields only two possible states, $\bar{c}_1=(a_1-1,a_2,\cdots,a_{r-1}-1,a_r+1,0,\cdots,0)$ and $\bar{c}_2=(a_1-1,a_2,\cdots,a_{r-1},a_r-1,1,0,\cdots,0)$. Since $\bar{a}_{l_0+1}\neq\bar{b}_{l_0+1}$, if $r>2$, without loss of generality, assume $\bar{a}_{l_0+1}=\bar{c}_1,\bar{b}_{l_0+1}=\bar{c}_2$.

Since any descending operation from $\bar{b}_{l_0+1}$ cannot change the $(r-1)$-th coordinate a_{r-1} and any descending operation from $\bar{a}_{l_0+1}$ cannot raise the $(r-1)$-th coordinate $a_{r-1}-1$, the $(r-1)$-th coordinate's of any descendent of $\bar{a}_{l_0+1}$ and any descendent of $\bar{b}_{l_0+1}$ cannot be the same.

Thus, a contradiction follows since $\bar{b}$ is a descendent of $\bar{a}_{l_0+1}$ and $\bar{b}_{l_0+1}$. For the case $r=2$, the number of clusters of any descendant of $\bar{c}_2$ is a_2, while that of $\bar{c}_1$ is bigger than or equal to a_2+1. The case $r=1$ is impossible. □

LEMMA 2.4. *For any $\bar{a}=(a_1,a_2,\cdots,a_l,0,0,\cdots,0)$, let $a_0=\sum_{i=1}^{l}ia_i$ and $\bar{a}_0=(a_0,0,0,\cdots,0)$, then $\bar{a}_0\succ\bar{a}$.*

PROOF. We can construct a sequence of descending operations as following.

$$\begin{aligned}\bar{a}_0=(a_0,0,0,\cdots,0)&=\left(\sum_{i=1}^{l}ia_i,0,0,\cdots,0\right)\\&\succ\left(\sum_{i=1}^{l}ia_i-2,1,0,0,\cdots,0\right)\cdots\\&\succ\left(\sum_{i=1}^{l}ia_i-2\sum_{i=2}^{l}a_i,\sum_{i=2}^{l}a_i,0,0,\cdots,0\right)\\&\succ\left(\sum_{i=1}^{l}ia_i-2\sum_{i=2}^{l}a_i-1,\sum_{i=2}^{l}a_i-1,1,0,0,\cdots,0\right)\end{aligned}$$

$$\succ \left(\sum_{i=1}^{l} ia_i - 2\sum_{i=2}^{l} a_i - 2, \sum_{i=2}^{l} a_i - 2, 2, 0, 0, \cdots, 0\right) \cdots$$

$$\succ \left(\sum_{i=1}^{l} ia_i - 2\sum_{i=2}^{l} a_i - \sum_{i=3}^{l} a_i, a_2, \sum_{i=3}^{l} a_i, 0, 0, \cdots, 0\right)$$

$$\succ \left(\sum_{i=1}^{l} ia_i - 2\sum_{i=2}^{l} a_i - \sum_{i=3}^{l} a_i - 1, a_2, \sum_{i=3}^{l} a_i - 1, 1, 0, 0, \cdots, 0\right) \cdots$$

$$\succ \left(\sum_{i=1}^{l} ia_i - 2\sum_{i=2}^{l} a_i - \sum_{i=3}^{l} a_i - \sum_{i=4}^{l} a_i, a_2, a_3, \sum_{i=4}^{l} a_i, 0, 0, \cdots, 0 \cdots\right)$$

$$\succ \left(\sum_{i=1}^{l} ia_i - 2\sum_{i=2}^{l} a_i - \sum_{i=3}^{l} a_i - \cdots \right.$$

$$\left. - \sum_{i=l-1}^{l} a_i + a_l, a_2, a_3, \cdots, a_l, 0, 0, \cdots, 0\right)$$

$$= (a_1, a_2, \cdots, a_l, 0, 0, \cdots, 0). \qquad \square$$

Note that, since in each operation step there are at most two different choices, $|\Lambda_n^m| \le 2^{m+1}$.

DEFINITION 2.5. For $\bar{a}, \bar{b} \in \Lambda_n^m$, such that $\bar{a} \succ\succ \bar{b}$, we define $s(\bar{a}, \bar{b})$ by the number of descending operations applied for $\bar{a}$ to get $\bar{b}$, i.e. if $\bar{a} \succ \bar{a}_1 \succ \bar{a}_2 \succ \cdots \succ \bar{a}_l \succ \bar{b}$, then $s(\bar{a}, \bar{b}) = l + 1$.

REMARK. The inverse of descending operation is unique in the sense that, if $\bar{b}_1 \succ \bar{a}$, $\bar{b}_2 \succ \bar{a}$ and $s(\bar{b}_1, \bar{a}) = s(\bar{b}_2, \bar{a})$, then $\bar{b}_1 = \bar{b}_2$.

Now recall $Z \doteq (Z_1, Z_2, \cdots, Z_n)$, the equilibrium zero range process with rate function g, and $Z_n^* = \max_{1 \le i \le n} Z_i$. From now on, we will assume there exists M such that $\frac{g(1)g(l-1)}{g(l)} \le M$ for all l.

THEOREM 2.6. *Suppose* $m \sim n^\beta, 0 < \beta < \frac{1}{2}$. *Then* $Prob(Z_n^* \ge 2) \to 0$ *as* $n \to \infty$.

PROOF. Let $\bar{a}_0 = (m, 0, 0, \cdots, 0)$ and let

$$G_k = \{\bar{b} : \bar{a}_0 \succ\succ \bar{b}, s(\bar{a}, \bar{b}) = k\},$$
$$A_{\bar{a}_0}^k = \{A_{\bar{b}} : \bar{b} \in G_k\},$$
$$\bar{A}_{\bar{a}_0}^k = \cup_{\bar{b} \in G_k} A_{\bar{b}}$$

for $k = 0, 1, 2, \cdots$. Then, for any $\epsilon > 0$,

$$\begin{aligned}
&P(Z_n^* \geq 2)\\
&= \nu(\cup_{k\geq 1}\bar{A}^k_{\bar{a}_0})\\
&= \frac{\mu(\cup_{k\geq 1}\bar{A}^k_{\bar{a}_0})}{\mu(\Omega_n^m)}\\
&\leq \frac{\mu(\cup_{k\geq 1}\bar{A}^k_{\bar{a}_0})}{\mu(\bar{A}^0_{\bar{a}_0})}\\
&\leq \frac{\sum_{k\geq 1}\mu(\bar{A}^k_{\bar{a}_0})}{\mu(\bar{A}^0_{\bar{a}_0})}\\
&\leq \frac{\mu(\bar{A}^1_{\bar{a}_0})}{\mu(\bar{A}^0_{\bar{a}_0})} + \frac{\mu(\bar{A}^1_{\bar{a}_0})}{\mu(\bar{A}^0_{\bar{a}_0})}\cdot\frac{\mu(\bar{A}^2_{\bar{a}_0})}{\mu(\bar{A}^1_{\bar{a}_0})} + \frac{\mu(\bar{A}^1_{\bar{a}_0})}{\mu(\bar{A}^0_{\bar{a}_0})}\cdot\frac{\mu(\bar{A}^2_{\bar{a}_0})}{\mu(\bar{A}^1_{\bar{a}_0})}\cdot\frac{\mu(\bar{A}^3_{\bar{a}_0})}{\mu(\bar{A}^2_{\bar{a}_0})} + \cdots.
\end{aligned}$$

Note that, for $k = 0, 1, 2, \cdots$, $|G_{k+1}| \leq 2|G_k|$, since in each operation step there are at most two different choices. Moreover, for any $\bar{a} \in G_k, \bar{b} \in G_{k+1}, \bar{a} \succ \bar{b}$, any case in Lemma 2.2 can be bounded by

$$\frac{\mu(A_{\bar{b}})}{\mu(A_{\bar{a}})} \leq 2M\frac{m^2}{n} < \frac{\epsilon}{4},$$

for large n. Therefore,

$$\frac{\mu(\bar{A}^{k+1}_{\bar{a}_0})}{\mu(\bar{A}^k_{\bar{a}_0})} = \frac{\sum_{\bar{b}\in G_{k+1}}\mu(A_{\bar{b}})}{\sum_{\bar{b}\in G_k}\mu(A_{\bar{b}})} < \frac{\epsilon}{4}\cdot 2 = \frac{\epsilon}{2},$$

and

$$P(Z_n^* \geq 2) < \frac{\epsilon}{2} + (\frac{\epsilon}{2})^2 + \cdots \leq \epsilon.$$

That is, $Prob(Z_n^* \geq 2) \to 0$ as $n \to \infty$. □

Let $Y^n = \sharp\{i : Z_i^n = 2\}$, i.e., the number of 2-clusters.

THEOREM 2.7. *If $m \sim n^{\frac{1}{2}}$, then $Prob(Z_n^* \geq 3) \to 0$ as $n \to \infty$ and $Y^n \to Pois(\frac{g(1)^2}{g(2)})$.*

The proof can be completed after proving several Lemmas concerning the tightness of the sequence.

LEMMA 2.8. *If $m \sim n^{\frac{1}{2}}$, then for any $\epsilon > 0$, small, there exists r such that $Prob(Z_n^* > r) < \epsilon$ for sufficiently large n.*

PROOF. For given ϵ small, choose $r > 2$ such that $\frac{2M}{r-1} < \frac{\epsilon}{4}$. Let

$$B = \{A_{\bar{a}} : \bar{a} \in \Lambda_m^n, \max(\bar{a}) = r\},$$

$$G^k = \{\bar{b} : \bar{a} \succ\succ \bar{b}, s(\bar{a}, \bar{b}) = k \text{ for some } \bar{a} \text{ such that } \max(\bar{a}) = r\},$$

$$B^k = \{A_{\bar{b}} : \bar{b} \in G^k\},$$

$$\bar{B}^k = \cup_{\bar{b} \in G_k} A_{\bar{b}},$$

for $k = 0, 1, 2, \cdots$. Note that G^k is uniquely defined by Lemma 2.3 and the Remark after Definition 2.5. Then $|G^{k+1}| \leq 2|G^k|$ and for any $\bar{a} \in G^k, \bar{b} \in G^{k+1}, \bar{a} \succ \bar{b}$, since a_{l-1} in equations (1) and (2) of Lemma 2.2 satisfies $a_{l-1} \leq \frac{m}{r-1}$,

$$\frac{\mu(A_{\bar{b}})}{\mu(A_{\bar{a}})} \leq 2M \frac{n^{1/2} n^{1/2}/(r-1)}{n} \leq \frac{2M}{r-1} < \frac{\epsilon}{4},$$

for large n. Therefore, similarly to Theorem 2.6,

$$\begin{aligned} P(Z_n^* > r) &\leq \frac{\mu(\cup_{k \geq 1} \bar{B}^k)}{\mu(\bar{B}^0)} \\ &\leq \frac{\sum_{k \geq 1} \mu(\bar{B}^k)}{\mu(\bar{B}^0)} \\ &\leq \frac{\mu(\bar{B}^1)}{\mu(\bar{B}^0)} + \frac{\mu(\bar{B}^1)}{\mu(\bar{B}^0)} \frac{\mu(\bar{B}^2)}{\mu(\bar{B}^1)} \cdots \\ &< \frac{\epsilon}{2} + (\frac{\epsilon}{2})^2 + \cdots \\ &\leq \epsilon, \end{aligned} \tag{2.4}$$

for the choice of r. □

LEMMA 2.9. *If $m \sim n^{\frac{1}{2}}$, then $Prob(Z_n^* \geq 3) \to 0$ as $n \to \infty$.*

PROOF. For given ϵ small, choose the same r in the proof of Lemma 2.8. Let $\bar{a}_i = (m-2i-2, i+1, 0, 0, \cdots, 0)$ and $\bar{b}_i = (m-2i-3, i, 1, 0, 0, \cdots, 0), i = 0, 1, 2, \cdots$.

Let

$$G_i = \{\bar{a} : \bar{b}_i \succ\succ \bar{a}, s(\bar{b}_i, \bar{a}) \geq 0\},$$

$$G_i^k = \{\bar{a} : \bar{b}_i \succ\succ \bar{a}, s(\bar{b}_i, \bar{a}) = k\},$$

$$B_i = \{A_{\bar{a}} : \bar{a} \in G_i\}$$

and let

$$B_i^k = \{A_{\bar{a}} : \bar{a} \in G_i^k\}, k = 0, 1, 2, \cdots.$$

Let $\bar{B}_i = \cup_{\bar{a} \in G_i} A_{\bar{a}}$ and $\bar{B}_i^k = \cup_{\bar{a} \in G_i^k} A_{\bar{a}}, k = 0, 1, 2, \cdots$.

First, consider the case $i \leq \epsilon n^{1/2}/8M \doteq C$. Then

$$\begin{aligned} \nu(\cup_{i \leq C} \bar{B}_i) &= \frac{\mu(\cup_{i \leq C} \bar{B}_i)}{\mu(\Omega_n^m)} \\ &\leq \frac{\mu(\cup_{i \leq C} \bar{B}_i)}{\mu(\cup_{i \leq C} A_{\bar{a}_i})} \\ &= \frac{\sum_{i \leq C} \mu(\bar{B}_i)}{\sum_{i \leq C} \mu(A_{\bar{a}_i})}. \end{aligned} \tag{2.5}$$

Note that for $k = 0, 1, 2, \cdots$, $|G_i^{k+1}| \leq 2|G_i^k|$ and for any $\bar{a} \in G_i^k, \bar{b} \in G_i^{k+1}, \bar{a} \succ \bar{b}$, since $a_{l-1} \leq C$ in equations (1) and (2) of Lemma 2.2,

$$\frac{\mu(A_{\bar{b}})}{\mu(A_{\bar{a}})} \leq 2M \frac{n^{1/2} \epsilon n^{1/2}/8M}{n} < \frac{\epsilon}{4}, \text{ and } \frac{\mu(\bar{b}_i)}{\mu(\bar{a}_i)} \leq 2M \frac{n^{1/2} \epsilon n^{1/2}/8M}{n} < \frac{\epsilon}{4}$$

for large n. Therefore, similarly to Theorem 2.6

$$\begin{aligned} \frac{\mu(\bar{B}_i)}{\mu(A_{\bar{a}_i})} &= \frac{\mu(\cup_{k \geq 0} \bar{B}_i^k)}{\mu(A_{\bar{a}_i})} \\ &\leq \frac{\sum_{k \geq 0} \mu(\bar{B}_i^k)}{\mu(A_{\bar{a}_i})} \\ &< \frac{\mu(\bar{B}_i^0)}{\mu(A_{\bar{a}_i})} + \frac{\mu(\bar{B}_i^0)}{\mu(A_{\bar{a}_i})} \frac{\mu(\bar{B}_i^1)}{\mu(\bar{B}_i^0)} + \cdots \\ &\leq \left(\frac{\epsilon}{2}\right) + \left(\frac{\epsilon}{2}\right)^2 + \left(\frac{\epsilon}{2}\right)^3 + \cdots \\ &\leq \epsilon. \end{aligned} \tag{2.6}$$

Therefore,

$$\nu(\cup_{i\le C}\bar{B}_i) < \epsilon.$$

Now for $i > C$, note that, from equation (4) of Lemma 2.2,

$$\frac{\mu(A_{\bar{a}_i})}{\mu(A_{\bar{a}_r})} \le \frac{\mu(A_{\bar{a}_{r+1}})}{\mu(A_{\bar{a}_r})}\cdot\frac{\mu(A_{\bar{a}_{r+2}})}{\mu(A_{\bar{a}_{r+1}})}\cdots\frac{\mu(A_{\bar{a}_i})}{\mu(A_{\bar{a}_{i-1}})} \le \frac{r!(2M)^i}{i!},$$

for large n.

Let

$$\begin{aligned} H_i &= \{\bar{a} : \max(\bar{a}_i) \le r\},\\ C_i &= \{A_{\bar{a}} \in B_i : \bar{a} \in H_i\},\\ \bar{C}_i &= \cup_{\bar{a}\in H_i} A_{\bar{a}}. \end{aligned}$$

For any $\bar{a} \in \bar{C}_i$, there exists $s \le m$ and $\bar{c}_k$'s such that

$$\bar{a}_i \succ \bar{b}_i = \bar{c}_1 \succ \bar{c}_2 \cdots \succ \bar{c}_s = \bar{a}.$$

Note that among the ratio's $\frac{\mu(\bar{b}_i)}{\mu(\bar{a}_i)}$, $\frac{\mu(\bar{c}_{k+1})}{\mu(\bar{c}_k)}, k = 1, 2, \cdots, s-1$, which are bounded by $2M$ for large n, at most r^2 terms are bigger than $\epsilon/4$, since only possible cases are $l \le r$ and the maximum cluster size $\le r$ in Lemma 2.2 (1). Hence $\mu(\bar{a}) \le \mu(\bar{b}_i)(2M)^{r^2}$, for small ϵ, e.g., $\frac{\epsilon}{4} \le 1$.

Therefore, for large n, since $|\Lambda_m^n| \le 2^{m+1}$ as described after Lemma 2.4,

$$\begin{aligned} \frac{\mu(\cup_{i>C}\bar{C}_i)}{\mu(A_{\bar{a}_r})} &\le \frac{\sum_{i>C}\mu(\bar{C}_i)}{\mu(A_{\bar{a}_r})}\\ &\le \sum_{i>C}\frac{\mu(A_{\bar{a}_i})}{\mu(A_{\bar{a}_r})}\cdot\frac{\mu(\bar{C}_i)}{\mu(A_{\bar{a}_i})}\\ &\le \sum_{i>C}\frac{r!(2M)^i}{i!}(2M)^{r^2}2^{m+1}\\ &\le m\max_{i>C}\frac{r!(2M)^i}{i!}(2M)^{r^2}2^{m+1}\\ &\sim \max_{i>C}\frac{r!(2M)^i e^i}{\sqrt{2\pi i}i^i}(2M)^{r^2}2^{m+1} \text{ (by Sterling formula)}\\ &\le \frac{d_1(d_2)^m}{i_0^{i_0}}, \end{aligned} \tag{2.7}$$

for some constants d_1, d_2 independent of n and $i_0 > C$. Recall $m \sim n^{1/2}, C = \epsilon n^{1/2}/8M$, and conclude that the final term tends to 0, as $n \to \infty$. Now the proof is completed by Lemma 2.8. □

LEMMA 2.10. *If $m \sim n^{\frac{1}{2}}$, for any $\epsilon > 0$ there exists r such that $Prob\{Y^n \geq r\} \leq \epsilon$.*

PROOF. It suffices to prove that for the same r in the previous Lemma,

$$\frac{\mu(\cup_{i>r} A_{\bar{a}_i})}{\mu(A_{\bar{a}_r})} < \epsilon.$$

Indeed,

$$\begin{aligned}\frac{\mu(\cup_{i>r} A_{\bar{a}_i})}{\mu(A_{\bar{a}_r})} &= \frac{\mu(A_{\bar{a}_{r+1}})}{\mu(A_{\bar{a}_r})} + \frac{\mu(A_{\bar{a}_{r+1}})}{\mu(A_{\bar{a}_r})} \cdot \frac{\mu(A_{\bar{a}_{r+2}})}{\mu(A_{\bar{a}_{r+1}})} + \cdots \\ &\leq \left(\frac{\epsilon}{4}\right) + \left(\frac{\epsilon}{4}\right)^2 + \cdots \\ &< \epsilon.\end{aligned}$$

□

PROOF OF THEOREM 2.7. Tightness is done from Lemma 2.9 and Lemma 2.10. Now

$$\begin{aligned}\frac{\mu(A_{\bar{a}_l})}{\mu(A_{\bar{a}_0})} &= \frac{g(1)^2 m(m-1)(n-1)}{g(2)(n-m+1)(n-m+2)} \\ &\cdot \frac{g(1)^2(m-2)(m-3)}{2g(2)(n-m+2)} \cdot \frac{g(1)^2(m-4)(m-5)}{3g(2)(n-m+3)} \\ &\cdots \frac{g(1)^2(m-2l+2))(m-2l+1)}{lg(2)(n-m+l)} \to \left(\frac{g(1)^2}{g(2)}\right)^l \frac{1}{l!}.\end{aligned} \tag{2.8}$$

That is,

$$\frac{\text{Prob}\{Y^n = l\}}{\text{Prob}\{Y^n = 0\}} \to \left(\frac{g(1)^2}{g(2)}\right)^l \frac{1}{l!},$$

and the proof is done. □

THEOREM 2.11. *If $m \sim n^\beta, \frac{1}{2} < \beta < 1$, then for any l, $Prob\{Y^n \leq l\} \to 0$ as $n \to \infty$.*

PROOF. With the same notation of the proof of Lemma 2.9,

$$\frac{\mu(A_{\bar{a}_{i+1}})}{\mu(A_{\bar{a}_i})} \geq \frac{g(1)^2(m-2i)(m-2i+1)}{g(2)(i+1)(n-m+i+1)}, \quad i \geq 1, \tag{2.9}$$

and

$$\frac{\mu(A_{\bar{a}_1})}{\mu(A_{\bar{a}_0})} \geq \frac{g(1)^2 m(m-1)(n-1)}{g(2)(n-m+1)(n-m+2)}$$

The both terms are made to be larger then $2/\epsilon$ for sufficiently large n. Therefore,

$$\begin{aligned}
\frac{\mu(\cup_{i=0}^{l} A_{\bar{a}_i})}{\mu(\Omega_n^m)} &\leq \frac{\mu(\cup_{i=0}^{l} A_{\bar{a}_i})}{\mu(A_{\bar{a}_{l+1}})} \\
&\leq \frac{\mu(A_{\bar{a}_l})}{\mu(A_{\bar{a}_{l+1}})} + \frac{\mu(A_{\bar{a}_l})}{\mu(A_{\bar{a}_{l+1}})} \cdot \frac{\mu(A_{\bar{a}_{l-1}})}{\mu(A_{\bar{a}_l})} + \cdots \qquad (2.10) \\
&\leq \left(\frac{\epsilon}{2}\right) + \left(\frac{\epsilon}{2}\right)^2 + \cdots \\
&\leq \epsilon.
\end{aligned}$$

□

References

[1] D. Aldous, *Deterministic and stochastic models for coalescence (aggregation, coagulation): A review of the mean-field theory for probabilists*, Bernoulli **5** (1999), 3–48.

[2] R. Arratia and S. Tavaré, *Independent process approximations for random combinatorial structures*, Adv. Math. **104** (1994), 90–154.

[3] I. Jeon, *Existence of gelling solutions for coagulation fragmentation equations*, Comm. Math. Phys. **194** (1998), 541–567.

[4] I. Jeon and P. March, *Condensation transition for zero range invariant measures*, In Stochastic models. Proceedings of the International Conference on Stochastic Models in Honor of Professor Donald A. Dawson (Luis G. Gorostiza, B. Gail Ivanoff, eds.) (2000), 233–244.

[5] I. Jeon, P. March, and B. Pittel, *Size of the largest cluster under Zero-range invariant measures*, Ann. Probab. **28** (2000), 1162–1194.

[6] T. M. Liggett, *Interacting Particle Systems*, Springer-Verlag, New York, 1985.

[7] F. Spitzer, *Interaction of Markov processes*, Adv. Math. **5** (1970), 246–290.

Department of Mathematics
The Catholic University of Korea
Bucheon 420-832, Korea
E-mail: injeon@catholic.ac.kr