A first principles derivation of animal group size distributions

Qi Ma, Anders Johansson, David J.T. Sumpter

www.elsevier.com/locate/yjtbi

A first principles derivation of animal group size distributions

Qi Ma, Anders Johansson, David J.T. Sumpter

PII: S0022-5193(11)00232-3
DOI: doi:10.1016/j.jtbi.2011.04.031
Reference: YJTBI6462

To appear in: *Journal of Theoretical Biology*

Received date: 27 October 2010
Revised date: 28 April 2011
Accepted date: 29 April 2011

Cite this article as: Qi Ma, Anders Johansson and David J.T. Sumpter, A first principles derivation of animal group size distributions, *Journal of Theoretical Biology*, doi:10.1016/j.jtbi.2011.04.031

A first principles derivation of animal group size distributions

Qi Ma*, Anders Johansson, David J. T. Sumpter

Mathematics Department, Uppsala University, Box 480, 751 06 Uppsala, Sweden

Abstract

Several empirical studies have shown that the animal group size distribution of many species can be well fit by power laws with exponential truncation. A striking empirical result due to H-S Niwa is that the exponent in these power laws is one and the truncation is determined by the average group size experienced by an individual. This distribution is known as the logarithmic distribution. In this paper we provide first principles derivations of the logarithmic distribution and other truncated power laws using a site-based merge and split framework. In particular, we investigate two such models. Firstly, we look at a model in which groups merge whenever they meet but split with a constant probability per time step. This generates a distribution similar, but not identical to the logarithmic distribution. Secondly, we propose a model, based on preferential attachment, that produces the logarithmic distribution exactly. Our derivation helps explain why logarithmic distributions are so widely observed in nature. The derivation also allows us to link splitting and joining behavior to the exponent and truncation parameters in power laws.

1. Introduction

Animals are often found in groups. Fish school, birds flocks and insects swarms are ubiquitous examples. Being in groups benefits individuals in several ways. Amongst other things, it can enhance their foraging efficiency, reduce their chance of being captured by a predator, and help to conserve

*Corresponding author.
Email address: qi@math.uu.se (Qi Ma)

energy (Foster et al., 1981; Parrish, 1989; Barbraud et al., 2001; Sumpter, 2010). However, animal groups do not always have a typical size, and group sizes often have large variation both among and within species (Gerard et al., 2002; Bonabeau et al., 1999). This property brings up several interesting questions: How are animal group sizes distributed? Are there any common patterns for these distributions? How do different distributions form?

A wide range of models have been suggested for distribution of animal group sizes. The earliest such models proposed a single stable group size, around which the size of the groups may fluctuate (Sibly, 1983; Beauchamp et al., 2004; Clark et al., 1986; Mottley et al., 2000). For example, Caraco proposed the Poisson and negative binomial distributions as one or, respectively, two parameter models of group size distributions (Caraco et al., 1980). Such distributions have a single peak at a group size somewhat larger than one and a narrow variation around this maximum.

Observed group size distributions of many animal species do not follow such Poisson or negative binomial distributions (Gerard et al., 2002; Bonabeau et al., 1999). Most importantly, the variation in group sizes is usually much wider than predicted by these distributions. The geometric distribution is a special case of the negative binomial, obtained by maximizing the variance. In his influential review, Okubo predicted that group sizes should follow a geometric distribution and presented a number of empirical cases where this relationship held (Okubo, 1986). However, even the geometric distribution fails to capture the large variation in group sizes observed for many species. In particular, several studies have shown that many species follow power law distributions over a number of orders of magnitude (Bonabeau et al., 1995, 1999; Sjöberg et al., 2000).

A natural question emerges about how these distributions arise from interactions between individual animals. Bonabeau and Dagorn proposed a model for animal grouping based on a single assumption: if groups meet they always merge to form a larger group (Bonabeau et al., 1995, 1999). Their model predicts power law distributions of group sizes, which again appeared consistent with some observational data of fish and mammals. However, in their model individuals need to be continuously added in order to get a power law. Furthermore, unless they add spatial structure to the model the power law exponent is always -2. Even with spatial structure, where such models give power laws with exponents between $-4/3$ and $-3/2$ (Takayasu et al., 1988), the dynamics of their model are difficult to motivate from a biological perspective. Although they suggest that power laws should be truncated

by faster decreasing functions such as an exponential function, no natural explanation on the cutoff is given. In particular, Bonabeau and Dagorn did not provide a method for relating their model assumptions to the point in the distribution at which the cutoff should occur.

Niwa proposed a site-based model (Niwa, 2003), which can be described as follows. Assume that space is divided into s sites on which a total of Φ individuals are initially randomly distributed. Every site is either empty or occupied by one group. At each discrete time step, (1) each group with size larger than 1 has a probability p of splitting into a pair of groups. When a group splits, the size of the two components is chosen uniformly at random; (2) all groups move to a new randomly chosen site. If two groups move to a same site, they merge to a new group with a size equal to the sum of the two groups. The same rule holds if three or more groups meet.

Niwa predicted the following stable group size distribution for his model:

$$W(n) \sim n^{-1} \exp\left(-\frac{n}{N_P}\left(1 - \frac{\exp(-\frac{n}{N_P})}{2}\right)\right) \tag{1}$$

where $W(n)$ is the probability density function of group of sizes. Niwa noted that the final term $(1 - \exp(-n/N_P))$ makes little difference to the quality of data fitting compared to a simpler distribution

$$W(n) \sim n^{-1} \exp(-\frac{n}{N_P}) \tag{2}$$

Here, if n is considered to be a continuous variable, then the single parameter N_P denotes the expected size of groups in which an arbitrary individual engages, i.e.

$$N_P = \frac{\int n^2 W(n) dn}{\int n W(n) dn} \tag{3}$$

This is a spectacularly simple yet powerful result. Simply by measuring the average group size experienced by an individual, Niwa is able to recreate the whole distribution of group sizes. Furthermore, Niwa showed that this distribution fits well with data of several fish species, and to a lesser degree that of several mammal species. More recently, Griesser et al. (2011) have shown that it fits data for house sparrows too.

There are however a number of limitations with Niwa's derivation of the above distribution. Firstly, he used simulation results to establish how fluctuations in group size change as a function of group size itself. This relationship

is then used in a stochastic differential equation approximation of the site-based model. However, since this initial result is obtained by simulation, the further results based on the stochastic differential equation lack a rigorous foundation. Secondly, given the stochastic differential equation, the derivation of the potential function does not appear correct. Specifically, equations (8), (9) and (11) in Niwa's paper do not appear to follow. Thirdly, Niwa states self-consistency of the model by calculating N_P in continuous sense, while it is more natural to use a discrete distribution. Animals are discrete entities. The discrete calculation gives different result from the continuous one and equation (3) fails to hold.

Given the excellence of Niwa's distribution in explaining observations and its one-parameter simplicity, it is important that we have a clear derivation linking individual behavior of animals to the predicted group size distribution. In this paper we give several such derivations and at the same time address the above issues in Niwa's original paper.

To make this derivation we use a discrete analogue of a general continuous split-merge model, also called Coagulation-Fragmentation Processes (CFP), discussed by Gueron and Levin (Gueron et al., 1995). They study the following general evolution equation for the density distribution of groups,

$$\begin{aligned}\frac{\partial \tilde{f}(x)}{\partial t} =& -p(x)\tilde{f}(x) - \int_0^\infty \tilde{f}(x)\tilde{f}(z)\psi(x,z)dz \\ &+ \frac{1}{2}\int_0^x \tilde{f}(y)\tilde{f}(x-y)\psi(y,x-y)dy + \int_x^\infty \tilde{f}(y)\varphi(y,x)dy\end{aligned} \tag{4}$$

where $\tilde{f}(x) = f(x,t)$ denotes number of groups of size x, $p(x)$ is the rate of split for groups of size x, $\psi(x,y) = \psi(y,x)$ is a symmetric function denotes merge rate of group of size x and group of size y. The first two terms in (4) account for decrease caused by split of groups of size x and groups of size x merge with another group. $\varphi(x,y) = \varphi(x,x-y)$ denotes the rate of a group of size x splits to two groups whose size are respectively y and $x-y$. The last two terms in (4) account for increment caused by merge of smaller groups to become an x-sized group, and larger groups splitting to become size x.

We derive such an evolution equation for a slightly different version of Niwa's merge and split site-based model and find that it gives a distribution similar to but not identical to that proposed by Niwa. We further discuss several other models which produce a discrete version of Niwa's distribution.

2. The logarithmic distribution

We begin by re-stating equation (2) as a discrete distribution and discuss some of its properties. It is more natural to use a discrete distribution because animal groups consist of integer number of individuals. Equation (2) is a truncated power law with exponent -1 and an exponential tail from a cutoff at N_P. In discrete form this can be written as

$$W(n) = \frac{Ka^n}{n} \tag{5}$$

which is now the probability that a group is of size n. The parameter $K = -1/\ln(1-a)$ is a normalization constant and a is a positive constant smaller than one. This distribution is known as **logarithmic distribution**, originally described by R.A. Fisher for abundance of different species in a random sample of an animal population (Fisher et al., 1943). Figure 1a and 1b show a typical example of this distribution on linear and log-log scales, respectively.

For animal groups, a key property of the logarithmic distribution is its relationship to the average group size experienced by an individual, denoted as N_P . Following Niwa, we define

$$N_P = \frac{\sum_{n=1}^{\infty} n^2 W(n)}{\sum_{n=1}^{\infty} nW(n)} = \frac{\sum_{n=1}^{\infty} na^n}{\sum_{n=1}^{\infty} a^n} = \frac{1}{1-a} \tag{6}$$

Note that the mean group size $\mathbf{E}(n) = Ka/(1-a)$ is always less than N_P. Thus the size of the group in which a (uniformly) randomly chosen individual is in is likely to be larger than the size of a randomly chosen group (see figure 1a,b). Equation (6) further implies that the group size experienced by an individual is geometrically distributed. In other words, the size-biased logarithmic distribution equals a geometric distribution with parameter $(1-a)$ (Patil et al., 1978) .

Furthermore, by equation (6) we know that, unlike in equation (2),

$$a \neq \exp(-1/N_P)$$

although when N_P is large, $a \approx \exp(-1/N_P)$. In any case, the parameters in equation (5) can be expressed simply in terms of average group size experienced by a random individual, i.e.

$$a = 1 - \frac{1}{N_P} \text{ and } K = \frac{1}{\ln(N_P)} \tag{7}$$

These relationships make fitting of this distribution to data straightforward. In order to perform such a fitting one need only estimate N_P directly from data and then compare the overall fit of

$$W(n) = \frac{1}{\ln(N_P)} \frac{(1 - 1/N_P)^n}{n} \tag{8}$$

to the empirical distribution (Griesser et al., 2011).

3. Merge and split model

We now investigate a simpler version of the site-based model proposed by Niwa for how groups split and merge (Niwa, 2003, 2004). Assume a system with fixed total population Φ and s sites. Each site is either empty or occupied by one group. In time interval τ, each group with size larger than one splits to two groups with probability $p\tau$, i.e. p is the instantaneous rate of splitting. The two component groups both move to randomly chosen empty sites. With rate q, each group moves to a randomly chosen site. If this new site is occupied, the moving group members are added to the group at the new site.

We ran computer simulations of the above model using a variation of Gillespie algorithm (Gillespie, 1977). The stable group size distribution for one such simulation is shown in Figure 1(c,d). Repeated runs of the simulation suggested that the group size distribution of this model was independent of initial condition. For these parameter values, the distribution from the simulation was similar to that of Niwa's distribution. Estimating N_P from the simulation and substituting this into equation (8) gives a distribution which, at least qualitatively, looks like the stable group size distribution (i.e. compare figures 1a and 1c).

In order to investigate the nature of the relationship between logarithmic distribution and the merge & split model, we now derive deterministic evolution equations for our stochastic merge and split model. The correspondence between stochastic coagulation-fragmentation models and such deterministic models are studied by Gueron (1998) and Durrett et al. (1999). Let $f(n,t)$ be the number of groups with size n at time t, $Z(t) = \sum_{n=1}^{\infty} f(n,t)$ denotes the total number of groups, $W(n,t)$ is the fraction of groups with size n, therefore $f(n,t) = Z(t)W(n,t)$. For large random systems, W can be treated as the density distribution of group size. Time evolution of $f(n,t)$ is the result

of the balance between increase and decease of the number of groups in size n caused by both splitting and merging. This can be expressed by four terms:

1. Increment by the merge of groups with size i and $n-i$. A group of size i ($i < n$) moves to a site occupied by a group of size $n-i$, they merge to an n-sized group. This event happens at rate $f(i,t)qf(n-i,t)/s$. Groups whose size is less than n all have the chance to merge to an n-sized group, so the rate at which two groups merge to become a group of size n is
$$\sum_{i=1}^{n-1} \frac{q}{s} f(i,t) f(n-i,t)$$
2. Increment from splitting of groups whose size are larger than n. The rate that a group of size n created by splitting of group with size i ($i > n$) is $2f(i,t)p/(i-1)$. Adding up the rate for all groups larger than n, we get
$$\sum_{i=n+1}^{\infty} \frac{2f(i,t)p}{i-1}$$
3. Decrease by a group of size n splitting to two smaller components, the rate of this event is $pf(n,t)$.
4. Decrease by groups of size n merging with another group, which can happen in three ways: a group of size n moves to a site occupied by a group whose size is not n, the corresponding rate of this event is $f(n,t)q(Z(t)-f(n,t))/s$; a group of size n moves to a site occupied by an n-sized group, this will lead to 2 groups' lost for n-sized groups, therefore decreases $f(n,t)$ by $2qf(n,t)(f(n,t)-1)/s$; a group with size other than n moves and lands on a site occupied by an n-sized group, occurs at rate $q(Z(t)-f(n,t))f(n,t)/s$. Adding up these three possibilities gives a decrease in $f(n,t)$ of
$$-\frac{2q}{s} f(n,t)(Z(t)-1)$$

Combining the above four terms, we can obtain the evolution equation for $f(n,t)$, i.e:

$$\begin{aligned} f(n,t+\tau) - f(n,t) = & \sum_{i=1}^{n-1} qf(i,t)\frac{f(n-i,t)}{s}\tau + \sum_{i=n+1}^{\infty} \frac{2pf(i,t)}{i-1}\tau \\ & - pf(n,t)\tau - 2\frac{q}{s} f(n,t)(Z(t)-1)\tau \end{aligned} \tag{9}$$

for $n = 2, 3, 4, ...$ Since groups with only one individual can neither be formed by merging nor can they split to smaller groups, the evolution of $f(1)$ is determined by two terms. Specifically,

$$f(1, t+\tau) - f(1,t) = \sum_{i=2}^{\infty} \frac{2pf(i,t)}{i-1}\tau - \frac{2q}{s} f(1,t)(Z(t)-1)\tau \qquad (10)$$

gives the evolution of groups of size one. Finally,

$$Z(t+\tau) - Z(t) = p(Z(t) - f(1,t))\tau - qZ(t)\frac{Z(t)-1}{s}\tau \qquad (11)$$

gives the mean field evolution of the total group number $Z(t)$. We note that, interpreting our model in terms of Gueron and Levin's (Gueron et al., 1995) evolution equation, i.e. equation (4), the merging rate of groups of size i and $n-i$ merge to a group of size n is $\psi(i, n-i) = 2q/s$, and the splitting rate of a group of size n splitting to groups of size i and $n-i$ is $\varphi(n,i) = 2p/(n-1)$.

We now assume that the above equations have a unique stable group size distribution, namely $f(n,t) = f(n)$ and $cov(f(n), f(m)) = 0$ when $t \to \infty$. As a result, $Z = \sum_{n=1}^{\infty} f(n)$ is constant at steady state. We also assume the system is large enough so that $Z \approx Z - 1$ and since $f(1) = ZW(1)$, we get at steady state

$$Z_* = \frac{ps}{q}(1 - W(1)) \qquad (12)$$

Simulation results confirm that the total number of groups converges to and then fluctuates around the theoretical value of Z_* for a wide range of p values (results not shown).

Treating the total group number $Z = Z_*$ as constant allows us to simplify the evolution equations. Substituting Z_* in to the right hand side of equations (9) and (10) and still assuming $Z_* \approx Z_* - 1$ gives the following iteration expression for $f(n)$ at stationary state,

$$f(1) = \frac{1}{1 - W(1)} \sum_{i=2}^{\infty} \frac{f(i)}{i-1} \qquad (13)$$

$$f(2) = \frac{ps}{q} W(1)(1 - W(1))^2 \left(\frac{W(1)+2}{5 - 2W(1)}\right) \qquad (14)$$

$$f(n) = \frac{(n-1)(3-W(1)-W(1)^2)}{(3-2W(1))(n-1)+2} f(n-1) + \frac{q(n-1)}{ps[(3-2W(1))(n-1)+2]} \sum_{i=1}^{n-2} f(i)(f(n-i)-f(n-1-i)) \tag{15}$$

Note that $f(1) = ZW(1) = ps(1-W(1))W(1)/q$. Thus $f(n)$ $(n = 2, 3, ...)$ can be expressed entirely in terms of $W(1)$ and the model parameters p, q and s.

Using the conservation condition $\sum_{n=1}^{\infty} nf(n) = \Phi$, we can get a numerical solution for $W(1)$. Specifically, for any given set of parameter values, we assume there is a unique $W(1)$ which is consistent with the conservation condition. In practice this appears to hold. Thus given this value for $W(1)$ we can then calculate all $f(n)$ accordingly. The calculated $W(n) = f(n)/Z$ for various values of p are shown in figure 2 and compared to simulation outcomes. The numerical solution of the evolution equation accurately reflects the simulation result over a wide range of parameter values.

Equations (13) to (15) are not consistent with the logarithmic distribution (i.e. equation 5). If they were we should have the ratio

$$R(n) \equiv \frac{nW(n)}{(n-1)W(n-1)} = a \tag{16}$$

for some constant a. Such a solution simply does not hold for equation (15). We thus reject the idea that the logarithmic distribution represented by equation (5) is the exact solution to our merge and split model.

To see the difference between the logarithmic distribution and the real solution of our model, we compare $R(n)$ to: (1) its value estimated from the numerical solution of equations (13) to (15); (2) that arising from our discussion of logarithmic distribution, namely $a = 1 - 1/N_P$; and (3) to the value $a = \exp(-1/N_P)$ suggested by Niwa and consistent with equation (2). Figure 3 displays this difference for various splitting rates. When p is small, $R(n)$ from the simulation is approximately constant and equal to both $\exp(-1/N_P)$ and $1-1/N_P$, except for at very small n. As groups split faster, $R(n)$ is no longer a constant, but still approaches a constant close to $1-1/N_P$ as n becomes larger. When p is large, intuitively there will be more small groups, and N_P is smaller. Here the approximation by a constant becomes poorer and further away from $1 - 1/N_P$. However, in this case $\exp(-1/N_P)$ is visibly larger than $1-1/N_P$ (figure 3d) and as a result $\exp(-1/N_P)$ gives a

better match to $R(n)$ calculated from the simulations. This result goes some way to explaining why equation (2) gave a good approximation of the group size distribution produced by his site-based model (Niwa, 2003).

Given the importance of the parameter N_P it would be useful to have this quantity as a function of the site-based model's parameters, namely p, q, Φ and s. Niwa (2004) gave an estimation of N_P as

$$N_P = 4c_0 \frac{q\Phi}{sp} \tag{17}$$

where $c_0 = 1.54$ was estimated numerically from simulations. Figure 4a compares N_P from the simulation with equation (17). We see that this gives a good fit, based on one fitted parameter.

We can remove the need to fit a parameter such as c_0 in this relationship, by estimating N_P directly from equation (8). We use $Z_* = \Phi/\mathbf{E}(n) = ps(1 - W(1))/q$ and substitute $W(1) = Ka$ along with equations (7) to get an implicit expression

$$N_P = \frac{ps(N_P - 1)^2}{ps(N_P)\ln N_P - \Phi q(\ln N_P)^2} \tag{18}$$

N_P can then be estimated by solving this equation numerically. Similarly, we can estimate the mean group size by solving a numerically and then substituting in to $\mathbf{E}(n) = a/[(a-1)\ln(1-a)]$. The comparison of simulation and this solution for N_P and that of the mean group size for different splitting rates p are shown in figure 4. Once again, the logarithmic distribution gives reasonably good estimations for both N_P and the mean group size. To summarize, while equation (18) is more complicated than equation (17) it has the advantage that it does not require an additional fitted parameter.

To investigate how system parameters affect N_P, we rewrite equation (18) as

$$\frac{q\Phi}{sp} = \frac{N_P}{\ln N_P} - \frac{(N_P - 1)^2}{N_P(\ln N_P)^2} \tag{19}$$

When N_P is relatively large, i.e. $N_P - 1 \approx N_P$, the right side of the above equation approximately equals to

$$G(N_P) = \frac{N_P}{\ln N_P} - \frac{N_P}{(\ln N_P)^2} \tag{20}$$

When N_P is large, over any particular order of magnitude, N_P increases much faster than $\ln N_P$. Furthermore, although $(\ln N_P)^2$ changes faster than

$\ln N_P$, the second term in equation (20) is smaller than the first. As a result, $G(N_P)$ is roughly a linear function of N_P over reasonably small ranges. This property can be contrasted with the result showed in equation (17), namely the truncation size of the truncated power law distribution grows roughly linearly with the total population when p, q and s are constant (figure 5). Our results supports this suggestion over small ranges. However, over larger ranges, i.e. several orders of magnitude, $G(N_P)$ is not exactly a linear function.

4. Other derivations of the logarithmic and related distributions

There are alternative ways of deriving the logarithmic distribution (equation 5) from first principles. The first point to note is that the logarithmic distribution is a limiting case of the negative binomial distribution (Fisher et al., 1943). If an observed variable follows the negative binomial distribution

$$f(k,p) = \frac{(k+n-1)!}{(k-1)!n!} p^n (1-p)^k \tag{21}$$

then taking the limiting case $k = 0$, and assuming that group sizes are never zero, we recover equation (5) with $a = p$. The negative binomial distribution has been used for modeling animal group sizes (Caraco et al., 1980; Cohen, 1972), but in general the focus has been on parameter values between $k = 1$, which corresponds to the geometric distribution and $k \to \infty$ which corresponds to the Poisson distribution. Thus truncated power laws, of which the logarithmic distribution is an example, can be put in the same framework as other models of animal group size distributions.

While the merge-split model above is not exactly consistent with logarithmic distribution, it is possible to 'backwards engineer' such a merge and split model. Note that from equation (5) we require

$$\frac{W(n)}{W(n-1)} = \frac{a(n-1)}{n} \tag{22}$$

We now discuss an evolution equation which fulfills this property.

The main idea in creating such a distribution is the use of preferential attachment (Yule, 1925; Barabasi et al., 1999; Newman, 2001) with splitting. In particular, larger groups have both higher merge and split rates. Assume

as before a system with conserved population Φ. We now insist that groups of size 1, decide to merge with other groups with rate q, and choose the group to merge with by preferential attachment. So that the probability it joins a group of size n is

$$\frac{nf(n,t)}{\Phi}$$

Groups with size $n > 1$ never merge with other groups (except those consisting of single individuals) but split with rate pn, i.e. larger groups split more often. When these groups split they do so in to n groups of size 1, so that all individuals in the group are then on their own. This mechanism gives,

$$\frac{\partial f(n,t)}{\partial t} = -q\frac{nf(n,t)}{\Phi}f(1,t) + q\frac{(n-1)f(n-1,t)}{\Phi}f(1,t) - pnf(n,t) \quad (23)$$

for $n = 2, 3, ...,$ and

$$\frac{\partial f(1,t)}{\partial t} = -\sum_{n=2}^{\infty} q\frac{nf(n,t)}{\Phi}f(1,t) - 2q\frac{(f(1,t)-1)}{\Phi}f(1,t) + \sum_{n=2}^{\infty} pn^2 f(n,t) \quad (24)$$

At equilibrium, equation (23) equals to 0, we obtain

$$\frac{f(n)}{f(n-1)} = \frac{(n-1)}{n}\frac{1}{1+\frac{p\Phi}{qf(1)}} \quad (25)$$

Since $f(n,t) = Z(t)W(n,t)$, and in equilibrium $Z = \sum f(n)$ is constant, the above equation is identical to equation (22) with

$$a = \frac{1}{1+\frac{p\Phi}{qf(1)}} \quad (26)$$

This derivation gives a direct link between logarithmic distribution and a simple preferential attachment mechanism. Finally, we identify

$$N_P = 1 + \frac{qf(1)}{p\Phi} \quad (27)$$

as the average group size experienced by an individual. Here $f(1)$ is a constant determined by system parameters. To show this, we rewrite equation (23) in equilibrium as

$$nf(n) = \frac{q(n-1)f(n-1)f(1)}{qf(1)+q\Phi} \quad (28)$$

Substituting the above equation into conservation condition $\Phi = \sum_{n=1}^{\infty} nf(n) = f(1) + \sum_{n=2}^{\infty} nf(n)$, we can get

$$f(1)^2 + \frac{p\Phi}{q} f(1) - \frac{p}{q}\Phi^2 = 0 \tag{29}$$

which leads to

$$f(1) = \frac{\Phi}{2}(\sqrt{\frac{p^2}{q^2} + \frac{4p}{q}} - \frac{p}{q}) \tag{30}$$

and

$$N_P = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{q}{p}} \tag{31}$$

We thus find N_P is specified only by the ratio of system parameters p and q. As long as the total population Φ is large enough, N_P is independent of Φ. This property distinguishes our preferential attachment model from the merge and split model in the previous section, where equation (19) shows a clear relationship between population density and N_P.

Logarithmic series can be categorized to a truncated power law with power exponent -1. For more general truncated power laws with other exponents, Gueron et al. (1995) gave detailed derivation in the CFP framework with some special forms of merging and splitting functions. They solve equation (4) analytically for three cases with the splitting and merging rate in the following form: the rate that a group of size i and a group of size $n-i$ merge to become a group of size n is $\psi(i, n-i) = \alpha a(i)a(n-i)$, the rate of a n-sized group splitting to a group of size i and $n-i$ is $\varphi(n, i) = 2\beta a(n)$, where $a(n)$ is some nonnegative function, α and β are positive constants and $\alpha/\beta < 2$. The three cases they solved are where: (1) $a(n) = 1$; (2) $a(n) = n$; (3) $a(n) = 1/n$. Via Laplace transform, Gueron and Levin derive the solution of (4) as:

$$f(x) = 2\frac{\beta}{\alpha}\left(\frac{1}{a(n)}\right)\exp(-\lambda n) \tag{32}$$

on the condition

$$\lim_{x\to 0} \frac{a(n)}{n^2} > 0$$

where λ is a constant determined by population size. We can now see that the logarithmic distribution is actually a specific example of equation (32), where $a(n) = n$. Gueron and Levin's distribution gives a very general and

natural way of producing truncated power laws. For example $a(n) = n^2$ gives group size distribution

$$f(n) = \frac{2\beta}{\alpha}\frac{1}{n^2}e^{-\lambda n}$$

The Gueron & Levin approach requires that the CFP is reversible for existence and uniqueness of the stationary distribution. CFP reversibility results from their assumption that $\psi(i,j) = b(i)b(j)$ and $\varphi(i,j) = b(i+j)$. Durrett et al. (1999) generalized the condition on the merging and splitting rate to their ratio satisfying the following form:

$$q(i,j) = \frac{\psi(i,j)}{\varphi(i,j)} = \frac{a(i+j)}{a(i)a(j)}$$

without losing the reversibility of the CFP for some positive function $a(i)$. However, this provides only a small number of merge-split models which are reversible and none of the models we have discussed in this paper satisfy this property. Our preferential attachment model has merging rate

$$\psi(i, n-i) = \begin{cases} \frac{q}{\Phi} i(n-i) & \text{i=1} \\ 0 & \text{otherwise} \end{cases}$$

and splitting rate

$$\varphi(n,i) = pn.$$

Still, it is possible that there are a whole range of models which fit the evolution equation (32) and produce distributions similar or identical to truncated power laws. Our numerical investigations suggest that CFP can produce stable distributions even when reversibility is not satisfied.

5. Discussion

A large number of recent theoretical studies have derived power law distributions from assumptions about how individuals interact (Newman, 2001, 2005; Sornette, 2004; Ma et al., 2008; Takayasu et al., 1988; Barabasi et al., 1999). On the whole these studies are concerned with producing distributions with a power law tail rather than, as we have done here, simultaneously deriving both the power law and an exponential decay above some point of truncation. The exponential decay is often viewed as an inconvenient necessity, required when fitting models to data because all natural systems

have some physical limit to their size. In many data sets involving power laws, however, the exponential decay should have a genuine physical, biological or sociological interpretation (Newman, 2001). For example, Kéfi et al. (2007) found that vegetation patch sizes are power law distributed with an exponential decay which reflects grazing pressure. The animal group size distributions, we have focussed on here, are another such example, where there is a clear truncation point in the distribution when groups become large (Bonabeau et al., 1999; Niwa, 2003; Griesser et al., 2011).

The fact that we can link the merging and splitting of groups with the logarithmic distribution means that it has a number of useful properties. Indeed, it is the average group size experienced by an individual, N_P, which is central to the entire model. Some previous models (Bonabeau et al., 1995, 1999; Sjöberg et al., 2000) suggested animal group size distributions should be fitted by truncated power laws but these failed to explain where the truncation should occur or the behavioral background of the truncation size. Both the merge and split model and the preferential attachment model we discuss in this paper can be used to calculate N_P directly (through equations 18 and 31, respectively). The truncation parameter a can then be calculated from N_P.

We have shown that the logarithmic distribution can be derived from mechanisms based on joining through preferential attachment, with splitting rate increasing with group size. Furthermore, the merge and split model in section 2 gives approximately the logarithmic distribution. The fact that data generated from this model is well fit by equation (5), raises the possibility that there are a whole range of site-based mechanisms that produce something close to the logarithmic distribution. This may explain why the logarithmic distribution is widespread in natural populations (Niwa, 2003; Griesser et al., 2011).

While the merge & split model and preferential attachment model provide different derivations of the same distribution, these models produce completely different scaling relations between N_P and population size. In the merge & split model, N_p depends on the total population size Φ roughly linearly over small ranges when other system parameters are fixed. In the preferential attachment model, N_P is independent with the total population. This property may apply to model selection for empirical data. If the experienced group size grows roughly linearly with the system size when environmental factors do not change, probably groups in the system follow the merge and split mechanism. Otherwise if the experienced group size does

not change with the total population, there is good chance that individuals chose the group to join by preferential attachment.

The research presented here has built upon two results provided by Gueron and co-workers (Gueron, 1998; Durrett et al., 1999; Gueron et al., 1995). Firstly, Gueron & Levin provided a clear framework to describe general CFPs, i.e. equation (4). We have used a discretized version of this framework to describe our models. Secondly, for symmetrical splitting rate and merging rate Gueron & Levin generalized logarithmic distribution to other truncated power laws.

The logarithmic distribution is a limiting case of two important distributions, truncated power laws (Bonabeau et al., 1995, 1999; Sjöberg et al., 2000) and the negative binomial distribution (Caraco et al., 1980; Cohen, 1972; Okubo, 1986). In this sense, there is a stronger link between the different types of distributions used to model animal groups than that has previously been suggested (Gerard et al., 2002; Bonabeau et al., 1999). We maintain that the logarithmic distribution should be the first model of choice in fitting experimental data from fission-fusion groups, precisely because it can be related to both merge/split and preferential attachment processes. The Gueron-Levin distribution (equation 32) or the negative binomial distribution (equation 21) then provide two useful generalizations.

Acknowledgements: *We thank Michael Griesser for providing the empirical data that drove us to understand more about animal group size models. This research was funded by the European Research Council grant (IDCAB) and the Chinese Scholarship Council.*

References

Foster, W. A. & Treherne J. E. 1981 Evidence for the dilution effect in the selfish herd from fish predation on a marine insect. *Nature* 293, 466-467.

Parrish, J. K. 1989 Layering with depth in a heterospecific fish aggregation. *Environ. Biol. Fishes* 26(2), 79-85.

Barbraud, C. & Weimerskirch, H. 2001 Emperor penguins and climate change. *Nature* 411, 183-186.

Sumpter, D. J.T. 2010 Collective Animal Behavior. Princeton University Press.

Gerard, J. F., Bideau, E., Maublanc, M. L., Loisel, P. & Marchal, C. 2002 Herd Size in Large Herbivores: Encoded in the Individual or Emergent? *Biol. Bull.* 202, 275-282.

Beauchamp, G. & Fernández-Juricic, E. 2005 The group-size paradox: effects of learning and patch departure rules. *Behav. Ecol.* 16(2), 352-357.

Clark, C. W. & Mangel, M. 1986 The evolutionary advantages of group foraging. *Theoret. Pop. Biol.* 30(1), 45-75.

Sibly R.M., 1983 Optimal group size is unstable. *Anim. Behav.* 31, 947-948.

Mottley, K. & Giraldeau, L. A. 2000 Experimental evidence that group foragers can converge on predicted producerscrounger equilibria. *Anim. Behav.* 60, 341-350.

Caraco, T., Martindale, S. & Pulliam, H. R. 1980 Avian flocking in the presence of a predator. *Nature* 285, 400-401.

Okubo, A. 1986 Dynamical aspects of animal grouping: Swarms, schools, flocks and herds. *Adv. Biophys.* 22, 1-94.

Bonabeau, E. & Dagorn, L. 1995 Possible universality in the size distribution of fish schools. *Phys. Rev. E* 51, 5220-5223.

Bonabeau, E., Dagorn, L., & Freon, P. 1999 Scaling in animal group-size distributions. *Proc. Natl. Acad. Sci. USA* 96, 4472-4477.

Takayasu, H., Nishikawa, I. & Tasaki, H. 1988 Power-law mass distribution of aggregation systems with injection. *Phys. Rev. A* 37, 3110-3117.

Sjöberg, M., Albrectsen, B. & Hjältén, J. 2000 Truncated power laws: a tool for understanding aggregation patterns in animals? *Ecol. Lett.* Q, 90-94.

Gillespie, D. T. 1977 Exact Stochastic Simulation of Coupled Chemical Reactions. *J. Phys. Chem.* 81(25), 2340-236.

Gueron, S. 1998 The steady-state distributions of coagulation-fragmentation processes. *J. Math. Biol.* 37, 1-27.

Durrett, R. , Granovsky, B. L. & Gueron, S. 1999 The Equilibrium Behavior of Reversible Coagulation-Fragmentation Processes. *J. Theoret. Probab.* 12(2), 447-474.

Niwa, H. S. 2003 Power-law versus exponential distributions of animal group sizes. *J. theor. Biol* 224, 451-457.

Niwa, H. S. 2004 Space-irrelevant scaling law for fish school sizes. *J. theor. Biol* 228, 347-357.

Griesser, M., Ma, Q., Webber, S., Bowgen, K. & Sumpter, D. J. T. 2011 Understanding animal group-size distributions, in submission.

Gueron, S. & Levin, S. A. 1995 The Dynamics of Group Formation. *Math. Biosci.* 128, 243-264.

Fisher, R. A., Corbet A. S. & Willams, C. B. 1943 The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* 12, 42-58.

Cohen, J. E. 1972 Markov Population Processes as Models of Primate Social and Population Dynamics. *Theoret. Pop. Biol.* 3, 119-134.

Yule, G. U. 1925 A Mathematical Theory of Evolution, based on the Conclusions of Dr. J. C. Willis, F.R.S.. *Philos. Trans. R. Soc. London Ser. B* 213, 21-87.

Patil, G. P. & Rao, C. R. 1978 Weighted Distributions and Size-Biased Sampling with Applications to Wildlife Populations and Human Families. *Biometrics* 34(2), 179-189.

Barabasi, A. L. & Albert, R. 1999 Emergence of scaling in random networks. *Science* 286, 509-512.

Newman, M. E. J. 2001 Clustering and preferential attachment in growing networks. *Phys. Rev. E* 64, 025102.

Newman, M. E. J. 2001 The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. USA* 98(2), 404-409.

Newman, M. E. J. 2005 Power laws, Pareto distributions and Zipfs law. *Contemp. Phys.* 46(5), 323-351.

Sornette, D. 2004 Critical Phenomena in Natural Sciences, Springer, Heidelberg.

Ma, Q., Chen, Y. W., Tong, H. & Di, Z. R. 2008 Production, depreciation and the size distribution of firms. *Physica A* 387, 3209-3217.

Kéfi, S., Rietkerk, M., Alados, C. L., Pueyo, Y., Papanastasis, V. P., ElAich, A. & Ruiter, P. C. 2007 Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. *Nature* 449, 213-217.

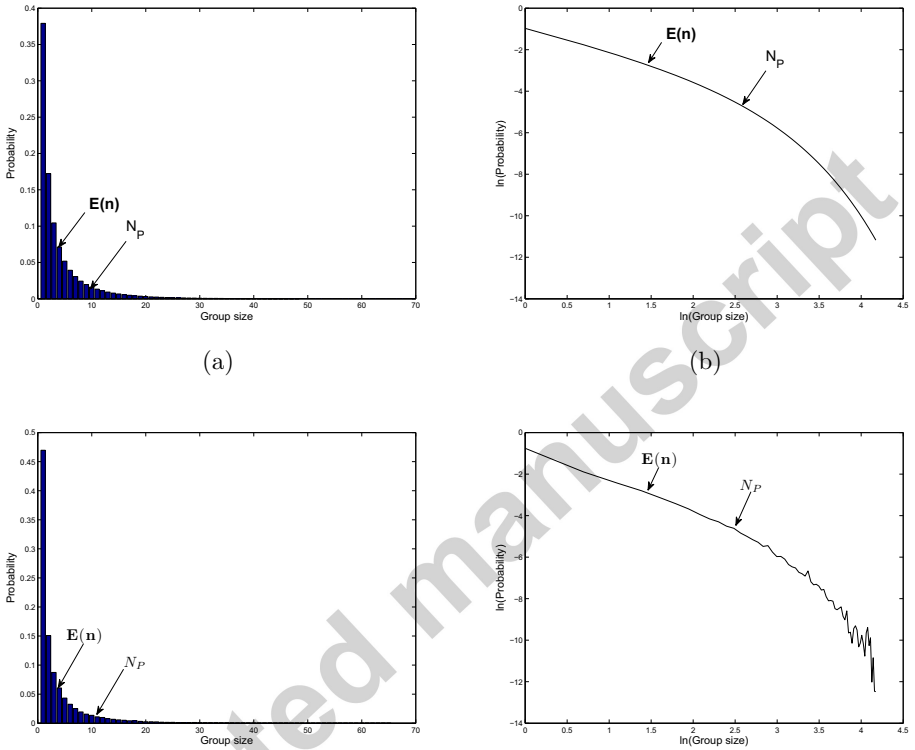
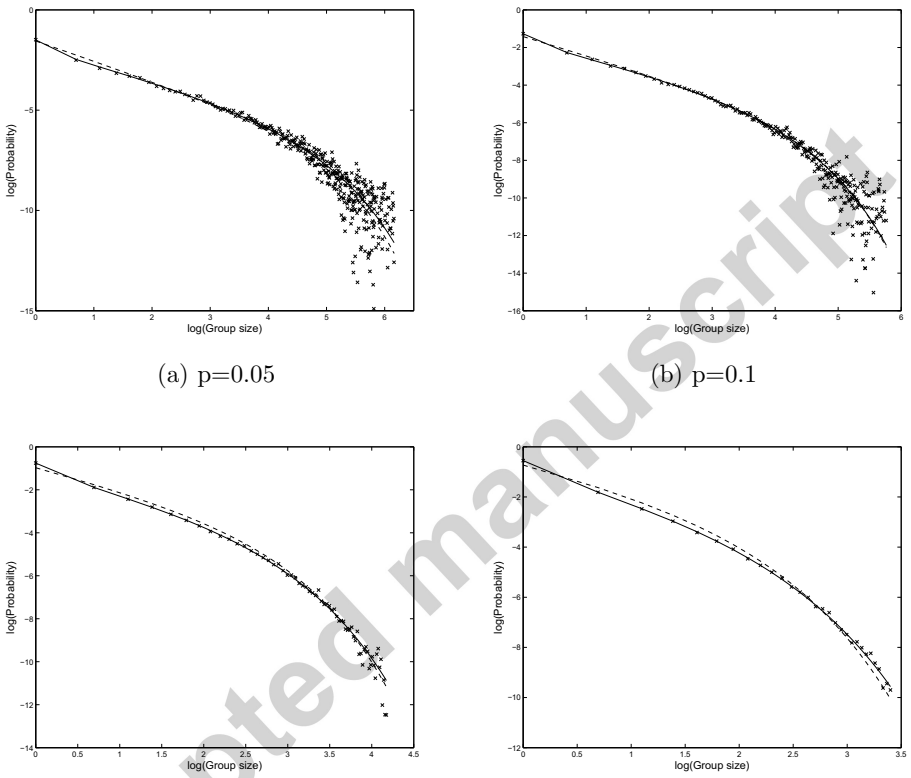


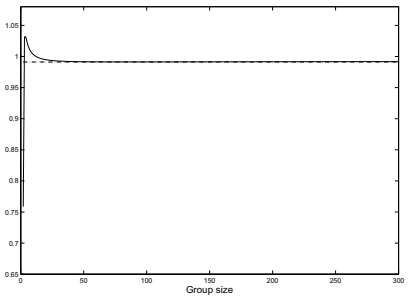
Figure 1: (a) Histogram of group size distribution in the form of equation (8), with N_P set equal to that estimated from the simulation in (c), i.e. $N_P = 11$. (b) Same data as in (a), plotted in log-log scale. (c) Histogram of the stable group size distribution by simulation of the merge and split model. Total population $\Phi = 10000$, splitting rate $p = 0.5$, moving rate $q = 1$, total number of sites $s = 10000$, simulation time step T=400000. Every group has size 10 for the initial state. (d) Same data as in (c), plotted in log-log scale;



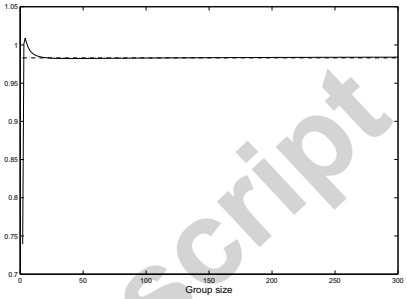
(a) p=0.05 (b) p=0.1

(c) p=0.5 (d) p=1

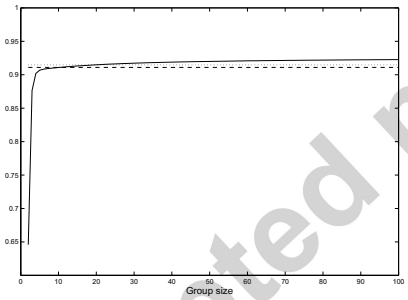
Figure 2: Comparison of simulated group size distribution (the crosses, with total population $\Phi = 10000$, moving rate $q = 1$, total number of sites $s = 10000$, simulation time step T=400000. Every group has size 10 for the initial state), evolution equation iteration (the solid line), and Niwa's distribution as in equation(5) (the dashed line).



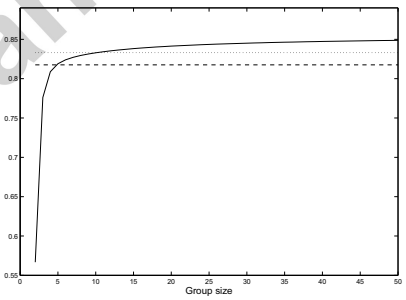
(a) p=0.05



(b) p=0.1

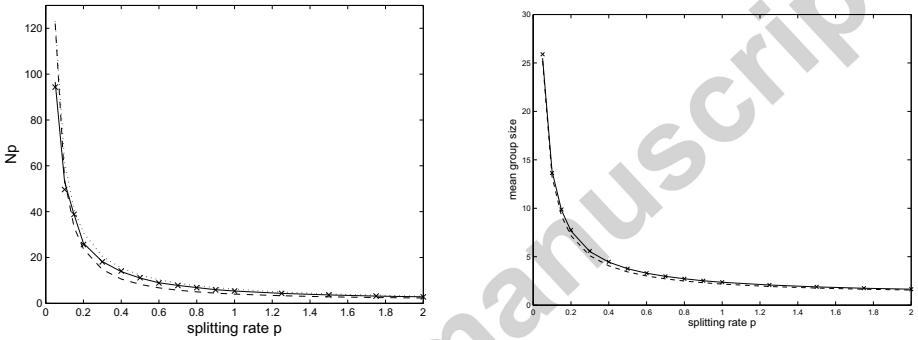


(c) p=0.5



(d) p=1

Figure 3: Comparison of $nW(n)/(n-1)W(n-1)$ from numerical solution (the solid line) with $a = \exp(-1/N_P)$ (the dotted line) and $a = 1 - 1/N_P$ (the dashed line). N_P used here is estimated from the numerical solution. $\Phi = 10000$, s=10000, $q = 1$.



(a) N_P vs splitting rate (b) Mean group size vs splitting rate

Figure 4: N_P and mean group size versus splitting rate p. $\Phi = 10000$, s=10000, $q = 1$. The crosses denote simulation results, solid line is the corresponding value derived numerically from the evolution equations. Dashed line is calculated by using logarithmic distribution with $W(1) = Ka$, and dotted line in (a) is the estimation proposed by Niwa in the form of equation (17).

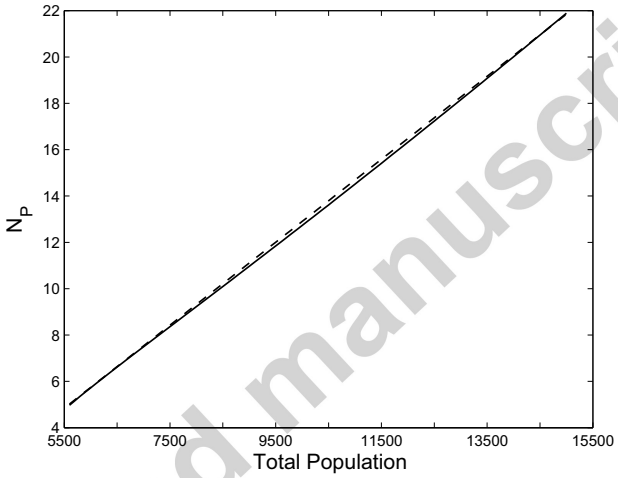


Figure 5: Scaling relation between N_P and total population size. The solid line is total population Φ versus the corresponding solution N_P of equation (18) when $p = 0.2$, $q = 1$ and $s = 15000$. The dashed line is the linear fitting of the solid line.

Research Highlights:

- First principles derivations of the logarithmic distribution and other truncated power laws using a site-based merge and split framework.
- Link the truncation size to the merging and splitting behaviors of animal groups.
- The logarithmic distribution and truncated power laws can be obtained by preferential attachment mechanisms.
- The logarithmic distribution is a limiting case of truncated power law and negative binomial distribution.