Coordination number statistics of cluster formation

COLIN P. REYNOLDS¹, DAVID ROBINSON¹, DIRK G. A. L. ARTS¹, MARK WILSON¹, WILLIAM W. SAMPSON² and ROEL P. A. DULLENS¹

¹ Department of Chemistry, Physical and Theoretical Chemistry Laboratory - South Parks Road, Oxford, OX1 3QZ, UK
² School of Materials, University of Manchester - Manchester, M13 9PL, UK

received 24 June 2016; accepted in final form 3 November 2016
published online 23 November 2016

PACS 82.70.Dd – Colloids
PACS 83.80.Gv – Electro- and magnetorheological fluids
PACS 02.50.-r – Probability theory, stochastic processes, and statistics

Abstract – We establish a direct relation between the distribution of coordination numbers for individual particles and the mean cluster size using simple statistical considerations. We test our probabilistic theory against an experimental model system of superparamagnetic colloidal particles, which form chain-like clusters upon the application of an external magnetic field. We find that the experimental cluster size distribution is well described by a shifted geometric distribution, consistent with a Smoluchowski aggregation scheme. Computer simulations of an atomistic model show that ordering of the positions of the particles prior to aggregation alters the cluster size distribution. These simulations corroborate the validity of our coordination number statistics approach to aggregation processes and demonstrate that the cluster formation process can be interpreted as a compound Poisson process.

Introduction. – Self-assembly is an umbrella term for a fascinating range of processes by which initial components build into a complex structure via a chemical or physical change. From complex inorganic macromolecules [1] to the tertiary structure of polypeptides [2], self-assembled materials represent an extremely wide variety of systems [3,4]. However, controlling self-assembly remains a formidable challenge, not least because there are usually many components present which lead to a range of different interactions and competing driving forces [5]. To exploit self-assembled structures for the development of useful materials and tunable devices, it is a prerequisite to have a deep understanding of the fundamental structural and dynamical processes underlying the self-assembly.

The coordination number of a constituent entity is a central quantity in self-assembling systems as this characterises the local structure around that entity [6–12]. In granular packings, the importance of the coordination number on the rigidity of the structure has also long been recognised [13]. In molecular systems, the local structure is strongly affected by the geometry of molecules and their directional interactions result in highly specific self-assembled structures [14]. Accordingly, controlling the structure in self-assembling mesoscopic systems, such as nanoparticles and colloids, requires the ability to control the coordination number, resulting in considerable effort in developing directional interactions in these systems [15–22]. However, somewhat surprisingly, the statistics characterising the coordination number distribution, its evolution during self-assembly and its relation to the final structure has never been addressed to our knowledge.

In this letter, we use simple statistical considerations to provide a full prediction of the distribution of coordination numbers for individual particles in a chain-forming system. Our theory is tested against experiments in which we follow the chain formation process in a system of superparamagnetic colloidal particles upon the application of an external magnetic field [23–26], an excellent model system to establish the relation between cluster formation and coordination number statistics. The validity and generality of our approach is further confirmed using atomistic computer simulations in which the starting configuration of a chain-forming system is systematically varied from random to ordered. Ordering the starting positions of
the particles results in a distinct cluster size distribution indicating that chain growth is a compound Poisson process. In principle our probabilistic approach can be applied to more complex self-assembling systems [27–29], providing a general perspective on self-assembly based on local geometry as characterised by the coordination number.

Theory. – The growth of clusters results in an evolving distribution of cluster sizes characterised by a discrete probability distribution, \( P(s) \). From this distribution, coupled with knowledge of the structures of the clusters, we may elaborate the distribution of coordination numbers of the particles as a function of time. Here, we analyse one of the simplest cluster-forming systems: chain formation of superparamagnetic spheres in an external magnetic field [23–26].

As a starting point for our theory we need a functional form for the cluster size distribution, \( P(s) \), where \( s \) is the chain length defined for the set \( \{ s: s \in \mathbb{N}^+ \} \) as the smallest possible cluster is a single particle. We will show experimentally that \( P(s) \) for our chain forming system is accurately described by a shifted geometric distribution. As such, we will develop the theory specifically for this distribution. It is, however, important to note that the concept is general and \( P(s) \) may be replaced with an alternative distribution appropriate to another growth process. We also note that the shifted geometric distribution is known to result from a Smoluchowski aggregation scheme [30].

As such, the shifted geometric distribution is a probable candidate distribution for a range of clustering processes. The discrete probability distribution of a shifted geometric is given by

\[
P(s) = \frac{1}{\mu} \left( 1 - \frac{1}{\mu} \right)^{s-1},
\]

such that \( P(s) \) gives the probability of choosing a chain of size \( s \) when randomly selecting a cluster and \( \sum_{s=1}^{\infty} P(s) = 1 \). Here, \( \mu \) represents the mean of the distribution, \( \mu = \sum_{s=1}^{\infty} s P(s) \).

The probability that a particle \( i \) has a coordination number \( c \) is conditional on \( s \), the size of the cluster that the particle belongs to. We denote this probability \( P(c|i; s) \). The probability that a particle \( i \), chosen at random, is in a chain of size \( s \) is denoted \( P(i; s) \). The product of the two, \( P(c \cap i; s) = P(c|i; s)P(i; s) \), is the probability of randomly selecting a particle that is in a cluster of size \( s \) and has a coordination number \( c \). It follows that the probability that a particle, chosen at random, has a coordination \( c \) is given by the sum over all cluster sizes of this probability:

\[
P_c = \sum_{s=1}^{\infty} P(c \cap i; s) = \sum_{s=1}^{\infty} P(c|i; s)P(i; s).
\]

The probability of finding a particle in a cluster of a particular size, \( P(i; s) \), may be elaborated from the probability distribution, \( P(s) \), as

\[
P(i; s) = \frac{s P(s)}{\sum_{s=1}^{\infty} s P(s)} = \frac{s}{\mu^2} \left( 1 - \frac{1}{\mu} \right)^{s-1},
\]

which is a normalised probability distribution, \( \sum_{s=1}^{\infty} P(i; s) = 1 \). Specifically in the case of linear chains, where only zero coordinate, one coordinate, and two coordinate particles exist, for \( s \geq 2 \), \( P(c|i; s) \) may be described in terms of the cluster sizes as \( P(0|i; s) = 0 \), \( P(1|i; s) = 2/s \) and \( P(2|i; s) = (s-2)/s \). Single particles are isolated and have a coordination number of 0. In other words, for \( s = 1 \), \( P(0|i; s) = 1 \), \( P(1|i; s) = 0 \) and \( P(2|i; s) = 0 \). For larger chains the particles corresponding to the termini of the chains each have one neighbour and the other particles, forming the body of the chain, have two neighbours. Note that this model neglects the possibility of cruciform defects [31], laterally aggregated chains, and other deviations from simple chains.

The probability that a particle has coordination number \( c \), \( P_c \), is now easily found. The fraction of zero coordinate particles is simply the probability of isolated particles:

\[
P_0 = P(i; s = 1) = \frac{1}{\mu^2}.
\]

The fractions of particles with coordination numbers 1 and 2 follow from eq. (2):

\[
P_1 = \sum_{s=2}^{\infty} \frac{2 P(i; s)}{s} = \frac{2(\mu-1)}{\mu^2},
\]

and

\[
P_2 = \sum_{s=2}^{\infty} \frac{(s-2) P(i; s)}{s} = \frac{(\mu-1)^2}{\mu^2},
\]

such that \( \sum_{c=0}^{\infty} P_c = 1 \). From eqs. (4)–(6) we obtain the mean and variance of the coordination number as

\[
\bar{c} = \frac{2(\mu-1)}{\mu},
\]

and

\[
\sigma_c^2 = \frac{\bar{c}}{\mu},
\]

respectively. Note that \( \lim_{\mu \rightarrow 1} \bar{c} = 0 \), \( \lim_{\mu \rightarrow 1} \sigma_c^2 = 0 \), \( \lim_{\mu \rightarrow 0} \bar{c} = 2 \) and \( \lim_{\mu \rightarrow 0} \sigma_c^2 = 0 \), as required. Importantly, eq. (7) tells us that there is a one-to-one relationship between the mean coordination number and the mean chain length such that knowledge of either allows the full statistics of the coordination number distribution to be calculated.

Experimental and simulation details. – We compare our theory with experiments carried out using monodisperse superparamagnetic spheres with a diameter of 3.0 μm (Dynabeads®, Life Technologies). Dilute suspensions with packing fractions \( \phi < 0.05 \) are confined by gravity to a monolayer at the bottom of a quartz
Results and discussion. – From the image analysis of snapshots, fig. 1(a)–(c), we obtain the chain lengths and the corresponding chain length distribution at all times. For a geometric distribution, eq. (1), plotting $P(s/\mu)$ against $s/\mu$ reduces the distributions at different times to approximate the probability density function of the exponential distribution—the continuous analogue of the geometric distribution. We indeed find very good agreement between our experimental data, which is averaged over 25 different runs, and the exponential distribution at all times, as shown in fig. 1(d). Low values of $P(s/\mu)$ are only sampled by larger cluster sizes. Due to the limited field of view, and the observation time we only probe a range of mean cluster sizes. As such, the exponential distribution slightly overestimates the observed probability of low values of $s/\mu$. The inset in fig. 1(d) shows the variance of the chain length against the mean chain length, which agrees very well with the prediction that follows from eq. (1): $\sigma_s = \mu \sqrt{1 - 1/\mu}$, again corroborating the applicability of the geometric distribution.

Next, we determine the fraction of particles with specific coordination numbers for our experiments and compare these with the theoretical predictions, eqs. (4)–(6). Figure 2(a) shows the fraction of particles with $c = 0, 1, 2$ and $c > 2$ as a function of the mean chain length $\mu$. The agreement between the experiments and the predictions from eqs. (4)–(6) is excellent for all mean cluster sizes. Only for larger mean cluster sizes, corresponding to later time, a very small deviation becomes apparent. This is due to the presence of a small fraction of particles with $c > 2$, which is not described in terms of simple chain formation. This assertion is confirmed by plotting the
Fig. 2: (Colour online) (a) The fraction of particles with coordination number $c$, $P_c$, for $c = 0, 1, 2$ and $c > 2$ as a function of the mean chain length $\mu$. Closed symbols correspond to experiments; solid lines represent eqs. (4)–(6). (b) Variance of the coordination numbers $\sigma^2_c$ plotted against the mean coordination number $\bar{c}$. Black symbols correspond to experiments and grey ones to simulations; the solid line is the prediction from eqs. (7) and (8):

$$\sigma^2_c = \bar{c} - \bar{c}^2/2.$$

The variance of the coordination number $\sigma^2_c$ against the mean coordination number $\bar{c}$—but now neglecting the particles with $c > 2$—alongside the theoretical prediction following from eqs. (7) and (8): $\sigma^2_c = \bar{c} - \bar{c}^2/2$ (see fig. 2(b)). Excellent agreement between experiment and theory is observed, especially given the fact that there are no fitting parameters.

Our computer simulation of the chain growth process starting from a random starting configuration shows the same coordination number statistics, as confirmed by plotting $\sigma^2_c(c)$ as a function of $\bar{c}$ in fig. 2(b). Importantly, this shows that the coordination number statistics hold for both the colloidal (experimental) and atomistic (simulation) systems, despite the fact that the overdamped dynamics exhibited by the experimental system are very different from those of the simulation, where the presence of momentum leads to a qualitatively different particle dynamics.

The shifted geometric distribution, eq. (1), arises from simple Smoluchowski aggregation schemes [30], though it is known that at higher concentrations the geometric distribution is not predicted [34–36]. Qualitatively, it is informative to also consider the aggregation mechanism as a compound Poisson process, from which a geometric distribution is the simplest distribution that can arise. Due to the initial random spatial arrangement of a sparse monolayer of particles, the application of a magnetic field leads to a Poisson distribution of cluster lengths as particles begin to associate with other particles in their local environment. Subsequently, the chains, randomly distributed in the plane, aggregate due to the interactions of chain ends [37], which represents a second independent Poisson process. Hence, in this scheme the cluster length results from a compound Poisson process. While this model is not an exact description of our experiments, as it implies a clear separation in time of the initial and subsequent aggregation events, and treats all subsequent aggregation events as the same, the compound Poisson interpretation highlights a key concept: the length of a chain is derived from a random number of aggregation events of clusters of random length, which is entwined with the initial spatial distribution of particles.

The simulations allow us to reveal how the chain growth, and hence the resulting coordination number statistics, is affected by the initial particle configuration. The solid line in fig. 3 shows the probability of coordination number 1 as given by eq. (5) along with simulation data associated with different degrees of ordering of the initial particle positions. Clearly, larger deviations from eq. (5) are seen for increasingly ordered starting configurations. Particles starting from a perfectly ordered lattice have a tendency to “pair up” and the doublets that result still have a large degree of ordering and can thus pair up again. This explains both the observed excess of $P_1$ around $\mu = 2$, and its subsequent deficit for $\mu > 2.5$, relative to the prediction using the shifted geometric distribution, eq. (5). The difference
We illustrate this in fig. 4(a), where we show $P$ indicated by the dashed line in panel (a). Note that $s$ponding to these three experimental systems at $t = 500$ s, as indicated by the dashed line in panel (a).

is a consequence of a more uniform chain length distribution that arises when the assumption in the compound Poisson process of independent initial particle positions is violated. We note that already a small amount of disorder in the initial particle positions significantly interferes with the pairing process observed for perfect lattice and hence randomises the configuration very rapidly, leading to a distribution of chain lengths. As a result, the compound Poisson process is recovered very rapidly for increasing amounts of disorder in the starting configuration and as the mean chain length increases.

Finally, we show how the coordination number statistics can be probed in the time domain, which is the natural “reaction coordinate” when studying cluster formation. By using the time dependence of the mean cluster size

$$\mu = a + bt^z,$$

with $a = 1.12$, $b = 3.1 \cdot 10^4$, $z = 0.61$ and $t^* = \phi^{2.44}B^{1.61}t$ [32], we can directly convert the dependence of $P_c$ on $\mu$ into a time dependence for any $\phi$ and $B$. We illustrate this in fig. 4(a), where we show $P_0$, $P_1$ and $P_2$ as a function of time for three different combinations of $\phi$ and $B$ alongside the theoretical prediction.

The corresponding snapshots at $t = 500$ s are shown in fig. 4(b)–(d). As expected the sample with the highest packing fraction ($\phi = 0.024$) and at the highest magnetic field ($B = 1.88$ mT, $\lambda = 350$) exhibits the fastest aggregation, while the sample at $\phi = 0.013$ and $B = 0.63$ mT ($\lambda = 40$) shows the slowest chain formation. The sample at $\phi = 0.013$ and $B = 1.88$ mT is an intermediate case with the same packing fraction as the slowest case and the same field as the fastest case. This shows that increasing either $\phi$ or $B$ independently has the expected effect. Excellent agreement between experiments and theory is observed, which implies that our statistical description holds, regardless of whether the packing fraction or the magnetic field is varied.

**Conclusions.** – We have developed a direct relation between the coordination number statistics and the mean cluster size using simple statistical considerations. We confirm our model in a chain forming experimental model system of super-paramagnetic colloidal particles in an external magnetic field. Our approach has been confirmed by atomistic computer simulations by systematically varying the starting configuration, thereby affecting the nature of the chain formation process. We have observed that the cluster size distribution is well described by a shifted geometric distribution and interpret the cluster formation as a compound Poisson process from an initial configuration of randomly distributed particles. Our results imply that the coordination number distribution fully characterises the cluster size distribution and vice versa at all times in the case of geometrically distributed linear chains. Our probabilistic approach of analysing cluster formation through coordination number statistics is applicable to more complex systems and thus provides a general geometric perspective on self-assembly.

***

The authors thank the EPSRC (EP/J001902/1 and EP/J00331X/1) for financial support.

REFERENCES

[1] Wilson M., Kumar A., Sherrington D. and Thorpe M. F., Phys. Rev. B, 87 (2013) 214108.
[2] Gething M.-J. and Sambrook J., Nature, 355 (1992) 33.
[3] Mann S., Nat. Mater., 8 (2009) 781.
[4] Philip D. and Stoddart J. F., Angew. Chem., Int. Ed., 35 (1996) 1154.
[5] Lehn J.-M., Science, 295 (2002) 2400.
[6] Arkus N., Manoharan V. N. and Brenner M. P., Phys. Rev. Lett., 103 (2009) 118303.
[7] Meng G., Arkus N., Brenner M. P. and Manoharan V. N., Science, 327 (2010) 560.
[8] Campbell A. I., Anderson V. J., van Duineveldt J. S. and Bartlett P., Phys. Rev. Lett., 94 (2005) 208301.
[9] van Slooeneveld M. M., de Villeneuve V. W. A., Dullens R. P. A., Aarts D. G. A. L., Leunissen M. E. and Kegel W. K., J. Phys. Chem. B, 113 (2009) 4560.

[10] Ruizcka B., Zaccarelli E., Zulian L., Angelini R., Sztkuci M., Moussaïd A., Narayanan T. and Sciortino F., Nat. Mater., 10 (2011) 56.

[11] Markova O., Albrecht J., Munro E. and Lenne P.-F., Phys. Rev. E, 90 (2014) 022301.

[12] Damasceno P. F., Engel M. and Glotzer S. C., Science, 337 (2012) 453.

[13] Liu A. J. and Nagel S. R., Nature, 396 (1998) 21.

[14] Lehn J.-M., Supramolecular Chemistry: Concepts and Perspectives (Wiley-VCH) 1995.

[15] Glotzer S. C. and Solomon M. J., Nat. Mater., 6 (2007) 557.

[16] Wang Y., Wang Y., Breed D. R., Manoharan V. N., Feng L., Hollingsworth A. D., Weck M. and Pine D. J., Nature, 491 (2012) 51.

[17] Wolters J. R., Avvisati G., Hagemans F., Vissers T., Kraft D. J., Dijkstra M. and Kegel W. K., Soft Matter, 11 (2015) 1067.

[18] Sabapathy M., Christdoss Pushpam S. D., Basavaraj M. G. and Mani E., Langmuir, 31 (2015) 1255.

[19] Kraft D. J., Groenewold J. and Kegel W. K., Soft Matter, 5 (2009) 3823.

[20] Kraft D. J., Ni R., Smallenburg F., Hermes M., Yoon K., Weitz D. A., van Blaaderen A., Groenewold J., Dijkstra M. and Kegel W. K., Proc. Natl. Acad. Sci. U.S.A., 109 (2012) 10787.

[21] Sacanna S., Korpics M., Rodriguez K., Colón-Meléndez L., Kim S.-H., Pine D. J. and Yi G.-R., Nat. Commun., 4 (2013) 1688.

[22] Chen Q., Bae S. C. and Granick S., Nature, 469 (2011) 381.

[23] De Gennes P. and Pincus P., Phys. Kondens. Mater., 11 (1970) 189.

[24] Helgesen G., Skjeltorp A., Mors P. M., Botet R. and Jullien R., Phys. Rev. Lett., 61 (1988) 1736.

[25] Skjeltorp A. T., J. Appl. Phys., 57 (1985) 3285.

[26] Promislov J. H. E., Gast A. P. and Fermigier M., J. Chem. Phys., 102 (1995) 5492.

[27] Beltran-Villegas D. J. and Bevan M. A., Soft Matter, 7 (2011) 3280.

[28] Manoharan V. N., Elssesser M. T. and Pine D. J., Science, 301 (2003) 483.

[29] Jiang S., Chen Q., Tripathy M., Luijten E., Schweizer K. S. and Granick S., Adv. Mater., 22 (2010) 1660.

[30] Sonntag H., Strenge K., Vincent B. and Vincent B., Coagulation Kinetics and Structure Formation (Springer) 1987.

[31] Fürst E. M. and Gast A. P., Phys. Rev. Lett., 82 (1999) 4130.

[32] Reynolds C. P., Klop K. E., Lavergne F. A., Morrow S. M., Aarts D. G. A. L. and Dullens R. P. A., J. Chem. Phys., 143 (2015) 214903.

[33] Fermigier M. and Gast A. P., J. Colloid Interface Sci., 154 (1992) 522.

[34] Bossis G., Iskakova L., Kostenko V. and Zubarev A., Physica A, 390 (2011) 2655.

[35] Bossis G., LANÇON P., Meunier A., Iskakova L., Kostenko V. and Zubarev A., Physica A, 392 (2013) 1567.

[36] Mimouni Z. and Wattis J., Physica A, 388 (2009) 1067.

[37] Fürst E. M. and Gast A. P., Phys. Rev. E, 62 (2000) 9616.