Analytic results on the polymerisation random graph model

Kryven, I.

CrossMark

Analytic results on the polymerisation random graph model

Ivan Kryven[1]

Abstract The step-growth polymerisation of a mixture of arbitrary-functional monomers is viewed as a time-continuos random graph process with degree bounds that are not necessarily the same for different vertices. The sequence of degree bounds acts as the only input parameter of the model. This parameter entirely defines the timing of the phase transition. Moreover, the size distribution of connected components features a rich temporal dynamics that includes: switching between exponential and algebraic asymptotes and acquiring oscillations. The results regarding the phase transition and the expected size of a connected component are obtained in a closed form. An exact expression for the size distribution is resolved up to the convolution power and is computable in subquadratic time. The theoretical results are illustrated on a few special cases, including a comparison with Monte Carlo simulations.

Keywords Random graph · Connected components · Polymerisation · Molecular network

Mathematics Subject Classification 05C80 · 82D60

1 Introduction

The chemical graph theory is the branch of mathematical chemistry that applies graph theory to mathematical modelling of chemical processes. This theory centres its attention on the concept of a molecular graph, which identifies atoms (or monomers) as vertices and chemical bonds as edges. This structure, finite or infinite, is usu-

---

✉ Ivan Kryven
  i.kryven@uva.nl

1   University of Amsterdam, Science Park 904, The Netherlands

ally defined a priori, e.g. molecular graphs describing structural isomers or Euclidian graphs describing crystal nets [5,27]. The graph-theoretical invariants of such chemical objects are known to be strongly correlated with physical properties of the resulting materials. These invariants include but are not restricted to: Wiener index, average shortest path, shape index, centric index, and connectivity index [7,21,22,26]. Not all molecular topologies can be described by a single graph, but rather by a probability measure over graphs [17,19]. This scope covers (hyper-)branched polymers, cross-linked polymers, molecular networks, and gels to name a few. A branch of graph theory that operates with probability distributions over graphs—random graph theory—has little documented applications to chemistry at present.

Consider a chemical system where each monomer has a predefined functionality, that is the maximum number of neighbours in the network. If the spatial positioning of the monomers is disregarded, the monomers can be represented as vertices in a graph model. From this perspective, the polymerisation process is a random graph process that respects the limitations induced by the chemistry, for instance, the bound on the vertex degree. The fact that this chemical system can be well described by graph theory is already hinted by a broad range of analogues to graph-theoretical terminology that exists in polymer chemistry: vertex (*monomer*), degree bound (*functionality*), graph (*polymer network*), tree (*branched polymer*), connected component (*polymer molecule*), giant component (*gel*), density (*conversion*), etc.

In this paper a random graph process is introduced to model an evolving molecular network. The degree distribution of this random graph is defined by a time-continuous evolution equation that mimics the chemistry of the step-growth polymerisation process. This process starts with disconnected vertices and progresses up to the point where no new edge can be placed. The degree of each vertex is bounded, but different bounds may be defined for distinct vertices. Therefore, we distinguish between the degree—actual number of incident edges, and the functionality—pre-imposed bound on the number of incident edges. At each time step, the probability that a vertex receives an edge is proportional to the difference between the vertex's functionality and degree. The share of vertices in each functionality class is pre-defined, and constitutes the only input parameter for the random graph model.

Most of the available studies target narrow special cases of this system and pursue results with a distinct reasoning from the graph-theoretical one. Important contributions include: Hamilton–Jacobi formalism as applied to dynamic graphs with globally bounded degrees [1], results on the grabbing-particle system [4], open-form analytical results for non-phase-transiting systems [11], combinatorial analysis for monomers bearing identical groups [6], closed-form analytical [31] and numerical [17] results for trifunctional vertices in a directed topology, analytical results for mixture of bi- and trifunctional vertices [12], analytical results on phase transition in evolving directed graphs [14], and stochastic simulations on molecular networks [16]. The random graph model is also related to many processes outside polymer chemistry. For instance, Smoluchowski coagulation equation with a multiplicative kernel governs the dynamics of component-size distribution of the polymerisation random graph with trifunctional vertices. Only in this special case, the analytical expression for component sizes is available also after the phase transition, for a review on Smoluchowski coagulation see Refs. [3,29]. In probability theory, the gambler's ruin problem for infinite number

of games is equivalent to finding criteria for the phase transition in the polymerisation random graph with vertices not exceeding degree three [10].

The rest of the paper is organised as follows. First, a differential-difference equation describing evolution of the degree distribution due to the step-growth polymerisation process is formulated and solved in time. Then, given the time dependent degree distribution, the emergence of the giant component is analysed. This includes results on the edge density at which the giant component appears and the criterion on the functionality distribution that admit emergence of the giant component at finite time. Furthermore, the size distribution of connected components is resolved and expressions for the expected component size are given. Finally, the theoretical results are discussed for a few special cases. The theory is also compared against the size distributions that were generated by a Monte Carlo simulation.

2 Evolution process for the degree distribution

This paper studies infinite graphs as a model for a polymer network: a chemical system composed of randomly interconnected identical units. In the infinite graph, degree distribution $u(n), n = 0, 1, 2, \ldots$ is the probability that a randomly sampled vertex has n adjacent edges [24]. Since a degree of a vertex cannot be arbitrary large in a chemical system, each vertex is assigned a bound on its degree, $m = 0, 1, 2, \ldots$. To copy the chemical terminology, we refer to this bound as the functionality [28]. So that one may speak of a two-variate distribution $u(n, m), \ n, m = 0, 1, 2, \ldots$ as the probability to sample a vertex with degree n and functionality m, such that $u(n, m) = 0$ for $n > m$. We will now construct an evolutionary process for $u(k)$ that mimics the step-growth polymerisation of multifunctional monomers. This linking process starts with disconnected vertexes, that is the probability to sample a vertex of degree zero is $d(0, k) = 1$, and the process ends when one samples a vertex with $n = m$ with probability one. The precise rule of assigning a new edge is the following conceptualisation of the step-growth polymerisation process: on each time step, one samples two candidate vertices with probability proportional to $(m - n)u(n, m)$ and connects them with an edge. So that

$$\{(n_1, m_1), (n_2, m_2)\} \rightarrow \{(n_1 + 1, m_1), (n_2 + 1, m_2)\}, \ n_1 \leq m_1, \ n_2 \leq m_2, \quad (1)$$

where (n_1, m_1) and (n_2, m_2) are the configurations of the candidate vertices. This linking process may be viewed as a generalisation of the linking process with constant degree bounds (all vertices have the same functionality m) as introduced in Ref. [1], Eq. (3). An alternative way of introducing (1) is by writing the corresponding reaction mechanism for monomer species $M_{n,m}$:

$$M_{n_1,m_1} + M_{n_2,m_2} \xrightarrow{(m_1-n_1)(m_2-n_2)} M_{n_1+1,m_1} + M_{n_2+1,m_2}. \quad (2)$$
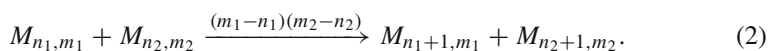
Both notations (1) and (2) are equivalent and correspond to the following Kolmogorov forward equation governing the evolution of $u(n, m)$,

$$\frac{\partial}{\partial t} u(n, m, t) = \left((m - n + 1)u(n - 1, m, t) - (m - n)u(n, m, t) \right) \times$$
$$\sum_{m=0}^{\infty} \sum_{n=0}^{m} \left(m\, u(n, m, t) - n\, u(n, m, t) \right); \tag{3}$$

where at $t = 0$, $u(n, m, t)$ satisfies the following initial conditions,

$$u(0, m, 0) = f_m,$$
$$u(n, m, 0) = 0, \ n > 0. \tag{4}$$

In this equation, the probability to sample a vertex of functionality m is constant over time, $\sum_{n=0}^{\infty} u(n, m, t) = f_m$, $\sum_{m=1}^{\infty} f_m = 1$, and f_m is treated as the only parameter of the model. The sum written in the second line of Eq. (3) represents the expected number of unused but potentially available edges and can be viewed as a difference of two partial moments, $\mu_{01}(t) - \mu_{10}(t)$, where

$$\mu_{ij}(t) = \sum_{m=0}^{\infty} \sum_{n=0}^{m} n^i\, m^j\, u(n, m, t). \tag{5}$$

The edge density, $c(t) \in [0, 1]$, is a ratio of expected number of edges at time t to the expected number of edges at the end of the process:

$$c(t) = \frac{\mu_{10}(t)}{\mu_{01}}.$$

It is convenient to use $c(t)$ as an alternative measure of the progress. The differential equation (3) falls into the class of linear population balance equations. This class of equations frequently appears as a model for many chemical and biological problems where it is usually approached numerically [18,20]. In the current case, it is possible to find an analytical solution of (3) by transforming the equation to the domain of generating functions, solving the corresponding partial differential equation, and applying the inverse transform.

Let us rewrite (3) in terms of a univariate generating function,

$$U(x, m, t) = \sum_{n=0}^{m} x^n u(n, m, t), \ |x| < 1, \ x \in \mathbb{C}.$$

Taking generating function transform on both sides of Eq. (3) leads to a partial differential equation (PDE),

$$\begin{cases} \frac{\partial}{\partial t} U(x, m, t) = \left((mx - m)U(x, m, t) + (x - x^2)\frac{\partial}{\partial x}U(x, m, t) \right)(\mu_{01}(t) - \mu_{10}(t)) \\ U(x, m, 0) = f_m, \ |x| < 1. \end{cases} \tag{6}$$

The first partial moments appearing in (6) can be related to the generating function $U(x, m, t)$,

$$
\mu_{01}(t) = \sum_{m=0}^{\infty} m U(1, m, t);
$$

$$
\mu_{10}(t) = \sum_{m=0}^{\infty} \frac{\partial}{\partial x} U(x, m, t)|_{x=1};
$$

(7)

Substituting (7) into (6) we obtain a system of ordinary differential equations for the partial moments,

$$
\begin{cases}
\mu'_{10}(t) = \left(\mu_{01}(t) - \mu_{10}(t)\right)^2, \\
\mu'_{01}(t) = 0,
\end{cases}
$$

(8)

that is subject to initial conditions $\mu_{10}(0) = 0$, $\mu_{01}(0) = \mu_{01}$. Solving (8) gives

$$
\mu_{10}(t) = \frac{\mu_{01}^2 t}{1 + \mu_{01} t};
$$

$$
\mu_{01}(t) = \mu_{01};
$$

(9)

Now, having explicit expressions for $\mu_{10}(t), \mu_{01}(t)$ at hand, allows us to write the solution of PDE (6),

$$
U(x, m, t) = \left(\frac{1 + \mu_{01} t}{1 + \mu_{01} t \, x}\right)^{-m} f_m,
$$

(10)

which, in turn, generates $u(n, m, t)$,

$$
u(n, m, t) = \binom{m}{n} (\mu_{01} t)^n (1 + \mu_{01} t)^{-m} f_m, \quad n \le m.
$$

(11)

The latter expression can be reformulated in terms of edge density $c(t)$ instead of time. To do this, it is enough to realise that $c(t) = \frac{\mu_{10}(t)}{\mu_{01}} = \frac{\mu_{01} t}{1 + \mu_{01} t}$ and $(\mu_{01} t)^n (1 + \mu_{01} t)^{-m} = (\mu_{01} t)^n (1 + \mu_{01} t)^{-n} (1 + \mu_{01} t)^n (1 + \mu_{01} t)^{-m} = \left(\frac{\mu_{01} t}{1 + \mu_{01} t}\right)^n \left(\frac{1}{1 + \mu_{01} t}\right)^{m-n} = \left(\frac{\mu_{01} t}{1 + \mu_{01} t}\right)^n \left(1 - \frac{\mu_{01} t}{1 + \mu_{01} t}\right)^{m-n}$ so that Eq. (11) transforms to

$$
u(n, m, t) = \binom{m}{n} c^n(t) \left(1 - c(t)\right)^{m-n} f_m, \quad n \le m.
$$

(12)

Expressions (11), (12) satisfy the initial conditions (4), whereas in the limiting case of $t \to \infty$, the degree distribution and the distribution of maximal functionalities coincide:

$$
\begin{cases}
\lim_{t \to \infty} u(n, m, t) = f_m, & n = m; \\
\lim_{t \to \infty} u(n, m, t) = 0, & n < m.
\end{cases}
$$

The actual degree distribution u(n), is found by summating u(n, m, t) over functionalities m,

$$u(n, t) = \sum_{m=1}^{\infty} u(n, m, t).$$ (13)

Here, we employed the fact, that u(n, m, t) = 0, for $n > m$. Degree distribution $u(n, t)$ evolves form the Kronecker's delta function, δ_n at $t = 0$ to f_m in the limit of $t \to \infty$. The moments of the degree distribution, $\mu_i = \sum_{n=0}^{\infty} n^i u(n, t) = \mu_{i0}$ can be directly found from summation of Eq. (13). For instance the expressions for the first three moments read,

$$\mu_1(t) = \frac{\mu_{01}^2 t}{1 + \mu_{01} t},$$

$$\mu_2(t) = \frac{\mu_{01}^2 t (1 + \mu_{02} t)}{(1 + \mu_{01} t)^2},$$ (14)

$$\mu_3(t) = \frac{\mu_{01}^2 t (1 - 3\mu_{01} t + 4\mu_{02} t)}{1 + \mu_{01} t^2}.$$

3 Global properties of the network, the giant component

Up to this point we have discussed only local properties, i.e. the way the graph can be seen from a viewpoint of a single vertex. However, in a randomly interconnected system, local properties, as for instance the degree distribution, play a decisive role in defining the global properties of the graphs itself. An important finding that allows us to connect the the two worlds is the result by Molloy and Reed on the existence of the giant component [23]: there exists a component of the same order of magnitude as the whole graph (the giant component) iff,

$$\sum_{n=1}^{\infty} n(n - 2) u(n, t) > 0,$$

while the equality is reached exactly at the phase transition point. This phase transition condition can be rewritten in terms of moments (7),

$$\mu_{20}(t) - 2\mu_{10}(t) = 0.$$ (15)

Substituting the analytical expression for moments (9) into Eq. (15) we obtain the phase transition time (or the gelation time in the chemical terminology),

$$t_g = \frac{1}{\mu_{02} - 2\mu_{01}}.$$ (16)

Similarly, the edge density at the phase transition (or gel conversion) is written out as

$$c_g = \frac{\mu_{01} t_g}{1 + \mu_{01} t_g} = \frac{\mu_{01}}{\mu_{02} - \mu_{01}}. \tag{17}$$

From the last relation (17) we can see that the system features the phase transition in a finite time only when $\mu_{02} - 2\mu_{01} > 0$. If the inequality is replaced by an equality ($\mu_{02} = 2\mu_{01}$), then the phase transition will be approached asymptotically at $t \to \infty$. This brings us to the following, especially important for its chemical context,

Corollary *Let M monomer species of functionalities $m = 1, \ldots, M$ and fractions $f_1, f_2, \ldots, f_M$, $\sum\limits_{m=1}^{M} f_m = 1$ react at constant rate k_p, then the system features the phase transition in a finite time if and only if*

$$\sum_{m=1}^{M} m^2 f_m - 2 \sum_{m=1}^{M} m f_m > 0. \tag{18}$$

If the phase transition occurs, then it occurs at the following time and edge density,

$$t_g = \left(k_p \sum_{m=1}^{M} (m^2 - 2m) f_m \right)^{-1},$$

and

$$c_g = \left(\sum_{m=1}^{M} (m^2 - m) f_m \right)^{-1} \sum_{m=1}^{M} m f_m. \tag{19}$$

As special cases of this corollary, the following statements hold true.

1. If all monomers have the same functionality m, then the phase transition is reached in a finite time only if $m \geq 3$ (i.e. m is the smallest positive integer satisfying $m^2 - 2m > 0$).
2. Adding (or removing) monomers of functionality two does not affect phase transition time t_g, whereas it does alter the edge density at the phase transition, c_g.
3. Adding sufficient amount of f_1 to any system will prevent the phase transition.
4. Consider a system that consists of two species: monomers with functionality m that are present at fraction f_m and monomers with functionality one, that are present at fraction $f_1 = 1 - f_m$. The system does not go through the phase transition in finite time if,

$$f_1 > \frac{m^2 - 2m}{m^2 - 2m + 1}. \tag{20}$$

5. When all monomers have functionality m, the polymerisation leads to an infinite network at edge density

$$c_g = \frac{m}{m^2 - m} = \frac{1}{m - 1}.$$

The latter equation was derived by Flory [9]. Although Flory did not consider non-constant functionality, somewhat later, he conjectured that the equation can be generalised for a mixture of arbitrary functional monomers if $m - 1$ were replaced "by the appropriate average, weighted according to the numbers of functional groups." (see [8], p. 353).

4 Size distribution of connected components

For the sake of brevity we drop time argument t where it leads to no confusion, and refer to the degree distribution, as given in Eq. (13), by simply $u(n)$ or by its generating function,

$$U(x) = \sum_n x^n u(n), \quad |x| \leq 1, \quad x \in \mathbb{C}. \tag{21}$$

We will now apply the theory from Refs. [24,25] to recover other non-local properties of the polymer network.

When talking about a property of a randomly sampled vertex in an infinite graph it is important to specify what is exactly the sampling rule. Up to this point, we considered the case when every vertex has equal chances to be sampled. Consider a different strategy to choose a vertex: suppose one samples *an edge* at random, so that every edge has equal probability to be sampled. Then, one of incident to this edge vertices is chosen and the edge itself is removed. We will refer to this vertex as the biased vertex. Let $u_1(n)$ denotes the probability that a biased vertex has n incident edges. Then,

$$u_1(n) = \frac{(n+1)u(n+1)}{\sum\limits_{n=1}^{\infty} n u(n)},$$

and the corresponding generating function is

$$U_1(x) = \frac{U'(x)}{U'(x)|_{x=1}}. \tag{22}$$

A connected component is a subset of vertices in a graph, such that every couple of vertices is connected with a path. Let $w(n)$ denotes the probability that a randomly sampled node belongs to a connected component of size n. Similarly to definition of $u_1(n)$, let $w_1(n)$ denotes the probability that a biased vertex belongs to a connected component of size n. Newman et al. [24] noticed that the generating functions for $u_1(n)$ and $w_1(n)$ are related by a functional equation

$$W_1(x) = x \sum_{n=0}^{\infty} u_1(n) W_1^n(x), \tag{23}$$

where $W_1(x)$ generates $w(n)$ and $U_1(x)$ generates $u_1(n)$. This equation has a straightforward interpretation: the equation unfolds the generating function for $w_1(n)$ as a sum over all configurations of a biased vertex. Each configuration occurs with probability

$u_1(n)$ and involves n biased sub-components of size $w_1(n)$. Furthermore, the sum in Eq. (23) can be in itself viewed as the definition of the generating function. So that one may write,

$$W_1(x) = x U_1\Big(W_1(x)\Big).$$ (24)

Following a similar logic to derivation of (24), the generating function for $w(n)$ reads

$$W(x) = x U\Big(W_1(x)\Big).$$ (25)

Due to Lagrange inversion principle [2], the system of functional equations (24), (25) has a unique solution. Furthermore, the formal expression for $w(n)$ can be written out in terms of convolution powers[15],

$$w(n,t) = \begin{cases} \frac{\mu_{01}^2 t}{(1+\mu_{01}t)(n-1)} u_1^{*n}(n-2), & n > 1, \\ u(0) & n = 1. \end{cases}$$ (26)

Here $u_1^{*n}(n)$ denotes the convolution power,

$$u(k)^{*n} = u(k)^{*n-1} * u(k),$$ (27)

where

$$f(k) * g(k) = \sum_{i+j=k} f(i)g(j), \ i, j, k \geq 0.$$

On practice, the exact numerical values of (26) can be computed by making use if the convolution theorem and evaluating (27) with the fast Fourier transform algorithm. Such numerical routine results in $\mathcal{O}(n \log n)$ multiplicative operations. If all vertices have the same functionality m, $f_m = 1$, then $w(n)$ is simply given by,

$$w(n,t) = \frac{\mu_{01}^2 t}{(1+\mu_{01}t)(n-1)} \binom{n(m-1)}{n-2} (1+\mu_{01}t)^{-n(m-3)} (\mu_{01}t)^{n-2}.$$ (28)

The restrictions imposed by chemistry of the polymerisation system guarantee that $u(n) = 0$ for some $n > n_{max}$. This class of degree distributions features a defined asymptotic behaviour of $w(n)$ at large $n \gg 1$, see Ref. [15]. Namely,

$$\lim_{n\to\infty} \frac{w(n)}{w_\infty(n)} = 1,$$

where

$$w_\infty(n) = C_1(t) e^{-C_2(t)n} n^{-3/2},$$ (29)

and the coefficients are given by

$$C_1(t) = \frac{\mu_1^2(t)}{\sqrt{2\pi\left(\mu_1(t)\mu_3(t) - \mu_2^2(t)\right)}} = \frac{\mu_{01}^2\sqrt{t}}{\sqrt{2\pi(\mu_{02} - \mu_{01})(2 + 3\mu_{01}t - \mu_{02}t)}},$$

$$C_2(t) = \frac{\left(\mu_2(t) - 2\mu_1(t)\right)^2}{2\left(\mu_1(t)\mu_3(t) - \mu_2^2(t)\right)} = \frac{(1 - (\mu_{02} - 2\mu_{01})t)^2}{2t(\mu_{02} - \mu_{01})(2 + 3\mu_{01}t - \mu_{02}t)}.$$

In the latter transformation we made use of the expressions of the moments (14).

One may see that at the phase transition, when $t = t_g$, the coefficient in the exponential function in (29) vanishes and the asymptote switches to the power law decay.

5 Expected size of connected components

It is important to note, that $w(n)$ describes only finite components. Before the phase transition, a randomly sampled node belongs to a finite component with probability one, therefore $\sum_{n=1} w(n) = W(1) = 1$. After the phase transition, when $t > t_g$, the probability that a randomly sampled node belongs to a finite component is smaller than one and $w(n)$ fails to be normalised:

$$W(1) = \sum_n w(n) = 1 - g_f,$$

where g_f is the the probability that a randomly sampled vertex belongs to the giant component (or gel fraction): $g_f = 0$ for $t < t_g$ and $g_f \in [0, 1]$ for $t > t_g$. Plugging $x = 1$ into (24) one obtains,

$$g_f = 1 - W(1) = 1 - U(r_0), \tag{30}$$

where $r_0 := W_1(1)$ is the smallest positive fixed point of $U_1(x)$,

$$r_0 = U_1(r_0). \tag{31}$$

We will now derive the expression for the expected size of connected component, as given by

$$M_w := \frac{\sum_{n=1} nw(n)}{\sum_{n=1} w(n)} = \frac{W'(1)}{W(1)}.$$

Let $t < t_g$, then $W(1) = W_1(1) = 1$ and evaluating $W_1'(1)$ from Eq. (24) gives,

$$W_1'(1) = \text{U}(W_1(1)) + \text{U}'(W_1(1))W_1'(1) = 1 + \text{U}_1'(1)W_1'(1) = \frac{1}{1 - \text{U}_1'(1)}.$$

Similarly, evaluating $W'(1)$ from Eq. (25) gives,

$$\begin{aligned} M_w &= W'(1) = 1 + \text{U}'(1)W_1'(1) = 1 + \frac{\text{U}'(1)}{1 - \text{U}_1'(1)} = 1 - \frac{\mu_1^2(t)}{\mu_2(t) - 2\mu_1(t)} \\ &= 1 + \frac{\mu_{01}^2 t}{1 + 2\mu_{01}t - \mu_{02}t}, \quad t < t_g. \end{aligned} \tag{32}$$

The latter transformation is made realising that $\text{U}'(1) = \mu_1(t)$, $\text{U}_1'(1) = (\mu_2(t) - \mu_1(t))/\mu_1(t)$ and the moments of the degree distribution are as defined by Eqs. (14).

Let $t > t_g$, then $W_1(1) = r_0 \neq 1$ and evaluating $W_1'(1)$ from Eq. (24) gives the following equality,

$$W_1'(1) = \text{U}_1(W_1(1)) + \text{U}_1'(W_1(1))W_1'(1) = r_0 + \text{U}_1'(r_0)W_1'(1), \tag{33}$$

so that

$$W_1'(1) = \frac{r_0}{1 - \text{U}_1'(r_0)}.$$

Evaluating $W'(1)$ from Eq. (25), gives

$$W'(1) = \text{U}(W_1(1)) + \text{U}'(W_1(1))W_1'(1) = W(1) + \text{U}'(r_0)\frac{r_0}{1 - \text{U}_1'(r_0)}.$$

Now, realising that according to Eq. (22), $\text{U}'(r_0) = \text{U}'(1)\text{U}_1(r_0) = \mu_1 r_0$, one obtains:

$$\begin{aligned} M_w &= \frac{W'(1)}{W(1)} = \frac{1 - g_f + \frac{\mu_1 r_0^2}{1 - \text{U}_1'(r_0)}}{1 - g_f} = 1 + \frac{\mu_1 r_0^2}{(1 - g_f)(1 - \text{U}_1'(r_0))} = \\ &= 1 + \frac{\mu_{01}^2 r_0^2 t}{(1 + \mu_{01}t)(1 - g_f)(1 - \text{U}_1'(r_0))}, \quad t > t_g. \end{aligned} \tag{34}$$

Together, Eqs. (32) and (34) define the expected component size before and after the phase transition, that is at $t \in [0, t_g) \cup (t_g, \infty)$. Precisely at the phase transition, $t = t_g$, the expected component size diverges, as $(t - t_g)^{-1}$. So that

$$\lim_{t \to t_g} \frac{M_w(t)}{(t - t_g)^{-1}} = \mathcal{O}(1).$$

This happens due to a different type of the asymptotical behaviour of the size distribution at the phase transition, see Eq. (29).
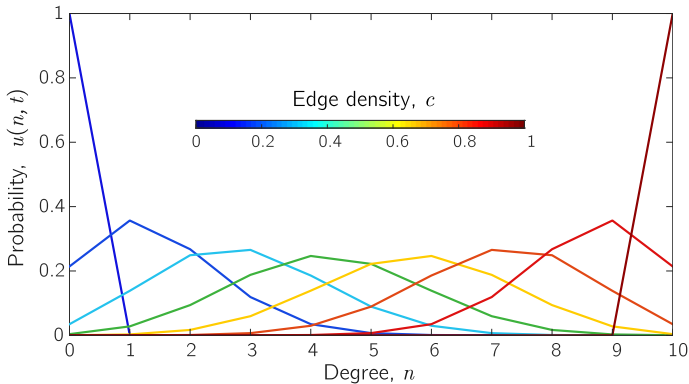
Fig. 1 Evolution of the degree distribution for $f_{10} = 1$

6 Interpretation of the results and examples

The present paper introduces a model for studying polymer networks composed of multifunctional monomers that polymerise according to the step-growth mechanism (2). This model associates a vertex with a monomer and an edge with a chemical bond between two such monomers in the network. A resulting topology of the polymer network is viewed as a random graph defined by its degree distribution. Initial fractions of monomers of different functionalities f_m are directly related to molar concentrations of monomer species. The reaction kinetics is formalised by the master equation (3) and yields an analytical expression for the degree distribution at any point of time (11). Although the master equation (3) has a unit rate, an arbitrary reaction rate can be modelled by simply scaling time variable t in a linear fashion. An example of a degree distribution evolving in time is given in Fig. 1. In this example, the initial condition of the kinetic model is chosen to be $f_{10} = 1$, that corresponds to pure 10-functional monomers. In the given context, both, initial and terminal degree distributions are Kronecker's delta functions positioned correspondingly at $m = 0$ and $m = 10$.

A deeper analysis reveals that when initial concentrations of monomers satisfy condition (18), the random graph develops a giant component at time t_g that is given by Eq. (16). This event is related to the fact that the molecular network undergoes a phase transition. Such phase transition is called gelation, and is a well-documented chemical phenomenon that signifies transition from liquid-like to solid-like state in soft matter[30,32]. Figure 2 presents two examples showing how t_g is influenced by varying f_m. The figure illustrates the fact that addition of one- and two- functional vertices may be used to control the timing of the phase transition: addition of two-functional vertices postpones the emergence of the giant component in terms of c_g, whereas t_g remains invariant; addition of one-functional vertices may entirely prevent it.

The size distribution of connected components, as given in Eq. 26, is interpreted as the molecular weight distribution, whereas the asymptote (29) might serve as a good way to approximate the latter if rapid computations are required. Evolution of
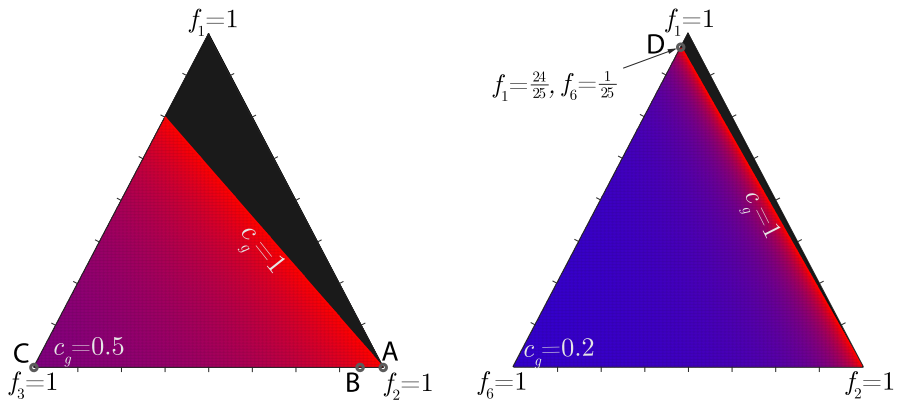
Fig. 2 The edge density at phase transition, c_g, is plotted as a function of concentration in barycentric coordinates for two sets of monomer functionalities: (*left*) the only non-zero concentrations are f_1, f_2, f_3, (*right*) the only non-zero concentrations are f_1, f_2, f_6. The black area corresponds to the configurations that does not feature the phase transition. The points (A, B, C, D) refer to special cases discussed in the paper
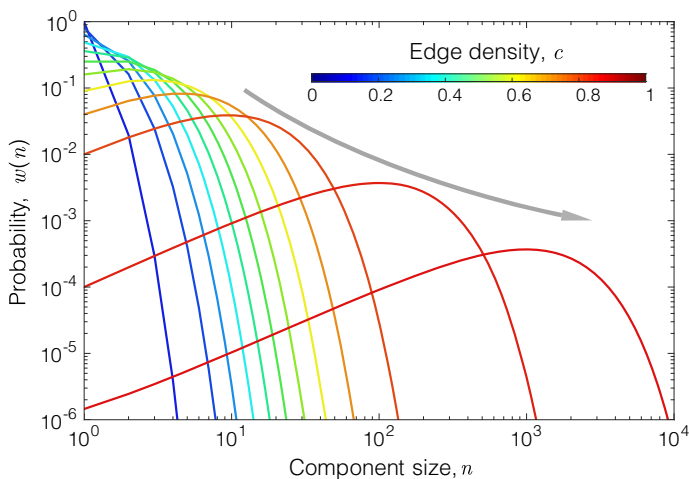

Fig. 3 Evolution of the size distribution of connected components for a system with $f_2 = 1$ and various values of the edge density The giant component emerges asymptotically at infinite time $(c \to 1)$

the expected number of this distribution, also known in the chemical literature as number-average molecular weight, is given by Eqs. (32), (34).

More examples of phase transitioning systems, as obtained for a few instances of functionality distribution f_m, follow below. These examples are supplemented with a MATALB code that reproduces the size distribution and the corresponding expected value for an arbitrary functionality distribution and the process time [13].

Example 1 We consider vertices with at most degree 2, that is $f_2 = 1$. Graphs generated by such a process are always linear and, according to (18), the giant component is reachable only asymptotically at $t \to \infty$. Furthermore, a small perturbation,
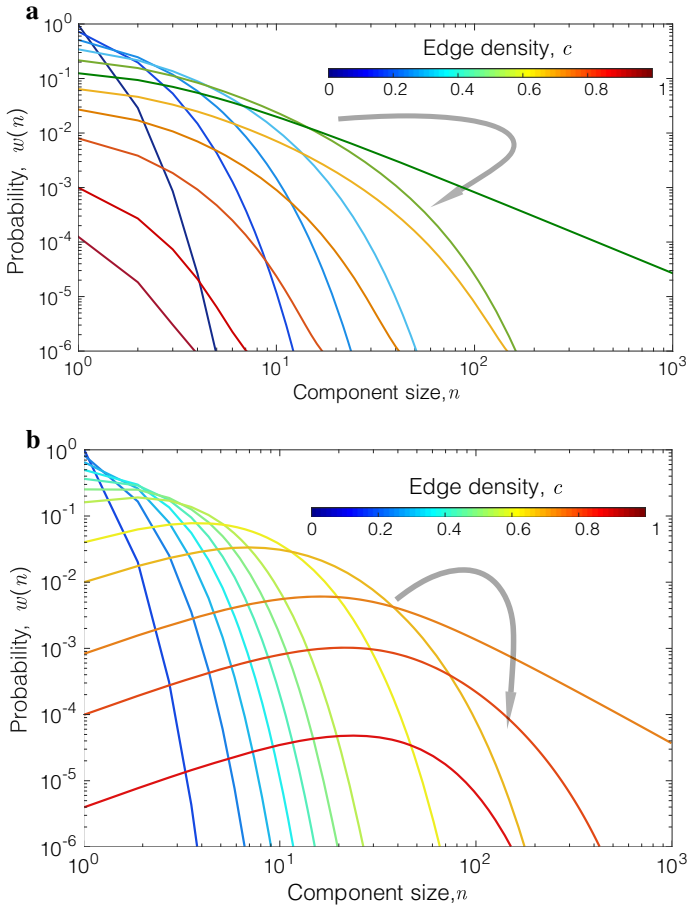
Fig. 4 Evolution of the size distribution of connected components for: **a** a system with $f_3 = 1$, phase transition at $c_g = \frac{1}{2}$, and **b** a system with $f_2 = \frac{49}{50}$, $f_3 = \frac{1}{50}$), phase transition at $c_g = \frac{101}{104}$
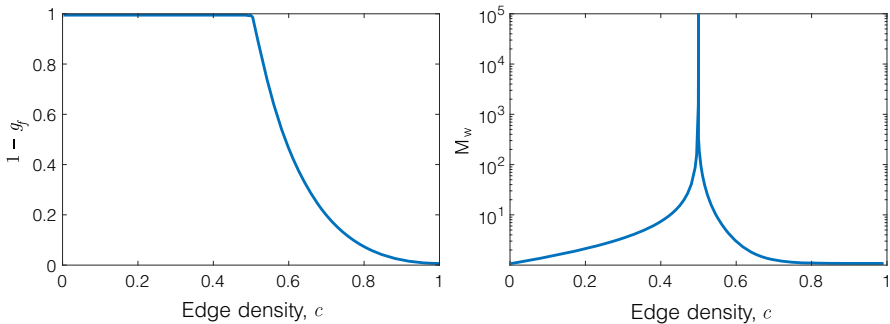

Fig. 5 Emergence of the giant component in a system with $f_3 = 1$ that features phase transition at $c = \frac{1}{2}$. *Left* probability that a randomly sampled node belongs to a finite-size connected component. *Right* the expected size of connected components features a singularity at the phase transition

$f_1 = \varepsilon$, $f_2 = 1 - \varepsilon$, prevents emergence of the giant component even at infinite time (see points A at barycentric plot of configurations, Fig. 2). The component-size distribution is illustrated in Fig. 3. One may notice the constant "drift" (as indicated with an arrow) of the distribution towards larger values of components sizes. The distribution features the exponential asymptote at any $t > 0$.

Example 2 In this example we consider a system with $f_3 = 1$. This random graph consists of three-functional vertices and features the phase transition at edge density $c = \frac{1}{2}$ (configuration C in Fig. 2). The component-size distribution is illustrated in Fig. 4a. Asymptotically, when $n \to \infty$, the component-size distribution switches
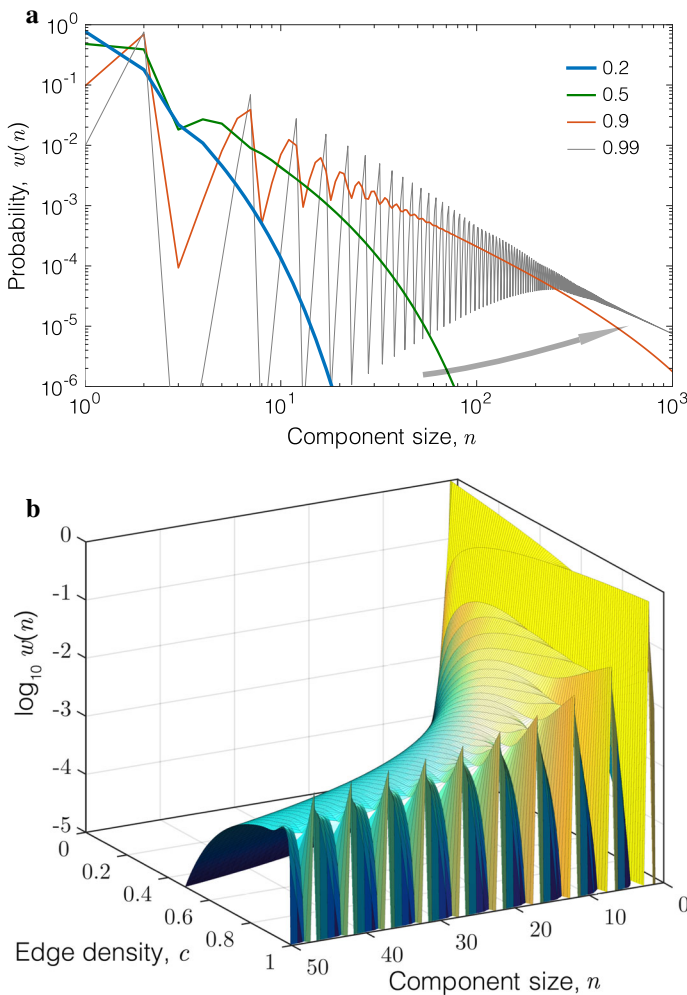

Fig. 6 The size distributions of connected components for a system with $f_1 = \frac{24}{25}$, $f_6 = \frac{1}{25}$ as predicted by the theory. **a** The size distributions at a few instances of time. **b** A surface representing the evolution of the size distribution during the whole time-continuous process, $c \in [0, 1]$.

between exponential decay ($0 < c < 0.5$), algebraic decay ($c = 0.5$), and back to exponential decay again ($0.5 < c < 1$). Prior to the phase transition, the distribution 'drifts' to the right (expected component size becomes larger), and swings back to small expected component sizes at the end of the process. When edge density traverses the critical point $c_g = \frac{1}{2}$, the probability that a randomly sampled node belongs to finite-size component departs from one and the expected component size features a singularity, see Fig. 5.

As shown in Fig. 2, one may postpone the phase transition so it occurs anywhere between 0.5 and 1 by adding vertices of functionality 2 to the system. For instance, a mixture of vertices with functionalities two and three having fractions $f_2 = \frac{49}{50}$ and $f_3 = \frac{1}{50}$, as denoted by point B in Fig. 2, postpones the phase transition to $c_g = \frac{101}{104} \approx 0.97$. The evolution of the size distribution for this case is depicted in Fig. 4b.

While vertices of degree two postpone the phase transition, vertices of degree one may prevent it by "consuming" all available edges in a single connected component and thus locking its size finite. For this reason vertices of degree one are called *termination agents* within the chemical context. Depending on what is the degree of the other species, the probability of randomly selecting a component may feature regular oscillations. For instance, in a dense, $c = 1$, mixture of m-functional and one-functional vertices, connected components can take their sizes only from

$$n \in \{2\} \cup \{km - k + 2 \mid k = 1, 2, \dots \}.$$

Here we rely on the fact that non-giant components do not contain cycles[25]. Whereas when edge density $c < 1$, the sizes of connected components are not restricted to this set and, as is demonstrated in the next example, the transition of the size distribution from $c < 1$ to $c = 1$ is non-trivial.
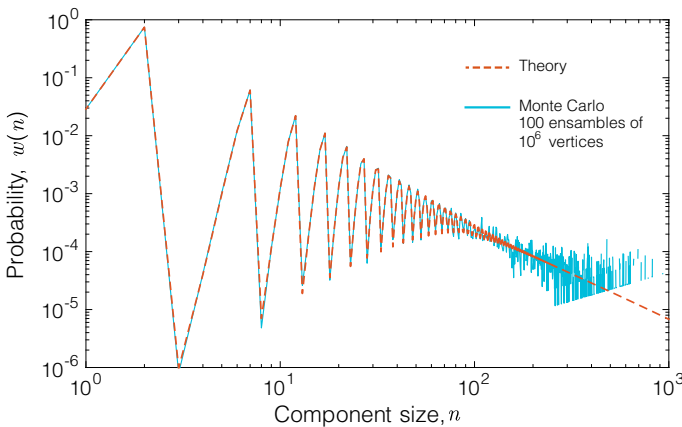

Fig. 7 The size distribution of connected components for a system with $f_1 = \frac{24}{25}$, $f_6 = \frac{1}{25}$ at edge density $c = 0.97$ is obtained with two different methods: (*red line*) the theory; (*blue line*) Monte Carlo simulations of a network with 10^6 vertices. The simulation data is averaged over 100 simulation runs (Color figure online)

Example 3 We consider a mixture of one- and six-functional vertices present with fractions $f_1 = \frac{24}{25}$, $f_6 = \frac{1}{25}$. This distribution of functionalities features the phase transition at $c = 1$. As illustrated in Fig. 6a, the size distribution decays monotonically at low edge densities, but switches to oscillations as c approaches 1. The switch itself is gradual as can be seen in Fig. 6b. In Fig. 7, the theoretical results are compared to component-size distribution generated by Monte Carlo (MC) computations. The theory and MC data are in a perfect agreement; however, despite extensive size of MC computations (100 ensembles of size 10^6 vertices), the MC resolution in the tail of the distributions remains poor.

References

1. E. Ben-Naim, P. Krapivsky, Dynamics of random graphs with bounded degrees. J. Stat. Mech. Theory Exp. **2011**(11), P11,008 (2011)
2. F. Bergeron, G. Labelle, P. Leroux, *Combinatorial Species and Tree-Like Structures* (Cambridge University Press, Cambridge, 1998)
3. J. Bertoin, V. Sidoravicius, The structure of typical clusters in large sparse random configurations. J. Stat. Phys. **135**(1), 87–105 (2009)
4. J. Bertoin, V. Sidoravicius, M.E. Vares, A system of grabbing particles related to galton-watson trees. Random Struct. Algorithms **36**(4), 477–487 (2010)
5. D. Bonchev, *Chemical Graph Theory: Introduction and Fundamentals*, vol. 1 (CRC Press, Boca Raton, 1991)
6. D. Durand, C.M. Bruneau, Statistics of random macromolecular networks, 1. Stepwise polymerization of polyfunctional monomers bearing identical reactive groups. Macromol. Chem. Phys. **183**(4), 1007–1020 (1982)
7. E. Estrada, Generalization of topological indices. Chem. Phys. Lett. **336**(3), 248–252 (2001)
8. P. Flory, *Principles of Polymer Chemistry* (Cornell University Press, New York, 1953), p. 353
9. P.J. Flory, Molecular size distribution in three dimensional polymers. I. Gelation. J. Am. Chem. Soc. **63**(11), 3083–3090 (1941)
10. G. Harik, E. Cantú-Paz, D.E. Goldberg, B.L. Miller, The gambler's ruin problem, genetic algorithms, and the sizing of populations. Evol. Comput. **7**(3), 231–253 (1999)
11. L.T. Hillegers, J.J. Slot, Step-growth polymerized systems of general type "$A_{f_i} B_{g_i}$": generating functions and recurrences to compute the MSD. Macromol. Theory Simul. **24**(3), 248–259 (2015)
12. P.D. Iedema, M. Dreischor, K.D. Hungenberg, Y. Orlov, Predicting the change of MWD caused by interchange reactions during melt-mixing of linear and branched polycondensates (AB_2). Macromol. Theory Simul. **21**(9), 629–647 (2012)
13. I. Kryven, Git repository with supporting code in matlab. https://github.com/ikryven/PolyRandGrpah
14. I. Kryven, Emergence of the giant weak component in directed random graphs with arbitrary degree distributions. Phys. Rev. E **94**(1), 012–315 (2016)
15. I. Kryven, General expression for the component size distribution in infinite configuration networks. Phys. Rev. E **95**(5), 052–303 (2017)
16. I. Kryven, J. Duivenvoorden, J. Hermans, P.D. Iedema, Random graph approach to multifunctional molecular networks. Macromol. Theory Simul. **25**(5), 449–465 (2016)

17. I. Kryven, P. Iedema, Predicting multidimensional distributive properties of hyperbranched polymer resulting from ab_2 polymerization with substitution, cyclization and shielding. Polymer **54**(14), 3472–3484 (2013)
18. I. Kryven, P. Iedema, Deterministic modelling of copolymer microstructure: composition drift and sequence patterns. Macromol. React. Eng. **9**(3), 285–306 (2014)
19. I. Kryven, P.D. Iedema, Transition into the gel regime for crosslinking radical polymerisation in a continuously stirred tank reactor. Chem. Eng. Sci. **126**, 296–308 (2015)
20. I. Kryven, S.Röblitz, C. Schütte, Solution of the chemical master equation by radial basis functions approximation with interface tracking. BMC Syst. Biol. **9**(1), 67 (2015)
21. Z. Mihalić, D. Veljan, D. Amić, S. Nikolić, D. Plavšić, N. Trinajstić, The distance matrix in chemistry. J. Math. Chem. **11**(1), 223–258 (1992)
22. B. Mohar, T. Pisanski, How to compute the wiener index of a graph. J. Math. Chem. **2**(3), 267–277 (1988)
23. M. Molloy, B. Reed, The size of the giant component of a random graph with a given degree sequence. Comb. Probab. Comput. **7**(03), 295–305 (1998)
24. M. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010)
25. M. Newman, S. Strogatz, D. Watts, Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E **64**(2), 026–118 (2001)
26. M. Randić, Generalized molecular descriptors. J. Math. Chem. **7**(1), 155–168 (1991)
27. D. Rouvray, R. King, *Topology in Chemistry: Discrete Mathematics of Molecules* (Elsevier, Amsterdam, 2002)
28. W.H. Stockmayer, Theory of molecular size distribution and gel formation in branched-chain polymers. J. Chem. Phys. **11**(2), 45–55 (1943)
29. J.A. Wattis, An introduction to mathematical models of coagulation-fragmentation processes: a discrete deterministic mean-field approach. Physica D **222**(1), 1–20 (2006)
30. H.H. Winter, M. Mours, *Rheology of Polymers Near Liquid–Solid Transitions* (Springer, Berlin, 1997), pp. 165–234
31. Z. Zhou, D. Yan, Distribution function of hyperbranched polymers formed by AB_2 type polycondensation with substitution effect. Polymer **47**(4), 1473–1479 (2006)
32. R. Ziff, G. Stell, Kinetics of polymer gelation. J. Chem. Phys. **73**(7), 3492–3499 (1980)