Theoretical and Simulation Results for a 2-type Network Evolution Model*

István Fazekas, Attila Barta

University of Debrecen, Faculty of Informatics, Hungary
fazekas.istvan@inf.unideb.hu
barta.attila@inf.unideb.hu

Abstract

A continuous-time network evolution model is studied. The basic units of the model are edges and triangles. The evolution of the units is governed by a continuous-time branching process. The asymptotic behaviour of the model is studied. It is proved that the number of edges and triangles have the same magnitude on the event of non-extinction, and it is $e^{\alpha t}$, where α is the Malthusian parameter.

Keywords: Random graph, network, branching process, Malthusian parameter

AMS Subject Classification: 05C80, 90B15, 60J85

1. Introduction

Network theory is one of the most popular research topics of our age. It studies both real-life networks and theoretical models. Networks are described by graphs. The nodes of the network are the vertices of the graph, and the connections are the edges. One of the most famous models is the preferential attachment model introduced by Albert and Barabási. It is a discrete time model (that is the evolution

events occur at time $n = 1, 2, \ldots$) and it describes connections of two nodes. The meaning of connection can be cooperation or any interaction. Therefore the connections of more than two nodes are also important. For example, Backhausz and Móri in [1] describe three-interactions, Fazekas and Porvázsnyik in [6] N-interactions, or Fazekas and Perecsényi in [5] star-like connections. Continuous-time network evolution models seem to be more difficult but more realistic models than the discrete time ones. In [8] a continuous-time branching process is applied to govern the evolution mechanism. In [4] we extended the results of [8] for 3 interaction models. There we applied the general theory of branching processes [7]. In this paper we describe our model and present our limit theorems. The mathematical proofs are quite long, so they will be published in another paper (see [3]). In the paper at hand we show numerical and simulation results supporting our mathematical theorems.

2. The Model

In this paper we shall study the following random graph evolution model. Parameter $t \in [0, \infty)$ will denote the time. At the initial time $t = 0$ we start with a single object, it can be either an edge or a triangle. We call this object the ancestor. This ancestor object produces offspring objects which can be also edges or triangles. Then these offspring objects also produce their offspring objects, and so on. The reproduction times of any object, including the ancestor, are given by its own Poisson process with rate 1. We assume that during the evolution, the reproduction processes of different objects are independent. The reproduction processes of the triangles are independent copies of the generic triangle's reproduction mechanism. Similarly, the reproduction processes of the edges are independent copies of the reproduction mechanism of the generic edge.

Consider the generic edge and its Poisson process $\Pi_2(t)$. When the Poisson process jumps, then a new vertex appears an it is connected to our generic edge. The probability that this new vertex is connected to the generic edge by one new edge is r_1, $0 \le r_1 \le 1$. The end point of this new edge is chosen from the two vertices of the generic edge uniformly at random. So in this step an edge produces one edge. The other possibility is that the new vertex is connected to both vertices of the generic edge. It has probability $r_2 = 1 - r_1$. In this case the offspring of the generic edge is a triangle consisting of the generic edge and the two new edges. It is important to fix that in this case the generic edge and the new triangle will produce offspring, but the two new edges will not do it.

For the generic triangle the reproduction is the following. Let $\Pi_3(t), t \ge 0$, denote the Poisson process with rate 1 corresponding to our triangle. The jumping points of $\Pi_3(t)$ are the birth times of the generic triangle. At every birth time a new vertex is added to the graph which can be connected to our generic triangle with j edges ($j = 1, 2, 3$). Let p_j denote the probability that the new vertex will be connected to j vertices of our generic triangle. The vertices to be connected to the new vertex are chosen uniformly at random. It follows from the definition of

the above evolution process that at each birth step we always add 1 new vertex. With probability p_1 the generic triangle gives birth to a new edge. However, in the remaining cases we count only the new triangles and not the new edges. So, with probability p_2 the generic triangle produces one child triangle, moreover, with probability p_3, the generic triangle produces three children triangles.

We shall consider that any edge is of type 2 object and it will be denoted by subscript 2, while for a triangle it will be denoted by 3. So let us denote by $\xi_{i,j}(t)$ the number of type j offspring of the type i generic object up to time t $(i, j = 2, 3)$. As usual, $\xi_{i,j}$, $i, j = 2, 3$, are point processes. Then

$$\xi_2(t) = \xi_{2,2}(t) + \xi_{2,3}(t)$$

is the number of all offspring (either edge or triangle) of the generic edge up to time t. Similarly,

$$\xi_3(t) = \xi_{3,2}(t) + \xi_{3,3}(t)$$

is the number of all offspring (either edge or triangle) of the generic triangle up to time t.

Let us denote by $\tau_3(1), \tau_3(2), \dots$ the birth times of the generic triangle and let $\varepsilon_3(1), \varepsilon_3(2), \dots$ be the corresponding total litter sizes. That is the generic triangle bears $\varepsilon_3(i)$ children being either triangles or edges at the ith birth event. Then $\varepsilon_3(1), \varepsilon_3(2), \dots$ are independent identically distributed discrete random variables with distribution $\mathbb{P}\left(\varepsilon_3(i) = j\right) = q_j, j \geq 1$. We have

$$\mathbb{P}\left(\varepsilon_3(i) = 1\right) = q_1 = p_1 + p_2, \ \ \mathbb{P}\left(\varepsilon_3(i) = 3\right) = q_3 = p_3,$$

$$\mathbb{P}\left(\varepsilon_3(i) = j\right) = q_j = 0, \ \ \text{if} \ \ j \notin \{1, 3\}.$$

Throughout the paper we assume that $p_1 < 1$, because otherwise there were no reproduction of the triangles. We assume that the litter sizes are independent of the birth times.

Let the finite, non-negative random variable λ_3 be the life-length of the generic triangle. We assume that the reproduction terminates at the death of the individual, therefore $\xi_3(t) = \xi_3(\lambda_3)$ for $t > \lambda_3$. Then the reproduction process of a triangle can be given by

$$\xi_3(t) = \sum_{\tau_3(i) \leq t \wedge \lambda_3} \varepsilon_3(i) = S_3\left(\Pi_3\left(t \wedge \lambda_3\right)\right),$$

where $\Pi_3(t)$ is the Poisson process, $S_3(n) = \varepsilon_3(1) + \cdots + \varepsilon_3(n)$ gives the total number of offspring of the generic triangle before the $(n+1)$th birth event and $x \wedge y$ denotes the minimum of $\{x, y\}$.

The survival function of the life-length. Let $L_3(t)$ denote the distribution function of λ_3. Then the survival function of a triangle's life-length is

$$1 - L_3(t) = \mathbb{P}\left(\lambda_3 > t\right) = \exp\left(-\int_0^t l_3(u)\,\mathrm{d}u\right),$$

106

where $l_3(t)$ is the hazard rate of the life-length λ_3. We assume that the hazard rate depends on the total number of offspring, so that

$$l_3(t) = b + c\xi_3(t)$$

with positive constants b and c.

Let λ_2 be the life-length of the generic edge. Then $\xi_2(t) = \xi_2(\lambda_2)$ for $t > \lambda_2$. As the edge always gives birth to one offspring (which can be an edge or a triangle), so the total number of offspring of the generic edge is

$$\xi_2(t) = \Pi_2(t \wedge \lambda_2),$$

where $\Pi_2(t)$ is the Poisson process.

Let $L_2(t)$ denote the distribution function of λ_2. Then the survival function of the life-length of an edge is

$$1 - L_2(t) = \exp\left(-\int\limits_0^t l_2(u)\,\mathrm{d}u\right),$$

where we assume that the hazard rate of the life-length λ_2 is of the form $l_2(t) = b + c\xi_2(t)$.

We have to mention that we do not delete any triangle or edge when it dies, because its edges and vertices can belong to other triangles or edges, too. So we consider a dead triangle or edge as an inactive object not producing new offsprings.

3. Theoretical Results

Here we summarize the theoretical results of our paper [3].

3.1. The Mean Offspring Number

Let us denote by $m_{i,j}(t) = \mathbb{E}\xi_{i,j}(t)$ the expectation of the number of type j offspring of a type i mother until time t.

Proposition 3.1. *For any $t \geq 0$ we have*

$$m_{2,2}(t) = r_1 F(t), \quad m_{2,3}(t) = r_2 F(t),$$

where

$$F(t) = \int\limits_0^t (1 - L_2(s))\,\mathrm{d}s = \int\limits_0^t e^{-(b+1)s} e^{\frac{1-e^{-cs}}{c}}\,\mathrm{d}s = \frac{1}{c}\int\limits_0^{1-e^{-ct}} (1-u)^{\frac{b+1}{c}-1} e^{\frac{u}{c}}\,\mathrm{d}u.$$

$$\mathbb{E}\lambda_2 = \frac{1}{c}\int\limits_0^1 (1-u)^{\frac{b+1}{c}-1} e^{\frac{u}{c}}\,\mathrm{d}u.$$

For any $t \geq 0$ we have

$$m_{3,2}(t) = p_1 G(t), \quad m_{3,3}(t) = (p_2 + 3p_3)G(t),$$

where

$$G(t) = \int_0^t (1 - L_3(s))\, ds = \int_0^t e^{-s(b+1)} e^{\frac{3(p_1+p_2)\left(1 - e^{-cs}\right) + p_3\left(1 - e^{-3cs}\right)}{3c}}\, ds$$

$$= \frac{1}{c} \int_0^{1 - e^{-ct}} (1 - u)^{\frac{b+1}{c} - 1} e^{\frac{u}{3c}\left(p_3 u^2 - 3p_3 u + 3\right)}\, du.$$

$$\mathbb{E}\lambda_3 = \frac{1}{c} \int_0^1 (1 - u)^{\frac{b+1}{c} - 1} e^{\frac{u}{3c}\left(p_3 u^2 - 3p_3 u + 3\right)}\, du.$$

$0 < \mathbb{E}\lambda_2, \mathbb{E}\lambda_3 < \infty$ because $b \geq 0$.

Let

$$m^*_{i,j}(\kappa) = \int_0^\infty e^{-\kappa t} m_{i,j}(dt), \quad i, j = 2, 3,$$

be the Laplace transform of $m_{i,j}$.

Proposition 3.2. *For any $\kappa \geq 0$ we have*

$$m^*_{2,2}(\kappa) = r_1 A(\kappa), \quad m^*_{2,3}(\kappa) = r_2 A(\kappa),$$

where

$$A(\kappa) = \int_0^\infty e^{-\kappa s} e^{-(b+1)s} e^{\frac{1 - e^{-cs}}{c}}\, ds = \frac{1}{c} \int_0^1 (1 - u)^{\frac{\kappa + b + 1}{c} - 1} e^{\frac{u}{c}}\, du.$$

For any $\kappa \geq 0$ we have

$$m^*_{3,2}(\kappa) = p_1 B(\kappa), \quad m^*_{3,3}(\kappa) = (p_2 + 3p_3)B(\kappa),$$

where

$$B(\kappa) = \int_0^\infty e^{-\kappa s} e^{-s(b+1)} e^{\frac{3(p_1+p_2)\left(1 - e^{-cs}\right) + p_3\left(1 - e^{-3cs}\right)}{3c}}\, ds$$

$$= \frac{1}{c} \int_0^1 (1 - u)^{\frac{\kappa + b + 1}{c} - 1} e^{\frac{u}{3c}\left(p_3 u^2 - 3p_3 u + 3\right)}\, du.$$

3.2. The Perron Root and the Malthusian Parameter

Let

$$M(\kappa) = \begin{pmatrix} m_{2,2}^*(\kappa) & m_{2,3}^*(\kappa) \\ m_{3,2}^*(\kappa) & m_{3,3}^*(\kappa) \end{pmatrix}$$

be the matrix of the Laplace transforms. By direct calculation we obtain that the characteristic roots of $M(\kappa)$ are

$$\varrho_{1,2}(\kappa) = \frac{(p_2+3p_3)B(\kappa)+r_1A(\kappa)\pm\sqrt{((p_2+3p_3)B(\kappa)-r_1A(\kappa))^2+4p_1B(\kappa)r_2A(\kappa)}}{2}. \qquad (3.1)$$

The greater of these values is called the Perron root, that is

$$\varrho(\kappa) = \varrho_1(\kappa) = \frac{(p_2+3p_3)B(\kappa)+r_1A(\kappa)+\sqrt{((p_2+3p_3)B(\kappa)-r_1A(\kappa))^2+4p_1B(\kappa)r_2A(\kappa)}}{2}$$

is the Perron root.

We assume that the process is supercritical, that is $\varrho_1(0) > 1$. Now, that value of κ for which the Perron root is equal to 1 is called the Malthusian parameter. In the usual notation, α is the Malthusian parameter if $\varrho(\alpha) = 1$. We assume the existence of the Mathusian parameter. From relation $\varrho(\alpha) = 1$ and (3.1) we obtain, that for the Malthusian α we have

$$r_1A(\alpha)(p_2 + 3p_3)B(\alpha) - (r_1A(\alpha) + (p_2 + p_3)B(\alpha)) = r_2A(\alpha)p_1B(\alpha) - 1. \quad (3.2)$$

We shall need the eigenvectors of $M(\alpha)$. So let α be the Malthusian parameter and let $(v_2, v_3)^\top$ be the right eigenvector of $M(\alpha)$ satisfying condition $v_2 + v_3 = 1$. Then direct calculation shows, that

$$v_2 = \frac{(r_1 - 1)A(\alpha)}{(2r_1 - 1)A(\alpha) - 1}, \qquad v_3 = \frac{r_1A(\alpha) - 1}{(2r_1 - 1)A(\alpha) - 1}. \qquad (3.3)$$

Again let α be the Malthusian parameter and let $(u_2, u_3)^\top$ be the left eigenvector of $M(\alpha)$ satisfying condition $u_2v_2 + u_3v_3 = 1$. Direct calculation shows, that

$$u_2 = \frac{p_1B(\alpha)((2r_1-1)A(\alpha)-1)}{p_1B(\alpha)(r_1-1)A(\alpha)-(r_1A(\alpha)-1)^2}, \qquad u_3 = \frac{(1-r_1A(\alpha))((2r_1-1)A(\alpha)-1)}{p_1B(\alpha)(r_1-1)A(\alpha)-(r_1A(\alpha)-1)^2}. \qquad (3.4)$$

3.3. Results

In the following theorem we shall need the next formulae.

$$D(\alpha) = \sum_{l,j=2}^{3} u_l v_j \left(-m_{l,j}^*(\alpha)\right)'.$$

Here u_i and v_i are from equations (3.4) and (3.3). Moreover, by Proposition 3.1 or by Proposition 3.2, we have that

$$\left(-m_{2,2}^*(\alpha)\right)' = r_1\left(-A'(\alpha)\right), \quad \left(-m_{2,3}^*(\alpha)\right)' = r_2\left(-A'(\alpha)\right)$$

$$\left(-m^*_{3,2}(\alpha)\right)' = p_1\left(-B'(\alpha)\right), \quad \left(-m^*_{3,3}(\alpha)\right)' = (p_2 + 3p_3)\left(-B'(\alpha)\right),$$

where

$$-A'(\alpha) = \int\limits_0^\infty s e^{-\alpha s} e^{-(b+1)s} e^{\frac{1-e^{-cs}}{c}} \, \mathrm{d}s = -\frac{1}{c^2} \int\limits_0^1 \ln(1-u)\,(1-u)^{\frac{\alpha+b+1}{c}-1} e^{\frac{u}{c}} \, \mathrm{d}u,$$

$$-B'(\alpha) = \int\limits_0^\infty s e^{-\alpha s} e^{-s(b+1)} e^{\frac{3(p_1+p_2)\left(1-e^{-cs}\right)+p_3\left(1-e^{-3cs}\right)}{3c}} \, \mathrm{d}s$$

$$= -\frac{1}{c^2} \int\limits_0^1 \ln(1-u)\,(1-u)^{\frac{\alpha+b+1}{c}-1} e^{\frac{u}{3c}\left(p_3 u^2 - 3p_3 u + 3\right)} \, \mathrm{d}u.$$

We turn to the limit of the edges and triangles.

Proposition 3.3. *Assume that (3.2) has a finite positive solution α. Assume that $0 \le r_1 < 1$, $0 < p_1 \le 1$ and it is excluded, that both $r_1 = 0$ and $p_1 = 1$ are satisfied at the same time.*

Let $_2E(t)$ and $_3E(t)$ denote the number of all edges being born up to time t if the ancestor of the population was an edge and triangle, respectively. Then

$$\lim_{t\to\infty} e^{-\alpha t}\,_iE(t) = {_iW}\frac{v_i u_2}{\alpha D(\alpha)}$$

almost surely for $i = 2, 3$.

Let $_2\tilde{E}(t)$ and $_3\tilde{E}(t)$ denote the number of all edges alive at time t if the ancestor of the population was an edge and triangle, respectively. Then

$$\lim_{t\to\infty} e^{-\alpha t}\,_i\tilde{E}(t) = {_iW}\frac{v_i u_2 A(\alpha)}{D(\alpha)}$$

almost surely for $i = 2, 3$.

Let $_2T(t)$ and $_3T(t)$ denote the number of all triangles being born up to time t if the ancestor of the population was an edge and triangle, respectively. Then

$$\lim_{t\to\infty} e^{-\alpha t}\,_iT(t) = {_iW}\frac{v_i u_3}{\alpha D(\alpha)}$$

almost surely for $i = 2, 3$.

Let $_2\tilde{T}(t)$ and $_3\tilde{T}(t)$ denote the number of all triangles alive at time t if the ancestor of the population was an edge and triangle, respectively. Then

$$\lim_{t\to\infty} e^{-\alpha t}\,_i\tilde{T}(t) = {_iW}\frac{v_i u_3 B(\alpha)}{D(\alpha)}$$

almost surely for $i = 2, 3$.

4. Numerical and Simulation Results

4.1. The Simulation of the Model

To get a closer look on the theoretical results, we made some simulations about them. We generated our code in Julia language [2]. We chose Julia, because of the great implementation of priority queues. The simulation time of our code was significantly faster in Julia than in other programming languages. There were 4 kind of objects in our simulations: a birth event for an edge and a triangle and a death event for an edge and a triangle. We put all objects in a priority queue with the priority of its occurrence time, because we can pop out the element with the lowest priority. If it is a death event, we just modify the number of objects being born or being alive. If it is a birth event for an object we can handle its birth process with the predefined parameters b, c, r_1, r_2, p_1, p_2, p_3. In the birth process we generated an exponential time step for the next birth step of our object. After that we checked if the object is still alive by calculating the survival function. If the object is dead, we put a death event in the queue and move to the next one. If it is alive, then we generate the offspring and put them into the priority queue with the calculated birth time priorities. After this step we moved to the next event.

4.2. Support for the Theoretical Results

Our first step in the simulation was to check if the parameters have a root for Equation 3.1. In Figure 1 we see that $\varrho(\kappa) - 1$ has a root at $\kappa = 0.9133$ with the parameters $r_1 = 0.1$, $r_2 = 0.9$ $p_1 = 0.2$, $p_2 = 0.6$, $p_3 = 0.2$, $b = 0.25, c = 0.25$. This numerical value can be considered as the Malthusian parameter α.
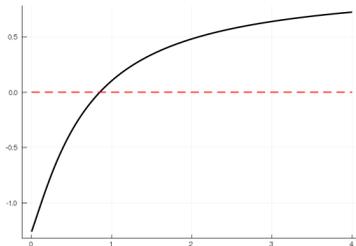

Figure 1. Searching for α with $\varrho(\kappa) - 1$.

If the conditions were given by the existence of the Malthusian parameter, we simulated $2^{20} = 1,048,576$ events and calculated the number of edges and triangles up to time t. We have to mention that all the events are processed till the end time of the simulation thanks to structure of the priority queue. For the above demonstration we used the parameter set $r_1 = 0.1$, $r_2 = 0.9$, $p_1 = 0.2$, $p_2 = 0.6$, $p_3 = 0.2$, $b = 0.25$, $c = 0.25$ again. On Figure 2 and Figure 3 the ancestor object was a triangle. Here we show $_3E(t)$ by Figure 2a, $_3\tilde{E}(t)$ by Figure 2b, $_3T(t)$

by Figure 3a and $_3\tilde{T}(t)$ by Figure 3b for the same simulation. On each plot the simulation result can be obtained by a straight line on a logarithmic scale and the theoretical result with a dashed linear line with α slope. The parallel lines show us a good support for our Proposition 3.3.
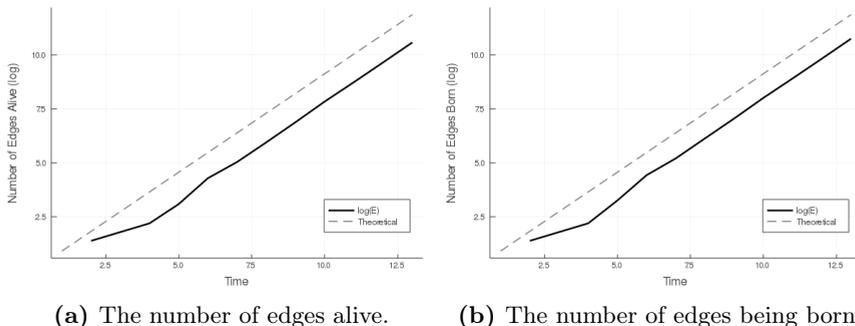

(a) The number of edges alive. **(b)** The number of edges being born.

Figure 2. Simulation results for the number of edges, where the ancestor is a triangle.

(a) The number of triangles alive. **(b)** The number of triangles being born.
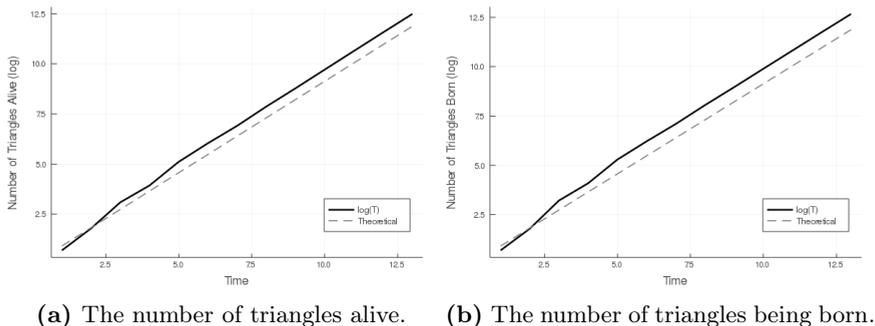
Figure 3. Simulation results for the number of triangles, where the ancestor is a triangle.

For the random variable $_iW$ of Proposition 3.3 we have got only some general information, but nothing about the distribution. Therefore we simulated 500 models with the same parameters until the slope stabilized at a large fixed t and saved the last values for the number of edges and triangles being born and being alive. After it we multiplied each result by the correct constant values of Proposition 3.3.

$$_iW \sim e^{-\alpha t}{}_iE(t)\frac{\alpha D(\alpha)}{v_i u_2} \sim e^{-\alpha t}{}_i\tilde{E}(t)\frac{D(\alpha)}{v_i u_2 A(\alpha)} \sim$$

$$\sim e^{-\alpha t}{}_iT(t)\frac{\alpha D(\alpha)}{v_i u_3} \sim e^{-\alpha t}{}_i\tilde{T}(t)\frac{D(\alpha)}{v_i u_3 B(\alpha)}$$

112

Figure 4 demonstrates the histograms of the previously described simulations with a triangle ancestor $(_3W)$ and parameters $r_1 = 0.1$, $r_2 = 0.9$, $p_1 = 0.2$, $p_2 = 0.6$, $p_3 = 0.2$, $b = 0.25$, $c = 0.25$.
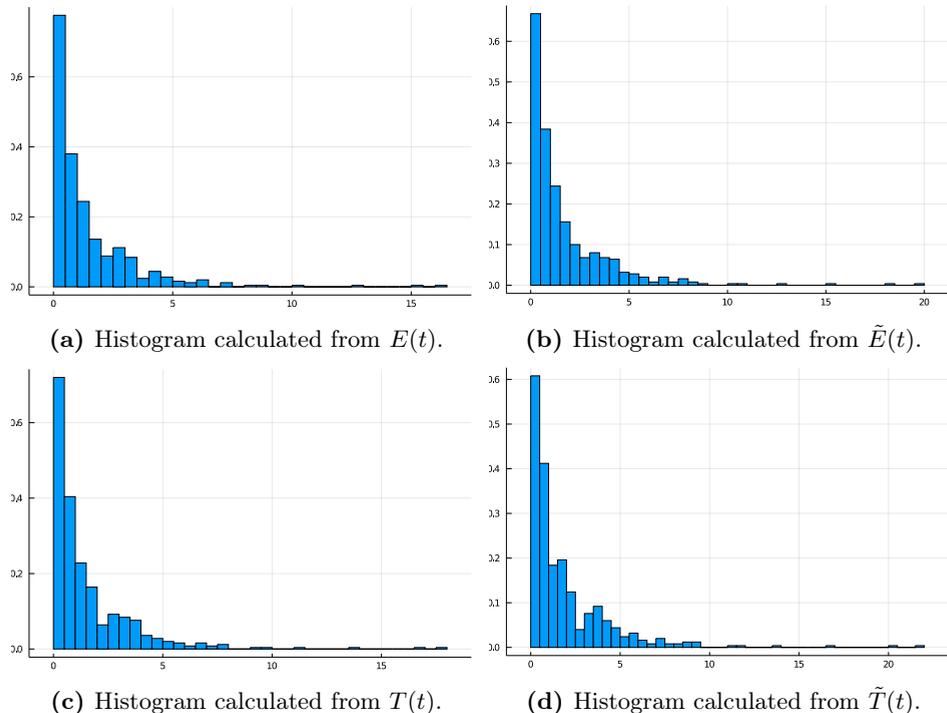

(a) Histogram calculated from $E(t)$.

(b) Histogram calculated from $\tilde{E}(t)$.

(c) Histogram calculated from $T(t)$.

(d) Histogram calculated from $\tilde{T}(t)$.

Figure 4. Histogram of $_3W$ calculated from the simulations.

It was also important to know that the experimented $_3W$ has the same distribution for the different kinds of extractions. For this purpose we made the Kolmogorov-Smirnov test for each pair of extractions. Table 1 shows us the p-values. In the table E denotes the empirical distribution of $_3W$ came from the number of edges being born, $\tilde{E}$ denotes the empirical distribution of $_3W$ came from the number of edges being alive, T denotes the empirical distribution of $_3W$ came from the number of triangles being born and $\tilde{T}$ denotes the empirical distribution of $_3W$ came from the number of edges being alive.

Table 1. p-values for the Kolmogorov-Smirnov tests.

	$\tilde{E}$	T	$\tilde{T}$
E	0.8127	0.9938	0.4676
$\tilde{E}$		0.9938	0.9938
T			0.8127

References

[1] Á. Backhausz, T. F. Móri: *A random graph model based on 3-interactions*, Ann. Univ. Sci. Budapest. Sect. Comput. 36 (2012), pp. 41–52.

[2] J. Bezanson, A. Edelman, S. Karpinski, V. Shah: *Julia: A fresh approach to numerical computing*, SIAM review 59 (2017), pp. 65–98.

[3] I. Fazekas, A. Barta: *A continuous-time network evolution model describing 2- and 3-interactions*, Manuscript (2020).

[4] I. Fazekas, A. Barta, C. Noszály, B. Porvázsnyik: *A continuous-time network evolution model describing 3-interactions*, Manuscript (2020).

[5] I. Fazekas, C. Noszály, A. Perecsényi: *The N-star network evolution model*, J. Appl. Probab. 56 (2019), pp. 416–440.

[6] I. Fazekas, B. Porvázsnyik: *Scale-free property for degrees and weights in an N-interactions random graph model*, J. Math. Sci. (N.Y.) 214 (2016), pp. 69–82.

[7] P. Jagers: *Branching Processes with Biological Applications*, Wiley (1975).

[8] T. F. Móri, S. Rokob: *A random graph model driven by time-dependent branching dynamics*, Annales Univ. Sci. Budapest., Sect. Comp. 46 (2017), pp. 191–213.