Lab 6 - CSN: Network dynamics

Pierre-Antoine Porte
porte.pierreantoine@gmail.com

Rodrigo Arias Mallo
rodarima@gmail.com

December 15, 2017

1 Introduction

In this session, we are going to generate some data by using 3 different variations of the dynamical principles in Barabasi-Albert models (BA models in the future). Those principles are: vertex growth and preferential attachment. The different variations of models we will implements are:

- **A**: Vertex growth and preferential attachment (original).
- **B**: Vertex growth and uniform random attachment.
- **C**: Suppressed growth and preferential attachment.

Those generator models were simulated and the stored data will let us to analyze mathematical properties. In this report we will show, discuss and explain the results as well as the details of the implementation.

2 Results

For each generation model (A, B and C), two metrics are analyzed, the distribution of the degrees of the nodes (D), and the evolution of a node as the graph grows over time (T). We track the evolution of 4 different nodes selected from different points in the simulation. In the figures from 16 to 15 the best models are plotted along with the data. The notation for each model has been slightly changed, to avoid confusion. All models from session 2 where renamed with a T as prefix, and those from the session 3 a D as prefix instead. Those prefixes let us distinguish between say the model $D1$ and $T1$, as one is modeling the degree distribution, and the other the degree over time. The models for the evolution of degree over time can be shown in table 1, the models for the degree distribution are used in the minimum log likelihood form, and the only change is $\gamma = 3$ in model 2 (see table 2 of session 2). The tables 2 to 6 use AIC to measure $\Delta = AIC - AIC_{best}$ of each model.

Model	Function	Parameters
T0	$f(n) = at$	a
T1	$f(t) = (t/2)^b$	b
T2	$f(t) = at^b$	a, b
T3	$f(t) = ae^{ct}$	a, c
T4	$f(t) = a \log t$	a
T5	$f(t) = at^b e^{ct}$	a, b, c
T1+	$f(t) = (t/2)^b + d$	b, d
T2+	$f(t) = at^b + d$	a, b, d
T3+	$f(t) = ae^{ct} + d$	a, c, d
T4+	$f(t) = a \log t + d$	a, d
T5+	$f(t) = at^b e^{ct} + d$	a, b, c, d

Table 1: The list of models to fit the degree over time.

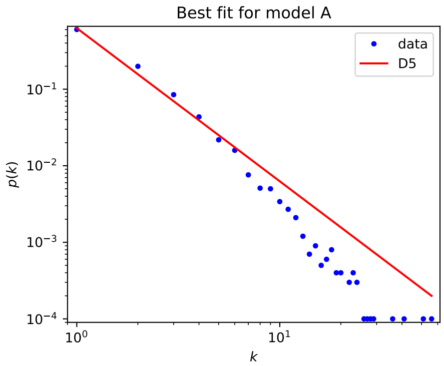

Figure 1: Distribution degree for model A

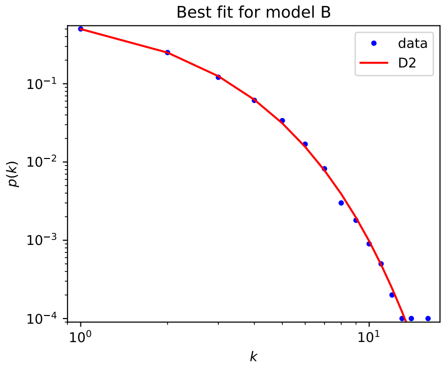

Figure 2: Distribution degree for model B

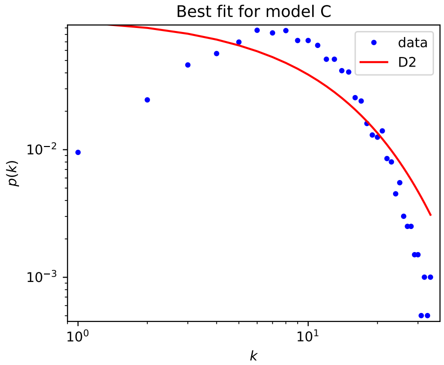

Figure 3: Distribution degree for model C

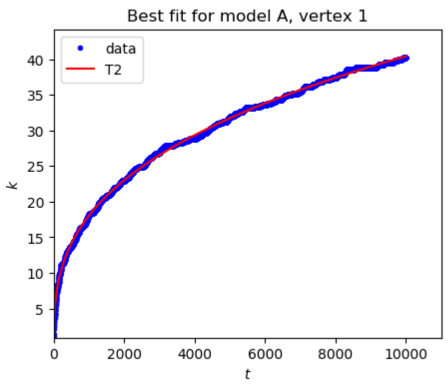

Figure 4: Degree over time for model A with vertex at $t = 1$

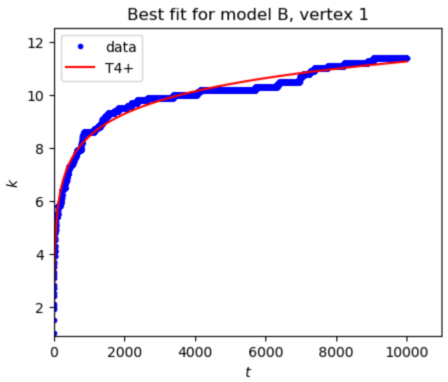

Figure 5: Degree over time for model B with vertex at $t = 1$

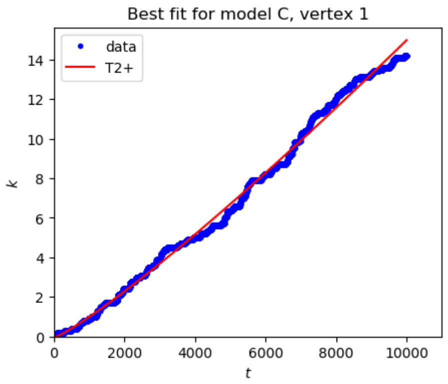

Figure 6: Degree over time for model C with vertex at $t = 1$

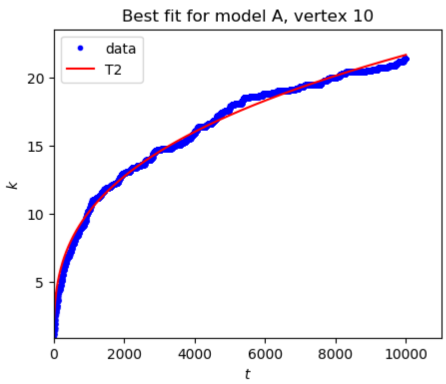

Figure 7: Degree over time for model A with vertex at $t = 10$

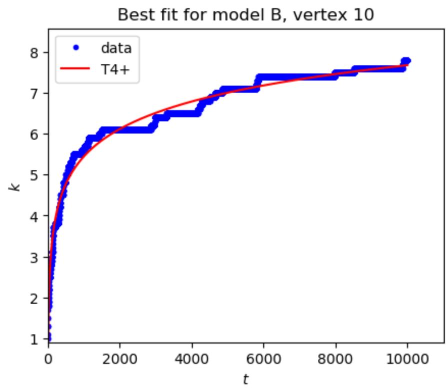

Figure 8: Degree over time for model B with vertex at $t = 10$

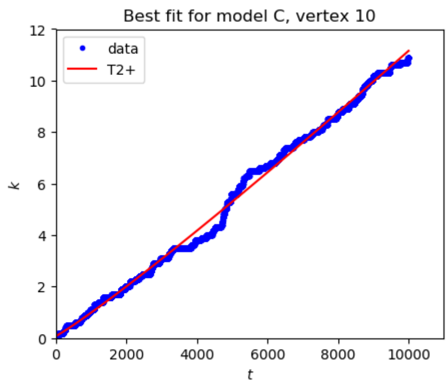

Figure 9: Degree over time for model C with vertex at $t = 10$

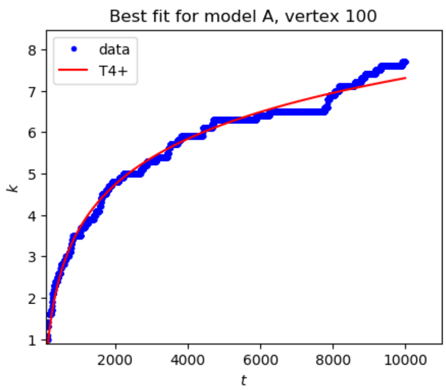

Figure 10: Degree over time for model A with vertex at $t = 100$

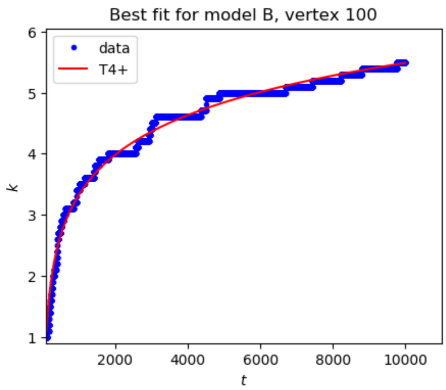

Figure 11: Degree over time for model B with vertex at $t = 100$

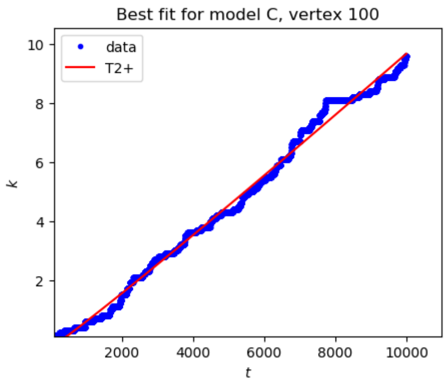

Figure 12: Degree over time for model C with vertex at $t = 100$

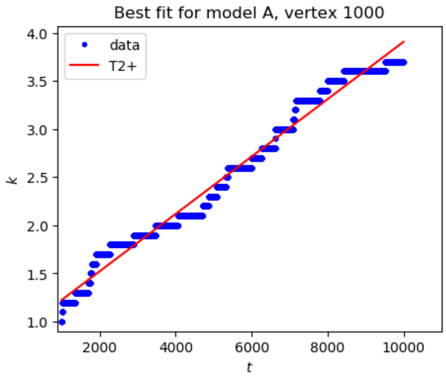

Figure 13: Degree over time for model A with vertex at $t = 10000$

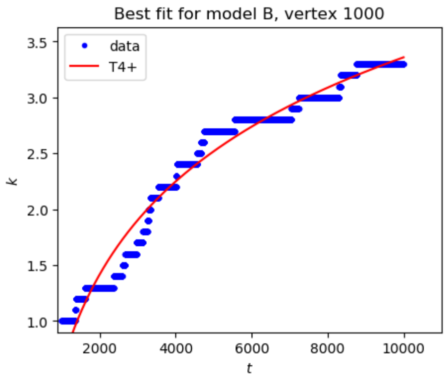

Figure 14: Degree over time for model B with vertex at $t = 1000$

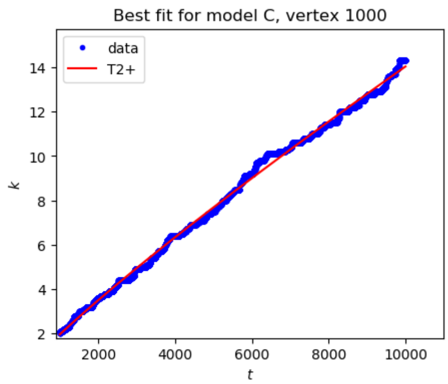

Figure 15: Degree over time for model C with vertex at $t = 1000$

3 Discussion

We used m_0 and n_0 always with the same values (which can be however different for each model). We never compared the models with different m_0 or n_0 for the same model. It could have indicate us if the model behave differently given different graph in input. We could have went further and do this but we preferred to focus on the analysis of our different models with the input specified in the methods section.

3.1 Model A

Vertex degree over time We checked if the power-law dependency with 1/2 exponent gives the best fit to all the time series as required by the statement. This power-law dependency is not the best fit. This power-law is model T1, which has an AIC higher than model T2 (see tables 3 to 6). In fact the best model is model T2, defined as $f(t) = at^b$, which is the exponential growth. It makes sense because with one free parameter we can better optimize this model to fit the data, by adapting b to somewhat close to 1/2 as expected in model T1, however we may be overestimating the data. If we look at table 8 to 11, we see that $b \approx 0.5$ only for $t_i = 1000$ (table 11).

3.2 Model B

Vertex degree over time For table 3 to 6, best model is as suggested by the statement, the logarithmic model T4. Indeed, we have 0 as Δ for the vertex degree over time for this model and for every t_i chosen to look at the vertex.

Degree distribution The best model for the degree distribution for the Barabasi Albert model without preferential attachment is the geometric one (model D1). Actually only the variant of the geometric model is good (model D1+), as we can see in table 2. The model D1 has a Δ too high. However by replacing correctly the generated data from model D1, we now have a Δ of 0. Thus, the geometric model variant is the best model.

3.3 Model C

Vertex degree over time As expected by statement, the degree over time should fit a linear scale. By computing the AIC and making the Δ (see tables 3 to 6) we have seen that the linear model was good. Also, we are really confident by saying it's linear when we look at the plot generated for model C, for every t_is. However we find that model T0, T0+, T2 and T2+ are the best. Model T2 is represented as at^b, as b is close to 1 as we can see in tables 3 to 6, which explain this good fit for model T2 and T2+.

Degree distribution As stated by the statement, the degree distribution for this model should be closer to a binomial distribution. Indeed it looks like it, but we found out it looked even more of a displaced Poisson with on a different scale. If we took $\lambda = 2$, then we would have a faster increasing and decreasing Poisson which is what we want. However, this Poisson has a mean of $\lambda = 2$. Therefore, it does not fit our data which has a mean of approximately 10. So we displaced the Poisson and adjusted the scale. The lasting problem was that due do this scale and displacement, our model never produces data ≈ 0 whereas the generated model data had a lot of values ≈ 0.

In the end we chose to use a Poisson distribution as we did in session 2. It's not a really good fit but it's still the best fit we have. You can see this plot in figure

We also made sure that the distribution giving the best fit in not a power-law, it's looking more of a Gaussian one. However since it's not symmetric it was hard to model the data using a normal law, for example.

4 Methods

For the model A and B the initial graph was an empty graph with only one vertex. For C we used an unconnected graph with $t_{\max}$ vertices. Because we have no vertex growth, the vertices are not increasing and $n_0 = n_{tmax}$. For the three models we used $m_0 = 1$ as the initial number of edges added at each step, changing this parameter can affect the final results, but was not tested.

We measured the growth of the vertex degree over time and the degree distribution for each model. The vertex degree was measured over the time for $t_i \in 1, 10, 100, 1000$ successively.

We used python for generating the models, to store the results and to analyze the data. For each BA model with the letter M a folder in `data/model`M`/` contains all the results produced from this model. Inside, the degree sequence is stored in the file `dseq.txt`, the degree distribution in `dd.txt` and for each T in the arrival time, we produced `dt_`t_i`.txt` tracing the degree of the vertex arriving at time t.

4.1 Generating model C

While trying to define how to make the preferential attachment for model C we faced a problem. We were choosing from the edges to be linked to our vertex in an array of probability p (we didn't used the stubs method), with the degree of the node over the sum of the degrees of all nodes as:

```
p[i] = vertex.degree() / sum(graph.all_degrees())
```

However with $m0 = 5$ and only 4 vertices connected, we cannot choose 5 vertices, and it would not work. To mimic the stub solution, we added one virtual degree to each unconnected node. Therefore our array of probability p was computed as:

```
if vertex.degree == 0:
    p[i] = 1 / sum(graph.all_degrees()) + sum(number_of_nodes_with_degree_0)
else:
    p[i] = vertex.degree() / (sum(graph.all_degrees()) +
        sum(number_of_nodes_with_degree_0))
```

In the end the degree was represented with the number of stubs.

A Tables

	A	B	C
D1	7245.741	1308.885	776.040
D2	1097.608	0.000	0.000
D3	606.012	2505.972	6112.841
D4	359.648	2456.367	6110.839
D5	0.000	1899.960	5988.746

Table 2: Δ for the degree distribution.

	A	B	C
T0	69 534.390	59 969.865	10 786.775
T1	45 188.426	48 383.218	38 204.833
T2	0.000	4088.919	33.872
T3	57 486.358	23 686.609	65 632.205
T4	57 130.622	10.271	49 999.142
T0+	45 767.062	22 835.391	4019.623
T1+	27 730.356	25 665.773	26 144.944
T2+	9832.537	10 660.418	0.000
T3+	∞	∞	∞
T4+	51 900.497	0.000	45 217.274

Table 3: Δ for the vertex degree over time for $t_i = 1$.

	A	B	C
T0	50 581.204	56 957.191	3294.824
T1	28 319.859	44 284.884	35 679.447
T2	0.000	4270.384	77.463
T3	75 169.798	75 611.683	65 467.831
T4	36 243.653	5879.067	48 564.528
T0+	27 779.459	22 417.994	1463.430
T1+	12 283.879	15 052.096	23 894.852
T2+	1779.080	9451.757	0.000
T3+	∞	∞	∞
T4+	24 327.113	0.000	41 655.822

Table 4: Δ for the vertex degree over time for $t_i = 10$.

	A	B	C
T0	45 373.458	56 667.047	6775.516
T1	25 127.170	41 502.234	34 445.263
T2	2465.286	11 882.066	1083.024
T3	68 989.828	77 000.991	62 248.713
T4	28 472.120	25 393.086	46 399.114
T0+	20 393.046	28 456.942	343.071
T1+	9261.143	20 574.043	20 084.117
T2+	1800.322	14 737.800	0.000
T3+	∞	∞	∞
T4+	0.000	0.000	38 696.113

Table 5: Δ for the vertex degree over time for $t_i = 100$.

	A	B	C
T0	24 212.569	23 931.847	15 817.628
T1	8639.541	3786.391	34 773.465
T2	4130.729	3099.317	41.092
T3	58 036.196	54 148.007	69 546.084
T4	31 443.880	26 504.386	49 306.819
T0+	0.021	9269.596	2438.358
T1+	5672.115	2499.846	13 376.474
T2+	0.000	238.319	0.000
T3+	∞	∞	∞
T4+	13 880.155	0.000	44 696.015

Table 6: Δ for the vertex degree over time for $t_i = 1000$.

Param	A	B	C
$D1$, λ	1.593	1.593	10.000
$D2$, q	0.500	0.500	0.100
$D4$, γ	2.189	2.078	2.000
$D5$, γ	2.000	2.000	2.000
$D5$, $k_{\max}$	20.000	20.000	20.000

Table 7: Parameters for degree distribution models fitting.

Param	A	B	C
$T0$, a	0.005	0.002	0.001
$T1$, a	0.434	0.134	0.110
$T2$, a	1.616	3.348	0.000
$T2$, b	0.350	0.133	1.166
$T3$, a	14.000	8.637	0.000
$T3$, c	0.000	0.000	-0.527
$T4$, a	3.792	1.222	0.837
$T4$, d_1	-0.849	1.117	-0.871
$T0+$, a	0.003	0.000	0.002
$T0+$, d	14.000	8.496	-0.735
$T1+$, a	0.370	0.072	0.178
$T1+$, d	5.000	5.000	-5.000
$T2+$, a	0.800	0.394	0.000
$T2+$, b	0.416	0.300	1.155
$T2+$, d	3.969	5.251	-0.058
$T3+$, a	0.289	0.289	0.289
$T3+$, c	0.853	0.853	0.853
$T3+$, d	-10.000	-10.000	-10.000
$T4+$, a	4.955	1.230	2.020
$T4+$, d_1	-0.552	1.413	7.886
$T4+$, d_2	-10.000	-0.060	-10.000

Table 8: Parameters for the vertex degree over time for $t_i = 1$.

Param	A	B	C
$T0$, a	0.003	0.001	0.001
$T1$, a	0.237	0.093	0.087
$T2$, a	1.036	1.795	0.001
$T2$, b	0.330	0.159	1.070
$T3$, a	−0.330	−0.331	−0.352
$T3$, c	−1.363	−1.363	−0.566
$T4$, a	2.079	0.820	0.693
$T4$, d_1	−9.843	−9.408	−9.875
$T0+$, a	0.001	0.000	0.001
$T0+$, d	10.366	5.424	−0.220
$T1+$, a	0.172	0.036	0.139
$T1+$, d	4.885	4.303	−3.958
$T2+$, a	0.800	0.314	0.000
$T2+$, b	0.355	0.300	1.088
$T2+$, d	0.798	2.893	0.076
$T3+$, a	0.289	0.289	0.289
$T3+$, c	0.853	0.853	0.853
$T3+$, d	−10.000	−10.000	−10.000
$T4+$, a	3.238	0.970	1.874
$T4+$, d_1	−8.242	−1.528	11.972
$T4+$, d_2	−10.000	−1.258	−10.000

Table 9: Parameters for the vertex degree over time for $t_i = 10$.

Param	A	B	C
$T0$, a	0.001	0.001	0.001
$T1$, a	0.083	0.064	0.074
$T2$, a	0.422	0.735	0.000
$T2$, b	0.313	0.220	1.122
$T3$, a	-0.500	-0.500	-0.500
$T3$, c	-0.500	-0.500	-0.500
$T4$, a	0.718	0.563	0.568
$T4$, d_1	-15.000	-15.000	-15.000
$T0+$, a	0.000	0.000	0.001
$T0+$, d	3.674	3.410	-0.463
$T1+$, a	0.059	0.033	0.128
$T1+$, d	1.864	2.366	-4.056
$T2+$, a	0.498	0.289	0.001
$T2+$, b	0.300	0.300	1.047
$T2+$, d	-0.328	1.054	-0.304
$T3+$, a	0.289	0.289	0.289
$T3+$, c	0.853	0.853	0.853
$T3+$, d	-10.000	-10.000	-10.000
$T4+$, a	1.614	0.931	1.757
$T4+$, d_1	50.000	-10.000	-10.000
$T4+$, d_2	-7.569	-3.089	-10.000

Table 10: Parameters for the vertex degree over time for $t_i = 100$.

Param	A	B	C
$T0$, a	0.000	0.000	0.001
$T1$, a	0.036	0.034	0.120
$T2$, a	0.016	0.025	0.005
$T2$, b	0.596	0.537	0.864
$T3$, a	-0.500	-0.500	-0.500
$T3$, c	-0.500	-0.500	-0.500
$T4$, a	0.306	0.300	1.013
$T4$, d_1	-15.000	-15.000	-15.000
$T0+$, a	0.000	0.000	0.001
$T0+$, d	0.925	1.003	0.928
$T1+$, a	0.041	0.037	0.183
$T1+$, d	-0.360	-0.224	-4.921
$T2+$, a	0.000	0.334	0.004
$T2+$, b	1.001	0.300	0.880
$T2+$, d	0.926	-1.835	0.132
$T3+$, a	0.289	0.289	0.289
$T3+$, c	0.853	0.853	0.853
$T3+$, d	-10.000	-10.000	-10.000
$T4+$, a	1.293	1.222	2.171
$T4+$, d_1	50.000	50.000	-10.000
$T4+$, d_2	-8.397	-7.904	-10.000

Table 11: Parameters for the vertex degree over time for $t_i = 1000$.

B Figures with all models

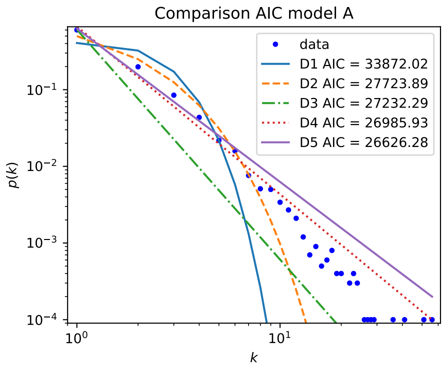

Figure 16: Distribution degree for model A

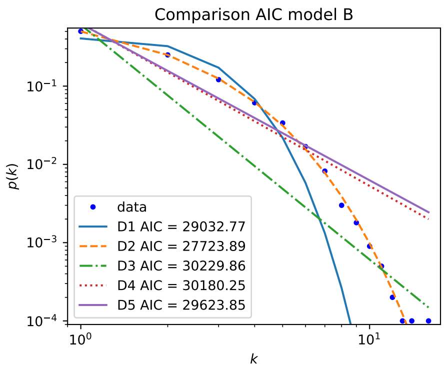

Figure 17: Distribution degree for model B

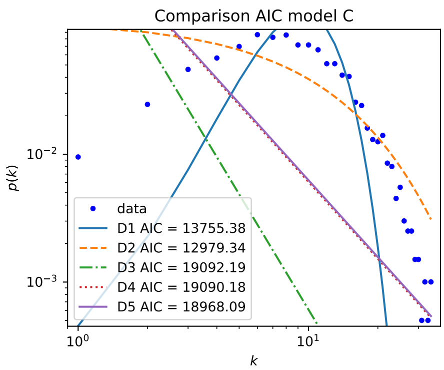

Figure 18: Distribution degree for model C

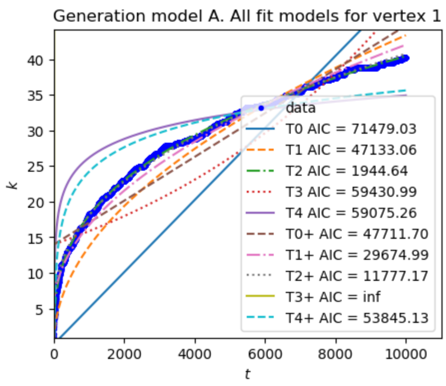

Figure 19: Degree over time for model A with vertex at $t = 1$

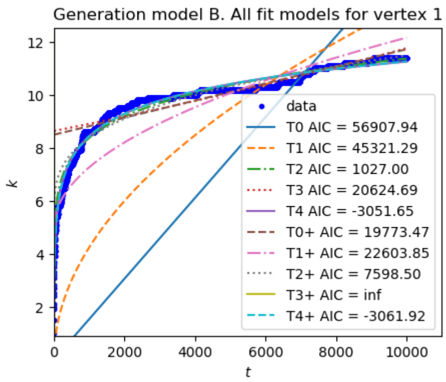

Figure 20: Degree over time for model B with vertex at $t = 1$

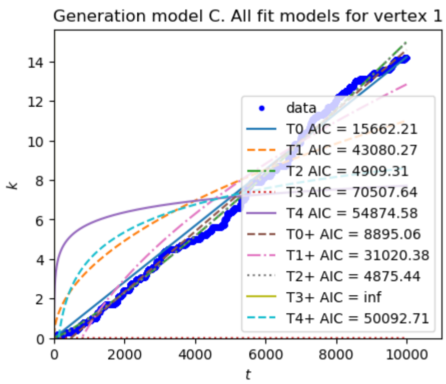

Figure 21: Degree over time for model C with vertex at $t = 1$

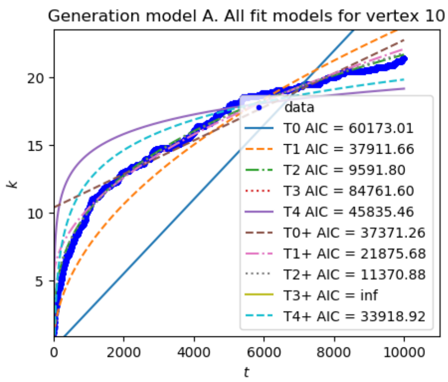

Figure 22: Degree over time for model A with vertex at $t = 10$

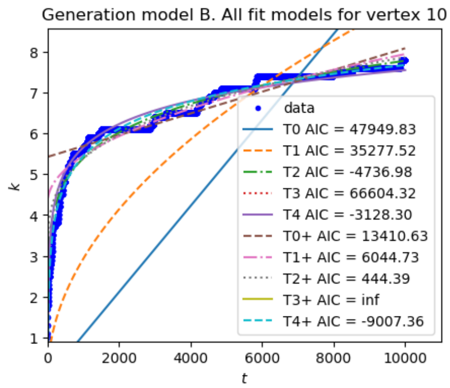

Figure 23: Degree over time for model B with vertex at $t = 10$

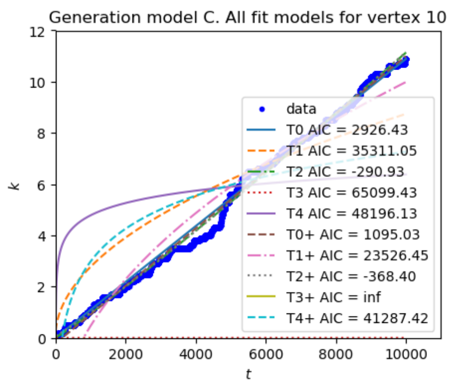

Figure 24: Degree over time for model C with vertex at $t = 10$

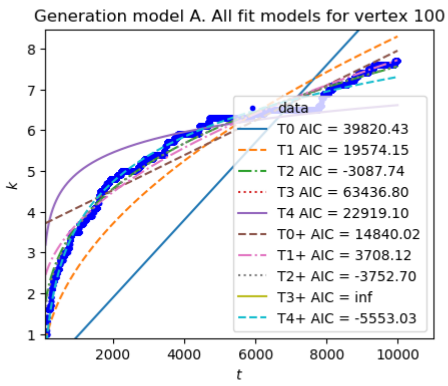

Figure 25: Degree over time for model A with vertex at $t = 100$

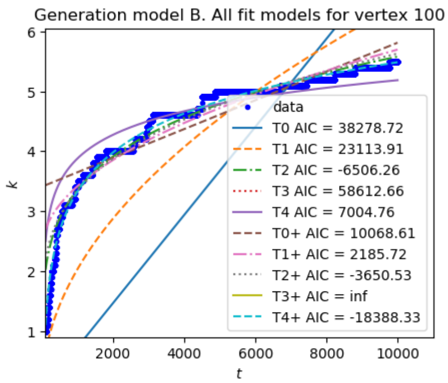

Figure 26: Degree over time for model B with vertex at $t = 100$

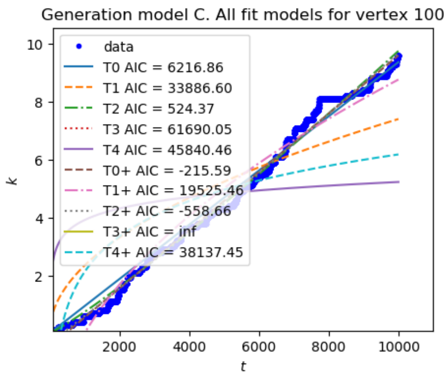

Figure 27: Degree over time for model C with vertex at $t = 100$

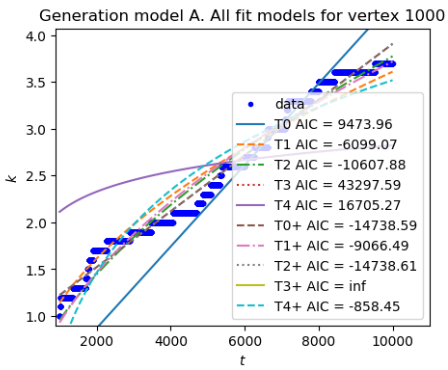

Figure 28: Degree over time for model A with vertex at $t = 10000$

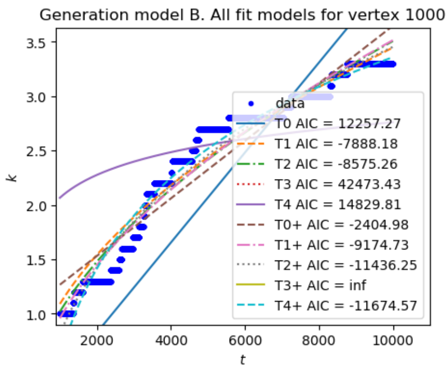

Figure 29: Degree over time for model B with vertex at $t = 1000$

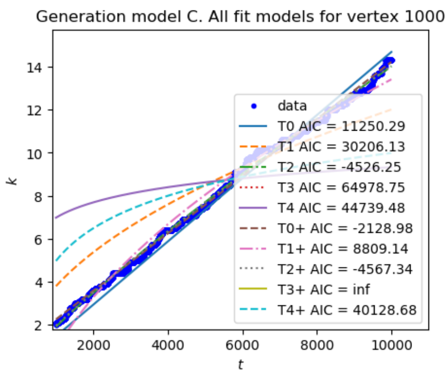

Figure 30: Degree over time for model C with vertex at $t = 1000$