

MASTER IN DATA SCIENCE

Complex and Social Networks Laboratory

Network dynamics

Raffaele D'Agostino

raffaele.d.agostino@estudiantat.upc.edu

Giulia Pincioli

giulia.pincioli@estudiantat.upc.edu

Project Report

Academic Year 2025/2026

December 4, 2025

Contents

1	Introduction	2
2	Results	2
2.1	BA with Preferential Attachment	2
2.1.1	Scaling of $k_i(t)$	2
2.1.2	Rescaled Data Collapse	3
2.1.3	Model Selection	3
2.1.4	Degree Distribution	4
2.2	BA with Random Attachment	5
2.2.1	Scaling of $k_i(t)$	5
2.2.2	Rescaled Data Collapse	5
2.2.3	Best Model Fits	6
2.2.4	Degree Distribution	7
2.3	No-Growth Preferential Attachment	8
2.3.1	Time evolution of $k_i(t)$	8
2.3.2	Model Selection and Best Model Fit	8
2.3.3	Degree Distribution	8
3	Discussion	9
3.1	BA with Preferential Attachment	9
3.1.1	Scaling of $k_i(t)$	9
3.1.2	Rescaled Data Collapse	10
3.1.3	Model Selection	10
3.1.4	Degree Distribution	11
3.2	BA with Random Attachment	12
3.2.1	Scaling of $k_i(t)$	12
3.2.2	Rescaled Data Collapse	12
3.2.3	Model Selection	13
3.2.4	Degree Distribution	13
3.3	No-Growth Preferential Attachment	14
3.3.1	Time evolution of $k_i(t)$	14
3.3.2	Model Selection	14
3.3.3	Degree Distribution	15
3.4	Conclusions	15
4	Methods	16
4.1	Networks initialization	16
4.2	Models algorithms	16
4.3	Procedure of experiments	18
4.4	On the fitting of degree distributions	18
4.5	The discrete Gaussian model	19

1 Introduction

This lab session focuses on understanding the dynamical principles behind the Barabási–Albert (BA) model and its variants through simulation and statistical analysis. The goals are to deepen understanding of network growth mechanisms, improve simulation skills, and apply the model selection techniques introduced in lab session 2 and 4.

The standard BA model is defined by two mechanisms: vertex growth, where one new vertex is added at each time step, and preferential attachment, where the probability of connecting to an existing vertex is proportional to its degree. The model depends on the parameters n_0 (initial number of vertices) and m_0 (number of edges created by each new vertex), and the initial configuration can affect the evolution of the network.

We analyse three versions of the model:

1. the original BA model (growth + preferential attachment)
2. a variant where preferential attachment is replaced by random attachment, and
3. a variant with no vertex growth but preferential attachment.

These combinations allow us to isolate the effect of each mechanism on network structure.

For each model, we study (i) how vertex degrees grow over time and (ii) the degree distribution at a final time $t_{max} = 10^5$. Required outputs include the degree sequence at t_{max} and four time series for vertices arriving at times 1, 10, 100, and 1000.

To evaluate theoretical predictions, we perform model selection by fitting several candidate growth functions (linear, power-law, exponential, logarithmic, and variants with additive constants) using non-linear regression and comparing them through AIC. This framework allows us to determine which functional forms best describe the simulated degree evolution in each model.

2 Results

2.1 BA with Preferential Attachment

2.1.1 Scaling of $k_i(t)$

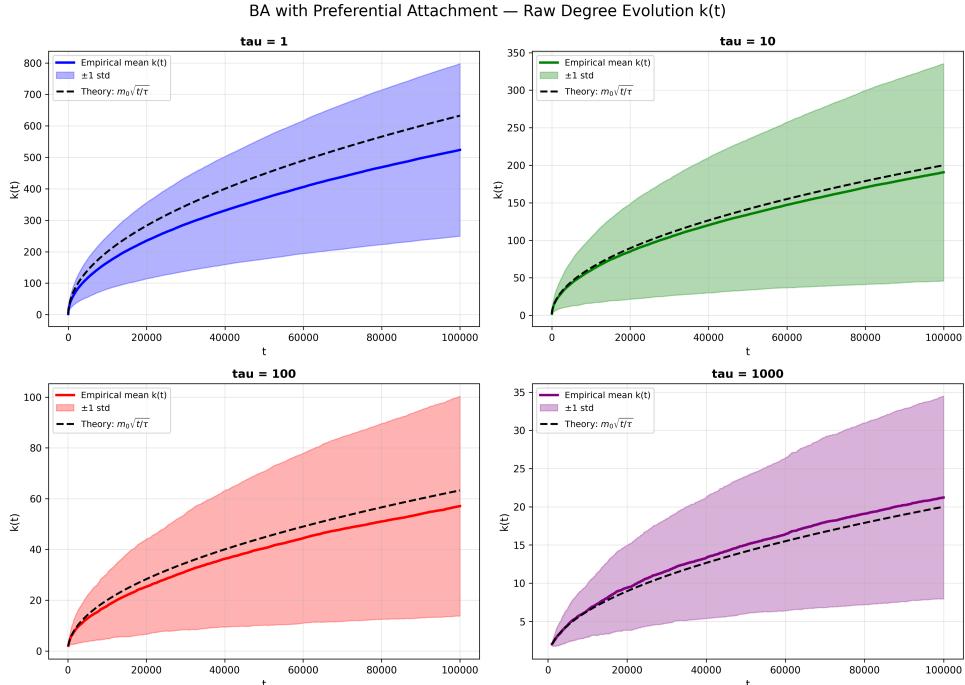


Figure 1: Empirical mean degree trajectories $k(t)$ with ± 1 standard deviation for vertices arriving at $\tau = 1, 10, 100$, and 1000 , shown together with the theoretical mean-field prediction $m_0\sqrt{t/\tau}$.

Figure 1 displays the evolution of the degree $k(t)$ for vertices arriving at four different times ($\tau = 1, 10, 100$, and 1000). For each arrival time, the figure shows the empirical mean degree across all simulation runs, together with the corresponding ± 1 standard deviation band. The theoretical mean-field prediction $m_0 \sqrt{\frac{t}{\tau}}$ is plotted alongside the empirical curves for reference. The trajectories for different arrival times are presented in separate panels, each covering the full time interval from the vertex's arrival up to t_{max} .

2.1.2 Rescaled Data Collapse

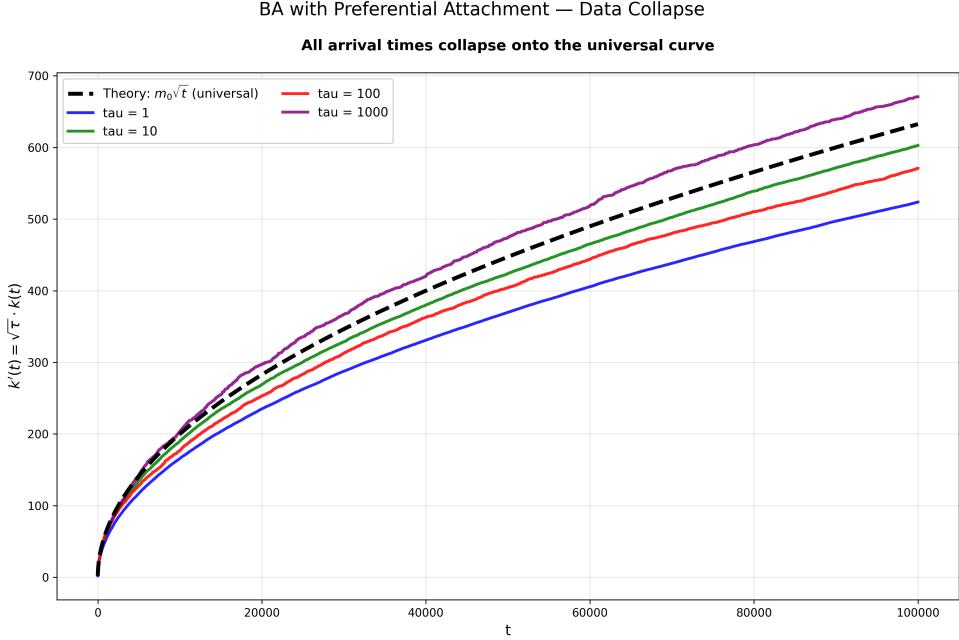


Figure 2: Rescaled degree trajectories $k'(t) = \sqrt{\tau} k(t)$ for vertices arriving at $\tau = 1, 10, 100$, and 1000 , plotted together with the universal mean-field prediction $m_0 \sqrt{t}$.

Figure 2 shows the rescaled degree trajectories obtained by applying the transformation $k' = \sqrt{\tau} k(t)$ to the empirical averages for the four arrival times considered. The figure displays the complete time series for each τ , along with the theoretical universal curve $m_0 \sqrt{t}$. All curves are plotted on the same axes to allow direct comparison of their shapes and relative alignment. The rescaled trajectories for $\tau = 1, 10, 100$ and 1000 cover the interval from their respective arrival times up to t_{max} , illustrating how the rescaling modifies the relative positions of the degree curves.

2.1.3 Model Selection

Table 1: Model selection summary for the BA preferential attachment model, showing the best-fitting model and its estimated parameters for each arrival time τ .

τ	Best Model	AIC	RMSE	a	b	d
1	Model 2+	87174.43	0.374	1.644 ± 0.00039	0.5005 ± 0.000019	0.844 ± 0.011
10	Model 2+	-11449.29	0.228	0.5686 ± 0.00023	0.5045 ± 0.000032	0.882 ± 0.0064
100	Model 2+	-44319.89	0.193	0.1818 ± 0.00021	0.4995 ± 0.000093	-0.0407 ± 0.0057
1000	Model 2+	-245978.81	0.069	0.07472 ± 0.000095	0.4929 ± 0.00010	-0.453 ± 0.0025

Table 1 reports the AIC values, RMSE values, and estimated parameters for the winning model at each τ . In all four cases, Model 2+ provides the best fit to the observed data according to the AIC criterion.

Figure 3 presents the best-fitting model for each arrival time based on the AIC criterion. For each τ , the empirical mean degree trajectory is plotted together with the curve corresponding to the selected model over the interval $t \in [\tau, t_{max}]$.

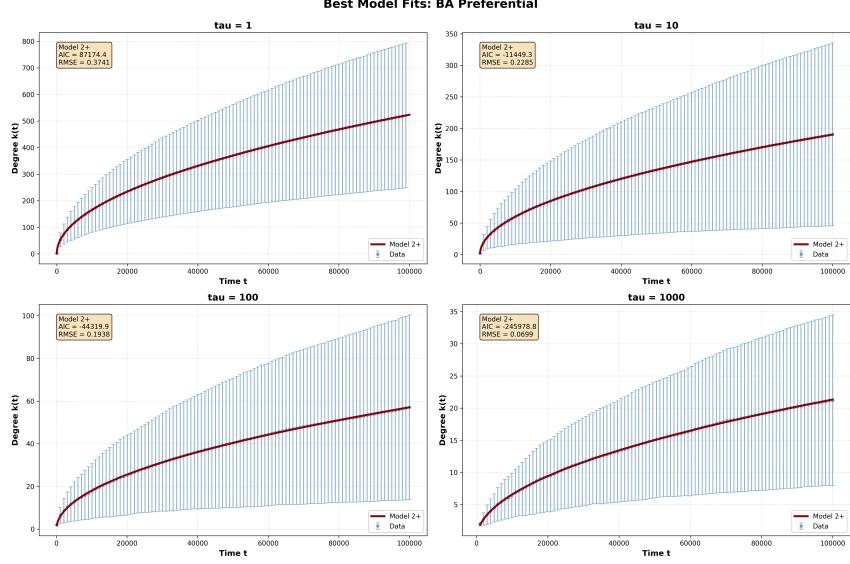


Figure 3: Best-fitting curves for the BA preferential attachment model at $\tau = 1, 10, 100$, and 1000 , shown together with the empirical mean degree trajectories.

2.1.4 Degree Distribution

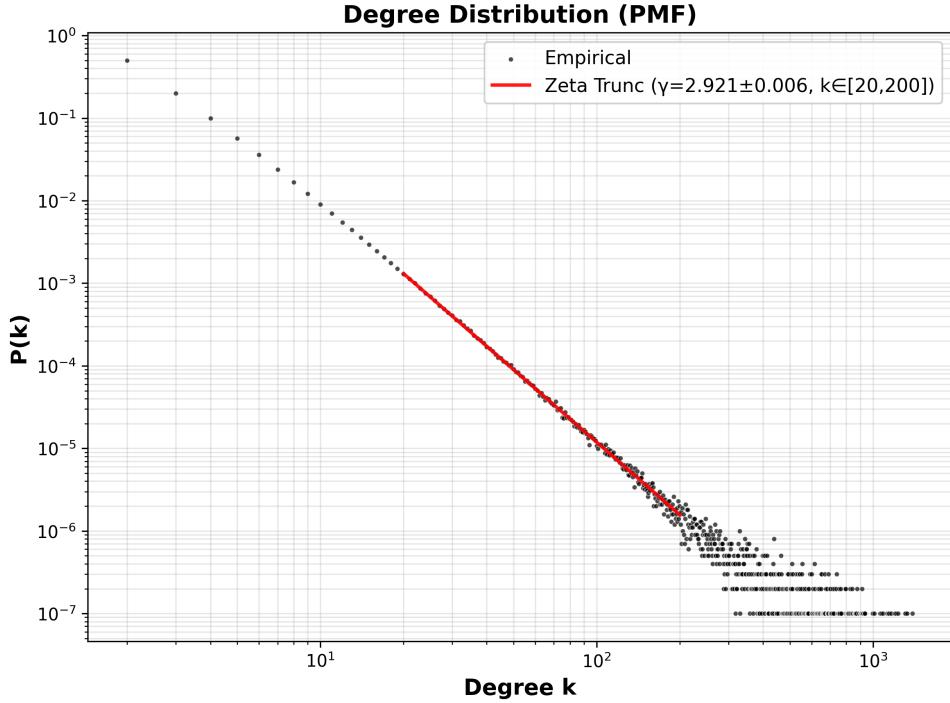


Figure 4: Empirical degree distribution at t_{\max} together with the fitted truncated zeta distribution on the interval $k \in [20, 200]$. The scale of the axis is log-log in order to highlight the power law.

Figure 4 shows the empirical probability mass function of the degree distribution at the final simulation time on logarithmic axes. Each point represents the relative frequency of nodes with degree k . The red curve represents the truncated zeta distribution fitted on the selected interval $k \in [20, 50]$. The plot displays both the empirical tail of the distribution and the fitted model evaluated over the same k range.

Table 2 reports the maximum-likelihood parameter estimates and associated AIC scores for several candidate models. The table includes Poisson, geometric, truncated geometric, zeta, truncated zeta, and Gaussian models. For each model, the estimated parameters, AIC value, and ΔAIC relative to the best

Table 2: Maximum-likelihood estimates and AIC values for several candidate models fitted to the empirical degree distribution over the interval $k \in [20, 200]$.

Model	Parameters	AIC	ΔAIC
poisson	$\lambda = 36.3139 \pm 0.0160$	2329384.18	1269566.99
displaced geometric	$q = 0.9717 \pm 0.0001$	1286145.16	226327.98
geometric trunc	$q = 0.9422 \pm 0.0001, k \leq 200$	1080382.56	20565.37
zeta	$\gamma = 2.92341 \pm 0.0051, k_{\min} = 20$	1097033.83	37216.64
zeta trunc	$\gamma = 2.9210 \pm 0.0058, k \in [20, 200]$	1059817.18	0.00
gaussian	$\mu = 20.00 \pm 0.14, \sigma = 29.04 \pm 0.06$	1159300.38	99483.19

model are listed.

2.2 BA with Random Attachment

2.2.1 Scaling of $k_i(t)$

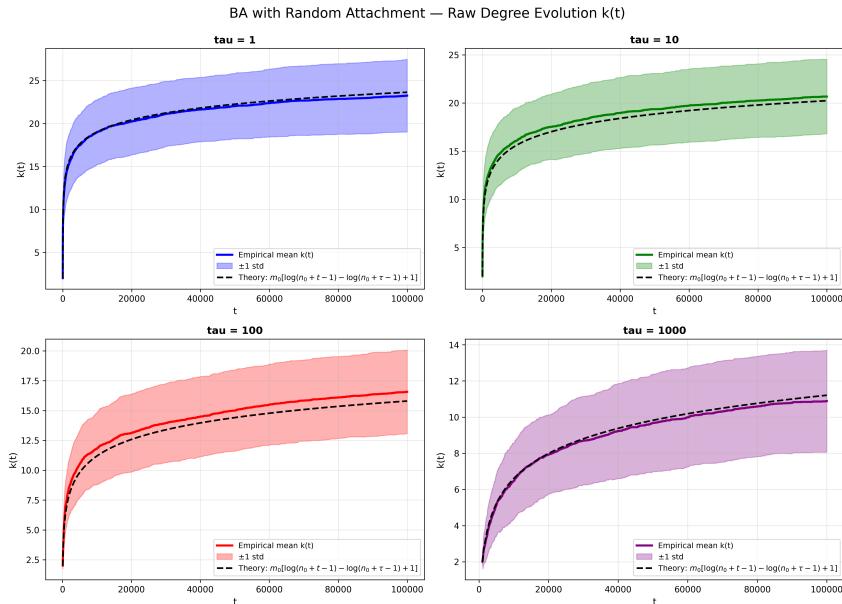


Figure 5: Empirical degree trajectories $k(t)$ with ± 1 standard deviation for vertices arriving at $\tau = 1, 10, 100$, and 1000 , plotted together with the theoretical curve $m_0[\log(n_0 + t - 1) - \log(n_0 + \tau - 1) + 1]$.

Figure 5 displays the empirical mean degree evolution for vertices arriving at $\tau = 1, 10, 100$, and 1000 in the BA model with random attachment. For each arrival time, the figure includes the mean degree $k(t)$ across simulation runs together with the corresponding ± 1 standard deviation band. The theoretical expression $m_0[\log(n_0 + t - 1) - \log(n_0 + \tau - 1) + 1]$ is plotted alongside the empirical curves for reference. Each panel shows the complete recorded trajectory for a single arrival time, using consistent axes to allow visual comparison across the four cases.

2.2.2 Rescaled Data Collapse

Figure 6 shows the rescaled degree trajectories for the random-attachment model. For each arrival time $\tau = 1, 10, 100$, and 1000 , the empirical time series $k_i(t)$ is transformed using the expression $k'i(t) = k_i(t) + m_0 \log(n_0 + \tau - 1) - m_0$. The transformed curves are plotted on the same axes together with the theoretical expression $m_0 \log(n_0 + t - 1)$, allowing direct comparison of the rescaled trajectories across arrival times. All available time points from $t = \tau$ to t_{\max} are shown.

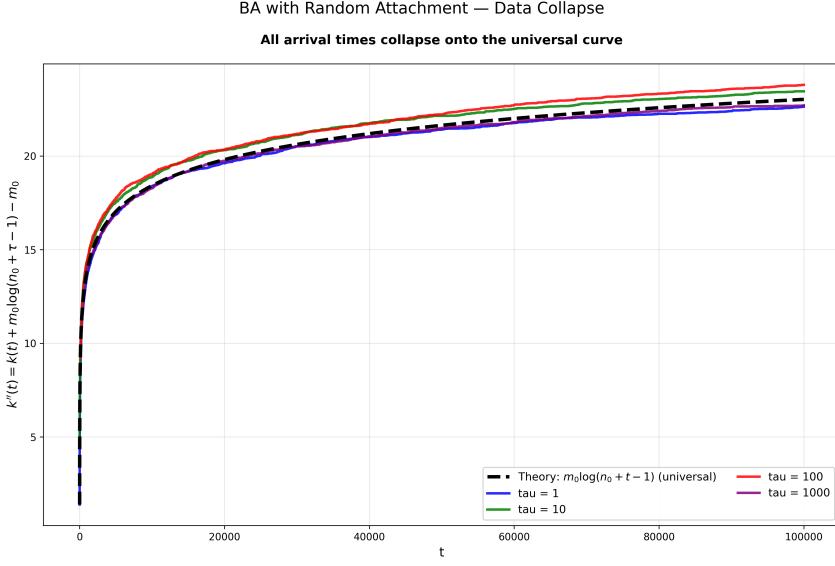


Figure 6: Rescaled degree trajectories $k'_i(t) = k_i(t) + m_0 \log(n_0 + \tau - 1) - m_0$ for vertices arriving at $\tau = 1, 10, 100$, and 1000 , plotted together with the universal theoretical curve $m_0 \log(n_0 + t - 1)$.

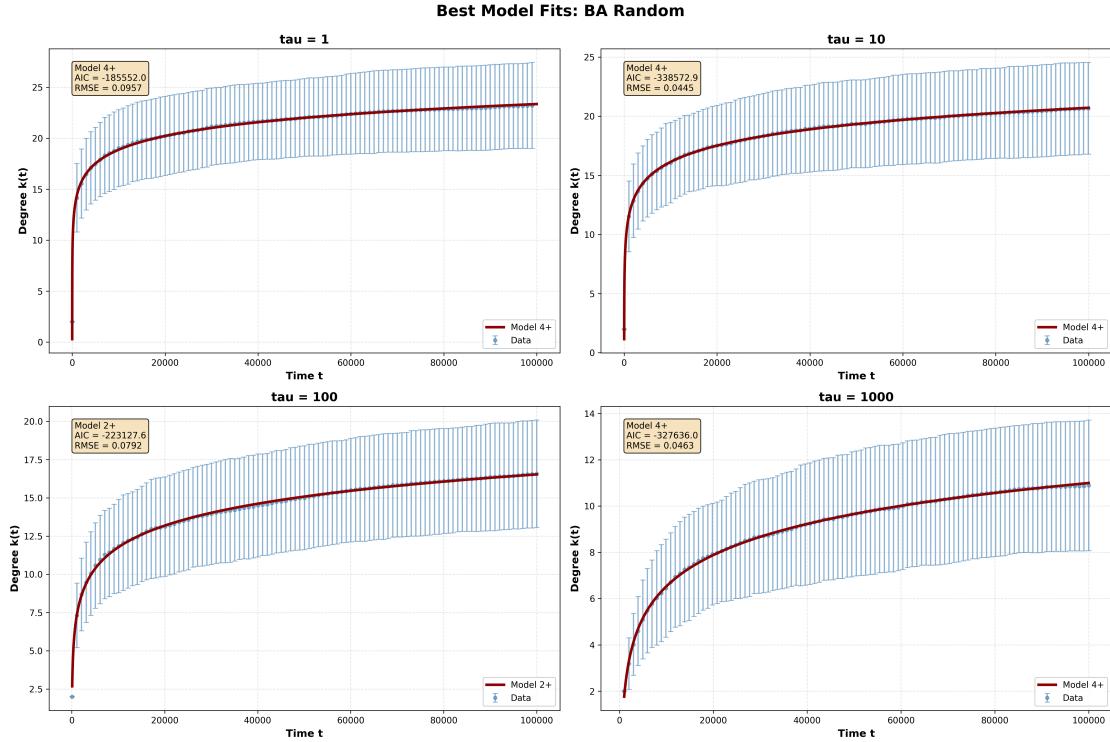


Figure 7: Best-fitting models for the BA random attachment model at $\tau = 1, 10, 100$, and 1000 . For each arrival time, the empirical mean degree trajectory is shown together with the model selected by AIC.

2.2.3 Best Model Fits

Figure 7 presents the best-fitting model for each arrival time according to the AIC criterion. Each panel shows the empirical mean trajectory of $k(t)$ together with the corresponding fitted model over the interval $t \geq \tau$. The four arrival times $\tau = 1, 10, 100$, and 1000 are displayed separately. For each case, the fitted curve and its AIC and RMSE values are shown as reported by the model-selection procedure.

Table 3 summarizes the outcomes of the model-selection procedure for the BA random attachment model. For each arrival time τ , the table lists the model with the lowest AIC value, together with its RMSE and estimated parameters. The table includes the fitted coefficients and standard errors for all

Table 3: Model selection results for the BA random attachment model. For each arrival time τ , the table reports the best-fitting model (together with its AIC, RMSE, and estimated parameters).

τ	Best Model	AIC	RMSE	a	b	d (or d_1)	d_2
1	Model 4+	-185552.03	0.0957	1.95 ± 0.0003	—	-0.28 ± 0.30	0.93 ± 0.003
10	Model 4+	-338572.91	0.0445	1.987 ± 0.0001	—	-4.679 ± 0.05	-2.167 ± 0.002
100	Model 2+	-223127.60	0.0792	128.98 ± 1.62	0.014 ± 0.0002	-134.83 ± 1.64	—
1000	Model 4+	-327636.00	0.046	1.925 ± 0.0003	—	-174.83 ± 1.24	-11.164 ± 0.003

parameters that belong to the selected model.

2.2.4 Degree Distribution

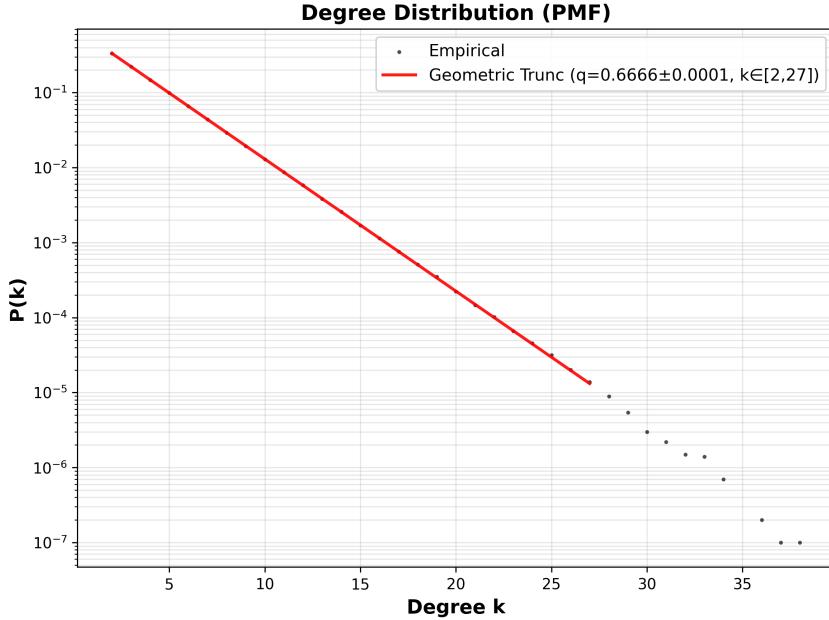


Figure 8: Empirical degree distribution for the BA random attachment model, shown on logarithmic axes together with the fitted truncated geometric distribution over the interval $k \in [2, 27]$.

Figure 8 displays the empirical degree distribution of the random-attachment model at the final simulation time. The plot shows the relative frequency of nodes with degree k on logarithmic axes. The red curve corresponds to the truncated geometric distribution fitted to the empirical data in the interval $k \in [2, 27]$. The fitted distribution has been rescaled for visualization so that it aligns with the empirical mass within the shown range. Table 4 reports the maximum-likelihood parameter estimates and AIC

Table 4: Maximum-likelihood estimates and AIC values for candidate models fitted to the degree distribution of the BA random attachment model over the interval $k \in [2, 27]$.

Model	Parameters	AIC	ΔAIC
poisson	$\lambda = 3.9993 \pm 0.0006$	43653690.89	5468353.31
displaced_geometric	$q = 0.6666 \pm 0.0001$	38185343.6	6.08
geometric_trunc	$q = 0.6666 \pm 0.0001$, $k \leq 27$	38185337.58	0.00
zeta	$\gamma = 2.2799 \pm 0.0004$, $k_{\min} = 2$	40405347.82	2220010.24
zeta_trunc	$\gamma = 2.0914 \pm 0.0005$, $k \in [2, 27]$	39745644.51	1560306.93
gaussian	$\mu = 2.000 \pm 0.001$, $\sigma = 3.343 \pm 0.002$	39839463.45	1654125.87

scores for all candidate models fitted to the empirical degree distribution over the interval $k \in [2, 27]$. The models considered include geometric, truncated geometric, displaced geometric, zeta, truncated zeta, and Gaussian distributions. For each model, the fitted parameters, AIC value, and ΔAIC relative to the best-fitting model are shown.

2.3 No-Growth Preferential Attachment

2.3.1 Time evolution of $k_i(t)$

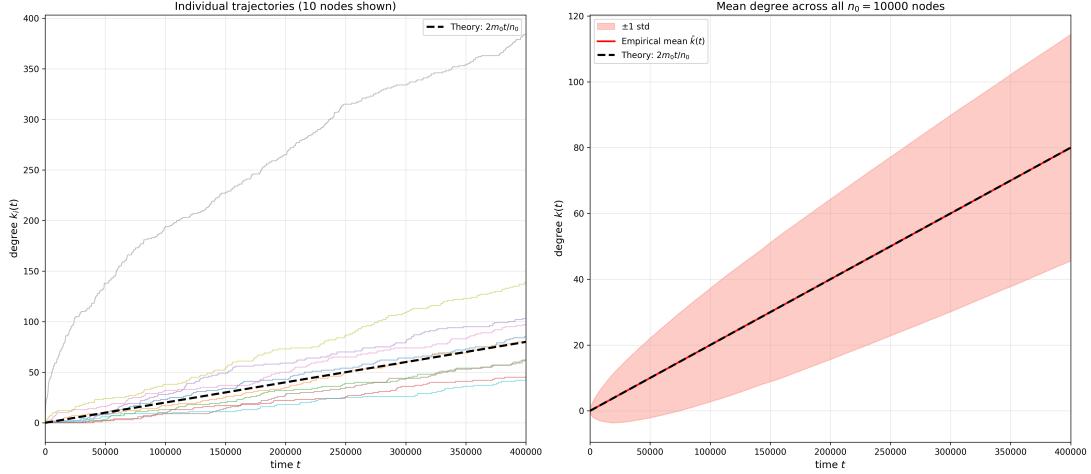


Figure 9: Degree evolution in the no-growth preferential attachment model ($n_0 = 10000$, $m_0 = 1$). Left: individual degree trajectories for a sample of 10 nodes, together with the theoretical prediction $k(t) = 2m_0t/n_0$. Right: empirical mean degree $\bar{k}(t)$ across all nodes with ± 1 standard deviation, shown together with the same theoretical curve.

Figure 9 displays the raw degree evolution in the no-growth preferential attachment model with $n_0 = 10000$ and $m_0 = 1$. The left panel shows individual degree trajectories for a set of 10 representative nodes, plotted together with the theoretical prediction $k(t) = 2m_0t/n_0$. The right panel presents the empirical mean degree $\bar{k}(t)$ over all nodes as a function of time, accompanied by a ± 1 standard-deviation band and the same theoretical line. Both panels use the full time window $t \in [0, 400000]$ as generated by the simulation.

2.3.2 Model Selection and Best Model Fit

Table 5: Model selection results for the no-growth preferential attachment model with $n_0 = 10000$ and $m_0 = 1$.

Model	AIC	RMSE	Parameters
Model 0	-24712313.51	$< 10^{-7}$	$a = 0.00020 \pm < 10^{-7}$

Table 5 reports the AIC value, RMSE, and parameter estimate for the model best fitted to the time-evolution data. The fitting range is restricted to $t \geq 10000$, as indicated in the model-selection procedure. Numerical results for the model include the estimated coefficients and its standard errors. Figure 10 shows the empirical mean degree trajectory in the no-growth preferential attachment model over the interval $t \in [10000, 400000]$, together with the ± 1 standard deviation band. The best-fitting model selected by AIC is plotted alongside the empirical curve. The theoretical prediction $k(t) = 2m_0t/n_0$ is also displayed for comparison. The fitting was performed only on the interval $t \geq 10000$, as specified in the model-selection procedure.

2.3.3 Degree Distribution

Figure 11 shows the empirical degree distribution of the no-growth preferential attachment model at the final time. The red curve corresponds to the Gaussian distribution fitted to the data in the interval $k \in [25, 100]$. For visualization purposes, the fitted curve is normalized to match the empirical mass within the shown range.

Table 6 lists the maximum-likelihood parameter estimates and AIC scores for all fitted distributions. The fitting was performed over the interval $k \in [25, 100]$, and the table includes the Poisson, geometric, truncated geometric, zeta, truncated zeta, and Gaussian models. For each model, the estimated parameters, AIC value, and Δ AIC relative to the best-fitting model are reported.

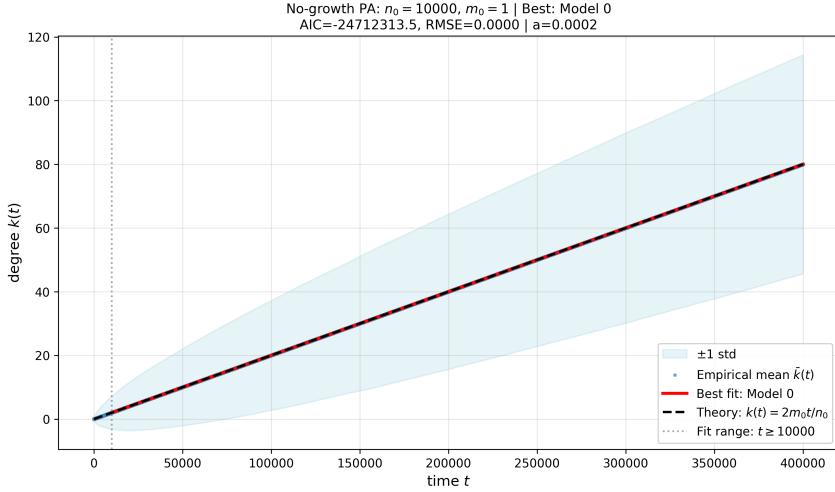


Figure 10: Model-selection result for the no-growth preferential attachment model with $n_0 = 10000$ and $m_0 = 1$. The plot shows the empirical mean degree $\bar{k}(t)$ with ± 1 standard deviation, together with the best-fit model according to AIC. The theoretical prediction $k(t) = 2m_0t/n_0$ is shown for reference.

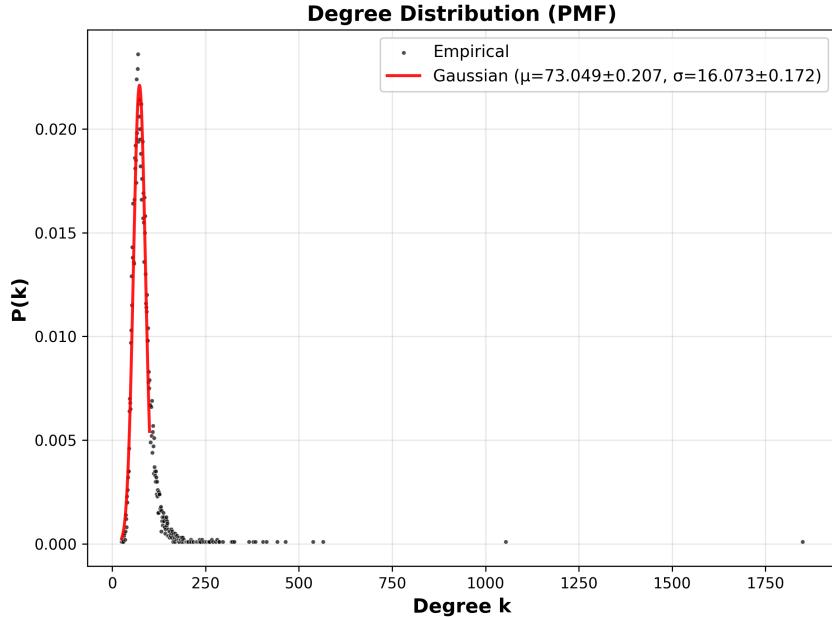


Figure 11: Empirical degree distribution for the no-growth preferential attachment model ($n_0 = 10000$, $m_0 = 1$) at the final simulation time. The histogram together with the fitted Gaussian distribution over the interval $k \in [25, 100]$.

3 Discussion

3.1 BA with Preferential Attachment

3.1.1 Scaling of $k_i(t)$

The mean-field prediction for the Barabási–Albert model states that the degree of a vertex arriving at time t_i grows as

$$k_i \approx m_0 \sqrt{\frac{t}{t_i}}$$

Our simulations are consistent with this square-root behaviour for all arrival times. In the raw degree plots (Figure 1), the empirical mean trajectories follow the same concave, sublinear shape as the theoretical curve for later arrivals.

Table 6: Maximum-likelihood parameter estimates and AIC values for models fitted to the degree distribution of the no-growth preferential attachment model over the interval $k \in [25, 100]$. Estimates marked with an asterisk (*) are reported without uncertainty because the Hessian matrix was numerically unstable, causing error estimates to diverge. Anyway, these models were not ultimately selected.

Model	Parameters	AIC	ΔAIC
poisson	$\lambda = 71.5579 \pm 0.0917$	77279.07	8215.89
displaced_geometric	$q = 0.9790^*$	82513.91	13450.74
geometric_trunc	$q = 0.9790^*, k \leq 100$	82513.91	13450.74
zeta	$\gamma = 1.8816 \pm 0.0088, k_{\min} = 25$	109186.43	40123.25
zeta_trunc	$\gamma = 0.1000 \pm 0.0277, k \in [25, 100]$	73984.07	4920.89
gaussian	$\mu = 73.05 \pm 0.21, \sigma = 16.07 \pm 0.17$	69063.18	0.00

Small but systematic deviations from the theory do appear, and they differ across arrival times. For the earliest vertex ($\tau = 1$) and intermediate arrivals ($\tau = 100$), the empirical curve lies below the theoretical curve. The curves for $\tau = 10$ and $\tau = 1000$ align more closely with the theoretical line.

These discrepancies occur because the mean-field expression is derived using the approximation

$$\frac{dk_i}{dt} = \frac{m_0 k_i}{2t-1} \xrightarrow{t \rightarrow \infty} \frac{m_0 k_i}{2t}$$

That is an asymptotic approximation that becomes accurate only when the arrival time is large. Early vertices experience strongly discrete dynamics and stronger finite-size effects, while later vertices (especially $\tau = 1000$) enter the network when the total degree is already large, making the continuum approximation more accurate.

Overall, the simulations support the theoretical claim that the BA model produces square-root growth of vertex degrees for sufficiently large times, with arrival-time-dependent scaling and strongly heteroscedastic fluctuations.

3.1.2 Rescaled Data Collapse

To test the predicted scaling behaviour of the BA model, we rescaled the degree trajectories according to

$$k'_i(t) = \sqrt{\tau} k_i(t)$$

so that, under the mean-field approximation, all curves should follow the universal law

$$k'_i(t) \approx m_0 \sqrt{t}$$

independently of the arrival time τ . The rescaled plot (Figure 2) shows that this transformation substantially reduces the differences between vertices arriving at different times and produces a clearer common trend for later arrival times.

However, as expected, the collapse is imperfect, particularly for early arrivals. The rescaled trajectory for the earliest vertex ($\tau = 1$) shows systematic deviations from the theoretical curve, especially at small t . This occurs for the same reason discussed above: the approximation $dk_i/dt \approx m_0 k_i/(2t)$ breaks down at early times, when discrete attachment events dominate and the asymptotic scaling has not yet been established. For $\tau = 100$, similar but weaker deviations persist. In contrast, the curves for $\tau = 10$ and $\tau = 1000$ collapse much better onto the theoretical line, as these vertices experience dynamics closer to the asymptotic regime throughout most of their evolution.

These systematic offsets for small τ demonstrate that finite-time corrections to the mean-field prediction are significant and arrival-time dependent. Despite these discrepancies, the overall alignment after rescaling supports the existence of an underlying square-root growth pattern that becomes increasingly accurate as t and τ grow large.

3.1.3 Model Selection

To determine which functional form best describes the growth of vertex degrees in the BA model, we fitted all models prescribed in the lab (Models 0 through 4 and their "+" variants) to the empirical mean trajectories using non-linear least squares. Model selection was performed using the AIC criterion specified in the lab instructions, and the fitting was carried out on the interval $t \geq \tau$, as required. Table 1

reports, for each arrival time τ , the best-fitting model together with its AIC value, RMSE, and parameter estimates with standard errors. The corresponding fitted curves are shown in Figure 3, where the selected model is plotted alongside the empirical mean trajectory and its variability.

Across all four arrival times ($\tau = 1, 10, 100, 1000$), the selected model is consistently Model 2+, which has the form

$$f(t) = at^b + d$$

The parameter estimates reveal that the fitted exponent b remains extremely close to $\frac{1}{2}$ for every τ , with values ranging from approximately 0.49 to 0.50. This is precisely the exponent predicted by the mean-field growth law

$$k_i(t) \sim t^{1/2}$$

indicating that the fitted model successfully captures the dominant scaling behaviour.

The additive constant d plays a crucial role in improving the fit quality. Since the fitting range starts at $t = \tau$ rather than at $t = 0$, and since each vertex enters the network with an initial degree $k_i(\tau) = m_0$, the pure power-law form at^b cannot accurately describe the data near the initial time. The offset parameter d effectively accounts for this initial condition and for the fact that we are fitting the absolute degree $k_i(t)$ rather than the degree increment $k_i(t) - k_i(\tau)$. This explains why the "+" variants systematically outperform their non-shifted counterparts in the AIC ranking. In Figure 3, this is reflected in the close alignment between the fitted Model 2+ curves and the empirical mean trajectories across all panels, particularly in the region close to $t = \tau$.

The fitted curves provide excellent descriptions of the data over the entire fitting range $t \geq \tau$. By including the offset parameter d , Model 2+ effectively captures both the asymptotic square-root scaling and the boundary condition at $t = \tau$, yielding an accurate and flexible representation of the empirical growth dynamics.

Overall, the model-selection analysis confirms that the BA preferential attachment model is well described by a generalized power-law growth function with exponent $b \approx \frac{1}{2}$. The consistency of this result across all arrival times demonstrates that the square-root scaling is robust, even when finite-time effects introduce systematic deviations from the idealized mean-field expression.

3.1.4 Degree Distribution

The degree distribution at the final time was examined by fitting several candidate models on the empirical data using maximum-likelihood estimation over the tail interval $k \in [20, 200]$. This choice is consistent with the theoretical requirement that the BA model's power-law behaviour holds only asymptotically. For small degrees, the distribution is strongly affected by discreteness and early-time finite-size effects, while for very large degrees the empirical counts are sparse and sampling noise dominates. Selecting an intermediate tail region is therefore necessary for obtaining meaningful estimates of the exponent and comparing models using AIC.

Table 2 reports the MLE parameter estimates and AIC values for all models prescribed in the lab. Among all fitted distributions (Poisson, geometric, truncated geometric, Gaussian, zeta, and truncated zeta), the truncated zeta distribution achieves the lowest AIC, making it the preferred model for the empirical tail. The corresponding fitted exponent is

$$\gamma = 2.921 \pm 0.0058$$

which is close to, but slightly below, the mean-field prediction $\gamma = 3$ for the Barabási–Albert model. This deviation is expected and comes from the fact that the theoretical value $\gamma = 3$ is derived under the continuum approximation, which assumes $t \rightarrow \infty$ and neglects discrete attachment dynamics. As explained in Section 4.4, the power-law exponent is exact only in the asymptotic limit, and finite-time simulations necessarily produce corrections. Our choice of $t_{\max} = 10^5$ represents a practical compromise between computational cost and asymptotic accuracy, but it is not large enough to fully eliminate finite-time effects.

Figure 4 shows the empirical PMF on logarithmic axes together with the fitted truncated zeta distribution. For visualization purposes only, the fitted distribution has been renormalized to match the empirical mass within the displayed interval. This adjustment affects only the vertical placement of the fitted curve and has no influence on the MLE estimates or the AIC values used for model selection. The fitted line aligns well with the empirical tail across the chosen interval, illustrating the good agreement between the truncated zeta model and the data.

Finally, it is worth noting that the clarity of the power-law behaviour is facilitated by the initial condition used in the simulations as discussed in Section 4.1.

Overall, the degree-distribution analysis confirms that the simulated network exhibits the characteristic heavy-tailed behaviour predicted by the Barabási–Albert model. The truncated zeta distribution provides the best statistical description of the tail, and the estimated exponent $\gamma \approx 2.921$ is consistent with the theoretical value $\gamma = 3$ when finite-time corrections are taken into account. Reaching the exact asymptotic value would require simulations at significantly larger time scales.

3.2 BA with Random Attachment

3.2.1 Scaling of $k_i(t)$

In the random-attachment variant of the BA model, each new edge selects its target uniformly among existing vertices, independently of their degree. The theoretical prediction for the degree evolution of a vertex arriving at time τ follows from integrating the mean-field equation

$$\frac{dk_i}{dt} = \frac{m_0}{n_0 + t - 1}$$

which yields

$$k_i(t) \approx m_0[\log(n_0 + t - 1) - \log(n_0 + \tau - 1) + 1]$$

This expression is the one implemented in the plotting code and shown as the dashed curve in Figure 5.

The empirical results match this theoretical form closely for all four arrival times $\tau = 1, 10, 100, 1000$. In each panel, the empirical mean trajectory follows the predicted logarithmic growth shape, and the discrepancy between simulation and theory remains small across the full time window. The effect of the arrival time appears mainly as a vertical shift: earlier vertices have had more opportunities to receive edges and therefore maintain a larger degree at every time t . This is consistent with the theoretical dependence on the difference of logarithms in the formula above.

Across the four arrival times, the empirical curves remain close to the theoretical prediction throughout the simulation range, indicating that the mean-field approximation for random attachment is highly accurate even at moderate times. Unlike in the preferential-attachment case, where finite-size effects can produce visible deviations for early vertices, the random-attachment model does not rely on an approximation such as $2t - 1 \approx 2t$, and the continuum solution remains valid essentially for all t . Consequently, the empirical trajectories adhere tightly to the theoretical logarithmic form, with only minor fluctuations driven by finite sampling.

Overall, the raw degree evolution in the random-attachment model exhibits excellent agreement with the theoretical prediction, with consistent behaviour across all arrival times.

3.2.2 Rescaled Data Collapse

Starting from the theoretical expression

$$k_i(t) = m_0[\log(n_0 + t - 1) - \log(n_0 + \tau - 1) + 1]$$

we can remove the dependence on the arrival time τ by defining the transformed variable

$$k'_i(t) = k_i(t) + m_0 \log(n_0 + \tau - 1) - m_0$$

Under this transformation, all vertices are predicted to satisfy

$$k'_i(t) \approx m_0 \log(n_0 + t - 1)$$

which no longer contains τ and therefore represents the universal growth law of the model.

Figure 6 shows the empirical rescaled trajectories for the four considered arrival times $\tau = 1, 10, 100$, and 1000, plotted together with the theoretical curve $m_0 \log(n_0 + t - 1)$. After rescaling, the degree trajectories for all arrival times closely overlap across the entire time range, indicating that the transformation has successfully removed the vertical offsets present in the raw data. The figure shows that all curves follow the same logarithmic shape, with only small fluctuations around the theoretical line. These fluctuations are associated with stochastic variability in edge selection but remain modest and consistent across arrival times.

In contrast to the preferential-attachment model, where the rescaling does not fully eliminate differences for the earliest and latest arrival times, the random-attachment model exhibits an almost perfect collapse. The residual variability is minor and mainly due to the finite number of simulation runs and the

finite size of the network. This behaviour is consistent with theory: unlike the preferential-attachment case, the random-attachment model does not rely on a continuum approximation of the form $\frac{dk_i}{dt} \propto k_i/t$, and the expression for $k_i(t)$ is exact apart from the discrete nature of attachment events. As a result, the empirical trajectories follow the theoretical prediction very closely, and the rescaling transformation performs as expected.

Overall, the successful collapse of the trajectories provides strong evidence that the logarithmic growth law accurately describes the temporal evolution of degrees in this model.

3.2.3 Model Selection

Model selection was carried out by fitting all prescribed models (Models 0–4 and their “+” variants) to the empirical mean trajectories for each arrival time τ , using AIC to determine the best description of the data. The results are reported in Table 3, and the corresponding fitted curves appear in Figure 7.

The theoretical degree-growth law for random attachment is

$$k_i(t) = m_0[\log(n_0 + t - 1) - \log(n_0 + \tau - 1) + 1]$$

which is logarithmic in t . Consistent with this prediction, the AIC procedure selects Model 4+, the logarithmic model

$$f(t) = a \log(t + d_1) + d_2$$

for $\tau = 1, 10$ and 100 , where the empirical trajectories most closely follow the theoretical range. The fitted parameters d_1 and d_2 capture the horizontal and vertical shifts associated with the arrival time τ .

For $\tau = 100$, the best-fitting model is Model 2+,

$$f(t) = at^b + d$$

but with very small exponents ($b \approx 0.01$), meaning that the function increases extremely slowly. Over the simulated time window, such a shallow power law effectively approximates a logarithm. The AIC therefore favours Model 2+ in cases where the empirical curve spans a degree range that is well matched by its flexible vertical shift and small exponent.

Across all arrival times, the selected models reproduce the empirical trajectories with low RMSE, and the fitted curves in Figure 7 track the data closely. The overall pattern is fully consistent with the theoretical prediction that degrees grow logarithmically under random attachment: either the explicit logarithmic model (Model 4+) or a shallow power-law approximation (Model 2+) provides the best fit depending on the arrival time.

3.2.4 Degree Distribution

The degree distribution of the random-attachment model was analysed using maximum-likelihood estimation over the interval $k \in [2, 27]$, as prescribed by the lab guidelines. This interval excludes the lowest-degree region, where no element was recorded, and avoids the extreme upper tail, where finite-size sampling noise makes estimation unreliable. The fitted distributions and their corresponding AIC values are reported in Table 4, and the empirical histogram together with the fitted model is shown in Figure 8.

Across all tested models (displaced geometric, truncated geometric, zeta, truncated zeta, and Gaussian) the best-fitting model according to AIC is the truncated geometric distribution, with parameter

$$q = 0.6666 \pm 0.0001$$

This result is consistent with the theoretical prediction for random attachment: each new edge chooses its target uniformly among existing vertices, which leads to an exponential (geometric) degree distribution rather than a power law. In particular, the continuum mean-field argument shows that under uniform attachment, the probability that a vertex receives an edge at time t is

$$\Pr(\text{increment at } t) = \frac{1}{n_0 + t - 1}$$

independent of its degree. This memoryless structure produces degree increments that behave like repeated Bernoulli trials, resulting in a geometric stationary distribution in the limit of large networks.

According to the theoretical derivation [1], the continuum limit predicts an exponential degree distribution $p(k) \propto \exp(-k/m_0)$, which corresponds to a geometric distribution with parameter $q_{\text{theory}} =$

$\exp(-1/m_0)$. For $m_0 = 2$, this yields $q_{\text{theory}} \approx 0.606$, which is approximately 10% lower than the observed value $q \approx 0.667$. This discrepancy may be attributed to finite-size effects in the simulated network. Despite this moderate quantitative difference, the qualitative agreement confirms the exponential nature of the distribution and validates the absence of preferential attachment.

The truncated geometric model outperforms the full geometric distribution only slightly, indicating that the truncation primarily accommodates the finite upper bound $k_{\max} = 27$ imposed in the fitting window. The poor performance of the zeta and truncated zeta models reflects the fact that random attachment does not generate heavy-tailed behaviour: their AIC values are several orders of magnitude worse than those of the geometric models, confirming the absence of scale-free structure in this version of the model.

In Figure 8, the fitted truncated geometric curve has been rescaled for visualization to match the empirical mass within the selected interval. This vertical normalization is performed only for plotting purposes and does not affect the likelihood calculations or the model-selection outcome. The shape of the fitted distribution closely follows the empirical frequencies over the entire fitting range, illustrating the quality of the geometric approximation.

Overall, the results confirm that the BA random-attachment model produces a thin-tailed, exponential-like degree distribution, in agreement with theoretical expectations. Uniform attachment eliminates the preferential reinforcement mechanism responsible for the emergence of heavy-tailed power laws, and the model-selection analysis quantitatively validates this behaviour.

3.3 No-Growth Preferential Attachment

3.3.1 Time evolution of $k_i(t)$

In the no-growth preferential-attachment model, the number of vertices is fixed at n_0 , and each time step adds a new edge between two existing nodes selected proportionally to their degree. Since all vertices are present from $t = 0$, there is no arrival-time dependence; all nodes evolve under identical conditions. The mean-field argument gives a deterministic prediction for the average degree:

$$\langle k(t) \rangle = \frac{2m_0}{n_0}t$$

with $m_0 = 1$ and $n_0 = 10000$, yielding a theoretical slope

$$a = \frac{2}{10000} = 0.0002$$

Figure 9 illustrates these dynamics. The left panel shows ten individual trajectories, which fluctuate substantially but follow the same long-term linear trend. The right panel presents the empirical mean degree with a ± 1 standard deviation band. The mean aligns exactly with the theoretical prediction, while the variance increases over time due to cumulative stochastic effects. Because no node enjoys an intrinsic advantage and all are present from the start, dispersion reflects randomness rather than structural asymmetry.

3.3.2 Model Selection

In the no-growth preferential attachment model, all vertices are present from the beginning and the network size remains fixed. Each new edge contributes two degree units, and the preferential rule implies that, on average, degree increments are distributed proportionally across all vertices. Summing over all vertices leads directly to the mean-field prediction

$$\langle k(t) \rangle = \frac{2m_0}{n_0}t$$

which is a strictly linear function of time. For $m_0 = 1$ and $n_0 = 10000$, this yields a theoretical slope of $a = 0.0002$.

The model-selection analysis confirms this prediction unambiguously. Model 0, the simple linear model

$$f(t) = at$$

achieves the lowest AIC and an RMSE of almost zero. The fitted parameter is

$$a_{fit} = 0.0002$$

matching the theoretical slope to machine precision.

The results demonstrate that linear growth fully captures the dynamics of $k_i(t)$ in the no-growth scenario. Unlike models with node arrival, no rescaling or universal collapse is required; all vertices evolve under the same rule from the same starting point, and the degree distribution of increments is homogeneous.

3.3.3 Degree Distribution

In the no-growth preferential attachment model, the number of vertices remains fixed and edges are repeatedly added between existing nodes according to a preferential rule. Because no new vertices enter the network, the mechanism that produces heavy-tailed degree distributions in the classical BA model is absent. Instead, degree increments accumulate within a closed population, leading to behaviour closer to a central-limit regime than to a power law.

The empirical degree distribution at the final time has a unimodal, bell-shaped form, as shown in Figure 11. To characterise this distribution, maximum-likelihood estimation was performed on the interval $k \in [25, 100]$, avoiding the noisy upper tail. Among all fitted models (Poisson, geometric, truncated geometric, zeta, truncated zeta, Gaussian), the Gaussian distribution achieves the lowest AIC, with parameters

$$\mu = 73.05 \pm 0.21, \quad \sigma = 16.07 \pm 0.17$$

This indicates that the distribution is symmetric and concentrated around a typical degree, in clear contrast to the broad, heavy-tailed behaviour of the classical growing BA model.

This result aligns with theoretical expectations. In the no-growth setting, each vertex receives degree increments at rates proportional to its current degree, but because the vertex set is fixed and large, these increments behave like many small, weakly dependent contributions. Over time, their cumulative effect approaches a normal distribution. The absence of preferential amplification across newly arriving nodes prevents the emergence of a scale-free tail.

The zeta and truncated zeta models perform poorly in the AIC comparison, reflecting the lack of power-law structure, while the Poisson and geometric models fail to capture the wider and more symmetric shape of the empirical distribution. Overall, the Gaussian model provides a clear and quantitatively supported description of the degree distribution in the no-growth preferential attachment model, consistent with the theoretical understanding of this regime.

3.4 Conclusions

This project investigated the dynamical and structural properties of three network models: the classical BA preferential attachment model, the BA random-attachment model, and a no-growth preferential attachment model. Across all analyses (degree evolution, scaling behaviour, model selection, and stationary degree distributions) the empirical results aligned closely with the theoretical expectations specific to each mechanism.

In the preferential attachment model, degree trajectories grew according to the mean-field law $k_i(t) \sim t^{1/2}$ (Figures 1–2). Model selection consistently chose the shifted power-law form (Model 2+) with exponent $b \approx 1/2$ (Table 1), confirming the predicted scaling. The degree distribution developed a heavy-tailed form well described by a truncated zeta distribution (Figure 4), matching the theoretical exponent $\gamma \approx 3$.

The random-attachment model behaved differently in all respects. Its degree evolution followed the logarithmic law predicted by theory, and the rescaling transformation produced an almost perfect collapse across all arrival times (Figure 6). Model selection alternated between explicit logarithmic forms (Model 4+) and shallow power-law approximations (Model 2+), both compatible with slow sublinear growth (Table 3). The corresponding degree distribution was light-tailed and fitted best by a truncated geometric distribution (Figure 8), as expected in the absence of preferential attachment.

In the no-growth preferential model, the degree evolution was purely linear, with empirical slopes matching the theoretical value $2m_0/n_0$ to machine precision (Figure 9). Model selection confirmed that the simple linear model (Model 0) was decisively preferred (Table 5). Without network growth, preferential attachment no longer produces heavy-tailed outcomes; instead, the degree distribution becomes bell-shaped and is best fitted by a Gaussian (Figure 11), consistent with cumulative stochastic fluctuations in a fixed population.

These results confirms the results of [1]: both growth and preferential attachment are required to produce scale-free behaviour, and modifying either mechanism fundamentally alters the resulting network structure.

4 Methods

4.1 Networks initialization

In the Barabási-Albert model with preferential and random attachment, we initialize the network with a fully connected graph of $n_0 = m_0 = m$ nodes, where m represents the number of edges each newly added node forms with existing nodes. This choice minimizes finite-size effects arising from the seed network: all initial nodes start with identical degree, ensuring that no node gains an artificial advantage in the tail of the degree distribution. In this way, the characteristic power-law distribution with exponent $\gamma \approx 3$ emerges clearly without spurious accumulations in the tail region.

If instead we chose a complete graph with $n_0 > m$, the initial nodes would exhibit heterogeneous degrees from the outset. These highly connected seed nodes would tend to dominate the tail of the degree distribution, creating an artificial peak in the high-degree region that does not reflect an intrinsic property of the model, but rather an artifact of the initial graph configuration. Under preferential attachment, this phenomenon is clear: the initial nodes receive a disproportionate share of new connections due to their already elevated degree, producing a more pronounced peak in the right tail of the distribution.

To investigate network dynamics under preferential attachment without growth, the initial graph is constructed with a large number of nodes $n_0 \gg m_0$, where $m_0 = 1$. The theoretical rationale for setting n_0 much larger than m_0 is to minimize saturation effects: when the network is not growing, a sufficiently large n_0 avoids rapid homogenization and allows diverse degree evolution through preferential linking. Thus, initializing with a high node count enables the emergence and observation of dynamic processes in preferential attachment without artificial constraints from limited network size.

4.2 Models algorithms

Barabási-Albert model with preferential attachment The preferential attachment mechanism is implemented using a stub list representation, where each edge (u, v) contributes two stubs to the list. The probability of selecting a node is proportional to its degree, naturally encoded by the frequency of its stubs.

Algorithm 1 Barabási-Albert with Preferential Attachment

```

1: Input:  $n_0$  (initial nodes),  $m_0$  (edges per new node),  $t_{\max}$  (time steps)
2: Output: Degree sequence and vertex time series
3: Initialize  $N_{\text{final}} \leftarrow n_0 + t_{\max}$ 
4: Initialize degree array  $\text{degrees}[0 \dots N_{\text{final}} - 1] \leftarrow 0$ 
5: Initialize empty stub list  $\text{stubs} \leftarrow []$ 
6: ▷ Build initial complete graph  $G_{n_0}$ 
7: for  $u = 0$  to  $n_0 - 1$  do
8:   for  $v = u + 1$  to  $n_0 - 1$  do
9:      $\text{stubs.append}(u, v)$  ▷ Add both endpoints
10:     $\text{degrees}[u] \leftarrow \text{degrees}[u] + 1$ 
11:     $\text{degrees}[v] \leftarrow \text{degrees}[v] + 1$ 
12:  end for
13: end for ▷ Growth phase with preferential attachment
14: for  $t = 1$  to  $t_{\max}$  do
15:    $v_{\text{new}} \leftarrow n_0 + t - 1$  ▷ New node arrives
16:    $\text{targets} \leftarrow \emptyset$ 
17:    $L \leftarrow |\text{stubs}|$ 
18:   while  $|\text{targets}| < m_0$  do
19:      $v \leftarrow \text{stubs}[\text{random}(0, L - 1)]$  ▷ Sample from stubs
20:     if  $v \neq v_{\text{new}}$  then
21:        $\text{targets} \leftarrow \text{targets} \cup \{v\}$ 
22:     end if
23:   end while
24:   for each  $v \in \text{targets}$  do
25:      $\text{degrees}[v] \leftarrow \text{degrees}[v] + 1$ 
26:      $\text{degrees}[v_{\text{new}}] \leftarrow \text{degrees}[v_{\text{new}}] + 1$ 
27:      $\text{stubs.append}(v, v_{\text{new}})$  ▷ Add new stubs
28:

```

```

29:   end for
30:   Record degrees for analysis
31: end for
32: return degree sequence and time series

```

Complexity Analysis: The initialization of G_{n_0} requires $O(n_0^2)$ operations. Each time step involves sampling m_0 targets from the stub list with expected $O(m_0)$ attempts (assuming few collisions), and updating the stub list in $O(m_0)$ time. The total complexity is $O(n_0^2 + t_{\max} \cdot m_0)$. The stub list grows as $O(t_{\max} \cdot m_0)$, making sampling operations $O(1)$ on average.

Barabási-Albert model with random attachment Random attachment selects target nodes uniformly at random, independent of their degree. This serves as a null model to contrast with preferential attachment dynamics.

Algorithm 2 Barabási-Albert with Random Attachment

```

1: Input:  $n_0$  (initial nodes),  $m_0$  (edges per new node),  $t_{\max}$  (time steps)
2: Output: Degree sequence and vertex time series
3: Initialize  $N_{\text{final}} \leftarrow n_0 + t_{\max}$ 
4: Initialize degree array  $\text{degrees}[0 \dots N_{\text{final}} - 1] \leftarrow 0$ 
5:  $\text{degrees}[0 \dots n_0 - 1] \leftarrow n_0 - 1$                                 ▷ Complete graph initialization
6:                                                               ▷ Growth phase with random attachment
7: for  $t = 1$  to  $t_{\max}$  do
8:    $v_{\text{new}} \leftarrow n_0 + t - 1$                                          ▷ New node arrives
9:    $N_{\text{current}} \leftarrow n_0 + t - 1$                                          ▷ Current network size
10:  targets  $\leftarrow$  random sample of  $\min(m_0, N_{\text{current}})$  nodes from  $\{0, \dots, N_{\text{current}} - 1\}$ 
11:  for each  $v \in \text{targets}$  do
12:     $\text{degrees}[v] \leftarrow \text{degrees}[v] + 1$ 
13:  end for
14:   $\text{degrees}[v_{\text{new}}] \leftarrow |\text{targets}|$ 
15:  Record degrees for analysis
16: end for
17: return degree sequence and time series

```

Complexity Analysis: Initialization is $O(n_0)$. Unlike the preferential attachment variant, the random attachment algorithm does not require explicit construction of the initial complete graph's edge structure. Since target selection is uniform and independent of degree, only the degree array needs initialization, which can be accomplished by setting $\text{degrees}[0 \dots n_0 - 1] = n_0 - 1$ in linear time. Each time step requires selecting m_0 random targets without replacement using Python's `random.sample()`, which takes $O(m_0)$ expected time for small m_0 relative to the network size. The total complexity is $O(n_0 + t_{\max} \cdot m_0)$, which is slightly more efficient than the preferential attachment variant due to the absence of a growing stub list that must be maintained for degree-proportional sampling.

No growth model with preferential attachment This model maintains a fixed network size n_0 while continuously adding edges according to preferential attachment. At each time step, a random node i is selected, and m_0 new edges are formed to nodes chosen preferentially based on their current degree.

Algorithm 3 No Growth with Preferential Attachment

```

1: Input:  $n_0$  (fixed nodes),  $m_0$  (edges per step),  $t_{\max}$  (time steps)
2: Output: Degree evolution matrix  $\text{k\_all}[n_0 \times (t_{\max} + 1)]$ 
3: Initialize degree array  $\text{degrees}[0 \dots n_0 - 1] \leftarrow 0$ 
4: Initialize adjacency lists  $\text{adj}[i] \leftarrow \emptyset$  for  $i = 0, \dots, n_0 - 1$ 
5: Initialize degree history  $\text{k\_all}[i, 0] \leftarrow 0$  for all  $i$                                 ▷ Temporal evolution with preferential attachment
6: for  $t = 1$  to  $t_{\max}$  do
7:    $i \leftarrow$  random integer in  $\{0, \dots, n_0 - 1\}$                                          ▷ Select initiating node
8:   for  $\ell = 1$  to  $m_0$  do
9:     candidates  $\leftarrow \{j \in \{0, \dots, n_0 - 1\} : j \neq i \text{ and } j \notin \text{adj}[i]\}$ 
10:    if candidates  $= \emptyset$  then

```

```

12:         continue                                ▷ Skip if no valid targets
13:     end if
14:
15:      $s \leftarrow \sum_{j \in \text{candidates}} \text{degrees}[j]$            ▷ Preferential selection
16:     if  $s > 0$  then
17:         Select  $j \in \text{candidates}$  with probability  $\propto \text{degrees}[j]$ 
18:     else
19:         Select  $j \in \text{candidates}$  uniformly at random          ▷ Handle zero-degree case
20:     end if
21:                                         ▷ Add edge  $(i, j)$ 
22:      $\text{adj}[i] \leftarrow \text{adj}[i] \cup \{j\}$ 
23:      $\text{adj}[j] \leftarrow \text{adj}[j] \cup \{i\}$ 
24:      $\text{degrees}[i] \leftarrow \text{degrees}[i] + 1$ 
25:      $\text{degrees}[j] \leftarrow \text{degrees}[j] + 1$ 
26: end for
27:      $\text{k\_all}[:, t] \leftarrow \text{degrees}$                       ▷ Record current degree distribution
28: end for
29: return time array and degree evolution matrix

```

Complexity Analysis: At each time step, constructing the candidate set requires checking adjacency lists, which takes $O(n_0 + \langle k \rangle)$ where $\langle k \rangle$ is the average degree. Computing preferential probabilities requires summing over candidates in $O(n_0)$. This process repeats m_0 times per step. The total complexity is $O(t_{\max} \cdot m_0 \cdot n_0)$. As the network densifies over time, the candidate set shrinks, potentially improving performance but also risking saturation when most node pairs are already connected.

4.3 Procedure of experiments

All simulations were implemented in Python and conducted following the procedures outlined below.

Barabási–Albert model with preferential attachment. We performed 100 independent simulation runs, each initialized with $n_0 = m_0 = 2$ nodes and evolved to a final network size of $N = 10^5$ nodes. In each run, we tracked the degree evolution $k_i(t)$ of four specific vertices arriving at times $\tau = 1, 10, 100$, and 1000, recording their degrees at regular time intervals throughout the simulation. The theoretical predictions for individual vertex growth were verified by computing the mean trajectories and standard deviations across all 100 runs for each value of τ . For the degree distribution analysis, we aggregated the final degree sequences from all 100 runs into a single pooled dataset. This aggregation is justified because the shape of the degree distribution is preserved across independent realizations of the same model, while pooling significantly improves statistical power and reduces sampling noise.

Barabási–Albert model with random attachment. The same simulation protocol described above was applied to the random attachment variant of the BA model, using identical parameters ($n_0 = m_0 = 2$, $N = 10^5$, 100 runs) and tracking the same arrival times $\tau = 1, 10, 100, 1000$.

No-growth model with preferential attachment. For the no-growth scenario, we initialized a fixed network of $n_0 = 10000$ disconnected nodes and evolved the system by adding one edge at a time using preferential attachment with parameter $m_0 = 1$. At each time step, one node was selected uniformly at random, and it formed a connection to another node chosen with probability proportional to degree. The simulation was run until $t_{\max} = 40 \times n_0$ time steps. The theoretical prediction for the mean degree growth was verified by computing the network-averaged mean degree $\langle k(t) \rangle$ over all nodes as a function of time. The final degree distribution at $t = t_{\max}$ was analyzed by fitting a Gaussian distribution to the empirical degree counts.

4.4 On the fitting of degree distributions

Degree distribution fits are not performed over the entire range of observed degrees, but rather over a restricted interval $[k_{\min}, k_{\max}]$ chosen to maximize the validity of the theoretical model. This approach is theoretically justified by the asymptotic nature of the expected distributions and the presence of finite-size effects at extremes of the degree range.

For the Barabási-Albert model with preferential attachment, the exact degree distribution is taken from the book of Barabási-Albert themselves [1], and it is given by:

$$p_k = \frac{2m(m+1)}{k(k+1)(k+2)}$$

where m is the number of edges added by each new node and $k \geq m$. This expression asymptotically approaches a power law $p_k \sim k^{-3}$ for large k . However, this holds asymptotically as $k \rightarrow \infty$, and deviations occur at small degrees due to discrete effects and the finite initial network structure. At the opposite extreme, for very large k in finite networks, a natural cutoff emerges (as stated e.g. in the book of Barabási-Albert [1]): the maximum expected degree k_{\max} scales with network size N , but beyond a certain threshold the number of nodes per degree bin becomes so small that the observed distribution is dominated by sampling noise. The literature explicitly addresses this phenomenon through truncated power laws or power laws with exponential cutoffs. In our analysis, we select the interval $[k_{\min}, k_{\max}]$ by visual inspection of the log-log degree distribution, choosing the region where the curve exhibits clear linear behavior while excluding both the low-degree regime (discrete and initialization effects) and the extreme high-degree tail (finite-size and sampling artifacts). For geometric and Gaussian distributions arising in the random attachment and no-growth models, respectively, we do not impose a stringent lower cutoff k_{\min} , as these distributions are not purely asymptotic and exhibit their characteristic behavior across the full degree range. Similarly, the upper cutoff remains relevant primarily for finite-size and sampling considerations rather than theoretical validity.

4.5 The discrete Gaussian model

In addition to the models considered in Lab 2, we introduce the discrete Gaussian as a new candidate model for fitting degree distributions. This model is particularly relevant for networks generated by no-growth dynamics, where degree distributions tend to concentrate around a characteristic scale.

The discrete Gaussian distribution is defined over the discrete support $k \in \{k_{\min}, k_{\min} + 1, \dots, k_{\max}\}$ with probability mass function:

$$P(k | \mu, \sigma) = \frac{1}{Z} \exp\left(-\frac{(k - \mu)^2}{2\sigma^2}\right)$$

where $\mu \in \mathbb{R}$ is the mean parameter, $\sigma > 0$ is the standard deviation, and Z is the normalization constant ensuring that the distribution sums to unity over the discrete support:

$$Z = \sum_{k=k_{\min}}^{k_{\max}} \exp\left(-\frac{(k - \mu)^2}{2\sigma^2}\right)$$

Unlike the continuous Gaussian, which is defined over \mathbb{R} , the discrete Gaussian is inherently bounded by the finite degree range $[k_{\min}, k_{\max}]$ observed in the network. This truncation is natural in finite networks and ensures proper normalization without requiring asymptotic approximations.

Parameter estimation and model selection Model parameters (μ, σ) are estimated by minimizing the negative log-likelihood (NLL):

$$\text{NLL}(\mu, \sigma) = - \sum_{k=k_{\min}}^{k_{\max}} n_k \log P(k | \mu, \sigma)$$

where n_k denotes the observed count of nodes with degree k . To avoid numerical underflow when computing the normalization constant Z for large degree ranges or small σ , we employ the log-likelihood:

$$\log Z = \log \sum_{k=k_{\min}}^{k_{\max}} \exp\left(-\frac{(k - \mu)^2}{2\sigma^2}\right) = a + \log \sum_{k=k_{\min}}^{k_{\max}} \exp\left(-\frac{(k - \mu)^2}{2\sigma^2} - a\right)$$

where $a = \max_{k \in [k_{\min}, k_{\max}]} \left\{-\frac{(k - \mu)^2}{2\sigma^2}\right\}$ is chosen to stabilize the computation. This ensures numerical stability even when the unnormalized probabilities span several orders of magnitude.

Model comparison is performed using the Akaike Information Criterion (AIC), which balances goodness-of-fit against model complexity:

$$\text{AIC} = 2p + 2 \cdot \text{NLL}$$

where p is the number of free parameters ($p = 2$ for the discrete Gaussian: μ and σ).

References

- [1] Albert-László Barabási. *Network Science*. Cambridge University Press, 2016.