An Information-Theoretic Analysis on Temporal Graph Evolution

Amirmohammad Farzaneh
Department of Engineering Science
University of Oxford
Oxford, UK
amirmohammad.farzaneh@eng.ox.ac.uk

Justin P. Coon
Department of Engineering Science
University of Oxford
justin.coon@eng.ox.ac.uk

Abstract

In this paper, we present a novel model termed Network Evolution Chains for simulating the temporal dynamics of networks. Our model's design is tailored to enable comprehensive analysis through information theory. We establish that this model creates a stationary and ergodic stochastic process, thus facilitating the application of the asymptotic equipartition property. This breakthrough paves the way for a thorough information-theoretic investigation into network behavior, encompassing the definition of typical sequences, future state prediction, and beyond.

1 Introduction

Graphs are frequently employed to represent networks in various scientific domains. Over time, several random graph models have been introduced to mimic diverse network types. Among these, two pivotal models emerged in 1959 (Erdős and Rényi [1960], Gilbert [1959]), effectively capturing properties of networks such as social (Robins et al. [2007]), electric (Holmgren [2006]), and biological networks (Saul and Filkov [2007]). However, as networks are dynamic, a model encompassing their evolving topology is essential for accurate analysis. Network evolution has been extensively studied, yielding models like the Barabási–Albert model (Dorogovtsev and Mendes [2002]), while Network Evolution Models simulate network changes (Banks and Carley [1996], Toivonen et al. [2009]). Though these models answer key network science questions, our focus is quantifying network evolution for efficient information transmission. Existing models often overlook network contraction, seen in scenarios like online social network departures. Here, we introduce an information-theoretic Network Evolution Model capable of simulating various networks' dynamics.

Central to information theory, finding probable sequences of a random variable aids data compression. Asymptotic Equipartition Property (AEP) defines typical sequences, highly probable occurrences. This concept drives us to seek an analogous approach for graphs, essential for scenarios requiring maximal topology compression. Often, we possess a network's current state, aiming to compress future states. Thus, we propose a network evolution model, both accurately simulating networks and amenable to information theory analysis. While Zhao et al. [2011] explores entropy rate in growing networks, differences exist in our model's entropy interpretation (section 3).

We start by defining the proposed model, and then continue by studying the model using information-theoretical tools such as the Shannon-McMillan-Breiman theorem and entropy rate. After extracting information-theoretic parameters of the model, we move on to introduce two expansions to the proposed model. The first expansion allows us to study the network in the continuous-time domain, and the second one lets the user focus on functions of the network rather than the network topology.

Temporal Graph Learning Workshop @ NeurIPS 2023, New Orleans.

2 Model Definition

We will start by defining the proposed temporal graph evolution model, which we have named Network Evolution Chain.

Definition 2.1 (Network Evolution Chain)**.** A stochastic process $\{G_i\}$ is called a Network Evolution Chain (NEC) if

- G_1 is a graph with exactly 1 node.
- G_i $(i > 1)$ is a graph created from G_{i-1}. To this end, a transition scheme is first chosen among the following options: deletion, addition, and same. The probability of choosing deletion, addition, or same is $r(G_{i-1})$, $a(G_{i-1})$, and $s(G_{i-1})$, respectively. Therefore, we have $r(G_{i-1}) + a(G_{i-1}) + s(G_{i-1}) = 1$. Based on the chosen transition scheme, one of the following procedures will be done for creating G_i:
  - **Deletion:** G_i is G_{i-1} with one of its nodes deleted according to a probability distribution.
  - **Addition:** G_i is generated by adding a node to G_{i-1}. The edges between the new node and the existing ones are added according to a random graph generation model. In this random graph generation model, all possible edge connections between the new node and the old ones should have a non-zero probability.
  - **Same:** G_i is the same as G_{i-1}.
- There exists an integer $n_{max} \geq 1$ such that if G_i is a graph with n_{max} nodes then we have $a(G_i) = 0$. For all other graphs $a(G_i) \neq 0$.
- If G_i is a graph with only one node, then we have $r(G_i) = 0$. For all other graphs $r(G_i) \neq 0$.
- For all graphs we have $s(G_i) \neq 0$.

Fig 1 illustrates four possible consecutive steps of an instance of an NEC defined based on definition 2.1.

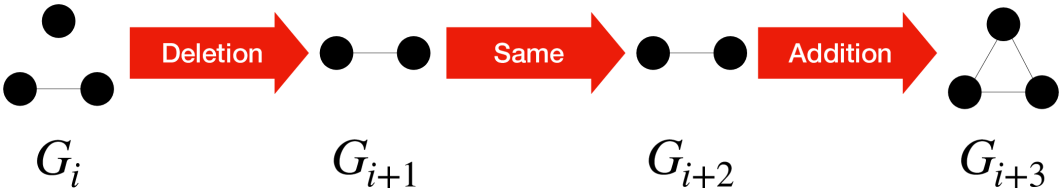

Figure 1: **Example of four consecutive graphs in an NEC.**

It can be seen that setting the parameters of an NEC according to the behaviour of real-life networks will allow us to simulate their behaviours. For instance, the model used for adding a new node can be a preferential attachment model to preserve the power law in the network that we want to simulate. The reason for the limitations that we have imposed in the definition of NECs is for us to be able to apply the Asymptotic Equipartition Property to it. We will introduce this property and perform the information-theoretic analysis of the model in the next section.

3 Information-Theoretic Analysis of the Model

We begin this section by restating the general Asymptotic Equipartition Porperty.

Theorem 1 (AEP: Shannon-McMillan-Breiman Theorem)**.** Cover and Thomas [2006, Thm. 16.8.1] If H is the entropy rate of a finite-valued stationary ergodic process $\{X_i\}$, then

$$-\frac{1}{n}\log p(X_1, X_2, ..., X_n) \to H \quad \textit{with probability } 1.$$

Therefore, to be able to use general AEP, we need to prove that the stochastic process formed by an NEC is a statioinary and ergodic process. The following lemmas, and theorems provide us with the proof for this claim.

Lemma 2. Cover and Thomas [2006, Ch. 4] "If the finite-state Markov chain $\{X_i\}$ is irreducible and aperiodic, the stationary distribution is unique, and from any starting distribution, the distribution of X_n tends to the stationary distribution as $n \to \infty$."

Corollary 2.1. Pakes [1969, Theorem 2] A finite-state, irreducible, and aperiodic Markov chain will form an ergodic process.

Proof. Theorem 2 from Pakes [1969] provides a condition under which an irreducible and aperiodic Markov chain is ergodic. It can be easily shown that if the Markov chain is finite-state, it fulfils the specified condition in the mentioned theorem and the Markov chain will be ergodic. □

Theorem 3. A Network Evolution Chain tends towards a stationary and ergodic stochastic process as $n \to \infty$.

Proof. Using lemma 2 and corollary 2.1, it suffices for us to show that an NEC is a finite-state, irreducible and aperiodic Markov chain. We will talk about each of these in turn. Firstly, based on its definition, an NEC is in fact a Markov chain, as G_i is a random graph that only depends on G_{i-1}. Furthermore, this Markov chain is also finite-state. Notice that the definition of an NEC does not allow us to have graphs with more than n_{max} nodes in the chain. Therefore, our state space consists of all graphs whose number of nodes is less than or equal to n_{max} and greater than or equal to 1, which is finite. An NEC is also irreducible. To prove this, we must present a manner in which we can go from any state to another. Suppose we want to go from graph g_1 to graph g_2, and both of these graphs are in the state space. If $g_1 = g_2$, we are done. Otherwise, we first remove nodes from g_1 one by one until we reach a graph with only one node. Notice that this is possible because $r(g) \neq 0$ for $\{g|g \in S, |V(g)| \neq 1\}$ where S shows the state space of possible graphs and $|V(g)|$ shows the number of nodes in graph g. Afterwards, we can add nodes to our one node graph one by one until we build g_2. This is also possible because $a(g) \neq 0$ for $\{g|g \in S, |V(g)| \neq n_{max}\}$, and the fact that recreating the exact edge connections of g_2 has a non-zero probability. To prove that an NEC is aperiodic, notice that the shortest path we can take from a graph to itself is of length 1. This is because for all $g \in S$ we have $s(g) \neq 0$. Since the largest common factor of any number and 1 is 1, we have proved that this chain is aperiodic. Now that we have proved that an NEC fulfils all the conditions in lemma 2 and corollary 2.1, we can conclude that an NEC will tend towards a stationary and ergodic process as $n \to \infty$.

□

Now that we have proven that the AEP can be applied to NEC, we are equipped with the necessary tools to perform an information-theoretic analysis of the model. Namely, we will start by calculating the entropy rate of the model, and defining typical sequences. These sequences can help us both in compression and the prediction of the future states of the network.

we can use theorem 1 and write

$$-\frac{1}{n}\log p(G_1, \ldots, G_n) \to H(G) \quad with\ probability\ 1, \tag{1}$$

where $H(G)$ shows the entropy rate of the stochastic process $\{G_i\}$.

Based on equation 1, we can define the typical sets of an NEC.

Definition 3.1. Cover and Thomas [2006, Ch. 3] The typical set $A_\epsilon^{(n)}$ with respect to a defined NEC is the set of all sequences $(G_1, G_2, ..., G_n) \in S^n$ with the property

$$2^{-n(H(G)+\epsilon)} \leq p(G_1, G_2, ..., G_n) \leq 2^{-n(H(G)-\epsilon)}.$$

It can be proved that "the typical set has probability nearly 1, all elements of the set are nearly equiprobable, and the number of elements in the set is nearly 2^{nH}." (Cover and Thomas [2006, Ch. 3]). To calculate the entropy rate of an NEC, we use the following lemma.

Lemma 4. Cover and Thomas [2006, Theorem 4.2.1] For a stationary stochastic process $\{X_i\}$, the following limit exists and is equal to the entropy rate of the process:

$$H(X) = \lim_{n\to\infty} H(X_n|X_{n-1}, X_{n-2}, ..., X_1)$$

We can now use lemma 4 and calculate the entropy rate of an NEC as below.

$$
\begin{aligned}
H(G) &= \lim_{n\to\infty} H(G_n|G_{n-1}, G_{n-2}, ..., G_1) && \text{(2a)}\\
&= \lim_{n\to\infty} H(G_n|G_{n-1}) && \text{(2b)}\\
&= \lim_{n\to\infty} H(G_n|G_{n-1}, T) \\
&\quad + H(T|G_{n-1}) - H(T|G_{n-1}, G_n) && \text{(2c)}\\
&= \lim_{n\to\infty} H(G_n|G_{n-1}, T) + H(T|G_{n-1}) && \text{(2d)}
\end{aligned}
$$

Equality 2a is the definition of entropy rate based on lemma 4. Equality 2b holds because of the fact that G_n is only dependent on G_{n-1}. Equality 2c comes from taking into account the three states of addition, deletion, and remaining the same. T is defined as a random variable that indicates which of the progression methods we have chosen for going from G_{n-1} to G_n. In other words, we have $T \in \{addition, deletion, same\}$. Finally, equality 2d holds because having both G_{n-1} and G_n, there will be no uncertainty about which transition scheme has been chosen for going from G_{n-1} to G_n and therefore we will have $H(T|G_{n-1}, G_n) = 0$. For calculating $H(T|G_{n-1})$, we can write

$$
H(T|G_{n-1}) = -\sum_{g\in S} p(g)(a(g)\log a(g) + r(g)\log r(g) + s(g)\log s(g)). \tag{3}
$$

In equation 3, $p(g)$ shows the stationary distribution of graph g. Now that $H(T|G_{n-1})$ has been calculated, we move on to calculate $H(G_n|G_{n-1}, T)$.

$$
\begin{aligned}
H(G_n|G_{n-1}, T) =& \sum_{g\in S} p(g)a(g)H(G_n|G_{n-1} = g, T = addition) \\
&+ \sum_{g\in S} p(g)r(g)H(G_n|G_{n-1} = g, T = deletion) \\
&+ \sum_{g\in S} p(g)s(g)H(G_n|G_{n-1} = g, T = same) && \text{(4a)}\\
=& \sum_{g\in S} p(g)a(g)H(G_n|G_{n-1} = g, T = addition) \\
&+ \sum_{g\in S} p(g)r(g)H(G_n|G_{n-1} = g, T = deletion) && \text{(4b)}
\end{aligned}
$$

Equality 4a comes from conditional entropy and equality 4b comes from the fact that we have $H(G_n|G_{n-1} = g, T = same) = 0$ because there remains no uncertainty when a graph stays the same as it was before. Notice that equation 4b can not be more simplified because the node addition and deletion models are not known.

We can now provide the following equation for calculating the entropy rate of a Network Evolution Chain.

$$
\begin{aligned}
H(G) = &-\sum_{g\in S} p(g)(a(g)\log a(g) + r(g)\log r(g) + s(g)\log s(g)) \\
&+ \sum_{g\in S} p(g)a(g)H(G_n|G_{n-1} = g, T = addition) \\
&+ \sum_{g\in S} p(g)r(g)H(G_n|G_{n-1} = g, T = deletion) \qquad (5)
\end{aligned}
$$

Calculating the entropy rate of a Network Evolution Chain opens the door to further analysis using information theory. Having the entropy rate of the network quanitifies the amount of change we expect to see in the network from one time step to another. Having typical sequences of length n defined will help us in determining the most probable outcomes of the future time steps, in addition to knowing when the network is diverging from its trajectory and exhibiting abnormal behaviour.

4 Expansions to NEC

An NEC can be used to model the dynamics of graph topologies through time, where this dynamic is fully analyzable using information theory. However, this analysis might have its own limitations. Firstly, an NEC takes discrete steps in time, whereas we might be interested in analysing the temporal evolution of the network in a more continuous manner. Secondly, we might sometimes be interested in studying the evolution of certain properties of the graph rather than the entire graph. Because of these reasons, we introduce two expansions to NEC: Continuous-Time NEC, and Network Property Chains (NPC). In this section, we will introduce and demonstrate these two processes.

4.1 Continuous-Time NEC

We have observed that Network Evolution Chains have the ability to simulate the topology dynamics of networks. However, they fail to provide us with a simulation for the network in continuous time. In this section, we provide two methods to overcome this limitation.

4.1.1 The $s(g)$ Method

The first method uses the value of $s(g)(g \in S)$ to simulate the continuous time domain. The idea is that we first quantize time and consider the time between each step of the chain to be a constant τ. Now, for each $g \in S$, we can set $s(g)$ according to how much time we expect g to remain unchanged. Consider $L(g)$ to be the random variable that shows the time that graph g stays unchanged and $Q(g)$ to be the random variable that shows the number of consecutive steps that staying the same is chosen for graph g. We can write

$$\begin{aligned}\mathbb{E}[L(g)] &= \tau(\mathbb{E}[Q(g)] + 1) \\ &= \tau(1 + \sum_{i=0}^{\infty} i(s(g))^i(1 - s(g))) \\ &= \frac{\tau}{1 - s(g)}.\end{aligned} \tag{6}$$

After choosing our desired value for the time quantization variable τ, we can use equation 6 to set the values for $s(g)$ for all $g \in S$. This way, the expected time that each network stays unchanged will match our desired value.

4.1.2 The Continuous Time Markov Chain Method

Continuous Time Markov Chains (CTMC) (Ross et al. [1996, Ch. 5]) are an extension to Markov chains with the goal to simulate Markov chains in the continuous time domain. Therefore, we can use Continuous Time Markov Chains to simulate an NEC in continuous time.

Definition 4.1. A Continuous Time Network Evolution Chain is a Continuous Time Markov Chain $G(t)$ in which the states and the transition probabilities are the same as a Network Evolution Chain and additionally, $\lambda(g)$ is chosen for all $g \in S$ as the parameter for the exponential distribution used for determining the time we stay at state g.

We can see that the states are generated in the same manner as explained for the discrete time NEC. Furthermore, we remain in G_i for τ_i seconds. τ_i is generated using an exponential distribution with parameter $\lambda(G_i)$. We therefore have to choose a value $\lambda(g)$ for all $g \in S$. This value is what determines how much time we wait in each state according to an exponential distribution with its expected value equal to $1/\lambda(g)$.

Suppose that we label the graphs in S from 1 to $|S|$ according to any ordering. Let p_{ij} denote the probability of going from state i to j in the respective NEC. We use the following notation from chapter 5 of Ross et al. [1996] to provide a result about the stationary distribution of Continuous Time Network Evolution Chains.

$$p_{ij}(t) = p(G(t + s) = j | G(s) = i)$$

$$r_{ij} = \begin{cases} \lambda(i)p_{ij}, & \text{if } i \neq j \\ -\lambda(i), & \text{if } i = j \end{cases}$$

p_{ij} represents the probability that the process goes to state j when it leaves state i. Based on the definition, $p_{ij}(t)$ represents the probability of the process with current state of i moves to state j after time t has passed, and r_{ij} represents the rate of going from state i to state j. To assist us with the calculations, the matrices $\boldsymbol{R}$, $\boldsymbol{P}(t)$, and $\boldsymbol{P}'(t)$ are defined. The element in row i and column j of these matrices are, respectively, r_{ij}, $p_{ij}(t)$, and $p'_{ij}(t)$ (Ross et al. [1996, Ch. 5]). It is stated in Ross et al. [1996] that the following equation is the solution for finding the stationary distribution of a CTMC.

$$\boldsymbol{P}(t) = e^{\boldsymbol{R}t} = \sum_{t=0}^{\infty} (\boldsymbol{R}t)^i / i! \tag{7}$$

where $\boldsymbol{R}^0$ is the identity matrix.

Equation 7 provides us with the solution to find the stationary continuous time distributions. However, it may be computationally infeasible to calculate the distribution using this equation. To overcome this challenge, Ross et al. [1996] presents a method to estimate the answer to equation 7 (Ross et al. [1996, pg. 250]).

4.2 Network Property Chains (NPC)

We have introduced Network Evolution Chains in order to simulate the way networks evolve over time. When a certain network is simulated using an NEC, we have the whole graph of the network at each step as our random variable and there will be no information loss. However, if we want to only look at certain properties of the network, it may not be efficient to store the whole graphs in our chain. In this section, we introduce functions of Network Evolution Chains as a method for extracting specific information from the networks in the chain. To this end, we define Network Property Chains.

Definition 4.2 (Network Property Chain (NPC))**.** Consider the Network Evolution Chain $G = \{G_1, G_2, ..., G_k\}$. We define $Y = \{Y_1, Y_2, ..., Y_k\}$ as a Network Property Chain of G if there exists a function f so that we have

$$\forall i \in \{1, 2, ..., k\} \qquad Y_i = f(G_i).$$

It can be seen that an NPC is defined using an NEC and a function f. We can use Network Property Chains to look for certain properties and behaviors of a network. Specifically, we choose f so that it returns the property of its input that we are interested in.

Notice that a function of a Markov chain is not necessarily a Markov chain. However, the following theorem will aid us in analysing Network Property Chains without needing them to be Markov chains.

Theorem 5. A Network Property Chain is a stationary and ergodic stochastic process.

Proof. We first prove that an NPC is a stationary process. Consider P_i as the set of all $g \in S$ for which we have $f(g) = Y_i$. We can calculate the probability $p(Y_i, ..., Y_{i+k})$ using the following equation.

$$\begin{aligned} p(Y_i, Y_{i+1}, ..., Y_{i+k}) &= p(Y_i) \prod_{j=i+1}^{i+k} p(Y_j | Y_{j-1}, ..., Y_i) \\ &= p(P_i) \prod_{j=i+1}^{i+k} \sum_{g_1 \in P_{j-1}} \sum_{g_2 \in P_j} p(g_2 | g_1) \end{aligned} \tag{8}$$

As we know that an NEC is a stationary process, there are no terms in equation 8 which are time-variant and therefore an NPC is also stationary.

To prove that an NPC is ergodic, we use Birkhoff's ergodic theorem (Birkhoff [1931]). For using Birkhoff's theorem to prove ergodicity, we need to show that all functions of an NPC satisfy Birkhoff's

condition. Firstly, all functions of an NPC are also a function of its respective NEC. To see this, notice that if the function f which is used for building the NPC is injective, then this statement is easily concluded and if f is surjective, then any function of the NPC can be seen as a function of its respective NEC where the function takes on equal values for all g_i and g_j for which we have $f(g_i) = f(g_j)$. Additionally, we know that an NEC is in fact ergodic and all functions of it satisfy the condition of Birkhoff's theorem. Therefore, an NPC is also an ergodic process.

□

Theorem 5 shows that Network Property Chains are stationary and ergodic processeses. Therefore, we can apply the Asymptotic Equipartition Property to these chains as well. We will end up with the following equation after applying the AEP to NPC.

$$-\frac{1}{k}\log p(Y_1, Y_2, ..., Y_k) \to H(Y) \quad \text{as} \quad k \to \infty \tag{9}$$

Applying the AEP to an NPC gives us the opportunity to define typical sequences for Network Property Chains. This means that we will be able to obtain the typical ways that properties of the network will evolve over time. For example, assume that we are interested in the number of triangles in the network. We will define $f(g), g \in S$ so that it returns the number of triangles in g. The typical set of $\{Y_i\}$ will then contain the most probable evolution chains of the number of triangles in the network.

Notice that as an NPC is not necessarily a Markov chain, calculating its entropy rate can not be done in the same way as NECs. However, we will be able to estimate its entropy rate using the AEP. To this end, we first need $p(Y_1, Y_2, ..., Y_k)$, which can be calculated using equation 8. However, equation 8 may be very difficult to compute, particularly for NECs with a large n_{max}. To assist us with the computation, we use the concept of a Hidden Markov Model (HMM) (Rabiner and Juang [1986]). HMMs are very useful for simulating Network Property Chains. As mentioned earlier, one of the goals for introducing NPCs was to remove the need for storing the whole graphs. In an HMM, the assumption is that the original Markov chain is not available (it is hidden), and we can only observe another chain that is based on our original Markov chain. The transition probabilities for the Hidden Markov Model (G_k, Y_k) are the same as the transition probabilities for the NEC $\{G_i\}$. Assume the set P to be the image of S under f. For the emission probabilities of the HMM for all $g \in S$ and $y \in P$ we have

$$p(y|g) = \begin{cases} 1, & \text{if } f(g) = y \\ 0, & \text{otherwise} \end{cases}.$$

Now that we have defined our Hidden Markov Model, we can make use of its properties to efficiently calculate the probability of the observed sequence. HMMs provide us with the forward algorithm (Rabiner [1989]) which is used to efficiently calculate the probability of an observed sequence. Therefore, our solution for calculating $p(Y_1, Y_2, ..., Y_k)$ is to first define the Hidden Markov Model (G_k, Y_k) and then use the forward algorithm. This will provide us with the sufficient tool to calculate the probability of observed sequences, estimate the entropy rate of the Network Property Chain, and also recognize typical sequences.

5 Conclusion

In conclusion, this paper has introduced the Network Evolution Chain (NEC) model, a novel framework for simulating the temporal dynamics of evolving networks. Beyond its capability to effectively replicate a wide array of real-life networks, we have demonstrated its compatibility with the Asymptotic Equipartition Property, thereby enabling the extraction of valuable information-theoretic features. These features can be harnessed in diverse applications, including data compression and future state prediction. Furthermore, the two expansions of the Network Evolution Chains extend its utility, allowing it to simulate continuous-time changes and functions of networks. Looking ahead, we are eager to explore the potential of integrating the proposed model into dynamic graph neural networks, further enhancing its applicability in dynamic network analysis.

References

David L. Banks and Kathleen M. Carley. Models for network evolution. *The Journal of Mathematical Sociology*, 21(1-2):173–196, 1996. doi: 10.1080/0022250X.1996.9990179.

George D Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(12):656–660, 1931.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, USA, 2006. ISBN 0471241954.

Sergey N Dorogovtsev and Jose FF Mendes. Evolution of networks. *Advances in physics*, 51(4): 1079–1187, 2002.

Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

E. N. Gilbert. Random graphs. *Ann. Math. Statist.*, 30(4):1141–1144, 12 1959. doi: 10.1214/aoms/1177706098.

Åke J Holmgren. Using graph models to analyze the vulnerability of electric power networks. *Risk analysis*, 26(4):955–969, 2006.

Anthony G Pakes. Some conditions for ergodicity and recurrence of markov chains. *Operations Research*, 17(6):1058–1061, 1969.

Lawrence Rabiner and Biinghwang Juang. An introduction to hidden markov models. *IEEE ASSP Magazine*, 3(1):4–16, 1986.

Lawrence R Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.

Garry Robins, Pip Pattison, Yuval Kalish, and Dean Lusher. An introduction to exponential random graph (p*) models for social networks. *Social Networks*, 29(2):173–191, 2007. ISSN 0378-8733. doi: https://doi.org/10.1016/j.socnet.2006.08.002. Special Section: Advances in Exponential Random Graph (p*) Models.

Sheldon M Ross, John J Kelly, Roger J Sullivan, William James Perry, Donald Mercer, Ruth M Davis, Thomas Dell Washburn, Earl V Sager, Joseph B Boyce, and Vincent L Bristow. *Stochastic processes*, volume 2. Wiley New York, 1996.

Zachary M. Saul and Vladimir Filkov. Exploring biological network structure using exponential random graph models. *Bioinformatics*, 23(19):2604–2611, 07 2007. ISSN 1367-4803. doi: 10.1093/bioinformatics/btm370.

Riitta Toivonen, Lauri Kovanen, Mikko Kivelä, Jukka-Pekka Onnela, Jari Saramäki, and Kimmo Kaski. A comparative study of social network models: Network evolution models and nodal attribute models. *Social Networks*, 31(4):240–254, 2009. ISSN 0378-8733. doi: https://doi.org/10.1016/j.socnet.2009.06.004.

Kun Zhao, Arda Halu, Simone Severini, and Ginestra Bianconi. Entropy rate of nonequilibrium growing networks. *Physical Review E*, 84(6):066113, 2011.