Bayesian Inference of the Exponential Degree Distribution Complex Networks

Deqiang Zou (Department of Mathemetics, Jilin University, China)

Abstract

The exponential degree distribution has been found in many real world complex networks other than the power law, based on which, the accurate identification of exponential patterns is a significant consequence for exactly modeling complex systems. Nonetheless, the commonly used methods of analyzing exponential complex networks are barely approved in that somehow produce inaccurate estimate of exponents over the exponential distribution. In this paper, we introduce generalized random graph theory as the framework to account for the degree distribution of the complex network which follows an exponential, analyzing the derivation of the exponent with rigorous proofs. And then, we take advantage of Bayes inference to acquaint the optimal exponent both in simulated data and empirical data with the help of Markov Chain Monte Carlo (MCMC). Eventually, we conduct the Kolmogorov-Smirnov tests of goodness-of-fit to exponential distributions in empirical data, in which, the exponential exponents are estimated with Bayes-MCMC, presenting an efficient and robust analytical way in complex networks.

Keywords : Complex network Generalized random graph Exponential degree distribution Bayes-MCMC Kolmogorov-Smirnov test

1 Introduction

A central claim in modern network science is that real world networks are typically "scale free", meaning that the fraction of nodes with degree k follows a power law, decaying like $k^{-\alpha}$, often with $2 < \alpha < 3$. Whereas, the empirical evidence for this claim derives from a relatively small number of real world networks. Across the domain, we find that scale free networks are rare, with only 4% exhibiting the strongest-possible evidence of scale free structure and 52% exhibiting the weakest-possible evidence[1]. The results undermine the universality of scale-free networks and reveal that real world networks possess rich structural diversity. This trend leads us to gradually regard the other degree distribution networks which may be dug up the ponderable information to study real world behaviors, and exponential is expected. The exponential network is also widely studied whose node's connectivity distribution is exponential[2, 3]. The exponential degree distribution has been found in many real world complex networks as long as under the constraints of geographical conditions or topological characteristics[4], such as the Email Network of University at Rovira iVirgili in Spain[5], neural network of the C.elegans[6], the North American Power Grid Network[7], the Worldwide Marine Transportation Network[8] etc.

It deserves to mention that the acquaintance of degree distribution is of great significance to studying real world behaviors. Just like the North American Power Grid Network, which plays an important role among the operation and the development of cities, also in people's daily life. While the power grid is increasing on scale and complexity, more potential danger factors come out, like the frequent large power grid cascading blackouts. During 2003, the power grids had experienced a series of large blackouts, coming up in the USA and Canada. Just blackout on the 14th August, it broke up 50 million people and 61,200 MW among eight states span, causing a sharp economic loss of around 25-30 billion dollars[9, 10]. Blackouts seem to be arresting of reminding people of taking account of the power grid for granted and ponder on how to constrain their taking place and the emergency strategies to lowing loss in blackouts. That is studied, and blackouts are easily caused with the failure of a certain component which always appears on overload lines, concentration areas of electricity, or load intensive areas of the system, and then gradually expand, resulting in the system collapsing rapidly[11]. The effective solution is to set up a group of power stations distributed among the hub nodes, in order to share the load and low the risk of cascading failure. Thus it becomes much important to acquaint the transmission of electric energy among power grid network to strengthen the stability of power supplements in case of the electric emergency. Whereas performing an analytic description of the power transmission integrated over the entire power grid is daunting, instead we mainly acquaint the transmission mechanism and the inherent vulnerability of the power grid network by means of analyzing the topological structure. Once acquainted with the degree distribution of the power grid network, we would propose specific and effective ways to improve the robustness and far always from blackout risk[12, 13]. As well as the Worldwide Marine Transportation Network, which represnts the merchandise trade and economics development all over the world. For the shipping, the congestion in hub ports always makes ships wait available berth outside for several days, which would cost companies much high expense. If only they arranged their sailing schedules, they would concern about the congestion of berths with the topological structure of marine transpotation. Thus newcome shipping links are not intended to sail into the potential cluster of ports which are analyzed by degree distribution through topological structure, so as to improve efficiency and low economic loss[14].

A complex network follows an exponential degree distribution expressed as

$$P(x) \propto e^{-\frac{x}{\bar{x}}} \tag{1}$$

or purely

$$P(x) \propto e^{-\lambda x} \tag{2}$$

where $\lambda = \frac{1}{\bar{x}}$ is a constant parameter, $\bar{x}$ on behalf of the average degree of the complex network.

The current studies have been lacking points on estimating the exponential exponent if compared with the power law, but we could draw lessons from pow law research as much as possible. Generally, the methods in common making parameter estimation would take simple graphical methods, such as the direct linear fit of the $x - logy$ plot of the full raw histogram of the data[15], the linear fit to logarithmically binned histograms[16] etc. But, the easy graphical nature methods always intend to mask their basic inaccuracy[17], that estimated exponent would fluctuate greatly and hardly converge to a certain. Until Clauset, Shalizi and Newman proposed a numeric method of identifying power law phenomena, which is based on the maximum likelihood estimation[18]. The article introduces that few empirical phenomena obey power laws for all degrees, always among the fraction fell behind—tail, also for other complex networks. They take maximum likelihood estimation to the fraction data acquiring a probability distribution that fits the power law model a lot. We modify it to exactly fit the exponential in a logical way. Assuming the data are drawn from an exponential distribution exactly for $x \geq x_{min}$, of which probability distribution fits the exponential model as much as possible above x_{min}, and then we could derive maximum likelihood estimators of the exponent:

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n}(x_i - x_{min})} \tag{3}$$

where $\lambda = \frac{1}{E[X]}$, theorically $\lambda \geq 0$. In practice, the estimated exponents lack consideration of independence, thus one would only expect weak objectivity in the application. It makes improvident sense if we just took the fraction in that way which may ignore certain information. In most cases, the Pareto principle is appropriate this time[19]. Although not all degrees, promise at about 20% enough on the tail.

Complex networks are closely related to graph theory, especially for object modeling, which could be mostly attributed to various kinds of graphs build[20]. Thus the study on the topological structure of complex networks could fully refer to graph theory, and take advantage of standard graph-theoretical terminology. We propose a graph-based method to estimate the scaling parameter of exponential degree distribution complex networks, which regards the real world complex network as the generalized random graph, and then we explore the Bayesian process with abundant theoretical tools of generalized random graph. Based on the MCMC algorithm, we would effectively sample from its posterior distribution and acquire the target exponent dealing with the complicated integration process. This method could be applied to a large amount of real-world complex networks, even increasing complexity. Meanwhile, we prove that our graph-based method is a bit more plausible of fitting the data of complex networks with the KS-test even though compared.

The entire text is organized as follows. In Sec.2, we will introduce the theory of generalized random graph. In Sec.3, the Bayesian inference with MCMC will be deprived for the exponent, and discuss the optimal widget. We will make a critical comparison with maximum likelihood estimation on the simulated data, and evaluate the effectiveness of Bayes-MCMC with Kolmogorov-Smirnov test on the empirical data in Sec.4.

2 Generalized random graph theory

Generalized random graph(GRG) is one of the inhomogeneous random graphs, which is firstly introduced by Britton, Deijfen，Martin-L¨ etc[21]. In GRG, each vertex $i \in [n]$ would receive a weight w_i. The weight can be considered as the tendency of the vertex to possess edges and would turn out to be quite close to the expected degree of the vertex. Given the weights, edges are existing independently, but the occupation probabilities of different edges are not identical, which are moderated by the weights of the vertices. In the GRG model, the edge probability of the edge between vertices i and j, for $i \neq j$, is equal to

$$p_{i,j} = \frac{w_i w_j}{l_n + w_i w_j} \tag{4}$$

where l_n is the total weights of all vertices, given by

$$l_n = \sum_{i \in [n]} w_i \tag{5}$$

Literally, the topology of the generalized random graph sensitively depends upon the choice of the vertex weights $\boldsymbol{w} = (w_i)_{i \in [n]}$, which empirical distribution function can be defined by

$$F_n(x) = \frac{1}{n} \sum_{i \in [n]} \mathbb{1}_{\{w_i \leq w\}} \tag{6}$$

F_n is the distribution of the weight of a uniformly chosen vertex in $[n]$. Thus the inhomogeneity is modeled by adding vertex weights in this model, and means that vertices with high weights are more likely to have more neighbors than vertices with small weights—just degree. Vertices with extremely high weights could act as the hubs observed in many real world networks[22].

We aim to account for the exponential degree distribution of the complex network in the GRG model. The degree distribution could only converge while the vertex weights are sufficiently regular. We always assume that the vertex weights satisfy the following regularity conditions, which turn out to imply convergence of the degree distribution in GRG[22]:

Condition 1 *(Regularity conditions for vertex weights) There exists a distribution function F such that, as $n \to \infty$ the following conditions hold:*
(a) Weak convergence of vertex weight. As $n \to \infty$,

$$W_n \xrightarrow{d} W \tag{7}$$

where W_n and W have distribution functions F_n and F, respectively. Equivalently, for any x for which $x \mapsto F(x)$ is continuous,

$$\lim_{n \to \infty} F_n(x) = F(x) \tag{8}$$

(b) Convergence of average vertex weight.

$$\lim_{n \to \infty} E[W_n] = E[W] \tag{9}$$

where W_n and W have distribution functions F_n and F, respectively.

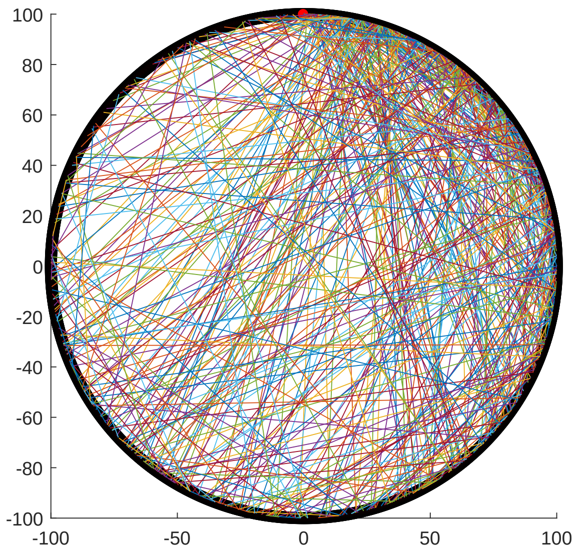

Figure 1: $GRG_n(\boldsymbol{w})$ for $(w_i)_{i\in[n]}$ as in Eq.(13) with n=2000,$\lambda = 2.5$.(For clarity in the schematic, but even smaller λ is used in the context)

An example of the generalized random graph for conducting F to have the exponential distribution for which,

$$F(x) = \begin{cases} 0 & ,x \leq b \\ 1 - e^{-\lambda(x-b)} & ,x > b \end{cases} \qquad ,b = \frac{1}{\lambda}log(\frac{\lambda+1}{\lambda}) \tag{10}$$

that b is fully considered in order to make degree distribution of the generalized random graph converges to exponential. Then

$$[1-F]^{-1}(u) = -\frac{log\,u}{\lambda} + b \tag{11}$$

take

$$w_i = [1-F]^{-1}(i/n) \tag{12}$$

so that

$$w_i = -\frac{log(i/n)}{\lambda} + b \tag{13}$$

Definition 1 *(Mixed Poisson Distribution) A random variable X possesses a mixed Poisson distribution with mixing distribution F when, for every $k \in N_0$*

$$\mathbb{P}(X = k) = E[e^{-W}\frac{W^k}{k!}] \tag{14}$$

where W is a random variable with distribution function F.

Theorem 2.1 *(Degree of uniformly chosen vertex in GRG) Assume that distribution function F satisfies* ***Conditions*** *(a)(b),then*
(1)the degree of a uniformly chosen vertex converges in distribution to a mixed poisson random variable with mixing distribution F;
(2)the degree of m uniformly chosen vertices in [n] are asymptotically independent.

A $GRG_n(\boldsymbol{w})$ is generated when we regard w_i as the weight of vertice, which distribution function is defined by Eq.(10), and just see the instance in Fig.1. By Theorem 2.1, we can ealisy prove that the degree converges in distribution to a mixed possion random variable with mixing distribution F.

Proof *To prove the above statement, we just need to prove F satisfy* ***Conditions*** *(a)(b) in Theorem 2.1*:

(a)

$$F_n(x) = \frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{w_i \le F(x)\}} = \frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{[1-F]^{-1}(i/n)\le x\}} = \frac{1}{n}\sum_{j=0}^{n-1} \mathbb{1}_{\{[1-F]^{-1}(1-\frac{j}{n})\le x\}}$$

$$= \frac{1}{n}\sum_{j=0}^{n-1} \mathbb{1}_{\{F^{-1}(\frac{j}{n})\le x\}} = \frac{1}{n}\sum_{j=0}^{n-1} \mathbb{1}_{\{\frac{j}{n}\le F(x)\}} = \frac{1}{n}(\lfloor nF(x)\rfloor + 1) \wedge 1$$

of which $j = n - i$*, when* $n \to \infty$,

$$\lim_{n\to\infty} F_n(x) = F(x)$$

Condition(a) is true.

(b)

$$\begin{aligned}
E[W_n] &= \frac{1}{n}\sum_{i=1}^{n}[-\frac{log(\frac{i}{n})}{\lambda} + b] \\
&= -\frac{1}{n\lambda}\sum_{i=1}^{n} log(\frac{i}{n}) + b \\
&= -\frac{1}{n\lambda} log(\frac{n!}{n^n}) + b \\
&\sim -\frac{1}{n\lambda} log(\frac{\sqrt{2\pi n}}{e^n}) + b \\
&= \frac{n - log\sqrt{2\pi n}}{n\lambda} + b
\end{aligned}$$

$$\begin{aligned}
E[W] &= \int_b^\infty w dF(w) \\
&= \lambda \int_b^\infty w e^{-\lambda(w-b)} dw \\
&= \frac{1}{\lambda} + b
\end{aligned}$$

As $n \to \infty$,

$$\lim_{n\to\infty} E[W_n] = E[W]$$

Condition(b) is true. This completes the argument. ■

Theorem 2.2 *When* $(w_i)_{i\in[n]}$ *are i.i.d. random variables with distribution function F which defined by Eq.10, the degree* D_i *of vertex i converges in distribution to a mixed poisson F. As the probability distribution listed as* $p(k), k \ge 0$*, there is a constant* $\bar{k}$*, which exactly makes* $p(k) \sim e^{\frac{k}{\bar{k}}}$ *only if k is large enough.*

Proof *Therefore we give a rigorous theoretical proof:*

$$\begin{aligned}
p(k|\lambda) &= E[e^{-w}\frac{w^k}{k!}] \\
&= \int_b^\infty e^{-w}\frac{w^k}{k!} dF(w) \\
&= \frac{\lambda e^{\lambda b}}{k!}\int_b^\infty w^k e^{-(\lambda+1)w} dw \\
&= \frac{\lambda e^{\lambda b}}{k!}(\frac{k!}{(\lambda+1)^{k+1}} - \int_0^b w^k e^{-(\lambda+1)w} dw) \\
&= \frac{\lambda e^{\lambda b}}{k!}(Q_1 - Q_2)
\end{aligned} \tag{15}$$

For Q_2, it is easy now to deduce that,

$$0 \leq Q_2 = \int_0^b w^k e^{-(\lambda+1)w} dw \leq b^k \int_0^b e^{-(\lambda+1)w} dw$$

while k is large enough,

$$\lim_k \frac{\lambda e^{\lambda b}}{k!} Q_2 = 0$$

so that,

$$\begin{aligned} p(k|\lambda) &\sim \frac{\lambda e^{\lambda b}}{k!} Q_1 \\ &= \frac{\lambda e^{\lambda b}}{k!} \frac{k!}{(\lambda+1)^{k+1}} \\ &= e^{-\frac{k}{\frac{1}{log(\lambda+1)}}} \end{aligned} \tag{16}$$

apparently,

$$p(k) \sim e^{-\frac{k}{\bar{k}}} \tag{17}$$

where $\bar{k} = \frac{1}{log(\lambda+1)}$ is an exactly exponential degree distribution. ■

In above process of proof, we have gotten the general format of $p(x = k|\lambda)$ in Eq.15. But to acquire a better estimating effect in the pratical work, we would exactly refine it as the following,

$$\begin{aligned} p(k|\lambda) &= \frac{\lambda e^{\lambda b}}{k!}(Q_1 - Q_2) \\ &= \frac{\lambda e^{\lambda b}}{k!}(\frac{k!}{(\lambda+1)^{k+1}} - \int_0^b w^k e^{-(\lambda+1)w} dw) \\ &= \frac{\lambda e^{\lambda b}}{k!}(\frac{k!}{(\lambda+1)^{k+1}} + \frac{1}{\lambda+1}\int_0^b w^k de^{-(\lambda+1)w}) \\ &= \frac{\lambda e^{\lambda b}}{k!}(\frac{k!}{(\lambda+1)^{k+1}} + \frac{1}{\lambda+1}(b^k e^{-(\lambda+1)b} - k\int_0^b w^{(k-1)} e^{-(\lambda+1)w} dw)) \\ &= \frac{\lambda e^{\lambda b}}{k!}[\frac{k!}{(\lambda+1)^{k+1}} + \frac{1}{\lambda+1}(b^k e^{-(\lambda+1)b} + \frac{k}{\lambda+1}(b^{k-1}e^{-(\lambda+1)b} - (k-1)\int_0^b w^{k-2}e^{-(\lambda+1)w} dw))] \\ &= \quad \ldots\ldots \\ &= \frac{\lambda e^{\lambda b}}{k!}\{\frac{k!}{(\lambda+1)^{k+1}} + e^{-(\lambda+1)b}[\sum_{j=0}^{k-1} \frac{b^{k-j}\Gamma(k+1)}{(\lambda+1)^{j+1}\Gamma(k-j+1)} + \frac{\Gamma(k+1)}{(\lambda+1)^{k+1}}] - \frac{\Gamma(k+1)}{(\lambda+1)^{k+1}}\} \\ &= \frac{\lambda}{e^b}(\sum_{j=0}^{k-1} \frac{b^{k-j}}{(\lambda+1)^{j+1}\Gamma(k-j+1)} + \frac{1}{(\lambda+1)^{k+1}}) \end{aligned} \tag{18}$$

3 Bayes inference for component

In fact, there are few empirical phenomena that obey exponential for the entire degree in complex networks, more often the exponential just only for the tail, that is. Therefore it is imperative to choose up a good tail on behalf of the entire degree information. In most cases, the Pareto principle is more applicable here that 20% data primely interpret all data and then expect a steady estimation, exactly in accordance with the real world complex networks. The main steps as follows:

(a) Randomly choose a certain number of vertices from the complex network;

(b) Rank the choosen vertices by their data of degree;

(c) Truncate the vertices where degree are at about 80%;

(d) Explore the back 20% vertices which follow the exponential.

Construct a vertice set $[n] = \{1, 2, ..., n\}$, and then denote the degree of vertices $i \in [n]$ with d_i, taking $D = (d_1, d_2, ..., d_m)$ which satisfy $d_1 < d_2 < ... < d_m$. By Theorem 2.1(2), we infer that the degree of vertices are asymptotically independent if chosen uniformly, then allow us to write

$$P(K|\lambda) \sim \prod_{i=1}^{m} p(k_i|\lambda) \tag{19}$$

Corollary 1 *Assuming that the degree of t vertices which are uniformly chosen from an exponential distribution are exactly for $k_i \geq x_{min}$ $(i \in [t])$, and then, the degrees of t vertices $K = (k_1, k_2, ..., k_t)$ are also asymptotically independent.*

$$P(K|\lambda) \sim \prod_{i=1}^{t} p(k_i|k_i \geq x_{min}) \tag{20}$$

Proof *Assuming that case $x_{min} = a$, we denote the degrees of n vertices by $X = (X_1, X_2, ..., X_n)$, and, $\mathbb{C} = \{i_1, i_2, ..., i_t\}$ are t vertices which are randomly and uniformly chosen from $\mathcal{B} = \{i : k_i \geq a\}$. Take $Y = (Y_1, Y_2, ..., Y_t)$ as a subcolumn of X, but Y satisfies $Y_j \geq a$, $j = 1, 2, ..., t$, for $Y_1 = X_{i_1}, Y_2 = X_{i_2}, ..., Y_t = X_{i_t}$,*

$$\begin{aligned}
P(Y_1 = k_1, Y_2 = k_2, ..., Y_t = k_t) &= P(X_{i_1} = k_1, X_{i_2} = k_2, ..., X_{i_t} = k_t | X_{i_1} \geq a, X_{i_2} \geq a, ..., X_{i_t} \geq a) \\
&= \frac{P(X_{i_1} = k_1, X_{i_2} = k_2, ..., X_{i_t} = k_t)}{P(X_{i_1} \geq a, X_{i_2} \geq a, ..., X_{i_t} \geq a)} \\
&\sim \prod_{j=1}^{t} \frac{p(X_{i_j} = k_j)}{p(X_{i_j} \geq a)} \\
&= \prod_{j=1}^{t} p(X_{i_j} = k_j | X_{i_j} \geq a) \\
&= \prod_{j=1}^{t} p(Y_j = k_j)
\end{aligned}$$

This complete the proof. ■

Then, the likelihood fuction Eq.(19) would be transformed to

$$\begin{aligned}
P(K|\lambda) &\sim \prod_{i=1}^{m} p(k_i|\lambda) \\
&\sim \prod_{i=1}^{t} p(k_i|k_i \geq x_{min}) \\
&= \frac{\prod_{i=1}^{t} p(k_i|\lambda)}{[1 - \sum_{j=1}^{x_{min}-1} p(k = j|\lambda)]^t}
\end{aligned} \tag{21}$$

3.1 Bayes process

Bayes theorm is that combine evidence and prior probability to acquire the posterior probability. For this part, we would introduce a prior λ, then

$$P(\lambda|K) = \frac{P(\lambda)P(K|\lambda)}{P(K)} \tag{22}$$

$P(K|\lambda)$ is conditional likelihood fuction given by Eq.(21), and $P(K)$ is normalization factor.

Assume that prior λ follows a natural distribution,

$$\pi(\lambda) = e^{-\lambda} \tag{23}$$

with the prior probability $\pi(\lambda)$ and conditional joint probability $P(K|\lambda)$, we now easily write down the posterior probability $\pi(\lambda|K)$

$$\begin{aligned}
\pi(\lambda|K) &= \frac{P(K|\lambda)\pi(\lambda)}{\int P(K|\lambda)\pi(\lambda)d\lambda} \\
&= \frac{\frac{\prod_{i=1}^{t} p(k_i|\lambda)}{[1-\sum_{j=1}^{x_{min}-1} p(k=j|\lambda)]^t} e^{-\lambda}}{\int \frac{\prod_{i=1}^{t} p(k_i|\lambda)}{[1-\sum_{j=1}^{x_{min}-1} p(k=j|\lambda)]^t} e^{-\lambda} d\lambda}
\end{aligned} \tag{24}$$

then

$$\begin{aligned}
\hat{\lambda} = E_{\pi(\lambda|K)}(\lambda) &= \int \lambda\pi(\lambda|K)d\lambda \\
&= \int \lambda e^{-\lambda} \frac{\frac{\prod_{i=1}^{t} p(k_i|\lambda)}{[1-\sum_{j=1}^{x_{min}-1} p(k=j|\lambda)]^t}}{\int \frac{\prod_{i=1}^{t} p(k_i|\lambda)}{[1-\sum_{j=1}^{x_{min}-1} p(k=j|\lambda)]^t} e^{-\lambda} d\lambda} d\lambda
\end{aligned} \tag{25}$$

Whereas consider a lot about Eq.(18) and Eq.(25), we find it complex to work out the expectation of exponent λ in Eq.(25). We intend to acquire the value of $\hat{\lambda}$ through numerical simulation by sampling the posterior probability $\pi(\lambda|K)$ with MCMC.

3.2 MCMC

Markov Chain Monte Carlo is the way of random vector sampling in higher dimensions[23], which basic idea is to construct a Markov Chain for what purpose that the stationary distribution is the posterior distribution of the goal parameter λ, and the Monte Carlo integral is performed based on the samples which taken out from the stationary distribution, so that effectively deals with the complicated distribution form, like Eq.(24).

Detailed balance is the sufficient condition for the stationary distribution of Markov Chain, which

$$\pi(\lambda|K)Q(\lambda \longmapsto \lambda^*) = \pi(\lambda^*|K)Q(\lambda^* \longmapsto \lambda) \tag{26}$$

there in,

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & \dots & Q_{1n} \\ Q_{21} & Q_{22} & \dots & Q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{m1} & Q_{m2} & \dots & Q_{mn} \end{bmatrix}$$

$Q = [Q_{ij}]$ is the state transition matrix. Whereas Q is matched with $\pi(\lambda|K)$ which somehow brings difficulty, we would never encounter that in Hasting-Metropolis, that is.

Claim 1 *Take a general proposed distribution density $q = [q_{ij}]$, and introduce the accceptance rate α, then*

$$\pi(\lambda|K)q(\lambda \longmapsto \lambda^*)\alpha(\lambda \longmapsto \lambda^*) = \pi(\lambda^*|K)q(\lambda^* \longmapsto \lambda)\alpha(\lambda^* \longmapsto \lambda) \tag{27}$$

there in

$$\alpha(\lambda, \lambda^*) = min(1, \frac{\pi(\lambda^*|K)q(\lambda^* \longmapsto \lambda))}{\pi(\lambda|K)q(\lambda \longmapsto \lambda^*)})$$

Proof *We can properly use α to accomplish the detailed balance*

$$\begin{aligned}
\pi(\lambda|K)q(\lambda \longmapsto \lambda^*)\alpha(\lambda \longmapsto \lambda^*) &= \pi(\lambda|K)q(\lambda \longmapsto \lambda^*)min(1, \frac{\pi(\lambda^*|K)q(\lambda^* \longmapsto \lambda))}{\pi(\lambda|K)q(\lambda \longmapsto \lambda^*)}) \\
&= min(\pi(\lambda|K)q(\lambda \longmapsto \lambda^*), \pi(\lambda^*|K)q(\lambda^* \longmapsto \lambda)) \\
&= \pi(\lambda^*|K)q(\lambda^* \longmapsto \lambda)min(1, \frac{\pi(\lambda|K)q(\lambda \longmapsto \lambda^*)}{\pi(\lambda^*|K)q(\lambda^* \longmapsto \lambda)}) \\
&= \pi(\lambda^*|K)q(\lambda^* \longmapsto \lambda)\alpha(\lambda^* \longmapsto \lambda)
\end{aligned}$$

this proves the claim. ■

From now, we are already qualified to carry out this algorithm, as follows:

Algorithm 1 Hasting-Metropolis algorithm

Initialize: $\lambda^{(0)} = \lambda_0$

for iterations $i = 0, 1, 2, ...$ **do**

  Propose the quantities:

  $u \sim Uniform[0, 1]$

  $\lambda^* \sim q(\lambda|\lambda^{(i-1)})$

  $\alpha(\lambda, \lambda^*) = min(1, \frac{\pi(\lambda^*|K)q(\lambda^* \longmapsto \lambda))}{\pi(\lambda|K)q(\lambda \longmapsto \lambda^*)})$

  if $u \leq \alpha$ **then**

    $\lambda^{(i)} = \lambda^*$

  else

    $\lambda^{(i)} = \lambda^{(i-1)}$

  end if

end for

The Algorithm 1 would produce n standard samples $\lambda^{(1)}, \lambda^{(2)}, ..., \lambda^{(n)}$ which intend to work out the Eq.(25) and acquire the exponent $\hat{\lambda}$,

$$\begin{aligned}
\hat{\lambda} = E_{\pi(\lambda|K)}(\lambda) &= \int \lambda \pi(\lambda|K) d\lambda \\
&= \int \lambda e^{-\lambda} \frac{\frac{\prod_{i=1}^{t} p(k_i|\lambda)}{[1-\sum_{j=1}^{x_{min}-1} p(k=j|\lambda)]^t}}{\int \frac{\prod_{i=1}^{t} p(k_i|\lambda)}{[1-\sum_{j=1}^{x_{min}-1} p(k=j|\lambda)]^t} e^{-\lambda} d\lambda} d\lambda \\
&\approx \frac{1}{n} \sum_{i=1}^{n} \lambda^{(i)}
\end{aligned} \tag{28}$$

With the law of large numbers, the $\hat{\lambda}$ intends to approach the true λ, if only n is large enough and in accordance with the independence principle. In addition, since the initial stage of Markov Chain fluctuates a lot, we are supposed to pick in samples at an appropriate state, not the early stage.

For probably existing a drawback of algorithm efficiency, we consider improving it with a proper proposal distribution density. Given the Table 1 below, we attempt a series of proposal distribution densities and then evaluate them. All of them make good performances, Nonetheless, $q(\lambda) = \frac{\partial log(\pi(\lambda|K))}{\partial \lambda}$ is a theoretically optimal choice which comes from Hybrid Metropolis Hasting.

Table 1: Evaluations among various proposal ditribution densities.

$q(\lambda)$	Evaluation	$q(\lambda)$	Evaluation
$U(0,1)$	great	$Beta(1,3)$	great
$N(0,1)$	great	e^{λ}	great
$Gamma(1,2)$	great	$\frac{\partial log(\pi(\lambda|K))}{\partial \lambda}$	great

4 Experiment

4.1 Analysis on the simulated data

We conduct the numerical simulation experiments for 9 exponential distribution networks among different sizes and then analyze the accuracy of Bayes-MCMC. For MCMC, once we do 10000 iterations with removing the initial 500 which may be in the fluctuant zone, results as follows:

Table 2: The results among different size exponential networks simulations. λ_{MLE} and λ_{MCMC} depict the estimated exponents with different methods. For each data set, r referring to the number of randomly chosen vertices from the complex network, a is the lower bound of the exponential behavior tail which truncated at about 80%, and m is the length of the tail.

No.	Data set λ	Nodes	r	m	a	Method	
						λ_{MLE}	λ_{MCMC}
(a)	0.12	2000	2000	58	33	0.125	0.122
(b)	0.13	5000	2500	63	30	0.136	0.133
(c)	0.14	5000	2500	63	28	0.150	0.143
(d)	0.16	8000	5000	91	24	0.173	0.162
(e)	0.24	8000	5000	91	15	0.274	0.240
(f)	0.22	12000	7500	108	17	0.238	0.216
(g)	0.37	18000	9500	227	0	0.396	0.369
(h)	0.45	18000	9500	227	0	0.227	0.447
(i)	0.79	25000	10000	304	0	1.074	0.787

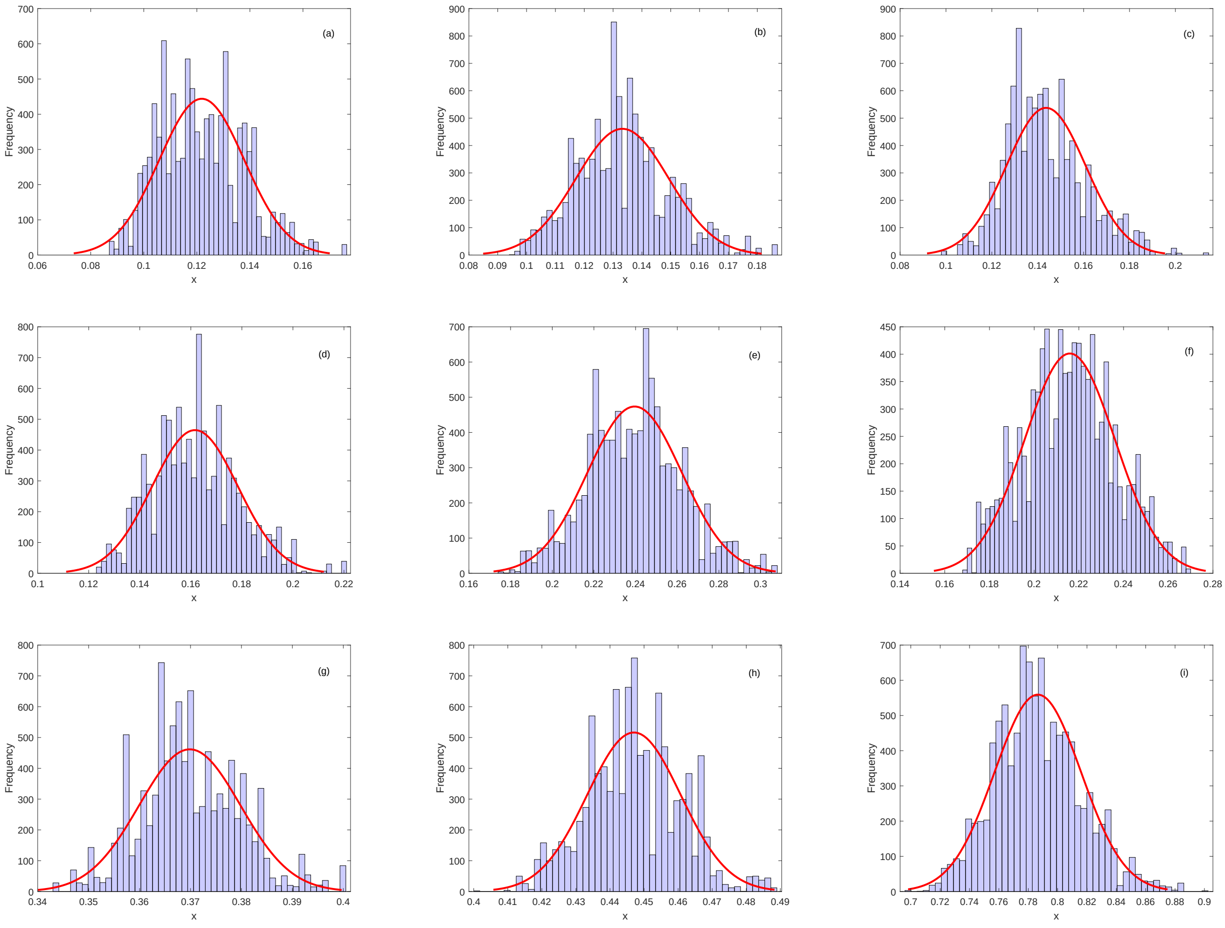

Figure 2: The simulating visualizations of Bayes-MCMC in 9 cases.

In Table 2, we take advantage of the control variates method to hold fairness of the comparison between the two methods(Often in physics). Without loss of generality, it shows that the Bayes-MCMC gives a more effective estimation. In practice, although that's enough, we find the Bayes-MCMC only produces good results for approximately $\lambda \leq 0.3$ when we truncate the degree at about 80% in complex networks, just $(a) \sim (f)$. This phenomenon is plausible that the exponential networks possess exactly sparse tails of which degree are not obviously separated if λ were greater. For $\lambda \geq 0.3$, in which possess the bad tails, we just rank the degree and take the degree out in accordance with the proportion to estimate the exponent, just $(g) \sim (i)$. And surprisingly, the second way shows much robustness, even for $\lambda \leq 0.3$. One plight is that the exponential networks possess super tails which may bring explosions of calculation when $\lambda \leq 0.1$, requiring the strong computing power.

In Table 3, we evaluate their computing performances as λ in different ranges that MLE shows a slight advantage in computing speed, while MCMC takes an edge in accuracy, even more. And in fact, it is unnecessary to truncate the points in simulating cases because our simulative complex networks are entirely exponential over all degree, but other cases in real world networks.

Table 3: The empirical judgment of MLE and MCMC methods as λ in different ranges through large amounts of experiments. '$\surd$' indicates that the method effectively produces plausible exponents, even robust; '$-$' indicates that the method produces general exponents with fluctuation; '$\times$' indicates that the method always produces bad exponents.

λ	MLE	$MCMC_{truncate}$	$MCMC_{proportion}$
$0 < \lambda \leq 0.1$	$-$	$\times$	$\surd$
$0.1 < \lambda \leq 0.3$	$-$	$\surd$	$\surd$
$0.3 < \lambda$	$\times$	$\times$	$\surd$

In real world networks, we find λ always locate in the range $0 \leq \lambda \leq 1.0$, clustering below 0.3, though there are occasional exceptions. The pattern of manifestation is small-world attribution which turns out to be widespread in biological, social man-made systems etc, referring to [24].

Remark 1 *The above λ we estimate are exponents with regard to the weights, but the exponents of the exponential complex networks approximately satisfy $\lambda_{degree} = log(\lambda + 1)$, that is. Do not fixate on the format, because we have acquired the accurate one, more in Sec.2.*

4.2 Application to the real world data

We apply the Bayes-MCMC method we describe to the following data sets from various real world complex networks, including physics, biology, correspondence and transportation. The exponential hypothesis turns out to be, statistically speaking if possible, a reasonable description of the data. In other words, the data are compatible with the hypothesis that they are drawn from an exponential distribution, though do not exclude contingency. We collect 4 data sets,as follows:

(A)Arxiv High Energy Physics Theory category network[25, 26]

(B)The human disease network[25, 27]

(C)The Email Network of University at Rovira iVirgili[25, 28]

(D)The North American Power Grid Network[24, 25]

Given observed data sets and exponential hypotheses from which the data are drawn, we intend to figure out whether our hypotheses are plausible. The standard approach is Kolmogorov-Smirnov test which conducts goodness-of-fit curves and generates p-values that quantify the plausibility of the hypotheses,

$$H0 : F = F_0 \qquad H1 : F \neq F_0$$

Once we calculated p-values, we intend to decide whether they are small enough to reject exponential hypotheses, or plausible that complex networks follow the exponential. In our experiments, we make a bit conservative decisions that the exponential hypotheses are merely rejected when $p \leq 0.05$, where complex networks poorly fitting in the exponential model are rare events—Reject H0.

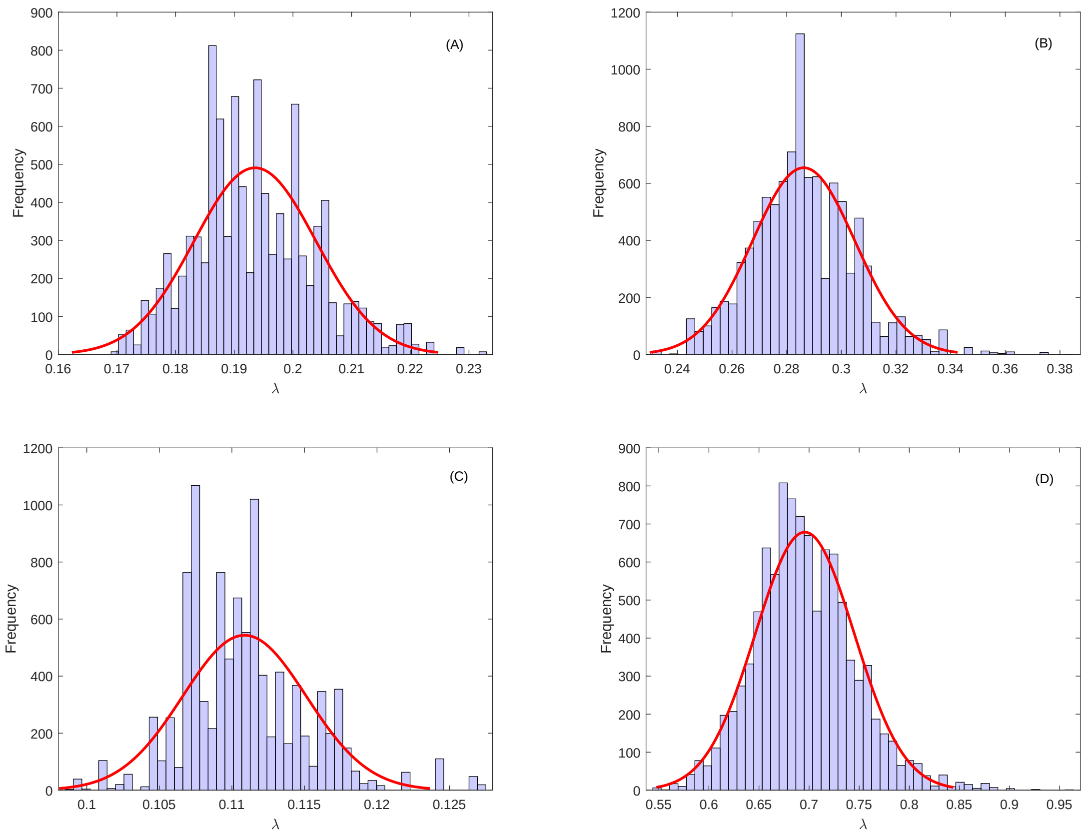

Figure 3: The empirical visualizations of Bayes-MCMC in 4 cases.

Table 4: Bayes-MCMC method on real world data sets.

No.	Data set	Nodes	x_{min}	n_{tail}	λ	λ_{degree}	p
(A)	Ca-Hepth	68745	10	1433	0.194	0.177	0
(B)	Bio-Deseasome	516	6	134	0.286	0.252	0.03
(C)	ENURV	1133	14	241	0.111	0.105	0.92
(D)	NAPGN	4941	1	4941	0.695	0.528	0.21

Based on the results, we account for the exponential models are great fitting to the data sets (C)(D), simultaneously meaning that the degree of complex networks follow the exponential, but not (A)(B). Although not enough evidence for exactly exponential, we just find the graphics of Ca-Hepth and Bio-Dieasome exhibit the exponential-like degree distribution, may be as references.

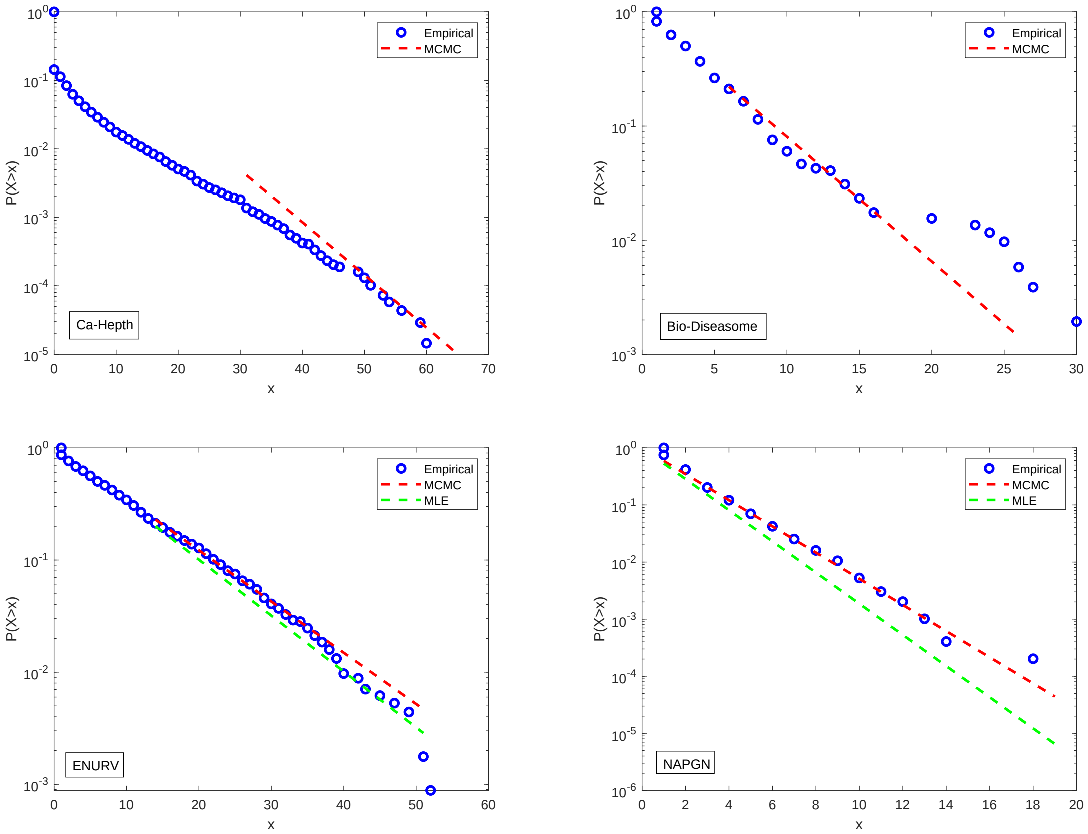

Figure 4: Bayes-MCMC and MLE on CDFs $P(x)$ fitting in 4 empirical data sets.

Remark 2 *In the final Kolmogorov-Smirnov test, it is taken for granted that large amounts of samples always bring anomaly points which screw up the whole test. We generally sample from the fitting curve to conduct the test, or the partial ways.*

5 Conclusion and prospect

In conjunction with the theory of GRG and Bayes, we provide a cutting edge view to study the exponent of exponential degree distribution complex networks. Without loss of generality, Bayes-MCMC method produces exactly effective and robust results, and that is an advancing step.

Nonetheless, this paper comes up with two prospects: First, there is no criterion defining how a good tail is. Second, the optimal exponent λ is going to theoretically cluster which presents the highest cell in MCMC, though it is not obvious in our work.

Reference

[1] Anna D. Broido and Aaron Clauset. Scale-free networks are rare. *Nature Communications*, 10(1):1017–1027, 2019.

[2] A. Lachgar and A. Achahbar. Network growth with preferential attachment and without ”rich get richer” mechanism. *International Journal of Modern Physics C*, 27(2), 2016.

[3] Liu Jing-Zhou and Tang Yi-Fa. An exponential distribution network. *Chinese Physics*, 14(4):643–645, mar 2005.

[4] Zhengwen Lai, Wenjun Xiao, Meisheng Li, and Zhanying Zhang. An exponential distribution complex network model constructed by degree sequence length iteration. volume 1, pages 267–271, 2017.

[5] R. Guimerà, L. Danon, A. Díaz-Guilera, F. Giralt, and A. Arenas. Self-similar community structure in a network of human interactions. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 68(6):651031–651034, 2003.

[6] Lav R. Varshney, Beth L. Chen, Eric Paniagua, David H. Hall, and Dmitri B. Chklovskii. Structural properties of the caenorhabditis elegans neuronal network. *PLoS Computational Biology*, 7(2):e1001066, 2011.

[7] Réka Albert, István Albert, and Gary L. Nakarado. Structural vulnerability of the north american power grid. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 69(2), 2004.

[8] Deng Wei-Bing, Guo Long, Li Wei, and Cai Xu. Worldwide marine transportation network: Efficiency and container throughput. *Chinese Physics Letters*, 26(11):118901, 2009.

[9] Pjm statement commends the u.s.-canada power system outage task force, 2003.

[10] Christopher W. Anderson, Joost R. Santos, and Yacov Y. Haimes. A risk-based input-output methodology for measuring the effects of the august 2003 northeast blackout. *Economic Systems Research*, 19(2):183–204, 2007.

[11] Xiangying Xie, Dayan Ma, Yudong Lu, and Xin Luo. Improvement strategy for robustness of power grid complex network by distributed photovoltaic power system. *E3S Web of Conferences*, 143:2019, 2020.

[12] Paolo Crucitti, Vito Latora, and Massimo Marchiori. A topological analysis of the italian electric power grid. *Physica A-statistical Mechanics and Its Applications*, 338(1-2):92–97.

[13] Adilson E. Motter and Ying-Cheng Lai. Cascade-based attacks on complex networks. *Physical Review E Statistical Nonlinear and Soft Matter Physics*, 66(6):065102.

[14] Yihong Hu and Daoli Zhu. Empirical analysis of the worldwide maritime transportation network. *Physica A: Statistical Mechanics and its Applications*, 388(10):2061–2071, 2009.

[15] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[16] A.-L. Barab´asi R. Albert, H. Jeong. *Nature*, 401(130), 1999.

[17] M. L. Goldstein, S. A. Morris, and G. G. Yen. Problems with fitting to the power-law distribution. *The European Physical Journal B*, 41(2):255–258, 2004.

[18] Aaron Clauset, Cosma R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009.

[19] Mohammad R. Mesbahi, Amir M. Rahmani, and Mehdi Hosseinzadeh. Highly reliable architecture using the 80/20 rule in cloud computing datacenters. *Future Generation Computer Systems*, 77:77–86, 2017.

[20] S. Antunovic, T. Kokan, T. Vojkovic, and D. Vukicevic. Exponential generalised network descriptors. *ADVANCES IN MATHEMATICS OF COMMUNICATIONS*, 13(3):405–420, 2019.

[21] Tom Britton, Maria Deijfen, and Anders Martin-Löf. Generating simple random graphs with prescribed degree distribution. *Journal of Statistical Physics*, 124(6):1377–1397, 2006.

[22] Remco van der Hofstad. *Random graphs and complex networks*, volume 1. 2016.

[23] Siddhartha Chib and Edward Greenberg. Understanding the metropolis-hastings algorithm. *The American Statistician*, 49(4):327–335, 1995.

[24] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.

[25] Ryan A. Rossi and Nesreen K. Ahmed. The network data repository with interactive graph analytics and visualization. In *AAAI*, 2015.

[26] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 177–187. ACM, 2005.

[27] Kwang-Il Goh, Michael E Cusick, David Valle, Barton Childs, Marc Vidal, and Albert-László Barabási. The human disease network. *Proceedings of the National Academy of Sciences*, 104(21):8685–8690, 2007.

[28] A. Diaz-Guilera F. Giralt R. Guimera, L. Danon and A. Arenas. *Physical Review E*, 68:065103(R), 2003.