DISEI - Università degli Studi di Firenze

Working Paper Series - Economics

Statistical Equilibrium Models for Sparse Economic Networks

L . Bargigli

Working Paper N. 25/2013

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze, Italia
www.disei.unifi.it

The findings, interpretations, and conclusions expressed in the working paper series are those of the authors alone. They do not represent the view of Dipartimento di Scienze per l'Economia e l'Impresa, Università degli Studi di Firenze

Statistical equilibrium models for sparse economic networks *

Leonardo Bargigli[a]

[a] Dipartimento di Scienze per L'Economia e l'Impresa
Via delle Pandette 32, 50127 Firenze (Italy)
leonardo.bargigli@unifi.it

November 25, 2013

Abstract

Real markets can be naturally represented as networks, and they share with other social networks the fundamental property of sparsity, whereby agents are connected by $l = \mathrm{O}(n)$ relationships. The exponential networks model introduced by Park and Newman can be extended in order to deal with this property. When compared with alternative statistical models of a given real network, this extended model provides a better statistical justification for the observed network values. Consequently, it provides more reliable maximum entropy estimates of partially known networks than previously known ME techniques.

1 Introduction

Since the outbreak of the global crisis, major policy considerations have elicited an increased push towards better models of large scale market interaction. These models should be able to cope both with agents' heterogeneity and with their reciprocal interdependency, in order to provide a better agreement with empirical evidence and a greater forecasting ability. With respect to these efforts, the topological properties of markets, such as the density of links between economic agents, generally play a puzzling role. The standard

*The author acknowledges the financial support from the European Community Seventh Framework Programme (FP7/2007-2013) under Socio-economic Sciences and Humanities, grant agreement no. 255987 (FOC-II).

microeconomic perspective, in fact, dictates that the more diversification, the better for economic agents [1]. This statement requires economic networks, and especially financial networks like interbank credit markets, to be dense, i.e. to have a number of connections or *connectivity* $l = \mathrm{O}(n^2)$ while in many cases they are found to be sparse, i.e. to have a connectivity $l = \mathrm{O}(n)$ (see, for instance, [10] or [5]).

Specific hypotheses, such as costly information gathering or sunk costs, must be adopted in order to explain why economic agents may prefer to interact only with a small neighborhood. These hypotheses may not be fully satisfactory in many circumstances, e.g. when the goods offered on the market are almost perfect substitutes (all the more so when they are actually identical like in the credit market), or when information on market participants is publicly available. In this paper I adopt the complementary view of building a class of statistical equilibrium market models which are expected to have a given connectivity $\bar{l}$. Statistical means that the commodity flow w_{ij} between any couple of agents (i, j) is viewed as the realization of a random variable, defined over a discrete nonnegative domain. The market as a whole is nothing more than the collection of all these variables, that can be represented as a random matrix W with entries which are statistically independent but non necessarily equally distributed. Each realization of W represents a possible state or configuration of the market, and the collection of all these market states, together with a probability distribution over states $P(W)$, is called a *statistical ensemble*. Equilibrium means instead that, if the market is allowed to relax without external disturbances, it will converge to the stable probability distribution $P^*(W)$ which is obtained by solving the model itself footnoteThere is a bijective relationship between a given model and the corresponding ensemble, since solving a given model means actually to find the particular P^* which is consistent with the constraints of that model. For this reason the terms "model" and "market ensemble" or, as we will see below, "network ensemble" can be used as equivalents..

Starting from this background, the transition to complex networks theory is very natural. In fact, the representation of markets outlined above is nothing different from the matrix representation of a network G [1]. Thanks to this parallelism, we can take advantage of previous contributions from the complex networks literature. In particular, Park and Newman [14] have proposed

[1]For this reason, in this paper the terms "market" and "network" are used as equivalents even if, strictly speaking, a market is a weighted network. If G is a binary network, its links can take only binary values, and thus its matrix representation is given by the matrix A with binary entries. For convenience, in this paper binary and weighted networks are labeled respectively fermionic and bosonic networks by analogy with the terminology of statistical physics [14].

a general methodology for building ensembles of networks, with a fixed number n of nodes, satisfying linear and non linear constraints over the expected values of network observables. They show, in analogy with equilibrium statistical physics, that P^* is a Boltzmann-Gibbs probability distribution over network configurations. Since the Boltzmann-Gibbs distribution belongs to the exponential family of probability distributions, networks belonging in the resulting ensemble are labeled as *exponential networks*.

Park and Newman provide a solution for exponential networks when the constraints are represented either by the degree or by the strength distributions [2]. Given our previous market representation, strengths represent the expected supply or demand of the agents, while degrees represent their expected number of buy or sell relationships with other agents. From the definition of connectivity (see note 2), we see that only in the latter case l is fixed. Thus we are left with the unsatisfactory alternative between a bosonic model with uncontrolled connectivity and a fermionic model that can't describe, by construction, a market configuration. The authors of [9] have shown that it is possible to overcome this limitation, i.e. to derive a statistical ensemble where both the expected strength and degree distributions are fixed.

The rest of the paper is organized as follows. In sec. 2 I recall the main results presented in [14]. The critical step for any empirical application of exponential networks is to find the solution of a large non linear system of equations, which provides the values of the Lagrange multipliers acting as parameters for P^*. For this reason, in the same section I show how it is possible, by employing iteratively standard Newton methods from a suitable starting point, to compute the solution of such system. In sec. 3, employing the results of [9], I present two models. The first one is an intermediate model satisfying both a constraint over connectivity and a constraint over the expected strength distribution (sec. 3.1). Subsequently, in sec. 3.2 I present the complete model of [9], where both the strength and degree distributions are used as constraints. Sec. 4 is dedicated to a comparison of alternative models as tools for explaining a given observed network. Finally, sec. 5 concludes.

---

[2]By strength of a node i in a bosonic network we define the sum $w_i = \sum_{j \neq i} w_{ij}$. If W is asymmetric, i.e. G is directed, we need to distinguish between the out-strength and in-strength of the node i. The degree of a node i, instead, is defined over the binary matrix A as the sum $d_i = \sum_{j \neq i} a_{ij}$. If A is asymmetric, again we need to distinguish between the out-degree and in-degree of the node i. From the definition of d_i we get $l = \sum_i d_i$. Thus connectivity is a function of nodes' degrees. In the case of bosonic networks, it's useful to define $v = \sum_i w_i$, where v is said to be the *volume* of G.

2 Exponential Networks

The rationale of the approach of Park & Newman is straightforward: they seek a systematic way to generate random graphs displaying a set of desired properties $\{x_i\}$. Usually these properties are taken from a real network: for instance, we would like to test if a random network with those properties displays other observed properties of the real network. In principle, their approach can be employed to isolate the fundamental properties of an observed network, which can be defined as those that cannot be justified by random interaction alone [3]. For instance, it is well known from the complex network literature that the degree and strength distributions that we generally observe in real networks cannot be derived from a statistical equilibrium model. That's the main reason why these distributions usually enter into the set of desired properties mentioned above, acting as constraints in the model.

Since we work with random networks, our observables are naturally formulated in terms of statistics computed over the network ensemble. In the simplest case, we wish to equalize the ensemble average $\langle x_i \rangle$ with some empirical estimate of such average $\bar{x}_i$. Since the observables depend on network realizations, we need to weight the ensemble average against the probability $P(G)$ of observing a given realization G:

$$\langle x_i \rangle = \sum_{G \in \mathcal{G}} P(G) x_i(G) = \bar{x}_i \tag{1}$$

Since the $x_i(G)$ are a given, we need to specify a parameter dependent functional shape of $P(G)$ in order to solve the system. By adopting the basic concepts of equilibrium statistical mechanics we obtain a solution for this task by maximizing the following Lagrangean:

$$\mathcal{L} = S + \lambda(1 - \sum_G P(G)) + \sum_i \theta_i \left(\bar{x}_i - \sum_G P(G) x_i(G) \right) \tag{2}$$

Gibbs entropy $S = -\sum_G P(G) \ln P(G)$ is maximized, under the given constraints, for the distribution satisfying $\partial \mathcal{L} / \partial P(G) = 0$, i.e.

$$\ln P(G) + 1 + \lambda + \sum_i \theta_i x_i(G) = 0 \tag{3}$$

Rearranging and taking antilogs:

[3] In this sense, statistical ensembles act as null models for real networks (see below, sec. 4). A subsequent elaboration in this direction is given by [15] and the related analysis of world trade [6].

$$P(G) = \frac{e^{-H(G)}}{Z} \tag{4}$$

where $H(G) \equiv \sum_i \theta_i x_i(G)$ is the graph Hamiltonian which, thanks to matrix representation of G (see sec. 1), can be rewritten in terms of the matrix W or A. $Z \equiv e^{(\lambda+1)}$ is the partition function: from the normalization constraint we easily obtain that $Z = \sum_G e^{-H(G)}$. Once we have obtained the functional shape of P^* with the help of eq. (4), the model is said to be solved when the values of the parameters $\{\theta_i\}$, which fully determine P^*, are obtained from the system (1). In fact, it is possible to show that, if we adopt the Boltzmann-Gibbs distribution (4), then the system (1) provides the maximum likelihood estimates for the parameters $\{\theta_i\}$ [8].

When the constrained observables are the out- and in-degree distributions of a fermionic directed network, the main quantities of the exponential model read:

$$\begin{aligned}
H(G) &= \sum_i \sum_{j\neq i} [(\lambda_i + \theta_j) a_{ij}] = \sum_i \sum_{j\neq i} \Lambda_{ij} a_{ij} \\
Z &= \prod_i \prod_{j\neq i} \left(1 + e^{-\Lambda_{ij}}\right) \\
F &= -\ln Z = -\sum_i \sum_{j\neq i} \ln\left(1 + e^{-\Lambda_{ij}}\right) \\
p_{ij} &= \langle a_{ij} \rangle = \frac{\partial F}{\partial \Lambda_{ij}} = \frac{1}{e^{\Lambda_{ij}} + 1}
\end{aligned}$$

Since the a_{ij} are bernoullian, fermionic exponential networks may be labeled more compactly as *Bernoulli networks*. When the constrained observables are the out- and in-strength distributions of a bosonic directed network, the main quantities of the exponential model read:

$$
\begin{aligned}
H(G) &= \sum_i \sum_{j\neq i} \left[(\lambda_i + \theta_j) w_{ij}\right] = \sum_i \sum_{j\neq i} \Lambda_{ij} w_{ij} \\
Z &= \prod_i \prod_{j\neq i} \sum_{w_{ij}=0}^{\infty} e^{-\Lambda_{ij} w_{ij}} = \prod_i \prod_{j\neq i} \left(\frac{1}{1 - e^{-\Lambda_{ij}}}\right) \\
F &= \sum_i \sum_{j\neq i} \ln\left[1 - e^{-\Lambda_{ij}}\right] \\
\langle w_{ij} \rangle &= \frac{\partial F}{\partial \Lambda_{ij}} = \frac{e^{-\Lambda_{ij}}}{1 - e^{-\Lambda_{ij}}}
\end{aligned}
$$

It's also possible to write down $P^*(G)$:

$$
P^*(G) = \prod_i \prod_{j\neq i} \left(1 - e^{-\Lambda_{ij}}\right) e^{-\Lambda_{ij} w_{ij}} = \prod_i \prod_{j\neq i} P^*(w_{ij}) \tag{5}
$$

It is easy to see that the $P^*(w_{ij})$ are non identical but independent geometric distributions with $p_{ij} = 1 - e^{-\Lambda_{ij}}$. For this reason in the following we label bosonic exponential networks more compactly as *geometric networks*. Substituting the last equation into the constraints we obtain the following specialization of system (1):

$$
\begin{aligned}
\sum_{j\neq i} \frac{x_i y_j}{1 - x_i y_j} &= \bar{w}_i \quad \forall i \in [1, \dots n] \\
\sum_{j\neq i} \frac{y_i x_j}{1 - y_i x_j} &= \bar{w}_i \quad \forall i \in [n+1, \dots 2n]
\end{aligned}
$$

where $x_i = e^{-\lambda_i} (\forall i \in [1, \dots n])$ and $y_i = e^{-\theta_{i-n}} (\forall i \in [n+1, \dots 2n])$.

When we use numerical optimization with standard Newton methods to solve the system, it's very likely that solutions with $\langle w_{ij} \rangle < 0$ (i.e. $x_i y_j > 1$) are selected. In order to avoid this problem it's helpful to start with an approximate solution such that $x_i y_j < 1 \quad \forall i, j$. If $x_i y_j \ll 1$ the following *sparse limit approximation* holds:

$$
\frac{x_i y_j}{1 - x_i y_j} \approx x_i y_j \tag{6}
$$

In this case the system may be rewritten

$$\sum_{j\neq i} x_i y_j = \bar{w}_i \quad \forall i \in [1, \dots n] \tag{7}$$

$$\sum_{j\neq i} y_i x_j = \bar{w}_i \quad \forall i \in [n+1, \dots 2n] \tag{8}$$

It is also well known that we can derive an explicit solution of the approximated system if we allow for self-loops. In this case by adding equivalently one of the two set of constraints we obtain (see note 2)

$$\sum_i x_i \sum_j y_j = \bar{v} \tag{9}$$

Substituting into the ith constraint

$$x_i \frac{\bar{v}}{\sum_i x_i} = \bar{w}_i \quad \forall i \in [1, \dots n]$$

$$y_i \sum_i x_i = \bar{w}_i \quad \forall i \in [n+1, \dots 2n]$$

Substituting into the approximated expression for $\langle w_{ij} \rangle$ we obtain

$$x_i y_j = \frac{\bar{w}_i \bar{w}_j}{\bar{v}} \qquad \forall i \in [1, \dots n] \wedge \forall j \in [n+1, \dots 2n] \tag{10}$$

from which

$$x_i = \frac{\bar{w}_i}{\sqrt{\bar{v}}} \quad \forall i \in [1, \dots n]$$

$$y_j = \frac{\bar{w}_j}{\sqrt{\bar{v}}} \quad \forall j \in [n+1, \dots 2n]$$

In order to obtain that $x_i y_j \ll 1$, if we are not lucky enough to have a distribution that directly provides the result, we can divide both sides of the system (7) - (8) by a constant k [4]. Then using the approximation (6) we obtain the new system

[4] The sparse limit approximation holds also for Bernoulli networks. In this limit, the exponential fermionic model is coincident with the so-called expected degree model [3]. On the other hand, the former model is always more general than the latter in the sense that it does not impose constraints on the degree distribution unless we wish to work in the sparse limit. In fact, since eq. (10) in the fermionic case becomes $a_{ij} = \frac{\bar{d}_i \bar{d}_j}{\bar{l}}$, we see that in order for a_{ij} to be bernoullian we need to suppose $\bar{d}_i \bar{d}_j \leqslant \bar{l} \quad \forall (i,j)$. This constraint is always required by the expected degree model. For an deeper analysis of this issue, see [8] and [15].

$$k\sum_{j\neq i}\frac{x_i y_j}{1-x_i y_j}=\bar{w}_i \quad \forall i\in[1,\dots n]$$
$$k\sum_{j\neq i}\frac{y_i x_j}{1-y_i x_j}=\bar{w}_i \quad \forall i\in[n+1,\dots 2n]$$

After having computed numerically the solution of this system, we may reach the solution of the original one by iterating for $k\downarrow 1$.

3 Conditional Geometric Networks

In [9] the exponential model is extended in order to comply with constraints where both the expected strength and degree distributions are fixed. In this section, I review these results in order to define two distinct but related network ensembles. In the first ensemble the expected strength distribution and the expected connectivity of a bosonic network are fixed. The second ensemble is the one of [9], where both the strength and degree distributions are fixed. Since $P^*(w_{ij}|a_{ij})$ still follows a geometric distribution in these models, they can be labeled as *conditional geometric networks*. Although for expository convenience we present the results for the undirected case, they can be immediately extended to directed networks. As an example, we provide the solution of these models using data taken from a real directed bosonic network.

3.1 Fixed sparsity and strength distributions

The main quantities for the model with fixed strength distribution and sparsity read:

$$
\begin{aligned}
H(G) &= \sum_{i>j} [(\lambda_i + \lambda_j) w_{ij} + \theta a_{ij}] = \sum_{i>j} [\Lambda_{ij} w_{ij} + \theta a_{ij}] \\
Z &= \prod_{i>j} \sum_{w_{ij}=0}^{\infty} e^{-[\Lambda_{ij} w_{ij} + \theta a_{ij}]} = \prod_{i>j} \left(1 + \sum_{w_{ij}=1}^{\infty} e^{-[\Lambda_{ij} w_{ij} + \theta]} \right) = \\
&= \prod_{i>j} \left(1 + e^{-\theta} \frac{e^{-\Lambda_{ij}}}{1 - e^{-\Lambda_{ij}}} \right) \\
F &= - \sum_{i>j} \ln \left[1 + e^{-\theta} \frac{e^{-\Lambda_{ij}}}{1 - e^{-\Lambda_{ij}}} \right] \\
\langle w_{ij} \rangle &= \frac{\partial F}{\partial \Lambda_{ij}} = \frac{e^{-\theta}}{1 + (e^{-\theta} - 1)\, e^{-\Lambda_{ij}}} \quad \frac{e^{-\Lambda_{ij}}}{1 - e^{-\Lambda_{ij}}} \\
\langle l \rangle &= \frac{\partial F}{\partial \theta} = \sum_{i>j} \frac{e^{-(\Lambda_{ij} + \theta)}}{1 + (e^{-\theta} - 1)\, e^{-\Lambda_{ij}}} = \sum_{i>j} \langle a_{ij} \rangle
\end{aligned}
$$

where

$$
a_{ij} = \begin{cases} 1 & \text{if } w_{ij} > 0 \\ 0 & \text{if } w_{ij} = 0 \end{cases}
$$

For this model the system (1) becomes

$$
\sum_{j \neq i} \frac{y}{1 + (y-1)\, x_i x_j} \quad \frac{x_i x_j}{1 - x_i x_j} = \bar{w}_i \quad \forall i \in [1, \dots n] \tag{11}
$$

$$
\sum_i \sum_{j \neq i} \frac{x_i x_j y}{1 + (y-1)\, x_i x_j} = \bar{l} \tag{12}
$$

which may be solved numerically to provide the values of the parameters. The equilibrium probability distribution is

$$
\begin{aligned}
P^*(W \cap A) &= \prod_{i>j} \frac{e^{-[\Lambda_{ij} w_{ij} + \theta a_{ij}]}}{Z} = \\
&= \prod_{i>j} \frac{1 - e^{-\Lambda_{ij}}}{1 + (e^{-\theta} - 1) e^{-\Lambda_{ij}}} e^{-[\Lambda_{ij} w_{ij} + \theta a_{ij}]} \\
&= \prod_{i>j} P^*(w_{ij} \cap a_{ij})
\end{aligned}
$$

From the latter expression we derive $P^*(w_{ij}|a_{ij})$:

$$P^*(w_{ij}|a_{ij}) = \frac{P^*(w_{ij} \cap a_{ij})}{P^*(a_{ij})} = = \frac{P^*(w_{ij} \cap a_{ij})}{\langle a_{ij} \rangle}$$
$$= \begin{cases} (1 - e^{-\Lambda_{ij}})e^{-[\Lambda_{ij}(w_{ij}-1)]} & \text{if } a_{ij} = 1 \\ \delta(w_{ij} - 0) & \text{if } a_{ij} = 0 \end{cases}$$

Thus, if $a_{ij} = 1$ we still obtain a geometric distribution with $w_{ij} \in \{1, 2, \dots\}$ and conditional expectation

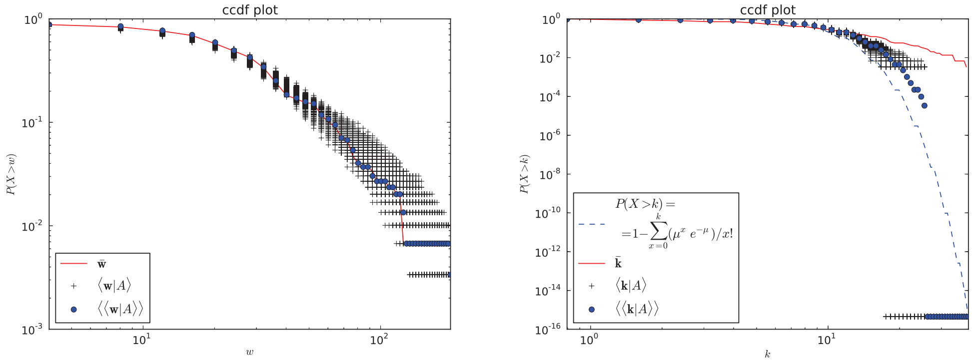
$$\langle w_{ij}|a_{ij} \rangle = \frac{\langle w_{ij} \rangle}{\langle a_{ij} \rangle} = \begin{cases} \dfrac{1}{1 - e^{-\Lambda_{ij}}} & \text{if } a_{ij} = 1 \\ 0 & \text{if } a_{ij} = 0 \end{cases} \tag{13}$$

We remark that the $\langle a_{ij} \rangle$ obtained by solving conditional geometric models define fermionic models which are different from the Bernoulli networks of sec. 2. These fermionic networks are labelled as *conditional Bernoulli networks*.

In order to test this model, the system (11) - (12) was solved taking as parameters the connectivity and the out- and in-strength distributions of the neural network of the nematode C. Elegans, compiled by Watts & Strogatz [18]. Fig. (1) shows that the resulting expected out-strength distribution follows the original strength data, while the out-degree distribution is placed, so to speak, halfway between a Poisson distribution and the original degree data. Analogous behavior is obtained for the in-strength and in-degree distributions.

3.2 Fixed strength and degree distributions

The main quantities for the model of [9] read:



(a) (b)

Figure 1: Out-strength distribution (Panel a) and out-degree distribution (Panel b). In both panels the red continuous lines represent the ccdf computed from original data of the C.Elegans dataset, the black crosses represent the ccdf computed from the conditional expectations $\langle \mathbf{w}|A\rangle$ given by eq. (13), the blue circles represent the ccdf computed from the average of the different $\langle \mathbf{w}|A\rangle$.

$$H(G) = \sum_{i>j} \left[(\lambda_i + \lambda_j) w_{ij} + (\theta_i + \theta_j)\, a_{ij}\right] = \sum_{i>j} \left[\Lambda_{ij} w_{ij} + \Theta_{ij} a_{ij}\right]$$

$$Z = \prod_{i>j} \sum_{w_{ij}=0}^{\infty} e^{-[\Lambda_{ij} w_{ij} + \Theta_{ij} H(w_{ij})]} = \prod_{i>j} \left(1 + \sum_{w_{ij}=1}^{\infty} e^{-[\Lambda_{ij} w_{ij} + \Theta_{ij}]}\right) =$$

$$= \prod_{i>j} \left(1 + e^{-\Theta_{ij}} \frac{e^{-\Lambda_{ij}}}{1 - e^{-\Lambda_{ij}}}\right)$$

$$F = -\sum_{i>j} \ln \left[1 + e^{-\Theta_{ij}} \frac{e^{-\Lambda_{ij}}}{1 - e^{-\Lambda_{ij}}}\right]$$

$$\langle w_{ij}\rangle = \frac{\partial F}{\partial \Lambda_{ij}} = \frac{e^{-\Theta_{ij}}}{1 + \left(e^{-\Theta_{ij}} - 1\right) e^{-\Lambda_{ij}}} \quad \frac{e^{-\Lambda_{ij}}}{1 - e^{-\Lambda_{ij}}}$$

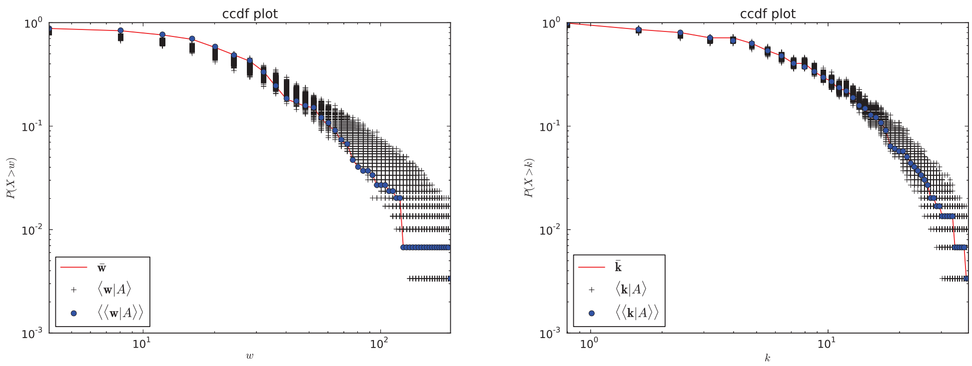
$$\langle a_{ij}\rangle = \frac{\partial F}{\partial \Theta_{ij}} = \frac{e^{-(\Lambda_{ij} + \Theta_{ij})}}{1 + \left(e^{-\Theta_{ij}} - 1\right) e^{-\Lambda_{ij}}}$$

where a_{ij} is defined as in the previous model. From the last equations we may derive the following non-linear system

$$\sum_{j\neq i} \frac{y_i y_j}{1+(y_i y_j-1)\,x_i x_j} \quad \frac{x_i x_j}{1-x_i x_j} = \bar{w}_i \quad \forall i \in [1,\dots n] \tag{14}$$

$$\sum_{j\neq i} \frac{x_i x_j y_i y_j}{1+(y_i y_j-1)\,x_i x_j} = \bar{k}_i \quad \forall i \in [1,\dots n] \tag{15}$$

which may be solved numerically to provide the values of the parameters. It's easy to see that the expressions for $P(w_{ij} \cap a_{ij})$, $P(w_{ij}|a_{ij})$ and $\langle w_{ij}|a_{ij}\rangle$ have the same shape of the previous model where Θ_{ij} is substituted for θ. The system (14) - (15) was solved using data taken from the C.Elegans dataset. Fig. (2) shows that the resulting expected out-strength and out-degree distributions follow the original data. Analogous behavior is obtained for the in-strength and in-degree distributions.

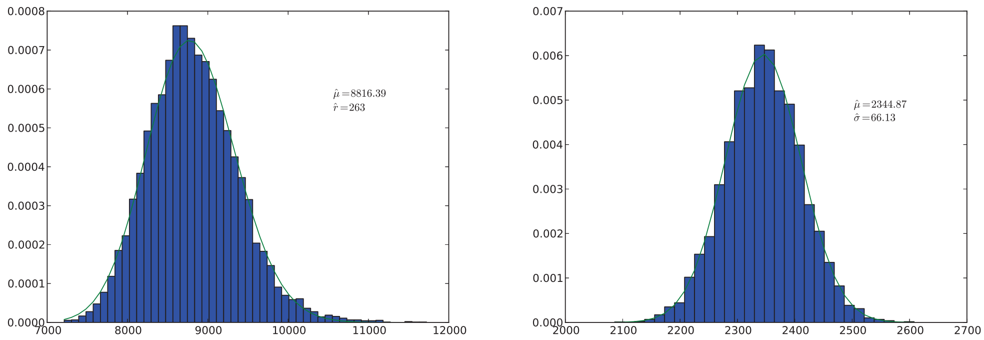


(a) (b)

Figure 2: Out-strength distribution (Panel a) and out-degree distribution (Panel b). In both panels the red continuous lines represent the ccdf computed from original data of the C.Elegans dataset, the black crosses represent the ccdf computed from the conditional expectations $\langle \mathbf{w}|A\rangle$ given by eq. (13), the blue circles represent the ccdf computed from the average of the different $\langle \mathbf{w}|A\rangle$.

From Fig.(3) we see that l, when treated as independent variable, is well approximated by a normal distribution, while v, when the strengths are conditioned to a given topology, is approximated by a negative binomial, although not perfectly since the w_{ij} are independent but non identically distributed geometric variables. Also the equilibrium distribution is easily adapted from the previous model:

$$P^*(W \cap A) = \prod_{i>j} P^*(w_{ij} \cap a_{ij}) = \prod_{i>j} \frac{1 - e^{-\Lambda_{ij}}}{1 + (e^{-\Theta_{ij}} - 1)e^{-\Lambda_{ij}}} e^{-[\Lambda_{ij} w_{ij} + \Theta_{ij} a_{ij}]} \quad (16)$$



(a) (b)

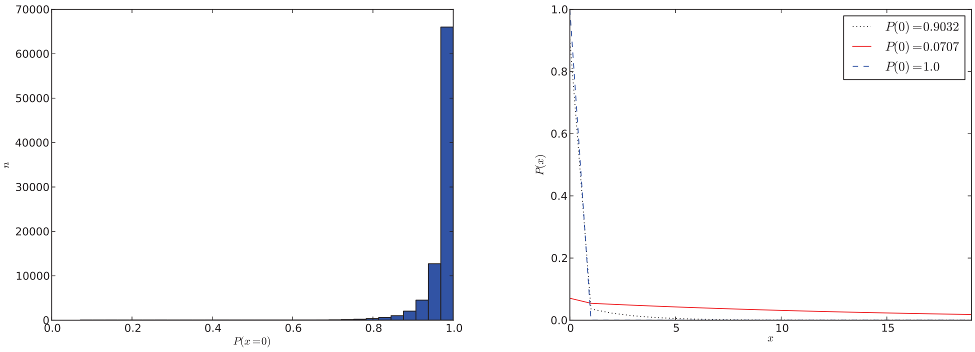
Figure 3: Panel (a): volume (v) distribution with negative binomial fit ($\bar{v} = 8,819$). Panel (b): connectivity (l) distribution with normal fit ($\bar{l} = 2,345$).

Eq.(16) can be used to compute the pmfs $P_{ij}(x)$. From the Panel (a) of Fig. (4) we see that for a large majority of (i,j) we have that $P_{ij}(0) = 1$, a fact which is consistent with the sparsity of the original network. Panel (b) shows instead that the pmfs vary from a delta-like function to a discontinuous geometric-like distribution.

4 Comparison of models

Once we have at hand the solutions of different models for a given observed network, it becomes possible to compare the performance of those models as predictors of the real network from which they are derived. In particular, we would like to have a testbed by which we can select among different models the one which better explains the observed network. Broadly speaking, this task is accomplished either by testing globally the null hypothesis that the observed network belongs to a given statistical ensemble, or by testing simultaneously the null hypotheses that the w_{ij} are distributed according to $P^*(w_{ij})$. In both cases, the statistical ensemble acts as a null model for the observed data [5].

[5] This approach has been recently applied to the analysis of world trade, see [6] and references therein.



(a) (b)

Figure 4: Joint distribution: values of $P_{ij}(x = 0) \quad \forall i, j$ (Panel a); $P_{ij}(x)$ for different values of $P(0)$ (Panel b).

In this comparison, I try to follow both the paths just mentioned. For this purpose I use all the models solved above: Bernoulli (aka exponential fermionic networks); geometric (aka exponential bosonic networks); conditional geometric (CG1) and conditional Bernoulli (CG2) of sec. 3.1; conditional geometric (CG2) and conditional Bernoulli (CG2) of sec. 3.2. Furthermore, I introduce now an additional ensemble which can be labeled as *binomial networks*. In fact, in this ensemble the w_{ij} are binomially distributed with parameters $\bar{v}$ and p_{ij} [2]. In the undirected case, the parameters p_{ij} are obtained by solving the following maximum entropy problem, which is easily adapted to the directed case:

$$\max_{\mathbf{p}} g(\mathbf{p}) = -\sum_{i \neq j} p_{ij} \ln p_{ij} \tag{17}$$

subject to the following constraints:

$$\sum_{j \neq i} p_{ij} = \bar{r}_i$$
$$\sum_i \sum_{j \neq i} p_{ij} = 1$$

where $\bar{r}_i = \bar{w}_i / \bar{v}$.

If we allow for self-loops we can obtain the following explicit solution [2]:

$$p_{ij} = \frac{\bar{w}_i \bar{w}_j}{\bar{v}^2} \tag{18}$$

from which we get

$$\langle w_{ij} \rangle = \frac{\bar{w}_i \bar{w}_j}{\bar{v}} \tag{19}$$

The identity between the latter expression and eq. (10) is remarkable, but there are obvious differences: in the first place, the underlying distribution of w_{ij} is changed from geometric to binomial; in the second place, no approximation is involved here, so that we don't need to impose the constraint $\langle w_{ij} \rangle \ll 1$.

The fundamental property of binomial networks is that their expectation matrix $\langle W \rangle$ is close to a rank-1 matrix, and becomes exactly rank-1 if we allow for self-loops (in fact, in this case, the rows and columns of $\langle W \rangle$ are all linearly dependent). For this reason the model is usually employed in order to define a "community free" expectation of a bosonic, possibly directed, network with given strength distribution. In fact, for the binomial ensemble, Newman's modularity function [13] (which is increasing in the expected strength of communities within the network inasmuch as it is increasing in the rank of $\langle W \rangle$) is expected to be close to zero, with its expectation becoming exactly equal to zero if we allow for selfloops [6].

Strictly speaking, this property is shared neither from the exponential networks of section 2 nor from the conditional geometric networks of section 3. From fig. (5) we see that these models, and especially the one of section 3.2, display a number of relatively large normalized singular values[7] when compared with binomial networks, although in general the behavior of all models looks pretty similar.

Since the $\langle w_{ij} \rangle$, in all these ensembles, are a function only of the strength and degree distributions, the expectation matrices cannot be too different[8]. The Frobenius distance matrices reported in table 1 confirm this claim, since we see that all the normalized expectation matrices $\langle K \rangle$ are equally distant from real data. Further we see that the binomial model is equally distant

[6]Thus, the binomial model is closely related to the growing field of community detection in complex networks. For a detailed review of this topic, see [7].

[7]The normalization is obtained by introducing, e.g. for a undirected bosonic network, the normalized matrix $K = D^{-\frac{1}{2}} W D^{-\frac{1}{2}}$ where D is a diagonal matrix with elements $\{w_1, w_2, \ldots, w_n\}$. It' possible to show that, if $\sigma_0(M)$ denotes the largest singular value of a matrix M, then $\sigma_0(K) = 1\ \forall K$. For fermionic networks normalization becomes $K = D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$, with D diagonal matrix with entries $\{d_1, d_2, \ldots, d_n\}$.

[8]In particular, the normalized expectation matrix $\langle K \rangle$ has in all models the same first projection $K_0 = \sigma_0 u_0 \otimes v_0$, where u_0 and v_0 are the singular vectors associated with σ_0 and $\otimes$ denotes the outer product. In fact, K_0 depends only on the expected strength or degree distributions [2], which remain the same in all models. If the other projections don't contribute too much, we can assume $\langle K \rangle \approx K_0$.

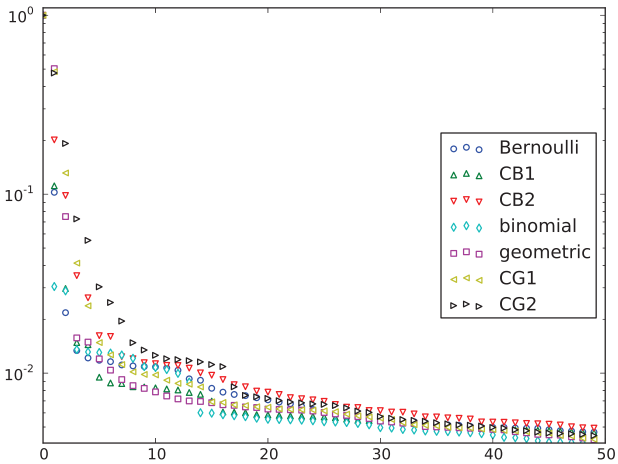


Figure 5: Comparison of the 50 largest normalized singular values of $\langle W \rangle$ across different null models of the C.Elegans network.

from the others, and that the geometric and CM1 models are much closer to each other than the CM2 model. Regarding fermionic models, we see instead that Bernoulli and CB2 networks are closer than CB1 networks. The latter fact is not surprising, since CB1 networks are not constrained by the degree distribution.

As a first approximation, the global approach mentioned above can be pursued by comparing the observed squared deviation of W, defined as $S = \sum_i \sum_{j\neq i}(w_{ij} - \langle w_{ij} \rangle)^2$ with its expected value $\Sigma = \sum_i \sum_{j \neq i} \sigma^2(w_{ij})$ using the Chebyshev inequality

$$P(S \geqslant \lambda) \leqslant \frac{\Sigma}{\lambda} \tag{20}$$

for $\lambda \geqslant \Sigma$. In some cases we can obtain sharper bounds employing the normalized matrix K [9]. Since in this case $k_{ij} \leqslant 1$, we can use the following symmetric inequalities [4]:

$$P(X \leqslant \langle X \rangle - \lambda) = P(X \geqslant \langle X \rangle + \lambda) \leqslant \exp\left(-\frac{\lambda^2}{2(\|X\|^2 + \lambda/3)}\right) \tag{21}$$

[9] Of course, it is always possible to compute sharper p-values once we know at least some approximation of the probability distribution of $X = \sum_i \sum_{j\neq i} w_{ij}$, like for instance we did in fig. 3 (a). Otherwise it's always possible to compute the probability distributions from montecarlo simulations.

	Real data	Binomial	Geometric	CG1	CG2
Real data	0	5.5437	5.5820	5.5773	5.5066
Binomial	5.5437	0	0.5072	0.5020	0.5214
Geometric	5.5820	0.5072	0	0.1171	0.7654
CG1	5.5773	0.5020	0.1171	0	0.7424
CG2	5.5066	0.5214	0.7654	0.7424	0

	Real data	Bernoulli	CB1	CB2
Real data	0	5.3034	5.3234	5.2948
Bernoulli	5.3034	0	0.4036	0.1931
CB1	5.3234	0.4036	0	0.4321
CB2	5.2948	0.1931	0.4321	0

Table 1: Frobenius distance matrix of the normalized expectation matrices $\langle K \rangle$ for the listed bosonic and fermionic models and of real normalized data K from the C.Elegans network.

where $X = \sum_i \sum_{j\neq i} w_{ij}$ and $\|X\| = \sqrt{\sum_i \sum_{j\neq i} \langle w_{ij}^2 \rangle}$.

While the other models follow a known distribution, for the conditional geometric models we need to compute $\langle w_{ij}^2 \rangle$. In the case of CM2 networks the second moment $\langle w_{ij}^2 \rangle$ reads

$$\langle w_{ij}^2 \rangle = \frac{(1+e^{-\Lambda_{ij}})}{1+(e^{-\Theta_{ij}}-1)e^{-\Lambda_{ij}}} \; \frac{e^{-(\Lambda_{ij}+\Theta_{ij})}}{(1-e_{-\Lambda})^2} \tag{22}$$

with straightforward adaptation for the CM1 model (eq. (22) is derived in the appendix).

In table 2 we test the null hypothesis that the C.Elegans network belongs to the ensembles $\mathcal{G}$ defined by the different null models. The p-values are computed using eqs. (20) and (21) with the original and normalized data respectively. We see that, with these approximate bounds, the null hypothesis is rejected for the Binomial model only. It's noteworthy that the test over normalized variables appears to be more powerful only for the latter model while the opposite holds for all the others. What is really interesting in the normalized case is that S is more or less constant, a behavior which is consistent with our previous remark that the expectations of the different models are not very dissimilar, while Σ changes drastically across bosonic models, and gets close to S for the conditional geometric models. This fact shows that the latter provide a better statistical justification of the real network in two ways: i) to a lesser extent, by adapting the $\langle w_{ij} \rangle$ as shown from fig. 5; ii) to a greater extent, by introducing a distribution which assigns a higher probability to the observed w_{ij} against $\langle w_{ij} \rangle$ as we see from table 2. In the latter mechanism the connectivity constraint is essential, as it is shown from the fact that the Bernoulli model, where this constraint holds by construction, performs very well too.

	Original variables (W)			Normalized variables (K)		
Model	S	Σ	pvalue	S	Σ	pvalue
Bernoulli	2,123.42	2,082.77	0.9808	28.1269	27.9893	0.6246
Cond. Bernoulli (1)	2,184.33	2,200.28	-	28.3386	30.9062	0.6493
Cond. Bernoulli (2)	2,105.92	2,078.69	0.9870	28.0353	27.8394	0.6245
Binomial	91,949.37	8,816.54	0.0958	30.7336	6.3428	0.1560
Geometric	165,138.96	98,297.77	0.5952	31.1588	10.9522	0.3056
Cond. Geometric (1)	156,061.16	194,025.69	-	31.1067	30.5524	0.6218
Cond. Geometric (2)	77,860.65	124,797.55	-	30.3234	30.5087	0.6290

Table 2: Statistical tests against alternative null models ($H_0 : G \in \mathcal{G}$, where G is the C.Elegans network).

As it turns out, a more accurate comparison of the null models is obtained by following the second path outlined above, i.e. by using the known probability distributions of the w_{ij} to test at once the null hypotheses that the observed values are drawn from those distributions. This procedure is nothing different from the so-called statistical validation of links, proposed by [16]. Thus, if the null hypothesis is rejected, we say that the link is statistically validated against the given null model. In order to validate the links of an observed network with respect to a given null model, we must handle the CDF $F(x)$ of that model[10]. In fact, a link with integer weight x is validated when $F(x-1) \geqslant 1-\alpha$. Usually α is a threshold defined with the help of Bonferroni correction. In our case, we set $\alpha = \dfrac{0.01}{n(n-1)}$. From table 3 we see that none of the links of the original network are validated against the conditional geometric models, confirming the previous claim that the latter fit better with real data than other models for bosonic networks. Regarding fermionic networks, we observe instead that the Bernoulli model performs in this setting as well as the two more complex alternatives.

[10]For the derivation of CDFs in the case of geometric and conditional geometric models see the appendix.

Model	Uncorrected ($\alpha = 1e^{-2}$)	Bonferroni ($\alpha = 1.14e^{-07}$)
Bernoulli	102	0
Cond. Bernoulli (1)	102	0
Cond. Bernoulli (2)	100	0
Binomial	86,878	275
Geometric	1,016	120
Cond. Geometric (1)	527	0
Cond. Geometric (2)	399	0

Table 3: Statistically validated links of the C.Elegans dataset against different null models.

5 Conclusive Remarks

All the models presented in sec. 2 and 3 rely on the same basic principle of equilibrium statistical mechanics, the maximization of Gibbs entropy. When we maximize the Lagrangean (2), we obtain the most likely distribution P^* which is consistent with the given constraints. In fact, P^* is associated by construction with the greatest number of market configurations. As a consequence, if the market is undisturbed by outside shocks, it will converge to P^* from any initial distribution P_0. In particular, when the model is constrained as in sec. 3.2, P^* has the peculiar property to make it most likely for market participants to allocate their expected supply and demand through their expected number of buy and sell relationship. In this sense, the conditional geometric model of sec. 3.2 realizes efficiently agents' expectations, in whichever way the latter are obtained. We can say that this model is *neutral* with respect to expectations.

On the other hand, the results of sec. 4 show that entropy maximization by itself is not sufficient to turn a model into a good predictor of real networks. For instance, it is well known that some ME models, like binomial networks, return unrealistic dense networks [12]. Instead, conditional geometric networks overcome this major weakness by construction, thus providing a significant improvement with respect to known ME techniques. In particular, in sec. 4 I have compared conditional geometric networks with other models obtained through entropy maximization with reference to the same real network. This comparison suggests that the conditional geometric model is able to provide more reliable ME estimates of unknown markets or networks whenever both the strength and degree distributions of agents are known. [11][17]

References

[1] Franklin Allen and Douglas Gale. Financial contagion. *Journal of Political Economy*, 108(1):1–33, 2000.

[2] Leonardo Bargigli and Mauro Gallegati. Random digraphs with given expected degree sequences: A model for economic networks. *Journal of Economic Behavior & Organization*, 78(3):396 – 411, 2011.

[3] Fan Chung and Linyuan Lu. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics*, 6:125–145, 2002.

[4] Fan Chung and Linyuan Lu. Concentration inequalities and martingale inequalities ? a survey. *Internet Mathematics*, 3(1):79–127, 2006.

[5] G. De Masi, Y. Fujiwara, M. Gallegati, B. Greenwald, and J. E. Stiglitz. An analysis of the japanese credit network. 2009.

[6] G. Fagiolo, T. Squartini, and D. Garlaschelli. Null Models of Economic Networks: The Case of the World Trade Web. *ArXiv e-prints*, December 2011.

[7] Santo Fortunato. Community detection in graphs. *Physics Reports*, 486(3-5):75–174, 2010.

[8] Diego Garlaschelli and Maria I. Loffredo. Maximum likelihood: Extracting unbiased information from complex networks. *Phys. Rev. E*, 78(1):015101, Jul 2008.

[9] Diego Garlaschelli and Maria I. Loffredo. Generalized bose-fermi statistics and structural correlations in weighted networks. *Phys. Rev. Lett.*, 102(3):038701, Jan 2009.

[10] Giulia Iori, Giulia De Masi, Ovidiu Vasile Precup, Giampaolo Gabbi, and Guido Caldarelli. A network analysis of the italian overnight money market. *Journal of Economic Dynamics and Control*, 32(1):259 – 278, 2008. ¡ce:title¿Applications of statistical physics in economics and finance¡/ce:title¿.

[11] R. Mastrandrea, T. Squartini, G. Fagiolo, and D. Garlaschelli. Enhanced network reconstruction from irreducible local information. *ArXiv e-prints*, July 2013.

[12] I. Mastromatteo, E. Zarinelli, and M. Marsili. Reconstruction of financial networks for robust estimation of systemic risk. *Journal of Statistical Mechanics: Theory and Experiment*, 3:11, March 2012.

[13] M. E. J. Newman. Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, 74:036104, 2006.

[14] Juyong Park and M. E. J. Newman. Statistical mechanics of networks. *Phys. Rev. E*, 70:066117, Dec 2004.

[15] Tiziano Squartini and Diego Garlaschelli. Analytical maximum-likelihood method to detect patterns in real networks. *New Journal of Physics*, 13(8):083001, 2011.

[16] M. Tumminello, S. Miccichè, F. Lillo, J. Piilo, R. N. Mantegna, and E. Ben-Jacob. Statistically Validated Networks in Bipartite Complex Systems. *PLoS ONE*, 6, March 2011.

[17] Stefano Viaggiu, Andrea Lionetto, Leonardo Bargigli, and Michele Longo. Statistical ensembles for money and debt. *Physica A: Statistical Mechanics and its Applications*, 391(20):4839 – 4849, 2012.

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

6 Appendix

6.1 Derivation of equation (22)

In order to obtain $\langle w_{ij}\rangle$ for the GM2 model (sec. 3.2) we can use the following general relationship:

$$-\frac{\partial^2 F}{\partial \Lambda_{ij}^2} = \sigma^2(w_{ij}) \tag{23}$$

In fact, remembering that $Z = \sum_0^\infty e^{-\Lambda_{ij} w}$, we have that

$$\begin{aligned}
-\frac{\partial^2 F}{\partial \Lambda_{ij}^2} &= -\frac{\partial}{\partial \Lambda_{ij}}\left(\frac{1}{Z}\sum_0^\infty w e^{-\Lambda_{ij} w}\right) = \\
&= \frac{1}{Z}\sum_0^\infty w^2 e^{-\Lambda_{ij} w} - \frac{1}{Z^2}\left(\sum_0^\infty w e^{-\Lambda_{ij} w}\right)^2 = \\
&= \langle w_{ij}^2\rangle - \langle w_{ij}\rangle^2
\end{aligned}$$

which implies

$$\langle w_{ij}^2 \rangle = -\frac{1}{Z}\frac{\partial}{\partial \Lambda_{ij}}\left(\sum_0^\infty w e^{-\Lambda_{ij} w}\right) \tag{24}$$

In our case, we obtain

$$\begin{aligned}
\langle w_{ij}^2 \rangle &= -\frac{1-e^{-\Lambda_{ij}}}{1+(e^{-\Theta_{ij}}-1)e^{-\Lambda_{ij}}} \;\frac{\partial}{\partial \Lambda_{ij}}\left[\frac{e^{-(\Theta_{ij}+\Lambda_{ij})}}{(1-e^{\Lambda_{ij}})^2}\right] \\
&= \frac{(1+e^{-\Lambda_{ij}})}{1+(e^{-\Theta_{ij}}-1)e^{-\Lambda_{ij}}} \;\frac{e^{-(\Lambda_{ij}+\Theta_{ij})}}{(1-e_{-\Lambda})^2}
\end{aligned}$$

6.2 Derivation of the CDF and quantile function

For the GM2 model (sec. 3.2) the following CDF is obtained:

$$\begin{aligned}
F(x) &= \sum_{w=0}^{x} \frac{1-e^{-\Lambda_{ij}}}{1+(e^{-\Theta_{ij}}-1)e^{-\Lambda_{ij}}} e^{-[\Lambda_{ij} w + \Theta_{ij} H(w)]} = \\
&= \frac{1-e^{-\Lambda_{ij}}}{1+(e^{-\Theta_{ij}}-1)e^{-\Lambda_{ij}}}\left(1+e^{-\Theta_{ij}}\sum_{w=1}^{x} e^{-\Lambda_{ij} w}\right) = \\
&= \frac{1-e^{-\Lambda_{ij}}}{1+(e^{-\Theta_{ij}}-1)e^{-\Lambda_{ij}}}\left[1+e^{-\Theta_{ij}}\left(\frac{1-e^{-\Lambda_{ij}(x+1)}}{1-e^{-\Lambda_{ij}}}-1\right)\right] = \\
&= \frac{1-e^{-\Lambda_{ij}}+e^{-[\Lambda_{ij}+\Theta_{ij}]}\left(1-e^{-\Lambda_{ij} x}\right)}{1+(e^{-\Theta_{ij}}-1)e^{-\Lambda_{ij}}} = \\
&= F(0) + \langle a_{ij} \rangle \left(1-e^{-\Lambda_{ij} x}\right)
\end{aligned}$$

The quantile function is obtained simply by inversion

$$x = \begin{cases} 0 & \text{if } F \leqslant F(0) \\ \dfrac{\log\langle a_{ij}\rangle - \log\left(\langle a_{ij}\rangle + F(0) - F\right)}{\Lambda_{ij}} & \text{if } F > F(0) \end{cases} \tag{25}$$

Using eq. (25) it's possible to create random variables distributed according to the pmf defined by eq. (16). All of these functions have straightforward adaptations for the model of Sec. 3.1 . In the case of the model of section 2, instead, the CDF is geometric:

$$F(x) = 1 - e^{-\Lambda_{ij}(x+1)} \tag{26}$$

and the quantile function reads

$$x = \max\left(\frac{\log(F-1)}{\Lambda_{ij}} - 1, 0\right) \tag{27}$$