



# Compound Poisson models for weighted networks with applications in finance

Axel Gandy<sup>1</sup> · Luitgard A. M. Veraart<sup>2</sup>

Received: 7 June 2019 / Accepted: 4 May 2020 / Published online: 29 May 2020  
© The Author(s) 2020

## Abstract

We develop a modelling framework for estimating and predicting weighted network data. The edge weights in weighted networks often arise from aggregating some individual relationships between the nodes. Motivated by this, we introduce a modelling framework for weighted networks based on the compound Poisson distribution. To allow for heterogeneity between the nodes, we use a regression approach for the model parameters. We test the new modelling framework on two types of financial networks: a network of financial institutions in which the edge weights represent exposures from trading Credit Default Swaps and a network of countries in which the edge weights represent cross-border lending. The compound Poisson Gamma distributions with regression fit the data well in both situations. We illustrate how this modelling framework can be used for predicting unobserved edges and their weights in an only partially observed network. This is for example relevant for assessing systemic risk in financial networks.

**Keywords** Weighted directed networks · Compound Poisson distribution · Regression · Subnetwork prediction · Financial networks · Systemic risk

**JEL Classification** C02 · C46 · C53 · D85 · G32

## 1 Introduction

We provide a modelling framework that can be used to estimate and predict weighted network data. The edge weights in weighted networks often arise from aggregating some individual relationships between the nodes. For example, they can represent trades between financial

---

✉ Luitgard A. M. Veraart  
l.veraart@lse.ac.uk

Axel Gandy  
a.gandy@imperial.ac.uk

<sup>1</sup> Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

<sup>2</sup> Department of Mathematics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK

institutions in trading networks, see e.g. Gandy and Veraart [20] for a network of financial exposures arising from trading financial derivatives, or they can represent the supply of goods or services between different sectors in the economy modelled as an input–output network, see e.g. Acemoglu et al. [1]. Other applications arise for example in transport networks where the weights can represent the number of passengers travelling, see e.g. Barrat et al. [4], or in networks representing co-authorship in scientific publications, see also Barrat et al. [4], where the weights are a measure that accounts for the number of joint papers written in co-author networks. Motivated by this, we introduce a modelling framework for weighted directed networks based on the compound Poisson distribution.

We are interested in these weights and not just the topology of the underlying network, because in many applications the weights are fundamental for the behaviour of processes that can be observed on these networks. For example, in the 2007–2008 financial crisis, the interconnections between the financial institutions served as transmission channels for stress and losses that led to significant feedback and amplification mechanisms with severe consequences for the real economy. The magnitude of these losses is fundamentally linked to the weights of the edges in the network. This is clear from many studies on systemic risk in financial networks such as models looking at solvency contagion [16,42], contagion caused by marking-to-market effects [44], fire sales [7,9,11,12,22] or liquidity contagion [29]; see also Glasserman and Young [21], Capponi [6] for surveys.

The compound Poisson model class which we propose includes the size of the weights as an integral part of the model. Many financial networks are essentially the aggregation of several individual trades. This is why compound Poisson based models seem a natural choice. The networks would be resulting from a random number of individual items, that are themselves random.

Another feature of weighted networks is that they are heterogeneous. Financial networks are a prime example. Some nodes are strongly connected with a large number of trading partners, whereas others only trade with a small number of counterparties. In transport networks, we see similar effects. E.g., if the nodes are the cities and the weights are available seats on non-stop flights between two cities per day as in Barrat et al. [4], these networks are strongly heterogeneous.

We take account of this heterogeneity by allowing the nodes in the network to have individual characteristics, which we call *fitness*, with the interpretation that a larger fitness leads to a larger number of edges.

We model these fitness parameters using a regression framework (Sect. 2). In particular, we model some characteristics of the compound Poisson Gamma distribution (such as its mean, which represents the mean weight between two nodes in the network) as a suitable function of a fitness parameter that is associated with every node. By doing that, both the existence of an edge and also its weight is influenced by the fitness parameters associated with the nodes in between which the edge is formed. This enables us to reproduce several stylised facts of financial networks.

We apply the new model class to two different types of financial network data (Sect. 3): First, we consider networks that describe exposures based on a special type of financial derivative (Credit Default Swaps). Second, we consider networks that describe international lending relationships between financial institutions. We fit some models of our new model class to the empirical financial network data and find in general that they fit the data well. In particular, we find that in most cases the compound Poisson models that model both the expectation of the Poisson random variable and the expectation of the Gamma distribution via separate regression models perform best.

As an application, we show how the modelling framework can be used to predict unobserved parts of a larger network. For that, we take the empirical networks as given and assume that a subset of the edges is no longer observable. We fit several models from our framework to the observable part of the network and use the results to predict the unobserved edges. For the Credit Default Swap data we find that a model which only uses one regression for the mean of the Poisson distribution performs best. The Credit Default Swap data exhibit a rather traditional monotonic relationship between strengths and degrees in the network. For the international lending network the relationship between strengths and degrees is no-longer monotonic. In this case we find a clear advantages of using a model with both a regression for the mean of the Poisson distribution and a separate regression for the mean of the Gamma distribution. This type of analysis, namely predicting unobserved parts of a network, could be incorporated into a macro-prudential stress test for assessing systemic risk in partially observed financial networks.

## 1.1 Related literature

Network models have been developed for a wide range of applications, for example in biology, information science and economics. The seminal model by Erdős and Rényi [17] (henceforth ER) considers a network of  $n$  nodes and assumes that every pair of nodes is connected with probability  $p \in [0, 1]$ . To account for properties of empirical networks, a wide selection of models has been suggested, see Albert and Barabási [2] and Newman [33,34] for overviews.

The existing literature that analyses financial network data mainly focuses on the corresponding adjacency matrix or on the degree distribution. Financial network data have been studied for various countries, e.g. Austria [5], Brazil [10], Germany [43], Italy [25], Mexico [31], the Netherlands [24] and the UK [45]. Papers that do consider the weights of the network usually focus on the tail of the weights and find heavy tails, see e.g., Boss et al. [5] and Cont et al. [10]. The focus on adjacency matrices and degree distributions is also evident in the literature on core-periphery financial networks [13,18,24], as well as in the literature on reconstructing financial networks from partial information; see Gandy and Veraart [19].

A huge variety of fitness models for financial networks has been considered in the literature, see e.g. Jacobs and Clauset [26] and Gandy and Veraart [19]. Fitnesses are also sometimes referred to as *sociability parameter* [8], or *capacities* [35]. The statistics literature considers these fitness models in the context of *graphons* which are functions in two variables (fitnesses) determining the link existence probabilities between any two nodes [30,37,47].

The majority of fitness models use the fitnesses only to model the existence of the edges in a network but not its weights. To the best of our knowledge, Gandy and Veraart [19] is the only model that uses a fitness approach to model both the existence and the weight of an edge in a (financial) network. This is also what we suggest in this paper. In contrast to the model considered in Gandy and Veraart [19] we can allow for a wider class of models for the weights of the distribution of the edges. This is because in the present paper we fit a network model to observed network data and do not try to reconstruct a network from observed aggregates of the network. The statistical inference for the former problem seems to be more easily tractable than for the latter which allows us to consider a wider class of probability distributions for the financial network.

Compound Poisson models for networks have been considered before but in a slightly different context. For non-weighted networks, for example, Ranola et al. [41] and Norros and Reittu [35] propose models in which number of edges is modelled using a Poisson distribution that depends on fitness parameters. In contrast to these approaches, we consider

weighted networks and the fitness parameters do not just influence the existence of edges but also their weights.

Exponential random graph models [23,38] are another popular approach for statistical inference of networks. While these models do not consider weighted edges, there are some proposals for extensions to weighted random graphs, see for example the *generalized exponential random graph model (GERGM)* by Wilson et al. [46] who specify a joint distribution for an exponential family of graphs with edge weights. They provide a Metropolis-Hasting method to estimate the model and apply it to several real-world networks one of which is also an international lending network of the type that we consider in our empirical study as well.

In our models, we develop stochastic (probabilistic) models for random weighted graphs (the financial networks)—so the random object is the graph itself. This is different from the field of probabilistic graphical models [28] and from the field of high-dimensional random graph estimation [32], where graphs are used to help describe dependencies between components of a multivariate random variable. There the graph is not an (observable) random object—it is a property of the random object.

## 2 Compound Poisson models

### 2.1 Definitions

In the following, we introduce a new model class for *weighted* and *directed* graphs consisting of a fixed number  $n \in \mathbb{N}$  of nodes. Furthermore, we assume that the edges are modelled as random variables. A *network* consisting of  $n \in \mathbb{N}$  nodes is given by a matrix  $L = (L_{ij})_{i,j \in \{1,2,\dots,n\}}$ , where the  $L_{ij}$  are random variables modelling the weight of the directed edge from node  $i$  to node  $j$ . A weight of 0 indicates that the corresponding edge is not present. This definition of a network allows for at most one weighted directed edge between two nodes. In practice, these weights are often aggregates of several individual relationships between the nodes, which motivates our model choice.

We propose using a compound Poisson Gamma distribution for these weights, with parameters given by a regression model. A compound Poisson Gamma distribution can be defined via the random variable

$$X = \sum_{v=1}^N S_v,$$

where  $N \sim \text{Poisson}(\lambda)$  and  $S_v \sim \text{Gamma}(\alpha, \mu_S)$ ,  $v = 1, \dots, N$ , are independent, where  $\text{Poisson}(\lambda)$  is the Poisson distribution with mean  $\lambda$  and  $\text{Gamma}(\alpha, \mu_S)$  is the Gamma distribution with shape parameter  $\alpha$  and mean  $\mu_S$ .<sup>1</sup> Then  $\text{Var}(S_v) = \mu_S^2/\alpha$  and  $\mathbb{E}[S_v^2] = \mu_S^2 \frac{1+\alpha}{\alpha}$ .

It is well known that  $\mathbb{E}[X] = \mathbb{E}[N]\mathbb{E}[S_v] = \lambda\mu_S$  and  $\text{Var}[X] = \mathbb{E}[N]E[S_v^2] = \lambda\mu_S^2 \frac{1+\alpha}{\alpha}$ . This can be seen as a special case of the so-called Tweedie distribution Jorgensen [27] and Dunn and Smyth [15], which is usually parametrised via its mean  $\mu$  and parameters  $\phi, p$  such that  $\mathbb{E}[X] = \mu$  and  $\text{Var}[X] = \phi\mu^p$ .<sup>2</sup>

<sup>1</sup> In particular, the probability density function of  $S_v$  is given by  $f(s) = \Gamma(\alpha)^{-1} \left(\frac{\alpha}{\mu_S}\right)^\alpha x^{\alpha-1} \exp(-x \frac{\alpha}{\mu_S})$  for  $x \in (0, \infty)$ .

<sup>2</sup> We can set  $\mu = \lambda\mu_S$ ,  $p = 2$ ,  $\phi = \frac{1+\alpha}{\alpha\lambda}$ . Then, indeed  $\mathbb{E}[X] = \mu$  and  $\text{Var}(X) = \lambda\mu_S^2 \frac{1+\alpha}{\alpha} = \phi\mu^p$ .

Our network  $L = (L_{ij})_{1 \leq i, j \leq n}$  will be modelled as independent random variables having a compound Poisson Gamma distributions, with parameters defined via a regression.<sup>3</sup> We will propose two ways of doing this—the first (CPNet1) will model  $\mu_{ij} := \mathbb{E}[L_{ij}]$  via regression and the second (CPNet2) will model both the mean of  $N$ , i.e.  $\lambda$  and the mean of  $S_v$ , i.e.  $\mu_S$ , via regression. The numbers 1 and 2 in the names of CPNet1 and CPNet2 indicate how many regressions are embedded in the model.

The parameters of CPNet1 are chosen as follows. The shape parameter of the Gamma distribution is a fixed constant  $\alpha$ . As mentioned before, we would like to define the overall mean via regression—thus we want to achieve  $\mathbb{E}[L_{ij}] = \mu_{ij}$  for given  $\mu_{ij}$ . That leaves flexibility on how to define the means of the Poisson and Gamma part of the distribution. We resolve this by imposing a second moment condition, namely  $\mathbb{V}\text{ar}[L_{ij}] = \phi \mu_{ij}^p$ , where  $p = \frac{\alpha+2}{\alpha+1}$ . This ensures that every element of  $L$  will follow a Tweedie distribution with parameters  $\mu_{ij}, \phi, p$ , with  $p \in (1, 2)$ .

**Definition 2.1** (CPNet1) Let  $p \in \mathbb{N}, X \in \mathbb{R}^{n \times p}$ , let  $\theta = (\beta_1, \dots, \beta_p, \alpha, \phi) \in \mathbb{R}^p \times (0, \infty)^2$  and let  $l : \mathbb{R}^2 \rightarrow (0, \infty)$ . Then we say that the matrix  $L$  has a Compound Poisson Gamma Network regression model for the mean (CPNet1) if for all  $i, j \in \{1, \dots, n\}$ ,

$$L_{ij} = \sum_{v=1}^{N_{ij}} S_{ij}^v,$$

where  $N_{ij} \sim \text{Poisson}(\lambda_{ij}), S_{ij}^v \sim \text{Gamma}(\alpha, \mu_{ij}^S), \lambda_{ij} = \frac{1}{\phi\alpha} \mu_{ij}^{\frac{\alpha}{\alpha+1}}, \mu_{ij}^S = \phi\alpha \mu_{ij}^{\frac{1}{\alpha+1}}$ , with  $\mu_{ij} = l(f_i, f_j)$  and

$$f_i = \sum_{v=1}^p X_{iv} \beta_v.$$

In the above,  $X_{ij}, i \in \{1, \dots, n\}, j \in \{1, \dots, p\}$ , are the elements of the design matrix. The variable  $f_i$  can be interpreted as “fitness” of node  $i$ .

We refer to  $l$  in the definition above as a *link function*. Examples for link functions are  $l(x, y) = \exp(x + y), l(x, y) = \max(\exp(x), \exp(y))$  and  $l(x, y) = \exp(x) + \exp(y)$ . We would usually choose link functions that are monotonically non-decreasing in each of their arguments. This then implies that higher values of the fitnesses imply higher means of the corresponding compound Poisson distributions.

**Example 2.2** (CPNet1F model) One example of a model that falls into the CPNet1 model class is the model that we refer to as CPNet1F model, which we will use in our empirical analysis later. It is defined by setting  $p = n, X = I_n \in \mathbb{R}^{n \times n}$  and  $l : \mathbb{R}^2 \rightarrow (0, \infty)$  with  $l(x, y) = \exp(x + y)$ . Hence, it has  $n + 2$  parameters given by the vector  $\theta = (\beta_1, \dots, \beta_n, \alpha, \phi) \in \mathbb{R}^n \times (0, \infty)^2$ . In this model, the fitness parameter satisfies  $f_i = \beta_i$  and the overall mean of the edge from  $i$  to  $j$  is given by  $\mathbb{E}[L_{ij}] = \mu_{ij} = l(f_i, f_j) = \exp(\beta_i + \beta_j)$ . The parameter of the Poisson distribution is then given by  $\lambda_{ij} = \frac{1}{\phi\alpha} \mu_{ij}^{\frac{\alpha}{\alpha+1}} = \frac{1}{\phi\alpha} (\exp(\beta_i + \beta_j))^{\frac{\alpha}{\alpha+1}}$ , the shape parameter of the Gamma distribution is given by  $\alpha$  and the mean of the Gamma distribution is given by  $\mu_{ij}^S = \phi\alpha (\exp(\beta_i + \beta_j))^{\frac{1}{\alpha+1}}$ . Hence, we see that both the mean of the Poisson and the mean of the Gamma distributions are controlled simultaneously by the fitness parameter  $(\beta_1, \dots, \beta_n)$  and the parameters  $\alpha$  and  $\phi$ .

<sup>3</sup> We might have additional information of the network such as for example that no self-loops exists. If that is the case we set  $L_{ii} = 0$  for all  $i \in \{1, \dots, n\}$ .

Next we consider CPNet2, which is a model in which both the mean of the Poisson distribution and the mean of the Gamma distribution are modelled separately via regression. The shape parameter  $\alpha$  of the Gamma distribution is again a fixed constant.

**Definition 2.3** (CPNet2) For  $n \in \mathbb{N}$ , for  $k \in \{N, S\}$  let  $p_k \in \mathbb{N}$ ,  $X^k \in \mathbb{R}^{n \times p_k}$  and  $l^k : \mathbb{R}^2 \rightarrow (0, \infty)$  and let  $\theta = (\beta_1^N, \dots, \beta_{p_N}^N, \beta_1^S, \dots, \beta_{p_S}^S, \alpha) \in \mathbb{R}^{p_N+p_S} \times (0, \infty)$ .

Then a network  $L$  consisting of  $n$  nodes follows a *Compound Poisson Gamma network model with links on lambda and the mean of the Gamma distribution*(CPNet2) if  $L$  is given by

$$L_{ij} = \sum_{v=1}^{N_{ij}} S_{ij}^v,$$

where for  $i, j = 1, \dots, n$ ,  $N_{ij} \sim \text{Poisson}(l^N(f_i^N, f_j^N))$  and for  $v = 1, \dots, N_{ij}$ ,  $S_{ij}^v \sim \text{Gamma}(\alpha, l^S(f_i^S, f_j^S))$ , with

$$f_i^k = \sum_{j=1}^{p_k} X_{ij}^k \beta_j^k, \quad k \in \{N, S\}; i = 1, \dots, n.$$

In the above,  $X_{ij}^k, i \in \{1, \dots, n\}, j \in \{1, \dots, p_k\}, k \in \{N, S\}$  are the elements of the design matrices. Examples for link functions are as above. The variables  $f_i^N$  and  $f_i^S$  can be interpreted as fitnesses of node  $i$ , one affecting the Poisson part of the model, the other the Gamma part of the model.

**Example 2.4** (CPNet2FPG model) We introduce the CPNet2FPG as an example of a CPNet2 model. It has fitness-based parameters on both the Poisson and the Gamma part of the model, i.e.  $p_N = n, X^N = I_n, p_S = n, X^S = I_n, l^S(x, y) = l^N(x, y) = \exp(x + y)$ . It thus has  $2n + 1$  parameters, namely  $2n$  fitness parameters and the shape parameter of the Gamma distribution  $\alpha$ . In particular, the fitness parameters for the Poisson distribution are given by  $f_i^N = \beta_i^N, i \in \{1, \dots, n\}$ . Hence, the mean of the Poisson distribution used to model the edge from  $i$  to  $j$  is given by  $\exp(\beta_i^N + \beta_j^N)$ . Furthermore, the fitness parameters for the Gamma distribution are given by  $f_i^S = \beta_i^S, i \in \{1, \dots, n\}$ . Hence, the mean of the Gamma distribution for the edge from  $i$  to  $j$  is given by  $\exp(\beta_i^S + \beta_j^S)$ . The parameters for the Poisson distribution are different from the parameters used for the Gamma distribution. This will enable us to model the existence of edges independently of the weights of the edges as we will discuss later.

## 2.2 Motivation behind the choice of compound Poisson distributions

One motivation behind the compound Poisson models (Definitions 2.1 and 2.3) is that many weighted networks consist of multiple directed edges between the nodes and these are then aggregated to obtain one network with at most one directed edge between each node. For example, consider a network of bilateral exposures on individual CDS. Each bilateral exposure consists in fact of several separate transactions, as described in Peltonen et al. [39].

Another motivation comes from the fact that many financial networks do not automatically net exposures between counterparties. Using the compound Poisson distribution independently for both possible directional exposures allows for exposures in both directions, as well as for no exposure at all between counterparties.

Furthermore, our basic framework has enough flexibility to match important features of the link existence distribution and of the exposure distribution. The compound Poisson Gamma distribution has three parameters (one for the Poisson part and two for the Gamma part), thus enabling us to match 3 properties such as the probability of no link, as well as the mean and the variance of the exposure.

### 2.3 Interpretation as fitness models

These new models were inspired by the classical fitness models (see Sect. 1.1) that assign fitnesses to every node which then determines the link existence probabilities for every edge. We, however, take a broader view by considering a general regression framework that enables us to characterise more general features of the random graph. In particular, our regression framework incorporates fitness models as special cases but with the additional feature that fitnesses are used to characterise properties of the weights of edges in addition to the existence of edges.

To see how CPNet1 can be interpreted as a classical fitness model (in which no regression is used to determine the fitness parameter), we can set  $X = I_n$ , where  $I_n$  is the  $n \times n$  identity matrix in CPNet1. Then,  $f_i = \sum_{j=1}^p X_{ij} \beta_j = \beta_i$  for all  $i \in \{1, \dots, n\}$ . Hence, the overall mean of  $L_{ij}$  is given by  $\mu_{ij} = l(f_i, f_j) = l(\beta_i, \beta_j)$  which can be interpreted as a fitness model for the mean of the weighted edges where  $\beta_i, i \in \{1, \dots, n\}$  are the fitnesses.

Similarly, we can set  $X^N = X^S = I_n$  in CPNet2. Then,  $l^N(f_i^N, f_j^N) = l^N(\beta_i^N, \beta_j^N)$  can be interpreted as a fitness model for the mean of the Poisson distribution where  $\beta_i^N, i \in \{1, \dots, n\}$  are the fitnesses and  $l^S(f_i^S, f_j^S) = l^S(\beta_i^S, \beta_j^S)$  can be interpreted as is a fitness model for the mean of the Gamma distribution with fitnesses  $\beta_i^S, i \in \{1, \dots, n\}$ .

Both CPNet1 and CPNet2 could be extended to give every node an in-fitness and an out-fitness. For example, in CPNet2, we could have 4 instead of 2 design matrices, i.e. for  $k \in \{N, S\}$  and  $l \in \{\text{in}, \text{out}\}$  we have  $X^{k,l} \in \mathbb{R}^{n \times p_{k,l}}$  and corresponding fitnesses

$$f_i^{k,l} = \sum_{j=1}^{p_{k,l}} X_{ij}^{k,l} \beta_j^{k,l}$$

that then define the model via

$$N_{ij} \sim \text{Poisson}(l^N(f_i^{N,\text{out}}, f_j^{N,\text{in}})) \quad \text{and} \quad S_{ij}^v \sim \text{Gamma}(\alpha, l^S(f_i^{S,\text{out}}, f_j^{S,\text{in}})).$$

Similarly, one could define an extension of CPNet1 with in-fitness and out-fitness.

As discussed in our literature review fitness models have been studied before and it has been shown that they can also be used to construct degree distributions with heavy tails, see e.g. Gandy and Veraart [19]. These results carry over to our class of compound Poisson models, since as one can see from the formulae for the link existence probabilities (1) one can model a wide range of link behaviour with an appropriate choice of fitness parameters and link functions  $l$ .

### 2.4 Expected degrees and strengths

Next, we derive formulae for the existence and non-existence of edges, for the expected in- and out-degrees and the expected in- and out-strengths in the new models. In particular, we show that only the parameters of the Poisson distribution determine the link existence

probabilities of the edges (together with the link function  $l$  in CPNet1 or  $l^N$  in CPNet2). The distribution used for the individual  $S_{ij}^v$  only matters for the actual weights along the edges and these weights are then also influenced by the parameters of the Poisson distribution.

**Proposition 2.5** *Let  $L$  be CPNet1 as in Definition 2.1 and let  $\tilde{L}$  be CPNet2 as in Definition 2.3. Then, for any  $i, j \in \{1, \dots, n\}$ ,*

1. *the probability for the non-existence and the existence of a directed edge from  $i$  to  $j$  is given by*

$$\begin{aligned} \mathbb{P}(L_{ij} = 0) &= \exp\left(-\frac{1}{\phi\alpha}l(f_i, f_j)^{\frac{\alpha}{\alpha+1}}\right), \quad \mathbb{P}(L_{ij} > 0) = 1 - \mathbb{P}(L_{ij} = 0), \\ \mathbb{P}(\tilde{L}_{ij} = 0) &= \exp\left(-l^N(f_i^N, f_j^N)\right), \quad \mathbb{P}(\tilde{L}_{ij} > 0) = 1 - \mathbb{P}(\tilde{L}_{ij} = 0); \end{aligned} \tag{1}$$

2. *the expected in- and out-degrees are given by*

$$\begin{aligned} \mathbb{E}[d^{in}(L)_i] &= \mathbb{E}\left[\sum_{j=1}^n \mathbb{I}_{\{L_{ji}>0\}}\right] = n - \sum_{j=1}^n \exp\left(-\frac{1}{\phi\alpha}l(f_j, f_i)^{\frac{\alpha}{\alpha+1}}\right), \\ \mathbb{E}[d^{out}(L)_i] &= \mathbb{E}\left[\sum_{j=1}^n \mathbb{I}_{\{L_{ij}>0\}}\right] = n - \sum_{j=1}^n \exp\left(-\frac{1}{\phi\alpha}l(f_i, f_j)^{\frac{\alpha}{\alpha+1}}\right), \\ \mathbb{E}[d^{in}(\tilde{L})_i] &= n - \sum_{j=1}^n \exp\left(-l^N(f_j^N, f_i^N)\right), \\ \mathbb{E}[d^{out}(\tilde{L})_i] &= n - \sum_{j=1}^n \exp\left(-l^N(f_i^N, f_j^N)\right); \end{aligned} \tag{2}$$

3. *the expected in- and out-strengths are given by*

$$\begin{aligned} \mathbb{E}[s^{in}(L)_i] &= \mathbb{E}\left[\sum_{j=1}^n L_{ji}\right] = \sum_{j=1}^n l(f_j, f_i), \\ \mathbb{E}[s^{out}(L)_i] &= \mathbb{E}\left[\sum_{j=1}^n L_{ij}\right] = \sum_{j=1}^n l(f_i, f_j), \\ \mathbb{E}[s^{in}(\tilde{L})_i] &= \mathbb{E}\left[\sum_{j=1}^n \tilde{L}_{ji}\right] = \sum_{j=1}^n l^N(f_j^N, f_i^N)l^S(f_j^S, f_i^S), \\ \mathbb{E}[s^{out}(\tilde{L})_i] &= \mathbb{E}\left[\sum_{j=1}^n \tilde{L}_{ij}\right] = \sum_{j=1}^n l^N(f_i^N, f_j^N)l^S(f_i^S, f_j^S). \end{aligned} \tag{3}$$

The results follow directly from the definition of the new models and properties of compound Poisson distributions and therefore we omit the proof.

When comparing the expected strengths to the degrees, i.e., formula (3) to (2) we see the main difference between CPNet1 and CPNet2. In CPNet1 the same model parameters determine the magnitude of the degrees and the strengths. In CPNet2 there are additional model parameters  $f_i^S$ , and link functions  $l_i^S, i \in \{1, \dots, n\}$  that influence the strengths of



the nodes but not the degrees. Hence, if there is no clear monotonic relationship between strengths and degrees this can be captured with the model class CPNet2. We will discuss this in more detail in our empirical case study.

## 2.5 Special cases: Erdős–Rényi and core-periphery model

Both CPNet1 and CPNet2 reduce to the classical Erdős–Rényi random graph model for the existence of edges for special choices of the model parameters. Indeed, in CPNet1, if we set all parameters  $f_i$  to the same value, say  $x$ , then from (1) we see that all link existence probabilities are identical and hence the CPNet1 model reduces to the classical ER model for the existence of the edges. The same holds for CPNet2 if all parameters  $f_i^N$  are set to the same value.

We can also reproduce a core-periphery structure with our new model classes. One could, for example, choose two fitnesses  $x^{\text{core}} \geq x^{\text{periphery}}$  and assign all nodes  $i$  in the core the fitness  $f_i = x^{\text{core}}$  and all nodes  $i$  in the periphery the fitness  $f_i = x^{\text{periphery}}$ . This can be achieved by setting  $\beta_1 = x^{\text{core}}$ ,  $\beta_2 = x^{\text{periphery}}$ ,  $p_k = 2$ ,  $X_{i1}^k = 1$  if  $i$  is in the core and 0 otherwise and  $X_{i2}^k = 0$  if  $i$  is in the core and 1 otherwise. Then for any function  $l^k$  that is non-decreasing in its first two arguments, one would obtain the highest probability for existing edges between two members of the core and the lowest between two members of the periphery. From (2) it is also clear that nodes in the core would have higher expected in- and out-degrees compared to nodes in the periphery. This approach could be generalised by considering possibly more than two types of vertices as in the stochastic block models for random graphs.

## 2.6 Possible applications of the models

Our modelling framework can be used to deal with missing information in network models. For example, situations in which a financial network is only partially observed and one would like to fill in the remaining parts. In contrast to the literature on network reconstruction, see e.g. Gandy and Veraart [19,20] we do not assume that the row and column sums of the network matrix  $L$  are observed and the individual entries need to be estimated, but we have a situation in mind in which the row and column sums are not observable but some individual entries of the matrix are observable. In such a situation one could fit our new model class to the available data and predict the missing entries from the fitted model. We will demonstrate how this can be done in our empirical case study.

An alternative application would be that one observes a network in the past (on one or several occasions) and fits the new model class to these observations. One then uses these results to predict a network in the future.

Alternatively, one might be in a situation that one observes a network that is related to a network of interest, e.g., a derivative exposure network corresponding to Credit Default Swap exposures written on a given reference entity is observed (for example where the reference entity is a UK company) but one is interested in the same type of network written on a different reference entity (for example a non-UK company) and would like to make predictions about this network.

All these possible application areas could arise in the context of macro-prudential stress testing for systemic risk analysis in financial networks. To be able to conduct a macro-prudential stress test one needs to consider the financial system as a whole and analyse potential feedback and amplification mechanisms between the market participants. Often,

the connections that give rise to such feedback mechanisms are not fully observable and therefore one will need to rely on statistical and simulation methods to deal with the missing information. This is where our compound Poisson model class can be used. In Gandy and Veraart [19,20] it was demonstrated how a network reconstruction method can be used in a macroprudential stress test if the network of interest is not fully observable. As mentioned before, in these papers the assumption was that the network matrix itself was not observable but its row and column sums were. Here we assume that a subset of the network is observable, and we use the subset to estimate a statistical model that will then be used to predict the missing edges in the original network that is not fully observable.

### 3 Empirical case studies

We will now fit the new class of compound Poisson models to two different data sets of financial networks. The first data set contains financial networks representing exposures due to financial derivatives and the second data set contains financial networks representing cross-border lending activities. In addition to the compound Poisson models with regression introduced in this paper, we will compare the fit to some alternative models for financial networks. We compare the performance of the models in-sample in Sect. 3.4 (using the Akaike information criterion (AIC)) and out-of-sample in Sect. 3.5 (using cross-validation).

#### 3.1 Data description: derivative exposure network

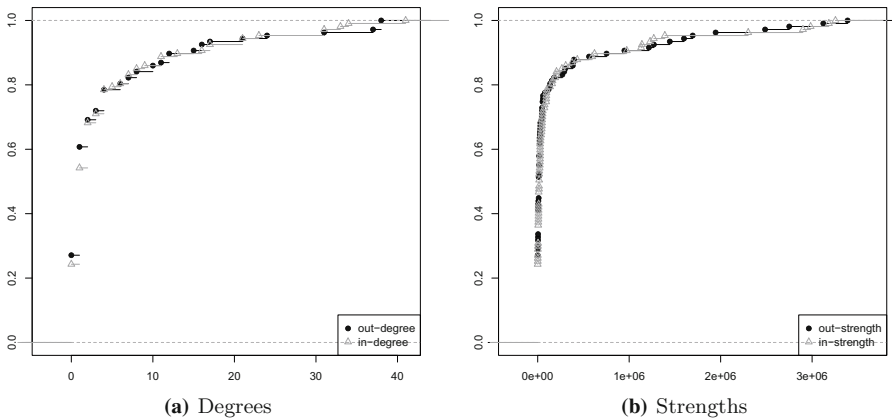
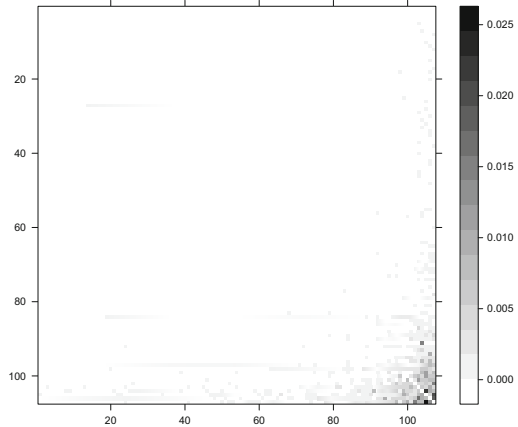
First, we consider a data set that contains a snapshot of roughly 134,000 outstanding positions in Credit Default Swaps (CDS) referencing 89 different UK institutions, taken in the second half of 2011. We will refer to them as *CDS data*. The data come from the Depository Trust and Clearing Corporation's (DTCC) Trade Information Warehouse (TIW) and were supplied to us by the Bank of England with anonymized counterparties. These data were also considered in Gandy and Veraart [20]. As described there, these data record for each reference entity, both counterparties of a position (buyer and seller) and the notional amount. We only consider positions for which the notional amounts are quoted in EUR. The notional amount "represents the par amount of credit protection bought or sold, equivalent to debt or bond amounts, and is used to derive the coupon payment calculations for each payment period and the recovery amounts in the event of a default" [14, p. 3]. From these data, we construct for each UK reference entity being referenced a network between buyers and sellers describing the total outstanding positions in credit default swaps referencing this particular institution. This leads to 89 networks in total. Sometimes, for a given reference entity, a pair of buyer and seller is listed more than once which corresponds to outstanding positions for different maturities. For these cases, we just add up all the multiple entries to obtain the total weight for such an edge. In the following we consider (an arbitrary selection of) 5 of these networks—we refer to them as CDS\_A, where  $A \in \{1, \dots, 5\}$ . Table 1 provides some summary statistics for these five networks.

Figure 1 contains a plot of one of these networks consisting of 107 nodes. The network matrix has been normalised such that the sum of all entries of the matrix equals 1. We see that there is a strong clustering of exposures in the lower right corner representing mainly exposures between dealers in this network. This network represents a very typical financial network exhibiting some core-periphery structure.

**Table 1** Network characteristics for the five CDS networks

Network name	CDS_1	CDS_2	CDS_3	CDS_4	CDS_5
Number of nodes	32	20	33	62	107
Density	0.21	0.43	0.15	0.09	0.04

**Fig. 1** Normalized exposure network for CDS\_5 (row: protection seller, column: protection buyer)

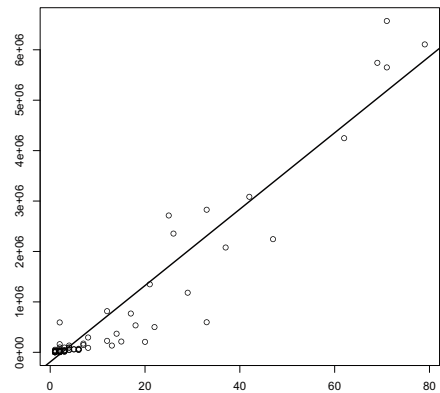


**Fig. 2** Empirical cumulative distribution function of the in- and out-degrees (left) and the in- and out-strengths for the CDS\_5 network

In the following we consider some more descriptive statistics to understand some properties of the network. Figure 2 shows the empirical cumulative distribution functions of the in- and out-degrees (left) and the in- and out-strengths (right). In general, we find that this network appears to be symmetric with almost no difference between the in- or out-degrees and the in- or out-strengths.

To understand the relationship between strengths and degrees, we consider Fig. 3 which shows a scatter plot of the strengths against the degrees. There is a clear tendency for nodes with high degrees to also have high strengths. We fit a simple linear model to the observed total strengths (in- + out-strengths) using an intercept  $\beta_0$  and slope parameter  $\beta_1$  where we use the

**Fig. 3** Relationship between strengths and degrees and fitted regression line for the CDS\_5 network



total degree as explanatory variables. In particular, we set  $\text{strength}_i = \beta_0 + \beta_1 \text{degree}_i + \epsilon_i$ ,  $i \in \{1, \dots, n\}$  where  $\epsilon_i$  is the error term. The regression line is also included in Fig. 3 and we see that the linear relationship seems to describe the data reasonably well. Hence, for this data set, a model that associates higher weights with more links seems to be appropriate. We will see that this can be achieved by our compound Poisson models.

Similar monotonic relationships between strengths and degrees have been found in other networks. For example, Barrat et al. [4] finds in an analysis of a world-wide airport network (in which nodes represent airports, and the weighted edges represent the number of available seats on direct flights between these airports) that the average strength of a node with degree  $k$  increases with the degree proportional to  $k^b$  for some parameter  $b$ .

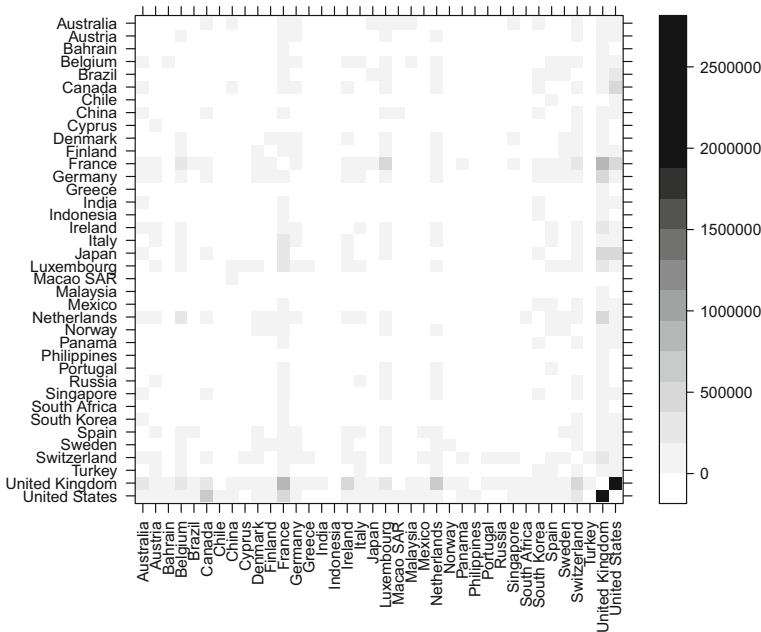
### 3.2 Data description: international lending network

As the second example, we consider data from the Bank for International Settlements that they collect as part of their *locational banking statistics (LBS)*. We will refer to them as *LBS data*. These data are publicly available.<sup>4</sup> These data contain information on claims and liabilities of financial institutions aggregated on a country level. From these data, we chose the 38 countries that report their financial activities to the BIS, see Table 2. These reporting countries have financial interactions with 521 other countries or groups of countries, but we only consider the trading activities between the 38 reporting countries that we have chosen. Hence, all the networks we consider in this case study contain 38 nodes. The data set contains additional information such as the sector of the counterparty (e.g. bank or non-bank). To keep the analysis tractable we have chosen the highest level of aggregation, i.e., we do not differentiate based on additional information available in the data set. The data are reported quarterly (starting from 1977 for some countries) and we have chosen the first quarter (Q1) of the following three years: 2000, 2009, 2018. For these three time points, we construct three networks as follows. First, we consider the network of claims reported by the reporting countries, i.e., the individual entries of the network  $L_{ij}^{(c)}$  represent the cross-border outstanding claims from country  $i$  to country  $j$  in million USD ( $i, j \in \{1, \dots, 38\}$ ) that were reported by the reporting countries. If pairs of countries occur multiple times, we add the corresponding positions. Second, we consider the network of liabilities reported by

<sup>4</sup> We used the data set available under the Link *Locational banking statistics* on the website [https://www.bis.org/statistics/full\\_data\\_sets.htm](https://www.bis.org/statistics/full_data_sets.htm).

**Table 2** Selected countries of the BIS reporting countries

Australia	Austria	Bahrain	Belgium	Brazil
Canada	Chile	China	Cyprus	Denmark
Finland	France	Germany	Greece	India
Indonesia	Ireland	Italy	Japan	Luxembourg
Macao SAR	Malaysia	Mexico	Netherlands	Norway
Panama	Philippines	Portugal	Russia	Singapore
South Africa	South Korea	Spain	Sweden	Switzerland
Turkey	United Kingdom	United States		

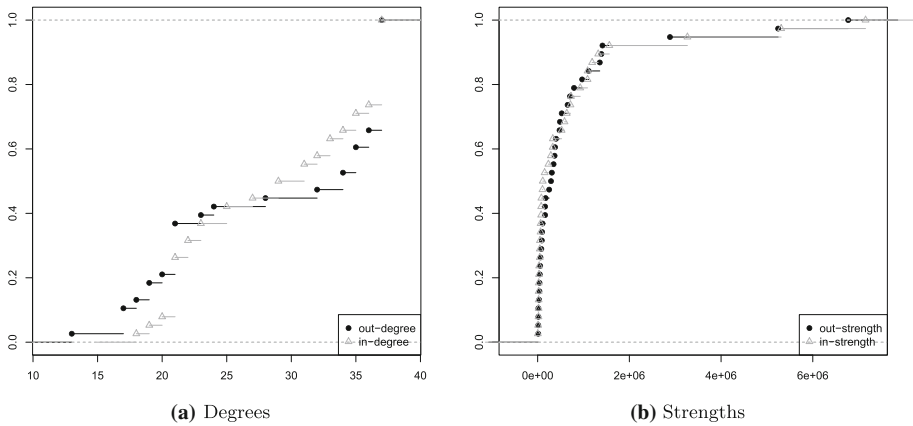


**Fig. 4** Total cross-border liabilities between the LBS reporting countries in million USD

the reporting countries. Hence,  $L_{ij}^{(l)}$  represent the cross-border outstanding liabilities from country  $i$  to country  $j$  in million USD ( $i, j \in \{1, \dots, 38\}$ ) reported by the reporting countries. Again, if there are multiple entries for pairs of countries we just add up these positions. Third, we consider a combined network of claims and liabilities given by  $L_{ij}^{(c)} = L_{ij}^{(l)} + L_{ji}^{(c)}$ . We will refer to the resulting 9 matrices as LBS\_A\_B, where  $A \in \{2000, 2009, 20018\}$  denotes the year and  $B \in \{L, C, T\}$  denotes whether the network it the network of liabilities (L) (from the perspective of the reporting country), claims (C) (from the perspective of the reporting country) or the combined network (T).

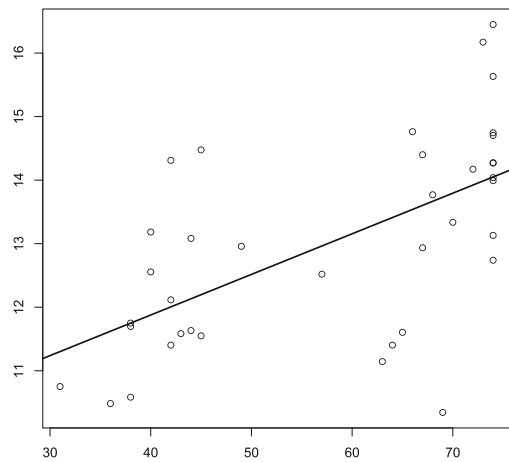
The BIS data have been used in a wide range of studies, see e.g. Oatley et al. [36] for another application.

Figure 4 contains a plot of one of these networks consisting of 38 nodes and representing the combined network of claims and liabilities reported in Q1 2018 (in million USD). We see that the United States and the United Kingdom are dominating the picture by being essentially



**Fig. 5** Empirical cumulative distribution function of the in- and out-degree (left) and the in- and out-strength for the LBS data from 2018 Q1

**Fig. 6** Relationship between log-strength and degree and fitted regression line for the LBS data from 2018 Q1



the only two countries that are connected to almost all other countries. Furthermore, the cross-border liabilities from the UK to the USA and vice-versa are by orders of magnitude larger than liabilities between any other pair of countries. When observing the data over time (not reported here) we find that the cross-border liabilities seem to concentrate and the UK and the USA have become more important over time, both when measured in terms of their degrees and their strengths. This concentration has been analysed further in Aldasoro and Ehlers [3].

In the following, we investigate the relationship between strength and degree in the LBS data. Figure 5 shows the empirical cumulative distribution functions of the in- and out-degrees (left) and the in- and out-strengths (right). Similarly to the results for the CDS data, we find that this network appears to be quite symmetric. The in-strength seem to be quite similar to the out-strengths and the same holds for the in- and out-degrees. In contrast to the CDS data, now (in-/out-) degrees seem to exhibit a different pattern compared to (in-/out-) strengths. In particular, as one can see from the empirical cumulative distribution functions the in- and out-degree distributions appear to be bimodal which is not the case for the distribution of the in- and out-strengths. To illustrate this difference further we again look at a scatter plot

of the log-strengths against the degrees in Fig. 6. We fit a regression line which still exhibits a positive slope indicating that there is still some tendency for nodes with higher strengths to be associated with nodes that have higher degrees, with considerable scatter around the regression line. We clearly see that there are some nodes which have very high strengths but rather low degrees and nodes that have rather high degrees but low strengths.

On the one hand, there are four countries for which their total strength is greater or equal than the median strength of all countries and at the same time their total degrees are less or equal than the median degrees over all countries (this holds for China, Germany, Japan and Singapore). These countries have rather low degree despite their high strength. On the other hand, there are four countries whose total strength is less or equal than the median strength but their total degree is larger or equal than the median degree (this holds for Finland, Philippines, South Africa, South Korea). These countries have high degrees despite their small strengths. Hence, according to this informal “outlier” criterion  $8/38 \approx 21\%$  nodes are outliers. The same analysis for the CDS network only reveals  $11/107 \approx 10\%$  outliers. Hence, the LBS data have different features compared to the CDS data. In particular, ordering the nodes according to their strength does not coincide with ordering the nodes according to their degree. To be able to fit such a type of behaviour we need a model class that is flexible enough to at least accommodate a partial separation of the weights of an edge from the existence of an edge. We will in the following show how this can be achieved with the compound Poisson regression models.

### 3.3 Models in the comparisons

We now list the models that we consider in our comparisons for the empirical case study. Since our model classes CPNet1 and CPNet2 are very flexible, we choose several choices of models that fall within these two model classes. In addition, we consider some modelling approaches that do not fall within the classes CPNet1 and CPNet2 but appear to be a natural alternative modelling approach to the compound Poisson approach.

The first three models are models for homogeneous networks. The other models allow for differences between the nodes by introducing fitness parameters.

1. *ERE* uses an Erdős–Rényi network for the link existence probability and then has weights following an exponential distribution. Formally, its model parameter is  $\theta = (p, \lambda) \in [0, 1] \times (0, \infty)$  and independently,  $\mathbb{P}(L_{ij} > 0) = p$  and  $L_{ij}|L_{ij} > 0 \sim \text{Exponential}(\lambda)$ . This model has been used in Gandy and Veraart [19] as an a priori model in a Bayesian framework for network reconstruction. This model is not part of the CPNet1 or CPNet2 class.
2. *ERG* extends the previous model by allowing a Gamma distribution for the distribution of the weights. The parameter is now  $\theta = (p, \alpha, \beta) \in [0, 1] \times (0, \infty)^2$ ,  $\mathbb{P}(L_{ij} > 0) = p$  and  $L_{ij}|L_{ij} > 0 \sim \text{Gamma}(\alpha, \beta)$ . This model is also not part of the CPNet1 or CPNet2 class.
3. *Tweedie* is a special case of CPNet1 with  $p = 1$ ,  $X = (1, \dots, 1)^T$  and  $l(x, y) = \exp(x + y)$ . This implies that all entries follow the same Tweedie distribution.
4. *CPNet1F* is a fitness-based model from the CPNet1 family of models. It uses the fitness in the regression of the overall mean. To be precise, it sets  $p^N = n$ ,  $X = I_n$  and uses the link function  $l(x, y) = \exp(x + y)$ . It has  $n + 2$  parameters.
5. *CPNet2FP* is a CPNet2 model with fitness-based parameters on the Poisson part of the model only, i.e.  $p_N = n$ ,  $X^N = I_n$ ,  $p_S = 1$ ,  $X^S = (1, \dots, 1)^T$ . It uses the link functions  $l^S(x, y) = l^N(x, y) = \exp(x + y)$ . The distribution of the Gamma part of the model is

only controlled by a one-dimensional parameter for the mean and by the shape parameter. It has  $n + 2$  parameters in total.

6. *CPNet2FG* uses the regression on the Gamma part of CPNet2 only. It has  $n+2$  parameters, and uses the following settings:  $p_N = 1$ ,  $X^N = (1, \dots, 1)^T$ ,  $p_S = n$ ,  $X^S = I_n$ ,  $l^S(x, y) = l^N(x, y) = \exp(x + y)$ .
7. *CPNet2FPG* is a CPNet2 model with fitness-based parameters on both the Poisson and the Gamma part of the model, i.e.  $p_N = n$ ,  $X^N = I_n$ ,  $p_S = n$ ,  $X^S = I_n$ ,  $l^S(x, y) = l^N(x, y) = \exp(x + y)$ . It thus has  $2n + 1$  parameters.
8. *CPNet2FPGmax* is the same as CPNet2FPG but with a different link function on the Poisson part, namely  $l^N(x, y) = \max(\exp(x), \exp(y))$ . It also has  $2n + 1$  parameters.
9. *GlmF* is a fitness based model that is not part of the compound Poisson Gamma family. GlmF is a combination of a logistic regression and a Gamma generalized linear model. The existence of a link, i.e.  $\mathbb{I}_{\{L_{ij} > 0\}}$ , is defined via a logistic regression model and the weight of the link (conditional on existence), i.e.  $L_{ij} | L_{ij} > 0$  is defined via a Gamma regression with the inverse function as the link function. Both regressions use as predictors only the row and the column index, i.e. the linear predictor for  $\mathbb{I}_{\{L_{ij} > 0\}}$  in the logistic regression is  $\theta_i^{\text{logistic}} + \theta_j^{\text{logistic}}$  and for  $L_{ij} | L_{ij} > 0$  the Gamma regression it is  $\theta_i^\Gamma + \theta_j^\Gamma$ . The overall parameter vector is  $(\theta_1^{\text{logistic}}, \dots, \theta_n^{\text{logistic}}, \theta_1^\Gamma, \dots, \theta_n^\Gamma, \phi)$ , where  $\phi$  is the dispersion parameter in the Gamma regression. Thus, the dimension of the parameter of the model is  $2n + 1$ .

All models are implemented in R [40]. All models but GlmF get fitted by optimising the likelihood using general-purpose optimisers (*optim*). The likelihood of the Tweedie, CPNet1 and CPNet2 models are using the methods developed in Dunn and Smyth [15]. GlmF uses the *glm* function available in R.

### 3.4 In-sample results

We now assess the fit of the models in the empirical case studies. Table 3 gives the Akaike information criterion (AIC) of the models, which is given by  $-2l + 2k$ , where  $l$  is the maximised log-likelihood of the model and  $k$  is the number of parameters in the model. To ease comparisons, we have subtracted the AIC of the basic ERE model for all datasets. Smaller numbers indicate a better fit.

In addition to the data from the case studies, two simulated networks are included (ER8 and ER50); these are simulated from the ERE model with 8 and 50 nodes, with  $p = 0.3$  and  $\lambda = 0.2$ . As expected, the true underlying model (ERE) performs best. For the networks from the case studies the picture is different.

We find that ERG and Tweedie outperform ERE. However, compared to the fitness-based models, their performance is relatively poor.

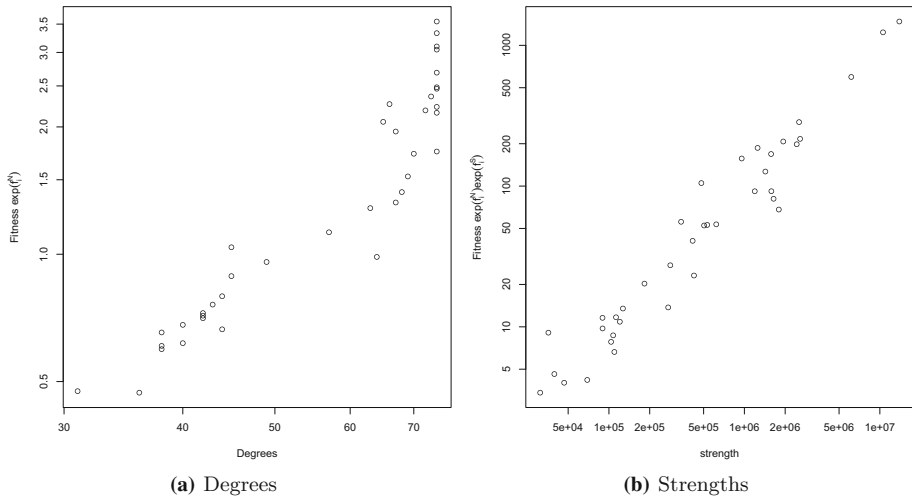
The models CPNet1F, CPNet2FP and CPNet2FG all have one fitness parameter per node and additionally two free parameters. For the LBS data, CPNet1F, which is modelling the overall mean in a regression, seems to be slightly better than modelling only the mean of the Poisson distribution (CPNet2FP) in most cases. For the CDS data modelling the mean of the Poisson distribution CPNet2FP is slightly better than modelling the overall mean (CPNet1F). In both cases, the results for modelling only the mean of the Gamma distribution (CPNet2FG) are worst.



**Table 3** Change in AIC compared to the basic ERE model for multiple methods on multiple data sets—smaller values indicate better fit

	ERE	ERG	Tweedie	CPNettIF	CPNet2FFP	CPNet2FG	CPNet2FPG	CPNet2FFPGmax	GlmF
ER8	<b>0</b>	2	2	<b>-0</b>	2	4	6	7	7
ER50	<b>0</b>	1	<b>-0</b>	53	58	54	113	98	122
CDS_1	0	-15	-14	-778	-803	-210	<b>-830</b>	-437	-818
CDS_2	0	-7	-8	-223	-287	-66	<b>-325</b>	-159	-310
CDS_3	0	1	1	-385	-497	-48	<b>-524</b>	-197	-511
CDS_4	0	-62	-60	-1677	-1696	-466	-1880	-1223	<b>-1933</b>
CDS_5	0	-52	-51	-2726	-2794	-591	-2972	-2076	<b>-3023</b>
LBS_2000_C	0	-436	-426	-1317	-1147	-988	<b>-1588</b>	-1566	-1469
LBS_2009_C	0	-1129	-1104	-2054	-1917	-1741	<b>-2468</b>	-2200	-2331
LBS_2018_C	0	-1945	-1897	-2886	-2673	-2691	<b>-3344</b>	-3109	-3063
LBS_2000_L	0	-879	-867	-1701	-1545	-1365	-1932	<b>-1953</b>	-1875
LBS_2009_L	0	-1455	-1426	-2531	-2205	-2184	<b>-2787</b>	-2729	-2648
LBS_2018_L	0	-3005	-2934	-4094	-3686	-3917	<b>-4539</b>	-4332	-4267
LBS_2000_T	0	-1347	-1308	-2940	-2869	-2182	-3371	<b>-3664</b>	-3505
LBS_2009_T	0	-2289	-2209	-3831	-3691	-3227	-4323	-4233	<b>-4451</b>
LBS_2018_T	0	-3827	-3677	-5301	-4961	-4947	-5775	<b>-5920</b>	-5555
Average	0	-1028	-999	-2024	-1919	-1535	<b>-2284</b>	-2106	-2227

Values in bold indicate the best fitting model for a given network



**Fig. 7** Degree and strength plotted against the fitted exponential fitness for the LBS data from 2018 Q1

Both CPNet2FPG and GLmF have two fitness parameters per node. They seem to be doing better than the models with just one fitness parameter. Their performance is somewhat comparable, with CPNet2FPG slightly outperforming GlmF overall.

CPNet2FPGmax seems to be doing slightly better for some of the networks, specifically the LBS networks with total liability.

Figure 7 illustrates the estimated fitnesses of one particular model fitted onto the LBS Data from 2018 Q1, the network used in Figs. 4, 5 and 6. The model we consider is CPNet2FPG which has  $2n + 1$  parameters. One plot shows the estimated exponential fitness  $\exp(f_i^N)$  of the Poisson part against the degrees of the nodes. The other plot shows the combined exponential fitnesses  $\exp(f_i^N) \exp(f_i^S)$  of the Poisson and the Gamma part against the strength. These demonstrate a good alignment and show that the fitnesses are indeed capturing the desired strength and degrees of the network.

### 3.5 Cross-validation results

Next, we use a cross-validation approach to compare the performance of the compound Poisson models. We partition the elements of the network matrix into tenfolds (roughly equally sized; all elements belong to exactly onefold) and they stay in these folds for the duration of the analysis. We fit the model using the data for ninefolds and then compute the log-likelihood (using the fitted model) of the remaining fold. We repeat the process for all folds and average the results. Hence, every observation is used to fit the model 9 times and is used to test the fit exactly once.

Table 4 presents the average log-likelihood in the testing fold. For the simulated data sets (ER8 and ER50), the true underlying model (ERE) does best as we would expect. For the CDS data sets, the CPNet2FP model seems to be doing best. For the LBS data set, it is one of the models with two fitness models—either CPNet2FPG or CPNet2FPGmax. For the CDS data sets, the CPNet1F data set seems to be doing badly.

Table 5 is also based on cross-validation, but with a different error criterion. It simulates from the fitted model 100 times and then reports the average accuracy (the proportion of

**Table 4** Average log-likelihood in the testing fold of a tenfold cross validation

	ERE	ERG	Tweedie	CPNet1F	CPNet2FP	CPNet2FG	CPNet2FFG	CPNet2FFPGmax
ER8	-9.6	-9.7	-9.7	-10.3	-10.4	-13.6	-25.1	-16.3
ER50	<b>-353.6</b>	-353.7	-353.7	-357.1	-357.2	-357.4	-361	-366.9
CDS_1	-304.6	-303.9	-303.9	-1.284e+06	<b>-275.7</b>	-305.4	-301.5	-305
CDS_2	-196.5	-196.2	-196.2	-4.161e+05	<b>-185.6</b>	-198.3	-230.5	-194.8
CDS_3	-212.6	-220.7	-212.6	-1.334e+12	<b>-200.5</b>	-228.2	-226.1	-234.1
CDS_4	-509.4	-506.1	-506.2	-9297	<b>-444.9</b>	-539.1	-544.2	-527.7
CDS_5	-787.6	-784.8	-785	-2457	<b>-690.7</b>	-981.7	-757	-808.3
LBS_2000_C	-457.1	-434.9	-435.4	-393.6	-401	-410.2	<b>-380.7</b>	-390.6
LBS_2009_C	-696.9	-639.8	-641.2	-611	-601.3	-634.5	<b>-582.1</b>	-618.2
LBS_2018_C	-903.6	-805.3	-807.7	-768.7	-770.8	-786.6	-750	-771.8
LBS_2000_L	-453.1	-408.3	-408.9	-374	-375.3	-393	-365	<b>-363.3</b>
LBS_2009_L	-3683	-582.3	-583.8	-533.9	-545.9	-553.5	<b>-525.4</b>	-536.5
LBS_2018_L	-1304	-1357	-776.1	-726.2	-739.6	-743.4	<b>-719.6</b>	-732.5
LBS_2000_T	-756.1	-688	-690	-612.5	-613.2	-650.5	-593.1	<b>-581.9</b>
LBS_2009_T	-1037	-921.2	-925.3	-859.3	-852.5	-888.8	-829.6	<b>-828.5</b>
LBS_2018_T	-1303	-1109	-1117	-1046	-1055	-1066	-1024	<b>-1020</b>

Higher values indicate a better fit. Values in bold indicate the best fitting model for a given network

**Table 5** Average accuracy in the cross validation (in per cent)

	ERE	ERG	Tweedie	CPNet1F	CPNet2FP	CPNet2FG	CPNet2FPG	CPNet2FPGmax
LBS_2000_C	62	62	62	74	<b>77</b>	62	<b>77</b>	75
LBS_2000_L	61	61	61	72	<b>76</b>	61	<b>76</b>	75
LBS_2000_T	51	51	51	71	81	50	83	<b>91</b>
LBS_2009_C	53	53	53	65	72	53	<b>73</b>	69
LBS_2009_L	38	53	53	66	<b>72</b>	53	71	68
LBS_2009_T	52	52	52	68	78	51	79	<b>88</b>
LBS_2018_C	50	50	50	58	<b>68</b>	50	<b>68</b>	63
LBS_2018_L	52	52	51	59	<b>69</b>	50	<b>69</b>	63
LBS_2018_T	66	66	65	72	79	65	81	<b>88</b>
ER50	<b>57</b>	<b>57</b>	<b>57</b>	<b>57</b>	<b>57</b>	<b>57</b>	<b>57</b>	<b>57</b>
ER8	50	50	51	53	<b>54</b>	49	53	52
CDS_1	67	67	67	86	<b>88</b>	67	<b>88</b>	74
CDS_2	51	51	51	69	78	51	79	64
CDS_3	74	75	75	85	<b>89</b>	74	<b>89</b>	80
CDS_4	84	84	84	92	<b>94</b>	84	<b>94</b>	87
CDS_5	92	92	92	95	<b>96</b>	92	<b>96</b>	93

Values in bold indicate the best fitting model(s) for a given network

elements that were correctly present/not present). The table reports the results in percent. Generally speaking, models that allow for a fitness parameter in the Poisson part of the model (such as CPNet2FP, CPNet2FPG or CPNet2FPGmax) are doing best. This is not surprising since in (1) we have seen that the link existence probabilities are directly determined by the fitnesses associated with the Poisson distribution.

## 4 Conclusion

We have introduced a new model class for directed and weighted random graphs with a fixed number of nodes in which each edge has a compound Poisson distribution for its weight. We have proposed different regression approaches to model features of the compound Poisson distribution. When fitting the new models to empirical network data we found that in most cases the compound Poisson models that model both the expectation of the Poisson random variable and the expectation of the Gamma distribution via separate regression models performed best (measured in terms of their AIC), i.e., the CPNet2 model class is preferable to the CPNet1 model class which itself is preferable to more basic Erdős–Rényi-type models.

In our tests on using these models for predicting subnetworks of a larger network we found the following. The CDS data exhibit a more traditional monotonic relationship between strengths and degrees. Consistent with this finding, the CPNet2 model class with only one regression for the mean of the Poisson distribution performed best for the CDS data. The LBS data do not have a monotonic relationship between strengths and degrees. Therefore, we found clear advantages of using the CPNet2 model with both a regression for the mean of the Poisson distribution and a separate regression for the mean of the Gamma distribution for the LBS data.

**Acknowledgements** We would like to thank the Bank of England for providing the data on the CDS exposures. In particular, we would like to thank Nicholas Vause (Bank of England) for his feedback and discussion. We would also like to thank the associate editor and two referees for their constructive feedback that helped us to improve the paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Acemoglu, D., Carvalho, V.M., Ozdaglar, A., Tahbaz-Salehi, A.: The network origins of aggregate fluctuations. *Econometrica* **80**, 1977–2016 (2012)
2. Albert, R., Barabási, A.-L.: Statistical mechanics of complex networks. *Revi. Mod. Phys.* **74**, 47 (2002)
3. Aldasoro, I., Ehlers, T.: Concentration in cross-border banking. *BIS Q. Rev.* June 2019, 1–11 (2019)
4. Barrat, A., Barthélemy, M., Pastor-Satorras, R., Vespignani, A.: The architecture of complex weighted networks. *Proc. Natl. Acad. Sci.* **101**, 3747–3752 (2004)
5. Boss, M., Elsinger, H., Summer, M., Thurner, S.: Network topology of the interbank market. *Quant. Finance* **4**, 677–684 (2004)

6. Capponi, A.: Systemic risk, policies, and data needs. In: Gupta, A., Capponi, A., Smith, J.C., Greenberg, H.J. (eds.) *Optimization Challenges in Complex, Networked and Risky Systems*, pp. 185–206. INFORMS, Catonsville (2016)
7. Capponi, A., Larsson, M.: Price contagion through balance sheet linkages. *Rev. Asset Pricing Stud.* **5**, 227–253 (2015)
8. Caron, F., Fox, E.B.: Sparse graphs using exchangeable random measures. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **79**, 1295–1366 (2017)
9. Cifuentes, R., Ferrucci, G., Shin, H.S.: Liquidity risk and contagion. *J. Eur. Econ. Assoc.* **3**, 556–566 (2005)
10. Cont, R., Moussa, A., Santos, E.B.: Network structure and systemic risk in banking systems. In: Fouque, J.-P., Langsam, J.A. (eds.) *Handbook on Systemic Risk*. Cambridge University Press, Cambridge (2010)
11. Cont, R., Schaanning, E.: Fire sales, indirect contagion and systemic stress testing. *Norges Bank Working Paper 2/2017* (2017)
12. Cont, R., Wagalath, L.: Fire sales forensics: measuring endogenous risk. *Math. Finance* **26**, 835–866 (2016)
13. Craig, B., Von Peter, G.: Interbank tiering and money center banks. *J. Financ. Intermed.* **23**, 322–347 (2014)
14. DTTC (2015). Trade information warehouse explanation of TIW reports. <https://www.dtccdata.com/~media/data-products-portal/files/resources/tiw/tiw-market-report-guide.pdf> (2015)
15. Dunn, P.K., Smyth, G.K.: Series evaluation of Tweedie exponential dispersion model densities. *Stat. Comput.* **15**, 267–280 (2005)
16. Eisenberg, L., Noe, T.H.: Systemic risk in financial systems. *Manag. Sci.* **47**, 236–249 (2001)
17. Erdős, P., Rényi, A.: On random graphs I. *Publ. Math. Debr.* **6**, 290–297 (1959)
18. Fricke, D., Lux, T.: Core-periphery structure in the overnight money market: evidence from the e-mid trading platform. *Comput. Econ.* **45**, 359–395 (2015)
19. Gandy, A., Veraart, L.A.M.: A Bayesian methodology for systemic risk assessment in financial networks. *Manag. Sci.* **63**, 4428–4446 (2017)
20. Gandy, A., Veraart, L.A.M.: Adjustable network reconstruction with applications to CDS exposures. *J. Multivar. Anal.* **172**, 193–209 (2019)
21. Glasserman, P., Young, H.P.: Contagion in financial networks. *J. Econ. Lit.* **54**, 779–831 (2016)
22. Greenwood, R., Landier, A., Thesmar, D.: Vulnerable banks. *J. Financ. Econ.* **115**, 471–485 (2015)
23. Holland, P.W., Leinhardt, S.: An exponential family of probability distributions for directed graphs. *J. Am. Stat. Assoc.* **76**, 33–50 (1981)
24. In't Veld, D., van Lelyveld, I.: Finding the core: network structure in interbank markets. *J. Bank. Finance* **49**, 27–40 (2014)
25. Iori, G., De Masi, G., Precup, O.V., Gabbi, G., Caldarelli, G.: A network analysis of the Italian overnight money market. *J. Econ. Dyn. Control* **32**, 259–278 (2008)
26. Jacobs, A.Z., Clauset, A.: A unified view of generative models for networks: models, methods, opportunities, and challenges. *ArXiv preprint arXiv:1411.4070* (2014)
27. Jorgensen, B.: Exponential dispersion models. *J. R. Stat. Soc. Ser. B (Methodol.)* **49**, 127–162 (1987)
28. Koller, D., Friedman, N.: *Probabilistic Graphical Models: Principles and Techniques*. MIT Press, Cambridge (2009)
29. Lee, S.H.: Systemic liquidity shortages and interbank network structures. *J. Financ. Stab.* **9**, 1–12 (2013)
30. Lovász, L.: *Large Networks and Graph Limits*, vol. 60. American Mathematical Society, Providence (2012)
31. Martínez-Jaramillo, S., Alexandrova-Kabadjova, B., Bravo-Benitez, B., Solórzano-Margain, J.P.: An empirical study of the Mexican banking system's network and its implications for systemic risk. *J. Econ. Dyn. Control* **40**, 242–265 (2014)
32. Meinshausen, N., Bühlmann, P.: High-dimensional graphs and variable selection with the lasso. *Ann. Stat.* **34**, 1436–1462 (2006)
33. Newman, M.: *Networks: An Introduction*. Oxford University Press, Oxford (2010)
34. Newman, M.E.: The structure and function of complex networks. *SIAM Rev.* **45**, 167–256 (2003)
35. Norros, I., Reittu, H.: On a conditionally Poissonian graph process. *Adv. Appl. Probab.* **38**, 59–75 (2006)
36. Oatley, T., Winecoff, W.K., Pennock, A., Danzman, S.B.: The political economy of global finance: a network model. *Perspect. Polit.* **11**, 133–153 (2013)
37. Orbanz, P., Roy, D.M.: Bayesian models of graphs, arrays and other exchangeable random structures. *IEEE Trans. Pattern Anal. Mach. Intell.* **37**, 437–461 (2015)
38. Park, J., Newman, M.E.: Statistical mechanics of networks. *Phys. Rev. E* **70**, 066117 (2004)
39. Peltonen, T.A., Scheicher, M., Vuillemeys, G.: The network structure of the CDS market and its determinants. *J. Financ. Stab.* **13**, 118–133 (2014)

40. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2018)
41. Ranola, J.M., Ahn, S., Sehl, M., Smith, D.J., Lange, K.: A Poisson model for random multigraphs. *Bioinformatics* **26**, 2004–2011 (2010)
42. Rogers, L.C.G., Veraart, L.A.M.: Failure and rescue in an interbank network. *Manag. Sci.* **59**, 882–898 (2013)
43. Upper, C., Worms, A.: Estimating bilateral exposures in the German interbank market: is there a danger of contagion? *Eur. Econ. Rev.* **48**, 827–849 (2004)
44. Veraart, L.A.M.: Distress and default contagion in financial networks. *Mathematical Finance*. <https://doi.org/10.1111/mafi.12247> (2020)
45. Wells, S.J.: Financial interlinkages in the United Kingdom's interbank market and the risk of contagion. Bank of England working paper (2004)
46. Wilson, J.D., Denny, M.J., Bhamidi, S., Cranmer, S.J., Desmarais, B.A.: Stochastic weighted graphs: flexible model specification and simulation. *Soc. Netw.* **49**, 37–47 (2017)
47. Wolfe, P.J., Olhede, S.C.: Nonparametric graphon estimation. ArXiv preprint [arXiv:1309.5936](https://arxiv.org/abs/1309.5936) (2013)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.