Joint Statistical Meetings 2021 – Virtual Conference

# Mutually exciting point process graphs for modelling dynamic networks

Imperial College London

**Francesco Sanna Passino**, Nick Heard
Department of Mathematics, Imperial College London
f.sannapassino@imperial.ac.uk

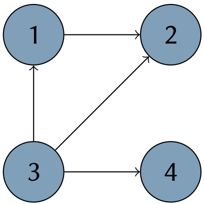
*11th August, 2021*

JSM 2021 Virtual Conference

# Dynamic graphs as point processes with dyadic marks

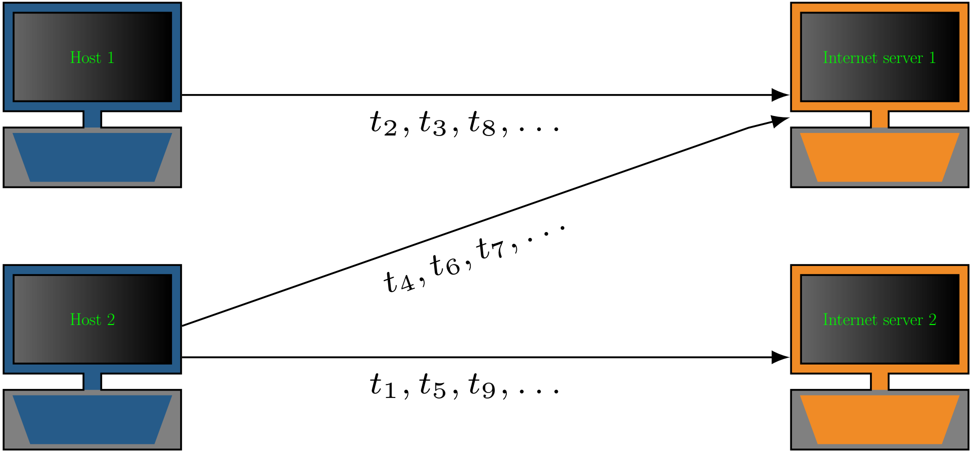
- Event data from dynamic networks are observed as triplets $(t_1, x_1, y_1), \dots, (t_m, x_m, y_m)$, where $0 \leq t_1 \leq t_2 \leq \dots$ are event times and the dyadic marks $(x_k, y_k)$ denote the source and destination nodes, each belonging to a set of nodes $V = \{1, \dots, n\}$ of size $n$.
- The sequence of graph edges $(x_1, y_1), \dots, (x_m, y_m)$ induces a directed *network adjacency matrix* $\mathbf{A} = \{A_{ij}\} \in \{0,1\}^{n \times n}$ where $A_{ij} = 1$ if node $i$ connected to node $j$ at least once during the entire observation period, and $A_{ij} = 0$ otherwise.

$$\begin{matrix} (1.3, 1, 2) \\ (1.7, 3, 1) \\ (1.9, 1, 2) \\ (2.2, 3, 4) \\ (2.7, 3, 2) \\ \vdots \end{matrix} \quad \Rightarrow$$

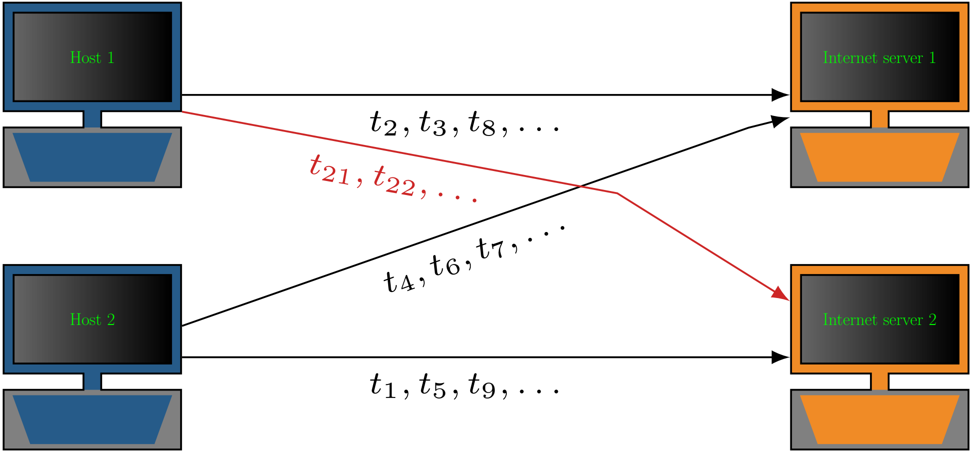



$$\Rightarrow \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

# Motivation: new links in cyber-security



# Motivation: new links in cyber-security



# Background

- Objective: propose a model which can calculate *anomaly scores* for *unobserved marks.*
- Motivation: computer network attacks tend to form previously unobserved connections.
- Related literature in cyber-security: Price-Williams and Heard, 2020, demonstrate that *self-exciting processes* have an excellent performance for modelling individual edges.
- Related methodology: new link prediction in networks. Latent position models (LPMs, Hoff, Raftery, and Handcock, 2002) postulate that the probability of a link is a function of *node-specific* latent features:

$$\mathbb{P}(A_{ij}=1 \mid \boldsymbol{x}_i, \boldsymbol{x}_j) = \kappa(\boldsymbol{x}_i, \boldsymbol{x}_j),\ \boldsymbol{x}_i, \boldsymbol{x}_j \in \mathbb{R}^d,$$

  where $\kappa(\cdot)$ is a kernel function. Conditional on the latent positions, LPMs naturally allow to calculate link probabilities for unobserved links.
- Model proposed in this work: *mutually exciting process* on *each edge*, parametrised only by *node-specific parameters.*

Virtual Conference JSM 2021

# Mutually exciting graphs (MEGs)

- MEGs are defined by a time-varying matrix of non-negative functions $\boldsymbol{\lambda}(t) = \{\lambda_{ij}(t)\}$.
- Each entry $\lambda_{ij}(t)$ represents the intensity of the counting process of events occurring on the edge $(i,j)$: $N_{ij}(t) = \sum_{k=1}^{m} \mathbb{1}_{[0,t]\times\{i\}\times\{j\}}(t_k, x_k, y_k)$.
- For generality, it is assumed that for each edge $(i,j)$ there exists a changepoint $\tau_{ij} \geq 0$ after which the edge becomes observable. In the simplest case, $\tau_{ij} = 0 \ \forall\, i, j$.
- Each entry of $\lambda_{ij}(t)$ is represented as an additive model with three components:
  - The first, denoted $\alpha_i(t)$, characterises the process of arrival times involving $i$ as source node;
  - The second, $\beta_j(t)$, corresponds to arrivals for which $j$ is the destination node;
  - The third, $\gamma_{ij}(t)$, is an interaction term, also be parameterised by node-specific parameters.

$$\lambda_{ij}(t) = \alpha_i(t) + \beta_j(t) + \gamma_{ij}(t), \quad t \geq \tau_{ij}. \tag{1}$$

- The intensity function resembles the link function used in *additive and multiplicative effect network* (AMEN) models for network adjacency matrices (Hoff, 2018).

# Main effects

- Define the source and destination counting processes as $N_i(t) = \sum_{k=1}^m \mathbb{1}_{[0,t]\times\{i\}}(t_k, x_k)$ and $N_j'(t) = \sum_{k=1}^m \mathbb{1}_{[0,t]\times\{j\}}(t_k, y_k)$.
- Let $\ell_{i1}, \ell_{i2}, \dots$ denote the event indices $\{k : x_k = i\}$ such that $i$ appears as source node, and $\ell_{j1}', \ell_{j2}', \dots$ denote the event indices $\{k : y_k = j\}$ for which $j$ is the destination node.
- To allow self-excitation of both source and destination nodes, the latent functions $\alpha_i(t)$ and $\beta_j(t)$ are assigned the following form:

$$\alpha_i(t) = \alpha_i + \sum_{k>N_i(t)-r}^{N_i(t)} \omega_i(t - t_{\ell_{ik}}), \quad \beta_j(t) = \beta_j + \sum_{k>N_j'(t)-r}^{N_j'(t)} \omega_j'(t - t_{\ell_{jk}'}),$$

where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n), \boldsymbol{\beta} = (\beta_1, \dots, \beta_n) \in \mathbb{R}_+^n$ are node-specific baseline intensity levels, and $\omega_i, \omega_i'$ are node-specific, non-increasing excitation functions from $\mathbb{R}_+$ to $\mathbb{R}_+$.
- For simplicity, the excitation functions assume the following scaled exponential form, for non-negative parameters $\boldsymbol{\mu}_i, \boldsymbol{\mu}_j', \boldsymbol{\phi}_i, \boldsymbol{\phi}_j' \in \mathbb{R}_+^n$:

$$\omega_i(t) = \mu_i \exp\{-(\mu_i + \phi_i)t\}, \quad \omega_j'(t) = \mu_j' \exp\{-(\mu_j' + \phi_j')t\}.$$

# Interactions

- Let $\ell_{ij1}, \ell_{ij2}, \dots$ be the indices $\{k : x_k = i, y_k = j\}$ of events observed on the edge $(i,j)$.
- The interaction term $\gamma_{ij}(t)$ in (1) assumes a similar form to the main effects, but with a background rate obtained as the inner product between two node-specific $d$-dimensional baseline parameter vectors $\boldsymbol{\gamma}_i, \boldsymbol{\gamma}_j^\prime \in \mathbb{R}_+^d$:

$$\gamma_{ij}(t) = \boldsymbol{\gamma}_i \cdot \boldsymbol{\gamma}_j^\prime + \sum_{k > N_{ij}(t)-r}^{N_{ij}(t)} \omega_{ij}(t - t_{\ell_{ijk}}),$$

- The inner product baseline is inspired by random dot product graphs (RDPGs; see, for example, Athreya et al., 2018) for link probabilities.
- For simplicity, the excitation function $\omega_{ij}(t)$ is expressed as a sum of scaled exponentials, parameterised by four node-specific, non-negative latent $d$-vectors $\boldsymbol{\nu}_, \boldsymbol{\nu}_j^\prime, \boldsymbol{\theta}_i, \boldsymbol{\theta}_j^\prime \in \mathbb{R}_+^d$:

$$\omega_{ij}(t) = \sum_{\ell=1}^{d} \nu_{i\ell}\nu_{j\ell}^\prime \exp\{-(\theta_{i\ell} + \nu_{i\ell})(\theta_{j\ell}^\prime + \nu_{j\ell}^\prime)t\}.$$

# MEGs: an example

- The integer parameter $r$ expresses *how many events* are taken into account in the intensity function. There are three limiting cases:
  - $r = 0$: Poisson process (the process is independent of previous events);
  - $r = 1$: Markov process (dependence only on the distance to the last event);
  - $r \to \infty$: Hawkes process (dependence on the entire history of the process).

**(a)** Main effect $\alpha_i(t)$ **(b)** Main effect $\beta_j(t)$ **(c)** Interaction $\gamma_{ij}(t)$

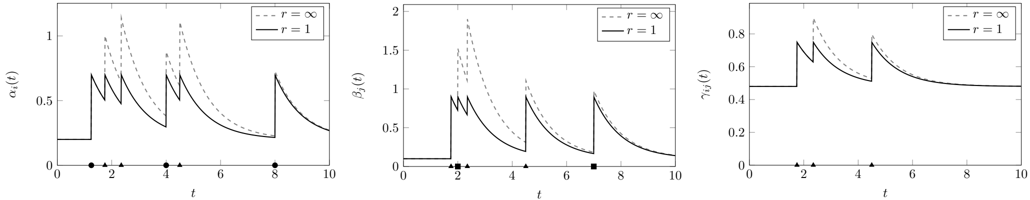



**Figure 1.** Cartoon of a 1-dimensional MEG model with $\alpha_i = 0.2,\ \mu_i = 0.5,\ \phi_i = 0.5,\ \beta_j = 0.1,\ \mu'_j = 0.8,\ \phi'_j = 0.2,\ \gamma_i = 0.8,\ \nu_i = 0.9,\ \theta_i = 1.1,\ \gamma'_j = 0.6,\ \nu'_j = 0.3,\ \theta'_j = 0.2$. Events with source node $i$ and destination node $j$ are denoted by triangles; other events with source node $i$ are are denoted with circles, and other events with destination node $j$ are denoted by squares.

# MEGs: the resulting edge intensity function

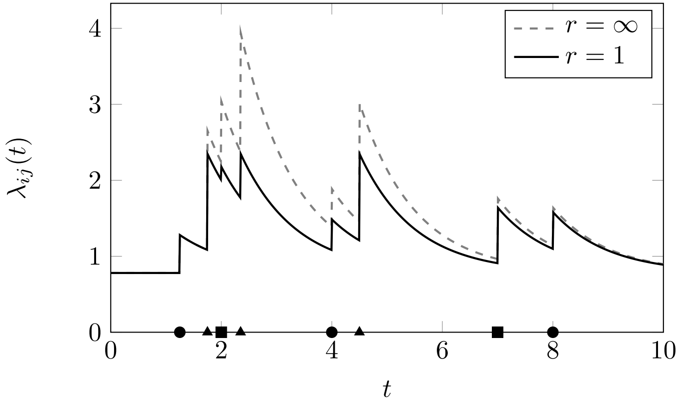



**Figure 2.** Cartoon of a 1-dimensional MEG model with $\alpha_i = 0.2,\ \mu_i = 0.5,\ \phi_i = 0.5,\ \beta_j = 0.1,\ \mu_j^\prime = 0.8,\ \phi_j^\prime = 0.2,\ \gamma_i = 0.8,\ \nu_i = 0.9,\ \theta_i = 1.1,\ \gamma_j^\prime = 0.6,\ \nu_j^\prime = 0.3,\ \theta_j^\prime = 0.2$. Events with source node $i$ and destination node $j$ are denoted by triangles; other events with source node $i$ are are denoted with circles, and other events with destination node $j$ are denoted by squares.

## Inference and goodness-of-fit

- **Inference: likelihood optimisation via gradient ascent.** For a sequence of observed events $\mathcal{H}_T = \{(x_1, y_1, t_1), \dots, (x_m, y_m, t_m)\}$, the log-likelihood is:

$$\log L(\mathcal{H}_T; \boldsymbol{\Psi}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \left\{ \sum_{k=1}^{n_{ij}} \log \lambda_{ij}(t_{\ell_{ijk}}) - \int_{\tau_{ij}}^{T} \lambda_{ij}(t) \mathrm{d}t \right\},$$

where $n_{ij}$ is the number of events observed on the edge $(i, j)$.

- **Goodness-of-fit: distribution of the $p$-values of out-of-sample events.** Given arrival times $t_1, \dots, t_{n_{ij}}$ on the edge $(i, j)$, the upper tail $p$-values are:

$$p_k = \exp\left\{ -\int_{t_{k-1}}^{t_k} \lambda_{ij}(s) \mathrm{d}s \right\}, k = 1, \dots, n_{ij}.$$

Under the null hypothesis of correct specification of the conditional intensity $\lambda_{ij}(t)$, the $p$-values are uniformly distributed.

# Results: Enron e-mail network

- The Enron e-mail network collection is a record of e-mails exchanged between the employees of Enron Corporation before its bankruptcy.
- These data have already been demonstrated to be well-modelled as self-exciting point processes by Fox et al., 2016.
- $34{,}427$ distinct triplets $(x_k, y_k, t_k)$, corresponding to messages exchanged between $n = 184$ employees between November 1998 and June 2002, forming a total of $3{,}007$ edges.
- Some of the emails are sent to multiple receivers, and only $18{,}031$ unique event times are observed, implying that on average each e-mail is sent to approximately $1.90$ nodes.
- Because an e-mail can have multiple recipients, and because the event times are recorded to the nearest second, the likelihood must be adapted with the arrivals modelled by an analogous discrete time process.
- The model is trained on $30{,}704$ e-mails, and tested on the remaining $3{,}723$ e-mails.
- In the training set, $2{,}720$ edges are observed, and $811$ in the test set, of which $287$ are *not* observed in the training period.

# Results: Enron e-mail data, $\tau_{ij} = t_{\ell_{ij1}}$

**Table 1.** Training and test KS scores on the Enron e-mail network for different configurations of the MEG model.

| **KS scores (train & test)** | | | Main effects $\alpha_i(\cdot)$ and $\beta_j(\cdot)$ ↓ | | | |
|---|---|---|---|---|---|---|
| $\tau_{ij}$ ↓ | Interactions $\gamma_{ij}(\cdot)$ ↓ | | Absent | Poisson ($r=0$) | Markov ($r=1$) | Hawkes ($r=\infty$) |
| $\tau_{ij} = t_{\ell_{ij1}}$ (MLE) | Absent | | – – | 0.4530 0.4133 | 0.3678 0.3484 | 0.4443 0.3586 |
| | Poisson ($r=0$) | $d=1$ | 0.4252 0.4221 | 0.3946 0.4179 | 0.3434 0.3574 | 0.4255 0.3560 |
| | | $d=5$ | 0.3490 0.3851 | 0.3498 0.3953 | 0.3165 0.3677 | 0.3491 0.3613 |
| | | $d=10$ | 0.3339 0.3763 | 0.3347 0.3688 | 0.3112 0.3470 | 0.3376 0.3575 |
| | Markov ($r=1$) | $d=1$ | 0.1662 0.2029 | 0.1491 0.1945 | 0.1305 0.1777 | 0.1702 0.1874 |
| | | $d=5$ | 0.0916 0.1875 | 0.0910 0.1684 | **0.0885 0.1628** | 0.0916 0.1746 |
| | | $d=10$ | **0.0885** 0.1743 | **0.0885** 0.1848 | **0.0885** 0.1696 | **0.0885** 0.1743 |
| | Hawkes ($r=\infty$) | $d=1$ | 0.2640 0.2755 | 0.2825 0.2887 | 0.2538 0.2637 | 0.2599 0.2871 |
| | | $d=5$ | 0.2304 0.2904 | 0.2284 0.2760 | 0.2271 0.2774 | 0.2420 0.2981 |
| | | $d=10$ | 0.2461 0.2923 | 0.2521 0.2865 | 0.2413 0.3091 | 0.2498 0.3129 |

- The MLE approach has a drawback: the $p$-values for the first observation on each edge are *always* 1. This implies that the KS scores are bounded below by $2720/30704 \approx 0.0885$ for the training set and $287/3723 \approx 0.0770$ for the test set.

# Results: Enron e-mail data, $\tau_{ij} = 0$

Table 2. Training and test KS scores on the Enron e-mail network for different configurations of the MEG model.

| KS scores (train & test) | | | Main effects $\alpha_i(\cdot)$ and $\beta_j(\cdot)\downarrow$ | | | |
|---|---|---|---|---|---|---|
| $\tau_{ij}\downarrow$ | Interactions $\gamma_{ij}(\cdot)\downarrow$ | | Absent | Poisson $(r=0)$ | Markov $(r=1)$ | Hawkes $(r=\infty)$ |
| $\tau_{ij}=0$ | Absent | | – – | 0.4530 0.4133 | 0.3678 0.3484 | 0.4443 0.3586 |
| | Poisson $(r=0)$ | $d=1$ | 0.7039 0.7926 | 0.6627 0.7753 | 0.6543 0.7148 | 0.7059 0.6050 |
| | | $d=5$ | 0.5623 0.7059 | 0.5646 0.7206 | 0.5748 0.7008 | 0.7060 0.6053 |
| | | $d=10$ | 0.5354 0.6853 | 0.5332 0.6739 | 0.5725 0.6952 | 0.7060 0.6059 |
| | Markov $(r=1)$ | $d=1$ | 0.3135 0.3324 | 0.3004 0.3326 | 0.3262 0.3240 | 0.2027 0.1999 |
| | | $d=5$ | 0.0760 0.1664 | 0.0825 0.1584 | 0.0855 0.1782 | 0.0495 **0.0924** |
| | | $d=10$ | 0.0775 0.1649 | 0.0793 0.1546 | 0.0816 0.1606 | **0.0402** 0.0971 |
| | Hawkes $(r=\infty)$ | $d=1$ | 0.2871 0.2486 | 0.2333 0.2449 | 0.2485 0.2379 | 0.1749 0.1991 |
| | | $d=5$ | 0.1939 0.2167 | 0.1885 0.2246 | 0.2010 0.2137 | 0.1467 0.1994 |
| | | $d=10$ | 0.2029 0.2395 | 0.2158 0.2470 | 0.2207 0.2339 | 0.1606 0.1943 |

- Comparison to alternative node-based models:
  - Poisson processes $\lambda_i(t)=\alpha_i$. KS score: $0.4088$;
  - Hawkes processes $\lambda_i(t)=\alpha_i+\sum_{k=1}^{N_i(t)}\mu_i\exp\{-(\mu_i+\phi_i)(t-t_{ik})\}$. KS score: $0.2499$;
  - Mutually exciting process with intensity $\lambda_i(t)=\alpha_i+\sum_{k=1}^{N_i'(t)}\mu_i\exp\{-(\mu_i+\phi_i)(t-t_{ik}')\}$ (Fox et al., 2016). KS score: $0.2806$. It could be inferred that users tend to respond to multiple e-mails in sessions, and not necessarily immediately after an individual e-mail is received.

# Results: Enron e-mail data, $\tau_{ij} = 1/A_{ij}$

**Table 3.** Training and test KS scores on the Enron e-mail network for different configurations of the MEG model.

| **KS scores (train & test)** | | | Main effects $\alpha_i(\cdot)$ and $\beta_j(\cdot)$ ↓ | | | |
|---|---|---|---|---|---|---|
| $\tau_{ij}$ ↓ | Interactions $\gamma_{ij}(\cdot)$ ↓ | | Absent | Poisson ($r=0$) | Markov ($r=1$) | Hawkes ($r=\infty$) |
| $\tau_{ij} = \begin{cases} 0, & A_{ij}=1 \\ \infty, & A_{ij}=0 \end{cases}$ | Absent | | – – | 0.4530 0.4133 | 0.3678 0.3484 | 0.4443 0.3586 |
| | Poisson ($r=0$) | $d=1$ | 0.5158 0.6038 | 0.4812 0.5864 | 0.3742 0.3602 | 0.4197 0.2808 |
| | | $d=5$ | 0.4269 0.5516 | 0.4309 0.5641 | 0.3553 0.3598 | 0.3938 0.2803 |
| | | $d=10$ | 0.4035 0.5413 | 0.4084 0.5565 | 0.3430 0.3537 | 0.3659 0.2810 |
| | Markov ($r=1$) | $d=1$ | 0.1950 0.2115 | 0.1600 0.2017 | 0.1504 0.1422 | 0.1309 0.1445 |
| | | $d=5$ | 0.0709 0.1222 | 0.0746 0.1008 | 0.0696 0.0917 | **0.0152** 0.0848 |
| | | $d=10$ | 0.0619 0.1029 | 0.0627 0.1079 | 0.0634 0.0836 | 0.0213 **0.0800** |
| | Hawkes ($r=\infty$) | $d=1$ | 0.1870 0.2084 | 0.1816 0.2049 | 0.1783 0.1747 | 0.1719 0.1879 |
| | | $d=5$ | 0.1377 0.1805 | 0.1374 0.1840 | 0.1391 0.1642 | 0.1553 0.2154 |
| | | $d=10$ | 0.1556 0.2023 | 0.1588 0.2046 | 0.1546 0.1863 | 0.1640 0.2082 |

- The best performance (KS score $0.0152$) is achieved when a Markov process is used for the interaction, with $d=5$ or $d=10$, combined with a Hawkes process for the main effects.
- In general, setting $\tau_{ij}$ using the adjacency matrix seems to outperform competing strategies for estimation of $\tau_{ij}$ in terms of KS scores.
- The interaction term should be included in the model.

Virtual Conference JSM 2021

# Results: ICL NetFlow data

- Many enterprises routinely collect network flow (NetFlow) data, representing summaries of connections between internet protocol (IP) addresses.
- The data consists of $1{,}951{,}067$ arrival times (in milliseconds), recorded in three weeks.
  - Sources: $n_1 = 173$ clients hosted within the Department of Mathematics at ICL;
  - Destinations: $n_2 = 6{,}083$ internet servers connecting on ports 80 and 443;
  - $156{,}186$ unique edges in total.
- Only edges such that the percentage of arrival times observed between 7am and 12am is larger than $99\%$, corresponding to the college opening hours, were considered.
- The MEG model is trained on the first two weeks of data, corresponding to $1{,}299{,}372$ events, and tested on $651{,}695$ events observed in the final week.
- The number of unique edges in the training set is $115{,}600$, and $70{,}408$ in the test set.
- Only $29{,}822$ edges are observed in both time windows, which implies that $40{,}586$ new edges are observed in the test set.

Virtual Conference JSM 2021

# Results: ICL NetFlow data, KS scores

**Table 4.** Kolmogorov-Smirnov scores on the ICL NetFlow data for different configurations of the MEG model.

| **KS scores (train & test)** | | Main effects $\alpha_i(\cdot)$ and $\beta_j(\cdot)$ ↓ | | | |
|---|---|---|---|---|---|
| Interactions $\gamma_{ij}(\cdot)$ ↓ | | Absent | Poisson ($r = 0$) | Markov ($r = 1$) | Hawkes ($r = \infty$) |
| Absent | | – – | 0.7351 0.7148 | 0.6678 0.6489 | 0.7312 0.6950 |
| Poisson ($r = 0$) | $d = 1$ | 0.7328 0.7157 | 0.7325 0.7150 | 0.6672 0.6480 | 0.7316 0.6960 |
| | $d = 5$ | 0.7295 0.7167 | 0.7313 0.7123 | 0.6673 0.6487 | 0.7275 0.6967 |
| | $d = 10$ | 0.7260 0.7174 | 0.7289 0.7140 | 0.6680 0.6493 | 0.7270 0.6969 |
| Markov ($r = 1$) | $d = 1$ | 0.2194 0.1723 | 0.2242 0.1657 | 0.2038 0.1440 | 0.1645 0.1281 |
| | $d = 5$ | 0.1024 0.1080 | 0.0896 0.0805 | **0.0728 0.0738** | 0.1041 0.0899 |
| | $d = 10$ | 0.0843 0.0764 | 0.0871 0.0761 | 0.0850 0.0843 | 0.1100 0.0883 |
| Hawkes ($r = \infty$) | $d = 1$ | 0.1080 0.0802 | 0.0747 0.1182 | 0.1082 0.0794 | 0.0884 0.1262 |
| | $d = 5$ | 0.1576 0.1819 | 0.1532 0.2126 | 0.1677 0.2143 | 0.2307 0.2383 |
| | $d = 10$ | 0.1584 0.1935 | 0.1546 0.2112 | 0.1619 0.2206 | 0.2388 0.2503 |

- $\tau_{ij} = 0$ if $A_{ij} = 1$, $\tau_{ij} = \infty$ if $A_{ij} = 0$.

# Results: ICL NetFlow data, Q-Q plots

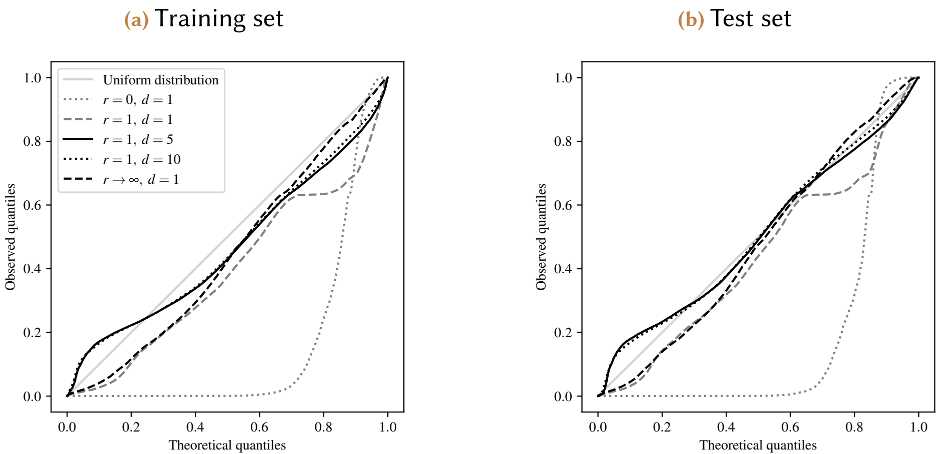
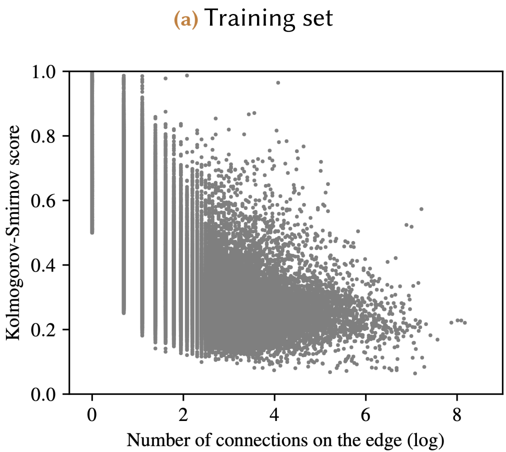



**Figure 3.** Q-Q plots for the training and test $p$-values obtained from different MEG models, with main effects $\alpha_i(t)$ and $\beta_j(t)$ with $r=1$, and different parameters for the interaction term $\gamma_{ij}(t)$, specified in the legend.

# Results: ICL NetFlow data, KS scores



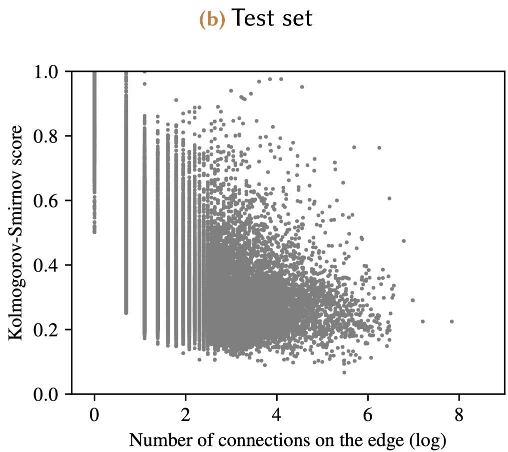




**Figure 4.** Scatterplot of the Kolmogorov-Smirnov scores, calculated for each edge, versus the logarithm of the total number of connections on the edge, for the best performing model in Table 4.

# Conclusion

- Mutually-exciting graphs (MEG), network-wide models for point processes with dyadic marks, have been proposed.
- Edge-specific intensities are obtained only via node-specific parameters, which is useful for large and sparse graphs.
- MEG is able to predict events observed on *new* edges.
- MEG greatly outperforms results previously obtained in the literature on the Enron e-mail network.
- More details in Sanna Passino and Heard, 2021. Scan the QR code to get the ar$\chi$iv preprint!
- *python* code on GitHub: `fraspass/meg`.

Sanna Passino, F. and N. A. Heard (2021). "Mutually exciting point process graphs for modelling dynamic networks". In: *arXiv e-prints*. arXiv: 2102.06527 [cs.SI].

# References


- Athreya, A. et al. (2018). "Statistical Inference on Random Dot Product Graphs: a Survey". In: *Journal of Machine Learning Research* 18.226, pp. 1–92.
- Fox, E.W. et al. (2016). "Modeling E-mail Networks and Inferring Leadership Using Self-Exciting Point Processes". In: *Journal of the American Statistical Association* 111.514, pp. 564–584.
- Hoff, P. D. (2018). "Additive and multiplicative effects network models". In: *arXiv e-prints*. arXiv: 1807.08038 [stat.ME].
- Hoff, P. D., A. E. Raftery, and M. S. Handcock (2002). "Latent space approaches to social network analysis". In: *Journal of the American Statistical Association* 97.460, pp. 1090–1098.
- Price-Williams, M. and N. A. Heard (2020). "Nonparametric Self-exciting Models for Computer Network Traffic". In: *Statistics and Computing* 30.2, pp. 209–220.
- Sanna Passino, F. and N. A. Heard (2021). "Mutually exciting point process graphs for modelling dynamic networks". In: *arXiv e-prints*. arXiv: 2102.06527 [cs.SI].


Virtual Conference JSM 2021