Modelling non-stationarity and shock resilience in financial temporal networks

Fabrizio Lillo

Universitá di Bologna and Scuola Normale Superiore, Pisa, Italy

Fields Institute

Toronto, June 26, 2024

This research is funded by the European Union Next Generation EU with grant number PNRRIR0000013 'SoBigData.it'.

Finanziato dall'Unione europea NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

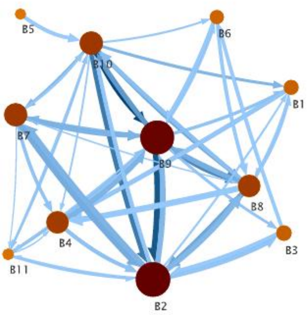
Outline

- Our case study: the interbank network
- The importance of statistical models: identifying macroscopic structure
- Statistical models for temporal networks
  - Modeling preferential lending
  - Lead-lag relations and modeling of resilience
  - Tackling non-stationarity with score driven models

Direct Relations Among Financial Institutions

Example: Weighted Directed Networks of Inter-Bank Loans

- Banks establish mutual credit relation with different maturity and/or collateral in order to fund themselves or to use excess liquidity
- Interbank markets allow banks to cope with specific liquidity shocks
- The interbank market is one of the important channels of propagation of shocks and systemic risk



Banks

$$\leftrightarrow \begin{pmatrix} 10^{10}€ & . & . & . & . \\ . & . & . & . & . \\ . & . & Y_{nk} & . & . \\ . & . & . & . & . \\ . & . & . & . & 10^{7}€ \end{pmatrix}$$

Banks

Picture from Giraitis et. al. 2016

The interbank market: stylized facts

- Very low connectivity: only $\sim 1\%$ of links are present
- Power law tailed distribution of in- and out- degree. Estimated exponent $\alpha \simeq 2 \div 3$
- The distribution of interbank exposures (node strenght) is also heavy-tailed.
- Low average distance between nodes,
- A disassortative mixing, i.e. the tendency of high degree nodes to connect with low degree nodes,
- Small clustering

e-MID: Electronic Market for Inter-Bank Deposits

Data Description

Inter-bank Over-Night Loans

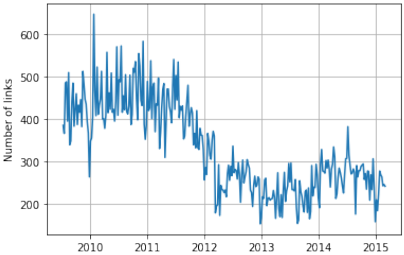
- 132 banks, 297 weekly networks between 2009 and 2015
- Sparse networks, density ≤ 0.08
- The network might be affected by shocks that hits a subset of nodes
- Example: Lehman default, LTRO, etc

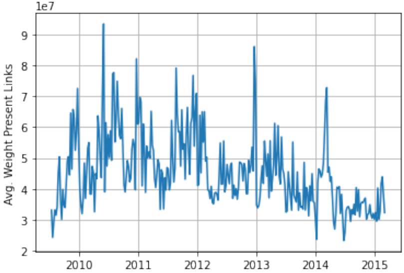
e-MID: Electronic Market for Inter-Bank Deposits

Data Description

Inter-bank Over-Night Loans

- 132 banks, 297 weekly networks between 2009 and 2015
- Sparse networks, density ≤ 0.08
- The network might be affected by shocks that hits a subset of nodes
- Example: Lehman default, LTRO, etc





Graph/Networks Notation

Binary Graph

- Pair $(\mathcal{V}, \mathcal{E})$
  - $\mathcal{V}$ set of N nodes
  - $\mathcal{E}$ set of pairs of nodes: M links
- Adjacency Matrix:

$$A_{ij} = \begin{cases} 1 & \textit{if } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}$$

- Can be **directed**, or **undirected**

Weighted

- Weighted Adjacency Matrix $Y_{ij} \in \mathcal{R}$ (for us only $\mathcal{R}^+$)
- $A_{ij} = \mathbb{1}_{Y_{ij}>0}$, indicator function $\mathbb{1}$

Graph/Networks Notation

Binary Graph

- Pair $(\mathcal{V}, \mathcal{E})$
  - $\mathcal{V}$ set of N nodes
  - $\mathcal{E}$ set of pairs of nodes: M links
- Adjacency Matrix:

$$A_{ij} = \begin{cases} 1 & \textit{if } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}$$

- Can be **directed**, or **undirected**

Weighted

- Weighted Adjacency Matrix $Y_{ij} \in \mathcal{R}$ (for us only $\mathcal{R}^+$)
- $A_{ij} = \mathbb{1}_{Y_{ij}>0}$, indicator function $\mathbb{1}$

Probabilistic models of networks

- A network probabilistic model is defined by a set $\mathcal{X}$ of graphs and a probability mass function

$$\mathbb{P}_\theta\left[\mathbf{A}\right] : \mathcal{X} \to [0,1]\,,$$

indexed by a vector of model parameters θ and such that $\sum_{\mathbf{A}\in\mathcal{X}} \mathbb{P}_\theta\left(\mathbf{A}\right) = 1$.

- For an arbitrary (regular enough) network function $F : \mathcal{X} \to \mathbb{R}$ defined on the set $\mathcal{X}$, the expected value of F on the ensemble $\mathcal{X}$ is defined as

$$\mathbb{E}_{\mathcal{X}}\left[F\right] = \sum_{\mathbf{X}\in\mathcal{X}} F\left(\mathbf{A}\right)\,\mathbb{P}_\theta\left[\mathbf{A}\right].$$

Probabilistic (or statistical) models of networks are useful for

- Assessing the significance of observed network structures
- Construct null models
- Hypothesis testing
- Identification of mechanisms of network formation
- Reconstruct networks from partial information
- Produce conditional scenario generations
- ....

Identifying core-periphery

The core is a subset of nodes that are maximally connected to other members of the core, while the periphery is the complementary subset made of nodes with no reciprocal connections, but only with the core.

A standard approach: consider a partition in two blocks $\mathcal{C}$ and an error matrix

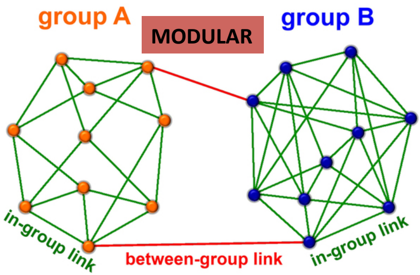
$$E(\mathcal{C}) = \begin{pmatrix} E_{CC} & E_{CP} \\ E_{PC} & E_{PP} \end{pmatrix} = \begin{pmatrix} N_c(N_c-1) - \sum_{i,j\in\mathcal{C}} A_{ij} & 0 \\ 0 & \sum_{i,j\notin\mathcal{C}} A_{ij} \end{pmatrix}$$

The optimal partition is given by

$$\mathcal{C}^* = \arg\min_{\mathcal{C}} \frac{E_{CC}+E_{PP}}{M}$$

This procedure always produces a core but does not say if it is significant!

What is the two-block large scale organization of a network ?



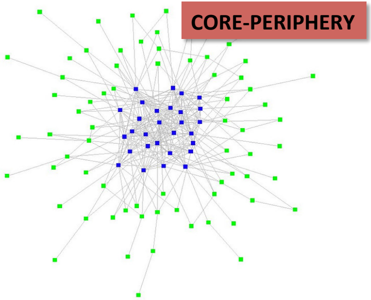
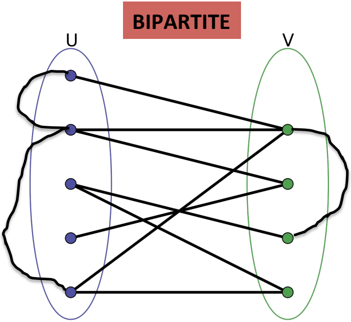
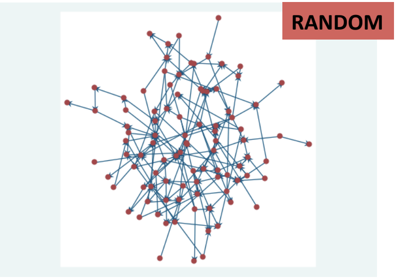


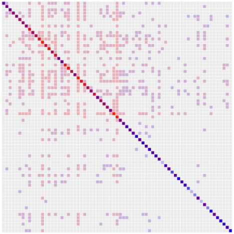
Figure 4 – Core/Periphery Network





Why is the inference a complicated problem?

Adjacency matrix of the e-MID interbank market. **The two matrices differ only for the sorting of rows and columns**. Color scale of links is proportional to nodes degree.



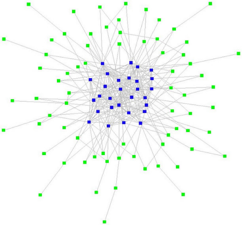
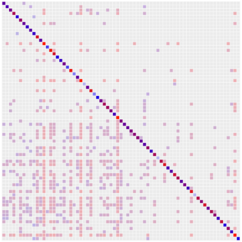
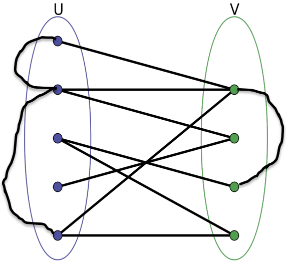


Figure 4 – Core/Periphery Network





Stochastic block model

- Generalizes Erdös-Renyi to many blocks.
- Nodes are divided in $m(=2)$ groups. Node i belongs to group $g_i \in \{1, .., m\}$.
- The probability that node i and j are linked is $p_{g_i g_j}$, independent from the other links
- The $m \times m$ matrix p is the **affinity matrix**
- We consider here the directed and weighted SBM with weights $Y_{ij} \in \mathbb{N}$

$$\mathcal{P}(Y|\boldsymbol{g}, p) = \prod_{(i,j)}^{N} \frac{p_{g_i g_j}^{Y_{ij}}}{Y_{ij}!} \exp\left(-p_{g_i g_j}\right). \tag{1}$$

- Inference
    - Message passing (aka belief propagation) based on cavity method
    - Parameter inference by minimizing the microcanonical entropy via Markov Chain Monte Carlo (Peixoto 2014)

Two block inference

Affinity matrix

$$p = \begin{pmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{pmatrix} \tag{2}$$

Depending on the ranking of the affinity matrix elements p_{ij} we have

- If $p_{11} > p_{22} > p_{12}$ (or $p_{22} > p_{11} > p_{12}$) the network has a **modular structure**,
- If $p_{11} > p_{12} > p_{22}$ (or $p_{22} > p_{12} > p_{11}$) the network has a **core-periphery structure**.
- If $p_{12} > p_{11} > p_{22}$ (or $p_{12} > p_{22} > p_{11}$) the network has a **bipartite structure**.

Application: the e-MID interbank market

Data: Italian electronic market for interbank deposits (e-MID). Overnight transactions from July 2009 to December 2014.

Weighted	Day				Week				Month			
Year	B	C	M	R	B	C	M	R	B	C	M	R
2010	53	0	6	41	92	0	0	8	100	0	0	0
2011	55	0	5	40	90	0	0	10	100	0	0	0
2012	41	0	9	50	55	0	0	45	75	0	0	25
2013	39	0	5	56	36	0	2	62	58	0	0	42
2014	52	0	4	44	80	0	0	20	100	0	0	0

Table: Percentages of inferred structures in the e-MID interbank market at different levels of aggregation in the 5 investigated years. The structures are bipartite (B), core-periphery (C), modular (M), and no structure (R).

- At all time scales the interbank market is bipartite (rather than core-periphery as suggested in the literature)
- In 2012-2013 the interbank has very often a random structure

Hidden node-specific parameter models

- Models with hidden node-specific parameters are flexible network formation mechanisms, able to generate a wide range of structural features.
- Each node i is associated with a (vector of) parameter(s) θ_i,
- Links are formed between nodes with a probability

$$p_{ij} = f(\theta_i, \theta_j) \tag{3}$$

- In Stochastic Block Model the θs are discrete variables identifying the group (block) the node belong to.
- In **fitness model** the θs are continuous variables associated with the degree (Caldarelli et al 2002)
- Mathematical connection to Lagrange multipliers of Maximum Entropy ensemble with constraints on degrees (configuration model).

Hidden node-specific parameter models

- Models with hidden node-specific parameters are flexible network formation mechanisms, able to generate a wide range of structural features.
- Each node i is associated with a (vector of) parameter(s) θ_i,
- Links are formed between nodes with a probability

$$p_{ij} = f(\theta_i, \theta_j) \tag{3}$$

- In Stochastic Block Model the θs are discrete variables identifying the group (block) the node belong to.
- In **fitness model** the θs are continuous variables associated with the degree (Caldarelli et al 2002)
- Mathematical connection to Lagrange multipliers of Maximum Entropy ensemble with constraints on degrees (configuration model).

Hidden node-specific parameter models

- Models with hidden node-specific parameters are flexible network formation mechanisms, able to generate a wide range of structural features.
- Each node i is associated with a (vector of) parameter(s) θ_i,
- Links are formed between nodes with a probability

$$p_{ij} = f(\theta_i, \theta_j) \tag{3}$$

- In Stochastic Block Model the θs are discrete variables identifying the group (block) the node belong to.
- In **fitness model** the θs are continuous variables associated with the degree (Caldarelli et al 2002)
- Mathematical connection to Lagrange multipliers of Maximum Entropy ensemble with constraints on degrees (configuration model).

Hidden node-specific parameter models

- Models with hidden node-specific parameters are flexible network formation mechanisms, able to generate a wide range of structural features.
- Each node i is associated with a (vector of) parameter(s) θ_i,
- Links are formed between nodes with a probability

$$p_{ij} = f(\theta_i, \theta_j) \tag{3}$$

- In Stochastic Block Model the θs are discrete variables identifying the group (block) the node belong to.
- In **fitness model** the θs are continuous variables associated with the degree (Caldarelli et al 2002)
- Mathematical connection to Lagrange multipliers of Maximum Entropy ensemble with constraints on degrees (configuration model).

Hidden node-specific parameter models

- Models with hidden node-specific parameters are flexible network formation mechanisms, able to generate a wide range of structural features.
- Each node i is associated with a (vector of) parameter(s) θ_i,
- Links are formed between nodes with a probability

$$p_{ij} = f(\theta_i, \theta_j) \tag{3}$$

- In Stochastic Block Model the θs are discrete variables identifying the group (block) the node belong to.
- In **fitness model** the θs are continuous variables associated with the degree (Caldarelli et al 2002)
- Mathematical connection to Lagrange multipliers of Maximum Entropy ensemble with constraints on degrees (configuration model).

Hidden node-specific parameter models

- Models with hidden node-specific parameters are flexible network formation mechanisms, able to generate a wide range of structural features.
- Each node i is associated with a (vector of) parameter(s) θ_i,
- Links are formed between nodes with a probability

$$p_{ij} = f(\theta_i, \theta_j) \tag{3}$$

- In Stochastic Block Model the θs are discrete variables identifying the group (block) the node belong to.
- In **fitness model** the θs are continuous variables associated with the degree (Caldarelli et al 2002)
- Mathematical connection to Lagrange multipliers of Maximum Entropy ensemble with constraints on degrees (configuration model).

Hidden node-specific parameter models

- Models with hidden node-specific parameters are flexible network formation mechanisms, able to generate a wide range of structural features.
- Each node i is associated with a (vector of) parameter(s) θ_i,
- Links are formed between nodes with a probability

$$p_{ij} = f(\theta_i, \theta_j) \tag{3}$$

- In Stochastic Block Model the θs are discrete variables identifying the group (block) the node belong to.
- In **fitness model** the θs are continuous variables associated with the degree (Caldarelli et al 2002)
- Mathematical connection to Lagrange multipliers of Maximum Entropy ensemble with constraints on degrees (configuration model).

Maximum Entropy Network Ensembles

Consider the following Maximum Entropy problem for the probability P of having a (static) network i

$$\max_{\{P\}} \left(-\sum_i P_i \ln P_i \right) \qquad s.t. \quad \sum_i P_i = 1, \qquad \sum_i P_i f_i^s = \bar{f}^s$$

where f_i^s $(s = 1, .., S)$ are certain network metrics of network i.

The solution is

$$P_j = \frac{\exp\left(\sum_s \theta_s f_j^s\right)}{\sum_j \exp\left(\sum_s \theta_s f_j^s\right)}$$

The Lagrange multipliers $\theta_1, ..., \theta_S$ are determined (in principle) by imposing the constraints $\sum_i P_i f_i^s = \bar{f}^s$.

Exponential Random Graph Models (ERGM)

Definition

- Set of network metrics $\{f_s(\mathbf{A})\}_{s=1}^{S}$
- Probability mass function (PMF)

$$P(\mathbf{A}) = \frac{e^{\sum_s \theta_s f_s(\mathbf{A})}}{\mathcal{K}(\theta)}$$

- For both directed and undirected binary networks

Examples of Metrics

- Number of links $f(\mathbf{A}) = \sum_{i,j} A_{ij} \rightarrow$ Erdos-Renyi model
- Degree sequence $f_i(\mathbf{A}) = \sum_{j \neq j} A_{ij} \rightarrow$ **fitness, configuration, beta model**.
- Number of triangles

Exponential Random Graph models naturally emerges from Maximum Entropy principle

Exponential Random Graph Models (ERGM)

Definition

- Set of network metrics $\{f_s(\mathbf{A})\}_{s=1}^{S}$
- Probability mass function (PMF)

$$P(\mathbf{A}) = \frac{e^{\sum_s \theta_s f_s(\mathbf{A})}}{\mathcal{K}(\theta)}$$

- For both directed and undirected binary networks

Examples of Metrics

- Number of links $f(\mathbf{A}) = \sum_{i,j} A_{ij} \rightarrow$ Erdos-Renyi model
- Degree sequence $f_i(\mathbf{A}) = \sum_{j \neq j} A_{ij} \rightarrow$ **fitness, configuration, beta model**.
- Number of triangles

Exponential Random Graph models naturally emerges from Maximum Entropy principle

Fitness Model for Binary Networks

A.K.A. Configuration Model or Beta Model

Choosing $f_i(\mathbf{A}) = \sum_{j \neq j} A_{ij}$ it is possible to show that:

Definition

- $2N$ parameters for nodes' heterogeneity:
  - $\overleftarrow{\theta}_i$ in-fitness - tendency to form incoming connections
  - $\overrightarrow{\theta}_i$ out-fitness - tendency to form outgoing connections

$$P(A_{ij} = 1) = \frac{1}{1 + e^{-\left(\overleftarrow{\theta}_i + \overrightarrow{\theta}_j\right)}}$$

- ERGM with 2 statistics per node: in and out degree

The Likelihood of the Directed Fitness Model

Log-Likelihood

The Log-likelihood function of $\theta \equiv \begin{pmatrix} \overleftarrow{\theta} \\ \overrightarrow{\theta} \end{pmatrix}$:

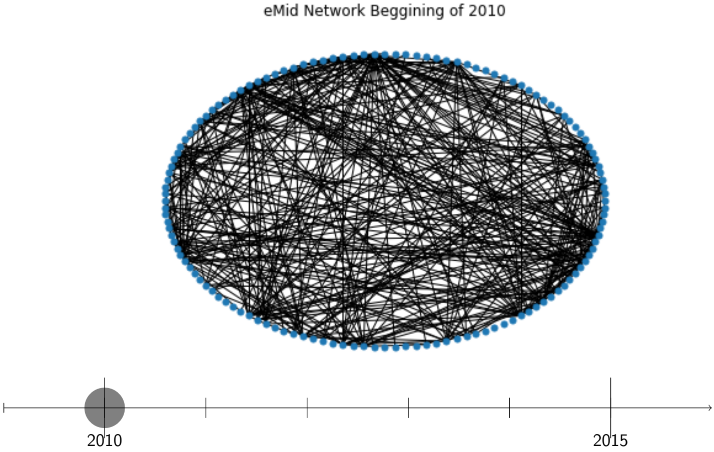
$$l(\theta) = \log P(\mathbf{A}|\theta) = \sum_{ij} \left\{ A_{ij} \left(\overleftarrow{\theta}_i + \overrightarrow{\theta}_j \right) - \log \left(1 + e^{\overleftarrow{\theta}_i + \overrightarrow{\theta}_j} \right) \right\}$$

Comments

- MLE estimation
- Large N single network asymptotic theory (valid in dense regime)

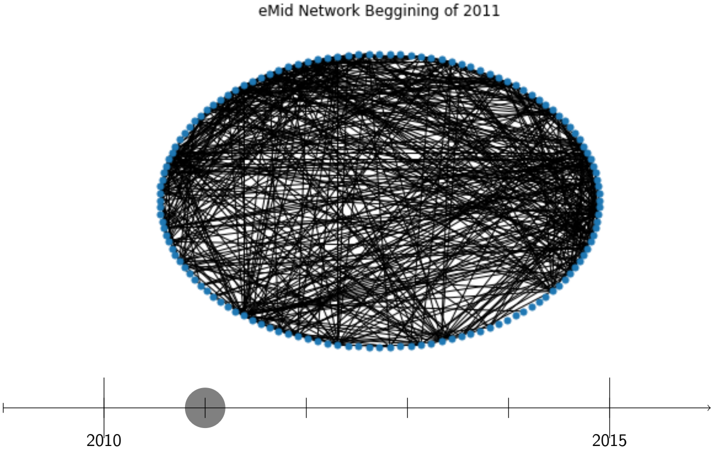
Temporal Networks in Discrete Time

Networks Evolve in Time



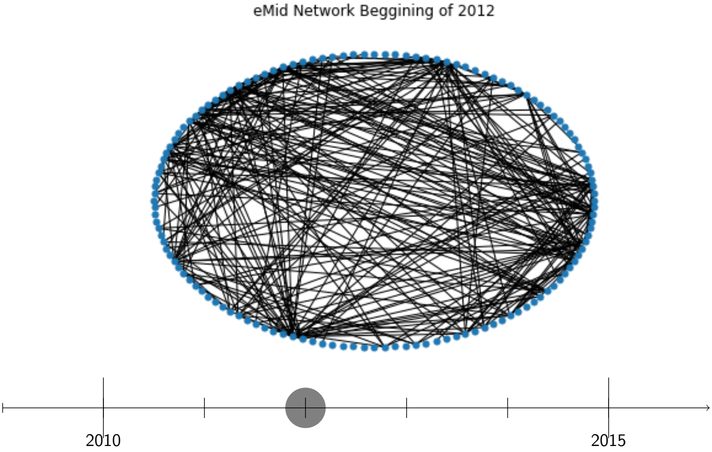
Temporal Networks in Discrete Time

Networks Evolve in Time



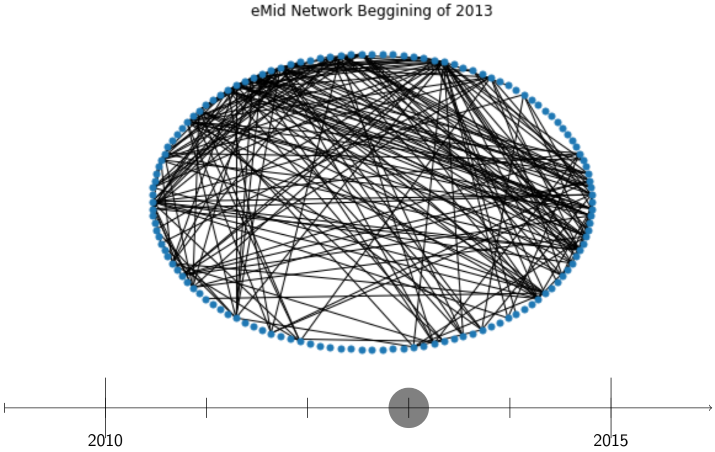
Temporal Networks in Discrete Time

Networks Evolve in Time



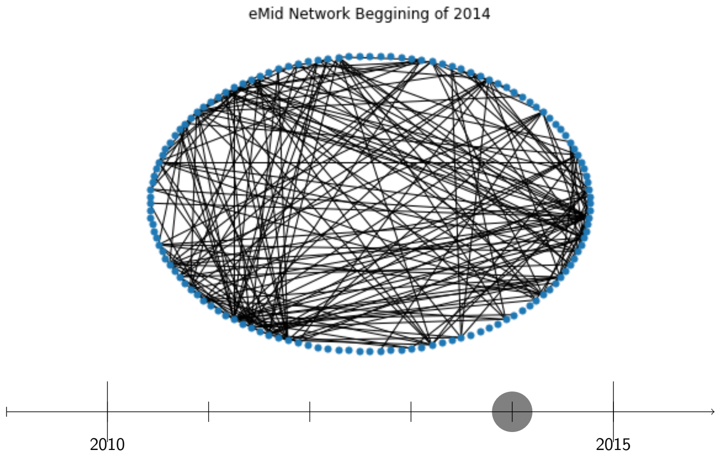
Temporal Networks in Discrete Time

Networks Evolve in Time



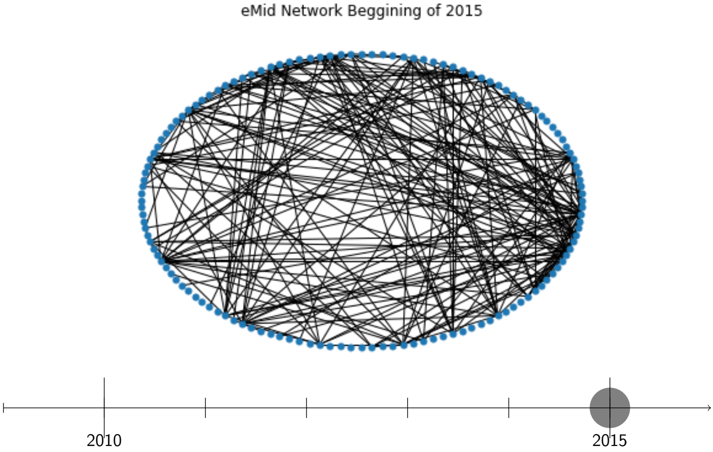
Temporal Networks in Discrete Time

Networks Evolve in Time



Temporal Networks in Discrete Time

Networks Evolve in Time

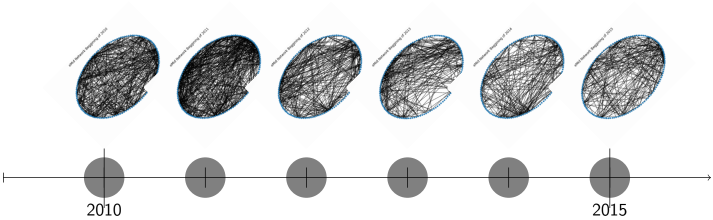


Temporal Networks in Discrete Time

Networks Evolve in Time

Discrete Time Description

$\{\mathbf{Y}^{(t)}\}^{t=0,1,\ldots,T}$



Dynamical networks

Idea: The temporal networks is described by the dynamics of the latent variables (fitness) $\theta_i \to \theta_i^{(t)}$.

Different approaches:

- Each θ_i evolves independently as an AR(1) model (Mazzarisi et al., EJOR 2020)

$$\theta_i^{(t)} = a + b\, \theta_i^{(t-1)} + \epsilon_i^{(t)}$$

- The θs evolve as a VAR model (Rizzini and Lillo, 2024)

$$\vec{\theta}^{(t)} = \vec{\mu} + \mathbf{B}\vec{\theta}^{(t-1)} + \epsilon^{(t)}$$

  allowing for interactions and leading to scenario generation (Impulse Response Analysis)

- Allow for generic dynamics of fitnesses (and networks) via Score Driven models which must be **filtered** from the data (Di Gangi, Bormetti, Lillo, 2022).

What drives network dynamics?

Many networks are inherently dynamic as links are created and destroyed through time.

- *Preferential* relations between nodes tend to preserve past links (If we were friends yesterday we will be friend today).
- *Node specific* properties can drive the evolution of the network topology (Two social persons are more likely to be friend).

- How the node characteristic and link persistence shape a network and how to account for the two linking mechanisms in a statistical model of temporal networks?
- A proper modeling of network dynamics allows performing short term link prediction.
- In general, the dynamics will be a combination of the two types of mechanisms.

What drives network dynamics?

Many networks are inherently dynamic as links are created and destroyed through time.

- *Preferential* relations between nodes tend to preserve past links (If we were friends yesterday we will be friend today).
- *Node specific* properties can drive the evolution of the network topology (Two social persons are more likely to be friend).

- How the node characteristic and link persistence shape a network and how to account for the two linking mechanisms in a statistical model of temporal networks?
- A proper modeling of network dynamics allows performing short term link prediction.
- In general, the dynamics will be a combination of the two types of mechanisms.

What drives network dynamics?

Many networks are inherently dynamic as links are created and destroyed through time.

- *Preferential* relations between nodes tend to preserve past links (If we were friends yesterday we will be friend today).
- *Node specific* properties can drive the evolution of the network topology (Two social persons are more likely to be friend).

- How the node characteristic and link persistence shape a network and how to account for the two linking mechanisms in a statistical model of temporal networks?
- A proper modeling of network dynamics allows performing short term link prediction.
- In general, the dynamics will be a combination of the two types of mechanisms.

What drives network dynamics?

Many networks are inherently dynamic as links are created and destroyed through time.

- *Preferential* relations between nodes tend to preserve past links (If we were friends yesterday we will be friend today).
- *Node specific* properties can drive the evolution of the network topology (Two social persons are more likely to be friend).

- How the node characteristic and link persistence shape a network and how to account for the two linking mechanisms in a statistical model of temporal networks?
- A proper modeling of network dynamics allows performing short term link prediction.
- In general, the dynamics will be a combination of the two types of mechanisms.

What drives network dynamics?

Many networks are inherently dynamic as links are created and destroyed through time.

- *Preferential* relations between nodes tend to preserve past links (If we were friends yesterday we will be friend today).
- *Node specific* properties can drive the evolution of the network topology (Two social persons are more likely to be friend).

- How the node characteristic and link persistence shape a network and how to account for the two linking mechanisms in a statistical model of temporal networks?
- A proper modeling of network dynamics allows performing short term link prediction.
- In general, the dynamics will be a combination of the two types of mechanisms.

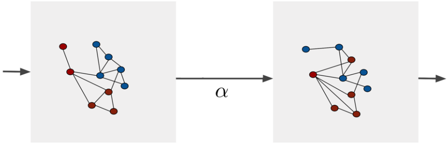
What drives network dynamics?

Many networks are inherently dynamic as links are created and destroyed through time.

- *Preferential* relations between nodes tend to preserve past links (If we were friends yesterday we will be friend today).
- *Node specific* properties can drive the evolution of the network topology (Two social persons are more likely to be friend).

- How the node characteristic and link persistence shape a network and how to account for the two linking mechanisms in a statistical model of temporal networks?
- A proper modeling of network dynamics allows performing short term link prediction.
- In general, the dynamics will be a combination of the two types of mechanisms.

Modeling link persistence

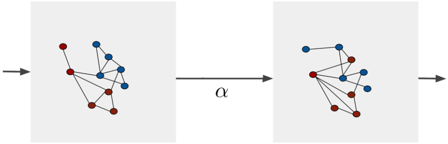


- We model the tendency of a link that does (or does not) exist at time $t-1$ to continue existing (or not existing) at time t.
- Discrete AutoRegressive DAR(1) model

$$P(\mathbf{A}^t|\mathbf{A}^{t-1},\alpha,\chi) = \prod_{i\neq j} \underbrace{\alpha_{ij}\delta_{A^t_{ij}A^{t-1}_{ij}}}_{\substack{\text{Copying the last observation} \\ \text{with probability } \alpha_{ij}}} + \underbrace{(1-\alpha_{ij})\chi_{ij}^{A^t_{ij}}(1-\chi_{ij})^{1-A^t_{ij}}}_{\substack{\text{Bernoulli trial with probability } 1-\alpha_{ij}; \\ \text{The link probability is } \chi_{ij}}}$$

- One independent process per link.
- The larger is α_{ij}, the more persistent is the link between i and j.

Modeling link persistence

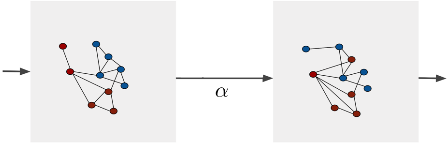


- We model the tendency of a link that does (or does not) exist at time $t-1$ to continue existing (or not existing) at time t.
- Discrete AutoRegressive DAR(1) model

$$P(\mathbf{A}^t|\mathbf{A}^{t-1},\alpha,\chi) = \prod_{i\neq j} \underbrace{\alpha_{ij}\delta_{A^t_{ij}A^{t-1}_{ij}}}_{\substack{\text{Copying the last observation}\\ \text{with probability } \alpha_{ij}}} + \underbrace{(1-\alpha_{ij})\chi_{ij}^{A^t_{ij}}(1-\chi_{ij})^{1-A^t_{ij}}}_{\substack{\text{Bernoulli trial with probability } 1-\alpha_{ij};\\ \text{The link probability is } \chi_{ij}}}$$

- One independent process per link.
- The larger is α_{ij}, the more persistent is the link between i and j.

Modeling link persistence

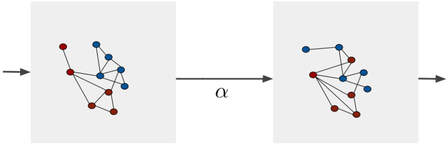


- We model the tendency of a link that does (or does not) exist at time $t-1$ to continue existing (or not existing) at time t.
- Discrete AutoRegressive DAR(1) model

$$P(\mathbf{A}^t|\mathbf{A}^{t-1},\alpha,\chi) = \prod_{i\neq j} \underbrace{\alpha_{ij}\delta_{A^t_{ij}A^{t-1}_{ij}}}_{\text{Copying the last observation with probability } \alpha_{ij}} + \underbrace{(1-\alpha_{ij})\chi_{ij}^{A^t_{ij}}(1-\chi_{ij})^{1-A^t_{ij}}}_{\text{Bernoulli trial with probability } 1-\alpha_{ij}\text{; The link probability is } \chi_{ij}}$$

- One independent process per link.
- The larger is α_{ij}, the more persistent is the link between i and j.

Modeling link persistence

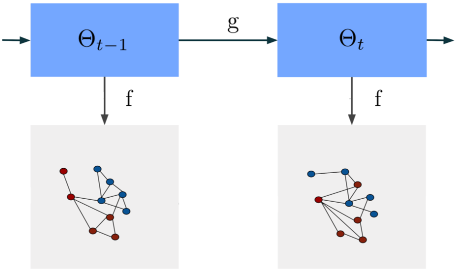


- We model the tendency of a link that does (or does not) exist at time $t-1$ to continue existing (or not existing) at time t.
- Discrete AutoRegressive DAR(1) model

$$P(\mathbf{A}^t|\mathbf{A}^{t-1},\alpha,\chi) = \prod_{i\neq j} \underbrace{\alpha_{ij}\delta_{A^t_{ij}A^{t-1}_{ij}}}_{\text{Copying the last observation with probability } \alpha_{ij}} + \underbrace{(1-\alpha_{ij})\chi_{ij}^{A^t_{ij}}(1-\chi_{ij})^{1-A^t_{ij}}}_{\text{Bernoulli trial with probability } 1-\alpha_{ij}\text{; The link probability is } \chi_{ij}}$$

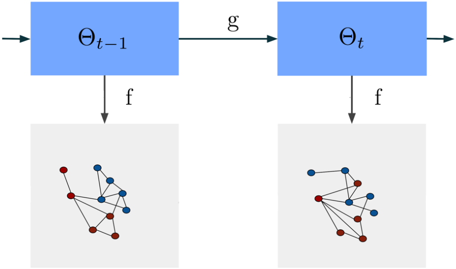
- One independent process per link.
- The larger is α_{ij}, the more persistent is the link between i and j.

Dynamics of node specific variables



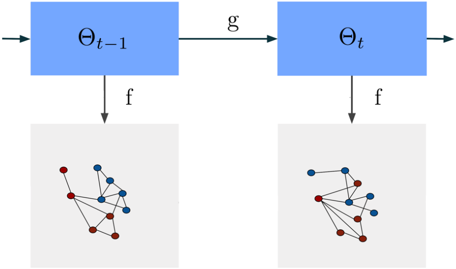
- One **latent** (dynamic) parameter θ_i^t per i. $\Theta^t = \{\theta_i^t\}_{i=1,..,N}$.
- The probability of a link between two nodes depends on their latent parameter through a link function.
- The parameter θ_i^t evolves in time and has temporal persistence.

Dynamics of node specific variables



- One **latent** (dynamic) parameter θ_i^t per i. $\Theta^t = \{\theta_i^t\}_{i=1,..,N}$.
- The probability of a link between two nodes depends on their latent parameter through a link function.
- The parameter θ_i^t evolves in time and has temporal persistence.

Dynamics of node specific variables



- One **latent** (dynamic) parameter θ_i^t per i. $\Theta^t = \{\theta_i^t\}_{i=1,..,N}$.
- The probability of a link between two nodes depends on their latent parameter through a link function.
- The parameter θ_i^t evolves in time and has temporal persistence.

Temporal fitness

- Each node i is characterized by a quantity θ_i^t, i.e. the **node fitness**. We assume that it follows an AR(1) process,

$$\theta_i^t = \phi_{0,i} + \phi_{1,i}\theta_i^{t-1} + \epsilon_i^t$$

$\phi_{0,i} \in \mathbb{R}$, $|\phi_{1,i}| < 1$ and $\epsilon_i^t \sim NID(0, \sigma_i)$ with $\sigma_i > 0$.

- We define the link probability at time t as:

$$P(A_{ij}^t = 1 | \theta_i^t, \theta_j^t) = \frac{1}{1 + e^{-(\theta_i^t + \theta_j^t)}}$$

- The larger θ_i^t is, the larger is the probability for all links incident to node i at time t (i.e. the in-degree).

Temporal fitness

- Each node i is characterized by a quantity θ_i^t, i.e. the **node fitness**. We assume that it follows an AR(1) process,

$$\theta_i^t = \phi_{0,i} + \phi_{1,i}\theta_i^{t-1} + \epsilon_i^t$$

$\phi_{0,i} \in \mathbb{R}$, $|\phi_{1,i}| < 1$ and $\epsilon_i^t \sim NID(0, \sigma_i)$ with $\sigma_i > 0$.

- We define the link probability at time t as:

$$P(A_{ij}^t = 1|\theta_i^t, \theta_j^t) = \frac{1}{1 + e^{-(\theta_i^t + \theta_j^t)}}$$

- The larger θ_i^t is, the larger is the probability for all links incident to node i at time t (i.e. the in-degree).

Temporal fitness

- Each node i is characterized by a quantity θ_i^t, i.e. the **node fitness**. We assume that it follows an AR(1) process,

$$\theta_i^t = \phi_{0,i} + \phi_{1,i}\theta_i^{t-1} + \epsilon_i^t$$

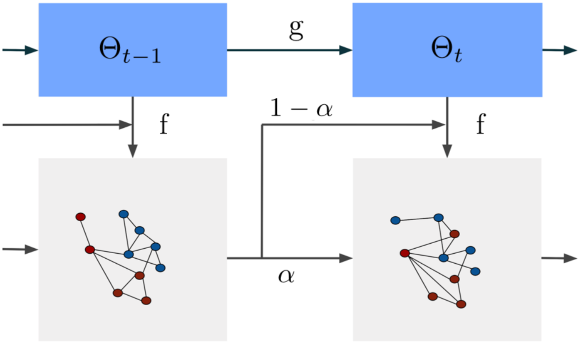
$\phi_{0,i} \in \mathbb{R}$, $|\phi_{1,i}| < 1$ and $\epsilon_i^t \sim NID(0, \sigma_i)$ with $\sigma_i > 0$.

- We define the link probability at time t as:

$$P(A_{ij}^t = 1 | \theta_i^t, \theta_j^t) = \frac{1}{1 + e^{-(\theta_i^t + \theta_j^t)}}$$

- The larger θ_i^t is, the larger is the probability for all links incident to node i at time t (i.e. the in-degree).

Dynamic fitness + link persistence



We combine the hidden dynamics of (**fitness**) with the mechanism of copying from the past (**link persistence**).

DAR-TGRG: fitness dynamics + link persistence

Mixture of the two linking mechanisms:

$$\begin{cases} (\theta_i^t|\theta_i^{t-1},\Phi_i) & \sim \mathcal{N}(\theta_i^t - \phi_{0,i} - \phi_{1,i}\theta_i^{t-1}, \sigma_i) \ \ \forall i = 1,...,N \\ P(\mathbf{A}^t|\mathbf{A}^{t-1},\Theta^t,\alpha) & = \prod_{i,j>i}\left(\alpha_{ij}\delta_{A_{ij}^t A_{ij}^{t-1}} + (1-\alpha_{ij})\frac{e^{A_{ij}^t(\theta_i^t+\theta_j^t)}}{1+e^{(\theta_i^t+\theta_j^t)}}\right) \end{cases}$$

- $\phi_{0,i} \in \mathbb{R}$, $|\phi_{1,i}| < 1$, $\sigma_i > 0$ and $\alpha_{ij} \in [0,1]$ $\forall i,j = 1,...,N$.
- Copying from the past with probability α_{ij} and a time evolving marginal described by the dynamic fitness model with probability $1-\alpha_{ij}$.
- α_{ij} ***disentangles*** the two mechanisms for each link.
- $3 \times N$ parameters Φ, $\binom{N}{2}$ parameters α.
- Expectation-Maximization (EM) algorithm for model estimation since likelihood is known.
- The estimation of parameters via EM is unbiased, while the Single Snapshot Inference is biased.

DAR-TGRG: fitness dynamics + link persistence

Mixture of the two linking mechanisms:

$$\begin{cases} (\theta_i^t | \theta_i^{t-1}, \Phi_i) & \sim \mathcal{N}(\theta_i^t - \phi_{0,i} - \phi_{1,i}\theta_i^{t-1}, \sigma_i) \;\; \forall i = 1, ..., N \\ P(\mathbf{A}^t | \mathbf{A}^{t-1}, \Theta^t, \alpha) & = \prod_{i,j>i} \left(\alpha_{ij} \delta_{A_{ij}^t A_{ij}^{t-1}} + (1 - \alpha_{ij}) \frac{e^{A_{ij}^t(\theta_i^t + \theta_j^t)}}{1 + e^{(\theta_i^t + \theta_j^t)}} \right) \end{cases}$$

- $\phi_{0,i} \in \mathbb{R}$, $|\phi_{1,i}| < 1$, $\sigma_i > 0$ and $\alpha_{ij} \in [0,1]$ $\forall i, j = 1, ..., N$.
- Copying from the past with probability α_{ij} and a time evolving marginal described by the dynamic fitness model with probability $1 - \alpha_{ij}$.
- α_{ij} ***disentangles*** the two mechanisms for each link.
- $3 \times N$ parameters Φ, $\binom{N}{2}$ parameters α.
- Expectation-Maximization (EM) algorithm for model estimation since likelihood is known.
- The estimation of parameters via EM is unbiased, while the Single Snapshot Inference is biased.

DAR-TGRG: fitness dynamics + link persistence

Mixture of the two linking mechanisms:

$$\begin{cases} (\theta_i^t|\theta_i^{t-1},\Phi_i) & \sim \mathcal{N}(\theta_i^t - \phi_{0,i} - \phi_{1,i}\theta_i^{t-1}, \sigma_i) \quad \forall i = 1,\dots,N \\ P(\mathbf{A}^t|\mathbf{A}^{t-1},\Theta^t,\alpha) & = \prod_{i,j>i}\left(\alpha_{ij}\delta_{A_{ij}^t A_{ij}^{t-1}} + (1-\alpha_{ij})\frac{e^{A_{ij}^t(\theta_i^t+\theta_j^t)}}{1+e^{(\theta_i^t+\theta_j^t)}}\right) \end{cases}$$

- $\phi_{0,i} \in \mathbb{R}$, $|\phi_{1,i}| < 1$, $\sigma_i > 0$ and $\alpha_{ij} \in [0,1]$ $\forall i,j = 1,\dots,N$.
- Copying from the past with probability α_{ij} and a time evolving marginal described by the dynamic fitness model with probability $1-\alpha_{ij}$.
- α_{ij} ***disentangles*** the two mechanisms for each link.
- $3 \times N$ parameters Φ, $\binom{N}{2}$ parameters α.
- Expectation-Maximization (EM) algorithm for model estimation since likelihood is known.
- The estimation of parameters via EM is unbiased, while the Single Snapshot Inference is biased.

DAR-TGRG: fitness dynamics + link persistence

Mixture of the two linking mechanisms:

$$\begin{cases} (\theta_i^t|\theta_i^{t-1},\Phi_i) & \sim \mathcal{N}(\theta_i^t-\phi_{0,i}-\phi_{1,i}\theta_i^{t-1},\sigma_i) \;\; \forall i=1,...,N \\ P(\mathbf{A}^t|\mathbf{A}^{t-1},\Theta^t,\alpha) & = \prod_{i,j>i}\left(\alpha_{ij}\delta_{A_{ij}^t A_{ij}^{t-1}}+(1-\alpha_{ij})\frac{e^{A_{ij}^t(\theta_i^t+\theta_j^t)}}{1+e^{(\theta_i^t+\theta_j^t)}}\right) \end{cases}$$

- $\phi_{0,i}\in\mathbb{R}$, $|\phi_{1,i}|<1$, $\sigma_i>0$ and $\alpha_{ij}\in[0,1]$ $\forall i,j=1,...,N$.
- Copying from the past with probability α_{ij} and a time evolving marginal described by the dynamic fitness model with probability $1-\alpha_{ij}$.
- α_{ij} ***disentangles*** the two mechanisms for each link.
- $3\times N$ parameters Φ, $\binom{N}{2}$ parameters α.
- Expectation-Maximization (EM) algorithm for model estimation since likelihood is known.
- The estimation of parameters via EM is unbiased, while the Single Snapshot Inference is biased.

DAR-TGRG: fitness dynamics + link persistence

Mixture of the two linking mechanisms:

$$\begin{cases} (\theta_i^t|\theta_i^{t-1},\Phi_i) & \sim \mathcal{N}(\theta_i^t - \phi_{0,i} - \phi_{1,i}\theta_i^{t-1}, \sigma_i) \;\; \forall i = 1,\ldots,N \\ P(\mathbf{A}^t|\mathbf{A}^{t-1},\Theta^t,\alpha) & = \prod_{i,j>i}\left(\alpha_{ij}\delta_{A_{ij}^t A_{ij}^{t-1}} + (1-\alpha_{ij})\frac{e^{A_{ij}^t(\theta_i^t+\theta_j^t)}}{1+e^{(\theta_i^t+\theta_j^t)}}\right) \end{cases}$$

- $\phi_{0,i} \in \mathbb{R}$, $|\phi_{1,i}| < 1$, $\sigma_i > 0$ and $\alpha_{ij} \in [0,1]$ $\forall i,j = 1,\ldots,N$.
- Copying from the past with probability α_{ij} and a time evolving marginal described by the dynamic fitness model with probability $1-\alpha_{ij}$.
- α_{ij} ***disentangles*** the two mechanisms for each link.
- $3 \times N$ parameters Φ, $\binom{N}{2}$ parameters α.
- Expectation-Maximization (EM) algorithm for model estimation since likelihood is known.
- The estimation of parameters via EM is unbiased, while the Single Snapshot Inference is biased.

DAR-TGRG: fitness dynamics + link persistence

Mixture of the two linking mechanisms:

$$\begin{cases} (\theta_i^t|\theta_i^{t-1},\Phi_i) & \sim \mathcal{N}(\theta_i^t - \phi_{0,i} - \phi_{1,i}\theta_i^{t-1}, \sigma_i) \ \ \forall i = 1, ..., N \\ P(\mathbf{A}^t|\mathbf{A}^{t-1},\Theta^t,\alpha) & = \prod_{i,j>i} \left(\alpha_{ij}\delta_{A_{ij}^t A_{ij}^{t-1}} + (1-\alpha_{ij}) \frac{e^{A_{ij}^t(\theta_i^t+\theta_j^t)}}{1+e^{(\theta_i^t+\theta_j^t)}} \right) \end{cases}$$

- $\phi_{0,i} \in \mathbb{R}$, $|\phi_{1,i}| < 1$, $\sigma_i > 0$ and $\alpha_{ij} \in [0,1]$ $\ \forall i,j = 1, ..., N$.
- Copying from the past with probability α_{ij} and a time evolving marginal described by the dynamic fitness model with probability $1-\alpha_{ij}$.
- α_{ij} ***disentangles*** the two mechanisms for each link.
- $3 \times N$ parameters Φ, $\binom{N}{2}$ parameters α.
- Expectation-Maximization (EM) algorithm for model estimation since likelihood is known.
- The estimation of parameters via EM is unbiased, while the Single Snapshot Inference is biased.

Goodness of fit of the EM algorithm

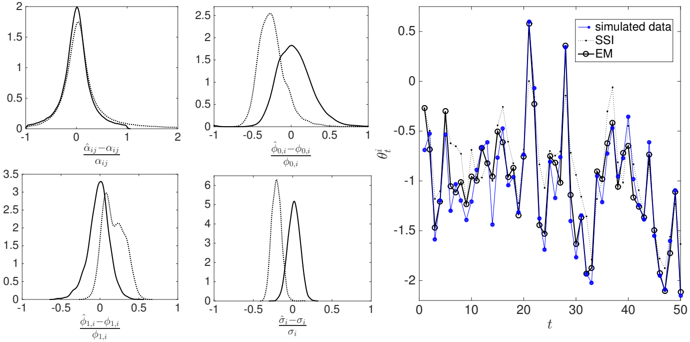
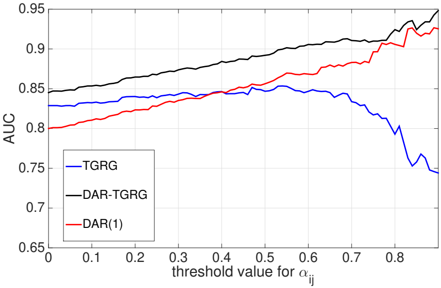


Figure: Left panels: distribution of relative errors in the estimation of α_{ij}, $\phi_{0,i}$, $\phi_{1,i}$ and σ_i. Right panel: latent dynamics for a generic θ_i^t compared with the inferred one according to EM.

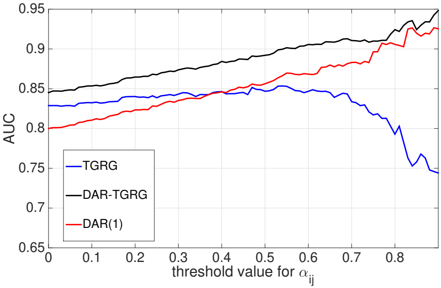
The estimation of parameters via EM is unbiased, while the Single Snapshot Inference (SSI) is biased

Out-of-sample link prediction for e-MID



- We derive closed form expression for the p-step ahead forecast for the presence of a link.
- Out of sample analysis for the e-MID interbank market.
- DAR-TGRG outperforms TGRG and DAR(1) network models in terms of Area Under Curve (AUC).
- On average, network topology is more important than link stability for link prediction in e-MID.

Out-of-sample link prediction for e-MID



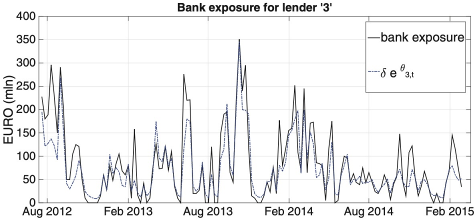
- We derive closed form expression for the p-step ahead forecast for the presence of a link.
- Out of sample analysis for the e-MID interbank market.
- DAR-TGRG outperforms TGRG and DAR(1) network models in terms of Area Under Curve (AUC).
- On average, network topology is more important than link stability for link prediction in e-MID.

Modeling shock propagation and resilience in directed temporal networks

- How does a networked system reacts to an exogenous shock and how does it relax back to the normal state?
- We study the resilience of a temporal network by proposing a modification of the impulse response analysis
- Rather than considering a shock on a network observable (e.g. a link disappears), we study **shocks on nodes' characteristics**
- **Idea**:
  - the node i propensity to create link is contained in the hidden variable (fitnesses) θ_i;
  - propose a model for the evolution of the θs;
  - study how the θs evolve as a consequence of an idiosyncratic or systemic shock on them.

Economic interpretations

- In a dynamic fitness model calibrated on the unweighted interbank network Mazzarisi et al (EJOR 2020) finds that $\exp(\theta_i)$ correlates closely with the in- or out-strength (not used in the calibration): $\theta_i \sim \log(bank\ \ exposure)$



- In zero-inflated gravity models of the World Trade Web, the probability of a link is

$$P(A_{ij} = 1) = \frac{\overline{GDP_i} \cdot \overline{GDP_j}}{1 + \overline{GDP_i} \cdot \overline{GDP_j}}$$

where $\overline{GDP_i}$ is the (normalized) GDP of country i. Thus $\theta_i = \log \overline{GDP_i}$

- A shock Δ on θ_i corresponds to a log (or percentage) change of Δ in bank exposure or GDP.

The model for fitness dynamics

- Vector $\vec{\theta}_t$ evolves according to a vector autoregressive model of order 1 (VAR(1)):

$$\vec{\theta}_t = \vec{\mu} + \mathbf{B}\vec{\theta}_{t-1} + \boldsymbol{w}_t, \tag{4}$$

  - $\vec{\mu}$ controls the mean value of the latent variables, thus the temporal average degree of each node
  - **B** plays the role of modeling the lagged interactions between latent variables:
    - the diagonal elements of **B** describe the temporal autocorrelations of the θs
    - the off-diagonal elements describe how the latent variable of a node at time t affects the latent variable of another node at time $t + 1$
  - $\boldsymbol{w}_t \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I})$ is the vector of white noise terms.
- Mazzarisi et al (EJOR 2020) considered a model with diagonal **B** (plus a mechanism for adding memory to each link)

Impulse Response Function

- We assume that an exogenous shock on the latent variables $\vec{\boldsymbol{\Delta\theta}}$ happens at time τ
- We define the (standard) impulse response function of the latent vector as

$$\boldsymbol{IRF}^{\theta}(t;\vec{\boldsymbol{\Delta\theta}}) = \mathbb{E}[\vec{\boldsymbol{\theta}}_{\tau+t}|\vec{\boldsymbol{\theta}}_{\tau} + \vec{\boldsymbol{\Delta\theta}}, \vec{\boldsymbol{\theta}}_{\tau-1}\ldots] - \mathbb{E}[\vec{\boldsymbol{\theta}}_{\tau+t}|\vec{\boldsymbol{\theta}}_{\tau}, \vec{\boldsymbol{\theta}}_{\tau-1}\ldots] \quad (5)$$

where the expectation is taken over the realization of the noise $\boldsymbol{w}_t$

- The IRF of the latent variables is

$$\boldsymbol{IRF}^{\theta}(t,\vec{\boldsymbol{\Delta\theta}}) = \mathbf{B}^{t}\vec{\boldsymbol{\Delta\theta}} \quad (6)$$

  - $\boldsymbol{IRF}^{\theta}(t,\vec{\boldsymbol{\Delta\theta}})$ is linear in $\vec{\boldsymbol{\Delta\theta}}$
  - $\boldsymbol{IRF}^{\theta}(t,\vec{\boldsymbol{\Delta\theta}})$ does not depend on $\vec{\boldsymbol{\theta}}_{\tau}, \vec{\boldsymbol{\theta}}_{\tau-1}\ldots$ but only on the shock vector $\vec{\boldsymbol{\Delta\theta}}$

Network Impulse Response Function

- However we are interested on the effect of the shock on the network, not on the latent variables
- The network is a complex object and the expected value of its adjacency matrix might not be informative of the effect of the shock
- We propose to define the impulse response function on a network metric $\mathcal{P}_t = f(\mathbf{A}_t)$ as

$$\boldsymbol{IRF}^{\mathcal{P}}(t;\vec{\boldsymbol{\Delta}}\boldsymbol{\theta}) = \mathbb{E}[\mathcal{P}_{t+\tau}|\vec{\boldsymbol{\theta}}_\tau + \vec{\boldsymbol{\Delta}}\boldsymbol{\theta}, \vec{\boldsymbol{\theta}}_{\tau-1}...] - \mathbb{E}[\mathcal{P}_{t+\tau}|\vec{\boldsymbol{\theta}}_\tau, \vec{\boldsymbol{\theta}}_{\tau-1}...] \quad (7)$$

- In the fitness model, the properties of the observable network is a function of the random vector $\vec{\boldsymbol{\theta}}$
- The Impulse Response Function on the network metric is a **non-linear** function of the shock $\vec{\boldsymbol{\Delta}}\boldsymbol{\theta}$ and depends on $\vec{\boldsymbol{\theta}}$.

Network Impulse Response Function

Proposition

If $\vec{\theta}_t$ follows a VAR(1) dynamics, the Impulse Response Function for the network metric $f(\mathbf{A}_t)$ is

$$\boldsymbol{IRF}^{\mathcal{P}}(t;\vec{\Delta\theta}) = \int \mathbb{E}\left[f(\mathbf{A}_t)|\vec{\theta}_t\right] \cdot$$
$$\cdot \left[\mathcal{N}\left(\vec{\theta}_t;\boldsymbol{\mu}_t + \mathbf{B}^{t-\tau}\vec{\Delta\theta},\boldsymbol{\Sigma}_t\right) - \mathcal{N}\left(\vec{\theta}_t;\boldsymbol{\mu}_t,\boldsymbol{\Sigma}_t\right)\right] d\vec{\theta}_t$$

where

$$\begin{aligned}\boldsymbol{\mu}_t &= (\mathbf{I}-\mathbf{B})^{-1}(\mathbf{I}-\mathbf{B})^t\vec{\mu} + \mathbf{B}^{t-\tau}\vec{\theta}_\tau \\ \boldsymbol{\Sigma}_t &= (\mathbf{I}-\mathbf{B}^2)^{-1}(\mathbf{I}-\mathbf{B}^2)^t\boldsymbol{\Sigma}\end{aligned}$$

are, respectively, the conditional mean and variance.

Notice that the expectation $\mathbb{E}\left[f(\mathbf{A}_t)|\vec{\theta}_t\right]$ refers to the *static* fitness model.

Network density with fitness model

- As a specific case, we consider as network metric the density,

$$\mathcal{P} = f(\mathbf{A}) = \delta = 2(\sum A_{ij})/(n(n-1))$$

- At each time t, the expected network density is given by

$$\mathbb{E}[\delta|\mathbf{A}] = \frac{2}{n(n-1)} \sum_{i>j} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{1+e^{-(\theta_i+\theta_j)}} p_{\boldsymbol{\theta}}(\theta_i, \theta_j) d\theta_i d\theta_j,$$

- When each θ_i is Gaussian with mean m, variance s^2 and correlation r, the expected density of a network is

$$\mathbb{E}[\delta] = I(2m, 2s^2(1+r))$$

  where

$$I(m, s^2) := \int_{-\infty}^{\infty} \frac{1}{1+e^{-x}} \frac{1}{s\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-m}{s}\right)^2} dx.$$

  is the logistic-normal integral.
- For small noise s it is

$$\mathbb{E}[\delta] \approx \frac{1}{1+e^{-2m}} + (1+\rho)\frac{e^{2m}(1-e^{2m})}{(1+e^{2m})^3} s^2 + O(s^4)$$

Mean field model

- All nodes have the same unconditional mean $\vec{\mu} = \mu\vec{\mathbf{1}}$
- Assume that matrix **B** has
    - diagonal elements equal to a
    - off-diagonal elements all equal to b
- $Var[\vec{\mathbf{w}}_t] = \sigma^2\mathbf{I}$
- Alternatively, the elements of **B** are equal to 1 with probability p and zero otherwise (sparse interaction).

Mean field model: properties

- The spectral radius of $\mathbf{B}$ is $\lambda_1 = a + b(n-1)$ and covariance stationarity of the VAR model is guaranteed when $|\lambda_1| < 1$
- The stationary value of $\vec{\theta}$ is

$$\vec{\theta}_S = \frac{\mu}{1-\lambda_1}\vec{1}$$

- For a generic $\mathbf{B}$ (not necessarily mean field or dense) it is

$$\vec{\theta}_S = (\mathbf{I}-\mathbf{B})^{-1}\vec{\mu} = \mu(\mathbf{I}-\mathbf{B})^{-1}\vec{1}$$

  i.e. the stationary value of the fitness of node i is its Katz centrality when the dynamical matrix $\mathbf{B}$, is seen as an adjacency matrix of a weighted network. The stationary degree of a node (driven by θ_S) is related to its Katz centrality in the matrix $\mathbf{B}$.

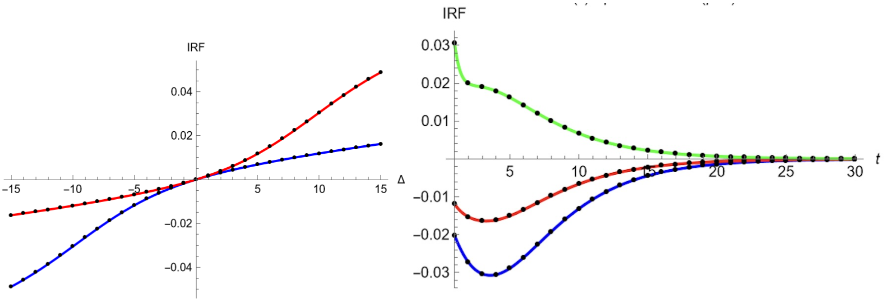
Network Impulse Response Function

- We consider a shock Δ of a mean field model affecting only the first node when all nodes have the same fitness θ_τ.
- The Network Impulse Response Function is analytically computed as

$$\begin{aligned} \boldsymbol{IRF}^{\delta}(t;\vec{\boldsymbol{\Delta}}\boldsymbol{\theta}) = {} & \frac{2}{n} I(\mu_{1,t} + \mu_{j\in V',t}, 2\sigma_t^2(1+\rho_t)) \\ & + \frac{n-2}{n} I(2\mu_{j\in V',t}, 2\sigma_t^2(1+\rho_t)) \\ & - I(2\mu_t, 2\sigma_t^2(1+\rho_t)). \end{aligned}$$

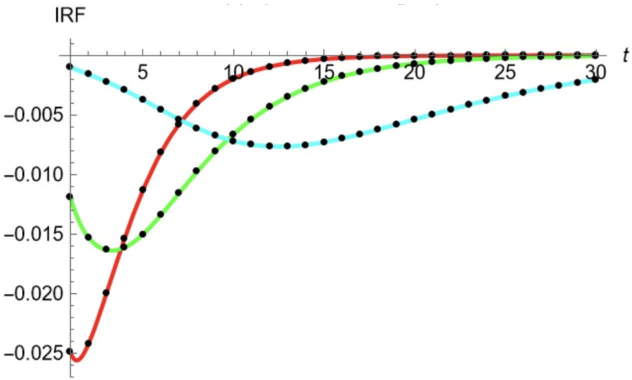
where $\mu_{1,t}$, $\mu_{j\in V',t}$, σ_t, and ρ_t are explicit functions of t, n, a, b (and p), σ, θ_τ, and Δ.

Dependence on shock size



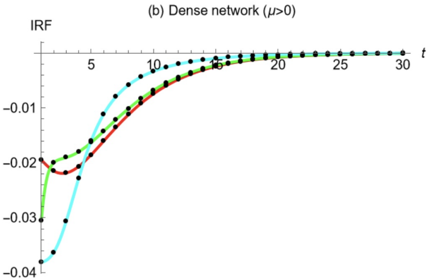
Left. IRF at time $t = 1$ as a function of Δ for sparse (red) and dense (blue) graphs.
Right. IRF on a sparse network for $\Delta = -20$ (blue), $\Delta = -10$ (red) and, $\Delta = 10$ (green).

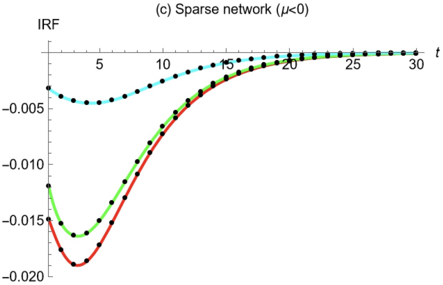
Dependence on max eigenvalue



IRF on a sparse network for $\lambda_1 = 0.69$ (red), $\lambda_1 = 0.79$ (green), and $\lambda_1 = 0.89$ (cyan)

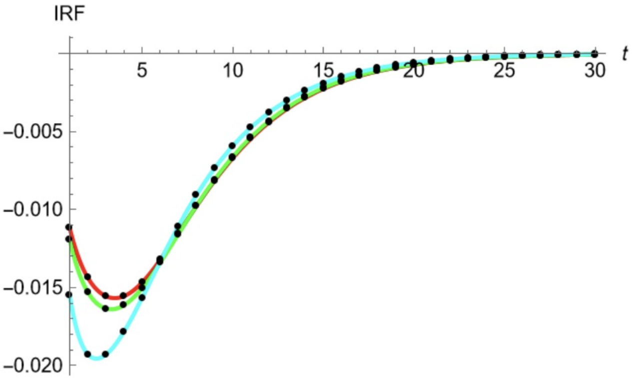
The max eigenvalue explains only part of the dynamics





IRF keeping $\lambda_1 = 0.79$ fixed: strong interaction b (red), medium interaction b (green) and weak small b (cyan).

The role of noise



IRF on a sparse network with $\sigma^2 = 0.01$ (red), $\sigma^2 = 0.1$ (green), and $\sigma^2 = 0.5$ (cyan).

Econometric estimation procedure

- The single snapshot MLE procedure (N-SSI) does not produce consistent estimations for of sparse networks (see Chatterjee et al.2011)
- We propose a new econometric approach to estimate the fitness dynamics based on the Kalman Filter (KF-SSI).
- The single snapshot estimated parameters $\vec{\Theta}_t$ are described by the state space model

$$\begin{cases} \vec{\tilde{\theta}}_t = & \vec{\tilde{\mu}} + \hat{\mathbf{B}}\vec{\tilde{\theta}}_{t-1} + \boldsymbol{w}_t, \\ \vec{\Theta}_t = & \mathbf{I}\vec{\tilde{\theta}}_t + \boldsymbol{v}_t \end{cases} \tag{8}$$

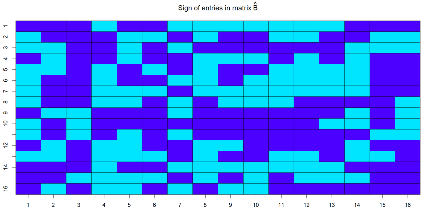
where $\mathbf{I}$ is the identity matrix, $\boldsymbol{v}_t \sim \mathcal{N}(0, \mathbf{R}_t)$, $\hat{\mathbf{B}}$ collects the estimated coefficients of matrix $\mathbf{B}$ in (4) and $\boldsymbol{w}_t \sim \mathcal{N}(0, \mathbf{Q}_t)$

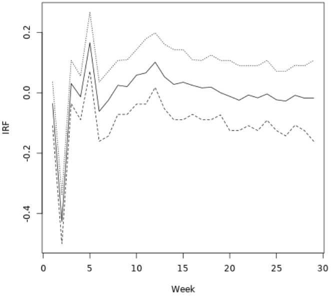
	$\theta_{i,t}$	a	b	μ	σ^2
N-SSI	0.5757569	0.3755679	0.010707	0.3219423	0.3755679
KF-SSI	0.4043644	0.1048899	0.02555794	0.1380771	0.145144

Table: The mean absolute relative error of the estimated parameters .

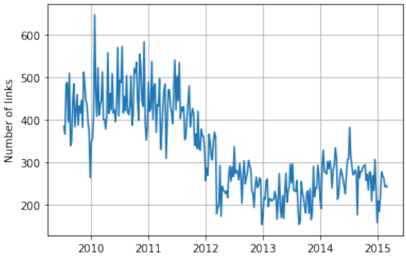
Empirical application

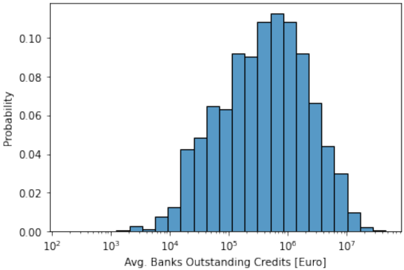
- Financial network of electronic Market of Interbank Deposit (e-MID)
- Weekly data of 8 banks in the period Jan-Oct to 2014
- Left: Mix of positive (dark blue) and negative (light blue) interactions. $\lambda_1 = 0.93$.
- Right: IRF of the density under a shock to a specific node **at a specific date**





Non stationary temporal networks





- The network dynamics is sometimes non stationary and strongly affected by external variables
- Not only topology, but also weights change with time, and the model seen so far are limited to binary networks

Score Driven Generalized Fitness Model

Modelling Choices

Contribution: Introduce a new class of weighted temporal networks model.

Key Elements

- Fitness model for degrees heterogeneity
- Zero augmentation to allow low density and modeling of weights
- Score driven, latent parameters for temporal evolution

Zero augmentation for weights

Zero Augmented Fitness Model

$$P\left(Y_{ij}^{(t)} = w\right) = \begin{cases} \frac{e^{-\left(\overleftarrow{\theta}_i + \overrightarrow{\theta}_j\right)}}{1+e^{-\left(\overleftarrow{\theta}_i + \overrightarrow{\theta}_j\right)}} & \textit{for} \quad w = 0 \\ \frac{1}{1+e^{-\left(\overleftarrow{\theta}_i + \overrightarrow{\theta}_j\right)}} g_{ij}\left(w\right) & \textit{for} \quad w > 0 \end{cases}$$

Positive Support Distribution for the Weights

- g_{ij} Gamma distribution s.t.

$$E\left[Y_{ij} | Y_{ij} > 0\right] = E_{w \sim g_{ij}}\left[w\right] = e^{\left(\overleftarrow{\lambda}_i + \overrightarrow{\lambda}_j\right)}$$

- Can handle discrete weights

Score Driven Models

Creal, Koopman and Lucas (2008), Harvey and Chakravarty (2008)

- $y^{(t)} \sim P\left(y^{(t)} | f^{(t)}\right)$, $y^{(t)}$ possibly matrix valued
- $f^{(t)}$ vector of K time varying parameters

$$f^{(t+1)} = w + b\, f^{(t)} + a\, \mathcal{I}^{(t)} \frac{\partial \log P\left(y^{(t)} | f^{(t)}\right)}{\partial f}$$

- w, a, b static parameters, $\mathcal{I}^{(t)}$ scaling matrix
- Recipe for T.V. parameters: $y^{(t)} \sim P\left(y^{(t)} | f\right) \quad \rightarrow \quad y^{(t)} \sim P\left(y^{(t)} | f^{(t)}\right)$
- MLE estimates

These are **Observation-driven models**: parameters evolve in time based on some nonlinear function of past observations

$$\text{Var}[f^{(t+1)}] > 0, \qquad \text{Var}[f^{(t+1)} | \mathcal{F}_t] = 0$$

Score Driven Gaussian Variance: GARCH

Generalized Auto-Regressive Conditional Heteroscedasticity

Consider $y_t = \sigma \epsilon^{(t)}$ where $\epsilon^{(t)} \sim N(0,1)$ hence $p(y_t|\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-y_t^2/2\sigma^2}$

$$\sigma^2_{t+1} = w + b\sigma^2_t + a\mathcal{I}^{(t)} \frac{\partial \log p(y_t|\sigma^2)}{\partial \sigma^2}$$

$$\frac{\partial \log p(y_t|\sigma^2)}{\partial \sigma^2} = \frac{\partial}{\partial \sigma^2}\left(-\frac{\log \sigma^2}{2} - y_t^2/2\sigma^2\right) = \frac{y_t^2 - \sigma^2}{2\sigma^4}$$

$$\mathcal{I} = 2\sigma^4 \quad \rightarrow \quad \sigma^2_{t+1} = w + b\sigma^2_t + a(y_t^2 - \sigma^2_t) = w + a' y_t^2 + b' \sigma^2_t$$

Many well known models can be obtained as SD (EGARCH, MEM, ACD ...)

The GARCH (and SD models) as a filter of a misspecified dynamics

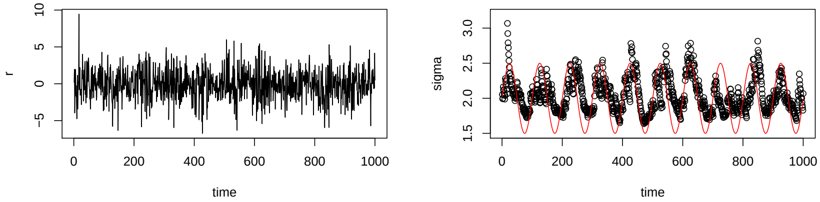
Simulate 1000 observations from the model

$$r_t = \sigma_t \epsilon_t \qquad \epsilon_t \sim \mathcal{N}(0,1) \qquad \sigma_t = 2 + 0.5 \sin \frac{\pi \ t}{100}$$

and we use

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$$

to **filter** the values of σ_t.



(Left) Artificially generated time series of returns.
(Right) Filtered (black) and real (red) values of σ_t.

Score Driven Generalized Fitness Model

Gamma Distributed Weights

Definition

$$P\left(Y_{ij}^{(t)} = w\right) = \begin{cases} \dfrac{e^{-(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)})}}{1+e^{-(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)}+)}} & \textit{for} \quad w = 0 \\ \dfrac{\left(\mu_{ij}^{(t)}\right)^{-\sigma}\Gamma(\sigma)^{-1}}{1+e^{-(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)})}} w^{(\sigma-1)} e^{-\frac{w}{\mu_{ij}^{(t)}}} & \textit{for} \quad w > 0\,. \end{cases}$$

with $\mu_{ij}^{(t)} = \sigma^{-1} e^{\left(\overleftarrow{\lambda}_i^{(t)} + \overrightarrow{\lambda}_j^{(t)}\right)}$ Score Details

- Maximum Likelihood Estimation
- It can be seen as a filter of misspecified dynamics
- Provides a real time estimation of latent parameters

Score Driven Generalized Fitness Model

Details: binary score

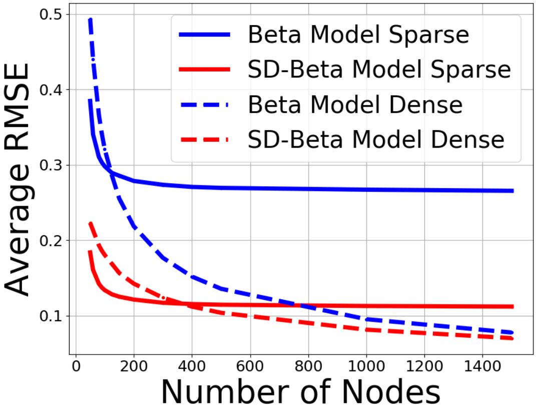
We let the fitness, both binary and weighted, evolve in time, following the score-driven recursive update rule with score

$$\frac{\partial \log P\left(\mathbf{Y}^{(t)} | f^{(t)}\right)}{\partial \overleftarrow{\theta}_k^{(t)\prime}} = \sum_j \left(A_{kj}^{(t)} - \frac{1}{1 + e^{-\left(\overleftarrow{\theta}_k^{(t)} + \overrightarrow{\theta}_j^{(t)}\right)}} \right)$$

$$\frac{\partial \log P\left(\mathbf{Y}^{(t)} | f^{(t)}\right)}{\partial \overrightarrow{\theta}_k^{(t)\prime}} = \sum_i \left(A_{ik}^{(t)} - \frac{1}{1 + e^{-\left(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_k^{(t)}\right)}} \right) \quad (9)$$

and does not depend on the choice of g

SD-Fitness (Beta) Model for large graphs



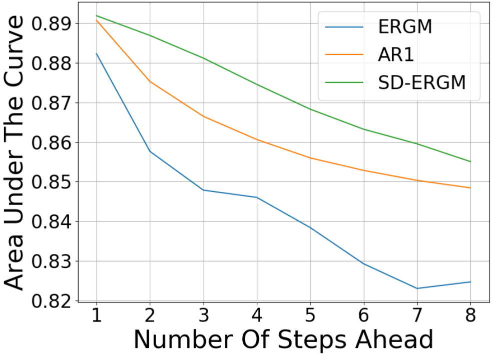
Average RMSE of the filtered parameters with respect to the simulated DGP in both the dense (dashed lines) and sparse (solid lines) regimes. The average RMSE from the ERGM is plotted in blue, while the one from the SD-ERGM in red.

Sequence of Static Vs Score Driven Models Comparison

Binary Link Prediction: Existence of a Link in the Future

Comparison of Binary Models

- e-Mid Data, rolling windows of 100 observations
- Forecast up to 8 steps ahead (roughly two months)



Sequence of Static Vs Score Driven Models Comparison

Weight Prediction

Forecast Weights of Existing Links

- Focus on the present links only
- One step ahead

Method	Score Driven	AR(1) Single Snapshot	Diebold-Mariano (p-value)
MSE Log.	0.859	0.882	1.73×10^{-7}
MAD Log.	0.726	0.737	1.21×10^{-7}

Comparison with Benchmark For Sparse Weighted Networks

Link Specific Regression Model from The literature

Localized Tobit Model of Giraitis et. al. 2016

- Separate modeling of each $N(N-1)$ links
- Censored Linear Regression: joint modeling of link and weights
- Define 6 functions of $\mathbf{Y}^{(t-1)}$ as regressors
  - $Y_{ij}^{(t-1)}$;
  - lagged total daily amount lent by i to all other banks except j ;
  - lagged total daily amount borrowed by j from all other banks except i ;
  - lagged total daily amount lent by j to all other banks except i ;
  - lagged total daily amount borrowed by i from all other banks except j ;
  - lagged total daily lending and borrowing not involving either i or j.
- Local likelihood estimate of the parameters

Comparison with Benchmark For Sparse Weighted Networks

Comparison in One Step Ahead Forecasting

Model	N_{par}	MSE Log.	AUC
Localized Tobit	$\sim N^2$	2.351	0.714
SD Gen. Fit.	$\sim N$	0.859	0.896

- Lower number of parameters gives SD model an edge
- Tobit does not decouple links' presence from their weights

Model	T_{train}	MSE Log.	MAD Log.	AUC
Localized Tobit	100	2.351	1.067	0.714
ZA Regression	100	2.785	1.240	0.830
SD Generalized Fitness	100	**0.859**	**0.726**	**0.896**

Localized Tobit model of Giraitis et al (2016) vs simple Zero Augmented regression that uses the same regressors vs score driven generalized fitness model.

External Regressors in SD Generalized Fitness Model

Example of Scalar $x^{(t)}$

- Binary part

$$p_{ij}^{(t)} = \frac{1}{1 + e^{-(\overleftarrow{\theta}_i^{(t)} + \overrightarrow{\theta}_j^{(t)} + x^{(t)}\beta_{bin})}}$$

- Weights

$$E\left[Y_{ij}{}^{(t)} | w > 0\right] = e^{\left(\overleftarrow{\lambda}_i^{(t)} + \overrightarrow{\lambda}_j^{(t)} + x^{(t)}\beta_w\right)}.$$

Case study:

impact of interest rates on the probability of observing each link and on the expected weight of observed links

EONIA as External Regressor on e-Mid

EONIA

is a measure of the effective interest rate prevailing in the euro inter-bank overnight market. It is computed as a weighted average of the interest rates on unsecured overnight contracts (Definition from `https://stats.oecd.org/`).

EONIA as External Regressor on e-Mid

	Constant Fitness	SD Fitness
BIC Bin	0.53×10^6	0.45×10^6
BIC Weight	1.46×10^6	1.46×10^6
AUC - Test Set	0.82	0.92
MSE Log. - Test Set	1.01	0.78
β_{bin}	0.69 ± 0.06	0.29 ± 0.05
β_w	0.022 ± 0.029	-0.13 ± 0.02

- Importance of including time varying fitness to improve goodness of fit, both in sample and out of sample
- The probability of observing a link is positively related with the interest rates, hence the lowering of interest rates tends to reduce the overall market interconnectdness, even taking into account bank specific effects captured by the fitness
- The weight of the observed overnight loans is negatively related with the average interest rate in the market
- Thanks to zero augmentation we separate the role of EONIA on links and weights

Summary

- Probabilistic models of networks are rich and flexible tools for investigating structures and dynamics' drivers
- We study how an external shock propagates on a temporal network using a VAR(1) dynamics of the fitness model
- We derive the closed form for the nonlinear Impulse Response Function of a generic shock
- Prominent role in the shock spreading played by the lagged interactions between latent variables
- We propose a method based on Kalman Filter to estimate the dynamics and the IRF
- We propose a generalization allowing for non-stationary dynamics of fitness and considering temporal weighted networks
- Superior performances in link and weight prediction w.r.t. static version and benchmark from the literature
- Model external dependency on binary and weighted part

Summary

- Probabilistic models of networks are rich and flexible tools for investigating structures and dynamics' drivers
- We study how an external shock propagates on a temporal network using a VAR(1) dynamics of the fitness model
- We derive the closed form for the nonlinear Impulse Response Function of a generic shock
- Prominent role in the shock spreading played by the lagged interactions between latent variables
- We propose a method based on Kalman Filter to estimate the dynamics and the IRF
- We propose a generalization allowing for non-stationary dynamics of fitness and considering temporal weighted networks
- Superior performances in link and weight prediction w.r.t. static version and benchmark from the literature
- Model external dependency on binary and weighted part

Summary

- Probabilistic models of networks are rich and flexible tools for investigating structures and dynamics' drivers
- We study how an external shock propagates on a temporal network using a VAR(1) dynamics of the fitness model
- We derive the closed form for the nonlinear Impulse Response Function of a generic shock
- Prominent role in the shock spreading played by the lagged interactions between latent variables
- We propose a method based on Kalman Filter to estimate the dynamics and the IRF
- We propose a generalization allowing for non-stationary dynamics of fitness and considering temporal weighted networks
- Superior performances in link and weight prediction w.r.t. static version and benchmark from the literature
- Model external dependency on binary and weighted part