Frontiers of Network Science
Fall 2023
Class 5: Seminar

Failures, Dynamics, Evolution and Control of the Global Risk Network By World Economic Forum

Boleslaw Szymanski
Xin Lin, Xiang Niu, Noemi Derzsy, Alaa Moussawi,
Jianxi Gao, and Gyorgy Korniss

NeST Center & SCNARC
Department of Computer Science
Department of Physics, Applied Physics and Astronomy
Rensselaer Polytechnic Institute, Troy, NY

Outline

- **Background information**
- **Model details and dynamics**

Definitions

- Global risks are defined each year by the World Economic Forum (WEF) which defines Global Risk Networks for its meeting in Davos since 2000. We are building models of this networks starting with the 2013 Global Risk Network* but using data from all networks.

- Each WEF report identifies **30-50 global risks** that are classified into five broad categories of **economic**, **environmental**, **geopolitical**, **societal** and **technological** risks.

- The report contains also an assessment of the risks potential **impact**, **interconnectedness** and **likelihood of materialization** in the next 10 years prepared by over a thousand industrial, governmental, and academic experts.

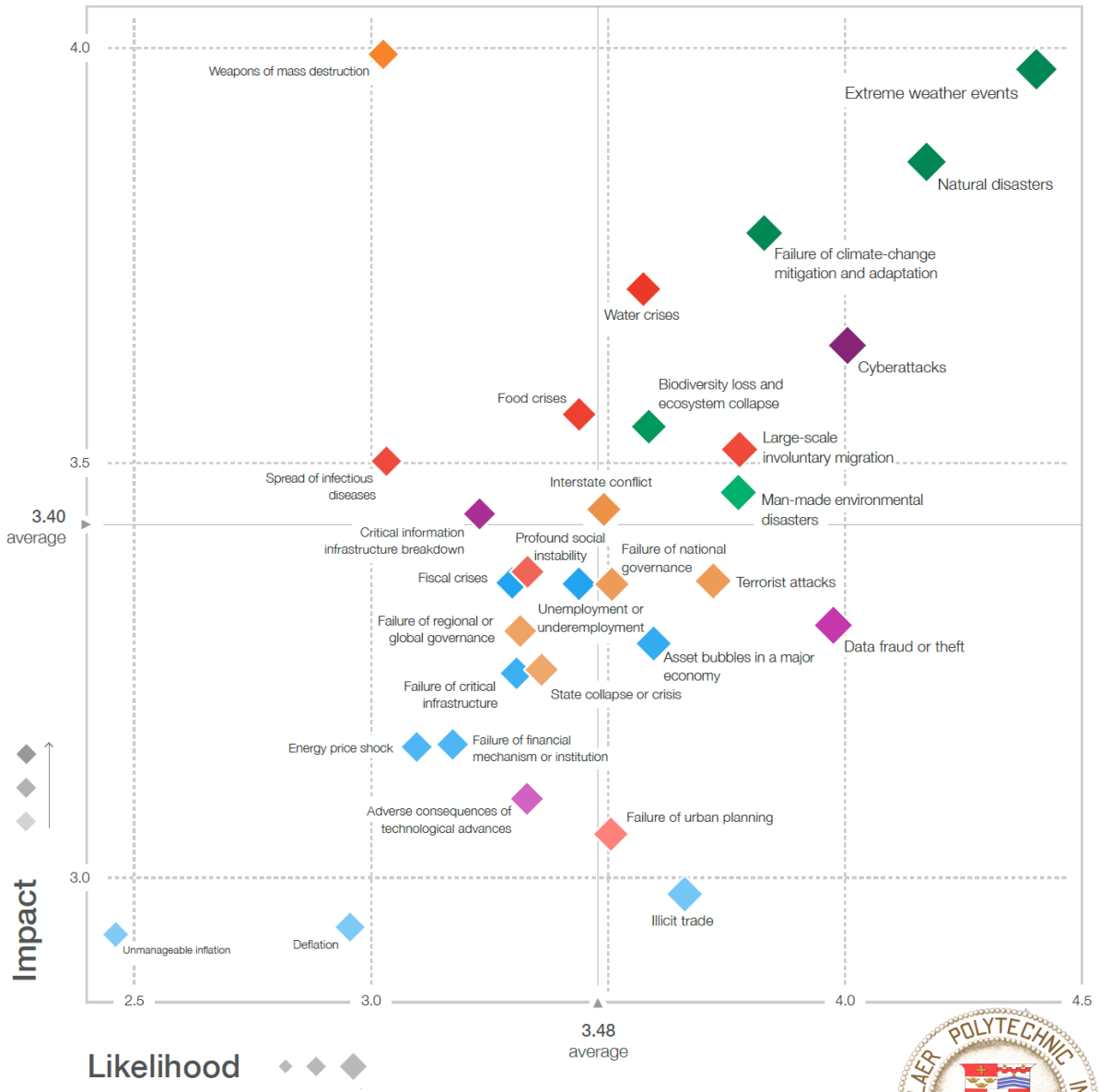
**World Economic Forum Global Risks Report* (2013)
http://www.weforum.org/reports/global-risks-2013-eighth=edition

An Example of 2018 Report Data*

Distribution of likelihood and impact for **30 global risks grouped into five categories** indicated by the color of risk node by GRR18*

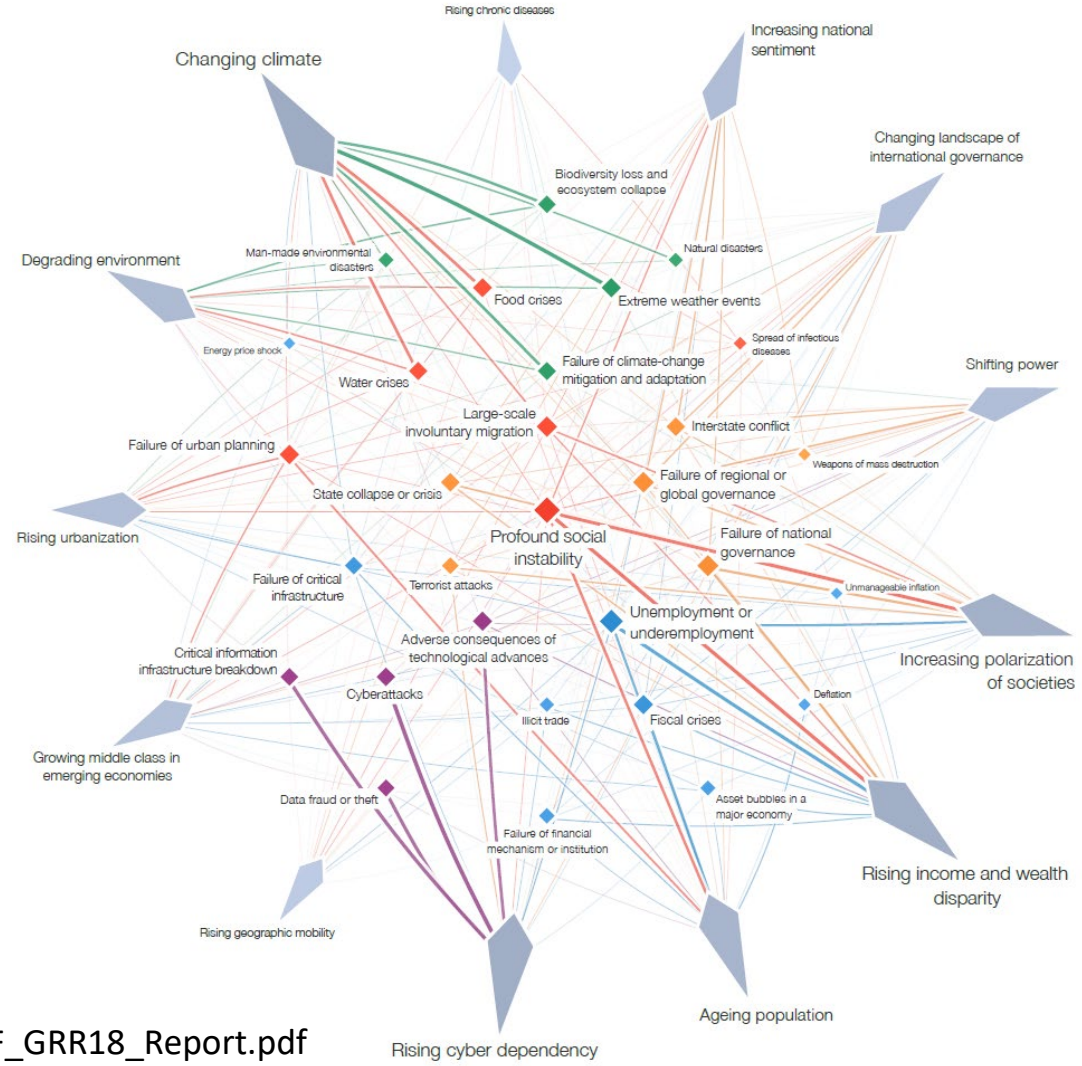
*

http://www3.weforum.org/docs/WEF_GRR18_Report.pdf



An Example of 2018 Report Data*

2018 Interconnections of Global Risks reported In GRR18*

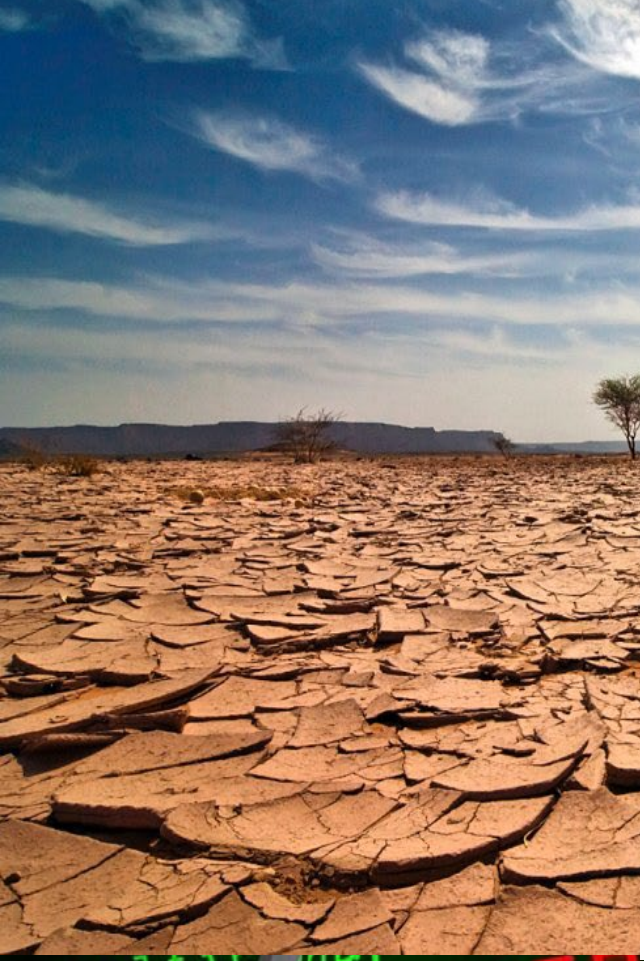


* http://www3.weforum.org/docs/WEF_GRR18_Report.pdf

SCNARC

RENSSELAER POLYTECHNIC INSTITUTE 1824

Global Risks



Weapons of Mass Destruction

Failures, Dynamics, Evolution and Control of the Global Risk Network by World Economic Forum

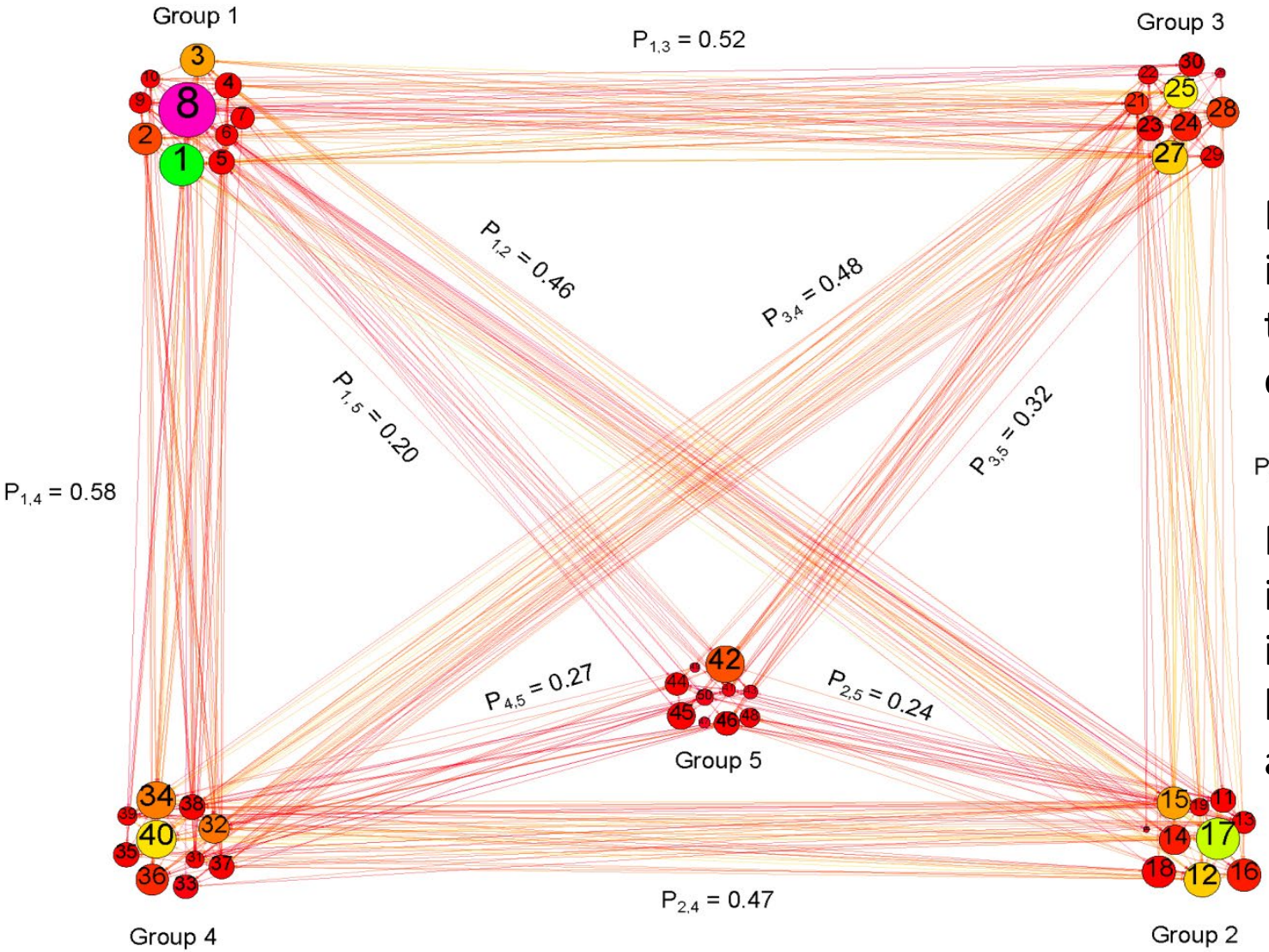
Dynamics of Global Risk Failures

- Each risk is a node in the global risk network and is assigned **likelihood of failure in the next 10 years** and the **failure economical impact** by the 20-25 **experts in this risk***.
- The risk network is a **Stochastic Block Model (SBM) graph**** with the **average node connectivity 20.6** and undirected weighted edges representing risks influence of nodes on each other.
- The **optimal model** can measure how decreasing the strength of some critical network nodes and edges can decrease probability of future failures.

*_World Economic Forum_ (2013) http://www3.weforum.org/tools/rrn/wef_grr/20130108/server/getrisks.json

**R. Holland, K. Laskey & S. Leinhardt, Stochastic blockmodels: first steps. _Soc. Networks_ 5:109–137 (1983)

Visualization of Inter-group Connectivity



Node's color indicate its total connectivity

Node's size indicates its internal likelihood of activation

The number of inter-group links indicates the strength of connectivity between a pair of groups

Alternating Renewal Processes*: Discrete Model

Each risk at time *t* is either at state *1* (**materialized** also called **active**) or state *0* (**not materialized** also called **inactive**) but can transition to a new state in the next time instance by one of the three processes.

Process 1. For a risk *i*, given that it is in state *0*, its **internal materialization** is a **latent** Poisson process with intensity λ_i^{int}.

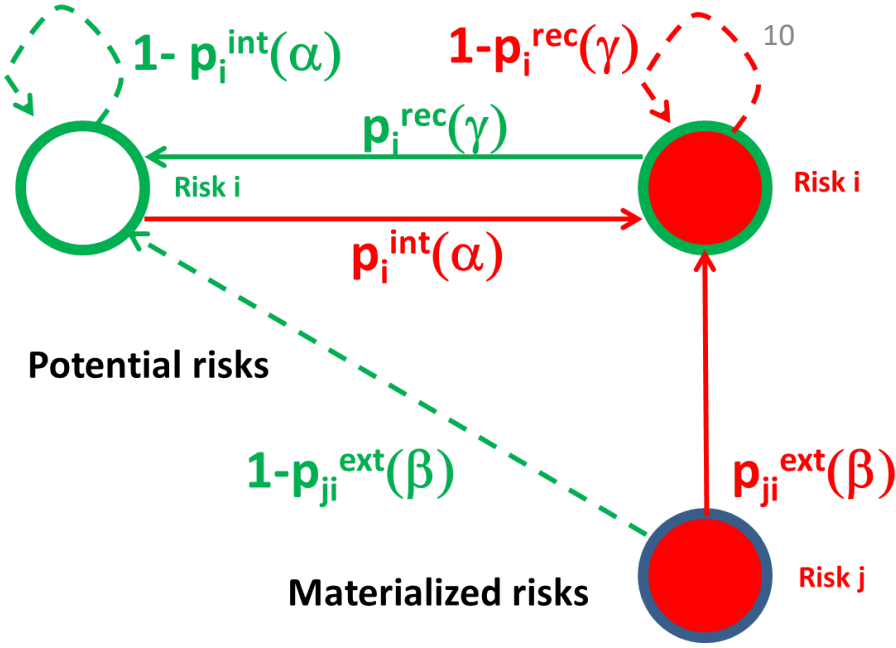
Process 2. Given that the risk *i* is in state *1*, its **recovery** from this state is an **observable** Poisson process with intensity λ_i^{rec}.

Process 3. Given that risk *i* is in state *0*, and risk *j* connected to it in state *1*, the materialization of risk *i* due to the **external influence** of risk *j* is a **latent** Poisson process with intensity λ_i^{ext} (which is independent of *j*).

Only cumulative effect of **Processes 1** and **2** is **observable**.

* D.R. Cox & H.HD. Miller, *The Theory Of Stochastic Processes* (Methuen, 1965)

Model Dynamics



The dynamics is Markovian and progresses by assuming that at each time step t:

A risk *i* that was inactive at time *t - 1* **materializes** internally with probability $p_i^{\text{int}} = 1 - e^{-\lambda_i^{\text{int}}}$

A risk *j* active at time *t - 1* causes a connected risk *i* inactive at time *t - 1* to **materialize** with probability $p_{ji}^{ext} = 1 - e^{-\lambda_i^{ext}} = p_i^{ext}$

A risk *i* that was active at time *t – 1* **recovers** from its active risk with probability

$$p_i^{rec} = 1 - e^{-\lambda_i^{rec}}$$

SCNARC

RENSSELAER POLYTECHNIC INSTITUTE 1824

Model Parameters

Normalizing the likelihood values, L_i, given in the range [*1, 5*] to their natural range of [*0, 1*], we obtain:

$$N_I = (L_i\text{-}1)/4$$

This normalized likelihood values N_i are in direct proportion to the expert assessment L_i and for our purposes captures the node's vulnerability to a failure.

Probability **of internally triggering** inactive risk is:

$$p_i^{\text{int}} = 1-(1-N_i)^{\alpha}$$

Probability of active risk **recovering** to inactive state is:

$$p_i^{rec} = 1-N_i^{\gamma}$$

Probability of **triggering externally** inactive risk *i* by active risk *j* is:

$$p_i^{ext} = 1-(1-N_i)^{\beta}$$

Outstanding Questions

- **Limits of parameter recovery**
- **Evolution of the WEF model over years**
- **Control of the global risk network**
- **Conclusion**

Model Parameters

Normalizing the likelihood values, L_i, given in the range [*1, 5*] to their natural range of [*0, 1*], we obtain:

$$N_I = (L_i\text{-}1)/4$$

This normalized likelihood values N_i are in direct proportion to the expert assessment L_i and for our purposes captures the node's vulnerability to a failure.

Probability **of internally triggering** inactive risk is:

$$p_i^{\text{int}} = 1-(1-N_i)^{\alpha}$$

Probability of active risk **recovering** to inactive state is:

$$p_i^{rec} = 1-N_i^{\gamma}$$

Probability of **triggering externally** inactive risk *i* by active risk *j* is:

$$p_i^{ext} = 1-(1-N_i)^{\beta}$$

Parameter Recovery Formulation

We have four parameters for each **Poisson process**. The relationship between the intensities and parameters are:

$$\lambda_i^{\text{int}} = -\alpha \ln(1 - N_i)$$
$$\lambda_i^{ext} = -\beta \ln(1 - N_i)$$
$$\lambda_i^{rec} = -\gamma \ln(N_i)$$

N_i reflects the properties of risk *i*.

To find the optimal values for the parameters, we implement parameter recovery on the historical data using **Maximum Likelihood Estimation (MLE).** The likelihood *L()* of observing a sequence $S(1) \to S(2), \cdots S(T-1) \to S(T)$ of risk materialization events, where *S(t)* denotes the state of all risks at step *t*, can be written as:

$$\ln L(S(1) \to S(2), \cdots, S(T-1) \to S(T)) = \sum_{t=1}^{T-1} \sum_{i=1}^{R} \ln p_i(t)^{s_i(t) \to s_i(t+1)}$$

where *T* is total size of historical data, *R* is the number of risks, $s_i(t)$ is the state of risk *i* at step t, and $p_i(t)^{s_i(t) \to s_i(t+1)}$ denotes the probability of transition for risk *i* at step *t*.

The optimal values for parameters are those that maximize the likelihood of observing the historical data with the assumed distribution.

Parameter Recovery Formulation

We have four parameters for each **Poisson process**. The relationship between the intensities and parameters are:

$$\lambda_i^{\text{int}} = -\alpha \ln(1 - N_i)$$
$$\lambda_i^{ext} = -\beta \ln(1 - N_i)$$
$$\lambda_i^{rec} = -\gamma \ln(N_i)$$

N_i reflects the properties of risk *i*.

To find the optimal values for the parameters, we implement parameter recovery on the historical data using **Maximum Likelihood Estimation (MLE).** The likelihood *L()* of observing a sequence $S(1) \to S(2), \cdots S(T-1) \to S(T)$ of risk materialization events, where *S(t)* denotes the state of all risks at step *t*, can be written as:

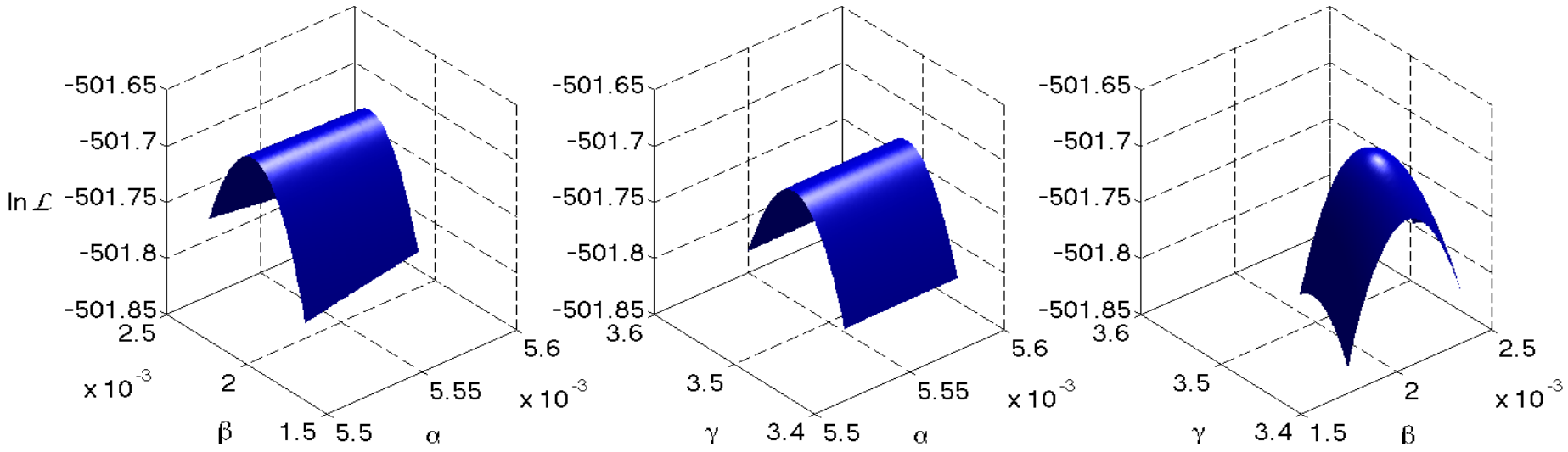
$$\ln L(S(1) \to S(2), \cdots, S(T-1) \to S(T)) = \sum_{t=1}^{T-1} \sum_{i=1}^{R} \ln p_i(t)^{s_i(t) \to s_i(t+1)}$$

where *T* is total size of historical data, *R* is the number of risks, $s_i(t)$ is the state of risk *i* at step t, and $p_i(t)^{s_i(t) \to s_i(t+1)}$ denotes the probability of transition for risk *i* at step *t*.

The optimal values for parameters are those that maximize the likelihood of observing the historical data with the assumed distribution.

MLE* Values of Parameters

- By scanning different combinations of α, β and γ over their respective acceptable ranges, and by computing the resulting log-likelihoods, we find with the desired precision the values of α, β and γ that maximize the likelihood of observing the data.
- The likelihood function is itself smooth with a unique maximum that guarantees that the found parameter values are indeed globally optimal for the model considered.
- **α=0.364≈4/11, β=0.14≈1/7, γ=427**



* Y. Pawitan, *In All Likelihood: Statistical Modelling And Inference Using Likelihood* (Clarendon, 2001)

Historical Data of Global Risks

- We collected data on the materialization of each risk over the period **2000 – 2017**.
- The source of data is news, magazine, academic articles and websites.
- There are **50 * 13 * 12 = 7800** data points for finding system parameters for 2013.

	Events	Influence	Period	Reference
Global Governance Failure (19)	**Crisis in Syria**	**Global and serious**	**2011.05 --- 2012.12**	http://en.wikipedia.org/wiki/Syrian_Civil_War http://www.usaid.gov/crisis/syria http://www.cbsnews.com/feature/syria-crisis/ http://www.bbc.com/news/world-middle-east-17258397
	Crisis in Libya	**Global and serious**	**2011.02 --- 2012.12**	http://en.wikipedia.org/wiki/Libyan_Civil_War http://www.nytimes.com/2013/10/11/opinion/libyas-security-crisis.html?gwh=3863288CE0B70560093F7E40D0254D5D&gwt=pay http://www.bbc.com/news/world-middle-east-12480844

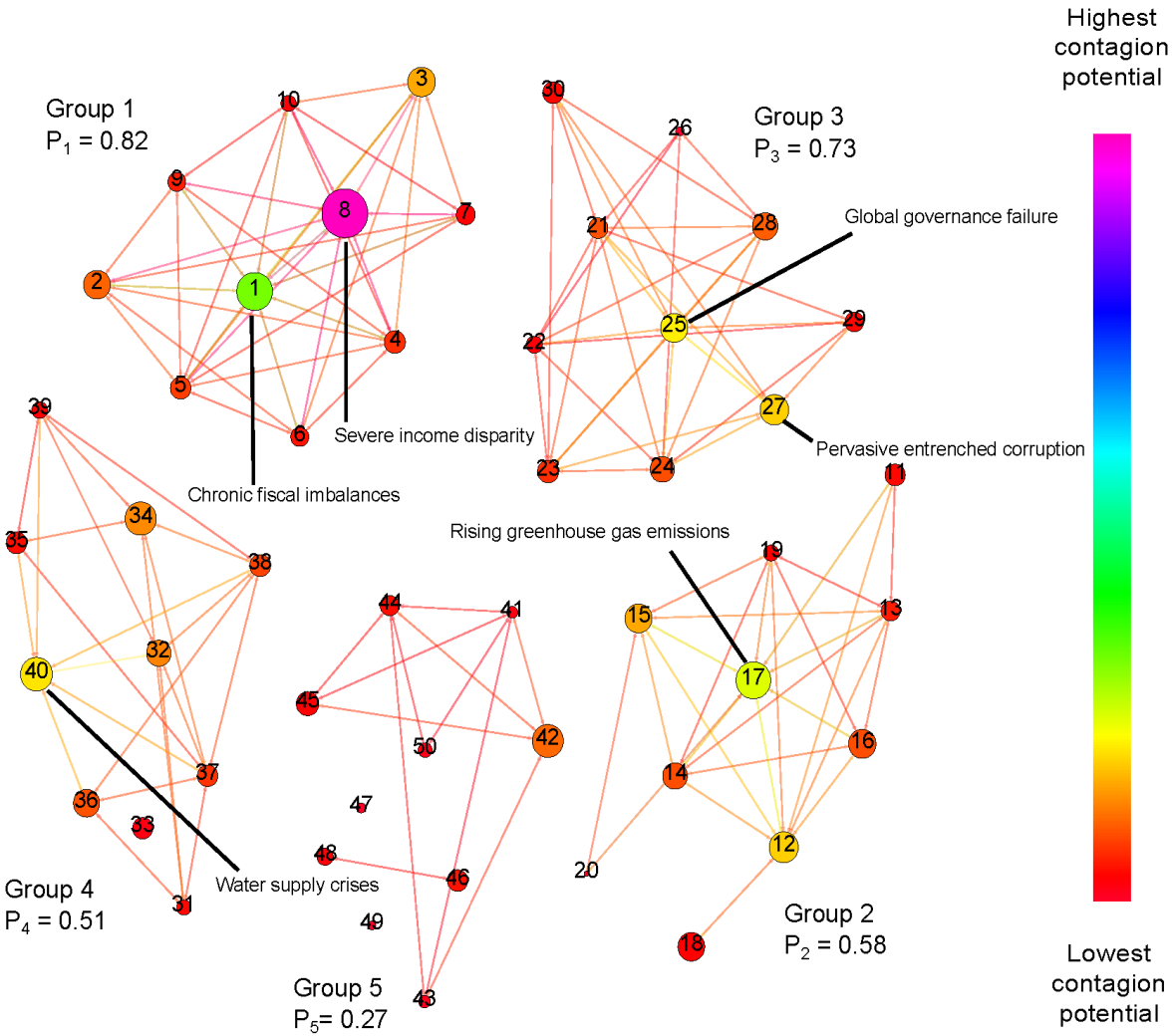
The optimal values of model parameters α, β, γ based on collected data

Year	Nodes	Edges	Avg. Degree	Edge Prob.	Avg. CC.	Diameter	α	β	γ
2013	50	515	20.60	0.42	0.61	3	0.364	0.140	427
2017	30	275	18.33	0.63	0.74	2	0.634	0.364	300

Contagion Potentials Versus Internal Vulnerability

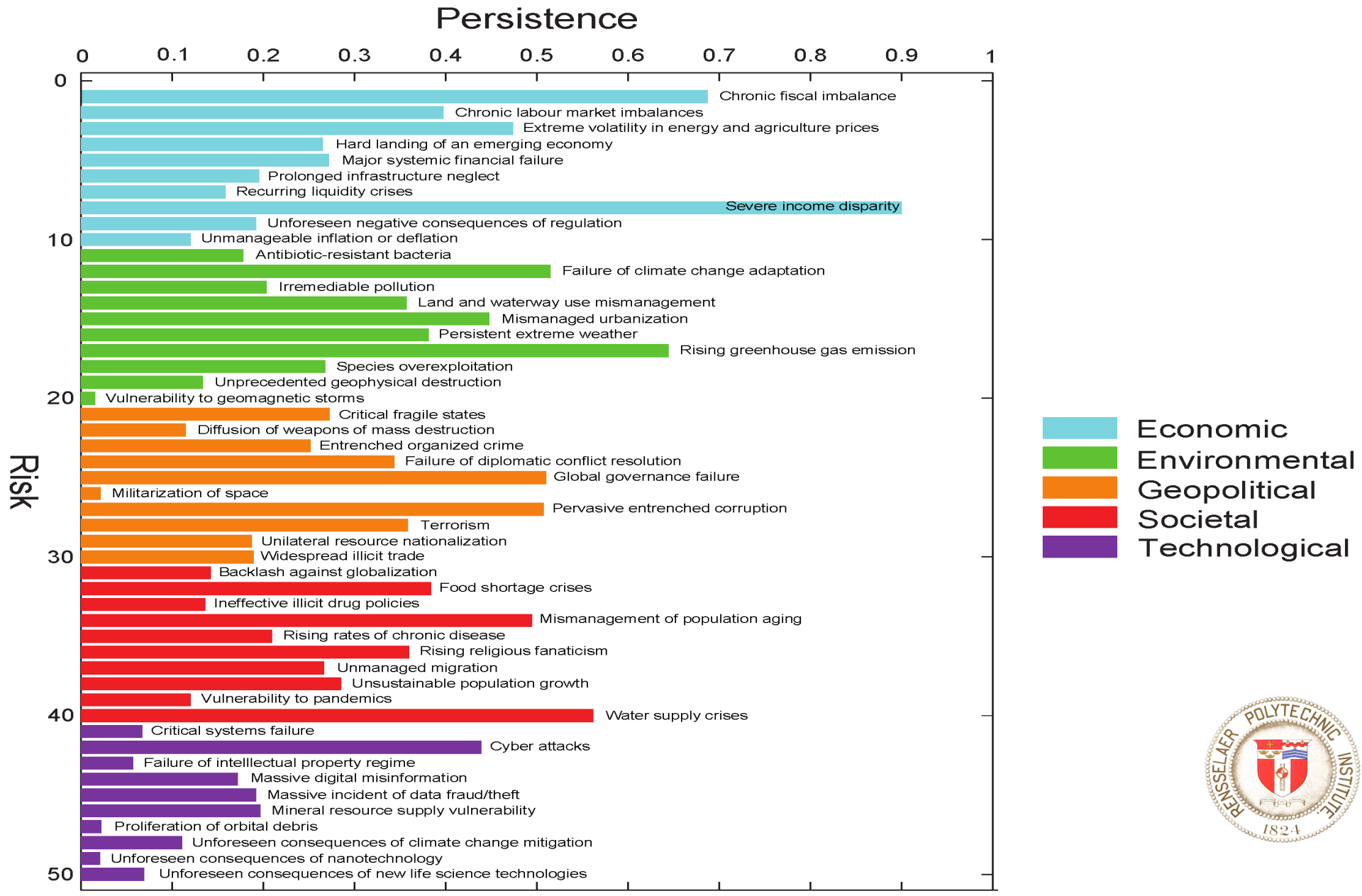
Five risks with the highest contagion potential in 2013 were:

8 -- Severe income disparity
25 – Global government failure
1 – Chronic fiscal imbalances
27 – Pervasive entrenched corruption
12 -- Failure of climate change adaptation



Network visualization showing the contagion potentials (indicated by color) and the internal failure probabilities (indicated by size) with the optimal parameters.

Asymptotic (Steady State) Risk-Persistence for 2013



Conclusions Based on the Network Model

- For 2013 data, we rank order the persistence of various risks during the lifetime of a cascade. Strikingly, risk 8 - "Severe income disparity" - was active for about 80% of the lifetime of a cascade on average, while in comparison, the second most persistently active risk - "Chronic fiscal imbalances" - was active for about 33% of the lifetime of a cascade on average.

- Decreasing the internal failure and external influence probabilities of global risks both contribute to the stability of the global economy, with reduction of internal failure probabilities contributing more effectively.

Measuring Quality of Predictions

- In the global risk network, **we recover latent (hidden) and explicit parameters of the model** from historical data and predict future activation of global risks, including cascades of such risk activations.
- The question arises **how reliable such parameter recovery is** and **how the recovery precision depends on the complexity of the model** and the length of its historical data.
- Here, this model is applied to **fire propagation in an artificial city** with modular blocks that can be assembled into a complex system.
- We simulate the fires in such cities of varying size, over varying periods of times and **use the Maximum Likelihood Estimation to recover the parameters** and compare them with the values assigned to them in simulations.

* X. Lin, A. Moussawi, G. Korniss, J.Z. Bakdash, and B.K. Szymanski, Limits of Risk Predictability In a Cascading Alternating Renewal Process Model, *Scientific Reports* **7**:6699, (2017)

Model Details

In the firehouse model, three types of houses are defined: **small, medium and large.**

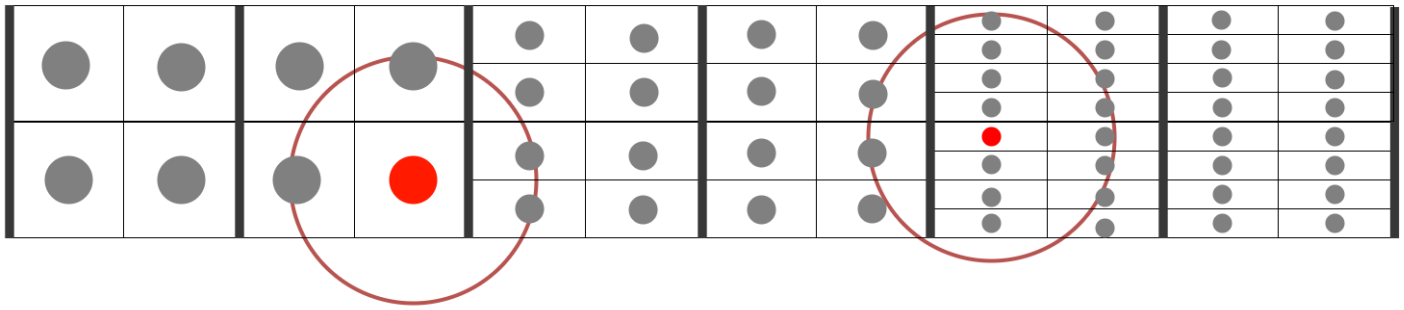
For each house, the quality of its building material and size of its lot are proportional to its size (type).

Each house fire worthiness properties, such as resistance to fires, ability to spread fire, etc. are determined by its type and the housing density in its neighborhood.

Large houses have a low probability of catching fire and a high ability to recover from burning, while medium and small houses have increasingly lower characteristics.

Each house has an influencing circle with a fixed radius, in which all the neighbors inside the circle are at risk of being catching fire from this house.

We can expand the city by adding blocks horizontally and vertically.

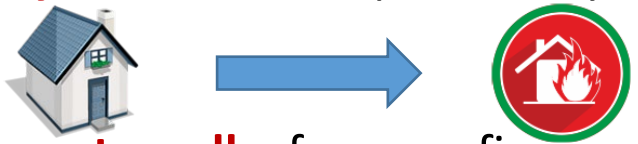


Model Dynamics - Probabilities

At time t, a house is in one of three states: **susceptible (0), on-fire (1), recovery(-1)**
The dynamics progresses at each time step t > 0 as follows:

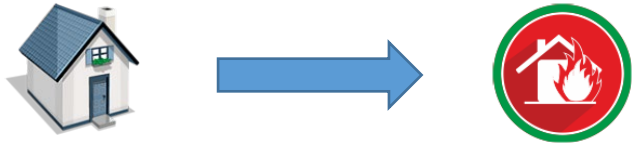
1. House *i* **susceptible** at time *t–1* **catches fire internally** at time *t* with probability:

$$p_i^{\text{int}} = 1 - e^{-\lambda_i^{\text{int}}}$$



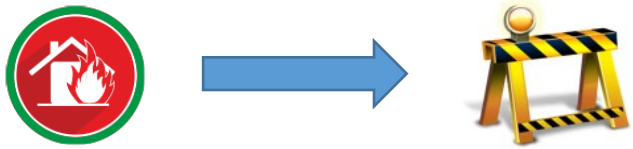
2. House *i* that was **susceptible** at time *t–1* **catches fire externally** from on-fire neighbor *j* at time *t* with probability:

$$p_{ji}^{ext} = 1 - e^{-\lambda_i^{ext}} = p_i^{ext}$$



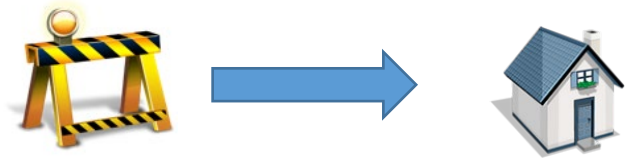
3. House *i* **on-fire** at time t–1 is extinguished and enter the **recovery** state at time *t* with probability:

$$p_i^{fire} = 1 - e^{-\lambda_i^{fire}}$$



4. House *i* **in-recovery** at time *t–1* becomes **susceptible** at time *t* with probability:

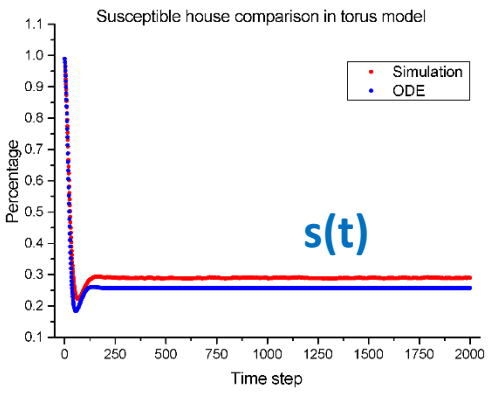
$$p_i^{rec} = 1 - e^{-\lambda_i^{rec}}$$

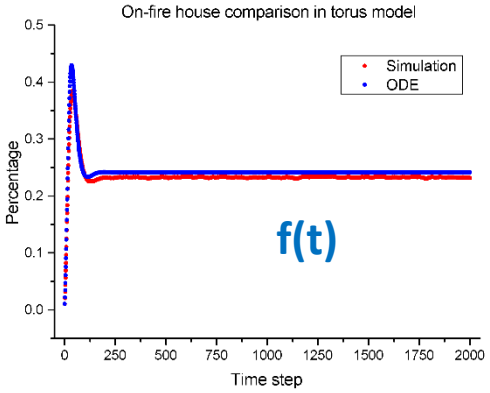


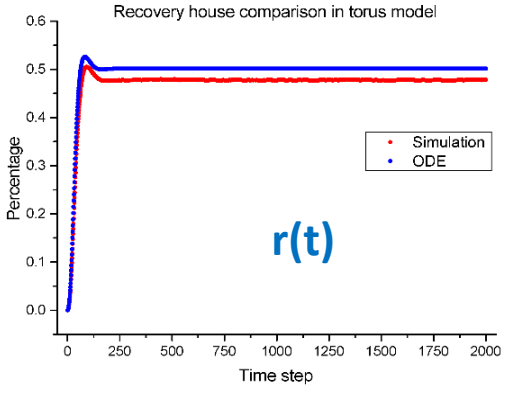
SCNARC

Model Comparison – Discrete vs. ODE

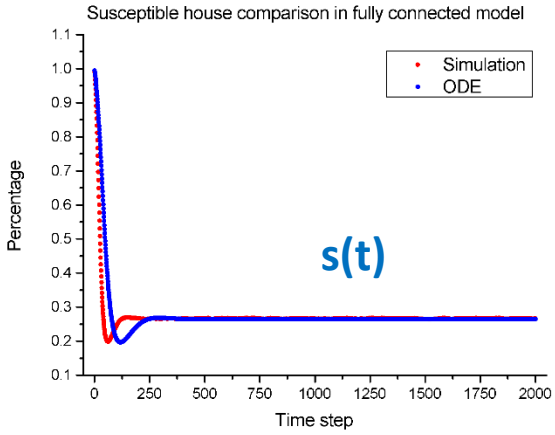
Three state comparisons in torus model

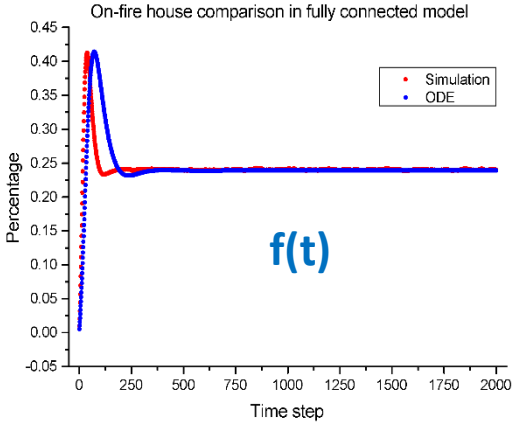


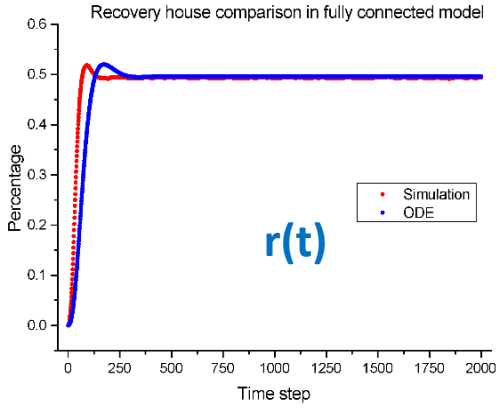




Three state comparisons in fully connected model

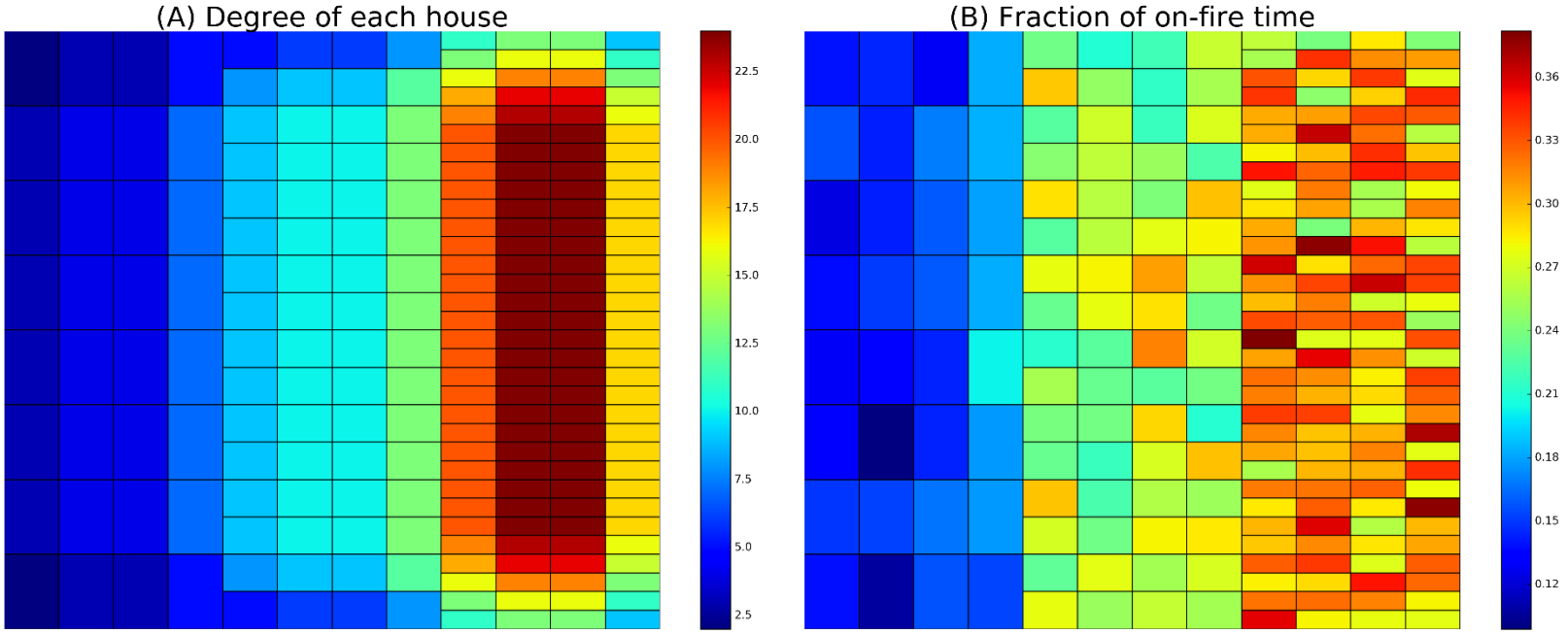






The higher the model connectivity, the better the match between discrete and ODE results.

Fraction of On-fire Time in Simulation



- Time steps of simulation: 10,000
- Counts at how many time steps each house is on fire
- Averaged 20 independent realizations

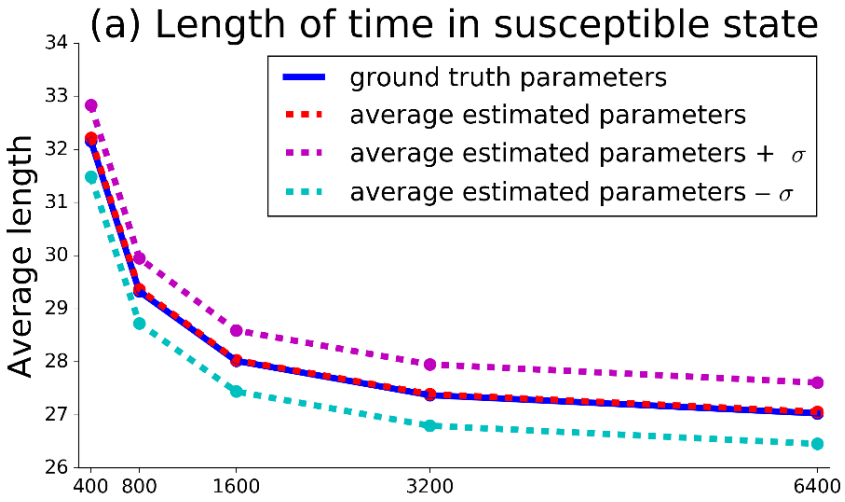
Measuring Performance of Estimated Parameters

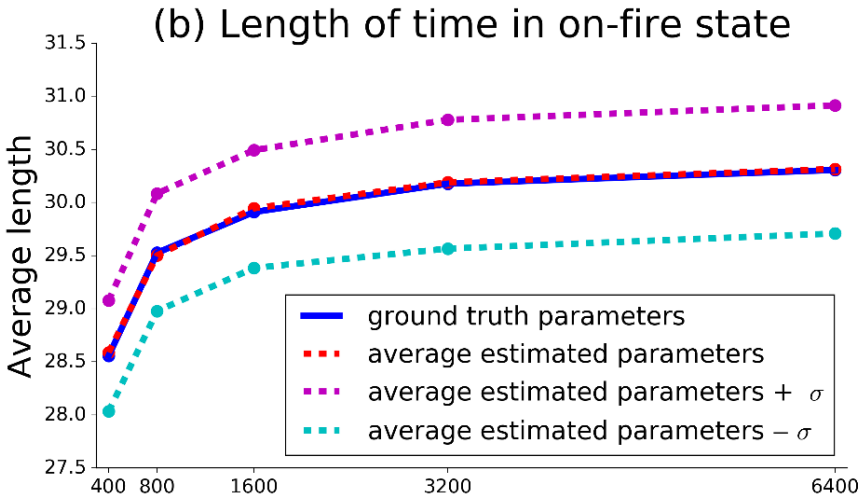
- We generate 125 sets of time series for 6400 steps using ground truth parameters: $\alpha = 0.008,\ \beta = 0.012,\ \gamma = 0.016,\ \delta = 0.032$ and compute estimated parameters from each of time series using Maximum Likelihood Estimation.
- Using sets of estimated parameters, we simulate multiple lengths of future periods: 400, 800, 1600, 3200 and 6400 and record the length of normal, on-fire and recovery state as well as the number of emerging fires during the period; results are averaged over 20 realizations.
- We compute the difference between these estimated parameters and the ground truth parameters and, determine, using Kolmogorov-Smirnov metric the $\pm\sigma$ distance between estimated parameters and ground truth parameters.
- Finally, we find the $\pm\sigma$ boundary of each set of estimated parameters by removing the largest 39 sets of results; the remaining results contain 68% of all sets of estimated parameter values, that are closest to the ground truth.

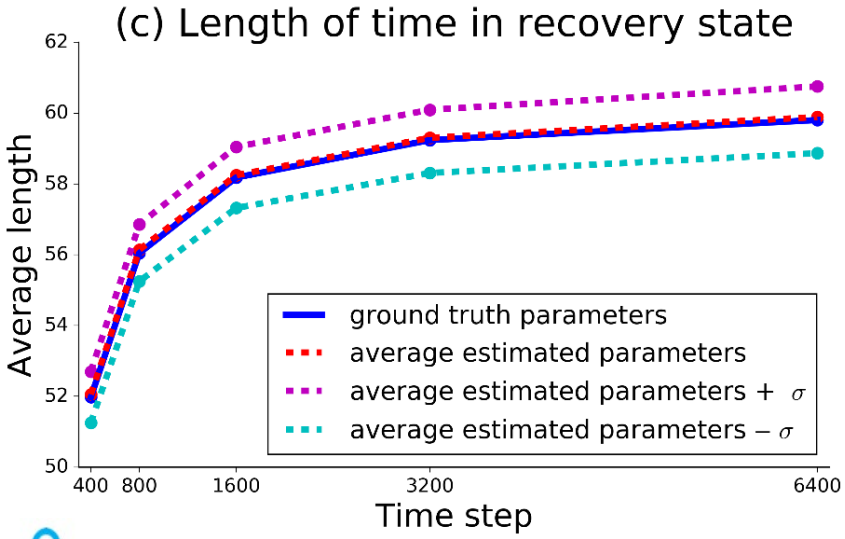
Measuring Performance of Estimated Parameters

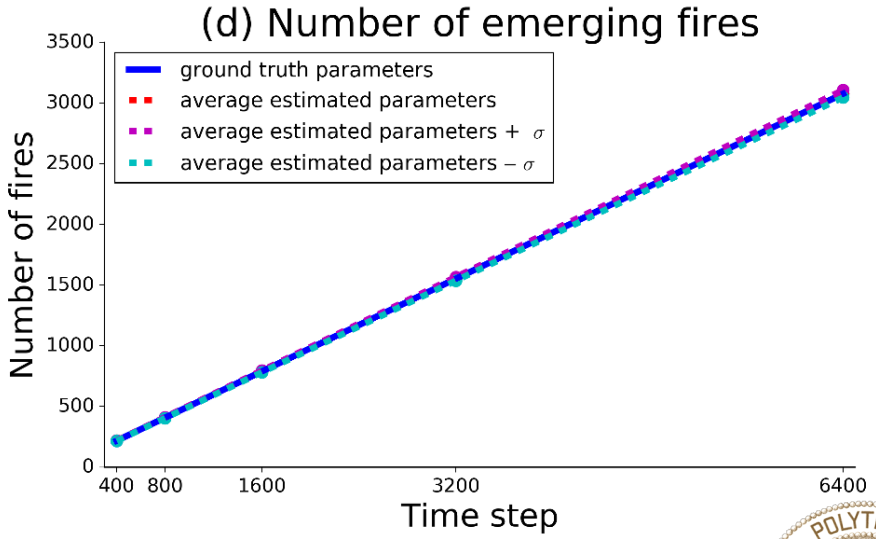
- We generate 125 sets of time series for 6400 steps using ground truth parameters: $\alpha = 0.008,\ \beta = 0.012,\ \gamma = 0.016,\ \delta = 0.032$ and compute estimated parameters from each of time series using Maximum Likelihood Estimation.
- Using sets of estimated parameters, we simulate multiple lengths of future periods: 400, 800, 1600, 3200 and 6400 and record the length of normal, on-fire and recovery state as well as the number of emerging fires during the period; results are averaged over 20 realizations.
- We compute the difference between these estimated parameters and the ground truth parameters and, determine, using Kolmogorov-Smirnov metric the $\pm\sigma$ distance between estimated parameters and ground truth parameters.
- Finally, we find the $\pm\sigma$ boundary of each set of estimated parameters by removing the largest 39 sets of results; the remaining results contain 68% of all sets of estimated parameter values, that are closest to the ground truth.

Performance of Estimated Parameters

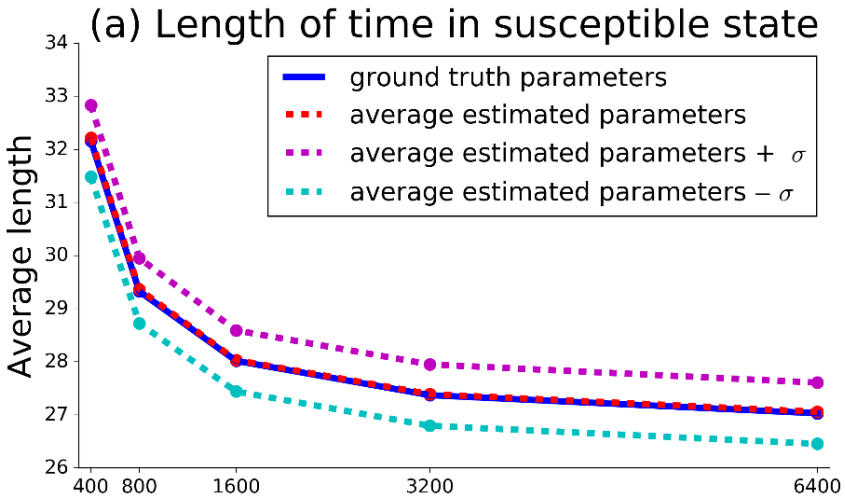


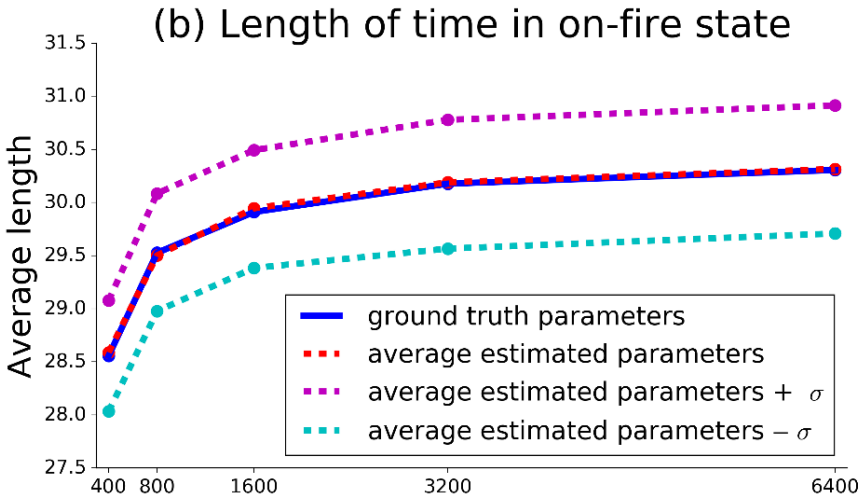


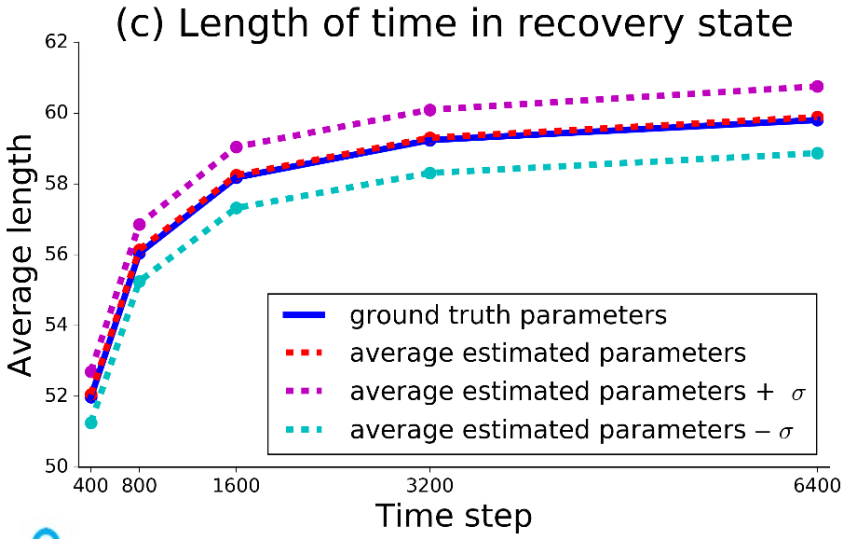


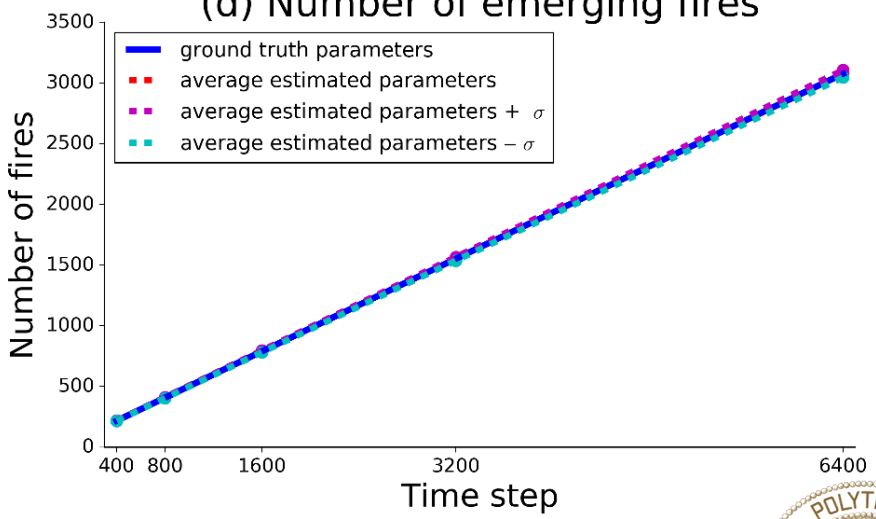


Performance of Estimated Parameters

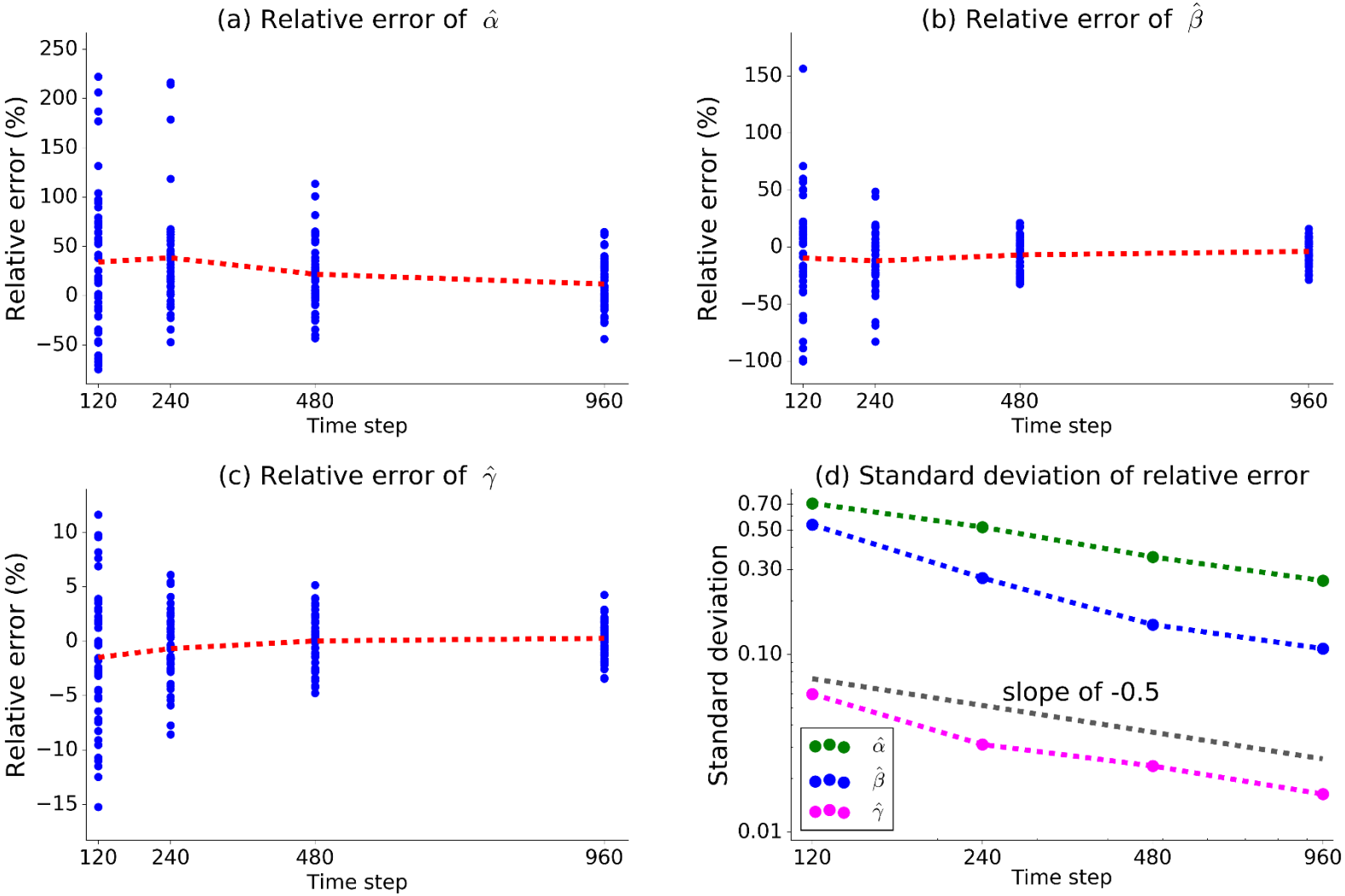






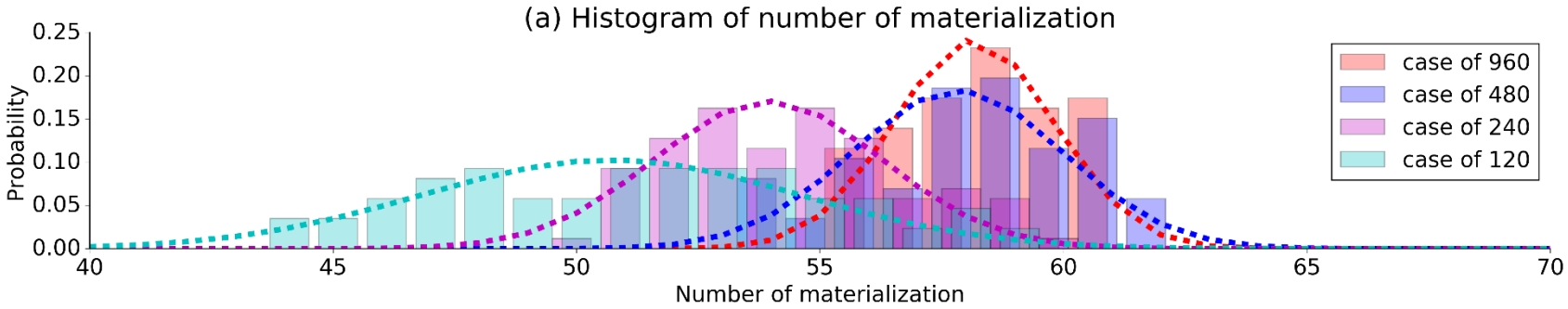


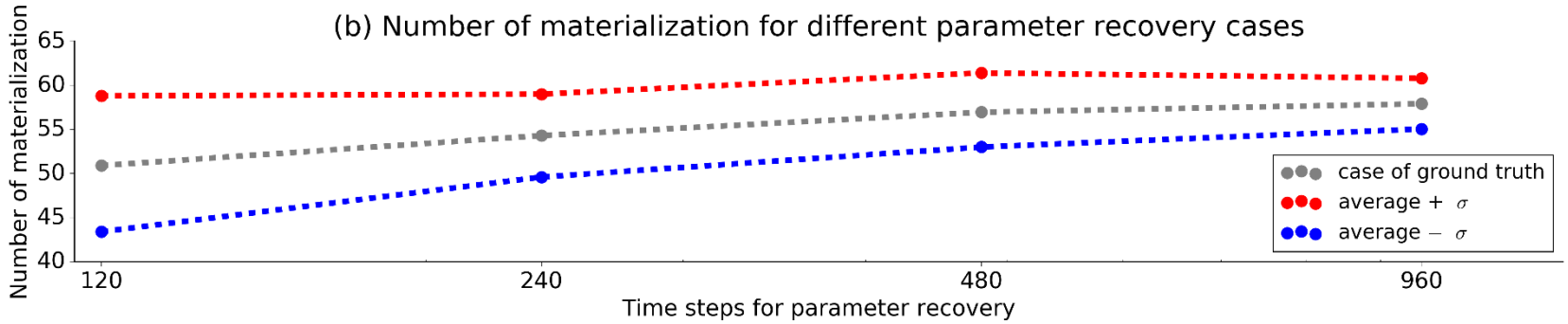
Parameter Recovery in Global Risk Network

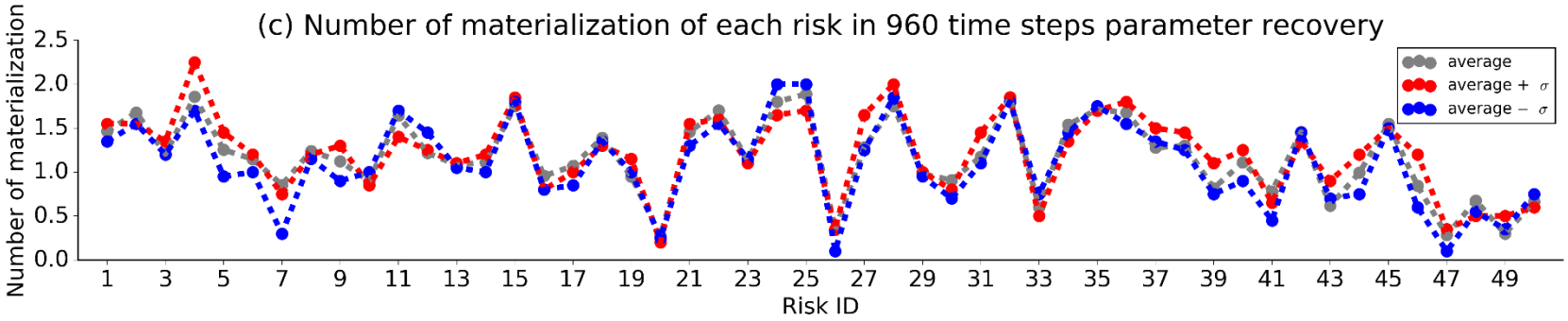


- Ground truth parameters: $\alpha = 0.364,\ \beta = 0.140,\ \gamma = 427$
- For each case, there are 50 realizations.

Parameter Recovery in Global Risk Network







- 125 sets of estimated parameters.
- Number of realizations is 20 in each scenario.

Conclusions and Future Plan

Conclusions

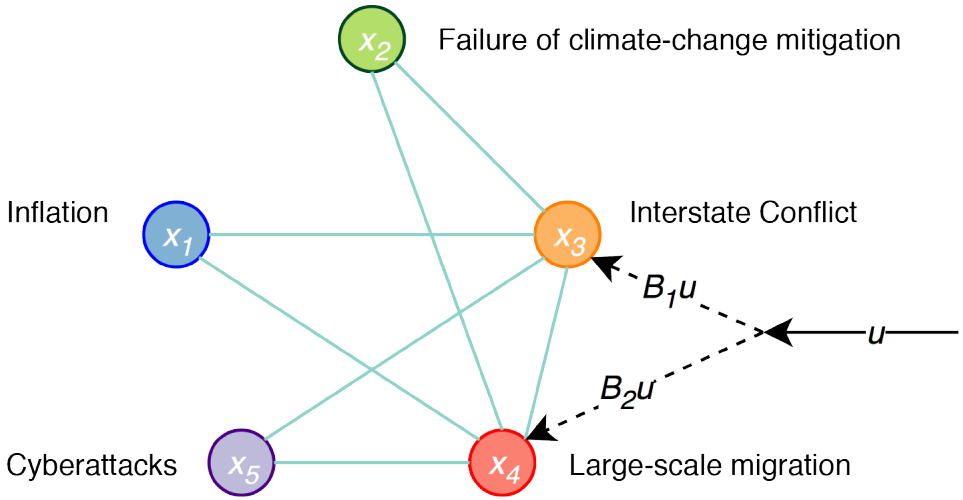
Studying the prediction limit of an interconnected network of risks using Alternative Renewal Process, we find that:

- Simulations of discrete and continuous (ODE) risk models match each other with precision improving with the increasing model connectivity.
- The parameter recovery performance improves and its error decreases when the volume of training data grows.
- **The relative error reduces asymptotically to zero with unlimited growth of training data.**

Future Plans

- Implementing parameter recovery for more complex models
- Measuring the prediction accuracy using statistical metrics
- **Studying applications combining human expert assessment with the stochastic computer predictions based on MLE recovered parameters for regional models.**

Risk control example



$$E = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}; \quad B_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}; \quad B_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}; \quad \vec{x}(k_0) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}; \quad \vec{x}(k_f) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}; \quad \vec{x}_s = \begin{bmatrix} 0.03 \\ 0.18 \\ 0.29 \\ 0.46 \\ 0.33 \end{bmatrix};$$

Nonlinear update: $\vec{x}(k+1) = F(\vec{x}(k)) + G(\vec{x}(k), E) + Bu(k);$ Linearization: $A_{ij} = \left. \frac{\partial F_i + \partial G_i}{\partial x_j} \right|_{\vec{x}=\vec{x}_s}$

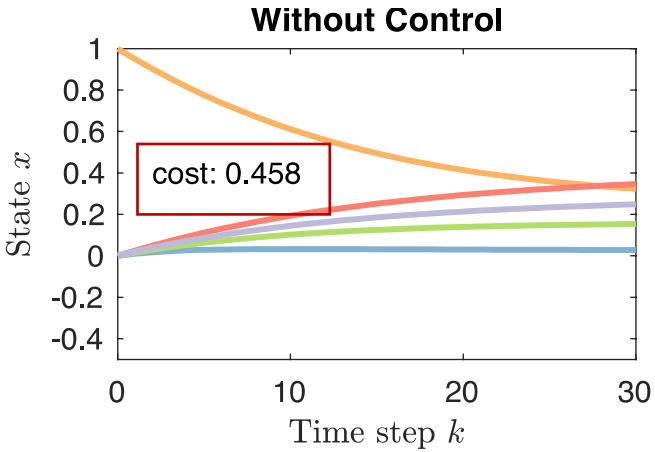
Linear update:

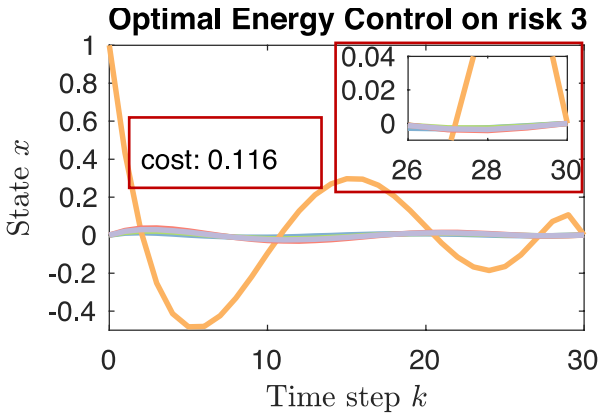
$$\Delta\vec{x}(k+1) = A\Delta\vec{x}(k) + B\vec{u}(k);$$
$$\vec{x}(k) = \Delta\vec{x}(k) + \vec{x}_s;$$

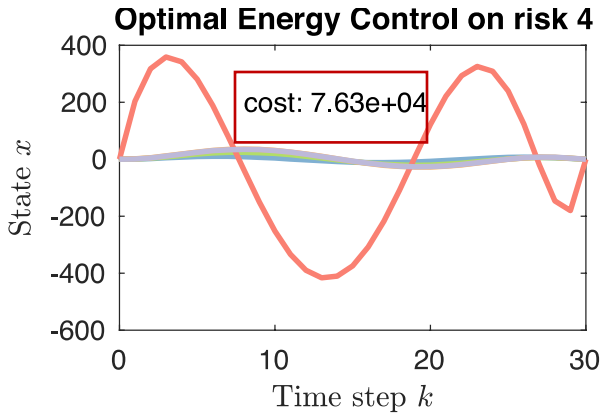
Energy: $\sum \vec{u}^T(k)\vec{u}(k);$

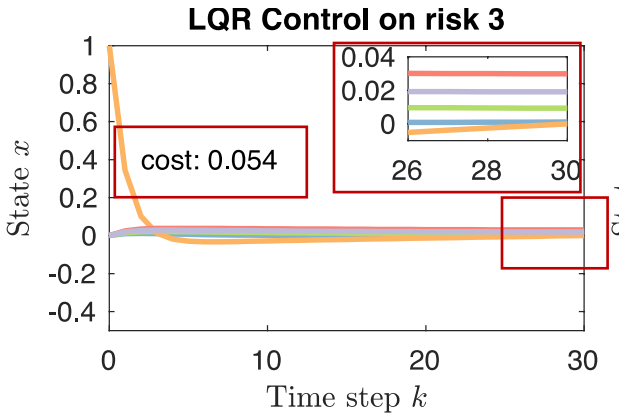
Cost: $\sum \vec{u}^T(k)\vec{u}(k) + \vec{x}^T(k)\vec{x}(k);$

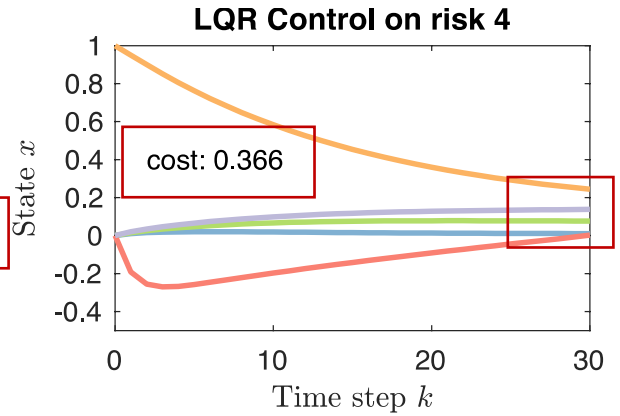
Reactive control example





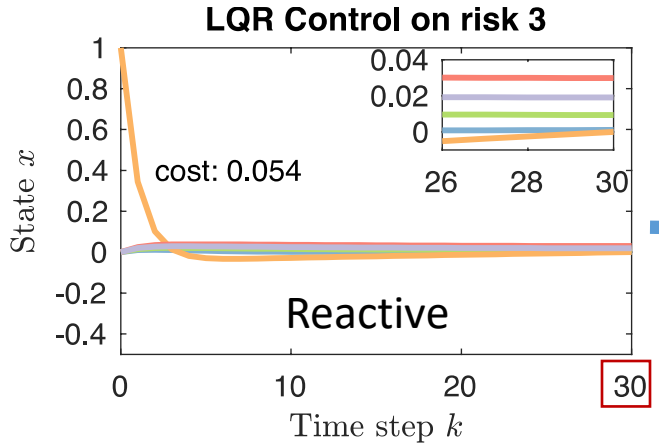


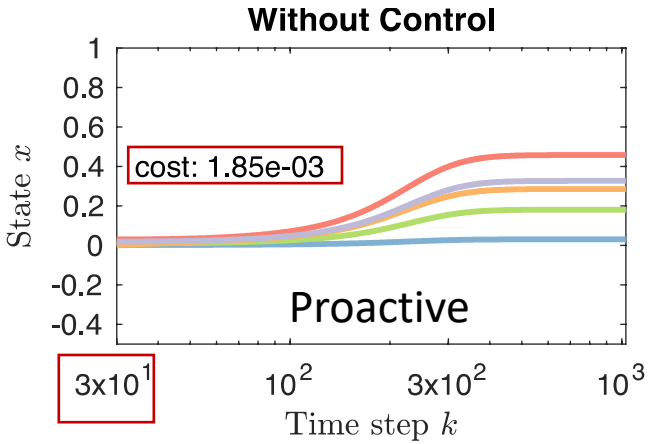


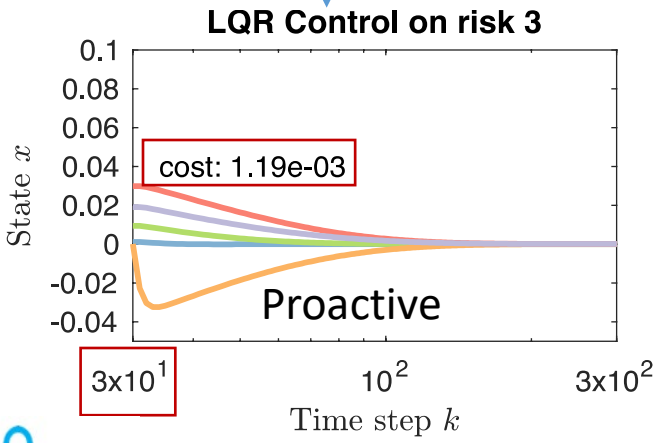


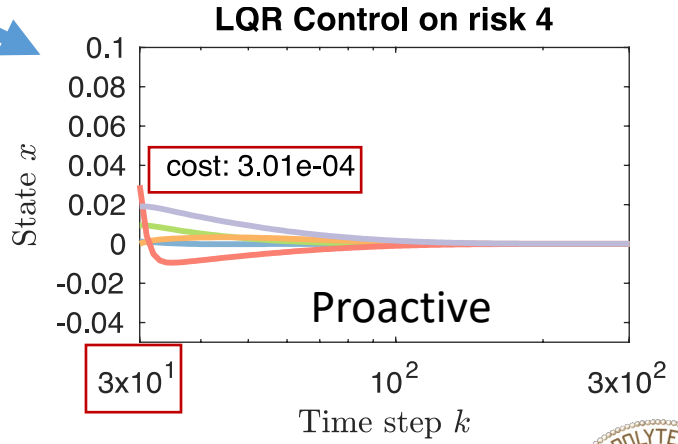
Control active risk 3 costs less than control a nonactive risk 4

Proactive control example





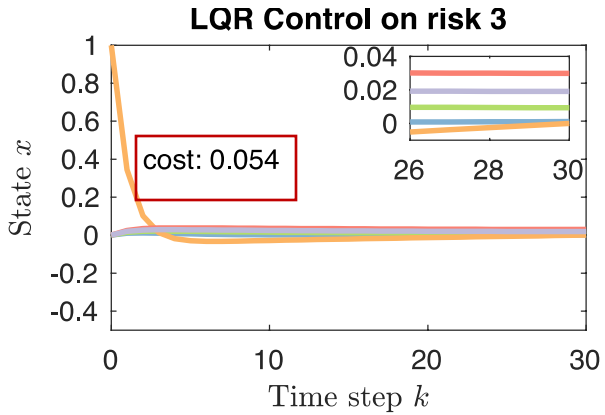


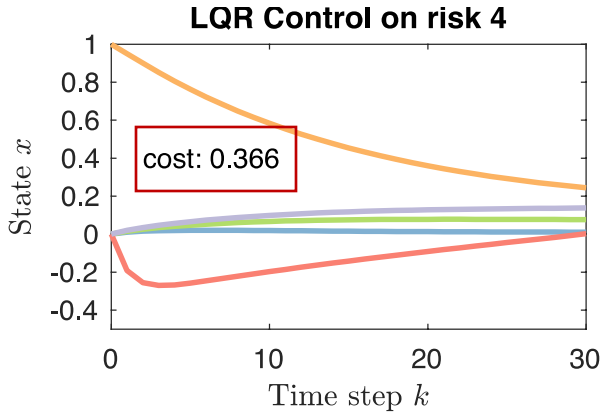


Reactive vs Proactive

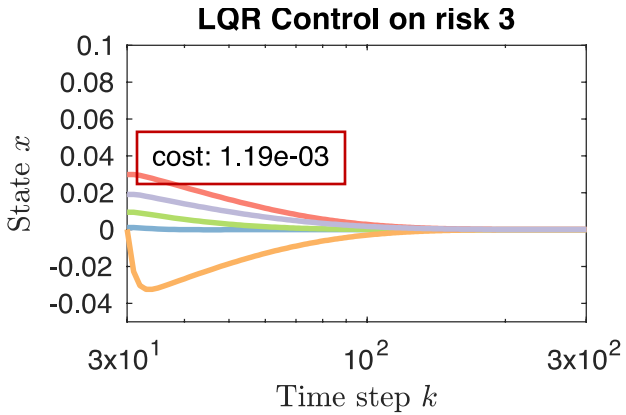
- Prevention is better than Governance in terms of cost.

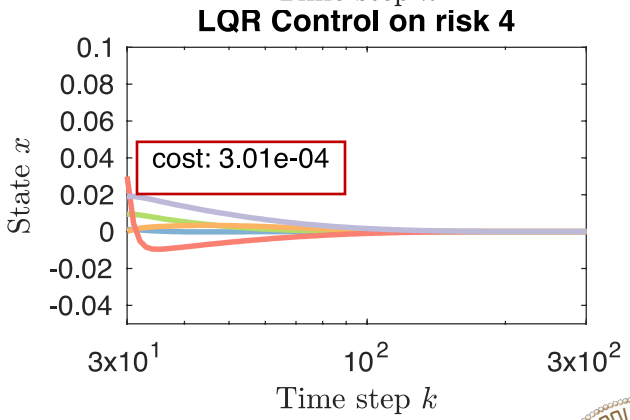
Reactive





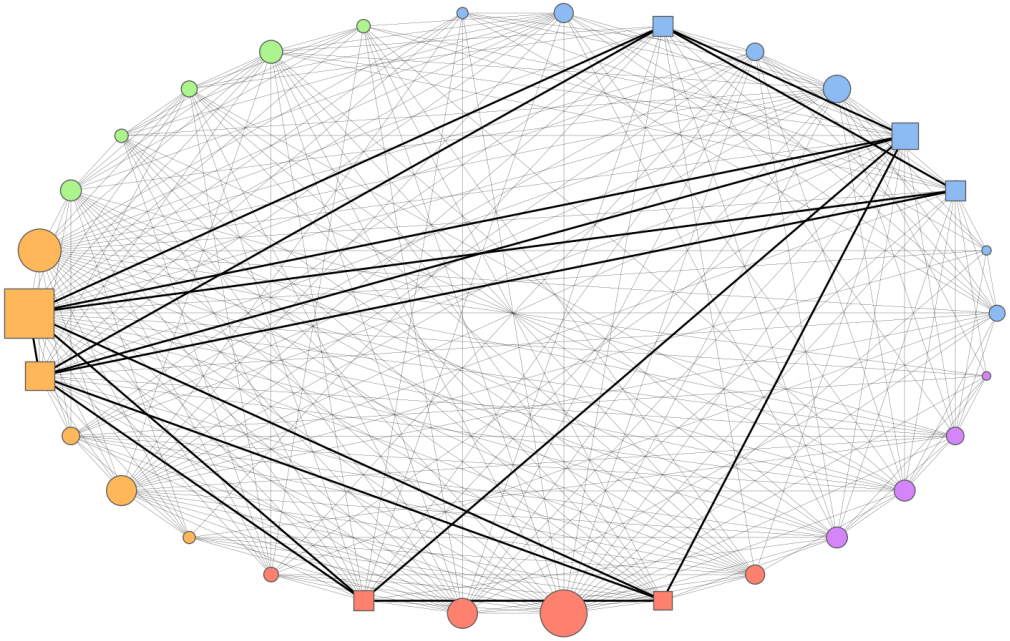
Proactive

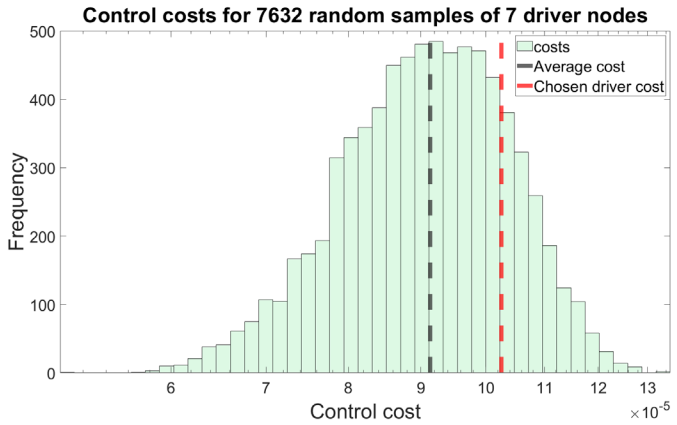


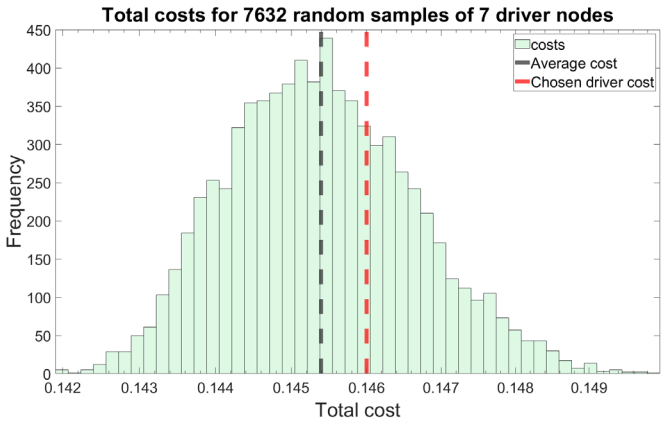


Example: COVID-19 Control

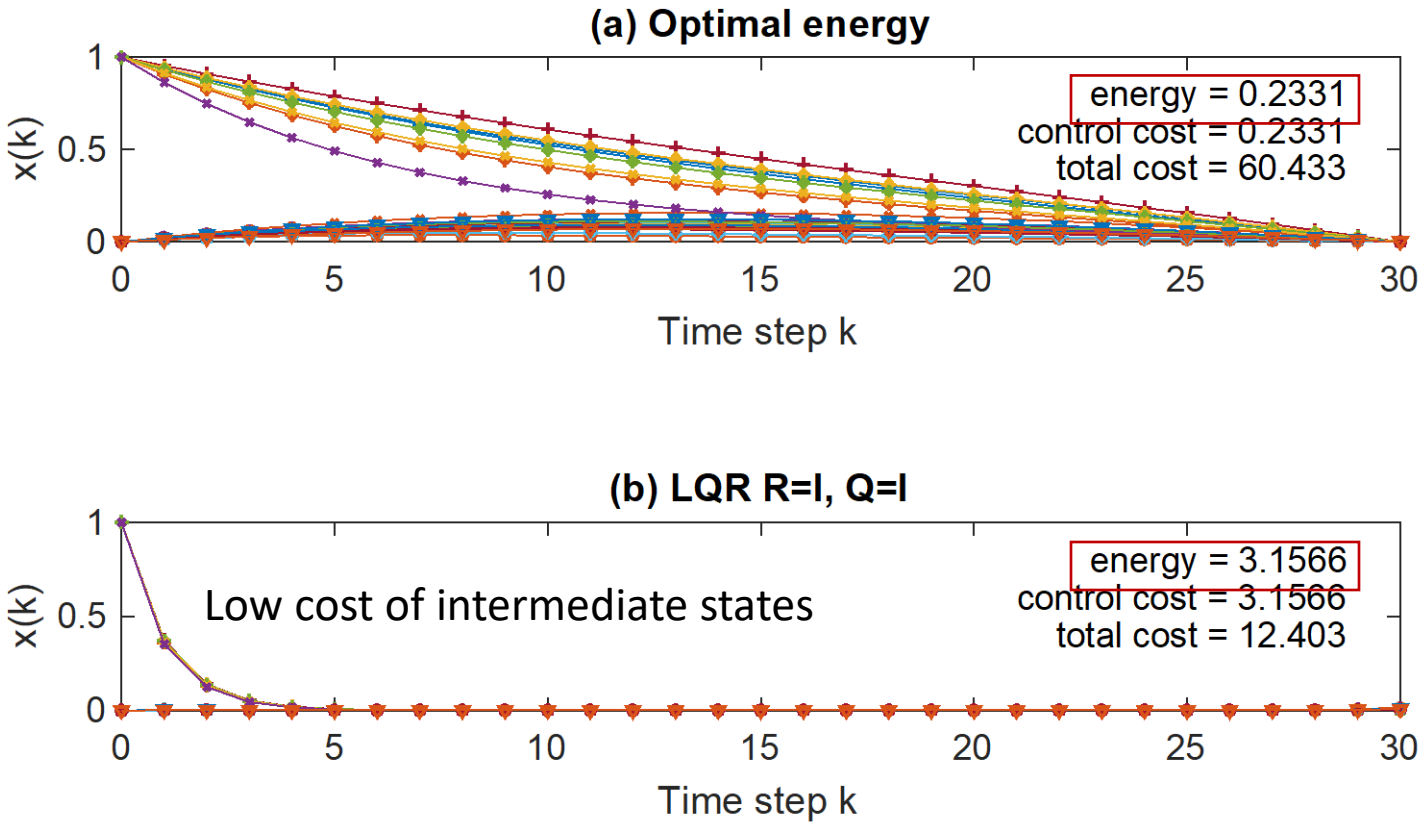
- Formally defined optimal control in the risk networks: $x(k+1) = F[x(k)] + G[x(k), E] + BU$
- Established a function of controllability index and corresponding optimal energy and conditions for nonnegative optimal control
- Provided a universal methodology of applying the LQR control in real world networked systems
- **Qualitative analysis of COVID-19 governmental policies**







Optimize functions



Contributions:

- Formally defined optimal control in the risk networks:

$$x(k+1) = F[x(k)] + G[x(k), E] + BU$$

- Established a function of controllability index and corresponding optimal energy:
  - Controllability index $\zeta = N/N_D$
  - Upper bound of optimal energy $\hat{J}_\epsilon = e^{10N/N_D}$

- Established condition for nonnegative optimal control: $N = N_D$

- Quantitively analyzed the tradeoffs between control and state costs in Reactive and Proactive phases:
  - Reactive: cost is almost linearly related to the controlled number of active risks
  - Proactive: cost is proportional to the potential risk activities
  - Prevention is better than Governance: the cost in the proactive phase is much smaller than that in the reactive phase

Contributions:

- Provided a universal methodology of applying the LQR control in real world networked systems:
  - Built a flight-delay network with five million flights record in 2015.
  - Built a delay cost matrix Q and aircraft cost matrix R according to official statistic data
- Provided significant results on flights control:
  - LQR control saves around 90% time for the customer and 70% cost for the society on average.
  - In over 5000 unique flights, almost every single one benefits from the LQR control.
- Provided significant results on airports control:
  - The small airport in the inland area benefits more than large international one in the coastal area
  - In over 300 airports, almost every single one benefits from the LQR control.
- Discovered that the airline ranking by simulated steady states in the CARP model are highly (above 0.8) correlated with Airline Quality Ranking.
- Submitted to:
  - **X. Niu**, C. Jiang, J. Gao, G. Korniss, and B. K. Szymanski. Data-driven control of networked risks with minimal cost. *Nature Communications*, 2019

Questions?