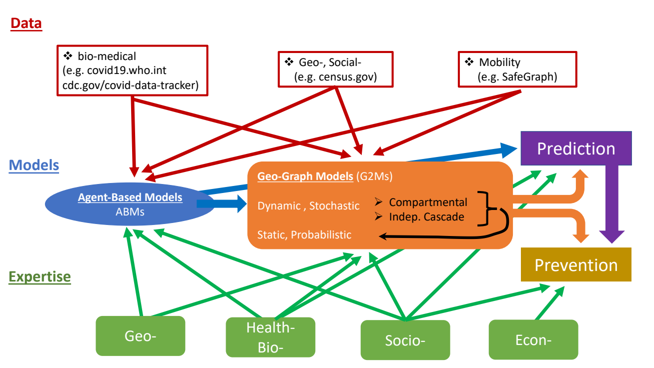
Graphical (and Agent Based) Models of Pandemic

Misha Chertkov Applied Math @ UArizona

Oct 27, 2021

Data-based Diagnosis of Networked Dynamical Systems

CCS2021 Satellite Symposium



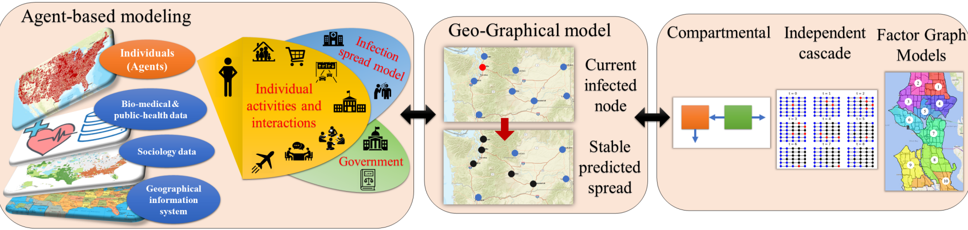
medRxiv:2021.02.24.21252390 & arXiv:2109.04517

- General Setting – Agent & **Graph** based models
  - available data
  - chain of agent- & graph-based models
  - chain of tasks: training -> **inference** -> **mitigation**

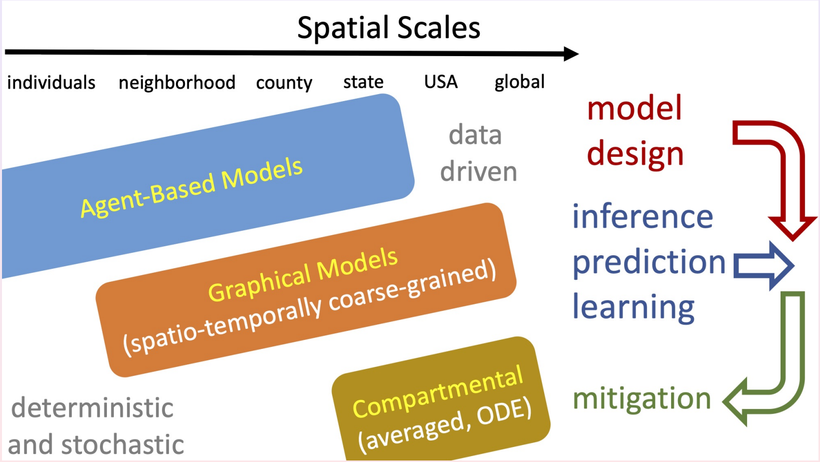
A lot of COVID-19 data ... never enough

Data Category	Sample Data Elements	Available Data Sources
Biological & Epidemiological	Cases data and number of deaths, vaccination rates, health care, loss of work, mental health	CDC COVID-19 Data Tracker ; National Hospital Care Survey ; Google Cloud Platform datasets for COVID-19 research ; The COVID Tracking Project ; The COVID-19 Data Hub on Tablauea; The Johns Hopkins COVID-19 GitHub
Geographic	Boundaries, roads, addresses	U.S. Census TIGER ; OpenStreetMap
Environmental	environmental factors	HealthData, ArcGISHub, EPA
Mobility (standard)	Home-to-work distance, commute frequency, geographic location indicators	Decennial Census ; American Community Survey ; Consumer Expenditure Survey ; SafeGraph ; Travel Survey .
Mobility (shopping, entertainment)	Frequency of travel for shopping, distance traveled, geographic location indicators	Decennial Census ; Consumer Expenditure Survey ; SafeGraph . Travel Survey .
Demographics	Age, income, education, race/ethnicity	Decennial Census ; American Community Survey ; Current Population Survey.
Employment	Employment status, work at home status, occupation	Decennial Census ; American Community Survey ; Current Population Survey ; Longitudinal Employer-Household Dynamics.
Size & Density	Population size and density of geographic units	Decennial Census
Behaviors	Social distancing, avoids indoor dining, practicing physical distancing	Safegraph ; NY Time COVID Data ; Household Pulse Survey;USC Understanding Coronavirus in America Survey.
Politics, Voting	Voting behavior and political party distribution	Current Population Survey Voting Supplement .
Workplace (Size, Density)	Distribution of number of employees per workplace; workplace density	Decennial Census,Longitudinal Employer-Household Dynamics.
School (Size, Density)	Distribution of number of students per school; average class size	Decennial Census ; Current Population Survey; American Community Survey ; National Center for Education Statistics Information System .

Hierarchy of Models



Hierarchy of Models

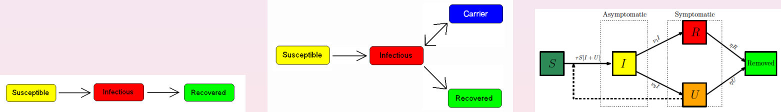


Compartmental models (no spatial extent)

Susceptible $\rightarrow$ Infectious $\rightarrow$ Recovered

$$\frac{dS}{dt} = -\frac{\beta S I}{N}$$
$$\frac{dI}{dt} = \frac{\beta S I}{N} - \gamma I$$
$$\frac{dR}{dt} = \nu I$$

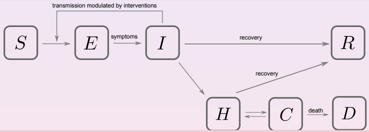
- N - population size
- β - rate of S $\rightarrow$ I transition
- γ - rate of I $\rightarrow$ R transition



Exemplary Compartmental Models (a,b) from wiki, (c) from Liu, etal (2020)

Compartmental Model of COVID-19 scenarios

- Sub-populations are indexed a, b
- Parameters **should be learned from data**
- $R_0 = 2.5$ - do nothing, 0.35 (or even low) - today; ξ_a - degree of isolation: control parameters



- Susceptible → Exposed → Infected → Hospitalized → Critical → Deceased
- Sub-populations are indexed a, b
- $R \approx 2.5$ - do nothing, 0.35 - today; ξ_a - degree of isolation: **control** parameters
- other parameters should be **learned from data** parameters

$$\frac{dS_a}{dt} = -\frac{\beta_a S_a \sum_b I_b}{N}$$

$$\frac{dE_a}{dt} = \frac{\beta_a S_a \sum_b I_b}{N} - \frac{E_a}{t_l}$$

$$\frac{dI_a}{dt} = \frac{E_a}{t_l} - \frac{I_a}{t_i}$$

$$\frac{dH_a}{dt} = \frac{(1-m_a)I_a}{t_i} + \frac{(1-f_a)C_a}{t_c} - \frac{H_a}{t_h}$$

$$\frac{dC_a}{dt} = \frac{c_a H_a}{t_h} - \frac{C_a}{t_c}$$

$$\frac{dR_a}{dt} = \frac{m_a I_a}{t_i} + \frac{(1-c_a)H_a}{t_h}$$

$$\frac{dD_a}{dt} = \frac{f_a C_a}{t_c}$$

$$\beta_a(t) = R_0 \xi_a \left(1 + \varepsilon \cos(2\pi(t - t_{max}))\right)/t_i$$

$R = R_0 S/N$ - reproduction rate

- (# of contacts when infectious)*(prob. of transmission per contact)

What is GOOD about compartmental models?

- Based on basic principles of "decease"-interaction (think “chemical" reaction)
- Working horse of epidemiology - all credible predictions and analysis have these models under the hood
- Allows some analytical analysis
  - early exponential growth
  - fixed (balance) points
- Simple ...

History

- First compartmental ordinary differential equation model for epidemiology published by Ronald Ross in 1911 for malaria (in India)
- More general overview of mass-action models for epidemiology published byKermack and McKendrick in 1927
- Substantial increase in use after reviews published by Anderson and May in 1991 and Hethcote in 2000

What is WRONG with compartmental models?

- Modeling time-varying infectivity **delays**
- Constructing realistic models with **spatial resolution**
- Defining limitations of deterministic models
  - finite size **fluctuations**
  - exogenous **fluctuations** (spontaneous hot spots)
- **Data** is fed through hand-adjustment of a small # of parameters, based on historical data (often for another decease, from another place, ... **20 years old science**)

Geography-Resolving Compartmental Models: Dynamical System Approach

Network Models (with spatial resolution)

- $\dot{S}_i = -\sum_j a_{ij} S_i I_j,\ \dot{I}_i = \sum_j a_{ij} S_i I_j - \gamma_i I_i,\ R_i = \gamma_i I_i$
  - Lajmanovich and Yorke (1976)
  - Meia, Mohagheghi, Zampieri and Bullo (2017)
  - Ma, Liu, and Olshevsky (2020)
- Can also include stochasticity
  - $\dot{S}_i = -\sum_j a_{ij} S_i I_j + \alpha_1 dW_1,\ \dot{I}_i = \sum_j a_{ij} S_i I_j - \gamma_i I_i - \alpha_1 dW_1 - \alpha_2 dW_2,\ R_i = \gamma_i I_i + \alpha_2 dW_2$
  - ... yet to be analyzed

Memory Effects (age of infection)

- Data Driven Modeling COVID-19 Dynamics in Illinois under Nonpharmaceutical Interventions
- link (Wang et al - 2020)
- No spatial extent – Illinois is aggregated in one node
- Modern version of Kermack-McKendrick (1927) age-of-infection model. Accounts for memory.
- Stochasticity is in data, i.e. in noisy observations - not in the model itself

Agent Based Models (ABM)

- Geographical Map
- Individuals/agents, households are resolved
- Age, job, ethnicity are accounted for
- Input: current, $t = 0$, status, rules for agents
- Output: spatio-temporal sample of the agents status

Agent vs Individual – differences are subtle

- (Imperial/Pitt) – agents don't move, interaction via spatial kernel
- UW/LANL - individuals move, interaction is local
- UVirginia — individuals move - based on samples of real people

- link MIDAS: Online Portal for COVID-19 Modeling Research

UW/LANL ABM (the only pre-pandemic open source)

- FLUTE – C/C++, GNU public license, Open MPI
  `https://www.epimodels.org/midas/flute.do`
- creates synthetic population
- hierarchy of "mixing" groups – household, neighbors, schools, kindergarden, employment, etc
- two time steps a day, infectious for 6 days (influenza)
- travel (long-distance, short term)
- infection, contact probability (the same site), R_0-related
- simulated different types of interventions [control - site-dependent rate of infection, isolation, etc]
- entire US, 280 mln individuals, 60-90 days

Many open-source codes became available since 03/2020

Paper	Date	Interventions	Has geography	Realistic geography
Covasim	July 2021	Masking, Quarantine, Vaccination, and Contact tracing	No	No
OpenABM-Covid19	July 2021	Masking, Quarantine, Vaccination, lockdown, and Contact tracing	Yes	No
DESSABNeT	May 2021	Quarantine and Masking	No	No
	February 2021	Quarantine and Masking	No	No
COVID-ABS	October 2020	Quarantine and Masking	No	No
COVI-AgentSim	October 2020	Quarantine, Masking, and Contact tracing	Yes	No
FACS	July 2020	Quarantine and Masking	Yes	Yes
INFEKTA	April 2020	Quarantine and lockdown	Yes	Yes

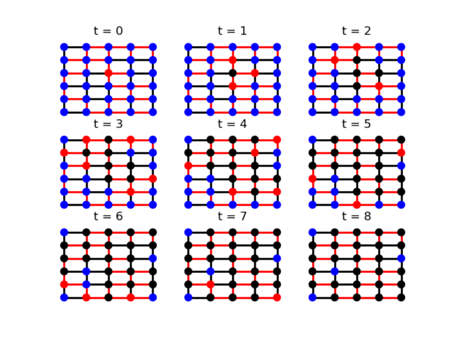
General Features of Agent Based Models

- Working Horse of Epidemiology
- Highly (may be too) detailed $\Leftarrow$ Public Health Expertise
- Very Heavy (HPC) – cannot explore many options, scenarios, mitigation strategies
- Provide Excellent (quasi-realistic) ground truth headway for REDUCED MODELS

What? Why and How? of Graphical Models of Pandemic

- Reduced, Macroscopic, Probabilistic vs High-Fidelity Microscopic = Agent Based Models (ABM)
- Efficient for Making Probabilistic Predictions (Inference)
- Dynamic (cascades transition probability) vs Static (statistically quasi-steady, derived from dynamic)
- Data-driven – set up based on the on-going extraction of data from open source resources, such as pandemic data repositories, U.S. Census, and Geographical Information System (GIS) sources, mobility data, etc.
  - Input: graph + pair-wise rates of transmission + initial infection pattern (e.g. super-spreader exogeneous event)
  - Exemplary Output: Heat map of marginal probabilities for neighborhoods to be (largely) infected in two weeks

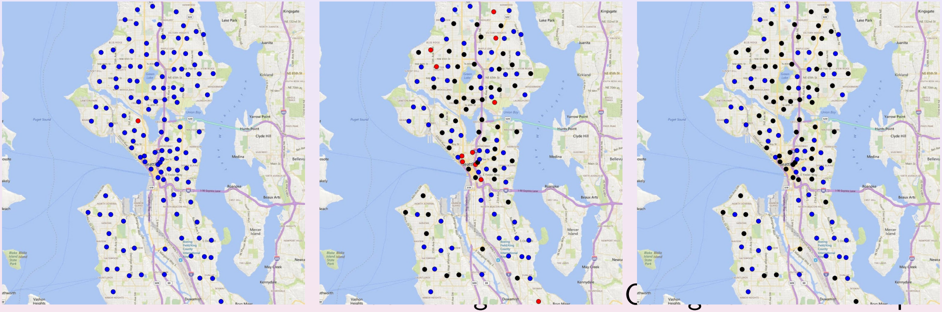
Independent Cascade (dynamic) model – our adaptation



- Edges are activated according $\Leftarrow$ transition probabilities
- A node is (I)nfectious for one time step only, then it becomes (R)emoved (black).
- Final (R)emoved pattern = connected component of the graph formed by initially (I)nfected node.

- (Kempe et al 2003) – network of "influences"
  - "Maximizing the Spread of Influence through a Social Network", 7110 citations
  - sub-modularity, convexity, bounds
- Was not used as a reduced dynamic GM of epidemiology

Independent Cascade for Seattle



Sources of Randomness

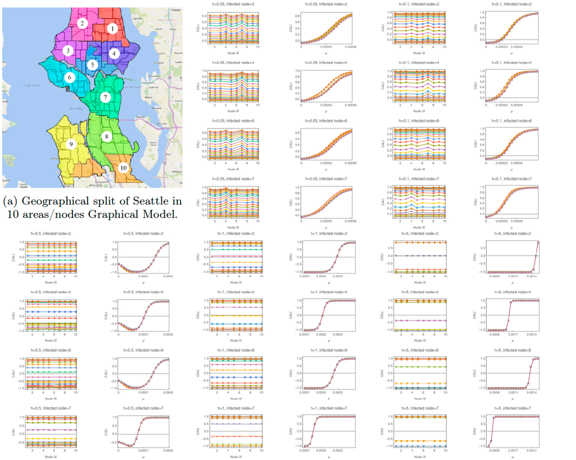
- Initial Seed
- Sample of activated edges (according to transition probabilities)

Graphical (Ising) Model of Pandemic

Final state of the Independent Cascade Model $\Leftarrow$ Ising model

- $$P(\boldsymbol{x}) = Z^{-1}\exp\left(\sum_{\{a,b\}\in\mathcal{E}} x_a J_{ab} x_b + \sum_{a\in\mathcal{V}} x_a H_a\right)$$
  $$Z \doteq \sum_{\boldsymbol{x}} \exp\left(\sum_{\{a,b\}\in\mathcal{E}} x_a J_{ab} x_b + \sum_{a\in\mathcal{V}} x_a H_a\right)$$
- $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ - graph of pair-wise links (edges) between neighborhoods (nodes)
- $\boldsymbol{J} \doteq (J_{ab} \doteq (\log(1+e^{g_{ab}}))/2 | \forall\{a,b\}\in\mathcal{E})$ – vector of the effective pair-wise connectivity of communities
- $\boldsymbol{H} = (H_a \doteq -(\log((1-x_a^{(in)})/2) + \beta)/2 | \forall a \in \mathcal{V})$ – vector of the single-community exogenous infection bias

CALI: Conditioned A-Posteriori Level of Infection



(a) Geographical split of Seattle in 10 areas/nodes Graphical Model.

FIG. 5: CALI evaluated (exact inference) over the 10 nodes GM of Seattle. Each pair of subfigures (CALI vs node id and CALI vs μ) corresponds to one set of experiments for fixed (and uniform) bias h and showing results for different values of the interaction factor and different initial seed of infection (infected nodes). See discussions in the text.

$$P(\boldsymbol{x}|\mathscr{G},\boldsymbol{J},\boldsymbol{h}) = \frac{\exp\left(-E(\boldsymbol{x}|\mathscr{G},\boldsymbol{J},\boldsymbol{h})\right)}{Z(\mathscr{G},\boldsymbol{J},\boldsymbol{h})},$$

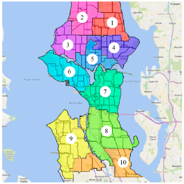
$$E(\boldsymbol{x}|\mathscr{G},\boldsymbol{J},\boldsymbol{h}) := \sum_{a\in\mathscr{V}} h_a x_a - \sum_{\{a,b\}\in\mathscr{E}} J_{ab} x_a x_b,$$

$$Z(\mathscr{G},\boldsymbol{J},\boldsymbol{h}) := \sum_{\boldsymbol{x}} \exp\left(-\sum_{a\in\mathscr{V}} h_a x_a + \sum_{\{a,b\}\in\mathscr{E}} J_{ab} x_a x_b\right),$$

$$\forall a \in \mathscr{V} \setminus \mathscr{V}^{(in)}: \quad m_a(\boldsymbol{x}^{(in)}) := \mathbb{E}\left[x_a|\boldsymbol{x}_{in}\right]$$

- Data Driven
  - census data
  - mobility data
  - CDC data
- Levels of granularity
  - City – Seattle
  - State – Wisconsin
  - Region – SouthWest

CALI: Conditioned A-Posteriori Level of Infection



(a) Geographical split of Seattle in 10 areas/nodes Graphical Model.

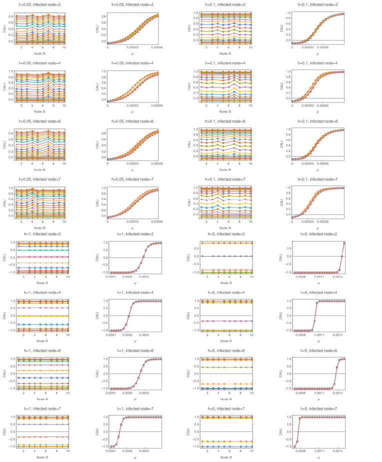


FIG. 5: CALI evaluated (exact inference) over the 10 nodes GM of Seattle. Each pair of subfigures (CALI vs node id and CALI vs μ) corresponds to one set of experiments for fixed (and uniform) bias h and showing results for different values of the interaction factor and different initial seed of infection (infected nodes). See discussions in the text.

$$P(\boldsymbol{x}|\mathcal{G},\boldsymbol{J},\boldsymbol{h}) = \frac{\exp\left(-E(\boldsymbol{x}|\mathcal{G},\boldsymbol{J},\boldsymbol{h})\right)}{Z(\mathcal{G},\boldsymbol{J},\boldsymbol{h})},$$

$$E(\boldsymbol{x}|\mathcal{G},\boldsymbol{J},\boldsymbol{h}) := \sum_{a\in\mathcal{V}} h_a x_a - \sum_{\{a,b\}\in\mathcal{E}} J_{ab} x_a x_b,$$

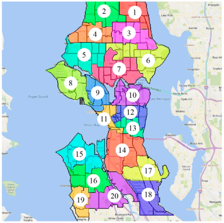
$$Z(\mathcal{G},\boldsymbol{J},\boldsymbol{h}) := \sum_{\boldsymbol{x}} \exp\left(-\sum_{a\in\mathcal{V}} h_a x_a + \sum_{\{a,b\}\in\mathcal{E}} J_{ab} x_a x_b\right),$$

$$\forall a \in \mathcal{V}\setminus\mathcal{V}^{(in)}: \quad m_a(\boldsymbol{x}^{(in)}) := \mathbb{E}\left[x_a|\boldsymbol{x}_{in}\right]$$

Setting:

- inference is difficult
- dependence on μ -scaling of interaction
- dependence h - community bias
- different inference methods
- calibration on 10- & 20-node model

CALI: Conditioned A-Posteriori Level of Infection



(a) Geographical split of Seattle in 20 areas/nodes Graphical Model.

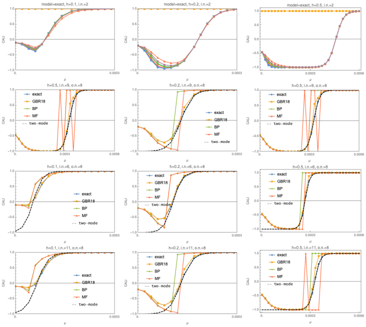


FIG. 6: CALI evaluated over the 20 nodes GM of Seattle, with the split (into nodes) shown on the left. Each column of the sub-plots corresponds to three different values of local bias (from smaller on the left to larger on the right). Top row shows CALI at all the observed nodes (o.n.) for the case when the initial infection is seeded at the node #2 calculated exactly (model=exact). Three bottom rows show CALI, computed with different methods (exact, GBR18,BP and MF – see discussion in the text), at the observed node #8, local bias correspondent to the respective column and injected at various nodes (#9, #6 and #11 at the second, third and fourth rows, respectively).

Results:

- sharp transition (nobody -> all)
- strong sensitivity to seed-location
- two-mode approximation – simple & efficient

CALI [Conditioned A-Posteriori Level of Infection] Index

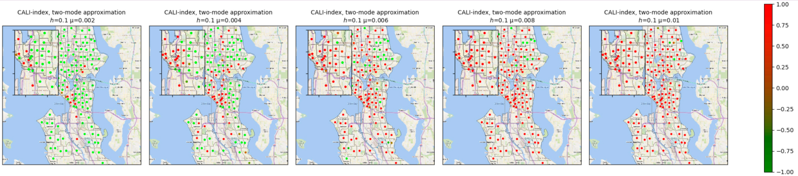


FIG. 7: CALI-index, computed within the two-mode approximation, is shown for the full (123-nodes) model of Seattle. It shows dependence of the overall (integrated over entire city) level of infection dependent on the position on the initial seed. Sub-figure on the left, correspondent to reduced pair-wise interaction factor, μ, shows that the infection spread succeed only if seeded at a few location visited often by non-residents. Increase in μ (sub-figures on the right) results in a much stronger spread of infection. See explanations in the text for further details.

- CALI Index – collective response
- Two-mode approximation (full-scale)

$$M(\boldsymbol{x}^{(in)}) := \sum_{a\in\mathscr{V}} \rho_a m_a(\boldsymbol{x}^{(in)})$$

- sharp transition (nobody -> all)
- strong sensitivity to seed-location
- methodology works ! $\Rightarrow$

CALI [Conditioned A-Posteriori Level of Infection] Index

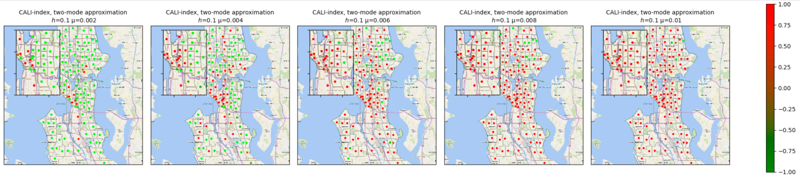


FIG. 7: CALI-index, computed within the two-mode approximation, is shown for the full (123-nodes) model of Seattle. It shows dependence of the overall (integrated over entire city) level of infection dependent on the position on the initial seed. Sub-figure on the left, correspondent to reduced pair-wise interaction factor, μ, shows that the infection spread succeed only if seeded at a few location visited often by non-residents. Increase in μ (sub-figures on the right) results in a much stronger spread of infection. See explanations in the text for further details.

- CALI Index – collective response
- Two-mode approximation (full-scale)

$$M(\boldsymbol{x}^{(in)}) := \sum_{a\in\mathscr{V}} \rho_a m_a(\boldsymbol{x}^{(in)})$$

- sharp transition (nobody -> all)
- strong sensitivity to seed-location
- methodology works ! $\Rightarrow$

Next Step

- Mitigation/Prevention

Prediction and Prevention of Pandemics

... via Graphical Model Inference and Convex Programming arXiv:2109.04517

Final state of the Independent Cascade Model $\Leftarrow$ Ising model

- $P(\boldsymbol{x}) = Z^{-1} \exp\left(-E(\boldsymbol{x}|\boldsymbol{J},\boldsymbol{h})\right)$
- $E(\boldsymbol{x}|\boldsymbol{J},\boldsymbol{h}) = -\sum\limits_{\{a,b\}\in\mathcal{E}} x_a J_{ab} x_b - \sum\limits_{a\in\mathcal{V}} x_a H_a$
- $Z \doteq \sum\limits_{\boldsymbol{x}} \exp\left(\sum\limits_{\{a,b\}\in\mathcal{E}} x_a J_{ab} x_b + \sum\limits_{a\in\mathcal{V}} x_a H_a\right)$
- $\mathcal{G} = (\mathcal{V},\mathcal{E})$ - graph of pair-wise links (edges) between neighborhoods (nodes)
- $\boldsymbol{J} \doteq (J_{ab} \doteq (\log(1+e^{g_{ab}}))/2|\forall\{a,b\}\in\mathcal{E})$ – vector of the effective pair-wise connectivity of communities
- $\boldsymbol{H} = (H_a \doteq -(\log((1-x_a^{(in)})/2)+\beta)/2|\forall a\in\mathcal{V})$ – vector of the single-community exogenous infection bias

Prediction and Prevention of Pandemics

... via Graphical Model Inference and Convex Programming arXiv:2109.04517

Final state of the Independent Cascade Model $\Leftarrow$ Ising model

- $P(\boldsymbol{x}) = Z^{-1} \exp\left(-E(\boldsymbol{x}|\boldsymbol{J}, \boldsymbol{h})\right)$
- $E(\boldsymbol{x}|\boldsymbol{J}, \boldsymbol{h}) = -\sum_{\{a,b\}\in\mathcal{E}} x_a J_{ab} x_b - \sum_{a\in\mathcal{V}} x_a H_a$
- **MAP-inference challenge:** Exact Easy (Max-LP), Approximate Super-Easy (two-mode)
  - $x^{(MAP)}(\mathcal{I}|\boldsymbol{J}, \boldsymbol{h}) = \arg\min_{\boldsymbol{x}} E(\boldsymbol{x}|\boldsymbol{J}, \boldsymbol{h})\Big|_{\forall a\in\mathcal{I}:\ x_a=+1}$
  - $\mathcal{R}(\mathcal{I}|\boldsymbol{J}, \boldsymbol{h}) = \{a \in \mathcal{V} | x^{(MAP)}(\mathcal{I}|\boldsymbol{J}, \boldsymbol{h}) = +1\}$
- If $\mathcal{R}$ is too large $\Rightarrow$ need to mitigate/prevent

Prediction and Prevention of Pandemics

... via Graphical Model Inference and Convex Programming arXiv:2109.04517

Final state of the Independent Cascade Model ⇐ Ising model

- $P(\boldsymbol{x}) = Z^{-1}\exp\left(-E(\boldsymbol{x}|\boldsymbol{J},\boldsymbol{h})\right)$
- $E(\boldsymbol{x}|\boldsymbol{J},\boldsymbol{h}) = -\sum_{\{a,b\}\in\mathcal{E}} x_a J_{ab} x_b - \sum_{a\in\mathcal{V}} x_a H_a$
- **MAP-inference challenge:**
  - $x^{(MAP)}(\mathcal{I}|\boldsymbol{J},\boldsymbol{h}) = \arg\min_{\boldsymbol{x}} E(\boldsymbol{x}|\boldsymbol{J},\boldsymbol{h})\Big|_{\forall a\in\mathcal{I}:\ x_a=+1}$
  - $\mathcal{R}(\mathcal{I}|\boldsymbol{J},\boldsymbol{h}) = \{a\in\mathcal{V}|x^{(MAP)}(\mathcal{I}|\boldsymbol{J},\boldsymbol{h}) = +1\}$
- If $\mathcal{R}$ is too large ⇒ need to mitigate/prevent
- **Prevention challenge**:
  - $(\boldsymbol{J}^{(corr)},\boldsymbol{h}^{(corr)}) = \arg\min_{(\boldsymbol{J},\boldsymbol{h})} C(\boldsymbol{J},\boldsymbol{h}),(\boldsymbol{J}^{(corr)},\boldsymbol{h}^{(corr)}))\Big|_{\forall\mathcal{I}\in\mathcal{Y}:\ |\mathcal{R}(\mathcal{I}|\boldsymbol{J},\boldsymbol{h})|\leq k}$

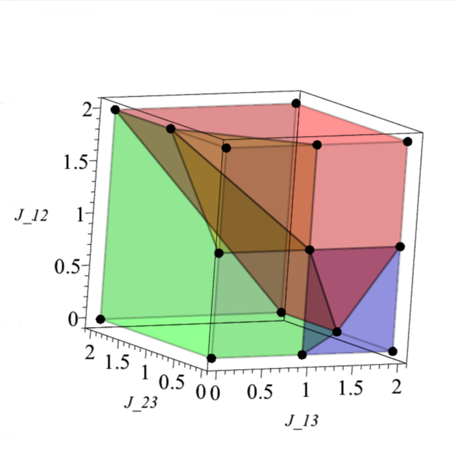
Prediction and Prevention of Pandemics

... via Graphical Model Inference and Convex Programming arXiv:2109.04517

Final state of the Independent Cascade Model $\Leftarrow$ Ising model

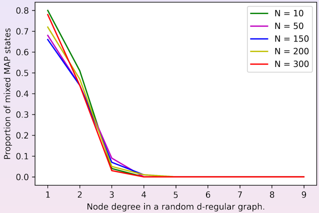
- $P(\boldsymbol{x}) = Z^{-1} \exp\left(-E(\boldsymbol{x}|\boldsymbol{J}, \boldsymbol{h})\right)$
- $E(\boldsymbol{x}|\boldsymbol{J}, \boldsymbol{h}) = -\sum_{\{a,b\}\in\mathcal{E}} x_a J_{ab} x_b - \sum_{a\in\mathcal{V}} x_a H_a$
- **MAP-inference challenge:**
  - $x^{(MAP)}(\mathcal{I}|\boldsymbol{J}, \boldsymbol{h}) = \arg\min_{\boldsymbol{x}} E(\boldsymbol{x}|\boldsymbol{J}, \boldsymbol{h})\Big|_{\forall a\in\mathcal{I}:\ x_a=+1}$
  - $\mathcal{R}(\mathcal{I}|\boldsymbol{J}, \boldsymbol{h}) = \{a \in \mathcal{V} | x^{(MAP)}(\mathcal{I}|\boldsymbol{J}, \boldsymbol{h}) = +1\}$
- If $\mathcal{R}$ is too large $\Rightarrow$ need to mitigate/prevent
- **Prevention challenge** = Projection to Feasible Space (Polytope):
  - $(\boldsymbol{J}^{(corr)}, \boldsymbol{h}^{(corr)}) = \arg\min_{(\boldsymbol{J},\boldsymbol{h})} C(\boldsymbol{J}, \boldsymbol{h}), (\boldsymbol{J}^{(corr)}, \boldsymbol{h}^{(corr)}))\Big|_{\forall \mathcal{I}\in\mathcal{Y}:\ |\mathcal{R}(\mathcal{I}|\boldsymbol{J},\boldsymbol{h})|\leq k}$

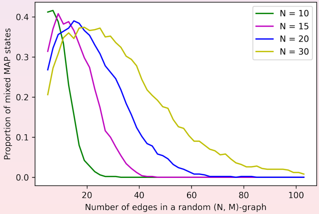
Two-Mode Approximation is Justified (for MAP)



- The case of triangle
- 2 pure + 2 mixed states

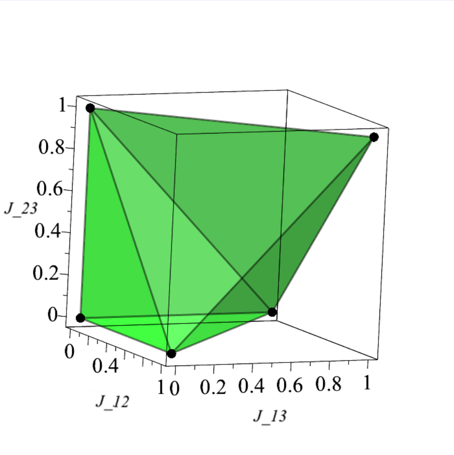
Two-Mode Approximation is Justified (for MAP)





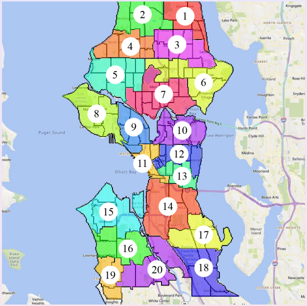
- Random Graphs $\Rightarrow$ mixed states are rare for large, dense graphs
- Urban Graphs (e.g. Seattle) are Dense
- Rural Graphs are NOT as Dense (show mixed states)

Prevention Challenge = Projection to Feasible Space



- Feasible Space = Polytope (illustration)

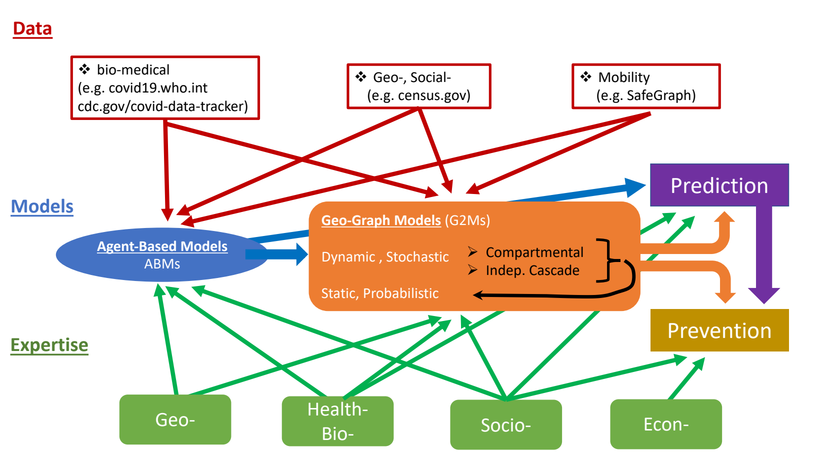
Prevention Challenge = Projection to Feasible Space



k	LP Constraints	Runtime	Cost
1	801	1.65s	41.69
2	991	3.04s	43.62
3	2131	10.90s	44.30
4	6976	100.08s	44.56

- Minimial Projection to the Feasible Polytope is Efficient (LP when the cost is l_1 norm)
- Weak dependence on the feasibility threshold (k)
- Working on learning the models (Ising or reacher model, two-state or mixed) and making the algorithms practical

Work in Progress & Path Forward



Work in Progress & Path Forward

Better and Richer Inference Models

- continuous-valued GM, e.g. Soft-Ising
- variational methods vs MCMC vs elimination

Learning Graphical Models

- Efficient and Exact with modern (pseudo-likelihood and interaction screening) approaches
- Adding NN – epidemiology-blind – learning as needed

Modeling Pipe Line: Data $\Rightarrow$ ABMs $\Rightarrow$ GMs

- Data (all kind) $\Rightarrow$ Learning New ABMs (rules) $\Rightarrow$ Generating synthetic samples $\Rightarrow$ Learning GMs $\Rightarrow$ Prediction= Inference

Work in Progress & Path Forward

Data Pipeline (to train Agent- and Graph- Models)

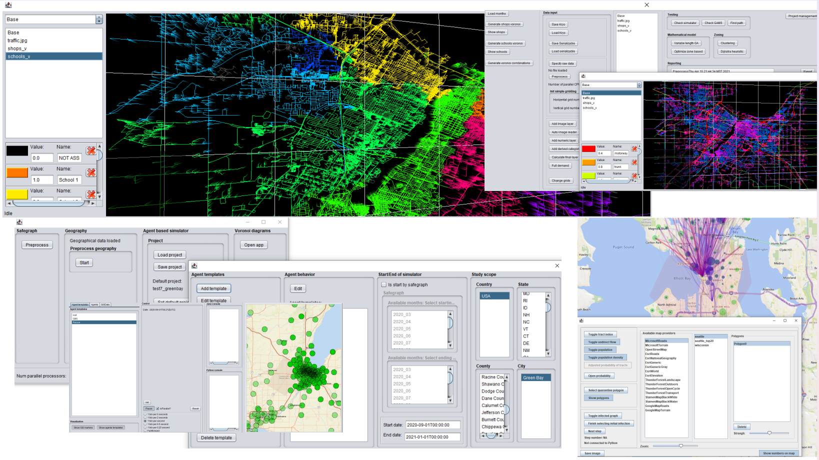
- **Safegraph** dataset shows cell phone trajectories of anonymous individuals which is used to model individual mobility.
  - Safegraph is a large dataset. The data is preprocessed and converted into a binary format to read and write faster.
- **Geographical information** of U.S.A which are represented as FIPS code and latitude/longitude are preprocessed.
- **Infection information** – by zipcode, air pollution, traffic load, and demographic data are added as overlay on the GIS data.
- **OpenStreetMap** dataset is used to calculate detailed travel time of individuals.

Work in Progress & Path Forward

Agent-Based software development (within the team)

- to pre-process **all the datasets** required for modeling
- to model & run ABMs both in **Java and Python**
- **user-friendly** templates and script for agent in Java or Python
- **Voronoi tessellation** – for different sets of points on map, representing travel time, to partition areas based on the needs (groceries, schools, etc.)

Work in Progress & Path Forward

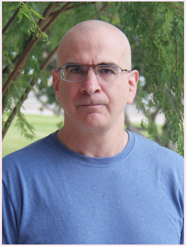


- rich variety of user interfaces and visualization tools are developed

Mikhail Krechetov (Skoltech)

Amir Mohammad Esmaieeli Sikaroudi (UA/CS)

Valentin Polishchuk (Linköping)



Alon Efrat (UA/CS)

Atsushi Nara (SDSU)

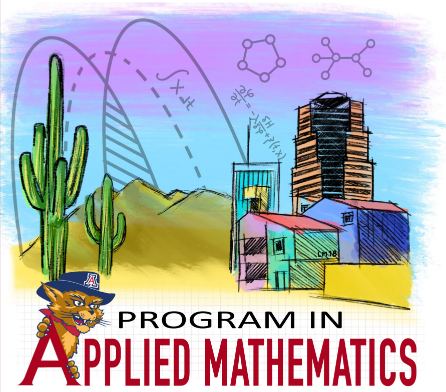
Eyal Oren (SDSU)

Support is Appreciated !!

- **Pandemic/GM**: NSF/Rapid award

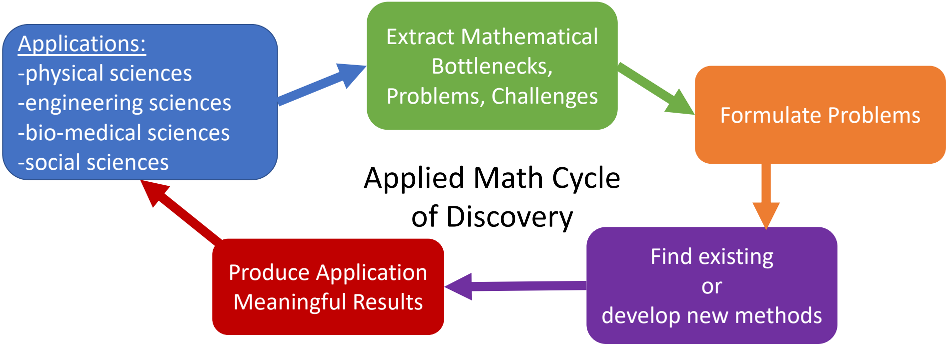
Thanks for your attention !

- Research focused, since 1976, one of the US first [dynamical systems, integrability, turbulence …]
- Interdisciplinary: 100+ professors/ 26 departments/ 8 colleges across UA campus (CoS & CoE & Optics – top 3)
- Mixing traditional @ contemporary Applied Math
- Graduate, Ph.D. focused, no terminal M.Sc.
- 60 Ph.D students (**13/**16/10 enrolled in **2021**/20/19)
- **3 Core Courses** (1st year -- Methods, Analysis, Algorithms) https://appliedmath.arizona.edu/students/new-core-courses
- Strong collaborations with Industry (e.g. Raytheon, Uber, Intel, Critical Path, etc) and National Labs (e.g. LANL, LLNL, NREL, NNSS, etc)
- 5 seminar/colloquium series – recorded and posted online
- Participation in many UA & National Edu Projects



http://appliedmath.arizona.edu/

How does Mathematics work with Applications @ UArizona?



Core courses provide hands on teaching of the AM-cycle methodology

- Training in **methods** (Math/APPL **581**), **theory** (Math/APPL **584**), **algorithms** (Math/APPL **589**)
- Math (quantitative) and Application-specific (qualitative) intuition