



OPEN

Prediction and prevention of pandemics via graphical model inference and convex programming

Mikhail Krechetov¹, Amir Mohammad Esmiaeeli Sikaroudi², Alon Efrat^{2,3}, Valentin Polishchuk⁴ & Michael Chertkov^{2,3,5}✉

Hard-to-predict bursts of COVID-19 pandemic revealed significance of statistical modeling which would resolve spatio-temporal correlations over geographical areas, for example spread of the infection over a city with census tract granularity. In this manuscript, we provide algorithmic answers to the following two *inter-related public health challenges of immense social impact which have not been adequately addressed* (1) *Inference Challenge* assuming that there are N census blocks (nodes) in the city, and given an initial infection at any set of nodes, e.g. any N of possible single node infections, any $N(N - 1)/2$ of possible two node infections, etc, what is the probability for a subset of census blocks to become infected by the time the spread of the infection burst is stabilized? (2) *Prevention Challenge* What is the minimal control action one can take to minimize the infected part of the stabilized state footprint? To answer the challenges, we build a Graphical Model of pandemic of the attractive Ising (pair-wise, binary) type, where each node represents a census tract and each edge factor represents the strength of the pairwise interaction between a pair of nodes, e.g. representing the inter-node travel, road closure and related, and each local bias/field represents the community level of immunization, acceptance of the social distance and mask wearing practice, etc. *Resolving the Inference Challenge* requires finding the Maximum-A-Posteriori (MAP), i.e. most probable, state of the Ising Model constrained to the set of initially infected nodes. (An infected node is in the $+1$ state and a node which remained safe is in the -1 state.) We show that almost all attractive Ising Models on dense graphs result in either of the two possibilities (modes) for the MAP state: either all nodes which were not infected initially became infected, or all the initially uninfected nodes remain uninfected (susceptible). This bi-modal solution of the Inference Challenge allows us to re-state the *Prevention Challenge* as the following *tractable convex programming*: for the bare Ising Model with pair-wise and bias factors representing the system without prevention measures, such that the MAP state is fully infected for at least one of the initial infection patterns, find the closest, for example in l_1 , l_2 or any other convexity-preserving norm, therefore prevention-optimal, set of factors resulting in all the MAP states of the Ising model, with the optimal prevention measures applied, to become safe. We have illustrated efficiency of the scheme on a quasi-realistic model of Seattle. Our experiments have also revealed useful features, such as sparsity of the prevention solution in the case of the l_1 norm, and also somehow unexpected features, such as localization of the sparse prevention solution at pair-wise links which are NOT these which are most utilized/traveled.

We follow our previous work¹ in justification for the use of the Graphical Models (GM) to study and mitigate pandemics. Therefore, we start from providing a brief recap of the prior literature on modeling of the epidemics, describe the logic which led us in¹ to the Ising Model (IM) formulation, and then state formally the inference and prevention problems addressed in the manuscript.

Difficulty in both predicting and neutralizing the spread of pandemics is a major social challenge of humanity. Technically speaking, we are yet to design a coherent data lifecycle for modeling and prevention both in terms of the global strategies and local tactics. To address the challenge, we must devise a hierarchy of spatio-temporal

¹Skolkovo Institute of Science and Technology, Moscow 121205, Russia. ²Department of Computer Science, University of Arizona, Tucson, AZ 85721, USA. ³Program in Applied Mathematics, University of Arizona, Tucson, AZ 85721, USA. ⁴Communications and Transport Systems, Linköping University, Norrköping 60174, Sweden. ⁵Department of Mathematics, University of Arizona, Tucson, AZ 85721, USA. ✉email: chertkov@arizona.edu

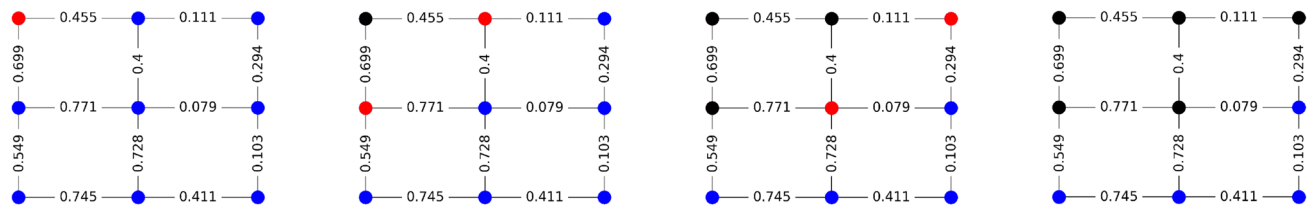


Figure 1. An exemplary random sequence (top-left to top-right to bottom-left to bottom-right) of the Independent Cascade Model (ICM) dynamics over 3×3 grid. Nodes colored red, blue, and black are **Infected**, **Susceptible**, and **Removed** at the respective stage of the dynamical process. This (shown) sample of the dynamic process terminates in 3 steps. Ising Model of Pandemic (IMP), which is the focal point of this manuscript, describes a regularized version of the ICM terminal state, where only two states (**S**—blue and **R**—black) are left. (See text for details).

models with different resolutions—from individual to community, county to the city, and from the moment a pathogen first enters our bodies, to days of disease development and to community transmission. Importantly, the models should be efficient in computing probabilistic predictions (for instance, offering the marginal probability heat map for the city neighborhoods to transition from the current/prior state of infection to the projected/a-posteriori state in 2 weeks).

Epidemiology and Mathematical Biology experts have relied in the past on a number of modeling approaches. The Agent-Based-Models (ABMs), introduced in epidemiology in 2004–2008^{2–7}, have complemented the earlier compartmental models^{8–11}. Using ABMs, even though not exclusive to epidemiology^{12,13}, became a breakthrough in the field, as they allowed to make a significant improvement in the quality of predictions, especially in the spatio-temporal resolution of how the disease spreads and how one can mitigate its spread. The models became and remained a core part of the epidemiology data life-cycle. (See for instance^{14,15} for most recent bibliography.) The ABMs provide a detailed prediction of how pandemics spread within counties, cities, and regions. A majority of the country-, city- or county-scale testbeds testing various mitigation strategies are resolved nowadays with ABMs. In particular, recently ABMs have been used extensively to inform public health in (non-pharmaceutical) interventions against the spread of COVID-19^{16–20}, and verify new strategies like test-trace-quarantine¹⁵, among many other applications.

There are two major problems with the modeling of pandemic. First, many parameters need to be calibrated on data. Second, even when calibrated for the current state of pandemic the models which are too detailed become impractical for making a forecast and for developing prevention strategies—both requiring checking multiple (forecast and/or prevention) scenarios. Using ABMs, which are clearly over-modeled (too detailed) is especially problematic in the context of the latter. For example, the open-source ABM solver FLUTE²¹ developed originally for modeling influenza, works with data that are acquired through Geographic Information Systems (GIS) on the scale of census tracts or communities, which is a very reasonable scale of spatial resolution to understand the dynamics of pandemics on a local scale. FLUTE populates each of the communities with thousands to millions of inhabitants in order to account for their daily patterns of travel. We believe that constructing effective Graphical Models (GM) of Pandemics with community-scale spatial resolution and then modeling pairwise (and possibly higher-order) epidemic interactions between communities directly, without introducing the thousands-to-millions of dummy agents, will complement (as discussed in the next paragraph), but also improve upon ABMs by being more efficient, robust and easier to calibrate.

An important, and possibly one of the first, Graphical Model (GM) of the COVID-19 pandemic was proposed in²². Dynamic bi-partite GMs connecting census tracts to specific Points Of Interest (non-residential locations that people visit such as restaurants, grocery stores and religious establishments) within the city and studying dynamics of the four-state (Susceptible, Exposed, Infectious and Removed) of a census tract (graph node) on the graph, were constructed in²² for major metro-area in USA based on the SafeGraph mobility data²³.

In fact, similar dynamic GMs, e.g. of the Independent Cascade Model (ICM) type^{24–28}, were introduced even earlier in the CS/AI literature in the context of modeling how the rumors spread over social networks (with a side reference on using ICM in epidemiology). As argued in²⁴ the Independent Cascade Models (ICMs) can be adapted to modeling pandemics. (Another interesting use of the ICM to model COVID-19 pandemic was discussed in²⁹.) In its minimal version, an ICM of Pandemic can be built as follows. Assume that the virus spreads in the community (census tract) sufficiently fast, say within five days—which is the estimate for the early versions of COVID-19 median incubation period. If an infected person enters a community/neighborhood but does not stay there, he infects others with some probability. If a single resident of the community becomes infected, all other residents are assumed infected as well (instantaneously). The model is a discrete-time dynamic model in which nodes in a network are in one of the three states: Susceptible, Infected, or Removed. The nodes represent communities/neighborhoods. A contact between an Infected community/node and another community which is Susceptible has an assigned probability of disease transmission, which can also be interpreted as the probability of turning the **S** state into **I** state. Consistently with what was described above, the network is represented as a graph, where nodes are tracts and edges, connecting two tracts, have an associated strength of interaction representing the probability for the infection to spread from one node to its neighbor. A seed of the infection is injected initially at random, for example, mimicking an exogenous super-spreader infection event in the area; examples could include political or religious gatherings. See Fig. 1 illustrating dynamics of the cascade model over 3-by-3 grid graph. Color coding of nodes is according to Susceptible = blue, Infected = red, Removed = black.

Given the starting infection configuration, each infected community can infect its graph-neighbor community during the next time step with the probability associated with the edge connecting the two communities. Then the infected community moves into the removed state. The attempt to infect each neighbor is independent of all other neighbors. This creates a cascading spread of the virus across the network. The cascade stops in a finite number of steps, thereby generating a random Removed pattern, shown in black in the Fig. 1, while other communities which were never infected (remain Susceptible) are shown in blue.

It was shown in¹ that with some regularization applied, statistics of the terminal state of the Cascade Model of Pandemic turns into a Graphical Model of the attractive Ising Model type.

This manuscript Road Map. Working with the Ising Model of Pandemic, we start the technical part of the manuscript by posing the *Inference/Prediction Challenge* in Section “[Ising model of pandemic](#)”. Here, the problem is stated, first, as the Maximum A-Posteriori over an attractive Ising Model, and we argue, following the approach which is classic in the GM literature, that problem can be re-stated as a tractable LP. We then proceed to Section “[Prevention challenge](#)” to pose the main challenge addressed in the manuscript—the *Prevention Challenge*—as the two-level optimization with inner step requiring resolution of the aforementioned Prediction Challenge. Aiming to reduce the complexity of the Prevention problem, we turn in Section “[Geometry of the MAP states](#)” to the analysis of the conditions in the formulation of the Prediction Challenge, describing the Safety domain in the space of the Ising Model parameters. We show the Safety domain is actually a polytope, even though exponential in the size of the system. We proceed in Section “[Projecting to the safe polytope](#)” with analysis of the Prevention Challenge, discussing the interpretation of the problem as a projection to the Safety Polytope from the polytope exterior, needed when the bare prediction suggests that system will be found with high probability outside of the Safety Polytope. Section “[Two polarized modes](#)” is devoted to approximation which allows an enormous reduction in the problem complexity. We suggest here that if the graph of the system is sufficiently dense, the resulting MAP solution may only be in one of the two polarized states (a) completely safe (no other nodes except the initially infected) pick the infection, or (b) the infection is spread over the entire system. We support this remarkable simplification by detailed empirical analysis and also by some theoretical arguments. Section “[Results](#)” is devoted to the experimental illustration of the methodology on the practical example of the Graphical Model of Seattle. The manuscript is concluded in Section “[Discussion](#)” with a brief summary and discussion of the path forward.

We conclude this introductory section with a general comment about relation between ABM and GM approaches. Focusing on GMs we did not mean to suggest that the GM approach is outright better than the ABM approach. In fact, we do believe that the two are complementary. ABM is more accurate but computationally heavy, while GM is coarser but computationally lighter. We are also convinced that it will be advantageous in the future to consider the two in tandem, e.g., with the parameters of the GM trained on the synthetic data generated by the ABM.

Ising model of pandemic

As argued in¹ the terminal state of a dynamic model generalizing the ICM model can be represented by the Ising Model of Pandemic (IMP), defined over graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where \mathcal{V} is the set of $N = |\mathcal{V}|$ nodes and \mathcal{E} is the set of undirected edges. The IMP, parameterized by the vector of the node-local biases, $h = (h_a | a \in \mathcal{V}) \in \mathbb{R}^N$, and by the vector of the pair-wise (edge) interactions, $J = (J_{ab} | \{a, b\} \in \mathcal{E})$, describes the following Gibbs-like probability distribution for a state, $x = (x_a = \pm 1 | a \in \mathcal{V}) \in 2^{|\mathcal{V}|}$, associated with \mathcal{V} :

$$P(x | J, h) = \frac{\exp(-E(x | J, h))}{Z(J, h)}, \quad (1)$$

where any node, $a \in \mathcal{V}$ can be found in either **S**—(susceptible, never infected) state, marked as $x_a = -1$, or **R**—(removed, i.e. infected prior to the termination) state, marked as $x_a = +1$. In Eq. (1), $E(x | J, h)$ and $Z(J, h)$ are model’s energy function and partition function respectively:

$$E(x | J, h) = \sum_{a \in \mathcal{V}} h_a x_a - \sum_{a, b \in \mathcal{V}} J_{ab} x_a x_b, \quad (2)$$

$$Z(J, h) = \sum_x \left(\sum_{a \in \mathcal{V}} h_a x_a - \sum_{a, b \in \mathcal{V}} J_{ab} x_a x_b \right). \quad (3)$$

In what follows, we will focus on finding the Maximum-A-Posteriori (MAP) state of the IMP conditioned to a particular initialization—setting a subset of nodes, $\mathcal{I} \subseteq \mathcal{V}$, to be infected. We coin the MAP problem *Inference challenge*.

Inference challenge

In this paper, we focus on finding the mode of distribution defined in Eq. (1). This problem is called maximum a posteriori (MAP) estimation, and the most probable states are called MAP states, or ground states (in physics literature). Moreover, in our setting, some nodes are initialized to plus one state (we denote this set of nodes by \mathcal{I}). The corresponding binary optimization problem is formulated as follows:

$$x^{(\text{MAP})}(\mathcal{I} | J, h) = \arg \min_x E(x | J, h) \Big|_{\forall a \in \mathcal{I}: x_a = +1}, \quad (4)$$

where we emphasize dependence of the MAP solution on the set of the initially infected nodes, \mathcal{I} .

Note that in general finding $x^{(\text{MAP})}$ is NP-hard³⁰. However if $J > 0$ element-wise, i.e. the Ising Model is attractive (also called ferromagnetic in statistical physics), Eq. (4) becomes equivalent to a tractable (polynomial in N) Linear Programming (see³¹ and references therein). In fact, the IMP is attractive, reflecting the fact that the state of a node is likely to be aligned with the state of its neighbor.

Let us also emphasize some other features of the IMP:

1. (G) should be thought of as an “interaction” graph of a city, reflecting transportation, commutes, and other forms of interactions between populations with the homes at the two nodes (census tracts) linked by an edge. The strength of a particular J_{ab} shows the level of interaction associated with the edge $\{a, b\}$.
2. A component, h_a , of the vector of local biases, h , is reflecting a -node specific factors such as immunization level, imposed quarantine, and degree of compliance with the public health measures (e.g., wearing masks and following other rules). Large negative/positive h_a shows that residents of the census tract associated with the node a are largely healthy/infected.

If solution of the Inference Challenge problem is such that the \mathbf{R} -subset of the MAP solution, $x^{(\text{MAP})}(\mathcal{I}|J, h)$, i.e.

$$\mathcal{R}(\mathcal{I}, J, h) = \left\{ a \in \mathcal{V} \mid x_a^{(\text{MAP})}(\mathcal{I}|J, h) = +1 \right\}, \quad (5)$$

is sufficiently large, we would like to mitigate the infection, therefore setting the Prevention Challenge discussed in the next Section.

Prevention challenge

Let us assume that modification of J and h are possible and consider the space of all feasible J and h . We will then identify *Safe Domain* as a sub-space of feasible J and h such that for all the initial sets of the initially infected nodes, \mathcal{I} , considered the resulting “infected” subset, $\mathcal{R}(\mathcal{I}, J, h)$, is sufficiently small. A more accurate definition of the Safe Domain follows. Then, we rely on the definition to formulate the control/mitigation problem coined Prevention Challenge. At this stage, we would also like to emphasize that studying the geometry of the Safe Domain is one of the key contributions of this manuscript.

Definition Consider IMP over $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and with the parameters (J, h) . Let us also assume that the set of initially infected nodes, \mathcal{I} , is drawn from the list, Υ . We say that (J, h) is in the k -*Safe Domain* if for every \mathcal{I} from Υ the number of \mathbf{R} -nodes in the MAP solution (4), is at most k , i.e.

$$\forall \mathcal{I} \in \Upsilon : |\mathcal{R}(\mathcal{I}, J, h)| \leq k, \quad (6)$$

where $\mathcal{R}(\mathcal{I}, J, h)$ is defined in Eq. (5).

Prevention challenge Given $(J^{(0)}, h^{(0)})$ describing the bare status of the system (city) which is not in the k -*Safe Domain*, and given the cost of the (J, h) change, $C((J, h); (J^{(0)}, h^{(0)}))$, what is least expensive change to $(J^{(0)}, h^{(0)})$ state of the system which is in the the k -*Safe Domain*? Formally, we are interested to solve the following optimization:

$$(J^{(\text{corr})}, h^{(\text{corr})}) = \arg \min_{(J, h)} C((J, h); (J^{(0)}, h^{(0)}))_{\text{Eq. (6)}}. \quad (7)$$

Expressing it informally, the Prevention Challenge seeks to identify a minimal correction (thus “corr” as the upper index) $(J^{(\text{corr})}, h^{(\text{corr})})$, which will move the system to the safe regime from the unsafe bare one, $(J^{(0)}, h^{(0)})$. The measures may include limiting interaction along some edges of the graph, thus modifying some components of J , or enforcing local biases, e.g., increasing level of vaccination, at some component of h .

Given that condition in Eq. (6) also requires solving Eq. (4) for each candidate (J, h) , the *Prevention challenge* formulation is a difficult two-level optimization. However, as we will see in the next Section, the condition in Eq. (6) (and thus the inner part of the aforementioned two-level optimization) can be re-stated as the requirement of being inside of a polytope in the (J, h) space. In other words, the (k) -*Safe Domain* is actually a polytope in the (J, h) space.

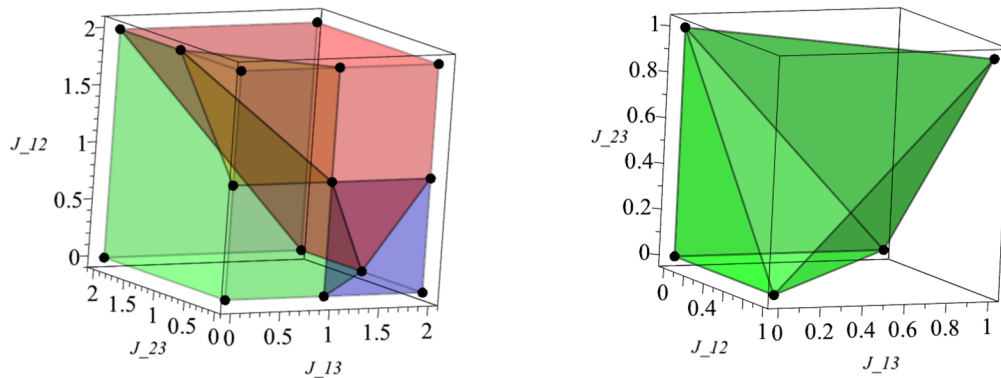
Geometry of the MAP states

Before solving the Prevention Challenge problem, we want to shed some light on the geometry of the MAP states. We work here in the space of all the Ising models over a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where each of the models is specified by (J, h) .

Proposition *Safe Domain of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $N = |\mathcal{V}|$ nodes is a polytope in the space of all feasible parameters, (J, h) , defined by an exponential in N number of linear constraints.*

Remark The Proposition allows us, from now on, to use *Safe Polytope* instead of the *Safe Domain*.

Proof of the Proposition The space of all the Ising models is divided into 2^N regions by the corresponding MAP states. Moreover, the boundary between any pair of neighboring regions is linear: consider two states $x^{(l)}$ and



(a) Geometry of the attractive Ising model illustrated on the example of a triangle graph (K_3) when a single node is infected. See explanations in the text.

(b) The Safe Polytope illustrated on the example of a triangle graph (K_3) with field vector $h = [-1, -1, -1]$. See explanations in the text.

Figure 2. Geometry of MAP states for Ising on K_3 .

$x^{(j)}$, and denote $(J, h)^{(i)}$ (resp. $(J, h)^{(j)}$) the set of all the Ising models with the MAP state $x^{(i)}$ (resp. $x^{(j)}$), then $(J, h)^{(i)}$ and $(J, h)^{(j)}$ are separated by the equation, $E(x^{(i)} | J, h) = E(x^{(j)} | J, h)$, which is linear in (J, h) . For a subset, $R \subseteq \mathcal{V}$, of nodes, let $x^{(R)}$ be the state in which, $x_a = +1, \forall a \in R, x_a = -1, \forall a \notin R$. Let $X^{(R)}$ be the set of all the MAP states, x , such that $\forall a \in R, x_a = +1$ (while other nodes, i.e. $b \in \mathcal{V} \setminus R$, are not constrained, $x_b = \pm 1$). Then the k -Safe Polytope, which we denote, $SP(k)$, is defined by at most $\sum_{k'=1}^k \binom{N}{k'} \cdot (2^{N-k'} - 1)$ linear inequalities:

$$SP(k) = \bigcap_{\substack{\forall R, |R| \leq k; \\ \forall x \in X^{(R)} \setminus x^{(R)}}} \left\{ (J, h) \mid E(x^{(R)} | J, h) > E(x | J, h) \right\}, \tag{8}$$

were some of these linear inequalities on the right hand side may be redundant.

Remark In the case of $k = 1$ (which, obviously, applies only if all the initial infections are at a single nodes, i.e. $\forall \mathcal{I} \in \Upsilon, |\mathcal{I}| = 1$), there are at most, $N \cdot (2^{N-1} - 1)$ linear inequalities.

We illustrate the geometry of the Ising model over the triangle graph (three nodes connected in a loop, K_3) in Fig. 2a and b. For both illustrations, we fix the h value to -1 at all the nodes, and we are thus exploring the remaining three degrees of freedom, J_{12}, J_{13}, J_{23} (since J is symmetric), which corresponds to exploring interactions within the class of attractive Ising models, $\forall a, b = 1, 2, 3 : J_{ab} \in \mathbb{R}_+$.

First, we consider the case when the only node $a = 1$ is infected. In this simple setting there are four possible MAP states, $(x_1, x_2, x_3) \in \{(+1, -1, -1), (+1, -1, +1), (+1, +1, -1), (+1, +1, +1)\}$, shown in Fig. 2a as green, blue, yellow and red, respectively. Finally, in Fig. 2b we plot the Safe Polytope $SP(1)$. We observe that the two “polarized” MAP states, $(+1, -1, -1)$ and $(+1, +1, +1)$, are seen most often among the samples, while domain occupied by the other two “mixed” MAP states, $(+1, -1, +1)$ and $(+1, +1, -1)$ is much smaller, with the two modes positioned on the interface between the two polarized states.

As will be shown below in the next Section, the polarization phenomena with only two “polarized” MAP states, which we coin in the following the two polarized modes, which we see on this simple triangle example, is generic for the attractive Ising model.

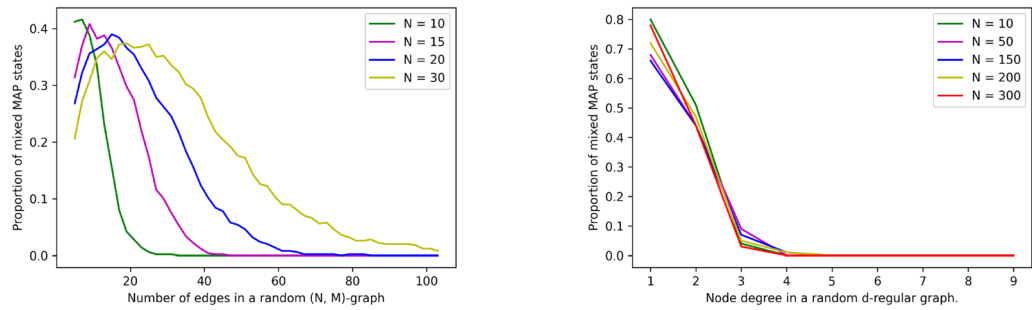
Two polarized modes

Definition Consider a particular subset of the initially infected nodes, \mathcal{I} (where thus, $\forall a \in \mathcal{I} : x_a = +1$). We call the MAP state of the model *polarized* when one of the following is true: (i) only initially infected nodes show $+1$ within the MAP solution, $\forall a \in \mathcal{I} : x_a = +1, \forall b \in \mathcal{V} \setminus \mathcal{I} : x_b = -1$ or (ii) all nodes within the MAP state show $+1, \forall a \in \mathcal{V} : x_a = +1$. We call a MAP state *mixed* otherwise.

Experimenting with many dense graphs, which are typical in the pandemic modeling of modern cities with extended infrastructures and multiple destinations visited by many inhabitants, we observe that the two polarized MAP states dominate generically, while the mixed states are extremely rare.

Figure 3a illustrates results of one our ensemble of random IMPs’ experiments. We, first, fix N to 20, pick M such that $M \leq N(N - 1)/2 = 190$ and then generate at random M edges connecting the 20 nodes. Then, for each of the random graphs (characterized by its own M) we generate 500 random samples of (J, h) , representing attractive Ising models. Finally, we find the MAP state for each IMP instance, count the number of mixed states and show the dependence of the fractions of the mixed states (in the sample set) in the Fig. (3a). A fast decrease of the proportion of the mixed states is observed with an increase in M .

Extension of these experiments (see Fig. 3b) suggests that when we consider an ensemble of IMPs over graphs with N nodes and the average degree $\alpha = O(1)$ which is sufficiently large (so that the graph is sufficiently dense) and increase N , we observe that the Mixed State Probability (MSP), or equivalently proportion of the



(a) Proportion of the mixed states in all samples for an ensemble of the (attractive) Ising Model of Pandemic over graphs with N nodes, shown as a function of the varying number of edges, M . Each shown point is the result of the averaging over 500 random instances of the (J, h) over the same graph. (See text for additional details.)

(b) Proportion of the mixed-to-polarized states (y-axis) vs α (x-axis) for an ensemble of the (attractive) Ising Model of Pandemic over d -regular graphs with N nodes, shown as a function of d . Each point is the result of averaging over 100 random instances of the (J, h) over different random graphs with the same node degree. (See text for additional details.)

Figure 3. Mixed-to-polarised transition experiments.

mixed-to-polarized states, decreases dramatically. Moreover, based on the experiments, we conjecture that the MSP decays to zero at $\alpha > \alpha_c$, but it saturates at $\alpha < \alpha_c$, where α_c is the threshold depending on the ensemble details. This threshold behavior is akin to the phase transition that occurred in many models of the spin glass theory³² and many models of the Computer Science and Theoretical Engineering defined over random graphs and considered in the thermodynamic limit, i.e. at $N \rightarrow \infty$. See e.g.³³ (application in the Information Theory, and specifically in the theory of the Low Density Parity Check Codes)³⁴ (applications in the Computer Science, and specifically for random SAT and related models) and references therein. We postpone further discussions of the conjecture for a future publication (see also brief discussion in Section “Discussion”).

We will continue discussion of the two-mode solution in the next Section.

Projecting to the safe polytope

In this Section we aim to summarize all the findings so far to resolve the Prevention Challenge formulated in Section “Prevention challenge”, specifically in Eq. (7) stating the problem as finding a minimal projection to the Safety Domain/Polytope from its exterior. The task is well defined, but in general, and as shown in Section “Geometry of the MAP states”, it is too complex—as the description of the Safety Polytope (number of linear constraints, required to define it) is exponential in the system size (number of nodes in the graph). However, the two-mode approximation, introduced in Section “Two polarized modes”, suggests a path forward: use the two-mode approximation and therefore remove all the linear constraints but one, separating the two polarized states.

Let us denote the two-mode approximation of the Safe Polytope by $\widehat{SP}(\Upsilon)$, where thus k in the original Safe Polytope, $SP(k)$, is replaced by the set Υ of all the initial infection patters. Then we write,

$$\widehat{SP}(\Upsilon) = \bigcap_{\mathcal{I} \in \Upsilon} \{ (J, h) \mid E(+1^{2N} \mid J, h) \geq E(x^{(\mathcal{I})} \mid J, h) \}, \tag{9}$$

where, $\forall a \in \mathcal{I} : x_a^{(\mathcal{I})} = +1$ and $\forall b \in \mathcal{V} \setminus \mathcal{I} : x_b^{(\mathcal{I})} = -1$. Eq. (9) represents a polytope stated in terms of the $|\Upsilon|$ constraints. In particular, if Υ accounts for all the initial infections, \mathcal{I} , of size not large than k , then $|\Upsilon| = \sum_{k'=1}^k \binom{N}{k'}$: the number of constraints grows exponentially in the maximal size of the initial infections, however the number of the constraints remains tractable for any $k = O(1)$. Replacing conditions in Eq. (7) by $\widehat{SP}(\Upsilon)$, defined in Eq. (9), one arrives at the following tractable (in the case of $k = O(1)$) convex optimization expression answering the Prevention Challenge approximately (within the two-mode approximation):

$$(\widehat{J}^{(\text{corr})}, \widehat{h}^{(\text{corr})}) = \arg \min_{(J, h)} C \left((J, h); J^{(0)}, h^{(0)} \right)_{\text{Eq. (9)}}. \tag{10}$$

This formula is the final result of this manuscript analytic evaluation. In the next Section we use Eq. (10), with $C(\cdot; \cdot)$ substituted by the l_1 -norm, l_2 -norm and their convex combination, to present the result of our experiments in a quasi-realistic setting describing a (hypothetical) pandemic attack and optimal defense, i.e., prevention scheme.

Results

Seattle data. We illustrate our methodology on a case study of the city of Seattle. (To verify scalability of the approach we have also experimented with larger system, e.g., models of New-York City and of the State of Wisconsin. This is still work in progress, partially reported in³⁵.) Seattle has 131 Census Tracts. Each Census Tract includes 1 to 10 Census Block Groups with 600 to 3000 residents and represents 1200 to 8000 population. Boundaries separating Census Tracts are designed to represent natural or urban landmarks and also to be persistent over time³⁶. To reduce complexity, we merge census tracts into 20 regions. See Fig. 4. To

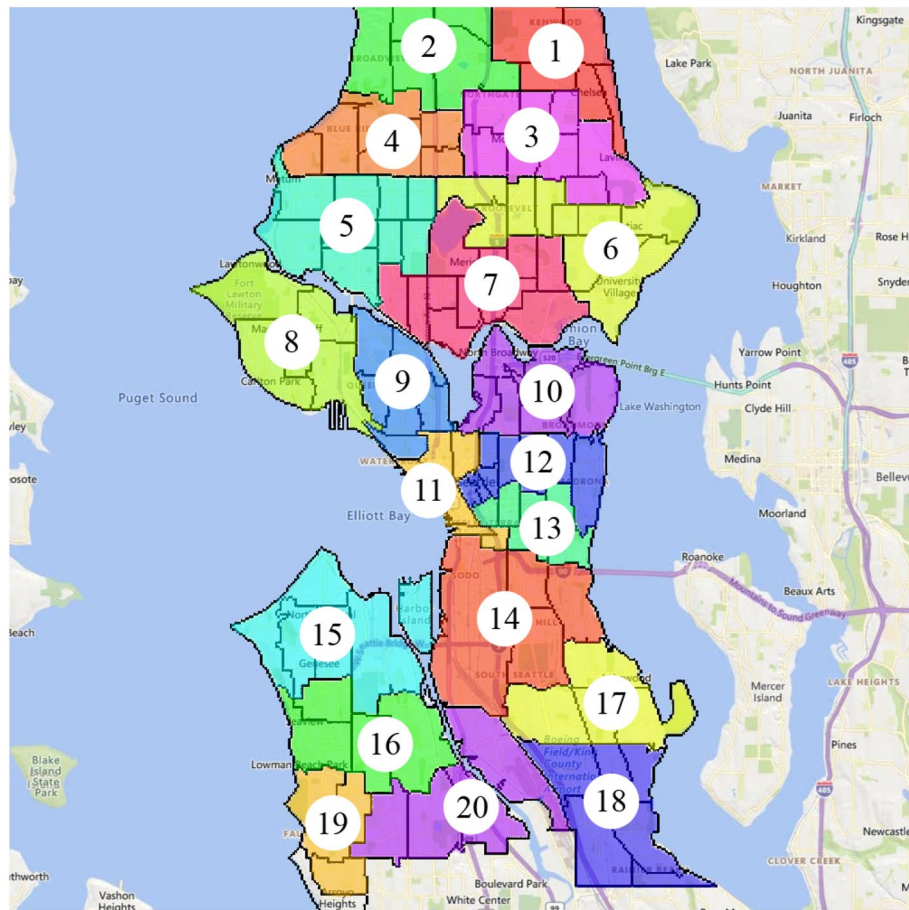


Figure 4. Seattle case study areas and census tracts³⁹.

prepare this splitting of Seattle into 20 regions/nodes, we utilize geo-spatial information from the TIGER/Line Shapefiles project provided by U.S. Census Bureau³⁷. The travel data of Seattle was extracted from the Safegraph dataset³⁸, which provides anonymized mobile tracking data. Each data point in the Safegraph database describes the number of visits from a Census Block to a specific point of interest represented by latitude and longitude. Mobility data associated with travelers crossing the boundaries of Seattle was ignored.

We then follow the methodology developed in¹ to combine the aggregated travel data with the epidemiological data, representing current state of infection in the area. This results in the estimation of the pair-wise interactions, J , parameterizing the Ising Model of Pandemic. We also come up with an exemplary (uniform over the system) local biases, h , completing the definition of the model. (We remind that the prime focus of the manuscript is on developing methodology which is AI sound and sufficiently general. Therefore, the data used in the manuscript are roughly representative of the situation of interest, however not fully practical.) We consider a situation with different levels of infection and chose $(J^{(0)}, h^{(0)})$ stressed enough, that is resulting in the prediction (answer to our Prediction Challenge), which lands the system in the dangerous domain—outside of the Safety Polytope.

Convex projection. In all of our experiments, we have used the general-purpose Gurobi optimization solver⁴⁰ to compute the MAP states and thus to validate the two-mode assumption. (We have also experimented with CVXPY⁴¹, but found it performing slower than Gurobi, at least over the relatively small samples considered in this proof of principles study. In the future, we plan to use existing, or developing new, LP solvers designed specifically for finding the MAP state of the attractive Ising model.) To illustrate our Prevention strategy, we took the Seattle data described above, and fed it as an input into the optimization (10), describing projection to the Safety Polytope, where $C(\cdot; \cdot)$ is substituted by the l_1 norm. CVXPY solver was used for this convex optimization task. Our code (python within jupyter notebook) is available at³⁵.

Table 1 shows results of our Prevention experiments on the Seattle data. We analyze, first, l_1 projection to $\widehat{SP}(\Upsilon)$ where Υ consists of all the initial infection patterns consisting of up to k nodes. In all of our experiments, the values of the field vector h (uniform across the system) was fixed to -1 . We observe that the number of constraints grows exponentially with k ; however, the cost of intervention remains roughly the same. We intend to analyze the results of this and other (more realistic) experiments in future publications aimed at epidemiology experts and public health officials.

Our initial reason for choosing l_1 in the experiments discussed above was based on empirical expectation that l_1 , as the closest of the convex options to l_0 , will enforce sparsity of the optimal reduction in the pair-wise

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

Table 1. Summary of our prevention experiments on the Seattle data. k , in the first column, is the maximal number of nodes in the initially infected patterns (all accounted for to construct the k -Safe Polytope). The second column shows number of linear constraints characterizing the k -Safe Polytope. Respective Run Time and Cost are shown in the 3rd and 4th column, where Cost shows the difference in l_1 norm between the $(J^{(0)}, h^{(0)})$, characterizing stressed but unmitigated regime, and the optimal prevention regime, resulting in $(\hat{J}^{(\text{corr})}, \hat{h}^{(\text{corr})})$ computed according to Eq. (10).

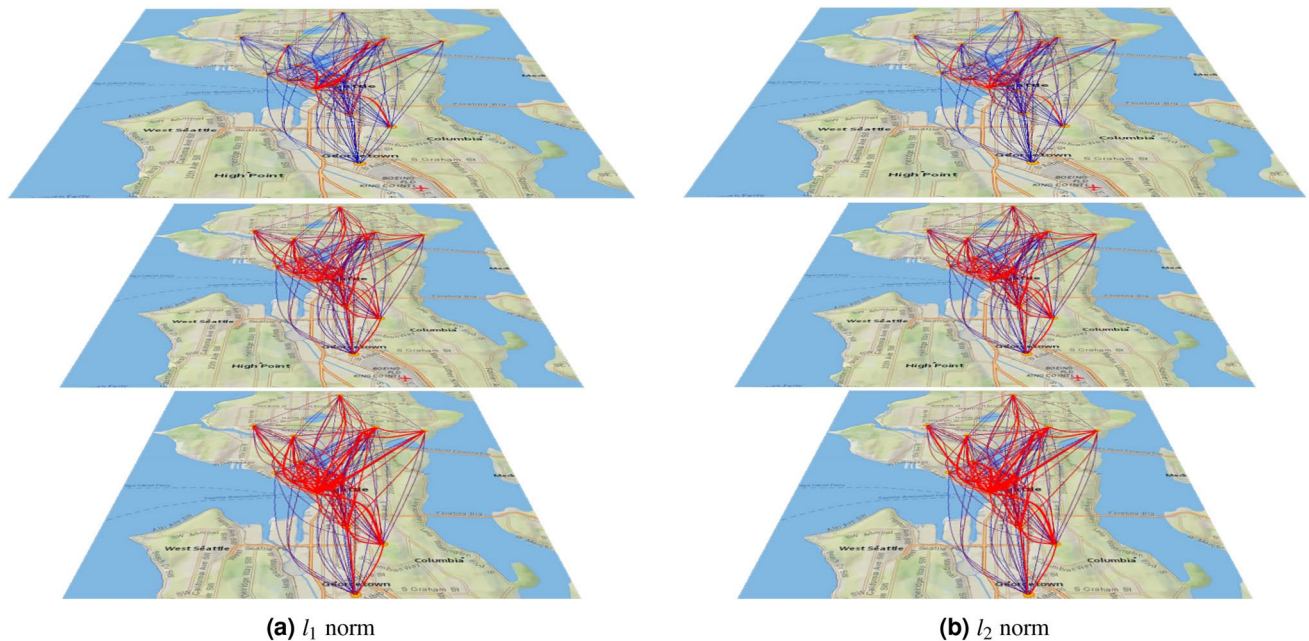
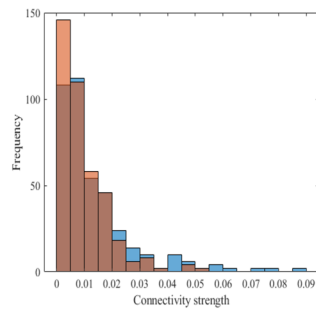


Figure 5. Comparison of the optimal prevention in the 20-node model of the Seattle city, correspondent to Eq. (7) with $k = 1$, computed for a particular instance of $J^{(0)}$ (leading to infection of the entire city, if not reduced) under l_1 (left sub-figure) and l_2 (right sub-figure). In each sub-graph initial instance, $J^{(0)}$, optimal correction, $\hat{J}^{(\text{corr})}$, and respective optimal reduction, $\hat{J}^{(\text{corr})} - J^{(0)}$, are shown at the bottom, middle and top layers respectively. Wider lines and red colors (in the red-to-dark blue color-coding) correspond to pair-wise links with higher values of J . (The Unfolding⁴² and Processing⁴³ software was used to create the figure).

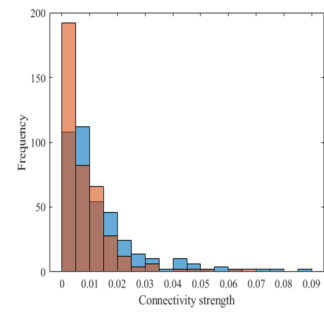
influences/traffic patterns. However, by choosing to work with the l_1 -norm, we certainly did not mean to suggest that other norms cannot be used, especially if propounded by public health and/or government experts—after all it is the professional judgment of these experts which is expected to drive practical choices of the cost function. Therefore, we also experimented with the l_2 norm and more generally with a convex mixture (interpolation between) l_1 and l_2 norms. These richer experiments are illustrated in Figs. 5 and 6.

Figure 5 compares initial infectious patterns and optimal mitigation in the l_1 (left sub-figure) and l_2 (right sub-figure) cases. The experiments suggest (consistently with the original hypothesis) that significant reduction (of influence/traffic) in the case of l_1 norm is observed at a fewer pair-wise links and, most interestingly, these reductions are NOT all along the links which are most intense (traveled less) in the original instance. On the contrary, in the l_2 case there are more edges with comparable reductions and these reductions are largely across the links which show heavy influences/traffic already in the bare cases (without correction) which require prevention.

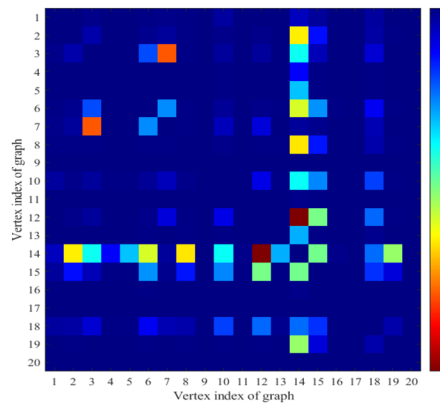
Figure 6 extends information provided in Fig. 5. Sub-figures (a) and (b) show histograms of the connectivity degree of the infected instance (blue) and corrected instance (orange) for the use case of the Seattle city from Fig. 5. We observe that there are fewer nodes of high degree in the corrected graph. Sub-figure (c) and Sub-figure (d) show the heat-map representation of the reduction graphs from the middle layers in Fig. 5, where color coding of the (i, j) element of the symmetric 20×20 matrix is chosen proportional to $(\hat{J}^{(\text{corr})} - J^{(0)})_{ij}$. Also, the histogram sub-figure show clearly that reduction mainly applies to immediate vicinity of the highest degree node in the downtown area of the Seattle city.



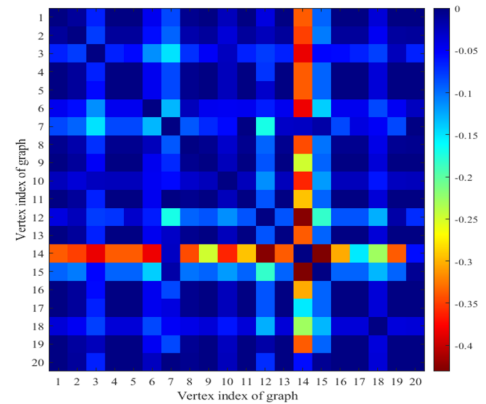
(a) Histogram of the connectivity degree in the case of the l_1 norm.



(b) Histogram of the connectivity degree in the case of the l_2 norm.



(c) Heat map of the reduction graph/matrix normalized by $J^{(0)}$ in the case of the l_1 norm, $\frac{(\hat{J}^{(corr)} - J^{(0)})}{\max(J^{(0)})}$



(d) Heat map of the reduction graph/matrix normalized by $J^{(0)}$ in the case of the l_2 norm, $\frac{(\hat{J}^{(corr)} - J^{(0)})}{\max(J^{(0)})}$

Figure 6. Sub-figure (a) and Sub-figure (b), correspondent to l_1 and l_2 cases respectively, present histogram of the connectivity degree of the infected instance (blue) and corrected instance (orange) for the case of the Seattle city shown in Fig. 5. Sub-figure (c) and Sub-figure (d) show the heat-map representation of the reduction graphs normalized by $J^{(0)}$ from the middle layers in Fig. 5, where color coding of the (i, j) element of the symmetric 20×20 matrix is chosen proportional to $(\hat{J}^{(corr)} - J^{(0)})_{ij}$. (The Matlab⁴⁴ was used to create the figure).

Discussion

In this manuscript, written specifically for an interdisciplinary audience of researchers in mathematical, physical and engineering sciences, we follow our prior work¹ (aimed primarily at experts in public health and epidemiology) and explain challenges in inference (prediction) and control (prevention) of pandemics. We use the language of GMs, which is one powerful tool in the modern arsenal of Data Science and AI, and state the Prediction Challenge as a MAP optimization over an attractive Ising model, which can be expressed generically as a solution of a tractable Linear Programming (LP). We then turn to the analysis of the prevention problem, which one formulates if the aforementioned prediction solution suggests that the probability of significant infection is above a pre-defined (by the public health experts) tolerance threshold. We show that in its simplest formulation, the prevention problem is equivalent to finding minimal l_1 projection to the safety polytope, where the latter is defined by solving the aforementioned prediction problem. In general, the polytope does not allow a description non-exponential in the size of the system. However, we suggested an approximation that allows to approximate the safety polytope efficiently - that is, linearly in the number of the initial infection patterns. The approximation is justified (empirically, with supporting theoretical arguments, however not yet backed by a mathematically rigorous theory) in the case when the interaction graph of the system (e.g., related to the system/city transportation and human-to-human interaction network) is sufficiently dense. We conclude by providing a quasi-realistic experimental demonstration on the GM of Seattle.

As a disclaimer but also a path forward, we would like to point out that this manuscript is not there yet where we would dream it to be in terms of practical use by health-care authorities/practitioners. However, the manuscript suggests an important, and yet missing in the literature, path towards the applications. We focused in the manuscript on developing methodology which we intend to make useful for practical purposes in the future. This will require accurate calibration of the pairwise and singleton terms in the Ising (Graphical) models, e.g., as mentioned above, based on “learning” the rates from much better developed ABMs. When such calibration of the parameters in the Ising model is done, the pair-wise terms (J) and the singleton terms (h) may be interpreted as corresponding to pair-wise “traffic” or “influences” between aggregated areas and singleton biases, e.g., related to degree of vaccination or compliance with the local public health mandates. Then, the outcome of this paper

methodology will become practical, e.g., resulting in (optimal) recommendations of the traffic/influence and local modifications to avoid a disastrous spread of infection.

We conclude with an incomplete list of technical challenges, presented in the order of importance (subjective), which need to be resolved to make the powerful GM approach to pandemic prediction and prevention practical:

- Build a hierarchy of Agent Based and Probabilistic Graphical Models which allow more accurate (than Ising model) representation of the infection patterns over geographical and community graphs. (See also our remark above about making the modeling more realistic.) The models may be both of the static (like Ising) or dynamic (like Independent Cascade Model) types. Extend the notion of the Safety Region (polytope) to the new GM of pandemics.
- Consider the case when the resolution of the Prediction Challenge problem returns a positive answer - most likely future state of the system is safe, and then develop the methodology which allows estimating the probability of crossing the safety boundary. In other words, we envision formulating and solving in the context of the GM a problem which is akin to the one addressed in⁴⁵: estimate the probability of finding the system outside of the Safety Polytope.
- Construct other (than two-mode) approximations to the Safety Polytope. In particular, approximations built on sampling of the boundaries of the safety polytope and learning (possibly reinforcement learning) are needed.
- Develop the asymptotic (thermodynamic limit) theory which allows validating (and/or correcting systematically) the efficient (two-mode and other) approximations of the Safety Polytope.

Received: 25 October 2021; Accepted: 20 April 2022

Published online: 09 May 2022

References

1. Chertkov, M. *et al.* Graphical models of pandemic. *Medrxiv* <https://doi.org/10.1101/2021.02.24.21252390> (2021).
2. Eubank, S. *et al.* Modelling disease outbreaks in realistic urban social networks. *Nature* **429**, 180–184. <https://doi.org/10.1038/nature02541> (2004).
3. Longini, I. *et al.* Containing pandemic influenza at the source. *Science* **309**, 1083–1087. <https://doi.org/10.1126/science.1115717> (2005).
4. Ferguson, N. *et al.* Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* **437**, 209–214. <https://doi.org/10.1038/nature04017> (2005).
5. Ferguson, N. *et al.* Strategies for mitigating an influenza pandemic. *Nature* **442**, 448–452. <https://doi.org/10.1038/nature04795> (2006).
6. Germann, T. C., Kadau, K., Longini, I. M. & Macken, C. A. Mitigation strategies for pandemic influenza in the united states. *Proc. Natl. Acad. Sci.* **103**, 5935–5940. <https://doi.org/10.1073/pnas.0601266103> (2006).
7. Halloran, M. *et al.* Modeling targeted layered containment of an influenza pandemic in the united states. *Proc. Natl. Acad. Sci.* **105**, 4639–4644. <https://doi.org/10.1073/pnas.0706849105> (2008).
8. Ross, R. *The Prevention of Malaria* (John Murray, 1910).
9. Kermack, W., McKendrick, A. & Walker, G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A* **115**, 700–721. <https://doi.org/10.1098/rspa.1927.0118> (1927).
10. Anderson, R. & May, R. *Infectious Disease of Humans: Dynamics and Control* (Oxford University Press, 1991).
11. Hethcote, H. The mathematics of infectious diseases. *SIAM Rev.* **42**, 599–653. <https://doi.org/10.1137/S0036144500371907> (2000).
12. Wikipedia. *Agent Based Models*. https://en.wikipedia.org/wiki/Agent-based_model (2020).
13. Downey, A. *Think Complexity: Complexity Science and Computational Modeling* 2nd edn. (O'Reilly Media Inc, 2018).
14. Lovasi, G. *et al.* *Population Health Methods: Agent Based Modeling* (Springer, 2020).
15. Kerr, C. *et al.* Covasim: An agent-based model of covid-19 dynamics and interventions. *PLOS Comput. Biol.* **17**, 1–32. <https://doi.org/10.1371/journal.pcbi.1009149> (2021).
16. Ferguson, N. *et al.* Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid-19 mortality and healthcare demand (2020).
17. Eubank, S. *et al.* Commentary on ferguson, et al. impact of non-pharmaceutical interventions (npis) to reduce covid-19 mortality and healthcare demand. *Bull. Math. Biol.* **82**, 52. <https://doi.org/10.1007/s11538-020-00726-x> (2020).
18. LANL. *Covid-19 Confirmed and Forecasted Case Data* (2020).
19. Maziarz, M. & Zach, M. Agent-based modelling for sars-cov-2 epidemic prediction and intervention assessment: A methodological appraisal. *J. Eval. Clin. Pract.* **26**, 1352–1360. <https://doi.org/10.1111/jep.13459> (2020).
20. Kaxiras, E. & Neofotistos, G. Multiple epidemic wave model of the covid-19 pandemic: Modeling study. *J. Med. Internet Res.* **22**, e20912 (2020).
21. Chao, D., Halloran, M., Obenchain, V. & Longini, J. I. M. Flute, a publicly available stochastic influenza epidemic simulation model. *PLoS Comput. Biol.* **6**, e1000656 (2010).
22. Chang, S. *et al.* Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* <https://doi.org/10.1038/s41586-020-2923-3> (2020).
23. SafeGraph. *Safegraph Social Distancing Metrics*. Safegraph Inc. <https://docs.safegraph.com/docs/social-distancing-metrics> (2021).
24. Kempe, D., Kleinberg, J. & Tardos, E. Maximizing the spread of influence through a social network. in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '03*, 137–146. <https://doi.org/10.1145/956750.956769> (Association for Computing Machinery, 2003).
25. Netrapalli, P. & Sanghavi, S. Learning the graph of epidemic cascades. in *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems, SIGMETRICS '12*, 211–222. <https://doi.org/10.1145/2254756.2254783> (Association for Computing Machinery, 2012).
26. Gomez-Rodriguez, M., Leskovec, J. & Krause, A. Inferring networks of diffusion and influence. *ACM Trans. Knowl. Discov. Data* <https://doi.org/10.1145/2086737.2086741> (2012).
27. Khalil, E., Dilkina, B. & Song, L. Cuttingedge: Influence minimization in networks. in *Workshop on Frontiers of Network Analysis Methods, Models, and Applications at NIPS* (2013).

28. Rosenfeld, N., Nitzan, M. & Globerson, A. Discriminative learning of infection models. in *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining, WSDM '16*, 563–572, <https://doi.org/10.1145/2835776.2835802> (Association for Computing Machinery, 2016).
29. Chen, Y., Lu, P., Chang, C. & Liu, T. A time-dependent sir model for covid-19 with undetectable infected persons. *IEEE Trans. Netw. Sci. Eng.* **7**, 3279–3294. <https://doi.org/10.1109/tNSE.2020.3024723> (2020).
30. Barahona, F. On the computational complexity of ising spin glass models. *J. Phys. A* **15**, 3241–3253. <https://doi.org/10.1088/0305-4470/15/10/028> (1982).
31. Živný, S., Werner, T. & Průša, D. The power of lp relaxation for map inference. *Advanced Structured Prediction*, 19–42 (The MIT Press, 2014).
32. Mezard, M., Parisi, G. & Virasoro, M. *Spin Glass Theory and Beyond* (World Scientific, 1986).
33. Richardson, T. & Urbanke, R. *Modern Coding Theory* (Cambridge University Press, 2008).
34. Mezard, M. & Montanari, A. *Information, Physics, and Computation* (Oxford University Press, 2009).
35. Supplementary Material and Source Code of the Manuscript. <https://github.com/mkrechetov/IsingMitigation>.
36. United States Census Bureau. *United States Census Bureau Glossary*. <https://www.census.gov/programs-surveys/geography/about/glossary.html> (2019).
37. United States Census Bureau. *Tiger Line Shapefiles Technical Documentation* (2021).
38. SafeGraph. *SafeGraph COVID-19 Data Consortium*. <https://www.safegraph.com/covid-19-data-consortium> (2021).
39. Office of Planning & Community Development. *Census Tract Map of Seattle*. <https://www.seattle.gov/Documents/Departments/OPCD/Demographics/GeographicFilesandMaps/2010CensusTractMap.pdf> (2010).
40. Gurobi Optimization, LLC. *Gurobi Optimizer Reference Manual* (2021).
41. CVXPY. *Convex Optimization for Everyone* (2021).
42. Unfolding maps. *Version 0.9.92* <https://github.com/tillnagel/unfolding> (2022).
43. Processing. *Version 3.5.4* <https://github.com/processing/processing> (2022).
44. MATLAB. *Version 9.10 (R2021a)* (The MathWorks Inc., 2022).
45. Owen, A., Maximov, Y. & Chertkov, M. Importance sampling the union of rare events with an application to power systems analysis. *Electron. J. Stat.* **13**, 231–254. <https://doi.org/10.1214/18-EJS1527> (2019).

Acknowledgements

This work was supported in part by NSF via #2027072 “RAPID:Infer and Control Global Spread of Corona-Virus with Graphical Models” project. The work of M.K. was supported by the Analytical center at Skoltech (subsidy agreement 000000D730321P5Q0002, Grant No. 70-2021-00145 02.11.2021). The work of V.P. was partially supported by the Swedish Research Council and the Swedish Transport Administration.

Author contributions

M.C. formulated the problem, M.K. developed the two-mode approximation and conducted the projection experiments, A.M.E.S. prepared Seattle data and conducted the experiments, A.E. and V.P. analyzed the results. All authors contributed to writing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.C.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2022