Optimal Deployment of Resources for Maximizing Impact in Spreading Processes

Andrey Y. Lokhov[1] and David Saad[2]

[1]Center for Nonlinear Studies and Theoretical Division T-4,
Los Alamos National Laboratory, Los Alamos, NM 87545, United States

[2]The Nonlinearity and Complexity Research Group,
Aston University, Birmingham B4 7ET, United Kingdom

The effective use of limited resources for controlling spreading processes on networks is of prime significance in diverse contexts, ranging from the identification of "influential spreaders" for maximizing information dissemination and targeted interventions in regulatory networks, to the development of mitigation policies for infectious diseases and financial contagion in economic systems. Solutions for these optimization tasks that are based purely on topological arguments are not fully satisfactory; in realistic settings the problem is often characterized by heterogeneous interactions and requires interventions over a finite time window via a restricted set of controllable nodes. The optimal distribution of available resources hence results from an interplay between network topology and spreading dynamics. We show how these problems can be addressed as particular instances of a universal analytical framework based on a scalable dynamic message-passing approach and demonstrate the efficacy of the method on a variety of real-world examples.

Spreading corresponds to omnipresent processes describing a vast number of phenomena in social, natural and technological networks [1–4] whereby information, viruses and failures propagate through their edges via the interactions between individual constituents. Spreading cascades have a huge impact on the modern world, be it negative or positive. An 11 minute power grid disturbance in Arizona and California in 2011 led to cascading outages and left 2.7 million customers without power [5]. As many as 579,000 people around the world could have been killed by the H1N1 influenza pandemic characterized by a rapid spreading through the global transportation networks [6]. The U.S. economy losses from the 2008 financial crisis resulted from cascading bankruptcies of major financial institutions are estimated at the level of $22 trillion [7]. Therefore, it is not surprising that efficient prediction and control of these undesired spreading processes are regarded as fundamental questions of paramount importance in developing policies for optimal placement of cascade-preventing devices in power grid, real-time distribution of vaccines and antidotes to mitigate epidemic spread, regulatory measures in interbanking lending networks and other modern world problems, such as protection of critical infrastructures against cyber-attacks and computer viruses [8].

On the other hand, spreading processes can also be considered beneficial. The ice bucket challenge campaign in social networks raised $115 million donations to the ALS association fighting the Amyotrophic Lateral Sclerosis, in particular due to a significant involvement of celebrities acting as "influencers" [9]. In the context of political campaigning, there are already winners [10, 11] and losers, and this division is likely to become more pronounced and critical in the future [12]. Winners are those who use communication and social networks effectively to set the opinions of voters or consumers, maximizing the impact of scarce resource such as activists or advertise-ments by applying control to the most influential groups of nodes at the right time; while losers will spend their resource sub-optimally relying on intuition and serendipity. Additional examples of domains where optimal resource allocation plays a crucial role in enhancing the effect of spreading include viral marketing campaigns [13], targeted chemically-induced control of dynamic biological processes [14]; drug discovery [15]; and even gaining military advantage through the propagation of disinformation [16]. All these applications share several important common properties such as restricted budget, finite-time windows for control interventions and the need for fast and scalable optimization algorithms which can be deployed in real time.

There exists a large body of work on optimal resource deployment in various spreading settings. A widely addressed formulation focuses on identifying influential spreaders, i.e. nodes that play important role in the dynamical process. Identification is often done by employing different centrality measures based on the topology of the underlying interaction network, including selection strategies based on high-degree nodes [17], neighbors of randomly selected vertices [18], betweenness centrality [19], random-walk [20], graph-partitioning [21], and k-shell decomposition [22], to name a few. It is quite natural that algorithms based exclusively on topological characteristics appear to have variable performance depending on particular network instances and dynamical models used [23, 24]. Another line of work consists in studying the NP-complete problem of network dismantling [25–27]: the underlying reasoning is that removal of nodes breaking the giant component to small pieces is likely to prevent the global percolation of the contagion. The localization of an optimal immunization set has been addressed using a belief propagation algorithm built on top of percolation-like equations for SIR (Susceptible, Infected, Recovered) and SIS (Susceptible, In-

fected, Survived) models [28], based on cavity method techniques developed previously for deterministic threshold models [29, 30]. This formulation is close to the problem of finding optimal *seeds*, i.e. the smallest set of initial nodes which maximizes the spread asymptotically [13]. It was rigorously analyzed [31, 32] for two simple diffusion models with a special submodularity property, Independent Cascade (IC) and Linear Threshold, and was shown to be NP-hard for both. A greedy algorithm based on a sampling subroutine has been explored for the IC model [33] in the setting of finite time horizon. For other spreading models the impact maximization problem at finite time and resources has been addressed in the setting of optimal control as reported in a recent survey [34]. However, only deterministic mean-field dynamics have been considered so far; this approximation ignores the topology of the specific network considered and yields non-distributed solutions to the control problem.

All of these techniques consider the problem of static (open-loop) resource allocation, preplanned at some initial time. A less explored direction consists in developing an online policy of assigning a limited remedial budget dynamically based on real-time feedback, also known as a closed-loop control. The impact of vaccination of the largest degree nodes or of those with the largest number of infected neighbors was investigated in [35, 36], while an alternative strategy is focused on the largest reduction in infectious edges [37]. Finally, an online policy based on the resolution of the minimal maxcut problem was introduced [38], where optimization is carried out with respect to the expected time to extinction of the SIS epidemic.

We introduce a general optimization framework which accommodates both dynamical and topological aspects of the problem and which allows for a broad range of objectives. The framework is principled, probabilistic, computationally efficient and incorporates the topological properties of the specific network under consideration. It facilitates the optimization of objective functions beyond the maximization or minimization of the spread, including: targeting specific nodes at specific times given a subset of accessible nodes; a limited global budget, possibly distributed over time; and an optimal dynamic vaccination strategy using the feedback from the spreading process. The problem is stated in a dynamical control setting with finite-time horizon that requires an explicit resolution of the dynamics, which is addressed via a distributed message-passing algorithm. We test the efficacy of the method on particular synthetic optimization problems as well as on a set of real-world instances.

I. RESULTS

A. Model

A large number of spreading models have been suggested in the literature to describe stochastic dynamical processes in epidemiology, information and rumor propagation, and cascades in biological and infrastructure networks [2–4]. They all share the same common features: the nodes transition from inactive to active state due to spontaneous activation mechanism associated with the nodes themselves, or due to interactions with active neighbours through the network edges. As an illustration of our approach, we have chosen a popular stochastic spreading process known as susceptible-infected-recovered, or SIR model, which is often used to describe propagation of infectious diseases or information spreading [2]. More precisely, we consider a generalized version of the discrete-time SIR model defined as follows. A node i in the interaction graph $G = (V, E)$, where V denotes the set of nodes, and E is the set of pairwise edges, at time step t can be found in either of three states σ_i^t: "susceptible" $\sigma_i^t = S$, "infected" $\sigma_i^t = I$ or "recovered" $\sigma_i^t = R$. At each time step, an infected (or, depending on the application domain, informed or active) individual i can transmit the activation signal to one of its susceptible (respectively, uninformed or inactive) neighbors j with probability α_{ij}, associated with the edge connecting them. Independently on the interaction between nodes a node i in the S state can turn active, assuming state I, with the probability $\nu_i(t)$, or spontaneously become recovered (uninterested, protected) with probability $\mu_i(t)$ at time step t. The first mechanism corresponds to a node activation due to an external influence such as advertisement in the context of information spreading. In the case of the epidemic spreading the second mechanism models the effect of vaccination: once a node goes to the protected R state, it becomes immune to the infection at all times. These probabilistic transmission rules at each time step t can be summarized using the following schematic rules:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j), \tag{1}$$

$$S(i) \xrightarrow{\nu_i(t)} I(i), \qquad S(i) \xrightarrow{\mu_i(t)} R(i). \tag{2}$$

In the definition of the dynamic rules (1) and (2), $\nu_i(t)$ and $\mu_i(t)$ represent control parameters we could manipulate with a certain degree of freedom defined by a particular instance of the problem. In what follows, we assume that the spreading couplings α_{ij} are known (or can be estimated) and are fixed in time. In some applications, α_{ij} may vary in time (e.g. this is true for temporal networks) or may represent a set of control parameters themselves. We outline such scenarios in the *Discussion* section; the optimization scheme presented below can be straightforwardly generalized to include the edge-related control parameters. However, for simplicity we will only

present optimization involving node-related control parameters.

To quantify the success of the spreading process, one may look for instance at the expected spread (the total number of infected nodes) at finial time horizon T, $\mathcal{S}(T)$, given by

$$\mathcal{S}(T) = \mathbb{E}\left[\sum_{i \in V} \mathbb{1}[\sigma_i^T = I]\right] = \sum_{i \in V} P_I^i(T), \qquad (3)$$

where the expectation is taken with respect to the realization of the stochastic dynamics and $P_I^i(T)$ denotes the marginal probability of node i to be found in the state I at time T. The quantities $P_S^i(T)$ and $P_R^i(T)$ can be defined in a similar way for the susceptible and recovered states, respectively. Hence, it is important to understand how to compute approximately the marginal probabilities $P_\sigma^i(t)$ on a given network, σ assuming the corresponding state; note that in the general case, an exact estimation of marginals in the SIR model is an NP-hard problem [39]. We use the recently introduced Dynamic Message-Passing (DMP) equations [40–42] which provide the estimates (asymptotically exact on sparse graphs) of the probabilities $P_\sigma^i(t)$ with a linear computational complexity in the number of edges and time steps. When applied to real-world loopy networks, the DMP algorithm typically yields a accurate prediction of the marginal probabilities as validated empirically [42] for a large class of spreading models on real-world networks. In the *Methods* section, we provide an intuitive derivation of the corresponding DMP equations for the generalized SIR model. An example of the DMP performance on real-world networks is provided in the Figure 1, where the method predictions are compared to values obtained through extensive Monte Carlo simulations of the SIR dynamics on a network of flights between major U.S. hubs (a detailed description of this data set is provided in the *Results* section and in the *Appendix A*). The accuracy of marginals estimation supports the use of the DMP equations at the core of our optimization algorithm.

B. Optimization framework

We formulate the dynamic allocation of resource as a general optimization problem with respect to an objective function $\mathcal{O}$ and a set of constraints associated with the budget of available resources $\mathcal{B}$, accessible values of control parameters $\mathcal{P}$, initial conditions $\mathcal{I}$ and the dynamical model equations $\mathcal{D}$. We employ the Lagrangian formulation of the constrained optimization problem:

$$\mathcal{L} = \underbrace{\mathcal{O}}_{\text{objective}} + \underbrace{\mathcal{B} + \mathcal{P} + \mathcal{I} + \mathcal{D}}_{\text{constraints}}. \qquad (4)$$

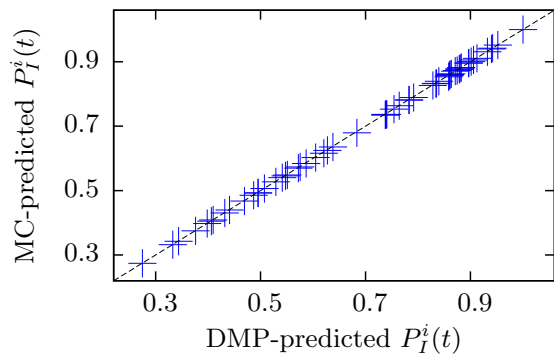Let us discuss the form of each term in the expression (4).

FIG. 1: Performance of DMP equations for the generalized SIR model on a network of flights between major US airports. The network represents $M = 383$ flight routes between the $N = 61$ largest US hubs. The weights α_{ij} are proportional to the average number of transported passengers on each route and are distributed in the interval $[0.05, 0.5]$; ν_i and μ_i are generated at random in the range $[0, 0.1]$. The scatter plot represents marginal probabilities $P_I^i(T)$ obtained from the DMP equations and by averaging over 10^7 Monte Carlo simulations. There is one randomly selected active node at the initial time and the dynamics is simulated for $t = 5$ time steps.

$\underline{\mathcal{O}}$ - Many objective functions of interest relate to the delivered information *at particular times* defined for each node. So for the general case we define:

$$\mathcal{O} = \mathbb{E}\left[\sum_{i \in U} \mathbb{1}[\sigma_i^{t_i} = I]\right] = \sum_i P_I^i(t_i), \qquad (5)$$

where t_i is the required activation time for node i and the sum is over the subset of nodes $U \subset V$ that is required to be activated. We refer to this general formulation as the *targeting* problem. The popular problem of maximizing the total spread $\mathcal{S}(T)$ is a special case whereby $U = V$ and $t_i = T$ for all $i \in V$.

$\underline{\mathcal{B}}$ - In many relevant situations, resources are not fully available at a given time, but rather become available on the fly, and their amount may vary across the time steps. For example, it takes some time to develop and produce the vaccines or the advertisement budget is allocated in stages depending on the success of the campaign. Hence, we define the budget constraints in the following form:

$$\sum_{i \in V} \nu_i(t) = B_\nu(t), \quad \sum_{i \in V} \mu_i(t) = B_\mu(t), \qquad (6)$$

where $B_\nu(t)$ and $B_\mu(t)$ denote the available total budget for the control parameters $\nu_i(t)$ (spontaneous infection) and $\mu_i(t)$ (recovery) at time t. The constraint $\mathcal{B}$ reads

$$\mathcal{B} = \sum_{t=0}^{T-1} \lambda_B^\nu(t) \left[\sum_{i \in V} \nu_i(t) - B_\nu(t)\right], \qquad (7)$$

with a similar expression for the parameters $\mu_i(t)$, where $\lambda_B^\nu(t)$ and $\lambda_B^\mu(t)$ are the associated Lagrange multipliers,

respectively. Clearly, one is not forced to use the whole available budget at each time step; in this case, we assume that $B_\nu(t)$ and $B_\mu(t)$ are reallocated accordingly at subsequent time steps. However, in cases where *specific* targeting times are not required, using monotonicity arguments, it is easy to show that it is always advantageous to use all available budget fully at each step for maximizing the impact at a later stage. Allocation of budget at the initial time only corresponds to the optimal *seeding* problem.

$\underline{\mathcal{P}}$ - In an unrestricted scenario, where all nodes are accessible, control parameters associated with node i, $\nu_i(t)$ and $\mu_i(t)$, may take arbitrary values from zero to one depending on total budget. However, in realistic situations access level to different nodes may differ: for example, only a subset $W \subseteq V$ of nodes may be controllable together with additional restrictions on parameter values. The parameter block $\mathcal{P}$ is introduced to enforce parameters $\nu_i(t)$ to take values in the range $[\underline{\nu_i^t}, \overline{\nu_i^t}]$ at each time step. This can be accomplished with the help of "barrier" functions, widely used in constrained optimization, assuming the form

$$\mathcal{P} = \epsilon \sum_{t=0}^{T-1} \sum_{i \in V} \left(\log \left[\nu_i(t) - \underline{\nu_i^t} \right] + \log \left[\overline{\nu_i^t} - \nu_i(t) \right] \right), \quad (8)$$

where ϵ is a small regularization parameter chosen to minimize the impact on the objective $\mathcal{O}$ in the regime of allowed $\nu_i(t)$ values, away from the borders. An equivalent expression can be written for the constraints on the $\mu_i(t)$ values.

$\underline{\mathcal{I} \text{ and } \mathcal{D}}$ - Finally, the constraints $\mathcal{I}$ and $\mathcal{D}$ enforce the given initial conditions and dynamics of the system via the associated Lagrange multipliers. For example, if no active individuals are present at initial time, then we set $P_I^i(0) = 0$ for all nodes using the constraint set $\mathcal{I}$; if some infected or recovered nodes are present, they assume an initial values 1 for the respective marginal probabilities. The set $\mathcal{D}$ encodes the evolution of the marginal probabilities with the DMP equations, as explained in the *Methods* section.

The extremization of the Lagrangian (4) is done as follows. Variation of $\mathcal{L}$ with respect to the dual variables (Lagrange multipliers) results in the DMP equations starting from the given initial conditions, while derivation with respect to the primal variables (control and dynamic parameters) results in a second set of equations, coupling the Lagrange multipliers and the primal variable values at different times. We solve the coupled systems of equations by forward-backward propagation, a widely used method for learning and optimization in artificial neural networks [43], detailed in the *Methods* section. This method has a number of advantages compared to other localized optimization procedures such as gradient descent and its variants. In particular, it is simple to implement, is of modest computational complexity, does

not require any adjustable parameters and is less prone to being trapped in local minima since the optimization is performed globally [44].

C. Targeting problem

We first demonstrate the approach using the general targeting problem, one of the new features of the suggested framework. In this toy example, we consider disinformation spreading on a small network extracted from the study of terrorists associations [46]. We assume that the spreading dynamics follows a particular case of the dynamical model with $\mu_i(t) = 0 \ \forall \ t$ and $i \in V$, corresponding to the Susceptible-Infected, or SI model with controlled spontaneous transition to the informed state I due to external influence via the control parameters $\nu_i(t)$. The activation of nodes is required in a predefined priority order, targeting selected nodes at specific times. The DMP-based optimization scheme converges to a unique optimal solution within a few forward-backward iterations as reported in Figure 2. The resources are allocated dynamically over time such that the activation path meets the targeting requirements: $P_I^i(t_i) > 0.95$ is achieved at all nodes, with the majority of nodes targeted with probability one.

Targeting is quite a general task and can provide algorithms to solve a number of related problems. For instance, identifying the origin of the spreading process from measurements at sparsely located sensors at different times [47] is a difficult problem that has been addressed by other approaches [48, 49] but can be equally viewed as optimally allocating a budget at time zero in order to target the sensor nodes at specific times that correspond to the times when measurements were taken.

D. Optimal seeding

The majority of existing algorithms [17–22, 25, 28, 31] have been designed to solve the seeding problem – finding an optimal set of nodes which would lead to the maximum number of activations at subsequent times. In the SI model of information spreading, even a single active node at initial time will ultimately lead to the activation of the whole connected network. However, a more interesting problem is the one of finding the best initial conditions which would lead to the maximum impact at finite time T. In the formulation involving control parameters $\nu_i(t)$, setting the initial conditions at time $t = 1$ is equivalent to distributing the activation budget at time $t = 0$ in the system where all nodes are at state S; optimal distribution of the budget at time $t = 0$ would thus lead through spontaneous infection to the maximum spread $\mathcal{S}(T+1)$.
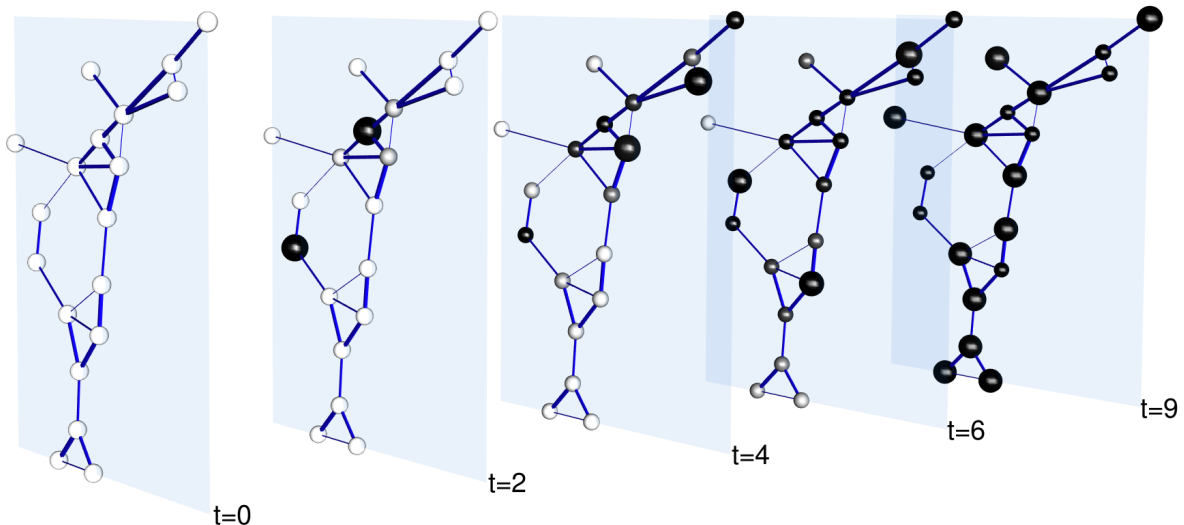
FIG. 2: **Optimal targeting with the DMP algorithm on a small network of terrorists associations**. Edge thickness indicates the strength of the corresponding pairwise transmission probability α_{ij}, generated uniformly at random in the interval $[0, 1]$. The size of nodes relates to the time activation requirements: large nodes should be activated by the corresponding time. In this example, two chosen nodes should be activated at time $t = 2$, another two nodes by time $t = 4$, three particular nodes by time $t = 6$ and all remaining nodes by time $t = 9$; available budget for each time step has been fixed to $B_\nu(t) = 0.1 \cdot N$. Color intensity (gradually from white to black) indicates the value of the marginal probabilities $P_I^i(t)$ which result from the dynamics using the optimal distribution of resources provided by the DMP algorithm. The visualization has been created using the MuxViz software [45].

For demonstration and validation purposes, we first consider the small network of relations between Slovene parliamentary parties in 1994; links represent estimated similarity relations based on a sociological survey of the parliament members who were asked to estimate the distance between each pair of parties in the political space [50]. Given the total campaigning budget $B_\nu(0)$, the goal is to maximize the total informational "influence" at time T by optimally distributing the "lobbying" budget at initial time. One advantage of this small test case is that we can check the validity of the scheme via an explicit symbolic solution of the DMP equations, obtaining a closed-form expression of the objective function $\mathcal{O} = \mathcal{S}(T + 1)$ that represents the final information spread as a function of the independent control parameters; a more in-depth description of this procedure is given in the *Appendix B*. As in the previous example, the forward-backward optimization scheme in this case quickly converges to a unique optimal solution starting from an arbitrary initial values of the control parameters. The ground-truth optimal values of parameters can be established by a direct maximization of the objective function $\mathcal{O}$ plotted in the Figure 3. The optimal solution is in full agreement with the solution of the forward-backward iteration scheme up to the insignificant domain border perturbations due to the finite value of ϵ, the regularization parameter that keeps values away from border values. This example validates our optimization procedure on this small scale problem.

Studies of the optimal seeding problem usually focus on a limited setting of homogeneous strength of links under deterministic dynamics and a search for the integer-valued deterministic budget deployment to specific nodes. To test the efficacy of the DMP-optimization approach on large scale instances we compare its performance to that of popular heuristics for this restricted setting. Although one should point out that our method addresses a broad range of problems and has not been optimized for this particular task, it is useful to assess its performance in the case where the structure of the ground-truth solution is known: for the deterministic SI spreading, it is clear that the initial distribution of seeds should target some combination of the high-degree nodes, and a number of well-performing centrality techniques [17, 22, 25] are known to select the respective combination (see the *Appendix B* for a detailed discussion of methods used for comparison).

Table I presents the normalized total spread for some of the best-performing centrality measures and the DMP algorithm after $T = 3$ time steps of the dynamics on different benchmark networks of various topologies and sizes. The transmission probabilities have been set to a uniform value $\alpha = 0.99$, and the total available seeding budget is equal to $B_\nu(0) = 0.05 \, N$. Note that the DMP-estimated marginals provide a natural and convenient measure for comparing the performance of different algorithms in the finite time horizon setting, especially on large graphs where running extensive Monte Carlo simu-

TABLE I: **Comparison of the DMP algorithm for the seeding problem in the setting of deterministic dynamics with popular well-performing heuristics on various real-world networks.** The left of the Table provides topological information on the networks considered [51–54]. On the right are presented values of the normalized total spread $\mathcal{S}(T)/N$ at time $T=3$ for the different algorithms: assignment to randomly-selected nodes, an adaptive version of the high-degree strategy of [17] (HDA), k-shell decomposition [22], Collective Influence CI_l [25] (with $l=2$ and $l=4$), uniform assignment and the DMP algorithm. For different test cases, solutions obtained by DMP span the range between delocalized and node-centric assignments and are on par with the best-performing centrality heuristics.

| Network	N	M	Random	HDA	k-shell	CI_2	CI_4	Uniform	DMP
Road EU	1174	1417	0.305	0.480	0.160	0.500	0.468	0.324	0.513
Protein	2361	6646	0.736	0.863	0.769	0.861	0.838	0.752	0.856
US Power Grid	4941	6594	0.367	0.602	0.209	0.605	0.565	0.397	0.601
GR Collaborations	5242	14484	0.565	0.644	0.296	0.660	0.658	0.634	0.710
Internet	22963	48436	0.880	0.998	0.969	0.996	0.994	0.891	0.972
Web-sk	121422	334419	0.645	0.833	0.239	0.751	0.734	0.699	0.837
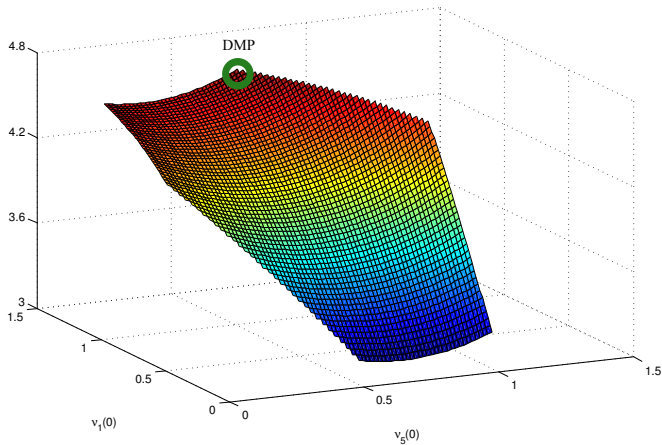
FIG. 3: **Validation of the DMP algorithm on a small network of relations between Slovene parliamentary parties with an explicit evaluation of the objective function.** Edge weights have been renormalized such that the maximum pairwise mutual "influence" receives the value $\alpha_{ij}^{\max} = 0.5$ and other weights are distributed proportionally to the survey data. We assume that only three nodes in this network belong to a controllable set $W = \{1, 3, 5\}$, and the total campaigning budget $B_\nu(0) = 1.5$ so that $\nu_3(0) = 1.5 - \nu_1(0) - \nu_5(0)$. The plotted surface represents the the total informational "influence" $\mathcal{O}(\nu_1(0), \nu_5(0))$ as a function of two independent seeding control parameters $\nu_1(0)$ and $\nu_5(0)$, computed via a symbolic solution of the DMP equations. Color variations reflect the change of the value of the objective, changing gradually from the minimum (blue) to the maximum (red). The DMP algorithm correctly recovers the parameter values that maximize the total spread at the finial time (the corresponding solution is marked by a green circle).

This suggests that the DMP algorithm performs well also for more general dynamic resource allocation problems with heterogeneous couplings, for which other principled methods do not exist. In terms of computational complexity, solving the dynamics with DMP is linear in T and $|E|$; the number of forward-backward iterations is typically small and can be controlled, as explained in the *Methods* section. This compares well against the other algorithms even in the considered restricted setting where taking into account the dynamics is not required, and allows one to use the DMP approach for very large real-world networks. Additional implementation details and remarks are given in the *Appendix B*.

Given that the problem is NP-hard, it is not surprising that the optimization landscape is much more complex in the case of large networks due to a presence of multiple solutions with comparable costs; the forward-backward iteration scheme no longer converges to a unique optimum as in the case of small networks considered before. Instead, the algorithm "jumps" between local optima that representing different control-parameter distributions that obey the budget constraints (6). This is an indication that it is arguably more appropriate to view the different seeding sets as a collective phenomenon, rather than assigning "influence" measure to individual nodes. In principle, several different initializations for $\nu_i(t)$ can be used to achieve the best solution; the results reported here correspond to the uniform starting values of the control parameters. Note that the initial distribution of $\nu_i(t)$ does not have to satisfy the budget condition (6), but the solutions obey the constraint already after the first forward-backward iteration.

lations is computationally prohibitive. Results presented in the Table I show that the DMP algorithm is close to the best-performing heuristics in all cases, showing a consistently good performance. Notice that our method does not rely explicitly on topological features such as targeting high-degree nodes, but instead explores a large space of parameters with impact on the full dynamic trajectory.

E. Online mitigation of epidemic spreading

To illustrate the suitability of the DMP algorithm to online deployment of resources in a dynamic setting with feedback we employ a prototypical example: developing an effective mitigation policy for confining an infectious

disease – a practical and challenging question of public concern. A SIR model with vaccination is an appropriate dynamic model in this case, where the $\nu_i(t)$ variables are set to zero, and the parameters $\mu_i(t)$ play the role of vaccination control, allowing the nodes to assume a protected state R. In contrast to the seeding problem, the initial conditions (origin of the epidemic) are specified in this setting and the vaccination budget has to be allocated dynamically according to the current state of the spreading process (monitored at each time step) in order to suppress the epidemic. The goal is to deploy the resources optimally so that the total number of infected nodes $\mathcal{S}(T)$ at the final time is minimized. The assumption of a time-distributed budget $B_\mu(t)$ is highly reasonable due to the restricted vaccine availability.

Previously developed real-time strategies for mitigating contagion on a given network [35, 37, 38] explored policies that are based on topological characteristics of the graph under the assumption of homogeneous transmission probabilities. The common denominator of existing approaches consists in local interventions which ensure the islanding of infected nodes. We generalize the methods [35, 38] to the case of heterogeneous transmission probabilities using a "high-risk" [36] ranking of nodes according to their probability of getting infected at the next time step. This measure is defined in our case as

$$P_i^t(S \to I) = 1 - \prod_{j \in \partial i} (1 - \alpha_{ji} \mathbb{1}[\sigma_j^t = I]), \qquad (9)$$

where ∂i denotes the set of neighbors of node i. A reasonable local intervention strategy for benchmarking consists in distributing the vaccination budget to priority nodes with a high-risk measure (9). This algorithm will be referred to as the *greedy* strategy.

Several policies can be conceived using the DMP optimization framework. As a reference, we consider the *planned* deployment of resources which does not take into account feedback from an actual realization of the process, but merely follows the solution of the dynamic resource allocation problem with a specified initial condition. Two other closed-loop strategies take into account the real-time information on the spreading process, using the seeding formulation as a subroutine: (a) The first, termed *DMP-greedy*, is close in spirit (but differs in the algorithmic implementation, based here on the DMP optimization framework) to the greedy algorithm and uses the current state of the epidemic as the initial condition, aiming to minimize the spread at the next time step only. (b) The second utilizes the full power of the DMP framework by exploiting the up-to-date information available to reinitialize the dynamics at each time step t to allocate the resources at the next time step $t+1$, by running the optimization procedure for the remaining $T-t$ time steps. This *DMP-optimal* policy is similar to the planned strategy, but takes advantage of the new information available from the realization of the process.

We compare these strategies for the case of infection spreading mediated by air traffic, which has been recognized to play an important role in recent world's pandemics [56]. As a particular example, we study the real-world transportation network of busiest flight routes between major U.S. airports, extracted from the Bureau of Transportation Statistics data (BTS) [55] and depicted in Figure 4 (a). We employ a plausible assumption that the infection transmission probability associated with a link between airports is proportional to the number of passengers carried along this route, see the *Appendix C* for a detailed description of the network and data used. The "vaccination" interventions on this network can be interpreted as quarantine measures taken in different airports using the updates on the new infected cases. In the simulations, we assume that the epidemic starts at the largest airport hub of Atlanta.

The comparison of different mitigation algorithms is given in Figure 4 (b), showing the average number of infected sites as a function of time under different mitigation strategies. As expected, the DMP-optimal scheme represents the best performing policy, which leads to stabilization of the expected number of infected nodes by $t=6$, at a lower level compared to the greedy algorithm that optimizes the spread at the next time step only. Notice that on a short time scale, the greedy algorithm has a slightly better performance, which represents a typical situation when localized and immediate optimal decisions lead ultimately to sub-optimal global optimization results.

II. DISCUSSION

We introduced an efficient, versatile and principled optimization framework for solving dynamic resource allocation problems in spreading processes, which allows for the synthesis of previously studied settings using a general targeting formulation. This probabilistic framework allows for the study of problems that involve a finite-time horizon, which requires an explicit solution of the dynamics, the targeting of specific nodes at given times and it accommodates scenarios where only a subset of the nodes is accessible. This is done in our scheme using the DMP equations for spreading processes. Although in this work we focused on the examples involving the discrete-time generalized SIR model, the approach can be straightforwardly applied to the case of continuous dynamics (the continuous formulation is expounded in the *Appendix D*) and to other spreading models, such as cascading and threshold models as well as rumor dynamics [42]. Another possible application area of the present framework relates to systems defined on temporal graphs, where network dynamics can be encoded into the time-dependent coefficients $\alpha_{ij}(t)$ within the existing framework.

Although we showed that the method can be employed in the case where transmission probabilities are uniform
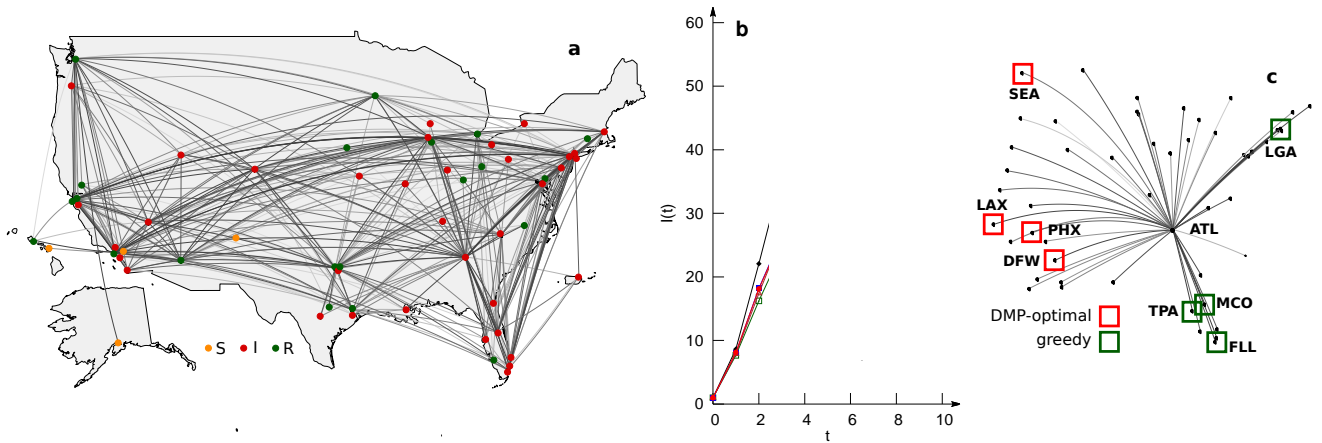
FIG. 4: **Online mitigation of air-traffic mediated epidemic on the network of flights between major U.S. hubs.** (a) A geographical layout of the air transportation sub-network extracted from the BTS data [55]. The transmission probabilities are indicated by the thickness of the corresponding edges, which is proportional to the aggregated traffic between airports. Different colors of airports (yellow, red and green) represent an outcome of a single realization of the spreading dynamics (nodes in the susceptible, infected and recovered states, respectively) under the DMP-optimal policy. (b) Comparisons of mitigation strategies showing the average number of infected sites as a function of time, averaged over 100 random realizations of the dynamics, as a result of applying different policies. In the simulations the epidemic starts at the largest airport hub of Atlanta; a budget of $B_\mu(t) = 0.5 \; N$ is available at each time step and the objective is to suppress the epidemic by $T = 10$. The DMP-optimal algorithm demonstrates the best performance in the number of infected nodes at time T. (c) An illustration of a radically different decisions taken by the DMP-optimal and greedy algorithms already at the first step of the optimization: the greedy policy chooses to vaccinate nodes which are most "in danger" at the next time step, while the decision done by the DMP-optimal scheme takes into account the forecasted evolution of the dynamics.

and only the detailed topology of the network is known, its major advantage consists in the ability to incorporate efficiently detailed information on transmission probabilities when such prior information is available, or can be either estimated (as in the examples of the Slovene political parties or flight transportation networks) or learned from observations of the dynamics [57].

The optimization method used is interesting in itself being based on changes to the entire trajectory instead of taking incremental improvement steps in the direction of the gradient; thus, the suggested algorithm results in large steps and arguably explores more effectively the parameter space. In spite of the global budget constraints involving all network nodes the resulting message-passing scheme is fast and distributed, requiring a number of operations which grows linearly in time and with respect to the number of edges in the network. An attractive property of the suggested framework is its versatility: instead of optimizing the spread given a fixed budget one can minimize the budget needed to meet certain requirements on the spread, imposed as a constraint in the Lagrangian formulation. Another interesting scenario is the optimization over the spreading parameters α_{ij}: this formulation is useful in the design of technological networks or for mitigation of an epidemic by removing and adding links in the graph. Finally, it would be interesting to apply the presented optimization scheme to the percolation-type equations describing the asymptotic $T \to \infty$ limit of the spreading dynamics with heterogeneous couplings.

Work on these research directions are underway.

III. METHODS

A. Dynamic message-passing equations

Dynamic message-passing belongs to the family of algorithms derived using the cavity method of statistical physics and may be given an interpretation of passing messages along the graph edges. The obtained marginals are exact on tree graphs, and asymptotically exact on sparse random networks. We provide an intuitive derivation of the DMP equations for the adopted generalized SIR model, defined by (1) and (2). On a given instance of a network, these equations allow one to compute the marginal probability distributions $P^i_\sigma(t)$, where $\sigma \in \{S, I, R\}$ denotes the node state. The first key equation reads:

$$P^i_S(t) = P^i_S(0) \left(\prod_{t'=0}^{t-1} (1 - \nu_i(t'))(1 - \mu_i(t')) \right) \prod_{k \in \partial i} \theta^{k \to i}(t). \tag{10}$$

It states the probability of node i to be susceptible at time t and is equal to the probability that i was in the S state at initial time $P^i_S(0)$ and remained so until time t. It neither changed states by following the ν and μ mechanisms (in brackets), nor by being infected by a neighbor

(final term on right); the dynamic message $\theta^{k \to i}(t)$ has a meaning of the probability that node k did not pass an activation message to node i until time t. Strictly speaking, Eq. (10) is only valid on a tree graph; only in this case $\theta^{k \to i}(t)$ are independent for all $k \in \partial i$, so that the corresponding probability is factorized as in (10). However, in practice the decorrelation assumption holds to a good precision even on general networks, even with small loops, see [42] for in-depth discussions and supporting numerical experiments. The quantities $\theta^{k \to i}(t)$ are updated as follows:

$$\theta^{k \to i}(t) = \theta^{k \to i}(t-1) - \alpha_{ki}\phi^{k \to i}(t-1), \qquad (11)$$

which corresponds to the fact that $\theta^{k \to i}(t)$ can only decrease if an activation signal is passed along the directed link (ki); the corresponding probability equals the product of α_{ki} and the dynamic variable $\phi^{k \to i}(t-1)$, which has a meaning of the probability that node k is in the state I at time t, but has not infected node i until time $t-1$. To simplify further explanations we introduce the dynamic messages $P_S^{k \to i}(t)$, $P_I^{k \to i}(t)$ and $P_R^{k \to i}(t)$, which denote the probabilities that node k is found at time t in the states S, I or R, respectively, conditioned on node i remaining in state S. Alternatively, these variables can be thought of as the probabilities of k being susceptible, infected or recovered on a *cavity graph*, on which node i has been removed. Formally,

$$P_S^{k \to i}(t) = P_S^k(0)\left(\prod_{t'=0}^{t-1}(1-\nu_k(t'))(1-\mu_k(t'))\right)\prod_{l \in \partial k \backslash i}\theta^{l \to k}(t), \tag{12}$$

which coincides with the expression (10), except that $\theta^{i \to k}(t)$ is not included in the product on the right ($\partial k \backslash i$ denotes the set of neighbors of k without i). We also have

$$P_R^{k \to i}(t) = P_R^{k \to i}(t-1) + \mu_k(t-1)P_S^{k \to i}(t), \qquad (13)$$

which expresses the monotonic increase of $P_R^{k \to i}(t)$ at each time step with the probability $\mu_k(t-1)P_S^{k \to i}(t)$, and

$$P_I^{k \to i}(t) = 1 - P_S^{k \to i}(t) - P_R^{k \to i}(t) \qquad (14)$$

due to the properties of linked probabilities. We are now ready to formulate the last relation which leads to the closure of the system of message-passing equations. The evolution of the message $\phi^{k \to i}(t)$ reads:

$$\phi^{k \to i}(t) = (1-\alpha_{ki})\phi^{k \to i}(t-1) + \Delta P_I^{k \to i}(t-1) \qquad (15)$$

where $\Delta P_I^{k \to i}(t-1) \equiv P_I^{k \to i}(t) - P_I^{k \to i}(t-1)$. The physical meaning of equation (15) is as follows: $\phi^{k \to i}(t)$ decreases if the activation signal is actually transmitted (first term) and increases if node k transitions to the state I at the current time step. Equations (10)-(15) can be iterated in time starting from the given initial conditions $\{P_S^i(0), P_I^i(0), P_R^i(0)\}_{i \in V}$, with

$$\theta^{i \to j}(0) = 1, \quad \phi^{i \to j}(0) = \delta_{\sigma_i^0, I} = P_I^i(0). \qquad (16)$$

The marginals $P_S^i(t)$ used throughout the text are obtained using (10), while $P_I^i(t)$ and $P_R^i(t)$ are computed via

$$P_R^i(t) = P_R^i(t-1) + \mu_i(t-1)P_S^i(t-1), \qquad (17)$$

$$P_I^i(t) = 1 - P_S^i(t) - P_R^i(t). \qquad (18)$$

The computational complexity of the DMP equations for solving the dynamics up to time T is given by $O(|E|T)$, where $|E|$ is the number of edges in the graph, which makes them scalable to sparse networks with millions of nodes.

B. Enforcing dynamical constraints and backward equations

The dynamics $\mathcal{D}$ and initial conditions $\mathcal{I}$ constraints are enforced in a similar way to that of $\mathcal{P}$ and the budget $\mathcal{B}$ constraints in Eqs. (7) and (8). To each generic dynamic variable $\xi^i(t)$ and message $\chi^{k \to i}(t)$ we associate the corresponding Lagrange multipliers $\lambda_i^\xi(t)$ and $\lambda_{k \to i}^\chi(t)$ which enforce the relation between dynamic variables at subsequent times. For instance, the evolution of the quantities $\{P_R^i(t)\}_{i \in V}$ in the Lagrangian $\mathcal{L}$ is enforced via the term

$$\sum_{i \in V}\sum_{t=0}^{T-1}\lambda_i^R(t+1)\left[P_R^i(t+1) - P_R^i(t) - \mu_i(t)P_S^i(t)\right].$$

Variation with respect to the dual variables $\lambda_i^\xi(t)$ and $\lambda_{k \to i}^\chi(t)$ returns the *forward* DMP equations (10)-(18), while setting to zero the derivative of $\mathcal{L}$ with respect to the primal dynamic variables yields the relations between the Lagrange multipliers at subsequent times, which we interpret as the *backward* dynamic equations in our scheme. Similarly to (10)-(15), the backward equations have a distributed message-passing structure with linear computational complexity $O(|E|T)$, and are used to update the values of control parameters $\nu_i(t)$ and $\mu_i(t)$ at each iteration, taking into account the budget requirements (6). Specifically, initializing the control parameters $\nu_i(t)$ and $\mu_i(t)$ to some arbitrary values (e.g., uniform over all nodes and times), we first propagate the DMP equations *forward* in time, up to the horizon T; then, using the existing primal parameter values we fix endpoint conditions for the dual parameters and propagate the equations for the dual parameters *backward* in time, updating the control parameters respecting the budget and variation constraints. These two steps are iterated for a predefined number of times or until global convergence of the process.

In the large-scale problems, where the algorithm explores the space of parameters by hopping from one solution to another, we choose a simple strategy: we run the forward-backward algorithm for several iterations for

a range of values of the regularization parameter ϵ which appears in the $\mathcal{P}$ block, and keep track of the best local optimum which provides the solution to the optimization problem after a maximum number of iterations (kept below the desired threshold which determines the computational complexity) is reached. The choice of ϵ impacts on the type of solution obtained: larger values of ϵ correspond to solutions where the budget is disseminated more uniformly across nodes, while smaller values lead to weight concentration on particular nodes. Depending on the application and the level of control over nodes, one type of solution can be preferred to another; this flexibility represents an attractive feature of the DMP algorithm. An explicit form of the Lagrangian for the problems considered in this work together with additional details is given in the *Appendices B and C*.

Appendix A: Model and DMP equations

All models considered in the main text represent variants of a SIR model with possible spontaneous infection and vaccination transitions. We refer to this model as *generalized* SIR model, because usually the spontaneous $S \to I$ and $S \to R$ transitions are not considered; instead, the standard SIR model contains a spontaneous recovery transition $I \to R$ which was not relevant for the examples considered in the paper and therefore was ignored. However, the inclusion of this transition in the DMP equations is very easy and has been done in [42]. Hence, the model we consider here is defined as follows: at each time step t, the transitions from the state S to I and R occur with the probabilities summarized in the Figure 5 for individual nodes and edges. In the discrete time setting considered throughout the work, it may occur that both transitions to I and R states are realized at the same time; in this case, we assume that the transition to the R state effectively takes place. Note that in the continuous time setting (see Section D), this effect is of a second order in the discretization step dt, and hence this tie-breaking rule is not required for sufficiently small dt.
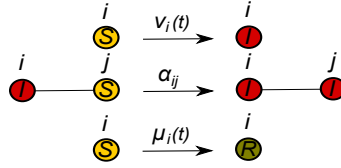

FIG. 5: Transition diagram summarizing dynamic rules in the generalized SIR model considered in this work. In the case of a simultaneous transition of a susceptible node (marked in yellow) to the states I (infected - red) and R (recovered - green) at time t, it assumes a recovered state at the next time step.

The DMP equations associated with this model, as well as the intuition behind them are described in the *Methods* section of the main text. We repeat them here for consistency:

$$P_S^{k \to i}(t) = P_S^k(0) \left(\prod_{t'=0}^{t-1} (1 - \nu_k(t')) \right) \left(\prod_{t'=0}^{t-1} (1 - \mu_k(t')) \right) \prod_{l \in \partial k \backslash i} \theta^{l \to k}(t), \tag{A1}$$

$$P_R^{k \to i}(t) = P_R^{k \to i}(t-1) + \mu_k(t-1) P_S^{k \to i}(t-1), \tag{A2}$$

$$\theta^{k \to i}(t) = \theta^{k \to i}(t-1) - \alpha_{ki} \phi^{k \to i}(t-1), \tag{A3}$$

$$\phi^{k \to i}(t) = (1 - \alpha_{ki}) \phi^{k \to i}(t-1) + \left[\left(P_S^{k \to i}(t-1) - P_S^{k \to i}(t) \right) - \left(P_R^{k \to i}(t) - P_R^{k \to i}(t-1) \right) \right]. \tag{A4}$$

The initial conditions are

$$\theta^{i \to j}(0) = 1, \qquad \phi^{i \to j}(0) = \delta_{\sigma_i^0, I} = P_I^i(0) = 1 - P_S^i(0). \tag{A5}$$

The marginal probabilities for nodes to be in the states S or I at time t are computed via

$$P_S^i(t) = P_S^i(0) \left(\prod_{t'=0}^{t-1} (1 - \nu_i(t')) \right) \left(\prod_{t'=0}^{t-1} (1 - \mu_i(t')) \right) \prod_{k \in \partial i} \theta^{k \to i}(t), \tag{A6}$$

$$P_R^i(t) = P_R^i(t-1) + \mu_i(t-1) P_S^i(t-1), \tag{A7}$$

$$P_I^i(t) = 1 - P_S^i(t) - P_R^i(t). \tag{A8}$$

As explained in the main text, the optimization problem is stated in the form of a Lagrangian to be extremized:

$$\mathcal{L} = \underbrace{\mathcal{O}}_{\text{objective}} + \underbrace{\mathcal{B} + \mathcal{D} + \mathcal{I} + \mathcal{P}}_{\text{constraints}}, \tag{A9}$$

where $\mathcal{O}$ is the objective function one would like to maximize, and $\mathcal{B}$, $\mathcal{D}$, $\mathcal{I}$ and $\mathcal{P}$ correspond to the constraints representing budget, dynamics, initial conditions and limits on the parameters, respectively. Below we discuss in detail how to obtain an approximate solution to the optimal control problem using this framework, for the examples of the spread maximization or minimization under different constraints. In all examples, we assume that budget constraints for spontaneous infections or vaccination take a global form over a subset $W \subseteq V$ of nodes in the network, and are specified at each time step:

$$\sum_{i \in W} \nu_i(t) = B_\nu(t), \qquad \sum_{i \in W} \mu_i(t) = B_\mu(t). \tag{A10}$$

Without loss of generality and for the sake of simplicity, in the equations below we assume that all nodes in the network are controllable, $W = V$; this case is the hardest in terms of the optimization procedure. Extending the derivation to the general case of any subset W is straightforward (an illustration for the case $W \neq V$ will be given in the Section B 3 below).

Appendix B: Maximizing information spread under special targeting policy

In this section, we write the detailed form of the Lagrangian for the targeting problem, explained in the main text: we assume that for each node $i \in V$ the activation is required at a predefined time t_i. We derive the corresponding forward and backward equations for the particular case of the generalized SIR model without a vaccination transition, i.e. assuming $\mu_i(t) = 0$ for all $i \in V$ and $t \in [0, T-1]$.

1. Lagrangian formulation

The Lagrangian in this case takes the following form:

$$
\begin{aligned}
\mathcal{L} = & \underbrace{\sum_{i \in V} \left(1 - P_S^i(t_i) \right)}_{\mathcal{O}} + \underbrace{\sum_{t=0}^{T-1} \lambda_B^\nu(t) \left[\sum_{i \in V} \nu_i(t) - B_\nu(t) \right]}_{\mathcal{B}} + \underbrace{\epsilon \sum_{t=0}^{T-1} \sum_{i \in V} \left(\log \left[\nu_i(t) - \underline{\nu_i^t} \right] + \log \left[\overline{\nu_i^t} - \nu_i(t) \right] \right)}_{\mathcal{P}} \\
& \left. \begin{aligned}
& + \sum_{i \in V} \sum_{t=0}^{T-1} \lambda_i^S(t+1) \left[P_S^i(t+1) - P_S^i(t)(1 - \nu_i(t)) \prod_{k \in \partial i} \frac{\theta^{k \to i}(t+1)}{\theta^{k \to i}(t)} \right] \\
& + \sum_{(ki) \in E} \sum_{t=0}^{T-1} \lambda_{k \to i}^S(t+1) \left[P_S^{k \to i}(t+1) - P_S^{k \to i}(t)(1 - \nu_k(t)) \prod_{l \in \partial k \setminus i} \frac{\theta^{l \to k}(t+1)}{\theta^{l \to k}(t)} \right] \\
& + \sum_{(ki) \in E} \sum_{t=0}^{T-1} \lambda_{k \to i}^\theta(t+1) \left[\theta^{k \to i}(t+1) - \theta^{k \to i}(t) + \alpha_{ki} \phi^{k \to i}(t) \right] \\
& + \sum_{(ki) \in E} \sum_{t=0}^{T-1} \lambda_{k \to i}^\phi(t+1) \left[\phi^{k \to i}(t+1) - (1 - \alpha_{ki}) \phi^{k \to i}(t) - P_S^{k \to i}(t) + P_S^{k \to i}(t+1) \right]
\end{aligned} \right\} \mathcal{D} \\
& \left. \begin{aligned}
& + \sum_{i \in V} \lambda_i^S(0) \left[P_S^i(0) - 1 + \delta_{\sigma_i^0, I} \right] + \sum_{(ki) \in E} \lambda_{k \to i}^S(0) \left[P_S^{k \to i}(0) - 1 + \delta_{\sigma_k^0, I} \right] \\
& + \sum_{(ki) \in E} \lambda_{k \to i}^\theta(0) \left[\theta^{k \to i}(0) - 1 \right] + \sum_{(ki) \in E} \lambda_{k \to i}^\phi(0) \left[\phi^{k \to i}(0) - \delta_{\sigma_k^0, I} \right].
\end{aligned} \right\} \mathcal{I}
\end{aligned}
$$

In this expression, the dual variables $\lambda_i^S(t)$, $\lambda_{k\to i}^S(t)$, $\lambda_{k\to i}^\theta(t)$ and $\lambda_{k\to i}^\phi(t)$ in $\mathcal{D}$ and $\mathcal{I}$ enforce the dynamics given by the DMP equations as well as the initial conditions at time zero, while $\lambda_B^\nu(t)$ is the corresponding Lagrange multiplier for the budget constraint in $\mathcal{B}$. The parameter constraint $\mathcal{P}$ has a form of the logarithmic barrier function, often used in constrained optimization, forcing each of the parameters $\nu_i(t)$ to take vales inside the interval $[\underline{\nu_i^t}, \overline{\nu_i^t}]$; for sufficiently small value of a positive coefficient ϵ, this regularization has a negligible impact on the objective function away from the extreme values.

Note that in the formulation above we have made an implicit assumption that $\alpha_{ki} < 1$ for all $(ki) \in E$, so that the ratios $\frac{\theta^{k\to i}(t+1)}{\theta^{k\to i}(t)}$ are correctly defined. In most situations, the case $\alpha_{ki} = 1$ for some $(ki) \in E$ is somewhat trivial, because there is no need for control for the nodes adjacent to these links. This assumption can be easily avoided by introducing other auxiliary variables in order to decouple $\theta^{k\to i}(t)$ at different times, which results in a slightly more complicated formulation.

2. Forward and backward equations

Once the form of the Lagrangian is established, we use a standard derivation with respect to primal and dual variables. The variation with respect to the multipliers $\lambda_i^S(t)$, $\lambda_{k\to i}^S(t)$, $\lambda_{k\to i}^\theta(t)$ and $\lambda_{k\to i}^\phi(t)$ give us back the direct DMP equations (A1)-(17), while the variation with respect to $\lambda_B^\nu(t)$ yields the cost constraint (A10). Setting the derivatives with respect to the primal variables to zero leads to the following set of dual equations:

$$\partial\mathcal{L}/\partial\phi^{k\to i}(t) = \left[\alpha_{ki}\lambda_{k\to i}^\theta(t+1) - (1-\alpha_{ki})\lambda_{k\to i}^\phi(t+1)\right]\mathbb{1}[t \neq T] + \lambda_{k\to i}^\phi(t) = 0, \tag{B1}$$

$$\partial\mathcal{L}/\partial P_S^{k\to i}(t) = \lambda_{k\to i}^S(t) + \lambda_{k\to i}^\phi(t)$$
$$- \left[\lambda_{k\to i}^S(t+1)(1-\nu_k(t))\prod_{l\in\partial k\backslash i}\frac{\theta^{l\to k}(t+1)}{\theta^{l\to k}(t)} + \lambda_{k\to i}^\phi(t+1)\right]\mathbb{1}[t\neq T] = 0, \tag{B2}$$

$$\partial\mathcal{L}/\partial\theta^{k\to i}(t) = \lambda_{k\to i}^\theta(t) - \lambda_{k\to i}^\theta(t+1)\mathbb{1}[t\neq T]$$
$$- \sum_{l\in\partial i\backslash k}\lambda_{i\to l}^S(t)P_S^{i\to l}(t-1)(1-\nu_i(t-1))\frac{1}{\theta^{k\to i}(t-1)}\prod_{m\in\partial i\backslash\{k,l\}}\frac{\theta^{m\to i}(t)}{\theta^{m\to i}(t-1)}\mathbb{1}[t\neq 0]$$
$$+ \sum_{l\in\partial i\backslash k}\lambda_{i\to l}^S(t+1)P_S^{i\to l}(t)(1-\nu_i(t))\frac{1}{\theta^{k\to i}(t)}\prod_{m\in\partial i\backslash l}\frac{\theta^{m\to i}(t+1)}{\theta^{m\to i}(t)}\mathbb{1}[t\neq T]$$
$$- \lambda_i^S(t)P_S^i(t-1)(1-\nu_i(t-1))\frac{1}{\theta^{k\to i}(t-1)}\prod_{m\in\partial i\backslash k}\frac{\theta^{m\to i}(t)}{\theta^{m\to i}(t-1)}\mathbb{1}[t\neq 0]$$
$$+ \lambda_i^S(t+1)P_S^i(t)(1-\nu_i(t))\frac{1}{\theta^{k\to i}(t)}\prod_{m\in\partial i}\frac{\theta^{m\to i}(t+1)}{\theta^{m\to i}(t)}\mathbb{1}[t\neq T] = 0. \tag{B3}$$

$$\partial\mathcal{L}/\partial P_S^i(t) = -\mathbb{1}[t=t_i] + \lambda_i^S(t) - \lambda_i^S(t+1)(1-\nu_i(t))\prod_{k\in\partial i}\frac{\theta^{k\to i}(t+1)}{\theta^{k\to i}(t)}\mathbb{1}[t\neq T] = 0. \tag{B4}$$

Finally, the variation with respect to the parameters $\nu_i(t)$ for $t < T$ gives

$$\partial\mathcal{L}/\partial\nu_i(t) = \lambda_B^\nu(t) + \sum_{l\in\partial i}\lambda_{i\to l}^S(t+1)P_S^{i\to l}(t)\prod_{m\in\partial i\backslash l}\frac{\theta^{m\to i}(t+1)}{\theta^{m\to i}(t)} + \lambda_i^S(t+1)P_S^i(t)\prod_{m\in\partial i}\frac{\theta^{m\to i}(t+1)}{\theta^{m\to i}(t)}$$
$$+ \frac{\epsilon}{\left(\nu_i(t) - \underline{\nu_i^t}\right)} - \frac{\epsilon}{\left(\overline{\nu_i^t} - \nu_i(t)\right)} = 0. \tag{B5}$$

This set of forward and backward equations is sufficient for determining $\{\underline{\nu}_i\}_{i\in V}$, using the following optimization procedure, analogous to [43, 44]. First, Initialize $\nu_i(t)$ to some initial values, for example using the uniform assignment $\nu_i(t) = 1/NT$ for all $i\in V$ and $t\in[0, T-1]$. Then repeat the following steps for a fixed number of iterations or until convergence:

1. Starting from the initial values for the dynamics variables (A5), propagate the DMP equations (A1)-(A8) forward, storing the values of dynamic messages at different times.

2. Use equations (B1)-(B5) for fixing the boundary values of Lagrange multipliers at time T:
   - (B1) assigns $\lambda_{k\to i}^\phi(T) = 0$;
   - (B2) gives $\lambda_{k\to i}^S(T) = 0$;

- (B4) gives $\lambda_i^S(T) = \delta_{t_i,T}$;
- (B5) sets $\lambda_B^\nu(T-1)$ and $\nu_i(T-1)$ through $\lambda_i^S(T)$ and $\lambda_{k\to i}^S(T)$. Writing (B5) as

$$\lambda_B^\nu(t) + \psi_i(t) + \frac{\epsilon}{\left(\nu_i(t) - \underline{\nu_i^t}\right)} - \frac{\epsilon}{\left(\overline{\nu_i^t} - \nu_i(t)\right)} = 0, \tag{B6}$$

we first express $\nu_i(T-1)$ for each $i \in V$ as a function of $\lambda_B^\nu(T-1)$ as a solution of the second order equation. When control parameters can take all possible values between 0 and 1, meaning that $\underline{\nu_i^t} = 0$ and $\overline{\nu_i^t} = 1$ for all $i \in V$ and $t \in [0, T-1]$, then there are two a priori possible solutions:

$$\nu_i^\pm(t) = \frac{\lambda_B^\nu(t) + \psi_i(t) - 2\epsilon \pm \sqrt{(\lambda_B^\nu(t) + \psi_i(t))^2 + 4\epsilon^2}}{2\lambda_B^\nu(t) + 2\psi_i(t)}. \tag{B7}$$

Assuming that ϵ is positive, we choose the solution with the positive sign in front of the square root which always leads to $0 < \nu_i^+(t) < 1$, thus expressing $\{\nu_i(T-1)\}_{i\in V}$ as a function of $\lambda_B^\nu(T-1)$. We then determine $\lambda_B^\nu(T-1)$ (and hence $\{\nu_i^\pm(T-1)\}_{i\in V}$) from the numerical solution of the budget constraint equation (A10).

- (B3) gives $\lambda_{k\to i}^\theta(T)$ through $\lambda_i^S(T)$, $\lambda_{k\to i}^S(T)$ and $\nu_i(T-1)$.

3. Using the computed values for dynamic messages at different times, propagate the equations (B1)-(B5) backward to compute the values of the Lagrange multipliers and parameters at all times (compute values at time t given values at times $t+1$):

- Eq. (B1) gives $\lambda_{k\to i}^\phi(t)$;
- solve (B2) and (B4) for evaluating $\lambda_{k\to i}^S(t)$ and $\lambda_i^S(t)$, correspondingly.
- compute $\lambda_B^\nu(t-1)$ and $\nu_i(t-1)$ using (B5) and (A10) in the same way as it has been done for $t = T-1$, see equation (B7);
- Eq. (B3) gives $\lambda_{k\to i}^\theta(t)$ through $\lambda_{k\to i}^\theta(t+1)$, $\lambda_i^S(t+1)$, $\lambda_{k\to i}^S(t+1)$, $\lambda_i^S(t)$, $\lambda_{k\to i}^S(t)$, $\nu_i(t)$ and $\nu_i(t-1)$.

4. Using the updated values of $\nu_i(t)$ in the step 1.

At each iteration, this procedure leads to "jumps" between local optima of the objective $\mathcal{O}$ verifying the budget constraint (A10), where we keep track of the best local optimum which provides the solution to the optimization problem after a maximum number of iterations is reached. Additionally, several different initializations for $\{\nu_i(t)\}_{i,t}$ and ϵ can be used in this scheme in order to achieve the best solution. Note that the initial distribution of $\{\nu_i(t)\}_{i,t}$ does not have to satisfy the budget condition (A10), but the solutions obey the constraint already after the first forward-backward iteration. In the case of a small number of nodes in the network N which corresponds to a simpler optimization landscape, we observed that the scheme quickly converges to a unique optimum starting from an arbitrary initial condition. In the next subsection, we describe tests of this scheme on a number of real-world networks.

3. Tests on real-world networks and comparisons with popular heuristics

For validating the optimization scheme and its performance against existing approaches, we test the algorithm on a particular case of the targeting setting: the optimal seeding problem which corresponds to optimizing the initial conditions. Indeed, the initial conditions at time one can be defined as an optimization over $\{\nu_i(0)\}_{i\in V}$ at time $t = 0$ at which nodes will have a spontaneous probability of switching to the infected state at time zero. This leads to a more general space of possible initial conditions, without restrictions to a selection of an integer number of infected seeds. A classical setting corresponding to the selection of a group of "influential spreaders" at initial time $t = 1$ can be recovered in our framework by imposing additional constraints to the domain of variation of $\{\nu_i(0)\}_{i\in V}$, restricting them to take only values close to zero or one (e.g. using an appropriate barrier potential).

a. Comparison with an explicit symbolic resolution of the DMP equations

As a validation example considered in the main text, we run the optimization scheme on a small network of relations between Slovene parliamentary parties in 1994, depicted in the Figure 6. The nodes in this network correspond to Slovene political parties in 1994, and the links represent the similarity relations estimated from a sociological survey of the parliament members which were asked to estimate the distance between pairs of parties in the political space [50]. We have renormalized the weights of edges in such a way that the maximum pairwise mutual "influence" receives the value $\alpha_{ij}^{\max} = 0.5$, and the other weights are distributed proportionally to the survey data. We use this real-world network for the first tests in the seeding problem, which
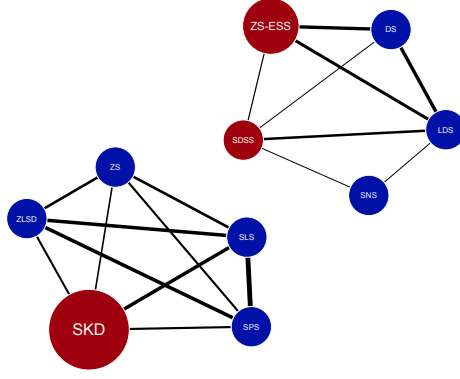
FIG. 6: A small network of relations between Slovene parliamentary political parties from the social study of [50]. The edge thickness is proportional to the estimated mutual "influence" α_{ij} between a pair of parties. Only red nodes are assumed to be controllable in this network. The sizes of the nodes reflect the optimal distribution of the seeding budget $B_\nu(0) = 1.5$ for maximizing the global impact: small-size nodes do not require any control, and the spontaneous activation probabilities associated with the medium and large-size nodes are equal to 0.5 and 1.0, respectively.

was defined as follows. We assumed that three nodes in this network belong to a controllable set $W = 1, 3, 5$, corresponding to the parties SKD, ZS-ESS and SDSS, correspondingly. Given the total campaigning budget $B_\nu(0) = 1.5$, the goal is to maximize the total informational "influence" at time $T = 3$ by finding the optimal initial distribution of the associated control parameters $\nu_1(0)$, $\nu_5(0)$ and $\nu_3(0) = 1.5 - \nu_1(0) - \nu_5(0)$.

As suggested in subsection B 3, the forward-backward optimization scheme in this case converges to a unique optimal solution in about 7 iterations with $\epsilon = 5 \cdot 10^{-4}$ starting from an arbitrary initial values of the control parameters, outputting the values $\nu_1^*(0) = 0.993217$ and $\nu_5^*(0) = 0.505758$. One advantage of this small test case is that we can check of the validity of the scheme via an explicit symbolic resolution of DMP equations, obtaining a closed-form expression of the objective function $\mathcal{O} = \sum_{i \in V} P_I^i(T)$, representing the final spread at time T as a function of the independent control parameters $\nu_1(0)$ and $\nu_5(0)$:

$$\begin{aligned}
\mathcal{O} = &- 0.0213192\nu_5(0)^3\nu_1(0) - 0.0113728\nu_5(0)^4 - 0.00149759\nu_5(0)^3 + 0.125857\nu_5(0)^2\nu_1(0) \\
&+ 1.19019\nu_5(0)^2 + 1.05873\nu_5(0)\nu_1(0) - 1.5701\nu_5(0) - 0.00130489\nu_1(0)^4 + 0.0456274\nu_1(0)^3 \\
&- 0.754842\nu_1(0)^2 + 1.74962\nu_1(0) - 0.00124248\nu_1(0)^3\nu_5(0)^2 - 0.00274613\nu_1(0)^2\nu_5(0)^3 \\
&- 0.00470875\nu_1(0)^2\nu_5(0)^2 + 0.00523158\nu_1(0)^3\nu_5(0) + 0.121085\nu_1(0)^2\nu_5(0) - 0.0017293\nu_1(0)\nu_5(0)^4 \\
&- 2.9969 \cdot 10^{-6}\nu_1(0)^4\nu_5(0)^2 + 7.3384 \cdot 10^{-5}\nu_1(0)^3\nu_5(0)^3 + 8.88018 \cdot 10^{-6}\nu_1(0)^4\nu_5(0) \\
&+ 0.000238133\nu_1(0)^2\nu_5(0)^4 + 0.000244127\nu_1(0)\nu_5(0)^5 - 0.000234529\nu_5(0)^5 + 8.23747 \cdot 10^{-5}\nu_5(0)^6 \\
&+ 2.10571 \cdot 10^{-7}\nu_5(0)^3\nu_1(0)^4 + 8.42285 \cdot 10^{-7}\nu_5(0)^4\nu_1(0)^3 + 1.26343 \cdot 10^{-6}\nu_5(0)^5\nu_1(0)^2 \\
&+ 8.42285 \cdot 10^{-7}\nu_5(0)^6\nu_1(0) + 2.10571 \cdot 10^{-7}\nu_5(0)^7 + 3.63985.
\end{aligned}$$

The optimal parameter values can be obtained by a direct maximization of $\mathcal{O}(\nu_1(0), \nu_5(0))$ plotted in Figure 2 of the main text. The optimal parameter values are given in this case by $\nu_1^*(0) = 1.0$ and $\nu_5^*(0) = 0.5$, see Figure 6, in full agreement with the solution of the forward-backward iteration scheme up to domain border perturbations due to the finite value of ϵ.

b. Description of popular heuristics used for comparisons on large real-world networks

In the main text, we described the performance of the DMP algorithm for the seeding problem on real-world networks of different nature, topology and size, and presented comparisons with well-preforming heuristics in the case of deterministic spreading in the SI model. Here we briefly describe the algorithms used for comparisons.

1. **Naive strategies**. As a reference, we present results obtained by a naive allocation of budget to randomly-selected nodes (**Random**) and a uniform allocation of the budget in a probabilistic way (**Uniform**).

2. **High degree adaptive (HDA)**. The algorithm is based on the idea that the best spreaders correspond to nodes with the highest degree [17], which is a valid and reasonable assumption for a large number of models, including the SI model considered here. The budget is iteratively attributed to the nodes with the current largest degree. Note the adaptive nature of the algorithm: once the node is selected, it is effectively removed from the network, and the degrees of all nodes

are recomputed accordingly; this strategy is much more efficient than the allocation of the budget to the high-degree nodes computed once and for all nodes in the original graph.

3. k-**shell decomposition**. In this algorithm, the influential spreaders are ranked according to their belonging to the k-core with maximum k. However, it has been noted in [22] that while this strategy is successful in identifying a single "influential" node, it performs badly when a group of nodes is selected, which is confirmed by our findings.

4. **Collective influence**. The relation between the destruction of the giant component and "optimal percolation" has been suggested in [25], where the authors have put forward another topological centrality measure called Collective Influence, defined as

$$\mathrm{CI}_l = (d_i - 1) \sum_{j \in \partial B(i,l)} (d_j - 1), \tag{B8}$$

where $\partial B(i, l)$ denotes the set of nodes at a distance l from node i. This topological characteristic results from mapping the spreading process asymptotically onto percolation [58] (in a particular instance of the Linear Threshold model [59]). The intuition behind this measure is that the "collective influence" of node i is not only given by its degree, but also by the contribution of the degrees of nodes at a certain distance from it. However, on many graphs the number of nodes in the set $\partial B(i, l)$ grows exponentially with l, which makes the computation of CI_l rather involved even for small values of l. Similarly to HDA, the algorithm is adaptive: after an allocation of resources to the node with the highest score, the measure is recomputed for all nodes. It is intuitive that in the case of a finite-time horizon objective, the best performance should be attained for $l \leq T$; indeed, we find that $l = 2$ realizes the best choice for the algorithm, with CI_4 leading to sub-optimal results in this near-deterministic spreading case.

Note that studying the minimization of the spread on networks is less interesting, because the optimal solution corresponds to activating nodes in clusters which are disconnected from the majority of nodes belonging to the giant component.

Appendix C: Offline and online mitigation of epidemic via optimal vaccination

Here we consider the problem of spread minimization, for example in the case of an undesired spreading process such as the propagation of epidemic. The minimization of the spread at time T is achieved by an effective distribution of vaccines to vulnerable susceptible nodes in a dynamical fashion. We use the variant of the generalized SIR model with a vaccination transition: the optimization parameters in this case are given by $\{\mu_i(t)\}_{i \in V}$ for $t \in [0, T-1]$, and we assume that the spontaneous transition to the I state is absent, meaning that $\nu_i(t) = 0$ for all i and t. Modifications with respect to the case considered previously in the Section B include:

1. We are interested in maximizing the objective $\sum_{i \in V} \left(1 - P_I^i(T)\right) = \sum_{i \in V} \left(P_S^i(T) + P_R^i(T)\right)$ at this time, so the objective has a different sign with respect to the previous case of the spread maximization;

2. Since now nodes can transit to the R state, we additionally need to keep track of the evolution of the quantities $P_R^i(t)$ and $P_R^{i \to j}(t)$;

3. Similarly to the formulation in the spread maximization, it is possible to include additional information on a list of target nodes which require priority protection, thus forcing the optimization to minimize the probability of infection on "key" nodes first;

4. In the setting of "online" mitigation of the spreading process one should reinitialize the optimization process at each time step once the new state of the network is available, getting a more accurate and updated account of the state of the system. One then solves the optimization problem of spread minimization at the horizon $t = T$ for the allocation of resources at the next time step only (seeding problem) using new data as initial conditions for the DMP equations. Running optimization for a few more time steps results in a vaccine distribution plan with forcasted estimates, which might be an important feature in realistic settings.

1. Lagrangian and forward-backward equations

The corresponding Lagrangian in the case of spreading minimization under vaccination constraints takes the following form:

$$\mathcal{L} = \underbrace{\sum_{i \in V} \left(P_S^i(T) + P_R^i(T) \right)}_{\mathcal{O}} + \underbrace{\sum_{t=0}^{T-1} \lambda_B^\mu(t) \left[\sum_{i \in V} \mu_i(t) - B_\mu(t) \right]}_{\mathcal{B}} + \epsilon \underbrace{\sum_{t=0}^{T-1} \sum_{i \in V} \left(\log \left[\mu_i(t) - \underline{\mu_i^t} \right] + \log \left[\overline{\mu_i^t} - \mu_i(t) \right] \right)}_{\mathcal{P}}$$

$$\left. \begin{aligned}
&+ \sum_{i \in V} \sum_{t=0}^{T-1} \lambda_i^S(t+1) \left[P_S^i(t+1) - P_S^i(t)(1 - \mu_i(t)) \prod_{k \in \partial i} \frac{\theta^{k \to i}(t+1)}{\theta^{k \to i}(t)} \right] \\
&+ \sum_{i \in V} \sum_{t=0}^{T-1} \lambda_i^R(t+1) \left[P_R^i(t+1) - P_R^i(t) - \mu_i(t) P_S^i(t) \right] \\
&+ \sum_{(ki) \in E} \sum_{t=0}^{T-1} \lambda_{k \to i}^S(t+1) \left[P_S^{k \to i}(t+1) - P_S^{k \to i}(t)(1 - \mu_k(t)) \prod_{l \in \partial k \backslash i} \frac{\theta^{l \to k}(t+1)}{\theta^{l \to k}(t)} \right] \\
&+ \sum_{(ki) \in E} \sum_{t=0}^{T-1} \lambda_{k \to i}^R(t+1) \left[P_R^{k \to i}(t+1) - P_R^{k \to i}(t) - \mu_k(t) P_S^{k \to i}(t) \right] \\
&+ \sum_{(ki) \in E} \sum_{t=0}^{T-1} \lambda_{k \to i}^\theta(t+1) \left[\theta^{k \to i}(t+1) - \theta^{k \to i}(t) + \alpha_{ki} \phi^{k \to i}(t) \right] \\
&+ \sum_{(ki) \in E} \sum_{t=0}^{T-1} \lambda_{k \to i}^\phi(t+1) \left[\phi^{k \to i}(t+1) - (1 - \alpha_{ki}) \phi^{k \to i}(t) - P_S^{k \to i}(t) + P_S^{k \to i}(t+1) - P_R^{k \to i}(t) + P_R^{k \to i}(t+1) \right]
\end{aligned} \right\} \mathcal{D}$$

$$\left. \begin{aligned}
&+ \sum_{i \in V} \lambda_i^S(0) \left[P_S^i(0) - \delta_{\sigma_i^0, S} \right] + \sum_{i \in V} \lambda_i^R(0) \left[P_R^i(0) - \delta_{\sigma_i^0, R} \right] + \sum_{(ki) \in E} \lambda_{k \to i}^S(0) \left[P_S^{k \to i}(0) - \delta_{\sigma_k^0, S} \right] \\
&+ \sum_{(ki) \in E} \lambda_{k \to i}^R(0) \left[P_R^{k \to i}(0) - \delta_{\sigma_k^0, R} \right] + \sum_{(ki) \in E} \lambda_{k \to i}^\theta(0) \left[\theta^{k \to i}(0) - 1 \right] + \sum_{(ki) \in E} \lambda_{k \to i}^\phi(0) \left[\phi^{k \to i}(0) - \delta_{\sigma_k^0, I} \right].
\end{aligned} \right\} \mathcal{I}$$

The variation of the Lagrangian with respect to the dual variables yield the forward DMP equations (A1)-(A8). The derivation of the backward equations follow the same principles as in the case of the SI model, with obvious modifications: change of the control parameters $\nu_i(t) \to \mu_i(t)$, additional equations

$$\partial \mathcal{L} / \partial P_R^i(t) = \mathbb{1}[t = T] + \lambda_i^R(t) - \lambda_i^R(t+1) \mathbb{1}[t \neq T] = 0, \tag{C1}$$

$$\partial \mathcal{L} / \partial P_R^{k \to i}(t) = \lambda_{k \to i}^R(t) - \lambda_{k \to i}^R(t+1) \mathbb{1}[t \neq T] + \lambda_{k \to i}^\phi(t) - \lambda_{k \to i}^\phi(t+1) \mathbb{1}[t \neq T] = 0, \tag{C2}$$

and additions to the previous backward equations due to the new terms

$$\Delta \partial \mathcal{L} / \partial P_S^i(t) = -\lambda_i^R(t+1) \mu_i(t) \mathbb{1}[t \neq T], \tag{C3}$$

$$\Delta \partial \mathcal{L} / \partial P_S^{k \to i}(t) = -\lambda_{k \to i}^R(t+1) \mu_k(t) \mathbb{1}[t \neq T], \tag{C4}$$

$$\Delta \partial \mathcal{L} / \partial \mu_i(t) = -\lambda_i^R(t+1) P_S^i(t) - \sum_{k \in \partial i} \lambda_{i \to k}^R(t+1) P_S^{i \to k}(t). \tag{C5}$$

2. Description of the data set used for tests

As a test example for the vaccination problem, we constructed a transportation network of busiest flight routes between major U.S. airports, extracted from the publicly available Bureau of Transportation Statistics data [55]. We used the data table providing the number of transported passengers by different companies between U.S. airports over the past several years. First of all, we extracted the sub-table of flights between 61 biggest airports in terms of the total number of emplaned passengers per year, including 30 "major" and 31 "largest" hubs according to the BTS classification. Then multiple entries corresponding to the same route have been aggregated, and the routes carrying less then 10% of the passengers transported by the busiest route have been pruned as less significant ones, primarily for rending the network reasonably sparse for a clear visualization. This resulted in a network with $M = |E| = 383$ edges. Finally, we have assigned the spreading couplings $\alpha_{ij\,(ij) \in E}$ proportionally to the number of carried passengers in such a way that the lightest route had the value $\alpha_{ij}^{\min} = 0.05$, and hence the busiest route received the value $\alpha_{ij}^{\max} \simeq 0.495$. This choice is based on a reasonable assumption that the probability of infection transmission along each link is proportional to the number of carried passengers on this route.

Appendix D: Case of continuous dynamics

The optimization procedure described in this work can be directly applied to the case of continuous dynamics, which might be more relevant in other applications. In the continuous case, the backward equations are obtained through the variation of the Lagrangian resulting in the continuous Euler-Lagrange equations. In this section, we illustrate the approach in the setting of maximizing the spread in the SI model using the continuous version of the DMP equations.

1. Maximization of spread at the time horizon T

In the continuous case, the marginal probability for node i to be in the state S at time t reads

$$P_S^i(t) = P_S^i(0) \left(\int_0^t dt' e^{-\nu_i(t')t'} \right) \prod_{k \in \partial i} \theta^{k \to i}(t). \tag{D1}$$

In the case of constant rates α_{ij}, we define the transmission function as $f_{ij}(t) = \alpha_{ij} e^{-\alpha_{ij}t}$. Then the functions $\theta^{i \to j}(t)$ are computed as follows [40]:

$$\theta^{i \to j}(t) = 1 - \int_0^t d\tau \, f_{ij}(\tau) \left[1 - P_S^i(0) \left(\int_0^{t-\tau} dt' e^{-\nu_i(t')t'} \right) \prod_{k \in \partial i \setminus j} \theta^{k \to i}(t - \tau) \right]$$

$$= e^{-\alpha_{ij}t} + P_S^i(0)\alpha_{ij}e^{-\alpha_{ij}t} \int_0^t d\tau \, e^{\alpha_{ij}\tau} \left(\int_0^\tau dt' e^{-\nu_i(t')t'} \right) \prod_{k \in \partial i \setminus j} \theta^{k \to i}(\tau). \tag{D2}$$

In order to compute the dynamic messages $\theta^{i \to j}(t)$, we can either integrate the Eq. (D2) numerically, or transform it into an ordinary differential equation by differentiating with respect to t:

$$\dot{\theta}^{i \to j}(t) = -\alpha_{ij}\theta^{i \to j}(t) + \alpha_{ij}P_S^i(0) \left(\int_0^t dt' e^{-\nu_i(t')t'} \right) \prod_{k \in \partial i \setminus j} \theta^{k \to i}(t), \tag{D3}$$

which represents the dynamics of $\theta^{i \to j}(t)$ (the dot notation represents $\frac{d}{dt}$) and can be solved numerically starting from initial conditions $\theta^{i \to j}(0) = 1$. To derive the dynamics of $P_S^i(t)$ one can differentiate (D1) to obtain:

$$\dot{P}_S^i(t) = P_S^i(0)e^{-\nu_i(t)t} \prod_{k \in \partial i} \theta^{k \to i}(t) + P_S^i(t) \sum_{k \in \partial i} \frac{\dot{\theta}^{k \to i}(t)}{\theta^{k \to i}(t)}. \tag{D4}$$

Maximizing $\mathcal{O} = \sum_{i \in V} P_I^i(T)$ is equivalent to minimizing

$$\hat{\mathcal{O}}(T) = \sum_{i \in V} P_S^i(T) = \int_0^T \sum_{i \in V} \dot{P}_S^i(t)dt \tag{D5}$$

under the constraints imposed by (D3), (D4) and the budget constraint (A10)

$$\sum_{i \in V} \nu_i(t) = B_\nu(t). \tag{D6}$$

All constraints will be imposed by the corresponding Lagrange multipliers. The Lagrangian to be extremized takes the form

$$\mathcal{L} = \int_0^T \left\{ \sum_{i \in V} \dot{P}_S^i(t) + \lambda_B^\nu(t) \left[\sum_{i \in V} \nu_i(t) - B_\nu(t) \right] \right.$$

$$+ \sum_{(i,j) \in E} \lambda_{i \to j}^\theta(t) \left[\dot{\theta}^{i \to j}(t) + \alpha_{ij} \theta^{i \to j}(t) - \alpha_{ij} P_S^i(0) \left(\int_0^t dt' e^{-\nu_i(t')t'} \right) \prod_{k \in \partial i \setminus j} \theta^{k \to i}(t) \right]$$

$$+ \left. \sum_{i \in V} \lambda_i^S(t) \left[\dot{P}_S^i(t) - P_S^i(0) e^{-\nu_i(t)t} \prod_{k \in \partial i} \theta^{k \to i}(t) - P_S^i(t) \sum_{k \in \partial i} \frac{\dot{\theta}^{k \to i}(t)}{\theta^{k \to i}(t)} \right] \right\} dt$$

$$= \int_0^T \left\{ \sum_{i \in V} \dot{P}_S^i(t) + \lambda_B^\nu(t) \left[\sum_{i \in V} \nu_i(t) - B_\nu(t) \right] \right.$$

$$+ \sum_{(i,j) \in E} \lambda_{i \to j}^\theta(t) \left[\dot{\theta}^{i \to j}(t) + \alpha_{ij} \theta^{i \to j}(t) - \alpha_{ij} \frac{P_S^i(t)}{\theta^{j \to i}(t)} \right]$$

$$+ \left. \sum_{i \in V} \lambda_i^S(t) \left[\dot{P}_S^i(t) - P_S^i(0) e^{-\nu_i(t)t} \prod_{k \in \partial i} \theta^{k \to i}(t) - P_S^i(t) \sum_{k \in \partial i} \frac{\dot{\theta}^{k \to i}(t)}{\theta^{k \to i}(t)} \right] \right\} dt \tag{D7}$$

Variational maximization of a Lagrangian

$$\mathcal{L} = \int_0^T F(t, y, \dot{y}) \, dt$$

where $F(t, y, \dot{y})$ is a function of some variable t, y and $\dot{y}$ its derivative with respect to t, is carried out by varying y with respect to a small perturbation $\epsilon \to 0$, throughout the t interval [60], to $y + \epsilon \delta y$. Extremizing $\mathcal{L}$ by setting $\frac{d\mathcal{L}}{d\epsilon} = 0$ and using integration by parts one obtains the optimization condition

$$\frac{d\mathcal{L}}{d\epsilon} \bigg|_{\epsilon=0} = \int_0^T \left[\frac{\partial F}{\partial y} - \frac{d}{dt} \frac{\partial F}{\partial \dot{y}} \right] \delta y \, dt + \left[\frac{\partial F}{\partial \dot{y}} \right]_0^T = 0 \ .$$

Both the Euler-Lagrange equation $\left[\frac{\partial F}{\partial y} - \frac{d}{dt} \frac{\partial F}{\partial \dot{y}} \right]$ and the boundary conditions $\left[\frac{\partial F}{\partial \dot{y}} \right]_0^T$ should be zero.

Varying (D7) with respect to $\delta P_S^i(t)$ and $\delta \theta^{i \to j}(t)$ results in the following differential equations:

$$\dot{\lambda}_i^S(t) = -\sum_{j \in \partial i} \frac{\lambda_{i \to j}^\theta \alpha_{ij}}{\theta^{j \to i}(t)} - \lambda_i^S(t) \sum_{k \in \partial i} \frac{\dot{\theta}^{k \to i}(t)}{\theta^{k \to i}(t)} - \delta_{t,0} \lambda_i^S(0) \prod_{k \in \partial i} \theta^{k \to i}(0), \tag{D8}$$

$$\dot{\lambda}_{i \to j}^\theta(t) = \lambda_{i \to j}^\theta(t) \alpha_{ij} + \lambda_{j \to i}^\theta(t) \alpha_{ji} \frac{P_S^j(t)}{[\theta^{i \to j}(t)]^2} \tag{D9}$$

$$- \lambda_j^S(t) \left\{ P_S^j(0) e^{-\nu_j(t)t} \prod_{k \in \partial j / i} \theta^{k \to j}(t) - \frac{\dot{P}_S^j(t)}{\theta^{i \to j}(t)} \right\} + \dot{\lambda}_j^S(t) \frac{P_S^j(t)}{\theta^{i \to j}(t)} \ ,$$

$$= \lambda_{i \to j}^\theta(t) \alpha_{ij} - \sum_{k \in \partial j / i} \frac{\lambda_{j \to k}^\theta(t) \alpha_{jk} P_S^j(t)}{\theta^{k \to j}(t) \theta^{i \to j}(t)} - \delta_{t,0} \lambda_j^S(0) P_S^j(t) \prod_{k \in \partial j / i} \theta^{k \to i}(0),$$

where the expression for $\dot{\theta}^{k \to i}(t)$ in both equations (D8) and (D9) are calculated on the basis of (D3) and the expressions for $\dot{\lambda}_j^S(t)$ and $\dot{P}_S^j(t)$ in (D9) are taken from (D8) and (D4), respectively, and were employed in the simplifications of (D9). These equations will be integrated back on the basis of end points.

Varying (D7) with respect to $\delta \nu_i(t)$ one obtains:

$$\lambda_B^\nu(t) = -t \lambda_i^S(t) P_S^i(0) e^{-\nu_i(t)t} \prod_{k \in \partial i} \theta^{k \to i}(t) \ , \tag{D10}$$

which can be re-written as

$$\nu_i(t) = -\frac{1}{t} \ln \left[-\frac{\lambda_B^\nu(t)}{t \lambda_i^S(t) P_S^i(0) \prod_{k \in \partial i} \theta^{k \to i}(t)} \right] \ . \tag{D11}$$

Using the condition $\sum_{i \in V} \nu_i(t) = B_\nu(t)$ we obtain

$$B_\nu(t) = -\frac{N}{t} \ln \lambda_B^\nu(t) + \frac{1}{t} \sum_{i \in V} \ln \left[-t \lambda_i^S(t) P_S^i(0) \prod_{k \in \partial i} \theta^{k \to i}(t) \right] \ . \tag{D12}$$

That leads to a straightforward solution for $\lambda_B^\nu(t)$

$$\lambda_B^\nu(t) = \exp\left\{ \frac{1}{N} \sum_{i \in V} \ln\left[-t \lambda_i^S(t) P_S^i(0) \prod_{k \in \partial i} \theta^{k \to i}(t) \right] - \frac{t}{N} B_\nu(t) \right\} , \tag{D13}$$

from which the $\nu_i(t)$ values can be calculated.

Remark: An alternative path to the derivation of (D11) and (D13) consists in directly using the normalized representation

$$\nu_i(t) = \frac{e^{-\beta_i(t)}}{\sum_{j \in V} e^{-\beta_j(t)}} B_\nu(t) .$$

instead of enforcing the budget constraint with the Lagrange multiplier $\lambda_B^\nu(t)$ in (D7). For the optimization with respect to $\beta_i(t)$ in this case it is then possible to use

$$\frac{\partial}{\partial \beta_k} = -\sum_{j \in V} [\delta_{jk} \nu_k(t) - \nu_k(t) \nu_j(t)] \frac{\partial}{\partial \nu_j} .$$

Optimization with respect to $\beta_i(t)$ gives

$$\nu_i(t) = \frac{\lambda_i^S(t) P_S^i(0) e^{-\nu_i(t) t} \prod_{k \in \partial i} \theta^{k \to i}(t)}{\sum_{j \in V} \lambda_j^S(t) P_S^j(0) e^{-\nu_j(t) t} \prod_{k \in \partial j} \theta^{k \to j}(t)} B_\nu(t) , \tag{D14}$$

which is equivalent to (D11) and (D13). ∎

Additionally, one should enforce the boundary conditions in the two sets of equations:

$$0 = \delta P_S^i(T) \left[1 + \lambda_i^S(T) \right] - \delta P_S^i(0) \left[1 + \lambda_i^S(0) \right] \tag{D15}$$

$$0 = \delta \theta^{i \to j}(T) \left[\lambda_{i \to j}^\theta(T) - \frac{P_S^j(T) \lambda_j^S(T)}{\theta^{j \to i}(T)} \right] - \delta \theta^{i \to j}(0) \left[\lambda_{i \to j}^\theta(0) - \frac{P_S^j(0) \lambda_j^S(0)}{\theta^{j \to i}(0)} \right] . \tag{D16}$$

Since $\delta P_S^i(0) = \delta \theta^{i \to j}(0) = 0 \ \forall i, j \in V$ it provides the end conditions:

$$\lambda_i^S(T) = -1 \tag{D17}$$

$$\lambda_{i \to j}^\theta(T) = \frac{P_S^j(T)}{\theta^{j \to i}(T)} . \tag{D18}$$

Note that $\lambda_i^S(T) = -1$ is a result of the minimization of $\hat{\mathcal{O}}(T)$; a maximization of $\mathcal{O}$ would provide a boundary condition of $\lambda_i^S(T) = 1$.

The optimization process should be carried out as follows:

1. Using the initial conditions $\theta^{i \to j}(0) = 1 \ \forall i$ and $P_S^i(0)$ according to the case studied, and some valid initial set of $\nu_i(t)$ (can be uniform), one can solve forward the equations (D3) and (D4) to obtain (and register) values throughout the dynamics $t = 0 \to T$.

2. Using the boundary conditions (D18) and (D17) one solve backward (D8) and (D9) and updates the values of $\nu_i(t)$ according to (D13) and (D11).

3. The process is repeated until it converges and the final $\nu_i(t)$ values represent the solution.

2. Continuous dynamics with targeted and accessible nodes

In this variant of the problem one targets specific nodes $i \in U$ where $U \subseteq V$ is the subset of all nodes V, aiming to maximize the impact at predefined times t_i, which may be different for each of the nodes. We will define $T \equiv \max_{i \in U} t_i$. We also assume one has access to a subset on the nodes $i \in W$ where $W \subseteq V$ and $W \cap U = \emptyset$. Again, maximizing $\mathcal{O}$ is equivalent to minimizing

$$\hat{\mathcal{O}} = \sum_{i \in U} P_S^i(t_i) = \sum_{i \in U} \int_0^{t_i} \dot{P}_S^i(t) \, dt . \tag{D19}$$

Since the budget for nodes $i \notin W$ is zero by definition $\nu_i(t) = 0, \forall i \notin W, \forall t$, and the cost constraint corresponding to (A10) becomes

$$\sum_{i \in W} \nu_i(t) = B_\nu(t) , \tag{D20}$$

where we can use the normalized representation where $\beta_i(t) = -\infty, \forall i \notin W, \forall t$,

$$\nu_i(t) = \frac{e^{-\beta_i(t)}}{\sum_{j \in V} e^{-\beta_j(t)}} B_\nu(t) \ .$$

The Lagrangian to be extremized takes the form (the budget constraint on accessible nodes (D19) is embedded in the corresponding ν variables through β, as remarked in Sec. D 1

$$
\begin{aligned}
\mathcal{L} &= \sum_{i \in U} \int_0^{t_i} \dot{P}_S^i(t) \, dt \\
&+ \int_0^T \Bigg\{ \sum_{(i,j) \in E} \lambda_{i \to j}^\theta(t) \left[\dot{\theta}^{i \to j}(t) + \alpha_{ij} \theta^{i \to j}(t) - \alpha_{ij} P_S^i(0) \left(\int_0^t dt' e^{-\nu_i(t')t'} \right) \prod_{k \in \partial i \setminus j} \theta^{k \to i}(t) \right] \\
&+ \sum_{i \in V} \lambda_i^S(t) \left[\dot{P}_S^i(t) - P_S^i(0) e^{-\nu_i(t)t} \prod_{k \in \partial i} \theta^{k \to i}(t) - P_S^i(t) \sum_{k \in \partial i} \frac{\dot{\theta}^{k \to i}(t)}{\theta^{k \to i}(t)} \right] \Bigg\} \, dt \\
&= \sum_{i \in U} \int_0^{t_i} \dot{P}_S^i(t) \, dt \\
&+ \int_0^T \Bigg\{ \sum_{(i,j) \in E} \lambda_{i \to j}^\theta(t) \left[\dot{\theta}^{i \to j}(t) + \alpha_{ij} \theta^{i \to j}(t) - \alpha_{ij} \frac{P_S^i(t)}{\theta^{j \to i}(t)} \right] \\
&+ \sum_{i \in V} \lambda_i^S(t) \left[\dot{P}_S^i(t) - P_S^i(0) e^{-\nu_i(t)t} \prod_{k \in \partial i} \theta^{k \to i}(t) - P_S^i(t) \sum_{k \in \partial i} \frac{\dot{\theta}^{k \to i}(t)}{\theta^{k \to i}(t)} \right] \Bigg\} \, dt
\end{aligned}
\tag{D21}
$$

Alternatively, the first term can be written as

$$\int_0^T \sum_{i \in U} \dot{P}_S^i(t) \Theta(t_i - t) \, dt.$$

Since the first term does not contribute to the Euler-Lagrange equation, varying (D21) with respect to $\delta P_S^i(t)$, $\delta \theta^{i \to j}(t)$ and $\delta \nu_i(t)$ results in the same differential equations for $\dot{\lambda}_i^S(t)$ (D8) and $\dot{\lambda}_{ij}(t)$ (D9).

Optimization with respect to $\beta_i(t)$ or $\nu_i(t)$ gives a similar expression for $i \in W$ (D14) but $\nu_i(t) = 0, \forall i \notin W$ and $\forall t$. Additionally, one should enforce the boundary conditions in the two sets of equations (D15) and (D16). For $j \in U$ (D15) and (D16) are the same as in the non-targeted case, but for $j \notin U$ there is no constant (of value 1) in (D15) leading to:

$$\lambda_i^S(T) = \begin{cases} -1 & \forall \, i \notin U \\ 0 & \forall \, i \notin U \end{cases} \tag{D22}$$

$$\lambda_{i \to j}^\theta(t_j) = \begin{cases} 0 & \forall \, i \notin U \\ \frac{P_S^j(t_j)}{\theta^{j \to i}(t_j)} & \forall \, j \in U \end{cases} . \tag{D23}$$

The optimization process should be carried out as before.

[1] Anderson, R. M., May, R. M. & Anderson, B. *Infectious diseases of humans: dynamics and control*, vol. 28 (Wiley Online Library, 1992).

[2] Boccaletti, Stefano and Latora, Vito and Moreno, Yamir and Chavez, Martin and Hwang, D.-U. Complex networks: Structure and dynamics. *Physics reports* **424**, 175–308 (2006).

[3] Rogers, E. M. *Diffusion of innovations* (Simon and Schuster, 2010).

[4] Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Reviews of modern physics* **87**, 925 (2015).

[5] Federal Energy Regulatory Commission and North American Electric Reliability Corporation Staff. Arizona - southern california outages on september 8, 2011: causes and recommendations. Tech. Rep. (2012).

[6] Dawood, F. S. *et al.* Estimated global mortality associated with the first 12 months of 2009 pandemic influenza a h1n1 virus circulation: a modelling study. *The Lancet infectious diseases* **12**, 687–695 (2012).

[7] U.S. Government Accountability Office. Financial regulatory reform: financial crisis losses and potential impacts of the dodd-frank act. Tech. Rep. (2012).

[8] Lokhov, A. Y., Lemons, N., McAndrew, T. C., Hagberg,

A. & Backhaus, S. Detection of cyber-physical faults and intrusions from physical correlations. In *ACM SIGKDD Workshop on Outlier Definition, Detection and Description on Demand (ODD 4.0)* (2016).

[9] Als association. `http://www.alsa.org/fight-als/ice-bucket-challenge.html`.

[10] Rutledge, P. How obama won the social media battle in the 2012 presidential campaign. *The National Psychologist* (2013).

[11] Epstein, R. & Robertson, R. E. The search engine manipulation effect (seme) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences* **112**, E4512–E4521 (2015).

[12] Margetts, H., John, P., Hale, S. & Yasseri, T. *Political Turbulence: How Social Media Shape Collective Action* (Princeton University Press, 2015).

[13] Domingos, P. & Richardson, M. Mining the network value of customers. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, 57–66 (ACM, 2001).

[14] Martin, K. R. *et al.* Computational model for autophagic vesicle dynamics in single cells. *Autophagy* **9**, 74–92 (2013).

[15] Csermely, P., Korcsmáros, T., Kiss, H. J., London, G. & Nussinov, R. Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review. *Pharmacology & therapeutics* **138**, 333–408 (2013).

[16] Jones, S. Army revives chindits as 'facebook warriors' for smart battle. *Financial Times* (2015).

[17] Pastor-Satorras, R. & Vespignani, A. Immunization of complex networks. *Phys. Rev. E* **65**, 036104 (2002).

[18] Cohen, R., Havlin, S. & Ben-Avraham, D. Efficient immunization strategies for computer networks and populations. *Phys. Rev. Lett.* **91**, 247901 (2003).

[19] Holme, P., Kim, B. J., Yoon, C. N. & Han, S. K. Attack vulnerability of complex networks. *Phys. Rev. E* **65**, 056109 (2002).

[20] Holme, P. Efficient local strategies for vaccination and network attack. *EPL (Europhysics Letters)* **68**, 908 (2004).

[21] Chen, Y., Paul, G., Havlin, S., Liljeros, F. & Stanley, H. E. Finding a better immunization strategy. *Phys. Rev. Lett.* **101**, 058701 (2008).

[22] Kitsak, M. *et al.* Identification of influential spreaders in complex networks. *Nature Physics* **6**, 888–893 (2010).

[23] Borge-Holthoefer, J. & Moreno, Y. Absence of influential spreaders in rumor dynamics. *Phys. Rev. E* **85**, 026116 (2012).

[24] Hébert-Dufresne, L., Allard, A., Young, J.-G. & Dubé, L. J. Global efficiency of local immunization on complex networks. *Scientific reports* **3** (2013).

[25] Morone, F. & Makse, H. A. Influence maximization in complex networks through optimal percolation. *Nature* (2015).

[26] Mugisha, S. & Zhou, H.-J. Identifying optimal targets of network attack by belief propagation. *arXiv preprint arXiv:1603.05781* (2016).

[27] Braunstein, A., Dall'Asta, L., Semerjian, G. & Zdeborová, L. Network dismantling. *arXiv preprint arXiv:1603.08883* (2016).

[28] Altarelli, F., Braunstein, A., DallAsta, L., Wakeling, J. R. & Zecchina, R. Containing epidemic outbreaks by message-passing techniques. *Phys. Rev. X* **4**, 021024 (2014).

[29] Altarelli, F., Braunstein, A., DallAsta, L. & Zecchina, R. Optimizing spread dynamics on graphs by message passing. *Journal of Statistical Mechanics: Theory and Experiment* **2013**, P09011 (2013).

[30] Guggiola, A. & Semerjian, G. Minimal contagious sets in random regular graphs. *Journal of Statistical Physics* **158**, 300–358 (2015).

[31] Kempe, D., Kleinberg, J. & Tardos, É. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 137–146 (ACM, 2003).

[32] Chen, W., Lakshmanan, L. V. & Castillo, C. Information and influence propagation in social networks. *Synthesis Lectures on Data Management* **5**, 1–177 (2013).

[33] Du, N., Song, L., Gomez-Rodriguez, M. & Zha, H. Scalable influence estimation in continuous-time diffusion networks. In *Advances in Neural Information Processing Systems*, 3147–3155 (2013).

[34] Nowzari, C., Preciado, V. M. & Pappas, G. J. Analysis and control of epidemics: A survey of spreading processes on complex networks. *IEEE Control Systems* **36**, 26–46 (2016).

[35] Borgs, C., Chayes, J., Ganesh, A. & Saberi, A. How to distribute antidote to control epidemics. *Random Structures & Algorithms* **37**, 204–222 (2010).

[36] Nian, F. & Wang, X. Efficient immunization strategies on complex networks. *Journal of theoretical biology* **264**, 77–83 (2010).

[37] Scaman, K., Kalogeratos, A. & Vayatis, N. A greedy approach for dynamic control of diffusion processes in networks. In *Tools with Artificial Intelligence (ICTAI), 2015 IEEE 27th International Conference on*, 652–659 (IEEE, 2015).

[38] Drakopoulos, K., Ozdaglar, A. & Tsitsiklis, J. N. An efficient curing policy for epidemics on graphs. *IEEE Transactions on Network Science and Engineering* **1**, 67–75 (2014).

[39] Shapiro, M. & Delgado-Eckert, E. Finding the probability of infection in an sir network is np-hard. *Mathematical biosciences* **240**, 77–84 (2012).

[40] Karrer, B. & Newman, M. E. J. Message passing approach for general epidemic models. *Phys. Rev. E* **82**, 016101 (2010).

[41] Shrestha, M. & Moore, C. Message-passing approach for threshold models of behavior in networks. *Phys. Rev. E* **89**, 022805 (2014).

[42] Lokhov, A. Y., Mézard, M. & Zdeborová, L. Dynamic message-passing equations for models with unidirectional dynamics. *Phys. Rev. E* **91**, 012811 (2015).

[43] Le Cun, Y., Touresky, D., Hinton, G. & Sejnowski, T. A theoretical framework for back-propagation. In *The Connectionist Models Summer School*, vol. 1, 21–28 (1988).

[44] Saad, D. & Rattray, M. Globally optimal parameters for on-line learning in multilayer neural networks. *Phys. Rev. Lett.* **79**, 2578 (1997).

[45] De Domenico, M., Porter, M. A. & Arenas, A. Muxviz: a tool for multilayer analysis and visualization of networks. *Journal of Complex Networks* cnu038 (2014).

[46] Krebs, V. E. Mapping networks of terrorist cells. *Connections* **24**, 43–52 (2002).

[47] Pinto, P. C., Thiran, P. & Vetterli, M. Locating the

source of diffusion in large-scale networks. *Phys. Rev. Lett.* **109**, 068702 (2012).

[48] Lokhov, A. Y., Mézard, M., Ohta, H. & Zdeborová, L. Inferring the origin of an epidemic with a dynamic message-passing algorithm. *Phys. Rev. E* **90**, 012801 (2014).

[49] Altarelli, F., Braunstein, A., Dall'Asta, L., Lage-Castellanos, A. & Zecchina, R. Bayesian inference of epidemics on networks via belief propagation. *Phys. Rev. Lett.* **112**, 118701 (2014).

[50] Doreian, P. & Mrvar, A. A partitioning approach to structural balance. *Social networks* **18**, 149–168 (1996).

[51] Šubelj, L. & Bajec, M. Robust network community detection using balanced propagation. *The European Physical Journal B* **81**, 353–362 (2011).

[52] Bu, D. *et al.* Topological structure analysis of the protein–protein interaction network in budding yeast. *Nucleic acids research* **31**, 2443–2450 (2003).

[53] Leskovec, J., Kleinberg, J. & Faloutsos, C. Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data (TKDD)* **1**, 2 (2007).

[54] Boldi, P., Codenotti, B., Santini, M. & Vigna, S. Ubi-Crawler: A scalable fully distributed web crawler. *Software: Practice & Experience* **34**, 711–726 (2004).

[55] Bureau of transportation statistics. `http://www.rita.dot.gov/bts/`.

[56] AJ, T., DJ, R. & SI, H. Global transport networks and infectious disease spread. *Advances in Parasitology* **62**, 293–343 (2006).

[57] Lokhov, A. Y. & Misiakiewicz, T. Efficient reconstruction of transmission probabilities in a spreading process from partial observations. *arXiv preprint arXiv:1509.06893* (2015).

[58] Newman, M. E. Spread of epidemic disease on networks. *Phys. Rev. E* **66**, 016128 (2002).

[59] Granovetter, M. Threshold models of collective behavior. *American journal of sociology* 1420–1443 (1978).

[60] Hildebrand, F. B. Methods of applied mathematics prentice-hall. *New Jersey. pp* 36–122 (1965).