Optimal control of networks in the presence of attackers and defenders

Ishan Kafle, Sudarshan Bartaula, Afroza Shirin, Isaac Klickstein, Pankaz Das, and Francesco Sorrentino

Citation: Chaos **28**, 051103 (2018); doi: 10.1063/1.5030899
View online: https://doi.org/10.1063/1.5030899
View Table of Contents: http://aip.scitation.org/toc/cha/28/5
Published by the American Institute of Physics

Optimal control of networks in the presence of attackers and defenders

Ishan Kafle,[1,a)] Sudarshan Bartaula,[1,b)] Afroza Shirin,[1,c)] Isaac Klickstein,[1,d)] Pankaz Das,[2,e)] and Francesco Sorrentino[1,f)]
[1]*Department of Mechanical Engineering, University of New Mexico, Albuquerque, New Mexico 87131, USA*
[2]*Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, New Mexico 87131, USA*

(Received 26 March 2018; accepted 19 April 2018; published online 15 May 2018)

We consider the problem of a dynamical network whose dynamics is subject to external perturbations ("attacks") locally applied at a subset of the network nodes. We assume that the network has an ability to defend itself against attacks with appropriate countermeasures, which we model as actuators located at (another) subset of the network nodes. We derive the optimal defense strategy as an optimal control problem. We see that the network topology as well as the distribution of *attackers* and *defenders* over the network affect the optimal control solution and the minimum control energy. We study the optimal control defense strategy for several network topologies, including chain networks, star networks, ring networks, and scale free networks. *Published by AIP Publishing.* https://doi.org/10.1063/1.5030899

Optimal control of networks is an area of recent interest in the literature, where focus has been placed on how the network topology and the position of *driver* and *target* nodes affect the optimal solution. Here, we study the different but related problem of optimally controlling a network under attack. We investigate the role of the network topology as well as of the distribution of *attackers* and *defenders* over the network. Some of our results are counter-intuitive, as we find that for small chain networks, star networks, and ring networks, the distance between a single attacker and a single defender is not the key factor that determines the minimum control energy. We also consider the case of a large scale-free network in the presence of a single attacker and multiple defenders for which we see that the minimum control energy varies over different orders of magnitude as the position of the attacker is changed over the network. For this case, we observe that the minimum distance between the attacker node and the defender nodes is a good predictor of the *strength* of an attack.

I. MODEL

Most infrastructure systems are networked by design, such as power grids,[1] road systems,[2] telecommunications,[3] water and sewer systems,[4] and many others. These networked systems are prone to disruption by either natural causes, such as extreme weather events and aging equipment, or purposeful attack, such as terrorism.[5,6] In power grid systems, small local failures have been known to cascade to blackouts affecting large swaths of a state or a country.[7] In a road system, small incidents can lead to large scale congestion.[8] Attacks on networked systems can be either structural, where links in the underlying graph are damaged or destroyed, or dynamical, where a disruptive term is added to the dynamical equations that govern the behavior of the system. While our approach can be extended to encompass both structural and dynamical attacks, for the sake of simplicity, here we focus on dynamical attacks. Some examples of dynamical attacks on networked systems are pollutants introduced in a hydraulic network[9] or the spreading of viruses in networked computers.[10]

We examine the behavior of both simple and complex dynamical networks, when they are attacked by one or more external signals which perturb the dynamics of the network nodes. To illustrate this situation, a ten node network where three of the nodes are under attack is shown in Fig. 1. We assume that the networks at hand have an ability to defend themselves against attacks with appropriate countermeasures, which we model as actuators located at a subset of the nodes in the network. In Fig. 1(e), nodes 1, 3, and 10 are attached to actuators, and so these nodes we define as *defenders* (equivalently *driver nodes* as they are defined in much of the complex network literature[11]). Defenders can take many different forms in the networked systems described above, such as traffic signals and GPS routing in road networks, purposely tripping lines in a power grid in case of shedding, or quarantining a portion of a computer network when attacked by viruses.

The network dynamics is described by a linear model

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + H\mathbf{w}(t) + B\mathbf{u}(t), \tag{1}$$

where $\mathbf{x}(t) = [x_1(t), .., x_n(t)]$ is the $n \times 1$ time-varying state vector, $\mathbf{u}(t) = [u_1(t), .., u_m(t)]$ is the $m \times 1$ time-varying control input vector, and $\mathbf{w}(t) = [w_1(t), .., w_q(t)]$ is the $q \times 1$ time-varying vector representing the attackers. Hereafter, we design the control input vector using the fixed-end point minimum energy control problem for a system described by the

[a)]Electronic mail: ikafle@unm.edu.
[b)]Electronic mail: sbartaula@unm.edu.
[c)]Electronic mail: ashirin@unm.edu.
[d)]Electronic mail: iklick@unm.edu.
[e)]Electronic mail: pankazdas@unm.edu.
[f)]Electronic mail: fsorrent@unm.edu.

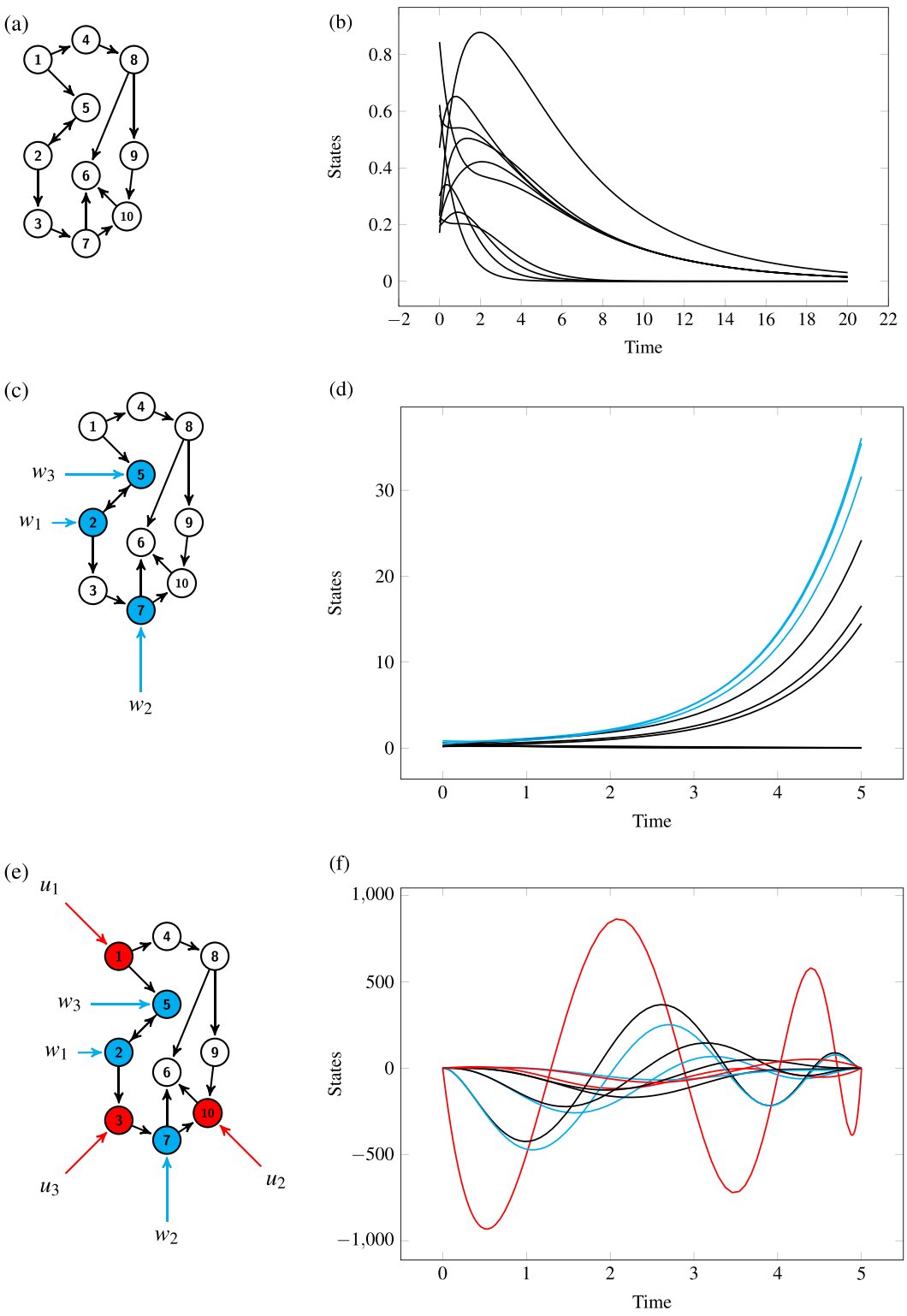

FIG. 1. (a) Network under no attack. (b) Time evolution of the network nodes under no attack. (c) Same network as in (a), with attackers located at nodes 2, 5, and 7. (d) Time evolution of the network nodes under attack. (e) Same network as in (c) with defenders located at nodes 1, 3, and 10. (f) Time evolution of the network nodes for the case that the network is attacked [same as in (d)] but a response from the defender nodes is also present.

linear dynamics shown in Eq. (1). Here, $A = \{a_{ij}\}$ is a square $n \times n$ real adjacency matrix which has non-zero elements a_{ij} if node i receives a signal from node j and is 0 otherwise. The $n \times m$ matrix B is the control input matrix and describes how the control inputs are connected to the nodes, i.e., the location of the defenders, namely, B_{ij} is different from zero if the control input j is attached to node i and is zero otherwise. The matrix H models how the attackers affect the network nodes, namely, H_{ij} is different from zero if attacker j is active on node i and is zero otherwise. The matrix A is Hurwitz and therefore by setting $\mathbf{w} = \mathbf{0}$ and $\mathbf{u} = \mathbf{0}$, the system asymptotically approaches the origin of state space, which represents the nominal healthy condition for the system.

In Figs. 1(a) and 1(b), we show a network which has stabilizing self-loops so that the adjacency matrix is Hurwitz. This ensures that after any perturbation of the states away from the origin, the states will return to the origin. Next, we add external attackers to the system attached to nodes 2, 5, and 7. In Fig. 1, we color nodes 2, 5, and 7 cyan as they are the attacked nodes. We assume the dynamics of each attacker is $w_j(t) = e^t w_j(0)$, where the initial conditions $w_j(0)$ are chosen as uniformly distributed random numbers in the interval (0, 1). As shown in Fig. 1(d), the time evolution of the states of those nodes directly attacked, and any nodes downstream such as nodes 3, 6, and 10 are now diverging. On the other hand, any nodes upstream of the attackers are not affected by the attack and so they will converge to the origin. To counter the attacks, we add external control inputs attached to the red nodes 1, 3, and 10, which we call defender nodes. Thanks to the control action exerted by the defender nodes, the attack can be mitigated and now all nodes return to the origin again. The particular control strategy implemented at the defender nodes to counter the attack will be described in what follows.

As a reference example, in this paper we consider a power grid, with nodes representing buses and edges representing transmission lines.[12] Both generation and a load can be present at each bus. We assume that some of the loads are vulnerable to attacks, in which case ancillary generation at the other nodes can be used to mitigate the effects of the attack. As explained in Sec. III, the dynamics of a power grid can be cast in the form of Eq. (1),

$$\begin{bmatrix} \dot{\boldsymbol{\delta}} \\ \dot{\boldsymbol{\theta}} \\ \dot{\boldsymbol{\omega}} \end{bmatrix} = A \begin{bmatrix} \boldsymbol{\delta} \\ \boldsymbol{\theta} \\ \boldsymbol{\omega} \end{bmatrix} + H \begin{bmatrix} 0 \\ \boldsymbol{P}^L \\ 0 \end{bmatrix} + B \begin{bmatrix} 0 \\ 0 \\ \boldsymbol{P}^{M'} \end{bmatrix}, \tag{2}$$

where the vector $\boldsymbol{\delta} = [\delta_1, \ldots, \delta_n,]$ contains information on the voltage phase angles at generator buses, the vector $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_n]$ describes the voltage phase angles at load buses, and the vector $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_n,]$ represents the frequency deviation at generator buses. The vector $\boldsymbol{P}^L = [P^L_1, \ldots, P^L_n]$ contains information on the power consumption at the load buses and the vector $\boldsymbol{P}^{M'} = [P^{M'}_1, \ldots, P^{M'}_n]$ represents the ancillary power generation (for more details, see Sec. III.)

In what follows we introduce the strong assumption that a known model for the attackers exists. This assumption could be more or less unrealistic depending on the application to which we are applying this methodology; however, our results are general as they can be applied to a variety of models for the attackers' strategy and as we will see, they are to some extent independent of the attackers' specific strategy. The type of attack strategies we consider is either one of the following functions: (i) constant, (ii) linearly increasing, or (iii) exponentially increasing, which can be modeled as

$$\dot{w}_i = s_i w_i + r_i, \tag{3}$$

where s_i and r_i are constants. We consider the following three cases:

(i) if $s_i = 0$ and $r_i = 0$ then the attack strategy is constant.
(ii) if $s_i = 0$ and $r_i > 0$ then the attack strategy is linearly increasing.
(iii) if $s_i > 0$ and $r_i = 0$ then the attack strategy is exponentially increasing.

By incorporating the model for the attackers' behavior, we can rewrite Eq. (1) as follows:

$$\dot{\tilde{\mathbf{x}}}(t) = \tilde{A}\tilde{\mathbf{x}}(t) + \tilde{\mathbf{r}} + \tilde{B}\mathbf{u}(t), \tag{4}$$

$$\mathbf{y}(t) = C\tilde{\mathbf{x}}(t), \tag{5}$$

where

$$\tilde{A} = \begin{bmatrix} A & \vdots & H \\ \cdots & \cdots & \cdots \\ 0 & \vdots & S \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} B \\ \cdots \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} I_n & \vdots & 0 \end{bmatrix},$$

$$\text{and } \tilde{\mathbf{r}} = \begin{bmatrix} 0 \\ \cdots \\ \mathbf{r} \end{bmatrix}. \tag{6}$$

Here, $\tilde{\mathbf{x}} = [\mathbf{x}^T, \mathbf{w}^T]^T$ is the $n+q$ vector containing the states of the network nodes and attackers, the behavior of which is assumed to be known, $\mathbf{y}(t) = [y_1(t), .., y_n(t)]$ is the $n \times 1$ time-varying vector of outputs, $S = \text{diag}\{s_1, .., s_q\}$ is the diagonal matrix that contains information on the attackers strengths, and the vector $\mathbf{r} = [r_1, \ldots, r_q]$ describes the attackers strategies [see Eq. (3)]. The matrix $\tilde{A}$ now has a block triangular structure and is non-Hurwitz, due to the attackers' dynamics. The matrix C relates the outputs $\mathbf{y}(t)$ to the state $\tilde{\mathbf{x}}(t)$. In this particular case, $\mathbf{y}$(t) coincides with $\mathbf{x}(t)$ in Eq. (1), i.e., it selects the states of the nodes but not those of the attackers.

When $\mathbf{u} = 0$, the time evolution of the network nodes deviates from the origin due to the influence of the attackers. The question we will try to address is the following: how can we design an optimal control input that in the presence of an attack, will set the state $\mathbf{x}(t_f) = 0$ at some preassigned time t_f. We assume that all the attacks take place simultaneously in the interval $[0, t_f]$. The time t_f can be thought of as the required time to neutralize the attackers, so that for $t > t_f$, the network has returned to its healthy state and the control action is no longer needed anymore. Here, without loss of generality, we assume the optimal control input to be the one that minimizes the energy function

$$E = \int_0^{t_f} \mathbf{u}^T(t)\mathbf{u}(t)dt. \tag{7}$$

The dynamics is linear and our objective function is quadratic (minimum energy). Therefore, applying the theory of linear quadratic optimal control (LQR),[13–15] the control input $\mathbf{u}^*(t)$ that satisfies the constraints and minimizes the control energy is[16]

$$\mathbf{u}^*(t) = B^T e^{\tilde{A}^T(t_f - t_0)} C^T (CWC^T)^{-1} \times \left[\mathbf{y}_f - Ce^{\tilde{A}^T(t_f - t_0)}\mathbf{x}_0 - CF(t_f)\tilde{\mathbf{r}}\right], \tag{8}$$

where $F(t_f) = \int_{t_0}^{t_f} e^{\tilde{A}(t_f - \tau)} d\tau$. The corresponding optimal energy is $E^* = \int_{t_0}^{t_f} \mathbf{u}^{*T}(t)\mathbf{u}^*(t)dt$. First, we define the controllability Gramian as a real, symmetric, semi-positive definite matrix

$$W = \int_{t_0}^{t_f} e^{\tilde{A}(t_f-\tau)}\tilde{B}\tilde{B}^T e^{\tilde{A}^T(t_f-\tau)}dt. \tag{9}$$

Following Ref. 16, the minimum control energy can be computed and is equal to

$$\begin{aligned} E^* &= (\mathbf{y}_f - Ce^{\tilde{A}(t_f-\tau)} - CF(t_f)\tilde{\mathbf{r}})^T(CWC^T)^{-1} \\ &\quad \times (\mathbf{y}_f - Ce^{\tilde{A}(t_f-\tau)} - CF(t_f)\tilde{\mathbf{r}})dt \\ &= \boldsymbol{\beta}^T W_p^{-1}\beta, \end{aligned} \tag{10}$$

where the vector $\boldsymbol{\beta} = Ce^{\tilde{A}(t_f-t_0)}\mathbf{x}_0 + CF(t_f)\tilde{\mathbf{r}} - \mathbf{y}_f$ is the control maneuver and $W_p = CWC^T$ is the $n \times n$ symmetric, real, non-negative definite output controllability Gramian. The smallest eigenvalue of the output controllability Gramian, μ_1, is nonzero if the system is output controllable. If this condition is satisfied, μ_1 usually dominates the expression for the minimum control energy.[16]

A. Effect of the attackers on output controllability Gramian

Consider the system under attack described by Eqs. (4) and (5). We write

$$e^{\tilde{A}t} = \begin{bmatrix} e^{At} & \vdots & F_1(t) \\ \cdots & \cdots & \cdots \\ 0 & \vdots & e^{St} \end{bmatrix}, \text{ where}$$

$$F_1(t) = \int_0^1 e^{(1-\tau)At}Hte^{\tau St}d\tau.^{17}$$

Moreover,

$$\tilde{B}\tilde{B}^T = \begin{bmatrix} B \\ \cdots \\ 0 \end{bmatrix}\begin{bmatrix} B^T & \vdots & 0^T \end{bmatrix} = \begin{bmatrix} BB^T & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & 0 \end{bmatrix}.$$

The controllability Gramian

$$\begin{aligned} W &= \int_{t_0}^{t_f} e^{\tilde{A}t}\tilde{B}\tilde{B}^T e^{\tilde{A}^T t}dt = \int_{t_0}^{t_f} \begin{bmatrix} e^{At} & \vdots & F_1(t) \\ \cdots & \cdots & \cdots \\ 0 & \vdots & e^{St} \end{bmatrix} \\ &\quad \times \begin{bmatrix} BB^T & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & 0 \end{bmatrix}\begin{bmatrix} e^{A^T t} & \vdots & 0 \\ \cdots & \cdots & \cdots \\ F_1^T(t) & \vdots & e^{S^T t} \end{bmatrix} dt \\ &= \int_{t_0}^{t_f} \begin{bmatrix} e^{At}BB^T e^{A^T t} & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & 0 \end{bmatrix} dt \\ &= \begin{bmatrix} W_p & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & 0 \end{bmatrix}. \end{aligned} \tag{11}$$

Note that $W_p \in \mathbb{R}^{n\times n}$ does not depend on the matrices S and E i.e., it is independent of the location of the attackers and the strength of the attackers. The output controllability Gramian

$$CWC^T = \begin{bmatrix} I & \vdots & 0 \end{bmatrix}\begin{bmatrix} W_p & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & 0 \end{bmatrix}\begin{bmatrix} I \\ \cdots \\ 0 \end{bmatrix} = W_p. \tag{12}$$

If the pair (A, B) is controllable, the matrix W_p is positive definite and thus invertible.[18,19]

B. Effect of the attackers on control maneuver

We have already defined the control maneuver as

$$\boldsymbol{\beta} = Ce^{\tilde{A}(t_f-t_0)}\mathbf{x}_0 + CF(t_f)\tilde{\mathbf{r}} - \mathbf{y}_f. \tag{13}$$

According to our assumptions, we set $\mathbf{y}_f = \mathbf{0}$ (target state coincides with the origin). The eigenvalue equation for the matrix $\tilde{A} = V\Lambda V^{-1}$, where the eigenvector matrix $V = [\mathbf{v}_1|\mathbf{v}_2|\cdots|\mathbf{v}_{n+q}]$ and the eigenvalue matrix $\Lambda = diag\{\lambda_1, \ldots, \lambda_q, \lambda_{q+1}, \ldots, \lambda_{q+n}\}$ where, $\lambda_1 \geq \cdots \geq \lambda_q \geq \lambda_{q+1} \geq \cdots \geq \lambda_{q+n}$. Note that because of the block diagonal structure of $\tilde{A}$ and the assumption that the matrix A is Hurwitz the first q eigenvalues of $\tilde{A}$, which correspond to the attackers dynamics, $\lambda_i = s_i$, $i = 1, \ldots, q$.

We write, $\mathbf{x_0} = \sum_i c_i\mathbf{v}_i = V\mathbf{c}$, where the vector $\mathbf{c} = [c_1, c_2, \ldots, c_{n+q}]$ and $F(t_f)\tilde{a}(r) = \sum_i g_i V_i = V\mathbf{c}$ where the vector $\mathbf{g} = [g_1, g_2, \ldots, g_{(n+q)}]$. Now from Eq. (13)

$$\boldsymbol{\beta} = Ce^{\tilde{A}(t_f-t_0)}\mathbf{x}_0 = C\left(\sum_{i=1}^{n+q} c_i e^{\lambda_i(t_f-t_0)}\mathbf{v}_i + \sum_{i=1}^{n+q} g_i J_i \mathbf{v}_i\right), \tag{14}$$

where $J_i = \frac{e^{\lambda_i(t_f-t_0)}-1}{\lambda_i}$. For large t_f, the above equation can be approximated as

$$\boldsymbol{\beta} \approx C\sum_{i=1}^{q}\left(c_i e^{s_i(t_f-t_0)} + g_i\frac{e^{s_i(t_f-t_0)}-1}{s_i}\right)\mathbf{v}_i. \tag{15}$$

We write

$$\boldsymbol{\beta} = \beta\boldsymbol{n}, \tag{16}$$

where $\boldsymbol{n}$ is a vector with norm equal to 1 having the same direction as $\boldsymbol{\beta}$. We see from Eq. (15) that for large t_f, the order of magnitude of $\boldsymbol{\beta}$ is determined by the number and strengths of attackers (i.e., s_i, $i = 1, .., q$).

We now express the symmetric matrix W_p in terms of its eigenvalues and eigenvectors. $W_p\boldsymbol{w}_i = \mu_i\boldsymbol{w}_i$, where $i = 1, \ldots, N$ so that $W_p^{-1} = \sum_{i=1}^{n}\mu_i^{-1}\boldsymbol{w}_i\boldsymbol{w}_i^T$.

Replacing W_p^{-1} into the Eq. (10)

$$\begin{aligned} E^* &= \boldsymbol{\beta}^T\sum_{i=1}^{n}\mu_i^{-1}\boldsymbol{w}_i\boldsymbol{w}_i^T\boldsymbol{\beta}, \quad \text{where } \mu_1 \leq \mu_2 \leq \ldots.. \leq \mu_N \\ &\approx \beta^2(\boldsymbol{n}^T\boldsymbol{w}_1)^2\mu_1^{-1}, \end{aligned} \tag{17}$$

where the approximation holds, when $\mu_1 \ll \mu_2$ and when $\boldsymbol{n}^T\boldsymbol{w}_1 \neq 0$. Thus, we can write

$$E^* \approx \beta^2 \mu_1^{-1} (\boldsymbol{n}^T \boldsymbol{w}_1)^2 = E_1 E_2 E_3, \tag{18}$$

where $E_1 = \beta^2$ corresponds to the strength of the attackers, $E_2 = \mu_1^{-1}$ does not depend on the attackers but depends on the network topology and the location of the defenders and $E_3 = (\boldsymbol{n}^T\boldsymbol{w}_1)^2$ depends on the distribution of attackers and defenders over the network. Note that the vector $\boldsymbol{w}_1$ is the eigenvector of W_p associated with its smallest eigenvalue. The term $\boldsymbol{n}^T\boldsymbol{w}_1$ measures the angle between two vectors both taken to have norm 1, thus $0 \leq (\boldsymbol{n}^T\boldsymbol{w}_1)^2 \leq 1$.

Now consider the simple ten node network in Fig. 1 and place the defenders on three nodes [nodes colored red in Fig. 1(e)]. In Fig. 2 we have considered the effect of different choices for the attackers. The smallest eigenvalue μ_1 of the output controllability Gramian W_p remained constant as the number and position of the attackers was varied.

Figure 2 also illustrates the case that the same network is subjected to attack changing only the position of the defenders, now at nodes 2, 7, and 9. Again we see that the minimum eigenvalue of the Gramian (μ_1) is independent of the number and position of the attackers. However, we see that μ_1 depends on the location of the defenders.

We come to the initial conclusion that we can determine for different networks, and different locations of attackers and defenders, the minimum control energy needed to control a network under attack in a preassigned time. Our main result is that the expression for the minimum control energy can be approximated as follows: $E^* \approx E_1E_2E_3$. While E_1 depends on the position of the attackers but not on the network topology, E_2 depends on the matrices A and B (on the Gramian), but not on the number, position, or strength of the attackers, and the quantity E_3 depends on the distribution of attackers and defenders over the network. This is investigated in more detail in Sec. II.

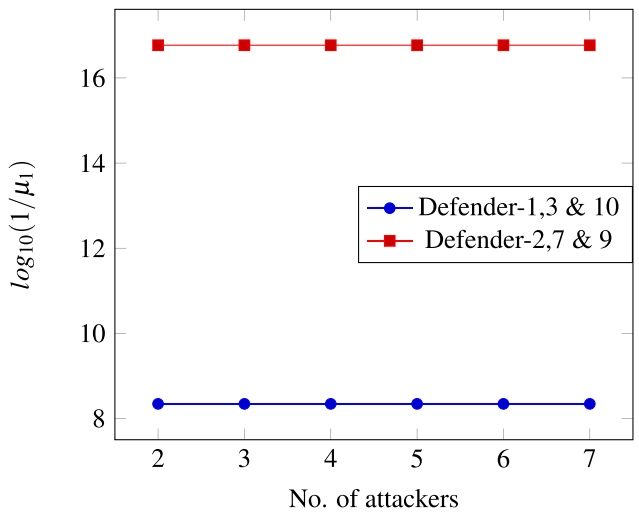

FIG. 2. log10($1/\mu_1$) for the network shown in Fig. 1 as we increased the number of nodes subject to attacks. Blue circle: Defenders are placed on nodes 1, 3, and 10; Red square: Defenders are placed on nodes 2, 7, and 9. Attackers are chosen in a random order, but ensuring that no node is both attacked and a defender.

II. ANALYSIS OF NETWORK TOPOLOGIES

In this section, we investigate how the control energy changes as we vary the position of attacked nodes and defenders over several networks. In all the simulations that follow, we set $A_{ij} = A_{ji} = 1$ if a connection exists between node i and j and $A_{ij} = A_{ji} = 0$ otherwise. We also set the matrices B and H to be composed of different versors as columns, which indicates each attacker and/or defender is localized at a given node (in particular, each attacker is attached to one and only one attacked node). We will first ensure for each of the networks that follow, the controllability is verified by a proper choice of defender nodes. Then, we study the effect of positioning the attackers on different network nodes. In order to compute the quantities $\boldsymbol{n}^T\boldsymbol{w}_1$ and β^2, we add a small random term to the entries on the main diagonal of the adjacency matrix A, $A_{ii} \leftarrow A_{ii} + \phi_i$, $i = 1, \ldots, N$, where ϕ_i is a random number uniformly chosen in the interval $[0, \varepsilon]$. This is done to ensure the pair (A, B) is controllable, see, e.g., Refs. 16 and 20.

A. Chain networks

Now we investigate how E_1 and E_3 vary in the six node bidirectional chain network shown in Fig. 3(a). We keep the position of the defender fixed at node 1 as indicated in Fig. 3. Then, we vary the position of the attacked nodes over the chain.

We see that the term $\boldsymbol{n}^T\boldsymbol{w}_1$, corresponding to E_3, generally increases when we increase the distance between the defender node and the attacked node. The term $\boldsymbol{n}^T\boldsymbol{w}_1$ is largest when the attacker is at node 6, i.e., the node which is farthest from the defender. Also, we see a small variation in the terms of β^2 as we change the position of the attacker as above. However, the effect of varying the position of the attacker on β^2 is less pronounced than on $\boldsymbol{n}^T\boldsymbol{w}_1$. Overall, these results are consistent with the previous studies on target control of networks where the control energy was found to increase with the distance between driver nodes and target nodes.[21]

Figure 4 shows the case that the defender is placed at the center node of the chain network. Here, we see that the quantity β^2 decreases as the distance from the defender node and the attacked node increases. However, the quantity $\mathbf{n}^T\mathbf{w}_1$ displays a much more complex and somehow surprising behavior, also distinctly different from that observed in Fig. 3. Namely, we see that the quantity $\mathbf{n}^T\mathbf{w}_1$ alternatively increases and decreases as the position of the attacker is moved over the chain. This type of behavior is different from what seen in the case of target control of networks.[21]

B. Star network

Consider the case of the star network in Fig. 5(a) with defender at node 1 and the position of the attacked node varied from node 2 to 9. We see that the value of $\boldsymbol{n}^T\boldsymbol{w}_1$ when the position of the attacked node is in the first layer of the star network (i.e., on nodes 2, 3, 4, and 5) is nearly constant over

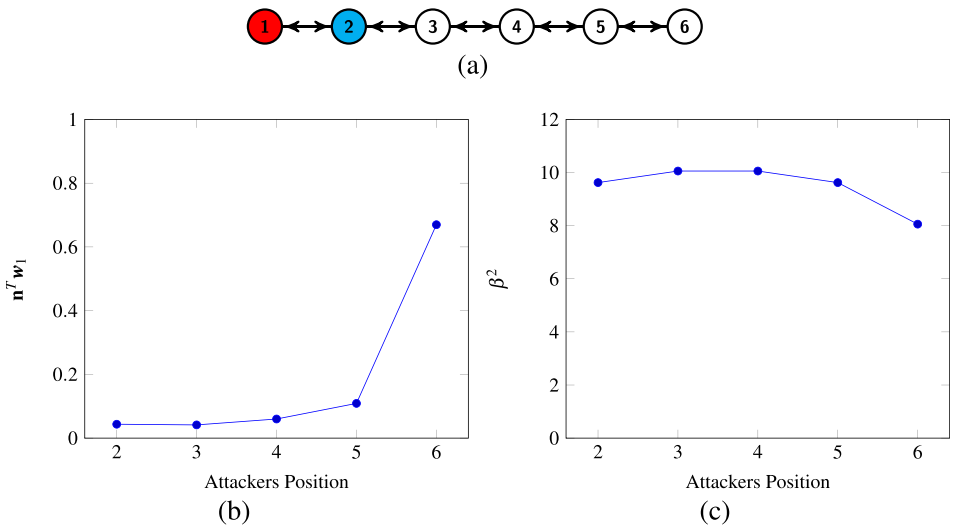

FIG. 3. (a) Bidirectional chain network. The defender node is in red and the attacked node is in cyan. (b) $\mathbf{n}^T\mathbf{w}_1$ versus the position of the attacker. (c) Plot of $\boldsymbol{\beta}^2$ as the position of the attacker is varied. We perform calculations setting t_f=1, $s_i = 2.5$, $r_i = 0$, and $\varepsilon = 10^{-2}$.

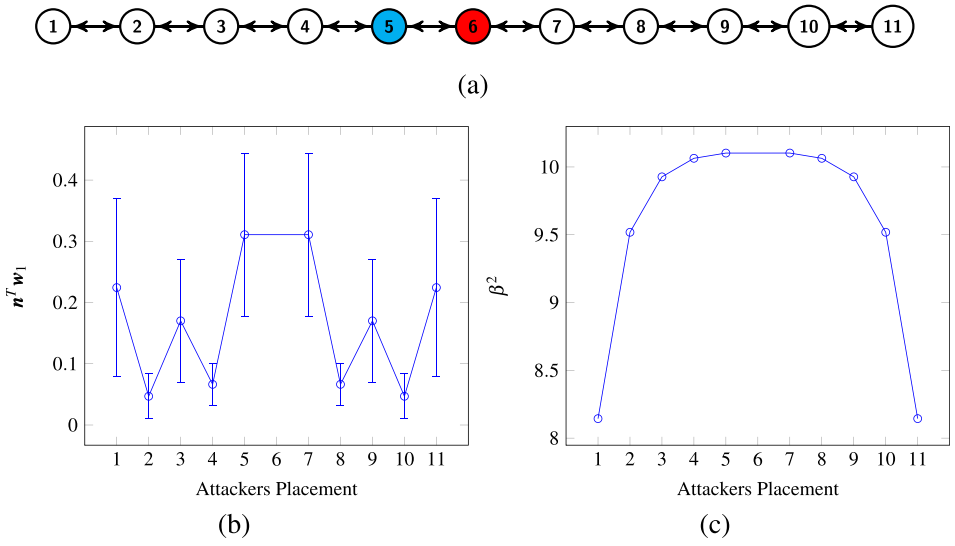

FIG. 4. (a) A chain network with defender at the center node. (b) Plot of $\boldsymbol{n}^T\boldsymbol{w}_1$ vs. the position of the attacked node. (c) Plot of β^2 vs. the position of the attacked node. We perform calculations setting t_f=1, $s_i = 2.5$, $r_i = 0$ and $\varepsilon = 10^{-2}$. The bars represent the standard deviation taken over 100 different realizations.

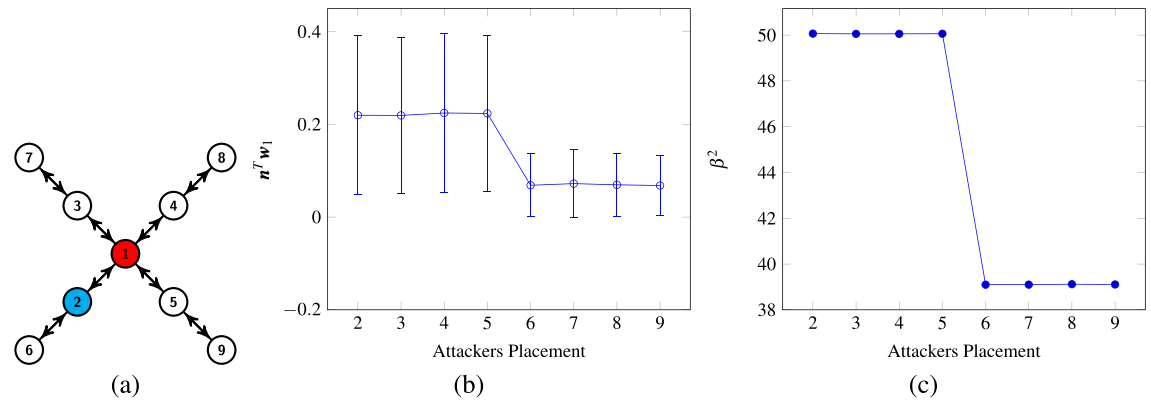

FIG. 5. (a) A star network. (b) Plot of $\boldsymbol{n}^T\boldsymbol{w}_1$ vs position of the attacker. (c) β^2 vs. the position of the attacked nodes. We perform calculations setting t_f=1, $s_i = 2.5$, $r_i = 0$ and $\varepsilon = 10^{-2}$. The bars represent the standard deviation taken over 100 different realizations.

that layer. When the attacker is on the second layer (i.e., on nodes 6, 7, 8, and 9), the value of $\boldsymbol{n}^T\boldsymbol{w}_1$ is also nearly constant. In Fig. 5(b), we see a similar pattern for β^2 as we saw for $\boldsymbol{n}^T\boldsymbol{w}_1$ in Fig. 5(b). The value of β^2 for the first layer is equal and so is for the second layer. However, when comparing the two layers, we see from both panels (b) and (c) in Fig. 5 that surprisingly the energy to control the star network decreases with the distance between the attacked node and the defender over the network.

C. Ring network

We now consider a small ring network of 8 nodes [shown in Fig. 6(a)]. The defender is at node 1 and the

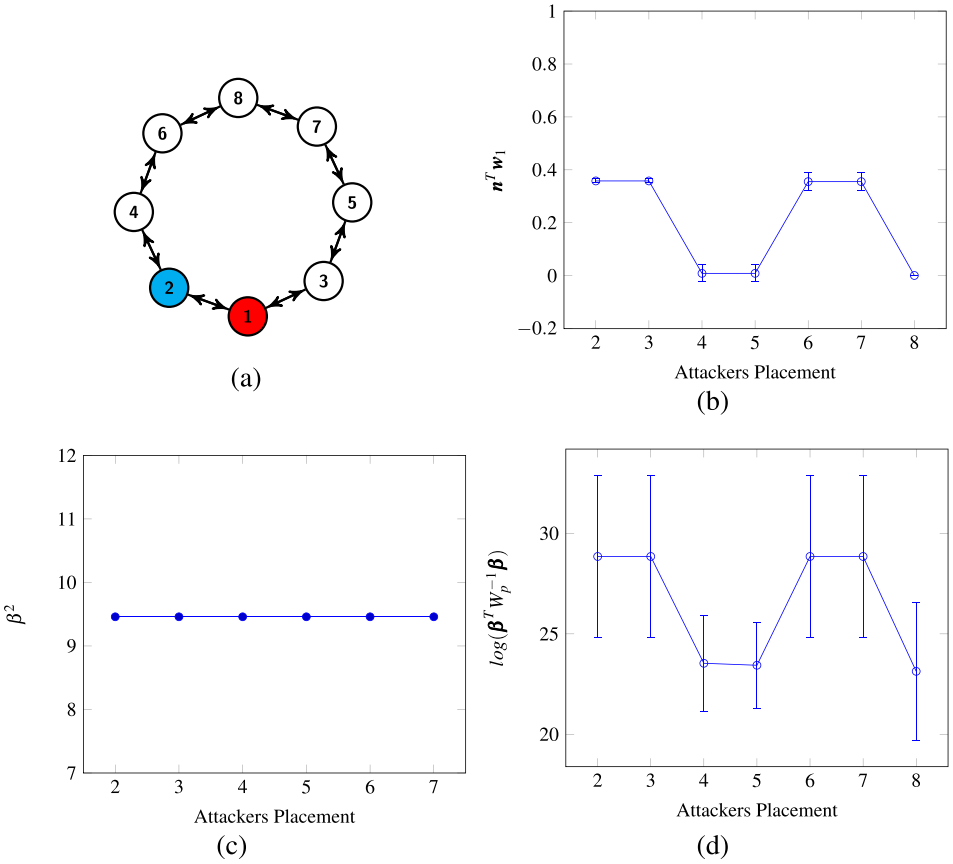

FIG. 6. (a) An eight node ring network with defender at node 1. (b) Plot of $\boldsymbol{n}^T\boldsymbol{w}_1$ vs the position of the attacked node. (c) Plot of β^2 vs the position of the attacked node. (d) Total energy E as the position of the attacked node is varied. The bars represent standard deviations over 100 different realizations. We perform calculations setting the final time t_f=1, $s_i = 2.5$, $r_i = 0$ and $\varepsilon = 10^{-2}$.

attacker can be at any other node. From Fig. 6(b), we see that the value of $\boldsymbol{n}^T\boldsymbol{w}_1$ varies with the distance between the attacked and defender nodes over the ring (nodes 2 and 3 at distance 1, nodes 4 and 5 at distance 2, nodes 6 and 7 at distance 3, and node 8 at distance 4). However, the variation is, once again, non-monotonic with respect to the distance. In particular, we do not see that the energy to control the attack monotonically increases with the distance between attacker and defender. Figure 6(c) shows that in this case β^2 is independent of the position of the attacker over the ring network. Figure 6(d) shows the total energy E^* from Eq. (10), which is consistent with Fig. 6(b).

D. Scale free networks

Here, we consider a 300 node Barabasi Albert scale free network[11] with average degree 2. We select 10% of the nodes to be defenders and position them so to ensure that the pair (A, B) is controllable using the algorithm described in Ref. 22. We then vary the choice of a single attacked node over the network, one by one, excluding the defender nodes. For each selection, we compute Δ, the minimum shortest distance over the network between the attacked node and the defender nodes

$$\Delta = \min_d \text{ shortest distance}(a, d), \tag{19}$$

where a indicates the attacked node and d the defender nodes. Each point in Fig. 7(a) indicates the value of $\boldsymbol{n}^T\boldsymbol{w}_1$ for a given choice of the attacked node versus the degree of the attacker. As can be seen, the quantity $\boldsymbol{n}^T\boldsymbol{w}_1$ varies over several orders of magnitude for different choices of the attacked nodes. In particular, certain nodes are *weak attackers* as the required control energy is particularly low when these nodes

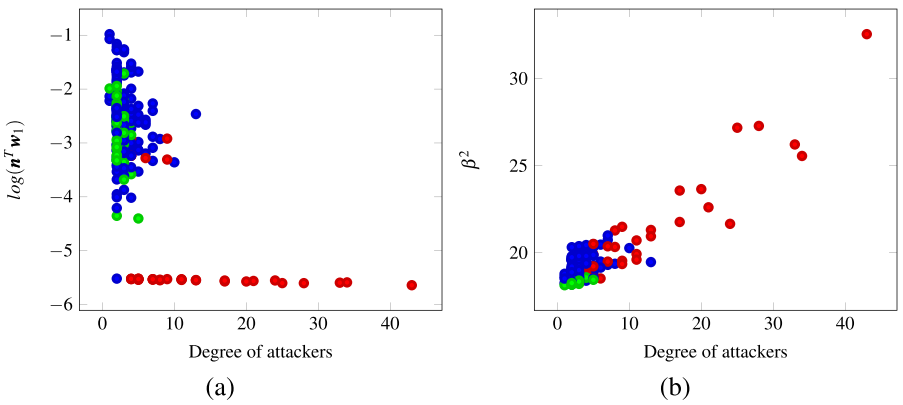

FIG. 7. (a) Plot of $\boldsymbol{n}^T\boldsymbol{w}_1$ vs the degree of the attacked node for a 300 node scale free network with average degree 2. (b) Plot of β^2 vs the degree of the attacked node. Green circle indicates attacked nodes with $\Delta = 3$. Blue circle indicates attacked nodes with $\Delta = 2$. Red circle indicates attacked nodes with $\Delta = 1$. We perform calculations setting the final time t_f=1 with $s_i = 2.5$, $r_i = 0$ and $\varepsilon = 10^{-2}$.

are subject to an attack. Figure 7(b) indicates the value of β^2 for the given choice of attacker node versus the degree of the attacker. The quantity β^2 increases as the degree of the attacked nodes increases and the quantity Δ decreases. While attacked nodes with $\Delta = 1$ tend to have a slightly higher value of β^2, the value of $\boldsymbol{n}^T\boldsymbol{w}_1$ is typically at least one order of magnitude lower, indicating that the minimum control energy E is much lower when these nodes are attacked. Overall, Fig. 7 shows that the degree of a node is not a good predictor for a weak attacker, as these are nodes of all possible degrees. However, the parameter Δ appears to be a good indicator for a weak attacker, as these have typically $\Delta = 1$, i.e., they are neighbors of at least one defender.

III. AN EXAMPLE OF APPLICATION OF THE ANALYSIS TO INFRASTRUCTURE NETWORKS

The application discussed in this section was presented in Ref. 12. There, they use the IEEE 39 bus system where a dynamic load altering attack is used to destabilize the system.

The power system dynamics can be described as follows:[12]

$$\begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & -M & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{\delta}} \\ \dot{\boldsymbol{\theta}} \\ \dot{\boldsymbol{\omega}} \\ \dot{\boldsymbol{\varphi}} \end{bmatrix} = \begin{bmatrix} 0 & 0 & I & 0 \\ 0 & 0 & 0 & -I \\ K^I + H^{GG} & H^{GL} & K^P + D^G & 0 \\ H^{LG} & H^{LL} & 0 & D^L \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta} \\ \boldsymbol{\theta} \\ \boldsymbol{\omega} \\ \boldsymbol{\varphi} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ I \end{bmatrix} \boldsymbol{P^L}. \tag{20}$$

Equation (20) can be rewritten as follows, after setting $\dot{\boldsymbol{\varphi}}$ to zero and replacing in the equations for the time evolution of $\boldsymbol{\delta}$, $\boldsymbol{\theta}$, and $\boldsymbol{\omega}$

$$\begin{bmatrix} \dot{\boldsymbol{\delta}} \\ \dot{\boldsymbol{\theta}} \\ \dot{\boldsymbol{\omega}} \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & D^{L^{-1}} & 0 \\ 0 & 0 & -M^{-1} \end{bmatrix} \begin{bmatrix} 0 & 0 & I \\ H^{LG} & H^{LL} & 0 \\ K^I + H^{GG} & H^{GL} & K^P + D^G \end{bmatrix} \times \begin{bmatrix} \boldsymbol{\delta} \\ \boldsymbol{\theta} \\ \boldsymbol{\omega} \end{bmatrix} + \begin{bmatrix} 0 \\ D^{L^{-1}} \\ 0 \end{bmatrix} \boldsymbol{P^L}. \tag{21}$$

Let us assume that we can add an ancillary generator in our power grid system to compensate for over- and under-frequency disruptions. Then, the mechanical power input $\boldsymbol{P_i^M}$ at the i generator with ancillary generation power $\boldsymbol{P_i^{M'}}$ is given by

$$\boldsymbol{P}_i^M = -\left(K_i^P \omega_i + K_i^P \int_0^t \omega_i + \boldsymbol{P}_i^{M'} \right). \tag{22}$$

Now the total power grid system with load attack on $\mathbf{P}^L$ and ancillary generation $\mathbf{P}^M$ can be written in the form of Eq. (22) as follows:

$$\begin{bmatrix} \dot{\boldsymbol{\delta}} \\ \dot{\boldsymbol{\theta}} \\ \dot{\boldsymbol{\omega}} \end{bmatrix} = A \begin{bmatrix} \boldsymbol{\delta} \\ \boldsymbol{\theta} \\ \boldsymbol{\omega} \end{bmatrix} + H \begin{bmatrix} 0 \\ \boldsymbol{P^L} \\ 0 \end{bmatrix} + B \begin{bmatrix} 0 \\ 0 \\ \boldsymbol{P^M} \end{bmatrix}, \tag{23}$$

where

$$\mathrm{A} = \begin{bmatrix} I & 0 & 0 \\ 0 & (D^L)^{-1} & 0 \\ 0 & 0 & -M^{-1} \end{bmatrix} \times \begin{bmatrix} 0 & 0 & I \\ H^{LG} & H^{LL} & 0 \\ K^I + H^{GG} & H^{GL} & K^P + D^G \end{bmatrix}$$

And

$$\mathrm{H} = \begin{bmatrix} 0 \\ (D^L)^{-1} \\ 0 \end{bmatrix},$$

$$\mathrm{B} = \begin{bmatrix} 0 \\ 0 \\ -M^{-1} \end{bmatrix}.$$

The matrix A is the system matrix, the matrix E determines the effect and position of the attackers, and the matrix B determines the effect and position of the defenders. Note that Eq. (23) is the same as Eq. (2).

IV. CONCLUSIONS

In this paper, we have studied an optimal control problem on networks, where a subset of the network nodes are attacked but the attack is mitigated by using available actuating capabilities at another subset of the network nodes. Compared with the previous work on optimal control of network,[11,16,20,21] we consider a situation in which the control action is implemented, along side external dynamics also affecting the network.

We envision this work to be relevant to critical infrastructure networks (such as power grids), which are susceptible to attacks. While our results assume knowledge of the attacker's strategy, which is often unavailable, our analysis can be used to design infrastructure networks that are resistant to attacks. This can be done by considering all possible attacks that can affect the network and for each case, compute the optimal control solution. We have studied how the minimum control energy varies as the position of the *attackers* and *defenders* is varied over different networks such as chain, star, ring, and scale free networks. Our main result is that the expression for the minimum control energy can be approximated by the product of three different quantities $E_1E_2E_3$. While E_1 depends on the position of the attackers but not on the network topology, E_2 depends on the matrices A and B (on the Gramian), and E_3 depends on the position of both the attacked nodes and defender nodes over the network.

In chain, star, and ring networks, we see that for a single attacker and a single defender, often the minimum control energy is not an increasing function of the distance between the attacked node and the defender node. However, for a scale free network with multiple defenders and a single attacker, we see that a good predictor for the strength of the

attack is provided by the quantity Δ (the minimum distance between the defender nodes and the attacked node over the network). Using the approach in Ref. 16, our work can be generalized to the case that the defense strategy only protects a subset of the network nodes.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation though NSF Grant No. CMMI-1400193, NSF Grant No. CRISP-1541148, ONR Grant No. N00014–16-1–2637, and DTRA Grant No. HDTRA1–12-1–0020.

[1]G. A. Pagani and M. Aiello, "The power grid as a complex network: A survey," Physica **392**(11), 2688–2700 (2013).
[2]H. Yang and M. G. H. Bell, "Models and algorithms for road network design: A review and some new developments," Transp. Rev. **18**(3), 257–278 (1998).
[3]M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as optimization decomposition: A mathematical theory of network architectures," Proc. IEEE **95**(1), 255–312 (2007).
[4]R. Prakash and U. V. Shenoy, "Targeting and design of water networks for fixed flow rate and fixed contaminant load operations," Chem. Eng. Sci. **60**(1), 255–268 (2005).
[5]R. Albert, H. Jeong, and A.-L. Barabsi, "Error and attack tolerance of complex networks," Nature **406**(6794), 378–382 (2000).
[6]P. Crucitti, V. Latora, M. Marchiori, and A. Rapisarda, "Error and attack tolerance of complex networks," Physica A **340**(1), 388–394 (2004).
[7]R. Albert, I. Albert, and G. L. Nakarado, "Structural vulnerability of the North American power grid," Phys. Rev. E **69**(2), 025103 (2004).
[8]M. G. H. Bell, U. Kanturska, J.-D. Schmcker, and A. Fonzone, "Attacker–defender models and road network vulnerability," Philos. Trans. R. Soc., London A **366**(1872), 1893–1906 (2008).
[9]A. Kessler, A. Ostfeld, and G. Sinai, "Detecting accidental contaminations in municipal water networks," J. Water Resources Plann. Manage. **124**(4), 192–198 (1998).
[10]R. Pastor-Satorras and A. Vespignani, "Epidemic dynamics and endemic states in complex networks," Phys. Rev. E **63**(6), 066117 (2001).
[11]Y.-Y. Liu, J.-J. Slotine, and A. Lszl Barabsi, "Controllability of complex networks," Nature **473**(7346), 167–173 (2011).
[12]S. Amini, F. Pasqualetti, and H. Mohsenian-Rad, "Dynamic load altering attacks against power system stability: Attack models and protection schemes," IEEE Trans. Smart Grid (published online, 2016).
[13]D. E. Kirk, *Optimal Control Theory: An Introduction* (Courier Corporation, 2012).
[14]D. G. Luenberger, *Introduction to Dynamic Systems: Theory, Models, and Applications* (Wiley, New York, 1979), Vol. 1.
[15]R. F. Stengel, *Optimal Control and Estimation* (Courier Corporation, 1986).
[16]I. Klickstein, A. Shirin, and F. Sorrentino, "Energy scaling of targeted optimal control of complex networks," Nat. Commun. **8**, 15145 (2017).
[17]L. Dieci and A. Papini, "Pad approximation for the exponential of a block triangular matrix," Linear Algebra Appl. **308**(1-3), 183–202 (2000).
[18]K. Murota and S. Poljak, "Note on a graph-theoretic criterion for structural output controllability," IEEE Trans. Automat. Control **35**(8), 939–942 (1990).
[19]W. J. Rugh and W. J. Rugh, *Linear System Theory* (Prentice Hall, Upper Saddle River, NJ, 1996), Vol. 2.
[20]G. Yan, G. Tsekenis, B. Barzel, J.-J. Slotine, Y.-Y. Liu, and A.-L. Barabsi, "Spectrum of controlling and observing complex networks," Nat. Phys. **11**(9), 779 (2015).
[21]I. Klickstein, I. Kafle, S. Bartaula, and F. Sorrentino, "Energy scaling with control distance in complex networks," preprint arXiv:1801.09642 (2018).
[22]F. L. Iudice, F. Garofalo, and F. Sorrentino, "Structural permeability of complex networks to control signals," Nat. Commun. **6**, 8349 (2015).