Hindawi
Complexity
Volume 2021, Article ID 5584123, 19 pages
https://doi.org/10.1155/2021/5584123

Research Article

Modeling Coupled Nonlinear Multilayered Dynamics: Cyber Attack and Disruption of an Electric Grid

Marc Mangel [1,2,3] **and Jimmie McEver** [4]

[1] *Department of Applied Mathematics, University of CA, Santa Cruz, CA 95060, USA*
[2] *Department of Biology, University of Bergen, Bergen 9020, Norway*
[3] *Puget Sound Institute, University of WA, Tacoma 98402, USA*
[4] *JHU Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723, USA*

Correspondence should be addressed to Marc Mangel; marcmangel@protonmail.com

Received 25 February 2021; Accepted 20 August 2021; Published 7 September 2021

Academic Editor: M. De Aguiar

We study the consequences of cyberattack, defense, and recovery in systems for which a physical system is enabled by a cyber system by extending previous applications of models from the population biology of disease to the cyber system and coupling the state of the cyber system to the physical system, using the synchronous model for the electric grid. In analogy to disease models in which individuals are susceptible, infected, or recovered, in the cyber system, components can be uncompromised and vulnerable to attack, uncompromised and temporarily invulnerable to attack, compromised, or reset and thus not able to contribute to the performance of the physical system. We model cyber defensive countermeasures in analogy to the adaptive immune system. We link the physical and cyber systems through a metric of performance of the physical system that depends upon the state of the cyber system using (i) a generic nonlinear relationship between the state of the cyber system and the performance of the physical system and (ii) the synchronous motor model of an electric grid consisting of a utility with many customers whose smart meters can become compromised, in which a steady state in the difference in rotor angles is the metric of performance. We use the coupled models, both of which have emergent properties, to investigate two situations. First, when an attacker that relies on stealth compromise is hidden until it is either detected during routine maintenance or an attack is initiated. The probability that compromise remains undetected declines with time and the level of compromise increases with time. Because of these dynamics, an optimal time of attack emerges, and we explore how it varies with parameters of the cyber system. Second, we illustrate one of the Electric Power Research Institute scenarios for the reverse engineering of Advanced Metering Infrastructure (AMI) by coupling the synchronous motor equations for the generator and utility to the model of compromise. We derive a canonical condition for grid failure that relates the level of compromise at the time of detection of compromise and the dissipation parameter in the synchronous motor model. We conclude by discussing the innovative aspects of our methods, which include (i) a fraction of decoy components in the cyber system, which are not connected to the rest of the cyber system or the physical system and thus do not spread compromise but increase the probability of detection of compromise, (ii) allowing components of the cyber system to return to the un-compromised state either temporarily invulnerable or immediately vulnerable, (iii) adaptive Defensive Counter Measures that respond in a nonlinear fashion to attack and compromise (in analogy to killer T cells of the immune system), (iv) a generic metric of performance of the physical system that depends upon the state of the cyber system, and (v) coupling a model of the electric grid to the model of compromise of the cyber system that leads to a condition for failure of the grid in terms of parameters of both compromise and the synchronous motor model, directions for future investigations, and connections to recent studies on broadly the same topics. We include a pseudocode as an Appendix and indicate how to obtain *R* script for the models from the first author.

1. Introduction

A recent survey [1] of cyberattack in complex systems with both physical and cyber components in which the cyber system enables functionality of the physical system emphasizes the importance of improving our ability to understand this class of systems, and that it is critical to go beyond the attack-defense dynamics in cyberspace, but to include how changes in the cyber domain generate changes in physical outcomes (and possibly vice versa). That is, one wishes to assure that the physical system remains in the desired operating regime or out of undesired ones and needs to discover methods to achieve this goal. Modern power and communication systems are canonical examples of such systems [2, 3] and in [1], the particular case of a smart grid subject disruption and stabilization through a Distributed Denial of Service (DDOS) or deception attacks that alters sensor information [4, 5] is treated.

In this paper, we extend the ideas in [1] in a new direction by explicitly coupling the models of attack dynamics and the physical functional response of the system under attack. This allows us to analytically explore the system dynamics of the coupled nonlinear systems, in particular how parameterized attacks of certain types and behaviors result in effects observed in performance and functionality of the resulting system (in this case, a smart power grid). In doing so, we develop a modeling approach that creates the ability for attackers to optimize the timing and pace of their attacks, and for defenders and designers to select defense strategies, deploy defensive assets, and make architectural and design choices tuned to the coupled cyber-physical system.

To do this, we build on existing nonlinear dynamic models of compromise of cyber systems that import ideas from the population biology of disease [6–11], characterize performance using either a nonlinear metric for a generic physical system (cf. [6–10]) or the nonlinear dynamics of an electric grid, and introduce counter measures having nonlinear dynamics based on the immune systems [12, 13].

Our goal is to develop a systems dynamics model to explore the roles of attack and maintenance rates, detection capability, and other design parameters on the dynamics, particularly how compromise of the cyber system propagates, and performance of the linked physical system responds. Since our goal is to understand the important variables and how they affect performance, we build a heuristic model [14] that is not specific to any particular situation but has commonalities with many complex systems. Using the model will help identify what to measure to be able to assess vulnerability to cyberattack, the consequences of attack on performance of the physical system, and to identify design tradeoffs and routes to defense.

We link the physical and cyber systems through a metric of performance of the physical system that depends upon the state of the cyber system, using (i) a generic nonlinear relationship between the state of the cyber system and the performance of the physical system and (ii) the synchronous motor model of an electric grid in which the load is a utility having consumers that used Advanced Metering Infrastructure (AMI; smart meters) that can be compromised.

In the next section, we develop the model for compromise of the cyber system, after which we introduce two metrics of performance of the physical system. We then illustrate the dynamics of the model for the cyber system, after which we treat two examples. First we show how the timing of attack when the attacker relies on stealth emerges from the nonlinear nature of both compromise and performance. We explore how performance of the physical system depends on parameters of compromise and performance and show how the optimal time of attack depends upon the parameters of the physical system and rates of external compromise and internal co-compromise.

Second, we explore compromise of Advanced Metering Infrastructure (AMI) and failure of the electric grid. AMI is an example of the Internet of Things, which increases the risk of cyber compromise because it increases the number of points of access for attackers [14–19]. In this example, we couple the model of compromise to the synchronous motor model of an electric grid to illustrate Electric Power Research Institute scenario AMI.27 [20] for reverse engineering of AMI, and a specific case of how compromising smart meters can lead to load-side failure of an electric grid. Because power grids are now considered critical infrastructure, data about them are often treated in a confidential manner [21, 22]. By using a generally accessible model for the grid, we are able to clearly see how compromise propagates and interacts with performance of the grid.

2. Materials and Methods

We envision a complex system consisting of a physical system that is enabled by a cyber system that takes commands, exchanges information, and more generally interacts with the external world. Attacks on the cyber system lead to compromise of its components, which has an effect on the physical system. Our goal is to provide a framework for modeling compromise in the cyber system, linking the cyber and physical systems (both generically and specifically [the electric grid]), and use the model to explore the dynamics of compromise, attack, and recovery of the cyber system and the related performance of the physical system.

We first describe the conceptual framework for the model of the cyber system (Section 2.1), after which we derive the equations for the dynamics of compromise and recovery (Section 2.2). We then show how the dynamics simplify when the attacker relies on stealth (Section 2.3), in which compromise is built until either it is discovered during regular maintenance or an attack is executed. We consider two metrics of performance (Section 2.4): (i) a generic metric that depends upon the number of uncompromised cyber components, the number of compromised cyber components, and the level of Defensive Counter Measures (DCM) and (ii) a synchronous motor model for the electric grid for which the metric of performance is the stability of the grid.

2.1. Conceptual Description of the Cyber Subsystem. An attack on a cyber system first requires external compromise: gaining access to components that interface with the external world. Once "inside" the cyber system, the second stage of internal co-compromise can commence, in which compromised components infect noncompromised components. Compromised components may immediately reduce the performance of the physical system or the adversary may hide compromise until ready to execute the attack.

In many cases, cyber components that interface with external world can be protected by external hardness [14] in which case antimalware prevents attackers from entering the cyber system. Cyber components that do not interface with the external world can similarly be protected from co-compromise by internal hardness [14]. External and internal hardness rely on a variety of mechanisms [21, 23, 24] that we do not model explicitly. However, it is now clear that external and internal hardness are necessary but not sufficient for both a reliable and resilient cyber system [21] and that resilience of the cyber system (and thus performance of the physical system) requires some form of Defensive Counter Measure (DCM) that returns the system to a state closer to the one before the attack. Protection from external compromise and internal co-compromise may not be effective (e.g., the installing antimalware does not defend) or may lose its effectiveness over time (e.g., the attacker discovers a way to circumvent the antimalware currently installed).

Thus, at any time, the cyber system consists of five classes of components (Figure 1): First, uncompromised and vulnerable components can be compromised either externally or internally. We allow a fraction η of these components to be decoys, with no functionality, but instrumented to detect compromise with high probability. Second, uncompromised cyber components that are currently invulnerable to either external or internal compromise are temporarily protected against malware, but as time progresses their antimalware software ages and is no longer effective. Third, compromised cyber components are infected by malware. Fourth, once compromised components are discovered, they are temporarily removed from the cyber system to be restored or reset later [21]. These components do not contribute to performance of the physical system. After some amount of time, components that are being reset return to the cyber system temporarily invulnerable (effective antimalware installed) or still vulnerable (either no antimalware installed or the installed antimalware is ineffective). Fifth, once compromise is detected, DCM are activated to discover and send compromised components into the resetting phase. Since DCM use cyber resources, we assume that their presence reduces the functionality of the physical system.

2.2. The Dynamics of Compromise, Defense, and Detection

2.2.1. Dynamics of the Cyber System Components. We assume that the cyber system has N components, with dynamics characterized by mass action in a mean field approximation [25–29].

> *Uncompromised and vulnerable cyber components*, denoted by $x_1(t)$ transition to become compromised components (i) because of external compromise at a rate proportional to their numbers and (ii) because of internal co-compromise by previously compromised cyber components at a rate proportional to the numbers of both kinds of components, accounting for the decoy cyber components. In addition, temporarily invulnerable cyber components lose their protection and compromised cyber components that are reset without any protection increase the number of uncompromised and vulnerable components. We assume that the rate of these transitions is proportional to their number. Hence the dynamics are

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -(c + c_s(1-\eta)y)x_1 + gx_2 + f_1(N - x_1 - x_2 - y). \tag{1}$$

On the right hand side of equation (1), η is dimensionless; c is the rate of external compromise so that the fraction of uncompromised cyber components surviving external compromise from t to $t + \Delta t$ is $e^{-c\Delta t}$; c_s is the rate of co-compromise so that when there are y compromised cyber components at time t, the fraction of uncompromised cyber components surviving co-compromise to $t + \Delta t$ is $e^{-c_s(1-\eta)y\Delta t}$; g is the rate at which temporarily invulnerable cyber components become vulnerable; and f_1 is the rate at which compromised components being reset return vulnerable to compromise.

Uncompromised and temporarily invulnerable cyber components, denoted by $x_2(t)$, transition to the uncompromised and vulnerable state as they lose protection. Compromised cyber components that reset with temporary protection from external or internal compromise transition to the uncompromised and temporarily invulnerable states. We assume that the rates of these transitions are proportional to the number of components so that the dynamics are

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = -gx_2 + f_2(N_1 - x_1 - x_2 - y). \tag{2}$$

In this equation, f_2 is the rate at which compromised cyber components are reset and returned to the cyber system temporarily invulnerable. If $f_1 \neq 0$ and $f_2 = 0$, then all cyber components that are reset return vulnerable to attack. On the other hand, if $f_1 = 0$ and $f_2 \neq 0$, then all cyber components return temporarily invulnerable to attack. In the most general case, both f_1 and f_2 are nonzero, so that returned cyber components have a mixture of vulnerabilities.

Compromised cyber components, denoted by $y(t)$, increase in number due to the processes in the first term on the right hand side of equation (1). Compromised cyber components transition to the resetting state due to regular maintenance during which compromised components may be discovered and when removed by DCM, denoted by $z(t)$. We assume that the rates of these transitions are proportional to their number for regular maintenance and to the their number and the level of DCM when DCM are activated. We let

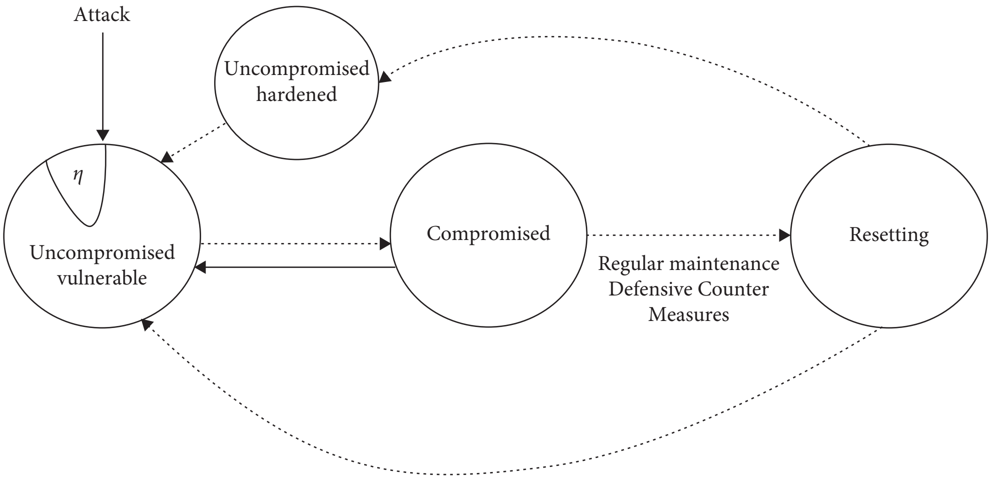


Figure 1: The cyber system contains components that are uncompromised and vulnerable, uncompromised and currently invulnerable (hardened), compromised, and resetting (and thus temporarily unavailable system; metaphorically "in the shop"). A fraction η of the cyber components (shown only for uncompromised and vulnerable components) are decoys that do not contribute to functionality of the physical system and are unable to co-compromise [20]. Dotted lines represent transitions from one stage to another and solid lines represent either external compromise or co-compromise.

$\mathcal{I}_{\mathrm{DCM}}(t)$ denote an indicator function that is 1 if DCM are activated at time t and is 0 otherwise. The dynamics of compromised cyber components are then

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \left(c + c_s(1-\eta)y\right)x_1 - \left(\mu_m + \mathcal{I}_{DCM}(t)\mu_{DCM}z\right)y, \tag{3}$$

where μ_m is the rate of transition of compromised cyber components to resetting during regular maintenance and μ_{DCM} is the rate of transition of compromised cyber components to resetting by DCM.

Defensive Counter Measures are intensive malware eradication efforts. We do not specify the mechanism of the DCM, since they are situation dependent [23]. DCM are active only after compromise is detected, which may occur during regular maintenance, through recognition of external compromise, or anomalous performance of the physical system [30–32]. After compromise is detected, we model the dynamics of DCM, denoted by $z(t)$, in analogy to T cells of the immune system [12, 13].

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\alpha + \gamma y}{1 + \beta z} - Mz. \tag{4}$$

If $\alpha = 0$ in equation (4), DCM are active only when compromised cyber components are present and in absence of compromised cyber components, the only steady state is $\overline{z} = 0$. This is the situation we assume for computations. When $\alpha > 0$, there is a positive steady-state level of DCM even when there is no compromise in the system, providing ready defense of the cyber system at the possible expense of performance of the physical system (see below). The parameter β controls the rate of increase of DCM; in particular, when $\beta > 0$, the rate of increase of DCM declines as the level of DCM increases.

Because cyber components are neither created nor destroyed, the numbers of uncompromised and vulnerable, uncompromised and temporarily invulnerable, compromised, and resetting components sums to N. This is captured by the second terms on the right hand sides of equations (1) and (2).

2.2.2. The Detection of Compromise. A variety of methods have been proposed for detecting compromise [30], but in general the detection of compromise will be imperfect, because regardless of the specific process (e.g., signal matching or anomalous behavior) there will be both false positives and false negatives [14]. We let $U(t)$ denote the probability that compromise remains undetected at time t.

We assume that the probability of discovering compromise in the next Δt units of time when the rate of external compromise is c and y cyber components are compromised is $\left(\left[\varepsilon_1(1-\eta) + \varepsilon_2\eta\right]y + \varepsilon_c c\right)\Delta t + o(\Delta t)$ where ε_1 and ε_2 are the rates of detection of compromised nondecoy and compromised decoy cyber components, respectively, and ε_c is the rate of detection of compromising activity when the rate of external compromise is c. In general, we anticipate that $\varepsilon_2 > \varepsilon_1$. For simplicity, for computations here we set $\varepsilon_c = 0$.

Assuming that compromise is undetected at the start, $U(0) = 1$; subsequently

$$\frac{\mathrm{d}U}{\mathrm{d}t} = -\left(\left[\varepsilon_1(1-\eta) + \varepsilon_2\eta\right]y + \varepsilon_c c\right)U. \tag{5}$$

Equation (5) generates a trajectory $U(t)$ for the probability that compromise is not detected by time t. We use this trajectory to characterize the stochastic process of detection. That is, from $U(t)$, we generate realizations of the time at which compromise is detected, and explore the consequences of different times of detection on the performance of the physical system.

2.2.3. Summary: a Mixed Deterministic-Stochastic Nonlinear Model. The full model for dynamics of the cyber system consists of equations (1)–(5). Because of equations (1)–(4), the model is inherently nonlinear. Because of equation (5), it is stochastic, even though the rest of the dynamics are deterministic. We implement this mixed deterministic-stochastic model by solving equations (1)–(5) with $\mathcal{I}_{\mathrm{DCM}} \equiv 0$ to generate the trajectory of $U(t)$. Given this trajectory, we generate K times of detection, denoted by $t_d(k)$, $k = 1, 2, \ldots, K$ by drawing uniformly distributed random variables and comparing them with $U(t)$. Then, for every $t_d(k)$, we solve equations (1)–(5) to determine the entire trajectory of compromise, detection, and recovery from compromise.

For the base case, we used these parameters: $N = 100, c = 0.005, c_s = 0.001 g = 0.1, f_1 = 0, f_2 = 3, \quad \mu_m = 0.05, \mu_{\mathrm{DCM}} = 0.2, \alpha = 0, \gamma = 0.1, \beta = 1.0, \quad M = 0.5, \varepsilon_1 = 0.005, \varepsilon_2 = \quad 0.05, \varepsilon_c = 0, \eta = 0$. All computations were done in *R* Studio Version 1.0.143 with underlying *R* version 3.6.1.

2.3. Simplification When the Attacker Relies on Stealth. When an adversary relies on stealth, compromise is hidden until the attack is executed. That is, cyber compromise may be ongoing for a long period of time and undetected [14, 21]. In general, the longer an attacker waits to initiate the attack, the more likely it is that compromise will be detected and removed [33–35], suggesting that an optimal time of attack will emerge from the dynamics of compromise and detection.

To model this situation, we assume that compromise (both external and internal) increases until (i) it is detected through regular maintenance or (ii) the attacker decides to launch an attack at time t_A. We assume that if compromise is detected before the attack is launched, the value to the attacker is 0. Otherwise, when the hidden compromise becomes active, assessing the value of the attack requires a metric of utility for the attack, which we consider to be (1) the number of compromised cyber components or (2) the reduction in performance of the physical system.

2.3.1. Dynamics before Attack or Detection. For simplicity, we assume all cyber components are initially vulnerable to compromise. Since compromise is hidden until attack or discovery, $x_2 = 0$ always and no cyber components are being reset. Consequently before attack or detection, only equations (1), (3), and (5) are relevant. That is, cyber components are either un-compromised and vulnerable or compromised, but compromised cyber components are not recognized until either detected or the attack is executed. We let x denote the number of un-compromised and vulnerable cyber components so that equations (1) and (3) reduce to

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= -(c + c_s(1-\eta)y)x, \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= (c + c_s(1-\eta)y)x. \end{aligned} \tag{6}$$

Since $y(t) = N - x(t)$

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -(c + c_s(1-\eta)(N-x))x, \tag{7}$$

which has solution

$$x(t) = \frac{\left((1 - c_s(1-\eta))Ne^{-(c+c_s(1-\eta))t}\right)}{1 - c_s(1-\eta)e^{-(c+c_s(1-\eta))t}}, \tag{8}$$

from which $y(t)$ follows directly.

2.3.2. The Value of Attack. The probability that compromise remains hidden until the time of attack t_A is $U(t_A)$, determined by the solution of equation (5). The value of an attack requires that we introduce a utility function for the attacker. We will consider two choices. First, the value of attack may be measured in terms of the number of compromised cyber components at the time of attack. Second, the value of attack may be measured in terms of reduced performance of the physical system after the attack is executed.

The value of attack measured in terms of the expected number of compromised cyber components is

$$V_{A_1}(t_A) = y(t_A)U(t_A). \tag{9}$$

Since $y(t)$ is an increasing function of time and $U(t)$ is a decreasing function of time, their product will have a peak, leading to an optimal time of attack and times of attack around that optimum that are "pretty" good, in the sense of giving nearly the same value as the value at the optimal time of attack.

Assessing the value of attack measured in reduced performance of the physical system due to compromised cyber components requires a metric $\phi(x, y, z)$ for the performance of the physical system when the cyber system has x uncompromised and y compromised components, and level of DCM z. We describe such a performance function in the next section.

Given $\phi(x, y, z)$, the value to the attacker is the difference between $\phi(N, 0, 0)$ – performance of the physical system in the absence of compromised cyber components and DCM–and $\phi(x(t_A), y(t_A), 0)$–the performance of the physical subsystem when there are $x(t_A)$ un-compromised cyber components and $y(t_A)$ compromised cyber components at the time of attack, and no active DCM. The expected value of this performance function is

$$V_{A_2}(t_A) = U(t_A)[\phi(N, 0, 0) - \phi(x(t_A), y(t_A), 0)]. \tag{10}$$

Other choices of utility functions are possible, although as will be seen in the next section, a generic metric of performance of the physical system can capture a wide range of utility functions, from threshold relationships to linear ones. We also note that the value of attack to the adversary may not be the same as the cost to the owner of the system.

2.4. Metrics of Performance of the Cyber Physical System. In general, the cyber and physical systems are coupled nonlinearly. We first consider a generic measure of performance in which performance of the physical system is a function of the number of uncompromised cyber components, the fraction of decoy cyber components, the number of compromised cyber components, and the level of DCM. Performance increases as the total number of uncompromised, nondecoy cyber components increases, and decreases as both the total number of compromised cyber components y increases and the level of DCM increases.

We then turn to the synchronous motor model of the electric grid, with stability as the metric [21, 32], although we will show that both reliability and resilience of the grid emerge from the model.

2.4.1. A Generic Metric of Performance. We model performance function of the physical system as follows. First, we assume that performance of the physical system increases with the number of uncompromised cyber components according to

$$\phi\left(X|x_{50},\sigma_x\right)=\frac{1}{1+e^{\left(x_{50}-X\right)/\sigma_x}}, \tag{11}$$

where x_{50} is the number of uncompromised cyber components at which performance of the physical system is 50% of its maximum value (set to 1 without loss of generality) and σ_x captures the dispersal in performance of the physical system (Figure 2).

Equation (11) can accommodate a variety of assumptions about performance of the physical system. For example, if performance of the physical system is approximately a linear function of the total number of uncompromised cyber components, then parameters giving the dotted line in Figure 2(c) are appropriate and x_{50} can be chosen to give the 50% point of performance. On the other hand, if the physical system is a communications system and only one or a few cyber components need to be uncompromised for a message to get through, then a small value of x_{50} is appropriate. In communications systems, performance will be determined by bandwidth, connectivity, and message accuracy, all of which can be compromised; mission performance metrics are accuracy, timeliness, and completeness of the information. Mapping these onto the sigmoidal functions requires detailed knowledge of the particular system and is thus beyond the scope of this paper.

We use a similar kind of sigmoid to characterize the reduction in performance due to compromise of the cyber system and assume that DCM use computational bandwidth, thus also reducing performance of the physical system. We assume that when the level of DCM is z, performance of the physical system is reduced by a factor $e^{-\omega z}$, where $\omega > 0$.

A generic metric of performance of the physical system when a fraction η of the cyber components are decoys, and there are $X = x_1 + x_2$ uncompromised cyber components, y compromised cyber components, and the level of DCM is z, is then

$$\phi(X,y,z)=\left[\frac{1}{1+e^{\left(x_{50}-(1-\eta)X\right)/\sigma_x}}\cdot\frac{1}{1+e^{\left((1-\eta)y-y_{50}\right)/\sigma_y}}\right]\cdot e^{-\omega z}. \tag{12}$$

If the attacker needs to compromise many cyber components to degrade performance of the physical system, we set y_{50} to a moderately large value. On the other hand, if the attacker can shut down the physical system by compromising a few cyber components, then y_{50} will be small. The $1-\eta$ terms in equation (12) capture that decoy components in the cyber system are not part of the functionality of the physical system.

For computations using this generic performance function in the baseline case, we used $x_{50} = 40, \sigma_x = 10$, $y_{50} = 20, \sigma_y = 10$, and $\omega = 0.2$.

2.4.2. The One Generator-One Load-Many Consumers Electric Grid. One approach to modelling complex electric grids is to use second-order oscillator equations derived from Kirchoff's laws that balance power at each node of the grid [33–55]. That is, a power grid is modeled as a system of electrically coupled devices that deliver power from generators to machines/loads via transmission lines and a key objective of management of electrical grids is that generators and loads are properly synchronized. When a load is too strong, loads are unevenly distributed, or a major disruption occurs, and a generator or load may lose synchrony [20]. If this is sufficiently strong, it may lead to grid failure [50–55].

In the synchronous motor model, the i^{th} generator or load is characterized by a power balance equation [53] of the form $P_i = P_{i,\text{diss}} + P_{i,\text{acc}} + P_{i,\text{transmitted}}$ where the terms on the right hand side are, respectively, the rate at which the motor dissipates energy, the rate at which it accumulates energy, and the rate at which it transmits energy. The state of each generator or load is described by the rotor angle, measured relative to a reference device rotating at a standardized frequency, and satisfying the swing equation [50–54]

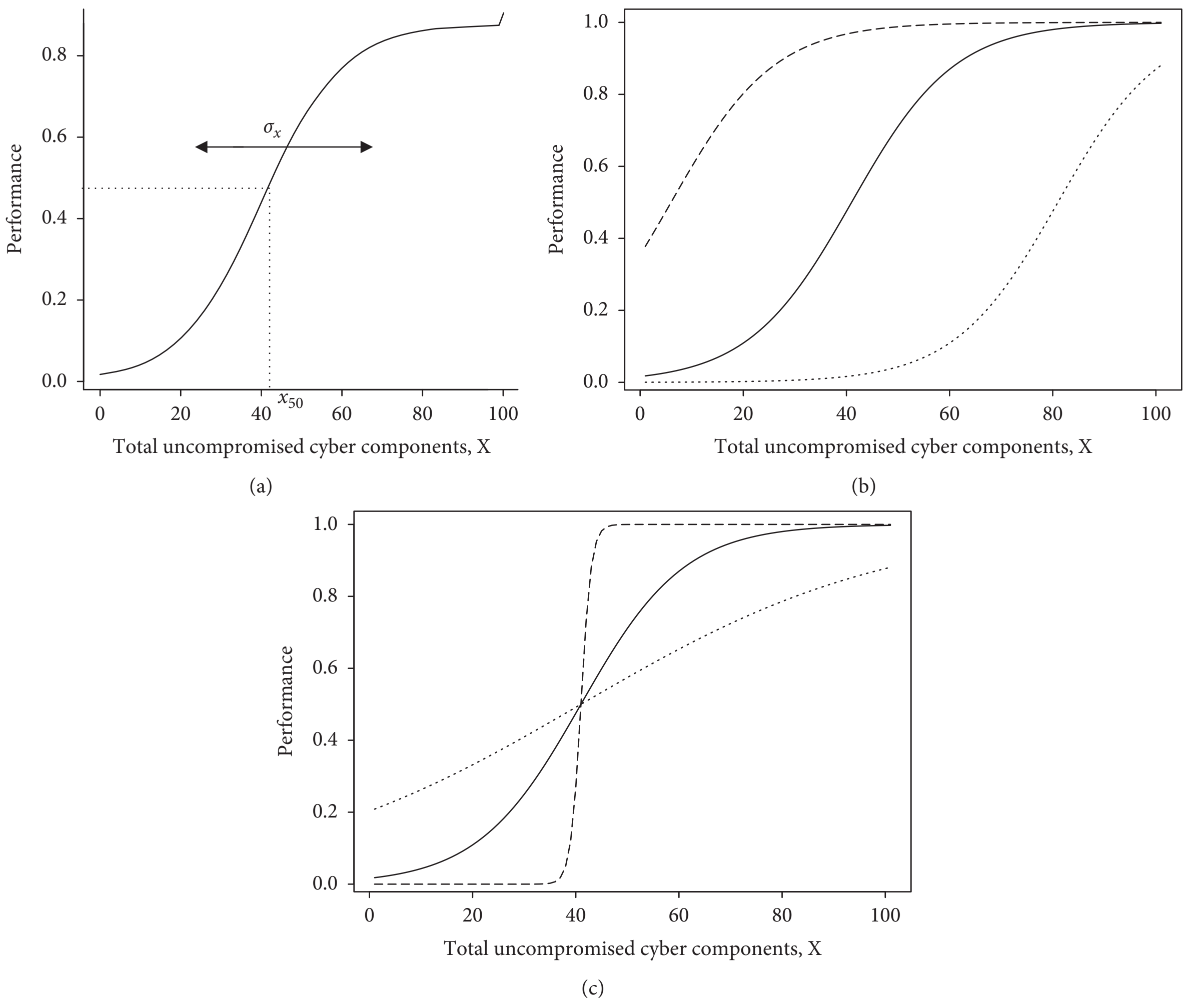
$$I_i\frac{\mathrm{d}^2\Theta_i}{\mathrm{d}t^2}+D_i\frac{\mathrm{d}\Theta_i}{\mathrm{d}t}=P_{i,\text{mech}}-P_{i,\text{elec}}, \tag{13}$$

where Θ_i is the deviation of the i^{th} rotor from the reference frequency, I_i is the moment of inertia of the rotor times the reference frequency, D_i is the coefficient of damping/friction as the rotor revolves, and $P_{i,\text{mech}}$ and $P_{i,\text{elec}}$ are, respectively, the mechanical power generated and the net electrical power transmitted through transmission lines.

A power grid is modeled as a coupled collection of equations similar to equation (13); and the deviation from the reference frequency for the i^{th} generator or load, denoted by θ_i, satisfies [53].

$$\frac{\mathrm{d}^2\theta_i}{\mathrm{d}t^2}=P_i-\delta_i\theta_i+\lambda\sum_{j\neq i}^{N}A_{ij}\sin\left(\theta_j-\theta_i\right), \tag{14}$$

where δ_i is the dissipation occurring in the i^{th} generator or load, λ is the coupling strength of the transmission line, and A_{ij} is a matrix whose entries are 1 if generators/loads i and j are connected, and 0 otherwise. $P_i > 0$ for generators and



(a)

(b)

(c)

Figure 2: (a) The generic performance function $1/(1 + e^{(x_{50} - X)/\sigma_x})$, which links the cyber and physical systems, depends on two parameters, x_{50} and σ_x. When $X = x_{50}$, the fraction is 0.5 so that x_{50} is the number of uncompromised cyber components at which performance of the physical system is 50% of its maximum value. The parameter σ_x captures the dispersal in performance. As σ_x declines, performance becomes more knife-edged; in the limit that $\sigma_x = 0$ performance is a step function that is 0 for $x < x_{50}$, 1/2 at $X = x_{50}$, and 1 for $X > x_{50}$. (b) Holding $\sigma_x = 10$, we show the sigmoid for $x_{50} = 5$, 40, or 80 (dashed, solid, and dotted lines, respectively). (c) Holding $x_{50} = 40$, we show the sigmoid for $\sigma_x = 1$, 10, or 30 (dashed, solid, and dotted lines, respectively).

$P_i < 0$ for loads. The power network will be in a steady state when consumption and generation of energy balance, so that if there are in total N generators and loads $\sum_{i=1}^{N} P_i = 0$.

In order to focus on compromise in the demand-side cyber subsystem, we use the simplest grid, as in [56], consisting of a single generator ($P_g = P_0$) and a single load ($P_l = -P_0$) (Figure 3) envisioned as a utility with many customers having AMI. The physical system consists of generators, substations, transformers, and towers and transmission lines; there are two cyber systems coupled to it. The first cyber system is the Supervisory Control And Data Acquisition (SCADA) [2, 37–39] system that collects measurements from substations and sends out control signals to equipment, substation automation or protection systems, and energy management systems. The second is the cyber system of consumers with AMI. To simplify the analysis, we focus on compromise of the latter cyber system and how compromise in it can destabilize the grid.

We let θ_g and θ_l denote the phase deviations for the generator and utility. In this case, A_{ij} reduces to a single value; we replace λA_{ij} by K, for simplicity set the dissipation parameters $\delta_g = \delta_l = \delta$, replace P_i by P_0 for the generator and $-P_0$ for the load, so that the equations for generator and load are

$$\begin{aligned} \frac{\mathrm{d}^2\theta_g}{\mathrm{d}t^2} &= P_0 - \delta\theta_g + K\sin\left(\theta_l - \theta_g\right), \\ \frac{\mathrm{d}^2\theta_l}{\mathrm{d}t^2} &= -P_0 - \delta\theta_l + K\sin\left(\theta_g - \theta_l\right). \end{aligned} \tag{15}$$

Subtracting the second equation from the first equation, we obtain a single second-order equation for the phase difference $\phi(t) = \theta_g(t) - \theta_l(t)$, written as the first-order system

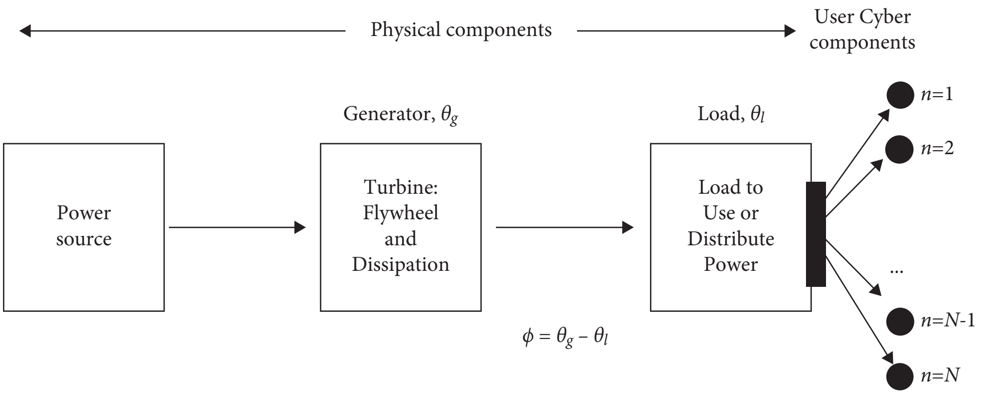


Figure 3: An electric grid consisting of one generator (with rotor angle θ_g) and one load (with rotor angle θ_l), in this case envisioned as a utility company that has N consumers, indexed by $n = 1, 2, 3, \ldots, N$, with AMI. Our analysis focuses on the consequences of compromise of the AMI rather than attacks on the cyber system of the power generating system.

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \nu,$$
$$\frac{\mathrm{d}\nu}{\mathrm{d}t} = 2P_0 - \delta\nu - 2K\sin(\phi). \tag{16}$$

For this model, the transmitted power is $\mathscr{P}_{\text{trans}} \propto K\sin(\phi)$ [53], and a natural metric of performance of this grid model is that the transmitted power is in a steady state. If $P_0 \le K$, this grid has a steady state at $\overline{\nu} = 0, \overline{\phi} = \arcsin(P_0/K)$. If $P_0 > K$, there is no steady state; instead the solution cycles. However, when $P_0 < K$ but close to K, if the dissipation coefficient is small enough, a periodic solution (and therefore unstable operation of the grid) coexists with a steady state. In particular [57–59], for each value of P_0, there is a critical value of the dissipation coefficient, which we denote by $\delta_c(P_0)$ so that if $\delta > \delta_c(P_0)$ then there is no periodic solution and the grid will be globally stable. However, when $0 < \delta < \delta_c(P_0)$, there is also a periodic solution ([52], Figure 2) so that the operation of the grid will be unstable; Rohden et al. [52] identify this situation with a power outage. Major power grids are often operated in the region in which K is on the order of P_0 to reduce the cost of overcapacity in transmission lines [53], which makes the possibility of transition from a steady state to oscillatory or chaotic behavior more likely.

For computations in the base case, we used $\phi(0) = \pi/4, \nu(0) = 0, K = 0.4$, and $\delta = 0.15$.

3. Results and Discussion

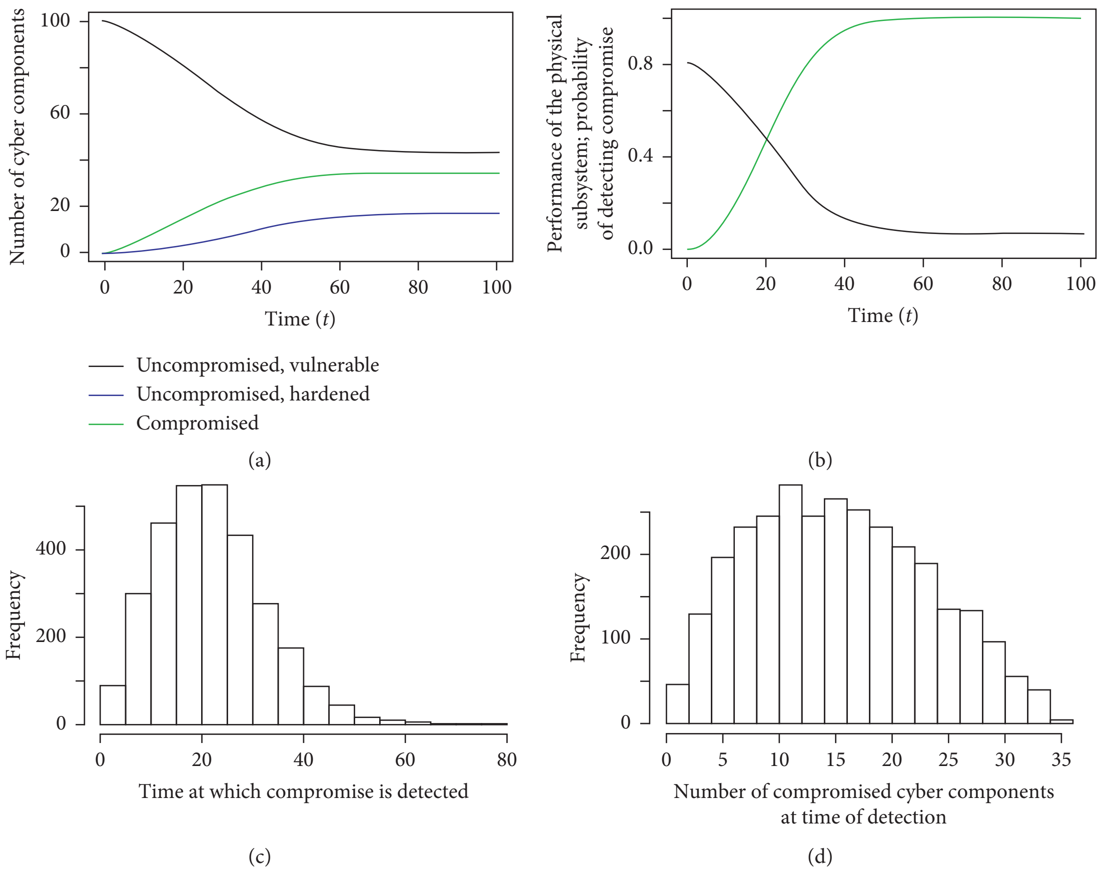
We begin (Section 3.1) by illustrating the basic dynamics of the model and associated performance of the physical system using the generic performance metric. These show that a steady-state level of compromise is reached, even with active defense, consistent with the observation that one should assume that cyber systems have been penetrated and focus on determining the consequences of the penetration [14]. Because of the mixed deterministic-stochastic nature of the model of compromise, we report distributions for the time of detection and level of compromise at the time of detection and then show realizations of the dynamics. As a sensitivity analysis (Section 3.2), we explore how the dynamics of compromise and performance of the physical system depends upon the fraction of decoy cyber components η. We then (Section 3.3) show results when an attacker relies on stealth, particularly how the optimal time of attack emerges and how it depends upon parameters of the system and the choice of metric for value to the attacker. We begin the section on the grid model by briefly summarizing (Section 3.4.1) the dynamics of the one generator-one load grid in the absence of compromise and then (Section 3.4.2) couple the model of compromise to the model of the grid. We illustrate EPRI scenario AMI.27 [20] for reverse engineering AMI by showing how smart meters sending misleading signals about power demand can lead to load-side failure of the electric grid and derive a canonical condition for failure of the grid in terms of the number of compromised components at the time of detection and the dissipation parameter of the synchronous motor model.

3.1. Dynamics of Compromise and Recovery.

3.1. Dynamics of Compromise and Recovery. For these results, we set $\eta = 0$. Absent DCM, compromise is removed only through regular maintenance. In Figure 4, we show the dynamics of the number of cyber components (Figure 4(a)), and performance of the physical system and the probability that compromise is detected (Figure 4(b)). We use the dynamics of the probability of detection to create stochastic realizations of the time of detection (Figure 4(c)) and the number of compromised cyber components at the time of detection of compromise (Figure 4(d)).

The cyber system reaches a steady state in which compromised cyber components co-exist with vulnerable and temporarily invulnerable cyber components. Consequently, performance of the physical system declines to a steady state.

Performance in Figure 4(b) can be compared directly to Figure 3.14 in [21], showing the time course of the fraction of the load delivered before, during, and after a cyberattack or Figures 2.1–2.3 in [32] characterizing the resilience of a system in response to a disruption/attack. When DCM become active, although performance of the



(a)

(b)

(c)

(d)

FIGURE 4: The dynamics of compromise and performance with DCM inactive when there are no decoy cyber components. (a) The numbers of un-compromised and vulnerable, un-compromised and temporarily invulnerable, and compromised cyber components. (b) The performance of the physical system (black) and the probability that compromise will be detected (green). Using the probability of detection curve in (b) we constructed 3000 realizations of (c) the time of detection of compromise and (d) the number of compromised cyber components at the time of detection of compromise.

physical system may increase because of the removal of compromise, performance of the physical system need not return to its previous state following disruption/attack (see below).

Conditioned on time of detection, the behavior of the full system is captured by equations (1)–(5), which we illustrate by choosing 8 times of detection that capture most of the range in Figure 4(c).

The cyber system recovers to more than 90% of its initial state (left panels in Figure 5). That is in the steady state, and regardless of the time of detection of compromise, there are uncompromised and vulnerable, uncompromised and currently hardened, and compromised cyber components (with the remainder resetting). Compromised components of the cyber system persist in the steady state because both external attack and internal co-compromise continue. The number of compromised components remains at a low level because DCM are maintained at a nonzero level (Figure 5, right hand panels, blue lines). The cost of a nonzero steady state of DCM is reduced performance. To illustrate this point, in Figure 6, we plot performance for the same times of detection as in Figure 5. Performance drops as compromise builds before detection of compromise and the minimum level of performance depends upon the time of detection of compromise. However, in the steady state, performance only returns to about 70% of its initial value.

3.2. The Role of Decoy Cyber Components. To explore the role of decoy cyber components, we swept over eight values of η, ranging from 0 to 0.25 (in the spirit of [60]). For each value of η, we determined 300 times of detection. Conditioned on both η and the time of detection, we computed the dynamics of the cyber system and the performance of the physical system.

Because of the formulation (Equations (1), (3), and (5)) as η increases, the probability of detecting compromise by a given time will increase, and the number of compromised cyber components at a given time will decline (Figures 7(a) and 7(b)). What cannot be predicted is that the standard deviation of both the time of detection and the number of compromised components at the time of detection of compromise decline as η increases.

In light of the results shown in Figures 5 and 6, we anticipate once DCM are activated, the system will recover

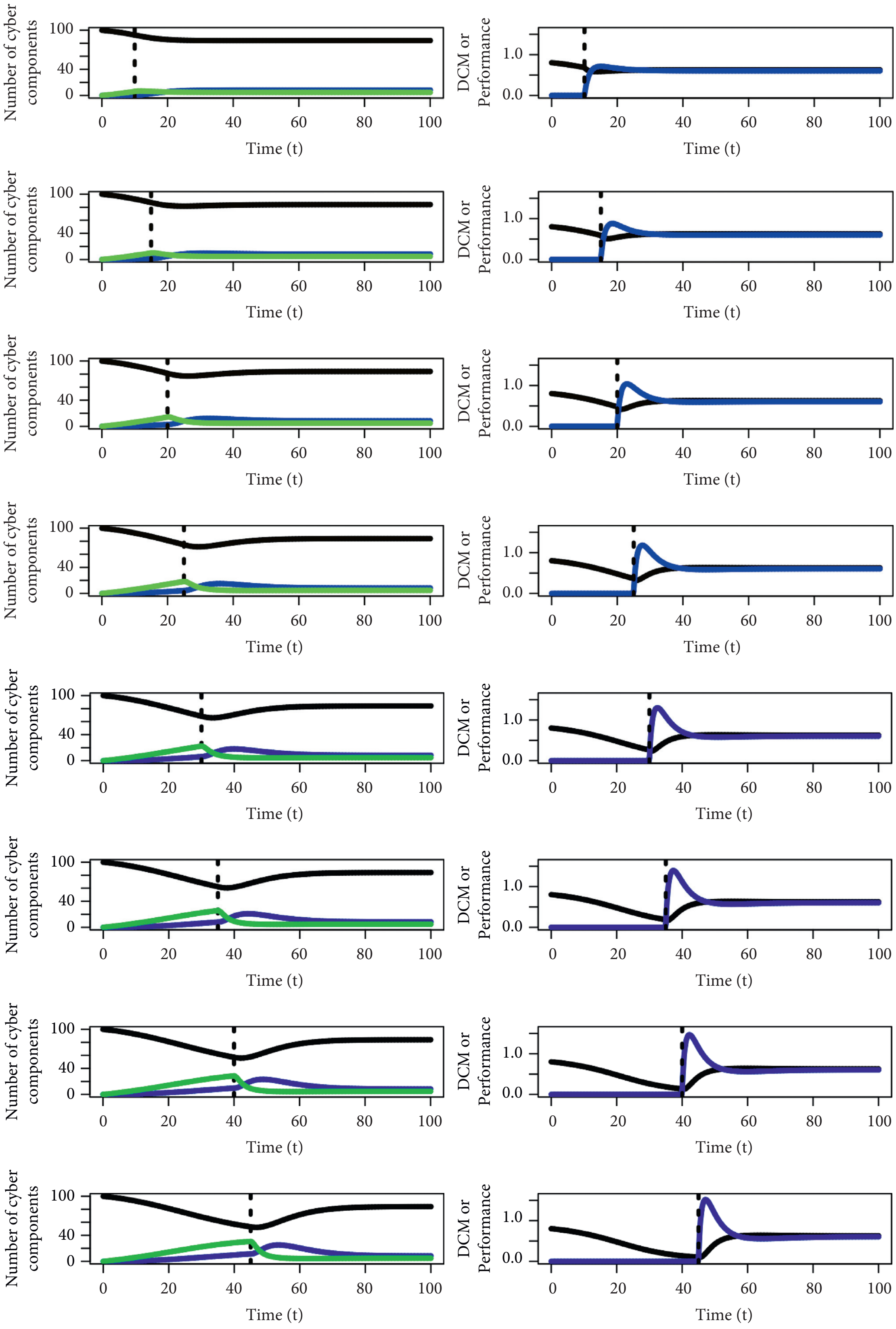


FIGURE 5: Dynamics of the cyber system and performance of the physical system when detection occurs at times drawn from the distribution in Figure 4(c), chosen to span most of the range of this distribution. The left hand column shows the number of uncompromised and vulnerable (black), uncompromised and temporarily hardened (blue), and compromised (green) cyber components as a function of time. The time of detection is denoted by a vertical dashed line. The right hand column shows performance of the physical system (black) and level of DCM (blue) as a function of time. As on the left, the vertical dotted line shows the time of detection.

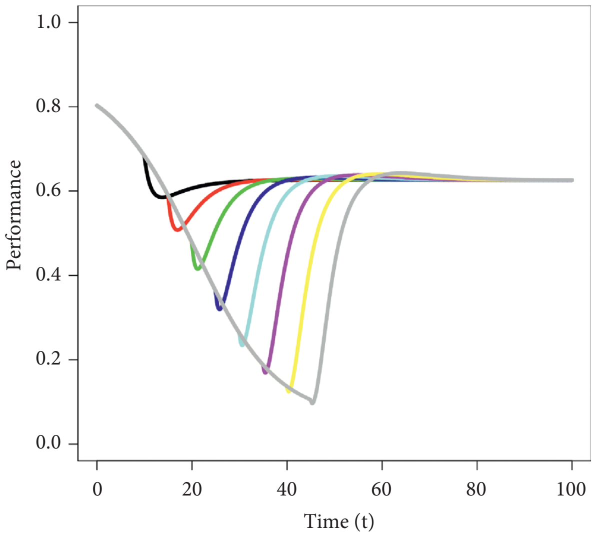


Figure 6: Performance of the physical system for the detection times in Figure 5. Although in all cases performance reaches the same steady state (about 70% of performance before compromise and activation of DCM), minimum performance varies considerably according to the time of detection of compromise.

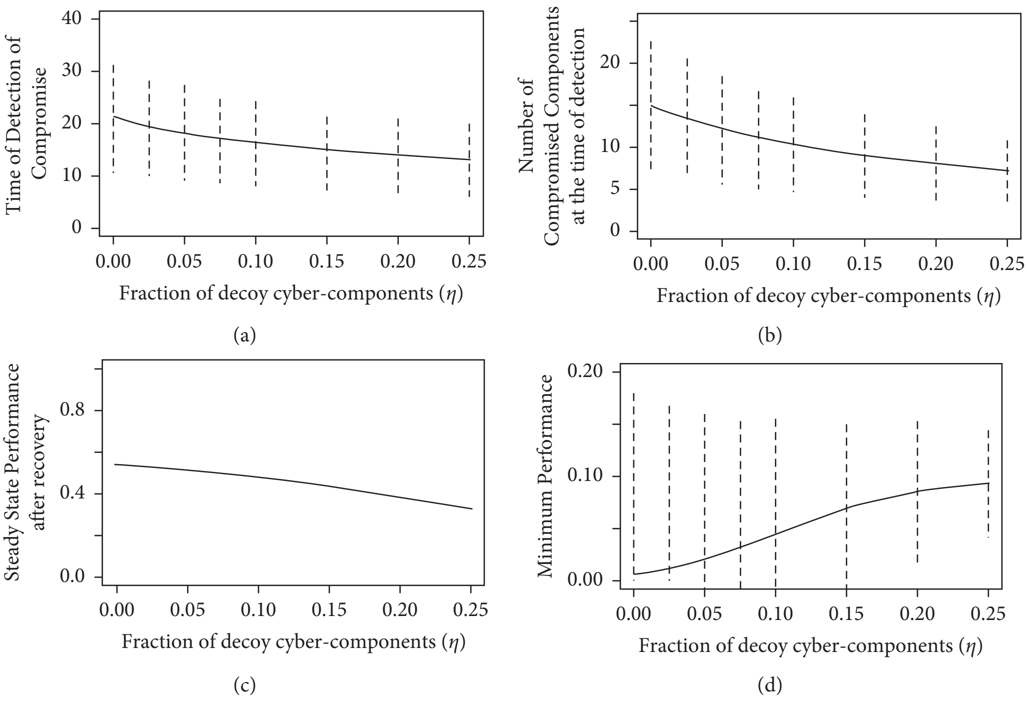


Figure 7: Sweeps over the fraction of decoy cyber components, η. (a) The mean (solid line) and standard deviation of the time of detection (dotted line) of the time at which compromise is detected. (b) The mean (solid) and standard deviation (dotted line) number of compromised cyber components at the time that cyber compromise is detected. (c) The steady-state performance of the physical system after recovery. In this case, we only show the mean because the standard deviation is much smaller than the mean. (d) The mean (solid line) and standard deviation (dotted line) of the minimum performance over the course of compromise and recovery.

and reach a steady state that is independent of both the time at which compromise is detected (the stochastic component) and varies in a deterministic way with η. The consequence (Figure 7(c)) is that although the mean of the steady-state performance of the physical system declines with η, the standard deviation of steady-state performance is virtually 0.

On the other hand, the minimum performance of the physical system depends upon the time at which compromise is detected (as in Figure 6), and thus on the value of η. Minimum performance is determined both by the level of compromise of cyber components at the time that compromise is detected and the response of DCM. In

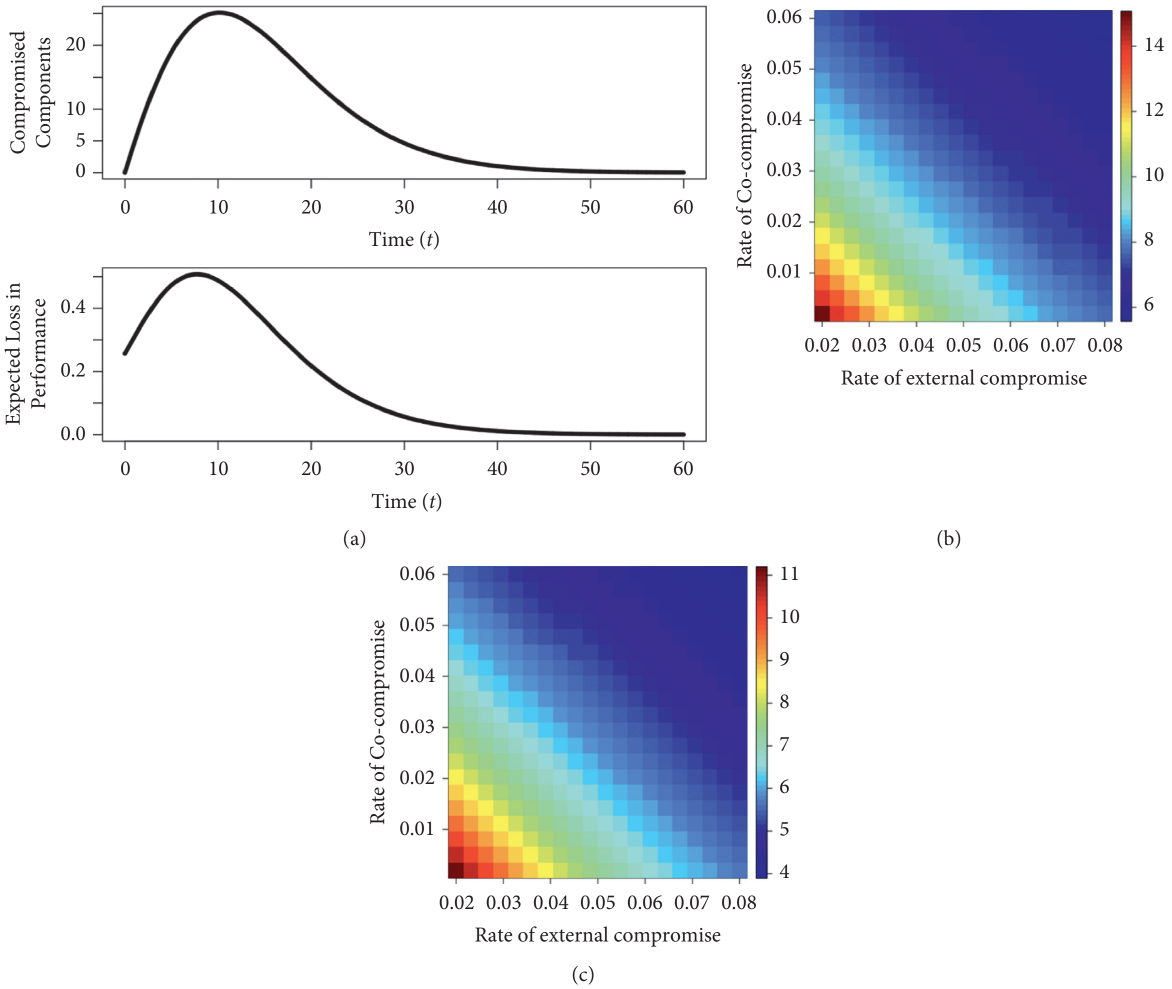


FIGURE 8: The value to an attacker when compromise is hidden until the time of attack. (a) The expected number of compromised cyber components (Equation (9)), upper panel, or expected loss in the performance of the physical system (Equation (10)), lower panel, as a function of the time at which the attack is executed, for the rate of external compromise $c = 0.05$ and rate of co-compromise $c_s = 0.02$. (b) The optimal time of attack across a range of values of c and c_s if the expected number of compromised cyber components is the metric of value to the attacker. (c) The optimal time of attack across a range of values of c and c_s when reduction in performance is the metric of value to the attacker.

Figure 7(d), we show the mean and standard deviation of the minimum performance of the physical system. The mean of minimum performance is an increasing function of η and the standard deviation of minimum performance is a decreasing function of η.

3.3. Stealth and the Optimal Time of Attack. When the attacker relies on stealth, the results become deterministic because the value to the attacker is an expectation over the time of detection In Figure 8(a), we show the time course of the two value functions. An optimal time of attack clearly emerges from these figures, as does a range of times of attack for which the value of attack is "pretty good" [61]. Conditioned on the other parameters, we predict the attack will occur earlier when value to the attacker is measured in terms of loss of performance of the physical system rather than the number of compromised cyber components. In Figures 8(b) and 8(c) we show the optimal time of attack across a range of values of c and c_s.

3.4. Load Side Failure of the One Generator-One Load Grid due to Compromise in AMI. We begin by summarizing the properties of the grid model in the absence of any compromise and then link it to the model of compromise.

3.4.1. Properties of the Grid Model in the Absence of Compromise. In the absence of compromise, the parameters $K = 0.4$, $\delta = 0.15$, and $P_0 = 0.25K, 0.35K$ or $0.55K$ and initial conditions $\phi(0) = \pi/4$ and $v(0) = 0$ lead to a steady state for equation (16). After transient initial dynamics, the steady state given by $\overline{v} = 0, \overline{\phi} = \arcsin(P_0/K)$ is reached. Since transmitted power is $\mathscr{P}_{\text{trans}} \propto K \sin(\phi)$ and the steady-state value of ϕ increases with P_0, transmitted power also increases with P_0.

As discussed above, as P_0 approaches K from below, instability may occur. For example, with all other parameters as above, for P_0 equal to $0.94K$ or $0.95K$, the grid reaches a steady state, but for $P_0 = 0.96K$ an oscillatory solution

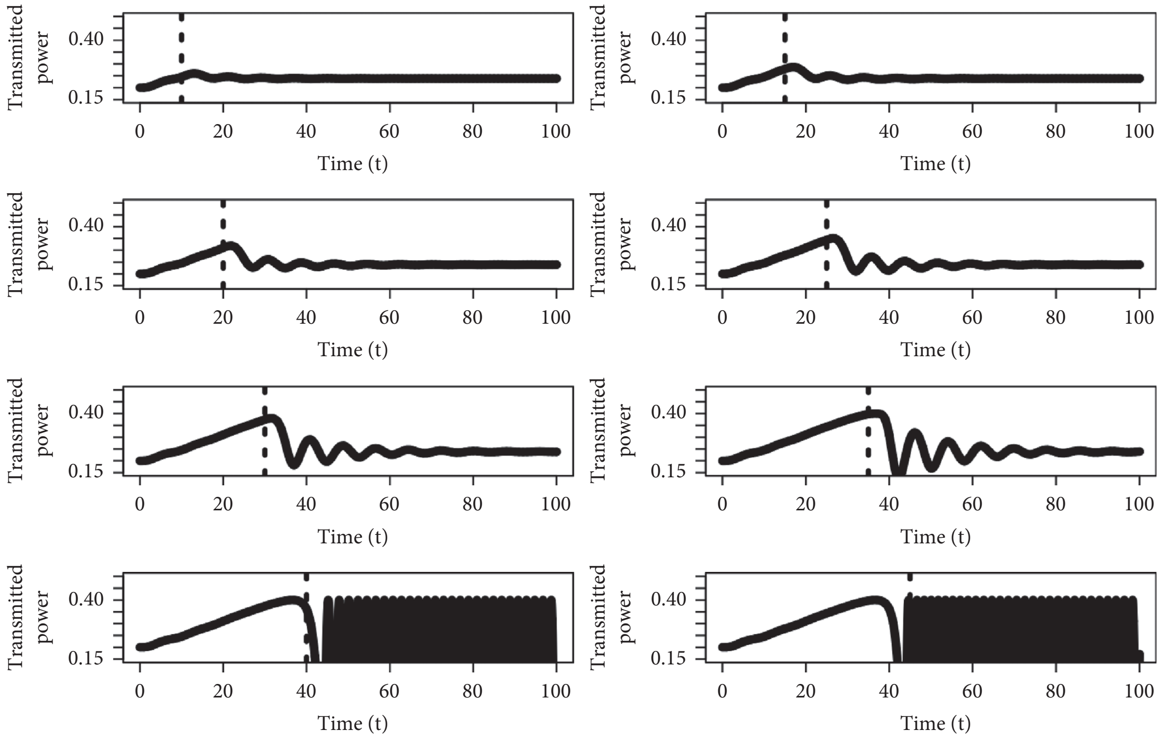


Figure 9: Transmitted power when the model for the grid and compromise are connected by assuming that P_0 in equation (16) is replaced by $P_0(1 + \varepsilon_d y(t))$ with $\varepsilon_d = 0.04$ representing how much one compromised AMI increases the demand for power, and $y(t)$ is the number of compromised cyber components. The dynamics of compromise and times of detection are the same as in shown in Figure 5–the vertical dotted line denotes the time at which compromise is detected.

develops. Conditioned on the other parameters, the critical value for δ falls between 0.177 (oscillations) and 0.178 (steady state approached).

These results suggest that one route to the instability of the electric grid (which we explore in detail below) is increased demands of power. To illustrate the concept without the complexity of the model for compromise and detection, we solved equation (16) for $\delta = 0.15, K = 0.4$ and $P_0 = 0.5K(1 + \varepsilon_p)$ where $\varepsilon_p = 0.75$, 0.8, 0.85, 0.9, and 0.95 characterizes the increase in demanded power. In this case, $P_0/K = 0.5(1 + \varepsilon_p) < 1$ so that a steady state exists. However, we expect instability before ε_p reaches 1. This is indeed the case–for a value of ε_p between 0.831 and 0.832, the grid loses stability. Furthermore, increasing δ makes the system more stable: when $\delta = 0.2$, the value of ε_p leading to instability falls between 0.8588 and 0.8589 and when $\delta = 0.25$ the value of ε_p leading to instability falls between 0.8844 and 0.8845. This is an example of criticality in physical systems as they operate closer to their capacity and become more fragile to perturbations. One of the roles of models such as we develop here is to help recognize and characterize potential modes of failure of the physical system.

3.4.2. Load Side Failure of the One Generator-One Load Grid due to Compromise of AMI. We couple the models for compromise and its detection and for the electric grid by replacing P_0 by $P_0(1 + \varepsilon_d y(t))$, where ε_d is how much one compromised AMI increases the demand for power, and $y(t)$, computed from equation (3), is the number of compromised AMI at time t. We assumed $\varepsilon_d = 0.04$ for computations. In Figure 9, we show transmitted power assuming that the dynamics of compromise and times of detection are the same as in shown in Figure 5.

The grid shows a signal of compromise with a secular, almost linear increase in the demand for power (the portion of the trajectories before the time of detection represented by the vertical dotted line). If detection occurs early enough (as in the first six panels), DCM are activated, compromise is reduced, and the grid returns to a stable steady state. However, as shown in the bottom two panels, if detection occurs too late, then the grid enters a region of instability even after DCM are activated.

To compute the risk to the grid due to compromise of AMI in the cyber system, we note that when power demanded by the load is $P_0(1 + \varepsilon_d y(t))$ and compromise is detected at time t_d, the condition for instability of the grid is

$$\varepsilon_d y(t_d) > \varepsilon_c(\delta), \tag{17}$$

which also can be written as

$$y(t_d) > \frac{\varepsilon_c(\delta)}{\varepsilon_d}. \tag{18}$$

Equation (18) is a canonical relationship linking compromise and load-side failure of the grid. Since detection of compromise is a stochastic process, the time of detection and

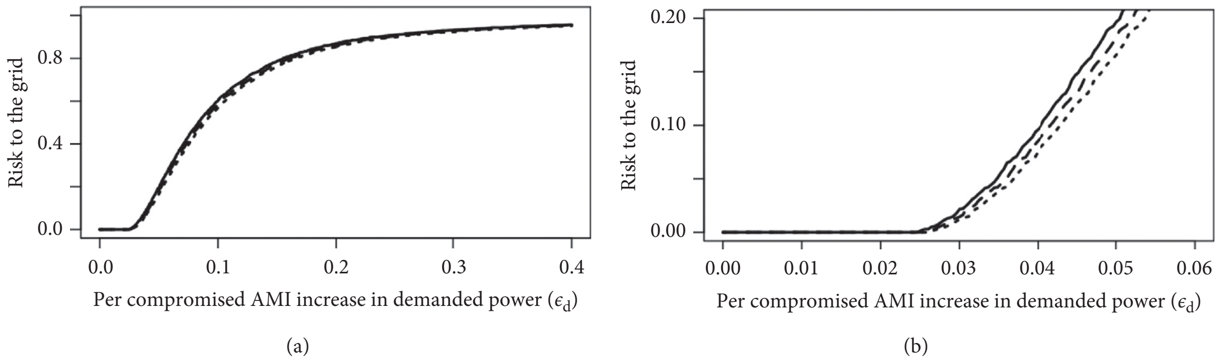


Figure 10: (a) Risk to the grid from compromise is determined by the increase in demanded power per-compromised AMI, ε_d, the critical value $\varepsilon_c(\delta)$ at which the grid becomes unstable when demanded power is $P_0(1+\varepsilon_c(\delta))$, and the level of compromise at the time of detection of compromise. We show risk to the grid over a range of values of ε_d for $\varepsilon_c(\delta) = 0.8315, 0.85885, 0.88445$ (solid, large dashed, and small dashed lines, respectively) (b) Expansion of the lower corner of (a).

level of compromise at the time of detection vary. Because detection of compromise is a random process, the level of compromise on the left-hand side of Equation (18) is a random variable. The risk of grid failure is the probability that the level of compromise on the left-hand side of Equation (18) exceeds the ratio on the right-hand side of that equation. In Figure 10, we show risk to the grid as a function of the per-compromised AMI increase in demanded power. If the per-compromised AMI increase in demanded power ε_d is small enough, the risk to the grid from compromise is virtually 0. There is a range in which the risk to the grid is a linear function of ε_d. For large enough values of ε_d, failure of the grid is guaranteed.

Thus, knowing the increase in demanded power by a compromised cyber component can be useful for predicting grid failure or not. For example, drawing a horizontal line in Figure 10(a) or 10(b) at the level of risk to the grid that can be tolerated and locating the value of ε_d at which the line intersects the curve will provide information on the value of per AMI increased demand in power that can be tolerated. The increase in demand is an anomalous operating condition, which may be used for detection of compromise ([21], Recommendation 4.10, pg 92).

4. Conclusions

Our work involves coupling of a set of nonlinear differential equations for the dynamics of compromise and defense of the cyber system with either (i) a nonlinear generic performance function to understand cyberattack via stealth or (ii) the nonlinear rotor equations for an electric grid to understand the role of load-side compromise on the stability of the grid. Doing so allowed us to explore how an optimal time of attack when using stealth emerges (Figures 8 and 9) and to derive a canonical condition (Equation (17) or (18)) for stability of the electric grid in the face of load-side compromise.

4.1. Novel Contributions of this Paper and Connection to Recent Other Work. The innovations of our work include (i) a fraction of decoy components in the cyber system, which are not connected to the rest of the cyber system or the physical system and thus do not spread compromise but increase the probability of detection of compromise, (ii) allowing components of the cyber system to return to the un-compromised state either temporarily invulnerable or immediately vulnerable, (iii) adaptive Defensive Counter Measures that respond adaptively to attack and compromise, (iv) a generic metric of performance of the physical system that depends upon the state of the cyber system, and (v) coupling a model of the electric grid to the model of compromise of the cyber system that leads to a condition for failure of the grid in terms of parameters of both compromise and the synchronous motor model.

4.2. Directions for Future Work. We discuss future work that is related to the model of cyber compromise and the generic performance function, the model of cyber compromise and the electric grid, and how the methods and results presented in our paper relate to recent studies [1, 62] on the general topic of security in systems with cyber and physical components.

4.2.1. Related to Cyber Compromise and the Generic Performance Function. Topics emerging from the model for cyber compromise and generic performance function warranting future investigation include: (i) The dynamics of compromise and recovery of the cyber system show that there is an opportunity for further investigation of operational resilience by planning for degradation of the cyber system when designing the linked physical system and (ii) when the attacker relies on stealth, a natural extension of our results is to explore how Figure 8 changes when the probability of detection depends upon the rate of external compromise, which sometimes generates a “wake” that is detectable.

4.2.2. Related to Cyber Compromise and the Electric Grid. We have shown how load-side compromise can destabilize an electric grid. If the metric of performance for an electric grid is the fraction of the load delivered, then equation (11) does not need any kind of rescaling to capture performance of the physical system, as can be seen by comparison of our Figure 2 with Figure 3.14 in [21].

Although defending the grid by disconnecting the utility from the generator is one route to defense, it is an extreme solution. On the other hand, developing a mechanism to recognize when individual consumers are demanding an anomalous amount of power and then disconnecting those consumers from the utility is a potential defense. Smart meters report on regular schedule with a defined message structure in a predictable way [21]. Signs of compromise or malfunction could include another reporting structure, reports at unusual times, or reported values that are outside of the usual range [18]. An implication of these results is that it is important to match the detection strategy to the willingness to accept risk when the underlying demand for power has a temporal pattern. That is, at times of high demand for power, compromise in the cyber components of AMI may lead to instability of the grid, suggesting that the strategy for detection of compromise should be adjusted to the pattern of demands for power. Models such as we have developed here can help in the allocation of resources for detection. To make this idea fully operational, one would include a detection model [30] with false positives (consumers are not requesting additional power but are seen to be doing so), false negatives (consumers with compromised AMI are not detected), and an economic model to assess the costs of power disruptions [2]. Doing so is beyond the scope of this paper.

At the level of the utility, one defense is to have a mechanism, via generalized governors or electric springs [63], that increases the value of δ as the perceived demand for power increases. In this case, the cost is that the utility operates less efficiently as δ increases. As in the case of disconnecting consumers, developing the requisite models is feasible but beyond the scope of this paper.

Another future step is to expand the number of generators and loads. The Zealand grid [53] is an excellent candidate for such a next step. Alternatively, one of the recommendations concerning resilience of electric grids is to rely on various types of micro-grids [21]. By their very structure, such grids will have many points of access that interface with the external world [51], thus raising an entirely new set of issues for which the models in this paper can be applied. Similarly, the development of consumer-generated solar power that can be moved to the electric company (thus changing the historical one way distribution of electricity from the power company to the consumer to bidirectional transfer of power) raises issues for which the models developed here can be adapted. Furthermore, as the electric grid becomes more and more dispersed, resistance to and recovery from cyberattacks will increasingly depend upon rapid or even real-time measurements and responses [22].

4.2.3. Connections to Other Recent Work. We now briefly discuss how the ideas in our paper link to [1] and an additional recent survey [62] of filtering in networked nonlinear systems. In [1], the authors describe approaches to tolerate cyberattacks (based approaches such as game theory or control theory); these can be explicitly modeled and tested for functional effectiveness of the cyber system using our approach. Our work complements that in [1] and in the future, our work can be used to explore the impact of cyberattack on a smart grid, examining the potential for access-and-maneuver types of attacks to disrupt system control and place the power system into ineffective and, in some cases, destructive states.

The survey in [62] raises a number of potential extensions of our work and directions for future exploration including.

(i) Our methods can be used for the analysis of particular communications and contract protocols, investigating within the system context which protocols are more or less resilient than others and the performance and effectiveness of different communications protocols/policies under different network conditions.

(ii) Using our methods, one can directly experiment with computational representation of more granular network-induced complexities, and validate the generality of the analytic method and associated assumptions.

(iii) A natural extension of our analytic methods is to treat other types or sources of volatility that may propagate or affect nodes in different ways.

(iv) Because of the explicit representation of the compromise of the cyber system, our methods can be used to explore how filtering methods can be used to improve response in the face of loss of nodes when nodes play both sensing and communication roles and to explore the consequences when sensor and communications functions are impacted asymmetrically [64].

(v) At a more theoretical level, our methods can be used to examine applicability of set-membership filtering to deal with system information censored or otherwise unavailable due to cyberattack [65].

4.3. Final Conclusion. In this paper, we provided a framework based on the population biology of disease for the analysis of compromise in complex systems with cyber and physical systems, gave examples of how the framework could be used for both a generic metric of performance and for a metric of performance involving the electric grid, and suggested opportunities for further explanation. Much remains to be done using this framework.

Appendix

A. Pseudocode for the Computations

The pseudocode for the computations given here follows the code available from the first author. Both the code and pseudocode are written in *R*, and organized in a way to allow other individuals to access the ideas directly. Thus, efficiency of computation has been sacrificed for clarity of development and presentation. In general, we adapt the form of the pseudocode from [65] and continue to use the mathematical notation from the paper, which is modified in clear ways in the actual code. Rscript can be obtained from the first author at marcmangel@protonmail.com.

A.1. Create the Generic Performance Function. To create the generic performance function in equation (12), use these steps.

(i) Specify the parameters $N, x_{50}, y_{50}, \sigma_x$, and σ_y.

(ii) Cycle over $x = 0$ to N.

(iii) Cycle over $y = 0$ to $N - x$.

(iv) For each combination of x and y, set $\phi(x) = e^{(x_{50}-x)/\sigma_x}/(1 + e^{(x_{50}-x)/\sigma_x})$ and $\phi(y) = e^{-(y_{50}-y)/\sigma_x}/(1 + e^{-(y_{50}-y)/\sigma_y})$

(v) The generic performance function in equation (12) for $z = 0$ is then the product $\phi(x) \cdot \phi(y)$.

A.2. Specify the Remainder of the Parameters. The remainder of the parameters are as follows:

(i) Those for the dynamics of compromise, defense, and detection (equations (1)–(5)): $N, c, c_s, \eta, g, f_1, f_2, \mu_m$, $\mu_{\text{DCM}}, \alpha, \beta, \gamma, M, \varepsilon_1, \varepsilon_2, \varepsilon_c$, and the time horizon T.

(ii) Those for the one-generator one-load model for an electric grid (Equation (16)): P_0, δ, and K.

A.3. Run the Model of Compromise in the Absence of Defensive Countermeasure. In order to produce the results shown in Figure 4(a), solve equations (1)–(3) and (5) in the absence of DCM (that is, set $z \equiv 0$). We used the 4th order Runge Kutta scheme in the package deSolve in R. The steps are as follows:

(i) Specify the time increment dt when solving equations (1)–(3) and (5).

(ii) Dimension the dynamic variables x_1, x_2, y, and U as vectors of length $T/\mathrm{d}t$.

(iii) Specify the initial conditions; if all components are initially uncompromised (as in this paper) these are $x_1 = N, x_2 = 0, y = 0$ and $U = 1$.

(iv) To produce Figure 4(b), solve the differential equations and link to the generic performance function from Section A.1.

A.4. Create the Distribution of Times at which Compromise is Detected and the Level of Compromise at the Time of Detection. In order to produce the results shown in Figures 4(c) and 4(d), follow the steps given below:

(i) Specify the number of replicates N_{sim} of the time of detection and create vectors t_{dec} and y_{dec} of length N_{sim}

(ii) Cycle n_{sim} from 1 to N_{sim}

(iii) For each n_{sim}, draw a random variable $\widetilde{U}$ uniformly distributed on [0, 1] and determine the time t for which $U(t) > \widetilde{U}$ and $U(t + \mathrm{d}t) \leq \widetilde{U}$. Set $t_{\text{dec}}(n_{\text{sim}})$ equal to this time and $y_{\text{dec}} = y(t_{\text{dec}})$, where $y(t)$ is computed from Section A. 3.

A.5. Run the Model of Compromise when Defensive Countermeasures are Activated. In order to produce the results shown in Figures 5 and 6 in which a range of times of detection is systematically evaluated, we now solve the full deterministic-stochastic model, which requires the indicator function $\mathcal{I}_{\text{DCM}}(t)$, which will be 0 for times less than the time of detection t_{dec} and 1 for times greater than it. We model this as the cumulative Gaussian distribution function, which is essentially a step function for small enough standard deviations. Thus, in the code, we compute by $\mathcal{I}_{\text{DCM}}(t)$ by adding one more equation to equations (1)–(5):

$$\frac{\mathrm{d}\mathcal{I}_{\text{DCM}}(t)}{\mathrm{d}t} = \frac{1}{\sqrt{2\pi\sigma_I^2}} \exp\left[-\frac{(t - t_{\text{dec}})^2}{2\sigma_I^2}\right], \tag{A.1}$$

where σ_I is the standard deviation (for computations, we used $\sigma_I = 0.1$) and t_{dec} is the time at which compromise is detected. Once equation (A.1) is appended to equations (1)–(5), one proceeds as follows:

(i) Specify the value of σ_I and the range of times of detection; for the results shown in the paper, we used the vector $t_{\text{dec}} = (10, 15, 20, 25, 30, 35, 40, 45)$.

(ii) Cycle over detection times.

(iii) Follow the same steps as in Section A.3 with equation (A.1) appended.

(iv) For each value of t_{dec}, confirm that the $\mathcal{I}_{\text{DCM}}(t)$ so generated is essentially a step function at t_{dec}.

A.6. Sweeping Over the Fraction of Decoy Components. In order to explore the role of decoy components (Figure 7), we convert η from a scalar to a vector; for computations we used $\eta = (0, 0.025, 0.05, 0.0075, 0.1, 0.15, 0.2, 0.25)$ and then proceed as follows:

(i) Convert the fraction of decoy components to a vector with the range to be explored.

(ii) Cycle over η

(iii) For each value of η repeat the steps in Sections A.3–A.5; doing so generates all of the data needed for Figure 7.

A.7. When the Attacker Relies on Stealth. When the attacker relies on stealth (Equation (8), Figure 8), we have an explicit solution for the number of uncompromised cyber components ($x(t)$ in equation (8)) and the number of compromised cyber components ($y(t) = N - x(t)$. This allows rapid exploration by sweeping over the rate of external compromise c and rate of co-compromise c_s, implemented with these steps.

(i) Specify the range over which the rates of compromise and co-compromise are explored and create or replace the scalars c and c_s by vectors. Specify a vector for the time t_A at which attack is initiated.

(ii) Cycle over the rate of external compromise c and the rate of co-compromise c_s

(iii) Cycle over time from $t = 0$ to $t = T$.

(iv) For each time, compute $x(t)$ from equation (8) and $y(t) = N - x(t)$.

(v) Cycle over each time of attack.

(vi) Compute the performance directly from equations (9), (10), and (12).

A.8. The One Generator-One Load Model for the Electric Grid in the Absence of Compromise. As noted in the text, the one generator-one load model may show instability depending on the parameter values in equation (16), for which the transmitted power is proportional to $K\sin(\phi)$. In order to explore the nature of this instability, proceed with these steps.

(i) Specify the scalars K and δ and a vector for the values of P_0 in equation (16).

(ii) As above, dimension vectors for v and ϕ of length T/dt and specify their initial values. Solve equation (16) (we used deSolve in R).

(iii) Repeat the previous step with different values of δ to numerically determine the value of δ at which an instability develops for given values of K and P_0.

(iv) In order to initially explore the failure of the grid due to increasing demand, choose values of K, δ, and P_0 for which the solution of equation (16) is stable. Create a vector ε_p that increases the demanded power.

(v) Cycle over ε_p and for each value of it, solve equation (16) as above to determine whether the grid has stable or oscillatory behavior.

A.9. Couple the Model of Compromise and the Mode of the Electric Grid. One is now in a position to couple the models for compromise and the electric grid, in which P_0 in equation (16) is replaced by $P_0(1 + \varepsilon_d y(t))$ where $y(t)$ is determined during the solution of equations (1)–(5) and proceeds with these steps.

(i) Specify the value of ε_d

(ii) Meld Sections A.5 and A.8 above, with P_0 in equation (16) being replaced by $P_0(1 + \varepsilon_d y(t))$. This will be sufficient to produce the results shown in Figure 9. To produce the results shown in Figure 10, one adds the steps.

(iii) Specify a vector for the increase in demanded power ε_d (which provides the x-axis in Figure 10) and critical values for ε_p determined from Section A.8; denote this vector by ε_c.

(iv) For each combination of time of detection t_{dec}, ε_d, and ε_c, find the fraction of $y(t_{\text{dec}})$ that exceeds the threshold given in equation (18).

Data Availability

All the results in this paper were generated using *R* version 3.6.1 (available at the CRAN website https://cran.r-project.org/) in RStudio version 1.0.143. *R* script for the models used in this paper can be obtained directly from the first author.

Conflicts of Interest

The authors have no conflicts of interest.

Acknowledgments

The authors thank Jeff Dunne for the suggestion of using a double logistic metric of performance, rather than a single logistic that depends only upon un-compromised cyber components; Sue Lee for modeling load-side failure of the electric grid due to compromised AMI; Paul Velez for the suggestion of adding the analogue of killer T cells; and colleagues in the Theoretical Ecology Group, Department of Biology, University of Bergen for comments in response to a talk by MM. The authors thank an anonymous referee for pointing out references [1, 62]. This work was performed as part of the employment of Jimmie McEver at the Johns Hopkins University Applied Physics Laboratory [JHU-APL]. Marc Mangel was supported by contract 15–1015 with JHU-APL.

References

[1] D. Ding, Q.-L. Han, X. Ge, and J. Wang, "Secure state estimation and control of cyber-physical systems: a survey," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 176–190, 2021.

[2] National Research Council, *Terrorism and the Electric Power Delivery System*, The National Academies Press, Washington, NJ, USA, 2012.

[3] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber-physical system security for the electric power grid," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 210–224, 2012.

[4] G. Casadei, D. Astolfi, A. Alessandri, and L. Zaccarian, "Synchronization in networks of identical nonlinear systems via dynamic dead zones," *IEEE Control Systems Letters*, vol. 3, no. 3, pp. 667–672, 2019.

[5] H. Yuan, Y. Xia, and H. Yang, "Resilient state estimation of cyber-physical system with multichannel transmission under DoS attack," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–12, 2020.

[6] B. K. Mishra and G. M. Ansari, "Differential epidemic model of virus and worms in computer network," *International Journal on Network Security*, vol. 14, pp. 88–94, 2012.

[7] B. K. Mishra and S. K. Pandey, "Fuzzy epidemic model for the transmission of worms in computer network," *Nonlinear Analysis: Real World Applications*, vol. 11, no. 5, pp. 4335–4341, 2010.

[8] J. Morris-King and H. Cam, "Ecology-inspired cyber risk model for propagation of vulnerability exploitation in tactical edge," in *Proceedings of Milcom 2015 Track 3-Cyber Security*

and Trusted Computing, pp. 336–341, Baltimore, MD, November 2015.

[9] B. Nguyen, "Modelling cyber vulnerability using epidemic models," in *Proceedings of the 7th International Conference on Simulation and Modeling Methodologies, Technologies and Applications (SIMULTECH 2017)*, pp. 232–239, Madrid, Spain, July, 2017.

[10] C. Zhan, C. K. Tse, and M. Small, "A general stochastic model for studying time evolution of transition networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 464, pp. 198–210, 2016.

[11] M. Mangel, "The theoretical biologist's toolbox," *Quantitative Methods for Ecology and Evolutionary Biology*, Cambridge University Press, Cambridge, UK, 2006.

[12] M. Nowak and R. M. May, *Virus Dynamics. Mathematical Principles of Immunology and Virology*, Oxford University Press, Oxford, UK, 2007.

[13] D. Wodarz, *Killer Cell Dynamics. Mathematical and Computational Approaches to Immunology*, Springer, New York, NY, USA, 2007.

[14] M. C. Libicki, L. Ablon, and T. Webb, *The Defender's Dilemma. Charting a Course toward Cybersecurity*, Rand Corporation, Santa Monica, CA, USA, 2015.

[15] F. Cleveland and A. Lee, *Cyber Security for DER Systems*, EPRI (Electric Power Research Institute), Palo Alto, CA, USA, 2013, https://smartgrid.epri.com/doc/der%20rpt%2007-30-13.pdf.

[16] SGIP [Smart Grid Interoperability Panel], *Introduction to NISTIR 7628 Guidelines for Smart Grid Cyber Security*, SGIP, Washington, NJ, USA, 2010, https://www.smartgrid.gov/files/nistir_7628_.pdf.

[17] A. Srivastava, T. Morris, T. Ernster, C. VEllaithurai, S. Pan, and U. Adhikari, "Modeling cyber-physical vulnerability of the smart grid with incomplete information," *IEEE Transactions on Smar Grid*, vol. 4, pp. 235–244, 2012.

[18] C.-C. Sun, A. Hahn, and C.-C. Liu, "Cyber security of a power grid: state-of-the-art," *International Journal of Electrical Power & Energy Systems*, vol. 99, pp. 45–56, 2018.

[19] M. P. Kokare and S. H. Pawar, "Energy monitoring system in electric grids: the role of advanced intelligent and IOT for future electric grid," in *Proceedings of the International Conference on Emerging Trends in Information Technology and Engineering (Ic-ETITE)*, Vellore, India, September 2020.

[20] NESCOR, "Electric sector failure scenarios and impact analyses," 2013, https://smartgrid.epri.com/doc/NESCOR%20failure%20scenarios09-13%20finalc.pdf.

[21] National Academies of Sciences, Engineering, and Medicine, *Enhancing the Resilience of the Nation's Electricity System*, The National Academies Press, Washington, NJ, USA, 2017.

[22] National Research Council, *The Resilience of the Electric Power Delivery System in Response to Terrorism and Natural Disasters: Summary of a Workshop*, The National Academies Press, Washington, NJ, USA, 2013.

[23] J. P. Carlin and G. M. Graff), *Dawn of the Code War and the Rising Global Cyber Threat*, Hachette Book Group, New York, NY, USA, 2018.

[24] E. Gartzke and J. R. Lindsay, "Weaving tangled webs: offense, defense, and deception in cyberspace," *Security Studies*, vol. 24, no. 2, pp. 316–348, 2015.

[25] P. G. Fennell and J. P. Gleeson, "Multistate dynamical processes on networks: analysis through degree-based approximation frameworks," *SIAM Review*, vol. 61, no. 1, pp. 92–118, 2019.

[26] D. Merl, L. R. Johnson, R. B. Gramacy, and M. Mangel, "A statistical framework for the adaptive management of epidemiological interventions," *PLoS One*, vol. 4, no. 6, p. e5807, 2009.

[27] D. Merl, L. R. Johnson, R. B. Gramacy, and M. Mangel, "amei: an R package for the adaptive management of epidemiological interventions," *Journal of Statistical Software*, vol. 32, no. 6, pp. 1–32, 2010.

[28] F. Ben-Ami, R. R. Regoes, and D. Ebert, "A quantitative test of the relationship between parasite dose and infection probability across different host-parasite combinations," *Proceedings of the Royal Society B: Biological Sciences*, vol. 275, no. 1636, pp. 853–859, 2008.

[29] R. R. Regoes, J. W. Hottinger, L. Sygnarski, and D. Ebert, "The infection rate of *Daphnia magna* by *Pasteuria ramosa* conforms with the mass-action principle," *Epidemiology and Infection*, vol. 131, no. 2, pp. 957–966, 2003.

[30] D. Ding, Q.-L. Han, Y. Xiang, X. Ge, and X.-M. Zhang, "A survey on security control and attack detection for industrial cyber-physical systems," *Neurocomputing*, vol. 275, pp. 1674–1683, 2018.

[31] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2715–2729, 2013.

[32] H. H. Willis and L. Loa, *Measuring the Resilience of Energy Distribution Systems*, https://www.rand.org/pubs/research_reports/RR883.html, RAND Corporation, Santa Monica, CA, USA, 2015, https://www.rand.org/pubs/research_reports/RR883.html.

[33] R. Axelrod and R. Iliev, "Timing of cyber conflict," *Proceedings of the National Academy of Sciences*, vol. 111, no. 4, pp. 1298–1303, 2014.

[34] B. Edwards, A. Furnas, S. Forrest, and R. Axelrod, "Strategic aspects of cyberattack, attribution, and blame," *Proceedings of the National Academy of Sciences*, vol. 114, no. 11, pp. 2825–2830, 2017.

[35] R. M. Lee, M. J. Assante, and T. Conway, *Analysis of the Cyber Attack on the Ukrainian Power Grid*, E-ISAC, Washington, NJ, USA, 2016.

[36] M. F. Wolff, P. G. Lind, and P. Maass, "Power grid stability under perturbation of single nodes: effects of heterogeneity and internal nodes," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 10, Article ID 103120, 2018.

[37] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation & Control*, John Wiley & Sons, New York, NY, USA, 1984.

[38] P. Kundur, *Power Stability and Control*, McGraw-Hill, New York, NY, USA, 1994.

[39] J. J. Grainger and W. D. Stevenson Jr., *Power System Analysis*, McGraw-Hill, New York, NY, USA.

[40] L. Cao, C. T. Tian, Z. Wang, X. Zhang, and Z. Liu, "Influence of stochastic perturbations on the cluster explosive synchronization of second-order Kuramoto oscillators on networks," *Physical Review E*, vol. 97, Article ID 22220, 2019.

[41] J. Gao and K. Efstathiou, "Self-consistent method and steady states of second-order oscillators," *Physical Review E*, vol. 98, Article ID 42201, 2018.

[42] F. Dörfler and F. Bullo, "Synchronization and transient stability in power networks and nonuniform Kuramoto oscillators," *SIAM Journal on Control and Optimization*, vol. 50, no. 3, pp. 1616–1642, 2012.

[43] B. Ermentrout, Y. Park, and D. Wilson, "Recent advances in coupled oscillator theory," *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 377, no. 2160, Article ID 20190092, 2019.

[44] K. Schmietendorf, J. Peinke, R. Friedrich, and O. Kamps, "Self-organized synchronization and voltage stability in networks of synchronous machines," *The European Physical Journal-Special Topics*, vol. 223, no. 12, pp. 2577–2592, 2014.

[45] S. Kettemann, "Delocalization of disturbances and the stability of ac electricity grids," *Physical Review*, vol. 94, Article ID 62311, 2016.

[46] P. Ji, T. K. D. Peron, P. J. Menck, F. A. Rodrigues, and J. Kurths, "Cluster explosive synchronization in complex networks," *Physical Review Letters*, vol. 110, no. 21, Article ID 218701, 2013.

[47] P. Ji, T. K. Peron, F. A. Rodrigues, and J. Kurths, "Analysis of cluster explosive synchronization in complex networks," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 90, Article ID 62810, 2014.

[48] T. K. DM. Peron, P. Ji, F. A. Rodrigues, and J. Kurths, "Effects of assortative mixing in the second-order Kuramoto model," *Physical Review*, vol. E91, Article ID 52805, 2015.

[49] M. Rohden, A. Sorge, M. Timme, and D. Witthaut, "Self-organized synchronization in decentralized power grids," *Physical Review Letters*, vol. 109, Article ID 64101, 2012.

[50] D. Manik, M. Rohden, H. Ronellenfitsch et al., "Network susceptibilities: theory and applications," *Physical Review*, vol. 95, Article ID 12319, 2017.

[51] M. Rohden, D. Jung, S. Tamrakar, and S. Kettemann, "Cascading failures in ac electricity grids," *Physical Review*, vol. 94, Article ID 32209, 2016.

[52] G. Filatrella, A. H. Nielsen, and N. F. Pedersen, "Analysis of a power grid using a Kuramoto-like model," *The European Physical Journal B*, vol. 61, no. 4, pp. 485–491, 2008.

[53] T. Nishikawa and A. E. Motter, "Comparative analysis of existing models for power-grid synchronization," *New Journal of Physics*, vol. 17, no. 1, Article ID 15012, 2015.

[54] D. Z. Szabó, P. Duck, and P. Johnson, "Optimal trading of imbalance options for power systems using an energy storage device," *European Journal of Operational Research*, vol. 1, 2020.

[55] S. Liu, S. Mashayekh, D. Kundur, T. Zourntos, and K. Butler-Purry, "A framework for modeling cyber-physical switching attacks in smart grid," *IEEE Transactions on Emerging Topics in Computing*, vol. 1, no. 2, pp. 273–285, 2013.

[56] M. Levi, F. C. Hoppensteadt, and W. L. Miranker, "Dynamics of the josephson junction," *Quarterly of Applied Mathematics*, vol. 36, no. 2, pp. 167–198, 1978.

[57] M. Gittelman, *The Chaotic Pendulum*, World Scientific, Singapore, 2010.

[58] H. Risken, "The fokker-planck equation," *Methods of Solution and Applications*, Springer-Verlag, Berlin, Germany, 1989.

[59] M. Mangel and S. B. Munch, "A life-history perspective on short- and long-term consequences of compensatory growth," *The American Naturalist*, vol. 166, no. 6, pp. E155–E176, 2005.

[60] M. Mangel, "Applied mathematicians and naval operators," *SIAM Review*, vol. 24, no. 3, pp. 289–300, 1982.

[61] J. Mao, Y. Sun, X. Yi, H. Liu, and D. Ding., "Recursive filtering of networked nonlinear systems: a survey," *International Journal of Systems Science*, vol. 52, no. 6, pp. 1110–1128, 2021.

[62] E. F. Areed, M. A. Abido, A. T. Al-Awami, and S. A. Hussain, "Electric spring average model development and dynamic analysis for demand-side management," *IET Smart Grid*, vol. 3, no. 2, pp. 226–236, 2020.

[63] H. Wang, S. Xie, B. Zhou, and W. Wang, "Non-fragile robust H ∞ filtering of Takagi-Sugeno fuzzy net-worked control systems with sensor failures," *Sensors*, vol. 20, 2020.

[64] J. Li, G. Wei, D. Ding, and Y. Li, "Set-membership filtering for discrete time-varying nonlinear systems with censored measurements under round-robin protocol," *Neurocomputing*, vol. 281, pp. 20–26, 2019.

[65] R. Hilborn and M. Mangel, *The Ecological Detective. Confronting Models with Data*, Princeton University Press, Princeton, NJ, USA, 1997.