9-2022

Mitigating Cascading Failures in Power Grids via Markov Decision-Based Load-Shedding With DC Power Flow Model

Pankaz Das
Marquette University

Rezoan Ahmed Shuvro
Marquette University

Kassie Povinelli
Marquette University

Francesco Sorrentino
University of New Mexico

Majeed M. Hayat
Marquette University, majeed.hayat@marquette.edu

IEEE Industrial Electronics Magazine, Vol. 16, No. 3 (September 2022): 4048-4059. DOI.

Mitigating Cascading Failures in Power Grids via Markov Decision-Based Load-Shedding with DC Power Flow Model

Pankaz Das
Department of Electrical and Computer Engineering, Marquette University, Milwaukee, WI
Rezoan A. Shurvo
Department of Electrical and Computer Engineering, Marquette University, Milwaukee, WI
Kassie Povinelli
Department of Electrical and Computer Engineering, Marquette University, Milwaukee, WI
Francesco Sorrentino
Department of Mechanical Engineering, University of New Mexico, Albuquerque, NM
Majeed M. Hayat
Department of Electrical and Computer Engineering, Marquette University, Milwaukee, WI

Abstract:

Despite the reliability of modern power systems, large blackouts due to cascading failures (CFs) do occur in power grids with enormous economic and societal costs. In this article, CFs in power grids are theoretically modeled proposing a Markov decision process (MDP) framework with the aim of developing optimal load-shedding (LS) policies to mitigate CFs. The embedded Markov chain of the MDP, established earlier to capture the dynamics of CFs, features a reduced state-space and state-dependent transition probabilities. We introduce appropriate actions affecting the dynamics of CFs and associated costs. Optimal LS policies are computed that minimize the expected cumulative cost associated with CFs. Numerical simulations on the IEEE 118 and IEEE 300 bus systems show that the actions derived by the MDP result in minimum total cost of CFs, compared to fixed and random policies. Moreover, the optimality of derived policies is validated by a CF simulation based on dc power flow for the IEEE 118 bus system. Therefore, such actions developed by the proposed theoretical MDP framework can serve as a baseline for devising optimal LS strategies to mitigate CFs in power grids.

SECTION I. Introduction

Power systems are notably vulnerable to cascading failures (CFs) in the power-transmission network, which can occur due to either natural or intentional events [1]–[5]. CFs in power systems are described as a sequence of correlated and uncontrolled failures of network components [e.g., nodes (e.g., generators, transformers, etc.) and edges (e.g., transmission-lines)] that successively weakens the power system [6]. In addition, the propagation of CFs in power grids is largely accelerated by the interdependencies between power and communication neworks [7], [8]. CFs often result in huge monetary losses and have broad societal impacts, as observed in the case of the North American blackouts [9] and other blackouts worldwide (see [10, Table 1]). Since we have limited control over the contingencies (natural or human-made intentional attacks) that initiate CFs in complex power grids, it is utmost important to develop a strategy to control CFs before they spread. Even with current sophisticated control mechanisms, CFs occur and cause large power blackouts. Thus, having a policy for possible contingencies, which depends on the evolution stages of CFs, is fundamental for mitigating CFs in power grids and protecting them from large blackouts. A policy, i.e., a sequence of corrective actions, to prevent the spread of CFs in power grids can be automated or human-made. Determination of the policy is usually based upon a given objective, such as maximizing the power delivery or minimizing the spread of a CF.

The main two corrective methods to mitigate CFs in power grids are generation redispatch and load-shedding (LS), along with some other measures [11]. The following considerations are the main motivations for our decision to focus on LS, which is also broadly investigated as a mechanism to maintain grid reliability (please see [12]–[16] and references therein). First, we assumed that a generation dispatch strategy is integrated as a part of the operator's control scheme (due to the dynamics of the economics (e.g., market price) of grid operations involved in generation dispatch). However, the economics of generation dispatch considering market prices are beyond the scope of this article. Second, since LS is a quick, simple, and reliable remedy to establish active power balance [11], [17], it is a commonly adopted strategy by grid operators for emergency control exemplified by cases when a CF propagates very fast and escalates on a very short time scale. In contrast, generation dispatch is typically used when there is sufficient time for the grid operators to diagnose a situation and take preventive actions (e.g., transmission loading relief and redispatch normally took 10–30 mins [18], [19]). Third and most importantly, LS affects the customers directly (due to power outage) but generation dispatch does not necessarily involve power outage. We tied power outage from LS in the MDP cost function and minimized customer impact induced by the outage from CF by prescribing optimal actions for LS.

In this article, we model the time-varying and state-dependent characteristics of LS actions in power grids using a Markov decision process (MDP) and develop optimal LS policies for the grid operators to mitigate CFs based on contingencies. A Markov chain (MC) was established in [20] to capture the dynamics of CFs in power grids. In

this article, in addition to capturing the dynamics of CFs, we propose an MDP model for mitigating CFs in power grids through dynamic actions. Specifically, the newly introduced actions in the MDP correspond to time-varying LS interventions during a CF. Moreover, state- and action-dependent costs corresponding to each action are introduced in the MDP and a tradeoff is established between the number of line failures (LFs) and the amount of LS. Most importantly, this article addresses the challenge of finding the optimal LS actions at each time step and for each state of the MDP in order to mitigate CFs.

The proposed model offers the following major contributions beyond existing stochastic decision models for mitigating CFs. First, it lays a framework, using the MDP, as a tool to dynamically mitigating CFs by considering the cost of LF an LS. The MDP model is engineering based, i.e., it can incorporate key physical and operating characteristics of the power grids, such as the loading level, the ability to implement LS, and the line capacity estimation error. Second, the definition of the discrete action space and the adoption of the reduced state space ensure the scalability and analytical tractability of the MDP as the dimension of the power grid topology increases. Third, and most importantly, we propose a cost function that balances the tradeoff between the amount of LS and the number of LFs.

SECTION II. Related Works

CFs occur in a variety of networked systems, such as power transmission grids [21]–[23], other cyber-physical systems, including interdependent networks, e.g., power grids and communication systems [7], [24], as well as supply networks [25]. Due to the relevance of this article to power systems, we will discuss the mitigation strategies that has been proposed for protecting power grids from CFs. A multiobjective optimization framework to optimally allocate link capacities for cascade-resilient power transmission network was proposed in [26]. Parandehgheibi *et al.* [27] proposed a two-phase LS policy for mitigating CFs in an interdependent power grid and communication network. The authors in [28] modeled power-communication interdependency, and proposed an algorithm for identifying the failed lines, and subsequently proposed a mitigation strategy to limit the cascade propagation by minimizing generator redispatch and LS. An optimal LS algorithm, based on a linear quadratic cost function, allowed computation of optimal LS amounts in both deterministic and stochastic settings [29]. Similarly, using convex relaxation of the ac power flow, an LS algorithm was proposed to mitigate the deterministic propagation of CFs in [30]. In contrast, in [31], a dc power flow-based hidden failure model for CFs in power grids and provided mitigation strategies; the model aimed to keep key grid parameters below a critical level. Yao *et al.* [32] used a Markovian tree search-based risk management scheme to reduce the risk of cascading outages. Similarly, a very recent work [33] determined the critical lines of a grid that are most likely responsible for instigating CFs using a Markovian influence graph generated from historical data and suggested a mitigation strategy by using resources to harden these critical lines. It was argued in [34] that hardening the critical lines may cause the remaining lines to become critical.

The authors in [35] proposed a metric for comparing a dc power flow-based model and a transient stability analysis-based model in terms of their performance in modeling CFs. Later, the authors formulated sequential attacks on the grid topology, using Q-learning based on a dc power flow model for CFs, and obtained effective attack strategies that maximize the number of LFs with a reduced number of attacks [36]. In addition, an interaction model is proposed for capturing the dynamics of cascading power grids that estimates the interaction matrix and the interaction network from CF simulations using ac power flow [37]. Similar to [37], we estimate the transition probability of the MC using data from CFs simulations, which are based on the optimal dc power flow.

There are also cascade mitigation strategies based on the game theory. For example, in [38], a stochastic zero-sum game was studied, where an attacker and a defender simultaneously play a game to maximize and minimize LS cost, respectively. In addition, authors in [39] proposed a Q-learning-based algorithm to mitigate

CFs in power grids and showed the convergence of the game to a Nash equilibrium. Similarly, in [40], the authors proposed a game-theoretic approach to provide an optimal criterion to design an industrial control system, which is resilient to CFs for a special form of a quadratic cost function. An advantage of the MDP model proposed in this article is that since we have estimated the transition probabilities of the MDP, the model can compute the policy rather than learning the policy through Q-learning.

Our group further generalized the MC model developed earlier in [20] to include the impact of followings on CFs: communication-network failures [8], human-operator errors [41], cyber threats [42] as well as community structures of grid networks [43]. However, although the above-mentioned MC model is shown to be effective in analyzing the cascade dynamics, previous work did not study the role of cascade mitigation actions. In addition, a recent work has formulated CFs using MDP and developed an online search method using an MDP-based Q-learning approach [44]. Another line of works is based on identifying the "risky fault chain" that characterizes a CF process [32], [45] and weakening these links to reduce the chance of a cascade occurrence. However, our approach is different as it derives the optimal policy offline for all the possible states of the grid by considering a cost function that accounts for both the positive and negative implications of LS. Finally, another distinguishing feature of our work is that it incorporates real physical operating characteristics of power grids in the model.

SECTION III. Modeling CFs via MDP

In this section, we first review germane aspects of the MC model for modeling CFs. We then formulate an MDP by incorporating actions and costs to devise optimal LS policies to mitigate CFs.

A. Stochastic Abstract-State Evolution (SASE) Model

The MC based model, termed in [20] as SASE, captures the stochastic dynamics of CFs in power grids using a finite state MC over a reduced state space defined by three key physical variables from the power grid. In particular, the MC is based on three state variables, namely, the number of failed lines, F, maximum capacity of the failed lines, $C^{\max}$, and the binary state variable, I, representing whether a state is absorbing $(I = 1)$ or nonabsorbing $(I = 0)$. Namely, if the power grid is in a nonabsorbing state then the failure cascade will continue. Conversely, the CFs terminate if the power grid is in an absorbing state. As such, the state of a power grid during a CF is represented as an abstract state variable $S = (F, C^{\max}, I)$ and a CF process is modeled as an MC over the state space of the abstract state variables. In addition, the SASE model assumes that the cascade-continue transition occurs as a result of single-LFs in the system. The single-failure-per-transition approximation is based upon the assumption that time is divided into sufficiently small intervals such that each interval can allow only a single failure event [20]. A detailed description of how the transition probabilities of the MC are computed is provided in Appendix A.

Note that this reduced state space, as utilized in many of our earlier and recent works [8], [20], [41], [42] as well as recent work by other groups [46], can represent a grid at a high level albeit providing significant computation efficiency and tractability. At the same time, since the transition probabilities of the high-level state variables (i.e., reduced state space) of the MC are learned from the detailed dc power flow applied to detailed simulations of CFs with knowledge of specific line outages and the load shed at all buses, they implicitly capture the role of the omitted variables that are eliminated in the state-space reduction. It is also important to keep in mind that despite the state-space reduction, the MC model is very effective in capturing historical trends of real-world CFs (see [20, Figs. 2, 14, and 15]). In what follows, we adopt the MC of the SASE model as the underlying MC of the MDP.

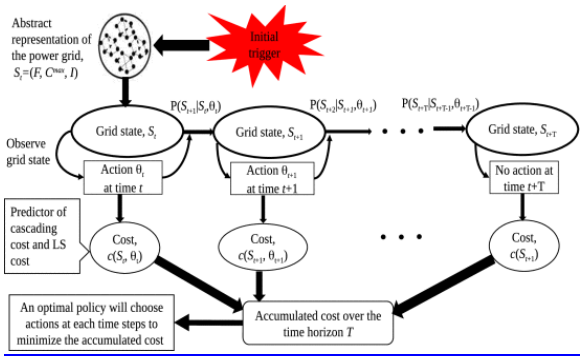

Fig. 1. Framework for the development of optimal LS policies using the MDP for mitigating CFs in power grids. Here, $S_t = (F_t, C_t^{\max}, I_t)$ represents the state of the power grid. The probability $\mathbf{P}(S_{t+1}|S_t, \theta_t)$ models the state-transition probability of transitioning form state S_t to state S_{t+1} due to a chosen action θ_t. Moreover, the cost of taking an action at state S_t is denoted by $c(S_t, \theta_t)$. An optimal policy minimizes the total accumulated cost over a predefined time horizon, T.

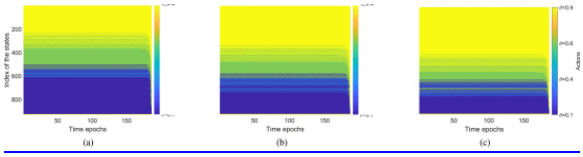

Fig. 2. Optimal MDP policies for different values of μ. The policy represents the actions θ for each state at each time epoch. Specifically, the vertical axis represents the index of the states of the MDP, and the horizontal axis represents time epochs at which actions are taken. Color represents the value of the optimal action (color legend on the right). By fixing a state on the vertical axis the horizontal axis shows all the optimal actions for all the time epochs. (a) Optimal MDP policy with $\mu = 0$. (b) Optimal MDP policy with $\mu = 0{:}5$. (c) Optimal MDP policy with $\mu = 1$.

B. MDP for the Mitigation of CFs

A finite-state MDP [47] can be defined using a five-tuple vector $\langle \mathcal{S}, \Theta, \mathcal{P}, \mathcal{R}, \gamma \rangle$, where $\mathcal{S}$ is the finite state-space of the MDP, Θ is the finite action-space such that a deterministic action at time t, θ_t, is defined as $\theta_t{:}\mathcal{S} \rightarrow \Theta$. $\mathcal{P}$ is a set of probability transition matrix of the MDP, namely, $\mathrm{P}_t(s_{t+1}|s_t, \theta_t) \in \mathcal{P}$ denotes the probability of transitioning to state s_{t+1} at time $t+1$ from state s_t at time t upon taking an action at at time t, and $\mathcal{R}$ represents the set of costs. An element of $\mathcal{R}$ is $c_t(s_t, \theta_t) \in \mathbb{R}$, which denotes the costs at time t for choosing an action θ_t while in state st. Finally, $\gamma \in [0,1]$ is a discount factor. The purpose of a finite-horizon MDP is to find optimal actions at each time instant for each state to minimize a cumulative cost over a time horizon, which is termed as *a policy*.

In this article, we formulate an MDP for developing optimal LS policy to mitigate CFs in the power grid by incorporating LS actions, costs associated with LS actions, and a discount factor with the underlying MC of the SASE model. Fig. 1 illustrates the framework of the development of LS policies using the MDP for mitigating CFs. We represent the power grid with an abstract state space denoted by $S_t = (F_t, C_t^{\max}, I_t)$. LS actions θ_t are considered in response to a CF that is instigated by an initial trigger. Specifically, the LS actions affect the dynamics of CFs, and the grid transitions to a state S_{t+1} with probability $\mathrm{P}(S_{t+1}|S_t, \theta_t)$. Note that the time for the generation of an LS action and the associated delay in implementing the LS actions is shorter than the time

interval between two failure events. This is because the LS actions are precomputed (offline) based on each possible state of the power grid and the associated implementation delay is assumed to be smaller than the time interval between two failure events. In addition, each LS action is associated with a cost. Here, we associate costs to predict the number of LFs and the amount of LS. In the following, we describe each element of the MDP.

1) State Space $\mathcal{S}$

In the MC of the SASE model, $\mathcal{S}$ is a finite collection of all $(F, C^{\max}, I)$. Let there be N transmission lines in the grid, and let the capacities of these transmission lines be quantized into a finite set of capacity values $\mathcal{C} = \{C_1, C_2, \dots, C_K\}$, with $|\mathcal{C}| = K$, where $|\cdot|$ denotes the cardinality of a set. Thus, since I is binary, the number of states of the MDP is $N_s = 2KN$. Note that tracking the number of LFs instead of the operational status (e.g., failed or functional) of each individual line reduces the complexity of the MDP model associated with tracking the dynamics of CFs from exponential to linear in N [20]. Even though our model does not track the sequence of failed lines, such detailed state information is indirectly captured by our MC model. Specifically, the transition probabilities of the MC are identified through learning from extensive physics-based and full-scale (including failure sequences) power-flow simulations of CFs on the IEEE 118, 300 grids.

2) Action Space Θ

We incorporate time-varying and state-dependent actions in the MDP, denoted by $\theta_t(s) \in \Theta$ for all $s \in \mathcal{S}$, which model interventions for preventing the propagation of CFs. Here, different from our previous work [20], we replace the static LS parameter θ with a time-varying and state dependent action function $\theta_t(s)$. In practice, these actions represent alteration of some parameters in the power grid, which can be implemented by sending control signals through the associated communication network, e.g., SCADA system. LS can be implemented by the grid operators to maintain a stable operation of the grid in the presence of an imbalance between power generation and demand [6], [12]. Since the grid operators have access to limited resources in practice, it is reasonable to assume a finite set of actions. For capturing a finite action space, we quantize θ in N_θ discrete values, denoted by $\theta^1, \theta^2, \dots, \theta^{N_\theta}$. Here, $|\Theta| = N_\theta$, and the action set can be expressed as $\Theta = \{\theta^1, \theta^2, \dots, \theta^{N_\theta}\}$. A policy, denoted by π, is a sequence of actions for all time instants $t = 1,2, \dots, T-1$, i.e., $\pi = (\theta_1, \theta_2, \dots, \theta_{T-1})$, where T is the time step when the cascade stops.

Note that a load-dispatch in a power-system during CFs is best represented by the fixed LS policy. This is because at the time of contingencies the operators implement LS following standard procedure to avoid line tripping. In the simulations described in the following Sections IV and V, θ is chosen from the set $\Theta = \{0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.9\}$, which allows consideration of power grids with various LS capabilities. For example, the lowest value $\theta_{\min} = 0.1$ indicates that not more than 90% of the loads are shedable. Such an assumption is realistic since all grids have some critical loads (such as emergency services like hospitals), for which LS is not applicable under any circumstances. On the other hand, the maximum value of θ is $\theta_{\max} = 0.9$, which indicates 10% of the loads are always shedable, as it is unrealistic to consider a power grid for which LS cannot be implemented at all. Note that the amount of LS is a function of θ and is done automatically by the optimal dc power flow. Specifically, $(1-\theta)$ is the maximum fraction of the load that can be shed uniformly from each load bus. However, the actual amount of load-shed on each bus is determined by the optimal dc power flow and is always less than or equal to $(1-\theta)$ (see Fig. 7 for details). Note that this type of strategy is more realistic than manual LS for the following two important reasons: 1) it captures the realistic power flow dynamics in power grids, which is important when dealing with a CF and 2) it implements the minimum LS from load buses obtained as the solution of the MDP. In practice, θ (more precisely, $1-\theta$) can be interpreted by the grid operators as a maximum amount of LS, as it informs on how much LS can be implemented at each load bus of the grid during a CF.

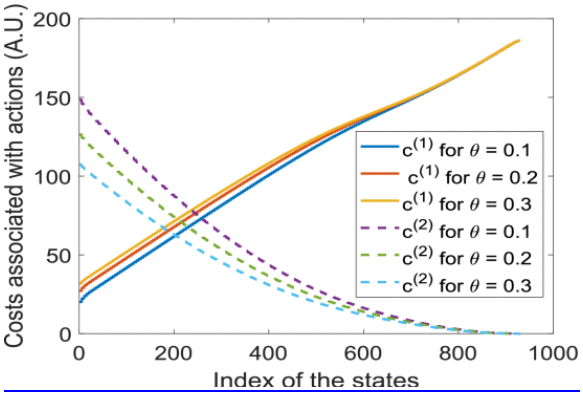

Fig. 3. Interplay between the LF cost $c^{(1)}$, and the LS cost, $c^{(2)}$, for various values of $\theta_t = \theta$. (A.U.=arbitrary unit.) The number of transmission LFs increases with the state index. Hence, $c^{(1)}$ ($c^{(2)}$) increases (decreases) with the state index.

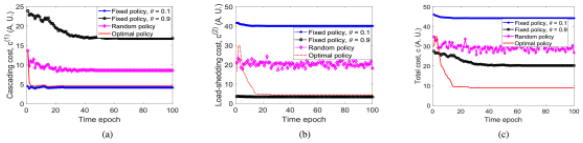

Fig. 4. Average cost given by the optimal policy and other policies. Note from Fig. (c) that the total cost is significantly reduced by the optimal policy.

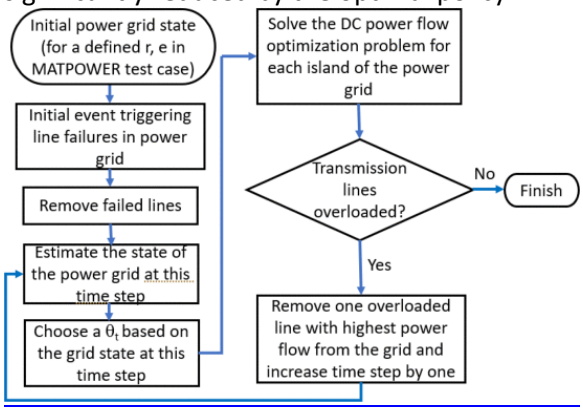

Fig. 5. Flowchart of the CF simulation by incorporating LS actions at each time iteration of the dc power flow.

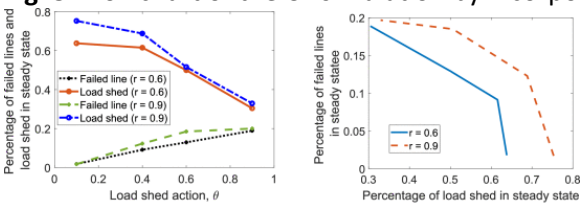

Fig. 6. Validation of the analytical LF and LS costs from CFs simulation using dc power flow. (a) Increase (decrease) of LFs (LF) with LS actions, (b) LFs decrease with the LS.

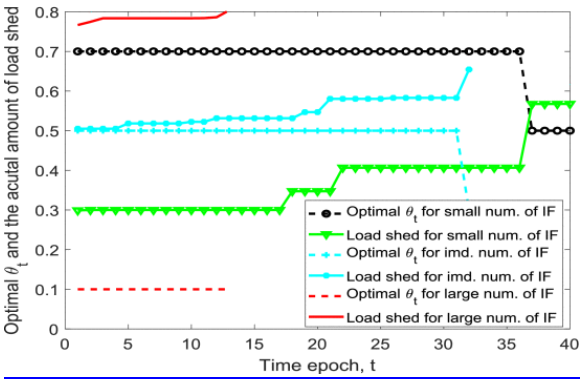

Fig. 7. Optimal LS actions and corresponding actual amount of load shed (in percentage) for the three cases of a small, intermediate, and large number of initial failures. Here, num. := number, imd. := intermediate, IF := initial failures.

3) Transition Probability P

Following [20], we assume that the transition probability matrix is time invariant, i.e., $p_{ss'}^{\theta} = \mathrm{P}\{S_{t+1} = s' | S_t = s, \theta_t = \theta\}$, where $\theta \in \Theta$. Hence, the transition from the current state to the next state depends on the current state and on the current action. Accordingly, for different values of θ, we obtain different transition matrices. In order to generate the transition matrix for each choice of θ, we used 1000 scenarios of CFs, starting from random initial disturbances with two or three random LFs while simulating the CFs using dc OPF [20]. Based on these large, simulated datasets, analytical formulation of the transition matrix of the MC was generated for any given θ using the technique reported in [20]. In the end, we obtain nine MC models, one for each θ, which are subsequently used for the analytical formulation of the MDP. For example, for each state of the MDP, defined by a failure event, we have nine possibilities for the transition matrix, which can be chosen from.

4) Cost Function c

We assign a cost to all actions, $\theta_t \in \Theta$, taken by the grid operators at state $S_t \in \mathcal{S}$, which is denoted by $c(S_t, \theta_t)$. The total cost associated with a policy $\boldsymbol{\pi} = (\theta_1, \theta_2, \dots, \theta_{T-1})$ over a time horizon T is defined as $R_T^{\boldsymbol{\pi}} = \sum_{t=1}^{T-1} c(S_t, \theta_t) + c(S_T)$, where $c(S_T)$ is the terminal cost that only depends on the state of the grid at the last time T. We assume this latter cost to be equal to the number of failed lines, F_{S_T}, at state S_T, i.e., c(ST)=FST, $S_t \in \mathcal{S}$. With a discount factor, γ, the expected total discounted cost (ETDC), $\overline{R}_{T,\gamma}^{\boldsymbol{\pi}}$, for a policy π, is defined as [47]

$$\begin{aligned} \overline{R}_{T,\gamma}^{\boldsymbol{\pi}} &= \mathrm{E}_{\boldsymbol{\pi}}\left[\sum_{t=1}^{T} \gamma^{t-1} c(S_t, \theta_t)\right] \\ &= \mathrm{E}_{S_1}\left[\mathrm{E}_{\boldsymbol{\pi}}[\sum_{t=1}^{T} \gamma^{t-1} c(S_t, \theta_t) | S_1]\right] \\ &= \sum_{s \in S} v_{T,\gamma}^{\boldsymbol{\pi}}(s) \mathrm{P}(S_1 = s) \end{aligned}$$

where $\mathrm{E}_{\boldsymbol{\pi}}$ and E_{S_1} are the expectations conditional on a policy π and initial state S_1, respectively. Furthermore, $v_{T,\gamma}^{\boldsymbol{\pi}}(s) = \mathrm{E}_{\boldsymbol{\pi}}\left[\sum_{t=1}^{T} \gamma^{t-1} c(S_t, \theta_t) | S_1 = s\right]$ is the state-value function of the policy π when there

are T states to traverse starting from state s. Clearly, for a given $S_1 = s$, minimizing the ETDC is equivalent to minimizing $v_{T,\gamma}^{\boldsymbol{\pi}}(s)$ [47]. For a given initial state $S_1 = s \in \mathcal{S}$, the objective is to find an optimal policy

$$\boldsymbol{\pi}^*(s) \quad = \arg\min_{\boldsymbol{\pi}} v_{T,\gamma}^{\boldsymbol{\pi}}(s).$$

(1)

We can express $v_{T,\gamma}^{\boldsymbol{\pi}}(s)$ as follows:

$$\begin{aligned} & v_{T,\gamma}^{\boldsymbol{\pi}}(s) \\ & = \mathrm{E}_{\boldsymbol{\pi}}[c(s,\theta)] + \gamma \mathrm{E}_{\boldsymbol{\pi}}\left[\sum_{t=2}^{T} \gamma^{t-2} c(S_t,\theta_t) | (S_1 = s, \theta_1 = \theta)\right] \\ & = \mathrm{E}_{\boldsymbol{\pi}}[c(s,\theta)] + \gamma \sum_{s \in S} v_{T-1,\gamma}^{\boldsymbol{\pi}}(s) \mathrm{P}\big(S_2 = s \big| (S_1 = s, \theta_1 = \theta)\big) \end{aligned}$$

(2)

where $v_{T-1,\gamma}^{\boldsymbol{\pi}}(s) = \mathrm{E}_{\boldsymbol{\pi}}\left[\sum_{t=2}^{T} \gamma^{t-2} c(S_t,\theta_t) | S_2 = s\right]$ is the state-value function of the policy π when there are $T-1$ states to traverse starting from state s. Note that $v_{0,\gamma}^{\boldsymbol{\pi}}(s) = F_s, s \in \mathcal{S}$, as defined by the terminal cost $c(S_T)$ above. Hence, it can be seen from (2) that the backward induction algorithm [47] can be used to solve (1) to find the optimal policy, which yields the optimal action at every time step for any states of the MDP.

SECTION IV. Cost Function for CFs

We define the primary objective of an optimal LS policy to be preventing subsequent LFs during a cascade. However, implementing LS will generate a cost in terms of shedding loads (e.g., industrial, residential customers) from the grid. Accordingly, we propose a cost function c that balances between the cost of transmission-LFs (i.e., cost of LFs), $c^{(1)}$, and the cost associated with LS, $c^{(2)}$. Specifically, we define the cost function $c(S_t,\theta_t)$ corresponding to an action $\theta_t \in \Theta$ at time t while the grid is in state $S_t \in \mathcal{S}$, as a linear combination of $c^{(1)}(s_t,\theta_t)$and $c^{(2)}(s_t,\theta_t)$

$$c(s_t,\theta_t) = w\alpha_1 c^{(1)}(s_t,\theta_t) + (1-w)\alpha_2 c^{(2)}(s_t,\theta_t)$$

(3)

where $w \in [0,1]$. The cost function (3) is designed to control the tradeoff between LS and the expected LFs: the weight parameter, [0, 1], is responsible for adjusting this tradeoff. Specifically, if we increase the LS then the expected number of LF decreases, since shedding loads from the grid reduces overloading of the lines, hence, reducing transmission-LFs. Conversely, if we decrease LS then the expected LFs (due to cascade propagation) will be higher in the steady state. By tuning w, it is possible to assign a larger weight to a term with respect to another one based on contingencies. In what follows, we give equal weight to both terms in (3) and assign $w = 0.5$. In practice, two constants α_1 and α_2 convert the two costs (with different units) into one of the same cost unit. The values of α_1 and α_2 can be estimated from the real costs of transmission-LFs and LS, respectively. Denoting $\beta_1 = w\alpha_1$ and $\beta_2 = (1-w)\alpha_2$ we rewrite (3)

$$c(s_t,\theta_t) = \beta_1 c^{(1)}(s_t,\theta_t) + \beta_2 c^{(2)}(s_t,\theta_t).$$

(4)

Two special cases of (4) are as follows: 1) $c^{(1)}(s_t, \theta_t) = p_{\text{stop}}(s_t, \theta_t)$ and $c^{(2)}(s_t, \theta_t) = 0$, for which maximizing the objective function of the MDP results in a stationary policy $\theta_t = \theta_{\text{min}} = 0.1$, and 2) $c^{(1)}(s_t, \theta_t) = \text{E}[R(s_t, \theta_t)]$ and $c^{(2)}(s_t, \theta_t) = 0$, where $\text{E}[R(s_t, \theta_t)] = \sum_{j=1}^{N_s} \lim_{t\to\infty} \text{P}_{sj}^{\theta}(t)\left(F_j - F_s\right)$ is the expected number of additional failed lines in the steady state of the MC starting from the state $s_t = s$ and taking action $\theta_t = \theta$. Also in this case minimizing the objective function of the MDP results in a stationary policy $\theta_t = \theta_{\text{min}} = 0.1$, which corresponds to aggressive LS. Now we describe some other choices β_1 and β_2 to illustrate how they influence the LS actions using as example the IEEE 118 bus system. Since $\beta_1 = w\alpha_1$ and $\beta_2 = (1-w)\alpha_2$, the values of β_1 and β_2 are automatically determined by the values of w, α_1, and α_2. In this article, $w = 0.5$, and α_1, α_2 indicate the importance of both LFs and LS (on the surviving lines) costs. In practice, the importance of these two terms can be different for two grid operational energy management systems. Consequently, the choice of these parameters largely depends on the tradeoff between the need for stopping a CF and the cost of LS and serving the maximum loads at the risk of having large CFs. Intuitively, for a very large β_1 the optimal policy is dominated by first term and vice versa when β_2 is very large. Table I shows some choices of β_1 and β_2 to illustrate how they influence the LS actions (details of the simulation setup are described in Section V). This is evaluated by a new metric, which shows the percentage of different LS actions, i.e., $\theta_t = 0.1, 0.4, 0.6, 0.9$, executed by the grid for all time epochs at all states of the IEEE 118 bus system. Notice the prevalence of extreme LS action, i.e., $\theta_t = 0.1$ when β_1 is large. In contrast, minimum LS is applied, i.e., θt = 0.9 when β_2 is equal to 1. Note also that unlike $\theta_t = 0.1$, the percentage of minimum LS action, $\theta_t = 0.9$, is never zero. This is because for large time epochs, for which the CFs stop due to either no additional LFs or due to the absence of surviving lines in the grid, minimal LS is the optimal action as this minimizes the LS cost.

TABLE I Percentage of Different LS Actions Executed by the Grid for All Time Epochs at All States of the IEEE 118 Bus System

β_1	β_2	$\theta_t = 0.1$	$\theta_t = 0.4$	$\theta_t = 0.6$	$\theta_t = 0.9$
0	1	0%	0%	0%	100%
0.2	0.8	2.11%	0.32%	1.68 %	95.89 %
0.5	0.5	39.71%	2.99%	23.22%	34.07 %
0.8	0.2	75.93%	3.21 %	20.32%	0.54%
1	0	99.57%	0%	0%	0.43%

The two special cost functions described earlier both set the cost of LS to zero. In practice, implementing LS has an associated cost for not serving loads (e.g., customers). However, since no state variables are assigned in the MC to track LS, we cannot track the amount of load shed directly. Although it is possible in principle to track such quantity, incorporating and tracking such quantity would make the state space of the embedded MC of the MDP model very large. As a result, the computational complexity of the MDP model would increase drastically. Instead, we propose a novel approach to quantify the cost of LS, which measures the LS cost in terms of the fraction of surviving lines at the steady state. This assumption is based on the following rationale: if there are no surviving lines in the grid then all the loads will be shed as the grid will not be able to serve loads without lines. Most importantly, we validate this rationale in Section V-C using load shed data from CF simulation on the IEEE 118 and IEEE 300 bus systems. Accordingly, we define the following cost function, which incorporates the associated cost of LS, for all $S_t \in \mathcal{S}$:

$$c(s_t,\theta_t) = \beta_1 \mathrm{E}\left[F_{s_t,\mathrm{end}}^{\theta_t}\right] + \beta_2(1-\theta_t)\left(1-\frac{\mathrm{E}\left[F_{s_t,\mathrm{end}}^{\theta_t}\right]}{N}\right)\left(N-F_{s_t}\right) \quad (5)$$

where $\mathrm{E}\left[F_{s_t=s,\mathrm{end}}^{\theta_t=\theta}\right] = \sum_{j=1}^{N_s} \lim_{t\to\infty} \mathrm{P}_{sj}^{\theta}(t)F_j$ is the expected number of failed lines in the steady state (denoted by *end* to signify the ending of a cascade) for choosing an action $\theta_t = \theta$ at state $s_t = s$ at time t and F_{S_T} is the number of failed lines at state st. This corresponds to setting in (4), $c^{(1)}(s_t,\theta_t) = \mathrm{E}\left[F_{s_t,\mathrm{end}}^{\theta_t}\right]$ and $c^{(2)}(s_t,\theta_t) = (1-\theta_t)\left(N-F_{s_t}\right)\left(1-\mathrm{E}\left[F_{s_t,\mathrm{end}}^{\theta}\right]/N\right)$. In the following, we provide a detailed description of each term appearing on the right-hand side of (5).

The first term accounts for the LFs cost, $c^{(1)}(s_t,\theta_t)$, which is the expected number of failed lines in the steady state. It is important to note that $\mathrm{E}\left[F_{s_t=s,\mathrm{end}}^{\theta_t=\theta}\right]$ is a function of the state s, which governs the capacity of the failed lines. Hence, two states with same number of initial failures but with different capacities will have different expected LFs, which captures the role of line capacities in the cost function. The second term represents the LS cost, $c^{(2)}(s_t,\theta_t)$, where (N−Fst) represents the number of survived lines at state s_t and $\left(1-\mathrm{E}\left[F_{s_t,\mathrm{end}}^{\theta_t}\right]/N\right)$ is the probability that a line will survive in the steady state due to the chosen action θ_t at state s_t. Hence, the product $\left(1-\mathrm{E}\left[F_{s_t,\mathrm{end}}^{\theta_t}\right]/N\right)\left(N-F_{s_t}\right)$ represents the expected fraction of survived lines in the steady state. The term $(1-\theta_t)$ represents the probability of taking an action corresponding to LS, which captures the cost of LS for different $\theta_t s$. In addition, since the quantity $\mathrm{E}\left[F_{s_t,\mathrm{end}}^{\theta_t}\right]/N$ impacts the LS, the passive cost of LS (i.e., LS not determined by the power system operator but caused by the transmission LF) is also captured in (5). Two extreme cases are possible by looking at (5) when $\theta_t = \theta = 0$ (i.e., full LS), the cost only accounts for shedding loads: $\left(1-\mathrm{E}\left[F_{s_t,\mathrm{end}}^{\theta_t}\right]/N\right)\left(N-F_{s_t}\right)$; when $\theta_t = \theta = 1$ (i.e., no LS), the associated cost of LS is zero. Clearly, for the action space, Θ, defined in Section III, the smallest $\theta_t = \theta_{\min} = 0.1$ results in minimum LFs cost and maximum LS cost; on the other hand, the largest $\theta_t = \theta_{\max} = 0.9$ results in maximum LF cost and minimum LS cost. While c_2 is directly dependent on c_1, c_1 is indirectly dependent on c_2 through the LS amount θ_t as well as the number of failed lines at that time F_s.

Hence, by using (5) in (2) and subsequently in (1), we see that the objective of the optimization (1) is to "minimize the cumulative expected LFs at the end of cascade with minimum possible cumulative expected LS, where cumulation of LFs and LS are taken over each time steps until the cascade stop time (T)." Moreover, since this is a finite horizon MDP, the policy argument, π, of (1) is not dependent on the discount factor, i.e., LS actions are independent of γ. Specifically, the objective function without a discount factor can also be used as a criterion for deriving the optimal policy for a finite-horizon MDP [47]. However, we have kept γ to maintain the generality of our analytical formulation. The discount factor will play an important role for the convergence of the objective function for infinite-horizon MDP, which we will be considered in future extensions of this work.

Finally, it is important to note that the LS can be achieved through small and large controllable loads, such as air conditioners, electric vehicles, refrigerators, and thermostats. In practice, the coefficient $c^{(2)}$ would be underestimated without consideration of the adjustment cost. We have, therefore, incorporated a cost, $c^{(a)}$, in

addition to the actual LS cost. Hence, the adjusted LS cost is $c^{(2)} = c^{(LS)}$ (the actual LS cost) + $c^{(a)}$ (the adjustment cost of shedding different types of loads). We assume the adjustment cost $c^{(a)}$ is a linear function of the amount of load-shed, i.e., $c^{(a)} = \mu \times$ (amount of load-shed), where $\mu = [0, 1]$ is a constant that determines the importance of the adjustment cost. The rationale for this assumption is that a large LS results in customers needing to turn down large loads, such as electric vehicles and air conditioners, as opposed to little adjustments (such as turning off a thermostat) in case of small LS. Fig. 2 shows optimal MDP policies for different values of μ for the IEEE 118 bus system (details of the simulation setup are described in Section V). We see that increasing μ increases the minimum LS region, i.e, the region occupied by the $\theta = 0.9$ action (the yellow-colored region in Fig. 2), due to an increase in the adjustment cost. In particular, for $\mu = 0, 0.5$, and 1 the number of times that a minimum LS action is chosen, i.e., $\theta = 0.9$, relative to the total number of actions in the optimal policy (i.e., $\theta = 0.9, 0.6, 0.4, 0.1$) are 0.34, 0.46, and 0.58, respectively.

SECTION V. Numerical Simulation

We test our cascading mitigation strategy on the IEEE 118 and 300 bus systems, which have 186 lines and 411 lines, respectively. Note that the IEEE 118 bus system provides a simplified approximation for the American (U.S. Midwest) power system. We assume that at each time iteration a single LF occurs. This is consistent with choosing each time interval to coincide with a relevant event, e.g., a transmission LF. Since there are 186 (411) lines in the IEEE 118 (300) bus system and one line may fail at each time step, we consider a time horizon $T = 186\ (411)$. Furthermore, since the transmission lines have finite power flow capacity, we quantize the capacities of the transmission lines into five categories-based upon power flows through the lines when the grid is operating at full generation capacity. The five capacities are chosen from the set $\mathcal{C} = \{20\ MW, 80\ MW, 200\ MW, 500\ MW, 800\ MW\}$ [48]. Hence, the total number of states of the MDP is $2|\mathcal{C}|N = 1860$ and $2|\mathcal{C}|N = 4110$ for IEEE 118 and 300 systems, respectively. To simplify the representation of the states of the MC (e.g., 1860 states and 4110 states for IEEE 118 and 300 grids, respectively), we label each state of the MC, $S_i = (F_i, C_i, I_i)$, with an integer index $i = 2|\mathcal{C}|(F_i - 1) + 2\left(C_i^{\text{index}} - 1\right) + I_i + 1$, where $C_i^{\text{index}} = 1,2, \ldots ,5$, for $C_i = 20,80, \ldots ,800$, respectively. Note that the indices of the grid-states are in order of increasing number of line failures and line-capacities, i.e., if the index is higher, either the number of line failure or the maximum capacity of the failed line is higher. Specifically, for two indices i and j: $i < j$ if (1) $F_i < F_j$ or (2) $F_i = F_j$ and $C_i < C_j$ or (3) $F_i = F_j$ and $C_i = C_j$ and Ii<Ij. As a result, 1860 states of the IEEE 118 grid are denoted by 1,2,…,1860. Moreover, the number of line failures F_i corresponding to a state index i is equal to $\lfloor i/2|\mathcal{C}| \rfloor$ if i is an integer multiple of $2|\mathcal{C}|$, or $F_i = \lfloor i/2|\mathcal{C}| \rfloor + 1$ otherwise. Here, $\lfloor x \rfloor$ is a floor function that returns the largest integer lower than or equal to x.

The MC is used to compute (5) for a given state st and action θ_t, which, in turn, requires the calculation of $\mathrm{E}\left[F_{s_t=s,\text{end}}^{\theta_t=\theta}\right]$. Specifically, this is done by performing a steady-state analysis of the developed MC, which enables us to find the limiting probability, $\lim_{t\to\infty} \mathrm{P}_{sj}^{\theta}(t)$, of reaching any state j given that we started from an initial state s by taking an action θ. Using the limiting probabilities, we can compute the expected number of failed lines at the steady state, i.e., $\mathrm{E}\left[F_{s_t=s,\text{end}}^{\theta_t=\theta}\right]$. Finally, given all the five-tuples of an MDP (described in Section III), the backward induction algorithm can numerically compute the optimal actions at every time step for any states of the finite-state MDP. Note also that we use $\gamma = 0.7$ in the MDP. We have verified that for other values of γ, the results change quantitatively but not qualitatively.

A. Optimal Policy

Fig. 3 shows the two costs $c^{(1)}$ and $c^{(2)}$ with respect to the state of the grid corresponding to three different actions. Note that for a given grid state, $c^{(1)}$ ($c^{(2)}$) decreases (increases) with LS (i.e., as θ_t decreases). This

indicates that the LFs can be reduced by LS. Besides, as the grid-states are in order of increasing LFs, $c^{(1)}$ ($c^{(2)}$) increases (decreases) with the grid-states.

The derived optimal policy is state-dependent and time dependent. Specifically, for the states corresponding to a lower expected number of failed lines (indices: 1–225), the policy always returns a minimum amount of LS (i.e., $\theta_t = \theta_{\max} = 0.9$). This is because the grid is not at risk of a cascade-escalation when the expected number of failed lines in the steady state is small; thus, shedding more load would generate unnecessary LS cost. However, for state indices 225–580, which correspond to larger expected numbers of failed lines, the policy suggests an incremental increase of LS. When the expected number of failed lines is large, the optimal policy consists in aggressive LS, i.e., $\theta_t = \theta_{\min} = 0.1$. This is because, as the grid loses most of its transmission lines, minimizing the cost associated with LFs is more important than minimizing the LS cost. To observe the time dependency, we found two regimes of the optimal policy: a first regime for small times corresponding to a constant optimal θ and a second regime for larger times corresponding to a decaying LS θ. In fact, we found that for large times, for which the CFs stop due to either no additional failed lines or due to the absence of surviving lines in the grid, minimal LS is the optimal action, as this minimizes the LS cost. For a larger power system, such as IEEE 300 bus systems, we found $\theta_t = \theta = 0.9,0.6,0.4,0.1$ for the state indices $1-650, 651-980, 981-1240, 1241-2090$, respectively, before it approaches $\theta = 0.9$ for large times. Clearly, this is slightly scaled version of the optimal policy for the IEEE 118 bus systems, which shows the scalability of the MDP proposed formulation.

In addition, the policy is obtained through an offline implementation of the optimal dc power flow for simulating CFs and collecting the data to calculate the transition matrix of the MDP. Hence, given a grid, the optimal policies are precalculated offline considering all the states a grid may occupy during a CF. The only variable entered for online policy implementation is the observed state (number of failed lines and the maximum capacity of the failed lines) during the progression of a cascade. As a result, after observing the state during a CF, it is straightforward to implement the actions corresponding to the optimal policy, which was pre-calculated offline. The computational complexity of the MDP is as follows. We have implemented the backward induction algorithm (see[47, p. 92]) to solve the MDP. Since the MDP is a finite-time horizon ($T = 186$ steps and 411 steps for the IEEE 118 and IEEE 300 bus systems, respectively) and since the state and action space are finite and discrete ($2NK$ states and N_θ actions, where N is the number of failed lines and K is the number of capacities), the backward induction computation requires $(T-1)(2NK)^2$ multiplications to compute each of the $N_\theta^{2NK(T-1)}$ policies [47]. Hence, for each policy, the computational complexity is linear in time and quadratic in the number of state variables. As a result, as the state and/or action space becomes larger, the computational complexity scales with the cardinality of state-action space as well as with the time-horizon. For example, we have computed the policy in 0.0254 min, 1.0589 min, and 11.6705 min for the IEEE 39, IEEE 118, and IEEE 300 bus systems, respectively, using a simple 4.0 GHz Intel Corei7 CPU (without parallel processing), which can be reduced by 8 times (i.e., 0.1323 min only for IEEE 118 grid) if the computation is distributed among 8 cores. In addition, we anticipate that the central controller of an actual power system will have large computing machines to handle large computations. Nonetheless, we found that the computation time for computing real-time actions is reasonably comparable to the historical data for the evolution time of CFs [28], [49].

B. Minimization of the Total CF Cost by the Optimal Policy

In this section, we explore the performance of our optimal policy in minimizing the cost function (5). Fig. 4(a)–(c) show the cost of LFs $c^{(1)}$, the LS cost $c^{(2)}$, and the total cost $c^{(1)} + c^{(2)}$ associated with a CF, respectively. We average the results over 1000 numerical experiments each one of which corresponds to a random number of initial failures. We see from Fig. 4(a) that a fixed policy with $\theta_t = \theta = 0.1$ gives minimum LF cost and maximum LS cost. On the other hand, we see from Fig. 4(b) that a fixed policy with $\theta_t = \theta = 0.9$ gives minimum LS cost

but maximum LF cost. Fig. 4(c) shows the total CF cost obtained by adding the LF cost and LS cost. We see that the total CF cost is significantly reduced by the optimal LS policy. Hence, the developed optimal LS policy can be used to mitigate CFs while using an optimal amount of LS. Similar patterns for all the three costs for all the policies were obtained for the IEEE 300 bus system (not shown). Also for this case, we saw that the optimal policy resulted in an improved performance with respect to all other policies (e.g., fixed and random policies).

C. Validation of the Optimal Policy

In this section, we validate our policy with extensive physics-based simulations of CFs, which incorporate physical power flow of realistic electric grids. Following many state-of-the-art stochastic models of CFs [4], [22], [24], [31], [32], [50], dc power flow models (see Appendix VI) have been used to simulate CFs due to their simplicity yet effectiveness in capturing realistic power flow redistribution during CFs in power grids. Note that even if we implement optimal power flow to distribute the power flow, the protection relay (circuit breaker or impedance protection relay) may still trip a line when the power flow through the line reaches a certain threshold of the actual transmission capacity of the line [20], [22]. This step is done as a safety measure. The threshold, which affects the actual capacity of a line, depends on various factors and mechanisms, such as smaller measure impedance by relay due to line overloading, ambient temperature, communication/control system problems [20]. Hence, even with the execution of optimal power-flow, line tripping may occur, which may, in turn, overload other lines and, hence, cause a cascade. In addition, the optimal dc power flow has no capability to predict the line failures at the steady state. Hence, it cannot be used to minimize the steady-state line failure through LS, which is considered here.

Note that the use of the ac power flow captures the CF phenomenon better than the dc power flow model. However, the ac power-flow model is nonlinear in the power flow optimization, as it includes the reactive power, variations of voltage magnitude, power losses, etc. Hence, the nonlinear ac power flow equations need to be solved iteratively, which increases the computational complexity compared to the linear and much faster dc power flow model OPF [51]. Second, due to their iterative nature, the ac power flow equations face divergence problems in their iterations, and they may not have a solution in certain cases in which the dc power-flow optimization does have a solution [51].

The flowchart in Fig. 5 shows the simulation steps for the CF with real power grid test cases. Various grid test cases with realistic physical parameters can be found in MATPOWER [52]. Here, time steps of the CF are assumed to be discrete, and to be consistent with the assumption of the MDP formulation that one line failure occurs at each time step. The simulation starts with an initial event at the first time step that fails some lines, which are chosen randomly to capture the stochastic nature of the initial event. After the initial failure, we estimate the state of the power grid at the first time step and choose an action. The optimal action, θ_t, depends on the state index at a given time. The value of LS capability, $1-\theta_t$, is implemented in the dc power flow, which redistributes power flow through all lines in every possible island of the grid after balancing power generation and demand. A line failure occurs if a line is overloaded. Note that the threshold (C_l) of the actual transmission line is not by itself the upper/lower bound of the line capacity in the optimal dc power flow calculation in Appendix B. Specifically, the lth is defined to be overloaded if the power flow through the lth line, f, satisfies the following condition: if $|f_l| \geq (1-e)C_l$, where e measures the line-capacity (C_l) estimation error by the control center (defined in Appendix VI). If there are multiple line failures, we choose a line that carries a relatively high-power flow and increase the time step by one. In addition, a small probability (0.01) of failure for neighboring lines following a line failure is incorporated in the dc power flow model to capture the probability that a protection relay may not work properly. Note that such mis-operations of the protection relays may also contribute to the occurrence of major blackouts [20]. Following a line failure, power redistribution occurs based on the optimal dc power flow. If there are additional lines overloaded at this time step, then the CF process continues and an action is taken at each time step until there are no more line-overloadings. Note that at each

time step the value of θ_t changes based on the state of the power grid (i.e., based on number of line failures and maximum capacity of the failed lines), which is determined by the MDP optimal policy devised in Section V-A.

The flowchart in Fig. 5 can also be used to describe how the derived policy can be used in practice. Specifically, the left-side blocks of the flowchart show how an operator can use the derived policy in the event of CFs. For example, operators usually know the state of the grid during normal operation, which provides the model parameters, r and e. Then, given an initial event, an operator would take the derived policy from the analytical model for that grid parameters and choose the action for that grid state from the derived policy. However, we are aware that the estimation of the transition probabilities can be a challenge to obtain, due to shortage and quality of CF data from a grid. Hence, in future extensions of this work, we will be using a Q-learning algorithm to learn the approximate policy instead of deriving the optimal policy. However, the proposed work will serve as a baseline to measure the performances of Q-learning; the latter is beyond the scope of this work.

1) Validation Results

Recall that the first term on the right-hand side of (5) represents the cost associated with LFs, which is supposed to increase as the LS action value θ_t increases. The second term on the right-hand side of (5) represents LS cost in terms of survived lines, which is supposed to decrease as θ_t increases. We begin by verifying the trend of analytical results for the first and second terms in (5) using the data of LFs and load shed from CFs simulation of the IEEE 118 grid. Fig. 6(a) shows that as we increase θ_t in the optimal dc power flow optimization, the percentage of LS decreases and the percentage of line failure increases. This observation provides the rationale for increasing θ_t in the analytical model to minimize LFs. Most importantly, validation of LS is crucial since in the analytical MDP formulation there are no state variables in the MDP to track the amount of load shed during a CF. That is the reason why we have used a fraction of the number of surviving lines in the steady state to account for the LS costs [i.e., the second term on the right-hand side of (5)]. From simulation, we find the amount of load shed, which shows a similar behavior with respect θ_t as the LS cost, $c^{(2)}$, of the analytical model. Moreover, Fig. 6(b) shows the impact of LS on minimizing LFs. Notice that increasing the amount of load shed corresponds to reducing LFs. This provides evidence that the two terms of the cost function (5) resemble the physical simulation qualitatively.

Next, we validate the results of the optimal policy that minimizes the total CF cost in terms of LFs and LS. We initiate the simulation of CFs by randomly failing some lines of different capacities from the IEEE 118 bus system. Table II shows the percentage of LFs and LS for different fixed values of θ_t, optimal policy (OP) and random policy (Rnd). The three subtables are for three different fractions of initial LF, i.e., small = 0.02, intermediate = 0.27, and large = 0.54. This makes our presentation of initial failures consistent with other works on CFs [7], [24]. Here, for small and large initial failures, the results given by the optimal policy correspond to fixed polcies $\theta_t = \theta = 0.9$ and $\theta_t = \theta = 0.1$, respectively. For intermediate numbers of initial failures, the costs given by the optimal policy fluctuate between $\theta_t = \theta = 0.4$ and $\theta_t = \theta = 0.6$. These observations are consistent with the optimal policy discussed in Section V-A.

TABLE II Percentage of Failed Line (FL) and Load Shed (LS) at Steady state for Different Policies With Small, Intermediate (Imd.), and Large Initial Failures

θ_t	small initial failure		Imd. initial failure		large initial failure	
	FL	LS	FL	LS	FL	LS
0.1	2.1%	76%	31 %	77%	58%	80%
0.4	12%	69%	39%	68%	61%	72%
0.6	18%	51%	42%	58%	62%	66%
0.9	19%	36%	44%	61%	61%	72%
OP	19 %	36%	41 %	60%	58%	80%

Rnd	7%	75%	36%	78%	61%	77%

OP = optimal policy, Rnd = random.

D. Examples of the Optimal LS Actions

For the implementation of LS, we apply a uniform LS on all buses as done [12]. In other words, after identifying the optimal LS actions, θ, we rely on the optimal dc power flow for its implementation. We set $\Theta = \{0.1, 0.3, 0.5, 0.7\}$ in order to show the effectiveness of our model for different values of LS. In addition, we choose a large value of the parameter r corresponding to a high level of stress for the grid, i.e., $r = 0.9$. This value corresponds to a situation in which the grid uses 90% of generation power to serve loads. Fig. 7 shows three instances of actions and LS corresponding to three different instances of initial failures. Notice that with 5 initial failures (few initial failures), $\theta_t = 0.7$ is the preferred LS action, i.e., the optimal policy corresponds to minimum LS. However, when the CF evolves to a larger number of LFs, θ_t decreases to 0.5, which corresponds to application of a larger LS to stop the cascade. In the case of 50 initial failures, $\theta_t = 0.5$ (intermediate LS) is the optimal action initially followed by $\theta_t = 0.3$. Finally, with 100 initial failures $\theta_t = 0.1$ (aggressive LS) is the optimal action. Note that the obtained optimal LS policies are in agreement with the proposed MDP model. We also see from the horizontal axis that the stopping time of the cascade decreases with increasing amount of LS. This implies that aggressive LS actions can stop the cascade sooner at the price of a larger amount LS. Furthermore, from Fig. 7, we see that the amount of actual load-shed increases with time, even when the optimal LS parameter, θ_t, is constant for some time epochs. This is because the optimal dc power flow optimizer performs LS incrementally to minimize the number of LFs.

It is important to note that a direct comparison of the results between the proposed work and related works [27]–[30] could not be presented in this article for the following reasons: 1) our MDP framework is built on the reduced state-space MC (which makes it scalable) and to the best of authors' knowledge the frameworks and formulations proposed in other works [27]–[30] are fundamentally different from our proposed formulations; 2) the model parameters are also different (e.g., LS parameters and evaluation metrics). Therefore, we took the general approach (see, e.g., [47], [53, Figs. 6 and 7]) of comparing the derived policy (based on our MDP approach) with random and fixed policies, and we have shown that the derived MDP policies always outperform the two alternatives. In addition, we performed a comparison between our analytical model and the dc power-flow model (see Table II and Fig. 7). This comparison is highly revealing because it shows an agreement between our analytical model and the CF model based on dc power-flow. In particular, this comparison demonstrates the success of our analytical model to predict the behavior seen from physics-based simulations of CFs, which gives our abstract model a practical quality.

SECTION VI. Conclusion

This article presents a novel theoretic framework based on an MDP to devise optimal LS policies in order to mitigate CFs in power grids. Specifically, we have judiciously incorporated state-dependent and time-varying actions into the MDP by introducing a realistic cost function that considers both positive and negative implications of a chosen action. We computed the optimal policies for the IEEE 118 and IEEE 300 bus systems, which was found to be very effective in reducing the size of CF but largely dependent on the LS costs. Numerical simulations have shown a significant advantage of our optimal policy compared to other policies in terms of minimizing the total CF cost associated with LFs as well as LS. An implementation of the policy on the IEEE 118 system-based on optimal dc power flow has shown excellent agreement with the MDP predictions. Notably, the proposed policy outperforms fixed and random policies in reducing the size of the CFs. This indicates an advantage of adopting the MDP derived optimal policy in mitigating CFs while using minimal LS. The developed optimal policy can be used as a baseline to support LS decisions by grid operators in order to mitigate the spread of CFs in real time. In addition, since the policy can be developed offline, it can also be used in designing a

cascade-preventive grid, which, upon incorporation, can control the CFs by taking LS decisions based on the state of the grid during a CF.

Appendix A Transition Probabilities of the MC

Following [20], [41], the transition probabilities of the MC were chosen to be dependent on the following three parameters:

1. $r \in [0,1]$, the ratio of power demand to power generation;
2. $e \in [0,0.5]$, the power flow estimation error through a transmission line;
3. $\theta \in [0,1]$, the LS parameter that determines the capability of implementing LS when needed.

A key component in the transition probability matrix is the state-dependent cascading-stop probability, $P_{\text{stop}}(S_i)$, i.e., the probability that the binary state variable $I = 1$ at state S_i. Mathematically, $P_{\text{stop}}(S_i)$ is formulated as (see (7) [20])

$$P_{\text{stop}}(S_i) = wP_{\text{stop}}^{(1)}(F_i) + (1-w)P_{\text{stop}}^{(2)}(C_i^{\max})$$

(6)

where $P_{\text{stop}}^{(1)}(F_i)$ is the cascade-stop probability given F_i, the number of line failures, and $P_{\text{stop}}^{(2)}(C_i^{\max})$ is the cascade-stop probability with the maximum capacity of failed lines being Cmaxi and $w = 0.5$. Note that $P_{\text{stop}}^{(1)}(F_i)$ and $P_{\text{stop}}^{(2)}(C_i^{\max})$ are parametrically modeled using the CF simulation data (see Figs. 5 and 6 and (8) and (9) for the parametric formulations [20]). By using $P_{\text{stop}}(S_i)$, the transition probability of the MC, $P_{\text{trans}}(S_i, S_j)$, going from state S_i to S_j, is given by (see[20, (6)])

$$\begin{aligned} P_{\text{trans}}(S_i, S_j) = \quad & \delta_{S_i^*, S_j} P_{\text{stop}}(S_i) \\ & + \left(1 - \delta_{S_i^*, S_j}\right)\left(1 - P_{\text{stop}}(S_i)\right) P_{\text{cont}}(S_i, S_j) \end{aligned}$$

(7)

where $\delta_{S_i^*, S_j} = 1$ when $S_j = S_i^*$ and $\delta_{S_i^*, S_j} = 0$ otherwise. Moreover, S_i^* denotes a state that indicates the end of CFs at the subsequent time step. In addition, $P_{\text{cont}}(S_i, S_j)$ is the probability that the cascade continues and transitions from state S_i to S_j. Similar to $P_{\text{stop}}(S_i)$, $P_{\text{cont}}(S_i, S_j)$ is also parametrically formulated by using data from CF simulations (see [20, (10) and Figs. 7 and 8]).

Appendix B DC Power Flow-Based CFs

We consider a generalized power transmission system with a set of generators and nongenerator buses denoted by $\mathcal{G}$ and $\mathcal{L}$, respectively. The buses are interconnected by m power transmission lines. Furthermore, let $g_i \geq 0$ be the generated power at generation bus $i \in \mathcal{G}$ (maximum generation value is $G_i^{\max}$), and $l_j \leq 0$ be the power consumption at the load bus $j \in \mathcal{L}$ (actual initial demand is L_j). The three grid operating characteristic parameters are described in Appendix A. A transmission-line in a power grid trips if the estimated power flow through the line is larger than its capacity. The dc power flow equations are [48]

$$\mathbf{f} = A\mathbf{p}$$

(8)

where $\mathbf{f} \in \mathbb{R}^m$ is a vector of length m that represents power flow through the m transmission lines, $\mathbf{p} \in \mathbb{R}^k$ is a vector that represents input power to the k buses in the grid (except the reference generator), and $A \in \mathbb{R}^{m \times k}$ is a matrix that consists of connectivities of buses through transmission lines and impedances of the lines. Since the system of (8) does not have a unique solution, a standard optimization approach is used to solve (8), which minimizes the following objective function [54]:

$$\min_{g_i, l_j} \left(\sum_{i \in \mathcal{G}} w_i^g g_i + \sum_{j \in \mathcal{L}} w_j^l l_j \right), \text{ subject to}$$

(a) DC power flow equations: $\mathbf{f} = A\mathbf{p}$

(b) Limits on the generator's power: $0 \le g_i \le G_i^{\max}, i \in \mathcal{G}$

(c) Limits on the controllable loads: $(1 - \theta_t)L_j \le b_j \le 0$

where $l_j = \theta_t L_j + b_j, j \in \mathcal{L}, b_j$ is to be determined,

(d) Limits on the power flow, f_l, through lines:

$|f_l| \le C_l, C_l =$ capacity of lth line, $l \in \{1, \dots, m\}$,

(e) Power generation and consumption constraints:

$$\sum_{i \in \mathcal{G}} g_i + \sum_{j \in \mathcal{L}} l_j = 0 \,.$$

(9)

Here, $w_i^g > 0$ and $w_j^l > 0$ weigh the power generation cost at the ith generator bus and the power consumption cost at the jth load bus, respectively. As in [20], we assume the LS cost is larger than the generation cost so that a load can only be curtailed when either the power generated is inadequate or the transmission capacity is limited. The solution of (9) provides amount of generation g_i, load consumption l_j (thus load-shed) and power flow through the lth line, f_l. Note that balancing of power generation and demand is done in our simulation before executing the dc power-flow calculation.

References

1. P. Hines, K. Balasubramaniam and E. C. Sanchez, "Cascading failures in power grids", *IEEE Potentials*, vol. 28, no. 5, pp. 24-30, Sep./Oct. 2009.
2. M. Panteli and P. Mancarella, "Modeling and evaluating the resilience of critical electrical power infrastructure to extreme weather events", IEEE Syst. J., vol. 11, no. 3, pp. 1733-1742, Sep. 2017.
3. G. Fu et al., "Integrated approach to assess the resilience of future electricity infrastructure networks to climate hazards", IEEE Syst. J., vol. 12, no. 4, pp. 3169-3180, Dec. 2018.
4. H. Sabouhi, A. Doroudi, M. Fotuhi-Firuzabad and M. Bashiri, "Electrical power system resilience assessment: A comprehensive approach", IEEE Syst. J., vol. 14, no. 2, pp. 2643-2652, Jun. 2020.
5. S. Poudel, A. Dubey and A. Bose, "Risk-based probabilistic quantification of power distribution system operational resilience", *IEEE Syst. J.*, vol. 14, no. 3, pp. 3506-3517, Sep. 2019.
6. J. Bialek et al., "Benchmarking and validation of cascading failure analysis tools", IEEE Trans. Power Syst., vol. 31, no. 6, pp. 4887-4900, Nov. 2016.
7. S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley and S. Havlin, "Catastrophic cascade of failures in interdependent networks", *Nature*, vol. 464, no. 7291, pp. 1025-1028, 2010.

8. M. Rahnamay-Naeini and M. M. Hayat, "Cascading failures in interdependent infrastructures: An interdependent Markov-chain approach", IEEE Trans. Smart Grid, vol. 7, no. 4, pp. 1997-2006, Jul. 2016.

9. P. Hines, J. Apt and S. Talukdar, "Large blackouts in north america: Historical trends and policy implications", *Energy Policy*, vol. 37, pp. 5249-5259, 2009.

10. H. Guo, C. Zheng, H.H.-C. Iu and T. Fernando, "A critical review of cascading failure analysis and modeling of power system", *Renewable Sustain. Energy Rev.*, vol. 80, pp. 9-22, 2017.

11. J. Bertsch, C. Carnal, D. Karlson, J. McDaniel and K. Vu, "Wide-area protection and power system utilization", Proc. IEEE, vol. 93, no. 5, pp. 997-1003, May 2005.

12. S. Pahwa, C. Scoglio, S. Das and N. Schulz, "Load-shedding strategies for preventing cascading failures in power grid", *Electric Power Compon. Syst.*, vol. 41, no. 9, pp. 879-895, 2013.

13. R. Faranda, A. Pievatolo and E. Tironi, "Load shedding: A new proposal", *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2086-2093, Nov. 2007.

14. F. Lin, "Worst-case load shedding in electric power networks", *IEEE Trans. Control Netw. Syst.*, vol. 6, no. 3, pp. 1269-1277, Sep. 2019.

15. Y. Zhang and S. Moura, "Stochastic optimal load shedding with heterogeneous load zones", Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf., pp. 1-5.

16. A. M. Bolhasan, N. T. Kalantari and S. N. Ravadanegh, "Load Shedding Emergency and Local Control", *Microgrid Architectures Control Protection Methods*, pp. 447-462, 2020.

17. M. Begovic, D. Novosel, D. Karlsson, C. Henville and G. Michel, "Wide-area protection and emergency control", *Proc. IEEE*, vol. 93, no. 5, pp. 876-891, May 2005.

18. "Reliability Coordination—Transmission Loading Relief (TLR), Standard IRO-006-4", 2007.

19. W. Ju, "Modeling simulation and analysis of cascading outages in power systems", 2018.

20. M. Rahnamay-Naeini, Z. Wang, N. Ghani, A. Mammoli and M. M. Hayat, "Stochastic analysis of cascading-failure dynamics in power grids", *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1767-1779, Jul. 2014.

21. S. Soltan, D. Mazauric and G. Zussman, "Analysis of failures in power grids", IEEE Trans. Control Netw. Syst., vol. 4, no. 2, pp. 288-300, Jun. 2017.

22. I. Dobson, B. A. Carreras, V. E. Lynch and D. E. Newman, "Complex systems analysis of series of blackouts: Cascading failure critical points and self-organization", *Chaos: An Interdiscipl. J. Nonlinear Sci.*, vol. 17, no. 2, 2007.

23. J. Kim, K. R. Wierzbicki, I. Dobson and R. C. Hardiman, "Estimating propagation and distribution of load shed in simulations of cascading blackouts", IEEE Syst. J., vol. 6, no. 3, pp. 548-557, Sep. 2012.

24. M. Korkali, J. G. Veneman, B. F. Tivnan, J. P. Bagrow and P. D. Hines, "Reducing cascading failure risk by increasing infrastructure network interdependence", *Sci. Rep.*, vol. 7, pp. 1-13, 2017.

25. K. Zhao, A. Kumar, T. P. Harrison and J. Yen, "Analyzing the resilience of complex supply network topologies against random and targeted disruptions", IEEE Syst. J., vol. 5, no. 1, pp. 28-39, Mar. 2011.

26. Y.-P. Fang, N. Pedroni and E. Zio, "Comparing network-centric and power flow models for the optimal allocation of link capacities in a cascade-resilient power transmission network", *IEEE Syst. J.*, vol. 11, no. 3, pp. 1632-1643, Sep. 2017.

27. M. Parandehgheibi, E. Modiano and D. Hay, "Mitigating cascading failures in interdependent power grids and communication networks", Proc. IEEE Int. Conf. Smart Grid Commun, pp. 242-247, 2014.

28. D. Z. Tootaghaj, N. Bartolini, H. Khamfroush, T. He, N. R. Chaudhuri and T. La Porta, "Mitigation and recovery from cascading failures in interdependent networks under uncertainty", *IEEE Trans. Control Netw. Syst.*, vol. 6, no. 2, pp. 501-514, Jun. 2019.

29. D. Bienstock, "Optimal control of cascading power grid failures", *Proc. 50th IEEE Conf. Decis. Control Eur. Control*, pp. 2166-2173, 2011.

30. M. Sinha, M. Panwar, R. Kadavil, T. Hussain, S. Suryanarayanan and M. Papic, "Optimal load shedding for mitigation of cascading failures in power grids", *Proc. 10th ACM Int. Conf. Future Energy Syst*, pp. 416-418, 2019.

31. J. Chen, J. S. Thorp and I. Dobson, "Cascading dynamics and mitigation assessment in power system disturbances via a hidden failure model", *Int. J. Elect. Power Energy Syst.*, vol. 27, no. 4, pp. 318-326, 2005.

32. R. Yao, K. Sun, F. Liu and S. Mei, "Management of cascading outage risk based on risk gradient and Markovian Tree search", *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4050-4060, Jul. 2018.

33. K. Zhou, I. Dobson, Z. Wang, A. Roitershtein and A. P. Ghosh, "A markovian influence graph formed from utility line outage data to mitigate large cascades", *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 3224-3235, Jul. 2020.

34. Y. Yang, T. Nishikawa and A. E. Motter, "Small vulnerable sets determine large network cascades in power grids", *Science*, vol. 358, no. 6365, 2017.

35. J. Yan, Y. Tang, H. He and Y. Sun, "Cascading failure analysis with DC power flow model and transient stability analysis", *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 285-297, Jan. 2015.

36. J. Yan, H. He, X. Zhong and Y. Tang, "Q-learning-based vulnerability analysis of smart grid against sequential topology attacks", *IEEE Trans. Inf. Forensics Secur.*, vol. 12, no. 1, pp. 200-210, Jan. 2017.

37. J. Qi, K. Sun and S. Mei, "An interaction model for simulation and mitigation of cascading failures", *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 804-819, Mar. 2015.

38. C. Y. Ma, D. K. Yau, X. Lou and N. S. Rao, "Markov game analysis for attack-defense of power networks under possible misinformation", *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1676-1686, May 2013.

39. W. Liao, S. Salinas, M. Li, P. Li and K. A. Loparo, "Cascading failure attacks in the power system: A stochastic game perspective", *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2247-2259, Dec. 2017.

40. Q. Zhu and T. Başar, "A dynamic game-theoretic approach to resilient control system design for cascading failures", *Proc. 1st Int. Conf. High Confidence Netw. Syst*, pp. 41-46, 2012.

41. Z. Wang et al., "Impacts of operators' behavior on reliability of power grids during cascading failures", *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6013-6024, Nov. 2018.

42. R. A. Shuvro, P. Das and M. M. Hayat, "Balancing smart grid's performance enhancement and resilience to cyber threat", *Proc*, pp. 235-241.

43. U. Nakarmi and M. Rahnamay-Naeini, "A Markov chain approach for cascade size analysis in power grids based on community structures in interaction graphs", *Proc. Int. Conf. Probabilistic Methods Appl. Power Syst*, pp. 1-6, 2020.

44. Z. Zhang, R. Yao, S. Huang, Y. Chen, S. Mei and K. Sun, "An online search method for representative risky fault chains based on reinforcement learning and knowledge transfer", *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 1856-1867, May 2020.

45. A. Wang, Y. Luo, G. Tu and P. Liu, "Vulnerability assessment scheme for power system transmission networks based on the fault chain theory", *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 442-450, Feb. 2011.

46. J. Guo, F. Liu, J. Wang, J. Lin and S. Mei, "Toward efficient cascading outage simulation and probability analysis in power systems", *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2370-2382, May 2018.

47. M. L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming, Hoboken, NJ, USA:Wiley, 2014.

48. J. D. Glover, M. S. Sarma and T. Overbye, Power System Analysis and Design SI Version, Boston, MA, USA:Cengage Learning, 2012.

49. E. Xypolytou, T. Zseby, J. Fabini and W. Gawlik, "Detection and mitigation of cascading failures in interconnected power systems", *Proc. IEEE PES Innov. Smart Grid Technol. Conf.*, pp. 1-6, 2017.

50. S. Yang, W. Chen, X. Zhang, C. Liang, H. Wang and W. Cui, "A graph-based model for transmission network vulnerability analysis", *IEEE Syst. J.*, vol. 14, no. 1, pp. 1447-1456, Mar. 2020.

51. H. Cetinay, S. Soltan, F. A. Kuipers, G. Zussman and P. Van Mieghem, "Comparing the effects of failures in power grids under the AC and DC power flow models", *IEEE Trans. Netw. Sci. Eng.*, vol. 5, no. 4, pp. 301-312, Oct.–Dec. 2018.

52. R. D. Zimmerman, C. E. Murillo-Sánchez and R. J. Thomas, "Matpower: Steady-state operations planning and analysis tools for power systems research and education", *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12-19, Feb. 2011.

53. D. Li and S. K. Jayaweera, "Machine-learning aided optimal customer decisions for an interactive smart grid", *IEEE Syst. J.*, vol. 9, no. 4, pp. 1529-1540, Dec. 2015.

54. B. A. Carreras, V. E. Lynch, I. Dobson and D. E. Newman, "Critical points and transitions in an electric power transmission model for cascading failure blackouts", *Chaos Interdiscipl. J. Nonlinear Sci.*, vol. 12, no. 4, pp. 985-994, 2002.