Failure Probability Constrained AC Optimal Power Flow

Anirudh Subramanyam [ID], *Member, IEEE*, Jacob Roth [ID], Albert Lam [ID], and Mihai Anitescu [ID], *Member, IEEE*

Abstract—**Despite cascading failures being the central cause of blackouts in power transmission systems, existing operational and planning decisions are made largely by ignoring their underlying cascade potential. This paper posits a reliability-aware AC Optimal Power Flow formulation that seeks to design a dispatch point which has a low operator-specified likelihood of triggering a cascade starting from any single component outage. By exploiting a recently developed analytical model of the probability of component failure, our Failure Probability-constrained ACOPF (FP-ACOPF) utilizes the system's expected first failure time as a smoothly tunable and interpretable signature of cascade risk. We use techniques from bilevel optimization and numerical linear algebra to efficiently formulate and solve the FP-ACOPF using off-the-shelf solvers. Extensive simulations on the IEEE 118-bus case show that, when compared to the unconstrained and N-1 security-constrained ACOPF, our probability-constrained dispatch points can significantly lower the probabilities of long severe cascades and of large demand losses, while incurring only minor increases in total generation costs.**

Index Terms—**AC optimal power flow, cascading failures.**

I. INTRODUCTION

A CASCADING failure in a power transmission system refers to a sequence of dependent outages of individual system components that successively disable parts of the grid, leading to a significant loss of served power or large blackout in the worst case. Accounts of major blackouts reveal that cascading failures are often triggered by an initial event that is largely unpredictable (such as extreme weather) but are sustained by subsequent events that are causally linked via Kirchhoff's laws and automatic control actions of protection devices. For example, the outage of a single component can lead to redistribution of power flows in the remainder of the network in a way that can cause large overcurrents on some transmission lines. This, in turn, may trigger protection relays to disconnect these lines automatically if the current flow exceeds some threshold rating, or it may lead to eventual thermal failure if the overcurrents remain sustained for a long time.

The large direct and indirect costs associated with blackouts, along with mandatory standards set forth by the North American Electric Reliability Corporation (NERC) to address cascading outages, have motivated the development of a plethora of tools for the simulation and analysis of cascading failures; see reviews [1]–[4] and references therein. These tools can be broadly classified as *(a)* complex network approaches that consider the pure topological properties of power networks while ignoring or simplifying the underlying physics (e.g., see [5], [6]), *(b)* high-level statistical models built either on historical/simulation data or on simplified power system physics (e.g., see [7], [8]), *(c)* quasi-steady-state methodologies that utilize DC or AC power flow (steady state) models, typically in conjunction with models of protection mechanisms or operator interventions [9]–[13], and rely on Monte Carlo or enumerative sampling (e.g., see [14]–[16]), *(d)* dynamics-based models in which the state evolution is resolved with a physics-based representation of the grid dynamics (e.g., see [17], [18]) to study post-fault transient dynamics [19], synchronization [20]–[22], or behavior over longer time horizons [23]–[25].

While these tools have proved immensely useful in deepening our understanding of cascading failures, their use for risk mitigation and decision-making have been largely restricted to limiting the propagation of a cascading failure, e.g., via controlled load shedding or intentional islanding, *after* severe contingencies have already occurred (e.g., see [26]–[28]). Existing practices for the prevention of cascading failures *before* they occur, by tuning and modifying the controllable properties of the power network, has largely relied upon $N-k$ security criteria and simulation-based contingency analyses as notional surrogates for reducing cascading likelihood.

Although simulation-based tools can influence long-term cascade mitigation solutions, such as line capacity or generator margin allocations and protection system enhancements (e.g., see [29]–[31]), they are fundamentally limited in preventing cascades in short-term operations such as economic dispatch or optimal power flow. This is because they do not provide any direct functional relationship between control parameters and cascade potential, or they entail expensive Monte Carlo sampling and numerical integration requirments and are thus challenging to incorporate within optimization algorithms.

This work proposes to incorporate in the classical ACOPF model, an analytical – as opposed to simulation-based – model of cascade severity that is an *explicit function* of the network properties and dispatch point. In contrast to existing methods, we aim to determine a dispatch point subject to the constraint that

the *probability of individual component failure* remains below an operator-prescribed threshold. Our model capitalizes upon results from [32], which in contrast to other approaches for simulating cascading failures, provides an *analytic expression* for the failure probability as a function of the dispatch point. This is achieved by modeling *Gaussian load and generation fluctuations* in the AC power flow dynamics, and interpreting the latter as the diffusion of a particle in an energy landscape subject to stochastic forcing. Large deviations theory then provides the means to analytically relate the failure probability to the underlying energy surface.

For a given system state (consisting of voltages and power flows) at equilibrium, the analytical expression for the failure probability of an individual component requires solving a nonlinear optimization problem that computes a "most likely" failure state starting from this equilibrium state. However, since the latter is contingent on the dispatch point, explicitly constraining the failure probability of an *individual component*, or of a *cascading sequence of multiple components*, is tantamount to solving a bilevel optimization problem. We settle for the former and demonstrate that constraining individual failure probabilities, which is equivalent to increasing the system's expected first failure time–or decreasing its *failure rate*– can be an effective surrogate for constraining the probability of a cascading failure sequence. Solving the bilevel model, however, is computationally challenging since it is nonlinear, nonconvex, and involves eigenvalues and determinants of high-dimensional matrices that scale with the network size. Nevertheless, we show that this eigenvalue- and determinant-constrained model can be reformulated entirely algebraically and solved efficiently using standard solvers, if we exploit the low-rank nature of the failure constraints along with the first- and second-order optimality conditions of the nested problem.

Our key finding is that it is possible to modify the dispatch point so as to satisfy a prescribed threshold of failure probability. Specifically, the system's expected first failure time, and hence, the probability of cascading over a given time horizon, can be lowered significantly without incurring vastly higher generation costs or load shedding.

We believe this is the first work that captures some notion of cascading risk in operational dispatch. Although it can be viewed as a probabilistic $N - 1$ variant, our approach offers several advantages over classical $N - k$ models. The first crucial difference is that, by tuning the system's *failure rate limit*, we can systematically and *smoothly* explore the trade-offs between cascade potential, dispatch costs and operator conservatism, without significant increase in computational complexity. In contrast, $N - k$ approaches must resort to a non-smooth control of k to achieve the same objective, while invariably incurring a sharp increase in combinatorial complexity. Another subtle, yet practically useful, advantage of our approach is its *interpretability*. Indeed, the benefit (in terms of reliability) of increasing k in $N - k$ approaches, is difficult to convey outside the domain. Our approach, on the other hand, allows the system operator to decide between a *failure rate limit* of $10^{-6}s^{-1}$ or $10^{-15}s^{-1}$ (for example), which is equivalent to deciding between observing the first failure once every 10^6 or 10^{15} seconds. This is a statement that is better aligned with the philosophy of regulatory constraints which tend to be in occurrences per unit of time.

The paper is organized as follows. Section II presents assumptions, and reviews the failure probability model of [32]. Section III presents the failure probability-constrained ACOPF model along with its reformulation. Section IV demonstrates the effectiveness of our method via extensive simulations, and Section V offers conclusions and directions for future work.

II. FAILURE PROBABILITY MODEL

A. Notation

We use $\mathcal{N} = \{1, \ldots, n_b\}$ to denote the set of buses, and $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$ to denote the set of transmission lines, where $l = (i, j) \in \mathcal{L}$ is a line from bus i and to bus j. We denote the set of generators by $\mathcal{G} = \{1, \ldots, n_g\}$, and let $\mathcal{G}_i$ be the set of generators connected to bus $i \in \mathcal{N}$; note that $\mathcal{G} = \cup_{i \in \mathcal{N}} \mathcal{G}_i$. For ease of notation, we define $\mathcal{N}' := \{i \in \mathcal{N} : \mathcal{G}_i = \emptyset\}$ to denote the set of non-generator buses, *i.e.*, those that are not connected to any generator and similarly, $\mathcal{L}' := \{(i, j) \in \mathcal{L} : \mathcal{G}_i = \emptyset \text{ or } \mathcal{G}_j = \emptyset\}$ to denote the set of lines that are connected to at least one non-generator bus. The nodal admittance matrix $Y = G + \sqrt{-1}B \in \mathbb{C}^{n_b \times n_b}$ with conductance and susceptance matrices $G \in \mathbb{R}^{n_b \times n_b}$ and $B \in \mathbb{R}^{n_b \times n_b}$, respectively.

We denote active and reactive power generations by $p_g, q_g \in \mathbb{R}^{n_g}$, active and reactive power demands by $p_d, q_d \in \mathbb{R}^{n_b}$, and net active and reactive powers by $p_{net}, q_{net} \in \mathbb{R}^{n_b}$, where we define $p_{net,i} := p_{d,i} - \sum_{k \in \mathcal{G}_i} p_{g,k}$ and $q_{net,i} := q_{d,i} - \sum_{k \in \mathcal{G}_i} q_{g,k}$ for $i \in \mathcal{N}$. The voltage magnitudes and angles are denoted by $V \in \mathbb{R}^{n_b}$ and $\theta \in \mathbb{R}^{n_b}$, and the generator angular velocities are denoted by $\omega \in \mathbb{R}^{n_g}$.

For a vector $z \in \mathbb{C}^n$, we use $\|z\|$ and z^* to denote its Euclidean norm and Hermitian transpose, respectively, and we use e^z and $\log(z)$ to denote the vectors $(e^{z_1}, \ldots, e^{z_n}) \in \mathbb{C}^n$ and $(\log(z_1), \ldots, \log(z_n)) \in \mathbb{C}^n$. Given vectors $z, \tilde{z} \in \mathbb{C}^n$, we use $z \circ \tilde{z}$ to denote the Hadamard product $(z_1\tilde{z}_i, \ldots, z_n\tilde{z}_n) \in \mathbb{C}^n$. For a matrix $A \in \mathbb{C}^{n \times n}$, we use $\rho(A)$, $\det(A)$, and $\mathrm{adj}(A)$ to denote its spectral radius, determinant and adjugate, respectively, and $A \succ 0$ ($\succeq 0$) to indicate that it is positive definite (semi-definite). For a given matrix $A \succeq 0$ and vector $z \in \mathbb{C}^n$, we use $\|z\|_A$ to denote $\sqrt{z^*Az}$.

In formulating the probability model, it will be convenient to divide the system state vector $(p_g, q_g, V, \theta, \omega) \in \mathbb{R}^m$ ($m := 3n_g + 2n_b$) into two distinct sets. After selecting an arbitrary generator bus $\sigma \in \mathcal{N} \setminus \mathcal{N}'$, $|\mathcal{G}_\sigma| = 1$, as the slack bus in the steady-state context and reference bus in the dynamics context, we collect all voltage magnitudes at non-generator buses, and phase angles and angular velocities at non-slack buses in the state vector $x = (\{V_i\}_{i \in \mathcal{N}'}, \{\theta_i\}_{i \in \mathcal{N} \setminus \{\sigma\}}, \{\omega_i\}_{i \in \mathcal{G} \setminus \mathcal{G}_\sigma}) \in \mathbb{R}^d$, $d = |\mathcal{N}'| + |\mathcal{N}| + |\mathcal{G}| - 2$. The remaining voltage components, angular velocities and power generations are aggregated in the vector $y = (\{V_i\}_{i \in \mathcal{N} \setminus \mathcal{N}'}, \theta_\sigma, \omega_\sigma, p_g, q_g) \in \mathbb{R}^{m-d}$. This partition of the system variables, illustrated in Table I, is purely for notational convenience. In particular, any specified value of y implicitly defines a set of state vectors x that solve the power flow equations, as we shall elaborate in Section II-C.

TABLE I
PARTITION OF SYSTEM VARIABLES INTO SUB-VECTORS x AND y

	Load buses $\mathcal{N}'$		Generator buses $\mathcal{N} \setminus (\mathcal{N}' \cup \{\sigma\})$				Slack bus σ			
	θ	V	ω	θ	V	p_g, q_g	ω	θ	V	p_g, q_g
x	✓	✓	✓	✓						
y					✓	✓	✓	✓	✓	✓

B. Assumptions

Motivated by the automatic control actions of protection relays, we assume that individual system components fail in a deterministic manner according to a component-specific state equation. Specifically, we assume that a component, such as a transmission line $l \in \mathcal{L}$, fails if the value of a certain function $\Theta_l : \mathbb{R}^m \mapsto \mathbb{R}$ at the current state (x, y) exceeds a critical value $\Theta_l^{\max}$. Furthermore, we assume that after failure, a component remains failed over the entire dispatch horizon of the ACOPF.

The algebraic state functions Θ can be any continuously differentiable function of the state vector (x, y), and hence, it can be used to model various component failures, such as undervoltage load shedding at buses or line surges in transformers. For the purpose of computing a dispatch point, however, we assume that once initiated, component failures occur only due to relay trips caused by current overloads in transmission lines. Suppose that I_l^{trip} denotes the emergency current rating of a line $l = (i, j) \in \mathcal{L}$. In this case, one can define $\Theta_l(x, y)$ to be the square of the magnitude of current flow in line l (which depends only on the voltages and phase angles at its terminal buses), and $\Theta_l^{\max}$ to be the square of the corresponding rating:

$$
\begin{aligned}
\Theta_l(x, y) &= \left(|Y_{ij}| \left| V_i e^{\mathrm{i}\theta_i} - V_j e^{\mathrm{i}\theta_j} \right| \right)^2 \\
&= |Y_{ij}|^2 \left(V_i^2 + V_j^2 - 2 V_i V_j \cos(\theta_i - \theta_j) \right), \\
\Theta_l^{\max} &= (I_l^{\text{trip}})^2.
\end{aligned} \tag{1}
$$

This is partly because it simplifies the exposition, and partly because the corresponding failure model has already been studied and validated against real cascade data in [32], [33]. We note, however, that the methodology does not preclude us from considering more general definitions of Θ. Importantly, the simulations in Section IV model several other protection mechanisms beyond current overloads.

In addition, we shall assume that *(i)* the network is lossless[1], *i.e.*, $G = 0$, and *(ii)* only the subset $\mathcal{L}'$ of transmission lines that connect to at least one non-generator bus have a nonzero likelihood of failing because of stochastic fluctuations in active and reactive power demand.

C. Analytical Model of Line Failures

We model the grid's electro-mechanical behavior using a system of stochastic differential equations (SDE), by introducing

[1]This assumption is not particularly restrictive since many high-voltage power transmission networks typically have resistance/reactance ratios below 0.2. Moreover, the standard DCOPF model for dispatch widely used in industry explicitly relies on this assumption [34]; in contrast, our model is far more general since we also consider nonlinearities, reactive powers and voltage magnitudes that the DCOPF model does not.

a scalar noise parameter $\tau > 0$ into the following variant of the standard, structure-preserving model [35]–[39]:

$$
\mathrm{d}x_t^\tau = (J - S) \nabla_x \mathcal{H}(x_t^\tau, y) \mathrm{d}t + \sqrt{2\tau S} \mathrm{d}W_t. \tag{2}
$$

Here, $x_t^\tau \in \mathbb{R}^d$ denotes the system state at time t, J and S represent appropriate system matrices, and W_t is a d-dimensional vector of independent Wiener processes. When $\tau = 0$, (2) models the deterministic dynamics of the state variables in x as a function of the input vector y. Specifically, it is a singularly-perturbed version [37] of the classical differential-algebraic structure-preserving model [40, Chapter 7]. When $\tau = 0$, (2) includes the standard second-order swing equations employed in the classical model, but augments its constant active power load modeling assumption by adding linear frequency-dependent load damping terms D_l to the active power loads (see, e.g., [41]) and voltage-dependent perturbation terms D_v to the reactive power loads. The parameters D_l and D_v control the rates at which the phase angles and load voltages approach the real and reactive power flow equation manifolds, respectively [36]. Crucially, (2) converges to the classical model as the noise parameter $\tau \to 0$ and damping terms $D_l, D_v \to 0$ but with the added advantages of global well-posedness (when $\tau = 0$) and modeling Gaussian perturbations in active and reactive power demands (when $\tau > 0$). This model has been widely used for transient stability analysis in both deterministic and stochastic regimes [35], [42].

The detailed differential equations encapsulated within (2), including definitions of J and S, can be found in [32], [35], [43]. The energy function $\mathcal{H}$ is then obtained as the first integral of the deterministic dynamics, leading to the gradient-based formulation (2). Notably, the integral admits the following closed-form expression :

$$
\begin{aligned}
\mathcal{H}(x, y) := &\tfrac{1}{2}\omega^\top M \omega + \tfrac{1}{2}(V \circ e^{\mathrm{i}\theta})^* Y (V \circ e^{\mathrm{i}\theta}) \\
&+ p_{net}^\top \theta + q_{net}^\top \log(V),
\end{aligned} \tag{3}
$$

where M is the $n_g \times n_g$ diagonal matrix of generator masses. Unlike (1), the right-hand side of (3) depends on all components of x and y. It is worth pointing out that (3) has been widely used as a Lyapunov function for transient stability analysis and dynamic control of the structure preserving swing dynamics model in deterministic settings [44], [45], and it is closely related to the widely-used transient energy function of [46]. Similar to the latter, which is used to measure distance to instability, we use (3) to measure the probability of the stochastically perturbed dynamics to reach a triggering surface.

We note that the vector y is assumed to have been fixed to appropriate values *a priori*, and we shall return to choosing a value for y in the subsequent section when we describe the ACOPF formulation. For now, observe that the structure of the energy function (3) ensures that, for any specified value of y, any point $x \in \mathbb{R}^d$ that satisfies $\nabla_x \mathcal{H}(x, y) = 0$ is a solution to the lossless power flow equations:

$$
p_{net,i} + \sum_{j \in \mathcal{N}} B_{ij} V_i V_j \sin(\theta_i - \theta_j) = 0, \quad i \in \mathcal{N}, \tag{4}
$$

$$
q_{net,i} - \sum_{j \in \mathcal{N}} B_{ij} V_i V_j \cos(\theta_i - \theta_j) = 0, \quad i \in \mathcal{N}. \tag{5}
$$

In particular, any local energy minimizer,

$$\bar{x}(y) \in \arg\min_{x \in \mathbb{R}^d} \mathcal{H}(x, y) \tag{6}$$

defines a feasible solution to (4), (5) and serves as a stable equilibrium for the network dynamics defined by y. To simplify notation, we drop the dependence of $\bar{x}$ on y in the remainder of the paper.

Under the SDE model (2) initialized at $x_0^\tau = \bar{x}$, the failure probability of a line $l \in \mathcal{L}'$ (considered in isolation from the rest of the network) can be quantified by its so-called *first passage time*. The latter is defined to be the first time at which the system state x_t^τ triggers the failure condition (1) for line l. Assuming $\Theta_l(\bar{x}, y) < \Theta_l^{\max}$, this is equivalent to

$$T_l^\tau(y) := \inf\{t > 0 : \Theta_l(x_t^\tau, y) \geq \Theta_l^{\max}\}. \tag{7}$$

By exploiting results from large deviations theory [47], [48], it was shown in [32] that, as $\tau \to 0$, $T_l^\tau(y)$ is an exponential random variable whose mean satisfies the following relation:

$$\lim_{\tau \to 0} \tau \log\left(\mathbb{E}\left[T_l^\tau(y)\right]\right) = \min_{x:\Theta_l(x_t^\tau, y) = \Theta_l^{\max}} \mathcal{H}(x, y) - \mathcal{H}(\bar{x}, y) \tag{8}$$

A point at which the minimum is obtained is called the *most likely failure point* and is defined as

$$x_l^\star(y) := \arg\min_{x \in \mathbb{R}^d}\left\{\mathcal{H}(x, y) : \Theta_l(x, y) = \Theta_l^{\max}\right\}, \tag{9}$$

where we have implicitly assumed $x_l^\star$ to be unique[2]. As before, we drop the dependence of $x_l^\star$ on y to simplify notation.

For finite values of τ, we can thus use relation (8) to compute a log-approximate *failure rate* (*i.e.*, the reciprocal of the mean failure time). In addition to this log-approximation, a subexponential correction to the failure rate was also obtained in [32] yielding the following expressions:

$$\lambda_l^\tau(y) := \mathsf{pf}_l^1(y) \times \mathsf{ef}_l(y) \tag{10}$$

$$\mathsf{pf}_l^1(y) := \mathsf{pf}_l^0(y) \times \left(1 + \frac{\tau}{\mathcal{H}(x_l^\star, y) - \mathcal{H}(\bar{x}, y)}\right). \tag{11}$$

$$\mathsf{pf}_l^0(y) := \|\nabla_x \mathcal{H}(x_l^\star, y)\|_S^2 \sqrt{\frac{\det \nabla_{xx}^2 \mathcal{H}(\bar{x}, y)}{2\pi\tau C_l^\star(y)}} \tag{12}$$

$$\mathsf{ef}_l(y) := \exp\left[-\frac{\mathcal{H}(x_l^\star, y) - \mathcal{H}(\bar{x}, y)}{\tau}\right], \tag{13}$$

where $C_l^\star(y)$ is a factor accounting for the curvature of the failure surface in the vicinity of $x_l^\star$, and is given by:

$$C_l^\star(y) := \nabla_x^\top \mathcal{H}(x_l^\star, y) \, \mathrm{adj}\left(X_l(y)\right) \nabla_x \mathcal{H}(x_l^\star, y) \tag{14}$$

$$X_l(y) := \nabla_{xx}^2 \mathcal{H}(x_l^\star, y) - \mu_l^\star \nabla_{xx}^2 \Theta_l(x_l^\star, y), \tag{15}$$

where $\mu_l^\star \in \mathbb{R}$ is the optimal Lagrange multiplier in the constrained optimization problem (9). In (10)–(13), pf and ef stand for the *pre-factor* and *exponential factor*, respectively. The latter follows directly from (8). The former is a subexponential correction that is obtained by quantifying how the probability distribution over the system's state space changes with

time. Specifically, from an initial *quasistationary ensemble* [48] of particles distributed as $p_{qst}^\tau(x) \propto \exp[-\tau^{-1}(\mathcal{H}(x, y) - \mathcal{H}(\bar{x}, y))]$, probability mass is lost over time through the surface $\{x \in \mathbb{R}^d : \Theta_l(x, y) = \Theta_l^{\max}\}$. A Laplace integral approximation of this probability yields the (zeroth order) failure rate: $\mathsf{pf}_l^0(y) \times \mathsf{ef}_l(y)$. However, for large values of τ, this zeroth order expression can lead to non-physical behavior as the $1/\sqrt{\tau}$ term in $\mathsf{pf}_l^0(y)$ may cause the failure rate to decrease as τ increases. Therefore, the first order correction in (11) is used to remedy this behavior, and it has demonstrated strong empirical performance in approximating the failure rate across a wide range of τ [32].

In summary, the distribution of failure times for line $l \in \mathcal{L}'$ can be well-approximated by $T_l^\tau(y) \sim \mathrm{Exp}(\lambda_l^\tau(y))$, and therefore, the probability of observing line l fail in the first t seconds can be given by

$$\mathbb{P}(T_l^\tau(y) \leq t) = 1 - \exp\left[-\lambda_l^\tau(y)t\right]. \tag{16}$$

A noteworthy feature of this model is that (16) gives a closed-form expression for the failure probability, thus avoiding the need for direct integration of the differential model (2).

III. FAILURE PROBABILITY-CONSTRAINED AC OPTIMAL POWER FLOW FORMULATION

A. Conceptual Formulation

The probability of transmission line failure (16) is a function of $y = (\{V_i\}_{i \in \mathcal{N} \setminus \mathcal{N}'}, \theta_\sigma, p_g, q_g)$, the voltage magnitudes at generator buses, the voltage angle at the slack bus and the active and reactive power generations. The choice of y also determines the equilibrium operating point $\bar{x}(y)$ through the power flow (4), (5); we shall denote these simply as $x = (\{V_i\}_{i \in \mathcal{N}'}, \{\theta_i\}_{i \in \mathcal{N} \setminus \{\sigma\}},)$ and along with the vector y, they constitute decision variables[3] in our Failure Probability-constrained ACOPF (FP-ACOPF) formulation:

$$\underset{x,y}{\text{minimize}} \quad \sum_{k \in \mathcal{G}} c_k(p_{g,k}) \tag{17}$$

subject to $(4)-(5)$,

$$p_{g,k}^{\min} \leq p_{g,k} \leq p_{g,k}^{\max}, \quad k \in \mathcal{G}, \tag{18}$$

$$q_{g,k}^{\min} \leq q_{g,k} \leq q_{g,k}^{\max}, \quad k \in \mathcal{G}, \tag{19}$$

$$V_i^{\min} \leq V_i \leq V_i^{\max}, \quad i \in \mathcal{N}, \tag{20}$$

$$\Theta_l(x, y) \leq (I_l^{\lim})^2, \quad l \in \mathcal{L}, \tag{21}$$

$$\lambda_l^\tau(y) \leq \lambda_l^{\lim} := -t_H^{-1}\log\left(1 - \epsilon_l^{\lim}\right), l \in \mathcal{L}' \tag{22}$$

The objective function (17) minimizes the cost of generation where c_k is a function representing the cost of generating active power $p_{g,k}$ at generator $k \in \mathcal{G}$. The constraints (18), (19) and (20) ensure that the power generations and voltage magnitudes stay within predefined limits. The constraint (21) limits the amount of current flow on the transmission lines, where we have distinguished the flow limit $I_l^{\lim}$ from the value I_l^{trip} used in the failure model (1). The latter quantity is the value at which

---

[2]This is not a strong assumption based on the validation in [32].

[3]Note that we do not include the frequency variables ω as decisions, since we can set $\omega = 0$ without loss of optimality.

protection relays would automatically disconnect the line, and it is typically much larger than the former, which is determined based on thermal considerations. Finally, constraint (22) imposes an upper bound on the failure rate of line $l \in \mathcal{L}'$, which by definition, is equivalent to delaying its expected failure time to be greater than $(\lambda_l^{\lim})^{-1}$.

Observe that, using expression (16), the failure rate constraint (22) is equivalent to a chance constraint which ensures that the probability of line $l \in \mathcal{L}'$ failing over a time horizon t_H is less than some operator-prescribed limit $\epsilon_l^{\lim}$ (say 1%):

$$\mathbb{P}(T_l^\tau(y) \leq t_H) \leq \epsilon_l^{\lim}, \quad l \in \mathcal{L}'. \quad (23)$$

Constraint (23) only limits the failure probability of a single line. While it can also be used as an approximation for limiting the probability of a cascading failure sequence involving more than one line[4], we highlight that assessing the true probability of a cascading sequence needs careful consideration and modeling of subsequent events, including system dynamics and protection behavior, which is outside the scope of the proposed optimization model. The individual failure rate constraint (22) is thus closer in spirit to a probabilistic $N-1$ constraint, as opposed to a full-fledged cascading failure constraint.

The presented model is a standard ACOPF formulation with the exception of the *failure rate constraint* (22). We now present an efficient reformulation of these rate constraints that can be incorporated into standard optimization solvers.

B. Reformulation of Failure Rate Constraints: Key Steps

Constraining the failure rate via (22) is nontrivial because it involves *(i)* the solution of a nested nonconvex optimization problem (9) to obtain $x_l^\star(y)$, and *(ii)* constraints involving determinants of large matrices, namely $\nabla_{xx}^2 \mathcal{H}(x, y)$ and $X_l(y)$, see (12) and (14). This section shows that these can be circumvented by exploiting the low-rank property of the line failure function Θ_l, and a partial Talyor expansion of the energy function $\mathcal{H}$ around the equilibrium operating point x.

1) Low Rank Factorization of Failure Function: We shall capitalize on the fact that, the Hessian of the failure function Θ_l admits a low-rank factorization of the form:

$$\nabla_{xx}^2 \Theta_l(x, y) = Q_l(x, y) K_l(x, y) Q_l^\top(x, y), \quad (24)$$

where $Q_l(x, y)$ is a $d \times r_l$ matrix, with $r_l \ll m$ and $K_l(x, y)$ is a $r_l \times r_l$ diagonal matrix. Observe that this low-rank structure is natural whenever Θ_l is a function of a small number of state variables, as is the case when Θ_l models the failure of transmission lines, system buses, generators or transformers.

Although such a low-rank factorization can be computed numerically (*e.g.*, using an eigenvalue decomposition), it can also

[4]For example, limiting the probability that line l_2 will fail after line l within time t_H can be enforced via:

$$\int_0^{t_H} \mathbb{P}\left(T_l^\tau(y) \in [t, t+\mathrm{d}t]\right) \mathbb{P}\left(T_{l_2}^\tau(y; Y') \in [t, t_H]\right) \mathrm{d}t \leq \epsilon^{\lim},$$

where $T_{l_2}^\tau(y; Y')$ is the failure time of line l_2 calculated using a modified admittance matrix Y' after removing line l. Since both probabilities in the above integral involve exponential random variables (16), it can be analytically expressed as a single algebraic constraint using the corresponding failure rates. Observe that this probability is upper bounded by the left-hand side of (23).

be obtained analytically in several cases, such as in the examples of Section II-B. To illustrate this, [49, Appendix A] presents explicit closed-form expressions for Q_l and K_l when Θ_l is the line overcurrent function (1). In particular, it shows that $r_l = 3$ for a line $l = (i, j)$ where both $i, j \in \mathcal{N}'$ are non-generator buses, whereas $r_l = 2$ otherwise. The fact that the rank $r_l \leq 3$ has nothing to do with the specific functional form of (1); in fact, $r_l \leq 3$ will always hold whenever Θ_l is a function only of V_i, V_j and the angle difference $\theta_i - \theta_j$.

2) Taylor Approximation of Energy Function: The value of the energy function at the most likely failure point, $\mathcal{H}(x_l^\star, y)$, is approximated using a quadratic Taylor polynomial centered around the equilibrium operating point x that is determined by the ACOPF formulation:

$$\mathcal{H}(x_l^\star, y) = \mathcal{H}(x, y) + \frac{1}{2}(x_l^\star - x)^\top \nabla_{xx}^2 \mathcal{H}(x, y)(x_l^\star - x) \quad (25)$$

where we have ignored the first-order term since the FP-ACOPF constrains x to satisfy the power flow (4), (5), which are equivalent to $\nabla_x \mathcal{H}(x, y) = 0$. We shall comment on the accuracy of (25) in the next subsection.

C. Reformulation of the 'Most Likely Failure Point' Problem

Equations (24) and (25) can be used to reformulate the nested nonconvex *most likely failure point* optimization problem (9). In particular, the following proposition shows that they can be used to generate tractable algebraic constraints under which the solution of (9) can be characterized by its first- and second-order optimality conditions

Proposition 1: Fix a line $l \in \mathcal{L}'$, and assume that $x, x_l^\star \in \mathbb{R}^d$ and $y \in \mathbb{R}^{m-d}$ are such that:
1) $\nabla_{xx}^2 \mathcal{H}(x, y) \succ 0$, and
2) the Taylor approximation (25) is applicable.

If $x_l^\star$ and $\mu_l^\star \geq 0$ satisfy (26)–(28), then $x_l^\star$ is a local solution of the optimization problem (9).

$$\nabla_{xx}^2 \mathcal{H}(x, y)(x_l^\star - x) = \mu_l^\star \nabla_x \Theta_l(x_l^\star, y) \quad (26)$$

$$\Theta_l(x_l^\star, y) = \Theta_l^{\max} \quad (27)$$

$$\mu_l^\star \rho(A_l) < 1, \quad (28)$$

where $A_l := K_l(x_l^\star, y) Q_l^\top(x_l^\star, y) \left[\nabla_{xx}^2 \mathcal{H}(x, y)\right]^{-1} Q_l(x_l^\star, y)$ is a $r_l \times r_l$ matrix.

Proof: See [49, Appendix B] of the preprint. □

Equations (26) and (27) are the first-order optimality conditions, whereas the spectral inequality (28) is derived from the second-order optimality condition of the nested problem (9). It is noteworthy that the matrix A_l in the left-hand side of the latter inequality is a small $r_l \times r_l$ matrix, and hence it is possible to reformulate the inequality in a completely algebraic form without requiring a numerical eigenvalue computation.

When $r_l = 2$, it can be verified (*e.g.*, using the characteristic polynomial of A_l) that inequality (28) is equivalent to:

$$\mu_l^\star \operatorname{tr}(A_l) - (\mu_l^\star)^2 \det(A_l) < 1, \quad (29)$$

where the trace and determinant can be computed in algebraic closed-form, since A_l is a 2×2 matrix.

When $r_l = 3$, it is nontrivial to get a tractable reformulation, since A_l need not be symmetric. We consider two cases.

- If $K_l(x_l^\star, y) \succ 0$, then $\mu_l^\star \rho(A_l) < 1$ is equivalent to $\mu_l^\star \rho(\hat{A}_l) < 1$, $\hat{A}_l := K_l^{\frac{1}{2}} Q_l^\top [\nabla_{xx}^2 \mathcal{H}(x, y)]^{-1} Q_l K_l^{\frac{1}{2}}$. This is a symmetric matrix, and hence, we can use Sylvester's criterion to enforce that all r_l leading principal minors of $I - \mu_l^\star \hat{A}_l$ (which are computable in algebraic closed-form) must be positive. A similar trick can be used if $K_l(x_l^\star, y)$ is negative definite.

- If $K_l(x_l^\star, y)$ is indefinite, then it is difficult to enforce the spectral inequality, and we settle for a relaxation which enforces a constraint on the sum of its r_l eigenvalues:

$$\mu_l^\star \operatorname{tr}(A_l) < r_l, \tag{30}$$

where again the trace can be computed in closed-form.

We close with some remarks on the assumptions of Proposition 1. The first assumption requires that the Hessian of the energy function at the equilibrium point x is positive definite. Note that it is satisfied whenever x is obtained via the energy minimization problem (6). Although using the power flow (4), (5) as a surrogate for the latter cannot guarantee this condition, we found that it was almost never violated in our experiments. With regards to the second assumption, the Taylor approximation (25) is guaranteed to be accurate only when $x_l^\star$ is near x. We provide empirical evidence in Section IV that the approximation is not severe.

D. Treatment of High-Dimensional Determinants

Using equations (10) and (11), the failure rate constraint (22) can be equivalently written as:

$$\lambda_l^\tau(y) = \mathsf{pf}_l^0(y) \cdot \mathsf{ef}_l(y) \cdot \left(1 + \frac{\tau}{\mathcal{H}(x_l^\star, y) - \mathcal{H}(\bar{x}, y)}\right) \le \lambda_l^{\lim}.$$

The formula for the prefactor $\mathsf{pf}_l^0(y)$ that appears in the above constraint involves determinants of (the typically large) $d \times d$ matrices, $\nabla_{xx}^2 \mathcal{H}(x, y)$ and $X_l(y)$, via (12) and (14) respectively. The presence of these determinants can slow down computation of the rate constraints and their gradients, and may result in ill-conditioning. Fortunately, (24) and (25) can be used to circumvent these issues. The following proposition uses the matrix-determinant lemma [50] to obtain a reformulation of the prefactor that avoids large determinants.

Proposition 2: Fix a line $l \in \mathcal{L}'$, and assume that $x, x_l^\star \in \mathbb{R}^d$ and $y \in \mathbb{R}^{m-d}$ satisfy the conditions of Proposition 1. If $x_l^\star$ and $\mu_l^\star \ge 0$ satisfy (26)–(28), then the prefactor and energy factor in (10) admit the following reformulation:

$$\mathsf{pf}_l^1(y) = \mathsf{pf}_l^0(y) \left(1 + \frac{2\tau}{\mu_l^\star \alpha_l}\right) \tag{31}$$

$$\mathsf{pf}_l^0(y) = \frac{(\mu_l^\star)^{3/2} \|\nabla_x \Theta_l(x_l^\star, y)\|_S^2}{\sqrt{2\pi\tau}\,(\alpha_l \det(W_l) + \beta_l)} \tag{32}$$

$$\mathsf{ef}_l(y) = \exp\left(-\frac{\mu_l^\star \alpha_l}{2\tau}\right), \tag{33}$$

where $W_l := I - \mu_l^\star A_l$ is a $r_l \times r_l$ matrix, A_l is as in Proposition 1, and the scalars $\alpha_l := (x_l^\star - x)^\top \nabla_x \Theta_l(x_l^\star, y)$ and $\beta_l := (x_l^\star - x)^\top Q_l(x_l^\star, y) \operatorname{adj}(W_l) K_l(x_l^\star, y) Q_l^\top(x_l^\star, y)(x_l^\star - x)$.

Proof: See [49, Appendix C] of the preprint. □

Unlike (12) and (14), (32) involves the determinant and adjugate of the small $r_l \times r_l$ matrix W_l. This allows efficient computation of the scalar β_l, and hence, of the overall failure rate and its gradient. Moreover, it avoids the ill-conditioning that plagues the original constraint, and allows expressing it in an algebraic optimization system.

E. Efficient Practical Implementation

We now highlight other key algorithmic enhancements that are necessary to improve the practical performance of the FP-ACOPF. First, since the absolute value of the failure rates $\lambda_l^\tau(y)$ can be very small, we found it beneficial to implement the rate constraint (22) in its log-form. Using (10), and (31)–(33), this is equivalent to:

$$\frac{3}{2}\log(\mu_l^\star) + \log\left(\|\nabla_x \Theta_l(x_l^\star, y)\|_S^2\right) - \frac{1}{2}\log\left(\alpha_l \det(W_l) + \beta_l\right)$$
$$+ \log\left(1 + \frac{2\tau}{\mu_l^\star \alpha_l}\right) - \frac{\mu_l^\star \alpha_l}{2\tau} \le \log(\sqrt{2\pi\tau}\lambda_l^{\lim}). \tag{34}$$

All intermediate terms involved in the left-hand side of (34) are computable in closed form, except for the matrix A_l which appears in the definition of $W_l := I - \mu_l^\star A_l$. A close examination of the formula for A_l in Proposition 1 reveals that this matrix can be efficiently constructed by first solving for the intermediate matrix $Z_l \in \mathbb{R}^{d \times r_l}$ as follows:

$$\nabla_{xx}^2 \mathcal{H}(x, y) \cdot Z_l = Q_l(x_l^\star, y), \tag{35}$$

and then setting $A_l = K_l(x_l^\star, y) Q_l^\top(x_l^\star, y) Z_l$. We therefore introduce explicit decision variables Z_l, along with (35) as constraints in the FP-ACOPF formulation. Notably, this allows the complete rate constraint (34), including all of its intermediate expressions, to be computed algebraically.

Implementing the rate constraint (22) for a particular line $l \in \mathcal{L}'$ thus involves introducing $dr_l + d + 1$ additional variables, namely $Z_l \in \mathbb{R}^{d \times r_l}, x_l^\star \in \mathbb{R}^d$ and $\mu_l^\star \ge 0$ and $dr_l + d + 3$ additional constraints, namely (26)–(28) and (34)–(35). In practice, the rate constraints (34) are binding for only a small fraction of critical lines. Therefore, adding $\mathcal{O}(dr_l)$ constraints for the remaining non-critical lines unnecessarily increases model complexity. To limit this growth, we add additional variables and constraints in an iterative manner, only for those lines $l \in \hat{\mathcal{L}}$ whose rate constraints (34) are found to be violated. Our algorithm can be described as follows.

Step 1: Set $\hat{\mathcal{L}} = \emptyset$.

Step 2: Define decision variables $x \in \mathbb{R}^d$, $y \in \mathbb{R}^{m-d}$, and $\{(x_l^\star, \mu_l^\star, Z_l) \in \mathbb{R}^d \times \mathbb{R}_{\ge 0} \times \mathbb{R}^{d \times r_l}\}$ for each $l \in \hat{\mathcal{L}}$. Compute an optimal solution $\hat{z}$ of the problem:
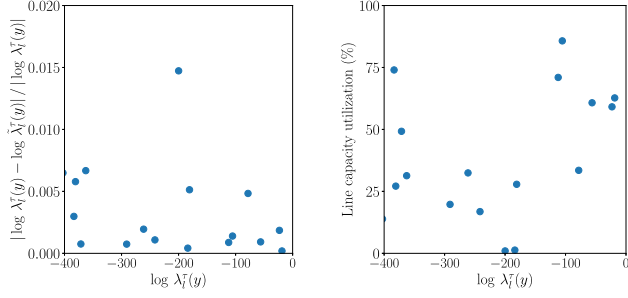
    minimize (17)

    subject to (4), (5), (18)−(21),

        (26)−(28), (34), (35) $\forall l \in \hat{\mathcal{L}}$.

Step 3: For all $l \in \mathcal{L}' \setminus \hat{\mathcal{L}}$ (possibly in parallel):

  1) Use $\hat{z}$ to compute candidate values of $x_l^\star, \mu_l^\star, Z_l$ by solving (26)–(28), and (35).

(a) Failure rate approximation error (b) Capacity utilization vs. failure rate

Fig. 1. Fig. 1(a) shows the relative log-error in approximating the true failure rate $\lambda_l^\tau(y)$ by $\tilde{\lambda}_l^\tau(y)$ in the rate-constrained model. Fig. 1(b) shows line capacity utilization as a function of its failure rate. Each dot corresponds to a transmission line and y is the classical (*i.e.*, $N-0$) ACOPF dispatch.

2) If inequality (34) is violated using $\hat{z}$ and the computed values of $x_l^\star, \mu_l^\star, Z_l$, update $\hat{\mathcal{L}} = \hat{\mathcal{L}} \cup \{l\}$ and save the computed values as warm-starts.

Step 4: If $\hat{\mathcal{L}}$ was updated, go to Step 2; else, stop and output $(\hat{x}, \hat{y})$ as the optimal FP-ACOPF solution.

IV. NUMERICAL SIMULATIONS

We now present results and insights obtained from several experiments.[5] The IEEE 118-bus test system with line limits $I_l^{\lim}$ obtained from PGLib [51], is used in all experiments except Section IV-B where we also include larger networks. We set branch and shunt conductances to zero to satisfy the lossless assumption,[6] remove transmission line taps for simplicity, and set the system matrices J and S that appear in (2) as per [43].[7] The numerical performance of the FP-ACOPF is largely unaffected by the choice of the noise parameter τ, the line failure threshold I_l^{trip}, and the network loading (p_d, q_d) levels. Therefore, Sections IV-A, IV-B, IV-C and IV-D use 'baseline' values of $\tau = 10^{-4}$, $I_l^{\text{trip}} = 1.10\,I_l^{\lim}$ and (p_d, q_d) values from MATPOWER [52]. Their values can, however, strongly affect cascading behavior; therefore, Section IV-E includes a discussion of cascading sensitivity to their values.

A. Analysis of Line Failure Rates

For a fixed dispatch point y, the true line failure rates $\lambda_l^\tau(y)$ are computed by solving problem (9) to obtain the most likely failure point $x_l^\star$ and its multiplier $\mu_l^\star$. However, the FP-ACOPF computes an estimate, which we denote as $\tilde{\lambda}_l^\tau(y)$, by solving (26)–(28) instead of (9). Fig. 1(a) shows the corresponding approximation error for the subset of lines with the highest failure rates. We observe that the relative log-error in the approximation is

TABLE II
SUMMARY OF NUMERICAL PERFORMANCE FOR THE 118-BUS TEST CASE

Formulation	$\max\limits_{l \in \mathcal{L}'} \lambda_l^\tau(y)$	Solve time (sec)	Cost diff (%)
$N-0$	7.8×10^{-9}	0.1	–
$N-1$	3.5×10^{-9}	73.0	0.04
$\lambda^{\lim} = 10^{-9}$	9.6×10^{-11}	2.7	0.00
$\lambda^{\lim} = 10^{-12}$	1.0×10^{-12}	2.8	0.00
$\lambda^{\lim} = 10^{-15}$	9.9×10^{-16}	2.3	0.01

less than 10^{-2}. Fig. 1(b) shows that lines with higher utilization of their flow capacity, defined as $\sqrt{\Theta_l(x,y)}/I_l^{\lim} \times 100\%$, do not necessarily have higher failure rates, thus illustrating that *the magnitude of line flow is not an effective surrogate for its failure rate*. For example the line with the second-largest capacity utilization has a negligible failure rate, whereas lines with the largest failure rates have capacity utilization close to 60%. For accuracy, we must thus work directly with the failure rates via (10).

B. Numerical and Economic Performance

Table II summarizes the computational time and generation cost of the failure probability-constrained, the $N-0$, and the $N-1$ security-constrained[8] dispatch points for the 118-bus test system. We observe that the FP-ACOPF solves within an order of magnitude of the time it takes to solve the $N-0$ ACOPF. On the other hand, the $N-1$ model is slower by more than two orders of magnitude, although this may be because all contingencies are included in a single optimization problem. In terms of generation cost, including the rate constraints results in a cost increase of less than 0.01% compared to the $N-0$ model, even with extremely stringent rate limits, without requiring any load shedding.

It is worth pointing out the higher cost of the $N-1$ model comes with a guarantee of $N-1$ reliability. Based on a contingency analysis of the FP-ACOPF dispatch points, we found that they remained feasible in 184 (out of 186) line contingencies, even though this was not explicitly imposed in their optimization. Further analysis found that they could provide feasible power flows under all contingencies, but failed only to satisfy the thermal line limits (21) in the two infeasible contingencies. In practice, operators may impose the failure rate constraints in each contingency of the $N-1$ ACOPF, thus ensuring both $N-1$ reliability and low failure probability.

Table III summarizes the computational times of the FP-ACOPF for various PGLib test cases of increasing size. Due to the distinct structures of these test networks, we impose failure rate limits that are a factor of 10 and 100 smaller than the maximum failure rate of their $N-0$ ACOPF dispatch point. For some of the cases, the imposed rate limits may be so stringent that the corresponding FP-ACOPF formulation can become infeasible. To recover a dispatch point in such cases, we first add load shedding (slack) variables $p_s, q_s \in \mathbb{R}^{n_b}$ to the power

TABLE III
SUMMARY OF NUMERICAL PERFORMANCE ACROSS PGLIB TEST CASES

Case	$N-0$	$\lambda^{\lim}=10^{-1}\times\max\limits_{l\in\mathcal{L}'}\lambda_l^{\tau}$ ($N-0$ dispatch)			$\lambda^{\lim}=10^{-2}\times\max\limits_{l\in\mathcal{L}'}\lambda_l^{\tau}$ ($N-0$ dispatch)		
	NLP time (sec)	NLP time (sec)	Eval time (sec)	Load shed (%)	NLP time (sec)	Eval time (sec)	Load shed (%)
118	0.1	0.7	0.7	0.00	0.6	1.1	0.00
200	0.1	2.8	2.0	0.79	8.2	2.3	2.10
300	0.3	3.3	4.7	0.00	2.9	5.8	0.00
500	0.6	5.2	17.7	0.00	10.5	18.4	0.00
1354	3.2	19.9	351.0	0.11	554.2	557.0	0.19

flow equations (4)–(5) as follows:

$$p_{net,i} + \sum_{j\in\mathcal{N}} B_{ij}V_iV_j\sin(\theta_i-\theta_j) = p_{s,i}, \;\; i\in\mathcal{N},$$

$$q_{net,i} - \sum_{j\in\mathcal{N}} B_{ij}V_iV_j\cos(\theta_i-\theta_j) = q_{s,i}, \;\; i\in\mathcal{N},$$

and then minimize them in the objective function (17) using a large quadratic penalty ($\phi = 10^6$) as follows:

$$\sum_{k\in\mathcal{G}} c_k(p_{g,k}) + \phi\sum_{i\in\mathcal{N}}\left(p_{s,i}^2 + q_{s,i}^2\right).$$

The column "Load shed" in Table III reports the absolute values of these variables (as a fraction of the total system load). The columns "NLP time" and "Eval time" report the total (wall-clock) time spent in Steps 2 and 3 of the algorithm described at the end of Section III, respectively.

We make the following observations from Table III. First, when compared to the $N-0$ ACOPF, the time spent to solve the FP-ACOPF formulation increases roughly by a factor of 10 for both loose rate limits (across all cases) as well as tighter rate limits (up to the 500-bus case). In contrast, for the tighter rate limit in the 1354-bus case, the time increases from roughly 3 sec to 560 sec. This increase can be explained by the observation that for this particular case, Step 2 was invoked 3 times, each solving a nonlinear problem with an increasing number of failure rate constraints. Second, the time spent in Step 3 to screen and evaluate violated failure rate constraints is an increasing function of the number of network lines. Third, a nonzero (but small ∼2%) amount of load must be shed in the 200-bus and 1354-bus cases to satisfy the imposed rate limits. However, this phenomenon is also true in the $N-1$ ACOPF where it is common to allow constraint violations to ensure feasibility [53]. Finally, we note that the "NLP time" can be reduced using decomposition techniques that are used to solve the $N-1$ ACOPF [54], whereas the "Eval time" can be reduced using a larger number of parallel computing processes. Indeed, although not shown in the table, increasing the number of processes from 32 to 64 decreases the evaluation time in the largest case from 557 sec to roughly 370 sec.

C. Cascading Failure Simulations of Different Dispatch Points

To compare the cascade potentials of the failure probability-constrained, the $N-0$, and the $N-1$ security-constrained ACOPF, we perform 1000 cascading failure simulations per dispatch point, using the Kinetic Monte Carlo (KMC) simulator

developed in [32]. In each simulation, the original network is randomly perturbed with two initial line contingencies[9].

For a given initial network state and dispatch point, KMC simulates cascading sequences of line failures by using the failure rate expression in (10) to identify the next likely line failure. Besides the initial dispatch point, KMC assumes that no operator interventions take place over the cascading time horizon, which we set to be one hour based on the typical time scale of operator actions [55], [56]. Moreover, it assumes that after a failure occurs, the system 're-equilibriates' to a new static equilibrium (satisfying the power flow equations) before the next failure occurs. That is, the system reaches a post-failure state (without triggering any intervening failures) in an amount of time that is small relative to the time until the next failure. This assumption is implicit in the majority of existing quasi-steady-state models [16] and numerical evidence in support of this assumption can be found in [32].

Each post-failure state satisfies the power flow equations and hence is a local minimizer (6) of the energy function, appropriately modified (via the admittance matrix) to account for the degraded network topology. Following each line failure, the slack bus is assumed to maintain power balance in the network, and buses that no longer connect to the slack bus (under the modified network topology) are disconnected, as well as buses whose voltage magnitudes fall below 0.9 p.u. The cumulative active demand p_d at disconnected buses serves as a measure of lost load. This enables fair comparison of different dispatch points, on the basis of time and severity until full system collapse or until the end of the cascading time horizon (one hour).

Fig. 2 shows the distributions of the number of failed lines and the amount of load shed for each dispatch point, in the form of survival functions, for baseline values of $\tau = 10^{-4}$, $I_l^{\text{trip}} = 1.10\, I_l^{\lim}$ and nominal MATPOWER (p_d, q_d). Their averages and standard deviations are summarized in Table IV.

Fig. 2(a) shows the probability that the total number of line failures is at least as large as a given value, as a proportion of all 1000 simulations. Here, we observe that all survival functions decrease as the number of line failures increases, and this decrease accelerates when the number of failures is large, indicating that the frequency of large blackouts decreases rapidly. Moreover, unlike the $N-1$ model, the FP-ACOPF dispatch can

---

[9]Initial contingencies are sampled as follows: *(i)* choose the first contingency uniformly at random from all available lines, *(ii)* sample a realization d from a Zipf distribution with parameters $s = 3$ and $N = 10$, *(iii)* group all lines that are topological distance d from the first initial contingency, and *(iv)* uniformly sample the second contingency from this group.

(a) Distribution of number of failed lines

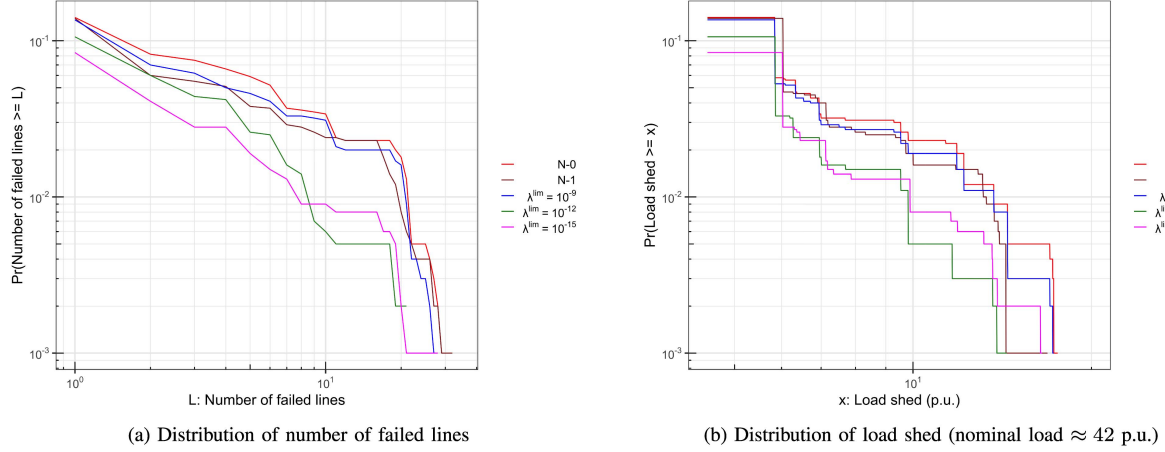(b) Distribution of load shed (nominal load $\approx$ 42 p.u.)

Fig. 2. Distribution of number of failed lines and load shed using the KMC simulator, considering all simulations.

TABLE IV
AVERAGE NUMBER OF FAILED LINES AND LOAD SHED (STANDARD DEVIATION IN PARENTHESES) OVER ALL KMC SIMULATIONS

Formulation	Number of failed lines	Load shed (p.u.)
$N - 0$	0.88 (3.52)	1.06 (2.85)
$N - 1$	0.71 (3.17)	1.02 (2.71)
$\lambda^{\mathrm{lim}} = 10^{-9}$	0.75 (3.21)	1.00 (2.72)
$\lambda^{\mathrm{lim}} = 10^{-12}$	0.39 (1.78)	0.71 (2.15)
$\lambda^{\mathrm{lim}} = 10^{-15}$	0.33 (1.92)	0.60 (2.11)

be tuned via the rate limits λ^{lim}, and for sufficiently stringent $\lambda^{\mathrm{lim}} \in \{10^{-12}, 10^{-15}\}$, it can reduce the average number of line failures by more than 50% when compared to the $N - 0$ and $N - 1$ models, as seen in Table IV.

Fig. 2(b) shows the distribution of load shed for each dispatch point as a probability of the load shed being at least as large as a given value, over all 1000 simulations. We find that the $N - 1$ security-constrained ACOPF performs as well as the FP-ACOPF in reducing the probability of very large load shedding, which is not surprising since the FP-ACOPF formulation explicitly constrains the probability of only the first line failure. However, unlike the former, the latter is also able to reduce the probability of small-to-intermediate demand losses, leading to a more than 30% reduction in average load shed for $\lambda^{\mathrm{lim}} \in \{10^{-12}, 10^{-15}\}$.

Fig. 3 shows the average time until a certain number of lines have failed. We find that, by delaying the system's first failure time, the FP-ACOPF is able to successfully delay–and hence, prevent–subsequent line failures which could lead to system collapse. For example, in the first 15 minutes after the first failure, FP-ACOPF with $\lambda^{\mathrm{lim}} = 10^{-15}$ shows about two times fewer lines failures compared to the $N - 1$ dispatch.

D. Validation and Benchmarking of Cascade Statistics

The goal of this section is to validate the simulation statistics from the previous section. We proceed to do this in three ways: 1) qualitative comparisons with historical cascades, 2) quantitative comparisons with an alternate AC-based cascade simulator that is based on an entirely different methodology, and finally, 3) validation of some of the key assumptions.
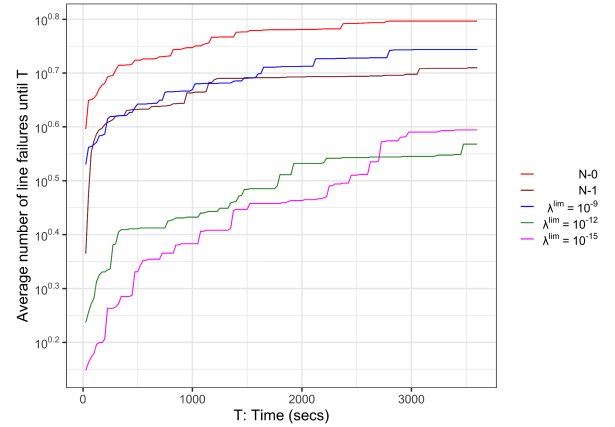

Fig. 3. Average timeline of line failures (averaged across simulations where at least one failure occurred).

1) Historical Cascades: The number of line failures and load shed in historical outage data are known to follow a heavy-tailed or Zipf distribution [16], [56]. This pattern can also be seen in Figs. 2(a) and (b); specifically, each survival function decreases roughly linearly in these log-log plots, which is characteristic of a power law distribution. Although this linear behavior is more pronounced in Fig. 2(a), it is also evident for an intermediate range of values in Fig. 2(b). In addition, the distinctive 'knee' in Fig. 2(a) has been observed in other cascading failure models as well [16].

Historical cascades also show a distinctive acceleration during their initial stages [56], [57]. This is somewhat evident in Fig. 3 where the $N - 1$ and FP-ACOPF dispatch points with $\lambda^{\mathrm{lim}} \in \{10^{-12}, 10^{-15}\}$ exhibit a rapid succession of failures in the beginning but then appear to taper off.

We note that even though our simulations cannot be directly compared to historical data (since they correspond to different systems and initiating events), the above observations indicate that the KMC cascading failure model captures important patterns that are representative of historical outages. A more thorough validation and comparison against real outage data can be found in [32].
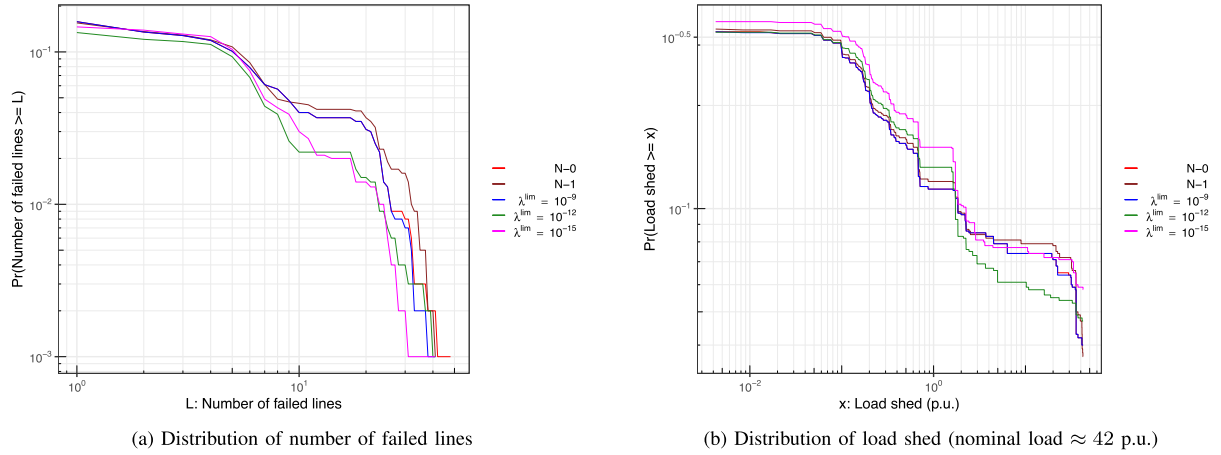
(a) Distribution of number of failed lines

(b) Distribution of load shed (nominal load ≈ 42 p.u.)

Fig. 4. Distribution of number of failed lines and load shed using the AC-CFM simulator, considering all simulations.

2) Comparison With an Alternate Cascading Failure Simulator: We now compare the cascade potentials of the various dispatch points using AC-CFM, an existing AC-based quasi-steady-state simulator developed in [57]. We perform 1000 simulations per dispatch point, where the original network (which we adjusted to be lossless) is randomly perturbed with the same two initial line contingencies as in Section IV-C. In contrast to the KMC simulator, AC-CFM uses a set of deterministic rules to model various protection mechanisms. In addition to line trips (which we adjusted per (1) with $I_l^{\text{trip}} = 1.10$) and under-voltage load shedding—mechanisms modeled by KMC—it also models under-frequency load (and over-frequency generator) shedding, and under- and over-excitation limiters.

Fig. 4(a) and (b) show the distribution of the number of failed lines and total load shed for each dispatch point, respectively, as now predicted by AC-CFM. For small values of the number of failed lines in Fig. 4(a), the survival functions have a similar shape but begin at different levels.

Examining the intercept of Fig. 4(a) shows that, relative to $N - 0$, the dispatch points corresponding to FP-ACOPF with $\lambda^{\text{lim}} \in \{10^{-12}, 10^{-15}\}$ demonstrate a 15% and 8% reduction, respectively, in the number of observed cascades with at least one failure, compared to a 2% reduction demonstrated by $N - 1$. In addition, Fig. 4(a) indicates that the FP-ACOPF dispatch points have lower failure probabilities for intermediate-to-large numbers of failed lines. Though no full system failures were observed, this pattern can also be found in Fig. 4(b) for load shed greater than 1 p.u, indicating that the FP-ACOPF dispatch points can also reduce the frequency of intermediate-to-large demand losses at the expense of possibly inducing a higher proportion of smaller losses.

Finally, Table V shows that across all simulations, the best FP-ACOPF dispatch point reduced the average number of line failures by more than 25% (and had lower variability) and reduced the average amount of load lost by more than 10%, compared to $N - 0$ ACOPF. In contrast, the $N - 1$ dispatch point *increased* the number of failures by 10% (with larger variability) and *increased* the amount of load lost by 8%.

Additional experiments across a range of AC-CFM parameters, including which protection mechanisms are modeled along

TABLE V
AVERAGE NUMBER OF FAILED LINES AND LOAD SHED (STANDARD DEVIATION IN PARENTHESES) OVER ALL AC-CFM SIMULATIONS

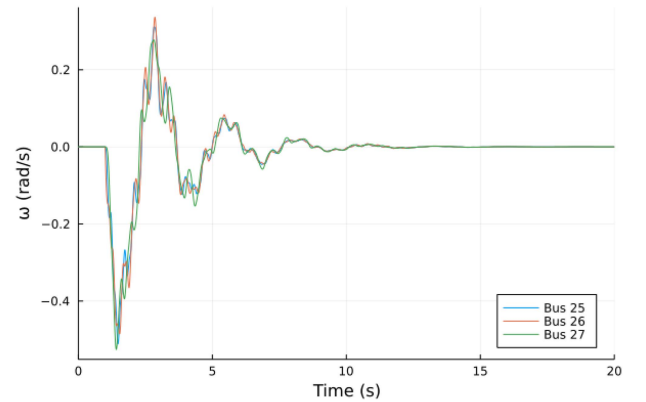| Formulation	Number of failed lines	Load shed (p.u.)
$N - 0$	1.48 (5.08)	2.94 (9.94)
$N - 1$	1.63 (5.62)	3.17 (10.41)
$\lambda^{\text{lim}} = 10^{-9}$	1.46 (4.92)	2.93 (9.93)
$\lambda^{\text{lim}} = 10^{-12}$	1.08 (3.96)	2.56 (9.46)
$\lambda^{\text{lim}} = 10^{-15}$	1.15 (3.80)	3.21 (10.58)

Fig. 5. Representative generator angular velocities (relative to system frequency) following the outage of line 96 at time $t = 1$ second.

with their thresholds, point to the fact that it is difficult to directly compare absolute estimates of failure probability with the KMC simulator. This is not surprising since they are based on different assumptions, and similar observations across a large range of simulators have been made in [16]. Nevertheless, we find it encouraging that the various dispatch points exhibit similar patterns (in particular in terms of their relative ordering with respect to cascade sizes) even when evaluated using a cascading methodology, AC-CFM, that is different from our KMC approach, and that FP-ACOPF can lead to lower failure probabilities under certain conditions.

3) Validation of Assumptions via Transient Stability Analysis: The KMC and AC-CFM simulators are both based on
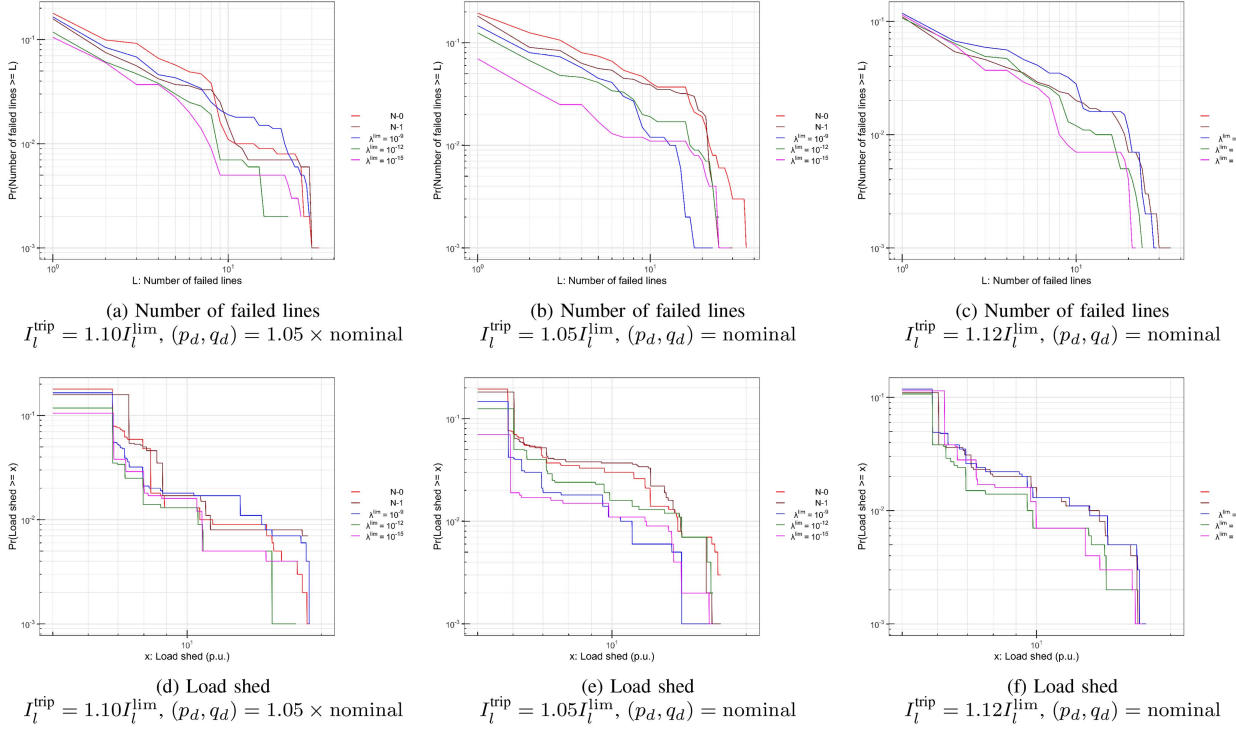
(a) Number of failed lines
$I_l^{\text{trip}} = 1.10 I_l^{\text{lim}}$, $(p_d, q_d) = 1.05 \times$ nominal

(b) Number of failed lines
$I_l^{\text{trip}} = 1.05 I_l^{\text{lim}}$, $(p_d, q_d) = $ nominal

(c) Number of failed lines
$I_l^{\text{trip}} = 1.12 I_l^{\text{lim}}$, $(p_d, q_d) = $ nominal

(d) Load shed
$I_l^{\text{trip}} = 1.10 I_l^{\text{lim}}$, $(p_d, q_d) = 1.05 \times$ nominal

(e) Load shed
$I_l^{\text{trip}} = 1.05 I_l^{\text{lim}}$, $(p_d, q_d) = $ nominal

(f) Load shed
$I_l^{\text{trip}} = 1.12 I_l^{\text{lim}}$, $(p_d, q_d) = $ nominal

Fig. 6.   Distribution of number of failed lines and load shed for various values of line failure threshold I_l^{trip} and system loads (p_d, q_d) using KMC.

quasi-steady-state methodologies. They implicitly assume not only that the system stabilizes to a new static equilibrium (satisfying the power flow equations), but also that this happens in a time frame that is small relative to the time until the next failure. However, these assumptions can fail because of system dynamics that are not modeled in these simulators. For example, generator desynchronization induced by an outage may cause the system to become unstable or to stabilize to another state than the one predicted by the simulators. To validate these assumptions, we use an existing tool [58] to perform stability analysis and verify: *(i)* whether the FP-ACOPF dispatch point is a stable operating point; *(ii)* whether the settling time following a failure (if the system remains stable) is relatively smaller than the expected next failure time as predicted by (16); and *(iii)* whether the new stable operating state is same as that predicted by (6).

To that end, we consider the FP-ACOPF dispatch point $(\lambda^{\text{lim}} = 10^{-9})$ and simulate system dynamics (using classical swing equations for generators and assuming constant active and reactive power outputs at load buses) for multiple fault scenarios, each corresponding to a line outage. We found that in roughly 89% of the (134 total) simulations, the system oscillations were damped to stable values in 27.3 seconds on average (with a maximum and standard deviation of 34.6 and 3.1 seconds, respectively), whereas only 11% of simulations were reported as being potentially unstable. Fig. 5 illustrates the typical response of representative system variables following a line outage where the system remained stable.

We also observed that, in all of the stable scenarios, the system stabilized to the same equilibrium state as the one predicted by (6) (appropriately modified to account for the failed line). On

the other hand, among the unstable scenarios where (6) also converged to a solution, only roughly half of them turned out to be dynamically unstable. These experiments indicate that the aforementioned quasi-steady-state assumptions do hold for the majority of line outages, although this is likely to be true only up to a certain point where the system has not degraded significantly. At the same time, they also highlight the challenges and need to model system dynamics more accurately in the context of cascading failures [18], [19].

E. Sensitivity Analysis

To study the sensitivity of our results to the choice of line failure threshold I_l^{trip} and loads (p_d, q_d), we repeated our analysis by varying them as follows: $I_l^{\text{trip}} \in \{1.05, 1.10, 1.12\} \times I_l^{\text{lim}}$ and $(p_d, q_d) \in \{1.0, 1.05\} \times$ nominal MATPOWER values, where the lower I_l^{trip} and higher (p_d, q_d) values were chosen so as to stress the system towards larger failures, and the higher I_l^{trip} value was chosen to elicit the opposite effect. Fig. 6 summarizes the KMC cascading failure simulations of the various dispatch points under these settings.

First, we observe that the distribution of the number of failed lines (Fig. 6(a)–(c)) for each dispatch point is an approximate power law in the range $L \in [1, 10]$, followed by a distinctive 'knee' similar to Fig. 2(a), indicating a significant decrease in the probability of large line failures. The distribution of the load shed (Fig. 6(d)–(f)) is also an approximate power law, although for a smaller range of load shedding values. Second, we observe that the FP-ACOPF generally outperforms the $N - 0$ and $N - 1$ ACOPF, both in terms of reduced probabilities of large line failures as well as load shedding. Specifically, for

higher load levels (Fig. 6(a) and (d)), we find that the FP-ACOPF dispatches with $\lambda^{\text{lim}} \in \{10^{-12}, 10^{-15}\}$ perform better than the other dispatch points; for lower line failure thresholds (Fig. 6(b) and (e)), the FP-ACOPF dispatches with $\lambda^{\text{lim}} \in \{10^{-9}, 10^{-15}\}$ perform significantly better, lowering failure probabilities by a factor of three on average, compared to both $N - 0$ and $N - 1$; and for higher line failure thresholds (Fig. 6(c) and (f)), the FP-ACOPF dispatch points with $\lambda^{\text{lim}} \in \{10^{-12}, 10^{-15}\}$ have the lowest failure probabilities. In this setting, we note that the survival functions of the $N - 0$ dispatch coincide precisely with that of $\lambda^{\text{lim}} = 10^{-9}$. In summary, we find that the improvement in performance in regards to cascading risk of FP-ACOPF relative to $N - 0$ and $N - 1$ ACOPF appears to hold for a range or parameter choices beyond the ones derived from PGLib [51] and MATPOWER [52] that we used for the assessments in the preceding subsections.

V. CONCLUSION AND FUTURE WORK

This paper takes the first steps towards quantifying and proactively reducing failure risk–and implicitly, cascading risk–in operational dispatch. Our proposed FP-ACOPF generalizes the standard ACOPF formulation by constraining the probability of component failures due to automatic relay trips, via analytic functions of the system state and stochastic load fluctuations derived from our previous work [32]. By using techniques from bilevel optimization and numerical linear algebra, we reformulated the probability constraints entirely algebraically allowing their solution using off-the-shelf nonlinear solvers. We empirically showed that constraining failure probabilities– which is equivalent to increasing the system's expected first failure time–is a safe approximation and effective surrogate for constraining the probability of a cascading failure sequence starting from any single component outage. Simulation outputs from two different cascading failure simulators provide evidence that FP-ACOPF can significantly outperform classical $N - 0$ and $N - 1$ security-constrained ACOPF models, in terms of reducing the probability of line outages and load shedding, without incurring significantly larger computational or economic costs.

We envision future work along several directions. From a methodological viewpoint, we need to extend the failure probability model and the corresponding optimization model to lossy networks. From a modeling viewpoint, more general multiple-component failure rate constraints, as well as extensions that combine classical $N - 1$ models with failure rate constraints need to be investigated. From a practical viewpoint, the model parameters (e.g., τ) need to be calibrated and the model itself needs to be further validated and compared against other established cascading simulators on real network data. We believe these extensions can open up several other use cases for our method including long-term planning decisions, line capacity allocations, and contingency screening.

REFERENCES

[1] R. Baldick et al., "Initial review of methods for cascading failure analysis in electric power transmission systems IEEE PES CAMS task force on understanding, prediction, mitigation and restoration of cascading failures," in Proc. IEEE Power Energy Soc. Gen. Meeting, 2008, pp. 1–8.

[2] M. Papic et al., "Survey of tools for risk assessment of cascading outages," in Proc. IEEE Power Energy Soc. Gen. Meeting, 2011, pp. 1–9.

[3] 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.

[4] H. Haes Alhelou, M. E. Hamedani-Golshan, T. C. Njenda, and P. Siano, "A survey on power system blackout and cascading events: Research motivations and challenges," Energies, vol. 12, no. 4, 2019, Art. no. 682.

[5] P. Hines and S. Blumsack, "A centrality measure for electrical networks," in Proc. 41st Annu. Hawaii Int. Conf. Syst. Sci., 2008, pp. 185–185.

[6] C. D. Brummitt, R. M. D'Souza, and E. A. Leicht, "Suppressing cascades of load in interdependent networks," Proc. Nat. Acad. Sci., vol. 109, no. 12, pp. E680–E689, 2012.

[7] I. Dobson, B. A. Carreras, and D. E. Newman, "A loading-dependent model of probabilistic cascading failure," Probability Eng. Informational Sci., vol. 19, no. 1, 2005, Art. no. 15.

[8] 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.

[9] X. Yu and C. Singh, "A practical approach for integrated power system vulnerability analysis with protection failures," IEEE Trans. Power Syst., vol. 19, no. 4, pp. 1811–1820, Nov. 2004.

[10] 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. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061505000232

[11] S. Mei, Y. Ni, G. Wang, and S. Wu, "A study of self-organized criticality of power system under cascading failures based on AC-OPF with voltage stability margin," IEEE Trans. Power Syst., vol. 23, no. 4, pp. 1719–1726, Nov. 2008.

[12] M. Anghel, K. A. Werley, and A. E. Motter, "Stochastic model for power grid dynamics," in Proc. 40th Annu. Hawaii Int. Conf. Syst. Sci., 2007, pp. 113–113.

[13] Y. Yang et al., "A dynamic cascading failure model integrating relay protection in power grid," in Proc. IEEE 16th Int. Conf. Automat. Sci. Eng., 2020, pp. 1331–1336.

[14] D. P. Nedic, I. Dobson, D. S. Kirschen, B. A. Carreras, and V. E. Lynch, "Criticality in a cascading failure blackout model," Int. J. Elect. Power Energy Syst., vol. 28, no. 9, pp. 627–633, 2006.

[15] 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.

[16] P. Henneaux et al., "Benchmarking Quasi-steady state cascading outage analysis methodologies," in Proc. IEEE Int. Conf. Probabilistic Methods Appl. to Power Syst., 2018, pp. 1–6.

[17] C. L. DeMarco, "A phase transition model for cascading network failure," IEEE Control Syst. Mag., vol. 21, no. 6, pp. 40–51, Dec. 2001.

[18] J. Song, E. Cotilla-Sanchez, G. Ghanavati, and P. D. Hines, "Dynamic modeling of cascading failure in power systems," IEEE Trans. Power Syst., vol. 31, no. 3, pp. 2085–2095, May 2016.

[19] B. Schäfer, D. Witthaut, M. Timme, and V. Latora, "Dynamically induced cascading failures in power grids," Nature Commun., vol. 9, no. 1, pp. 1–13, 2018.

[20] F. Dörfler and F. Bullo, "Synchronization and transient stability in power networks and non-uniform kuramoto oscillators," in Proc. Amer. Control Conf., 2010, pp. 930–937.

[21] T. Nishikawa and A. E. Motter, "Comparative analysis of existing models for power-grid synchronization," New J. Phys., vol. 17, no. 1, Jan. 2015, Art. no. 015012. [Online]. Available: https://doi.org/10.1088/1367-2630/17/1/015012

[22] M. Rohden, D. Jung, S. Tamrakar, and S. Kettemann, "Cascading failures in AC electricity grids," Phys. Rev. E, vol. 94, Sep. 2016, Art. no. 032209. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.94.032209

[23] D. Fabozzi and T. Van Cutsem, "Simplified time-domain simulation of detailed long-term dynamic models," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2009, pp. 1–8.

[24] S. K. Khaitan, C. Fu, and J. McCalley, "Fast parallelized algorithms for on-line extended-term dynamic cascading analysis," in *Proc. IEEE/PES Power Syst. Conf. Expo.*, 2009, pp. 1–7.

[25] S. Abhyankar and A. J. Flueck, "Real-time power system dynamics simulation using a parallel block-jacobi preconditioned newton-gmres scheme," in *Proc. High Perform. Comput., Netw. Storage Anal.*, 2012, pp. 299–305.

[26] D. Bienstock, "Optimal control of cascading power grid failures," in *Proc. 50th IEEE Conf. Decis. Control Eur. Control Conf.*, 2011, pp. 2166–2173.

[27] A. Esmaeilian and M. Kezunovic, "Prevention of power grid blackouts using intentional islanding scheme," *IEEE Trans. Ind. Appl.*, vol. 53, no. 1, pp. 622–629, Jan./Feb. 2017.

[28] L. Guo, C. Liang, A. Zocca, S. H. Low, and A. Wierman, "Localization & mitigation of cascading failures in power systems, Part III: Real-time mitigation," 2020, *Arxiv:2005.11319*.

[29] H. Wang and J. S. Thorp, "Optimal locations for protection system enhancement: A simulation of cascading outages," *IEEE Trans. Power Del.*, vol. 16, no. 4, pp. 528–533, Oct. 2001.

[30] R. C. Hardiman, M. Kumbale, and Y. V. Makarov, "An advanced tool for analyzing multiple cascading failures," in *Proc. Int. Conf. Probabilistic Methods Appl. Power Syst.*, 2004, pp. 629–634.

[31] 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.

[32] J. Roth, D. A. Barajas-Solano, P. Stinis, J. Weare, and M. Anitescu, "A kinetic monte carlo approach for simulating cascading transmission line failure," *Multiscale Model. Simul.*, vol. 19, no. 1, pp. 208–241, 2021. [Online]. Available: https://doi.org/10.1137/19M1306865

[33] H. Zheng and C. L. DeMarco, "A bi-stable branch model for energy-based cascading failure analysis in power systems," in *Proc. North Amer. Power Symp.*, 2010, pp. 1–7.

[34] B. Stott, J. Jardim, and O. Alsaç, "DC power flow revisited," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009.

[35] C. DeMarco and A. Bergen, "A security measure for random load disturbances in nonlinear power system models," *IEEE Trans. Circuits Syst.*, vol. 34, no. 12, pp. 1546–1557, Dec. 1987.

[36] A. R. Bergen and D. J. Hill, "A structure preserving model for power system stability analysis," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 1, pp. 25–35, Jan. 1981.

[37] C. L. DeMarco and A. R. Bergen, "Application of singular perturbation techniques to power system transient stability analysis," in *Proc. Int. Symp. Circuits Syst.*, 1984, pp. 597–601.

[38] C. L. DeMarco and A. R. Bergen, "Small noise in power system dynamics and its relation to security assessment," in *Proc. Amer. Control Conf.*, 1985, pp. 559–560.

[39] M. Pai, *Energy Function Analysis for Power System Stability*. Berlin, Germany: Springer Science & Business Media, 2012.

[40] P. W. Sauer and M. A. Pai, *Power System Dynamics and Stability*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 1997.

[41] A. Arif, Z. Wang, J. Wang, B. Mather, H. Bashualdo, and D. Zhao, "Load modeling–A review," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5986–5999, Nov. 2018.

[42] C. L. DeMarco and T. J. Overbye, "An energy based security measure for assessing vulnerability to voltage collapse," *IEEE Trans. Power Syst.*, vol. 5, no. 2, pp. 419–427, May 1990.

[43] H. Zheng and C. L. DeMarco, "A new dynamic performance model of motor stalling and FIDVR for smart grid monitoring/planning," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 1989–1996, Jul. 2016.

[44] C. DeMarco, "A new method of constructing lyapunov functions for power systems," in *Proc. IEEE Int. Symp. Circuits Syst.*, 1988, vol. 1, pp. 905–908.

[45] J. Sun, H. Zheng, Y. Chai, Y. Hu, K. Zhang, and Z. Zhu, "A direct method for power system corrective control to relieve current violation in transient with UPFCs by barrier functions," *Int. J. Elect. Power Energy Syst.*, vol. 78, pp. 626–636, 2016.

[46] A. Fouad and V. Vittal, "The transient energy function method," *Int. J. Elect. Power Energy Syst.*, vol. 10, no. 4, pp. 233–246, 1988.

[47] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, vol. 260, 3rd ed. Berlin, Germany: Springer-Verlag, 2012.

[48] P. Collet, S. Martínez, and J. San Martín, *Quasi-Stationary Distributions: General Results*. Berlin, Germany: Springer, 2013, pp. 17–29.

[49] A. Subramanyam, J. Roth, A. Lam, and M. Anitescu, "Failure probability constrained AC optimal power flow," 2020, *Arxiv:2011.02453*.

[50] D. A. Harville, *Matrix Algebra From a Statistician's Perspective*. New York, NY, USA: Springer, 2008.

[51] S. Babaeinejadsarookolaee *et al.*, "The power grid library for benchmarking AC optimal power flow algorithms (v21.07)," *arXiv e-prints*, Aug. 2019.

[52] R. D. Zimmerman, C. E. Murillo-Sanchez, 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] *ARPA-E Grid Optim. Competition*, "Challenge 1 problem formulation," 2019, Accessed: Dec. 8, 2021. [Online]. Available: https://gocompetition.energy.gov/challenges/challenge-1/formulation

[54] N. Chiang, C. G. Petra, and V. M. Zavala, "Structured nonconvex optimization of large-scale energy systems using PIPS-NLP," in *Proc. Power Syst. Comput. Conf.*, 2014, pp. 1–7.

[55] H. Ren and I. Dobson, "Using transmission line outage data to estimate cascading failure propagation in an electric power system," *IEEE Trans. Circuits Syst. II: Exp. Briefs*, vol. 55, no. 9, pp. 927–931, Sep. 2008.

[56] I. Dobson, "Estimating the propagation and extent of cascading line outages from utility data with a branching process," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2146–2155, Nov. 2012.

[57] M. Noebels, R. Preece, and M. Panteli, "AC cascading failure model for resilience analysis in power networks," *IEEE Syst. J.*, to be published, doi: 10.1109/JSYST.2020.3037400.

[58] A. Plietzsch *et al.*, "Powerdynamics. jl–an experimentally validated open-source package for the dynamical analysis of power grids," 2021, doi: 10.1016/j.softx.2021.100861.

Anirudh Subramanyam (Member, IEEE) received the bachelor's degree in chemical engineering from the Indian Institute of Technology, Bombay, India, and the Ph.D. degree in chemical engineering from Carnegie Mellon University. He is a Postdoctoral Researcher with Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, IL, USA. His research interests include computational methods for nonlinear and discrete optimization under uncertainty with applications in energy, transportation, and process systems.

Jacob Roth received the M.S. degree from the University of Chicago, Chicago, IL, USA. He is currently working toward the Ph.D. degree with the Industrial and systems engineering department, University of Minnesota, Minneapolis, MN, USA. He was a Predoctoral Appointee with Argonne National Laboratory, Lemont, IL, USA.

Albert Lam received the B.S. degree from the University of New South Wales, Sydney, NSW, Australia, and the M.S. degree from the University of Chicago, Chicago, IL, USA. He is a Predoctoral Appointee with Argonne National Laboratory, Lemont, IL, USA.

Mihai Anitescu (Member, IEEE) received the engineer diploma in electrical engineering from the Polytechnic University of Bucharest, Bucharest, Romania, in 1992, and the Ph.D. in applied mathematical and computational sciences from the University of Iowa, Iowa City, IA, USA, in 1997. He is a Senior Computational Mathematician with Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, IL, USA, and a Professor with the Department of Statistics, University of Chicago, Chicago, IL, USA. He specializes in the areas of numerical optimization, computational science, numerical analysis, and uncertainty quantification in which he has authored or coauthored more than 100 papers in scholarly journals and book chapters. He has been recognized for his work in applied mathematics by his selection as a SIAM Fellow in 2019.