Analyzing large frequency disruptions in power systems using large deviations theory

Brendan Patch
CWI
Amsterdam, The Netherlands
brendan.patch@cwi.nl

Bert Zwart
CWI, IEEE Member
Amsterdam, The Netherlands
bert.zwart@cwi.nl

Abstract—We propose a method for determining the most likely cause, in terms of conventional generator outages and renewable fluctuations, of power system frequency reaching a predetermined level that is deemed unacceptable to the system operator. Our parsimonious model of system frequency incorporates primary and secondary control mechanisms, and supposes that conventional outages occur according to a Poisson process and renewable fluctuations follow a diffusion process.

We utilize a large deviations theory based approach that outputs the most likely cause of a large excursion of frequency from its desired level. These results yield the insight that current levels of renewable power generation do not significantly increase system vulnerability in terms of frequency deviations relative to conventional failures. However, for a large range of model parameters it is possible that such vulnerabilities may arise as renewable penetration increases.

Index Terms—energy systems, power system frequency, renewable energy, stochastic processes, Ornstein–Uhlenbeck process, Poisson process, large deviations theory, Schilder's theorem

I. INTRODUCTION

It is advocated that substantial increase in output from wind, water, and solar energy sources is required to combat climate change [1]. As a consequence of the highly stochastic nature of the variability in the supply of these renewable energy sources [2], their integration into the existing power system is highly challenging.

In particular, the evolution of system frequency may no longer behave as predicted by traditional deterministic models. Moreover, widespread methods of assessing system stability, where worst case scenarios are considered in terms of combinations of conventional outages, may no longer apply [3].

Decisions based on these methods often only ensure that systems are able to withstand a single generator failure, but do not ensure robustness to multiple failures occurring within a short amount of time of each other. In the future the combination of a single generator failure and renewable fluctuations may be an issue as well, as power output of a wind farm can drop dramatically within seconds [2].

Since system frequency is the primary signal used by control mechanisms in modern power systems, it is crucial to understand how it evolves in a system with a high level of renewable energy penetration. In particular, it is useful to understand where the system's vulnerability is located, which translate into the most likely causes of frequency dips.

This motivates us to consider new methods for determining how significant system frequency dips occur. We consider a parsimonious model of aggregate system frequency that incorporates primary and secondary control mechanisms. We assume that conventional generator outages occur according to a Poisson process (PP) with rate λ and renewable generation fluctuates according to an Ornstein–Uhlenbeck (OU) diffusion with diffusion parameter σ (our methods generalize to more detailed models). Together, these processes can result in a loss of power generation that causes frequency to trend downwards.

We primarily focus on determining the most likely cause of a significant frequency dip, in terms of conventional outages and renewable fluctuations, as a function of the parameters λ and σ. We consider a setting where such excursions occur infrequently and as such constitute rare events. We propose to use large deviations (LD) theory [4] to formulate and solve an optimization problem. Its solution indicates how many conventional outages and how much renewable fluctuation most likely causes a frequency deviation.

Technically underpinning our work are LD results for diffusions [4], and a novel scaling of the PP governing conventional outages. The contraction principle is used to map the resulting LD principles into the equation governing power system frequency. The resulting approximation of rare event probabilities are computed using calculus of variations, methods for efficiently evaluating matrix exponentials, and numerical optimization techniques.

Our approach is markedly different from Monte Carlo simulation. In rare event settings, standard simulation methods are unreliable and must be carefully reformulated to be effective (see, e.g., [5]). Such implementations can be difficult to achieve in our dynamic continuous-time setting. Our work complements other studies investigating the most probable failure modes of modern power systems. For example, [6] uses instanton theory, and [7] formulates the model as a Lur'e system to simplify analysis. The LD approach that we propose has already been shown to be effective in related studies [8], [9] and is, to our knowledge, a novel method for studying anomalous fluctuations in power system frequency. Our LD approach gives more explicit answers than the instanton approach, as the associated scaling procedure washes out the details that do not contribute significantly to the rare event of interest.

Our study includes an extensive numerical investigation that demonstrates how our theoretical results can be utilized and provides key qualitative insights. We provide evidence that current levels of renewable power generation do not increase system vulnerability in terms of frequency deviations relative to conventional failures, but that it is possible that such vulnerabilities may arise as renewable penetration increases.

We also show how our method can be used to examine the influence of model parameters. For example characterizing how perturbing inertia does not cause the probability of system failure to change much for a large range of values of inertia, but that such perturbations can have large effects for some specific values of inertia.

The remainder of this paper is organized as follows. Sec. II presents our model of aggregate power system frequency. Sec. III provides our key theoretical result, which is used in Sec. IV to analyze the model. Proofs are contained in Sec. V. We conclude in Sec. VI with an outlook towards future research opportunities.

II. System Model

The aggregate power system frequency model we consider incorporates primary and secondary control, and is closely related to the single area models presented in [10]. The model captures stochastic load dynamics in terms of an OU process $(P(t))_{t\in[0,T]}$, which has mean reversal parameter ρ, diffusion parameter σ and stochasticity driven by the zero-mean-unit-variance Brownian motion (BM) $(W(t))_{t\in[0,T]}$, and an independent PP $(N(t))_{t\in[0,T]}$ with rate λ. The process P models deviations from net aggregate nominal power generation and consumption in the power system at buses housing stochastic renewable power generators or unpredictable demand side loads. The process N records the times of conventional power generator failures, allowing the deviation in aggregate power generation from conventional (non-renewable) sources to be quantified.

Throughout this paper, $\dot{f}$ and $\ddot{f}$ represent $\mathrm{d}f/\mathrm{d}t$ and $\mathrm{d}^2 f/\mathrm{d}t^2$, and $f^{(i)}$ represents higher order derivatives $\mathrm{d}^i f/\mathrm{d}t^i$. We assume that a system is operating in equilibrium at time zero and that we are interested in how frequency responds to conventional power disturbances and renewable energy fluctuations. Therefore, we model deviations from nominal values during a planning interval of $[0,T]$ using the equations

$$\mu\,\ddot{\theta}(t) + \alpha\,\dot{\theta}(t) = P(t) - \delta\,N(t) - \beta\,\theta(t)$$
$$\dot{P}(t) = -\rho\,P(t) + \sigma\,\dot{W}(t)\,, \tag{1}$$

where $\dot{\theta}$ tracks the per unit frequency deviation of the center of inertia of a power system. We assume $\theta(0)=\dot{\theta}(0)=P(0)=0$. Each conventional failure results in δ units of power being instantly removed from the system. The evolution of the system is governed by the inverse aggregate dimensionless droop coefficient α, the rate of change of frequency has inertia μ and is controlled by an integral controller with parameter β. These parameters implicitly capture the response of control

mechanisms in the system which ramp power up or down depending on frequency and the integral of frequency.

We are interested in the minimum system frequency during the planning interval, known as the *nadir*, which is given by the random variable $Z = \min_{t\in[0,T]}\dot{\theta}(t)$. Of particular interest is the most likely combination of renewable fluctuations (from their nominal values) and generating unit failures that result in the nadir reaching level $-\gamma$. The event $\{Z \leq -\gamma\}$ depends on the trajectories of the BM W and PP N in a complicated way. We are interested in the most likely trajectories of renewable fluctuations P and conventional outages N conditional on the event $\{Z \leq -\gamma\}$ occurring.

This is related to the instanton approach [6] which effectively maximizes the conditional (on $\{Z \leq -\gamma\}$) likelihood over trajectories of P and N, after discretizing time. Also of interest is the probability that the nadir reaches a prespecified level $-\gamma$, denoted by $Q(\gamma) = \mathbb{P}\,(Z \leq -\gamma)$. In the next section we show how these quantities can be analyzed in a tractable manner using LD theory.

III. A large-deviations approximation

This section shows how to obtain an approximation to the most-likely combination of renewable energy fluctuations and conventional generating unit outages associated with large frequency disturbance events $\{Z \leq -\gamma\}$. To that end, consider a scaled version of (1):

$$\mu\,\ddot{\theta}^{(\varepsilon)}(t) + \alpha\,\dot{\theta}^{(\varepsilon)}(t) = P^{(\varepsilon)}(t) - \delta\,N^{(\varepsilon)}(t) - \beta\,\theta^{(\varepsilon)}(t) \tag{2a}$$
$$\dot{P}^{(\varepsilon)}(t) = -\rho\,P^{(\varepsilon)}(t) + \sqrt{\varepsilon}\widetilde{\sigma}\,\dot{W}(t)\,, \tag{2b}$$

where failures occur at the scaled rate $\lambda = \exp(-\widetilde{\lambda}/\varepsilon)$, with $\widetilde{\sigma}$ and $\widetilde{\lambda}$ constants, and $\varepsilon > 0$ a scaling parameter, Let $\boldsymbol{\theta}^{(\varepsilon)} \equiv (\dot{\theta}^{(\varepsilon)},\dots,\theta^{(6)^{(\varepsilon)}})$ and $Z^{(\varepsilon)} = \min_{t\in[0,T]}\dot{\theta}^{(\varepsilon)}(t)$ with corresponding probability of a large frequency disturbance $Q^{(\varepsilon)}(\gamma) = \mathbb{P}\,(Z^{(\varepsilon)} \leq -\gamma)$. We will now show how to approximate this probability and identify the most likely combination of renewable energy fluctuations and conventional generating unit outages associated with disturbances.

Given the 7×7 matrix $\mathcal{A}$, which we define later, we show that the most relevant limiting trajectories to our study are elements of the family of functions defined by

$$y(t) = \exp\big(\mathcal{A}\,t\big)y(0)\,, \quad t \in [0,T]\,, \tag{3}$$

with

$$y(0) = \left[\begin{array}{ccccccc} \mathbb{0}_{1,2} & -\frac{k\delta}{\mu} & c_1 & c_2 & c_3 & -\frac{k\delta a_2}{\mu} + a_3 c_2 + \frac{\beta\rho^2\delta k}{\mu^2} \end{array}\right]^{\mathsf{T}}.$$

Here, $k \in \mathbb{N}$ and the free variables $c = (c_1, c_2, c_3) = c \in \mathbb{R}^3$. The ith coordinate y_i is corresponding to $\theta^{(i-1)^{(\varepsilon)}}$ as $\varepsilon \downarrow 0$.

Given $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{B}_3$ defined in this section, we have the following result, proven in Section V.

Theorem 1: As $\varepsilon \downarrow 0$

$$Q^{(\varepsilon)}(\gamma) = \exp\big(-\frac{1}{\varepsilon}J(c^\star, k^\star)\big(1 + \mathrm{o}(1)\big)\big)\,,$$

where $o(1) \to 0$ and $J(c^\star, k^\star)$ is the optimal value of

$$\min_{c \in \mathbb{R}^3, \, k \in \mathbb{N}} J(c, k), \qquad (4)$$
$$\text{s.t.} \quad \min_{0 \le t \le T} y_2(t) \le -\gamma,$$

where

$$J(c, k) = k\widetilde{\lambda} + \frac{1}{2}\widetilde{\sigma}^{-2} \Big\{ y(0)^\top \mathcal{B}_1 y(0) \\ + \mathcal{B}_2 y(0) + y(0)^\top \mathcal{B}_3 + (\rho\delta k)^2 T \Big\}. \qquad (5)$$

In this result $y_2(t)$ is computed using (3). The rate of decay of the noise in the scaled model (2), as given by J, provides information on the most likely way a large frequency dip occurs. Events which are less likely in the original model have a higher rate of decay and consequently become exponentially less likely in the scaled model as $\varepsilon \downarrow 0$. This allows the decay rate to be used to compare the likelihood of different events. To determine the most likely event, we let $k^\star$ and $c^\star$ be the optimizers in (4) and the corresponding trajectory of (3) be $y^\star$. The most likely way that a frequency dip occurs is by having k^* generator failures at the beginning of the interval $[0, T]$, and the renewable fluctuations follow the path

$$p^\star(t) = \mu y_3^\star(t) + \alpha y_2^\star(t) + \beta y_1^\star(t).$$

We conclude this section by providing expressions for $\mathcal{A}, \mathcal{B}_1, \mathcal{B}_2,$ and $\mathcal{B}_3$. Firstly

$$\mathcal{A} = \begin{bmatrix} \mathbb{O}_{7,1} & \begin{array}{c} \mathbb{I}_6 \\ [\ a_1 \ 0 \ a_2 \ 0 \ a_3 \ 0\ ] \end{array} \end{bmatrix},$$

$\mathbb{O}_{i,j}$ is a $i \times j$ matrix of zeros, $\mathbb{I}_i$ is a $i \times i$ identity matrix, with $a_1 = \mu^{-2}\beta^2\rho^2$, $a_2 = \mu^{-2}(2\beta\rho^2\mu - \beta^2 - \alpha^2\rho^2)$, $a_3 = \mu^{-2}(\alpha^2 + \rho^2\mu^2 - 2\mu\beta)$. Now, define $(\mathcal{H})_{ij} = h_i h_j$ with $h_1 = \beta\rho$, $h_2 = \beta + \rho\alpha$, $h_3 = \alpha + \rho\mu$, $h_4 = \mu$, $h_5 = h_6 = h_7 = 0$, and $h_8 = \rho\delta k$ and let

$$\mathcal{H} = \begin{bmatrix} \mathcal{H}_1 & \mathcal{H}_2 \\ \mathcal{H}_2^\top & (\rho\delta k)^2 \end{bmatrix},$$

where $\mathcal{H}_1$ and $\mathcal{H}_2$ are 7×7 and 7×1 submatrices. The matrices $\mathcal{B}_1, \mathcal{B}_2,$ and $\mathcal{B}_3$ are computed as follows. Firstly $\mathcal{B}_1 = \mathcal{B}_{12}^\top \mathcal{B}_{11}$, $\mathcal{B}_3 = \mathcal{B}_{12}^\top \mathcal{B}_{31}$ are 7×7 and 7×1 matrices with $\mathcal{B}_2, \mathcal{B}_{11}, \mathcal{B}_{12},$ and $\mathcal{B}_{31}$ corresponding to 1×7, 7×7, 7×7, and 7×1 submatrices that follow from evaluating the matrix exponentials

$$e^{\mathcal{Q}_1 T} = \begin{bmatrix} \mathcal{D} & \mathcal{B}_{11} \\ \mathbb{O}_{7,7} & \mathcal{B}_{12} \end{bmatrix}, \quad e^{\mathcal{Q}_2 T} = \begin{bmatrix} \mathcal{D} & \mathcal{B}_2 \\ \mathbb{O}_{1,7} & \mathcal{D} \end{bmatrix},$$
$$e^{\mathcal{Q}_3 T} = \begin{bmatrix} \mathcal{D} & \mathcal{B}_{31} \\ \mathbb{O}_{1,7} & 0 \end{bmatrix},$$

letting $\mathcal{D}$ denote entries which are discarded, with

$$\mathcal{Q}_1 = \begin{bmatrix} -\mathcal{A}^\top & \mathcal{H}_1 \\ \mathbb{O}_{7,7} & \mathcal{A} \end{bmatrix}, \quad \mathcal{Q}_2 = \begin{bmatrix} 0 & \mathcal{H}_2^\top \\ \mathbb{O}_{7,1} & \mathcal{A} \end{bmatrix},$$
$$\mathcal{Q}_3 = \begin{bmatrix} -\mathcal{A}^\top & \mathcal{H}_2 \\ \mathbb{O}_{1,7} & 0 \end{bmatrix}.$$
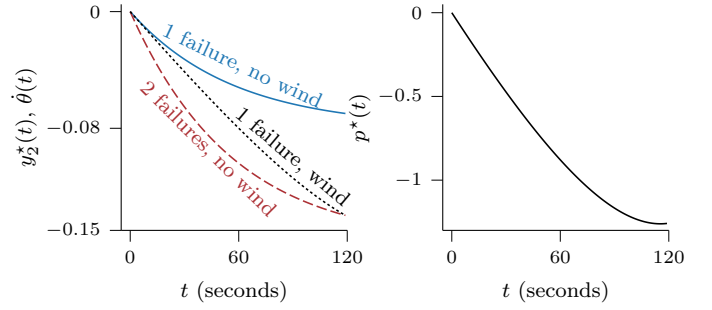
(a) Frequency deviation.   (b) Renewable fluctuation.

Fig. 1: (a) Most likely random trajectory to failure (dotted) when $\sigma = 0.2916$ and $\lambda = 10^{-3}$ compared with failure scenarios with one (solid) or two (dashed) conventional failures but no renewable fluctuations, and (b) associated renewable fluctuation for random system.

IV. EXPERIMENTAL RESULTS

This section contains three illustrative numerical investigations. Unless specified otherwise: $\mu = 12$, $\alpha = 12.5$, $\beta = 0.05$, and $\rho = 1/30$; these values are in line with the values used in [10]. Additionally, we take $\varepsilon = 0.1$ and $\delta = 1$. This value of δ corresponds to each conventional failure resulting in 1000MW being removed from the system, which is the typical loss protected against in Great Britain's power system [12]. Additionally, $T = 2$ minutes (i.e., 120 seconds).

We utilize the Python 3.7.5 SciPy minimize function using the 'SLSQP' routine to find $c^\star \in \arg \min_c J(c, k)$ for fixed k; this procedure is repeated for a range of k values and the best selected. We report on the resulting $k^\star$ and $p^\star(T)$ values, as described in the previous section. These are used to quantify the most likely number of generator outages and renewable fluctuation size associated with large frequency disturbances. The code used to perform these experiments is available at https://github.com/bpatch/power-system-frequency-deviations.

A. Most likely path to failure in contemporary power systems.

We begin our investigation by considering the effect on the frequency of a system subject to one or two conventional failures which is not subject to renewable fluctuations. In Fig. 1a the solid and dashed lines show how the frequency of such a system evolves when one or, respectively, two conventional generating units fail at $t = 0$. It can be seen that a single failure results in frequency reaching approximately -0.0699 within the planning interval and two failures results in the frequency reaching approximately -0.1397. The dotted line in this figure is discussed in the next subsection.

In [2] it is suggested that renewable fluctuations in the range $30-50\%$ of total capacity over a 2 minute interval at any individual wind park may occur once over several months. For a windfarm such as Hornsea Offshore on the Great Britain power system, which at times produces 737MW [12], this means that a reasonable rare renewable fluctuation is of the order 350MW and occurs in any two minute planning interval with probability of the order $1/43200 \approx 0.000023$ (since there

are approximately 43200 two minute planning intervals every two months). Our OU assumption for renewable fluctuations implies $P(2) \sim \text{Normal}(0, \frac{\sigma^2}{2\rho}(1 - e^{-4\rho}))$, i.e., approximately $P(2) \sim \text{Normal}(0, \sigma^2 1.8724)$ given our parameter values. This implies that $\sigma \approx 0.0628$ is a reasonable illustrative renewable diffusion coefficient value for current power systems (since this implies $\mathbb{P}(P(2) \leq -0.35) = 1/43200$). After consultation with practitioners, we found it reasonable to illustrate our method under the assumption that approximately two to three conventional generator failures will occur every three days, implying $\lambda = 10^{-3}$ is an appropriate order of magnitude. With these values of σ and λ our model and analysis suggests that the most likely way for a nadir of -0.1397 to occur within an arbitrary planning interval is from two conventional generator failures occurring in the planning interval, providing evidence that conventional failures are currently the primary cause of large frequency deviations.

In light of this it is reasonable to ask for what values of σ do renewable fluctuations threaten system security? We address this question in the next subsection.

B. Conventional outages vs. renewable fluctuations.

We now investigate whether the results from the previous subsection hold for a greater range of renewable fluctuation intensities and conventional failure rates. The plots in Fig. 2 explore $\sigma \in [0.03, 0.5]$ and $\lambda \in [10^{-10}, 10^{-1}]$. To apply Theorem 1, we set $\varepsilon = 0.1$.

Fig. 2a provides evidence that the effect of an increase in σ or λ on the probability of a frequency deviation to -0.1397 has a strong dependence on the respective value of λ or σ. It can, for example, be seen that that for a high value of either σ or λ we have $Q^{(\varepsilon)}(0.1397) \approx \exp(-0.8/\varepsilon) \approx 0.0003$ regardless of the value of the other parameter. Of note though is that in Fig. 2b we see that the cause of the failure probability taking this value is different depending on whether σ or λ takes on a large value. For high values of σ and low values of λ the nadir is most likely to be caused by high levels of renewable fluctuations, while for high values of λ and low values of σ the nadir is most likely to be caused by two conventional failures, but for high values of both parameters the nadir is most likely to be caused by a combination of conventional failures and renewable fluctuations. For any particular smaller value of λ, Fig. 2b shows that as σ increases there are distinct thresholds where renewable fluctuations replace conventional failures as the most likely cause of a large frequency deviation. A key strength of our framework is that it is able to provide system operators with an estimate of where these thresholds are.

Considering $\lambda = 10^{-3}$, as used in the previous subsection, we see that for $\sigma < 0.28$ the most likely cause of a frequency deviation to -0.1397 is two conventional generator failures. At this threshold value of σ (marked with a cross) there is a transition in system behavior and the most likely cause of such a frequency deviation becomes a single conventional generator failure accompanied by renewable generation fluctuations. To illustrate this, in Fig. 1a the dotted line corresponds to the most likely path to -0.1397 for a system which is subject

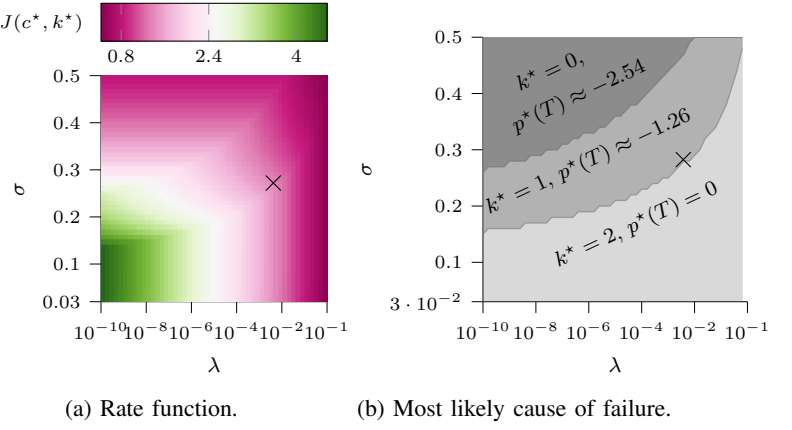(a) Rate function.   (b) Most likely cause of failure.

Fig. 2: (a) Rate function value for a large frequency deviation, and (b) most likely cause of a large frequency deviation. Both as a function of conventional generation failure rate (λ) and renewable fluctuation intensity (σ).
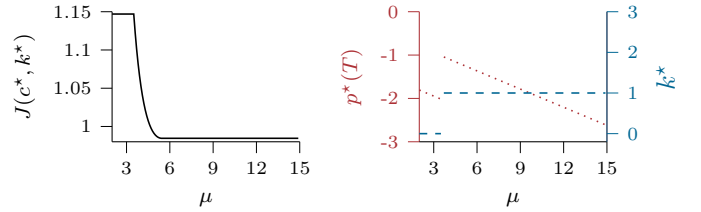

Fig. 3: Influence of inertia μ on LD decay rate $J(c^\star, k^\star)$ (left) and the most likely cause ($k^\star, p^\star(T)$) (right) with $\gamma = 0.1397$.

to conventional failures at rate $\lambda = 10^{-3}$ and renewable fluctuations with variability $\sigma = 0.28$. In Fig. 1b the associated most likely renewable fluctuation is $p^\star(T) \approx -1.27$ (i.e, a 1270MW power loss).

The experiments conducted in this subsection provide further evidence that conventional failures remain to be the primary threat to system security, and that only after substantial increases in renewable penetration would this source of stochasticity present a true threat to system security.

C. Inertia.

Inertia influences the likelihood and causes of large deviations. The curve in the left panel of Fig. 3 displays the LD decay rate $J(c^\star, k^\star)$ as a function of μ. Recall $Q^{(\varepsilon)}(0.1397) \approx \exp(-J(c^\star, k^\star)/\varepsilon)$ approximates $Q(0.1397)$, implying that as $J(c^\star, k^\star)$ decreases the probability of the frequency deviation increases. Therefore this curve provides evidence that there are ranged of inertia over which the probability of a large frequency deviation decreases as inertia is increased. In this example this range is close to a value of inertia where the most likely number of conventional failures increases from zero to one, and the most likely renewable deviation experiences a simultaneous decrease in magnitude from approximately -2 to approximately -1. Interestingly there are large ranges of inertia over which changes in inertia do not noticeably affect the probability of a large frequency disruption but where the magnitude of the associated renewable fluctuation increases.

V. PROOF OF MAIN RESULT

To prove Thm. 1 we first provide several lemmas. Our first lemma is a useful integral representation for the frequency in terms of the renewable fluctuations P and outages N.

Lemma 2: Frequency evolves according to

$$\dot{\theta}(t) = \int_0^t a(t-s)[P(s) - \delta N(s)]ds \qquad (6)$$

with $a(s) = x_3(x_1 e^{x_1 s} - x_2 e^{x_2 s})$, where $\zeta = \sqrt{\alpha^2 - 4\beta\mu}$,

$$x_1 = \frac{-\alpha + \zeta}{2\mu}, \quad x_2 = \frac{-\alpha - \zeta}{2\mu}, \quad x_3 = \delta\mu/\zeta.$$

Proof. Let $G = (\theta, \dot{\theta})$, then

$$dG(t) = \mathcal{M}G(0)dt + \begin{bmatrix} 0 \\ P(t) - \delta N(t) \end{bmatrix} dt$$

where

$$\mathcal{M} = \begin{bmatrix} 0 & 1 \\ -\alpha/\mu & -\beta/\mu \end{bmatrix}$$

so

$$G(t) = e^{\mathcal{M}t}G(0) + \int_0^t e^{\mathcal{M}_1(t-s)} \begin{bmatrix} 0 \\ P(s) - \delta N(s) \end{bmatrix} dt.$$

Let $a(s) = (e^{\mathcal{M}s})_{2,2}$. Since $G(0) = \begin{bmatrix} 0 & P(0) - \delta N(0) \end{bmatrix}^\top$ and $P(0) = 0$, we have that $a(s)$ satisfies (6) as required. $\square$

We next provide a useful monotonicity result.

Lemma 3: If the outage process $N(s)$ is replaced with $\widetilde{N}(s)$ such that $N(s) \le \widetilde{N}(s)$ for $s \in [0, T]$, then the associated frequency process $\dot{\theta}(s), s \in [0, T]$ satisfies $\dot{\theta}(s) \ge \dot{\widetilde{\theta}}(s), s \in [0, T]$. Consequently, given $N(T) = k$, the frequency process $\dot{\theta}(s)$ is minimized by choosing $N(s) = k$, $s > 0$.

Proof. Note that $x_1 > x_2$ in Lem. 6 implies $a(0) > 0$ and therefore we have $a(s) > 0$ for all $s \in [0, T]$. Hence the result follows from (6). $\square$

Lemma 4: Define $\xi_k^{(\varepsilon)}(t) = \mathbb{P}(N^{(\varepsilon)}(t) = k)$. As $\varepsilon \downarrow 0$,

$$\varepsilon \log \xi_k^{(\varepsilon)}(\nu) \to -\widetilde{\lambda}k$$

for all $s \in [0, T]$.

Proof. This follows from

$$\lim_{\varepsilon \to 0} \varepsilon \log \left(\frac{e^{-k\widetilde{\lambda}/\varepsilon} T^k}{k!} e^{-\exp(-\widetilde{\lambda}/\varepsilon)T} \right) = -k\widetilde{\lambda}.$$

$\square$

Define the operator $\phi(f) = \inf_{0 \le t \le T} \int_0^T a(t-s)f(s)ds$.

Lemma 5:

$$\lim_{\varepsilon \to 0} \varepsilon \log \mathbb{P} \left(\inf_{0 \le t \le T} \int_0^t a(t-s)P^{(\varepsilon)}(s)ds \le -\eta \right) = -I_g^\star(\eta),$$

With $I_g^\star(\eta) = \inf_{\substack{f \in H_0: \\ \phi(f) \le -\eta}} I_d(f)$, and $I_d(f)$ defined in (7).

Proof. Upon removing $N^{(\varepsilon)}$ from (2), $\{\theta^{(\varepsilon)}\}$ satisfies an LDP with good rate function

$$I_d(f) = \begin{cases} \frac{1}{2} \int_0^T \mathcal{L}(f, \dot{f}, \ddot{f}, f^{(3)})^2 ds, & \text{if } f \in H_0, \\ \infty, & \text{if } f \notin H_0, \end{cases} \qquad (7)$$

where $\mathcal{L}(f, \dot{f}, \ddot{f}, f^{(3)}) = (\mu f^{(3)} + (\alpha + \rho\mu)\ddot{f} + (\beta + \rho\alpha)\dot{f} + \beta\rho f)^2$ and

$$H_x = \{f \in C_x[0, T] : f(t) = x + \int_0^t g(s)ds, \; g \in L_2[0, T]\}$$

is the space of absolutely continuous functions with square integrable derivative. To see this, from (2b) and (2a),

$$\dot{W}(t) = \widetilde{\sigma}^{-1}(\dot{P}(t) + \rho P(t)), \qquad (8)$$

$$P(t) = \mu\ddot{\theta}(t) + \alpha\dot{\theta}(t) + \beta\theta(t) + \delta k, \qquad (9)$$

$$\dot{P}(t) = \mu\theta^{(3)}(t) + \alpha\ddot{\theta}(t) + \beta\dot{\theta}(t). \qquad (10)$$

Upon substituting (9) and (10) into (8) and rearranging,

$$\dot{W} = \widetilde{\sigma}^{-1}\left(\mu\theta^{(3)} + (\alpha + \rho\mu)\ddot{\theta} + (\beta + \rho\alpha)\dot{\theta} + \beta\rho\theta + \rho\delta k\right).$$

Since we removed N, we have $k = 0$. This characterizes the inverse of a continuous map. Using the contraction principle (see [4, Thm. 4.2.1]) and Schilder's Theorem (see [4, Thm. 5.2.3]), the appropriate rate function for $\{\theta^{(\varepsilon)}\}$ is (7). $\square$

Lemma 6:

$$\lim_{\varepsilon \to 0} \varepsilon \log \mathbb{P}\left(N^{(\varepsilon)}(T) \ge k + 1\right) = (k+1)\widetilde{\lambda}$$

Proof. The lower bound follows from Lemma 4; the upper bound from $\mathbb{P}\left(N^{(\varepsilon)}(T) \ge k + 1\right) \le e^{-(k+1)\widetilde{\lambda}/\varepsilon}$. $\square$

Lemma 7:

$$\varepsilon \log Q^{(\varepsilon)}(\gamma) \to -\inf_{k=0,1,\dots\bar{k}}[\widetilde{\lambda}k + I_g^\star(\gamma - f(0, k))].$$

with $f(\nu, k) = \delta k \int_\nu^T a(t-s)ds$ and $\bar{k} = \inf\{k : f(0, k) \ge \gamma\}$.

Proof. Note that $f(\nu, k)$ represents the decrease in frequency in $[\nu, T]$ if there have been k generator failures by time ν. We begin with the asymptotic lower bound. Let

$$\bar{p}^{(\varepsilon)}(t) = \int_0^t a(t-s)P^{(\varepsilon)}(s)ds,$$

then

$$Q^{(\varepsilon)}(\gamma) \ge \mathbb{P}\left(Z^{(\varepsilon)} \le -\gamma, \; N^{(\varepsilon)}(\nu) \le \bar{k}\right)$$

$$= \sum_{k=0}^{\bar{k}} \mathbb{P}\left(Z^{(\varepsilon)} \le -\gamma, \; N^{(\varepsilon)}(\nu) = k\right)$$

$$\ge \sum_{k=0}^{\bar{k}} \mathbb{P}\left(\inf_{0 \le t \le T} \bar{p}^{(\varepsilon)}(t) \le -(\gamma - f(\nu, k))\right) \xi_k^{(\varepsilon)}(\nu)$$

Using Lemma's 4 and 5 and the principle of the largest term, we obtain

$$\liminf_{\varepsilon \to 0} \varepsilon \log Q^{(\varepsilon)}(\gamma) \ge -\inf_{k \le \bar{k}}[\widetilde{\lambda}k + I_g^\star(\gamma - f(\nu, k))]$$

The lower bound now follows by letting $\nu \downarrow 0$.

For the upper bound, we write

$$Q^{(\varepsilon)}(\gamma) \le \mathbb{P}(Z^{(\varepsilon)} \le -\gamma, \; N^{(\varepsilon)}(T) \le \bar{k})\mathbb{P}(N^{(\varepsilon)}(T) \le \bar{k})$$

$$+ \mathbb{P}(Z^{(\varepsilon)} \le -\gamma, \; N^{(\varepsilon)}(T) > \bar{k})\mathbb{P}(N^{(\varepsilon)}(T) > \bar{k})$$

$$\le \mathbb{P}(Z^{(\varepsilon)} \le -\gamma, \; N^{(\varepsilon)}(T) \le \bar{k}) + \mathbb{P}(N^{(\varepsilon)}(T) > \bar{k})$$

The behavior of the second term is controlled by Lemma 6. Using a similar argument as in the proof of the lower bound, we can show for the first term

$$\liminf_{\varepsilon \to 0} \varepsilon \log \mathbb{P}\left(Z^{(\varepsilon)} \le -\gamma,\, N^{(\varepsilon)}(T) \le \bar{k}\right)$$
$$\ge -\inf_{k \le \bar{k}}[\widetilde{\lambda}k + I_g^\star(\gamma - f(\nu, k))].$$

Since $f(\nu, \bar{k}) \ge \gamma$, $I_g^\star(\gamma - f(\nu, \bar{\,})) = 0$ (no rare-event behavior of $P^{(\varepsilon)}$ is required) and therefore the first term dominates. $\square$

Proof of Theorem 1. As the dominant scenario is to have several outages at time 0, we can also represent the decay rate of the probability of interest as

$$\min_{f,k} \quad k\widetilde{\lambda} + I_{d,k}(f),$$
$$\text{s.t.} \quad \min_{0 \le t \le T} \dot{f}(t) \le -\gamma, \qquad (11)$$
$$f(0) = \dot{f}(0) = 0 \ \ddot{f}(0) = -k\delta/\mu.$$

Here, $I_{d,k}(f)$ is the same as $I_d(f)$ defined in (7), except that $\mathcal{L}$ is replaced by $\mathcal{L}(f, \dot{f}, \ddot{f}, f^{(3)}) = \big(\mu f^{(3)} + (\alpha + \rho\mu)\ddot{f} + (\beta + \rho\alpha)\dot{f} + \beta\rho f + \delta k\big)^2$. Optimality is determined by the Euler-Lagrange equation

$$\nabla_f \mathcal{L}_k - \frac{\mathrm{d}}{\mathrm{d}t}\left(\nabla_{f^{(1)}}\mathcal{L}_k\right) + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(\nabla_{f^{(2)}}\mathcal{L}_k\right) - \frac{\mathrm{d}^3}{\mathrm{d}t^3}\left(\nabla_{f^{(3)}}\mathcal{L}_k\right) = 0,$$

where $\nabla_f \mathcal{L}_i$ is the derivative of $\mathcal{L}_i$ with respect to the function f (see e.g., [13, p. 190]). This equation can be simplified to

$$\beta\rho^2\delta k - \mu^2 f^{(6)} + \left(\alpha^2 + \rho^2\mu^2 - 2\mu\beta\right)f^{(4)}$$
$$+ \left(2\beta\rho^2\mu - \beta^2 - \alpha^2\rho^2\right)\ddot{f} + \beta^2\rho^2 f = 0. \qquad (12)$$

Differentiation of (12) leads to

$$-\mu^2 f^{(7)} + \left(\alpha^2 + \rho^2\mu^2 - 2\mu\beta\right)f^{(5)}$$
$$+ \left(2\beta\rho^2\mu - \beta^2 - \alpha^2\rho^2\right)f^{(3)} + \beta^2\rho^2\dot{f} = 0, \qquad (13)$$

and the additional initial condition $f^{(6)}(0) = a_1 f(0) + a_2 \ddot{f}(0) + a_3 f^{(4)}(0) + \mu^{-2}\beta\rho^2\delta k$. Now, for $k = 1, \dots, 7$ define $y_k = f^{(k-1)}$, then due to (13), the solution to (11) is of the form (3), and the function f in (4) is fully specified by c and k. In this notation $\mathcal{L}_k$ becomes $\mathcal{L}_k\big(y_1, y_2, y_3, y_4\big) = \big(\mu y_4 + (\alpha + \rho\mu)y_3 + (\beta + \rho\alpha)y_2 + \beta\rho y_1 + \rho\delta k\big)^2$, so that given our explicit form of y. Therefore J in (4) equals

$$k\widetilde{\lambda} + \frac{1}{2}\widetilde{\sigma}^{-2}\left[\, y(0)^\top \ \ 1 \,\right] \int_0^T \mathrm{e}^{\mathcal{A}^\top t}\,\mathcal{H}\,\mathrm{e}^{\mathcal{A}t}\mathrm{d}t \left[\begin{array}{c} y(0) \\ 1 \end{array}\right],$$

where $\mathcal{H}$ is given in Sec. III. The second term in the above expression for J can be rewritten as

$$\frac{1}{2}\widetilde{\sigma}^{-2}\Big\{y(0)^\top \int_0^T \mathrm{e}^{\mathcal{A}t}\mathcal{H}_1\mathrm{e}^{\mathcal{A}t}\mathrm{d}t\, y(0) + \int_0^T \mathcal{H}_2^\top \mathrm{e}^{\mathcal{A}t}\mathrm{d}t\, y(0)$$
$$+ y(0)^\top \int_0^T \mathrm{e}^{\mathcal{A}^\top t}\mathcal{H}_2\mathrm{d}t + (\rho\delta k)^2 T\Big\}.$$

The three integrals can be evaluated using [14, Thm. 1] as

$$\int_0^T \mathrm{e}^{\mathcal{A}t}\mathcal{H}_1\mathrm{e}^{\mathcal{A}t}\mathrm{d}t = \mathcal{B}_1, \qquad \int_0^T \mathcal{H}_2^\top \mathrm{e}^{\mathcal{A}t}\mathrm{d}t = \mathcal{B}_2,$$
$$\int_0^T \mathrm{e}^{\mathcal{A}^\top t}\mathcal{H}_2\mathrm{d}t = \mathcal{B}_3,$$

where these matrices are defined in Sec. (III). Combining these together leads to (5). $\square$

VI. Outlook

We have illustrated how LD can be applied to investigate the interplay of conventional outage and renewable fluctuations. We have used a Gaussian process model noise, which might be conservative and therefore the impact of renewable fluctuations compared to conventional generating unit outages may be more significant than our analysis suggests; LD theory can handle more general types of noise. This and several other extensions are called for. In particular, the influence of the network topology is a key aspect to many frequency deviation questions; early computations of a network version of our model are promising. We also intend to extend the LD approach to incorporate finite turbine-time constants and deadband controls. Finally, our analysis can be used to enhance rare event simulation methods, as in [15].

References

[1] Gielen, D., Boshell, F., Saygin, D., Bazilian, M.D., Wagner, N. and Gorini, R., 2019. The role of renewable energy in the global energy transformation. *Energy Strategy Reviews*, 24, pp.38-50.

[2] Milan, P., Wächter, M. and Peinke, J., 2013. Turbulent character of wind energy. *Physical Review Letters*, 110(13), p.138701.

[3] Ahmadyar, A.S., Riaz, S., Verbic, G., Chapman, A. and Hill, D.J., 2017. A Framework for Frequency Stability Assessment of Future Power Systems: An Australian Case Study. *arXiv preprint arXiv:1708.00739*.

[4] Dembo, A. and Zeitouni, O., 1998. *Large Deviations Techniques and Applications*.

[5] Moriarty, J., Vogrinc, J. and Zocca, A., 2018, December. Frequency violations from random disturbances: an MCMC approach. In *2018 IEEE Conference on Decision and Control (CDC)* (pp. 1598-1603). IEEE.

[6] Chertkov, M., Pan, F. and Stepanov, M.G., 2010. Predicting failures in power grids: The case of static overloads. *IEEE Transactions on Smart Grid*, 2(1), pp.162-172.

[7] Lee, D., Aolaritei, L., Vu, T.L. and Turitsyn, K., 2019. Robustness against disturbances in power systems under frequency constraints. *IEEE Transactions on Control of Network Systems*, 6(3), pp.971–979.

[8] Nesti, T., Nair, J. and Zwart, B., 2019. Temperature overloads in power grids under uncertainty: A large deviations approach. *IEEE Transactions on Control of Network Systems*, 6(3), pp.1161-1173.

[9] Nesti, T., Zocca, A. and Zwart, B., 2018. Emergent failures and cascades in power grids: a statistical physics perspective. *Physical Review Letters*, 120(25), p.258301.

[10] Vorobev, P., Greenwood, D.M., Bell, J.H., Bialek, J.W., Taylor, P.C. and Turitsyn, K., 2019. Deadbands, Droop, and Inertia Impact on Power System Frequency Distribution. *IEEE Transactions on Power Systems*, 34(4), pp.3098-3108.

[11] Ganesh, A.J., O'Connell, N. and Wischik, D.J., 2004. *Big Queues*. Springer.

[12] Technical Report on the events of 9 August 2019. Available at https://www.nationalgrideso.com/document/152346/download.

[13] Courant, R. and Hilbert, D., 2008. *Methods of Mathematical Physics: Partial Differential Equations*. John Wiley & Sons.

[14] Van Loan, C., 1978. Computing integrals involving the matrix exponential. *IEEE Transactions on Automatic Control*, 23(3), pp.395-404.

[15] Wadman, W.S., Crommelin, D.T. and Zwart, B., 2016. A large-deviation-based splitting estimation of power flow reliability. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 26(4), pp.1-26.