Phasor Measurement Unit Change-Point Detection of Frequency Hurst Exponent Anomaly with Time-to-Event

J. Sia[1], E. Jonckheere[2,*], L. Shalalfeh[3], and P. Bogdan[4,†]

Abstract—The objective of this paper is real-time detection of a change-point in the baseline distribution of the frequency signal generated by Phasor Measurement Units (PMUs) that could indicate potential for voltage collapse, false data injection, or other security threats. The alarm flags an anomaly event when the cumulative sum (CUSUM) of the log-likelihood departures of the Hurst exponent of the frequency from its baseline statistic exceeds a threshold set up as a compromise between the conflicting objectives of minimum detection delay and acceptable False Alarm Rate. As main theoretical contribution, an extra protection layer is developed that provides an estimate of the time to the alarm event, giving the Transmission System Operator (TSO) or an autonomous agent more time to deploy proactive measures to avoid large catastrophic system states. The proposed method is illustrated by a retrospective analysis of the 2012 Indian blackout that reveals that 10-12 minutes before the voltage collapse, a significant increase in the Hurst exponent of the frequency could have been detected. Conceptually, this shows that PMUs provide significant value towards ensuring autonomous cognition in the power grid by enabling the abnormal events to be fault-detected in an unsupervised fashion.

Index Terms—Smart grid, frequency anomaly, Hurst exponent, Change-Point Detection

1 Introduction and Novel Contributions

Phasor Measurement Unit (PMU) technology enables a wide range of applications within the autonomous power grid design such as state estimation, voltage and transient stability analysis, oscillation monitoring, event and fault detection [24], situation awareness, and model validation [8]. In order to enable the dynamic management of the smart grid for higher efficiency, resiliency, stability and security, the ability to detect abnormal events and the change-point concurrent with abnormality are especially important to prevent voltage collapse and power system blackout [11] [12].

Fundamental for endowing the smart grid with autonomous cognitive intelligence is its ability to monitor and analyze in real-time the mathematical characteristics of the PMU signals and identify the change-points through rigorous and robust statistical techniques. Towards this end, a pioneering effort [27], [29] demonstrated that PMU time series (consisting of voltage magnitude, phase angle, and frequency) exhibit long-range memory and fractal characteristics. From a mathematical perspective, this long-range memory and fractal behavior are evaluated quantitatively through the Hurst exponent. The Hurst exponent determines whether the observations are obeying a short-range memory dynamics (an exponential decay of the autocorrelation function with the lag) or a long-range memory dynamics (a power law decay of the autocorrelation function with the lag). A Hurst exponent of 0.5 indicates a short-range memory behavior (implying independence between consecutive events). In contrast, the Hurst exponents observed in the PMU analysis were significantly greater than 0.5 and demonstrated to have long-range memory behavior [27].

1,2,4: Dept. of Elec. Eng., Univ. of Southern California, Los Angeles, CA 90089 {jsia,jonckhee,pbogdan}@usc.edu
3: German Jordanian University, Amman, Jordan, laith.shalalfeh@gju.edu.jo
**: corresp. author, supported by NSF CCF-1423624; †: supported by DARPA*

Moreover, an extension of this work has shown that the increasing trend in the Hurst exponent of the frequency time series is a good indicator of the proximity of a power system to blackout and thus can be used as an early warning signal [28] [29] [31]. These prior research efforts not only offer novel ways, next to the ROCOF (Range Of Change Of Frequency), for power systems operators to detect imminent danger given the PMU frequency time series, but also enable new mathematical and algorithmic strategies to monitor the state of the grid and predict the chance of an abnormal event with potentially catastrophic implications in real time.

Starting from the above-mentioned mathematical characteristics of the PMU data and with the goal of enabling mathematical strategies for designing an autonomous power grid, we propose as in [2] a novel and robust change-point detection (CPD) strategy capable of detecting frequency anomalies and the danger of a blackout ahead of time such that proactive measures can be taken. Essentially, CPD continuously monitors the departure of the signal from its baseline Hurst exponent characteristic and constructs a cumulative statistic that triggers the alarm if it crosses a threshold consistent with an acceptable false alarm rate.

From a formal statistical point of view, the situation where the frequency follows its "baseline Hurst exponent characteristic" is referred to as *Null Hypothesis*. From that perspective, a false alarm is a Type I error, that is, the Null Hypothesis is rejected when it is valid.

1.1 Contributions

In summary, we make the following novel contributions:

- The *unique* feature of the detection is that it is based on an anomaly in the Hurst exponent of the PMU frequency signal.
- The CUSUM approach is proposed in contrast with the Bayesian approach [2].

- The CUSUM algorithm is combined with an Expectation Maximization (EM) estimate of the current, possibly abnormal, Hurst exponent density parameter for which the data is a sufficient statistic.
- The CPD theory is shown to be amenable to the Itô calculus formulation for estimating the distribution of the time-to-event and its mean, where "event" is meant to be threshold crossing.
- The proposed CPD algorithmic detection strategy is evaluated on the *real* PMU dataset of the Indian 2012 blackout, which demonstrates that the blackout could have been predicted approximately 12 minutes in advance.

1.2 New results relative to earlier versions

The present paper is an expanded version of [31] with the following main additions:

- The chief experimental difference is that, here, the PMU observation windows of the Indian blackout have been spaced out to make the Hurst exponents feeding the CUSUM algorithm independent, while in [31] this theoretical requirement was disregarded. Moreover, a comparison of the cases of overlapping and non-overlapping, "spaced out," windows has been added (see Fig. 2 and Sec. 5.3).
- A significant addition compared with [31] is the evaluation of the *time-to-event*, that is, a continuously updated unbiased estimate of the time it would take for the alarm to ring should the present trend continue. This is an additional feature on top of the sole reliance on the threshold crossing alarm that might flag an anomaly when it might already be too late to respond.
- Another addition is the histogram of the time-to-event under the initialization $\mathrm{CUSUM} = 0$. This distribution is instrumental for the Type I error analysis.

This last item, along with the accrued level of mathematical rigor necessary to derive the time-to-event, has prompted us to focus here on the PMU applications while providing only sketches of the mathematical technicalities and their proofs, relegating the details to a companion paper [32].

2 Change-point Detection

2.1 Fundamentals

Consider an i.i.d. sequence $\{X_k\}_{k=1}^n$ with "normal" regime probability density function (PDF) p_0 from $k = 1$ up to and including $k = \lambda - 1$, and with "abnormal" PDF p_1 as of $k = \lambda$ up to and including n. Change-Point Detection (CPD) endeavors to find the change-point λ in the fastest possible way subject to an acceptable false alarm rate. There is a vast literature on the subject (see [3] and references cited therein), which can be partitioned into, on the one hand, the Shiryaev (Bayesian) procedure [30] and, on the other hand, Page's CUmulative SUM (CUSUM) (minimax) procedure [20]. In the Shiryaev procedure, λ is assumed to have an a priori distribution and the goal is to minimize the expected detection delay subject to a false alarm rate. In the CUSUM procedure, λ is deterministic, but unknown, and the goal is to minimize the worst case detection delay subject to an acceptable false alarm rate. Here we follow the CUSUM, consistently with the early work on application of CPD to security problems [4], [13], [26].

Regarding the Bayesian procedure, an evaluation of its cognitive intelligence implementation for detecting anomalies in the μPMU current magnitudes is available in [2]. The procedure works best with spatial anomalies [2, Table 2]. Here, we restrict ourselves to temporal rather than spatial anomalies.

Given a change-point λ, let P_λ denote the probability measure defined as p_0 on $\{X_k\}_{k=1}^{\lambda-1}$ and p_1 on $\{X_k\}_{k=\lambda}^n$. Let $\mathbb{E}_\lambda$ be the corresponding mathematical expectation. Let $\mathbb{E}_{p_{0,1}}$ be the mathematical expectation relative to the probability density $p_{0,1}$ on $\{X_k\}_{k=1}^n$. Note that $\mathbb{E}_\infty = \mathbb{E}_{p_0}$.

The Null Hypothesis H_0 that there has been no changes from λ up to and including n is rejected when the log-likelihood ratio CUSUM statistic,

$$Z_\lambda^n = \sum_{k=\lambda}^{n} \log \frac{p_1(x_k)}{p_0(x_k)}, \tag{1}$$

takes "too large" values. As opposed to the Bayesian procedure, the premise of the CUSUM is to consider the above statistic for the worst possible λ. Moreover, for the same reason as making the detection robust, the statistic is reset to 0 in case it takes negative values[1]:

$$U^n = \left\{ \max_{0\le\lambda\le n} Z_\lambda^n \right\}_+, \tag{2}$$

where $\{z\}_+ = \max\{0, z\}$. The alarm is triggered to indicate that H_0 is rejected when $U^n \ge h$, where the threshold h is set up to compromise between False Alarm Rate and detection delay.

The statistic U^n can be put in recursive form as follows:

$$\begin{aligned} U^{n+1} &= \left\{ \max_{0\le\lambda\le n+1} \left(Z_\lambda^n + \log \frac{p_1(x_{n+1})}{p_0(x_{n+1})} \right) \right\}_+ \\ &= \left\{ \max \left\{ U^n + \log \frac{p_1(x_{n+1})}{p_0(x_{n+1})}, \log \frac{p_1(x_{n+1})}{p_0(x_{n+1})} \right\} \right\}_+ . \end{aligned}$$

The first argument of the $\max\{\cdot,\cdot\}$ is the case where the $\max_\lambda$ is reached for $0 \le \lambda \le n$ while the second argument is the case where the $\max_\lambda$ is reached for $n+1$. Since U^n is forced to be nonnegative, the first term in the argument of $\max\{\cdot,\cdot\}$ is greater than or equal to the second. Hence, the recursion:

$$U^{n+1} = \max \left\{ 0, U^n + \log \frac{p_1(X_{n+1})}{p_0(X_{n+1})} \right\}, \quad U^0 = 0. \tag{3}$$

The decision that there has been a change is taken at the first time τ that the CUSUM statistic U^n equals, or goes above, the threshold h:

$$\tau(h) = \min\{n : U^n \ge h\}. \tag{4}$$

To derive the FAR, assume that, while p_0 prevails, $U^k < h$ for $0 \le k < n$ and $U^n \ge h$, hence a false alarm. The

1. Here, the exposition is simplified relative to the traditional one to avoid the formal argument justifying the reset of the recursive algorithm (3) to 0 in case U^{n+1} takes negative values [3, §2.2]. Another reason for this is to prepare the ground for simultaneous detection and identification of p_1.

expected time from $U^0 = 0$ to the false alarm is

$$T_{\mathrm{FA}}(h) = \mathbb{E}_{p_0}(\tau(h)|U^0 = 0).$$

The False Alarm Rate, confronted with its admissible upper bound, is

$$\mathrm{FAR}(h) = \frac{1}{T_{\mathrm{FA}}(h)} \leq \overline{\mathrm{FAR}}. \tag{5}$$

As proved in [32], $\mathbb{E}_{p_0}(\tau(h))$ is monotone decreasing with h. Therefore, the threshold is selected as $\underline{h} = \min\{h : \text{Eq. (5) holds}\}$. With this threshold, the Average Detection Delay is

$$\mathrm{ADD}_\lambda = \mathbb{E}_\lambda(\tau\left(\underline{h}\right) - \lambda : \tau\left(\underline{h}\right) \geq \lambda).$$

This CPD minimizes ADD_λ subject to $\mathrm{FAR}(h) \leq \overline{\mathrm{FAR}}$.

2.2 Novelty: time to alarm event

In the time to false alarm, the initial condition was set as $U^0 = 0$. Assume now that, at some time m before the alarm, $U^m = x < h$. Resetting m to 0, the time-to-event is defined as

$$T(x) = \mathbb{E}_{p_0}\left(\tau(h)|U^0 = x\right),$$

a generalization of T_{FA}.

2.3 Itô connection

To acquire an early intuitive understanding of how the algorithm works—and to connect with Itô calculus—let $p_{0,1}$ be Gauss-distributed as $\mathcal{N}(\mu_{0,1}, \sigma^2)$; in other words, the abnormality is a mere change of the mean, keeping the variance unchanged to simplify. It is then easily verified that

$$U^{n+1} = \left\{U^n - \frac{(\mu_0 - \mu_1)^2}{2\sigma^2} + \frac{\mu_1 - \mu_0}{\sigma}\left(\frac{X_{n+1} - \mu_0}{\sigma}\right)\right\}_+. \tag{6}$$

The intuitive idea is that U^{n+1} is the CUSUM of the deviation of the data X_{n+1} from what one would expect of the data under normal circumstances, that is, μ_0. The Null Hypothesis is tacitly introduced here by setting $\mathbb{E}(X_{n+1} - \mu_0) = 0$. Also observe the less intuitive drift term $-(\mu_0 - \mu_1)^2/2\sigma^2$.

During a reflection-free phase (no resetting to 0), assuming that $X_n \sim p_0$, this process will be rewritten as

$$U^{n+1} = U^n + b + sV^{n+1}, \quad U^0 = 0, \tag{7}$$

where

$$\begin{aligned} b &= -\frac{(\mu_0 - \mu_1)^2}{2\sigma^2}, \quad s = \frac{\mu_1 - \mu_0}{\sigma}, \\ V^{n+1} &= \frac{X_{n+1} - \mu_0}{\sigma} \Rightarrow V^{n+1} \sim \mathcal{N}(0, 1). \end{aligned} \tag{8}$$

It should already be clear from Eq. 7 that it is natural to approximate the CPD statistic between reflection and absorption with an Itô process:

$$dU^t = \boldsymbol{\beta}(U^t)dt + \boldsymbol{\sigma}(U^t)dB_t, \quad U^0 = x \in \mathcal{D}, \tag{9}$$

where the initial condition x is extended to the domain $\mathcal{D} = (0, h)$. Identification of the discrete and continuous time processes yields

$$\boldsymbol{\beta}\, dt = b, \quad \boldsymbol{\sigma}^2\, dt = s^2, \tag{10}$$

where dt is the inverse of the sampling rate. The relationship between the drift terms should be obvious; on the other hand, the identification of the incremental processes requires the counter-intuitive feature $\mathbb{E}(dB)^2 = dt$ of the Itô process.

Conversely and somewhat more formally, this identification can be justified by the Euler-Maruyama [16] and the Milstein [18] methods. The Euler-Maruyama method applied to (9) resembles closely the standard Euler method for ODEs but takes into account the extra stochastic term:

$$U^{n+1} = U^n + \boldsymbol{\beta}(t_n, U^n)\Delta t + \boldsymbol{\sigma}(t_n, U^n)\Delta B_n. \tag{11}$$

For the Gaussian p_0 and p_1 cases, the recursive form of the CUSUM change-point detection as stated in (6) resembles closely the Euler-Maruyama discretization expression in (11). Similar to (10), we have

$$\boldsymbol{\beta}(t_n, U^n)\Delta t = b, \quad \boldsymbol{\sigma}(t_n, U^n)\sqrt{\Delta t} = s. \tag{12}$$

Meanwhile, Milstein's method is the second-order numerical approximation for a stochastic differential equation (SDE):

$$\begin{aligned} U^{n+1} = U^n &+ \boldsymbol{\beta}(t_n, U^n)\Delta t + \boldsymbol{\sigma}(t_n, U^n)\Delta B_n + \\ &\frac{1}{2}\boldsymbol{\sigma}(t_n, U^n)\frac{\partial \boldsymbol{\sigma}(t_n, U^n)}{\partial U}[(\Delta B_n)^2 - \Delta t]. \end{aligned} \tag{13}$$

Note that from Itô's formula, we have the relation between the Brownian motion and time increments as $dB_n^2 = dt$. For the simulation, we use $\Delta B_n = z_n\sqrt{\Delta t}$, where z_n is sampled from a standard normal distribution $\mathcal{N}(0, 1)$. Interestingly, note that for state-independent noise, i.e. $\partial\boldsymbol{\sigma}(t_n, U^n)/\partial U = 0$, the two methods results in the same discretization solution.

2.4 Unknown "abnormal" density p_1

While it is fair to assume that the "normal" regime density p_0 is known, the same cannot be said for the "abnormal" regime density p_1. The latter is modeled as a distribution $f_\theta(x)$ parameterized by θ, and θ is adjusted consistently and *efficiently* with a random sampling $\mathbf{x} = \{x_i\}$ of p_1. Efficiency is the concept of sufficient statistic. A statistic $\sum_i S(x_i)$ is said to be *sufficient* for θ if the conditional distribution $f_\theta(\mathbf{x}|s)$ of the observations given $s = \sum_i S(x_i)$ is independent of θ; in other words, all pertinent information regarding θ is embedded in the sufficient statistic.

It is convenient to assign p_1 the general Koopman-Darmois exponential family of densities [6], [14]

$$f_\theta(x) = g(x)\, \mathrm{e}^{\theta S(x) - A(\theta)}, \tag{14}$$

where $S(x)$ defines a sufficient statistic, θ is the natural parameter, and $A(\theta)$ is a strictly upward concave normalization factor to make $f_\theta(x)$ a *probability* measure. For reason

explained soon, we will choose the Weibull distribution, which with fixed shape β and adjustable scale η parameters,

$$f_\theta(x) = \tfrac{\beta}{\eta}\left(\tfrac{x}{\eta}\right)^{\beta-1} \mathrm{e}^{-\left(\frac{x}{\eta}\right)^\beta}, \tag{15}$$

is a Koopman-Darmois distribution with sufficient statistic $S(x) = x^\beta$ and natural parameter $\theta = -1/\eta^\beta$.

2.5 Novelty: Simultaneous detection and estimation

In case of an imprecisely known abnormal density, instead of (2), we proceed from the *double* maximization [3, (§2, §4.3.1)],

$$U^n = \left\{ \max_{0\le\lambda\le n} \max_\theta Z_\lambda^n(\theta) \right\}_+, \tag{16}$$

where in the definition of $Z_\lambda^n(\theta)$, p_1 is replaced by f_θ. Again, in this "worst case" scenario, especially in case of stealthy attack [24] [31] [34], another justification for $\max_\theta$ is to assume that the density f_θ is *the worst possible given the data*. Whatever the motivation, the detection rule remains $U^n \ge h$, but with U^n now defined by (16) instead of (2).

The problem is that this approach does not easily lend itself to a recursive formulation. Indeed, it is easily seen that the heuristically defined statistic

$$\widehat{U}^{n+1} = \max\left\{0, \widehat{U}^n + \max_\theta \log\frac{f_\theta(X_{n+1})}{p_0(X_{n+1})}\right\}, \widehat{U}^0 = 0 \tag{17}$$

would only dominate the true statistic in the sense that $\widehat{U}^n \ge U^n$ and give overly conservative results with high false alarm rate, because the $\max_\theta$ is taken over every log-likelihood ratio rather than the sum of them. A remedy is to smooth over the $\arg\max_\theta$ by combining the last one at time $n+1$ with the previous ones:

$$\begin{aligned} \widetilde{U}^{n+1} &= \max\left\{0, \widetilde{U}^n + \log\frac{f_{\widetilde{\theta}^{n+1}}(X_{n+1})}{p_0(X_{n+1})}\right\}, \ \widetilde{U}^0 = 0, \\ \widetilde{\theta}^{n+1} &= \kappa\widetilde{\theta}^n + (1-\kappa)\arg\max_\theta \log\frac{f_\theta(X_{k+1})}{p_0(x_{k+1})}, \end{aligned} \tag{18}$$

where $0 < \kappa < 1$ is some smoothing gain.

Among the Koopman-Darmois distributions, the Weibull distribution has been retained since it is the distribution that needs the least amount of data to be correctly identified [7]. Therefore, p_1 will be correctly identified in the shortest amount of time, which is crucial in such a security application where time is the essence.

2.6 Time complexity

The computational complexity of the CUSUM procedure (3) is constant $O(1)$ at each iteration. When monitoring for a change-point in the PMU frequency Hurst exponents, the Hurst exponent calculation by Detrended Fluctuation Analysis (DFA) calculation is $O(N \log N)$, where N is the length of the time window considered at each iteration. Hence, the algorithmic strategy can be computed in real time.

3 Time-to-Alarm

Here, we address the mathematical questions that underline the time-to-event. For the sake of conciseness, proofs are sketched with details relegated to [32].

Theorem 1. *The average time to the threshold crossing event, $T(x) := \mathbb{E}_{p_0}(\min\{t \,|\, U^t > h\} \,|\, U^0 = x)$, is given by the following ordinary differential equation (ODE) subject to mixed Dirichlet-Neumann boundary conditions [22, §1.2], [33, §8.2]:*

$$\left(\frac{1}{2}\boldsymbol{\sigma}^2\frac{\partial^2}{\partial x^2} + \boldsymbol{\beta}\frac{\partial}{\partial x}\right)T(x) = -1, \tag{19a}$$

$$T(h) = 0, \quad \text{(Dirichlet)}, \tag{19b}$$

$$\left.\frac{\partial T(x)}{\partial x}\right|_{x=0} = 0, \quad \text{(Neumann)}. \tag{19c}$$

The complete proof is available in [32], but to make the exposition somewhat self-contained, we provide an outline. Following in the footsteps of [5], let $p(x,y,t)$ be the transition probability from the initial condition density $p(U^0) = \delta_x(U^0)$, a unit mass at x, to the density at a future time $p(U^t = y)$. As is well known, p is solution to the Fokker-Planck or Kolmogorov Forward Equation (KFE) [9, §X.5], [19, §8, p. 153]. From this transition, it is easily seen that $\int_0^h p(x,y,t)dy =: \pi(x,t)$ is the probability that the process starting at x has not yet reached h by time t. From there, the probability that the CUSUM process crosses the threshold between times t and $t+dt$ is seen to be $\pi(x,t) - \pi(x,t+dt) = -d\pi$. From there on, using integration by part and $\lim_{t\to\infty} p(x,y,t) = 0$, we get

$$T(x) = \mathbb{E}_{-\frac{\partial\pi}{\partial t}}(t) \tag{20a}$$

$$= \int_0^h \underbrace{\left(\int_0^\infty p(x,y,t)dt\right)}_{-G(x,y)} dy. \tag{20b}$$

$G(x,y)$ is the *Green function* of the Kolmogorov equation also subject to Dirichlet-Neumann boundary conditions:

$$\left(\frac{1}{2}\boldsymbol{\sigma}^2\frac{\partial^2}{\partial y^2} - \beta\frac{\partial}{\partial y}\right)G(x,y) = \delta(x-y), \tag{21a}$$

$$G(x,h) = 0, \quad \text{(Dirichlet)}, \tag{21b}$$

$$\left.\frac{\partial G(x,y)}{\partial y}\right|_{y=0} = 0, \quad \text{(Neumann)}. \tag{21c}$$

From here on, the development departs significantly from [5]. A delicate argument, developed in details in [32], shows that, up to a good approximation for small drifts, $G(x,y)$ depends on the absolute difference of arguments, that is, $G(x,y) = g(|x-y|)$, where $g \in W^{2,2}$, the Sobolev space of twice differentiable square integrable functions [15]. Therefore, $G(x,y) = G(y,x)$ and $\partial G/\partial y = -\partial G/\partial x$. Changing the partial differential operators in the Green function PDE, integrating both sides of Eq. (21)(a) as $\int_0^h(\cdot)dy$, and defining $T(x) = -\int_0^h G(x,y)dy$ yields the ODE for $T(x)$. The transfer of the Dirichlet-Neumann conditions from $G(x,y)$ to $T(x)$ is more delicate. Regarding Dirichlet, $G(x,h) = 0, \forall x$ yields $G(h,y) = 0, \forall y$ and thereofore $T(h) = 0$. Regarding the Neumann condition,

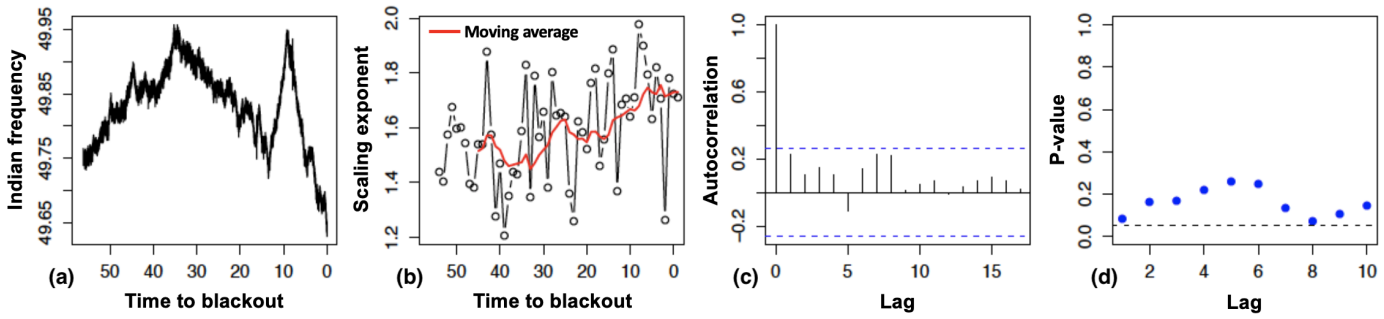


Fig. 1: (a) Indian blackout frequency time series, (b) scaling (Hurst) exponent time series obtained using the DFA procedure with 1 min time window and 1 min time shift, (c) autocorrelation, and (d) Ljung-box test for the Hurst exponent time series.

$T(x) = \int_0^h G(y,x)dy$ yields

$$T'(0) = \int_0^h \underbrace{\frac{\partial G(y,x)}{\partial x}\Big|_{x=0}}_{=0} dy = 0.$$

Simulation studies of Sec. 5 show that the proposed approximation is perfectly acceptable.

Corollary 1 ([32]). *The solution to the differential equation for $T(x)$ over the enlarged domain $(-\epsilon, 0)$ subject to the boundary condition $T'(-\epsilon) = 0$ is given by*

$$T(x) = -\frac{2x + e^{-\frac{2\boldsymbol{\beta}x}{\boldsymbol{\sigma}^2}}\boldsymbol{\sigma}^2 c_1 - 2\boldsymbol{\beta}c_2}{2\boldsymbol{\beta}}, \tag{22}$$

where the integration constants c_1, c_2 are evaluated from the boundary conditions:

$$c_1 = \frac{1}{\boldsymbol{\beta}} e^{-\frac{2\boldsymbol{\beta}\epsilon}{\boldsymbol{\sigma}^2}}, \quad c_2 = \frac{2h\boldsymbol{\beta} + e^{-\frac{2\boldsymbol{\beta}(h+\epsilon)}{\boldsymbol{\sigma}^2}}\boldsymbol{\sigma}^2}{2\boldsymbol{\beta}^2}.$$

Corollary 2 ([32]). *$T(x)$ is continuous as $\epsilon \downarrow 0$, $\forall x \geq 0$.*

Finally, setting $x = 0$ and $\epsilon = 0$ yields

$$T(0) = \frac{h}{2\boldsymbol{\beta}}\left(2 + \frac{\boldsymbol{\sigma}^2}{\boldsymbol{\beta}h}\left(e^{-\frac{2\boldsymbol{\beta}h}{\boldsymbol{\sigma}^2}} - 1\right)\right). \tag{23}$$

Upon second order approximation of $\exp(-2\beta h/\boldsymbol{\sigma}^2)$, the time $T(0)$ to alarm and hence an approximate lower bound on the threshold are easily found to be

$$T(0) \approx \frac{h^2}{\boldsymbol{\sigma}^2}, \quad h \geq \frac{\boldsymbol{\sigma}}{\sqrt{\overline{\text{FAR}}}}. \tag{24}$$

Remark 1. *The proof of Th. 1 is a generalization of [5, Th. 1] to the case where the drift $\boldsymbol{\beta}$ is nonvanishing. Moreover, here, the proof is based on the* explicit *construction of the transition probability $p(x, y, t)$, with the $\lim_{t\to\infty} p(x, y, t)$ convergence different from the one of [5].*

Remark 2. *The crucial approximation $G(x, y) \approx g(|x-y|)$ and its justification in [32] are believed to be new. The passage from the Dirichlet-Neumann conditions on $G(x, y)$ to those on $T(x)$ is believed to be new compared with [5].*

Remark 3. *The crucial approximation can be justified as follows: Since the medium is isotropic ($\boldsymbol{\beta}$ and $\boldsymbol{\sigma}$ independent of U^t), $g(|x - y|)$ is only an approximation of $G(x, y)$ because the boundary conditions are not symmetric. Because of the negative drift term, the density $p(U^t)$ is "pulled" to the left (the "head") while the "tail" of the distribution is less affected and the alarm is mostly triggered by the tail.*

Remark 4. *Eq.* (23) *is available in [3, Eq. 3.1.105], but proved via moment generating methods. However, here, Eq.* (23) *is derived from the general Eq.* (22)*. The latter is of interest for change-point detection as $T(U^t)$ tells the operator in how much time the alarm is expected to ring.*

4 Hurst exponent window of PMU frequency time series

Prior works show that PMU time series exhibit long-range memory and fractal characteristics that are quantified by the Hurst exponent [27], [29]. Moreover, an increasing trend in the Hurst exponent of the frequency is a good indicator, consistently with the ROCOF [29], of proximity of power system to blackout [28].

Detrended Fluctuation Analysis (DFA) [21] is one of the most reliable and robust methods to calculate the Hurst (scaling) exponent. The DFA procedure to compute the Hurst exponent series could lead to correlated and non-independent Hurst exponent samples especially when the PMU data windows have a large overlap as in [31]. To mitigate this issue, we consider reducing the size of the moving window and minimizing the amount of overlap between consecutive windows.

In the present paper, *contrary to* [31], the frequency Hurst exponents have been calculated inside a moving window of length 3,000 samples (1 minute) and a 1 minute time shift making the windows non-overlapping and hopefully generating independent Hurst exponent samples. We apply the DFA on the frequency series before the 2012 Indian blackout as shown in Fig. 1(a). The resulting Hurst exponents from the Indian frequency data are shown in Fig. 1(b). Since we aim to generate i.i.d. samples of Hurst exponents by reducing the overlap between windows, this results in a much smaller amount of Hurst exponent samples (56 samples). On the positive side, Fig. 1(c) shows that the autocorrelation function of the Hurst exponent series up to 20 samples is within the 95% confidence band (dotted line) for independence. From the individual autocorrelations, we can infer that the samples of Hurst exponents are uncorrelated.

Finally, we use the Ljung-Box test to verify the independence of the Hurst exponent samples. The test provides a more rigorous way to test the autocorrelations at multiple lags jointly. The Null hypothesis (H_0) is that the samples of Hurst exponent series are uncorrelated and independent. As shown in Fig. 1(d), the p-values of Ljung-Box statistic are higher than the significance level of 0.05 (dotted line), so we cannot reject the Null hypothesis (H_0), i.e. we do not have evidence that the samples are autocorrelated.

5 Change-Point Detection of PMU Frequency Anomaly

In this section, we will utilize the derivation from Sec. 2 and empirical results from Sec. 4 to implement the change-point detection on the PMU frequency data obtained before the Indian blackout, which occurred on the 30th and 31st of July 2012 [1]. Fig. 2(i-a) shows that the frequency time series collected before the blackout has a length of 167,600 samples (sampled at 50 samples/second) and spans approximately 56 minutes.

On the same Figs. 2(i-d),(iii,d) we also plot the time-to-alarm. It is noted that this time-to-alarm, $T(U^n)$, is slaved on U^n and reset to $T(0)$ when U^n is reset to 0. Also note that, for the sake of the argument, U^n is allowed to run beyond the threshold h where it is observed that $T(U^n) < 0$, where a negative time-to-event means that the threshold has already been crossed. As long as the threshold has not been crossed, $T(U^n) > 0$ as it should be. This indicates that, while theoretically $T(U^n)$ is just an approximation of the time-to-event, it is still accurate for the application being considered.

We follow Sec. 4 in calculating the Hurst exponent of the frequency data using DFA procedure with non-overlapping moving window length of 3000 samples (1 minute) and also a shift of 1 minute as shown in Fig. 2(i-b).

5.1 Empirical estimate for pre- and post-distributions

In [31], statistical tests on the PMU frequency time series in EPFL campus grid show that the Gaussian distribution is the best distribution to fit the histogram of the Hurst exponent of the frequency during normal operations.

From Sec. 4, we estimate the pre- and post-event distributions, p_0 and p_1, as Gaussian distributions based on the calculated Hurst exponent time series as shown Fig. 2(i-b). Since the true change-point occurrence for the 2012 Indian blackout is not known, we divide the time series into two regimes of operations by setting the cutoff at t = 40 min. (or 16 mins. before the blackout). We estimate empirically the means and standard deviations of the pre- and post-event distributions: the pre-event regime is assume to be within $t \in [0, 40]$ with $(\mu_0, \sigma_0) = (1.5487, 0.1681)$; the post-event regime is assume to be within $t \in [40, 56]$ with $(\mu_1, \sigma_1) = (1.7116, 0.1758)$.

From Sec. 2.3, we follow the recursive change-point detection formulation (6) for the Gaussian p_0 and p_1 case. Fig. 2(i-c) shows the change-point CUSUM statistic U^n at each time step based on the estimated p_0 and p_1 distributions. For an acceptable false alarm rate of $\overline{\text{FAR}} = 0.1$, the threshold value is set at $\underline{h} = 2.05$ with the U statistic crossing at $n = 41$ min. (15 mins. before the blackout). Fig. 2 (i-d) shows the mean time to alarm event $T(U^n)$.

The setting up of the threshold value to raise the alarm and the calculation of $T(U^n)$ are discussed in more detail at the next section.

5.2 Threshold for False Alarm Rate—Itô approach

TABLE 1: Summary of the crossing times corresponding to different admissible false alarm rates $\overline{\text{FAR}}$.

$\overline{\text{FAR}}$	threshold $\underline{h}$	crossing time τ
0.2	1.5990	37 min.
0.1	2.0470	41 min.
0.01	3.9499	47 min.
0.001	6.1674	49 min.

Here, with the objective of deriving the FAR, we follow Section 3 and compute the expected value of the first time the Itô process (9), starting from $U^0 = 0$, properly scaled as a model of (7), crosses the threshold h. Starting from (24) and using the identification recipes (10), and (8), it is found that

$$\mathbb{E}_{p_0}(\tau(h)) := T(0) = 2\left(\frac{\sigma}{\mu_1 - \mu_0}\right)^2 \left(e^h - (h+1)\right) dt \approx \left(\frac{\sigma}{\mu_1 - \mu_0}\right)^2 h^2\, dt. \quad (25)$$

To calculate the appropriate threshold $\underline{h}$, we follow the relationship (5) between the admissible upper bound for the false alarm rate $\overline{\text{FAR}}$ and the first crossing time of the Itô process $\mathbb{E}_{p_0}(\tau(h))$.

Using the same setup as in Sec. 5.1, Fig. 2(ii) shows the threshold values and the corresponding threshold crossing times for different false alarm rates. A larger $\overline{\text{FAR}}$ results in a lower threshold h and vice versa. In agreement with intuition, choosing a larger $\overline{\text{FAR}}$ results in a faster alarm trigger but at the cost of higher incidence of false alarms. Conversely, setting a lower $\overline{\text{FAR}}$ results in smaller chances of false alarm triggers but at the cost of a slower response to a true alarm. A summary of results is shown in Table 1.

5.3 Comparison between overlapping and nonoverlapping moving windows for DFA

The DFA procedure uses a PMU moving time window to estimate the Hurst exponent samples. The chosen window parameters (window length, number of box sizes per window, and window time shift) affect the accuracy of the Hurst exponent estimate. A longer window length and greater number of boxes generally increase the accuracy of the Hurst exponent estimate but at the cost of higher computational complexity. Additionally, choosing a smaller window time shift generates more data samples. If the window time shift is chosen too small compared to the window length, this could lead to overlapping moving windows that could lead to correlated data samples.

The results in Sec. 5.1 are obtained using non-overlapping moving window setup for the DFA procedure (as explained in Sec. 4) so as to satisfy the i.i.d. assumption on the data samples. We now compare this result to a DFA setup with overlapping moving window of length 49,400

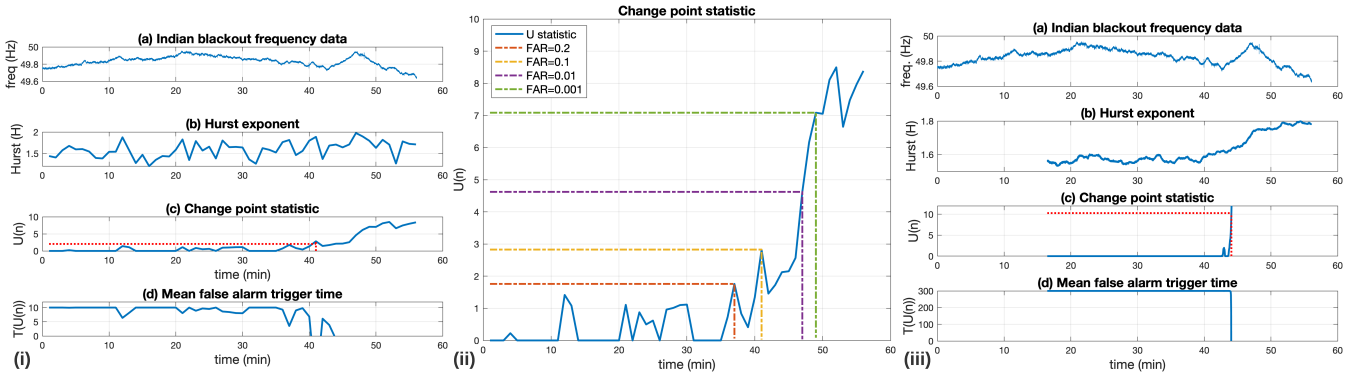


Fig. 2: 2012 Indian blackout change-point detection on the frequency Hurst exponent data. (i) Hurst exponent data obtained from *non-overlapping* time windows (1 min. window and 1 min. time shift), upper admissible false alarm rate $\overline{\text{FAR}} = 0.1$ and $\underline{h} = 2.05$ (red dashed-line), (ii) comparison of different $\overline{\text{FAR}}$ values and corresponding $\underline{h}$ values (dashed-lines), (iii) Hurst exponent data as obtained in [31] from an *overlapping* moving window of length 49,400 samples (16.5 minutes) and shift of 100 samples (0.033 mins), $\overline{\text{FAR}} = 0.1$ and $\underline{h} = 9.194$ (red dashed-line).

samples (16.5 minutes) and shift of 100 samples (0.033 mins) as in [31]. Fig. 2(iii-b) shows the Hurst exponent samples given the overlapping window setup. Similarly, the estimated pre- and post- distribution parameters in [31] are $(\mu_0, \sigma_0) = (1.5722, 0.0198)$ and $(\mu_1, \sigma_1) = (1.7327, 0.0582)$. Figs. 2(iii-c) and (iii-d) show the change-point CUSUM statistic U^n and the mean false alarm trigger time $T(U^n)$, respectively, at each time step based on the estimated p_0 and p_1 distributions. Choosing the same false alarm rate at $\overline{\text{FAR}} = 0.1$, the threshold value is at $\underline{h} = 9.19$, which resulted in no false alarms. The alarm is raised at $t = 44.07$ min. (11.93 mins. before the blackout).

The non-overlapping window setup with $dt = 1$ min. resulted in significantly fewer data samples ($n = 56$) compared to the overlapping window setup with $dt = 0.033$ min. which generated 1,190 samples. As a result, the case of non-overlapping window has a higher volatility in the Hurst exponent samples and the change-point statistic (compare Figs. 2(i-b) and (iii-b) and their corresponding standard deviations).

Comparison between Figs. 2(i) and 2(iii) makes a *strong case* for the CUSUM approach. In case of overlapping windows, the increase of the Hurst exponent is visually obvious and the CUSUM is just confirming this. However, in case of nonoverlapping windows, the trend in Fig. 2(i)(b) is not obvious but revealed by CUSUM in Fig. 2(i)(c).

Using the same admissible false alarm rate $\overline{\text{FAR}} = 0.1$ for both cases, the non-overlapping window setup has the crossing time occurring at $t = 41$ min. (compared to 44.77 min. for the overlapping window case). While it is worth noting that the non-overlapping case results in earlier true alarm trigger and no false alarm triggers, there are a few instances that the change-point statistic has come close to the threshold. In addition, the mean false alarm trigger time $T(0)$ is lower for the non-overlapping case.

5.4 Unknown post-distribution

For this case, the post distribution p_1 is assumed to be unknown and is assigned the Weibull distribution with natural parameter $\theta = -1/\eta^\beta$ as suggested in Sec. 2.4, Eq. (15). For a fixed shape β, the distribution is parameterized only by the scaling η as $f_\eta(x)$, where η is related to the mean as $\mathbb{E}(x) = \eta\Gamma(1 + 1/\beta)$.

Given this assignment, we use the recursive form of the change-point algorithm for the simultaneous detection and estimation as defined by (18). For the Weibull distribution, the $\arg\max_\eta$ term in (18) results in

$$\arg\max_\eta \log \frac{f_\eta(X)}{p_0(X)} = \arg\max_\eta f_\eta(X).$$

Taking the derivative with respect to η and equating to zero results in a simple expression $\eta_{\max} = x$, where x is the *numerical value* recorded. As such, η is re-adjusted every single step, which creates some oscillation in the statistic, which can be smoothed over by virtue of (18)(b).

Fig. 3(i) shows the PDF plots for the empirical estimates for p_0 and p_1, as used in Sec. 5.1, and the unknown PDF taking the form f_η. By setting $\eta_{\max} = X$, the *variable* rather than the numerical value, and plugging this in $f_\eta(X)$ yields $(\beta/X)\,\mathrm{e}^{-1}$, the envelope of the Weibull distributions.

One can choose the shape parameter β so that the distributions have higher peak values. However, setting β too high can lead to higher false alarm rates. Additionally, the range of values for the scaling parameter η, related to the mean, is limited to a minimum value of 1.65. By setting the minimum value for η, this ensures that f_η would not completely overlap with the known pre-event distribution p_0.

Fig. 3(ii) shows the change-point statistic U^n and the mean false alarm trigger time $T(U^n)$ for the simultaneous detection and estimation algorithm with $\beta = 7$ for the shaping parameter for the non-overlapping window case. For the same threshold value $\underline{h} = 2.05$, the algorithm resulted in 2 false alarms before it correctly raised the alarm at $t = 41$ min. (15 mins. before the blackout). Similar simulations were run for higher values of β and these resulted in an even higher numbers of false alarms. Setting β too low results in more sluggish response in raising the true alarm.

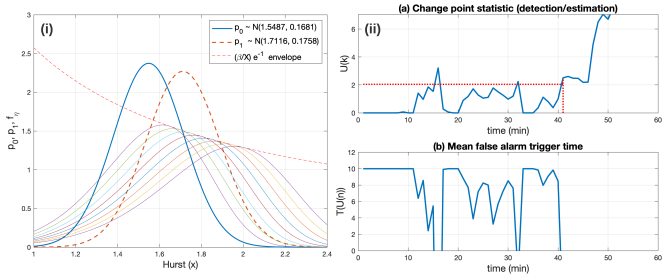


Fig. 3: (i) Pre- and post-event Gaussian distributions p_0 and p_1, estimated post-event family of Weibull distributions $f_\eta(x)$ with $\beta = 7$, and envelope of Weibull distributions. (ii) Change-point detection for simultaneous detection and estimation using p_0 and f_η as in (i) with $\overline{\text{FAR}} = 0.1$ and $\underline{h} = 2.05$ (red dashed-line).

5.5 First hitting time distribution for Type I error analysis

The probability that the data is misclassified as a rejection of the Null Hypothesis is computed from the distribution of the first threshold hitting time $T(0)$ consistently with an acceptable significance level. Should the data of the threshold crossing time generated by the CUSUM process be much to the left tail of $p_{T(0)}$, the probability of an erroneous rejection of the Null Hypothesis, a Type I error, is very low.

We conduct a Monte Carlo experiment to estimate empirically the first hitting time distribution [10] by running the process for $m = 10000$ trials and observing the instance that the change-point statistic U^n hit the threshold h for the first time. Experiment 1 resembles the setup for the overlapping Hurst exponent time window case with finer time steps [31]; the system parameters are as follows: sampling time $dt = 0.033$, $\mu_0 = 1.5722$, $\mu_1 = 1.7327$, $\sigma = 0.198$; the process drift and diffusion coefficients are $b = -0.33$ and $s = 0.81$ with a corresponding theoretical mean escape time $T(0) = 9.43$ as calculated from (23). The experimental mean escape time is 25.30. Figure 4(i) shows the first hitting time distribution for $h = 4.6$ and the five best distribution fittings. Both the gamma and exponential distributions result in the best fit (based on the SSE, AIC and BIC criteria). The resulting exponential distribution has mean parameter $\mu = 25.30$ $[24.81, 25.80]$. Additionally, we perform the 2-sample Kolmogorov-Smirnov (KS) test [17] between the empirical distribution and sampled exponential distribution. The resulting p-value of 0.2185 failed to reject the null hypothesis that the two sample distributions come from the same distribution.

We perform another Monte Carlo test (Experiment 2) for a larger sampling time $dt = 1$ resembling the non-overlapping time window case with the following system parameters similar to Sec. 5.1: $\mu_0 = 1.5487$, $\mu_1 = 1.7116$, $\sigma = 0.1681$; the resulting process drift and diffusion coefficients are $b = -0.47$ and $s = 0.97$ with a corresponding theoretical mean escape time $T(0) = 10.05$. Figure 4(ii) shows the first hitting time distribution for $h = 2.05$ with an average escape time of 41.25. The exponential and chi-squared are the best distribution fit. The resulting exponential distribution has mean parameter $\mu = 41.25$ $[40.45, 42.07]$.

	SSE	AIC	BIC
expon	0.0000	1,528.0760	-201,735.0363
gamma	0.0000	1,540.5422	-202,637.1384
exponpow	0.0002	1,789.4374	-175,481.6682
lognorm	0.0003	1,425.9009	-172,493.0399
alpha	0.0007	1,435.2831	-164,652.7068

	SSE	AIC	BIC
chi2	0.0000	1,758.0988	-192,035.0040
expon	0.0000	1,776.5707	-198,492.5345
erlang	0.0000	1,807.0330	-198,905.5504
gamma	0.0001	1,748.8264	-186,568.1270
exponpow	0.0001	2,087.1781	-183,208.2792

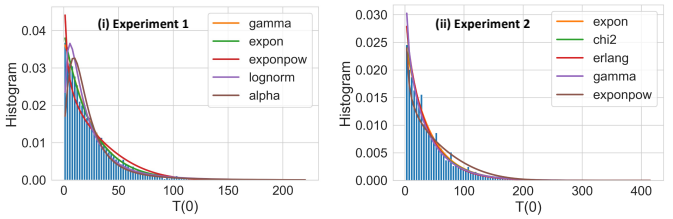


Fig. 4: First hitting time histograms and fitted distributions for Experiments 1 & 2. The exponential distribution is the best fit for both experiments.

6 Conclusion

In this paper, the change-point detection of PMU frequency anomalies that could indicate imminent voltage collapse was explored. More specifically, the detection is based on the CUSUM of the log-likelihood ratios of the deviations of the Hurst exponent from its baseline statistic. The case of unknown abnormal distribution of the Hurst exponent was dealt with by simultaneously running the CUSUM and identifying the scale parameter of the Weibull distribution model. In addition to the alarm triggered when danger is imminent, the operator is provided with a time to the alarm event. Retrospective analysis of the 2012 Indian blackout data has shown that the change-point detection algorithm could have anticipated the voltage collapse as early as 12 minutes before the blackout event.

The method could be improved and extended in several directions. Instead of artificially making the Hurst samples i.i.d. by spacing the observations with the drawback of a reduced data set, an Auto-Regressive (AR) model of the Hurst exponent sequence would be more appropriate. Also left for further research is the change-point detection of false data injection [23], [24], [25] or other security threats.

Lastly, the Type II error analysis could be developed based on the histograms of the distribution of the first hitting time distribution when the data driving the CUSUM is drawn from f_θ rather than p_0.

References

[1] Report on the Grid Disturbance on 30th July 2012 and Grid Disturbance on 31st July 2012. Technical report, Central Agency Regulatory Commision, 2012.

[2] A. Barua, D. Muthirayan, P. P. Khargonekar, and A. A. Al Faruque. Hierarchical temporal memory based one-pass learning for real-time anomaly detection and simultaneous data prediction in smart grids. *IEEE Transactions on Dependable and Secure Computing*, 19(3):1770–1782, 2022.

[3] Michèle Basseville and Igor V. Nikiforov. *Detection of Abrupt Changes - Theory and Application*. Englewood Cliffs, New Jersey, 2010.

[4] R. Blazek, H. Kim, B. Tozovskii, and A. Tartakovsky. A novel approach to detection of 'denial-of-service' attacks via adaptive sequential and batch sequential change-point detection. In *Proceedings of the 2001 IEEE Workshop on Information Assurance and Security*, pages 220–226, U.S. Military Academy, West Point, NY, June 5-6 2001.

[5] C. Caginalp and X. Chen. Analytical and nuerical results for an escape problem. *Arch. Rational Mech. Anal.*, 203:329–342, 2012.

[6] G. Darmois. Sur les lois de probabilité à estimation exhaustive. *Comptes Rendus de l'Académie des Sciences, Paris*, 200:1265–1266, 1936.

[7] M. Martin del Campo and E. Jonckheere. Bursting rate variability. *Frontiers in Physiology*, 12, December 2021.

[8] Zhaoyang Dong and Pei Zhang. *Emerging Techniques in Power System Analysis*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[9] W. Feller. *An introduction to probability theory and its applications—volume 2*. John Wiley & Sons, 1971.

[10] N Haydn, Y. Lacroix, and S. Vaienti. Hitting and return times in ergodic dynamical systems. *Annals of Probability*, 33(5):2043–2050, September 2005.

[11] Joe-Air Jiang, Jun-Zhe Yang, Ying-Hong Lin, Chih-Wen Liu, and Jih-Chen Ma. An Adaptive PMU Based Fault Detection/Location Technique for Transmission Lines - Part I : Theory and Algorithms. *IEEE Transactions on Power Delivery*, 15(4):486–493, 2000.

[12] Do In Kim, Tae Yoon Chun, Sung Hwa Yoon, Gyul Lee, and Yong June Shin. Wavelet-based event detection method using PMU data. *IEEE Transactions on Smart Grid*, 8(3):1154–1162, 2017.

[13] H. Kim, B. L. Rozovskii, and A. G. Tartakovsky. A nonparametric multichart CUSUM test for rapid detection of DOS attacks in computer networks. *International Journal of Computing & Information Sciences*, 2(3):149–158, December 2004.

[14] B. O. Koopman. On distributions admitting a sufficient statistic. *Trans. of the American Mathematical Society*, 39(3):399–409, May 1936.

[15] L. A. Lusternik and V. J. Sobolev. *Elements of Functional Analysis*. Hindustan Publishing Corporation, Delhi, 1974.

[16] Gisiro Maruyama. Continuous Markov processes and stochastic equations. *Rendiconti del Circolo Matematico di Palermo*, 4(1):48–90, 1955.

[17] Frank J. Massey Jr. The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951.

[18] G. Mil'shtejn. Approximate Integration of Stochastic Differential Equations. *Theory of Probability & Its Applications*, 19(3):557–562, 1975.

[19] B. Oksendal. *Stochastic Differential Equations*. Springer, Berlin, 1995.

[20] E. Page. Continuous inspection schemes. *Biometrika*, 41:100–115, 1954.

[21] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger. Mosaic organization of dna nucleotides. *Phys. Rev. E*, 49:1685–1689, Feb 1994.

[22] M. Renardy and R. C. Rogers. *An introduction to partial differential equations*, volume 13 of *Texts in Applied Mathematics*. Springer, 1993.

[23] H. Sedghi and E. Jonckheere. Statistical structure learning of smart grid for detection of false data injection. In *Power and Energy Society General Meeting (PES), 2013 IEEE*, pages 1–5, July 2013.

[24] H. Sedghi and E. Jonckheere. Statistical structure learning to ensure data integrity in smart grid. *IEEE Transaction on Smart Grid*, 6(4):1924–1933, 2015. Available at http://eudoxus2.usc.edu.

[25] Hanie Sedghi and Edmond Jonckheere. On the conditional mutual information in the Gaussian–Markov structured grids. In Giacomo Como, Bo Bernhardsson, and Anders Rantzer, editors, *Information and Control in Networks*, volume 450 of *Lecture Notes in Control and Information Sciences*, pages 277–297. Springer International Publishing, 2014. Available at `http://eudoxus2.usc.edu`.

[26] Khushboo Shah, Edmond Jonckheere, and Stephan Bohacek. Dynamic modeling of internet traffic for intrusion detection. *EURASIP Journal on Advances in Signal Processing*, 2007, 2007.

[27] L. Shalalfeh, P. Bogdan, and E. Jonckheere. Evidence of long range-dependence in power grid. In *Power and Energy Society General Meeting (PESGM)*, Boston, MA, July 2016. Available at `http://ee.usc.edu/ jonckhee`.

[28] L. Shalalfeh, P. Bogdan, and E. Jonckheere. Kendall's tau of frequency Hurst exponent as blackout proximity margin. In *IEEE International Conference on Smart Grid Communications*, Sydney, Australia, November 06-09 2016. 978-1-5090-4075-9/16; available at `http://ee.usc.edu/ jonckhee`.

[29] L. Shalalfeh, P. Bogdan, and E. A. Jonckheere. Fractional dynamics of PMU data. *IEEE Transactions on Smart Grid*, 12(3):2578–2588, May 2021. doi: 10.1109/TSG.2020.3044903.

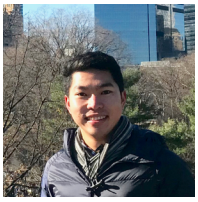
[30] A. N. Shiryaev. On optimum methods in quickest detection problems. *Theory Probab. Appli.*, pages 22–46, 1963.

[31] J. Sia, E. Jonckheere, L. Shalalfeh, and P. Bogdan. PMU Change Point Detection of Imminent Voltage Collapse and Stealthy Attacks. In *2018 IEEE Conference on Decision and Control*, pages 6812–6817, 2018.

[32] J. Sia, E. Jonckheere, L. Shalalfeh, and P. Bogdan. Stochastic calculus of change point detection with time-to-event. available as supplemental material, 2022.

[33] M. E. Taylor. *Partial Differential Equations II—Qualitative Studies of Linear Equations*, volume 116 of *Applied Mathematical Sciences*. Springer, New York, Berlin, 1997.

[34] A. Teixeira, G. Dan, H. Sandberg, and K H. Johansson. A cyber security study of a SCADA energy management system: Stealthy deception attacks on the state estimator. In *IFAC World Congress*, Sept. 2011.



Jayson Sia received his B.Eng. in Electrical Engineering, B.A. in Economics and M.S. in Electrical Engineering specialization in Control and Automation Engineering from the National University of Singapore. He was awarded the Annenberg PhD fellowship and is currently pursuing PhD in Electrical Engineering at the University of Southern California. His research interest includes control systems, network science and machine learning.

Edmond A. Jonckheere received his Electrical Engineering degree from the University of Louvain, Belgium, in 1973, the Dr.-Eng. in Aerospace Engineering from the Université Paul Sabatier, Toulouse, France, in 1975, and Ph.D. in Electrical Engineering from the University of Southern California in 1978. In 1973-1975, he was a Research Fellow of the European Space Agency. In 1979, he was with the Philips Research Laboratory, Brussels, Belgium. In 1980, he returned to the University of Southern California, where he is a Professor of Electrical Engineering and Mathematics, a member of the Centers for Applied Mathematical Sciences, and a member of the Quantum Information Science and Technology Center. He is a Life Fellow of the IEEE. His research interests include conventional versus quantum control, adiabatic quantum computations, wireless networking, and the power grid.

Laith Shalalfeh received the B.S. degree in electrical engineering from the University of Jordan in 2009, the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California in 2012 and 2017, respectively. Then, he joined the German Jordanian University, where he is currently an Associate Professor with the Energy Engineering Department. His research interests include electric vehicles, load modeling, voltage stability, phasor measurement units, and smart grid.

Paul Bogdan received the PhD degree from Carnegie Mellon University. He is a Jack Munushian Early Career Chair associate professor with the Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California. His research interests include the cyber-physical systems, control of complex time-varying networks, modeling and analysis of biological systems and swarms, new control algorithms for dynamical systems exhibiting multifractal characteristics, modeling biological or molecular communication, fractal mean field games, machine learning, artificial intelligence, performance analysis and design methodologies for manycore systems. His work has been recognized with a number of awards and distinctions, including the 2019 Defense Advanced Research Projects Agency (DARPA) Directors Fellowship, the 2018 IEEE CEDA Ernest S. Kuh Early Career Award, 2017 DARPA Young Faculty Award, 2017 Okawa Foundation Award, 2015 National Science Foundation CAREER Award, 2012 A.G. Jordan Award from Carnegie Mellon University for an outstanding PhD thesis and service, several best paper awards, including the 2013 Best Paper Award from the 18th Asia and South Pacific Design Automation Conference, 2012 Best Paper Award from the Networks-on-Chip Symposium, 2012 D.O. Pederson Best Paper Award from IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2012 Best Paper Award from the International Conference on Hardware/Software Codesign and System Synthesis, and the 2009 Roberto Rocca PhD Fellowship.