Kendall's Tau of Frequency Hurst Exponent as Blackout Proximity Margin

Laith Shalalfeh, Paul Bogdan, and Edmond Jonckheere
Dept. of Electrical Engineering, Univ. of Southern California, Los Angeles, CA 90089
{shalalfe,pbogdan,jonckhee}@usc.edu

***Abstract*—The rapid deployment of Phasor Measurement Units (PMUs) can keep the smart grid in a secure and reliable state. The large amount of data collected from the power grid by PMUs requires new algorithms to detect abnormal and potentially catastrophic events. In this paper, we introduce a novel method to assess the distance to blackout or other instability of the smart grid. Based on the existence of long-range correlation in the PMU data, we exhibit an increase in the frequency Hurst exponent—quantified by Kendall's tau rank correlation coefficient—before the blackout. High Kendall's tau of the frequency Hurst exponent is here proposed as an early-warning signal for blackout.**

I. Introduction

A Phasor Measurement Unit (PMU) is a device that provides synchrophasor and system frequency estimates, as well as other optional information such as calculated megawatts (MW) and megavars (MVAR) [1]. The PMUs have been introduced in the 1990s to support and overcome the drawbacks of the conventional Supervisory Control and Data Acquisition (SCADA) system. The drawbacks are related to the network security issues of the SCADA systems and the asynchronous data arrival due to the transmission delay and low sampling rate (one sample every 2-4 seconds).

The PMUs provide secure data by having their own dedicated communication network. The data is measured at high sampling rate (30-50 samples per second). Moreover, the PMU resolves the issue of data delay by GPS time stamping of the data measurements.

Due to the high sampling rate of the data measured by PMUs, extracting real-time useful information in a timely manner could be a challenge. The detection of events in the power grid using PMUs has been an active area of research [7], [8], [20]. Voltage collapse and more generally power system blackout—either accidental or malicious—are among the most severe events that cause power loss over a wide area of the power system. Anticipating such events is a high priority in smart grid research.

In 2003, one of the largest blackouts in US history hit the Northeastern and Midwestern parts of the United States, and the Canadian province of Ontario. The blackout left 55 million without electricity with total economic cost between \$7 and \$10 billion. Two large consecutive blackouts occurred in the Northern part of the Indian power grid on July 30, 2012 and July 31, 2012. These two blackouts affected 300 million and 600 million, respectively. Unfortunately, the classical loading margin [19] is static and while it has been argued that a dynamic approach is irrelevant [4], here we show that fractal dynamics is relevant and could provide another early warning of blackout.

In [16], we provided an evidence to the existence of long-range memory in the power system data collected using PMUs. We analyzed the data of voltage magnitude (V), frequency (f), and voltage phase angle (θ) collected from different locations in the Texas Synchrophasor Network. Using the Detrended Fluctuation Analysis (DFA) method [14], we were able to show that the voltage magnitude, frequency, and phase angle have scaling (Hurst) exponents higher than 0.5.

In the present paper, we first study the scaling properties of 2079 medium voltage (12 kV) PMU data sets (100,000 samples each) collected over four months from the École Polytechnique Fédérale de Lausanne (EPFL) campus network. We apply the DFA method on the voltage magnitude, frequency, and phase angle data sets to show the strong consistency of the scaling exponents of 120V and 69 kV PMU data collected from the Texas Synchrophasor Network [16] and the EPFL campus network. From our earlier work [16], it appears that scaling properties of the *frequency* do not depend on the voltage.

Secondly, we introduce a new method based on the change of fractal characteristics of the frequency data before the power system goes to blackout. We compare the Hurst exponents of frequency data collected from the EPFL campus network and frequency data collected before the 2012 Indian blackout. The increase in the Hurst exponent of the frequency time series can be used as a new early-warning signal of the proximity of the power system to a blackout.

The paper is organized as follows: Sec. II reviews the related work and contrasts it with our novel contribution. In Sec. III, we provide a description of the power system PMU data including its fractal characteristics. In Sec. IV, we introduce the early-warning signal and propose Kendall's tau of the frequency Hurst exponent as a measure of proximity to blackout. Sec. V is the conclusion.

II. Related work and novel contribution

Complex systems usually have critical thresholds before they go into sudden change from one state to another. Because of the complex structure of these systems, it is certainly not easy to build models that describe their dynamics accurately. However, such systems show a critical slowing down phenomenon as early-warning signal prior to the critical transition [9]. Critical slowing down means that the system needs more time to recover from a perturbation before critical

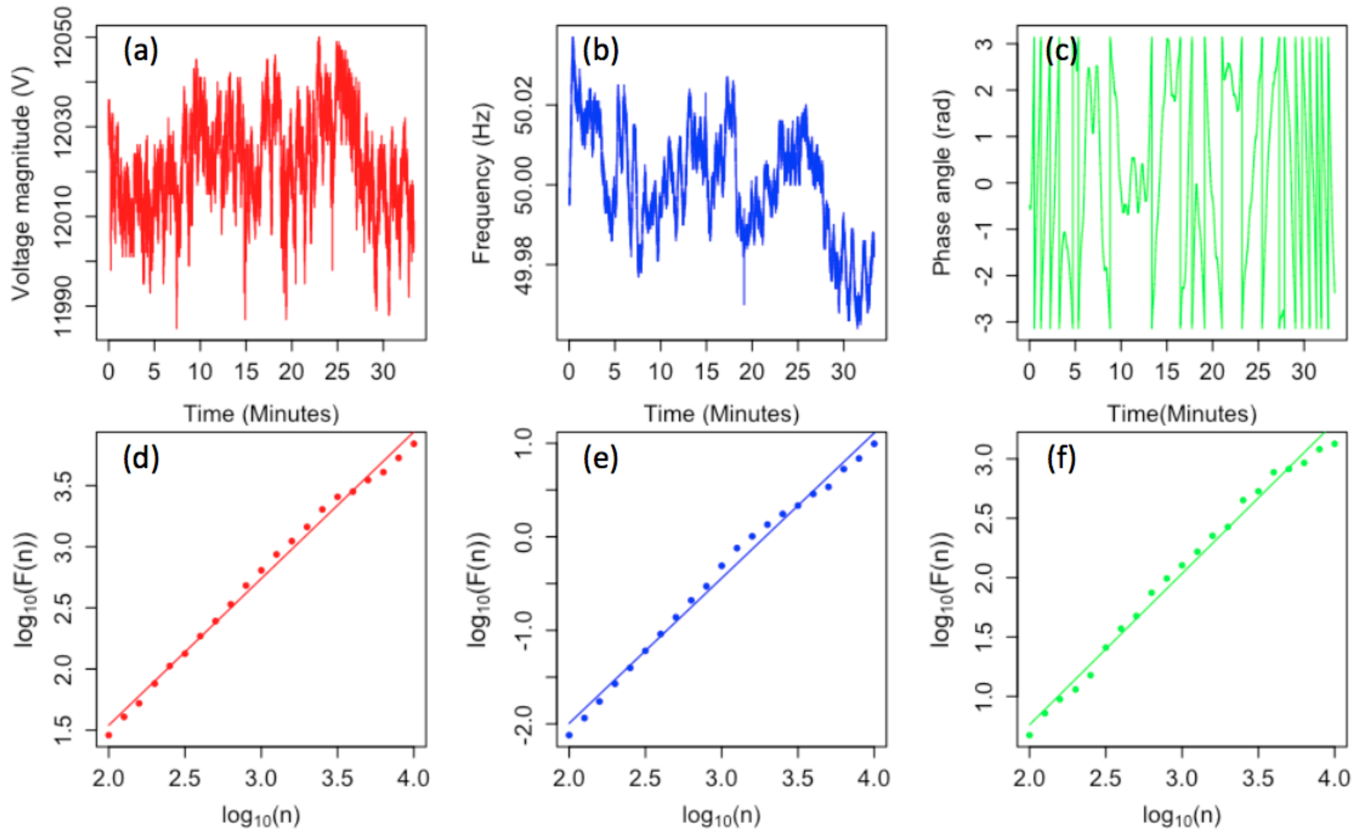

Fig. 1: Analysis of PMU#4 data recorded at EPFL on 01/15/2015 (8:00-8:33 AM) during normal operation. First row: raw data recorded; second row: corresponding log-log plot of rms fluctuation function versus the box size n. First (red) column: voltage magnitude; second (blue) column: frequency; third (green) column: phase angle

transition. The critical slowing down manifests itself as a sudden increase in the first auto-regressive (AR(1)) coefficient before the occurrence of the transition [12].

Critical slowing down has been used as an early-warning signal for critical transition in climate [9], environment [6], ecosystems [10], and power system [3]. In [3], the slowing down has been investigated by calculating the AR(1) coefficient of the frequency data collected before the Western Interconnect Blackout in 1996. Our contribution consists in using another frequency parameter as an indication of an imminent blackout. Specifically, our contributions are:

- First, we show mono-fractality of the PMU data (V, f, and θ) collected in EPFL campus during normal operation. We run the DFA over a PMU big data set collected over four months (January, April, June, and October) in 2015.
- Second, we introduce a new measure of the proximity to blackout in the power system by comparing the Kendall's tau of the Hurst exponent of the frequency data collected from the EPFL campus (normal condition) and before the 2012 Indian blackouts.

III. Power system data

In this section, we give descriptions of the PMU data (V, f, and θ) collected in the EPFL campus and review how the Hurst exponent of these PMU data is computed.

A. Description of the PMU data

The EPFL campus network has five PMUs installed throughout the campus to collect several data measurement of the power system variables. These variables are voltage magnitude (V), frequency (f), phase angle (θ), active power (P), and reactive power (Q). The PMU data measurements for several months in 2015 are available online [15]. The PMUs are installed at the medium voltage side with nominal voltage 12 KV and nominal frequency 50 Hz. The three time series of voltage magnitude, frequency, phase angle measured on 1/15/2015 (8:00-8:33 AM) are shown in Figs.1 (a-c).

In our analysis, we have chosen four months (January, April, June, and October) of 2015 to represent the different seasons in the year. Then, we picked four time series (100,000 samples each) per day of each of the power system variables (V, f, and θ). Two of the four time series were measured during day time (8:00-10:00 AM) and the other two measured during night time (8:00-10:00 PM). Based on the availability of data, we will analyze 693 time series of each of voltage magnitude (V), frequency (f), and phase angle (θ) with a total of 207.9 million data samples.

B. PMU data fractality

In [16], we have shown the existence of non-stationarity and mono-fractality of the PMU data sets collected from the Texas

Synchrophasor Network. Here, we analyze a larger number of PMU data from the EPFL campus to compare the results with PMU data from Texas. Also, the large size of the data gives some confidence in the calculated scaling exponents of the data. Using the DFA method, we found the scaling exponents of the three time series shown in Figs. 1 (a-c). The plots of the rms fluctuation function ($F(n)$) versus the window size (n) for each of the three power system variables are shown in Figs. 1 (d-f). The scaling exponent of the voltage magnitude (V) in Fig. 1 (d) is 1.20. The time series of frequency (f) is similar to Brownian noise with scaling exponent equal to 1.55, as shown in Fig.1 (e). The scaling exponent of the phase angle time series is around 1.27 as shown in Fig. 1 (f). The Hurst exponent histograms of the time series collected

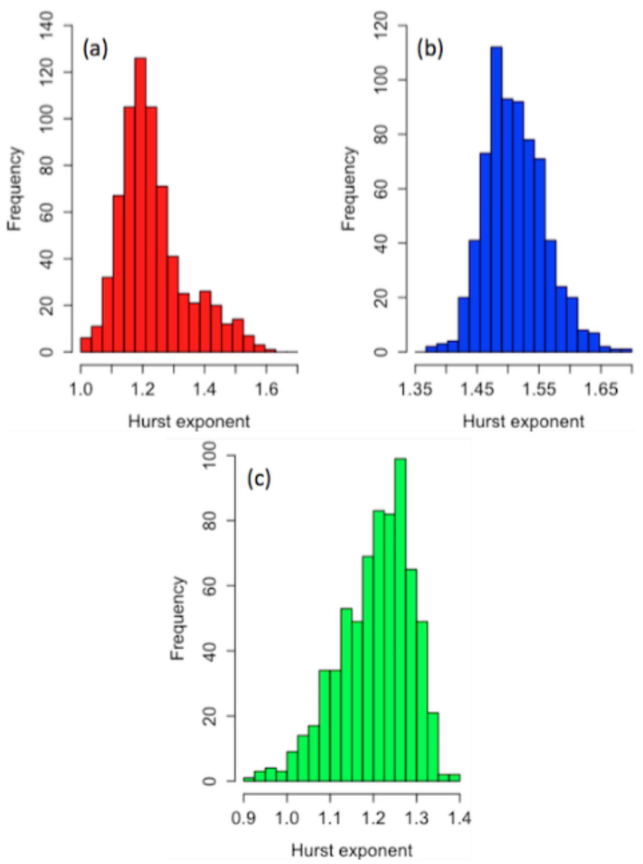

Fig. 2: Hurst exponent histograms of EPFL PMU#4 data during all 4 seasons (January, April, June, and October): (a) Voltage magnitude (b) Frequency (c) Phase angle

from EPFL are shown in Figs. 2 (a-c). The Hurst exponent histograms of the voltage magnitude, frequency, and phase angle are shown in red, blue, and green, respectively. Each of the Hurst exponent histograms represents the Hurst exponents of one of the power system variables (V, f, or θ) collected during the four months (January, April, June, or October).

The Hurst exponent histogram of voltage magnitude, shown in Fig. 2 (a), has mean 1.23 with standard deviation 0.11. The histogram of the frequency time series has a mean Hurst exponent 1.51 with standard deviation 0.05. Finally, the phase angle Hurst exponent has mean and standard deviation 1.21 and 0.08, respectively.

IV. Kendall's tau as Blackout Proximity Margin

The power blackout usually starts as instability in the voltage, frequency, or phase angle that leads to a major blackout. The power system blackout is traditionally explained using bifurcation theory [5], where the system at bifurcation point move from the stable region to the unstable one.

Among the many intertwined phenomena that ultimately lead to a blackout, a hitherto overlooked phenomenon is the "anti van der Pol" behavior characterized by the overall damping coefficient of the hidden feedback of the generator-transmission-distribution network [18] becoming more significant as the voltage decreases [17]. To be specific, consider a single generator with constant e.m.f. feeding a single load via a single transmission line. This simple generator-line-load system has a hidden feedback that has characteristic equation [17]:

$$1 + \mathbf{Z_L}\left(V_{\mathrm{L}}, w - j\sigma\right) \mathbf{Y_{Line}}\left(\sigma + j\omega\right) = 0. \tag{1}$$

In the above, V_L is the load voltage magnitude, $\mathbf{Y_{Line}}\left(\sigma + j\omega\right)$ is the classical circuit theoretical admittance of the line and $\mathbf{Z_L}\left(V_{\mathrm{L}}, \omega - j\sigma\right)$ is the non classical load impedance; specifically,

$$\mathbf{Z_L} = \frac{1}{K_p V_L^{p_v - 2}(\omega - j\sigma)^{p_w} - jK_q V_L^{q_v - 2}(\omega - j\sigma)^{q_w}},$$

where K_p, K_q, p_v, p_w, q_v, q_w are numerical coefficients copied from the Berg load model [2], which is here analytically extended to allow for a small damping.

If $\mathbf{Z_L}$ were a classical load, there would be no solution to Eq. (1). However, the nonclassical impedance of the load, which among other things depends on the voltage and encapsulates a load aggregation effect [18] manifesting itself in noninteger exponents of ω, makes solutions possible [17]. Next to the normal operation solution $\sigma(V_L = 1 \text{ p.u.}) = 0$, there could exist depending on the load characteristic voltage collapsing solutions of the form $\sigma(V_L < 1 \text{ p.u.}) < 0$. This is the "anti van der Pol" behavior.

Clearly a concept of "proximity margin" emerges here: Given a normal operation parameter set $(K_p, K_q, p_v, p_w, q_v, q_w)$, how close it is, in ℓ^1 or other norms, to a parameter set allowing for "bad" solutions to Eq. (1).

This "anti van der Pol" phenomenon has, however, revealed that a contributor to blackout is the noninteger property of the exponents $p_{v,w}, q_{v,w}$ of ω, which in turn can be reinterpreted in the time domain as fractional derivatives, making the system "complex" and subject to tipping point phenomena.

To test the possibility of existence of critical slowing down phenomena in the power system, in Sec. IV-A, we investigate the change in the AR(1) coefficient (short-range correlation) in the frequency time series before the 2012 Indian blackout. In Sec. IV-B, we introduce our new method to anticipate an imminent power system blackout based on the change of the Hurst exponent (long-range correlation) before the blackout. Then, in Sec. IV-C, we investigate the change in AR(1)

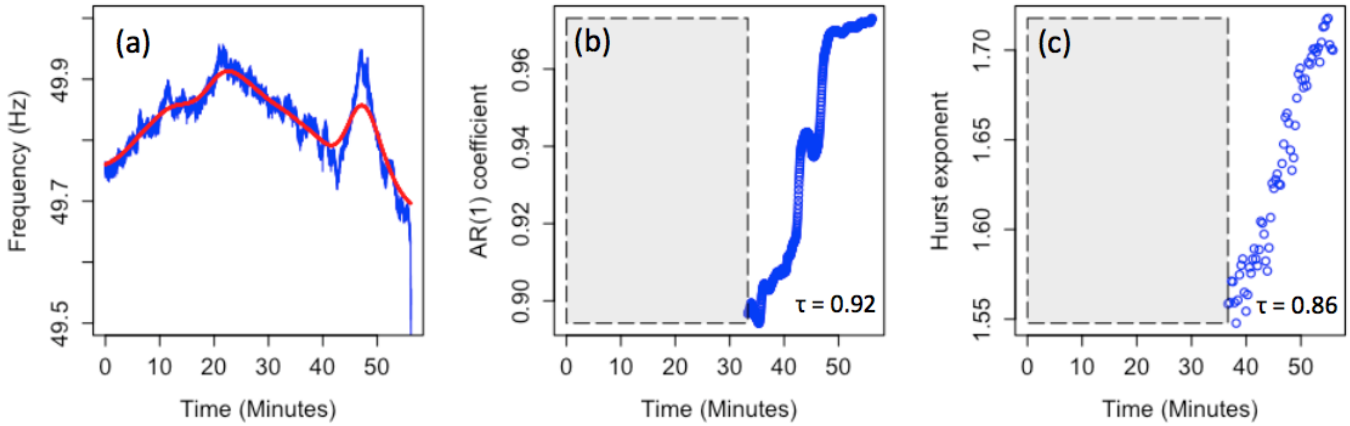

Fig. 3: (a) Raw (blue) and Gaussian kernel smoothed (red) frequency time series before 2012 Indian blackout (b) Autocorrelation coefficient at lag 1 of residual differences between raw and smoothed frequency time series (c) Hurst exponent of the raw frequency time series

coefficient and Hurst exponent over large data set of frequency time series collected from the EPFL campus during normal conditions.

A. *Kendall's tau of* AR(1)

The frequency time series, collected before the 2012 Indian blackout, is shown in blue color in Fig. 3(a). The length of the time series is 167,600 samples (∼56 minutes) and it is measured at sampling rate of 50 samples/second. The frequency time series is non-stationary [16]; therefore, we should remove the trends in the time series before calculating the AR(1) coefficient.

We remove the trends in the frequency time series using Gaussian convolution kernel [22]. We first find the smoothed function of the time series and then subtract the smoothed function from the original time series to get the residual. Choosing the bandwidth ($\sim 2.7\sigma$) of the Gaussian kernel correctly is a critical step to avoid underfitting (large bandwidth) and overfitting (small bandwidth). The bandwidth that has best smoothing for the frequency time series is 18,000 samples (6 minutes).

Using the command *ksmooth* in R software [21], we calculated the Gaussian kernel smoothed function of the frequency time series before the blackout as shown in red color (over blue frequency color) in Fig. 3(a). We have chosen the normal kernel with 18,000 samples bandwidth.

Now, we find the change in the AR(1) in the residual time series by calculating the AR(1) coefficient over a window (grey shaded area) of 100,000 samples (∼33 minutes). Then, we move the window 100 samples (2 seconds) to the right (toward the blackout) and calculate the AR(1) coefficient over the next window. Fig. 3(b) shows the change in the AR(1) coefficient of the residuals time series before the blackout. The AR(1) coefficient was calculated using *ar.ols* command in R software.

The AR(1) coefficient before the blackout shows an increase from 0.89 to 0.97. The increase starts around 33 minutes before the blackout, which could be an early-warning for power system blackout. The increase in the AR(1) coefficient can be quantified using Kendall's tau [13].

Kendall's tau is a rank correlation coefficient that is used to measure the ordinal association between two quantities. Assuming that we have n pairs of x and y $((x_1, y_1), (x_2, y_2), ..., (x_n, y_n))$. Kendall's tau is defined as:

$$\tau = \frac{\#\ of\ concordant\ pairs - \#\ of\ discordant\ pairs}{n(n-1)/2} \tag{2}$$

The pair is concordant if $x_i > x_j$ & $y_i > y_j$ or $x_i < x_j$ & $y_i < y_j$. On the other hand, the pair is discordant if $x_i > x_j$ & $y_i < y_j$ or $x_i < x_j$ & $y_i > y_j$. The range of Kendall's tau is between -1 and 1.

The Kendall tau of the AR(1) coefficient could go as high as 0.92 close to a blackout. With 685 data points, the Null Hypothesis of no trend in the AR(1) is rejected with $p = 2.2 \times 10^{-16}$. The very high confidence level means that the AR(1) increase is symptomatic of a dynamical shift in the grid. However, the observed increase in the autocorrelation could not be exclusively happening before the power system blackout. Indeed, in Sec. IV-C, we will show that such high confidence changes in the AR(1) coefficient are indeed happening over large data set of frequency time series collected from the EPFL campus network under normal conditions.

B. *Kendall's tau of the Hurst exponent as proximity margin to blackout*

In Sec. III-B, we have shown the existence of long-range memory in the frequency time series in the power system by analyzing a large number of data sets collected from the EPFL campus network. Our novel method to predict the power system blackout will be based on the change in long-range correlation (Hurst exponent) of frequency time series instead of the short range correlation at lag 1.

In Fig. 3(a), we have the frequency time series before the 2012 Indian blackouts for 56 minutes. We study the change in Hurst exponent of a moving window of length 110,000

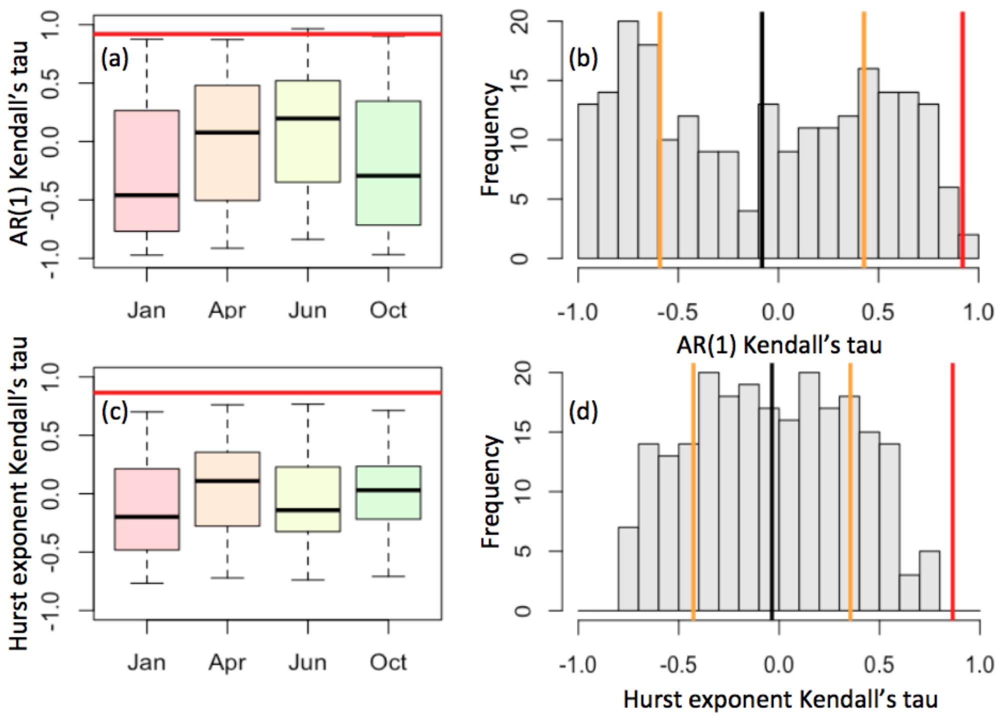

Fig. 4: (a) Box-plots of Kendall's tau of frequency AR(1) coefficient during January, April, June, and October (b) Histogram of Kendall's tau of frequency AR(1) coefficient during these four months (c) Box-plots of the Kendall's tau of frequency Hurst exponent during these four months (d) Histogram of Kendall's tau of frequency Hurst exponent during these four months. Due to blackout data scarcity, a red histogram and a test for random draws from different distributions couldn't be obtained.

samples ($\sim$37 minutes) and a shift of 900 samples (18 seconds). We calculated the Hurst exponent of each window using the Detrended Fluctuation Analysis (DFA) method, which is robust against the presence of trends and non-stationarity in the time series. Using *dfa* command in R software, the Hurst exponents of the frequency before the blackout were calculated as shown in Fig. 3(c). To calculate the Hurst exponent, we have used 21 window sizes with range from 100 to 10,000 samples.

The Hurst exponent before the blackout increases from 1.55 to 1.72. Kendall's tau of the Hurst exponent is 0.86. With 65 data points, the Null Hypothesis of no trend is rejected with $p = 2.2 \times 10^{-16}$, a very high confidence level. This increase in the Hurst exponent before the blackout can be a sign of the proximity of the power system to blackout.

Next, we need to show that the increase in the AR(1) coefficient and Hurst exponent before the 2012 Indian blackout is unique and can be used as reliable measures for the proximity of the power system to blackout. This goal will be achieved by calculating the change in the AR(1) coefficient and Hurst exponent of several frequency data sets collected from EPFL campus under normal conditions.

C. Kendall's tau of AR(1) coefficient versus Kendall's tau of the Hurst exponent under normal conditions

In this section, we use the Kendall's tau to quantify the change in the AR(1) coefficient and Hurst exponent of 230 frequency time series (180,000 samples each). These time series were collected during normal conditions from the EPFL campus. Kendall's tau close to 1 means that AR(1) coefficient or Hurst exponent has consistently increasing trend. However, Kendall's tau close to -1 indicates consistently decreasing trend. Then, we will compare the Kendall's tau of the frequency time series before blackout with the ones that are collected during normal conditions.

The AR(1) coefficient of 230 frequency data set is calculated over 100,000 samples window with 100 samples shift between consecutive windows. The Kendall's tau distributions of frequency AR(1) coefficient during each of the four months are shown in Fig. 4(a). The histogram of Kendall's tau of the frequency AR(1) coefficient for all the month is shown in Fig. 4(b). The histogram has -0.08 mean (black line) and 0.57 standard deviation (orange line) with range between -0.97 and 0.96. The histogram is very close to a uniform distribution over the range of the Kendall's tau from -1 to 1. Since the Kendall's tau of the frequency AR(1) coefficient before the Indian blackout is 0.92 (red line) and inside the range of the histogram of the EPFL data, there is no significant difference between blackout and normal data.

The Hurst exponent of each frequency data set is calculated over a 100,000 samples window with 1,000 samples shift between consecutive windows. Then, we calculate the Kendall's tau of the Hurst exponents for each of 230 frequency data sets. The distributions of the Kendall's tau of the frequency Hurst exponent for each of the four months are shown in Fig. 4(c). The histogram of the Kendall's tau for all the frequency data sets is shown in Fig. 4(d). The mean of the histogram is -

0.04 (black line) and the standard deviation is 0.39 (orange color). It is clear that the histogram of the Kendall's tau of frequency Hurst exponent is centered around 0 with range between -0.77 and 0.77. The Kendall's tau of frequency Hurst exponent before the Indian blackout is 0.86 (red line) and it is outside the histogram range of the normal EPFL data. That means there is a difference between the blackout and normal data and hence the Kendall's tau of frequency Hurst exponent is a good measure of the proximity to blackout.

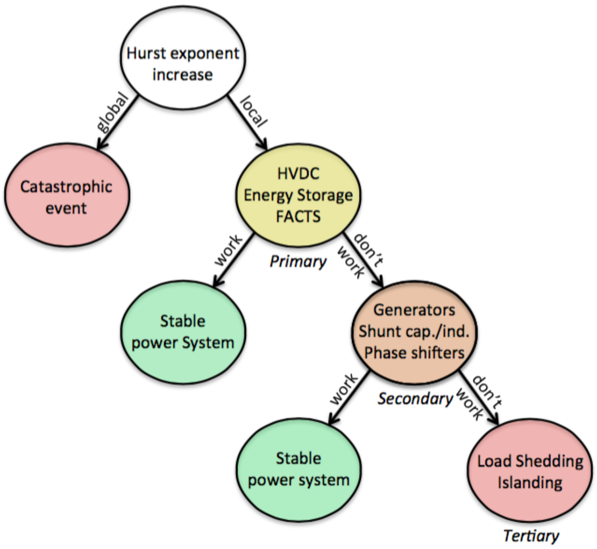

Fig. 5: Decision tree acting on the multi-layer decomposition of Fig. 5 of [11] as response to increase of Hurst exponent

V. Conclusion

A. Summary

In this paper, we first investigated the long-range memory in a large PMU data sets from the EPFL campus network. All the PMU data (V, f, and θ) showed a long-range correlation with average Hurst exponents approximately 1.23, 1.51, and 1.21, respectively. Next, we studied the existence of critical slowing down phenomenon in the frequency time series before the power system blackout. We provided an evidence that the increase in the autocorrelation coefficient at lag 1 before the blackout is not specific to blackout; however, the increase in the Hurst exponent of the frequency, more specifically having a high Kendall's tau of the frequency Hurst exponent, can be a better early-warning signal for power system blackout.

B. Future work: practical implementation in 3-level hierarchy

Should the Hurst exponent increase, the pressing question is whether it is global (catastrophic) or area-local (manageable) and what has caused the increase. Assuming that the area of Hurst exponent increase and the root cause are identified, the action would be to act as quickly as possible using FACTS, at the primary (milli second) layer Fig. 5 of [11], consistently with the fast sampling rate of PMUs. Should the problem not be rectified at this layer, we would attempt to rectify it by reactive power management at the secondary (seconds) layer. Finally, if the problem is not yet resolved at the secondary layer, the last attempt would be to resolve it at the tertiary (minutes) layer, by for example islanding, or load shedding, which unfortunately would affect some consumers. The overall process is depicted in the tree of Fig. 5.

Acknowledgment

P. Bogdan acknowledges the support by US National Science Foundation (NSF) under Grant CyberSEES 1331610 and CNS/CPS CAREER 1453860.

References

[1] IEEE Standard for Synchrophasor Data Transfer for Power Systems. *IEEE Std C37.118.2-2011 (Revision of IEEE Std C37.118-2005)*, pages 1–53, 2011.
[2] G. J. Berg. System and load behavior following loss of generation—experimental results and evaluation. *Proc. IEE*, 119(10):1483–1486, October 1972.
[3] E. Cotilla-Sanchez, P. Hines, and C. M. Danforth. Predicting critical transitions from time series synchrophasor data. *IEEE Transactions on Smart Grid*, 3(4), December 2012.
[4] I. Dobson. The irrelevance of electric power system dynamics for the loading margin to voltage collapse and its sensitivities. *NOLTA, IEICE*, 2(3):263–280.
[5] I. Dobson. Observations on the geometry of saddle node bifurcation and voltage collapse in electrical power systems. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 39(3):240 – 243, March 1992.
[6] J. M. Drake and B. D. Griffen. Early warning signals of extinction in deteriorating environments. *Nature*, 467, September 2010.
[7] D. I. Kim et al. Wavelet-based event detection method using pmu data. *IEEE Transactions on Smart Grid*, October 2015.
[8] JA Jiang et al. An adaptive PMU based fault detection/location technique for transmission lines part i: Theory and algorithms. *IEEE Transactions on Power Delivery*, 15(2), April 2000.
[9] V. Dakos et al. Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences*, 105:14308–14312, 2008.
[10] V. Dakos et al. Slowing down in spatially patterned ecosystems at the brink of collapse. *The American Naturalist*, 177(6), June 2011.
[11] B. Fardanesh. Future trends in power systems control. *IEEE Computer Applications in Power (CAP)*, 15(3):24–31, July 2002.
[12] A. R. Ives. Measuring resilience in stochastic systems. *Ecological Monographs,*, 65(2):217–233, May 1995.
[13] M. G. Kendall. A new measure of rank correlation. *Biometrika*, 30(81-93), 1938.
[14] C.-K. Peng et al. Mosaic organization of DNA nucleotides. *Phys Rev E*, 49(2), 1994.
[15] The École polytechnique fédérale de Lausanne. PMU data from EPFL campus. http://nanotera-stg2.epfl.ch/, Online; accessed March 8, 2016.
[16] 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. IEEE Power & Energy Society.
[17] L. Shalalfeh and E. Jonckheere. The existence of a voltage collapse solution in the static-dynamic gap. In *American Control Conference (ACC)*, pages 4126–4131, Boston, MA, July 2016. American Automatic Control Council.
[18] L. Shalalfeh and E. Jonckheere. Load aggregation effect in the power grid. December 2016. To appear, available at `http://eudoxus2.usc.edu`.
[19] A. Sode-Yome, N. Mithulananthan, and K. Y. Lee. A maximum loading margin method for static voltage stabilityu in power systems. *IEEE Transactions on Power Systems*, 21(2):799–808, May 2006.
[20] J. E. Tate and T. J. Overbye. Line outage detection using phasor angle measurements. *IEEE TRANSACTIONS ON POWER SYSTEMS*, 23(4), November 2008.
[21] R Development Core Team. R: A language and environment for statistical computing. https://www.r-project.org, Online; accessed May 22, 2016.
[22] G. S. Watson. Smooth regression analysis. *Sankhya A*, 26(4):359–372, 1964.