

Entropy Based Analysis of Delayed Voltage Recovery

Sambarta Dasgupta, *Student Member, IEEE*, Magesh Paramasivam, *Student Member, IEEE*,
Umesh Vaidya, *Member, IEEE*, and Venkataramana Ajjrapu, *Fellow, IEEE*

Abstract—In this work, Fault Induced Delayed Voltage Recovery (FIDVR) phenomenon has been characterized as the loss of Entropy in the voltage time series. The relationship between entropy and the voltage recovery rate has been established. It is shown that, the entropy can be an estimate of the recovery rate. Furthermore, the entropy computation has been done entirely from the statistical properties of the time series, without using any model information. Kullback-Leibler(KL) divergence, an entropy based quantity, has been demonstrated to capture both the rate and the level of recovery. A critical value of the KL divergence is computed, which can indicate possible violations of the WECC voltage performance criterion. The KL divergence is then applied to identify weak buses, and severe contingencies in power network.

Index Terms—Delayed Voltage Recovery, Entropy, Contingency analysis.

I. INTRODUCTION

U.S. Department of Energy Workshops on FIDVR recognized the growing concerns of utilities over FIDVR events and considers it as a national issue because of the increasing penetration of residential air conditioner (AC) loads [1], [2]. FIDVR related studies have got increased attention from industry and academic researchers in the recent past [3], [4]. Increasing efforts have been made to properly represent the behavior of induction motor loads in planning studies [5]. The major cause for FIDVR events is related to the dynamic behavior of induction motor loads, which tend to decelerate and stall following a large disturbance resulting in low voltages in a significant portion of the power system. The reactive requirement of the induction motor increases when the induction motor stalls and may prevent quick voltage recovery. The enhanced voltage instability predictor (VIP) method is used in [4] to identify FIDVR and short term voltage stability problems. The VIP method represents the power system using a Thevenin equivalent and the parameters are recursively estimated at the rate which the PMU data are measured. In [6], [7], slope based voltage recovery calculation is used to predict FIDVR events. Successive voltage measurements are used to calculate the slope of the voltage recovery trajectory. Using this slope the expected voltage recovery for the future time is predicted. The predicted voltage recovery is compared with the specified transient voltage recovery criteria for identifying FIDVR events. For the identification of FIDVR using steady

state screening of buses, MVA-Volt index is proposed in [8]. The MVA-volt index is summation over all system buses of load MVA times voltage depression from the nominal value during a three phase fault at any bus. The MVA-volt index can be perceived as an integral error type quantity.

The previously devised methods can be broadly classified into two categories - slope based techniques [6], [7], and integral error based methods [8]. Integral error based measure could not differentiate between two different voltage trajectories - one showing moderate recovery, the other showing fast recovery over a small period of time. On the other hand, any measure based on the slope or the derivative of the voltage trajectory will suffer in case of oscillations or sudden almost discontinuous changes in the voltage magnitude.

In this work, the entropy, a popular measure of complexity and uncertainty in the Statistical Mechanics [9], and Information Theory [10], is used to characterize the rate of recovery of the voltage signal. From the perspective of communication, the higher entropy of a signal implies, that more bits of data is needed to encode the signal [11]. Computation of the finite time entropy using model of the dynamical system has been pursued in [12]. In this work, a model-free method has been proposed to compute entropy from the time series data to characterize the nature of recovery. If the bus voltage recovers fast, it converges to the nominal value within a shorter time period. Henceforth, if the time series is viewed as an unknown variable, the corresponding probability distribution will have less of uncertainty, as the probability of finding the voltage values around the nominal value would be higher. The implication of this observation is that the entropy is an indicator of the rate of recovery.

The exact voltage level, where the voltage value finally converge, is also equally important along with the rate of recovery. To encompass both rate of recovery and level of it, Kullback-Leibler (KL) divergence, an entropy based measure of distance between distributions, is utilized [13]. Successively, a critical bound on the value of the divergence is computed, which meets the WECC voltage performance criteria. KL divergence is also used to identify critical buses and contingencies.

The contributions of the paper is as follows, - 1. demonstrating FIDVR as a phenomenon of loss of entropy of voltage signal, 2. Application of entropy based KL divergence to characterize both rate and level of recovery, 3. Using KL divergence to identify critical buses and contingencies. The rest of the paper is structured as follows. Section III describes how entropy and rate of recovery of voltage time series are related. Kullback-Leibler divergence based joint characterization of

Authors are with the Department of Electrical and Computer Engineering Iowa State University, Ames, Iowa 50011.

Email: dasgupta@iastate.edu, mageshp@iastate.edu, ugvaidya@iastate.edu, and vajjarap@iastate.edu.

rate and level of recovery is discussed in Section III-D. The applications of the proposed method are outlined in Section V. Finally, conclusions are drawn in Section VI.

II. PROBLEM DESCRIPTION

WECC transient voltage dip criterion states that for a Category B disturbance (single element outage), should not cause a transient voltage dip that is greater than 20% for more than 20 cycles at load buses, or exceed 25% at load buses or 30% at non-load buses at any time other than during the fault [14]. Also, the steady state voltage limit prescribes at steady state voltage values should be between the bound of 95% to 105%. Figure 1 summarizes the WECC voltage performance criteria. There are two important factors in the

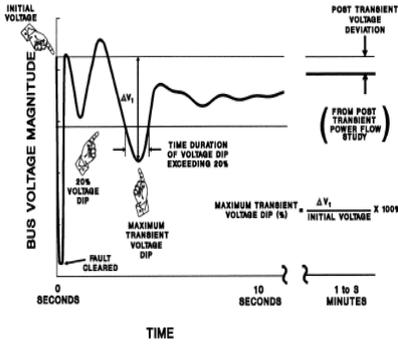


Fig. 1. WECC Voltage Performance Criteria [14].

recovery process - rate of recovery (how fast the waveform converges), and level of recovery (to which voltage level it converges). The WECC voltage performance criteria is a point-wise condition on the voltage waveform. It can only be inferred, whether or not the waveform satisfies the WECC criterion. It can not be used to identify the rate or level of recovery, and also can not be used to compare quantitatively recovery nature of several waveforms. The slope or integral based methods have limitations in quantifying the voltage recovery as an index for comparison. In the proposed method, first we show the rate of recovery is captured by entropy. Then we adopt a entropy based quantity, KL divergence, to simultaneously capture rate and level of recovery.

III. DENSITY-BASED APPROACH FOR THE CHARACTERIZATION OF FIDVR

We propose a density-based approach for the prediction of FIDVR. The basic idea behind the proposed approach is to construct a (probability) density function from the voltage time series data. Entropy of the density function is computed to determine the rate of recovery. However, we show that the entropy alone is not enough to determine if the recovery criteria as specified by WECC is violated or not. To address this problem, we propose to construct appropriate metric on the space of densities. In particular, we use *Kullback-Leibler* (KL) divergence to measure the distance between the given density and the density corresponding to the ideal voltage recovery. Based on the value of KL divergence we determine if there is a violation or not. We start this section

by first outlining the procedure for the construction of density function from a given time series data. Discussion on the computation of entropy to determine the rate of recovery and the KL divergence to determine the WECC violation follow in subsequent subsections.

A. Computation of density function from time series data

It is important to emphasize that the criteria as specified by WECC voltage performance criterion involves both temporal and magnitude information of the voltage time series data. WECC voltage performance criteria involves information where it is required for the voltage magnitude to have reached certain value within particular time interval. These two (temporal and magnitude) pieces of information can be combined with the help of density function. Mathematically speaking, the density function, $p : X \rightarrow \mathbb{R}$, is any nonnegative scalar value function with finite integral. Furthermore, the density function is said to be a probability density if the integral of the function over X is equal to one i.e.,

$$p(x) \geq 0, \text{ and } \int_X p(x)dx = 1. \quad (1)$$

We now provide a procedure for the construction of approximation density function p from the voltage time series data. Consider a voltage time series data where a fault occurs at $t = T_0$ and is cleared at $t = T_{cl}$, the voltage waveform starts to rise. The bus voltage $v_{\min} < v(t) < v_{\max}$ is observed from the time instant $t = T_{cl}$ to $t = T_f$. v_{\min} is the voltage level at the time instant when the fault is cleared and v_{\max} is the nominal value before fault (e.g. 1 pu). The interval (v_{\min}, v_{\max}) is divided into N intervals such that

$$[v_{\min}, v_{\max}] = \bigcup_{i=1}^N [v_i, v_{i+1}] = \bigcup_i D_i. \quad (2)$$

The time spent by the trajectory in the interval $[v_i, v_{i+1}]$ is denoted as Δt_i and defined as,

$$\Delta t_i := \int_{T_{cl}}^{T_f} \chi_{[v_i, v_{i+1}]}(v(t)) dt. \quad (3)$$

where $\chi_A(x)$ is the characteristic function of set A ,

$$\chi_A(x) = \begin{cases} 1 & \text{for } x \in A \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

We define

$$p_i := \frac{\Delta t_i}{T} = \frac{1}{T} \int_{T_{cl}}^{T_f} \chi_{[v_i, v_{i+1}]}(v(t)) dt. \quad (5)$$

It can be verified that sum of p_i over index i adds to one i.e.,

$$\sum_{i=1}^N p_i = \frac{1}{T} \int_{T_{cl}}^{T_f} \sum_{i=1}^N \chi_{[v_i, v_{i+1}]}(v(t)) dt = \frac{1}{T} \int_{T_{cl}}^{T_f} 1 dt = 1.$$

The p_i for $i = 1, \dots, N$ can now be used to approximate the probability density function $p(x)$ from Eq. (1). In particular,

$$\tilde{p} := (p_1, \dots, p_N)$$

is an approximation to $p(x)$, where the approximation essentially involves discretization of the space X into intervals D_i (2). The entries $p_i \approx \int_{D_i} p(x)dx$. Furthermore, in the limit as $N \rightarrow \infty$, the \tilde{p} will converge to p in weak sense.

B. Entropy and rate of recovery

The probability density function, p , and its approximation, \tilde{p} , can be used to compute entropy. Entropy is a measure of uncertainty and will be used to characterize the rate of recovery of voltage waveform. We start with the definition of entropy.

Definition 1 (Entropy): The entropy corresponding to a probability distribution function $p(x)$ is defined as follows:

$$H(p) := - \int_X p(x) \ln p(x) dx \quad (6)$$

The entropy corresponding to the approximation, \tilde{p} , is defined as

$$H(\tilde{p}) := - \sum_{i=1}^N p_i \ln p_i \quad (7)$$

Remark 1: It is to be noted that $p_i \ln p_i$ is taken as 0 for $p_i = 0$, using limit argument.

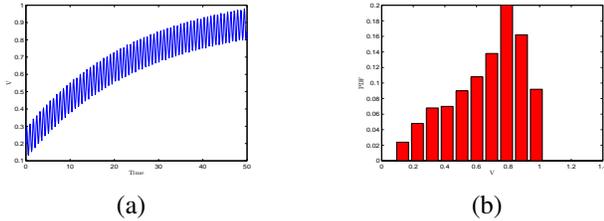


Fig. 2. (a) Voltage trajectory where $v(t) = 1 - 0.8e^{-\alpha t} + 0.09 \sin 8.3t$, $\alpha = 0.04$. (b) Corresponding probability distribution (Entropy = 2.1)

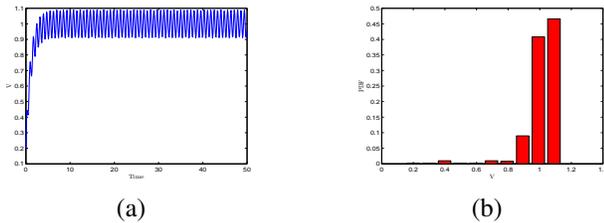


Fig. 3. (a) Voltage trajectory where $v(t) = 1 - 0.8e^{-\alpha t} + 0.09 \sin 8.3t$, $\alpha = 0.9$. (b) Corresponding probability distribution (Entropy = 1.1)

Entropy is a measure of uncertainty where larger value of entropy implies more uncertainty and vice versa. For example, the entropy corresponding to uniform probability distribution will be maximum and one corresponding to Dirac delta distribution (where the probability mass is concentrated at single point and hence certain) will be minimum. For a given voltage waveform if the voltage magnitude recovers fast then it converge to the nominal value quickly. The probability density function computed using the procedure outlined in Section III-A will be concentrated closer to the nominal value and hence will correspond to a smaller value of entropy. On the other hand, if the voltage magnitude recovers slowly to its nominal value then the density function corresponding to such a recovery will be more dispersed and hence will lead to larger value of entropy. So the entropy can be used as a measure of voltage recovery where larger value of entropy corresponds to slower recovery and vice versa. We now demonstrate using

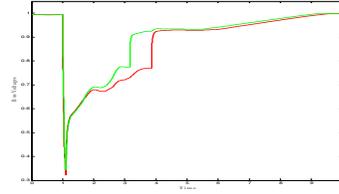


Fig. 4. Evolution of voltage magnitudes for buses 163 (red) and 164 (green).

simple example the connection between the rate of recovery and the entropy value.

Example 1: Consider voltage waveform of the form $v(t) = 1 - 0.8e^{-\alpha t} + 0.09 \sin 8.3t$. The rate of recovery is captured by parameter $\alpha > 0$. Larger value of α corresponds to faster recovery and vice versa. Fig. 2 and 3 demonstrate two different scenarios when $\alpha = 0.04$ and $\alpha = 0.9$. The time span for observing is 100 and in both cases voltage recovers from 0.2 pu to 1 pu. Comparing Figs. 2 and 3, we observe that the probability distribution corresponding to slower recovery (i.e., $\alpha = 0.04$) is more dispersed compared to the one with faster recovery (i.e., $\alpha = 0.9$). Hence the entropy value for the slower recovery is relatively larger ($H = 2.1$) compared to faster recovery ($H = 1.1$).

The following claim demonstrates that for the special case when the voltage recovery is exponential then the entropy decreases with the increase in recovery rate.

Claim 1: Let the bus voltage evolves exponentially with recovery rate α as described by the following equation,

$$v(t) = v_{max} - (v_{max} - v(0))e^{-\alpha t}.$$

Then, for sufficiently large T entropy is a decreasing function of the recovery rate, i.e.

$$\frac{\partial H(\alpha, T)}{\partial \alpha} < 0$$

where,

$$H(\alpha, T) := - \sum_{i=1}^N p_i \ln p_i, \quad \& \quad p_i = \frac{1}{T} \int_0^T \chi_{[v_i, v_{i+1})}(v(t)) dt.$$

We provide the proof in the Appendix.

C. Simulation Results for Entropy

In Figs. 4 and 5, we show the simulation results for IEEE 162 bus system. A three phase fault is created at bus 120 for six cycles and the fault is cleared by opening the line 5 – 120. For this fault, the evolution of bus voltage magnitudes for buses 163 and 164 is shown in Fig. 4. It can be observed that the bus voltage corresponding to 164 recovers faster compared to voltage at bus 163. Fig. 5 shows the probability distribution for the corresponding bus voltages. The entropy values are 1.7 and 1.5 for the buses 163 and 164 respectively. The entropy values are in agreement with the observed rate of recovery of the corresponding buses.

Although entropy serves as a good measure for the rate of recovery, it is insensitive to the final steady state value of bus voltage. In particular, consider a scenario where the

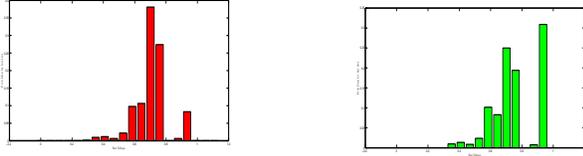


Fig. 5. Bus voltage probability distribution for buses 163 (red) and 164 (green)

two voltage waveform recovers with the same rate but they converge to two different steady state values. It is not difficult to show that the entropy values computing from the probability distribution function corresponding to these waveform will be the same. However, given the fact these waveforms converge to two different steady state values it is desirable that our proposed approach will be able to differentiate such cases. To elaborate this point further, we consider the following example case.

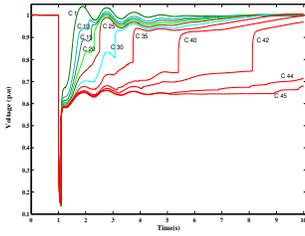


Fig. 6. Voltage recoveries corresponding to different % of induction motor loads. Labels starting with 'c' corresponds to % of induction motors

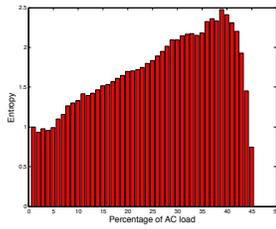


Fig. 7. Entropy values corresponding to voltage recoveries as % of AC load is increased from 0 - 45%

Example 2: It is known that the increase in percentage of AC load has a negative effect on the rate of recovery. Simulations are performed for varying percentage of AC load from 0 to 45% and the voltage time series data is observed. It can be observed from Fig. 6 that the recovery becomes slower with increase in AC load percentage. Figure 7 shows the entropy values plotted for different AC load percentage. It can be observed that the entropy does not show an increasing pattern over the entire range. In fact, for AC load larger than 40%, there is a decrease in entropy value. This is due to the fact, that for both the 0% and 45% AC load cases, recovery rates are same but the voltage recovers to two different levels. A complete characterization of the FIDVR phenomenon is only possible if both the rate and level of recovery are captured. In order to address this problem, we propose following modification of the entropy measure to capture not only the rate of recovery but also the final steady state value.

D. Kullback-Leibler divergence for joint characterization of rate and level of recovery

In order to circumvent the problem highlighted at the end of previous section, we propose the use of Kullback-Leibler (KL) divergence. KL divergence also known as relative entropy, is a popular measure of distance used in statistics and

information theory [13]. It is used to capture the difference between information contained in two different probability density functions and is defined as follows:

Definition 2: The Kullback-Liebler divergence or relative entropy between two probability density function $p(x)$ and $q(x)$ is denoted by $D(p \parallel q)$ as is given by following formula

$$D(p \parallel q) = \int_X p(x) \ln \frac{p(x)}{q(x)} dx \quad (8)$$

It can be shown that the KL divergence is always nonnegative and is zero if and only if $p = q$. However, it is not a true distance between two density function in true sense because it is not symmetric and does not satisfy triangular inequality. However it is convenient to think of KL divergence as a distance between two density functions. The KL divergence will be used for the purpose of characterizing rate as well as level of recovery of voltage signal. Towards this goal we first define a probability density function, $p^{ideal}(x)$, corresponding to ideal voltage recovery. The ideal density function will correspond to an voltage waveform which will recover instantaneously to a rated voltage value, say v_{max} , following a fault. Hence, p^{ideal} , will be a Dirac-delta function with all its mass concentrated at $x = v_{nom} = 1$ p.u. However, strictly speaking the Dirac-delta function does not qualify the definition of probability density function and hence we use following approximation for the ideal probability density function

$$p^{ideal}(x) = \hat{Z}^{-1} e^{-\Lambda(x-v_{nom})^2}, \quad \hat{Z} = \int_X e^{-\Lambda(x-v_{nom})^2} dx.$$

where \hat{Z} is the normalization constant and is introduced to ensure that $\int_X p^{ideal}(x) dx = 1$. The positive parameter $\Lambda > 0$ controls the concentration of the density near $x = v_{nom}$. For large value of λ more mass is concentrated near $x = v_{nom}$. Now let $p(x)$ be the probability density function corresponding to a particular voltage waveform with a given recovery. The objective is to compare the "distance" between $p(x)$ and $p^{ideal}(x)$ using KL divergence i.e., $D(p \parallel p^{ideal})$ to determine the recovery. We will show that the KL divergence will capture not only the rate but also the level of recovery. We now discuss the finite dimensional approximation of the KL divergence formula (8). Let \tilde{p} and \tilde{p}^{ideal} be the finite dimensional approximation of the density $p(x)$ and $p^{ideal}(x)$ respectively and of the form

$$\tilde{p} = (p_1, \dots, p_N), \quad \tilde{p}^{ideal} = (p_1^{ideal}, \dots, p_N^{ideal}).$$

where, \tilde{p} is constructed using the procedure outlined in Section III-A and

$$p_i^{ideal} := \frac{e^{-\lambda(N^*-i)^2}}{Z}, \quad Z := \sum_{i=1}^N e^{-\lambda(N^*-i)^2},$$

where, Z is the normalizing factor to maintain property of probability distribution p^{ideal} , and N^* is the index corresponding to the nominal voltage. The finite dimensional approximation of KL divergence is defined as follows.

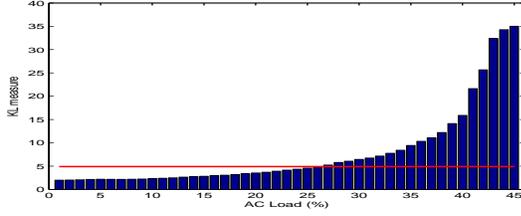


Fig. 8. KL divergence measure for different percentage of induction motor loads. Red line corresponds to the critical KL value

Definition 3: The KL divergence between \tilde{p} and \tilde{p}^{ideal} is denoted by \mathcal{K} and is defined as follows

$$\mathcal{K} := \sum_{i=1}^N p_i \ln \frac{p_i}{p_i^{ideal}}.$$

The KL divergence could be further simplified to,

$$\mathcal{K} = \ln Z + \sum_{i=1}^N p_i \ln p_i + \lambda \sum_{i=1}^N p_i (N^* - i)^2 \quad (9)$$

Now, inspecting (9), it can be observed that if the recovery is poor the KL divergence would be higher. If the voltage signal does not recover fast, p_i will be higher for smaller i . The weighting factor $(N^* - i)^2$ would be more for smaller i in comparison to larger ones. This would ensure the KL divergence would be more if the p_i is more for lower voltage levels. The KL divergence values are also plotted for various different AC load percentage in Fig. 8. It can be observed from the plot that the KL divergence shows an increasing trend with increase in percentage of AC load. This verifies the claim that KL divergence could be used to characterize both the rate and the level of recovery. In the next section, we compute the critical value of the KL divergence, which is to be satisfied in order to meet WECC criterion. The red horizontal line in the Fig. 8, shows the critical KL divergence and it can be observed that the recovery starts to violate the WECC criterion after the AC load percentage crosses 25.

IV. WECC CRITERION AND CRITICAL VALUE OF KL DIVERGENCE

The WECC criteria, as described in Section II, has been used to compute the critical value of the KL divergence. The computation involves identifying an envelope of the voltage signal, that satisfies the WECC criterion. The critical value is computed as the KL divergence of voltage curve, which forms the boundary of the envelope. The fault is cleared at T_{cl} and the voltage is observed till T_f . A pictorial representation of the voltage performance constraints is shown as dotted lines in fig.9. It is to be noted we have taken the envelope as an increasing function of time. As the KL divergence computation involves the probability density computation, this modification does not affect the the computation. The voltage value satisfies the following condition for two time instant T_1 and T_2 such that, $T_{cl} < T_1 < T_2 < T_f$,

$$\mathcal{E} := \begin{cases} v(t) \geq V_1, & T_{cl} \leq t < T_1, \\ v(t) \geq V_2, & T_1 \leq t < T_2, V_2 > V_1, \\ v(t) \geq V_3, & T_2 \leq t \leq T_f, V_3 > V_2 \end{cases} \quad (10)$$

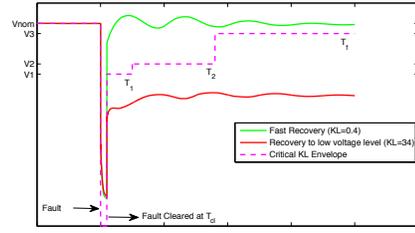


Fig. 9. Critical voltage performance envelope based on WECC criteria

where, $\Delta T_1 := T_1 - T_{cl}$, $\Delta T_2 := T_2 - T_{cl}$, and $\Delta T_f := T_f - T_{cl}$. Any voltage curve, which satisfies the above condition, falls within the envelope described by voltage T_1, T_2 , and V_1, V_2, V_3 . These quantities are treated as parameters, which could be chosen to prescribe the envelope. WECC criterion prescribes a specific values for the voltage levels and the time intervals. But the bound is computed for more general scenario, where appropriate values could be chosen for the parameters. The following proposition summarizes the critical value of the divergence,

Proposition 1: If a voltage trajectory $v(t)$ satisfies the condition, given by the envelope \mathcal{E} , then $\mathcal{K} \leq \mathcal{K}^*$ where,

$$\begin{aligned} \mathcal{K}^* := & \frac{1}{\Delta T_f} (\Delta T_1 \log \Delta T_1 + (\Delta T_2 - \Delta T_1) \log (\Delta T_2 - \Delta T_1)) \\ & + \frac{1}{\Delta T_f} (\Delta T_f - \Delta T_2) \log (\Delta T_f - \Delta T_2) + \log Z - \log \Delta T_f \\ & + \frac{\lambda}{\Delta T_f} (\Delta T_1 (1 - V_1)^2 + (\Delta T_2 - \Delta T_1) (1 - V_2)^2) \\ & + \frac{\lambda}{\Delta T_f} (\Delta T_f - \Delta T_2) (1 - V_3)^2. \end{aligned} \quad (11)$$

The proof of this proposition is provided in the appendix.

Remark 2: For $\Lambda = 450$ and $N = 50$ we get $\lambda = 0.18$. By putting appropriate values of V_i 's and ΔT_i 's corresponding to WECC criterion, the critical threshold of the KL divergence is obtained as $\mathcal{K}^* = 4.9$. It also can be noted that the critical value \mathcal{K}^* is not only a function of the parameters of the WECC voltage performance criterion, but also depends on the parameters λ , and N . The value \mathcal{K}^* is an increasing function of λ . It has been shown that, if the KL divergence is more than the \mathcal{K}^* , WECC criterion would be violated. Through extensive simulations, it is verified that the converse statement is also true in most cases, i.e. Almost all cases, satisfying the WECC criterion, would produce KL divergence less or equal \mathcal{K}^* . In order to achieve a more tighter sufficiency condition, the time interval can be divided into subdivisions, and the KL divergence can be computed for each one of these intervals. For each of these intervals, we can find corresponding critical KL values. In that case, if the KL divergence values for different intervals are within the prescribed critical bounds, the WECC criterion would almost surely be met.

V. APPLICATIONS

The framework developed for characterizing the voltage recovery phenomenon has been tested using IEEE 162 bus system and one of the applications of KL measure for system analysis has been presented in this section. The IEEE 162

bus test system has 17 generators, 111 loads, 34 shunts, and 238 branches. The power flow and dynamics data for the 162 bus system are available in [15]. For a more accurate load representation, the 22 load buses were stepped down through distribution transformers to the 12.47 kV level, and the new low voltage buses were assigned the numbers 163 through 174. To capture the dynamic behavior of motor loads, composite load model represented by CMDL [16] was used at the new representative load buses in the dynamic simulation studies.

The modified IEEE 162 bus system has N_B buses (184) and a total of N_C contingencies (316) of the type, a three phase fault at a bus which is cleared after 6 cycles by opening one of the transmission lines connected to the faulted bus are considered for simulation studies. The voltage time series corresponding to bus i and contingency j are stored in the vector $v_{ij}(t)$, $0 \leq t \leq T_f$, $i \in N_B$, $j \in N_C$. T_f represents the final simulation time instant which is chosen as 5 second for all the simulations.

Using the time domain simulation results (v_{ij}), KL divergence at each bus is computed for all contingencies and the results are stored in a matrix $\bar{K} \in \mathbb{R}^{N_B \times N_C}$. Figure 10 shows the pictorial representation of the \bar{K} matrix, where each row has KL measure corresponding to a particular bus for all the contingencies and each column has KL measure for all the buses corresponding to a particular contingency. The element K_{ij} corresponds to the KL divergence measure for i^{th} bus and j^{th} contingency.

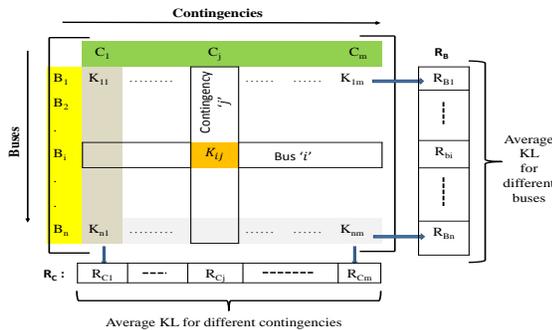


Fig. 10. Representation of \bar{K} matrix where each row has KL numbers of a particular bus for all contingencies and each column has KL numbers of all buses corresponding to a particular contingency

The recovery information from the time series has been captured in a scalar form using KL divergence measure. This will greatly reduce the burden of analyzing time series data especially when dealing with multiple contingencies and scenarios. For example, in our simulation studies, each contingency results in a time series data of length 2500 for every bus. By computing KL divergence the key information in the huge set of time series data is captured in a single number corresponding to each bus in the system. Further, these KL numbers can be used to ascertain the behavior of different contingencies and buses in the system.

The average value of the KL divergence for individual buses, for all contingencies could be used to determine the relative severity of an individual bus. Equation (12) is used to compute the average KL divergence value for all the buses, where $\vec{1} \in \mathbb{R}^{N_C}$ is a row vector of all entries equal to 1.

The i^{th} entry of the vector R_B contains the average value of the KL divergence for the i^{th} bus, when averaged over all contingencies.

$$R_B := \frac{1}{N_C} \vec{1} * \bar{K}, R_B \in \mathbb{R}^{N_C}, \vec{1} \in \mathbb{R}^{N_C} \quad (12)$$

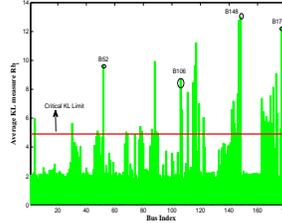


Fig. 11. Average value of KL divergence for different buses

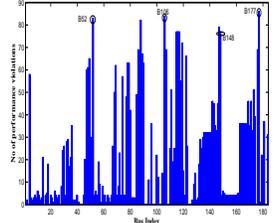


Fig. 12. Number of performance violations for different buses

Figure 11 shows the R_B vector plotted against different buses. The red horizontal line corresponds to the critical value of the KL divergence, which is derived as 4.9 (Section IV). The buses whose average KL values are greater than the critical KL value can be termed as severe buses. However it is not sufficient to conclude the severity of buses only based on the average KL values because of the masking effect. The number of contingencies that results in the performance violations at a particular bus is also important to decide the severity level of a particular bus. Figure 12 shows the number of performance violations at all the buses. For example, from fig.11, it is found that bus 148 has the highest average KL value and bus 106 has relatively less average KL value. However from fig.12, it is observed that more number of contingencies creates violation at bus 106 than at bus 148. This example shows that decisions solely based on average KL value or number of performance violations is not sufficient to arrive at critical buses.

TABLE I
TOP 8 SEVERE BUS IDS (SORTED IN DESCENDING SEVERITY)

sorted by	148	147	177	117	88	52	116	173
avg. KL								
sorted by	177	106	52	88	147	115	116	117
violations								

Table I shows the top 8 severe bus numbers, sorted based on average KL values (row 1) and number of performance violations (row 2). Although, there are common elements between the lists the ranking of buses is different in both the lists. There are multiple options to combine these two information (severity and number of violations) to arrive at the list of critical buses. When the average KL value is below critical and number of violations are very small, then such buses are termed as non-severe and they can be disregarded for further analysis. Buses having high average KL values and more number of performance violations (e.g Bus 177) can be grouped as the most critical buses. Such information can be very valuable in selecting the locations to be monitored for voltage performance violations. The above example has

showed that KL can be a useful measure to compare the voltage recovery behavior of different buses under different contingencies.

Similarly for deciding the critical contingencies both severity of contingency and number of performance violations have been considered. The severity of a contingency is defined as average KL values of all the buses corresponding to that contingency. Equation (13) is used to calculate the average KL value for different contingencies. The j^{th} entry in R_C vector captures the severity level for contingency j^{th} . The severity of different contingencies and the number of buses that violate the WECC performance criterion are shown in Fig. 13 and Fig. 14 respectively.

$$R_C := \frac{1}{N_B} \bar{K} * \vec{1}, R_C \in \mathbb{R}^{N_C}, \vec{1} \in \mathbb{R}^{N_B} \quad (13)$$

If the number of violations for a particular contingency is

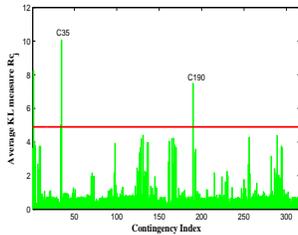


Fig. 13. Average value of KL divergence for different contingencies

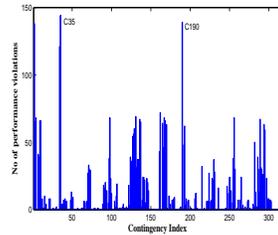


Fig. 14. Number of performance violations for different contingencies

small, the recovery is tolerable for that contingency provided the severity level is small. Contingencies that have lower severity value and small number of violations are termed as non-severe. A higher severity value and fewer number of violations signifies that those contingencies are severe for only certain buses. Contingencies that have lower severity value and has larger number of violations affects a wider region of the network (e.g C190 - Refer fig .13 & 14). The most critical contingencies are those with higher severity value and also it has large number of performance violations (e.g. C35).

TABLE II
TOP 8 SEVERE CONTINGENCY IDS (SORTED IN DESCENDING SEVERITY)

sorted by avg. KL	35	2	190	34	131	289	256	167
sorted by violations	35	190	2	34	162	131	4	98

Table II provides the indices of top 8 severe contingencies, sorted based on average KL value for contingencies and number of performance violations. The number of performance violations and severity level can be used to identify critical contingencies. Non-severe contingencies can be eliminated for further system level analysis.

VI. CONCLUSION

In this work, FIDVR has been shown as a phenomenon of loss of entropy in the voltage time series. KL divergence, an entropy based quantity, has been used to characterize both the

rate and level of voltage recovery. KL divergence provides a quantitative measure of voltage recovery phenomenon and is very useful for comparing different voltage waveforms. A critical value of the KL divergence based on WECC criteria has been derived. This will be very useful to identify the degree of WECC voltage performance in a quantitative way. KL divergence calculations can be extended to identify the degree of violations during the recovery phase and steady state settling phase. One of the applications of KL divergence for system level studies, identification of critical contingencies and buses, has been presented. Characterization of voltage waveforms using KL divergence can be very useful in system level studies dealing with large sets of data such as PMU data, planning studies etc. KL divergence, alongside statistical tools, can be used to classify contingencies and buses, according to nature of recovery. This classification of contingencies and buses would reduce substantially the further data processing, for planning and optimization.

REFERENCES

- [1] U. S. Department of Energy – NERC FIDVR Workshop, Apr.22, 2008 [Online]: <http://www.nerc.com/files/FIDVR-Conference-Presentations-4-22-08.pdf>
- [2] U. S. Department of Energy – NERC FIDVR Conference, Sep, 29 2009. [Online]: <http://www.nerc.com/files/FIDVR-Conference-Presentations-9-29-09.pdf>
- [3] NERC Transmission Issues Subcommittee and System Protection and Control Subcommittee – “*Fault-Induced Delayed Voltage Recovery*”, Technical reference paper, version 1.2, Princeton, NJ, June 2009.
- [4] M. Glavic, D. Novosel, E. Heredia, D. Kosterev, A. Salazar, F. Habibi-Ashrafi, and M. Donnelly, “See it fast to keep calm: Real-time voltage control under stressed conditions,” *IEEE Power and Energy Magazine*, vol. 10, no. 4, pp. 43–55, 2012.
- [5] R. Bravo, R. Yinger, D. Chassin, H. Huang, N. Lu, I. Hiskens, and G. Venkataramanan, “Load modelling transmission research - final project report,” Lawrence Berkeley National Laboratory., Tech. Rep., March 2010.
- [6] S. Halpin, K. Harley, R. A. Jones, and L. Taylor, “Slope-permissive under-voltage load shed relay for delayed voltage recovery mitigation,” *Power Systems, IEEE Transactions on*, vol. 23, no. 3, pp. 1211–1216, 2008.
- [7] H. Bai and V. Ajjarapu, “A novel online load shedding strategy for mitigating fault-induced delayed voltage recovery,” *Power Systems, IEEE Transactions on*, vol. 26, no. 1, pp. 294–304, 2011.
- [8] S. Halpin, R. A. Jones, and L. Taylor, “The mva-volt index: A screening tool for predicting fault-induced low voltage problems on bulk transmission systems,” *Power Systems, IEEE Transactions on*, vol. 23, no. 3, pp. 1205–1210, 2008.
- [9] L. K. Nash, *Elements of Statistical Thermodynamics*. Addison-Wesley, 1968.
- [10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Springer, 1991.
- [11] C. Shannon, “A mathematical theory of communication,” *ACM SIG-MOBILE Mobile Comput. Commun. Rev.*5.
- [12] G. Froyland and K. Padberg-Gehle, “Finite-time entropy: A probabilistic approach for measuring nonlinear stretching,” *Physica D: Nonlinear Phenomena*, vol. 241, no. 19, pp. 1612 – 1628, 2012.
- [13] S. Kullback and R. A. Leibler, “On information and sufficiency,” *Annals of Mathematical Statistics*, March 1951.
- [14] D.J. Shoup, J.J. Paserba, and C.W. Taylor. A survey of current practices for transient voltage dip/sag criteria related to power system stability. In *Power Systems Conference and Exposition, 2004. IEEE PES*, pages 1140 – 1147 vol.2, oct. 2004.
- [15] V. Vittal, “Transient stability test systems for direct stability methods,” *IEEE Transactions on Power Systems*, vol. 7, no. 1, pp. 37 –43, 1992.
- [16] *Siemens PTI Power Technologies Inc., PSS/E 33, Program Application Guide, Vol. II, May 2011.*

VII. APPENDIX

The proof of claim 1 is outlined below,

Proof: The voltage evolves according to the following equation,

$$v(t) = v_{max} - (v_{max} - v(0))e^{-\alpha t}.$$

where, $\alpha > 0$ is the rate of recovery. The time instant when voltage value reaches v_i is denoted as t_i i.e. $v(t_i) = v_i$.

This gives,

$$v_2 = v(t_2) = v_{max} - (v_{max} - v(0))e^{-\alpha t_2}$$

$$\alpha \Delta t_1 = \alpha t_2 = \ln \frac{v_{max} - v(0)}{v_{max} - v_2} = \ln \beta_1$$

It is to be noted that, $\beta_2 > 0$ which makes, $\Delta t_i > 0$. Now, for $i = 2, \dots, N-1$,

$$v_i = v(t_i) = v_{max} - (v_{max} - v(0))e^{-\alpha t_i}$$

$$v_{i+1} = v(t_{i+1}) = v_{max} - (v_{max} - v(0))e^{-\alpha t_{i+1}}$$

The above two equations imply,

$$\alpha \Delta t_i = \alpha(t_{i+1} - t_i) = \ln \frac{v_{max} - v_i}{v_{max} - v_{i+1}} = \ln \beta_i$$

where, $\beta_i = \frac{v_{max} - v_i}{v_{max} - v_{i+1}} > 0$. Finally the case is considered when $i = N$. It can be shown similarly,

$$\alpha \Delta t_N = \ln \frac{v_{max} - v_N}{v_{max} - v(T)} = \ln \beta_N$$

Finally, the following equation could be obtained.

$$\alpha \Delta t_i = \ln \beta_i \quad (14)$$

where,

$$\beta_i = \frac{v_{max} - v(0)}{v_{max} - v_2}, \quad i = 1$$

$$= \frac{v_{max} - v_i}{v_{max} - v_{i+1}}, \quad i = 2, \dots, N-1$$

$$= \frac{v_{max} - v_N}{v_{max} - v(T)}, \quad i = N.$$

From Equation (14), $p_i = \frac{1}{\alpha T} \cdot \ln \beta_i$

This implies,

$$H(\alpha, T) = - \sum_{i=1}^N p_i \ln p_i = - \sum_{i=1}^N \frac{\ln \beta_i}{\alpha T} (\gamma_i - \ln(\alpha T))$$

where, $\gamma_i = \ln(\ln \beta_i)$. Taking partial derivative of H w.r.t. α ,

$$\frac{\partial H(\alpha, T)}{\partial \alpha} = \frac{1}{\alpha^2 T} \cdot \sum_{i=1}^N \ln \beta_i \gamma_i + \sum_{i=1}^N \frac{\ln \beta_i}{\alpha^2 T} - \sum_{i=1}^N \frac{\ln \beta_i \ln(\alpha T)}{\alpha^2 T}$$

$$= \frac{1}{\alpha^2 T} \cdot \sum_{i=1}^{N-1} \gamma_i \ln \beta_i + \sum_{i=1}^{N-1} \frac{\ln \beta_i}{\alpha^2 T} - \sum_{i=1}^{N-1} \frac{\ln \beta_i \ln(\alpha T)}{\alpha^2 T}$$

$$+ \frac{1}{\alpha^2 T} \cdot \gamma_N \ln \beta_N + \frac{\ln \beta_N}{\alpha^2 T} - \frac{\ln \beta_N \ln(\alpha T)}{\alpha^2 T}$$

Now,

$$\beta_N = \frac{v_{max} - v_N}{v_{max} - v(T)} = e^{\alpha T} \frac{v_{max} - v_N}{v_{max} - v(0)}$$

$$\ln \beta_N = \alpha T + \ln \left(\frac{v_{max} - v_N}{v_{max} - v(0)} \right)$$

$$\ln \beta_N = \alpha T + c_N$$

This implies,

$$\frac{\partial H(\alpha, T)}{\partial \alpha} = \frac{1}{\alpha^2 T} \cdot \sum_{i=1}^N \ln \beta_i \gamma_i + \sum_{i=1}^N \frac{\ln \beta_i}{\alpha^2 T} - \sum_{i=1}^N \frac{\ln \beta_i \ln(\alpha T)}{\alpha^2 T}$$

$$= \frac{1}{\alpha^2 T} \left(\sum_{i=1}^{N-1} \gamma_i \ln \beta_i + \sum_{i=1}^{N-1} \ln \beta_i + \gamma_N c_N + c_N \right)$$

$$- \frac{\ln(\alpha T)}{\alpha^2 T} \left(c_N + \sum_{i=1}^{N-1} \ln \beta_i \right) + \frac{1 + \gamma_N}{\alpha} - \frac{\ln(\alpha T)}{\alpha}$$

$$= \frac{1}{\alpha^2 T} a_{N1} - \frac{\ln(\alpha T)}{\alpha^2 T} a_{N2} + a_{N3} - \frac{\ln(\alpha T)}{\alpha}$$

where, a_{N1} , a_{N2} , and, a_{N3} are constants. Clearly after sufficiently large T , the quantity $\frac{\partial H(\alpha, T)}{\partial \alpha}$ would be dominated by $-\frac{\ln(\alpha T)}{\alpha}$ because rest of the terms are either constant or inversely proportional to T (This contain both the terms proportional to $\frac{1}{\alpha T}$ and $\frac{\ln(\alpha T)}{\alpha T}$). After sufficiently large T , $\frac{\ln(\alpha T)}{\alpha T}$ will decay as it would be dominated by the αT in denominator). Finally for sufficiently large T ,

$$\frac{\partial H(\alpha, T)}{\partial \alpha} \approx - \frac{\ln(\alpha T)}{\alpha} < 0$$

This implies entropy decreases with faster recovery (higher α). Hence the proof. ■

Proof of Proposition 1:

Proof: Let us consider the voltage curve, that forms the envelope of \mathcal{E} .

$$v_*(t) = V_1, \quad T_{cl} \leq t < T_1,$$

$$v_*(t) = V_2, \quad T_1 \leq t < T_2, \quad V_2 > V_1,$$

$$v_*(t) = V_3, \quad T_2 \leq t \leq T_f, \quad V_3 > V_2.$$

The KL divergence, corresponding to $v_*(t)$, is given by \mathcal{K}^* , which is given by (11). The KL divergence for any voltage signal $v(t)$, that satisfies \mathcal{E} , the time spent on lower voltage levels would be less than that of $v_*(t)$. This gives, $\mathcal{K} \leq \mathcal{K}^*$. Next, \mathcal{K}^* is to be expressed in terms of the various parameters, specified by WECC. Let the probability corresponding to the $v_*(t)$ be denoted as p_i^* , where,

$$p_i^* = \frac{\Delta T_1}{\Delta T_f}, \quad v_i \leq V_1 \leq v_{i+1},$$

$$= \frac{\Delta T_2}{\Delta T_f}, \quad v_i \leq V_2 \leq v_{i+1},$$

$$= \frac{\Delta T_3}{\Delta T_f}, \quad v_i \leq V_3 \leq v_{i+1},$$

$$= 0, \quad \text{Otherwise.}$$

Plugging these values of the probability into 8, we get the expression for \mathcal{K}^* . ■