Data-Driven Joint Voltage Stability Assessment Considering Load Uncertainty: A Variational Bayes Inference Integrated With Multi-CNNs

Mingjian Cui, *Senior Member, IEEE*, Fangxing Li, *Fellow, IEEE*, Hantao Cui, *Senior Member, IEEE*, Siqi Bu, *Senior Member, IEEE*, and Di Shi, *Senior Member, IEEE*

***Abstract*—Few studies have focused on assessing the transient and steady-state voltage stability status of dynamic systems simultaneously. This motivated us to propose a new concept referred to as joint voltage stability assessment (JVSA). Towards this end, this paper proposes a novel data-driven JVSA method considering load uncertainty. It combines multiple convolutional neural networks (multi-CNNs) and a novel variational Bayes (VB) inference for better JVSA accuracy. First, the multi-CNN model is utilized to fast estimate the maximum voltage deviations during the transient and steady-state process. Uncertain load scenarios and system topology under N-1 contingency with are chosen as inputs of each CNN model. Second, estimated voltage deviations are put into the VB inference to automatically infer the transient and steady-state voltage stability status. To validate its effectiveness, numerical simulations are performed on the modified WECC 179-bus system by comparing with benchmark algorithms. It is demonstrated that the proposed data-driven JVSA method is more accurate and faster than the conventional VSA method.**

***Index Terms*—Fault-induced delayed voltage recovery, voltage stability assessment, load uncertainty, variational Bayes inference, convolutional neural network.**

I. Introduction

As DISTRIBUTED energy resources and utility-scale renewables reach increasingly high levels of penetration in transmission networks, voltage stability assessment (VSA) is encountering new challenges especially in the area of fault-induced delayed voltage recovery (FIDVR). In the literature, the VSA problem has been divided into two categories: short-term voltage stability (STVS) assessment and long-term voltage stability (LTVS) assessment. STVS assessment is capable of retaining synchronism subject to significant disturbances and faults. It aims to mathematically solve a set of non-linear and high-dimensional differential algebraic equations (DAEs) [1]–[3]. Research on transient VSA facilitates power system operators decision-making on and coordination of voltage control strategies. In most cases, the steady-state VSA is taken as a static security problem in the literature [4]. It is handled by solving the objective optimization model or security region. Steady-state voltage alarming limits are predefined based on normalized voltage deviation.

In the area of STVS research, Han *et al.* [5] formulated a multi-objective optimization model to improve STVS level for dynamic VAR planning. Yu *et al.* [6] developed a transient stability assessment (TSA) method using a temporal self-adaptive scheme. Pico *et al.* [7] assessed the STVS of multi-machine multi-converter power systems from simulations considering a phase locked loop (PLL), AC- and DC-side dynamics, and closed-loop controllers. Wang *et al.* [8] presented a projective integration method for the efficient STVS simulation of power systems with high penetration of distributed generators. Wang *et al.* [9] proposed a probabilistic methodology for integrated reliability evaluation considering resource adequacy and TSA in a unified framework. Dutta [10] described an approach utilizing clustering to extract common features in the voltage magnitude signals to identify outliers.

By comparison, in the area of LTVS research, Yan *et al.* [11] designed a composite security index based on voltage stability analysis and corresponding alarming limits. Li *et al.* [12] studied the impact of coupled transmission-distribution on VSA via physical analysis and numerical simulation considering the interactions between transmission and distribution networks. Oliveira *et al.* [13] utilized a multiway decision tree approach to assess power system operation steady-state security for multiple contingencies. Luo *et al.* [14] presented a multi-objective optimization model based on DC power flow to maximize the active power steady-state security region whilst minimizing the total generation cost.

Though many VSA strategies have been proposed and studied in the literature, there are still challenges to be addressed. Very little research simultaneously considers both STVS and LTVS assessment, though simultaneous assessment of voltage stability is a significant practical problem for real-time power system applications. In this paper, the joint voltage stability assessment

Manuscript received December 20, 2020; revised May 12, 2021 and August 17, 2021; accepted September 2, 2021. Date of publication September 13, 2021; date of current version April 19, 2022. This work was supported in part by SGCC Science & Technology Project SGCC-5700-201958523A-0-0-00. Paper no. TPWRS-02081-2020. *(Corresponding author: Fangxing Li.)*

Mingjian Cui and Fangxing Li are with the Department of Electrical Engineering and Computer Science, The University of Tennessee, Knoxville, TN 37996 USA (e-mail: mingjian.cui@ieee.org; fli6@utk.edu).

Hantao Cui is with the School of Electrical and Computer Engineering at Oklahoma State University, Stillwater, OK 74078 USA (e-mail: h.cui@okstate.edu).

Siqi Bu is with the Department of Electrical Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong (e-mail: Siqi.Bu@gmail.com).

Di Shi is with the GEIRI North America, San Jose, CA 95134 USA (e-mail: sdxjtu@gmail.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPWRS.2021.3111151.

Digital Object Identifier 10.1109/TPWRS.2021.3111151

(JVSA) problem is defined to simultaneously assess the system voltage stability during the transient and steady-state processes. The reasons of simultaneously assessing voltage stability are twofold:

i) Although much of the literature has focused on short-term (transient) VSA, LTVS (steady-state VSA) after the transient process has not yet been considered in those studies. It is highly possible that one contingency may cause the transient voltage instability even if the steady-state voltage is stable as the transient process is neglected (i.e., steady-state screening gives positive result while the transit process may diverge or go out of bound).
ii) Similarly, a contingency event may derive transient VSA while the steady-state voltage is unstable (i.e., steady-state screening gives non-convergent results that go out of bound). That is, one contingency may cause the steady-state voltage instability even if the transient voltage is stable.

It is significantly necessary to consider load uncertainty in VSA as the reactive power of load demand is highly related with the system voltage and its stability. Zhang *et al.* [15] discussed the voltage stability considering load parameters' uncertainty to obtain probabilistic characteristics of eigenvalues. Pierrou and Wang [16] studied the impact of stochastic load and renewable generation uncertainty on the dynamic voltage stability margin. Adusumilli and Kumar [17] studied VSA through continuation power flow analysis by considering uncertainties associated with active and reactive power injections at all buses in the system. Chi [18] studied the candidate bus selection for dynamic VAR planning towards voltage stability enhancement considering both wind and load uncertainties.

Recently, deep/machine learning-based data-driven methods with post-contingency measurements have received more attention with respect to VSA. The most important advantage of those methods is that they can provide significantly faster assessment speed [19], [20] when solving a specific dynamic simulation model with a large number of buses or complicated optimization models. A convolutional neural network (CNN) was designed in [11] to fast evaluate the system security status based on observation information. A long short-term memory (LSTM) network based recurrent neural network (RNN) was developed in [6] for TSA to learn temporal data dependencies of the input synchrophasor measurements. A hierarchical deep learning machine was developed in [21] to achieve both quantitative and qualitative online transient stability prediction.

To the best of the author's knowledge, this paper represents the first effort to study the JVSA problem using a data-driven approach simultaneously. In this paper, we seek to address two critical questions for VSA. (i) Is it possible to simultaneously cope with the transient and steady-state VSA problem considering load uncertainty? (ii) How can we improve the JVSA speed using state-of-the-art deep learning techniques? Towards this end, this paper proposes a novel data-driven fast JVSA approach considering load uncertainty. The main contributions and innovations of this paper include:

i) A multi-CNN model is proposed to quickly estimate the maximum voltage deviations during the transient and steady-state process. A novel loss function is designed to involve both transient and steady-state voltage deviations during iterations. The developed CNN model can significantly improve both the computational efficiency and accuracy for VSA. Voltage deviation variables in the dynamic process can be obtained by approximating solutions using the developed CNN function.
ii) We develop a novel variational Bayes (VB) inference to simultaneously assess both STVS and LTVS based on voltage deviation estimates. Better assessment accuracy can be obtained by a comparative study with several existing algorithms. The developed copula VB (CVB) algorithm takes all moments of transient and steady-state voltage deviations into account, while existing inference algorithms only consider the first and second moments of voltage deviations.
iii) The proposed method can quantitatively estimate the voltage deviation and then use the quantitative results to infer the final stability status. Unlike existing methods that only assess the stability status, not only the final stability status but also the critical stability status can be obtained by the proposed method. Thus, critical warnings can additionally be provided for power system operators.

The organization of this paper is as follows. In Section II, the methodology framework is introduced based on the joint transient and steady-state VSA index. Section III presents the multi-CNN model to fast estimate voltage deviation during the transient and steady-state process. Section IV describes the novel VSA method based on estimated voltage deviations. Case studies and result analysis performed on the WECC 179-bus system are discussed in Section V. Concluding remarks are summarized in Section VI.

II. Proposed Methodology Framework

Based on the post-fault voltage recovery performance, the WECC/NERC planning standards have proposed criteria to judge voltage stability status under one N-1 scenario. Inspired by these WECC/NERC standards, we propose taking advantage of the maximum percentage voltage deviation during the transient and steady-state processes for data-driven JVSA. The percentage voltage deviation ΔV_t^j of bus j at time t is:

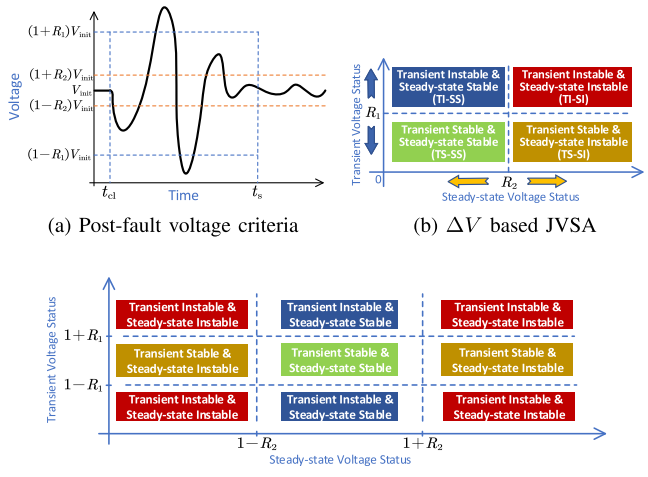
$$\Delta V_t^j = |(V_t^j - V_{\text{init}}^j)/V_{\text{init}}^j| \times 100\% \tag{1}$$

where V_{init}^j is the pre-fault initial voltage of bus j. V_t^j is the original voltage of bus j at time t. As shown in Fig. 1(a) (a case with LTVS but transient voltage instability), the following criteria has been considered to quantitatively assess both STVS and LTVS for post-fault voltage trajectories [22]:

$$\Delta V_t^j \le R_1, \ \ if \ \forall t, \ t_{\text{cl}} \le t \le t_{\text{s}} \tag{2a}$$

$$\Delta V_t^j \le R_2, \ \ if \ \forall t, \ t > t_{\text{s}} \tag{2b}$$

where t_{cl} is fault clearing time. t_{s} is the post-transient time and set at 3 seconds [23]. R_1 and R_2 are typical parameters to assess voltage stability (R_1 for STVS and R_2 for LTVS). In this paper, we define that if voltage deviation ΔV_t^j conforms to $\{\Delta V_t^j \le R_1\} \cap \{\Delta V_t^j \le R_2\}$, both STVS and LTVS can be obtained. If ΔV_t^j conforms to $\{\Delta V_t^j \le R_1\} \cap \{\Delta V_t^j > R_2\}$,



(a) Post-fault voltage criteria (b) ΔV based JVSA

(c) Original voltage V based JVSA

Fig. 1. Schematic diagram of JVSA comparing original voltage V with voltage deviation ΔV.

only STVS can be obtained. If ΔV_t^j conforms to $\{\Delta V_t^j > R_1\} \cap \{\Delta V_t^j \leq R_2\}$, only LTVS can be obtained. Otherwise, we can assess both transient and steady-state voltage instability. For the sake of simplicity, voltage deviation ΔV is more suitable for JVSA compared with original voltage V as shown in Figs. 1(b) and 1(c). In Fig. 1(b), the voltage stability statuses include four parts: Transient Stable & Steady-state Stable (TS-SS), Transient Stable & Steady-state Instable (TS-SI), Transient Instable & Steady-state stable (TI-SS), and Transient Stable & Steady-state Instable (TI-SI).

Remark 1: *Essentially, the maximum voltage deviation can be used to assess the voltage stability, that is,* $\Delta V_T = \max(\Delta V_t|_{t_{cl}}^{t_s})$ *and* $\Delta V_S = \max(\Delta V_t|_{t_s}^{\infty})$. *The criteria in (2a) and (2b) are simplified as TS-SS:* $\Delta V_T^s \leq R_1 \cap \Delta V_S^s \leq R_2$; *TS-SI:* $\Delta V_T^s \leq R_1 \cap \Delta V_S^s > R_2$; *TI-SS:* $\Delta V_T^s > R_1 \cap \Delta V_S^s \leq R_2$; *and TI-SI:* $\Delta V_T^s > R_1 \cap \Delta V_S^s > R_2$.

Based on the joint voltage deviation vector $\Delta \boldsymbol{V} = [\Delta V_{\mathrm{T}}, \Delta V_{\mathrm{S}}]$, we propose a data-driven JVSA methodology considering load uncertainty by integrating multi-CNNs and the CVB inference. The framework is shown in Fig. 2. Three major steps are briefly summarized, including:

Step 1: Multi-CNNs are used to fast estimate voltage deviation $\Delta \boldsymbol{V}$, including ΔV_{T} in the transient process and ΔV_{S} in the steady-state process. For multiple N-1 contingency scenarios (total number is N), there are $2N$ CNN models trained for JVSA. Inputs of each CNN include known parameters: bus active/reactive load and system topology, while the output of each CNN is the voltage deviation ΔV_{T} or ΔV_{S}, respectively.

Step 2: Based on estimated voltage deviation ΔV_{T} and ΔV_{S}, a CVB-based inference is designed to automatically assess both transient and steady-state voltage stability statuses. Inputs are the estimated voltage deviation ΔV_{T} and ΔV_{S}, while the output of CVB is the estimated voltage stability status (including TS-SS, TS-SI, TI-SS, and TI-SI).

Step 3: Estimated JVSA results under N-1 contingency scenarios considering load uncertainty are obtained and compared with several benchmarks to validate the effectiveness of the proposed data-driven JVSA methodology.

Detailed information about each step is described in the following sections.

III. Multi-CNNs Based Voltage Deviation Estimate

It has been demonstrated in the literature that machine learning methods can significantly improve both the computational efficiency and accuracy for both LTVS [11] and STVS assessment [10], [19], [21]. As an efficient regression machine learning method, a CNN has strong capabilities in processing grid-like topology data. Thus, system state variables (i.e., voltage deviation) in the dynamic process can be obtained by approximating solutions of DAEs using the CNN function. This is because power system topology has a grid-like structure and the voltage level at one bus is highly dependent on its adjacent buses.

Inspired by the mapping structure of the CNN in [11], [24], the input consists of the known parameters: active/reactive load $\boldsymbol{P}_{\mathrm{d}} = [P_{\mathrm{d}1}, P_{\mathrm{d}2}, \ldots, P_{\mathrm{d}n_{\mathrm{b}}}]$, $\boldsymbol{Q}_{\mathrm{d}} = [Q_{\mathrm{d}1}, Q_{\mathrm{d}2}, \ldots, Q_{\mathrm{d}n_{\mathrm{b}}}]$, and bus susceptance matrix $\boldsymbol{B}$. System topology is represented by the self-susceptance, which is the diagonal element of $\boldsymbol{B}$:

$$\mathrm{diag}\left(\mathrm{imag}\left(\boldsymbol{Y}\right)\right) = \mathrm{diag}\left(\boldsymbol{B}\right) = [b_{11}, b_{22}, \ldots, b_{n_{\mathrm{b}}n_{\mathrm{b}}}] \tag{3}$$

where n_{b} is the total number of system buses. $\boldsymbol{Y}$ is the bus admittance matrix. Thus, the input of the CNN is given by $[\boldsymbol{P}_{\mathrm{d}}; \boldsymbol{Q}_{\mathrm{d}}; \mathrm{diag}(\boldsymbol{B})]$, which is a $3 \times n_{\mathrm{b}}$ matrix.

The objective function of the CNN is to minimize a loss function to formulate an accurate regression model. The loss function of the CNN in this paper is designed as:

$$LOSS = \sum_{s=1}^{N\mathrm{s}} \left(\sum_{j=1}^{n_{\mathrm{b}}} \left(\Delta V_{j,s}^{*} - \Delta V_{j,s}\right)^2 / n_{\mathrm{b}} \right) / N\mathrm{s} \tag{4}$$

where $\Delta V_{j,s}^{*}$ is the expected voltage deviation output of the CNN. $N\mathrm{s}$ is the total number of training samples. A lower loss function value denotes higher accuracy of the training CNN model.

In a CNN model, the convolutional filter seeks to extract feature maps. The equation of feature extraction is given by:

$$I'(i,j) = \sum_{u=0}^{c-1} \sum_{v=0}^{c-1} I(u+i, v+j) \cdot \omega(u,v) + b \tag{5}$$

where $I(u,v)$ is a single unit calculated from the original input. $I'(i,j)$ is an updated single unit in feature map I of the convolutional filter. $\omega(u,v)$ is the weighted parameter in the filter. c is the size of the filter. b is the bias of the single unit. Note that as CNNs have been widely used in existing literature, we do not go into detail on CNNs in this paper. Detailed information on solving the parameter ω and the bias b can be found in [11] using the back-propagation algorithm. The structure of the developed CNN is illustrated in Fig. 3, where an example of estimating transient voltage deviation ΔV_{T}^s is shown.

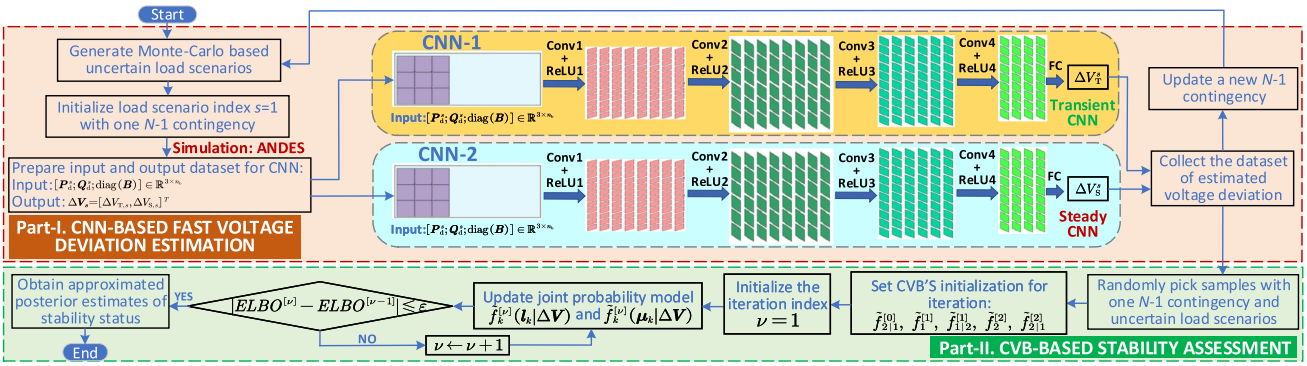


Fig. 2. Framework of the proposed data-driven JVSA methodology based on multi-CNNs and CVB inference.

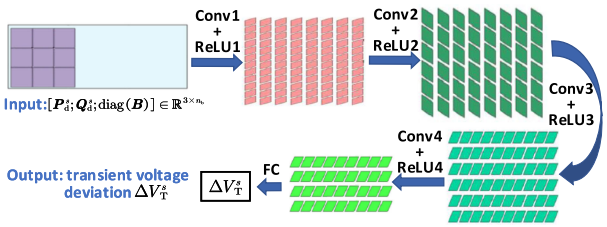


Fig. 3. CNN structure: an example of estimating transient voltage deviation $\Delta V_{\mathrm{T}}^{s}$.

Unlike existing CNN-based VSA methods, estimating voltage deviation proposed in this paper is more reasonable than directly predicting the stable status. This is because it is very challenging to judge the critical stability status which is close to the threshold R_1 or R_2 by only predicting the stability status. The quantitative voltage deviation can provide additional critical warnings for power system operators with fine granularity instead of a binary variable indicating stability status.

IV. CVB-Based VSA Methodology

After obtaining multiple voltage deviations under different load scenarios, manually assessing the voltage stability status of system buses one by one remains challenging. Thus, we have developed a CVB inference based assessment method integrated with the multi-CNN model.

The CVB algorithm is a relaxed form of mean-field approximation [25]. It utilizes the same tractable iterative scheme without being constrained by independent form, as compared to the conventional VB algorithm. It is also a free-form approximation and can automatically adapt to the true posterior distribution of voltage deviations until convergence. CVB can yield iterative closed-form solutions directly. The key advantage of the CVB algorithm is that it can return an optimal mixture of approximated distributions. It does not initially impose any particular form and provides estimates of higher-order moments as opposed to mere point estimates. As the true posterior distribution of voltage deviations is unknown but highly needed, we use the CVB algorithm to approximate it and consequently assess the voltage stability status.

Assuming that the estimated sample vector of voltage deviation is given by $\Delta \boldsymbol{V} = [\Delta \boldsymbol{V_1}, \ldots, \Delta \boldsymbol{V_s}, \ldots, \Delta \boldsymbol{V_{Ns}}] \in \mathbb{R}^{2\times Ns}$ and $\Delta \boldsymbol{V_s} = [\Delta V_{\mathrm{T},s}, \Delta V_{\mathrm{S},s}]^T \in \mathbb{R}^2$ under the sth load scenario $s \in \{1, 2, \ldots, Ns\}$. The estimated $\Delta \boldsymbol{V_s}$ belongs to one of the four voltage stability statuses k with equal probability $p_k \in \boldsymbol{p} = [p_1, p_2, p_3, p_4]^T$, i.e., $p_k = \frac{1}{4}$.

To determine the voltage stability status of one sample s (i.e., one N-1 contingency scenario), we define a temporal voltage stability status matrix $\boldsymbol{L} = [\boldsymbol{l}_1, \ldots, \boldsymbol{l}_s, \ldots, \boldsymbol{l}_{Ns}] \in \mathbb{I}^{4\times Ns}$ with labels $l_{k,s} \in \boldsymbol{l}_s = [l_{1,s}, l_{2,s}, l_{3,s}, l_{4,s}]^T =$ [TS-SS,TS-SI,TI-SS,TI-SI]^T. Given one boolean vector $\boldsymbol{\epsilon}_k$ with a kth non-zero element, i.e., $\boldsymbol{\epsilon}_k = [0, \ldots, 1, \ldots, 0]^T \in \mathbb{I}^4$, we set $\boldsymbol{l}_s = \boldsymbol{\epsilon}_k$ if $\Delta \boldsymbol{V_s}$ can determine the kth voltage stability status (i.e., TS-SS, TS-SI, TI-SS, and TI-SI). Considering load uncertainty, we set unknown means of voltage stability statuses as $\boldsymbol{\mu}_k \in \boldsymbol{\Lambda} = [\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, \boldsymbol{\mu}_4] \in \mathbb{R}^{2\times 4}$, such that $\boldsymbol{\mu}_k = [\mu_{\mathrm{T},k}, \mu_{\mathrm{S},k}]^T = \frac{1}{Ns} \times [\sum_s^{Ns} \Delta V_{\mathrm{T},s}, \sum_s^{Ns} \Delta V_{\mathrm{S},s}]^T$. Thus, we obtain the unknown parameters set $\boldsymbol{\Phi} = [\boldsymbol{\Lambda}, \boldsymbol{L}]$. Based on the conditional probability theory, the data-driven VSA problem is to solve:

$$\max_{\boldsymbol{l}_k, \boldsymbol{\mu}_k} f(\boldsymbol{\Phi}|\Delta \boldsymbol{V}) = \max_{\boldsymbol{l}_k, \boldsymbol{\mu}_k} f(\boldsymbol{\Lambda}, \boldsymbol{L}|\Delta \boldsymbol{V}) \tag{6}$$

Corollary 1: The joint posterior $f(\boldsymbol{\Lambda}, \boldsymbol{L}|\Delta \boldsymbol{V})$ is approximated by the CVB algorithm: $\tilde{f}^{[\nu]}(\boldsymbol{\Phi}|\Delta \boldsymbol{V}) = \tilde{f}^{[\nu-1]}(\boldsymbol{\Phi}_{\setminus k}|\boldsymbol{l}_k, \Delta \boldsymbol{V}) \tilde{f}_k^{[\nu]}(\boldsymbol{l}_k|\Delta \boldsymbol{V})$ at the νth iteration ($\nu \in \{1, 2, \ldots, \nu_c\}$), where $\tilde{f}_k^{[\nu]}(\boldsymbol{l}_k|\Delta \boldsymbol{V})$ is expressed in (7), $\tilde{f}_k^{[\nu]}(\boldsymbol{\mu}_k|\Delta \boldsymbol{V})$ is expressed in (8), and $\tilde{f}^{[\nu]}(\boldsymbol{l}_k|\boldsymbol{\Phi}_{\setminus k}, \Delta \boldsymbol{V}) = \tilde{f}^{[\nu]}(\boldsymbol{\Phi}|\Delta \boldsymbol{V}) / \tilde{f}_{\setminus k}^{[\nu]}(\boldsymbol{\Phi}_{\setminus k}|\Delta \boldsymbol{V})$.

$$\begin{aligned} \tilde{f}_k^{[\nu]}(\boldsymbol{l}_k|\Delta \boldsymbol{V}) &= \frac{f(\boldsymbol{l}_k|\Delta \boldsymbol{V})}{KL(\tilde{f}^{[\nu-1]}(\boldsymbol{\Phi}_{\setminus k}|\boldsymbol{l}_k, \Delta \boldsymbol{V}) || f(\boldsymbol{\Phi}_{\setminus k}|\boldsymbol{l}_k, \Delta \boldsymbol{V}))} \\ &= \exp \mathbb{E}_{\tilde{f}^{[\nu-1]}(\boldsymbol{\Phi}_{\setminus k}|\boldsymbol{l}_k, \Delta \boldsymbol{V})} \log \frac{f(\Delta \boldsymbol{V}, \boldsymbol{\Phi})}{\tilde{f}^{[\nu-1]}(\boldsymbol{\Phi}_{\setminus k}|\boldsymbol{l}_k, \Delta \boldsymbol{V})} \end{aligned} \tag{7}$$

$$\begin{aligned} \tilde{f}_k^{[\nu]}(\boldsymbol{\mu}_k|\Delta \boldsymbol{V}) &= \frac{f(\boldsymbol{\mu}_k|\Delta \boldsymbol{V})}{KL(\tilde{f}^{[\nu-1]}(\boldsymbol{\Phi}_{\setminus k}|\boldsymbol{\mu}_k, \Delta \boldsymbol{V}) || f(\boldsymbol{\Phi}_{\setminus k}|\boldsymbol{\mu}_k, \Delta \boldsymbol{V}))} \\ &= \exp \mathbb{E}_{\tilde{f}^{[\nu-1]}(\boldsymbol{\Phi}_{\setminus k}|\boldsymbol{\mu}_k, \Delta \boldsymbol{V})} \log \frac{f(\Delta \boldsymbol{V}, \boldsymbol{\Phi})}{\tilde{f}^{[\nu-1]}(\boldsymbol{\Phi}_{\setminus k}|\boldsymbol{\mu}_k, \Delta \boldsymbol{V})} \end{aligned} \tag{8}$$

Based on the probability theory, the true posterior $f(\boldsymbol{\Phi}|\Delta \boldsymbol{V}) = f(\boldsymbol{\Lambda}, \boldsymbol{L}|\Delta \boldsymbol{V})$ can be constituted by deriving CVB approximation $\tilde{f}^{[1]} = \tilde{f}_{2|1}^{[0]} \tilde{f}_1^{[1]} = \tilde{f}_{1|2}^{[1]} \tilde{f}_2^{[1]}$, where conditional

functions of f_1, f_2, $f_{1|2}$, and $f_{2|1}$ are given by:

$$\widetilde{f}(\boldsymbol{\Phi}|\Delta\boldsymbol{V}) = \underbrace{\widetilde{f}(\boldsymbol{\Lambda}|\boldsymbol{l}_s,\Delta\boldsymbol{V})}_{\widetilde{f}_{2|1}}\underbrace{\widetilde{f}(\boldsymbol{L}|\Delta\boldsymbol{V})}_{\widetilde{f}_1} = \underbrace{\widetilde{f}(\boldsymbol{L}_{\setminus s}|\boldsymbol{l}_s,\Delta\boldsymbol{V})}_{\widetilde{f}_{1|2}}\underbrace{\widetilde{f}(\boldsymbol{\Lambda},\boldsymbol{l}_s|\Delta\boldsymbol{V})}_{\widetilde{f}_2} \tag{9}$$

where $\widetilde{f}(\boldsymbol{\Phi}|\Delta\boldsymbol{V})$ is the approximated form of $f(\boldsymbol{\Phi}|\Delta\boldsymbol{V})$. The objective of the CVB algorithm is to minimize the Kullback-Leibler (KL) divergence $\mathrm{KL}_{\widetilde{f}||f} = \mathrm{KL}(\widetilde{f}(\boldsymbol{\Phi}|\Delta\boldsymbol{V})||f(\boldsymbol{\Phi}|\Delta\boldsymbol{V}))$ [26].

Remark 1: *This method is indirect but more feasible because it can circumvent the explicit form of $f(\boldsymbol{\Phi}|\Delta\boldsymbol{V})$. It would yield good approximation if $KL_{\widetilde{f}||f}$ could be set low enough.*

Considering the tractability, the initial CVB $\widetilde{f}^{[0]}_{2|1} = \widetilde{f}^{[0]}(\boldsymbol{\Lambda}|\boldsymbol{l}_j,\Delta\boldsymbol{V})$ as a restricted form of the true conditional $f_{2|1} = f(\boldsymbol{\Lambda}|\boldsymbol{L},\Delta\boldsymbol{V})$ is given by:

$$\widetilde{f}^{[0]}_{2|1} = \prod_{k=1}^{4}\widetilde{f}^{[0]}(\boldsymbol{\mu}_k|\boldsymbol{l}_s) = \prod_{k=1}^{4}\prod_{m=1}^{4}\mathcal{N}^{l_{m,s}}_{\boldsymbol{\mu}_k}(\widetilde{\boldsymbol{\mu}}^{[0]}_{k,m,s},\widetilde{\sigma}^{[0]}_{k,m,s}\mathbf{I}_2) \tag{10}$$

where $\mathcal{N}$ is the Gaussian distribution. $\mathbf{I}_2$ is the 2×2 identity covariance matrix. $\widetilde{\boldsymbol{\mu}}_{k,m,s}\in\mathbb{R}^2$ and $\widetilde{\sigma}^{[0]}_{k,m,s}$ are initial means and variances of $\widetilde{f}^{[0]}(\boldsymbol{\mu}_k|\boldsymbol{l}_s) = \prod_{m=1}^{K}\widetilde{f}^{[0]}(\boldsymbol{\mu}_k|l_{m,s})$.

By applying the CVB algorithm, $f_1 = f(\boldsymbol{L}|\Delta\boldsymbol{V})$ is approximated via $\widetilde{f}^{[0]}_{2|1}$ in (10), given by:

$$\widetilde{f}^{[1]}_1 = \widetilde{f}^{[1]}(\boldsymbol{L}|\Delta\boldsymbol{V}) = \exp\mathbb{E}_{\widetilde{f}^{[0]}_{2|1}}\log\frac{f(\Delta\boldsymbol{V},\boldsymbol{\Lambda},\boldsymbol{L})}{\widetilde{f}^{[0]}_{2|1}} = \prod_{k=1}^{4}\prod_{m=1}^{4}\left(\widetilde{\kappa}^{[1]}_{k,m,s}\prod_{s=1}^{Ns}\left(\widetilde{\gamma}^{[1]}_{k,m,i,s}\right)^{l_{k,i}}\right)^{l_{m,s}} \tag{11}$$

Comparing with (11), the form of $\widetilde{f}^{[1]}_{1|2}$ is identified as:

$$\widetilde{f}^{[1]}_{1|2} = \widetilde{f}^{[1]}\left(\boldsymbol{L}_{\setminus s}\mid\boldsymbol{l}_s,\Delta\boldsymbol{V}\right) = \prod_{i\neq s}Mu_{\boldsymbol{l}_i}(\widetilde{\boldsymbol{W}}^{[1]}_{i,s}\boldsymbol{l}_s) \tag{12}$$

where Mu denotes a multinomial distribution. $\widetilde{\boldsymbol{W}}^{[1]}_{i,s}$ is a left stochastic matrix.

By applying the CVB algorithm, $f_2 = f(\boldsymbol{\Lambda},\boldsymbol{l}_j\mid\boldsymbol{X})$ is approximated by $\widetilde{f}^{[2]}_2$ in $\widetilde{f}^{[2]} = \widetilde{f}^{[2]}_{2|1}\widetilde{f}^{[2]}_1 = \widetilde{f}^{[1]}_{1|2}\widetilde{f}^{[2]}_2$ via $\widetilde{f}^{[1]}_{1|2}$:

$$\widetilde{f}^{[2]}_2 = \widetilde{f}^{[2]}(\boldsymbol{\Lambda},\boldsymbol{l}_s|\Delta\boldsymbol{V}) = \exp\mathbb{E}_{\widetilde{f}^{[1]}_{1|2}}\log\frac{f(\Delta\boldsymbol{V},\boldsymbol{\Lambda},\boldsymbol{L})}{\widetilde{f}^{[1]}_{1|2}} = \widetilde{f}^{[2]}(\boldsymbol{\Lambda}|\boldsymbol{l}_j,\Delta\boldsymbol{V})\,\widetilde{f}^{[2]}(\boldsymbol{l}_j|\Delta\boldsymbol{V}) \tag{13}$$

$$\widetilde{f}^{[2]}_{2|1} = \widetilde{f}^{[2]}(\boldsymbol{\Lambda}|\boldsymbol{l}_j,\Delta\boldsymbol{V}) = \prod_{k=1}^{4}\prod_{m=1}^{4}\mathcal{N}^{l_{m,s}}_{\boldsymbol{\mu}_k}(\widetilde{\boldsymbol{\mu}}^{[1]}_{k,m,s},\widetilde{\sigma}^{[1]}_{k,m,s}\mathbf{I}_2) \tag{14}$$

Given that only one CVB marginal is updated per iteration ν, the approximated posterior estimates for voltage stability statuses $\boldsymbol{L}$ and $\boldsymbol{\Lambda}$ are expressed as:

$$\widehat{\boldsymbol{\Lambda}}(j) = \mathbb{E}_{\widetilde{f}^{[\nu_c]}_j(\boldsymbol{\Lambda}|\Delta\boldsymbol{V})}(\boldsymbol{\Lambda}) = \sum_{m=1}^{K}\widetilde{p}^{[\nu_c]}_{m,j}\widetilde{\boldsymbol{\Lambda}}^{[\nu_c]}_{m,j} \tag{15}$$

$$\widehat{\boldsymbol{l}}_i(j) = \arg\max_{\boldsymbol{l}_i}\tilde{f}^{[\nu_c]}_j(\boldsymbol{l}_i\mid\Delta\boldsymbol{V}) = \boldsymbol{\epsilon}_{\widehat{k}_i(j)} \tag{16}$$

where $\widetilde{\boldsymbol{\Lambda}}^{[\nu_c]}_j = [\widetilde{\boldsymbol{\mu}}^{[\nu_c]}_{1,j},\widetilde{\boldsymbol{\mu}}^{[\nu_c]}_{2,j},\ldots,\widetilde{\boldsymbol{\mu}}^{[\nu_c]}_{K,j}]$.

Remark 2 : CVB still belongs to a conditional structure at convergence though its initialization $\{\widetilde{\boldsymbol{\mu}}^{[0]}_{k,m,s},\widetilde{\sigma}^{[0]}_{k,m,s}\}$ is similar with that of conventional VB. Also, even if CVB is set $\widetilde{\boldsymbol{\mu}}^{[0]}_{k,m,s} = \widetilde{\boldsymbol{\mu}}^{[0]}_k$, $\tilde{\sigma}^{[0]}_{k,m,s} = \widetilde{\sigma}^{[0]}_k$, and $\tilde{f}^{[0]}_{2|1} = \tilde{f}^{[0]}(\boldsymbol{\Lambda}|\boldsymbol{l}_s,\Delta\boldsymbol{V})$, the conditional $\tilde{f}^{[\nu]}(\boldsymbol{\Lambda}|\boldsymbol{l}_s,\Delta\boldsymbol{V})$ is related to voltage stability status $\boldsymbol{l}_s$ during the iterations.

To set the stopping rule of iterations, the evidence lower bound (ELBO) is defined as:

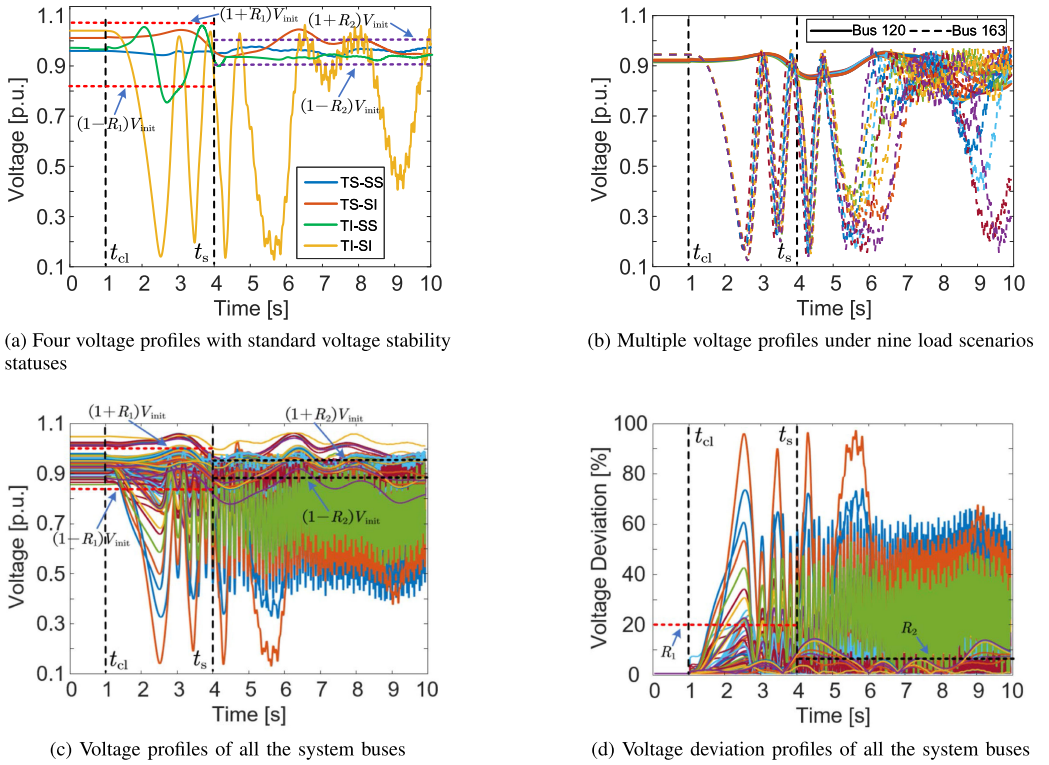
$$ELBO^{[\nu]} = -KL_{\tilde{f}^{[\nu]}(\boldsymbol{\Phi}|\Delta\boldsymbol{V})|f(\boldsymbol{\Phi},\Delta\boldsymbol{V})} + \log f(\Delta\boldsymbol{V}) \tag{17}$$

where $f(\Delta\boldsymbol{V})$ is a constant. $ELBO^{[\nu]}$ monotonically increases to a maximum at convergence $\nu{=}\nu_c$. That is, the iteration terminates as $|ELBO^{[\nu]} - ELBO^{[\nu_c]}| \le \varepsilon$, where ε is the ELBO threshold. The pseudocode of the developed CVB-based VSA methodology is given in Algorithm IV. Detailed information about the CVB inference can be found in [25].

Algorithm 1: The Developed CVB-based JVSA.

1. Set N-1 contingencies with uncertain load and prepare datasets of voltage deviation vector $\Delta\boldsymbol{V}$, voltage stability status matrix $\boldsymbol{L}$ and their probability $\boldsymbol{p}$, unknown stability status matrix $\boldsymbol{\Lambda}$.
2. **Initialization**: set the initial CVB $\widetilde{f}^{[0]}_{2|1}$ in (10), $\tilde{f}^{[1]}_1$ in (11), $\tilde{f}^{[1]}_{1|2}$ in (12), $\tilde{f}^{[2]}_2$ in (13), and $\tilde{f}^{[2]}_{2|1}$ in (14).
3. **for** Iteration ν from 1 to N_ν **do**
4. Update probability $\tilde{f}^{[\nu]}_k(\boldsymbol{l}_k|\Delta\boldsymbol{V})$ in (7) and $\tilde{f}^{[\nu]}_k(\boldsymbol{\mu}_k|\Delta\boldsymbol{V})$ in (8).
5. Calculate stopping metric $ELBO^{[\nu]}$ in (17);
6. **if** $|ELBO^{[\nu]} - ELBO^{[\nu-1]}| \le \varepsilon$ **then**
7. Terminate the iteration at step ν and set $\nu_c \leftarrow \nu$.
8. Obtain approximated posterior estimates $\widehat{\boldsymbol{\Lambda}}(j)$ (15) and $\widehat{\boldsymbol{l}}_i(j)$ (16):
9. $\widehat{\boldsymbol{\Lambda}}(j) = \mathbb{E}_{\tilde{f}^{[\nu_c]}_j(\boldsymbol{\Lambda}|\Delta\boldsymbol{V})}(\boldsymbol{\Lambda})$;
10. $\widehat{\boldsymbol{l}}_i(j) = \arg\max_{\boldsymbol{l}_i}\tilde{f}^{[\nu_c]}_j(\boldsymbol{l}_i\mid\Delta\boldsymbol{V})$.
11. **end**
12. **end**

To quantitatively evaluate the performance of different methods, three metrics including MSE ($MSE = \min\|\widehat{\boldsymbol{\Lambda}} - \boldsymbol{\Lambda}\|^2$), Purity ($Purity = \sum_{k=1}^{4}\frac{1}{Ns}\max\sum_{s=1}^{Ns}\delta(l_{k,s} = l_{m,s})$), and ELBO are used for comparison. The higher the Purity metric value, the higher the percentage of correct voltage stability statuses of samples. The Purity metric is a common measurement for percentage of successful VSA. The higher the ELBO metric value, the better the approximation of the correct voltage stability



(a) Four voltage profiles with standard voltage stability statuses

(b) Multiple voltage profiles under nine load scenarios

(c) Voltage profiles of all the system buses

(d) Voltage deviation profiles of all the system buses

Fig. 4. Voltage and voltage deviation profiles. (a) four classic voltage curves corresponding to four standard voltage stability statuses: TS-SS, TS-SI, TI-SS, and TI-SI and (b) multiple voltage profiles under uncertain load scenarios with N-1 contingency.

statuses of samples. The lower the MSE metric value, the better the estimate for VSA.

Furthermore, we have provided five metrics for better illustration, including Correlation Coefficient, Chi-Square goodness-of-fit, Normalized Root Mean Square Error (NRMSE), Maximum Absolute Error (MaxAE), and Mean Absolute Error (MAE). A lower value indicates a better estimate for most of the metrics, except for Correlation Coefficient. Detailed information about these metrics can be found in [27].

V. Case Studies and Results

A Python-based power system simulation software ANDES is used to model DAEs for the dynamic simulation [28]. The test case is performed on the WECC 179-bus 29-machine 263-branch system to retain the main dynamics of the entire system [29], [30]. Numerical simulations of the CNN and CVB are performed in the MATLAB R2019b environment. The ELBO threshold ε is set as 0.01. The maximum number of iterations N_ν is set as 200. The mean squared error (MSE) metric is used for comparison: $MSE = \min \|\widehat{\mathbf{\Lambda}} - \mathbf{\Lambda}\|^2$. The successful JVSA results are calculated using the Purity metric: $Purity = \sum_{k=1}^{4} \frac{1}{N_{\text{S}}} \max \sum_{s=1}^{N_{\text{S}}} \delta(l_{k,s} = l_{m,s})$. Typical parameters R_1 and R_2 for assessing voltage stability are set as R_1=20% and R_2=5%, which can be referred in [23]. Four representative system bus voltages are intentionally chosen to map the four voltage stability statuses, i.e., TS-SS, TS-SI, TI-SS, and TI-SI. The tripping line fault is set at 1 s. The fault clearing time t_{cl} is set at 1.1 s and the post-transient time t_{s} is set at 4 s. For both N-1 and N-2 contingencies, each transmission line is disconnected at 1 s and reconnected at 1.1 s. The maximum simulation time is set to 10 seconds. The sampling rate of measurement data is set to 30 samples per second. For one N-1 contingency, load uncertainty is considered by generating 1,000 load scenarios. That is, the real and reactive power of the load follows a normal distribution with a mean value and a standard deviation as one hundredth of the corresponding absolute value [31]. The total time of generating the dataset of one load scenario is approximately in the range of 90–120 seconds. The total time of generating the dataset of an N-1 contingency event with 1,000 load scenarios is approximately in the range of 25–33 hours. The training time of one CNN model for one N-1 scenario is approximately in the range of 30–40 seconds.

Fig. 4 shows some representative profiles of voltage and voltage deviation. In Fig. 4(a), four representative system buses are considered: Bus 14, Bus 120, Bus 163, and Bus 172. As can be seen, the solid blue line represents the voltage at Bus 14. Its status is both transient and steady-state stable, i.e., TS-SS. The solid yellow line represents the voltage at Bus 163. Its status is both transient and steady-state instable, i.e., TI-SI. The solid orange line represents the voltage at Bus 120. Its status is steady-state instable, i.e., TS-SI. The solid green line represents the voltage at Bus 172. Its status is transient instable, i.e., TI-SS. The STVS criteria $(1 + R_1)V_{\text{init}}$ is shown via the red dashed line between t_{cl} and t_{s}. The LTVS criteria $(1 + R_2)V_{\text{init}}$ is shown in the purple dashed line after t_{s}. Nine voltage profiles at Bus 120 and Bus 163 are shown in Fig. 4(b) under nine load scenarios. As can be seen, the voltage at Bus 120 (solid lines) is transient stable but steady-state instable, i.e., TS-SI, for nine scenarios. The voltage

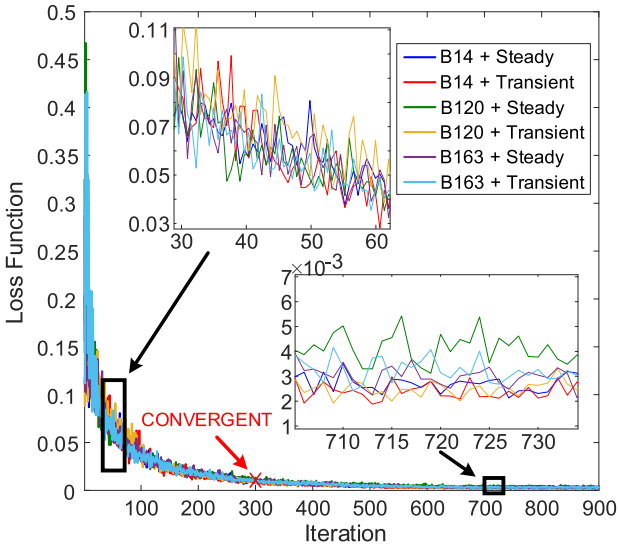


Fig. 5. Loss function values of multiple CNNs during iterations for estimating transient and steady-state voltage deviations ΔV_{T} and ΔV_{S} at three representative buses (Buses 14, 120, and 163).

TABLE I
REPRESENTATIVE PARAMETERS FOR TRAINING CNN MODEL

Maximum Number of Epochs	30
Initial Learning Rate	10^{-3}
Factor for Dropping the Learning Rate	0.1
Number of Epochs for Dropping the Learning Rate	20
Size of Mini-Batch	64
Factor for L_2 Regularization (Weight Decay)	10^{-4}

at Bus 163 is both transient and steady-state instable, i.e., TI-SI, for nine scenarios. Voltage profiles of all the system buses (179 buses) are illustrated in Fig. 4(c) under one N-1 contingency (that is, tripping Line 91–92 at 1 s). Voltage deviation profiles of all the system buses (179 buses) are illustrated in Fig. 4(d) under the same contingency. Comparing voltage profiles in Fig. 4(c) with voltage deviation profiles in Fig. 4(d), it is shown that voltage deviation is simpler for assessing the voltage stability status.

A. Validation of Voltage Deviation Estimation Performance

Taking the estimation of transient and steady-state voltage deviation at representative Buses 14, 120, and 163 as an example, Fig. 5 illustrates the loss function values of multi-CNNs during iterations. The maximum number of iterations is set as 900. As can be seen, the loss function values are significantly reduced at the beginning of the iterations with an exponential fall. The iterations are convergent at about 300 iterations with relatively small loss function values, which are close to zero. It is shown that all the multi-CNNs for voltage deviation estimates have relatively satisfactory convergence. Detailed information of architectures and representative parameters for training the CNN model is shown in Table I. Other default parameter values can be found in [32]. A representative structure for training the CNN model is shown in Fig. 6. As can be seen, there are one input layer, four convolution layers, four batch normalization layers, two average pooling layer, four ReLu layers, one drop

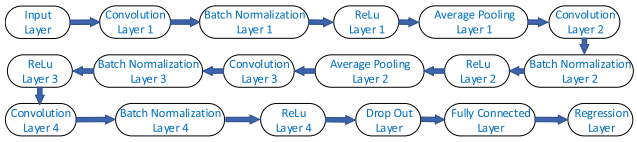


Fig. 6. Representative structure for training the CNN model.

TABLE II
RESULTS OF VOLTAGE DEVIATION ESTIMATION USING MULTI-CNNS FOR BOTH TRANSIENT AND STEADY-STATE PROCESS

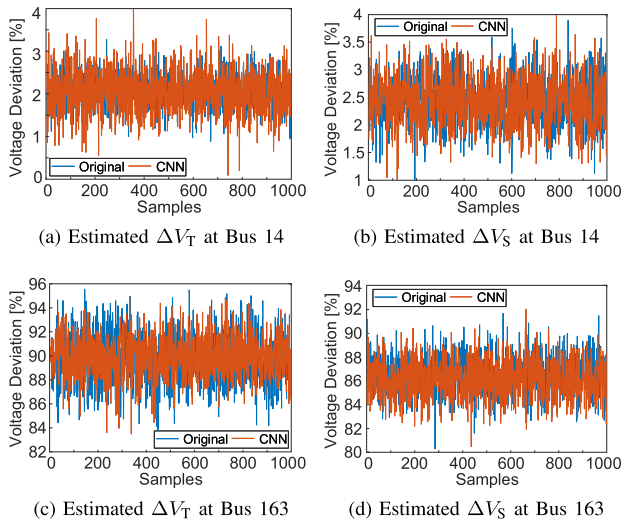
Bus Index	ΔV_{T} [%]	ΔV_{S} [%]	Bus Index	ΔV_{T} [%]	ΔV_{S} [%]
14	99.7	99.2	25	99.6	99.4
32	99.5	99.7	46	98.9	99.3
55	99.5	99.8	63	99.1	99.3
74	99.3	99.5	88	99.5	99.6
99	99.8	99.4	106	99.5	98.9
110	99.1	99.2	120	99.6	99.4
140	99.5	99.7	150	99.4	99.2
163	98.8	99.2	172	99.5	99.1

out layer, one fully connected layer, and one regression layer. To benchmark the performance of learners in our developed method, its standard architecture of CNN can be readily found for comparison from the existing examples, which have already been developed and embedded in MATLAB [33], [34]. A main branch with multiple convolutional, batch normalization, and ReLU layers are connected sequentially. Residual connections bypass the convolutional units of the main branch. Residual connections enable the parameter gradients to flow more easily from the output layer to the earlier layers of the network, which makes it possible to train deeper CNN.

To quantitatively validate the effectiveness of the developed multi-CNN model, Table II illustrates the voltage deviation estimation accuracy for both ΔV_{T} and ΔV_{S} at multiple system buses. There are sixteen system buses randomly chosen to calculate the estimation accuracy of multi-CNNs. As can be seen, the estimation accuracy is within the range of 98.8%–99.8%. This observation validates the effectiveness of the multi-CNN model we have developed. Thus, we utilize this model to fast estimate dynamic voltage deviation for the WECC 179-bus system.

Four examples of voltage deviation estimation results are illustrated in Fig. 7 using the developed multi-CNNs at Bus 14 and Bus 163. As can be seen, the estimated ΔV_{T} at Bus 14 is within the range of 0–4%, which is transient stable. The estimated ΔV_{S} at Bus 14 is within the range of 1–4%, which is steady-state stable. Thus, the voltage stability status at Bus 14 is TS-SS. Similarly, the estimated ΔV_{T} at Bus 163 is within the range of 82–96%, which is transient instable. The estimated ΔV_{S} at Bus 163 is within the range of 80–94%, which is steady-state instable. Thus, the voltage stability status at Bus 163 is TI-SI.

The difference between the proposed model and the traditional neural network model is that the former utilizes convolutional layers to extract features, while the latter utilizes the fully connected layers. Furthermore, CNN has multiple hidden layers for sufficient feature extraction, while in the traditional neural network there is only a simple hidden layer between the input and the output. The proposed method provides more accurate results than the traditional method. This is because the convolutional



(a) Estimated ΔV_{T} at Bus 14 (b) Estimated ΔV_{S} at Bus 14

(c) Estimated ΔV_{T} at Bus 163 (d) Estimated ΔV_{S} at Bus 163

Fig. 7. Estimated voltage deviation results of transient ΔV_{T} and steady-state ΔV_{S} using the developed multi-CNNs at Bus 14 and Bus 163.

kernels within the multi-CNNs deploy connectivity to extract better features for model regression. Also, the convolutional layers can significantly spare both computational and storage sources.

B. Validation of the Proposed JVSA Method

Based on the fast estimated voltage deviation using multi-CNNs, 100 samples are randomly picked for VSA using the developed CVB inference. This process is repeated using the Monte Carlo simulation 100 times. Thus, the total number of testing samples is 10,000 ($=100 \times 100$).

For one Monte Carlo simulation, Fig. 8 shows the convergent results of VSA based on estimated transient ΔV_{T} and steady-state ΔV_{S} using the developed CVB inference. Subfigures in this figure show the results of TS-SS, TS-SI, TI-SS, and TI-SI, respectively. The dashed green circles denote the contours of true normal distributions. The green plus signs denote the original samples that are randomly picked from the voltage deviation estimates using multi-CNNs. The black rectangles denote the mean vectors of one voltage stability status estimated by the k-means algorithm. The blue crosses denote the mean vectors of one voltage stability status estimated by the Expectation-Maximization-1 (EM_1) algorithm. The blue triangles denote the mean vectors of one voltage stability status estimated by the Expectation-Maximization-2 (EM_2) algorithm. The red asterisks denote the mean vectors of one voltage stability status estimated by the developed CVB algorithm. The methodologies of k-means, EM_1, EM_2, and VB algorithms are briefly introduced in Appendix A and Appendix B, respectively.

As can be seen in Fig. 8, both k-means and EM_1 algorithms provide very low VSA accuracy compared with other assessment algorithms. This is because they use the same point estimates for voltage stability statuses. Also, both EM_2 and VB algorithms provide low VSA accuracy compared with the developed CVB algorithm. This is because they use the models' probability of each voltage stability status as weighted credibility and make soft decisions during each iteration. However, the developed CVB algorithm can estimate the most accurate JVSA results, which are significantly close to the mean vector center of each voltage stability status. This is because the CVB algorithm takes all moments of transient and steady-state voltage deviations into account, while k-means only considers the first moment and EM only considers the first and the second moments of voltage deviation samples.

TABLE III
METRICS COMPARISON OF SUCCESSFUL VSA USING k-MEANS, EM_1, EM_2, VB, AND CVB

Metrics	k-means	EM_1	EM_2	VB	CVB
MSE [%]	7.32	6.37	6.87	4.87	3.46
Purity [%]	85.68	87.24	88.72	88.54	92.86
ELBO	-746.98	-751.82	-730.08	-735.07	-697.22
Corre. Coeff.	0.76	0.87	0.84	0.91	0.94
Chi-Square	2.54	1.68	1.98	1.56	1.21
NRMSE	0.18	0.15	0.16	0.15	0.09
MaxAE	0.35	0.29	0.31	0.27	0.18
MAE	0.15	0.13	0.09	0.08	0.06

TABLE IV
FNR METRIC OF FOUR STABILITY STATUSES BY USING THE PROPOSED METHOD

Metric	FNR [%]
TS-SS (real) *V.S.* TI-SI (predict)	0.35
TI-SS (real) *V.S.* TS-SI (predict)	0.16
TS-SI (real) *V.S.* TI-SS (predict)	0.24
TI-SI (real) *V.S.* TS-SS (predict)	0.15

Table III compares three metrics for successful VSA using four algorithms, i.e., k-means, EM_1, EM_2, VB, and CVB. As can be seen, the developed CVB algorithm obtains the lowest MSE, the highest Purity, and the highest ELBO. It is observed that the developed CVB is the best posterior approximation for the VSA model compared with k-means, EM_1, EM_2, and VB algorithms. Specifically, compared with k-means, EM_1, EM_2, and VB algorithms, CVB reduces the MSE metric by 52.73%, 45.68%, 49.64%, and 28.95%, respectively. It increases the Purity metric by 8.38%, 6.44%, 4.67%, and 4.88%, respectively, and it increases the ELBO metric by 6.66%, 7.26%, 4.5%, and 5.15%, respectively.

A false negative rate (FNR) metric is calculated when the assessment results are totally different from the real stability statuses. Specifically, the real status of one sample is TS-SS while the assessment method predicts a TI-SI status; the real status is TI-SS while a TS-SI status is predicted; the real status is TS-SI while a TI-SS status is predicted; and the real status is TI-SI while a TS-SS is predicted. FNR is a rate of calculated samples accounting for the total number. Table IV illustrates the FNR metric of four stability statuses by using the proposed method. As can be seen, the proposed method can provide relatively low FNR values for four stability statuses. However, it cannot guarantee the 100% assessment accuracy.

C. Effectiveness Analysis of N-2 Contingencies

In this case, we simulate the N-2 contingency by disconnecting Line 2 and Line 61 at 1 s and reconnecting them at 1.1 s, which is taken as a representative example for better illustration. For an N-2 contingency, load uncertainty is considered by generating 1,000 load scenarios. The real and reactive power

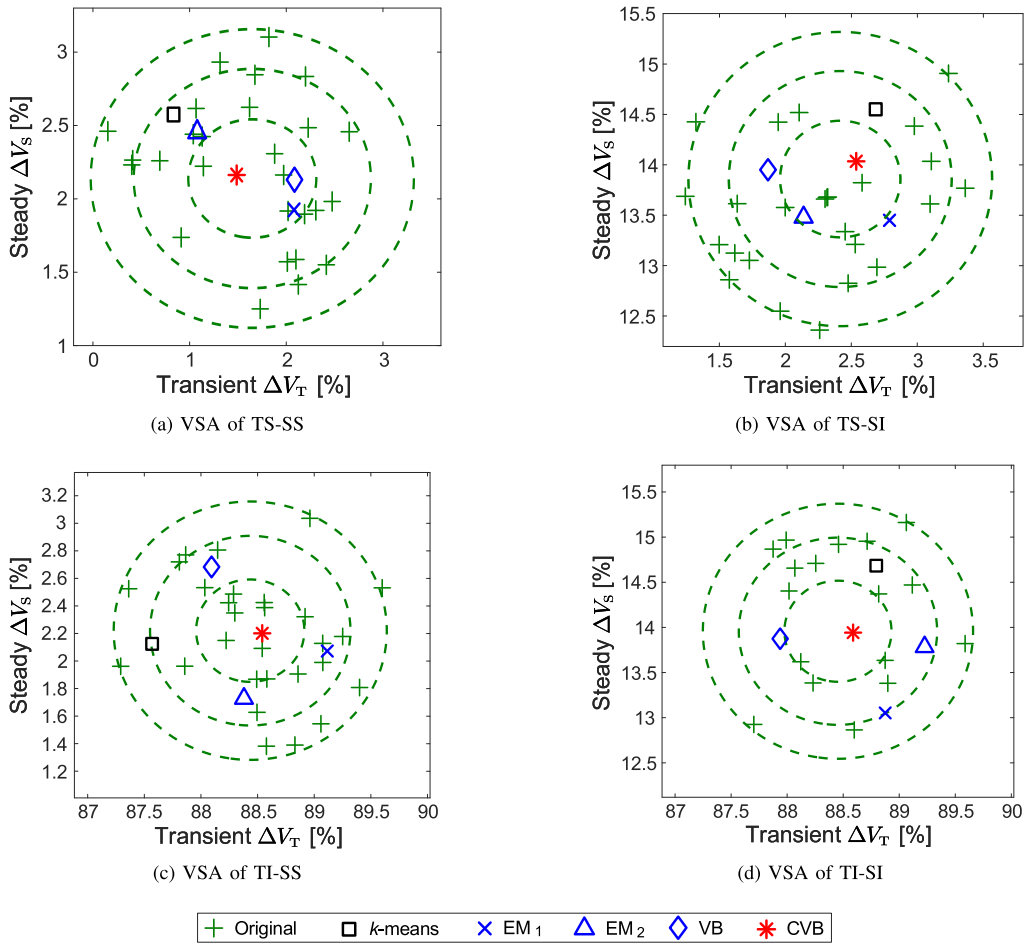


Fig. 8. VSA results based on estimated transient ΔV_T and steady-state ΔV_S using the developed CVB inference for N-1 contingency.

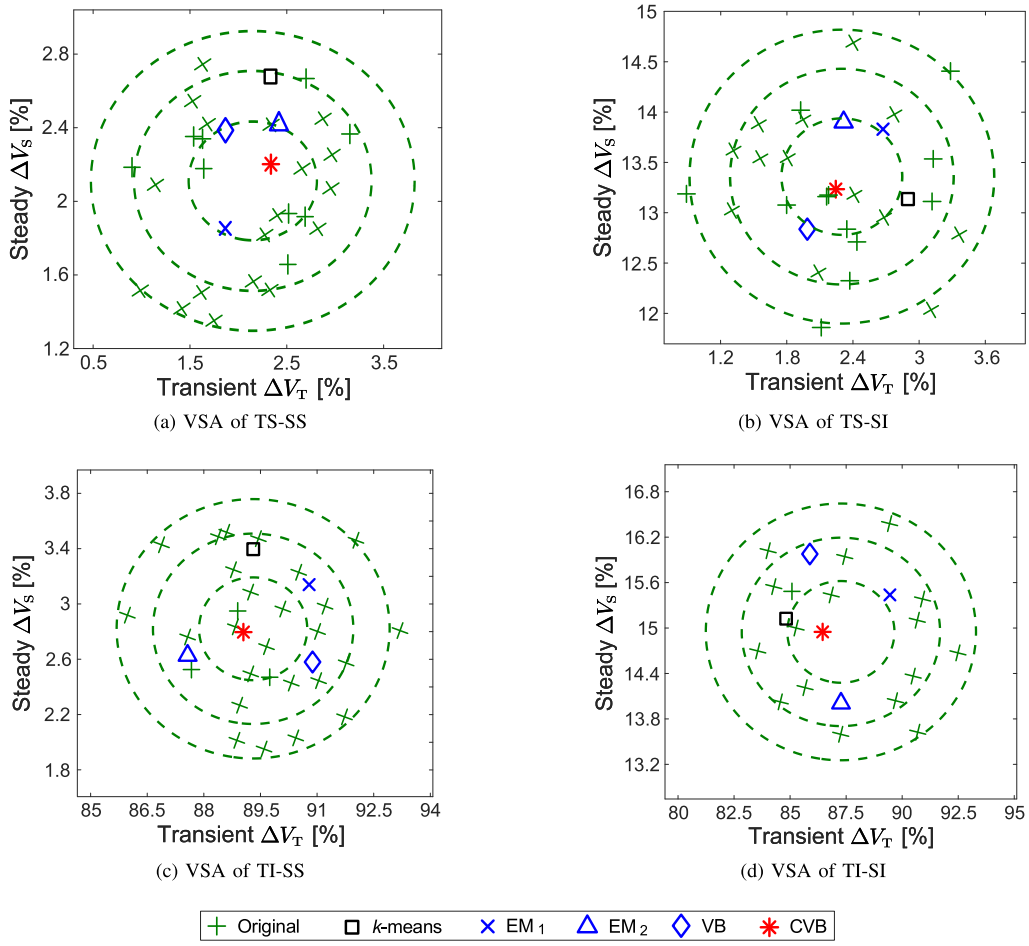
of the load follows a normal distribution with a mean value and a standard deviation as one hundredth of the corresponding absolute value. There are 100 samples randomly selected for VSA using the developed CVB inference. This process is repeated using the Monte Carlo simulation in 100 runs. The total number of testing samples is 10,000 ($=100 \times 100$). Taking a Monte Carlo simulation study as an example, Fig. 9 shows the converged results of VSA based on estimated transient ΔV_T and steady-state ΔV_S using the developed CVB inference.

As can be seen in Fig. 9, red asterisks (CVB) are significantly close to the mean vector center of each voltage stability status. It demonstrates that the developed CVB algorithm can estimate the most accurate JVSA results for N-2 contingencies. This observation is similar to the case under N-1 contingency. This is because the developed CVB algorithm considers all moments of transient and steady-state voltage deviations instead of merely taking the first and the second moments of voltage deviation samples into account.

D. Analysis of Computational Time

The most important finding is that the developed multi-CNN model is good at fast estimating voltage deviation through dynamic power systems, especially for a large system with hundreds of buses. Thus, we utilize this model to fast estimate the voltage stability status of the system. Actually, it is significantly time-consuming when the conventional model-based dynamic power system simulation is deployed. Specifically, assuming that we have one N-1 contingency and 100 load scenarios, the simulation time for obtaining dynamic voltage curves is $\sim$2 minutes for one scenario on the WECC 179-bus system. Thus, the total simulation time is $\sim$200 ($=2 \times 100$) minutes considering load uncertainty scenarios. This significantly large consuming time makes the conventional model-based simulations unpractical in the real-time applications. However, using the developed multi-CNN model can significantly reduce the simulation time for obtaining dynamic voltage, especially for multiple load uncertainty scenarios. Thus, it is significantly suitable for Monte Carlo simulation based dynamic simulations considering load uncertainty.

Table V compares the computational efficiency of the proposed data-driven JVSA method with the conventional assessment method. Two cases and four voltage stability statuses are used for comparison: one deterministic load scenario and 25 uncertain load scenarios. As can be seen, for one deterministic load scenario, the conventional VSA method consumes 102–114 seconds, most of which is taken by running the DAE-based power system simulation model. Similarly, for the case with 25 load scenarios, the computational time of the conventional VSA method is 2500–2800 seconds, which is significantly large



(a) VSA of TS-SS (b) VSA of TS-SI

(c) VSA of TI-SS (d) VSA of TI-SI

Fig. 9. VSA results based on estimated transient ΔV_T and steady-state ΔV_S using the developed CVB inference for N-2 contingency.

TABLE V
COMPUTATIONAL TIME COMPARISON BETWEEN THE PROPOSED DATA-DRIVEN JVSA METHOD AND THE CONVENTIONAL VSA METHOD

Cases	Stability Status	Conventional VSA method [s]	Proposed JVSA method [s]	Speedup
one load scenario	TSS	102	2.21	46
	TSSU	110	3.12	35
	TUSS	103	1.89	54
	TSU	114	2.04	56
25 load scenarios	TSS	2743	47.82	57
	TSSU	2534	43.25	59
	TUSS	2683	46.22	58
	TSU	2568	43.85	59

and unpractical for real-time VSA applications. However, the proposed data-driven JVSA method can significantly reduce the computational time. For the deterministic load scenario, the proposed JVSA method only takes 1.89–3.12 seconds for assessing four voltage stability statuses. For the case of 25 load scenarios, it only takes 43–48 seconds. This observation validates the computational effectiveness of the proposed data-driven fast VSA method.

The computational speed of the proposed method is significantly faster than the conventional method. This is because once the multi-CNN model is well-trained, it has formulated high dimensional mapping between input and output, and it can directly estimate voltage deviations for new instances with different uncertain loading conditions and system topology changes, without incurring the iterative calculation. This computation-free feature makes the multi-CNN model an advantageous tool for solving highly complex large-scale power system planning and operation problems, where the model-based conventional method can be excessively time- and resource-consuming. Also, assessment speed can be increased by designing a hardware-based architecture for energy-efficient CNN acceleration.

VI. CONCLUSION

In this paper, we developed a novel data-driven JVSA method by integrating the multi-CNN model and the VB inference. First, to fast estimate the maximum voltage deviations during the transient and steady-state process, a multi-CNN model is designed based on uncertain load scenarios and system topology under N-1 contingency. Second, voltage deviation estimates are utilized to fast assess the joint VSA for both transient and steady-state processes using the novel CVB inference. The effectiveness of the proposed data-driven JVSA method is validated by numerical simulations and comparisons on a modified WECC 179-bus system. It is demonstrated the proposed method can not only provide better VSA accuracy but also consume significantly less computational time.

In future work, N-3, $\cdots$, N-k contingencies with multiple faults and new topologies will be further studied by using the developed JVSA method to significantly improve the corresponding assessment accuracy and computational efficiency. Also, a theoretical method to determine the optimal number of convolutional layers and ReLu layers will be studied.

APPENDIX A
k-MEANS ALGORITHM

The objective model of k-means algorithm is expressed as:

$$\widehat{\boldsymbol{\Lambda}}^{[\nu]} = \arg\max_{\boldsymbol{\Lambda}} f(\boldsymbol{\Lambda} \mid \widehat{\boldsymbol{L}}^{[\nu-1]}, \Delta \boldsymbol{V}) \tag{18}$$

$$\widehat{\boldsymbol{L}}^{[\nu]} = \arg\max_{\boldsymbol{L}} f(\boldsymbol{L} \mid \widehat{\boldsymbol{\mu}}^{[\nu]}, \Delta \boldsymbol{V}) \tag{19}$$

The solution of k-means is given by:

$$\widehat{\boldsymbol{\mu}}_k^{[\nu]} = \overline{\boldsymbol{\mu}}_k(\widehat{\boldsymbol{L}}^{[\nu-1]}) = \sum_{i=1}^{N} \widehat{l}_{k,i}^{[\nu-1]} \Delta \boldsymbol{V}_i \Big/ \sum_{i=1}^{N} \widehat{l}_{k,i}^{[\nu-1]} \tag{20}$$

$$\widehat{k}_i^{[\nu]} = \arg\max_k \mathcal{N}_{\Delta \boldsymbol{V}_i}(\overline{\boldsymbol{\mu}}_k^{[\nu]}, \mathbf{I}_2) = \arg\min_k \|\Delta \boldsymbol{V}_i - \overline{\boldsymbol{\mu}}_k^{[\nu]}\|^2 \tag{21}$$

where k denotes the kth non-zero element in one boolean vector $\boldsymbol{\epsilon}_k = [0, \ldots, 1, \ldots, 0]^T \in \mathbb{I}^4$ and the voltage stability status label is $\boldsymbol{l} = \boldsymbol{\epsilon}_k$.

APPENDIX B
EM_1 AND EM_2 ALGORITHMS

The true posterior distribution of EM_1 is expressed as:

$$\widetilde{f}_{\mathrm{EM}_1}(\boldsymbol{\Lambda}, \boldsymbol{L} \mid \Delta \boldsymbol{V}) = f(\boldsymbol{\Lambda} \mid \widehat{\boldsymbol{L}}_{\mathrm{EM}_1}, \Delta \boldsymbol{V}) \delta[\boldsymbol{L} - \widehat{\boldsymbol{L}}_{\mathrm{EM}_1}] \tag{22}$$

where $\delta[\cdot]$ denotes the Kronecker delta function.

The objective model of EM_1 algorithm is expressed as:

$$\widehat{\boldsymbol{L}}_{\mathrm{EM}_1}^{[\nu]} = \arg\max_{\boldsymbol{L}} \mathbb{E}_{f(\boldsymbol{\Lambda} \mid \widehat{\boldsymbol{L}}_{\mathrm{EM}_1}^{[\nu-1]}, \Delta \boldsymbol{V})} \log f(\Delta \boldsymbol{V}, \boldsymbol{\Lambda}, \boldsymbol{L}) \tag{23}$$

The true posterior distribution of EM_2 is expressed as:

$$\widetilde{f}_{\mathrm{EM}_2}(\boldsymbol{\Lambda}, \boldsymbol{L} \mid \Delta \boldsymbol{V}) = f(\boldsymbol{L} \mid \widehat{\boldsymbol{\Lambda}}_{\mathrm{EM}_2}, \Delta \boldsymbol{V}) \delta[\boldsymbol{\Lambda} - \widehat{\boldsymbol{\Lambda}}_{\mathrm{EM}_2}] \tag{24}$$

The objective model of EM_2 algorithm is expressed as:

$$\widehat{\boldsymbol{\Lambda}}_{\mathrm{EM}_2}^{[\nu]} = \arg\max_{\boldsymbol{\Lambda}} \mathbb{E}_{f(\boldsymbol{L} \mid \widehat{\boldsymbol{\Lambda}}_{\mathrm{EM}_2}^{[\nu-1]}, \Delta \boldsymbol{V})} \log f(\Delta \boldsymbol{V}, \boldsymbol{\Lambda}, \boldsymbol{L}) \tag{25}$$

The solution of EM algorithms is given by:

$$\widetilde{\boldsymbol{\mu}}_k^{[\nu]} = \overline{\boldsymbol{\mu}}_k(\widehat{\boldsymbol{L}}_{\mathrm{EM}}^{[\nu-1]}),\ \widetilde{\sigma}_k^{[\nu]} = \bar{\sigma}_k(\widehat{\boldsymbol{L}}_{\mathrm{EM}}^{[\nu-1]}) \tag{26}$$

$$\widehat{k}_i^{[\nu]} = \arg\max_k \mathcal{N}_{\Delta \boldsymbol{V}_i}(\widehat{\boldsymbol{\mu}}_k^{[\nu]}, \mathbf{I}_2) / \exp((\widetilde{\sigma}_k^{[\nu]})^2) \tag{27}$$

REFERENCES

[1] P. Kundur *et al.*, "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1387–1401, Aug. 2004.

[2] N. Hatziargyriou *et al.*, "Definition and classification of power system stability revisited & extended," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3271–3281, Jul. 2021.

[3] S. M. Abbasi, F. Karbalaei, and A. Badri, "The effect of suitable network modeling in voltage stability assessment," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1650–1652, Mar. 2019.

[4] P. Mitra, V. Vittal, B. Keel, and J. Mistry, "A systematic approach to n-1-1 analysis for power system security assessment," *IEEE Power Energy Technol. Syst. J.*, vol. 3, no. 2, pp. 71–80, Jun. 2016.

[5] T. Han, Y. Chen, J. Ma, Y. Zhao, and Y.-Y. Chi, "Surrogate modeling-based multi-objective dynamic var planning considering short-term voltage stability and transient stability," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 622–633, Jan. 2018.

[6] J. James, D. J. Hill, A. Y. Lam, J. Gu, and V. O. Li, "Intelligent time-adaptive transient stability assessment system," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1049–1058, Jan. 2018.

[7] H. N. V. Pico and B. B. Johnson, "Transient stability assessment of multi-machine multi-converter power systems," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3504–3514, Sep. 2019.

[8] C. Wang, K. Yuan, P. Li, B. Jiao, and G. Song, "A projective integration method for transient stability assessment of power systems with a high penetration of distributed generation," *IEEE Trans. Smart Grid*, vol. 9, no. 1, pp. 386–395, Jan. 2018.

[9] Y. Wang, V. Vittal, M. Abdi-Khorsand, and C. Singh, "Probabilistic reliability evaluation including adequacy and dynamic security assessment," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 551–559, Jan. 2020.

[10] S. Dutta and T. J. Overbye, "Feature extraction and visualization of power system transient stability results," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 966–973, Mar. 2014.

[11] Y. Du, F. F. Li, T. Zheng, and J. Li, "Fast cascading outage screening based on deep convolutional neural network and depth-first search," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 2704–2715, Jul. 2020.

[12] Z. Li, Q. Guo, H. Sun, and J. Wang, "Impact of coupled transmission-distribution on static voltage stability assessment," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3311–3312, Jul. 2017.

[13] W. D. Oliveira, J. P. Vieira, U. H. Bezerra, D. A. Martins, and B. d. G. Rodrigues, "Power system security assessment for multiple contingencies using multiway decision tree," *Elect. Power Syst. Res.*, vol. 148, pp. 264–272, 2017.

[14] J. Luo, L. Shi, and L. Yao, "A multi-objective optimization model for active power steady-state security region analysis incorporating wind power," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2016, pp. 1–5.

[15] J. Zhang, C. Tse, K. Wang, and C. Chung, "Voltage stability analysis considering the uncertainties of dynamic load parameters," *IET Gener. Transm. Distrib.*, vol. 3, no. 10, pp. 941–948, 2009.

[16] G. Pierrou and X. Wang, "The effect of the uncertainty of load and renewable generation on the dynamic voltage stability margin," in *Proc. IEEE PES Innovative Smart Grid Technol. Europe*, 2019, pp. 1–5.

[17] B. S. Adusumilli and B. K. Kumar, "Modified affine arithmetic based continuation power flow analysis for voltage stability assessment under uncertainty," *IET Gener. Transm. Distrib.*, vol. 12, no. 18, pp. 4225–4232, 2018.

[18] Y. Chi, Y. Xu, and R. Zhang, "Candidate bus selection for dynamic VAR planning towards voltage stability enhancement considering copula-based correlation of wind and load uncertainties," *IET Gener. Transm. Distrib.*, vol. 15, no. 4, pp. 780–791, 2021.

[19] C. Ren and Y. Xu, "Transfer learning-based power system online dynamic security assessment: Using one model to assess many unlearned faults," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 821–824, Jan. 2020.

[20] R. Huang, Learning and Fast Adaptation for Grid Emergency Control via Deep Meta Reinforcement Learning. [Online]. Available: https://www.researchgate.net/publication/348363944_Learning_and_Fast_Adaptation_for_Grid_Emergency_Control_via_Deep_Meta_Reinforcement_Learning

[21] L. Zhu, D. J. Hill, and C. Lu, "Hierarchical deep learning machine for power system online transient stability prediction," *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 2399–2411, May 2020.

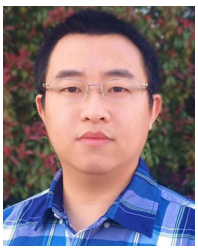
[22] R. Bravo *et al.*, Load Modeling Transmission Research. [Online]. Available: https://web.eecs.umich.edu/ hiskens/publications/LM_Final_Report.pdf

[23] H. Sun *et al.*, "Review of challenges and research opportunities for voltage control in smart grids," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2790–2801, Jul. 2019.

[24] F. Li and Y. Du, "From AlphaGo to power system AI: What engineers can learn from solving the most complex board game," *IEEE Power Energy Mag.*, vol. 16, no. 2, pp. 76–84, Apr. 2018.

[25] V. H. Tran, "Copula variational bayes inference via information geometry," 2018, *arXiv:1803.10998*.

[26] V. Šmídl and A. Quinn, *The Variational Bayes Method in Signal Processing*. Berlin, Germany: Springer Science & Business Media, 2006.
[27] M. Cui, C. Feng, Z. Wang, and J. Zhang, "Statistical representation of wind power ramps using a generalized Gaussian mixture model," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 261–272, Jan. 2018.
[28] *Python Software for Symbolic Power System Modeling and Numerical Analysis (ANDES)*. [Online]. Available: https://andes.readthedocs.io/en/stable/
[29] S. Maslennikov *et al.*, "A test cases library for methods locating the sources of sustained oscillations," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2016, pp. 1–5.
[30] H. Silva-Saravia, H. Pulgar-Painemal, D. A. Schoenwald, and W. Ju, "Adaptive coordination of damping controllers for enhanced power system stability," *IEEE Open Access J. Power Energy*, vol. 7, pp. 265–275, 2020.
[31] M. Cui, J. Wang, Y. Wang, R. Diao, and D. Shi, "Robust time-varying synthesis load modeling in distribution networks considering voltage disturbances," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4438–4450, Nov. 2019.
[32] The Mathworks, Inc. trainingoptions. [Online]. Available: https://www.mathworks.com/help/deeplearning/ref/trainingoptions.html
[33] The Mathworks, Inc, *Train Convolutional Neural Network for Regression*. [Online]. Available: https://www.mathworks.com/help/deeplearning/ug/train-a-convolutional-neural-network-for-regression.html
[34] The Mathworks, Inc, *Train Residual Network for Image Classification*. [Online]. Available: https://www.mathworks.com/help/deeplearning/ug/train-residual-network-for-image-classification.html



Mingjian Cui (Senior Member, IEEE) received the B.S. and Ph.D. degrees from Wuhan University, Wuhan, Hubei, China, all in electrical engineering and automation, in 2010 and 2015, respectively.

He is currently a Research Scientist with The University of Tennessee Knoxville. His research interests include renewable energy, power system operation, power system cybersecurity, power system data analytics, and machine learning. He has authored or coauthored more than 60 peer-reviewed publications. He has been the Editors of journals of IEEE TPWRS since 2021, IET GTD since 2021, and IEEE OAJPE since 2020. He is Best Reviewers of IEEE TSG 2018, IEEE TSTE 2019, and IEEE TSTE 2020.

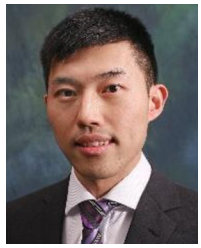
Fangxing Li (Fellow, IEEE) is also known as Fran Li. He received the B.S.E.E. and M.S.E.E. degrees from Southeast University, Nanjing, China, in 1994 and 1997, respectively, and the Ph.D. degree from Virginia Tech, Blacksburg, VA, USA, in 2001. From 2001 to 2005, he was with ABB Electric Systems Consulting (ESC), Raleigh, NC, USA. He is currently the James W. McConnell Professor in electrical engineering and the Campus Director of CURENT with the University of Tennessee, Knoxville, TN, USA.

His current research interests include renewable energy integration, demand response, distributed generation and microgrid, energy markets, and power system computing.

Prof. Li is currently the Editor-In-Chief (EIC) of IEEE OPEN ACCESS JOURNAL OF POWER AND ENERGY (OAJPE) and the Chair of IEEE PES PSOPE Committee. In the past, he was a Vice EIC of *Journal of Modern Power Systems and Clean Energy* (MPCE), a Consulting Editor and the Editor of IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, the Editor of IEEE TRANSACTIONS ON POWER SYSTEMS, the Editor of IEEE POWER AND ENERGY SOCIETY LETTERS, a Guest Editor of IEEE TRANSACTIONS ON SMART GRID, a Guest Editor of IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, and the Editorial Board Member of *CSEE Journal of Power and Energy Systems*. He was the past President from 2013to 2015 of North America Chinese Power Professional Association (NACPPA). He is also a Registered Professional Engineer (P.E.) in the state of North Carolina.

Hantao Cui (Senior Member, IEEE) received the B.S. and M.S. degrees from Southeast University, China, in 2011 and 2013, respectively, and the Ph.D. degree from the University of Tennessee, Knoxville, TN, USA, in 2018, all in electrical engineering.

He is currently a Research Assistant Professor with CURENT, the University of Tennessee, where he is also with the Department of Electrical Engineering and Computer Science. He has authored two Essential Science Indicators (ESI) highly cited articles. His research interests include power system modeling and simulation, high-performance computing, and software-hardware solutions for scientific computing. He is an Associate Editor for the *Journal of Modern Power Systems and Clean Energy*. He is recognized as an Outstanding Reviewer for 2019 by the IEEE TRANSACTIONS ON POWER SYSTEMS.



Siqi Bu (Senior Member, IEEEE) received the Ph.D. degree from the electric power and energy research cluster, The Queen's University of Belfast, Belfast, U.K., where he continued his Postdoctoral research work before entering industry. Then he was with the National Grid U.K. as an Experienced U.K. National Transmission System Planner and Operator. He is an Associate Professor with The Hong Kong Polytechnic University, Kowloon, Hong Kong, and also a Chartered Engineer with U.K. Royal Engineering Council, London, U.K.. His research interests include are power system stability analysis and operation control, including wind/solar power generation, PEV, HVDC, FACTS, ESS and VSG.

He is the Editor of IEEE ACCESS, *CSEE Journal OF Power and Energy Systems*, IEEE OPEN ACCESS JOURNAL OF POWER AND ENERGY, *Protection and Control of Modern Power Systems and PEERJ Computer Science*, and a Guest Editor of *IET Renewable Power Generation*, *Energies, IET Generation*, *Transmission & Distribution and Shock and Vibration*. He was the recipient of the various Prizes due to excellent performances and outstanding contributions in operational and commissioning projects during the employment with National Grid U.K. He is also the recipient of the Outstanding Reviewer Awards from IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, IEEE TRANSACTIONS ON POWER SYSTEMS, *Applied Energy, Renewable Energy*, *International Journal of Electrical Power & Energy Systems* and *Journal of Modern Power Systems and Clean Energy* respectively.

Di Shi (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Arizona State University, United States.

He is currently the Director of Fundamental RD Center and Department Head of the AI System Analytics Group, Global Energy Interconnection Research Institute North America (GEIRINA), San Jose, CA, USA. Prior to joining GEIRINA, He was a Research Staff Member with NEC Laboratories America. His research interests include PMU data analytics, AI, energy storage systems, IoT for power systems, and renewable integration. He is the Editor of the IEEE TRANSACTIONS ON SMART GRID and the IEEE POWER ENGINEERING LETTERS.