Sample Generation for Security Region Boundary Identification Based on Topological Features of Historical Operation Data

Xiaokang Wu, Wei Xu, and Feng Xue

***Abstract*—Since the scale and uncertainty of the power system have been rapidly increasing, the computation efficiency of constructing the security region boundary (SRB) has become a prominent problem. Based on the topological features of historical operation data, a sample generation method for SRB identification is proposed to generate evenly distributed samples, which cover dominant security modes. The boundary sample pair (BSP) composed of a secure sample and an unsecure sample is defined to describe the feature of SRB. The resolution, sampling, and span indices are designed to evaluate the coverage degree of existing BSPs on the SRB and generate samples closer to the SRB. Based on the feature of flat distribution of BSPs over the SRB, the principal component analysis (PCA) is adopted to calculate the tangent vectors and normal vectors of SRB. Then, the sample distribution can be expanded along the tangent vector and corrected along the normal vector to cover different security modes. Finally, a sample set is randomly generated based on the IEEE standard example and another new sample set is generated by the proposed method. The results indicate that the new sample set is closer to the SRB and covers different security modes with a small calculation time cost.**

***Index Terms*—Clustering analysis, principal component analysis (PCA), sample generation, security region boundary (SRB).**

I. Introduction

With the increasing penetration of renewable energy and the rapid development of electrical vehicles, energy storage devices, etc., uncertainties on both the supply and demand sides reduce the predictability of the operation states of the power system [1]. Traditional methods of determining the power limit of transmission sections based on given power adjustment directions can hardly reflect the secure operation boundary of the power system with high penetration of renewable energy [2], [3]. Security region boundary (SRB) methodology can determine the security and stability margin of the power system under uncertain power adjustment directions. According to the distance between the current operation point and the SRB, the security and stability status of the power system can be evaluated intuitively and quantitatively [4]-[6]. Therefore, constructing SRB can ensure the secure and stable operation of the power system with high uncertainties.

The research methods for constructing SRB include the analytic method and fitting method. The analytic method deduces the analytic expression of SRB near a critical operation point by linear approximation [7], [8], normal vector method [9], and hyperplane construction [6], [10]. To obtain the complete boundary information of SRB, the critical points containing all different security modes should be calculated. The fitting method constructs the SRB with hyperplane equations based on the distribution of enough critical operation points. The linear fitting method is the most direct way to constructing the SRB in the local range [11]. The convex hull fitting method is adopted to obtain the nonlinear part of SRB [12], [13].

Both the analytic method and fitting method need a large number of critical operation points to construct the complete SRB. The critical operation points are obtained by the pointwise method or the power injection space traversal method [14], [15]. However, the enormous calculation burden of generating critical operation points covering the entire SRB makes these methods difficult to meet the requirements of engineering applications. Based on historical operation data, the artificial intelligence (AI) methods have been widely applied to generate a large number of operation points with a small calculation time cost. AI methods are used to overcome the problems of too many similar samples and insufficient diversity in historical operation data, such as Wasserstein generative adversarial network (WGAN) [16], conditional generative adversarial network (CGAN) [17], long short-term memory (LSTM) [18], and deep neural network (DNN) [19]. Because of the "black box" nature of these models, the conclusions of AI methods have problems of low model transparency and weak interpretability. It is required to adopt another AI method to perform a fast security and stability assessment to identify the SRB. Additionally, there are very few research works providing a standard number and distribution status of the samples used to fit the SRB.

Manuscript received: May 17, 2023; revised: October 7, 2023; accepted: November 28, 2023. Date of CrossCheck: November 28, 2023. Date of online publication: March 28, 2024.

X. Wu is with the College of Energy and Electrical Engineering, Hohai University, Nanjing, China (e-mail: lken923@126.com).

W. Xu is with NARI Group Corporation, Nanjing, China (e-mail: 125481905@qq.com).

F. Xue (corresponding author) is with State Grid Electric Power Research Institute, Nanjing, China (e-mail: xue-feng@sgepri.sgcc.com.cn).

DOI: 10.35833/MPCE.2023.000321

Instead of generating critical operation points by traversing different power adjustment directions, this paper proposes a method that can generate evenly distributed samples to cover dominant security modes based on the topological features of SRB. The boundary sample pair (BSP) composed of a secure sample and an unsecure sample is defined to describe the feature of SRB. The resolution, sampling, and span indices are designed to evaluate the coverage degree of existing BSPs on the SRB and generate samples closer to the SRB. Based on the feature of flat distribution of BSPs over the SRB, the principal component analysis (PCA) is adopted to calculate the tangent vectors and normal vectors of SRB. Then, the sample distribution can be expanded along the tangent vector and corrected along the normal vector to cover different security modes with a small calculation time cost.

II. Characteristics of SRB

Taking the steady-state operation of the power system as an example, the power system can be formulated using a set of power flow equations and a set of inequalities describing the constraints as expressed in (1).

$$\begin{cases} \boldsymbol{F}(\boldsymbol{x})=\boldsymbol{0} \\ \boldsymbol{G}(\boldsymbol{x})>\boldsymbol{0} \end{cases} \tag{1}$$

The state vector $\boldsymbol{x}$, which can uniquely determine the power flow solution, is marked as an operation point. Secure operation points are those vectors that can be solved in the power flow equations $\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{0}$ and satisfy the constraint inequalities $\boldsymbol{G}(\boldsymbol{x})>\boldsymbol{0}$. The set of all the operation points meeting the power flow equations and constraint inequalities makes the steady-state security region.

To facilitate the engineering applications, it is preferred to focus on the SRB in the decision space, which means using a set of decision parameters instead of all the state parameters to make the parameter space. Taking studying the thermal security region as an example, the power injection space is usually used to determine the security status.

Within the scope of simplifying power flow equations using linear functions, the affine transformation can be applied to the state constraints to shape the SRB and the decision parameters [20]. When the constraint inequalities are composed of the variables of state vectors and constants of upper or lower boundaries, the shape of the SRB in the state vector space is a hypercuboid. Each hyperplane on the cuboid refers to a constraint inequality of the corresponding state variable. According to the affine transformation, the SRB in the decision space is a hyperpolyhedron. The hyperplanes on the hyper polyhedron and the hypercuboid are in one-to-one correspondence.

Considering the nonlinear parts in the power flow equations, the hyperplanes may be bent or twisted, but the conclusion to the topology characteristics of the SRB remains, i.e., within the acceptable range of the engineering application, the SRB in the decision space is a simply-connected and compact manifold without boundary. The operation points located within the manifold are secure samples, and those outside the manifold are unsecure ones.

The distribution of historical operation data in the decision space is a probability model. Without preventive actions, the distribution of the operation data and the SRB model are independent of each other. High uncertainty in power systems means that the probability model has a wider distribution and is more likely to intersect with the SRB. This paper proposes to study the samples around the intersection area. New samples are generated closer to the SRB. Then, the sample distribution is expanded based on the topology characteristics of the SRB to obtain samples in different fault modes that are likely to happen, but not included in the data set. The flowchart of the above procedure is shown in Fig. 1.

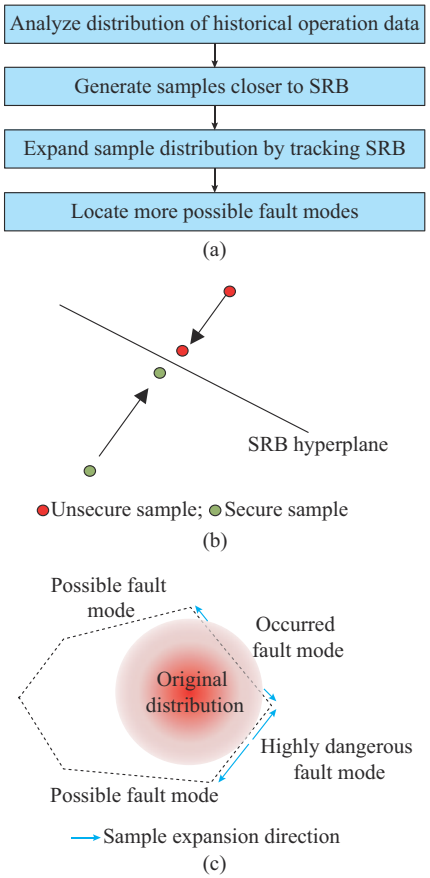

Fig. 1. Flowchart diagram of sample generation. (a) Flowchart of proposed method. (b) Diagram of generating samples closer to SRB. (c) Diagram of expanding sample distribution.

III. Definition of Boundary Samples

Unsecure samples in the historical operation data can be used to study the intersection region of the sample distribution and the unsecure region. If there are few unsecure samples in the historical operation data, it is feasible to create some unsecure samples according to the known security modes. By studying the topology relationship between secure samples and unsecure samples, part of the SRB that intersects with the historical sample distribution can be located.

Given an unsecure sample $\boldsymbol{B}$ in the decision space, for every secure sample $\boldsymbol{A}$, there must be at least one critical sample C on the SRB located on the line between samples $\boldsymbol{A}$

and $\boldsymbol{B}$. For every secure sample or unsecure sample, a critical sample can be potentially located. To reduce the calculation burden, not all the secure samples need to be studied. A few secure samples that are close enough to the unsecure samples can be selected to form the BSPs.

In a 2-dimensional decision space, the samples can be distributed as shown in Fig. 2. BSPs are linked with the blue dot line. A sufficient condition marking boundary samples out of the sample set in an N-dimensional decision space can be described as follows.

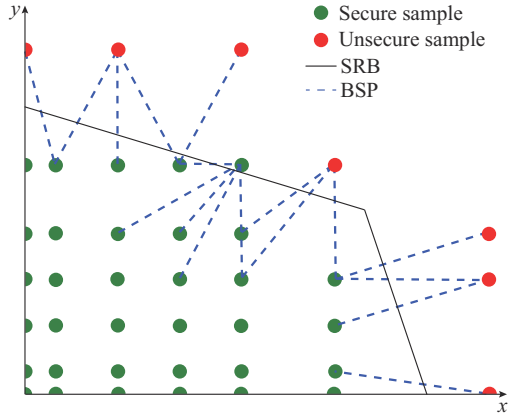

Fig. 2. Diagrammatic sketch for boundary samples around part of SRB.

For a secure sample, there should be an unsecure sample within a certain Euclidean distance in the decision space:

$$\begin{cases} \forall \boldsymbol{A} \in S,\ \exists B \in U \\ \text{s.t. } |\boldsymbol{A}-\boldsymbol{B}| < d \end{cases} \tag{2}$$

where S and U are the secure sample set and the unsecure sample set, respectively; and $\boldsymbol{A}$ and $\boldsymbol{B}$ can be described with the decision vectors. $\boldsymbol{A}$ and $\boldsymbol{B}$ are combined into a BSP, and the modulus of their difference should be less than a constant value d.

The constant value d is introduced to reduce the computation burden caused by the massive historical operation data. It is set to be the minimum value when all unsecure samples can be included in the BSPs, as expressed in (3).

$$d = \max(\min \boldsymbol{\Omega}_{u\times s}[i:]) \quad i = 1, 2, ..., u \tag{3}$$

where $\boldsymbol{\Omega}_{u\times s}$ is the matrix of the distances between the unsecure samples and secure samples; and s and u are the numbers of secure samples and unsecure samples, respectively.

IV. Evaluation of Historical Samples

BSPs cover most information of the SRB in historical operation data. Whether the BSPs are sufficient to deliver the information of the SRB can be assessed from three aspects as follows.

1) Length of BSP. A BSP is equivalent to a critical operation point if the length of BSP approaches 0. Such length should be as short as possible. The overall length of BSP is chosen to be the first index named resolution.

2) Distance between BSPs. The density of BSPs can be equivalent to that of critical operation points on the SRB. The overall distance between BSPs is used to represent the density, which is the second index named sampling.

3) Span of BSP distribution. The span of the BSP distribution represents the range of detectable SRB, which is the third index named span.

A. Resolution

The first index for sample resolution *RSL* can be derived as:

$$\frac{1}{RSL} = \frac{1}{d}\frac{1}{m}\sum_{i=1}^{m}|\boldsymbol{A}_i - \boldsymbol{B}_i| \tag{4}$$

where m is the number of BSPs; and $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ are the decision vectors of the secure sample and unsecure sample in the i^{th} BSP, respectively. The threshold value d is introduced to normalize the index.

Resolution can be supplemented with the interpolation method, i.e., inserting a new sample $\boldsymbol{C}_i$ between $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$. Simulate on $\boldsymbol{C}_i$ and determine its security status. Replace $\boldsymbol{A}_i$ with $\boldsymbol{C}_i$ if it is a secure sample, or replace $\boldsymbol{B}_i$ with $\boldsymbol{C}_i$ if it is an unsecure sample. The new BSP contributes better resolution than the original one.

A constant value e can be set to represent the maximum tolerated length of the BSPs. For any BSP satisfying $|\boldsymbol{A}_i - \boldsymbol{B}_i| > e$, generate a new sample $\boldsymbol{C}_i$ as:

$$\boldsymbol{C}_i = \boldsymbol{A}_i + (\boldsymbol{A}_i - \boldsymbol{B}_i)/2 \tag{5}$$

Simulate on $\boldsymbol{C}_i$ and a new BSP is formed. Repeat the procedure until all BSPs satisfying $|\boldsymbol{A}_i - \boldsymbol{B}_i| \le e$.

B. Sampling

To generate samples distributed evenly in the decision space, [15] proposed a procedure of generating samples to track SRB with the fixed step length and made examples in 2- and 3-dimensional spaces, where the SRB is the combination of lines and planes, respectively. But for an N-dimensional space, the SRB becomes the combination of $N-1$ hyperspaces, which indicates that the number of samples need to make a hypernet exponential increases, which caused the curse of dimensionality.

To avoid the calculation and evaluation seriously affected by dimensions, a list scheme called neighboring samples is proposed. The list scheme is defined according to the single linkage clustering procedures as follows.

Step 1: for samples in $\boldsymbol{S}$, form a distance matrix as $\boldsymbol{\Omega}^{(0)}_{s\times s}$.

Step 2: locate the minimum element of distance matrix $d_{\boldsymbol{L}_1,\boldsymbol{L}_2}$. Samples $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ make a neighboring sample pair with distance $l_{S,1}$. Combine $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ to a new cluster as $\boldsymbol{M}$.

Step 3: calculate the distance between $\boldsymbol{M}$ and another element $\boldsymbol{N}$ as $d_{\boldsymbol{M},\boldsymbol{N}} = \min(d_{\boldsymbol{L}_1,\boldsymbol{N}}, d_{\boldsymbol{L}_2,\boldsymbol{N}})$. Form a new distance matrix $\boldsymbol{\Omega}^{(1)}_{(s-1)\times(s-1)}$ with the cluster $\boldsymbol{M}$ and other elements maintained.

Step 4: repeat *Step 2* until all the elements are combined into one cluster. Neighboring samples of $\boldsymbol{S}$ are obtained.

The average distance between the neighboring BSPs is defined as the sampling index *SA*.

$$\frac{1}{SA} = \frac{1}{d}\frac{1}{m-2}\left(\sum_{i=1}^{s-1} l_{S,i} + \sum_{i=1}^{u-1} l_{U,i}\right) \tag{6}$$

where $l_{S,i}$ and $l_{U,i}$ are the distances between the neighboring secure and unsecure samples, respectively.

Instead of adding samples in the whole hyperspace, new samples are generated between the most distant neighboring samples, which are the most vacant areas of samples. BSPs under different security risks tend to be distributed separately, and hidden security risks may be located among them. Generating samples in the most vacant areas could detect these risks and make the overall distribution continuous.

For all neighboring samples $\boldsymbol{L}_i$ and $\boldsymbol{L}_j$ that satisfy $|\boldsymbol{L}_i-\boldsymbol{L}_j|>e$, new samples $\boldsymbol{C}_k$ are formed as:

$$\boldsymbol{C}_k=\boldsymbol{L}_i+k(\boldsymbol{L}_i-\boldsymbol{L}_j)/n \quad k=1,2,...,n \tag{7}$$

where n is the number of samples generated between the neighboring samples. By setting $n=|\boldsymbol{L}_i-\boldsymbol{L}_j|\backslash e$, new samples are ensured to have neighboring distances smaller than e.

C. Span

The span of the samples *SPAN* describes the range of BSP in the feasible decision space. The actual SRB is unknown yet, and the coverage percentage of BSP to SRB is not feasible, so a relative value of the coverage percentage of BSP to the decision space is created instead. Considering the effect of dimensionality, the percentages of dimensions are added instead of multiplied to calculate the volume.

$$SPAN=\sum_{i=1}^{N}\frac{a_{i,\max}-a_{i,\min}}{\eta_{i,\text{up}}-\eta_{i,\text{low}}} \tag{8}$$

where N is the number of dimensions; $a_{i,\max}$ and $a_{i,\min}$ are the maximum and minimum values of the parameters of all the BSPs, respectively; and $\eta_{i,\text{up}}$ and $\eta_{i,\text{low}}$ are the i^{th} upper and lower boundaries of the feasible decision space, respectively.

V. Boundary Sample Generation

It is proven that the SRB caused by a certain security mode can be approximated with a hyperplane [20]. The actual SRB is the combination of several hyperplanes. The BSPs can be classified into several sets according to their dominant security modes to construct the SRB. Each set is studied individually and the sample distribution is expanded along with the corresponding SRB until the simulation analysis of the new sample reveals that the security mode changes.

A. Tangent and Normal Vector Computation

To determine the directions of generating new samples and correcting parameters when the generated BSP loses track to the SRB, the tangent vectors and normal vector of the hyperplane are to be found.

Besides using analytic methods or sensitivity analysis to obtain the tangent vectors and normal vector of the SRB hyperplane, a statistical method based on the previous sample generation results is proposed. After resolution and sampling compensation, it is expected that the BSPs are flat distributed along the SRB. On this basis, the vectors from the center to the edges can be used as the tangent vectors and the normal vector can be obtained using principal component analysis (PCA).

PCA is a common statistical method for dimension reduction and feature extraction. In this paper, PCA is proposed to be used differently. The geometric interpretation of PCA is finding the largest projection of the distribution as the 1st principal component. Among the normal directions of the 1st principal component, the largest projection is found as the 2nd principal component. For flat distribution samples over a hyperplane, the least principal component should be the normal vector of this hyperplane. The detailed steps are as follows.

1) Divide $\boldsymbol{U}$ into $\boldsymbol{U}_1,\boldsymbol{U}_2,...,\boldsymbol{U}_r$ according to the security risks, where r is the number of possible security modes. According to the BSP relationship, $\boldsymbol{S}$ is also divided into $\boldsymbol{S}_1,\boldsymbol{S}_2,...,\boldsymbol{S}_r$. For each set, perform the following procedures to obtain the tangent vectors and the normal vector.

2) Compute the expectation of $\boldsymbol{U}_i$ as:

$$\boldsymbol{E}(\boldsymbol{U}_i)=\frac{1}{m}\sum_{j=1}^{m}\boldsymbol{B}_{i,j} \tag{9}$$

where m is the number of samples in $\boldsymbol{U}_i$; and $\boldsymbol{B}_{i,j}$ is the j^{th} unsecured sample in $\boldsymbol{U}_i$.

3) Find the samples with the largest k^{th} parameter among $\boldsymbol{U}_i$ noted as $\boldsymbol{B}_{i,\max,k}$. Compute $\boldsymbol{B}_{i,\max,k}-\boldsymbol{E}(\boldsymbol{U}_i)$ and its unit vector $\boldsymbol{u}_{i,\max,k}$. Repeat it for all dimensions and their minimum value. The tangent vectors are obtained.

4) Compute the covariance matrix $\boldsymbol{\Sigma}_i$ for $\boldsymbol{U}_i$ as:

$$\boldsymbol{\Sigma}_i=\boldsymbol{E}(\boldsymbol{U}_i-\boldsymbol{E}(\boldsymbol{U}_i))(\boldsymbol{U}_i-\boldsymbol{E}(\boldsymbol{U}_i))^{\text{T}} \tag{10}$$

The minimum eigenvalue and the corresponding eigenvector of $\boldsymbol{\Sigma}_i$ are noted as $w_{i,\min}$ and $\boldsymbol{v}_{i,\min}$, respectively. $\boldsymbol{v}_{i,\min}$ can be used as the normal vector of the SRB formed by the i^{th} security mode.

5) To make sure the normal vector points outwards the SRB, check if $\boldsymbol{v}_{i,\min}$ satisfies $(\boldsymbol{E}(\boldsymbol{U}_i)-\boldsymbol{E}(\boldsymbol{U}\cup\boldsymbol{S}))\boldsymbol{v}_{i,\min}<\boldsymbol{0}$, and let $\boldsymbol{v}_{i,\min}=-\boldsymbol{v}_{i,\min}$.

By performing the above procedures, the tangent vectors and normal vector of the SRB can be obtained.

B. Expanding Sample Distribution

With the previous work, new samples can be generated along the tangent vectors for $\boldsymbol{S}_i$ and $\boldsymbol{U}_i$. Taking unsecure samples in the k^{th} maximum direction of $\boldsymbol{U}_i$ as an example, $\boldsymbol{C}_{i,\max,k}$ can be generated with $\boldsymbol{B}_{i,\max,k}$ and $\boldsymbol{u}_{i,\max,k}$ as expressed in (11).

$$\boldsymbol{C}_{i,\max,k}=\boldsymbol{B}_{i,\max,k}+e\boldsymbol{u}_{i,\max,k} \tag{11}$$

Perform simulation on $\boldsymbol{C}_{i,\max,k}$. It is expected that it should be an unsecure sample. If the result is secure, correct the parameters with $\boldsymbol{v}_{i,\min}$ as expressed in (12).

$$\boldsymbol{C}'_{i,\max,k}=\boldsymbol{C}_{i,\max,k}+e\boldsymbol{v}_{i,\min} \tag{12}$$

Perform simulation on $\boldsymbol{C}'_{i,\max,k}$. It is expected that it should be an unsecure sample.

New samples generated from $\boldsymbol{S}_i$ and $\boldsymbol{U}_i$ on the same k^{th} parameter direction make a new BSP. Procedures in Section III can be applied to narrow its distance and generate more samples.

VI. Sample Generation Procedures

The sample generation procedure can be divided into two phases, namely index analysis and sample generation, as shown in Fig. 3.

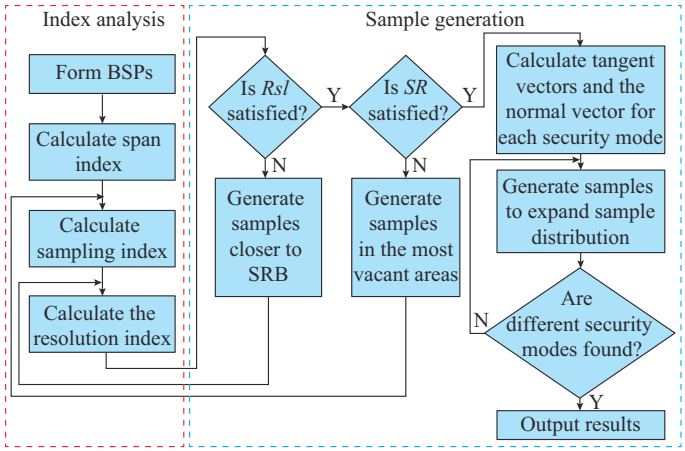

Fig. 3. Flowchart for index analysis and sample generation.

The overall process starts with a sample set consisting of a large number of secure samples and some unsecure samples covering the known security modes. The Euclidean distance matrix $\boldsymbol{\Omega}$ of the sample set is computed in the first step. Then, the original sample set can be evaluated with the indices.

Most of the historical operation points are secure samples and tend to be far from the SRB. To make sure that there are enough BSPs to start with, the threshold value of the initial boundary sample d is set to be a relatively large number, so that all the unsecure samples can be included. A large d leads to a low sample resolution index. The first step of sample generation is drawing samples closer to the SRB to improve the resolution index with the interpolation method. To identify changes in security modes of SRB and locate their positions, the second step is sampling supplement. The clustering analysis is adopted to link the neighboring BSPs and the interpolation method is adopted to generate samples in the most vacant areas.

With evenly distributed large number of sample pairs for each security mode, an applicable environment for PCA is established. After calculating the tangent vectors and normal vectors using PCA, the sample distribution can be expanded so that the security modes absent in the data can be discovered.

VII. Simulations

In this section, the proposed method is verified on the IEEE 39-bus system, which is designed to identify the thermal security region in the active power injection space with 2, 3, and 5 dimensions. Other two sample generation methods based on the WGAN and traversing SRB with fixed steps are also performed for comparison. The sample generation programs are built in Python. The security analysis of a single operation point is obtained using Pypower. The hardware for the calculation of the proposed method is a personal computer with an Intel Core i7-13700F 2.10 GHz CPU and 32 GB RAM. The WGAN is trained on NVIDIA GeForce GTX 3070Ti GPU.

A. Parameter Space Setup

To visualize the results and investigate the effectiveness of the proposed method, a 2-dimensional decision space is generated. The IEEE 39-bus system is divided into 3 zones according to power transmission directions, as shown in Fig. 4. The active power of generators in zone 1 is multiplied by a number denoted as P_1 and the active power of loads in zone 3 is multiplied by a number denoted as P_2. The difference in power is balanced by the slack generator at bus 31.

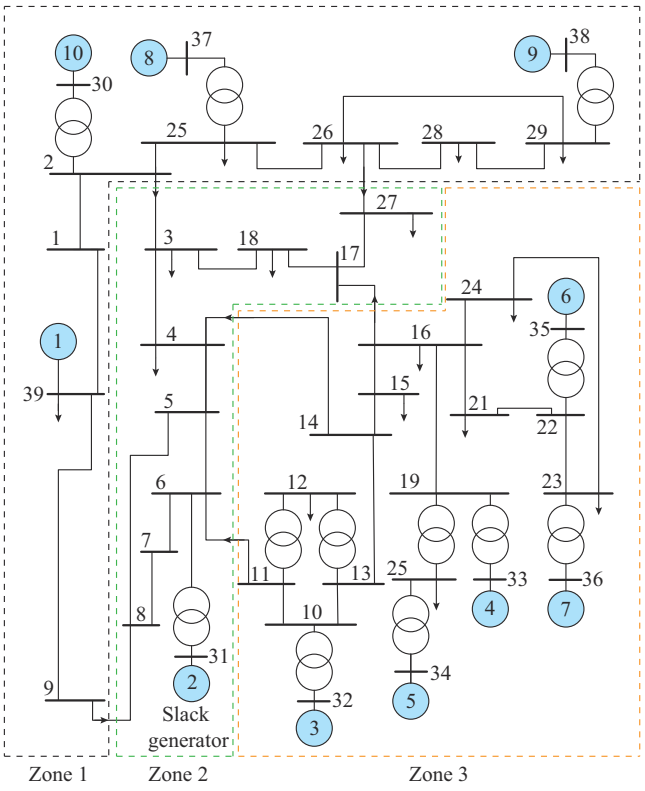

Fig. 4. Diagram of IEEE 39-bus system.

B. Steady-state Secure Region

The actual secure region is shown in Fig. 5(a), which is a polygon consisting of five straight lines, representing five different branches that could potentially fail the security check. The five secure modes are marked with the bus numbers on each side of the branch.

100 samples are normally randomly generated around (1, 1) as the original sample set, as shown in Fig. 5(b). The unsecure samples only consist of three secure modes, i. e., overcurrent occurs on branches 6-11, 2-3, and 2-25. 40 samples are selected from the 100 original samples as BSPs, as shown in Fig. 5(c).

The proposed method generates 92 samples based on BSPs, as shown in Fig. 6(a). During the sample expansion, the secure modes of overcurrent on branches 5-6 and 17-18 are discovered.

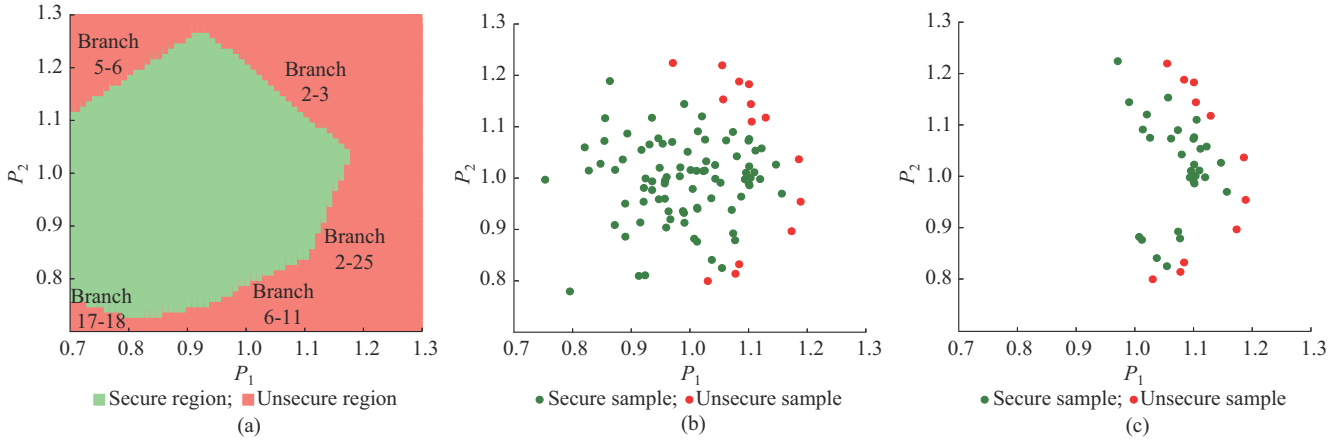

Fig. 5. 2-dimensional actual secure region and distributions of generated samples. (a) Actual secure region. (b) Normal randomly generated samples. (c) BSP distribution.

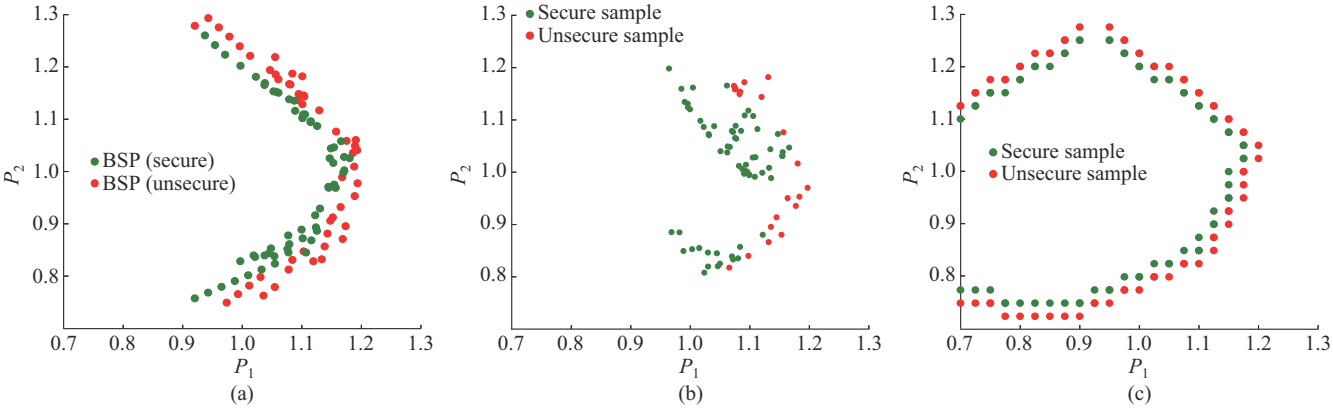

Fig. 6. Sample distribution of three sample generation methods in 2-dimensional space. (a) Proposed method. (b) WGAN. (c) Traversing method.

WGAN is adopted to generate samples using the BSPs as the training set. The model parameters of WGAN are shown in Table I. The 90 generated samples in 3 epochs after the D-loss converges are used as the sample generation results. The generated sample distribution using WGAN is shown in Fig. 6(b). The samples generated by WGAN cover three security modes of overcurrent on branches 6-11, 2-3, and 2-25.

TABLE I
MODEL PARAMETERS OF WGAN

Model parameter	Value
Discriminator training times in each epoch	5
Generator training times in each epoch	1
Batch size	30
Learning rate	0.0005
Epochs	3000

The indices of the above three methods in the 2-dimensional thermal security problem are compared, as shown in Table II. Note that the time cost of WGAN includes the training time cost and security check of the samples.

In the 2-dimensional thermal security problem, the traversing method is faster than the other two methods. It can generate samples that cover the entire SRB. The proposed method has advantages in resolution and sampling rate index and is also able to generate samples covering all five security modes. WGAN costs more training time and it only increases the sample with the same distribution as the BSPs, making the generated samples less worthy.

TABLE II
INDICES OF THREE SAMPLE GENERATION METHODS IN 2-DIMENSIONAL THERMAL SECURITY PROBLEM

Method	Time cost (s)	Number of security modes	Span	Resolution	Sampling rate
The proposed method	1.692	5	0.681	1.883	9.308
WGAN	16.945	3	0.520	1.531	8.992
Traversing method	1.060	5	0.875	1.942	8.636

C. Sample Generation in 3- and 5-dimensional Hyperspaces

A 3-dimensional hyperspace can be created by setting the increased rate of active power of generators in zone 1 as P_1, setting the increased rate of active power of loads in zone 2 as P_2, and setting the increased rate of active power of loads in zone 3 as P_3.

The actual security region is shown in Fig. 7(a), which is

a polygon consisting of six planes, representing six different branches that could potentially fail the security check. 500 samples are normally randomly generated around $(1, 1, 1)$ to be the original sample set, as shown in Fig. 7(b). The unsecure samples only consist of four security modes, i.e., overcurrent occurs on branches 6-11, 2-3, 5-6, and 2-25.

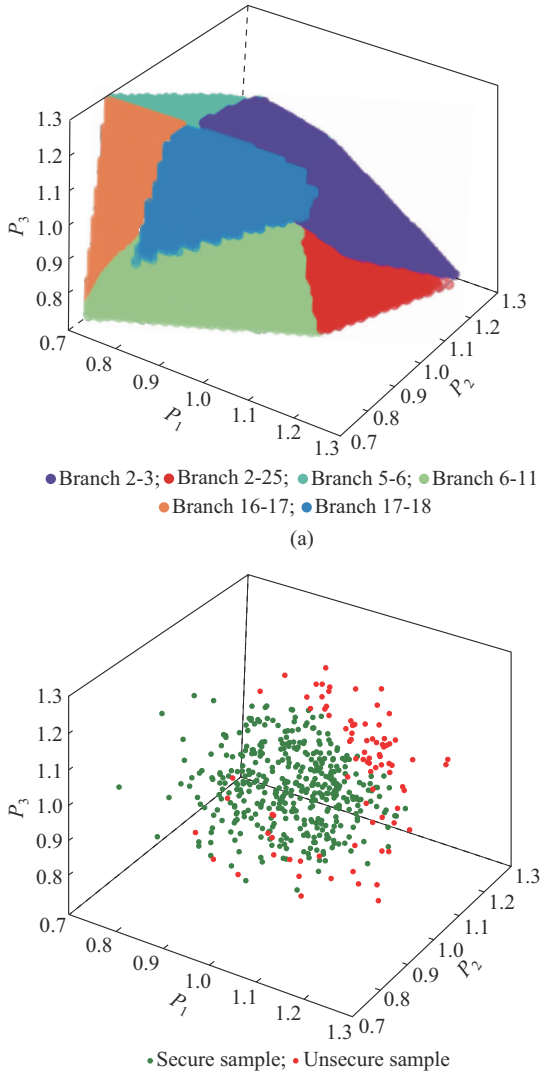

Fig. 7. 3-dimensional actual security region and distributions of generated samples. (a) Actual security region of different faults. (b) Normal randomly-generated samples.

Using the proposed method, 226 samples are generated, as shown in Fig. 8(a). By expanding the span, the overcurrent on 17-18 is discovered. By adjusting the batch size to 50 and using WGAN to generate samples, 250 samples are generated in the five batches after D-loss converges, as shown in Fig. 8(b). The generated samples consist of four security modes like the original sample set. The traversing method generates 3405 samples, as shown in Fig. 8(c). The generated samples cover all 6 security modes.

The indices of the above three methods in the 3-dimensional thermal security problem are compared, as shown in Table III.

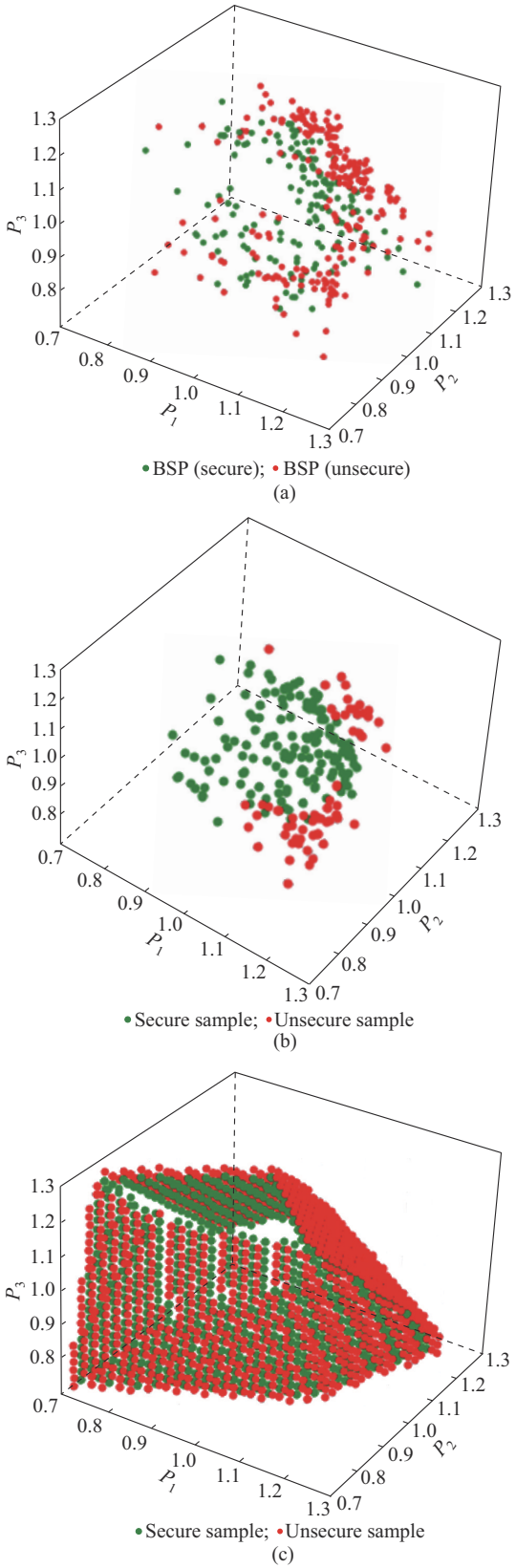

Fig. 8. Sample distribution of three sample generation methods in 3-dimensional space. (a) Proposed method. (b) WGAN. (c) Traversing method.

In the 3-dimensional thermal security problem, the traversing method covers the entire SRB and achieves better scores in the three indices, but takes much more time than other

two methods. The proposed method takes the least time and has decent indices. The problem is that one security mode is missing in the generated samples. By comparing Fig. 7(a) and Fig. 7(b), it can be concluded that the original sample distribution is distant from the SRB of the fault on branch 16-17, which means that the chance that the operation point shifting to fault on branch 16-17 is relatively small.

TABLE III
INDICES OF THREE SAMPLE GENERATION METHODS IN 3-DIMENSIONAL THERMAL SECURITY PROBLEM

Method	Time cost (s)	Number of security modes	Span	Resolution	Sampling rate
The proposed method	6.070	5	0.920	1.415	5.207
WGAN	11.303	4	0.715	1.368	4.939
Traversing method	27.527	6	0.920	1.524	12.444

By setting the increased rate of active power of generators in zone 1 as P_1, setting the increased rate of active power of loads in zone 1 as P_2, setting the increased rate of active power of loads in zone 2 as P_3, setting the increased rate of active power of generators in zone 3 as P_4, and setting the increased rate of active power of loads in zone 3 as P_5, a 5-dimensional decision space is created by applying the above procedures to generate samples.

Using the traversing method would generate millions of samples in the 5-dimensional decision space, and could take days for sample generation and evaluation. The calculation burden is getting massive as the dimensionality increases. As a result, only the proposed method and WGAN are compared in the 5-dimensional thermal security problem, as shown in Table IV.

TABLE IV
INDICES OF PROPOSED METHOD AND WGAN IN 5-DIMENSIONAL THERMAL SECURITY PROBLEM

Method	Time cost (s)	Number of security modes	Span	Resolution	Sampling rate
The proposed method	13.688	6	0.943	1.287	4.021
WGAN	12.745	4	0.673	1.368	4.302

In the 5-dimensional space, the proposed method takes longer time than WGAN. From the aspect of indices, the proposed method shows an advantage in span and covers more security modes, but achieves less resolution and sampling rate. The numbers of samples generated by the two methods are similar but the difference in their span indices is significant, which means the density of samples generated by WGAN is much higher than those by the proposed method. The difference in density is reflected in resolution and sampling rate, which are designed to measure the density across and along the SRB, respectively.

From the aspect of calculation burden, WGAN has great advantages. Little increase in training time is observed as the dimensionality increases. The calculation time of the proposed method is proportional to the dimensions and that of the traversing method exponentially increases with the dimensions.

From the aspect of generated sample evaluation, the traversing method can cover the entire SRB and find all the security risks. It is a better choice when the calculation burden can be ignored. WGAN generates samples with the least time cost. However, the generated samples share the same distribution as the original samples. The proposed method can expand the given sample distribution and track the SRB. Considering the nature of the historical operation data, it is a better way to generate samples to cover possible security modes at an acceptable calculation cost.

VIII. CONCLUSION

In this paper, three indices that evaluate the sufficiency of a sample set used for SRB identification are proposed. A sample generation procedure is proposed to improve the three indices and discover potential security modes. The procedure is verified on the IEEE 39-bus system and compared with the sample generation method using WGAN. The following conclusions can be drawn.

1) The sample generation method using historical operation data has the advantage of high efficiency, but handing over the entire process to AI methods such as WGAN can only mimic the distribution of existing samples and cannot expand the sample distribution and locate new security modes. The proposed method alternately conducts the statistical analysis of sample distribution and simulates the generated samples, which can effectively expand and optimize the sample distribution.

2) The sample generation method can generate samples quickly approaching the SRB and expand sample distribution to cover more power adjustment directions, which can describe the process of transformation between different faults and explore security risks missing in the original samples. Further research will focus on constructing the SRB from the generated samples and guiding the safety and security operation of power systems.

REFERENCES

[1] N. Hatziargyriou, J. Milanovic, C. Rahmann *et al*., "Definition and classification of power system stability-revisited & extended," *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 3271-3281, Jul. 2021.
[2] X. Li and Z. Guo, "Analysis of network transmission capability based on *N*–1 principle," *Automation of Electric Power Systems*, vol. 28, no. 9, pp. 28-30, May 2004.
[3] Z. Su, Z. Kang, W. Jiang *et al*., "An active power control strategy for large-scale cluster wind power with nested transmission section constraints considering the control delay effect," in *Proceedings of 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, Beijing, China, Oct. 2018, pp. 1-5.
[4] Y. Ji, Q. Xu, and Y. Xia, "Distributed robust energy and reserve dispatch for coordinated transmission and active distribution systems," *Journal of Modern Power Systems and Clean Energy*, vol. 11, no. 5, pp. 1494-1506, Sept. 2023.
[5] T. Yang and Y. Yu, "Steady-state security region-based voltage/var optimization considering power injection uncertainties in distribution grids," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2904-2911, May 2019.
[6] J. Xiao, Q. He, and G. Zu, "Distribution management system framework based on security region for future low carbon distribution sys-

tems," *Journal of Modern Power Systems and Clean Energy*, vol. 3, no. 4, pp. 544-555, Oct. 2015.
[7] A. Xue, W. Hu, S. Mei *et al*., "Comparison of linear approximation for the dynamic security region of power systems," *Automation of Electric Power Systems*, vol. 30, no. 5, pp. 9-13, Mar. 2006.
[8] C. Guo and Y. Yu, "Boundary of thermal security region in decision making space of power system," *Automation of Electric Power System*, vol. 37, no. 18, pp. 42-47, Sept. 2013.
[9] J. Xu, Y. Chen, Y. Fan *et al*., "Dynamic security regions of electric power system based on extended equal-area criterion," *Proceedings of the CSEE*, vol. 27, no. 31, pp. 20-26, Nov. 2007.
[10] F. Li, T. Niu, L. Xue *et al*., "Autonomous-synergic voltage security regions in bulk power systems," *Journal of Modern Power Systems and Clean Energy*, vol. 11, no. 3, pp. 686-692, Mar. 2023.
[11] Y. Zeng, J. Fan, Y. Yu *et al*., "Practical dynamic security region of bulk power systems," *Automation of Electric Power Systems*, vol. 28, no. 16, pp. 6-10, Aug. 2001.
[12] H. D. Nguyen, K. Dvijotham, and K. Turitsyn, "Constructing convex inner approximations of steady-state security regions," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 257-267, Jan. 2019.
[13] S. Chen, Z. Wei, G. Sun *et al*., "Convex hull based robust security region for electricity-gas integrated energy systems," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 1740-1748, May 2019.
[14] X. Li, T. Jiang, L. Bai *et al*., "Orbiting optimization model for tracking voltage security region boundary in bulk power grids," *CSEE Journal of Power and Energy Systems*, vol. 8, no. 2, pp. 476-487, Mar. 2022.
[15] J. Xiao, G. Zu, X. Gong *et al*., "Observation of security region boundary for smart distribution grid," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1731-1738, Jul. 2017.
[16] Y. Liao and Z. Wu, "Critical sample generation method for static voltage stability based on transfer learning and Wasserstein generative adversarial network," *Power System Technology*, vol. 45, no. 9, pp. 3722-3728, Sept. 2021.
[17] B. Tan, J. Yang, Q. Lai *et al*. "Data augment method for power system transient stability assessment based on improved conditional generative adversarial network," *Automation of Electric Power Systems*, vol. 43, no. 1, pp. 149-157, Jan. 2019.
[18] J. Chen, Y. Chen, F. Tian *et al*. "The method of sample generation for power grid simulation based on LSTM," *Proceedings of the CSEE*, vol. 39, no. 14, pp. 4129-4135, Jul. 2019.
[19] S. Wu, W. Hu, Z. Lu *et al*., "Power system flow adjustment and sample generation based on deep reinforcement learning," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 6, pp. 1115-1127, Nov. 2020.
[20] Y. Yu, Y. Liu, C. Qin *et al*., "Theory and method of power system integrated security region irrelevant to operation states: an introduction," *Engineering*, vol. 6, no. 7, pp. 754-777, Jul. 2020.

Xiaokang Wu received the M.E. degree from the University of Melbourne, Melbourne, Australia, in 2018. He is pursuing the Ph. D. degree at Hohai University, Nanjing, China. His research interests include power system stability and control, machine learning, and data mining applications in power system.

Wei Xu received the B. S. and Ph. D. degrees from Southeast University, Nanjing, China, in 2003 and 2009, respectively. His research interests include power system stability and control.

Feng Xue received the B. S. degree from Shanghai Jiao Tong University Shanghai, China, and the Ph.D. degree from Bath University, Bath, U.K., in 2008. He is currently a Professor at Hohai University, Nanjing, China. At the same time, he is working as the Chief Expert of NARI Group Corporation and the Director of the System Protection Laboratory. His research interests include power system security and risk analysis and dynamics.