INVESTIGATION OF DIFFERENT SIMULATION METHODS FOR EXTRACTING RARE FAILURE EVENTS IN POWER SYSTEM RELIABILITY EVALUATION

Vladislav Oboskalov[1,2], Sviatoslav Timashev[1], Amir Abdel Menaem[2,3], Rustam Valiev[2], Gennady Shmal[2]

•

[1]SEC «Reliability and Safety of Large Systems and Machines» UB RAS, RUSSİA
[2]Ural Federal University, Yekaterinburg, RUSSİA
[3]Electrical Engineering Department, Mansoura University, Mansoura, EGYPT
ashassan@mans.edu.eg

Abstract

In this paper, procedures for modeling rare events, which include power and energy shortages, are considered when analyzing the balance reliability of electric power systems using simulation methods. The primary goal is to investigate the use of iterative methods to form a sequence of nested probability subspaces with decreasing probability. Implementations of the subspace method are considered. A modification of the entropy method is proposed that involves a smooth adaptation of the indicator functions. The discussed procedures undergo a comparative analysis. The main disadvantage of the considered methods is their efficiency, which depends on the proper specification of parametric constants. The methods to improve the accuracy and convergence of iterative procedures are outlined.

Keywords: Monte Carlo simulation, power system reliability, simulation of rare events, cross-entropy method, probabilistic evaluation

I. Introduction

The main task of analyzing the reliability of an electric power system (EPS) is to assess the probabilistic indicators (probability, frequency, mathematical expectation, etc.) of power deficit (PD) in the EPS, the presence of which is considered a violation of the normal functioning of the EPS (failure) [1–3]. All technical systems, including EPS, exhibit a low probability of failure (a fraction of a percent) [4–5]. A technical system's reliability analysis is associated with the modeling and analysis of failure events. In this case, as a rule, failure occurs with an infinitely large combination of external events. Thus, in an EPS with two load nodes and a fixed maximum power of electric stations, a power deficit can manifest itself with an unacceptably large increase in the load of either the first, second, or both nodes. In this case, the excess of the total load over generation is continuous and hence has an infinitely uncertain value in implementation. The number of load and generation nodes in real EPS design circuits is measured in thousands. As a result, the probability of failure increases many times over, and identifying the most significant combinations of events leading to system failure becomes difficult.

The main mathematical apparatus for analyzing the reliability of EPS is currently the methods of statistical tests, and in particular the Monte Carlo method (MCM) [6–10]. These methods make it possible to model systems with complex functional relationships that cannot be described analytically, including taking into account the stochastic uncertainty associated with supply, demand, and the capacity of intersystem connections. The main drawback of the MCM is

known: to ensure the required accuracy of the results, a sufficiently large number of tests is required, since N is inversely proportional to the probability of the simulated events. In particular, when modeling an event with a probability of 10-3 and a modeling error of 1%, at least 106 statistical tests are required [11]. When analyzing the balance reliability of real EPS, the most significant are multiple failures (simultaneous failure of two power transmission lines, simultaneous failure of a power line and a power unit, failure of a power unit during peak loads, etc.), since single failures of EPS electrical equipment, as a rule, do not lead to limitations on power consumption. EPS are designed and constructed taking into account criterion N-1, in which a single failure of any element of the EPS should not lead to failure of the power supply process to consumers [12–14]. The probability of multiple independent failures is equal to the product of the failure probabilities of the failed elements. As a result, the probability of multiple failures takes values of 10-5–10-7. At the same time, to identify, model, and analyze post-emergency states of EPS using the MCM method, the number of tests is required to be at least 107. With a smaller number of tests, the identification of such rare events becomes unlikely, which leads to a significant error in the resulting reliability indicators [14].

Given that each test necessitates the execution of optimization calculations (optimal flow distribution), the reliability analysis of EPS will require several hours of computer time. Note that most general problems, such as choosing the optimal configuration of the electrical network during its development, involve the calculation of reliability indicators. Hence, the duration of calculating the reliability of the EPS development option as a separate calculation block is limited in time. The modern power system requires increasingly complex detail and expansion of the factors taken into account, which non-linearly increases the complexity of the task of assessing the reliability of the EPS. Taken together, this leads to the inexpediency, and sometimes even impossibility, of using a standard MCM to assess the reliability indicators of a power system. Modifications of the MCM are required, aimed at increasing the computational efficiency of statistical methods.

One of the ways to solve the problem of performing a large number of relatively similar calculations associated with modeling rare (with a failure probability of less than 10-5) events in a multidimensional probabilistic space is the idea of parallel computing. The technology for simultaneous use of several computers (or processors) has been actively developing over the past 30 years [15–17]. Parallel calculations certainly reduce the computational load of the MCM. However, the relationship between the number of parallel computing resources and computation time is similar to a linear function [17]. Hence, when modeling rare events, the positive effect of parallelization of calculations becomes less obvious. Another direction is the use of pattern methods, metamodels, artificial immunity, artificial neural networks, least-squares support vector machines [18–21], etc. These methods have also demonstrated a reduction in the MCM method's computational load. However, as a rule, for reliability evaluation tasks, they have not been tested, are focused on a certain class of problems, are often ineffective, and do not provide the required accuracy and reliability of identifying events with a probability of less than 10-4, which is typical for real energy systems. Hence, the problem of identifying rare events remains relevant. Iterative methods associated with the transformation of analyzed spaces, functions of probability distributions of random variables, and criterion functions are more universal in nature. The most common methods in this class are subspace and significant sampling methods [22–25], as well as cross-entropy methods [26–28].

This paper proposes and analyzes new procedures for determining the probabilistic parameters of rare events based on nested set methods, including mono and polycenters for forming an intermediate sample of significant events, as well as a modification of the cross-entropy method. The models and methods used in this study are described in Section III and IV, after Section II, where the mathematical formulation of rare event simulation is highlighted. Section V presents an example case study, and the results are analyzed to address the objectives of

the work. In Section VI, conclusions are drawn.

II. Mathematical formulation of Power Deficit as a rare event

Let $\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{m}}$be a random vector that combines all probabilistic input variables, L_i, G_i, etc. In the simplest case, when only available generation and load are taken into account and random states of system elements, ambient temperature, etc. are not taken into account, $\boldsymbol{x} = \{L_i, G_i,\ i = 1, \dots, n\}$, where i –where is the number of the electrical network node. In order to focus attention on the mechanisms for identifying rare events associated with failures of EPS operation, the set of random, probabilistically determined quantities in this work is limited only by the load and available generation.

The deficiency of the system is determined by the difference $D_\Sigma = L_\Sigma - G_\Sigma = \sum L_i - \sum G_i$, $L_\Sigma > G_\Sigma$. It can be estimated, for example, using the system power deficit (PD) function, $S = G_\Sigma - L_\Sigma$ or (taking into account the capacity of the electrical network) $S = \min(G_i - L_i, \forall i)$. When $\{L_i, G_i\}$ is independent, it often makes sense to consider a generalized random variable - the available generating capacity of a node, $r_i = G_i - L_i$. Its mean and variance: $\mu_{r_i} = \mu_{G_i} - \mu_{L_i};\ \sigma_{r_i}^2 = \sigma_{G_i}^2 + \sigma_{L_i}^2$. As a result, the number of control stochastic variables is reduced to the number of nodes in the electrical network. In this case, $D_\Sigma = -r_\Sigma = -\sum r_i,\ r_\Sigma < 0$.

In general, when analyzing local PDs, it is necessary to take into account the laws of power distribution in the electrical network. In this case, each node is characterized by the export $u_i, i = 1, \dots, n$ or import (-u_i) of power, determined by the dispatch control of the EPS and depending on the available power of the nodes $u_i = u_i(\boldsymbol{r})$. In this case, the local PD $D_i = u_i - r_i = u_i - G_i + L_i;\ D_i > 0$. With local PD, the system PD is fixed if at least one of the nodes has a local PD $D_\Sigma = \sum D_i$. To more accurately account for the scarcity of EPS, it is necessary to take into account active power losses in the network, which are determined by the distribution of power flows $\boldsymbol{z} = \{z_j, j = 1, \dots, m\}$ in the elements of the electrical network and are nonlinear in nature. Taking into account power losses is associated with the appearance of an additional condition $\pi_\Sigma = \sum u_i$. In this case, $D_\Sigma = L_\Sigma - G_\Sigma + \pi_\Sigma,\ \mathcal{D}_\Sigma \geq 0$.

Vector $\boldsymbol{z}$ is determined by solving the problem of optimal distribution of load power between power sources. In the simplest case, acceptable in reliability assessment problems, we can consider $\boldsymbol{z} = C\boldsymbol{u}$ where C is the flow distribution matrix [7]. As a result, it can be noted that the PD in an EES represents a complex (often determined algorithmically) functional dependence on a set of control variables, usually of a random nature.

The system failure state is fixed if some criterion function $\psi(\boldsymbol{x})$ specified on the set of control variables exceeds its inherent threshold value, $\psi(\boldsymbol{x}) < \psi_{lim}$. In practice as shown in Fig. 1, to identify failures, a function with a zero threshold is more often used, $S(\boldsymbol{x}) = \psi(\boldsymbol{x}) - \psi_{lim}$, which takes a negative value when the system fails, $S(\boldsymbol{x}) < 0$. In this case, the region of system states characterized as "failure":

$$H := \{\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{m}} : S(\boldsymbol{x}) < 0\}. \quad (1)$$

If we consider the presence of PD in the EPS as a failure, then the criterion function has the form:

$$S(\boldsymbol{r}) = \min\left(\sum r_i - \pi_\Sigma;\ r_i - u_i(\boldsymbol{r}), \forall i\right).$$

The criterion function corresponds to the indicator function:

$$J(\boldsymbol{x}) = \begin{cases} 1, & S(\boldsymbol{x}) < 0; \\ 0, & S(\boldsymbol{x}) \geq 0, \end{cases}$$

which allows you to express many logical constructions in analytical form. In particular, the probability $\mathbb{P}_H$ of a failure state (the presence of a power deficit in the EPS) can be defined as the mean of the indicator function $J(\boldsymbol{x})$, which in state $\boldsymbol{x}$ takes the value 1 in the presence of a failure in the EPS and 0 in its absence:

$$\mathbb{P}_H = \int J(\boldsymbol{x})\, f(\boldsymbol{x}) d\boldsymbol{x}, \tag{2}$$

where $f(\boldsymbol{x})$ – is the probability density function (PDF) of the multidimensional random variable $\boldsymbol{x}$ that determines the event under consideration. For a given PDF $f(\boldsymbol{x})$, the probability of a rare event occurring is determined by the region H, that satisfies the criterion $S(\boldsymbol{x}) < 0$.

$$\mathbb{P}_H = \int_{S(\boldsymbol{x})<0} f(\boldsymbol{x}) d\boldsymbol{x}.$$

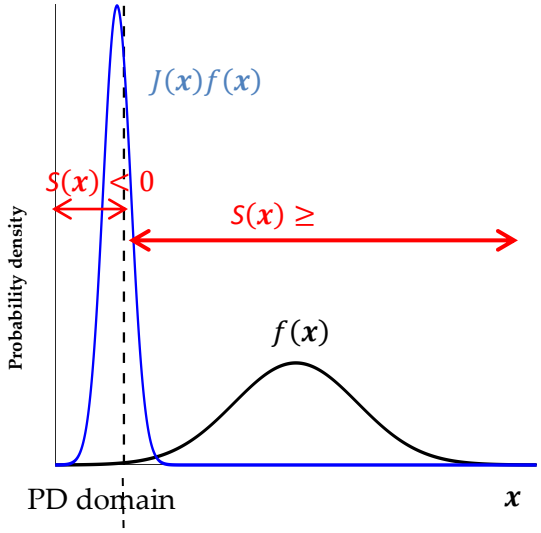

Figure 1: *Derivation of distorted probability distribution*

The PDF $f(\boldsymbol{x})$ of a random vector $\boldsymbol{x}$ of system states is almost impossible to represent in the form of an analytical expression, since it is the result of the convolution of a large number of system parameters of different types, with different parameters (loads, generations, binary variables that determine the state of system elements, solar radiation, wind speed, etc.). That is why the main method of analyzing the reliability of EPS at present is the simulation method (MCM), which allows, through repeated tests and their statistical processing, to obtain a fairly accurate estimate of the required values. In this case, as initial data, as a rule, the parameters of marginal distributions of initial RVs are set (for example, the load of node i of an EPS is described by a normal (Gaussian) distribution with a mathematical expectation μ_i and a standard deviation (STD) σ_i, $L_i \sim N(x, \mu_i, \sigma_i)$.

III. State subspace method

The performance of the MCM can be improved by applying variance reduction techniques such as significance sampling. The main idea of this group of methods is the formation of a sequence of subsets, $H_1 \supset H_2 \supset \cdots \supset H_m = H$ of the system state space, where each subsequent subspace increases the probability of identifying rare events and is determined on the basis of the previous one, forming a sequence of the Markov chain. In this case, the probability of the occurrence of a rare event can be written as follows:

$$\mathbb{P}_H = \Pr\left(H = \bigcap_{j=1}^{m} H_j\right) = \prod_{j=1}^{m} \Pr\left(H_j \middle| H_{j-1}\right).$$

Each subsequent subset is chosen so that the probability of the conditional event $\Pr\left(H_j \middle| H_{j-1}\right)$ would be large enough. As a result, a small probability is represented as a product of relatively large probabilities [23–25]. One of the ways to form $\left(H_j \middle| H_{j-1}\right)$ is to select a given share p_0 of the most significant events $X_b^j = \left(\boldsymbol{x}_1^j, \boldsymbol{x}_2^j, \ldots, \boldsymbol{x}_k^j\right)$, where $k = p_0 N$; N– is the sample size. The index b characterizes the maximum level of the criterion function $\psi(\boldsymbol{x})$ for the set X_b^{j-1}. The significance of

events is determined by the value of $\psi(\boldsymbol{x})$ – the smaller $\psi(\boldsymbol{x})$, the greater the significance of $\boldsymbol{x}$ (for $\psi(\boldsymbol{x}_i, i = 1,\dots k) < 0$, the entire area X_b^j consists of "significant" (deficit states of the EPS)). Significance differentiation determines the mechanism of formation of the set X_b^j. The complete sample H_{j-1} obtained from the data of step $j-1$ is ordered in ascending order $\psi(\boldsymbol{x})$: $\psi(\boldsymbol{x}_1) \le \cdots \le \psi(\boldsymbol{x}_N)$. The first $p_0 N$ events ($p_0 \cdot 100$ percentile of the function ψ(x)) determine the set X_b^j and the maximum value of the criterion function corresponding to it, $b_j = max\big(\psi(\boldsymbol{x}), \boldsymbol{x} \in X_b^j\big)$, which, in turn, is the basis for the formation of a new set $H_j = \big\{\boldsymbol{x}: \psi(\boldsymbol{x}) < b_j\big\}$. In this case, the value p_0 can be considered as the probability of the conditional event $\Pr\big(H_j \big| H_{j-1}\big) = p_0$.

The presented stage-by-stage process of formation of a set of rare events is characterized by a positive value $b_j > 0$ at all intermediate stages. This means that the set X_b^j contains both failure events, $\psi(\boldsymbol{x}) < 0$, and non-failure events $\psi(\boldsymbol{x}) > 0$, that is, the principle of selecting events is to exclude less significant events and expand the region of more significant events , $b_j < b_{j-1}$. Initially, $b_0 = \infty$, which means that all system states generated on the basis of marginal PDF belong to the analysis zone. However, at subsequent stages, when generating the analyzed set of pseudo-random states of the system, a restriction is introduced: $\psi(\boldsymbol{x}) < b_{j-1}$. As a rule, this is realized by replacing the PDF parameters of random variables with some new calculated values.

At the last stage $b_j < 0$. This means that all events in the truncated set with p_0 are failures. But failures can also be events that do not fall into the region determined by the percentile, $b_j < \psi(\boldsymbol{x}) < 0$. Here, the conditional probability is determined according to the relation: $Pr(H_m|H_{m-1}) = N(\psi(\boldsymbol{x}) < 0)/N$, where $N(\psi(\boldsymbol{x}) < 0) -$ is the number of sample elements of size N satisfying the requirement $\psi(\boldsymbol{x}) < 0$.

Depending on the algorithm, the total sample size $N^{(j)}$ at intermediate stages may differ from the specified N, however, the selection principle is preserved - only the p_0 −part of the set of system states formed for analysis is accepted for further consideration.

A. Monocenter for the formation of an intermediate sample

A relatively small set of X_b^{j-1} is only the basis for the formation of H_j. Its elements are determined according to the type and parameters of the pseudo-random number generation distribution function in step $j-1$. At stage j , these should be different parameters, with a greater degree of identification of a rare event. One of the possible options for the formation of a new sample is proposed to generate pseudo-random numbers distributed according to a normal distribution with mean $\boldsymbol{\mu}^j = \mathbb{E}\big(\boldsymbol{x}_1^{j-1}, \boldsymbol{x}_2^{j-1}, \dots, \boldsymbol{x}_k^{j-1}\big)$ and variance $\boldsymbol{D}^j = \mathbb{D}\big(\boldsymbol{x}_1^{j-1}, \boldsymbol{x}_2^{j-1}, \dots, \boldsymbol{x}_k^{j-1}\big)$. At the first stage, as $\boldsymbol{\mu}^{(1)}, \boldsymbol{D}^{(1)}$ the mean and variances of the considered set of initial random variables are taken. Since the vector with parameters $(\boldsymbol{\mu}^j, \boldsymbol{D}^j)$ is the best representative of the region X^j, it is reasonable to consider this vector as the center of the region H_j.

The new set of pseudo-random vectors generated at stage j with center $(\boldsymbol{\mu}^j, \boldsymbol{D}^j)$, in the general case, contains points that do not belong to H_j according to the criterion $H_j \subset H_{j-1}$, that is, do not satisfy the condition $\psi(\boldsymbol{x}) < b_{j-1}$. The solution to this problem is either the addition of the resulting set to N elements that satisfy the condition $\psi(\boldsymbol{x}) < b_{j-1}$, or a simple removal of unsatisfactory system states. In the latter case, the sample is reduced from N to N_j elements, but all elements of the remaining set belong to H_j. It should be noted that the number of elements to be removed is, as a rule, relatively small, and the reduction of the analyzed set has little effect on the statistical estimates of the desired parameters (in particular, on the probability and mean of PD).

Calculations show that the choice of $(\boldsymbol{\mu}^j, \boldsymbol{D}^j)$ as the center of formation of the set F_j leads to some overestimation of the probability of a rare event in the region of very low probabilities (order: 10-6 in relation to 10-7). More accurate is the choice of the center at the point $\boldsymbol{x}^{*j}$, corresponding to the maximum value at stage $j-1$ criterion function, $\psi(\boldsymbol{x}^{*j}) = b_{j-1}$. Here, it is a priori assumed that at least half of the new generation of system states will not satisfy the

condition $\psi(\boldsymbol{x}) < b_j$, but this increases the probability of taking into account those states that do not fall into the statistical sample with the center $(\boldsymbol{\mu}^j, \boldsymbol{D}^j)$. An increase in the proportion of deleted events is inextricably linked with the requirement to increase the sample size N. The shift of the center $(\boldsymbol{\mu}^{*j} = \boldsymbol{x}^{*j-1}, \boldsymbol{D}^{*j})$ of the sample relative to the MO leads to the need to correct the variance $\boldsymbol{D}^{*j} = \boldsymbol{D}^j + \left|\boldsymbol{x}^{*j-1} - \boldsymbol{\mu}^j\right|^2$.

B. Gaussian distribution- centered method

The monocentric approach assumes the concentration of the sample around some center, for example, according to a normal (Gaussian) distribution with mean at the center of the sample. However, the principle of selecting a peripheral region according to the probability p_0 and the asymmetry of the region of a rare event (the greater the load, the greater the PD) casts doubt on the validity of the Gaussian distribution at intermediate stages. Logically more justified here is a probability distribution, unknown in type, but represented by a set of (reference) points defined at the previous stage $C_{j-1} = \left(\boldsymbol{x}_1^{j-1}, \boldsymbol{x}_2^{j-1}, \dots, \boldsymbol{x}_k^{j-1}\right)$. Statistical modeling of this distribution is possible by representing the set H_j as a union of subsets centered at the reference points $H_j = \cup H_{js}\left(\boldsymbol{x}_s^{j-1}\right)$. As the STD σ_j when forming the set H_{js} we can consider the maximum distance between the points of the set C_{j-1}: $\sigma_j = \max\left|\boldsymbol{x}_k^{j-1} - \boldsymbol{x}_l^{j-1}\right|, \left(\boldsymbol{x}_k^{j-1}, \boldsymbol{x}_l^{j-1}\right) \in C_{j-1}$. This makes it possible to ensure the intersection of sets of points formed in multidimensional spheres with centers at points $\boldsymbol{x}_k^{j-1} \in C_{j-1}$, and therefore more fully and uniformly take into account the region C_{j-1}.

C. Adaptive Sampling Algorithm

The main problem of methods of nested subspaces of states is the dependence of the resulting data on the mechanism of formation of intermediate sets. In this case, situations are possible when the subsequent subspace practically does not change the criterion threshold $b_k \approx b_{k-1}$, which leads to the lack of convergence of the computational process for a given number of iterations. In this case, the resulting probability of a rare event becomes arbitrarily small. Hence, the main directions of research in this area are aimed at increasing the robustness of the methods. There are a fairly large number of proposals for the formation of intermediate sets [22-25]. Along with the monocentric methods described above, we proposed and tested the adaptive sampling method, the essence of which is the multiple adaptive correction of the parameters (mean and variance) of the distribution of the intermediate sample. Its steps are as follows:

1. Initialization: $k = 1;\ N_c = p_0 N;\ N_s = 1/p_0;\ \lambda = 0.6$
2. Generation according to the standard normal (Gaussian) distribution of N pseudo-random numbers: $U = \{\boldsymbol{u_1}, \dots, \boldsymbol{u_N}\}$.
3. Transformation of the set U into a matrix $X = \{\boldsymbol{x_1}, \dots, \boldsymbol{x_N}\}$ of named random variables (load, generation, state of the system elements), according to their marginal distribution laws $\boldsymbol{x}_i = \varphi_i(\boldsymbol{u}_i)$.
4. Definition of the vector of criterion functions $\boldsymbol{\psi} = \{\psi_i(\boldsymbol{x_i})\}$ and its sorting in ascending order of the function: $\psi_i(\boldsymbol{x_i}) \geq \psi_{i-1}(\boldsymbol{x_{i-1}})$
5. Determination on the set $\boldsymbol{\psi}$ of the percentile $(p_0 \cdot 100)$, the parameter $b_k \geq 0$ corresponding to it, and the set of named significant states of the system $H_{xk} = \{\boldsymbol{x_i}: \psi_i(\boldsymbol{x_i}) < b_k\}$.
6. Generation of a new sample of N random states of the system based on the reference states of the system.
7. Items 4-6 are repeated until $b_k \geq 0$. In this case, the conditional probability $Pr(H_k|H_{k-1}) = p_0$. At the last step m, for $b_m < 0$ the number N_m of elements is determined that satisfy the condition $\psi_{N_m}\left(\boldsymbol{x}_{N_m}\right) \leq 0,\ \psi_{N_m+1}\left(\boldsymbol{x}_{N_m+1}\right) > 0$. The resulting probability of a rare event is $\Pr(H) = p_0^{m-1} N_m / N$.
8. Determination of other analyzed probabilistic indicators (mean of PD, etc.).

IV. Cross entropy method

The cross-entropy method (CEM) [26-28] is based on replacing the real function $f(\boldsymbol{x})$ of the distribution density of the analyzed multivariate random variable, according to which the MCM is sampled, by some auxiliary PDF $q(\boldsymbol{x})$, which shifts the scope of analysis into the region of interest of a rare event. This approach is widely known in the calculation of integrals of complex functions by statistical methods, where the integrand is multiplied and divided by some PDF, completely defined on the integration interval under consideration:

$$\int_a^b f(x)dx = \int_a^b \frac{f(x)}{q(x)} q(x)dx = \mathbb{E}_q\big(W(x)\big) \cong \frac{1}{N}\sum_{i=1}^{N} W(x_i),$$

where $W(x) = f(x)/q(x)$. In this case, the problem of choosing an auxiliary function arises - it is necessary that the calculation procedure based on it be not only adequate in direction (shift of the analyzed area towards a rare event), but also efficient in terms of speed and convergence. In the presence of CDF, the probability of power shortage (2) can be represented as the mathematical expectation of the weighted indicator function $J(\boldsymbol{x})W(\boldsymbol{x})$, defined in a multidimensional space of random variables with a density function $q(\boldsymbol{x})$:

$$\mathbb{P}_H = \int J(\boldsymbol{x}) \frac{f(\boldsymbol{x})}{q(\boldsymbol{x})} q(\boldsymbol{x})d\boldsymbol{x} = \int J(\boldsymbol{x})W(\boldsymbol{x})q(\boldsymbol{x})dx = \mathbb{E}_q[J(\boldsymbol{x})W(\boldsymbol{x})], \quad (3)$$

where $W(\boldsymbol{x}) = f(\boldsymbol{x})/q(\boldsymbol{x})$ – is the weight function determined by the initial $f(\boldsymbol{x})$ and auxiliary $q(\boldsymbol{x})$ distribution densities. When choosing the most efficient function $q^*(\boldsymbol{x})$ in the class $\{q(\boldsymbol{x})\}$, the minimum of the variance $\mathbb{P}_H$ can serve as an optimization criterion:

$$\min_q \mathbb{V}_q\big[J(\boldsymbol{x})W\big(\boldsymbol{x}; q(\boldsymbol{x})\big)\big].$$

Theoretically, the best function leading to zero variance of the desired probability estimate $\mathbb{P}_H$ is the function [28]:

$$q^*(\boldsymbol{x}) = \frac{J(\boldsymbol{x})f(\boldsymbol{x})}{\int J(\boldsymbol{x})f(\boldsymbol{x})d\boldsymbol{x}} = \frac{J(\boldsymbol{x})f(\boldsymbol{x})}{\mathbb{P}_H}. \quad (4)$$

Since the optimal CDF depends on the unknown values $\mathbb{P}_H$ and $J(\boldsymbol{x})$, a direct analytical determination of $q^*(\boldsymbol{x})$ is impossible. A sufficiently good approximation makes it possible to obtain a FEM based on the successive refinement of the parameters $\boldsymbol{v}$ of the multidimensional distribution density $q(\boldsymbol{x}; \boldsymbol{v})$and the procedure for determining the resulting value is represented as a Markov chain with the choice of parameters at each step. To estimate the parameters $\boldsymbol{v}$, this method uses the results of intermediate statistical tests. The parameter vector $\boldsymbol{v}$ is determined by minimizing the cross entropy (KL divergence) [28]. KL-divergence determines the measure of proximity of two arbitrary PDFs: the optimal PDF $q^*(\boldsymbol{x})$ and its current estimate $q(\boldsymbol{x}; \boldsymbol{v})$. In the proposed work, the function $q(\boldsymbol{x}; \boldsymbol{v})$is represented by the density of a multidimensional normal distribution with the mathematical expectation $\boldsymbol{\mu}_q$ and the matrix of correlation moments Σ_q. In this setting, the degree of optimality, in essence, determines the vector of parameters $\boldsymbol{v} = [\boldsymbol{\mu}_q; \Sigma_q]$).

CEM solves the problem of optimization iteratively by determining a series of intermediate distribution densities $\{q(\boldsymbol{x}; \boldsymbol{v}_k),\ k = 1, \ldots, NT\}$, which, as shown in Fig. 2, gradually approach the target density $q^*(\boldsymbol{x})$, representing region of existence of a rare event. At step k, the optimal PDF $q^*(\boldsymbol{x})$can be represented by the estimate $q(\boldsymbol{x}; \boldsymbol{v}^*_{k-1})$ with the optimal parameters $\boldsymbol{v}^*_{k-1}$ obtained at the previous step. In this case, $W(\boldsymbol{x}; \boldsymbol{v}^*_{k-1}) = f(\boldsymbol{x})/q(\boldsymbol{x}; \boldsymbol{v}^*_{k-1})$. The area H_k of the intermediate set of system states is determined by the threshold b_k:

$$H_k = \{\boldsymbol{x}: \psi_k(\boldsymbol{x}) < b_k\},$$
$$J_k(\boldsymbol{x}) = \begin{cases} 1, & \psi_k(\boldsymbol{x_k}) < b_k; \\ 0, & \psi_k(\boldsymbol{x_k}) \geq b_k. \end{cases} \quad (5)$$

The threshold b_k is calculated as a θ-quantile (for example, a decile obtained during statistical tests and sorted from the smallest to the largest values of the threshold function $\psi_i(\boldsymbol{x_i})$. In this case, the simulation is performed according to the distribution density $q(\boldsymbol{x}; \boldsymbol{v}_{k-1})$ with parameters

$\boldsymbol{v}_{k-1}$.

The vector $\boldsymbol{v}^*$ obtained as a result of solving the optimization problem at step $k-1$ is considered as a new value of the vector of parameters of the optimal PDF $q^*(\boldsymbol{x})$, $\boldsymbol{v}_k = \boldsymbol{v}^*$. Starting from the initial vector of parameters $\boldsymbol{v}_0$ that is assumed to be equal to the initial parameters of the probability distributions of the analyzed random variables, each subsequent vector $\boldsymbol{v}_k$ is determined by the solution of the optimization problem, approaching the optimal distribution density $q^*(\boldsymbol{x}) = \lim_{k\to\infty}(q(\boldsymbol{x};\boldsymbol{v}_k))$, which, in turn, is the best estimate of the auxiliary PDF optimal in the CEM. The procedure is repeated until b_k becomes negative, or at least $\theta \in [0.01, 0.1]$ trials are located in the desired region of rare events [28]. If the number of steps to reach the termination criterion of the iterative process is m, then the resulting probability of a rare event is:

$$\hat{\mathbb{P}}_F = \frac{1}{N}\sum_{i=1}^{N} J(\boldsymbol{x}_i)W(\boldsymbol{x}_i;\boldsymbol{v}_{m-1}). \quad (6)$$

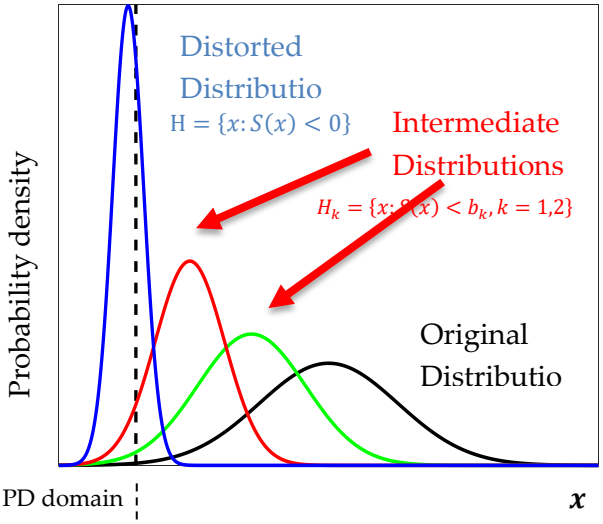

Figure 2: *Derivation of distorted probability distribution using CEM*

The classical CEM presented above requires an optimization procedure to be performed when determining the parameters $\boldsymbol{v}_k$ of the current PDF $q(\boldsymbol{x}, \boldsymbol{v}_k)$. In this case, $\boldsymbol{v}_k$ includes the mean $\boldsymbol{\mu}_k$ and the matrix of correlation moments Σ_k of those realizations x that belong to the area H_k, $\boldsymbol{x} \in H_k$. Expression (6) can be considered as the average value of the values $J_F(\boldsymbol{x}_i)$ with weights $W(\boldsymbol{x}_i)$. Since it refers to a probability, each component of the sum can be interpreted as the probability that the realization $\boldsymbol{x}_i$ belongs to the current region H_k of the rare event. Since the area $H_k$is defined by the indicator function $J_k(\boldsymbol{x}_i)$, then the mean of the available power of nodes in the area H_k is as follows:

$$\boldsymbol{\mu}_r^{(k)} = \left(\sum_{x_i \in H_k} \boldsymbol{x}_i W_i\right) / \left(\sum_{x_i \in H_k} W_i\right).$$

Matrix of correlation moments can be expressed as follows:

$$\Sigma_r^{(k)} = \left(\sum_{x_i \in H_k} W(\boldsymbol{x}_i)(\boldsymbol{x}_i - \boldsymbol{\mu}_r^{(k)})(\boldsymbol{x}_i - \boldsymbol{\mu}_r^{(k)})^T\right) / \left(\sum_{x_i \in H_k} W(\boldsymbol{x}_i)\right).$$

These parameters form the vector $\boldsymbol{v}_k$. With this approach, an optimization procedure is not required, which significantly reduces the duration of calculations without a significant decrease in the accuracy of the results.

V. Computational results

To compare the described procedures, calculations of the balance reliability indicators are performed for a five-node electrical circuit shown in Fig. 3, where in addition to the topology, load and generation expectations are presented. Standard deviations are taken equal to 10% of expectations. TLs, numbered in the order {1-5; 1-4; 4-5; 2-4; 2-3; 3-4}. TL have resistances respectively R = {10; 5; 5; 3; 1; 2}; X = {100; 50; 50; 50; 33; 10; 10}. The limiting capacity of all connections was taken equal to 500 MW.

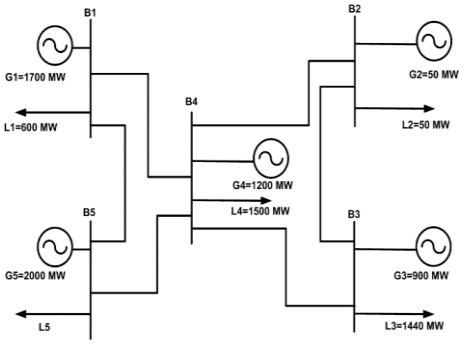

Figure 3: *Test circuit*

The calculation results are presented in Table I, where the methods are indicated: MCM – classical Monte Carlo method; Monocenter – subspace method with one point of formation of an intermediate subspace; Polycenter – the set H_{k-1}; is taken as reference points for the formation of the intermediate subspace H_k; SubSet– method of subspaces with changing parameters of probability distributions when forming an intermediate subspace; CEM– simplified cross-entropy method.

Table 1: *Computational results*

Methods	Pr, 10^{-5}	Pr, Cv	Pr_Eps, %	mD, 10^{-4}	mD, Cv	mD_Eps, %	t, sec
Convolution	2.23	0	0	1.21	0	0	0
MCM	2.23	0.07	0.1	1.21	0.10	0.1	19
Monocenter	2.24	0.6	0.5	1.48	0.90	22.1	0.44
Polycenter	2.29	0.65	2.9	1.52	0.92	25.7	0.19
SubSet	2.2	0.58	-1.4	1.41	0.92	16.7	3.21
CEM	2.3	0.54	3.3	1.24	0.03	2.5	0.12

The second column presents the calculated probability of the total power deficit of the EPS. It should be noted that all presented methods show an accuracy of a rare event acceptable for practical calculations - a maximum deviation of 3.3% (column Pr_Eps, %, CE method) with an exact probability value Pr = 2.23×10–5 can be considered insignificant for low probabilities. For comparison, it can be mentioned that with such probabilities, the widespread replacement of the binomial distribution with the Poisson distribution has a much larger error. The coefficient of variation presented in the Pr_Cv column shows that the spread of the resulting probabilities is relatively small for all methods.

The next 3 columns refer to the mean of PD. Here the spread (coefficient of variation (column mD, Cv)) of the resulting values is much greater. It should be noted that the relative error in calculating the mean of the PD exceeds 25%. Increasing accuracy is possible by increasing the volume of the intermediate sample. However, this leads to an increase in the duration of calculations (column t), which in the presented calculation procedures depends mainly on time-intensive transformations of probability distributions (uniform- Gaussian- individual (marginal)). If we assume that all random variables are described by the same normal distribution (with

different parameters), then this allows us to reduce the duration of calculations several times.

The results obtained allow us to recommend cross-entropy methods for practical use that provide the smallest scatter of the resulting indicators.

VI. Conclusion

The described methods for identifying rare events in the electric power industry refer to those events for which it is possible to determine a criterion function that changes its value depending on the distance from the desired events. Such functions (and events) in the electric power industry include power shortages, positive when there is a shortage and negative when there is a reserve of generating capacity. The main technology for identifying rare events is the use of Markov chains, where each subsequent event is determined on a set of events identified at the previous stage according to some criterion. Of the existing approaches to identifying rare events, two can be identified that are most suitable for technical systems: those based on the use of nested subspace technology and the transformation of distribution functions of random variables (including entropy methods). It can be noted that all the procedures considered make it possible to identify rare events with an accuracy acceptable for practical use. According to the robustness criterion, preference can be given to the adaptive sampling algorithm from the class of nested subspaces and the modified cross-entropy method from the class of distribution function transformations. The accuracy of the calculations largely depends on the settings of the methods, including the size of the test sample at the intermediate stages of forming significant events. In this case, the duration of calculations is determined not only by the sample size, but also by the need for probabilistic transformations of distribution functions.

References

[1] Abunima, H.; The, J.; Lai, C.; Jabir, H. A Systematic Review of Reliability Studies on Composite Power Systems: A Coherent Taxonomy Motivations, Open Challenges, Recommendations, and New Research Directions. Energies. –2018. –vol.11, 2417.

[2] Mohsen Firouzi, Abouzar Samimi, Abolfazl Salami, Reliability evaluation of a composite power system in the presence of renewable generations, Reliability Engineering & System Safety, Volume 222,2022,108396, ISSN 0951-8320,

[3] Oboskalov, V.; Abdel Menaem, A.; Valiev, R.; Mahnitko, A.; Varfolomejeva, R. Mathematical methods for probabilistic estimation of power shortage in concentrated electric power systems. Proceedings of the 10th International Scientific Symposium on Electrical Power Engineering, Elektroenergetika. –2019. – P. 128–132.

[4] Song, J.; Wei, P.; Valdebenito, M.; Signal, M. Active learning line sampling for rare event analysis. Mech. Syst. Signal. Process. 2020. – vol.147. 107113.

[5] Rubino, G.; Tuffin, B. Rare Event Simulation using Monte Carlo Methods. John Wiley & Sons, Inc, Hoboken, New Jersey. – 2009. – ISBN:9780470772690.

[6] Kroese, D.P.; Taimre, T.; Botev, Z.I. Handbook of Monte Carlo Methods. John Wiley & Sons, Inc, Hoboken, New Jersey. – 2011. – ISBN 9781118014967.

[7] Zhang, F.; Lu, Z.; Cui, L.; Song, S. Reliability sensitivity algorithm based on stratified importance sampling method for multiple failure modes systems. Chin. J. Aeronaut. – 2010. – vol.23. – P.660-669.

[8] Abdel Menaem, A.; Valiev, R.; Oboskalov, V.; Hassan, T.; Rezk, H.; Ibrahim, M. An efficient framework for adequacy evaluation through extraction of rare load curtailment events in composite power systems. Mathematics. –2021. –vol.8. – P.1-21

[9] Abdel Menaem, A.; Oboskalov, V.P. Comparing Three Methods for Solving Probabilistic Multi-Area Load Shedding Distribution. International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM 2020). –19-22 May 2020.

[10] Wang, Y.; Guo, C.; Wu, Q.; Dong, S. Adaptive sequential importance sampling technique for short-term composite power system adequacy evaluation. IET Gener. Transm. Distrib.–2014.– vol.8. – P. 730–741.

[11] González-Fernández, R.A.; Leite, A.M.; Silva, Da.; Resende, L.C.; Schilling, M.T. Composite systems reliability evaluation based on Monte Carlo simulation and cross-entropy methods. IEEE Trans. Power Syst. – 2013. – vol.28. – P.4598–4606.

[12] Lucarelli, G.; Yan, C.; Bie, Z.; Wang, C.; Ding, T.; Xiong, F. Three approaches to use Cross-Entropy in reliability evaluation of composite system with high penetration wind energy. 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC) –2016.– P. 227-231.

[13] Gonz´alez-Fern´andez, R.; Silva, A. Reliability assessment of time-dependent systems via sequential cross-entropy Monte Carlo simulation. IEEE Trans. Power Syst. –2011.– vol. 26. – P. 2381–2389.

[14] Silva, A.; Castro, J.; Billinton, R. Probabilistic assessment of spinning reserve via cross-entropy method considering renewable sources and transmission restrictions. IEEE Transactions on Power Systems –2018.–vol. 33. – P. 4574-4582.

[15] Ge, H.; Asgarpoor, S. Parallel Monte Carlo simulation for reliability and cost evaluation of equipment and systems. Electr. Power Syst. Res. – 2011.– vol.81.– P.347–356.

[16] Giuntoli, M.; Pelacchi, P.; Poli, D. Parallel computing of sequential Monte Carlo techniques for reliable operation of Smart Grids. Proc. - EUROCON 2015, doi:10.1109/EUROCON.2015.7313707.

[17] Martinez, J.A.; Guerra, G. A parallel Monte Carlo method for optimum allocation of distributed generation. IEEE Trans. Power Syst. 2014, 29, 2926–2933.

[18] Leite da Silva, A.M.; de Resende, L.C.; da Fonseca Manso, L.A.; Miranda, V. Composite reliability assessment based on Monte Carlo simulation and artificial neural networks. IEEE Trans. Power Syst. 2007, doi:10.1109/TPWRS.2007.901302.

[19] Urgun, D.; Singh, C.A hybrid Monte Carlo Simulation and Multi Label Classification Method for Composite System Reliability Evaluation. IEEE Trans. Power Syst. –2019. –vol.287. – P.8535–8545.

[20] Pindoriya, N.M.; Jirutitijaroen, P.; Srinivasan, D.; Singh, C. Composite reliability evaluation using Monte Carlo simulation and least squares support vector classifier. IEEE Trans. Power Syst. 2011. – vol.211. –P.6048–6056.

[21] Dheybi Cervan, Alberto M. Coronado, Jaime E. Luyo, Cluster-based stratified sampling for fast reliability evaluation of composite power systems based on sequential Monte Carlo simulation, International Journal of Electrical Power & Energy Systems,Volume 147,2023,108813.

[22] Hua, B.; Bie, Z.; Au, S.K.; Li, W.; Wang, X. Extracting rare failure events in composite system reliability evaluation via subset simulation. IEEE Trans. Power Syst. – 2015. –vol.30. – P.753–762.

[23] Y. Zhao, Y. Han, Y. Liu, K. Xie, W. Li and J. Yu, "Cross-Entropy-Based Composite System Reliability Evaluation Using Subset Simulation and Minimum Computational Burden Criterion," in IEEE Transactions on Power Systems, vol. 36, no. 6, pp. 5198-5209, Nov. 2021, doi: 10.1109/TPWRS.2021.3073478.

[24] Zhao Y. Composite Power System Reliability Evaluation Based on Enhanced Sequential Cross-Entropy Monte Carlo Simulation/ Y. Zhao, Y. Tang, W. Li, J.Y. Power// Proceedings of the IEEE Transactions on Power Systems. –2019. –vol. 34. –P. 3891–3901.

[25] Cai J, Xu Q, Cao M, Yang B. A novel importance sampling method of power system reliability assessment considering multi-state units and correlation between wind speed and load. Int J Electr Power Energy Syst 2019; 109:217 – 26. https://doi.org/ 10.1016/j.ijepes.2019.02.019.

[26] Tomasson, E.; Soder, L. Improved importance sampling for reliability evaluation of composite power systems. IEEE Trans. Power Syst. – 2017. –vol. 32. – P. 2426–2434.

[27] Biondini G. An introduction to rare event simulation and importance sampling. Big data analytics. Handbook of statistics, vol. 33. Elsevier; 2015. p. 29–68.**Ошибка! Источник ссылки не найден.Ошибка! Источник ссылки не найден.Ошибка! Источник ссылки не найден.**

[28] Kroese DP, Rubinstein RY, Glynn PW. Chapter 2 – The cross-entropy method for estimation. Handbook of statistics: machine learning: theory and applications. Handbook of Statistics, vol. 31. Elsevier; 2013. p. 19–34.