Cross-entropy based rare-event simulation for the identification of dangerous events in power systems

Florence Belmudes, Damien Ernst, *Member, IEEE,* and Louis Wehenkel, *Member, IEEE*

***Abstract*— We propose in this paper a novel approach for identifying rare events that may endanger power system integrity. This approach is inspired by the rare-event simulation literature and, in particular, by the cross-entropy (CE) method for rare-event simulation. We propose a general framework for exploiting the CE method in the context of power system reliability evaluation, when a severity function defined on the set of possible events is available. The approach is illustrated on the IEEE 30 bus test system when instability mechanisms related to static voltage security are considered.**

***Index Terms*— Power system security, identification of dangerous events, cross-entropy method.**

I. Introduction

Blackouts in power systems are rare, but they have such tremendous societal consequences that decreasing their probability of occurrence is of paramount importance. Common practice consists of designing and operating power systems in such a way that they are able to cope with a set of events such as the sudden loss of transmission elements, short-circuits, variations in the demand and generation patterns, etc. Typically, the (combinations of) events explicitly covered in security and reliability studies are those that are a priori supposed to have a rather high probability of occurrence, the rationale behind this choice being twofold. First, by focusing on a subset of events one can sufficiently reduce the complexity of the reliability/security evaluation procedure to make it practically feasible for large power systems. Second, it is assumed that ensuring the security of the system with respect to the chosen events also reduces its vulnerability with respect to the multitude of events not explicitly taken into account, and hence increases the system reliability. The increasing rates of large blackouts observed over the recent years may however question whether such an approach for assessing the security of a power system is still appropriate.

Within this context, the goal of our research is to propose a framework to identify in a “computationally feasible way” the events that should be considered by the TSOs in their reliability/security studies and, as a byproduct, to propose techniques for computing reliability indices based on a sample of reasonable size of such events. Before developing further the rationale behind this framework, let us remind here the causes of two major system-wide disturbances that have happened recently. One is the August 14th, 2003 blackout that has plunged the North-East part of the USA in the dark for several hours and the other is the November 4th, 2006 event which has affected West Germany as well as some areas in France and Belgium and other parts of the European interconnection. As for the former blackout, the sequence of events having lead to the collapse of the system is the following. First, some highly loaded transmission lines tripped due to some poor vegetation management. Second, this information has not been correctly displayed on the screen of an operator, who hence did not take appropriate actions. And then, the inadequate response of other dispatchers to this dangerous situation lead to a cascade of trippings eventually causing a major blackout [11]. The root cause behind the European quasi-blackout was from another puzzling nature. In order to allow a boat to cross a river, a transmission line had to be switched off. A system operator gave the green light for carrying this manoeuvre though the load on the power network at the time of the manoeuvre was significantly increased with respect to the load at the time at which the stability study took place. When the line was disconnected, several other nearby lines tripped on overload and initiated a cascade splitting the European interconnection into three islands, whose energy balance was restored on the edge thanks to fast load and generation shedding [1]. These two scenarios have really an “unfortunate nature” since their probability of inception, as it would have been estimated beforehand, is indeed extremely low. This presumably explains why the involved TSOs did not anticipate them in their security studies.

Because of the complexity of power systems, it is extremely difficult - and as a matter of fact impossible - to screen over all the plausible events to identify the dangerous ones. Thus, our main motivation in this paper is to propose an approach which could help identify highly severe (combinations of) events (i.e. those that would lead to blackouts or quasi-blackouts) without running a security analysis for every event. Our approach does not, as it is currently the case, focus on high-probability events as candidate interesting ones for security studies. Rather it aims to identify within an a priori defined extremely large set of events, those which have the potential to endanger the integrity of the system. The proposed approach is based on a reformulation of the issue of assessing the reliability of a power system as a rare-event simulation problem, which is then solved by using cross-entropy based techniques [15].

The rest of this paper is organised as follows. In the next section the cross-entropy (CE) method for rare-event simulation is introduced (Section II). Then, Section III discusses the application of this method for sampling dangerous events and estimating reliability indices. An illustrative experimental application of these concepts to a problem of static voltage security is reported in Section IV. Section V discusses the proposed approach with respect to existing work in reliability evaluation of power systems and conclusions are drawn in Section VI.

II. CE METHOD FOR RARE-EVENT SIMULATIONS

We start this section by presenting the main concepts related to the CE method for rare-event simulation. Afterwards, we explain how to build from these concepts some computationally efficient algorithms. Finally, we present a fully specified algorithm based on the CE method when the set of events is a subset of $\mathbb{R}^n$. The material of this section is largely borrowed from [15] to which we refer the reader for a complement of information.

A. Importance sampling for rare-event simulations

Let X be a random variable taking its value in some space $\mathcal{X}$ with a probability density function (pdf) $f(\cdot)$, let $S(\cdot)$ be a real-valued function defined on $\mathcal{X}$ and γ be a real number. In the rare-event simulation context, one needs to estimate the probability of occurrence l of an event $\{S(X) \geq \gamma\}$, i.e. to estimate the expression $E_{X\sim f(\cdot)}\left[I_{\{S(X)\geq\gamma\}}\right]$.[1]

In rare-event simulation problems, this probability is extremely low, say smaller than 10^{-6}, and estimating l with enough accuracy by relying on a Crude Monte-Carlo (CMC) estimator

$$\hat{l} = \frac{1}{N}\sum_{j=1}^{N} I_{\{S(X_j)\geq\gamma\}} \tag{1}$$

requires to draw a considerably large sample $X_1, X_2, \ldots, X_N$ from $f(\cdot)$. For example, estimating l, with a sample of size N leads to a standard error $\sigma_{\hat{l}} = \sqrt{\frac{l(1-l)}{N}}$. Hence, a sample size of $N \simeq 10^{10}$ is required in order to estimate $l \simeq 10^{-6}$ with relative error of 1% (i.e. with a standard error of $0.01l$).

An alternative to CMC is based on importance sampling. With such an approach, a random sample $X_1, X_2, \ldots, X_N$ is drawn from an importance sampling pdf $g(\cdot)$ and the probability of occurrence of the event is estimated via the following estimator[2]

$$\hat{l} = \frac{1}{N}\sum_{j=1}^{N} I_{\{S(X_j)\geq\gamma\}}\frac{f(X_j)}{g(X_j)}\,. \tag{2}$$

In this context, the most effective way to estimate l would be to adopt the "ideal" importance sampling pdf

$$g^*(X) = \frac{I_{\{S(X)\geq\gamma\}}f(X)}{l}\,. \tag{3}$$

Indeed, since l is constant, using this "ideal" importance sampling pdf $g^*(\cdot)$ (3) would lead to an estimator (2) having a zero variance. Consequently, we would need to produce only a one element sample to determine l.

The obvious difficulty is that $g^*(\cdot)$ depends on the unknown parameter l.

The main idea of the CE method for rare event simulation is to find inside an a priori given set $\mathcal{G}$ of pdfs defined on $\mathcal{X}$, an element $g(\cdot)$ such that its distance to the "ideal" sampling distribution is minimal. A convenient measure of "distance" between two pdfs $a(\cdot)$ and $b(\cdot)$ on $\mathcal{X}$ is the Kullback-Leibler divergence, which is also termed the cross-entropy between $a(\cdot)$ and $b(\cdot)$. The Kullback-Leibler divergence, which is not a distance in the formal sense since it is for example not symmetric, is defined as follows:

$$\mathcal{D}(a,b) = E_{X\sim a(\cdot)}\left[\ln\frac{a(X)}{b(X)}\right] \tag{4}$$

The CE method reduces the problem of finding an appropriate importance sampling pdf to the following optimization problem:

$$\arg\min_{g\in\mathcal{G}} \mathcal{D}(g^*, g)\,. \tag{5}$$

One can show through simple mathematical derivations that solving (5) is equivalent to solve:

$$\arg\max_{g\in\mathcal{G}} E_{X\sim f(\cdot)}\left[I_{\{S(X)\geq\gamma\}}\ln g(X)\right] \tag{6}$$

which does not depend explicitly on l anymore.

If l is not too small, CE-based algorithms for rare-event simulations estimate a good solution of (6) by solving its stochastic counterpart:

$$\arg\max_{g\in\mathcal{G}}\sum_{j=1}^{M} I_{\{S(X_j)\geq\gamma\}}\ln g(X_j) \tag{7}$$

where the sample $X_1, X_2, \ldots, X_M$ is drawn according to $f(\cdot)$. When l is too small, say $l < 10^{-6}$, which is often the case in rare-event simulation, the value of M one has to adopt for having a "good" stochastic counterpart may be prohibitively high and some specific iterative techniques need to be adopted to solve (6). The use of these techniques is often equivalent to solving a sequence of rare event problems using the same pdf $f(\cdot)$ and function S but with increasing values of γ converging to the value of γ related to the original problem.

Under some specific assumptions on $\mathcal{X}$, $f(\cdot)$ and $\mathcal{G}$, it is possible to solve analytically the optimization problem (7). This property is often exploited in the CE context.

For example, let us suppose that $\mathcal{X}$ is $\mathbb{R}^n$ and let us denote by $Gauss_{\mathbb{R}^n}(\cdot, v)$, where $v = [\mu,\sigma] \in \mathbb{R}^n \times \mathbb{R}^n$, the n-dimensional (diagonal) Gaussian pdf

$$Gauss_{\mathbb{R}^n}(x,v) = \Pi_{i=1}^{n}\frac{1}{\sigma[i]\sqrt{2\pi}}\exp\left(-\frac{x[i]-\mu[i]}{2\sigma[i]^2}\right)\,, \tag{8}$$

where $x[i]$ is the ith component of the random variable X and $\sigma[i]$ ($\mu[i]$) is the standard deviation (mean) of the n-dimensional pdf alongside the ith direction.

Then, one can show that if $f(\cdot)$ is a n-dimensional Gaussian pdf and $\mathcal{G}$ is the set of all n-dimensional Gaussian pdfs, the solution $Gauss_{\mathbb{R}^n}(\cdot, v^*)$ of (7) can be computed analytically:

$$\mu[i] = \frac{\sum_{j=1}^{M} I_{\{S(X_j)\geq\gamma\}}X_j[i]}{\sum_{j=1}^{M} I_{\{S(X_j)\geq\gamma\}}}\,, \tag{9}$$

$$\sigma[i] = \sqrt{\frac{\sum_{j=1}^{M} I_{\{S(X_j)\geq\gamma\}}(X_j[i]-\mu[i])^2}{\sum_{j=1}^{M} I_{\{S(X_j)\geq\gamma\}}}}\,. \tag{10}$$

[1] The function $I_{\{logical\_expression\}}$ is defined by $I_{\{logical\_expression\}} = 1$ if $logical\_expression = true$ and 0 otherwise. If $\mathcal{X}$ is finite, the expression $E_{X\sim f(\cdot)}\left[I_{\{S(X)\geq\gamma\}}\right]$ can be written equivalently as $\sum_{X\in\mathcal{X}} I_{\{S(X)\geq\gamma\}}f(X)$.

[2] assuming that $g(X) \neq 0$ whenever $I_{\{S(X)\geq\gamma\}}f(X) \neq 0$

Problem definition: A function $S : \mathcal{X} \to \mathbb{R}$ with $\mathcal{X} \subset \mathbb{R}^n$, a random variable $X \in \mathcal{X}$ taking its value in $\mathcal{X}$ with pdf $f(\cdot)$ and the value of γ.
Algorithm parameters: $v_1 = [\mu_1, \sigma_1]^T$, C, ϱ, N .

Output: An estimation of the (small) probability $E_{X\sim f(\cdot)}\left[I_{\{S(X)\geq\gamma\}}\right]$ and a pdf $Gauss_{\mathcal{X}}(\cdot, v)$ giving preference to the events x such that $\{S(x) \geq \gamma\}$.

Algorithm:

Step 1. Set t equal to 1. Set $nbElite$ equal to the largest integer inferior or equal to $\varrho \times C \times n$.
If $nbElite < 1$ then set $nbElite$ to 1.
Step 2. Set $\mathcal{X}_t$ equal to an empty set and r_t to an empty vector.
Step 3. Draw independently $C \times n$ elements according to the pdf $g_t(\cdot) = Gauss_{\mathcal{X}}(\cdot, v_t)$ and store them in $\mathcal{X}_t$.
Step 4. For every element $x \in \mathcal{X}_t$, compute $S(x)$ and add this value at the end of the vector r_t.
Step 5. Order the vector r_t in decreasing order and set $\hat{\gamma}_t = \min(\gamma, r_t[nbElite])$.

Step 6. Set $\mu_{t+1}[i] = \frac{\sum_{x\in\mathcal{X}_t} I_{\{S(x)\geq\hat{\gamma}_t\}} x[i] f(x)/g_t(x)}{\sum_{x\in\mathcal{X}_t} I_{\{S(x)\geq\hat{\gamma}_t\}} f(x)/g_t(x)}$ and $\sigma_{t+1} = \sqrt{\frac{\sum_{x\in\mathcal{X}_t} I_{\{S(x)\geq\hat{\gamma}_t\}} (x[i]-\mu_{t+1}[i])^2 f(x)/g_t(x)}{\sum_{x\in\mathcal{X}_t} I_{\{S(x)\geq\hat{\gamma}_t\}} f(x)/g_t(x)}}$ for $i = 1, 2, \ldots, n$ and set $v_{t+1} = [\mu_{t+1}, \sigma_{t+1}]$.

Step 7. If $\hat{\gamma}_t = \gamma$ then estimate l using the estimator $\frac{1}{N}\sum_{j=1}^{N} I_{\{S(X_j)\geq\gamma\}} \frac{f(X_j)}{Gauss_{\mathcal{X}}(X_j, v_{t+1})}$ where the samples X_j are drawn from $Gauss_{\mathcal{X}}(\cdot, v_{t+1})$ and return both $\hat{l}$ and the pdf $Gauss_{\mathcal{X}}(\cdot, v_{t+1})$. Otherwise, $t \leftarrow t + 1$ and go to **Step 2.**

Fig. 1. A fully specified cross-entropy based rare-event algorithm when the event space is a subset of $\mathbb{R}^n$.

B. Iterative CE based rare-event simulation algorithm

As mentioned in the previous subsection, when l is too small, one needs to draw a prohibitively high number of samples to obtain a "good" stochastic counterpart (7) of (6). This originates from the fact that to have a "good" stochastic counterpart, a sufficient number of samples X_j for which $\{S(X_j) \geq \gamma\}$ needs to be drawn. Iterative algorithms for solving accurately this stochastic counterpart are therefore used.

The rationale behind these algorithms is the following. First, let us observe that even if the pdf used to draw the sample used for building the stochastic counterpart was not drawn according to $f(\cdot)$ but well another pdf, called $h(\cdot)$, the stochastic counterpart could still have the same "meaning" provided that every term of the sum is weighted by a factor $\frac{f(X_j)}{h(X_j)}$. Therefore, if one can identify a pdf $h(X_j)$ for which the probability of occurrence of the event $\{S(X) \geq \gamma\}$ is not too small, it is very likely that the number of samples that will have to be drawn to obtain a good stochastic counterpart will have to be relatively small.

To identify such a pdf, it is of common practice to solve a sequence of rare-event problems differing only by the values of γ used. The first iteration of this sequence consists of a "rare-event" problem based on a small enough value of γ [3] so as to require only drawing a reasonable number of samples with f for having a "good" stochastic counterpart[4]. The pdf computed by solving this stochastic counterpart is in general more likely than f to generate events associated with high values of $S(\cdot)$. It is then used to generate the sample for solving the stochastic counterpart of the second rare-event problem, which differs only from the first one by a larger value of γ. This larger value of γ is itself defined from this same pdf to guarantee that by proceeding like this, not too many samples have to be drawn to solve accurately the stochastic counterpart at the second iteration. The algorithm proceeds similarly over the next iterations and stops when the rare-event problem solved is identical to the original one.

[3]The smaller the value of γ is, the higher the probability of the event $\{S(X_j) \geq \gamma\}$ is.

[4]Actually, it is not required to use f to draw the sample at the first iteration, provided that the stochastic counterpart is corrected appropriately.

C. A fully specified algorithm

Figure 1 gives the tabular version of a fully specified CE-based algorithm for rare-event simulation problems for which the set of events $\mathcal{X}$ is a bounded subset of $\mathbb{R}^n$. The rationale behind the algorithm is based on the iterative scheme described in previous subsection and is particularized to the case where $\mathcal{G}$, the set of pdfs in which one looks for an element which "stands" the closest to the "ideal" sampling distribution, is the set of Gaussian distributions "truncated" to values falling in $\mathcal{X}$. Similarly to the notation adopted for denoting non-truncated Gaussian distributions, the symbol $Gauss_{\mathcal{X}}(\cdot, v)$ is chosen to refer to truncated ones. The value of these truncated pdfs at $x \in \mathcal{X}$ is $\frac{Gauss_{\mathbb{R}^n}(x,v)}{E_{X\sim Gauss_{\mathbb{R}^n}(\cdot,v)}[I_{\{X\in\mathcal{X}\}}]}$.

At every iteration t, the algorithm proceeds as follows. First, it uses the pdf $Gauss_{\mathcal{X}}(\cdot, v_t)$ computed at the previous iteration and draws from this distribution a sample named $\mathcal{X}_t$. From this sample, it computes a value γ_t which is such that only a small fraction ϱ (ϱ is a parameter of the algorithm) of the elements of $x \in X_t$ lead to a value $S(x)$ larger or equal to γ_t. If the value of γ_t so computed turns out to be larger than γ, then it is replaced by γ. This value γ_t is then used together with $Gauss_{\mathcal{X}}(\cdot, v_t)$ to define the rare-event problem to be solved at iteration t. By defining the rare-event problem in this way, it is likely that the probability of the event

$\{S(X) \geq \gamma_t\}$ with $X \sim Gauss_{\mathcal{X}}(\cdot, v_t)$ is not too small. The stochastic counterpart of the optimization problem (see (7) and (6)) can therefore be defined by using a sample which is not too large. The algorithm described in Figure 1 uses the already drawn sample $\mathcal{X}_t$ to build this stochastic counterpart (where every term is weighted by $\frac{f(x)}{Gauss_{\mathcal{X}}(x,v_t)}$). An analytical solution, which is an approximation of the solution of the stochastic counterpart, is then used to compute the parameter $v_{t+1} = [\mu_{t+1}, \sigma_{t+1}]$ of the pdf used at the next iteration.

The algorithm stops when γ_t is equal to γ. Before stopping, the algorithm computes an estimate of l by exploiting the importance sampling estimator (2) with the importance sampling distribution $Gauss_{\mathcal{X}}(\cdot, v_{t+1})$. It returns both this estimate of l and $Gauss_{\mathcal{X}}(\cdot, v_{t+1})$, which is usually a pdf which gives a high preference to the events x such that $\{S(x) \geq \gamma\}$ is true.

D. Algorithm parameters

Before closing on the description of this CE algorithm, it is worth elaborating on the role of its parameters.

The parameter C determines the size of the samples $\mathcal{X}_t$ in a way that $|\mathcal{X}_t| = C \times n$. The rationale behind adopting a sample $\mathcal{X}_t$ whose cardinality is proportional to the dimension n of the event space $\mathcal{X}$ is that usually, the larger the dimension of event space space is, the larger the sample $\mathcal{X}_t$ has to be for the algorithm to behave well. A default value for this parameter C equal to 10 is adopted in this paper.

As explained before, the parameter ϱ determines the percentage of elements x of $\mathcal{X}_t$ which lead to true events $\{S(x) \geq \gamma_t\}$. A default value of 0.1 is chosen for ϱ.

Finally, the choice of the (diagonal) Gaussian family of importance sampling distributions is essentially guided by practical considerations, namely the fact that it is easy to draw samples from such distributions and the fact that it leads to closed-form solutions of the minimization problem (7).

III. CE METHOD FOR POWER SYSTEM ANALYSIS

The cross-entropy based algorithm for rare-event simulation described in the previous section determines, from a probability distribution $f(\cdot)$ defined over the set of events $\mathcal{X}$, a value γ and a real-valued function $S(x)$ defined over $\mathcal{X}$, the following information: a pdf biased towards events $x \in \mathcal{X}$ such that $S(x) \geqslant \gamma$, and an estimate of the probability that $S(x) \geqslant \gamma$.

In this paper, we suggest using this algorithm for two different types of problems met by power system engineers, namely the identification of dangerous events and the computation of reliability indices.

A. Identification of dangerous events

When studying the stability of a power system, identification of the dangerous events is often seen as the identification of the plausible (non-zero probability of inception) pairs "operating conditions-perturbation" that may endanger the stability of the system. For this task, specific stability evaluation tools able to assess whether a particular contingency will indeed lead or not to unacceptable operating conditions are used. Since the plausible pairs "operating conditions-perturbation" may be extremely numerous, screening each pair by the stability evaluation tool is generally not possible. Now, let us suppose that $S(x)$ is a function equal to 1 if x drives the system to unacceptable conditions and to 0 otherwise, and let us suppose that we use the CE method with a value of $\gamma = 1$, while $f(\cdot)$ is any pdf defined on the event space. By running the cross-entropy based algorithm on this problem with a family of importance sampling distributions $\mathcal{G}$, one can find in a computationally efficient way a pdf in $\mathcal{G}$ which concentrates on the dangerous events. By drawing samples from this latter pdf, one could therefore identify dangerous events with a high probability and therefore alleviate the computational burdens associated with running a security analysis on all the events or by sampling them from $f(\cdot)$.

B. Computation of reliability indices

The computation of a reliability index can be formulated in the following way. The power system can operate in different conditions (e.g., different values for the load, availability or not of some generation units) and perturbations can happen to the system (e.g., short-circuits). The probability that the system will operate in a specific condition as well as the probability of an event (that can be correlated to the operating conditions) are also given (the known or assumed to be known a priori). A tool to assess whether a given pair "operating conditions-perturbation" leads to unacceptable conditions or not is also available. From there, it is asked to compute the probability that the system may be driven to unacceptable conditions.

If the set of pairs "operating conditions-perturbation" were relatively small, this problem could be solved in a straightforward way by enumeration. However, it is generally not the case and efficient numerical tools have to be designed. By for example defining $S(x)$ as being a function equal to 1 if x drives the system to unacceptable conditions and to 0 otherwise, and using a value of $\gamma = 1$, it is straightforward to see that the problem of computation of the reliability index is equivalent to solving the rare-event problem $\{S(x) \geqslant \gamma\}$, which in turn can be solved efficiently by the cross-entropy based algorithms described previously.

C. Discussion

While this cross-entropy based framework for identifying dangerous events or computing reliability indices is certainly conceptually attractive, its successful application depends however on several factors that may be critical.

First, good models of the power system and appropriate security analysis tools have to be available. In the case of estimation of reliability indices, these models must also contain accurate probability distributions over the events, which may be particularly difficult to obtain.

Second, while the function $S(x)$ could take in principle binary values, these cross-entropy based algorithms usually behave better if, given any two non-dangerous events x and x', the difference between $S(x)$ and $S(x')$ gives some information about which element is the closest to the stability boundary, in a way that if x is the closest to this boundary, the difference is positive and negative otherwise. Let β the maximal value that $S(x)$ can take on the set of non-dangerous events and let α be greater than β. Then, by setting $S(x)$ equal to α when x is a dangerous event and γ to α, this information of distance with respect to the stability boundary greatly helps to cross-entropy algorithm to "drive" through the iterations the importance sampling pdfs to the dangerous events. Practically, rather than computing in the event space the stability boundary of the system and defining a topology in $\mathcal{X}$ for being able to compute distances, $S(x)$ is chosen as being an image of the severity of the event x. The more x is severe, the higher $S(x)$ is. This is the reason why we have chosen to name in this power system context the function $S(x)$, the severity function.

As last important technical point that we mention, there is the choice of $f(\cdot)$ that needs to be made when using the cross-entropy based approach for identifying dangerous events. Simulations have shown that this choice may greatly influence the convergence properties of the algorithm. However, we found out that a good choice for $f(\cdot)$ was to choose a pdf that was giving a good coverage over the event space. Another appropriate strategy, which is often used when cross-entropy methods are applied to combinatorial optimization, is to set all the terms $\frac{f(\cdot)}{Gauss_{\mathcal{X}}(\cdot,v_t)}$ equal to 1 at **Step 6** of the algorithm.

IV. Experiment for static voltage security

We have chosen to experiment the framework described in the previous section on the IEEE 30 bus system depicted at Figure 2, which has been vastly used as benchmark test system in the literature (see, e.g., [16]).

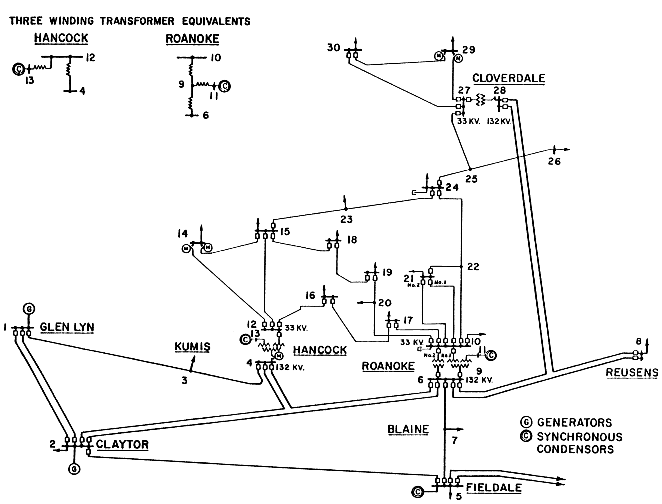

Fig. 2. One-line diagram of the IEEE 30 bus system.

The problem tackled here is the identification of dangerous events for static voltage security and more specifically for static loadability[5]. In this system, an event is defined as being an homothetic increase/decrease of the load with respect to the base case and we consider that such an event is acceptable if the corresponding demand level can be served by the available active and reactive generation capacity while respecting voltage constraints[6]. Computationally, we can thus determine whether an event is feasible by solving an active/reactive optimal power flow (OPF) problem (see below). The problem of identifying the dangerous events thus reduces to finding the set of homothetic variations of the load for which his OPF problem is not feasible anymore. By assuming that the coefficient x multiplying the load in the base case can vary in the interval $[0.25; 4]$, the problem of identifying the dangerous events consists therefore of finding the values of x in $\mathcal{X} = [0.25; 4]$ which lead to unfeasible OPF problems. By assuming that the lower x, the easier the system can operate within these constraints and by using a dichotomy approach to compute the maximum load the system can sustain, we found out that the set of dangerous events was $[2.1067; 4]$.

To apply our proposed approach for identifying dangerous events, a severity function needs to be adopted. The severity function chosen here is related to the algorithmic behaviour of an OPF solver which optimises the generation dispatch and the generator voltages to minimize generation cost. The OPF is based on an interior point algorithm (see, e.g. [6]) and starts iterating in the search space from a point corresponding to an optimised power system for the base case. The value of the severity function for a particular event x is chosen equal to the number of iterations to convergence, if convergence indeed occurs. The rationale behind choosing such a function $S(x)$ is related to the fact that one may reasonably assume that the higher the number of iterations is, the more the system is stressed and the closer it operates from its stability limits.

If convergence does not occur, it is assumed that no solution to the problem exists (that is that x is a dangerous event) and $S(x)$ is set equal to a large value, chosen equal to 1000. Note that a convergence case can never reach such a high value since the OPF automatically declares divergence after 500 iterations. Henceforth, an appropriate value of γ is any value in $]500; 1000]$. By running the algorithm described in Figure 1 (as previously indicated, we use $C = 10$ and $\rho = 0.1$ while $n = 1$) when choosing the initial pdf $f(\cdot) = Gauss_{[0.25;4]}(\cdot, [1, 0.5])$, we obtain the sequence of pdfs drawn on Figure 3. At the fifth iteration of the algorithm, the stopping criterion is reached ($\gamma_5 = \gamma$). As one can see, the pdf indeed evolves to give strong preference to the dangerous events.

When running the algorithm, the correction terms $\frac{f(X)}{Gauss_{\mathcal{X}}(X,v_t)}$ intervening in **Step 6** have been set equal to 1 since we were only interested in identifying pdfs giving preference to dangerous events, rather than computing reliability indices. We may however wonder whether the different distributions $Gauss_{\mathcal{X}}(\cdot, v_1)$, ..., $Gauss_{\mathcal{X}}(\cdot, v_5)$ are good importance sampling distributions when they are used to estimate the probability of the rare-event problem defined by the same $S(x)$ and γ as before but with a pdf $f'(\cdot)$ being a limited Gaussian distribution with much lower variance than

[5]In the previous section, the events we were referring to when discussing the application of cross-entropy based methods to the identification of dangerous events were described by a pair "operating conditions-perturbation" while here they can be described by operating conditions only. This can be seen as a particular case for which the set of possible perturbations does not contain any element.

[6]The set of events has been chosen relatively simple to allow an identification of the dangerous events using some common engineering practice.

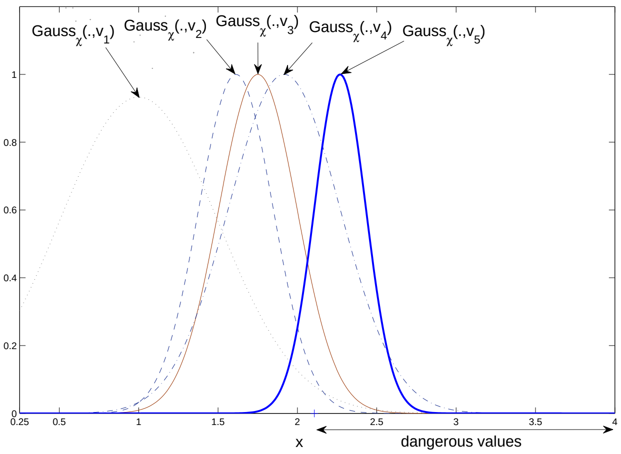

Fig. 3. A typical run of the algorithm.

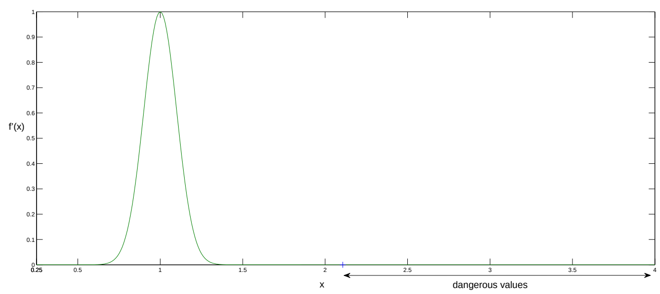

Fig. 4. Probability density function of the rare-event problem.

f and equal mean (see Figure 4). This leads to a rare-event probability estimation problem since under f' the probability that x is higher than 2.1067 is extremely low ($6.7468.10^{-27}$).

To answer this question, we have computed (using 1000 replications of the experiment) for these different importance sampling pdfs the standard deviation of the estimator given by Eqn (2) with N chosen equal to 20. As shown in Table I, the algorithm produces a sequence of pdfs which, when used for importance sampling, leads to a sequence of estimators with increasing accuracy.

V. Related work

The CE based framework for power system analysis proposed in this paper can be used specifically for two different tasks: the identification of probability distributions for sampling dangerous events and the computation of reliability indices. Many research papers focus on the problem of evaluation of reliability indices for specific power system problems. As way of example, in the bibliography on the application of probability methods in power system reliability evaluation over the period 1996-1999 (see [4]) over 150 references are given. It has long been recognised by power system engineers that crude Monte-Carlo (MC) simulations for evaluating the probability of blackouts was computationally inefficient and numerous techniques were proposed to address this problem. For example, Reference [2] proposes to combine, in the context of distribution systems, MC simulations with some analytical approaches. Paper [12] proposes to exploit artificial neural networks based on the learning vector quantization algorithm to make MC techniques more computationally efficient for loss of load probability calculations. Importance sampling as well as other variance reduction techniques have also been recurrently proposed in the power system literature as an enhancement of MC methods (see, e.g., [13], [3], [14]). The approach proposed in this paper can be seen as an importance sampling technique where the sampling distribution is built by using algorithms proposed in the rare-event simulation literature, and, more precisely, those based on the cross-entropy method. Researchers in power systems have already used tools from the rare-event simulation literature to compute reliability indices (see [8]), but, to the best of our knowledge, never those exploiting the cross-entropy method. This method has however already been exploited in the context of power system combinatorial optimization (see [9]).

For identifying probability distributions targeting dangerous events, the CE based approach proposed in this paper will only have to run in principle a security analysis for a relatively small set of events. Viewed in this light, it can be seen as an approach for rapidly identifying dangerous events in a power system and can therefore be apparented to the significant body of work related to contingency filtering/screening in power systems (see, e.g., [10], [5], [7]). Most of the approaches for contingency filtering however rely on deterministic algorithms while the one proposed in this paper is a stochastic one. The importance sampling distribution determined by the rare-event simulation algorithm can also be seen as a classifier for dangerous/non-dangerous events: if it associates a low probability to an event, then it is a non-dangerous one and, otherwise, a dangerous one. In this respect, the approach proposed has some similarities with the many works where classifiers for assessing the degree of severity of power system scenarios are built (see, e.g., [7], [18], [17]).

TABLE I
Standard error of the estimator (2) for different importance sampling distributions.

Importance sampling distribution	$\sigma_{\hat{l}}$
$Gauss_{\mathcal{X}}(\cdot, v_1)$	$4.897.10^{-27}$
$Gauss_{\mathcal{X}}(\cdot, v_2)$	$3.674.10^{-27}$
$Gauss_{\mathcal{X}}(\cdot, v_3)$	$3.549.10^{-27}$
$Gauss_{\mathcal{X}}(\cdot, v_4)$	$2.996.10^{-27}$
$Gauss_{\mathcal{X}}(\cdot, v_5)$	$1.803.10^{-27}$

VI. Conclusion

This paper has proposed a framework for identifying dangerous events and as a byproduct for efficiently computing reliability indices in a power system. The proposed framework is relying on the cross-entropy method for rare-event simulation and its application to power system problems requires to be able to associate to each event a severity value. The approach was illustrated on a problem of static voltage security and, even if preliminary, the results were encouraging.

The proposed approach essentially aims at deriving from an initial sampling distribution a sampling distribution which focuses on events of high severity in a way which is only mildly depending on the initial distribution. Therefore, this approach may help to uncover unanticipated blackout scenarios whose probability could then be elucidated by further analysis. We believe that such an approach helping to uncover possible blackout scenarios would foster further work in order to improve current probabilistic models used in power system reliability studies, in particular by forcing power system engineers to better assess the probability of the combinations of events leading to these situations. The identification of dangerous events that are usually ignored when planning and/or operating power systems will also raise questions concerning the design of strategies to mitigate their effects or decrease their probability of inception. In such a context, we believe it would also be interesting to design techniques able to identify whether there are some common modes of instability behind these numerous dangerous events that may potentially be discovered by such an approach. Then, the knowledge of these modes could be exploited to take specific actions (e.g., transmission investments, modification of market rules) for increasing the security of the system.

From a technical viewpoint, our experiments also suggest several research directions. At first, for a given stability problem, it would be interesting to identify which type of severity function would lead to the best results. Some simulations, not reported here, have shown that the performances of the proposed framework is indeed strongly correlated to the nature of this function. Second, in our opinion, it would also be possible to improve the performances of existing cross-entropy based methods by designing alternative algorithms for the stochastic counterpart problem, whose solution is at the heart of the cross-entropy method.

While the framework presented in this paper is certainly attractive and can potentially lead to the development of new tools for power system analysis, a careful evaluation of the performances of these techniques with respect to other apparented techniques, especially in the field of importance sampling, would be interesting.

Acknowledgment

Damien Ernst is a Research Associate of the Belgian FNRS (French acronym for National Fund of Scientific Research) from which he acknowledges the financial support.

References

[1] Final Report, System Disturbance on 4 November 2006. Technical report, UCTE, 2007.

[2] R.N. Allan and M.G. Da Silva. Evaluation of reliability indices and outage costs in distribution systems. *IEEE Transactions on Power Systems*, 10(1):413–419, 1995.

[3] K. Bae and J.S. Thorp. An importance sampling application: 179 bus WSCC system under voltage based hidden failures and misoperations. In *Proceedings of the Thirsty-First Annual Hawaii International Conference on System Sciences*, page 39. IEEE Computer Society, 1998.

[4] R. Billinton, M. Fotuhi-Firuzabad, and L. Bertlin. Bibliography on the application of probability methods in power system reliability evaluation 1996-1999. *IEEE Transactions on Power Systems*, 16(4):595–602, 2001.

[5] F. Capitanescu, M. Glavic, D. Ernst, and L. Wehenkel. Contingency filtering techniques for preventive security constrained optimal power flow. *IEEE Transactions on Power Systems*, 22(4):1690–1697, 2007.

[6] F. Capitanescu, M. Glavic, D. Ernst, and L. Wehenkel. Interior-point based algorithms for the solution of optimal power flow problems. *Electric Power Systems Research*, 77(5-6):508–517, 2007.

[7] K.W. Chan, R.W. Dunn, A.R. Daniels, J.A. Padget, A.O. Ekwue, P.H. Buxton, and M.J. Rawlins. On-line dynamic-security contingency screening and ranking. *IEE Proceedings- Generation, Transmission and Distribution*, 144(2):132–138, 1997.

[8] Q. Chen. *The Probability, Identification and Prevention of Rare-Events in Power Systems*. PhD thesis, Iowa State University, 2004.

[9] D. Ernst, M. Glavic, G.B. Stan, S. Mannor, and L. Wehenkel. The cross-entropy method for power system combinatorial optimization problems. In *Proceedings of the IEEE Power Tech Conference*, 2007.

[10] D. Ernst, D. Ruiz-Vega, M. Pavella, P. Hirsch, and D. Sobajic. A unified approach to transient stability contingency filtering, ranking and assessment. *IEEE Transactions on Power Systems*, 16(3):435–444, 2001.

[11] H. Laffaye. Les Blackouts Récents : Enseignements et Expérience RTE. Technical report, RTE, 2004.

[12] X. Luo, C. Singh, and A.D. Patton. Power system reliability evaluation using learning vector quantization and Monte-Carlo simulations. *Electrical Power Systems Research*, 66(2), 2003.

[13] M.V.F. Pereira, M.E.P. Maceira, G.C. Oliveira, and L.M.V.G. Pinto. Combining analytical models and monte-carlo techniques in probabilistic power system analysis. *IEEE Transactions on Power Systems*, 7:265–272, 1992.

[14] J.H. Pickels and I.H. Russel. Importance sampling for power system security assessment. In *Proceedings of the Third International PMAPS Conference*, pages 47–52, 1991.

[15] R.Y. Rubinstein and D.P. Kroese. *The Cross-Entropy Method. A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning*. Information Science and Statistics. Springer, 2004.

[16] M. Shaaban, Y. Ni, and F. Wu. Transfer capability computations in deregulated power systems. In *Proceedings of the 33rd Hawaii International Conference on System Sciences*, page 5, 2000.

[17] L. Wehenkel, C. Lebrevelec, M. Trotignon, and J. Batut. Probabilistic design of power-system special stability controls. *Control Engineering Practice*, 7(2):183–194, 1999.

[18] L. Wehenkel, T. Van Cutsem, and M. Ribbens-Pavella. An artificial intelligence framework for on-line transient stability assessment of power systems. *IEEE Transactions on Power Systems*, PWRS-4:789–800, 1989.

VII. Biographies

Florence Belmudes (M'07) received her M.Sc. from SUPELEC (France) in 2007. She is now a Ph.D. student in the Systems and Modeling Research Unit of the University of Liège. Her main research interests are in the field of power system analysis and control.

Damien Ernst (M'98) received the M.Sc. and Ph.D. degrees from the University of Liège in 1998 and 2003, respectively. He is currently a Research Associate of the FNRS (Belgian National Fund of Scientific Research) and he is affiliated with the Systems and Modeling Research Unit of the University of Liège. Damien Ernst spent the period 2003-2006 with the University of Liège as a Postdoctoral Researcher of the FNRS and held during this period positions as visiting researcher at CMU, MIT and ETH. He spent the academic year 2006-2007 working at SUPELEC (France) as professor. His main research interests are in the field of power system dynamics, optimal control, reinforcement learning and design of dynamic treatment regimes.

Louis Wehenkel (M'93) graduated as an Electrical Engineer (electronics) in 1986 and received the Ph.D. degree in 1990, both at the University of Liège, where he is full Professor of Electrical Engineering and Computer Science. His research interests lie in the fields of stochastic methods for systems and modeling, machine learning and data mining, with applications in power systems planning, operation and control and bioinformatics.