



HAL
open science

An entropy-driven method for exploring extreme and unexpected accident scenarios in the risk assessment of dynamic engineered systems

Pietro Turati, Nicola Pedroni, Enrico Zio

► To cite this version:

Pietro Turati, Nicola Pedroni, Enrico Zio. An entropy-driven method for exploring extreme and unexpected accident scenarios in the risk assessment of dynamic engineered systems. European Safety and Reliability Conference ESREL, Sep 2015, Zurich, Switzerland. hal-01270954

HAL Id: hal-01270954

<https://hal.science/hal-01270954>

Submitted on 8 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An entropy-driven method for exploring extreme and unexpected accident scenarios in the risk assessment of dynamic engineered systems

P. Turati & N. Pedroni

Chair on Systems Science and the Energetic Challenge, Foundation EDF, CentraleSupélec, Paris, France.

E. Zio

Energy Department, Politecnico di Milano, Milan, Italy.

Chair on Systems Science and the Energetic Challenge, Foundation EDF, CentraleSupélec, Paris, France.

ABSTRACT: In Integrated Deterministic Probabilistic Safety Assessment (IDPSA), the end state reached by the dynamic engineered system under analysis, as outcome of an accident scenario, depends on the failure sequence in the scenario, but also on the magnitude and timing of the failures. Including these additional dimensions can render the analysis infeasible, as the size of the system state space explodes with the large number of components and physical quantities involved in realistic cases. In this paper, we address the problem of automatically and efficiently probing the space of the event sequences in a dynamic system. In particular, an entropy-driven technique is proposed for intelligently focusing the exploration efforts preferably on those scenarios that are most interesting for the scope of the analysis. To this aim, the relevance of a given accident scenario is defined by two original indicators: (i) the criticality of the corresponding system end states, related to their (un)likelihood of occurrence and/or their (extreme) consequences; (ii) the level of “unexpectedness” of the corresponding event sequences, with respect to the prior knowledge of the analyst. The method is applied to a simple case study of a dynamic system for gas transmission.

1 INTRODUCTION

The response of a dynamic engineered system under nominal and accident conditions can be studied by resorting to mathematical models and computer codes that describe and reproduce the dynamic behavior of the system of interest (Smidts & Devooght, 1992), (Gao & Dougal, 2002), (Hansen et al., 2004). When the system is complex, as is typically the case in practice, simulation is used. The outcomes of the system model simulations allow studying, analyzing and understanding the possible evolutions (i.e. scenarios) that the system can undergo and the corresponding end states. In this respect, a framework of combination of event tree analysis (for the identification of event sequences) and the mathematical models of the dynamic of the physical phenomena developing along the sequences, allows determining the end states that can be reached and deriving causality relations among events and processes, in a comprehensive way that account for both the logic of the system and the physic of the process therein (Siu, 1994), (Aldemir, 2013).

In this respect, works on Dynamic Event Trees (DETs) (Cojazzi, 1996), (Hsueh & Mosleh, 1996), (Labeau et al., 2000), (Cepin & Mavko, 2002), (Kloos & Peschke, 2006), (Hakobyan et al., 2008),

have highlighted that, indeed, the end states reached by a system as a result of an accident scenario do not depend only on the order of occurrence of the events in the sequence, but also on the time at which these events occur and their magnitude (Devooght & Smidts, 1992), (Aldemir, 2013), (Di Maio et al., 2014). For example, it has been shown that the same order of occurrence of events (i.e. the same scenario) may lead to different end states, depending on the different occurrence times (Di Maio et al., 2014). On the other hand, the introduction of the time dimension into the analysis leads to a dramatic increase in the size of the system state space, which makes its thorough exploration impossible in practical cases of real engineered system characterized by thousands of components and associated physical quantities. Furthermore, the computational cost associated to the simulation of a single dynamic scenario can be very high: for example, the computer code RELAP can take long times for simulating the thermal-hydraulics of nuclear systems under specific accident conditions (Idaho National Laboratory, 2005), (Fong et al., 2009), (Perez et al., 2011). Indeed, in IDPSA the computational cost is always an issue (Rutt et al., 2006), (Catalyurek et al., 2010), (Zio, 2014).

On the other hand, in spite of the above challenges, IDPSA is considered the way to account for

time-dependences in the development of accident sequences and to probe the scenario space for identifying undiscovered plant unreliability, unexpected scenarios, surprises (Zio, 2014). For this latter point, in the present paper, an automatic procedure for efficiently probing the space of the event sequences in a dynamic system, which relies on an entropy-based criterion (Shannon, 2001), is proposed. More specifically, given an event tree of the logic of the accident sequences for the system of interest, a function of the entropy is used as a measure of the uncertainty in the end states (i.e. in the possible outcomes) of each scenario; then, intuitively, the more uncertain the outcome, the higher the number of simulations needed to probe the space for the event sequences associated to that scenario (Hu et al., 2004). Thus, the proposed procedure utilizes the entropy measure for efficiently allocating the simulation efforts among the different scenarios of the system event tree, with the specific objective of favoring those characterized by higher uncertainty. The approach consists of two steps: (i) one of the possible scenarios is chosen according to the entropy measure and (ii) a simulation of the system is run, conditioned to the selected scenario.

The performance of the proposed method has been verified on a simple, but representative, case of a dynamic system made by a gas transmission pipe (actively controlled by a valve), which is connected in series to two pipes in parallel; all the components are subject to stochastic failures described by proper probability distributions.

The rest of the paper is organized as follows. In Section 2, a mathematical formulation of the method is provided. In Section 3, suggestions on how to implement the method are given. In Section 4, the case study and the associated results are presented. Finally, in Section 5, concluding remarks and directions to future research are given.

2 NEGATIVE ENTROPY-DRIVEN METHOD FOR EXPLORING SCENARIOS IN DYNAMIC EVENT TREES

Within the context of interest of the present paper, a scenario defines a specific sequence (order) of events in the life evolution of the dynamic system, which may involve a particular group of components, safety functions or actions (e.g. mechanical failures, activation of safety systems and human decisions). A same scenario (i.e. same sequence of events) may bring the dynamic system to different final outcomes, hereafter also called End States (ESs), for different times of occurrence of the events. For example, consider a scenario of reduction of gas flow in a network due to a failure of a valve: the ES of the amount of gas supplied to the

consumers within a given time depends on the time of valve failure.

In the context of the present work, Dynamic Event Tree (DET) is used as the logical modeling technique to derive, by means of simulations, the scenarios that can arise in the life evolution of a dynamic system as a result of a sequence of successes and failures of different components and functions within the system itself (Mercurio et al, 2009). Softwares for DET analyses such as Dynamic Logical Analytical Methodology (DYLAM) (Cojazzi, 1996), Accident Dynamic Simulator (ADS) (Hsueh & Mosleh, 1996) and Monte Carlo Dynamic Event Tree (MCDET) (Kloos & Peschke, 2006), are available and, in principle, *all* possible *scenarios* that can occur in the system can be extracted. However, *not all* possible different *time sequences* within a given scenario can be explored, by reason of the extremely high computational costs needed for simulating all of them (Di Maio et al., 2014).

To contribute to the solution of this issue, we here propose a technique for efficiently exploring the variability in the ESs that can be reached within each scenario, due to *different event timings*. The driving criterion for the search is the following: the higher is the uncertainty (i.e. the lack of information) of the analyst about the possible ESs of a scenario, the higher is the “exploration effort” allocated for that scenario. The scheme of the algorithm consists of two steps: 1) *selection* of the scenario that is worth to be deeply explored according to an “information gain” criterion; 2) *simulation* of the evolution of the system (*conditioned* to the selected *scenario*) and updating of the knowledge about the selected scenario according to the end state reached by the simulation.

2.1 Scenario selection

Assume that for each a-priori available scenario S_j only a limited number N_j^{ES} of ESs can be reached. Each ES_i with $i = 1, \dots, N_j^{ES}$ has its unknown probability of occurring μ_{ji} . Thus, the outcome of a simulation of the system behavior conditioned to a certain scenario S_j follows a multinomial distribution with unknown parameter vector $\boldsymbol{\mu}_j = (\mu_{j1}, \dots, \mu_{jN_j^{ES}})$. In this view, our knowledge about the system and, in particular, about scenario S_j increases with the confidence about its parameters $\boldsymbol{\mu}_j$; in other words, our knowledge increases the more we know how many ESs can be reached (N_j^{ES}) and with what probability $(\mu_{j1}, \dots, \mu_{jN_j^{ES}})$. Within a Bayesian framework (Bayes, 1763), the uncertainty of the analyst about the vector of probabilities $\boldsymbol{\mu}_j$ can be described by a prior distribution on the parameter vector $\boldsymbol{\mu}_j$: the characteristics (i.e. the parameters) of such distribution can be updated when new pieces of information about the end states become available from new system simulations per-

formed. In this context, the Negative Differential Entropy $I(\boldsymbol{\mu}_j)$ (MacKay, 1992), (Loredo & Chernoff, 2003), (Hu et al., 2004), (Hu, 2005), (Cover & Thomas, 2012) has been considered as measure of the current information available about parameter $\boldsymbol{\mu}_j$ of scenario S_j , where the higher $I(\boldsymbol{\mu}_j)$, the higher the information about parameter $\boldsymbol{\mu}_j$:

$$\begin{aligned} I(\boldsymbol{\mu}_j) &= E \left[\log \left(f(\boldsymbol{\mu}_j) \right) \right] \\ &= \int \log \left(f(\boldsymbol{\mu}_j) \right) f(\boldsymbol{\mu}_j) d\boldsymbol{\mu}_j, \end{aligned} \quad (1)$$

with $f(\cdot)$ being the probability density function of the parameter vector $\boldsymbol{\mu}_j$. Since the value of the negative differential entropy is affected by the unit of measure, it is important to consider dimensionless quantities or assure that the different parameters involved are coherent (i.e. they have the same order of magnitude) (MacKay, 1992).

When new information about the end states is available (i.e. when a new system simulation is run whose end state is y), the probability distribution $f(\boldsymbol{\mu}_j)$ of parameter vector $\boldsymbol{\mu}_j$ can be updated in a Bayesian framework (Bayes, 1763) to get the corresponding posterior $f(\boldsymbol{\mu}_j|y)$. Then, the information $I(\boldsymbol{\mu}_j|y)$, after the simulation outcome y is obtained, becomes:

$$\begin{aligned} I(\boldsymbol{\mu}_j|y) &= E \left[\log \left(f(\boldsymbol{\mu}_j|y) \right) \right] \\ &= \int \log \left(f(\boldsymbol{\mu}_j|y) \right) f(\boldsymbol{\mu}_j|y) d\boldsymbol{\mu}_j. \end{aligned} \quad (2)$$

The increment of information, i.e. $\Delta I(\boldsymbol{\mu}_j, y) = I(\boldsymbol{\mu}_j|y) - I(\boldsymbol{\mu}_j)$, can be used to represent the information gain due to the last ES y reached by simulation. Finally, for taking into account the randomness of the simulation process, the expected information gain $E[\Delta I(\boldsymbol{\mu}_j, Y)]$ provided by a new simulation is considered as driving criterion for choosing the scenario which, on average, provides more additional information:

$$E[\Delta I(\boldsymbol{\mu}_j, Y)] = \int \Delta I(\boldsymbol{\mu}_j, y) g(y) dy, \quad (3)$$

with $g(y)$ being the probability density function of the outcomes of the system simulation. The next simulation to run can, then, be selected by considering the scenario S_j^* which presents the highest value of expected information gain:

$$S_j^* = \underset{j}{\operatorname{argmax}} E[\Delta I(\boldsymbol{\mu}_j, Y)]. \quad (4)$$

For the sake of clarity, in what follows an example is proposed where an explicit expression for the increment of information is derived with reference to the prior Dirichlet distribution $\boldsymbol{\mu}_j \sim \text{Dir}(\mathbf{u}_j)$. This modeling choice for the parameter uncertainty is straightforward when the distribution of the ESs is characterized by a Multinomial distribution. Indeed,

being the Dirichlet distribution the Multinomial conjugated one, the update of the prior distribution after the simulation of a new outcome y consists in updating the parameter vector \mathbf{u}_j . In particular u_{ji} represents how many times the ES_i has been reached in scenario S_j . From Equation (1), it follows (Lazo & Rathie, 1978), (Hu, 2005) that:

$$\begin{aligned} I(\boldsymbol{\mu}_j|\mathbf{u}_j) &= -\ln \left(\frac{\prod_{i=1}^{N_j^{ES}} \Gamma(u_{ji})}{\Gamma \left(\sum_{i=1}^{N_j^{ES}} u_{ji} \right)} \right) \\ &+ \sum_{i=1}^{N_j^{ES}} (u_{ji} - 1) \cdot \left(\Psi(u_{ji}) - \Psi \left(\sum_{i=1}^{N_j^{ES}} u_{ji} \right) \right), \end{aligned} \quad (5)$$

where $\Psi(\cdot)$ is the digamma function. For integer values of parameters u_{ji} , Equation (5) can be rewritten as:

$$\begin{aligned} I(\boldsymbol{\mu}_j|\mathbf{u}_j) &= -\ln \left(\frac{\prod_{i=1}^{N_j^{ES}} \Gamma(u_{ji})}{\Gamma \left(\sum_{i=1}^{N_j^{ES}} u_{ji} \right)} \right) \\ &- \sum_{i=1}^{N_j^{ES}} (u_{ji} - 1) \cdot \left(\sum_{m=u_{ji}}^{-1+\sum u_{ji}} \frac{1}{m} \right), \end{aligned} \quad (6)$$

where $\Gamma(\cdot)$ represents the Gamma function. Note that, for integer values of the input, $\Gamma(\cdot)$ has a factorial trend: then, even with a relatively small input argument, it may exceed the machine precision (e.g. $171! > 1.7977 \cdot 10^{308}$, which is the largest positive floating-point number in the Matlab software used in our calculations). However, assuming that a new simulation is run and its outcome $y = i$, then:

$$\begin{aligned} \Delta I(\boldsymbol{\mu}_j, y = i) &= I(\boldsymbol{\mu}_j|\mathbf{u}_j, y = i) - I(\boldsymbol{\mu}_j|\mathbf{u}_j) \\ &= \ln \left(\frac{\sum_{i=1}^{N_j^{ES}} u_{ji}}{u_{ji}} \right) + \frac{N_j^{ES}}{\sum_{i=1}^{N_j^{ES}} u_{ji}} \\ &- \frac{1}{u_{ji}} - \sum_{m=u_{ji}+1}^{\sum u_{ji}} \frac{1}{m}, \end{aligned} \quad (7)$$

which contains no Gamma function and, thus, presents less numerical issues associated to the evaluation of factorials. In the following, for the sake of clarity, but with no loss of generality, all the discussion is developed with reference to the above mentioned multinomial-Dirichlet framework.

2.2 System simulation and updating of the parameters

After selecting a scenario S_j according to the criterion outlined in Section 2.1, a simulation of the system is run conditioned to S_j . The ES y thereby reached is used for updating, within the Bayesian framework, the parameters of the Dirichlet distribu-

tion: the more simulations are run conditioned to S_j , the higher is the confidence gained about the parameter vector \mathbf{u}_j of the associated Dirichlet distribution and, consequently, about $\boldsymbol{\mu}_j$ itself. Since the Dirichlet distribution is the conjugate of the Multinomial one (Gelman et al., 2014), the corresponding posterior is still a Dirichlet distribution with properly updated parameters. In particular, for updating on the basis of a single simulated ES y , we have:

$$\boldsymbol{\mu}_j|y = i, \mathbf{u}_j \sim \text{Dir}(u_{j1}, \dots, u_{ji} + 1, \dots, u_{jN_j^{ES}}) \quad (8)$$

3 IMPLEMENTATION ISSUES AND GUIDELINES

3.1 Prior knowledge and adaptive parameter vector

If prior information regarding a scenario is available, it can be embedded in the prior (i.e. in the Dirichlet parameters). Otherwise, in absence of prior knowledge, uninformative priors should be adopted for leaving the algorithm free of gathering information and knowledge from the future simulations. The “type” of prior knowledge that most affects the information gain is the number N_j^{ES} of End States (ESs) that can be reached by simulating a scenario S_j . This is exemplified in Table 1. Let us assume that we have six scenarios S_1, \dots, S_6 and six end states ES_1, \dots, ES_6 , e.g., representing two possible safe conditions (ES_1, ES_2) and two near misses conditions (ES_3, ES_4), i.e., outcomes that have almost reached the respective failure conditions (ES_5, ES_6). We also know that not all the scenarios can reach all the end states: for example, S_1 can reach only the safe condition ES_1 , whereas, e.g. S_4 can reach ES_1, \dots, ES_4 . We assume, in addition, that each scenario has been explored with the same number of simulations (e.g. 60 in this case). Then, the corresponding average information gains for the six scenarios (that can reach a different number N_j^{ES} of ESs) are reported in Table 1. It can be seen that the larger N_j^{ES} , the higher the expected information gain, given the same number of simulations already run.

Table 1. Expected increment of information about scenarios $S_j, j = 1, \dots, 6$ for different priors $f(\boldsymbol{\mu}_j)$ given the same number of simulations. Each row corresponds to a scenario each of which can reach a different number of ESs (i.e. from 1 to 6).

	ES_1	ES_2	ES_3	ES_4	ES_5	ES_6	$E[\Delta I(\boldsymbol{\mu}, Y \mathbf{u})]$
S_1	60	0	0	0	0	0	0.000
S_2	30	30	0	0	0	0	0.008
S_3	20	20	20	0	0	0	0.016
S_4	15	15	15	15	0	0	0.025
S_5	12	12	12	12	12	0	0.033
S_6	10	10	10	10	10	10	0.041

On the other side, if we do not know how many ESs a given scenario can reach and we a priori assume that a certain scenario can reach all the ESs (even though in reality it cannot), what emerges from Table 2 is that, given the same number of simulations run for each scenario, the expected information gains are not so different.

Table 2. Expected increment of information about scenarios $S_j, j = 1, \dots, 6$ for different priors $f(\boldsymbol{\mu}_j)$ given the same number of simulations. Each row corresponds to a scenario, each of which can reach a different number of ESs (i.e. from 1 to 6). However, we assume an initial uninformative prior for all the scenarios considered.

	ES_1	ES_2	ES_3	ES_4	ES_5	ES_6	$E[\Delta I(\boldsymbol{\mu}, Y \mathbf{u})]$
S_1	61	1	1	1	1	1	0.032
S_2	31	31	1	1	1	1	0.033
S_3	21	21	21	1	1	1	0.034
S_4	16	16	16	16	1	1	0.035
S_5	13	13	13	13	13	1	0.036
S_6	11	11	11	11	11	11	0.037

Thus, in order to avoid considering reachable ESs that are not, an adaptive Dirichlet parameter vector is here proposed in this paper. At the beginning of the analysis, the prior Dirichlet vector contains only the parameters of those ESs that the analyst is sure the system can reach; then, adaptively, once a new ES has been reached for the first time, a new component is added to the Dirichlet parameter vector.

3.2 Stopping criteria

In many cases, the size of the system and the variability of its behavior (in practice, the number of ESs a scenario can reach and the corresponding probabilities), are not known a priori. On the other hand, the computational cost associated to a system simulation is typically known (e.g. in terms of average time for each simulation). In this view, we assume that the total computational effort is imposed as a constraint: then, the objective of the proposed method is to assure an efficient and intelligent distribution of the system simulations among the different scenarios, in order to perform the deepest possible exploration of the system behavior (which is a priori unknown). It must be noticed that the proposed method does not guarantee that the whole event space is probed. Indeed, if the computational capacity available (in practice, the total number of simulations that can be run) is small compared to the size of the system state space, only a limited number of end states can be explored for each scenario, anyway.

3.3 Uniform time space probing

Typically, the mathematical model describing system evolution consists of a system of differential equations whose parameters depend on the state of the components. Thus, the variability of the model output (i.e. of the system ESs) is dependent upon the time at which the stochastic transitions among the states of the components occur. However, simulating the occurrence of such random events using the real probability distributions of the components transition times could lead the system simulator to visit many times the same ESs (i.e. those with higher probability of occurrence) and very few times those ESs that are rare, but still possible. In order to avoid this problem, we propose exploring the time state space of each scenario by *uniformly* sampling on the support defined by the selected scenario, being aware of modifying the original probability distribution of the ESs. For the sake of clarity, consider a scenario S_j that involves the change of states of two components (namely, A and B) at the corresponding time T_A and T_B in a specific order within the mission time T_{Miss} (e.g. in this case, $T_A < T_B \leq T_{Miss}$): this exemplary situation is illustrated in Figure 1, where the domain of the scenario S_j is represented by the triangular shaded area. Then, in order to run a simulation of the dynamic system evolution, we sample the pair (T_A, T_B) from the joint uniform distribution on the support defined by S_j (see shaded area in Figure 1). For this purpose, we resort to a Markov Chain Monte Carlo (MCMC) Gibbs sampling (Robert and Casella, 2005), which can be easily implemented.

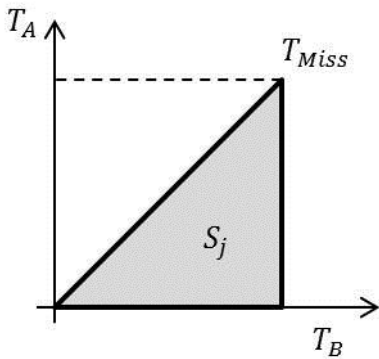


Figure 1. Support of the occurrence time of the events A and B defining scenario S_j , where $T_A < T_B \leq T_{Miss}$.

4 CASE STUDY

The case study under analysis is a gas transmission subnetwork composed of two pipes in parallel and another one in series. The input of each pipe is controlled by a valve. The whole block diagram is shown in Figure 2, where each pair valve-pipe is considered as a single block.

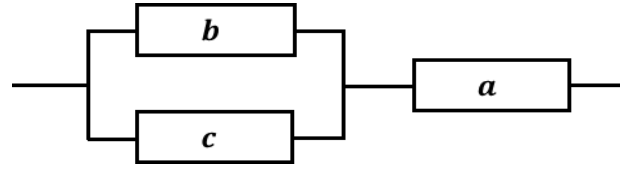


Figure 2. Block diagram of the system under analysis.

Each pipe can transmit gas with a maximum flow rate of $[\phi_a, \phi_b, \phi_c] = [8, 5, 5] \cdot 10^4 \text{ m}^3/\text{day}$, for pipes a, b, c , respectively. A control system adjusts the opening of the valves in order to guarantee the equilibrium between the input and output flows. Figure 3 shows the event tree containing all the scenarios that can occur in the system. If one of the pipes in parallel breaks, the control system immediately closes the corresponding valve and increases the flow rate of the remaining pipe to the maximum, in order to compensate for the diminished flow. No reparation strategies are considered. The system presents 8 possible scenarios with different operating conditions: *i) safe*, i.e. all pipes are functioning correctly; *ii) overloaded*, i.e. one of the pipes in parallel is closed; *iii) broken*, i.e. no gas is provided by the system.

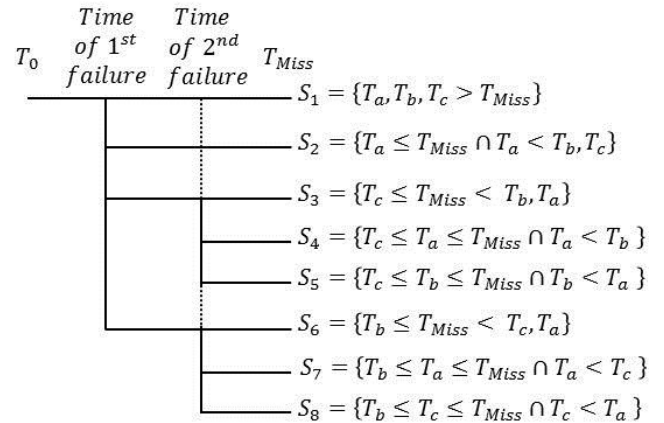


Figure 3 Event tree representation of the 8 scenario that can occur, where T_a, T_b, T_c are the times of failures of components a, b, c , respectively, and T_{Miss} is the mission time.

The ESs for each scenario have been defined and classified on the basis of two output variables: *i) the amount of Gas provided in Safe Conditions (GSC)*, i.e. when all the components are functioning correctly; *ii) the amount of Gas provided in Overloaded Conditions (GOC)*, i.e. when one of the two pipes in parallel is down and the remaining one works at its maximum flow rate. With respect to that, GSC_{max} and GOC_{max} indicate the maximum quantities of gas that can be provided within the mission time $T_{Miss} = 900d$, in safe and overload conditions, respectively, i.e. $GSC_{max} = \phi_a \cdot T_{Miss}$ and $GOC_{max} = \max(\phi_b, \phi_c) \cdot T_{Miss}$. The ESs are, then, divided into six classes according to the criteria reported in Figure 4. For example, $ES_4 = \{GSC_{max}/3 < GSC \leq 2GSC_{max}/3 \cap 0 \leq GOC \leq GOC_{max}/3\}$, which means that the system

gously, $\{S_5, ES_5\}$ (which are very difficult to be “spotted out” by a crude MC sampling and which obviously “determine” the value of NFE) occur in scenarios S_4 or S_5 that do not give rise to many ESs. As shown for an analytical representative situation in Table 4, the entropy-driven exploration tends to “visit” more frequently those scenarios that can reach more ESs and less frequently those that give rise to few end states (this is expected, since the highest computational effort is devoted to those scenarios whose outcomes are “more uncertain”).

Table 4. Average number of simulated scenarios that end in configuration $\{S, ES\}$, given a computational effort of 500 simulations with flow rate parameters $[\phi_a, \phi_b, \phi_c] = [8, 3, 5] \cdot 10^4 \text{ m}^3/\text{day}$, for 1000 experiments.

	ES_1	ES_2	ES_3	ES_4	ES_5	ES_6	Tot
S_1	0.0	0.0	0.0	0.0	0.0	16.1	16.1
S_2	13.8	0.0	0.0	13.5	0.0	13.5	40.8
S_3	0.0	19.9	0.0	13.3	6.6	19.9	59.7
S_4	26.0	12.9	0.0	22.9	1.0	8.2	71
S_5	26.1	12.9	0.0	22.6	0.9	8.2	70.7
S_6	0.0	0.0	13.1	0.0	13.2	13.2	39.5
S_7	22.5	22.3	11.4	22.5	11.2	11.2	111.1
S_8	22.5	22.4	11.4	22.3	11.3	11.2	111.1

Thus, in a configuration like the one exemplified, the behavior of the entropy-driven approach is not so different from that of a crude MC.

5 CONCLUSIONS

The complexity of modern engineered systems is challenging the classical methods of risk analysis for discovering vulnerability and identifying possible accident scenarios, in order to prevent them and/or prepare for protection and mitigation of their consequences. On the other hand, given the large impact that accidents may have in certain industries, there is a responsible will to know as much as possible of these scenarios and leave out only little of unexpected. Under these premises, IDPSA seems a valid approach for improving safety. In this context, with the objective of thoroughly exploring the response of dynamic systems subject to stochastic state transitions, an original entropy-driven method has been proposed to efficiently distribute the system simulation efforts among the possible scenarios. The proposed method allocates more simulation efforts to visiting those scenarios that can reach a larger number of ESs, i.e. those scenarios whose

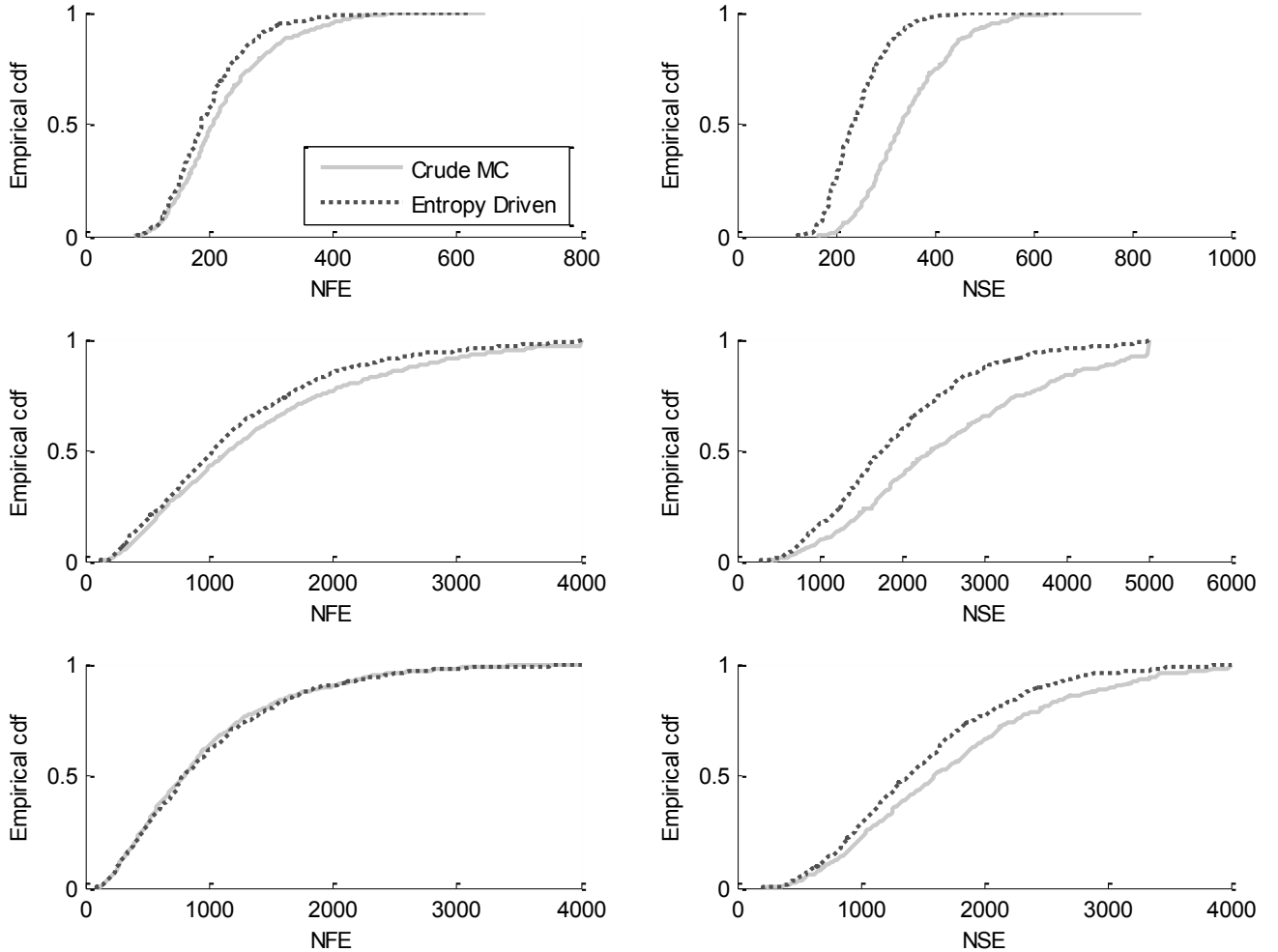


Figure 5. Empirical cdfs of the NFE (left) and of the NSE (right) for crude MC (light line) and for the entropy-driven method (dark dashed line) with flow rate parameters $[\phi_a, \phi_b, \phi_c] = [8, 5, 5] \cdot 10^4 \text{ m}^3/\text{day}$ (top), $[\phi_a, \phi_b, \phi_c] = [8, 3.67, 5] \cdot 10^4 \text{ m}^3/\text{day}$ (middle), $[\phi_a, \phi_b, \phi_c] = [8, 3, 5] \cdot 10^4 \text{ m}^3/\text{day}$ (bottom).

outcomes are more uncertain.

Application to a simple case study has shown that the performance is promising: the expected number of simulations needed to explore all possible end states of the system is consistently lower than that of a standard Monte Carlo technique, in the majority of the cases considered. However, the exploration performance decreases when the rarest events occur in scenarios that can reach a small number of ESs. Future research works will focus on the possibility of embedding specific preferences of the analyst in the exploration method, such as the interest in exploring those scenarios which can reach a specific ES, e.g. extremely unlikely or extremely critical from a risk perspective. Other points of interest are the opportunity of (i) including the uncertainty brought into the analysis by possible different failure magnitudes and (ii) assessing the probability of the events emerged during the exploration.

6 REFERENCES

- Aldemir, T. 2013. "A survey of dynamic methodologies for probabilistic safety assessment of nuclear power plants" *Annals of Nuclear Energy* 52: 113-124.
- Bayes, T. 1763. "An essay towards solving a problem in the doctrine of chances" *Philosophical Transactions of the Royal Society of London* 53: 370-418.
- Catalyurek, U., Rutt, B., Metzroth, K., Hakobyan, A., Aldemir, T., Denning, R., Dunagan S. & Kunsman, D. 2010. "Development of a code-agnostic computational infrastructure for the dynamic generation of accident progression event trees" *Reliability Engineering & System Safety* 95(3): 278-294.
- Čepin, M. & Mavko, B. 2002. "A dynamic fault tree" *Reliability Engineering & System Safety* 75(1): 83-91.
- Cojazzi, G. 1996. "The DYLAM approach for the dynamic reliability analysis of systems" *Reliability Engineering & System Safety* 52(3): 279-296.
- Cover, T. M., & Thomas J. A. 2012. *Elements of information theory*. John Wiley & Sons.
- Devooght, J. & Smidts C. 1992. "Probabilistic reactor dynamics—I: the theory of continuous event trees" *Nuclear science and engineering* 111(3): 229-240.
- Di Maio, F., Baronchelli, S. & Zio E. 2014. "A Computational Framework for Prime Implicants Identification in Noncoherent Dynamic Systems" *Risk Analysis* 35(1):142-156.
- Fong, C. J., Apostolakis, G. E., Langewisch, D. R., Hejzlar, P., Todreas, N. E., & Driscoll, M. J. 2009. "Reliability analysis of a passive cooling system using a response surface with an application to the flexible conversion ratio reactor" *Nuclear Engineering and Design* 239(12): 2660-2671.
- Gao, L., Shengyi L. & Dougal, R. A. 2002. "Dynamic lithium-ion battery model for system simulation" *Components and Packaging Technologies, IEEE Transactions on* 25(3): 495-505.
- Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. 2014. *Bayesian data analysis*. Vol. 2. Chapman & Hall/CRC.
- Hakobyan, A., Aldemir, T., Denning, R., Dunagan, S., Kunsman, D., Rutt, B. & Catalyurek, U. 2008. "Dynamic generation of accident progression event trees" *Nuclear Engineering and Design* 238(12): 3457-3467.
- Hansen, A. D., Jauch, C., Sørensen, P. E., Iov, F. & Blaabjerg, F. 2004. *Dynamic wind turbine models in power system simulation tool DlgSILENT*.
- Hsueh, K.-S. & Mosleh A. 1996. "The development and application of the accident dynamic simulator for dynamic probabilistic risk assessment of nuclear power plants" *Reliability Engineering & System Safety* 52(3): 297-314.
- Hu, Y., Groen, F., & Mosleh, A. 2004. "An entropy-based exploration strategy in dynamic PSA" *Probabilistic Safety Assessment and Management PSAM 7*:2391-2397.
- Hu, Y. 2005. "A guided simulation methodology for dynamic probabilistic risk assessment of complex systems" PhD Thesis, University of Maryland.
- Idaho National Laboratory, 2005. *RELAP5-3D© Code Manual*. Idaho Falls, Idaho.
- Kloos, M. & Peschke J. 2006. "MCDET: A probabilistic dynamics method combining Monte Carlo simulation with the discrete dynamic event tree approach" *Nuclear science and engineering* 153(2): 137-156.
- Labeau, P.-E., Smidts, C. & Swaminathan, S. 2000. "Dynamic reliability: towards an integrated platform for probabilistic risk assessment" *Reliability Engineering & System Safety* 68(3): 219-254.
- Lazo, A. V., & Rathie, P. 1978. "On the entropy of continuous probability distributions" *Information Theory, IEEE Transactions on* 24(1): 120-122.
- Loredo, T. J., & Chernoff, D. F. 2003. "Bayesian adaptive exploration" *Statistical challenges in astronomy*. Springer New York: 57-70.
- MacKay, D. 1992. "Information-based objective functions for active data selection" *Neural computation* 4(4): 590-604.
- Mercurio, D., Podofilini, L., Zio, E., & Dang, V. N. 2009. "Identification and classification of dynamic event tree scenarios via possibilistic clustering: application to a steam generator tube rupture event" *Accident Analysis & Prevention* 41(6): 1180-1191.
- Perez, M., Reventos, F., Batet, L., Guba, A., Tóth, I., Mieuxset, T., Bazin, P., De Crécy, A., Borisov, S., Skorek, T., Glaeser, H., Joucla, J., Probst, P., Ui, A., Chung, B.D., Oh, D.Y., Pernica, R., Kyncl, M., Macek, J., Manera, A., Freixa, J., Petrucci, A., D'Auria F. & Del Nevo, A. 2011. "Uncertainty and sensitivity analysis of a LBLOCA in a PWR Nuclear Power Plant: Results of the Phase V of the BEMUSE programme" *Nuclear Engineering and Design* 241(10): 4206-4222.
- Robert, C. P. & Casella G. 2005. *Monte Carlo statistical methods*. Springer Berlin.
- Rutt, B., Catalyurek, U., Hakobyan, A., Metzroth, K., Aldemir, T., Denning, R., Dunagan, S. & Kunsman, D. 2006. "Distributed dynamic event tree generation for reliability and risk assessment" *Challenges of Large Applications in Distributed Environments, IEEE*: 61-70.
- Shannon, C. E. 2001. "A mathematical theory of communication" *ACM SIGMOBILE Mobile Computing and Communications Review* 5(1): 3-55.
- Siu, N. 1994. "Risk assessment for dynamic systems: an overview" *Reliability Engineering & System Safety* 43(1): 43-73.
- Smidts, C. & Devooght, J. 1992. "Probabilistic reactor dynamics II: A Monte Carlo study of a fast reactor transient" *Nuclear science and engineering* 111(3): 241-256.
- Zio, E. 2014. "Integrated deterministic and probabilistic safety assessment: Concepts, challenges, research directions" *Nuclear Engineering and Design* 280: 413-419.