

Entropy-driven Monte Carlo simulation method for approximating the survival signature of complex infrastructures

Francesco Di Maio^a, Chiara Pettorossi^a, Enrico Zio^{a,b}

^aEnergy Department, Politecnico di Milano, Via La Masa 34, Milano, 20156, Italy

^bMINES Paris-PSL, CRC, 06560 Sophia Antipolis, France

Abstract: The reliability of critical infrastructures, such as power distribution networks, is of key importance for modern societies. The reliability of such complex systems can, in principle, be assessed by Monte Carlo simulation. However, the size and complexity of these systems, and the rarity of the failure events, can make the calculations quite demanding. Survival signature can help to address this issue, as it allows modelling the structure of the system separately from the probabilistic modelling for the reliability assessment. However, survival signature calculation suffers from the curse of dimensionality. Then, in this work, we propose a novel approach to approximate the survival signature of a system, which stands on the use of entropy to drive the sampling by Monte Carlo simulation (MCS) towards non-trivial system structure configurations, so as to save computational cost. The approach is exemplified by calculating the reliability of a generic synthetic multi-component network and the feasibility of its application is shown on a real-world network.

Keywords: Reliability, Survival signature, Critical infrastructures, Monte Carlo simulation, Entropy

1. INTRODUCTION

Infrastructures, such as power grids and transportation networks, are systems with complex network topology whose reliability assessment can be done by traditional methods, such as fault trees or reliability block diagrams, only if the size is relatively small or if the system can be decomposed in simple parts [1]. For realistically large networks instead, the reliability problem complexity increases exponentially with the network size [2]. To cope with this problem, Monte Carlo Simulation (MCS) can be used to generate a large number of replicas of the system random process of transition across its reachable states, recording system failures to eventually estimate its reliability [3]. However, systems with large numbers of reliable components require large computational efforts to obtain accurate estimates, and this can make the reliability problem practically intractable [4].

In recent years, the concept of system signature has been introduced to assist the reliability assessment of large systems made up of many components that fail with independent and identically distributed (*iid*) failure times [5]. Generalization of the signature to systems with multiple types of components (i.e., not *iid*) has been obtained by the adoption of the concept of survival signature [6], which is attractive for application to real-world networks. The advantage of the survival signature in reliability assessment is that it allows the separation of the system structure from the probabilistic information. This facilitates the consideration of dependencies, common cause failures and imprecisions in components characterization, offering a basis for analyzing complex networks efficiently [7]. Simulation approaches have been successfully applied to approximate the probabilistic time-dependent part of the reliability, both in the case of repairable and non-repairable components [8]. For the structural part instead, the approaches proposed over the years (including binary decision diagrams [9]), perform well only for systems with a limited number of components, whereas they fail as the network size increases.

Recently, percolation theory in combination with MCS has been proposed [10] to evaluate the survival signature of systems for which the full combinatorial evaluation would be intractable. Percolation theory is useful to study the macroscopic failure behavior of network systems in relation to the microscopic states of their components, denoted as nodes, by searching the connections between the

two end-nodes of the network [11]. The failure of a node (i.e., component) is modeled by its removal from the network and, as the number of removed nodes increases, the network undergoes the transition from being connected to being disconnected. Nevertheless, the computational cost related to MCS remains too high when dealing with large numbers of non-trivial network configurations, especially when the number of components and component types that make up the complex system increases [10]. In this paper, we present an automatic procedure for the approximation of the survival signature which combines percolation theory with entropy-driven MCS [12]. Once the values of the survival signature for trivial configurations are found by percolation, MCS is driven by the entropy function to carry out the survival signature assessment. Specifically, entropy is used to drive the simulation process towards the most unknown regions of the survival signature, so as to efficiently allocate the simulation efforts in the procedure of approximation [13].

The remainder of the paper is structured as follows. Section 2 presents an overview of the concept of survival signature and survival function. In Section 3, the entropy-driven MCS method for the survival signature approximation is presented. The application of the method is illustrated in Section 4 on a synthetic generic network and, then, applied to a real-world network, specifically, the electricity transmission network of Great Britain. In Section 5 some conclusions are drawn.

2. SURVIVAL SIGNATURE

Let $\underline{x} = (x_1, x_2, \dots, x_m) \in \{0,1\}^m$ be a Boolean vector describing the state of a system of m components, where each entry is $x_i = 1$ if the i -th component functions or $x_i = 0$ if not. The system reliability structure function is denoted as $\varphi : \{0,1\}^m \rightarrow \{0,1\}$, and it is defined for all 2^m possible \underline{x} , such that $\varphi(\underline{x}) = 1$ if the system functions and $\varphi(\underline{x}) = 0$ if the system does not function in correspondence of state vector \underline{x} . We assume that the system is coherent, i.e., the structure function is not decreasing if the number of functioning components increases, which also implies $\varphi(\underline{0}) = 0$ and $\varphi(\underline{1}) = 1$, i.e., the system functions if all components function and it is failed if all components are failed. Let S_l denote the set of state vectors \underline{x} with exactly l of the m components functioning, so that $\sum_{i=1}^m x_i = l$.

The survival signature $\Phi(l)$, with $l = \{0, 1, \dots, m\}$, is defined in Eq. (1) as the probability that the system functions given that exactly l components function:

$$\Phi(l) = \binom{m}{l}^{-1} \sum_{\underline{x} \in S_l} \varphi(\underline{x}) \quad (1)$$

If the m components are partitioned into $K \geq 2$ different types, with m_k components of the specific type k , so that $\sum_{k=1}^K m_k = m$, the state vector can be written as $\underline{x} = (\underline{x}^1, \underline{x}^2, \dots, \underline{x}^K)$, with $\underline{x}^k = (x_1^k, x_2^k, \dots, x_{m_k}^k)$ representing the states of the components of type k . The survival signature becomes $\Phi(l_1, l_2, \dots, l_K)$, as in Eq. (2), with $l_k = 0, 1, \dots, m_k$ for $k = 1, 2, \dots, K$, and it is defined as the probability that the system functions given that l_k out of m_k components of type k function, for each $k = 1, 2, \dots, K$ [6]:

$$\Phi(l_1, l_2, \dots, l_K) = \left[\prod_{k=1}^K \binom{m_k}{l_k} \right]^{-1} \times \sum_{\underline{x} \in S_{l_1, l_2, \dots, l_K}} \varphi(\underline{x}) \quad (2)$$

Since the state vectors \underline{x}^k with exactly l_k components functioning are $\binom{m_k}{l_k}$, where $l_k = \sum_{i=1}^{m_k} x_i^k$, the set of all the allowed combinations of all components is denoted as S_{l_1, l_2, \dots, l_K} , with magnitude $\prod_{k=1}^K \binom{m_k}{l_k}$; also, assuming that the failure times of the m_k components of type k are *iid*, all the state vectors $\underline{x}^k \in S_{l_k}$ are equally likely to occur.

In practice, the analytical calculation of Eq. (2) for the entries (l_1, l_2, \dots, l_K) requires the structure function $\varphi(\underline{x})$ to be evaluated $\binom{m_1}{l_1} \cdot \binom{m_2}{l_2} \cdot \dots \cdot \binom{m_K}{l_K}$ times, leading to 2^m structure function evaluations in total, a combinatorial explosion giving rise to a potentially significant computational burden for large systems. Eventually, the generalized survival signature is a multidimensional array with size $(m_1 + 1) \times (m_2 + 1) \times \dots \times (m_K + 1)$ (including the case $l_k = 0$ in which none of the components of type k are working).

Considering the probability that the system functions at time $t > 0$, the survival function of the system, i.e., the reliability, can be calculated as in Eq. (3) [6]:

$$P(T_s > t) = \sum_{l_1=0}^{m_1} \sum_{l_2=0}^{m_2} \dots \sum_{l_K=0}^{m_K} \Phi(l_1, l_2, \dots, l_K) P\left(\bigcap_{k=1}^K \{C_t^k = l_k\}\right) \quad (3)$$

where $C_t^k \in \{0, 1, \dots, m_k\}$ is the number of components of type k that function at time $t > 0$, described by the probability $P(C_t^k)$. It is important to notice that in Eq. (3) the structure of the system is kept separated from the probabilistic information (i.e., the failure time distributions of the components), which is the main advantage of the survival signature approach. Indeed, even if the survival signature needs to be calculated for all $\prod_{k=1}^K (m_k + 1)$ different $\underline{l} = (l_1, l_2, \dots, l_K)$ vectors, this calculation has to be done only once and, then, coupled with the probabilistic part. If the cumulative distribution function (CDF) $F_k(t)$ for component type k is known, and the failure times of the different component types are independent, the probabilistic part can be simplified as in Eq. (4):

$$P\left(\bigcap_{k=1}^K \{C_t^k = l_k\}\right) = \prod_{k=1}^K P(C_t^k = l_k) = \dots = \prod_{k=1}^K \left(\binom{m_k}{l_k} F_k(t)^{l_k} [1 - F_k(t)]^{m_k - l_k} \right) \quad (4)$$

The method proposed in this work addresses the quantification of the structure part of the survival function, $\Phi(l_1, l_2, \dots, l_K)$, whereas for what concerns the probabilistic part, $P(\bigcap_{k=1}^K \{C_t^k = l_k\})$, the reader is invited to refer to [4], [14]–[16] for efficient Monte Carlo sampling techniques, such as line sampling, subset simulation and Adaptive Kriging Monte Carlo Sampling (AK-MCS), applied for generating random component failure times.

3. ENTROPY-DRIVEN MONTE CARLO SIMULATION METHOD

In this Section, a novel method for the survival signature approximation of large infrastructure networks is presented. An overview of the method is shown in Fig. 1, in terms of the main steps of *i)* input definition, *ii)* percolation theory application and *iii)* entropy-driven MCS application.

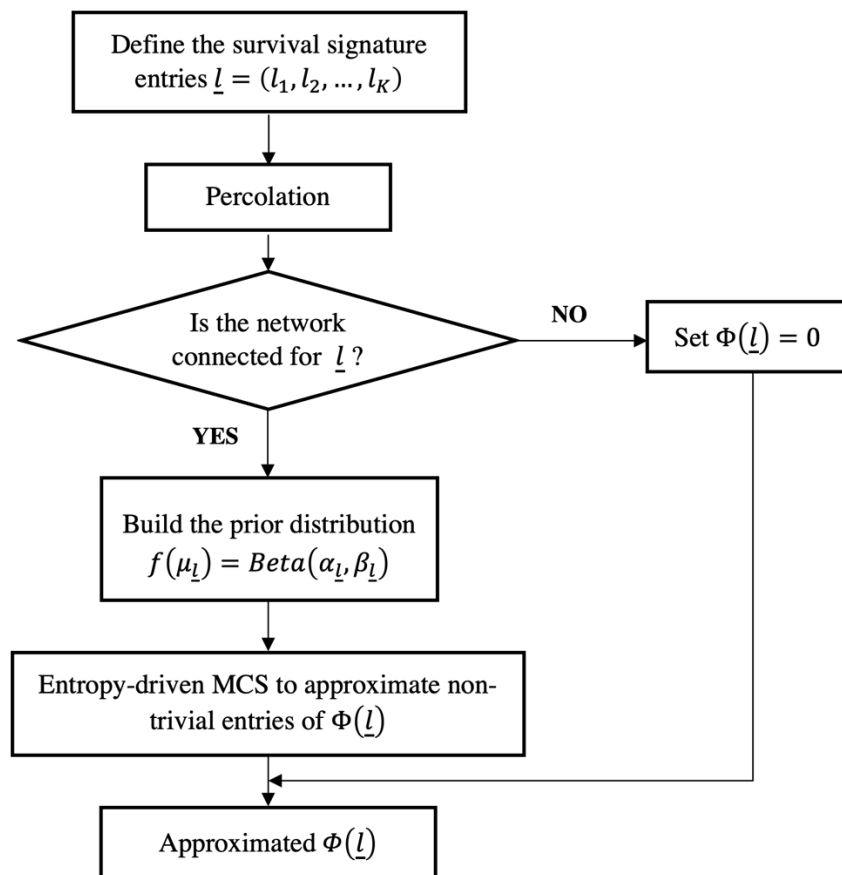


Figure 1 - Flowchart of the methodology proposed for the survival signature approximation.

3.1. Input definition

The survival signature has to be computed for each entry (l_1, l_2, \dots, l_K) , which means for each combination of the number of working components among the K types of components. To evaluate the system state for each $\underline{x} \in S_{l_1, l_2, \dots, l_K}$, for each (l_1, l_2, \dots, l_K) , the structure function, which is a Boolean function incorporating all the path sets of functioning nodes/components for which the system functions [17], is to be used. It is worth noting that the evaluation of the system state for an entry \underline{l} defines a

Bernoulli process, since the structure function can either result in a functioning state for the system, i.e., $\varphi(\underline{x}) = 1$, or in a non-functioning state, i.e., $\varphi(\underline{x}) = 0$, and so:

$$p(\underline{l}) \sim \text{Bernoulli}(\mu_{\underline{l}}) \quad (5)$$

where $\underline{l} = (l_1, l_2, \dots, l_K)$. In this way to every \underline{l} is associated a probability distribution $p(\underline{l})$, with parameter $\mu_{\underline{l}}$. In general, $\mu_{\underline{l}}$ is an uncertain parameter, described in a Bayesian framework by a prior Beta distribution, i.e., $\mu_{\underline{l}} \sim \text{Beta}(\alpha_{\underline{l}}, \beta_{\underline{l}})$ [18]. It is straightforward that the parameter $\mu_{\underline{l}}$ corresponds to the survival signature of Eq. (4) with entry \underline{l} , $\Phi(\underline{l}) = \mu_{\underline{l}}$, and, as shown in [18], $\alpha_{\underline{l}}$ represents the number of successes, i.e., the number of times the system is functioning, whereas $\beta_{\underline{l}}$ represents the number of failures, i.e., the number of times that the system is not functioning.

3.2. Percolation theory

Percolation theory is used to calculate the trivial solutions of Eq. (2), i.e., those for which the number of non-functioning components is sufficient to determine a non-functioning network. The percolation process consists in:

- computing the critical fraction of non-functioning components f_c for which the network makes a transition from functioning to non-functioning, as follows:

$$f_c = 1 - \frac{1}{k-1} \quad (6)$$

where $k = \langle d^2 \rangle / \langle d \rangle$, d is the node degree, $\langle d \rangle$ and $\langle d^2 \rangle$ are the first and second moments of the degree distribution for the network (this information is usually retrievable from the network adjacency matrix defining the connections among the network nodes), and

- finding the trivial entries that are those for which there is no network connectivity and whose survival signature can be set to zero, here taken as the entries of the survival signature for which the fraction of functioning components is lower than $1 - f_c$, i.e., those satisfying Eq. (7):

$$\sum_{k=1}^K l_k < (1 - f_c) \cdot \sum_{k=1}^K m_k \implies \Phi(\underline{l}) \approx 0 \quad (7)$$

3.3. Entropy-driven Monte Carlo simulation

Entropy $I(\mu)$ is defined as in Eq. (8), [12], where $f(\mu)$ is a prior probability distribution with uncertain parameter μ :

$$I(\mu) = E [\log(f(\mu))] = \int \log(f(\mu))f(\mu)d\mu \quad (8)$$

In a Bayesian framework, when a new evidence y is available for μ , the prior distribution $f(\mu)$ is updated to get the posterior distribution $f(\mu|y)$; thus, entropy becomes $I(\mu|y)$ and the information gain is [18]:

$$\Delta I(\mu|y) = I(\mu|y) - I(\mu). \quad (9)$$

which represents the increment of information for the characterization of that specific pdf. Due to stochasticity of the simulation process, it is preferable to consider the expected information gain $E[\Delta I(\mu|y)]$, instead of $\Delta I(\mu|y)$ [12].

Entropy is here used to evaluate the remaining entries \underline{l} . The entropy-driven MCS is performed as follows:

Step 1. Estimate the prior probability distributions $f(\mu_{\underline{l}})$ associated with the remaining entries

$\underline{l} = (l_1, l_2, \dots, l_K)$ by collecting a limited number N of random samples $\underline{x}_{\underline{l}} \in S_{l_1, l_2, \dots, l_K}$ of the components states vector and evaluating $\varphi(\underline{x}_{\underline{l}})$ that results in the system state when \underline{l} components are functioning. If, for a given entry \underline{l} , $\varphi(\underline{x}_{\underline{l}})$ is always the same, i.e., equal to 1, meaning that the system always reaches a functioning end-state, or equal to 0, vice versa,

the confidence on the prior distribution $\text{Beta}(\alpha_{\underline{l}}, \beta_{\underline{l}})$ is high [18] and we assume $\Phi(\underline{l}) = 1$ in the former case and $\Phi(\underline{l}) = 0$ in the latter. A pseudo-code for the retrieval of the prior distribution for each survival signature entry is shown in Fig. 2;

```

Require N: number of samples;  $\varphi$ : structure function;  $\underline{l} = (l_1, \dots, l_K)$  survival signature entry.

 $(\alpha, \beta) \leftarrow 0$                                      %Initialize beta distribution parameters
while  $n \leq N$                                        %Loop over number of samples to collect
     $n \leftarrow n + 1$                                  %Update number of samples
     $\underline{x}_{\underline{l}}$                                %Random state vector for  $\underline{l}$ 
    if  $\varphi(\underline{x}_{\underline{l}}) = 1$ 
         $\alpha \leftarrow \alpha + 1$                    %Update functioning configurations counter
    elseif  $\varphi(\underline{x}_{\underline{l}}) = 0$ 
         $\beta \leftarrow \beta + 1$                    %Update non-functioning configurations counter
    end
end

```

Figure 2 – Pseudocode of the pre-exploration algorithm (Step 1), to be performed for each non-trivial entry.

Step 2. For the remaining entries \underline{l} , compute the expected entropy gain of each distribution

$f(\mu_{\underline{l}}) = \text{Beta}(\alpha_{\underline{l}}, \beta_{\underline{l}})$ using Eq. (10), as shown in [18]:

$$E[\Delta I(\mu_{\underline{l}})] = L_W \cdot I(\mu_{\underline{l}} | \varphi(\underline{x}_{\underline{l}}) = 1) + L_F \cdot I(\mu_{\underline{l}} | \varphi(\underline{x}_{\underline{l}}) = 0) - I(\mu_{\underline{l}}) \quad (10)$$

$$L_W = \frac{\alpha_{\underline{l}}}{\alpha_{\underline{l}} + \beta_{\underline{l}}} \quad (11)$$

$$L_F = \frac{\beta_{\underline{l}}}{\alpha_{\underline{l}} + \beta_{\underline{l}}} \quad (12)$$

Step 3. Select the entry \underline{l}^* with the largest value of the expected information gain:

$$\underline{l}^* = \underset{\underline{l}}{\operatorname{argmax}} E[\Delta I(\mu_{\underline{l}})] \quad (13)$$

and collect a random sample $\underline{x}_{l^*} \in S_{l_1^*, l_2^*, \dots, l_K^*}$ from its probability distribution $p(\underline{l}^*)$, so that $\varphi(\underline{x}_{l^*}) = 1$ or $\varphi(\underline{x}_{l^*}) = 0$;

Step 4. Calculate the posterior distribution $f(\mu_{l^*} | \varphi(\underline{x}_{l^*}))$, whose parameters become:

- $(\alpha_{l^*}, \beta_{l^*}) = (\alpha_{l^*} + 1, \beta_{l^*})$ if the system is functioning,
- $(\alpha_{l^*}, \beta_{l^*}) = (\alpha_{l^*}, \beta_{l^*} + 1)$ if the system is non-functioning,

and update the corresponding expected entropy gain $E[\Delta I(\mu_{l^*} | \varphi(\underline{x}_{l^*}))]$ by means of Eq. (10). As we shall see in what follows, every piece of evidence (or sample) collected narrows the resulting posterior distribution: as the information content increases, the entropy value increases and the expected entropy gain decreases. Fig. 3 shows an example of expected entropy gain decreasing with increasing parameters of the Beta distribution.

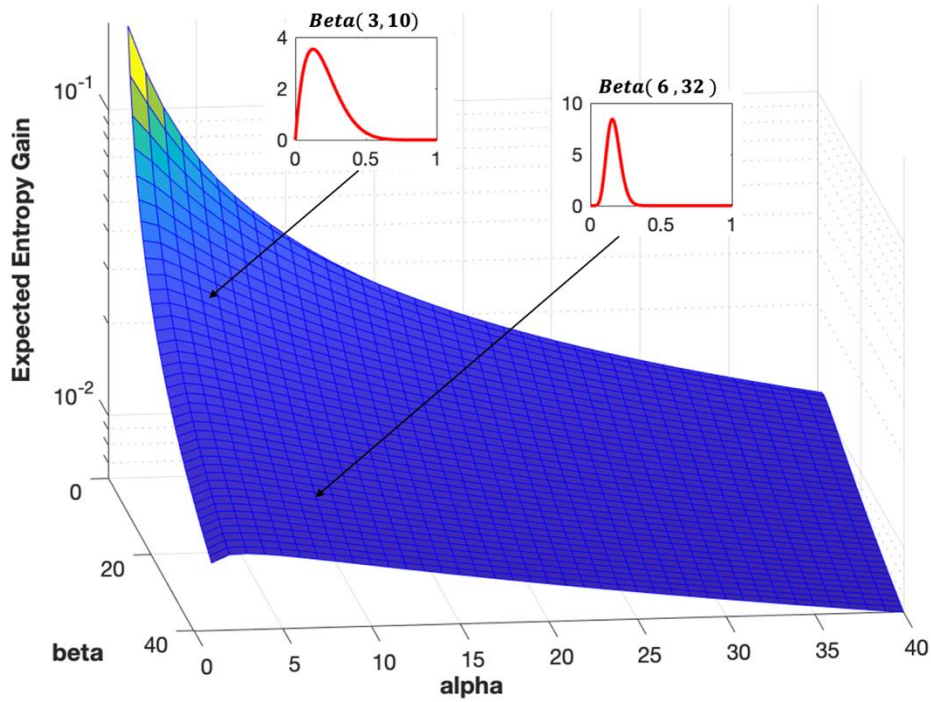


Figure 3 – Expected entropy gain for a generic Beta distribution.

Step 5. Compute the survival signature for the entry \underline{l}^* , $\Phi(\underline{l}^*)$, using Eq. (14):

$$\mu_{\underline{l}} = \frac{\alpha_{\underline{l}}}{\alpha_{\underline{l}} + \beta_{\underline{l}}} = \Phi(\underline{l}) \quad (14)$$

where $\alpha_{\underline{l}} + \beta_{\underline{l}} = N_{\underline{l}}$ is the number of instances of $\underline{x}_{\underline{l}} \in S_{\underline{l}}$ collected for the specific entry \underline{l} ;

Step 6. Compute the variance of the estimate of the survival signature for entry \underline{l}^* as in Eq.

(15):

$$V_{\underline{l}} = \frac{1}{N_{\underline{l}} - 1} \sum_{i=1}^{N_{\underline{l}}} |\Phi_i(\underline{l}) - \bar{\Phi}(\underline{l})|^2 \quad (15)$$

where $\bar{\Phi}(\underline{l})$ is the mean of $\Phi(\underline{l})$.

Repeat Steps 2-6 until $V_{\underline{l}}$ is smaller than V , a threshold value set by the analyst. A pseudo-code for the implementation of the Steps 2-6 is shown in Fig. 4.

Require \underline{l} : set of survival signature entries to approximate; (α, β) : prior beta distribution parameters;
 φ : structure function; V : variance threshold

```

gain  $\leftarrow$  entropy_gain ( $\alpha, \beta$ )           % Compute expected entropy gain for all entries

while  $v \leq V$ 
     $\underline{l}^* \leftarrow \max(\mathbf{gain})$            % Select next entry to investigate
     $\underline{x}_{\underline{l}^*}$            % Generate random state vector
    if  $\varphi(\underline{x}_{\underline{l}^*}) = 1$ 
         $\alpha^* \leftarrow \alpha^* + 1$ 
    elseif  $\varphi(\underline{x}_{\underline{l}^*}) = 0$ 
         $\beta^* \leftarrow \beta^* + 1$ 
    end

     $gain(\underline{l}^*) \leftarrow \text{entropy\_gain}(\alpha^*, \beta^*)$            % Update gain for current entry
     $n^* \leftarrow \alpha^*, \beta^*$            % Update number of samples for current entry
     $\Phi(\underline{l}^*) \leftarrow \alpha^*/n^*$            % Update survival signature for current entry
     $v(\underline{l}^*) \leftarrow \text{var}(\Phi(\underline{l}^*))$            % Evaluate stopping criteria
     $\Phi_{prior}(\underline{l}^*) \leftarrow \Phi(\underline{l}^*)$            % Update prior estimate for survival signature
end

```

Figure 4 – Pseudocode of the entropy-driven MCS (Steps 2-6).

4. STUDY CASES

4.1. The synthetic multi-component network

We consider the network in Fig. 5 that has $m = 6$ components, of $K = 2$ types, $k = 1, 2$, and specifically $m_1 = m_2 = 3$ components of each type. The failure times of components of type 1 follow an exponential distribution with $\lambda = 1$, in arbitrary units of inverse time, whereas the failure times of components of type 2 follow a Weibull distribution with scale parameter $a = 1$ and shape parameter $b = 2$.

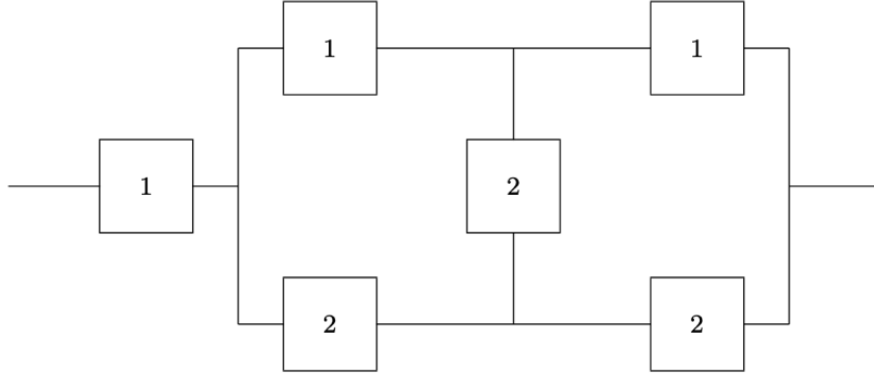


Figure 5 – The synthetic multi-component network considered as case study (the numbers indicate the components types).

The survival signature $\Phi(l_1, l_2)$ for all $(l_1, l_2) \in \{0, 1, 2, 3\}$ (which indicate the number of functioning components of type $k = 1$ and $k = 2$, respectively) is given in matrix form, whose dimension is $(m_1 + 1) \times (m_2 + 1) = (3 + 1) \times (3 + 1) = 4 \times 4$, in Table 1. This result can be calculated analytically using Eq. (2), once the structure function of the system is available, and it is used hereafter as benchmark for the crude MCS and the entropy-driven MCS approach proposed.

Table 1 - Survival signature for the network in Fig. 5.

$l_1 \backslash l_2$	0	1	2	3
0	0	0	0	0
1	0	0	1/9	1/3
2	0	0	4/9	2/3
3	1	1	1	1

To proceed with the entropy-driven MCS approach, all the possible combinations of $(l_1, l_2) \in \{0, 1, 2, 3\}$ are defined, and the critical fraction of non-functioning components f_c for percolation is computed, resulting in $f_c = 0.6$. By means of Eq. (7), the identified trivial entries are $(l_1, l_2) = \{(0,0); (0,1); (0,2); (1,0); (1,1); (2,0)\}$. For the remaining non-trivial entries, the probability distributions $p(l_1, l_2)$ are quantified within the pre-exploration step (Step 1), in which $N = 20$ samples

are collected for each entry and used to obtain the informative distributions shown in Fig. 6. The entries $(l_1, l_2) = \{(0,3); (2,1); (3,0); (3,1); (3,2); (3,3)\}$ result in a consistent outcome, peaked in either 0 or 1: the corresponding survival signature is assigned equal to 0 for $(l_1, l_2) = \{(0,3); (2,1)\}$ and equal to 1 for $(l_1, l_2) = \{(3,0); (3,1); (3,2); (3,3)\}$.

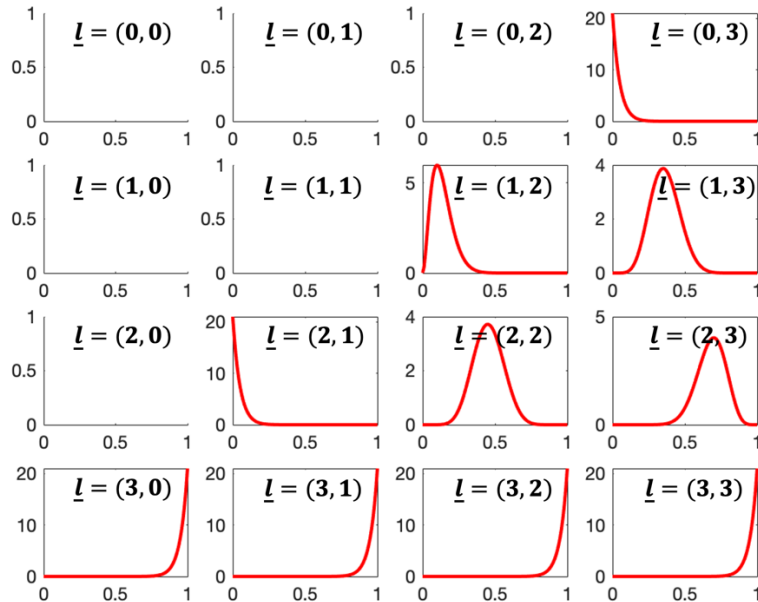


Figure 6 - Beta distributions for every survival signature entry after pre-exploration (Step 1).

Steps 2-6 are, thus, performed setting $V = 0.005$ for the remaining $(l_1, l_2) = \{(1,2); (1,3); (2,2); (2,3)\}$. $V = 0.005$ guarantees that the survival signature estimate reaches a good level of convergence, as shown in Fig. 7 for the entry $(l_1, l_2) = (1,3)$.

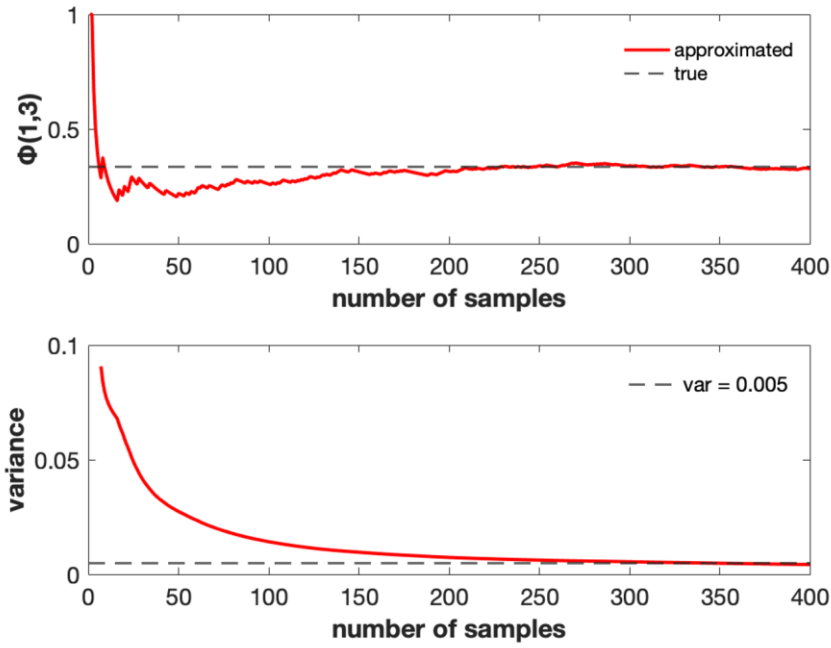


Figure 7 – $\Phi(1,3)$ vs number of samples (top). Variance of $\Phi(1,3)$ vs number of samples (bottom).

The evolution of the entropy-driven MCS for the four remaining entries, in terms of expected entropy gain and parameters α and β , is illustrated in Fig. 8 (left), where each stepwise line is referring to the simulation of a specific entry and it oscillates about a straight line that represents the $\alpha - \beta$ ratio corresponding to the exact survival signature for the entry, explicitly shown in Fig. 8 (right) for the entry $(l_1, l_2) = (1,3)$. The oscillations are related to the randomness of the simulation process.

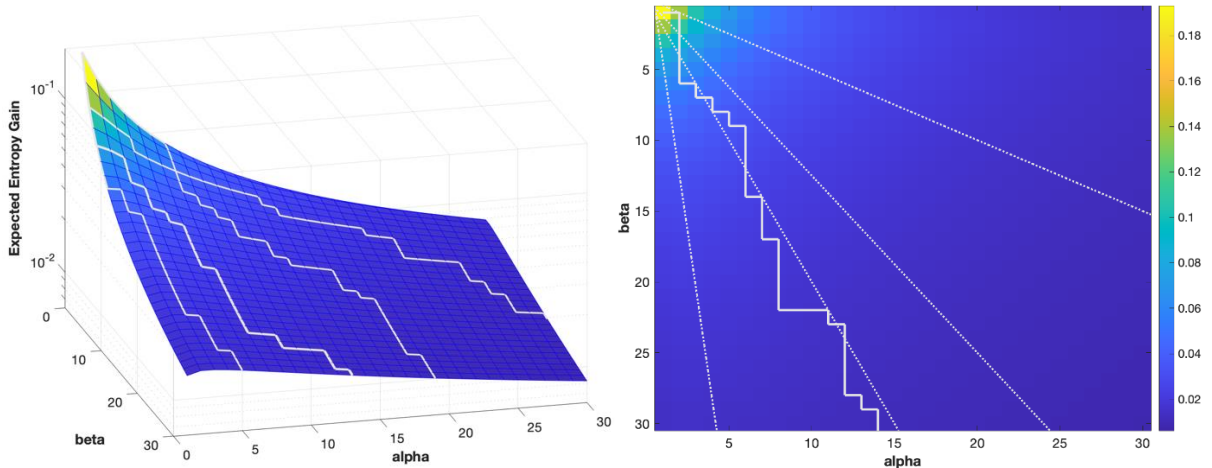


Figure 8 – Expected entropy gain evolution through the entropy-driven MCS as function of the Beta distribution parameters, 3D (left) and 2D (right). On the right, the continuous line represents the evolution for entry $(l_1, l_2) = (1,3)$.

The survival signature approximation resulting from the application of the entropy-driven MCS method is given in Table 2.

Table 2 – Survival signature for the network in Fig. 5 approximated by entropy-driven MCS.

$l_1 \setminus l_2$	0	1	2	3
0	0	0	0	0
1	0	0	0.0951	0.3317
2	0	0	0.4434	0.6691
3	1	1	1	1

The result is benchmarked with that obtained by the crude MCS method that is separately performed and stopped when the variance is smaller than $V = 0.005$ for each entry (l_1, l_2) of the survival signature, as for the entropy-driven MCS approach. To further verify the results, in light, also, of the stochasticity of both the crude MCS method and the entropy-driven MCS method, the approximation of the survival signature has been performed 500 times; for each of these, the number of required samples for the approximation has been collected and its distribution plotted in the boxplot of Fig. 9: the median of the samples required to approximate the whole survival signature for the entropy-driven MCS method is 794.7, whereas for the crude MCS is 831.2; the maximum number of samples is 2498 for the entropy-driven MCS method, whereas it is 2811 for the crude MCS.

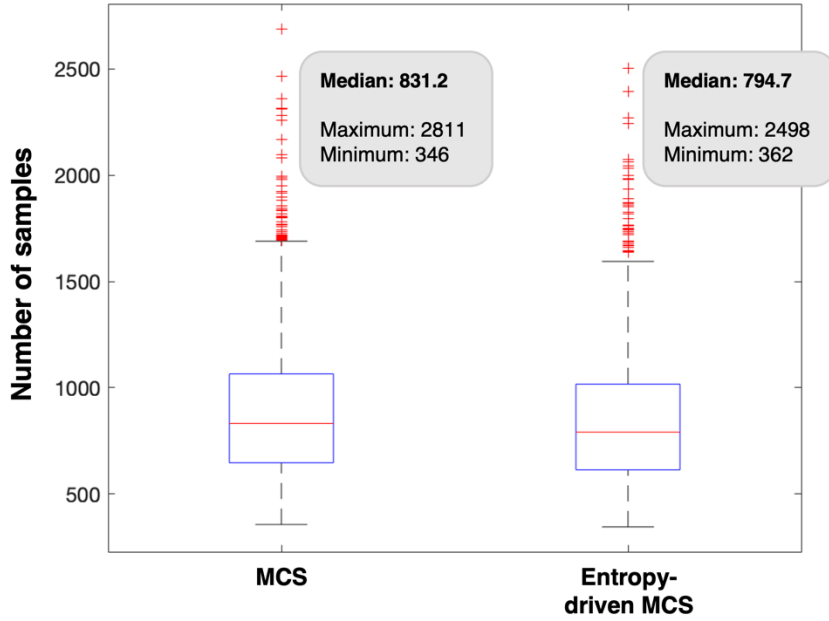


Figure 9 – Boxplot for the number of samples required to approximate the survival signature using the crude MCS method (left) and the entropy-driven MCS method (right).

Not only the computational savings are confirmed, but also the accuracy of the approximation is proven, as shown in Table 3, where the survival signatures computed analytically using Eq. (2), approximated by means of the crude MCS and approximated by the entropy-driven MCS are listed: the average relative error over the estimate provided by the entropy-driven MCS is 3.87%, smaller than the error of 4.11% obtained by the crude MCS.

Table 3 – Survival signature of the network of Fig. 5 computed analytically with Eq. (2) (left), approximated by means of crude MCS (center), and approximated by entropy-driven MCS (right). The average relative error for the approximated survival signature is given.

l_1	l_2	$\Phi(l_1, l_2)$ - Analytical	$\Phi(l_1, l_2)$ - MCS	$\Phi(l_1, l_2)$ - Entropy-driven MCS
0	[0,1,2,3]	0	0	0
[1,2]	[0,1]	0	0	0
1	2	0.1111	0.0945	0.0951
1	3	0.3333	0.3303	0.3317
2	2	0.4444	0.4417	0.4434
2	3	0.6667	0.6623	0.6691
3	[0,1,2,3]	1	1	1

Average relative

error [%]:

-

4.11

3.87

The system reliability can be finally estimated by means of the survival function, as in Eq. (3):

$$P(T_s > t) = \sum_{l_1=0}^3 \sum_{l_2=0}^3 \Phi(l_1, l_2) \binom{3}{l_1} [1 - e^{-t}]^{3-l_1} [e^{-t}]^{l_1} \binom{3}{l_2} [1 - e^{-t^2}]^{3-l_2} [e^{-t^2}]^{l_2} \quad (16)$$

In Fig. 10, the reliability (i.e., the survival function) for a mission time $T_m = 2$ years is shown for the analytical solution (continuous line), the crude MCS method (crosses) and the entropy-driven MCS (dots).

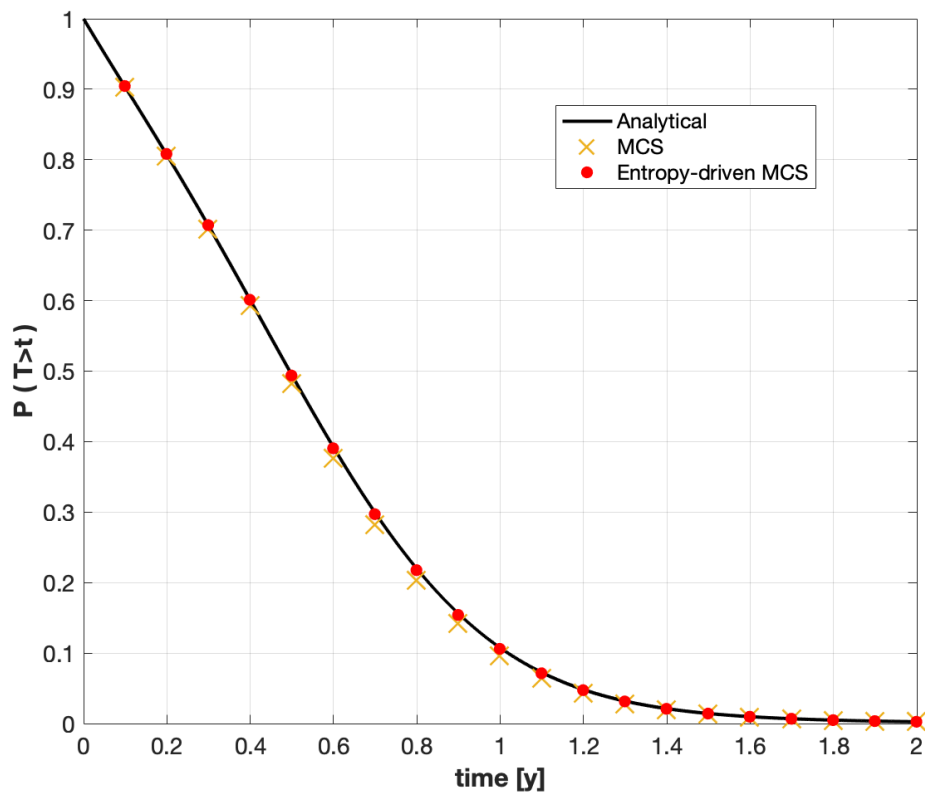


Figure 10 – Survival function of the system in Fig. 5 computed with: analytical survival signature (continuous line), survival signature approximated by the crude MCS method (crosses) and survival signature approximated by the entropy-driven MCS method (dots).

4.2. The electricity transmission network of Great Britain

We consider the topological model of the electricity transmission network of Great Britain shown in Fig. 11 [19]. The network comprises of $m = 29$ components, of $K = 2$ types, $k = 1, 2$, $m_1 = 5$ and $m_2 = 24$. We assume that the failure times of components of type 1 follow an exponential distribution with $\lambda = 1$, in arbitrary units of inverse time, and the failure times of components of type 2 follow a Weibull distribution with parameters $a = 1$ and $b = 2$.

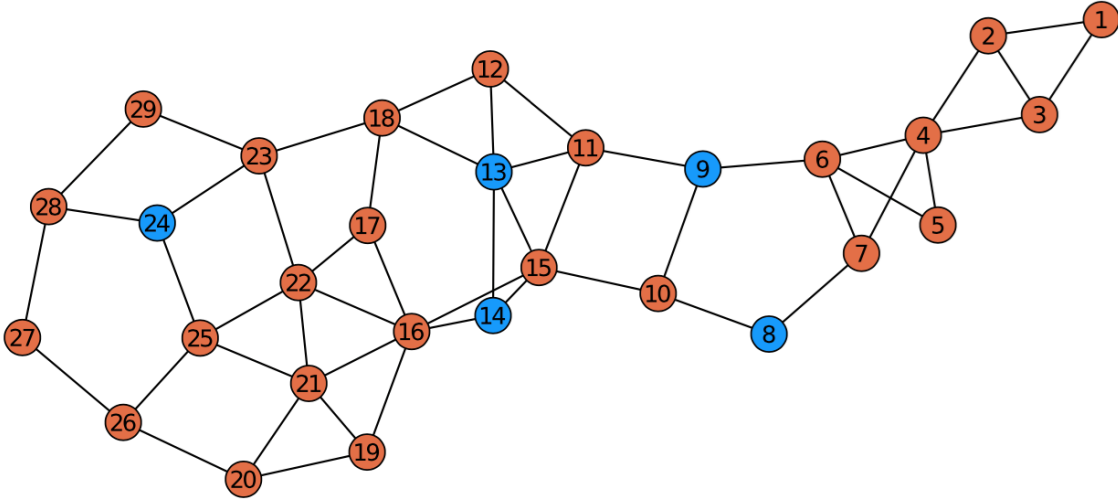


Figure 11 – Topology of the electricity transmission network of Great Britain. Components of type 1 are shown in blue, components of type 2 are shown in orange (adapted from [19]).

The network survival signature is an array with dimension 6×25 , and its exact, analytical, computation is provided in [10]. To determine the structure function for the network, the efficiency metric defined in [20], is used:

$$E(G) = \frac{1}{n(n-1)} \sum_{i \neq j \in G} \frac{1}{d_{i,j}} \quad (17)$$

where G is the network with n nodes and $d_{i,j}$ is the length of the shortest path between two nodes i and j , computed by means of the Floyd-Warshall algorithm [21]. To assess the network state, it is assumed that the network is non-functioning if the loss of efficiency due to the failing components exceeds 50%, as in Eq. (18):

$$\text{if } \frac{E(G(\underline{x}))}{E(G)} < 0.5 \Rightarrow \varphi(\underline{x}) = 0 \quad (18)$$

whereas the network is functioning, i.e., $\varphi(\underline{x}) = 1$, otherwise.

The entropy-driven MCS of Section 3 is applied to the network in Fig. 11. In particular, once the percolation threshold is computed ($f_c = 0.64$) and the trivial entries are solved, $N = 100$ samples are used in the pre-exploration, Step 1, and the entropy-driven MCS is performed for driving the simulation among the remaining entries, setting $V = 0.005$ as stopping criterion.

The survival signature is given in Table 4, where the exact (analytical) survival signature for the network is listed, along with the results provided by the crude MCS and the entropy-driven MCS approach.

Table 4 – Survival signature of the network of Fig. 11 computed analytically with Eq. (2) (left), approximated by means of crude MCS (center), and approximated with entropy-driven MCS (right). The average relative error for the approximated results is given.

l_1	l_2	$\Phi(l_1, l_2)$ - Analytical	$\Phi(l_1, l_2)$ - MCS	$\Phi(l_1, l_2)$ - Entropy- driven MCS
[0-5]	[0-12]	0	0	0
0	[13-22]	0	0	0
1	[13-17]	0	0	0
2	[13-15]	0	0	0
3	[13-14]	0	0	0
4	13	0	0	0
5	13	1.202e-06	0	0
4	14	1.835e-06	0	0
5	14	5.155e-04	1.998e-05	1.998e-05
3	15	1.682e-06	0	0
4	15	3.954e-04	0	1.998e-05
5	15	0.0165	0.0097	0.0103
2	16	2.712e-07	0	0
3	16	1.796e-04	9.99e-06	0
4	16	0.0147	0.0083	0.0073
5	16	0.1492	0.1228	0.1168
2	17	3.265e-05	0	0

3	17	0.0104	0.0042	0.0047
4	17	0.1497	0.1197	0.1199
5	17	0.4631	0.4158	0.4192
1	18	1.486e-06	0	0
2	18	0.0055	0.0018	0.0015
3	18	0.1387	0.1091	0.1097
4	18	0.4550	0.4142	0.4149
5	18	0.7371	0.7028	0.7051
1	19	0.0017	2.276e-04	8.869e-05
2	19	0.1150	0.0857	0.0899
3	19	0.4358	0.3979	0.4018
4	19	0.7284	0.6938	0.6916
5	19	0.9129	0.8863	0.8942
1	20	0.0738	0.0518	0.0538
2	20	0.3807	0.3469	0.3502
3	20	0.7187	0.6700	0.6772
4	20	0.9129	0.8873	0.8849
5	20	0.9890	0.9789	0.9793
1	21	0.2482	0.2327	0.2340
2	21	0.6734	0.6151	0.6204
3	21	0.9281	0.8938	0.8922
4	21	0.9924	0.9845	0.9843
5	21	0.9995	0.9998	0.9996
1	22	0.5529	0.4371	0.4393
2	22	0.9435	0.9044	0.9051
3	22	0.9971	0.9944	0.9919
0	23	0.2917	0	0
1	23	0.9667	0.9224	0.9270
[4-5]	22	1	1	1
[2-5]	23	1	1	1
[0-5]	24	1	1	1

Average relative error [%]:	-	10.08	10.10
------------------------------------	---	-------	-------

To further compare the performance of the entropy-driven MCS method with that of the crude MCS the procedure is repeated 100 times: for each of these, the number of samples are collected and plotted

in the boxplots of Fig. 12. The median of the number of samples required for the approximation with the entropy-driven MCS is 1.620×10^4 , whereas with the crude MCS it is 1.903×10^4 ; the maximum number of samples is 1.896×10^4 for the entropy-driven MCS method, and 2.169×10^4 for the crude MCS. Again, the computational savings are confirmed and the average relative errors are comparable, being 10.10% for the entropy-driven MCS and 10.08% for the crude MCS.

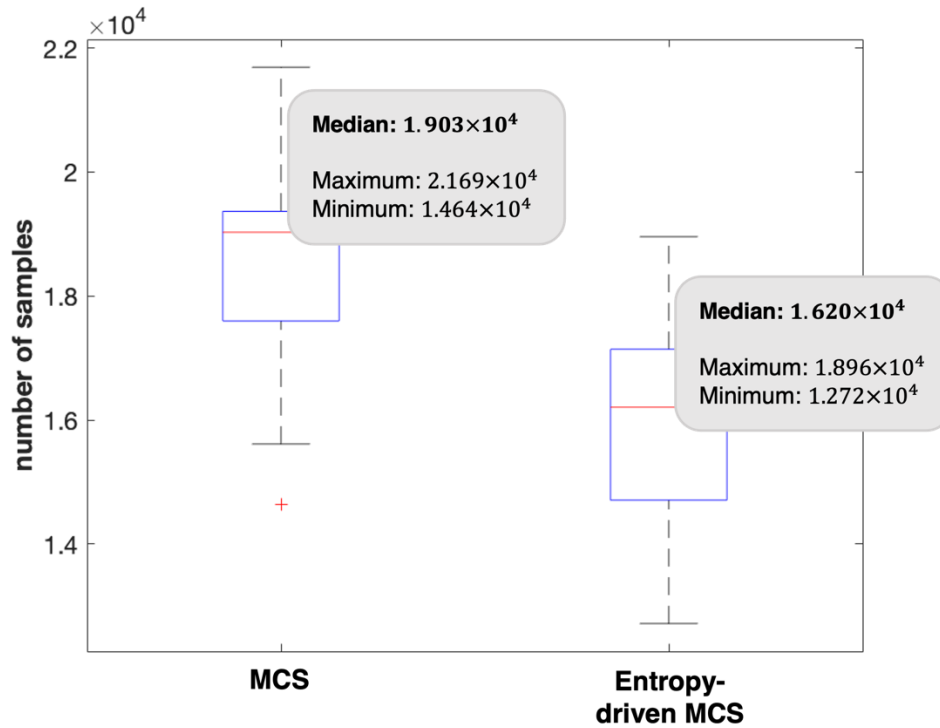


Figure 12 – Boxplot for the number of samples required to approximate the survival signature for the network in Fig. 11 using the crude MCS method (left) and the entropy-driven MCS method (right).

Eventually, the network reliability is computed by means of the survival function, Eq. (3):

$$P(T_s > t) = \sum_{l_1=0}^5 \sum_{l_2=0}^{24} \Phi(l_1, l_2) \binom{5}{l_1} [1 - e^{-t}]^{5-l_1} [e^{-t}]^{l_1} \binom{24}{l_2} [1 - e^{-t^2}]^{24-l_2} [e^{-t^2}]^{l_2} \quad (17)$$

In Fig. 13 the network reliability for a mission time $T_m = 1$ year is given for the analytical survival signature (continuous line), the survival signature computed by means of the crude MCS method (crosses) and by means of the entropy-driven MCS (dots).

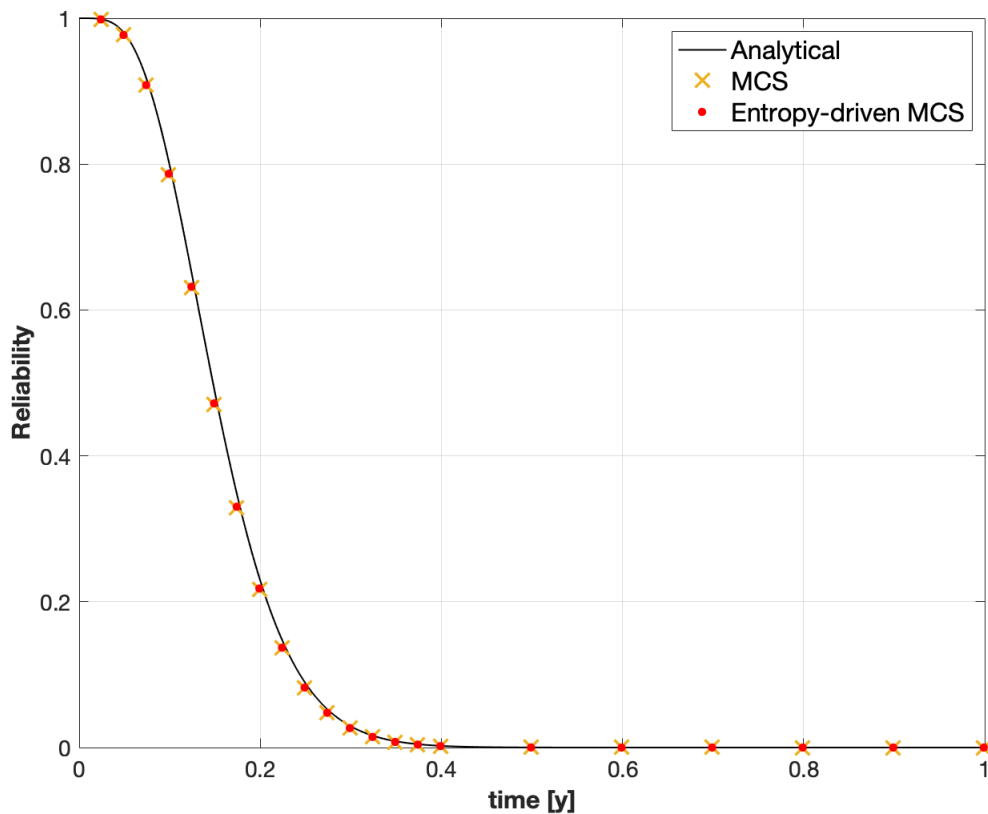


Figure 13 – Survival function of the system in Fig. 11 computed with: analytical survival signature (continuous line), survival signature approximated by the crude MCS method (crosses) and survival signature approximated by the entropy-driven MCS method (dots).

A satisfying agreement between the curve obtained using the approximated survival signature and the one computed analytically is found for both the entropy-driven MCS and the crude MCS, which demonstrates that the method proposed is suitable.

5. CONCLUSIONS

The survival signature has proven to be a practical method for the reliability assessment of complex systems with multiple types of components. However, methods for its exact computation are applicable only to small networks and efficient simulation methods are needed to provide an approximation, for enabling application to large complex systems. A new method for efficiently computing the survival signature has been presented in this paper. The method is based on the entropy-driven MCS and its feasibility to reduce the computational cost while guaranteeing accuracy is proven on two networks, for which the analytical survival signatures can be evaluated and benchmarked with those approximated by

crude MCS and by entropy-driven MCS. The beneficial computational effect of the entropy-driven MCS method over the crude MCS method has turned out to be larger, as expected, for the real-world network, i.e., the electricity network of Great Britain, than for the synthetic small network: this is because the larger the number of survival signature entries, the more the reduction of the sample size (and computational cost) when entropy is used as driving criterion of the simulation.

References

- [1] N. L. Dehghani, S. Zamanian, and A. Shafieezadeh, “Adaptive network reliability analysis: Methodology and applications to power grid,” *Reliability Engineering and System Safety*, vol. 216, 2021, doi: 10.1016/j.ress.2021.107973.
- [2] Y. Kim and W.-H. Kang, “Network reliability analysis of complex systems using a non-simulation-based method,” *Reliability Engineering and System Safety*, vol. 110, pp. 80–88, 2013, doi: 10.1016/j.ress.2012.09.012.
- [3] E. Zio, *The Monte Carlo Simulation Method for System Reliability and Risk Analysis*. London: Springer London, 2013. doi: 10.1007/978-1-4471-4588-2.
- [4] E. Zio and N. Pedroni, *Reliability estimation by advanced monte carlo simulation, simulation methods for reliability and availability of complex systems*. London: Springer, pp. 3-39, 2010. doi: 10.1007/978-1-84882-213-9_1.
- [5] Samaniego FJ, *System Signatures and their Applications in Engineering Reliability (International Series in Operations Research & Management Science)*. New York: Springer, pp. 21-36, 2007. doi: <https://doi.org/10.1007/978-0-387-71797-5>.
- [6] F. P. A. Coolen and T. Coolen-Maturi, Generalizing the signature to systems with multiple types of components. *Complex systems and dependability. Advances in intelligent and soft computing*, vol. 170, Berlin, Heidelberg: Springer, pp. 115-30, 2013. doi: 10.1007/978-3-642-30662-4-8.

- [7] X. Huang and F. P. A. Coolen, "Reliability sensitivity analysis of coherent systems based on survival signature," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 232, no. 6, pp. 627–634, 2018, doi: 10.1177/1748006X18754974.
- [8] E. Patelli, G. Feng, F. P. A. Coolen, and T. Coolen-Maturi, "Simulation methods for system reliability using the survival signature," *Reliability Engineering and System Safety*, vol. 167, pp. 327–337, 2017, doi: 10.1016/j.ress.2017.06.018.
- [9] S. Reed, "An efficient algorithm for exact computation of system and survival signatures using binary decision diagrams," *Reliability Engineering and System Safety*, vol. 165, pp. 257–267, 2017, doi: 10.1016/j.ress.2017.03.036.
- [10] J. Behrendorf, T.-E. Regenhardt, M. Broggi, and M. Beer, "Numerically efficient computation of the survival signature for the reliability analysis of large networks," *Reliability Engineering and System Safety*, vol. 216, 2021, doi: 10.1016/j.ress.2021.107935.
- [11] D. Li, Q. Zhang, E. Zio, S. Havlin, and R. Kang, "Network reliability analysis based on percolation theory," *Reliability Engineering and System Safety*, vol. 142, pp. 556–562, 2015, doi: 10.1016/j.ress.2015.05.021.
- [12] P. Turati, N. Pedroni, and E. Zio, "An entropy-driven method for exploring extreme and unexpected accident scenarios in the risk assessment of dynamic engineered systems," in *Safety and Reliability of Complex Engineered Systems - Proceedings of the 25th European Safety and Reliability Conference, ESREL 2015*, 2015, pp. 761–769. doi: 10.1201/b19094-102.
- [13] T. J. , & C. D. F. Loredo, "Bayesian Adaptive Exploration," *Statistical Challenges in Astronomy*, pp. 57–70, 2003.
- [14] M. de Angelis, E. Patelli, and M. Beer, "Line sampling for assessing structural reliability with imprecise failure probabilities," in *Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management - Proceedings of the 2nd International Conference on Vulnerability and Risk Analysis and Management, ICVRAM 2014 and the 6th International Symposium on Uncertainty Modeling a*, 2014, pp. 915–924. doi: 10.1061/9780784413609.093.

- [15] L. Puppo, N. Pedroni, A. Bersano, F. di Maio, C. Bertani, and E. Zio, "Failure identification in a nuclear passive safety system by Monte Carlo simulation with adaptive Kriging," *Nuclear Engineering and Design*, vol. 380, 2021, doi: 10.1016/j.nucengdes.2021.111308.
- [16] E. Patelli and G. Feng, Efficient simulation approaches for reliability analysis of large systems. *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pp. 129-140, 2016. doi: 10.1007/978-3-319-40596-4_12.
- [17] E. Zio, *An introduction to the basics of reliability and risk analysis*. World Scientific Publishing, Singapore, pp. 128-130, 2007.
- [18] Y. Hu, "A guided simulation methodology for dynamic probabilistic risk assessment of complex systems," PhD Thesis, University of Maryland, pp. 78-89, 2005. uri: <http://hdl.handle.net/1903/2472>.
- [19] M. K. Bukhsh WA, "Network data of real transmission networks," 2013. Published online at: <https://www.maths.ed.ac.uk/optenergy/NetworkData/index.html>.
- [20] E. Zio, "From complexity science to reliability efficiency: A new way of looking at complex network systems and critical infrastructures," *International Journal of Critical Infrastructures*, vol. 3, no. 3-4, pp. 488-508, 2007, doi: 10.1504/IJCIS.2007.014122.
- [21] R. W. Floyd, "Algorithm 97: Shortest path," *Commun ACM*, vol. 5, no. 6, p. 345, 1962, doi: 10.1145/367766.368168.