Ensemble of Artificial Neural Networks for Approximating the Survival Signature of Critical Infrastructures

Francesco Di Maio[a], Chiara Pettorossi[a], Enrico Zio[a,b]

[a]Energy Department, Politecnico di Milano, Via La Masa 34, Milano, 20156, Italy

[b]MINES Paris-PSL, CRC, 06560 Sophia Antipolis, France

Abstract: Survival signature can be useful for the reliability assessment of critical infrastructures. However, analytical calculation and Monte Carlo Simulation (MCS) are not feasible for approximating the survival signature of large infrastructures, because of the complexity and computational demand due to the large number of components. In this case, efficient and accurate approximations are sought. In this paper we formulate the survival signature approximation problem as a missing data problem. An ensemble of Artificial Neural Networks (ANNs) are trained on a set of survival signatures obtained by MCS. The ensemble of trained ANNs is, then, used to retrieve the missing values of the survival signature. A numerical example is worked out and recommendations are given to design the ensemble of ANNs for large-scale, real-world infrastructures. The electricity grid of Great Britain, the New England power grid (IEEE 39-Bus Case), the reduced Berlin metro system and the approximated American Power System (IEEE 118-Bus Case) are, then, eventually, analyzed as particular case studies.

Keywords: Critical infrastructures, Reliability, Survival Signature, Monte Carlo simulation, Artificial Neural Networks, Ensemble.

---

1. INTRODUCTION

Water and gas distribution systems, energy transmission systems, transportation and telecommunication systems are critical infrastructures (CIs) [1], [2]. The reliability assessment of CIs is fundamental [3], [4], but, given their complexity [5], it cannot be carried out with classical methods.

Recently, the survival signature has been introduced in reliability assessment [6] for modelling dependencies [4], [7], [8], handling imprecise probabilities in a Bayesian scheme [9]–[11] and dealing with the complexity of the CIs [12]–[15]. The survival signature allows decoupling the structural information of the system from the probabilistic one [6], [16], which can be a major advantage for the reliability assessment of large systems. On the other hand, its analytical definition for systems of realistic size is difficult, due to the combinatorial nature of the problem [17]. Monte Carlo Simulation (MCS) combined with percolation theory [18] has been used to obtain an approximation of the survival signature for large systems. Entropy has been used to guide MCS in a way to efficiently allocate the simulation efforts over the whole survival signature [12]. Nonetheless, the search for every combination of functioning components of a large system makes the approximation of the survival signature computationally prohibitive.

In this paper, we formulate the survival signature approximation problem as a missing data problem and present a novel method that does not require the estimation of every survival signature entry (i.e., every path set of functioning components that render the system functioning). Once percolation [19] is used to determine the trivial entries of the survival signature, i.e., those corresponding to non-functioning network system configurations, MCS [20] is used to estimate only a limited number of survival signature entries, thus obtaining sparse information on the survival signature. The survival signature entries that have not been estimated by MCS, are the missing data to be retrieved. For this, an ensemble of Artificial Neural Networks (ANNs) [21] are trained on the available, incomplete data of the sparse survival signature [22], which forms the training dataset. The ANN has been selected to approximate the survival signature because it is a well-known, widely spread tool in ML method. It is here used to demonstrate the validity of the methodology proposed with no claim of being the best ML method. The survival signature can, then, be completed by the estimates provided by the trained ensemble, with a reasonable computational effort. To quantify the accuracy of the ANN-based approximation, the confidence intervals of the estimates, the average confidence interval width and the coverage value are calculated as synthetic metrics [23], [24]. This allows quantifying the uncertainty on the survival signature estimate.

The proposed ANN-based method is developed with reference to a numerical example, which allows giving some guiding recommendations for the definition of the ensemble of ANNs. A set of realistic large scale CIs, i.e., the electricity grid of Great Britain [25], the New England power grid (IEEE 39-Bus Case) [26], the Berlin metro system [20] and the American Electric Power System (IEEE 118-Bus Case) [27] are, then, processed with the proposed method. Indeed, the MCS approximation of the survival signature has proven quite time consuming [12], [20] and the results obtained with the proposed method are benchmarked with those of a MCS method available in literature [18], confirming the superiority in reducing the computational time, while providing good coverage and accurate reliability assessment.

The remainder of the paper is organized as follows. In Section 2, a brief recall of the concept of the survival signature is provided for completeness. In Section 3, the novel method combining percolation, MCS and the ensemble of ANNs for its approximation is presented. The feasibility of the method is proven in Section 4 by means of a numerical example followed by the application to complex, real world infrastructures in Section 5, where a comparative study with MCS is also conducted. Finally, in Section 6 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, so that the generic element $x_i = 1$ if the *i-th* component functions or $x_i = 0$ if not. When 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 revised as $\underline{x} = (\underline{x}^1, \underline{x}^2, \dots, \underline{x}^K)$, with $\underline{x}^k = \left(x_1^k, x_2^k, \dots, x_{m_k}^k\right)$ representing the states of the components of type *k*. Let $\varphi : \{0,1\}^m \rightarrow \{0,1\}$ be the system reliability structure function, 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. We assume that the system is coherent, i.e., the structure function is not decreasing if the number of functioning components increases.

Then, the survival signature for the system $\Phi(l_1, l_2, \dots, l_K)$ is given in Eq. (1), 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]:

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

This results from the fact that the state vectors $\underline{x}^k$ with exactly l_k components functioning are $\binom{m_k}{l_k}$ and that, consequently, the set of all the allowed combinations of all components, denoted as $S_{l_1, l_2, \dots, l_K}$, has magnitude $\prod_{k=1}^{K} \binom{m_k}{l_k}$. Also, all the state vectors $\underline{x}^k \in S_{l_k}$ are equally likely to occur, under the assumption that the failure times of the m_k components of type k are *iid*. Eq. (1) implies that the analytical calculation of the survival signature for the entry $(l_1, l_2, \dots, l_K)$ requires evaluating $\binom{m_1}{l_1} \cdot \binom{m_2}{l_2} \cdot \dots \cdot \binom{m_K}{l_K}$ times the system structure function $\varphi(\underline{x})$, which results in a combinatorial explosion when large systems are considered. Eventually, the generalized survival signature is a multidimensional array with dimension $(m_1 + 1) \times (m_2 + 1) \times \dots \times (m_K + 1)$ (including the case $l_k = 0$ in which none of the components of type k are working), so that the total number of entries $\underline{l} = (l_1, l_2, \dots, l_K)$, with $l_k = 0, 1, \dots, m_k$ for each $k = 1, 2, \dots, K$, to be considered is $M = (m_1 + 1) \cdot (m_2 + 1) \cdot \dots \cdot (m_K + 1)$.

The survival signature of Eq. (1) allows to compute the survival function, that gives the reliability of the system. This can be calculated as in Eq. (2) [6]:

$$P(T_s > t) = \sum_{l_1=0}^{m_1} \sum_{l_2=0}^{m_2} \dots \sum_{l_K=0}^{m_K} \Phi(l_1, l_2, \dots, l_K)\, P\left(\bigcap_{k=1}^{K} \{C_t^k = l_k\}\right) \qquad (2)$$

where $P(C_t^k)$ describes the probability that the number of components of type k that function at time $t > 0$ is $C_t^k \in \{0, 1, \dots, m_k\}$. Eq. (2) shows that in the reliability calculation the system structure can be analyzed separately from the probabilistic information (e.g., the failure time distributions of the components), which is the main advantage of using the survival signature. Furthermore, when the cumulative distribution function $F_k(t)$ for components of type *k* is known, and the failure times of the different component types are independent, the evaluation of the probabilistic part is straightforward, Eq. (3):

$$P\left(\bigcap_{k=1}^{K}\{C_t^k = l_k\}\right) = \prod_{k=1}^{K} P(C_t^k = l_k) = \cdots = \prod_{k=1}^{K}\left(\binom{m_k}{l_k} F_k(t)^{\,m_k - l_k}[1 - F_k(t)\,]^{l_k}\right) \tag{3}$$

3. THE ANN-BASED METHOD

Fig. 1 shows an overview of the proposed method, which comprises of the sequential application of: *i)* percolation theory, *ii)* MCS and *iii)* an ensemble of ANNs.

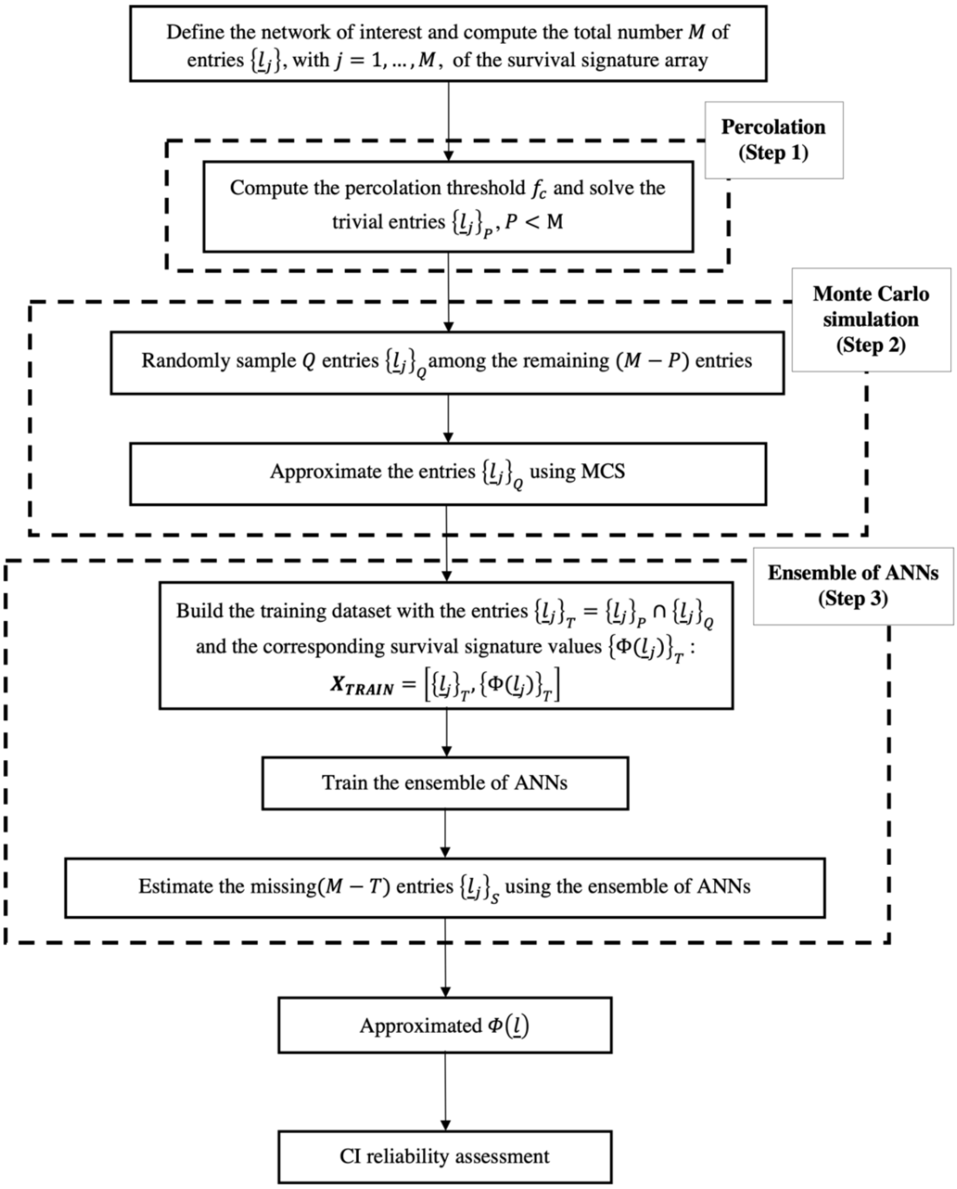

Figure 1 - Flowchart of the novel ANN-based method for the survival signature approximation.

Let us consider a critical infrastructure, i.e., a network system with K component types, m_k components of the type k, $k = 1, \dots, K$: its survival signature is a K dimensional array with M entries in total, $M = (m_1 + 1) \cdot (m_2 + 1) \cdot \dots \cdot (m_K + 1)$. All such entries $\underline{l} = (l_1, l_2, \dots, l_K)$ are grouped in the set $\{\underline{l}_j\}$ with $j = 1, \dots, M$. Assume that among this set of entries, a subset $\{\underline{l}_j\}_P$ with $P < M$ is solved by percolation, i.e., $\{\underline{l}_j\}_P$ is the subset of trivial entries (see Section 3.1 for more details) whose dimension is P. Then, randomly select a number Q among the remaining entries of the survival signature (as low as possible, for reasons that will become clear in what follows) to be approximated by MCS (see Section 3.2 for more details). Then, Q is the dimension of the subset comprising the entries $\{\underline{l}_j\}_Q$, and $\{\underline{l}_j\}_P \cap \{\underline{l}_j\}_Q = 0$. The entries $\{\underline{l}_j\}_T = \{\underline{l}_j\}_P \cup \{\underline{l}_j\}_Q$ sparsely populate the survival signature, and, together with the survival signature values $\{\Phi(\underline{l}_j)\}_T$, form the training dataset of dimension T. Fig. 2 shows a sparse and incomplete survival signature array for a generic system with $K = 2$ components: in black, the T entries making up the training dataset and in white, the missing entries to be approximated by the ensemble of ANNs.

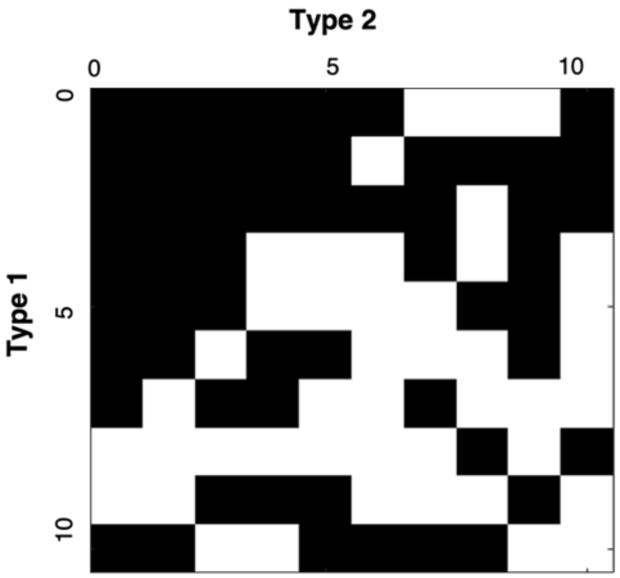

Figure 2 – Sparse survival signature for a network system with K=2 component types: 10 components of type 1 and 10 components of type 2. In black the T available entries, and in white the S missing entries.

After training (see Section 3.3 for more details), the ensemble of ANNs is used to estimate the missing data, i.e., the subset of entries $\{\underline{l}_j\}_S$ with dimension $S = M - T$, so that $\{\underline{l}_j\}_T \cap \{\underline{l}_j\}_S = 0$ and $\{\underline{l}_j\}_T \cup \{\underline{l}_j\}_S = \{\underline{l}_j\}_M$.

3.1. Step 1: Percolation theory

Percolation theory is used to calculate the trivial solutions of Eq. (1) with respect to the failure of the network system [19]. The trivial entries of the survival signature, forming the subset $\{\underline{l}_j\}_P$ of dimension P, are those configurations containing a set of non-functioning components sufficient to render non-functioning the network system [28]. The percolation process consists in:

- computing the critical fraction of non-functioning components f_c, which guarantees (conservatively) that the non-functioning probability of the network is negligible, by means of the Molloy-Reed Criterion [29]:

$$f_c = 1 - \frac{1}{\langle d^2 \rangle / \langle d \rangle - 1} \tag{4}$$

where 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 system, respectively, which are obtained from the network adjacency matrix (see Eq. 6 below),

- finding the trivial entries of the survival signature, that are those for which there is no network system connectivity and whose survival signature can be set to zero, i.e., those for which the fraction of functioning components is lower than $1 - f_c$, as shown in Eq. (5) [18]:

$$\sum_{k=1}^{K} l_k < (1 - f_c) \cdot \sum_{k=1}^{K} m_k \quad \Rightarrow \quad \Phi(\underline{l}) \approx 0 \tag{5}$$

For a network system with N nodes, the adjacency matrix is a $N \times N$ matrix $\{a_{ij}\}$ whose entries are:

$$\begin{aligned} a_{ij} &= 1 \text{ if there is a connection between node } i \text{ and } j \\ &= 0 \text{ otherwise} \end{aligned} \tag{6}$$

In Fig. 3, the survival signature array for a generic network system with $K = 2$ components is shown after percolation is performed: in black, the P trivial entries found by percolation, and in white, the remaining entries to be approximated.

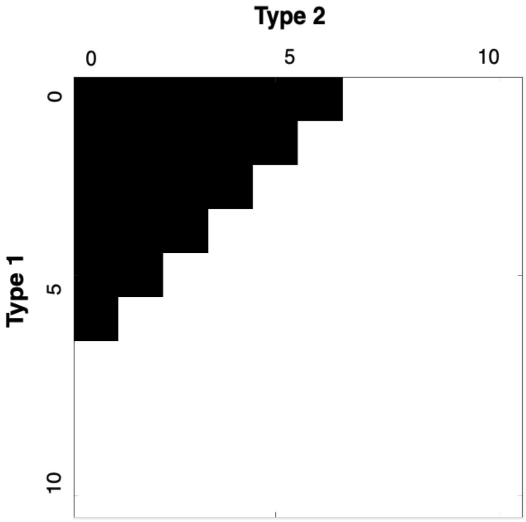

Figure 3 - Survival signature for a generic network system with K=2 component types, 10 components of type 1 and 10 components of type 2. In black the P trivial entries solved by percolation and in white the remaining entries.

3.2. Step 2: Monte Carlo simulation (MCS)

Let us consider the entry $\underline{l} = (l_1, l_2, \dots, l_K) \in \{\underline{l}_j\}_Q$, that is the subset of dimension Q containing the entries to be approximated by MCS. For the MCS approximation of the entry $\underline{l} = (l_1, l_2, \dots, l_K)$, random network system configurations $\underline{x_l} \in S_{l_1, l_2, \dots, l_K}$ are generated and evaluated by the structure function, $\varphi(\underline{x_l})$, with respect to the fact that the system is functioning or non-functioning in such configurations. For this, the efficiency metric $E(G)$ of network system G with N nodes is used [30]:

$$E(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in G} \frac{1}{r_{ij}} \tag{8}$$

where $\{r_{ij}\}$ is the matrix of the shortest path lengths, whose entry r_{ij} gives the length of the shortest path between nodes i and j, calculated using the Floyd-Warshall algorithm [31].

Indeed, for large network systems the structure function is usually not available and so it is assumed that, if the loss of efficiency in a given configuration $\underline{x_l}$ exceeds 50%, the network system is non-functioning and the structure function $\varphi(\underline{x_l})$ is set to 0:

$$if \ \frac{\mathrm{E}\left(G(\underline{x_l})\right)}{\mathrm{E}(G)} < 0.5 \Rightarrow \varphi(\underline{x_l}) = 0 \tag{7}$$

whereas the network system is functioning, i.e., $\varphi(\underline{x_l}) = 1$, otherwise. This threshold value has been adopted from [18], where it has proven to provide good results.

The algorithm for the MCS approximation of a survival signature entry is given in Fig. 4.

Require $\underline{l} = (l_1, l_2, \dots, l_K)$ survival signature entry; n: number of samples; E: efficiency function; G: network system.

w, counter, $\Phi \leftarrow 0$ %Initialize variables

while counter $\leq$ n %Loop over number of samples to collect

counter $\leftarrow$ counter + 1 %Update number of samples

$\underline{x_l}$ %Random state vector for $\underline{l}$

if $\frac{E(G(\underline{x_l}))}{E(G)} \geq 0.5$

w $\leftarrow$ w + 1 %Update working configurations counter

end

end

Φ = w/counter %Compute survival signature

Figure 4 – Algorithm for the approximation of a survival signature entry, $\Phi(l_1, l_2, \dots, l_K)$.

The procedure for generating a random network system configuration $\underline{x_l}$ for the entry $\underline{l} = (l_1, l_2, \dots, l_K)$ consists in randomly permuting the components order in each state vector $\underline{x}^k$ of the components of type k and set as functioning ($x_i^k = 1$) the first l_k elements of the vector, for each component type $k \in \{1,2, \dots, K\}$ [18]. For the sake of clarity, consider a network system made up of $m_1 = 6$ components of type $k = 1$, so that the state vector for component type $k = 1$ is $\underline{x}^1 = (x_1^1, x_2^1, x_3^1, x_4^1, x_5^1, x_6^1)$, and that a random system configuration with $l_1 = 4$ components functioning has to be sampled. Then, given a random permutation for $\underline{x}^1$, e.g., $\underline{x}^1{}_{perm} = (x_1^1, x_5^1, x_6^1, x_3^1, x_2^1, x_4^1)$, set the first $l_1 = 4$ elements of the permuted vector as functioning, i.e., $\{x_1^1, x_5^1, x_6^1, x_3^1\} = 1$, and the remaining as non-functioning, i.e., $\{x_2^1, x_4^1\} = 0$. Eventually, the sampled state vector is $\underline{x}^1{}_{sample} = (1, 0, 1, 0, 1, 1)$. The same procedure is repeated for the other component types. The number of random samples n, defined in the algorithm of Fig. 4, is selected by the analyst, and it is related to the magnitude of the set $S_{l_1,l_2,\dots,l_K}$, that is $\binom{m_1}{l_1} \cdot \binom{m_2}{l_2} \cdot \dots \cdot \binom{m_K}{l_K}$. In fact, this influences the accuracy of the survival signature approximation, since the larger the sampling space the larger the sample size n required to obtain an accurate approximation.

3.3. Step 3: Ensemble of Artificial Neural Networks

In this sub-Section, the ensemble of ANNs approach for estimating the missing entries of the survival signature of a network system is described. A typical ensemble for the survival signature approximation is given in Fig. 5.

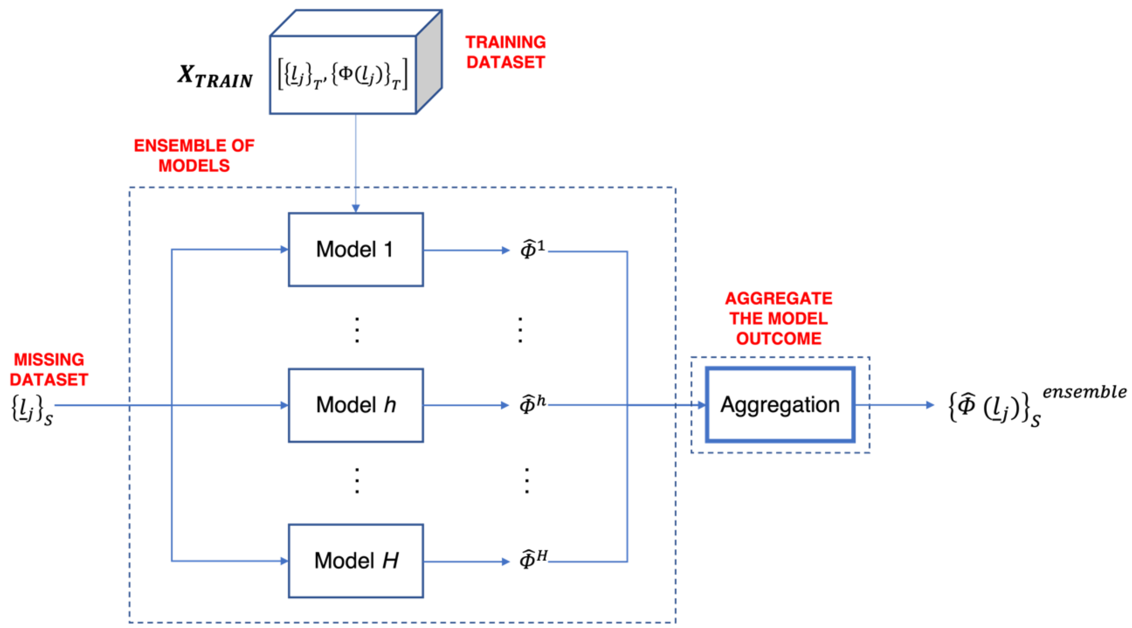

Figure 5 – Scheme of the ensemble of ANNs models used for the approximation of the survival signature of network systems.

The ensemble comprises of H alternative ANN models. In particular, Feedforward Neural Networks (FNNs) are considered with different architectures, i.e., numbers of neurons, n_{neuron}, and/or numbers of layers, n_{layer}. A training dataset $\boldsymbol{X_{TRAIN}} = \left[\{\underline{l}_j\}_T, \{\Phi(\underline{l}_j)\}_T\right]$, where the survival signature entries in the subset $\{\underline{l}_j\}_T$ are the inputs to be mapped into the output $\{\Phi(\underline{l}_j)\}_T$, with $j = 1, \dots, M$, is used to train the ensemble. In our application this is performed, without loss of generality, using the Levenberg-Marquardt algorithm available in the Deep Learning Toolbox in Matlab.

Once the ANNs in the ensemble are trained, they are used to provide the output estimate $\{\widehat{\Phi}(\underline{l}_j)\}_S$ of the missing entries in the subset $\{\underline{l}_j\}_S$. For each entry $\underline{l} = (l_1, l_2, \dots, l_K)$, the aggregated $\widehat{\Phi}(\underline{l})$ is computed as the average value of $\widehat{\Phi}(\underline{l})^h$:

$$\widehat{\Phi}(\underline{l}) = \frac{1}{H}\sum_{h=1}^{H} \widehat{\Phi}(\underline{l})^h \tag{9}$$

where $\widehat{\Phi}(\underline{l})^h$, with $h = 1, \dots, H$, is the output of the single, generic h-th model, i.e., the survival signature estimate of such single model. Each estimate $\widehat{\Phi}(\underline{l})$ comes with an associated error $\sigma_{\underline{l}}^2$, which is the ensemble model error due to the random initialization of the ANNs parameters and to the uncertainty related to the limited training dataset (i.e., the uncertainty of the MCS approximation) [21]. Hence, for an entry $\underline{l}$, a confidence interval, defined as the interval $[\widehat{\Phi}(\underline{l})^{low}, \widehat{\Phi}(\underline{l})^{up}]$ such that the probability is equal to α that the true survival signature $\Phi(\underline{l})$ falls within the interval, can be computed as follows [24]:

$$[\widehat{\Phi}(\underline{l})^{low}, \widehat{\Phi}(\underline{l})^{up}] = \widehat{\Phi}(\underline{l}) \pm C_{dof}^{\alpha}\sqrt{\sigma_{\underline{l}}^2} \tag{10}$$

where C_{dof}^{α} is the $(1-\alpha)/2$ quantile of the Student *t*-distribution with number of degrees of freedom equal to the number of the ensemble models, i.e., H.

In what follows, for evaluating the accuracy of the estimated survival signature, two metrics are considered [23]:

- the average confidence interval width, and
- the coverage, i.e., the fraction of true vs estimated survival signature entries which fall within the constructed confidence intervals.

4. NUMERICAL EXAMPLE

In Fig. 6, we plot a 21×21 matrix representing the survival signature of a generic network system with $K = 2$ component types, with $m = 40$ components in total, $m_1 = 20$ of type 1 and $m_2 = 20$ of type 2, respectively, which has to be approximated by the ANN-based method presented in the previous Section. To proceed with the analysis, one needs to set the dimensions P, Q and S of the subsets $\{\underline{l}_j\}_P$, $\{\underline{l}_j\}_Q$ and $\{\underline{l}_j\}_S$, with $j = 1, \ldots, M$, the type of ANNs composing the ensemble and their architecture, specifically the number of layers, n_{layer}, and number of neurons per layer, n_{neuron}.

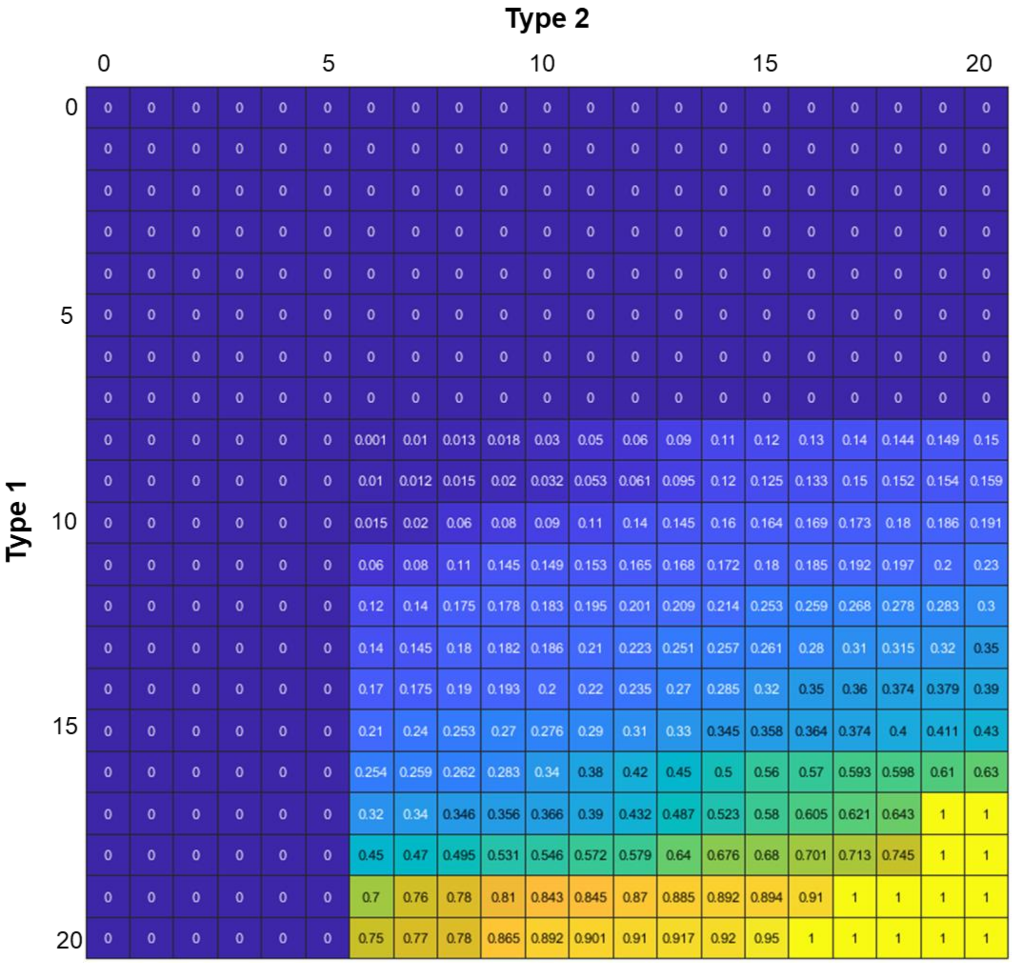

Figure 6 – Survival signature of a generic network system with 20 components of type 1 and 20 components of type 2.

The survival signature in Fig. 6 has $M = 441$ entries in total. P is the dimension of the subset of survival signature entries estimated by percolation, $\{\underline{l}_j\}_P = \{(l_1, l_2)_j\}_P$, which is fixed once the percolation threshold f_c for the network system is set. The subset $\{\underline{l}_j\}_Q = \{(l_1, l_2)_j\}_Q$ comprises the survival signature entries to be approximated by MCS with dimension Q. $\{(l_1, l_2)_j\}_T = \{(l_1, l_2)_j\}_P \cup \{(l_1, l_2)_j\}_Q$ is the subset of survival signature entries which, together with the corresponding survival signature values $\{\Phi(l_1,l_2)_j\}_T$, has been used as training dataset for the ANNs in the ensemble: $\boldsymbol{X_{TRAIN}} = \left[\{(l_1,l_2)_j\}_T, \{\Phi(l_1,l_2)_j\}_T\right]$, where T is the size of the dataset.

As explained earlier, the ANNs used in the ensemble are FNNs, which differ in the number of neurons, n_{neuron}, and/or layers, n_{layer}, and have been trained with the Levenberg-Marquardt algorithm available in the Deep Learning Toolbox in Matlab. Once the ensemble has been trained, given the missing entries in input, i.e., those in the subset $\{\underline{l}_j\}_S = \{(l_1,l_2)_j\}_S$, it provides in output the estimates $\{\widehat{\Phi}(l_1,l_2)_j\}_S$ of the survival signature $\{\Phi(l_1,l_2)_j\}_S$, and the corresponding confidence intervals (computed with a confidence level $\alpha = 90\%$) and coverage.

Since examining different values of Q is equivalent to examining different values of T, being P fixed and $T = P + Q$, the optimal dimension of the subset $\{\underline{l}_j\}_Q$ is searched by training the FNNs of the ensemble for different T. The results are given in Table 1, where the MSE, the average confidence interval width and the coverage are listed.

Table 1 – MSE, average confidence interval width and coverage of the estimated survival signature for different dataset dimensions T (Ensemble of ANNs: 1 layer FNNs, 15 to 25 neurons).

T	MSE	Average confidence interval width	Coverage [%]
221	8.35E-03	1.47E-01	86.4
265	3.15E-03	1.23E-01	88.7
309	4.35E-03	1.17E-01	90.2
353	2.85E-03	1.05E-01	85.8
397	4.28E-03	1.07E-01	79.6

The MSE is small for all scenarios considered, which indicates that the method is capable of estimating the survival signature, and the small confidence interval widths show that the ensemble of ANNs provides accurate results [23]. However, the ensemble trained with the dataset corresponding to $T = 309$ is the only one for which the coverage is larger than the confidence level: this means that $T = 309$ is the optimal size for the training dataset (i.e., $\frac{T}{M} \cdot 100 \cong 70\%$ of the survival signature entries). Then, as said, Q is determined as $Q = T - P$ and $S = M - T = 132$ (i.e., $S/M \cong 30\%$). These dimensions of the subsets of entries will be adopted also in Section 5 for the approximation of the survival signature of real-world CIs.

To optimize the architecture of the FNNs in the ensemble, a grid optimization approach is adopted: single layers FNNs, with n_{neuron} ranging from 1 to 40, are trained. The performance of each FNN is evaluated by means of the average mean squared error (MSE), as shown in Fig. 7.

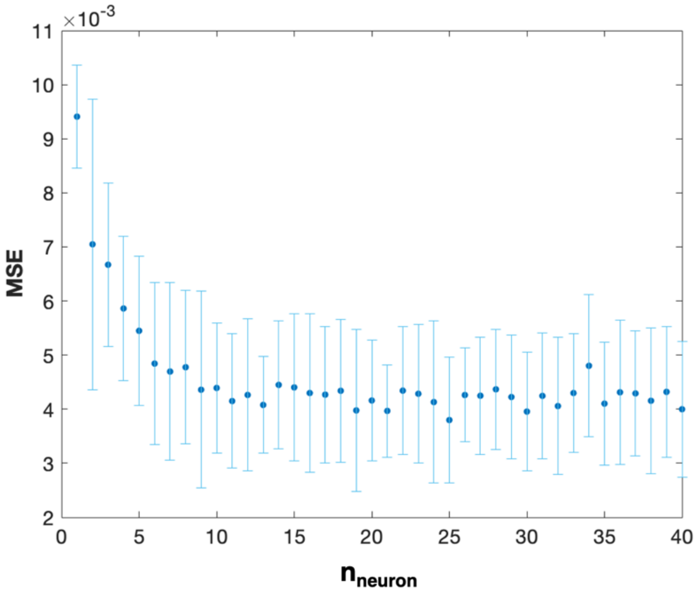

Figure 7 – MSE vs n_{neuron} for 1 layer FNNs, for the approximation of the survival signature in Fig. 6.

It can be seen that single layer FNNs with $n_{neuron} \geq 10$ provide survival signature estimates with MSE $\cong 3.5 \times 10^{-3} \div 4.5 \times 10^{-3}$. Nonetheless, an ensemble of ANNs has been selected here to capture the uncertainty in the estimates. As a general constraint, to limit the number of hyperparameters to be estimated (and the corresponding computational time) when training the FNNs, the number of hidden neurons should be kept small [32]. Then, the following ensembles of FNNs are selected, trained and their performances compared:

a) 11 FNNs: 1 layer FNNs, 15 to 25 neurons ($n_{layer} = 1$, $n_{neuron} = [15, 25]$);

b) 36 FNNs: 2 layers FNNs, 15 to 20 neurons ($n_{layer} = 2$, $n_{neuron} = [15, 20]$);

c) 47 FNNs: 1 layer FNNs, 15 to 25 neurons ($n_{layer} = 1$, $n_{neuron} = [15, 25]$); 2 layers FNNs, 15 to 20 neurons ($n_{layer} = 2$, $n_{neuron} = [15, 20]$);

d) 132 FNNs: 1 layer FNNs, 15 to 25 neurons ($n_{layer} = 1$, $n_{neuron} = [15\ ,25]$); 2 layers FNNs, 15 to 25 neurons ($n_{layer} = 2$, $n_{neuron} = [15\ ,25]$);

In Table 2, the different ensembles are compared with respect to the following metrics: *i)* MSE, *ii)* average confidence interval width, *iii)* coverage probability and *iv)* ensemble training time.

Table 2 – MSE, average confidence interval width, coverage and training time for the ensembles a), b), c), d).

Ensemble	MSE	Average confidence interval width	Coverage [%]	Training time [s]
a)	2.71E-03	1.06E-01	90.2	3.5
b)	2.91E-03	1.17E-01	90.9	15.2
c)	1.98E-03	1.07E-01	93.4	18.1
d)	2.11E-03	1.21E-01	94.1	33.3

The MSE is in the order of 10^{-3} for every ensemble, which is small enough for our purposes, and lower than the MSE of the best single FNN, provided in Fig.7. The average confidence interval width confirms that all the four ensembles provide estimates with high, but similar, accuracy, which is also verified by the coverage value, that is larger than the fixed confidence level ($\alpha = 90\%$). Concerning the computational time, the ensemble of ANNs with the smallest training time, i.e., ensemble a), made up of $H = 11$ FNNs with $n_{layer} = 1$ and $n_{neuron} = [15\ ,25]$, is selected as the optimal architecture for a network system with $m = 40$ components.

The survival signature used in this numerical example is a toy example representative of the survival signature of a generic system with $m = 40$ components and $k = 2$ component types, as described at the beginning. With similar reasoning, starting from generic survival signatures of network systems, we can, then, draft general recommendations for defining the optimal FNNs architecture in the ensemble. As we shall see in the next Section, these will be confirmed by other network systems considered as case studies. From a trial-and-error procedure on survival signatures of network systems of different topology, size and complexity, the following recommendations for a network system with m components were deducted:

- Number of layers $n_{layer} = k$, with $10^k \leq m < 10^{k+1}$ (for the network system analyzed: $m = 40$ components, so that $k = 1$ and $n_{layer} = 1$ is used). For example, if $10 \leq m < 100$ set $n_{layer} = 1$, if $100 \leq m < 1000$ set $n_{layer} = 2$, and so on;
- Number of neurons per layer $n_{neuron} = \left[\frac{m}{2 \cdot {n_{layer}}^2} - 5, \frac{m}{2 \cdot {n_{layer}}^2} + 5\right]$ (for the network system analyzed: $\frac{m}{2 \cdot {n_{layer}}^2} = 20$ and $n_{neuron} = [20 - 5, 20 + 5] = [15, 25]$ is used).

For practical purposes, the complete workflow for estimating the survival signature of Fig. 6 using the ANN-based method is summarized:

Step 1. Compute the total number of survival signature entries, $M = 441$. Set the percolation threshold, $f_c = 0.6$, which leads to $P = 136$ survival signature entries solved by percolation;

Step 2. Compute the dimension T of the training dataset for optimal training: $T = M \cdot 70\% \cong 309$. Randomly select $Q = T - P = 173$ entries to be approximated by means of MCS ($N = 1000$ samples are used);

Step 3. Build the training dataset with the sparse survival signature available and train the ensemble of ANNs. The ensemble comprises FNNs with $n_{layer} = 1$ $(10 \leq m < 100)$ and $n_{neuron} = [15, 25]$ $(\frac{m}{2 \cdot {n_{layer}}^2} = 20)$. The missing entries, $S = M - T = 132$, are estimated and the corresponding confidence intervals are computed.

The resulting survival signature is presented in Fig. 8, with the corresponding confidence intervals.

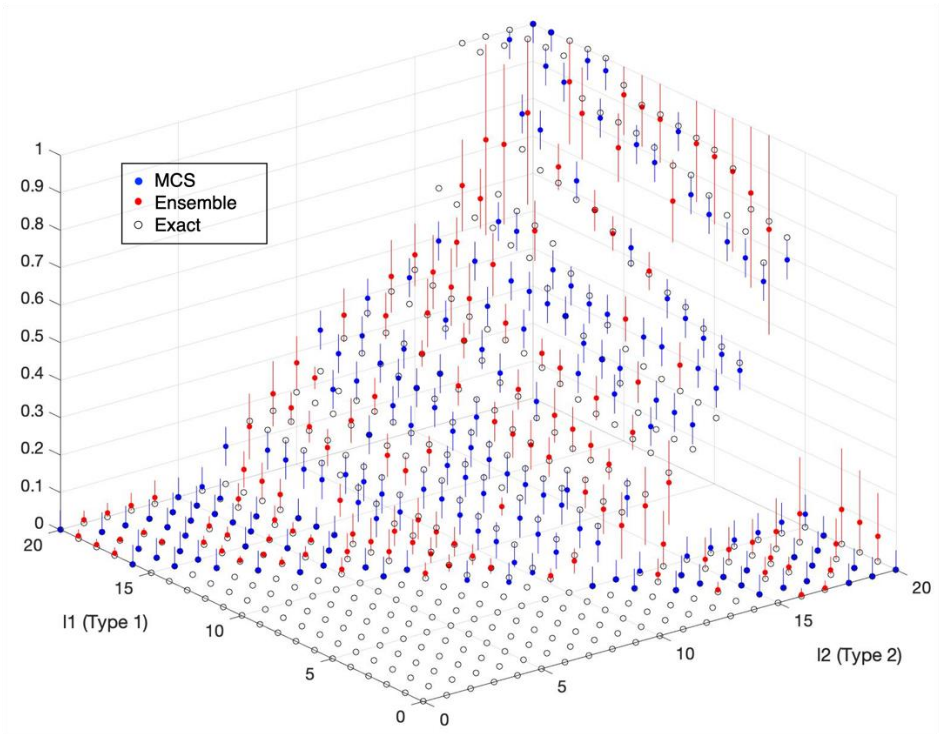

Figure 8 – 3D plot of the survival signature of Fig. 6 approximated with the ANN-based method: in blue the Q survival signature entries estimated using MCS and in red the S survival signature entries estimated with the ensemble of FNNs, both with the corresponding confidence intervals (the entries solved by percolation are neglected). The exact survival signature is given in black.

It is worth noting that the confidence intervals of the estimates provided by the ensemble (red dots), which have been computed by means of Eq. (10), are not constant in width: these are small for low values of l_1 or l_2, and increase for larger values of the parameters, due to the large variability (i.e., steepness) of the survival signature that might be difficult to be interpreted. On the other hand, the width of the confidence intervals for the MCS estimates (blue dots) is constant (except for the estimates close to 0 and 1), since related only to the number of samples used, and it is smaller than the one for the ensemble estimates: the average width for the former results is $8.40E-02$, whereas for the latter it is $1.20E-01$, and, on average, the coverage is 90.2%. Overall, these values are small and confirm that the method is suitable and provides accurate approximation of the survival signature.

5. APPLICATIONS

In what follows, the ANN-based method is used for approximating the survival signature of real-world, large-scale CIs: the electricity grid of Great Britain [25], the New England power grid (IEEE 39-Bus Case) [26], the Berlin metro system [20] and the American Electric Power System (IEEE 118-Bus Case) [27]. Eventually, the results are benchmarked with those provided by the MCS method [18].

5.1. The electricity transmission network of Great Britain

We consider the electricity transmission network of Great Britain [25], whose topological model is shown in Fig. 9. The network system comprises of $m = 29$ components which have been separated in $K = 2$ types, according to their bus type: load-controlled buses are grouped in component type 1, whereas voltage-controlled buses are grouped in component type 2, resulting in $m_1 = 5$ 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 $b = 2$.

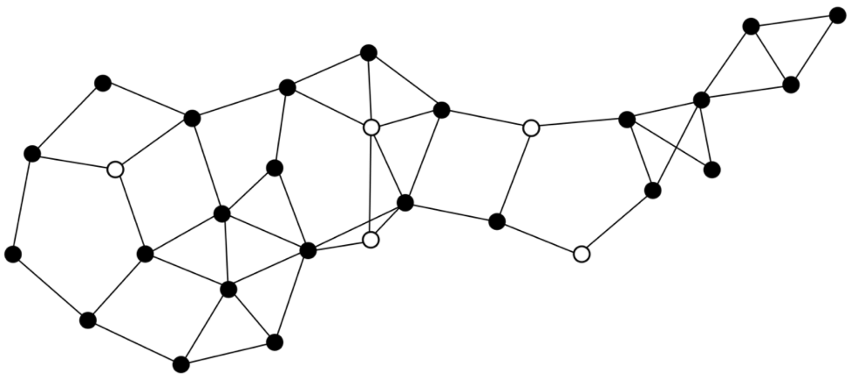

Figure 9 – Topology of the electricity transmission network of Great Britain. Components of type 1 (load-controlled buses) are shown in white, components of type 2 (generator-controlled buses) are shown in black (adapted from [18]).

The network system survival signature is an array with dimension 6×25. The ANN-based method is used to obtain an approximation of the survival signature. First, the percolation threshold is computed to be $f_c = 0.64$; this results in $P = 51$ of the $M = 150$ entries of the survival signature being solved. Subsequently, $Q = 54$ of the remaining entries are randomly selected and approximated by means of MCS, so that a sparse survival signature array comprising 70% of the entire survival signature is available. The ensemble of ANNs is defined using the setting in Section 4: it consists of FNNs with

$n_{layer} = 1$ and $n_{neuron} = [10 , 20]$, $H = 11$ FNNs in total. The ensemble is trained with the available data and used to estimate the missing $S = 45$ entries of the survival signature. The confidence intervals for the survival signature estimates are built setting a confidence level $\alpha = 90\%$: the average confidence interval width is 2.31E-01, the coverage is 96.6%.

This result is compared with that of the crude MCS method. In particular, as an indication of the computational cost, the time needed to approximate the survival signature using the MCS method is compared to the time needed with the ANN-based method: the former results in 22.6 s, the latter in 12.4 s, with 45.2% time saving. The training time for the ensemble of ANNs is 3.48 s.

The complete survival signature, estimated with the ANN-based method, is used to compute the reliability of the network system (see Eq. (2)) for a mission time $T_m = 1$, in arbitrary units of time. In Fig. 10, this is benchmarked with the reliability obtained using the survival signature approximated with MCS [18]. For illustrative purposes, the plot of the error bounds for the two alternative methods is also given, computed using the 5^{th} and the 95^{th} percentiles of the survival signature estimates.

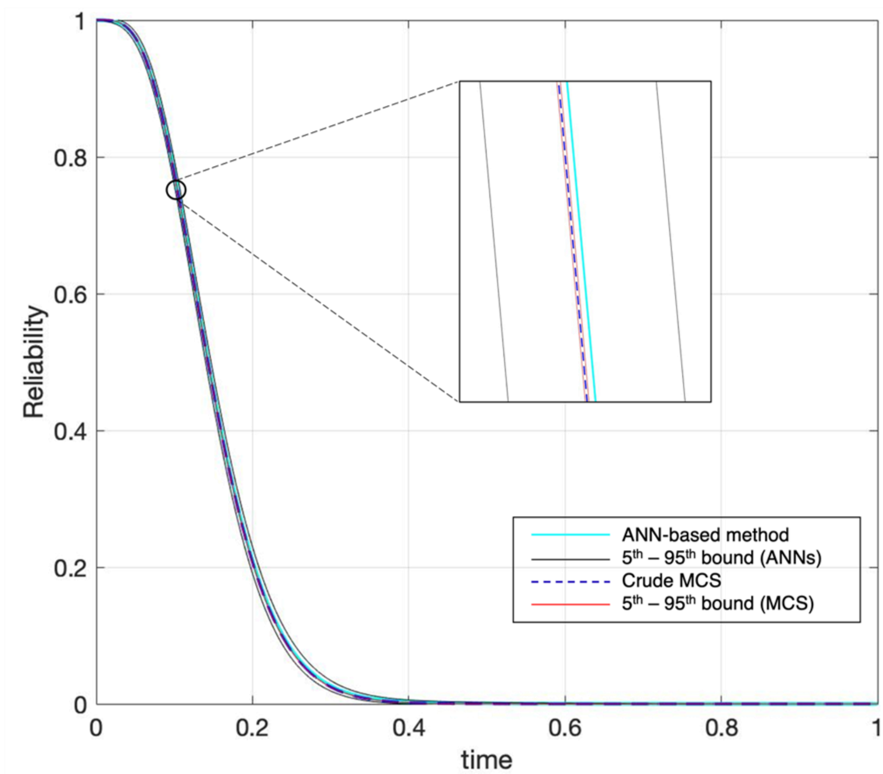

Figure 10 – Survival function of the system in Fig. 9 computed with the survival signature approximated by the crude MCS method (blue dashed line), with corresponding error bounds in red, and approximated by the ANN-based method (light blue), with corresponding error bounds in black.

The two reliability curves have very similar behavior. For the crude MCS the error bounds are smaller than the ones provided by the ANN-based method, as expected. Overall, on an absolute scale, the result provided by the novel method is accurate, which is also confirmed by the Kullback-Leibler divergence D_{KL} between the two reliability functions: $D_{KL}(rel(t)_{MCS} \parallel rel(t)_{ANN}) = 2.51E-02$. D_{KL} gives a measure of the error between the network system reliability estimated with the ANN-based method, $rel(t)_{ANN}$ and that estimated by MCS [18], $rel(t)_{MCS}$.

5.2. IEEE 39-Bus System: New England power system

The IEEE 39-bus power system test case (Fig. 11) represents a simplified model of the New England power grid [26]. The network system has $m = 39$ nodes in total, which can be separated in $K = 2$ types, specifically load buses, assigned to component type 1, and generator buses, assigned to component type 2. This results in $m_1 = 29$ nodes of type 1, whose failure times follow an exponential distribution with $\lambda = 1$, in arbitrary units of inverse time, and $m_2 = 10$ nodes of type 2, whose failure times follow a Weibull distribution with parameters $a = 1$ and $b = 2$.

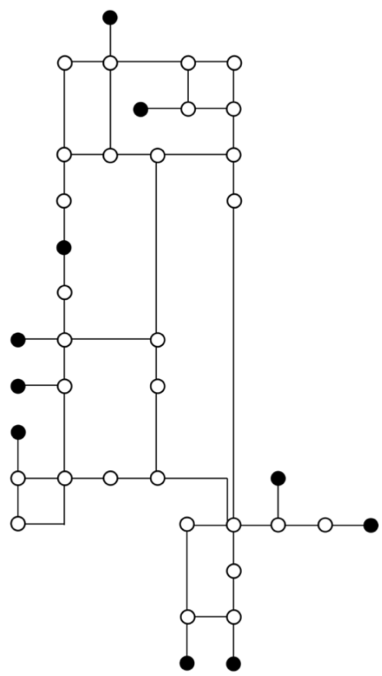

Figure 11 – Topology of the IEEE 39-Bus power system. Components of type 1 (load buses) are shown in white, components of type 2 (generator buses) are shown in black (adapted from [25]*).*

The network system survival signature is an array of dimension 30×11 and the ANN-based method is used for its approximation. The percolation threshold for this network system is $f_c = 0.4461$, which allows solving $P = 187$ of the $M = 330$ entries of the survival signature. To obtain a sparse survival signature matrix describing 70% of the entire survival signature, $Q = 44$ entries among the remaining entries are randomly selected and approximated by means of MCS. The ensemble of ANNs used to estimate the missing $S = 99$ entries consists of FNNs with $n_{layer} = 1$ and $n_{neuron} = [15, 25]$, on the basis of the analysis in Section 4. The confidence intervals for the survival signature estimates are built with a level $\alpha = 90\%$: the average confidence interval width results in 6.38E-02 and the coverage value is 94.5%.

A comparative study with the crude MCS method [18] is presented. The time needed to approximate the survival signature using MCS is 65.6 s, against 26.8 s for the ANN-based method, for a 59.2% time saving. The training time for the ensemble of ANNs is 3.44 s.

The network system reliability is computed by means of the survival signature estimated with the ANN-based method, for a mission time $T_m = 1$ in arbitrary units of time, and it is benchmarked with the reliability computed using the survival signature approximated by MCS (Fig. 12). The error bounds are plotted for both methods.

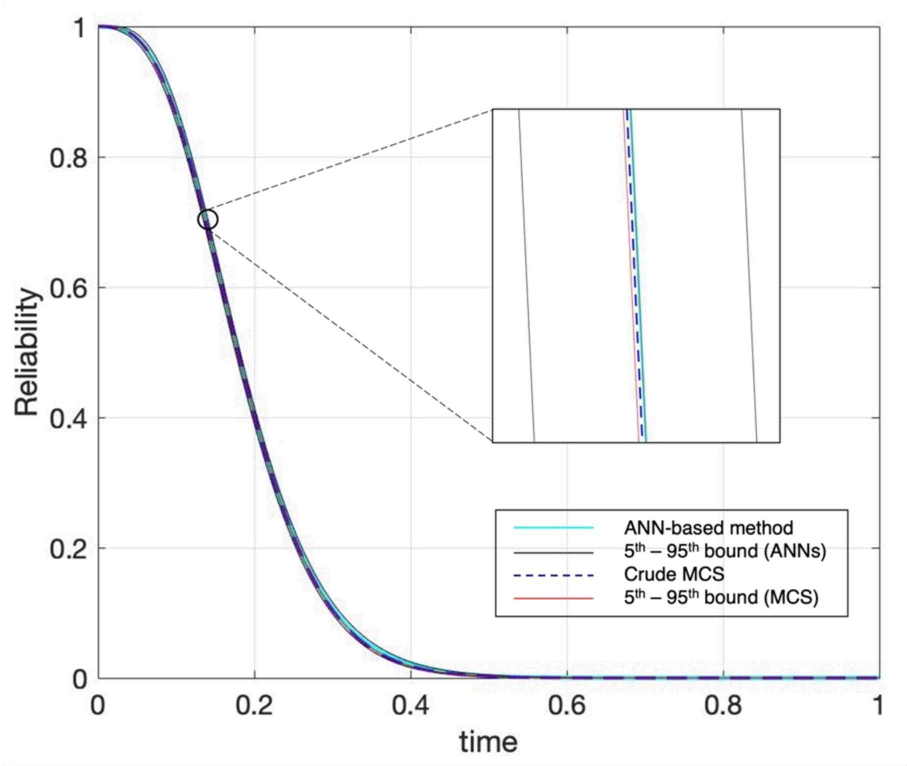

Figure 12 – Survival function of the system in Fig. 11 computed with the survival signature approximated by the crude MCS method (blue dashed line), with corresponding error bounds in red, and approximated by the ANN-based method (light blue), with corresponding error bounds in black.

Fig. 12 clearly shows that the two reliability curves match well. A more detailed analysis outlines that the error bounds of the ANN-based method are larger than the MCS ones, but such difference is negligible on the overall reliability estimation and this is confirmed by the Kullback-Leibler divergence for the two reliability functions, $D_{KL}(rel(t)_{MCS} \parallel rel(t)_{ANN}) = 1.46$ E-03.

5.3. Reduced Berlin metro network

We consider the simplified representation of the Berlin metro system [20] given in Fig. 13. The network system comprises of $m = 50$ stations, which have been divided in $K = 2$ types, based on their number of connections: stations with a maximum of three connections are grouped in type 1, $m_1 = 22$, and stations with more than three connections are grouped in type 2, $m_2 = 28$. We assume that stations of type 1 have disservice failure times following an exponential distribution with $\lambda = 1$ in arbitrary units of inverse time, whereas stations of type 2 have disservice failure times following a Weibull distribution with parameters $a = 1$ and $b = 2$.

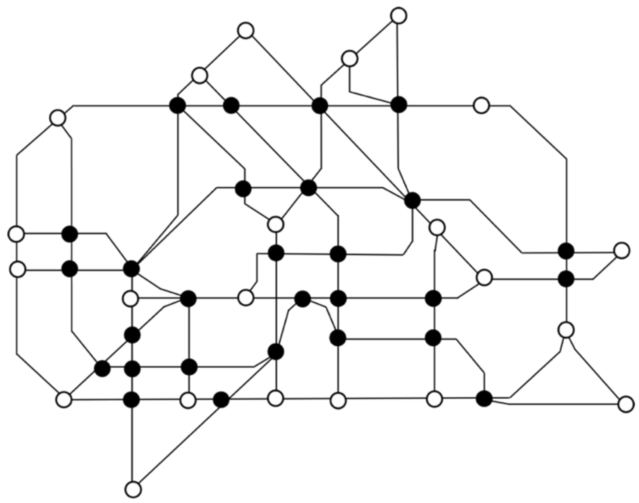

Figure 13 – Simplified representation of the Berlin metro system with 50 stations: those of type 1 are in white, those of type 2 are in black (adapted from [20]).

The network system survival signature is an array with dimension 23×29 and it is approximated using the ANN-based method presented in Section 3. The percolation threshold for the network system is $f_c = 0.6642$, which is used to solve $P = 153$ of the $M = 667$ entries of the survival signature. Among the remaining entries, $Q = 314$ are randomly selected and approximated by MCS. The ensemble of ANNs, comprising FNNs with $n_{layer} = 1$ and $n_{neuron} = [20\ ,30]$, is trained and used to retrieve the missing $S = 200$ entries. The confidence intervals for the survival signature estimates are built with $\alpha = 90\%$, with a resulting average width of 1.58E-02 and a coverage value of 96.5%.

For completeness, a comparative study with the crude MCS method is given. The time needed to approximate the survival signature using MCS is compared with the time needed using the ANN-based method: the former is 652.71 s, the latter is 351.92 s, with 46.1% time saving. The training time for the ensemble of ANNs is 5.56 s.

The survival signature estimated with the ANN-based method is used to compute the network system reliability (Eq. (2)) for a mission time $T_m = 1$ in arbitrary units of time. In Fig. 14, this is benchmarked to the reliability obtained using the survival signature approximated by MCS, and, for illustrative purposes, the error bounds, in terms of 5th and 95th percentiles of the survival signature, are given for the two methods.

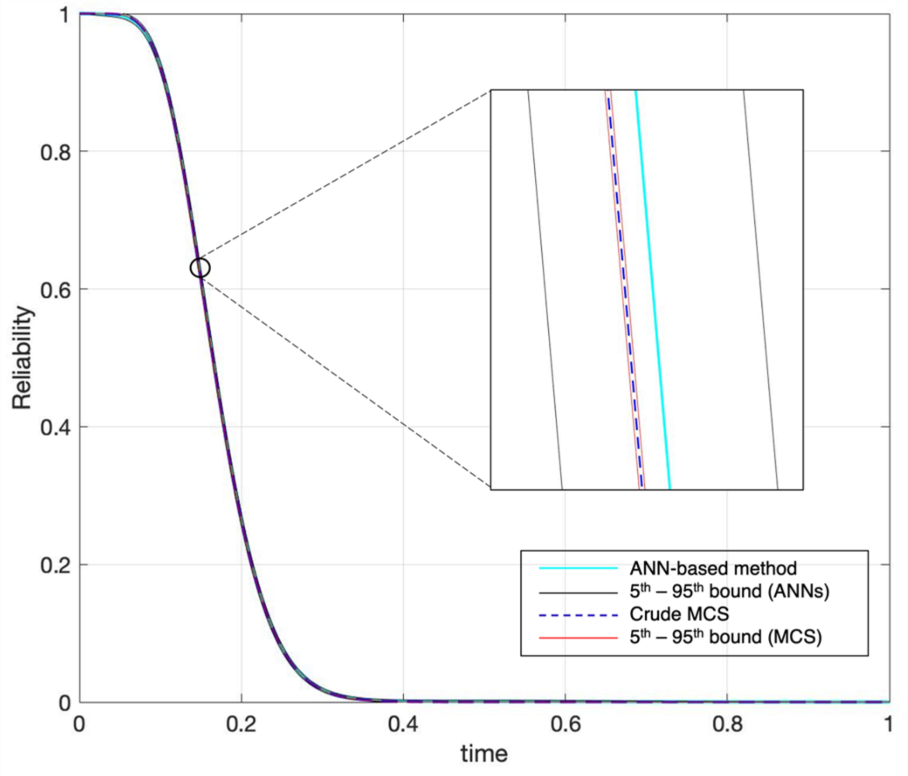

Figure 14 – Survival function of the system in Fig. 13 computed with the survival signature approximated by the crude MCS method (blue dashed line), with corresponding error bounds in red, and approximated by the ANN-based method (light blue), with corresponding error bounds in black.

There is a clear agreement between the two reliability functions, and the difference in the error bounds width is not discernible on the overall reliability evaluation. The accuracy of the ANN-based method in approximating the survival signature, and, eventually, the network system reliability, is also confirmed by the value of the Kullback-Leibler for the two functions, $D_{KL}(rel(t)_{MCS} \parallel rel(t)_{ANN})$ = 6.43E-04.

5.4. IEEE 118-Bus Case: American Electric Power System

We consider the network system in Fig. 15, which represents a portion of the American Electric Power System (IEEE 118-Bus System) as presented in [27] (2003 version). The network system comprises of $m = 118$ buses, which have been divided in $K = 2$ types according to their bus type. In detail, $m_1 = 64$ nodes are load-controlled buses and $m_2 = 54$ nodes are voltage-controlled buses (i.e., the buses which generators are connected to). For the purpose of this work, we assume that components of type 1 have failure times following an exponential distribution with $\lambda = 1$ in arbitrary units of inverse time, whereas components of type 2 have failure times following a Weibull distribution with parameters $a = 1$ and $b = 2$.

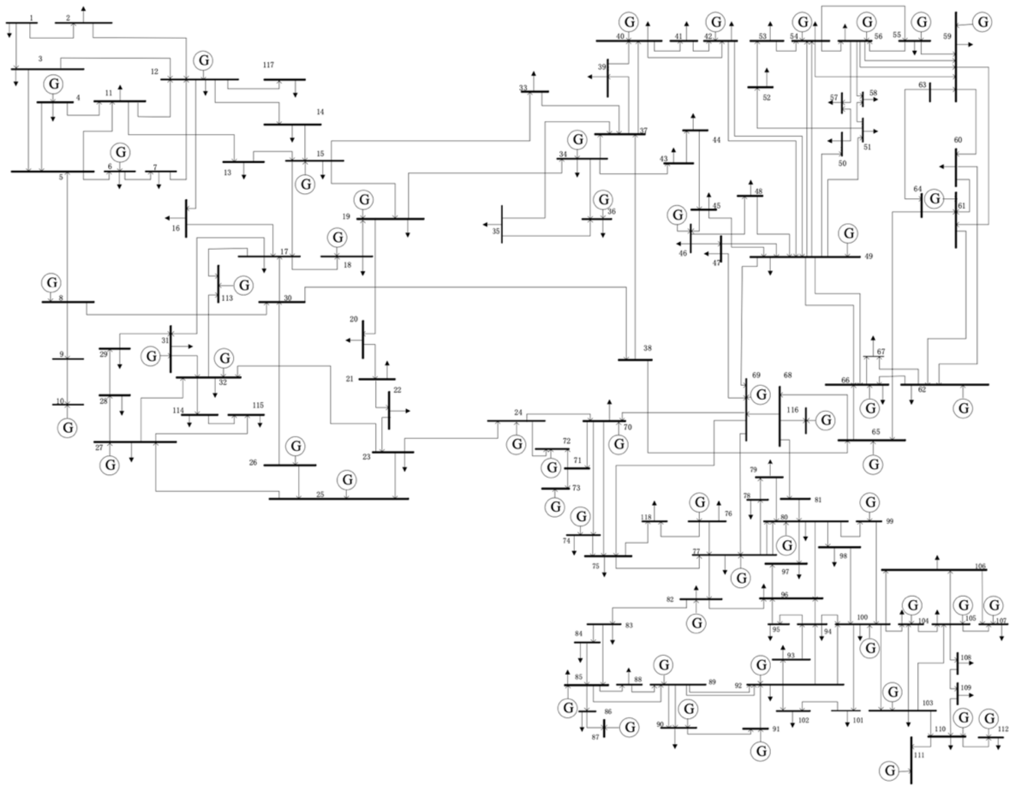

Figure 15 – IEEE 118-Bus System Case [27] (2003 version*). Generators denoted by the symbol "G".*

The network system survival signature is an array of dimension 65×45 and it is here approximated with the ANN-based method. The percolation threshold is $f_c = 0.6500$, which is used to solve $P = 903$ of the $M = 3575$ entries of the survival signature. $Q = 1600$ among the remaining entries are randomly selected and approximated by MCS, so that a sparse survival signature array comprising 70% of the entire survival signature is available. The ensemble of ANNs consisting of FNNs with $n_{layer} = 2$ and $n_{neuron} = [10\ , 20]$, $H = 121$ FNNs in total, is trained and used to estimate the missing 30% of the survival signature entries, i.e., $S = 1072$ entries. The confidence intervals for the survival signature estimates are built with a confidence level $\alpha = 90\%$: the average confidence interval width is 3.7E-03 and the coverage is 91.6%.

The result provided by the ANN-based method is compared to that of the crude MCS. The time needed to approximate the survival signature using MCS is 34779.5 s, whereas that using the ANN-based method is 17149.7 s, with 50.7% time saving. The training time for the ensemble is 64.2 s.

Eventually, the survival signature estimated with the ANN-based method is used to compute the reliability of the network system for a mission time $T_m = 1$ in arbitrary units of time, and this is benchmarked with the reliability obtained using the survival signature approximated by MCS (Fig. 16). The error bounds computed using the 5^{th} and 95^{th} percentiles of the survival signature estimates are provided.

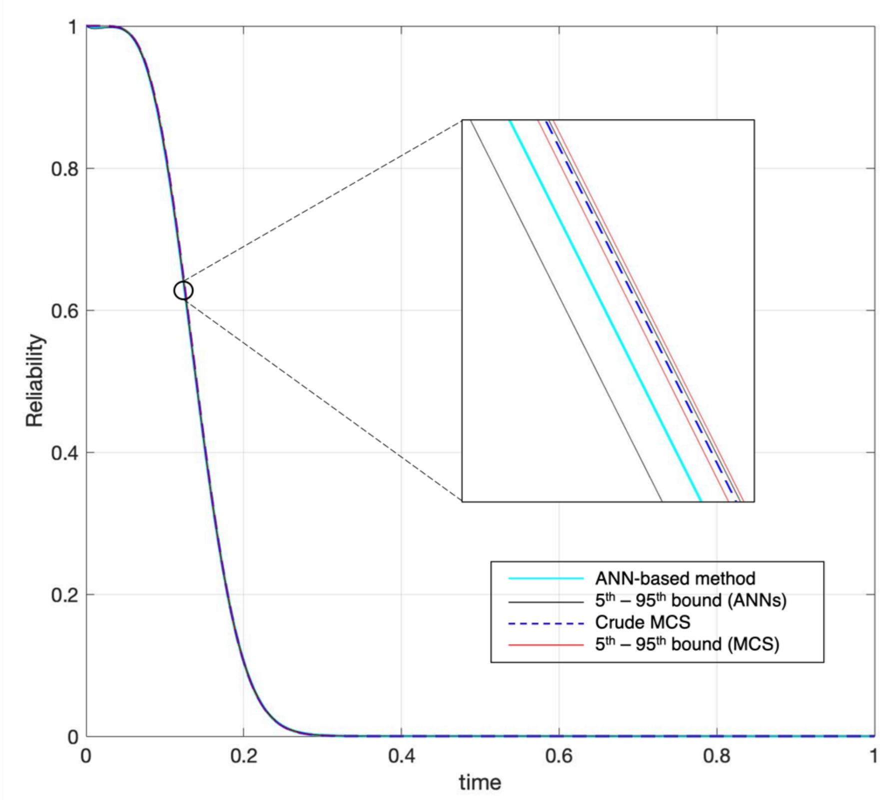

Figure 16 – Survival function of the system in Fig. 15 computed with the survival signature approximated by the crude MCS method (blue dashed line), with corresponding error bounds in red, and approximated by the ANN-based method (light blue), with corresponding error bounds in black.

Fig. 16 shows the good agreement between the two reliability functions. The error bounds for the MCS method are smaller than for the ANN-based method, but such difference is not influencing the overall accuracy, and this is confirmed by the Kullback-Leibler divergence of the two reliability functions, $D_{KL}(rel(t)_{MCS} \parallel rel(t)_{ANN})$ = 2.32 E-04.

6. CONCLUSION

In this work, the reliability assessment of CIs has been considered. To cope with the complexity of these systems, the survival signature is used for rendering feasible the reliability calculation. However, the computation of the survival signature suffers from the curse of dimensionality and its calculation is unacceptably demanding for large-scale CIs. To circumvent the issue, the estimation of the survival signature has been formulated as a missing data problem and a novel method for its efficient approximation has been developed, based on the combination of percolation theory, MCS and ensemble of ANNs. The ANNs of the ensemble have different number of layers and of neurons, and are trained on partial survival signature information obtained with a limited number of MCS; they are trained to retrieve the missing data, i.e., for estimating the missing entries of the survival signature. The original contribution of the work lies in that it allows avoiding the full exploration of the network system state space for approximating the survival signature, either analytically or by means of simulations. Instead, the survival signature is estimated by an ensemble of ANNs which is trained on a limited set of survival signatures obtained by MCS, implying a significant reduction of the computational cost, which is of major importance when large and complex network systems are analyzed.

The feasibility of the method proposed has been shown on a numerical example, which has also allowed to indicate general guidance for selecting the architecture and elements of the ensemble of ANNs. The practical applicability of the method has been shown on four real-world CIs models, i.e., the electricity transmission network of Great Britain, the New England power system (IEEE 39-Bus Case), the Berlin metro network and the American electric power system (IEEE 118-Bus Case). In all cases, relevant time savings have been achieved, without compromising the accuracy on the estimates. The value of the Kullback-Leibler divergence between the reliability calculated with MCS and the ANN-based method is listed in Table 3, for all the network systems analyzed.

Table 3 – Kullback-Leibler divergence D_{KL} for the network systems presented in Section 5.

Case Study	$D_{KL}(rel(t)_{MCS} \parallel rel(t)_{ANN})$
1	2.51E-02
2	1.46 E-03
3	6.43E-04
4	2.32 E-04

The bottleneck of the method from the point of view of the computational effort remains the MCS, and future work will be devoted to address this issue.

Acronyms

ANNs	Artificial Neural Networks
CIs	Critical Infrastructures
FNNs	Feedforward Neural Networks
MCS	Monte Carlo Simulation
MSE	Mean Squared Error

Symbols

$\underline{x}$	System Boolean state vector
x_i	Boolean state of the *i*-th component of the system (system of one component type)
$\underline{x}^k$	Boolean state vector of components of type k of the system
x_i^k	Boolean state of the *i*-th component of type k of the system
m	Total number of components of the system
m_k	Number of components of type k of the system
k	Index of component types, $k = 1, \dots K$
K	Number of component types in the system
φ	Structure function of the system
l_k	Number of functioning components of type k

$\underline{l}$	Vector of the number of functioning components of each type
$S_{l_1,l_2,\ldots,l_K}$	Set of all the Boolean state vector satisfying $\underline{l} = l_1, l_2, \ldots, l_K$
Φ	Survival signature
M	Total number of survival signature entries
$P(T_s > t)$	System survival function at time t
C_t^k	Number of components of type k functioning at time t
$F_k(t)$	Cumulative distribution function for the failure time of components of type k
j	Index of survival signature entry, $j = 1, \ldots, M$
$\{\underline{l}_j\}$	Set of all the survival signature entries
$\{\underline{l}_j\}_P$	Subset of the survival signature entries solved by percolation
P	Magnitude of the subset of the survival signature entries solved by percolation
$\{\underline{l}_j\}_Q$	Subset of the survival signature entries solved by MCS
Q	Magnitude of the subset of the survival signature entries solved by MCS
$\{\underline{l}_j\}_T$	Subset of the survival signature entries in the training dataset
T	Magnitude of the subset of the survival signature entries in the training dataset
$\{\underline{l}_j\}_S$	Subset of the survival signature entries solved by the ensemble of ANNs
S	Magnitude of the subset of the survival signature entries solved by the ensemble of ANNs
f_c	Percolation threshold
d	Network system node degree
$\langle d \rangle$	First momentum of the node degree distribution over the network system
$\langle d^2 \rangle$	Second momentum of the node degree distribution over the network system
N	Number of nodes in the network system
$\{a_{ij}\}$	Adjacency matrix
$E(G)$	Efficiency of the network system G
r_{ij}	Shortest path length between node i and j
$\underline{x}_{\underline{l}}$	Generic system Boolean state vector for a system with $\underline{l}$ components functioning
n	Number of samples for MCS

X_{TRAIN}	Training dataset for the ensemble of ANNs
$\widehat{\Phi}(\underline{l})^h$	Estimated survival signature for entry $\underline{l}$ obtained by the h-th model of the ensemble
h	Index of ensemble models, $h = 1, \dots, H$
H	Number of single models in the ensemble
$\widehat{\Phi}(\underline{l})$	Estimated survival signature for entry $\underline{l}$ obtained by the ensemble
$\sigma_{\underline{l}}^2$	Estimate error variance
$\widehat{\Phi}(\underline{l})^{low}$	Lower estimate of the survival signature for entry $\underline{l}$
$\widehat{\Phi}(\underline{l})^{up}$	Upper estimate of the survival signature for entry $\underline{l}$
α	Predefined confidence level
C_{dof}^{α}	$(1-\alpha)/2$ quantile of a Student t-distribution with a number of degree of freedom
n_{layer}	Number of layers of the FNNs in the ensemble
n_{neuron}	Number of neurons in a layer of the FNNs in the ensemble
λ	Parameter of the exponential distribution of the components failure time
a	Shape parameter of the Weibull distribution of components failure time
b	Scale parameter of the Weibull distribution of components failure time
T_m	System mission time
D_{KL}	Kullback-Leibler divergence
$rel(t)_{MCS}$	Reliability estimated with MCS
$rel(t)_{ANN}$	Reliability estimated with the ensemble of ANNs

References

[1] J. Johansson, H. Hassel, E. Zio, “Reliability and vulnerability analyses of critical infrastructures: Comparing two approaches in the context of power systems,” *Reliab Eng Syst Saf*, vol. 120, pp. 27–38, 2013, doi: 10.1016/j.ress.2013.02.027.

[2] S. M. Rinaldi, J. P. Peerenboom, T. K. Kelly, "Identifying, understanding, and analyzing critical infrastructure interdependencies," *IEEE Control Systems Magazine*, vol. 21, no. 6, pp. 11–25, 2001, doi: 10.1109/37.969131.

[3] K. M. Zuev, M. Beer, "Special Issue on Complex Engineered Networks: Reliability, Risk, and Uncertainty," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 3, no. 2, 2017, doi: https://doi.org/10.1115/1.4036240.

[4] J. R. Piacenza, S. Proper, M. A. Bozorgirad, C. Hoyle, I. Y. Tumer, "Robust Topology Design of Complex Infrastructure Systems," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 3, no. 2, Jun. 2017, doi: https://doi.org/10.1115/1.4036152.

[5] I. Eusgeld, C. Nan, S. Dietz, "System-of-systems approach for interdependent critical infrastructures," *Reliab Eng Syst Saf*, vol. 96, no. 6, pp. 679–686, 2011, doi: 10.1016/j.ress.2010.12.010.

[6] F. P. A. Coolen, T. Coolen-Maturi, *Generalizing the signature to systems with multiple types of components*, vol. 170 AISC. Complex systems and dependability. Advances in intelligent and soft computing, Berlin, Heidelberg: Springer, vol. 170, pp. 115-30, 2012. doi: 10.1007/978-3-642-30662-4-8.

[7] X. Zhuang, T. Yu, L. Shen, B. Guo, "Dynamic Reliability Evaluation of Nonrepairable Multistate Weighted k-Out-of-n System With Dependent Components Based on Copula," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 4, no. 4, 2018, doi: https://doi.org/10.1115/1.4039243.

[8] X. Liu, Y.-P. Fang, E. Ferrario, E. Zio, "Resilience Assessment and Importance Measure for Interdependent Critical Infrastructures," *ASCE-ASME Journal of Risk and*

Uncertainty in Engineering Systems, *Part B: Mechanical Engineering*, vol. 7, no. 3, 2021, doi: https://doi.org/10.1115/1.4051196.

[9] F. P. Coolen, T. Coolen-Maturi, "Modelling uncertain aspects of system dependability with survival signatures," *Advances in Intelligent Systems and Computing*, vol. 307, pp. 19–34, 2015, doi: 10.1007/978-3-319-08964-5_2.

[10] G. Feng, E. Patelli, M. Beer, F. P. A. Coolen, "Imprecise system reliability and component importance based on survival signature," *Reliab Eng Syst Saf*, vol. 150, pp. 116–125, 2016, doi: 10.1016/j.ress.2016.01.019.

[11] L. J. M. Aslett, F. P. A. Coolen, S. P. Wilson, "Bayesian Inference for Reliability of Systems and Networks Using the Survival Signature," *Risk Analysis*, vol. 35, no. 9, pp. 1640–1651, 2015, doi: 10.1111/risa.12228.

[12] F. Di Maio, C. Pettorossi, E. Zio, "Entropy-driven Monte Carlo simulation method for approximating the survival signature of complex infrastructures," *Reliab Eng Syst Saf*, vol. 231, 2023, https://doi.org/10.1016/j.ress.2022.108982.

[13] A. Scherb, L. Garre, D. Straub, "Reliability and Component Importance in Networks Subject to Spatially Distributed Hazards Followed by Cascading Failures," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 3, no. 2, 2017, doi: https://doi.org/10.1115/1.4036091.

[14] J. Behrensdorf, M. Broggi, M. Beer, "Reliability Analysis of Networks Interconnected With Copulas," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 5, no. 4, 2019, doi: https://doi.org/10.1115/1.4044043.

[15] X. Liu, E. Ferrario, E. Zio, "Resilience Analysis Framework for Interconnected Critical Infrastructures," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems,*

Part B: Mechanical Engineering, vol. 3, no. 2, 2017, doi: https://doi.org/10.1115/1.4035728.

[16] S. Miro, T. Willeke, M. Broggi, J. R. Seume, M. Beer, "Reliability Analysis of an Axial Compressor Based on One-Dimensional Flow Modeling and Survival Signature," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 5, no. 3, 2019, doi: https://doi.org/10.1115/1.4043150.

[17] S. Reed, "An efficient algorithm for exact computation of system and survival signatures using binary decision diagrams," *Reliab Eng Syst Saf*, vol. 165, pp. 257–267, 2017, doi: 10.1016/j.ress.2017.03.036.

[18] J. Behrensdorf, T.-E. Regenhardt, M. Broggi, M. Beer, "Numerically efficient computation of the survival signature for the reliability analysis of large networks," *Reliab Eng Syst Saf*, vol. 216, 2021, doi: 10.1016/j.ress.2021.107935.

[19] D. Li, Q. Zhang, E. Zio, S. Havlin, R. Kang, "Network reliability analysis based on percolation theory," *Reliab Eng Syst Saf*, vol. 142, pp. 556–562, 2015, doi: 10.1016/j.ress.2015.05.021.

[20] J. Behrensdorf, S. Brandt, M. Broggi, M. Beer, "Efficient Approximation of the Survival Signature for Large Networks," *Proceedings of the 6th International Symposium on Reliability Engineering and Risk Management*, 2018, doi: 10.3850/978-981-11-2726-7_CRR14.

[21] L. K. Hansen, P. Salamon, "Neural Network Ensembles," *IEEE Trans Pattern Anal Mach Intell*, vol. 12, no. 10, pp. 993–1001, 1990, doi: 10.1109/34.58871.

[22] M. Compare, F. di Maio, E. Zio, F. Carlevaro, S. Mattafirri, "Improving scheduled maintenance by missing data reconstruction: A double-loop Monte Carlo approach," *Proc Inst Mech Eng O J Risk Reliab*, vol. 230, no. 5, pp. 502–511, 2016, doi: 10.1177/1748006X16651988.

[23] S. Al-Dahidi, P. Baraldi, E. Zio, M. Lorenzo, “Bootstrapped ensemble of artificial neural networks technique for quantifying uncertainty in prediction of wind energy production,” *Sustainability (Switzerland)*, vol. 13, no. 11, 2021, doi: 10.3390/su13116417.

[24] P. Baraldi, F. Mangili, E. Zio, “Ensemble of bootstrapped models for the prediction of the remaining useful life of a creeping turbine blade,” *PHM 2012 - 2012 IEEE Int. Conf.on Prognostics and Health Management: Enhancing Safety, Efficiency, Availability, and Effectiveness of Systems Through PHM Technology and Application, Conference Program*, 2012, doi: 10.1109/ICPHM.2012.6299506.

[25] WA Bukhsh, K. McKinnon, “Network data of real transmission networks,” 2013. https://www.maths.ed.ac.uk/optenergy/NetworkData/index.html.

[26] T. Athay, R. Podmore, S. Virmani, “A practical method for the direct analysis of transient stability,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, no. 2, pp. 573–584, 1979, doi: 10.1109/TPAS.1979.319407.

[27] “Index of Data Illinois Institute of Technology.” http://motor.ece.iit.edu/data/

[28] R. S. Wilkov, “Analysis and Design of Reliable Computer Networks,” *IEEE Transactions on Communications*, vol. 20, no. 3, pp. 660–678, 1972, doi: 10.1109/TCOM.1972.1091214.

[29] M. Molloy, B. Reed, “A critical point for random graphs with a given degree sequence,” *Random Structures & Algorithms*, vol. 6, no. 2–3, pp. 161–180, 1995, doi: 10.1002/rsa.3240060204.

[30] 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.

[31] R. W. Floyd, “Algorithm 97: Shortest path,” *Commun ACM*, vol. 5, no. 6, p. 345, 1962, doi: 10.1145/367766.368168.

[32] E. Zio, “A study of the bootstrap method for estimating the accuracy of artificial neural networks in predicting nuclear transient processes,” *IEEE Trans Nucl Sci*, vol. 53, no. 3, pp. 1460–1478, 2006, doi: 10.1109/TNS.2006.871662.