Integrated deterministic and probabilistic safety assessment of a superconducting magnet cryogenic cooling circuit for nuclear fusion applications

Integrated Deterministic and Probabilistic Safety Assessment of a Superconducting Magnet Cryogenic Cooling Circuit for Nuclear Fusion Applications

R. Bellaera[1], R. Bonifetto[1], F. Di Maio[2], N. Pedroni[1], L. Savoldi[1], R. Zanino[1], E. Zio[2,3,4]

[1] *NEMO group, Dipartimento di Energia, Politecnico di Torino, Torino, Italy*
[2] *Dipartimento di Energia, Politecnico di Milano, Italy*
[3] *MINES ParisTech, PSL Research University, CRC, Sophia Antipolis, France*
[4] *Eminent scholar at Kyung Hee University, Republic of Korea*

ABSTRACT: The most promising configuration of a nuclear energy fusion system is the tokamak, the largest of which, called ITER, is under construction in Cadarache, France, which uses a complex system of superconducting magnets to generate a field of several tesla (T), aimed at confining the plasma in the toroidal chamber where nuclear fusion reactions occur. For industrial development, the safety of nuclear fusion systems has to be proved and verified by a systematic analysis of operational transients and accidental conditions. Although the final aim of fusion reactors is to reach steady state operation, present-day tokamaks present complex dynamic features, as their operation is based on the transformer principle with a subset of the superconducting magnets operating in a pulsed mode, to inductively generate plasma currents of the order of several MA. We adopt the framework of Integrated Deterministic and Probabilistic Safety Assessment (IDPSA), for identifying the component failures that may cause a Loss-Of-Flow-Accident (LOFA) in the cooling circuit of a superconducting magnet for fusion applications. Post-processing of the simulated scenarios for the identification of the abnormal transients is performed in an unsupervised manner resorting to a spectral clustering approach embedding a Fuzzy-C Means (FCM) that is compared with an Extended Symbolic Aggregate approXimation (ESAX) from the literature that also resorts to the FCM for the classification.

The proposed approach turns out to be more efficient than ESAX in the identification of clusters and "prototypical states" of abnormal system behavior. Results show that none of the identified scenarios (even those leading to a LOFA) are critical for the ITER central solenoid module integrity, in the mode of operation considered.

KEYWORDS: nuclear fusion, superconducting magnets, cryogenic cooling circuit, IDPSA, spectral clustering, Extended Symbolic Aggregate approXimation.

1 INTRODUCTION

Nuclear fusion reactors will use superconducting magnets to generate the powerful magnetic field needed to confine the plasma in the shape of a torus, where Deuterium-Tritium fusion reactions occur at a temperature of the order of 10^8 K. In ITER, the largest tokamak fusion reactor under construction at Cadarache (France) (ITER, 2014), the three principal magnetic sub-systems are (see Figure 1):

- the Toroidal Field (TF) coils, which operate in DC and generate the magnetic field component along the big circumference of the torus (*toroidal* direction), which can reach several T;
- the Central Solenoid, which operates in pulsed mode and is the primary of a transformer which induces in the secondary, i.e. the plasma, a very large current (several MA), which in turn generates the component of the magnetic field along the short circumference of the torus (*poloidal* direction);
- the Poloidal Field (PF) coils, which operate with time dependent currents defining the so-called plasma scenario, and generate a vertical component of the magnetic field which gives the desired shape (e.g., elliptical with some added triangularity) to the plasma cross section.

All coils need to be cooled at cryogenic temperatures in order to avoid their quench (i.e., the progressive and potentially divergent loss of the current carrying capability in superconducting state, with associated Ohmic heat generation) during operation. For example, the ITER SC coils are cooled by supercritical helium (SHe) at a pressure of 0.5–0.6 MPa and temperature of about 4.5 K (Mitchell et al., 2008). Dedicated cryogenic cooling loops remove the heat load from the magnets, releasing it to saturated liquid helium (LHe) pools, which act as thermal buffers in the transfer of the peak load to the refrigerator (Hoa et al., 2012).

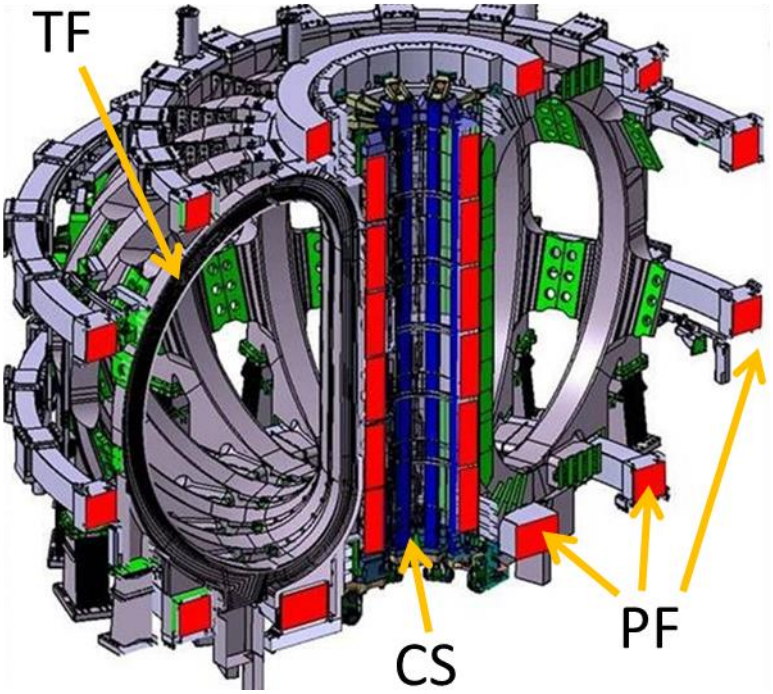


Figure 1 The ITER TF magnets (ITER).

The safety of nuclear fusion systems has to be proved and verified by a systematic analysis of the system behavior under operational transients and accidental conditions [e.g., loss of coolant accidents (Rivas et al., 2015)], and the corresponding (safety) issues and gaps have to be highlighted (Perrault, 2016; Wu et al., 2016). Indeed, contamination from radioactive sources (e.g., tritium and the materials activated by the neutrons produced by the fusion reactions) must be avoided for the operators and public safety, and for the environment (Taylor and Cortes, 2014; Taylor et al., 2017). Also, the high cost of the superconducting magnet system, as well as the essential need to guarantee at least a lifetime coincident with that of the plant, calls for its protection and integrity (Mitchell et al., 2012).

One challenge to the analysis of these issues is that the operation of tokamaks presents *complex dynamic features*, as seen above. Then, in the safety analysis the *order* and *timing* of the failure events occurring in an accident scenario, the *magnitude* of the failures and the *values* of the process variables at the time of failures occurrence could be critical in determining the response of the system (Zio and Di Maio, 2009).

The Integrated Deterministic and Probabilistic Safety Assessment (IDPSA) framework (Kirschenbaum et al., 2009) allows such an analysis by combining (deterministic) phenomenological models of the system dynamics with (probabilistic) failure models (Zio and Di Maio, 2010). IDPSA techniques have been successfully applied to analyze nuclear fission systems: see, e.g., (Galushin and Kudinov, 2015) for an application to a hypothetical LOCA transient in typical French 900 MWe PWR; (Jankovsky et al., 2018) for the dynamic analysis of a sodium

cooled fast reactor; and (Grishchenko et al., 2019) for the uncertainty quantification of a steam explosion scenario in a Nordic type BWR. In this work, IDPSA is employed, to analyze the response to abnormal transient conditions of the cooling system of a SC magnet for nuclear fusion applications, namely a *single* ITER Central Solenoid Module (CSM) in a reference (cold) *test facility* (Spitzer et al., 2015). The analysis focuses on identifying the failures that may lead the system to a Loss-Of-Flow Accident (LOFA). The state-of-the-art thermal-hydraulic code 4C (Savoldi Richard et al., 2010), validated against different types of transients characterized by a broad range of time scales [e.g., safety discharge (Zanino et al., 2011) and quench propagation (Zanino et al., 2018) in the ITER TF coils, and cooldown of the ITER CS coils (Bonifetto et al., 2017)] is employed for the (deterministic) simulation of the system behavior. Multiple Value Logic (MVL) is adopted for the description of the component failures at different times and with *different* (discrete) *magnitudes* (Di Maio et al., 2015). The combinatorial explosion of the MVL scenarios generated must be post-processed to extract relevant information on the response of the system in relation to the different states of the components. To post-process the scenarios, clustering is adopted. Several clustering algorithms have been proposed and used in practice, such as Fuzzy C-Means (FCM) (Bezdek, 1981) and Symbolic Aggregate approXimation (SAX) (Lin et al., 2003) to find the prototypical behaviours of systems working in abnormal conditions (Di Maio et al., 2012). In this work, we use Spectral Clustering (Von Luxburg, 2007) embedding a FCM classifier, which has been proven to be a powerful unsupervised approach in various dynamic failure analysis (Baraldi et al., 2013), as it will be explained in Section 3.2.1.

The scenarios are modelled here as MVL scenarios, simulated by 4C and eventually grouped by the Spectral Clustering. From the clusters, relevant safety insights are obtained (i.e., the component failures that lead the system into a LOFA). A comparison with reference to the Silhouette and Davies-Bouldin indexes (Davies and Bouldin, 1979), explained in Section 3.2.3, is performed with the ESAX method (Butler and Kazakov, 2015), that, as well as Spectral Clustering, embeds the FCM for the classification task.

The paper is organized as follows. In Section 2, a description of a single ITER CSM in the cold test facility is presented together with the corresponding simulator; the different regimes of system operation are also described. In Section 3, the IDPSA methodology is presented: in particular, the component failures causing the deviations from the nominal CSM conditions in the test facility are described, together with the use of MVL

for generating accident scenarios; also, the spectral clustering and the ESAX algorithms for scenario post-processing are explained in detail. In Section 4, the results are presented and critically analyzed. Finally, conclusions are summarized in Section 5.

2 DESCRIPTION OF THE PHYSICAL SYSTEM AND PRESENTATION OF THE NUMERICAL SIMULATOR

The ITER Central Solenoid (CS) system allows inducing the current in the plasma and maintaining it during long plasma pulses. It is composed of 6 CS modules (CSM) vertically stacked, see Figure 2a. Each module is currently being manufactured and will be individually tested; it is composed of 7 pancake-wound conductors, specifically 6 hexa-pancakes and 1 quad-pancake. Each pancake is cooled in parallel, resulting in 40 parallel cooling channels per module, each featuring 14 turns. The He inlets are located at the coil bore, while the outlets are at the outer side of the magnet, see again Figure 2a. All the pancakes of each module are electrically connected in series through suitable joints (Libeyre et al., 2015).

The main components of a reference facility for the CSMs *cold tests*, reported in Figure 2b, are the He refrigerator, producing the supercritical He (SHe) at 4.5K used to cool the magnets during the tests, and the test chamber where the module will be put into a cryostat. Besides these two main components, several manifolds, pipelines, control valves (CV) of the heat exchangers (HXs) of the cryoplant, and a liquid He (LHe) thermal buffer (enclosed in a suitable cryo-distribution cold box) will connect the refrigerator to the coil. A dedicated power supply system will provide a current up to ~50kA, while a control system ensures that all the (electrical, thermal-hydraulic, etc.) operating parameters are within the given operational range. Finally, a protection system constantly monitors several variables and is ready to intervene to protect the coil integrity in case of failure.

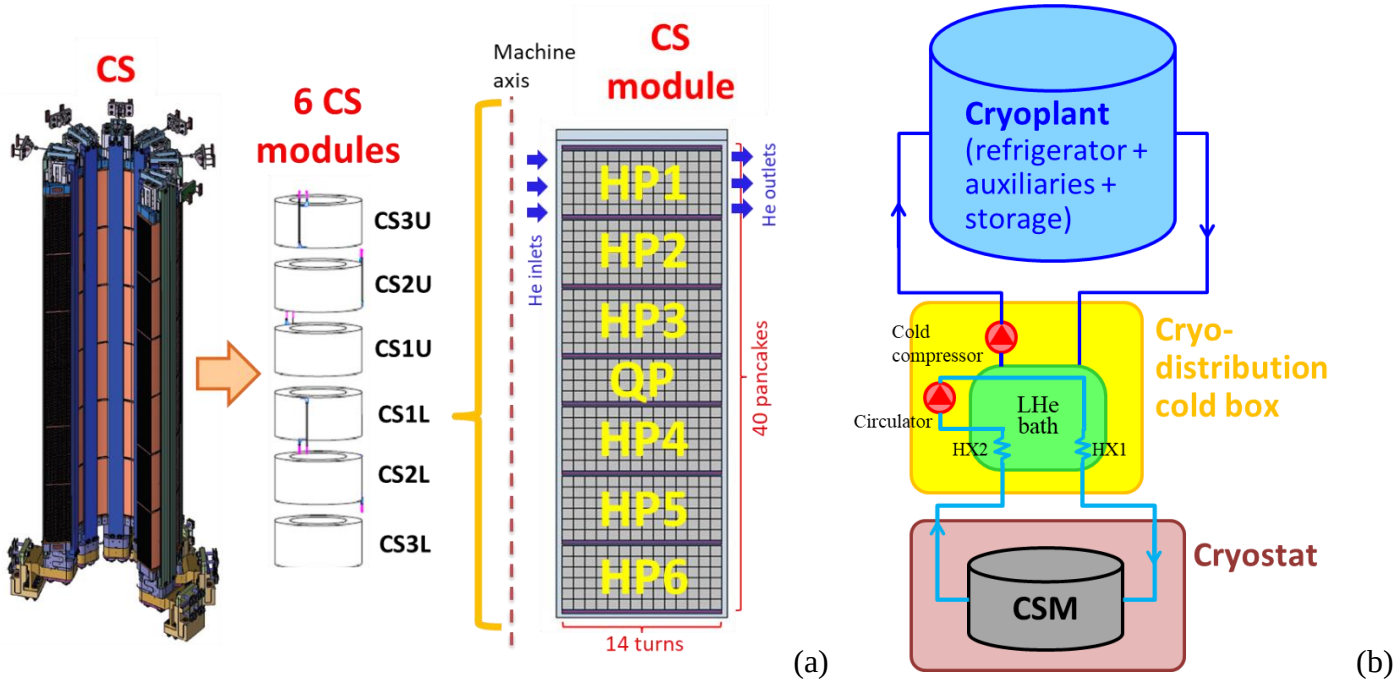


Figure 2 (a) Sketch of the ITER CS and its subcomponents (U = upper, L = lower, HP = hexa-pancake, QP = quad-pancake), partially taken from (Libeyre et al., 2015); (b) scheme of the facility for the CSM cold tests taken as reference in the present work.

From the hydraulic point of view, the analysis reported here is focused on the cooling loop of a single CSM, similar to that of the reference test facility. This closed loop provides SHe at ~4.5K to the inlets of the 40 hydraulic paths, collects the (warmer) SHe at their outlets and drives it by means of a cold circulator to two HXs, where the heat removed from the coil is transferred to a LHe buffer. The He evaporated in the buffer is, eventually, extracted and cooled down by the refrigerator.

The 4C thermal hydraulic code (Savoldi Richard et al., 2010) is used to model both the coil and its SHe cooling loop. Figure 3 shows a scheme of the loop model (where circles are pressure taps and open triangles are flow meters). The SHe at the outlet of the cold circulator is cooled in HX1 to remove the heat generated by the compression process. Then, it is driven to the coil inlets through CV1 (fully open in normal operation) and a supply cryoline (cryoline1). At the coil outlet, the SHe reaches HX2 through a return cryoline (cryoline2) and CV2 (fully open in normal operation). The main input parameters of each component model are reported in Table 1. For the details of the model of each component, please refer to (Bonifetto et al., 2012). A realistic characteristic of the cold circulator has been implemented in the circuit model, as described in detail in (Zanino et al., 2013) for another system.

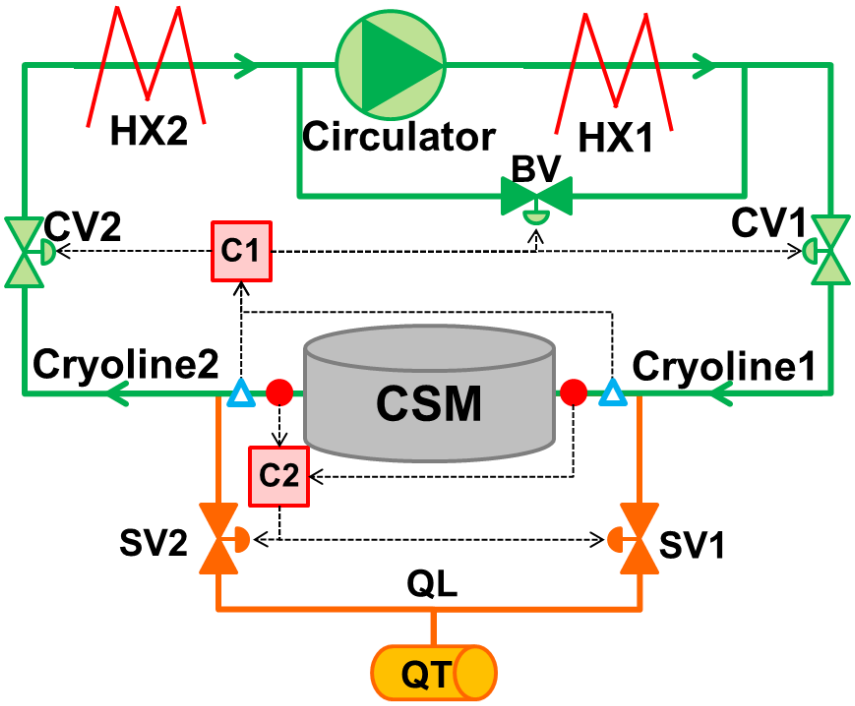


Figure 3 Scheme of the model of the SHe cooling loop of the CSM: HX = heat exchanger, CV = control valve, BV = bypass valve, C = controller, SV = safety valve, QL = quench line, QT = quench tank.

Table 1. Main input parameters of the circuit components models (L = length, D = diameter, K_v = flow coefficient).

Component	L [m]	D [mm]	# of parallel pipes	K_V [m^3/h]
HX1,2	31	20	11	-
Cryoline1	28	46	-	-
Cryoline2	24	46	-	-
CV1,2, BV, SVin, SVout	-	-	-	71

When the LOFA occurs, the protection system behaves similarly to that adopted in ITER as described in (Savoldi et al., 2018) for the Toroidal Field (TF) coils (see Figure 1). In order to protect the CS, a *controlled discharge* of the CS circuits is carried out in ITER, consisting of a *current ramp down* of about 30s (ITER_D_K7G8GN v2.1, 2014), driven by the plasma control system. Notice that the current variation causes AC losses in the SC cables, which induce a *heat deposition*. In the reference system, a similar fast discharge is induced by the protection system in case of LOFA. As far as the cryoplant is concerned, the basic circuit control in case of a LOFA includes the isolation of the circulator from the coil by means of the full closure of both CVs

and the opening of the by-pass valve (BV) to equalize the pressure at the circulator suction and discharge, thus preventing any damage to the pump itself as fail-safe condition. This action is taken by controller C1, as in Figure 3, when a SHe mass flow rate below 10% of the nominal value is measured both at the inlet and at the outlet of the CSM, after a validation time of 1s as in (Savoldi et al., 2018) where it was assumed of the same order of magnitude of the validation time of the primary quench detection system (based on voltage measurements). A different choice of this validation time will possibly lead to different consequences of the series of failures: the parametric analysis on this value, even if required, is outside the scope of the present work. Notice that the "nominal value" of the mass flow rate considered here corresponds to the nominal CSM conditions *during a test* in the reference facility, *not* to the "future" normal operating conditions in ITER. In case of excessive pressurization at the coil boundaries, two safety valves (SVs) at the CSM inlet and outlet open, driven by the Proportional-Integral-Derivative (PID) controller C2 (gain = 1×10^{-7} Pa^{-1}, integration time = 0.2 s, derivation time = 1 s), with set-point 1.8 MPa (Savoldi Richard et al., 2012); the controller parameters and set point are assumed equal to those in (Savoldi et al., 2018). When the SV opens, the He is released in a quench tank (QT), at a pressure of 0.42 MPa, by means of suitable quench lines (QL).

The detailed CSM model solves the 1D transient mass, momentum and energy conservation equations for the Supercritical Helium (SHe), computing the temperature, pressure and velocity distribution in each of the two regions (cable bundle and central, low impedance channel) of each pancake, and two transient heat conduction equations separately for the strands and the jacket, as described in detail in (Zanino et al., 1995). Then, the inter-turn and inter-pancake thermal coupling between adjacent turns and pancakes, respectively, are computed considering the insulation as a thermal resistance to evaluate the heat transfer between neighboring conductors (Savoldi et al., 2000).

2.1 *System operation regimes*

During the tests, the facility and the CSM will be operated in different regimes, which can be briefly described as:

a. Cold mode standby operation (e.g. during night or weekend): the CSM is not charged and kept at a nominal temperature of ~4.5 K. No dangerous consequences are expected due to a LOFA when the system is in these conditions, so this regime is not analyzed.

b. Cold mode experimental operation, i.e. during cold tests, with the coil inlet temperature taken as the nominal one: the CSM is charged. In this regime, the temperature can incidentally increase above the so-called *current sharing temperature* T_{CS} and may lead to a *quench*, i.e., to a loss of the superconducting state and to the consequent fast zero-voltage, local Joule heat deposition that induces thermal stresses seriously damaging the conductor, causing a degradation of its performance or, in the worst case, the loss of integrity of the conductor. The local increase of the conductor temperature above T_{CS} can be detected by measuring the voltage at the extremities of each pancake.

The objective of the analysis is the dynamic response to abnormal conditions of the system in operating mode b. Random series of failures are simulated in order to identify the component failures that drive the system into a LOFA and/or other relevant transients. A set of Z = 3 variables are monitored as "critical indicators" of the state of the system during the transients: the cooling helium pressure p_{cin} at the inlet of the CS magnet, the voltage ΔV measured at the coil extremities and the hot-spot temperature T_{hs} of the strands of the CSM. They are stored in Z different matrixes $\bar{\bar{X}}^k$ (with k = 1, …, Z) each one of size [N, L], where L is the duration of the transients. Actually, if $p_{cin} >$ 25MPa, the conductors can be damaged (ITER_D_2NBKXY v1.2, 2009); also, if $\Delta V >$ 0.1V for more than 1s, it means that the superconducting state of the magnet is irreversibly lost and a fast current discharge must be initiated by the protection system in order to protect the magnet.

3 INTEGRATED DETERMINISTIC AND PROBABILISTIC SAFETY ASSESSMENT (IDPSA)

In classical approaches of safety assessment, a small set of accidental scenarios is analyzed with sequences whose *order* of events is *pre-determined* by the analyst based on his/her *experience* and the failure magnitude is mostly set to a *most conservative* value (Zio, 2014). In complex systems and processes, this may lead to failing to identify some (*a-priori unknown*) vulnerable sequences, which would, then, remain uncovered (Aldemir, 2013). Several examples are given in the open literature, where classical approaches failing to *a priori* identifying vulnerable sequences are shown with respect to fission nuclear power plant Medium Break Loss of Coolant Accident (MLOCA) scenarios (Karanki and Dang, 2016), Steam Generator (SG) accidental scenarios (Di Maio et al., 2016), Station Black Out (SBO), Emergency Operation Plans (EOP) and fire scenarios (Kloos et al., 2018), and feedwater system failure scenarios in a lead fast reactor (Turati et al., 2018). To overcome this problem, the Integrated Deterministic and Probabilistic Safety Analysis (IDPSA) of the evolution of the system

response along an accident scenario (Zio et al., 2010) combines (*deterministic*) *phenomenological* models of system dynamics with (*probabilistic*) process models to account for the *order* and *timing* of the failure events, the *magnitude* of the failures and the *values* of the process variables at the time of event occurrence (Karanki et al., 2017).

The IDPSA methodology comprises three main steps:

1. A set of different accident scenarios is generated. In this paper, Multiple Value Logic (MVL) (Garibba et al., 1985) is adopted to describe the components fail occurrence at any time along the scenario and with different (discrete) magnitudes (Di Maio et al., 2015) (Section 3.1 below);
2. For each scenario, the dynamic response of the system is *deterministically* simulated by the validated 4C Code (Savoldi Richard et al., 2010) (Section 2 above);
3. The scenarios generated and *post-processed* allows getting relevant information on the response of the system and the states of the components. Clustering algorithms can be used to *group* the different scenarios according to a measure of similarity between them. The obtained clusters show "prototypical behaviors" of the system in abnormal conditions and safety insights can be gained from their analysis e.g., the components failures that lead the system into a LOFA. The Spectral Clustering algorithm (Baraldi et al., 2013) is adopted (Section 3.2).

3.1 *Accidental scenarios generation*

The different accident scenarios are generated assuming that the following component failures can occur at uniformly distributed random times in the time horizon [0, 600] s:

1. the Centrifugal Pump (CP) reduces exponentially the rotational speed, directly affecting the mass flow rate that can drop to i) 75%; ii) 50%; iii) 25%; iv) 0% of the nominal value.
2. the two CVs can fail in three different modes: i) stuck (open) at the nominal position; ii) stuck closed at 50% of the nominal position; iii) stuck totally closed.
3. the BV and the two SVs can fail in three different modes: i) stuck (closed) in nominal position; ii) stuck open at 50% of the flow area; iii) stuck totally open.

Notice that in the present analysis only mechanical components failures are considered for which spurious activation is neglected, whereas the instrumentation, control and protection systems are assumed to work always properly. In other words, whenever a LOFA occurs, such hazardous condition is immediately detected, the

controlled discharge of the CS circuits is successfully carried out and both controllers C1 and C2 correctly send out the signals aimed at protecting the integrity of the cooling circuit (see Section 2). In addition, notice that the identification of the components' failure modes listed above has been carried out by engineering judgment. This can be considered sufficient for the purpose of the present study, showing the applicability of IDPSA methodologies to a nuclear fusion system. The use of systematic and structured approaches, e.g., the Failure Mode Effect and Criticality Analysis (FMECA), would be recommended in the safety analysis of real, complex systems.

It has been shown by simulation that the system reaches a steady state condition in ~100 seconds, independently of the failure occurred. Therefore, we consider a mission time T_M = 700 s. On the one hand, the length of this time window allows "injecting" an amount of random failures in the system that guarantees a satisfactorily deep exploration of its (failure) state-space; on the other hand, it allows keeping the overall computational burden relatively manageable for our purposes (the simulation of one transient takes *on average* 55 hours on an Intel(R) Xeon(R)^{TM} CPU X5355 @2.66 GHz).

A Multiple Value Logic (MVL) scheme presented in (Bellaera et al., 2018) has been adopted to describe the N = 100 accidental scenarios generated by random sampling the stochastic (discrete) time (t) of occurrence of component failures, their (discrete) magnitude (m). The random components failures of a generated scenario are represented in an MVL sequence vector that contains the discretized time and magnitudes values and the order (*ord*) of occurrence in the sequence: [m_{CP}, t_{CP}, ord_{CP}, m_{CV1}, t_{CV1}, ord_{CV1}, m_{CV2}, t_{CV2}, ord_{CV2}, m_{BV}, t_{BB}, ord_{BV}, m_{SV1}, t_{SV1}, ord_{SV1}, m_{SV2}, t_{SV2}, ord_{SV2}] (Di Maio et al., 2017), where, for all components CP, CV1, CV2; BV, SV1, SV2, the discretization of the time (t) and magnitude (m) are as follows:

- time (t) : the label t = 1, 2, 3, 4, 5 and 6, is used to distinguish failures occurring in the intervals [0, 100] s, [101, 200] s, [201, 300] s, [301, 400] s, [401, 500] s, [501, 600] s, respectively (the choice of this discretization is based on the characteristics of the system mentioned above, i.e., the fact that it reaches a steady state condition in ~ 100 s); t = 0 means that the component does *not* fail within T_M and the value "NaN" is used to identify the respective (non-)failure order (*ord*) in the sequence vector of the accidental scenario. Notice that, the failure order (*ord*) is not redundant in the MVL representation, because it allows discriminating scenarios in which different components fail in the *same* time *interval t.*
- Magnitude (m) discretization:

- the CP failure magnitude is indicated with the label m_{CP} = 1, 2, 3 or 4 for failure states corresponding to an exponential decrease of the rotational speed down to 75%, 50%, 25% and 0% of the nominal value, respectively; if m_{CP} = 0, the component does not fail;
- for each CV, the magnitude is indicated by the label m_{CV} =1, 2 or 3 if the component stays stuck open at the nominal position, stuck completely closed at 50% of the nominal position and stuck closed, respectively; if m_{CV} = 0, the component works correctly;
- the BV magnitude is indicated by the label m_{BV} = 1, 2 or 3 if the component stays stuck in closed nominal position, stuck open at 50% of nominal flow area and stuck totally open, respectively; if m_{BV} = 0, the component does not fail;
- for each SV, the magnitude is indicated by the label m_{SV} =1, 2 or 3 if the component stays stuck closed in nominal position, stuck open at 50% of nominal flow area and stuck totally open, respectively; if m_{SV} = 0, the component works.

As an example, the accidental sequence vector [4, 6, 5, 0, 0, NaN, 1, 2, 1, 2, 4, 3, 2, 5, 4, 1, 3, 2] represents a scenario where the components states are as follows: the CP fails completely to 0% of the nominal value at a time in [501,600] s (fifth event occurring along the sequence); the CV1 correctly works throughout T_M; the CV2 fails stuck open at the nominal position at a time in [101,200] s (first event occurring along the sequence); the BV fails stuck open at 50% of the nominal flow area (third event along the sequence) at a time in [301,400] s; the SV1 fails stuck open at 50% of the flow area at a time in [401, 500] s (fourth event along the sequence); finally, the SV2 fails stuck closed in nominal position at a time in [201,300] s (second event along the sequence).

Finally, it is worth mentioning that other methods can be used to efficiently generate accident scenarios for the analysis of complex systems, containing a large number of components. In those cases, *intelligent* sampling techniques could be adopted to *preferentially* guide the exploration of the (large) system state-space towards the critical regions of interest (i.e., abnormal scenarios), making the best use of the information and knowledge gained at previous steps and iterations of the search (see the *adaptive* framework proposed by (Turati et al., 2017) as an example).

3.2 *Post-processing of accidental transients*

To mine information from the scenarios generated, Spectral Clustering (described in Section 3.2.1) is applied. A comparison with other method, ESAX (described in Section 3.2.2), is also performed.

3.2.1 *Spectral Clustering*

Spectral Clustering (Strang and Nguyen, 1996) allows classifying *N* different objects into *C* clusters by calculating a similarity measure μ between them. The similarity measure is case dependent: in the present application, it is based on the functional characteristics of the *Z* =3 (i.e., p_{cin}, ΔV and T_{hs})-dimensional transients of duration *L*. In line with (Zio et al., 2010), $\bar{\bar{X}}^k$ (*k*=1, …, Z) [*N*, *L*] are normalized in the range [0.2, 0.8] to obtain $\bar{\bar{Y}}^k$ [*N*, *L*]. For each *i-th* and *j-th* generic transient (with *j* and *i* = 1,…, *N*) we compute the pointwise Euclidean distance δ_{ij} (Zio et al., 2010):

$$\delta_{ij}=\sum_{k=1}^{Z}\sum_{l=1}^{L}\left| y_{il}^{k}-y_{jl}^{k}\right| \tag{1}$$

The pointwise distance δ_{ij} is mapped into an "approximately zero" fuzzy set (FS), to get a gradual transition of the similarity measure (Joentgen et al., 1999). In this study, the bell-shaped FS is used (Dubois et al., 1988):

$$\mu_{ij} = e^{-\left(-\frac{\ln(\alpha\prime)}{\beta^2}\delta_{ij}^2\right)} \tag{2}$$

where the similarity measure μ_{ij} between each *i-th* and *j-th* generic transients (with *j*, *i* = 1,…, *N*) is close to 0 when the evolution of the *i*-th and *j*-th transients are very different, whereas μ_{ij} is close to 1 when they are similar. The parameters α' and β are arbitrary and are set by the analyst: the larger the ratio $-\frac{\ln(\alpha\prime)}{\beta^2}$ (in our case, equal to 1.9×10^{-8}) the stronger the definition of the similarity (Zio et al., 2010).

By doing so, a similarity matrix $\bar{\bar{W}}$ of size [*N*, *N*] is defined, whose generic element μ_{ij} represents the similarity measure between the *i-th* and *j-th* trajectory. Then, the diagonal matrix $\bar{\bar{D}}$ is computed, whose single elements $d_1, d_2, \dots, d_N$ are:

$$d_i = \sum_{j=1}^{N} \mu_{ij} \ , i = 1, \dots, N \tag{3}$$

Defining $\bar{\bar{L}} = \bar{\bar{D}} - \bar{\bar{W}}$ and being $\bar{\bar{I}}$ the identity matrix of size [*N*, *N*], the normalized Laplacian matrix $\bar{\bar{L}}_{sym}$ can be defined as:

$$\bar{\bar{L}}_{sym} = \bar{\bar{D}}^{-1/2}\overline{\overline{LD}}^{-1/2} = \bar{\bar{I}} - \bar{\bar{D}}^{-1/2}\overline{\overline{WD}}^{-1/2} \quad (4)$$

To find the optimal number of clusters *C* in which the *N* transients will be clustered, the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ and the corresponding eigenvector $\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_N$ of the matrix $\bar{\bar{L}}_{sym}$ are computed. According to the eigengap heuristic theory (Mohar, 1997), *C* is set equal to the number of eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_C$ that are much smaller than λ_{C+1} (Von Luxburg, 2007). This guarantees that the similarity of the *i-th* trajectory with other trajectories is well captured in the *C*-dimensional vector $\bar{u}_i$ corresponding to the *i-th* row of the matrix $\bar{\bar{U}}$. Finally, the matrix $\bar{\bar{T}}$ (*N*, *C*), in which the first *C* eigenvectors $\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_C$ are stored, is fed to the Fuzzy-C-Means (FCM) clustering algorithm to assign each *i-th* transient in the corresponding cluster (Alata et al., 2008). The rows of matrix $\bar{\bar{T}}$ are normalized as shown in Eq. 5, which is proven to enhance clusters identification (Von Luxburg, 2007):

$$t_{ic} = \frac{u_{ic}}{(\sum_{c=1}^{C} u_{ic}^2)^{\frac{1}{2}}} \quad (5)$$

For each *i*-th trajectory (i.e., transient), the FCM algorithm provides its membership m_{ic} (whose values varies between 0 and 1) to all clusters, *c* = 1, 2, ..., *C*. To guarantee that the trajectories are assigned to each cluster without misclassification, the *i-th* trajectory is assigned to cluster *c* if the corresponding membership $m_{ic} > 0.7$. Mathematical details about the FCM algorithm are found in (Bezdek, 1981).

3.2.2 *Extended SAX*

The Extended Symbolic Aggregate approXimation (ESAX) (Lkhagva et al., 2006) is an algorithm that allows symbolically represent the gaussian Z-dimensional time series of data $\bar{\bar{X}}$ into a series of symbols $\bar{\bar{P}}$ (in our case, integer numbers), that will be used for calculating a measure that will be used within the FCM for clustering. The mapping of the time series into symbols can be done reducing the length *L* of $\bar{\bar{X}}$ into *n* intervals, and the range of values of $\bar{\bar{X}}$ into α intervals, as follows:

Step 1: Choose *n* and α (in our case *n* = 4 and α = 19);

Step 2: Build *Z* matrices $\bar{\bar{B}}^k$ (*k*=1, …, *Z*) [*N*, *L*] from $\bar{\bar{X}}$, whose generic element b_{il}^k is:

$$b_{il}^{k} = \frac{x_{il}^{k} - E^{k}}{\sigma^{k}}, i = 1, \dots, N \; and \; l = 1, \dots, L \tag{6}$$

where

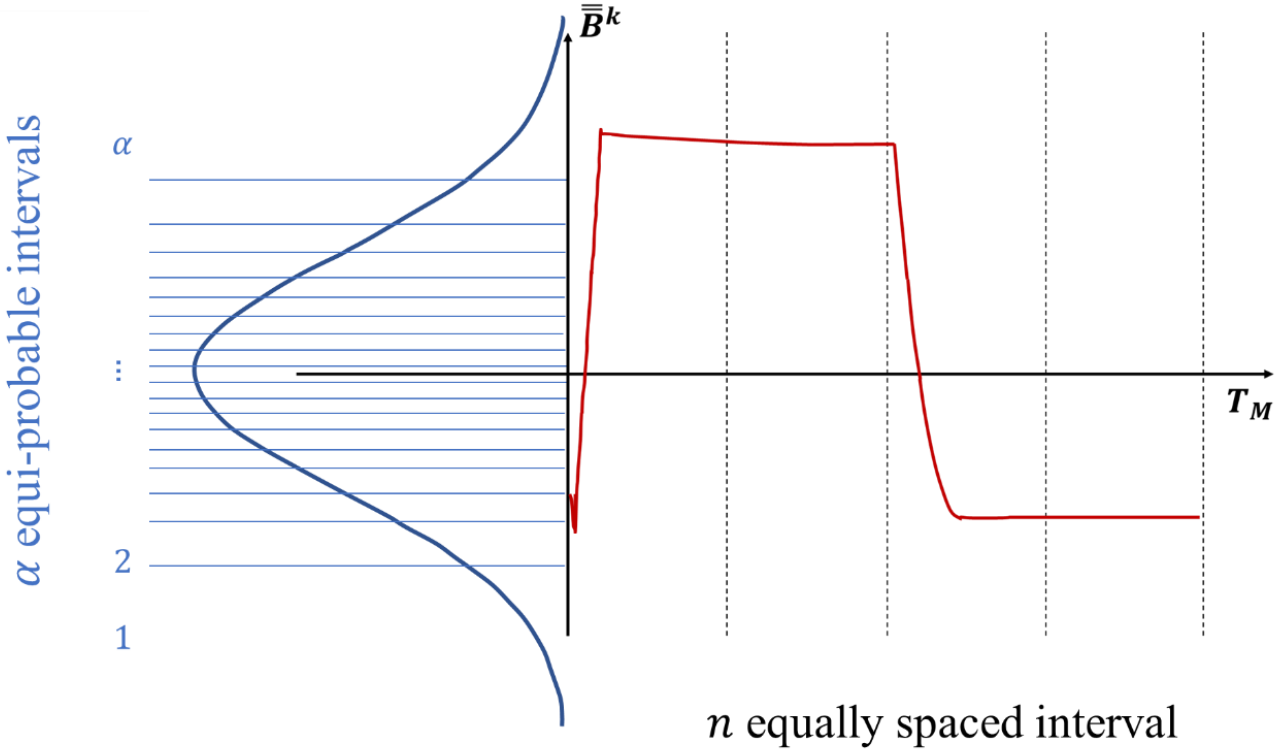
x_{il}^{k} is the generic element of the matrix $\bar{\bar{X}}$

$$E^{k} = \frac{1}{N\,L}\sum_{i=1}^{N}\sum_{l=1}^{L} x_{il}^{k} \tag{7}$$

$$\sigma^{k} = \sqrt{\frac{\sum_{i=1}^{N}\sum_{l=1}^{L}\left(x_{il}^{k} - E^{k}\right)^{2}}{N\,L}} \tag{8}$$

Step 3: Partition the time interval [0, T_M] into *n* equal sized intervals (see Figure 4);

Step 4: Partition the range of values of $\bar{\bar{B}}$ into α equi-probable intervals and assign a symbol to each region (see Figure 4);



$$\bar{D}_{i}^{k} = [19\ 2\ 12\ 19\ 19\ 19\ 19\ 3\ 10\ 3\ 3\ 3]$$

Figure 4 Mapping of a generic i-th transient of the $\bar{\bar{B}}$ (solid line) in the corresponding sequence of symbols $\bar{D}_{i}^{k}$.

Step 5: Build *Z* matrices $\bar{\bar{D}}^{k}$ (*k*=1, ..., *Z*) [*N*, 3 *n*] from $\bar{\bar{B}}^{k}$, whose generic element d_{ip}^{k} is:

$$d_{ip}^{k} = symb\left(\begin{cases} \max\left(\bar{B}_{iq}^{k}\right) & for\ p = 1, 4, \dots, 3n-2 \\ \text{mean}\left(\bar{B}_{iq}^{k}\right) & for\ p = 2, 5, \dots, 3n-1 \\ \min\left(\bar{B}_{iq}^{k}\right) & for\ p = 3, 6, \dots, 3n \end{cases}\right), i = 1, \dots, N \quad (9)$$

where $\bar{B}_{iq}^{k}$ (q=1, …, n) are the values of the *i-th* transient in the *q-th* interval of $\bar{\bar{B}}^{k}$ and *symb( )* is a function that assigns symbols according to the discretization on the Step 4 (also shown in Figure 4);

Step 6: Map $\bar{\bar{B}}$ into a matrix $\bar{\bar{P}}$ [N, 3 n Z] as timely ordered sequence of symbols of the Z matrices $\bar{\bar{D}}^{k}$. For clarity, the generic *i-th* $\bar{\bar{P}}$ in which Z=3 and n=1 is:

$$\bar{P}_i = [\ d_{i1}^{1}\ d_{i1}^{2}\ d_{i1}^{3}\ d_{i2}^{1}\ d_{i2}^{2}\ d_{i2}^{3}\ d_{i3}^{1}\ d_{i3}^{2}\ d_{i3}^{3}\ ] \quad (10)$$

Step 7: Calculate $\bar{\bar{M}}$ [N, N], using the distance measure *MINDIST* (Lin et al., 2003):

$$M_{ij} = MINDIST\left(\bar{P}_i, \bar{P}_j\right) \equiv \sqrt{\frac{L}{n}} \sqrt{\sum_{m=1}^{3\,n\,Z} \left(dist\left(p_{im}, p_{jm}\right)\right)^2} \quad (11)$$

where $\bar{P}_i$ and $\bar{P}_j$ are the *i-th* and *j-th* generic rows of the matrix $\bar{\bar{P}}$, and the $dist( )$ function is a lookup table (Mandelli et al., 2013) that derives from the Euclidean distance.

Step 8: Apply the FCM to $\bar{\bar{M}}$ for classifying the N transients into C clusters: as before, for Spectral Clustering, the *i-th* transient is assigned to the *c-th* cluster if the corresponding membership $m_{ic} > 0.7$.

3.2.3 *Validation of the results*

The goodness of the clusters identified is quantified in terms of their separation and compactness, as measured by the following internal validity indexes (Al-Dahidi et al., 2015):

- Silhouette (Rousseeuw, 1987), that measures the similarity of the data belonging to the same cluster and the dissimilarity to those in the other clusters. The Silhouette index varies in the interval [-1, 1] and should be maximized (Al-Dahidi et al., 2018);
- Davies-Bouldin (DB) (Davies and Bouldin, 1979), that is based on the ratio of within-cluster and between-cluster distances. The DB index ranges in the interval [0, ∞) and should be minimized (Al-Dahidi et al., 2018).

Large Silhouette and small DB values indicate that the obtained clusters are well separated and compacted. In what follows, the two indexes have been calculated to i) validate the choice on the number of the optimal

number of clusters *C* and ii) to compare the performance of the two methods described, i.e. spectral clustering and ESAX.

4 RESULTS

The Spectral Clustering algorithm described in Section 3.2.1 has been applied to classify in C clusters the N scenarios randomly generated. The matrix $\overline{\overline{W}}$ is shown in Figure 5. It contains the degree of similarity μ_{ij} (2) between the $N = 100$ simulated transients: brighter areas mean higher similarity. Figure 6 shows the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_N$ of the matrix $\overline{\overline{L}}_{sym}$ (see Eq. 4) obtained by transforming $\overline{\overline{W}}$ (calculated with the ratio $-\frac{\ln(\alpha)}{\beta^2}$ equal to $1.9 \cdot 10^{-8}$). The first C=7 eigenvalues are remarkably smaller than the eighth eigenvalue: therefore, according to the eigengap heuristic theory (Mohar, 1997), the number of clusters C is set equal to 7 (Baraldi et al., 2013).

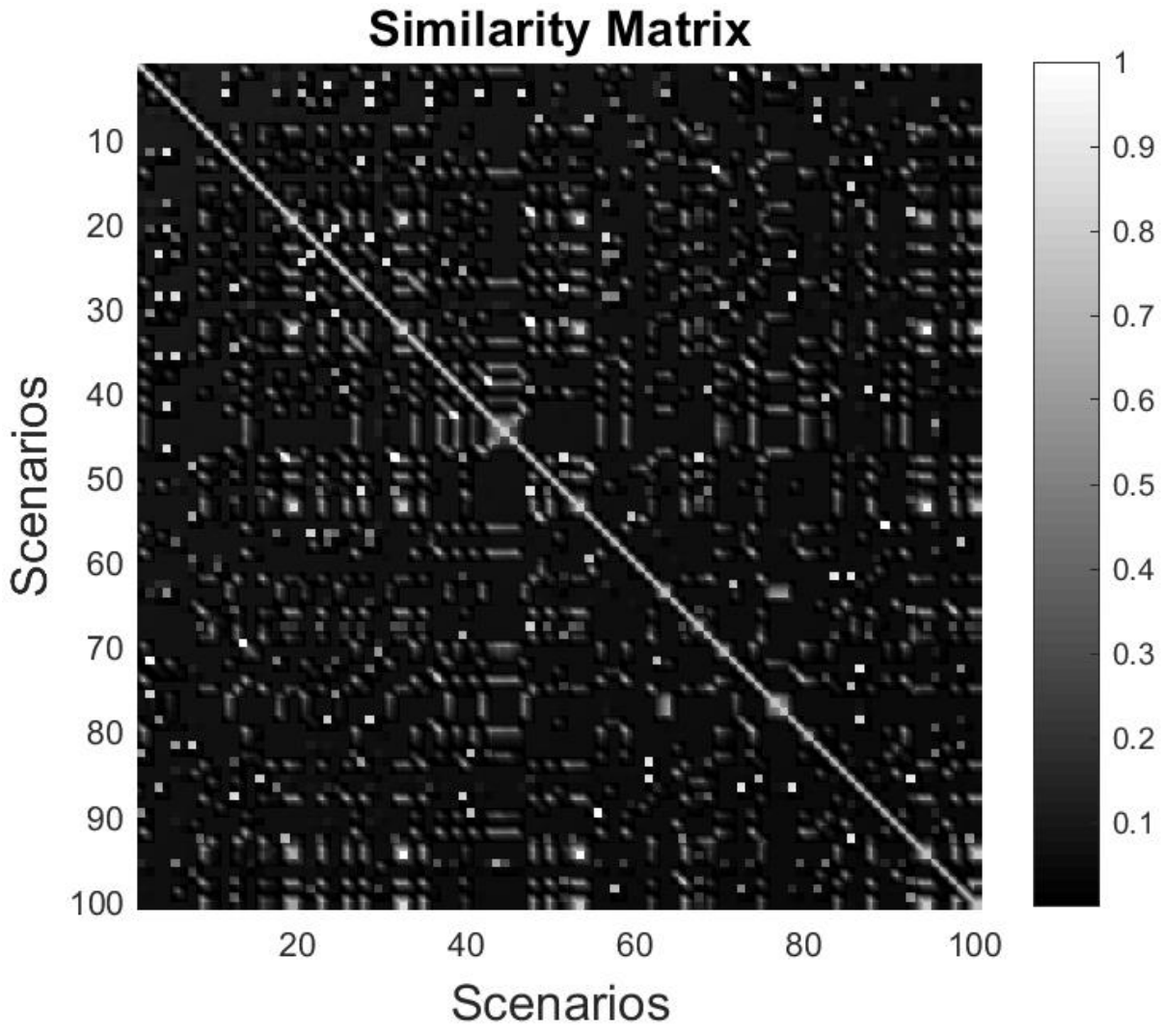


Figure 5 Similarity matrix $\overline{\overline{W}}$ containing the degree of similarity μ_{ij} (2) between the N = 100 simulated transients: brighter areas mean higher similarity.

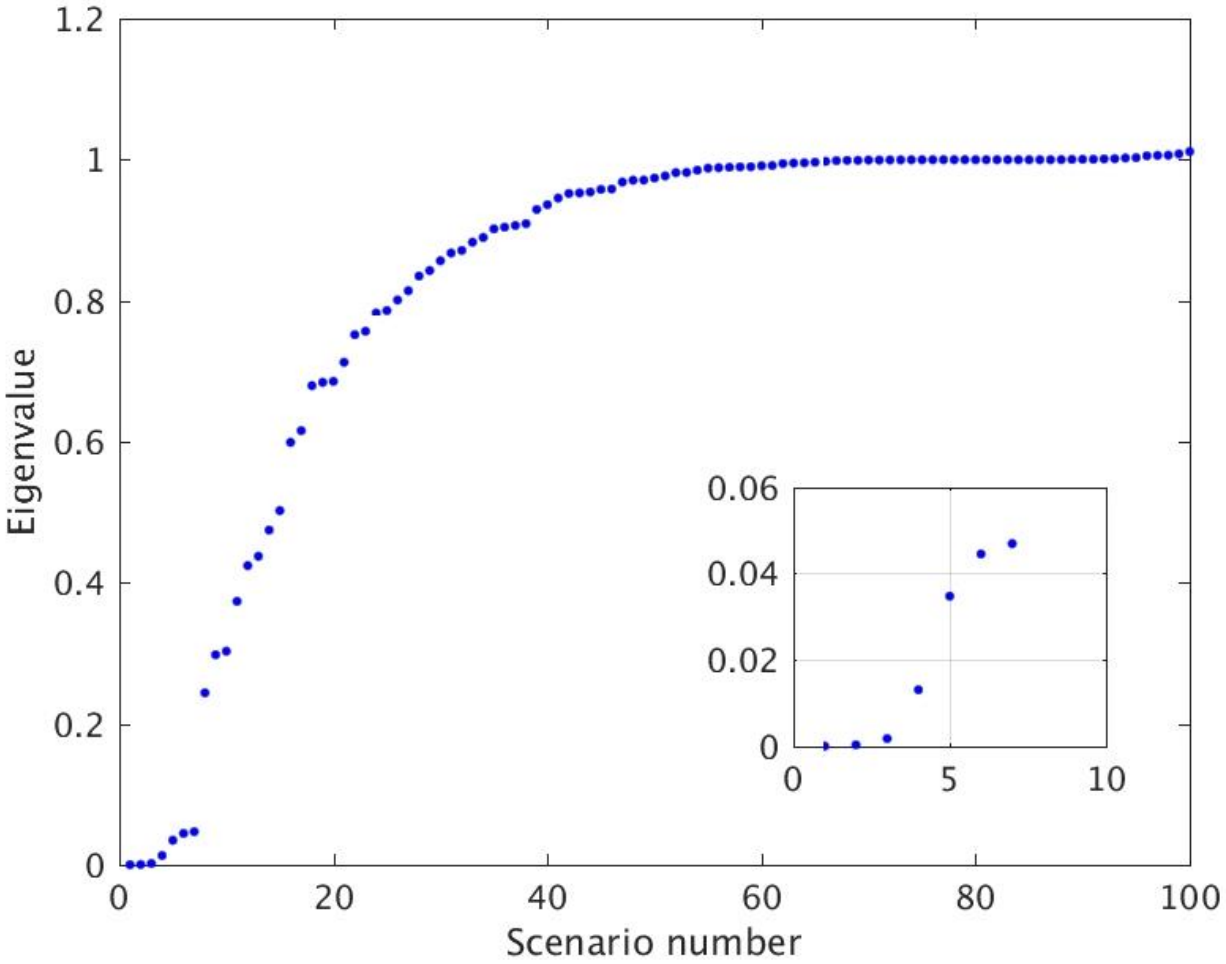


Figure 6 The N = 100 eigenvalues of the matrix $\bar{\bar{L}}_{sym}$: the first C = 7 eigenvalues are very close to zero whereas the eighth is larger: the number of clusters C is, then, set equal to 7.

Once the number of clusters has been found, the assignment of the different trajectories to the most proper cluster can be carried out. The *i-th* transient is assigned to the *c-th* cluster if the membership $m_{ic} > 0.7$. This forces the scenarios belonging to the same cluster to be very similar but, at the same time, some of them may not be assigned to any of the *C* clusters (in our case 2 transients).

In Figure 7, the similarity matrix $\bar{\bar{W}}$ is shown, properly shuffled according to the $C = 7$ clusters obtained using the FCM (i.e., the most similar scenarios are grouped together and highlighted by thick solid lines).

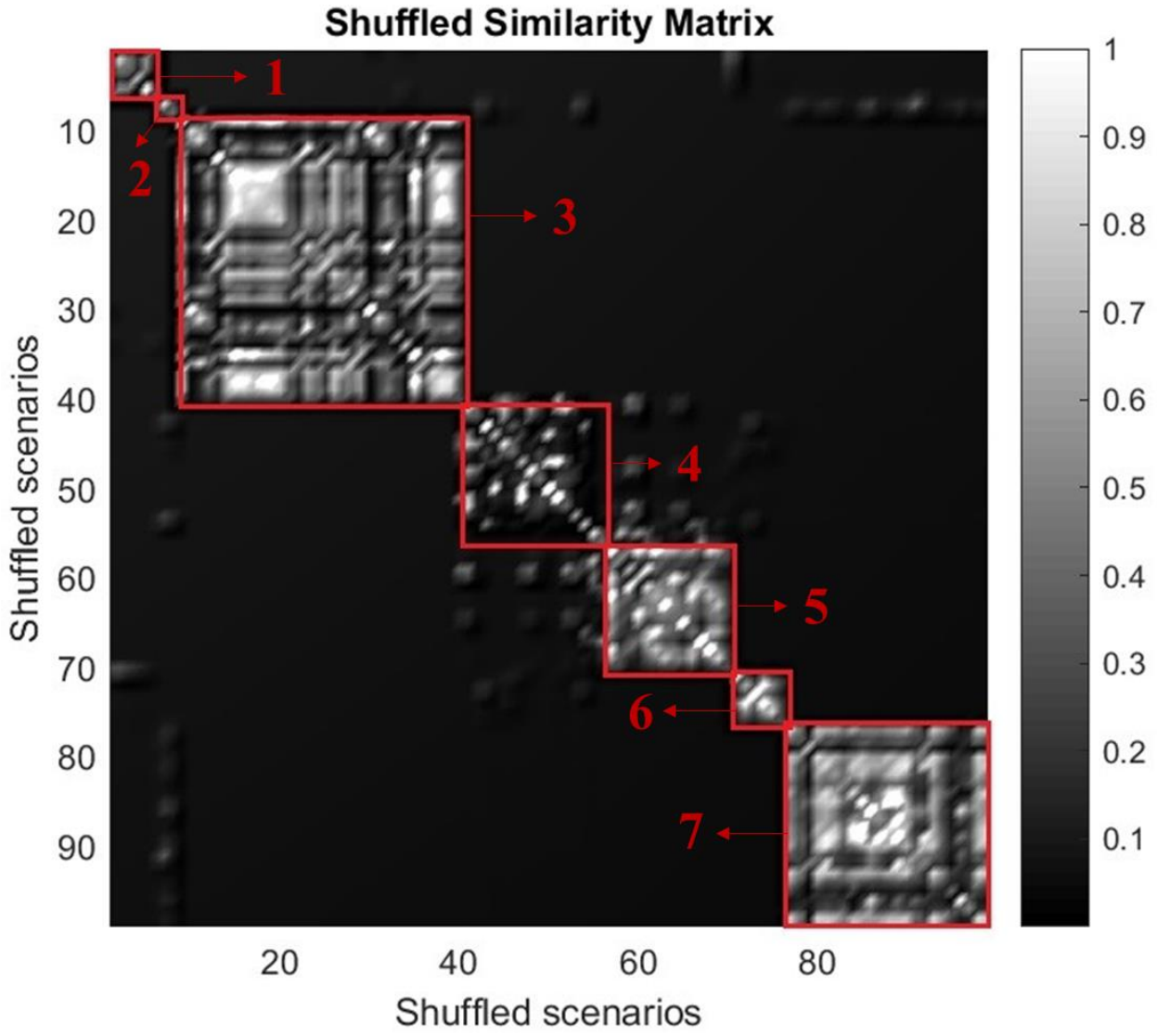


Figure 7 The shuffled similarity matrix $\bar{\bar{W}}$. The C = 7 clusters are represented by the brighter areas along the diagonal of the matrix and highlighted by thick solid lines.

Figure 8, shows the evolution of the safety features (monitored variables) in the different clusters (on the left, ΔV and p_{cin} and on the right T_{hs} and p_{cin}). The following features can be highlighted:

- Cluster 1 is characterized by transients in which LOFA occurs at medium time (i.e., between [201, 300, 400] s), that is jointly characterized by ΔV dropping to zero following an exponential decrease, and p_{cin} increasing during the current ramp.
- In cluster 2, the ΔV initially increases (without relevant safety concerns), then drops to zero, whereas p_{cin} and the T_{hs} are stable along the transient.
- Clusters 3 and 6 are similar to cluster 1, except that the LOFA occurs, respectively, at early and late times (i.e., between [1, 200] s and [401, 600] s, respectively).

- An initial increase of p_{cin} and ΔV can be acknowledged for transients belonging to cluster 4, then followed by a decrease at medium time.
- Cluster 5 groups transients with initial increase of p_{cin} followed by a decrease with the LOFA occurrence at early time.
- Finally, transients of cluster 7 are those for which the LOFA occurrence is avoided despite the failure of components: they show an increase in ΔV and T_{hs}, whereas p_{cin} is constant to its nominal value.

The characteristics of each clusters are listed in Table 2, where the maximum values of p_{cin} and T_{hs} over all the transients of each cluster are also reported.

From the system viewpoint, the results obtained show that *none* of the MVL scenarios is *critical* for the CS module *integrity* (even those leading to a LOFA): in particular, in all the cases considered, the p_{cin} is kept below the safety threshold of 1.8 MPa, i.e., with a positive safety margin and ΔV keeps below the threshold of 0.1 V, which implies T_{hs} does not exceed the current sharing temperature T_{CS} and the magnet does *not* lose its SC properties. This is mainly due to the fact that the CS magnet has been analyzed in "cold mode experimental operation".

Table 2 Summary of the characteristics of the Z = 3 (p_{cin}, T_{hs} and ΔV) safety features for each cluster.

Cluster	Max p_{cin} *[MPa]*	Max T_{hs} *[K]*	ΔV	LOFA (occurrence)
1	0.65	5.5	Nominal value	Yes, at medium time
2	0.5	9	Increases before LOFA	Yes, at late time
3	0.65	5.5	Nominal value	Yes, at early time
4	0.55	12	Increases before LOFA	Yes, at medium time
5	0.65	6	Nominal value	Yes, at early time
6	0.6	6	Nominal value	Yes, at late time
7	0.45	6	Increases (below limit)	No

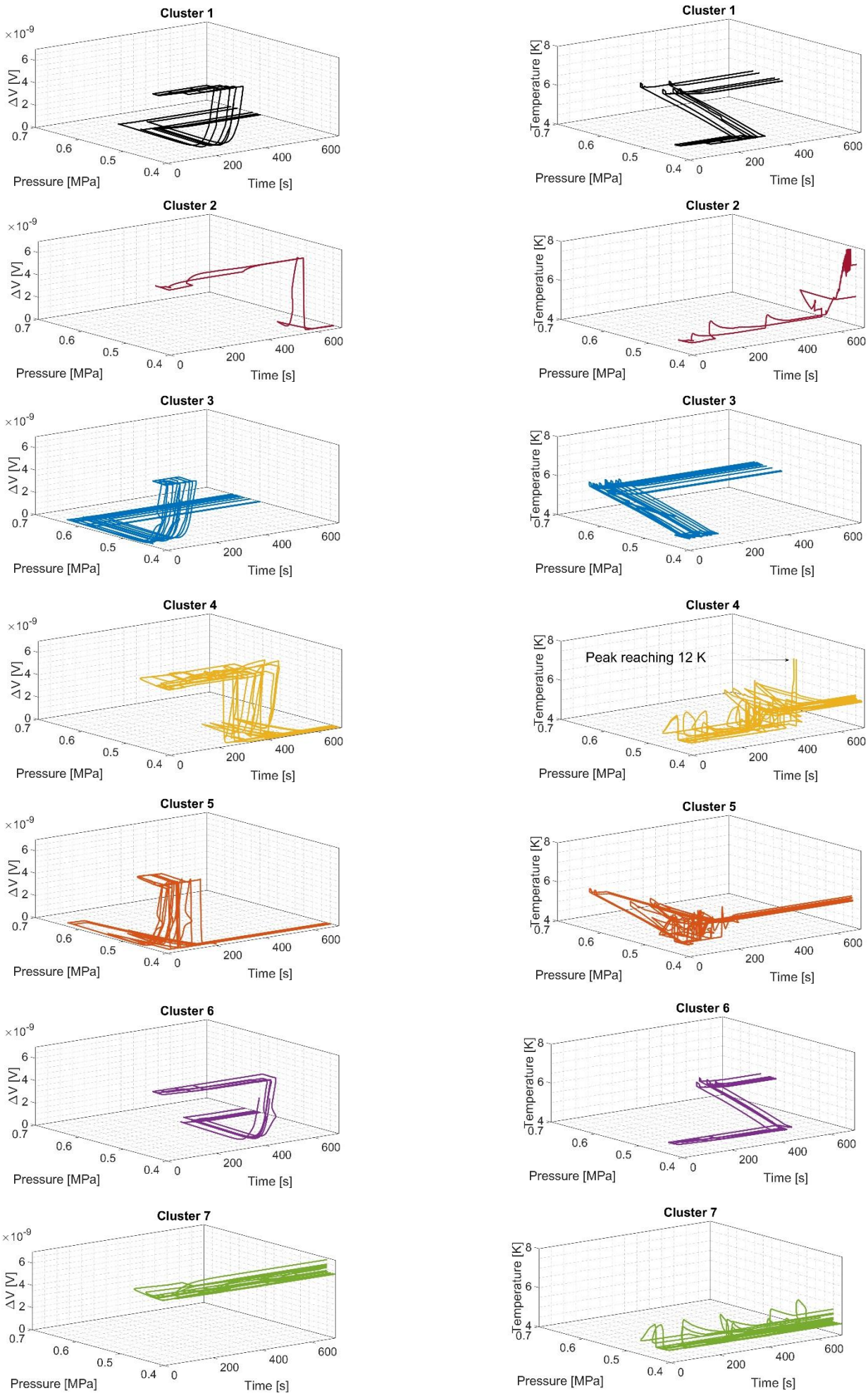


Figure 8 Time evolution of the three safety features (Pressure, Temperature and voltage ΔV) in the 7 clusters identified.

For the sake of clarity and by way of example, in the following one representative transient belonging to cluster 4 (i.e., the one showing the largest values of T_{hs}) is analyzed in detail. The aim is to provide the reader with few physical insights about the effect of (some) components failures on: (i) the time evolution of the system safety features; (ii) the occurrence of the LOFA; and (iii) the intervention of the control and protection systems. The MVL sequence vector that represents the selected scenario is the following: [3, 3, 3, 3, 4, 4, 1, 2, 2, 0, 0, NaN, 0, 0, NaN, 3, 1, 1]. The first failure that happens in the representative transient is the spurious opening of the SV2 (normally closed) at t = 0s. This valve connects the outlet of the CSM to the QT. Since the pressure at the CSM outlet (i.e., 0.376 MPa) is ~0.044 MPa lower than the nominal pressure of the QT (i.e., 0.420 MPa), LHe back-flows from the QT to the circuit as a consequence of the opening of the SV2. This causes an obvious increase in the CSM outlet pressure p_{cout} (not shown), a peak of more than 1 K in T_{hs} (Figure 9 (a)) and a slight increase (~10^{-10} V, indeed negligible) in ΔV (Figure 9 (c)). Notice that, since the CP seems to be working correctly, because it fails later according to the MVL scenario, the nominal pressure drop between the inlet and the outlet of the component is maintained. Therefore, the increase in the outlet coil pressure p_{cout} is followed by an equivalent (~0.044 MPa) increase in the inlet coil pressure p_{cin} (Figure 9 (b)). However, the pressure difference between the QT and the CSM outlet is ~0.044 MPa, so that a small mass of LHe flows from the QT to the circuit, resulting in a limited heating of the SHe, which leads to the stabilization of T_{hs} at the nominal value approximately after 20s. At t = 45s the CV2 fails open, i.e., in nominal position, therefore no variation in the three safety features (p_{cin}, T_{hs} and ΔV) can be seen. The drop in p_{cin} at t = 130s is instead the consequence of the failure of the CP, whose rotational speed decreases exponentially down to 25% of its nominal value. Finally, at t = 295s the CV1 fails closed, thus preventing the SHe from cooling the CSM: consequently, the mass flow rate at the inlet and outlet of the CSM falls below the 10% of its nominal value and the LOFA occurs. This condition is immediately detected and the control and protection systems drive the actions required to protect the integrity of the magnet and of the hydraulic circuit (see Section 2). When the current is removed from the CSM, ΔV exponentially decrease to zero, whereas T_{hs} and p_{cin} increase of ~0.8 K and 0.04 MPa, respectivley, due to the deposition of heat induced by AC losses in the SC cables. When the current is completely removed from the CSM, p_{cin} stabilizes at the QT pressure (0.416 MPa) and T_{hs} reaches a value that is higher than the nominal value, i.e., 5.35 K (due to the heat generated in the coil during the controlled discharge).

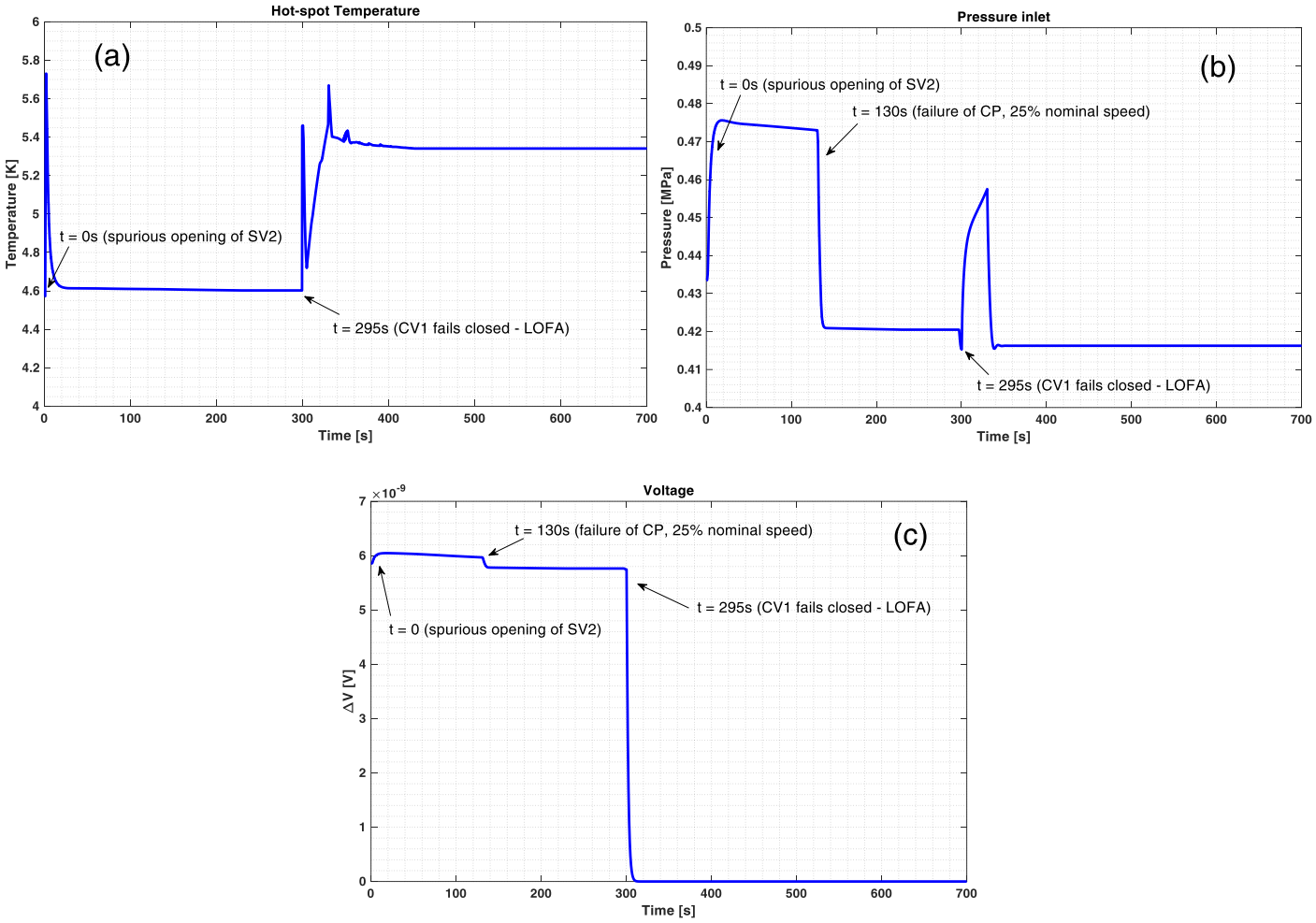


Figure 9 Evolution of the three safety features, i.e., hot-spot temperature T_{hs} (a), pressure at the inlet of the coil p_{cin} (b) and voltage ΔV (c), in one representative transient belonging to Cluster 4.

The phenomena described by the $N = 100$ randomly simulated transients, that have been then grouped in clusters 1-7, are related to the component failures that occur and are described in the corresponding MVL sequences (listed in APPENDIX A). To find the "prototypical sequence" of failure events for each cluster, a bubble plot (Figure 10) is used for visualization of the characteristics shared by the transients in each cluster. The principle underlying the identification of the prototypical sequence with the bubble plot is the following. The bubble dimension depends on the number of times the state of a component appears in the MVL sequences belonging to a cluster. Therefore, the prototypical state of a given component in a given cluster (corresponding to the larger bubble in the plot) is the one that appears most frequently in the MVL sequences of the transients of that cluster. Table 3 presents a summary of the analysis. For example, Cluster 4 shows the largest values of T_{hs}, that reaches 12 K (see Figure 8, 4th row, right, determined by the prototypical failures highlighted in bold in Table 3): for this reason, a more detailed analysis of this cluster is presented in the following. Failures that lead the system to evolve as transients belonging to cluster 4 are those of CV2 typically failing stuck open at early time

and as first along the sequence of events, and of BV typically failing stuck open at early time at 50% of nominal flow area and as second along the sequence. Moreover, SV1 fails more frequently stuck open at 50% of nominal flow area at early time, whereas SV2 is not relevant to characterize the cluster transients, because all its three possible states (i. stuck closed in nominal position, ii. stuck open at 50% of the flow area, and iii. stuck totally open) are equally present in the MVL of the Cluster 4 transients. CV1 typically fails stuck close at 50% of the nominal mass flow rate at early time or stuck totally close at medium time; CP most likely slows down to 25% of the nominal rotational speed (see Figure 10, 2nd row, right). As it can be seen in Figure 10, 2nd row, right, the largest bubbles are associated to BV and CV2, which may be thus considered the most important components of Cluster 4 (they are marked by an asterisk in Table 3).

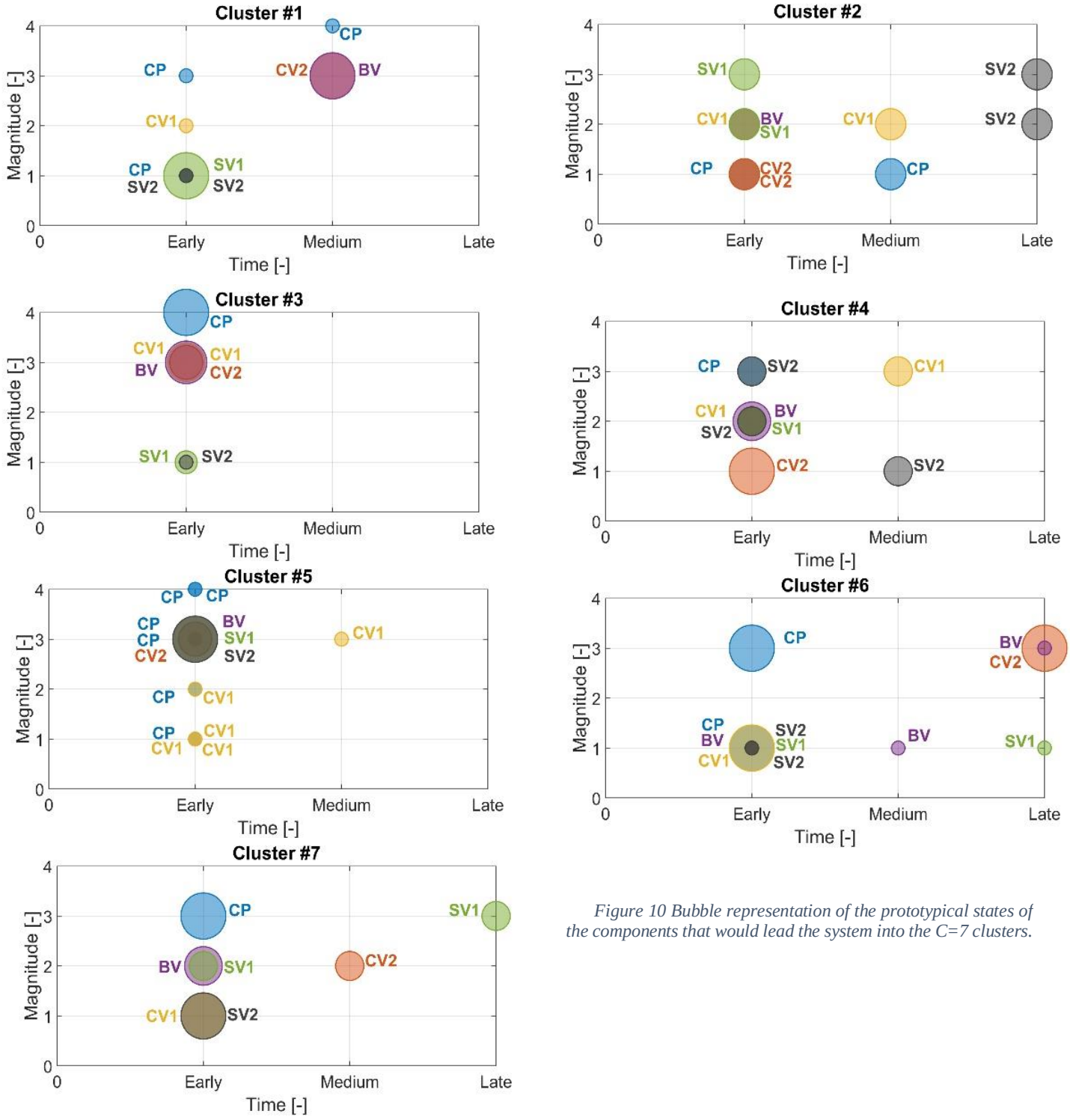


Figure 10 Bubble representation of the prototypical states of the components that would lead the system into the C=7 clusters.

Table 3 Summary of the "prototypical sequences" of components failure events for each cluster.

Cluster	Most important components	Magnitude	Time	Order
1	CV2	3	Medium	3
	BV	3	Medium	3
	SV1	1	Early	1
2	CP	1	Early/medium	-
	SV1	2/3	Early	-
	SV2	2/3	Late	-
3	CP	4	Early	1
	CV1	3	Early	1

	BV	3	Early	1
4	**CV2***	**1**	**Early**	**1**
	BV*	**2**	**Early**	**2**
	CP	**3**	**Early**	**2**
	CV1	**2-3**	**Early/Medium**	**2-3**
	SV1	**2**	**Early**	**1-2**
5	BV	3	Early	2
	SV1	3	Early	1
	SV2	3	Early	1
6	CP	1/3	Early	2
	CV1	1	Early	3
	CV2	3	Late	4
7	CP	3	Early	1
	CV1	1	Early	1
	SV2	1	Early	1

The prototypical accidental sequence identified in Cluster 4 is in accordance with the (prototypical) behavior of the three safety parameters considered: thus, in what follows, a deeper analysis of the states of the components in Cluster 4 is carried out, in order to explain more clearly the physical causes of the safety parameters behaviors. With respect to that, for the sake of representation clarity, in Figure 11 (top, left and right, and bottom), the time evolutions of the three safety features (i.e., pressure, hot-spot temperature and voltage, respectively) in Cluster 4 are also plotted separately.

As already mentioned, the pressure at the inlet of the coil, p_{cin} (Figure 11, top right), does not suffer variations that are significant to the safety of the system (it remains between 0.4 and 0.54 MPa). On the other hand, it should be noticed that it gets close to the pressure of the QT, because at least one of the two SVs is typically open or partially open in Cluster 4. The occurrence of this failure is, frequently, at the beginning of the accidental sequence, as first or second event (see Figure 10, 2nd row, right).

With respect to the high values of the hot-spot temperature, T_{hs} (Figure 11, top left) (e.g., several spikes exceeding 6-7 K and one reaching 12 K), the following considerations can be done. The BV stuck open at 50% of the nominal mass flow rate is the most frequent failure in Cluster 4. As consequence of this failure, a lower quantity of SHe is available to cool down the CSM and, thus, the hot-spot temperature increases. On the other hand, notice that the LOFA cannot occur as the consequence of the sole failure of BV stuck open at 50%. Actually, in Cluster 4, after the BV failure the LOFA typically occurs because *one* component between the two CVs or the CP fails with the highest possible magnitude (i.e., CV1 or CV2 fails stuck completely closed or CP reduces its rotational speed to a complete stop). Other "prototypical" failures in Cluster 4 determine a reduction in the

helium mass flow rate that cools down the CSM, explaining the hot-spot temperature increase. These failures concern the same components that lead to the occurrence of the LOFA and typically occur at medium time in the accident scenario, but with a lower magnitude. In particular, the CVs fail stuck close at 50% of the nominal mass flow rate, whereas the CP reduces its rotational speed from the nominal value to 75%-25%.

Finally, the voltage ΔV at the extremities of each hydraulic channel (Figure 11, bottom) typically presents a slight increase before the occurrence of the LOFA (consistent with the conductor temperature increase) and, then, a sudden decrease when the current is removed from the CSM.

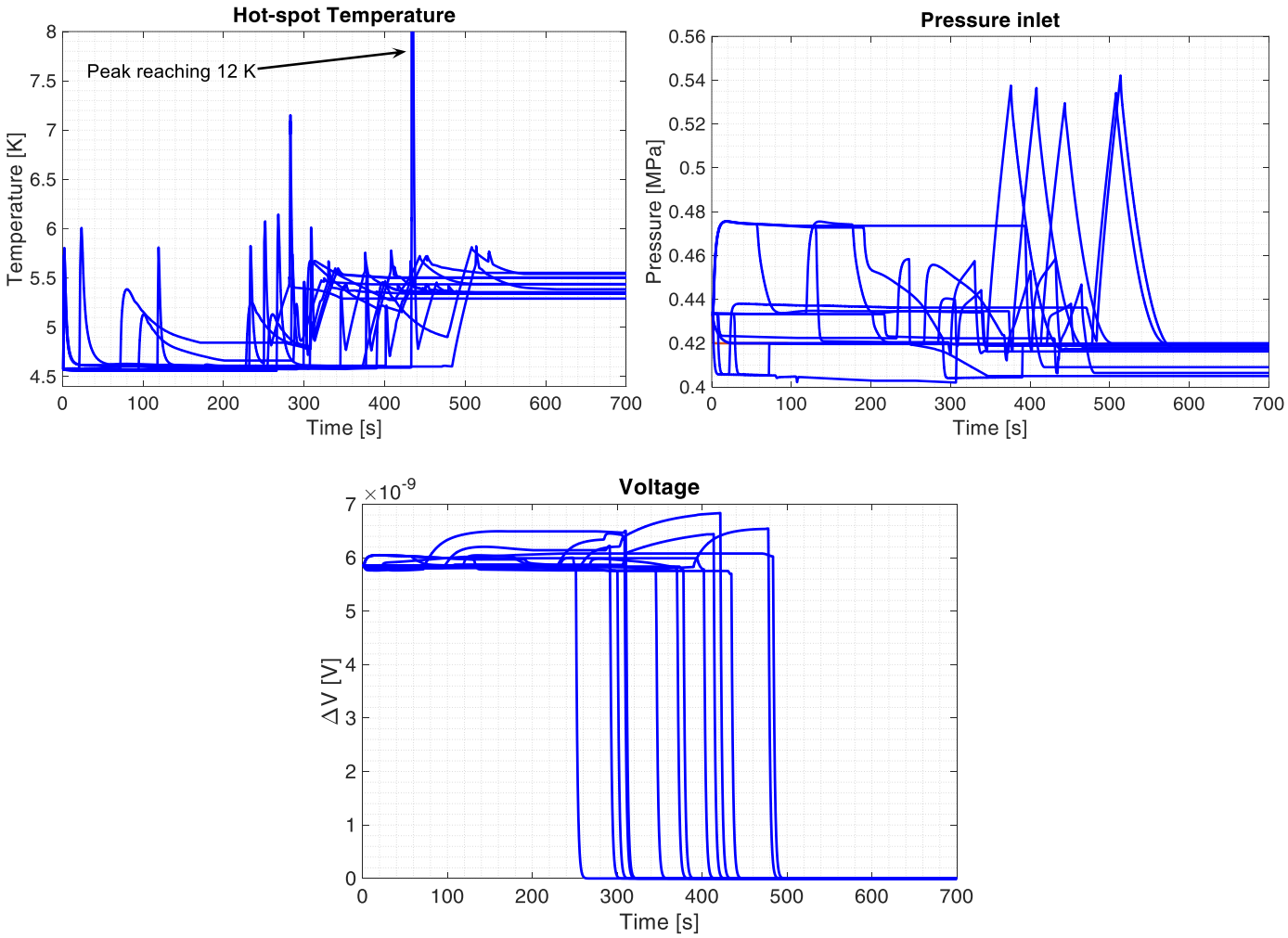


Figure 11 Evolution of the three safety features, i.e., hot-spot temperature T_{hs} (top left), pressure at the inlet of the coil p_{cin} (top right) and voltage ΔV (bottom), in Cluster 4.

In conclusion, the most important outcome of the analysis reported above is the identification of clusters and "prototypical states" of abnormal system behavior (Figure 8 and Table 3). The relevance lies in the fact that such prototypes can serve as a *basis* for: (i) a timely classification of new (developing) scenarios as 'safe' or 'faulty'; (ii) the identification of critical system components that are more likely to lead the system to failure,

and (iii) the consequent prioritization of inspection/maintenance actions on such relevant components (as highlighted, e.g., for CV2 and BV with reference to Cluster 4). However, it has to be acknowledged that for inspection/maintenance prioritization purposes, the analyst should also consider the *likelihood* of scenarios. In this paper, scenarios are (by construction) equally likely, whereas in reality, there can be very large differences between their probabilities/frequencies: this has an obvious impact on components' *importance/criticality ranking* and, thus, on the outcome of the corresponding inspection/maintenance prioritization process.

4.1 *Comparison with ESAX*

The ESAX has been used to group the N available MVL scenarios. The number of clusters C is found by calculating the Silhouette and the DB index for the set of values of the number of clusters $[2, 10]$. Figure 12 shows the Silhouette value (circles) and the DB index (squares): they are maximized and minimized, respectively, for C = 7 (in detail, the corresponding values are 0.95 and 0.40, respectively). Also in this case, each *i-th* transient is assigned to the cluster C if the membership m_{ic}> 0.7 (13 scenarios cannot, then, be assigned to any of the C clusters).

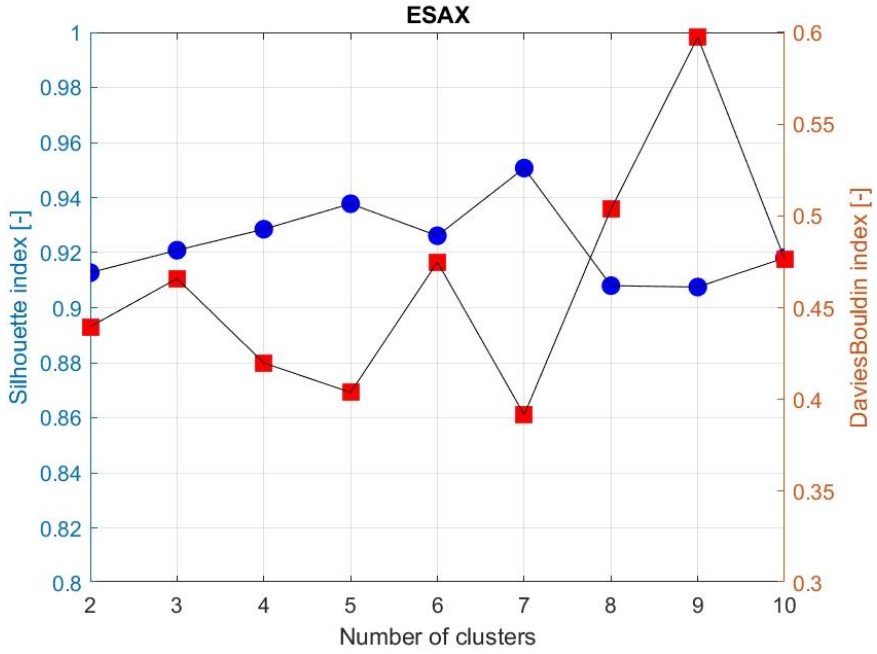


Figure 12 Silhouette (circles) and DB index (squares) values obtained for different number of clusters by ESAX

The results are shown in Figure 13. The clusters obtained are very similar to those of Spectral Clustering. A notable difference can be found in cluster 2: for the ESAX, the LOFA occurs at late time with ΔV close to the

nominal value before the exponential decrease, an initial increase of p_{cin} that then goes to nominal values and no significant variations in T_{hs}; for Spectral Clustering the ΔV initially increases (without relevant safety concerns) then drops to 0 and p_{cin} is stable along the whole transient.

Figure 13 The time evolution of the three safety features in the 7 clusters using ESAX.

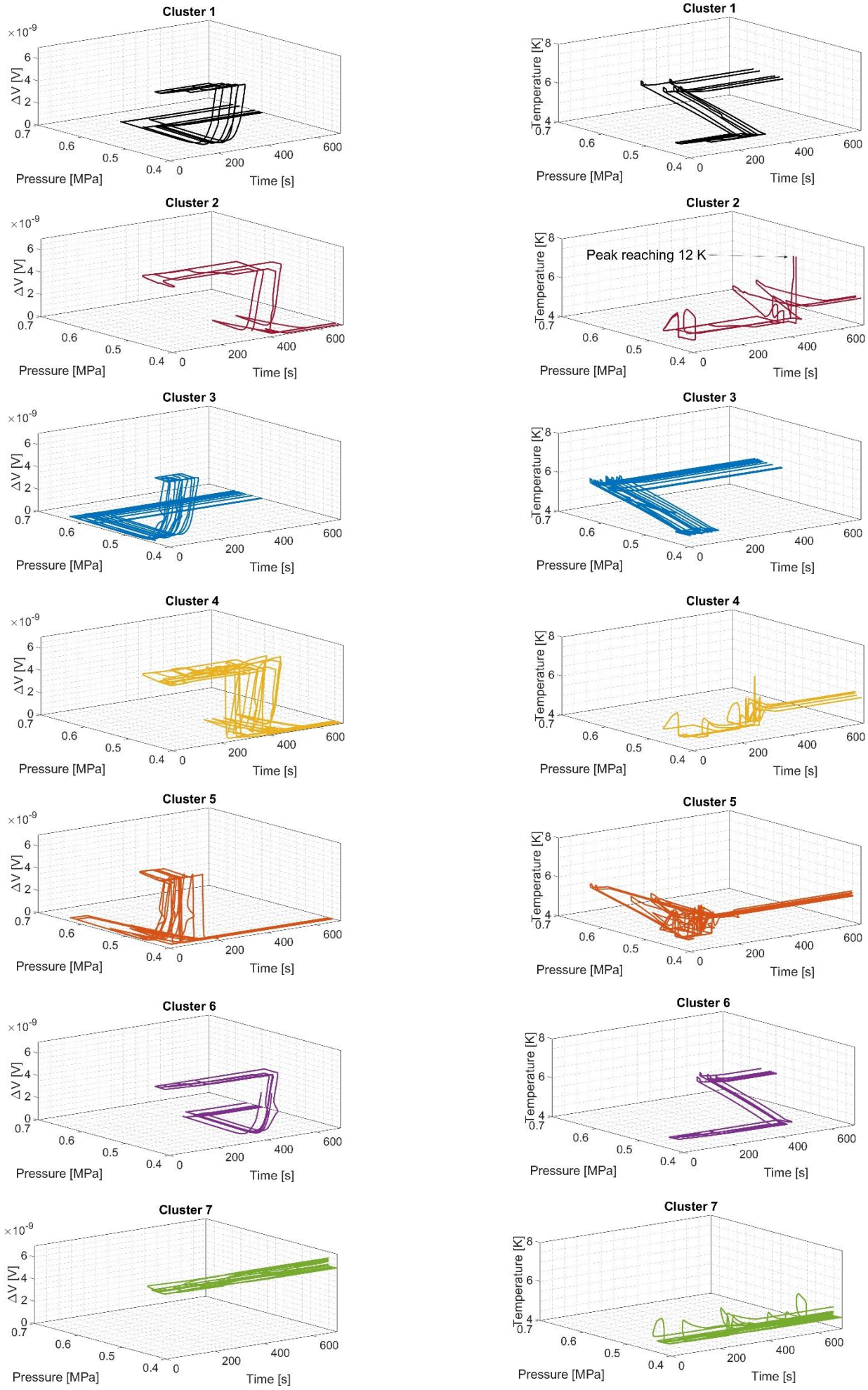
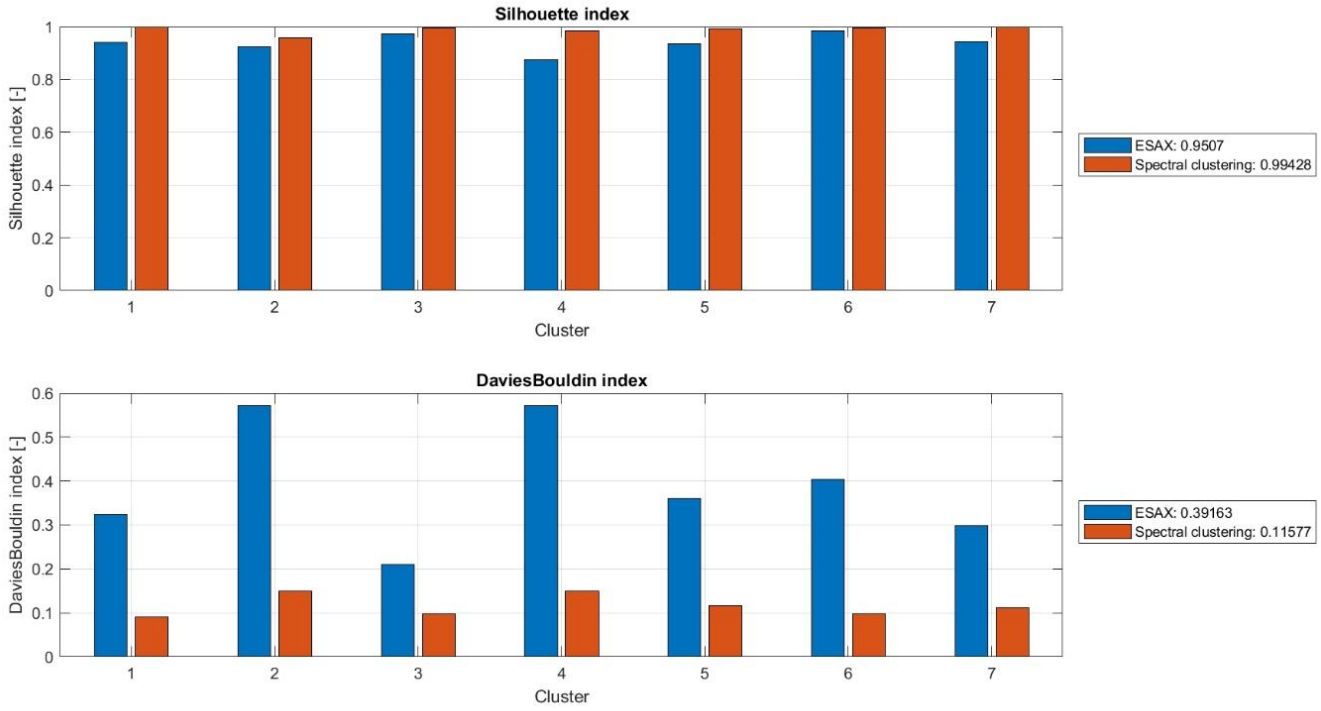


Figure 14 shows the values of the Silhouette and DB indexes for each of the $C = 7$ cluster for Spectral Clustering and ESAX. The best clusters are obtained by Spectral Clustering: for Spectral Clustering, Silhouette values are almost equal to 1 for all clusters, and the DB values are close to 0 (as mentioned above, large Silhouette and small DB values indicate that the obtained clusters are well separated and compacted), whereas for ESAX, Silhouette values are smaller and DB values much larger than those obtained by Spectral Clustering. Moreover, in this latter case, DB values are also far away from 0. This implies higher reliability and "clustering capabilities" of Spectral Clustering with respect to ESAX in the present case study, which suggests its application also in future works.

Figure 14 Silhouette and DB index values for ESAX and Spectral Clustering



5 CONCLUSIONS

In this paper, we have considered the safety analysis of the simplified cooling system of a *single* module of the ITER Central Solenoid in a cold test facility, to identify the component failures that may drive the system into a Loss-Of-Flow Accident.

A *reduced set* of $N = 100$ Multiple Value Logic sequences has been randomly sampled and the corresponding transients post-processed to identify prototypical system behaviors. Spectral Clustering, combined to FCM, has proved capable of dealing with the (unsupervised) identification of prototypical behaviors of abnormal transients. For validation of the results, Spectral Clustering has been compared to ESAX. Silhouette and DB indexes have shown the superiority of Spectral Clustering in this case, because the obtained clusters are more separated and compacted than those of ESAX. These satisfactory results, supported by those previously obtained in other dynamic failure analysis settings (Baraldi et al., 2013), suggest the use of Spectral Clustering and FCM also in future applications to nuclear fusion systems.

From the system point of view, the results have shown that *none* of the MVL scenarios is *critical* for the CS module integrity and superconductive property (even those leading to a LOFA), which is mainly because the CS magnet has been analyzed in "cold mode experimental operation". A more realistic analysis should be done under reduced temperature *margin* circumstances because of an operating temperature closer to the T_{CS}, due e.g. to the AC losses induced during current ramp similar to those experienced in the tokamak operation: such severe operating conditions will be considered in future research.

Another important outcome of the analysis is the identification of clusters and "prototypical states" of abnormal system behavior, which can serve as a *basis* for: (i) a *timely classification* of new (developing) scenarios as 'safe' or 'faulty'; (ii) the identification of *critical system components* that are more likely to lead the system into failure, and (iii) the *prioritization* of *inspection/maintenance actions* on such relevant components. For example, for transients belonging to Cluster 4, it can be suggested to start the inspection/maintenance actions from the analysis of the CV2 and BV.

The results have, thus, confirmed the capability of IDPSA methodologies of capturing the dynamic features of an engineering system in abnormal conditions. In particular, the capability of identifying prototypical behaviors of accident scenarios, verified on nuclear fission systems in previous works, e.g., (Di Maio et al., 2011), has been successfully tested for the first time on a component of a fusion reactor (i.e., a tokamak). The encouraging performances registered could make IDPSA a rather attractive framework for the analysis of nuclear fusion systems (characterized by strongly *dynamic features*) and *possibly* one worth considering for extended adoption in risk assessment applications, provided that the numerous possible accident scenarios and outcomes can be handled computationally in an efficient way.

In this light, future work will be possibly devoted to the construction of *fast-running* Reduced Order Models and/or metamodels, able to reproduce approximately the behavior of the detailed, long-running codes at a *lower computational cost*.

A conclusive consideration is in order with respect to the intended *role* of the proposed approach in the risk assessment of a nuclear system. Indeed, the classification of abnormal conditions should be complemented by: (i) a structured and systematic method (e.g., FMECA) to identify the components' *failure modes* and relevant accident *initiating events*; (ii) an advanced sampling technique for *efficiently* and *exhaustively* exploring the high-dimensional state space; (iii) the calculation of the *probabilities/frequencies* of the generated accident scenarios, for a proper *quantitative evaluation* of the *overall* system *risk.*

ACKNOWLEDGMENTS

The authors would like to thank the three anonymous reviewers for their helpful and constructive comments that greatly contributed to improving the quality of the paper.

6 REFERENCES

Alata, M., Molhim, M., Ramini, A., 2008. Optimizing of Fuzzy C-Means Clustering. World Acad. Sci. Eng. Technol. 2, 224–229.

Al-Dahidi, S., Di Maio, F., Baraldi, P., Zio, E., Seraoui, R., 2015. A Novel Ensemble Clustering for Operational Transients Classification with Application to a Nuclear Power Plant Turbine. Int. J. Progn. Heal. Manag. 6, 1–21.

Al-Dahidi, S., Maio, F. Di, Baraldi, P., Zio, E., Seraoui, R., 2018. A framework for reconciliating data clusters from a fleet of nuclear power plants turbines for fault diagnosis. Appl. Soft Comput. 69, 213–231. https://doi.org/https://doi.org/10.1016/j.asoc.2018.04.044

Aldemir, T., 2013. A survey of dynamic methodologies for probabilistic safety assessment of nuclear power plants. Ann. Nucl. Energy 52: 113–124.

Alpert, C.J., Yao, S., 1995. Spectral Partitioning: The More Eigenvectors, The Better. Discret. Appl. Math 90.

Baraldi, P., Di Maio, F., Zio, E., 2013. Unsupervised Clustering for Fault Diagnosis in Nuclear Power Plant Components. Int. J. Comput. Intell. Syst. 6, 764–777. https://doi.org/10.1080/18756891.2013.804145

Bellaera, R. Bonifetto, R., Di Maio, F., Pedroni, N., Savoldi, L., Zanino, R., Zio, E., "Integrated Deterministic and Probabilistic Safety Assessment of the Cooling Circuit of a Superconducting Magnet for Nuclear Fusion Applications", in: Stein Haugen, Anne Barros, Coen van Gulijk, Trond Kongsvik, Jan Erik Vinnem (Eds.), Safety and Reliability – Safe Societies in a Changing World: Proceedings of ESREL 2018, June 17-21, 2018, Trondheim, Norway, pp. 2161-2168; CRC Press, London, UK, 2018, ISBN 9781351174657.

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms.

Bonifetto, R., Brighenti, A., Isono, T., Martovetsky, N., Kawano, K., Savoldi, L., Zanino, R., 2017. Analysis of the cooldown of the ITER central solenoid model coil and insert coil. Supercond. Sci. Technol., 30: 015015.

Bonifetto, R., Casella, F., Savoldi Richard, L., Zanino, R. 2012. Dynamic modeling of a supercritical helium closed loop with the 4C code. AIP Conference Proceedings (1434): 1743-1750.

Butler, M., Kazakov, D., 2015. SAX Discretization Does Not Guarantee Equiprobable Symbols. IEEE Trans. Knowl. Data Eng. 27, 1162–1166. https://doi.org/10.1109/TKDE.2014.2382882

Davies, D.L., Bouldin, D.W., 1979. A Cluster Separation Measure. IEEE Trans. Pattern Anal. Mach. Intell. PAMI-1, 224–227. https://doi.org/10.1109/TPAMI.1979.4766909

Di Maio, F., Baronchelli, S., Vagnoli, M., Zio, E., 2017. Determination of Prime Implicants by Differential Evolution for the Dynamic Reliability Analysis of Non-Coherent Nuclear Components. Annals of Nuclear Energy, 102, pp. 91–105, 2017.

Di Maio, F., Baronchelli, S., Zio, E., 2015. A computational framework for prime implicants identification in non-coherent dynamic systems. Risk Analysis 35(1): 142–156.

Di Maio, F., Hu, J., Tse, P., Pecht, M., Tsui, K., Zio, E., 2012. Ensemble-approaches for clustering health status of oil sand pumps. Expert Syst. Appl. 39, 4847–4859. https://doi.org/10.1016/j.eswa.2011.10.008

Di Maio, F., Secchi, P., Vantini, S., Zio, E., 2011. Fuzzy C-means clustering of signal functional principal components for post-processing dynamic scenarios of a nuclear power plant digital instrumentation and control system. IEEE Trans. Reliab. 60, 415–425. https://doi.org/10.1109/TR.2011.2134230

Di Maio, F., Vagnoli, M., Zio, E., 2016. Transient Identification by Clustering based on Integrated Deterministic and Prob-abilistic Safety Analysis Outcomes. Annals of Nuclear Energy, Volume 87, pp. 217–227, 2016.

Dubois, D., Prade, H., Testemale, C., 1988. Weighted Fuzzy Pattern Matching. Fuzzy Sets Syst. 28, 313–331. https://doi.org/10.1016/0165-0114(88)90038-3

Galushin, S., Kudinov, P., 2015. Scenario Grouping and Classification Methodology for Postprocessing of Data Generated by Integrated Deterministic-Probabilistic Safety Analysis. Science and Technology of Nuclear Installations, Article ID 278638, 13 pages, http://dx.doi.org/10.1155/2015/278638.

Garibba, S., Guagnini, E., Mussio, P., 1985. Multiple-Valued Logic Trees: Meaning and Prime Implicants. IEEE Transactions On Reliability 34: 463–472.

Grishchenko, D., Galushin, S., Kudinov, P., 2019. Failure domain analysis and uncertainty quantification using surrogate models for steam explosion in a Nordic type BWR. Nuclear Engineering and Design, Volume 343, Pages 63-75.

Hoa, C., Bon-Mardion, M., Bonnay, P., Charvin, P., Cheynel, J.N., Lagier, B., Michel, F., Monteiro, L., Poncet, J.M., Roussel, P., Rousset, B., Vallcorba-Carbonell, R., 2012. Investigations of pulsed heat loads on a forced flow supercritical helium loop – Part A: experimental set up. Cryogenics; 52(7–9): 340–8.

ITER - the way to new energy [Online]. - November 30, 2018. - http://www.iter.org/.

ITER, 2014. Central Interlock System Strategy for ITER Mag-net Protection: Machine Protection Functions. Report ITER_D_K7G8GN v2.1, January 24, 2014.

ITER_D_2NBKXY v1.2, 2009. ITER Design Description Document: Magnets – Conductors, 09/09/2009.

ITER_D_K7G8GN v2.1, 2014. Central Interlock System Strategy for ITER Magnet Protection: Machine Protection Functions, January 24, 2014.

Jankovsky, Z. K., Denman, M. R., Aldemir, T., 2018. Dynamic event tree analysis with the SAS4A/SASSYS-1 safety analysis code. Annals of Nuclear Energy, Volume 115, Pages 55-72.

Joentgen, A., Mikenina, L., Weber, R., Zimmermann, H.-J., 1999. Dynamic fuzzy data analysis based on similarity be-tween functions. Fuzzy Sets Syst. 105, 81–90. https://doi.org/10.1016/S0165-0114(98)00337-6

Karanki, DR, Dang, VN, 2016. Quantification of Dynamic Event Trees–A comparison with event trees for MLOCA scenario. Reliability Engineering & System Safety, 147, pages 19-31.

Karanki, DR, Rahman, S, Dang, VN, Zerkak, O, 2017. Epistemic and aleatory uncertainties in integrated deterministic and probabilistic safety assessment: Tradeoff between accuracy and accident simulations. Reliability Engineering & System Safety, Volume 162, Pages 91-102.

Kirschenbaum, J., Bucci, P., Stovsky, M., Mandelli, D., Aldemir, T., Yau, M., Guarro, S., Ekici, E., Arndt, S.A., 2009. A benchmark system for comparing reliability modeling approaches for digital instrumentation and control systems. Nucl. Technol. 165: 53–95.

Kloos, M., Berner, N., Peschke, J., Scheuer, J., 2018. CHAPTER 3: MCDET: A Tool for Integrated Deterministic Probabilistic Safety Analyses. In: Modern Nuclear Energy Analysis Methods - Advanced Concepts in Nuclear Energy Risk Assessment and Management, pp. 87-131.

Kohonen, T., 1990. The self-organizing map. Proc. IEEE 78, 1464–1480. https://doi.org/10.1109/5.58325

Lee, JH, Yilmaz, A, Denning, R, Aldemir, T, 2018. Use of Dynamic Event Trees and Deep Learning for Real-Time Emergency Planning in Power Plant Operation. Nuclear technology, DOI: 10.1080/00295450.2018.1541394, 8 pages.

Libeyre, P., Cormany, C., Dolgetta, N., Gaxiola, E., Jong, C., Lyraud, C., Reiersen, W., Everitt, D., Martovetsky, N., Rosenblad, P., Cole, M., Freudenberg, K., Sheng Liu, Smith, J., Jing Wei, Lin Wang, Xiaowu Yu, Xiaoyu Dong, Jijun Xin, Chao Li, Wangwang Zheng, Chao Fang 2015. Status of design and manufacturing of the ITER Central Solenoid and Correction Coils. Proceedings of IEEE 26th Symposium on Fusion Engineering (SOFE): 1-8.

Lin, J., Keogh, E., Lonardi, S., Chiu, B., 2003. A symbolic representation of time series, with implications for streaming algorithms. Proc. 8th ACM SIGMOD Work. Res. issues data Min. Knowl. Discov. - DMKD '03 2. https://doi.org/10.1145/882085.882086

Lkhagva, B., Suzuki, Y., Kawagoe, K., 2006. Extended sax: Extension of symbolic aggregate approximation for financial time series data representation. DEWS2006 4A-i8.

Mandelli, D., Smith, C., Yilmaz, A., Aldemir, T., 2013. Mining nuclear transient data through symbolic conversion. Int. Top. Meet. Probabilistic Saf. Assess. Anal. 2013, PSA 2013 3.

Mitchell, N., Bessette, D., Gallix, R., Jong, C., Knaster, J., Libeyre, P., Sborchia, C., Simon, F. 2008. The ITER magnet system. IEEE Trans. Appl. Supercond. 18: 435–440.

Mitchell, N., Devred, A., Libeyre, P., Lim, B., Savary, F. 2012. The ITER magnets: design and construction status. IEEE Trans. Appl. Supercond. 22: 4200809.

Mohar, B., 1997. Some applications of Laplace eigenvalues of graphs, in: G. Hahm and G. Sabidussi (Eds.), Graph Symmetry: Algebraic Methods and Applications. pp. 225–275. https://doi.org/10.1.1.17.467

Perrault, D., 2016. Safety issues to be taken into account in designing future nuclear fusion facilities. Fusion Engineering and Design 109–111: 1733–1738.

Rivas, J.C., Dies, J., Fajarnés, X., 2015. Revisiting the analysis of passive plasma shutdown during an ex-vessel loss of coolant accident in ITER blanket. Fusion Engineering and Design 98–99: 2206–2209.

Rousseeuw, P.J., 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Savoldi Richard, L., Bessette, D., Bonifetto, R. and Zanino, R. 2012. Parametric analysis of the ITER TF fast discharge using the 4C code. IEEE Trans. Appl. Supercond. 22(3): Art. no. 4704104.

Savoldi Richard, L., Casella, F., Fiori, B., Zanino, R. 2010. The 4C Code for the Cryogenic Circuit Conductor and Coil modeing in ITER. Cryogenics (50): 167-176.

Savoldi, L., Bonifetto, R., Pedroni, N., Zanino, R. 2018. Analysis of a protected Loss Of Flow Accident (LOFA) in the IT-ER TF coil cooling circuit. IEEE Transactions on Applied Superconductivity 28(3): 4202009.

Savoldi, L., Bonifetto, R., Zanino, R., 2017. Analysis of a loss-of-flow accident (LOFA) in a tokamak superconducting Toroidal Field Coil. In: Safety and Reliability – Theory and Applications, Proceedings of the ESREL 2017 Conference; 18-22 June 2017; Portoroz, Slovenia; pp. 67-74; ISBN 9781138629370.

Savoldi, L., Zanino, R. 2000. Thermal-hydraulic analysis of TCS measurement in conductor 1A of the ITER central solenoid model coil using the M&M code. Cryogenics (40): 593-604.

Spitzer, J., Stephens, A., Schaubel, K, Smith, J., Norausky, N., Khumthong, K., Gattuso, A. 2015. ITER Central Solenoid Module fabrication program. Proceedings of IEEE 26th Symposium on Fusion Engineering (SOFE): 1-6.

Strang, G., Nguyen, T., 1996. Wavelets and filter banks. Wellesley, MA : Wellesley-Cambridge Press.

Taylor, N., Ciattaglia, S., Boyer, H., Coombs, D., Zhou Jin, X., Liger, K., Mora, J.C., Mazzini, G., Pinna, T., Urbonavicius, E., 2017. Resolving safety issues for a demonstration fusion power plant. Fusion Engineering and Design 124: 1177–1180.

Taylor, N., Cortes, P. 2014. Lessons learnt from ITER safety & licensing for DEMO and future nuclear fusion facilities. Fusion Engineering and Design 89: 1995–2000

Turati, P., Cammi, A., Lorenzi, S., Pedroni, N., Zio, E. 2018. Adaptive simulation for failure identification in the Advanced Lead Fast Reactor European Demonstrator. Progress in Nuclear Energy 103: 176-190.

Turati, P., Pedroni, N., Zio, E. 2017. Simulation-based exploration of high-dimensional system models for identifying un-expected events. Reliability Engineering and System Safe-ty 165: 317-330.

Von Luxburg, U., 2007. A tutorial on spectral clustering. Stat. Comput. 17, 395–416. https://doi.org/10.1007/s11222-007-9033-z

Wu, Y., Chen, Z., Hu, L., Jin, M., Li, Y., Jiang, J., Yu, J., Alejaldre, C., Stevens, E., Kim, K., Maisonnier, D., Kalashnikov, A., Tobita, K., Jackson D., and Perrault, D., 2016. Identification of safety gaps for fusion demonstration reactors. Nature Energy 1: Article number: 16154.

Zanino, R., Bonifetto, R., Heller, R., Savoldi, L., 2011. Validation of the 4C Thermal-Hydraulic Code against 25 kA Safety Discharge in the ITER Toroidal Field Model Coil (TFMC). IEEE Trans. Appl. Supercond. 21: 1948-1952.

Zanino, R., Bonifetto, R., Brighenti, A., Isono, T., Ozeki, H., Savoldi, L., 2018. Prediction, experimental results and analysis of the ITER TF Insert Coil quench propagation tests, using the 4C code. Supercond. Sci. Technol. 31: 035004.

Zanino, R., Bonifetto, R., Hoa, C., Savoldi Richard, L. 2013. 4C modeling of pulsed-load smoothing in the HELIOS facility using a controlled by-pass valve. Cryogenics (57): 31-44.

Zanino, R., De Palo, S., Bottura, L. 1995. A two-fluid code for the thermohydraulic transient analysis of CICC superconducting magnets. Journal of Fusion Energy (14): 25-40.

Zio, E. 2014. Integrated deterministic and probabilistic safety assessment: Concepts, challenges, research directions. Nu-clear Engineering and Design 280: 413-419.

Zio, E., Di Maio, F., 2009. Processing dynamic scenarios from a reliability analysis of a nuclear power plant digital instrumentation and control system. Ann. Nucl. Energy 36(9): 1386-1399.

Zio, E., Di Maio, F., 2010. A data-driven fuzzy approach for predicting the remaining useful life in dynamic failure scenarios of a nuclear power plant. Reliab. Eng. Syst. Saf. 95: 1: 49-57.

Zio, E., Di Maio, F., Stasi, M., 2010. A data-driven approach for predicting failure scenarios in nuclear systems. Ann. Nucl. Energy 37: 482–491.

7 APPENDIX A

For the sake of completeness, all the N = 100 MVL sequence vectors representing the simulated scenarios are reported in the following Table: the failure magnitude (m), time (t) and order (ord) are indicated for all the

components, i.e., CP, CV1, CV2, BV, SV1 and SV2 (see Figure 3). The seventh row (highlighted in bold) corresponds to the MVL sequence taken as reference in Section 3.1.

CP			**CV1**			**CV2**			**BV**			**SV1**			**SV2**		
m	***t***	***ord***	***m***	***t***	***ord***	***m***	***t***	***ord***	***m***	***t***	***ord***	***m***	***t***	***ord***	***m***	***t***	***ord***
0	0	NaN	1	3	3	3	4	4	2	3	2	3	2	1	0	0	NaN
0	0	NaN	0	0	NaN	0	0	NaN	1	5	3	1	5	2	1	4	1
4	4	2	3	6	5	3	4	1	3	6	4	2	5	3	1	6	6
1	6	6	2	3	2	1	2	1	3	5	5	3	3	3	1	5	4
3	2	2	1	5	5	3	4	4	3	2	1	1	6	6	2	2	3
1	6	6	2	4	3	1	6	5	2	4	2	2	5	4	1	2	1
4	**6**	**5**	**0**	**0**	**NaN**	**3**	**2**	**1**	**2**	**4**	**3**	**2**	**6**	**4**	**1**	**3**	**2**
1	2	2	3	4	3	1	1	1	0	0	NaN	2	5	4	0	0	NaN
3	6	5	2	3	3	0	0	NaN	3	1	1	3	4	4	2	3	2
0	0	NaN	1	6	5	1	5	3	3	2	2	1	1	1	2	5	4
2	5	4	2	6	5	3	4	3	2	2	2	1	2	1	0	0	NaN
4	6	4	0	0	NaN	0	0	NaN	2	5	3	3	3	1	3	4	2
3	3	1	3	3	2	0	0	NaN	0	0	NaN	0	0	NaN	2	5	3
0	0	NaN	1	3	2	0	0	NaN	2	1	1	0	0	NaN	0	0	NaN
3	5	3	0	0	NaN	0	0	NaN	1	6	4	2	4	1	1	4	2
3	1	1	3	2	2	2	3	3	3	5	5	0	0	NaN	1	5	4
2	4	5	3	2	1	2	2	2	0	0	NaN	1	4	4	2	3	3
3	6	4	3	5	3	3	5	2	0	0	NaN	0	0	NaN	3	3	1
4	3	3	3	1	1	2	3	4	2	4	5	2	1	2	3	6	6
3	3	2	0	0	NaN	3	2	1	0	0	NaN	3	4	3	0	0	NaN
2	3	3	2	2	1	0	0	NaN	3	4	5	3	3	2	2	3	4
4	5	4	2	5	5	1	1	2	0	0	NaN	2	2	3	2	1	1
4	1	1	2	3	2	0	0	NaN	0	0	NaN	1	5	3	0	0	NaN
3	5	3	2	6	5	1	5	4	2	5	2	2	2	1	3	6	6
1	4	4	1	4	5	3	5	6	1	3	3	3	1	1	3	2	2
0	0	NaN	3	2	2	3	1	1	0	0	NaN	1	6	3	2	6	4
1	3	2	1	6	5	2	6	4	1	2	1	1	6	3	0	0	NaN
4	6	4	0	0	NaN	3	4	1	3	6	3	0	0	NaN	1	6	2
3	2	3	2	1	1	3	6	6	2	2	2	2	4	5	1	4	4
1	3	4	2	1	1	1	1	2	0	0	NaN	2	2	3	2	6	5
3	2	4	0	0	NaN	1	4	5	3	1	2	3	2	3	2	1	1
4	2	1	2	3	2	1	6	4	0	0	NaN	1	6	5	1	5	3
1	6	5	0	0	NaN	1	5	3	3	1	1	1	5	4	2	5	2
1	5	4	1	1	1	2	4	3	1	3	2	3	6	5	0	0	NaN
2	1	2	3	1	1	1	5	5	3	5	6	1	2	3	1	3	4
0	0	NaN	3	4	3	2	1	2	2	4	4	1	5	5	2	1	1
0	0	NaN	1	1	1	2	2	2	1	6	5	3	4	4	1	3	3
1	2	3	1	2	2	3	2	4	0	0	NaN	2	2	1	0	0	NaN
1	6	4	2	1	1	1	3	2	1	5	3	1	6	5	2	6	6
3	4	3	3	3	1	2	3	2	2	5	5	2	4	4	0	0	NaN
1	1	1	1	4	4	0	0	NaN	1	1	3	2	5	5	2	1	2
4	3	3	3	5	5	2	4	4	0	0	NaN	3	3	2	3	2	1
2	4	3	0	0	NaN	1	4	4	2	3	2	0	0	NaN	1	3	1
1	5	3	1	3	2	0	0	NaN	0	0	NaN	2	1	1	0	0	NaN
2	3	2	1	5	4	0	0	NaN	0	0	NaN	3	4	3	1	1	1
3	4	2	0	0	NaN	2	5	4	1	2	1	1	5	3	0	0	NaN
4	4	4	3	5	5	1	1	2	2	6	6	3	1	3	3	1	1
0	0	NaN	1	5	5	2	4	4	3	2	2	1	1	1	3	2	3
0	0	NaN	2	3	2	3	5	4	3	4	3	1	1	1	0	0	NaN

CP			CV1			CV2			BV			SV1			SV2		
m	***t***	***ord***	***m***	***t***	***ord***	***m***	***t***	***ord***	***m***	***t***	***ord***	***m***	***t***	***ord***	***m***	***t***	***ord***
4	1	1	0	0	NaN	1	6	5	1	5	3	1	5	2	3	6	4
1	3	3	2	5	6	1	1	2	3	4	5	2	1	1	3	4	4
2	5	5	3	2	3	2	1	2	3	1	1	0	0	NaN	2	3	4
4	1	1	1	6	6	2	5	5	2	3	2	2	3	3	2	4	4
4	1	1	1	5	5	3	3	3	2	4	4	3	5	6	2	2	2
1	1	1	3	5	6	1	4	4	3	3	3	1	4	5	1	1	2
3	3	1	2	3	2	1	4	3	1	4	4	0	0	NaN	2	5	5
3	3	3	3	4	4	1	2	2	3	5	5	0	0	NaN	3	1	1
4	6	2	3	6	3	0	0	NaN	3	6	4	3	6	5	3	2	1
1	6	3	0	0	NaN	0	0	NaN	1	5	2	1	5	1	1	6	4
3	2	2	3	6	4	3	4	3	2	1	1	3	6	5	0	0	NaN
0	0	NaN	2	4	4	2	1	1	3	3	2	0	0	NaN	3	4	3
2	1	1	1	6	5	2	5	4	3	2	2	2	6	6	1	3	3
2	6	4	0	0	NaN	3	5	2	0	0	NaN	3	4	1	1	6	3
3	4	5	3	3	4	2	2	1	2	3	3	2	3	2	0	0	NaN
2	2	1	1	3	3	2	6	4	0	0	NaN	0	0	NaN	1	3	2
2	3	3	3	3	2	1	5	4	3	1	1	0	0	NaN	1	6	5
1	6	3	0	0	NaN	0	0	NaN	3	4	2	0	0	NaN	1	2	1
4	4	3	1	5	5	3	5	4	1	3	1	1	3	2	0	0	NaN
3	4	6	3	2	2	2	4	5	2	2	1	3	3	4	3	3	3
1	4	4	1	2	1	0	0	NaN	2	3	2	2	4	3	0	0	NaN
3	4	1	0	0	NaN	0	0	NaN	0	0	NaN	2	5	2	1	6	3
4	3	1	0	0	NaN	3	4	3	0	0	NaN	0	0	NaN	2	3	2
2	5	4	2	6	5	1	1	1	0	0	NaN	2	3	2	2	4	3
0	0	NaN	1	3	1	0	0	NaN	0	0	NaN	2	5	2	0	0	NaN
2	2	1	3	2	2	1	5	6	1	3	3	1	5	5	1	5	4
2	6	5	2	2	3	0	0	NaN	3	1	2	2	1	1	2	4	4
0	0	NaN	1	1	1	2	4	4	3	1	3	0	0	NaN	2	1	2
2	4	4	1	4	3	2	6	6	3	1	2	1	4	5	2	1	1
4	4	6	1	3	5	2	3	4	1	1	1	1	2	3	2	2	2
0	0	NaN	0	0	NaN	1	5	1	2	5	2	3	5	3	0	0	NaN
1	3	3	0	0	NaN	1	2	2	0	0	NaN	0	0	NaN	3	1	1
2	6	3	1	2	2	3	6	4	0	0	NaN	0	0	NaN	1	2	1
3	1	1	0	0	NaN	2	4	3	0	0	NaN	3	4	2	0	0	NaN
0	0	NaN	3	2	2	2	5	4	3	2	1	2	4	3	0	0	NaN
1	2	2	1	3	3	1	3	4	3	5	5	1	1	1	0	0	NaN
3	3	3	1	3	1	3	4	4	1	3	2	1	5	5	0	0	NaN
0	0	NaN	1	4	3	1	2	1	2	2	2	3	4	5	2	4	4
0	0	NaN	1	3	2	3	1	1	1	4	5	2	3	4	1	3	3
4	3	3	2	4	4	1	3	2	3	5	5	1	6	6	1	1	1
3	1	1	2	3	2	1	3	3	0	0	NaN	3	5	5	1	4	4
3	2	2	3	5	4	2	1	1	1	4	3	1	5	5	2	6	6
0	0	NaN	0	0	NaN	0	0	NaN	2	5	2	0	0	NaN	1	3	1
3	6	5	1	6	6	3	3	2	2	5	4	2	4	3	1	3	1
3	3	2	3	1	1	0	0	NaN	1	5	3	3	5	5	2	5	4
4	2	1	0	0	NaN	0	0	NaN	1	6	3	1	4	2	0	0	NaN
1	1	2	2	4	5	1	2	4	2	1	1	3	1	3	3	6	6
3	2	2	2	3	3	3	5	4	1	1	1	0	0	NaN	3	6	5
4	3	1	1	5	2	0	0	NaN	1	6	3	0	0	NaN	0	0	NaN
4	5	5	2	1	2	1	1	1	1	2	4	1	6	6	2	1	3
0	0	NaN	0	0	NaN	0	0	NaN	3	1	1	0	0	NaN	2	5	2

Table A.1. List of all the N = 100 MVL sequence vectors representing the scenarios simulated in the present work