Interpretable Anomaly Detection in the LHC Main Dipole Circuits with Non-Negative Matrix Factorization

Christoph Obermair, Andrea Apollonio, Zinour Charifoulline, Lukas Felsberger, Marvin Janitschke, Franz Pernkopf, Emmanuele Ravaioli, Arjan Verweij, Daniel Wollmann, and Mariusz Wozniak

Abstract—CERN's Large Hadron Collider (LHC) with its eight superconducting main dipole circuits has been in operation for over a decade. During this time, relevant operational parameters of the circuits, including circuit current, voltages across magnets and their coils, and current to ground, have been recorded. These data allow for a comprehensive analysis of the circuit characteristics, the interaction between their components, and their variation over time. Such insights are essential to understand the state of health of the circuits and to detect and react to hardware fatigue and degradation at an early stage.

In this work, a systematic approach is presented to better understand the behavior of the main LHC dipole circuits following fast power aborts. Non-negative Matrix Factorization is used to model the recorded frequency spectra as common sub-spectra, by decomposing the recorded data as a linear combination of basis vectors, which are then related to hardware properties. The loss in reconstructing the recorded frequency spectra allows to distinguish between normal and abnormal magnet behavior. In the case of abnormal behavior, the analysis of the sub-spectra properties enables to infer possible hardware issues. Following this approach, five dipole magnets with abnormal behavior were identified, of which one was confirmed to be damaged. As three of the other four identified magnets share similar sub-spectra characteristics, they are also treated as potentially critical. These results are essential for preparing targeted magnet measurements and may lead to preventive replacements.

Index Terms—Large Hadron Collider, Quench Protection, Non-negative Matrix Factorization, Machine Learning

I. Introduction

THE LHC is the world's highest energy particle accelerator, relying on 1232 superconducting main dipole magnets to bend the high-energy particle beams along its circumference. These dipole magnets are powered through eight separate circuits of 154 magnets each. To reach the nominal field of 8.0 T and a current of 11.85 kA, each magnet is cooled down to 1.9 K with superfluid helium. At this temperature, the magnet is superconducting. A resistive

Christoph Obermair is with CERN, CH, 1211 Meyrin, Switzerland, and also with Graz University of Technology, AT, 8010 Graz, Austria (e-mail: christoph.obermair@cern.ch).

Andrea Apollonio, Zinour Charifoulline, Lukas Felsberger, Emmanuele Ravaioli, Arjan Verweij, Daniel Wollmann, Mariusz Wozniak are with CERN, CH, 1211 Meyrin, Switzerland.

Marvin Janitschke is with CERN, CH, 1211 Meyrin, Switzerland, and also with the University of Rostock, DE, 18051 Rostock, Germany.

Franz Pernkopf is with Graz University of Technology, AT, 8010 Graz, Austria.

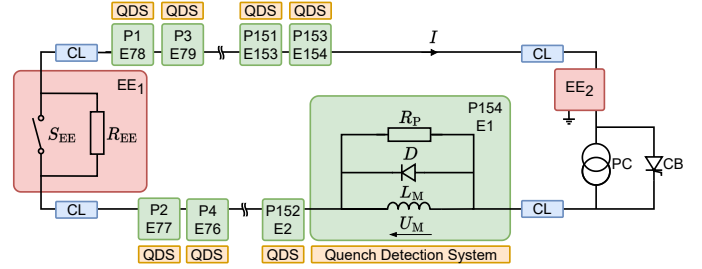The code for this project is publicly available at https://github.com/cobermai/lhc_anomaly_detection

Fig. 1. Schematic view of the main dipole circuit, including the Power Converter (PC), the Crowbar (CB) and the Current Leads (CL). The Quench Detection System (QDS) triggers a Fast Power Abort (FPA), which deactivates the PC and activates the energy extraction systems. Furthermore, it triggers the discharge of the quench heaters (QH) in the respective magnet, if a quench is detected. The two Energy Extraction systems EE_1 and EE_2 consist of a switch S_{EE}, and an Energy Extraction Resistance R_{EE}. The circuit is grounded at the center of the resistor R_{EE} in the EE_2 system. The magnet with inductance L_M and the by-pass Diode D with a Parallel Resistance R_P are in a liquid helium cryostat. Magnets are labeled by their Physical position (P) from the left to the right. The Electrical positions (E) are counted clockwise along the electrical connection starting from the PC. The numbering shown here is representing the circuits in sectors 12, 34, 56, and 78. In sectors 23, 45, 67, and 81 the electrical labels are inverted, as the PC is on the left side of the circuit.

transition in a superconducting magnet, also called quench, results in local heating in the superconducting cables and high voltage transients in the magnet, which can possibly damage the magnet if not appropriately managed. In case of a quench or other powering failures in the circuits, a system of protection elements is in place to safely dissipate the energy in the quenched magnets and extract the remaining energy from the circuit [1]. This process is referred to as a Fast Power Abort (FPA) event. To better understand the data recorded during a FPA event, the LHC main dipole circuits and their protection system are explained in more detail below.

Figure 1 shows a schematic view of a main dipole circuit with its 154 magnets, each represented by a magnet inductance L_M [2]. For this analysis, the magnets are counted along their physical position from the left to the right or clockwise along the electrical connection starting from the Power Converter (PC). In case the PC is switched off, the current I circulating in the circuit by-passes the PC via the Crowbar (CB). The Current Leads (CL) indicate the transition between the cold superconducting part of the circuit and the warm, normal conducting part of the circuit.

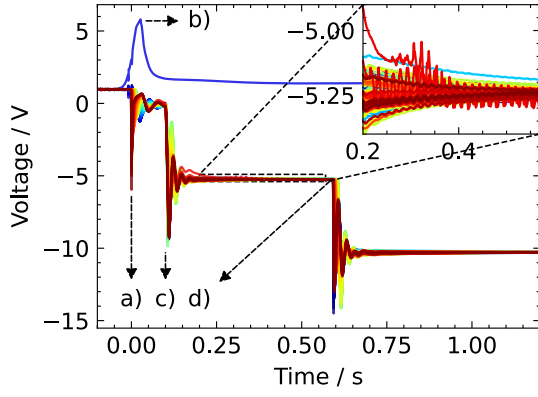The protection system includes a Quench Detection

Fig. 2. Voltages U_M across the 154 main dipole magnets of sector 78 following a quench in the magnet with the electrical position 141 on 31.03.2021 with its different phases: a) the FPA is triggered at 0 s, the QHs for magnet 141 are activated, and the PC is deactivated; b) after around 0.03 s the by-pass diode of the quenched magnet becomes conductive; c) the first Energy Extraction system EE_1 is activated about 0.1 s after the FPA trigger; d) the second Energy Extraction system EE_2 is activated approximately 0.5 s after the first one. The blue curve shows the voltage across the quenched magnet, while the remaining curves represent the voltages across the other 153 magnets of the circuit. The voltage signals of one of the analyzed plateaus are shown in the magnified view on the right top corner.

System (QDS), which detects a quench and triggers the appropriate protection actions [3], [4]. Upon the detection of a quench in a dipole magnet, the PC is switched off and the Quench Heaters (QH) of this magnet are activated. QHs are resistive strips attached to the outer surface of each magnet coil [5]. They ensure protection by distributing the magnet's stored energy more uniformly over the quenched magnet windings [6]. The by-pass Diode D diverts current from the quenched magnet. This restricts the quenching magnet to only absorb its stored magnetic energy, not the energy of the entire circuit. The Parallel Resistance R_P installed across each magnet, smoothens transient voltages during this process [7]. To avoid the circuit's energy to solely discharge in the diode of the quenched magnet, the switches S_{EE} in both Energy Extraction (EE) systems are sequentially activated [8]. They direct the circuit current towards the Resistances R_{EE}, which extracts the circuit's energy within around 300 s. The voltages U_M measured over the 154 magnets during a FPA event are shown in Fig. 2. These signals are voltage transients that contain information about the behavior of the electrical circuit and its components [9].

A quench is a routine occurrence, during training periods aimed at increasing the peak magnetic field in the superconducting magnets [10], and rarely occurs during operation. Once a quench emerges, it is frequently accompanied by secondary quenches. Secondary quenches result from electromagnetic perturbations milliseconds after the initial quench [2] or from thermal propagation in the helium tens of seconds later [11]. Following a quench or a secondary quench, the magnet is exposed to local heating, high voltages, and thermal expansion, depending on the circuit's energy level [12]. The magnets are designed with additional margins for this case, but defects can still occur. Certain hardware failures in the superconducting circuits

can notably impact the availability of the LHC, potentially resulting in months of downtime. The understanding of normal and abnormal circuit dynamics helps to ensure safe quench mitigation and to detect precursors of hardware failures, allowing to schedule preventive maintenance.

To better understand the circuit dynamics, local frequency responses of selected main dipole magnets have been measured and evaluated [13], [14]. FPA events have been deliberately triggered in all main dipole circuits to better understand the voltage transients in the circuits in the absence of a quench [15]. The main dipole circuits have also been extensively studied with electrical simulations with Simulation of Transient Effects in Accelerator Magnets (STEAM) framework [16], [17].

While these simulations account for much of the circuit's behavior and measurements, some aspects - such as the voltage transients observed in the magnified view of Fig. 2 following activation of the EE systems - cannot be captured entirely by simulation. In this time window, secondary quenches frequently occur due to electromagnetic perturbations [2]. A better understanding of the circuit behavior, and in particular the voltage transients after triggering the EE systems, allows the development of mitigation strategies to reduce the number of these electromagnetically induced quenches and the risk associated with them.

The presented research aims to provide insights into the propagation and physical process explaining the observed frequency spectra of the magnet voltage after activation of the EE systems. Normal and abnormal behavior in these frequency spectra is detected and characterized.

The detection of normal and abnormal behavior and their associated physical processes is carried out by Non-negative Matrix Factorization (NMF). The choice of NMF is motivated by recent successful applications of data-driven models to predict quenches [18], to classify QH failures [19], or to model the voltage across magnets [20].

NMF aims at providing interpretable results, as the lack of interpretability is a frequent criticism to other data-driven methods [21], [22]. The method was originally used to decompose pictures of human faces into coherent components like eyes, mouth etc. [23]. The decomposed components are additive and are therefore easy to understand by humans. NMF has been successfully applied to discover molecular patterns in genes [24], to separate different sources of a mixed acoustic signal [25], and to derive properties of galaxies from astronomical observations [26]. In the context of this research, NMF is used to decompose the frequency spectra of the voltages recorded in the LHC's main dipole circuits during FPA events (see Fig. 2) and understand the physical processes causing them. The loss introduced by the frequency decomposition for each FPA event allows detecting and interpreting abnormal behavior in the circuits.

The remainder of the paper is structured as follows. In Section II, an overview of NMF within the context of this study is given. In Section III, the results are presented by showing possible causal relationships between the distinct frequencies in the circuit and the circuit hardware. In addition, five abnormal FPA events are highlighted, and their

characteristic frequencies are interpreted. In Section III-E, the risk of abnormal FPA events for the different hardware components of the machine is elaborated. Finally, Section IV summarizes the results.

II. METHODOLOGY

A. Available Data and Preprocessing

This subsection explains the selection and pre-processing of the measured voltage signals. After the FPA is triggered (point a. in Fig. 2), the first EE system is activated 0.1 s later (Fig. 2 b.). The second EE system is activated after further 0.5 s (Fig. 2 d.). The two periods analyzed in this study are the two voltage plateaus [0.2; 0.575] and [0.7; 1.075] seconds after the triggering of the FPA. These were chosen because the observed frequency spectra in the magnet voltages are not reconstructed by the existing simulation models, which are based on the current knowledge of the circuit's behaviour. Each of these plateaus covers $P = 154$ voltage signals recorded with a sampling rate of 1068 Hz over a length of 0.375 s.

All data used for this study have been recorded after 2017, as the activation times of the EE systems have been kept unchanged since then. In total $Q = 699$ distinct FPA events have been used. These events are split into three categories: 48 events do not contain a quench, 494 events contain a single quench of one magnet in the circuit, and 157 events contain at least one secondary quench due to electromagnetic perturbations. The frequency spectra of the latter events deviate strongly from the others, therefore, only the 48 events without a quench and the 494 events with a single quench are compared to derive anomalies in Subsection II-E. In contrary, the spectral components of all $Q = 699$ events are interpreted in Subsection II-D. Secondary quenches due to thermal propagation in the helium do not affect the frequency spectra in the voltage plateaus, as they occur at a later stage. Events with those secondary quenches are treated like events with a single quench.

The 699 events with voltage signals from 154 magnets for each of the two plateaus after the activation of the EE systems yield a total of $M = 2 \cdot P \cdot Q = 215292$ distinct voltage signals. Each of these $M = 215292$ voltage signals is transformed into a frequency spectrum with $N = 200$ data points via a Fast Fourier Transformation (FFT), an efficient algorithm for computing the discrete Fourier transform [27]. The Nyquist criterion allows showing frequencies of up to 534 Hz [28]. In order to mitigate spectral leakage seven window functions including a window-specific amplitude correction are compared. The window functions are: Rectangular, Hanning, Hamming, Bartlett, Blackman, Flat-top, and Tukey [29]. An exponential trend $\bar{\mathbf{x}}$ of the form

$$\bar{\mathbf{x}} = Ae^{-\mathbf{t}/\tau} + C, \tag{1}$$

is fitted with least squares [30] and subtracted from each individual voltage signal $\mathbf{x}$ with timestamps $\mathbf{t}$. A corresponds to the amplitude of the decay, τ to the decay's time constant, and C to the offset. This exponential trend, which is best visible in the voltage signal with the highest amplitude in the
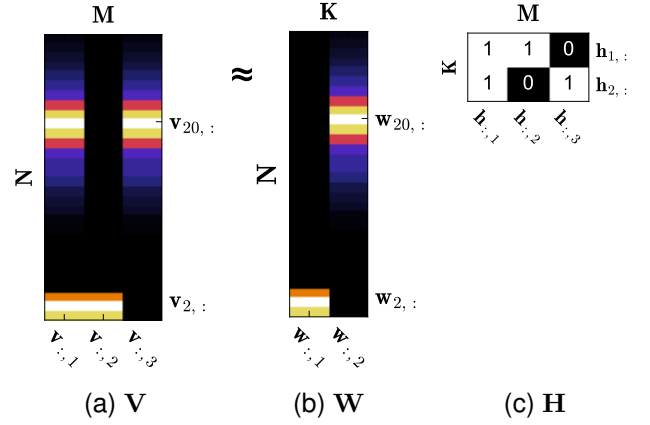

Fig. 3. Example of the NMF decomposition $\mathbf{V} \approx \mathbf{WH}$. In this example, $M = 3$ frequency spectra with $N = 30$ data points are approximated with $K = 2$ spectral components. The color spectrum indicates the voltage amplitudes and ranges from black 0 to white 1. The vertical index count i starts at the bottom.

magnified view of Fig. 2, corresponds to non-linear effects in the magnets [2]. Applying these pre-processing steps yields a dataset composed of M frequency spectra from Q FPA events, with N data points in each frequency spectrum that is processed by NMF.

B. Non-negative Matrix Factorization

Using N and M defined above, let $\mathbf{V}$ be the input matrix, with entries $v_{i,j}$ for $i = 1, .., N$ and $j = 1, .., M$. NMF decomposes the $N \times M$ matrix $\mathbf{V}$ into a product of a $N \times K$ matrix $\mathbf{W}$ and a $K \times M$ matrix $\mathbf{H}$ such that:

$$\mathbf{V} \approx \mathbf{WH}. \tag{2}$$

Here, $\mathbf{W}$ represents the spectral components and $\mathbf{H}$ their weights. The parameter K defines the number of spectral components. All elements $w_{i,k}$ and $h_{k,j}$ of the matrices $\mathbf{W}$ and $\mathbf{H}$ are constrained to be non-negative, leading to the additive nature of the NMF decomposition [23].

Figure 3 illustrates this behavior for a simplified example: the $K = 2$ spectral components can be added to reconstruct the $M = 3$ frequency spectra. For simplicity, the shorthand representation ":" is used to identify all entries from one dimension, e.g. $[\mathbf{W}]_{i=[1,...,N],k=1} = \mathbf{w}_{:,1}$. The two spectral components $\mathbf{w}_{:,1}$ and $\mathbf{w}_{:,2}$, have their maximum of 1 at $i = 2$ and $i = 20$, respectively.

Optimizing $\mathbf{W}$ and $\mathbf{H}$ involves minimizing an element-wise similarity metric $d_*(\cdot)$ between the input $v_{i,j}$ and the reconstructed input $\hat{v}_{i,j} = \sum_k^K w_{i,k}h_{k,j}$. Three widely used similarity metrics are:

1) Squared Eucleadian (Eu) distance [31]:

$$d_{Eu}(v_{i,j}, \hat{v}_{i,j}) = \|v_{i,j} - \hat{v}_{i,j}\|^2 \tag{3}$$

2) Generalized Kullback-Leibler (KL) divergence [32]:

$$d_{KL}(v_{i,j}, \hat{v}_{i,j}) = v_{i,j}\log\frac{v_{i,j}}{\hat{v}_{i,j}} - v_{i,j} + \hat{v}_{i,j} \tag{4}$$
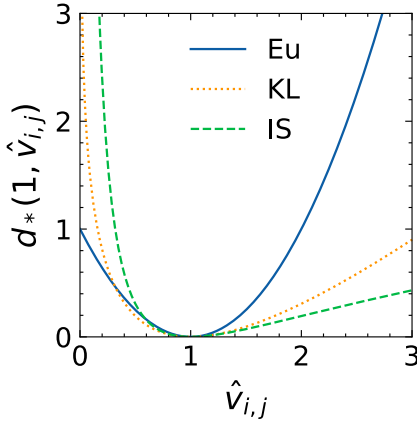
Fig. 4. The Squared Eucleadian (Eu) distance, the Generalized Kullback-Leibler (KL) divergence, and the Itakura-Saito (IS) divergence as a function of $\hat{v}_{i,j}$, assuming $v_{i,j} = 1$.

3) Itakura-Saito (IS) divergence [33]:

$$d_{IS}(v_{i,j}, \hat{v}_{i,j}) = \frac{v_{i,j}}{\hat{v}_{i,j}} - \log \frac{v_{i,j}}{\hat{v}_{i,j}} - 1 \tag{5}$$

Figure 4 shows the three metrics as a function of $\hat{v}_{i,j}$ with $v_{i,j} = 1$. The Eu-distance squares the absolute difference, demonstrated by this example: $d_{Eu}(1, 2) = d_{Eu}(100, 101)$. Figure 4 (Eu) therefore shows a quadratic function, equally sensitive to $\hat{v}_{i,j}$ greater or less than one. The KL-divergence reflects the relative entropy, corresponding to the energy in a system. This causes an increased sensitivity for under-estimation and a decreased sensitivity for over-estimation of the reconstruction $\hat{v}_{i,j}$ [32]. This is reflected in Fig. 4 (KL) by a larger distance metric for $\hat{v}_{i,j} < 1$ as compared to $\hat{v}_{i,j} > 1$. The IS-divergence is scale-invariant [34] as it compares relative differences, illustrated by $d_{IS}(1, 2) = d_{IS}(100, 200)$. The effect of sensitivity observed for the KL-divergence is amplified for the IS-divergence, as is visible in Fig. 4 (IS) [33]. The advantages and weaknesses of these properties will be discussed in Section III-A. Using $d_*(\cdot)$, the reconstruction loss is obtained by $\sum_i^N \sum_j^M d_*(v_{i,j}, \hat{v}_{i,j})$.

All values of $\mathbf{W}$ and $\mathbf{H}$ are initialized with average non-negative double Singular Value Decomposition (SVD) [35]. The term "double" is derived from the use of SVD in approximating both matrices $\mathbf{W}$ and $\mathbf{H}$. This method leads to faster convergence and more robust spectral components compared to random initialization. Any zero values derived by SVD are replaced by the global average of $\mathbf{V}$ as they would otherwise remain at zero during the consecutive multiplicative updates. These multiplicative updates of $w_{i,k}$ and $h_{k,j}$ are specific to each of the three similarity measures [25]:

4) Squared Eucleadian (Eu) distance:

$$w_{i,k} \leftarrow w_{i,k} \frac{\sum_j v_{i,j} h_{k,j}}{\sum_j \hat{v}_{i,j} h_{k,j}}, \quad h_{k,j} \leftarrow h_{k,j} \frac{\sum_i v_{i,j} w_{i,k}}{\sum_i \hat{v}_{i,j} w_{i,k}} \tag{6}$$

5) Generalized Kullback-Leibler (KL) divergence:

$$w_{i,k} \leftarrow w_{i,k} \frac{\sum_j \frac{v_{i,j}}{\hat{v}_{i,j}} h_{k,j}}{\sum_j h_{k,j}}, \quad h_{k,j} \leftarrow h_{k,j} \frac{\sum_i \frac{v_{i,j}}{\hat{v}_{i,j}} w_{i,k}}{\sum_i w_{i,k}} \tag{7}$$

6) Itakura-Saito (IS-divergence) divergence:

$$w_{i,k} \leftarrow w_{i,k} \sqrt{\frac{\sum_j \frac{v_{i,j}}{\hat{v}_{i,j}} \frac{h_{i,j}}{\hat{v}_{i,j}}}{\sum_j \frac{h_{i,j}}{\hat{v}_{i,j}}}}, \quad h_{i,k} \leftarrow h_{i,k} \sqrt{\frac{\sum_i \frac{v_{i,j}}{\hat{v}_{i,j}} \frac{w_{i,j}}{\hat{v}_{i,j}}}{\sum_i \frac{w_{i,j}}{\hat{v}_{i,j}}}}, \tag{8}$$

No NMF regularization [36], [37] is applied to avoid the risk of regularizing spectral components with small amplitudes and to minimize the number of parameters to be optimized that would unavoidably be added with regularization.

C. Spectral Component Identification

In this subsection the methodology to derive the final spectral components $\mathbf{W}$, their number K, and their corresponding weights $\mathbf{H}$ is described. 19 different numbers of spectral components are investigated ($K = 2, ..., 20$). The exact spectral components to which the NMF algorithm converges, also depend on the choice of the three distance measures from Eq. 3-5, and the seven types of distinct window functions. All three parameters are referred to as hyperparameters in the remainder of this paper. In total, 399 possible combinations of parameters exist.

The number of spectral components K determines the resolution of the factorization. Choosing a larger K results in a reduced reconstruction loss. However, in the context of this project, separating the measured frequency spectra into common spectral components aims at representing different physical processes. Hence, more spectral components are only desirable if they can be mapped to separate physical processes. Ideally, one physical process should be represented by one spectral component. To choose K accordingly, an additional performance measure is introduced, based on prior research [39], [40].

This performance measure calculates the mean pairwise Chebyshev distance between column pairs of spectral components and column pairs of their weights. The Chebyshev distance shows the maximum value of the absolute differences between two vectors [38]. For the spectral components, the average Chebyshev distance over all $(K-1)!$ possible pairs of spectral components is used as the performance measure $\bar{d}_{Ch}$. An example to calculate $\bar{d}_{Ch}$ for the spectral components in Fig. 3b is shown below.

	$\mathbf{w}_{:,1}$	$\mathbf{w}_{:,2}$		$\mathbf{w}_{:,1}$	$\mathbf{w}_{:,2}$	$\mathbf{w}_{:,1}^{\text{new}}$
$i = 20$	1	0		0	1	0
$i = 2$	0	1		1	0	1
	$\bar{d}_{Ch} = 1$			$\bar{d}_{Ch}^{\text{new}} = \frac{1+1+0}{3} = \frac{2}{3}$		

Since $\mathbf{W}$ has only two columns, there is one possible column pair of spectral components, resulting in $\bar{d}_{Ch} = \max(|\mathbf{w}_{:,1} - \mathbf{w}_{:,2}|) = 1$. If the columns in Fig. 3b are subtracted, this is evident as their absolute difference is $w_{2,1} = w_{20,2} = 1$. Suppose K is increased by one, and the additional component $\mathbf{w}_{:,1}^{\text{new}}$, happens to be identical to $\mathbf{w}_{:,1}$. In this case $\mathbf{w}_{:,1}^{\text{new}}$ represents the same physical process as $\mathbf{w}_{:,1}$. Consequently, $\max(|\mathbf{w}_{:,1} - \mathbf{w}_{:,1}^{\text{new}}|)$ is zero, leading to a decreased $\bar{d}_{Ch}^{\text{new}}$ on the right side of the example calculation above. This example shows that adding more spectral components K, which are not expected to come from different physical processes, gets penalized by the introduced performance metric.

The performance measure is derived similarly for the spectral component weights $\mathbf{H}$ in Fig. 3c:

	$\mathbf{h}_{:,1}$	$\mathbf{h}_{:,3}$	$\mathbf{h}_{:,3}$		$\mathbf{h}_{:,1}$	$\mathbf{h}_{:,3}$	$\mathbf{h}_{:,3}$
$k=1$	1	1	0		0.5	0.5	0
$k=2$	1	0	1		1	0	1
$k^{\text{new}}=1$	–	–	–		0.5	0.5	0

$$\bar{d}_{Ch} = \frac{1+1+1}{3} = 1 \qquad \bar{d}_{Ch}^{\text{new}} = \frac{1+1+0.5}{3} = \frac{5}{6}$$

For the existing $\mathbf{H}$ in Fig. 3c this is shown on the left side of the example calculation. There are three possible column pairs where the Chebyshev distance of *e.g.* the first column pair is calculated by $\max(|\mathbf{h}_{:,1}-\mathbf{h}_{:,3}|) = 1$. If the spectral component $\mathbf{w}_{:,1}^{\text{new}}$ is added, the spectral component weights are adjusted to obtain the same reconstructions. This is illustrated by the right side of the example calculation above. The Chebyshev distance of $\max(|\mathbf{h}_{:,1} - \mathbf{h}_{:,3}|)$ is reduced to 0.5, affecting the average $\bar{d}_{Ch}^{\text{new}} = \frac{5}{6}$ accordingly. Again, the performance measure indicates that K should not be increased. For computational efficiency, this performance metric is not evaluated for all $(M-1)!$ combinations of $M = 215292$ frequency spectra, but for $M^* = 1000$ randomly chosen frequency spectra. In addition to the performance metric, the final choice of spectral components is based on manual inspection of the identified components. This will be discussed in Subsection III-A.

D. Spectral Component Interpretation

This subsection describes the method of identifying the physical process behind a spectral component. For this purpose, the location of the maximum, the average of the maximum amplitude, and the weight propagation of the spectral components are analyzed and discussed. These allow relating a spectral component to hardware behavior in the LHC main dipole circuits [41].

For each FPA event q, the location of the maximum is defined as the magnet position with the highest of the $P = 154$ weights of a spectral component k. The average weight at this position over a selection of FPA events is defined as the average maximum amplitude. A distinction is made between maxima near the quenched magnet, the PC, or the two EE systems.

From the magnet position at which the maximum occurs, the spectral component can propagate to its physical or electrical neighbors. This results from the fact that physical magnet neighbors can experience electromagnetic coupling due to gaseous helium flow between adjacent cryostats, or instrumentation cables and other equipment being installed in their close vicinity, even if they are not directly electrically connected. The magnets are therefore labeled both in their physical and electrical order (see Fig. 1).

The type and direction of the propagation give insight into the mutual interaction of circuit components during a FPA. A distinction is made between the following propagation types:

- Propagation along the electrical position of the circuit: If the weights $\mathbf{H}$ of a spectral component decrease continuously in both directions with the electrical positions, the spectral component propagates through the electrical wiring.

- Propagation along the physical position of the circuit: If the weights $\mathbf{H}$ decrease with the physical position, the spectral component propagates through the helium or the mechanical connection between the magnets. Propagation in helium is usually limited to three physically close magnets, which are installed together in adjacent cryostats belonging to the same cryogenic cell. It is further possible that the instrumentation wires or power cables of nearby physical neighbors cause interference and generate noise, which also propagates along the physical position.

- Artifact of the QDS measurement unit: Lastly, the propagation can depend on effects in the QDS measurement unit. One QDS measurement unit measures the voltage signals on one to three electrically close magnets and one reference magnet. If the QDS measurement unit's input position affects the spectral component, it is assumed that the voltage signal is generated by the measurement unit's electronics, not present directly at the magnet [3], [4].

In addition to the characteristics of the spectral components, several correlation parameters are considered to interpret the spectral components. The most relevant of these are:

- The magnet manufacturer: Each of the three manufacturers used slightly different materials and fabrication methods to produce the magnets, which results in slightly different magnet behavior.

- The sector number: The circuit layout and hardware component manufacturers vary slightly across each of the eight sectors.

- The amplitude and the ramp rate of the circuit current at the time of the FPA trigger: These one relate to the stored energy in the circuit and affect the voltage amplitude after the triggering of the FPA. The ramp rate dI/dt refers to the change of the current over time.

- The FPA event type: The presence of a magnet quench during the analyzed time period affects the frequency spectra significantly. The activation of the QHs induces new voltages, and the diode opening leads to additional transient voltages visible in the frequency spectra.

In the results section, the listed propagation types and correlation parameters are determined and relationships are established for each spectral component to identify the spectral components' underlying potential physical processes.

E. Anomaly Detection

FPA events are abnormal when a frequency spectrum cannot be reconstructed well with the learned spectral components. For this purpose, for each FPA event $q = 1, ..., Q$ the maximum reconstruction loss $\hat{d}_q$ over all signals in the FPA event is calculated. This reconstruction loss depends on the chosen hyperparameters. Those are the combinations of distance measure $d_*(\cdot)$, the number of spectral components K, and the selected FFT window functions. Hence, anomalies with abnormal behavior are estimated across all hyperparameter combinations.

To make the maximum reconstruction loss $\hat{d}_q$ of different combinations of hyperparameters comparable, the probability

distribution over $\hat{d}_q$ is calculated. The probability distribution over $\hat{d}_q$ for each hyperparameter combination is assumed to be gamma distributed,

$$f(\hat{d}_q; \alpha, \theta) = \frac{1}{\Gamma(\alpha)\theta^\alpha} \hat{d}_q^{\alpha-1} e^{-\frac{\hat{d}}{\theta}}, \qquad (9)$$

where Γ is the Gamma function. The parameters α and θ are derived through a maximum likelihood estimation [42] to fit the distribution of the maximum reconstruction losses over all FPA events for one combination of hyperparameters. For this distribution, a p-value z of an event can be defined by the probability of obtaining a $\hat{d}'_q$ at least as high as the observed $\hat{d}_q$,

$$z = \int_{\hat{d}'_q}^{\infty} f(\hat{d}_q; \alpha, \theta) d\hat{d}_q. \qquad (10)$$

The distribution fit is performed for all possible combinations of hyperparameters. Therefore, each event has as many p-values as there are combinations of hyperparameters. Abnormal events are then considered as those for which the median of all p-values is low. This yields an anomaly identification strategy that is robust to the choice of hyperparameters.

As mentioned earlier, the 48 events without a quench and the 494 events with a single quench are considered for anomaly detection. Another restriction concerns the selection of distance measures. Here, the IS-divergence is not considered for anomaly detection. Anomalies that indicate critical hardware faults are expected to have dominant amplitudes, but the IS-divergence compares relative amplitudes. The IS-divergence is, however, still relevant for identifying components explaining physical processes behind anomalies. Thus, for anomaly detection, only two distance measures are used: The Euclidean distance and KL-divergence. Together with the seven FFT window functions and the 19 different numbers of spectral components, 266 combinations of hyperparameters are used for anomaly detection.

F. Anomaly Interpretation

Anomalies are interpreted by considering the spectral components' weights in the FPA voltages. If the weights of certain spectral components are higher in a FPA event with high reconstruction loss, these spectral components might be associated with the anomaly. Thus, with the knowledge gained in Section II-D, also the physical process underlying the spectral component can be attributed to the anomaly. This is used to check whether the anomaly is pointing at a hardware problem.

III. RESULTS

A. Spectral Component Selection

Contrary to anomaly detection, all 399 combinations of hyperparameters and all 699 FPA events, are used to derive the spectral components for interpretation. Figure 5 shows the resulting mean pairwise Chebyshev distance $\bar{d}_{Ch}$ of the (a) spectral components $\mathbf{W}$ and their (b) corresponding weights $\mathbf{H}$, as a function of the number of spectral components
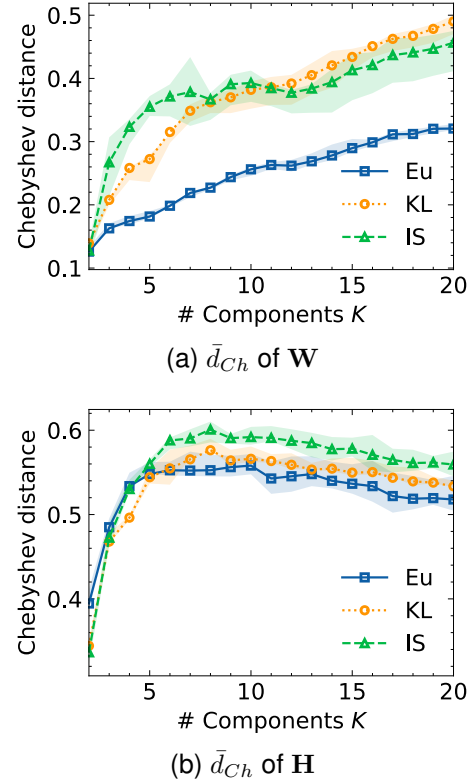

(a) $\bar{d}_{Ch}$ of $\mathbf{W}$

(b) $\bar{d}_{Ch}$ of $\mathbf{H}$

Fig. 5. The mean pairwise Chebyshev distance $\bar{d}_{Ch}$ for (a) the spectral components $\mathbf{W}$ and (b) their weights $\mathbf{H}$ as a function of the number of extracted spectral components K. Compared are the NMF distance measures: The *Squared Euclidean (Eu) distance*, the *Generalized Kullback-Leibler (KL) divergence*, and the *Itakura-Saito (IS) divergence*. The curves indicate the mean over seven different FFT windows, with the first and third quartiles defining the lower and upper confidence intervals, respectively.

K for the three distance measures in Eq. 3-5. The average $\bar{d}_{Ch}$ values across seven distinct FFT windows are shown, with the first and third quartiles forming the lower and upper confidence intervals, respectively. In the range of around seven components, there is a local maximum visible, except for the Eu-distance and KL-divergence in Fig. 5a. Although the $\bar{d}_{Ch}$ curves in Fig. 5a further increase after this extreme point, they decrease in Fig. 5b. The IS-divergence shows the best performance for less than eleven spectral components in Fig. 5a and for more than four spectral components in Fig. 5b. In these regions, the effect of considering relative amplitudes by the IS-divergence is reflected: Also frequencies with small amplitudes are well reconstructed, which increases the diversity of the spectral components and their weights. A similar behaviour is observed for the KL-divergence.

Compared to the number of spectral components or the distance measure, the impact of the FFT window function on $\bar{d}_{Ch}$ is lower, as shown by the relatively narrow confidence intervals in Fig. 5a and 5b. To choose the ideal FFT window function, the FFT window function with the highest $\bar{d}_{Ch}$ at the local maxima of the curves at $K = 7$ is selected. At $K = 7$, the Chebyshev distance is highest if the FFT is calculated using a Hanning window function (not shown in plot). Hence, frequency spectra, derived with a Hanning FFT window function and reconstructed with seven components,

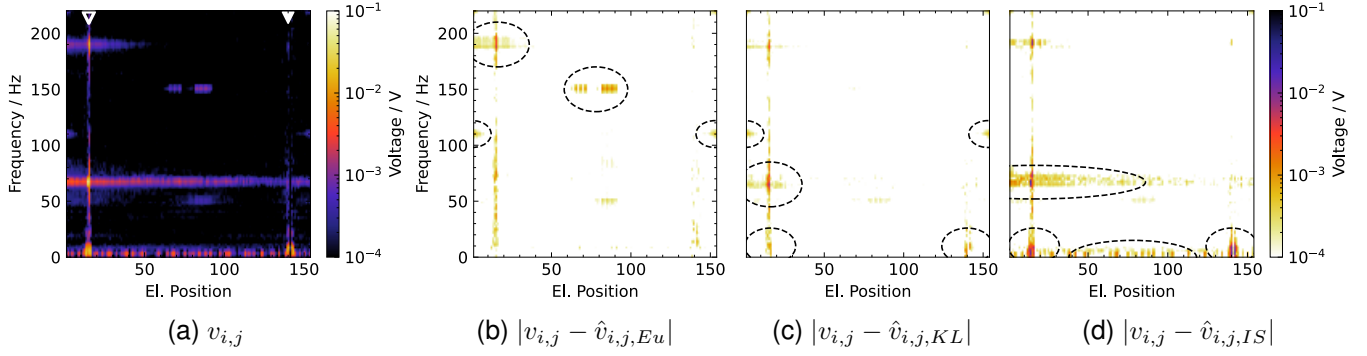| (a) $v_{i,j}$ | (b) $|v_{i,j} - \hat{v}_{i,j,Eu}|$ | (c) $|v_{i,j} - \hat{v}_{i,j,KL}|$ | (d) $|v_{i,j} - \hat{v}_{i,j,IS}|$ |

Fig. 6. Comparison of three different distance measures with a FPM. An FPM shows the frequency and amplitude of the voltage signal as a function of the electrical position. (a) shows a FPA event with a quench in sector 78 on 31.03.2021 with an FPM. The solid white arrow marks the quenched magnet, while the empty arrow marks its physical neighbors. This FPA event was reconstructed with $K = 7$ and (b) the Eu-distance, (c) the KL-divergence, and (d) the IS-divergence. To facilitate visual comparison, the values from the input FPA event in (a) are subtracted from those of the reconstructed events. The absolute values of the subtractions are shown as an FPM in (b-d). Circled areas highlight incomplete reconstructions.

are shown in more detail in Fig. 6.

To illustrate the propagation of frequencies in a FPA event, Frequency-Position Maps (FPMs) are used. An example of such an FPM is shown in Fig. 6a, where the frequencies occurring in the 154 voltage signals, measured [0.2; 0.575] seconds after the triggering of the FPA, are plotted as a function of the electrical position for the FPA event on 31.03.2021 in sector 78. The FPM in Fig. 6a shows the processed and Fourier-transformed voltage signals from Fig. 2 as frequency spectra values $v_{i,j}$. During this event, a quench occurred at the electrical position 141 (white solid arrow in Fig. 6a). The electrical positions 14 and 15 are the physical circuit neighbors of the quenched magnet (white empty arrow in Fig. 6a).

Figures 6b-6d, show the absolute difference $|v_{i,j} - \hat{v}_{i,j,*}|$ between the NMF reconstructions $\hat{v}_{i,j,*}$ and the input $v_{i,j}$ for the different distance measures discussed above. Only frequencies below 220 Hz are included for better visibility. The example aims at comparing the reconstructions for Eu-distance, KL-divergence, and IS-divergence using $K = 7$ and a Hanning FFT window function. The colored spots in the FPMs (Figs. 6b-6d) show electrical positions and frequencies with a reconstruction difference. The darker the color of the points, the larger is the reconstruction difference. If the reconstruction is identical to the input and the reconstruction difference is zero, the plots would be completely white.

In Fig. 6b, small reconstruction differences are visible in the low-frequency range for the Eu-distance. Voltage amplitudes in this range are generally higher. Reconstructions optimized with the Eu-distance, demonstrate superior performance in reconstructing lower frequencies as compared to the two other measures. At 110 Hz, 150 Hz, and 180 Hz significant reconstruction differences are visible by dark spots, highlighted by dashed ellipses. At these frequencies, the amplitudes are lower and are therefore not taken into account by the Eu-distance.

In comparison, for the KL-divergence more significant reconstruction differences are visible by the dark spots in the low-frequency range in Fig. 6c. Instead, the reconstruction differences at 150 Hz and 180 Hz are smaller than for the

Eu-distance measure. This can be explained by the fact that the KL-divergence reflects the relative entropy. Hence, instead of reconstructing the low frequencies with high amplitude more accurately, the entropy is optimized by reconstructing frequencies with lower amplitudes as well.

In Fig. 6d, the IS-divergence leads to reconstructions where both 110 Hz and 150 Hz oscillations have been captured, as there are no significant reconstruction differences visible there. The scale-invariance [34] of the IS-divergence makes it ideal to reproduce also frequencies with low amplitude. However, the low-frequency range, where significant reconstruction differences are visible, has been reconstructed poorly by the IS-divergence.

Based on the interpretation of Fig. 6 and further visual inspections, four Eu-distance components and three IS-divergence components that capture the frequency spectra best were selected for further analysis. Figure 7 shows how the selected spectral components are used to reconstruct the frequency spectra in the voltage signal, measured [0.2; 0.575] seconds after the triggering of the FPA event. Figures 7b-7h show the contribution $\hat{v}_{i,j}$ of each of the seven selected spectral components j to the reconstruction of the frequency spectra $v_{i,j}$ of the FPA event in Fig. 7a. In all FPMs, the frequencies and amplitudes of the voltage signals are shown as a function of the electrical position. The amplitude is displayed logarithmically as a color in the range 10^{-4} V to 10^{x} V, where x is the maximum amplitude of the spectral component in this event. This x is indicated in the caption of each FMP with the spectral component number j. For the reconstruction of the sub-spectra with the different components, the following can be observed: High amplitudes in the low-frequency range, with their maxima at (b) 3 Hz, (c) 6 Hz, (d) 20 Hz, and (e) 66 Hz, are reconstructed by the Eu-distance spectral components. Lower amplitudes in the high-frequency range are reconstructed with the IS-distance spectral components, having their maxima at (f) 150 Hz and (g) 478 Hz. Lastly, a broadband spectrum, spanning vertically over the whole frequency range, can be reconstructed by the IS-distance spectral component (h). Due to the additive nature of NMF the Eu-distance reconstruction in Figs. 7b-7e and the
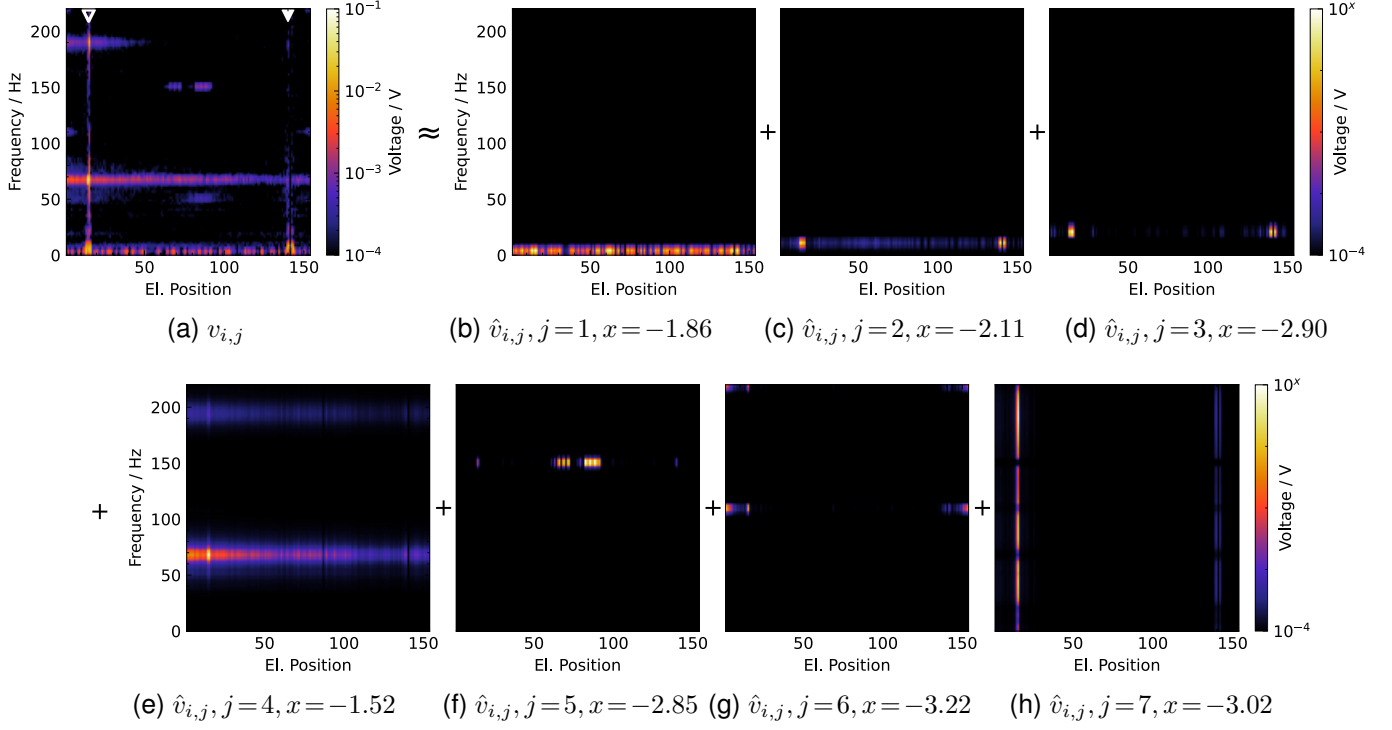
Fig. 7. Frequency and amplitude of the identified seven spectral components as a function of the electrical position. (a) shows the FPM of the frequencies occurring in the voltage signal, measured [0.2; 0.575] seconds after the triggering of the FPA of sector 78 on 31.03.2021. (b-h) show $\hat{v}_{i,j}$ for each spectral component j, used to reconstruct the initial FPM in (a). For better visibility, the maximum of the color axis is scaled with 10^x V. Additionally, the frequency range is restricted to 0-220 Hz, in which the majority of the spectral components occur.

IS-divergence reconstructions in Figs. 7f-7h can be added to reconstruct the input in Fig. 7a. The propagation and physical process of each spectral component are discussed in the next subsection.

B. Spectral Component Interpretation

Seven spectral components have been identified and are considered important to describe the overall frequency response of the LHC's main dipole circuits during FPAs. In the following, their characteristics and the potential underlying physical processes are discussed one by one. A summary of the discussed spectral components is given in Table I, where the columns show the characteristics described in Section II-D.

- Spectral component one (SC1) is visible in Fig. 7b in the bright horizontal frequency band at 3 Hz. There are particularly bright spots at positions 15 and 141 which have a different explanation than the remaining bright spots.
  The magnets at positions 15 and 141 are the physical and electrical neighbors of the quenched magnet, respectively. Considering all FPA events with a quench, the average maximum amplitude at the physical and electrical neighbors of the quenched magnet is 62 mV. In FPA events without a quench, the average maximal amplitude is 5 mV. It can be concluded that the physical process causing SC1 is the quench of a magnet. It can be assumed that the quenching magnet causes electromagnetic perturbations, which are propagating

through instrumentation wires and the connected QDS measurement units. Interestingly, the 3 Hz frequency amplitude is one order of magnitude larger if the quenched magnet was produced by manufacturer one. These perturbations are important because they are most likely the origin of high-energy secondary quenches in neighboring magnets tens of milliseconds after the primary quench occurred.
  The remaining bright spots are introduced in the pre-processing steps. Deviations of the signal x from the exponential trend x̄ (see Eq. 1), are interpreted as oscillation by the FFT. These oscillations are part of SC1 but do not originate from a physical process.
- Spectral component two (SC2) is visible in Fig. 7c by two bright points at 6 Hz at the electrical positions around 15 and 141.
  These are the locations of the physical and electrical neighbors of the quenched magnet. As for SC1, it can be assumed that the physical processes causing SC2 are electromagnetic perturbations induced by the quenched magnet. This assumption is supported by the fact that the average maximum voltage of FPA events is 36 mV, while for events without quench the average voltage is 100 µ V. Similarly to SC1, the frequency amplitudes are one order of magnitude larger if the quenched magnet was produced by manufacturer one.
- Spectral component three (SC3), in Fig. 7d, shows a similar pattern to SC2. Bright spots are visible at the physical and electrical neighbors of the quenched

TABLE I
CHARACTERISTICS OF SPECTRAL COMPONENTS.

Spectral Component	Dominant Frequencies	Location of Maximum	Average Maximum Amplitude	Propagation	Possible Physical Process
SC1 Fig. 7b	3 Hz	Phys. & el. neighbors of quenched magnet	62 mV	Phys. & el. position	Electromagnetic perturbation
		Evenly distributed within circuit	5 mV	Evenly distributed within circuit	Preprocessing
SC2 Fig. 7c	6 Hz	Phys. & el. neighbors of quenched magnet	36 mV	Phys. & el. position	Electromagnetic perturbation
SC3 Fig. 7d	20 Hz	Constant across all magnets	14 mV	Constant across all magnets	Diode induced oscillation
		Phys. & el. neighbors of quenched magnet	4 mV	Phys. & el. position	Electromagnetic perturbation
		Phys. & el. neighbors of power converter	1 mV	Phys. & el. position	Leftover voltage waves traveling along the chain of magnets by magnet impedance
SC4 Fig. 7e	66 Hz 184 Hz 302 Hz	Phys. neighbors of quenched magnet	73 mV	El. position	Oscillations caused by quench
SC5 Fig. 7f	150 Hz	El. neighbors of EE systems	1 mV	Position in QDS measurement unit	Artifact of the QDS measurement unit
SC6 Fig. 7g	107 Hz 220 Hz 260 Hz 370 Hz 478 Hz	Phys. & el. neighbors of PC	690 μV	El. position	Passive hardware elements of PC in sector 78
SC7 Fig. 7h	Broadband spectrum	Phys. & el. neighbors of quenched magnet	4 mV	Phys. & el. position	Quench heater induced oscillation
		Phys. & el. neighbors of quenched magnet	1 mV	Phys. position	Quench dependent oscillations

magnet. However, additional physical processes appear when comparing different events where either no quench occurs, a single quench occurs, or a diode opened due to a secondary quench in the analyzed time window.

In events with an additional diode opening during an EE plateau, the average maximum amplitudes are 14 mV, where the values show little variance. Here, the diode opening induces a 20 Hz oscillation in the circuit, which is constant across all magnets. No diode opened during the FPA event shown in Fig. 7d, which is why the bright spots are not visible across all electrical positions.

In events with a single quench the average maximum amplitudes are highest at the physical and electrical neighbors, with 4 mV. This effect can be seen in Fig. 7d. Similar to SC1 and SC2, the physical process causing SC3 is electromagnetic perturbations originated at by the quenched magnet.

In events with no quenches the amplitude of SC3 is highest at the magnet close to the PC with an average maximum amplitude of 1 mV. Here, the amplitude gradually decreases and is lowest at the first EE system. This can be traced back to the quench-independent leftover voltage waves traveling along the chain of magnets as governed by the magnet impedance. This effect is observed for all events and is proportional to the amplitude and the ramp rate of the circuit current at the moment of the FPA trigger [17]. In Fig. 7d the frequencies caused by the quench are more prominent,

making this process not observable with the given color range.

- Spectral component four (SC4) is visible in Fig. 7e and shows a bright spot at 66 Hz at positions 14 and 15. This shows that the locations of its voltage maxima are in the physical neighbors of the quenched magnet. From there the bright spot is gradually getting darker in both directions, indicating that the oscillation is propagating along the electrical direction.

  A similar pattern is observed at 184 Hz, but in a darker color. In addition, SC4 is high at 302 Hz. The amplitude of this approximate 3rd and 5th harmonic scales indirectly proportional to the number of the nth harmonic.

  While the exact physical process of SC4 remains elusive, it is expected that it is emphasized by a quench. This expectation is supported by comparing FPA events with and without a quench. In events without a quench, the average maximum amplitude is two orders of magnitude lower than for events with a quench (73 mV vs. 730 μV).

- Spectral component five (SC5) appears as a double horizontal band at 150 Hz near the center around the electrical position 77 in Fig. 7f. The bright spots of the band occur at exactly the same input of each QDS measurement unit. This indicates that SC5 is introduced by the electronics of the QDS measurement unit. SC5 only occurs in FPA events in sectors 12, 45, 67, 78, and 81. No specific hardware component has been identified as the cause of this behavior. It occurs in the same

position, regardless of the number of quenches or the quench position, with an average maximum amplitude of 1 mV.

- Spectral component six (SC6) is visible at 107 Hz and 220 Hz as spots originating at the electrical positions 1 and 154 of Fig. 7g. The magnets at these electrical positions are installed close to the power converters of the circuit. The amplitude of SC6 decreases with increasing distance from the power converter. In addition, SC6 has high amplitudes at 260 Hz, 370 Hz, and 478 Hz, not visible in Fig. 7g due to the restricted range of the frequency axis.

  During the EE plateaus the PC is deactivated, indicating that SC6 originates from passive hardware components in the PC. SC6 only occurs in sector 78. No exact hardware component has been identified that could explain this behavior. It always occurs in the same electrical positions, regardless of the quench position, with an average maximum amplitude of 690 µV.

- Spectral component seven (SC7) in Fig. 7h shows one vertical line with high amplitude at the electrical positions 14 and 15, and one line with low amplitude at the electrical positions around 140 and 142. Both lines have interruptions at frequencies already reconstructed by other spectral components. These vertical lines indicate a broadband spectrum in magnets physically close to the quench. In the time domain, this broadband spectrum corresponds to a spike. In a previous analysis of faults in a subsystem of the LHC protection system, such spikes were used as indicators for intermittent short circuits [19]. Hence, SC7 might also be a critical indicator of an intermittent short circuit in the magnet.

  In FPA events with a single quench the average maximum amplitude of this broadband spectrum is 1 mV. For FPA events without quenches, the average maximum amplitude of SC7 is significantly smaller (<100 µV). This shows that SC7 depends on the quench.

  A different physical process is observed during events with secondary quenches, where the average maximum amplitude is 4 mV if the QHs of one of the additionally quenched magnets are activated during the EE plateaus. In this case, SC7 propagates along the physical and electrical position of all magnets in the circuit and is likely induced by the inductive part of the QH strips. No QHs were activated during the FPA event shown in Fig. 7d, which is why this physical process cannot be observed there.

C. Anomaly Detection

In this subsection, the selected anomalies from FPA events with low median p-values following the definitions given in Section II-E are presented. FPA events with low median p-values are labeled based on an LHC specific four-digit identifier of the quenched magnet, marked by a '#'.

Figure 8 shows a boxplot, a conventional method to illustrate statistical data characteristics [43], of the ten FPA events with the lowest median p-value. For each anomalous

FPA event, the quenched magnet is given on the x-axis and a box represents the statistical distribution of p-values over 266 different combinations of hyperparameters (the IS-divergence is not used). The box represents the range between the first and the third quartile, where the line in the middle represents the median. The outer limits further indicate the variability of p-values, which are obtained by subtracting 1.5 times the box length from the first quartile and adding it to the third quartile, respectively. The y-axis is plotted logarithmically, therefore, the first quartile's outer limit is cut, if it is zero.
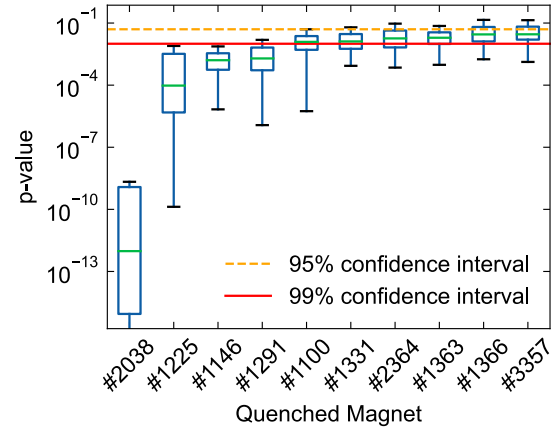

Fig. 8. Boxplot of the 10 FPA events with the smallest median p-values. For each event, the quenched magnet is given on the x-axis. Each box extends from the first to the third quartile, with a horizontal line at the median. The lines extending from the box further show the variability of p-values. They cover a range of 1.5 times the box length, either subtracted from the first quartile or added to the third quartile. The orange dashed line indicates the confidence interval of 95% at a p-value of 0.05, while the red vertical line indicates the 99% confidence interval at a p-value of 0.01.

The four FPA events with the quenched magnets #2038, #1225, #1146, and #1291 state a median p-value smaller than 1%. In the boxplot for those FPA events, the first and third quartiles are below the 99% confidence interval, and the outer box limits are below 95%. For the six other FPA events in Fig. 8, both quartiles and outer limits are above their respective 99% and 95% intervals. Based on this classification, only the four events with a median p-value of less than 1% are therefore referred to as anomalies. In the following subsection, these anomalies are discussed in more detail.

D. Anomaly Interpretation

To better understand the characteristics of the anomalies, the spectral component weights in the anomalous FPA events are compared to those in normal FPA events. Notably, the weights of SC7 are elevated in three of the four identified FPA events. As discussed above, SC7 represents a broadband spectrum and has an average maximum amplitude of 1 mV for FPA events with a single quench (see Tab. I). However, with 820 µV only the maximum amplitude in the anomalous FPA event where the magnet #1146 quenched is similar to this average. For the FPA events where the magnets #2038, #1225, and #1291 quenched the amplitudes of this spectral component are 240 mV, 80 mV, and 210 mV, respectively. These amplitudes are more than 80 times higher than the

maximum average amplitude for other FPA events. In no other FPA event with a single quench the values are this high. It is inferred that a high SC7 in the quenched magnet is a sufficient criterion for identifying an anomaly.

Having identified the criticality of a high SC7 amplitude in the quenched magnet, the previously excluded 157 FPA events with secondary quenches are also examined for this characteristic. One FPA event with an amplitude of 1200 mV, more than 1200 times higher than the average maximum, stands out. Therefore, this FPA event, where a quench occurred at the magnet #2421, is also referred to as an anomaly in the course of this analysis.

In magnet #1146, the SC7 amplitude is not elevated during the quench. Instead, the low p-value in this FPA event is associated with a high SC1 amplitude of 1500 mV in the physical neighbors of the quenched magnet. As discussed in Subsection III-B, this component represents electromagnetic perturbations propagating through instrumentation wires and the connected QDS measurement units. Based on current knowledge, these electromagnetic perturbations are not associated with a hardware fault.

E. Recommended Maintenance Actions

One of the four quenched magnets, with a significantly increased SC7 amplitude during a FPA event, has developed an intermittent short circuit during the FPA event on 25.04.2021. As such an intermittent short circuit is a critical event, the other three magnets #1225, #1291, and #2421 are also treated as potentially critical and will be checked by transient voltage measurement. If an intermittent short circuit cannot be excluded during the transient voltage measurement, these magnets could be replaced in one of the next maintenance stops of the LHC. In any case, the electronics of the QDS measurement units of these magnets should be exchanged in order to exclude measurement errors.

Transient measurements will also be performed on the magnet #1146. These measurements will provide further information about electromagnetic perturbations.

Tab. II summarizes the five discussed anomalies, sorted by their median p-value in the first column. The second column shows the quenched magnet, the affected circuit and the date of the related FPA event. The third column summarizes main findings of the FPA event and states the recommended maintenance actions.

IV. CONCLUSION

In this study, the voltages measured across the 1232 magnets in the eight LHC main dipole circuits were analyzed to understand the normal and abnormal behavior of the circuits. Specifically, the amplitude and propagation of the frequency spectra measured at the magnets in 699 Fast Power Abort (FPA) events were investigated using Non-negative Matrix Factorization (NMF).

This allowed the extraction of seven spectral components that define normal behavior, occurring in the measured voltages during a FPA event. Analyzing the spectral components' distribution and propagation across the circuit

TABLE II
LIST OF DETECTED ANOMALIES WITH RECOMMENDED MAINTENANCE ACTIONS IN THE REMARKS COLUMN.

Median p-value	FPA Event	Remarks
8×10^{-11}	#2038 Sector 78 25.04.2021	High SC7 in #2038 (240 mV) Exchanged on 25.04.2021 due to intermittent short circuit
7×10^{-5}	#1225 Sector 45 12.05.2021	High SC7 in #1225 (80 mV) Additional measurements Hardware replacements
1×10^{-3}	#1146 Sector 34 06.05.2021	High SC1 (1500 mV) Additional measurements
2×10^{-3}	#1291 Sector 12 14.05.2021	High SC7 in #1291 (210 mV) Additional measurements Hardware replacements
-	#2421 Sector 34 20.04.2021	High SC7 in #2421 (1200 mV) Additional measurements Hardware replacements

and across FPA events provided a deeper understanding of the mutual interaction of hardware components and allowed identifying the potential physical processes causing the spectral components. It was shown that spectral components one to three, with maxima at 3 Hz, 6 Hz, 20 Hz, are induced by the quench due to electromagnetic perturbations. Their amplitudes are one order of magnitude higher when the quenched magnet was produced by manufacturer one. Spectral component four shows a 66 Hz oscillation induced by the quench. With maxima at 150 Hz and 478 Hz, components five and six are independent of the quench and were attributed to artifacts of the QDS measurement unit, and to passive elements in the power converter of one individual circuit. Spectral Component seven (SC7) shows a broadband spectrum, induced by the quench. As previous studies showed that such broadband spectra can be an indicator of short circuits [19], SC7 could also indicate an intermittent short circuit in the magnet.

Five magnets with abnormal behavior during FPA events were detected using the reconstruction loss of NMF and the SC7 amplitude at the quenched magnet. One of these magnets was replaced on 25.04.2021 after a short circuit was detected following the FPA event. Similarly to the replaced magnet, three of the four remaining magnets showed an elevated SC7 amplitude during their quench, which is more than 80 times higher than normal. Dedicated transient measurements will be performed on these magnets and the electronics of their QDS measurement unit should be replaced. If an intermittent short circuit still cannot be excluded, the three magnets could be replaced in one of the next maintenance stops of the LHC to prevent weeks of unplanned LHC downtime. In the magnet which did not show a high SC7 during the FPA event in the quenched magnet, data do not indicate a hardware fault. Instead, an abnormally high spectral component one was observed, which will also be evaluated by transient measurements. The presented methodology has proven to be a powerful tool to describe the normal behavior of the circuits systematically and to detect abnormal behavior indicating potential hardware fatigue and degradation of hardware components in the circuit.

References

[1] J. Schultz, "Protection of Superconducting Magnets," *IEEE Transactions on Applied Superconductivity*, vol. 12, no. 1, pp. 1390–1395, 2002.

[2] E. Ravaioli, K. Dahlerup-Petersen, F. Formenti, V. Montabonnet, M. Pojer, R. Schmidt, A. Siemko, M. Solfaroli Camillocci, J. Steckert, H. Thiesen, and A. Verweij, "Impact of the Voltage Transients After a Fast Power Abort on the Quench Detection System in the LHC Main Dipole Chain," *IEEE Transactions on Applied Superconductivity*, vol. 22, no. 3, pp. 9002504-9002504, 2012.

[3] R. Denz and F. Rodriguez-Mateos, "Electronic Systems for the Protection of Superconducting Elements in the LHC," *IEEE Transactions on Applied Superconductivity*, vol. 16, no. 2, pp. 1725-1728, 2006.

[4] R. Denz, K. Dahlerup-Petersen, F. Formenti, K. H. Meß, A. Siemko, J. Steckert, and L. Walckiers, "Upgrade of the Protection System for Superconducting Circuits in the LHC," *Proceedings of the Particle Accelerator Conference*, vol. 23, pp. 244-246, 2009.

[5] F. Rodriguez-Mateos, P. Pugnat, S. Sanfilippo, R. Schmidt, A. Siemko, and F. Sonnemann, "Quench Heater Experiments on the LHC Main Superconducting Magnets," *Proceedings of the Particle Accelerator Conference*, vol. 1-3, pp. 2154-2156, 2000.

[6] F. Sonnemann, "Resistive Transition and Protection of LHC Superconducting Cables and Magnets," Ph.D. thesis, RWTH Aachen University, 2001.

[7] F. Bourgeois and K. Dahlerup-Petersen, "Methods and Results of Modeling and Transmission Line Calculations of the Superconducting Dipole Chains of CERN's LHC Collider," *LHC Project Report 497*, CERN, 2001.

[8] K. Dahlerup-Petersen, F. Rodriguez-Mateos, R. Schmidt, F. Sonnemann, "Energy Extraction for the LHC Superconducting Circuits," *Proceedings of the Particle Accelerator Conference*, Vol. 5, pp. 3448-3450, 2002.

[9] E. Ravaioli, K. Dahlerup-Petersen, F. Formenti, J. Steckert, H. Thiesen, and A. Verweij, "Modeling of the Voltage Waves in the LHC Main Dipole Circuits," *IEEE Transactions on Applied Superconductivity*, vol. 22, no. 3, pp. 9002704-9002704, 2012.

[10] A. Verweij, V. Baggiolini, A. Ballarino, B. Bellesia, F. Bordry, A. Cantone, M. Casas Lino, A. Serra, C. Trello, N. Catalan Lasheras, Z. Charifoulline, G. Coelingh, K. Dahlerup-Petersen, G. D'Angelo, R. Denz, S. Feher, R. Flora, M. Gruwé, V. Kain, and M. Zerlauth, "Performance of the Main Dipole Magnet Circuits of the LHC during Commissioning," *Proceedings of European Particle Accelerator Conference*, vol. 11, pp. 2473–2475, 2008.

[11] L. Coull, D. Hagedorn, G. Krainz, F. Rodriguez-Mateos, R. Schmidt, et al., "Quench Propagation Tests on the LHC Superconducting Magnet String," *LHC Tech. Rep.*, no. 70, 1996.

[12] A. Siemko, "Magnet Quench Process," *CERN Tech. Rep.*, no. 567209, 2001.

[13] R. Saederup, "Local Transfer Function Measurement Data Analysis," *CERN Tech. Rep.*, no. 2675917, 2021.

[14] E. Ravaioli, A. P. Verweij, and H. H. J. ten Kate, "Unbalanced Impedance of the Aperture Coils of Some LHC Main Dipole Magnets," *IEEE Transactions on Applied Superconductivity*, vol. 23, no. 3, pp. 4000104-4000104, 2013.

[15] E. Ravaioli, "First Analyses and Conclusions on the Fast Power Abort Measurements in Sector 67 During the Long Shut-Down 2010-11," *CERN Tech. Rep.*, no. 1137620, 2011.

[16] STEAM website, https://espace.cern.ch/steam/ (accessed Oct. 16, 2023).

[17] M. Janitschke, "Framework for automatic superconducting magnet model generation & validation against transients measured in LHC magnets," Master's thesis, Technical University of Berlin, 2021.

[18] D. Hoang, C. Boffo, N. Tran, S. Krave, S. Kazi, S. Stoynev, and V. Marinozzi, "Intelliquench: an adaptive machine learning system for detection of superconducting magnet quenches," *IEEE Transactions on Applied Superconductivity*, vol. 31, no. 5, pp. 1–5, 2021.

[19] Z. Charifoulline, L. Bortot, R. Denz, F. Rodriguez Mateos, A. Siemko, J. Steckert, A. Verweij, and G. Willering, "Overview of the performance of quench heaters for high-current LHC superconducting magnets," *IEEE Transactions on Applied Superconductivity*, vol. 27, no. 4, pp. 1–5, 2016.

[20] M. Wielgosz, A. Skoczeń, and M. Mertik, "Using LSTM recurrent neural networks for monitoring the LHC superconducting magnets," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 867, no. 1, pp. 40–50, 2017.

[21] B. Goodman and S. Flaxman, "European Union regulations on algorithmic decision-making and a 'right to explanation'," *AI Magazine*, vol. 38, no. 3, pp. 50-57, 2017.

[22] J.M. Be'itez, J.L. Castro, and I. Requena, "Are artificial neural networks black boxes?," *IEEE Transactions on Neural Networks*, vol. 8, no. 5, pp. 1156-1164, 1997.

[23] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788-791, 1999.

[24] J.-P. Brunet, P. Tamayo, T. R. Golub, and J. P. Mesirov, "Metagenes and molecular pattern discovery using matrix factorization," *Proceedings of the National Academy of Sciences*, vol. 101, no. 12, pp. 4164–4169, 2004.

[25] H. Sawada, H. Kameoka, S. Araki, and N. Ueda, "Multichannel extensions of non-negative matrix factorization with complex-valued data," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 5, pp. 971-982, 2013.

[26] M. R. Blanton and S. Roweis, "K-corrections and filter transformations in the ultraviolet, optical, and near-infrared," *The Astronomical Journal*, vol. 133, no. 2, pp. 734-754, 2007.

[27] H. J. Nussbaumer, *The Fast Fourier Transform*. Springer, 1981.

[28] A. V. Oppenheim, *Discrete-time Signal Processing*, 1999.

[29] Fredric J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, 1978.

[30] Å. Björck, *Numerical Methods for Least Squares Problems*. SIAM, 1996.

[31] D. Lee and H. Sebastian Seung, "Algorithms for Non-negative Matrix Factorization," *Advances in Neural Information Processing Systems*, vol. 13, MIT Press, 2000.

[32] R. Kompass, "A generalized divergence measure for nonnegative matrix factorization," *Neural Computation*, vol. 19, no. 3, pp. 780-791, 2007.

[33] Y. Cao, P.P.B. Eggermont, and S. Terebey, "Cross Burg entropy maximization and its application to ringing suppression in image reconstruction," *IEEE Transactions on Image Processing*, vol. 8, no. 2, pp. 286-292, 1999.

[34] C. Févotte, N. Bertin, and J.-L. Durrieu, "Nonnegative matrix factorization with the Itakura-Saito divergence: With application to music analysis," *Neural Computation*, vol. 21, no. 3, pp. 793-830, 2009.

[35] C. Boutsidis and E. Gallopoulos, "SVD based initialization: A head start for nonnegative matrix factorization," *Pattern Recognition*, vol. 41, no. 4, pp. 1350-1362, 2008.

[36] Y.-X. Wang and Y.-J. Zhang, "Nonnegative matrix factorization: A comprehensive review," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 6, pp. 1336-1353, 2012.

[37] R. Peharz and F. Pernkopf, "Sparse nonnegative matrix factorization with $\ell 0$-constraints," *Neurocomputing*, vol. 80, no. 1, pp. 38-46, 2012.

[38] Cyrus D. Cantrell, *Modern Mathematical Methods for Physicists and Engineers*. Cambridge University Press, 2000. ISBN 0-521-59827-3.

[39] T. Smets, N. Verbeeck, M. Claesen, A. Asperger, G. Griffioen, T. Tousseyn, W. Waelput, E. Waelkens, and B. De Moor, "Evaluation of distance metrics and spatial autocorrelation in uniform manifold approximation and projection applied to mass spectrometry imaging data," *Analytical Chemistry*, vol. 91, no. 9, pp. 5706-5714, 2019.

[40] M. Suresh and I. R. Varughese, "Effect of Similarity Measures on Repetitive Audio Source Separation," *International Journal of Scientific Engineering and Technology Research*, vol. 3, no. 30, pp. 5933-5934, 2014.

[41] J. Pearl and D. Mackenzie, *The Book of Why: The New Science of Cause and Effect*. Basic Books, 2018.

[42] R. J. Rossi, *Mathematical Statistics: An Introduction to Likelihood Based Inference*, John Wiley & Sons, 2018.

[43] J. W. Tukey, *Exploratory Data Analysis*. Addison-Wesley, 1977.