FERMILAB-TM-2834-TD

Exploration of signal processing methods for superconducting magnet and quench data

Ghanshyam Bhatt[1]
Advisor: Stoyan Stoynev

August 2024

Fermi National Accelerator Laboratory, VFP-2024 Final report

[1]Email address: gbhatt@tnstate.edu

Exploration of signal processing methods for superconducting magnet and quench data

Abstract

Quenching is the phenomenon of a superconducting magnetic material carrying current transitioning into a regular conducting material. This may cause severe and irreparable damage to the superconductor due to Joule heating. The Magnet Department at Fermi National Accelerator Laboratory (FNAL) has acquired experimental data through quench antenna arrays that are recorded when the quench is detected. These data are in terms of voltage signals that are sampled at $100kHz$ for several minutes. There are multiple channels and each channel provides a data set of more than 20 million observations, while there is one channel, called the trigger channel which shows the time when quench is detected. Despite some advancements that were made including machine learning, data complexity still shadows the progress. In this work, we studied a multi-resolution analysis of the quench antenna data through the Haar wavelet transform. In particular, we applied the maximally overlapped discrete wavelet transform (MODWT) of a suitable level L to the given data and then projected it onto the wavelet basis. This decomposes a given signal (Original data) $x \in \mathbb{R}^N$ into $L+1$ subspaces of $\mathbb{R}^N$. One of the subspaces called the approximation, captures the trend of the signal, and the others, called the details, capture the fluctuations at different frequency bands. This decomposition provides a clear trend of the data at a suitable level and also various activities (spikes) are seen in the details of the decomposition at every level. These spikes might reveal some information about the quench under investigation but in any case, give information about magnet behavior. Also, this decomposition is seen to be very useful in removing noise present in the data due to the source or mechanism of the experiment.

Acknowledgements

- This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.
- This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Visiting Faculty Program (VFP).

1 Introduction

Superconducting magnets are used widely in particle accelerators since they carry large current densities with no electrical resistance. They generate very intense magnetic fields. These generated fields are then responsible for keeping particle beams stable and precisely aligned, or squeezing the beams closer together when they enter a particle detector. Thus, superconducting magnets play an extremely important role in ensuring the normal operation of particle accelerators. The resistance of the magnet coil windings is zero in the superconducting state, so it requires no energy to maintain the current flow. If the coil temperature rises above the superconductivity threshold, the windings suddenly develop a finite resistance. The circulating current passing through this elevated coil resistance creates heat, which quickly results in further expansion of the resistive zone. This irreversible phenomenon is called *quench*. Having large stored energy, magnets can be damaged by quenches due to localized heating, high voltage, or large force transients. The loss has a huge price. The available quench protection systems can detect quenches only after they happen.

Fermi National Accelerator Laboratory (FNAL) has been developing multiple versions of flexible quench antennas (flex-QA), including some specially optimized for high sensitivity and/or high resolution, to characterize quench events and transients during current ramping in superconducting magnets [1]. Flex-QA arrays can be used to pinpoint quench locations with remarkable precision, although the time development of signals was not fully characterized. Still, acoustic sensors, initially developed by Lawrence Berkeley National Laboratory (LBNL) [2] were also used in many magnet tests and are becoming a standard tool for magnet characterization. Both quench antenna and acoustic sensors can gather data during the whole current ramp (to quench) of a magnet and the combined data provide extremely rich and detailed information about magnet behavior and, possibly, quenches. Authors in [3] have established that there could be quench precursors that exist in acoustic data by investigating a dynamic learning setup using Auto-Encoder (AE) Deep Neural Network (DNN). Those manifest as anomalies in acoustic data during current ramps. While some of the current research on magnet data and applications at FNAL can be found in [1], [3], and [4], what is missing from all studies so far is a detailed mathematical investigation on signal properties and anomalies in the large amount of data available.

Understanding the quench antenna data information content for both ramp and quench data using mathematical tools for signal processing, with similarities to [5], is the objective of the proposed work.

Research Objectives

The fact that data from quench antenna and acoustic data have a lot of features during magnet ramp(s) and quench(es) has been established. Finding relevant features in data or relations between "events" in data or different parts of data is still in the early steps. An important aspect of those developments is signal processing and exploring mathematical tools to extract maximal useful information. The research objective of this project is to further study and advance the topic of data processing with the aim being to find proper mathematical tools and techniques to describe the processes we observe in terms of relations over time or space, eventually linking earlier magnet behavior to actual quenches. We will apply maximally overlapped discrete wavelet transform on the quench antenna data.

Notations

Throughout the paper $\mathbb{Z}$, and $\mathbb{R}$ are used to denote the set of integers and the set of real numbers respectively. N denotes the length of the data, F_s denotes the sampling frequency, and L denotes the level of wavelet decomposition. A scaling function is denoted by ϕ while ψ denotes the wavelet function. The space $L^2(\mathbb{R})$ is a usual Hilbert space but here it can be considered as a space of finite energy signals.

2 Method of Wavelet Analysis

2.1 Continuous Wavelet Transform

Wavelet functions provide time and frequency information of a signal together unlike the Fourier transform which provides localized frequency information. This transformation yields a multiresolution of the signal [8]. A wavelet is a function $\psi \in L^2(\mathbb{R})$ with zero average. A wavelet family $\psi_{a,b}$ is the set of functions generated by scaling and translation of a given mother wavelet ψ, viz.

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right), \tag{1}$$

where $a, b \in \mathbb{R}, a \neq 0$ are the scaling and translation parameters respectively.
The *continuous wavelet transform* (CWT) of a signal $f(t) \in L^2(\mathbb{R})$ is defined as

$$W_\psi f(a,b) := |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t)\psi^*\left(\frac{t-b}{a}\right) dt =: \langle f, \psi_{a,b}\rangle, \tag{2}$$

where ψ^* is the complex conjugate of the function ψ. This inner product captures the correlation between the function $f(t)$ and the wavelet family $\psi_{a,b}$ at each scale a and translation parameter b. The numbers $\langle f, \psi_{a,b}\rangle$ are called the wavelet coefficients. If a signal has a major component within the current scale, the CWT coefficients calculated within this scale in the time-scale domain would be relatively large. This dilation and translation of wavelet make it an applicable transformation because it can reveal transient features in both time and frequency domains. This transformation is invertible, so a perfect reconstruction of the signal $f(t)$ is possible. Theoretically, this transformation works perfectly but in practice, there are infinitely many coefficients and there is redundancy in the representation of a signal.

Let ϕ be such that its integer translates $\phi_{0,k}(x) := \{\phi(x-k),\ k \in \mathbb{Z}\}$ form an orthonormal basis for some closed subspace $V_0 \in L^2(\mathbb{R})$, i.e. $V_0 = \overline{span\{\phi(t-k),\ k \in \mathbb{Z}\}}$. The function ϕ is called the *scaling function* and the space V_0 is called the *approximation space.* It turns out that the family of functions $\phi_{j,k} := 2^{\frac{-j}{2}}\phi(2^{-j}t - k),\ \ k \in \mathbb{Z}$ forms an orthonormal basis for V_j, where $V_j = \overline{span\{\phi_{j,k},\ k \in \mathbb{Z}\}}$. This is an approximation space at scale at scale 2^j. This generates a sequence of closed subspaces of $L^2(\mathbb{R})$

$$.....V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots. \tag{3}$$

This is what constitutes *a multi-resolution analysis* (MRA) of the space $L^2(\mathbb{R})$ [8]. We assume that the signal is in the space V_0. In this case it is said to be at a scale $2^0 = 1$. We can approximate

the signal by projecting the signal onto the space V_1 as well. Since $V_1 \subset V_0$, and the set $\phi(t-k),\ \ k \in \mathbb{Z}$ forms an orthogonal basis for V_0, some extra functions are needed to have an orthogonal basis for V_1. These extra functions are wavelet functions $\psi(t-k),\ \ k \in \mathbb{Z}$. This wavelet family spans a space $W_1 \subset V_0$, so that $V_0 = V_1 \bigoplus W_1$. This approximation is at scale $2^1 = 2$. This is called level 1 approximation. We therefore get a sequence of subspaces $\cdots W_2 \subset W_1 \subset W_0 \subset W_{-1} \subset \cdots$, such that $V_j = V_{j+1} \bigoplus W_{j+1}$. By iteration of the above, we obtain Since $V_0 = V_j \bigoplus W_j \bigoplus W_{j-1} \bigoplus \cdots \bigoplus W_2 + \bigoplus W_1$. So the function $f(t) \in V_0$ can be represented as

$$f(t) = \sum_{k\in\mathbb{Z}} a_k^{j+1}\phi(2^{-(j+1)}t - k) + \sum_{i=1}^{j}\sum_{k\in\mathbb{Z}} b_k^i \psi(2^{-i}t - k),$$

where a_k^{j+1} and b_k^{j+1} are the scaling and wavelet coefficients. Thus the function $f(t)$ can be decomposed as

$$f(t) = f_j + \sum_{i=1} w_j, \tag{4}$$

where f_j is the projection onto V_j and w_j is the projection onto W_j. The simplest wavelet is the well-known *Haar wavelet*, whose scaling function ϕ, and the wavelet function ψ are given respectively by

$$\phi(x) = \begin{cases} 1 & \text{if } 0 \le x < 1, \\ 0 & \text{otherwise,} \end{cases} \qquad \psi(x) = \begin{cases} 1 & \text{if } 0 \le x < 1/2, \\ -1 & \text{if } 1/2 \le x < 1. \end{cases}$$

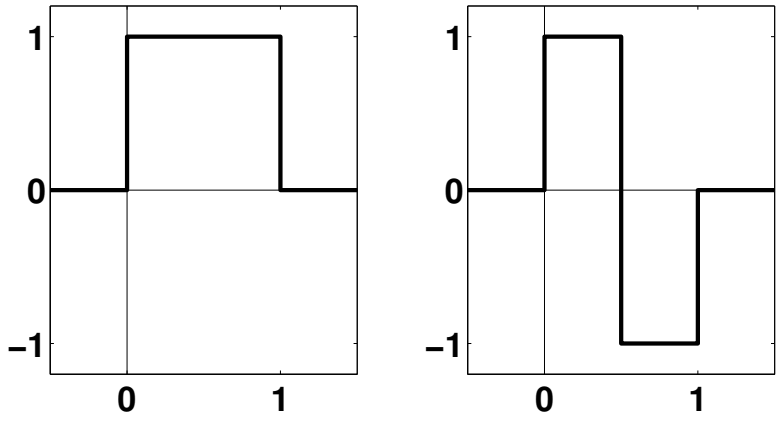

Figure 1: Haar scaling function and Haar wavelet function

2.2 Discrete Wavelet Analysis

The *discrete wavelet transform* (DWT) is a sampled wavelet transform applicable to digital signals, it provides a non-redundant representation of the signal and this can be easily implemented. For

this, we choose a mother wavelet function ψ and the corresponding scaling function ϕ. Throughout this paper, we adopt the convention that the given signal lies in a space $V_0 \in \mathbb{R}^N.$, where N is the length of the signal. This signal can be decomposed as $V = V_1 \bigoplus W_1$, and recursively

$$V_0 = V_L \bigoplus W_L \bigoplus W_{L-1} + \cdots \bigoplus W_1,$$

provides the $L-$ level decomposition of the given space V_0. The space V_L captures the average trend of the signal at a specified scale, and the subspaces $W_1, \cdots W_L$ are related to the changes of averages over specific scales. The parameter L determines the scale and magnitude of the corresponding scaling function. The parameter k determines the time displacement of the wavelet. The wavelet coefficients capture the details as the signal is approximated by the scaling function. The subspace V_L maintains the domain properties of the signal and the subspaces W_j preserve its properties in the frequency domain. The space W_j contains the frequency information of the signal in the frequency band $\frac{F_s}{2^{j+1}} \leq \omega \leq \frac{F_s}{2^j}$, F_s is the sampling frequency. The space V_L contains the frequency information for the frequency range $0 \leq \frac{F_s}{2^{L+1}}$. The decomposition level is guided by the signal and the wavelet chosen. We are concerned with only the Haar wavelet in this study for which the maximum level is given by $Log_2 N$ [8] but it's not necessary to the maximum leve

2.3 Discrete Haar Wavelet Transform

Discrete Haar wavelet transform is a 2 tap wavelet transform. The scaling and the wavelet filter coefficients are denoted by $h_{j,k}$, and $g_{j,k}$, where j denotes the level and k denotes translation. The level 1 scaling filters are $h_{1,0} = h_{1,1} = \frac{1}{\sqrt{2}}$, and the wavelet filters are $g_{1,0} = \frac{1}{\sqrt{2}}$, and $g_{1,1} = -\frac{1}{\sqrt{2}}$ [8].

Example 1 *Let $N = 8$, let $x = (x_0, x_1, \cdots, x_7)^T \in \mathbb{R}^8$ be a signal, where T denotes the transpose. Its level 1 Discrete Haar transform $\widetilde{W_1}x$ provides the approximation coefficients a_1 and detail d_1 which are given by $\widetilde{W_1}x = (S_1x, W_1x)^T = (a_1, d_1)^T$, where*

$$\widetilde{W_1} = \left(\begin{array}{cccccccc} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \hline \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & & & \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{array}\right) = \left(\begin{array}{c} S_1 \\ \hline W_1 \end{array}\right)$$

This decomposition projects the vector x onto two 4 dimensional subspaces of $\mathbb{R}^8$. It turns out that the projection onto the scaling and wavelet spaces as in the equation (4) is given by [8],

$$A_1 = S_1^T S_1 x, \text{ and } D_1 = W_1^T W_1 x, \text{ so that } x = A_1 + D_1.$$

For the level 2 transform, there are $2^2 = 4$ coefficients. They are $h_{2,0} = h_{2,1} = h_{2,3} = h_{2,4} = \frac{1}{2}$, $g_{2,0} = g_{2,1} = \frac{1}{2}$, and $g_{2,1} = g_{2,3} = -\frac{1}{2}$. W_2x provides the approximation coefficients a_2 and detail

d_1 *and* d_2 *which are given by given by* $\widetilde{W}_2 x = (S_2 x, W_2 x, W_1 x)^T = (a_2, d_2, d_1)^T$, *where*

$$\widetilde{W}_2 = \left(\begin{array}{cccccccc} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \hline \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \hline \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{array}\right) = \left(\begin{array}{c} S_2 \\ \hline W_2 \\ \hline W_1 \end{array}\right)$$

At the third level, there are $2^3 = 4$ *coefficients. The scaling coefficients are* $h_{3,0} = h_{3,1} = h_{3,2} = h_{3,3} = h_{3,4} = h_{3,5} = h_{3,6} = h_{3,7} = \frac{1}{2\sqrt{2}}$, $g_{3,0} = g_{3,1} = g_{3,2} = g_{3,3} = \frac{1}{2\sqrt{2}}$, *and* $g_{3,4} = g_{3,5} = g_{3,6} = g_{3,7} = -\frac{1}{2\sqrt{2}}$. $W_3 x$ *provides the approximation coefficients* a_3 *and detail* d_1, d_2, *and* d_3 *which are given by given by* $\widetilde{W}_1 x = (S_3 x, W_3 x, W_2 x, W_1 x)^T = (a_3, d_3, d_2, d_1)^T$, *where*

$$\widetilde{W}_3 = \left(\begin{array}{cccccccc} \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \\ \hline \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} \\ \hline \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \hline \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{array}\right) = \left(\begin{array}{c} S_3 \\ \hline W_3 \\ \hline W_2 \\ \hline W_1 \end{array}\right)$$

This decomposition projects the vector x *onto four* 4 *subspaces of* $\mathbb{R}^8$. *They are a scaling subspace (one dimensional), and* 3 *wavelet subspaces (four, two, and one dimensional each) and it turns out that the projection onto the scaling and wavelet spaces as in (4) is given by [8],*

$$A_3 = S_3^T S_3 x, \;\; D_3 = W_3^T W_3 x, \;\; D_2 = W_2^T W_2 x, D_1 = W_1^T W_1 x,$$

so that the multiresolution decomposition is given by

$$x = A_3 + \sum_{i=1}^{3} D_i \tag{5}$$

Note 1 *Since* $N = length(x) = 2^3 = 8$, *we can perform only* $L = 3$ *levels.* A_3 *provides the trend of the signal and* D_i *represents the details at those three levels for* $i = 1, 2, 3$. *If the signal is of length (*$N = 2^n$*) for some* n, *then* n *levels of decomposition can be performed. At each level of decomposition, the length of the source vector (signal) goes down by half. For example the vector* $x \in \mathbb{R}^8$ *but the length of* a_1 *and* d_1 *is* 4. *The length of* d_2 *and* a_2 *are 2 each and the length of* a_3, d_3 *are 1 each.*

2.4 Maximal Overlapping Discrete Wavelet Transform

Maximal Overlapping Discrete Wavelet Transform (MODWT) is similar to DWT except there are a few things that make it a better choice for the antenna data analysis of this project. The DWT for level L requires the data to be of size 2^L, while MODWT does not, but this comes at some computational price [8] because MODWT does not downsample and provides the vector of the same length at each level, unlike DWT. Other differences are provided in [8]. MODWT produces a $L \times N$ matrix at the L^{th} level where N is the length of the signal. Each row provides the frequency information in a different frequency band. MODWT partitions a signal's energy across the detail and scaling coefficients. For example, at the first level, MODWT provides a $2 \times N$ matrix where the second row provides the average of points called the (approximation, A_1) and the first row provides the fluctuation of the data from the average (details, D_1). A third-level MODWT provides a $3 \times N$ matrix, of which the third row provides the approximation and the first two rows provide the details at two different scales. The scaling coefficients for the Haar MODWT are given by [8]

$$\tilde{h}_{j,k} = \frac{h_{j,k}}{2^{j/2}}, \text{ and } \tilde{g}_{j,k} = \frac{g_{j,k}}{2^{j/2}}, \tag{6}$$

where $h_{j,k}$ and $g_{j,k}$ are the filter coefficients of the Haar DWT algorithm outlined above. The number of filter coefficients at each level given by $L_j = (2^j - 1)(M - 1) + 1$, where M is the filter length [8]. In our case, it's Haar wavelet so the width of the filter is $M = 2$, so the filter length is $L_j = 2^j$. These filters convolve the original signal and the MODWT pyramid algorithm generates the approximation and detail coefficient. For clarity, here are the matrices that perform the first and the second-level MODWT.

Example 2 *At level* 1, *MODWT produces a* $2 \times N$ *matrix of which the first and the second rows are given by* $(W_1 x)^T$ *and* $(S_1 x)^T$, *respectively where* W_1, *and* S_1 *are given by the following* $(n \times N)$ *matrices*

$$W_1 = \begin{pmatrix} \frac{1}{2} & 0 & \cdots & \cdots & 0 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & 0 & \cdots & \cdots & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots\cdots & 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \cdots\cdots & 0 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad S_1 = \begin{pmatrix} \frac{1}{2} & 0 & \cdots & \cdots & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & \cdots & \cdots & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots\cdots & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \cdots\cdots & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

For level 2 *there are* $2^2 = 4$ *coefficients. They are* $h_{2,0} = h_{2,1} = h_{2,3} = h_{2,4} = \frac{1}{4}$, $g_{2,0} = g_{2,1} = \frac{1}{4}$, *and* $g_{2,1} = g_{2,3} = -\frac{1}{4}$. *The level* 2 *transform matrix (approximation) is given by*

$$S_2 = \begin{pmatrix} \frac{1}{4} & 0 & \cdots & \cdots & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & 0 & \cdots & \cdots & 0 & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & \cdots & \cdots & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & \cdots & \cdots & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & \cdots & \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \cdots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

The wavelet matrix W_2 can be considered similarly. At the third level, there are $2^3 = 4$ coefficients. The scaling coefficients are $h_{3,0} = h_{3,1} = h_{3,2} = h_{3,3} = h_{3,4} = h_{3,5} = h_{3,6} = h_{3,7} = \frac{1}{8}$. So the scaling coefficients provide the averages of 8 entries from the signal at a time. An important thing to notice is that the first few entries (in this case it is $2^2 - 1 = 3$) take the points from the tail of the data. At level 3, it takes the 7 points from the end for convolution.

Note 2 *We are going to consider level $L = 16$ below. So $2^{16} - 1$ points at the beginning and end of the reconstructed signal will be misleading because of the boundary conditions. The circular boundary condition of the MODWT algorithm assumes that the data is periodic i.e. the given data $(x_0, x_1, x_2, \cdots x_{N-1})$ is assumed to be like $(....x_{N-2}, x_{N-1}, x_0, x_1, \cdots x_{N-1})$. In reflection boundary conditions, it is assumed to be like $(...x_2, x_1, x_0, x_1, x_2, \cdots, x_{N-1})$. The reflection boundary requires more computations, therefore we mainly performed the circular boundary conditions here.*

Our results are based on circular boundary conditions but the data after the quench is detected is modified (replaced by the value at the quench time) as explained below. An important feature of MODWT is that the length of the vectors at each level is the same as the length of the given data N. The multi-resolution analysis (projection of the original signal into several wavelet subspaces) as in the equation (3) of the signal is then carried out by MODWTMRA. At level L decomposition of the signal using MODWT, the MODWTMRA provides the projection of the signal x onto the $L + 1$ subspaces of $\mathbb{R}^N$ as in the equation (4)

$$x = \sum_{j=1}^{L} Dj + A_L. \tag{7}$$

2.5 Objective of the Project

The fact that data from quench antenna and acoustic data have a lot of features during magnet ramp(s) and quench(es) has been established. Finding relevant features in data or relations between "events" in data or different parts of data is still in the early steps. Understanding the quench antenna data information content for quench data using mathematical tools for signal processing is the objective of the proposed work. The research objective of this project is to further study and advance the topic of data processing with the aim being to find proper mathematical tools and techniques to describe the processes we observe in terms of relations over time or space, eventually linking earlier magnet behavior to actual quenches. Any anomalies present in the data and mapping them to the "suspicious events" is the primary goal of this project.

3 Scientific Approach

Since Haar wavelets have short support and we are looking for rapidly changing features in our data, we prefer choosing the Haar wavelets in this project. MODWT is defined for all sample sizes except its highly redundant and non-orthogonal transform. Since the size of the quench antenna data is huge, MODWT is preferred here. Several levels are run through the code but most of the results presented here are based on level $L = 16$. Here the signal lies in $\mathbb{R}^N$. This produces a matrix of size $17 \times N$. The last row captures the averages on a specified scale. In our case, it captures the

Table 2: Frequency bands of wavelet decomposition levels

Level	Frequency Bands (Hz)	Filtered by	Remarks
1	$50k - 25k$	D_1	mostely noise
2	$25k - 12.5K$	D_2	Some spikes, some noise
3	$12.5k - 6.25k$	D_3	Some spikes, some noise
4	$6.25k - 3.125k$	D_4	
5	$3.125k - 1.5625k$	D_5	
6	$1.5625k - 781.25$	D_6	
7	$781.25 - 390.625$	D_7	
8	$390.625 - 195.3125$	D_8	
9	$195.3125 - 97.6562$	D_9	
10	$97.6562 - 48.8281$	D_{10}	
11	$48.8281 - 24.4140$	D_{11}	
12	$24.4140 - 12.207$	D_{12}	
$13 - 16$.....	...	$D_{13} - D_{16}$	small spikes only

averages of 2^{16} points at a time (if we perform only up to level 10, this will take the averages of 2^{10} points at a time). This last row which is the projection onto the averaging space, is denoted by A_{16} throughout this document. The first row of the output represents the component of the signal with the highest frequency filtered (in this case it is $[\frac{F_s}{4}, \frac{F_s}{2}]$, where F_s is the sampling frequency). This is the projection of the signal into one of the sub-spaces (high frequency) spanned by wavelet vectors, it is denoted by D_1 throughout this paper. The second row (D_2) represents the component of the signal with frequency at the next level $[\frac{F_s}{8}, \frac{F_s}{4}]$). The other rows $D_3 - D_{16}$ represent the component of the signal in the frequency range as given in the table 2. So the given signal can be represented as in (4)

$$x = A_{16} + D_{16} + D_{15} + \cdots + D_1. \tag{8}$$

According to the Nyquist rule of sampling the highest frequency that can be accurately represented is less than one-half of the sampling rate. According to Mallat's algorithm and this rule the maximum frequency that can be filtered out by the wavelet decomposition is provided in the table 2.

3.1 Structure of Quench Antenna Data

The geometric structure of the data-acquiring channels and their mechanism is provided in [1]. They are a result of 57 experiments on MBHSM03 superconducting magnet that have been performed at the Fermi National Accelerator Lab [6]. In each current ramp, the current was increased from zero until the quench was detected and the voltage signals were recorded at a sampling rate of $100kHz$ for hundreds of seconds. In most cases, the current rate was abruptly changed from $50A/s$ to $20A/s$ near $7.5kA$. The channels are called IN, OUT, LE, and RE channels for each ramp. The data are stored in two groups with each group containing 40 channels indexed from $Voltage\_0$ to $Voltage\_39$ [7]. We are going to use the names as they appear in [7] in this paper. The IN/OUT channels (20) are numbered from IN_T1_T20–IN_T10_T11 and

OUT_T10_T11–OUT_T1_T20. The LE/RE channels (20 each) are numbered LE_T1– LE_T20, and RE_T1–RE_T20. A complete description of the geometry and naming is given in [1, 7]. The quench detection time in each ramp is calculated from the trigger channel "$Voltage\_32$." The channel data used in this work are adjusted to have the quench detection at $t = 0$. For example, the data taken the trigger channel and "IN_T2_T19" channel in Ramp 3 are shown in the figure 2 after adjusting quench detection time to be $t = 0s$.

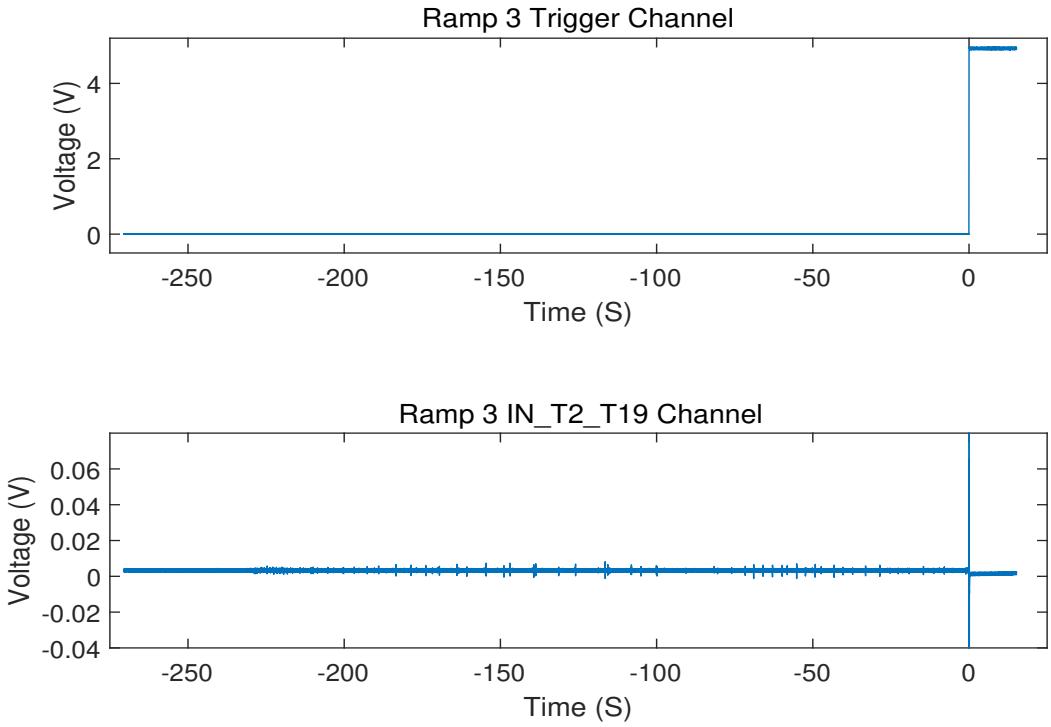

Figure 2: Top is the Ramp 3 trigger Channel and the bottom is IN_T2_T19 channel's original data.

Observation 1:

The channel LE_T10 in ramp 3 is decomposed into 17 subspaces of $\mathbb{R}^N$ by $L(= 16)$ levels of Haar transform. Here, the original length of the signal is $N = 28,434,432$. All entries in the signal after the quench detection time are replaced by the value of the signal at $t = 0s$. In this channel, this value is $-.1053$. The graphs in the figures 3 and 4 illustrate the complete Haar Multiresolution analysis of this signal. The plot for D_1 shows the highest frequency it can filter and so on as given in the table 2. This plot shows that there is a very high frequency, possibly noise, present in the data. As the level increases, the MRA shows that the noise settles down. $D_9 - D_{10}$ show that there is significant noise present in the signal in those frequency bands but that may be from source. After that, there are some samll spikes seen in $D_{11} - D_{16}$. The data in A_{16} captures the trend of the signal. The spike near $t = -100s$ which is seen in the original data seems very clear in the trend, that's originating from the ramping current as reported in [7] too. The spike at $t = -50s$

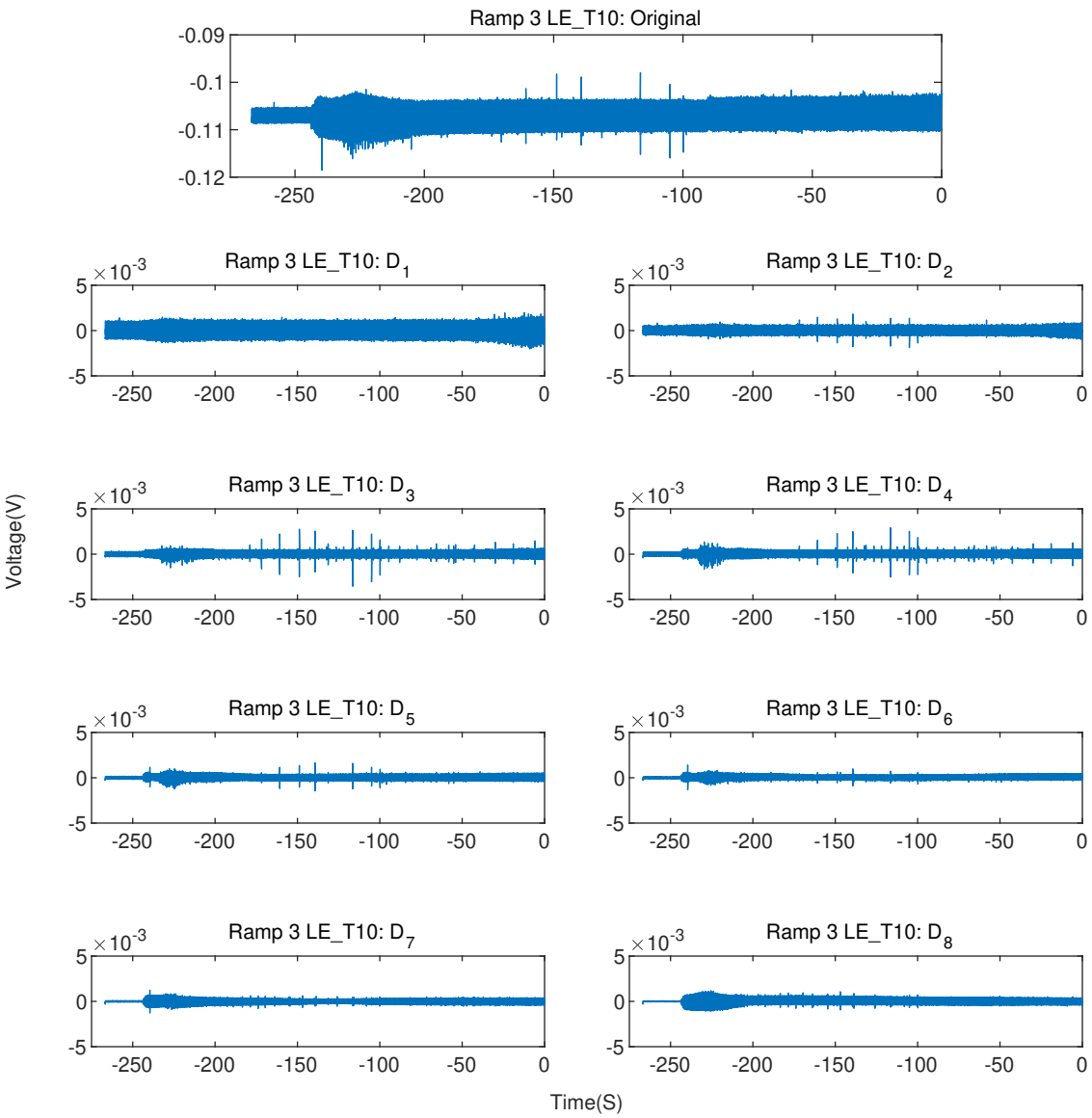

Figure 3: Ramp 3 LE_T10 Channel: The original data and its Haar multi-resolution analysis from D_1 to D_8.

was indicated in [7] as an important anomalous event. It's seen by wavelet decomposition too. We note that this decomposition satisfies the equation (7).

Observation 2:

The MRA allows us to compress the signal. The Figure 6 shows the original signal, the trend signal and the trend with some higher level details. The fourth plot is the signal with trend and details $D_{11} - D_{16}$, ignoring the other levels completely, i.e. ignoring all high frequency content. The third and fourth plot in this figure show the presence of noise in D_{12} and D_{13}. It turns out that the noise

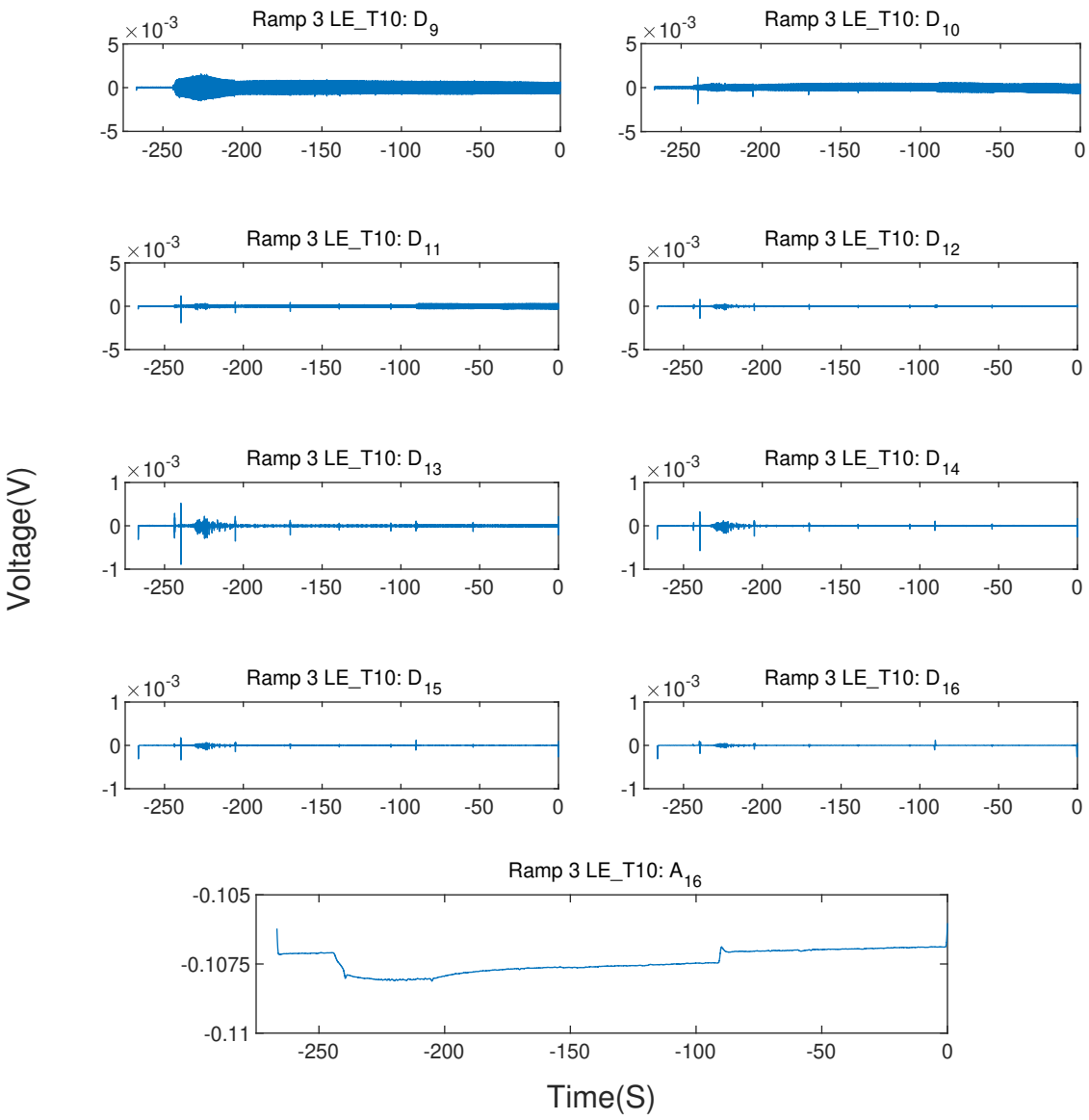

Figure 4: Ramp 3 LE_T10 Channel: Haar multi-resolution analysis of the data from D_9-D_{16} and the trend A_{16} of the data.

gets worse as we go from D_{16} to D_1. The spike near $t = -100s$ as well as $t = -50s$ are clearer here together with other spikes.

Observation 3:

The details at D_8 and D_9 are probably noisy due to the source. After removing all high frequency components, the figure 6 shows some spikes occurring at a regular interval. It was noted in [7] that there are spikes occurring at a constant but extremely low frequencies extremely low frequency of every $35\,s$. The figure 6 shows that they are occurring at a regular interval of approximately $35\,s$.

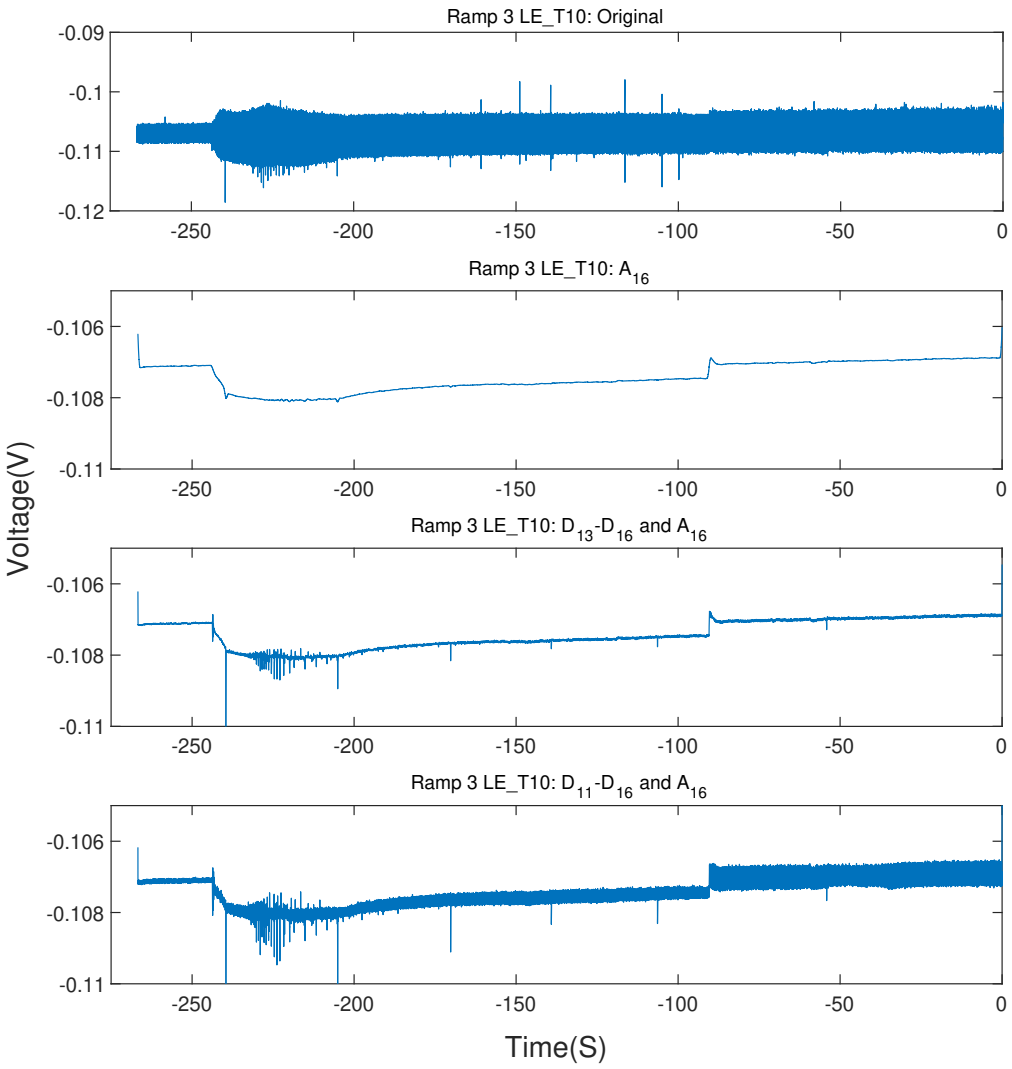

Figure 5: Ramp 3 LE_T10 Channel: The top plot is the original data. The second plot is the trend signal A_{16}. The third plot is the trend with details $D_{13} - D_{16}$, and the fourth plot is the trend with details $D_{11} - D_{16}$

It looks like these spikes are contributed by low frequency part of the signal.

Observation 4

Since details D_i (for lower values of i) are noisy, a hard threshold given by $\lambda_i = \sigma_i\sqrt{2Log(N)}$ is applied at each level i, where σ_i is the standard deviation of the values in D_i. The data up to $t = 0s$ are considered for this standard deviation (i.e. N used here is shorter than the original N). Figure 7 shows D_2, D_3, D_4 and D_5 after applying a hard threshold of λ_i. In this case D_1 completely vanishes. D_2 has some spikes left near $t = -225s$ and $t = 0s$. Actually D_1 shows this pattern too but λ_1 is large enough to nullify it. The figure 7 shows the values in D_2-D_5 after the hard threshold was applied.

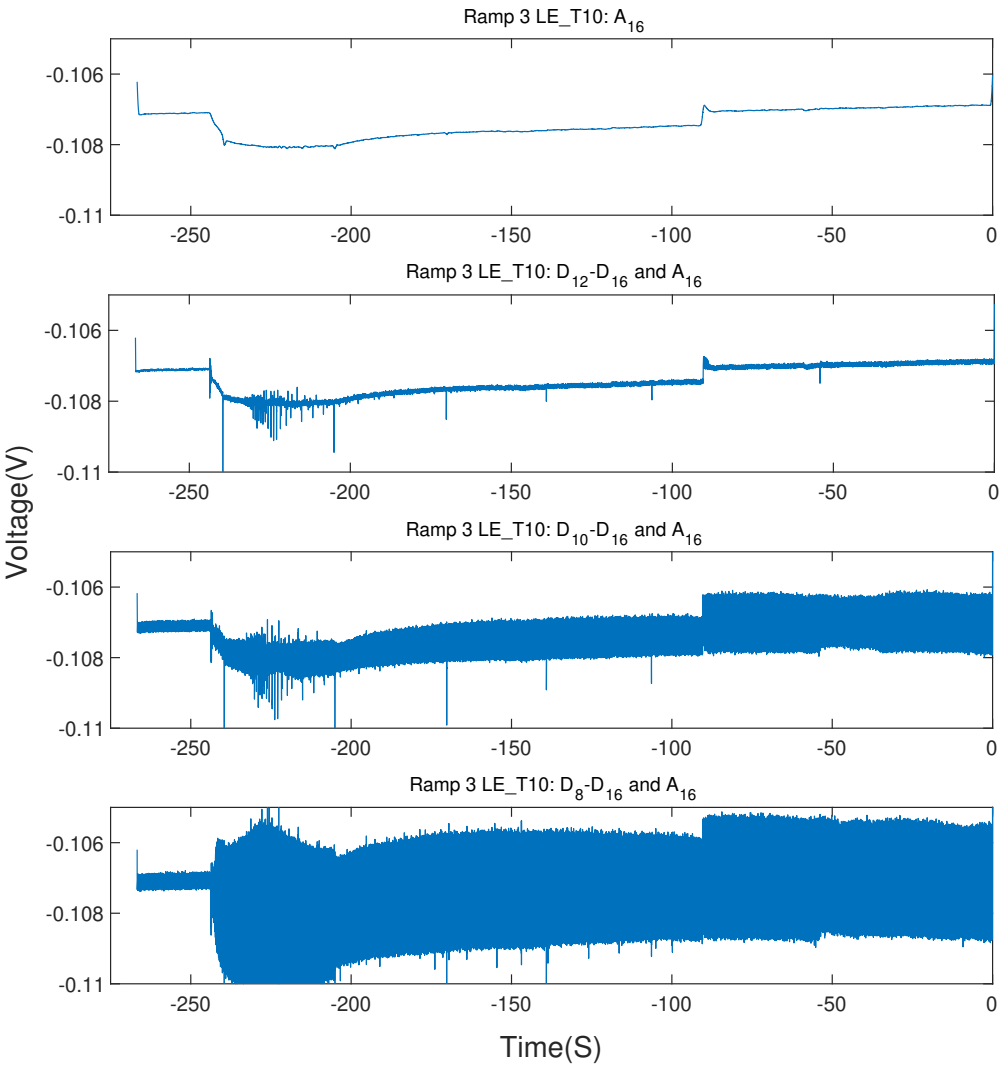

Figure 6: Ramp 3 LE_T10 Channel: The top plot is the trend A_{16}. The second plot is the trend A_{16} and $D_{12} - D_{16}$. The third plot is the trend A_{16} and $D_{10} - D_{16}$, and the fourth plot is the trend with details $D_8 - D_{16}$.

The figure 8 shows the compressed signal. The first plot in the figure 8 is the trend signal A_{16}, the second plot shows the sum of all D_i after applying the above mentioned threshold and the third plot shows the compressed signal i.e. A_{16} and all $D_i's$ added together (after the application of threshold). The harmonics seen in the compressed signal near $t = 0s$ are mainly from $D_2 - D_3$, (high frequency) and those seen near $t = -225s$ are from all levels.

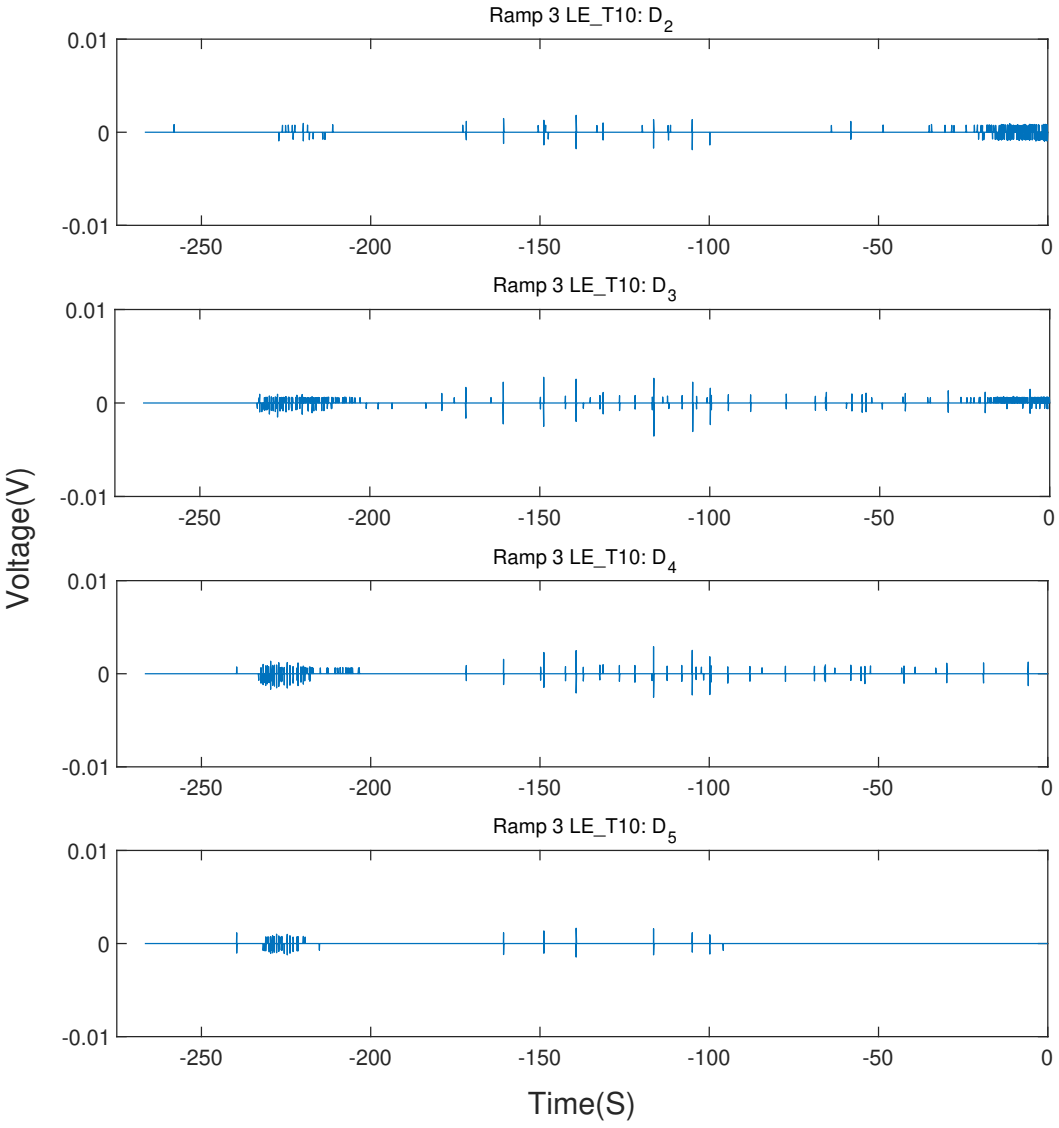

Figure 7: Ramp 3 LE_T10 Channel: D_2, D_3, D_4, and D_5 after applying a hard threshold.

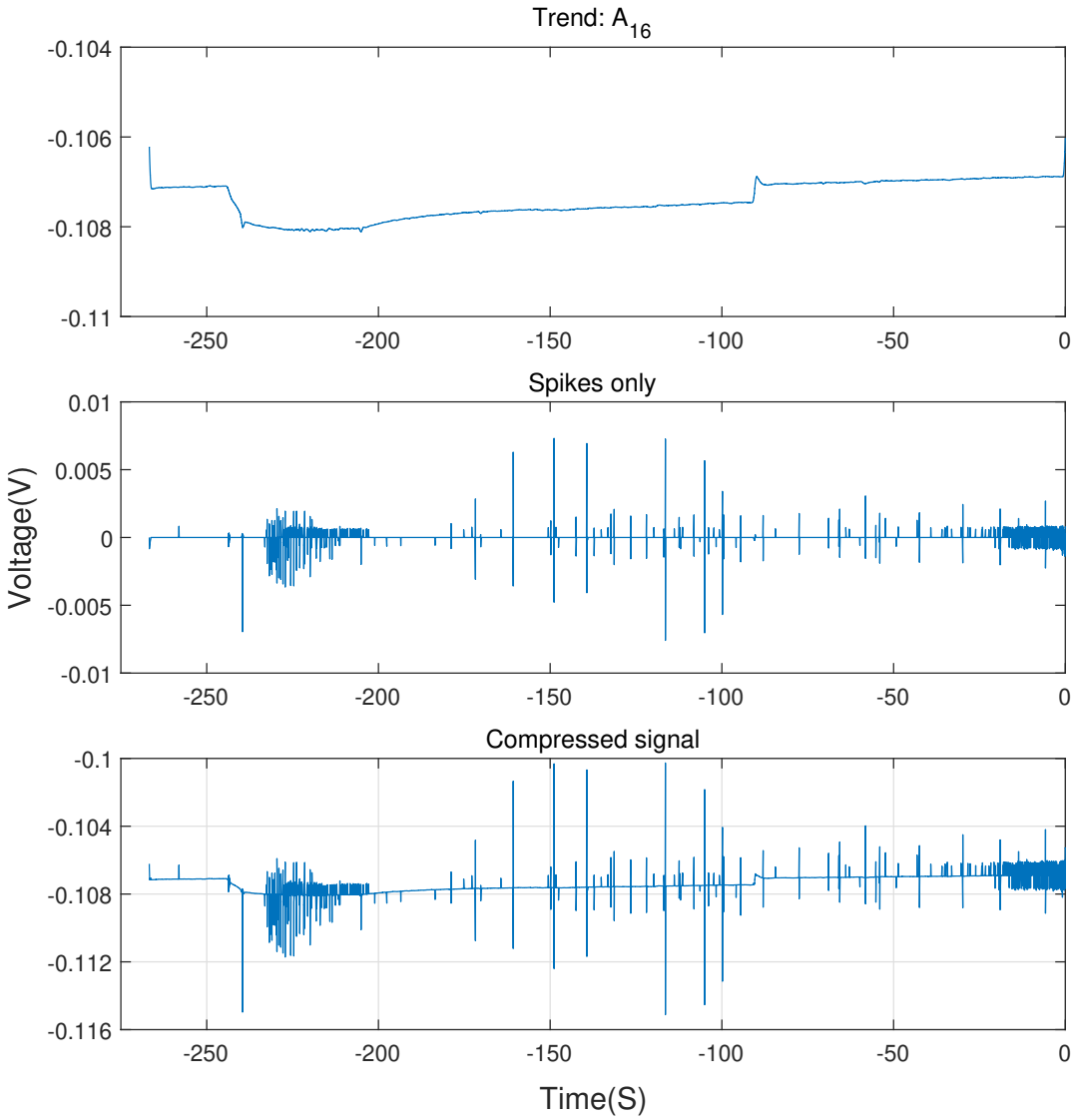

Figure 8: Ramp 3 LE_T10 Channel: Top plot is the trend signal, the second plot is the details only after compression and the third plot is the reconstructed (compressed) signal after applying the threshold.

4 Future work

This work was based on Haar wavelets. The choice of the wavelet is an important question for the future work. The number of decomposition levels chosen is another point to be considered, and handing of the boundary conditions of the data are a few things that need some clarification. The threshold applied for compression is an issue to be investigated too. The behavior of the spikes across all channels and their relation to the quench under consideration is yet to be investigated.

5 Conclusions

The analysis of the quench antenna data is an ongoing research. We analyzed the quench antenna data using the Haar wavelets to understand the anomaly of the data when the quench is detected. A very large signal has been decomposed into several subspaces depending on the frequency bands, thereby allowing us to compress the signal leaving behind some spikes that are present in the signal. This work also supports the earlier work [7] based on single or double moving averages as it provides the trend at each level of decomposition. The wavelet analysis presented here shows that there is noise present in high frequency bands that can be cleaned up and some spikes that are present at all frequency bands are potentially interesting events and it is interesting to see if some spikes are predominantly in specific bands. The analysis also reveals the location of these spikes. These spikes may be related to the quench and may be indicators of the magnet behavior. The result presented here may also be useful for machine learning algorithms.

References

[1] S. Stoynev and J. DiMarco, *Flex-PCB Quench Antenna Developments at FNAL,* IEEE Trans. Appl. Supercond., vol. 32, no. 6, pp. 1-5, Sept. 2022, Art no. 9500205, doi: 10.1109/TASC.2022.3146821.

[2] M. Marchevsky, G. Sabbi, H. Bajas, and S. Gourlay, *Acoustic emission during quench training of superconducting accelerator magnets,* Cryogenics, 49, 2015, 10.1016/j.cryogenics.2015.03.005.

[3] D. Hoang et al., *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, doi: 10.1109/TASC.2021.3058229.

[4] S. Stoynev et al., *Effect of CLIQ on Training of HL-LHC Quadrupole Magnets,* IEEE Transactions on Applied Superconductivity, vol. 34, no. 5, pp. 1-6, Aug. 2024, Art no. 4900606, doi: 10.1109/TASC.2023.3341871.

[5] M. Calvi, L. Agrisani, L. Bottura, A. Masi and A. Siemko, *On the Use of Wavelet Transform for Quench Precursors Characterization in the LHC Superconducting Dipole Magnets,* IEEE Transactions on Applied Superconductivity, vol. 16, no. 2, pp. 1811-1814, 2006, doi: 10.1109/TASC.2006.871308.

[6] S. Stoynev and J. DiMarco, *Assessment and Performance of Flexible Quench Antenna Array Diagnostics for Superconducting Magnets,* IEEE Transactions on Applied Superconductivity(5), 33:1-5, 2023, https://doi.org/10.1109/tasc.2023.3238677

[7] D. Peng, *Analysis of superconducting magnet quench antenna data,* https://inspirehep.net/literature/2808848.

[8] D. B. Percival, and A. T. Walden, *Wavelet Methods for Time Series Analysis,* Cambridge University Press; 2000 vol 4.