

Intelliquench: an adaptive machine learning system for detection of superconducting magnet quenches.

D. Hoang, C. Boffo, N. Tran, S. Krave, S. Kazi, S. Stoynev, and V. Marinozzi.

Abstract—In superconducting magnets, the irreversible transition of a portion of the conductor to resistive state is called a “quench.” Having large stored energy, magnets can be damaged by quenches due to localized heating, high voltage, or large force transients. Unfortunately, current quench protection systems can only detect a quench after it happens, and mitigating risks in Low Temperature Superconducting (LTS) accelerator magnets often requires fast response (down to ms). Additionally, protection of High Temperature Superconducting (HTS) magnets is still suffering from prohibitively slow quench detection. In this study, we lay the groundwork for a quench prediction system using an auto-encoder fully-connected deep neural network. After dynamically trained with data features extracted from acoustic sensors around the magnet, the system detects anomalous events seconds before the quench in most of our data. While the exact nature of the events is under investigation, we show that the system can “forecast” a quench before it happens under magnet training conditions through a randomized experiment. This opens up the way of integrated data processing, potentially leading to faster and better diagnostics and detection of magnet quenches.

Index Terms—quench, quench detection, machine learning, real-time system, superconducting magnets, online learning.

I. INTRODUCTION

Superconducting magnets are used widely in particle accelerators. Because of their ability to carry large current densities with no electrical resistance, they can 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 to ensure normal operation of particle accelerators.

Spontaneous quenches are an integral part of the “training” [1] of superconducting accelerator magnets and can also occur during nominal operations. Quench origins can be attributed to various phenomena [2]. Superconducting accelerator magnets are designed and operated considering their vulnerability to quench effects. Quench protection systems (QPSs) are used to discharge the magnet after a quench happens by dissipating magnetic energy either onto the magnet’s thermal mass or externally, on a dump resistor [3]. Currently, QPSs can only detect the quench *after* it happens, which requires the systems to have very short response time $\mathcal{O}(ms)$. This motivates

the development of quench prediction systems in order to minimize the quench’s damage in superconducting magnets.

Deep Learning (DL) is a sub-field of machine learning that focuses on learning features from data through multiple layers of abstraction [4]. Several DL techniques have been explored for anomaly prediction & detection of superconducting magnet quenches [5]–[7]. However, the techniques are still in their early development phase, and there is yet to be a practical setup and a well-defined procedure to be used for real-time prediction of magnet quenches. Furthermore, magnet quench’s behavior is still not fully-understood, which is why there has been significant uncertainty as to whether “quench precursors” exist, and in what kind of data they would appear. Therefore, in this study, we want to tackle the mentioned problems by investigating a dynamic learning setup using Auto-Encoder (AE) Deep Neural Network (DNN) for anomaly detection in magnets’ acoustic data. In particular, our contributions are the following:

- 1) We propose a *dynamic learning* auto-encoder DNN methodology, which allows us to have a real-time adaptive system for anomaly detection in current ramps to quench data under magnet training conditions.
- 2) Our algorithm detected anomalous events in 77% of the quenches in a randomized experiment, with the events’ time all fall within 15 seconds of the quench. This hints at the existence of quench precursors in magnet’s acoustic signals seconds before the quench.

The codes were developed using the Numpy [8], Matplotlib [9], and Scikit-learn [10] software packages. The DNN models are implemented using Keras [11].

II. BACKGROUND

A. Magnet data

Magnet training is an integral process to produce superconducting accelerator magnets. This happens because stresses in the winding exceed the epoxy’s fracture stress, which causes the magnet to initially quench at a lower level than their design current. After consecutive quenches, friction partially locks the coils in a new and more secure state, allowing the conductors to withstand higher levels of Lorentz forces. Thus, during the training process, a magnet is repeatedly ran at a low current, and then at a slightly increasing current until a natural quench happens. Typically, a magnet is trained until reaching a plateau. In this study, we use acoustic data from two superconducting magnets being tested at Fermilab—MQXFS1d [12]–[15] and MDPCT1b [16], [17]. MQXFS1d is a quadruple magnet that is 1.5m long, and MDPCT1b is a 1m

D. Hoang is with the Department of Physics, Rhodes College, Memphis, TN, 38112 USA (e-mail: minhduc8199@gmail.com)

C. Boffo, N. Tran, S. Krave, S. Stoynev and V. Marinozzi are with Fermi National Accelerator Laboratory, Batavia, IL 60510 USA .

S. Kazi is with the Massachusetts Institute of Technology, Cambridge, MA 02139 USA

FERMILAB-PUB-21-035-CMS-SCD-STUDENT-TD

long dipole magnet. For both magnets training ramps were at 20 A/s close to quench current.

B. Acoustic sensors for magnet diagnostics

Acoustic emission (AE) detection is a non-invasive technique which allows for localization and characterization of mechanical events in superconducting magnets. Thus, it has been widely adapted for non-destructive evaluation of mechanical stability in various engineering and manufacturing settings [18]–[22].

In this study, the available data consisted of acoustic data from sensors placed at the ends of the magnet, as well as data showing the current as a function of time. For the MQXFS1d magnet, two sensors are on one end of the magnet, known as the lead end. The other three sensors are on the other end of the magnet, known as the return end. Sensors setup in MDPCT1b magnet is similar to MQXFS1d’s setup, with only two sensors on two sides of the magnet. MQXFS1d’s data is sampled at a rate of 100 kHz, i.e. one data point every 10 μ s. MDPCT1b’s data was initially sampled at 1MHz, but was then filtered and down-sampled to 100kHz due to high frequency noise. Data were recorded during the whole current ramps to quench for most of the ramps.

C. Deep neural networks as auto-encoders

Deep learning approaches have been applied successfully in several anomaly detection tasks of time series data [23]. In this study, a fully-connected DNN architecture, depicted in Fig. 1, is used as an auto-encoder/decoder [24]. The first half of the DNN takes in input features and transform them into a low-dimensional representation, so-called latent space. The second half of the DNN reconstructs the original inputs by up-sampling the data from a lower-dimensional space.

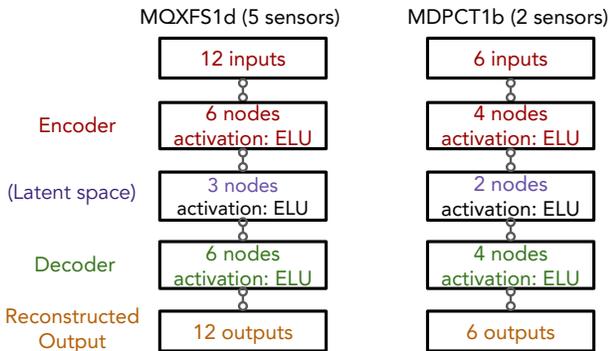


Fig. 1. DNN architecture used on MQXFS1d and MDPCT1b acoustic data in this study. The network architectures usually involve a small number of neurons at the middle of the network, which serve as the latent-space bottle neck. In the latent space the original inputs are compressed into low-dimensional representations. The difference between the number of outputs is due to different number of sensors attached to each magnet.

Eventually, the reconstructed outputs are evaluated by calculating its Root Mean Square Error (RMSE), which is defined as following:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y^{(i)} - \hat{y}^{(i)})^2} \quad (1)$$

where $y^{(i)}$ and $\hat{y}^{(i)}$ are the input features and their predicted counterpart, respectively. The formula is calculated for N data points, which depend on the size of the data set that is used to train and test the model. During the training process, the model is optimized such that the error is minimized. The trained model can then be used for “inference,” during which the model evaluate the error on unseen data points. For the inference process, however, we employ a modified scheme to calculate the error for convenience in triggering as well as in incorporating previous knowledge:

$$L = \log_{10}(L_{\text{inference}}) - \log_{10}(\text{median}(L_{\text{training}})) \quad (2)$$

Where $L_{\text{inference}}$ is the inference error of the data point, and $\text{median}(L_{\text{training}})$ is the median of the training error array for all training data points. As can be seen in Fig.2, when an anomaly happens, the anomalous data points deviate from the “main distribution” of normal data points in the latent space, causing the error to rise significantly. Thus, we can detect these events by setting a threshold of the inference error.

Regarding input features to the model, the process of calculating each sensor’s statistical features are depicted in Fig. 3. Raw acoustic data are first taken from the magnets. The standard deviation and mean of the amplitude are then calculated from each sensor’s raw signal data by using a rolling window with size 20 ms and a step size of 100 μ s. The window’s time label is assigned to the time label of the last data point. This gives 10 streaming features for the MQXFS1D magnet and 4 streaming features for the MDPCT1b magnet. Additionally, in order to accentuate on coherent signal across all the sensors, we also add the product of the standard deviations and the product of the mean of amplitude as input features. Thus, MQXFS1D and MDPCT1b magnet respectively have 12 and 6 streaming input features in total. After all the data are generated, during the training process, all the inputs are scaled to range (0, 1) using Scikit-learn’s MinMaxScaler routine [10]. The same scaling factor is then also applied to testing (unseen) data during the inference process.

Latent space of an anomalous event in MDPCT1b data

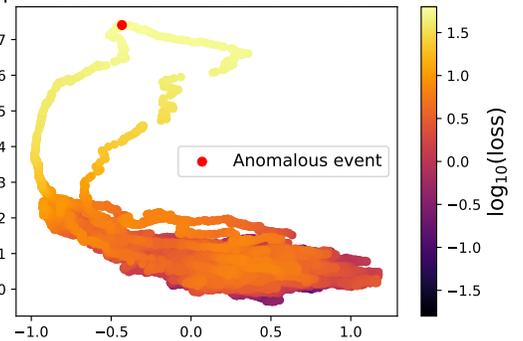


Fig. 2. Latent space data points near an anomalous event in MDPCT1b data. The anomalous data points tend to significantly deviate from the normal data point cluster, causing the inference error (loss) to increase significantly.

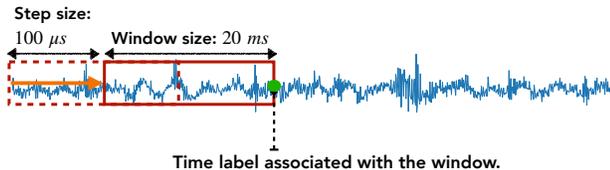


Fig. 3. Sliding window process used in this study. A window with size equivalent to 20 ms in time axis is sequentially moved through the data with steps of 100 μ s. The time label associated with the statistics calculated for each window is assigned using the last data point in the window.

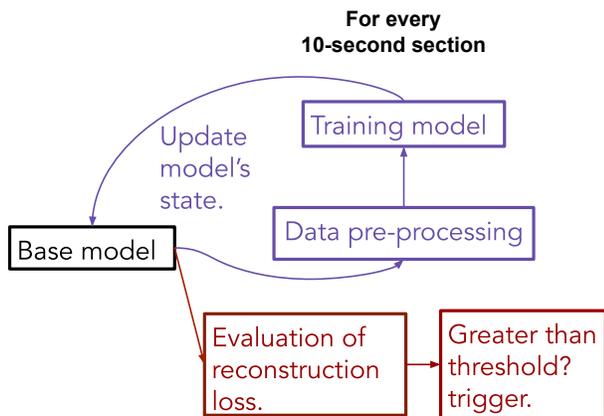


Fig. 4. Dynamic learning workflow used in this study. For every 10-second section of new data, there are two processes happen simultaneously: training and triggering. Training makes the model adaptable with changing noise level, and triggering helps the model to make decisions in real-time.

D. Dynamic learning

Typically, DL model deployment is "static", meaning it is first trained on a specific set of data, and then remains unchanged during the inference process. However, as mentioned in [5], this deployment method makes it hard to set a consistent threshold for all the magnet current ramps, since each ramp has a different scale of the loss. Additionally, acoustic noise level in a magnet tends to increase over time as the magnet reaches higher currents. Thus, a "dynamic" algorithm is thus needed to dynamically adapt to the changing noise level, which in turn also helps highlight anomalous events.

Figure 4 demonstrates our workflow for dynamic learning. Essentially, we want to update the model every 10 seconds, which is the amount of time we estimate to sufficiently train the model in real-time. For each 10-second section of data, there are two simultaneous processes: training and evaluating. During the training process, the data is pre-processed using the procedure described by previous section. Then, the model is trained using the data and a new set of weights is subsequently used to update the base model after the training session. This gives the model the ability to adapt to gradual changes of the noise level during current ramps. Additionally, when the training process is being executed, the evaluation process runs in parallel to make predictions and comparisons of reconstruction loss with the threshold as new data comes in.

In addition to using a dynamic learning algorithm, we also employ a simple threshold setting, where the trigger threshold decreases over time until it reaches the previous quench's

maximum current. There are two main motivations for this. First, the data is taken under magnet training conditions, which means that the current is continuously ramped up. Thus, as oppose to a magnet under normal operating conditions, the superconducting magnet's state is continuously brought closer to its critical surface and the Minimum Quench Energy decreases [25]. As a result, the lower the current, the more "anomalous" the event has to be to quench the magnet. Second, at lower current up until it reaches the previous ramp current (due to "mechanical memory"), the magnet can be very quiet. This implies that a non-significant signal can blow up the reconstruction loss. Therefore, setting a higher threshold at lower current can prevent the system from triggering false-positively. The initial threshold keeps decreasing for some amount every 10 seconds until it stays constant after the maximum current of the previous quench is reached.

III. EXPERIMENTS

After designing the algorithm, we carried out offline experiments using a simulated Python framework. The framework evaluates how the system would theoretically work under real-time conditions during magnet training.

A. Initial observations in MQXFS1d data

In the MQXFS1d magnet, there are acoustic data available for 14 current ramps to quench. It should be noted that the algorithm is only activated after the current reaches 9/10 the maximum current of the previous quench. This is because, in a magnet training context, the quench would not appear at significantly lower current compared with the maximum current of the previous quench.

By setting the initial \log_{10} threshold to 5.0 and the final threshold to be 2.6, the algorithm was able to detect several anomalies seconds before the majority of MQXFS1D quenches. The anomaly points in each quench are presented in Fig.5. It should be pointed out that it is uncertain whether these anomalous events are actually related to the quench. That being said, Table 5, which shows a detailed summary of the events, indicates that the algorithm found the anomalous events within 25 seconds of the quench in 79% of MQXFS1d's quenches. This provides hints to the existence of quench precursors, which have yet to be defined in the literature.

TABLE I
SUMMARY OF DETECTED ANOMALOUS EVENTS IN MQXFS1D'S CURRENT RAMPS.

Characterization of events	Count	Percentage
Within 25 seconds of quench	11	79 %
Entirely outside 25 seconds of quench.	1	7 %
Both inside and outside 25 seconds of quench	1	7 %
No events	1	7 %
Total	14	100 %

Anomaly events in MQXFS1d quenches (\log_{10} threshold 5.0 \rightarrow 2.6)

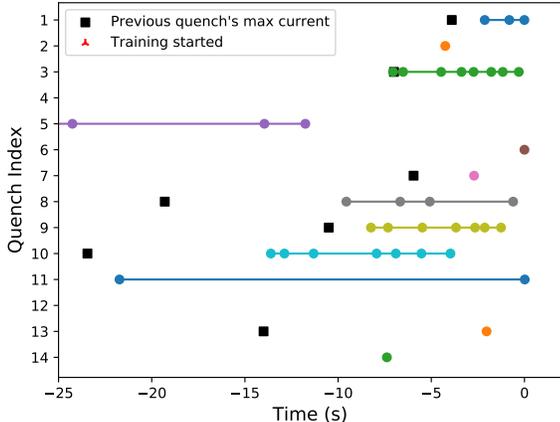


Fig. 5. Anomaly events detected in each of MQXFS1d quenches in the last 25 seconds of the quenches. Majority of anomalous events in MQXFS1d fall within this range. The quench index renumbers current ramps but keeps their order (ramps with no acoustic data are excluded).

B. Verification randomized experiment on MDPCT1b data

To further verify the validity of our algorithm, we also use the algorithm on MDPCT1b data in a more randomized manner with data available for over 50 current ramps. Since threshold setting is the only manually-set parameter in our algorithm, we first randomly picked out 6 quenches (with index 7, 15, 35, 37, 38, and 50) to pick appropriate initial and final thresholds. After appropriate thresholds (\log_{10} threshold of 3.0 and a final threshold 1.6) are chosen, we run the algorithm for the rest of MDPCT1b's quenches and observe the results to verify that we can detect anomalies in this manner. With this threshold setting, anomalous events are detected within 15 seconds of all the randomly selected quenches.

The results on all ramps are depicted in figure 6, which shows that anomalies are detected in 40/52 (77%) of the quenches. Figure 7 further shows the time-to-quench distribution of all the anomalies detected, nearly 100% of which fall within 15 seconds before the quench happens.

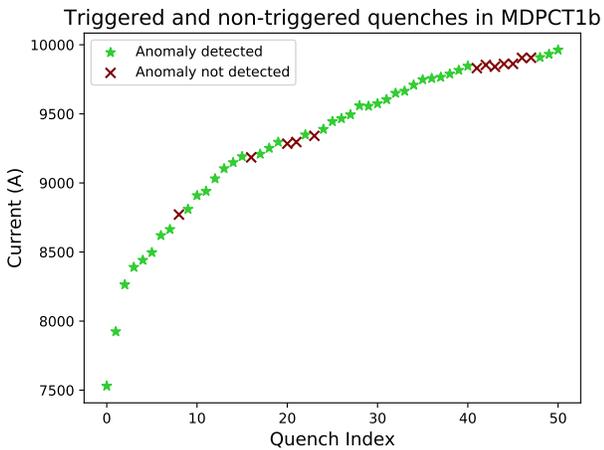


Fig. 6. Anomaly events detected in all of MDPCT1b current ramps [16] using threshold 3.0 and final threshold 1.6 fitted on 6 randomly selected quenches. Anomalous events are detected in the majority (77%) of quenches.

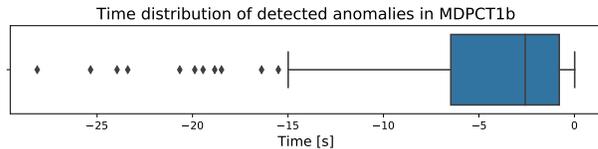


Fig. 7. Time distribution of anomalous events detected in MDPCT1b current ramps. The distribution lies within 15 seconds to the quenches, with many events clustering just within a few seconds of the quench.

IV. CONCLUSION

In this work, we introduced a dynamic learning & threshold procedure using AE-DNN to “predict” quenches under magnet training conditions. Our randomized experiment shows that we can detect anomalies near the quench in a large number of current ramps using the system. This hints at the existence of quench precursors in acoustic data, which then can be utilized in an anomaly detection system such as ours.

As for limitation and future work, the analysis is performed using exclusively magnet training data, due to the rarity of quenches happening under normal operating conditions. That said, studying magnet training quenches is still necessary, as it should bear resemblance to quenches under normal operation. Additionally, while anomalous events are detected near the quench, its exact nature and relationship with the quench are not yet determined. Another limitation is that all the results in this study are produced using a simulated Python framework, which might not perfectly reproduce a real-time scenario. It should also be noted that there are other types of data that can be useful. However, acoustics was the only type taken at high rate (100 kHz or more), and for the whole magnet ramp. Other types of data are available for this study, but at much lower ramp rate (at most 8 kHz) and for limited time (± 1 s around quench detection time). Therefore, they were excluded from this study.

In future work, we plan to better characterize the anomalous events we detected to determine its relationship with the quench. Future more diverse data types such as voltages at high DAQ rate will help to better characterize events, which would also help us to propose useful expert features for the AE-DNN model. It would also be interesting to implement the system on hardware to test its performance under real-life scenarios. Finally, more sophisticated signal-processing deep learning models [26], [27], albeit being more computationally expensive for real-time applications, are also worth being explored as well.

ACKNOWLEDGEMENTS

The magnet data in this paper were made available to us by the LARP, AUP and HL-LHC teams, and the FNAL US-MDP teams. D.Hoang and S.Kazi are supported by the SIST and SULI internships sponsored 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.

REFERENCES

- [1] P. Seidel, Ed., *Applied superconductivity Handbook on devices and applications*. Germany: Wiley-VCH, 2015.
- [2] A. Devred, “Quench origins,” *AIP Conference Proceedings*, vol. 249, no. 2, pp. 1262–1308, 1992. [Online]. Available: <https://aip.scitation.org/doi/abs/10.1063/1.41993>
- [3] R. Denz, “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] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. [Online]. Available: <https://doi.org/10.1038/nature14539>
- [5] 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, pp. 40–50, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S016890021730668X>
- [6] M. Wielgosz, A. Skoczeń, and E. De Matteis, “Protection of superconducting industrial machinery using rnn-based anomaly detection for implementation in smart sensor,” *Sensors*, vol. 18, no. 11, p. 3933, Nov 2018. [Online]. Available: <http://dx.doi.org/10.3390/s18113933>
- [7] A. S. Nawaz, S. Pfeiffer, G. Lichtenberg, and H. Schlarb, “Self-organized critical control for the european xfel using black box parameter identification for the quench detection system,” in *2016 3rd Conference on Control and Fault-Tolerant Systems (SysTol)*, 2016, pp. 196–201.
- [8] C. R. Harris *et al.*, “Array programming with numpy,” *Nature*, vol. 585, no. 7825, pp. 357–362, 2020. [Online]. Available: <https://doi.org/10.1038/s41586-020-2649-2>
- [9] J. D. Hunter, “Matplotlib: A 2d graphics environment,” *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.
- [10] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [11] F. Chollet *et al.* (2015) Keras. [Online]. Available: <https://github.com/fchollet/keras>
- [12] P. Ferracin *et al.*, “The hl-lhc low- β quadrupole magnet mqxf: From short models to long prototypes,” *IEEE Transactions on Applied Superconductivity*, vol. 29, no. 5, pp. 1–9, 2019.
- [13] H. Bajas *et al.*, “Test result of the short models mqxf3 and mqxf5 for the hl-lhc upgrade,” *IEEE Transactions on Applied Superconductivity*, vol. 28, no. 3, pp. 1–6, 2018.
- [14] M. Anerella *et al.*, “MQXFS1 Quadrupole Fabrication Report,” 7 2017.
- [15] M. Anerella *et al.*, “MQXFS1 Quadrupole Design Report,” 1 2016.
- [16] A. Zlobin *et al.*, “Reassembly and test of high-field nb3sn dipole demonstrator mdpct1,” *Submitted for publication*, 2020.
- [17] A. Zlobin *et al.*, “Development and first test of the 15 t nb3sn dipole demonstrator mdpct1,” *IEEE Transactions on Applied Superconductivity*, vol. PP, pp. 1–1, 02 2020.
- [18] O. Tsukamoto *et al.*, “Identification of quench origins in a superconductor with acoustic emission and voltage measurements,” *Applied Physics Letters*, vol. 39, no. 2, pp. 172–174, Jul. 1981.
- [19] O. Tsukamoto and Y. Iwasa, “Acoustic emission triangulation of disturbances and quenches in a superconductor and a superconducting magnet,” *Applied Physics Letters*, vol. 40, no. 6, pp. 538–540, 1982. [Online]. Available: <https://doi.org/10.1063/1.93135>
- [20] O. Tsukamoto and Y. Iwasa, “Sources of acoustic emission in superconducting magnets,” *Journal of Applied Physics*, vol. 54, no. 2, pp. 997–1007, 1983. [Online]. Available: <https://doi.org/10.1063/1.332027>
- [21] M. Marchevsky and others., “Acoustic emission during quench training of superconducting accelerator magnets,” *Cryogenics*, vol. 69, pp. 50–57, 2015.
- [22] M. Pappé, “Discussion of acoustic emission of a superconducting solenoid,” *IEEE Transactions on Magnetics*, vol. 19, no. 3, pp. 1086–1089, May 1983.
- [23] C. Zhang *et al.*, “A deep neural network for unsupervised anomaly detection and diagnosis in multivariate time series data,” *CoRR*, vol. abs/1811.08055, 2018. [Online]. Available: <http://arxiv.org/abs/1811.08055>
- [24] P. Baldi, “Autoencoders, unsupervised learning, and deep architectures,” in *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, ser. Proceedings of Machine Learning Research, I. Guyon, G. Dror, V. Lemaire, G. Taylor, and D. Silver, Eds., vol. 27. Bellevue, Washington, USA: PMLR, 02 Jul 2012, pp. 37–49. [Online]. Available: <http://proceedings.mlr.press/v27/baldi12a.html>
- [25] M. N. Wilson and R. Wolf, “Calculation of minimum quench energies in rutherford cables,” *IEEE Transactions on Applied Superconductivity*, vol. 7, no. 2, pp. 950–953, 1997.
- [26] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, p. 1735, 1997. [Online]. Available: <http://relay.rhodes.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,url,uid,cookie&db=pbh&AN=9710215021&site=ehost-live>
- [27] H. Lim, J. Park, and Y. Han, “Rare sound event detection using 1d convolutional recurrent neural networks,” in *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2017 Workshop*, 2017, pp. 80–84.