

Efficient and scalable detection of acoustic emission multiplets for fatigue crack growth monitoring

Alessandra Panerai¹ Andrea Bernasconi¹ and Michele Carboni¹

Department of Mechanical Engineering, Politecnico di Milano
Via la Masa 1, 20156 Milano, Italy
e-mail : alessandra.panerai@polimi.it

Abstract

The presence of repeating self-similar acoustic emission signals, known as *multiplets*, has been linked to the phenomenon of fatigue crack growth, making their detection an extremely useful way to assess damage. However, the high computational cost and memory requirements of existing multiplet detection algorithms restrict their applicability to only relatively small datasets, posing a strong limitation to the exploitation of multiplet detection to fatigue monitoring.

This work proposes two complementary strategies to increase the computational efficiency of multiplet detection algorithms. It is shown that they can be easily implemented with no loss to multiplet detection accuracy. The strategies are then applied to a large fatigue dataset consisting of more than one million acoustic emission signals.

Overall, the proposed strategies enable efficient and scalable multiplet detection, extending its applicability to large-scale and potentially real-time fatigue monitoring scenarios.

Keywords: Acoustic emission multiplets, Fatigue, Composites, Adhesive bonded joints, Clustering algorithms

1. Introduction

The initiation and progressive growth of cracks in structural components under cyclic loads remains one of the leading causes of failure in many industries, for which effective Structural Health Monitoring techniques are sorely needed. Acoustic Emission (AE) is a good candidate for monitoring fatigue crack growth rate, as the release of elastic energy associated with crack advancement generates transient elastic waves that can be detected by surface-mounted sensors [1]. However, reliably quantifying damage based on AE activity remains challenging, since a wide range of other mechanisms, as well as environmental and operational noise, can also produce AE signals [2].

Correctly identifying which signals are caused by fatigue crack growth is challenging: signals may be clustered based on various time- or frequency-domain characteristics (amplitude, peak frequency, duration, and more) [3], but assigning physical meaning to these clusters typically requires a pre-existing library of labeled signals [4]. Moreover, the features of damage-related AE signals, including those associated with fatigue crack growth, depend not only on the material, but also on propagation path, sensor location and coupling conditions – as a result, feature based characterization of damage signals is often difficult to translate from one configuration to another [5,6].

Recent developments [7–9] have uncovered a powerful tool for monitoring fatigue crack growth: the detection of the so-called *acoustic multiplets*. Multiplets are groups of repeated signals whose waveforms are nearly identical to one another, implying that they share the same source mechanism and approximately the same location. It has been shown [7,10] that the presence of multiplets is a signature of the incremental advancement of fatigue cracks, where signals are triggered at each loading cycle. The fact that multiplets can be identified not by specific features in the signals, but by the similarities between the signals within a single test means that fatigue crack growth can be detected without any prior knowledge about the characteristics of damage signals. In addition, the self-similarity of multiplet signals can be exploited to improve localization [7], further increasing monitoring performance.



Despite these advantages, multiplet detection remains computationally demanding. Though some strategies have been introduced to increase computational efficiency [8], both the computational cost and memory requirements of current multiplet detection algorithms remain prohibitive for large datasets, which are typical of fatigue tests.

In this work, two strategies are proposed to substantially reduce the computational cost and memory requirements of multiplet detection algorithms. The proposed approaches are straightforward to implement and do not compromise multiplet detection accuracy. Their effectiveness is demonstrated on a large fatigue dataset comprising more than one million AE signals, a scale that would have been impractical to analyze using conventional multiplet detection methods.

2. Multiplet detection algorithm and proposed efficiency strategies

2.1. Conventional multiplet detection algorithm

Multiplet detection is based on the identification of clusters of acoustic emission signals exhibiting a high degree of waveform similarity. In the approach originally proposed by de la Selle et al. [8], schematically shown in Figure 1, the similarity between two signals is quantified by cross-correlation, from which a *dissimilarity* metric is defined, which takes values between 0 (for identical signals) and 1 (for totally dissimilar signals). Thus, a $N \times N$ dissimilarity matrix is constructed.

The clustering is then performed by DBSCAN (Density Based Spatial Clustering of Applications with Noise), an unsupervised algorithm which takes as input the dissimilarity matrix, representing the distances between signals, and two parameters, the maximum allowed dissimilarity, *epsilon*, between signals belonging to the same multiplet, and the minimum number of signals *MinPts* required to form a cluster. Signals assigned to a cluster are identified as part of a multiplet, and therefore classified as being due to fatigue crack growth, while unclustered signals are considered to be due to noise or other damage phenomena. Thus, the detection of damage signals is done based on waveform similarity only, not predefined signal features.

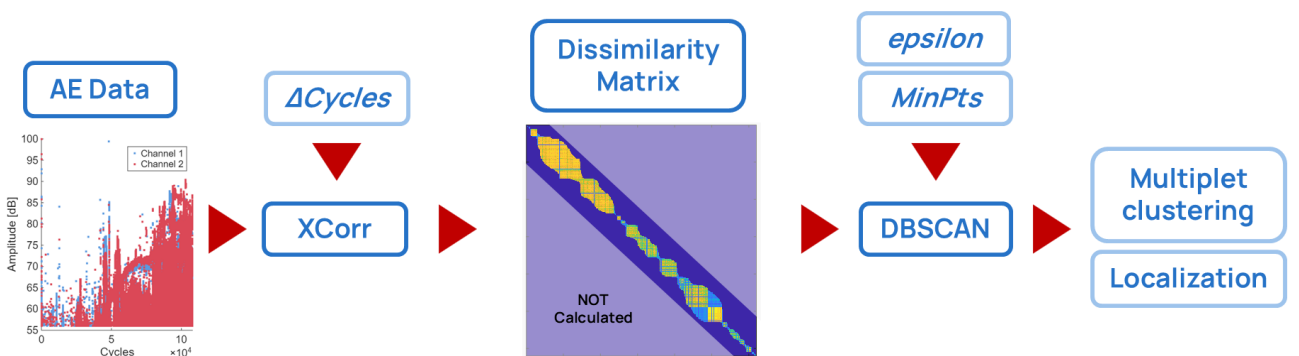


Figure 1 - Multiplet detection algorithm.

The main limitation of this approach lies in the construction and handling of the dissimilarity matrix. In its most general form, the number of required cross-correlations scales as $O(N^2)$, quickly becoming prohibitive for large datasets. To mitigate this issue, de la Selle et al. [8,11] introduced a time-based filter, whereby each signal is compared only to signals occurring within a limited time window, here expressed as a corresponding number of fatigue cycles ΔCycles . All other

dissimilarity values are set to 1. While this approach significantly reduces the number of required calculations, both computational cost and memory requirements remain high.

2.2. Amplitude-based filtering

To further reduce the computational cost, it is here proposed to introduce, alongside the time-based ΔCycles filtering, an additional amplitude-based filter. When acoustic emission activity is intense, the number of signals within a given time window can be very high. Under these conditions, time-based filtering alone may still result in a large number of cross-correlations. However, ΔCycles usually has as a lower limit, as multiplets may have silent zones of a few cycles. Thus, it is needed to reduce the number of comparisons made without reducing ΔCycles .

In this approach, each signal is compared only with other signals whose peak amplitudes fall within a prescribed range around its own amplitude, as shown in Figure 2. Since signals belonging to the same multiplet are generated by the same source mechanism and propagate along similar paths, they are expected to exhibit not only similar waveforms, but also comparable amplitudes. Amplitude-based filtering therefore restricts the similarity search to signals that are more likely to belong to the same multiplet while excluding most likely irrelevant comparisons.

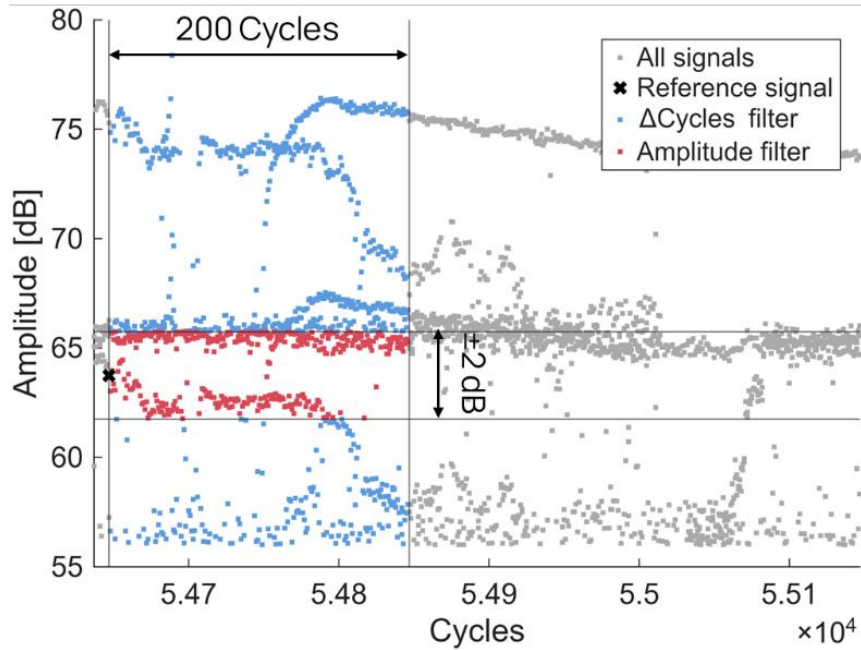


Figure 2 - Example application of amplitude-based filtering.

Other signal features, such as energy and duration, were also considered for this purpose. However, these quantities were found to be more sensitive to background noise and acquisition conditions and did not provide improved performance compared to peak amplitude.

This strategy directly reduces the number of cross-correlations to be computed when applied in combination with the time-based ΔCycles restriction. Its impact on computational cost is assessed quantitatively in Section 4.

2.3. Block-wise clustering

Even with time-based and amplitude-based restrictions, storing and handling the dissimilarity matrix for very large datasets can easily exceed the memory capabilities of standard computing environments. To address this limitation, a block-wise clustering strategy is here proposed.

In this approach, schematically shown in Figure 3, the full dataset is divided into consecutive blocks of a fixed time duration. For each block, the dissimilarity matrix is computed only for the signals contained within that block, and DBSCAN clustering is performed independently. This strategy is not aimed at reducing the total number of cross-correlations, but rather at limiting the maximum size of the dissimilarity matrix that must be stored and processed at any given time. As such, it is particularly well suited to long-duration fatigue tests and to real-time monitoring applications.

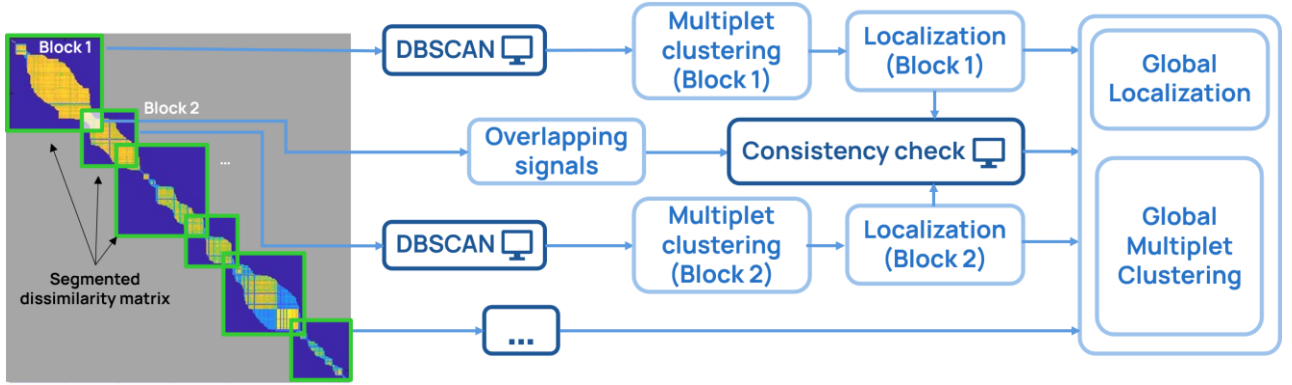


Figure 3 - Proposed block-wise clustering algorithm.

Within each block, both time-based and amplitude-based filters may be employed to reduce computational time. To avoid the loss of information at block boundaries, adjacent blocks are overlapped by a duration equal to ΔCycles . As shown in Figure 3, a signal within an overlap is clustered twice, and so may be assigned to clusters in both blocks. To ensure consistency, the average waveform of each cluster is stored and used to compare cluster assignments across overlapping blocks. The following rules are applied for signals in the overlap region:

1. If it is not assigned to any cluster in either block, it is retained as unclustered.
2. If it is assigned to a cluster in only one of the two blocks, it is assigned to that cluster.
3. If it is assigned to clusters in both blocks and the corresponding cluster averages have a dissimilarity below epsilon, the signal is arbitrarily assigned to the cluster in the second block.
4. If it is assigned to clusters in both blocks and the corresponding cluster averages have a dissimilarity exceeding epsilon, the signal is considered ambiguous and removed from both clusters.

It should be noted that rule three is applied in the case when the two clusters belong to the same multiplet which continues from one block to the other. It would be wholly possible to instead coalesce the two clusters, allowing clusters to span several blocks. While this would better reflect the actual multiplet behaviour, it was chosen in the present work to keep the clusters intentionally small in order to have a high degree of waveform similarity within each cluster. Indeed, multiplets evolve over time as crack growth progresses and the source location shifts, so that signals at the beginning and end of long multiplet can be quite different [7]. Maintaining smaller, more homogeneous clusters is preferable when multiplets similarity is exploited to improve time-of-arrival estimation for source localization, as described in [7] – nevertheless, if desired, block-wise clustering can be used to detect longer multiplets by modifying rule three accordingly.

3. Experimental setup

The proposed strategies were validated using an acoustic emission dataset acquired during fatigue testing of an adhesively bonded composite joint, originally presented in [12]. The specimen consisted of a single lap joint, shown in Figure 4a, tested in a uni-axial MTS810 servo-hydraulic testing machine with a 100 kN capacity. The test was conducted at a loading frequency of 5Hz and a load ratio $R=0.05$. The joint consisted of twill woven carbon fiber laminate adherends bonded with 3M 9323 A/B adhesive, as shown in Figure 4b [13]. The load level was selected to ensure an expected fatigue life of at least 500,000 cycles, and the test was carried out for 100,000 cycles. Due to the joint geometry, crack growth generally occurs from both sides of the overlap region, as illustrated in Figure 4c and e.

Acoustic emission monitoring was performed using two Vallen Systeme VS150-M resonant piezoelectric sensors with a nominal resonant frequency of 150 kHz, equipped with AEP5 34 dB preamplifiers and connected to the Vallen Systeme multi-channel system AMSY-6. The sample rate was set to 5 MHz and the acquisition threshold to 56 dB. The resulting AE activity recorded during the test is shown in Figure 4d. A total of $N = 1,082,495$ acoustic emission signals were acquired and used for the subsequent analysis.

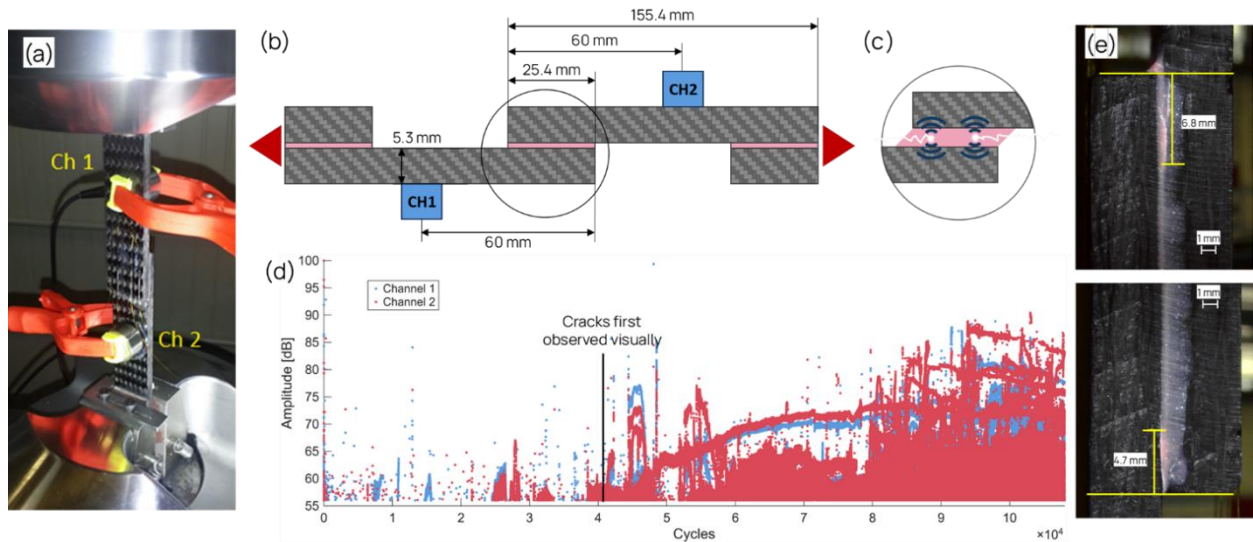


Figure 4 - (a) Fatigue testing setup; (b) Specimen Geometry; (c) detail of (b) showing crack growth; (d) recorded AE raw data; (e) Crack growth at the end of the test.

4. Results and discussion

4.1. Computational efficiency and validation on a reduced dataset

To assess the impact of the proposed strategies on both computational efficiency and clustering quality, a reduced portion of the dataset corresponding to 500 loading cycles (5832 AE signals) was selected. This subset was analyzed using four different approaches:

1. Conventional clustering using only the time-based ΔCycles restriction,
2. Clustering with amplitude-based filtering,
3. Block-wise clustering without amplitude filtering,
4. Block-wise clustering combined with amplitude-based filtering.

The parameters used for each method are summarized in Table 1. The values of ΔCycles , epsilon and MinPts were kept constant and were set to 50, 0.07 and 5, respectively. For methods 2 and 4, the amplitude filter was set to ± 2 dB. For the block-wise approaches (methods 3 and 4), a block size of 150 cycles and an overlap of 50 cycles – matching ΔCycles – were employed.

Table 1 also reports the total computational time required for each method, including both dissimilarity matrix computation and DBSCAN clustering. Using either amplitude filtering or block-wise clustering alone reduced computational time by approximately a factor of four, by decreasing the number of required cross-correlations or limiting memory usage, respectively. When both strategies were combined, a dramatic reduction in computational time was achieved, with clustering completed about 40x faster than with the conventional approach.

Method	ΔCycles	Amplitude threshold	Block size	Computational time	N° of clusters
Conventional clustering	50	n/a	n/a	482 s	8
Amplitude filtering	50	± 2 dB	n/a	108 s	11
Block-wise clustering	50	n/a	150 cycles	159 s	44
Block-wise with amp. filtering	50	± 2 dB	150 cycles	12 s	44

Table 1 - Comparison of clustering methods

As expected (see section 2.3), block-wise clustering leads to an increase in the number of detected clusters, as longer multiplets are divided into multiple shorter clusters. This behavior is illustrated in Figure 5, which compares the results obtained using the conventional approach (method 1) and the combined block-wise and amplitude-filtered strategy (method 4). While multiplets identified by the conventional approach appear as long cluster spanning many cycles (Figure 5a), the block-wise splits these into several shorter clusters (Figure 5b). This, however, does not affect multiplet detection. Figure 5c and d show a representative cluster identified by method 1 and the corresponding four clusters obtained by method 4, respectively. Thanks to amplitude-based filtering, outliers visible in the conventional clustering are excluded from the block-wise clusters. Importantly, all clustered signals are consistently localized within the bonded region of the joint, where fatigue crack propagation is known to occur (Figure 5e and f).

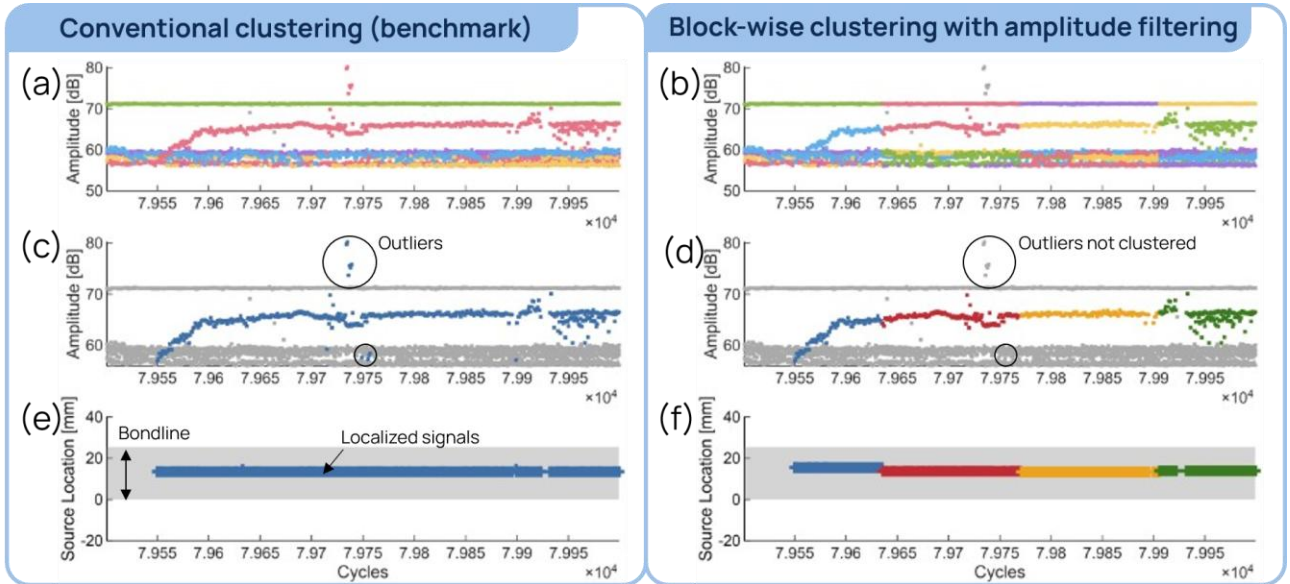


Figure 5 - Comparison of conventional clustering and block-wise clustering with amplitude filtering: (a),(b) clustering; (c) representative cluster obtained by conventional clustering; (d) corresponding clusters obtained by block-wise clustering; (e),(f) comparison of source localization of the clusters shown in (c) and (d), respectively.

These results indicate that the proposed strategies significantly improve computational efficiency without compromising the identification of multiplets.

4.2. Application to large-scale fatigue crack growth monitoring

Having verified that the proposed strategies significantly reduce computational cost without compromising the identification of multiplets, the combined approach was then applied to the entire fatigue dataset of over 1 million signals. The parameters employed were kept identical to those reported in Table 1. The total computational time required to cluster the entire dataset was approximately 300 minutes.

After clustering, source localization was performed both on the entire dataset, and on multiplet signals. The method proposed in [7] was employed for determining the time-of-arrival for multiplet localization. Figure 6 shows the localization results obtained by both approaches. The point at which crack initiation was first visually observed (cycle 41,000) in the specimen is also reported. Aside from a small cluster of signals at around 25,000 cycles, sustained multiplet activity begins at the onset of visible crack initiation, strengthening the association of multiplet presence to fatigue crack growth.

While traditional localization is highly scattered, especially in the later stages of the test, the great majority of multiplet signals is localized within the adhesive bond. Due to the heterogeneous nature of composite substrates, local variations in sound speed, and inherent measurement uncertainties, neither localization approach was accurate enough to provide a meaningful measure of the crack length. Nevertheless, multiplet detection and localization was able to both detect the presence of fatigue crack growth and locate the affected area, which could subsequently be targeted for inspection using other methods. These results demonstrate that efficient multiplet detection can be applied to large-scale fatigue monitoring scenarios that were previously computationally prohibitive.

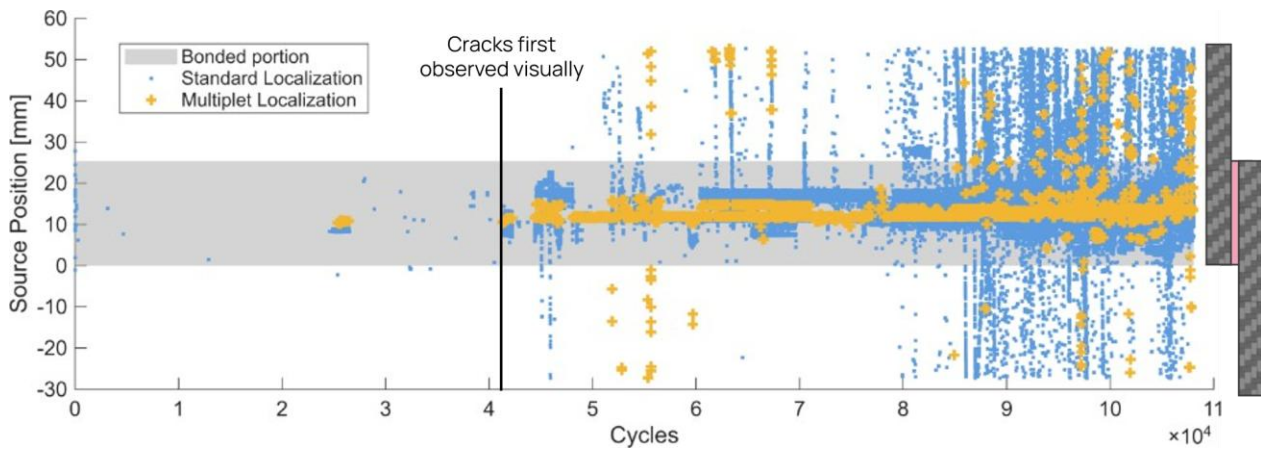


Figure 6 - Comparison of source localization results obtained using all acoustic emission signals and using only multiplet signals.

5. Conclusions

This work addressed the applicability of multiplet detection algorithms to large acoustic emission datasets and proposed two complementary strategies to tackle both the computational cost and memory requirements. Validation on a small dataset showed that their combined approach produced an order of magnitude improvement compared to state-of-the-art algorithms with no loss in clustering accuracy. The combined approach was successfully applied to a large fatigue dataset comprising more than one million acoustic emission signals, a scale that would have been computationally infeasible using conventional methods. Results showed that multiplet-based localization provides a more robust indication of the damaged region than localization performed on the full acoustic emission dataset. Overall, the proposed strategies enable efficient and scalable multiplet detection, extending their applicability to large-scale and potentially real-time fatigue monitoring scenarios.

Acknowledgements

The authors gratefully acknowledge Mr. A. Grossi for his active help with the experimental work.

References

- [1] Saeedifar M, Zarouchas D. Damage characterization of laminated composites using acoustic emission: A review. *Compos B Eng* 2020;195:108039. <https://doi.org/10.1016/J.COMPOSITESB.2020.108039>.
- [2] Masmoudi S, El Mahi A, Turki S, El Guerjouma R. Mechanical behavior and health monitoring by Acoustic Emission of unidirectional and cross-ply laminates integrated by piezoelectric implant. *Applied Acoustics* 2014;86:118–25. <https://doi.org/10.1016/J.APACOUST.2014.04.011>.
- [3] Roundi W, El Mahi A, El Gharad A, Rebiere JL. Acoustic emission monitoring of damage progression in Glass/Epoxy composites during static and fatigue tensile tests. *Applied Acoustics* 2018;132:124–34. <https://doi.org/10.1016/J.APACOUST.2017.11.017>.
- [4] Godin N, Huguet S, Gaertner R, Salmon L. Clustering of acoustic emission signals collected during tensile tests on unidirectional glass/polyester composite using supervised and unsupervised classifiers. *NDT & E International* 2004;37:253–64. <https://doi.org/10.1016/J.NDTEINT.2003.09.010>.
- [5] Ni QQ, Iwamoto M. Wavelet transform of acoustic emission signals in failure of model composites. *Eng Fract Mech* 2002;69:717–28. [https://doi.org/10.1016/S0013-7944\(01\)00105-9](https://doi.org/10.1016/S0013-7944(01)00105-9).
- [6] Oz FE, Ersoy N, Lomov S V. Do high frequency acoustic emission events always represent fibre failure in CFRP laminates? *Compos Part A Appl Sci Manuf* 2017;103:230–5. <https://doi.org/10.1016/J.COMPOSITESA.2017.10.013>.

- [7] Deschanel S, Ben Rhouma W, Weiss J. Acoustic emission multiplets as early warnings of fatigue failure in metallic materials. *Sci Rep* 2017;7:1–10. <https://doi.org/10.1038/S41598-017-13226-1>;TECHMETA.
- [8] de la Selle T, Weiss J, Deschanel S. Acoustic multiplets detection based on DBSCAN and cross-correlation. *Mech Syst Signal Process* 2024;211:111149. <https://doi.org/10.1016/J.YMSSP.2024.111149>.
- [9] de la Selle T, Réthoré J, Weiss J, Lachambre J, Deschanel S. Signatures of fatigue crack growth from acoustic emission repeaters. *Eng Fract Mech* 2024;110388. <https://doi.org/10.1016/J.ENGFRACMECH.2024.110388>.
- [10] De La Selle T, Deschanel S, Weiss J, Lachambre J, Réthoré J, De La Selle T, et al. Detection and characterization of fatigue crack growth by acoustic emission and digital image correlation 2022;255. <https://doi.org/10.34894/VQ1DJA>.
- [11] De La Selle T, Deschanel S, Weiss J. ACOUSTIC EMISSION CLASSIFICATION BY CROSS-CORRELATION 2023. <https://doi.org/10.58286/27633>.
- [12] Bernasconi A, Martulli LM, Carboni M. Fatigue crack growth analysis in composite bonded joints by back face distributed strain sensing and comparison with X-ray microtomography. *Int J Fatigue* 2022;154:106526. <https://doi.org/10.1016/J.IJFATIGUE.2021.106526>.
- [13] 3MTM Scotch-WeldTM Structural Epoxy Adhesive EC-9323 B/A 2022. https://multimedia.3m.com/mws/media/2157921O?&fn=EC-9323-BA-Structural-Epoxy-Paste-Adhesive_TDS_Feb22_ref259_R2.pdf (accessed February 2, 2023).