

Deinterleaving RADAR Emitters With Optimal Transport Distances

MANON MOTTIER 

GILLES CHARDON 

FRÉDÉRIC PASCAL , Senior Member, IEEE

Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systèmes, Gif-sur-Yvette, France

Detection and identification of emitters provide vital information for defensive strategies in electronic intelligence. Based on a received signal containing pulses from an unknown number of emitters, this article introduces an unsupervised methodology for deinterleaving RADAR signals based on a combination of clustering algorithms and optimal transport distances. The first step involves separating the pulses with a clustering algorithm under the constraint that the pulses of two different emitters cannot belong to the same cluster. Then, as the emitters exhibit complex behavior and can be represented by several clusters, we propose a hierarchical clustering algorithm based on an optimal transport distance to merge these clusters. A variant is also developed, capable of handling more complex signals. Finally, the proposed methodology is evaluated on simulated data provided through a realistic simulator. Results show that the proposed methods are capable of deinterleaving complex RADAR signals.

Manuscript received 8 September 2023; revised 20 December 2023; accepted 9 February 2024.

DOI. No. 10.1109/TAES.2024.3367287

Refereeing of this contribution was handled by M. Y. Lanzerotti.

This work was supported by ATOS, which contributed both data and RADAR expertise.

Authors' address: Manon Mottier, Gilles Chardon, and Frédéric Pascal are with the Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systèmes, 91190 Gif-sur-Yvette, France, E-mail: (manon.mottier@centralesupelec.fr; gilles.chardon@centralesupelec.fr; frederic.pascal@centralesupelec.fr). (*Corresponding author: Manon Mottier.*)

0018-9251 © 2024 IEEE

I. INTRODUCTION

In a passive listening system [1], [2], [3], [4], [5], an electronic support measure system intercepts signals containing interleaved time domain pulses from an unknown number of emitters. Pulse description words (PDW) represent pulses containing various parameters, also called features, such as frequency, pulse width, angle of arrival, etc. Over the years, many methods have been developed to deinterleave a RADAR signal, i.e., to separate the pulses from different emitters. At first, the systems were less sophisticated than now; simple and easily separable characteristics represented emitters as transmitted continuously on a single frequency (little noise, few adverse countermeasures). Early work from the 80s mainly used frequency, direction of arrival, and time of arrival (TOA) to deinterleave RADAR signals [3]. After filtering the frequency and the direction of arrival, the TOA is used to find the pulse repetition interval (PRI). The PRI corresponds to the difference in emission between two successive pulses and is usually different between RADAR emitters. This work has been a foundation for many authors who have contributed to enriching RADAR literature. Most of the following methods are based on the search of patterns in the PRI, for example, by using the cumulative difference histogram [6] or the sequential difference histogram [7].

Searching for the PRI pattern is straightforward for simple RADAR emitters because they transmit continuously with a single PRI. Nowadays, emitters transmit less regularly and can have characteristics of agility such as emitting on several frequencies with one or more PRIs. Some methods have adapted to cope with these changes [8]. Deep learning models have been increasingly requested to manage the sophistication and emission complexity of emitters [9], [10], [11], [12], [13], [14], [15], [16], [17]. Deep Learning models require considerable data and are difficult to parameterize. These methods are primarily built in a supervised framework, and their results strongly depend on the quality of the simulator used. Most of these methods are based on processing the PRI data. When the signal is poorly estimated, it is possible to have missing pulses, which can lead to a poor reconstruction of the PRI and, therefore, failure to regroup the pulses correctly. Indeed, due to the complexity of the problem, supervised approaches, such as deep learning techniques, generally failed to provide good deinterleaving performance. However, methods based on clustering pulse features are based on models too simple to deal with the variety and complexity of RADAR signals [18], [19], [20]. Indeed, these methods cannot identify RADAR emitters that emit on several different frequencies, pulse width, etc.

To tackle more complex cases (e.g., agile emitters, noisy signals, etc.), we propose two simple unsupervised methods based on optimal transport distances to deinterleave a signal without considering the PRI. The first method is based on an algorithm presented at a conference [21] (see also [22] for an application of optimal transport to RADAR emitter classification), which now integrates cluster processing with

a few pulses and a more efficient decision model to cut the dendrogram during the hierarchical clustering. Then, we propose a variant capable of handling emitters with similar characteristics at a low signal-to-noise ratio. Empirical results are given to evaluate and compare the robustness of the proposed methods under challenging cases (complex emitters characteristics, noisy data, etc.). Note that, for such problems in electronic warfare, data are often provided by ministry of defence (MoD) services/companies in the defense domain. To cope with such restrictions, this work provides a data simulator that mimics real data.

The rest of this article is organized as follows. The realistic simulated dataset is presented in Section II. Section III introduces the proposed methodology with the basic notion of optimal transport, followed by an improved version of this algorithm in Section IV. The simulation results are provided in Section V. Finally, Section VI concludes this article.

II. DATA DESCRIPTION

This section introduces the structure of the data considered in this work. Due to the complexity of working with real and labeled data (for defense security matters), the proposed deinterleaving methods are tuned, parameterized, and validated on simulated data. The complexity of the simulated data represents the main challenge in RADAR signal deinterleaving. We have created a RADAR signal simulator based on signal theory and telecommunication engineering. The RADAR simulator assumes the receiver is static, has a known detection threshold, and is omnidirectional, making the direction of arrival unusable. Several characteristics (or features) describe each emitter class as follows.

- 1) *Pulses Parameters*: Frequency, pulse width, level, direction of arrival, and PRI pattern.
- 2) *Technical Emitter Parameters*: Transmitted power, scan period, antenna directivity, and spatial distance from the receptor.

As the received signal is noisy, the features are perturbed by estimation errors. In addition, some low-energy pulses may be lost or truncated.

There is a wide variety of emitters. Some may have elementary parameters, such as emitting on a single frequency, while others have sophisticated pulse characteristics, e.g., agility in frequency, multiple pulse repetition of intervals, etc., or with rarer appearance periods. From these heterogeneous emitter profiles, an interleaved signal is created. The simulator controls the number of emitters in the signal and the desired number of pulses. A listening system collects the data without information about the environment. The diversity of the simulated signals results in the acquisition of single- or multisensor labeled signals, signals with frequency or time modulation, and signals comprising measurement errors, outliers, missing data, or non-Gaussian noise. The Doppler effect, fundamental in classical RADAR processing, is neglected here, as the frequency shifts corresponding to realistic target velocities are

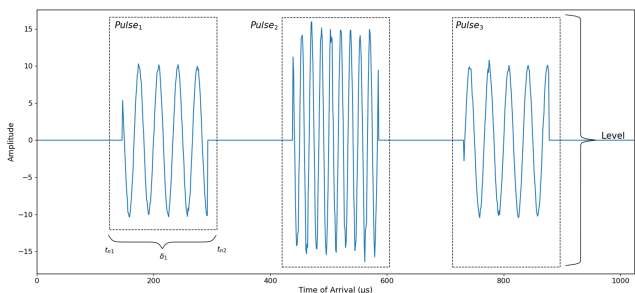


Fig. 1. Example of three simulated pulses from an emitter.

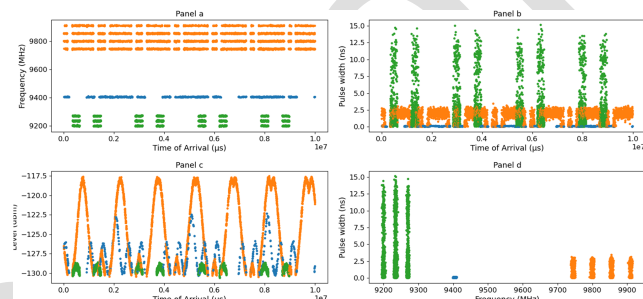


Fig. 2. Example of a simulated signal gathering 8917 pulses of three emitters, identified by a color.

negligible compared to the difference between frequencies used by different emitters. Another source of uncertainty in the measurements, multipath propagation, manifests itself by additional pulses, sharing the same frequency and duration as the original pulses, with decreased power and slightly shifted in time. As the proposed methods are based on the unaffected frequency and duration of the pulses, multipath propagation will be neglected.

A short simulated signal is shown in Fig. 1, containing three pulses. After the signal acquisition, the pulses are segmented, analyzed, and described by features. In this work, the n th acquired pulse is described using the following four characteristics.

- 1) Frequency, $F (f_n)$.
- 2) Pulse width, $PW (pw_n)$.
- 3) Level, $G (g_n)$.
- 4) Time of Arrival, $TOA (t_n)$.

An additional feature can be added: the difference of TOA (DTOA) δ_n , defined as the interval of time between two successive pulses belonging to the same emitter (in the case of a unique emitter, $\delta_n = t_n - t_{n-1}$). As the emitters are not separated at this stage, two successive pulses are not guaranteed to be transmitted by the same emitter. As a consequence, DTOA cannot be used before deinterleaving the signal.

Other features (waveform, frequency modulation, etc.) will not be considered here. Our approach considers only the four abovementioned features because they are always available and relatively well estimated. Fig. 2 shows a simulated signal containing 8917 pulses of three different emitters, each represented by a color (blue, green, and orange). Each point represents a pulse. Frequency, level,

Algorithm 1: Hierarchical Agglomerative Clustering Using Optimal Transport Distances to Deinterleave Emitter Pulses - HACOT.

Data: Set of pulses (X)

Features: Frequency (f_n), Pulse width (pw_n), Level (g_n), Time of arriva (t_n)

Parameters: Minimum number of points to form a cluster (minPts), Confidence level (α), Statistical test (*test*), Nonparametric method (*method*), Threshold (λ)

Procedure:

- 1) Pulses separation: apply HDBSCAN for all x_i of X from f_n and pw_n according to minPts: C (Set of clusters)
- 2) Cluster aggregation: apply hierarchical agglomerative clustering based on optimal transport distances presented in Algorithm 2 considering α , *test*, *method*, and λ for all c_i of C from t_n and g_n : Y (Set of aggregated clusters)

Result: Y , Deinterleaved sets of pulses.

and pulse width are plotted in function of time in panels Fig. 2(a)–(c), respectively. In addition, in Fig. 2(d), pulses are represented in the (f_n, pw_n) plane. These representations highlight several challenging characteristics of RADAR signals as follows.

- 1) Several emitters can be active at the same time.
- 2) Emitters can transmit continuously and regularly on different frequencies, as shown by emitter 1 in the (f_n, t_n) plane, which transmits on four different frequencies or emitter 2 on three frequencies.
- 3) Pulse widths can be severely misestimated, as low-power pulses can be split during the segmentation stage, spreading the pulses over 360 ns.
- 4) In general, estimated features are noisy.

In particular, several clusters in the (f_n, pw_n) plane can represent a given emitter, necessitating appropriate processing to regroup these clusters specifically for clustering-based methods in this plane. Our proposal introduces methods based on optimal transport distances capable of capturing and managing this phenomenon.

III. HIERARCHICAL AGGLOMERATIVE CLUSTERING COMBINED WITH OPTIMAL TRANSPORT DISTANCES

As pointed out in the previous section, one of the challenges in deinterleaving is the dispersion through several groups of pulses from an emitter in the (f_n, pw_n) plane. To overcome this problem, an extension of a two-step strategy proposed by the authors in [21] is introduced and highlighted in Algorithm 1. First, a clustering algorithm separates the pulses in the (f_n, pw_n) plane. Then, a hierarchical agglomerative clustering algorithm combined with optimal transport distances groups these clusters as an emitter can be represented by several clusters.

A. Pulses Separation With HDBSCAN in 2 Dimensions

The first step of the method is to apply a clustering algorithm to separate the pulses only from two highly discriminating and reliable characteristics: 1) frequency and 2) pulse width. It is important to note that not merging two different emitters in the same cluster is crucial. Thus, the proposed approach will overestimate the number of clusters at this step. Several algorithms were tested to determine the most suitable for the method, such as K-MEANS [23], hierarchical agglomerative clustering (HAC) [24], Gaussian mixture models (GMM) [25], Bayesian Gaussian mixtures models [26], or density-based algorithms (DBSCAN [27], OPTICS [28], HDBSCAN [29]). K-MEANS and GMM require assumptions about the number of clusters to be identified. Conventional methods, such as the Elbow method, unsupervised metrics, or data projection can set this number. These methods are effective in handling uncomplicated signals and emitters with simple characteristics. However, deinterleaving algorithms are designed for electronic warfare and unsupervised frameworks, where emitters often have complex and similar characteristics, thus rendering these methods unsuitable. Manually fixing the number of clusters for detection is unfeasible because each signal differs.

We propose to use the hierarchical density-based spatial clustering of applications with noise algorithm (HDBSCAN) [29], an unsupervised clustering algorithm based on a hierarchical version of the density-based spatial clustering of applications with noise (DBSCAN) [27]. The choice of the unsupervised HDBSCAN algorithm is motivated by its capacity to identify clusters with different densities and shapes, the few parameters to optimize, its ability to detect the clusters without setting the number of clusters to be detected beforehand, and its capacity to clean the signal by detecting outliers.

The main idea is to group points that live in the same dense area from a vector $X = \{(f_1, pw_1), (f_2, pw_2), \dots, (f_N, pw_N)\}$ grouping frequency and pulsewidth pulses from a signal of N pulses. Clustering results strongly depend on the hyperparameter determining the minimum size of a group that can be considered a cluster. HDBSCAN has been configured to overestimate the number of clusters returned so as not to mix the pulses of several emitters in the same cluster. Comparisons were made to find a tradeoff to fix this threshold. HDBSCAN is applied to normalized data using the quantile method in the (f_n, pw_n) plane because it preserves the data distribution as well as possible.

The results of the clustering performed by HDBSCAN in the (f_n, pw_n) plane are shown for the simulated dataset presented in Section II in Fig. 3. The algorithm identifies 11 clusters and a class of outliers (−1) with heterogeneous sizes ranging from 86 to more than 1500 pulses. The (f_n, pw_n) plane shows the plane where the clustering was performed and that the clustering results are consistent; the black boxes highlight the present emitters. The pulses from emitter 0 transmitting around 958 MHz have been correctly grouped into a single cluster. Conversely, HDBSCAN splits

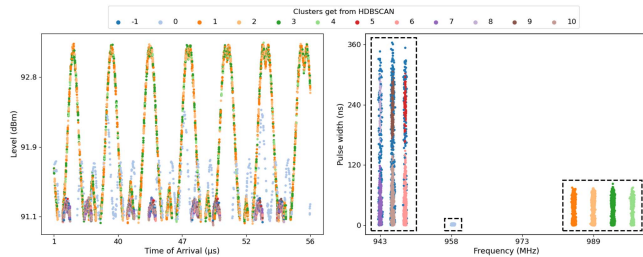


Fig. 3. HDBSCAN outputs obtained from frequency and pulse width. The algorithm identifies 11 clusters and an outliers class (−1), and the boxes highlight the pulses of the three emitters.

the emitter pulses above 981 MHz and below 951 MHz into several clusters. Indeed, both emitters transmit on several frequencies. In addition, low-power pulses of the lower frequency emitter are split at the segmentation stage, adding short pulse widths in the data.

The (g_n, t_n) plane on the left highlights the distribution of these clusters along the lobes.

B. Cluster Fusion With Hierarchical Agglomerative Clustering Using Optimal Transport Distances

As shown in Fig. 3, an emitter can be split over several clusters. Therefore, the second step of the method is designed to group these clusters according to shared characteristics. In particular, we assume that the clusters of a given emitter are simultaneously active, as illustrated in Fig. 3, where the pulses of an emitter are distributed over the lobes. From the HDBSCAN clusters, hierarchical agglomerative clustering was performed [30] using the optimal transport distances [31], [32] from TOA and level, as illustrated in Algorithm 2. The rest of the article will show that optimal transport distances are particularly well suited to our problem.

To define optimal transport distances between clusters, each cluster is represented by a probability measure describing its distribution from TOA weighted by level

$$\tau = \frac{1}{Z} \sum_{p=1}^P g_p \delta_{t_p} \quad (1)$$

with P , the number of pulses in the cluster and $Z = \sum_{p=1}^P g_p$, the total energy in the cluster. In order to decrease the computational complexity of the method and avoid numerical problems, the probability measures used in practice are obtained from data histograms. The range of TOAs is partitioned in B intervals $[b_{i-1}, b_i)$. The measure τ in (1) is replaced by

$$\bar{\tau} = \sum_{i=1}^B \bar{g}_i \delta_{c_i} \quad (2)$$

where $c_i = \frac{1}{2}(b_{i-1} + b_i)$, and $\bar{g}_i$ is the energy of the pulses in the interval, normalized so that $\sum_{i=1}^B \bar{g}_i = 1$.

The number of the bins is fixed according to the Freedman—Diaconis rule [33]: $B = 2 \frac{\text{IQ}(N)}{\sqrt[3]{N}}$ with N the signal size and IQ the interquartile range of the signal. This

Algorithm 2: Cluster Aggregation With Hierarchical Agglomerative Clustering Based on Optimal Transport Distances.

Data: Set of clusters (C)

Features: Level (g_n), Time of arrival (t_n)

Parameters: Threshold (λ), Nonparametric method ($method$), Confidence level (α), Statistical test ($test$)

Procedure:

- 1) Cluster significance analysis: separation of clusters according to λ : $D = \{c_i \in C \mid |c_i| \geq \lambda\}$ and $E = C \setminus D$
- 2) Determination of a hierarchical structure aggregating the clusters using optimal transport distances from D : *structure*
 - While** $\text{len}(D) \neq 1$ **do**
 - For all d_i of D , representation by a measure from t_n weighted by g_n : τ_i
 - Computation of the distance for all d_i of D using the optimal transport: $d(\tau_i, \tau_j)$
 - Aggregation of the two closest clusters: $(i, j)^*$
 - Updating D
- end**
- 3) Dendrogram pruning according to the decisional model presented in Algorithm 3 considering *method* and *test*: set of aggregated clusters, F
- 4) Dealing with excluded clusters E :
 - Estimation of the probability density for all f_i in F from t_n using *method*
 - Association of all e_i of E to F by maximum likelihood estimation

Output: F , aggregated clusters.

method is less sensitive to outliers because it is based on the width of the IQ range, which is less influenced by outliers than the mean and standard deviation. In addition, it better adapts to nonnormal or skewed distributions by considering the real dispersion of the signal. This method adjusts for the sample size, avoiding over or under-determination of the number of bins (other methods could be considered). Finally, the bin values are defined uniformly concerning the TOA values of the signal.

We now recall the definition of optimal transport distances, or Wasserstein-Kantorovitch distances, in the particular case of discrete probability measures. Let us consider two discrete probability distributions

$$\nu = \sum_{n=1}^N a_n \delta_{x_n} \text{ and } \mu = \sum_{m=1}^M b_m \delta_{y_m}$$

with $\mathbf{a} = (a_1, \dots, a_N)^T \in \mathbf{R}_+^N$, such that $\sum_{n=1}^N a_n = 1$, and $\mathbf{b} = (b_1, \dots, b_M)^T \in \mathbf{R}_+^M$, such that $\sum_{m=1}^M b_m = 1$.

A transport plan $\mathbf{P}$ between ν and μ is defined by its coefficients P_{nm} , representing the amount of mass taken

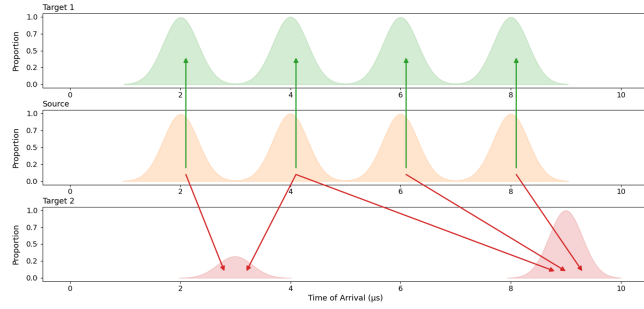


Fig. 4. Example of three simulated TOAs distribution. Source and target 1 belong to the same emitter and are characterized by a low transport cost. Conversely, source and target 2 come from two different emitters, implying a high transportation cost due to the difference in transmission times.

from x_n to y_m . Consider the cost function $c(\cdot, \cdot)$. In this work, we use the L_2 -norm: $c(x, y) = \|x - y\|_2$.

$C_{nm} = c(x_n, y_m)$ represents the cost of transporting a unit of mass from x_n to y_n , and the total cost $C(\mathbf{P})$ of a transport plan is

$$C(\mathbf{P}) = \sum_{n=1}^N \sum_{m=1}^M C_{nm} P_{nm}. \quad (3)$$

The consistency of the transport plan $\mathbf{P}$ with ν and μ is guaranteed by $\mathbf{P} \mathbf{1}_M = \mathbf{a}$, $\mathbf{P}^T \mathbf{1}_N = \mathbf{b}$. The optimal transport plan $\mathbf{P}^*$ is defined as the minimizer of (3) under the following constraints:

$$\mathbf{P}^* = \operatorname{argmin}_{\mathbf{P} \in \mathbf{R}_+^{N \times M}} C(\mathbf{P}) \text{ subject to } \mathbf{P} \mathbf{1}_M = \mathbf{a}, \mathbf{P}^T \mathbf{1}_N = \mathbf{b}. \quad (4)$$

The optimal transport distance between ν and μ is then defined by $d(\nu, \mu) = C(\mathbf{P}^*)$. Fig. 4 displays an example of applying optimal transport to define a distance between three simulated TOA distributions. Source and Target 1 belong to the same emitter and are active simultaneously, requiring very few moves to make them match, so low transport costs characterize them. Conversely, Source and Target 2 come from two different emitters, implying a high transportation cost due to the difference in transmission times.

In this particular case of univariate distributions, the distance can be computed using the cumulative distribution functions or quantile functions of the distributions. Note that the formulation (4) is more general, as it can be extended to multivariate distributions and can be modified to deal with unbalanced distributions [34].

The optimal transport distance is used in an agglomerative clustering algorithm; the two clusters with the smallest optimal transport distance are aggregated

$$(i^*, j^*) = \operatorname{argmind} (\tau_i, \tau_j). \quad (5)$$

After merging, the distances between merged clusters and other clusters are updated. Then, the process is repeated until the clusters are fully aggregated. To obtain accurate estimates of the distance between clusters, each cluster must have sufficient pulses, as a probability distribution represents them. This threshold was set to 100, considering

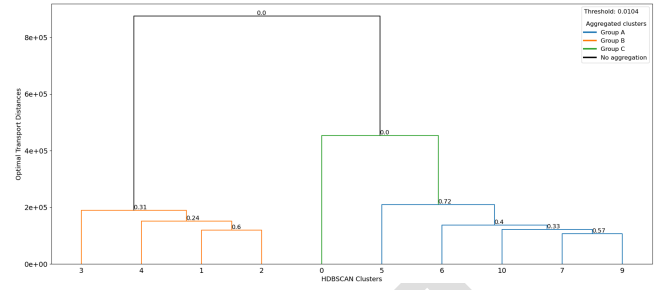


Fig. 5. Dendrogram representing the aggregations of hierarchical agglomerative clustering combined with optimal transport distances. The p -values of the Kolmogorov–Smirnov test from t_n evaluating the aggregated groups are displayed at each aggregation, while each color identified the final group.

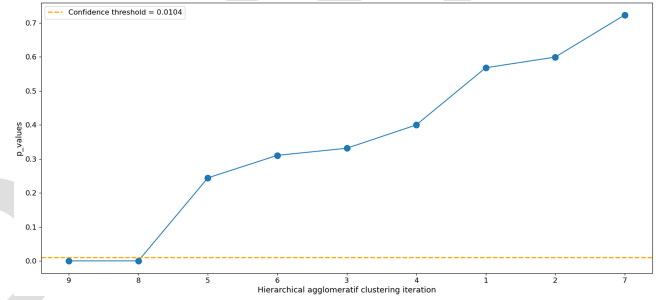


Fig. 6. p -values sorted according to the Kolmogorov–Smirnov test, and the orange line highlights the estimated confidence level.

the results we have seen on several signals. Clusters having fewer pulses than this threshold are excluded from the hierarchical approach but will be processed separately using an alternative approach afterward. Here, cluster 8, with 86 pulses, is set aside.

Fig. 5 represents the aggregation at each step of the hierarchical agglomerative clustering combined with the optimal transport distances, with heights of the edges representing the optimal transport distances calculated at each aggregation. The final clustering is obtained by stopping the fusion of clusters at appropriate stages. Indeed, prematurely stopping the fusions may result in separating several emitters' pulses into multiple groups, and late fusion results in grouping pulses of different emitters in a single cluster.

This final clustering is obtained by considering each fusion by increasing distance and deciding if this fusion should actually be performed by testing if the involved TOAs are likely to be sampled from the same TOA probability distribution. If the fusion is not performed, the later fusions involving these clusters are not considered. We propose to use the Kolmogorov–Smirnov test [35] to decide if a fusion is to be performed. For each fusion, the p -value of the hypothesis that the two clusters have been drawn from the same distribution is computed. Fusions with a p -value higher than a confidence threshold are performed.

This threshold is estimated by sorting the p -values in increasing order and identifying a breakpoint. The p -values for the test case considered here are plotted in Fig. 6, and the threshold is fixed at $\alpha_{ks} = 0.0104$, indicated by the horizontal line.

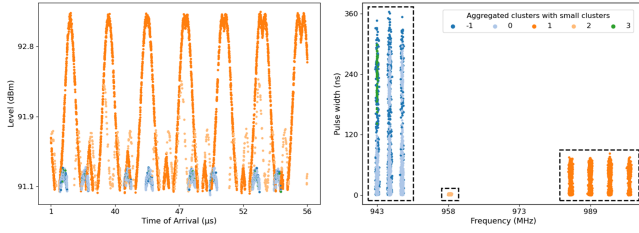


Fig. 7. Aggregated cluster obtained following the dendrogram from the hierarchical agglomerative clustering using optimal transport distances. Each color identified an aggregated cluster with an outliers class (-1).

Algorithm 3: Improved Decisional Model for Pruning the Dendrogram Based on Statistical Test.

Data: Set of clusters (D), Hierarchical structure ($structure$)

Features: Time of arrival (t_n)

Parameters: Statistical test ($test$), Confidence level (α)

Initialisation: Create a list with all aggregations labels from $structure$ and an empty list: N and L

Procedure:

```

while  $len(N) \neq 0$  do
    Selection of the first two aggregated clusters according to
     $structure: \{i, j\}$  in  $N$ 
    Representation by a measure from  $t_n: \theta_i$  and  $\theta_j$ 
    Computation of  $test$  between  $\theta_i$  and  $\theta_j$  and get  $p$ -value
    if  $p\text{-value} < \alpha$  then
        Remove all aggregations from  $N$  using these clusters
    else
        Add  $\{i, j\}$  in  $L$  and del  $\{i, j\}$  from  $N$ 
    end
end
Aggregate  $l_i$  in  $L$  according  $structure: F$ 
Output:  $F$ : Set of aggregated clusters

```

Given the dendrogram in Fig. 5, Clusters 7 and 9 are first evaluated and are represented by a measure describing their distributions, as in (1). The associated p -value is 0.57, which is higher than the threshold, meaning clusters 7 and 9 belong to the same emitter and are aggregated. The process is repeated until the top of the tree is reached, or a nonsignificant p -value is returned. The p -value between clusters 0 and 5, 6, 10, 7, and 9 is 0, which means the clusters should not belong to the same emitter. The process stops, and the next iteration is evaluated. The decisional model identifies three aggregated groups.

Compared to the method proposed in [21], where fusions were stopped at a given transport optimal distance, obtained by considering three unsupervised metrics, the fusions are here analyzed locally. This allows us to handle more complicated cases, e.g., with heterogeneous emitters, where a fusion should be performed, even if the associated optimal transport distance is higher than a fusion for which fusion was not performed.

Finally, Fig. 7 shows the result of cluster agglomeration. The emitter pulses around 958 MHz and above 981 MHz have been perfectly grouped into two sets of pulses. Conversely, the lower frequency emitter is still split into two

clusters, as one of the clusters has been excluded from the processing.

C. Dealing With Excluded Clusters

As previously mentioned, cluster 3 was set aside from the analysis due to its insufficient number of pulses. Clusters excluded from hierarchical clustering are now associated with an extensive cluster obtained after the grouping phase. The process is done in two steps as follows.

- 1) Estimation of the probability density of TOA of the extensive clusters using a kernel density estimator [36], [37] with Gaussian kernel.
- 2) Association of an excluded cluster to an extensive cluster by maximum likelihood estimation.

We used a nonparametric model to estimate the distribution parameters and their density function. Among these methods, the most commonly used are that of histograms [37], [38], K-nearest neighbors [39], [40], kernel-density estimation (KDE) [41], or NEURAL NETWORKS [42]. We chose to use the KDE because it allows us to build a density from each point and, therefore, to better consider the distribution's behavior. For a sample of points ($x_0, x_1, \dots, x_N$) belonging to a cluster distribution f , the kernel density estimator is written as

$$\hat{f}(x) = \frac{1}{Nh} \sum_{n=1}^N K\left(\frac{x - x_n}{h}\right) \quad (6)$$

with N the number of pulses, K the kernel, and h the smoothing parameter (bandwidth). The KDE mainly depends on two parameters: 1) the choice of the kernel and 2) the bandwidth. For a Gaussian kernel, by assuming the TOA can be represented by a Gaussian distribution, the optimal bandwidth can be obtained from the Scott method's [43] and is given by

$$h = \left(\frac{4}{3n}\right)^{1/5} \hat{\sigma} \approx 1.06 \hat{\sigma} n^{-1/5} \quad (7)$$

with $\hat{\sigma}$ the standard deviation of the cluster. As the distribution of the TOA for a given emitter is multimodal, we assume that the distribution of the TOA is a mixture of Gaussian densities. Temporal clustering with HDBSCAN for each extensive cluster is applied to identify each appearance and estimate the variance of each subcluster. The parameter $\hat{\sigma}$ in the bandwidth estimation is then obtained by averaging the variance of each subcluster. Finally, for a given excluded cluster, the likelihood of belonging to an extensive cluster i is given by

$$L_i = \prod_{n=1}^N \hat{f}_i(t_n) \quad (8)$$

with t_n the TOA. The excluded cluster is assigned to the extensive cluster, maximizing the likelihood

$$i^* = \operatorname{argmax}_i L_i. \quad (9)$$

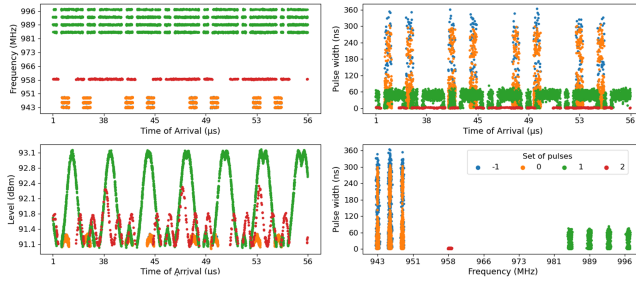


Fig. 8. Final grouping after applied hierarchical agglomerative clustering combined with optimal transport distances and kernel density estimation on excluded clusters. Each color identified a set of pulses with an outliers class (−1).

With the likelihood of the considered cluster belonging to the extensive cluster i . Cluster 3 will finally be associated with Cluster 2.

Finally, Fig. 8 displays the result of cluster agglomeration. The general hierarchical agglomerative clustering combined with the optimal transport distances and the excluded clusters processing correctly separates the pulses in the signal, groups them into three sets of pulses, and can handle clusters with few pulses. The outlier rate indicates that 4.5% of the pulses are lost during the deinterleaving process; HDBSCAN excluded these pulses in step 1.

D. Computational Complexity

The computational complexity of the first method can be estimated in function of the number N of pulses, the number of clusters C identified by HDBSCAN, and the number of bins B . The complexity of the first stage, dominated by HDBSCAN, is $O(N^2)$ [44]. In the second stage, C^2 distances are computed at the beginning of the hierarchical clustering, each costing $O(B)$ using the formulation of the Wasserstein distance as a distance between quantile functions. The number of updated distances during the clustering is also in $O(C^2)$.

The actual clustering, with complexity in $O(C^3)$, can be neglected under the reasonable assumption that the number of bins B is larger than the number of clusters. The $C - 1$ Komolgorov–Smirnov tests and the treatment of the excluded clusters are negligible compared to the initial distance computation.

The total complexity of the algorithm is thus $O(N^2 + C^2B)$. This complexity can be rewritten in functions of parameters related to the characteristics of the emitters. With T the length of the signal, K the number of emitters, $\bar{R}$ the mean rate of impulses, $\bar{C}$ the mean number of clusters per emitter (i.e., an evaluation of the complexity of the emitter), the complexity is $O(K^2(T^2\bar{R}^2 + \bar{C}^2B))$.

IV. IMPROVED HIERARCHICAL AGGLOMERATIVE CLUSTERING COMBINED WITH OPTIMAL TRANSPORT DISTANCES

In some cases, pulses associated with two different emitters can be grouped in the same cluster by HDBSCAN. This confusion can be caused by the emitters having similar

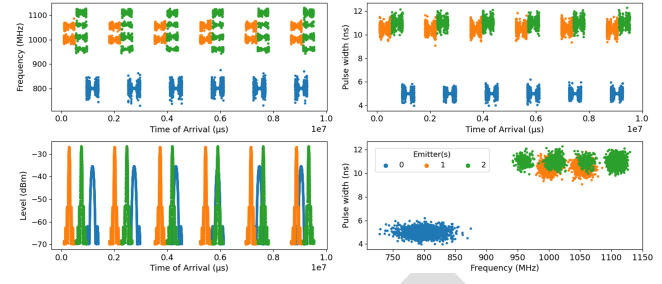


Fig. 9. Example of a simulated signal gathering the pulses of three emitters identified by a color.

Algorithm 4: Improved Hierarchical Agglomerative Clustering Using Optimal Transport Distances to Deinterleave Emitter Pulses - IHACOT.

Data: Set of pulses (X)

Features: Frequency (f_n), Pulse width (pw_n), Level (g_n), Time of arrival (t_n)

Parameters: Minimum number of points to form a cluster (minPts), Confidence level (α), Statistical test (test), Nonparametric method (method), Threshold (λ), distance (distance)

Procedure:

- 1) Pulses separation: apply HDBSCAN for all x_i of X from t_n , f_n , and pw_n according to minPts : C (Set of clusters)
- 2) Pre-clusters aggregation: hierarchical agglomerative clustering based on distance for all c_i of C from $\bar{f}_n$ and $\bar{pw}_n$: P (Set of preaggregated clusters)
- 3) Cluster aggregation: apply hierarchical agglomerative clustering based on optimal transport distances presented in Algorithm 2 considering α , test , method and λ for all p_i of P from t_n and g_n : Y (Set of aggregated clusters)

Result: Y , Deinterleaving sets of pulses.

characteristics and inaccurate estimation of the parameters of the pulses. This section proposes a variant of the method to alleviate this issue. Detailed in Algorithm 4, it differs from the previous method by adding the TOA in the first clustering steps.

Fig. 9 shows an example of a simulated signal gathering pulses from three emitters. In the (f_n, t_n) plane, the first emitter is easily identifiable and characterized by low frequencies (around 800 MHz). In comparison, the pulses of the two other emitters share the same frequency bands apart from 900 MHz. The pulse width spread and the simultaneity of emission of the two emitters emitting above 900 MHz, presented, respectively, in the (pw_n, t_n) and (g_n, t_n) planes, are not sufficient to differentiate them. The spreading of the pulses is caused by the presence of noise in the signal; this result is confirmed by looking at the pulses spread in the (f_n, pw_n) plane with the pulses overlapping, leading to an inseparability of these two emitters.

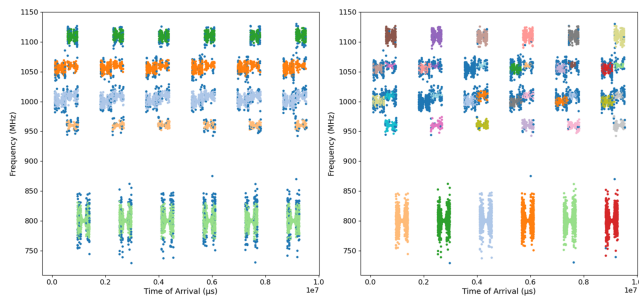


Fig. 10. Comparison of HDBSCAN clustering, respectively, performed in two dimensions on the left plot and three dimensions on the right plot.

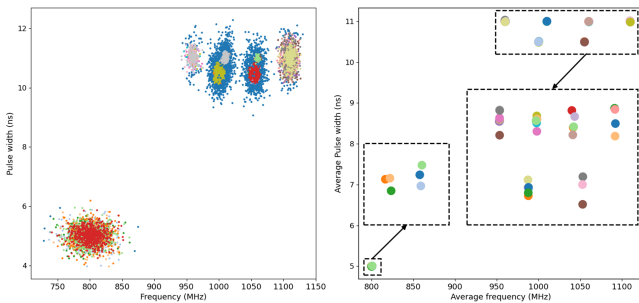


Fig. 11. HDBSCAN clustering results performed in three dimensions in the (f_n, pw_n) plane on the left and cluster averages in the (f_n, pw_n) plane on the right. HDBSCAN detects 42 clusters with an outliers class (−1) identified by colors.

A. Grouping of Pulses With the Same Characteristics in the (t_n, f_n, pw_n) Plane With HDBSCAN

A 3-D vector $X = \{(t_1, f_1, pw_1), (t_2, f_2, pw_2), \dots, (t_N, f_N, pw_N)\}$ grouping TOA, frequency, and pulsewidth pulses from a signal of N pulses is passed as input in HDBSCAN to be clustered.

Incorporating the last feature in the clustering improves emitters' pulse discrimination, as shown in Fig. 10. The clustering results obtained in two dimensions from the (f_n, pw_n) plane are displayed on the left plot; HDBSCAN identifies five clusters with an outliers class. The 2-D clustering mixed the pulses of the two emitters at 1000 and 1060 MHz. Conversely, adding the third dimension improves the separation, as highlighted by the right plot; the algorithm identifies 42 clusters with a class of outliers (−1).

B. Estimation of Frequencies and Pulse Width Means From Previous Clusters

The clusters obtained in three dimensions by HDBSCAN displayed on the left graph of Fig. 11 are overlapped. We compute the previous clusters' frequency and pulse width averages to obtain a new representation of each cluster in the $(f_{\text{mean}}, pw_{\text{mean}})$ plane, as highlighted in the right plot. Each dot represents the clusters' frequency and pulse width averages. This new representation improves cluster separation and avoids overlapping cluster pulses.

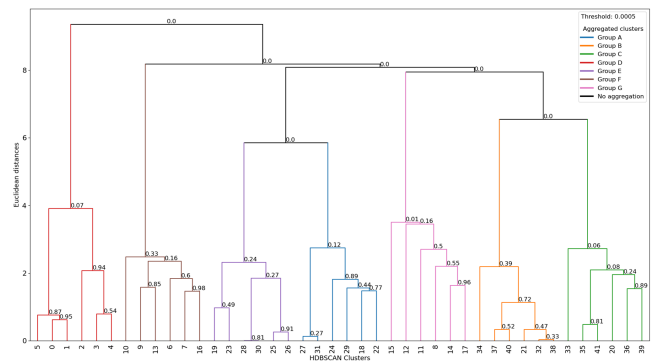


Fig. 12. Dendrogram representing the aggregations of hierarchical agglomerative clustering. The p -values of the Kolmogorov-Smirnov test from f_n evaluating the aggregated groups are displayed at each aggregation, while each color represents the aggregated cluster.

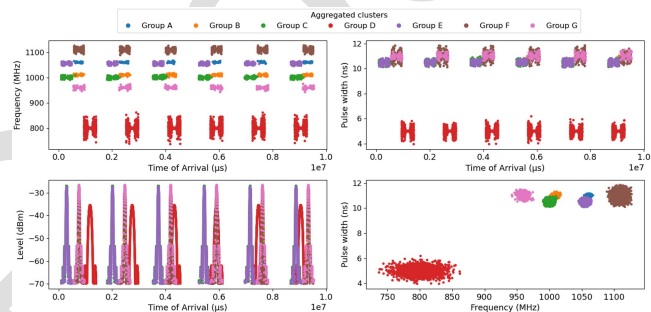


Fig. 13. Aggregated cluster obtained following the dendrogram from hierarchical agglomerative clustering represented by a color.

C. Grouping Clusters With the Same Characteristics From Frequency.

At this step, each cluster is represented by a frequency and pulse width averages: $(f_{\text{mean}}, pw_{\text{mean}})$. From this new representation, a classical hierarchical agglomerative clustering using Euclidean distances is applied to group clusters with the same characteristics. As previously, Algorithm 3 is applied, and fusions are stopped using the Kolmogorov-Smirnov test, checking that the frequency and pulse widths of the clusters follow the same distribution. Alternatively, the Student test [45] can also be used.

The dendrogram in Fig. 12 displays the groupings made from the pulses' characteristics by applying Algorithm 3, where the branches' heights are the optimal transport distance adjusted to improve the dendrogram. Here, the aggregated groups obtained using the Student test coincide with those obtained with the Kolmogorov-Smirnov test. The decision model identified seven aggregated clusters, and the results are presented in Fig. 13.

D. Clusters Agglomeration With Similar Temporal Characteristics With Optimal Transport Distances

The last step of this algorithm is similar to the previous method, based on hierarchical clustering with optimal transport distances. Fig. 14 shows the final three sets of pulses that our method identified and a class of outliers (−1). The

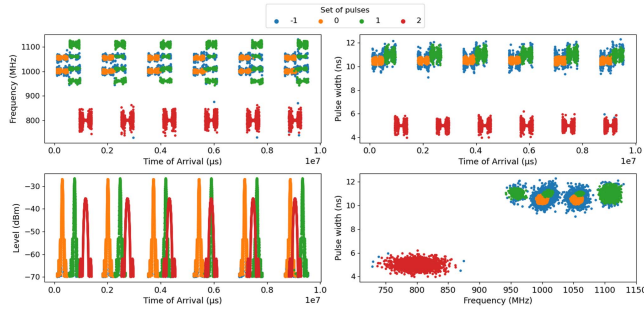


Fig. 14. Final grouping after applied hierarchical agglomerative clustering combined with optimal transport distances and kernel density estimation on excluded clusters. Each color identified a set of pulses with an outliers class (−1).

TABLE I
Simulated Emitters Characteristics

Emitter	Frequency (MHz)	Pulse width (ns)	PRI (μ s)
0	1025, 1050, 1075, 1100	15.3	70, 110
1	825, 884	15.2	85, 86, 87, 88, 89, 9, 91, 92, 93, 94, 104, 106, 108, 11.
2	860	15.3	95, 96, 97, 98, 99, 100.
Std	2.14	1e-3	2.55

methodology could separate and regroup emitters' pulses with similar characteristics in the presence of noise; HDBSCAN excluded about 9% of pulses.

E. Computational Complexity

The complexity of this second method is the sum of the complexity of HDBSCAN, the precluster aggregation, and the cluster aggregation. The complexity of HDBSCAN and the cluster aggregation is identical to the first method, with C now being the number of preaggregation clusters. The precluster aggregation is a simple hierarchical clustering, with complexity $O(P^3)$, with P the number of clusters identified by HDBSCAN. P can be estimated by $P = CTS$, where S is the average number of sweep by second. Compared to the previous algorithm, the complexity is augmented by $O(C^3T^3S^3)$, or $O(\bar{C}^3K^3T^3S^3)$.

V. RESULTS

The two proposed methods are tested on a simulated signal with three emitters, with characteristics given in Table I. An example with 10044 pulses is shown in Fig. 15. The characteristics of the emitters, with multiple frequencies, similar pulse widths, and PRIs, are chosen to highlight the robustness of the methods concerning adverse cases. The methods' performance is assessed using the adjusted rand index (ARI) [46]. As the emitters can emit on several frequencies with various pulse widths, methods based on clustering on these features will fail to group all pulses of a given emitter.

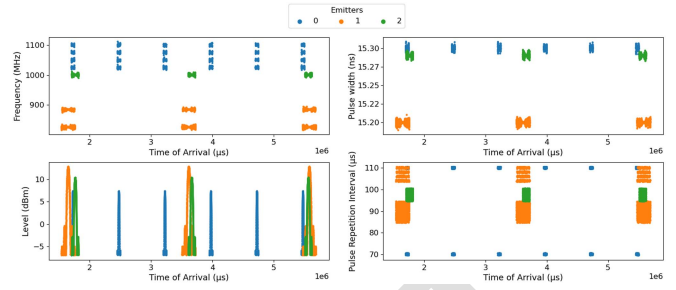


Fig. 15. Simulated signal gathering 10044 pulses from three emitters. Each color identified an emitter.

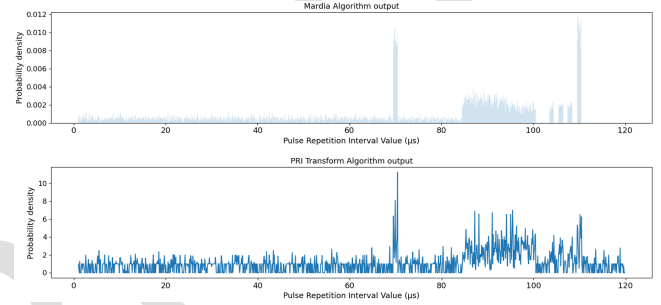


Fig. 16. Results of Mardia and PRI transform algorithms to deinterleave the signal of the Fig. 16 from the PRI.

A. Comparison With PRI-Based Methods

Two approaches were selected to be compared to the two deinterleaving methods developed. One conventional method, based on the analysis of PRI histograms [6] and another, more recent, based on the PRI transform [47], have been selected. Methods based on deep learning models have been deliberately set aside as they require a lot of data for the training step and perform inadequately with few data, while the proposed method is unsupervised. The outputs of these methods, shown in Fig. 16, cannot be used to estimate the number of emitters and separate their pulses. Indeed, the complex PRI patterns of each emitter and the similarity of their PRIs are such that both methods fail.

Fig. 16 represents the result of Mardia's and PRI Transform method applied to our previous signal. The calculated PRI on the Mardia's method values visually oscillates around 70 μ s. We also observe a block of PRI between 80 and 110 μ s. The same results are observable in the PRI Transform method but are less pronounced. The application of the PRI Transform method is significantly time-consuming due to the length of the signal; we restricted the search for PRI values to a predefined range instead of considering all TOA values. Despite knowing the PRI values to detect, we failed to find the emitter's emission patterns. As emitters share emission patterns, using these methods to estimate the present PRI values and detect the emitters in this signal is unfeasible.

To facilitate understanding, HACOT and IHACOT identify the deinterleaving methods based on the 2-D and 3-D clustering developed and presented in Sections III and IV. The methods are performed based on applying

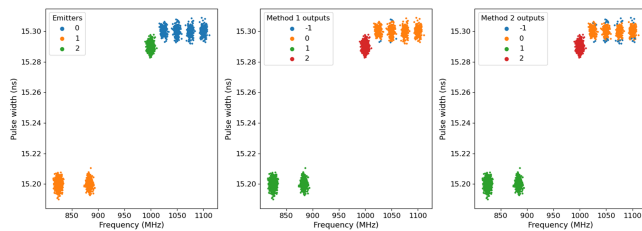


Fig. 17. Results of the HACOT and IHACOT developed in Sections III and IV to deinterleave the signal of the Fig. 15 based on Kolmogorov–Smirnov test for all grouping phases.

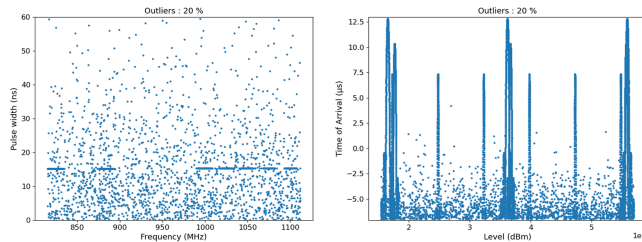


Fig. 18. Pulses representation in the (f_n, pw_n) and (t_n, g_n) planes according 20% of outliers added in the initial signal.

the Kolmogorov–Smirnov test on all grouping phases, and results are displayed in Fig. 17. The emitters have different characteristics, and the pulses are sufficiently spread in the plane, leading to a fast and correct deinterleaving of the three emitters' pulses. HACOT detects 0.01% outliers against 0.5% for IHACOT.

B. Robustness to Outliers

The performance of the methods with varying proportions of outliers is evaluated. The proportion of outliers added to the signal varies between 0 and 90%. The performance criteria are averaged over 50 realizations. The outliers were generated uniformly between the range values for frequency and TOA. Alternatively, an exponential law generates pulse widths and level values because it favors the appearance of low values, making the outliers' values more realistic, as illustrated in Fig. 18, which represents the previous signal in the (f_n, pw_n) and (g_n, t_n) planes with 20% of outliers added.

The previously mentioned metrics used to evaluate the algorithm's performances are presented in Fig. 19. IHACOT is evaluated through two approaches: 1) using only the Kolmogorov–Smirnov test for the cluster regrouping phases (IHACOT-KS) and 2) by mixing the Student test for the first grouping and the Kolmogorov–Smirnov test for the second grouping (IHACOT-Mix). IHACOT-KS and IHACOT-Mix give higher results, as the ARI indicates. The better performance of these methods is due to the 3-D-clustering, which cleans the data more efficiently by detecting outliers, leading to more reliable and stable results. The result is confirmed by the following graph: the number of pulses classified as outliers by HDBSCAN is higher for IHACOT-KS and IHACOT-Mix than HACOT-KS as the outliers added in the signal increase. Moreover,

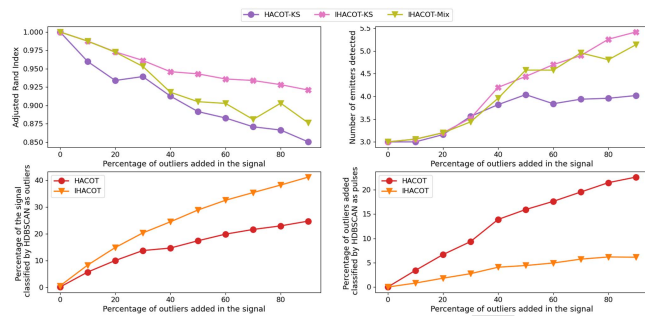


Fig. 19. Performance of deinterleaving methods according to the outliers rate added with ARI, number of emitters detected, part of the signal classified by HDBSCAN as outliers, and part of outliers added in the signal classified in the sets of pulses. Each curve identifies a method.

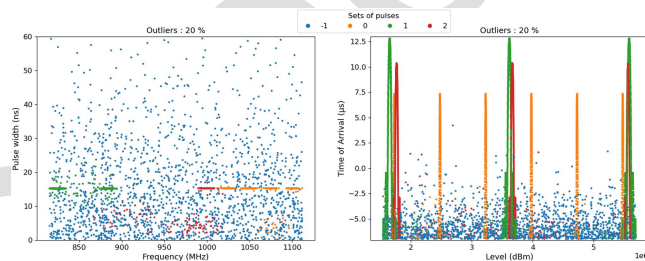


Fig. 20. Results of IHACOT with Kolmogorov–Smirnov for each grouping phase applied with 20% of outliers added in the (f_n, pw_n) and (t_n, g_n) plane. Each color represents a detected emitter, and -1 is the outliers class.

the final sets of pulses obtained from IHACOT-KS and IHACOT-Mix are composed of fewer outliers than those from HACOT-KS, as illustrated by the bottom-right plot, meaning the 3-D clustering provides less noisy sets of pulses. IHACOT-KS and IHACOT-MIX overestimate the number of emitters detected in the signal, as shown in the top-right plot. For our problem, we prioritized the overestimated emitter detected rather than underestimating them and having mixed pulses in a single cluster. IHACOT-KS and IHACOT-Mix seem more adapted to make deinterleaving.

Fig. 20 show the results of IHACOT-KS applied on Fig. 18. Note that some outliers corresponding to low-intensity pulses are grouped in a set of pulses. By employing 3-D clustering, the pulses emitted were successfully separated, avoiding mixing pulses from multiple emitters into one group despite adding outliers in the signal.

C. Sensitivity to Estimation Errors

Similarly, we applied the same process to evaluate the performance of our algorithms by varying the noise level. The noise level in the estimated parameters of the PDW is varied by multiplying the baseline level given in Table I with a factor (*noise coefficient*) ranging from 1 to 10. Likewise, 50 simulated signals per noise level were generated in the outliers experimentation to average the results and obtain consistent results.

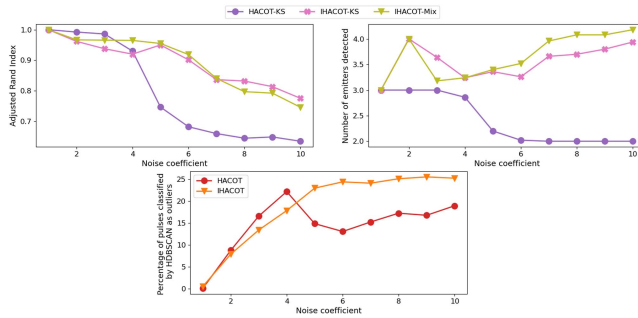


Fig. 21. Performance of deinterleaving methods according to the noise added with ARI, number of emitters detected, and part of pulses classified by HDBSCAN as outliers. Each curve identifies a method.

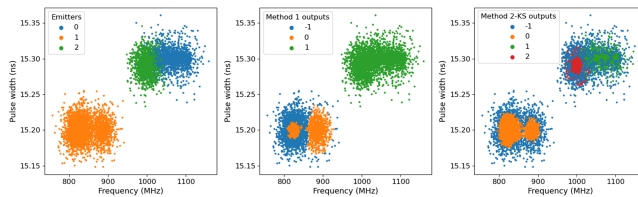


Fig. 22. Deinterleaving results in the plane (f_n, pw_n) obtained from HACOT-KS and IHACOT-KS applied on Fig. 15 with the baseline level multiplied by the coefficient 40.

Fig. 21 shows the evolution of previously used metrics, averaged over 50 realizations, applied on Fig. 15 with different noise levels. ARI is high for all methods, and their results are similar. IHACOT-KS and IHACOT-Mix give better results than HACOT-KS. Methods based on 3-D clustering classify more pulses as outliers, leading to better separation and avoiding mixing pulses in the numerous 3-D detected clusters, as shown at the bottom. Finally, as illustrated in the top-right plot, HACOT-KS mixes the emitter's pulses given a certain noise level, while IHACOT-KS and IHACOT-Mix tend to overestimate them slightly.

An application of HACOT-KS and IHACOT-KS is shown in Fig. 22. The left plot represents the ground truth, and the middle and the right plots are the results of HACOT-KS and IHACOT-KS. Each color represents an emitter and (-1) the outliers. IHACOT-KS separates the pulses better by applying a 3-D clustering, whereas HACOT-KS fails to separate the pulses of the two emitters around 15.3 ns.

VI. CONCLUSION

This article proposes two deinterleaving methods capable of managing complex emitters without considering the PRI pattern. Hierarchical agglomerative clustering combined with optimal transport distances is used to identify emitters using several frequencies, presented agility characteristics, or having multiple operation modes. Here, the assumption is that the pulses of a given emitter, with varied

parameters, are received at similar times. The PRI is not used, making the methods robust with respect to complex PRI patterns. However, using frequency and pulse width in the first clustering step is straightforward, specifically when the listening perimeter concerns a port or an airport, and the signal is noisy; a second approach, including preclustering made from the TOA, frequency, and pulse width, is initiated, to separate pulses better, before applying the previous hierarchical agglomerative clustering using optimal transport distances developed to overcome these limitations. The two resulting algorithms are flexible, playing with various hyperparameters (e.g., threshold, kernels, etc.) that are adjusted with the datasets. Finally, simulation results have shown the interest of the proposed methods in terms of accuracy. Furthermore, in addition to the robustness concerning complex emission patterns, the proposed approaches are also robust with respect to the estimation of the parameters and outliers.

The deinterleaving approaches are based on accessible and reliable features. The deinterleaving results could be improved by including additional features, such as the direction of arrival, which is not always accessible but highly discriminating. A complementary algorithm could be developed based on temporarily available features and combined with the principal methodology to improve the deinterleaving of a signal. The Kolmogorov–Smirnov and Student tests only apply between two data distributions during the hierarchical approach to prune the dendrogram, implying that at least two emitters are present in the deinterleaving signal. Other tests or methods could be explored to consider cases where the intercepted pulses correspond to only one emitter or combine other relevant tests to improve the pruning. The similarity between emitters strongly influences the approaches, specifically when the emitters share common characteristics. In the electronic warfare context, operators deliberately select listening areas; specific areas like airports or ports are not typically favored for these activities, although these locations can potentially pick up many emitters' signals. Our considerations have not accounted for scenarios where a hundred or more emitters might be within the surveillance area but could be subject to extending this research. During the hierarchical approach, the threshold evaluating the significance of clusters is deliberately set at 100, considering the results we have seen on several signals. We are currently working on the development of a statistical method to set this threshold more confidently. The main idea will be to compare the distribution of the cluster, grouping all its pulses with distributions by removing pulses first using the Kolmogorov–Smirnov test to determine if the original distribution is statistically different from the distribution with suppressed pulses to fix the threshold and several descriptive statistics (comparison of means, standard deviations, kurtosis, skewness, etc.). Finally, applying the hierarchical clustering algorithm using the optimal transport distances assumes that the clusters belonging to an emitter are simultaneously active, which in some instances is not valid.

731 REFERENCES

- [1] M. I. Skolnik, *Radar Handbook*, New York, NY, USA: McGraw-Hill Education, 2008.
- [2] R. Wiley, *ELINT: The Interception and Analysis of Radar Signals*, Norwood, MA, USA: Artech, 2006.
- [3] D. R. Wilkinson and A. W. Watson, "Use of metric techniques in ESM data processing," *IEE Proc. F (Commun., Radar Signal Process.)*, vol. 132, no. 4, pp. 229–232, 1985.
- [4] C. Davies and P. Hollands, "Automatic processing for ESM," *IEE Proc. F (Commun., Radar Signal Process.)*, vol. 129, no. 3, pp. 164–171, 1982.
- [5] V. Clarkson and P. S. Ray, "Parallel processing for differencing and histogramming," Electronics Research Lab Adelaide (Australia), Adelaide, SA, Australia, Tech. Rep. ERL-0692-RR, 1993.
- [6] H. Mardia, "New techniques for the deinterleaving of repetitive sequences," *IEE Proc. F (Radar Signal Process.)*, vol. 136, no. 4, pp. 149–154, 1989.
- [7] D. Milojević and B. Popović, "Improved algorithm for the deinterleaving of radar pulses," *IEE Proc. F (Radar Signal Process.)*, vol. 139, no. 1, pp. 98–104, 1992.
- [8] K. Nishiguchi and M. Kobayashi, "Improved algorithm for estimating pulse repetition intervals," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 407–421, Apr. 2000.
- [9] Z.-M. Liu and P. S. Yu, "Classification, denoising, and deinterleaving of pulse streams with recurrent neural networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 4, pp. 1624–1639, Aug. 2019.
- [10] Z. Zhou, G. Huang, H. Chen, and J. Gao, "Automatic radar waveform recognition based on deep convolutional denoising auto-encoders," *Circuits, Syst., Signal Process.*, vol. 37, no. 9, pp. 4034–4048, 2018.
- [11] M. Zhu, S. Wang, and Y. Li, "Model-based representation and deinterleaving of mixed radar pulse sequences with neural machine translation network," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 3, pp. 1733–1752, Jun. 2022.
- [12] Z.-M. Liu, "Pulse deinterleaving for multifunction radars with hierarchical deep neural networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 57, no. 6, pp. 3585–3599, Dec. 2021.
- [13] M. A. Nuhoglu, Y. K. Alp, M. E. C. Ulusoy, and H. A. Cirpan, "Image segmentation for radar signal deinterleaving using deep learning," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 1, pp. 541–554, Feb. 2023.
- [14] P. Sharma and K. K. Sarma, "PRI modulation classification in EW systems using deep learning," in *Proc. Int. Conf. Emerg. Glob. Trends Eng. Technol.*, 2022, pp. 453–462.
- [15] S. Yang et al., "Deep contrastive clustering for signal deinterleaving," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 60, no. 1, pp. 252–263, Feb. 2024.
- [16] Z. Kang, Y. Zhong, Y. Wu, and Y. Cai, "Signal deinterleaving based on U-Net networks," in *Proc. 8th Int. Conf. Comput. Commun. Syst.*, 2023, pp. 62–67.
- [17] H. Xiang, F. Shen, and J. Zhao, "Deep ToA mask-based recursive radar pulse deinterleaving," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 2, pp. 989–1006, Apr. 2023.
- [18] X. Dong, Y. Liang, and J. Wang, "Distributed clustering method based on spatial information," *IEEE Access*, vol. 10, pp. 53143–53152, 2022.
- [19] M. Scherrek and B. Rigling, "Online estimation of radar emitter cardinality via Bayesian nonparametric clustering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 57, no. 6, pp. 3791–3800, Dec. 2021.
- [20] S. Scholl and S. Brüggewirth, "Incremental deinterleaving of radar emitters," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 38, no. 8, pp. 26–36, Aug. 2023.
- [21] M. Mottier, G. Chardon, and F. Pascal, "Deinterleaving and clustering unknown RADAR pulses," in *Proc. IEEE Radar Conf.*, 2021, pp. 1–6.
- [22] M. Mottier, G. Chardon, and F. Pascal, "RADAR emitter classification with optimal transport distances," in *Proc. 30th Eur. Signal Process. Conf.*, 2022, pp. 1871–1875.
- [23] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm," *J. Roy. Stat. Soc. Ser. C (Appl. Statist.)*, vol. 28, no. 1, pp. 100–108, 1979.
- [24] S. C. Johnson, "Hierarchical clustering schemes," *Psychometrika*, vol. 32, no. 3, pp. 241–254, 1967.
- [25] G. J. McLachlan and K. E. Basford, *Mixture Models: Inference and Applications to Clustering*, New York, NY, USA, M. Dekker, 1988, vol. 38.
- [26] S. J. Roberts, D. Husmeier, I. Rezek, and W. Penny, "Bayesian approaches to Gaussian mixture modeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1133–1142, Nov. 1998.
- [27] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowl. Discov. Data Mining*, vol. 96, no. 34, 1996, pp. 226–231.
- [28] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, "OPTICS: Ordering points to identify the clustering structure," *ACM Sigmod Rec.*, vol. 28, no. 2, pp. 49–60, 1999.
- [29] R. J. Campello, D. Moulavi, and J. Sander, "Density-based clustering based on hierarchical density estimates," in *Proc. Pacific-Asia Conf. Knowl. Discov. Data Mining*, 2013, pp. 160–172.
- [30] S. Chakraborty, D. Paul, and S. Das, "Hierarchical clustering with optimal transport," *Statist. Probability Lett.*, vol. 163, 2020, Art. no. 108781.
- [31] C. Villani, *Optimal Transport: Old and New*, Berlin, Germany: Springer, 2009, vol. 338.
- [32] N. Bonneel, M. V. D. Panne, S. Paris, and W. Heidrich, "Displacement interpolation using Lagrangian mass transport," in *Proc. SIGGRAPH Asia Conf.*, 2011, pp. 1–12.
- [33] D. Freedman and P. Diaconis, "On the histogram as a density estimator: L. 2 theory," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 57, no. 4, pp. 453–476, 1981.
- [34] T. Séjourné, G. Peyré, and F.-X. Vialard, "Chapter 12 - unbalanced optimal transport, from theory to numerics," in *Handbook of Numer. Anal.*, vol. 24, E. Trélat and E. Zuazua, Eds., Amsterdam, The Netherlands: Elsevier, 2023, pp. 407–471.
- [35] F. J. Massey Jr, "The Kolmogorov-Smirnov test for goodness of fit," *J. Amer. Stat. Assoc.*, vol. 46, no. 253, pp. 68–78, 1951.
- [36] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *Ann. Math. Statist.*, vol. 27, pp. 832–837, 1956.
- [37] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1065–1076, 1962.
- [38] D. W. Scott, "On optimal and data-based histograms," *Biometrika*, vol. 66, no. 3, pp. 605–610, 1979.
- [39] Y. Mack and M. Rosenblatt, "Multivariate k-nearest neighbor density estimates," *J. Multivariate Anal.*, vol. 9, no. 1, pp. 1–15, 1979.
- [40] K. Fukunaga and L. Hostetler, "Optimization of k nearest neighbor density estimates," *IEEE Trans. Inf. Theory*, vol. 19, no. 3, pp. 320–326, May 1973.
- [41] D. W. Scott, R. A. Tapia, and J. R. Thompson, "Nonparametric probability density estimation by discrete maximum penalized-likelihood criteria," *Ann. Statist.*, vol. 8, pp. 820–832, 1980.
- [42] M. Magdon-Ismael and A. Atiya, "Neural networks for density estimation," in *Proc. Annu. Conf. Neural Inf. Process. Syst.*, 1998, pp. 522–528.
- [43] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*, Hoboken, NJ, USA: Wiley, 2015.
- [44] R. J. G. B. Campello, D. Moulavi, A. Zimek, and J. Sander, "Hierarchical density estimates for data clustering, visualization, and outlier detection," *ACM Trans. Knowl. Discov. Data*, vol. 10, no. 1, Jul. 2015, Art. no. 5.
- [45] Student, "The probable error of a mean," *Biometrika*, vol. 6, no. 1, pp. 1–25, 1908.
- [46] L. Hubert and P. Arabie, "Comparing partitions," *J. Classification*, vol. 2, pp. 193–218, 1985.
- [47] K. Nishiguchi, "A new method for estimation of pulse repetition intervals," in *National Convention Record of IECE of Japan*, 1983.



Manon Mottier received the M.Sc. degree in economic engineering and data analysis, CY Cergy Paris University, Pontoise, France, in 2017, the M.Sc. degree in information processing and computer science from Sorbonne University, Paris, France, in 2018, and the Ph.D. degree in mathematics and signal processing from the Signals and Systems Laboratory, CentraleSupélec, Paris-Saclay University, Gif-sur-Yvette, France, in 2024.

Her Ph.D. research focused on integrating artificial intelligence for radar signal processing. More specifically, She initially worked on two unsupervised approaches using optimal transport distances to deinterleave a signal containing pulses from an indeterminate number of emitters. Subsequently, she employed the optimal transport distance framework to develop an algorithm that could distinguish and identify these emitters by referencing a comprehensive database. She is currently an R&D Engineer in signal processing and artificial intelligence in an industrial research laboratory. Her research interests include artificial intelligence (optimal transport, unsupervised approach, deep learning, etc.) and signal and image processing, specializing in electronic warfare for military applications.



Gilles Chardon received dual Engineering degrees in electrical engineering from Ecole Polytechnique, Palaiseau, France, and Telecom Paris-Tech, Paris, France, in 2009, the M.Sc. degree in acoustics, signal processing, and computer science applied to music (ATIAM) and the Ph.D. degree in acoustics from Sorbonne Université, Paris, in 2009 and 2012, respectively.

He is currently an Associate Professor with CentraleSupélec, Gif-sur-Yvette, France. His research interests include acoustical source localization and radar signal processing.



Frédéric Pascal (Senior Member, IEEE) received the master's degree in applied statistics (probabilities, statistics and applications: signal, image et networks) from the University Paris VII - Jussieu, Paris, France, in 2003, and the Ph.D. degree in signal processing from the University of Paris X Nanterre, Nanterre, France, in 2006.

From 2006 to 2008, he was a Postdoc with the SATIE Lab, CNRS, ENS Cachan, Gif-sur-Yvette, France. Between 2008 and 2011, and 2012 and 2013, he was an Assistant Professor

and Associate Professor with SONDR, CentraleSupélec, Gif-sur-Yvette. Between 2013 and 2014, he was a Visiting Associate Professor with ECE Department, National University of Singapore, Singapore. Since January 2014, he has been a Full Professor with the L2S Laboratory, CentraleSupélec, University Paris-Saclay. Between 2017 and 2018, he was the Head of the "Signals and Statistics" group of L2S. Since 2019, he has been the Coordinator of the activities in artificial intelligence at CentraleSupélec and is the Givaudan Chair on data sciences. Since 2017, he has been on the Executive Committee of the DATAIA Institute as the Program Coordinator, and he was appointed as the Co-Director in 2021. Since 2023, he has been the Director of the DATAIA Institute, University Paris-Saclay. He is the author and coauthor of more than 200 papers in the top journals and conferences in signal processing, image processing, and statistics. His research interests include estimation, detection, and classification for statistical signal processing, and radar and image processing applications.

Dr. Pascal was the Chair of the EURASIP SAT in Theoretical and Methodological Trends in Signal Processing (TMTSP) from 2015 to 2017. He was a Member of the IEEE Signal Processing Society SAM Technical Committee (2015–2021). Between 2019 and 2021, he was the Vice Chair of the Data Science Initiative of the IEEE Signal Processing Society. He was an Associate Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING (2015–2018), *EURASIP Journal on Advances in Signal Processing* (2015–2021), and *Elsevier Signal Processing* (2018–2021).