UNICA UNIVERSITÀ DEGLI STUDI DI CAGLIARI

This is the Author's accepted manuscript version of the following contribution:

David Macii, Paolo Attilio Pegoraro, Carlo Sitzia, "Clustering and Fusion of Distributed Multitone Frequency Measurement Data for Enhanced Grid-level Power Quality Monitoring," in IEEE Transactions on Instrumentation and Measurement, vol. 74, 2025, Art. no. 9004014.

The publisher's version is available at:
http://dx.doi.org/10.1109/TIM.2025.3556911

When citing, please refer to the published version.

© 2025 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This full text was downloaded from UNICA IRIS https://iris.unica.it/

Clustering and Fusion of Distributed Multitone Frequency Measurement Data for Enhanced Grid-level Power Quality Monitoring

David Macii, *Senior Member, IEEE*, Paolo Attilio Pegoraro, *Senior Member, IEEE* and Carlo Sitzia, *Member, IEEE*

***Abstract*—The increasing diffusion of both nonlinear loads and renewable-based distributed energy resources may cause serious power quality problems due to the injection of significant harmonics and interharmonics. To mitigate some of these problems, the controllers of smart Power Electronic Converters (PECs) can embed techniques for harmonics' compensation. However, to maximize their performance, the most relevant narrowband disturbances should be detected and their frequency estimated with high accuracy. Given that harmonics and interharmonics maintain consistent frequencies across different points in the grid, this paper introduces a flexible technique that leverages multiple distributed estimates to accurately identify critical narrowband disturbances propagating throughout the system. The proposed approach relies on four main steps: the initial two are designed for implementation in various measurement devices to identify local interferers, while the latter two are intended to operate on a central aggregator. This central unit collects multitone distributed frequency and amplitude estimates to cluster and fuse data corresponding to the same interferers. A key novelty of the proposed approach lies in the clustering algorithm that relies on an adaptive threshold depending on the actual measurement uncertainty of the available data. This method effectively filters out potential outliers, reducing the risk of misdetections, before calculating the uncertainty-weighted average of frequency values within each cluster. Several Monte Carlo and grid-level simulations confirm the excellent performance of the proposed approach in identifying critical sinusoidal disturbances, thus paving the way to better distributed compensation strategies to be implemented in PECs.**

***Index Terms*—Power Electronic Converters (PECs), harmonics and interharmonics detection, synchronized measurement, ESPRIT, frequency estimation, DBSCAN, data fusion.**

I. Introduction

The increasing penetration of both renewable-based Distributed Energy Resources (DERs) (e.g., wind turbines and Photovoltaic generators) as well as strongly nonlinear loads [1] may cause critical Power Quality (PQ) issues at distribution level, especially due to large harmonics and interharmonics [2], [3]. Such components can lead to equipment overheating, improper operation of protective devices, and increased losses. Therefore, identifying the common narrowband disturbances spreading across a distribution system (DS) is vital for maintaining PQ and ensuring efficient grid operation.

The high-efficiency grid-supporting and grid-forming Power Electronics Converters (PECs) play a pivotal role for flexible and safe integration of DERs [4]. Despite the implementation of various techniques to reduce the effects of nonlinear distortion [5], [6], such devices may inject significant voltage and current harmonics and interharmonics into the grid [7], [8]. These disturbances not only degrade PQ, but might also jeopardize grid stability [9]–[11], causing the sudden disconnection of one or more DERs even if advanced fault ride-through functions and anti-islanding systems are used [12].

To further enhance the compensation of the harmonic and interharmonic distortion levels caused by PECs and other nonlinear devices, more advanced techniques are needed, e.g., by using the data potentially collected and shared by multiple devices [13], [14]. Indeed, it is shown in [15], [16] that the use of collaborative PECs has the potential to improve both stability and PQ, especially if advanced time synchronization solutions are adopted to align the timestamps of data measured by the devices deployed throughout the grid [17]. However, the effectiveness of harmonics and interharmonics compensation techniques is still highly dependent on the accurate and timely identification of the most critical tones, which is a challenging problem for various reasons, i.e.,

- The presence of measurement points in DSs is usually rather limited.
- The local detection of harmonics and interharmonics is affected by unavoidable misdetection risks [18], [19].
- The number of significant harmonic and interharmonic interferers is usually unknown a priori.
- The harmonic frequencies are only partially known (they may suffer from off-nominal deviations), whereas interharmonic frequencies are completely unknown.

As far as the first issue is concerned, a polymorphic, fine-grained measurement infrastructure at the distribution level is highly desirable to support advanced monitoring applications such as harmonic state estimation (HSE) [20], or harmonic sources identification (see, for example, [21], [22]), especially in the case of unknown harmonic impedance [23]. While the widespread deployment of dedicated measurement devices is often impractical, a promising approach involves integrating data acquisition, measurement, and time synchronization

This work was funded by the European Union - Next Generation EU, Mission 4, Component 2, through the Italian Ministerial grant PRIN 2022 "Smart grid-connected power converters based on advanced synchrophasor-inspired harmonics measurements for holistic integration of renewable energy sources (POWERHERO)," n. 20224X2AYH, CUP F53D23000490006.

D. Macii is with the Department of Industrial Engineering, University of Trento, Trento, Italy (e-mail: david.macii@unitn.it).

P. A. Pegoraro and C. Sitzia are with the Department of Electrical and Electronic Engineering of the University of Cagliari, 09123 Cagliari, Italy (email: [paolo.pegoraro,carlo.sitzia]@unica.it).

functions directly into the control units of PECs and other smart grid (SG) devices. This integration not only reduces costs but also simplifies installation challenges. For instance, in [24] a Phasor Measurement Unit (PMU) is implemented into the control unit of a smart inverter; in [25] a PQ-framed aggregation algorithm is integrated into a smart meter; in [26] a technique for PQ parameters estimation is implemented in the Field Programmable Gate Array (FPGA) of a power meter; in [27] a distribution-oriented harmonic PMU is proposed; finally, in [28] a smart substation is presented.

Even assuming that the number of measurement devices is sufficient to detect the critical harmonics and interharmonics spreading across the grid, to the best of Authors' knowledge, no existing algorithm is comparable to that presented in this paper. In fact, typical HSE techniques rely on the assumption that the number of harmonics to be measured is known a priori [29], the interharmonics are usually not included in the signal model and, last but not least, the frequencies of the various sinusoidal disturbances are assumed known and are not corrected by the estimation algorithm.

The solution presented in this paper addresses all the aforementioned challenges by utilizing a recently proposed technique for effective and reliable harmonic and interharmonic identification [19]. Additionally, it incorporates a clustering and data fusion approach that:

1) Reduces the risk of both false and missed detections;
2) Enhances the amplitude and frequency estimation accuracy of detected narrowband components, thereby facilitating a distributed compensation of such disturbances.

Compared with the approach preliminarily presented at IEEE AMPS 2024 [30], which introduced the idea of grid-level disturbances frequency support estimation, the main technical advancements are summarized below.

- A customized version of the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [31] is used to create sets of frequency and amplitude data collected in different points of the grid and associated to the same narrowband components. Also, the threshold for clustering is not fixed as in [30], but it depends on the uncertainty of an Estimation of Signal Parameters via the Rotational Invariance Technique (ESPRIT)-based multitone frequency estimator, thereby improving both algorithm flexibility and robustness.
- The theoretical background of the paper has been thoroughly revised to integrate new concepts and provide a deeper description of the proposed algorithm.
- Several additional simulation results are shown to demonstrate the enhanced capabilities of the proposed method.

The rest of this paper is organized as follows. In Section II, the various steps of the proposed algorithm and the related theoretical background are described in detail. In Section III the algorithm performance are analyzed extensively considering both Monte Carlo (MC) simulations and a realistic Medium Voltage (MV) distribution network. Finally, Section IV provides some closing remarks and an outlook of future research.

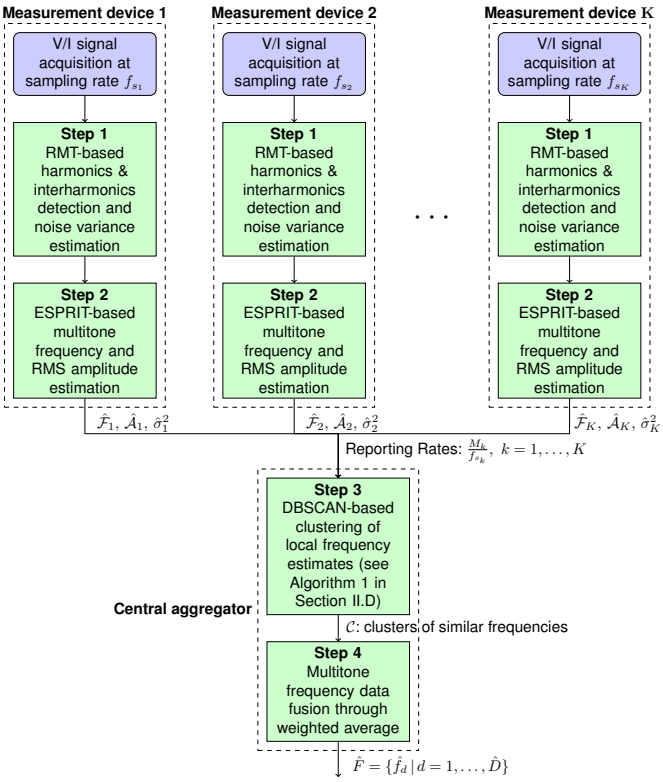

Fig. 1. Overview of the overall grid-level multitone detection and frequency estimation technique.

II. Theoretical Model and Algorithm Description

The proposed algorithm is based on the signal model described in Section II-A, and it consists of four steps, as qualitatively shown in Fig. 1. The first step is implemented locally in each measurement device and it utilizes the Random Matrix Theory (RMT) to detect the harmonics and interharmonics emerging from the noise floor in a given point of the grid. The details of this step are recalled in Section II-B. The second step of the algorithm is also local to each measurement device and it relies on the ESPRIT to determine frequency and Root Mean Square (RMS) amplitude of all the detected sinusoidal components [32] (see Section II-C).

The third step relies on a customized DBSCAN approach to cluster the common frequency components detected by different devices at the same time. The details of this step are reported in Section II-D. Finally, in the last step, the frequency data associated to the same tone (i.e. within the same cluster) and the same time instant are merged into a single value through a weighted average keeping into account the Signal-to-Noise Ratio (SNR) of each estimated component (see Section II-E). It is important to emphasize that the last two steps are assumed to run on a central coordinator collecting and processing the local measurement results.

A. Signal Model and Singular Vector Decomposition

Let K be the number of measurement points in a grid. If the voltage or current waveforms are acquired by the

$k-$th measurement device with a sampling rate f_{s_k}, a generic discrete-time data record of length M_k can be modeled as

$$x_k(n) = \sum_{d=1}^{D_k} \sqrt{2} A_{d,k} \cos\left(2\pi \frac{f_{d,k}}{f_{s_k}} n + \varphi_{d,k}\right) + \epsilon_k(n) \quad (1)$$

where index $n = (r-1)M_k+1, \ldots, rM_k$ (with $r \in \mathbb{N}^+$) refers to the elements of disjoint data records of size M_k, D_k is the (unknown) number of significant narrowband components at measurement point k, with RMS amplitude $A_{d,k}$, frequency $f_{d,k}$ and initial phase $\varphi_{d,k}$, respectively, (arranged in ascending order of frequency), and $\epsilon_k(\cdot)$ denotes the normally-distributed additive wideband random noise with zero mean and variance σ_k^2. If the samples of (1) are arranged into the column vector

$$\mathbf{x}_k(n) = [x_k(rM_k)\, x_k(rM_k-1) \cdots x_k((r-1)M_k+1)]^T \quad (2)$$

where ^{T} is the transpose operator, it can be shown that the autocorrelation matrix of (2) can be decomposed into singular values (SVs) and singular vectors as follows, i.e. [32],

$$\mathbf{R}_k = E\{\mathbf{x}_k(n)\mathbf{x}_k^T(n)\} = \mathbf{S}_k \mathbf{\Lambda}_k \mathbf{S}_k^T = \left[\mathbf{S}_{2D_k}\, \mathbf{S}_{M_k-2D_k}\right] \begin{bmatrix} \frac{M_k}{2}\mathbf{P}_k + \sigma_k^2 \mathbf{I}_{2D_k} & 0 \\ 0 & \sigma_k^2 \mathbf{I}_{M_k-2D_k} \end{bmatrix} \begin{bmatrix} \mathbf{S}_{2D_k}^T \\ \mathbf{S}_{M_k-2D_k}^T \end{bmatrix} \quad (3)$$

where:

- $E\{\cdot\}$ is the expectation operator;
- $\mathbf{\Lambda}_k$ is the diagonal matrix comprising the SVs of matrix (3). Such values can be grouped into two subsets: the largest $2D_k$ ones are approximately proportional to the elements of diagonal matrix $\mathbf{P}_k = \text{diag}\left(A_{1,k}^2, A_{1,k}^2, \ldots, A_{D_k,k}^2, A_{D_k,k}^2\right)$. Each pair of equal values corresponds to the power of a specific sinusoidal tone, while the remaining $M_k - 2D_k$ values are equal to the wideband noise power.
- The columns of matrix $\mathbf{S}_k$ are the orthonormal singular vectors associated with the corresponding elements of $\mathbf{\Lambda}_k$. Again, the first $2D_k$ columns of $\mathbf{S}_k$ are gathered into matrix $\mathbf{S}_{2D_k}$ and refer to the sinusoidal tones, while the remaining singular vectors in $\mathbf{S}_{M_k-2D_k}$ refer to the noise SVs.
- Finally, $\mathbf{I}$ represents an identity matrix with a size as specified in the subscript.

The SVs and singular vectors from (3) can be used to detect the sinusoidal tones in (1), as it will be explained in the following subsection. It is important to note that both $\mathbf{R}_k$ and all the terms on the right-hand side of (3) vary with each data record and should, therefore, be indexed by n. However, to simplify the notation and enhance readability, this indexing is omitted here and in the following sections as well.

B. Step 1: Local RMT-based Harmonics and Interharmonics Detection and Noise Variance Estimation

The singular value decomposition (SVD) in (3) reveals that, if the SVs of the autocorrelation matrix $\mathbf{R}_k$ were known and ordered properly, D_k could be determined by just counting the number of SVs exceeding the noise variance and then dividing the result by two. However, in practice $\mathbf{R}_k$ can be estimated using only a finite number of data records of given length. Thus, the ordered SVs exhibit a smoothed rather than a staircase profile, which depends on: i) the number L_k of data vectors (2) used to estimate the autocorrelation matrix; ii) their length M_k and iii) the correlation between them. As a result, the smallest estimated SV associated with a sinusoidal tone is lower than expected one and, at the same time, the largest estimated SV of the noise subspace is greater than σ_k^2. The reduction of the gap between the estimated SVs associated with the signal subspace and wideband noise subspace, respectively, tends to hinder the sinusoidal tones detection process. In fact, the threshold to discriminate the sinusoidal tones from the noise floor must be set with care. Even though such a gap tends to broaden as the values of M_k and L_k grow, in practice these parameters must be as low as possible to keep both latency and computational burden within reasonable limits. Thus, if L_k consecutive records of M_k samples each are considered (this is done to keep the additive noise sequences in different records uncorrelated), the maximum likelihood estimator of $\mathbf{R}_k$ is simply

$$\hat{\mathbf{R}}_k = \frac{1}{L_k} \sum_{l=0}^{L_k-1} \mathbf{x}_k(n - l \cdot M_k)\mathbf{x}_k^T(n - l \cdot M_k) \quad (4)$$

where l is the index referring to one of the L_k data records. If $D_k = 0$, the data vectors used to build (4) would consist just of uncorrelated noise samples. In this case, the largest estimated noise SV follows a Tracy-Widom distribution [33]. However, if one or more sinusoidal tones are included in (1) (i.e. $D_k > 0$), the statistical distribution of the largest estimated noise SV is biased. Therefore, both the estimated variance of the noise floor and the threshold for detecting possible narrowband components have to be adjusted iteratively to correct this bias. To this end, the following algorithm can be used [19].

1) The SVD is applied to the estimated autocorrelation matrix $\hat{\mathbf{R}}_k$ to obtain an expression similar to (3).
2) The SVs in $\hat{\mathbf{\Lambda}}_k = \text{diag}\left(\hat{\lambda}_{1,k}, \ldots, \hat{\lambda}_{M_k,k}\right)$ are arranged in a descending order.
3) Given that the fundamental tone is certainly present in (1), the initial guess for the estimated number of sinusoidal tones is 1. Thus, all the diagonal elements of $\hat{\mathbf{\Lambda}}_k$ other than the greatest pair are assumed to belong to the noise vector subspace. As a consequence, the variance of the noise floor can be estimated by averaging such elements, while correcting the bias due to the fundamental tone iteratively [33]. Since the actual number of sinusoidal tones is unknown a priori, the noise variance estimator might wrongly include SVs associated to one or more harmonics or interharmonics. For this reason, the compensation approach described above must be repeated iteratively to exclude progressively all the pairs of SVs associated with some sinusoidal tone, until the first SV of the noise subspace is actually found. In particular, if $j \geq 1$ is the iteration index of this loop, the noise variance can be estimated as follows, i.e., [33]

$$\hat{\sigma}_{k_{j,t}}^2 = \frac{1}{M_k - 2j} \left[\sum_{m=2j+1}^{M_k} \hat{\lambda}_{m,k} - \sum_{i=1}^{2j} \tau_{i,j,k,t}\right] \quad (5)$$

where $\tau_{i,j,k,t}$ (for $i = 1, \dots, 2j$ and $t = 0, 1, \dots$) are the correction terms of the bias due to the estimated SVs associated to the first j tones. Since any correction term corresponds to the largest possible positive real-valued root of the second-order algebraic equation [33]

$$\tau_{i,j,k,t}^2 - \tau_{i,j,k,t}\left[\hat{\lambda}_{i,k} + \hat{\sigma}_{k_{j,t-1}}^2\left(1 - \frac{M_k - 2j}{L_k}\right)\right] + \hat{\lambda}_{i,k}\hat{\sigma}_{k_{j,t-1}}^2 = 0 \tag{6}$$

equations (5)-(6) must be solved iteratively in a loop (indexed by t) until when either the larger root no longer changes significantly (e.g., changes are $< 1\,\%$) or the roots of (6) are no longer real numbers. When the termination condition of the loop is reached, the best possible estimated noise variance for a given number j of detected sinusoidal components is the value obtained in the last iteration, i.e., $\hat{\sigma}_{k_j}^2 = \hat{\sigma}_{k_{j,t}}^2$.

4) Afterwards, the following largest SV in $\hat{\mathbf{\Lambda}}_k$ (i.e., $\hat{\lambda}_{2j+1,k}$) is compared with a threshold that depends on the statistical distribution of the SV that is assumed to belong to the noise subspace. Accordingly, the following hypothesis test is performed:

$$\begin{aligned} H_0 &: \hat{\lambda}_{2j+1,k} \leq T_j \\ H_1 &: \hat{\lambda}_{2j+1,k} > T_j \end{aligned} \tag{7}$$

where

$$T_j = \hat{\sigma}_{k_j}^2 \left(\mu_{L_k, M_k - 2j} + q(\alpha)\xi_{L_k, M_k - 2j}\right) \tag{8}$$

is indeed the narrowband signal detection threshold updated after iteration j, while μ_{L_k,M_k-2j}, ξ_{L_k,M_k-2j} and $q(\alpha)$ are, respectively, the centering parameter, the scale parameter and the $\alpha-$quantile of the Tracy-Widom distribution [34]. Note that α denotes the given probability of wrongly detecting a sinusoidal tone (namely, the acceptable false alarm probability). Also, in [33] it is shown that if $\alpha \ll 1$, the parameters of (8) can be computed in a closed form as follows, i.e.,

$$\begin{aligned} \mu_{L_k,M_k-2j} &= \frac{1}{L_k}\left(\sqrt{L_k - \tfrac{1}{2}} + \sqrt{M_k - 2j - \tfrac{1}{2}}\right)^2 \\ \xi_{L_k,M_k-2j} &= \sqrt{\frac{\mu_{L_k,M_k-2j}}{L_k}}\left(\frac{1}{\sqrt{L_k - \frac{1}{2}}} + \frac{1}{\sqrt{M_k - 2j - \frac{1}{2}}}\right)^{\frac{1}{3}} \\ q(\alpha) &\approx \left(-\tfrac{3}{2}\cdot\log(4\sqrt{\pi}\cdot\alpha)\right)^{\frac{2}{3}} \end{aligned} \tag{9}$$

5) If the null hypothesis H_0 is rejected (i.e., if H_1 is true), j is incremented by 1 and the algorithm starts over. Otherwise, the algorithm ends returning the number of detected sinusoidal tones (i.e., $\hat{D}_k = j$) along with the noise variance estimate $\hat{\sigma}_k^2 = \hat{\sigma}_{k_j}^2$.

C. Step 2: Local ESPRIT-based Multitone Frequency and RMS Amplitude Estimation

Once the number of sinusoidal tones detected by the $k-$th measurement point is known, the corresponding singular vectors (namely the first $2\hat{D}_k$ columns of matrix $\hat{\mathbf{S}}_k$ that result from the SVD of (4)) can be extracted to create the submatrix $\hat{\mathbf{S}}_{2D_k}$. This submatrix can be used to build the following two $(M_k - 1) \times 2\hat{D}_k$ matrices, i.e., $\hat{\mathbf{S}}'_{D_k} = [\mathbf{I}_{M_k-1}\ \mathbf{0}]\hat{\mathbf{S}}_{2D_k}$ and $\hat{\mathbf{S}}''_{D_k} = [\mathbf{0}\ \mathbf{I}_{M_k-1}]\hat{\mathbf{S}}_{2D_k}$. Since the columns of $\hat{\mathbf{S}}_{2D_k}$ consist of pairs of sine and cosine functions sharing the same frequencies as the detected tones, matrix $\hat{\mathbf{S}}''_{D_k}$ is related to $\hat{\mathbf{S}}'_{D_k}$ by a rotation matrix Ψ, i.e., $\hat{\mathbf{S}}''_{D_k} = \hat{\mathbf{S}}'_{D_k}\Psi$ [32]. Therefore, the rotation matrix results from

$$\Psi = (\hat{\mathbf{S}}'^T_{D_k}\hat{\mathbf{S}}'_{D_k})^{-1}\hat{\mathbf{S}}'^T_{D_k}\hat{\mathbf{S}}''_{D_k} \tag{10}$$

The SVs $e^{\omega_1}, e^{-\omega_1} \dots, e^{\omega_{2\hat{D}_k}}, e^{-\omega_{2\hat{D}_k}}$ of Ψ exhibit unit magnitude and their arguments are the positive and negative normalized angular frequencies of the $\hat{D}_k$ detected tones. If such arguments are arranged in ascending order, the frequencies of the detected tones expressed in Hz are finally given by

$$\hat{f}_{d,k} = \frac{f_{s_k}}{2\pi}\omega_{\left(\hat{D}_k + d\right)} \quad \text{for } d = 1, \dots, \hat{D}_k \tag{11}$$

Note that, in general, for a given index d, the value of $\hat{f}_{d,k}$ in (11) might be completely different from the value of $f_{d,k}$ in (1), since some of the sinusoidal tones in the $k-$th signal might be either undetected or wrongly detected.

The RMS amplitude of the sinusoidal tones can be finally estimated from the first $2\hat{D}_k$ SVs of $\hat{\mathbf{\Lambda}}_k$ resulting from the SVD of (4). This estimation is performed after associating the correct pair of SVs to each frequency. Specifically, let $m = g_k(d)$ represent the function that returns the index of the first SV of matrix $\hat{\mathbf{\Lambda}}_k$ corresponding to the $d-$th estimated frequency $\hat{f}_{d,k}$. Recalling that a pair of approximately equal SVs is associated to the same tone, it follows from (3) that

$$\hat{A}_{d,k} = \sqrt{\frac{2}{M_k}\left(\frac{\hat{\lambda}_{m,k} + \hat{\lambda}_{(m+1),k}}{2} - \hat{\sigma}_k^2\right)}, \quad d = 1, \dots, \hat{D}_k \tag{12}$$

In (12), the estimation bias caused by the noise variance is properly compensated and the SVs corresponding to the same tone (that in principle should be the same as shown in (3)) are instead averaged to reduce the estimation uncertainty affecting $\hat{\lambda}_{m,k}$ and $\hat{\lambda}_{(m+1),k}$. In the rest of the paper, the frequency and amplitude values of the detected sinusoidal tones shall be arranged into two sets, i.e.,

$$\begin{aligned} \hat{\mathcal{F}}_k &= \left\{\hat{f}_{d,k} \,|\, d = 1, \dots, \hat{D}_k\right\} \\ \hat{\mathcal{A}}_k &= \left\{\hat{A}_{d,k} \,|\, d = 1, \dots, \hat{D}_k\right\} \end{aligned} \tag{13}$$

The elements of these sets, along with the estimated noise variance $\hat{\sigma}_k^2$ for $k = 1, \dots, K$, are finally sent to the central aggregator device for data clustering and fusion.

D. Step 3: DBSCAN-based Clustering of Local Frequency Estimates

The clustering step relies on a customized DBSCAN algorithm [31]. The reason of this choice is twofold. First, as explained above, the number of sinusoidal tones (and thus the number of classes) is unknown a priori, which makes other clustering techniques (e.g., K-means) hard to apply. Second, the centroid of different clusters (i.e., the actual frequency of the various sinusoidal tones) has to be estimated. The detailed pseudo-code of the proposed DBSCAN-based clustering technique is reported in Algorithm 1. The algorithm

Algorithm 1: Customized DBSCAN-based clustering of the pairs of indexes that refer to similar frequencies at different measurement points (see Step 3 of Fig. 1).

Input : Estimated quantities from K measurement devices (see Step 2 in Fig. 1):
- $\hat{f}_{d,k} \in \hat{\mathcal{F}}_k$, for $d = 1, \ldots, \hat{D}_k$, $k = 1, \ldots, K$;
- $\hat{A}_{d,k} \in \hat{\mathcal{A}}_k$, for $d = 1, \ldots, \hat{D}_k$, $k = 1, \ldots, K$;
- $\hat{\sigma}_k^2$, $k = 1, \ldots, K$.

Minimum points per cluster: $N_{\text{pts}} = 2$.

Output: $\mathcal{C}$: Set of clusters including the pairs of indexes that refer to similar frequencies at different measurement points.

foreach $k = 1$ *to* K **do**
  foreach $\hat{f}_{d,k} \in \hat{\mathcal{F}}_k$ **do**
    EstiVar: Compute the approximate variance of $\hat{f}_{d,k}$ for given $\hat{A}_{d,k}$ and $\hat{\sigma}_k^2$ values;
  end
end

DBSCAN Radius: Compute the DBSCAN clustering radius:
$U_f = \frac{10\sqrt{2}}{K} \sum_{k=1}^{K} \frac{1}{\hat{D}_k} \frac{f_{s_k}}{\sqrt{\pi^2 L_k M_k^2}} \sum_{d=1}^{\hat{D}_k} \frac{\hat{\sigma}_k}{\hat{A}_{d,k}}$

Initialization:
- Let $\mathcal{I} = \{(d,k) \mid \hat{f}_{d,k} \in \hat{F}_k,\ k = 1, \ldots, K\}$;
- Mark all index pairs in $\mathcal{I}$ as unvisited.
- Initialize empty list of clusters: $\mathcal{C} \leftarrow \emptyset$.

foreach $(d,k) \in \mathcal{I}$ **do**
  if (d,k) *is unvisited* **then**
    Mark (d,k) as visited;
    `// Find neighbors of` $\hat{f}_{d,k}$ `within` U_f
    $\mathcal{N} \leftarrow \{(d',k') \in \mathcal{I} \mid |\hat{f}_{d,k} - \hat{f}_{d',k'}| \leq U_f\}$;
    if $|\mathcal{N}| \geq N_{pts}$ **then**
      Create a new cluster $\mathcal{C}_i$;
      Add (d,k) to $\mathcal{C}_i$;
      `// Expand the cluster by exploring neighbors recursively`
      while $\mathcal{N}$ *is not empty* **do**
        Remove a point (d',k') from $\mathcal{N}$;
        if (d',k') *is unvisited* **then**
          Mark (d',k') as visited;
          `// Find neighbors of` $\hat{f}_{d',k'}$ `within` U_f
          $\mathcal{N}' \leftarrow \{(d'',k'') \in \mathcal{I} \mid |\hat{f}_{d',k'} - \hat{f}_{d'',k''}| \leq U_f\}$;
          if $|\mathcal{N}'| \geq N_{pts}$ **then**
            Add all points in $\mathcal{N}'$ to $\mathcal{N}$;
          end
        end
        if $(d',k') \notin \mathcal{C}_i$ **then**
          Add (d',k') to $\mathcal{C}_i$;
        end
      end
      Add $\mathcal{C}_i$ to $\mathcal{C}$;
    else
      Mark (d,k) as noise;
    end
  end
end
Return $\mathcal{C}$

relies on the assumption that all the elements of sets $\hat{\mathcal{F}}_k$ and $\hat{\mathcal{A}}_k$, for $k = 1, \ldots, K$ are time-stamped with the same time tag. In brief, the algorithm clusters iteratively the pair of indexes (d,k) associated to the frequency estimates collected from different measurement devices when they differ by no more than the expanded uncertainty U_f. If $\mathcal{I} = \{(d,k) \mid \hat{f}_{d,k} \in \hat{F}_k,\ k = 1, \ldots, K\}$ is the set including the index pairs of the collected frequency data, first the unvisited elements of $\mathcal{I}$ are progressively marked as visited; then the index pairs of the neighbors of each $\hat{f}_{d,k}$ value are found and placed into a set $\mathcal{N}$. If the number of elements of $\mathcal{N}$ is at least equal to a given minimum number of points N_{pts}, a new cluster $\mathcal{C}_i$ is created. Then, the cluster is expanded recursively by adding the pair of indexes of all the other frequency estimates that differ by less than $\pm U_f$ until no new points can be added. The index pairs that do not meet the clustering criteria are marked as noise. This process is repeated until all data are processed, yielding the final set of clusters $\mathcal{C}$. The elements of $\mathcal{C}$ are indeed the clusters $\mathcal{C}_i$ (for $i = 1, \ldots, \hat{D}$) of the index pairs associated to the $\hat{D}$ sinusoidal tones that are detected in at least N_{pts} measurement points of the grid. Note that in general $\hat{D}$ may differ from any $\hat{D}_k$. The clustering algorithm may wrongly regard the index pairs of existing sinusoidal tones as noise or it may incorrectly split into two or more classes the index pairs of different local frequency estimates that instead refer to the same sinusoidal component.

Two parameters of the algorithm require some further explanation. First, in this paper $N_{\text{pts}} = 2$. This means that a cluster is valid if the frequency estimates collected in at least 2 measurement points are consistent. This setting provides higher sensitivity than the solution in [30], where a sinusoidal tone was instead regarded as critical when detected by the majority of the K measurement devices.

The second crucial parameter that requires some further explanation is U_f. This parameter should depend indeed on the estimation uncertainty of the multitone frequency values collected from all measurement devices. In this respect, it is worth recalling that the covariance matrix of the ESPRIT multitone frequency estimator was analytically derived in [35]. Unfortunately, the exact general expression of the diagonal elements of that matrix is rather complex. However, if the values of L_k, M_k and SNR are large enough, and if the frequencies of the various sinusoidal components are not excessively close to one another, the variance of a classic ESPRIT-based, multitone frequency estimator for a given (d,k) pair is approximately $\text{var}(\hat{f}_{d,k}) \approx \frac{f_{s_k}^2}{2\pi^2} \frac{1}{L_k M_k^2} \frac{\hat{\sigma}_k^2}{\hat{A}_{d,k}^2}$.

Since DBSCAN requires a common range for all clusters, as a rule of thumb, the value U_f can be set about one order of magnitude (i.e., about 10 times) greater than the average standard uncertainty of the difference between a generic pair of ESPRIT-based frequency estimates, i.e.,

$$U_f = \frac{10\sqrt{2}}{K} \sum_{k=1}^{K} \frac{1}{\hat{D}_k} \frac{f_{s_k}}{\sqrt{2\pi^2 L_k M_k^2}} \sum_{d=1}^{\hat{D}_k} \frac{\hat{\sigma}_k}{\hat{A}_{d,k}}, \qquad (14)$$

where the $\sqrt{2}$ coefficient in the numerator relies on the assumption that the random uncertainty contributions affecting the frequency data collected from different measurement devices are likely uncorrelated. Consequently, if these contributions have equal variance, they combine in quadrature.

The adoption of the DBSCAN-based approach described above offers significant advantages compared to the basic heuristic method initially used in [30]. Indeed, in the earlier approach, the frequency estimates collected from each measurement point were compared with those of a predefined list of plausible frequencies and a cluster was formed only when more than half of the measurement devices detected

a tone with a frequency within $\pm 1.5\,\text{Hz}$ of one of the assumed frequencies. In this new version instead, no a-priori information is needed, since both the cluster centroids and the clustering threshold are updated adaptively using the collected data. Thus, the proposed DBSCAN-based clustering approach offers three key advantages, i.e.,

- *Improved detection reliability:* the risk of local false detections is reduced, as DBSCAN designates one-point sets as outliers, excluding them from the data fusion step.
- *Enhanced detection sensitivity:* by forming clusters when at least two devices detect the same tone, even minimal consensus among devices leads to tone identification for subsequent data fusion (see Section II-E).
- *Higher flexibility:* the uncertainty-based threshold given by (14) makes the proposed approach more flexible and more suitable in dynamic conditions than using a fixed constant threshold like in [30].

E. Step 4: Multitone Frequency Data Fusion

Once the clusters are formed, the frequency values within each cluster can be merged to obtain a more accurate estimate of the spectral support of the harmonics and interharmonics spreading across the grid. Recalling that the measurement uncertainty of the local frequency estimates is inversely proportional to the SNR (which is also confirmed by the results in [19]), the data fusion of the frequency values within the $i-$th cluster relies on the following weighted average, i.e.,

$$\hat{f}_i = \frac{\sum_{(d,k)\in\mathcal{C}_i} \hat{f}_{d,k} \frac{\hat{A}^2_{d,k}}{\hat{\sigma}^2_k}}{\sum_{(d,k)\in\mathcal{C}_i} \frac{\hat{A}^2_{d,k}}{\hat{\sigma}^2_k}}, \quad i = 1, \ldots, \hat{D} \tag{15}$$

where the index pair $(d,k) \in \mathcal{C}_i$ refers to the $d-$th voltage or current waveform collected by the $k-$th measurement device.

F. Computational Complexity

As already mentioned, the proposed algorithm requires four steps, two of which are executed locally while the others run in the aggregator. Thus, the computational burden affects local devices and the aggregator differently. The computational complexity of the steps 1 and 2 on the $k-$th measurement device depends mainly on i) covariance matrix estimation (whose complexity is $\mathcal{O}(L_k M_k^2)$), ii) SVD computation ($\mathcal{O}(M_k^3)$) and iii) ESPRIT calculation (whose complexity increases cubically with D_k, but just linearly with M_k). Further details on the complexity of these steps can be found in [19].

The data clustering and fusion steps running on the central aggregator are instead much lighter from the computational point of view. The former step depends on DBSCAN, which has a worst-case complexity of $\mathcal{O}(T^2)$ with $T = \sum_{k=1}^{K} \hat{D}_k$. The complexity of (15) in the final step increases linearly with the number of sum of products to compute the weighted average within each cluster. Therefore, in the worst case (namely if all measurement devices detect the same sinusoidal tones), the computational complexity is $\mathcal{O}(K\hat{D})$.

III. Tests and Results

In this Section two kinds of tests are performed. First, the correct operation of the overall algorithm as well as the clustering and data fusion performance are analyzed through MC simulations in different conditions. Afterwards, the results of grid-level simulations of a three-phase MV DSs are reported.

A. Monte Carlo Simulations

The MC simulations assume that $K = 20$ measurement devices are deployed throughout grid. This number of devices (in this section single-phase measurements are considered) is large enough to perform a reliable performance evaluation of the clustering and data fusion steps, and it is also in line with the number of devices resulting from optimal placement strategies in small DSs [36] and with the number of measurements in Section III-B. For the sake of simplicity, but with no loss of generality, the sampling rate of all measurement devices is assumed to be the same (i.e., $f_{s_k} = f_s$, for $k = 1, \ldots, K$). However, to analyze the impact of different sampling rates on algorithm results, three different values were considered in simulations, i.e., 5 kHz, 10 kHz and 15 kHz. These values are well above the Nyquist frequency when low-order harmonics are considered and they are also consistent with the typical sampling rates of different kinds of measurement devices for SG monitoring. As far as the data record size for local multitone detection is concerned, it is advisable to collect and process samples spanning slightly more than one fundamental cycle, as explained in [19]. In this way, the initial phase of the collected waveforms in subsequent records is swept quite linearly between $[0, 2\pi]$, which is essential to estimate (3) with good accuracy, while keeping the computational burden as low as possible. Hence, assuming that the nominal frequency is 50 Hz, $M_k = M = \{134, 268, 401\}$ is set depending on whether f_s is 5 kHz, 10 kHz or 15 kHz, respectively.

Also, the total number of data records L_k used to compute (3) and in (14) is set equal to M_k because the theoretical asymptotic detection limit to detect a signal component emerging from the noise floor is $\sigma_k^2\sqrt{M_k/L_k}$ [34], and this value tends to the actual noise variance if $L_k \approx M_K \to \infty$.

The MC simulations were performed including a linearly increasing number of harmonics from order 2 to 25, which are the main harmonics in the EN Standard 50160:2022 [37]. The harmonics' amplitude in different simulation runs was varied randomly with uniform probability between $0.5\,\%$ and $10\,\%$ of the fundamental amplitude, assuming frequency deviations within $\pm 10\,\%$ of the nominal value. It is important to note that such testing conditions are broader than those compliant with the limits reported in [37]. Specifically, the maximum amplitude and frequency values are the same as those specified in the IEC/IEEE 60255-118-1:2018 for M-class PMUs [38], although the proposed approach does not refer to PMUs, nor it assumes to rely on PMU data only.

Two alternative SNR levels were used for MC simulations, i.e., 50 dB and 60 dB. The corresponding noise levels are rather high, but they align with those adopted in previous studies [39], [40], and allow an effective analysis of the multitone detection capability of the proposed algorithm.

TABLE I
SIMULATION PARAMETERS

Parameter	Symbol	Value
No. of measurement points	K	20
Sampling rates	f_{s_k}	5, 10 or 15 kHz, $\forall k = 1, \ldots, K$
No. of samples in each record	M_k	134, 268 or 401, $\forall k = 1, \ldots, K$
No. of records in each test	L_k	134, 268 or 401, $\forall k = 1, \ldots, K$
No. of sinusoidal tones	D_k	From 2 to 25
Actual frequency values	$f_{d,k}$	$d \cdot 50$ Hz$\mp 10\%$, $d = 1, \ldots, D_k$
Actual relative amp. values	$A_{d,k}$	$[0.5\,\%, 10\,\%]$, $d = 1, \ldots, D_k$
Signal-to-noise ratio	SNR	50 dB or 60 dB
False detection rate	α	$0.1\,\%$

The local acceptable false detection rate α was empirically set to 0.1 %, as this parameter in principle should be as low as possible, but excessively low rates may increase significantly the risk of not detecting harmonics smaller than 1 %.

The MC simulation settings are summarized in Table I. Each simulation was executed 50 times, assuming that K waveforms (including harmonics and wideband noise with known SNR) are sampled at one of the rates specified in Table I over about 5 s. It was not possible to run more than 50 iterations for each test due to the memory constraints of MATLAB workspace. In any case, the total number of data per harmonic is statistically significant (in the order of several thousands after clustering).

The results of several significant simulations are shown in Fig. 2(a)-(d), which consists of 4 pairs of box-and-whisker plots, for different sampling rates and SNR values, i.e., (a) $f_s = 5$ kHz and SNR = 60 dB; (b) $f_s = 10$ kHz and SNR = 60 dB; (c) $f_s = 10$ kHz and SNR = 50 dB; (d) $f_s = 15$ kHz and SNR = 50 dB.

In each subfigure, the upper box-and-whisker plot shows the statistical range of variability of the Macro F1 scores when the number of harmonics to be detected (and therefore the number of clusters) grows for different sampling rates and SNR values. Therefore, these plots highlight the performance of the DBSCAN-based clustering algorithm. The Macro F1 score is a typical indicator of multi-class classification performance. It is computed as the arithmetic mean of the F1 scores over all classes, treating each one of them equally, regardless of its size. Assuming that the actual number of sinusoidal tones is the same throughout the grid (i.e., $D_k = D$, $\forall k$) in the case at hand the Macro F1 score results from

$$\text{Macro F1} = \frac{1}{D}\sum_{d=1}^{D} \text{F1}_d = \frac{1}{D}\sum_{d=1}^{D} 2\frac{\text{Pc}_d \text{Rc}_d}{\text{Pc}_d + \text{Rc}_d} \qquad (16)$$

where $\text{Pc}_d = \frac{\text{TP}_d}{\text{TP}_d + \text{FP}_d}$ and $\text{Rc}_d = \frac{\text{TP}_d}{\text{TP}_d + \text{FN}_d}$ denote the precision and recall in classifying the harmonic of order d, and TP_d, FP_d, and FN_d represent the amount of true positives, false positives, and false negatives, respectively. The upper box-and-whisker plots in Fig. 2(a)-(d) demonstrate that the Macro F1 scores generally remain above 0.97 with no discernible trend and a limited number of outliers, regardless of the number of harmonics, the sampling rate and the SNR values. These results confirm that the customized DBSCAN-based clustering approach is generally effective, robust and scalable, thereby enhancing the already excellent multitone identification performance of steps 1 and 2 of the overall algorithm. Nevertheless, further results, not reported for brevity, show that the Macro F1 scores may suddenly decrease when the number of sinusoidal tones grows and either the SNR or the sampling rate become excessively low. A clue of such behavior is visible in the upper box-and-whisker plot in Fig. 2(a), namely when $f_s = 5$ kHz and 24 or 25 harmonics are considered.

The lower box-and-whisker plots in Fig. 2(a)-(d) display instead the ranges of the frequency estimation errors after data fusion when 25 harmonics are comprised in (1). Observe that except for the fundamental frequency case (i.e., when $d = 1$) whose frequency errors are within ± 1 mHz, the frequency errors of the other harmonics are more than one order of magnitude greater and they may become more than double when the SNR is decreased by 10 dB. Again, the number of outliers is reasonable and apparently the estimation errors are quite independent of both the harmonic order and the sampling rate. This result was expected, because the harmonics are much smaller than the fundamental and their amplitude changes randomly within the same range at all frequencies. In conclusion, the results depicted in Fig. 2(a)-(d) confirm that the data fusion step ensures accurate multitone frequency estimation, even with low SNR or reduced sampling rates.

B. Grid-level Simulations

The MV DS used as a test case for grid-level simulations is shown in Fig. 3. A Real-Time Digital Simulator (RTDS) was used to perform tests in realistic and reproducible operating conditions as well as to foster a future integration with hardware prototypes [41]. In this regard, three scenarios are considered for testing.

- In the first one (called "Scenario A") the behavior of the DS is simulated in a basic configuration, namely, using the data and the topology summarized in the Appendix.
- In the second one ("Scenario B") an additional bank of capacitors storing up to 628 kVAR of reactive power is deployed at node 6.
- In "Scenario C", the reactive power injection at bus 13 is increased by 42.5 kVAR with respect to the baseline.

In all scenarios, the three-phase voltage and current waveforms generated by RTDS are regarded as the reference values. Five measurement points on the MV side of the grid are selected to acquire the waveforms. They are highlighted with red dots in Fig. 3, and they are located at nodes 5, 7, 9, 12 and 14. Each measurement device is supposed to collect the samples of the three-phase bus voltage and line current (indicated by the red arrow in Fig. 3). Therefore, in this case 5 measurement devices are considered, collecting the samples of $K = 15$ voltage and current signals. Based on the results shown in Section II-B, $f_{s_k} = f_s = 10$ kHz and $M_k = L_k = 268$ for $k = 1, \ldots, K$ are set, as these settings ensure excellent clustering performance and high accuracy with a reasonable computational burden. In each scenario, repeated tests were performed and the estimates are updated approximately every fundamental cycle. The duration of each test was about 34 s.

In all scenarios, both sinusoidal disturbances and wideband noise are injected into the grid, In particular,

- Five odd harmonics (i.e., the 3rd, 5th, 7th, 9th and 11th) of varying amplitude are considered in all tests. In some of them an interharmonic component at 75 Hz is added.

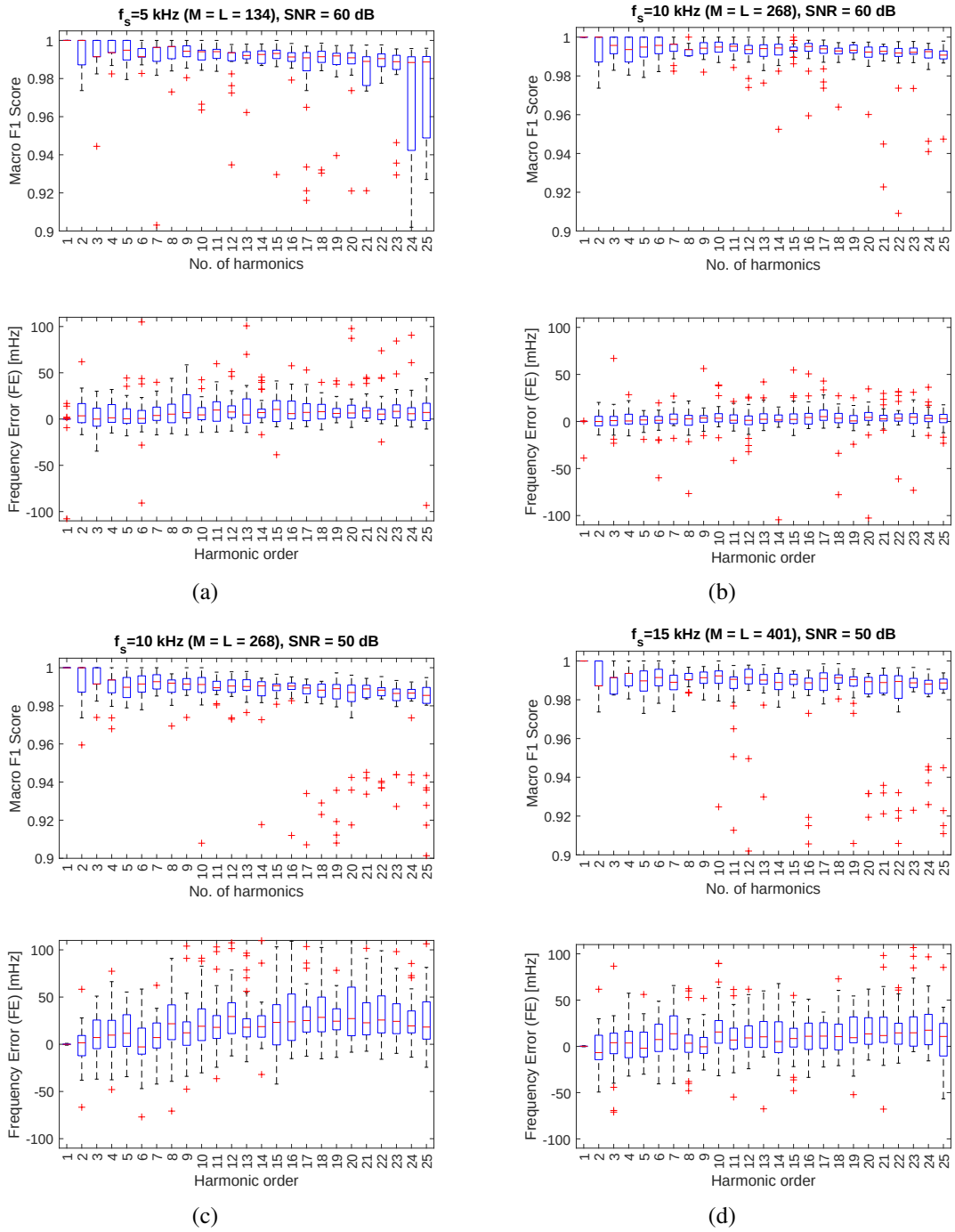

Fig. 2. Box-and-whiskers plots of the macro F1 score, and the harmonic frequency estimation errors as function of the number of harmonics and the harmonic order, respectively, in four cases, i.e., for (a) $f_s = 5$ kHz and SNR = 60 dB; (b) $f_s = 10$ kHz and SNR = 60 dB; (c) $f_s = 10$ kHz and SNR = 50 dB and (d) $f_s = 15$ kHz and SNR = 50 dB.

- The SNR is not always constant throughout the grid, but in all cases it is kept constant over time.

In the first series of tests (i.e., with harmonics only) SNR $= 60$ dB in all measurement points. In Scenario A, the maximum amplitudes relative to the fundamental of the 3rd, 5th, 7th, 9th and 11th voltage and current harmonics are $1.3\,\%$, $1\,\%$, $0.7\,\%$, $0.5\,\%$ and $0.4\,\%$ and $1.4\,\%$, $1\,\%$, $0.65\,\%$, $0.45\,\%$ and $0.3\,\%$, respectively. These values are compatible with the limits specified in the EN Standard 50160:2022 [37], and they are small enough to stress the performance of the clustering step of the algorithm. Note that, in this example, the harmonics amplitude decreases with the harmonic order and that these values are generally smaller than those used to run the MC simulations in Section III-A. Therefore, the range of frequency errors is expected to be larger than in Fig. 2(b). The box-and-whiskers plots of the voltage and current harmonics' frequency estimation errors before and after data clustering and fusion are shown in Fig. 4(a)-(b), respectively. Both in this figure and in the discussion below, the results obtained with the proposed algorithm are referred to as "Data Fusion" results, while the

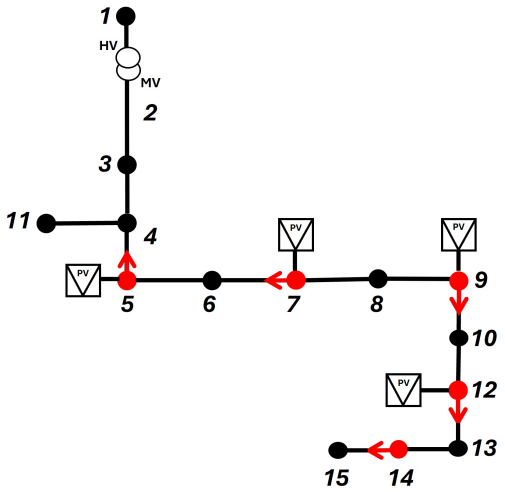

Fig. 3. MV distribution system simulated in RTDS to test the performance of the proposed algorithm under realistic operating conditions.

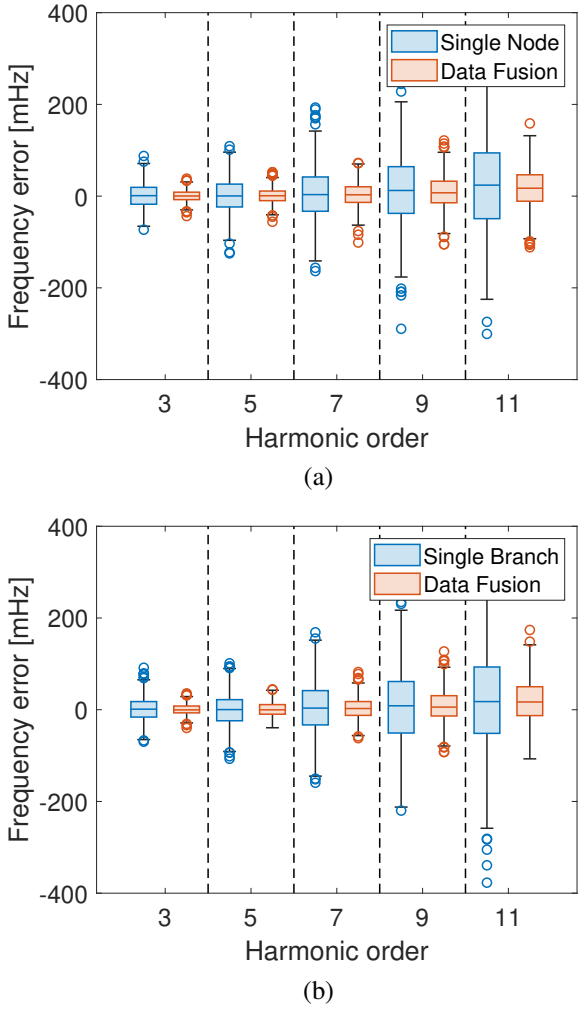

Fig. 4. Frequency estimation errors of five (a) voltage and (b) current harmonics in "Scenario A". The results of the clustering and data fusion algorithm are compared with local frequency estimates at node 9 (for voltage waveforms) and between nodes 14 and 15 (for current waveforms).

local estimates (given simply by the average of the frequency estimates of the three phases) are labeled as "Single Node" or "Single Branch" estimates depending on whether voltage or current measurements are considered. In the case at hand, the "Single Node" frequency errors refer to node 9, while the "Single Branch" frequency errors refer to the line between nodes 14 and 15. However, no significant differences were observed when other nodes or lines are considered.

As a first comment, it is evident that in both cases the proposed data fusion approach significantly improves harmonic frequency estimation accuracy, with no multimodal results. In addition, it is also evident that, as the amplitude of harmonics decreases, the frequency estimation errors grows, as expected due to the dependence of the estimation uncertainty on $\frac{1}{\mathrm{SNR}}$ [35]. Nevertheless, the data fusion strategy significantly lowers the frequency errors, compared to the "Single Node" or "Single Branch" cases. For instance, focusing on the 3rd, 7th and 11th voltage harmonic components, the 75th percentiles of the absolute frequency errors decrease from 30.1, 65.4 and 117.2 mHz to 13.5, 28.9 and 51.8 mHz, respectively, (thus more than halving the errors), while in the case of current harmonics, the 75th percentiles of the absolute frequency errors of the 3rd and 11th harmonics decrease by more than 50 %. In general, the 90 % confidence interval is also more than halved after data clustering and fusion are applied. These considerations can be generalized to every single measurement point, too.

In addition, the results obtained clustering and merging the data collected from a subset of the measurement points confirm that the decrease of the frequency estimation errors achieved through data fusion depends on the quality and quantity of the available information, such as the actual number of measurement points K. Focusing, for instance, on the 9th harmonic for the current waveforms, the 75th percentile of the absolute frequency errors grows from 79.3 mHz to 93.5 mHz when three measurement devices are considered and just the local estimates collected from nodes 5, 9 and 14 are used for data clustering and fusion. Nevertheless, the proposed method, also in this case, demonstrates its applicability and potentiality.

To better illustrate the performance of the proposed approach and, thus, to further generalize the previous conclusions, Fig. 5 shows the same comparison as in Fig. 4(a), but assuming that, in Scenario B, a 75 Hz voltage interharmonic component of relative amplitude 0.25 % is injected at node 2. The new component corresponds to a fractional harmonic order of 1.5 in Fig. 5. The different operating conditions in Scenario B affect the harmonic distortion level and, thus, the obtained estimates. In fact, the frequency estimation uncertainty of the 5th harmonic component is now much larger than before (see Fig. 4(a)). Nevertheless, again the benefits of the clustering and data fusion approach are rather evident. Indeed, considering all the detected components (including the interharmonic one, which is the lowest and most challenging), the average relative reduction of the 75th percentile of the absolute frequency errors is about 69 %.

The performance of the proposed technique was also assessed by computing the errors for each test using, as evalua-

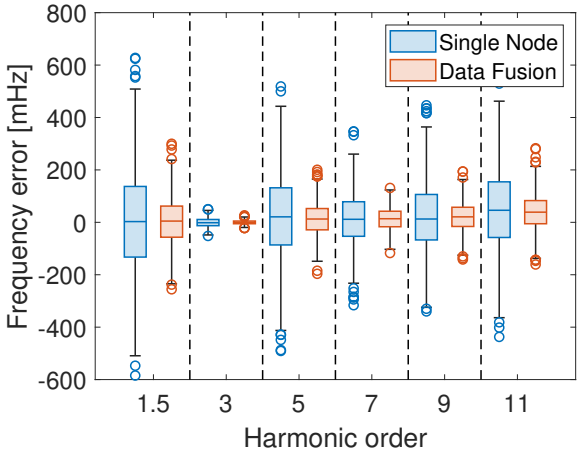

Fig. 5. Frequency estimation errors of five voltage harmonics and one interharmonic component at 75 Hz in "Scenario B". The results of the clustering and data fusion algorithm are compared with the local frequency estimates at node 9.

tion metrics, the root mean square estimation errors (RMSEs) and the mean absolute errors (MAEs) of the frequency of all the detected sinusoidal tones. It is important to recall that the present analysis does not include sources of uncertainty other than the wideband noise of the signal acquisition stage. For instance, the uncertainty contributions of instrument transformers (ITs) are not considered because the aim here is not to evaluate the overall measurement uncertainty, but rather to assess the benefits of data fusion instead of using local measurements only. Furthermore, different ITs could be considered. In a future-proof perspective, Low Power ITs (LPITs) have almost flat and wide bandwidths up to tens or hundreds of kHz [42], [43] and no significant impact on the proposed data fusion is thus expected. The impact of inductive ITs is instead two-fold: on one hand, the frequency response can alter the amplitudes of disturbances, but their frequencies are not affected [27]; on the other hand, small nonlinear effects can create additional disturbances [44]. In both cases, the estimation of the spectral support of voltage and current waveforms must rely on component identification and Data Fusion can significantly help.

Table II illustrates an example of a quantitative comparison of the RMSE and MAE values obtained in Scenario A for different SNR levels at different measurement points. In particular, the SNR of the waveforms collected at nodes 7, 9 and 12 was set to 65 dB, while the SNR of the waveforms collected at nodes 5 and 14 was assumed to be 10 dB lower (i.e., 55 dB). The results obtained with the voltage and current waveforms and considering the data fusion approach are compared, for sake of brevity, with the RMSE and MAE values of the local estimates at nodes 5, 7 and 12 only. The effect of different noise levels at different measurement points is well visible. Indeed, the local frequency estimates related to both voltage and current at node 5 are clearly less accurate than those performed at nodes 7 and 12, as expected. For instance, the RMSE values, as well as the MAE values, of the harmonic frequencies estimated at nodes 7 and 12 (where the SNR is 10 dB

TABLE II
LOCAL AND DATA-FUSION-BASED RMSE AND MAE VALUES OF VOLTAGE AND CURRENT HARMONIC FREQUENCIES IN SCENARIO A, WHEN DIFFERENT SNR LEVELS ARE CONSIDERED (I.E., SNR = 65 dB AT NODES 5, 7 AND 12 AND SNR = 55 dB AT NODES 9 AND 14)

Harmonic order	Component frequency [Hz]	Method	RMSE [mHz] Voltages	RMSE [mHz] Currents	MAE [mHz] Voltages	MAE [mHz] Currents
3	150	Data Fusion	8.25	7.59	6.59	6.07
		Point 5	46.20	43.38	37.02	34.96
		Point 7	14.86	13.65	11.87	10.77
		Point 12	15.18	12.46	12.33	10.01
5	250	Data Fusion	11.85	10.40	9.37	8.39
		Point 5	60.61	57.07	48.71	45.73
		Point 7	20.72	19.10	16.49	15.41
		Point 12	21.14	18.06	17.00	14.33
7	350	Data Fusion	17.19	16.00	13.70	12.76
		Point 5	87.58	93.80	68.46	76.00
		Point 7	29.71	26.85	23.84	21.52
		Point 12	32.25	28.08	25.61	22.42
9	450	Data Fusion	24.83	23.88	19.80	19.00
		Point 5	122.84	135.40	97.76	108.2
		Point 7	40.35	41.12	32.05	32.98
		Point 12	47.59	41.83	38.22	33.64
11	550	Data Fusion	31.05	32.07	24.86	25.56
		Point 5	157.80	184.30	123.94	147.16
		Point 7	48.74	52.99	38.94	41.69
		Point 12	59.07	58.17	47.30	46.85

higher) are about 70 % lower than those at node 5. Again, the advantage of using the proposed data fusion approach is clear. In fact, focusing on the best local RMSE values associated with the detected voltage and current harmonics (i.e., those at node 7 except for the 3rd and 5th current harmonics), the uncertainty reduction due to data fusion ranges from a minimum of 36.3 % to a maximum of 44.5 %, thus proving the ability of the proposed method to weigh properly the local estimates according to their uncertainty. Same conclusions can be drawn considering MAE.

To further analyze the impact of low SNR levels on the proposed data fusion strategy, a test (in the same scenario of Table II and not reported here for sake of brevity) was performed considering the SNR of the waveforms collected at nodes 7, 9 and 12 equal to 60 dB, and the SNR of the waveforms collected at nodes 5 and 14 equal to 50 dB, respectively. In the current waveforms case, the RMSE values obtained with Data Fusion for the 3rd, 7th and 11th harmonic component increase from 7.59, 16.00 and 32.07 mHz (see Table II) to 13.54, 28.94 and 63.39 mHz, with an average increment of about 85 %. However, the proposed method improves the local estimates, on average, by about 83 %, 41 % and 41 % when nodes 5, 7 and 12 are considered, respectively.

A deeper comparison is presented in Table III, where an additional 75-Hz interharmonic interferer of relative amplitude 0.3 % was injected at node 2 in scenarios A, B and C. All the other settings (i.e., about noise and harmonics) are the same as those used to build Table II. As expected, the interharmonic frequency estimation is less accurate than that of the other components. The main reason is again the smaller amplitude of this component as well as the lower spectral distance from the closest tones [35]. Therefore, in such cases, performing

TABLE III
LOCAL AND DATA-FUSION-BASED RMSE AND MAE VALUES OF VOLTAGE AND CURRENT HARMONIC AND INTERHARMONIC FREQUENCIES IN SCENARIOS A, B, AND C WITH DIFFERENT SNR LEVELS (I.E., SNR = 65 dB AT NODES 5, 7 AND 12 AND SNR = 55 dB AT NODES 9 AND 14)

Harmonic order	Component frequency [Hz]	Method	RMSE [mHz] Voltages	RMSE [mHz] Currents	MAE [mHz] Voltages	MAE [mHz] Currents	Harmonic order	Component frequency [Hz]	Method	RMSE [mHz] Voltages	RMSE [mHz] Currents	MAE [mHz] Voltages	MAE [mHz] Currents
Scenario A													
1.5	75	Data Fusion	64.27	55.88	52.19	44.89	7	350	Data Fusion	16.57	15.43	13.24	12.34
		Point 5	404.55	373.88	323.57	298.80			Point 5	84.63	91.49	66.04	73.58
		Point 7	110.75	102.61	89.04	81.68			Point 7	28.41	25.73	22.73	20.45
		Point 12	118.76	97.41	94.62	78.46			Point 12	31.14	26.98	24.63	21.34
3	150	Data Fusion	9.13	7.95	7.34	6.29	9	450	Data Fusion	24.84	23.84	20.07	19.11
		Point 5	51.98	46.04	41.14	36.76			Point 5	124.18	140.12	99.50	113.10
		Point 7	15.59	14.37	12.39	11.50			Point 7	39.60	40.05	31.75	31.77
		Point 12	16.31	13.60	12.96	10.80			Point 12	47.65	41.93	37.74	33.78
5	250	Data Fusion	12.70	11.55	10.09	9.34	11	550	Data Fusion	32.01	33.22	25.56	26.27
		Point 5	67.71	62.38	54.34	50.02			Point 5	163.01	196.98	128.18	156.90
		Point 7	21.85	20.23	17.49	16.20			Point 7	51.24	53.43	40.58	41.65
		Point 12	23.35	19.92	18.66	15.99			Point 12	61.55	60.33	49.10	48.83
Scenario B													
1.5	75	Data Fusion	60.20	54.49	48.33	43.45	7	350	Data Fusion	30.60	26.67	24.51	21.23
		Point 5	363.78	311.22	291.19	248.09			Point 5	136.21	65.09	108.78	51.32
		Point 7	105.17	100.58	84.73	80.82			Point 7	50.03	51.85	39.65	41.86
		Point 12	110.03	96.35	88.35	77.39			Point 12	56.54	50.28	45.27	39.90
3	150	Data Fusion	5.50	4.39	4.40	3.54	9	450	Data Fusion	39.44	36.92	31.71	29.79
		Point 5	32.56	12.90	25.86	10.31			Point 5	182.98	117.12	145.55	94.06
		Point 7	9.15	8.82	7.27	7.00			Point 7	63.08	65.55	50.00	52.24
		Point 12	9.77	8.18	7.75	6.47			Point 12	75.34	67.92	57.99	54.57
5	250	Data Fusion	47.69	23.76	38.04	18.79	11	550	Data Fusion	51.48	52.24	40.56	41.93
		Point 5	159.23	28.37	128.66	22.13			Point 5	221.79	178.30	174.92	141.69
		Point 7	83.85	79.43	67.18	63.99			Point 7	78.90	83.52	62.48	65.43
		Point 12	88.46	75.27	70.90	59.99			Point 12	93.93	92.25	74.69	72.59
Scenario C													
1.5	75	Data Fusion	67.96	57.76	53.88	46.18	7	350	Data Fusion	16.10	14.21	13.07	11.37
		Point 5	415.26	375.61	332.73	301.61			Point 5	84.23	86.81	66.59	69.36
		Point 7	115.99	104.03	90.94	83.47			Point 7	28.22	27.97	22.82	22.34
		Point 12	119.19	98.44	93.51	78.50			Point 12	31.36	25.16	24.89	20.39
3	150	Data Fusion	9.35	7.95	7.38	6.25	9	450	Data Fusion	24.16	22.37	19.04	17.87
		Point 5	48.24	45.55	38.50	36.16			Point 5	116.73	138.98	94.03	109.82
		Point 7	15.50	14.64	12.53	11.66			Point 7	37.75	38.58	30.41	30.64
		Point 12	17.42	12.63	13.85	10.06			Point 12	48.07	40.06	38.34	31.64
5	250	Data Fusion	12.40	10.93	10.03	8.79	11	550	Data Fusion	34.32	30.94	27.12	24.53
		Point 5	65.78	65.76	51.94	52.82			Point 5	161.83	191.9	129.92	151.01
		Point 7	20.95	20.61	16.94	16.39			Point 7	52.48	52.44	41.82	41.56
		Point 12	23.56	18.55	18.85	14.75			Point 12	62.70	57.06	49.38	45.91

data fusion is even more important. In general, it is possible to conclude that the combined effect of low SNR levels and small sine wave amplitudes, although detrimental, makes the benefit of the data fusion approach more evident.

Minor RMSE and MAE differences can be observed when comparing the values of Table II with those of Table III (Scenario A). This is explanatory of the potential of clustering and data fusion; indeed, the presence of an interharmonic interferer does not deteriorate the performance of the method presented in Section II. In all scenarios, this component is detected with an average RMSE of about 60 mHz (and an average MAE of about 48 mHz). On the other hand, in Scenario A, the harmonic frequency RMSE values after data fusion are reduced on average by about $81.4\,\%$ and $83.1\,\%$ compared to the local RMSE values at node 5 for voltage and current waveforms, respectively.

In Scenarios B and C similar conclusions can be drawn, although the actual level of improvement may exhibit significant differences. In Scenario B, the RMSE values after data fusion on average are about $69\,\%$, $44\,\%$ and $46\,\%$ lower than those of local estimates at nodes 5, 7 and 12, respectively, while, for Scenario C, the average RMSE values are instead decreased by $82\,\%$, $42\,\%$ and $44\,\%$, respectively. The varying frequency estimation accuracy of voltage or current harmonics of the same order observed in Scenarios B and C is basically due to the different amplitude of the various sinusoidal components, which are indeed influenced by the specific operating conditions of the simulated grid in either case.

Further tests have been carried out to test the proposed algorithm in the same scenarios and configurations of Table III

when the photovoltaic (PV) generators and inverters of the grid detailed in Appendix work in different operating modes. In particular, the tests were repeated with the inverters working without maximum power point tracking (MPPT). These tests (not reported here for sake of brevity) confirm the results of Table III with only slight variations due to the differences in the operating conditions. It is important to highlight that different inverter models and operation modes can influence the tones to be detected, but they do not affect the general benefits, nor the effectiveness of the proposed method.

Additional specific tests have been carried out also to evaluate the performance of the algorithm when the communication between the measurement devices and the central aggregator is affected by packet losses. A packet loss means that the measurement data record produced by a specific device and corresponding to a given timestamp is lost and, thus, it can not be included in the data fusion stage (15). Synchronization loss or excessive data latency would also result in unusable measurements. Thus, their effect is the same as losing data. To replicate this situation, for each measurement device, a lossy packet transmission was modelled according to the well-known two-state Gilbert-Elliot model [45]. A representative example of the results obtained assuming a $5\,\%$ error rate and an expected average burst length of 2 is discussed below.

The RMSEs achieved in Scenario A for voltage waveforms without packet losses (i.e., like in Table III) are compared in Table IV with the results obtained considering possible packet losses according to the above model. Table IV shows just slight differences in the performance obtained with and without packet losses. Increasing the expected packet error rate leads to an increase of lost measurements and, thus, to performance degradation. In this regard, if the data collected from one or more measurement devices are lost completely, the results become similar to those obtained in the previously presented tests with a varying K.

TABLE IV
IMPACT OF PACKET LOSSES ON DATA-FUSION-BASED RMSE VALUES OF VOLTAGE HARMONIC AND INTERHARMONIC FREQUENCIES IN SCENARIO A WHEN DIFFERENT SNR LEVELS ARE CONSIDERED (I.E., SNR $= 65\,$dB AT NODES 5, 7 AND 12 AND SNR $= 55\,$dB AT NODES 9 AND 14)

Harmonic order	Component frequency [Hz]	RMSE [mHz]	
		No Losses	Losses
1.5	75	64.27	67.03
3	150	9.13	9.74
5	250	12.70	13.34
7	350	16.57	16.96
9	450	24.84	25.72
11	550	32.01	33.12

To evaluate the real-time feasibility of the algorithm, the execution times of the various steps were measured using a notebook equipped with an Intel Core™ i7-11370H CPU running at 3.30 GHz, 16 GB of RAM, and Windows 11. The routines were implemented in MATLAB 2024b with no ad-hoc optimization. The reported execution times are indicative and could potentially be significantly reduced by porting the code to C or C++. Currently, with the settings described at the beginning of this section, the local steps require an average processing time of 12 ms, whereas clustering and data fusion steps take about 1.2 ms, on average. Such execution times confirm the considerations on complexity reported in Section II-F and demonstrate that achieving from tens to a hundred estimates per second is feasible.

IV. CONCLUSIONS

Modern power DSs are nowadays experiencing a large penetration of renewable-based generation and nonlinear loads, which may inject significant harmonics and interharmonics into the grid. However, since these disturbances maintain consistent frequencies across the grid, potentially they can be better identified by merging the measurement data collected in different points. This paper has explored and validated the idea of improving both harmonics and interharmonics detection capability and their frequency estimation accuracy by sharing, clustering and fusing multiple distributed measurement data, thus paving the way to future advanced compensation strategies. The presented technique combines the benefits of an effective multitone detection algorithm based on RMT (followed by an ESPRIT-based frequency and amplitude estimator) with the flexibility of a customized DBSCAN clustering approach supporting the data fusion of the frequency values, which are finally averaged using weights depending on the local measurement uncertainty. One key advantage of the proposed technique is that both the local thresholds for sinusoidal component detection and the clustering range are adjusted adaptively as a function of the collected data, with no need for heuristic or preliminary tuning.

Multiple Monte Carlo and grid-level simulations based on a three-phase MV DS confirm the good robustness and remarkable estimation accuracy of the proposed approach when different sinusoidal tones and noise levels are considered. In addition, a deeper performance analysis reveals that the frequency estimation uncertainty of common voltage and current harmonics can be decreased by more than $50\,\%$ compared with local estimates, even when just a few measurement points are monitored and regardless of the presence of some interharmonics or heterogeneous noise levels. Future research will focus on two main aspects. First, a deeper analysis of the different uncertainty contributions affecting clustering and data fusion is of interest. Afterwards, the role of the proposed data fusion strategy for both harmonic state estimation and distributed harmonic compensation shall be investigated.

APPENDIX
DISTRIBUTION SYSTEM USED AS A TEST CASE

The DS considered in Section III as a test case is an MV three-phase network that was simulated in RTDS considering three types of distribution lines. The parameters of these lines are summarized in (A.1)-(A.3), where the impedance matrices are reported in terms of [Ω/km].

$$\mathbf{Z}_{\mathrm{T1}}=\begin{bmatrix} 1.7099+2.8482i & 0.2970+0.829i & 0.2970+0.829i \\ 0.2970+0.829i & 1.7099+2.8482i & 0.2970+0.829i \\ 0.2970+0.829i & 0.2970+0.829i & 1.7099+2.8482i \end{bmatrix} \quad \text{(A.1)}$$

TABLE V
TOPOLOGICAL INFORMATION OF THE GRID UNDER TEST

Branch	From	To	Type	Length [km]
2	2	3	T1	3.5
3	3	4	T1	1.5
4	4	11	T3	3.5
5	4	5	T1	2
6	5	6	T2	3.5
7	6	7	T1	1.0
8	7	8	T2	6.5
9	8	9	T1	1.75
10	9	10	T1	10
11	10	12	T3	10
12	12	13	T1	10
13	13	14	T1	5
14	14	15	T1	3.6

TABLE VI
ACTIVE AND REACTIVE POWER OF THE LOADS

Node	P [kW]	Q [kVAR]
2	0.0	0.0
3	0.0	0.0
4	−425	−264
5	0.0	0.0
6	0.0	0.0
7	0.0	0.0
8	−587	27.5
9	−582	0.0
10	0.0	0.0
11	0.0	0.0
12	0.0	0.0
13	−86.4	48
14	−86.4	48
15	−62.4	15.3

$$\mathbf{Z}_{\text{T2}}=\begin{bmatrix} 0.9338+1.5554i & 0.1622+0.4528i & 0.1622+0.4528i \\ 0.1622+0.4528i & 0.9338+1.5554i & 0.1622+0.4528i \\ 0.1622+0.4528i & 0.1622+0.4528i & 0.9338+1.5554i \end{bmatrix} \quad \text{(A.2)}$$

$$\mathbf{Z}_{\text{T3}}=\begin{bmatrix} 0.9338+1.3554i & 0.1622+0.4228i & 0.1422+0.4528i \\ 0.1622+0.4228i & 0.9118+1.4554i & 0.1622+0.3528i \\ 0.1422+0.4528i & 0.1622+0.3528i & 0.9168+1.5554i \end{bmatrix} \quad \text{(A.3)}$$

The system includes a High-Voltage (HV)/MV transformer with rated voltages of 110 kV and 20 kV at the primary and secondary side, respectively. The details of the grid topology are shown in Fig. 3, while the nominal load and power injections are reported in Table V and Table VI, respectively.

Additionally, the grid was simulated in RTDS with four PV generators connected to specific nodes with Maximum Power Point Tracking capability, as illustrated in Fig. 3. The average solar irradiance and the environmental temperature are assumed to be $1000\,\mathrm{W/m^2}$ and 25 °C, respectively.

REFERENCES

[1] I. Khan, A. S. Vijay, and S. Doolla, "Nonlinear load harmonic mitigation strategies in microgrids: State of the art," *IEEE Syst. J.*, vol. 16, no. 3, pp. 4243–4255, Sep. 2022.

[2] C.-I. Chen, Y.-C. Chen, Y.-C. Chin, and C.-H. Chen, "Integrated power-quality monitoring mechanism for microgrid," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6877–6885, Sep. 2018.

[3] S. Sepasi, C. Talichet, and A. S. Pramanik, "Power quality in microgrids: A critical review of fundamentals, standards, and case studies," *IEEE Access*, vol. 11, pp. 108 493–108 531, Oct. 2023.

[4] J. Hu, Y. Shan, K. W. Cheng, and S. Islam, "Overview of power converter control in microgrids—challenges, advances, and future trends," *IEEE Trans. Power Electron.*, vol. 37, no. 8, pp. 9907–9922, Mar. 2022.

[5] X. Liang and C. Andalib Bin-Karim, "Harmonics and mitigation techniques through advanced control in grid-connected renewable energy sources: A review," *IEEE Trans. Ind. Appl.*, vol. 54, no. 4, pp. 3100–3111, Apr. 2018.

[6] P. Imgart, M. Beza, M. Bongiorno, and J. R. Svensson, "An overview of grid-connection requirements for converters and their impact on grid-forming control," in *2022 24th European Conf. on Power Electron. and Appl. (EPE'22 ECCE Europe)*, Sep. 2022, pp. 1–10.

[7] F. Nejabatkhah, Y. W. Li, and H. Tian, "Power quality control of smart hybrid AC/DC microgrids: An overview," *IEEE Access*, vol. 7, pp. 52 295–52 318, Apr. 2019.

[8] N. Ashraf *et al.*, "Investigation of the power quality concerns of input current in single-phase frequency step-down converter," *Appl. Sci.*, vol. 12, no. 7, pp. 1–18, Apr. 2022.

[9] W. U. K. Tareen *et al.*, "Mitigation of power quality issues due to high penetration of renewable energy sources in electric grid systems using three-phase APF/STATCOM technologies: A review," *Energies*, vol. 11, no. 6, June 2018.

[10] T. K. Gimenes, M. P. C. da Silva, J. J. G. Ledesma, and O. H. Ando, "Impact of distributed energy resources on power quality: Brazilian scenario analysis," *Elec. Power Syst. Res.*, vol. 211, p. 108249, Oct. 2022.

[11] Z. M. Ali, M. Ćalasan, F. Jurado, and S. H. E. Abdel Aleem, "Complexities of power quality and harmonic-induced overheating in modern power grids studies: Challenges and solutions," *IEEE Access*, vol. 12, pp. 151 554–151 597, Oct. 2024.

[12] I. V. Banu *et al.*, "Passive anti-islanding protection for three-phase grid-connected photovoltaic power systems," *Int. J. of Electr. Power & Energy Syst.*, vol. 148, pp. 1–10, June 2023.

[13] J. Benzaquen, M. Miranbeigi, P. Kandula, and D. Divan, "Collaborative autonomous grid-connected inverters: Flexible grid-forming inverter control for the future grid," *IEEE Electrific. Mag.*, vol. 10, no. 1, pp. 22–29, Mar. 2022.

[14] A. M. dos Santos Alonso, D. I. Brandao, T. Caldognetto, F. P. Marafão, and P. Mattavelli, "A selective harmonic compensation and power control approach exploiting distributed electronic converters in microgrids," *Int. J. of Electr. Power & Energy Syst.*, vol. 115, pp. 1–15, Feb. 2020.

[15] P.-T. Cheng, C.-A. Chen, T.-L. Lee, and S.-Y. Kuo, "A cooperative imbalance compensation method for distributed-generation interface converters," *IEEE Trans. Ind. Appl.*, vol. 45, no. 2, pp. 805–815, Mar. 2009.

[16] A. Borrell, M. Velasco, J. Miret, A. Camacho, P. Martí, and M. Castilla, "Collaborative voltage unbalance elimination in grid-connected AC microgrids with grid-feeding inverters," *IEEE Trans. Power Electron.*, vol. 36, no. 6, pp. 7189–7201, Nov. 2021.

[17] D. Fontanelli and D. Macii, "Accurate time synchronization in PTP-based industrial networks with long linear paths," in *IEEE ISPCS 2024*, Sep. 2010, pp. 97–102.

[18] K. Daniel, L. Kütt, M. N. Iqbal, N. Shabbir, H. A. Raja, and M. U. Sardar, "A review of harmonic detection, suppression, aggregation, and estimation techniques," *Appl. Scien.*, vol. 14, no. 23, pp. 1–25, Oct. 2024.

[19] X. Shan, D. Macii, D. Petri, and H. Wen, "Enhanced IpD2FT-based synchrophasor estimation for M class PMUs through adaptive narrowband interferers detection and compensation," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–14, Jan. 2024.

[20] W. Zhou, O. Ardakanian, H.-T. Zhang, and Y. Yuan, "Bayesian learning-based harmonic state estimation in distribution systems with smart meter and DPMU data," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 832–845, Jan. 2020.

[21] D. Carta, C. Muscas, P. A. Pegoraro, A. V. Solinas, and S. Sulis, "Compressive sensing-based harmonic sources identification in smart grids," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, Jan. 2021.

[22] F. Xu, C. Wang, K. Guo, Q. Shu, Z. Ma, and H. Zheng, "Harmonic sources' location and emission estimation in underdetermined measurement system," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, May 2021.

[23] Q. Shu, H. Wang, C. Wang, and J. Li, "A multiharmonic sources localization algorithm based on ICA and posterior harmonic admittance," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–10, Dec. 2023.

[24] R. Nicolosi, L. Piegari, and A. Benigni, "A smart PV inverter controller with PMU capability," in *2016 Clemson Univ. Power Syst. Conf. (PSC)*, Mar. 2016, pp. 1–7.

[25] M. M. Albu, M. Sănduleac, and C. Stănescu, "Syncretic use of smart meters for power quality monitoring in emerging networks," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 485–492, Jan. 2017.

[26] L. De Oro Arenas, G. de Azevedo e Melo, and C. A. Canesin, "A methodology for power quantities calculation applied to an FPGA-based

smart-energy meter," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, Jan. 2021.
[27] P. Castello, C. Laurano, C. Muscas, P. A. Pegoraro, S. Toscani, and M. Zanoni, "Harmonic synchrophasors measurement algorithms with embedded compensation of voltage transformer frequency response," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, Jan. 2021.
[28] D. D. Giustina *et al.*, "Smart grid automation based on IEC 61850: An experimental characterization," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 8, pp. 2055–2063, Aug. 2015.
[29] S. K. Jain and S. Singh, "Harmonics estimation in emerging power system: Key issues and challenges," *Elec. Power Syst. Res.*, vol. 81, no. 9, pp. 1754–1766, Sep. 2011.
[30] C. Sitzia, D. Macii, and P. A. Pegoraro, "Towards a coordinated spectral support estimation of sinusoidal disturbances in smart grids through synchronized measurements," in *IEEE 14th International Workshop on Applied Measurements for Power Systems (AMPS)*, Sep. 2024.
[31] 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. of the Second International Conference on Knowledge Discovery and Data Mining*, ser. KDD'96. AAAI Press, Aug. 1996, p. 226–231.
[32] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, July 1989.
[33] S. Kritchman and B. Nadler, "Non-parametric detection of the number of signals: Hypothesis testing and random matrix theory," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3930–3941, Oct. 2009.
[34] ——, "Determining the number of components in a factor model from limited noisy data," *Chemom. and Intell. Lab. Systems.*, vol. 94, no. 1, pp. 19–32, Nov. 2008.
[35] P. Stoica and T. Soderstrom, "Statistical analysis of MUSIC and subspace rotation estimates of sinusoidal frequencies," *IEEE Trans. Signal Process.*, vol. 39, no. 8, pp. 1836–1847, Aug. 1991.
[36] A. Aflaki, M. Gitizadeh, A. A. Ghasemi, and K. E. Okedu, "Optimal placement of measuring devices for distribution system state estimation using Dragonfly algorithm," *Math. Probl. in Eng.*, vol. 2022, no. 1, pp. 1–14, Apr. 2022.
[37] *CENELEC Standard - Voltage Characteristics of Electricity Supplied by Public Distribution Systems*, European Committee for Electrotechnical Standardization (CENELEC) CEI EN 50160, Dec. 2022.
[38] *IEEE/IEC International Standard - Measuring relays and protection equipment - Part 118-1: Synchrophasor for power systems - Measurements*, IEC/IEEE 60255-118-1:2018, Dec. 2018.
[39] M. Luiso, D. Macii, P. Tosato, D. Brunelli, D. Gallo, and C. Landi, "A low-voltage measurement testbed for metrological characterization of algorithms for phasor measurement units," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 10, pp. 2420–2433, Oct. 2018.
[40] G. Frigo, A. Derviškadić, A. Bach, and M. Paolone, "Statistical model of measurement noise in real-world PMU-based acquisitions," in *Int. Conf. on Smart Grid Synchr, Meas. and Analytics (SGSMA)*, May 2019.
[41] RTDS Technologies Inc., "Real Time Digital Simulator (RTDS)," https://www.rtds.com.
[42] G&W Electric, "LPITs," https://www.gwelectric.com/products/high-accuracy-sensors/all-products/.
[43] A. Mingotti, L. Peretto, and R. Tinarelli, "A smart frequency domain-based modeling procedure of Rogowski coil for power systems applications," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6748–6755, Apr. 2020.
[44] G. D'Avanzo *et al.*, "Theory and experimental validation of two techniques for compensating VT nonlinearities," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, Feb. 2022.
[45] E. N. Gilbert, "Capacity of a burst-noise channel," *The Bell Syst. Tec. Jour.*, vol. 39, no. 5, p. 1253–1265, Sep. 1960.

David Macii (M' 06, SM' 2014) received the M.S. degree in Electronics Engineering and the Ph.D. degree in Information Engineering from the University of Perugia, Perugia, Italy, in 2000 and in 2003, respectively. He was a visiting researcher at the University of Westminster, London, UK in 2002, at the Advanced Learning and Research Institute of the University of Lugano, Lugano, Switzerland, between 2004 and 2005 and a Fulbright Research Scholar at the Berkeley Wireless Research Center, University of California at Berkeley, Berkeley, USA, between 2009 and 2010. Currently, he is an Associate Professor with the Department of Industrial Engineering of the University of Trento, Trento, Italy. He is author and co-author of about 180 papers published in books, scientific journals, and international conference proceedings. Since 2024 he has been Deputy Editor-in-Chief and Special Issue Editor of the journal "Measurement". His research interests are focused on measurement and estimation techniques based on digital signal processing for mechatronics applications, industrial electronics and power systems.

Paolo Attilio Pegoraro (M'06, SM'19) received the M.Sc. (*summa cum laude*) degree in telecommunication engineering and the Ph.D. degree in electronic and telecommunication engineering from the University of Padova, Padua, Italy, in 2001 and 2005, respectively. From 2015 to 2018 he was an Assistant Professor with the Department of Electrical and Electronic Engineering, University of Cagliari, Cagliari, Italy, where he is currently Associate Professor. He has authored or co-authored over 170 scientific papers. His current research interests include the design of new measurement techniques for modern power networks, with attention to synchronized measurements and state estimation. Dr. Pegoraro is a member of IEEE IMS TC 39 (Measurements in Power Systems) and of IEC TC 38/WG 47. He is also a member of IEEE EPPC WG on Energy. He is an Associate Editor in Chief of the IEEE Transactions on Instrumentation and Measurement and the General Chair of the IEEE International Workshop on Applied Measurements for Power Systems (AMPS).

Carlo Sitzia (Member, IEEE) received the M.S. degree in Electrical Engineering (cum laude) and the Ph.D. degree in Industrial Engineering from the University of Cagliari, Cagliari, Italy, in 2020 and 2024, respectively.

He has been a Postdoctoral Research Fellow with the Electrical and Electronic Measurements Group, Department of Electrical and Electronic Engineering, University of Cagliari. In 2025, he joined Politecnico di Milano as Postdoctoral Research Fellow. His research focuses on parameter estimation and fault location for modern power networks, phasor measurements, and instrument transformers.