

# Improving Real-world Measurement-based Phase Identification in Power Distribution Feeders with a Novel Reliability Criteria Assessment

Milad Izadi, *Student Member, IEEE* and Hamed Mohsenian-Rad, *Fellow, IEEE*

**Abstract**—This paper is concerned with solving the phase identification problem in a real-world smart grid project; where there is only a few smart meters available on each of the five power distribution feeders in the test site in Riverside, CA. The main idea is to develop and use two reliability criteria that can identify the most reliable components in a broken-down phase identification analysis; thereby significantly improving the accuracy of phase identification. The proposed method consists of three steps. The results from field implementation reveal the accuracy and consistency of the proposed method in practice, in correctly and reliability identifying the phase connectivity.

**Keywords:** Phase identification, data-driven method, sliding window, reliability criteria, smart meters, power distribution.

## I. INTRODUCTION

Accurate phase identification is critical in order to properly monitor and operate power distribution networks and to maintain updated three-phase network models [1].

Traditionally, in practice, phase identification is done by manually inspecting the phase connectivity in the field. However, manual inspection is labor-intensive and costly.

An alternative to manual inspection is to use *measurement-based* methods. A popular approach is to examine the correlation across the voltage measurements on different phases, e.g., see [2], [3]. In [4], correlation analysis is done based on linear regression. In [5], a mixed integer programming model is developed to identify phase connectivity using power measurements. In [6], a correlation-based method is designed using the voltage phasor measurements from micro-phasor measurement units. In [7], a k-means clustering algorithm is used for phase identification. In [8], a spectral and saliency analysis is presented to extract the frequency domain features from load profiles for the purpose of phase identification.

In practice, if we have access to *only a few* sensors, then the methods in [2]–[6], such as the use of correlation analysis, are more applicable. On the other hand, the methods in [7], [8], often require access to several sensors; for example to achieve accurate clustering results. However, the problem with correlation-based methods is lack of high accuracy, except for when we have access to phase angle measurements, as in [6]; but this is often *not* the case in practice.

In this paper, we are concerned with solving the phase identification problem in a real-world test site in Riverside, CA. The goal is to significantly improve the accuracy of the measurement-based methods, despite having access to *only a few* smart meters on each power distribution feeder.

The proposed method uses a sliding window to *break-down* the phase identification problem on each day into

The authors are with the Department of Electrical and Computer Engineering, University of California, Riverside, CA, USA; e-mails: {mizadi, hamed}@ece.ucr.edu. This work was supported by DoE grant EE 0008001. The corresponding author is H. Mohsenian-Rad.

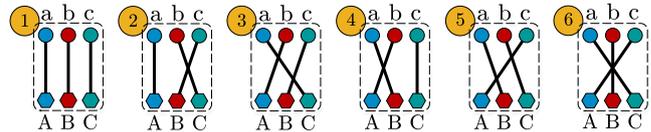


Fig. 1. Six possible phase connection configurations between two three-phase systems. Phases A, B, and C are the known reference phase labeling; and Phases a, b, and c are the unknown phase labeling. The configuration number from 1 to 6 is marked at the top left corner for each configuration.

several smaller problems to help evaluate the reliability of the phase identification analysis. Subsequently, the results are examined through two *reliability criteria*, to identify the most reliable components to make the ultimate phase identification conclusions. The proposed method in this paper is rather general; as it provides a new approach to measurement-based phase identification to *explicitly use reliability assessment* to enhance the results. The proposed method is applied to the real-world measurements from smart meters on five power distribution feeders in Riverside, CA. The results illustrate the high accuracy and consistency of the proposed method in identifying the correct phase connection configuration.

## II. PROBLEM STATEMENT

The objective of the phase identification problem is to determine the phase connectivity for a given bus or a given device in a power distribution network. Phase identification is done in reference to phase labeling in another bus. In a three-phase power system, there are six possible phase configurations; as shown in Fig. 1. The three phases of the bus or device with *unknown* phase connectivity are denoted by Phases a, b, and c. The goal is to identify phase connectivity with respect to a *reference* phase labeling, denoted by Phases A, B, and C. For example, if Phase a of a meter with unknown phase connectivity is connected to Phase C of another meter with known phase connectivity, Phase b is connected to Phase A, and Phase c is connected to Phase B; then the two smart meters are connected according to configuration #3.

### A. Field Experiment

In this paper, our focus is on *measurement-based* phase identification. Here, phase identification is done solely by examining the voltage measurements at the meter with unknown phase labeling and comparing them with the voltage measurements at the meter with reference phase labeling.

The real-world field measurements that are used in this study are collected from 24 smart meters across five power distribution feeders in Riverside, CA. An example for the three-phase voltage measurements at two smart meters are shown in Fig. 2 over a period of 24 hours. Both meters are on the *same* power distribution feeder. The objective of the

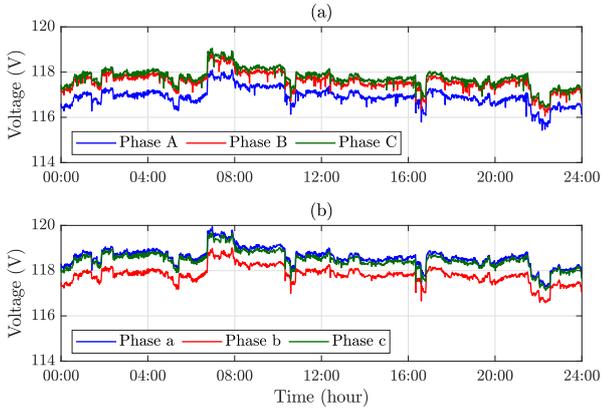


Fig. 2. The time series of three-phase voltage measurements on April 5, 2020 by two smart meters on the same distribution feeder: (a) smart meter 2 which acts as the reference; (b) smart meter 1 with unknown phase labeling.

phase identification in this example is to determine the phase connection configuration between these two smart meters.

Ultimately, our goal is to do phase identification for *all* the smart meters that are on the *same* power distribution feeder.

### B. Baseline Method

A common approach in practice to solve the measurement-based phase identification problem is to examine and *maximize* the *correlations* between the two sets of time-series. They include the voltage measurements on Phases a, b, and c, at the meter with unknown phase labeling, denoted by  $v_1^a(t)$ ,  $v_1^b(t)$ , and  $v_1^c(t)$ , respectively. They also include the voltage measurements on Phases A, B, and C, at the meter with reference phase labeling, denoted by  $v_2^A(t)$ ,  $v_2^B(t)$ , and  $v_2^C(t)$ , respectively. In this regard, let us define the three-phase correlation coefficient between the three-phase voltage measurements at smart meter 1 with unknown phase labeling a, b, c and the three-phase voltage measurements at smart meter 2 with the reference phase labeling A, B, C as follows:

$$\Psi_{a,b,c} = \frac{1}{3} \left\{ \rho\left(v_1^a(t), v_2^A(t)\right) + \rho\left(v_1^b(t), v_2^B(t)\right) + \rho\left(v_1^c(t), v_2^C(t)\right) \right\}, \quad (1)$$

where  $\rho(\cdot, \cdot)$  returns the correlation coefficient between two *single-phase* voltage measurements. It should be noted that, (1) is calculated over a time interval denoted by  $T$ , e.g.,  $T = \text{one day} = 24 \text{ hours}$ . We can similarly obtain:

$$\Psi_{a,c,b}, \Psi_{c,a,b}, \Psi_{b,a,c}, \Psi_{b,c,a}, \Psi_{c,b,a}. \quad (2)$$

The three-phase correlation coefficients in (1) and (2) are corresponding to the six different phase connection configurations in Fig. 1. According to the baseline method, we can identify the phase connection configuration of the smart meter with the unknown phase labeling as follows:

$$\{a^*, b^*, c^*\} = \arg \max_{i,j,k \in \{a,b,c\}} \Psi_{i,j,k}. \quad (3)$$

Putting it simply, in the baseline method, the correct phase connection configuration between two meters is the configuration in which the three-phase correlation is maximized.

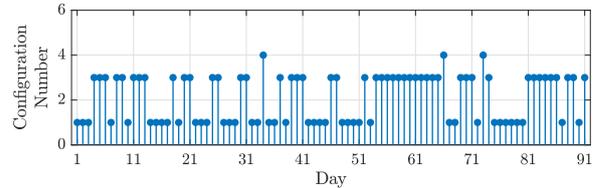


Fig. 3. The connection configuration number between smart meters 1 and 2 that is reported by the baseline phase identification method on a daily basis over a period of three months, from April 1, 2020 till June 30, 2020.

### C. The Challenge

Again consider the real-world three-phase voltage measurements in Fig. 2. Suppose we apply the baseline method in (1)-(3) in Section II-B to conduct phase identification on a daily basis for three months. The results are shown in Fig. 3. We can see that, the baseline method identifies Configuration #1 as the phase identification solution on 41 days. The baseline method also identifies Configuration #3 on 47 days. On the remaining 3 days, the baseline method identifies Configuration #4 as the phase identification solution. Noted that, the correlation coefficients are calculated for the entire day, i.e., over  $T = 24 \text{ hours}$ . Accordingly, the baseline method identifies one configuration for each day.

From the results in Fig. 3, it is clear that the baseline method results in significant *inconsistency* in the phase identification solution. Importantly, the phase connectivity did *not* change from one day to another. The *true phase connectivity remained the same* throughout this three-month period. Yet, the baseline method reported different phase configurations on different days. This makes the results useless; because the operator remains unclear about the correct phase connectivity.

## III. PROPOSED METHOD

In this section, we use real-world measurements to make the case that, the poor performance of the conventional baseline method is due to the *variable reliability* of the analysis in (1)-(3). That is, the phase identification results that are obtained based on (1)-(3) are *not* always reliable. Therefore, if we can somehow develop a data-driven method that properly evaluates the reliability of the results in (1)-(3), then we can ultimately develop a phase identification method that makes use of *only the most reliable cases*; thus providing us with a much better phase identification performance.

The development of our new method involves three steps. First, we break-down the phase identification problem into several small window sizes. This will allow us to examine the analysis in (1)-(3) in a significantly higher granularity. Second, we introduce two reliability criteria that determine how reliable the analysis in (1)-(3) is in each small window during the day. Accordingly, we also introduce a threshold-based mechanism to discard the results from unreliable time windows. Finally, we obtain the final phase configuration for the entire day from the selected reliable time windows.

#### A. Step 1: Granular Window-based Analysis

Suppose a sliding window with time-duration  $\Delta$ , where  $\Delta \ll T$ , and a speed of  $\gamma$  moves along the time series. Given the duration of the time series  $T$ , the number of windows is:

$$N = \left\lceil \frac{T - \Delta}{\gamma} \right\rceil + 1, \quad (4)$$

where  $[\cdot]$  returns the integer part. For example, if we assume that the duration of the time window is one hour, i.e.,  $\Delta = 1$  hour, and the window moves every 30 minutes, i.e.,  $\gamma = 0.5$  hour; then the total number of windows that can cover the entire day, i.e.,  $T = 24$  hours, is  $N = 47 = [(24-1)/0.5] + 1$ . Noted that, the first window is from time 00:00 to time 01:00; and the last window is from time 23:00 to time 24:00.

In Step 1, we use the baseline method in Section II-B to identify the phase connection configuration at each time window. Let  $w$  denote the  $w$ th window, where  $w \in \mathcal{N} = \{1, \dots, N\}$ . We start from the first window and solve the problem in (1)-(3) for *each window*, all the way to the last window on each day. It should be noted that, in each window, the correlation analysis is done over the length of the window, i.e., which is equal to  $\Delta$ . The phase configuration for every window, i.e., from window 1 to window  $N$ , is obtained as

$$\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_w, \dots, \mathcal{K}_N, \quad (5)$$

respectively. Here,  $\mathcal{K}_w$  denotes the configuration number at window  $w$ ; and it takes a number from 1 to 6, as in Fig. 1.

Fig. 4(a) shows the configuration number at each window; based on the same two sets of three-phase voltage measurements that we saw in Fig. 2 in Section II-B. We can still see some inconsistencies even within the smaller windows on the same day. This indicates that solving the phase identification problem in smaller windows does *not* necessary improve the consistency of the results. This calls for scrutinizing the results in each small window in order to select only the most reliable cases, as we will explain in the next step.

### B. Step 2: Threshold-Based Reliability Assessment

In this step, we examine the phase identification results at each time window in order to identify the most reliable windows. We do so by introducing two criteria.

**Criteria 1:** First, recall from Section II-B that the baseline method works by *maximizing* the three-phase correlation coefficient; i.e., by solving the optimization problem in (3). Accordingly, for each window  $w$ , let  $\Psi_w$  denote the *maximum* three-phase correlation coefficient that is obtained by solving the problem in (3). On each day, we can obtain

$$\Psi_1, \Psi_2, \dots, \Psi_w, \dots, \Psi_N. \quad (6)$$

The value of  $\Psi_w$  is one possible index for the reliability of the phase identification result in window  $w$ . A higher  $\Psi_w$  means a higher correlation among the voltage measurements at the selected phase connection configuration during window  $w$ . Therefore, as the first criteria in Step 2, we consider the phase identification result at time window  $w$  to be reliable *only if* the following inequality holds:

$$\Psi_w > \psi_{\text{Threshold}}, \quad (7)$$

where  $\psi_{\text{Threshold}}$  denotes the *correlation threshold*, i.e., the minimum required value for  $\Psi_w$  in order for the phase identification result at window  $w$  to be acceptable, as far as the maximum achieved correlation coefficient is concerned.

**Criteria 2:** Next, at each time window  $w$ , we look at the *difference* between the *highest* three-phase correlation coefficient and the *second highest* three-phase correlation coefficient. Recall from Criteria 1 that  $\Psi_w$  denotes the *highest* three-phase correlation coefficient that is obtained by solving

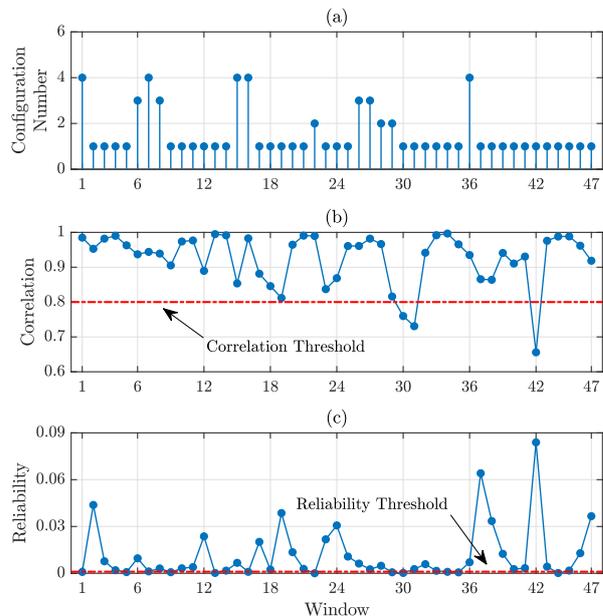


Fig. 4. The results of the phase identification between smart meters 1 and 2 over 47 time windows on April 5, 2020: (a) the connection configuration number that is obtained by solving the maximization problem in (3); (b) correlation index; (c) reliability index. The correlation threshold and the reliability threshold are marked with red horizontal dashed lines.

problem (3) at time window  $w$ . Let  $\Phi_w$  denote the *second highest* three-phase correlation coefficient. We define:

$$\mathcal{R}_w = \Psi_w - \Phi_w. \quad (8)$$

On each day, we can obtain:

$$\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_w, \dots, \mathcal{R}_N. \quad (9)$$

The value of  $\mathcal{R}_w$  is another index for the reliability of the phase identification result in window  $w$ . A higher  $\mathcal{R}_w$  means a higher *distinction* between the connection configuration that is selected by solving the problem in (3) versus the connection configuration that has the second highest three-phase correlation coefficient. Thus, as the second criteria in Step 2, we consider the phase identification result at window  $w$  to be reliable *only if* the following inequality holds:

$$\mathcal{R}_w > R_{\text{Threshold}}, \quad (10)$$

where  $R_{\text{Threshold}}$  denotes the *reliability threshold*, i.e., the minimum required value for  $\mathcal{R}_w$  in order for the phase identification result at window  $w$  to be acceptable, as far as the margin of the maximization in problem (3) is concerned.

Figs. 4(b) and (c) show the correlation and the reliability at each time window corresponding to the connection configurations in Fig. 4(a). As we can see, the reliability index is quite high at the time windows where the configuration is #1, which is the correct configuration. However, the reliability index is quite low at the time windows where the configuration is *not* #1. The correlation is mostly high at every window. This shows the importance of using *both* Criteria 1 and Criteria 2 in selection of the most reliable time windows.

### C. Step 3: Final Configuration Determination

On each day, let us define set  $\mathcal{S}$  as the set of all time windows that satisfy *both* Criteria 1 and Criteria 2. For example, if we assume that the correlation threshold is  $\psi_{\text{Threshold}} = 0.80$  and the reliability threshold is  $R_{\text{Threshold}} = 0.001$  for the

---

**Algorithm 1** Phase Identification Method

---

```
1: // Step 1:
2: for each window  $w \in \mathcal{N} = \{1, \dots, N\}$  do
3:   Solve the maximization problem in (1)-(3).
4: end for
5: // Step 2:
6:  $\mathcal{S} = \{\}$ .
7: for each window  $w \in \mathcal{N}$  do
8:   if correlation criteria in (7) holds then
9:     if reliability criteria in (10) holds then
10:       $\mathcal{S} = \mathcal{S} \cup \{w\}$ .
11:   end if
12: end if
13: end for
14: // Step 3:
15: Obtain the final configuration  $\mathcal{K}^*$  using (12)
16: Obtain  $\Psi^*$  and  $\mathcal{R}^*$  using (13) and (14), respectively.
```

---

analysis in Fig. 4, then the number of the windows that satisfy both Criteria 1 and Criteria 2 is 34. The set of the selected time windows in this example is obtained as

$$\mathcal{S} = \{2, 3, 4, 6, 7, 8, 10, 11, 12, 14, 15, 17, \dots, 21, 23, \dots, 28, 32, 33, 36, \dots, 41, 43, 45, 46, 47\}. \quad (11)$$

We obtain the final connection configuration for the entire day from the results during the time windows in set  $\mathcal{S}$ :

$$\mathcal{K}^* = \text{mode}\{\mathcal{K}_w \mid w \in \mathcal{S}\}, \quad (12)$$

where  $\text{mode}\{\cdot\}$  returns the configuration number that appears most among the time windows in set  $\mathcal{S}$ . For example, for the analysis in Fig. 4, we have  $\mathcal{K}^w = 1$  in 26 time windows (76%);  $\mathcal{K}^w = 2$  in one time window (3%);  $\mathcal{K}^w = 3$  in four time windows (12%); and  $\mathcal{K}^w = 4$  in three time windows (9%). Thus, we obtain  $\mathcal{K}^* = 1$ . In this example, although we initially started with inconsistent phase identification results at different time windows, as we saw in Fig. 4(a), the analysis of the correlation index in Fig. 4(b) and the reliability index in Fig. 4(c), resulted in focusing only on the time windows in set  $\mathcal{S}$ , where the phase identification results are *very consistent*.

Also, we obtain a representative correlation and a representative reliability for the final configuration:

$$\Psi^* = \text{mean}\{\Psi_w \mid \mathcal{K}_w = \mathcal{K}^*, w \in \mathcal{S}\}, \quad (13)$$

$$\mathcal{R}^* = \text{mean}\{\mathcal{R}_w \mid \mathcal{K}_w = \mathcal{K}^*, w \in \mathcal{S}\}, \quad (14)$$

where  $\text{mean}\{\cdot\}$  returns the mean value. For the example in Fig. 4, we have:  $\Psi^* = 0.93$  and  $\mathcal{R}^* = 0.016$ .

#### D. Algorithm

The process in Steps 1-3 is summarized in Algorithm 1. It is worth pointing out that, the approach in Algorithm 1 is rather general and it may work even if we replace the baseline method in (1)-(3) with other phase identification methods.

## IV. REAL-WORLD CASE STUDIES

In this section, we assess the performance of the proposed phase identification method using real-world smart meter data on a test site with 24 smart meters that are installed across 5 power distribution feeders, as we explained in Section

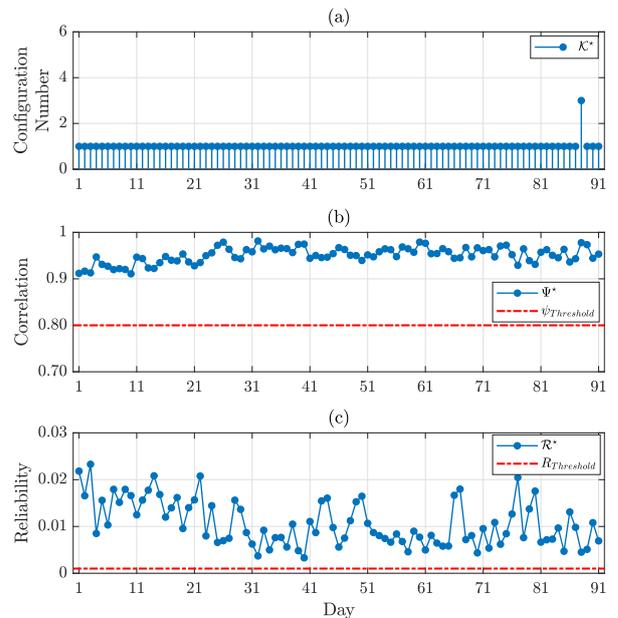


Fig. 5. The results of the phase identification between smart meters 1 and 2 over three months using the proposed method: (a) configuration number; compare it with Fig. 3; (b) correlation; (c) reliability.

II-A. The measurements are obtained over a period of three months, from April 1, 2020 till June 30, 2020. Unless stated otherwise, we assume that the duration of each window is one hour, i.e.,  $\Delta = 1$  hour, and the window moves every half an hour, i.e.,  $\gamma = 0.5$  hour. Also, unless stated otherwise, we assume that  $\psi_{\text{Threshold}} = 0.80$  and  $R_{\text{Threshold}} = 0.001$ .

#### A. Performance Comparison

Again consider the real-world three-phase voltage measurements in Fig. 2. Recall that the performance of the baseline method was *very inconsistent*, as we saw in Fig. 3. Next, we apply the proposed method in Algorithm 1 to repeat phase identification across the same two smart meters. The results are shown in Fig. 5. As we can see, the results are *very consistent* during the entire three months. Algorithm 1 always results in connection configuration #1 except on one day. This indicates 98.9% accuracy. Also, the correlation is always over 0.9, and the reliability is always more than 0.001. These results confirm the accuracy and consistency of the proposed method in identifying the correct phase configuration with high correlation and high reliability. On the contrary, as we saw in Fig. 3, the baseline method identifies Configuration #1 on only 41 days out of three months, which is 45% accuracy.

Table I shows the accuracy of phase identification results using the baseline method and the proposed method across the five feeders. As we can see, the proposed method has always higher accuracy compared with the baseline method across the five feeders. Notice that, the *worst-case* accuracy of the proposed method is 87.9%; while the *worst-case* accuracy of the baseline method is 39.6%; and there are also other cases where the accuracy of the baseline method is as low as 45.1% and 54.9%. It should be noted that, the distance of the smart meters affect the performance of phase identification. When their distance is higher, the performance of the baseline method is poor; yet the performance of the proposed method is still very good. For example, consider smart meters 20 and 23 that are installed with some distance from each other. The

TABLE I  
COMPARING PHASE IDENTIFICATION ACCURACY

Feeder	Smart Meters	Baseline Method	Proposed Method
1	1 and 2	45.1%	98.9%
	1 and 3	100%	100%
	1 and 4	100%	100%
	1 and 5	39.6%	87.9%
2	6 and 7	98.9%	100%
	6 and 8	98.9%	98.9%
	6 and 9	98.9%	100%
	6 and 10	61.5%	100%
3	11 and 12	93.4%	100%
	13 and 14	100%	100%
4	13 and 15	100%	100%
	13 and 16	100%	100%
	13 and 17	100%	100%
	13 and 18	100%	100%
	13 and 19	100%	100%
5	20 and 21	96.7%	100%
	20 and 22	100%	100%
	20 and 23	54.9%	91.2%
	20 and 24	100%	100%
<b>Average</b>		<b>88.8%</b>	<b>98.9%</b>

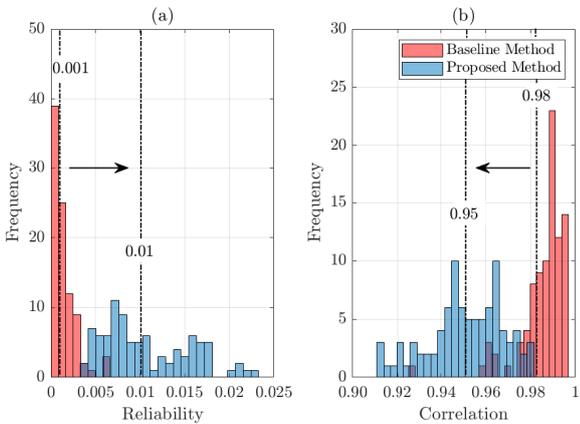


Fig. 6. Distribution of (a) the reliability; (b) the correlation of phase identification results using the baseline method and the proposed method. The means of the distributions are marked with dashed lines.

accuracy of the baseline method is 54.9%, and the accuracy of the proposed method is 91.2%, which is 66.1% improvement.

### B. Enhanced Reliability

Fig 6 shows the distribution of the correlation and the reliability of phase identification results using the baseline method and the proposed method. We can see that the proposed method has much higher reliability but slightly lower correlation, compared with the baseline method. The average reliability and the average correlation over the entire three-months period are marked on the dash lines. We can conclude that, not only the phase identification results are much more accurate as we saw in Section IV-A, the results demonstrate also a much higher reliability and robustness.

### C. Sensitivity Analysis

1) *Impact of Parameters of Sliding Window:* Table II shows the accuracy of Algorithm 1 when we try different choices of window duration  $\Delta$  and window speed  $\gamma$ . Note that, the window speed must always be smaller than the window duration in order to cover the entire day. Therefore, the window speed is set based on the different factors of window duration in Table II. As we can see, the high accuracy is achieved when the duration of the window is small. The

TABLE II  
IMPACT OF THE PARAMETERS OF THE SLIDING WINDOW

Window Speed: $\gamma$	Window Duration: $\Delta$ (hour)			
	1	2	4	8
$1/8 \times \Delta$	98.9%	96.7%	92.3%	79.1%
$1/4 \times \Delta$	97.8%	96.7%	94.5%	80.2%
$1/2 \times \Delta$	98.9%	95.6%	95.6%	76.9%
$1 \times \Delta$	98.9%	97.8%	93.4%	76.9%

TABLE III  
IMPACT OF THRESHOLDS IN CRITERIA 1 AND 2

Reliability threshold: $R_{\text{Threshold}}$	Correlation Threshold: $\psi_{\text{Threshold}}$			
	0.80	0.85	0.90	0.95
0.001	98.9%	98.9%	98.9%	98.9%
0.005	98.9%	98.9%	97.8%	95.6%
0.01	100%	100%	98.9%	81.3%
0.02	90.1%	87.9%	82.4%	25.3%

results are poor for longer windows. Also, the accuracy slightly changes based on the speed of the window. We can conclude that the proposed phase identification method works well for a sliding window that has a short duration.

2) *Impact of Thresholds in Criteria 1 and 2:* Table III shows the accuracy of Algorithm 1 when we try different correlation threshold  $\psi_{\text{Threshold}}$  and different reliability threshold  $R_{\text{Threshold}}$ . As we can see, the higher accuracy is achieved when both  $\psi_{\text{Threshold}}$  and  $R_{\text{Threshold}}$  are small. The results are poor for larger thresholds. We can conclude that the proposed phase identification method works well for small thresholds.

## V. CONCLUSIONS

A new data-driven method was proposed to solve the phase identification problem based on real-world measurements from a small number of smart meters. In this method, a cross correlation analysis is done across different time intervals using a sliding window; and then a reliability assessment is conducted using two threshold-based criteria in order to make conclusions based on the most reliable segments in the available data. The accuracy, reliability, and the sensitivity to the choice of parameters were tested across five real-world power distribution feeders. On average, the accuracy in phase identification improves from 88.8% to 99.0%. The worst-case accuracy improves drastically from 39.6% to 87.9%.

## REFERENCES

- [1] S. J. Pappu, N. Bhatt, R. Pasumarthy, and A. Rajeswaran, "Identifying topology of low voltage distribution networks based on smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5113–5122, Sep. 2018.
- [2] H. Pezeshki, , and P. J. Wolfs, "Consumer phase identification in a three phase unbalanced LV distribution network," in *IEEE PES ISGT Europe*, Berlin, Germany, 2012, pp. 1–7.
- [3] W. Luan, J. Peng, M. Maras, J. Lo, and B. Harapnuk, "Smart meter data analytics for distribution network connectivity verification," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1964–1971, Jul. 2015.
- [4] T. A. Short, "Advanced metering for phase identification, transformer identification, and secondary modeling," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 651–658, Jun. 2013.
- [5] V. Arya, D. Seetharam, S. Kalyanaraman, K. Dontas, C. Pavlovski, S. Hoy, and J. R. Kalagnanam, "Phase identification in smart grids," in *IEEE SmartGridComm*, Brussels, Belgium, 2011, pp. 25–30.
- [6] M. H. F. Wen, R. Arghandeh, A. von Meier, K. Poolla, and V. Li, "Phase identification in distribution networks with micro-synchrophasors," in *IEEE PES General Meeting*, Denver, CO, USA, 2015.
- [7] F. Olivier, A. Sutura, P. Geurts, R. Fonteneau, and D. Ernst, "Phase identification of smart meters by clustering voltage measurements," in *Power Syst. Comput. Conf. (PSCC)*, Dublin, Ireland 2018, pp. 1–8.
- [8] M. Xu, R. Li, and F. Li, "Phase identification with incomplete data," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2777–2785, Jul. 2018.