A Data-Driven Time-Synchronization Method to Convert Power Quality Waveform Measurements into Synchro-Waveforms

Zong-Jhen Ye, *Student Member, IEEE* and Hamed Mohsenian-Rad, *Fellow, IEEE*

***Abstract*—Synchronized waveform measurements, also known as *synchro-waveforms*, are gaining increasing attention in advanced power systems monitoring in recent years. However, there is a major gap in this field in practice. Although many utilities *do* have an existing infrastructure to measure voltage and current waveforms, such as by using conventional power quality sensors, those existing waveform measurements are *not* time-synchronized. In this paper, we propose novel methods to achieve *data-driven* time-synchronization among the conventional power quality sensors solely based on the analysis of the time-series of waveform measurements using statistical and optimization techniques. This will enable utilities to take advantage of the emerging applications of synchro-waveforms *without* the need to replace or retrofit their existing sensors. Experimental results confirm the high performance of the proposed methods.**

***Keywords*— Synchro-waveforms, synchronized power system waveforms, real-world data, power quality sensors, data-driven method, time synchronization, statistical analysis, optimization.**

I. Introduction

Synchro-waveforms, i.e., time-synchronized waveform measurements, have emerged recently as a new measurement technology with diverse applications in power system monitoring and situational awareness; e.g., see the overview in [1]. The device to measure synchro-waveforms is often referred to as a Waveform Measurement Unit (WMU). The common approach to achieve precise time-synchronization among WMUs is to use the Global Positioning System (GPS); see [2].

A. Motivation: A Practical Challenge

WMU installations are still very rare in practice. In fact, although there is already a wide range of sensor technology deployments by many utilities that *do* provide voltage and current waveform measurements, such measurements are *not* time synchronized. One of such technologies that is currently widely used in practice is the power quality sensor. The majority of the existing power quality sensors that have been deployed by many utilities over the past two decades are *not* equipped with precise time-synchronization capabilities such as GPS; due to the high cost. In fact, it was observed in a recent study in [3] that the time stamps from power quality sensors can be off by a few milliseconds to a few seconds.

As a result, while many utilities *do* have various existing power quality sensors to measure voltage and current waveforms, they are *unable* to use such existing measurements as synchro-waveforms due to lack of proper time-synchronization. Thus, most utilities are still currently *unable* to support the emerging applications of synchro-waveforms.

The authors are with the University of California, Riverside, CA, USA. The corresponding author is H. Mohsenian-Rad. E-mail: hamed@ece.ucr.edu.

B. Our Approach

In this paper we seek to address the above open problem. Specifically, we seek to answer the following question: *is it possible to take the waveform measurements from the existing power quality installations that are not time-synchronized and somehow convert them through an automated process into synchro-waveforms with some reasonable time accuracy, solely by analyzing the waveform measurements themselves?*

To answer the above question, we will develop novel methods to *estimate* an unknown *synchronization factor* between two power quality sensors. Our methods are *data driven* and based on real-world measurements. They utilize the measurement data itself, such as the shape of the captured waveforms.

C. Related Literature

Two recent overviews of the emerging field of synchronized waveform measurements are available in [1], [2]. The link to a related IEEE Task Force are available in [4]. Some recent applications of synchro-waveforms include: event and fault location identification [5]–[7], wildfire monitoring [3], power system oscillation monitoring [8], and dynamic modeling of the sub-cycle behavior of inverter-based resources [9].

The use of GPS to time-synchronize waveform measurements is discussed in [2]. The use of time-synchronization protocols and their communications requirements are discussed in [10]. Achieving time-synchronization when GPS signal is momentarily lost by WMUs is discussed in [11] based on the analysis of events and circuits models under some specific network topologies on a power distribution system.

II. Problem Statement

Consider two power quality sensors on a power network. They record voltage and current waveforms during *power quality events*. Accordingly, the waveform measurements that are captured by these two sensors are *event-triggered* [12, Section 4.2]. In practice, a wide range of events can trigger waveform measurements, such as voltage sags and swells, faults, and changes in harmonics. The two power quality sensors are *not* time synchronized. That means, although each sensor has a *local clock* to time-stamp its own measurements, the two local clocks are *not* synchronized with each other.

The objective in this paper is to examine the various waveform measurements that are recorded by the these two power quality sensors to estimate the offset among their local clocks and hence to time-synchronize their waveform measurements.

Next, we will break down the above problem into two specific sub-problems. We will develop methods to solve these two sub-problems in Sections III and IV, respectively.

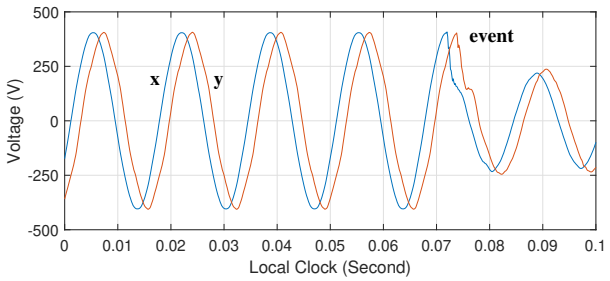

Fig. 1. Two time-series of voltage waveform measurements of the same phase of the power system but from two different power quality sensors at two different locations. The two sensors are *not* time synchronized.

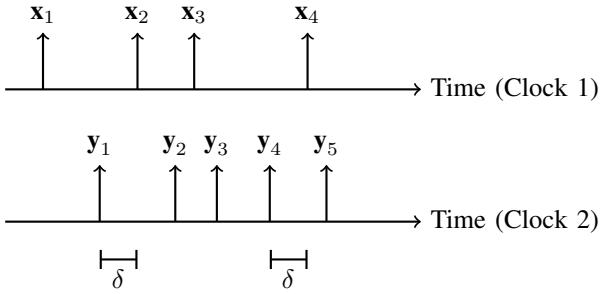

Fig. 2. Top arrows indicate the timing of the events and the corresponding time-series of the waveform captures at the first power quality sensor. Bottom arrows indicate the timing of the events and the corresponding time-series of the waveform captures at the second power quality sensor. The amount of the time-synchronization factor between the two sensors is shown by δ.

A. Problem 1

Consider the following two time-series of waveform measurements that are recorded by the two power quality sensors during the same physical event in the power system:

$$\begin{aligned} \mathbf{x} &= x[1], \ldots, x[n], \\ \mathbf{y} &= y[1], \ldots, y[n], \end{aligned} \tag{1}$$

where n is the number of measurement samples. An example is shown in Fig. 1. Only one phase and only voltage waveforms are shown in this figure. Both measurements have the same sampling rate at 128 samples per cycle. Six cycles of waveform measurements are shown here; therefore, we have $n = 6 \times 128 = 768$ samples. The two measurements are *not* time synchronized. In fact, there is $\delta = 2$ milliseconds offset between the two measurements. We refer to δ as the *synchronization factor* between the two measurements.

If δ is known, then we can shift one of the two time-series $\mathbf{x}$ and $\mathbf{y}$ by δ and turn the two time-series into synchro-waveforms. However, the challenge is that δ is *not* known.

Accordingly, Problem 1 is concerned with estimating δ from all the available pairs of waveform measurements $\mathbf{x}$ and $\mathbf{y}$.

B. Problem 2

In Problem 1, we inherently assumed that we know which pair of time-series of waveform measurements from the two power quality sensors correspond to the *same* physical phenomena/event. In other words, we assumed that $\mathbf{x}$ and $\mathbf{y}$ are already known to be related to each other due to capturing the waveform measurements during the same physical event.

However, in practice, such pair-wise correspondence between the waveform measurements from two power quality sensors is *not* known. Such pair-wise relationship needs to be established first, before we can attempt to solve Problem 1.

A power quality sensor records the waveforms only if it detects an event. In this regard, the first power quality sensor captures the waveform $\mathbf{x}$ only if it detects an event at its own location. Similarly, the second power quality sensor captures the waveform $\mathbf{y}$ only if it detects an event at its own location.

Importantly, not every physical event triggers waveform capture by every power quality sensor. There are many physical events that are captured only by one power quality sensor. Any such event that is captured by the first power quality sensor would not be in correspondence to any event that is captured by the second power quality sensor, and vice versa. Furthermore, in practice, δ can be large. Therefore, one cannot easily find the pairs of waveform time-series from the two sensors that correspond to each other. Finally, even if the two waveform time-series from the two sensors correspond to the same physical event, the way that the event is manifested in the two waveform time-series may *not* be similar.

An example is shown in Fig. 2. Four events are captured by the first power quality sensor, and five events are captured by the second power quality sensor. The two sensors are *not* time synchronized, i.e., $\delta \neq 0$. Let $\mathcal{X}$ denote the set of all time-series of waveform measurements from the first sensor. Similarly, let $\mathcal{Y}$ denote the set of all time-series of waveform measurements from the second sensor. We have:

$$\begin{aligned} \mathcal{X} &= \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}, \\ \mathcal{Y} &= \{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4, \mathbf{y}_5\}. \end{aligned} \tag{2}$$

The second event at the first sensor and the first event at the second sensor correspond to the same physical cause; which has been major enough to affect the waveforms at the locations of *both* sensors. The fourth event at the first sensor and the fourth event at the second sensor also correspond to the same physical cause. The rest of the events are captured only by one of the two sensors. Accordingly, we shall solve Problem 1 based on pair $\mathbf{x} = \mathbf{x}_2, \mathbf{y} = \mathbf{y}_1$ and pair $\mathbf{x} = \mathbf{x}_4, \mathbf{y} = \mathbf{y}_4$.

The above alignments are *not* known in advance. Accordingly, Problem 2 is concerned with identifying which pairs of the waveform time-series from the two power quality sensors should be used for the purpose of solving Problem 1.

III. Method to Solve Problem 1

Consider Problem 1 in Section II-A and the example in Fig. 1. Recall that the two waveform measurements in this figure are *not* time synchronized. Synchronization factor δ, i.e., the offset between the local clocks of the two sensors, is *unknown*.

A. Measuring Two Quantities in Waveforms

While we cannot directly measure δ from Fig. 1, we can measure the following two quantities, as illustrated in Fig. 3:

1) The difference in time between the *zero-crossing points* at the two waveforms as measured by the local clocks at the two power quality sensors. This is denoted by Δ_{zero}.

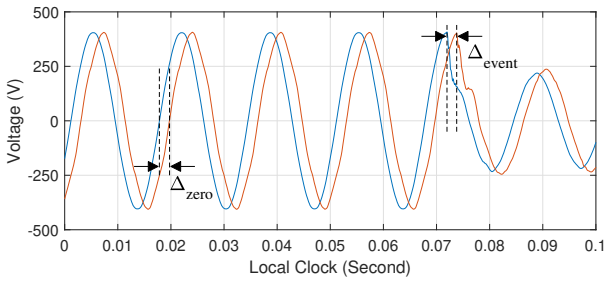

Fig. 3. Calculating the values of Δ_{zero} and Δ_{event} for the two time-series of voltage waveform measurements in Fig. 1.

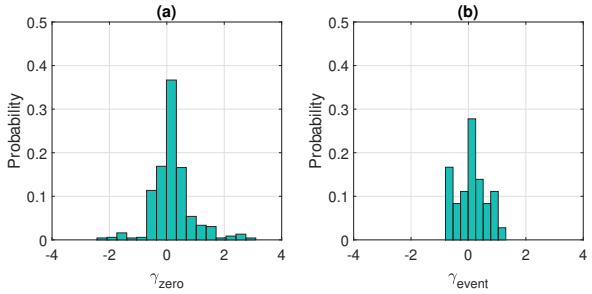

Fig. 4. Experimental results to obtain the distribution of (a) γ_{zero} (n = 246) and (b) γ_{event} (n = 36). The x-axis is in milliseconds in both figures.

2) The difference in time between the *event-starting points* at the two waveforms as measured by the local clocks at the two power quality sensors. This is denoted by Δ_{event}.

These quantities are both related to the unknown synchronization factor δ. With regards to the first quantity, we have:

$$\begin{aligned}\Delta_{\text{zero}} = \delta &+ \text{Phase Difference} \\ &+ \text{Error in Calculating Zero Crossing Time.}\end{aligned} \tag{3}$$

With regards to the second quantity, we have:

$$\begin{aligned}\Delta_{\text{event}} = \delta &+ \text{Propagation Delay Difference} \\ &+ \text{Error in Calculating Event Start Time.}\end{aligned} \tag{4}$$

Next we explain the terms that are mentioned in (3) and (4).

The phase difference in (3) commonly exists between any two locations on a power network. In fact, it is related to the concept of Relative Phase Angle Difference (RPAD) which is often used in the analysis of power systems based on phasor measurements, e.g., see [12, Section 3.4]. The amount of phase difference depends on the characteristics of the circuits and the operating conditions in the power system. Accordingly, it is *not* a fixed amount and it can vary at different moments.

The propagation delay difference in (4) also commonly exists between any two locations on a power network, *relative* to the location of the event. That is, when an event, such as a fault, occurs at a location on the circuit, its impact may *not* be seen precisely at the same time if the electric distance is not the same between the location of the event and the locations of the two power quality sensors. Importantly, the propagation delay difference is not fixed and it depends on the electrical distance of *each* event relative to the two sensor locations.

The error terms in (3) and (4) are self-explanatory. The challenges that may arise in calculating the zero crossing time are discussed in [13, Appendix]. The challenges in calculating the event start time are also discussed in [12, Section 4.2]. These challenges can affect calculating the values of Δ_{zero} and Δ_{event}. Importantly, these errors are not fixed. They can vary depending on the shape of the waveforms, such as the presence of harmonics or the shape of the event signature.

B. Data-Driven Characteristics of Δ_{zero} *and* Δ_{event}

Next, let us define the following random variables:

$$\gamma_{\text{zero}} = \Delta_{\text{zero}} - \delta \quad \text{and} \quad \gamma_{\text{event}} = \Delta_{\text{event}} - \delta. \tag{5}$$

From (3), γ_{zero} is the random variable that results from phase difference and error in calculating zero-crossing times. From (4), γ_{event} is the random variable that results from propagation delay difference and error in calculating event start time.

Using WMUs in Riverside, California, the distributions of γ_{zero} and γ_{event} are derived based on the real-world data. The results are shown in Figs. 4(a) and (b). Note that the distribution is similar on each phase. The distribution in Fig. 4(a) is based on calculating the difference in zero-crossing times among the waveform time-series that are simultaneously measured by two *time-synchronized* WMUs at two different locations on the same power network. And the distribution in Fig. 4(b) is based on calculating the difference in event start times among the waveform time-series from the same two WMUs. The use of two time-synchronized WMUs was necessary in this analysis to set $\delta = 0$, such that we can measure γ_{zero} and γ_{event} by measuring Δ_{zero} and Δ_{event}, respectively.

Interestingly, both of the two distributions in Fig. 4 have mean values close to zero. In particular, we have:

$$E\{\gamma_{\text{zero}}\} = 0.19 \quad \text{milliseconds} \tag{6}$$

and

$$E\{\gamma_{\text{event}}\} = 0.10 \quad \text{milliseconds}, \tag{7}$$

which are very small compared to the range of γ_{zero} and γ_{event}. Note that, from Fig. 4, γ_{zero} can be as low as -2.43 milliseconds and as high as 3.10 milliseconds. Also, γ_{event} can be as low as -0.81 milliseconds and as high as 1.32 milliseconds.

Based on the above results, we can approximately *cancel out* the impact of the added terms in (3) and (4) to estimate δ by taking the expected value of Δ_{zero} and the expected value of Δ_{event}. Accordingly, we can approximately derive that

$$\delta \approx E\{\Delta_{\text{zero}} + \kappa T\} \tag{8}$$

and

$$\delta \approx E\{\Delta_{\text{event}}\}, \tag{9}$$

where T denotes the period of each AC cycle, such as 16.667 milliseconds in North America. The expected value $E\{\Delta_{\text{zero}}\}$ is calculated based on all instances of measuring Δ_{zero}, i.e., based on all pairs of waveform time-series $\mathbf{x}$ and $\mathbf{y}$ that are available. Similarly, the expected value $E\{\Delta_{\text{event}}\}$ is calculated based on all instances of measuring Δ_{event} in Problem 1.

In (8), notation κ is an *unknown* that can be a distinct integer number for *every* reading of Δ_{zero}. This is due to the *ambiguity* that inherently exists in calculating the difference between zero-crossing points in $\mathbf{x}$ and $\mathbf{y}$ due to the *repeating* nature of the waveform cycles. Importantly, the value of Δ_{zero}

would be the same if δ increases or decreases by multiples of T. Notice that, the term κT is *inside* the expected value. Thus, there can be a different κ for every Δ_{zero} that we measure.

To resolve the above ambiguity, for a given Δ_{zero}, we introduce the following *complex* number to capture the *angular* difference between the zero-crossing points in $\mathbf{x}$ and $\mathbf{y}$:

$$z_{\text{zero}} = \exp\left(\sqrt{-1}\, 2\pi \Delta_{\text{zero}}/T\right). \tag{10}$$

We can now replace (8) with the following expression, where the expected value is in the domain of *complex* numbers:

$$\delta \approx (T/2\pi)\, \angle E\{z_{\text{zero}}\} + \kappa T. \tag{11}$$

Notice that the term κT is now *outside* the expected value.

C. Joint Optimization to Estimate δ and κ

We propose to estimate the synchronization factor δ, as well as to estimate κ to resolve the above ambiguity, by solving the following optimization problem over both unknown variables:

$$\underset{\delta,\kappa}{\text{minimize}} \left\| \begin{bmatrix} \delta \\ \delta \end{bmatrix} - \begin{bmatrix} E\{\Delta_{\text{event}}\} \\ \frac{T}{2\pi}\angle E\{\exp\left(\sqrt{-1}\, \Delta_{\text{zero}} \frac{2\pi}{T}\right)\} + \kappa T \end{bmatrix} \right\|_2, \tag{12}$$

where κ is an integer variable. Here, we use Δ_{event} as a reference to identify k; because the repeating nature of the waveform cycles does *not* affect calculating the difference between the event starting points. It only affects calculating the difference between the zero crossing points. Furthermore, we select δ in (12) according to a trade-off between the two approximations in (9) and (11) using the Euclidean norm.

IV. Method to Solve Problem 2

Next, consider Problem 2 in Section II-B. In this section, we propose a method to solve this problem. Our method involves pair-wise calculation of dissimilarities among the waveform time-series in sets $\mathcal{X}$ and $\mathcal{Y}$; followed by conducting an optimization to identify the waveform time-series from the two power quality sensors that correspond to the same event.

A. Pair-wise Dissimilarity Assessment

For any given waveform $\mathbf{x}_i \in \mathcal{X}$ and any given waveform $\mathbf{y}_j \in \mathcal{Y}$, let us define a *dissimilarity index* as follows:

$$\sigma_{ij} = f(\mathbf{x}_i, \mathbf{y}_j), \tag{13}$$

where $f(\cdot,\cdot)$ denotes a *dissimilarity function* which calculates the similarity between two waveform time-series. For example, we can define the similarity function as follows [3]:

$$f(\mathbf{x},\mathbf{y}) = \min_{m_1,m_2} \left\|\Delta\mathbf{x}[m_1 : m_1{+}w] - \Delta\mathbf{y}[m_2 : m_2{+}w]\right\|_2, \tag{14}$$

where m_1 and m_2 are integer numbers between 1 and $n{-}T{-}w$ and w is the length of a sliding window. Furthermore, we denote the *differential waveforms* corresponding to $\mathbf{x}$ and $\mathbf{y}$ as follows, respectively [12, Section 4.2.5]:

$$\begin{aligned} \Delta\mathbf{x} &= \mathbf{x}[CT+1:n] - \mathbf{x}[1:n-CT], \\ \Delta\mathbf{y} &= \mathbf{y}[CT+1:n] - \mathbf{y}[1:n-CT]. \end{aligned} \tag{15}$$

The above differential waveforms are the difference between the original waveform measurements and the *delayed* versions of the *same* waveform measurements, where the delay is C cycles. In this paper, we set $C = 3$ to assure separating the duplicate event signatures in differential waveforms; see [12, p. 168]. Defining dissimilarity based on differential waveforms has multiple advantages, such as removing the impact of background load to focus only on the waveform signature that is caused by the event. An example for the extraction of $\Delta\mathbf{x}$ and $\Delta\mathbf{y}$ from the raw waveform measurements $\mathbf{x}$ and $\mathbf{y}$ is shown in Fig. 5. Note that, whenever needed, one can use the *per-unit* representation of the waveforms before extracting $\Delta\mathbf{x}$ and $\Delta\mathbf{y}$ to have comparable event signatures for the purpose of checking dissimilarities in the minimization in (14).

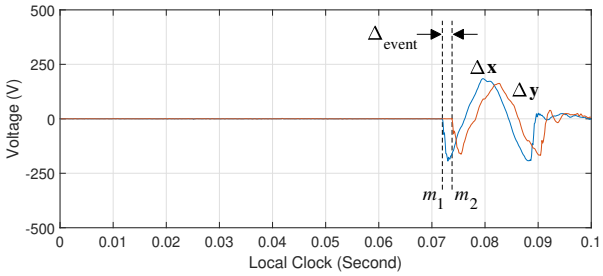

Fig. 5. The differential waveforms values of $\Delta\mathbf{x}$ and $\Delta\mathbf{y}$ corresponding to the two time-series of voltage waveform measurements in Fig. 1.

The minimization in (14) calculates the difference between any window of length w in $\Delta\mathbf{x}$ and any window of length w in $\Delta\mathbf{y}$ to identify the smallest dissimilarity (i.e., the *highest similarity*) between the two differential waveforms. The window size w can be set to $T/4$, $T/2$, or T; see [3].

B. Binary Optimization for Pair-wise Waveform Matching

Let b_{ij} denote a binary decision variable to indicate whether waveform $\mathbf{x}_i$ and waveform $\mathbf{y}_j$ correspond to the same physical event. This variable can be defined as follows:

$$b_{ij} = \begin{cases} 1, & \text{if } \mathbf{x}_i \text{ and } \mathbf{y}_j \text{ correspond to the same event,} \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

In general, we must have:

$$\begin{aligned} \sum_j b_{ij} \le 1, \quad \forall\, \mathbf{x}_i \in \mathcal{X}, \\ \sum_i b_{ij} \le 1, \quad \forall\, \mathbf{y}_j \in \mathcal{Y}. \end{aligned} \tag{17}$$

From (17), each waveform capture at the first power quality sensor can correspond to *at most* one waveform capture at the second power quality sensor. Similarly, each waveform capture at the second power quality sensor can correspond to *at most* one waveform capture at the first power quality sensor. Of course, it is possible that a waveform capture on one power quality sensor is *not* corresponding to any waveform capture at the other power quality sensor. This would be the case where the waveform capture was due to a local phenomena that did not trigger waveform capture at the other power quality sensor.

Next, we note that, although the two power quality sensors are not time-synchronized, the local clock at each power quality sensor can help us determine the *order* of the waveform

captures at each power quality sensor. For example, from (2), we know that $\mathbf{x}_1$ was captured before $\mathbf{x}_2$; $\mathbf{x}_2$ was captured before $\mathbf{x}_3$; and $\mathbf{x}_3$ was captured before $\mathbf{x}_4$. We must maintain the orders of the events when we identify which waveform from set $\mathcal{X}$ corresponds to which waveform (if any) from set $\mathcal{Y}$. For example, it does not make sense to pair $\mathbf{x}_1$ and $\mathbf{y}_4$ while we also pair $\mathbf{x}_3$ and $\mathbf{y}_2$; because this would violate the order of the waveforms in sets $\mathcal{X}$ and $\mathcal{Y}$. To address this issue, for each i and each j, we propose to use the following constraints:

$$b_{lk} \leq 1 - b_{ij}, \quad \forall l, k : l < i,\ k > j. \tag{18}$$

If $b_{ij} = 1$, i.e., if $\mathbf{x}_i$ and $\mathbf{y}_j$ are deemed to correspond to the same event, then (18) becomes $b_{lk} \leq 0$; which means we would have $b_{lk} = 0$ for any pair of $\mathbf{x}_l$ and $\mathbf{y}_k$ for which $l < i$ and $k > j$. This would assure maintaining the order of the events that are captured by the two power quality sensors. If $b_{ij} = 0$, i.e., if $\mathbf{x}_i$ and $\mathbf{y}_j$ are *not* deemed to correspond to the same event, then (18) becomes $b_{lk} \leq 1$; which would *not* impose any constraint on binary variable b_{lk} for any l or k.

If we seek to pair exactly M waveform captures across the two power quality sensors, then we must have:

$$\sum_i \sum_j b_{ij} = M, \tag{19}$$

where M is an integer number that is upper bounded by the number of members in set $\mathcal{X}$ and the number of members in set $\mathcal{Y}$. That is, M is upper bounded by $\min\{|\mathcal{X}|, |\mathcal{Y}|\}$. For example, if we set $M = 3$, then we require identifying exactly three pairs of the time-series of waveform measurements across the two power quality sensors. The remaining $|\mathcal{X}| - M$ waveform captures from the first power quality sensor and the remaining $|\mathcal{Y}| - M$ waveform captures from the second power quality sensor are accordingly assumed to correspond to some local events that are *not* captured by both sensors.

In order to address Problem 2, we need to obtain the value of b_{ij} for any $\mathbf{x}_i \in \mathcal{X}$ and any $\mathbf{y}_j \in \mathcal{Y}$. For notional simplicity, we stack up all such decision variables into vector $\mathbf{b}$. In this regard, we propose the following optimization problem:

$$\begin{aligned} \underset{\mathbf{b}}{\text{minimize}} \quad & \sum_i \sum_j b_{ij}\sigma_{ij} \\ \text{subject to} \quad & \text{Eqs. (17) and (18) and (19).} \end{aligned} \tag{20}$$

The above optimization problem is an integer linear program (ILP). It can be solved using solvers such as CPLEX [14].

For a given M, let us denote the optimal objective value of the optimization problem in (20) by $g(M)$. We can show that function $g(M)$ is *non-decreasing* over M. That is, we have:

$$g(M) \leq g(M+1), \quad \forall M = 1, \dots \min\{|\mathcal{X}|, |\mathcal{Y}|\} - 1. \tag{21}$$

The proof of the above inequality is evident, given the fact that increasing M results in adding more non-negative dissimilarity terms to the optimal objective value in (20); which directly results in a non-decreasing behavior in function $f(M)$.

Based on (21), we propose to repeatedly solve (20) by starting from $M = 1$ and incrementing M until we either reach M's upper bound at $\min\{|\mathcal{X}|, |\mathcal{Y}|\}$ or we exceed a threshold ϵ on the optimal objective value of the optimization problem in (20), i.e., until the following stopping criteria is reached:

$$g(M) \geq \epsilon. \tag{22}$$

The above condition stops increasing M when the *total dissimilarities* in the selected pairs of the waveform measurements exceeds the threshold, suggesting that the captured waveform measurements from the two sensors that are being paired are *not* captured during the same event in the power system. Based on the experimental results, we propose to set $\epsilon = 1.5$.

V. Experimental Case Study

An experiment is done based on the measurements from two power quality sensors in California to evaluate the performance of the proposed methods. Twenty waveform captures are considered from the two sensors on one day, where:

$$|\mathcal{X}| = |\mathcal{Y}| = 10. \tag{23}$$

Only five events at each power quality sensor are also captured by the other power quality sensor. The rest of the events are local events. Specifically, events $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{x}_8$, $\mathbf{x}_{10}$ at the first power quality sensor correspond to events $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$, $\mathbf{y}_4$, $\mathbf{y}_5$ at the second power quality sensor. The rest of the events are local events: $\mathbf{x}_4$, $\mathbf{x}_5$, $\mathbf{x}_6$, $\mathbf{x}_7$, $\mathbf{x}_9$, $\mathbf{y}_6$, $\mathbf{y}_7$, $\mathbf{y}_8$, $\mathbf{y}_9$, $\mathbf{y}_{10}$.

The above information is available to us; because both sensors are equipped with a GPS synchronization module. This provides us with the *ground truth* about the available waveform measurements. However, for the purpose of our case study, we assume that not only the above knowledge is *not* available to us, but also there is an *unknown* offset between the local clocks at the two sensors, i.e., $\delta \neq 0$ and δ is unknown.

First we solve the optimization in (14) for each pair of the available waveform measurements from the two sensors. The results are shown in Table I. Notice that if we solely consider the values of σ_{ij} in each row or each column then we *cannot* correctly identify which events at the two sensors correspond to each other. For example, from the first row in Table I, the waveform measurements in $\mathbf{x}_1$ are most similar to the waveform measurements in $\mathbf{y}_4$; however, $\mathbf{x}_1$ and $\mathbf{y}_4$ do *not* correspond to the same physical event. Similarly, from the first column in Table I, the waveform measurements in $\mathbf{y}_1$ are most similar to the waveform measurements in $\mathbf{x}_8$; however, $\mathbf{y}_1$ and $\mathbf{x}_8$ do *not* correspond to the same physical event.

The remedy is to rather solve the systematic optimization problem in (20). The optimal results are obtained as follows:

$$\begin{aligned} M = 1: &\quad b_{84} = 1 \\ M = 2: &\quad b_{84} = b_{22} = 1 \\ M = 3: &\quad b_{84} = b_{22} = b_{11} = 1 \\ M = 4: &\quad b_{84} = b_{22} = b_{11} = b_{33} = 1 \\ M = 5: &\quad b_{84} = b_{22} = b_{11} = b_{33} = b_{105} = 1, \end{aligned} \tag{24}$$

where b_{105} is the binary decision variable between $\mathbf{x}_{10}$ and $\mathbf{y}_5$. The values of $g(M)$ for $M = 1, \dots, 10$ are as follows: 0.1769, 0.3848, 0.6089 0.8678, 1.3102, 1.9196, 2.5293, 3.1574, 3.8387, 4.6571. These numbers confirm the non-decreasing property of $g(M)$ in (21). Importantly, *all* the

TABLE I
RESULTS ON ESTIMATING σ_{ij} FOR ANY PAIR OF EVENTS BETWEEN TWO POWER QUALITY SENSORS THAT ARE NOT TIME SYNCHRONIZED

σ_{ij}		Event Number at Power Quality Sensor 2									
		1	2	3	4	5	6	7	8	9	10
Event Number at Power Quality Sensor 1	1	**0.2241**	0.2133	0.4132	0.1900	0.4863	0.6680	0.4153	0.5663	0.4397	0.3772
	2	0.2178	**0.2079**	0.3588	0.1830	0.4523	0.6577	0.3658	0.5796	0.4258	0.3898
	3	0.4518	0.3947	**0.2589**	0.3885	0.3696	0.7340	0.5508	0.5729	0.5649	0.5285
	4	0.5325	0.5073	0.5724	0.5559	0.6114	0.6527	0.5706	0.6454	0.5515	0.5419
	5	0.4734	0.4206	0.3659	0.4105	0.4827	0.6997	0.4754	0.6667	0.4978	0.4809
	6	0.5603	0.5151	0.5249	0.5228	0.6193	0.7174	0.4886	0.6056	0.5337	0.5201
	7	0.4700	0.4349	0.4890	0.4422	0.5086	0.6630	0.4626	0.5787	0.4896	0.4382
	8	0.2107	0.1962	0.4198	**0.1769**	0.5087	0.6955	0.6955	0.6349	0.4913	0.4497
	9	0.3710	0.3430	0.4129	0.3288	0.4739	0.7283	0.5517	0.6362	0.5348	0.5506
	10	0.5060	0.4896	0.2821	0.4856	**0.4424**	0.7319	0.6013	0.7605	0.5952	0.5779

TABLE II
TEST RESULTS TO ESTIMATE δ AND κ IN PROBLEM 1

Ground Truth	Data-Driven Optimization	
δ (ms)	δ (ms)	κ
3.5	3.1	0
8.7	8.3	1
-13.1	-13.5	-1
46.9	46.5	3

results in (24) are *correct*. Furthermore, when $M = 6$, we have $g(M) > \epsilon = 1.5$, which stops the process of incrementing M, leaving us to choose $M = 5$; which is indeed the *correct* number of pairs of waveform measurements that correspond to the same physical event. Therefore, our solution approach to solve Problem 2 (see Section IV-B) is successful.

Next, we use the results in the last row in (24) to calculate Δ_{zero} and Δ_{event} for the following pairs of waveforms between the two power quality sensors: for $\mathbf{x}_8$ and $\mathbf{y}_4$, for $\mathbf{x}_2$ and $\mathbf{y}_2$, for $\mathbf{x}_1$ and $\mathbf{y}_1$, for $\mathbf{x}_3$ and $\mathbf{y}_3$, and for $\mathbf{x}_{10}$ and $\mathbf{y}_5$. Of course, the values of Δ_{zero} and Δ_{event} will depend on the value of the synchronization factor δ, i.e., the offset between the clocks of the two power quality sensors. For the purpose of our experiment, we assume the following values:

$$\delta = 3.5 \text{ ms},\ \delta = 8.7 \text{ ms},\ \delta = -13.1 \text{ ms},\ \delta = 46.9 \text{ ms}. \quad (25)$$

In each case, we impose an offset equal to one of the above values to the waveform measurements. Accordingly, we obtain the values of Δ_{zero} and Δ_{event} in each case. Importantly, we obtain Δ_{zero} and Δ_{event} for each phase; thus extracting a total of $15 = 5 \times 3$ values for Δ_{zero} and $15 = 5 \times 3$ values for Δ_{event}. The results are then used to optimally identify the unknown values of δ and κ by solving the optimization problem in (12).

The results of the optimization are shown in Table II. Both δ and κ are estimated very reasonably in all cases. Therefore, we can conclude that our solution approach to solve Problem 1 (see Section II-A) is successful. It is worth pointing out that the results in Table II are based on the measurements on one day only. If we utilize more measurements, we anticipate to further improve the accuracy of the results.

VI. CONCLUSIONS AND FUTURE WORK

The practical results in this paper can greatly help utilities to directly turn the measurements from their *existing* power quality sensors to synchro-waveforms. Accordingly, they can take advantage of the growing use cases of synchro-waveforms *without* the need to replace or retrofit their existing sensors.

Experimental case studies confirmed the high performance of the proposed methods in achieving data-driven time-synchronization among conventional power quality sensors solely based on investigating their waveform measurements using statistical analysis and novel optimization techniques.

The study in this paper can be extended in various directions. First, other choices of the dissimilarity (or similarity) functions between the two time-series of waveform measurements can be used, such as based on correlation or convolution. Second, the methods can be extended to incorporate unsynchronized waveform measurements from more than two power quality sensors to enhance performance due to the redundancy in measurements. Third, additional case studies can be conducted based on weeks or months of test data.

REFERENCES

[1] H. Mohsenian-Rad and W. Xu, "Synchro-waveforms: A window to the future of power systems data analytics," *IEEE Power and Energy Magazine*, vol. 21, no. 5, pp. 68–77, Sep. 2023.

[2] A. F. Bastos, S. Santoso, W. Freitas, and W. Xu, "Synchrowaveform measurement units and applications," in *Proc. of the IEEE PES General Meeting*, Atlanta, GA, Jul. 2019.

[3] H. Mohsenian-Rad, A. Shahsavari, and M. Majidi, "Analysis of power quality events for wildfire monitoring: Lessons learned from a California wildfire," in *Proc. of the IEEE PES ISGT*, San Juan, Puerto Rico, 2023.

[4] https://ieee-synchrowaveform.engr.ucr.edu/.

[5] M. Izadi and H. Mohsenian-Rad, "Synchronous waveform measurements to locate transient events and incipient faults in power distribution networks," *IEEE Trans. on Smart Grid*, vol. 12, pp. 4295–4307, 2021.

[6] I. Khan, H. Sun, K. Kim, J. Guo, and D. Nikovski, "Combined detection and localization model for high impedance fault under noisy condition," in *Proc. of the IEEE PES General Meeting*, Orlando, FL, Jul. 2023.

[7] M. Mansour Lakouraj, H. Hosseinpour, H. Livani, and M. Benidris, "Waveform measurement unit-based fault location in distribution feeders via short-time matrix pencil method and graph neural network," *IEEE Trans. on Industry Applications*, vol. 59, no. 2, pp. 2661–2670, 2023.

[8] T. Zaman, Z. Feng, M. Syed, B. Pilscheur, D. Flynn, and G. Burt, "Multimode synchronous resonance detection in converters dominated power system using synchro-waveforms," in *Proc. of the International Conference on Electricity Distribution*, Rome, Italy, Jun 2023.

[9] F. Ahmadi-Gorjayi and H. Mohsenian-Rad, "Data-driven models for sub-cycle dynamic response of inverter-based resources using WMU measurements," *IEEE Trans. on Smart Grid*, pp. 4125–4128, Sep 2023.

[10] M. Alberto *et al.*, "Newly implemented real-time PQ monitoring for transmission 4.0 substations," *Electric Power Systems Research*, vol. 204, p. 107709, Mar 2022.

[11] Z.-J. Ye, M. Izadi, M. Farajollahi, and H. Mohsenian-Rad, "A remedy to losing time synchronization at D-PMUs, H-PMUs, and WMUs in event location identification in power distribution systems," *IEEE Trans. on Smart Grid*, vol. 15, no. 1, pp. 651–654, 2024.

[12] H. Mohsenian-Rad, *Smart Grid Sensors: Principles and Applications*. Cambridge University Press, UK, Apr. 2022.

[13] IEEE PES Working Group on Power Quality Data Analytics, "Electric signatures of power equipment failures," Technical Report, 2018.

[14] https://www.ibm.com/products/ilog-cplex-optimization-studio.