# Statistical Model of Measurement Noise in Real-World PMU-based Acquisitions

G. Frigo*, A. Derviškadić*, A. Bach†, M. Paolone*

*École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland
†École Normale Supérieure de Cachan, Université Paris-Saclay, Paris, France
mail: guglielmo.frigo@epfl.ch

*Abstract*—In this paper, we present a processing technique to determine the statistical distribution of additive measurement noise in real-world acquisitions, with specific reference to Phasor Measurement Unit (PMU) applications in Active Distribution Networks (ADNs). The proposed approach identifies the power signal fundamental component, as well as harmonic and inter-harmonic interferences, and models the measurement noise as a Gaussian random variable. First, we describe the algorithm main stages and the criteria for the most suitable parameter setting. Then, we carry out a numerical validation inspired by IEEE Std. C37.118.1 test conditions. Finally, we validate the proposed approach on real-world measurements acquired by a PMU in the distribution network of EPFL campus.

*Index Terms*—Noise measurement, Statistical analysis, Taylor-Fourier Transform, Phasor Measurement Unit (PMU)

## I. Introduction

In recent years, Phasor Measurement Units (PMUs) are increasingly employed as a distributed and flexible measurement infrastructure for electrical power network monitoring and control [1]–[3]. In this context, the IEEE Std C37.118.1 [4] and its recent amendment [5], briefly IEEE Std, define the PMU requirements in terms of estimation accuracy and reporting latency in both static and dynamic test conditions. To this end, the IEEE Std introduces a set of challenging test conditions, intended to evaluate the PMU estimates' reliability in the presence of different uncertainty sources, like interfering components or time-varying trends of the fundamental amplitude, frequency or initial phase.

Unfortunately, as regards the resilience against both narrow- and wide-band disturbances, the IEEE Std tests do not represent a plausible approximation of the distortion and noise levels typically experienced in modern power networks [6]. In particular, the *Harmonic distortion* test considers a steady-state condition where the fundamental component is characterized by constant amplitude, frequency and initial phase, and a single harmonic component produces a Total Harmonic Distortion (THD) equal to 1 or 10%. Similar considerations hold for the *Out-of-band interference* test, where the distortion level due to a single inter-harmonic component is limited to 10% only. In terms of measurement noise, the IEEE Std does not even provide any specific requirement, thus limiting the noise sources to the quantization and non-linearity effects inherent in any device that processes digitized data [7].

Before deploying PMUs in real-world applications, it would be recommended to test them in plausible operating conditions. As a consequence, the development of statistical models for measurement noise along with harmonic and inter-harmonic components still represents an open issue. In this context, recent literature has discussed the statistical distribution of noise [9] and time-varying distortion levels [10] in real medium voltage networks. However, the proposed results tend to be application-dependent and lack of a parametric model to be used in the assessment of PMU metrological performance.

In this paper, we propose a processing method to infer a thorough statistical model of the measurement noise from real-world current and voltage waveforms, as acquired by PMUs in distribution networks. From a mathematical point of view, we model this uncertainty source as the summation of two additive and uncorrelated components. In particular, we approximate the high-frequency spurious oscillations by means of a white Gaussian random variable, whereas we represent the low-frequency spectrum by means of a stochastic random walk. In this way, we are able to model not only a generic wide-band measurement noise, but also the non-linear effects of low-power instrument transformers.

The proposed approach consists of four main stages. First, we employ the Iterative Interpolated DFT (i-IpDFT) algorithm [11] to accurately determine the fundamental frequency. Based on this information, we design a Taylor-Fourier Multifrequency (TFM) model [12] that accounts for the first 50 harmonic components and enables us to recover the signal dynamic evolution over time [13]. By construction, the difference between acquired waveform and TFM-based estimates should account only for measurement noise and other (inter-harmonic) components, that were not previously included in the TFM-model. This working hypothesis can be verified *a posteriori* by inspecting the statistical properties of the obtained noise model.

In order to mitigate the effect of inter-harmonic components, we apply a peak-selection routine to the difference signal. If a significant spectral tone is detected, the corresponding frequency is included in the TFM-model and the estimate is repeated. Otherwise, we assume that the difference signal consists only of noise, and thus we process it in order to distinguish low- and high-frequency noise components. On one side, we apply a Savitzky-Golay filter to extract a smooth

signal baseline [14]. On the other side, we identify a Gaussian random variable that better fits the residual distribution.

The possible applications of this noise modeling technique are numerous. On one side, most state and line parameter estimation techniques require a prior knowledge of the measurement noise variance. This information is typically assumed as known by the modeler, but is hardly retrievable in practice. In this context, the presented routine can provide an accurate and frequently updated noise characterization and thus avoid discrepancy between model inputs and real measurements [15], [16]. To the best of the Authors' knowledge, some analysis have been carried out on the noise affecting PMU estimates [17], [18], but a thorough and rigorous modeling routine for the actual measurement noise on the acquired waveforms (and thus applicable to other smart meters) is still missing.

On the other side, the present paper validates the proposed method by using point-on-wave data coming from PMUs installed in a real-world distribution feeder, and quantifies the noise level typical of medium voltage distribution networks. Such values can be used by PMU developers when testing the appropriateness of their synchrophasor estimation technique for distribution installations.

The paper is organized as follows. In Section II, we describe the processing method for the extraction of measurement noise from real-world acquired waveform. Section III provides a thorough characterization of the routine accuracy in numerically simulated scenarios, inspired by the IEEE Std test conditions. In Section IV, we evaluate the effect of improper noise modeling through the simulation of a non-normal distributed measurement noise, and in Section V, we apply the proposed technique to an experimental dataset of real-world waveforms directly acquired in the EPFL campus [19]. Finally, Section VI provides some closing remarks.

## II. PROPOSED STATISTICAL ANALYSIS TECHNIQUE

In this Section, we present the processing method for the identification of measurement noise statistical properties. First, we define the signal model, then we describe the algorithm main stages and discuss the criteria for the most suitable setting of the algorithm parameters. It is worth noticing that the proposed algorithm has been developed based on the given signal model, whose validity has been *a posteriori* verified by means of real-world experimental acquisition. Nevertheless, in the following Section, we investigate also the uncertainty contribution due to possible discrepancy between the signal model and the acquired waveform.

*a) Signal model:* Let us consider a noiseless sinusoidal signal, affected by narrow-band interfering components:

$$x(t) = A \cdot (1 + \gamma_A(t)) \cdot \cos(2\pi f t + \varphi + \gamma_\varphi(t)) + \beta \quad (1)$$

where $A$, $f$, and $\varphi$ represent the amplitude, frequency and initial phase of the fundamental component, respectively. The time-varying terms $\gamma_A$ and $\gamma_\varphi$ account for any dynamic evolution of amplitude and phase (e.g. modulations, linear

ramps, quadratic trends, etc.). The additive term $\beta$ models the contribution of both harmonic and inter-harmonic components.

The acquisition process introduces further disturbances that can be modeled as additive uncorrelated[1] stochastic variables:

$$y[nT_s] = x[nT_s] + \eta[nT_s] + \varepsilon[nT_s], \qquad n = 1, \ldots N_s \quad (2)$$

where the sample index $n$ ranges from 1 to $N_s$, i.e. the length of the considered observation interval. For this analysis, we assume the sampling time $T_s$ to be uniform and the corresponding sampling rate $F_s = 1/T_s$ to exceed the signal Nyquist rate by almost one order of magnitude. It is thus reasonable to say that the acquired signal $y$ is not affected by neither jitter phase noise nor aliasing. The two additive terms $\eta$ and $\varepsilon$ account for low- and high-frequency noise components. In more detail, $\eta$ represents the time-evolution of the signal baseline level (i.e. the DC component in the context of power network) and is modeled as an one-dimensional random walk, whereas $\varepsilon$ is associated to the wide-band measurement noise and is assumed to be known in terms of distribution type (i.e., the identification problem unknowns are the distribution parameters). The combination of these two terms enables us to suitably reproduce many realistic operating conditions, where the acquired signal is affected by several uncertainty sources, e.g. quantization noise or transducer non-linearity.

*b) Identification method:* First of all, given the acquired signal $y$, we compute and then subtract its mean value $\mathbb{E}[y]$. This operation allows for significantly reducing spectral leakage effects due to DC components.

Once removed any constant offset, we identify the fundamental frequency $f$ by means of an Iterative Interpolated DFT (i-IpDFT) approach. In particular, the i-IpDFT performs a three-point interpolation of the DFT bins centered around the nominal fundamental frequency (i.e. 50 Hz) and implements an enhanced compensation routine for the mitigation of both long- and short-range spectral leakage effects [11]. In particular, the i-IpDFT algorithm adopts 2 iterations to estimate and compensate the interference coming from the negative frequency component (long-range spectral leakage) and 28 iterations to mitigate the effect of neighboring interfering terms, like harmonic or inter-harmonic components (short-range spectral leakage). It is thus reasonable to expect that the obtained frequency estimate is accurate and robust even in the presence of high distortion levels.

Based on the fundamental frequency information, we design a Taylor-Fourier Multifrequency (TFM) model $B$ that accounts for the spectral contribution of the signal harmonic support $\mathcal{S}_h$ [12] . In this regard, the TFM performance depends on two main model parameters. On one side, the maximum harmonic order $N_h$ defines the actual signal bandwidth, as we limit our spectral analysis up to $N_h \cdot f$ Hz. On the other side, the maximum order $N_o$ accounts for the number of derivative terms to be included in the Taylor-series expansion. Recent

---

[1]Also the assumption of non-correlation between signal and noise components has been verified *a posteriori* on the experimentally acquired waveforms.

results have shown how a $2^{nd}$-order model allows for an accurate estimation of the entire harmonic support in both static and dynamic conditions [13].

It should be noticed that the definition of the signal spectral support $S_h$ does not require the interpolation process to be repeated for each harmonic component. Differently from [20], we consider a scenario where the narrow-band components are further affected by wide-band disturbances due to measurement noise and acquisition system non-linearities. In this case, the estimation accuracy of the harmonic terms by means of an i-IpDFT approach might degrade significantly, particularly if the harmonic amplitudes are comparable with the noise level.

As reported in Algorithm 1, given the TFM model $B$, it is possible to recover the time-domain trend of the selected frequency component by means of a simple Least-Squares (LS) approximation [21][2]. At this point, the estimation residual $r$ should account only for additive uncorrelated components and possible inter-harmonic components. In order to discriminate between these two contributions, we perform a peak search on the residual DFT. For each detected peak that exceeds the significance threshold $\theta_{ih}$ (whose setting is discussed in Sec. III.c), the peak frequency is included in the signal support $S_h$, the TFM model is correspondingly updated, and the LS-based signal recovery is repeated.

Once removed any narrow-band component, we further process the residual $r$, i.e. the measurement noise, to provide a more detailed description of its statistical properties. First, we apply the Savitzky-Golay filter $\mathcal{H}_{lp}$ to extract a smooth approximation $\eta$ of the low-frequency components. For the sake of completeness, this time-varying baseline model includes also the previously computed DC offset $\mathbb{E}[y]$. As regards the high-frequency components $\varepsilon$, instead, we assume

---

[2]By adopting a LS-based estimation approach, we assume that the residual, i.e. the estimated noise, is additive, uncorrelated and normally distributed.

---

**Algorithm 1** Measurement Noise Identification

**Input:** $y$, $N_h$, $N_o$, $\theta_{ih}$, $\mathcal{H}_{lp}$
**Output:** $r$, $\varepsilon$, $\eta$, $\mu$, $\sigma^2$
1. signal mean removal: $y = y - \mathbb{E}[y]$
2. fundamental frequency: $f = \text{i-IpDFT}(y)$
3. harmonic support: $S_h = \{f_h = f \cdot h,\ h = 1, \dots N_h\}$
4. Taylor-Fourier model: $B = \text{TFM}(S_h, N_o)$
5. dynamic signal recovery: $\hat{y} = (B^\dagger B)^{-1} B^\dagger y$
6. estimation residual spectrum: $\rho = \text{fft}(r) = \text{fft}(y - \hat{y})$
7. interharmonic peak search: $S_{ih} = \text{findpeaks}(\rho, \theta_{ih})$
   **if** $S_{ih} \neq \emptyset$
8.     support update: $S_h = S_h \cup S_{ih}$
9.     repeat routine steps from 4. to 7.
   **end**
10. residual de-trending: $\eta = \text{filter}(r, \mathcal{H}_{lp}) + \text{E}[y]$
11. random contribution: $\varepsilon = r - \eta$
12. statistical description: $\varepsilon \longrightarrow \mathcal{N}(\mu, \sigma^2)$

---

the probability density function to be Gaussian, and thus thoroughly describe it through its mean $\mu$ and variance $\sigma$.

*c) Parameter setting:* The proposed algorithm presents several different parameters that can be suitably tuned according to the expected operating conditions or the desired estimation accuracy requirements.

As regards the acquired signal $y$, the observation interval length $N_s$ directly affects the final reconstruction accuracy. In fact, a larger sample number corresponds to a finer frequency resolution, and thus enables us to reduce scalloping loss and spectral leakage effects. As suggested by IEC Std. 61000-4 [22], the observation interval should include ten cycles of the nominal system frequency (in the present case, 50 Hz). In the following, we set the sampling frequency and the observation interval length equal to 50 kHz and 200 ms, respectively.

It should be also noticed that the i-IpDFT algorithm applies a preliminary Hanning weighing function to further mitigate both long- and short-range leakage effects. In this way, in the signal DFT, each component is characterized by a wider main lobe and rapidly-decreasing side-lobes. The first one limits the scalloping loss, whereas the second one reduces mutual interference between adjacent components [23].

The structure of the TFM model depends on the selection of $N_h$ and $N_o$. In concordance with the *Harmonic distortion* test, the maximum harmonic order is set equal to 50. In other words, we assume that the actual signal bandwidth does not exceed 2500 Hz. As regards the derivative terms, instead, it is common practice to limit the Taylor-series expansion up to the $2^{nd}$ order. In this way, in fact, it is possible to estimate the Rate-of-Change-of-Frequency (ROCOF) as the second time-derivative of the signal phase. In this context, it is worth observing that the inclusion of higher-order derivative terms allows for a more accurate reconstruction of dynamic conditions. On the other hand, it is also more easily subject to noise injections [24].

As regards the significance threshold $\theta_{ih}$, the most suitable setting requires a prior knowledge of the additive noise level. Unfortunately, this information is unavailable and the user has to find an optimal trade-off between support completeness and avoiding the misinterpretation of noisy spikes as actual spectral components. An improper setting of this parameter might lead to misinterpret the noise superposed to the signal spectrum as an actual narrow-band spectral component. Nevertheless, it is reasonable to say that similar occurrences are very unlikely. In fact, the proposed approach adopts a window length of 200 ms that corresponds to a spectral resolution of 5 Hz. Differently from the PMU-based measurement context, such a refined value guarantees limited spectral leakage and interference effects among adjacent components. As a result, the discrimination between noise floor and narrow-band tones is less challenging. In particular, in this paper we set $\theta_{ih}$ equal to 3% of the fundamental component amplitude, i.e. nearly one third of the expected inter-harmonic amplitude as defined by the *Out-of-band interference* test [4].

Finally, the Savitzky-Golay filter $\mathcal{H}_{lp}$ is characterized by two main parameters: the filter length and the polynomial

order [14]. In the following, the optimal trade-off between de-noising and reconstruction accuracy has been heuristically found by a filter length of 20 ms, i.e. one cycle of the nominal fundamental frequency, and a polynomial order equal to 3. In other words, we assume that the low-frequency trend $\eta$ is nearly-stationary within the considered window length and its curvature or inflexion can be adequately approximated by a cubic polynomial function.

## III. NUMERICAL VALIDATION

In this Section, we present the validation of the algorithm performance by means of numerical simulations carried out in Matlab 2018 programming environment.

---

**Algorithm 2** Test Waveform Generation

---

**Signal model:** $x(t) = A \cdot \cos(2\pi f t + \varphi) + \beta(t) + \eta(t) + \varepsilon(t)$
1. fundamental fixed parameters: $A = 1$ pu, $\varphi = 0$ rad
2. fundamental frequency: $f \in [45, 55]$ Hz, $\Delta f = 1$ Hz
**Harmonic addition:** $\beta(t) = \sum_h A_h \cdot \cos(2\pi f_h t + \varphi_h)$
3. harmonic frequencies: $f_h = h \cdot f, \qquad h = 2, \ldots 50$
4. harmonic init. phases: $\varphi_h \in \mathcal{U}(0, \pi), \quad h = 2, \ldots 50$
5. harmonic amplitudes: $A_h = A \cdot \text{THD}/50 + \mathcal{N}(0, 0.005)$
   where THD $= [10, 40]\%$, incremental step 5%
**Inter-harmonic tone:** $\beta(t) = \beta(t) + A_{ih} \cdot \cos(2\pi f_{ih} t + \varphi_{ih})$
6. inter-harmonic amplitude: $A_{ih} = 0.1$ pu
7. inter-harmonic init. phase: $\varphi_{ih} \in \mathcal{U}(0, \pi)$ rad
8. inter-harmonic frequency: $f_{ih} \in [10, 25] \cup [75, 90]$ Hz
**Low-frequency trend:** $\eta \in \{$ offset, walk $\}$
9. noisy offset trend: $\eta(t) = 2 \cdot 10^{-4} + \mathcal{N}(0, 1 \cdot 10^{-8})$
10. random walk trend: $\eta(t) = \eta(t-1) + \mathcal{N}(0, 1 \cdot 10^{-6})$
**Measurement noise:** $\varepsilon \in \mathcal{N}(0, \sigma^2) \rightarrow \text{SNR} = [40, 60]$ dB

---

In this regard, Algorithm 2 summarizes the main steps for generating the test waveforms for a power system with nominal rated frequency of 50 Hz. First, we define the fundamental component, whose frequency ranges between 45 and 55 Hz, whereas amplitude and initial phase are set equal to 1 pu and 0 rad, respectively. Then, we add the harmonic components, up to the $50^{th}$ order. In more detail, the harmonic amplitudes and initial phases are modeled as random variables: the amplitudes belong to a Gaussian distribution and produce an overall THD that ranges between 10% and 40%, whereas the phases are uniformly distributed between 0 and $\pi$ rad.

In a specific test, we also evaluate the robustness of the proposed algorithm in the presence of out-of-band interferences. To this end, we add an inter-harmonic component characterized by a generic initial phase, and an amplitude of 0.1 pu, i.e. corresponding to a Total Inter-Harmonic Distortion (TIHD) of 10%. In concordance with the *Out-of-band interference* test, we vary the inter-harmonic frequency between [10, 25] and [75, 90] Hz, and we repeat the test with different fundamental frequency values, i.e. $f \in [47.5, 50, 52.5]$ Hz.

Once defined the spectral composition of the test waveform, we simulate the additive measurement noise. As regards the low-frequency trends, we consider two alternative models: a
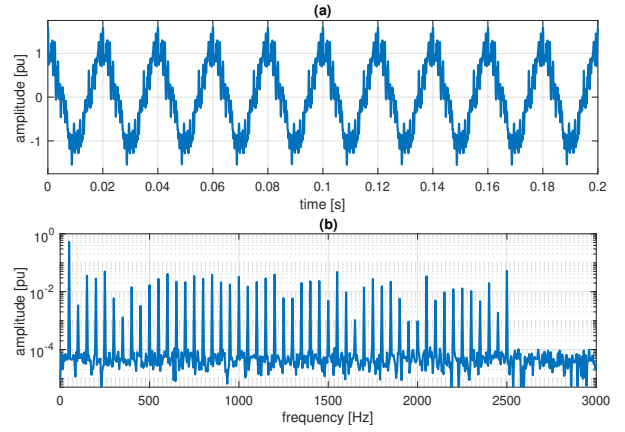


Fig. 1. Test waveform with THD = 25% and SNR = 45 dB in time (a) and frequency (b) domain.
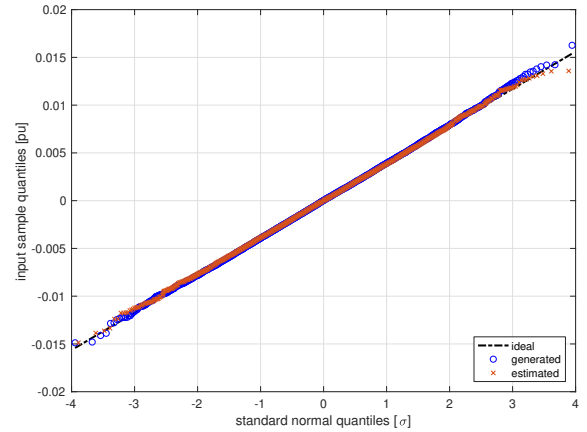


Fig. 2. Quantile-quantile plot of the original (blue) and estimated (red) noise with respect to the ideal Gaussian distribution (black).

stationary offset of $2 \cdot 10^{-4}$ pu that is corrupted by an 80 dB white Gaussian noise, or a random walk whose steps are characterized by a variance of $1 \cdot 10^{-6}$ pu. The wide-band noise contribution, instead, is modeled as an uncorrelated Gaussian random variable, whose variance is scaled to reproduce a SNR ranging from 40 to 60 dB.

Fig. 1 represents a test waveform, where THD and SNR are set to 25% and 45 dB, respectively, and the fundamental frequency is equal to 50 Hz. In the upper graph, it is worth noticing how the time-domain trend of the fundamental component is gravely distorted by both noise and spurious contributions. In the lower graph, we present the signal DFT in the bandwidth of interest [0, 3000] Hz, where the harmonic components are characterized by randomly-varying amplitudes.

In order to assess the algorithm estimation accuracy, Fig. 2 compares the quantile-quantile plots associated to the generated and estimated measurement noise in blue and red line, respectively. In concordance with the assumption of additive Gaussian noise, the two distributions are nearly coincident and most of the samples lay on the ideal Gaussian
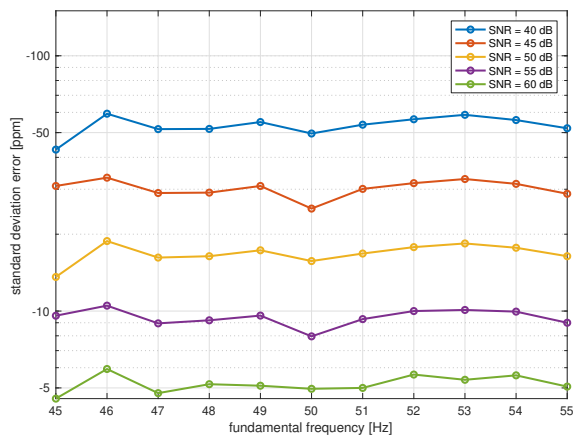
Fig. 3. Standard deviation estimation error as function of the fundamental frequency for different SNR values. For this analysis, THD is fixed at 40%.
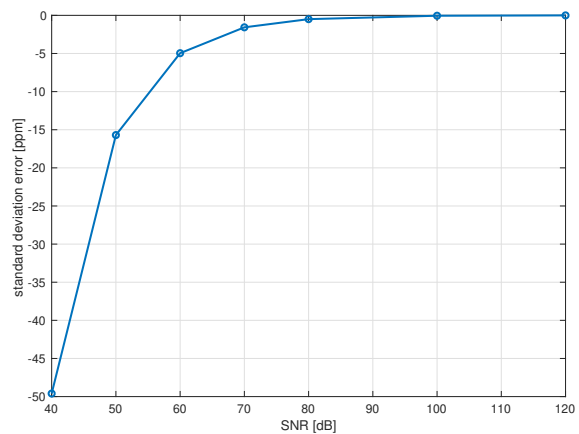


Fig. 4. Standard deviation estimation error as function of the SNR for nominal fundamental frequency (50 Hz). For this analysis, THD is fixed at 40%.

distribution and are included within the range $\pm 3\sigma$. In this context, a more quantitative metric is given by the discrepancy between the estimated and generated standard deviations (that is equal to -27.91 ppm, in the present case).

In the first test, we evaluate the algorithm estimation accuracy as function of the fundamental frequency with different SNR levels ranging from 40 to 60 dB. For this analysis, we set THD equal to 40%, whereas we neglect the effect of low-frequency trend and inter-harmonic interference. As shown in Fig. 3, the proposed algorithm tends to slightly under-estimate the noise standard deviation, with a maximum error of -60 ppm with SNR equal to 40 dB. On the other hand, this discrepancy proves to be inversely proportional to the SNR. As the noise level decreases, the algorithm correctly identifies the fundamental and harmonic contributions and is thus able to accurately recover the original measurement noise.

In order to better understand the accuracy dependence on noise level, we perform the same test for a wider range of SNR values, i.e. from 40 up to 100 dB. In order to minimize the contribution due to spectral leakage, for this analysis we set the fundamental frequency equal to 50 Hz. In Fig. 4, the standard deviation error follows an exponential decay trend. The optimal performance is met for SNR larger than 75 dB, when the estimation error can be considered as negligible.

In Table I, we report the standard deviation errors as function of THD in the presence of three different low-frequency trends, i.e. none, noisy offset, and random walk. For this analysis, we set the SNR equal to 40 dB and we consider three different values of the fundamental frequency, i.e. $f = [45, 50, 55]$ Hz, that span the entire PMU pass-bandwidth. It is worth noticing that the estimation accuracy is independent from the harmonic distortion level and the low-frequency trend. As expected, the best performance is obtained in nominal conditions ($f = 50$ Hz), whereas the long-range spectral leakage from the negative frequency components results in an accuracy deterioration (e.g., for $f = 45$ Hz).

In the final test, we characterize the algorithm robustness towards inter-harmonic interferences. For this analysis, we fix

TABLE I
STANDARD DEVIATION ERRORS [PPM] AS FUNCTION OF THD

| Frequency [Hz] | Low-Freq. Trend | THD [%] | | | |
|---|---|---|---|---|---|
| | | 10 | 20 | 30 | 40 |
| 45 | none | -54.953 | -54.950 | -54.946 | -54.912 |
| | noisy offset | -54.964 | -54.958 | -54.949 | -54.931 |
| | random walk | -54.962 | -54.956 | -54.951 | -54.933 |
| 50 | none | -49.766 | -49.751 | -49.738 | -49.629 |
| | noisy offset | -49.783 | -49.764 | -49.757 | -49.648 |
| | random walk | -49.796 | -49.789 | -49.783 | -49.676 |
| 55 | none | -52.032 | -52.026 | -52.018 | -52.012 |
| | noisy offset | -52.139 | -52.131 | -52.125 | -52.112 |
| | random walk | -52.244 | -52.227 | -52.216 | -52.203 |

THD and SNR equal to 40% and 40 dB, respectively, and we vary the fundamental frequency within the range [47.5, 50, 52.5] Hz. In Fig. 5, we show the standard deviation error as function of the inter-harmonic frequency. Once more, the estimation accuracy proves to be independent from the additive distortion contributions, and the standard deviation errors are comparable with the results of the previous tests.

## IV. MODELING UNCERTAINTY

In this Section, we evaluate the uncertainty contribution introduced by an improper modeling of the statistical distribution of the measurement noise. In other words, we discuss the algorithm estimation accuracy when the assumption of uncorrelated white Gaussian noise is not verified.

To this end, we consider a test waveform where the wide-band measurement noise is taken from a Rayleigh probability density function. In general, this statistical model accounts for a two-dimensional vector, whose coordinates vary according to independent and identically distributed zero-mean Gaussian distributions, and is typically employed to represent fading or scattering phenomena.

For this analysis, we set the fundamental component amplitude, frequency, and initial phase equal to 1 pu, 50 Hz, and 0 rad, respectively. In concordance with the
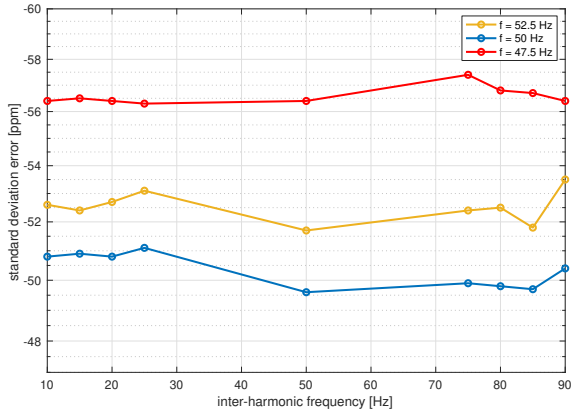
Fig. 5. Standard deviation estimation error as function of the inter-harmonic frequency (TIHD = 10%) with different fundamental frequency values: 47.5 (blue), 50 (red), 52.5 Hz (yellow).
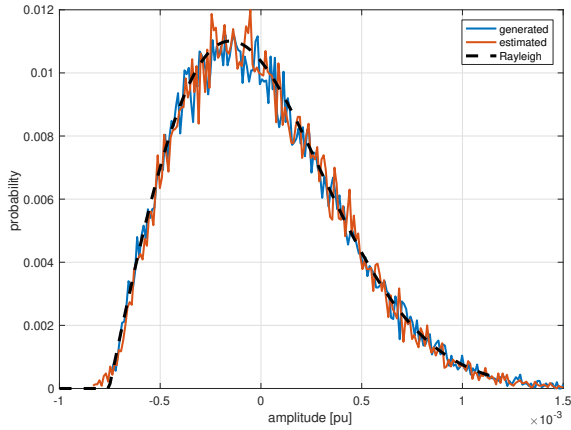


Fig. 6. Histograms of the generated and estimated noise (in blue and red line, respectively), compared with the probability density function of the employed Rayleigh distribution (in black dashed line).

previous numerical validation, the narrow-band interference contribution consists of the first 50 harmonic terms with a THD of 40%. As regards the additive measurement noise, we consider a Rayleigh-distributed random variable, whose scale parameter is set equal to 280 ppm. Furthermore, a continuous offset is added in order to center the probability density function around zero.

In Fig. 6, we show the histogram representations associated to the generated and estimated measurement noise. It is worth noticing how the statistical distributions are nearly coincident and accurately approximate the ideal Rayleigh function.

In this regard, Table II reports the standard deviation error[3] as function of the harmonic distortion level. Independently from the considered THD value, the proposed approach tends to under-estimate the actual noise standard deviation. Nevertheless, the introduced error can be reasonably considered negligible as it does not exceed 7 ppm.

---

[3]By construction, the standard deviation of a Rayleigh-distributed random variable is directly proportional to its scale factor.

| THD [%] | Std Dev Err [ppm] | SNR [dB] | ER [dB] |
|---------|-------------------|----------|---------|
| 10 | -6.526 | 65.23 | 61.58 |
| 20 | -6.541 | 65.29 | 61.65 |
| 30 | -6.540 | 65.35 | 61.71 |
| 40 | -6.571 | 65.56 | 61.89 |

Even if the noise modeling assumption is not verified in the present case, it is still possible to extract useful information regarding the noise energy content by computing the traditional SNR index. In this context, Table II compares the SNR values provided by Matlab routines with a similar index derived directly from the algorithm estimates. In particular, we define the Energy Ratio (ER) index, that is computed as the ratio between the energy associated to the recovered signal $\hat{y}$ and to the estimated measurement noise $\varepsilon$. It is worth noticing that ER slightly under-estimates SNR by less than 5 decibels. Nevertheless, both the indexes suggest an equivalent resolution of 10 bits for the simulated test waveforms.

Based on the obtained results, it is reasonable to say that the proposed approach guarantees a remarkable estimation accuracy also in the presence of non-normally distributed measurement noise. In particular, we are able to recover the original distribution of noise amplitude, and to infer significant power quality indexes, representative of the noise variation range and energy content.

## V. EXPERIMENTAL VALIDATION

The experimental validation of the proposed algorithm has been carried out by means of point-on-wave data acquired in a real-scale PMU installation [19]. Specifically, the data refer to a 20 kV substation of EPFL campus, coupled with a battery energy storage system (BESS) (see Fig. 7). The busbar is composed of 2 line departures connecting neighbor substations, 2 MV/LV 630 kVA transformers supplying electricity to office buildings, and one cable connecting the BESS.

The *waveform recorder* functionality has been integrated in the PMU developed in EPFL-DESL laboratory (for more details refer to [25]). The hardware platform is based on the National Instruments compactRIO 9068, an embedded industrial controller with a dual-core ARM Cortex-A9 processor and equipped with a customized Linux-RT OS, a user-programmable Field Programmable Gate Array (FPGA) Artix-7 equipped with a 40 MHz free-running clock and reconfigurable IO modules. The device has been equipped with a stationary GPS unit NI-9467 for the synchronization to UTC-time, accurate to within ± 100 ns and two analog input modules NI-9215 characterized by an input range of ± 10 V and a sampling frequency of 50 kHz. Medium voltage signals are acquired by dedicated instrument transformer manufactured by Altea Solutions (CS-50-I current class 0.2 and VS-24-C voltage class 0.5 sensor).
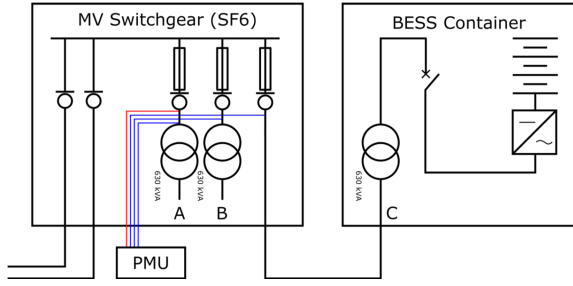
Fig. 7. Experimental setup of PMU-based waveform acquisition in EPFL campus [19].
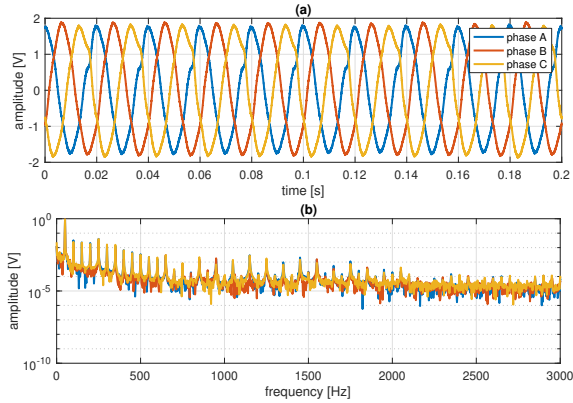


Fig. 8. Real waveforms acquired in the EPFL campus distribution network in time (a) and frequency (b) domain.
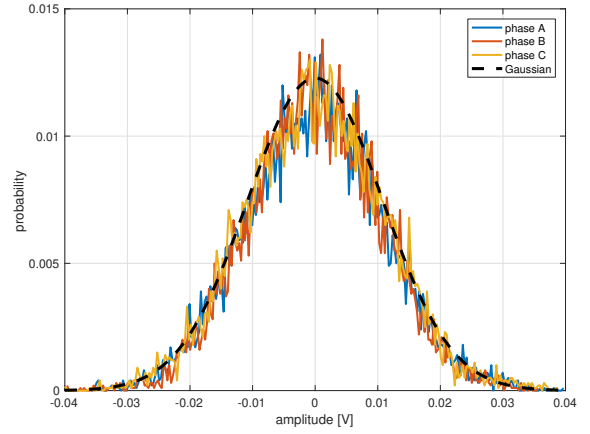


Fig. 9. Histograms of the estimated noise in the three phases (in blue, red and yellow line, respectively), compared with the Gaussian distribution derived from the mean and standard deviation of phase A noise (in black dashed line).

TABLE III
EPFL CAMPUS DATA - NOISE MODEL STATISTICAL PROPERTIES

| Phase | | Mean [V] | Std Dev [V] | | SNR [dB] | ER [dB] |
|---|---|---|---|---|---|---|
| A | | $6.95 \cdot 10^{-07}$ | 0.0108 | | 34.27 | 32.35 |
| B | | $4.35 \cdot 10^{-05}$ | 0.0104 | | 35.18 | 34.89 |
| C | | $-4.15 \cdot 10^{-05}$ | 0.0110 | | 34.62 | 32.13 |

The BESS is composed by Lithium titanate oxide (LTO) cells arranged in series and parallel to form 9 battery racks connected to a DC bus whose voltage oscillates between 590 and 810 V. The DC bus is interfaced with the medium-voltage grid through a four quadrant DC-AC converter and a 0.3/21 kV, 630 kVA transformer [26].

In this context, the present analysis considers a three-phase voltage waveform, characterized by an average THD and SNR approximately equal t0 20% and 35 dB, respectively. As shown in Fig. 8, the entire observed spectrum is affected by harmonic components, whose amplitude is slowly decreasing and becomes almost negligible after the $40^{th}$ order.

Since the adopted input model does not employ any low-pass filter, the acquired waveforms might be affected by aliasing effects. In order to quantify this distortion source, we compare the energy associated to the time-domain acquired waveform and to the DFT coefficients included in the Nyquist bandwidth, i.e. between 0 and 25 kHz. We repeated this analysis on all the three phases, and we obtained a worst-case discrepancy equal to 11.67 ppm. As a consequence, it is reasonable to say that the aliasing effects are negligible in the considered measurement setup.

In Fig. 9, we show the histogram representation of the estimated measurement noise. Based on the comparison with the ideal normal probability density function derived from phase A noise, it is reasonable to say that the assumption of purely additive Gaussian noise is verified. As further proof,

we perform a one-sample *t-test* on the estimated measurement noise, that confirms the null hypothesis of normal distribution with a probability exceeding 99% for all the three phases.

Table III reports the mean and standard deviation as derived from the algorithm estimates. In this regard, it is worth noticing that phase B and C presents a larger DC component, whereas the standard deviation values are comparable in the three phases. In concordance with the previous Section analysis, we compare SNR and ER indexes: the proposed algorithm proves to be able of providing rather accurate approximation of the real-world measurement noise, with an estimation error lower than 1 dB.

Finally, we verify *a posteriori* the assumption of uncorrelated noise components. To this end, we compute the mutual correlation index between the recovered signal $\hat{y}$, and the estimated low-frequency trend $\eta$ and the wide-band measurement noise $\varepsilon$. As shown in Table IV, the recovered signal is almost uncorrelated with both the noisy contributions, with a correlation index in the order of 0.8% and 1% for $\eta$ and $\varepsilon$, respectively. It should be noticed that the two noisy contributions are characterized by a slightly larger similarity, even if the correlation index does not exceed 4.6%.

In conclusion, the present experimental validation has confirmed the likelihood of our modeling assumptions. On one side, the measurement noise can be accurately approximated by an additive white Gaussian random variable. On the other side, the informative and noisy components prove to be statistically independent and uncorrelated.

TABLE IV
EPFL CAMPUS DATA - NOISE MODEL CORRELATION PROPERTIES

| Phase | Corr($\hat{y}, \varepsilon$) | Corr($\hat{y}, \eta$) | Corr($\varepsilon, \eta$) |
|-------|--------|--------|--------|
| A | 0.0147 | 0.0056 | 0.0339 |
| B | 0.0133 | 0.0034 | 0.0456 |
| C | 0.0116 | 0.0081 | 0.0338 |

## VI. Conclusions

In this paper, we introduced a novel processing technique to determine the statistical distribution parameters of additive measurement noise in real-world acquisitions. The proposed method is intended to provide a useful tool to evaluate the noise properties in experimental scenarios, as well as a predictor of realistic noise variance values, that can be used in many distribution network modeling and control applications. A detailed representation of the noise statistical properties allows for more accurate state estimations and enables us to identify possible non-linear effects in the acquisition system.

First, we described in detail the main stages of the processing technique and explained the criteria for the most suitable parameter setting. In this regard, we particularly focused on the accurate estimation of fundamental and interfering components in off-nominal conditions.

Then, we carried out a numerical validation to evaluate the estimation accuracy as function of SNR, THD, low-frequency trends, and possible inter-harmonic components.

To account for possible improper modeling assumptions, we also simulated a test waveform corrupted by a Rayleigh-distributed measurement noise. We proved the accuracy of the recovered measurement noise, also when the assumption of purely additive white Gaussian noise is not verified.

Finally, we validated the proposed algorithm on a point-on-wave data acquired on a PMU installation in EPFL campus, as representative of a real-world distribution network. This test enables us to evaluate the algorithm robustness in the presence of high harmonic distortion and noise level, and verify a posteriori the initial modeling assumptions. Even in this challenging scenario, the obtained results confirmed the accuracy of the recovered noise model.

## References

[1] I. Kamwa, S. R. Samantaray and G. Joos, *"Compliance Analysis of PMU Algorithms and Devices for Wide-Area Stabilizing Control of Large Power Systems,"* in IEEE Trans. on Pow. Sys., vol. 28, no. 2, pp. 1766-1778, May 2013.

[2] C. Muscas, M. Pau, P. Pegoraro, S. Sulis, F. Ponci, A. Monti, *"Multiarea Distribution System State Estimation,"* in IEEE Trans. on Instr. and Meas., vol. 64, no. 5, pp. 1140-1148, May 2015.

[3] M. Pignati, L. Zanni, P. Romano, R. Cherkaoui, M. Paolone, *"Fault Detection and Faulted Line Identification in Active Distribution Networks Using Synchrophasors-Based Real-Time State Estimation,"* in IEEE Trans. on Pow. Del., vol. 32, no. 1, pp. 381-392, Feb. 2017.

[4] *"IEEE Standard for Synchrophasor Measurements for Power Systems,"* in IEEE Std C37.118.1-2011, Dec. 2011.

[5] *"IEEE Standard for Synchrophasor Measurements for Power Systems – Amendment 1: Modification of Selected Performance Requirements,"* in IEEE Std C37.118.1a-2014, April 2014.

[6] Paulo F. Ribeiro, *"Time-Varying Waveform Distortions in Power Systems,"* Wiley-IEEE Press, 2010.

[7] M. Bertocco, C. Narduzzi, P. Paglierani, D. Petri, *"A Noise Model for Digitized Data,"* in IEEE Trans. on Instr. and Meas., vol. 49, no. 1, pp. 83-86, Feb. 2000.

[8] F. Milano, *"Power System Modelling and Scripting,"* Springer, 2010.

[9] Z. Tao, Y. Xiaoxian, Z. Baohui, N. H. Xu, F. Xiaoqun and L. Changxin, *"Statistical Analysis and Modeling of Noise on 10-kV Medium-Voltage Power Lines,"* in IEEE Trans. on Pow. Del., vol. 22, no. 3, pp. 1433-1439, July 2007.

[10] M. Chakir, I. Kamwa and H. Le Huy, *"Extended C37.118.1 PMU Algorithms for Joint Tracking of Fundamental and Harmonic Phasors in Stressed Power Systems and Microgrids,"* in IEEE Trans. on Pow. Del., vol. 29, no. 3, pp. 1465-1480, June 2014.

[11] A. Derviškadić, P. Romano and M. Paolone, *"Iterative-Interpolated DFT for Synchrophasor Estimation: A Single Algorithm for P- and M-Class Compliant PMUs,"* in IEEE Trans. on Instr. and Meas., vol. 67, no. 3, pp. 547-558, March 2018.

[12] M. Bertocco, G. Frigo, C. Narduzzi, C. Muscas and P. A. Pegoraro, *"Compressive Sensing of a Taylor-Fourier Multifrequency Model for Synchrophasor Estimation,"* in IEEE Trans. on Instr. and Meas., vol. 64, no. 12, pp. 3274-3283, Dec. 2015.

[13] C. Narduzzi, M. Bertocco, G. Frigo and G. Giorgi, *"Fast-TFM—Multifrequency Phasor Measurement for Distribution Networks,"* in IEEE Trans. on Instr. and Meas., vol. 67, no. 8, pp. 1825-1835, Aug. 2018.

[14] S. R. Krishnan and C. S. Seelamantula, *"On the Selection of Optimum Savitzky-Golay Filters,"* in IEEE Trans. on Signal Processing, vol. 61, no. 2, pp. 380-391, Jan.15, 2013.

[15] A. Angioni, T. Schlösser, F. Ponci and A. Monti, *"Impact of Pseudo-Measurements From New Power Profiles on State Estimation in Low-Voltage Grids,"* in IEEE Trans. on Instr. and Meas., vol. 65, no. 1, pp. 70-77, Jan. 2016.

[16] L. Zanni, J. Le Boudec, R. Cherkaoui and M. Paolone, *"A Prediction-Error Covariance Estimator for Adaptive Kalman Filtering in Step-Varying Processes: Application to Power-System State Estimation,"* in IEEE Trans. on Control Systems Technology, vol. 25, no. 5, pp. 1683-1697, Sept. 2017.

[17] M. Brown, M. Biswal, S. Brahma, S. J. Ranade and H. Cao, *"Characterizing and Quantifying Noise in PMU Data,"* 2016 IEEE Power and Energy Society General Meeting (PESGM), Boston, MA, 2016, pp. 1-5.

[18] Z. Jin, H. Zhang, *"Noise Characteristics and Fast Filtering of Synchronized Frequency Measurement in Low Voltage Grid,"* 2016 IEEE Smart Energy Grid Engineering (SEGE), Oshawa, ON, 2016, pp. 398-403.

[19] *"Point-On-Wave Data Collected on the 20 kV Network of EPFL-Campus,"* [Online] https://github.com/DESL-EPFL/Point-on-wave-Data-of-EPFL-campus-Distribution-Network, Sep. 2018.

[20] D. Gallo, R. Langella and A. Testa, *"Desynchronized Processing technique for Harmonic and interharmonic analysis,"* in IEEE Trans. on Pow. Del., vol. 19, no. 3, pp. 993-1001, July 2004.

[21] J. A. de la O Serna, *"Dynamic Phasor Estimates for Power System Oscillations,"* in IEEE Trans. on Instr. and Meas., vol. 56, no. 5, pp. 1648-1657, Oct. 2007.

[22] *"IEC 61000 Electromagnetic compatibility (EMC) - Part 4-1: Testing and measurement techniques,"* in IEC 6100-4-1, Apr. 2016.

[23] A. Testa, D. Gallo, R. Langella, *"On the Processing of Harmonics and Interharmonics: Using Hanning Window in Standard Framework,"* in IEEE Trans. on Pow. Del., vol. 19, no. 1, pp. 28-34, Jan. 2004.

[24] G. Frigo, A. Derviškadić, C. Narduzzi and M. Paolone, *"Synchrophasor-Based ROCOF Measurements: Feasibility in Real-World Scenarios,"* in 2018 IEEE International Workshop on Applied Measurements for Power Systems Proceedings (AMPS), Bologna, 2018, [accepted].

[25] P. Romano and M. Paolone, *"Enhanced Interpolated-DFT for Synchrophasor Estimation in FPGAs: Theory, Implementation, and Validation of a PMU Prototype,"* in IEEE Trans. on Instr. and Meas., vol. 63, no. 12, pp. 2824-2836, Dec. 2014.

[26] F. Sossan, E. Namor, R. Cherkaoui and M. Paolone, *"Achieving the Dispatchability of Distribution Feeders Through Prosumers Data Driven Forecasting and Model Predictive Control of Electrochemical Storage,"* in IEEE Trans. on Sust. Energy, vol. 7, no. 4, pp. 1762-1777, Oct. 2016.