CHALMERS

Chalmers Publication Library

A Method to Evaluate Harmonic Model-Based Estimations under Non-White Measured Noise

This document has been downloaded from Chalmers Publication Library (CPL). It is the author´s version of a work that was accepted for publication in:

(article starts on next page)

A Method to Evaluate Harmonic Model-Based Estimations under Non-White Measured Noise

Cuong D. Le, Math H. J. Bollen, and Irene Y. H. Gu

Abstract—Automatic extracting information from power-system event recordings requires applications of signal-processing estimation techniques whose performance has been verified under white noise. This paper proposes a method to test these techniques under real power-system noise, which is very different from white noise, to evaluate their application feasibility. The first part of the paper describes the evaluation method used to evaluate the techniques in a statistical sense and a method to extract noise from measured power-system recordings. The second part of the paper focuses on the evaluation of a number of harmonic model-based techniques under non-white noise, including: Kalman filter, MUSIC, ESPRIT, and segmentation algorithms. The paper shows that for the Kalman filter, a very high order with high computational burden is necessary only if high frequency components are of interest. The application of MUSIC, ESPRIT, and the segmentation algorithms under natural power-system noise is shown to be feasible.

Index Terms—harmonic analysis, performance evaluation, power quality, signal-processing applications.

I. Introduction

SIGNALprocessing techniques are widely applied in power engineering, especially those related to the estimation of harmonics/interharmonics [1]-[2]. A number of currently used estimation techniques in power engineering employ a harmonic model under the assumption that the observation signal is a combination of different harmonics plus white noise [3]. However, measurements show that the spectrum of real noise is far from that of white noise, and this may impact the performance of the estimation. Some properties of power-system noise can be found, e.g., in [4]. Thus, it is desired to evaluate the performance of these techniques under real (measured) noise.

The main difficulty to evaluate the estimation techniques applied in power engineering is the lack of the *"ground truth"* signal to compare with. If synthetic data with white noise is used, the parameters of the signal are predefined which are the needed reference but the noise is not realistic. On the other hand, if the measurement data is used, the model is exposed to the real measurement noise (non-white) but there is no reference to compare with since the content of the measurement signal is unknown. A previous study [5] proposes a method to create semi-synthetic data sequences by embedding the measured noise into the synthetic data through superimposing

real measurements with synthetic data. Such semi-synthetic data sequences are used in [5] to evaluate the performance of estimated frequencies and amplitudes of damped sinusoids from a frequency estimation method, ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) [6], which is able to detect and quantify harmonic and inter-harmonic components in a time-varying voltage or current signal [7].

In this study, we introduce a systematic method to evaluate the performance of harmonic model-based estimation techniques under real power-system noise. The first application of the proposed method is to find the model order small enough to have less computational burden but acceptable accuracy (i.e. trade-off between model order and computational demand). This is based on the observation that the higher the model order, the closer the noise spectrum moves towards white (as more harmonics are shifted from the noise subspace to the signal subspace), hence the more accurate the estimation becomes but the heavier the computational burden is. The second application is to evaluate the model-based estimation techniques where the observed noise deviates from the white noise assumption in the model, and to obtain statistic properties of their performance under different measured (color) noise levels and realizations. The method is applied but not limited to a number of estimation techniques under harmonic models, including Kalman filters, ESPRIT [6], MUSIC (Multiple Signal Classification) [8], and a segmentation algorithm [9]. The performance is evaluated using semi-synthetic data (i.e., synthetic signal plus measured power-system noise).

In the first case study, the impact of model order selection on the estimation accuracy of model-based techniques for harmonic estimation is studied as there is a trade-off between the estimation accuracy and computational burden. To study the impact of the model order on the estimation accuracy, we evaluate the rms error of the Kalman filter applied to both clean and noisy signals. In the second case study, three estimation methods are evaluated under different signal-to-noise ratios (SNRs). Each method is evaluated using numerous realizations of semi-synthetic data sequences (i.e., different measured noise sequences are superimposed into a synthetic signal). The three methods studied are the Kalman filter-based event segmentation, the ESPRIT and MUSIC based harmonic estimation methods.

The remaining part of the paper is organized as follow: Section II briefly describes the evaluation method. The measurement noise extraction method is presented in Section III, and the simulation results of the study are presented in Section IV. Conclusions and future work are drawn in Section V.

II. THE EVALUATION METHOD

The evaluation method is a statistically-based method. A large number of measurements of voltage and/or current in the power systems are collected. In principle, the measurements should be done at the locations where the methods and applications are likely to be applied. Since the measured noise also contains the dominant fundamental voltage or current that is typically at least an order of magnitude larger than the true power-system background noise of interest, pre-filtering is first applied to remove or suppress the fundamental. The result from the pre-filter to each measurement sequence is used as one noise realization.

Each noise realization is then added to a synthetic signal with predefined parameters to create the so-called semi-synthetic data that is used to evaluate the selected estimation techniques. The output of the estimation technique is compared with the predefined parameters of the synthetic input signal. The process is repeated for all noise realizations to obtain statistical properties for the performance of the selected estimation techniques. The estimated values are compared with the predefined parameters in the synthetic signal and a number of metrics for performance evaluations of the estimation technique are calculated. The performance metrics used in our study include mean values, rms errors, and standard deviation of the estimation error.

III. NOISE EXTRACTION

A. Signal and Noise Model

Under the harmonic model, a signal is modeled as the sum of K harmonics by (1).

$$y(n) = \sum_{m=1}^{M} s_m(n) + v(n) = \sum_{m=1}^{M} A_m e^{j\omega_m n} + v(n) \quad (1)$$

where $s_m(n)$ is the mth harmonic component and $v(n)$ is the model noise assumed to be white. Under this model, apart from the K harmonics of interest, the remaining of the semi-synthetic signal $y(n)$ is considered as noise which means that the second and all higher order harmonics in the input signal are treated as noise if the model order is one. It is obvious that in measured power system data, the noise is not white as there always exist some harmonics. Thus, it is important to examine the impact of the model order as well as different noise levels and noise realizations.

B. Noise Extraction

As described in Section I, the step of noise extraction is aimed at maximally removing 50/60-Hz fundamental voltage/current in each measured noise sequence. An optimal linear-phase finite impulse response (FIR) filter using the Parks-McClellan algorithm ([10]) is used to remove the fundamental from the measurement data to obtain different realizations of noise sequences. Fig. 1 and Fig. 2 show an example of noise extraction, where the waveforms and FFT spectra before and after the noise extraction are included.
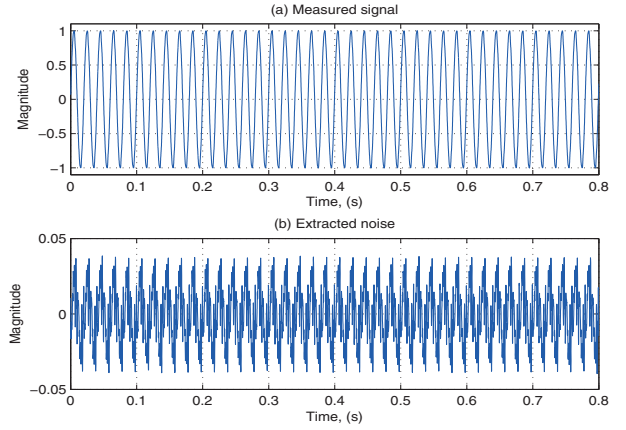

Fig. 1. Waveforms of measured noise embedded in 50-Hz voltage (top) and the extracted noise (bottom). Note the difference in vertical scale.
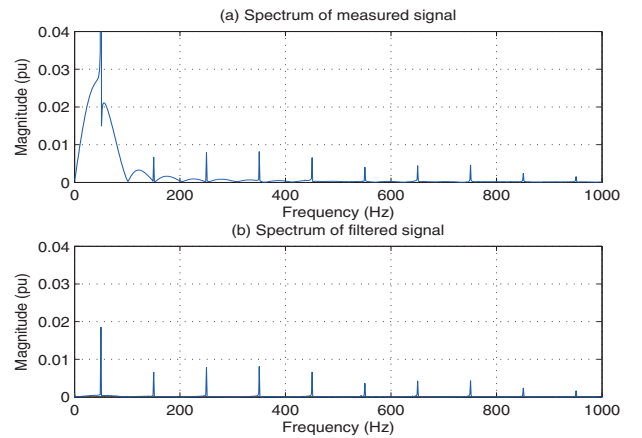

Fig. 2. Spectra of the measured noise in 50-Hz voltage (correspond to Fig. 1(a)) and the extracted noise (correspond to Fig. 1(b)).

From Fig. 2 one can see that apart from the fundamental component (whose magnitude is unity but cut off for better display in Fig. 2(a), there is no significant difference in the harmonic components of the measured signal and the extracted noise at frequencies above 200 Hz. However, the broadband components around the fundamental frequency disappear from the extracted noise due to the bandpass filter.

In order to quantify this realistic noise, a signal-to-noise ratio with a pure 50-Hz sine wave used as the reference signal is defined in (2).

$$SNR = 10 \log_{10} \frac{P_{signal}}{P_{noise}} \quad (2)$$

where P_{signal} and P_{noise} are the power of the reference sine wave and the extracted noise, respectively. In this study, a set of 366 voltage recordings are used to extract 366 noise realizations; and SNRs for these noise realizations are calculated according to (2) and presented in the histogram in Fig. 3.

The histogram shows that the defined SNR for the extracted noise sequences (the original extracted noise without scaling, denoted by natural noise sequences) varies from 32 dB to
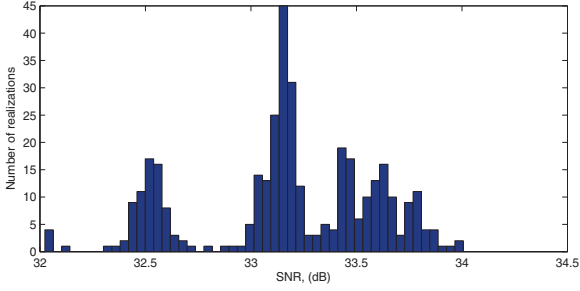
Fig. 3. Histogram of SNRs of 366 extracted noise realizations.

34 dB and the mean value is 33.17 dB.

IV. SIMULATIONS AND RESULTS

In this part we present the evaluation results from the study for the Kalman filter, ESPRIT, MUSIC, and a segmentation algorithm.

A. Evaluate the impact of harmonic model order using Kalman filter-based estimation on quasi-stationary signals

The impact of model order on the performance of Kalman filters under the harmonic model on quasi-stationary signal is studied for two cases: one is for a clean synthetic voltage containing only the fundamental component, and another is a noisy semi-synthetic voltage obtained by superimposing the real measurement noise into the clean synthetic voltage. To quantify the performance, the rms error in (3) is used.

$$e_{rms} = \sqrt{\frac{\sum_{i=1}^{L}[s(i) - \hat{s}(i)]^2}{L}} \tag{3}$$

where $s(i)$ and $\hat{s}(i)$ are the ith sample of the clean signal and the estimated signal, respectively, and L is the length of the signal. Fig. 4 shows the rms errors as a function of the harmonic model order.
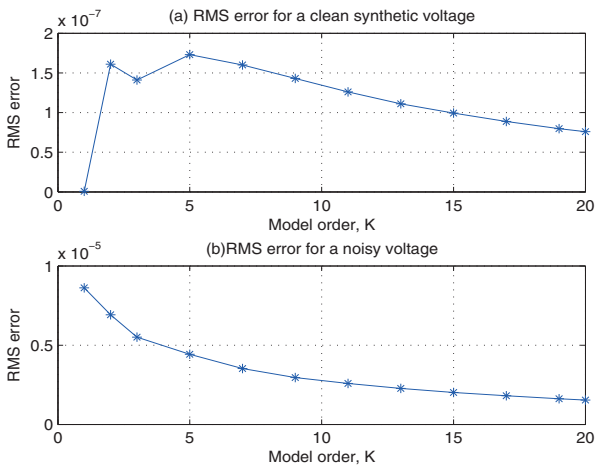

Fig. 4. rms errors vs. model order. Note that there is a factor 50 difference in vertical scale between the two graphs.

From Fig. 4(a) one can see that the rms error for the clean signal reaches its minimum value at $K = 1$ since the signal only contains the fundamental component. In the case of using semi-synthetic data, the rms error in Fig. 4(b) decreases with the increase of model order due to the fact that more harmonics in the noise are moved to the signal subspace. Although, in both cases, the rms errors are rather small, the error for a noisy voltage is a factor of 50 higher than the one for a clean synthetic voltage. The model order K, thus, in the case of application for quasi-stationary noisy signals can be freely selected depending on the number of harmonics to be estimated. The very high model order with heavy computational burden is not necessary if only several power harmonic components are of interest.

B. Evaluate the performance of ESPRIT and MUSIC using semi-synthetic data

In order to evaluate the performance of ESPRIT and MU-SIC, a synthetic signal consisting of two sinusoids (50 Hz, 1 pu and 150 Hz, 0.3 pu) is generated. The set of 366 extracted noise realizations are added to this synthetic signal to create a set of 366 semi-synthetic noisy signals for each selected SNR. These semi-synthetic sequences are then used to evaluate the harmonic model using ESPRIT and MUSIC algorithms. The means and the standard deviations of the estimated parameters from these 366 noisy signal with different noise levels (SNRs) are then calculated. The results for frequency estimation using the two algorithms are plotted in Fig. 5 and Fig. 6 for the 50-Hz component and the 150-Hz component, respectively.
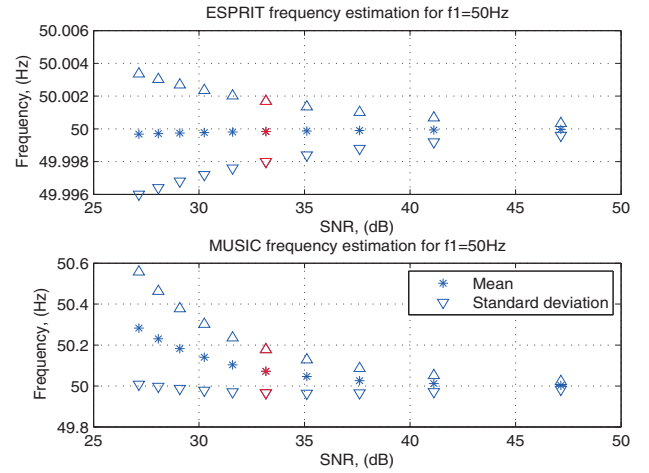

Fig. 5. Frequency estimation results for the 50-Hz component. The red points represent the case of SNR= 33.17 dB which corresponds to the case the extracted noise is directly added to the synthetic signal (natural noise).

From the figure, one can see that the estimated frequencies converge to the ground-truth frequencies with the increase of SNR (or the decrease of noise level). A systematic over-estimation is observed for MUSIC and a slight under-estimation is observed for ESPRIT. Table I shows the estimated parameters of the two algorithms at two selected noise levels: SNR = ∞ (no noise), and SNR = 33.17 dB (natural
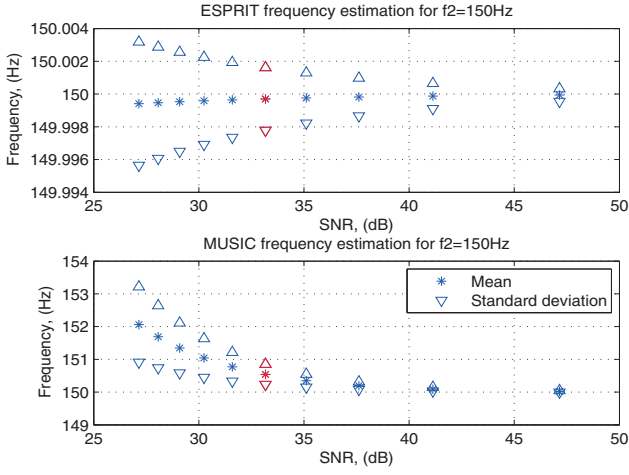
Fig. 6. Frequency estimation results for the 150-Hz component.

TABLE I
FREQUENCY ESTIMATES FOR NO-NOISE AND NATURAL-NOISE SCENARIOS
(μ: MEAN VALUE; σ: STANDARD DEVIATION; F1, F2: FREQUENCIES).

-	SNR=∞ (no noise)		SNR = 33.17 dB (natural noise)			
-	f1	f2	$\mu(f1)$	$\sigma(f1)$	$\mu(f2)$	$\sigma(f2)$
MUSIC	49.99	149.99	50.08	0.12	150.92	0.32
ESPRIT	50.00	150.00	49.99	8e-4	150.00	7e-4

TABLE II
AMPLITUDE ESTIMATES FOR NO-NOISE AND NATURAL-NOISE SCENARIOS
(μ: MEAN VALUE; σ: STANDARD DEVIATION; A1, A2: AMPLITUDES).

-	SNR=∞ (no noise)		SNR = 33.17 dB (natural noise)			
-	A1	A2	$\mu(A1)$	$\sigma(A1)$	$\mu(A2)$	$\sigma(A2)$
MUSIC	0.9765	0.318	0.9755	0.0130	0.3182	0.0024
ESPRIT	1.00	0.30	0.9984	0.0129	0.3018	0.0014

noise). In a similar way, amplitude estimation results are presented in Fig. 7, Fig. 8, and Table II.
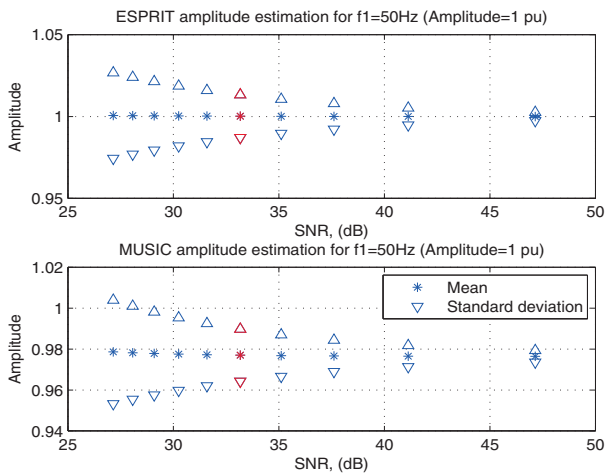

Fig. 7. Amplitude estimation results for the 50-Hz component.

From the results, one can see that for the dominant fundamental component, the mean values tend to be stable at
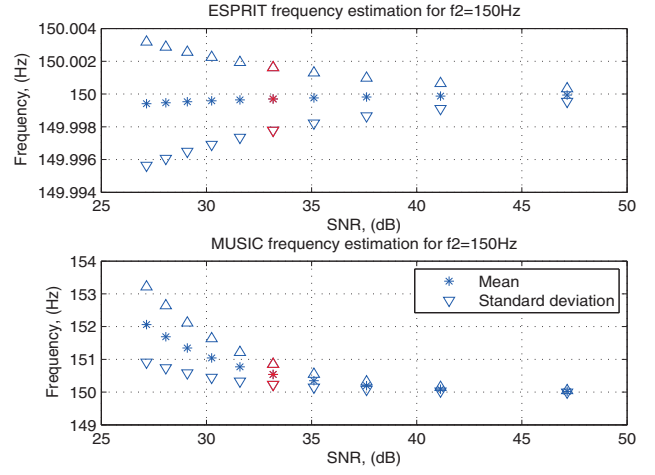

Fig. 8. Amplitude estimation results for the 150-Hz component.

TABLE III
FREQUENCY AND AMPLITUDE ESTIMATES CORRESPONDING TO THE
MAXIMUM ERROR UNDER NATURAL NOISE SCENARIO.

-	f1	f2	A1	A2
MUSIC	50.56	152.43	0.957	0.4265
	(1.11%)	(1.62%)	(4.303%)	(42%)
ESPRIT	49.98	149.98	1.022	0.3067
	(0.04%)	(0.01%)	(2.2%)	(2.23%)

different noise levels but the standard deviations increase with increase of noise level. For the third harmonic component, overestimation is observed for both algorithms. This is due to the contribution of the third harmonic component in the noise. MUSIC shows consistent underestimation for the dominant component and overestimation for the weak component.

In order to observe the performance at the natural noise scenario, the estimates of the two algorithms under this noise level for 366 realizations are presented in Fig. 9, Fig. 10. The estimates corresponding to the maximum error are given in Table III.
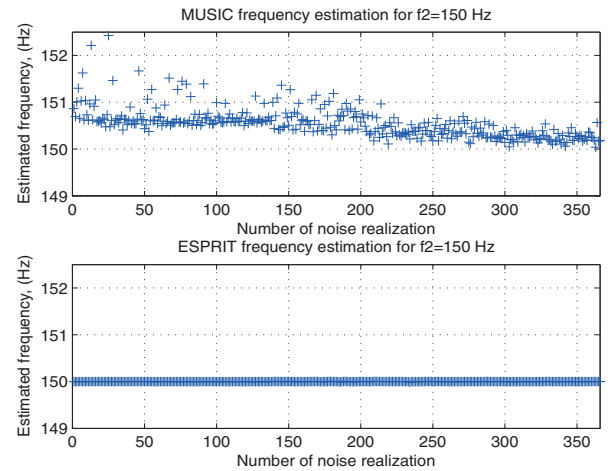

Fig. 9. Frequency estimates of 150-Hz component at the natural noise level.
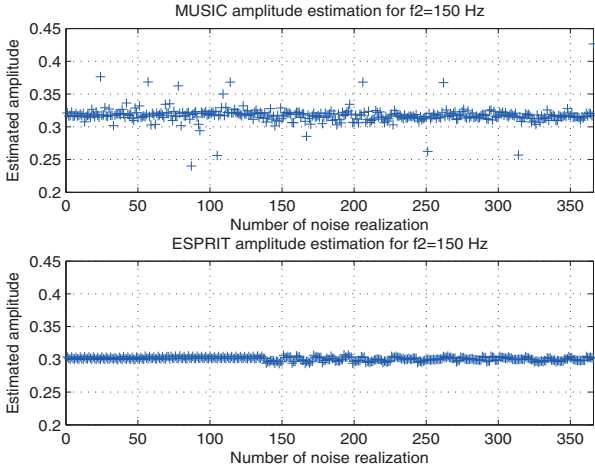
Fig. 10. Amplitude estimates of 150-Hz component at the natural noise level.

Fig. 11. Estimation errors for segmentation algorithm.

The results show that ESPRIT shows very good and stable performance over different noise realizations. The aim of this study is, however, not to compare the performance of the two algorithms as the settings for the two algorithms are different: the signal subspace dimension is half of the data length for ESPRIT and 10 (equal to the number of sinusoids to be estimated) for MUSIC. The aims are to illustrate the evaluation method using semi-synthetic data and to show that the performance of both ESPRIT and MUSIC is good enough under real power-system noise. Table III shows that ESPRIT gives very good performance at the natural noise level; the maximum error is 0.04% for frequency estimation and 2.23% for amplitude estimation. MUSIC gives higher error in most of the cases and the maximum error is observed for the case of amplitude estimation for the weak signal. This is due to the quite low signal subspace dimension setting.

C. Evaluate the performance of segmentation algorithm

The proposed evaluation method is also applied to the segmentation algorithm proposed in [9]. This segmentation algorithm calculates a forward (causal) and backward (anti-causal) version of the Kalman filter residuals to very accurately detect the location of transition segments. For the purpose of this study, the set of 366 extracted noise realizations is added to a synthetic fast transition obtained by multiplying a sine wave with a step transition at time T_0 to create 366 transition sequences. The next step is to apply the segmentation algorithm to these sequences to evaluation its performance. The difference (error E) between T_0 and the estimated transition instant ($\hat{T}_0$) is used to evaluate the performance. In the case of no noise, the error E is 5 samples (0.5 ms as the sampling frequency is 10 kHz). The evaluation results are shown in Fig. 11.

The top plot in Fig. 11 shows that the mean error in most of the cases are very close to 5 samples which is the error of the case of no noise. The standard deviations show systematic decrease with increase of SNR. The estimates in the case of natural noise are presented in the bottom plot of Fig. 11 which vary from 2 samples to 8 samples. The results show
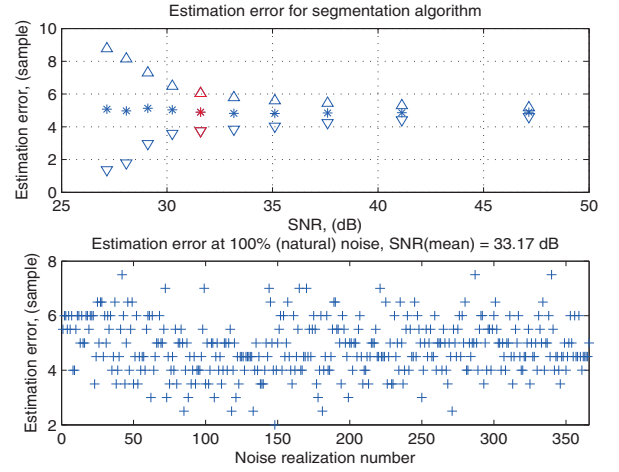
that the segmentation algorithm is not significantly impacted by the presence of noise. In contrast, the presence of noise somewhat reduces the estimation error; this is due to the fact that when the noise increases, the model unfit increases which leads to increases in the detection parameter making it easier to detect the transition point. However, this does not imply that the overall performance of the segmentation algorithm is better with the presence of noise. At higher noise levels, the possibility of false alarm increases (Reference [9] is recommended for complete understanding of this issue).

V. CONCLUSIONS AND FUTURE WORK

A systematic method has been proposed to evaluate the performance of signal-processing techniques in power-system applications. The method combines synthetic signals with a large number of measured noise realizations. The method is motivated by the real world scenarios of non-white noise, where we propose to use synthetic data embedded in measurement noise to make the data close to the real world situation. Since non-white noise does not fit to the white noise assumption in the harmonic model, performance evaluation is conducted and shown in this paper in terms of mean and variance of the estimation under different estimation techniques in several applications. This evaluation method could be used to evaluate the performance of any estimation techniques applied to power-system voltages and currents.

The study shows that the algorithms behave differently under measured noise. In the case that Kalman filter is applied to quasi-stationary signal under harmonic models, the rms errors between the input and the estimated signals are rather small. If a few harmonics of power frequency are interested, it is not necessary to have high model order to reduce computational load. Further, mean estimates and standard deviations for semi-synthetic sequences under different SNRs and realizations are studied to show the impact of measured noise in power systems on the performance of ESPRIT and MUSIC algorithms. The results show that both ESPRIT and MUSIC give estimations with acceptable error under natural noise scenario and their application to power-system data is still feasible. In the case

of evaluation of the segmentation algorithm, the influence of the measured noise on the performance of the algorithm is not significant. The study presented in this paper has also shown that it is possible to evaluate signal-processing algorithms under realistic noise situations and that such an evaluation provides additional information beyond evaluating using white noise.

A large number of noise realizations have been used in this study to evaluate the algorithms but they all are collected from one measurement point. It is more interesting to observe the performance of the algorithm subjected to noise sequences obtained from various locations and voltage levels. Another issue to study next is to look at the influence of the broadband components around the fundamental frequency by applying different filtering techniques to remove the fundamental frequency.

REFERENCES

[1] M. H. J. Bollen and I. Y. H. Gu, *Signal processing of power quality disturbance*, A John Wiley & Sons, Inc., 2006.

[2] A. Testa, M.F. Akram, et al., "Interharmonics: Theory and Modeling," *IEEE Trans. Power Delivery*, vol. 22, no. 4, pp. 2335-2348, 2007.

[3] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, A John Wiley & Sons, Inc., 1996.

[4] T. Zheng, X. Yang, et al., "Research of noise characteristics for 10-kV medium-voltage power lines," *IEEE Trans. Power Delivery*, vol. 22, no. 1, pp. 142-157, 2007.

[5] A. Tjäder, I. Y. H. Gu, et al., "Performance evaluation for frequency estimation of transients using the ESPRIT: measure noise versus white noise," 13th *Int. Conf. Harmonics and Quality of Power (ICHQP)*, Wollongong, NSW, 2008.

[6] A. Paulraj, R. Roy, et al, "Estimation of signal parameters via rotational invariance techniques - Esprit," *Circuits, Systems and Computers,* 19th *Asilomar Conference on*, 1985.

[7] I. Y. H. Gu, M. H. J. Bollen, "Estimating interharmonics by using sliding-window ESPRIT," *IEEE Trans. Power Delivery*, vol. 23, no. 1, pp. 13-23, 2008.

[8] R. Schmidt, "Multiple emitter location and signal parameter estimation," *Antennas and Propagation, IEEE Transactions on*, vol. 34, no. 3, pp. 276-280, 1986.

[9] C. D. Le, I. Y. H. Gu, and M. H. J. Bollen, "Joint causal and anti-causal segmentation and location of transitions in power disturbances," *IEEE PES General Meeting*, Minneapolis, USA, 2010.

[10] T. Parks and J. McClellan, "Chebyshev approximation for nonrecursive digital filters with linear phase," *Circuit Theory, IEEE Transactions on*, vol. 19, no. 2, pp. 189-194, 1972.