TRANSIENT DISTURBANCE DETECTION FOR POWER SYSTEMS WITH A GENERAL LIKELIHOOD RATIO TEST

Xiufeng Song†, Zafer Sahinoglu‡, and Jianlin Guo‡
†Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT
‡Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA

***Abstract*—A voltage/current transient typically caused by islanding and switching operations is treated as an adverse phenomenon that degrades power quality, and it may cause damage to electrical equipment. Therefore, a reliable system should effectively detect and monitor a transient disturbance. In this paper, the transient detection problem is formulated as a binary hypothesis test: normal signal (null) vs. transient (alternative). The sampled data is described by a sinusoid under the null hypothesis, while a sum of damped sinusoids is utilized to model the alternative one. As no prior knowledge is imposed on complex amplitudes, frequencies, or damping factors in signal modeling, the general likelihood ratio test (GLRT) is employed to fulfill the task. To reduce computational complexity, the maximum likelihood estimator is replaced by ESPRIT for parameter estimation. Probability of detection of 0.98 is achieved at a SNR of 27dB and probability of false alarm of 0.0005.**

***Index Terms*—Power system transients, detection algorithms, maximum likelihood estimation**

I. INTRODUCTION

A transient refers to a short time voltage/current disturbance from one steady-state into another, and it is attributed to a sudden load and generation mismatch on the distribution network [1]–[7]. Even though some natural phenomena such as lightning can occasionally cause transient, majority of them are internally generated: load switching, breaker switching, fuse disconnection, short-circuit, and islanding [1]. The amount of voltage/current change is case dependent. For example, a short-circuit may bring on a large current increase in a faulted area of the system. The cumulative effect of transients is treated as a major threat to delicate semiconductors, which have been extensively integrated into modern electric systems. As a consequence, the whole power system nowadays becomes less tolerant of transient disturbances [2].

Transient is one of the unfavorable factors that degrade the power quality, and it has attracted many efforts in the last twenty years. For example, a power transient due to load switching is comprehensively summarized in [1], and spectrum analysis of transients is presented in [2]. A general categorization of transients based on their reasons is discussed in [3]. The difference between the voltage transient and voltage sag is investigated in [4]. Statistically, those works are either nonparametric or parametric. The former does not impose any structure conjecture on observations, and classic approaches include Fourier and wavelet transforms base spectrum analysis [2] [8]. Nonparametric approaches are computationally friendly; however, they may subject to the constraint of low resolution or to the difficulty of signature explanation [2] [4]. Since electric circuits can be described by differential equations, a sum of damped sinusoids becomes a popular approach to parametrically model power transients [3]–[6]. A parametric approach offers an opportunity to accurately extract individually (damped) sinusoidal components from contaminated observations, and it may further benefit diagnose and failure identification of a power system [7].

The previous publications focus on theoretical analysis, parameter estimation, and classification of transients, while this paper addresses *transient detection*, which has not been widely attended to the best of our knowledge. The detection problem will be parametrically investigated. Intuitively, detection action is to statistically conjecture whether the interested part of a power system is normal via properly utilizing noisy observations. Decision results can be used to trigger protection signals, or, they can provide an insight to the stability of a power system [2].

In this paper, we formulate the transient detection problem as a binary hypothesis test, and derive a general likelihood ratio test (GLRT). The GLRT involves the maximum likelihood estimation (MLE) of unknown model parameters; its computational complexity is high. Therefore, we suggest the utilization of the ESPRIT algorithm [9] to replace the MLE in parameter estimation. As a supplementary contribution, we prove that the ESPRIT approach—a typically stationary signal processing tool—is theoretically applicable for the parameter estimation of the nonstationary sum of damped sinusoids. This analysis has not appeared anywhere else based on our knowledge. Note that our method does not require any prior knowledge of the fundamental frequency, initial phase, or amplitude. Therefore, its performance is preserved under nonideal situations such as frequency deviation.

The rest of this paper follows: Section II states the binary detection problem and gives the signal model; Section III implements the GLRT; numerical results are shown in Section IV, while the conclusions are drawn after that.

II. PROBLEM STATEMENT

We are interested in a single-phase power system in this paper, and the extension to a three-phase system is straightforward. Under normal conditions, the sampled voltage/current signal is sinusoid

$$y(n) = a_0 e^{j\omega_0 n} + v(n), \tag{1}$$

where n is an integer representing the sampling index, a_0 denotes the *complex amplitude* (a product of the amplitude and the initial phase) of the fundamental frequency component, $v(n)$ represents the receiver noise, and ω_0 denotes the normalized fundamental frequency. The analogy fundamental frequency $\tilde{\omega}_0$ can be obtained via

$$\tilde{\omega}_0 = \frac{\omega_0}{\Delta t}, \tag{2}$$

where Δt denotes the sampling interval. Note that as signal models for voltage and current are mathematically identical, no separate discussion will be further offered.

If the power system suffers from disturbances such as load switching and islanding, the voltage/current signal will undergo certain kind of transient. Generally, these transients can be modeled as a sum of damped sinusoids [3]–[5]

$$y(n) = \sum_{i=1}^{M} a_i e^{-\gamma_i n} e^{j\omega_i n} + v(n), \tag{3}$$

where $\gamma_i > 0$ and ω_i denote the normalized damping factor and the normalized frequency, respectively, for the ith components, a_i is the complex amplitude, and $M > 0$ counts the number of sinusoids. The analogy damping factor $\tilde{\gamma}_i$ and transient frequency $\tilde{\omega}_i$ can be similarly obtained as that in (2). Note that we set $e^{j\omega_i - \gamma_i} \neq e^{j\omega_k - \gamma_k}$ for $i \neq k$ in (3) to avoid model trivialness.

The transient detection problem can now be formulated as a binary hypothesis test

$$\begin{aligned} \mathcal{H}_0 &: y(n) = a_0 e^{j\omega_0 n} + v(n) \\ \mathcal{H}_1 &: y(n) = \sum_{i=1}^{M} a_i e^{-\gamma_i n} e^{j\omega_i n} + v(n). \end{aligned} \tag{4}$$

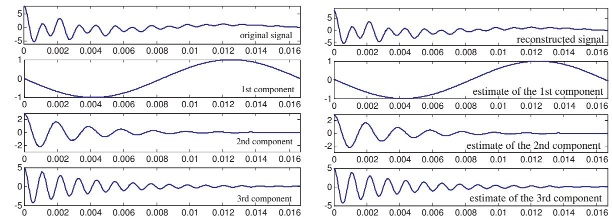
Suppose that the total number of samples is N. Defining $\boldsymbol{y} = [y(0), y(1), \cdots, y(N-1)]^T$ and $\boldsymbol{v} = [v(0), v(1), \cdots, v(N-1)]^T$, hence the binary hypothesis test can be compactly expressed in matrix format

$$\begin{aligned} \mathcal{H}_0 &: \boldsymbol{y} = a_0 \boldsymbol{e}_0 + \boldsymbol{v} \\ \mathcal{H}_1 &: \boldsymbol{y} = \boldsymbol{C}(N, M)\boldsymbol{a} + \boldsymbol{v}, \end{aligned} \tag{5}$$

where $\boldsymbol{a} = [a_1, a_2, \cdots, a_M]^T$ collects complex amplitudes, $\boldsymbol{e}_0 = [1, e^{j\omega_0}, \cdots, e^{j\omega_0(N-1)}]^T$, and

$$\boldsymbol{C}(N, M) = \begin{bmatrix} 1 & \cdots & 1 \\ e^{j\omega_1 - \gamma_1} & \cdots & e^{j\omega_M - \gamma_M} \\ \vdots & \ddots & \vdots \\ e^{(j\omega_1 - \gamma_1)(N-1)} & \cdots & e^{(j\omega_M - \gamma_M)(N-1)} \end{bmatrix} \tag{6}$$

is a full rank Vandermonde matrix [10], where N and M denote the total numbers of rows and columns, respectively.



(a) The original signal and its three components (b) The reconstructed signal and its three estimated components

Fig. 1. Parameter estimation of a sum of damped sinusoids with ESPRIT algorithm : (a) the original signal and its three components, (b) the reconstructed signal and the estimates of its three components. The first component is responsible for the fundamental frequency, which is set as 60 Hz for this example. The SNR between the first component and the noise is 30 dB.

III. General Likelihood Ratio Test (GLRT)

In this paper, $\boldsymbol{a}$, ω_i, γ_i, and M are unknown. Let $\boldsymbol{v}$ be a zero-mean complex Gaussian random vector with distribution

$$\boldsymbol{v} \sim \mathcal{CN}(\boldsymbol{0}, \delta^2 \boldsymbol{I}_N), \tag{7}$$

where δ is responsible for the standard deviation. The detection problem can be solved via GLRT

$$\frac{\max\limits_{\boldsymbol{\theta}_0} \; f(\boldsymbol{y}|\boldsymbol{\theta}_0, \mathcal{H}_0)}{\max\limits_{\boldsymbol{\theta}_1} \; f(\boldsymbol{y}|\boldsymbol{\theta}_1, \mathcal{H}_1)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} T, \tag{8}$$

where T is the threshold, $f(\cdot|\cdot)$ denotes the likelihood function, $\boldsymbol{\theta}_0 = \{a_0, \omega_0\}$ is the parameter set for the null hypothesis, and $\boldsymbol{\theta}_1 = \{\boldsymbol{a}, \omega_i, \gamma_i, M, \text{ where } 1 \leq i \leq M\}$ denotes that for the alternative one.

A. Estimation of $\boldsymbol{\theta}_0$ for $\mathcal{H}_0$

As the likelihood function of $\boldsymbol{y}$ under the null hypothesis is

$$f(\boldsymbol{y}|\boldsymbol{\theta}_0, \mathcal{H}_0) = \frac{1}{\pi^N \sigma^{2N}} e^{-\frac{(\boldsymbol{y} - a_0\boldsymbol{e}_0)^H(\boldsymbol{y} - a_0\boldsymbol{e}_0)}{\sigma^2}}, \tag{9}$$

the maximum likelihood estimate of $\boldsymbol{\theta}_0$ can be obtained via

$$\begin{aligned} \hat{\boldsymbol{\theta}}_0 &= \arg\max_{\boldsymbol{\theta}_0} \; f(\boldsymbol{y}|\boldsymbol{\theta}_0, \mathcal{H}_0) \\ &= \arg\max_{\boldsymbol{\theta}_0} \; \ln f(\boldsymbol{y}|\boldsymbol{\theta}_0, \mathcal{H}_0) \\ &= \arg\min_{\boldsymbol{\theta}_0} \; (\boldsymbol{y} - a_0\boldsymbol{e}_0)^H(\boldsymbol{y} - a_0\boldsymbol{e}_0), \end{aligned} \tag{10}$$

Apparently, (10) is a nonlinear quadratic least-square problem, and gradient based approaches can be used to obtain a feasible estimate.

B. Estimation of $\boldsymbol{\theta}_1$ for $\mathcal{H}_1$

As the number of damped sinusoids is unknown, model order estimation must be performed before the MLE. Various techniques are available in literature [11], and the maximum description length (MDL) based one is adopted in this paper. Define the partial data segment vector as

$$\boldsymbol{r}(n, K) = [y(n), y(n+1), \cdots, y(n+K-1)]^T, \tag{11}$$

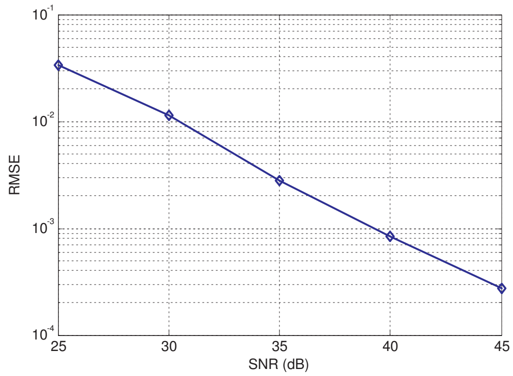


Fig. 2. The reconstruction root mean squared error (RMSE) as a function of SNR for the ESPRIT algorithm. 1000 Monte Carlo runs are used to obtain each point of the curve.

where n indicates the start index, and K represents its length, and hence the estimate of the covariance matrix is

$$\hat{\boldsymbol{R}} = \frac{1}{N-K+1} \sum_{n=0}^{N-K} \boldsymbol{r}(n,K)\boldsymbol{r}^H(n,K), \qquad (12)$$

where $(N-K+1)$ denotes the total number of segment vectors. Suppose that the eigenvalues of $\hat{\boldsymbol{R}}$ are

$$\zeta_1 \geq \zeta_2 \geq \cdots \geq \zeta_K, \qquad (13)$$

and then the MDL criterion is formulated as [12]

$$\mathrm{MDL}(p) = -\ln \left(\frac{\prod_{i=p+1}^{K} \zeta_i^{\frac{1}{K-p}}}{\frac{1}{K-p}\sum_{i=p+1}^{K} \zeta_i} \right)^{(K-p)(N-K+1)} + \frac{1}{2}p(2K-p)\ln(N-K+1), \qquad (14)$$

where $0 \leq p \leq K-1$. If K is properly selected, namely $K > M$, $\mathrm{MDL}(p)$ is firstly decreasing and then increasing. The optimal estimate of M can be uniquely obtained via

$$\hat{M} = \left(\arg\min_p \ \mathrm{MDL}(p) \right) - 1. \qquad (15)$$

Practically, an initial guess of K is desirable as one does not know the exact value of M. If the obtained $\mathrm{MDL}(p)$ is a monotonically decreasing function of p, an increasing adjustment for K is necessary.

With $\hat{M}$, the likelihood function under the alternative hypothesis becomes

$$f(\boldsymbol{y}|\boldsymbol{\theta}_1, \mathcal{H}_1) = \frac{1}{\pi^N \sigma^{2N}} e^{-\frac{(\boldsymbol{y}-\boldsymbol{C}(N,\hat{M})\boldsymbol{a})^H(\boldsymbol{y}-\boldsymbol{C}(N,\hat{M})\boldsymbol{a})}{\sigma^2}}, \qquad (16)$$

and hence the MLE of $\boldsymbol{\theta}_1$ can be obtained via

$$\begin{aligned}\hat{\boldsymbol{\theta}}_1 &= \arg\max_{\boldsymbol{\theta}_1} \ f(\boldsymbol{y}|\boldsymbol{\theta}_1, \mathcal{H}_1) \\ &= \arg\min_{\boldsymbol{\theta}_1} \ (\boldsymbol{y} - \boldsymbol{C}(N,\hat{M})\boldsymbol{a})^H(\boldsymbol{y} - \boldsymbol{C}(N,\hat{M})\boldsymbol{a}).\end{aligned}$$

Obviously, the optimization problem here is also a nonlinear least-quare.

C. Parameter Estimation with ESPRIT

The MLE involves the search of nonlinear parameter space, and its computational complexity is high. Here, we suggest to use the ESPRIT algorithm to replace the MLE in the GLRT. It is well know that the ESPRIT algorithm is suitable for the stationary rather than nonstationary signal processing [9]. We show that ESPRIT still works for parameter estimation of a sum of damped sinusoids, even though the signal itself is nonstationary. Due to a page limitation, we can't share our theoretical proof here.

The realization of ESPRIT algorithm is quite standard [3] [5] [9]. A brief description of its implementation follows:

ESPRIT Algorithm

1) Obtain the estimate of covariance matrix $\hat{\boldsymbol{R}}$ via (12).
2) Obtain the eigenvalues, say ζ_i, and their corresponding eigenvectors, say $\boldsymbol{s}_i$, of $\hat{\boldsymbol{R}}$, where ζ_i's are in a decreasing order.
3) Define $\boldsymbol{S} = [\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_{\hat{M}}]$, and then calculate the its up-submatrix $\boldsymbol{S}_u = [\boldsymbol{I}_{(K-1)} \ \boldsymbol{0}]\boldsymbol{S}$ and its down-submatrix $\boldsymbol{S}_d = [\boldsymbol{0} \ \boldsymbol{I}_{(K-1)}]\boldsymbol{S}$.
4) Obtain the eigenvalues of matrix $(\boldsymbol{S}_u^H\boldsymbol{S}_u)^{-1}\boldsymbol{S}_u^H\boldsymbol{S}_d$, say ξ_i's. Then, the estimates of the normalized frequency and the damping factor both corresponding to ξ_i are
$$\hat{\omega}_i = \mathrm{angle}(\xi_i) \text{ and } \hat{\gamma}_i = -\ln(|\xi_i|), \qquad (17)$$
respectively.
5) Employ $\hat{\omega}_i$'s and $\hat{\gamma}_i$'s to construct the estimate of $\hat{\boldsymbol{C}}(N,M)$, and hence the estimate of $\boldsymbol{a}$ can be calculated via least-square
$$\hat{\boldsymbol{a}} = \left(\hat{\boldsymbol{C}}^H(N,M)\hat{\boldsymbol{C}}(N,M)\right)^{-1}\hat{\boldsymbol{C}}^H(N,M)\boldsymbol{y}. \qquad (18)$$

With estimates of frequencies, complex amplitudes, and damping factors, the original signal can be reconstructed from its contaminated samples. To demonstrate the effectiveness of the ESPRIT algorithm, an example will be given to infer the individual component of signal mixture

$$y(t) = e^{j2\pi\times 60t} + 3e^{j2\pi\times 510t - 300t} + 5e^{j2\pi\times 912t - 250t} + v(t), \qquad (19)$$

where the unit of $y(t)$ can be either volts or amperes. The sampling interval is $\Delta t = 1/5000$, and the signal-to-noise ratio (SNR) corresponding the fundamental component $e^{j2\pi\times 60t}$ is 30 dB. The original and reconstructed signals together with their components are depicted in Fig. 1, where one cycle of the fundamental frequency is utilized, while the reconstruction error is illustrated in Fig. 2. From those figures, we see that original and estimated signals are very similar. The ESPRIT algorithm does work well.

D. Performance Analysis

The probabilities of false alarm and detection are two fundamental factors to measure the performance of a detector. Unfortunately, both the null and the alternative hypotheses contain unknown parameters that cannot be analytically reachable. Therefore, an exact formula of the probability of false alarm or of detection is unavailable. As an alternative, the receiver operation characteristic (ROC) [13] becomes a tool to numerically insight the detection performance.

IV. Numerical Results

A. Single Damped Sinusoid

This subsection examines the detection performance of the GLRT in discriminating a single sinusoid and a single damped sinusoid. The binary hypothesis are specified as

$$\begin{aligned} \mathcal{H}_0 &: y(t) = e^{j2\pi\times 50t} + v(t) \\ \mathcal{H}_1 &: y(t) = a_1 e^{j2\pi\times 50t - \gamma_1 t} + v(t). \end{aligned} \tag{20}$$

To guarantee a fair comparison, the energy of $e^{j2\pi\times 50t}$ and $a_1 e^{j2\pi\times 50t-\gamma_1 t}$ is equally preset:

$$\int_0^T 1 dt = \int_0^T |a_1|^2 e^{-2\gamma_1 t} dt, \tag{21}$$

which yields

$$a_1 = \sqrt{\frac{2\gamma_1 T}{1 - e^{-2\gamma_1 T}}} \tag{22}$$

after ignoring the initial phase, where $[0, T]$ denotes the time span of sampled signals. In computation, the sampling interval Δt is set as $1/500$ second, and T is chosen as 0.02 second (one cycle). The detection performance is shown with the help of ROC curves: Fig. 3a illustrates various ROC curves with different damping factors, while Fig. 3b depicts various ROC curves with different SNR values. It is not surprise to see that the detection performance is improved with the increase of γ_1, as the Kullback-Leiber divergence between their probability density functions (pdfs) will be enlarged if γ_1 increases. Similarly, enhancing the SNR will improve the detection performance.

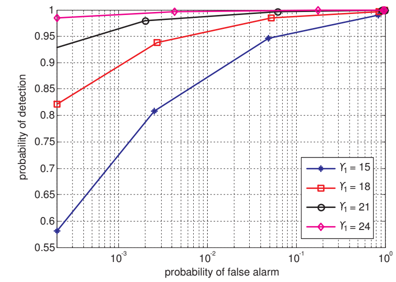
It is worth to mention that the damping factor for a real power system is in general much larger than 15, and the SNR can be over 20 dB [2]. Therefore, this approach would become even better in handling real situations.

B. Multiple Damped Sinusoids

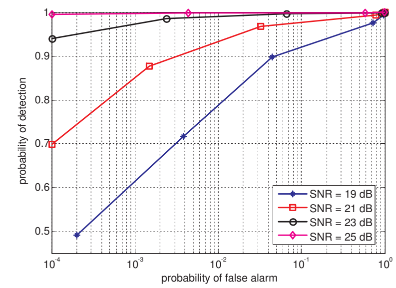
This subsection demonstrates the performance of the GLRT in detecting multiple damped sinusoids. The binary hypothesis are specified as

$$\begin{aligned} \mathcal{H}_0 &: y(t) = e^{j2\pi\times 50t} + v(t) \\ \mathcal{H}_1 &: y(t) = e^{j2\pi\times 50t} + 0.2 e^{j2\pi\times 213t - 300t} \\ &\qquad + 0.1 e^{j2\pi\times 479t - 210t} + v(t). \end{aligned} \tag{23}$$

The sampling interval Δt is set as $1/500$ second, and sampling span is chosen as 0.01 second (half cycle). The ROC curves with different SNR values are shown in Fig. 4, where the detection performance is obviously improved with the increase of SNR. Clearly, the ratio between the power of the two damped sinusoids and that of the fundamental frequency is small for the alternative hypothesis; this demonstrates that the GLRT works well for weak transient detection.



(a) ROC curves with different damping factors



(b) ROC curves with different SNR values

Fig. 3. ROC curves with different (a) damping factors and (b) different average SNR values. The SNR for (a) is set as 20 dB, while the damping factor for (b) is chosen as 2. Each curve is statistically obtained via 10000 Monte Carlo simulations.

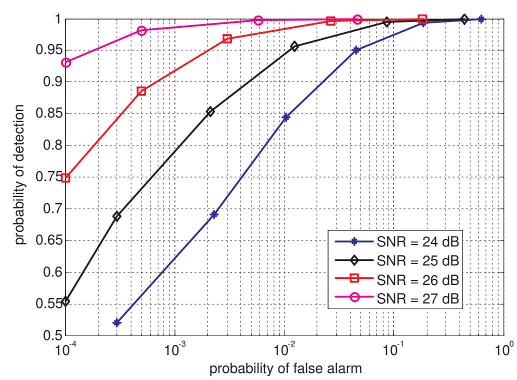


Fig. 4. The ROC curves in detecting multiples damped sinusoids, where half cycle of the fundamental frequency is employed.

V. Conclusions

The power grid is not a stationary network, and frequently internal and external circuit changes will result in voltage/current transients. Those transients will degrade the power quality and can cause potential damage to dedicate electrical devices. A reliable system should properly monitor and detect them. In this paper, the transient detection problem is investigated with binary hypothesis test, where the null hypothesis is modeled as a sinusoid, while the alternative one is treated as a sum of damped sinusoids. As the model parameters are assumed unknown, a GLRT realization is used. Some implement issues including computational complexity have been covered in this paper, and numerical results reveal its good performance.

REFERENCES

[1] L. van der Sluis, *Transients in Power Systems*. John Wiley & Sons, Inc, 2001.
[2] M. H. J. Bollen and I. Y. H. Gu, *Signal Processing of Power Quality Disturbances*. John Wiley & Sons, Inc, 2006.
[3] M. H. J. Bollen, E. Styvaktakis, and I. Y. H. Gu, "Categorization and analysis of power system transients," *IEEE Trans. Power Del.*, vol. 20, no. 3, pp. 2298–2306, Jul. 2005.
[4] N. Ruiz-Reyes, P. Vera-Candeas, and F. Jurado, "Discrimination between transient voltage stability and voltage sag using damped sinusoids-based transient modeling," *IEEE Trans. Power Del.*, vol. 20, no. 4, pp. 2644–2650, Oct. 2005.
[5] W. K. A. Najy, H. H. Zeineldin, A. H. K. Alaboudy, and W. L. Woon, "A Bayesian passive islanding detection method for inverter-based distributed generation using ESPRIT," *IEEE Trans. Power Del.*, vol. 26, no. 4, pp. 2687–2696, Oct. 2011.
[6] R. Boyer and K. Abed-Meraim, "Damped and delayed sinusoidal model for transient signals," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1720–1730, May 2005.
[7] N. Perera and A. D. Rajapakse, "Development and hardware implementation of a fault transients recognition system," *IEEE Trans. Power Del.*, vol. 27, no. 1, pp. 40–52, Jan. 2012.
[8] J. Yin, C. P. Diduch, and L. Chang, "Islanding detection using proportional power spectral density," *IEEE Trans. Power Del.*, vol. 23, no. 2, pp. 776–784, Apr. 2008.
[9] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.
[10] H. Lutkepohl, *Handbook of Matrices*. New York: Wiley, 1996.
[11] P. Stoica and Y. Selen, "Model-order selection: a review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, Jul. 2004.
[12] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 33, no. 2, pp. 387–392, Feb. 1985.
[13] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 2000.