

Estimation of Modal Damping in Power Networks

Mark Glickman, Peter O'Shea, and Gerard Ledwich, *Senior Member, IEEE*

Abstract—This paper describes a new Fourier-based sliding window method for estimating the damping of exponentially decaying modes that occur in power networks as a result of electric disturbances. The key innovation in the new method is the use of multiple orthogonal sliding windows rather than just a single pair of sliding windows. The use of these multiple orthogonal windows allows least-squares averaging strategies to be used, enabling lower variance estimates to be obtained as a result. A statistical analysis is provided, and simulations are presented to illustrate the effectiveness of the new method. The technique is also applied to real power system data.

Index Terms—Interconnected power systems, least-squares methods, parameter estimation, power system monitoring, power system transients, singular value decomposition.

I. INTRODUCTION

POWER networks consist of interconnected generators that share power to meet the demand. Various oscillating modal components can occur within the power system due to line tripping, changes in the load, and various faults. “Negative damping” corresponds to the scenario of a growing oscillation, whereas “positive damping” corresponds to a decaying oscillation. Negatively damped oscillations might lead to power system failures. If damping factor estimation is performed regularly, and modes with dangerously low (or even negative) damping are detected early, then appropriate control measures can be applied.

There are a number of existing approaches for estimating damping from power system observation records [1]–[6]. One of these approaches is to use Fourier-based “sliding window” algorithms that estimate the damping factor via Fourier transformation of two consecutive time windows [5], [6]. This approach exploits the fact that the ratio of the Fourier amplitudes in consecutive time windows is dependant on the rate of modal decay. A second approach to damping factor estimation from power system observations is the use of the so-called “parametric methods.” One of the most well-known parametric methods within the power system sector is Prony’s method. In this method, the electric disturbance modes are modeled as the output of a linear time-invariant filter driven by complex white noise, and the modal parameters are determined by formulating and solving a set of linear prediction equations [7], [8]. Discussions on the application of Prony’s method to various power systems scenarios are also provided in [9] and [10]. One of the key advantages of Prony’s method over conventional

Manuscript received November 15, 2005; revised November 16, 2006. Paper no. TPWRS-00720-2005.

The authors are with the School of Engineering Systems, Queensland University of Technology, Brisbane, Australia (e-mail: m.glickman@student.qut.edu.au; pj.oshea@qut.edu.au; g.ledwich@qut.edu.au).

Digital Object Identifier 10.1109/TPWRS.2007.901122

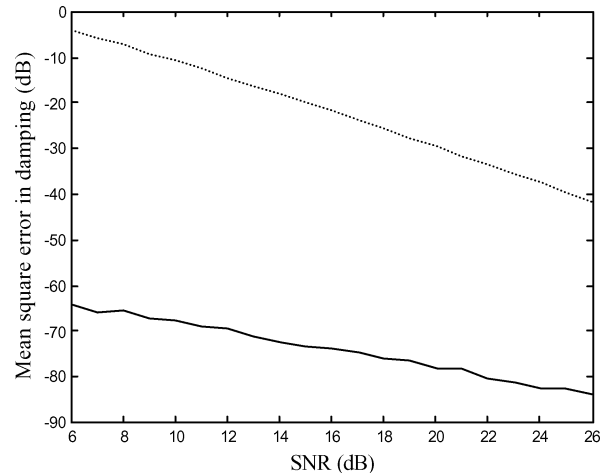


Fig. 1. Comparison of Kumaresan–Tufts damping estimate MSEs (full) and basic Prony damping estimate MSEs (dashed) for a single mode.

Fourier-based algorithms is the ability to process modes that are spectrally very close. On the other hand, Prony’s method does not perform well at low signal-to-noise ratios (SNRs).

Prony’s method by itself tends to have quite poor numerical conditioning. Kumaresan and Tufts showed that one could improve the conditioning (and therefore the accuracy of damping and frequency estimates) by using two specific modifications to the original Prony algorithm [7]. The first of these modifications was to use singular value decomposition (SVD) techniques to solve the linear prediction equations rather than standard matrix inversion procedures. The second modification was to make the linear prediction model, L , especially high. Note that there is a trade-off with the proposed modifications of Kumaresan and Tufts, because although increasing L tends to improve accuracy, it also increases the computational burden significantly. Fig. 1 illustrates the statistical advantage of using the Kumaresan–Tufts (KT) approach rather than a simple Prony analysis. The figure shows the damping estimate mean-square errors (MSEs) versus SNR for a single mode with both the KT and basic Prony methods. For the simulations, the modal signal was assumed to have the following form:

$$z_r(n) = z_s(n) + \varepsilon(n) = A \exp [j(\omega_0 n + \phi) - \alpha n] + \varepsilon(n) \quad (1)$$

with amplitude, A ; damping, α ; modal frequency, ω_0 ; initial phase, ϕ ; and $\varepsilon(n)$ being additive complex Gaussian noise. In the simulations for Fig. 1, the parameters were set to $\alpha = 0.005$, $\omega_0 = 0.0982$ rad/s, $\phi = 0$, and $A = 1$. The sampling rate was assumed without loss of generality to be 1, and 200 simulation runs were performed. The total number of samples in the data record, N , was 300, and in the KT method, the linear prediction model order, L_{KT} , was set to 180.

This paper presents a *new* sliding window method that incorporates multiple *orthogonal* sliding windows to estimate the damping of decaying oscillating modes. The orthogonality of the windows effectively enables the results from each of the multiple windows to be treated as being independent. For this reason, least-squares averaging techniques can be used to advantageously combine the results obtained from the multiple sliding windows. The new algorithm is compared with existing sliding window methods [6], and it is seen that the new multiple sliding window method has advantages over existing sliding window methods. An extension of the new method is also proposed for analyzing closely spaced modes. This extension is seen to perform comparatively well also. It should be noted that an earlier version of the new method and some preliminary results were presented in [11].

The new method is presented in Section II in a signal processing context, along with supporting simulations. Section III presents an extension of the new method to be able to process closely spaced modes, with simulations for this extension again being provided. In Section IV, the relevance of the new method to power system applications is considered. The method is applied to 1) a simulated four-machine power system scenario and 2) a real data scenario. Section V is devoted to conclusions. The Appendix contains a statistical analysis of the new method.

II. SLIDING MULTIPLE WINDOW METHOD

A. Background: The Basic Sliding Window Method

It is assumed that the oscillating mode of interest is a complex exponential of the form given in (1).

In the sliding window algorithms proposed in [5] and [6], two windows were applied to the signal at different time positions to obtain the following windowed signals:

$$y_{r1}(n) = z_r(n)w(n) = y_{s1}(n) + \varepsilon(n)w(n) \quad 0 \leq n \leq N_w - 1 \quad (2)$$

and

$$y_{r2}(n) = z_r(n + N_g)w(n) = y_{s2}(n) + \varepsilon(n + N_g)w(n) \quad 0 \leq n \leq N_w - 1 \quad (3)$$

where $w(n)$ denotes the window; $y_{s1}(n) = z_s(n)w(n)$; $y_{s2}(n) = z_s(n + N_g)w(n)$; $N_g =$ Number of samples between windows; $N_w =$ Number of samples in both windows. In this paper, it will be assumed that the two windows are consecutive in time (i.e., $N_w = N_g$). In the algorithm in [5], the window was restricted to being rectangular, whereas in [6], it was assumed to be a “smooth” Kaiser window. Note that a rectangular window starts and ends abruptly, as shown in Fig. 2(a). A “smooth” window, on the other hand, tapers smoothly on and off [see Fig. 2(c)]. The advantage of a smooth window is that its Fourier transform has negligibly low side-lobes and, therefore, has negligible spectral content outside its “main-lobe.” See, for example, Fig. 2(d), which is the Fourier transform (magnitude) of a smooth Kaiser window. By contrast, the Fourier transform

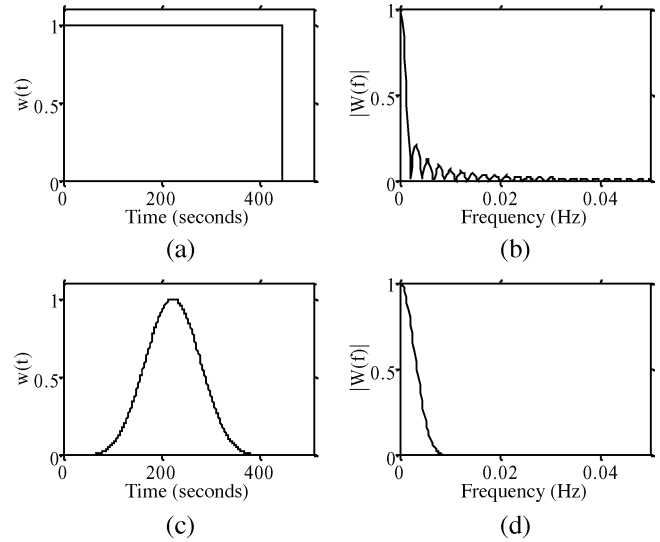


Fig. 2. (a) Rectangular window. (b) Fourier transform magnitude of a rectangular window. (c) Smooth (Kaiser) window. (d) Fourier transform magnitude of a smooth (Kaiser) window.

(magnitude) of a rectangular window is shown in Fig. 2(b). The high spectral side-lobes of the rectangular window are readily apparent. In sliding window analysis of multiple modes, these side-lobes can appear to be multiple modes. The interference effect from these side-lobes is therefore difficult to deal with [6].

Because of the interference effect described above, smooth windows are generally preferred in spectral analysis problems. Examining Fig. 2(c), though, it is apparent that smooth windows have a disadvantage. There is a tapering off (i.e., a de-emphasis of data) at the beginning and end of the window. This effectively gives rise to a “lost data” effect. In the analysis of decaying oscillations, the tapering off at the start of a smooth window is particularly problematic. This is because the start of the record is where the SNR tends to be highest. As will be seen later in this paper, this disadvantage can be mitigated with the use of *multiple* smooth windows to “recover” lost data.

For the basic sliding window estimation method, it is first necessary to estimate the frequency of the oscillating mode. This can be done with the following estimator:

$$\hat{\omega}_0 = \arg \max_{\omega} \left[\sum_{n=0}^{N-1} z_r(n) \exp(-j\omega n) \right] \quad (4)$$

Note that in practical power systems, the modal frequencies will not always be constant. There can be step changes after faults, and there can also be frequency drift [9], particularly during times where large load variations occur. The paper in [2] describes a method for estimating modal parameters that slowly change. The work in [12] outlines a method for detecting when substantial and sudden modal changes occur in the power system. If such changes do occur, though, it is probably better to wait until the system stabilizes before attempting to estimate modal parameters. That is, it is recommended that the algorithms in this paper be used only when the power system dynamics are approximately piece-wise stationary.

As explained in the introduction, damping can be estimated by taking two consecutive time windows and examining the relative strength of the Fourier transform in both windows. An estimator based on this approach is given by

$$\hat{\alpha} = \frac{1}{N_g} \log \left[\operatorname{Re} \left(\frac{Y_{r1}(\omega_0)}{Y_{r2}(\omega_0) \exp(-j\omega_0 N_g)} \right) \right] \quad (5)$$

where $Y_{r1}(\omega_0)$ and $Y_{r2}(\omega_0)$ are the Fourier transforms of the two windowed signals, $y_{r1}(n)$ and $y_{r2}(n)$ which are defined in (2) and (3), respectively. $\operatorname{Re}(\cdot)$ denotes the real part. The estimator in (5) can be further expressed as

$$\hat{\alpha} = \frac{1}{N_g} \log \left[\operatorname{Re} \left(\frac{Y_{s1}(\omega_0) + q_1}{Y_{s2}(\omega_0) \exp(-j\omega_0 N_g) + q_2} \right) \right] \quad (6)$$

where $Y_{s1}(\omega_0)$ and $Y_{s2}(\omega_0)$ are the Fourier transforms of $y_{s1}(n)$ and $y_{s2}(n)$, respectively; q_1 and q_2 are the perturbations to the spectral amplitudes, $Y_{s1}(\omega_0)$ and $Y_{s2}(\omega_0)$, respectively, due to the noise on the observation. It is assumed in this paper that q_1 and q_2 are small compared with $Y_{s1}(\omega_0)$ and $Y_{s2}(\omega_0)$, and that as a consequence, the bias of $\hat{\alpha}$ is very close to zero. Then the mean square error (MSE) and variance of $\hat{\alpha}$ are both given by

$$\operatorname{var}(\hat{\alpha}) \approx \operatorname{MSE}(\hat{\alpha}) = E [|\hat{\alpha} - \alpha|^2] \quad (7)$$

where $E[\cdot]$ denotes expected value. The variance of $\hat{\alpha}$ is derived in the Appendix. The final result for the variance is given in (41) of the Appendix.

In addition to the damping and frequency estimators defined already, one can define an estimator for the complex amplitude (i.e., amplitude and phase) of the mode. This estimate is

$$\hat{A} e^{j\hat{\phi}} = \frac{Y_{s1}(\omega_0)}{\sum_{n=0}^{N_w-1} w(n) e^{-\alpha n}} \quad (8)$$

B. Sliding Multiple Window Method

To achieve accurate damping factor estimates, it is obviously desirable to reduce the variance of the damping estimate(s). One can reduce the variance somewhat by optimizing the window length as described in [6]. It is proposed in this paper to obtain further reductions in the variance of $\hat{\alpha}$ by using *multiple orthogonal* sliding windows [13]. The following paragraphs elaborate on the concept of “orthogonal windows.”

It is possible to create an orthogonal set of K smooth windows, each with a main-lobe of bandwidth, $4\pi B$ rad/s, such that each of these K smooth windows has zero correlation with all the other windows in both the time and frequency domain. The set of windows is formed with an eigen-decomposition process, as described formally in [13]. MATLAB supports the generation of the smooth orthogonal windows via the “dps” command within the Signal Processing Toolbox. Fig. 3 shows the five orthogonal windows generated for $N_w B = 4$. It is apparent that the window in Fig. 3(a) would strongly de-emphasise the data at the start and end of a record, but the windows in Fig. 3(c)–(e) would give relatively more emphasis to the start and end of the data. Shrewd use of all the windows

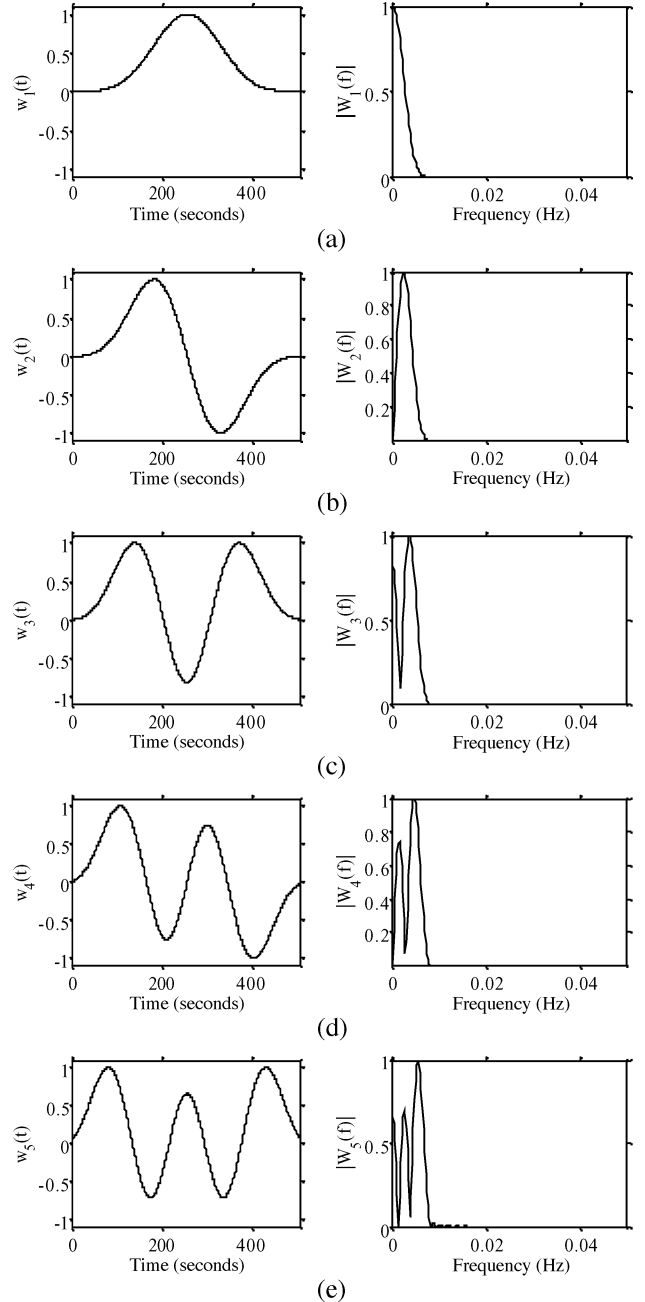


Fig. 3. Five orthogonal time windows and their Fourier transforms ($N_w B = 4$).

actually allows the data in all parts of the data record to be more effectively used. The strategy for shrewdly using all windows is outlined below.

Instead of using just one pair of windows to obtain just *one* estimate of α , one can use K pairs of *orthogonal* windows to yield K estimates of α . The K pairs of windows of the observed data are defined by

$$y_{r1}^k(n) = z_r(n) w_k(n) \quad (\text{for } 0 \leq n \leq N_w - 1, k = 1, 2, \dots, K) \quad (9)$$

$$y_{r2}^k(n) = z_r(n + N_g) w_k(n) \quad (\text{for } 0 \leq n \leq N_w - 1, k = 1, 2, \dots, K). \quad (10)$$

$w_k(n)$ is the k th orthogonal window. $\hat{\alpha}_k$ will be used to denote the k th estimate for damping.

Because of the orthogonality of the K pairs of windows, the noise on these estimates will be uncorrelated, as long as the spectrum of the additive noise is flat over the (fairly narrow) bandwidth, $4\pi B$ (a typical value for B is $4/N_W$). Because of the uncorrelated nature of the estimates, least-squares averaging can be used to yield an overall damping factor estimate with reduced variance [11]. Assume that the K pairs of orthogonal windows are used to form a vector of K damping factor estimates, $r = [\hat{\alpha}_1 \cdots \hat{\alpha}_K]^T$. From this vector, it is necessary to create a single estimate of α . The overall estimate of α that has the minimum MSE is given by [14]

$$\hat{\alpha} = (X^T C^{-1} X)^{-1} X^T C^{-1} r \quad (11)$$

where $X = [1 \ 1 \ \cdots \ 1]^T$, and C is the covariance matrix for r . C is defined by

$$C = \begin{bmatrix} E[(\alpha - \hat{\alpha}_1)(\alpha - \hat{\alpha}_1)] & \cdots & E[(\alpha - \hat{\alpha}_1)(\alpha - \hat{\alpha}_K)] \\ \vdots & \ddots & \vdots \\ E[(\alpha - \hat{\alpha}_K)(\alpha - \hat{\alpha}_1)] & \cdots & E[(\alpha - \hat{\alpha}_K)(\alpha - \hat{\alpha}_K)] \end{bmatrix}. \quad (12)$$

Now all the estimates in the r vector are obtained from windows that are orthogonal. Therefore, all estimates in the r vector are uncorrelated. Thus, all elements in the covariance matrix except those on the diagonal will be equal to zero. Thus

$$C \approx \begin{bmatrix} \text{var}(\hat{\alpha}_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \text{var}(\hat{\alpha}_K) \end{bmatrix}. \quad (13)$$

The entries for the elements on the diagonal can be obtained using the result in (41) of the Appendix. The variance of the final damping factor estimate is [14]

$$\text{var}(\hat{\alpha}) = (X^T C^{-1} X)^{-1}. \quad (14)$$

C. Summary of the Sliding Multiple Window Method

Step 1) Estimate the frequency of the oscillating mode

$$\hat{\omega}_0 = \arg \max_{\omega} \left[\left| \sum_{n=0}^{N-1} z_r(n) \exp(-j\omega n) \right| \right]. \quad (15)$$

Step 2) Form K pairs of windows of the observed data according to (9) and (10).

Step 3) Form K estimates of damping factor according to

$$\hat{\alpha}_k = \frac{1}{N_g} \log \left[\text{Re} \left(\frac{Y_{r_1}^k(\hat{\omega}_0)}{Y_{r_2}^k(\hat{\omega}_0) \exp(-j\hat{\omega}_0 N_g)} \right) \right] \quad (\text{for } k = 1, 2, \dots, K) \quad (16)$$

where $Y_{r_1}^k(\hat{\omega}_0)$ and $Y_{r_2}^k(\hat{\omega}_0)$ are the Fourier transforms of $y_{r_1}^k(n)$ and $y_{r_2}^k(n)$, respectively.

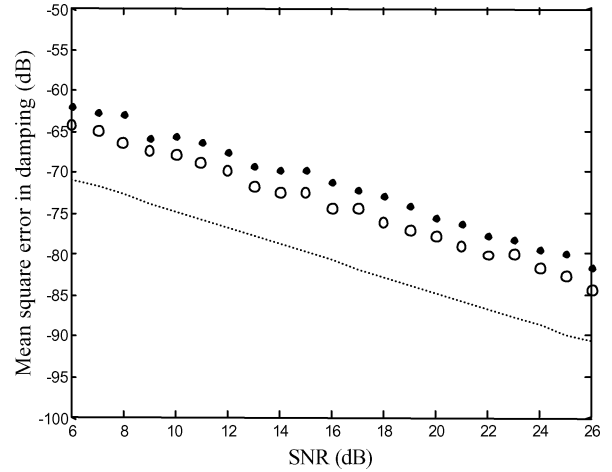


Fig. 4. Comparison of multiple sliding window method MSEs (circles), basic sliding window method MSEs (dots), and CR lower bound (dashed).

Step 4) Use least-squares averaging to estimate the overall damping factor estimate

$$\hat{\alpha} = (X^T C^{-1} X)^{-1} X^T C^{-1} r \quad (17)$$

where:

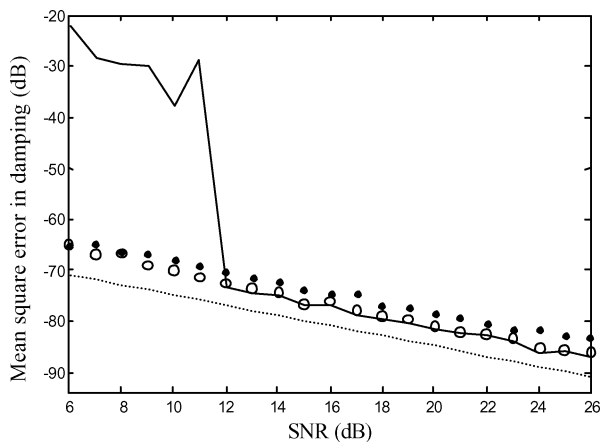
$$C = \begin{bmatrix} \text{var}(\hat{\alpha}_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \text{var}(\hat{\alpha}_K) \end{bmatrix} \quad (18)$$

and $\text{var}(\hat{\alpha}_k)$ is as given in (41) of the Appendix.

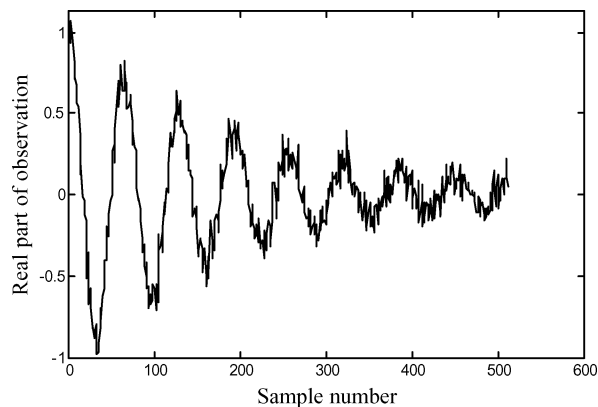
Note that the least-squares averaging procedure in (17) and (18) requires knowledge of α and A . These quantities, however, are not known *a priori*. There are a number of ways to get around this difficulty. One way is to first compute an estimate of α just using the first window. It is possible for this estimate to then be used to compute an updated estimate of α via (17) and (18). An alternative that can be used when there is less certainty about the characteristics of the additive noise is to simply weight each of the $\hat{\alpha}_k$ by the mean value of the k th window and then average all $\hat{\alpha}_k$ for $k = 1, \dots, K$. Complex amplitude estimates can be found via (8).

D. Simulations to Compare Multiple Sliding Window Method With Basic Sliding Window Method

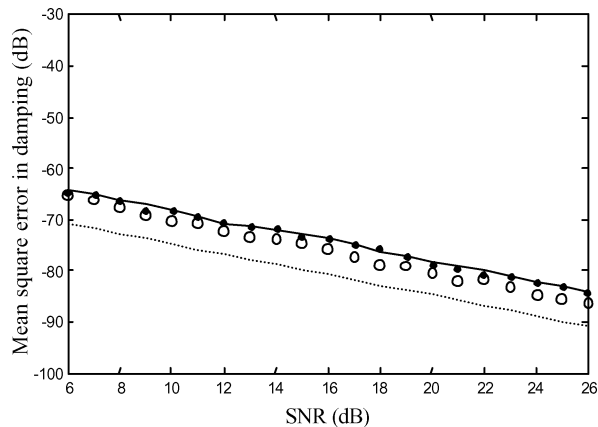
1) *Additive White Noise:* A signal of the form of (1) was synthesized, with additive white Gaussian noise of various different noise powers being used. Simulations were then performed so as to compare the multiple sliding window method with the basic sliding window method. The following simulation parameters were used: number of samples, $N = 512$; amplitude, $A = 1$; phase, $\phi = 0$ rad; damping, $\alpha = 0.005$; frequency, $\omega_0 = 0.0982$ rad/s; number of windows in the multiple sliding window method, $K = 5$; window bandwidth, $N_W B = 4$ Hz; number of samples between windows, $N_g = 64$ samples; and number of samples in both windows, $N_w = 64$ samples. In the basic sliding window method, the window used was the first of the orthogonal windows (i.e., the first-order discrete Slepian function [13]). For Figs. 4 and 5, the multiple sliding



(a)



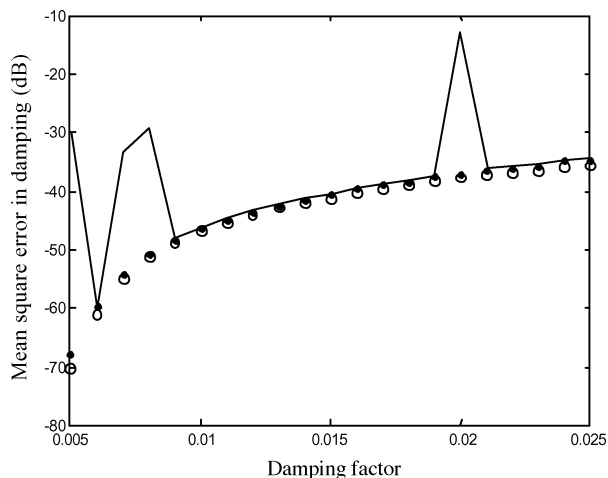
(b)



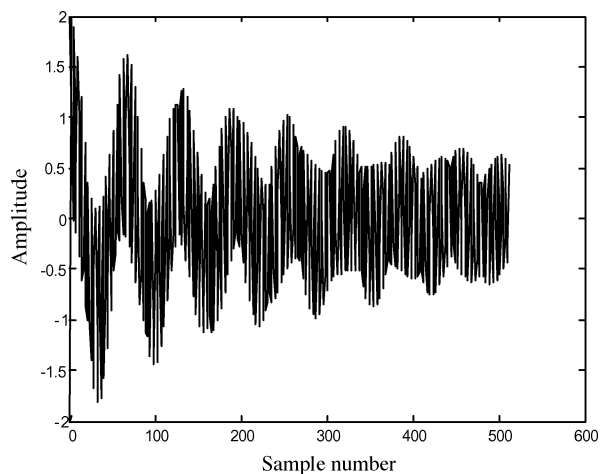
(c)

Fig. 5. (a) MSEs for KT (full), high-resolution multiple sliding window (circles), high-resolution basic sliding window (dots) methods, and the CR bound (dashed). Number of samples in KT method = 500. (b) Noisy time domain signal used to create MSEs in Fig. 5(a). (c) MSEs for KT, high-resolution multiple sliding window, and high-resolution basic sliding window methods. Number of samples for KT method = 300.

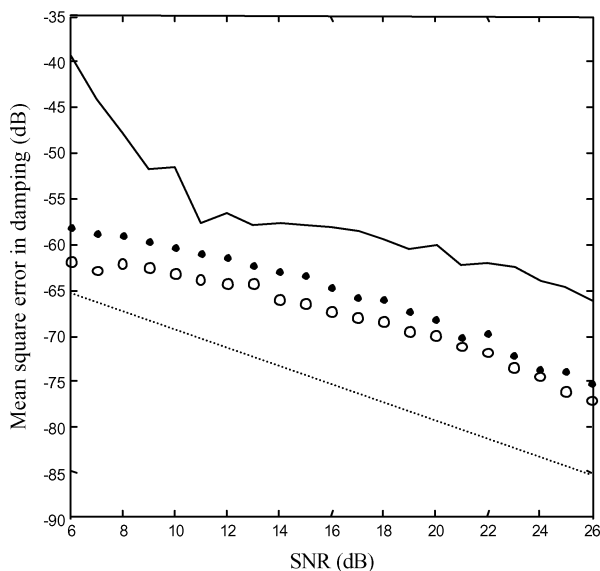
window-based damping estimate MSEs are depicted with circles, the basic sliding window damping estimate MSEs are depicted with dots, and the theoretical minimum damping estimate MSEs [i.e., the Cramer–Rao (CR) bound] are shown as a dashed line. The results of the simulations are shown in Fig. 4. It is seen that the multiple sliding window yields lower MSE estimates than the basic sliding window method. The difference in MSEs is about 2.5 dB.



(d)



(e)



(f)

Fig. 5 (cont'd). (d) Comparison of KT (full), high-resolution multiple sliding window (circles), and high-resolution basic sliding window methods (dots) for white noise. SNR held constant and damping factor varied. Number of samples in KT method = 500. (e) Noiseless time domain signal used for simulations for Fig. 5(f). (f) MSEs for KT, high-resolution multiple sliding window, and high-resolution basic sliding window methods. Number of samples in KT method = 300.

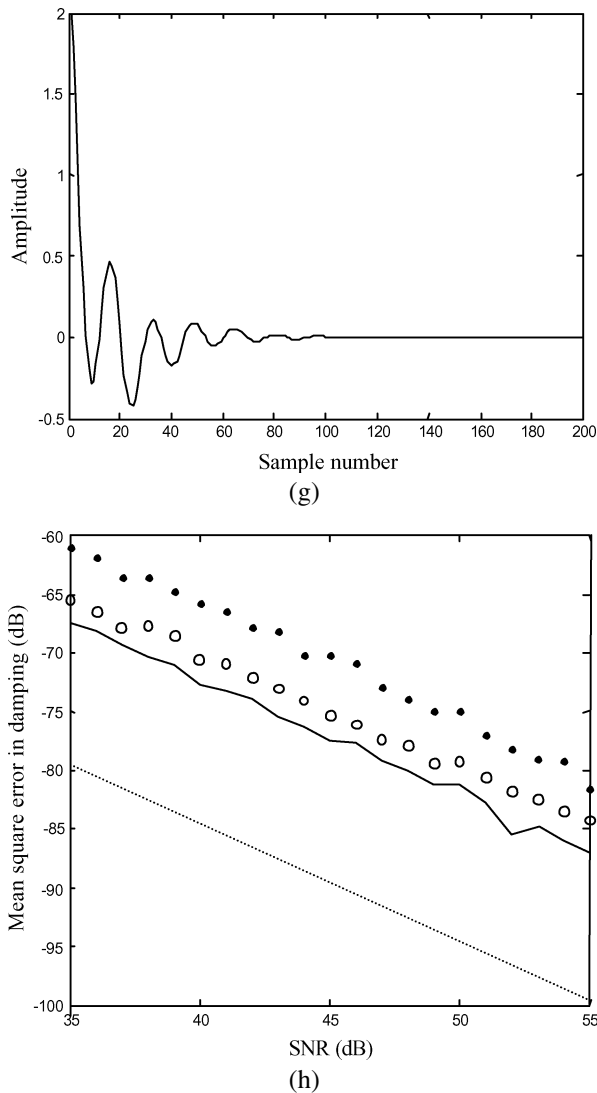


Fig. 5 (cont'd). (g) Noiseless time domain signal used for simulations. (h) Comparison of KT (full), high-resolution multiple sliding window (circles), and high-resolution basic sliding window methods (dots) for white noise. Number of samples for KT method, $N_{KT} = 60$.

2) *Additive Colored Noise*: Simulations were repeated for additive narrowband colored noise, with the noise occupying the same band as the signal. The colored noise was generated by passing the white noise through a narrowband filter (with bandwidth equal to 1/16 of the total band) centered on the modal frequency of interest. Thus, the noise was flat over 1/16 of the total band and zero elsewhere. All other parameters were kept the same as specified in the previous paragraph. When the MSEs were plotted against the SNR of the original unfiltered white noise, the resulting plot was extremely similar to the plot in Fig. 4. That is, there was little difference in the estimate MSEs whether the noise was wideband or narrowband. This is expected, since any noise outside of the relatively narrow bandwidth of the orthogonal windows has virtually no effect on the sliding window damping estimate. This is one of the key reasons for using the smooth orthogonal windows—they have very low side-lobes and therefore suppress information outside a narrow spectral region of interest. Thus, noise or other modes that are of no interest can be suppressed as required.

III. HIGH-RESOLUTION SLIDING WINDOW METHOD

The algorithm in Section II is an improvement on the basic sliding window algorithm but is still limited in a number of ways. First, it is unable to process multiple closely spaced modes. That is, it is a “low-resolution” technique. Second, because the sliding window spectrum is only evaluated at two time positions, the resulting accuracy is limited. For these reasons, it is appropriate to try to extend the *multiple* sliding window algorithm in Section II. The extended algorithm will be referred to as the high-resolution multiple sliding window algorithm. The extension draws its inspiration from the high-resolution *basic* sliding window method in [15], which dealt with an extension of the *basic* sliding window method to be able to do high-resolution processing. There are a number of goals in doing the extension. These goals are 1) to be able to process multiple closely spaced modes, 2) to be able to achieve accurate damping factor estimates, and 3) to be able to do the damping factor estimation in a computationally efficient manner.

In order to achieve goal 1), it is proposed to evaluate the sliding window spectrum at more than two time positions. As will be seen later, the additional windowed spectra tend to give new information that makes possible the analysis of multiple closely spaced modes.

To achieve goal 2), recourse will be made to filtering techniques. Such techniques are common in power system modal analysis, since they can eliminate noise and modes that are not of interest. Because of the nature of the modal analysis application, it is important to do the filtering as shrewdly as possible. That is, because one is dealing with modes that are decaying, it is important to be able to do the filtering without distorting (or losing) too much of the information at the onset of the mode (where the SNR tends to be highest). For this reason, it is critical to use filters that have both good frequency selectivity and short impulse responses. The orthogonal window functions of Slepian were designed for precisely that purpose. That is, they arose from trying to solve the problem of optimally concentrating filter bandwidth for a given length of impulse response [16]. The use of multiple sliding orthogonal windows is inherently, therefore, a very effective filtering technique for limited duration sequences. The estimates from the K different filters are statistically combined to give further improvements in accuracy.

To achieve goal 3), it is proposed to first use the “spectrogram” command in MATLAB’s Signal Processing Toolbox. (The command is invoked K times, with each invocation corresponding to one of the K orthogonal windows.) The output from each of the K “spectrogram” commands represents a bank of band-pass filtered sliding window time series, with the various filters in the bank being centered on the “bins” (discrete frequencies) of the sliding FFTs. The use of the “spectrogram” command is very efficient because a large number of filtering operations can be implemented simultaneously, thereby enabling multiple modes at different frequencies to be processed efficiently. Further efficiencies are introduced by explicitly introducing a common technique from the field of communications—namely, down-conversion. Because the time series outputs from the “spectrogram” commands (at any given frequency) are all band-pass filtered, they are narrow-band signals. It is well known that for narrowband band-pass signals, no

information is lost by translating the signal's spectral content to "base-band" (achieved by simply multiplying the time series by $\exp(-j\omega_0 n)$). Furthermore, narrowband signals that are translated to base-band have very low bandwidth and hence need quite low sample rates (according to the Nyquist theorem [17]). That is, the sliding window Fourier transform only needs to be sampled (evaluated) at infrequent stepped intervals. No information loss occurs if the stepping interval is $< 1/(2B)$ s. Because of the relatively low sample rates, the effective data record lengths are small, and processing is computationally efficient. If the sampling rate reduction is G , then the reduction in computation as compared to the KT method will typically be of the order of G^3 . There will be additional computation in the new method due to the spectrogram computations, but this would normally be quite small compared to the KT computation. The new multiple sliding window method would be more complex than the basic sliding window method by a factor of about K .

In this section, the approach outlined in the previous paragraphs is described mathematically. Assume that there are, say, M modes, that are closely spaced in a particular spectral region, centered around ω_0 .

Step 1) Determine K sliding window Fourier transforms, $F_k(n, \omega)$

$$F_k(n, \omega) = \sum_{m=0}^{N_w-1} z_r(m+n)w(m)e^{-j\omega m} \quad (19)$$

where $n = vN_g$, $v = 0, 1, 2, \dots, V$, where $V > M + 1$, $k = 1, \dots, K$. MATLAB's "specgram" command can be used to create the $F_k(n, \omega)$, for $k = 1, \dots, K$.

Step 2) For each mode (or modal cluster) of interest, determine an initial frequency estimate, $\hat{\omega}_{0i}$, by finding the frequency at which the appropriate Fourier transform region is maximized. Let the number of modes present in the frequency region of interest be M .

Step 3) Downshift the sliding window time series to base-band

$$F_{dk}(n) = F_k(n, \hat{\omega}_{0i})e^{-j\hat{\omega}_{0i}n}, \quad k = 0, 1, \dots, K. \quad (20)$$

Step 4) For each of the K time series, $F_{dk}(n)$, use the KT algorithm to estimate the frequencies and damping factors of the M modes. Then perform a weighted average of the K damping factor estimates to obtain an overall damping factor estimate

$$\hat{\alpha}_m = \frac{\sum_{k=1}^K \alpha_{m_k} u_k}{\sum_{k=1}^K u_k}, \quad m = 0, 1, \dots, M \quad (21)$$

where $\hat{\alpha}_{m_k}$ is the estimate of the damping of the m th mode based on the k th orthogonal window, $\hat{\alpha}_m$ is the "overall" estimate of the damping of the m th mode, and u_k is the mean value of the k th window. The

overall damping estimate for each mode is obtained via the "weighted average" procedure in (21), rather than via a covariance based procedure [as in (17) and (18)] because reliable covariance information is typically difficult to determine accurately for multiple modes.

If one has data from multiple sites, one can follow the approach in [9] and [18]. That is, linear prediction equations arising from the different sites are set up and grouped together in a single matrix equation that is then solved. Singular value decomposition can be used in the solution of these equations to improve robustness.

A. Simulations

Although the above method has particular application to solving multiple mode problems, it can also be very effective for single mode analysis. A number of different modal scenarios were simulated to evaluate performance. These different scenarios are described below. In all scenarios, five orthogonal windows were used, the window bandwidth was 0.1811 rad/sample, and MSEs were computed by averaging the results from 200 trials.

1) *Scenario I: A Single Mode in White Noise. Damping Factor is Held Constant, and SNR is Varied:* The simulation parameters used were: $M = 1$, number of samples in a window and step time both equal to 56 samples; amplitude equal to 1, damping factor equal to 0.005, frequency equal to 0.0982 rad/s, phase equal to 0 rad, $V = 9$, order of linear prediction in high-resolution multiple window method, $L_{HRMW} = 5$, order of linear prediction in KT method, $L_{KT} = 350$, number of samples fed into the KT method, $N_{KT} = 500$. The results are shown in Fig. 5(a). Above about 12 dB, the new high-resolution method is seen to have similar performance to the KT method, provided that large values (e.g., $L_{KT} > 200$) are used for the prediction order in the KT method. With such large values of L_{KT} , the KT method is very computationally intensive (requiring $O(L_{KT}^3)$ operations). Below about 12 dB, the new method is more resilient than the KT method. This is not entirely surprising—Fourier-based methods have a reputation for being resilient at low SNRs. The new method is also lower in MSE than the basic window high-resolution method.

Note that the SNR threshold for the KT method can be reduced (for decaying modes) by discarding the samples toward the end of the data record. This is effective because the latter part of the data record is where the SNR is poorest. This is readily apparent in Fig. 5(b), the time domain plot of a noisy version of the signal under analysis. While the discarding of information at the end of the data record does reduce the SNR threshold, the inherent loss of information also reduces the accuracy. This is illustrated in Fig. 5(c), which shows a new set of damping estimation simulations for Scenario 1 but with only 300 samples fed into the KT algorithm [as opposed to 500 samples for Fig. 5(a)]. Comparison between Fig. 5(a) and (c) shows that the discarding of samples in the KT method has reduced the SNR threshold but has also reduced the accuracy a little above threshold.

2) *Scenario II: A Single Mode in White Noise. SNR is Held Constant, and Damping Factor is Varied:* The simulation parameters used were: SNR = 10 dB,

$N = 500$, $M = 1$, amplitude = 1, damping factor varying from 0.005 to 0.025, frequency = 0.0982 rad/s and phase = 0 rad, $V = 9$, window length = stepping time = $\text{round}(56 * \text{scale factor})$, where $\text{scale factor} = (\log(0.0821)/\sigma^2)/500$, KT method's linear prediction order = $\text{round}(350 * \text{scale factor})$, order of linear prediction in high resolution multiple sliding window method = $\text{round}(K * \text{scale factor})$. Note that $\text{round}(\cdot)$ denotes rounding to the nearest integer. Simulation results are presented in Fig. 5(d). The KT method appears to be more vulnerable to failure at this comparatively low SNR of 10 dB.

3) *Scenario III: Two Closely Spaced Modes in White Noise. SNR is Varied:* The simulation parameters used were: $M = 2$, number of samples in a window and step time both equal to 56 samples; amplitudes equal to 1, damping factors equal to 0.005 and 0.001, respectively, frequencies equal to 0.0982 and 0.1104 rad/s, respectively, phases equal to 0 rad. The noiseless signal is shown in Fig. 5(e). Other parameters were $V = 9$, order of linear prediction in multiple window high-resolution method, $L_{HRMW} = 5$, order of linear prediction in KT method, $L_{KT} = 180$, number of samples fed into the KT method, $N = 300$. The results for estimating the damping of the 0.005 damping factor mode are shown in Fig. 5(f). The high-resolution multiple window method yields the lowest MSE estimates.

4) *Scenario VI: Two Heavily Damped Modes in White Noise. SNR is Varied:* The simulation parameters used were: $M = 2$, number of samples in a window, $N_w = 32$ samples; stepping time, $N_g = 4$ samples, amplitudes equal to 1, damping factors equal to 0.05 and 0.075, respectively, frequencies equal to 0.0982 and 2.356 rad/s, respectively, phases equal to 0 rad. The noiseless signal is shown in Fig. 5(g). Other parameters were $V = 9$, order of linear prediction in multiple window high resolution method, $L_{HRMW} = 5$, order of linear prediction in KT method, $L_{KT} = 30$, number of samples fed into the KT method, $N_{KT} = 60$. The results for estimating the damping of the 0.05 damping factor mode are shown in Fig. 5(h). The KT method gives the best results for this case of heavily damped modal analysis. This is probably due to the fact that it is difficult to effectively filter heavily damped modes (which have very limited duration because of their rapid decay), and filtering is an inherent part of the sliding window algorithms.

The simulations in this section have shown that the new algorithm can be useful in a number of circumstances, although it is clearly not better in all circumstances. It is least suited to analyzing weak modes. It is recommended that the new approach be used in conjunction with existing techniques such as the KT method. One of the major benefits of the new method is that the multiple windows allow several different estimates to be obtained from the one data record, and these different estimates can be used for cross-validation purposes.

IV. APPLICATION TO SIMULATED SYSTEM AND REAL DATA

A. Simulated Power System Example

A power system model was devised to reflect to some extent the power system inter-area configuration of mainland Australia. This model is shown in Fig. 6 and involves a four machine problem “strung in line” with interconnecting impedances

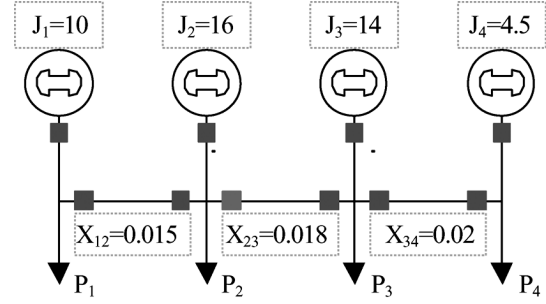


Fig. 6. Simulated power system model.

between nodes i and j denoted by X_{ij} in Fig. 6. The machines are modeled with simple classical models with inertia J_i , input power Pm_i , load power P_i , and angle δ_i . The equation relating these various quantities is

$$J_i \ddot{\delta}_i = Pm_i - P_i - \sum_j \frac{\sin(\delta_i - \delta_j)}{X_{ij}}. \quad (22)$$

Consider the case where the input powers are taken to be equal to the mean of the local load powers. Also assume that the load power variations are modeled as a random process obtained by integrating white noise. The variations are considered to be sufficiently small that the machine angle perturbations are small and $\sin(\delta) \approx \delta$. Then

$$J_i \ddot{\delta}_i \approx -\Delta P_i - \sum_j \frac{(\delta_i - \delta_j)}{X_{ij}}. \quad (23)$$

Despite the simplicity of the modeling, the specifications in Fig. 6 and (23) above give rise to a set of spectra that are close to those actually observed in the Australian system. Several algorithms were applied to the problem of estimating the damping of the inter-area modes. Ambient data (i.e., data from a simulated power system in normal operation) were used for the damping estimation. The wide band load variations in the simulated data resulted in angle difference variations of around 1 degree in the frequency band of interest, which is similar to the kind of noise levels observed in the Australian system. Approximately 3 h of simulated data were used to do the analysis, with data being “sampled” at ten samples per second. The autocorrelation function of the modal disturbance oscillations in a center of area framework from one of the machines in Fig. 6 is shown in Fig. 7. The results of analyzing this signal are shown in Table I. From the table, one can see that the algorithm using multiple windows performs best. This performance advantage is seen even more clearly in Table II, where the MSEs (in dB) are shown for each of the modes. Note that having high-quality estimates is beneficial generally in power systems from a control perspective. Control actions should ideally be based on statistical tests, and the lower the variance of damping estimates, the fewer the tests needed to ensure that any potential corrective control actions are fully justified. Additionally, when damping estimates are formed from ambient data, low variance estimators can give useful information with shorter data records than high variance estimators, thereby allowing control actions to be taken sooner.

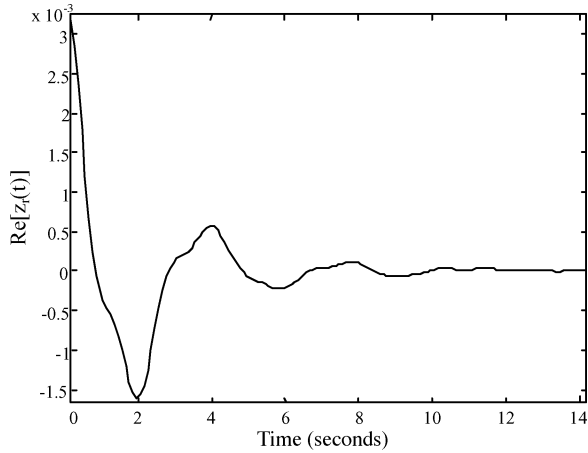


Fig. 7. Autocorrelation of data from the power system model in Fig. 6.

TABLE I
DAMPING ESTIMATES

	Damping estimates		
	Mode 1	Mode 2	Mode 3
True values	0.4768	0.1868	0.5070
Basic sliding window method	0.4672	0.1871	0.5060
High resolution sliding multiple window method	0.4768	0.1871	0.5070
Basic Prony method	0.4672	0.1871	0.5060
Kumaresan-Tufts method	0.4855	0.1838	0.4969

TABLE II
DAMPING ESTIMATES MSE

	Damping MSE (dB)		
	Mode 1	Mode 2	Mode 3
Basic high resolution sliding window method	-61.4099	-68.3039	-66.3011
High resolution sliding multiple window method	-89.4409	-69.4026	-93.4376
Basic Prony method	-40.3642	-69.3563	-59.5874
Kumaresan-Tufts method	-41.1757	-50.4140	-39.8667

B. Real Power System Example

Real power system modal oscillation data in the form of power flow on the major lines between states were acquired from Blackwall substation in near Brisbane in Queensland, Australia. The oscillation was initiated by a 300-MW braking resistor test for 0.2 s at Gladstone power station 800 km away. The measurements were done at a distance to reduce the affect of Gladstone local disturbances and oscillations. The test was able to create a measurable disturbance on the inter-area mode between Queensland and the southern states. The autocorrelation function of the oscillation data (after the brake is removed), and the Fourier transform thereof are shown in Figs. 8 and 9, respectively. Two modes are apparent in the spectrum. The larger modal component (0.3414 Hz) is the inter-area mode. Acquired data was sampled at the rate of five samples per second. In order to accommodate the assumption of unity sampling rate stated earlier in this paper, all frequencies within the analysis were expressed as “normalized frequencies”; i.e., the

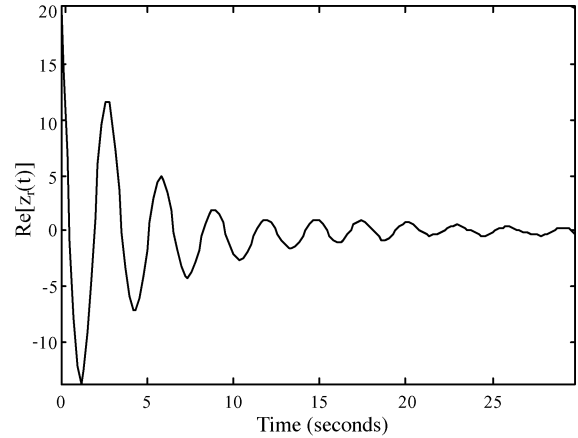


Fig. 8. Real data (time domain autocorrelation).

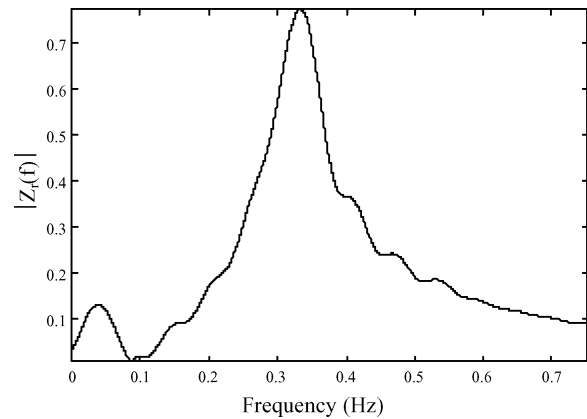


Fig. 9. Real data (frequency domain)

true frequencies scaled by 1/5. A sliding window analysis was applied with the following parameters: number of windows, $K = 5$; number of samples in each window and number of samples between each window = 48. Two consecutive time windows were used for both the basic and multiple sliding window algorithms described in Section II. The different modes, being well separated, were processed separately. The basic Prony and KT methods were also used to analyze the signal, with the first 96 samples of the data record being fed in. Frequency and damping estimates for the two modes were obtained for all methods, and then the amplitudes and phases were obtained using the standard approach [7].

With real data, one does not know the true estimates, and so one cannot determine the error in the damping (mean squared or otherwise). One can, however, measure the mean-square difference between the actual and estimated signal. This quantity is known as the “residual power,” and it is effectively a measure of the MSE of the overall estimation process. While it is not as direct a measure of the effectiveness of damping estimation, it can be a useful measure of the effectiveness of the estimation method. This residual power is shown for all of the methods in Table III. It is seen that the residual power is lowest for the sliding multiple window method. The Prony and KT methods perform poorly in this situation, possibly because they are not as resilient to model mismatches as Fourier-based methods.

TABLE III
POWER OF DIFFERENCE BETWEEN ACQUIRED AND ESTIMATED SIGNALS

Method	Residual (dB)
Sliding basic window method	-10.5630
Sliding multiple window method	-11.4828
Basic Prony method	-7.1774
Kumaresan-Tufts method	-0.2654

V. CONCLUSION

A new Fourier-based sliding window algorithm has been presented for estimating the damping of oscillating modes. The new method uses multiple orthogonal windows along with least-squares error minimization techniques. Simulations show that the algorithm can give quite useful results in various situations. It does not always outperform existing methods such as the KT method, though. The new method has also been used to analyze the output from a real braking resistor test and has been found to perform well. Because it is difficult to find a method that works well in all possible scenarios, it is recommended that the new method be used as one of a number of possible analysis tools.

APPENDIX

A. Damping Factor Estimation From Two Sliding Windows

Assume that a single modal component exists in the observation [as defined in (1)] and that two windowed signals are created from this component as per (2) and (3). If the additive noise power in the observation is zero, then the ratio of the Fourier transforms of the two windowed signals (evaluated at the oscillation frequency) is

$$\left(\frac{Y_{r1}(\omega_0)}{Y_{r2}(\omega_0)}\right) = \left(\frac{Y_{s1}(\omega_0)}{Y_{s2}(\omega_0)}\right) \quad (24)$$

$$= \left(\frac{\sum_{n=0}^{N_w-1} Ae^{(-\alpha+j\omega_0)n+j\phi}w(n)e^{-j\omega_0n}}{\sum_{n=0}^{N_w-1} Ae^{(-\alpha+j\omega_0)(n+N_g)+j\phi}w(n)e^{-j\omega_0n}}\right) \quad (25)$$

$$= e^{(\alpha-j\omega_0)N_g} \quad (26)$$

Therefore, the damping factor is defined by

$$\alpha = \frac{1}{N_g} \ln \left[\frac{Y_{s1}(\omega_0)}{Y_{s2}(\omega_0)e^{-j\omega_0N_g}} \right] \quad (27)$$

The amplitude and phase is given by

$$Ae^{j\phi} = \frac{Y_{s1}(\omega_0)}{S_W} \quad (28)$$

where

$$S_W = \sum_{n=0}^{N_w-1} w(n)e^{-\alpha n} \quad (29)$$

In practice, the noise power on the observation will be nonzero. One cannot, in general, then, determine an *exact* value for the damping factor. One can only obtain an *estimate* of the damping. Motivated by (26), the estimate for the damping factor (assuming $\hat{\omega}_0 = \omega_0$) is defined as

$$\hat{\alpha} = \frac{1}{N_g} \ln \left[\text{Re} \left\{ \frac{Y_{r1}(\omega_0)}{Y_{r2}(\omega_0)e^{-j\omega_0N_g}} \right\} \right] \quad (30)$$

$$= \frac{1}{N_g} \ln \left[\text{Re} \left\{ \frac{Y_{s1}(\omega_0) + q_1}{(Y_{s2}(\omega_0) + q_2)e^{-j\omega_0N_g}} \right\} \right] \quad (31)$$

where

$$q_1 = \sum_{n=0}^{N_w-1} \varepsilon(n)w(n)e^{-j\omega_0n} \quad (32)$$

and

$$q_2 = \sum_{n=0}^{N_w-1} \varepsilon(n + N_g)w(n)e^{-j\omega_0n} \quad (33)$$

$$\hat{\alpha} = \frac{1}{N_g} \ln \left[\text{Re} \left\{ \frac{Ae^{j\phi}S_W + q_1}{(Ae^{(-\alpha+j\omega_0)N_g+j\phi}S_W + q_2)e^{-j\omega_0N_g}} \right\} \right] \quad (34)$$

$$= \frac{1}{N_g} \ln \left[e^{\alpha N_g} \text{Re} \left\{ \frac{1 + q_1(Ae^{j\phi}S_W)^{-1}}{1 + q_2(Ae^{j\phi}S_W e^{-\alpha N_g + j\omega_0 N_g})^{-1}} \right\} \right] \quad (35)$$

Assume that q_1 and q_2 are small compared with $Y_{s1}(\omega_0)$ and $Y_{s2}(\omega_0)$. Then, using the fact that $1/(1+x) \approx 1-x$ and $\log(1+x) \approx x$ if x is small

$$\hat{\alpha} \approx \frac{\ln \left[e^{\alpha N_g} \text{Re} \left\{ \left(1 + \frac{q_1}{Ae^{j\phi}S_W}\right) \left(1 - \frac{q_2}{Ae^{j\phi}S_W e^{(-\alpha+j\omega_0)N_g}}\right) \right\} \right]}{N_g} \quad (36)$$

$$\approx \alpha + \frac{1}{N_g} \left[\text{Re} \left\{ \left(q_1 - q_2 e^{(\alpha-j\omega_0)N_g} \right) (Ae^{j\phi}S_W)^{-1} \right\} \right] \quad (37)$$

Since the windows are assumed to be consecutive, q_1 and q_2 are uncorrelated and the variance of $\hat{\alpha}$ is

$$\text{var}(\hat{\alpha}) \approx \frac{\text{var}(q_1) + \text{var}(q_2)e^{2\alpha N_g} + \text{cov}(q_1 q_2 e^{2\alpha N_g})}{2(N_g A \cdot S_W)^2} \quad (38)$$

Now if the noise is stationary over the observation interval and white, the following simplifications can be made:

$$\text{var}(q_1) = \text{var}(q_2) = \sigma^2 \left| \sum_{n=0}^{N_w-1} (w(n)^2) \right| \quad (39)$$

$$\text{var}(\hat{\alpha}) \approx \frac{\sigma^2}{2N_g^2 A^2} \left[(1 + e^{2\alpha N_g}) \left| \sum_{n=0}^{N_w-1} (w(n)^2) \right| (S_W)^{-2} \right] \quad (40)$$

$$= \frac{\sigma^2 (1 + e^{2\alpha N_g}) \left| \sum_{n=0}^{N_w-1} (w(n)^2) \right|}{2N_g^2 A^2 \left| \sum_{n=0}^{N_w-1} w(n)e^{-\alpha n} \right|^2} \quad (41)$$

REFERENCES

- [1] M. G. Anderson, N. Zhou, J. W. Pierre, and R. W. Wies, "Bootstrap-based confidence interval estimates for electromechanical modes from multiple output analysis of measured ambient data," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 943–950, May 2005.
- [2] R. W. Wies, J. W. Pierre, and D. J. Trudnowski, "Use of least-mean squares (LMS) adaptive filtering technique for estimating low-frequency electromechanical modes in power systems," in *Proc. IEEE Power Eng. Soc. General Meeting*, Jun. 2004.
- [3] J. F. Hauer, J. Hughes, and D. Trudnowski, A Dynamic Information Manager For Networked Monitoring Of Large Power Systems, EPRI Rep. TR-112031, May 1999.
- [4] J. F. Hauer and J. G. DeSteele, A Tutorial on Detection and Characterization of Special Behavior in Large Electric Power Systems, PNNL draft report, Sep. 30, 1999, revised for general distribution as PNNL Rep. PNNL-14655, Jul. 2004.
- [5] K. K.-P. Poon and K.-C. Lee, "Analysis of transient stability swings in large interconnected power systems by Fourier transformation," *IEEE Trans. Power Syst.*, vol. 3, no. 4, pp. 1573–1581, Nov. 1988.
- [6] P. O'Shea, "The use of sliding spectral windows for parameter estimation in power system disturbance monitoring," *IEEE Trans. Power Syst.*, vol. 15, no. 4, pp. 1261–1267, Nov. 2000.
- [7] S. L. Marple, *Digital Spectral Analysis: With Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [8] R. Kumaresan and D. Tufts, "Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-30, pp. 833–840, Dec. 1982.
- [9] D. J. Trudnowski, J. M. Johnson, and J. F. Hauer, "Making Prony analysis more accurate using multiple signals," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 226–231, Feb. 1999.
- [10] Bonneville Power Administration web site (last accessed Jul. 20, 2006; see, in particular, WAMS Information sub-site). [Online]. Available: <http://www.bpa.gov/corporate/>.
- [11] M. Glickman and P. O'Shea, "Damping estimation of electric disturbances in distributed power systems," in *Proc. 7th IASTED Int. Conf. Signal and Image Processing*, Aug. 2005.
- [12] R. A. Wiltshire, P. O'Shea, and G. Ledwich, "Monitoring of individual modal damping changes in multi-modal power systems," *Australian J. Elect. Electron. Eng.*, vol. 2, pp. 217–221, 2005.
- [13] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. IEEE*, vol. 70, no. 9, pp. 1055–1096, Sep. 1982.
- [14] S. Kay, "A fast and accurate single frequency estimator," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-37, pp. 1987–1990, Dec. 1989.
- [15] P. O'Shea, "A high-resolution spectral analysis algorithm for power-system disturbance monitoring," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 676–680, Aug. 2002.
- [16] D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions Fourier analysis and uncertainty-I," *Bell Syst. Tech. J.*, vol. 40, pp. 43–64, 1961.
- [17] S. S. Haykin, *Communication Systems*. New York: Wiley, 2001.
- [18] D. J. Trudnowski, J. M. Johnson, and J. F. Hauer, "SIMO system identification from measured ringdowns," in *Proc. Amer. Control Conf.*, Jun. 1998, vol. 5, pp. 2968–2972.



Mark Glickman received the B.Eng. degree in electronic engineering from Royal Melbourne Institute of Technology (RMIT), Melbourne, Australia, in December 2002. He is currently pursuing the Ph.D. degree in power system disturbance monitoring at Queensland University of Technology, Brisbane, Australia.

His research interests include power system monitoring, parameter estimation, and signal processing of acquired power signals.

Peter O'Shea received the B.E., Dip.Ed., and Ph.D. degrees from the Queensland University of Technology (QUT), Brisbane, Australia, in 1978, 1983, and 1991, respectively.

He worked as an Engineer at the Overseas Telecommunications Commission for three years, at the University of Queensland's Department of Electrical Engineering for four years, at QUT's School of Electrical and Electronic Systems Engineering for three years, and at the School of Electrical and Computer Systems Engineering, Royal Melbourne Institute of Technology (RMIT), Melbourne, Australia, for seven years. He has recently rejoined the staff of the QUT's School of Electrical and Electronic Systems Engineering. His interests are in signal processing for communications and power systems, reconfigurable computing, and the use of multimedia in engineering education.

Dr. O'Shea received the Faculty of Engineering and President's Awards for teaching at RMIT.

Gerard Ledwich (M'73–SM'92) received the Ph.D. degree in electrical engineering from the University of Newcastle, Newcastle, Australia, in 1976.

He has been Chair Professor in electrical asset management at Queensland University of Technology (QUT), Brisbane, Australia, since 1998. He was Head of electrical engineering at the University of Newcastle from 1997 to 1998. Previously, he was associated with the University of Queensland from 1976 to 1994. His interests are in the areas of power systems, power electronics, and controls.

Prof. Ledwich is a Fellow of I.E.Aust.