A New Kalman Filter-Based Algorithm for Adaptive Coherence Analysis of Non-stationary Multichannel Time Series

Z. G. Zhang and S. C. Chan
Department of Electrical & Electronic Engineering, the University of Hong Kong, Pokfulam Road, Hong Kong
{zgzhang; scchan}@eee.hku.hk

***Abstract*—This paper proposes a new Kalman filter-based algorithm for multichannel autoregressive (AR) spectrum estimation and adaptive coherence analysis with variable number of measurements. A stochastically perturbed k-order difference equation constraint model is used to describe the dynamics of the AR coefficients and the intersection of confidence intervals (ICI) rule is employed to determine the number of measurements adaptively to improve the time-frequency resolution of the AR spectrum and coherence function. Simulation results show that the proposed algorithm achieves a better time-frequency resolution than conventional algorithms for non-stationary signals.**

I. Introduction

Coherence analysis is a popular spectral analysis technique for analyzing the correlation and synchronization between different spectral components in two time series. Methods for coherence analysis can be broadly classified into non-parametric and parametric approaches according to the spectrum estimation methods used. Non-parametric coherence methods are based on non-parametric spectrum analysis techniques, such as Fourier transform [1], wavelet transform [2] and Lomb periodogram [3]. Similarly, it is possible to perform coherence analysis using parametric spectrum estimation techniques. Coherence analysis based on autoregressive (AR) model was proposed in [4] and [5]. In [5], a multichannel AR model for multivariate non-stationary time series was employed to evaluate the entire spectral density matrix and hence the coherence function. The recursive least-squares (RLS) method was used to track slowly time-varying AR coefficient matrix, from which the spectral density matrix and coherences can be computed.

In this paper, we propose a new Kalman filter-based algorithm for computing the AR coefficient matrices and coherence function with variable number of measurements. Kalman filter is a generalization of the RLS algorithm and it allows prior information of the system dynamics be incorporated into the estimating process. In the context of AR parameter estimation, the system dynamic is given by a linear state-space model, while the observations or measurements are derived from the state of the AR process. In the RLS algorithm, the AR parameters are assumed to be slowly time-varying and the estimation is based solely on the observations. On the other hand, the proposed Kalman filter-based algorithm employs a variable number of measurements and a stochastically perturbed k-order difference equation constraint model to describe the dynamics of the AR coefficients. The basic idea of using variable number of measurements is to improve the time-frequency resolution through a better tradeoff between bias and variance. Basically, a measurement window of appropriate length has to be chosen in order to reduce the variance of estimation due to additive noise, while avoiding excessive bias for non-stationary time series. The intersection of confidence intervals (ICI) rule [3, 7, and 8] is employed to determine the window size and hence the number of measurements adaptively to improve the time-frequency resolution of the AR spectrum and coherence function. Simulation results show that the proposed algorithm achieves a better time-frequency resolution than conventional algorithms for non-stationary signals.

This paper is organized as follows. Section II briefly reviews the basic of multichannel AR model. Our new Kalman filter with variable measurements algorithm is presented in Section III. Section IV is devoted to the adaptive parametric spectrum/coherence estimation from the multichannel AR model. Simulation results and comparison are given in Section V. Conclusions are drawn in Section VI.

II. Multichannel AR Process

Given an M-channel P-order multichannel AR process $y_{m,n}$ ($n=1,\cdots,N$ and $m=1,\cdots,M$), where N is the number of data samples and M is the number of channels:

$$y_n = \sum_{p=1}^{P} A_p y_{n-p} + \Phi_n , \quad (1)$$

where $y_n = [y_{1,n}, y_{2,n}, \cdots, y_{M,n}]^T$ is the measurement vector at time instant n, and A_p, $p=1,\cdots,P$, are $M\times M$ AR coefficient matrices. Φ_n is an M-dimensional zero mean Gaussian white noise with the covariance matrix Σ.

Let $\boldsymbol{\mathcal{A}}(z) = \mathbf{I}_M - \sum_{p=1}^{P} A_p z^p$, where $\mathbf{I}_M$ is the $M\times M$ identity matrix, the spectral density matrix of this multichannel AR process can be computed as [5]:

$$F(\theta) = \frac{1}{d} \boldsymbol{\mathcal{A}}^{-1}(z_\theta) \Sigma [\boldsymbol{\mathcal{A}}^{-1}(z_\theta^{-1})]^T , \quad (2)$$

where $z_\theta = \exp(-2j\pi\theta / d)$, d is the sampling rate and $\theta \in [0, d/2)$.

If the AR coefficient matrices is time-varying, i.e. $A_p = A_{p,n}$, (2) gives rise to a time-frequency representation:

$$F(n,\theta) = \frac{1}{d} \boldsymbol{\mathcal{A}}^{-1}(n, z_\theta) \Sigma [\boldsymbol{\mathcal{A}}^{-1}(n, z_\theta^{-1})]^T , \quad (3)$$

where $\boldsymbol{\mathcal{A}}(n,z) = \mathbf{I}_M - \sum_{p=1}^{P} A_{p,n} z^p$.

0-7803-9390-2/06/$20.00 ©2006 IEEE

The coherence function between any two channels (say the i-th and the j-th components of the signal vector) of the time series is defined as

$$C_{i,j}(n,\theta) = \frac{F_{i,j}(n,\theta)}{\sqrt{F_{i,i}(n,\theta)F_{j,j}(n,\theta)}}. \quad (4)$$

A key step in the parametric coherence analysis is therefore the estimation or tracking of the time-varying AR coefficient matrices $A_{p,n}$. As mentioned earlier, the RLS algorithm was used to estimate the AR coefficients in [5]. In this paper, we shall introduce a dynamic to the AR coefficients and use the Kalman filter to improve the tracking performance.

III. Kalman Filter with Variable Measurements

Consider a conventional linear state-space model as follows:

$$x(n) = \mathbf{F}(n)x(n-1) + w(n), \quad (5)$$

$$y(n) = \mathbf{H}(n)x(n) + \varepsilon(n), \quad (6)$$

where $x(n)$ and $y(n)$ are respectively the state vector and the observation vector at time instant n. $\mathbf{F}(n)$ and $\mathbf{H}(n)$ are respectively the state transition matrix and the observation matrix. The state noise vector $w(n)$ and the observation noise vector $\varepsilon(n)$ are zero mean Gaussian noise with covariance matrices $\mathbf{Q}(n)$ and $\mathbf{R}(n)$ respectively. An optimal state estimator in the least mean squares criterion for the above state-space model can be computed by the Kalman filter.

Let $\hat{x}(n/k)$ ($k = n-1, n$) represent the estimator of $x(n)$ given the measurements up to time instant k $\{y(j), j \le k\}$, and $\mathbf{P}(n/k)$ is the corresponding error covariance matrix of $\hat{x}(n/k)$. The standard Kalman filter recursions are given by:

$$\hat{x}(n+1/n) = \mathbf{F}(n)\hat{x}(n/n), \quad (7)$$

$$\mathbf{P}(n+1/n) = \mathbf{F}(n)\mathbf{P}(n/n)\mathbf{F}(n)^T + \mathbf{Q}(n), \quad (8)$$

$$e(n) = y(n) - \mathbf{H}(n)\hat{x}(n/n-1), \quad (9)$$

$$\mathbf{K}(n) = \mathbf{P}(n/n-1)\mathbf{H}(n)^T \cdot [\mathbf{H}(n)\mathbf{P}(n/n-1)\mathbf{H}(n)^T + \mathbf{R}(n)]^{-1}, \quad (10)$$

$$\hat{x}(n/n) = \hat{x}(n/n-1) + \mathbf{K}(n)e(n), \quad (11)$$

$$\mathbf{P}(n/n) = [\mathbf{I} - \mathbf{K}(n)\mathbf{H}(n)]\mathbf{P}(n/n-1), \quad (12)$$

where $e(n)$ is the prediction error of the observation vector, and its covariance matrix is $\mathbf{H}(n)\mathbf{P}(n/n-1)\mathbf{H}(n)^T + \mathbf{R}(n)$, as shown in (10). Recently, Đurović and Kovačević [9] proposed a new robust Kalman filter frame using the equivalence between the Kalman filter and a particular least-squares (LS) regression problem. Combining (5) and (6) together, we get the following equivalent linear model:

$$\begin{bmatrix} \mathbf{I} \\ \mathbf{H}(n) \end{bmatrix} x(n) = \begin{bmatrix} \mathbf{F}\hat{x}(n-1/n-1) \\ y(n) \end{bmatrix} + \mathbf{E}(n), \quad (13)$$

where $\mathbf{E}(n) = \begin{bmatrix} \mathbf{F}[x(n-1) - \hat{x}(n-1/n-1)] + w(n-1) \\ -\varepsilon(n) \end{bmatrix}$ and $E[\mathbf{E}(n)\mathbf{E}^T(n)] = \begin{bmatrix} \mathbf{P}(n/n-1) & 0 \\ 0 & \mathbf{R}(n) \end{bmatrix} = \mathbf{S}(n)\mathbf{S}^T(n)$. $\mathbf{S}(n)$ can be computed from the UD factorization or Cholesky decomposition of $E[\mathbf{E}(n)\mathbf{E}^T(n)]$. By multiplying both sides of (13) by $\mathbf{S}^{-1}(n)$, we get the following linear regression:

$$\mathbf{Y}(n) = \mathbf{X}(n)\beta(n) + \xi(n), \quad (14)$$

where $\mathbf{X}(n) = \mathbf{S}^{-1}(n)\begin{bmatrix} \mathbf{I} \\ \mathbf{H}(n) \end{bmatrix}$, $\mathbf{Y}(n) = \mathbf{S}^{-1}(n)\begin{bmatrix} \mathbf{F}\hat{x}(n-1) \\ y(n) \end{bmatrix}$, $\beta(n) = x(n)$, and $\xi(n) = -\mathbf{S}^{-1}(n)\mathbf{E}(n)$. Note that $\mathbf{E}(n)$ is whitened by $\mathbf{S}^{-1}(n)$ and the residual $\xi(n)$ satisfies $E[\xi(n)\xi^T(n)] = \mathbf{I}$. It can be seen that (14) is a standard LS regression problem with solution:

$$\hat{\beta}(n) = \hat{x}(n/n) = (\mathbf{X}^T(n)\mathbf{X}(n))^{-1}\mathbf{X}^T(n)\mathbf{Y}(n), \quad (15)$$

and the covariance matrix of estimating $\beta(n)$ is

$$E[(\beta(n) - \hat{\beta}(n))(\beta(n) - \hat{\beta}(n))^T] = \mathbf{P}(n/n) = (\mathbf{X}^T(n)\mathbf{X}(n))^{-1}. \quad (16)$$

In other words, the Kalman filter can also be thought of as the solution to a weighted LS problem with $\hat{\beta}(n) = \hat{x}(n/n)$ and $\mathbf{P}(n/n) = \mathrm{cov}(\hat{\beta}(n))$. Using (14)–(16), we obtain an equivalent Kalman filtering algorithm based on LS estimation.

To derive the proposed Kalman filter with variable measurement equations, let's rewrite (14) as

$$\left\{\mathbf{S}^{-1}(n)\begin{bmatrix} \mathbf{F}\hat{x}(n-1) \\ y(n) \end{bmatrix}\right\} = \left\{\mathbf{S}^{-1}(n)\begin{bmatrix} \mathbf{I} \\ \mathbf{H}(n) \end{bmatrix}\right\} \cdot x(n) + \xi(n). \quad (17)$$

The lower part of the equation is equivalent to a conventional LS estimation of $x(n)$ from the available measurement. The upper part is a regularization term that imposes a smoothness constraint from the state dynamic into the LS problem. If $\mathbf{F}$ is an identity matrix, (17) is equivalent to the LMS algorithm with some kind of diagonal loading. Another observation is that only one measurement is used to update the state vector. Hence, the bias error will be low especially when the system is fast time-varying. On the other hand, if the system is time-invariant or slowly time-varying, including more measurements can help to reduce the estimation variance. These observations motivate us to develop a new Kalman filter algorithm with variable number of measurements to achieve the best bias-variance tradeoff for time-series analysis.

Suppose the block of measurements used for the state estimate lie in a symmetry window centered at $y(n)$: $[y(n-L), \cdots, y(n), \cdots, y(n+L)]$, where L is the one-side window size and so $h = 2L+1$ is the total number of measurements. Including all the measurements in (14) gives:

$\mathbf{Y}(n)=\mathbf{S}^{-1}(n)\left[\{\mathbf{F}\hat{x}(n-1)\}^{T}, y(n-L),\cdots, y(n),\cdots, y(n+L)\right]^{T}$ and $\mathbf{X}(n)=\mathbf{S}^{-1}(n)\left[\mathbf{I},\mathbf{H}^{T}(n-L),\cdots,\mathbf{H}^{T}(n),\cdots,\mathbf{H}^{T}(n+L)\right]^{T}$. Note that $\mathbf{S}^{-1}(n)$ is obtained from $\begin{bmatrix}\mathbf{P}(n/n-1) & \mathbf{0}\\ \mathbf{0} & diag\{\mathbf{R}(n-L),\cdots,\mathbf{R}(n),\cdots,\mathbf{R}(n+L)\}\end{bmatrix}$ in the new algorithm. $\mathbf{E}(n)$ is whitened by $\mathbf{S}^{-1}(n)$ and the residual $\xi(n)$ satisfies $E[\xi(n)\xi^{T}(n)]=\mathbf{I}$.

The linear LS problem (14) in block-update form can be solved using (15). A method for choosing the window length h will be discussed in the next section.

IV. Adaptive Coherence Analysis

In [10, 11], a stochastically perturbed k -order difference equation constraint model is used to describe the change or dynamics of the AR coefficients. That is to say, the AR coefficients $A_{p,n}$ are modeled as k -order AR processes. For convenience, k is assumed to be 1 in this paper and we have:

$$A_{p,n}=A_{p,n-1}+\boldsymbol{\delta}_{p,n}, \tag{18}$$

where $\boldsymbol{\delta}_{p,n}$ is used to describe the change of $A_{p,n}$ and it is assumed to be an $M\times M$ zero mean Gaussian white noise matrix with covariance $\mathbf{Q}_{\delta}(p,n)=\mathbf{Q}_{\delta}(n)$, $p=1,2,\cdots,P$. We define the state matrix $\mathbf{A}(n)=[A_{1,n},A_{2,n},\cdots,A_{P,n}]^{T}$ and state noise matrix $\boldsymbol{\Delta}(n)=[\delta_{1,n},\delta_{2,n},\cdots,\delta_{P,n}]^{T}$ to obtain the following dynamic state function similar to the state equation (5) in the state-space model:

$$\mathbf{A}(n)=\mathbf{F}\mathbf{A}(n-1)+\boldsymbol{\Delta}(n), \tag{19}$$

where $\mathbf{F}=diag(\underbrace{\mathbf{I}_M,\cdots,\mathbf{I}_M}_{P})$ is the state transition matrix, and the covariance matrix of state noise $\boldsymbol{\Delta}(n)$ is $\mathbf{Q}_{\Delta}(n)=diag(\underbrace{\mathbf{Q}_{\delta}(n),\cdots,\mathbf{Q}_{\delta}(n)}_{P})$.

From (1), we let $\mathbf{Y}(n)=[y_{1,n},y_{2,n},\cdots,y_{M,n}]$ be the measurement vector and the observation matrix be $\mathbf{H}(n)=[y_{1,n-1},\cdots y_{M,n-1},\cdots,y_{1,n-p},\cdots y_{M,n-p}]$, the following observation function similar to (6) can be obtained:

$$\mathbf{Y}(n)=\mathbf{H}(n)\mathbf{A}(n)+\Psi(n), \tag{20}$$

where $\Psi(n)=\Phi_n^{\ T}$ is the observation noise vector with covariance matrix $\mathbf{R}_{\Psi}(n)$. If the noises added to every channel are all i.i.d. zero mean white Gaussian noise with variance σ^2, the covariance matrix Σ of Φ_n will be simplified to $\Sigma=diag(\underbrace{\sigma^2,\cdots,\sigma^2}_{M})$ and $\mathbf{R}_{\Psi}(n)$ becomes $M\sigma^2$. Equations (19) and (20) together constitute the state-space equations for our multichannel AR model. So, Kalman filter or our proposed Kalman filter algorithm can be used to track the state matrix $\mathbf{A}(n)$. After the AR parameters $\mathbf{A}(n)$ are estimated using the proposed Kalman filter algorithm, the instantaneous spectrum matrix and coherence can be calculated by (3) and (4), respectively.

If we use our Kalman filter algorithm to estimate the AR coefficients $A_{p,n}$ with a block of measurements $y(n-i)$, $i=-L,\cdots,L$, the corresponding spectrum and coherence will be affected by the number of measurements $h=2L+1$. If h is given a small value, fewer measurements will be used to estimate the AR coefficients so that the spectrum and coherence will have a good time resolution. In other words, for fast varying time series, a small block size is preferred. On the contrary, when a large block size is chosen, more measurements will be employed to solve the AR coefficients. As a result, the time resolution will be reduced, but the variances of the AR coefficient estimation will decrease and it gives a higher frequency resolution of the spectrum and coherence. Consequently, if the instantaneous frequency of the time series changes slowly, a larger block size should be used. When $h=1$, the Kalman filter with multi-measurements algorithm will reduce to the conventional Kalman filter.

Similar to the non-parametric Lomb spectrum and coherence [3], we will use the intersection of confidence intervals (ICI) rule to choose the number of measurements adaptively in time-frequency plane. The basic idea of the proposed method is to calculate a series of spectrums or coherences with a series of h 's first. The ICI rule will examine a sequence of confidence intervals of the spectrums or coherences to determine the optimal number of measurements h^{+} and the corresponding adaptive spectrum or coherence. The details of ICI rule and adaptive spectrum/coherence are omitted to save space, and more information can be found in [3, 7, and 8].

V. Simulation Results

The performances of the proposed algorithms are evaluated using computer simulations. A two-channel ($M=2$) non-stationary sinusoidal signal $y(t)$ with duration of 200 seconds (sampling rate $d=1$) was generated as follows:

$$y_1(n)=\begin{cases}\sin(0.4\pi n)+w_1(n) & for \quad 1\le n<100s\\ \sin(0.8\pi n)+w_1(n) & for \quad 100\le n\le 200s\end{cases},$$

$$y_2(n)=\begin{cases}\sin(0.4\pi n)+w_2(n) & for \quad 1\le n<100s\\ \sin(0.8\pi n+\pi/3)+w_2(n) & for \quad 100\le n\le 200s\end{cases},$$

where $w_1(t)$ and $w_2(t)$ are two i.i.d. zero mean white Gaussian noise added to the two channels and the SNR is 20 dB. We can see that these two components are correlated as follows: (1) in phase for the first half of time; (2) differ by a phase difference of $\pi/3$ for the second half of time.

The order of the multichannel AR process can be chosen using the AIC rule, and here P is assumed to be 2 for simplicity. The number of measurements h 's used to compute a series of spectrum matrices and coherences are 3, 5, 9, and 17. The state noise covariance matrix $\mathbf{Q}_{\Delta}(n)$ and the observation noise covariance matrix $\mathbf{R}_{\Psi}(n)$ can be estimated using the algorithm in [9], and the estimate of Σ in (3) follows the method in [5].

To achieve a better visualization of the coherence, a masking operation on the coherence can be applied:

$$C_{i,j}(n,\theta)=\begin{cases}\dfrac{P_{i,j}(n,\theta)}{\sqrt{P_{i,i}(n,\theta)P_{j,j}(n,\theta)}} & P_{i,i}(n,\theta)>\lambda, \\ & \text{and } P_{j,j}(n,\theta)>\lambda \\ 0 & \text{else}\end{cases}, \qquad (21)$$

where λ is a threshold to zero out those locations of $P_{i,i}$ or $P_{j,j}$ having very small amplitudes. Here, the threshold λ is given as 10% of the maximum value of $P_{i,i}$ and $P_{j,j}$.

Fig. 1 shows $P_{1,1}$ in the entire spectrum matrix. We can see that the spectrum component based on our Kalman filter with fewer measurements has a high time resolution but a rather unsatisfactory frequency resolution, while the spectrum with more measurements has the opposite property. The adaptive spectrum with ICI rule using adaptive number of measurements achieves good time and frequency resolution at the same time. The similar behaviors are also observed in the coherence function (Fig. 2). The time-varying coherent structure of the two signal components is clearly identified from the adaptive coherence in the two segments: first half at 0.2Hz, second half at 0.4Hz, and the jump discontinuity is located at $n=100$.

VI. Conclusion

A new Kalman filter-based algorithm for multichannel AR spectrum estimation and adaptive spectrum/coherence analysis with variable number of measurements has been presented. A stochastically perturbed difference equation constraint model is used to describe the dynamics of the AR coefficients and the ICI rule is employed to determine the number of measurements adaptively to improve the time-frequency resolution of the AR spectrum and coherence. The method can be applied to wide range of multichannel non-stationary signals such as EEG data, and the details will be left for future work.

References

[1] R. Challis and R. Kitney, "Biomedical signal processing (in four parts). III. The power spectrum and coherence function," *Med. Biol. Eng. Comput.*, vol. 29, pp. 225 - 241, 1991.

[2] J. P. Lachaux, A. Lutz, D. Rudrauf, D. Cosmelli, M. Le Van Quyen, J. Martinerie, and F. Varela, "Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence," *Neurophysiol Clin.*, vol. 32, issue 3, pp. 157 - 174, 2002.

[3] S. C. Chan and Z. G. Zhang, "Adaptive coherence analysis of nonstationary time series using the adaptive Lomb coherence," *in Proc. IEEE SSP workshop*, July 2005.

[4] L. H. Koopmans, *The Spectral Analysis of Time Series*, New York: Acadamic Press, 1994.

[5] E. Möller, B. Schack, M. Arnold, and H. Witte, "Instantaneous multivariate EEG coherence analysis by means of adaptive high-dimensional autoregressive models," *Journal of Neuroscience Methods*, vol. 105, no. 2, pp. 143 - 158, 2001.

[6] P. Wahlberg and G. Lantz, "Approximate time-variable coherence analysis of multichannel signals," *Multidimensional Systems and Signal Processing*, vol. 13, no. 3, pp. 237 - 264, 2002.

[7] L. Stanković, and V. Katkovnik, "The Wigner distribution of noisy signals with adaptive time-frequency varying window," *IEEE Trans. Signal Processing*, vol. 47, no. 4, pp. 1099 - 1108, 1999.

[8] Z. G. Zhang, and S. C. Chan, "Robust adaptive Lomb periodogram for time-frequency analysis of signals with sinusoidal and transient components," *in Proc. IEEE ICASSP 2005*, vol. 4, pp. 493 - 496, March 2005.

[9] Ž. M. Durović and B. D. Kovačević, "Robust estimation with unknown noise statistics," *IEEE Trans. Automat. Contr.*, vol. 44, no. 6, pp. 1292 - 1296, 1999.

[10] G. Kitagawa, and W. Gersch, "A smoothness priors time-varying AR coefficient modeling of nonstationary covariance time series," *IEEE Trans. Automat. Contr.*, vol. 30, no. 1, pp. 48 - 56, Jan 1985.

[11] G. Kitagawa, and W. Gersch, "A smoothness priors long AR model method for spectral estimation," *IEEE Trans. Automat. Contr.*, vol. 30, no. 1, pp. 57 - 65, Jan 1985.

[12] J. Proakis, C. Rader, F. Ling, C. Nikias, M. Moonen, and I. Proudler, *Algorithms for Statistical Signal Processing*, Englewood Cliffs: Prentice Hall, 2002.

[13] D. G. Manolakis, V. K. Ingle, and S. M. Kogan, *Statistical and Adaptive Signal Processing*, New York: McGraw-Hill, 2000.

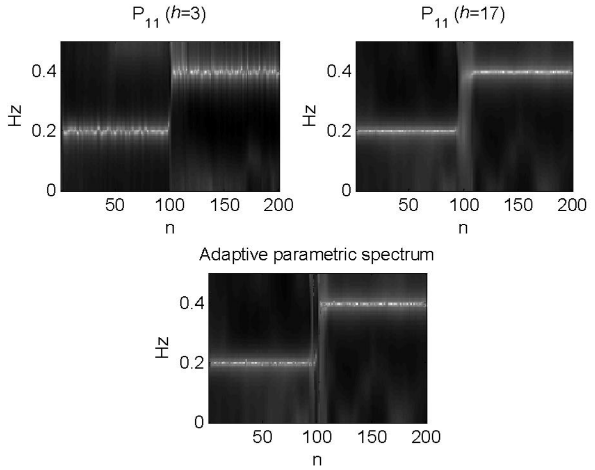

Fig. 1. Parametric spectrum $P_{1,1}$ using different measurements, and adaptive spectrum.

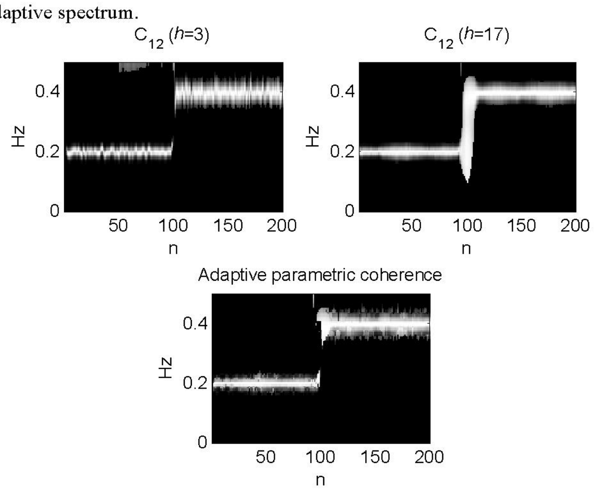

Fig. 2. Parametric coherence $C_{1,2}$ using different measurements, and adaptive coherence.