[11] H. B. Lee, "Eigenvalues and eigenvectors of covariance matrices for signals closely spaced in frequency," *IEEE Trans. Signal Processing*, vol. 40, pp. 2518–2535, 1992.
[12] D. W. Tufts and I. P. Kirsteins, "Rapidly adaptive nulling of interference," in *High-Resolution Methods in Underwater Acoustics*, M. Bouvet and G. Bienvenu, Eds. New York: Springer Verlag, 1991, ch. 6.
[13] A. Haimovich, "The eigencanceler: Adaptive radar by eigenanalysis methods," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, pp. 532–542, 1996. Also A. M. Haimovich and Y. Bar-Ness, "An eigenanalysis interference canceler," *IEEE Trans. Signal Processing*, vol. 39, pp. 76–84, 1991.
[14] H. Cox and R. Pitre, "Robust DMR and multi-rate adaptive beamforming," in *Proc. 31st Asilomar Conf. Signals, Syst., Comput.*, 1997, pp. 920–924.

Reply to "Comments on 'Theory and Application of Covariance Matrix Tapers for Robust Adaptive Beamforming'"

Joseph R. Guerci

It comes as no surprise that one of the original inventors of the covariance matrix taper (CMT) concept, and its use in adaptive pattern "robustification," continues to deepen our knowledge of CMT's in relation to alternative methods. The comments and observations in [1] go a long way toward highlighting and contrasting the fundamental differences (and similarities) between CMT's and derivative constraints.

It is important, however, to emphasize some key differences. 1) The CMT approach does not require knowledge of the interference steering vectors (thereby eliminating an entire estimation step). Although this requirement can be relaxed in the derivative constraints method for the finite sample, high JNR case [1], in practice, a method whose performance does not hinge on these conditions is clearly desirable. 2) The CMT approach provides more control over the notch widening process as it depends on a continuous notch width parameter [1], [2]. This is in contrast to the derivative constraints method that depends on the number of constraints included (discontinuous).

Finally, the CMT framework has recently given rise to an entirely new class of structured covariance estimation techniques applicable to a broad range of subspace leakage problems [3], [4].

We would also like to point out an errata for [2]. A matrix inverse sign is missing on the covariance matrices appearing in the denominators of (3), (4), and (14). In addition, in the third paragraph of Section II-D, "Theorem 2" should read "Theorem 1."

References

[1] M. Zatman, "Comments on 'Theory and application of covariance matrix tapers for robust adaptive beamforming'," *IEEE Trans. Signal Processing*, vol. 48, pp. 1796–1800, June 2000.
[2] J. R. Guerci, "Theory and application of covariance matrix tapers for robust adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 47, pp. 977–985, Apr. 1999.
[3] J. R. Guerci and J. S. Bergin, "Principal components, covariance matrix tapers, and the interference modulation problem," in *Proc. Adaptive Sensor Array Process. (ASAP) Workshop*. Lexington, MA, Mar. 10–12.
[4] ——, Principal components, covariance matrix tapers, and the subspace leakage problem, submitted for publication.

Manuscript received December 1, 1999; revised December 31, 1999. The associate editor coordinating the review of this paper and approving it for publication was Dr. Alex B. Gershman.

The author is with the Special Projects Office, Defense Advanced Research Projects Agency (DARPA), Arlington, VA 22203 USA (e-mail: jguerci@darpa.mil).

Publisher Item Identifier S 1053-587X(00)04077-0.

On Periodic Autoregressive Processes Estimation

Sophie Lambert-Lacroix

***Abstract*—We consider the autoregressive estimation for periodically correlated processes, using the parameterization given by the partial autocorrelation function. We propose an estimation of these parameters by extending the sample partial autocorrelation method to this situation. The comparison with other methods is made. Relationships with the stationary multivariate case are discussed.**

***Index Terms*—Autoregressive estimation, periodically correlated processes, sample partial autocorrelation, stationary multivariate processes.**

I. Introduction

It is well known that partial autocorrelation coefficients are the basis of most methods for autoregressive (AR) estimation. For scalar stationary processes, these coefficients are central in Burg's technique [2] and the technique of the residual energy ratio (RER) [7]. They are also used in maximum likelihood methods [11], [16] and can be directly estimated in a natural way [4]. The class of periodically correlated processes, which were introduced by Gladysev [10], is quite useful in many signal processing problems, e.g., [9] and references therein. They are not only of interest in their own right but because of their connection with multivariate covariance stationary processes. They also provide much insight into these processes and facilitate their modeling. Three methods are of major interest for estimating the second-order properties of periodic autoregressive (PAR) processes: on the one hand an extension of the Yule-Walker's method [15] and on the other hand two different extensions of Burg's technique [1], [18].

In this correspondence, we are mainly concerned with extension to the periodic situation of the method of the sample partial autocorrelation (SPAC) in [4] and that of RER [7]. The methods are compared, based on both conceptual and numerical view points. Furthermore, we consider the relationships between theses latter approaches and those associated with stationary multivariate processes.

II. Periodic Autoregressive Models

A random process $X(\cdot)$ indexed on $\mathbb{Z}$ with $EX(t) = 0,\ t \in \mathbb{Z}$, is called periodically correlated [10] if there exists a smallest $T > 0$ such that

$$R(t,s) = E(X(t)\overline{X(s)}) = \langle X(t), X(s)\rangle = R(t+T, s+T)$$

for every $(t,s) \in \mathbb{Z}^2$. We consider, in the above formula, the Hermitian product $\langle\cdot,\cdot\rangle$ defined by the expectation because it is convenient

Manuscript received October 23, 1998; revised Movember 17, 1999. The associate editor coordinating the review of this paper and approving it for publication was Dr. Jean Jacques Fuchs.

The author is with the Laboratoire LMC-IMAG, Université Joseph Fourier, Grenoble, France (e-mail: Sophie.Lambert@imag.fr).

Publisher Item Identifier S 1053-587X(00)03290-6.

here to use a geometrical approach. The process $X(\cdot)$ is said to be a periodic autoregressive process of period T and order $(p_1,\ldots,p_T)$ (PAR$(p_1,\ldots,p_T)$) if there exist some constants $a_t(k),\ k=1,\ldots,p_t$ such that

$$\sum_{k=0}^{p_t} a_t(k)X(t-k)=\varepsilon(t),\quad a_t(0)=1,\quad a_t(p_t)\neq 0 \tag{1}$$

where $\varepsilon(\cdot)$ is the innovation process, and $a_{t+T}(k)=a_t(k), k=1,\ldots,p_t$. The PAR model parameters are given by T filters $a_k(\cdot)$ and T residual variances $\sigma_k^2(\sigma_{k+T}^2=\sigma_k^2)$. Otherwise, these models can be parameterized by the first autocovariance coefficients $R(t,t-k),\ k=0,\ldots,p_t,\ t=1,\ldots,T$, using the analog of the Yule-Walker equations. In [6] (see also [12]), it is shown that these first coefficients are not always those of a PAR model. This occurs when there exists $i\in[1,\ldots T]$ such that $p_{i+1}>p_i+1$, where $p_{T+1}=p_1$. In such a case, a procedure based on using partial autocorrelation function (PACF) $\beta(\cdot,\cdot)$ allows us to check the existence of a PAR model. This motivates our choice of this parameterization. Let us recall the definition of this function. The $(t-s)$th-order forward partial innovations are denoted by $\varepsilon^f(t;s)$ with variance $\sigma^{f\,2}(t;s)$. Setting $\varepsilon^f(t;t)=X(t)$, the associated normalized innovations are defined, for $s\leq t$, by $\eta^f(t;s)=\varepsilon^f(t;s)/\sigma^f(t;s)$ with the convention $0^{-1}=0$. The backward innovations, which are obtained by reversing the time index, are indexed with b. We set $\beta(t,t)=R(t,t)$ and, for $s<t,\ \beta(t,s)$ is given by

$$\overline{\beta(s,t)}=\beta(t,s)=\langle\eta^f(t;s+1),\eta^b(s;t-1)\rangle. \tag{2}$$

This function characterizes the second-order properties of nonstationary processes [6], [12] but is easily identifiable in comparison with $R(\cdot,\cdot)$, which must be non-negative definite. Precisely, for $t\neq s$, the magnitude of $\beta(t,s)$ generally is strictly less than 1, the equality to 1 corresponding to linear relationships, namely, for $s<t,\ |\beta(t,s)|=1$ if and only if s is the largest integer such that $X(t)$ belongs to the set $\mathcal{L}\{X(s),\ldots,X(t-1)\}$ and our convention leads to $\beta(t,s-k)=\beta(t+k,s)=0$ for $k\geq 1$. Such a process is said to be *locally deterministic*.

Finally, $X(\cdot)$ is periodically correlated of period T if and only if its PACF satisfies $\beta(t+T,s+T)=\beta(t,s)$ for all $(t,s)\in\mathbb{Z}^2$. Therefore, the PAR model (1) is characterized by $\beta_t(n)=\beta(t,t-n),\ t=1,\ldots,T$, with $\beta_t(n)=0$ for $n>p_t$ [6], [12]. The one-to-one correspondence between $R(\cdot,\cdot)$ and $\beta(\cdot,\cdot)$ is realized with the periodic Levinson-Durbin (PLD) algorithm (cf. for instance [17] and [18]).

III. Sample Partial Autocorrelation Method

Let $X(1),\ldots,X(m)$ be a sequence coming from a PAR$(p_1,\ldots,p_T)$ model with order and period supposed to be known. As in the stationary case [4], the PACF method is based on a natural geometrical analysis of the sample data. In the periodic case, the sequences of length $t-s+1$, leads us to introduce the vector subspace of $C^{m_t+1},\ m_t=[(m-t)/T]$ ($[x]$ is the integer part of x) generated by the vectors

$$X_{m_t}(u)=[X(u),X(u+T),\ldots,X(u+m_tT)]^T \qquad u=s,\ldots,t.$$

Using the usual Hermitian product

$$\left\langle\vec{X}_{m_t+1}(u),\vec{X}_{m_t+1}(v)\right\rangle_e=\frac{\sum_{j=0}^{m_t}X(u+jT)\overline{X(v+jT)}}{m_t+1}$$

we obtain

$$E\left\{\left\langle\vec{X}_{m_t+1}(u),\vec{X}_{m_t+1}(v)\right\rangle_e\right\}=\langle X(u),X(v)\rangle=R(u,v).$$

In words, the sequence $\{\vec{X}_{m_t+1}(s),\ldots,\vec{X}_{m_t+1}(t)\}$ presents, in the mean, the same structure as the sequence $\{X(s),\ldots,X(t)\}$. Then, the sample partial autocorrelation coefficient, which is denoted by $\hat{\beta}_{\mathrm{spac}}(t,s)$, is the "partial correlation" between $\vec{X}_{m_t+1}(s)$ and $\vec{X}_{m_t+1}(t)$ in the set $\{\vec{X}_{m_t+1}(s),\ldots,\vec{X}_{m_t+1}(t)\}$, according to this analogy. We set $\hat{\beta}_{\mathrm{acpe}}(s,s)=\|X_{m_s}(s)\|_e^2,\ s=1,\ldots,T$, and for $0<t-s\leq p_t\leq m_t$, $\hat{\beta}_{\mathrm{spac}}(t,s)$ is given by (2), replacing $\langle\cdot,\cdot\rangle$ by $\langle\cdot,\cdot\rangle_e$ and the partial innovations by the prediction errors obtained by the least squares criterion. When the process is stationary $(T=1)$, the PACF coefficients, so estimated, correspond to those of the not symmetrized version of [4]. In the periodic situation, a Cholesky factorization is required for each coefficient $\hat{\beta}_{\mathrm{spac}}(t,s)$, namely, $\sum_{i=1}^{T}p_i$. However, when $m=NT$, an algorithm [17] more efficient than the Cholesky factorization permits us to compute the quantities $\hat{\beta}_{\mathrm{spac}}(t,s)$.

In order to extend the RER method to the periodic situation, we introduce the integers $t_i,\ i=1,\ldots,T$ in such a way that $t_i=kT+i,\ k\in\mathbb{N}$, and $0<t_i-p_i\leq T$. Then, $\hat{\beta}_{\mathrm{rer}}(t_i,t_i)=\|\vec{X}_{m_{t_i}+1}(t_i)\|_e^2$, and for $0<t_i-s\leq p_i\leq m_{t_i}$, $\hat{\beta}_{\mathrm{rer}}(t_i,s)$ is the "partial correlation" between $\vec{X}_{m_{t_i}+1}(t_i)$ and $\vec{X}_{m_{t_i}+1}(s)$ in the set $\{\vec{X}_{m_{t_i}+1}(s),\ldots,\vec{X}_{m_{t_i}+1}(t_i)\}$. Notice that this method needs at most T Cholesky factorizations since depending on the PAR model order, some coefficients can be determined from the same factorization.

IV. Comparisons with Other Approaches

The Yule-Walker method investigated in [15] keeps the PAR$(p_1,\ldots,p_T)$ model associated with the usual biased autocovariance estimates. Contrary to the stationary case, this model is not always defined unless $p_{i+1}\leq p_i+1,\ i=1,\ldots,T$ (see Section II) and the procedure in [6] (see also [12]) for checking the existence of the model allows us to obtain the estimates of the partial autocorrelation coefficients and those of the model parameters.

Burg-type generalization methods are based on the following result. For $s<t,\ \beta(t,s)$ is the β value for which

$$\|\eta^f-\beta\eta^b\|^2+\|\eta^b-\bar{\beta}\eta^f\|^2$$

is minimum, where $\eta^f=\eta^f(t;s+1)$, and $\eta^b=\eta^b(s;t-1)$. From the sample data, this criterion is applied in a recursive model order fashion. The quantities $\hat{\eta}^f$ and $\hat{\eta}^b$ are then defined from the estimates obtained at the previous stages. The difference between the two methods follows from a different choice for the residual variances estimates in the definition of the errors $\hat{\eta}^f$ and $\hat{\eta}^b$.

Precisely, these methods and the SPAC one can be recast in the following general framework. We set $\hat{\beta}(t,t)=\|\vec{X}_{m_t+1}(t)\|_e^2$ for $t=1,\ldots,T$ and, for $s\in[1,\ldots,T]$, $0<t-s\leq p_t$, the estimate $\hat{\beta}(t,s)$ is obtained by a two-stage procedure.

i) Choose the coefficients $\tilde{a}_t^f(t-s-1,\cdot)$ and $\tilde{a}_{t-1}^b(t-s-1,\cdot)$ of the filters giving the $(t-s-1)$th-order prediction errors

$$\tilde{\varepsilon}_{m_t+1}^f(t;s+1)=\sum_{j=0}^{t-s-1}\tilde{a}_t^f(t-s-1,j)\vec{X}_{m_t+1}(t-j)$$

$$\tilde{\varepsilon}_{m_t+1}^b(s;t-1)=\sum_{j=0}^{t-s-1}\tilde{a}_{t-1}^b(t-s-1,j)\vec{X}_{m_t+1}(s+j)$$

and compute the corresponding sample elements

$$\tilde{\sigma}^{f2}_{m_t+1}(t;s+1) = \left\|\tilde{\varepsilon}^{f}_{m_t+1}(t;s+1)\right\|^2_e$$

$$\tilde{\sigma}^{b2}_{m_t+1}(s;t-1) = \left\|\tilde{\varepsilon}^{b}_{m_t+1}(s;t-1)\right\|^2_e,$$

$$\tilde{\delta}_{m_t+1}(t,s) = \left\langle \tilde{\varepsilon}^{f}_{m_t+1}(t;s+1), \tilde{\varepsilon}^{b}_{m_t+1}(s;t-1)\right\rangle_e.$$

ii) Choose the estimators $\tilde{\sigma}^{f2}(t;s+1)$ and $\tilde{\sigma}^{b2}(s;t-1)$ of the residual variances.

Then, $\hat{\beta}(t,s)$ is given by

$$\frac{2\tilde{\delta}_{m_t+1}(t,s)}{\frac{\tilde{\sigma}^{b}(s;t-1)}{\tilde{\sigma}^{f}(t;s+1)}\tilde{\sigma}^{f2}_{m_t+1}(t;s+1) + \frac{\tilde{\sigma}^{f}(t;s+1)}{\tilde{\sigma}^{b}(s;t-1)}\tilde{\sigma}^{b2}_{m_t+1}(s;t-1)}.$$

The generalizations of Burg's technique take, in the first stage i), the filters associated with the preceding estimated values by the PLD algorithm equations. Boshnakov's method uses, in stage ii), the estimators $\hat{\sigma}^{f2}(t;s+1)$ and $\hat{\sigma}^{b2}(s;t-1)$ defined by these coefficients, whereas the Sakai's method selects the empirical elements introduced in i). The SPAC method chooses the filters given by the least squares criterion in i) and the empirical residual variances for ii).

For Burg-type methods, we can point out that the constraints in the recursive construction of the prediction error filters constitute a drawback with respect to SPAC or RER methods. On the other hand, when the data result from a locally deterministic process, the values of magnitude 1 are almost surely estimated by SPAC or RER methods since the filters are those given by the least squares criterion. Nevertheless, the filters that establish the relation between the components of the process $X(\cdot)$ are not generally defined by $\hat{\beta}(\cdot,\cdot)$ and must be estimated differently. The three other methods introduced in the not locally deterministic case only remain valid but without guaranteeing that the estimated structure be the one of a locally deterministic process. In the stationary situation, except for the RER method, all procedures operate in a recursive model order fashion. Surprisingly, in the periodic situation, the SPAC method is the only one that still satisfies this property in all cases. For the Burg technique generalizations, this property is not always satisfied, depending on the considered model order and the way it is increased. Indeed (see the PLD Algorithm), the $\beta(t,s)$ estimate depends, through the prediction errors filters, on all values of $\hat{\beta}(u,v)$, $v \le u$ with $(u,v) \in [s,\ldots,t]^2 \setminus (t,s)$. The same result holds for the Yule-Walker method.

V. Multivariate Approaches

Recall that starting from a periodically correlated process $X(\cdot)$ of period T, the T-multivariate process $Y_j(t) = X(j+T(t-1))$, $j = 1,\ldots,T$ is stationary and vice versa [10]. Furthermore, $X(\cdot)$ is PAR$(p_1,\ldots,p_T)$ if and only if [15] $Y(\cdot)$ is stationary autoregressive (AR) of order $p = \max_j[(p_j - j)/T] + 1$. When $m = NT$, the sequence $X(1),\ldots,X(m)$ provides $Y(1),\ldots,Y(N)$, and the multivariate autoregressive estimate methods lead to scalar ones and vice versa. The fundamental difference between both approaches is that the periodic structure corresponding to the multivariate method is the one of a PAR model with $p_i = pT + i - 1$. The multivariate methods estimate only a subclass of models among the whole stationary AR(p) processes, and we consider such models to compare both approaches. Dégerine [5] gives a general framework fitting most estimation methods that use partial autocorrelation matrices. We restrict the study to the case of matrices of "triangular" type [3], [17] that appear naturally in the correspondence with the periodically correlated processes. Therefore, the framework of [5] reduces to the analog of the one proposed for the scalar approaches. It consists of the choice of prediction errors filters and that of the residual covariance matrices estimators. According to the analogy between multivariate and scalar frameworks, the generalization of Burg's technique proposed by Nuttall [14] (see also [19]) corresponds to the Boshnakov technique [1], and that of Morf *et al.* [13] is related to the Sakai method [18]. However, these methods are not equivalent because the resulting estimated periodic structures differ. Our method is equivalent to the one defined by the multivariate SPAC one [5], [17] resulting from the same kind of choices in both stages. This also is true for the Yule-Walker method and the RER extension, which do not fit the scalar general framework.

TABLE I
Comparison Between Yule-Walker's and SPAC Methods; $m = 50$, $n_r = 2000$.

	$\beta(\cdot,\cdot)$	$\rho(\cdot,\cdot)$	$R(t,t)$	$a.(\cdot)$	$\sigma^2_\cdot$
YW bias	0.136	0.080	0.025	0.251	0.172
YW $\sqrt{\text{MSE}}$	0.211	0.171	1.517	0.388	0.261
SPAC bias	0.044	0.054	0.025	0.040	0.034
SPAC $\sqrt{\text{MSE}}$	0.161	0.169	1.517	0.261	0.066

PAR(4,4) model with

$\beta(1,\cdot)$: 3.00, 0.60, 0.85, -0.55, -0.65,

$\beta(2,\cdot)$: 2.00, -0.75, 0.60, -0.70, -0.50.

VI. Simulation Results

Without loss of generality, we consider the real case with $m = NT$. In the simulations results below, the former model is a regular PAR$(4,4)$: $|\beta(t,s)| \le 0.85$. The latter is a PAR$(6,6,6)$ model nearly locally deterministic: Some values of $\beta(\cdot,\cdot)$ are close to ± 1 in order to emphasize the differences between the methods. Comparison is made through estimation of the parameters $\beta_t(n)$, $\rho_t(n) = R(t,t-n)/\sqrt{R(t,t)R(t-n,t-n)}$, $R(t,t), a_t(n)$, and σ^2_t, $t = 1,\ldots,T$, $n = 1,\ldots,p_t$. All the characteristics of the simulations are recalled in each table, where n_r is the replicate index. The bias and the square root of the mean square error (MSE) of estimators are summarized as follows: For a parameter vector $\theta(\cdot)$ of dimension d ($d = \sum_{t=1}^{T} p_t$ for $\beta.(\cdot)$, $\rho.(\cdot)$, $a.(\cdot)$, and $d = T$ for $R(\cdot,\cdot)$, $\sigma^2_\cdot$), we give

$$\left\{\frac{1}{d}\sum_{i=1}^{d}[\text{bias}[\hat{\theta}_i(k)]]^2\right\}^{\frac{1}{2}}, \quad \left\{\frac{1}{d}\sum_{i=1}^{d}\text{MSE}[\hat{\theta}_i(k)]\right\}^{\frac{1}{2}}.$$

The SPAC method is compared with the Yule-Walker one for short data records ($m = 50$) from the PAR$(4,4)$ model. As in the stationary case, we observe the shortcoming due to the bias introduced in the Yule-Walker method, especially for short data records. Table I shows that the two methods are very close in estimating $\rho(\cdot,\cdot)$. Otherwise, the Yule-Walker method leads to very bad results, especially for the model parameters. In this case, the other methods give results similar to those of the SPAC method. On the other hand, we have noticed that this difference between both methods increases with a nearly locally deterministic model, even for long data records. Notice that if we consider the PAR$(4,0)$ model whose PACF coefficients not equal to zeroare given by that of the PAR$(4,4)$ process, the Yule-Walker method did not provide a solution in 26,7 percentage of time for data records of length 50 and for 2000 repetitions.

The SPAC method is compared with RER and the extension of Burg's technique. The closeness of the model to the locally deterministic case, the data record length, and the kind of parameters under study play an essential role in the difference between these methods. Burg's technique's generalizations behave similarly, whereas the RER method is

TABLE II
COMPARISON BETWEEN BURG'S, RER AND SPAC METHODS; $m = 105$, $n_r = 2000$.

	$\beta(\cdot,\cdot)$	$\rho(\cdot,\cdot)$	$R(t,t)$	$a.(\cdot)$	$\sigma^2_\cdot 10^2$
SA. bias	0.124	0.118	0.040	0.523	4.098
SA. $\sqrt{\text{MSE}}$	0.208	0.235	5.288	0.738	6.097
BO. biais	0.125	0.118	0.040	0.531	4.153
BO. $\sqrt{\text{MSE}}$	0.208	0.235	5.288	0.747	6.164
RER bias	0.066	0.119	0.039	0.017	0.499
RER $\sqrt{\text{MSE}}$	0.156	0.238	5.333	0.251	0.799
SPAC bias	0.065	0.118	0.040	0.017	0.475
SPAC $\sqrt{\text{MSE}}$	0.154	0.235	5.288	0.256	0.799

$PAR(6,6,6)$ model with

$\beta(1,\cdot)$: 7.60, -0.80, -0.30, 0.85, -0.95, 0.80, 0.25,

$\beta(2,\cdot)$: 2.00, 0.50, 0.90, -0.90, 0.70, -0.60, 0.45,

$\beta(3,\cdot)$: 4.50, 0.00, -0.90, 0.30,-0.80, 0.55, -0.90.

a very good approximation of the SPAC one. Table II indicates that the four methods are equivalent in $\rho(\cdot,\cdot)$ estimation. Otherwise, the Burg-type extensions give very bad results for the other parameters. This seems to be a consequence of the presence of constraints in the recursive filters construction. This clearly appears in the $\sigma^2_\cdot$ case, where the bias corresponds to overestimation of residual variances of each period. It can be expected that this failure has an increased effect when the model order is large.

VII. CONCLUSION

The SPAC and RER methods are extended to the periodically correlated processes case. They are compared with the Yule-Walker method and two extensions of Burg's method. Simulation results show that they eliminate some shortcomings of the other methods. On the other hand, we also consider the relationship between these different approaches and those related to the stationary multivariate process. The advantage of scalar approaches is to avoid use of matrices and to allow estimation of autoregressive models of any order.

REFERENCES

[1] G. Boshnakov, "Periodically correlated sequences: Some properties and recursions," Luleå Univ., Luleå, Sweden, Res. Rep. 1, Div. Quality Technol. Stat., 1994.

[2] J. P. Burg, "Maximum Entropy Spectral Analysis," Ph.D. dissertation, Dept. Geophys.. Standford Univ., Stanford, CA, 1975.

[3] S. Dégerine, "Canonical partial autocorrelation function of a multivariate time series," *Ann. Statist.*, vol. 18, no. 2, pp. 961–971, 1990.

[4] ——, "Sample partial autocorrelation function," *IEEE Trans. Signal Processing*, vol. 41, pp. 403–407, Jan. 1993.

[5] ——, "Sample partial autocorrelation function of a multivariate time series," *J. Multivariate Anal.*, vol. 50, no. 2, pp. 294–313, 1994.

[6] ——, Partial autocorrelation function of a nonstationary time series with application to periodically correlated processes, preprint, 1998.

[7] B. W. Dickinson, "Autoregressive estimation using residual energies ratios," *IEEE Trans. Inform. Theory*, vol. IT–24, pp. 503–506, Apr. 1978.

[8] ——, "Estimation of partial correlation matrices using Cholesky decomposition," *IEEE Trans. Automat. Contr.*, vol. AC–24, pp. 302–305, Apr. 1979.

[9] W. A. Gardner, Ed., *Cyclostationarity in Communications and Signal Processing*. New York: IEEE, 1994.

[10] E. G. Gladysev, "Periodically random sequences," *Sov. Math.*, vol. 2, pp. 385–388, 1961.

[11] S. M. Kay, "Recursive maximum likelihood estimation of autoregressive processes," *IEEE Trans. Signal Processing*, vol. 41, pp. 56–65, Jan. 1993.

[12] S. Lambert-Lacroix, "Fonction d'autocorrélation partielle des processus à temps discret non stationnaires et applications," Ph.D. dissertation, J. Fourier Univ., Grenoble, France, 1998.

[13] M. Morf, A. Vieira, D. Lee, and T. Kailath, "Recursive multichannel maximum entropy spectral estimation," *IEEE Trans. Geosci. Electron.*, vol. GE–16, pp. 85–94, 1978.

[14] A. H. Nuttall, "Positive definite spectral estimate and stable correlation recursion for multivariate linear predictive spectral analysis," Naval Underwater Syst. Cent., New London, CT, NUSC Tech. Doc. 5729, Nov. 14, 1976.

[15] M. Pagano, "On periodic and multiple autoregressions," *Ann. Stat.*, vol. 6, no. 6, pp. 1310–1317, 1978.

[16] D. T. Pham, "Maximum likelihood estimation of autoregressive model by relaxation on the reflection coefficients," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP–38, pp. 175–177, Jan. 1988.

[17] ——, Fast scalar algorithms for multivariate partial correlations and their empirical counterparts, Grenoble, France, Rap. Recherche, LMC (URA 397), RR 903-M, 1992.

[18] H. Sakai, "Circular lattice filtering using Pagano's method," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 279–287, Feb. 1982.

[19] O. N. Strand, "Multichannel complex maximum entropy (autoregressive) analysis," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 634–640, July 1977.

Nonuniform M-Band Wavepackets for Transient Signal Detection

Seema Kulkarni, V. M. Gadre, and Sudhindra V. Bellary

***Abstract*—In this paper, we present a scheme to detect significantly overlapping transients buried in white Gaussian noise. A nonuniform M-band wavepacket decomposition algorithm using M-band, translation-invariant wavelet transform (NMTI) is developed, and its application to transient signal detection is discussed. The robustness of the NMTI-based detector is illustrated.**

***Index Terms*—M-band, receiver operating characteristics, transient, wavepackets.**

I. INTRODUCTION

The transient signal detection problem has been studied using various techniques with different assumptions regarding *a priori* information about the signal like time of arrival, duration, time-bandwidth product and relative bandwidth, signal model, etc. [1]–[3]. In case of no prior knowledge, wavepackets and their variations have shown better performance due to their property to capture the transient signal in a

Manuscript received January 5, 1999; revised September 23, 1999. This work was supported in part by the Naval Research Board of India, CDAC Pune, and the All-India Council of Technical Education. The associate editor coordinating the review of this paper and approving it for publication was Dr. Lal C. Godara.

S. Kulkarni is with the Centre for Development of Advanced Computing (CDAC), Pune, India.

V. M. Gadre and S. V. Bellary are with the Department of Electrical Engineering, Indian Institute of Technology, Powai, Mumbai, India.

Publisher Item Identifier S 1053-587X(00)04048-4.