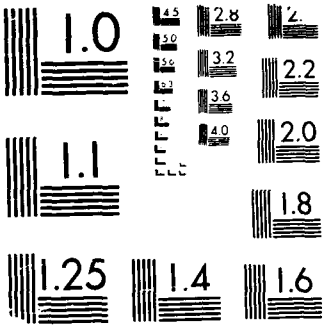
AD-A193 857 STATISTICAL ANALYSIS OF DYADIC STATIONARY PROCESSES(U) 1/1
PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS
M TANIGUCHI ET AL. SEP 87 TR-87-40 AFOSR-TR-88-0390
UNCLASSIFIED N00014-85-K-0292 F/G 12/3 NL



MICROCOPY RESOLUTION TEST CHART

DTIC FILE COPY

AFOSR-TR- 88-0398

STATISTICAL ANALYSIS OF DYADIC
STATIONARY PROCESSES*

M. Taniguchi, L.C. Zhao,
P.R. Krishnaiah and Z.D. Bai

Center for Multivariate Analysis
University of Pittsburgh

Technical Report No. 87-40

AD-A193 057

Center for Multivariate Analysis

University of Pittsburgh

DTIC
ELECTE
MAY 0 3 1988
S D
H

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

88 5 02 281

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR- 88-0398

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle): Statistical Analysis of Dyadic Stationary Processes

5. TYPE OF REPORT & PERIOD COVERED: Journal ~~Technical~~ - September 1987

6. PERFORMING ORG. REPORT NUMBER: 87-40

7. AUTHOR(s): M. Taniguchi, L.C. Zhao, P.R. Krishnaiah and Z.D. Bai

8. CONTRACT OR GRANT NUMBER(s): N00014-85-K-0292 (Navy) F49620-85-C-0008 (AIR FORCE)

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Center for Multivariate Analysis, Fifth Floor Thackeray Hall, University of Pittsburgh, Pittsburgh, PA 16260

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304 A6

11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research; Air Force Office of Scientific Research

12. REPORT DATE: September 1987

13. NUMBER OF PAGES: 30

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): AFOSR/NM BCI 410 BAFB DC 20332-6448

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. *Key words and phrases:* dyadic stationary process; information criterion; likelihood ratio criterion; quasi-maximum likelihood estimator; Walsh spectral density.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper we consider a multiple dyadic stationary process with the Walsh spectral density matrix $f_{\underset{\sim}{\theta}}(\lambda)$, where $\underset{\sim}{\theta}$ is an unknown parameter vector. We define a quasi-maximum likelihood estimator $\hat{\underset{\sim}{\theta}}$ of $\underset{\sim}{\theta}$, and give the asymptotic distribution of $\hat{\underset{\sim}{\theta}}$ under appropriate conditions. Then we

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20 Abstract (continued)

propose an information criterion which determines the order of the model, and show that this criterion gives the consistent order estimate. As for a finite order dyadic autoregressive model, we propose a simpler order determination criterion, and discuss its asymptotic properties in detail. This criterion gives strong consistent order estimate. Also detections of signals for dyadic stationary processes will be discussed. In Section 6 we discuss testing whether an unknown parameter $\underline{\theta}$ satisfies a linear restriction. Then we give the asymptotic distribution of the likelihood ratio criterion under the null hypothesis.

(2)

STATISTICAL ANALYSIS OF DYADIC STATIONARY PROCESSES*

M. Taniguchi, L.C. Zhao,
P.R. Krishnaiah and Z.D. Bai

Center for Multivariate Analysis
University of Pittsburgh

Technical Report No. 87-40

September 1987

Center for Multivariate Analysis
Fifth Floor Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

DTIC
ELECTE
MAY 0 3 1988
S D
H

* This work is supported by Contract N00014-85-K-0292 of the Office of Naval Research and Contract F49620-85-C-0008 of the Air Force Office of Scientific Research. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notations hereon.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

STATISTICAL ANALYSIS OF DYADIC
STATIONARY PROCESSES*

ABSTRACT

In this paper we consider a multiple dyadic stationary process with the Walsh spectral density matrix $f_{\underset{\sim}{\theta}}(\lambda)$, where $\underset{\sim}{\theta}$ is an unknown parameter vector. We define a quasi-maximum likelihood estimator $\hat{\underset{\sim}{\theta}}$ of $\underset{\sim}{\theta}$, and give the asymptotic distribution of $\hat{\underset{\sim}{\theta}}$ under appropriate conditions. Then we propose an information criterion which determines the order of the model, and show that this criterion gives the consistent order estimate. As for a finite order dyadic autoregressive model, we propose a simpler order determination criterion, and discuss its asymptotic properties in detail. This criterion gives strong consistent order estimate. Also detections of signals for dyadic stationary processes will be discussed. In Section 6 we discuss testing whether an unknown parameter $\underset{\sim}{\theta}$ satisfies a linear restriction. Then we give the asymptotic distribution of the likelihood ratio criterion under the null hypothesis.

AMS 1980 Subject Classifications: Primary 60K99, secondary 62M99.

Key words and phrases: dyadic stationary process; information criterion; likelihood ratio criterion; quasi-maximum likelihood estimator; Walsh spectral density.

* This work is supported by Contract N00014-85-K-0292 of the Office of Naval Research and Contract F49620-85-C-0008 of the Air Force Office of Scientific Research. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notations hereon.

1. INTRODUCTION

There has been much discussion of Walsh spectral analysis for dyadic stationary processes. Morettin ~~(1974)~~ investigated some asymptotic properties of the finite Walsh transforms of dyadic stationary processes. Nagai ~~(1977)~~ gave the spectral representations for dyadic stationary processes. If we consider finite dyadic linear models then the greatest differences between dyadic stationary processes and ordinary stationary ones appear. Nagai ~~(1980)~~ and Nagai and Taniguchi ~~(1987)~~ established that a dyadic autoregressive and moving average ~~(DARMA)~~ process of finite order can be expressed as a dyadic autoregressive ~~(DAR)~~ process of finite order and also as a dyadic moving average ~~(DMA)~~ process of finite order. Nagai and Taniguchi (1987) discussed the principal component analysis of a multiple dyadic process, and also the canonical correlation analysis. Morgettin ~~(1981)~~ gave a convenient survey for Walsh spectral analysis.

~~In~~ this paper ~~we~~ considers a multiple dyadic stationary process with ~~the~~ Walsh spectral density matrix $f_{\underset{\sim}{\theta}}(\lambda)$, where $\underset{\sim}{\theta}$ is an unknown parameter vector. We define a quasi-maximum likelihood estimator $\hat{\underset{\sim}{\theta}}$ of $\underset{\sim}{\theta}$, and give the asymptotic distribution of $\hat{\underset{\sim}{\theta}}$ under appropriate conditions. In Section 3 we propose an information criterion which determines the order of the model, and show that this criterion gives the consistent order estimate. In Section 4, for a finite order dyadic autoregressive model, we propose a simpler order determination criterion, and show that the estimated order has strong consistency. Also some interesting examples are given in the identification problem for Walsh spectrum. In Section 5 we consider a signal detection model for dyadic process of finite order, and show that this model is equivalent to a dyadic moving average model. In Section 6

INSPECTED 4

Dist | Avail and/or Special
A-1

we discuss a testing problem whether the unknown parameter $\underset{\sim}{\theta}$ satisfies a linear restriction. Then we give the asymptotic distribution of the likelihood ratio criterion under the null hypothesis.

2. DYADIC STATIONARY PROCESSES AND ESTIMATION THEORY

First we introduce some basic concepts and notations. Denote by T the set of all nonnegative integers. Let x and y be nonnegative real numbers and have the following binary expansions:

$$x = \sum_{\ell=-\infty}^{\infty} x_\ell 2^\ell \quad \text{with} \quad x_\ell = 0 \text{ or } 1,$$

$$y = \sum_{\ell=-\infty}^{\infty} y_\ell 2^\ell \quad \text{with} \quad y_\ell = 0 \text{ or } 1.$$

Then the dyadic addition $\oplus$ is defined by

$$x \oplus y = \sum_{\ell=-\infty}^{\infty} |x_\ell - y_\ell| 2^\ell.$$

A stochastic process (possibly vector process) $\{Y(t): t \in T\}$ is said to be dyadic stationary if the joint distribution of $Y(t_1), Y(t_2), \ldots, Y(t_n)$ is the same as that of $Y(t_1 \oplus t), Y(t_2 \oplus t), \ldots, Y(t_n \oplus t)$ for every finite set of integers $\{t_1, \ldots, t_n\}$ and for every integer t. We denote by $\{W(t,\lambda): 0 \leq \lambda \leq 1\}$, $t = 0,1,\ldots$ the system of Walsh functions. The properties of Walsh functions are well known:

(i) for each $t \in T$ and $\lambda \in [0,1]$, the value of $W(t,\lambda)$ is only +1 or -1,

(ii) for any $s, t \in T$,

$$W(t,\lambda)W(s,\lambda) = W(t \oplus s, \lambda), \quad \text{a.e. } \lambda,$$

(iii) for each $t \in T$ and $\lambda \in [0,1]$,

$$W(t,\lambda)W(t,\mu) = W(t, \lambda \oplus \mu), \quad \text{a.e. } \mu.$$

(See Morettin (1974).)

Let $Y(t) = \left(Y_1(t), \ldots, Y_q(t)\right)'$; $t \in T$ be a q-dimensional dyadic stationary process with zero mean and k-th order cumulants

$$c_{a_1 \ldots a_k}(t_1, \ldots, t_{k-1}) = \text{cum}\{Y_{a_1}(t_1 \oplus t_k), Y_{a_2}(t_2 \oplus t_k), \ldots, Y_{a_k}(t_k)\},$$

$t_1, \ldots, t_{k-1} \in T$. We denote the covariance matrices

$$\Gamma(t_1) = \{c_{a_1 a_2}(t_1)\}, \ q\times q \text{ matrices.}$$

The statistic

$$d^{(N)}(\lambda) = \sum_{t=0}^{N-1} Y(t)W(t,\lambda) \tag{2.1}$$

is called the finite Walsh transform of $\{Y(t): t = 0,1,\ldots,N-1\}$. Throughout this paper we assume that $N = 2^m$, with m a nonnegative integer and denote $d^{(N)}(\lambda) = \left(d_1^{(N)}(\lambda),\ldots,d_q^{(N)}(\lambda)\right)'$. Here we assume the following.

ASSUMPTION 1. For every k and $j = 1,2,\ldots,k-1$,

$$\sum_{t_1=0}^{\infty} \cdots \sum_{t_{k-1}=0}^{\infty} |c_{a_1\ldots a_k}(t_1, \ldots, t_{k-1})||t_j| < \infty, \tag{2.2}$$

for all $a_1, \ldots, a_k$.

Then the Walsh spectral density matrix and the Walsh cumulant spectrum of order k of $\{Y(t)\}$ are defined by

$$f(\lambda) = \sum_{t=0}^{\infty} \Gamma(t)W(t,\lambda),$$

and

$$f_{a_1\ldots a_k}(\lambda_1, \ldots, \lambda_{k-1})$$

$$= \sum_{t_1} \cdots \sum_{t_{k-1}} c_{a_1\ldots a_k}(t_1, \ldots, t_{k-1}) \prod_{j=1}^{k-1} W(t_j,\lambda_j), \tag{2.3}$$

respectively. The following proposition is due to Morettin (1974).

PROPOSITION 1. Under Assumption 1,

$$\mathrm{cum}\{d_{a_1}^{(N)}(\lambda_1), \ldots, d_{a_k}^{(N)}(\lambda_k)\}$$

$$= D_N(\lambda_1 \oplus \ldots \oplus \lambda_k)\{f_{a_1 \ldots a_k}(\lambda_1, \ldots, \lambda_{k-1}) + O(N^{-1})\}, \tag{2.4}$$

where $D_N(\lambda) = \sum_{t=0}^{N-1} W(t,\lambda)$, and the term $O(N^{-1})$ is uniform with respect to $\lambda_1, \ldots, \lambda_k$. □

Although we do not assume the Gaussianity of $\{Y(t)\}$, we can make the Gaussian likelihood function L of $\{Y(0), \ldots, Y(N-1)\}$, formally, and approximate L. That is, we get

$$\log L \cong -\frac{N}{2}\int_0^1 \{\log\det f_{\underset{\sim}{\theta}}(\lambda) + \mathrm{tr}\, I_N(\lambda) f_{\underset{\sim}{\theta}}(\lambda)^{-1}\} d\lambda + \text{constant}. \tag{2.5}$$

where the fitted Walsh spectral density matrix of $\{Y(t)\}$ is parameterized as $f_{\underset{\sim}{\theta}}(\lambda)$, $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_r)' \in \Theta \subset \mathbb{R}^r$, and

$$I_N(\lambda) = F_N(\lambda) F_N(\lambda)' = \{I_{ab}(\lambda)\}, \quad \text{say},$$

$$F_N(\lambda) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} Y(t) W(t,\lambda).$$

Thus we estimate $\underset{\sim}{\theta}$ by the value $\hat{\underset{\sim}{\theta}}$ which minimize

$$D(f_{\underset{\sim}{\theta}}, I_N) = \int_0^1 \{\log\det f_{\underset{\sim}{\theta}}(\lambda) + \mathrm{tr}\, I_N(\lambda) f_{\underset{\sim}{\theta}}(\lambda)^{-1}\} d\lambda, \tag{2.6}$$

with respect to $\underset{\sim}{\theta}$. Henceforth we call $\hat{\underset{\sim}{\theta}}$ quasi-maximum likelihood estimator of $\underset{\sim}{\theta}$. To discuss the asymptotic properties of $\hat{\underset{\sim}{\theta}}$, the following lemma is a keystone.

LEMMA 1. Let $\phi_j(\lambda) = \{\phi_{ab}^{(j)}(\lambda)\}$, $j = 1,\ldots,r$, be $q\times q$ matrix-valued continuous functions on $[0,1]$ such that $\phi_j(\lambda) = \phi_j(\lambda)'$. Under Assumption 1 we can show that

(1) $$\lim_{N\to\infty} \int_0^1 \operatorname{tr} I_N(\lambda)\phi_j(\lambda)d\lambda = \int_0^1 \operatorname{tr} f(\lambda)\phi_j(\lambda)d\lambda, \quad \text{in } p, \tag{2.7}$$

(2) the quantities

$$A_j = \sqrt{N} \int_0^1 \operatorname{tr}\{\big(I_N(\lambda) - f(\lambda)\big)\phi_j(\lambda)\}d\lambda, \quad j = 1,\ldots,r,$$

have, asymptotically, a normal distribution with zero mean vector and covariance matrix V whose (j,m) element is

$$2\int_0^1 \operatorname{tr}\{f(\lambda)\phi_m(\lambda)f(\lambda)\phi_j(\lambda)\}d\lambda$$
$$+ \sum_{a,b,c,d=1}^{q} \iint_0^1 \phi_{ba}^{(j)}(\lambda)\phi_{dc}^{(m)}(\mu)f_{abcd}(\lambda,\lambda,\mu)d\lambda d\mu. \tag{2.8}$$

Proof. Notice that

$$A_j = \sqrt{N} \sum_{a,b=1}^{q} \int_0^1 \{I_{ab}(\lambda) - f_{ab}(\lambda)\}\phi_{ba}^{(j)}(\lambda)d\lambda.$$

By Proposition 1, we have

$$E(I_{ab}(\lambda) - f_{ab}(\lambda)) = O(\tfrac{1}{N}),$$

where $O(\frac{1}{N})$ is uniform with respect to λ. Hence we obtain

$$E(A_j) = O(N^{-1/2}). \tag{2.9}$$

Since

$$\begin{aligned}\operatorname{cum}\{I_{ab}(\lambda),\ I_{cd}(\mu)\} &= \frac{1}{N^2}\{\operatorname{cum}\big(d_a(\lambda),\ d_d(\mu)\big)\operatorname{cum}\big(d_b(\lambda),\ d_c(\mu)\big) \\ &\qquad + \operatorname{cum}\big(d_a(\lambda),\ d_c(\mu)\big)\operatorname{cum}\big(d_b(\lambda),\ d_d(\mu)\big) \\ &\qquad + \operatorname{cum}\big(d_a(\lambda),\ d_b(\lambda),\ d_c(\mu),\ d_d(\mu)\big)\} \\ &= \frac{1}{N^2}\{D_N^2(\lambda \oplus \mu)[f_{ad}(\lambda)f_{bc}(\lambda) + f_{ac}(\lambda)f_{bd}(\lambda)] \\ &\qquad + D_N(\lambda \oplus \lambda \oplus \mu \oplus \mu)f_{abcd}(\lambda,\lambda,\mu) + O(\tfrac{1}{N})\}\end{aligned}$$

and

$$D_N(\lambda \oplus \lambda \oplus \mu \oplus \mu) = D_N(0) = N,$$

we have

$$\begin{aligned}\operatorname{cum}(A_j,A_m) &= N\int_0^1\!\!\int_0^1 \sum_{a,b,c,d=1}^{q} \phi_{ba}^{(j)}(\lambda)\phi_{dc}^{(m)}(\mu)\operatorname{cum}\big(I_{ab}(\lambda),\ I_{cd}(\mu)\big)\,d\lambda_1 d\lambda_2 \\ &= \sum_{a,b,c,d=1}^{q}\Bigg[\int_0^1\!\!\int_0^1 \phi_{ba}^{(j)}(\lambda)\phi_{dc}^{(m)}(\mu)f_{abcd}(\lambda,\lambda,\mu)\,d\lambda\, d\mu \\ &\qquad + 2\int_0^1 \phi_{ba}^{(j)}(\lambda)f_{ac}(\lambda)f_{bd}(\lambda)\,d\lambda\Big[\frac{1}{N}\int_0^1 \phi_{dc}^{(m)}(\mu)D_N^2(\lambda\oplus\mu)\,d\mu\Big]\Bigg] + O(\tfrac{1}{N}).\end{aligned} \tag{2.10}$$

Note

$$D_N(\lambda \oplus \lambda) = \begin{cases} N & \text{if} \quad |\lambda \oplus \mu| < \frac{1}{N}, \\ 0 & \text{otherwise,}\end{cases}$$

we get that by the continuity of $\phi_{dc}^{(m)}(\mu)$,

$$\frac{1}{N}\int_0^1 \phi_{dc}^{(m)}(\mu)D_N^2(\lambda \oplus \mu)\,d\mu = \begin{cases} \phi_{dc}^{(m)}(\lambda) + o(1), & \text{if } \frac{1}{2N} < \lambda < 1 - \frac{1}{2N}, \\ O(1), & \text{otherwise.}\end{cases}$$

Substituting the above into (2.10), we get

$$\mathrm{cum}(A_j, A_m) = 2\int_0^1 \mathrm{tr}\{f(\lambda)\phi_m(\lambda)f(\lambda)\phi_j(\lambda)\}d\lambda$$

$$+ \sum_{a,b,c,d=1}^{q} \iint_0^1 \phi_{ba}^{(j)}(\lambda)\phi_{dc}^{(m)}(\mu) f_{abcd}(\lambda,\lambda,\mu)d\lambda d\mu + o(1). \qquad (2.11)$$

Thus (2.9) and (2.11) imply our result (1). Also (2.11) gives the asymptotic variance (2.8). As for the asymptotic normality of A_j, we have only to evaluate $J(J \geq 3)$th order cumulant $\mathrm{cum}\{A_{i_1}, A_{i_2}, \ldots, A_{i_J}\}$. Here, without loss of generality, we evaluate it for scalar process.

By Theorem 2.3.2, p.21, Brillinger (1975), we have

$$\mathrm{cum}\left(d_{11}(\lambda_1)d_{12}(\lambda_1), \ldots, d_{J1}(\lambda_J)d_{J2}(\lambda_J)\right)$$

$$= \sum_{\nu} \mathrm{cum}\left(d_{ji}(\lambda_j), (j,i) \in \nu_1\right) \ldots \mathrm{cum}\left(d_{ji}(\lambda_j), (j,i) \in \nu_S\right) \qquad (2.12)$$

where the summation runs over all indecomposable partitions $\nu = \nu_1 \cup \ldots \cup \nu_s$ of the set $\{(j,i), j = 1,2,\ldots,J, i = 1,2\}$, (the definition of indecomposability can be found on p.20, Brillinger (1975)). By indecomposability of the partitions, each ν_n contains at least two elements, we have

$$S \leq J/2.$$

By Proposition 1, we have

$$\mathrm{cum}\left(d_{ji}(\lambda_j), (j,i) \in \nu_1\right) \ldots \mathrm{cum}\left(d_{ji}(\lambda_j), (j,i) \in \nu_S\right) = O\left(\prod_{n=1}^{S} D_N\left(\bigoplus_{(j,i)\in\nu_n} \lambda_j\right)\right).$$

Since

$$D_N(\lambda) = \begin{cases} N, & \text{if } 0 \leq \lambda < \frac{1}{N}, \\ 0, & \text{otherwise,} \end{cases}$$

we have, for $J \geq 2$

$$\int_0^1 \cdots \int_0^1 \prod_{n=1}^{s} D_N\Big(\bigoplus_{(j,i)\in\nu_n} \lambda_j\Big) d\lambda_1 \cdots d\lambda_J$$

$$= \int_0^1 \cdots \int_0^1 D_N(\mu_1 \oplus \mu_2) D_N(\mu_2 \oplus \mu_3) \cdots D_N(\mu_s \oplus \mu_1) d\mu_1 \cdots d\mu_s = O(N),$$

and for $J = 1$

$$\int_0^1 \cdots \int_0^1 D_N(\lambda_1 \oplus \cdots \oplus \lambda_J) d\lambda_1 \cdots d\lambda_J = O(1).$$

Thus,

$$\int_0^1 \cdots \int_0^1 \mathrm{cum}\big(d_{11}(\lambda_1) d_{12}(\lambda_1), \ldots, d_{J1}(\lambda_J) d_{J2}(\lambda_J)\big) d\lambda_1 \cdots d\lambda_j = O(N),$$

and consequently

$$\mathrm{cum}(A_{i_1}, \ldots, A_{i_J}) = N^{-J/2} \int_0^1 \cdots \int_0^1 \phi_{i_1}(\lambda_1) \cdots \phi_{i_J}(\lambda_J)$$

$$\mathrm{cum}\big(d_{11}(\lambda_1) d_{12}(\lambda_1), \ldots, d_{J1}(\lambda_J) d_{J2}(\lambda_J)\big) d\lambda_1 \cdots d\lambda_J$$

$$= O(N^{-J/2+1}),$$

which implies the asymptotic normality. □

Suppose $f(\lambda)$ is the spectral density of a stationary process and $\{f_{\underset{\sim}{\theta}}(\lambda)\}$ is a family of fitted spectral densities which are parameterized by $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^r$, where Θ is a compact set in $\mathbb{R}^r$. We define a pseudo-true value $\bar{\underset{\sim}{\theta}}$ of $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^r$, by a value which minimizes

$$D(f_{\underset{\sim}{\theta}}, f) = \int_0^1 \{\log \det f_{\underset{\sim}{\theta}}(\lambda) + \mathrm{tr}\, f(\lambda) f_{\underset{\sim}{\theta}}(\lambda)^{-1}\} d\lambda$$

with respect to $\underset{\sim}{\theta} \in \Theta$.

ASSUMPTION 2. The fitted model $f_{\underset{\sim}{\theta}}(\lambda)$ is twice continuously differentiable with respect to $\underset{\sim}{\theta} \in \Theta$.

ASSUMPTION 3. If $\underset{\sim}{\theta} \neq \underset{\sim}{\theta}^*$, then $f_{\underset{\sim}{\theta}}(\lambda) \neq f_{\underset{\sim}{\theta}^*}(\lambda)$ on a set of positive Lebesgue measure. The matrix

$$M_f(\underset{\sim}{\theta}) = \int_0^1 \frac{\partial^2}{\partial\underset{\sim}{\theta}\partial\underset{\sim}{\theta}'} [\log \det f_{\underset{\sim}{\theta}}(\lambda) + \operatorname{tr} f_{\underset{\sim}{\theta}}(\lambda)^{-1} f(\lambda)] d\lambda, \tag{2.13}$$

is nonsingular for all $\underset{\sim}{\theta} \in \Theta$, and $M_f = M_f(\bar{\underset{\sim}{\theta}})$.

The first statement of Assumption 3 is an identifiaility condition. In Section 4 some nonidentifiable examples will be given. Then we have the following theorem.

THEOREM 1. Let $\{Y(t)\}$ be a q-dimensional dyadic stationary process with mean zero and the spectral density $f(\lambda)$. Suppose that Assumptions 1-3 are satisfied, and that $\bar{\underset{\sim}{\theta}}$ exists uniquely and lies in Int Θ. Then

(i) $\lim_{N\to\infty} \hat{\underset{\sim}{\theta}} = \bar{\underset{\sim}{\theta}}$ in P,

(ii) the distribution of the vector $\sqrt{N}\{\hat{\underset{\sim}{\theta}} - \bar{\underset{\sim}{\theta}}\}$, as $N \to \infty$, tends to the normal distribution with mean zero and covariance matrix $M_f^{-1} V M_f^{-1}$, where $V = \{V_{jm}\}$ is a r×r matrix such that

$$V_{jm} = 2\int_0^1 \operatorname{tr}[f(\lambda) \frac{\partial}{\partial\theta_j} \{f_{\underset{\sim}{\theta}}(\lambda)\}^{-1} f(\lambda) \frac{\partial}{\partial\theta_m} \{f_{\underset{\sim}{\theta}}(\lambda)\}^{-1}]_{\underset{\sim}{\theta}=\bar{\underset{\sim}{\theta}}} d\lambda$$

$$+ \sum_{a,b,c,d=1}^{q} \int_0^1\int \left\{\frac{\partial f_{\underset{\sim}{\theta}}^{(b,a)}(\lambda)}{\partial\theta_j} \cdot \frac{\partial f_{\underset{\sim}{\theta}}^{(d,c)}(\lambda)}{\partial\theta_m}\right\}_{\underset{\sim}{\theta}=\bar{\underset{\sim}{\theta}}} f_{abcd}(\lambda,\lambda,\mu) d\lambda d\mu,$$

where $f_{\underset{\sim}{\theta}}^{(b,a)}(\lambda)$ is the (b,a)-th element of $f_{\underset{\sim}{\theta}}(\lambda)^{-1}$.

Proof. From the definitions of $\hat{\underset{\sim}{\theta}}$ and $\bar{\underset{\sim}{\theta}}$, we have

$$\frac{\partial}{\partial\underset{\sim}{\theta}} D(f_{\underset{\sim}{\theta}}, I_N)_{\underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}} = 0, \tag{2.14}$$

$$\frac{\partial}{\partial \underset{\sim}{\theta}} D(f_{\underset{\sim}{\theta}}, f)_{\underset{\sim}{\theta}=\bar{\underset{\sim}{\theta}}} = 0. \tag{2.15}$$

Expanding (2.14) around $\bar{\underset{\sim}{\theta}}$, we have

$$0 = \frac{\partial}{\partial \underset{\sim}{\theta}} D(f_{\bar{\underset{\sim}{\theta}}}, I_N) + \tilde{M}_f(\underset{\sim}{\theta}^*)(\hat{\underset{\sim}{\theta}} - \bar{\underset{\sim}{\theta}}), \tag{2.16}$$

where $\underset{\sim}{\theta}^*$ lies on the straight section with end points $\bar{\underset{\sim}{\theta}}$ and $\hat{\underset{\sim}{\theta}}$, and

$$\tilde{M}_f(\underset{\sim}{\theta}^*) = \frac{\partial^2}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}^T} D(f_{\underset{\sim}{\theta}^*}, I_N).$$

By Lemma 1, we have

$$\frac{\partial}{\partial \underset{\sim}{\theta}} D(f_{\bar{\underset{\sim}{\theta}}}, I_N) \to 0, \quad \text{in } P$$

and

$$\tilde{M}_f(\underset{\sim}{\theta}) \to M_f(\underset{\sim}{\theta}), \quad \text{in } P \quad \text{for each } \underset{\sim}{\theta} \in \Theta.$$

By Assumptions 2 and 3, absolute values of eigenvalues of $M_f(\theta)$ have a positive lower bound for all $\underset{\sim}{\theta} \in \Theta$. Hence when n is large enough, with a probability arbitrarily near to one, so do the absolute values of eigenvalues of $\tilde{M}_f(\underset{\sim}{\theta})$. By (2.16) we have

$$\hat{\underset{\sim}{\theta}} \to \bar{\underset{\sim}{\theta}}, \quad \text{in } P$$

and consequently

$$\tilde{M}_f(\underset{\sim}{\theta}^*) \to M_f(\bar{\underset{\sim}{\theta}}), \quad \text{in } P.$$

Then the limiting distribution of $\sqrt{N}(\hat{\underset{\sim}{\theta}} - \bar{\underset{\sim}{\theta}})$ is equivalent to that of

$$\begin{aligned} &-M_f^{-1}\sqrt{N}\,\frac{\partial}{\partial \underset{\sim}{\theta}} D(f_{\bar{\underset{\sim}{\theta}}}, I_N) \\ &\quad = -M_f^{-1}\sqrt{N}\int_0^1 \frac{\partial}{\partial \underset{\sim}{\theta}} \{\log \det f_{\bar{\underset{\sim}{\theta}}}(\lambda) + \operatorname{tr} f_{\bar{\underset{\sim}{\theta}}}(\lambda)^{-1} I_N(\lambda)\} d\lambda \\ &\quad = -M_f^{-1}\sqrt{N}\int_0^1 \frac{\partial}{\partial \underset{\sim}{\theta}} [\operatorname{tr} f_{\bar{\underset{\sim}{\theta}}}(\lambda)^{-1}\{I_N(\lambda) - f(\lambda)\}] d\lambda, \end{aligned} \tag{2.17}$$

by (2.15). Again applying Lemma 1 to (2.17), we have completed the proof. □

Remark. If the true Walsh spectral density matrix $f(\lambda) = f_{\underset{\sim}{\theta}}(\lambda)$, the pseudo-true value is equal to the true value, i.e., $\bar{\underset{\sim}{\theta}} = \underset{\sim}{\theta}$ (see Hosoya and Taniguchi (1982)).

3. MODEL SELECTION OF WALSH SPECTRAL MODELS

In the previous section we assumed that the order of the unknown parameter vector of the Walsh spectral model $f_{\underset{\sim}{\theta}}(\lambda)$ is known. However, in the actual situation, we must estimate the order of dim $\underset{\sim}{\theta} = r$ from the data. Here we assume that the process $\{Y(t)\}$ has the Walsh spectral density matrix $f_{\underset{\sim}{\theta}_r}(\lambda)$, $\underset{\sim}{\theta}_r = (\theta_1, \ldots, \theta_r)'$, where $\underset{\sim}{\theta}_r$ is an unknown parameter vector. (We use suffix r to stress the dimension.) Then we fit the Walsh spectral model $f_{\underset{\sim}{\theta}_k}(\lambda)$, $0 \leq k \leq L$, where L is a preassigned upper limit to the order. We determine the true order r by the value $\hat{k}$ which minimizes the following criterion:

$$A(k) = D(f_{\hat{\underset{\sim}{\theta}}_k}, I_N) + \frac{kC_N}{N} \quad \text{for} \quad k = 0,1,\ldots,L, \tag{3.1}$$

where $C_N \to \infty$ and $C_N/N \to 0$ as $N \to \infty$. For this estimated order $\hat{k}$ we have

THEOREM 2. Suppose that all the assumtions in Section 2 for $f(\lambda) = f_{\underset{\sim}{\theta}_r}(\lambda)$ and $f_{\underset{\sim}{\theta}}(\lambda) = f_{\underset{\sim}{\theta}_k}(\lambda)$ are satisfied. Then $\lim_{N\to\infty} \hat{k} = r$, in p.

Proof. From (2.16) we have

$$\sqrt{N}(\hat{\underset{\sim}{\theta}}_k - \bar{\underset{\sim}{\theta}}_k) = -\tilde{M}_f^{-1}(\underset{\sim}{\theta}_k^*)\sqrt{N} \int_0^1 \frac{\partial}{\partial \underset{\sim}{\theta}_k} \left[\operatorname{tr} f_{\bar{\underset{\sim}{\theta}}_k}(\lambda)^{-1}\{I_N(\lambda) - f(\lambda)\}\right] d\lambda \tag{3.2}$$

which tends to normal by Theorem 1. Thus we have, for any sequence of positive numbers $\tilde{C}_N \to \infty$,

$$P[\| \sqrt{N}(\hat{\underset{\sim}{\theta}}_k - \bar{\underset{\sim}{\theta}}_k)\| > \tilde{C}_N] = o(1), \tag{3.3}$$

where $\|\cdot\|$ is the Euclidian norm. Taking $\tilde{C}_N = \sqrt[4]{C_N}$, we obtain

$$\hat{\underset{\sim}{\theta}}_k - \bar{\underset{\sim}{\theta}}_k = O_p(\sqrt[4]{C_N}/\sqrt{N}). \tag{3.4}$$

Expanding around $\underset{\sim}{\theta} = \hat{\underset{\sim}{\theta}}_k$, and noting (3.4) we can see that

$$D(f_{\bar{\underset{\sim}{\theta}}_k}, I_N) = D(f_{\hat{\underset{\sim}{\theta}}_k}, I_N) + (\bar{\underset{\sim}{\theta}}_k - \hat{\underset{\sim}{\theta}}_k)' \left.\frac{\partial D(f_{\underset{\sim}{\theta}_k}, I_N)}{\partial \underset{\sim}{\theta}_k}\right|_{\underset{\sim}{\theta}_k = \hat{\underset{\sim}{\theta}}_k}$$

$$+ \frac{1}{2}(\bar{\underset{\sim}{\theta}}_k - \hat{\underset{\sim}{\theta}}_k)' \tilde{M}_f(\underset{\sim}{\theta}_k^*)(\underset{\sim}{\theta}_k - \hat{\underset{\sim}{\theta}}_k). \tag{3.5}$$

Since $\left.\frac{\partial D(f_{\underset{\sim}{\theta}_k}, I_N)}{\partial \underset{\sim}{\theta}_k}\right|_{\underset{\sim}{\theta}_k = \hat{\underset{\sim}{\theta}}} = \phi$, we have

$$D(f_{\hat{\underset{\sim}{\theta}}_k}, I_N) = D(f_{\bar{\underset{\sim}{\theta}}_k}, I_N) - \frac{1}{2}(\hat{\underset{\sim}{\theta}}_k - \bar{\underset{\sim}{\theta}}_k)' \tilde{M}_f(\underset{\sim}{\theta}_k^*)(\hat{\underset{\sim}{\theta}}_k - \bar{\underset{\sim}{\theta}}_k). \tag{3.6}$$

As first step we show that

$$P(\hat{k} < r) \to 0 \quad \text{as} \quad N \to \infty. \tag{3.7}$$

For $k < r$, we evaluate

$$P_1 = P\big(A(k) < A(r)\big) = P\{D(f_{\hat{\underset{\sim}{\theta}}_k}, I_N) - D(f_{\hat{\underset{\sim}{\theta}}_r}, I_N) < \frac{(r-k)C_N}{N}\}.$$

Using the relation (3.6), the above probability is approximated as

$$P\{D(f_{\bar{\underset{\sim}{\theta}}_k}, I_N) - D(f_{\bar{\underset{\sim}{\theta}}_r}, I_N)$$

$$< \frac{(r-k)C_N}{N} + \frac{1}{2}(\hat{\underset{\sim}{\theta}}_k - \bar{\underset{\sim}{\theta}}_k)' \tilde{M}_f(\underset{\sim}{\theta}_k^*)(\hat{\underset{\sim}{\theta}}_k - \bar{\underset{\sim}{\theta}}_k)$$

$$- \frac{1}{2}(\hat{\underset{\sim}{\theta}}_r - \bar{\underset{\sim}{\theta}}_r)' M_f(\underset{\sim}{\theta}_r^*)(\hat{\underset{\sim}{\theta}}_r - \bar{\underset{\sim}{\theta}}_r)\}. \tag{3.8}$$

Using Lemma 1 the left-hand side of the above $\{\cdot\}$ converges to $D(f_{\bar{\underset{\sim}{\theta}}_k}, f) - D(f_{\bar{\underset{\sim}{\theta}}_r}, f)$, which is strictly positive for $k < r$. On the other hand, by (3.4), the right-hand side of $\{\cdot\}$ converges to zero in probability which implies the

probability $P_1 \to 0$ as $N \to \infty$. As second step we show

$$P(\hat{k} > r) \to 0 \quad \text{as} \quad N \to \infty. \tag{3.9}$$

We have for $k > r$,

$$P_2 = P\{A(k) < A(r)\} = P\{D(f_{\hat{\underset{\sim}{\theta}}_k}, I_N) - D(f_{\hat{\underset{\sim}{\theta}}_r}, I_N) < \frac{(r-k)C_N}{N}\}.$$

Using the relation (3.6), the above probability is approximated as

$$P\{D(f_{\bar{\underset{\sim}{\theta}}_k}, I_N) - D(f_{\bar{\underset{\sim}{\theta}}_r}, I_N) - \frac{1}{2}(\hat{\underset{\sim}{\theta}}_k - \bar{\underset{\sim}{\theta}}_k)'\tilde{M}_f(\underset{\sim}{\theta}_k^*)(\hat{\underset{\sim}{\theta}}_k - \bar{\underset{\sim}{\theta}}_k)$$

$$+ \frac{1}{2}(\hat{\underset{\sim}{\theta}}_r - \bar{\underset{\sim}{\theta}}_r)'\tilde{M}_f(\underset{\sim}{\theta}_r^*)(\hat{\underset{\sim}{\theta}}_r - \bar{\underset{\sim}{\theta}}_r) < \frac{(r-k)C_N}{N}\}. \tag{3.10}$$

Because $f_{\bar{\underset{\sim}{\theta}}_k}(\lambda) = f_{\bar{\underset{\sim}{\theta}}_r}(\lambda)$, for $k \geq r$, we can see that

$$D(f_{\bar{\underset{\sim}{\theta}}_k}, I_N) - D(f_{\bar{\underset{\sim}{\theta}}_r}, I_N) = 0. \tag{3.11}$$

While, by (3.4), we can see that

$$-\frac{1}{2}(\hat{\underset{\sim}{\theta}}_k - \bar{\underset{\sim}{\theta}}_k)'\tilde{M}_f(\underset{\sim}{\theta}_k^*)(\hat{\underset{\sim}{\theta}}_k - \bar{\underset{\sim}{\theta}}_k) + \frac{1}{2}(\hat{\underset{\sim}{\theta}}_r - \bar{\underset{\sim}{\theta}}_r)'\tilde{M}_f(\underset{\sim}{\theta}_r^*)(\hat{\underset{\sim}{\theta}}_r - \bar{\underset{\sim}{\theta}}_r)$$

is at most of order $O_p(\frac{\sqrt{C_N}}{N})$. However the right-hand side of $\{\cdot\}$ in (3.10) is $\frac{(r-k)C_N}{N}$, $(r < k)$, which implies $P_2 \to 0$, as $N \to \infty$. Thus we have completed the proof. $\square$

4. DETERMINATION OF THE ORDER OF DYADIC AUTOREGRESSIVE MODELS

In the previous sections we could proceed in fairly analogous ways to those used in the ordinary stationary processes. However if we consider finite parametric models, for examples, dyadic autoregressive process of finite order (DAR-process), dyadic moving average process of finite order (DMA-process) and dyadic autoregressive moving average process of finite order (DARMA-process), then there exist the greatest differences between dyadic stationary processes and ordinary stationary ones. That is, it is known that these DAR, DMA and DARMA are equivalent in the sense that DAR or DARMA-process of finite order can be expressed as DMA-process of finite order (see Nagai (1980) or Nagai and Taniguchi (1987)).

In this section, for a finite order dyadic autoregressive model, we can propose a simpler order determination criterion. Then we show that this criterion gives strong consistent order estimate.

A q-dimensional dyadic stationary process $\{Y(t), t \in T\}$ is called a dyadic autoregressive process, if it can be expressed by

$$\sum_{j=0}^{p} A_j Y(t \oplus j) = \varepsilon(t), \quad t \in T, \tag{4.1}$$

where

(i) A_j's are $q \times q$ matrices, $A_0 = I_q$, and $p = 2^r - 1$, where r is a non-negative integer,

(ii) $\varepsilon(t)$, $t \in T$, are i.i.d. random vectors such that

$$E\varepsilon(t) = 0, \; E\varepsilon(t)\varepsilon(t)' = G > 0, \tag{4.2}$$

(iii) $$\det \Phi(\lambda) \neq 0, \quad \text{a.e.}\lambda, \tag{4.3}$$

where $$\Phi(\lambda) = \sum_{j=0}^{p} A_j W(j,\lambda).$$

If (4.2) and (4.3) hold true, then the Walsh spectral density of $\{Y(t)\}$ is

$$f(\lambda) = \Phi(\lambda)^{-1}G\{\Phi(\lambda)^{-1}\}'. \tag{4.4}$$

We call a DAR process (4.1) irreducible if there does not exist such a matrix

$$\Phi_1(\lambda) = \sum_{j=0}^{2^{r-1}-1} K_j W(j,\lambda),$$

which satisfies

$$f(\lambda) = \Phi_1(\lambda)^{-1}G\{\Phi_1(\lambda)^{-1}\}', \quad \text{a.e. } \lambda. \tag{4.5}$$

Especially, for an irreducible DAR process (4.1), there exists a t_0, $2^{r-1} \leq t_0 \leq 2^r - 1$, such that $A_{t_0} \neq 0$.

For an irreducible DAR model (4.1), p is called the order of the model. For simplicity, such a model is written as DAR(p). Note that in the above definition, the order of the model (4.1) is defined as $p = 2^r - 1$, not as $\max\{t: A_t \neq 0\}$. The advantage of such a definition is that it suits to the Walsh spectrum analysis, and is convenient for estimating the parameters of the model. To see this, consider the following two scalar irreducible DAR model:

$$X(t) + X(t \oplus 1) + \alpha X(t \oplus 2) = \varepsilon(t),$$

and

$$Y(t) + Y(t \oplus 1) + \alpha Y(t \oplus 3) = \varepsilon(t), \quad t \in T,$$

where $\alpha \neq 0$, $\alpha \neq \pm 2$, $\varepsilon(t)$'s are i.i.d. with $E\varepsilon(t) = 0$, $E\varepsilon(t)^2 = \sigma^2$. It is easily seen that they have the same Walsh spectral density

$$\sigma^2[1 + W(1,\lambda) + \alpha W(2,\lambda)]^{-2}.$$

But if we define the order of the model as $\max\{t: A_t \neq 0\}$, then their order

may be 2 and 3 respectively. Obviously such a definition is not convenient for Walsh spectrum analysis. It is easy to see that these two models are not essentially different. For a q×q matrix $A = (a_{ij}, 1 \le i, j \le q)$, denote $\|A\| = \sum_{i,j=1}^{q} |a_{ij}|$. To determine the order $p = 2^r - 1$ of the irreducible model (4.1), we suggest the following criterion:

$$L_N(k) = \frac{1}{2^k} \sum_{n=0}^{2^k-1} \left\| \frac{1}{N} \sum_{t=0}^{N-1} Y(t) Y\left(t \oplus (2^k + n)\right)' \right\|^2 - \frac{C_N}{N}, \tag{4.6}$$

where $Y(0), \ldots, Y(N-1)$ are the observations of the model (4.1), $N = 2^m$ with m positive integer, and C_N satisfies the following conditions:

$$\lim_{N\to\infty} \frac{C_N}{N} = 0 \quad \text{and} \quad \lim_{N\to\infty} \frac{C_N}{\log \log N} = \infty. \tag{4.7}$$

Define

$$\hat{r}_N = \max\{k \ge 0 : L_N(k-1) > 0,\ L_N(k) < 0\}. \tag{4.8}$$

where $L_N(-1) = 1$ for convenience. We can use $\hat{r}_N$ as an estimate of the true value r of the model (4.1). We have the following

THEOREM 3. If the model (4.1) is irreducible and (i), (ii) and (iii) are satisfied, then

$$\lim_{N\to\infty} \hat{r}_N = r, \quad \text{a.s.} \tag{4.9}$$

Proof. Suppose that p is the true order of the model (4.1) and $p = 2^r - 1$. According to Nagai and Taniguchi (1987), if $\det\{\Phi(\lambda)\} \neq 0$, then $\{Y(t): t \in T\}$ is a DMA-process written by

$$Y(t) = \sum_{j=0}^{2^r-1} K_{j \oplus t} \varepsilon(j), \quad t \in T. \tag{4.10}$$

Put $\Gamma(n) = EY(0)Y'(n)$. By (4.10) and the condition (ii), it is easily seen that for any n,

$$\left\|\frac{1}{N}\sum_{t=0}^{N-1} Y(t)Y'(t\oplus n) - \Gamma(n)\right\| = O\left(\sqrt{\frac{\log\log N}{N}}\right), \quad \text{a.s.} \tag{4.11}$$

as $N \to \infty$, (e.g., Petrov (1975)). By (4.10), for $n \geq 2^r$, $\Gamma(n) = 0$. Thus, if $k \geq r$, then

$$L_N(k) = O\left(\frac{\log\log N}{N}\right) - \frac{C_N}{N}, \quad \text{a.s.} \tag{4.12}$$

as $N \to \infty$. From this and $\frac{C_N}{\log\log N} \approx \infty$, it follows that with probability one for N large,

$$L_N(k) < 0, \quad k \geq r. \tag{4.13}$$

If $r = 0$, the theorem is proved.

Now assume that $r > 0$. We have

$$\lim_{N\to\infty} \sum_{n=0}^{2^{r-1}-1} \left\|\frac{1}{N}\sum_{t=0}^{N-1} Y(t)Y\left(t\oplus(2^{r-1}+n)\right)'\right\|^2$$

$$= \sum_{n=0}^{2^{r-1}-1} \|\Gamma(2^{r-1}+n)\|^2, \quad \text{a.s.} \tag{4.14}$$

We proceed to prove that

$$\sum_{n=0}^{2^{r-1}-1} \|\Gamma(2^{r-1}+n)\|^2 > 0. \tag{4.15}$$

Otherwise, we have

$$f(\lambda) = \sum_{\ell=0}^{2^{r-1}-1} \Gamma(\ell)W(\ell,\lambda), \quad \lambda \in [0,1]. \tag{4.16}$$

Put $h = 2^{r-1}-1$, $\lambda_j = j/(h+1)$, $j = 0,1,\ldots,h$. We know that for all $\ell \le h$, $W(\ell,\lambda) = W(\ell,\lambda_j)$ for $\lambda \in [\lambda_j, \lambda_{j+1})$. From this it is easily seen that $f(\lambda)$ takes only at most $h + 1$ different values, say, $f(\lambda_0),\ldots,f)\lambda_h)$. By (4.4), $G > 0$ and $\Phi(\lambda) \neq 0$, it is easily seen that $f(\lambda_j) > 0$, $j = 0,1,\ldots,h$. Hence we can write

$$G = G^{1/2}(G^{1/2})', \quad f(\lambda_j) = f^{1/2}(\lambda_j)f^{1/2}(\lambda_j)',$$

$j - 0,1,\ldots,h$. Put

$$H_{h+1} = \begin{bmatrix} W(0,\lambda_0) & W(1,\lambda_0) & \ldots, & W(h,\lambda_0) \\ W(0,\lambda_1) & W(1,\lambda_1) & \ldots, & W(h,\lambda_1) \\ \vdots & \vdots & & \vdots \\ W(0,\lambda_h) & W(1,\lambda_h) & \ldots, & W(h,\lambda_h) \end{bmatrix}.$$

Then $H'_{h+1}H_{h+1} = (h+1)I_{h+1}$. Thus the matrix equation

$$(H_{h+1} \otimes I_q)\begin{pmatrix} B_0 \\ \vdots \\ B_h \end{pmatrix} = \begin{pmatrix} G^{-1/2}f^{1/2}(\lambda_0) \\ \vdots \\ G^{-1/2}f^{1/2}(\lambda_h) \end{pmatrix} \tag{4.17}$$

has a unique solution $(B'_0, \ldots, B'_h)$, where B_j's are all $q\times q$ matrices. From (4.17) we can see that

$$\Big(\sum_{\ell=0}^{h} B_\ell W(\ell,\lambda_j)\ G\big(\sum_{\ell=0}^{h} B_\ell W(\ell,\lambda_j)\big)' = f(\lambda_j), \quad j = 0,1,\ldots,h, \tag{4.18}$$

which implies

$$\eta(\lambda)G\eta(\lambda)' = f(\lambda), \quad \lambda \in [0,1] \tag{4.19}$$

where

$$\eta(\lambda) = \sum_{\ell=0}^{h} B_\ell W(\ell,\lambda). \tag{4.20}$$

By Nagai and Taniguchi (1987) there exists

$$\Phi_1(\lambda) = \sum_{\ell=0}^{2^{r-1}-1} K_\ell W(\ell,\lambda)$$

such that

$$\Phi_1(\lambda)\eta(\lambda) = I_q, \quad \text{a.e. } \lambda. \tag{4.21}$$

Thus we have

$$f(\lambda) = \Phi_1(\lambda)^{-1} G\{\Phi_1(\lambda)^{-1}\}', \quad \text{a.e. } \lambda, \tag{4.22}$$

which contradicts our irreducibility assumption. Now (4.15) has been proved. By (4.14), (4.15), (4.6) and (4.7), with probability one for large N,

$$L_N(r-1) > 0, \quad r > 0. \tag{4.23}$$

Noting (4.13) and (4.23), with probability one for large N, we have

$$\hat{r}_N = r. \tag{4.24}$$

Remark. The following scalar process $\{Y(t);\ t \in T\}$ is a reducible DAR process:

$$X(t) + X(t\oplus 1) + X(t\oplus 2) - X(t\oplus 3) = \varepsilon(t), \quad t \in T,$$

where $\varepsilon(t)$'s are i.i.d. with $E\varepsilon(t) = 0$ and $E\varepsilon(t)^2 = \sigma^2$. Then

$$\Phi(\lambda) = 1 + W(1,\lambda) + W(2,\lambda) - W(3,\lambda),$$

but

$$\frac{\sigma^2}{\{\Phi(\lambda)\}^2} = \frac{\sigma^2}{4}.$$

5. DETECTION OF SIGNALS FOR DYADIC STATIONARY PROCESSES

In this section we shall consider a signal detection model for dyadic process of finite order, and show that this model is equivalent to a dyadic moving average model. Then we can apply the results in Sections 3 and 4 to our model. That is, we can determine the order of the signal dyadic process from the data.

Let $\{Y(t)\}$ be a q-dimensional dyadic stationary process defined by

$$Y(t) = \sum_{j=0}^{p} A(j)S(t \oplus j) + \xi(t), \tag{5.1}$$

where $S(t)$ is an r-dimensional dyadic stationary signal process, and $\xi(t)$ is a q-dimensional dyadic stationary noise process, and $A(j)$ are $q \times r$-matrices. Here we assume that $\{S(t)\}$ is an r-dimensional DARMA(s,h)-process defined by

$$\sum_{j=0}^{s} B(j)S(j \oplus t) = \sum_{j=0}^{h} C(j)U(j \oplus t), \quad t = 0,1,\ldots, \tag{5.2}$$

where $\{U(t)\}$ is an r-dimensional white noise process. Also $\{\xi(t)\}$ is a q-dimensional DARMA(ℓ,m)-process defined by

$$\sum_{j=0}^{\ell} D(j)\xi(j \oplus t) = \sum_{j=0}^{m} F(j)V(j \oplus t), \quad t = 0,1,\ldots, \tag{5.3}$$

where $\{V(t)\}$ is a q-dimensional white noise process which is independent of $\{U(t)\}$. We assume that all the coefficients $\{A(j)\}$, $\{B(j)\}$, $\{C(j)\}$, $\{D(j)\}$ and $\{F(j)\}$ are completely specified by an unknown parameter vector $\underset{\sim}{\theta}$ with dimension k, and that

$$\begin{aligned} &\det\{\sum_{j=0}^{s} B(j)W(j,\lambda)\} \neq 0, \\ &\det\{\sum_{j=0}^{\ell} D(j)W(j,\lambda)\} \neq 0, \quad \text{for all} \quad 0 \leq \lambda \leq 1. \end{aligned} \tag{5.4}$$

Denote by Walsh spectral representations of $\{Y(t)\}$, $\{S(t)\}$, $\{\xi(t)\}$, $\{U(t)\}$ and $\{V(t)\}$ by $Y(t) = \int_0^1 W(t,\lambda)dZ_Y(\lambda)$, $S(t) = \int_0^1 W(t,\lambda)dZ_S(\lambda)$, $\xi(t) = \int_0^1 W(t,\lambda)dZ_\xi(\lambda)$, $U(t) = \int_0^1 W(t,\lambda)dZ_U(\lambda)$ and $V(t) = \int_0^1 W(t,\lambda)dZ_V(\lambda)$, respectively. The relation (5.1) can be written as

$$dZ_Y(\lambda) = \Psi_A(\lambda)dZ_S(\lambda) + dZ_\xi(\lambda), \tag{5.5}$$

briefly, where $\Psi_A(\lambda) = \sum_{j=0}^{p} A(j)W(j,\lambda)$. While we can have

$$dZ_S(\lambda) = \Psi_B(\lambda)^{-1}\Psi_C(\lambda)dZ_U(\lambda), \tag{5.6}$$

$$dZ_\xi(\lambda) = \Psi_D(\lambda)^{-1}\Psi_F(\lambda)dZ_V(\lambda), \tag{5.6}$$

where $\Psi_B(\lambda) = \sum_{j=0}^{s} B(j)W(j,\lambda)$, $\Psi_C(\lambda) = \sum_{j=0}^{h} C(j)W(j,\lambda)$, $\Psi_D(\lambda) = \sum_{j=0}^{\ell} D(j)W(j,\lambda)$ and $\Psi_F(\lambda) = \sum_{j=0}^{m} F(j)W(j,\lambda)$. Finally, we get

$$\begin{aligned} dZ_Y(\lambda) &= \Psi_A(\lambda)\Psi_B(\lambda)^{-1}\Psi_C(\lambda)dZ_U(\lambda) + \Psi_D(\lambda)^{-1}\Psi_F(\lambda)dZ_V(\lambda) \\ &= [\Psi_A(\lambda)\Psi_B(\lambda)^{-1}\Psi_C(\lambda) : \Psi_D(\lambda)^{-1}\Psi_F(\lambda)] \begin{bmatrix} dZ_U(\lambda) \\ dZ_V(\lambda) \end{bmatrix}. \end{aligned} \tag{5.8}$$

By Nagai and Taniguchi (1987), under (5.4), we can get the following representations

$$\Psi_B(\lambda)^{-1} = \sum_{j=0}^{s'} B(j)W(j,\lambda)$$

and

$$\Psi_D(\lambda)^{-1} = \sum_{j=0}^{\ell'} D(j)W(j,\lambda),$$

where $s' = 2^a - 1$, $\ell' = 2^b - 1$, a and b are the minimum nonnegative integers which satisfy $s \le 2^a - 1$, $\ell \le 2^b - 1$, respectively. Thus it is easy to see that $\Psi_A(\lambda)\Psi_B(\lambda)^{-1}\Psi_C(\lambda)$ and $\Psi_D(\lambda)^{-1}\Psi_F(\lambda)$ can be written as finite linear

combinations of Walsh functions

$$\sum_{j=0}^{d_1} G(j)W(j,\lambda), \qquad \sum_{j=0}^{d_1} H(j)W(j,\lambda),$$

respectively. So the relation (5.8) implies that our model (5.1) is equivalent to a DMA-process of finite order. We can assume that our $\{Y(t)\}$ has the finite DMA type Walsh spectral density matrix $f_{\underset{\sim}{\theta}}(\lambda)$. To determine the order of dim Θ, we can use the criterion

$$A(k) = D(f_{\hat{\underset{\sim}{\theta}}_k}, I_N) + \frac{kC_N}{N},$$

given in (3.1). Of course we can use $A(k)$ to determine the number of signals.

6. TEST OF HYPOTHESIS FOR LINEAR RESTRICTION OF PARAMETERS

Let $\{Y(t)\}$ be a scalar-valued dyadic stationary process with Walsh spectral density $f_{\underset{\sim}{\theta}}(\lambda)$ depending on an unknown parameter $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_p)'$. We assume that $\{Y(t)\}$ satisfies all the assumptions in Theorem 1. The first problem is to test a composite hypothesis H_0: $\underset{\sim}{\theta}_2 = \underset{\sim}{\theta}_{20}$, against H: $\underset{\sim}{\theta}_2 \neq \underset{\sim}{\theta}_{20}$, where $\underset{\sim}{\theta}' = (\underset{\sim}{\theta}_1', \underset{\sim}{\theta}_2')$, $\underset{\sim}{\theta}_1' = (\theta_1, \ldots, \theta_\ell)$, $\underset{\sim}{\theta}_2' = (\theta_{\ell+1}, \ldots, \theta_p)$ and $\underset{\sim}{\theta}_{20}' = (\theta_{\ell+1,0}, \ldots, \theta_{p,0})$, a specified vector and $(\underset{\sim}{\theta}_1', \underset{\sim}{\theta}_{20}') \in \text{Int}\,\Theta$. Although we do not assume the Gaussianity of $\{Y(t)\}$, we can formally make the following log-likelihood ratio criterion

$$G = 2 \log L = N\{D(f_{(\hat{\underset{\sim}{\theta}}_1, \hat{\underset{\sim}{\theta}}_2)}, I_N) - D(f_{(\tilde{\underset{\sim}{\theta}}_1, \underset{\sim}{\theta}_{20})}, I_N)\}, \tag{6.1}$$

where $\hat{\underset{\sim}{\theta}}' = (\hat{\underset{\sim}{\theta}}_1', \hat{\underset{\sim}{\theta}}_2')$ is the quasi-maximum likelihood estimator for $\underset{\sim}{\theta}$ under H, and $\tilde{\underset{\sim}{\theta}}_1$ is that for $\underset{\sim}{\theta}_1$ under H_0. Put $v = \sqrt{N}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})$, $w = \sqrt{N}(\underset{\sim}{\theta}_1 - \underset{\sim}{\theta}_1)$ and $u' = (w', \phi')$. Expanding in a Taylor expansion around $\hat{\underset{\sim}{\theta}}$, we have

$$\begin{aligned}
-G &= N\{D(f_{(\tilde{\underset{\sim}{\theta}}_1, \underset{\sim}{\theta}_{20})}, I_N) - D(f_{(\hat{\underset{\sim}{\theta}}_1, \hat{\underset{\sim}{\theta}}_2)}, I_N)\} \\
&= \frac{1}{2}(u - v)' \frac{\partial^2 D(f_{\hat{\underset{\sim}{\theta}}}, I_N)}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}'} (u - v)\left(1 + o_p(1)\right) \\
&= \frac{1}{2}(u - v)' \frac{\partial^2 D(f_{\underset{\sim}{\theta}}, I_N)}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}'} (u - v)\left(1 + o_p(1)\right) \\
&= \frac{1}{2}(u - v)' M_f (u - v)'\left(1 + o_p(1)\right).
\end{aligned} \tag{6.2}$$

From Theorem 1 we have

$$v = -M_f^{-1}\sqrt{N}\, \frac{\partial}{\partial \underset{\sim}{\theta}} D(f_{\underset{\sim}{\theta}}, I_N)\left(1 + o_p(1)\right). \tag{6.3}$$

Similarly we have

$$u = -L_f \sqrt{N} \frac{\partial}{\partial \underset{\sim}{\theta}} D(f_{\underset{\sim}{\theta}}, I_N)\left(1 + o_p(1)\right), \tag{6.4}$$

where

$$L_f = \begin{bmatrix} I_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix},$$

and

$$E\frac{\partial^2}{\partial\underset{\sim}{\theta}_1\partial\underset{\sim}{\theta}_1'}D(f_{(\underset{\sim}{\theta}_1,\underset{\sim}{\theta}_{20})}, I_N) = I_{11} + O(N^{-1}).$$

From (6.2), (6.3) and (6.4) we have

$$\begin{aligned} -G &= \frac{1}{2}\sqrt{N}\frac{\partial D(f_{\underset{\sim}{\theta}},I_N)}{\partial\underset{\sim}{\theta}'}[M_f^{-1} - L_f]M_f[M_f^{-1} - L_f]\sqrt{N}\frac{\partial D(f_{\underset{\sim}{\theta}},I_N)}{\partial\underset{\sim}{\theta}}\left(1 + o_p(1)\right) \\ &= \frac{1}{2}\sqrt{N}\frac{\partial D(f_{\underset{\sim}{\theta}},I_N)}{\partial\underset{\sim}{\theta}'}[M_f^{-1} - L_f]\sqrt{N}\frac{\partial D(f_{\underset{\sim}{\theta}},I_N)}{\partial\underset{\sim}{\theta}}\left(1 + o_p(1)\right). \end{aligned} \tag{6.5}$$

Here we put the following assumptions:

Assumption 4. The process $\{Y(t)\}$ is a scalar linear dyadic stationary process represented as

$$Y(t) = \sum_{j=0}^{\infty} A_j e(t \oplus j), \tag{6.6}$$

where $\sum_{j=0}^{\infty}|A_j| < \infty$ and the $e(t)$'s are independent random variables.

Assumption 5. The unknown parameter $\underset{\sim}{\theta}$ of $f_{\underset{\sim}{\theta}}(\lambda)$ is innovation-free, i.e.,

$$\frac{\partial}{\partial\underset{\sim}{\theta}}\int_0^1\{f_{\underset{\sim}{\theta}}(\lambda)^{-1}f_{\underset{\sim}{\theta}_0}(\lambda)\}d\lambda = \phi. \tag{6.7}$$

(See Hosoya and Taniguchi (1982).)

LEMMA 2. Suppose that Assumptions 1-5 are satisfied, and that $\sum_{j=0}A_jW(j,\lambda) \neq 0$ for all $\lambda \in [0,1]$. For an innovation-free parameter $\underset{\sim}{\theta}$ we have

$$\sqrt{\frac{N}{2}}\frac{\partial D(f_{\underset{\sim}{\theta}},I_N)}{\partial\underset{\sim}{\theta}} \xrightarrow{\mathcal{L}} N(\phi_p, M_f). \tag{6.8}$$

Proof. Using a similar argument to Hosoya and Taniguchi (1982), we can see that

$$\int_0^1\int \frac{\partial f_{\underset{\sim}{\theta}}(\lambda)}{\partial\theta_j}\cdot\frac{f_{\underset{\sim}{\theta}}(\lambda)}{\partial\theta_m} f_4(\lambda,\lambda,\mu)d\lambda d\mu = 0,$$

for $j,m = 1,\ldots,p$, where $f_4(\cdot)$ is the fourth order cumulant spectral density. Putting $\phi_j(\lambda) = \frac{\partial}{\partial\theta_j} f_{\underset{\sim}{\theta}}(\lambda)$ in Lemma 1, we have the desired result. □

Applying Lemma 2 to (6.5) we have

THEOREM 4. Suppose that Assumptions 1-5 are satisfied. Then the distribution of $-G$ under H_0 tends to $\chi^2(p-\ell)$ as $N \to \infty$.

Now we consider a more general test of hypothesis.

$$H_0\colon B\underset{\sim}{\theta} = u_{20} \quad \text{against} \quad H\colon B\underset{\sim}{\theta} \neq u_{20},$$

where B is a $(p-\ell)\times p$ matrix with rank $B = p-\ell$, and $u_{20}' = (u_{\ell+1,0},\ldots,u_{p,0})$. Then there exists an $\ell\times p$ matrix A such that

$$\begin{pmatrix} A \\ B \end{pmatrix}\underset{\sim}{\theta} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = u(\underset{\sim}{\theta}),$$

where $\det\binom{A}{B} \neq 0$. Let $\hat{\underset{\sim}{\theta}}$ be the quasi-maximum likelihood estimator of $\underset{\sim}{\theta} \in \Theta$, then $u(\hat{\underset{\sim}{\theta}}) = \hat{u}$. Then the likelihood ratio criterion of testing

$$H_0\colon u_2 = u_{20} \quad \text{against} \quad H\colon u_2 \neq u_{20}$$

is given by

$$\tilde{G} = N\{D(f_{(\hat{u}_1,\hat{u}_2)}, I_N) - D(f_{(\tilde{u}_1,u_{20})}, I_N)\}, \tag{6.9}$$

where $\tilde{u}_1$ is the quasi-maximum likelihood estimator of u_1 under H_0. Then we have

THEOREM 5. Suppose that Assumptions 1-5 are satisfied. Then the distributions of $-\tilde{G}$ under H_0 tends to $\chi^2(p-\ell)$ as $N \to \infty$.

REFERENCES

[1] BRILLINGER, D.R. (1975). *Time Series; Data Analysis and Theory.* Holt, Rinehart & Winston.

[2] HOSOYA, Y. and TANIGUCHI, M. (1982). A central limit theorem for stationary processes and the parameter estimation of linear processes. *Ann. Statist.* **10**, 132-153.

[3] MORETTIN, P.A. (1974). Walsh function analysis of a certain class of time series. *Stochastic Processes and their Applications* **2**, 183-193.

[4] MORETTIN, P.A. (1981). Walsh spectral analysis. *SIAM Review* **23**, 279-291.

[5] NAGAI, T. (1977). Dyadic stationary processes and their spectral representations. *Bull. Math. Statist.* **17**, 65-73.

[6] NAGAI, T. (1980). On finite parametric linear models of dyadic stationary processes. *Bull. Math. Statist.* **19**, 45-53.

[7] NAGAI, T. and TANIGUCHI, M. (1987). Walsh spectral analysis of multiple dyadic stationary processes and its applications. *Stochastic Processes and their Applications* **24**, 19-30.

[8] PETROV, V.V. (1975). *Sums of Independent Random Variables.* Springer-Verlag.

END
DATE
FILMED
DTIC
JULY 88