LEAST-SQUARES AND MAXIMUM-LIKELIHOOD TFAR PARAMETER ESTIMATION FOR NONSTATIONARY PROCESSES

Michael Jachan, Gerald Matz, and Franz Hlawatsch

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria
phone: +43 1 58801 38914, fax: +43 1 58801 38999, email: michael.jachan@tuwien.ac.at

ABSTRACT

Time-frequency-autoregressive (TFAR) models allow the parsimonious modeling of underspread nonstationary random processes and are physically meaningful due to their formulation in terms of delays and Doppler frequency shifts. Here, we derive least-squares (LS) and maximum-likelihood (ML) methods for TFAR parameter estimation as well as approximative LS and ML methods specifically suited for the underspread case. We show that the LS, underspread LS, and underspread ML estimators are equivalent to estimators based on linear time-frequency Yule-Walker (TFYW) equations. The exact ML estimator, on the other hand, requires numerical maximization but yields better estimation accuracy than TFYW techniques. We also discuss the application of block-based TFAR estimation to the spectral analysis of signals with arbitrary length.

1. INTRODUCTION

1.1. Review of the TFAR Model

Time-frequency-autoregressive (TFAR), TF-moving-average (TF-MA), and TFARMA models have been introduced as parsimonious models for nonstationary finite-length random processes $x[n]$, $n = 0, \ldots, N-1$ [1,2]. These models are physically intuitive because of their formulation in terms of *time shifts* (delays) $(\mathbb{T}^m x)[n] = x[n-m]$ and *frequency shifts* (Doppler shifts) $(\mathbb{M}^l x)[n] = e^{j\frac{2\pi}{N}ln}x[n]$.

The TFAR$(M_A, L_A; L_B)$ model [1], in particular, is defined by the input-output relation

$$x[n] = -\sum_{m=1}^{M_A}\sum_{l=-L_A}^{L_A} a_{m,l}\,(\mathbb{S}_{m,l}\,x)[n] + \sum_{l=-L_B}^{L_B} b_{0,l}\,(\mathbb{M}^l\,e)[n]. \quad (1)$$

Here, $\mathbb{S}_{m,l} \triangleq \mathbb{M}^l\mathbb{T}^m$ is the TF shift operator acting as $(\mathbb{S}_{m,l}\,x)[n] = e^{j\frac{2\pi}{N}ln}x[n-m]$; $e[n]$ is stationary white innovations noise with variance 1; M_A and L_A denote the delay and Doppler model orders, respectively; L_B denotes the Doppler model order of the TFMA part (see below); the $M_A(2L_A+1)$ constants $a_{m,l}$, $m = 1, \ldots, M_A$, $l = -L_A, \ldots, L_A$ are the TFAR parameters; and the $2L_B + 1$ constants $b_{0,l}$, $l = -L_B, \ldots, L_B$ are TFMA parameters. For simplicity and to enable the application of efficient FFT algorithms, we will think of our signals $x[n]$ as being periodic with period N, so $\mathbb{T}^m$ is actually a cyclic time shift operator (we could write more explicitly $(\mathbb{T}^m x)[n] = x[(n-m) \bmod N]$). The second term in (1) is a nonrecursive component that corresponds to a degenerate, zero-delay TFMA$(0, L_B)$ model. We can write this latter component as $b_0[n]e[n]$ with $b_0[n] = \sum_{l=-L_B}^{L_B} b_{0,l}\,e^{j\frac{2\pi}{N}nl}$. The factor $b_0[n] \geq 0$ models a time-varying innovations variance $|b_0[n]|^2$ that cannot be modeled by the pure TFAR part.

1.2. State of the Art and Contribution

A central problem is the estimation of the TFAR parameters $a_{m,l}$ from a single realization of the nonstationary process $x[n]$. We have previously proposed the TF-Yule-Walker (TFYW) method and the underspread TFYW method for TFAR parameter estimation [1]. These methods are motivated by the Yule-Walker method for stationary AR estimation [3]. The underspread TFYW method exploits the underspread property satisfied by most practical processes to achieve a significant reduction of computational complexity.

Both for stationary AR models and for nonstationary (time-varying) AR models using a basis expansion of the time-varying AR coefficients, least-squares (LS) and maximum-likelihood (ML) methods for parameter estimation have been proposed in the literature [3, 4]. In this paper, we develop LS and ML methods for (nonstationary) TFAR parameter estimation as well as "underspread versions" of these methods. We show that the LS method is equivalent to the TFYW method and the underspread LS and ML methods are equivalent to the underspread TFYW method. We finally compare the performance of the ML estimator to the underspread TFYW method and describe a simple but practically useful method to extend TFAR spectrum estimators to signals of arbitrary length.

Our paper is organized as follows. The LS and ML methods for TFAR parameter estimation along with their underspread versions are developed in Sections 2 and 3, respectively. Simulation results are presented in Section 4.

2. LEAST-SQUARES ESTIMATORS

In this section, we develop the LS method for TFAR parameter estimation as well as an approximate LS method that is suited for underspread TFAR models. We will show that these methods can be formulated as TFYW methods.

2.1. The LS TFAR Estimator

We start by rewriting the TFAR model (1) as

$$x[n] = -\sum_{m=1}^{M_A} \boldsymbol{p}^T[n]\,\boldsymbol{\theta}_m\,x[n-m] + b_0[n]e[n]\,,$$

with the length $(2L_A+1)$ vectors $\boldsymbol{\theta}_m \triangleq \begin{bmatrix} a_{m,-L_A} \cdots a_{m,L_A} \end{bmatrix}^T$ and $\boldsymbol{p}[n] \triangleq \begin{bmatrix} e^{-j\frac{2\pi}{N}L_A n} \cdots e^{j\frac{2\pi}{N}L_A n} \end{bmatrix}^T$. A further stacking gives

$$x[n] = -\boldsymbol{s}_x^T[n]\,\boldsymbol{\theta} + b_0[n]e[n]\,,$$

with the length $M_A(2L_A + 1)$ vectors $\boldsymbol{\theta} \triangleq \begin{bmatrix} \boldsymbol{\theta}_1^T \cdots \boldsymbol{\theta}_{M_A}^T \end{bmatrix}^T$ and $\boldsymbol{s}_x[n] \triangleq \begin{bmatrix} \boldsymbol{p}^T[n]\,x[n-1] \cdots \boldsymbol{p}^T[n]\,x[n-M_A] \end{bmatrix}^T$ (note that in the definition of $\boldsymbol{s}_x[n]$ we actually use $x[(n-m)\bmod N]$). Finally, defining the length N vectors $\boldsymbol{x} \triangleq \begin{bmatrix} x[0] \cdots x[N-1] \end{bmatrix}^T$ and $\boldsymbol{u} \triangleq \begin{bmatrix} b_0[0]e[0] \cdots b_0[N-1]e[N-1] \end{bmatrix}^T$ as well as the $N \times M_A(2L_A+1)$ matrix $\boldsymbol{S}_x \triangleq \begin{bmatrix} \boldsymbol{s}_x[0] \cdots \boldsymbol{s}_x[N-1] \end{bmatrix}^T$, we obtain the following linear regression model

$$\boldsymbol{x} = -\boldsymbol{S}_x\boldsymbol{\theta} + \boldsymbol{u}\,.$$

Under the assumptions that $N \geq M_A(2L_A+1)$ and that $\boldsymbol{S}_x$ has full column-rank, the LS estimator (LSE) for $\boldsymbol{\theta}$ is given by [3, 5]

$$\hat{\boldsymbol{\theta}}_{LS} \triangleq \arg\min_{\boldsymbol{\theta}} \|\boldsymbol{x} - \boldsymbol{S}_x\boldsymbol{\theta}\|^2 = -(\boldsymbol{S}_x^H\boldsymbol{S}_x)^{-1}\boldsymbol{S}_x^H\boldsymbol{x}\,. \quad (2)$$

Because $\boldsymbol{u}$ is zero-mean, the TFAR LSE $\hat{\boldsymbol{\theta}}_{\mathrm{LS}}$ is unbiased [5]. We note that the stationary AR LSE [3] is reobtained for $L_{\mathrm{A}} = 0$.

The LS method does not yield estimates of the zero-delay TFMA parameters $b_{0,l}$, $l = -L_{\mathrm{B}}, \ldots, L_{\mathrm{B}}$. A method for estimating these parameters is discussed in [1].

2.2. Equivalence to the TFYW Estimator

We next develop a formulation of the TFAR LSE (2) in terms of the *ambiguity function*

$$\hat{A}_x[m,l] \triangleq \mathbb{F}_{n \to l}\{\hat{r}_x[n,m]\} = \sum_{n=0}^{N-1} \hat{r}_x[n,m]\, e^{-j\frac{2\pi}{N}ln}.$$

Here, $\hat{r}_x[n,m] \triangleq x[n]\,x^*[n-m]$, $n = 0, \ldots, N-1$, $m = -N/2, \ldots, N/2-1$ is an unbiased estimator of the cyclic autocorrelation $r_x[n,m] \triangleq \mathrm{E}\{x[n]\,x^*[n-m]\}$ (recall that this is short for $\mathrm{E}\{x[n]\,x^*[(n-m) \bmod N]\}$).

Defining the $M_{\mathrm{A}}(2L_{\mathrm{A}}+1) \times M_{\mathrm{A}}(2L_{\mathrm{A}}+1)$ matrix $\hat{\boldsymbol{A}}_x \triangleq \boldsymbol{S}_x^H \boldsymbol{S}_x$ and the length $M_{\mathrm{A}}(2L_{\mathrm{A}}+1)$ vector $\hat{\boldsymbol{a}}_x \triangleq \boldsymbol{S}_x^H \boldsymbol{x}$, (2) can be rewritten as

$$\hat{\boldsymbol{A}}_x \boldsymbol{\theta} = -\hat{\boldsymbol{a}}_x. \tag{3}$$

The matrix $\hat{\boldsymbol{A}}_x$ is a block matrix with matrix blocks of size $(2L_{\mathrm{A}}+1) \times (2L_{\mathrm{A}}+1)$ that are given by

$$\hat{\boldsymbol{A}}_x^{(m,m')} = \sum_{n=0}^{N-1} \boldsymbol{p}^*[n]\boldsymbol{p}^T[n]\, x[n-m']x^*[n-m],$$
$$m, m' = 1, \ldots, M_{\mathrm{A}},$$

where $\boldsymbol{p}^*[n]\boldsymbol{p}^T[n]$ is a $(2L_{\mathrm{A}}+1) \times (2L_{\mathrm{A}}+1)$ Toeplitz matrix with $\left[\boldsymbol{p}^*[n]\boldsymbol{p}^T[n]\right]_{l,l'} = e^{-j\frac{2\pi}{N}(l-l')n}$, $l, l' = -L_{\mathrm{A}}, \ldots, L_{\mathrm{A}}$. Hence, the matrix blocks $\hat{\boldsymbol{A}}_x^{(m,m')}$ are themselves Toeplitz with elements

$$\left[\hat{\boldsymbol{A}}_x^{(m,m')}\right]_{l,l'} = \sum_{n=0}^{N-1} \hat{r}_x[n-m', m-m']\, e^{-j\frac{2\pi}{N}(l-l')n}$$
$$= \hat{A}_x[m-m', l-l']\, e^{-j\frac{2\pi}{N}m'(l-l')},$$
$$l, l' = -L_{\mathrm{A}}, \ldots, L_{\mathrm{A}}. \tag{4}$$

In a similar way, it can be seen that $\hat{\boldsymbol{a}}_x = \left[\hat{\boldsymbol{a}}_x^{(1)T} \cdots \hat{\boldsymbol{a}}_x^{(M_{\mathrm{A}})T}\right]^T$ with the length $(2L_{\mathrm{A}}+1)$ vectors $\hat{\boldsymbol{a}}_x^{(m)}$ given by

$$\left[\hat{\boldsymbol{a}}_x^{(m)}\right]_l = \hat{A}_x[m,l], \qquad l = -L_{\mathrm{A}}, \ldots, L_{\mathrm{A}}. \tag{5}$$

The matrix $\hat{\boldsymbol{A}}_x$ is a "Toeplitz-block" (TB) matrix, which is a permutation of the block-Toeplitz (BT) matrix involved in the TFYW equations [1]. This permutation is due to the different stacking order (with respect to m and l) that was used in [1] to formulate the TFYW equations. Use of the (mathematically equivalent) BT stacking of [1] allows an efficient inversion of the BT matrix with complexity $\mathcal{O}(M_{\mathrm{A}}^3 L_{\mathrm{A}}^2)$ by means of the Akaike algorithm [6].

Comparing the expressions (4) and (5) to the corresponding expressions involved in the TFYW equations [1], it is readily shown that, in spite of the different stacking, the LS equation (3) is equivalent to the TFYW equations. Thus, the LSE is mathematically equivalent to the TFYW method proposed in [1]. Note that this equivalence holds in the strict sense only if the *cyclic* autocorrelation estimator $\hat{r}_x[n,m]$ is used.

2.3. The Underspread TFAR LSE

We can write $\hat{\boldsymbol{A}}_x$ as the Hadamard (elementwise) product $\hat{\boldsymbol{A}}_x = \hat{\boldsymbol{U}}_x \odot \boldsymbol{P}$, with an $M_{\mathrm{A}}(2L_{\mathrm{A}}+1) \times M_{\mathrm{A}}(2L_{\mathrm{A}}+1)$ Toeplitz/block-Toeplitz (TBT) matrix $\hat{\boldsymbol{U}}_x$ and an $M_{\mathrm{A}}(2L_{\mathrm{A}}+1) \times M_{\mathrm{A}}(2L_{\mathrm{A}}+1)$ TB phase matrix $\boldsymbol{P}$. We have $\hat{\boldsymbol{U}}_x = \mathrm{toep}\{\hat{\boldsymbol{U}}_x^{(m)}\}_{m=M_{\mathrm{A}}-1, \ldots, -M_{\mathrm{A}}+1}$

with the $(2L_{\mathrm{A}}+1) \times (2L_{\mathrm{A}}+1)$ Toeplitz blocks $[\hat{\boldsymbol{U}}_x^{(m)}]_{l,l'} = \hat{A}_x[m, l-l']$, $l, l' = -L_{\mathrm{A}}, \ldots, L_{\mathrm{A}}$. Furthermore, $\boldsymbol{P}$ has M_{A} identical matrix rows of the form $\left[\boldsymbol{P}_1 \cdots \boldsymbol{P}_{M_{\mathrm{A}}}\right]$ where $[\boldsymbol{P}_m]_{l,l'} = e^{-j\frac{2\pi}{N}m'(l-l')}$, $l, l' = -L_{\mathrm{A}}, \ldots L_{\mathrm{A}}$.

The *underspread* case is defined by $M_{\mathrm{A}}L_{\mathrm{A}} \ll N$. Here, the elements $e^{-j\frac{2\pi}{N}m'(l-l')}$ of $\boldsymbol{P}$ are approximately 1 and hence $\hat{\boldsymbol{A}}_x = \hat{\boldsymbol{U}}_x \odot \boldsymbol{P} \approx \hat{\boldsymbol{U}}_x$. It follows that the TFAR LSE $\hat{\boldsymbol{\theta}}_{\mathrm{LS}} = -\hat{\boldsymbol{A}}_x^{-1}\hat{\boldsymbol{a}}_x$ (cf. (3)) is approximated by the *underspread TFAR LSE* (ULSE)

$$\hat{\boldsymbol{\theta}}_{\mathrm{ULS}} \triangleq -\hat{\boldsymbol{U}}_x^{-1}\hat{\boldsymbol{a}}_x. \tag{6}$$

The TBT matrix $\hat{\boldsymbol{U}}_x$ can be inverted m-recursively by means of the Wax-Kailath algorithm [7], which is roughly twice as fast as the Akaike algorithm for inverting BT matrices. Thus, the underspread approximation yields a significant reduction of complexity. The m-recursive implementation of the ULSE can easily be combined with the nested order estimation procedures proposed in [2].

Comparing the ULSE in (6) with the underspread approximation to the TFYW equations described in [1], it is readily seen that the two are equivalent. Thus, the ULSE is identical to the underspread TFYM method of [1]. Finally, for $L_{\mathrm{A}} = 0$ the ULSE degenerates to the stationary AR LSE just as the LSE does.

3. MAXIMUM-LIKELIHOOD ESTIMATORS

The ML estimators developed in this section are based on an innovations system representation of the TFAR process $x[n]$. The input-output relation (1) can be expressed as $(\mathbb{A}x)[n] = (\mathbb{B}_0 e)[n]$ with the causal linear time-varying systems

$$\mathbb{A} \triangleq \sum_{m=0}^{M_{\mathrm{A}}} \sum_{l=-L_{\mathrm{A}}}^{L_{\mathrm{A}}} a_{m,l}\, \mathbb{S}_{m,l}, \qquad \mathbb{B}_0 \triangleq \sum_{l=-L_{\mathrm{B}}}^{L_{\mathrm{B}}} b_{0,l}\, \mathbb{M}^l,$$

where $a_{0,l} \triangleq \delta[l]$ (i.e., $\mathbb{A}$ is a *monic* system). Thus, the innovations system representation of the TFAR process (1) is

$$x[n] = (\mathbb{H}e)[n] \quad \text{with} \quad \mathbb{H} \triangleq \mathbb{A}^{-1}\mathbb{B}_0.$$

This is short for

$$x[n] = \sum_{m=0}^{N/2-1} h[n,m]\, e[n-m], \qquad n = 0, \ldots N-1,$$

where $h[n,m]$ is the impulse response of the innovations system $\mathbb{H}$. In matrix-vector notation, this can also be written as

$$\boldsymbol{x} = \boldsymbol{H}\boldsymbol{e}, \tag{7}$$

with length N vectors $\boldsymbol{x}$ and $\boldsymbol{e}$. Neglecting cyclic components, the $N \times N$ matrix $\boldsymbol{H}$ with elements $[\boldsymbol{H}]_{n,n'} = h[n, n-n']$ is lower triangular because $\mathbb{H}$ is a causal system. In the following, we combine the parameter vectors $\boldsymbol{\theta}$ and $\boldsymbol{b}_0 \triangleq [b_{0,-L_{\mathrm{B}}} \cdots b_{0,L_{\mathrm{B}}}]^T$ into the vector $\boldsymbol{\eta} = [\boldsymbol{\theta}^T \boldsymbol{b}_0^T]^T$, which determines $\boldsymbol{H}$.

3.1. The ML TFAR Estimator

Assuming the TFAR process $x[n]$ to be circularly symmetric complex Gaussian, the probability density function of $\boldsymbol{x}$ is given by [5]

$$p(\boldsymbol{x}; \boldsymbol{\eta}) = \frac{1}{\pi^N \det \boldsymbol{R}}\, e^{-\boldsymbol{x}^H \boldsymbol{R}^{-1} \boldsymbol{x}},$$

with the correlation (covariance) matrix $\boldsymbol{R} = \mathbb{E}\{\boldsymbol{x}\boldsymbol{x}^H\} = \boldsymbol{H}\boldsymbol{H}^H$. Our notation emphasizes that p depends on the parameter vector $\boldsymbol{\eta}$; this dependence is via $\boldsymbol{R}$. The log-likelihood function (LLF) $l(\boldsymbol{\eta}; \boldsymbol{x}) \triangleq \log p(\boldsymbol{x}; \boldsymbol{\eta})$ is, up to a constant term that is irrelevant,

$$l(\boldsymbol{\eta}; \boldsymbol{x}) = -\log \det \boldsymbol{R} - \boldsymbol{x}^H \boldsymbol{R}^{-1} \boldsymbol{x}. \tag{8}$$

Because $\mathbb{A}$ is monic, the elements on the diagonal of the lower triangular matrix $\boldsymbol{H}$ are $b_0[n]$, and thus $\log \det \boldsymbol{R} = \log \det \boldsymbol{H}\boldsymbol{H}^H = \sum_{n=0}^{N-1} \log b_0^2[n]$. Furthermore, using (7), we have $\boldsymbol{x}^H \boldsymbol{R}^{-1}\boldsymbol{x} = \boldsymbol{x}^H \boldsymbol{H}^{-H}\boldsymbol{H}^{-1}\boldsymbol{x} = \boldsymbol{e}^H \boldsymbol{e} = \|\boldsymbol{B}_0^{-1}\boldsymbol{A}\,\boldsymbol{x}\|^2$, where $\boldsymbol{A}$ and $\boldsymbol{B}_0$ are the $N \times N$ matrices corresponding to $\mathbb{A}$ and $\mathbb{B}_0$, respectively. Note that the inversion of $\boldsymbol{B}_0$ is trivial because $\boldsymbol{B}_0^{-1} = \operatorname{diag}\{\frac{1}{b_0[n]}\}_{n=0}^{N-1}$. Thus, the LLF can be expressed as

$$l(\boldsymbol{\eta};\boldsymbol{x}) = -\sum_{n=0}^{N-1} \log b_0^2[n] - \|\boldsymbol{B}_0^{-1}\boldsymbol{A}\,\boldsymbol{x}\|^2.$$

The ML estimator (MLE) $\hat{\boldsymbol{\eta}}_{\text{ML}} \triangleq \arg\max_{\boldsymbol{\eta}} l(\boldsymbol{\eta};\boldsymbol{x})$ is therefore given as

$$\hat{\boldsymbol{\eta}}_{\text{ML}} = \arg\max_{\boldsymbol{\eta}} \left\{ -\sum_{n=0}^{N-1} \log b_0^2[n] - \|\boldsymbol{B}_0^{-1}\boldsymbol{A}\,\boldsymbol{x}\|^2 \right\},$$

where $b_0[n]$ depends on the $b_{0,l}$'s as $b_0[n] = \sum_{l=-L_{\text{B}}}^{L_{\text{B}}} b_{0,l}\, e^{j\frac{2\pi}{N}nl}$.

The LLF can be maximized by any standard numerical method. The iterative maximization can be initialized by the TFAR ULSE estimator (6), i.e., $\hat{\boldsymbol{\eta}}^{(0)} \triangleq [\hat{\boldsymbol{\theta}}_{\text{ULS}}^T\ \hat{\boldsymbol{b}}_0^T]^T$, where $\hat{\boldsymbol{b}}_0$ is calculated as described in [1]. In the ith iteration, the maximization algorithm requires $l(\hat{\boldsymbol{\eta}}^{(i)};\boldsymbol{x})$, i.e., the LLF evaluated at the current parameter estimate $\hat{\boldsymbol{\eta}}^{(i)}$, and returns an improved parameter vector $\boldsymbol{\eta}^{(i+1)}$.

3.2. The Underspread TFAR MLE

A simplification of the MLE can be achieved by an underspread approximation. In what follows, we will use the time-varying transfer function (TVTF) of a linear time-varying system $\mathbb{K}$,

$$T_{\mathbb{K}}[n,k] \triangleq \mathop{\mathbb{F}}_{m\to k}\{k[n,m]\} = \sum_{m=-N/2}^{N/2-1} k[n,m]\, e^{-j\frac{2\pi}{N}km},$$

where $k[n,m]$ is the impulse response of $\mathbb{K}$.

The quadratic form $\boldsymbol{x}^H \boldsymbol{R}^{-1}\boldsymbol{x}$ occurring in the LLF can be written as an inner product in the TF domain:

$$\boldsymbol{x}^H \boldsymbol{R}^{-1}\boldsymbol{x} = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{k=-N/2}^{N/2-1} T_{\mathbb{R}^{-1}}[n,k]\, \widehat{C}_x^*[n,k], \qquad (9)$$

where $T_{\mathbb{R}^{-1}}[n,k]$ is the TVTF of the inverse of the correlation operator $\mathbb{R} = \mathbb{H}\mathbb{H}^\dagger$ ($\mathbb{H}^\dagger$ denotes the adjoint of $\mathbb{H}$) and $\widehat{C}_x[n,k]$ is the Rihaczek distribution [8],

$$\widehat{C}_x[n,k] \triangleq \mathop{\mathbb{F}}_{m\to k}\{\hat{r}_x[n,m]\} = \mathop{\mathbb{F}}_{m\to k}\mathop{\mathbb{F}}_{l\to n}^{-1}\{\hat{A}_x[m,l]\}.$$

For an underspread system $\mathbb{H}$, the TVTF of $\mathbb{R}^{-1} = (\mathbb{H}\mathbb{H}^\dagger)^{-1} = \mathbb{H}^{-\dagger}\mathbb{H}^{-1}$ can be approximated as $T_{\mathbb{R}^{-1}}[n,k] = T_{\mathbb{H}^{-\dagger}\mathbb{H}^{-1}}[n,k] \approx |T_{\mathbb{H}^{-1}}[n,k]|^2$ [9]. It can furthermore be shown that the TVTF of $\mathbb{H}^{-1}$ can be expressed (exactly) as

$$T_{\mathbb{H}^{-1}}[n,k] = \frac{T_{\mathbb{A}}[n,k]}{T_{\mathbb{B}_0}[n]}, \qquad (10)$$

where

$$T_{\mathbb{A}}[n,k] = N \mathop{\mathbb{F}}_{m\to k}\mathop{\mathbb{F}}_{l\to n}^{-1}\{a_{m,l}\}, \qquad T_{\mathbb{B}_0}[n] = b_0[n]. \qquad (11)$$

With (10) and (11), our underspread approximation for $T_{\mathbb{R}^{-1}}[n,k]$ becomes $T_{\mathbb{R}^{-1}}[n,k] \approx \left|\frac{T_{\mathbb{A}}[n,k]}{b_0[n]}\right|^2$, and hence (9) gives

$$\boldsymbol{x}^H \boldsymbol{R}^{-1}\boldsymbol{x} \approx \frac{1}{N} \sum_{n=0}^{N-1} \sum_{k=-N/2}^{N/2-1} \frac{|T_{\mathbb{A}}[n,k]|^2 \widehat{C}_x^*[n,k]}{b_0^2[n]}. \qquad (12)$$

Inserting (12) as well as the identity $\log \det \boldsymbol{R} = \sum_{n=0}^{N-1} \log b_0^2[n]$ into (8) shows that $l(\boldsymbol{\eta};\boldsymbol{x})$ can be approximated by

$$\tilde{l}(\boldsymbol{\eta};\boldsymbol{x}) \triangleq -\sum_{n=0}^{N-1} \left[\log b_0^2[n] + \frac{\sum_{k=-N/2}^{N/2-1} |T_{\mathbb{A}}[n,k]|^2 \widehat{C}_x^*[n,k]}{Nb_0^2[n]} \right].$$

Maximization of $\tilde{l}(\boldsymbol{\eta};\boldsymbol{x})$ can only be done numerically. To simplify the problem, we assume that the TFAR innovations process $b_0[n]e[n]$ is *stationary* with a given variance σ_e^2. This corresponds to $b_0[n] \equiv \sigma_e$ or equivalently $b_{0,l} = \sigma_e\delta[l]$. The approximate LLF then reduces to

$$\tilde{l}(\boldsymbol{\theta},\sigma_e^2;\boldsymbol{x}) = -N \log \sigma_e^2 - \frac{1}{N\sigma_e^2} \sum_{n=0}^{N-1} \sum_{k=-N/2}^{N/2-1} |T_{\mathbb{A}}[n,k]|^2 \widehat{C}_x^*[n,k]. \qquad (13)$$

Setting the derivative of (13) with respect to σ_e^2 equal to zero yields the innovations variance

$$\widehat{\sigma_e^2} = \frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{k=-N/2}^{N/2-1} |T_{\mathbb{A}}[n,k]|^2 \widehat{C}_x^*[n,k]. \qquad (14)$$

This still depends on the unknown $a_{m,l}$'s. Inserting (14) into (13), we obtain the partially optimized approximate LLF

$$\tilde{l}(\boldsymbol{\theta};\boldsymbol{x}) = -N \log\left(\frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{k=-N/2}^{N/2-1} |T_{\mathbb{A}}[n,k]|^2 \widehat{C}_x^*[n,k]\right) - N.$$

The "underspread MLE" (UMLE) for $\boldsymbol{\theta}$ is hence obtained as

$$\hat{\boldsymbol{\theta}}_{\text{UML}} \triangleq \arg\max_{\boldsymbol{\theta}} \tilde{l}(\boldsymbol{\theta};\boldsymbol{x})$$
$$= \arg\min_{\boldsymbol{\theta}} \left\{ \sum_{n=0}^{N-1} \sum_{k=-N/2}^{N/2-1} |T_{\mathbb{A}}[n,k]|^2 \widehat{C}_x^*[n,k] \right\}.$$

We thus differentiate $\sum_{n=0}^{N-1} \sum_{k=-N/2}^{N/2-1} |T_{\mathbb{A}}[n,k]|^2 \widehat{C}_x^*[n,k]$ with respect to a_{m_0,l_0} and equate the result to zero:

$$\sum_{n=0}^{N-1} \sum_{k=-N/2}^{N/2-1} \widehat{C}_x^*[n,k] \frac{\partial(T_{\mathbb{A}}[n,k]\,T_{\mathbb{A}}^*[n,k])}{\partial a_{m_0,l_0}}$$
$$= \sum_{n=0}^{N-1} \sum_{k=-N/2}^{N/2-1} \widehat{C}_x^*[n,k]\, T_{\mathbb{A}}^*[n,k] \frac{\partial T_{\mathbb{A}}[n,k]}{\partial a_{m_0,l_0}}$$
$$= \sum_{n=0}^{N-1} \sum_{k=-N/2}^{N/2-1} \widehat{C}_x^*[n,k]\, T_{\mathbb{A}}^*[n,k]\, e^{-j\frac{2\pi}{N}m_0 k} e^{j\frac{2\pi}{N}nl_0}$$
$$= \left(\mathop{\mathbb{F}}_{n\to l_0}\mathop{\mathbb{F}}_{k\to m_0}^{-1} \widehat{C}_x[n,k]\, T_{\mathbb{A}}[n,k] \right)^* = 0, \qquad (15)$$

where we have used (11). Because multiplication in the (n,k)-domain corresponds to 2-D convolution in the (m,l)-domain, (15) is equivalent to the following equation for the UMLE $\hat{\boldsymbol{\theta}}_{\text{UML}}$:

$$\sum_{m=0}^{M_{\text{A}}} \sum_{l=-L_{\text{A}}}^{L_{\text{A}}} \hat{A}_x[m_0-m,l_0-l]\, \hat{a}_{m,l} = 0,$$
$$m_0 = 1,\ldots,M_{\text{A}},\ l_0 = -L_{\text{A}},\ldots,L_{\text{A}}.$$

This equation is the underspread TFYW equation of [1] and equivalent to the equation (6) defining the ULSE. Hence, the UMLE is identical to the ULSE and also identical to the underspread TFYW method [1].
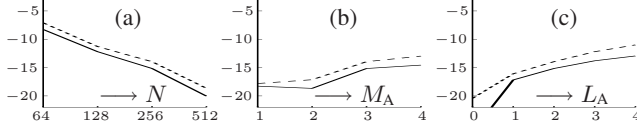
Figure 1: *Normalized MSE (in dB) of the MLE (solid line) and the underspread TFYW estimator (dashed line) versus (a) block length N, (b) delay order M_A, (c) Doppler order L_A.*

4. SIMULATION RESULTS

4.1. Performance of MLE and Underspread TFYW Estimator

We compare the performance of the exact MLE discussed in Section 3.1 and the underspread TFYW estimator (equivalently, the ULSE of Section 2.3 and the UMLE of Section 3.2) by means of simulations. We considered a TFAR$(3, 2)$ process with $b_{0,l} = \sigma_e \delta[l]$ (i.e., stationary innovations process, $L_B = 0$) and length $N = 256$. The 15 complex-valued TFAR parameters $a_{m,l}$ and the positive innovations variance σ_e^2 were estimated from a single process realization $x[n]$. This estimation was repeated for 100 realizations of the process and the empirical mean-square error (MSE) of each estimator was obtained by averaging over all parameters and all realizations. We then repeated the experiment for other processes constructed by varying N or M_A or L_A. The iterative maximization required for calculating the MLE was performed by means of the MATLAB function `fmincon.m`.

Figure 1 shows the normalized MSEs. It is seen that the MLE outperforms the underspread TFYW estimator by (typically) about 1–2 dB. The MSE generally increases for larger model orders M_A, L_A and for smaller block length N.

4.2. TFAR Spectral Analysis of Arbitrarily Long Signals

The TFAR model can be used for time-varying spectral analysis. The TFAR spectrum estimate of a finite-length signal $x[n]$, $n = 0, \ldots, N-1$ is defined as (cf. (10))

$$S_x[n,k] \triangleq \left| T_{\hat{\mathbb{H}}}[n,k] \right|^2 \approx \left| \frac{T_{\hat{\mathbb{B}}_0}[n]}{T_{\hat{\mathbb{A}}}[n,k]} \right|^2,$$

where $T_{\hat{\mathbb{A}}}[n,k] = N \, \mathbb{F}_{m \to k} \mathbb{F}^{-1}_{l \to n}\{\hat{a}_{m,l}\}$ and $T_{\hat{\mathbb{B}}_0}[n] = \hat{b}_0[n]$ (cf. (11)) are estimates of the TVTFs $T_{\mathbb{A}}[n,k]$ and $T_{\mathbb{B}_0}[n]$, respectively.

The finite block length, cyclic formulation used in the definition of the TFAR model enables application of efficient FFT techniques. Practical complexity considerations imply a restriction of the signal length N. Signals of arbitrary length can be analyzed by applying our TFAR estimators to successive signal blocks of length N and concatenating the resulting time-varying TFAR spectra. Discontinuities and aliasing artifacts (caused by the cyclic formulation) at the block boundaries can be reduced by using overlapping blocks and windowing, as described in what follows.

The overall signal is segmented into blocks of length N with N_o samples overlap. Each signal segment is multiplied by a raised cosine window of duration N with N_o samples rolloff (this reduces the amount of aliasing artifacts and artificial nonstationarities at the block boundaries). The concatenation of the TFAR spectrum estimates obtained for the individual segments then uses an overlap-add approach. Here, each spectrum estimate is multiplied (with respect to n) by the same raised-cosine window as before in order to produce smooth transitions within the overlap intervals.

Figure 2 shows results of TFAR spectral analysis for a "long" signal (length 1024) consisting of three FM components and white noise (SNR $= 10\,$dB). Note that this signal—even without the noise—does not conform to the TFAR model. The block length was $N = 256$ and the overlap length N_o was chosen as 0, 15, and 31. The TFAR orders M_A, L_A, L_B were determined by the MDL order-estimation criterion described in [2]. It is seen that the overlapping-windowing technique significantly reduces discontinuities at the block boundaries.
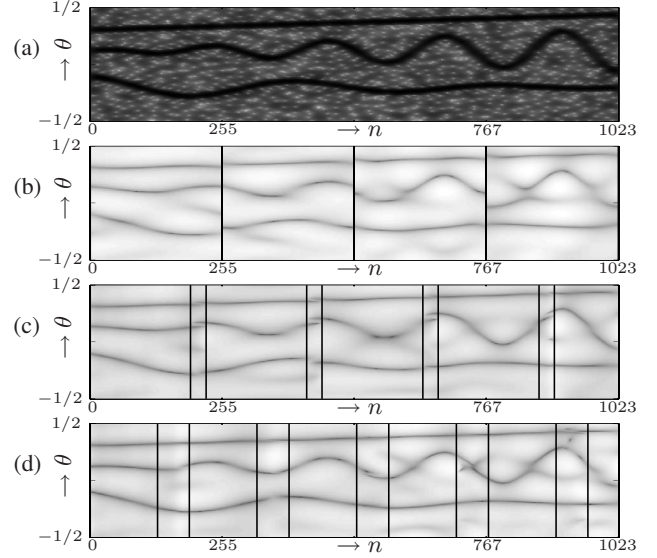

Figure 2: *Results of overlapped/windowed TFAR spectral analysis: (a) Spectrogram of the noisy signal, (b)–(d) estimated TFAR spectra for (b) no overlap, (c) 15 samples overlap, (d) 31 samples overlap. (Horizontal axis is time, vertical axis is normalized frequency, vertical lines indicate overlap intervals.)*

5. CONCLUSIONS

The least-squares (LS) and maximum-likelihood (ML) TFAR parameter estimators derived in this paper extend the known LS and ML techniques for stationary AR estimation to the nonstationary case. A reduction of computational complexity can be achieved by exploiting the underspread property satisfied by most practical processes. We showed that the LS, underspread LS, and underspread ML estimators are equivalent to estimators based on the time-frequency Yule-Walker (TFYW) equations, whereas the exact ML estimator can yield a performance improvement over TFYW techniques.

REFERENCES

[1] M. Jachan, G. Matz, and F. Hlawatsch, "Time-frequency-autoregressive random processes: Modeling and fast parameter estimation," in *Proc. IEEE ICASSP-2003*, vol. VI, (Hong Kong), pp. 125–128, April 2003.

[2] M. Jachan, G. Matz, and F. Hlawatsch, "TFARMA models: Order estimation and stabilization," in *Proc. IEEE ICASSP-2005*, (Philadelphia, PA), pp. 301–304, March 2005.

[3] S. M. Kay, *Modern Spectral Estimation*. Englewood Cliffs (NJ): Prentice Hall, 1988.

[4] Y. Grenier, "Time-dependent ARMA modeling of nonstationary signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, pp. 899–911, Aug. 1983.

[5] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs (NJ): Prentice Hall, 1993.

[6] H. Akaike, "Block Toeplitz matrix inversion," *SIAM J. Appl. Math.*, vol. 24, pp. 234–241, March 1973.

[7] M. Wax and T. Kailath, "Efficient inversion of Toeplitz-block Toeplitz matrix," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, pp. 1218–1221, Oct. 1983.

[8] P. Flandrin, *Time-Frequency/Time-Scale Analysis*. San Diego (CA): Academic Press, 1999.

[9] G. Matz and F. Hlawatsch, "Time-frequency transfer function calculus (symbolic calculus) of linear time-varying systems (linear operators) based on a generalized underspread theory," *J. Math. Phys., Special Issue on Wavelet and Time-Frequency Analysis*, vol. 39, pp. 4041–4071, Aug. 1998.