The Extended Least Squares Criterion: Minimization Algorithms and Applications

Arie Yeredor, *Member, IEEE*

***Abstract*—The least squares (LS) estimation criterion on one hand, and the total LS (TLS), constrained TLS (CTLS) and structured TLS (STLS) criteria on the other hand, can be viewed as opposite limiting cases of a more general criterion, which we term "Extended LS" (XLS). The XLS criterion distinguishes measurement errors from modeling errors by properly weighting and balancing the two error sources. In the context of certain models (termed "pseudo-linear"), we derive two iterative algorithms for minimizing the XLS criterion: One is a straightforward "alternating coordinates" minimization, and the other is an extension of an existing CTLS algorithm. The algorithms exhibit different tradeoffs between convergence rate, computational load, and accuracy. The XLS criterion can be applied to popular estimation problems, such as identifying an autoregressive (AR) with exogenous noise (ARX) system from noisy input/output measurements or estimating the parameters of an AR process from noisy measurements. We demonstrate the convergence properties and performance of the algorithms with examples of the latter.**

***Index Terms*—AR modeling, ARX system identification, constrained total least squares (CTLS), extended least squares (XLS), least squares, structured total least squares (STLS), total least squares (TLS).**

I. Introduction

IN A VAST variety of problems in engineering, it is desired to estimate unknown parameters $\boldsymbol{\theta}$ from data measurements $\boldsymbol{x}$. The parameters and the data are generally related by an approximate set of model equations

$$\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{\theta}) \approx \boldsymbol{0}. \tag{1}$$

The inequality in (1) may often be attributed to two possibly distinct mechanisms: model mismatch and measurement inaccuracies. Model mismatch encompasses inaccuracies that exist in (1) even when exact measurements are used. Measurement inaccuracies, on the other hand, account for possible deviations of the measured data $\boldsymbol{x}$ from the true (unknown) data, say $\tilde{\boldsymbol{x}}$, with which (1) is exact (in the absence of model mismatch). In other words, (1) may be broken down into two distinct (in) equalities:

$$\boldsymbol{g}(\tilde{\boldsymbol{x}}, \boldsymbol{\theta}) \approx \boldsymbol{0} \tag{2a}$$
$$\boldsymbol{x} - \tilde{\boldsymbol{x}} \approx \boldsymbol{0} \tag{2b}$$

where $\tilde{\boldsymbol{x}}$ is a vector of "presumed" ("accurate" or "latent") data so that (2a) accounts only for model mismatch, whereas (2b) accounts only for measurement inaccuracies. While measurement inaccuracies can always be attributed to model errors, they are often caused by unrelated mechanisms. In such cases, the distinction in (2a) and (2b) is justified.

The well-known least squares (LS) estimation approach seeks parameters $\boldsymbol{\theta}$, which, together with the given measurements $\boldsymbol{x}$, minimize the (possibly weighted) Euclidean norm of $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{\theta})$:

$$\min_{\boldsymbol{\theta}} \{\boldsymbol{g}^T(\boldsymbol{x}, \boldsymbol{\theta})\boldsymbol{W}\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{\theta})\} \Rightarrow \hat{\boldsymbol{\theta}}_{\text{LS}} \tag{3}$$

where $\boldsymbol{W}$ is some symmetric positive-definite weight matrix. $\hat{\boldsymbol{\theta}}_{\text{LS}}$ is actually the value that brings $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{\theta})$ as close as possible to its presumed value of $\boldsymbol{0}$. However, no substitution of the observed data is implied, and thus, (2b) is completely satisfied (with equality).

On the other hand, some well-known existing modifications of LS can often be regarded as taking the opposite approach. Specifically, they attempt to find a vector of "presumed" data $\tilde{\boldsymbol{x}}$ and an associated parameters vector $\boldsymbol{\theta}$ with which the model equation (2a) is satisfied (with equality) while keeping to a minimum the deviation of the "presumed" data $\tilde{\boldsymbol{x}}$ from the observed data $\boldsymbol{x}$.

To illustrate that, consider the case where the model equation is given by $\boldsymbol{x} \approx \boldsymbol{X}\boldsymbol{\theta}$ with measurements contained in both $\boldsymbol{x}$ and $\boldsymbol{X}$. In the notation of (1), such a model is specified by

$$\boldsymbol{x} - \boldsymbol{X}\boldsymbol{\theta} \approx \boldsymbol{0}. \tag{4}$$

The linear LS solution $\hat{\boldsymbol{\theta}}_{\text{LS}} = (\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{x}$ yields the value of $\boldsymbol{\theta}$ with which (4) is most closely satisfied without change in the measurements $\boldsymbol{x}$, $\boldsymbol{X}$. However, if perturbations in $\boldsymbol{X}$ as well as in $\boldsymbol{x}$ are allowed, the estimation approach assumes the form of a total least squares (TLS) problem (see [1]–[3] and references therein). In various engineering applications, structural constraints (e.g., Hankel, Toeplitz, etc.) are imposed on the augmented matrix $[\boldsymbol{x} \vdots \boldsymbol{X}]$. Since the basic TLS solution is incapable of addressing structural constraints, various modifications thereof were proposed and termed "constrained" (or "structured") TLS, in various application-specific contexts (e.g., [4]–[9] in addition to earlier works such as Steiglitz–McBride's [10] or IQML [11] methods, which have been shown to address an equivalent problem [12], [13]), as well as in more general forms, such as Cadzow's constrained TLS (CTLS) [14]–[16] and De Moor's structured TLS (STLS) [9], [17], [18].

However, as illustrated above, the use of TLS-based methods bears an implicit assumption that the model is exact. For example, in estimating the parameters of superimposed exponential signals (or, equivalently, the impulse response of a linear

Manuscript received June 8, 1999; revised September 22, 2000. The associate editor coordinating the review of this paper and approving it for publication was Prof. Philippe Loubaton.

The author is with the Department of Electrical Engineering—Systems, Tel-Aviv University, Tel-Aviv, Israel (e-mail: arie@eng.tau.ac.il).

Publisher Item Identifier S 1053-587X(01)00074-5.

1053–587X/00$10.00 © 2000 IEEE

system such as in [9]), it is assumed that the underlying noiseless signal satisfies an exact linear prediction equation. In the problem of identifying the parameters of a linear time-invariant (LTI) system from noisy input/output observations [which is also known as the "errors-in-variable" (EIV) model], it is assumed that the underlying (noiseless) input and output are related precisely by the system's model equation. The only implied errors are in the available measurements.

These assumptions are often either unjustified or can be removed deliberately to broaden the scope onto a more general problem. For example, if the composite noiseless exponential signal does not satisfy its respective linear prediction equation, it can be regarded as a stochastic auto-regressive (AR) process. If the LTI system does not satisfy its difference equation precisely, it can be regarded as an AR with exogenous disturbance (ARX) model (see, e.g., [19]). When ordinary LS or TLS-based methods are applied to these models, the resulting estimates become highly erroneous due to the misfit of these estimation approaches to the problems in hand.

The more comprehensive extended least squares (XLS) criterion proposed in this paper accounts for both error sources and discriminates data measurement errors from model errors. As such, it can address these problems as well and, furthermore, introduce proper balancing of the two error sources.

The possibility of accounting for model errors separately from measurement errors has been addressed by others, either explicitly, such as by Fuller [20] in the context of EIV models, which is termed "EIV models with an error in the equation," or implicitly in application-specific contexts, such as by Lim and Oppenheim [21] in estimating the poles of noisy speech. More recently, De Moor and Lemmerling proposed a similar "misfit versus latency" approach [22] in the context of linear systems identification. However, frameworks for the general problem formulation and efficient minimization algorithms have yet to be established. It is the purpose of this paper to fill this void.

We note in passing that like LS and TLS, the XLS criterion is purely based on deterministic considerations. However, LS and TLS often have statistical interpretations. For example, when the model errors (or the measurement errors) are assumed zero-mean Gaussian, LS (or STLS, respectively) is known to have the statistical interpretation of the maximum-likelihood (ML) criterion. The XLS criterion also has an interesting statistical interpretation that will not be pursued here. See [23]–[25].

This paper is organized as follows: In Section II, we formulate the XLS criterion and identify existing criteria as limiting cases thereof. In Section III, we define a family of "pseudo-linear" models with which efficient iterative algorithms for XLS minimization can be derived. Such algorithms are developed and outlined in Section IV and then illustrated in application examples in Section V. Our concluding remarks comprise Section VI.

II. Extended Least Squares (XLS) Criterion

Forming the concatenation of (2a) and (2b), we get

$$\tilde{\boldsymbol{g}}(\boldsymbol{x}, \tilde{\boldsymbol{x}}, \boldsymbol{\theta}) \triangleq \begin{bmatrix} \boldsymbol{g}(\tilde{\boldsymbol{x}}, \boldsymbol{\theta}) \\ \boldsymbol{x} - \tilde{\boldsymbol{x}} \end{bmatrix} \approx \boldsymbol{0}. \tag{5}$$

The XLS estimate of $\boldsymbol{\theta}$ is obtained by applying the LS criterion to $\tilde{\boldsymbol{g}}$. However, since $\tilde{\boldsymbol{x}}$ is obviously unknown, we have to minimize with respect to $\tilde{\boldsymbol{x}}$ as well, obtaining as a byproduct the XLS estimate of the presumed data

$$\min_{\tilde{\boldsymbol{x}}, \boldsymbol{\theta}} \left\{ \tilde{\boldsymbol{g}}^T(\boldsymbol{x}, \tilde{\boldsymbol{x}}, \boldsymbol{\theta}) \tilde{\boldsymbol{W}} \tilde{\boldsymbol{g}}(\boldsymbol{x}, \tilde{\boldsymbol{x}}, \boldsymbol{\theta}) \right\} \Rightarrow \hat{\boldsymbol{\theta}}_{\text{XLS}} \left(+\hat{\tilde{\boldsymbol{x}}}_{\text{XLS}} \right) \tag{6}$$

where $\tilde{\boldsymbol{W}}$ is the extended weight matrix. Often, a block-diagonal $\tilde{\boldsymbol{W}}$ would be chosen as

$$\tilde{\boldsymbol{W}} = \begin{bmatrix} \boldsymbol{W}_g & 0 \\ 0 & \boldsymbol{W}_x \end{bmatrix} \tag{7}$$

where $\boldsymbol{W}_g$ and $\boldsymbol{W}_x$ fit the dimensions of $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{\theta})$ and $\boldsymbol{x}$, respectively, so that the XLS minimization (6) may assume the following form:

$$\begin{aligned} &\min_{\tilde{\boldsymbol{x}}, \boldsymbol{\theta}} \left\{ \boldsymbol{g}^T(\tilde{\boldsymbol{x}}, \boldsymbol{\theta}) \boldsymbol{W}_g \boldsymbol{g}(\tilde{\boldsymbol{x}}, \boldsymbol{\theta}) + (\boldsymbol{x} - \tilde{\boldsymbol{x}})^T \boldsymbol{W}_x (\boldsymbol{x} - \tilde{\boldsymbol{x}}) \right\} \\ &\quad \Rightarrow \hat{\boldsymbol{\theta}}_{\text{XLS}} \left(+\hat{\tilde{\boldsymbol{x}}}_{\text{XLS}} \right). \end{aligned} \tag{8}$$

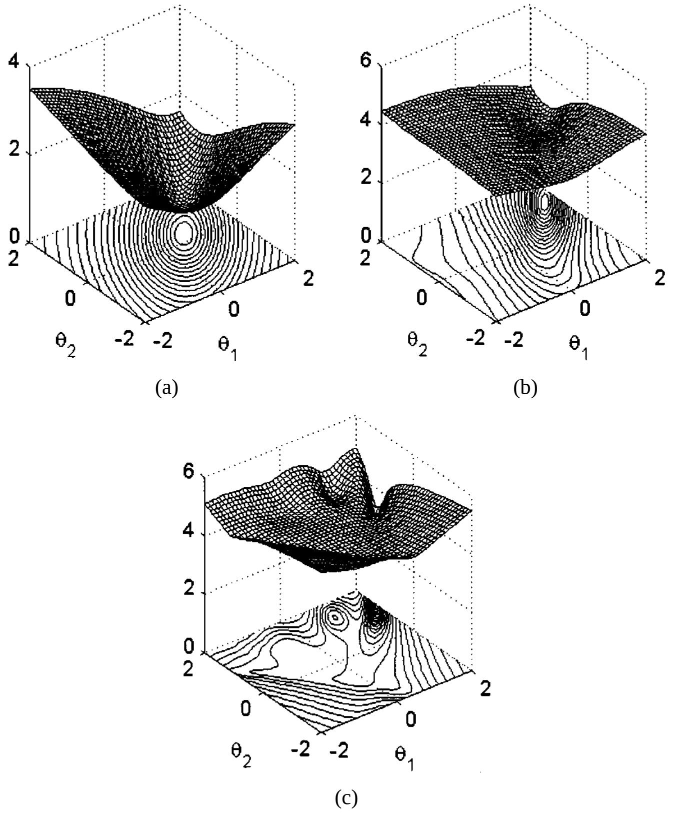
When $\boldsymbol{W}_g \ll \boldsymbol{W}_x$,[1] the minimization with respect to $\tilde{\boldsymbol{x}}$ is dominated by the second term and is attained near $\tilde{\boldsymbol{x}} \approx \boldsymbol{x}$ so that $\hat{\boldsymbol{\theta}}_{\text{XLS}}$ practically minimizes the first term with $\tilde{\boldsymbol{x}}$ replaced by $\boldsymbol{x}$. Obviously, this coincides with the LS estimate, which is thus identified as a limiting case that "blames" all the inconsistency in (1) on model mismatch.

When, on the other hand, $\boldsymbol{W}_x \ll \boldsymbol{W}_g$, minimization is attained when the first term is nearly zeroed out. Thus, in the second term, the minimal perturbation of the measured data is sought, for which the model equations can be completely satisfied with some value of $\boldsymbol{\theta}$. This approach is, in a sense, the opposite of LS, as it blames all the inconsistency on measurement inaccuracies.

The selection of the weight matrices often also affects the complexity of the minimization problem, as well as the existence of local minima. When $\boldsymbol{W}_g \ll \boldsymbol{W}_x$, the problem is more "LS-like" and is therefore "easier" and usually (in the context of "pseudo-linear models" to be defined immediately) has a unique global minimum. As $\boldsymbol{W}_g$ increases, the problem becomes more "STLS-like," and local minima begin to emerge. This behavior is demonstrated in Fig. 1. The model used is an AR(2) process in noise. The (log) criterion surface is drawn as a function of the two free parameters $\boldsymbol{\theta} = [\theta_1 \;\; \theta_2]^T$ (following further minimization with respect to $\tilde{\boldsymbol{x}}$ at each $\boldsymbol{\theta}$). The weight matrices used were $\boldsymbol{W}_x = \boldsymbol{I}$ and $\boldsymbol{W}_g = w_g \boldsymbol{I}$ (where $\boldsymbol{I}$ denotes the identity matrix) with $w_g = 0.2, 1, 5$ for Fig. 1(a)–(c), respectively (all three figures were calculated using the same data $\boldsymbol{x}$ generated with $\theta_1 = 0.2$, $\theta_2 = -0.15$). It is seen that with $w_g = 0.2$, the surface is unimodal, smooth, and nearly parabolic, whereas an increase in w_g "wrinkles" the surface and creates local minima.

Careful selection of $\boldsymbol{W}_g$ and $\boldsymbol{W}_x$ would normally reflect the optimal sharing of inconsistency between model mismatch and measurement errors. When statistical assumptions as to the nature of the model mismatch and measurement noises are incorporated, specific selection of weights reflects specific statistical

[1] We use "$\ll$" loosely to compare matrices of possibly unequal dimensions in the sense of comparing eigenvalues. All the eigenvalues of the smaller matrix are much smaller than all the eigenvalues of the other.



(a)

(b)

(c)

Fig. 1. Log of the criterion surface for an AR(2) process in noise. The surface is drawn as a function of the two parameters, following further minimization with respect to the "latent" data $\tilde{\boldsymbol{x}}$ at each point. The figures demonstrate the effect of the relative weights attributed to the model ($\boldsymbol{W}_g = w_g \boldsymbol{I}$) and to the data ($\boldsymbol{W}_x = \boldsymbol{I}$). As w_g increases, the surface "wrinkles," moving from a parabolic "LS-like" surface to a multimodal "STLS-like" surface. All three figures were drawn using the same data $\boldsymbol{x}$. $\boldsymbol{W}_g$ =0.2, 1, and 5 in Fig. 2(a)-(c), respectively.

interpretations for the obtained estimate. These interpretations, as well as a statistical analysis of performance, can be found in [23]–[25].

III. Pseudo-Linear Model

The presentation of the XLS criterion above did not restrict the structure of the model functions $\boldsymbol{g}(\tilde{\boldsymbol{x}}, \boldsymbol{\theta})$. In order to pursue computationally appealing algorithms for the required minimization, we now restrict our discussion to pseudo-linear models, which are applicable in a variety of engineering problems such as ARX system identification or AR parameter estimation.

By the term "pseudo-linear models," in this context, we refer to model functions $\boldsymbol{g}(\tilde{\boldsymbol{x}}, \boldsymbol{\theta})$ that are linear in the data $\tilde{\boldsymbol{x}}$ for any fixed value of the parameters $\boldsymbol{\theta}$, and vice versa. It is easy to observe that $\boldsymbol{g}(\tilde{\boldsymbol{x}}, \boldsymbol{\theta})$ complies with that definition iff it can be expressed in the form

$$\boldsymbol{g}(\tilde{\boldsymbol{x}}, \boldsymbol{\theta}) = \boldsymbol{A}(\tilde{\boldsymbol{x}})\boldsymbol{v}(\boldsymbol{\theta}) \tag{9}$$

where

$$\boldsymbol{A}(\tilde{\boldsymbol{x}}) = \boldsymbol{A}_0 + \sum_{n=1}^{N} \tilde{x}_n \boldsymbol{A}_n \tag{10a}$$

where $\boldsymbol{A}_0, \boldsymbol{A}_1, \cdots \boldsymbol{A}_N$ is a set of constant matrices, $\tilde{x}_n$ are the elements of $\tilde{\boldsymbol{x}}$, and

$$\boldsymbol{v}(\boldsymbol{\theta}) = \begin{bmatrix} v_0 \\ \boldsymbol{\theta} \end{bmatrix} \tag{10b}$$

where v_0 is some constant. The role of the free matrix $\boldsymbol{A}_0$ and the free constant v_0 is to provide optional terms that are purely linear in $\boldsymbol{\theta}$ and $\tilde{\boldsymbol{x}}$, respectively. In the absence of such terms in the model, $\boldsymbol{A}_0$ or v_0 are set to zero. Without loss of generality, we may also set $v_0 = 1$ when v_0 is nonzero. In that way, all the scaling is taken care of in the $\boldsymbol{A}(\tilde{\boldsymbol{x}})$ term. Note that if $v_0 = 0$, the model equations can be made exact with a trivial choice $\boldsymbol{\theta} = \boldsymbol{0}$, which renders such cases uninteresting, as the XLS criterion (6) would be minimized (zeroed out) by setting in addition $\tilde{\boldsymbol{x}} = \boldsymbol{x}$. Since we are only interested in interesting cases, we assume $v_0 = 1$.

IV. Algorithms for Minimization of the XLS Criterion

Despite the restriction of the general model to be pseudo-linear, no closed-form expression for minimization is currently known. In this section, we develop two iterative algorithms for minimizing the XLS criterion under the assumption of a "pseudo-linear model" (9).

We assume the simplified form (8) of the XLS minimization criterion. The cost function under minimization $C(\tilde{\boldsymbol{x}}, \boldsymbol{\theta})$ is therefore given by

$$\begin{aligned} C(\tilde{\boldsymbol{x}}, \boldsymbol{\theta}) &= \boldsymbol{g}^T(\tilde{\boldsymbol{x}}, \boldsymbol{\theta})\boldsymbol{W}_g \boldsymbol{g}(\tilde{\boldsymbol{x}}, \boldsymbol{\theta}) + (\boldsymbol{x} - \tilde{\boldsymbol{x}})^T \boldsymbol{W}_x (\boldsymbol{x} - \tilde{\boldsymbol{x}}) \\ &= \begin{bmatrix} 1 \; \boldsymbol{\theta}^T \end{bmatrix} \boldsymbol{A}^T(\tilde{\boldsymbol{x}}) \boldsymbol{W}_g \boldsymbol{A}(\tilde{\boldsymbol{x}}) \begin{bmatrix} 1 \\ \boldsymbol{\theta} \end{bmatrix} + (\boldsymbol{x} - \tilde{\boldsymbol{x}})^T \boldsymbol{W}_x (\boldsymbol{x} - \tilde{\boldsymbol{x}}). \end{aligned} \tag{11}$$

Thus, for pseudo-linear models, $C(\tilde{\boldsymbol{x}}, \boldsymbol{\theta})$ is quadratic in $\tilde{\boldsymbol{x}}$ for any fixed $\boldsymbol{\theta}$, and vice versa. The first algorithm exploits that fact by alternating between minimization with respect to $\tilde{\boldsymbol{x}}$ given a previous value of $\boldsymbol{\theta}$ and minimization with respect to $\boldsymbol{\theta}$, given the previous value of $\tilde{\boldsymbol{x}}$.

While analytically appealing and plain in concept, the alternating coordinates minimization requires $O(N^2)$ or even $O(N^3)$ multiply-add operations per iteration. We therefore propose another minimization algorithm that requires $O(N)$ operations per iteration. It dwells on the close relation of XLS minimization to an existing algorithm mentioned earlier, namely, Cadzow's CTLS ([14]–[16]), whose optimization problem was identified as a limiting case of our more general XLS problem. Our proposed algorithm is thus termed extended CTLS (ECTLS). The computational relief is attained at the cost of a drawback inherited by ECTLS from CTLS in the form of possible off-minimum convergence. We elaborate on that drawback in the sequel.

Throughout this presentation, we will assume that $\boldsymbol{A}(\tilde{\boldsymbol{x}})$ is $p \times (q+1)$ $(p > q)$ so that $\boldsymbol{\theta}$ is $q \times 1$. The data vectors $\tilde{\boldsymbol{x}}$ and $\boldsymbol{x}$ are assumed $N \times 1$.

A. Alternating Coordinates Minimization (ACM) Algorithm

The alternating coordinates minimization (ACM) strategy alternates between minimization with respect to the "coordinates" $\tilde{\boldsymbol{x}}$ and $\boldsymbol{\theta}$.

The quadratic minimization always results in a unique global minimum (assuming the other coordinates fixed) so that in each iteration the value of $C(\tilde{\boldsymbol{x}}, \boldsymbol{\theta})$ is guaranteed not to increase (usually to decrease). Since $C(\tilde{\boldsymbol{x}}, \boldsymbol{\theta})$ is bounded below (e.g., by zero), convergence is guaranteed. Furthermore, the convergence point is guaranteed to be a (local) minimum (or, at most, a saddle-point, being minimum with respect to $\tilde{\boldsymbol{x}}$ and $\boldsymbol{\theta}$ separately but not jointly) since these are the only possible stationary points of the alternating minimization operation.

We will now derive explicit expressions for the minimization in each phase. Given $\tilde{\boldsymbol{x}}$, the quadratic expression in $\boldsymbol{\theta}$ is readily available as the first term in (11), which is uniquely minimized by

$$\hat{\boldsymbol{\theta}}(\tilde{\boldsymbol{x}}) = -\boldsymbol{B}^{-1}(\tilde{\boldsymbol{x}})\boldsymbol{b}(\tilde{\boldsymbol{x}}) \tag{12}$$

where $\boldsymbol{B}(\tilde{\boldsymbol{x}})$ $(q \times q)$ and $\boldsymbol{b}(\tilde{\boldsymbol{x}})$ $(q \times 1)$ are defined via the following partition of the $(q+1) \times (q+1)$ matrix $\boldsymbol{A}^T(\tilde{\boldsymbol{x}})\boldsymbol{W}_g\boldsymbol{A}(\tilde{\boldsymbol{x}})$:

$$\boldsymbol{A}^T(\tilde{\boldsymbol{x}})\boldsymbol{W}_g\boldsymbol{A}(\tilde{\boldsymbol{x}}) = \begin{bmatrix} b(\tilde{\boldsymbol{x}}) & \boldsymbol{b}^T(\tilde{\boldsymbol{x}}) \\ \boldsymbol{b}(\tilde{\boldsymbol{x}}) & \boldsymbol{B}(\tilde{\boldsymbol{x}}) \end{bmatrix}. \tag{13}$$

Given $\boldsymbol{\theta}$, extraction of the quadratic expression in $\tilde{\boldsymbol{x}}$ from (11) is a little more subtle. Define the following vectors:

$$\boldsymbol{t}_n(\boldsymbol{\theta}) \triangleq \boldsymbol{A}_n \begin{bmatrix} 1 \\ \boldsymbol{\theta} \end{bmatrix} \qquad n = 0, 1, \cdots N \tag{14}$$

where $\boldsymbol{A}_n$ are the model matrices of (10a). Substituting (14) and (10a) into (11), $C(\tilde{\boldsymbol{x}}, \boldsymbol{\theta})$ can be written in the form

$$\begin{aligned} C(\tilde{\boldsymbol{x}}, \boldsymbol{\theta}) = &\left[\boldsymbol{t}_0(\boldsymbol{\theta}) + \sum_{n=1}^{N} \tilde{x}_n \boldsymbol{t}_n(\boldsymbol{\theta})\right]^T \boldsymbol{W}_g \left[\boldsymbol{t}_0(\boldsymbol{\theta}) + \sum_{n=1}^{N} \tilde{x}_n \boldsymbol{t}_n(\boldsymbol{\theta})\right] \\ &+ (\boldsymbol{x} - \tilde{\boldsymbol{x}})^T \boldsymbol{W}_x (\boldsymbol{x} - \tilde{\boldsymbol{x}}). \end{aligned} \tag{15}$$

Constructing the matrix

$$\boldsymbol{T}(\boldsymbol{\theta}) \triangleq \left[\boldsymbol{t}_1(\boldsymbol{\theta}) \vdots \boldsymbol{t}_2(\boldsymbol{\theta}) \vdots \cdots \vdots \boldsymbol{t}_N(\boldsymbol{\theta})\right] \tag{16}$$

so that $\sum_{n=1}^{N} \tilde{x}_n \boldsymbol{t}_n(\boldsymbol{\theta}) = \boldsymbol{T}(\boldsymbol{\theta})\tilde{\boldsymbol{x}}$, the criterion (15) may be rewritten as

$$\begin{aligned} C(\tilde{\boldsymbol{x}}, \boldsymbol{\theta}) = &\, c(\boldsymbol{x}, \boldsymbol{\theta}) + 2\left[\boldsymbol{t}_0^T(\boldsymbol{\theta})\boldsymbol{W}_g\boldsymbol{T}(\boldsymbol{\theta}) - \boldsymbol{x}^T\boldsymbol{W}_x\right]\tilde{\boldsymbol{x}} \\ &+ \tilde{\boldsymbol{x}}^T\left[\boldsymbol{T}^T(\boldsymbol{\theta})\boldsymbol{W}_g\boldsymbol{T}(\boldsymbol{\theta}) + \boldsymbol{W}_x\right]\tilde{\boldsymbol{x}} \end{aligned} \tag{17}$$

where $c(\boldsymbol{x}, \boldsymbol{\theta})$ contains all the terms that are independent of $\tilde{\boldsymbol{x}}$. The minimizing $\tilde{\boldsymbol{x}}$ is then given by

$$\hat{\tilde{\boldsymbol{x}}} = \left[\boldsymbol{T}^T(\boldsymbol{\theta})\boldsymbol{W}_g\boldsymbol{T}(\boldsymbol{\theta}) + \boldsymbol{W}_x\right]^{-1}\left[\boldsymbol{W}_x\boldsymbol{x} - \boldsymbol{T}^T(\boldsymbol{\theta})\boldsymbol{W}_g\boldsymbol{t}_0(\boldsymbol{\theta})\right]. \tag{18}$$

The iterative ACM algorithm therefore assumes the following form.

Alternating Coordinates Minimization (ACM)

```
Preliminaries:
The measurements: $\boldsymbol{x}$
$\boldsymbol{A}_0, \boldsymbol{A}_1, \cdots \boldsymbol{A}_N$–the set of $p \times (q+1)$ matrices of
  the pseudo-linear model.
Model and data weight matrices: $\boldsymbol{W}_g$ and
  $\boldsymbol{W}_x$ (resp.)
Initial guess: $\hat{\tilde{\boldsymbol{x}}}^{[0]}$ (Initializing with $\hat{\boldsymbol{\theta}}^{[0]}$
  is also possible, running the algorithm
  from part II)
Algorithm:
For $k = 1, 2, \cdots$ repeat until convergence:
  I. Minimize with respect to $\boldsymbol{\theta}$:
    Construct
        $\boldsymbol{A}^{[k]} = \boldsymbol{A}_0 + \sum_{n=1}^{N} \hat{\tilde{x}}_n^{[k-1]} \boldsymbol{A}_n$
    Form the partition:
        $\begin{bmatrix} b^{[k]} & \boldsymbol{b}^{[k]^T} \\ \boldsymbol{b}^{[k]} & \boldsymbol{B}^{[k]} \end{bmatrix} = \boldsymbol{A}^{[k]^T}\boldsymbol{W}_g\boldsymbol{A}^{[k]}$
    Obtain $\hat{\boldsymbol{\theta}}^{[k]}$:
        $\hat{\boldsymbol{\theta}}^{[k]} = -\boldsymbol{B}^{[k]^{-1}}\boldsymbol{b}^{[k]}$
  II. Minimize with respect to $\tilde{\boldsymbol{x}}$:
    Let
        $\boldsymbol{v}^{[k]} = \left[1 \; \hat{\boldsymbol{\theta}}^{[k]^T}\right]^T$
    Construct
        $\boldsymbol{t}_0^{[k]} = \boldsymbol{A}_0\boldsymbol{v}^{[k]}, \quad \boldsymbol{T}^{[k]} = \left[\boldsymbol{A}_1\boldsymbol{v}^{[k]} \vdots \boldsymbol{A}_2\boldsymbol{v}^{[k]} \vdots \cdots \vdots \boldsymbol{A}_N\boldsymbol{v}^{[k]}\right]$
    Obtain $\hat{\tilde{\boldsymbol{x}}}^{[k]}$:
        $\hat{\tilde{\boldsymbol{x}}}^{[k]} = \left[\boldsymbol{T}^{[k]^T}\boldsymbol{W}_g\boldsymbol{T}^{[k]} + \boldsymbol{W}_x\right]^{-1}\left[\boldsymbol{W}_x\boldsymbol{x} - \boldsymbol{T}^{[k]^T}\boldsymbol{W}_g\boldsymbol{t}_0^{[k]}\right]$
Upon convergence ($k = K$), set $\boldsymbol{\theta}_{\mathrm{XLS}} = \hat{\boldsymbol{\theta}}^{[K]}$.
  As a by-product, $\hat{\boldsymbol{x}}_{\mathrm{XLS}} = \hat{\tilde{\boldsymbol{x}}}^{[K]}$.
```

Various convergence criteria may be used, e.g., monitor the amount of the change in $\hat{\boldsymbol{\theta}}$, $\hat{\tilde{\boldsymbol{x}}}$, or $C(\hat{\tilde{\boldsymbol{x}}}, \hat{\boldsymbol{\theta}})$ between iterations and compare it with a reasonably small threshold.

In Fig. 2, we demonstrate typical convergence paths, as well as sensitivity to initialization, of the ACM algorithm, using the same AR(2) model and data as in Fig. 1. Iterations are presented as connected points in the 2-D parameter space, using different initialization values (placed on a grid), which are denoted by "○." The terminal values are denoted by "□." Fig. 2(a) and (b) show the results with $w_g = 0.2, 5$, respectively. Contour lines of the XLS criterion under minimization are superimposed. It is seen that when w_g is small, the problem is LS-like, and a unique (global) minimum is reached in a few iterations, with the minimum's vicinity approached in the first iteration, regardless of initialization. However, when w_g is large, the problem is more STLS-like, and several minima can be reached, depending on initialization. Moreover, more iterations are generally required, and the vicinity of the minimum is reached gradually.

When no particular initial guess for $\boldsymbol{\theta}$ is available, it is reasonable to use $\hat{\tilde{\boldsymbol{x}}}^{[0]} = \boldsymbol{x}$.

B. XLS Minimization via the Extended CTLS (ECTLS) Algorithm

The minimization algorithm developed in this subsection is based on an extension of the existing CTLS algorithm that was originally proposed in 1991 by Cadzow ([14]–[16]) for solving the "constrained total least squares" problem. For the sake of readers unfamiliar with CTLS, we will briefly review the CTLS problem and solution strategy and proceed to formulate a more general problem, which we term the ECTLS problem, proposing a general iterative solution thereof. We will then tie the knot between ECTLS and XLS by transforming the XLS criterion into an ECTLS problem, which warrants application of our ECTLS algorithm in XLS minimization.

1) CTLS Problem: The CTLS problem is aimed at finding the nearest rank-deficient approximation of a given matrix $\boldsymbol{A}$ so that the approximating matrix $\boldsymbol{B}$ belongs to a specified set of "properly structured" matrices (e.g., Hankel, Toeplitz, Block, etc.).

Explicitly stated, let $\mathcal{B}$ denote a set of "properly structured" $p \times \tilde{q}$ matrices (we use $\tilde{q}$ as an arbitrary dimension to be later related to q via $\tilde{q} = q + 2$). Given a $p \times \tilde{q}$ data matrix $\boldsymbol{A}$, find another matrix $\boldsymbol{B} \in \mathcal{B}$ and a nonzero $\tilde{q} \times 1$ vector $\boldsymbol{y}$ such that $\|\boldsymbol{B} - \boldsymbol{A}\|_F^2$ is minimized[2] subject to $\boldsymbol{By} = \boldsymbol{0}$:

$$\min_{\boldsymbol{B}\in\mathcal{B},\, \boldsymbol{y}\neq\boldsymbol{0}} \|\boldsymbol{B} - \boldsymbol{A}\|_F^2 \quad s.t.\text{:} \quad \boldsymbol{By} = \boldsymbol{0} \tag{19}$$

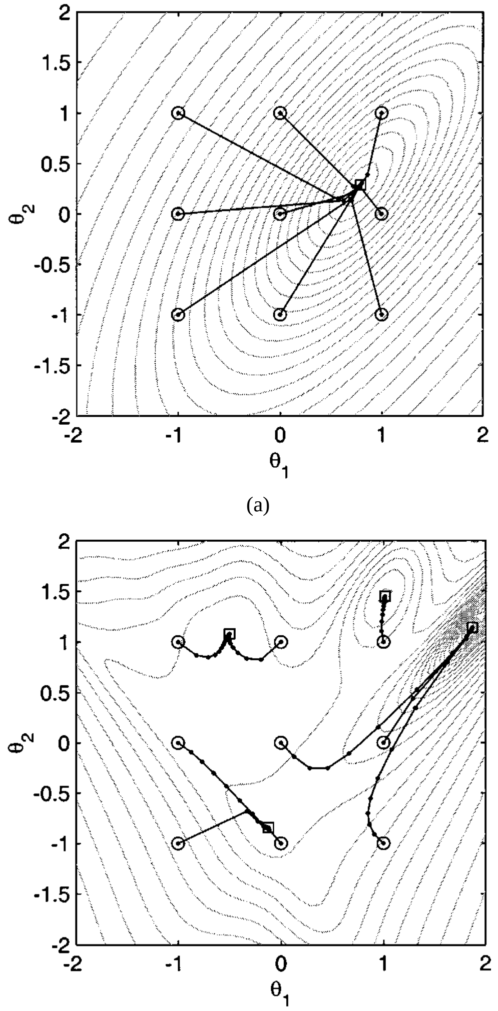
possibly with an additional constraint on the scale of $\boldsymbol{y}$.

2) CTLS Solution: The CTLS solution, which is based on Cadzow's concept of a composite property mapping ([26]), is an iterative process in which each iteration involves two phases. In the first phase of the first iteration, the structural constraint is ignored, and the nearest rank-deficient approximation $\boldsymbol{B}'$ of $\boldsymbol{A}$ is found using the TLS approach. While $\boldsymbol{B}'$ satisfies the linear constraint, it may not be properly structured. Therefore, in the second phase, the linear constraint $\boldsymbol{By} = \boldsymbol{0}$ is ignored, and a "properly structured" $\boldsymbol{B}'' \in \mathcal{B}$ nearest approximation to $\boldsymbol{B}'$ is found.

$\boldsymbol{B}''$ may not satisfy the linear constraint, but among all "properly structured" matrices, it is the closest to $\boldsymbol{B}'$, which, on its part, is the closest to $\boldsymbol{A}$ and satisfies that constraint. The idea is thus to repeat the process with the role of the original matrix $\boldsymbol{A}$ taken over by the most recent approximation $\boldsymbol{B}''$.

Under some regularity conditions, it was shown in [16] that this algorithm is guaranteed to converge to a matrix (and associated vector) that satisfies both the structural and the linear constraints. However, that matrix is not guaranteed to be the nearest to $\boldsymbol{A}$ among all matrices that share that property. In other words, the obtained solution is not necessarily at a true minimum—although it is often "close enough" in the sense that its deviation from a true minimum is often negligible with respect to inherent statistical estimation errors associated with applications that involve the CTLS problem.

3) ECTLS Problem: Our ECTLS problem extends the CTLS problem to accommodate an additional fixed linear constraint on the vector $\boldsymbol{y}$. Specifically, it is assumed that in addition to the $p \times \tilde{q}$ matrix $\boldsymbol{A}$, a full-rank $s \times \tilde{q}$ matrix $\boldsymbol{T}$ is specified, and the CTLS problem is to be solved subject to $\boldsymbol{Ty} = \boldsymbol{0}$. The necessity of this extension will unfold as we transform the XLS minimization problem into an ECTLS problem in the sequel.



(a)

(b)

Fig. 2. Typical convergence paths of the ACM algorithm for an AR(2) process in noise. Fig. 2(a) and (b) use the same data, with $w_g = 0.2,\ 5$, respectively. "○" denote initial values, and "□" denote terminal values. Intermediate results are denoted by points (connected). With small w_g, the problem is "easier," and fewer iterations are required. With large w_g, more iterations are required, and local minima can be reached.

Explicitly stated, let $\mathcal{B}$ denote the set of all "properly structured" $p \times \tilde{q}$ matrices. Given the $p \times \tilde{q}$ data matrix $\boldsymbol{A}$ and an $s \times \tilde{q}$ ($s < \tilde{q}$) full-rank matrix $\boldsymbol{T}$, find another $p \times \tilde{q}$ matrix $\boldsymbol{B} \in \mathcal{B}$ and a nonzero $\tilde{q} \times 1$ vector $\boldsymbol{y}$ such that $\|\boldsymbol{B} - \boldsymbol{A}\|_F^2$ is minimized, subject to $\boldsymbol{By} = \boldsymbol{0}$ and $\boldsymbol{Ty} = \boldsymbol{0}$

$$\min_{\boldsymbol{B}\in\mathcal{B},\, \boldsymbol{y}\neq\boldsymbol{0}} \|\boldsymbol{B} - \boldsymbol{A}\|_F^2 \quad s.t.\text{:} \quad \boldsymbol{By} = \boldsymbol{0}, \quad \boldsymbol{Ty} = \boldsymbol{0} \tag{20}$$

possibly with an additional constraint on the scale of $\boldsymbol{y}$.

4) ECTLS Solution: We indicate up front that in principle, the ECTLS problem can be easily transformed into a CTLS problem by redefining the set of allowable matrices $\mathcal{B}$ and the target matrix $\boldsymbol{A}$ such that the $\boldsymbol{Ty} = \boldsymbol{0}$ constraint is contained by $\boldsymbol{By} = \boldsymbol{0}$. This is easily attained by augmenting all matrices

[2] The Frobenius norm of a $p \times q$ matrix $\boldsymbol{P}$ is defined as $\|\boldsymbol{P}\|_F^2 \triangleq \sum_{i=1}^{p}\sum_{j=1}^{q} P_{i,j}^2 = Trace\{\boldsymbol{PP}^T\} = Trace\{\boldsymbol{P}^T\boldsymbol{P}\}$.

in the set by $\boldsymbol{T}$. However, that appealing strategy has a significant drawback. When the ordinary CTLS algorithm is invoked using the augmented version, the convergence rate becomes extremely slow. This can be partly explained by the fact that the augmented part is extraneous to the other data involved in $\boldsymbol{A}$ and hardly interacts with it throughout the iterative process.

We therefore pursue a less trivial solution strategy, which properly incorporates the additional constraint into the minimization process. As in CTLS, we propose a two-phase iterative solution.

We turn now to develop the solution for the first phase. Given a data matrix $\boldsymbol{A}$ and a constraints matrix $\boldsymbol{T}$, we seek the nearest matrix $\boldsymbol{B}'$ and a nonzero vector $\boldsymbol{y}$ such that $\boldsymbol{B}'\boldsymbol{y} = \boldsymbol{0}$ and $\boldsymbol{T}\boldsymbol{y} = \boldsymbol{0}$. We add an arbitrary constraint on the scale of $\boldsymbol{y}$, e.g., $\boldsymbol{y}^T\boldsymbol{y} = 1$, in order to produce a unique solution. Obviously, such a constraint is immaterial to the original minimization problem, which is scale invariant in $\boldsymbol{y}$. We begin by forming the Lagrangian

$$\begin{aligned} &L(\boldsymbol{B}', \boldsymbol{y}, \boldsymbol{l}, \boldsymbol{\xi}, \lambda) \\ &\quad = \tfrac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{\tilde{q}}(A_{i,j}-B'_{i,j})^2 + \boldsymbol{l}^T\boldsymbol{B}'\boldsymbol{y} - \boldsymbol{\xi}^T\boldsymbol{T}\boldsymbol{y} + \tfrac{1}{2}\lambda\left(\boldsymbol{y}^T\boldsymbol{y}-1\right) \end{aligned} \tag{21}$$

where $A_{i,j}$ and $B'_{i,j}$ denote the (i, j)th elements of $\boldsymbol{A}$ and $\boldsymbol{B}'$, respectively, and where $\boldsymbol{l}, \boldsymbol{\xi}$, and λ are Lagrange multipliers. Differentiating with respect to all the variables and equating zero, we get

$$\frac{\partial L}{\partial B'_{n,m}} = 0 \;\Rightarrow\; \boldsymbol{B}' - \boldsymbol{A} + \boldsymbol{l}\boldsymbol{y}^T = \boldsymbol{0} \tag{22}$$

$$\frac{\partial L}{\partial \boldsymbol{y}} = \boldsymbol{0}^T \;\Rightarrow\; \boldsymbol{l}^T\boldsymbol{B}' - \boldsymbol{\xi}^T\boldsymbol{T} + \lambda\boldsymbol{y}^T = \boldsymbol{0}^T \tag{23}$$

$$\frac{\partial^T L}{\partial \boldsymbol{l}} = \boldsymbol{0} \;\Rightarrow\; \boldsymbol{B}'\boldsymbol{y} = \boldsymbol{0} \tag{24}$$

$$\frac{\partial^T L}{\partial \boldsymbol{\xi}} = \boldsymbol{0} \;\Rightarrow\; \boldsymbol{T}\boldsymbol{y} = \boldsymbol{0} \tag{25}$$

$$\frac{\partial L}{\partial \lambda} = 0 \;\Rightarrow\; \boldsymbol{y}^T\boldsymbol{y} = 1. \tag{26}$$

Postmultiplying (23) with $\boldsymbol{y}$ and using (24)–(26) yields $\lambda = 0$. Substituting $\boldsymbol{B}' = \boldsymbol{A} - \boldsymbol{l}\boldsymbol{y}^T$ (22) into (24) and (23), respectively, we get

$$\boldsymbol{A}\boldsymbol{y} = (\boldsymbol{y}^T\boldsymbol{y})\boldsymbol{l} \tag{27a}$$

$$\boldsymbol{A}^T\boldsymbol{l} = \left(\boldsymbol{l}^T\boldsymbol{l}\right)\boldsymbol{y} + \boldsymbol{T}^T\boldsymbol{\xi}. \tag{27b}$$

Defining $\sigma^2 \triangleq \boldsymbol{l}^T\boldsymbol{l}$ and $\boldsymbol{x} \triangleq \boldsymbol{l}/\sigma$, using $\boldsymbol{y}^T\boldsymbol{y} = 1$, and rescaling $\boldsymbol{T}$ (in order to get rid of leading nuisance coefficients, which are irrelevant to the constraint $\boldsymbol{T}\boldsymbol{y} = \boldsymbol{0}$), we get

$$\boldsymbol{A}\boldsymbol{y} = \sigma\boldsymbol{x} \tag{28a}$$

$$\boldsymbol{A}^T\boldsymbol{x} = \sigma\boldsymbol{y} + \boldsymbol{T}^T\boldsymbol{\xi} \tag{28b}$$

which are to be solved subject to $\boldsymbol{y}^T\boldsymbol{y} = 1, \boldsymbol{x}^T\boldsymbol{x} = 1, \boldsymbol{T}\boldsymbol{y} = \boldsymbol{0}$.

Substituting (28a) into (28b) and rescaling $\boldsymbol{T}$ once again, we obtain

$$\left(\boldsymbol{A}^T\boldsymbol{A}\right)\boldsymbol{y} = \sigma^2\boldsymbol{y} + \boldsymbol{T}^T\boldsymbol{\xi}. \tag{29}$$

Now, let $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots \lambda_{\tilde{q}})$ and $\boldsymbol{U}$ be the eigenvalues and eigenvectors matrices, respectively, of $\boldsymbol{A}^T\boldsymbol{A}$, such that $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$ with $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{I}$ (diag$(\cdot)$ is in the MATLAB© convention, and $\boldsymbol{I}$ denotes the $\tilde{q}\times\tilde{q}$ identity matrix). Then, (29) can be rearranged to read

$$\boldsymbol{y} = \boldsymbol{U}\left(\boldsymbol{\Lambda} - \sigma^2\boldsymbol{I}\right)^{-1}\boldsymbol{U}^T\boldsymbol{T}^T\boldsymbol{\xi}. \tag{30}$$

Recalling the constraint $\boldsymbol{T}\boldsymbol{y} = \boldsymbol{0}$, we deduce that all possible solutions must correspond to values of σ^2 such that the matrix $\boldsymbol{T}\boldsymbol{U}(\boldsymbol{\Lambda} - \sigma^2\boldsymbol{I})^{-1}\boldsymbol{U}^T\boldsymbol{T}^T$ is singular (since $\boldsymbol{\xi} = \boldsymbol{0} \Rightarrow \boldsymbol{y} = \boldsymbol{0}$ is unacceptable). In other words, the equation that σ^2 must satisfy is

$$\left|\boldsymbol{T}\boldsymbol{U}\left(\boldsymbol{\Lambda} - \sigma^2\boldsymbol{I}\right)^{-1}\boldsymbol{U}^T\boldsymbol{T}^T\right| = 0 \tag{31}$$

where $|\cdot|$ denotes the determinant of the enclosed matrix. We will show immediately that normally, there are $\tilde{q} - s$ solutions in σ^2 for (31). Once σ^2 is determined and (28a) and (28b) are solved, the resulting Frobenius norm is given by [see (22)]

$$\|\boldsymbol{B} - \boldsymbol{A}\|_F^2 = \left\|\boldsymbol{l}\boldsymbol{y}^T\right\| = \sum_{i=1}^{p}\sum_{j=1}^{\tilde{q}} l_i^2 y_j^2 = \sigma^2\sum_{i=1}^{p} x_i^2 \sum_{j=1}^{\tilde{q}} y_j^2 = \sigma^2. \tag{32}$$

We therefore seek the value of σ^2 that is the smallest among all the possible $\tilde{q} - s$ solutions.

We now turn to discuss the solution of (29), which corresponds to a solution of (31). Let the QR decomposition of $\boldsymbol{T}^T$ be given by

$$\boldsymbol{T}^T = \left[\boldsymbol{Q}_1 \,\vdots\, \boldsymbol{Q}_2\right]\begin{bmatrix}\boldsymbol{R}\\ \boldsymbol{0}\end{bmatrix} = \boldsymbol{Q}_1\boldsymbol{R} \tag{33}$$

where $\boldsymbol{Q} = [\boldsymbol{Q}_1 \,\vdots\, \boldsymbol{Q}_2]$ is $\tilde{q}\times\tilde{q}$ unitary, $\boldsymbol{Q}_1$ is $\tilde{q}\times s$, $\boldsymbol{Q}_2$ is $\tilde{q}\times(\tilde{q}-s)$, $\boldsymbol{R}$ is $s\times s$ nonsingular upper triangular, and $\mathbf{o}$ denotes a $(\tilde{q} - s)\times s$ all-zeros matrix. Then, $\boldsymbol{T}\boldsymbol{y} = \boldsymbol{0}$ iff $\boldsymbol{y}$ belongs to the null space of $\boldsymbol{Q}_1^T$, which is spanned by the columns of $\boldsymbol{Q}_2$. In other words, $\boldsymbol{T}\boldsymbol{y} = \boldsymbol{0}$ iff there exists a vector $\boldsymbol{\eta}$ such that $\boldsymbol{y} = \boldsymbol{Q}_2\boldsymbol{\eta}$. We may therefore rewrite (29) as

$$\left(\boldsymbol{A}^T\boldsymbol{A}\right)\boldsymbol{Q}_2\boldsymbol{\eta} = \sigma^2\boldsymbol{Q}_2\boldsymbol{\eta} + \boldsymbol{Q}_1\boldsymbol{R}\boldsymbol{\xi}. \tag{34}$$

Premultiplying by $\boldsymbol{Q}_2^T$, we get

$$\left[\boldsymbol{Q}_2^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{Q}_2\right]\boldsymbol{\eta} = \sigma^2\boldsymbol{\eta} \tag{35}$$

which identifies the $\tilde{q}-s$ solutions for σ^2 as the $\tilde{q}-s$ eigenvalues of the matrix $\boldsymbol{Q}_2^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{Q}_2$ and further identifies $\boldsymbol{\eta}$ as the corresponding eigenvectors. We wish to select the smallest eigenvalue, which would in turn yield the smallest perturbation under the specified constraints.

To complete the solution, once σ^2 and $\boldsymbol{y}$ are found in either method, we proceed by normalizing $\boldsymbol{y}$ so that $\boldsymbol{y}^T\boldsymbol{y} = 1$ and then apply (22) and (27a) to retrieve $\boldsymbol{B}'$:

$$\boldsymbol{B}' = \boldsymbol{A} - \boldsymbol{l}\boldsymbol{y}^T = \boldsymbol{A}(\boldsymbol{I} - \boldsymbol{y}\boldsymbol{y}^T). \tag{36}$$

In the second phase of each iteration, we want to find a properly structured matrix $\boldsymbol{B}'' \in \mathcal{B}$ that is nearest to $\boldsymbol{B}'$. In order to provide a concise algorithm for the second phase, we use the following characterization of the "properly structured" set $\mathcal{B}$, encompassing a variety of structures such as Toeplitz, Hankel, partly Toeplitz, partly Hankel, circulant, partly circulant, sparse, partly constant etc.:

$$\mathcal{B} = \left\{ \boldsymbol{B} | \boldsymbol{B} = \boldsymbol{B}_0 + \sum_{n=1}^{N} b_n \boldsymbol{B}_n, \qquad b_1,\, b_2,\, \cdots b_N \in \mathcal{R} \right\} \tag{37}$$

where the constant matrices $\boldsymbol{B}_0,\, \boldsymbol{B}_1,\, \cdots \boldsymbol{B}_N$ are restricted not to have overlapping nonzero elements, or, in other words, satisfy the *distinction* property

$$\boldsymbol{B}_n \odot \boldsymbol{B}_m = \mathbf{0} \qquad n \neq m \quad n, m = 0,\, 1,\, \cdots N \tag{38}$$

where $\odot$ denotes Hadamard (i.e., element-wise) multiplication. Note that the distinction property implies, for any linear combination of the matrices $\boldsymbol{B}_0,\, \boldsymbol{B}_1,\, \cdots \boldsymbol{B}_N$,

$$\left\| \sum_{n=0}^{N} b_n \boldsymbol{B}_n \right\|_F^2 = \sum_{n=0}^{N} b_n^2 \|\boldsymbol{B}_n\|_F^2. \tag{39}$$

The distinction property is naturally met in many structure sets $\mathcal{B}$. For example, to describe the set of all Toeplitz matrices, each of the "building-block" matrices $\boldsymbol{B}_1,\, \boldsymbol{B}_2,\, \cdots \boldsymbol{B}_N$ would contain all-ones entries along its corresponding diagonal and be all-zero elsewhere.

Based on (39), it is easily deduced that the matrix $\boldsymbol{B}'' \in \mathcal{B}$ that is nearest to a given matrix $\boldsymbol{B}'$ would be given by

$$\boldsymbol{B}'' = \boldsymbol{B}_0 + \sum_{n=1}^{N} \bar{b}_n \boldsymbol{B}_n \tag{40}$$

where the coefficients $\bar{b}_0,\, \bar{b}_1,\, \cdots \bar{b}_N$ are given by

$$\bar{b}_n = \frac{\mathbf{1}_p^T [\boldsymbol{B}_n \odot \boldsymbol{B}'] \mathbf{1}_{\tilde{q}}}{\|\boldsymbol{B}_n\|_F^2} \qquad n = 1,\, 2,\, \cdots N \tag{41}$$

where $\mathbf{1}_r$ denotes a vector of r ones.

In most applications, the matrices $\boldsymbol{B}_1,\, \boldsymbol{B}_2,\, \cdots \boldsymbol{B}_N$ are extremely sparse. Moreover, all their nonzero elements are usually 1s. In such cases, the calculation of the coefficient in (41) is more simple than it appears because it does not involve any true multiplications but merely denotes averaging elements of $\boldsymbol{B}'$. For example, in the Toeplitz case, each of the elements along diagonals of $\boldsymbol{B}''$ would equal the arithmetic average of all elements along the corresponding diagonal in $\boldsymbol{B}'$.

That concludes the description of our iterative solution to the general ECTLS optimization problem. We now proceed to link to the original XLS problem by identifying XLS as a special case of an ECTLS problem to which we may apply our ECTLS algorithm.

5) XLS Minimization as an ECTLS Problem: Recall that the pseudo-linear model is specified by a set of "building block" matrices $\boldsymbol{A}_0,\, \boldsymbol{A}_1,\, \cdots \boldsymbol{A}_N$ such that $\boldsymbol{A}(\tilde{\boldsymbol{x}})$ is a linear combination thereof, whose coefficients are the elements of $\tilde{\boldsymbol{x}}$. We will assume that the set of matrices is *distinct*. Similarly, we denote by $\boldsymbol{A}(\boldsymbol{x})$ the "measured" model matrix, which is a linear combination of the same "building block" matrices, using the measured data $x_1,\, x_2,\, \cdots x_N$ as coefficients:

$$\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}_0 + \sum_{n=1}^{N} x_n \boldsymbol{A}_n. \tag{42}$$

Let us further define the following.

- The extended $p \times (q+2)$ matrix $\boldsymbol{Y}(\boldsymbol{x}) \triangleq [\mathbf{0} \vdots \boldsymbol{A}(\boldsymbol{x})]$.
- A vector of scaled residual model errors, $\boldsymbol{\epsilon} \triangleq \mu \boldsymbol{g}(\tilde{\boldsymbol{x}},\, \boldsymbol{\theta})$, where μ is an arbitrary nonzero constant.
- Another extended matrix $\tilde{\boldsymbol{Y}}(\boldsymbol{\epsilon},\, \tilde{\boldsymbol{x}}) \triangleq [\boldsymbol{\epsilon} \vdots \boldsymbol{A}(\tilde{\boldsymbol{x}})]$.

The Frobenius norm of the difference between $\boldsymbol{Y}(\boldsymbol{x})$ and $\tilde{\boldsymbol{Y}}(\boldsymbol{\epsilon},\, \tilde{\boldsymbol{x}})$ is then given by

$$\begin{aligned} \|\boldsymbol{Y}(\boldsymbol{x}) - \tilde{\boldsymbol{Y}}(\boldsymbol{\epsilon},\, \tilde{\boldsymbol{x}})\|_F^2 &= \|\boldsymbol{\epsilon}\|^2 + \left\| \sum_{n=1}^{N} (\tilde{x}_n - x_n) \boldsymbol{A}_n \right\|_F^2 \\ &= \mu^2 \|\boldsymbol{g}(\tilde{\boldsymbol{x}},\, \boldsymbol{\theta})\|^2 + \sum_{n=1}^{N} (\tilde{x}_n - x_n)^2 f_n \end{aligned} \tag{43}$$

where $f_n \triangleq \|\boldsymbol{A}_n\|_F^2$ denotes the Frobenius norm of $\boldsymbol{A}_n$, and where (39) was used for the last transition since the set $\boldsymbol{A}_1,\, \boldsymbol{A}_2,\, \cdots \boldsymbol{A}_N$ is assumed distinct.

Therefore, the Frobenius norm of the difference can be expressed as

$$\begin{aligned} &\|\boldsymbol{Y}(\boldsymbol{x}) - \tilde{\boldsymbol{Y}}(\boldsymbol{\epsilon},\, \tilde{\boldsymbol{x}})\|_F^2 \\ &\quad = \boldsymbol{g}^T(\tilde{\boldsymbol{x}},\, \boldsymbol{\theta}) \boldsymbol{W}_g \boldsymbol{g}(\tilde{\boldsymbol{x}},\, \boldsymbol{\theta}) + (\tilde{\boldsymbol{x}} - \boldsymbol{x})^T \boldsymbol{W}_x (\tilde{\boldsymbol{x}} - \boldsymbol{x}) \end{aligned} \tag{44}$$

where $\boldsymbol{W}_g = \mu^2 \boldsymbol{I}$, and $\boldsymbol{W}_x = \mathrm{diag}(f_1,\, f_2,\, \cdots f_N)$. Thus, minimization of the Frobenius norm is equivalent to the minimization of the XLS criterion (8), with specific weight matrices $\boldsymbol{W}_g$ and $\boldsymbol{W}_x$ (recall that μ is a free parameter, which can be chosen arbitrarily).

However, in order for the minimization to maintain the relation between $\boldsymbol{\epsilon}$ and $\boldsymbol{g}(\boldsymbol{x},\, \boldsymbol{\theta})$, an additional constraint has to be imposed on the matrix $\tilde{\boldsymbol{Y}}(\boldsymbol{\epsilon},\, \tilde{\boldsymbol{x}})$, namely

$$\tilde{\boldsymbol{Y}}(\boldsymbol{\epsilon},\, \tilde{\boldsymbol{x}}) \begin{bmatrix} -1/\mu \\ 1 \\ \boldsymbol{\theta} \end{bmatrix} = \mathbf{0}. \tag{45}$$

Thus, the XLS minimization criterion may be expressed as the minimization of the Frobenius distance between the extended matrices $\boldsymbol{Y}(\boldsymbol{x}) = [\mathbf{0} \vdots \boldsymbol{A}(\boldsymbol{x})]$ and $\tilde{\boldsymbol{Y}}(\boldsymbol{\epsilon},\, \tilde{\boldsymbol{x}}) = [\boldsymbol{\epsilon} \vdots \boldsymbol{A}(\tilde{\boldsymbol{x}})]$, subject to (45), where μ is a free weighting parameter, and $\boldsymbol{\epsilon}$, $\tilde{\boldsymbol{x}}$, and $\boldsymbol{\theta}$ are the free variables for minimization.

We therefore identify $\boldsymbol{Y}(\boldsymbol{x})$ as the ECTLS target matrix (which is denoted $\boldsymbol{A}$ in the ECTLS problem formulation). $\tilde{\boldsymbol{Y}}(\boldsymbol{\epsilon},\, \tilde{\boldsymbol{x}})$ is the approximating matrix (which is denoted $\boldsymbol{B}$ in ECTLS), which has to be "properly structured," i.e., has to be expressed as a linear combination of "building-block" matrices as in (37). The number of "building-block" matrices for ECTLS, which is denoted as $N_{\text{ECTLS}} + 1$, is actually $N + 1 + p$, where $N + 1$ is the number of "building-block" matrices $\boldsymbol{A}_0, \boldsymbol{A}_1, \cdots \boldsymbol{A}_N$ in the XLS model, and p is the column dimension of $\boldsymbol{Y}(\boldsymbol{x})$. We now explicitly identify these $N + 1 + p$ "building-block" matrices $\boldsymbol{B}_0, \boldsymbol{B}_1, \cdots \boldsymbol{B}_{N+p}$:

$\boldsymbol{B}_0, \boldsymbol{B}_1, \cdots \boldsymbol{B}_N$ are versions of $\boldsymbol{A}_0, \boldsymbol{A}_1, \cdots \boldsymbol{A}_N$ that are extended by an additional zeros column on the left

$$\boldsymbol{B}_n = \left[\mathbf{0} \,\vdots\, \boldsymbol{A}_n\right] \qquad n = 0, 1, \cdots N \tag{46}$$

where $\mathbf{0}$ denotes a $p \times 1$ vector of zeros.

$\boldsymbol{B}_{N+1}, \boldsymbol{B}_{N+2}, \cdots \boldsymbol{B}_{N+p}$ are additional "building-blocks" that are formally required to accommodate the p elements of $\boldsymbol{\epsilon}$ into the first column of $\boldsymbol{Y}(\boldsymbol{\epsilon}, \tilde{\boldsymbol{x}})$ without structural constraint

$$\boldsymbol{B}_{N+n} = \left[\boldsymbol{e}_n \,\vdots\, \mathbf{0}\right] \qquad n = 1, 2, \cdots p \tag{47}$$

where $\boldsymbol{e}_n$ denotes the nth column of the $p \times p$ identity matrix, and $\mathbf{0}$ denotes a $p \times (q + 1)$ all-zeros matrix.

Therefore, $\tilde{\boldsymbol{Y}}(\boldsymbol{\epsilon}, \tilde{\boldsymbol{x}})$ can be easily expressed as required by

$$\tilde{\boldsymbol{Y}}(\boldsymbol{\epsilon}, \tilde{\boldsymbol{x}}) = \left[\boldsymbol{\epsilon} \,\vdots\, \boldsymbol{A}(\tilde{\boldsymbol{x}})\right] = \boldsymbol{B}_0 + \sum_{n=1}^{N} \tilde{x}_n \boldsymbol{B}_n + \sum_{i=1}^{p} \epsilon_i \boldsymbol{B}_{N+i} \tag{48}$$

where $\tilde{x}_n$ denotes the nth element of $\tilde{\boldsymbol{x}}$, and ϵ_i denotes the ith element of $\boldsymbol{\epsilon}$.

Naturally, if $\boldsymbol{A}_0, \boldsymbol{A}_1, \cdots \boldsymbol{A}_N$ is a distinct set of matrices, so is the set $\boldsymbol{B}_0, \boldsymbol{B}_1, \cdots \boldsymbol{B}_{N+p}$.

We identify the ECTLS vector $\boldsymbol{y} \triangleq [-1/\mu \; 1 \; \boldsymbol{\theta}^T]^T$, where μ is a free nonzero processing parameter, and $\boldsymbol{\theta}$ is the desired vector of model parameters. This is where the additional linear constraint on $\boldsymbol{y}$ is introduced, which warrants the use of the ECTLS solution rather than the more simple CTLS. The ECTLS constraint matrix $\boldsymbol{T}$ is identified as $\boldsymbol{T} = [\mu \; 1 \; \mathbf{0}^T]$, where $\mathbf{0}$ denotes a $q \times 1$ all-zeros vector.

Note that $\boldsymbol{T}$ may be extended to accommodate additional linear constraints into the XLS problem without further conceptual complications. This feature can be exploited to address problems where some of the elements of $\boldsymbol{\theta}$ (or some linear combinations thereof) are known. Such are, for example, the problem of ARX system identification or AR signal modeling when some of the poles are known *a priori*. A similar concept has been introduced into TLS and STLS problems by Dowling and DeGroat [27], [28] as well as Chen *et al.* [3], [29]–[31]. For example, if the additional linear constraint on $\boldsymbol{\theta}$ is of the form $\boldsymbol{P\theta} = \boldsymbol{q}$, then $\boldsymbol{T}$ would be augmented as

$$\boldsymbol{T} = \begin{bmatrix} \mu & 1 & \mathbf{0}^T \\ \mathbf{0} & -\boldsymbol{q} & \boldsymbol{P} \end{bmatrix}. \tag{49}$$

The free processing parameter μ controls the XLS weighting by determining the relation between the weight matrices $\boldsymbol{W}_g = \mu^2 \boldsymbol{I}$ and $\boldsymbol{W}_x = \text{diag}(f_1, f_2, \cdots f_N)$. The Frobenius norms $f_1, f_2, \cdots f_N$ of nearly all the matrices $\boldsymbol{A}_0, \boldsymbol{A}_1, \cdots \boldsymbol{A}_N$ are often similar so that $\boldsymbol{W}_x$ is close to a scaled version of $\boldsymbol{I}$. If this is not the case, a simple scaling operation may be applied to $\boldsymbol{A}_0, \boldsymbol{A}_1, \cdots \boldsymbol{A}_N$ (and properly absorbed into the data), which would make $\boldsymbol{W}_x$ a scaled identity matrix, if desired.

We now summarize the XLS minimization via the ECTLS approach.

Extended Constrained Total Least Squares (ECTLS) Algorithm

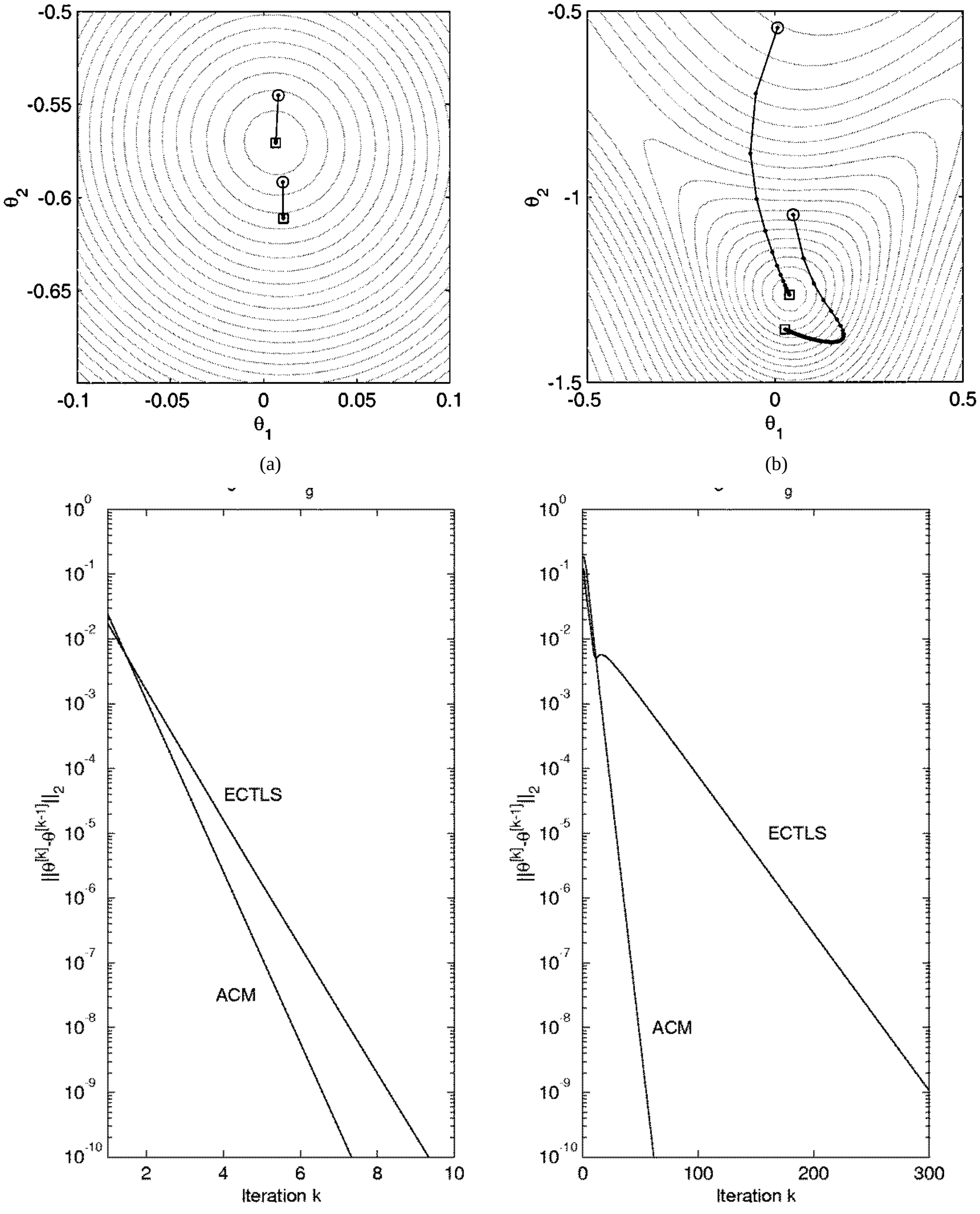
```
Preliminaries:
The measurements: $\boldsymbol{x} = [x_1 \; x_2 \; \cdots \; x_N]^T$
$\boldsymbol{A}_0, \boldsymbol{A}_1, \cdots \boldsymbol{A}_N$—the distinct set of $p \times (q + 1)$
  matrices of the pseudo-linear model.
A weight-balancing parameter: $\mu$
Preparation/Initialization:
Construct the linear constraint matrix $\boldsymbol{T} =$
  $[\mu \; 1 \; \mathbf{0}^T]$,
where 0 denotes a vector of $q$ zeros [If
  desired, extend $\boldsymbol{T}$ to accommodate additional linear constraints on $\boldsymbol{\theta}$—see
  (49)].
Denote by $s$ the resulting row-dimension of
  $\boldsymbol{T}$ (normally $s = 1$).
Find the QR decomposition of $\boldsymbol{T}^T$:

  $$\boldsymbol{T}^T = \left[\boldsymbol{Q}_1 \,\vdots\, \boldsymbol{Q}_2\right] \begin{bmatrix} \boldsymbol{R} \\ \mathbf{0} \end{bmatrix} = \boldsymbol{Q}_1 \boldsymbol{R}$$

  where $\boldsymbol{Q} = [\boldsymbol{Q}_1 \;\vdots\; \boldsymbol{Q}_2]$ is unitary, $\boldsymbol{Q}_1$ is
  $(q + 2) \times s$, $\boldsymbol{Q}_2$ is $(q + 2) \times (q + 2 - s)$, $\boldsymbol{R}$ is
  $s \times s$ nonsingular upper triangular, and 0
  denotes a $(q + 2 - s) \times s$ all zeros matrix.
Construct the matrix: $\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}_0 + \sum_{n=1}^{N} x_n \boldsymbol{A}_n$
and set: $\boldsymbol{Y}^{[0]} = [\mathbf{0} \,\vdots\, \boldsymbol{A}(\boldsymbol{x})]$
where 0 denotes a $p \times 1$ all-zeros vector.
Algorithm:
For $k = 1, 2, \cdots$ repeat until convergence:
  Compute $\boldsymbol{R}^{[k]} = \boldsymbol{Q}_2^T \boldsymbol{Y}^{[k-1]^T} \boldsymbol{Y}^{[k-1]} \boldsymbol{Q}_2$.
  Find the smallest eigenvalue $\lambda^{[k]}$ and
  corresponding eigenvector $\boldsymbol{\eta}^{[k]}$ of $\boldsymbol{R}^{[k]}$
  ($\boldsymbol{\eta}^{[k]}$ is assumed to have a unit norm).
  Let $\boldsymbol{y}^{[k]} = \boldsymbol{Q}_2 \boldsymbol{\eta}^{[k]}$.
  Let $\boldsymbol{B'}^{[k]} = \boldsymbol{Y}^{[k-1]}(\boldsymbol{I} - \boldsymbol{y}^{[k]} \boldsymbol{y}^{[k]^T})$
  Let

    $$\hat{\hat{x}}_n^{[k]} = \left(\mathbf{1}_p^T \left[\left[\mathbf{0} \,\vdots\, \boldsymbol{A}_n\right] \odot \boldsymbol{B'}^{[k]}\right] \mathbf{1}_{q+2}\right) \Big/ \|\boldsymbol{A}_n\|_F^2$$
    $$n = 1, 2, \cdots N$$

  where 0 is a vector of $p$ zeros, and $\mathbf{1}_r$
  is a vector of $r$ ones.
  Let $\boldsymbol{\epsilon}^{[k]} = \boldsymbol{B'}^{[k]}(\vdots, 1)$ (i.e., $\boldsymbol{\epsilon}^{[k]}$ is the first
  column of $\boldsymbol{B'}^{[k]}$)
  Let $\boldsymbol{Y}^{[k]} = [\boldsymbol{\epsilon}^{[k]} \,\vdots\, \boldsymbol{A}_0 + \sum_{n=1}^{N} \hat{\hat{x}}_n^{[k]} \boldsymbol{A}_n]$.
upon convergence ($k = K$),
$\boldsymbol{\theta}_{\text{XLS}} = \boldsymbol{y}^{[K]}(3{:}\, q + 2)/\boldsymbol{y}^{[K]}(2)$ (i.e., 3rd thru last
  elements of $\boldsymbol{y}^{[k]}$ divided by the second).
As a byproduct, $\hat{\hat{\boldsymbol{x}}}_{\text{XLS}} = [\hat{\hat{x}}_1^{[K]}, \hat{\hat{x}}_2^{[K]}, \cdots \hat{\hat{x}}_N^{[K]}]^T$.
```



(a) (b)

(c) (d)

Fig. 3. Typical convergence patterns of the ACM and ECTLS algorithms for an AR(2) process in noise. ACM was initialized "naturally" with $\hat{\hat{\boldsymbol{x}}}^{[0]} = \boldsymbol{x}$ (No initialization is required for ECTLS). (a) and (b) "○" initial values, "□" terminal values. Intermediate results are denoted by points (connected). (c) and (d) Norm of the update $\|\hat{\boldsymbol{\theta}}^{[k]} - \hat{\boldsymbol{\theta}}^{[k-1]}\|_2$ versus iteration number (k). In both cases, ACM converged to a true minimum, and ECTLS did not. $w_g = 0.2$ in (a) and (c) and $w_g = 5$ in (b) and (d).

The same convergence criteria that were mentioned in the context of ACM may be used. Note that there is no choice of initialization. The convergence properties of ECTLS are inherited from CTLS. Convergence is guaranteed, although the point of convergence is not guaranteed to be a true minimum. The major advantage of ECTLS is its relative computational simplicity per iteration, typically $O(N)$ versus $O(N^2)$ or even $O(N^3)$ in ACM.

In Fig. 3(a) and (b), we compare typical convergence paths of the ACM and ECTLS algorithms (using the same data) for an AR(2) model with different weighting values of w_g. ACM was used with the "natural" initialization $\hat{\hat{\boldsymbol{x}}}^{[0]} = \boldsymbol{x}$. The first iteration is denoted by "○" and the terminal iteration by "□." In Fig. 3(c) and (d), convergence rates are demonstrated for the same cases, respectively, in terms of the norm of the update $\|\hat{\boldsymbol{\theta}}^{[k]} - \hat{\boldsymbol{\theta}}^{[k-1]}\|_2$ versus the iteration number k. It is seen that although both algorithms display linear convergence, ECTLS has a break point in the $w_g = 5$ case [Fig. 3(d)], from which its convergence rate slows down. We do not have a solid explanation for this unfortunate behavior (which was observed to be typical of ECTLS with large w_g). Note, in addition, that as expected, ECTLS does not converge to a true minimum (in contrast to ACM).

V. Application Example

In this section, we demonstrate the application of the XLS minimization algorithms to the problem of estimating the parameters of an AR process from noisy observations.

We begin by formulating the estimation problem. Let $\tilde{x}_n$, $n = 1, 2, \cdots$ be an AR process of known order q [AR(q)]. The underlying model may be written as

$$\tilde{x}_n \approx -\sum_{k=1}^{q} \theta_k \tilde{x}_{n-k} \qquad n = 1, 2, \cdots \tag{50}$$

where $\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \cdots \ \theta_q]^T$ is the vector of unknown parameters. The initial conditions $\tilde{x}_0, \tilde{x}_{-1}, \cdots \tilde{x}_{-q+1}$ are assumed known (however, it is possible to accommodate unknown initial conditions as well).

In the case of strict equality in (50), $\tilde{x}_n$ is a deterministic process, which (for nonzero initial conditions) is usually comprised of a linear combination of exponential signals (if all the poles of its Z-transform are of single multiplicity). If, for example, all the initial conditions are zero except for $\tilde{x}_0 = 1$, then $\tilde{x}_n$ is the impulse response of an LTI all-poles system. The problem of estimating $\boldsymbol{\theta}$ from noisy observations of $\tilde{x}_n$ in that case has been treated extensively using STLS or equivalent techniques (e.g., [9], [11]). Generally, however, the deviations of (50) from equality are unknown, and $\tilde{x}_n$ is therefore a stochastic AR process.

Assume now that we are given N inaccurate (noisy) measurements x_n of $\tilde{x}_n$

$$x_n \approx \tilde{x}_n \qquad n = 1, 2, \cdots N \tag{51}$$

and wish to estimate $\boldsymbol{\theta}$ from $x_1, x_2, \cdots x_N$ (with known initial conditions $\tilde{x}_0, \tilde{x}_{-1}, \cdots \tilde{x}_{-q+1}$).

A. Formulation as an XLS Problem

We identify the measured data $\boldsymbol{x} = [x_1 \ x_2 \ \cdots \ x_N]^T$, the underlying data $\tilde{\boldsymbol{x}} = [\tilde{x}_1 \ \tilde{x}_2 \ \cdots \ \tilde{x}_N]^T$, and the pseudo-linear model that was obtained by rewriting (50) in matrix form as

$$\begin{aligned} \boldsymbol{g}(\tilde{\boldsymbol{x}}, \boldsymbol{\theta}) &= \boldsymbol{A}(\tilde{\boldsymbol{x}})\boldsymbol{v}(\boldsymbol{\theta}) \\ &= \begin{bmatrix} \tilde{x}_1 & \tilde{x}_0 & \cdots & \tilde{x}_{-q+1} \\ \tilde{x}_2 & \tilde{x}_1 & \ddots & \tilde{x}_{-q+2} \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{x}_N & \tilde{x}_{N-1} & \cdots & \tilde{x}_{N-q} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \boldsymbol{\theta} \end{bmatrix} \approx \boldsymbol{0}. \end{aligned} \tag{52}$$

The "building-block" $N \times (q+1)$ matrices $\boldsymbol{A}_0, \boldsymbol{A}_1, \cdots \boldsymbol{A}_N$ are all Toeplitz matrices constructed as follows. $\boldsymbol{A}_0$ is all-zeros except for the upper-right triangle, such that the $(1, m)$th element is the initial value $\tilde{x}_{2-m}$ (for $m = 2, 3, \cdots q+1$). The rest are all-zeros, except for a diagonal of $1 - s$, beginning at the $(n, 1)$th element for $\boldsymbol{A}_n$, $n = 1, 2, \cdots N$.

B. Application of the Minimization Algorithms

We will now identify matrices, vectors, and operations involved in the two minimization algorithms when applied to the problem. In order to simplify the presentation, we will assume the known initial conditions to be all zeros.

1) ACM: We will apply the ACM algorithm with weight matrices $\boldsymbol{W}_x = \boldsymbol{I}, \boldsymbol{W}_g = w_g\boldsymbol{I}$, where $\boldsymbol{I}$ denotes the $N \times N$ identity matrix, and w_g is a positive processing parameter.

Referring to the ACM algorithm, $\boldsymbol{A}^{[k]}$ would be an $N \times (q+1)$ Toeplitz matrix with an all-zeros upper right triangle (due to the zero initial conditions) and with the $(n, 1)$th element being $\hat{\tilde{x}}_n^{[k-1]}$, $n = 1, 2, \cdots N$. Consequently, the desired partition of $\boldsymbol{A}^{[k]^T}\boldsymbol{W}_g\boldsymbol{A}^{[k]}$ assumes the form

$$\boldsymbol{A}^{[k]^T}\boldsymbol{W}_g\boldsymbol{A}^{[k]} = w_g N \begin{bmatrix} \hat{r}^{[k-1]} & \hat{\boldsymbol{r}}^{[k-1]^T} \\ \hat{\boldsymbol{r}}^{[k-1]} & \hat{\boldsymbol{R}}^{[k-1]} \end{bmatrix} \tag{53}$$

where $\hat{r}^{[k-1]}$, $\hat{\boldsymbol{r}}^{[k-1]}$, and $\hat{\boldsymbol{R}}^{[k-1]}$ denote empirical correlations based on the last iteration's values $\hat{\tilde{x}}_1^{[k-1]}, \hat{\tilde{x}}_2^{[k-1]}, \cdots \hat{\tilde{x}}_N^{[k-1]}$ as follows:

$$\hat{r}^{[k-1]} = \frac{1}{N}\sum_{n=1}^{N}\left(\hat{\tilde{x}}_n^{[k-1]}\right)^2 \tag{54a}$$

and

$$\hat{\boldsymbol{r}}^{[k-1]} = \frac{1}{N}\sum_{n=1}^{N}\hat{\tilde{x}}_n^{[k-1]}\hat{\tilde{\boldsymbol{x}}}_n^{[k-1]} \tag{54b}$$

$$\hat{\boldsymbol{R}}^{[k-1]} = \frac{1}{N}\sum_{n=1}^{N}\hat{\tilde{\boldsymbol{x}}}_n^{[k-1]}\hat{\tilde{\boldsymbol{x}}}_n^{[k-1]^T} \tag{54c}$$

where $\hat{\tilde{\boldsymbol{x}}}_n^{[k-1]} \triangleq [\hat{\tilde{x}}_{n-1}^{[k-1]} \ \hat{\tilde{x}}_{n-2}^{[k-1]} \ \cdots \ \hat{\tilde{x}}_{n-q}^{[k-1]}]^T$ $(n = 1, 2, \cdots N)$ using the zero initial conditions when required for $n \leq q$.

Thus, the first phase (minimization with respect to $\boldsymbol{\theta}$) yields

$$\hat{\boldsymbol{\theta}}^{[k]} = -\hat{\boldsymbol{R}}^{[k-1]^{-1}}\hat{\boldsymbol{r}}^{[k-1]} \tag{55}$$

which coincides with the ordinary LS, or Yule–Walker estimate of $\boldsymbol{\theta}$, based on the previous iteration's estimate of $\tilde{\boldsymbol{x}}$.

In the second phase of each iteration, the minimization with respect to $\tilde{\boldsymbol{x}}$ requires the construction of $\boldsymbol{T}^{[k]}$ from the previous estimate $\hat{\boldsymbol{\theta}}^{[k]}$: $\boldsymbol{T}^{[k]}$ would be an $N \times N$ Toeplitz matrix, which is all-zeros, except for $q+1$ diagonals. The $(n, 1)$th element is 1 for $n = 1$ (for the main diagonal) and $\hat{\theta}_{n-1}^{[k]}$ for $n = 2, 3, \cdots q$. Since $\boldsymbol{t}_0^{[k]}$ is all-zeros (due to the zero initial conditions), we have

$$\hat{\tilde{\boldsymbol{x}}}^{[k]} = \left[w_g\boldsymbol{T}^{[k]^T}\boldsymbol{T}^{[k]} + \boldsymbol{I}\right]^{-1}\boldsymbol{x}. \tag{56}$$

Actually, in a statistical framework, this second phase can be interpreted as Wiener filtering (estimation) of $\hat{\tilde{\boldsymbol{x}}}$ from $\boldsymbol{x}$. Under such an interpretation, it is assumed that the autocorrelation of $\tilde{\boldsymbol{x}}$ is $[w_g\boldsymbol{T}^{[k]^T}\boldsymbol{T}^{[k]}]^{-1}$ (and thus depends on the recent estimate $\hat{\boldsymbol{\theta}}^{[k]}$) and that the additive noise is white with unit variance (implied by the arbitrary use of $\boldsymbol{W}_x = \boldsymbol{I}$).

Consequently, the ACM algorithm can be viewed as alternating between optimum filtering of the signal (when its required statistics are deduced from the recent estimate of the parameters) and LS estimate of the parameters (assuming the recent estimate of the signal to be noiseless). That approach was applied, e.g., by Lim and Oppenheim [21], in the context of estimating the poles of noisy speech. It is also reminiscent of the estimate-maximize (EM) algorithm (e.g., [32]), although it is essentially different from EM in the estimation phase. In fact, in a proper statistical context, whereas EM leads to the maximum likelihood (ML) estimate of $\boldsymbol{\theta}$, ACM leads to an estimate termed "joint maximum-*a posteriori* maximum likelihood" (JMAP-ML), whose interpretations and statistical properties are further discussed in [23]–[25].

The computational load per iteration can be shown to be roughly $2N^2 + (q+1)N$ multiply-add operations ("flops"), assuming proper exploitation of the Toeplitz structures involved.

2) ECTLS: We apply the ECTLS algorithm with the weight-balancing parameter $\mu = \sqrt{(q+1)w_g}$, where w_g is the weight used in ACM above. This causes the minimized costs to be nearly equivalent, with the exception of some end effects. The factor $q+1$ is due to the Frobenius norm of $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots \boldsymbol{A}_{N-q}$. The end effect is due to the smaller Frobenius norms of $\boldsymbol{A}_{N-q+1}$ through $\boldsymbol{A}_N$.

The application of ECTLS follows the specified algorithm without special structural observations. Note only that the apparently cumbersome computation of $\hat{x}_n^{[k]}$ reduces, in our case, to the following simple averaging:

$$\hat{x}_n^{[k]} = \begin{cases} \dfrac{1}{q+1}\displaystyle\sum_{m=0}^{q} {B'}^{[k]}_{n+m,\, m+2} & n = 1, 2, \cdots N-q \\ \dfrac{1}{N-n+1}\displaystyle\sum_{m=0}^{N-n} {B'}^{[k]}_{n+m,\, m+2} & n = N-q+1 \\ & N-q+2, \cdots N \end{cases} \tag{57}$$

where $B'^{[k]}_{i,j}$ denotes the (i, j)th element of $\boldsymbol{B'}^{[k]}$.

The computational load per iteration is roughly $(q^2 + 6q + 8)N + O(q^3)$ flops, which is only $O(N)$, as opposed to $O(N^2)$ for ACM.

C. Simulations Results

We demonstrate performance and convergence rates for the two algorithms in estimating the parameters of AR(4) and AR(10) processes from $N = 200$ noisy samples, emphasizing dependence on the weight w_g.

For each process the true parameters were set to reflect poles located symmetrically on a circle of radius $\rho < 1$ in the Z-plane:

$$1 + \sum_{k=1}^{q} \theta_k z^{-k} = \prod_{k=1}^{q} \left(1 - \rho e^{j2\pi(k-0.5/q)} z^{-1}\right) \tag{58}$$

where q is the AR order (4 or 10). The processes were generated using

$$\tilde{x}_n = -\sum_{k=1}^{q} \theta_k \tilde{x}_{n-k} + w_n \qquad n = 1, 2, \cdots N \tag{59}$$

with zero initial conditions, where w_n is a zero-mean white Gaussian sequence with variance $\sigma_w^2 = 4^2$. The observations were $x_n = \tilde{x}_n + v_v\ n = 1, 2, \cdots N$ where v_n is also a zero-mean white Gaussian process, statistically independent of w_n, with variance $\sigma_v^2 = 1^2$. These conditions reflect a relatively high signal to noise ratio (SNR), rendering the problem more "LS"-like, which warrants a relatively small w_g.

We show the empirical total mean squared error (mse) in estimating $\boldsymbol{\theta}$, calculated as $(1/q)\sum_{k=1}^{q}(\hat{\theta}_k - \theta_k)^2$, vs. the weight w_g. Results are shown for the AR(4) and AR(10) processes in Fig. 4(a) and (b), respectively. All results shown are based on 100 Monte-Carlo trials, always showing the average of 98 trials, with two outliers discarded (occasional outliers were due to convergence to local minima). The results attained by ACM/ECTLS are denoted by dots / circles, respectively.

It is interesting to note that while the two are different (since ECTLS does not attain a true XLS minimum, in contrast to ACM), the performance of ACM is nearly matched by ECTLS but with a smaller w_g. In fact, it has also been observed that ECTLS estimates obtained using any w_g are close to estimates obtained by ACM (from the same data) using a higher value of w_g. A possible explanation is the following. In the second phase of each ECTLS iteration, when the nearest structured matrix is found, no change is applied to the (unstructured) first column (containing "model errors"); only the other columns (containing implied "measurement errors") are changed. Consequently, the accumulated perturbations in the first column (applied only in the first phase of each iteration) is inherently smaller than the accumulated perturbations in the data (applied in both phases of each iteration). This inherently unbalanced total perturbation is thus equivalent to using a higher w_g (higher penalty for "model errors").

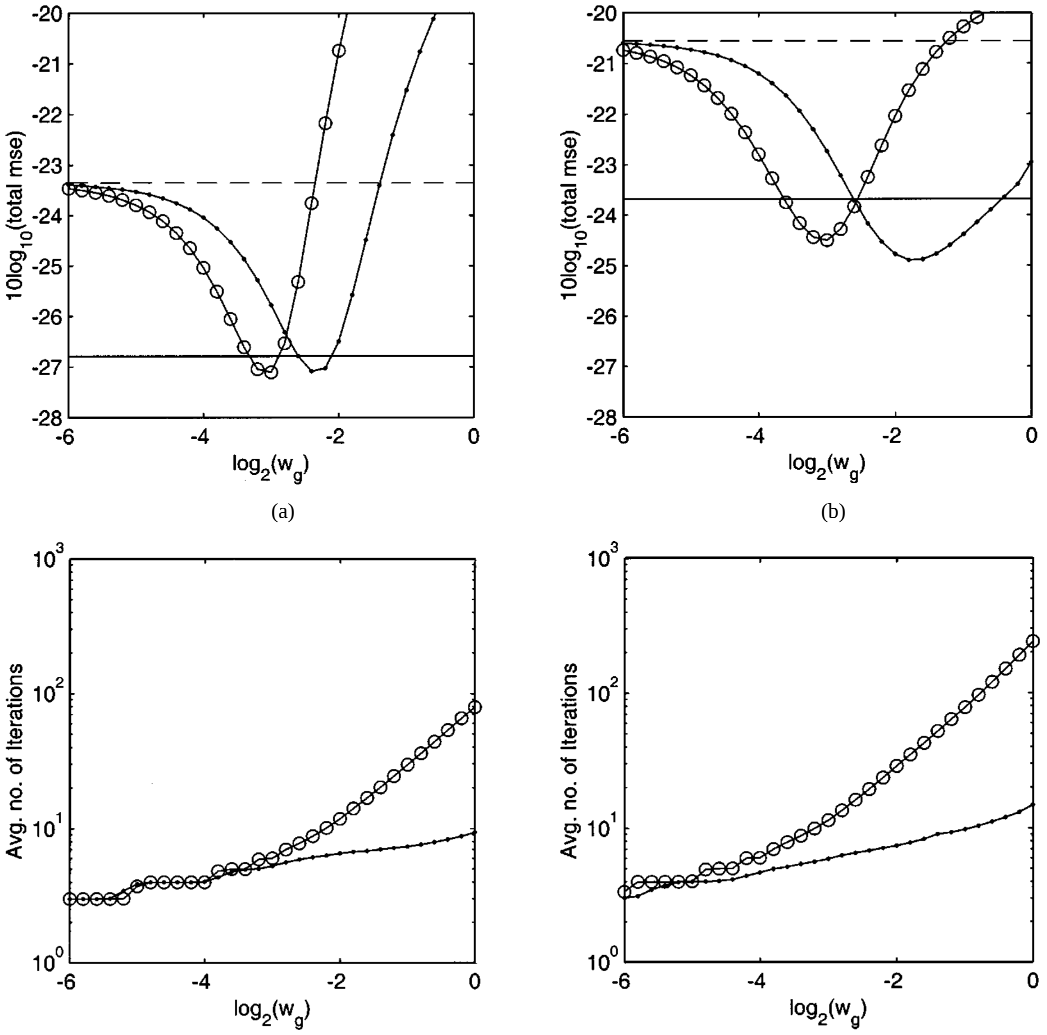
For reference, we also show the (empirical) MSE attained (using the same data) by the simple (biased) LS estimate, as well as by the asymptotically optimal ML estimate.[3] Both the LS and ML estimates are, of course, independent of w_g, but we show them as horizontal dashed/solid (resp.) lines for across-the-board comparison. For fair comparison, the two worst results out of 100 (corresponding to occasional outliers for ML, but not for LS) were excluded from the average.

As expected, when $w_g \ll 1$, the XLS estimate (calculated by either ACM or ECTLS) coincides with LS. On the other hand, when $w_g \gg 1$, the problem turns STLS, whose solution, due to the zero initial conditions, is $\tilde{x}_n = 0\ \forall n$ with no information on $\boldsymbol{\theta}$, approaching infinite MSE. With proper choice of w_g, the XLS estimate can outperform the ML estimate (with finite data). Note, however, that the optimal value of w_g is *not* σ_v^2/σ_w^2 (yet, in further simulations (not presented here), the optimal w_g was shown to be monotonic in σ_v^2/σ_w^2).

In Fig. 4(c) and (d), we show the averaged number of iterations required by both algorithms to converge to updates smaller in norm than 10^{-5}. As w_g increases, the minimization becomes more difficult, and the required number of iterations increases. This problem is aggravated as the dimension of $\boldsymbol{\theta}$ increases. ECTLS is more vulnerable than ACM in this respect, but recall that its computational load per iteration is significantly smaller.

Further simulations results (which are not presented here) indicate that under worse SNR conditions, when the optimal w_g increases, the performance of the algorithms is degraded both in terms of MSE and of convergence speed. Thus, the XLS criterion and associated algorithms are more useful in "nearly-LS"

[3] We used the EM algorithm [32] to calculate the ML estimate, which is the maximizer (with respect to $\boldsymbol{\theta}$) of the probability density function $f(x_1, x_2, \cdots x_N;\ \boldsymbol{\theta})$ (with σ_w^2 and σ_v^2 known).



(a)

(b)

(c)

(d)

Fig. 4. Performance and convergence of ACM (".") and ECTLS ("o") for (a), (c) AR(4) and (b), (d) AR(10) processes in noise with poles located symmetrically on a radius-ρ circle in the z plane (and zero initial conditions). (a), (c) ρ= 0.95. (b), (d) ρ= 0.99. $N = 200$ samples, 100 Monte Carlo trials. Performance is in terms of the total mse. LS (dashed) and ML (solid) performance displayed for reference (all algorithms used the same data; all results reflect averages of 98 trials with two outliers discarded.). Convergence is in terms of average number of iterations. Results are presented versus the XLS weight w_g.

situations as a substitute to the biased LS estimate. It has also been observed that the relative improvement over LS and ML increases as the poles approach the unit circle.

VI. Conclusion

We proposed the XLS criterion for discriminating measurement errors from model errors. The discrimination warrants proper weighting of implied deviations from the measured data and from the specified model. In the limiting cases, the XLS criterion coincides with LS on one hand and with CTLS/STLS on the other hand.

In the context of pseudo-linear models, two iterative minimization algorithms for computing the XLS estimate were proposed, differing in computational load and convergence properties. ACM offers versatile weighting in the form of arbitrary positive-definite weight matrices. ECTLS is confined to diagonal scaling weights only but provides inherent incorporation of linear constraints on the parameters, if necessary. It requires a lower computational effort per iteration but may have a significantly slower convergence rate and may converge to an off-minimum solution.

When the dimensionality of the problem increases, the usefulness of the criterion (and associated algorithms) is gradually confined to "nearly-LS" or "high SNR" conditions, where the measurement errors are much smaller than the model errors. When the conditions approach the opposite ("nearly-STLS") situation, the complexity of the minimization increases. Convergence rates become slower, and false (local) minima are more frequently encountered.

The performance of the XLS estimate depends on proper selection of the weights. It can be outperform the LS and ML estimates with short data records. We stress, however, that the intuitively appealing selection of weights to be the inverses of the noises' covariance matrices (in the spirit of Gauss–Markov theorem for ordinary LS) is *not* optimal (in the MSE sense). Guidelines for optimal selection of weights, as well as further statistical interpretations and analysis of XLS, may be found in [23]–[25].

REFERENCES

[1] G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," *SIAM J. Numer. Anal.*, vol. 17, no. 4, pp. 883–893, 1979.

[2] ——, *Matrix Computations*, 2nd ed. Baltimore, MD: John Hopkins Univ. Press, 1989.

[3] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*, *Frontiers in Applied Mathematics Series*. Philadelphia, PA: SIAM, 1991, vol. 9.

[4] D. Tufts, R. Kumaresan, and I. Kirsteins, "Data adaptive signal estimation by singular value decomposition of a data matrix," *Proc. IEEE*, vol. 70, pp. 684–685, 1982.

[5] T. J. Abatzoglou and J. M. Mendel, "Constrained total least squares," in *Proc. ICASSP*, 1987, pp. 1485–1488.

[6] Y. Hua and T. P. Sarkar, "On the total least squares linear prediction method for frequency estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 2186–2189, 1990.

[7] T. J. Abatzoglou, J. M. Mendel, and G. A. Harada, "The constrained total least squares technique and its application to harmonic superresolution," *IEEE Trans. Signal Processing*, vol. 39, pp. 1070–1087, 1991.

[8] S. Van Huffel, "Enhanced resolution based on minimum variance estimation and exponential data modeling," *Signal Process.*, vol. 33, no. 3, pp. 333–355, 1993.

[9] B. De Moor, "Total least squares for affinely structured matrices and the noisy realization problem," *IEEE Trans. Signal Processing*, vol. 42, pp. 3104–3113, Nov. 1994.

[10] K. Steiglitz and L. E. NcBride, "Techniques for the identification of linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-10, pp. 461–464, 1965.

[11] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1081–1089, Aug. 1986.

[12] J. H. McClellan and D. Lee, "Exact equivalence of the Steiglitz-McBride iteration and IQML," *IEEE Trans. Signal Processing*, vol. 39, pp. 509–512, Feb. 1991.

[13] P. Lemmerling, B. De Moor, and S. Van Huffel, "On the equivalence of constrained total least squares and structured total least squares," *IEEE Trams. Signal Processing*, vol. 44, pp. 2908–2911, Nov. 1996.

[14] J. A. Cadzow and D. M. Wilkes, "Enhanced sinusoidal and exponential modeling," *SVD Signal Process. 11: Algorithms, Anal., Applicat.*, pp. 335–352, 1991.

[15] J. A. Cadzow and D. M. Wilkes, "Constrained total least squares: Signal and system modeling," in *Proc. ICASSP*, vol. 5, 1992, pp. 277–280.

[16] J. A. Cadzow, "Total least squares, matrix enhancement, and signal processing," *Digital Signal Process.*, vol. 4, pp. 21–39, 1994.

[17] B. De Moor, "Structured total least squares and L2 approximation problems," *Linear Algebra Applicat., Special Issue on Numerical Linear Algebra Methods in Control, Signals, Systems*, vol. 188/189, pp. 163–205, 1993.

[18] P. Lemmerling, "Structured total least squares: analysis, algorithms and applications," Ph.D. dissertation, Dept. Elect. Eng., Katholieke Univ., Leuven, Belgium, 1999.

[19] L. Ljung, *System Identification*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[20] W. A. Fuller, *Measurement Error Models*. New York: Wiley, 1987.

[21] J. S. Lim and A. V. Oppenheim, "All-pole modeling of degraded speech," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, pp. 197–210, Mar. 1978.

[22] B. De Moor and P. Lemmerling, "A new trade-off in linear system identification: Misfit versus latency," in *Proc. MTNS*, Padova, Italy, 1998.

[23] A. Yeredor, "The extended least squares criterion for discriminating measurement errors from model errors: Algorithms, applications, analysis," Ph.D. dissertation, Tel-Aviv University, Tel-Aviv, Israel, Fac. Eng., Dept. Elect. Eng.—Systems, 1997.

[24] A. Yeredor and E. Weinstein, "The extended least squares and the joint maximum *a posteriori*–maximum likelihood estimation criteria," in *Proc. ICASSP*, vol. 4, Phoenix, AZ, USA, 1999, pp. 1813–1816.

[25] A. Yeredor, "The joint MAP - ML estimation criterion and its relation to ML and to extended least squares," *IEEE Trans. Signal Processing*, vol. 48, pp. 3484–3492, Dec. 2000.

[26] J. A. Cadzow, "Signal enhancement: A composite property mapping algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 49–62, Jan. 1988.

[27] E. M. Dowling, R. D. DeGroat, and D. A. Linebarger, "Total least squares with linear constrains," in *Proc. ICASSP*, vol. 5, 1992, pp. 341–344.

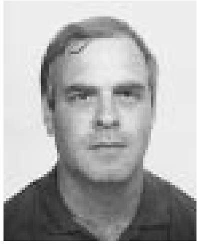
[28] ——, "Exponential parameter estimation in the presence of known components and noise," *IEEE Trans. Antennas Propagat.*, vol. 42, pp. 590–599, May 1994.

[29] H. Chen, S. Van Huffel, D. Van Ordmondt, and R. De Beer, "Parameter estimation with prior knowledge of known signal poles for the quantification of NMR spectroscopy data in the time domain," *J. Magn. Reson. A*, vol. 119, pp. 225–234, 1996.

[30] H. Chen, S. Van Huffel, and J. Vandewalle, "Improved methods for exponential parameter estimation in the presence of known poles and noise," *IEEE Trans. Signal Processing*, vol. 45, pp. 1390–1393, May 1997.

[31] S. Van Huffel and H. Zha, "The restricted total least squares problem: Formulation, algorithms and properties," *SIAM J. Matrix Anal. Appl.*, vol. 12, pp. 292–309, 1991.

[32] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Ann. R. Stat. Soc.*, vol. Ser.3g, pp. 1–38, 1977.



Arie Yeredor (M'99) was born in Haifa, Israel, in 1963. He received the B.Sc. (summa cum laude) and Ph.D. degrees in electrical engineering from Tel-Aviv University, Tel-Aviv, Israel, in 1984 and 1997, respectively.

From 1984 to 1990, he was with the Israeli Defense Forces (Intelligence Corps), where he was in charge of advanced research and development activities in the fields of statistical and array signal processing. Since 1990, he has been with NICE Systems, Inc., where he holds a consultant position in the fields of speech and audio processing and emitter location algorithms. He is also with the Department of Electrical Engineering—Systems, Tel-Aviv University, where he is currently a Faculty Member, teaching courses in parameter estimation, statistical signal processing, and digital signal processing. His research interests include estimation theory, statistical signal processing, and blind source separation.