A Unified Approach to Structured Covariances: Fast Generalized Sliding Window RLS Recursions for Arbitrary Basis

Ricardo Merched, *Senior Member, IEEE*

***Abstract*—This paper extends the existing fast RLS recursions originally intended to exponentially windowed problems for general models, to a generalized sliding window formulation (GSWRLS). From a matrix algebra perspective, we show explicitly how the displacement rank of the underlying inverse covariance matrix associated to any operator is defined as a function of the number of window breakpoints and how the fast GSWRLS calculates these rank factors in a fast manner. The recursions hold *regardless* of the (first order) data structure induced and show that fast fixed order and order recursive RLS algorithms can still be obtained for unwindowed data matrices exhibiting a fixed arbitrary relation between successive regressors. Our approach highlights the existence of a certain degree of freedom inherent to structured data matrices induced by general models, showing that efficient representations of their inverse covariances are not limited to factor circulants, but rather constructed from any arbitrary operator. These Bezoutians, usually expressed via reproducing kernel relations, can be represented exactly in matrix form, along with a precise correspondence to variables of a GSWRLS. As a fallout, we obtain a *vector* relation stating the so-called *minimality* property, for extended models and windows, as opposed to analogous generating function arguments normally seen in original approaches. These results pave the way to a more general framework of polynomial Vandermonde covariance decompositions which arise naturally via a proper choice of *recurrence related polynomials*. This has further impact on several signal processing applications, including superfast realization of equalizers in communications scenarios.**

***Index Terms*—Sliding window RLS algorithm, orthonormal model, lattice, regularized least-squares.**

I. Introduction

STRUCTURED problems are present in numerous signal processing and communications scenarios. Frequently, the data model involved in a given problem induces a particular structure in the corresponding solution, which in turn can be exploited in order to decrease the computational effort that would be required to realize the original "unstructured" formula. Common examples arise in the theory of fast recursive least-squares (RLS) algorithms [1], where for $M \times M$ matrices, instead of relying on $\mathcal{O}(M^3)$ Gauss elimination computations, recursive solutions require $\mathcal{O}(M^2)$ by exploiting the sequential least-squares structure of a given data, or even $\mathcal{O}(M)$ multiplications per iteration, by relying on additional data structure [8], [10]–[12]. A fast algorithm is by itself one way of proving the so-called *displacement structure*, tracing back the works [13]–[29] for inversion of Toeplitz matrices. For example, the displacement rank of any Toeplitz form does not exceed 2, a fact that extends similarly to other structures. The displacement rank is propagated through its *Bezoutian*[1], that is, the inverse of a structured matrix whose efficient representation can be exploited in numerous ways towards complexity reduction.

The concept of displacement has been vastly studied in the adaptive context in terms of successive time-varying inverse covariance matrices, related through sequential rank-one updates and/or downdates, in a more general sliding window context. We shall refer to such matrices (which include a regularization term as well) as a *covariance Bezoutian*, which is simply the inverse of a highly structured deterministic covariance. For instance, in a pre-windowed adaptive filter setup, a widely known result is that the displacement rank of any covariance Bezoutian with respect to a shift operator is equal to 3, while it is equal to 4, for non-prewindowed data models [5]–[7]. The celebrated fast fixed-order [8] and order-recursive RLS algorithms aforementioned [1] are common examples that make use of rank-3 displacements [39], while the *Generalized Window Fast Transversal Filter* (GWFTF) of [30] is an example that relies on a rank-4 displacement. The intrinsic property of these Bezoutians finds application not only in adaptive contexts, but also in non-adaptive computation of equalizers and receivers in channel estimation-based scenarios. For instance, I have shown in [32] and [36] that the rank-3 and rank-4 displacement properties of inverse covariances can be exploited via fixed-order RLS recursions in order to compute MIMO decision feedback equalizers (DFE), and via lattice recursions in the computation of SISO block DFEs in [35] (see also [37]).

The notion of displacement gave a deeper explanation to fast RLS recursions with shift data structure, and allowed us to show more recently that fast fixed-order and order-recursive RLS algorithms hold similarly for any prewindowed data structure whose defining basis functions satisfy recurrence relations

Manuscript received April 10, 2012; revised February 11, 2013; accepted June 27, 2013. Date of publication August 15, 2013; date of current version November 04, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Suleyman S. Kozat. This paper was funded in part by FAPERJ.

The author is with the Signal Processing Laboratory (LPS), Department of Electronics and Computer Engineering, Federal University of Rio de Janeiro, Rio de Janeiro 21945-970, Brazil (e-mail: r.merched@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2013.2278150

[1]This terminology is due to Sylvester [17] in the context of polynomial root localization, referring back to the studies of Euler and Bézout in elimination theory [15], [16], while the displacement equation was first introduced by Cayley [18], following a generating function description, and the works of Hermite [19] in polynomial stability. At that time, their concern was not related to the complexity issues of matrix-vector multiplication or inversion operations (see Olshevsky *et al.* and also [20]).

in [10], [11]. In these references, this was exemplified via a special case of Szegö polynomials [34] orthogonal on the unit circle, proving that even infinite-impulse-response (IIR) filters give rise to fast algorithms with complexity orders that reach much lower levels compared to the ones obtained by the usual tapped-delay-line. In fact, for pre-windowed data, the displacement rank of these covariances in regularized least-squares (LS) or minimum-mean-square-error (MMSE) solutions is shown not to exceed 3, regardless of the (first-order) basis functions generating the underlying data.

The ultimate goal of this presentation is to take a holistic approach on the representation of (time-varying) structured windowed covariances considering arbitrary operators. Our approach is underpinned by the development of a GSWFTF algorithm for *extended* data matrix structures constructed from successive related input regressors. This is because the algorithm itself naturally yields the generators of a displacement equation in a fast, causal manner. The results of this work are closely connected to the development of Bezoutian representations of [2]–[4]. We highlight the importance and applications of the algorithm development itself, and explain how it will be connected to the subjects of Bezoutian decompositions discussed in the related paper [2], as well as its impact in communication applications.

The fast *Generalized Sliding-Window Recursive Least Squares* (GSWRLS) algorithm [30] is known by its superior tradeoff among tracking, robustness, and convergence speed when compared to existing rectangular sliding window (SWRLS) based algorithms. Whereas the SWRLS completely forgets data beyond a length L of past input samples, quickly incurring in numerical errors accumulation, the GSWRLS algorithm makes use of only partial downdate recursions, thus preserving the exponential window decay towards infinity. The resulting window exhibits a characteristic "tail" beyond L samples, which helps regularizing the least-squares (LS) problem, therefore contributing to better conditioning and stability of sliding-window adaptive filters. Compared to the SWRLS and the *Affine Projection Algorithm* (APA) [31], the GSWRLS is able to retain both desired robustness and fast convergence features within a single algorithm. Note that the APA solves an underdetermined system of equations due to a short window length, and yet both SWRLS and the APA are subject to ill-conditioning and noise enhancement effects [30].

Now, the GSWRLS of [30] has been derived for shift data structures only, and no counterpart is available for general models. By general model, we mean any data model such that two successive regression vectors can be related by a constant matrix, say, Ψ. While Ψ is arbitrary, and the fast transversal recursions are obtained regardless of its structure, we shall assume later that it belongs to a class of operators induced by *recurrence related polynomials*, since this guarantees that the complexity of matrix-vector multiplications involving Ψ will be efficient and linear in the filter order. Changing basis representation brings several benefits, including better numerical conditioning, reduced computational complexity, and compact representation of models. For example, rational bases allows substitution of long FIR filters by shorter compact infinite impulse response (IIR) models (see [10] and the references therein). This can represent large computational savings, since in a variety of signal processing and communications applications, such as echo cancelation and equalization, one is often challenged with the problem of training or equalizing long tapped-delay line filters. Unlike the conventional IIR adaptive methods, which present serious problems of stability, local minima and slow convergence, the use of IIR bases offers a stable and global solution, due to the fixed poles location. The resulting algorithm then becomes fast of $\mathcal{O}(M')$ operations per iteration, where $M' < M$. Moreover, these IIR bases can be simultaneously chosen in such way that Ψ is also unitary, which implies perfect numerical conditioning for the computations involving this matrix.

In this paper, we extend the GSWRLS to such generic models, which we also refer to as *Extended Generalized Sliding Window Fast Transversal Filter* (EGSWFTF). Several recursions analogous to the ones encountered in the standard fast RLS theory are derived, as well as new updates and downdates not available in the context of the existing GSWRLS for shift data structure [30]. While some of these recursions may be familiar to the reader acquainted with the GSWRLS recursions, their derivation in our context contains a level of generality not seen in the original algorithms for shift-data, and will be fundamental for the connections with the framework of covariance decompositions pursued in [2] (as well as for the development of future order-recursive algorithm counterparts). This will allow us to pursue a rescuing mechanism for the EGSWFTF, which improves the stability behavior of the recursions.

This paper is organized as follows. Starting from a generalized window formulation, Section II introduces the GSWRLS problem for a data structure induced by a first order relation between two successive regressors. Without loss of generality, the EGSWFTF algorithm is obtained assuming a window with a single breakpoint after L samples. In Section III, the conditions for exact initialization of the EGSWFTF are established regardless of the (first-order) input model. Until this point, no specific structure for the input model is defined. This will be determined in Section IV, which gives a unified treatment on recurrence-related polynomials, and shows how the choice of input basis functions affects the efficiency and numerical conditioning of the underlying EGSWFTF recursions. The bridge between the input basis and the displacement theory is approached in Section V. It is shown how the displacement operator is constructed as a composition of the model operator and any arbitrary companion matrix, while the corresponding displacement equation in terms of the Kalman and prediction vectors of this problem are explicitly given. In Section VI, we solve the displacement equation for the constructed operators, and show how the reproducing kernel of the Ricatti variable, normally written in polynomial form, is instead obtained explicitly in vector form. This result, along with the additional relations of Section VII, complete the set of minimality conditions of the EGSWFTF algorithm, for which we propose a simple rescue mechanism. Finally, simulations considering an extended orthonormal model illustrate the benefits of the proposed algorithm in Section VIII[2].

Notation: We denote by $(\,)^*$ the conjugate and transpose of a vector. Since we will be dealing with row (time) updates and

[2]This work was partially published in [43].

downdates, as well as column (order) updates and downdates (i.e., order recursive variables), we shall write, e.g., $H_{M,N}$, as the order M data matrix at time N. A third index will be necessary, which refers to the updated or downdated variable, e.g., $P_{M,N,L}$ corresponds to the Ricatti variable at time N, defined via a length L window for an order M LS problem, and similarly for $u_{M,N}$, $e_{M,L}(N), \ldots$, and so on. We shall use $\mathrm{diag}(\cdot)$ to capture the diagonal elements of a matrix into a vector, as well as $\mathrm{diag}(\{x_m\}_{m=0}^{M-1})$ to perform the reverse operation. With an abuse of notation, we shall use this operator in the case of block diagonal matrices as well. The notation Φ^{-*} is used to denote $[\Phi^{-1}]^*$. The notation $\angle\phi$ is the phase of ϕ. We use $^\#$ to denote complex conjugation followed by order reversal of the coefficients of a column or row vector, and e.g., $A^\#$ for columnwise reversal in case of matrices. We refer to

$$Z_\theta = \begin{bmatrix} 0 & 0 & \cdots & 0 & \theta_0 \\ 1 & 0 & \cdots & 0 & \theta_1 \\ \vdots & \ddots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \theta_{M-1} \end{bmatrix} \tag{1}$$

as the companion matrix associated to the coefficient vector $\theta = \mathrm{col}\{\theta_0, \theta_1, \ldots, \theta_{M-1}\}$, which collapses to the so called θ_0—*factor circulant* operator, when $\{\theta_i\} = 0$, for $i \neq 0$. The *shift* operator is such that $\theta_i = 0, \forall i$.

II. Fast Generalized Sliding-Window RLS Filter for Extended Models

We assume that the reader is acquainted with the role of the displacement theory in LS adaptive algorithms in its state-of-the-art. We refer to the recent works [10]–[12], where a general framework for exploiting data structure in RLS problems has been introduced. The central idea in these references is to show that fast recursions propagate the displacement rank of the Ricatti variable corresponding to any prewindowed data matrix, whenever a relation among the successive rows of this matrix can be established. More specifically, we denote the individual rows of a data matrix $\mathcal{H}_{M,N}$ by $\{u_{M,k}\}$, i.e., $\mathcal{H}_{M,N}^* = [\, u_{M,0}^* \quad u_{M,1}^* \quad \cdots \quad u_{M,N}^* \,]$, where two successive (row) vectors $\{u_{M,N}, u_{M,N+1}\}$, of order M, are given by

$$\begin{aligned} u_{M,N} &= [\, u_0(N) \quad u_1(N) \quad \cdots \quad u_{M-1}(N) \,] \\ &= [\, u_{M-1,N} \quad u_{M-1}(N) \,], \\ u_{M,N+1} &= [\, u_0(N+1) \quad u_1(N+1) \quad \cdots \quad u_{M-1}(N+1) \,] \\ &= [\, u_0(N+1) \quad \breve{u}_{M-1,N+1} \,]. \end{aligned}$$

In tapped-delay-line models we have $\breve{u}_{M-1,N+1} = u_{M-1,N}$, so that $\mathcal{H}_{M,N}$ becomes *Toeplitz-like*. More generally, let the entries of $\{u_{M,N}, u_{M,N+1}\}$ be related as

$$\breve{u}_{M-1,N+1} = u_{M-1,N}\Psi_{M-1} \tag{2}$$

where Ψ_{M-1} is $(M-1) \times (M-1)$. Our development allows for a very general data structure, in the following senses: (*i*) The successive rows of $\mathcal{H}_{M,N}$ are related by a fixed matrix Ψ_M; (*ii*) The data structure is not necessarily pre-windowed; (*iii*) The structure of $\mathcal{H}_{M,N}$ can be interrupted at several points via scaling, while being exponentially weighted at the same time during each interval. That is, let I_{L_k} denote a size L_k identity matrix, and consider the $(N+1) \times (N+1)$ weighting matrix

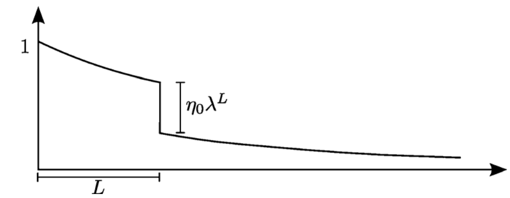

Fig. 1. Generalized window.

$$\begin{aligned} \mathcal{W}_{L_0,\ldots,L_K} \triangleq \mathrm{diag}(\bar{\eta}_K I_{L_K} \ \cdots \ \bar{\eta}_1 I_{L_1} \ \bar{\eta}_0 I_{L_0}) \\ \cdot\mathrm{diag}(\lambda^N \ \cdots \ \lambda \ 1) \end{aligned} \tag{3}$$

where the scalars $\{\bar{\eta}_0, \bar{\eta}_1, \ldots, \bar{\eta}_K\}$, determine the levels of downdating employed at each $L_0, L_1, \ldots, L_K$ past input blocks of data, respectively. Without loss of generality, we shall assume a prewindowed data model with one level of weighting, i.e., $\mathcal{W}_{L_0} = \mathcal{W}_L$, denoting the window length from the current time instant until its first breakpoint. Extension to K levels is straightforward. Fig. 1 illustrates the generalized window, for $\eta_0 \triangleq 1 - \bar{\eta}_1$.

Now, define the $M \times M$ inverse

$$P_{M,N,L} \triangleq \left(\Pi_M^{-1} + \mathcal{H}_{M,N}^* \mathcal{W}_L \mathcal{H}_{M,N}\right)^{-1}, \tag{4}$$

where Π_M is a positive definite regularization matrix. We shall assume that the effect of regularization in (4), in case it exists, is taken into account in the definition of the data matrix itself. That is, one can factor Π_M^{-1} in (4) as $\Pi_M^{-1} = \mathcal{A}_M^* \overline{\mathcal{W}}_Q \mathcal{A}_M$, so that defining the new data matrix as

$$H_{M,N} \triangleq \begin{bmatrix} \mathcal{A}_M \\ \mathcal{H}_{M,N} \end{bmatrix}, \quad \text{where } \mathcal{A}_M \triangleq \begin{bmatrix} u_{M,-Q} \\ \vdots \\ u_{M,-1} \end{bmatrix}, \tag{5}$$

$P_{M,N,L}$ can be equivalently written as $P_{M,N,L} \triangleq (H_{M,N}^* W_L H_{M,N})^{-1}$. Also, $\mathcal{W}_L$ is replaced by W_L, defined similarly to (3), but with size $N' = N + Q + 1$. Of course, this factorization is highly non-unique; the rows $u_{M,-i}, i = 1, \ldots, Q$, in $\mathcal{A}_M$ are simply interpreted as fictitious data, whose structure will be chosen accordingly[3]—see Section III. Still, we shall continue to use the index N in $H_{M,N}$, in order to denote time, instead of N', which denotes the row index.

A. Exploiting Structure Sequentially

Our goal here is to provide the vector relations involving the Kalman gains which are key in making use of data structure. We

[3]It is intuitive to conclude that the displacement rank of the corresponding Bezoutian is no longer 3 as in the case of a single exponentially weighted window, since structure will be broken at several points. Some adaptive algorithms are simply special cases of the above construction; for instance, RLS corresponds to $K = 0$ or $K = 1$ (for non-pre-windowed); the GSWRLS algorithm is equivalent to picking $K = 1$ for a pre-windowed scenario. In the latter, it is clear that the displacement rank of $P_{M,N,L}$ becomes 4.

allow these relations to take into account a non-prewindowed scenario, in which the first row of $H_{M,N}$ is not zero. This would be the case on a non-adaptive scenario where there is no control on the regularization term. In this case, from (2), the relation the data matrices $\breve{H}_{M,N}$ and $H_{M,N}$ through successive instants (i.e., assuming that N is associated with time), we have that

$$\begin{aligned}\breve{H}_{M,N} &= \begin{bmatrix} u_{M,-Q} \\ H_{M,N-1} \end{bmatrix} \Psi_M \\ &= \begin{bmatrix} 0 \\ H_{M,N-1} \end{bmatrix} \Psi_M + \begin{bmatrix} \breve{u}_{M,-Q} \\ 0_{N-1,M} \end{bmatrix}, \end{aligned} \tag{6}$$

with $H_{M,-1} = u_{M,-Q}$, and where for compactness of notation we shall denote $u_{M,-Q} \triangleq u$. The notation $\breve{}$ is due to the fact that in order-recursive problems, a certain column update may lead to a matrix structure that does not necessarily corresponds to data originated within the M taps of input network. That is, when updating $H_{M-1,N}$ to $\breve{H}_{M,N}$ by appending a new column to the right, the augmented matrix can assume any structure, so that $H_{M,N} \neq \breve{H}_{M,N}$. The last column of Ψ_M will be denoted by ψ, which in turn defines the last column of $\breve{H}_{M,N}$—see (41) further ahead.

Now consider the following definitions (7)–(12), described on the top of the next page, involved in the fast generalized window recursions defined from $P_{M,N,L}$ and the coefficient matrix $\breve{P}_{M,N,L} = (\breve{H}^*_{M,N} W_L \breve{H}_{M,N})^{-1}$, obtained from (6).

$$\begin{aligned} k^d_{M,N,L-1} &\triangleq P_{M,N,L} u^*_{M,N-L+1} \\ &= -\lambda^{L-1}\eta_0 P_{M,N,L-1} u^*_{M,N-L+1}\gamma^{-d}_{M,L-1}(N), \\ \bar{k}^d_{M,N,L-1} &\triangleq k^d_{M,N,L-1}\gamma^{d/2}_{M,L-1}(N) \end{aligned} \tag{7}$$

$$\begin{aligned} \breve{k}^d_{M,N,L-1} &\triangleq \breve{P}_{M,N,L} \breve{u}^*_{M,N-L+1} \\ &= -\lambda^{L-1}\eta_0 \breve{P}_{M,N,L-1} \breve{u}^*_{M,N-L+1}\breve{\gamma}^{-d}_{M,L-1}(N), \\ \bar{\breve{k}}^d_{M,N,L-1} &\triangleq \breve{k}^d_{M,N,L-1}\breve{\gamma}^{d/2}_{M,L-1}(N) \end{aligned} \tag{8}$$

$$\begin{aligned} k_{M,N,L} &\triangleq \lambda^{-1} P_{M,N-1,L-1} u^*_{M,N} \\ &= P_{M,N,L} u^*_{M,N}\gamma^{-1}_{M,L}(N), \\ \bar{k}_{M,N,L} &\triangleq k_{M,N,L}\gamma^{1/2}_{M,L}(N) \end{aligned} \tag{9}$$

$$\begin{aligned} \breve{k}_{M,N,L} &\triangleq \lambda^{-1} \breve{P}_{M,N-1,L-1} \breve{u}^*_{M,N} \\ &= \breve{P}_{M,N,L} \breve{u}^*_{M,N}\breve{\gamma}^{-1}_{M,L}(N), \\ \bar{\breve{k}}_{M,N,L} &\triangleq \breve{k}_{M,N,L}\breve{\gamma}^{1/2}_{M,L}(N) \end{aligned} \tag{10}$$

$$\begin{aligned} \gamma^{-d}_{M,L-1}(N) &\triangleq -\lambda^{-(L-1)}\eta_0^{-1} + u_{M,N-L+1} P_{M,N,L} u^*_{M,N-L+1} \\ &= [-\lambda^{L-1}\eta_0(1+\lambda^{L-1}\eta_0 u_{M,N-L+1} \\ &\quad \times P_{M,N,L-1} u^*_{M,N-L+1})]^{-1} \end{aligned} \tag{11}$$

$$\begin{aligned} \gamma^{-1}_{M,L}(N) &\triangleq 1 + \lambda^{-1} u_{M,N} P_{M,N-1,L-1} u^*_{M,N} \\ &= (1 - u_{M,N} P_{M,N,L} u^*_{M,N})^{-1}. \end{aligned} \tag{12}$$

The variables $\bar{k}^d_{M,N,L-1} \triangleq k^d_{M,N,L-1}\gamma^{d/2}_{M,L-1}(N)$ and $\bar{k}_{M,N,L} \triangleq k_{M,N,L}\gamma^{1/2}_{M,L}(N)$ in (7) and (9) are the normalized Kalman gains corresponding to the updating and downdating LS problems respectively, with respect to their likelihood variables (11), (12). Similar definitions hold for the ones arising from $\breve{P}_{M,N,L}$ in (8) and (10).

One way to exploit structure is to obtain a vector update for the Kalman vectors $\{\breve{k}^d_{M,N,L}, \breve{k}_{M,N,L}\}$, defined in (7) and (9), instead of ones relying on matrix-vector multiplications based recursions. This can be achieved, e.g., in a *causal* manner, from (6) as $\breve{P}^{-1}_{M,N,L} = \Psi^*_M(P^{-1}_{M,N-1,L} + \lambda^{N'} u^* u)\Psi_M$, which implies that $\{\breve{P}_{M,N,L}, P_{M,N-1,L}\}$ are related as

$$\breve{P}_{M,N,L} = \Psi^{-1}_M P_{M,N-1,L}\Psi^{-*}_M - \bar{\breve{k}}^{d_o}_{M,N}\bar{\breve{k}}^{d_o *}_{M,N} \tag{13}$$

where $\bar{\breve{k}}^{d_o}_{M,N} \triangleq \bar{\breve{k}}^{d_o}_{M,N,L-1}$, with $\breve{k}^{d_o}_{M,N} \triangleq \Psi^{-1}_M P_{M,N-1,L-1} u^*$, $\bar{\breve{k}}^{d_o}_{M,N} \triangleq \breve{k}^{d_o}_{M,N-1}\gamma^{d_o/2}_{M,L-1}(N)$, and likelihood variable $\gamma^{-d_o}_{M,L-1}(N) \triangleq \lambda^{-N'} + u P_{M,N-1,L} u^*$ for $L = N+1$. Using (7) and (9), the relations for $\{k_{M,N-1,L}, \breve{k}_{M,N,L}\}$ and $\{k^d_{M,N-1,L}, \breve{k}^d_{M,N,L}\}$ are readily established from (13) as

$$\breve{k}_{M,N,L} = \Psi^{-1}_M k_{M,N-1,L} - \left(\bar{\breve{k}}^{d_o *}_{M,N} u^*_{M,N}\right)\bar{\breve{k}}^{d_o}_{M,N} \tag{14}$$

$$\breve{k}^d_{M,N,L} = \Psi^{-1}_M k^d_{M,N-1,L} - \left(\bar{\breve{k}}^{d_o *}_{M,N} u^*_{M,N-L+1}\right)\bar{\breve{k}}^{d_o}_{M,N}. \tag{15}$$

Of course, for $\lambda < 1$, the effect of $u^* u$ disappears with time. In the adaptive filtering context, because u is arbitrary chosen a priori, we can set $u = 0$, which eliminates the need of recursions for $\bar{\breve{k}}^{d_o *}_{M,N} u^*_{M,N}$ and $\bar{\breve{k}}^{d_o *}_{M,N} u^*_{M,N-L+1}$ in (14) and (15). In an non-adaptive scenario, additional recursions for propagating $\bar{\breve{k}}^{d_o}_{M,N}$ can be extended along the lines of [36], where those were derived for shift-data structures.

Relations (14) and (15) are in turn key to performing fast update of the Kalman vectors, and require two main ingredients for this purpose: (*i*) That we obtain order updates for these quantities; (*ii*) That these matrix-vector multiplications are $\mathcal{O}(M)$ efficient as well.

B. Forward and Backward Prediction Updates and Downdates

We now derive all the updating equations that do not rely on data structure. In a sliding window scenario, the coefficient matrix $P_{M,N,L}$ is time-(row) updated as

$$P_{M,N,L} = \lambda^{-1} P_{M,N-1,L-1} - \bar{k}_{M,N,L}\bar{k}^*_{M,N,L} \tag{16}$$

while its downdated recursion is given by

$$P_{M,N,L-1} = P_{M,N,L} - \bar{k}^d_{M,N,L-1}\bar{k}^{d*}_{M,N,L-1}. \tag{17}$$

In the latter, the fraction $\eta_0\lambda^{L-1}$ of $u_{M,N-L+1}$ is removed from the covariance $P^{-1}_{M,N,L}$. Hence, the solution $w_{M,N,L}$ to

$$\min_w \|y_N - H_{M,N} w\|^2_{W_L} \tag{18}$$

can be recursively computed irrespective of data structure via

$$\begin{aligned} w^d_{M,N-1,L-1} &= w_{M,N-1,L} + k^d_{M,N-1,L-1}\gamma^d_{M,L-1}(N) \\ &\quad \times [d(N-L) - u_{M,N-L} w_{M,N-1,L}] \\ e_{M,L}(N) &= \gamma_{M,L}(N)\epsilon_{M,L}(N), \\ \epsilon_{M,L}(N) &= d(N) - u_{M,N} w^d_{M,N-1,L} \\ w_{M,N,L} &= w^d_{M,N-1,L-1} + \breve{k}_{M,N,L} e_{M,L}(N) \end{aligned}$$

where $\{w^d_{M,N-1,L-1}, \gamma^d_{M,L-1}(N)\}$ denote the solution and likelihood variable in the downdated problem.

Similar recursions hold for the forward and backward prediction problems with analogous expressions for $\{\breve{\gamma}^{-d}_{M,L-1}(N), \breve{\gamma}^{-1}_{M,L}(N)\}$. The likelihood factors relate the forward *a priori* errors $\alpha^d_{M-1,L-1}(N-2) \triangleq u(N-L-1) - u_{M-1,N-L-1}w^f_{M-1,N-2,L}$ and $\alpha_{M-1,L}(N-1) \triangleq u(N) - u_{M-1,N-1}w^f_{M-1,N-2,L-1}$ as well as the backward *a priori* errors $\beta^d_{M-1,L-1}(N-1) \triangleq u_{M-1}(N-L) - u_{M-1,N-L}w^b_{M-1,N-1,L-1}$ and $\beta_{M-1,L}(N) \triangleq u_{M-1}(N) - u_{M-1,N}w^b_{M-1,N-1,L}$ to their *a posteriori* versions (defined by replacing the prediction vectors with their time updates):

$$\begin{aligned} f^d_{M,L}(N) &= -\lambda^{-L+1}\eta_0^{-1}\gamma^d_{M,L}(N)\alpha^d_{M,L}(N),\\ b^d_{M,L}(N) &= -\lambda^{-L+1}\eta_0^{-1}\gamma^d_{M,L}(N)\beta^d_{M,L}(N) \end{aligned} \tag{19}$$

$$\begin{aligned} f_{M,L}(N) &= \gamma_{M,L}(N)\alpha_{M,L}(N),\\ b_{M,L}(N) &= \gamma_{M,L}(N)\beta_{M,L}(N). \end{aligned} \tag{20}$$

The gain vector $k^d_{M,N-1,L}$ is forward order-updated as

$$k^d_{M,N-1,L-1} = \begin{bmatrix} 0 \\ k^d_{M-1,N-1,L-1} \end{bmatrix} + \frac{\alpha^d_{M-1,L-1}(N-1)}{\xi^f_{M-1,L}(N-1)} \begin{bmatrix} 1 \\ -w^f_{M-1,N-1,L} \end{bmatrix}. \tag{21}$$

The likelihood factor and minimum cost in the above recursion satisfy the following updates:

$$\gamma^d_{M,L-1}(N) = \gamma^d_{M-1,L-1}(N)\left[1 - \gamma^d_{M-1,L-1}(N) \times \left|\alpha^d_{M-1,L-1}(N)\right|^2 \Big/ \xi^f_{M-1,L-1}(N)\right] \tag{22}$$

$$\xi^f_{M-1,L-1}(N) = \xi^f_{M-1,L}(N) + \gamma^d_{M-1,L-1}(N)\left|\alpha^d_{M-1,L-1}(N)\right|^2 \tag{23}$$

which, when combined yield

$$\gamma^d_{M,L-1}(N-1) = \gamma^d_{M-1,L-1}(N-1)\frac{\xi^f_{M-1,L}(N-1)}{\xi^f_{M-1,L-1}(N-1)}. \tag{24}$$

Similarly, the forward order update for $k_{M,N-1,L}$ is given by

$$k_{M,N-1,L} = \begin{bmatrix} 0 \\ k_{M-1,N-1,L} \end{bmatrix} + \frac{\alpha_{M-1,L}(N-1)}{\lambda\xi^f_{M-1,L-1}(N-2)} \begin{bmatrix} 1 \\ -w^f_{M-1,N-2,L-1} \end{bmatrix}. \tag{25}$$

with corresponding likelihood variable and minimum cost satisfying

$$\gamma_{M,L}(N-1) = \gamma_{M-1,L}(N-1) - |f_{M-1,L}(N-1)|^2/\xi^f_{M-1,L}(N-1) \tag{26}$$

$$\xi^f_{M-1,L}(N-1) = \lambda\xi^f_{M-1,L-1}(N-2) + \alpha^*_{M-1,L}(N-1)f_{M-1,L}(N-1), \tag{27}$$

which, combined, yield

$$\gamma_{M,L}(N-1) = \gamma_{M-1,L}(N-1)\frac{\lambda\xi^f_{M-1,L-1}(N-2)}{\xi^f_{M-1,L}(N-1)}. \tag{28}$$

The forward prediction vector is row-downdated as

$$w^f_{M-1,N-2,L-1} = w^f_{M-1,N-2,L} + k^d_{M-1,N-2,L-1}\gamma^d_{M-1,L-1}(N-2)\alpha^d_{M-1,L-1}(N-2) \tag{29}$$

and time-updated according to

$$w^f_{M-1,N-1,L} = w^f_{M-1,N-2,L-1} + k_{M-1,N-1,L}f_{M-1,L}(N-1). \tag{30}$$

Analogously, we can easily write the following backward prediction update and downdate counterparts. For example,

$$\breve{\gamma}_{M,L}(N) = \lambda\gamma_{M-1,L}(N)\xi^b_{M-1,L-1}(N-1)/\,\xi^b_{M-1,L}(N) \tag{31}$$

$$\gamma^d_{M,L-1}(N) = \gamma^d_{M-1,L-1}(N)\xi^b_{M-1,L}(N)/\,\xi^b_{M-1,L-1}(N). \tag{32}$$

The remaining equations are already included in the algorithm listing for simplicity. The algorithm is completed via the relation between the likelihood variables of the downdate problems $\{\breve{\gamma}^{-d}_{M,L}(N), \gamma^{-d}_{M,L}(N-1)\}$ and the ones of the update problem $\{\breve{\gamma}^{-1}_{M,L}(N), \gamma^{-1}_{M,L}(N-1)\}$. In view of the two alternative expressions (11) and (12), these can be related as

$$\begin{aligned} \breve{\gamma}^{-d}_{M,L-1}(N) &= \gamma^{-d}_{M,L-1}(N-1),\\ \breve{\gamma}^{-1}_{M,L}(N) &= \gamma^{-1}_{M,L}(N-1), \quad \text{or} \end{aligned} \tag{33}$$

$$\begin{aligned} \breve{\gamma}^{d}_{M,L-1}(N) &= \gamma^{d}_{M,L-1}(N-1),\\ \breve{\gamma}_{M,L}(N) &= \gamma^{-1}_{M,L}(N-1). \end{aligned} \tag{34}$$

Also, the order-updates for their inverses can be obtained by multiplying (21) and (25) from left and right by $u_{M,N-L+1}$ and $u^*_{M,N-1}$, to get

$$\gamma^{-1}_{M,L}(N-1) = \gamma^{-1}_{M-1,L}(N-1) + \frac{\lambda^{-1}|\alpha_{M-1,L}(N-1)|^2}{\xi^f_{M-1,L-1}(N-1)} \tag{35}$$

$$\breve{\gamma}^{-1}_{M,L}(N) = \gamma^{-1}_{M-1,L}(N) + \frac{\lambda^{-1}|\beta_{M-1,L}(N)|^2}{\xi^b_{M-1,L-1}(N)}, \tag{36}$$

as well as the respective order updates

$$\gamma^{-d}_{M,L-1}(N-1) = \gamma^{-d}_{M-1,L-1}(N-1) + \left|\alpha^d_{M-1,L-1}(N-1)\right|^2 \Big/ \xi^f_{M-1,L}(N-1) \tag{37}$$

$$\breve{\gamma}^{-d}_{M,L-1}(N) = \gamma^{-d}_{M-1,L-1}(N) + \left|\beta^d_{M-1,L-1}(N)\right|^2 \Big/ \xi^b_{M-1,L}(N). \tag{38}$$

III. State-Space Description and Initial Conditions

The correct initial conditions are easily set by relying on the interpretation that the effect of Π_M is due to the existence of fictitious input data, i.e., $\Pi_M^{-1} = \mathcal{A}_{M,L}^* \overline{\mathcal{W}}_Q \mathcal{A}_{M,L}$, in which case we can construct $\mathcal{A}_{M,Q}$ by applying an impulse of magnitude $\sqrt{\mu}\lambda^{-(Q-1)/2}$ to the underlying network. This results in the following structure for $\mathcal{A}_{M,Q}$:

$$\mathcal{A}_{M,Q} \triangleq \begin{bmatrix} u_{M,-Q} \\ \vdots \\ u_{M,-1} \end{bmatrix} = \begin{bmatrix} \sqrt{\mu}\lambda^{-\frac{Q-1}{2}} & \times & \times & \times \\ 0 & \times & \times & \times \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \times & \times & \times \end{bmatrix}.$$

As a consequence, Π_M^{-1} assumes the following form:

$$\Pi_M^{-1} = \mathcal{A}_{M,Q}^* \mathcal{W}_Q \mathcal{A}_{M,Q} = \sum_{i=-Q}^{-1} \lambda^{-(i+1)} u_{M,i}^* u_{M,i} \quad (39)$$

$$= \begin{bmatrix} \Pi_{M-1}^{-1} & \breve{c}_{M-1} \\ \breve{c}_{M-1}^* & \breve{\pi}_M \end{bmatrix}$$

$$= \Psi_M^* \begin{bmatrix} \mu & \bar{c}_{M-1}^* \\ \bar{c}_{M-1} & \lambda^{-1}\Pi_{M-1}^{-1} \end{bmatrix} \Psi_M + u^* u, \quad (40)$$

where $\mu \triangleq \xi_M^f(-2)$. Now consider the following partitions:

$$\Psi_M = \begin{bmatrix} p_{M-1} & \psi_0 \\ \bar{\Psi}_{M-1} & \bar{\psi} \end{bmatrix} \quad \text{and} \quad u = [u_{M-1}\ u(M)]. \quad (41)$$

where $\bar{\psi} = [\psi_1\ \psi_2 \cdots \psi_{M-1}]^T$. Thus, expanding (39), we find that Π_{M-1}^{-1} should satisfy the equations shown in (42)–(44) at the bottom of the page. In an adaptive scenario, we can set $u = 0$. Then, if $\lambda^{-1/2}T$ is stable and the pair $(\lambda^{-1/2}\bar{\Psi}_{M-1}^*, \sqrt{\mu}p_{M-1}^*)$ is controllable, the Lyapunov equation (42) admits a unique positive definite solution Π_{M-1}, which is given by

$$\Pi_{M-1}^{-1} = \mu \sum_{k=0}^{\infty} \lambda^{-k} \bar{\Psi}_{M-1}^{k*} p_{M-1}^* p_{M-1} \bar{\Psi}_{M-1}^k. \quad (45)$$

The above condition is in perfect agreement with the stability and controllability arguments from a state-space point of view. Indeed, a state-space description for the regressor $\breve{u}_{M,N}$ can be written as:

$$\begin{cases} \breve{u}_{M-1,N+1} = \breve{u}_{M-1,N} \bar{\Psi}_{M-1} \lambda^{-\frac{1}{2}} + x_0(N) p_{M-1} \\ y(N) = \breve{u}_{M-1,N} w + v(N) \end{cases} \quad (46)$$

with solution $\breve{u}_{M-1,N}^*$

$$\breve{u}_{M-1,N}^* = \sum_{i=0}^{n} \lambda^{-\frac{i}{2}} \bar{\Psi}_{M-1}^{i*} p_{M-1}^* x_0^*(N-i-1) + \lambda^{-\frac{n}{2}} \bar{\Psi}_{M-1}^{*n} \breve{u}_{M-1}^*$$

$$= \mathcal{R}_{\bar{\Psi}_{M-1}^*, p_{M-1}^*} x_{0,N-1}^{\#} + \lambda^{-\frac{n}{2}} \bar{\Psi}_{M-1}^{n*} \breve{u}_{M-1}^* \quad (47)$$

so that, with $n \to \infty$, the associated system controllability matrix is given by (48), shown at the bottom of the page.

All remaining variables thus follow from this solution and are included in the initialization step of Table I, listing what we refer to as the *Extended Generalized Sliding-Window Fast Transversal Filter* (EGSWFTF) for generic models. The least-squares solution $w_{M,N,L}$ is propagated via the last six recursions in the algorithm listing. As we have mentioned, we assume we have control on the regularization term, so we set $u = 0$ in (39).

Remark 1: The existence of the solution to the Lyapunov equation above depends on the structure of the induced operator Ψ. We shall return to this state-space description in Section IV, in the context of recurrence-related input basis functions.

Remark 2: Note that for an order-M LS problem, traditional derivations propagate the prediction and Kalman vectors as order-M quantities. Here, we see that in fact these quantities are only required to be order $M-1$, as expected from the displacement equation of $P_{M,N,L}$—see (83) further ahead.

Remark 3: As we shall see later on, in most cases of interest the operator Ψ_M arises from efficient recurrence polynomial relations, so that its inverse will be well defined. In addition, because of that, multiplication of its transpose Ψ_M^*, or its inverse Ψ_M^{-1} by a vector are easily obtained from the corresponding transpose or inverse realizations well known from systems theory, in light of the so-called *Horner polynomials* [23]. A unified approach considering recurrence related polynomial basis will be presented in Section IV. In the most interesting cases, these are $\mathcal{O}(M)$ efficient as well.

IV. Unified Approach to Recurrence Polynomials and Hessenberg Structures

Consider a *transversal* system realization based on arbitrary basis functions $\{Q_m(z)\}$ as illustrated in Fig. 2.

Although the set $\{Q_m(z)\}$ can be interrelated in several ways, as we have mentioned, in this paper we shall focus on first order relations only, in which case structure will be induced by a constant matrix Ψ_M relating two successive regressors as in (2). Different recurrence relations will give rise to different patterns for Ψ_M.

Note that we can represent Fig. 2 equivalently as a tapped-delay-line followed by a matrix transformation $\mathcal{B}_Q$, say, $\breve{u}_{M,n} = r_{M,n}\mathcal{B}_Q^*$, which performs a particular change of basis described according to this representation—see Fig. 3. The tapped-delay-line case is such that $\mathcal{B}_Q = I$.

An $L \times M$ the data matrix $\breve{H}_{M,N}$ can thus be expressed in terms of a Toeplitz-like matrix $\breve{T}_{M,N}$ as

$$\breve{H}_{M,N} \triangleq \begin{bmatrix} \breve{u}_{M,N-L+1} \\ \vdots \\ \breve{u}_{M,N} \end{bmatrix} = \begin{bmatrix} x_{0,N} & \breve{H}_{M-1,N} \end{bmatrix}$$

$$= \breve{T}_{M,N} \mathcal{B}_Q^* = \begin{bmatrix} r_{M,N-L+1} \\ \vdots \\ r_{M,N} \end{bmatrix} \mathcal{B}_Q^* \quad (49)$$

$$\begin{cases} \Pi_{M-1}^{-1} - \lambda^{-1}\bar{\Psi}_{M-1}^* \Pi_{M-1}^{-1} \bar{\Psi}_{M-1} = \mu p_{M-1}^* p_{M-1} + u_{M-1}^* u_{M-1} & (42) \\ \breve{c}_{M-1} = \mu \psi_0^* p_{M-1}^* + \lambda^{-1} T \Pi_{M-1}^{-1} \bar{\psi} + u^*(M) u_{M-1} & (43) \\ \breve{\pi}_M = \mu |\psi_0|^2 + \lambda^{-1} \bar{\psi}^* \Pi_{M-1}^{-1} \bar{\psi} + |u(M)|^2. & (44) \end{cases}$$

$$\mathcal{R}_{\bar{\Psi}_{M-1}^*, p_{M-1}^*} = \left[p_{M-1}^* \quad \lambda^{-1/2} \bar{\Psi}_{M-1}^* p_{M-1}^* \quad \lambda^{-1} \bar{\Psi}_{M-1}^{*2} p_{M-1}^* \quad \cdots \quad \lambda^{-(n-1)/2} \bar{\Psi}_{M-1}^{n-1*} p_{M-1}^* \right]. \quad (48)$$

TABLE I
THE *EXTENDED GENERALIZED SLIDING-WINDOW FAST TRANSVERSAL FILTER* (EGSWFTF) FOR GENERIC DATA STRUCTURES

Initialization

μ is a small positive; Π_{M-1}^{-1} is given by (45)
$\zeta_{M,L}^{f}(-2) = \mu$
$\breve{c}_{M-1} = \mu\psi_0^* p_{M-1} + \lambda^{-1} T \Pi_{M-1}^{-1}\bar{\psi}$
$\zeta_{M,L}^{b}(-1) = \mu|\psi_0|^2 + \lambda^{-1}\|\Pi_{M-1}^{-1/2}\bar{\psi}\|^2 - \|\Pi_{M-1}^{1/2}\breve{c}_{M-1}\|^2$
$k_{M-1,-2,L-1}^{d} = \Pi_{M-1} u_{M-1,-L-1}^{*}$
$k_{M-1,-1,L} = \lambda^{-1}\Pi_{M-1} u_{M-1,-1}^{*}$
$w_{M-1,-2,L}^{f} = \Pi_{M-1}\bar{c}_{M-1}$
$w_{M-1,-1,L}^{b} = \Pi_{M-1}\breve{c}_{M-1}$
$\gamma_{M-1,L-1}^{d}(-2) = -\lambda^{L-1}\eta_0$
$\gamma_{M-1,L}(-1) = 1$
$w_{M,-1} = 0$

For $N \geq 0$, repeat:

$\alpha_{M-1,L-1}^{d}(N-2) = u(N-L-1) - u_{M-1,N-L-1} w_{M-1,N-2,L}^{f}$
$k_{M,N-2,L-1}^{d} =$
$\begin{bmatrix} 0 \\ k_{M-1,N-2,L-1}^{d} \end{bmatrix} + \frac{\alpha_{M-1,L-1}^{d*}(N-2)}{\zeta_{M-1,L}^{f}(N-2)} \begin{bmatrix} 1 \\ -w_{M-1,N-2,L}^{f} \end{bmatrix}$
$\zeta_{M-1,L-1}^{f}(N-2) =$
$\zeta_{M-1,L}^{f}(N-2) + \gamma_{M-1,L-1}^{d}(N-2)|\alpha_{M-1,L-1}^{d}(N-2)|^2$
$w_{M-1,N-2,L-1}^{f} =$
$w_{M-1,N-2,L}^{f} + k_{M-1,N-2,L-1}^{d}\gamma_{M-1,L-1}^{d}(N-2)\alpha_{M-1,L-1}^{d}(N-2)$
$\gamma_{M,L-1}^{d}(N-2) = \gamma_{M-1,L-1}^{d}(N-2)\frac{\zeta_{M-1,L}^{f}(N-2)}{\zeta_{M-1,L-1}^{f}(N-2)}$
$\alpha_{M-1,L}(N-1) = u(N) - u_{M-1,N-1} w_{M-1,N-2,L-1}^{f}$
$f_{M-1,L}(N-1) = \gamma_{M-1,L}(N-1)\alpha_{M-1,L}(N-1)$
$k_{M,N-1,L} = \begin{bmatrix} 0 \\ k_{M,N-1,L} \end{bmatrix} + \frac{\alpha_{M-1,L}^{*}(N-1)}{\lambda\zeta_{M-1,L-1}^{f}(N-2)} \begin{bmatrix} 1 \\ -w_{M-1,N-2,L-1}^{f} \end{bmatrix}$
$\zeta_{M-1,L}^{f}(N-1) =$
$\lambda\zeta_{M-1,L-1}^{f}(N-2) + \alpha_{M-1,L}^{*}(N-1) f_{M-1,L}(N-1)$
$w_{M,N-1,L}^{f} = w_{M,N-2,L-1}^{f} + k_{M,N-1,L} f_{M-1,L}(N-1)$
$\gamma_{M,L}(N-1) = \gamma_{M-1,L}(N-1)\frac{\lambda\zeta_{M-1,L-1}^{f}(N-2)}{\zeta_{M-1,L}^{f}(N-1)}$
$\breve{k}_{M,N-1,L-1}^{d} = \Psi_M^{-1} k_{M,N-2,L-1}^{d}$
$\breve{k}_{M,N,L} = \Psi_M^{-1} k_{M,N-1,L}$
$\breve{\gamma}_{M,L-1}^{d}(N-1) = \gamma_{M,L-1}^{d}(N-2)$
$\breve{\gamma}_{M,L}(N) = \gamma_{M,L}(N-1)$
$\nu_{M-1,L-1}^{d}(N-1) = k_{M,N-1,L-1}^{d}(M)$
$k_{M-1,N-1,L-1}^{d} = \breve{k}_{1:M-1,N-1,L-1}^{d} + \nu_{M-1,L-1}^{d}(N-1) w_{M-1,N-1,L}^{b}$
$\beta_{M-1,L-1}^{d}(N-1) = \zeta_{M-1,L}^{b}(N-1)\nu_{M-1,L-1}^{d*}(N-1)$
$\gamma_{M-1,L-1}^{d}(N-1) = \breve{\gamma}_{M,L-1}^{d}(N-1)\cdot$
$\quad (1 - \breve{\gamma}_{M,L-1}^{d}(N-1)\beta_{M-1,L-1}^{d}(N-1)\nu_{M-1,L-1}^{d*}(N-1))^{-1}$
$\zeta_{M-1,L-1}^{b}(N-1) =$
$\zeta_{M-1,L}^{b}(N-1) + \gamma_{M-1,L-1}^{d}(N-1)|\beta_{M-1,L-1}^{d}(N-1)|^2$
$w_{M-1,N-1,L-1}^{b} =$
$w_{M-1,N-1,L}^{b} + k_{M-1,N-1,L-1}^{d}\gamma_{M-1,L-1}^{d}(N-1)\beta_{M-1,L-1}^{d}(N-1)$
$\nu_{M-1,L}(N) = k_{M,N,L}(M)$
$k_{M-1,N,L} = \breve{k}_{1:M-1,N,L} + \nu_{M-1,L}(N) w_{M-1,N-1,L-1}^{b}$
$\beta_{M-1,L}(N) = \lambda\zeta_{M-1,L-1}^{b}(N-1)\nu_{M-1,L}^{*}(N)$
$\gamma_{M-1,L}(N) = \breve{\gamma}_{M,L}(N)/(1 - \breve{\gamma}_{M,L}(N)\beta_{M-1,L}(N)\nu_{M-1,L}^{*}(N))$
$b_{M-1,L}(N) = \gamma_{M-1,L}(N)\beta_{M-1,L}(N)$
$\zeta_{M-1,L}^{b}(N) = \lambda\zeta_{M-1,L-1}^{b}(N-1) + b_{M-1,L}(N)\beta_{M-1,L}^{*}(N)$
$w_{M-1,N,L}^{b} = w_{M-1,N-1,L-1}^{b} + k_{M-1,N,L} b_{M-1,L}(N)$

$\epsilon_{M,L-1}^{d}(N-1) = d(N-L) - u_{M,N-L} w_{M,N-1,L}$
$e_{M,L-1}^{d}(N-1) = -\lambda^{-L}\eta^{-1}\gamma_{M,L-1}^{d}(N)\epsilon_{M,L-1}^{d}(N-1)$
$w_{M,N-1,L-1}^{d} = w_{M,N-1,L} + k_{M,N-1,L-1}^{d} e_{M,L-1}^{d}(N-1)$
$\epsilon_{M,L}(N) = d(N) - u_{M,N} w_{M,N-1,L}^{d}$
$e_{M,L}(N) = \gamma_{M,L}(N)\epsilon_{M,L}(N)$
$w_{M,N,L} = w_{M,N-1,L-1}^{d} + \breve{k}_{M,N,L} e_{M,L}(N)$

where $x_{0,N} = [x_0(N-L+1) \cdots x_0(N-1)\ x_0(N)]^T$. By virtue of the delay line of Fig. 3, $\breve{T}_{M,N}$ exhibits a Toeplitz-like structure. We shall make use of this interpretation further ahead.

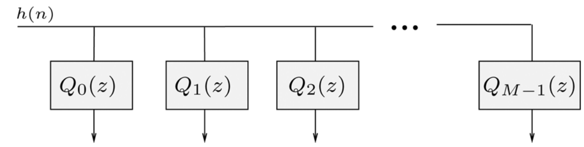

Fig. 2. Transversal realization based on general basis.

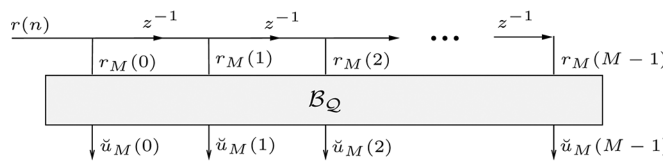

Fig. 3. Transversal realization as a change of basis.

It turns out that the choice of basis functions is key for the $\mathcal{O}(M)$ efficiency of the EGSWRLS recursions of Table I. In other words, note that although the EGSWFTF is obtained regardless of the data structure induced in Ψ_M, the updates of $\breve{k}_{M,N-1,L-1}^{d}$ and $\breve{k}_{M,N,L}$ require efficient multiplications by the inverse Ψ_M^{-1} which in turn requires that Ψ_M itself possesses a particular structure. The point is that when $\{Q_m(z)\}$ are constructed from recurrence-related polynomials, it can be shown that Ψ assumes in general a *Hessenberg* structure [2].

In the following, I unify the theory of structured matrix operations with its algorithmic realization, so that the use of different extended basis is justified.

A. Two-Term, Three-Term, and M-Term Recurrence Relations

There are numerous ways in which we can construct the polynomial basis $\{Q_m(z)\}$. They can be realized via two-terms, three-terms, or generally, via M-*term* recurrence relations, each particular one serving to a different purpose, including better numerical conditioning, reduced computational complexity, compact representation of models, and so on. In this section, we shall examine a few important recurrence relations that will lead to the solution of current open problems and connections.

Consider the (shifted) M-term recurrence relation

$$Q_0(z) = 1,\ Q_1(z) = \delta_0 z^{-1} Q_0(z) \tag{50}$$

$$Q_m(z) = (\bar{a}_{m-2,m-1} + \delta_{m-1} z^{-1}) Q_{m-1}(z) + \bar{a}_{m-3,m-1} Q_{m-2}(z) + \cdots + \bar{a}_{0,m-1} Q_1(z) \tag{51}$$

$$\breve{\Upsilon}_M(z) = \psi_0 Q_0(z) z^{-1} + \psi_1 Q_1(z) z^{-1} + \cdots + \psi_{M-1} Q_{M-1}(z) z^{-1} = \Upsilon_M(z) z^{-1} \tag{52}$$

for $m = 0, 1, \ldots, M-1$. The polynomial $\Upsilon_M(z)$ has also been referred to as the *master polynomial* associated to $\{Q_m(z)\}$ [28]. This results in the upper triangular structure for matrix Ψ_M shown in (53) at the bottom of the next page, where the M-th column $\psi = \mathrm{col}(\psi_0, \psi_1, \ldots, \psi_{M-1})$ defined in (41) contains the coefficients of the last polynomial. This structure can be easily verified via back-substitution of the polynomials $Q_{m-1}(z), Q_{m-2}(z)$, etc., into the definition of $Q_m(z)$. For instance, for $m = 3$, we obtain

$$zQ_3(z) = \delta_2 Q_2(z) + \bar{a}_{12}\delta_1 Q_1(z) + (\bar{a}_{12}\bar{a}_{01} + \bar{a}_{02}) Q_0(z). \tag{54}$$

In other words, the M-term recursion (50) can always be transformed into another M-term recurrence relation of the form

$$zQ_m(z) = a'_{m-1,m}Q_{m-1}(z) + a'_{m-2,m}Q_{m-2}(z) + \cdots + a'_{0,m}Q_0(z), \quad (55)$$

where the $a'_{m-k,m}, k = 1, \ldots, M-1$ are defined by the elements of Ψ_M.

Observe that in principle we do not need to assume any particular structure for the last column of Ψ_M. This is in agreement with what we mentioned regarding the notation of the order-updated matrix $\breve{H}_M$ in the beginning of Section II.A. The data matrix H_M will be perfectly defined, as long as its last column is mapped onto the last column of $\breve{H}_{M,N}$, which can be arbitrary chosen. It represents a degree of freedom in connection with systems theory not fully noticed or exploited in earlier approaches to these problems. For instance, if the latter is chosen as the tap filter weight vector $[w(0)\ w(1)\ \cdots\ w(M-1)]^T$, then $\Upsilon_M(z) = W(z)$.

Several facts regarding the connections of the state-of-the-art in recurrence polynomials with the new results pursued herein are in order:

1) Relation Between Ψ_M and the Confederate Structure: For the M-term recurrence relation in (51), the matrix $\Phi^{-1}_{M,\theta} = \Psi^{-1}_M Z_\theta$ becomes a *Hessenberg* matrix, which has been referred to as a *confederate matrix* associated to the system of polynomials $\{Q_m(z)\}$ (see [21]–[23] and its references). The confederate matrix has several useful properties, and in particular, its eigenvalues are directly related to $\Upsilon_M(z)$, which is a free polynomial. It is defined, e.g., in [23] under the notation "H_Q", where the role of the companion form Z_θ is fixed, and associated to the coefficients of the system transfer function $G(z)$. That is, therein their definition differs from ours with respect to its last column. We provide a broader view instead, suggesting that it can be suitably designed in order to construct operators $\Phi_{M,\theta}$ with eigenvalues placed at any desired location. Moreover, in [21]–[23], this association is such that the highest polynomial order as we see on the right hand side of (55) is always in terms of $Q_{m-1}(z)$, since in these references only FIR basis functions are considered. Here, we further allow for a first-order *rational* transfer function relating two successive basis functions, so that an additional m-th order term may arise in (55). More specifically, assume that $\{Q_m(z)\}$ are generated according to the following two-term recurrence relation:

$$Q_m(z) = \left(\frac{\delta_m z^{-1} - \bar{a}_m}{1 - a_m z^{-1}}\right) Q_{m-1}(z), \quad Q_0(z) = 1, \quad \bar{a}_1 = 0. \quad (56)$$

This results in the Hessenberg structure for Ψ_M shown in (57) at the bottom of the page.

Now, the two-term recurrence of the form (56) can also be transformed into a M-term recurrence relation. Indeed, cross-multiplying the terms of (56) and backsubstitution of the lower order polynomials into $Q_m(z)$, we verify that it satisfies the following M-term recurrence:

$$zQ_m(z) = a_m Q_m(z) + \bar{a}_{m-1,m}Q_{m-1}(z) + \bar{a}_{m-2,m}Q_{m-2}(z) + \ldots + \bar{a}_{0,m}Q_0(z), \quad (58)$$

for $m = 1, \ldots, M-1$, where $a_{m-k,m} = \bar{a}_{m-k} - \delta_{m-k}a_{m-k-1}$. The only difference compared to (55) is the presence of an additional m-th order term $Q_m(z)$ on the r.h.s. of (58), as a result of the IIR nature of these bases. We remark that in this case, $\Phi^{-1}_{M,\theta}$ is no longer Hessenberg, even though it is related to one.

Observe that for Hessenberg matrices, and when $Q_1(z) = cz^{-1}$, the $(0,0)$ element of $\bar{\Psi}_M$ becomes zero, so that $\bar{\Psi}^M_M = 0$. As a result, the state regressor $\breve{u}^*_{M,N}$ given in (47) becomes

$$\breve{u}^*_{M,N} = \sum_{i=0}^{M-1} \lambda^{-\frac{i}{2}} \bar{\Psi}^{i*}_M p^*_M x^*_0(N-i-1) + \lambda^{-\frac{M}{2}} \bar{\Psi}^{*M}_M \breve{u}^*_{M,1} \quad (59)$$

$$= \mathcal{R}_{\bar{\Psi}^*_M, p^*_M} x^{\#}_{0,N-1} \quad (60)$$

$$\Psi_M = \begin{bmatrix} \delta_0 & \bar{a}_{01}\delta_0 & (\bar{a}_{12}\bar{a}_{01} + \bar{a}_{02})\delta_0 & (\bar{a}_{23}(\bar{a}_{12}\bar{a}_{01} + \bar{a}_{02}) + \bar{a}_{13}\bar{a}_{01} + \bar{a}_{03})\delta_0 & \cdots & \psi_0 \\ 0 & \delta_1 & \bar{a}_{12}\delta_1 & (\bar{a}_{23}\bar{a}_{12} + \bar{a}_{13})\delta_1 & \cdots & \psi_1 \\ 0 & 0 & \delta_2 & \bar{a}_{23}\delta_2 & \cdots & \psi_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \ddots & \psi_{M-2} \\ 0 & 0 & 0 & \cdots & & \psi_{M-1} \end{bmatrix} \quad (53)$$

$$\Psi_M = \begin{bmatrix} \delta_1 & -\bar{a}_2\delta_1 & \cdots & -\bar{a}_{M-2}\cdots\bar{a}_2\delta_1 & \psi_0 \\ a_1 & \delta_2 - \bar{a}_2 a_1 & \cdots & \bar{a}_{M-2}\cdots\bar{a}_3(\delta_2 - \bar{a}_2 a_1) & \psi_1 \\ 0 & a_2 & \cdots & -\bar{a}_{M-2}\cdots\bar{a}_4(\delta_3 - \bar{a}_3 a_2) & \psi_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & (\delta_{M-1} - \bar{a}_{M-1}a_{M-2}) & \psi_{M-2} \\ 0 & 0 & \cdots & a_{M-1} & \psi_{M-1} \end{bmatrix} \quad (57)$$

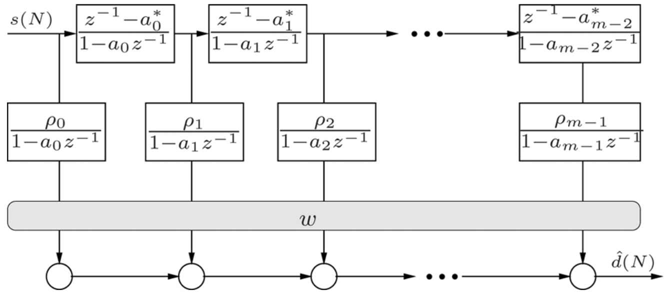

Fig. 4. Transversal orthonormal structure based on IIR Szegö bases.

with the controllability matrix shown in the first equation at the bottom of the page.

Now, substituting $\breve{u}^*_{M,N} = \mathcal{B}_Q s^*_{M,N} = \mathcal{B}_Q x^{\#}_{0,N-1}$ into (47) and using the persymmetry property of Toeplitz matrices, i.e., $\breve{T}_{M,N} = I^{\#}\breve{T}^T_{M,N}I^{\#}$, we verify that $\mathcal{B}_Q = \mathcal{R}_{\bar{\Psi}^*_M, p^*_M}$. Recall that the EGSWFTF algorithm is only feasible in case it can be properly initialized, and depends on the submatrices $\{\bar{\Psi}_{M-1}, p^*_{M-1}\}$ defined in (41). We thus see that for recurrence related basis, regardless of the choice for ψ, as long as $a_i < \sqrt{\lambda}$, and the pair $(\lambda^{-1/2}\bar{\Psi}_{M-1}, \sqrt{\mu}p^*_{M-1})$ is controllable, the structure of Ψ_M guarantees that the Lyapunov equation (42) for the initialization of the EGSWFTF in Table I has a unique solution, and given by $\Pi^{-1}_{M-1} = \mu\sum_{k=0}^{M-1}\lambda^{-k}\bar{\Psi}^{k*}_{M-1}p^*_{M-1}p_{M-1}\bar{\Psi}^k_{M-1} = \mu\mathcal{B}_Q\mathcal{B}^*_Q$. The condition $Q_1(z) = cz^{-1}$ simply implies that $\breve{u}_{M-1,0} = 0$, so that controllability is inferred from the rank of $\mathcal{B}_Q$.

2) Szegö Orthonormal Polynomials on the Unit Circle: In [11], by choosing $\delta_m = \rho_m/\rho_{m-1}, \bar{a}_m = \rho_m a^*_{m-1}/\rho_{m-1}$, with $\rho_m = \sqrt{1-|a_m|^2}$, we obtain an important class of Szegö polynomial basis which is rational and orthonormal on the unit circle, illustrated in Fig. 4:

$$Q_m(z) = \frac{\rho_m}{1-a_m z^{-1}}\prod_{k=0}^{m-1}\frac{z^{-1}-a^*_k}{1-a_k z^{-1}}, \quad a_0 = 0, |a_k| < 1, \tag{61}$$

It is an extension of the so-called *Laguerre* expansion based on a single pole $a_m = a$. It can be verified that with the border conditions $a_0 = a_M = 0$, (57) collapses to a *unitary Hessenberg* matrix (say, for M even), shown at the bottom of the page.

It can be verified that Ψ_M is a unitary matrix, i.e., $\Psi_M\Psi^*_M = I$. Hence, the EGSWFTF algorithm based on (61), besides the benefit of robustness due to an abrupt change in the window parameter η_0, fast convergence inherent to RLS, and compactness of modeling, it additionally provides perfect numerical conditioning.

B. Signal Flow Graph Connections

Referring to the two-term IIR recurrence relation in (56), assume we pick $\Upsilon_M(z) = Q_M(z)$, with $a_M = 0$. Since the input states $u_{M,N-1}$ are updated to $\breve{u}_{M,N}$ with only M operations through Ψ_M, it is obvious that applying the same structure to any arbitrary vector implies a one sample filtering step through the same network that originated Ψ_M in first place. Moreover, the inverse operation $\Psi^{-1}_M = \Psi^*_M$ can be simply realized by reversing the signal-flow graph of the original network, which also requires a one sample filtering through all inverse transfer functions $(1-a_m z^{-1})/(\delta_m z^{-1} - \bar{a}_m)$. For example, let $\breve{k}_{M,N}(m)$ denote the m-th entry of $\breve{k}_{M,N}$ defined in Table I. It is thus obtained from the entries $k_{M,N-1}(m)$ in $\mathcal{O}(M)$ operations as

$$\breve{k}_{M,N}(m-1) = \frac{1}{\delta_m}[\bar{a}_m k_{M,N-1}(m-1) - \breve{k}_{M,N}(m) + a_m k_{M,N-1}(m)] \tag{62}$$

Note that this is an overall $\mathcal{O}(M)$ operation, which therefore justifies the EGSWFTF algorithm as truly fast.

This simple arguments are a consequence of the well known theory of realization of digital filters, which has been reemerged more recently in the context of fast structured matrix operations and factorizations. This is easily seen by invoking their *dual* realizations and its defining *Horner-like* polynomials—see, for instance, [23]. The procedure for obtaining the Horner-like polynomials is known, and follows similarly the arguments that led to Ψ^{-1}_M, realized by reversing the signal flow graph directions of the original network, and identifying the Horner-like polynomials $\{\tilde{R}_k(z^{-1})\}$ as the partial transfer functions seen from the input to the tapped-delays inputs in the dual system. These dual polynomials easily realize inverses, transpositions, and matrix factorizations efficiently, and are paramount to the connections presented in this work[4]. Just like recurrence related polynomial basis, the associated Horner-like polynomials are

[4]A recent sequence of papers on structured matrices has cast real-orthogonal and Szegö polynomials as special cases of a broader class so-called $(H,1)$-*quasi-separable polynomials*, which possess efficient polynomial evaluation algorithms as an extension of the above Horner rule, as well as the Clenshaw rule for real-line orthogonal polynomials and the Ammar-Gragg-Reichel rule for Szegö polynomials [25]–[28].

$$\mathcal{R}_{\bar{\Psi}^*_M, p^*_M} = [\,p^*_M \quad \lambda^{-1/2}\bar{\Psi}^*_M p^*_M \quad \lambda^{-1}\bar{\Psi}^{*2}_M p^*_M \quad \cdots \quad \lambda^{-(M-1)/2}\bar{\Psi}^{M-1*}_M p^*_M\,].$$

$$\Psi_M = \begin{bmatrix} \rho_1\rho_0 & -\rho_2\rho_0 a^*_1 & \rho_3\rho_0(a_2a_1)^* & \cdots & \rho_{M-1}\rho_0(a_{M-2}\cdots a_2a_1)^* & -\rho_0(a_{M-1}\cdots a_2a_1)^* \\ a_1 & \rho_2\rho_1 & -\rho_3\rho_1 a^*_2 & \cdots & -\rho_{M-1}\rho_1(a_{M-2}\cdots a_2)^* & \rho_1(a_{M-1}\cdots a_2)^* \\ 0 & a_2 & \rho_3\rho_2 & \cdots & \rho_{M-1}\rho_2(a_{M-2}\cdots a_3)^* & -\rho_2(a_{M-1}\cdots a_3)^* \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \rho_{M-1}\rho_{M-2} & -\rho_{M-2}a^*_{M-1} \\ 0 & 0 & 0 & \cdots & a_{M-1} & \rho_{M-1} \end{bmatrix}$$

computed through recurrences involving any number of terms as well. In a more general case of the M-term recursion of (50), the corresponding dual polynomials satisfy the M-term relations

$$\begin{aligned}\tilde{R}_0(z) &= 1,\ R_1(z) = \tilde{\delta}_0 \phi_M z^{-1} R_0(z) \quad (63)\\ \tilde{R}_k(z) &= (\tilde{a}_{k-2,k-1} + \tilde{\delta}_{k-1} z^{-1})\tilde{R}_{k-1}(z)\\ &\quad + \tilde{a}_{k-3,k-1}\tilde{R}_{k-2}(z) + \cdots\\ &\quad + \tilde{a}_{0,k-1}\tilde{R}_1(z) + \phi_{M-k} \quad (64)\end{aligned}$$

where $\{\phi_{M-k}\}$ are the coefficients of the master polynomial of order $M+1$, i.e.,

$$\tilde{\Upsilon}_{M+1}(z) = \phi_0 Q_0(z) + \phi_1 Q_1(z) + \cdots + \phi_M Q_M(z) \quad (65)$$

with $\tilde{\delta}_k = \delta_{M-k}$ and $\tilde{a}_{k,j} = \frac{\delta_{M-j}}{\delta_{M-k}} \bar{a}_{M-j,M-k}$. In the monomial case, these polynomials, denoted by $\tilde{\mathcal{P}} = \{\tilde{h}_k(z)\}$ are simply called *Horner polynomials*, and satisfy the recursion

$$\tilde{h}_k(z) = z^{-1}\tilde{h}_{k-1}(z) + \phi_{M-k}, \quad \tilde{h}_0(z) = \phi_M. \quad (66)$$

V. Unified Displacement and Array Relations

In the following, we specify a class of operators $\{\Phi, \Gamma\}$ that will produce a low rank factorization of $P_{M,N,L}$, where its generators are explicitly defined, regardless of data structure. As a byproduct, we obtain a fast array version of the EGSWFTF. Extension to more general non-Hermitian cross-variances is straightforward.

A. Fast GSWRLS Array Algorithm

The propagation of low rank factors that define the Bezoutian in our scenario can be equivalently expressed as an array algorithm, and easily extended to a generalized window with several breakpoints. To see this, define the following quantities,

$$\begin{aligned}\tilde{P}_{M-1,N,L} &\triangleq \begin{bmatrix} P_{M-1,N,L} & 0\\ 0 & 0\end{bmatrix},\\ \tilde{k}_{M,N,L} &\triangleq \begin{bmatrix} \bar{k}_{M,N,L}\\ 0\end{bmatrix} \quad \text{and}\\ \tilde{k}^d_{M,N,L} &\triangleq \begin{bmatrix} \bar{k}^d_{M,N,L}\\ 0\end{bmatrix}, \quad (67)\end{aligned}$$

as well as the normalized vectors

$$\begin{aligned}\bar{w}^f_{M,N} &= \frac{1}{\xi^{f/2}_M(N)}\begin{bmatrix} 1\\ -w^f_{M,N}\end{bmatrix},\\ \bar{w}^b_{M,N} &= \frac{1}{\xi^{b/2}_M(N)}\begin{bmatrix} -w^b_{M,N}\\ 1\end{bmatrix}. \quad (68)\end{aligned}$$

Now, let Z_θ be the any *companion* matrix defined as (1), and consider the block decompositions shown in (69)–(70) at the bottom of the page. Before continuing, we point out that the choice of Z_θ is arbitrary. This extra degree of freedom in the choice of the vector θ will be crucial to the construction of suitable operators [see (81) and (82)], and will lead to efficient and useful representations of $P_{M,N,L}$.

Now, consider the time update for $P_{M,N,L}$

$$\begin{aligned}P_{M,N,L} = \lambda^{-1}P_{M,N-1,L} - \lambda^{-1}\bar{k}^d_{M,N-1,L-1}\bar{k}^{d*}_{M,N-1,L-1}\\ -\bar{k}_{M,N,L}\bar{k}^*_{M,N,L}. \quad (71)\end{aligned}$$

Substituting (69) and (70) into (13), and using the updating relation (71), we get

$$\begin{aligned}&\tilde{P}_{M-1,N,L} - \Psi_M^{-1} Z_\theta \tilde{P}_{M-1,N-1,L} Z^*_\varsigma \Psi_M^{-*}\\ &= \lambda^{-1}\Big[\Psi_M^{-1}\bar{w}^f_{M-1,N-2,L}\bar{w}^{f*}_{M-1,N-2,L}\Psi_M^{-*}\\ &\quad -\bar{w}^b_{M-1,N-1,L}\bar{w}^{b*}_{M-1,N-1,L}\Big]\\ &\quad - \lambda^{-1}\bar{\bar{k}}^{d_o}_{M,N}\bar{\bar{k}}^{d_o*}_{M,N}\\ &\quad - \tilde{k}_{M-1,N,L}\tilde{k}^*_{M-1,N,L}\\ &\quad - \lambda^{-1}\tilde{k}^d_{M-1,N-1,L-1}\tilde{k}^{d*}_{M-1,N-1,L-1}\\ &\quad + \Psi_M^{-1} Z_\theta\Big[\tilde{k}_{M-1,N-1,L}\tilde{k}^*_{M-1,N-1,L}\\ &\quad +\lambda^{-1}\tilde{k}^d_{M-1,N-2,L-1}\tilde{k}^{d*}_{M-1,N-2,L-1}\Big] Z^*_\varsigma\Psi_M^{-*}\\ &= \Psi_M^{-1}\bar{w}^f_{M-1,N-1,L}\bar{w}^{f*}_{M-1,N-1,L}\Psi_M^{-*}\\ &\quad - \bar{w}^b_{M-1,N,L}\bar{w}^{b*}_{M-1,N,L} - \bar{\bar{k}}^{d_o}_{M,N}\bar{\bar{k}}^{d_o*}_{M,N} \quad (72)\end{aligned}$$

which implies the following rank-5 recursion, in terms of "squared" quantities:

$$\begin{aligned}&- \Psi_M^{-1} Z_\theta\Big[\tilde{k}_{M-1,N-1,L}\tilde{k}^*_{M-1,N-1,L}\\ &\quad +\lambda^{-1}\tilde{k}^d_{M-1,N-2,L-1}\tilde{k}^{d*}_{M-1,N-2,L-1}\Big] Z^*_\varsigma\Psi_M^{-*}\\ &\quad + \lambda^{-1}\Psi_M^{-1}\bar{w}^f_{M-1,N-2,L}\bar{w}^{f*}_{M-1,N-2,L}\Psi_M^{-*}\\ &\quad - \lambda^{-1}\bar{w}^b_{M-1,N-1,L}\bar{w}^{b*}_{M-1,N-1,L}\\ &\quad - \lambda^{-1}\bar{\bar{k}}^{d_o}_{M,N}\bar{\bar{k}}^{d_o*}_{M,N}\\ &\quad = \tilde{k}_{M-1,N,L}\tilde{k}^*_{M-1,N,L}\\ &\qquad + \lambda^{-1}\tilde{k}^d_{M-1,N-1,L-1}\tilde{k}^{d*}_{M-1,N-1,L-1}\\ &\qquad + \Psi_M^{-1}\bar{w}^f_{M-1,N-1,L}\bar{w}^{f*}_{M-1,N-1,L}\Psi_M^{-*}\\ &\qquad - \bar{w}^b_{M-1,N,L}\bar{w}^{b*}_{M-1,N,L} - \bar{\bar{k}}^{d_o}_{M,N}\bar{\bar{k}}^{d_o*}_{M,N}. \quad (73)\end{aligned}$$

Equation (73) represents a norm preserving relation, which implies the array equation shown in (74) at the bottom of the

$$\begin{cases}\breve{P}_{M,N,L} = \tilde{P}_{M-1,N,L} + \bar{w}^b_{M-1,N,L}\bar{w}^{b*}_{M-1,N,L} & (69)\\ P_{M,N-1,L} = Z_\theta\tilde{P}_{M-1,N-1,L}Z^*_\varsigma + \bar{w}^f_{M-1,N-1,L}\bar{w}^{f*}_{M-1,N-1,L}, & (70)\end{cases}$$

page, where $\varnothing_{M,N}$ is a *J-unitary* matrix defined explicitly in terms of the square-roots of the minimum costs, likelihood variables, and prediction quantities involved in the EGSWFTF algorithm, with $J = \mathrm{diag}(1,-1,1,-1,-1)$. It can be obtained explicitly from the normalized version of updates and downdates presented in the previous section, which in turn define (implicitly) the hyperbolic rotations characterizing these recursions in array form. For instance, (21) can be normalized by (24), yielding

$$\bar{k}^d_{M,N-1,L-1} = \frac{\xi^{f/2}_{M-1,L}(N-1)}{\xi^{f/2}_{M-1,L-1}(N-1)} Z_\theta \tilde{k}^d_{M-1,N-1,L-1} - \lambda^{L-1}\eta_0 f^{d''}_{M-1,L-1}(N-1)\bar{w}^f_{M-1,N-1,L} \tag{75}$$

and

$$f^{d''}_{M-1,L-1}(N-1) \triangleq \frac{f_{M-1,L-1}(N-1)}{\gamma^{d/2}_{M-1,L-1}(N-1)\xi^{f/2}_{M-1,L-1}(N-1)}. \tag{76}$$

Similarly, defining

$$f''_{M-1,L}(N-1) \triangleq \frac{f_{M-1,L}(N-1)}{\gamma^{1/2}_{M-1,L}(N-1)\xi^{f/2}_{M-1,L}(N-1)} \tag{77}$$

gives

$$\bar{k}_{M,N-1,L} = \frac{\lambda^{1/2}\xi^{f/2}_{M-1,L-1}(N-2)}{\xi^{f/2}_{M-1,L}(N-1)} Z_\theta \tilde{k}_{M-1,N-1,L} + \lambda^{-1/2} f''_{M-1,L}(N-1)\bar{w}^f_{M-1,N-2,L-1}. \tag{78}$$

The normalized forward prediction and normalized quantities are row-downdated as

$$\bar{w}^f_{M-1,N-2,L-1} = \frac{\gamma^{d/2}_{M,L-1}(N-2)}{\gamma^{d/2}_{M-1,L-1}(N-2)} \bar{w}^f_{M-1,N-2,L} - \eta_0\lambda^L Z_\theta \tilde{k}^d_{M-1,N-2,L-1}, f^{d''}_{M-1,L-1}(N-2) \tag{79}$$

and time-updated as

$$\bar{w}^f_{M-1,N-1,L} = \frac{\gamma^{1/2}_{M,L}(N-1)}{\lambda^{1/2}\gamma^{1/2}_{M-1,L}(N-1)} \bar{w}^f_{M-1,N-2,L-1} - Z_\theta \tilde{k}_{M-1,N-1,L} f''_{M-1,L}(N-1). \tag{80}$$

Analogously, we can easily write normalized *backward prediction* update and downdate counterparts; Rearrangement of these relations and some algebra will lead to the explicit structure of $\varnothing_{M,N}$, which is not our main concern here.

B. Displacement Equations

The array relations of the previous section can now be combined with the order updates of the covariances $P_{M,N-1,L}$ and $\breve{P}_{M,N,L}$ of (69) and (70), yielding the following general result:

Theorem 1 (Displacement Equation of Arbitrarily Structured Covariances): Consider a matrix Ψ_M constructed from any first-order recurrence relation. Let $\{Z_\theta, Z_\varsigma\}$ be any arbitrary pair of companion matrices. Then, the following compositions for displacement operators

$$\left\{\Phi_{M,\theta} \triangleq \Psi_M Z_\theta^{-1},\, \Phi_{M,\varsigma} \triangleq \Psi_M Z_\varsigma^{-1}\right\}, \tag{81}$$

and

$$\left\{\breve{\Phi}_{M,\theta} \triangleq Z_\theta^{-1}\Psi_M,\, \breve{\Phi}_{M,\varsigma} \triangleq Z_\varsigma^{-1}\Psi_M\right\} \tag{82}$$

satisfy the displacement equations with respect to the coefficient matrices $P_{M,N,L}$ and $\breve{P}_{M,N,L}$:

$$\begin{aligned}
&\nabla_{\{\Phi^{-1}_{M,\theta},\Phi^{-*}_{M,\varsigma}\}}(P_{M,N,L}) \\
&\triangleq P_{M,N,L} - \lambda\Phi^{-1}_{M,\theta}P_{M,N,L}\Phi^{-*}_{M,\varsigma} \\
&= -Z_\theta \breve{\bar{k}}^{d_o}_{M,N}\breve{\bar{k}}^{d_o*}_{M,N}Z^*_\varsigma + \bar{w}^f_{M-1,N-1,L}\bar{w}^{f*}_{M-1,N-1,L} \\
&\quad - \lambda Z_\theta \bar{w}^b_{M-1,N,L}\bar{w}^{b*}_{M-1,N,L}Z^*_\varsigma \\
&\quad + \Phi^{-1}_{M,\theta}\Big[\tilde{k}^d_{M-1,N-1,L-1}\tilde{k}^{d*}_{M-1,N-1,L-1} \\
&\quad + \lambda\tilde{k}_{M-1,N,L}\tilde{k}^*_{M-1,N,L}\Big]\Phi^{-*}_{M,\varsigma}
\end{aligned} \tag{83}$$

and

$$\begin{aligned}
&\nabla_{\{\breve{\Phi}_{M,\theta},\breve{\Phi}^*_{M,\varsigma}\}}(\breve{P}_{M,N,L}) \\
&\triangleq \breve{P}_{M,N,L} - \lambda^{-1}\breve{\Phi}_{M,\theta}\breve{P}_{M,N,L}\breve{\Phi}^*_{M,\varsigma} \\
&= \bar{w}^b_{M-1,N,L}\bar{w}^{b*}_{M-1,N,L} \\
&\quad - \lambda^{-1}Z_\theta^{-1}\bar{w}^f_{M-1,N-1,L}\bar{w}^{f*}_{M-1,N-1,L}Z^{-*}_\varsigma \\
&\quad - \lambda^{-1}\tilde{k}^d_{M-1,N-1,L-1}\tilde{k}^{d*}_{M-1,N-1,L-1} \\
&\quad - \tilde{k}_{M-1,N,L}\tilde{k}^*_{M-1,N,L} \\
&\quad + \lambda^{-1}\breve{\Phi}_{M,\theta}\breve{\bar{k}}^{d_o}_{M,N-1}\breve{\bar{k}}^{d_o*}_{M,N-1}\breve{\Phi}^*_{M,\varsigma}
\end{aligned} \tag{84}$$

respectively.

Proof: From the normalized fast recursions, substituting $Z_\theta\tilde{P}_{M-1,N,L}Z^*_\varsigma = P_{M,N,L} - \bar{w}^f_{M-1,N,L}\bar{w}^{f*}_{M-1,N,L}$ into (72), and using (71), we obtain (83). Likewise, substituting (69) into (72), we obtain (84).

Note that in the adaptive filtering context, by setting $u = 0$, we obtain a rank-4 displacement rank, regardless of the structure induced in Ψ_M. Also, both displacement equations (83) and (84) can be used to represent the decomposition of a certain inverse covariance $P_{M,N,L}$, since one can always interchange the roles of $\breve{P}_{M,N,L}$ and $P_{M,N,L}$ with a suitable choice for the first and last columns of their corresponding data matrices.

$$\begin{bmatrix} \frac{1}{\sqrt{\lambda}}\Psi^{-1}_{M+1}Z_\theta\tilde{k}^d_{M,N-2,L-1} & \Psi^{-1}_{M+1}Z_\theta\tilde{k}_{M,N-1,L} & \frac{1}{\sqrt{\lambda}}\Psi^{-1}_{M+1}\bar{w}^f_{M,N-2,L} & \frac{1}{\sqrt{\lambda}}\bar{w}^b_{M,N-1,L} & \frac{1}{\sqrt{\lambda}}\breve{\bar{k}}^{d_o}_{M,N}\end{bmatrix}\varnothing_{M,N} = \begin{bmatrix}\frac{1}{\sqrt{\lambda}}\tilde{k}^d_{M,N-1,L-1} & \tilde{k}_{M,N,L} & \Psi^{-1}_{M+1}\bar{w}^f_{M,N-1,L} & \bar{w}^b_{M,N,L} & \breve{\bar{k}}^{d_o}_{M,N}\end{bmatrix} \tag{74}$$

VI. Solution of the Displacement Equation and Frequency Domain Representation for Arbitrary Operators

In this section, we show that an eigenvector based representation for the time-varying covariance $P_{M,N,L}$ is not limited to tapped-delay-lines, but holds for any first order induced operator. This result brings two significant consequences:

1. It paves the way to efficient representation of Bezoutians based on *polynomial Vandermonde* matrices arising from recurrence-related models; This in turn allows the development of superfast representations of such general covariances, corresponding to other than the ones originated from tapped-delay line models [2].
2. It will provide as a corollary, the minimality condition stated it terms of the generators of $P_{M,N,L}$ *as a vector relation*, as opposed to the generating function approach normally seen in prewindowed RLS algorithms. We remark that this holds for a generalized window and for any data model that induces a displacement equation.

Lemma 1 (Eigenvector Representation of Covariance Bezoutians): Let $P_{M,N,L}$ be the inverse covariance matrix arising in a generalized window LS formulation for an arbitrary data model induced by Ψ_M, and assume that $\{Z_\theta, Z_\varsigma\}$ are chosen such that the eigenvalues of $\{\Phi_{M,\theta}, \Phi_{M,\varsigma}\}$ in (81) are the M roots of arbitrary scalars $-1/\phi_0$ and $-\varrho_0 \in \mathbb{C}$ respectively, i.e., $\bar{z}_1(m) = \phi e^{j\frac{2\pi m}{M}}$, where $\phi = |\phi_0|^{-1/M} e^{j(\angle -\phi_0^{-1})/M}$, and $\bar{z}_2(m) = \varrho e^{j\frac{2\pi m}{M}}$, where $\varrho = |\varrho_0|^{1/M} e^{j(\angle -\varrho_0^{-1})/M}$, with $\phi_0 \varrho_0^* \neq \lambda^{-M}$. Let $\{V_1, V_2\}$ be their corresponding eigenvector matrices, and define $\bar{D}_{\phi\varrho^*} \triangleq \operatorname{diag}(\{\phi\varrho^*/\lambda^m\}_{m=0}^{M-1})$. Then, $P_{M,N,L}$ can be factored as

$$
\begin{aligned}
P_{M,N,L} = &\frac{M}{1-\lambda^M \phi_0 \varrho_0^{-*}} V_1^{-1} \\
&\times \Big(\Lambda_{\phi,\bar{w}^f_{M,N-1,L}} F_M \bar{D}_{\phi\varrho^*} F_M^* \Lambda^*_{\varrho,\bar{w}^f_{M,N-1,L}} \\
&- \lambda \Lambda_{\phi,Z_\theta \bar{w}^b_{M,N,L}} F_M \bar{D}_{\phi\varrho^*} F_M^* \Lambda^*_{\varrho,Z_\varsigma \bar{w}^b_{M,N,L}} \\
&+ \Lambda_{\phi,\Phi^{-1}_{M,\theta}\tilde{k}^d_{M,N-1,L}} F_M \bar{D}_{\phi\varrho^*} F_M^* \Lambda^*_{\varrho,\Phi^{-1}_{M,\varsigma}\tilde{k}^d_{M,N-1,L}} \\
&+ \lambda \Lambda_{\phi,\Phi^{-1}_{M,\theta}\tilde{k}_{M,N,L}} F_M \bar{D}_{\phi\varrho^*} F_M^* \Lambda^*_{\varsigma,\Phi^{-1}_{M,\varsigma}\tilde{k}_{M,N,L}} \\
&- \Lambda_{\phi,Z_\theta \bar{\bar{k}}^{d_o}_{M,N}} F_M \bar{D}_{\phi\varrho^*} F_M^* \Lambda^*_{\varrho,Z_\varsigma \bar{\bar{k}}^{d_o}_{M,N}} \Big) V_2^{-*}
\end{aligned} \tag{85}
$$

where F_M is the DFT matrix and

$$
\begin{aligned}
\Lambda_{\phi,\bar{w}^f_{M,N-1,L}} &= \operatorname{diag}\left(V_1 \bar{w}^f_{M,N-1,L}\right), \\
\Lambda_{\phi,Z_\theta \bar{w}^b_{M,N}} &= \operatorname{diag}\left(V_1 Z_\theta \bar{w}^b_{M,N,L}\right),
\end{aligned} \tag{86}
$$

$$
\begin{aligned}
\Lambda_{\phi,\Phi^{-1}_{M,\theta}\tilde{k}^d_{M,N-1,L}} &= \operatorname{diag}\left(V_1 \Phi^{-1}_{M,\theta} \tilde{k}^d_{M,N-1,L}\right), \\
\Lambda_{\phi,\Phi^{-1}_{M,\theta}\tilde{k}_{M,N,L}} &= \operatorname{diag}\left(V_1 \Phi^{-1}_{M,\theta} \tilde{k}_{M,N,L}\right),
\end{aligned} \tag{87}
$$

with analogous definitions for the variables with dependency on ς.

Proof: See Appendix.

Corollary (Minimality Condition in the V_1-Domain): The vectors $\{\bar{w}^f_{M,N-1,L}, Z_\theta \bar{w}^b_{M,N,L}, \Phi^{-1}_{M,\theta}\tilde{k}_{M,N,L}, \Phi^{-1}_{M,\theta}\tilde{k}^d_{M,N-1,L}, Z_\theta \bar{\bar{k}}^{d_o}_{M,N}\}$ which define $P_{M,N,L}$ are related in the V_1-domain as follows:

$$
\begin{aligned}
\Lambda_{\phi,Z_\theta \bar{w}^b_{M,N,L}} \Lambda^*_{\varrho,Z_\varsigma \bar{w}^b_{M,N,L}} &+ \lambda^{-1} \Lambda_{\phi,Z_\theta \bar{\bar{k}}^{d_o}_{M,N}} \Lambda^*_{\varrho,Z_\varsigma \bar{\bar{k}}^{d_o}_{M,N}} \\
= \lambda^{-1} &\Lambda_{\phi,\bar{w}^f_{M,N-1,L}} \Lambda^*_{\varrho,\bar{w}^f_{M,N-1,L}} \\
&+ \Lambda_{\phi,\Phi^{-1}_{M,\theta}\tilde{k}_{M,N,L}} \Lambda^*_{\varrho,\Phi^{-1}_{M,\varsigma}\tilde{k}_{M,N,L}} \\
&+ \lambda^{-1} \Lambda_{\phi,\Phi^{-1}_{M,\theta}\tilde{k}^d_{M,N-1,L}} \Lambda^*_{\varrho,\Phi^{-1}_{M,\varsigma}\tilde{k}^d_{M,N-1,L}}
\end{aligned} \tag{88}
$$

where $\Lambda_{\cdot,\cdot}$ are defined in (86) and (87).

Proof: Choosing $\varrho_0 = \lambda^{M/2}/\phi_0^*$ in (A-4), yields $\bar{D}_{\phi\varrho^*} = I$, and (88) follows.

A special case of the above relation appeared in [42], [44] in a pure matrix algebra scenario for $\lambda = 1$, and in the simple monomial basis case; Interestingly, in the latter, such relation has never been linked to its reproducing kernel counterpart in the context of RLS problems, for which it becomes a rank-3 relation (see, e.g., [10] and the references therein). Here, we further express the spectral factorization aforementioned in terms of the actual vector quantities, instead of generating functions.

Equation (A-4) brings a generalization with respect to Bezoutians representation for arbitrary polynomial basis, which holds even for infinite impulse response basis functions. Equation (88) provides a key result to all fast RLS algorithms for generalized window and data structures, and shows explicitly, how the normalized backward prediction filter is related to the normalized forward prediction and Kalman gain vectors in *any V_1-domain*. In particular, for tapped-delay-line models, $\{Z_\theta, Z_\varsigma\}$ become $\{\theta_0 = -\phi_0, \varphi_0 = -\varrho_0\}$—factor circulants, respectively, so that choosing $\phi_0 = \varrho_0 = 1$, the vector $\bar{w}^b_{M,N,L}$ is obtained via spectral factorization.

We highlight that at this point, the eigenvector factorization (85) assumes that operators with the properties $\Phi^M_{M,\theta} = -1/\phi_0 I, \Phi^M_{M,\varsigma} = -\varrho_0 I$ can be constructed, and implicitly, that $\{V_1, V_2\}$ are efficient transformations. In [2], we show that recurrence related polynomials along with the existing degree of freedom in the companion structure Z_θ in fact allows us to achieving this condition.

VII. Minimal Representation in the EGSWFTF

The error propagation in fast RLS algorithms is originated in the backward prediction part of the recursions. This can be understood by representing the propagated quantities of the prediction section as the states $\mathbf{x}(N)$ of a nonlinear system, say

$$
\mathbf{x}(N+1) = \mathcal{F}[\mathbf{x}(N), s(N+1)] \tag{89}
$$

where $s(N)$ is the input signal, and $\mathcal{F}$ is a memoryless nonlinearity that depends on the algorithm used. In the case of the GSWFTF algorithm, the states are

$$
\begin{aligned}
\mathbf{x}(N) = \Big\{ &w^f_{M-1,N-1,L}, w^b_{M-1,N,L}, \xi^f_{M-1,L}(N-1), \\
&\xi^b_{M-1,L}(N), \bar{k}^d_{M-1,N-1,L-1}, \bar{k}_{M-1,N,L}, \\
&\gamma^d_{M-1,L-1}(N-1), \gamma_{M-1,L}(N) \Big\}.
\end{aligned} \tag{90}
$$

Now consider the perturbed system

$$
\mathbf{x}'(N+1) = \mathcal{F}[\mathbf{x}'(N), s(N+1)] + \Delta(N) \tag{91}
$$

where $\Delta(N)$ is due to quantization. The state error $\mathbf{x}(N) - \mathbf{x}'(N)$ will remain bounded if (89) is exponentially stable for all states $\mathbf{x}(N)$ contained in a certain stability region $S_i(N)$ (the solution manifold), and if the perturbation Δ does not push $\mathbf{x}'(N)$ outside $S_i(N)$. Now let $S_f(N)$ be the stability domain of the perturbed system (91). An algorithm is said to be *backward consistent* if the computed solution of a problem is the exact solution to a perturbed problem. The procedure for stability analysis is to check if $S_f(N) \subset S_i(N)$ for all N, in which case its time recursions will be exponentially stable (see [39]).

The answer to whether the fast fixed-order RLS filters are stable or not relies on the fact that these represent systems with *non-minimal* dimension, in which case $S_f(N) \subset S_i(N)$ as shown in [39], [41] for the FIR case, and more generally in [10] for orthonormal models. We would like to define a similar stability domain $S_i(N)$ for the GSWFTF algorithm considering general operators, by specifying the minimal components of the state vector $\mathbf{x}(N)$.

The first (main) minimality condition was already stated in (88), where $\bar{w}^b_{M,N,L}$ is seen uniquely determined from $\{\bar{w}^f_{M,N-1,L}, Z_\theta \bar{w}^b_{M,N,L}, \Phi^{-1}_{M,\theta}\tilde{k}_{M,N,L}, \Phi^{-1}_{M,\theta}\tilde{k}^d_{M,N-1,L}, Z_\theta \breve{\bar{k}}^{d_o}_{M,N}\}$ as the spectral factor of the right-hand-side of this equation, that has all its zeros outside the unit circle and all its poles inside the unit circle[5]. The quantity $\xi^b_M(N)$ is inferred by normalizing the last entry of $\bar{w}^b_{M,N,L}$ to unity (in the orthonormal basis representation). Now, in order to completely characterize the minimal components of the GSWFTF algorithm, we further need to establish one last relation.

Equations (28) and (31) can be combined by using the fact that $\gamma_{M+1,L}(N) = \breve{\gamma}_{M+1,L}(N-1)$ to yield

$$\gamma_{M,L}(N) = \gamma_{M,L}(N-1)\frac{\xi^f_{M,L-1}(N-2)}{\xi^f_{M,L}(N-1)}\frac{\xi^b_{M,L}(N)}{\xi^b_{M,L-1}(N-1)}. \tag{93}$$

Likewise, (24) and (32) can be combined, since $\gamma^d_{M+1,L-1}(N) = \breve{\gamma}^d_{M+1,L-1}(N-1)$. This gives

$$\frac{\xi^f_{M,L-1}(N-2)}{\xi^b_{M,L-1}(N-1)} = \frac{\gamma^d_{M,L-1}(N-2)}{\gamma^d_{M,L-1}(N-1)}\frac{\xi^f_{M,L}(N-2)}{\xi^b_{M,L}(N-1)}, \tag{94}$$

[5]The minimality condition in the adaptive filtering case is usually stated via a reproducing kernel relation, i.e.,

$$P(z_1, z_2^*) = \frac{1}{1 - 1/z_1 z_2^*}[\bar{W}^f(z_1)\bar{W}^{f*}(z_2) - \bar{W}^b(z_1)\bar{W}^{b*}(z_2) + \bar{K}(z_1)\bar{K}^*(z_2) + \bar{K}^d(z_1)\bar{K}^{d*}(z_2)], \tag{92}$$

where the bivariate polynomial $P(z_1, z_2^*)$ extends to general basis functions, the *Christoffel-Darboux* formula [34] originally intended to Toeplitz matrices (see also [10]). The generating functions $\{W^{f*}(z_2), W^{b*}(z_2), \bar{K}^{d*}(z_2), \bar{K}^*(z_2)\}$ are obtained similarly from the forward and backward prediction quantities and Kalman gains. It has been the key to demonstrate (here written according to our new algorithm solution) that, when $L = N$, in which case $\bar{K}^d(z) = 0$, only $2M - 1$ degrees of freedom are necessary to represent $P_{M,N,L} = P_{M,N}$, a crucial observation for the theory of fast RLS algorithms. Until today, with the exception of the fast QR-RLS based on the Gray Markel lattice structure of [40], all fast RLS algorithms are known to be *non-minimum*, meaning that, for an order M filter, these versions do not enforce the condition of (88), and therefore propagate redundant variables in representing $P_{M,N}$. This extra redundancy causes all fast RLS recursions to eventually become unstable in the light of backward consistency and minimality concepts discussed in [39], [41].

so that substituting (94) into (93) we obtain

$$\frac{\gamma_{M,L}(N)}{\gamma_{M,L}(N-1)} = \frac{\gamma^d_{M,L-1}(N-2)}{\gamma^d_{M,L-1}(N-1)}\frac{\xi^b_{M,L}(N)}{\xi^b_{M,L}(N-1)} \times \frac{\xi^f_{M,L}(N-2)}{\xi^f_{M,L}(N-1)} \tag{95}$$

Finally, solving for $\gamma_{M,L}(N)$, we arrive at

$$\xi^b_{M,L}(N) = \sigma_{M,L}\gamma_{M,L}(N)\gamma^d_{M,L-1}(N-1)\xi^f_{M,L}(N-1), \tag{96}$$

where $\sigma_{M,L} = \frac{\xi^b_{M,L}(-1)}{\gamma^d_{M,L-1}(-2)\xi^f_{M,L}(-2)}$. Observe that (96) generalizes the well known relation $\xi^b_M(N) = \lambda^{-M}\xi^f_M(N-1)\gamma_M(N)$, obtained for shift-data structures and exponentially weighted window.

Equations (88) and (96) imply that in fact only $3M - 1$ degrees of freedom are needed to represent the LS solution. That is, the set S_i is represented by the variables such that (*i*) the spectral factorization (88) is verified with respect to any arbitrary basis functions, (*ii*) with $\bar{w}^b_{M,N,L}$ having all its zeros outside the unit circle and all poles inside the unit circle and (*iii*) the likelihood variable $\gamma_{M-1,L}(N)$ obtained from (96), satisfies $0 \le \gamma_{M-1,L}(N) < 1$. In this case, the minimal components of the state vector $\mathbf{x}(N)$ are

$$\Big\{w^f_{M-1,N-1,L}, \xi^f_{M-1,L}(N-1), \bar{k}^d_{M-1,N-1,L-1}, \bar{k}_{M-1,N,L}, \gamma^d_{M-1,L-1}(N-1)\Big\}$$

and the error between the actual and computed quantities $\{w^b_{M-1,N,L}, \xi^b_{M-1,L}(N), \gamma_{M-1,L}(N)\}$ is interpreted as a perturbation that leads the state outside the stability domain S_i.

The *rescuing mechanism* proposed in [10] represent a rough way of projecting the state $\mathbf{x}'(N)$ back onto the manifold S_i, once $\gamma_{M-1,L}(N)$ becomes negative. It amounts to monitoring the quantity (we describe the procedure here for the EGSWFTF algorithm) $1 - \breve{\gamma}_{M,L}(N)\beta_{M-1,L}(N)\nu^*_{M-1,L}(N)$. If it is positive, the algorithm continues its flow. Otherwise, we restart the algorithm as follows (*Rescuing Procedure*):

$$\begin{cases} u_{M,N} = 0, \quad w_{M,N} = w_{M,N-1} \\ \gamma^d_{M-1,L-1}(N-1) = -\lambda^{L-1}\eta_0; \ \gamma_{M-1,L}(N) = 1 \\ \xi^b_{M-1,L}(N) = -\lambda^{L-1}\eta_0\sigma_{M-1,L}\xi^f_{M-1,L}(N-1) \\ w^f_{M-1,-2,L} = \Pi_{M-1}\bar{c}_{M-1}; \ w^b_{M-1,-1,L} = \Pi_{M-1}\breve{c}_{M-1} \end{cases}$$

Recall that the original rescue mechanism proposed in [9] re-initializes the backward prediction minimum cost as $\xi^b_{M-1,L}(N) = \xi^b_{M-1,L}(-1)$ instead.

We remark that the original approach proposed for addressing the stability problem of the FTF algorithm was extended in [30] (see the references therein) for the GSWFTF algorithm, considering shift data. Although the resulting algorithm is claimed to be numerically stable, the method of analysis employed and the corresponding stabilized solution are valid only under some restrictive conditions, and instability can still occur in practice.

VIII. Simulations

Consider a *transversal* system of an orthonormal IIR realization with basis functions $\{Q_m(z)\}$ given by (61) which is illustrated in Fig. 4. In this section, we verify the ability of

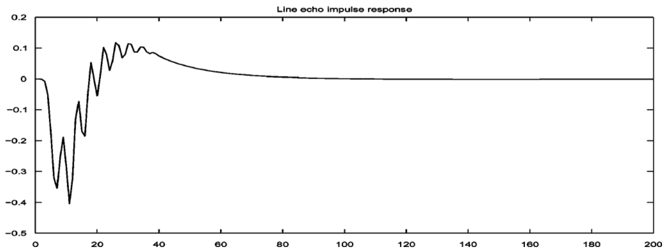

Fig. 5. Line echo path, 200 samples.

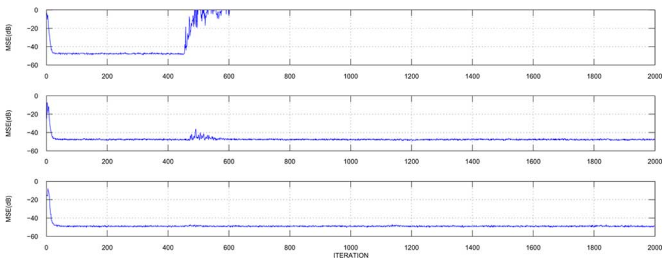

Fig. 6. EGSWFTF for $M = 8$ taps and $L = 20$ setup: (a) with no rescuing ($\eta_0 = 1$); (b) with rescuing ($\eta_0 = 1$); (c) with rescuing ($\eta_0 = 0.3$).

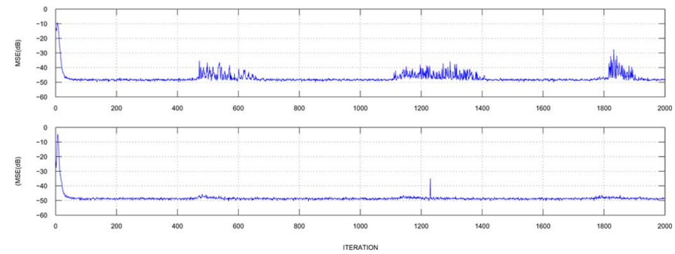

Fig. 7. ESWFTF with rescuing for $M = 10$ taps and $L = 8$. (a) ($\eta_0 = 0.7$); (c) ($\eta_0 = 0.3$).

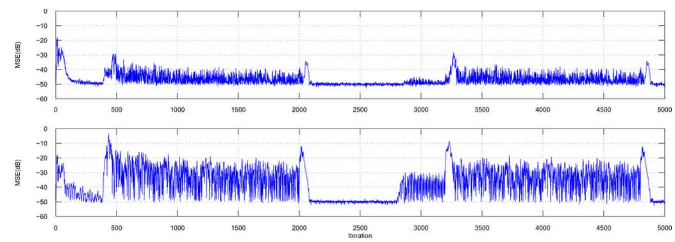

Fig. 8. Comparison between the GSWFTF for the FIR and for a Laguerre model with pole $a = 0.3$, considering a $L = 30$, $M = 45$ taps and $\eta_0 = 0.7$.

the rescue mechanism developed to the proposed extended GSWFTF algorithm to maintain stability, considering this compact representation.

⧫ *Experiment 1 (Comparison with the SWFTF and effect of rescuing)*: We test the new adaptive filter with $M = 8$ taps, $L = 20$ and $\lambda = 0.95$ in Matlab® precision, and considering a well behaved first order autoregressive input with pole at 0.9. From our discussion in Section IV.A, as long as $a_i < \sqrt{\lambda}$, the structure of Ψ_M guarantees that the Lyapunov equation (42) for the initialization of the GSWFTF in Table I has a unique solution. We thus set the poles values randomly, such that this condition is satisfied, and considered first an exact modeling scenario. As can be seen in Fig. 6, after 450 iterations, the EGSWFTF without rescue mechanism becomes unstable (a). For the sake of comparison, we plot in Fig. 6(b) the learning curve of the conventional (rectangular) SWFTF employing a rescuing mechanism, which corresponds to $\eta_0 = 1$. As we can see, the SWFTF exhibits numerical problems even in the case of rescuing. Note that while both extended SWFTF and the EGSWFTF algorithms have the same number of minimal states, the generalized window shape contributes to better conditioning (Although not shown here, this can be observed even in the *lattice* implementation of the SWRLS algorithm, when compared to the GSWFTF). The MSE performance of the EGSWFTF with $\eta_0 = 0.3$ is depicted in Fig. 6(c). We observe that the existence of a window "tail" beyond L past input samples helps in the stability of the FTF recursions, so that the use of rescuing in the latter case makes the EGSWFTF more robust in finite precision compared to the one employed in the standard SWFTF.

⧫ *Experiment 2 (Effect of different η_0 for L and M under rescuing)*: In this experiment, we examine the stability performance of the EGSWFTF for a short window length, say, $L = 8$, and $M = 10$-tap orthonormal model filter. Fig. 7 illustrates the EGSWFTF for two different cutoff levels, $\eta_0 = 0.7$ and $\eta_0 = 0.3$. Note that in the former, the performance approaches the one of a SWFTF, which as we expect, will show quick divergence, specially for $L < M$. Note that the choice $\eta_0 = 0.3$ allows a better tradeoff, in the sense that even in the case of a short window, the EGSWFTF with rescuing shows stability improvement.

⧫ *Experiment 3 (Compactness of Modeling of the EGSWFTF)*: Finally, the main purpose for using the orthonormal IIR basis of (61) is 1) the ability to represent long impulse response systems by replacing an adaptive FIR filter with one employing a reduced order EGSWFTF realization; 2) to maintain the numerical conditioning of the input signal, given that it remains unchanged with respect to a monomial basis. Hence, we compare the learning curves considering the identification of the line echo path of Fig. 5, for the EGSWFTF algorithm based on an $M = 45$-tap FIR model, i.e., for $Q_m(z) = z^{-m}$, and the one based on the orthonormal IIR realization of same order. For simplicity, we set all poles of the network as $a_m = a = 0.3$, after a few trial and error attempts to optimize its value. An optimization method for the poles choice is beyond the scope of this work, and will be addressed in a forthcoming publication. The input to the network is a composite source signal (CSS), which can be considered a ill-behaved input to the EGSWFTF. Chances that divergence occurs in this case are much higher, and the system is also slightly undermodeled, since most of the path energy is concentrated in the first 60 taps. Fig. 8 illustrates the learning curves of the EGSWFTF for $\lambda = 0.999$ and $\eta_0 = 0.7$. We clearly see that, under the same model order, the improved robustness of the FTF filter employing the *Laguerre* expansion (top) when compared to the FIR filter (bottom).

IX. Summary and Concluding Remarks

We have moved beyond the existing fast exponentially weighted RLS algorithms for extended models, and showed that generalized sliding window counterparts are also feasible. Different choices for the operator Ψ_M may lead to better numerical conditioning, reduced computational complexity, and compact representation of models.

The contributions herein include explicit fast transversal recursions for extended models and generalized window and new updates not available for shift-data structure, along with the minimality condition stated explicitly in terms of the EGSWFTF quantities. The latter is obtained from the covariance Bezoutian decomposition in exact correspondence with the GSWFTF state variables, also yielding a *rescue mechanism* as in the exponentially weighted RLS case. The generalized window, which combines the robustness of the APA and the convergence speed of RLS recursions can now be utilized in the context of extended models, as for instance, IIR orthonormal basis functions, which is specially useful in non-stationary environments.

While the EGSWFTF recursions are adaptive, they exert direct impact on the computation of *non-adaptive* scalar and block transmission equalization techniques, which can be formulated by solving the displacement equation with respect to particular bases [4]. The usefulness of changing basis representation stems from compactness of models and efficient superfast realizations, for which the computation of the displacement generators in connection with the Kalman recursions in both cases was unavailable, even for tapped-delay-line models.

We highlight that such DFT Bezoutian representations can be achieved regardless of the input basis that yields Ψ_M, due to the additional degree of freedom provided by the companion matrices $\{Z_\theta, Z_\varsigma\}$. This will be particularly discussed in a sequel of this work, where we shall extend the DFT expressions obtained herein to arbitrary transformations, by pointing out close connections with recurrence related polynomials and new general Bezoutian representations. More specifically, arguments show how the choice of free companion structures along with recurrence related basis representation yields an exact filterbank decomposition, from the solution of the corresponding displacement equation. Proper choices for the pair $\{\breve{\Phi}_{M,\theta}, \breve{\Phi}_{M,\varsigma}\}$ will thus lead to representations of highly structured inverses, extending the standard DFT formulas to more general transformations. We showed that the minimality condition holds even for arbitrary domain transformations, and provides a vector relation among the generators of the structured inverse.

APPENDIX

Iterating (83) $M-1$ times, we obtain

$$\begin{aligned} P_{M,N,L} = \lambda^M \Phi_{M,\theta}^{-M} P_{M,N,L} \Phi_{M,\varsigma}^{-M*} &- \sum_{m=0}^{M-1} \lambda^m \Phi_{M,\theta}^{-m} Z_\theta \breve{\bar{k}}^{d_o}_{M,N} \breve{\bar{k}}^{d_o*}_{M,N} Z_\varsigma^* \Phi_{M,\varsigma}^{-m*} \\ &+ \sum_{m=0}^{M-1} \lambda^m \Phi_{M,\theta}^{-m} \bar{w}^f_{M,N-1,L} \bar{w}^{f*}_{M,N-1,L} \Phi_{M,\varsigma}^{-m*} \\ &- \sum_{m=0}^{M-1} \lambda^{m+1} \Phi_{M,\theta}^{-m} Z_\theta \bar{w}^b_{M,N,L} \bar{w}^{b*}_{M,N,L} Z_\varsigma^* \Phi_{M,\varsigma}^{-m*} \\ &+ \sum_{m=0}^{M-1} \lambda^m \Phi_{M,\theta}^{-m-1} \tilde{k}^d_{M-1,N,L-1} \tilde{k}^{d*}_{M-1,N,L-1} \Phi_{M,\varsigma}^{-m-1*} \\ &+ \sum_{m=0}^{M-1} \lambda^{m+1} \Phi_{M,\theta}^{-m-1} \tilde{k}_{M-1,N,L} \tilde{k}^*_{M-1,N,L} \Phi_{M,\varsigma}^{-m-1*}. \end{aligned} \quad \text{(A-1)}$$

Now, observe that all terms inside the summations on the r.h.s. of (A-1) are of the form $\bar{\Phi}_{M,\theta}^{-m} bb^* \bar{\Phi}_{M,\varsigma}^{-m*}$, for some column vector b, and let $\bar{W}_M = \operatorname{diag}(1, \lambda, \lambda^2, \ldots, \lambda^{M-1})$. Then, in terms of the eigenvector matrices $\{V_1, V_2\}$, it can be written as

$$\begin{aligned} &\sum_{m=0}^{M-1} \lambda^m V_1^{-1} \Lambda_{z_1}^{-m} V_1 bb^* V_2^* \Lambda_{z_2}^{-m*} V_2^{-*} \\ &\qquad = V_1^{-1} \Lambda_{\phi,b} V_{\mathcal{P}}(\bar{z}_1) \bar{W}_M V_{\mathcal{P}}^*(\bar{z}_2) \Lambda_{\varrho,b}^* V_2^{-*} \end{aligned} \quad \text{(A-2)}$$

for $\Lambda_{\phi,b} = \operatorname{diag}(V_1 b), \Lambda_{\varrho,b} = \operatorname{diag}(V_2 b)$ and $\{V_{\mathcal{P}}(\bar{z}_1), V_{\mathcal{P}}(\bar{z}_2)\}$ are the monomial Vandermonde matrices, e.g.,

$$V_{\mathcal{P}}(\bar{z}_1) = \begin{bmatrix} 1 & z_1^{-1}(0) & \cdots & z_1^{-(M-1)}(0) \\ 1 & z_1^{-1}(1) & \cdots & z_1^{-(M-1)}(1) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & z_1^{-1}(M-1) & \cdots & z_1^{-(M-1)}(M-1) \end{bmatrix}, \quad \text{(A-3)}$$

where $\bar{z}_1(m) = \phi e^{j\frac{2\pi m}{M}}$, with $\phi = |\phi_0|^{-1/M} e^{j\frac{\angle -\phi_0^{-1}}{M}}$, and $\bar{z}_2(m) = \varrho e^{j\frac{2\pi m}{M}}$, with $\varrho = |\varrho_0|^{1/M} e^{j\frac{\angle -\varrho_0^{-1}}{M}}$, denote the eigenvalues of $\Phi_{M,\theta}^{-1}$ and $\Phi_{M,\varsigma}^{-1}$ respectively. Because it is well known that $V_{\mathcal{P}}(\bar{z}_1) = \sqrt{M} F_M D_\phi, D_\phi \triangleq \operatorname{diag}(\{\phi^{-m}\}_{m=0}^{M-1})$, with analogous expressions for $V_{\mathcal{P}}(\bar{z}_2)$, the r.h.s. of (A-2) becomes

$$M V_1^{-1} \Lambda_{\phi,\bar{w}^f_{M,N-1,L}} F_M D_\phi \bar{W}_M D_{\varrho^*} F_M^* \Lambda^*_{\varrho,\bar{w}^f_{M,N-1,L}} V_2^{-*}.$$

Finally, since $\Phi_{M,\theta}^{-M} = -\phi_0 I$, and $\Phi_{M,\varsigma}^{-M} = -1/\varrho_0 I$ we obtain

$$\begin{aligned} &(1 - \lambda^M \phi_0 \varrho_0^{-*}) P_{M,N,L} \\ &= M V_1^{-1} \Bigg(\Lambda_{\phi,\bar{w}^f_{M,N-1,L}} F_M \bar{D}_{\phi\varrho^*} F_M^* \Lambda^*_{\varrho,\bar{w}^f_{M,N-1,L}} \\ &\quad - \lambda \Lambda_{\phi, Z_\theta \bar{w}^b_{M,N,L}} F_M \bar{D}_{\phi\varrho^*} F_M^* \Lambda^*_{\varrho, Z_\varsigma \bar{w}^b_{M,N,L}} \\ &\quad + \Lambda_{\phi, \Phi_{M,\theta}^{-1} \tilde{k}^d_{M,N-1,L}} F_M \bar{D}_{\phi\varrho^*} F_M^* \Lambda^*_{\varrho, \Phi_{M,\varsigma}^{-1} \tilde{k}^d_{M,N-1,L}} \\ &\quad + \lambda \Lambda_{\phi, \Phi_{M,\theta}^{-1} \tilde{k}_{M,N,L}} F_M \bar{D}_{\phi\varrho^*} F_M^* \Lambda^*_{\varrho, \Phi_{M,\varsigma}^{-1} \tilde{k}_{M,N,L}} \\ &\quad - \lambda^N \Lambda_{\phi, Z_\theta \breve{\bar{k}}^{d_o}_{M,N}} F_M \bar{D}_{\phi\varrho^*} F_M^* \Lambda^*_{\varrho, Z_\varsigma \breve{\bar{k}}^{d_o}_{M,N}} \Bigg) V_2^{-*}. \end{aligned} \quad \text{(A-4)}$$

This gives (85), where $\Lambda_{.,.}$ are defined in (86) and (87).

ACKNOWLEDGMENT

The author thanks FAPERJ for funding.

REFERENCES

[1] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Hoboken, NJ, USA: Wiley, 2003.

[2] R. Merched, "A unified approach to structured covariances: Polynomial vandermonde bezoutian representations," in *Proc. EUSIPCO*, Bucharest, Romania, Aug. 2012, pp. 1860–1864, ISSN 2076-1465.

[3] R. Merched, "Exact trigonometric superfast inverse covariance representations," in *Proc. ICNC*, San Diego, 2013, pp. 490–495, EUA.

[4] R. Merched, "A unified approach to reduced-redundancy transceivers: Superfast linear and block-iterative generalized decision feedback equalizers," *IEEE Trans. Signal Process.*, vol. 61, no. 17, pp. 4214–4229, 2013.

[5] B. Friedlander, "Lattice filters for spectral estimation," *Proc. IEEE*, vol. 70, pp. 990–1017, 1982.

[6] D. Alpay, P. DeWilde, and H. Dym, "On the existence and construction of solutions to the partial lossless inverse scattering problem with applications to estimation theory," *IEEE Trans. Inf. Theory*, vol. 35, no. 6, pp. 1184–1205, Nov. 1989.

[7] P. A. Regalia and F. Desbouvries, "Displacement structures of covariance matrices, lossless systems, and numerical algorithm design," *SIAM J. Matrix Anal. Appl.*, vol. 16, no. 2, pp. 536–564, Apr. 1995.

[8] L. Ljung, M. Morf, and D. Falconer, "Fast calculation of gain matrices for recursive estimation schemes," *Int. J. Control*, vol. 27, pp. 1–19, Jan. 1978.

[9] D. W. Lin, "On the digital implementation of the fast Kalman algorithms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, pp. 998–1005, Oct. 1984.

[10] R. Merched and A. H. Sayed, "Extended fast fixed order RLS adaptive filtering," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3015–3031, Dec. 2001.

[11] R. Merched, "Extended RLS-lattice adaptive filters," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2294–2309, Sept. 2003.

[12] R. Merched, "On the extended RLS lattice adaptive filter variants: Error-feedback, normalized, and array-based algorithms," *EURASIP J. Appl. Signal Process.*, pp. 1235–1250, May 2005.

[13] T. Kailath, A. Vieira, and M. Morf, "Inverses of Toeplitz operators, innovations and orthogonal polynomials," *SIAM Rev.*, vol. 20, pp. 106–119, 1978.

[14] T. Kailath, S.-Y. Kung, and M. Morf, "Displacement ranks of matrices and linear equations," *J. Math. Anal. Appl.*, vol. 68, pp. 395–407, 1979.

[15] L. Euler, Introduction in Analysin Infinitorum, Tomus Primis, Lausanne, 1748.

[16] É. Bézout, "Recherches sur le degré deséquations résultantes de l'évanouissement des inconnues et sur les moyens qu'on doit employer pour trouver ces équations," (in French) *Histoire de l'Académie royale des sciences, partie Mémoires*, pp. 288–338, 1764.

[17] I. Sylvester, "On a theory of syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest Algebraical common measure," *Philos. Trans. Roy. Soc. London*, vol. 143, pp. 407–548, 1853.

[18] A. Cayley, "Note sur la méthode d'élimination de Bezout," (in French) *J. für die Reine und Angewandte Mathematik*, vol. 53, pp. 366–376, 1857, Berlin.

[19] C. Hermit, "Extrait d'une lettre de Mr. Ch. Hermite de Paris á Mr. Borchardt de Berlin, sur le nombre des racines d'une équation algibrique comprises entre des limits donnés," (in French) *J. für die Reine und Angewandte Mathematik*, vol. 52, pp. 39–51, 1856, Berlin.

[20] H. Lev-Ari, Y. Bistritz, and T. Kailath, "Generalized Bezoutians and families of efficient zero-location procedures," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 170–186, Feb. 1991.

[21] T. Kailath and V. Olshevsky, "Displacement structure approach to discrete trigonometric transform based preconditioners of G. Strang and T. Chan types," *Calcolo*, no. 33, pp. 191–208, 1996.

[22] T. Kailath and V. Olshevsky, "Displacement structure approach to polynomial Vandermonde and related matrices," *Linear Algebra Appl.*, vol. 261, pp. 49–90, 1997.

[23] V. Olshevsky, "Eigenvector computation for almost unitary Hessenberg matrices and inversion of Szego-Vandermonde matrices via Discrete Transmission lines," *Linear Algebra Appl.*, vol. 285, pp. 37–67, 1998.

[24] Y. Eidelman, I. Gohberg, I. Koltracht, and V. Olshevsky, "A Bjorck-Pereyra-type algorithm for Szego-Vandermonde matrices based on properties of unitary Hessenberg matrices," *Linear Algebra Appl.*, vol. 420, pp. 634–647, 2007.

[25] T. Bella, V. Olshevsky, and P. Zhlobich, "Signal flow graph approach to inversion of (H,m)-quasiseparable-Vandermonde matrices and new filter structures," *Linear Algebra Appl.*, vol. 432, pp. 2032–2051, Apr. 2010.

[26] T. Bella, V. Olshevsky, and P. Zhlobich, "Classifications of recurrence relations via subclasses of (H,m)-quasiseparable matrices," *Numerical Linear Algebra in Signals, Systems and Control*, ser. Lecture Notes in Electrical Engineering, vol. 80, pp. 23–53, 2011.

[27] T. Bella, Y. Eidelman, I. Gohberg, and V. Olshevsky, "Characterization of $(H, 1)$ quasiseparable matrices and their subclasses via recurrence relations and signal flow graphs," [Online]. Available: http://www.math.uconn.edu/~olshevsky/papers/characterization.pdf preprint

[28] T. Bella, Y. Eidelmany, I. Gohbergy, V. Olshevsky, E. Tyrtyshnikovz, and P. Zhlobich, "A Traub-like algorithm for Hessenberg-quasiseparable-Vandermonde matrices of arbitrary order," *Operator Theory, Adv. Appl.*, vol. 199, pt. II, pp. 127–154, 2010.

[29] G. Heinig and K. Rost, *Algebraic Methods for Toeplitz-Like Matrices and Operators*. Berlin, Germany: Akademie-Verlag and Birkhäuser, 1984.

[30] K. Maouche and D. T. M. Slock, "Performance analysis and FTF version of the generalized sliding window recursive least-squares (GSWRLS) algorithm," in *Proc. 29th IEEE Annu. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct. 1995, vol. 1, pp. 685–689.

[31] S. L. Gay and S. Tavanthia, "The fast affine projection algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 1995, vol. 5, pp. 3023–3026.

[32] R. Merched and N. R. Yousef, "Fast computation of finite-length MIMO MMSE decision feedback equalizers," in *Proc. Asilomar Conf.*, Dec. 2004, pp. 1781–1784.

[33] J. Maroulas and S. Barnett, "Polynomials with respect to a general basis. I. Theory," *J. Math. Anal. Appl.*, vol. 72, pp. 177–194, 1979.

[34] G. Szegö, *Orthogonal Polynomials*, 4th ed. Providence, RI, USA: Amer. Math. Soc., 1975, pp. 42–44.

[35] R. Merched and I. S. G. Figueiredo, "Block precoder-based energy constrained DFE," in *Proc. Int. Symp. Circuits Syst.*, Kos, Greece, May 2006, pp. 2057–2060.

[36] R. Merched, "Fast computation of constrained decision feedback equalizers," *IEEE Trans. Signal Process.*, vol. 55, pp. 2446–2457, Jun. 2007.

[37] R. Merched, "Fast algorithms in slow and high Doppler mobile environments," *IEEE Trans. Signal Process.*, vol. 9, no. 9, pp. 2890–2901, Sept. 2010.

[38] R. Merched, "Application of superfast algorithms to pilot-based channel estimation schemes," in *Proc. SPAWC*, Recife, Brazil, Jul. 2008, pp. 141–145.

[39] P. A. Regalia, "Numerical stability issues in fast least-squares adaptation algorithms," *Opt. Eng.*, vol. 31, pp. 1144–1152, Jun. 1992.

[40] P. A. Regalia, "Numerical stability properties of a QR-based fast least-squares algorithm," *IEEE Trans. Signal Process.*, vol. 41, no. 6, pp. 2096–2109, Jun. 1993.

[41] D. T. M. Slock, "The backward consistency concept and round-off error propagation dynamics in recursive least-squares algorithms," *Opt. Eng.*, vol. 31, pp. 1153–1169, Jun. 1992.

[42] I. Gohberg and V. Olshevsky, "Circulants, displacements and decomposition of matrices," *Integral Equat. Oper. Theory*, vol. 15, pp. 853–863, May 1992.

[43] R. Merched, "Fast generalized sliding window RLS recursions for IIR recurrence related basis functions," in *Proc. DSP Conf.*, Corfu, Greece, Jul. 2011, pp. 1–6.

[44] G. Ammar and P. Gader, "New decompositions of the inverse of a Toeplitz matrix," in *Signal Process., Scattering, Operator Theory, Proc. Int. Symp. (MTNS-89)*, Boston, MA, USA, 1990, vol. III, pp. 421–428, Birkhauser.

[45] G. Heinig and K. Rost, "DFT representations of Toeplitz-plus-Hankel Bezoutians with application to fast matrix vector multiplication," *Linear Algebra Appl.*, vol. 284, pp. 157–175, Nov. 1998.

Ricardo Merched (S'94–M'01–SM'06) received the B.S. and the M.Sc. degrees from the Federal University of Rio de Janeiro (UFRJ), Brazil, and the Ph.D. degree from University of California, Los Angeles (UCLA) in 2001. He became Professor at the Department of Electrical and Computer Engineering at UFRJ in 2002. He was a visiting professor at University of California, Irvine, and at Unik, University Graduate Center in Oslo, during 2006 and 2007. Dr. Merched was an Associate Editor of the *IEEE Transactions on Circuits and Sistems I*, the *IEEE Transactions on Signal Processing Letters*, and the *EURASIP, European Journal on Advances in Signal Processing*. His current main interests include adaptive filtering algorithms, multirate systems, efficient digital signal processing techniques for MIMO equalizer architectures in wireless and wireline communications, RADAR imaging, and biomedical imaging applications, as well as low complexity solutions for these problems. He holds six US patents on efficient digital signal processing algorithms for channel estimation and equalization.