Geometrical aspects of linear prediction algorithms

François Desbouvries
Département Signal et Image
Institut National des Télécommunications
9 rue Charles Fourier, 91011 Evry, France
Francois.Desbouvries@int-evry.fr

Keywords: Schur and Levinson-Szegö algorithms, Schur complements, spherical trigonometry, analytic interpolation theory, Lobachevski geometry

Abstract

In this paper, an old identity of G. U. Yule among partial correlation coefficients is recognized as being equal to the cosine law of spherical trigonometry. Exploiting this connection enables us to derive some new (and potentially useful) relations among partial correlation coefficients. Moreover, this observation provides new (dual) non-Euclidean geometrical interpretations of the Schur and Levinson-Szegö algorithms.

1 Introduction

Linear prediction and interpolation is a major tool in time series analysis and in signal processing. In this context, the Schur and Levinson-Szegö algorithms compute the partial autocorrelation function of a wide-sense stationary process. As such, they have found a large variety of electrical engineering applications.

Let us briefly recall the history of these algorithms. At the beginning of the century, Schur, Carathéodory and Toeplitz were active in such fields as analytic function theory, Toeplitz forms and moment problems. In 1917, Schur developped a recursive algorithm for checking whether a given function $s(z) = \sum_{k=0}^{\infty} s_k z^k$ is analytic and bounded by one in the unit disk [1]. Such functions are characterized by a sequence of parameters of modulus less than one (the Schur parameters) which are computed recursively from the power series coefficients s_k by an elegant algorithm. On the other hand, Carathéodory and Toeplitz showed that $c(z) = c_0 + 2\sum_{k=1}^{\infty} c_k z^k$ is analytic and has positive real part for $|z| < 1$ if and only if the Toeplitz forms $\sum_{i,j=0}^{n} a_i b_j^* c_{j-i}$, with $c_{-n} = c_n^*$, are positive for all n. Let

$$s(z) = \frac{c(z) - c_0}{c(z) + c_0} \iff c(z) = c_0 \frac{1 + s(z)}{1 - s(z)} \quad ; \qquad (1)$$

since $|s(z)| \leq 1$ if and only if $c(z)$ has positive real part, the Schur algorithm implicitely enables to test whether a Toeplitz form is positive.

On the other hand, Toeplitz forms were studied independently by Szegö, who introduced a set of orthogonal polynomials with respect to an (absolutely continuous) positive measure on the unit circle. These polynomials obey a two-terms recursion [2] involving a set of parameters of modulus bounded by one, which later on were recognized to be equal to the Schur parameters [3]. In the 1940s, Toeplitz forms received a revived interest in view of their natural occurrence in the Kolmogorov-Wiener prediction and interpolation theory of stationary processes (see e.g., [4, ch. 10], as well as the survey paper [5] and the references therein). Working on Wiener's solution of the continuous time prediction problem, Levinson proposed a fast algorithm for solving Toeplitz systems; later on, the Levinson recursions were recognized as being the recurrence relations of Szegö.

Finally, there was an intense activity in these fields beginning in the late 70s, mainly towards the development of fast algorithms for numerical linear algebra, on the one hand, and in the domain of analytic interpolation theory, on the other hand. Through these new developments and extensions, new connections with other mathematical topics and disciplines were developed, including among others displacement rank theory, $J-$lossless transfer functions, modern analytic function theory and operator theory. The literature on these connections and extensions is vast; the reader may refer for instance to the papers [6] [7] [8] [9] and books [10] [11].

The mathematical environment of these algorithms is thus very rich, and these various interactions have already been thoroughly investigated by many researchers. In this wealthy context, our contribution in this paper consists in exhibiting new unnoticed connections with spherical trigonometry.

As far as geometry is concerned, the Lobachevski geometry was already known to be a natural environment of the Schur and Levinson-Szegö algorithms, since the core of these algorithms mainly consists in a linear fractional transformation leaving the unit circle invariant. However, a new point of view is obtained when considering the algorithms (via positive definite Toeplitz forms) in the particular context of their application to linear prediction. Then, up to an appropriate normalization, the Schur and Levinson-Szegö algorithms become trigonometric identities in a spherical triangle. Since the real projective 2-space $\mathbb{P}^2$ is the quotient space obtained from the sphere by identifying antipodal points, we see that the alternate non-Euclidean geometry with constant curvature (i.e., the elliptic one) is indeed another natural geometrical environment of the Schur and Levinson-Szegö algorithms as well.

Let us briefly outline the underlying mechanisms leading to this new interpretation. Let $\{X_i\}$ be zero-mean square-integrable random variables, $\hat{X}_j^{1:n}$ the best linear mean-square estimate of X_j in terms of $\{X_i\}_{i=1}^n$, and $\tilde{X}_j^{1:n} = X_j - \hat{X}_j^{1:n}$ the corresponding estimation error. The partial correlation coefficient (or parcor) of X_0 and X_{n+1}, given $\{X_i\}_{i=1}^n$, is defined as $\rho_{0,n+1}^{1:n} = [E(\tilde{X}_0^{1:n})^2]^{-1/2} E(\tilde{X}_0^{1:n}\tilde{X}_{n+1}^{1:n})[E(\tilde{X}_{n+1}^{1:n})^2]^{-1/2}$. It is bounded by 1 in magnitude and is classicaly interpreted as the correlation coefficient of X_0 and X_{n+1}, once the influence of $\{X_i\}_{i=1}^n$ has been removed. In 1907, G. U. Yule [12] showed that the parcors could be computed recursively :

$$\rho_{0,n+1}^{1:n} = \frac{\rho_{0,n+1}^{1:n-1} - \rho_{0,n}^{1:n-1}\,\rho_{n,n+1}^{1:n-1}}{\sqrt{1-(\rho_{0,n}^{1:n-1})^2}\,\sqrt{1-(\rho_{n,n+1}^{1:n-1})^2}} \ . \qquad (2)$$

It happens that this well known formula is formally equal to the fundamental cosine law of spherical trigonometry :

$$\cos A = \frac{\cos a - \cos b \cos c}{\sin b \, \sin c} \,, \qquad (3)$$

which gives an angle of a spherical triangle, in terms of its three sides (see figure 1). This observation establishes an unexpected link between statistics and time-series analysis, on the one hand, and spherical trigonometry (a branch of trigonometry), on the other hand.

In former papers [13] [14], spherical trigonometry was shown also to admit a close connection with the electrical engineering topic of recursive least-squares adaptive filtering which, as linear regression analysis, is a mean square approximation problem. Now, the Schur and Levinson-Szegö algorithms can be written as algebraic recursions within a covariance matrix or its inverse; due to the identification (2) = (3), they admit a connection with spherical trigonometry as well.

Indeed, the source of such analogies is that (time- or order-) recursive least-squares algorithms can be developed from projection identities. In linear regression, one recognizes that the mean-square error to be minimized is a distance, so the projection theorem can be applied in the Hilbert space generated by the random variables. Introducing a new variable in the regression problem amounts to updating a projection operator, and the problem can indeed be described in terms of projections in a space generated by three vectors. But three unit-length vectors form a tetrahedron in 3D-space, and deriving projective identities in a normalized tetrahedron results in deriving trigonometric relations in the spherical triangle determined by this tetrahedron (see figure 1, and [14] for details).

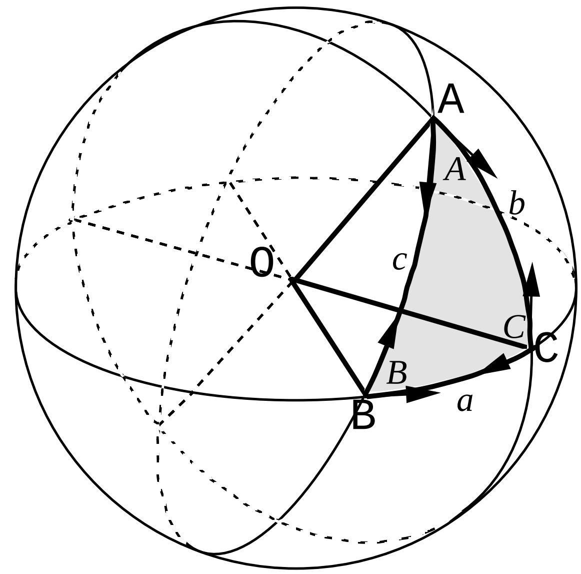

Figure 1: The spherical triangle ABC.

Let us now turn to the organization of this paper. Non-Euclidean hyperbolic aspects of the Schur algorithm are implicit in [15] but do not seem otherwise to be well known. Yet the Lobachevski geometry is, by construction, an essential feature of the algorithm, which deserves to be better appreciated. More precisely, we show in section 2 that Schur's layer-peeling type solution to the Carathéodory problem necessarily makes use of automorphisms of the unit disk which, on the other hand, happen to be the direct isometries of the Lobachevski plane.

The last two sections are devoted to the new geometrical interpretations in terms of spherical trigonometry. So in section 3, we relate recursive regressions within a set of $(n+1)$ random variables, algebraic manipulations in a covariance matrix or in its inverse, and spherical trigonometry. We show that adding (resp. removing) a new variable in the regression problem which, in terms of Schur complements on the covariance matrix (resp. on its inverse), amounts to using the quotient property [16, p. 279], corresponds in terms of spherical trigonometry to applying the law of cosines (resp. the polar law of cosines).

Lastly in section 4, we further assume that the random variables are taken out of a stationary time series, and we use the results of section 3 to interpret in parallel the Schur and Levinson-Szegö algorithms in terms of spherical trigonometry. The Schur (resp. Levinson-Szegö) relations consist in two Schur complement recursions (in the forward and backward sense) in the original covariance matrix (resp. in its inverse), and can indeed be interpreted in dual spherical triangles. Up to an appropriate normalization, the Schur (resp.

order decreasing Levinson-Szegö) recursions coincide with two coupled occurrences of the law of cosines (resp. of the polar law of cosines), and the Levinson-Szegö recursions with two coupled occurrences of the polar five elements formula.

2 Non-Euclidean *(hyperbolic)* geometrical aspects of the Schur and Levinson-Szegö algorithms

In this section, we first briefly recall the mechanisms underlying the Schur algorithm. We next show that the choice by Schur of a recursive solution to the Carathéodory problem naturally sets the algorithm in a non-Euclidean hyperbolic environment.

2.1 The Schur algorithm

The Carathéodory analytic interpolation problem consists in finding all functions s such that 1. $s(z) = \sum_{k=0}^{n} a_k z^k + \mathcal{O}(z^{n+1})$, and 2. $s \in \mathcal{S} \stackrel{def}{=} \{f(z)$ analytic in $|z| < 1,$ and $|f(z)| \leq 1$ for $|z| < 1\}$. In 1917, Schur proposed a "layer-peeling" type algorithm [1] (i.e., in which the interpolation data are processed recursively) which we briefly recall.

Let us first consider the case where there is only one interpolation point a_0. Due to the maximum principle, the problem has no solution if $|a_0| > 1$, and admits the unique solution $s(z) = a_0$ if $|a_0| = 1$. If $|a_0| < 1$, let

$$s_1(z) = \frac{1}{z}\frac{s(z) - a_0}{1 - a_0^* s(z)} \quad \Leftrightarrow s(z) = \frac{z s_1(z) + a_0}{1 + a_0^* z s_1(z)} \quad . \tag{4}$$

Then the key property of this transformation is that $s \in \mathcal{S} \quad \Leftrightarrow s_1 \in \mathcal{S}$, so that

$$(s \in \mathcal{S} \text{ and } s(0) = a_0) \Leftrightarrow (s(z) = \frac{z s_1(z) + a_0}{1 + a_0^* z s_1(z)} \text{ and } s_1 \in \mathcal{S}). \tag{5}$$

In the case of a single interpolation point a_0, equation (5) provides a parametrization of all solutions to the Carathéodory problem in terms of an arbitrary Schur function s_1. A new interpolation point a_1 can be accomodated by further restricting this set of possible functions s_1. From equation (5), we see that ($s \in \mathcal{S}$ and has interpolation constraints $(a_0 \cdots a_n)$) if and only if ($s_1 \in \mathcal{S}$ and has interpolation constraints $(a_0^1 \cdots a_{n-1}^1)$), in which the a_i^1 depend on the data a_i. In particular, $\left.\frac{s'(z)}{1!}\right|_{z=0} = a_1$ if and only if $s_1(0) = a_0^1$: we are thus led back to the same metric constrained interpolation problem, but now of order $n-1$. These considerations lead by induction to the Schur algorithm [1] :

$$s_0(z) = s(z);\ \text{if } |s_p(0)| < 1, s_{p+1}(z) = \frac{1}{z}\frac{s_p(z) - s_p(0)}{1 - s_p^*(0)s_p(z)}. \tag{6}$$

[1] In §2.3 we will deal with geometrical aspects of the recursion (6). This is why the regular case only is considered here; all details of the general case can be found in [1].

2.2 The Schur mechanism and automorphisms of the unit circle

Let us now analyse the design of the Schur algorithm in terms of automorphisms of the unit disk. Automorphisms of a domain G are bi-holomorphic mappings of G : $Aut(G) \stackrel{def}{=} \{f \in \mathcal{H}(G,G)$, s.t. f^{-1} exists, and $f^{-1} \in \mathcal{H}(G,G)\}$, where $\mathcal{H}(G_1, G_2)$ is the set of holomorphic functions of G_1 onto G_2. The Schur class $\mathcal{S}$ coincides with $\mathcal{H}(\mathbb{D}, \overline{\mathbb{D}})$ where $\mathbb{D}$ is the open unit disk.

The mapping (4) can be decomposed into two steps :

$$s_1'(z) = \frac{s(z) - a_0}{1 - a_0^* s(z)} \quad \Longleftrightarrow s(z) = \frac{s_1'(z) + a_0}{1 + a_0^* s_1'(z)} \quad , \tag{7}$$

and $s_1(z) = \frac{1}{z} s_1'(z)$. Since the transform $s \mapsto s_1'$ in (7) is a linear fractional transformation (LFT), we begin with recalling elementary (Euclidean) properties of LFTs [17] [18]. Let $M\ddot{o}b$ denote the Möbius group $M\ddot{o}b \stackrel{def}{=} \{z \mapsto \frac{az+b}{cz+d}$, with $ad - bc \neq 0\}$ ($ad - bc = 0$ corresponds to the trivial situation where the mapping is either constant or undefined). Then the mapping

$$\begin{array}{llll} \phi: & GL(2,\mathbb{C}) & \to & M\ddot{o}b \\ & M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} & \mapsto & \phi_M, \text{ with } \phi_M(z) = \frac{az+b}{cz+d} \end{array}$$

is a group homomorphism. Since $\phi_M = \phi_{\lambda M}$ for all $\lambda \in \mathbb{C}^\star$, there is no loss of generality in supposing that $det(M) = 1$. From now on we shall thus restrict to $SL(2,\mathbb{C})$. Then the kernel of ϕ reduces to I and $-I$, and $M\ddot{o}b$ is isomorphic to the associated quotient group : $M\ddot{o}b \cong (SL(2,\mathbb{C})/\pm I)$.

Now, $\frac{az+b}{cz+d} = \frac{a}{c} + \frac{bc-ad}{c^2}\frac{1}{z+d/c}$ (resp. $= \frac{a}{d}z + \frac{b}{d}$) if $c \neq 0$ (resp. $c = 0$), so any LFT is a succession of translations, inversions, rotations, and/or homotheties. Since all these geometrical transformations preserve circles, a LFT maps any circle in the complex plane (possibly of infinite radius) into another circle (possibly of infinite radius). In particular, ϕ_M maps the unit circle $\mathbb{T}$ onto itself if and only if M belongs to the subgroup $SU(1,1)$ of $SL(2,\mathbb{C})$ consisting of $\Sigma-$unitary matrices, with $\Sigma = diag(+1,-1)$:

$$SU(1,1) = \{\begin{bmatrix} a & b \\ b^* & a^* \end{bmatrix} \text{ with } |a|^2 - |b|^2 = 1\}$$

$$= \{M \in SL(2,\mathbb{C}),\ \text{s.t. } M\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} M^H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\}.$$

Let $\mathcal{N}$ be the set of such Möbius transforms; $\mathcal{N}$ is a subgroup of $M\ddot{o}b$, and

$$\begin{array}{lll} \mathcal{N} & \stackrel{def}{=} & \{z \mapsto \frac{az+b}{b^*z + a^*},\ \text{s.t. } |a|^2 - |b|^2 = 1\} \\ & \cong & (SU(1,1)/\pm I). \end{array}$$

Finally, observe that any $\phi_M \in \mathcal{N}$ also maps the interior $\mathbb{D}$ of the unit-circle onto itself (and similarly for its exterior).

We now turn back to the Schur algorithm. Since $|a_0| < 1$, we have

$$\begin{aligned} & (s(z) \in \mathcal{H}(\mathbb{D}, \overline{\mathbb{D}}) \text{ and } s(0) = a_0) && (8)\\ \Longleftrightarrow\ & (s(z) \in \mathcal{H}(\mathbb{D}, \mathbb{D}) \text{ and } s(0) = a_0) && (9)\\ \Longleftrightarrow\ & (s_1'(z) \in \mathcal{H}(\mathbb{D}, \mathbb{D}) \text{ and } s_1'(0) = 0) && (10)\\ \Longleftrightarrow\ & s_1(z) \in \mathcal{H}(\mathbb{D}, \overline{\mathbb{D}}) \ , && (11)\end{aligned}$$

whence (5). Equivalence (8)/(9) is due to the maximum principle and (10)/(11) to the Schwarz lemma. As for (9)/(10), it holds because $s_1' = \phi_M \circ s$, where the mapping $\phi_M : z \mapsto \phi_M(z) = \frac{z-a_0}{1-a_0^* z} = \frac{(z-a_0)/\sqrt{1-|a_0|^2}}{(1-a_0^* z)/\sqrt{1-|a_0|^2}}$ belongs to $\mathcal{N} \subset Aut(\mathbb{D})$. In fact, it is interesting to notice that (7) was the only possible choice, because, as is well known [19] [20], the automorphisms of $\mathbb{D}$ indeed coincide with the group of LFTs which leave the unit-disk invariant :

$$\mathcal{N} = Aut(\mathbb{D}) \ .$$

2.3 Hyperbolic geometry of the Schur and Levinson-Szegö algorithms

We now turn from analytical to geometrical considerations. These are obtained naturally in the framework of the theory developped in [21], which aims at describing the holomorphic structure of a domain G of $\mathbb{C}^n$ in terms of geometric properties of the space (G, d_G). If the distance d_G is chosen such that any holomorphic mapping of G onto itself is a contraction, then automorphisms of G are isometries with respect to d_G and can thus be interpreted (in the spirit of F. Klein) as rigid motions with respect to the geometry specified by d_G.

The Schwarz-Pick lemma [17] [20] [21] [22] provides a nice illustration of this general methodology to the present situation; as expected, we shall meet non-Euclidean hyperbolic geometry, since G reduces to the unit disk $\mathbb{D}$, which is a Euclidean model of the Lobachevski plane (see e.g. [23]). Let $d_M(.,.)$ denote the Möbius distance in $\mathbb{C}$: for all $z, z' \in \mathbb{C}$, $d_M(z, z') = \left|\frac{z-z'}{1-z'^* z}\right|$. The Schwarz-Pick lemma states that any function f belonging to $\mathcal{H}(\mathbb{D}, \mathbb{D})$ is a contraction with respect to the Möbius distance :

$$\begin{aligned} & f \in \mathcal{H}(\mathbb{D}, \mathbb{D}) \Rightarrow \\ & \text{for all } z, z' \in \mathbb{D}, \quad \left|\frac{f(z)-f(z')}{1-(f(z'))^* f(z)}\right| \leq \left|\frac{z-z'}{1-z'^* z}\right| ; \end{aligned}$$

equality holds if and only if f belongs to $Aut(\mathbb{D})$.

Let us now turn back to the discussion at the end of section 2.2. Schur had to chose the functional $s \mapsto s_1'$ (or, in general, $s_p \mapsto z s_{p+1}$) within $Aut(\mathbb{D})$, and automorphisms of $\mathbb{D}$ preserve the Möbius distance d_M, and thus the Poincaré distance $d_P(.,.)$, with $d_P(z, z') = \log \frac{1+d_M(z,z')}{1-d_M(z,z')}$. More precisely, they are well known to coincide with the direct isometries of the Lobachevski plane $(\mathcal{H}^2, d_P)$ [17] [18] [2] :

$$\mathcal{N} = Aut(\mathbb{D}) \cong \{ \text{ direct isometries of } (\mathcal{H}^2, d_P) \} \ .$$

[2] The full isometry group of $(\mathcal{H}^2, d_P)$ is obtained by including the map $z \mapsto z^*$ as a generator.

This geometry is thus, by construction, the natural geometrical environment of the Schur algorithm.

Finally, let us briefly consider the Levinson-Szegö algorithm. It is not a solution to an analytic interpolation problem, but can nevertheless be rephrased (via the Schur-Cohn stability test) in the framework of section 2.1 [24], and thus shares the same geometrical environment. For let $a(z) = \sum_{i=0}^n a_i^n z^i$, $b(z) = z^n(a(1/z^*))^*$ and $f(z) = b(z)/a(z)$. Then $f(z)$ is rational and has modulus 1 on $\mathbb{T}$. So, by the maximum modulus theorem, ($f(z)$ is analytic in $\mathbb{D}$ and $|f(z)| = 1$ on $\mathbb{T}$) if and only if ($|f(z)| \leq 1$ in $\mathbb{D}$ and $|f(z)| = 1$ on $\mathbb{T}$), i.e., if and only if f is a rational bounded function of the lossless type (a Blashke product). But this can be checked via the Schur algorithm, because $f = f_0$ is a Blashke product of order n if and only if $|f_p(0)| < 1$ for $0 \leq p \leq n-1$ and $|f_n(0)| = 1$; the recursions coincide with the order-decreasing Levinson-Szegö algorithm.

3 Schur Complements in $R_{0:n}/R_{0:n}^{-1}$ and spherical trigonometry

From now on we shall deal with spherical trigonometry aspects of the Schur and Levinson-Szegö algorithms. In this intermediate section, we first recall some elementary projective identities. We next bring back from spherical trigonometry some relations among parcors which we will refer to in sections 3.3 and 4. In the view of section 4, we eventually consider Schur complementation in a covariance matrix or in its inverse, because Schur complements provide the connection between the Schur and Levinson-Szegö algorithms and spherical trigonometry. The reason why is that algebraically the elementary recursion with pivot $t_{k,k}$: $t_{i,j} \rightarrow t'_{i,j} = t_{i,j} - t_{i,k} t_{k,k}^{-1} t_{k,j}$, reduces to a cosine law when normalized by $\sqrt{t_{i,i}}\ \sqrt{t_{j,j}}$.

3.1 Partial correlation coefficients, recursive projections and Yule's parcor identitity

Our geometrical results are based on the properties of orthogonal projectors and can thus be formalized in any Hilbert space. However, the natural framework in this paper is the space $\mathbb{L}^2(\Omega, \mathcal{A}, P)$ of complex, zero-mean, square-integrable random variables defined on $(\Omega, \mathcal{A}, P)$, endowed with the inner product $(X, Y) = E(XY^*)$.

Let $P_{\mathcal{M}}$ denote the orthogonal projector on the Hilbert space $\mathcal{H}(\mathcal{M})$ generated by $\mathcal{M}$, $P_{\mathcal{M}}^{\perp} = I - P_{\mathcal{M}}$, $\hat{A}^{\mathcal{M}}$ the projection of A onto $\mathcal{M}$, and $\tilde{A}^{\mathcal{M}} = A - \hat{A}^{\mathcal{M}}$. $\overline{X}$ denotes normalization to unit norm : $\overline{X} = X(X,X)^{-1/2}$. The (sometimes called total) correlation coefficient $\rho_{A,B}$ (resp. partial correlation coefficient $\rho_{A,B}^{\mathcal{M}}$) of A and B (resp. of A and B, with respect to a commun subspace $\mathcal{M}$) is defined as $\rho_{A,B} = (\overline{A}, \overline{B}) = (\rho_{B,A})^*$ (resp. $\rho_{A,B}^{\mathcal{M}} = (\overline{\tilde{A}^{\mathcal{M}}}, \overline{\tilde{B}^{\mathcal{M}}}) = (\rho_{B,A}^{\mathcal{M}})^*$).

Let us now consider recursive projections. It is well known that

$$P_{\mathcal{M},A} = P_{\mathcal{M}} + P_{\tilde{A}^{\mathcal{M}}} \ , \ P_{\mathcal{M},A}^{\perp} = P_{\mathcal{M}}^{\perp} - P_{\tilde{A}^{\mathcal{M}}} \qquad (12)$$

where $P_{\mathcal{M},A}$, say, is the orthogonal projector onto the closed subspace generated by $\mathcal{M}$ and A. These identities are of utmost importance in RLS adaptive filtering as well as in Kalman filtering. From (12), it is easy to show that

$$\tilde{B}^{\mathcal{M},A} = \tilde{B}^{\mathcal{M}} - (\tilde{B}^{\mathcal{M}}, \tilde{A}^{\mathcal{M}})(\tilde{A}^{\mathcal{M}}, \tilde{A}^{\mathcal{M}})^{-1}\tilde{A}^{\mathcal{M}} \quad , \tag{13}$$

which gives the useful relations

$$(\tilde{B}^{\mathcal{M},A}, \tilde{B}^{\mathcal{M},A})/(\tilde{B}^{\mathcal{M}}, \tilde{B}^{\mathcal{M}}) = 1 - |\rho^{\mathcal{M}}_{A,B}|^2 \tag{14}$$

and

$$(\overline{\tilde{A}^{\mathcal{M},B}}, \overline{\tilde{B}^{\mathcal{M},A}}) = -\rho^{\mathcal{M}}_{A,B} \ . \tag{15}$$

In order to get spherical trigonometry relations, we need consider three projection residuals. From (13) we have

$$\begin{aligned} (\tilde{C}^{\mathcal{M},A}, \tilde{B}^{\mathcal{M},A}) &= (\tilde{C}^{\mathcal{M}}, \tilde{B}^{\mathcal{M}}) \\ &- (\tilde{C}^{\mathcal{M}}, \tilde{A}^{\mathcal{M}})(\tilde{A}^{\mathcal{M}}, \tilde{A}^{\mathcal{M}})^{-1}(\tilde{A}^{\mathcal{M}}, \tilde{B}^{\mathcal{M}}) \end{aligned} \tag{16}$$

Dividing by $(\tilde{C}^{\mathcal{M}}, \tilde{C}^{\mathcal{M}})^{1/2}(\tilde{B}^{\mathcal{M}}, \tilde{B}^{\mathcal{M}})^{1/2}$, and using (14), we get

$$\rho^{\mathcal{M},A}_{C,B} = \frac{\rho^{\mathcal{M}}_{C,B} - \rho^{\mathcal{M}}_{C,A}\rho^{\mathcal{M}}_{A,B}}{\sqrt{1-|\rho^{\mathcal{M}}_{C,A}|^2}\sqrt{1-|\rho^{\mathcal{M}}_{A,B}|^2}} \ , \tag{17}$$

which is formally equal to (3) (at least in the real case), up to a straightforward identification of variables.

3.2 New relations among parcors induced by spherical trigonometry

Let us first briefly recall some facts from spherical trigonometry (and in particular the duality principle). Three points $\mathtt{A}, \mathtt{B}$ and $\mathtt{C}$ on the sphere $(0, 1)$ determine the spherical triangle $\mathtt{ABC}$, which consists of the 3 arcs of great circles $\mathtt{AB}$, $\mathtt{AC}$ and $\mathtt{BC}$ obtained by intersecting the sphere and the planes $\mathtt{OAB}$, $\mathtt{OAC}$ and $\mathtt{OBC}$ (see figure 1). A spherical triangle has 6 elements : the 3 sides a, b and c, and the 3 angles A, B and C. The side a, say, is defined as the angle $\widehat{\mathtt{BOC}}$ and is equal to the length of the arc $\mathtt{BC}$. The angle A, say, is defined as the dihedral angle between the planes $\mathtt{OAB}$ and $\mathtt{OAC}$, and is also equal to the angle made by the tangents to the spherical triangle $\mathtt{ABC}$ at point $\mathtt{A}$.

We now turn to the duality principle of spherical trigonometry. Let $\mathtt{A}'$ be the pole (with respect to the equator passing through $\mathtt{B}$ and $\mathtt{C}$) which is in the same hemisphere as $\mathtt{A}$; $\mathtt{B}'$ and $\mathtt{C}'$ are defined similarly. The spherical triangle $\mathtt{A'B'C'}$ is the polar triangle of $\mathtt{ABC}$. In $\mathtt{A'B'C'}$, the elements a' and A', say, are equal respectively to $\pi - A$ and $\pi - a$ (see e.g., [25] [26]). So, for any spherical trigonometry formula there exists a dual relation, obtained by replacing (a, b, c, A, B, C) by $(\pi - A, \pi - B, \pi - C, \pi - a, \pi - b, \pi - c)$, respectively.

There are three degrees of freedom in a spherical triangle, so there cannot be more than three distinct relations among the six elements. All the spherical trigonometry relations can thus be derived from the three cosine laws obtained by permuting variables into (3). Now, the identification (3) = (17) enables us to hint that some spherical trigonometry relations might hold as well when transposed to the parcors framework, and indeed they do. Similarly, they can all be derived from (17); however, though purely algebraic, these relations are not necessarily intuitive.

Among any 4 elements there exists one and only one relation. These 15 relations are the 3 cosine laws, the 3 cosine laws in the polar triangle, the 3 self-dual sine formulas and the 6 self-dual cotangent formulas. They all have a parcor equivalent. However, there are many different relations among any 5 elements (or between the 6), and it always seems possible to find new ones. Thus we give only one of them, the five elements formula. For sake of brevity, proofs are omitted.

3.2.1 The cosine law in the polar triangle

In the polar triangle, the cosine law reads :

$$\cos a = \frac{\cos A \ + \ \cos B \cos C}{\sin B \ \sin C} \ . \tag{18}$$

Similarly, (17) admits the polar version :

$$\rho^{\mathcal{M}}_{C,B} = \frac{\rho^{\mathcal{M},A}_{C,B} + \rho^{\mathcal{M},B}_{C,A}\ \rho^{\mathcal{M},C}_{A,B}}{\sqrt{1-|\rho^{\mathcal{M},B}_{C,A}|^2}\sqrt{1-|\rho^{\mathcal{M},C}_{A,B}|^2}} \ , \tag{19}$$

which was already known to Yule [12, (19) p. 93].

3.2.2 The sine law

The spherical triangle self-dual sine law is the following formula :

$$\frac{\sin A}{\sin a} = \frac{\sin B}{\sin b} = \frac{\sin C}{\sin c} \ . \tag{20}$$

Similarly, the following relation holds among parcors :

$$\sqrt{\frac{1-|\rho^{\mathcal{M},A}_{B,C}|^2}{1-|\rho^{\mathcal{M}}_{B,C}|^2}} = \sqrt{\frac{1-|\rho^{\mathcal{M},B}_{C,A}|^2}{1-|\rho^{\mathcal{M}}_{C,A}|^2}} = \sqrt{\frac{1-|\rho^{\mathcal{M},C}_{A,B}|^2}{1-|\rho^{\mathcal{M}}_{A,B}|^2}} \ . \tag{21}$$

3.2.3 The cotangent formulas

These are the 6 self-dual formulas obtained by permuting variables into the equation

$$\cot b \sin a = \cos C \cos a + \sin C \cot B \ . \tag{22}$$

Similarly, the following relation among parcors holds :

$$\begin{aligned} \left(\rho^{\mathcal{M}}_{A,C}/\sqrt{1-|\rho^{\mathcal{M}}_{A,C}|^2}\right)\sqrt{1-|\rho^{\mathcal{M}}_{B,C}|^2} \ = \ \rho^{\mathcal{M},C}_{A,B}\,\rho^{\mathcal{M}}_{B,C} \\ +\sqrt{1-|\rho^{\mathcal{M},C}_{A,B}|^2}\left(\rho^{\mathcal{M},B}_{A,C}/\sqrt{1-|\rho^{\mathcal{M},B}_{A,C}|^2}\right) \end{aligned} \tag{23}$$

3.2.4 The five elements formula

These are the 6 formulas obtained by permuting variables into

$$\cos b \sin c = \sin b \cos A \cos c + \sin a \cos B \ . \tag{24}$$

The dual equations are

$$\cos B \sin C = -\sin B \cos a \cos C + \sin A \cos b \ . \tag{25}$$

Similarly, the following relations among parcors hold :

$$\begin{aligned} \rho^{\mathcal{M}}_{C,A}\sqrt{1-|\rho^{\mathcal{M}}_{B,A}|^2} &= \\ \sqrt{1-|\rho^{\mathcal{M}}_{C,A}|^2}\rho^{\mathcal{M},A}_{C,B}\rho^{\mathcal{M}}_{B,A} &+ \sqrt{1-|\rho^{\mathcal{M}}_{C,B}|^2}\rho^{\mathcal{M},B}_{C,A} \end{aligned} \tag{26}$$

$$\begin{aligned} \rho^{\mathcal{M},B}_{C,A}\sqrt{1-|\rho^{\mathcal{M},C}_{B,A}|^2} &= \\ -\sqrt{1-|\rho^{\mathcal{M},B}_{C,A}|^2}\rho^{\mathcal{M}}_{C,B}\rho^{\mathcal{M},C}_{B,A} &+ \sqrt{1-|\rho^{\mathcal{M},A}_{C,B}|^2}\rho^{\mathcal{M}}_{C,A} \end{aligned} \tag{27}$$

3.3 Schur Complements in $R_{0:n}/R^{-1}_{0:n}$ and spherical trigonometry

We now consider covariance matrices and their inverses. Let $\{X_i\}_{i=0}^n$ be scalar random variables. For $p \leq q$, let $X_{p:q} = [X_p \cdots X_q]^T$. In all this section, we will assume that : $\{X_i\}_{i=0}^n$ belong to $\mathbb{L}^2(\Omega, \mathcal{A}, P)$, and that the covariance matrix $R_{0:n} = E(X_{0:n}X_{0:n}^H)$ of $X_{0:n}$ is invertible.

In the sequel, the general notation $\widetilde{X_i}^{\mathcal{H}(X_{l_1},\cdots,X_{l_k})}$ of section 3.1 is simplified to $\tilde{X}_i^{l_1,\cdots,l_k}$. Let $p \leq r, s \leq q$. Since we will essentially use contiguous sets of indices (without loss of generality), we also replace $\tilde{X}_i^{\mathcal{H}(\{X_m\}_{p\leq m\leq q})}$, $\tilde{X}_i^{\mathcal{H}(\{X_m\}^{m\neq r}_{p\leq m\leq q})}$ and $\tilde{X}_i^{\mathcal{H}(\{X_m\}^{m\neq r, m\neq s}_{p\leq m\leq q})}$ respectively by $\tilde{X}_i^{p:q}$, $\tilde{X}_i^{[p:q]\backslash r}$ and $\tilde{X}_i^{[p:q]\backslash r,s}$. Similar notations are adopted for the correlation coefficients, so that the partial correlation coefficient (of order $q-p$) $\rho_{X_i,X_j}^{\mathcal{H}(\{X_m\}^{m\neq r}_{p\leq m\leq q})}$, say, is denoted simply by $\rho_{i,j}^{[p:q]\backslash r}$. In our conventions, the order of the secondary (upper) indices p and q is meaningful : $\tilde{X}_i^{p:q}$, $\rho_{i,j}^{p:q}$ (and later on $R_{i,j}^{p:q}$, $r_{i,j}^{p:q}$, $P_{i,j}^{p:q}$ and $p_{i,j}^{p:q}$) reduce respectively to X_i, $\rho_{i,j}$, $R_{i,j}$, $r_{i,j}$, $P_{i,j}$ and $p_{i,j}$ if $p > q$. In this way the notation changes continuously from the total to the partial situation. For instance, there is no conceptual need to distinguish between total and partial correlation coefficients since a total correlation coefficient is simply a parcor of order 0.

In this section, we shall first recall (and slightly extend) some results [27] [28] [29] [30] giving the covariance matrix of $X_{0:n}$ (resp. its inverse) in terms of covariances of the random variables $\{X_i\}_{i=0}^n$ (resp. of the random variables $\{\tilde{X}_i^{[0:n]\backslash i}\}_{i=0}^n$). We thus get lemmas 3.1 and 3.2, which are generalized to theorem 3.1 by considering Schur complements in $R_{0:n}$ and in $R^{-1}_{0:n}$. Lastly these recursions receive a spherical trigonometry interpretation. We begin with the following elementary results.

Lemma 3.1 *Let $R_{0:n} = (\ r_{i,j}\ )_{i,j=0}^n$ and $P_{0:n} = R^{-1}_{0:n} = (\ p_{i,j}\ )_{i,j=0}^n$. Then for all $i, j \in [0 \cdots n]$,*

$$r_{i,j} = (X_i\ ,\ X_j)\ , \tag{28}$$

$$p_{i,j} = \left(\frac{\tilde{X}_i^{[0:n]\backslash i}}{(\tilde{X}_i^{[0:n]\backslash i}, \tilde{X}_i^{[0:n]\backslash i})}, \frac{\tilde{X}_j^{[0:n]\backslash j}}{(\tilde{X}_j^{[0:n]\backslash j}, \tilde{X}_j^{[0:n]\backslash j})}\right). \tag{29}$$

Lemma 3.2 *Let $R_{0:n} = (\ r_{i,j}\ )_{i,j=0}^n$ and $P_{0:n} = R^{-1}_{0:n} = (\ p_{i,j}\ )_{i,j=0}^n$. Then for all $i, j \in [0 \cdots n]$,*

$$\frac{r_{i,j}}{\sqrt{r_{i,i}r_{j,j}}} = (\overline{X_i}, \overline{X_j}) = \rho_{i,j}\ , \tag{30}$$

$$\frac{p_{i,j}}{\sqrt{p_{i,i}p_{j,j}}} = (\overline{\tilde{X}_i^{[0:n]\backslash i}}, \overline{\tilde{X}_j^{[0:n]\backslash j}}) = -\rho_{i,j}^{[0:n]\backslash i,j}. \tag{31}$$

We are now ready to extend lemma 3.2. Let $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ with A invertible. Then the Schur complement (M/A) of A in M is defined as $(M/A) = D - CA^{-1}B$. The following theorem encompasses and generalizes lemma 3.2 (which corresponds to the particular case $p = -1$) :

Theorem 3.1 *Let $R_{0:n} = (r_{i,j})_{i,j=0}^n$ and $P_{0:n} = R^{-1}_{0:n} = (p_{i,j})_{i,j=0}^n$. Let moreover $R^{0:p}_{0:n} = (r^{0:p}_{i,j})_{i,j=p+1}^n$ (resp. $P^{0:p}_{0:n} = (p^{0:p}_{i,j})_{i,j=p+1}^n$) be the Schur complement of $R_{0:p}$ in $R_{0:n}$ (of the $(p+1)\times(p+1)$ top left corner $[I_{p+1}0]P_{0:n}[I_{p+1}0]^T$ of $P_{0:n}$ in $P_{0:n}$). For $p=-1$, we set $R^{0:p}_{0:n} = R_{0:n}$, $P^{0:p}_{0:n} = P_{0:n}$, $r^{0:p}_{i,j} = r_{i,j}$, $p^{0:p}_{i,j} = p_{i,j}$, and $\rho^{0:p}_{i,j} = \rho_{i,j}$. Then for all $p \in [-1, 0, \cdots, n-1]$, and for all $i, j \in [p+1 \cdots n]$,*

$$r^{0:p}_{i,j}/\sqrt{r^{0:p}_{i,i}r^{0:p}_{j,j}} = \rho^{0:p}_{i,j}\ , \tag{32}$$

$$p^{0:p}_{i,j}/\sqrt{p^{0:p}_{i,i}p^{0:p}_{j,j}} = -\rho^{[p+1:n]\backslash i,j}_{i,j}\ . \tag{33}$$

We now turn to the connection with the spherical trigonometry cosine laws :

Corollary 3.1 *Up to normalization, an elementary (i.e. rank 1) Schur complement step on $R^{0:p}_{0:n}$ (resp. on $P^{0:p}_{0:n}$) performs the law of cosines (3) (resp. the polar law of cosines (18)) : For all $p \in [-1, 0, \cdots, n-2]$, and for all $i, j \in [p+2 \cdots n]$,*

$$\rho^{0:p+1}_{i,j} = \frac{\rho^{0:p}_{i,j} - \rho^{0:p}_{i,p+1}\rho^{0:p}_{p+1,j}}{\sqrt{1-|\rho^{0:p}_{i,p+1}|^2}\sqrt{1-|\rho^{0:p}_{p+1,j}|^2}}, \tag{34}$$

$$\rho^{[p+2:n]\backslash i,j}_{i,j} = \frac{\rho^{[p+1:n]\backslash i,j}_{i,j} + \rho^{[p+1:n]\backslash i,p+1}_{i,p+1}\rho^{[p+1:n]\backslash p+1,j}_{p+1,j}}{\sqrt{1-|\rho^{[p+1:n]\backslash i,p+1}_{i,p+1}|^2}\sqrt{1-|\rho^{[p+1:n]\backslash p+1,j}_{p+1,j}|^2}}. \tag{35}$$

Proof :

The Schur complementation step $R^{0:p}_{0:n} \to (R^{0:p}_{0:n}/r^{0:p}_{p+1,p+1}) = R^{0:p+1}_{0:n}$ reads componentwise : $r^{0:p+1}_{i,j} = (\tilde{X}_i^{0:p}, \tilde{X}_j^{0:p}) - (\tilde{X}_i^{0:p}, \tilde{X}_{p+1}^{0:p})\ (\tilde{X}_{p+1}^{0:p}, \tilde{X}_{p+1}^{0:p})^{-1}$

$(\tilde{X}_{p+1}^{0:p}, \tilde{X}_j^{0:p}) = (\tilde{X}_i^{0:p+1}, \tilde{X}_i^{0:p+1})$, due to (16). Normalizing as in section 3.1 we get (34).

We next consider (35). Similarly, from the quotient property of Schur complements (see e.g. [16]), we have $(P_{0:n}^{0:p}/p_{p+1,p+1}^{0:p}) = P_{0:n}^{0:p+1}$. But this equality reads componentwise

$$p_{i,j}^{0:p} - p_{i,p+1}^{0:p}(p_{p+1,p+1}^{0:p})^{-1}p_{p+1,j}^{0:p} = p_{i,j}^{0:p+1}.$$

Dividing by $\sqrt{p_{i,i}^{0:p}}\sqrt{p_{j,j}^{0:p}}$, and using (33), we get $\rho_{i,j}^{[p+1:n]\backslash i,j} + \rho_{i,p+1}^{[p+1:n]\backslash i,p+1}\rho_{p+1,j}^{[p+1:n]\backslash p+1,j} = \sqrt{p_{i,i}^{0:p+1}/p_{i,i}^{0:p}}$ $\rho_{i,j}^{[p+2:n]\backslash i,j}$ $\sqrt{p_{j,j}^{0:p+1}/p_{j,j}^{0:p}}$. Remarking from (29) and from the equality $P_{0:n}^{0:p} = P_{p+1:n}$ that $p_{i,i}^{0:p} = (\tilde{X}_i^{[p+1:n]\backslash i}, \tilde{X}_i^{[p+1:n]\backslash i})^{-1}$, and using (14), we see that $\sqrt{p_{i,i}^{0:p+1}/p_{i,i}^{0:p}} = \sqrt{1-|\rho_{i,p+1}^{[p+1:n]\backslash i,p+1}|^2}$. We thus get (35), which is the polar cosine law (19) = (18). ■

4 Non-Euclidean *(spherical)* geometrical aspects of the Schur and Levinson-Szegö algorithms

From now on, we shall further assume that $[X_0 X_1 \cdots X_n] = [X_t X_{t-1} \cdots X_{t-n}]$, where $\{X_t, t \in \mathbb{Z}\}$ is a zero-mean, discrete time, wide sense stationary time series. As a consequence, $R_{0:n}$ is a Toeplitz matrix. For simplicity, let us denote $R_{0:n}$ by R_n and $r_{i,j}$ by r_{j-i}. The parcors satisfy a shift-invariance property : for all $i, j, k, p, q \in \mathbb{Z}, p \leq q,\ i, j \notin \{p, \cdots, q\}$, $\rho_{i,j}^{p:q} = \rho_{i+k,j+k}^{p+k:q+k}$. Among all correlation coefficients (total or partial), the function $\{\rho(p) = \rho_{0,p}^{1:p-1}\}_{p \in \mathbb{N}^\star}$ (with $\rho(1) = \rho_{0,1}$, as in theorem 3.1) is the partial autocorrelation function of the process.

Let us now turn back to the Schur and Levinson-Szegö algorithms. In this final section, we shall write the commun (lattice) recursions of both algorithms as two Schur complement recursions (in the forward and backward directions), but acting on the covariance matrix (in the Schur case) or on its inverse, i.e., on the covariance matrix of the normalized interpolation process (in the Levinson-Szegö case). From section 3.3, the link with spherical trigonometry will follow immediately : up to normalization, the Schur (resp. inverse Levinson-Szegö) algorithm performs the law of cosines (resp. the polar law of cosines). This is a new feature of the classical duality of the Schur and Levinson-Szegö algorithms. As for the Levinson-Szegö algorithm, it is an implementation of the polar five elements formula.

4.1 Spherical geometry of the Schur algorithm

The new (spherical) geometrical interpretation of the algorithm stems from the connection between the Schur algorithm and linear regression. Let us thus initialize (6), via (1), with

$$s_0(z) = \frac{r_1 z + r_2 z^2 + \cdots}{r_0 + r_1 z + r_2 z^2 + \cdots} \ .$$

In this case, for all $p \geq 1$ the Schur parameter $s_p(0)$ is equal to the (partial) correlation coefficient $\rho_{0,p}^{1:p-1}$. It is convenient to write the algorithm in vector form [31] [32] : for $p \geq 1$,

$$\begin{bmatrix} u_{p-1}^0 & v_{p-1}^1 \\ u_{p-1}^1 & v_{p-1}^2 \\ \vdots & \vdots \\ u_{p-1}^k & v_{p-1}^{k+1} \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} 1 & -s_p(0) \\ -s_p^*(0) & 1 \end{bmatrix} = \begin{bmatrix} u_p^0 & 0 \\ u_p^1 & v_p^1 \\ \vdots & \vdots \\ u_p^k & v_p^k \\ \vdots & \vdots \end{bmatrix}, \quad (36)$$

with initialization $u_0^0 = r_0$, and $u_0^i = v_0^i = r_i$ for $i > 0$.

From the point of view of analytic interpolation theory, which was that of section 2.1, this p^{th} step of the algorithm incorporates the new data r_p in the covariance extension problem. This problem is recursive and "hierarchical" by nature : given $(r_0, \cdots, r_{p-1})$ such that R_{p-1} is positive definite (> 0), $R_p > 0$ if and only if r_p belongs to a disk (of decreasing radius $r_0 \prod_{i=1}^{p-1}(1-|\rho_{0,i}^{1:i-1}|^2)$), the center of which depends on $(r_0, \cdots, r_{p-1})$. So for all $k \geq 0$, the row number k of (36) integrates the contribution of the correlation lag r_p in the subsequent (possible) compatibibility of r_{p+k} with $(r_0, \cdots, r_{p+k-1})$. In particular, the row number zero tells whether r_p is compatible with the data $(r_0, \cdots, r_{p-1})$ via the following test : assuming that $R_{p-1} > 0$, $R_p > 0$ if and only if $|s_p(0) = v_{p-1}^1/u_{p-1}^0| < 1$.

This progressive incorporation of the constraints $r_2, \cdots, r_p, \cdots$ in the analytic interpolation problem corresponds to the progressive incorporation of the random variables $X_{t-1}, \cdots, X_{t-p+1}, \cdots$ in the linear prediction problem, and thus to the progressive updating of the associated projection operator (this, of course, is nothing but the classical lattice or Gram-Schmidt interpretation of the Schur algorithm [33]). To see this, let us rewrite the Schur algorithm in terms of projective identities. It is easily seen (by induction) that for $k \geq 0$, the two recursions of the row number k of (36) are two coupled occurrences of the same identity (16) :

$$\begin{aligned} &\left[(\tilde{X}_{t-p}^{t-p+1:t-1}, \tilde{X}_{t-p-k}^{t-p+1:t-1})(\tilde{X}_t^{t-p+1:t-1}, \tilde{X}_{t-p-k}^{t-p+1:t-1})\right] \\ &\times \begin{bmatrix} 1 & -\frac{(\tilde{X}_t^{t-p+1:t-1}, \tilde{X}_{t-p}^{t-p+1:t-1})}{(\tilde{X}_{t-p}^{t-p+1:t-1}, \tilde{X}_{t-p}^{t-p+1:t-1})} \\ -\frac{(\tilde{X}_{t-p}^{t-p+1:t-1}, \tilde{X}_t^{t-p+1:t-1})}{(\tilde{X}_t^{t-p+1:t-1}, \tilde{X}_t^{t-p+1:t-1})} & 1 \end{bmatrix} \\ &= \left[(\tilde{X}_{t-p}^{t-p+1:t}, \tilde{X}_{t-p-k}^{t-p+1:t})(\tilde{X}_t^{t-p:t-1}, \tilde{X}_{t-p-k}^{t-p:t-1})\right] \quad . \quad (37) \end{aligned}$$

Since all these quantities are covariances of estimation errors, they reduce to parcors when appropriately normalized; so a connection of the recursive equations (37) with spherical trigonometry is expected.

In fact, both equations are easily seen to be Schur complement recursions in $R_{0:n}^{1:p-1} \stackrel{def}{=} ((\tilde{X}_{t-i}^{t-p+1:t-1}, \tilde{X}_{t-j}^{t-p+1:t-1}))_{i,j=0}^n$. These two Schur complementation steps correspond to augmenting the set of variables $\{X_{t-i}\}_{i=1}^{p-1}$ in the projective space in its two (contiguous) opposite directions : the forward one X_t

and the backward one X_{t-p}. Because of stationarity, the resulting quantities still are covariances of estimation residuals with respect to the same subspace, because the right hand side of (37) also reads $[(\tilde{X}_{t-p-1}^{t-p:t-1}, \tilde{X}_{t-p-k-1}^{t-p:t-1})$ $(\tilde{X}_{t}^{t-p:t-1}, \tilde{X}_{t-p-k}^{t-p:t-1})]$, and the two coefficients in the transformation matrix reduce to $-\rho_{0,p}^{1:p-1}$ and $-(\rho_{0,p}^{1:p-1})^*$. From the discussion in section 3, the link with spherical trigonometry is immediate :

Theorem 4.1

Up to normalization, an elementary step of the Schur algorithm performs two coupled occurrences of the law of cosines : for all $p \geq 1$, and for all $k \geq 0$,

$$
\left[\ \rho_{p,p+k}^{1:p-1}\ ,\ \rho_{0,p+k}^{1:p-1}\ \right]
\left[\begin{array}{cc}
\frac{1}{\sqrt{1-|\rho_{p,0}^{1:p-1}|^2}} & \frac{-\rho_{0,p}^{1:p-1}}{\sqrt{1-|\rho_{0,p}^{1:p-1}|^2}} \\
\frac{-\rho_{p,0}^{1:p-1}}{\sqrt{1-|\rho_{p,0}^{1:p-1}|^2}} & \frac{1}{\sqrt{1-|\rho_{0,p}^{1:p-1}|^2}}
\end{array}\right] =
$$

$$
\left[\ \rho_{p,p+k}^{0:p-1}\sqrt{1-|\rho_{0,p+k}^{1:p-1}|^2}, \rho_{0,p+k}^{1:p}\sqrt{1-|\rho_{p,p+k}^{1:p-1}|^2}\ \right]. \quad (38)
$$

<u>Proof.</u>

Divide (37) by $(\tilde{X}_{t}^{t-p+1:t-1}, \tilde{X}_{t}^{t-p+1:t-1})^{1/2}(\tilde{X}_{t-p-k}^{t-p+1:t-1}, \tilde{X}_{t-p-k}^{t-p+1:t-1})^{1/2}$, which is equal to $(\tilde{X}_{t-p}^{t-p+1:t-1}, \tilde{X}_{t-p}^{t-p+1:t-1})^{1/2}$ $(\tilde{X}_{t-p-k}^{t-p+1:t-1}, \tilde{X}_{t-p-k}^{t-p+1:t-1})^{1/2}$, and use (14). ■

4.2 Spherical geometry of the Levinson-Szegö algorithm

We now turn to the spherical geometry of the Levinson-Szegö algorithm. Remember from Theorem 3.1 that successive Schur complements in $R_{0:n}$ (resp. in $R_{0:n}^{-1}$) correspond to an increase (resp. a reduction) in the number of variables in the regression problem. So, as was already the case at the end of section 2, the comparison with the Schur algorithm indeed proves easier when dealing with order-decreasing recursions.

Let us introduce the forward $j^{th}-$order linear prediction coefficients $\{-a_i^j\}_{i=1}^j$ by $\tilde{X}_t^{t-j:t-1} = \sum_{i=0}^j a_i^j X_{t-i}$ with $a_0^j = 1$. From the Wiener Hopf equations, and Theorem 3.1, we get $a_i^j = (\tilde{X}_{t-i}^{[t-j:t]\backslash t-i}, \tilde{X}_{t}^{[t-j:t]\backslash t})/$ $(\tilde{X}_{t-i}^{[t-j:t]\backslash t-i}, \tilde{X}_{t-i}^{[t-j:t]\backslash t-i})$. So the order-decreasing Levinson-Szegö recursions read :

$$
\left[\underbrace{\frac{(\tilde{X}_{t-k}^{[t-p:t]\backslash t-k}, \tilde{X}_{t}^{t-p:t-1})}{(\tilde{X}_{t-k}^{[t-p:t]\backslash t-k}, \tilde{X}_{t-k}^{[t-p:t]\backslash t-k})}}_{a_k^p}\ \ \underbrace{\frac{(\tilde{X}_{t-k}^{[t-p:t]\backslash t-k}, \tilde{X}_{t-p}^{t-p+1:t})}{(\tilde{X}_{t-k}^{[t-p:t]\backslash t-k}, \tilde{X}_{t-k}^{[t-p:t]\backslash t-k})}}_{(a_{p-k}^p)^*}\right]
$$

$$
\times\left[\begin{array}{cc}
\frac{1}{1-|\rho_{p,0}^{1:p-1}|^2} & \frac{\rho_{0,p}^{1:p-1}}{1-|\rho_{0,p}^{1:p-1}|^2} \\
\frac{\rho_{p,0}^{1:p-1}}{1-|\rho_{p,0}^{1:p-1}|^2} & \frac{1}{1-|\rho_{0,p}^{1:p-1}|^2}
\end{array}\right] =
$$

$$
\left[\underbrace{\frac{(\tilde{X}_{t-k}^{[t-p+1:t]\backslash t-k}, \tilde{X}_{t}^{t-p+1:t-1})}{(\tilde{X}_{t-k}^{[t-p+1:t]\backslash t-k}, \tilde{X}_{t-k}^{[t-p+1:t]\backslash t-k})}}_{a_k^{p-1}}\ ,\right.
$$

$$
\left.\underbrace{\frac{(\tilde{X}_{t-k}^{[t-p:t-1]\backslash t-k}, \tilde{X}_{t-p}^{t-p+1:t-1})}{(\tilde{X}_{t-k}^{[t-p:t-1]\backslash t-k}, \tilde{X}_{t-k}^{[t-p:t-1]\backslash t-k})}}_{(a_{p-k}^{p-1})^*}\right] \quad (39)
$$

(the first equation is valid for $1 \leq k \leq p$, with $p \geq 2$, and the second for $0 \leq k \leq p-1$ with $p \geq 2$).

These equations are Schur complement recursions in $R_{0:p}^{-1}$; they correspond to reducing the set of variables $\{X_{t-i}\}_{i=0}^p$ in the projective space in its two (extremum) opposite directions : the forward one X_t and the backward one X_{t-p}. From the discussion in section 3.3, we thus expect that, after appropriate normalization of the covariances of the estimation errors, (39) will reduce to some spherical trigonometry polar law.

This hint is enforced when looking at the random variables in the left hand side of (37) and (39). Let $\mathcal{Z} = (\{X_{t-i}\}_{0 \leq i \leq p}, X_{t-m})$, $(A, B, C) = (X_t, X_{t-p}, X_{t-m})$, and $\mathcal{M} = \mathcal{H}(\mathcal{Z} \setminus \{X_t, X_{t-p}, X_{t-m}\})$. So $\mathcal{M} = \mathcal{H}(\{X_{t-i}\}_{1 \leq i \leq p-1})$ if $m \geq p$, and $\mathcal{M} = \mathcal{H}(\{X_{t-i}\}_{1 \leq i \leq p-1}^{i \neq m})$ if $1 \leq m \leq p-1$. Then (37) can be visualized as projective identities within the tetrahedron $(\overline{\tilde{A}^{\mathcal{M}}}, \overline{\tilde{B}^{\mathcal{M}}}, \overline{\tilde{C}^{\mathcal{M}}})$, and (39) as projective identities within the tetrahedron $(\overline{\tilde{A}^{\mathcal{M},B,C}}, \overline{\tilde{B}^{\mathcal{M},C,A}}, \overline{\tilde{C}^{\mathcal{M},A,B}})$, which can easily be shown [14] to be the polar tetrahedron of $(\overline{\tilde{A}^{\mathcal{M}}}, \overline{\tilde{B}^{\mathcal{M}}}, \overline{\tilde{C}^{\mathcal{M}}})$.

Theorem 4.2

Up to normalization, an elementary step of the order-decreasing Levinson-Szegö algorithm (resp. of the Levinson-Szegö algorithm) performs two coupled occurrences of the polar law of cosines (resp. of the polar five elements formula) :
For all $p \geq 2$, and for all $1 \leq k \leq p-1$,

$$
\left[\rho_{k,0}^{[0:p]\backslash k,0}, \rho_{k,p}^{[0:p]\backslash k,p}\right]
\left[\begin{array}{cc}
\frac{1}{\sqrt{1-|\rho_{p,0}^{[0:p]\backslash p,0}|^2}} & \frac{\rho_{0,p}^{[0:p]\backslash 0,p}}{\sqrt{1-|\rho_{0,p}^{[0:p]\backslash 0,p}|^2}} \\
\frac{\rho_{p,0}^{[0:p]\backslash p,0}}{\sqrt{1-|\rho_{p,0}^{[0:p]\backslash p,0}|^2}} & \frac{1}{\sqrt{1-|\rho_{0,p}^{[0:p]\backslash 0,p}|^2}}
\end{array}\right] =
$$

$$
\left[\rho_{k,0}^{[0:p]\backslash k,0,p}\sqrt{1-|\rho_{k,p}^{[0:p]\backslash k,p}|^2}, \rho_{k,p}^{[0:p]\backslash k,p,0}\sqrt{1-|\rho_{k,0}^{[0:p]\backslash k,0}|^2}\right], \quad (40)
$$

and

$$
\left[\rho_{k,0}^{[0:p]\backslash k,0,p}\sqrt{1-|\rho_{k,p}^{[0:p]\backslash k,p}|^2}, \rho_{k,p}^{[0:p]\backslash k,p,0}\sqrt{1-|\rho_{k,0}^{[0:p]\backslash k,0}|^2}\right]
$$

$$
\left[\begin{array}{cc}
\frac{1}{\sqrt{1-|\rho_{p,0}^{[0:p]\backslash p,0}|^2}} & \frac{-\rho_{0,p}^{[0:p]\backslash 0,p}}{\sqrt{1-|\rho_{0,p}^{[0:p]\backslash 0,p}|^2}} \\
\frac{-\rho_{p,0}^{[0:p]\backslash p,0}}{\sqrt{1-|\rho_{p,0}^{[0:p]\backslash p,0}|^2}} & \frac{1}{\sqrt{1-|\rho_{0,p}^{[0:p]\backslash 0,p}|^2}}
\end{array}\right] = \left[\rho_{k,0}^{[0:p]\backslash k,0}, \rho_{k,p}^{[0:p]\backslash k,p}\right]. \quad (41)
$$

Proof :

Using (15), (39) can be rewritten as :

$$\left[\frac{(\tilde{X}_t^{[t-p:t]\backslash t}, \tilde{X}_t^{[t-p:t]\backslash t})^{1/2}}{(\tilde{X}_{t-k}^{[t-p:t]\backslash t-k}, \tilde{X}_{t-k}^{[t-p:t]\backslash t-k})^{1/2}}(-\rho_{k,0}^{[0:p]\backslash k,0}) \right. ,$$

$$\left. \frac{(\tilde{X}_{t-p}^{[t-p:t]\backslash t-p}, \tilde{X}_{t-p}^{[t-p:t]\backslash t-p})^{1/2}}{(\tilde{X}_{t-k}^{[t-p:t]\backslash t-k}, \tilde{X}_{t-k}^{[t-p:t]\backslash t-k})^{1/2}}(-\rho_{k,p}^{[0:p]\backslash k,p})\right]$$

$$\times \begin{bmatrix} \frac{1}{1-|\rho_{p,0}^{[0:p]\backslash p,0}|^2} & \frac{\rho_{0,p}^{[0:p]\backslash 0,p}}{1-|\rho_{0,p}^{[0:p]\backslash 0,p}|^2} \\ \frac{\rho_{p,0}^{[0:p]\backslash p,0}}{1-|\rho_{p,0}^{[0:p]\backslash p,0}|^2} & \frac{1}{1-|\rho_{0,p}^{[0:p]\backslash 0,p}|^2} \end{bmatrix} =$$

$$\left[\frac{(\tilde{X}_t^{[t-p+1:t]\backslash t}, \tilde{X}_t^{[t-p+1:t]\backslash t})^{1/2}}{(\tilde{X}_{t-k}^{[t-p+1:t]\backslash t-k}, \tilde{X}_{t-k}^{[t-p+1:t]\backslash t-k})^{1/2}}(-\rho_{k,0}^{[0:p]\backslash k,0,p}) \right. ,$$

$$\left. \frac{(\tilde{X}_{t-p}^{[t-p:t-1]\backslash t-p}, \tilde{X}_{t-p}^{[t-p:t-1]\backslash t-p})^{1/2}}{(\tilde{X}_{t-k}^{[t-p:t-1]\backslash t-k}, \tilde{X}_{t-k}^{[t-p:t-1]\backslash t-k})^{1/2}}(-\rho_{k,p}^{[0:p]\backslash k,p,0})\right] .$$

Next divide by $(\tilde{X}_t^{t-p+1:t-1}, \tilde{X}_t^{t-p+1:t-1})^{1/2}/(\tilde{X}_{t-k}^{[t-p:t]\backslash t-k}, \tilde{X}_{t-k}^{[t-p:t]\backslash t-k})^{1/2}$, which is equal to $(\tilde{X}_{t-p}^{t-p+1:t-1}, \tilde{X}_{t-p}^{t-p+1:t-1})^{1/2}/ (\tilde{X}_{t-k}^{[t-p:t]\backslash t-k}, \tilde{X}_{t-k}^{[t-p:t]\backslash t-k})^{1/2}$. Using (14), we get (40) = (41). ■

5 Conclusion

In this paper, we addressed non-Euclidean geometrical aspects of the Schur and Levinson-Szegö algorithms. We showed that the Lobachevski geometry is, by construction, the natural geometrical environment of these algorithms, since they call for automorphisms of the unit disk. By considering the algorithms in the particular context of their application to linear prediction, we next gave them a new interpretation in terms of spherical trigonometry. The role of Schur complementation in linear regression analysis was emphasized, because of the natural link between this basic algebraic mechanism and the spherical trigonometry cosine laws. Lastly, the Schur (resp. Levinson-Szegö) algorithm received a direct (resp. polar) spherical trigonometry interpretation, which is a new feature of the classical duality of both algorithms.

Finally, let us briefly mention that these interpretations provide the algorithms with structural constraints of a geometrical nature. The Lobachevski invariants are the Poincaré distance and the cross ratio (because of the use of linear fractional transformations), and those of spherical trigonometry are expressed by the relations among parcors given in section 3.2. These constraints could hopefully be used in the design of practical algorithms; this point is currently under investigation.

References

[1] I. Schur, *Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind*, Journal für die Reine und Angewandte Mathematik, Vol. 147, pp. 205-32, 1917; and 148, pp. 122-45, 1918. English translation : *On power series which are bounded in the interior of the unit circle*, in Operator Theory, Advances and Applications 18 (I. Gohberg, ed.), part I, pp. 31-59; and part II, pp. 61-88, Birkhäuser Verlag, Basel, 1986.

[2] G. Szegö, *Orthogonal polynomials*, American Mathematical Society, Providence, 1939.

[3] N. I. Akhiezer, *The classical moment problem and some related questions in analysis*, Hafner publishing Company, New York, 1965.

[4] U. Grenander and G. Szegö, *Toeplitz forms and their applications*, University of California Press, Berkeley, 1958.

[5] T. Kailath, *A view of three decades of linear filtering theory*, IEEE Transactions on Information Theory, Vol. 20, N. 2, pp. 146-81, 1974.

[6] P. Delsarte, Y. Genin, Y. Kamp and P. Van Dooren, *Speech modeling and the trigonometric moment problem*, Philips J. Res., Vol. 37, N. 5/6, pp. 277-92, 1982.

[7] P. Dewilde and H. Dym, *Schur recursions, error formulas, and convergence of rational estimators for stationary stochastic sequences*, IEEE Transactions on Information theory, Vol. 27, N. 4, pp. 446-61, 1981.

[8] P. Dewilde and H. Dym, *Lossless inverse scattering, digital filters, and estimation theory*, IEEE Transactions on Information theory, Vol. 30, N. 4, pp. 644-61, 1984.

[9] I. Gohberg (ed.), *I. Schur methods in Operator Theory and Signal Processing*, Operator theory : advances and applications, Vol. 18, Birkhäuser Verlag, Basel 1986.

[10] C. Foias and A. E. Frazho, *The Commutant lifting approach to interpolation problems*, Operator Theory : advances and applications, Vol. 44, Birkäuser Verlag, Basel 1990.

[11] T. Constantinescu, *Schur parameters, factorization and dilation problems*, Operator Theory : advances and applications, Vol. 82, Birkäuser Verlag, Basel 1991.

[12] G. U. Yule, *On the theory of correlation for any number of variables, treated by a new system of notations*, Proc. Roy. Soc., Vol. 79A, pp. 182-93, 1907. also in Statistical papers of George Udny Yule, Griffin, London, pp. 85-96, 1978.

[13] F. Desbouvries, *Spherical Trigonometry, Yule's PARCOR Identity and FRLS fully-normalized lattice*, IEEE Transactions on Signal Processing, Vol. 44, N. 2, pp. 427–430, 1996.

[14] F. Desbouvries, *Trigonométrie sphérique, identité de Yule entre PARCORs et algorithmes MCRR*, Traitement du Signal, Vol. 13, N. 4, pp. 303-18, 1996

[15] C. Chang and T. T. Georgiou, *Geometric aspects of the Carathéodory extension problem*, Linear algebra and its applications, Vol. 203-04, pp. 209-51, 1994.

[16] D. V. Ouelette, *Schur complements and statistics*, Linear Algebra and its Applications, Vol. 36, pp. 187-295, 1981.

[17] C. Carathéodory, *Theory of functions of a complex variable*, Chelsea Pub. Company, New York 1954.

[18] B.A. Dubrovin, A.T. Fomenko and S.P. Novikov, *Modern geometry - Methods and applications. Part I*, 2^{nd} edition, Springer Verlag, New York 1993.

[19] R. Remmert, *Theory of complex functions*, Springer-Verlag, New York 1991.

[20] L. V. Ahlfors, *Conformal invariants - Topics in geometric function theory*, McGraw Hill book company, New York 1973.

[21] M. Jarnicki and P. Pflug, *Invariant distances and metrics in complex analysis*, Walter de Gruyter, Berlin 1993.

[22] R. B. Burckel, *An introduction to classical complex analysis. Volume 1*, Birkäuser Verlag, Basel 1979.

[23] J. G. Ratcliffe, *Foundations of hyperbolic manifolds*, Springer-Verlag, New York 1994.

[24] Y. Genin, *An introduction to the modern theory of positive functions and some of its today applications to signal processing, circuits and systems problems*, in Proceedings of the ECCTD Conference, Paris, France, pp. 195-233, 1987.

[25] L. M. Kells, W. F. Kerns and J. R. Bland, *Plane and Spherical Trigonometry*, McGraw Hill, New York, 1951.

[26] E. G. Kogbeliantz, *Fundamentals of mathematics from an advanced viewpoint. Vol. 4: solid geometry and spherical trigonometry*, Gordon and Breach, New York, 1969.

[27] T. W. Anderson, *An introduction to multivariate statistical analysis*, 2^{nd} ed., Wiley, New York, 1984.

[28] H. Cramer, *Mathematical methods of statistics*, 13^{th} edition, Princeton University press, Princeton, NJ, 1974.

[29] J. Whittaker, *Graphical models in applied multivariate statistics*, John Wiley and sons, Chichester, 1990.

[30] S. S. Wilks, *Mathematical statistics*, John Wiley and sons, New York, 1962.

[31] T. Kailath, *A theorem of I. Schur and its impact on modern signal processing*, in Operator Theory: Advances and Applications, Vol. 18 (I. Gohberg, ed.), pp. 9-30, Birkäuser Verlag, Basel, 1986.

[32] T. Kailath, *Signal processing applications of some moment problems*, Proceedings of Symposia in Applied Mathematics 37, pp. 71-109, American Mathematical Society, 1987.

[33] B. Friedlander, *Lattice filters for adaptive processing*, Proceedings of the IEEE, Vol. 70, N. 8, pp. 829-68, 1982.