

New Square-Root Factorization of Inverse Toeplitz Matrices

Bo Wahlberg, *Fellow, IEEE* and Petre Stoica, *Fellow, IEEE*

Abstract—Square-root (in particular, Cholesky) factorization of Toeplitz matrices and of their inverses is a classical area of research. The Schur algorithm yields directly the Cholesky factorization of a symmetric Toeplitz matrix, whereas the Levinson algorithm does the same for the inverse matrix. The objective of this letter is to use results from the theory of rational orthonormal functions to derive square-root factorizations of the inverse of an $n \times n$ positive definite Toeplitz matrix. The main result is a new factorization based on the Takenaka-Malmquist functions, that is parameterized by the roots of the corresponding auto-regressive polynomial of order n . We will also discuss briefly the connection between our analysis and some classical results such as Schur polynomials and the Gohberg-Semencul inversion formula.

Index Terms—Toeplitz matrix, Square-root and Cholesky factorization, AR processes, Rational orthonormal functions.

I. INTRODUCTION

Consider the symmetric positive-definite Toeplitz matrices (for given covariances $\{r_k \in \mathbf{R}\}$ and for $j = 1, \dots, (n+1)$)

$$R_j = \begin{bmatrix} r_0 & r_1 & \cdots & r_{j-1} \\ r_1 & r_0 & \cdots & r_{j-2} \\ \vdots & \ddots & \ddots & \vdots \\ r_{j-1} & \cdots & r_1 & r_0 \end{bmatrix} > 0. \quad (1)$$

Let $\{a_{j,1}, \dots, a_{j,j}\}$ and σ_j^2 be the solution to

$$R_{j+1} \begin{bmatrix} 1 \\ a_{j,1} \\ \vdots \\ a_{j,j} \end{bmatrix} = \begin{bmatrix} \sigma_j^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad j = 0 \dots n. \quad (2)$$

This implies that $\{a_{j,k}, k = 1 \dots j\}$ are the coefficients of the auto-regressive (AR) process of order j , with innovation variance σ_j^2 , whose covariances at lags 0 to j exactly match r_0 to r_j . Define

$$L_{j+1} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{j,1} & 1 & \cdots & 0 \\ a_{j,2} & a_{j-1,1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{j,j} & a_{j-1,j-1} & \cdots & a_{1,1} & 1 \end{bmatrix}. \quad (3)$$

Copyright (c) 2008 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org

B. Wahlberg is with the Automatic Control Lab and ACCESS, School of Electrical Engineering, KTH, SE-100 44 Stockholm, Sweden. (e-mail: bo.wahlberg@ee.kth.se.)

P. Stoica is with the Department of Information Technology, Uppsala University, SE-75105 Uppsala, Sweden (e-mail: ps@it.uu.se)

This work was partially supported by the Swedish Research Council and the Linnaeus Center ACCESS at KTH

Then

$$R_{j+1}^{-1} = L_{j+1} D_{j+1}^{-1} L_{j+1}^T, \quad (4)$$

where

$$D_{j+1} = \text{diag}(\sigma_j^2, \sigma_{j-1}^2, \dots, \sigma_0^2)$$

is a diagonal matrix, see e.g. [1]-[2], and the many references therein. This well-known result, which forms the foundation of the Levinson algorithm, shows that $L_n D_n^{-1/2}$ is a lower triangular Cholesky factor of the inverse Toeplitz matrix R_n^{-1} . The factorization in (4) is a function of the coefficients of all AR models of orders 0 to $n-1$. In contrast to (4), the Gohberg-Semencul (G-S) formula for R_{n+1}^{-1} , see, e.g., [3],[2], and [4], (which is not of the Cholesky or, more generally, square-root factorization type) is a function only of the coefficients of the AR model of order n , i.e. $\{a_{n,1}, \dots, a_{n,n}, \sigma_n^2\}$. There is also a similar formula for R_n^{-1} , see e.g. [2], which again is an explicit function of the coefficients of the AR model of order n .

The objective of this letter is to show how the theory of rational orthonormal functions can be used for finding a class of square-root factorizations of R_n^{-1} or, if desired, of R_{n+1}^{-1} . In particular, using the Takenaka-Malmquist functions, [5, page 18], we obtain an apparently new square-root factorization of R_n^{-1} that is an explicit function of the roots of the n th order AR polynomial. We should stress that, similarly to the G-S theory, our main contribution consists of a theoretical inversion formula, rather than a computational inversion algorithm.

II. PRELIMINARIES ON RATIONAL ORTHONORMAL FUNCTIONS

Define the normalized n th order AR polynomial associated with the given covariances $\{r_k\}_{k=0}^n$:

$$A_n(z) = \frac{1}{\sigma_n} (z^n + a_{n,1} z^{n-1} + \dots + a_{n,n}), \quad (5)$$

and let

$$v(z) = [z^{n-1} \ z^{n-2} \ \dots \ 1]^T, \quad V(z) = \frac{v(z)}{A_n(z)}. \quad (6)$$

Since the covariances are given by

$$r_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i\omega k}}{|A_n(e^{i\omega})|^2} d\omega, \quad k = 0, 1, 2, \dots,$$

we have

$$R_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} V(e^{i\omega}) V^*(e^{i\omega}) d\omega,$$

where the superscript $*$ denotes the conjugate transpose. The following function, where $z_1, z_2 \in \mathbf{C}$, will play an important role in the analysis of this letter:

$$\begin{aligned}\Delta(z_1, z_2) &= \frac{v^T(1/z_1)R_n^{-1}v(z_2)}{A_n(1/z_1)A_n(z_2)} \\ &= V^T(1/z_1) \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} V(e^{i\omega})V^*(e^{i\omega})d\omega \right]^{-1} V(z_2).\end{aligned}$$

The function $\Delta(z_1, z_2)$ is called the reproducing kernel for the space spanned by the elements of $V(z)$. A key observation is that $\Delta(z_1, z_2)$ is invariant under a change of coordinates $\tilde{V}(z) = TV(z)$, $V(z) = T^{-1}\tilde{V}(z)$, where T is an $n \times n$ invertible matrix. Indeed

$$\begin{aligned}\Delta(z_1, z_2) &= \tilde{V}^T(1/z_1)T^{-T} \\ &\times \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} T^{-1}\tilde{V}(e^{i\omega})\tilde{V}^*(e^{i\omega})T^{-T}d\omega \right]^{-1} T^{-1}\tilde{V}(z_2) \\ &= \tilde{V}^T(1/z_1) \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{V}(e^{i\omega})\tilde{V}^*(e^{i\omega})d\omega \right]^{-1} \tilde{V}(z_2).\end{aligned}$$

A natural idea is then to use an ortho-normal basis for the space spanned by $\{z^k/A_n(z), k = 0, \dots, (n-1)\}$, for which

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{V}(e^{i\omega})\tilde{V}^*(e^{i\omega})d\omega = I,$$

(I denotes the identity matrix) and, therefore,

$$\Delta(z_1, z_2) = \tilde{V}^T(1/z_1)\tilde{V}(z_2).$$

The theory of rational ortho-normal (ON) functions is well developed. Two books on this subject are [6] and [5], which both contain extensive reference lists. Let $\{\xi_j\}_{j=1}^n$ be the roots of $A_n(z) = 0$, i.e.

$$A_n(z) = \frac{1}{\sigma_n} \prod_{j=1}^n (z - \xi_j).$$

Furthermore, assume that the roots satisfy $|\xi_j| < 1$, which is always the case under (1), see e.g. [1]. Applying the Gram-Schmidt procedure to the functions $\{1/(z - \xi_j)^k\}, k = 1 \dots n_j$ (where n_j is the multiplicity of root j) over all roots (these functions span the same space as the elements of $V(z)$), (6) leads to the Takenaka-Malmquist (T-M) ON functions, [5],

$$B_k(z) = \frac{\sqrt{1 - |\xi_k|^2}}{z - \xi_k} \prod_{i=1}^{k-1} \left[\frac{1 - \xi_i^* z}{z - \xi_i} \right], \quad k = 1, 2, \dots, n. \quad (7)$$

The T-M construction is valid for arbitrary root configurations with $|\xi_j| < 1$. The special case of $A(z) = (z - a)^n$, i.e. repeated roots in $\xi_j = a, -1 < a < 1$, leads to the well-known Laguerre functions, which further simplify to the delay functions $B_k(z) = z^{-k}$ for $\xi_j = 0$. In the case of complex conjugate roots, say ξ_j and $\xi_{j+1} = \xi_j^*$, define $(z - \xi_j)(z - \xi_{j+1}) = z^2 + b_j(c_j - 1)z - c_j$ with $-1 < b_j < 1$ and $-1 < c_j < 0$. Then take

$$\begin{aligned}B_j(z) &= \frac{\sqrt{1 - c_j^2}(z - b_j)}{z^2 + b_j(c_j - 1)z - c_j} \prod_{i=1}^{j-1} \left[\frac{1 - \xi_i^* z}{z - \xi_i} \right] \\ B_{j+1}(z) &= \frac{\sqrt{1 - c_j^2}\sqrt{1 - b_j^2}}{z^2 + b_j(c_j - 1)z - c_j} \prod_{i=1}^{j-1} \left[\frac{1 - \xi_i^* z}{z - \xi_i} \right]\end{aligned}$$

to obtain ON functions with real coefficients, [5]. For repeated complex roots, the functions obtained by this construction are often referred to as Kautz functions. The vector function made from $\{B_k(z)\}$ is $\tilde{V}(z) = [B_1(z), \dots, B_n(z)]^T$ and thus

$$\Delta(z_1, z_2) = \sum_{k=1}^n B_k(1/z_1)B_k(z_2) \quad (8)$$

Another set of ON functions can be constructed from the so-called Schur polynomials $A_k(z)$,

$$A_k(z) = \frac{1}{\sigma_k} (z^k + a_{k,1}z^{k-1} + \dots + a_{k,k}), \quad k = 0, \dots, n, \quad (9)$$

as follows, [5, Section 2.6.2],

$$B_k(z) = \frac{A_{k-1}(z)}{A_n(z)}, \quad k = 1, \dots, n,$$

where $\{a_{j,1}, \dots, a_{j,j}, \sigma_j\}$ are the solutions to (2) for $j = 0$ to n . There are simple algorithms to find $A_k(z)$ directly from $A_n(z)$ using Schur iterations.

In a somewhat different context, we remark on the fact that an important result in the theory of rational ON functions is the so-called Christoffel-Darboux formula, see [5]

$$\begin{aligned}\tilde{V}^T(1/z_1)\tilde{V}(z_2) &= \frac{H(1/z_1)H(z_2) - 1}{1 - z_2/z_1}, \\ H(z) &= \frac{z^n A_n(1/z)}{A_n(z)}.\end{aligned} \quad (10)$$

We will make use of (10) to briefly show how to connect the square-root factorization approach of this letter to the G-S formula for R_n^{-1} .

III. NEW SQUARE-ROOT FACTORIZATION

Using expression (8), we have that

$$\begin{aligned}v^T(1/z_1)R_n^{-1}v(z_2) &= A_n(1/z_1)A_n(z_2)\Delta(z_1, z_2) \\ &= \sum_{k=1}^n C_k(1/z_1)C_k(z_2),\end{aligned} \quad (11)$$

$$C_k(z) = A_n(z)B_k(z) \quad (12)$$

It is possible to recover the matrix R_n^{-1} from (11) by noting that the vectors $v(e^{i\omega_n k}), \omega_n = 2\pi/n$, for $k = 0, \dots, (n-1)$ are orthogonal (a well known result from the Discrete Fourier Transform (DFT) theory). Define the $n \times n$ DFT matrix

$$W_n = \frac{1}{\sqrt{n}} \left[v(e^{i\omega_n 0}) \ v(e^{i\omega_n 1}) \ \dots \ v(e^{i\omega_n (n-1)}) \right],$$

for which $W_n^* W_n = I$. We also define the row vectors

$$\gamma_k = \frac{1}{\sqrt{n}} \left[C_k(e^{i\omega_n 0}) \ C_k(e^{i\omega_n 1}) \ \dots \ C_k(e^{i\omega_n (n-1)}) \right].$$

Using this notation we can write from (12):

$$\begin{aligned} W_n^* R_n^{-1} W_n &= \sum_{k=1}^n \gamma_k^* \gamma_k \Rightarrow \\ R_n^{-1} &= \sum_{k=1}^n [\gamma_k W_n^*]^* \gamma_k W_n^* = [\Gamma_n W_n^*]^* \Gamma_n W_n^*, \\ \Gamma_n &= \begin{bmatrix} \gamma_n \\ \gamma_{n-1} \\ \vdots \\ \gamma_1 \end{bmatrix}. \end{aligned} \quad (13)$$

We have thus derived a square-root factorization¹ of R_n^{-1} . Note that the functions $C_k(z) = A_n(z)B_k(z)$ are polynomials of degree $n-1$,

$$C_k(z) = c_{k,1}z^{n-1} + c_{k,2}z^{n-2} + \dots + c_{k,n},$$

for all choices of minimal (of dimension n) ON bases that span the same space as $\{z^k/A_n(z)\}_{k=1}^{(n-1)}$. This means that we can write $\Gamma_n = U_n^T W_n$ with

$$U_n = \begin{bmatrix} c_{n,1} & \dots & c_{1,1} \\ \vdots & \ddots & \vdots \\ c_{n,n} & \dots & c_{1,n} \end{bmatrix}. \quad (14)$$

and therefore the factorization (13) can be re-written

$$R_n^{-1} = U_n U_n^T. \quad (15)$$

The simplest case is that of Schur ON polynomials, see (9), for which $C_k(z) = A_{k-1}(z)$. For the Takenaka-Malmquist functions we have

$$\begin{aligned} C_k(z) &= \frac{\sqrt{1-|\xi_k|^2}}{\sigma_n} \prod_{i=1}^{k-1} [1 - \xi_i^* z] \prod_{i=k+1}^n [z - \xi_i], \\ k &= 1, \dots, n, \end{aligned} \quad (16)$$

with the modification described in the previous section to handle complex-valued roots.

To summarize, let $\{B_k(z)\}$ be a rational ON basis for the space spanned by $\{z^k/A_n(z)\}_{k=0}^{(n-1)}$, and define $C_k(z) = A_n(z)B_k(z)$, $k = 1 \dots n$ and U_n as in (14). The key result of this letter is the general square-root factorization formula (15).

In the Schur polynomials case (15) reduces to the well-known formula (3) with $L_n D_n^{-1/2} = U_n$. On the other hand for the Takenaka-Malmquist functions we obtain a novel factorization of R_n^{-1} that depends only on the zeros of $A_n(z)$ and on σ_n , c.f. (16).

Note that it is easy to modify the result (15) to factorize R_{n+1}^{-1} , instead of R_n^{-1} , still using only $A_n(z)$. Let

$$v_{n+1}(z) = [z^n \ z^{n-1} \ z^{n-2} \ \dots \ 1]^T, \quad S(z) = \frac{v_{n+1}(z)}{A_n(z)},$$

¹By a matrix square-root factorization of M we mean $M = NN^*$. This includes the Cholesky factorization of M , for which N has a lower or upper triangular structure. It also includes the symmetric (Hermitian) matrix square root R of M , for which $M = RR$.

for which

$$R_{n+1} = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(e^{i\omega}) S^*(e^{i\omega}) d\omega.$$

By extending $\{B_1(z) \dots B_n(z)\}$ with $B_{n+1}(z) = 1$, we obtain an ON basis for the $n+1$ dimensional space spanned by $\{z^k/A_n(z)\}_{k=0}^n$. This is so because $B_{n+1}(z) = 1$ is orthogonal to all $B_k(z)$ (which all have relative degree one). Define as before $C_k(z) = A_n(z)B_k(z)$, $k = 1 \dots n$, and in addition $C_{n+1}(z) = A_n(z)$, and let:

$$C_{n+1}(z) = c_{n+1,1}z^n + c_{n+1,2}z^{n-2} + \dots + c_{n+1,n+1}.$$

Define

$$\bar{U}_n = \begin{bmatrix} c_{n+1,1} & 0 & \dots & 0 \\ c_{n+1,2} & c_{n,1} & \dots & c_{1,1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n+1,n+1} & c_{n,n} & \dots & c_{1,n} \end{bmatrix}. \quad (17)$$

Using the same arguments as for the case of R_n now gives

$$R_{n+1}^{-1} = \bar{U}_n \bar{U}_n^T.$$

This result can be viewed as a consequence of applying the Levinson algorithm to go from dimension n to $n+1$.

Finally, we note that the use of the Christoffel-Darboux formula leads to:

$$\begin{aligned} v^T(1/z_1) R_n^{-1} v(z_2) &= A_n(1/z_1) A_n(z_2) \Delta(z_1, z_2) \\ &= \frac{(z_2/z_1)^n A_n(z_1) A_n(1/z_2) - A_n(1/z_1) A_n(z_2)}{1 - z_1 z_2}. \end{aligned} \quad (18)$$

Using the DFT technique of this letter to recover R_n^{-1} from (18) will result in the Gohberg-Semencul formula. The connection between the Christoffel-Darboux and Gohberg-Semencul formulas is discussed in, e.g. [4].

IV. APPLICATION OF THE NEW FACTORIZATION FORMULA

Let $\{y(t)\}_{t=1}^n$ be a sequence of a stochastic signal whose power spectral density is $\Phi(e^{i\omega})$, $\omega \in [-\pi, \pi)$. Also let R_n denote a generic Toeplitz (covariance) matrix associated with an arbitrary m th order AR process with poles strictly inside the unit circle and with $m < n$ (m being a fixed integer that does not depend on n). Under the Gaussian assumption, the dominant term of the negative log-likelihood function that can be used to fit an m th order AR model to $\{y(t)\}_{t=1}^n$ is given by, see e.g., [1]:

$$Y^T R_n^{-1} Y, \quad Y = [y(1) \ y(2) \ \dots \ y(n)]^T. \quad (19)$$

Because (19) is a complicated function of the AR parameters, the minimization of (19) is often replaced by that of

$$\begin{aligned} \sum_{t=m+1}^n [\bar{A}_m(q) y(t-m)]^2, \\ \bar{A}_m(q) = q^m + a_1 q^{m-1} + \dots + a_m, \end{aligned} \quad (20)$$

where q is the forward shift operator. Notice that $\bar{A}(z) = \sigma_m A_m(z)$, due to the normalization of the coefficient of z^m . Minimizing (20) is a quadratic least-squares optimization problem in the coefficients of \bar{A}_m with a closed form

solution. In what follows, we make use of the new square-root factorization of R_n^{-1} introduced in this letter, see (15), to prove that²:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}\{Y^T R_n^{-1} Y\} = \frac{1}{\sigma_m^2} \mathbb{E}[\bar{A}_m(q)y(t)]^2, \quad (21)$$

which provides support to the fact that (20) is a good approximation of (19) asymptotically (as n increases). Using (15) and (16) yields:

$$\begin{aligned} \mathbb{E}\{Y^T R_n^{-1} Y\} &= \mathbb{E}\left\{\sum_{k=1}^n [C_k(q)y(t-n)]^2\right\} \\ &= \sum_{k=1}^n \frac{1}{2\pi} \int_{-\pi}^{\pi} |C_k(e^{i\omega})|^2 \Phi(e^{i\omega}) d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |A_n(e^{i\omega})|^2 \Phi(e^{i\omega}) \sum_{k=1}^n |B_k(e^{i\omega})|^2 d\omega, \quad (22) \end{aligned}$$

where, because we assume the process to be of order m , $A_n(z) = A_m(z)z^{n-m}$, and thus $A_n(z)$ has $n-m$ roots at $z=0$. This implies that $B_k = z^{-k}$, $k = m+1, \dots, n$ both for the Takenaka-Malmquist functions and for the Schur basis. Hence,

$$\sum_{k=1}^n |B_k(e^{i\omega})|^2 = \sum_{k=1}^m |B_k(e^{i\omega})|^2 + n - m, \quad (23)$$

where an expression for the first term on the right hand side is most easily obtained using the Takenaka-Malmquist functions,

$$|B_k(e^{i\omega})|^2 = \frac{1 - |\xi_k|^2}{|e^{i\omega} - \xi_k|^2}$$

Hence, we have

$$\begin{aligned} \frac{1}{n} \mathbb{E}\{Y^T R_n^{-1} Y\} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |A_m(e^{i\omega})|^2 \Phi(e^{i\omega}) \\ &\times \left(\frac{1}{n} \sum_{k=1}^m |B_k(e^{i\omega})|^2 + \frac{n-m}{n} \right) d\omega \\ &\rightarrow \frac{1}{\sigma_m^2} \mathbb{E}[\bar{A}_m(q)y(t)]^2, \quad n \rightarrow \infty, \quad (24) \end{aligned}$$

which proves (21). Note that the function (23) is actually a fundamental scale factor for the asymptotic variance of the corresponding autoregressive spectral density estimate as shown in [7]. The papers [8] and [9] contain further results in this direction. In fact, these results inspired our derivation of the square-root factorization (15). The Schur version of this formula was used in [10] to establish convergence rates for quadratic forms associated with inverse Toeplitz matrices as n tends to infinity.

Another possible application of the new square root factorization formula for R_n^{-1} , which is left for future research, is to establishing a novel relationship between the Capon and the AR method in the spectral analysis of stationary signals. See e.g. [11] for an introduction to this subject and some known relationships.

V. CONCLUSIONS

Toeplitz matrices and their inverses play a fundamental role both in the implementation and in the analysis of parametric estimation methods, and elsewhere. The Levinson algorithm and the Schur algorithm (as well as the Gohberg-Semencul formula) are classical results on how to find factorizations of such matrices. In this letter we have derived a class of square-root factorizations of inverse Toeplitz matrices, using rational ON functions. In particular, we have shown that the use of Takenaka-Malmquist functions in our framework leads to an apparently new square-root factorization formula. Interestingly, the magnitude of the roots of the autoregressive polynomial appear to play a role for the existence of this factorization, similar to that played by the magnitude of the reflection coefficients in the Levinson case.

REFERENCES

- [1] T. Söderström and P. Stoica, *System Identification*. Prentice Hall International, Hemel Hempstead, 1994.
- [2] M. H. Gutknecht and M. Hochbruck, "The stability of inversion formulas for Toeplitz matrices," *Linear Algebra and its Applications*, vol. 223-224, pp. 307–324, 1995, honoring Miroslav Fiedler and Vlastimil Ptak.
- [3] G. Heinig and K. Rost, *Algebraic Methods for Toeplitz-like Matrices and Operators*. Akademie-Verlag, Berlin and Birkhuser, Basel/Stuttgart, 1984.
- [4] T. Kailath and J. Chun, "Generalized Gohberg-Semencul formulas for matrix inversion," *Oper. Theory Adv. Appl.*, vol. 40, p. 231246, 1989.
- [5] P. Heuberger, B. Wahlberg, and P. Van den Hof, Eds., *Modelling and Identification with Rational Orthogonal Basis Functions*. Springer Verlag, London, 2005.
- [6] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad, *Orthogonal Rational Functions*. Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, UK, 1999.
- [7] L.-L. Xie and L. Ljung, "Variance expressions for spectra estimated using auto-regressions," *Journal of Econometrics*, vol. 118, pp. 247–256, 2004.
- [8] B. Ninness and H. Hjalmarsson, "Variance error quantifications that are exact for finite model order," *IEEE Transactions on Automatic Control*, vol. 49, no. 8, pp. 1275–1291, Aug. 2004.
- [9] H. Hjalmarsson and B. Ninness, "Least-squares estimation of a class of frequency functions," *Automatica*, vol. 42, no. 4, pp. 589–600, May 2006.
- [10] E. Hannan and B. Wahlberg, "Convergence rates for inverse Toeplitz matrix forms," *Journal of Multivariate Analysis*, vol. 31, pp. 127–135, 1989.
- [11] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Prentice Hall, New Jersey, 2005.

² \mathbb{E} denotes the expectation operator