Matrix spectral factorization for SA4 multiwavelet

Vasil Kolev1 · Todor Cooklev2 · Fritz Keinert3

Received: 6 December 2016 / Revised: 7 August 2017 / Accepted: 9 August 2017 / Published online: 22 August 2017 © Springer Science+Business Media, LLC 2017

Abstract In this paper, we investigate Bauer’s method for the matrix spectral factorization of an $r$-channel matrix product filter which is a half-band autocorrelation matrix. We regularize the resulting matrix spectral factors by an averaging approach and by multiplication by a unitary matrix. This leads to the approximate and exact orthogonal SA4 multiscaling functions. We also find the corresponding orthogonal multiwavelet functions, based on the QR decomposition.

Keywords SA4 orthogonal multiwavelet · Matrix spectral factorization · Bauer’s method · Cholesky factorization · Toeplitz matrix · QR factorization

1 Introduction

If the finite impulse response (FIR) system $H(z) = C_0 + C_1 z^{-1} + \cdots + C_n z^{-n}$ is a matrix polynomial which represents a causal multifilter bank, the product filter bank is given by $P(z) = H(z) H^*(z)$, where $H^*(z) = H^T(z^{-1})$. The coefficients $C_k$ are $r \times r$ matrices. Matrix Spectral Factorization (MSF) describes the problem of finding $H(z)$, given $P(z)$. MSF is still quite unknown in the signal processing community. Because of its numerous possible applications, in particular in multiple-input and multiple-output (MIMO) commu-
cations, image processing, multidimensional control theory, and others, it deserves to receive greater attention. We are interested in MSF as a tool for designing orthogonal multiwavelets.

It is known that MSF is possible as long as \( P(z) \) is positive semidefinite for \( z \) in the unit circle in the complex plane (Ephremidze et al. 2009; Hardin et al. 2004). We call \( P(z) \) **degenerate** if it is not strictly positive definite, that is, \( \det(P(z)) = 0 \) at one or more points.

A number of numerical approaches to MSF have been proposed, but they usually cannot handle the degenerate case. Unfortunately, those are exactly the cases of greatest interest for applications in the construction of multiwavelets.

We use Bauer’s method, as described by Youla and Kazanjian (1978), which is based on the Cholesky factorization of a block Toeplitz matrix. This algorithm can handle the degenerate case, but convergence is quite slow.

We study the SA4 multiwavelet as a test case, and as a benchmark for future applications. Starting from a known \( H(z) \), we compute \( P(z) \) and factor it again.

It is easy to see that MSF is not unique. If \( H_1(z) \) is any factor, then \( H_2(z) = H_1(z)U \) is also a factor for any orthogonal matrix \( U \), or more generally \( H_2(z) = H_1(z)U(z) \) for paraunitary \( U(z) \). For the factor produced by Bauer’s method, the matrix \( C_0 \) is always lower triangular, so we are forced to use further modifications to recover the original SA4.

In this paper, we only concentrate on modifications which speed up convergence and produce factors close to original SA4. This leads to two factorizations: the approximate SA4 multifilter, which is close to SA4, and the exact SA4 multifilter, which is equal to SA4 except for insignificant roundoff. In future papers, we plan to impose desirable conditions directly on the spectral factor, without working towards a known result.

The main contributions of this paper are

- Considering a novel method for computing the multichannel spectral factorization of degenerate matrix polynomial matrices;
- Speeding up the slow convergence of the algorithm;
- Minimizing the numerical errors which appear in the algorithm;
- Demonstrating by example that this approach can be used to construct an orthogonal symmetric multiwavelet filter;
- Comparing the frequency responses and experimental numerical performance of the approximate and exact SA4 multiwavelets.

The rest of the paper is organized as follows. In Sect. 2 we give some background material, state the problem, and describe related research. We also introduce matrix product filter theory, and present the product filter of the SA4 multiscaling function as a benchmark test case. In Sect. 3 we consider Bauer’s method, and discuss its numerical behavior. In Sect. 4 we describe the approach for obtaining the approximate and exact SA4 multiscaling and multiwavelet functions. The performance analysis of these multiwavelets is shown in Sect. 5. Section 6 gives conclusions.

## 2 Background and problem statement

We will use the following notation conventions, illustrated with the letter ‘a’:

- \( a \) lowercase letters refer to scalars or scalar-valued functions;
- \( \mathbf{a} \) lowercase bold letters are vectors or vector-valued functions;
- \( A \) uppercase letters are matrices or matrix-valued functions.

The symbols \( I \) and \( 0 \) denote the identity and zero matrices of appropriate size, respectively.
We will be using polynomials in a complex variable \( z \) on the unit circle, but all coefficients will be real-valued. Thus, the complex conjugate \( A^*(z) \) of a matrix polynomial \( A(z) = \sum_k A_k z^k \) is given by

\[
A^*(z) = A^T(z^{-1}) = \sum_k A_k^T z^{-k}.
\]

### 2.1 Matrix product filters

A two-channel multiple-input multiple-output (MIMO) filter bank of multiplicity \( r \) is shown in Fig. 1. Here \( x(z) \) is the input signal, a sequence of \( r \)-vectors, and the analysis multifilters are

\[
H(z) = C_0 + C_1 z^{-1} + \cdots + C_n z^{-n},
\quad G(z) = D_0 + D_1 z^{-1} + \cdots + D_n z^{-n},
\]

where \( C_k, D_k \) are matrices of size \( r \times r \). Likewise, \( E(z) \) and \( F(z) \) are the synthesis multifilters, and \( \hat{x}(z) \) is the output vector signal. For simplicity, we assume that all coefficients are real.

The input–output relation of this filter bank is

\[
\hat{x}(z) = \frac{1}{2} \left[ E(z) H(z) + F(z) G(z) \right] x(z) + \frac{1}{2} \left[ E(z) H(-z) + F(z) G(-z) \right] x(-z).
\]

It is convenient to write this equation in matrix form as

\[
\begin{bmatrix}
\hat{x}(z) \\
\hat{x}(-z)
\end{bmatrix} = \frac{1}{2} \begin{bmatrix}
E(z) & F(z) \\
E(-z) & F(-z)
\end{bmatrix} \begin{bmatrix}
H(z) & H(-z) \\
G(z) & G(-z)
\end{bmatrix} \begin{bmatrix}
x(z) \\
x(-z)
\end{bmatrix}.
\]

In general, the design of the multifilter bank requires four multifilters: two on the analysis and two on the synthesis side. In perfect-reconstruction orthogonal MIMO filter banks, the analysis modulation matrix

\[
M(z) = \begin{bmatrix}
H(z) & H(-z) \\
G(z) & G(-z)
\end{bmatrix}
\]

is paraunitary, that is

\[
M^*(z) M(z) = M(z) M^*(z) = I. \tag{1}
\]

In this case we can define the synthesis filters in terms of the analysis filters:

\[
E(z) = H^*(z), \quad F(z) = G^*(z).
\]

Equation (1) can also be expressed as

\[
H(z) H^*(z) + H(-z) H^*(-z) = I, \\
G(z) G^*(z) + G(-z) G^*(-z) = I, \\
H(z) G^*(z) + H(-z) G^*(-z) = 0,
\]

![Fig. 1 Two-channel multifilter bank](image-url)
or in terms of the coefficients
\[ \sum_k C_k C_{k+2\ell} = \sum_k D_k D_{k+2\ell} = \delta_{0\ell} I, \]
\[ \sum_k C_k D_{k+2\ell} = 0, \]
for any integer \( \ell \).

Given \( H(z) \), the Hermitian matrix polynomial
\[ P(z) = H(z)H^*(z) = \left( \sum_{k=0}^{n} C_k z^{-k} \right) \left( \sum_{k=0}^{n} C_k^T z^k \right) = \sum_{k=-n}^{n} P_k z^k, \]
is called the matrix lowpass product filter (or scalar product filter in the case \( r = 1 \)). The coefficients \( P_k \) satisfy \( P_{-k} = P_k^T \cdot P(z) \) is a half-band polynomial filter (Cooklev et al. 1996), that is, it satisfies the equation
\[ P(z) + P(-z) = 2I. \tag{2} \]
This implies that \( P_0 = I \), and \( P_{2\ell} = 0 \) for all nonzero integers \( \ell \).

2.2 Multiwavelet theory

2.2.1 Multiscaling and multiwavelet functions

Consider the iteration of a MIMO filter bank along the channel containing \( H(z) \). After \( k \) iterations, the equivalent filters will be
\[ H^{(k)}(z) = \prod_{j=0}^{k} H \left( z^{2^j} \right) = H \left( z^{2^k} \right) H \left( z^{2^{k-1}} \right) \cdots H(z), \]
\[ G^{(k)}(z) = G \left( z^{2^k} \right) \prod_{j=0}^{k-1} H \left( z^{2^j} \right) = G \left( z^{2^k} \right) H \left( z^{2^{k-1}} \right) \cdots H(z), \]
or in the time domain
\[ C_j^{(k)} = \sum_j C_j^{(k-1)} C_{j-2^{k-1}}, \]
\[ D_j^{(k)} = \sum_j D_j^{(k-1)} D_{j-2^{k-1}}, \]
where \( C_j^{(0)} = C_j, D_j^{(0)} = D_j \).

The filterbank coefficients are associated with function vectors \( \varphi = [\phi_0, \phi_1, \ldots, \phi_{r-1}]^T \), called the multiscaling function, and \( \psi = [\psi_0, \psi_1, \ldots, \psi_{r-1}]^T \), called the multiwavelet function. These functions satisfy the recursion equations
\[ \varphi(t) = \sqrt{2} \sum_{k=0}^{n} C_k \varphi(2t - k), \]
\[ \psi(t) = \sqrt{2} \sum_{k=0}^{n} D_k \varphi(2t - k). \]
The support of $\varphi, \psi$ is contained in the interval $[0, n]$, but it could be strictly smaller than this interval. See Massopust et al. (1996) for more details.

2.2.2 Multiwavelet transform

The input signal $x$, at level $j = 0$, is a sequence of $r$-vectors in $\ell_2$. Thus, in the 2-channel case,

$$x = \{x_k\}_{k=-\infty}^{\infty}, \quad x_k = \begin{bmatrix} c_0^0[k] \\ c_1^0[k] \end{bmatrix}, \quad \text{with } \sum_k \|x_k\|^2 < \infty.$$

In the case of a scalar-valued signal, it is necessary to vectorize the input first. A pre-filtering step must be inserted at the beginning of the analysis filter bank, and a corresponding post-filter must be incorporated at the end.

The multiwavelet transform of $x$ consists in computing recursively, for levels $j = 1, \ldots, J$, the approximation signal

$$\begin{bmatrix} e_0^j[\ell] \\ e_1^j[\ell] \end{bmatrix} = \sum_k C_{k-2\ell} \begin{bmatrix} c_0^{j-1}[k] \\ c_1^{j-1}[k] \end{bmatrix},$$

and the detail signal

$$\begin{bmatrix} d_0^j[\ell] \\ d_1^j[\ell] \end{bmatrix} = \sum_k D_{k-2\ell} \begin{bmatrix} c_0^{j-1}[k] \\ c_1^{j-1}[k] \end{bmatrix}.$$

In the orthogonal case, the inverse can be computed iteratively in a straightforward way:

$$\begin{bmatrix} c_0^j[k] \\ c_1^j[k] \end{bmatrix} = \sum_\ell C_{k-2\ell}^T \begin{bmatrix} e_0^{j-1}[\ell] \\ e_1^{j-1}[\ell] \end{bmatrix} + \sum_\ell D_{k-2\ell}^T \begin{bmatrix} d_0^{j-1}[\ell] \\ d_1^{j-1}[\ell] \end{bmatrix}.$$

The filters coefficients $C_k, D_k$ completely characterize the multiwavelet filter bank.

2.2.3 Multiwavelet properties

Pre- and postfiltering Since multifilters are MIMO, they operate on several streams of input data rather than one. Therefore, input data needs to be vectorized. There are many prefilters available for various applications, but three kinds are preferred:

- Oversampling—Repeated Row Preprocessing
- Critical Sampling—Matrix Preprocessing
- Embedded Orthogonal Symmetric Prefilter Bank

In Strela and Walden (2000), the first and second approach is presented for the GHM and CL multiwavelets. In practical applications, a popular choice is based on the Haar transform (Cotronei et al. 1998; Tan et al. 1999). Haar pre- and postfilters have the advantage of simultaneously possessing symmetry and orthogonality, and no multiplication is needed if one ignores the scaling factor. The problem of constructing symmetric prefilters is considered in detail in Hsung et al. (2003), with proposed solutions for the DGHM and CL multifilters.

The third technique is applied in Gan and Ma (2005) and Hsung et al. (2007). They find a three-tap prefilter for the DGHM multifilter by searching among the parameters which minimize the first-order absolute moment of the filter coefficients.

Similarly, at the filter bank output, a postfilter is needed to reconstruct the signal. Pre- and postprocessing is not needed for scalar wavelets.
Balancing order An orthogonal multiscaling function is said to be \textit{balanced of order }\( q \) \textit{if} the signals \( u_k = [\ldots, (-2)^k, (-1)^k, 0^k, 1^k, 2^k, \ldots]^T \) are preserved by the operator \( L^T \) for \( k = 0, 1, \ldots, q - 1 \). That is, \( L^T u_k = 2^{-k} u_k \),

where

\[
L = \begin{bmatrix}
\cdots & C_0 & C_1 & \cdots & C_n \\
C_0 & C_1 & \cdots & C_n & \cdots
\end{bmatrix}.
\]

Multiwavelets that do not satisfy this property are said to be \textit{unbalanced}. See Lebrun and Vetterli (1998), Lebrun and Vetterli (2001), Li and Peng (2011) and Plonka and Strela (1998) for more details. Balanced multiwavelets do not require preprocessing.

Approximation order In the scalar case, a certain approximation order refers to the ability of the low-pass filter to reproduce discrete-time polynomials, while the wavelet annihilates discrete-time polynomials up to the same degree. Since many real-world signals can be modeled as polynomials, this property typically leads to higher coding gain (CG).

In the case of non-balanced multiwavelets, appropriate preprocessing is required to take advantage of approximation orders.

The approximation order of multiscaling and multiwavelet functions can be determined using the following result established in Wu et al. (2010). A multiscaling function provides approximation order \( p \) iff there exist vectors \( u_k \in \mathbb{R}^r \), \( 0 \leq k < p \), \( u_0 \neq 0 \), which satisfy

\[
\sum_{\ell=0}^{k} \frac{1}{(2i)^\ell} \binom{k}{\ell} u_{k-\ell}^T D^\ell H(0) = 2^{-k} u_k^T,
\]

\[
\sum_{\ell=0}^{k} \frac{1}{(2i)^\ell} \binom{k}{\ell} u_{k-\ell}^T D^\ell H(\pi) = 0,
\]

where \( D^\ell \) is the derivative of order \( \ell \), and the \textit{masks} of multiscaling and multiwavelet functions are

\[
H(\omega) = \frac{1}{\sqrt{2}} \sum_{k=0}^{n} C_k e^{ik\omega}, \quad G(\omega) = \frac{1}{\sqrt{2}} \sum_{k=0}^{n} D_k e^{ik\omega}.
\]

To fully characterize the multifilter bank with respect to approximation order we can add the condition

\[
\sum_{\ell=0}^{k} \frac{1}{(2i)^\ell} \binom{k}{\ell} u_{k-\ell}^T D^\ell G(0) = 0.
\]

\textit{Good Multifilter Properties (GMPs)} A multiwavelet system can be represented by an equivalent scalar filter bank system (Tham et al. 2000). The \( r \) equivalent scalar filters are, in fact, the \( r \) polyphases of the corresponding multifilter. This relationship motivates a new measure called \textit{good multifilter properties} (GMPs) (Tham et al. 1998, 2000, (Def. 1)), which characterizes the magnitude response of the equivalent scalar filter bank associated with a multifilter. GMPs provide a set of design criteria imposed on the scalar filters, which can be translated directly to eigenvector properties of the designed multiwavelet filters.
An orthogonal multiwavelet system has GMPs of order at least \((1, 1, 1)\) if the following conditions are satisfied (Tham et al. 2000)

\[
\begin{align*}
H(0)e(0) &= e(0) \\
H(r\pi)e(\pi) &= 0 \\
G(0)e(0) &= 0
\end{align*}
\]

where \(H(\omega)\) and \(G(\omega)\) are masks of the lowpass and highpass filters (see Eq. (3)), and

\[
e(\omega) = [1, e^{-i\omega}, e^{-2i\omega}, \ldots, e^{-(r-1)i\omega}]^T.
\]

A class of symmetric–antisymmetric biorthogonal multiwavelet filters which possess GMPs was introduced in Tan et al. (1999). A GMP order of at least \((1, 1, 1)\) is critical for ensuring no frequency leakage across bands, hence improving compression performance. Multifilters possessing GMPs have better performance than those which do not possess GMPs (Li and Yang 2010), due to better frequency responses for energy compaction, greater regularity and greater approximation order of the corresponding wavelet/scaling functions. GMPs help prevent both DC and high-frequency leakage across bands; this contributes to reduced smearing, blocking, ringing artifacts, and also helps to prevent checkerboard artifacts in reconstructed images for image coding.

The SA4 multiwavelet, introduced below in Sect. 3.3, is a member of a one-parameter family of orthogonal multiwavelets with four coefficient matrices (Tham et al. 2000). Different members of the family have different approximation order, but all of them have a GMP order of at least \((1, 1, 1)\). This manifests itself in the smooth decay to zero of the magnitude responses near \(\omega = \pi\), compared to the following 4-tap orthogonal multiwavelet filters:

- The GHM multiwavelet (Donovan et al. 1996) has symmetric orthogonal scaling functions and an approximation order of 2 for its filter length 4.
- Chui and Lian’s CL multiwavelet (Chui and Lian 1996) has the highest possible approximation order of 3 for its filter length 3.
- Jiang’s multiwavelet JOPT4 (Jiang 1998b) has optimal time–frequency localization for its filter length.

Although the GHM and CL systems are the most commonly used orthogonal multiwavelet systems, and have higher approximation order than the SA4 system, they do not satisfy GMPs. Smoothness The smoothness of \(\varphi, \psi\) can be characterized by Sobolev regularity \(S\); this is discussed in Cohen et al. (1997), Micchelli and Sauer (1997) and Jiang (1998a). The regularity estimate is related to both the eigenvalues and the corresponding eigenvectors of the transition operator and to spectral radius of the transition operator. Symmetry and antisymmetry A function \(g(t)\) is symmetric about a point \(c\) if \(g(c-t) = g(c+t)\) for all \(t\). It is antisymmetric if \(g(c-t) = -g(c+t)\).

For scaling functions and wavelets of compact support, the only possible symmetry is about the center of their support. For multiscaling and multiwavelet functions, this could be a different point for different components. In the scalar case, the scaling function cannot be antisymmetric, since it must have a nonzero integral. In the multiwavelet case, some components of \(\varphi\) can have antisymmetry.

### 2.3 Problem statement

The design of perfect reconstruction multifilter banks and multiwavelets remains a significant problem. In the scalar case, spectral factorization of a half-band filter that is positive definite
on the unit circle was, in fact, the first design technique, suggested by Smith and Barnwell (1986). This provides the motivation to extend the technique to the multiwavelet case. Matrix Spectral Factorization (MSF) is more challenging than in the scalar case, and has not been investigated previously.

The present paper considers the application of MSF to the problem of finding an orthogonal lowpass multfilter $H(z)$, given a product filter $P(z)$. Since every orthogonal multiscaling filter $H(z)$ is a spectral factor of some matrix product filter, this can be used as a tool for designing new filters.

With an eye towards future applications, we are interested in constructing multifilters with desirable properties, especially GMPs. This implies that the determinant of $P(z)$ will have at least one zero of higher multiplicity on the complex unit circle. That is, $P(z)$ will be degenerate.

As a test case, we use a product filter $P(z)$ which satisfies the half-band condition (2), is derived from an $H(z)$ with more than 2 taps (i.e. $n \geq 3$), and has good regularity properties and GMPs. Our benchmark test case will be the 2-channel orthogonal SA4 multiwavelet.

While there are many matrix spectral factorization algorithms (Cooklev et al. 1996; Ježek and Kučera 1985; Lawton 1993), most cannot handle the degenerate case. We use Bauer’s method, based on the Toeplitz method of spectral factorization of the Youla and Kazanjian algorithm (Bauer 1956; Youla and Kazanjian 1978). This method can handle the degenerate case; however, convergence becomes very slow.

It is easy to see that matrix spectral factorization is not unique. If $H_1(z)$ is a factor of a given $P(z)$, then $H_2(z) = H_1(z)U$ is also a factor for any orthogonal matrix $U$, or more generally $H_2(z) = H_1(z)U(z)$ for paraunitary $U(z)$. Other operations may be permissible in some cases.

In our experiments with Bauer’s algorithm, the initial factors do not correspond to smooth functions. We apply regularization techniques, both to speed up convergence and to work towards recovering the original filter in this test case. In this manner, we derive two filters which we call the approximate and exact SA4 multiwavelets.

2.4 Previous work

MSF is especially central to problems in polynomial optimal estimation and control. For control theory problems defined by state-variable models, MSF is equivalent to solving a Riccati equation.

Polynomial matrix factorization for a biorthogonal multiwavelet with the help of an orthogonal projector is presented in Aliev et al. (1992). The orthogonal projector with respect to the unit circle is the solution of the algebraic Riccati equation, which allows us to use the Newton–Raphson scheme.

The multidimensional spectral factorization in Avelli and Trentelman (2008) considers the following important problems: Finding the polynomial spectral factors, formulating the spectral factorization problem in term of linear matrix inequalities, and finding how the rational spectral factors connect it to a linear eigenvalue problem.

Based on analogue control theory, a new approach using negative feedback for solving the scalar spectral factorization problem is applied in Moir (2011). It is unique in that it uses convolution of a symmetric causal and non-causal impulse response as part of the feedback path of a closed-loop servo, which until now has been untried. The algorithm should have many applications using polynomial matrices.

Rational matrix factorization theory is connected with infinite impulse response (IIR) systems and their more general case of spectral factorization. A finite impulse response
(FIR) system has an impulse response with compact support, while the impulse response of an IIR system continues indefinitely. There are a number of important differences between FIR and IIR.

In control and some communication applications, mostly IIR systems are considered. They can meet design requirements with lower order than would be required for an FIR system, and they are computationally more efficient (Rosen and Howell 2011). However, they have the potential for limit cycle behavior when idle, due to the feedback system in conjunction with quantization. IIR filters are also difficult to control and may be unstable.

On the other hand, FIR systems can be easier to design, and stability is guaranteed. They can also easily be made to be linear phase.

Baggio’s thesis (2014) provides a solution in discrete time to the multivariate rational spectral factorization method devised by Youla in his celebrated paper (Youla 1961). The thesis also discusses a natural extension of the spectral factorization problem, the $J$-spectral factorization approach for non-singular polynomial matrices, and the resulting spectral factors. This plays a crucial role in control and estimation theory.

In Bose (2017), a 2D rational spectral factorization is obtained via the Hilbert transform of a 2D power spectrum density.

Paper (Grinshpan et al. 2016) considers a method for construction of normalized polynomial matrices with determinantal representations. It uses the root method of rational matrix spectral factorization of a positive semidefinite 1D and 2D FIR and IIR matrix-valued polynomial, either on the circle (trigonometric polynomial) or on the real line (algebraic polynomial). The method employs a polynomial factorization technique for an appropriately formed stable Bezoutian matrix in one variable. The obtained spectral factors lead to minimal realization of the transfer function in steady space representation. Note that real-zero polynomials and Hermitian determinantal representations lead to easy representation of product filter of the 1D and $n$D-channel digital filter banks.

A steady space representation based on the use of the transfer function was systematically discussed for the first time in the early book (Bart et al. 1979) about control theory of stochastic systems (a MIMO system that will produce different outputs for a given inputs). This plays a very important role in applied mathematics and engineering applications as filtering and prediction in control theory (Willems 1971; Youla 1961). A wonderful explanation of the spectral factorization for nonsingular positive rational matrices is given in Vandewalle and Dewilde (1975). The algorithm produces a minimal rational spectral factor that is decomposed into elementary factors.

The general spectral factorization problem for an arbitrary polynomial rational matrix, such as rational spectral density functions in transfer functions of control theory, is considered in Baggio and Ferrante (2016a, b) and Ferrante and Picci (2017). They obtain stochastically minimal spectral factors corresponding to solutions of minimal complexity (i.e. minimal McMillan degree) of the spectral factorization problem. It is shown with a constructive proof in Baggio and Ferrante (2016b) that an arbitrary rational matrix function that is positive semidefinite on the unit circle admits a spectral factorization where the poles and zeroes of the rational spectral factors lie in prescribed regions. Two types of spectral factors are considered: causal spectral factors, where all the poles lie inside the unit circle, and outer or minimal phase spectral factors, where the zeroes lie inside the unit circle. It is also proved that the outer spectral factor is essentially unique.

Such a factorization corresponds to an optimal control and network synthesis problem. For instance, a stochastically minimal spectral factor for a rank deficient polynomial matrix obtained by the discrete-time spectral factorization is a solution for anti-causal realization having all its zeroes in the (closed) unit disk (Baggio and Ferrante 2016b).
The stochastically minimal (outer) spectral factors of a given rational spectral density are estimated in Baggio and Ferrante (2016a), without requiring any additional system-theoretic assumptions. As noted there, the key idea is the essential uniqueness of the spectral factors. In the scalar case, this result is straightforward.

The approach in Abdou et al. (2015) computes the moving average parameters from an estimated power spectral density, and uses that to estimate the outer factor. In the general matrix case, however, a rational function can feature a pole and a zero at the same point, so that the result appears to be quite difficult to derive. This was left open as a conjecture in Baggio and Ferrante (2016b). The proof (Baggio and Ferrante 2016a) makes use of a very elegant parameterization of rational all-pass functions.

A general characterization and complete factorization theory for state-space representations is given in Ferrante and Picci (2017). The left and right spectral factors of an arbitrary square all-pass rational function are classified. A classical result in systems and control theory (Willems 1971; Youla 1961) states that rational spectral factorization can be cast in terms of linear matrix inequalities. In Ferrante and Picci (2017), this is applied to some homogeneous Riccati equations.

Chapter 14 in Davis (2002) discusses recursive algorithms, in both continuous time and discrete time, for the solution of the algebraic Riccati equation. The algorithms are based on quasilinearization and Newton’s method. Convergence is proved. Theorem 8 describes in detail an algorithm for numerical spectral factorization, for continuous and discrete time. The discrete time version is based on optimal gain iteration, and is computationally attractive for spectral factorization of a multivariable spectrum.

At present, rational spectral factorization usually uses polynomial matrices with non-zero matrix coefficients. An essential characteristic of our problem, different from applications in control theory, is that the matrix product filter always consists of pairs of zero-matrix coefficients. For instance, in the two-channel case with decimation factor 2, all matrices with even subscripts are zero.

When poles are zeroes, then the MSF of a matrix product filter can be considered as a special case of the general discrete-time spectral factorization problem. In Zorzi and Sepulchre (2016), an identification procedure for autoregressive (AR) Gaussian stationary stochastic processes is derived. The author exploits the MSF of the maximal-entropy covariance extension and the sparse plus low-rank (S+L) decomposition of the manifest concentration matrix, to propose an efficient formulation of the identification problem.

Spectral factorization is used in Zorzi (2015), where a new multivariate divergence family between spectral densities in the context of optimal prediction is considered. It is interesting to note that this family is not even known in the scalar case. MSF leads to the spectral factor of the normalized innovation process, right multiplied by an all-pass function (Oppenheim and Schafer 2009).

MSF admits finding the unique solution of the dual problem in Zorzi (2014). Of course, a sensible and efficient method based on spectral factorization techniques may be employed. Future research can potentially address Bauer’s method for similar dual problem as in Zorzi (2014).

MSF also plays a crucial role in the solution of various applied problems for MIMO systems in communications and control engineering (Wang et al. 2015), construction of different 2D filters from a minimum-phase 2D spectral factor (Bose 2017), different methods for obtaining a product filter from a desirable prototype FIR filter (Kalathil and Elias 2015), and IIR prototype filters based on the fast Fourier transform (FFT) (Yin et al. 2015) for the design of filter banks.
MSF has been applied to designing minimum phase FIR filters and the associated all-pass filter (Hansen et al. 2010), quadrature-mirror filter banks (Charoenlarpnopparut 2007), MIMO systems for optimum transmission and reception filter matrices for precoding and equalization (Fischer 2005), precoders (Du et al. 2013), and many other applications.

The analysis of linear systems corresponding to a given spectral density function was first established by Wiener, who used linear prediction theory of multi-dimensional stochastic processes (Wiener and Masani 1957). The method of Wilson for MSF was developed for applications in the prediction theory of multivariate discrete time series, with known (or estimated) spectral density (Wilson 1972). Many numerical approaches to MSF have been proposed, for example Cooklev et al. (1996), Ježek and Kučera (1985) and Lawton (1993). A survey of such algorithms is given in Sayed and Kailath (2001); however, this does not include algorithms for the degenerate case.

It is well known that all MSF methods have difficulties in this situation, and some of them cannot handle zeroes on the unit circle at all. For example, Kučera’s algorithm, an otherwise popular matrix spectral factorization algorithm, has this limitation, and is not suitable for our purpose.

Bauer’s method was successful applied in Roux (1986) to the Radon projection. In that paper, factorization of the autocorrelation matrix of the Radon projection of a minimum phase pseudo-polynomial restored the coefficients of the original pseudo-polynomial with an accuracy around $10^{-17}$ after 10,000 recursions on a 64-bit floating point processor.

### 3 Matrix spectral factorization

The following theorem answers the existence question for MSF.

**Theorem 1** (Matrix form of the Riesz-Fejér lemma (Ephremidze et al. 2009; Hardin et al. 2004)) A matrix polynomial

$$P(z) = \sum_{k=-n}^{n} P_k z^k$$

can be factored as $P(z) = H(z)H^*(z)$, where

$$H(z) = \sum_{k=0}^{n} C_k z^{-k}$$

is an r-channel causal polynomial matrix, if and only if $P(z)$ is symmetric positive semidefinite for $z$ on the complex unit circle.

#### 3.1 Bauer’s method

If $H(z)$ and $P(z)$ are known, we can construct doubly infinite block Toeplitz matrices $L$ and $T$ from their coefficients, by setting $L_{ij} = C_{i-j}$, $T_{ij} = P_{j-i}$. $T$ is symmetric and block banded with bandwidth $n$; $L$ is block lower triangular, also with bandwidth $n$.

The relation $P(z) = H(z)H^*(z)$ corresponds to $T = LL^T$, that is, $L$ is a Cholesky factor of $T$. 
In Bauer’s method, we pick a large enough integer \( f \), and truncate \( T \) to (block) size \( f \times f \):

\[
T^{(f)} = \begin{pmatrix}
P_0 & P_1 & \cdots & P_n \\
P_{-1} & P_0 & \cdots & P_n \\
\vdots & \vdots & \ddots & \vdots \\
P_{-n} & \cdots & \ddots & P_n \\
P_{-n} & \cdots & \ddots & P_0
\end{pmatrix}
\]

If \( T^{(f)} \) is positive definite, we can compute the Cholesky factorization \( T = L^{(f)} L^{(f)T} \), where

\[
L^{(f)} = \begin{pmatrix}
C_0^{(1)} & C_1^{(2)} & \cdots & C_n^{(n+1)} \\
C_{(n+1)}^{(1)} & \ddots & \cdots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
C_{n+1}^{(f)} & \cdots & C_1^{(f)} & C_0^{(f)}
\end{pmatrix}
\]

It is shown in Youla and Kazanjian (1978) that this factorization is possible, and that \( C_k^{(f)} \to C_k \) as \( f \to \infty \).

Bauer’s method works even in the case of highly degenerate \( P(z) \). Unfortunately, convergence in this case is very slow.

### 3.2 Numerical behavior

For chosen size \( f \), the matrix \( T^{(f)} \) is of size \( 2f \times 2f \). Its singular values are all in the range \([0, 2]\). The first \((f-1)\) singular values are close to 2, then \( \sigma_f = \sigma_{f+1} = 1 \), and the remaining \((f-1)\) singular values are close to 0. See Fig. 2.

Fig. 2 A sharp drop of the singular values of the Toeplitz matrix \( T \) and of its spectral factor \( L^{SF} \), for \( f = 20 \); (filled circles) singular values of \( T \); (hollow circles) singular values of \( L^{SF} \).
The numerical stability of the factorization algorithm is good, until the higher singular values get too close to 0.

A bigger problem is the slow convergence. For example, in the scalar case $p(z) = z^{-1} + 2 + z$, where the factor is $h(z) = 1 + z^{-1}$, the value on the diagonal in row $f$ is

$$\sqrt{1 + \frac{1}{f}} \approx 1 + \frac{1}{2f},$$

which converges only very slowly to the limit 1. This simple problem has a double zero of the determinant on the unit circle. For higher order zeroes convergence speed becomes even worse.

### 3.3 Testing Bauer’s method

A flowchart for testing Bauer’s method for MSF is shown in Fig. 3. The steps are

- **Step 1**: Choice of the benchmark multiscaling function $H(z)$;
- **Step 2**: Construction of the matrix lowpass product filter $P(z)$;
- **Step 3**: Find a spectral factor $L_{SF}^{(f)} = L^{(f)}$, and assess its quality;
- **Step 4**: Do different kinds of postprocessing to obtain spectral factors $H^{(i)}$, $i = 1, 2, 3$;
- **Step 5**: Compare the obtained factors with the benchmark $H(z)$.

Step 1 We consider the well-known orthogonal SA4 multiwavelet (Chui and Lian 1996; Huo and Miao 2012; Tham et al. 2000) as a benchmark testing case.

The recursion coefficients $C_k$ and $D_k$ of SA4 depend on a parameter $t$. They are

$$C_0 = \alpha \begin{pmatrix} 1 & t \\ 1 & -t \end{pmatrix}; \quad C_1 = \alpha \begin{pmatrix} t^2 & t \\ -t^2 & t \end{pmatrix}; \quad C_2 = \alpha \begin{pmatrix} t^2 & -t \\ -t^2 & t \end{pmatrix}; \quad C_3 = \alpha \begin{pmatrix} 1 & -t \\ -1 & t \end{pmatrix};$$

$$D_0 = \alpha \begin{pmatrix} t & -1 \\ t & 1 \end{pmatrix}; \quad D_1 = \alpha \begin{pmatrix} -t & t^2 \\ t & -t^2 \end{pmatrix}; \quad D_2 = \alpha \begin{pmatrix} -t & -t^2 \\ t & t^2 \end{pmatrix}; \quad D_0 = \alpha \begin{pmatrix} t & 1 \\ -t & -1 \end{pmatrix};$$

where $\alpha = 1/(\sqrt{2}(1 + t^2))$. In this paper, we use the filter with $t = 4 + \sqrt{15}$, which leads to $\alpha = (4 - \sqrt{15})/(8\sqrt{2})$. Numerical values of $C_k$, $D_k$ are listed in Table 1.

These coefficients generate the 2-band compactly supported orthogonal multiscaling function $\phi = [\phi_0, \phi_1]^T$ and multiwavelet function $\psi = [\psi_0, \psi_1]^T$, all with support $[0, 3]$. Graphs of these functions can be found in Fig. 8 in a later chapter.

---

**Fig. 3** Steps in testing matrix spectral factorization using Bauer’s method
Table 1  Coefficients of the original SA4 multiscaling and multiwavelet functions

| n  | $C_n$                              | $D_n$                              |
|----|------------------------------------|------------------------------------|
| 0  | 0.011226792152545                  | 0.088388347648318 −0.088388347648318 0.011226792152545 |
| 1  | 0.011226792152545                  | −0.088388347648318 −0.088388347648318 −0.011226792152545 |
| 2  | 0.695879989034003                  | 0.088388347648318 0.088388347648318 −0.695879989034003 |
| 3  | 0.011226792152545                  | −0.088388347648318 0.088388347648318 −0.011226792152545 |

Step 2 The symbol of the product lowpass multifilter $P$ has the form

$$P(z) = P_3^T z^{-3} + P_1^T z^{-1} + P_0 + P_1 z + P_3 z^3,$$

where

$$P_0 = I,$$

$$P_1 = P_{-1}^T = \frac{1}{64} \begin{pmatrix} 4\sqrt{15} + 17 & 4\sqrt{15} + 16 \\ -4\sqrt{15} - 16 & -4\sqrt{15} - 17 \end{pmatrix},$$

$$P_2 = 0,$$

$$P_3 = P_{-3}^T = \frac{1}{64} \begin{pmatrix} 15 - 4\sqrt{15} & 4\sqrt{15} - 16 \\ 16 - 4\sqrt{15} & 4\sqrt{15} - 15 \end{pmatrix}.$$

Numerical values of $P_k$ are given in Table 2.

The impulse responses of the product filter (see Fig. 4) have the character of nearly Haar-type scaling and wavelet functions. The frequency responses have good selectivity.

Step 3 The matrix $T^{(f)}$ in this case looks like this:

$$T^{(f)} = \begin{pmatrix} P_0 & P_1^T & 0 & P_3^T \\ P_1 & P_0 & P_1^T & 0 \\ 0 & \ldots & \ldots & \ldots \\ P_3 & \ldots & \ldots & P_3^T \\ \ldots & \ldots & \ldots & \ldots \\ P_3 & 0 & P_1 & P_0 \end{pmatrix}$$

For any fixed $f$, the coefficients from the last row of $L^{(f)}$ are approximate factors of $P(z)$. Suppressing the dependence of $L^{SF}$ on $f$ in the notation, we define the filter

$$L^{SF}(z) = \sum_{k=0}^{n} C_k^{(f)} z^{-k}.$$
Table 2 The matrix coefficients $P_k$ of the half-band product filter $P$

| $k$ | $P_k$ |
|-----|-------|
| 0   | 1     | 0   |
| 1   | 0.507686459137964 | 0.492061459137964 |
| 3   | 0.007938540862036 | 0.007686459137964 |

Fig. 4 The impulse response (left) and frequency response (right) of the two components of the product filter $P = [P_0; P_1]^T$

and the matrix

$$L^{SF} = \begin{bmatrix} C_{0}(f) & C_{1}(f) & C_{2}(f) & C_{3}(f) \\ L_{00} & L_{01} & L_{02} & L_{03} & L_{04} & L_{05} & L_{06} & L_{07} \\ L_{10} & L_{11} & L_{12} & L_{13} & L_{14} & L_{15} & L_{16} & L_{17} \end{bmatrix}.$$  \hspace{1cm} (5)

This notation will be used again later.

To measure the quality of the factorization, we compute

$$\Delta P(z) = P(z) - L^{SF}(z)L^{SF*}(z) = \sum_{k=-n}^{n} \Delta P_k z^k.$$  

The residual is then

$$\Delta P = \max_{k=-n, \ldots, n} \max_{ij} \left| (\Delta P_k)_{ij} \right|.$$  

In numerical experiments, the minimal error $\Delta P$ is achieved for $f = 81$; the coefficients $C_{k}^{(81)}$ are given in Table 3. The numerical error $\Delta P^{(81)} = 3.169 \times 10^{-9}$ for Bauer’s method is much better than that of Wilson’s method for MSF, which cannot be lower than $10^{-5}$ Ephremidze et al. 2016. The impulse response of the obtained spectral factor is non-regular, but the frequency response has good selectivity. This approximate factor does not quite provide perfect reconstruction. This is addressed in more detail in Sect. 5.2.

Steps 4 and 5 will be covered in Sect. 4.
### Table 3

Coefficients of the lowpass spectral factor $L^{(f)}(z)$ for $f = 81$, with $\Delta p^{(81)} = 3.169 \times 10^{-9}$

| $k$ | $C_k^{(81)}$          | $D_k$          |
|-----|----------------------|----------------|
| 0   | 0.094428373297668    | 0.022310801334572 |
| 1   | -0.091754813647953   | 0.0679731045265413 |
| 2   | 0.175943193428843    | -0.702056243347588 |
| 3   | -0.165030422671601   | 0.021002135513698 |

#### 4 Construction of approximate and exact SA4 multiwavelets

Above, we showed how to obtain the approximate spectral factor $L^{SF}$ for a given size $f$. Given any factor of $P$, other possible factors can be found by postmultiplication with suitable orthogonal matrices, or in some cases by other manipulations.

In Sect. 4.1 we use a simple averaging process and column reversal to produce a lowpass filter that is similar to, but not identical to, the original SA4 multiscaling function. We call this the approximate SA4 multiwavelet.

In Sect. 4.2, we add a rotation postfactor and an averaging process to produce another multiwavelet which is even closer to the original SA4 multiwavelet. We call this the exact SA4 multiwavelet.

The newly derived multiscaling filters are denoted by

$$H^{(1)} = \begin{bmatrix} C_0^{(1)} & C_1^{(1)} & C_2^{(1)} & C_3^{(1)} \end{bmatrix}$$

for the approximate SA4 multiwavelet, and

$$H^{(3)} = \begin{bmatrix} C_0^{(3)} & C_1^{(3)} & C_2^{(3)} & C_3^{(3)} \end{bmatrix}$$

for the exact SA4 multiwavelet. ($H^{(2)}$ is an intermediate step in the calculation of $H^{(3)}$).

In each case, we find the corresponding multiwavelet filter by using a QR decomposition (Golub and Loan 2013). The multiscaling filter is factorized as

$$Q \begin{bmatrix} C_0 & C_1 & C_2 & C_3 \end{bmatrix} = R \begin{bmatrix} | & | & | & | \end{bmatrix} \begin{bmatrix} I \end{bmatrix}$$

The coefficients $D_k$ are then obtained from the third and fourth columns of $Q$.

We define the absolute error by

$$\Delta H^{(k)} = H^{(k)} - H^{(k)} = \begin{bmatrix} \Delta C_0^{(k)} & \Delta C_1^{(k)} & \Delta C_2^{(k)} & \Delta C_3^{(k)} \end{bmatrix},$$

where $\Delta C_n^{(k)} = C_n - C_n^{(k)}$ for $k = 1, 2, 3$, and likewise for $\Delta C^{(k)}$.

To measure the deviation of the new filters from the original, we introduce the mean square error (MSE) of the multiscaling and multiwavelet functions,

$$\text{MSE-MF}^{(k)} = \sum_{ij} |\Delta H_{ij}^{(k)}|^2, \quad \text{MSE-MwF}^{(k)} = \sum_{ij} |\Delta G_{ij}^{(k)}|^2.$$
and the maximal absolute errors (MAE) of the matrix coefficients, the multiscaling function and the multiwavelet function as

\[
\begin{align*}
\text{MAE-MC}_{(k)}^\ell &= \max_{ij} \left| \left\lfloor \Delta C^{(k)}_{\ell} \right\rfloor_{ij} \right| , \\
\text{MAE-MF} &= \max_{ij} \left| \Delta H^{(k)}_{ij} \right| = \max_{ij} \left| H_{ij} - H^{(k)}_{ij} \right| , \\
\text{MAE-MwF} &= \max_{ij} \left| \Delta G^{(k)}_{ij} \right| = \max_{ij} \left| G_{ij} - G^{(k)}_{ij} \right| .
\end{align*}
\]

### 4.1 Approximate SA4 multiwavelet

The spectral factor \( L^{SF}_{LF} \) obtained from Bauer’s method leads to non-symmetrical and non-regular functions. A simple algorithm can be used to symmetrize and regularize the scaling functions. Below, we are using the notation from Eq. (5).

First, we average the absolute values of the first and fourth, as well as the second and third, matrix coefficients of the spectral factor:

\[
\begin{align*}
L^{SF}_{0,6} &= \frac{1}{4} \left( \left| L^{SF}_{00} \right| + \left| L^{SF}_{10} \right| + \left| L^{SF}_{06} \right| + \left| L^{SF}_{16} \right| \right) , \\
L^{SF}_{1,7} &= \frac{1}{4} \left( \left| L^{SF}_{01} \right| + \left| L^{SF}_{11} \right| + \left| L^{SF}_{07} \right| + \left| L^{SF}_{17} \right| \right) , \\
L^{SF}_{2,4} &= \frac{1}{4} \left( \left| L^{SF}_{02} \right| + \left| L^{SF}_{12} \right| + \left| L^{SF}_{04} \right| + \left| L^{SF}_{14} \right| \right) , \\
L^{SF}_{3,5} &= \frac{1}{4} \left( \left| L^{SF}_{05} \right| + \left| L^{SF}_{15} \right| + \left| L^{SF}_{05} \right| + \left| L^{SF}_{15} \right| \right) .
\end{align*}
\]

and construct average matrices

\[
\begin{bmatrix}
L^{SF}_{2,4} & L^{SF}_{3,5} \\
L^{SF}_{2,4} & L^{SF}_{3,5}
\end{bmatrix}
\quad \quad \quad 
\begin{bmatrix}
L^{SF}_{0,6} & L^{SF}_{1,7} \\
L^{SF}_{0,6} & L^{SF}_{1,7}
\end{bmatrix}.
\]

Second, the matrix coefficients \( C_1^{(1)} \) and \( C_3^{(1)} \) are obtained by multiplying the averaged matrices from the right by \( \mathcal{J} = \text{diag}(1, -1) \), but keeping the signs, while the matrix coefficients \( C_0^{(1)} \) and \( C_2^{(1)} \) are obtained by multiplying \( C_1^{(1)} \) and \( C_3^{(1)} \) on the left and right by \( U = \text{antidiag}(1, 1) \):

\[
\begin{align*}
C_3^{(1)} &= \text{sign} \left( L^{SF}_{3} \right) \begin{bmatrix}
L^{SF}_{0,6} & L^{SF}_{1,7} \\
L^{SF}_{0,6} & L^{SF}_{1,7}
\end{bmatrix} \cdot \mathcal{J} , \\
C_0^{(1)} &= U \cdot C_3^{(1)} \cdot U , \\
C_1^{(1)} &= \text{sign} \left( L^{SF}_{1} \right) \begin{bmatrix}
L^{SF}_{2,4} & L^{SF}_{3,5} \\
L^{SF}_{2,4} & L^{SF}_{3,5}
\end{bmatrix} \cdot \mathcal{J} , \\
C_2^{(1)} &= U \cdot C_1^{(1)} \cdot U .
\end{align*}
\]

This produces coefficients quite close to the original coefficients in (4).

In order to determine the best approximate solution, we investigate the influence of the size \( f \) of the Toeplitz matrix \( T^{(f)} \) for the numerical errors MAE-MF\(^{(1)}\), MSE-MC\(^{(1)}\), and MSE-MC\(^{(1)}\). These errors are shown in Fig. 5 in logarithmic scale, for \( f \) up to 100. Note that MSE-MC\(^{(1)}\) = MSE-MC\(^{(1)}\), and MSE-MC\(^{(1)}\) = MSE-MC\(^{(1)}\), by construction.
Fig. 5 Dependence of maximum absolute error MAE-MF(1) and mean squared errors
MSE-MC(1) = MSE-MC(1),
MSE-MC(1) = MSE-MC(1) on matrix size f (log. scale)

Table 4 The minimum MAEs and MSEs for the approximate and exact SA4 multiscaling and multiwavelet functions, the size f and the corresponding angle θ

| f   | MAE-MF(1) | MAE-MwF(1) | MSE-MF(1) | MSE-MwF(1) |
|-----|-----------|------------|-----------|------------|
| 21  |           |            | 4.2797 × 10⁻³ | 4.6139 × 10⁻⁶ |
| 15,400 | MAE-MF(2) | MAE-MwF(2) | 1.220 × 10⁻⁴ | 7.376 × 10⁻⁹ |
| 12,042 | MAE-MF(3) | MAE-MwF(3) | 8.285 × 10⁻⁸ | 3.835 × 10⁻¹⁵ |

The important minimal errors MAE-MF(1) and MSE-MwF(1) for f = 21 are tabulated in Table 4; the matrix coefficients of are listed in Table 5. Obviously, the development of the errors shows the convergence of the algorithm, as well as the dependency of the errors on the size f. It also shows that the main error comes from the matrix coefficients C₀¹ and C₁², rather than C₀¹ and C₂¹.

The error MSE-MF(1) behaves essentially the same way as MAE-MC(1). Unfortunately, the convergence of MAE-MC(1) and MAE-MC(1) stalls at a relatively large number. Thus, the approximate SA4 multiwavelet is not very accurate. Better matrix coefficients are achieved in the next subsection.

4.2 Exact SA4 multiwavelet

The second algorithm leads to more regular scaling functions. We multiply the coefficients of L₅F from the right by the unitary matrix U(θ)
The errors MSE-MF of the very slow convergence; despite right multiplication with the unitary matrix, desirable 3 \theta \text{ in the angle there is one overshoot, after that it decreases very slowly. The minimum errors are obtained at} f \text{ in Fig. 7. Again, we are using the notation from Eq. (5).}

Table 5 Coefficients of the approximate SA4 multiscaling and multiwavelet functions for \( f = 21 \)

| n  | \( C_n^{(3)} \)            | \( D_n^{(3)} \) |
|----|---------------------------|-----------------|
| 0  | 0.011165766264837         | -0.08852225597447 | 0.011165766264837 |
| 0  | -0.08852225597447         | -0.08852225597447 | -0.011165766264837 |
| 1  | 0.088718768987217         | 0.088718768987217 | -0.691600252880066 |
| 1  | -0.691600252880066        | -0.691600252880066 | -0.691600252880066 |
| 2  | 0.088718768987217         | 0.088718768987217 | -0.691600252880066 |
| 2  | -0.691600252880066        | -0.691600252880066 | -0.691600252880066 |
| 3  | -0.08852225597447         | -0.08852225597447 | -0.011165766264837 |
| 3  | 0.08852225597447          | 0.08852225597447  | -0.011165766264837 |

\[
C_k^{(2)} = C_k^{SF} \cdot U(\theta) = C_k^{SF} \cdot \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad k = 0, \ldots, 3.
\]

This produces the approximate factor \( H^{(2)} \).

The angle \( \theta \) (in radians) is calculated as the average value

\[ \theta = \frac{1}{2} \left( \cos^{-1}(\text{even}) + \sin^{-1}(\text{odd}) \right), \]

where

\[ \text{even} = \frac{1}{4} \left( |L_{00} + L_{04} + L_{10} + L_{14}| + |L_{02} + L_{06} + L_{12} + L_{16}| \right) \]

and

\[ \text{odd} = \frac{1}{4} \left( |L_{01} + L_{05} + L_{11} + L_{15}| + |L_{03} + L_{07} + L_{13} + L_{17}| \right). \]

Again, we are using the notation from Eq. (5).

Figure 6 shows the dependence of the angle \( \theta \) on the size \( f \) (log. scale). It can be seen that in the angle there is one overshoot, after that it decreases very slowly. The minimum errors are obtained at \( f = 15,400 \), with the values shown in Table 4. This shows the influence of the very slow convergence; despite right multiplication with the unitary matrix, desirable precision cannot be an achieved. The errors MSE-MF\(^{(2)}\) and MAE-MF\(^{(2)}\) for \( H^{(2)} \) are shown in Fig. 7.

To increase the numerical precision and obtain the exact SA4 multiscaling function, we again apply an averaging approach.

\[
H^{(3)} = \left[ C^{(3)}_0, C^{(3)}_1, C^{(3)}_2, C^{(3)}_3 \right],
\]

\[
\begin{align*}
(C^{(3)}_0)_{ij} &= \text{sign}\left((C^{(2)}_0)_{ij}\right) \cdot \frac{1}{2} \left( \left| (C^{(2)}_0)_{ij} \right| + \left| (C^{(2)}_3)_{ij} \right| \right), \\
(C^{(3)}_1)_{ij} &= \text{sign}\left((C^{(2)}_1)_{ij}\right) \cdot \frac{1}{2} \left( \left| (C^{(2)}_1)_{ij} \right| + \left| (C^{(2)}_2)_{ij} \right| \right), \\
(C^{(3)}_2)_{ij} &= \text{sign}\left((C^{(2)}_2)_{ij}\right) \cdot \frac{1}{2} \left( \left| (C^{(2)}_0)_{ij} \right| + \left| (C^{(2)}_3)_{ij} \right| \right), \\
(C^{(3)}_3)_{ij} &= \text{sign}\left((C^{(2)}_3)_{ij}\right) \cdot \frac{1}{2} \left( \left| (C^{(2)}_0)_{ij} \right| + \left| (C^{(2)}_3)_{ij} \right| \right).
\end{align*}
\]
The errors for $H^{(3)}$ are shown in Fig. 7, with minima shown in Table 4. Let us consider the accuracy of the multiscaling functions obtained at the leading local minima of the error $\text{MAE-MF}^{(3)}$. There are 3 minima, as can be seen in the zoomed part of Fig. 7. The first local minimum is at the value $f = 4269$, the second at $f = 9139$, the third at $f = 12,042$. This means that increasing the size of the Toeplitz matrix leads to slow improvement in the precision of the exact multiscaling function. Nevertheless, due to applying the averaging method, the minimal errors in $H^{(3)}$ and $G^{(3)}$ for $f = 12,042$ are smaller by 2 or 3 orders of magnitude than the errors in $H^{(2)}$ and $G^{(2)}$, and smaller by 5 orders of magnitude than the errors in $H^{(1)}$ and $G^{(1)}$.

We choose $H^{(3)}$ for $f = 12,042$ for the coefficients of the exact SA4 multiwavelet, given in Table 6. They are almost the same as the original coefficients in (4).
Table 6  Coefficients of the exact SA4 multiscaling and multiwavelet functions for \( f = 12.042 \)

| \( n \) | \( C_n^{(3)} \) | \( D_n^{(3)} \) |
|------|--------|--------|
| 0    | 0.01122679201336641 | -0.08838843285624247 |
|      | -0.08838843013543155 | -0.08838843029442495 |
| 1    | 0.6958799459541121  | 0.08838843089794699 |
|      | -0.6958799676294172 | -0.6958799673174058 |
| 2    | 0.6958799459541121  | 0.08838843089794699 |
|      | -0.6958799676294172 | -0.6958799673174058 |
| 3    | -0.01122679228175923 | -0.08838843285624247 |

5 Performance analysis

5.1 Comparison with other multiwavelets

Table 7 gives the results of our experiments with the approximate and exact SA4 multiwavelets, compared with other well-known orthogonal and biorthogonal filter banks. Comparisons include coding gain (CG), Sobolev smoothness (S) (Cotronei et al. 1998), symmetry/antisymmetry, and length. All of the multiwavelets considered have symmetry/antisymmetry.

The CG for orthogonal transforms is a good indication of the performance in signal processing. It is the ratio of arithmetic and geometric means of channel variances \( \sigma_i^2 \):\

\[
CG = \frac{1}{r} \sum_{i=1}^{r} \sigma_i^2 / \left( \prod_{i=1}^{r} \sigma_i^2 \right)^{1/r}
\]

Coding gain is one of the most important factors to be considered in many applications. It is always greater than or equal to 1; greater values are better. CG is equal to 1 if all the variances are equal, which means that it is not possible to clearly distinguish between the smooth and the detailed components of the multiwavelet transformation coefficients. To estimate CG, the variance is computed using a first order Markov model AR(1) with intersample autocorrelation coefficient \( \rho = 0.95 \) (Downie and Silverman 1998).

The Sobolev exponent \( S \) of a filter bank measures the \( L^2 \)-differentiability of the corresponding multiscaling function \( \varphi = [\phi_0, \phi_1]^T \), and thus also the multiwavelet function \( \psi = [\psi_0, \psi_1]^T \). It is completely determined by the multiscaling symbol \( H(z) \).

The obtained Sobolev regularity and CGs of the approximate and exact SA4 multiwavelets are equal: \( S^{(1)} = S^{(3)} = 0.9919 \), and \( CG^{(1)} = CG^{(3)} = 3.7323 \) dB. This means that for some applications we can use the SA4 multiwavelet.

According to Table 7, the approximate and exact SA4 multiwavelets are better than most commonly used filter banks. They also allow an economical lifting scheme for future implementation.

Figure 8 shows a comparison of the approximate and exact SA4 multiwavelets in time domain and impulse response. The influence of the lower precision of the approximate SA4 multiwavelet can be seen in the time domain image.

The magnitude of the approximate SA4 in the time domain is observably smaller than for the exact SA4, while the frequency responses are essentially identical (see Fig. 8). Therefore,
Table 7  Comparison of coding gain (CG), Sobolev regularity (S), symmetry/antisymmetry (S/A), and length, for various multiwavelets

| Multifilter                                               | CG    | S   | S/A | Length |
|-----------------------------------------------------------|-------|-----|-----|--------|
| Biorthogonal Hermitian cubic spline (bih34n, Strela 1998) | 1.51  | 2.5 | Yes | 3/5    |
| Biorthogonal (dual to Hermitian cubic spline) (bih32s, Averbuch et al. 2007) | 1.01  | 2.5 | Yes | 3/5    |
| Integer Haar (Cheung and Po 2001)                        | 1.83  | 0.5 | Yes | 2      |
| Chui–Lian (CL, Chui and Lian 1996)                       | 2.06  | 1.06| Yes | 3      |
| Biorthogonal (dual to Hermitian cubic spline) (bih32s, Averbuch et al. 2007) | 2.42  | 0.61| Yes | 5/3    |
| Biorthogonal (from GHM) (bihm2, Averbuch et al. 2007)    | 2.43  | 0.5 | Yes | 2/6    |
| Biorthogonal (from GHM, dual to bigham2) (bigham6, Averbuch et al. 2007) | 3.53  | 2.5 | Yes | 6/2    |
| Biorthogonal (dual to Hermitian cubic spline) (bih52s, Turcajová 1998) | 3.69  | 0.83| Yes | 5/3    |
| Approximate SA4                                          | 3.73  | 0.99| Yes | 4      |
| Exact SA4 (Shen et al. 2000)                             | 3.73  | 0.99| Yes | 4      |
| Geronimo–Hardin–Massopust (GHM, Geronimo et al. 1994)    | 4.41  | 1.5 | Yes | 4      |

in applications where the frequency response is important, we can use the approximate multiscaling function.

5.2 Influence of inaccuracy of filter coefficients with approximate SA4 multiwavelet

The approximate SA4 multiwavelet \( H^{(1)} \) is not quite an orthogonal perfect reconstruction filter. In this section, we explore whether it can still be useful in applications.
5.2.1 1D signal (no additional processing)

The influence of the inaccuracy of the matrix filter coefficients in the case of the approxi-
mate SA4 multiwavelet is shown by 1D applications with no additional processing. By “no
additional processing” we mean that we simply decompose and reconstruct a signal through
several levels. No compression, denoising, or other processing is done.

We use Haar balancing pre- and postfilters (Cotronei et al. 1998; Strela et al. 1999)
\[ Q = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \]
whose small length preserves the time localization of the multiwavelet decomposition, sim-
plicity, orthogonality, and symmetry.

Results for the balanced version are shown in Fig. 9b.

The error in orthogonality of the approximate balanced SA4 multiwavelet is
\[ \Delta H = I - Q \cdot H^{(1)} \cdot (H^{(1)})^T Q^T = 0.0117 I. \]

For the quality measure, we decompose and reconstruct five normed test signals of length
\( N = 2^7 \), \( s \) (without noise) and \( \hat{s} = s + \epsilon \) (with noise). The number of decomposition levels
is 6.

The noise components \( \epsilon_i \) are independent identically distributed random variables with
mean 0 and standard deviation \( \sigma \). We use maximum absolute errors (MAE)
\[ \Delta s = \max_i |\text{reconstructed } s_i - \text{original } s_i|, \]
where \( s_i \) refers to the \( i \)th test signal.

The test signals are ‘Cusp’, ‘HiSine’, ‘LoSine’, ‘Piece-regular’, and ‘Piece-polynomial’,
implemented in the Matlab environment. See Fig. 9a.

For the test signal ‘Cusp’, increasing the number of decomposition and reconstruction
levels \( j \) leads to a linear increase of MAE. From the obtained small differences (about
\( 3.5 \times 10^{-4} \)) in error between noisy and noiseless signal, it follows that the presence of noise
does not make much difference in signals of this type.

For the test signal ‘HiSine’ we observe nearly constant MAEs for the noisy and noiseless
signals with increasing level \( j \); the difference between the noisy and noiseless signal is in
the interval \( [2 - 5.3] \times 10^{-3} \). Again, the presence of noise has only a minimal influence.

For the test signal ‘LoSine’, both noisy and noiseless MAEs show a linear increase up to
the 3rd level, with only a small increase at the 4th level and 5th level of \( -5.7 \times 10^{-3} \) and
\( 3.6 \times 10^{-3} \), as shown in Fig. 9b. Therefore, after the third level the influence of noise is weak.

For the test signal ‘Piece-regular’, the MAE grows quadratically with increasing \( j \). The
difference between noisy and noiseless signals is smallest at the 1st level (\( 6 \times 10^{-5} \)) and largest
at the 3rd level (\( 1.2 \times 10^{-3} \)). For the test signal ‘Piece-polynomial’, the MAE grows non-
uniformly and non-linearly with increasing \( j \). The difference between noisy and noiseless
signals is smallest at the 1st level (\( 3.8 \times 10^{-4} \)) and largest at the 4th level (\( 2 \times 10^{-3} \)). Thus,
for both (‘Piece-polynomial’ and ‘Piece-regular’) the influence of noise is minimal, when
using the approximate SA4 multiwavelet.

5.2.2 2D signal (no additional processing)

The performance of the approximate SA4 multiwavelet was tested by decomposing and
reconstructing several gray level images, through 6 or 7 levels. Figure 10 shows the details
for four of the images, with the quality measure $PSNR = \log_{10}(255^2/MSE)$ dB. The images used are ’Lena’, ’Peppers’, ’Girlface’, and ’Barbara’.

The approximate balanced SA4 multiwavelet applied to the four images leads to an exponential decrease of the PSNRs. According to these results, it is preferable to use no more than 3 or 4 levels of decomposition. After that, the reconstructed image has very low PSNRs and visibly worse quality. In some applications, such as big data archives, higher levels of decomposition may be useful.

### 5.3 Image denoising

In this section, we compare the balanced and non-balanced version of the exact SA4 multiwavelet with the GHM multiwavelet (Downie and Silverman 1998) and the Chui–Lian
Fig. 10 Test images of size (256 × 256 pixels) a lena, b peppers, and (512 × 512 pixels) c house, d girlface, e barbara, and f zelda

Fig. 11 The PSNRs of four images for 2D decomposition and reconstruction without additional processing of 1 through 6 or 7 levels with the approximate SA4 multiwavelet

multiwavelet CL (Downie and Silverman 1998), by considering image denoising with vector hard thresholding (Donoho and Johnstone 1994), using 2–6 levels. The images are ‘Lena’, ‘Zelda’, and ‘House’, of size 512 × 512 pixels, with white additive Gaussian noise with variance σ = 10. See Fig. 10.

The multiwavelet coefficients of the white noise are reduced at each level, but uniformly distributed within each level. Therefore, the best approach to denoising is to find an appropriate threshold value at each level.

The exact SA4 multiwavelet, both balanced and non-balanced, achieved the maximal PSNRs. The PSNR differences for the both versions of the test images are 3.57 dB at the
Fig. 12 Comparative analysis of PSNRs for image denoising with the exact SA4 multiwavelet and vector hard threshold through 2–6 levels of the images ‘House’, ‘Zelda’, and ‘Lena’ with 512 × 512 pixels and white additive Gaussian noise with variance $\sigma = 10$. Left non-balanced multiwavelets; right balanced multiwavelets

1st level and 4.06 dB at the 6th level for “Lena”; 4.17 dB (1st level) and 5.13 dB (6th level) for “Zelda”, 5.09 dB (1st level) and 6.53 dB (6th level) for “House”. Although balancing of multiwavelets destroys the symmetry, it leads to increasing PSNRs and better image denoising with the exact SA4 multiwavelet for the three test images for the both version multiwavelets, while for the non-balanced GHM and CL multiwavelets PSNRs decrease (see Fig. 12a). According to PSNRs, image decomposition and reconstruction through 3 levels with the approximate SA4 multiwavelet is comparable to image denoising with the non-balanced SA4 multiwavelet through 4 levels, while 2 levels are comparable with the balanced multiwavelet of 1 level (see Figs. 11 and 12). Therefore, in image denoising it is preferable to use the exact SA4 multiwavelet.

Obviously, the PSNRs of the exact balanced SA4 multiwavelet are better than the well-known orthogonal multiwavelets GHM and CL. They are useful in many applications.

6 Conclusion

In this paper, we consider for the first time the problem of obtaining an orthogonal multiscaling function by matrix spectral factorization from a degenerate polynomial matrix. We show benchmark testing of MSF, and apply Bauer’s method to factoring the product filter of the SA4 multiwavelet. In addition, we show how to remove numerical errors and improve the properties of the factors obtained from Cholesky factorization, leading to fast convergent algorithms.

A very important part is obtaining the key angle $\theta$ in explicit form. Based on the proposed averaging approach, we develop two filter banks, the approximate and the exact SA4 orthogonal multiwavelets.

Experimental results have shown that the performances of the resulting multiwavelets are better than those of the Chui–Lian multiwavelet and biorthogonal multifilters, and are highly comparable to that of longer multiwavelets. Theoretical analyses for the influence of the size $f$ of the Toeplitz matrix are considered, as well as simple 1D and 2D applications.

After comparing both types of multifilters, we concluded that the proposed averaging approach is a better way to remove numerical errors and find the exact SA4 multiwavelet filter bank. It is important to note that the performance of the balanced exact SA4 multiwavelet
for image denoising is better than the well-known orthogonal multiwavelets GHM and CL, which are longer.

A future research direction is the construction of other known or new orthogonal multiwavelet filters, based on MSF of their product filters. Finally, another future research direction is to apply Bauer’s method for spectral factorization, for instance by studying the theory of all-pass functions (Baggio and Ferrante 2016a, b; Ferrante and Picci 2017), AR Gaussian stationary stochastic processes (Zorzi and Sepulchre 2016), multivariate spectral densities (Zorzi 2014, 2015), and so on.

Acknowledgements The authors would like to thank the three anonymous referees for their critical review and helpful suggestions that allowed improving the exposition of the manuscript.

References

Abdou, A., Turcu, F., Grivel, E., Diversi, R., & Ferré, G. (2015). Identifying an autoregressive process disturbed by a moving-average noise using inner–outer factorization. Signal, Image and Video Processing, 9(1), 235–244.

Aliiev, F., Bordyug, B. A., & Larin, V. (1992). Discrete generalized algebraic Riccati equations and polynomial matrix factorization. Systems & Control Letters, 18(1), 49–59.

Avelli, D. N., & Trentelman, H. (2008). Algorithms for multidimensional spectral factorization and sum of squares. In Special issue devoted to selected papers presented at the 13th conference of the international linear algebra society, vol. 429, no. 5, pp. 1114–1134.

Averbuch, A. Z., Zheludev, V. A., & Cohen, T. (2007). Multiwavelet frames in signal space originated from Hermite splines. IEEE Transactions on Signal Processing, 55(3), 797–808.

Baggio, G. (2014). Spectral factorization of rational matrix valued functions. Master’s thesis, University of Padova.

Baggio, G., & Ferrante, A. (2016). On minimal spectral factors with zeroes and poles lying on prescribed regions. IEEE Transactions on Automatic Control, 61(8), 2251–2255.

Baggio, G., & Ferrante, A. (2016). On the factorization of rational discrete-time spectral densities. IEEE Transactions on Automatic Control, 61(4), 969–981.

Bart, H., Gohberg, I., & Kaashoek, M. (1979). Minimal factorization of matrix and operator functions. Operator theory: Advances and applications. Basel: Birkhäuser.

Bauer, F.L. (1956). Beiträge zur Entwicklung numerischer Verfahren für programmgesteuerte Rechenanlagen. II. Direkte Faktorisierung eines Polynoms. Bayer. Akad. Wiss. Math.-Nat. Kl. S.-B. 1956, 163–203 (1957)

Bose, N. K. (2017). Applied multidimensional systems theory (2nd ed.). Basel, Switzerland: Springer International Publishing.

Charoenlarpnopparut, C. (2007). One-dimensional and multidimensional spectral factorization using Gröbner basis approach. In 2007 Asia–Pacific conference on communications (pp. 201–204).

Cheung, K. W., & Po, L. M. (2001). Integer multiwavelet transform for lossless image coding. In Proceedings of 2001 international symposium on intelligent multimedia, video and speech processing, 2001 (pp. 117–120).

Chui, C. K., & Lian, J. A. (1996). A study of orthonormal multi-wavelets (Selected keynote papers presented at 14th IMACS World Congress, Atlanta, 1994). Applied Numerical Mathematics, 20(3), 273–298.

Cohen, A., Daubechies, I., & Plonka, G. (1997). Regularity of refinable function vectors. Journal of Fourier Analysis and Applications, 3(3), 295–324.

Cooklev, T., Nishihara, A., Kato, M., Sablathash, M. (1996). Two-channel multifilter banks and multiwavelets. In IEEE international conference on acoustics, speech, and signal processing conference proceedings ICASSP-96 (vol. 5, pp. 2769–2772).

Cotronei, M., Montefusco, L. B., & Puccio, L. (1998). Multiwavelet analysis and signal processing. IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, 45(8), 970–987.

Davis, J. H. (2002). Foundations of deterministic and stochastic control. Basel: Birkhäuser.

Donoho, D. L., & Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. Biometrika, 81(3), 425–455.

Donovan, G. C., Geronimo, J. S., Hardin, D. P., & Massopust, P. R. (1996). Construction of orthogonal wavelets using fractal interpolation functions. The SIAM Journal on Mathematical Analysis, 27(4), 1158–1192.
Downie, T. R., & Silverman, B. W. (1998). The discrete multiple wavelet transform and thresholding methods. *IEEE Transactions on Signal Processing*, 46(9), 2558–2561.

Du, B., Xu, X., & Dai, X. (2013). Minimum-phase FIR precoder design for multicasting over MIMO frequency-selective channels. *Journal of Electronics (China)*, 30(4), 319–327.

Ephremidze, L., Janashia, G., & Lagvilava, E. (2009). A simple proof of the matrix-valued Fejér-Riesz theorem. *Journal of Fourier Analysis and Applications*, 15(1), 124–127.

Ephremidze, L., Saied, F., & Spitkovsky, I. (2016). On algorithmization of Janashia–Lagvilava matrix spectral factorization method. *IEEE Transactions on Information Theory*. arXiv:1606.04909 (submitted).

Ferrante, A., & Picci, G. (2017). Representation and factorization of discrete-time rational all-pass functions. *IEEE Transactions on Automatic Control*, 62(7), 3262–3276.

Fischer, R. F. (2005). Sorted spectral factorization of matrix polynomials in MIMO communications. *IEEE Transactions on Communications*, 53(6), 945–951.

Gan, L., & Ma, K. K. (2005). On minimal lattice factorizations of symmetric–antisymmetric multifilterbanks. *IEEE Transactions on Signal Processing*, 53(2, part 1), 606–621.

Geronimo, J. S., Hardin, D. P., & Massopust, P. R. (1994). Fractal functions and wavelet expansions based on several scaling functions. *The Journal of Approximation Theory*, 78(3), 373–401.

Golub, G. H., & Van Loan, C. F. (2013). *Matrix computations*. Johns Hopkins studies in the mathematical sciences (4th ed.). Baltimore: Johns Hopkins University Press.

Grinshpan, A., Kalužný-Verbovetskyi, D. S., Vinnikov, V., & Woerdeman, H. J. (2016). Stable and real-zero polynomials in two variables. *Multidimensional Systems and Signal Processing*, 27(1), 1–26.

Hansen, M., Christensen, L. P. B., & Winther, O. (2010). Computing the minimum-phase filter using the QL-factorization. *IEEE Transactions on Signal Processing*, 58(6), 3195–3205.

Hardin, D. P., Hogan, T. A., & Sun, Q. (2004). The matrix-valued Riesz lemma and local orthonormal bases in shift-invariant spaces. *Advances in Computational Mathematics*, 20(4), 367–384.

Hsung, T. C., Lun, D. P. K., Shum, Y. H., & Ho, K. C. (2007). Generalized discrete multiwavelet transform with embedded orthogonal symmetric prefilter bank. *IEEE Transactions on Signal Processing*, 55(6), 5619–5629.

Hsung, T. C., Sun, M. C., Lun, D. K., & Siu, W. C. (2003). Symmetric prefilters for multiwavelets. *IEE Proceedings-Vision, Image and Signal Processing*, 150(1), 59–68.

Hu, G., & Miao, L. (2012). Cycle-slip detection of GPS carrier phase with methodology of SA4 multi-wavelet transform. *Chinese Journal of Aeronautics*, 25(2), 227–235.

Ježek, J., & Kučera, V. (1985). Efficient algorithm for matrix spectral factorization. *Automatica*, 21(6), 663–669.

Jiang, Q. (1998). On the regularity of matrix refinable functions. *The SIAM Journal on Mathematical Analysis*, 29(5), 1157–1176.

Jiang, Q. (1998). Orthogonal multiwavelets with optimum time–frequency resolution. *IEEE Transactions on Signal Processing*, 46(4), 830–844.

Kalathil, S., & Elias, E. (2015). Prototype filter design approaches for near perfect reconstruction cosine modulated filter banks—A review. *Journal of Signal Processing Systems*, 81(2), 183–195.

Lawton, W. (1993). Applications of complex valued wavelet transforms to subband decomposition. *IEEE Transactions on Signal Processing*, 41(12), 3566–3568.

Lebrun, J., & Vetterli, M. (1998). Balanced multiwavelets theory and design. *IEEE Transactions on Signal Processing*, 46(4), 1119–1125.

Lebrun, J., & Vetterli, M. (2001). High-order balanced multiwavelets: Theory, factorization, and design. *IEEE Transactions on Signal Processing*, 49(9), 1918–1930.

Li, B., & Peng, L. (2011). Balanced multiwavelets with interpolatory property. *IEEE Transactions on Signal Processing*, 20(5), 1450–1457.

Li, Y. F., & Yang, S. Z. (2010). Construction of paraunitary symmetric matrices and parametrization of symmetric orthogonal multiwavelet filter banks. *Acta Mathematica Sinica (China Series)*, 53(2), 279–290.

Massopust, P. R., Ruch, D. K., & Van Fleet, P. J. (1996). On the support properties of scaling vectors. *Applied and Computational Harmonic Analysis*, 3(3), 229–238.

Micchelli, C. A., & Sauer, T. (1997). Regularity of multiwavelets. *Advances in Computational Mathematics*, 7(4), 455–545.

Moir, T. J. (2011). A control theoretical approach to the polynomial spectral-factorization problem. *Circuits, Systems, and Signal Processing*, 30(5), 987–998.

Oppenheim, A. V., & Schafer, R. W. (2009). *Discrete-time signal processing* (3rd ed.). Upper Saddle River: Prentice Hall Press.

Plonka, G., & Strela, V. (1998). Construction of multiscaling functions with approximation and symmetry. *The SIAM Journal on Mathematical Analysis*, 29(2), 481–510.
Rosen, S., & Howell, P. (2011). Signals and systems for speech and hearing. Research in the sociology of organizations (2nd ed.). Bingley, UK: Emerald Group Publishing Ltd.
Roux, J. L. (1986). 2D spectral factorization and stability test for 2D matrix polynomials based on the radon projection. In IEEE international conference on acoustics, speech, and signal processing, ICASSP ’86 (vol. 11, pp. 1041–1044).
Sayed, A. H., & Kailath, T. (2001). A survey of spectral factorization methods. Numerical linear algebra techniques for control and signal processing. Numerical Linear Algebra with Applications, 8(6–7), 467–496.
Shen, L., Tan, H. H., & Tham, J. Y. (2000). Symmetric–antisymmetric orthonormal multiwavelets and related scalar wavelets. Applied and Computational Harmonic Analysis, 8(3), 258–279.
Smith, M., & Barnwell, T. (1986). Exact reconstruction techniques for tree-structured subband coders. IEEE Transactions on Acoustics, Speech, and Signal Processing, 34(3), 434–441.
Strela, V. (1998). A note on construction of biorthogonal multi-scaling functions. In Wavelets, multiwavelets, and their applications (San Diego, CA, 1997), Contemp. Math. (vol. 216, pp. 149–157). Providence: American Mathematical Society.
Strela, V., Heller, P. N., Strang, G., Topiwala, P., & Heil, C. (1999). The application of multiwavelet filterbanks to image processing. IEEE Transactions on Image Processing, 8(4), 548–563.
Tham, J. Y., Shen, L. X., & Tan, H. H. (1999). Good multifilter properties: A new tool for understanding multiwavelets. In Proceedings of international conference on imaging, science, systems and technology CISST-98, Las Vegas, USA (pp. 52–59).
Tham, J. Y., Shen, L., Lee, S. L., & Tan, H. H. (2000). A general approach for analysis and application of discrete multiwavelet transforms. IEEE Transactions on Signal Processing, 48(2), 612–618.
Turcajová, R. (1998). Hermite spline multiwavelets for image modeling. In Proceedings of SPIE 3391, wavelet applications V Orlando (vol. 46, pp. 46–56).
Vandewalle, J., & Dewilde, P. (1975). On the minimal spectral factorization of nonsingular positive rational matrices. IEEE Transactions on Information Theory, 21(6), 621–634.
Wang, Z., McWhirter, J. G., Weiss, S. (2015). Multichannel spectral factorization algorithm using polynomial matrix eigenvalue decomposition. In 2015 49th Asilomar conference on signals, systems and computers (pp. 1714–1718).
Wiener, N., & Masani, P. (1957). The prediction theory of multivariate stochastic processes. I. The regularity condition. Acta Mathematica, 98, 111–150.
Wilkins, J. (1971). Least squares stationary optimal control and the algebraic riccati equation. IEEE Transactions on Automatic Control, 16(6), 621–634.
Wilson, G. T. (1972). The factorization of matricial spectral densities. SIAM Journal on Applied Mathematics, 23, 420–426.
Wu, G., Li, D., Xiao, H., & Liu, Z. (2010). The M-band cardinal orthogonal scaling function. Applied Mathematics and Computation, 215(9), 3271–3279.
Yin, S. S., Zhou, Y., & Chan, S. C. (2015). An efficient method for designing of modulated filter banks with causal-stable IIR filters. Journal of Signal Processing Systems, 78(2), 187–197.
Youla, D. (1961). On the factorization of rational matrices. IRE Transactions on Information Theory, 7(3), 172–189.
Youla, D. C., & Kazanjian, N. N. (1978). Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle. IEEE Transactions on Circuits and Systems, CAS–25(2), 57–69.
Zorzi, M. (2014). A new family of high-resolution multivariate spectral estimators. IEEE Transactions on Automatic Control, 59(4), 892–904.
Zorzi, M. (2015). Multivariate spectral estimation based on the concept of optimal prediction. IEEE Transactions on Automatic Control, 60(6), 1647–1652.
Zorzi, M., & Sepulchre, R. (2016). AR identification of latent-variable graphical models. IEEE Transactions on Automatic Control, 61(9), 2327–2340.