A

Mathematical Basics

A.1 Inverses of Structured Matrices

This subsection presents the inverses of structured matrices which occur frequently in this book. In the following, the matrices $\boldsymbol{A}, \boldsymbol{D} \in \mathbb{C}^{M \times M}$ and $\boldsymbol{C}, \boldsymbol{F} \in \mathbb{C}^{N \times N}$ are Hermitian and invertible, and $\boldsymbol{B} \in \mathbb{C}^{N \times M}$, $N, M \in \mathbb{N}$.

A.1.1 Inverse of a Schur Complement

If the matrix $\boldsymbol{F}$ is the *Schur complement* (see, e. g., [210]) of $\boldsymbol{C}$, i. e.,

$$\boldsymbol{F} = \boldsymbol{C} - \boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{B}^{\mathrm{H}}, \tag{A.1}$$

where the matrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ are partitions of a larger matrix (cf. Equation A.4), its inverse computes as

$$\boldsymbol{F}^{-1} = \boldsymbol{C}^{-1} + \boldsymbol{C}^{-1}\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{C}^{-1}, \tag{A.2}$$

with the matrix

$$\boldsymbol{D} = \boldsymbol{A} - \boldsymbol{B}^{\mathrm{H}}\boldsymbol{C}^{-1}\boldsymbol{B}, \tag{A.3}$$

being the Schur complement of $\boldsymbol{A}$. In [210], the inversion of $\boldsymbol{F}$ is presented as the *matrix inversion lemma*.

A.1.2 Inverse of a Partitioned Matrix

If the Hermitian and non-singular matrix $\boldsymbol{X} \in \mathbb{C}^{(M+N)\times(M+N)}$ can be partitioned according to

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B}^{\mathrm{H}} \\ \boldsymbol{B} & \boldsymbol{C} \end{bmatrix}, \tag{A.4}$$

its inverse computes as

$$\boldsymbol{X}^{-1} = \begin{bmatrix} \boldsymbol{D}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{F}^{-1} \\ -\boldsymbol{F}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1} & \boldsymbol{F}^{-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}^{-1} & -\boldsymbol{D}^{-1}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{C}^{-1} \\ -\boldsymbol{C}^{-1}\boldsymbol{B}\boldsymbol{D}^{-1} & \boldsymbol{F}^{-1} \end{bmatrix}, \tag{A.5}$$

with the Schur complements $\boldsymbol{D}$ and $\boldsymbol{F}$ as given in Equations (A.3) and (A.1), respectively. The first line of Equation (A.5) is the well-known *inversion lemma for partitioned matrices* (see, e. g., [210]). Note that the second equality of Equation (A.5) holds if

$$\boldsymbol{A}^{-1}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{F}^{-1} = \boldsymbol{D}^{-1}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{C}^{-1}, \tag{A.6}$$

which can be easily shown by using Equation (A.2) and the inverse

$$\boldsymbol{D}^{-1} = \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{F}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}, \tag{A.7}$$

of the Schur complement $\boldsymbol{D}$ of $\boldsymbol{A}$.

Combining Equation (A.7) with Equation (A.5) yields the following alternative representation of the inverse:

$$\boldsymbol{X}^{-1} = \begin{bmatrix} \boldsymbol{A}^{-1} & \boldsymbol{0}_{M\times N} \\ \boldsymbol{0}_{N\times M} & \boldsymbol{0}_{N\times N} \end{bmatrix} + \begin{bmatrix} \boldsymbol{A}^{-1}\boldsymbol{B}^{\mathrm{H}} \\ -\boldsymbol{I}_N \end{bmatrix} \boldsymbol{F}^{-1} \begin{bmatrix} \boldsymbol{B}\boldsymbol{A}^{-1} & -\boldsymbol{I}_N \end{bmatrix}. \tag{A.8}$$

A.2 Matrix Norms

Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{C}^{N\times M}$, $N, M \in \mathbb{N}$, and $x \in \mathbb{C}$. Then, $\|\cdot\|$ is a *matrix norm* if the following conditions are fulfilled (e. g., [237]):

1. $\|\boldsymbol{X}\| \geq 0$, and $\|\boldsymbol{X}\| = 0$ only if $\boldsymbol{X} = \boldsymbol{0}_{N\times M}$,
2. $\|\boldsymbol{X} + \boldsymbol{Y}\| \leq \|\boldsymbol{X}\| + \|\boldsymbol{Y}\|$, and
3. $\|x\boldsymbol{X}\| = |x| \, \|\boldsymbol{X}\|$.

The next three subsections review some examples of matrix norms which are of great importance for this book.

A.2.1 Hilbert–Schmidt or Frobenius norm

The most important matrix norm is the *Hilbert–Schmidt* or *Frobenius norm* defined as

$$\|\boldsymbol{X}\|_{\mathrm{F}} = \sqrt{\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} |x_{n,m}|^2} = \sqrt{\operatorname{tr}\{\boldsymbol{X}^{\mathrm{H}}\boldsymbol{X}\}}, \tag{A.9}$$

where $x_{n,m}$ is the element in the $(n+1)$th row and $(m+1)$th column of $\boldsymbol{X}$ and $\operatorname{tr}\{\cdot\}$ denotes the trace of a matrix. If $M = 1$, i. e., the matrix $\boldsymbol{X} \in \mathbb{C}^{N\times M}$

reduces to the vector $\boldsymbol{x} \in \mathbb{C}^N$, the Frobenius norm $\|\boldsymbol{x}\|_\mathrm{F}$ is equal to the *Euclidean vector norm* or *vector 2-norm* $\|\boldsymbol{x}\|_2$, i. e.,

$$\|\boldsymbol{x}\|_2 = \sqrt{\sum_{n=0}^{N-1} |x_n|^2} = \sqrt{\boldsymbol{x}^\mathrm{H}\boldsymbol{x}} = \|\boldsymbol{x}\|_\mathrm{F}\,, \tag{A.10}$$

where x_n denotes the $(n+1)$th entry in $\boldsymbol{x}$.

A.2.2 A-norm

If $\boldsymbol{A} \in \mathbb{C}^{N\times N}$ is Hermitian and positive definite, we define the $\boldsymbol{A}$-norm of the matrix $\boldsymbol{X} \in \mathbb{C}^{N\times M}$ as

$$\|\boldsymbol{X}\|_{\boldsymbol{A}} = \sqrt{\mathrm{tr}\left\{\boldsymbol{X}^\mathrm{H}\boldsymbol{A}\boldsymbol{X}\right\}}. \tag{A.11}$$

If $\boldsymbol{A} = \boldsymbol{I}_N$, the $\boldsymbol{A}$-norm and the Frobenius norm are identical, i. e., $\|\boldsymbol{X}\|_{\boldsymbol{I}_N} = \|\boldsymbol{X}\|_\mathrm{F}$. Note that the $\boldsymbol{A}$-norm is not defined for arbitrary matrices $\boldsymbol{A}$.

A.2.3 2-norm

Compared to the Frobenius or $\boldsymbol{A}$-norm which are *general matrix norms*, the *2-norm* of a matrix is an *induced matrix norm*, i. e., it is defined with respect to the behavior of the matrix as an operator between two normed vector spaces. The 2-norm of $\boldsymbol{X}$ reads as

$$\|\boldsymbol{X}\|_2 = \sup_{\boldsymbol{a}\neq\boldsymbol{0}_M} \frac{\|\boldsymbol{X}\boldsymbol{a}\|_2}{\|\boldsymbol{a}\|_2} = \sup_{\substack{\boldsymbol{a}\in\mathbb{C}^M \\ \|\boldsymbol{a}\|_2=1}} \|\boldsymbol{X}\boldsymbol{a}\|_2\,, \tag{A.12}$$

and is equal to the largest singular value $\sigma_\mathrm{max}(\boldsymbol{X})$ of $\boldsymbol{X}$, i. e., $\|\boldsymbol{X}\|_2 = \sigma_\mathrm{max}(\boldsymbol{X})$.

With the 2-norm, one can express the *condition number* $\kappa(\boldsymbol{X})$ of a matrix $\boldsymbol{X}$ as follows:

$$\kappa(\boldsymbol{X}) = \frac{\sigma_\mathrm{max}(\boldsymbol{X})}{\sigma_\mathrm{min}(\boldsymbol{X})} = \|\boldsymbol{X}\|_2 \left\|\boldsymbol{X}^\dagger\right\|_2\,, \tag{A.13}$$

where $\sigma_\mathrm{min}(\boldsymbol{X})$ denotes the smallest singular value of $\boldsymbol{X}$ and $\boldsymbol{X}^\dagger = (\boldsymbol{X}^\mathrm{H}\boldsymbol{X})^{-1}\boldsymbol{X}^\mathrm{H}$ is the left-hand side *Moore–Penrose pseudoinverse* [167, 168, 184, 97] of $\boldsymbol{X}$. If $\boldsymbol{X}$ is square and non-singular, $\kappa(\boldsymbol{X}) = \|\boldsymbol{X}\|_2 \left\|\boldsymbol{X}^{-1}\right\|_2$.

A.3 Chebyshev Polynomials

The *Chebyshev polynomial* $T^{(n)}(x)$ of the first kind and degree $n \in \mathbb{N}_0$ [28] is defined as

$$T^{(n)}(x) = \frac{1}{2}\left(\left(x + \left(x^2 - 1\right)^{\frac{1}{2}}\right)^n + \left(x - \left(x^2 - 1\right)^{\frac{1}{2}}\right)^n\right) \in \mathbb{R}, \quad x \in \mathbb{R}. \tag{A.14}$$

With the inverse cosine defined as [1]

$$\arccos(x) = -\mathrm{j}\,\log\left(x + \left(x^2 - 1\right)^{\frac{1}{2}}\right) \in \mathbb{C}, \tag{A.15}$$

and due to the fact that $(x+(x^2-1)^{1/2})(x-(x^2-1)^{1/2}) = 1$, Equation (A.14) can be rewritten as

$$T^{(n)}(x) = \frac{1}{2}\left(\exp\left(\mathrm{j}\,n\arccos(x)\right)\right) + \exp\left(-\mathrm{j}\,n\arccos(x)\right))\,. \tag{A.16}$$

Finally, with *Euler's formula* $\exp(z) = \cos(z) + \mathrm{j}\,\sin(z)$, the cosine can be replaced by $\cos(z) = (\exp(\mathrm{j}\,z) + \exp(-\mathrm{j}\,z))/2$, and Equation (A.16) reads as

$$T^{(n)}(x) = \cos\left(n\arccos(x)\right) \in \mathbb{R}, \quad x \in \mathbb{R}, \tag{A.17}$$

which is an alternative expression of the Chebyshev polynomial of the first kind and degree n. A third representation of the Chebyshev polynomial $T^{(n)}(x)$ is given by the recursion formulas (e. g., [1])

$$T^{(n)}(x) = 2xT^{(n-1)}(x) - T^{(n-2)}(x), \quad n \geq 2, \tag{A.18}$$

$$T^{(2n)}(x) = 2T^{(n),2}(x) - 1, \quad n \geq 0, \tag{A.19}$$

initialized with $T^{(0)}(x) = 1$ and $T^{(1)}(x) = x$. Figure A.1 depicts exemplarily the Chebyshev polynomials $T^{(n)}(x)$ of degree $n \in \{2, 3, 8\}$.

It follows the most important properties of Chebyshev polynomials of the first kind. From Equation (A.17), we see that $|T^{(n)}(x)| \leq 1$ for $x \in [-1, 1]$ and $T^{(n)}(0) = 1$. Besides, the Chebyshev polynomial $T^{(n)}(x)$ is even or odd, if n is even or odd, respectively (cf. Equation A.14). Finally, it can be easily verified that

$$T^{(n)}(x_i) = (-1)^i \quad \text{with} \quad x_i = \cos\left(\frac{i\pi}{n}\right), \quad i \in \{0, 1, \ldots, n\}, \tag{A.20}$$

are the $n + 1$ local extrema of the Chebyshev polynomial $T^{(n)}(x)$ if x is restricted to the region $[-1, 1]$. Note that $T^{(n)}(-1) = (-1)^n$ and $T^{(n)}(1) = 1$ are no longer local extrema of $T^{(n)}(x)$ if $x \in \mathbb{R}$ (cf. also Figure A.1). Since a polynomial in $\mathbb{R}$ of degree n has at most $n - 1$ extrema and due to the fact that all of the extrema of $T^{(n)}(x)$ lie in $[-1, 1]$ and $|T^{(n)}(\pm 1)| = 1$, the values $|T^{(n)}(x)|$ for x outside of $[-1, 1]$ are larger than one, i. e., $|T^{(n)}(x)| > 1$ for $|x| > 1$.

A.4 Regularization

Given the operator $A : \mathcal{X} \to \mathcal{B}, x \mapsto b$, between the abstract vector spaces $\mathcal{X}$ and $\mathcal{B}$, and an output vector b, the *inverse problem* (e. g., [110]) is to find the

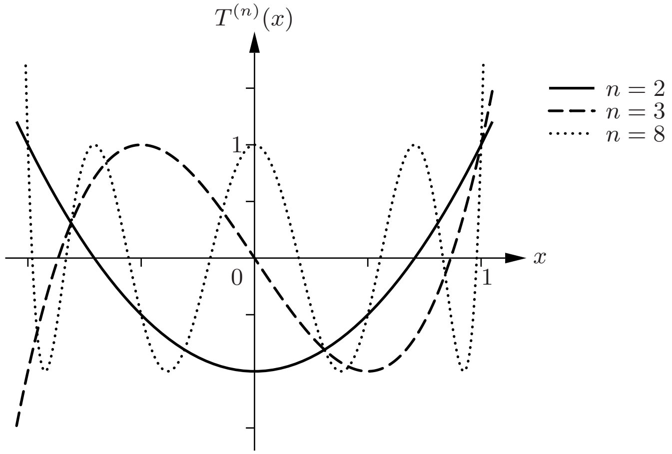

Fig. A.1. Chebyshev polynomials $T^{(n)}(x)$ of the first kind and degree $n \in \{2, 3, 8\}$

corresponding input vector x such that the application of A to x results in b. Although the operator can be either linear or non-linear, we restrict ourselves in the following to the linear case which is the focus of this book. One example for a linear inverse problem is to solve the *Fredholm integral equation of the first kind* (see, e. g., [99]), i. e.,

$$\int_0^1 A(s,t)x(t)\,\mathrm{d}t = b(s), \quad 0 \leq s \leq 1, \tag{A.21}$$

with respect to the unknown function $x : [0,1] \to \mathbb{C}, t \mapsto x(t)$. Here, the square integrable kernel function $A : [0,1] \times [0,1] \to \mathbb{C}, (s,t) \mapsto A(s,t)$, and the right-hand side function $b : [0,1] \to \mathbb{C}, s \mapsto b(s)$, are known functions. Another example of a linear inverse problem where the vector spaces have the finite dimension NM, is solving the system

$$\sum_{m=0}^{N-1} a_{n,m} \bar{\boldsymbol{x}}_m^{\mathrm{T}} = \bar{\boldsymbol{b}}_n^{\mathrm{T}}, \quad n \in \{0, 1, \ldots, N-1\}, \quad \text{or} \quad \boldsymbol{A}\boldsymbol{X} = \boldsymbol{B}, \tag{A.22}$$

of linear equations with multiple right-hand sides with respect to the unknown matrix $\boldsymbol{X} = [\bar{\boldsymbol{x}}_0, \bar{\boldsymbol{x}}_1, \ldots, \bar{\boldsymbol{x}}_{N-1}]^{\mathrm{T}} \in \mathbb{C}^{N \times M}$, $N, M \in \mathbb{N}$. The transpose of the vector $\bar{\boldsymbol{x}}_m \in \mathbb{C}^M$ denotes the $(m+1)$th row of the matrix $\boldsymbol{X}$. Again, the system matrix $\boldsymbol{A} \in \mathbb{C}^{N \times N}$ whose element in the $(n+1)$th row and $(m+1)$th column is denoted as $a_{n,m} = \boldsymbol{e}_{n+1}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{e}_{m+1}$, and the right-hand side matrix $\boldsymbol{B} = [\bar{\boldsymbol{b}}_0, \bar{\boldsymbol{b}}_1, \ldots, \bar{\boldsymbol{b}}_{N-1}]^{\mathrm{T}} \in \mathbb{C}^{N \times M}$ are known. In the sequel, we restrict ourselves to the system of linear equations as an example of an inverse problem. However, the following derivation can be easily extended to the general inverse problem defined above.

Solving an inverse problem can be challenging if the problem is either *ill-posed* or *ill-conditioned*. Hadamard [104] defined a problem as ill-posed if no unique solution exists or if it is not a continuous function in the right-hand side, i. e., an arbitrarily small error in $\boldsymbol{B}$ can cause an arbitrarily large error in the solution $\boldsymbol{X}$. A problem is denoted as ill-conditioned if the condition number of the system matrix $\boldsymbol{A}$ (cf. Appendix A.2) is very large, i. e., a small error in $\boldsymbol{B}$ causes a large error in the solution $\boldsymbol{X}$. Especially, in floating or fixed point arithmetic, ill-conditioned problems lead to numerical instability.

The method of *regularization* (e. g., [110]) introduces additional assumptions about the solution like, e. g., smoothness or a bounded norm, in order to handle ill-posed or ill-conditioned problems. In the remainder of this section, we discuss *Tikhonov regularization* [234, 235] in order to illustrate the basic idea of regularizing the solution of an inverse problem. With the system of linear equations defined in Equation (A.22), the Tikhonov regularized solution $\boldsymbol{X}_{\mathrm{reg}} \in \mathbb{C}^{N\times M}$ can be found via the following optimization:

$$\boldsymbol{X}_{\mathrm{reg}} = \operatorname*{argmin}_{\boldsymbol{X}\in\mathbb{C}^{N\times M}} \left(\|\boldsymbol{A}\boldsymbol{X} - \boldsymbol{B}\|_{\mathrm{F}}^2 + \alpha \|\boldsymbol{X}\|_{\mathrm{F}}^2 \right). \tag{A.23}$$

Here, $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm as defined in Appendix A.2, and $\alpha \in \mathbb{R}_{0,+}$ is the so-called *Tikhonov factor*. Finally, the regularized solution computes as

$$\boldsymbol{X}_{\mathrm{reg}} = \left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} + \alpha \boldsymbol{I}_N\right)^{-1} \boldsymbol{A}^{\mathrm{H}}\boldsymbol{B}. \tag{A.24}$$

It remains to choose the factor α properly. Note that if we set $\alpha = 0$, the regularized solution in Equation (A.24) reduces to the *Least Squares* (LS) *solution*. Besides heuristic choices for α, it can be derived via a *Bayesian approach* if the solution and the right-hand side are assumed to be multivariate random variables (see, e. g., [235]).

The *filter factors* are a well-known tool to characterize regularization. If $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{H}}$ denotes the *eigen decomposition* (see, e. g., [131]) of the system matrix (cf. Equation 3.74), the exact solution of Equation (A.22) can be written as

$$\boldsymbol{X} = \boldsymbol{A}^{-1}\boldsymbol{B} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{B}. \tag{A.25}$$

Then, the regularized solution is defined as

$$\boldsymbol{X}_{\mathrm{reg}} = \boldsymbol{U}\boldsymbol{F}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{B}, \tag{A.26}$$

where the diagonal matrix $\boldsymbol{F} \in \mathbb{R}_+^{N\times N}$, in the following denoted as the *filter factor matrix*, comprises the filter factors. The diagonal elements in $\boldsymbol{F}$ attenuate the inverse eigenvalues in $\boldsymbol{\Lambda}^{-1}$ in order to suppress the eigenmodes which have the highest influence on the error in the solution. In the case of Tikhonov regularization, we get (cf. Equation A.24)

$$\boldsymbol{X}_{\text{reg}} = \boldsymbol{U}\left(\boldsymbol{\Lambda}^2 + \alpha \boldsymbol{I}_N\right)^{-1} \boldsymbol{\Lambda} \boldsymbol{U}^{\text{H}} \boldsymbol{B}, \tag{A.27}$$

leading to the filter factor matrix

$$\boldsymbol{F} = (\boldsymbol{\Lambda}^2 + \alpha \boldsymbol{I}_N)^{-1} \boldsymbol{\Lambda}^2, \quad \text{with} \quad f_i = \boldsymbol{e}_{i+1}^{\text{T}} \boldsymbol{F} \boldsymbol{e}_{i+1} = \frac{\lambda_i^2}{\lambda_i^2 + \alpha} \in \mathbb{R}_+, \tag{A.28}$$

being the Tikhonov filter factors. We observe that

$$\lim_{\lambda_i \to 0} f_i = \lim_{\lambda_i} \frac{\lambda_i^2}{\lambda_i^2 + \alpha} = 0, \tag{A.29}$$

i. e., the filter factors f_i corresponding to small eigenvalues are very small. If we remember that especially errors in the small eigenvalues cause a large error in the solution due to their inversion in Equation (A.25), we see that the suppression of the eigenmodes corresponding to small eigenvalues, by the Tikhonov filter factors decreases the error in the regularized solution $\boldsymbol{X}_{\text{reg}}$.

A.5 Vector Valued Function of a Matrix and Kronecker Product

The *vector valued function* of the matrix $\boldsymbol{A} \in \mathbb{C}^{N \times M}$, $N, M \in \mathbb{N}$, in the following denoted as the vec-*operation*, is defined as (e. g., [17])

$$\text{vec}\{\boldsymbol{A}\} = \begin{bmatrix} \boldsymbol{a}_0 \\ \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_{M-1} \end{bmatrix} \in \mathbb{C}^{NM}, \tag{A.30}$$

where the vector $\boldsymbol{a}_m = \boldsymbol{A}\boldsymbol{e}_{m+1} \in \mathbb{C}^N$, $m \in \{0, 1, \ldots, M-1\}$, is the $(m+1)$th column of $\boldsymbol{A}$.

The *Kronecker product* between the matrix $\boldsymbol{A} \in \mathbb{C}^{N \times M}$ and the matrix $\boldsymbol{B} \in \mathbb{C}^{Q \times P}$, $Q, P \in \mathbb{N}$, is defined as (e. g., [17])

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{0,0}\boldsymbol{B} & a_{0,1}\boldsymbol{B} & \cdots & a_{0,M-1}\boldsymbol{B} \\ a_{1,0}\boldsymbol{B} & a_{1,1}\boldsymbol{B} & \cdots & a_{1,M-1}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N-1,0}\boldsymbol{B} & a_{N-1,1}\boldsymbol{B} & \cdots & a_{N-1,M-1}\boldsymbol{B} \end{bmatrix} \in \mathbb{C}^{NQ \times MP}, \tag{A.31}$$

where the scalar $a_{n,m} \in \mathbb{C}$, $n \in \{0, 1, \ldots, N-1\}$, $m \in \{0, 1, \ldots, M-1\}$, denotes the element in the $(n+1)$th row and $(m+1)$th column of $\boldsymbol{A}$, i. e., $a_{n,m} = \boldsymbol{e}_{n+1}^{\text{T}} \boldsymbol{A} \boldsymbol{e}_{m+1}$.

It follows a summary of the most important properties of the vec-operation and the Kronecker product which are used in this book. A detailed derivation of the formulas can be found in [17]. If $\boldsymbol{A} \in \mathbb{C}^{N \times M}$, $\boldsymbol{B} \in \mathbb{C}^{Q \times P}$, $\boldsymbol{C} \in \mathbb{C}^{M \times S}$, $\boldsymbol{D} \in \mathbb{C}^{P \times R}$, $\boldsymbol{E} \in \mathbb{C}^{S \times R}$, $\boldsymbol{F} \in \mathbb{C}^{N \times N}$, and $\boldsymbol{G} \in \mathbb{C}^{M \times M}$, $S, R \in \mathbb{N}$, with the latter two matrices being non-singular, i. e., the inverse matrices exist, it holds:

$$(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = \boldsymbol{A}\boldsymbol{C} \otimes \boldsymbol{B}\boldsymbol{D} \in \mathbb{C}^{NQ \times SR}, \tag{A.32}$$

$$\operatorname{vec}\{\boldsymbol{A}\boldsymbol{C}\boldsymbol{E}\} = \left(\boldsymbol{E}^{\mathrm{T}} \otimes \boldsymbol{A}\right) \operatorname{vec}\{\boldsymbol{C}\} \in \mathbb{C}^{NR}, \tag{A.33}$$

$$(\boldsymbol{F} \otimes \boldsymbol{G})^{-1} = \boldsymbol{F}^{-1} \otimes \boldsymbol{G}^{-1} \in \mathbb{C}^{NM \times NM}. \tag{A.34}$$

A.6 Square Root Matrices

Let $\boldsymbol{A} \in \mathbb{C}^{N \times N}$, $N \in \mathbb{N}$, be a Hermitian and positive semidefinite matrix with rank R. Then, there exists the *eigen decomposition* (see, e. g., [131])

$$\boldsymbol{A} = \sum_{i=0}^{R-1} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{H}}, \tag{A.35}$$

with the *eigenvalues* $\lambda_i \in \mathbb{R}_+$, $i \in \{0, 1, \ldots, R-1\}$, and the corresponding orthonormal *eigenvectors* $\boldsymbol{u}_i \in \mathbb{C}^N$, i. e., $\boldsymbol{u}_i^{\mathrm{H}} \boldsymbol{u}_j = \delta_{i,j}$, $i, j \in \{0, 1, \ldots, R-1\}$. Here, $\delta_{i,j}$ denotes the *Kronecker delta* being one only for $i = j$ and zero otherwise. In the following, we review two different definitions of the *square root matrix* of $\boldsymbol{A}$. Note that we assume for all definitions that the square root matrix is Hermitian.

The *general square root matrices* $\sqrt{\boldsymbol{A}} \in \mathbb{C}^{N \times N}$ of $\boldsymbol{A}$ with $\sqrt{\boldsymbol{A}}\sqrt{\boldsymbol{A}} = (\sqrt{\boldsymbol{A}})^2 = \boldsymbol{A}$ are defined as

$$\sqrt{\boldsymbol{A}} := \sum_{i=0}^{R-1} \pm\sqrt{\lambda_i} \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{H}}. \tag{A.36}$$

Due to the fact that the signs of the square roots $\sqrt{\lambda_i}$ can be chosen arbitrarily without changing the product $(\sqrt{\boldsymbol{A}})^2 = \boldsymbol{A}$, there exists 2^R different general square root matrices.

If we are interested in a unique definition of the square root matrix of $\boldsymbol{A}$, we can restrict it to be not only Hermitian but also positive semidefinite (see, e. g., [131]), i. e., we choose the one of the general square root matrices of $\boldsymbol{A}$ where the eigenvalues are non-negative. The unique *positive semidefinite square root matrix* reads as

$$\sqrt[\oplus]{\boldsymbol{A}} := \sum_{i=0}^{R-1} \sqrt{\lambda_i} \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{H}} \in \mathbb{C}^{N \times N}. \tag{A.37}$$

A.7 Binary Galois Field

The set $\mathbb{F}$ together with the operations '$\oplus$' and '$\odot$', in the following denoted as $(\mathbb{F}, \oplus, \odot)$, is a *field* if $a, b, c \in \mathbb{F}$ satisfy the following conditions (e. g., [118]):

1. *Closure of* $\mathbb{F}$ *under* '$\oplus$' *and* '$\odot$':
   $a \oplus b \in \mathbb{F}$ and $a \odot b \in \mathbb{F}$.
2. *Associative law of both operations*:
   $a \oplus (b \oplus c) = (a \oplus b) \oplus c$ and $a \odot (b \odot c) = (a \odot b) \odot c$.
3. *Commutative law of both operations*:
   $a \oplus b = b \oplus a$ and $a \odot b = b \odot a$.
4. *Distributive law of* '$\odot$' *over* '$\oplus$':
   $a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$.
5. *Existence of an additive and a multiplicative identity*:
   There exists the *zero element* or *additive identity* $n \in \mathbb{F}$ such that $a \oplus n = a$, and the *unit element* or *multiplicative identity* $e \in \mathbb{F}$ such that $a \odot e = a$.
6. *Existence of additive and multiplicative inverses*:
   For every $a \in \mathbb{F}$, there exists an *additive inverse* $\bar{a}$ such that $a \oplus \bar{a} = n$ and if $a \neq n$, there exists a *multiplicative inverse* $\underline{a}$ such that $a \odot \underline{a} = e$.

The most famous example of a field is $(\mathbb{R}, +, \cdot)$ with the common addition '$+$', the common multiplication '$\cdot$', the zero element 0, the unit element 1, the additive inverse $-a$ of $a \in \mathbb{R}$, and the multiplicative inverse $1/a$ of $a \in \mathbb{R} \backslash \{0\}$.

The field $(\mathbb{F}, \oplus, \odot)$ with a finite set $\mathbb{F}$ is called *Galois field* (e. g., [118]). If q is the cardinality of $\mathbb{F}$, the respective Galois field is denoted as $GF(q)$. For example, the field $(\{0,1\}, \boxplus, \boxdot)$, i. e., the set $\{0,1\}$ together with the modulo 2 addition '$\boxplus$' and the modulo 2 multiplication '$\boxdot$' as defined in Table A.1, is the *binary Galois field* $GF(2)$ (cf., e. g., [163, 118]) which plays a fundamental role in digital communications (see Chapter 5).

Table A.1. Modulo 2 addition '$\boxplus$' and modulo 2 multiplication '$\boxdot$'

$\boxplus$	0	1
0	0	1
1	1	0

$\boxdot$	0	1
0	0	0
1	0	1

Note that in $GF(2)$, 0 is the zero element since $a \boxplus 0 = a$, $a \in \{0,1\}$, 1 the unit element since $a \boxdot 1 = a$, 0 the additive inverse of 0 since $0 \boxplus 0 = 0$, 1 the additive inverse of 1 since $1 \boxplus 1 = 0$, 1 the multiplicative inverse of 1 since $1 \boxdot 1 = 1$. Recall that the zero element has no multiplicative inverse. Clearly, the two elements of the binary Galois field $GF(2)$ could be also represented by other symbols than '0' and '1' if the operations in Table A.1 are defined adequately.

B

Derivations and Proofs

B.1 Inversion of Hermitian and Positive Definite Matrices

The procedure to invert a Hermitian and positive definite matrix $\boldsymbol{A} \in \mathbb{C}^{N \times N}$, $N \in \mathbb{N}$, is as follows (see, e. g., [218, 136]):

1. *Cholesky Factorization* (CF):
   Compute the lower triangular matrix $\boldsymbol{L} \in \mathbb{C}^{N \times N}$ such that $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^{\mathrm{H}}$.
2. *Lower Triangular Inversion* (LTI):
   Compute the inverse $\boldsymbol{L}^{-1} \in \mathbb{C}^{N \times N}$ of the lower triangular matrix $\boldsymbol{L}$.
3. *Lower Triangular Product* (LTP):
   Construct the inverse $\boldsymbol{A}^{-1}$ via the matrix-matrix product $\boldsymbol{A}^{-1} = \boldsymbol{L}^{-1,\mathrm{H}}\boldsymbol{L}^{-1}$.

First, the CF is given by Algorithm 2.1. which has a computational complexity of (cf. Equation 2.23)

$$\zeta_{\mathrm{CF}}(N) = \frac{1}{3}N^3 + \frac{1}{2}N^2 + \frac{1}{6}N \quad \text{FLOPs}. \tag{B.1}$$

Second, in order to compute the inverse of the lower triangular matrix $\boldsymbol{L}$, we have to solve the system $\boldsymbol{L}\boldsymbol{L}^{-1} = \boldsymbol{I}_N$ of linear equations according to the lower triangular matrix $\boldsymbol{L}^{-1}$. For $i \in \{0, 1, \ldots, N-1\}$ and $j \in \{0, 1, \ldots, i\}$, the element in the $(i+1)$th row and $(j+1)$th column of $\boldsymbol{L}\boldsymbol{L}^{-1}$ can be written as

$$[\boldsymbol{L}]_{i,:} \left[\boldsymbol{L}^{-1}\right]_{:,j} = [\boldsymbol{L}]_{i,j:i} \left[\boldsymbol{L}^{-1}\right]_{j:i,j} = \delta_{i,j}, \tag{B.2}$$

with the *Kronecker delta* $\delta_{i,j}$ being one only for $i = j$ and zero otherwise, and the notation for submatrices defined in Subsection 2.2.1 and Section 1.2, respectively. In Equation (B.2), we exploit the fact that the non-zero entries of the $(i+1)$th row of $\boldsymbol{L}$ and the $(j+1)$th column of $\boldsymbol{L}^{-1}$ whose product yields the element of the $(i+1)$th row and $(j+1)$th column of $\boldsymbol{L}\boldsymbol{L}^{-1}$, overlap only from the $(j+1)$th element to the $(i+1)$th element $(j \leq i)$. All elements of the $(i+1)$th row of $\boldsymbol{L}$ with an index higher than i and all elements of the $(j+1)$th column of $\boldsymbol{L}^{-1}$ with an index smaller than j are zero due to the lower triangular structure of $\boldsymbol{L}$ and $\boldsymbol{L}^{-1}$. Remember that the inverse of a lower triangular matrix is again lower triangular. If we extract the term $[\boldsymbol{L}]_{i,i}[\boldsymbol{L}^{-1}]_{i,j}$ from the sum given by Equation (B.2), and resolve the resulting equation according to $[\boldsymbol{L}^{-1}]_{i,j}$, we get

$$\left[\boldsymbol{L}^{-1}\right]_{i,j} = \begin{cases} \dfrac{1}{[\boldsymbol{L}]_{i,i}}, & j = i, \\ -\dfrac{[\boldsymbol{L}]_{i,j:i-1}\left[\boldsymbol{L}^{-1}\right]_{j:i-1,j}}{[\boldsymbol{L}]_{i,i}}, & j < i, \ i \in \mathbb{N}. \end{cases} \tag{B.3}$$

Algorithm B.1. computes the inverse of the lower triangular matrix $\boldsymbol{L}$ based on Equation (B.3) where the number in angle brackets denotes the number of FLOPs required to perform the respective line of the algorithm. Recall from Subsection 2.2.1 that the matrix-matrix product of an $n \times m$ matrix by an $m \times p$ matrix requires $np(2m-1)$ FLOPs.

Algorithm B.1. Lower Triangular Inversion (LTI)

for $i = 0, 1, \ldots, N-1$ **do**
  $\boldsymbol{L}^{-1} \leftarrow \boldsymbol{0}_{N \times N}$
  $[\boldsymbol{L}^{-1}]_{i,i} \leftarrow 1/[\boldsymbol{L}]_{i,i}$ $\langle 1 \rangle$
  for $j = 0, 1, \ldots, i-1 \ \wedge \ i \in \mathbb{N}$ **do**
    $[\boldsymbol{L}^{-1}]_{i,j} \leftarrow -[\boldsymbol{L}]_{i,j:i-1}[\boldsymbol{L}^{-1}]_{j:i-1,j}/[\boldsymbol{L}]_{i,i}$ $\langle 2(i-j) \rangle$
  end for
end for

The total number of FLOPs needed for LTI performed by Algorithm B.1., calculates as

$$\zeta_{\mathrm{LTI}}(N) = \sum_{i=0}^{N-1} \left(1 + \sum_{\substack{j=0 \\ i \in \mathbb{N}}}^{i-1} 2(i-j) \right) = \frac{1}{3}N^3 + \frac{2}{3}N. \tag{B.4}$$

Third, it remains to compute the matrix-matrix product $\boldsymbol{L}^{-1,\mathrm{H}}\boldsymbol{L}^{-1}$ which is the wanted inverse $\boldsymbol{A}^{-1}$. The element in the $(i+1)$th row and $(j+1)$th column of $\boldsymbol{A}^{-1}$ can be written as

$$\left[\boldsymbol{A}^{-1}\right]_{i,j} = \left[\boldsymbol{L}^{-1,\mathrm{H}}\right]_{i,:}\left[\boldsymbol{L}^{-1}\right]_{:,j} = \left[\boldsymbol{L}^{-1}\right]_{j:N-1,i}^{\mathrm{H}}\left[\boldsymbol{L}^{-1}\right]_{j:N-1,j}, \tag{B.5}$$

if $i \in \{0, 1, \ldots, N-1\}$ and $j \in \{i, i+1, \ldots, N-1\}$. In Equation (B.5), we used the lower triangular property of $\boldsymbol{L}^{-1}$, i. e., for each column with index i or j, the elements with an index smaller than i or j, respectively, are zero. Since $j \geq i$, the index j determines the first non-zero addend in the sum given by Equation (B.5). If we exploit additionally the Hermitian structure of $\boldsymbol{A}^{-1}$,

Algorithm B.2. Lower Triangular Product (LTP)

for $i = 0, 1, \ldots, N-1$ **do**
  for $j = i, i+1, \ldots, N-1$ **do**
    $[\boldsymbol{A}^{-1}]_{i,j} \leftarrow [\boldsymbol{L}^{-1}]^{\mathrm{H}}_{j:N-1,i}[\boldsymbol{L}^{-1}]_{j:N-1,j}$ $\langle 2(N-j)-1 \rangle$
    if $j \neq i$ **then**
      $[\boldsymbol{A}^{-1}]_{j,i} = [\boldsymbol{A}^{-1}]^{*}_{i,j}$
    end if
  end for
end for

we get Algorithm B.2.. Note that the inverse operation does not change the Hermitian property.

The number of FLOPs required to perform the LTP according to Algorithm B.2., is

$$\zeta_{\mathrm{LTP}}(N) = \sum_{i=0}^{N-1} \sum_{j=i}^{N-1} (2(N-j)-1) = \frac{1}{3}N^3 + \frac{1}{2}N^2 + \frac{1}{6}N, \tag{B.6}$$

which is the same computational complexity as the one of the CF. Finally, the number of FLOPs needed to invert a Hermitian and positive definite $N \times N$ matrix (*Hermitian Inversion*, HI) is given by the sum of FLOPs in Equations (B.1), (B.4), and (B.6), i. e.,

$$\boxed{\begin{aligned} \zeta_{\mathrm{HI}}(N) &= \zeta_{\mathrm{CF}}(N) + \zeta_{\mathrm{LTI}}(N) + \zeta_{\mathrm{LTP}}(N) \\ &= N^3 + N^2 + N. \end{aligned}} \tag{B.7}$$

With this knowledge, it is interesting to discuss the question whether it is computational cheaper to solve systems of linear equations, i. e., $\boldsymbol{AX} = \boldsymbol{B} \in \mathbb{C}^{N \times M}$, $N, M \in \mathbb{N}$, based on the Cholesky factorization as described in Subsection 2.2.1, or via calculating $\boldsymbol{X} = \boldsymbol{A}^{-1}\boldsymbol{B} \in \mathbb{C}^{N \times M}$ with the explicit inversion of $\boldsymbol{A}$ as presented above. Whereas the Cholesky based method requires $\zeta_{\mathrm{MWF}}(N, M)$ FLOPs as given in Equation (2.28), the HI based approach has a total computational complexity of $\zeta_{\mathrm{HI}}(N) + MN(2N-1)$ FLOPs where the additional term arises from the required multiplication of $\boldsymbol{A}^{-1} \in \mathbb{C}^{N \times N}$ by $\boldsymbol{B} \in \mathbb{C}^{N \times M}$. With some straightforward manipulations, it can be easily shown that

$$\zeta_{\mathrm{HI}}(N) + MN(2N-1) > \zeta_{\mathrm{MWF}}(N, M) \quad \Leftrightarrow \quad M < \frac{2}{3}N^2 + \frac{1}{2}N + \frac{5}{6}. \tag{B.8}$$

Thus and due the fact that $N < 2N^2/3 + N/2 + 5/6$ for all $N \in \mathbb{N}$, the Cholesky based algorithm is computationally cheaper than the HI based method if $M \leq N$. Further, for values of M which fulfill $N < M < 2N^2/3 + N/2 + 5/6$, it remains computationally less expensive. However, in the case where $M > 2N^2/3 + N/2 + 5/6$ or if the inverse of the system matrix $\boldsymbol{A}$ is explicitly needed, it is recommended to use the HI based approach to solve systems of linear equations.

B.2 Real-Valued Coefficients of Krylov Subspace Polynomials

Consider the approximation of the solution $\boldsymbol{x} \in \mathbb{C}^N$ of the system of linear equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \in \mathbb{C}^N$ where $\boldsymbol{A} \in \mathbb{C}^{N\times N}$ is Hermitian and positive definite, in the D-dimensional Krylov subspace

$$\mathcal{K}^{(D)}\left(\boldsymbol{A}, \boldsymbol{b}\right) = \operatorname{range}\left\{\left[\boldsymbol{b}\ \boldsymbol{A}\boldsymbol{b} \cdots\ \boldsymbol{A}^{D-1}\boldsymbol{b}\right]\right\}. \tag{B.9}$$

In other words, the solution $\boldsymbol{x}$ is approximated by the polynomial

$$\boldsymbol{x} \approx \boldsymbol{x}_D = \Psi^{(D-1)}(\boldsymbol{A})\boldsymbol{b} = \sum_{\ell=0}^{D-1} \psi_\ell \boldsymbol{A}^\ell \boldsymbol{b} \in \mathbb{C}^N. \tag{B.10}$$

The polynomial coefficients ψ_ℓ are chosen to minimize the squared $\boldsymbol{A}$-norm of the error vector $\boldsymbol{e}'_D = \boldsymbol{x} - \boldsymbol{x}'_D$ with $\boldsymbol{x}'_D = \sum_{\ell=0}^{D-1} \psi'_\ell \boldsymbol{A}^\ell \boldsymbol{b}$ (cf. Section 3.4 and Appendix A.2), i. e.,

$$(\psi_0, \psi_1, \ldots, \psi_{D-1}) = \operatorname*{argmin}_{\left(\psi'_0, \psi'_1, \ldots, \psi'_{D-1}\right)} \left\|\boldsymbol{e}'_D\right\|_{\boldsymbol{A}}^2. \tag{B.11}$$

Setting the partial derivatives of

$$\begin{aligned}\left\|\boldsymbol{e}_D\right\|_{\boldsymbol{A}}^2 = \boldsymbol{x}^{\mathrm{H}}\boldsymbol{x} - \boldsymbol{x}^{\mathrm{H}} \sum_{\ell=0}^{D-1} \psi'_\ell \boldsymbol{A}^\ell \boldsymbol{b} - \sum_{\ell=0}^{D-1} \psi_\ell^{\prime,*} \boldsymbol{b}^{\mathrm{H}} \boldsymbol{A}^\ell \boldsymbol{x} \\ + \sum_{\ell=0}^{D-1}\sum_{m=0}^{D-1} \psi_\ell^{\prime,*} \boldsymbol{b}^{\mathrm{H}} \boldsymbol{A}^{\ell+m} \boldsymbol{b} \psi'_m,\end{aligned} \tag{B.12}$$

with respect to $\psi_\ell^{\prime,*}$, $\ell \in \{0, 1, \ldots, D-1\}$, equal to zero at $\psi'_m = \psi_m$, yields the set of linear equations:

$$\sum_{m=0}^{D-1} \boldsymbol{b}^{\mathrm{H}} \boldsymbol{A}^{\ell+m} \boldsymbol{b} \psi_m = \boldsymbol{b}^{\mathrm{H}} \boldsymbol{A}^\ell \boldsymbol{x}, \quad \ell \in \{0, 1, \ldots, D-1\}, \tag{B.13}$$

or

$$\sum_{m=0}^{D-1} \boldsymbol{b}^{\mathrm{H}} \boldsymbol{A}^{\ell+m} \boldsymbol{b} \psi_m = \boldsymbol{b}^{\mathrm{H}} \boldsymbol{A}^{\ell-1} \boldsymbol{b}, \quad \ell \in \{0, 1, \ldots, D-1\}, \tag{B.14}$$

if we remember that $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$. Due to the fact that $\boldsymbol{A}$ is Hermitian and positive definite, the matrices $\boldsymbol{A}^i$, $i \in \mathbb{N}_0$, are also Hermitian and positive definite, and the terms $\boldsymbol{b}^{\mathrm{H}} \boldsymbol{A}^i \boldsymbol{b}$ are real-valued and positive. Thus, all elements of the linear system in Equation (B.14) are real-valued and positive and finally, the optimal coefficients $\psi_\ell \in \mathbb{R}$ for all $\ell \in \{0, 1, \ldots, D-1\}$ (cf. also [67]).

In the limiting case where $D = N$, the solution $\boldsymbol{x}$ is exactly represented by the optimal linear combination of the Krylov basis vectors using only real-valued coefficients, i. e., $\boldsymbol{x} = \boldsymbol{x}_N$. However, we are not operating in a real-valued vector space because we still need complex-valued coefficients in order to reach any $\boldsymbol{x}' \in \mathbb{C}^N$ with the linear combination of the Krylov basis vectors $\boldsymbol{b}$, $\boldsymbol{A}\boldsymbol{b}$, ..., and $\boldsymbol{A}^{N-1}\boldsymbol{b}$. Especially, if $\boldsymbol{A}\boldsymbol{x}' \neq \boldsymbol{b}$, minimizing the squared $\boldsymbol{A}$-norm of the error between $\boldsymbol{x}'_N$ and $\boldsymbol{x}'$ leads to Equation (B.13) where $\boldsymbol{x}$ is replaced by $\boldsymbol{x}'$. In this case, Equation (B.13) cannot be simplified to Equation (B.14) since $\boldsymbol{b} \neq \boldsymbol{A}\boldsymbol{x}'$, and the right-hand sides are no longer real-valued. Thus, the optimal coefficients ψ_ℓ are complex-valued for all $\ell \in \{0, 1, \ldots, D-1\}$. Only the optimal solution $\boldsymbol{x}$ of the system of linear equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ can be represented exactly by the polynomial $\Psi^{(N-1)}(\boldsymbol{A})\boldsymbol{b}$ with real-valued coefficients because of the matching between $\boldsymbol{x}$ and the special choice of the Krylov basis vectors according to the system matrix $\boldsymbol{A}$ and the right-hand side $\boldsymbol{b}$.

If the Hermitian and positive definite system of linear equations has multiple right-hand sides, the matrix coefficients have no special structure, i. e., they are in general neither Hermitian nor are their elements real-valued.

B.3 QR Factorization

The *QR factorization* [94, 237] of the matrix $\boldsymbol{A} \in \mathbb{C}^{N \times M}$, $N, M \in \mathbb{N}$, $N \geq M$, is defined as the decomposition $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$ where $\boldsymbol{Q} \in \mathbb{C}^{N \times M}$ is a matrix with mutually orthonormal columns, i. e., $\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{Q} = \boldsymbol{I}_M$, and $\boldsymbol{R} \in \mathbb{C}^{M \times M}$ is an upper triangular matrix. To access the factors $\boldsymbol{Q}$ and $\boldsymbol{R}$ of the QR factorization of $\boldsymbol{A}$, we define the operator 'QR(·)' such that $(\boldsymbol{Q}, \boldsymbol{R}) = \mathrm{QR}(\boldsymbol{A})$.

The derivation of the QR factorization is very similar to the one of the block Arnoldi algorithm in Section 3.2. However, instead of orthogonalizing the basis of the Krylov subspace, the QR factorization orthogonalizes the columns of the matrix $\boldsymbol{A}$. If we recall the *modified Gram–Schmidt procedure* [94, 237] of Equation (3.15) and apply it to orthogonalize the $(j+1)$th column of $\boldsymbol{A}$, i. e., $[\boldsymbol{A}]_{:,j}$, $j \in \{0, 1, \ldots, M-1\}$, with respect to the already orthonormal columns of $[\boldsymbol{Q}]_{:,0:j-1}$, we get

$$[\boldsymbol{Q}]_{:,j}[\boldsymbol{R}]_{j,j} = \left(\prod_{\substack{i=0 \\ j \in \mathbb{N}}}^{j-1} \boldsymbol{P}_{\perp [\boldsymbol{Q}]_{:,i}} \right) [\boldsymbol{A}]_{:,j}, \tag{B.15}$$

where $\boldsymbol{P}_{\perp [\boldsymbol{Q}]_{:,i}} = \boldsymbol{I}_N + [\boldsymbol{Q}]_{:,i}[\boldsymbol{Q}]_{:,i}^{\mathrm{H}} \in \mathbb{C}^{N \times N}$ is a projector onto the subspace which is orthogonal to the $(i+1)$th column of $\boldsymbol{Q}$, i. e., $[\boldsymbol{Q}]_{:,i}$, $i \in \{0, 1, \ldots, j-1\}$, $j \in \mathbb{N}$. Again, we used the notation for submatrices as defined in Subsection 2.2.1 and Section 1.2, respectively. Due to the fact that each column of $\boldsymbol{A}$ is only orthogonalized with respect to the j already computed orthonormal vectors, the entries of the lower triangular part of $\boldsymbol{R} = \boldsymbol{Q}^{\mathrm{H}}\boldsymbol{A}$ are equal to zero.

Algorithm B.3. performs the QR factorization based on the modified Gram–Schmidt procedure. Immediately after $[\boldsymbol{Q}]_{:,j}$ is known, the projector $\boldsymbol{P}_{\perp[\boldsymbol{Q}]_{:,j}} = \boldsymbol{I}_N + [\boldsymbol{Q}]_{:,j}[\boldsymbol{Q}]_{:,j}^{\mathrm{H}}$ is applied to each column of $[\boldsymbol{A}]_{:,j+1:M-1}$ (cf. inner **for**-loop of Algorithm B.3.). Here, the columns of $\boldsymbol{A}$ which are no longer needed in the subsequent iterations are overwritten to save storage. We refer to [94, 237] for a detailed derivation of Algorithm B.3..

Algorithm B.3. QR factorization

for $j = 0, 1, \ldots, M-1$ **do**
  $[\boldsymbol{R}]_{j,j} \leftarrow \|[\boldsymbol{A}]_{:,j}\|_2$ $\langle 2N \rangle$
  $[\boldsymbol{Q}]_{:,j} \leftarrow [\boldsymbol{A}]_{:,j} / [\boldsymbol{R}]_{j,j}$ $\langle N \rangle$
  for $i = j+1, j+2, \ldots, M-1 \wedge j < M-1$ **do**
    $[\boldsymbol{R}]_{j,i} \leftarrow [\boldsymbol{Q}]_{:,j}^{\mathrm{H}} [\boldsymbol{A}]_{:,i}$ $\langle 2N-1 \rangle$
    $[\boldsymbol{A}]_{:,i} \leftarrow [\boldsymbol{A}]_{:,i} - [\boldsymbol{Q}]_{:,j} [\boldsymbol{R}]_{j,i}$ $\langle 2N \rangle$
  end for
end for

In each line of Algorithm B.3., the number in angle brackets denotes the number of FLOPs required to perform the corresponding line of the algorithm. Remember from Section 2.2 that the scalar product of two n-dimensional vectors needs $2n-1$ FLOPs. Finally, the total number of FLOPs which are necessary to perform the QR factorization based on the modified Gram–Schmidt procedure according to Algorithm B.3., computes as

$$\begin{aligned} \zeta_{\mathrm{QR}}(N, M) &= \sum_{j=0}^{M-1} \left(3N + \sum_{\substack{i=j+1 \\ j<M-1}}^{M-1} (4N-1) \right) \\ &= (2M+1)MN - \frac{1}{2}(M-1)M. \end{aligned} \tag{B.16}$$

B.4 Proof of Proposition 3.3

The columns of $\boldsymbol{Q}$, viz., the vectors $\boldsymbol{q}_j \in \mathbb{C}^N$, $j \in \{0, 1, \ldots, D-1\}$, are orthonormal due to the Gram–Schmidt procedure. If D is an integer multiple of M, i. e., $\mu = M$, Proposition 3.3 is true because the block Lanczos–Ruhe algorithm is equal to the block Lanczos procedure by construction and $\mathcal{K}_{\mathrm{Ruhe}}^{(D)}(\boldsymbol{A}, \boldsymbol{q}_0, \boldsymbol{q}_1, \ldots, \boldsymbol{q}_{M-1}) = \mathcal{K}^{(D)}(\boldsymbol{A}, \boldsymbol{Q}_0)$. It remains to prove Proposition 3.3 for $D \neq dM$, i. e., the vectors $\boldsymbol{q}_j$, $j \in \{0, 1, \ldots, D-1\}$, are given by the linear combination

$$\begin{aligned}\boldsymbol{q}_j &= \Psi_{0,j}^{(\iota(j)-1)}(\boldsymbol{A})\boldsymbol{q}_0 + \Psi_{1,j}^{(\iota(j)-1)}(\boldsymbol{A})\boldsymbol{q}_1 + \ldots + \Psi_{\varrho(j)-1,j}^{(\iota(j)-1)}(\boldsymbol{A})\boldsymbol{q}_{\varrho(j)-1} \\ &\quad + \Psi_{\varrho(j),j}^{(\iota(j)-2)}(\boldsymbol{A})\boldsymbol{q}_{\varrho(j)} + \Psi_{\varrho(j)+1,j}^{(\iota(j)-2)}(\boldsymbol{A})\boldsymbol{q}_{\varrho(j)+1} + \ldots + \Psi_{M-1,j}^{(\iota(j)-2)}(\boldsymbol{A})\boldsymbol{q}_{M-1} \\ &= \sum_{k=0}^{\varrho(j)-1} \Psi_{k,j}^{(\iota(j)-1)}(\boldsymbol{A})\boldsymbol{q}_k + \sum_{\substack{k=\varrho(j)\\ \varrho(j)<M}}^{M-1} \Psi_{k,j}^{(\iota(j)-2)}(\boldsymbol{A})\boldsymbol{q}_k \in \mathbb{C}^N, \end{aligned} \tag{B.17}$$

where the polynomials $\Psi_{k,j}^{(\iota(j)-1)}(\boldsymbol{A})$, $k \in \{0, 1, \ldots, \varrho(j)-1\}$, of degree $\iota(j)-1$ and the polynomials $\Psi_{k,j}^{(\iota(j)-2)}(\boldsymbol{A})$, $k \in \{\varrho(j), \varrho(j)+1, \ldots, M-1\}$, of degree $\iota(j)-2$ are defined analogous to Equation (3.3).[1] The graphs of the functions $\iota(j) = \lceil (j+1)/M \rceil \in \{1, 2, \ldots, d\}$, $d = \lceil D/M \rceil$, which is one plus the maximal polynomial degree, and $\varrho(j) = j + 1 - (\iota(j)-1)M \in \{1, 2, \ldots, M\}$ which is the number of polynomials with maximal degree $\iota(j)-1$, are depicted in Figure B.1.

We prove by induction on j. For $j \in \{0, 1, \ldots, M-1\}$, $\boldsymbol{q}_0 = \Psi_{0,0}^{(0)}(\boldsymbol{A})\boldsymbol{q}_0$ with $\Psi_{0,0}^{(0)}(\boldsymbol{A}) := \boldsymbol{I}_N$, $\boldsymbol{q}_1 = \Psi_{0,1}^{(0)}(\boldsymbol{A})\boldsymbol{q}_0 + \Psi_{1,1}^{(0)}(\boldsymbol{A})\boldsymbol{q}_1$ with $\Psi_{0,1}^{(0)}(\boldsymbol{A}) := \boldsymbol{0}_{N\times N}$ and $\Psi_{1,1}^{(0)}(\boldsymbol{A}) := \boldsymbol{I}_N$, ..., and $\boldsymbol{q}_{M-1} = \sum_{k=0}^{M-1} \Psi_{k,M-1}^{(0)}(\boldsymbol{A})\boldsymbol{q}_k$ with $\Psi_{k,M-1}^{(0)}(\boldsymbol{A}) := \boldsymbol{0}_{N\times N}$ for $k < M-1$ and $\Psi_{M-1,M-1}^{(0)}(\boldsymbol{A}) := \boldsymbol{I}_N$. Assume that Equation (B.17) is true for all $M-1 \le i \le j$, $j \in \{M-1, M, \ldots, D-2\}$.[2] Then, we have to show that it is also true for $i = j+1$. Use Equation (3.25) and replace m by $j - M + 1$ to get

$$\boldsymbol{q}_{j+1} h_{j+1,j-M+1} = \boldsymbol{A}\boldsymbol{q}_{j-M+1} - \sum_{\substack{i=j-2M+1\\ i\in\mathbb{N}_0}}^{j} \boldsymbol{q}_i h_{i,j-M+1}. \tag{B.18}$$

With the relations $\iota(j-M+1) = \iota(j+1) - 1$ and $\varrho(j-M+1) = \varrho(j+1)$,[3] Equation (B.18) may be rewritten as

[1] The additional index j in the notation of the polynomials is necessary since the polynomial coefficients can change even if the degree of the polynomial is not increasing.

[2] Since we have already proven that Equation (B.17) holds for $j \in \{0, 1, \ldots, M-1\}$, it remains to show the induction from j to $j+1$ beginning from $j = M-1$. This strategy is necessary because Equation (3.25) which we use to derive the induction, only holds for $j \in \{M-1, M, \ldots, D-2\}$.

[3] It holds that $\iota(j-M+1) = \lceil (j-M+2)/M \rceil = \lceil (j+2)/M \rceil - 1 = \iota(j+1) - 1$ and therefore, $\varrho(j-M+1) = j - M + 2 - (\iota(j-M+1)-1)M = j - M + 2 - (\iota(j+1)-2)M = j + 2 - (\iota(j+1)-1)M = \varrho(j+1)$.

$$\begin{aligned}\boldsymbol{q}_{j+1} h_{j+1,j-M+1} = \boldsymbol{A} \left(\sum_{k=0}^{\varrho(j+1)-1} \Psi_{k,j-M+1}^{(\iota(j+1)-2)}(\boldsymbol{A})\boldsymbol{q}_k + \sum_{\substack{k=\varrho(j+1)\\ \varrho(j+1)<M}}^{M-1} \Psi_{k,j-M+1}^{(\iota(j+1)-3)}(\boldsymbol{A})\boldsymbol{q}_k \right) \\ - \sum_{\substack{i=j-2M+1\\ i\in\mathbb{N}_0}}^{j} h_{i,j-M+1} \left(\sum_{k=0}^{\varrho(i)-1} \Psi_{k,i}^{(\iota(i)-1)}(\boldsymbol{A})\boldsymbol{q}_k \right. \\ \left. + \sum_{\substack{k=\varrho(i)\\ \varrho(i)<M}}^{M-1} \Psi_{k,i}^{(\iota(i)-2)}(\boldsymbol{A})\boldsymbol{q}_k \right), \quad \text{(B.19)}\end{aligned}$$

if we recall the polynomial representation of $\boldsymbol{q}_j$ in Equation (B.17). First consider the case where $j = (\ell+1)M-1$, $\ell \in \{0,1,\ldots,d-2\}$, i. e., $\iota(j) = \ell+1$, $\iota(j+1) = \iota(j)+1 = \ell+2$, $\varrho(j) = M$, and $\varrho(j+1) = 1$ (cf. Figure B.1). Hence,

$$\begin{aligned}\boldsymbol{q}_{j+1} &= \underbrace{h_{j+1,j-M+1}^{-1} \boldsymbol{A}\Psi_{0,j-M+1}^{(\iota(j+1)-2)}(\boldsymbol{A})\boldsymbol{q}_0}_{k=0,\ \text{degree}\ \iota(j+1)-1} + \underbrace{h_{j+1,j-M+1}^{-1} \sum_{k=1}^{M-1} \boldsymbol{A}\Psi_{k,j-M+1}^{(\iota(j+1)-3)}(\boldsymbol{A})\boldsymbol{q}_k}_{k\in\{1,2,\ldots,M-1\},\ \text{maximal degree}\ \iota(j+1)-2} \\ &\quad \underbrace{- \sum_{\substack{i=j-2M+1\\ i\in\mathbb{N}_0}}^{j} \frac{h_{i,j-M+1}}{h_{j+1,j-M+1}} \left(\sum_{k=0}^{\varrho(i)-1} \Psi_{k,i}^{(\iota(i)-1)}(\boldsymbol{A})\boldsymbol{q}_k + \sum_{\substack{k=\varrho(i)\\ \varrho(i)<M}}^{M-1} \Psi_{k,i}^{(\iota(i)-2)}(\boldsymbol{A})\boldsymbol{q}_k \right)}_{k\in\{0,1,\ldots,M-1\},\ \text{maximal degree}\ \iota(j)-1=\iota(j+1)-2} \\ &= \Psi_{0,j+1}^{(\iota(j+1)-1)}(\boldsymbol{A})\boldsymbol{q}_0 + \sum_{k=1}^{M-1} \Psi_{k,j+1}^{(\iota(j+1)-2)}(\boldsymbol{A})\boldsymbol{q}_k, \qquad \text{(B.20)}\end{aligned}$$

which confirms Equation (B.17) for the special case of $j = (\ell+1)M-1$. In Equation (B.20), only the terms with $k = 0$ have a maximal degree of $\iota(j+1)-1$ which justifies the last equality. It remains to find the polynomial representation of $\boldsymbol{q}_{j+1}$ for $j \neq (\ell+1)M-1$, $\ell \in \{0,1,\ldots,d-2\}$. In this case, $\iota(j+1) = \iota(j)$ and $\varrho(j+1) = \varrho(j)+1$ (cf. Figure B.1), and Equation (B.19) reads as

$$
\begin{aligned}
\boldsymbol{q}_{j+1}h_{j+1,j-M+1} = & \underbrace{\sum_{k=0}^{\varrho(j)-1} \boldsymbol{A}\Psi_{k,j-M+1}^{(\iota(j+1)-2)}(\boldsymbol{A})\boldsymbol{q}_k - \sum_{i=(\iota(j)-1)M}^{j} h_{i,j-M+1} \sum_{k=0}^{\varrho(i)-1} \Psi_{k,i}^{(\iota(i)-1)}(\boldsymbol{A})\boldsymbol{q}_k}_{k \in \{0,1,\ldots,\varrho(j)-1\},\ \text{degree}\ \iota(j)-1=\iota(j+1)-1} \\
& + \underbrace{\boldsymbol{A}\Psi_{\varrho(j),j-M+1}^{(\iota(j+1)-2)}(\boldsymbol{A})\boldsymbol{q}_{\varrho(j)}}_{k=\varrho(j),\ \varrho(j)<M,\ \text{degree}\ \iota(j+1)-1} + \underbrace{\sum_{\substack{k=\varrho(j)+1 \\ \varrho(j)<M-1}}^{M-1} \boldsymbol{A}\Psi_{k,j-M+1}^{(\iota(j+1)-3)}(\boldsymbol{A})\boldsymbol{q}_k}_{k \in \{\varrho(j)+1,\varrho(j)+2,\ldots,M-1\},\ \varrho(j)<M-1,\ \text{degree}\ \iota(j+1)-2} \\
& \underbrace{- \sum_{\substack{i=j-2M+1 \\ i\in\mathbb{N}_0}}^{(\iota(j)-1)M-1} h_{i,j-M+1} \sum_{k=0}^{\varrho(i)-1} \Psi_{k,i}^{(\iota(i)-1)}(\boldsymbol{A})\boldsymbol{q}_k}_{k \in \{0,1,\ldots,M-1\},\ \text{maximal degree}\ \iota(j)-2=\iota(j+1)-2} \\
& \underbrace{- \sum_{\substack{i=j-2M+1 \\ i\in\mathbb{N}_0}}^{j} h_{i,j-M+1} \sum_{\substack{k=\varrho(i) \\ \varrho(i)<M}}^{M-1} \Psi_{k,i}^{(\iota(i)-2)}(\boldsymbol{A})\boldsymbol{q}_k.}_{k \in \{0,1,\ldots,M-1\},\ \text{maximal degree}\ \iota(j)-2=\iota(j+1)-2}
\end{aligned}
\tag{B.21}
$$

Since the degree of the polynomial with $k = \varrho(j) = \varrho(j+1) - 1$ is increased by one, only the terms with $k \in \{0, 1, \ldots, \varrho(j+1) - 1\}$ have the maximal degree of $\iota(j+1) - 1$, and we can write

$$
\boldsymbol{q}_{j+1} = \sum_{k=0}^{\varrho(j+1)-1} \Psi_{k,j+1}^{(\iota(j+1)-1)}(\boldsymbol{A})\boldsymbol{q}_k + \sum_{\substack{k=\varrho(j+1) \\ \varrho(j+1)<M}}^{M-1} \Psi_{k,j+1}^{(\iota(j+1)-2)}(\boldsymbol{A})\boldsymbol{q}_k, \tag{B.22}
$$

which completes the proof.

B.5 Correlated Subspace

Consider the observation of the unknown vector random sequence $\boldsymbol{x}[n] \in \mathbb{C}^M$, $M \in \mathbb{N}$, with the mean $\boldsymbol{m}_{\boldsymbol{x}} \in \mathbb{C}^M$ via the vector random sequence $\boldsymbol{y}[n] \in \mathbb{C}^N$, $N \in \mathbb{N}$, $N \geq M$, with the mean $\boldsymbol{m}_{\boldsymbol{y}} \in \mathbb{C}^N$ (cf. Section 2.1). For simplicity, we assume in the following considerations that the random sequences are stationary. However, the results can be easily generalized to the non-stationary case by introducing a time-dependency as in Section 2.1. The *normalized matrix matched filter* [88] or *normalized matrix correlator* $\boldsymbol{M} \in \mathbb{C}^{N\times M}$ with orthonormal columns is defined to maximize the *correlation measure* $\mathrm{Re}\{\mathrm{tr}\{\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{x}}\}\}$, i. e., the real part of the sum of cross-covariances between each element of

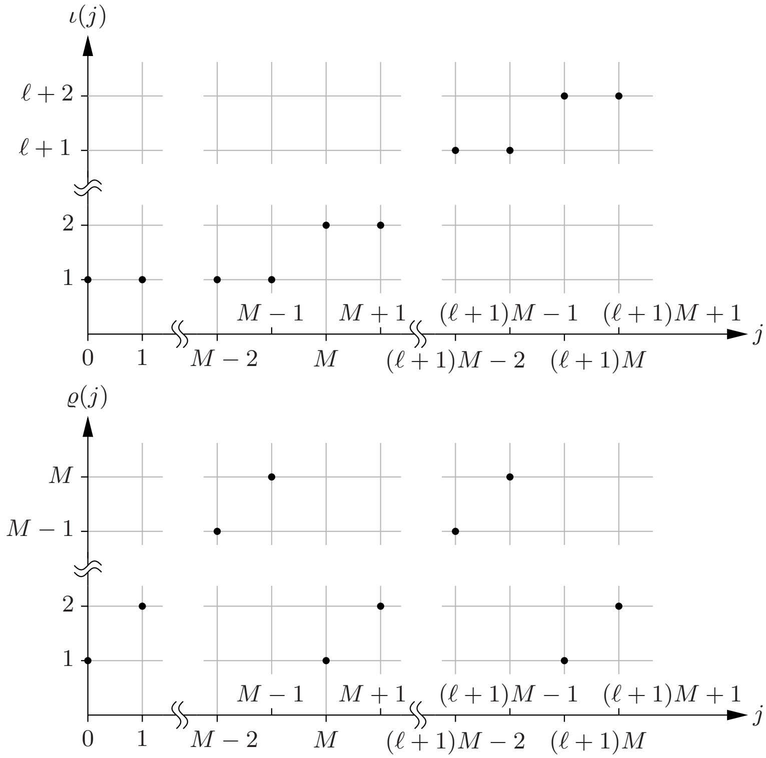

Fig. B.1. Graphs of functions $\iota(j)$, i. e., one plus the maximal polynomial degree, and the number $\varrho(j)$ of polynomials with maximal degree $\iota(j) - 1$

$\boldsymbol{x}[n]$ and the corresponding element of its output $\hat{\boldsymbol{x}}[n] = \boldsymbol{M}^{\mathrm{H}}\boldsymbol{y}[n] \in \mathbb{C}^M$.[4] The cross-covariance matrix between $\hat{\boldsymbol{x}}[n]$ and $\boldsymbol{x}[n]$ reads as $\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{x}} = \boldsymbol{M}^{\mathrm{H}}\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}} = \boldsymbol{M}^{\mathrm{H}}(\boldsymbol{R}_{\boldsymbol{y},\boldsymbol{x}} - \boldsymbol{m}_{\boldsymbol{y}}\boldsymbol{m}_{\boldsymbol{x}}^{\mathrm{H}})$.

In the sequel, we are interested in the subspace $\mathcal{M} \in G(N, M)$ spanned by the columns of $\boldsymbol{M} \in \mathbb{C}^{N \times M}$ where $G(N, M)$ denotes the *Grassmann manifold* (e. g., [13]), i. e., the set of all M-dimensional subspaces in $\mathbb{C}^N$. We denote $\mathcal{M}$ as the *correlated subspace* because the projection of the observed vector random sequence $\boldsymbol{y}[n]$ on $\mathcal{M}$ still implies the signal which is maximal correlated to $\boldsymbol{x}[n]$, i. e., which maximizes $\mathrm{Re}\{\mathrm{tr}\{\boldsymbol{C}_{\hat{\boldsymbol{x}},\boldsymbol{x}}\}\}$. If we define the *subspace correlation measure* $\rho(\mathcal{M}') \in \mathbb{R}$ of the subspace $\mathcal{M}' \in G(N, M)$ as

[4] Note that we choose the real part instead of the absolute value of the sum of cross-covariances to ensure that the phase of the solution is unique. This choice forces the output $\hat{\boldsymbol{x}}[n]$ of the normalized matrix correlator to be in-phase with $\boldsymbol{x}[n]$ motivated by the trivial case where $\hat{\boldsymbol{x}}[n] = \boldsymbol{x}[n]$.

$$\rho\left(\mathcal{M}'\right) = \max_{\boldsymbol{\Psi}' \in \mathbb{C}^{M\times M}} \operatorname{Re}\left\{\operatorname{tr}\left\{\boldsymbol{\Psi}'^{,\mathrm{H}}\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}\right\}\right\} \quad \text{s. t.} \quad \operatorname{range}\left\{\boldsymbol{M}'\right\} = \mathcal{M}' \text{ and } \boldsymbol{\Psi}'^{,\mathrm{H}}\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{M}'\boldsymbol{\Psi}' = \boldsymbol{I}_M, \tag{B.23}$$

where $\boldsymbol{M}' \in \mathbb{C}^{N\times M}$ is an arbitrary basis of the M-dimensional subspace $\mathcal{M}'$ and $\boldsymbol{\Psi}' \in \mathbb{C}^{M\times M}$ is chosen to maximize the cost function as well as to ensure that the columns of the matrix $\boldsymbol{M}'\boldsymbol{\Psi}' \in \mathbb{C}^{N\times M}$ are orthonormal to each other, the correlated subspace $\mathcal{M}$ is the subspace which maximizes $\rho(\mathcal{M}')$ over all subspaces $\mathcal{M}' \in G(N, M)$, i. e.,

$$\mathcal{M} = \operatorname*{argmax}_{\mathcal{M}' \in G(N,M)} \rho\left(\mathcal{M}'\right). \tag{B.24}$$

Before solving the optimization in Equation (B.24), we simplify the subspace correlation measure $\rho(\mathcal{M}')$ given by Equation (B.23). Partially differentiating the real-valued Lagrangian function

$$L\left(\boldsymbol{\Psi}', \boldsymbol{\Lambda}\right) = \operatorname{Re}\left\{\operatorname{tr}\left\{\boldsymbol{\Psi}'^{,\mathrm{H}}\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}} - \left(\boldsymbol{\Psi}'^{,\mathrm{H}}\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{M}'\boldsymbol{\Psi}' - \boldsymbol{I}_M\right)\boldsymbol{\Lambda}\right\}\right\}, \tag{B.25}$$

$\boldsymbol{\Lambda} \in \mathbb{C}^{M\times M}$, with respect to $\boldsymbol{\Psi}'^{,*}$ and setting the result equal to $\mathbf{0}_{M\times M}$ for $\boldsymbol{\Psi}' = \boldsymbol{\Psi}$, yields

$$\left.\frac{\partial L\left(\boldsymbol{\Psi}', \boldsymbol{\Lambda}\right)}{\partial \boldsymbol{\Psi}'^{,*}}\right|_{\boldsymbol{\Psi}'=\boldsymbol{\Psi}} = \frac{1}{2}\left(\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}} - \boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{M}'\boldsymbol{\Psi}\left(\boldsymbol{\Lambda} + \boldsymbol{\Lambda}^{\mathrm{H}}\right)\right) \stackrel{!}{=} \mathbf{0}_{M\times M}. \tag{B.26}$$

With the Hermitian matrix $\tilde{\boldsymbol{\Lambda}} := \boldsymbol{\Lambda} + \boldsymbol{\Lambda}^{\mathrm{H}} \in \mathbb{C}^{M\times M}$, we get the solution

$$\boldsymbol{\Psi}\tilde{\boldsymbol{\Lambda}} = \left(\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{M}'\right)^{-1}\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}. \tag{B.27}$$

Note that $\tilde{\boldsymbol{\Lambda}}$ is not necessarily invertible. It remains to choose $\tilde{\boldsymbol{\Lambda}}$ such that $\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{M}'\boldsymbol{\Psi} = \boldsymbol{I}_M$. Multiplying the latter equation with $\tilde{\boldsymbol{\Lambda}}$ on the left-hand and right-hand side, and replacing $\boldsymbol{\Psi}\tilde{\boldsymbol{\Lambda}}$ based on Equation (B.27) yields

$$\tilde{\boldsymbol{\Lambda}}^2 = \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}^{\mathrm{H}}\boldsymbol{M}'\left(\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{M}'\right)^{-1}\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}. \tag{B.28}$$

With the Hermitian general square root matrix as defined in Appendix A.6 and the projector $\boldsymbol{P}_{\mathcal{M}'} = \boldsymbol{M}'(\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{M}')^{-1}\boldsymbol{M}'^{,\mathrm{H}}$ projecting onto the subspace $\mathcal{M}'$, i. e., the subspace spanned by the columns of $\boldsymbol{M}'$, we get

$$\tilde{\boldsymbol{\Lambda}} = \sqrt{\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}^{\mathrm{H}}\boldsymbol{P}_{\mathcal{M}'}\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}}. \tag{B.29}$$

On the other hand, if we multiply Equation (B.27) by $\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{M}'$ on the left-hand side and exploit again the fact that $\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{M}'\boldsymbol{\Psi} = \boldsymbol{I}_M$, it follows

$$\tilde{\boldsymbol{\Lambda}} = \boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}. \tag{B.30}$$

Thus, at the optimum, we can replace the term $\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{M}'^{,\mathrm{H}}\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}$ in Equation (B.23) by $\tilde{\boldsymbol{\Lambda}}$ as given in Equation (B.29). However, the maximum defined by Equation (B.23) can only be achieved by taking the non-negative eigenvalues of the general square root matrix which is the same as using the unique positive semidefinite square root matrix instead of the general square root matrix (cf. Appendix A.6). Thus, the subspace correlation measure can finally be written as

$$\rho(\mathcal{M}') = \operatorname{tr}\left\{ \sqrt[\oplus]{\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}^{\mathrm{H}} \boldsymbol{P}_{\mathcal{M}'} \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}} \right\}, \tag{B.31}$$

and the correlated subspace is the solution of the optimization

$$\boxed{\mathcal{M} = \underset{\mathcal{M}' \in G(N,M)}{\operatorname{argmax}} \operatorname{tr}\left\{ \sqrt[\oplus]{\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}^{\mathrm{H}} \boldsymbol{P}_{\mathcal{M}'} \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}} \right\}.} \tag{B.32}$$

In Equation (B.31), we omitted the real part operator because the square root matrix $\sqrt[\oplus]{\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}^{\mathrm{H}} \boldsymbol{P}_{\mathcal{M}'} \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}}$ is by definition Hermitian and positive semidefinite (cf. Appendix A.6), thus, its trace is non-negative real-valued.

It remains to solve the optimization defined by Equation (B.32). With the parameterization of $\boldsymbol{P}_{\mathcal{M}'} = \bar{\boldsymbol{M}}' \bar{\boldsymbol{M}}'^{,\mathrm{H}}$ where $\bar{\boldsymbol{M}}' \in \mathbb{C}^{N \times M}$ and $\bar{\boldsymbol{M}}'^{,\mathrm{H}} \bar{\boldsymbol{M}}' = \boldsymbol{I}_M$, the optimal parameter $\bar{\boldsymbol{M}}$ whose columns span the correlated subspace $\mathcal{M}$, i. e., $\mathcal{M} = \operatorname{range}\{\bar{\boldsymbol{M}}\}$, is obtained via the optimization

$$\bar{\boldsymbol{M}} = \underset{\bar{\boldsymbol{M}}' \in \mathbb{C}^{N \times M}}{\operatorname{argmax}} \operatorname{tr}\left\{ \sqrt[\oplus]{\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}^{\mathrm{H}} \bar{\boldsymbol{M}}' \bar{\boldsymbol{M}}'^{,\mathrm{H}} \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}} \right\} \quad \text{s. t.} \quad \bar{\boldsymbol{M}}'^{,\mathrm{H}} \bar{\boldsymbol{M}}' = \boldsymbol{I}_M. \tag{B.33}$$

With the singular value decomposition of $\bar{\boldsymbol{M}}'^{,\mathrm{H}} \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}} = \boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}^{\mathrm{H}}$ where $\boldsymbol{U} \in \mathbb{C}^{M \times M}$ as well as $\boldsymbol{V} \in \mathbb{C}^{M \times M}$ are unitary matrices, and $\boldsymbol{\Sigma} \in \mathbb{R}_{0,+}^{M \times M}$ is diagonal, it follows

$$\operatorname{tr}\left\{ \sqrt[\oplus]{\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}^{\mathrm{H}} \bar{\boldsymbol{M}}' \bar{\boldsymbol{M}}'^{,\mathrm{H}} \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}} \right\} = \operatorname{tr}\left\{ \sqrt[\oplus]{\boldsymbol{V} \boldsymbol{\Sigma}^2 \boldsymbol{V}^{\mathrm{H}}} \right\} = \operatorname{tr}\{\boldsymbol{\Sigma}\}, \tag{B.34}$$

i. e., in Equation (B.33), the matrix $\bar{\boldsymbol{M}}$ maximizes the sum of singular values of $\bar{\boldsymbol{M}}'^{,\mathrm{H}} \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}$ over all $\bar{\boldsymbol{M}}' \in \mathbb{C}^{N \times M}$. Since $\boldsymbol{U}^{\mathrm{H}} \bar{\boldsymbol{M}}'^{,\mathrm{H}} \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}} \boldsymbol{V} = \boldsymbol{U}^{\mathrm{H}} \boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}^{\mathrm{H}} \boldsymbol{V} = \boldsymbol{\Sigma}$, we get further[5]

$$\operatorname{tr}\{\boldsymbol{\Sigma}\} = \operatorname{Re}\left\{\operatorname{tr}\left\{\boldsymbol{V} \boldsymbol{U}^{\mathrm{H}} \bar{\boldsymbol{M}}'^{,\mathrm{H}} \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}\right\}\right\} = \operatorname{Re}\left\{\operatorname{tr}\left\{\breve{\boldsymbol{M}}'^{,\mathrm{H}} \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}\right\}\right\}, \tag{B.35}$$

where $\breve{\boldsymbol{M}}' := \bar{\boldsymbol{M}}' \boldsymbol{U} \boldsymbol{V}^{\mathrm{H}} \in \mathbb{C}^{N \times M}$ and the real part operator can be introduced because the sum of singular values is by definition real-valued. Due to the fact that $\bar{\boldsymbol{M}}'^{,\mathrm{H}} \bar{\boldsymbol{M}}' = \boldsymbol{I}_M$, it holds also $\breve{\boldsymbol{M}}'^{,\mathrm{H}} \breve{\boldsymbol{M}}' = \boldsymbol{I}_M$. Clearly, the matrix $\boldsymbol{U} \boldsymbol{V}^{\mathrm{H}}$ which transforms $\bar{\boldsymbol{M}}'$ to $\breve{\boldsymbol{M}}'$ depends on $\bar{\boldsymbol{M}}'$. However, since we are finally interested in the range space of $\bar{\boldsymbol{M}}'$ and it holds that $\operatorname{range}\{\bar{\boldsymbol{M}}'\} = \operatorname{range}\{\breve{\boldsymbol{M}}'\}$, the matrix

[5] Note that $\operatorname{tr}\{\boldsymbol{A}\boldsymbol{B}\} = \operatorname{tr}\{\boldsymbol{B}\boldsymbol{A}\}$ if $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ and $\boldsymbol{B} \in \mathbb{C}^{n \times m}$.

$$\breve{\boldsymbol{M}} = \underset{\breve{\boldsymbol{M}}' \in \mathbb{C}^{N \times M}}{\operatorname{argmax}} \; \operatorname{Re}\left\{\operatorname{tr}\left\{\breve{\boldsymbol{M}}'^{,\mathrm{H}} \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}\right\}\right\} \quad \text{s. t.} \quad \breve{\boldsymbol{M}}'^{,\mathrm{H}} \breve{\boldsymbol{M}}' = \boldsymbol{I}_M, \tag{B.36}$$

is also a basis of the correlated subspace $\mathcal{M}$. Differentiating the corresponding Lagrangian function

$$L\left(\breve{\boldsymbol{M}}', \boldsymbol{\Lambda}\right) = \operatorname{Re}\left\{\operatorname{tr}\left\{\breve{\boldsymbol{M}}'^{,\mathrm{H}} \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}} - \left(\breve{\boldsymbol{M}}'^{,\mathrm{H}} \breve{\boldsymbol{M}}' - \boldsymbol{I}_M\right) \boldsymbol{\Lambda}\right\}\right\}, \tag{B.37}$$

$\boldsymbol{\Lambda} \in \mathbb{C}^{M \times M}$, with respect to $\breve{\boldsymbol{M}}'^{,*}$ and setting the result to $\mathbf{0}_{N \times M}$ for $\breve{\boldsymbol{M}}' = \breve{\boldsymbol{M}}$, yields

$$\breve{\boldsymbol{M}}\left(\boldsymbol{\Lambda} + \boldsymbol{\Lambda}^{\mathrm{H}}\right) = \boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}. \tag{B.38}$$

Again, we are only interested in the subspace spanned by the columns of $\breve{\boldsymbol{M}}$. Due to the fact that the Hermitian matrix $\boldsymbol{\Lambda} + \boldsymbol{\Lambda}^{\mathrm{H}} \in \mathbb{C}^{M \times M}$ does not change the subspace, it holds that $\operatorname{range}\{\breve{\boldsymbol{M}}\} = \operatorname{range}\{\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}\}$ and the correlated subspace is finally given by

$$\boxed{\mathcal{M} = \operatorname{range}\left\{\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}\right\}.} \tag{B.39}$$

C

Abbreviations and Acronyms

ARQ	Automatic Repeat reQuest
AV	Auxiliary Vector
AWGN	Additive White Gaussian Noise
BCD	Block Conjugate Direction
BCG	Block Conjugate Gradient
BCJR	Bahl–Cocke–Jelinek–Raviv
BER	Bit Error Rate
BGMRES	Block Generalized Minimal RESidual
Bi-CG	Bi-Conjugate Gradient
Bi-CGSTAB	Bi-Conjugate Gradient STABilized
bit	binary digit
BLR	Block Lanczos–Ruhe
BPSK	Binary Phase Shift Keying
BS	Backward Substitution
CDMA	Code Division Multiple-Access
cf.	confer (Latin, compare)
CF	Cholesky Factorization
CG	Conjugate Gradient
CME	Conditional Mean Estimator
CS	Cross-Spectral
CSI	Channel State Information
DFD	Decision Feedback Detector
DFE	Decision Feedback Equalization
DS	Direct Sequence
EDGE	Enhanced Data rates for GSM Evolution
e. g.	exempli gratia (Latin, for example)
et al.	et alii (Latin, and others)

EXIT	EXtrinsic Information Transfer
FEC	Forward Error Correction
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FLOP	FLoating point OPeration
FS	Forward Substitution
GMRES	Generalized Minimal RESidual
GPS	Global Positioning System
GSC	Generalized Sidelobe Canceler
GSM	Global Systems for Mobile communications
HI	Hermitian Inversion
i. e.	id est (Latin, that is)
i. i. e.	independent and identically distributed
INA	Institute for Numerical Analysis
ISI	InterSymbol Interference
KLT	Karhunen–Loève Transform
LLR	Log-Likelihood Ratio
LMS	Least Mean Squares
LS	Least Squares
LTI	Lower Triangular Inversion
LTP	Lower Triangular Product
MA	Multiple-Access
MAI	Multiple-Access Interference
MAP	Maximum A Posteriori
MI	Mutual Information
MIMO	Multiple-Input Multiple-Output
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
M-PSK	M-ary Phase Shift Keying
M-QAM	M-ary Quadrature Amplitude Modulation
MSE	Mean Square Error
MSMWF	MultiStage Matrix Wiener Filter
MSVWF	MultiStage Vector Wiener Filter
MSWF	MultiStage Wiener Filter
MUSIC	MUltiple SIgnal Classification
MVDR	Minimum Variance Distortionless Response
MWF	Matrix Wiener Filter
OVSF	Orthogonal Variable Spreading Factor
PC	Principal Component
QPSK	QuadriPhase Shift Keying

RC	Reduced-Complexity
RCMWF	Reduced-Complexity Matrix Wiener Filter
RLS	Recursive Least Squares
SIMO	Single-Input Multiple-Output
SNR	Signal-to-Noise Ratio
TV	Time-Variant
UCLA	University of California, Los Angeles
VCG	Vector Conjugate Gradient
VCS	Vector Cross-Spectral
viz.	videlicet (Latin, namely)
VPC	Vector Principal Component
VWF	Vector Wiener Filter
WCDMA	Wideband Code Division Multiple-Access
WF	Wiener Filter

References

1. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, Dover, 1972.
2. H. Akaike, “Block Toeplitz Matrix Inversion,” *SIAM Journal on Applied Mathematics*, vol. 24, no. 2, pp. 234–241, March 1973.
3. S. P. Applebaum and D. J. Chapman, “Adaptive Arrays with Main Beam Constraints,” *IEEE Transactions on Antennas and Propagation*, vol. AP-24, no. 5, pp. 650–662, September 1976.
4. W. E. Arnoldi, “The Principle of Minimized Iterations in the Solution of the Matrix Eigenvalue Problem,” *Quarterly of Applied Mathematics*, vol. 9, no. 1, pp. 17–29, January 1951.
5. M. E. Austin, “Decision-Feedback Equalization for Digital Communication over Dispersive Channels,” Technical Report 437, MIT Lincoln Lab, Lexington, MA, August 1967.
6. O. Axelsson, “Solution of Linear Systems of Equations: Iterative Methods,” in *Sparse Matrix Techniques (Lecture Notes in Mathematics)*, V. A. Barker, Ed., vol. 572, pp. 1–51. Springer, 1977.
7. L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate,” *IEEE Transactions on Information Theory*, vol. IT-20, no. 2, pp. 284–287, March 1974.
8. G. Bauch and N. Al-Dhahir, “Reduced-Complexity Space-Time Turbo-Equalization for Frequency-Selective MIMO Channels,” *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 819–828, October 2002.
9. P. Beckmann, *Orthogonal Polynomials for Engineers and Physicists*, The Golem Press, 1973.
10. W. Bell, “Signal Extraction for Nonstationary Time Series,” *The Annals of Statistics*, vol. 12, no. 2, pp. 646–664, June 1984.
11. Benoit, “Note sur une méthode de résolution des équations normales provenant de l’application de la méthode des moindres carrés à un système d’équations linéaires en nombre inférieur à celui des inconnues – Application de la méthode à la résolution d’un système defini d’équations linéaires, Procédé du Commandant Cholesky (French, Note on a Method for Solving Normal Equations Occuring at the Application of the Least Squares Method to a System of Linear Equations Whose Number is Smaller than the Number of Unknowns – Application of the Method to Solve the Definite System of Linear Equations,

Approach of Commandant Cholesky)," *Bulletin Géodésique (French, Geodetical Bulletin)*, vol. 2, pp. 67–77, April/May/June 1924.

12. C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon Limit Error-Correcting Coding and Decoding: Turbo-Codes (1)," *IEEE International Conference on Communications*, pp. 1064–1070, May 1993.
13. W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Elsevier, 2002.
14. M. Botsch, G. Dietl, and W. Utschick, "Iterative Multi-User Detection Using Reduced-Complexity Equalization," in *Proceedings of the 4th International Symposium on Turbo Codes & Related Topics and the 6th International ITG-Conference on Source and Channel Coding (TURBO-CODING-2006)*, April 2006.
15. J. Brehmer, G. Dietl, M. Joham, and W. Utschick, "Reduced-Complexity Linear and Nonlinear Precoding for Frequency-Selective MIMO Channels," in *Proceedings of the 2004 IEEE 60th Vehicular Technology Conference (VTC 2004-Fall)*, September 2004, vol. 5, pp. 3684–3688.
16. J. Brehmer, M. Joham, G. Dietl, and W. Utschick, "Reduced Complexity Transmit Wiener Filter Based On a Krylov Subspace Multi-Stage Decomposition," in *Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, May 2004, vol. 4, pp. 805–808.
17. J. W. Brewer, "Kronecker Products and Matrix Calculus in System Theory," *IEEE Transactions on Circuits and Systems*, vol. CAS-25, no. 9, pp. 772–781, September 1978.
18. C. Brezinski, "André Louis Cholesky," *Bulletin de la Société Mathématique de Belgique (French, Bulletin of the Mathematical Society of Belgium)*, pp. 45–50, December 1996, Supplement to volume 3, Numerical Analysis, A Numerical Analysis Conference in Honour of Jean Meinguet.
19. S. ten Brink, "Designing Iterative Decoding Schemes with the Extrinsic Information Transfer Chart," *AEÜ International Journal of Electronics and Communications*, vol. 54, no. 6, pp. 389–398, November 2000.
20. S. ten Brink, "Convergence Behavior of Iteratively Decoded Parallel Concatenated Codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727–1737, October 2001.
21. S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative Demapping and Decoding for Multilevel Modulation," in *Proceedings of the Global Telecommunications Conference (GLOBECOM '98)*, November 1998, vol. 1, pp. 579–584.
22. S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative Demapping for QPSK Modulation," *Electronics Letters*, vol. 34, no. 15, pp. 1459–1460, July 1998.
23. S. Burykh and K. Abed-Meraim, "Reduced-Rank Adaptive Filtering Using Krylov Subspace," *EURASIP Journal on Applied Signal Processing*, vol. 12, pp. 1387–1400, 2002.
24. K. A. Byerly and R. A. Roberts, "Output Power Based Partial Adaptive Array Design," in *Proceedings of the 23rd Asilomar Conference on Signals, Systems & Computers*, October 1989, vol. 2, pp. 576–580.
25. J. Capon, "High-Resolution Frequency-Wavenumber Spectrum Analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, August 1969.
26. R. H. Chan and M. K. Ng, "Conjugate Gradient Methods for Toeplitz Systems," *SIAM Review*, vol. 38, no. 3, pp. 427–482, September 1996.

27. P. S. Chang and A. N. Willson, Jr., “Analysis of Conjugate Gradient Algorithms for Adaptive Filtering,” *IEEE Transactions on Signal Processing*, vol. 48, no. 2, pp. 409–418, February 2000.
28. P. L. Chebyshev, “Sur les questions de minima qui se rattachent à la représentation approximative des fonctions (French, On the Minima-Problem Occuring at the Approximate Representation of Functions),” *Mémoires de l’Académie Impériale des Sciences de Saint Pétersbourg. Sciences Mathématiques, Physiques et Naturelles. 1. Partie, Sciences Mathématiques et Physiques. 6. Série (French, Memoirs of the Imperial Academy of Sciences of Saint Petersburg. Mathematical, Physical, and Natural Sciences. 1. Part, Mathematical and Physical Sciences. 6. Series)*, vol. 7, pp. 199–291, 1859.
29. H. Chen, T. K. Sarkar, S. A. Dianat, and J. D. Brulé, “Adaptive Spectral Estimation by the Conjugate Gradient Method,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-34, no. 2, pp. 272–284, April 1986.
30. W. Chen, U. Mitra, and P. Schniter, “On the Equivalence of Three Reduced Rank Linear Estimators with Applications to DS-CDMA,” *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2609–2614, September 2002.
31. S. Choi, Y. U. Lee, and K. Hirasawa, “Real-Time Design of a Smart Antenna System Utilizing a Modified Conjugate Gradient Method for CDMA-Based Mobile Communications,” in *Proceedings of the 1997 IEEE 47th Vehicular Technology Conference (VTC 1997-Spring)*, May 1997, pp. 183–187.
32. S. Chowdhury and M. D. Zoltowski, “Application of Conjugate Gradient Methods in MMSE Equalization for the Forward Link of DS-CDMA,” in *Proceedings of the 2001 IEEE 54th Vehicular Technology Conference (VTC 2001-Fall)*, October 2001, pp. 2434–2438.
33. S. Chowdhury, M. D. Zoltowski, and J. S. Goldstein, “Reduced-Rank Adaptive MMSE Equalization for the Forward Link in High-Speed CDMA,” in *Proceedings of the 43rd Midwest Symposium on Circuits and Systems*, August 2000, vol. 1, pp. 6–9.
34. P. Cifuentes, W. L. Myrick, S. Sud, J. S. Goldstein, and M. D. Zoltowski, “Reduced Rank Matrix Multistage Wiener Filter with Applications in MMSE Joint Multiuser Detection for DS-CDMA,” in *Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’02)*, May 2002, vol. 3, pp. 2605–2608.
35. B.-I. Clasen, “Sur une nouvelle méthode de résolution des équations linéaires et sur l’application de cette méthode au calcul des déterminants (French, On a New Method for Solving Systems of Linear Equations and Its Application to the Computation of Determinants),” *Annales de la Société Scientifique de Bruxelles (French, Annals of the Scientific Society of Brussels)*, vol. 12, no. 2, pp. 251–281, 1888.
36. P. Concus, G. H. Golub, and D. P. O’Leary, “A Generalized Conjugate Gradient Method for the Numerical Solution of Elliptic Partial Differential Equations,” in *Sparse Matrix Computations*, J. R. Bunch and D. J. Rose, Eds., pp. 309–332. Academic Press, 1976.
37. J. W. Cooley, “How the FFT Gained Acceptance,” in *A History of Scientific Computing*, S. G. Nash, Ed., pp. 133–140. ACM Press, Addison-Wesley, 1990.
38. J. W. Cooley and J. W. Tukey, “An Algorithm for the Machine Calculation of Complex Fourier Series,” *Mathematics of Computation*, vol. 19, no. 2, pp. 297–301, April 1965.

39. J. W. Daniel, “The Conjugate Gradient Method for Linear and Nonlinear Operator Equations,” *SIAM Journal on Numerical Analysis*, vol. 4, no. 1, pp. 10–26, 1967.
40. S. Darlington, “Linear Least-Squares Smoothing and Prediction, With Applications,” *Bell Systems Technical Journal*, vol. 37, pp. 1221–1294, September 1958.
41. A. Dejonghe and L. Vandendorpe, “A Comparison of Bit and Symbol Interleaving in MMSE Turbo-Equalization,” in *Proceedings of the 2003 IEEE International Conference on Communications (ICC 2003)*, May 2003, vol. 4, pp. 2928–2932.
42. G. Dietl, “Conjugate Gradient Implementation of Multi-Stage Nested Wiener Filter for Reduced Dimension Processing,” Diploma thesis, Munich University of Technology, May 2001.
43. G. Dietl, M. Botsch, F. A. Dietrich, and W. Utschick, “Robust and Reduced-Rank Matrix Wiener Filter Based on the Block Conjugate Gradient Algorithm,” in *Proceedings of the 6th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2005)*, June 2005, pp. 555–559.
44. G. Dietl and P. Breun, “Introduction to Reduced-Rank Matrix Wiener Filters,” Technical Report TUM-LNS-TR-04-02, Munich University of Technology, April 2004.
45. G. Dietl, P. Breun, and W. Utschick, “Block Krylov Methods in Time-Dispersive MIMO Systems,” in *Proceedings of the 2004 IEEE International Conference on Communications (ICC 2004)*, June 2004, vol. 5, pp. 2683–2688.
46. G. Dietl, I. Groh, and W. Utschick, “Low-Complexity MMSE Receivers Based on Weighted Matrix Polynomials in Frequency-Selective MIMO Systems,” in *Proceedings of the 8th International Symposium on Signal Processing and Its Applications (ISSPA 2005)*, August 2005, vol. 2, pp. 699–702.
47. G. Dietl, M. Joham, P. Kreuter, J. Brehmer, and W. Utschick, “A Krylov Subspace Multi-Stage Decomposition of the Transmit Wiener Filter,” in *Proceedings of the ITG Workshop on Smart Antennas, 2004*, March 2004.
48. G. Dietl, M. Joham, W. Utschick, and J. A. Nossek, “Long-Term Krylov-Prefilter for Reduced-Dimension Receive Processing in DS-CDMA Systems,” in *Proceedings of the 2002 IEEE 56th Vehicular Technology Conference (VTC 2002-Fall)*, September 2002, pp. 492–495.
49. G. Dietl, C. Mensing, and W. Utschick, “Iterative Detection Based on Reduced-Rank Equalization,” in *Proceedings of the 2004 IEEE 60th Vehicular Technology Conference (VTC 2004-Fall)*, September 2004, vol. 3, pp. 1533–1537.
50. G. Dietl, C. Mensing, and W. Utschick, “Iterative Detection Based on Widely Linear Processing and Real-Valued Symbol Alphabets,” in *Proceedings of the Eleventh European Wireless Conference 2005*, April 2005, vol. 1, pp. 74–78.
51. G. Dietl, C. Mensing, W. Utschick, J. A. Nossek, and M. D. Zoltowski, “Multi-Stage MMSE Decision Feedback Equalization for EDGE,” in *Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’03)*, April 2003, vol. 4, pp. 509–513.
52. G. Dietl and W. Utschick, “On Reduced-Rank Aproaches to Matrix Wiener Filters in MIMO Systems,” in *Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2003)*, December 2003.

53. G. Dietl and W. Utschick, “Low-Complexity Equalizers – Rank Versus Order Reduction,” in *Proceedings of the 7th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2006)*, 2006.
54. G. Dietl and W. Utschick, “MMSE Turbo Equalisation for Real-Valued Symbols,” *European Transactions on Telecommunications*, vol. 17, no. 3, pp. 351–359, May/June 2006.
55. G. Dietl and W. Utschick, “Complexity Reduction of Iterative Receivers Using Low-Rank Equalization,” *IEEE Transactions on Signal Processing*, vol. 55, no. 3, pp. 1035–1046, March 2007.
56. G. Dietl, M. D. Zoltowski, and M. Joham, “Recursive Reduced-Rank Adaptive Equalization for Wireless Communications,” in *Proceedings of SPIE – The International Society for Optical Engineering*, April 2001, vol. 4395, pp. 16–27.
57. G. Dietl, M. D. Zoltowski, and M. Joham, “Reduced-Rank Equalization for EDGE via Conjugate Gradient Implementation of Multi-Stage Nested Wiener Filter,” in *Proceedings of the 2001 IEEE 54th Vehicular Technology Conference (VTC 2001-Fall)*, October 2001, pp. 1912–1916.
58. F. Dietrich, G. Dietl, M. Joham, and W. Utschick, “Robust and Reduced Rank Space-Time Decision Feedback Equalization,” in *Smart Antennas – State of the Art*, EURASIP Book Series on Signal Processing & Communications, chapter 10, part I: Receiver. Hindawi Publishing Corporation, 2005.
59. P. Ding, M. D. Zoltowski, and M. Firmoff, “Preconditioned Conjugate Gradient Based Fast Computation of Indirect Decision Feedback Equalizer,” in *Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’04)*, May 2004, vol. 4, pp. 1001–1004.
60. B. Dong and X. Wang, “Sampling-Based Soft Equalization for Frequency-Selective MIMO Channels,” *IEEE Transactions on Communications*, vol. 53, no. 2, pp. 278–288, February 2005.
61. C. Douillard, M. Jézéquel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, “Iterative Correction of Intersymbol Interference: Turbo-Equalization,” *European Transactions on Telecommunications*, vol. 6, no. 5, pp. 507–511, September/October 1995.
62. A. Duel-Hallen, “A Family of Multiuser Decision-Feedback Detectors for Asynchronous Code-Division Multiple-Access Channels,” *IEEE Transactions on Communications*, vol. 43, no. 2/3/4, pp. 421–434, February/March/April 1995.
63. C. Dumard and T. Zemen, “Double Krylov Subspace Approximation for Low Complexity Iterative Multi-User Decoding and Time-Variant Channel Estimation,” in *Proceedings of the 6th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2005)*, June 2005, pp. 328–332.
64. C. Eckart and G. Young, “The Approximation of One Matrix by Another of Lower Rank,” *Psychometrika*, vol. 1, no. 3, pp. 211–218, September 1936.
65. P. Elias, “Coding for Noisy Channels,” in *IRE WESCON Convention Record, Part 4*, 1955, pp. 37–47.
66. J. Erhel and F. Guyomarc’h, “An Augmented Conjugate Gradient Method for Solving Consecutive Symmetric Positive Definite Linear Systems,” *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1279–1299, 2000.
67. V. T. Ermolayev and A. G. Flaksman, “Signal Processing in Adaptive Arrays Using Power Basis,” *International Journal of Electronics*, vol. 75, no. 4, pp. 753–765, 1993.

68. V. Faber and T. Manteuffel, "Necessary and Sufficient Conditions for the Existence of a Conjugate Gradient Method," *SIAM Journal on Numerical Analysis*, vol. 21, no. 2, pp. 352–362, April 1984.
69. R. Fletcher, "Conjugate Gradient Methods for Indefinite Systems," in *Proceedings of the Dundee Conference on Numerical Analysis, 1975, Numerical Analysis (Lecture Notes in Mathematics)*, G. A. Watson, Ed. 1976, vol. 506, pp. 73–89, Springer.
70. G. D. Forney, Jr., "The Viterbi Algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, March 1973.
71. G. D. Forney, Jr., "Convolutional Codes II: Maximum Likelihood Decoding," *Information Control*, vol. 25, no. 3, pp. 222–266, July 1974.
72. L. Fox, H. D. Huskey, and J. H. Wilkinson, "Notes on the Solution of Algebraic Linear Simultaneous Equations," *The Quarterly Journal of Mechanics and Applied Mathematics*, vol. 1, pp. 149–173, 1948.
73. Z. Fu and E. M. Dowling, "Conjugate Gradient Projection Subspace Tracking," *IEEE Transactions on Signal Processing*, vol. 45, no. 6, pp. 1664–1668, June 1997.
74. C. F. Gauss, "Disquisitio de elementis ellipticis Palladis ex oppositionibus annorum 1803, 1804, 1805, 1807, 1808, 1809 (Latin, Disquisition on the Elliptical Elements of Pallas Based on the Oppositions in the Years 1803, 1804, 1805, 1807, 1808, and 1809)," in *Werke (German, Works)*, vol. 6, pp. 1–24. Königliche Gesellschaft der Wissenschaften zu Göttingen (German, The Royal Society of Sciences in Göttingen), 1874, Originally published in 1810.
75. C. F. Gauss, "Theoria interpolationis methodo nova tractata (Latin, Interpolation Theory Based on a Newly Derived Method)," in *Werke (German, Works)*, vol. 3, pp. 256–327. Königliche Gesellschaft der Wissenschaften zu Göttingen (German, The Royal Society of Sciences in Göttingen), 1876.
76. C. F. Gauss, "Gauss an Gerling (German, Gauss to Gerlin). Göttingen, December 26, 1823," in *Werke (German, Works)*, vol. 9, pp. 278–281. Königliche Gesellschaft der Wissenschaften zu Göttingen (German, The Royal Society of Sciences in Göttingen), B. G. Teubner, 1903.
77. W. Gautschi, "On Generating Orthogonal Polynomials," *SIAM Journal on Scientific and Statistical Computing*, vol. 3, no. 3, pp. 289–317, 1982.
78. A. Gersho and T. L. Lim, "Adaptive Cancellation of Intersymbol Interference for Data Transmission," *Bell Systems Technical Journal*, vol. 60, no. 11, pp. 1997–2021, Nov. 1981.
79. T. R. Giallorenzi and S. G. Wilson, "Multiuser ML Sequence Estimator for Convolutionally Coded Asynchronous DS-CDMA Systems," *IEEE Transactions on Communications*, vol. 44, no. 8, pp. 997–1008, August 1996.
80. T. R. Giallorenzi and S. G. Wilson, "Suboptimum Multiuser Receivers for Convolutionally Coded Asynchronous DS-CDMA Systems," *IEEE Transactions on Communications*, vol. 44, no. 9, pp. 1183–1196, September 1996.
81. K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, Jr., and C. E. Wheatley, III, "On the Capacity of a Cellular CDMA System," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 2, pp. 303–312, May 1991.
82. A. Glavieux, C. Laot, and J. Labat, "Turbo Equalization over a Frequency Selective Channel," in *Proceedings of the International Symposium on Turbo Codes and Related Topics*, September 1997, pp. 96–102.

83. J. S. Goldstein, J. R. Guerci, and I. S. Reed, "An Optimal Generalized Theory of Signal Representation," in *Proceedings of the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '99)*, March 1999, vol. 3, pp. 1357–1360.
84. J. S. Goldstein and I. S. Reed, "A New Method of Wiener Filtering and Its Application to Interference Mitigation for Communications," in *Proceedings of the IEEE Military Communications Conference (MILCOM 1997)*, November 1997, vol. 3, pp. 1087–1091.
85. J. S. Goldstein and I. S. Reed, "Reduced-Rank Adaptive Filtering," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 492–496, February 1997.
86. J. S. Goldstein and I. S. Reed, "Subspace Selection for Partially Adaptive Sensor Array Processing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 2, pp. 539–543, April 1997.
87. J. S. Goldstein and I. S. Reed, "Theory of Partially Adaptive Radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 4, pp. 1309–1325, October 1997.
88. J. S. Goldstein, I. S. Reed, D. E. Dudgeon, and J. R. Guerci, "A Multistage Matrix Wiener Filter for Subspace Detection," in *Proceedings of the IT Workshop on Detection, Estimation, Classification, and Imaging*, February 1999.
89. J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A Multistage Representation of the Wiener Filter Based on Orthogonal Projections," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2943–2959, November 1998.
90. J. S. Goldstein, I. S. Reed, L. L. Scharf, and J. A. Tague, "A Low-Complexity Implementation of Adaptive Wiener Filters," in *Proceedings of the 31st Asilomar Conference on Signals, Systems & Computers*, November 1997, vol. 1, pp. 770–774.
91. J. S. Goldstein, I. S. Reed, and P. A. Zulch, "Multistage Partially Adaptive STAP CFAR Detection Algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 2, pp. 645–661, April 1999.
92. G. H. Golub and D. P. O'Leary, "Some History of the Conjugate Gradient and Lanczos Algorithms: 1948–1976," *SIAM Review*, vol. 31, no. 1, pp. 50–102, March 1989.
93. G. H. Golub and R. Underwood, "The Block Lanczos Method for Computing Eigenvalues," in *Mathematical Software III*, J. R. Rice, Ed., pp. 361–377. Academic Press, 1977.
94. G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, 1996.
95. A. Gorokhov and M. van Dijk, "Optimized Labeling Maps for Bit-Interleaved Transmission with Turbo Demodulation," in *Proceedings of the 2001 IEEE 53rd Vehicular Technology Conference (VTC 2001-Spring)*, May 2001, vol. 2, pp. 1459–1463.
96. A. Greenbaum and Z. Strakoš, "Predicting the Behavior of Finite Precision Lanczos and Conjugate Gradient Computations," *SIAM Journal on Matrix Analysis and Applications*, vol. 13, no. 1, pp. 121–137, January 1992.
97. T. N. E. Greville, "The Pseudoinverse of a Rectangular or Singular Matrix and Its Application to the Solution of Systems of Linear Equations," *SIAM Review*, vol. 1, no. 1, pp. 38–43, January 1959.
98. L. J. Griffiths and C. W. Jim, "An Alternative Approach to Linearly Constrained Adaptive Beamforming," *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 1, pp. 27–34, January 1982.

99. C. W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Research Notes in Mathematics; 105. Pitman, 1984.
100. I. Groh and G. Dietl, “Reduced-Rank MIMO Receivers Based on Krylov Subspaces and Random Matrices,” Technical Report TUM-LNS-TR-05-01, Munich University of Technology, April 2005.
101. I. Groh, G. Dietl, and W. Utschick, “Preconditioned and Rank-Flexible Block Conjugate Gradient Implementations of MIMO Wiener Decision Feedback Equalizers,” in *Proceedings of the 2006 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’06)*, May 2006, vol. 4, pp. 441–444.
102. J. R. Guerci, J. S. Goldstein, and I. S. Reed, “Optimal and Adaptive Reduced-Rank STAP,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 2, pp. 647–663, April 2000.
103. D. Guo, S. Shamai (Shitz), and Sergio Verdú, “Mutual Information and Minimum Mean-Square Error in Gaussian Channels,” *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, April 2005.
104. J. Hadamard, *Lectures on Cauchy’s Problem in Linear Partial Differential Equations*, Yale University Press, 1923.
105. J. Hagenauer, “Forward Error Correcting for CDMA Systems,” in *Proceedings of the 4th International Symposium on Spread Spectrum Techniques and Applications (ISSSTA 1996)*, September 1996, vol. 2, pp. 566–569.
106. J. Hagenauer, “The Turbo Principle: Tutorial Introduction and State of the Art,” in *Proceedings of the 2nd International Symposium on Turbo Codes & Related Topics*, September 1997, pp. 1–11.
107. J. Hagenauer, E. Offer, and L. Papke, “Iterative Decoding of Binary Block and Convolutional Codes,” *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 429–445, March 1996.
108. M. Hanke, *Conjugate Gradient Type Methods for Ill-Posed Problems*, Longman, 1995.
109. E. J. Hannan, “Measurement of a Wandering Signal Amid Noise,” *Journal of Applied Probability*, vol. 4, no. 1, pp. 90–102, April 1967.
110. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, 1998.
111. S. Haykin and M. Moher, *Modern Wireless Communications*, Prentice Hall, 2005.
112. S. S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 1986.
113. C. Heegard and S. B. Wicker, *Turbo Coding*, Kluwer, 1999.
114. M. T. Heideman, D. H. Johnson, and C. S. Burrus, “Gauss and the History of the Fast Fourier Transform,” *IEEE Acoustics, Speech, and Signal Processing (ASSP) Magazine*, vol. 1, no. 4, pp. 14–21, October 1984.
115. M. R. Hestenes, “Pseudoinverses and Conjugate Gradients,” *Communications of the ACM*, vol. 18, no. 1, pp. 40–43, 1975.
116. M. R. Hestenes, “Conjugacy and Gradients,” in *A History of Scientific Computing*, S. G. Nash, Ed., pp. 167–179. ACM Press, Addison-Wesley, 1990.
117. M. R. Hestenes and E. Stiefel, “Methods of Conjugate Gradients for Solving Linear Systems,” *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, December 1952.
118. H. Heuser and H. Wolf, *Algebra, Funktionalanalysis und Codierung (German, Algebra, Functional Analysis, and Coding)*, B. G. Teubner, 1986.

119. J. D. Hiemstra and J. S. Goldstein, “Robust Rank Selection for the Multistage Wiener Filter,” in *Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’02)*, May 2002, vol. 3, pp. 2929–2932.
120. M. L. Honig and J. S. Goldstein, “Adaptive Reduced-Rank Residual Correlation Algorithms for DS-CDMA Interference Suppression,” in *Proceedings of the 32nd Asilomar Conference on Signals, Systems & Computers*, November 1998, vol. 2, pp. 1106–1110.
121. M. L. Honig and J. S. Goldstein, “Adaptive Reduced-Rank Interference Suppression Based on the Multistage Wiener Filter,” *IEEE Transactions on Communications*, vol. 50, no. 6, pp. 986–994, June 2002.
122. M. L. Honig and R. Ratasuk, “Large-System Performance of Iterative Multiuser Decision-Feedback Detection,” *IEEE Transactions on Communications*, vol. 51, no. 8, pp. 1368–1377, August 2003.
123. M. L. Honig and Y. Sun, “Performance of Iterative Multiuser Decision-Feedback Receivers,” in *Proceedings of the 2001 IEEE Information Theory Workshop (ITW 2001)*, September 2001, pp. 30–32.
124. M. L. Honig, V. Tripathi, and Y. Sun, “Adaptive Decision Feedback Turbo Equalization,” in *Proceedings of the International Symposium on Information Theory (ISIT 2002)*, June/July 2002, p. 413.
125. M. L. Honig and M. K. Tsatsanis, “Adaptive Techniques for Multiuser CDMA Receivers,” *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 49–61, May 2000.
126. M. L. Honig, G. Woodward, and P. D. Alexander, “Adaptive Multiuser Parallel-Decision-Feedback with Iterative Decoding,” in *Proceedings of the International Symposium on Information Theory (ISIT 2000)*, June 2000, p. 335.
127. M. L. Honig, G. K. Woodward, and Y. Sun, “Adaptive Iterative Multiuser Decision Feedback Detection,” *IEEE Transactions on Wireless Communications*, vol. 3, no. 2, pp. 477–485, March 2004.
128. M. L. Honig and W. Xiao, “Large System Performance of Reduced-Rank Linear Filters,” in *Proceedings of the 37th Annual Allerton Conference on Communications, Systems, and Computing*, October 1999.
129. M. L. Honig and W. Xiao, “Adaptive Reduced-Rank Interference Suppression with Adaptive Rank Selection,” in *Proceedings of the IEEE Military Communications Conference (MILCOM 2000)*, October 2000, vol. 2, pp. 747–751.
130. M. L. Honig and W. Xiao, “Performance of Reduced-Rank Linear Interference Suppression,” *IEEE Transactions on Information Theory*, vol. 47, no. 5, pp. 1928–1946, July 2001.
131. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1993.
132. H. Hotelling, “Analysis of a Complex of Statistical Variables into Principal Components,” *Journal of Educational Psychology*, vol. 24, no. 6/7, pp. 417–441, 498–520, September/October 1933.
133. A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Dover, 1964.
134. C. Hu and I. S. Reed, “Space-Time Adaptive Reduced-Rank Multistage Wiener Filtering for Asynchronous DS-CDMA,” *IEEE Transactions on Signal Processing*, vol. 52, no. 7, pp. 1862–1877, July 2004.

135. Y. Hua, M. Nikpur, and P. Stoica, "Optimal Reduced-Rank Estimation and Filtering," *IEEE Transactions on Signal Processing*, vol. 49, no. 3, pp. 457–469, March 2001.
136. R. Hunger, "Floating Point Operations in Matrix-Vector Calculus," Technical Report TUM-LNS-TR-05-05, Munich University of Technology, 2005.
137. B. R. Hunt, "Deconvolution of Linear Systems by Constrained Regression and Its Relationship to the Wiener Theory," *IEEE Transactions on Automatic Control*, vol. 17, no. 5, pp. 703–705, October 1972.
138. C. G. J. Jacobi, "Ueber [sic] eine neue Auflösungsart der bei der Methode der kleinsten Quadrate vorkommenden lineären [sic] Gleichungen (German, On a New Method for Solving Linear Equations Occuring at the Least Squares Method)," *Astronomische Nachrichten (German, Astronomical News)*, vol. 22, no. 523, pp. 297–306, 1845.
139. C. G. J. Jacobi, "Über ein leichtes Verfahren die in der Theorie der Säcularstörungen [sic] vorkommenden Gleichungen numerisch aufzulösen (German, On a Simple Approach of Solving Numerically the Equations Occuring at the Theory of Secular Perturbance)," *Journal für die reine und angewandte Mathematik (German, Journal for Pure and Applied Mathematics, Crelle's Journal)*, vol. 30, pp. 51–94, 1846.
140. M. Joham, Y. Sun, M. D. Zoltowski, M. Honig, and J. S. Goldstein, "A New Backward Recursion for the Multi-Stage Nested Wiener Filter Employing Krylov Subspace Methods," in *Proceedings of the IEEE Military Communications Conference (MILCOM 2001)*, October 2001, vol. 2, pp. 28–31.
141. M. Joham and M. D. Zoltowski, "Interpretation of the Multi-Stage Nested Wiener Filter in the Krylov Subspace Framework," Technical Report TUM-LNS-TR-00-6, Munich University of Technology, November 2000, Also: Technical Report TR-ECE-00-51, Purdue University.
142. W. Jordan, *Handbuch der Vermessungskunde (German, Handbook of Surveying)*, Verlag der J. B. Metzler'schen Buchhandlung, 1888.
143. T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, 1999.
144. S. Kangshen, J. N. Crossley, and A. W.-C. Lun, *The Nine Chapters on the Mathematical Art*, chapter 8: Rectangular Arrays, pp. 386–438, Oxford University Press, 1999.
145. S. Kaniel, "Estimates for Some Computational Techniques in Linear Algebra," *Mathematics of Computation*, vol. 22, pp. 369–378, 1966.
146. A. Kansal, S. N. Batalama, and D. A. Pados, "Adaptive Maximum SINR RAKE Filtering for DS-CDMA Multipath Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 9, December 1998.
147. K. Karhunen, "Zur Spektraltheorie stochastischer Prozesse (German, On the Spectral Theory of Stochastic Processes)," *Suomalaisen Tiedeakatemian Toimituksia, Annales Academiæ Scientiarum Fennicæ. Serja A, I. Mathematica-Physica (Finish, Latin, Annals of the Finish Academy of Sciences. Series A, I. Mathematics-Physics)*, vol. 34, pp. 1–7, 1946.
148. P. Karttunen and R. Baghaie, "Conjugate Gradient Based Signal Subspace Mobile User Tracking," in *Proceedings of the 1999 IEEE 49th Vehicular Technology Conference (VTC 1999-Spring)*, May 1999, pp. 1172–1176.
149. W. Karush, "Convergence of a Method of Solving Linear Problems," *Proceedings of the American Mathematical Society*, vol. 3, no. 6, pp. 839–851, December 1952.

150. S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, 1993.
151. S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Prentice Hall, 1998.
152. A. N. Kolmogorov, "Sur l'interpolation et extrapolation des suites stationnaires (French, On the Interpolation and Extrapolation of Stationary Sequences)," *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences (French, Proceedings of the Weekly Sessions of the Academy of Sciences)*, vol. 208, pp. 2043–2045, 1939.
153. A. N. Kolmogorov, "Интерполирование и экстраполирование стационарных случайных последовательностей (Russian, Interpolation and Extrapolation of Stationary Random Sequences)," Известия академии наук СССР, Серия математическая *(Russian, Bulletin of the Academy of Sciences of the USSR, Mathematical Series)*, vol. 5, pp. 3–14, 1941.
154. T. P. Krauss, M. D. Zoltowski, and G. Leus, "Simple MMSE Equalizers for CDMA Downlink to Restore Chip Sequence: Comparison to Zero-Forcing and RAKE," in *Proceedings of the 2006 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '00)*, June 2000, vol. 5, pp. 2865–2868.
155. H. Krim and M. Viberg, "Two Decades of Array Signal Processing Research," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, July 1996.
156. C. Lanczos, "An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators," *Journal of Research of the National Bureau of Standards*, vol. 45, no. 4, pp. 255–282, October 1950.
157. C. Lanczos, "Solution of Systems of Linear Equations by Minimized Iterations," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 1, pp. 33–53, July 1952.
158. C. Laot, R. Le Bidan, and D. Leroux, "Low-Complexity MMSE Turbo Equalization: A Possible Solution for EDGE," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 965–974, May 2005.
159. C. Laot, A. Glavieux, and J. Labat, "Turbo Equalization: Adaptive Equalization and Channel Decoding Jointly Optimized," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 9, pp. 1744–1752, September 2001.
160. K. Li and X. Wang, "EXIT Chart Analysis of Turbo Multiuser Detection," *IEEE Transactions on Wireless Communications*, vol. 4, no. 1, pp. 300–311, January 2005.
161. L. Li, A. M. Tulino, and S. Verdú, "Design of Reduced-Rank MMSE Multiuser Detectors Using Random Matrix Methods," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 986–1008, June 2004.
162. P. Li, W. Utschick, and J. A. Nossek, "A New DS-CDMA Interference Cancellation Scheme for Space-Time Rake Processing," in *Proceedings of the European Wireless 1999 & ITG Mobile Communications*, October 1999, pp. 385–389.
163. S. Lin and D. J. Costello, Jr., *Error Control Coding*, Prentice Hall, 2004.
164. M. Loève, "Fonctions aléatoires du second ordre (French, Random Functions of Second Order)," in *Processus stochastiques et mouvement brownien (French, Stochastic Processes and the Brownian Motion)*, Paul Lévy, Ed., pp. 299–352. Gauthier-Villars, 1948, Partly published in *Comptes Rendus de l'Académie des Sciences (French, Proceedings of the Academy of Sciences)*, vol. 220, pp. 295,380–382, 1945, vol. 222, pp. 469,628,942, 1946.

165. D. G. Luenberger, *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, 1965.
166. M. Moher, "An Iterative Multiuser Decoder for Near-Capacity Communications," *IEEE Transactions on Communications*, vol. 46, no. 7, pp. 870–880, July 1998.
167. E. H. Moore, "On the Reciprocal of the General Algebraic Matrix," *Bulletin of the American Mathematical Society*, vol. 26, pp. 394–395, 1920.
168. E. H. Moore and R. W. Barnard, *General Analysis*, vol. I of *Memoirs of the American Philosophical Society*, American Philosophical Society, 1935.
169. S. Moshavi, E. G. Kanterakis, and D. L. Schilling, "Multistage Linear Receivers for DS-CDMA Systems," *International Journal of Wireless Information Networks*, vol. 3, no. 1, pp. 1–17, 1996.
170. M. S. Mueller and J. Salz, "A Unified Theory of Data-Aided Equalization," *Bell Systems Technical Journal*, vol. 60, no. 9, pp. 2023–2038, Nov. 1981.
171. R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley, 1982.
172. R. R. Müller and S. Verdú, "Design and Analysis of Low-Complexity Interference Mitigation on Vector Channels," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 8, pp. 1429–1441, August 2001.
173. W. L. Myrick, M. D. Zoltowski, and J. S. Goldstein, "Anti-Jam Space-Time Preprocessor for GPS Based on Multistage Nested Wiener Filter," in *Proceedings of the IEEE Military Communications Conference (MILCOM 1999)*, October/November 1999, vol. 1, pp. 675–681.
174. W. L. Myrick, M. D. Zoltowski, and J. S. Goldstein, "GPS Jammer Suppression with Low-Sample Support Using Reduced-Rank Power Minimization," in *Proceedings of the 10th IEEE Workshop on Statistical Signal and Array Processing (SSAP 2000)*, August 2000, pp. 514–518.
175. D. P. O'Leary, "The Block Conjugate Gradient Algorithm and Related Methods," *Linear Algebra and Its Applications*, vol. 29, pp. 293–322, 1980.
176. H. Omori, T. Asai, and T. Matsumoto, "A Matched Filter Approximation for SC/MMSE Iterative Equalizers," *IEEE Communications Letters*, vol. 5, no. 7, pp. 310–312, July 2001.
177. R. Otnes and M. Tüchler, "EXIT Chart Analysis Applied to Adaptive Turbo Equalization," in *Proceedings of the Nordic Signal Processing Symposium*, October 2002.
178. D. A. Pados and S. N. Batalama, "Low-Complexity Blind Detection of DS/CDMA Signals: Auxiliary-Vector Receivers," *IEEE Transactions on Communications*, vol. 45, no. 12, pp. 1586–1594, December 1997.
179. D. A. Pados and S. N. Batalama, "Joint Space-Time Auxiliary-Vector Filtering for DS/CDMA Systems with Antenna Arrays," *IEEE Transactions on Communications*, vol. 47, no. 9, pp. 1406–1415, September 1999.
180. D. A. Pados and G. N. Karystinos, "An Iterative Algorithm for the Computation of the MVDR Filter," *IEEE Transactions on Signal Processing*, vol. 49, no. 2, pp. 290–300, February 2001.
181. D. A. Pados, F. J. Lombardo, and S. N. Batalama, "Auxiliary-Vector Filters and Adaptive Steering for DS/CDMA Single-User Detection," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 6, November 1999.
182. C. C. Paige and M. A. Saunders, "LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares," *ACM Transactions on Mathematical Software*, vol. 8, no. 1, pp. 43–71, March 1982.

183. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1991.
184. R. Penrose, “A Generalized Inverse for Matrices,” *Mathematical Proceedings of the Philosophical Society*, vol. 51, pp. 406–413, 1955.
185. B. Picinbono and P. Chevalier, “Widely Linear Estimation with Complex Data,” *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, August 1995.
186. R. L. Pickholtz, D. L. Schilling, and L. B. Milstein, “Theory of Spread-Spectrum Communications – A Tutorial,” *IEEE Transactions on Communications*, vol. COM-30, no. 5, pp. 855–884, May 1982.
187. R. L. Plackett, “Some Theorems in Least Squares,” *Biometrika*, vol. 37, no. 1/2, pp. 149–157, June 1950.
188. H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer, 1988.
189. H. V. Poor and S. Verdú, “Probability of Error in MMSE Multiuser Detection,” *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 858–871, May 1997.
190. B. Porat and B. Friedlander, “Blind Equalization of Digital Communications Channels Using High-Order Moments,” *IEEE Transactions on Signal Processing*, vol. 39, pp. 522–526, February 1991.
191. R. Prasad and T. Ojanperä, “An Overview of CDMA Evolution Towards Wideband CDMA,” *IEEE Communications Surveys & Tutorials*, vol. 1, no. 1, pp. 2–29, 1998, http://www.comsoc.org/pubs/surveys.
192. R. Price and P. E. Green, Jr., “A Communication Technique for Multipath Channels,” *Proceedings of the IRE*, vol. 46, no. 3, pp. 555–570, March 1958.
193. J. G. Proakis, *Digital Communications*, McGraw-Hill, 1995.
194. H. Qian and S. N. Batalama, “Data Record-Based Criteria for the Selection of an Auxiliary Vector Estimator of the MMSE/MVDR Filter,” *IEEE Transactions on Communications*, vol. 51, no. 10, pp. 1700–1708, October 2003.
195. V. Ramon and L. Vandendorpe, “Predicting the Performance and Convergence Behavior of a Turbo-Equalization Scheme,” in *Proceedings of the 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’05)*, March 2005.
196. M. C. Reed, C. B. Schlegel, P. D. Alexander, and J. A. Asenstorfer, “Iterative Multiuser Detection for CDMA with FEC: Near-Single-User Performance,” *IEEE Transactions on Communications*, vol. 46, no. 12, pp. 1693–1699, December 1998.
197. J. K. Reid, “The Use of Conjugate Gradients for Systems of Equations Processing ‘Property A’,” *SIAM Journal on Numerical Analysis*, vol. 9, no. 2, pp. 325–332, June 1972.
198. D. Reynolds and X. Wang, “Low-Complexity Turbo-Equalization for Diversity Channels,” *Signal Processing*, vol. 81, no. 5, pp. 989–995, 2001.
199. D. C. Ricks and J. S. Goldstein, “Efficient Architectures for Implementing Adaptive Algorithms,” in *Proceedings of the 2000 Antenna Applications Symposium*, September 2000, pp. 29–41.
200. A. Ruhe, “Implementation Aspects of Band Lanczos Algorithms for Computation of Eigenvalues of Large Sparse Symmetric Matrices,” *Mathematics of Computation*, vol. 33, no. 146, pp. 680–687, April 1979.

201. Y. Saad, “On the Lanczos Method for Solving Symmetric Linear Systems with Several Right-Hand Sides,” *Mathematics of Computation*, vol. 48, no. 178, pp. 651–662, April 1987.
202. Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, Halstead Press – *out of print*, 1992.
203. Y. Saad, “Analysis of Augmented Krylov Subspace Methods,” *SIAM Journal on Matrix Analysis and Applications*, vol. 18, no. 2, pp. 435–449, April 1997.
204. Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, 2003.
205. Y. Saad and M. H. Schultz, “GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems,” *SIAM Journal on Scientific and Statistical Computing*, vol. 7, no. 3, pp. 856–869, July 1986.
206. Y. Saad and H. A. van der Vorst, “Iterative Solution of Linear Systems in the 20th Century,” *Journal of Computational and Applied Mathematics*, vol. 123, no. 1–2, pp. 1–33, November 2000.
207. M. Sadkane, “Block-Arnoldi and Davidson Methods for Unsymmetric Large Eigenvalue Problems,” *Numerische Mathematik (German, Numerical Mathematics)*, vol. 64, pp. 195–212, 1993.
208. M. Sadkane, “A Block Arnoldi-Chebyshev Method for Computing the Leading Eigenpairs of Large Sparse Unsymmetric Matrices,” *Numerische Mathematik (German, Numerical Mathematics)*, vol. 64, pp. 181–193, 1993.
209. E. L. Santos and M. D. Zoltowski, “On Low Rank MVDR Beamforming Using the Conjugate Gradient Algorithm,” in *Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’04)*, May 2004, vol. 2, pp. 173–176.
210. L. L. Scharf, *Statistical Signal Processing*, Addison-Wesley, 1991.
211. L. L. Scharf, “The SVD and Reduced Rank Signal Processing,” *Signal Processing*, vol. 25, no. 2, pp. 113–133, 1991.
212. L. L. Scharf, L. T. McWhorter, E. K. P. Chong, J. S. Goldstein, and M. D. Zoltowski, “Algebraic Equivalence of Conjugate Gradient Direction and Multistage Wiener Filters,” in *Proceedings of the 11th Annual Workshop on Adaptive Sensor Array Processing (ASAP)*, March 2003, pp. 388–392.
213. L. L. Scharf and D. W. Tufts, “Rank Reduction for Modeling Stationary Signals,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no. 3, pp. 350–355, March 1987.
214. R. O. Schmidt, “Multiple Emitter Location and Signal Parameter Estimation,” *IEEE Transactions on Antennas and Propagation*, vol. AP-34, no. 3, pp. 276–280, March 1986, Originally published in *Proceedings of the Rome Air Development Center (RADC) Spectrum Estimation Workshop*, October 1979.
215. M. K. Schneider and A. S. Willsky, “Krylov Subspace Estimation,” *SIAM Journal on Scientific Computing*, vol. 22, no. 5, pp. 1840–1864, 2001.
216. R. A. Scholtz, “The Origins of Spread-Spectrum Communications,” *IEEE Transactions on Communications*, vol. COM-30, no. 5, pp. 822–854, May 1982.
217. F. Schreckenbach, N. Görtz, J. Hagenauer, and G. Bauch, “Optimization of Symbol Mappings for Bit-Interleaved Coded Modulation with Iterative Decoding,” *IEEE Communications Letters*, vol. 7, no. 12, pp. 593–595, December 2003.
218. H. R. Schwarz, H. Rutishauser, and E. Stiefel, *Numerik symmetrischer Matrizen (German, Numerics of Symmetric Matrices*, B. G. Teubner, 1968.

219. J. A. Scott, “An Arnoldi Code for Computing Selected Eigenvalues of Sparse, Real, Unsymmetric Matrices,” *ACM Transactions on Mathematical Software*, vol. 21, no. 4, pp. 432–475, December 1995.
220. L. Seidel, “Ueber [sic] ein Verfahren, die Gleichungen, auf welche die Methode der kleinsten Quadrate führt, sowie lineäre [sic] Gleichungen überhaupt, durch successive [sic] Annäherung aufzulösen (German, On an Approach for Solving Equations Obtained From the Least Squares Method or Linear Equations in General by Successive Approximation),” *Abhandlungen der mathematisch-physikalischen Classe [sic] der königlich bayerischen Akademie der Wissenschaften (German, Essay of the Mathematical-Physical Class of the Royal Bavarian Academy of Sciences)*, vol. 11, pp. 81–108, 1874.
221. G. M. A. Sessler and F. K. Jondral, “Low Complexity Polynomial Expansion Multiuser Detector with Strong Near-Far Resistance,” in *Proceedings of the 2002 IEEE 56th Vehicular Technology Conference (VTC 2002-Fall)*, September 2002, vol. 3, pp. 1869–1872.
222. G. M. A. Sessler and F. K. Jondral, “Low Complexity Polynomial Expansion Multiuser Detector for CDMA Systems,” *IEEE Transactions on Vehicular Technology*, vol. 54, no. 4, pp. 1379–1391, July 2005.
223. V. Simoncini and E. Gallopoulos, “An Iterative Method for Nonsymmetric Systems with Multiple Right-Hand Sides,” *SIAM Journal on Scientific Computing*, vol. 16, no. 4, pp. 917–933, July 1995.
224. V. Simoncini and E. Gallopoulos, “Convergence Properties of Block GMRES and Matrix Polynomials,” *Linear Algebra and Its Applications*, vol. 247, pp. 97–119, 1996.
225. A. van der Sluis and H. A. van der Vorst, “The Rate of Convergence of Conjugate Gradients,” *Numerische Mathematik (German, Numerical Mathematics)*, vol. 48, pp. 543–560, 1986.
226. H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, Prentice Hall, 1986.
227. Z. Strakoš, “On the Real Convergence Rate of the Conjugate Gradient Method,” *Linear Algebra and Its Applications*, vol. 154–156, pp. 535–549, 1991.
228. G. Strang, “A Proposal for Toeplitz Matrix Calculations,” *Studies in Applied Mathematics*, vol. 74, no. 2, pp. 171–176, 1986.
229. S. S. Sud and J. S. Goldstein, “A Low Complexity Receiver for Space-Time Coded CDMA Systems,” in *Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’02)*, May 2002, vol. 3, pp. 2397–2400.
230. Y. Sun and M. L. Honig, “Performance of Reduced-Rank Equalization,” in *Proceedings of the International Symposium on Information Theory (ISIT 2002)*, June/July 2002.
231. Y. Sun, M. L. Honig, and V. Tripathi, “Adaptive, Iterative, Reduced-Rank Equalization for MIMO Channels,” in *Proceedings of the IEEE Military Communications Conference (MILCOM 2002)*, 2002, vol. 2, pp. 1029–1033.
232. Y. Sun, V. Tripathi, and M. L. Honig, “Adaptive Turbo Reduced-Rank Equalization for MIMO Channels,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 2789–2800, November 2005.
233. A. Thorpe and L. Scharf, “Reduced Rank Methods for Solving Least Squares Problems,” in *Proceedings of the 23rd Asilomar Conference on Signals, Systems & Computers*, October 1989, vol. 2, pp. 609–613.

234. A. N. Tikhonov, "Solution of Incorrectly Formulated Problems and the Regularization Method," *Soviet Mathematics Doklady*, vol. 4, pp. 1035–1038, 1963.
235. A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, Winston, 1977.
236. J. Todd, "The Prehistory and Early History of Computation at the U. S. National Bureau of Standards," in *A History of Scientific Computing*, S. G. Nash, Ed., pp. 251–268. ACM Press, Addison-Wesley, 1990.
237. L. N. Trefethen and D. Bau, III, *Numerical Linear Algebra*, SIAM, 1997.
238. W. F. Trench, "An Algorithm for the Inversion of Finite Toeplitz Matrices," *Journal of the Society for Industrial and Applied Mathematics*, vol. 12, no. 3, pp. 515–522, September 1964.
239. M. Tüchler, R. Koetter, and A. Singer, "Turbo Equalization: Principles and New Results," *IEEE Transactions on Communications*, vol. 50, no. 5, pp. 754–767, May 2002.
240. M. Tüchler, A. Singer, and R. Koetter, "Minimum Mean Squared Error Equalization Using A-priori Information," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 673–683, March 2002.
241. D. W. Tufts and R. Kumaresan, "Singular Value Decomposition and Improved Frequency Estimation Using Linear Prediction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-30, no. 4, pp. 671–675, August 1982.
242. D. W. Tufts, R. Kumaresan, and I. Kirsteins, "Data Adaptive Signal Estimation by Singular Value Decomposition of a Data Matrix," *Proceedings of the IEEE*, vol. 70, no. 6, pp. 684–685, June 1982.
243. G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals," *IEEE Transactions on Information Theory*, vol. 28, no. 1, pp. 55–67, January 1982.
244. H. L. Van Trees, *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*, Wiley, 2002.
245. M. K. Varanasi and T. Guess, "Optimum Decision Feedback Multiuser Equalization with Successive Decoding Achieves the Total Capacity of the Gaussian Multiple-Access Channel," in *Proceedings of the 31st Asilomar Conference on Signals, Systems & Computers*, November 1997, vol. 2, pp. 1405–1409.
246. S. Verdú, "Minimum Probability of Error for Asynchronous Gaussian Multiple-Access Channels," *IEEE Transactions on Information Theory*, vol. IT-32, no. 1, pp. 85–96, January 1986.
247. S. Verdú, *Multiuser Detection*, Cambridge University Press, 1998.
248. A. J. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," *IEEE Transactions on Information Theory*, vol. IT-13, no. 2, pp. 260–269, April 1967.
249. V. V. Voevodin, "The Question of Non-Self-Adjoint Extension of the Conjugate Gradients Method is Closed," *USSR Computational Mathematics and Mathematical Physics*, vol. 23, no. 2, pp. 143–144, 1983.
250. D. V. Voiculescu, K. J. Dykema, and A. Nica, *Free Random Variables*, American Mathematical Society, 1992.
251. H. A. van der Vorst, "Bi-CGSTAB: A Fast and Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmmetric Linear Systems," *SIAM Journal on Scientific and Statistical Computing*, vol. 13, no. 2, pp. 631–644, March 1992.
252. H. A. van der Vorst, "Krylov Subspace Iteration," *Computing in Science & Engineering*, vol. 2, no. 1, pp. 32–37, January/February 2000.

253. H. F. Walker, "Implementation of the GMRES Method Using Householder Transformations," *SIAM Journal on Scientific Computing*, vol. 9, no. 1, pp. 152–163, 1988.
254. X. Wang and H. V. Poor, "Iterative (Turbo) Soft Interference Cancellation and Decoding for Coded CDMA," *IEEE Transactions on Communications*, vol. 47, no. 7, pp. 1046–1061, July 1999.
255. G. A. Watson, "An Algorithm for the Inversion of Block Matrices of Toeplitz Form," *Journal of the Association for Computing Machinery*, vol. 20, no. 3, pp. 409–415, July 1973.
256. M. E. Weippert, J. D. Hiemstra, J. S. Goldstein, and M. D. Zoltowski, "Insights from the Relationship Between the Multistage Wiener Filter and the Method of Conjugate Gradients," in *Proceedings of the Sensor Array and Multichannel Signal Processing Workshop*, August 2002, pp. 388–392.
257. B. Widrow and M. E. Hoff, Jr., "Adaptive Switching Circuits," in *IRE WESCON Convention Record, Part 4*, 1960, pp. 96–104.
258. B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode, "Adaptive Antenna Systems," *Proceedings of the IEEE*, vol. 55, no. 12, pp. 2143–2159, December 1967.
259. N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, MIT Press, 1949.
260. N. Wiener and E. Hopf, "Über eine Klasse singulärer Integralgleichungen (German, On a Class of Singular Integral Equations)," *Sitzungsberichte der Königlich Preussischen [sic] Akademie der Wissenschaften zu Berlin (German, Proceedings of the Royal Prussian Academy of Sciences in Berlin)*, pp. 696–706, 1931.
261. G. Woodward and M. L. Honig, "Performance of Adaptive Iterative Multiuser Parallel Decision Feedback with Different Code Rates," in *Proceedings of the 2001 IEEE International Conference on Communications (ICC 2001)*, June 2001, vol. 3, pp. 852–856.
262. G. Woodward, R. Ratasuk, and M. L. Honig, "Multistage Multiuser Decision Feedback Detection for DS-CDMA," in *Proceedings of the 1999 IEEE International Conference on Communications (ICC 1999)*, June 1999, vol. 1, pp. 68–72.
263. G. Woodward, R. Ratasuk, M. L. Honig, and P. B. Rapajic, "Minimum Mean-Squared Error Multiuser Decision-Feedback Detectors for DS-CDMA," *IEEE Transactions on Communications*, vol. 50, no. 12, pp. 2104–2112, December 2002.
264. W. Xiao and M. L. Honig, "Convergence Analysis of Adaptive Full-Rank and Multi-Stage Reduced-Rank Interference Suppression," in *Proceedings of the 2000 Conference on Information Sciences and Systems*, March 2000, vol. 1, pp. WP2–6–WP2–11.
265. W. Xiao and M. L. Honig, "Large System Transient Analysis of Adaptive Least Squares Filtering," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2447–2474, July 2005.
266. X. Yang, T. K. Sarkar, and E. Arvas, "A Survey of Conjugate Gradient Algorithms for Solution of Extreme Eigen-Problems of a Symmetric Matrix," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 10, pp. 1550–1556, October 1989.

267. D. M. Young, “A Historical Review of Iterative Methods,” in *A History of Scientific Computing*, S. G. Nash, Ed., pp. 180–194. ACM Press, Addison-Wesley, 1990.
268. N. Young, *An Introduction to Hilbert Space*, Cambridge University Press, 1988.
269. M. Zoltowski and E. Santos, “Matrix Conjugate Gradients for the Generation of High-Resolution Spectograms,” in *Proceedings of the 37st Asilomar Conference on Signals, Systems & Computers*, November 2003, vol. 2, pp. 1843–1847.
270. M. D. Zoltowski, W. J. Hillery, S. Özen, and M. Fimoff, “Conjugate-Gradient-Based Decision Feedback Equalization with Structured Channel Estimation for Digital Television,” in *Proceedings of SPIE – The International Society for Optical Engineering*, April 2002, vol. 4740.
271. M. D. Zoltowski, T. P. Krauss, and S. Chowdhury, “Chip-Level MMSE Equalization for High-Speed Synchronous CDMA in Frequency Selective Multipath,” in *Proceedings of SPIE – The International Society for Optical Engineering*, April 2000, vol. 4045, pp. 187–197.
272. P. A. Zulch, J. S. Goldstein, J. R. Guerci, and I. S. Reed, “Comparison of Reduced-Rank Signal Processing Techniques,” in *Proceedings of the 32nd Asilomar Conference on Signals, Systems & Computers*, November 1998, vol. 1, pp. 421–425.

Index

2-norm, 8
  definition, 179
$\boldsymbol{A}$-conjugate, 47
  direction matrix, *see* direction matrix
$\boldsymbol{A}$-norm, 8, 46, 47, 54, 55, 59, 190, 191
  definition, 179
$\boldsymbol{A}$-conjugate, 47
vec-operation, 8, 62
  definition, 183
  properties, 184

a posteriori, *see* LLR
a priori, *see* LLR
additive white Gaussian noise, *see* AWGN
adjusted, 136
  auto-covariance matrix, 131–133, 139, 146, 147
  mean, 135
  mean vector, 131, 139, 147
  MWF, *see* MWF
  reduced-rank MWF, *see* reduced-rank MWF
  variance, 136
affine filter, 4, 15, 139
Arnoldi, 2, 33
  block, *see* block Arnoldi
  eigenvalue problems, 37, 39
ARQ, 115
automatic repeat request, *see* ARQ
auxiliary vector, *see* AV
AV, 72, 93
  GSC, *see* GSC
  MSWF, *see* MSWF
AWGN, 125, 142
  channel, 114, 118, 149

backward recursion, *see* MSMWF
backward substitution, *see* BS
Bahl–Cocke–Jelinek–Raviv, *see* BCJR
band Hessenberg matrix, 37, 40, 41
  definition, 37
  eigenvalue problems, *see* eigenvalue problems
  Hermitian, 42, 99, 100
Bayes parameter estimation, 1, 14–15, 182
BCD, 47–51, 53
BCG, 2, 5, 35, 46, 73, 133
  algorithm, 54
  computational complexity
    comparison, 107–108
    formula, 106
  derivation, 46–54
  dimension-flexible, 5, 66–68
    algorithm, 68
  filter factors, 5
    derivation, 60–64
    example, 66
    formula, 63, 64
  Krylov subspace, 52–53
  MSMWF connection, *see* MSMWF
  MSMWF implementation, *see* MSMWF
  preconditioning, 60
  rate of convergence, 5
    derivation, 54–59
    superlinear, 59–60

upper bound, 59
regularization, 5, 60–66, 171
BCJR, 118, 138, 143
BER, 7, 114, 119, 121, 141, 152
of MSMWF, *see* MSMWF
of MSVWF, *see* MSVWF
of MWF, *see* MWF
of Turbo receiver, *see* Turbo receiver
of VWF, *see* VWF
relationship to MI, 145, 156
BGMRES, 35, 46
Bi-CG, 34
stabilized, *see* Bi-CGSTAB
Bi-CGSTAB, 34
bi-conjugate gradient, *see* Bi-CG
bi-conjugate gradient stabilized, *see* Bi-CGSTAB
bi-orthogonalization
CG, *see* Bi-CG
Lanczos, *see* Lanczos
bias–variance trade-off, 26
binary phase shift keying, *see* BPSK
bit error rate, *see* BER
bitwise interleaving, *see* interleaver
block Arnoldi, 2, 5
algorithm, 40
block Lanczos connection, *see* block Lanczos
derivation, 38–41
eigenvalue problems, 37
MSMWF connection, *see* MSMWF
system of linear equations, 46
block conjugate direction, *see* BCD
block conjugate gradient, *see* BCG
block generalized minimal residual, *see* BGMRES
block interleaver, *see* interleaver
block Lanczos, 2, 5
algorithm, 43
block Arnoldi connection, 41–42
derivation, 41–43
dimension-flexible, *see* BLR
eigenvalue problems, 37
MSMWF connection, *see* MSMWF
Ruhe, *see* BLR
system of linear equations, 46
block Lanczos–Ruhe, *see* BLR
block Toeplitz, 18
low-complexity, 166
block tridiagonal matrix, 85–87
definition, 85
blocking matrix, 6, 71, 98
definition, 75, 79
dependency on, *see* MSMWF
interpretation, 76
scaled orthonormal, 89, 91, 103
BLR, 5, 98, 133
algorithm, 45
computational complexity
comparison, 107–108
formula, 106
derivation, 43–46
MSMWF implementation, *see* MSMWF
system of linear equations, 46
BPSK, 118, 120, 141, 146
BS, 18, 21
algorithm, 22
computational complexity, 22

CDMA, 71, 72, 113, 114
DS, *see* DS-CDMA
wideband, *see* WCDMA
cdma2000, 113
CF, 2, 4, 14, 18, 24, 187
algorithm, 20
computational complexity, 21
derivation, 18–21
optimal blocking matrices, 88
outer product version, 21
system of linear equations, 105, 189
CG, 2, 5, 33, 55, 72
as direct method, 34
as iterative method, 34
augmented, 35
bi-, *see* Bi-CG
block, *see* BCG
computational complexity, 162
filter factors, 5, 60
derivation, 64–65
example, 65–66
formula, 64
inventors, 34
MSWF connection, *see* MSWF
MSWF implementation, *see* MSWF
rate of convergence
derivation, 56–59
superlinear, 59–60

upper bound, 58
regularization, 60
stabilized bi-, *see* Bi-CGSTAB
vector, *see* VCG
channel coding, *see* encoder
channel encoder, *see* encoder
channel state information, *see* CSI
channel type
AWGN, *see* AWGN
MA-SIMO, *see* MA-SIMO
Porat, *see* Porat
Proakis, *see* Proakis
channel-shortening, 115
Chebyshev polynomial, 57, 59
derivation, 179–180
chip, 115, 121, 122
rate, 123, 144
Cholesky factorization, *see* CF
CME, 1, 14–15
code division multiple-access, *see* CDMA
complex multivariate normal distribution, *see* normal distribution
complex normal distribution, *see* normal distribution
complexity isoset, 163
definition, 163
complexity–performance chart, 7, 141
derivation, 162–163
computational complexity
FLOP, *see* FLOP
of BCG, *see* BCG
of BLR, *see* BLR
of BS, *see* BS
of CF, *see* CF
of CG, *see* CG
of CS, *see* CS
of FS, *see* FS
of HI, *see* HI
of LTI, *see* LTI
of LTP, *see* LTP
of MSMWF, *see* BCG *or* BLR
of MSVWF, *see* CG
of MSWF, *see* MSWF
of MWF, *see* MWF
of PC, *see* PC
of QR factorization, *see* QR factorization
of RCMWF, *see* RCMWF
of VCG, *see* CG
of VCS, *see* CS
of VPC, *see* PC
of VWF, *see* VWF
of WF, *see* WF
condition number, 60, 65
definition, 179
conditional mean estimator, *see* CME
conjugate gradient, *see* CG
convergence rate, *see* rate of convergence
convolutional encoder, 153
derivation, 116–118
generator matrix, 117
generator sequence, 116, 117
memory, 116
correlated subspace, 6, 77, 78, 86, 95, 98
derivation, 195–199
correlation measure, 195, 197, 198
definition, 196
correlation-subtraction architecture, *see* MSWF
correlator, *see* matrix correlator
CS, 3, 5, 71, 132
computational complexity, 31, 107
derivation, 30–31
eigensubspace, 3, 30, 31
EXIT characteristic, 153–155
metric, 30
MI performance, 156
MMSE of, 31
MSE performance, 156–158
prefilter formula, 31
regularization, 31
vector, *see* VCS
CSI, 141
estimated, 165, 171
perfect, 164, 171

decision feedback detector, *see* DFD
decision feedback equalizer, *see* DFE
decoder, 127
derivation, *see* MAP
MAP, *see* MAP
ML, *see* ML
soft-input hard-output, 4
soft-input soft-output, 4, 6
demapper, *see* soft demapper
detector
decision feedback, *see* DFD

ML, *see* ML
MMSE, *see* MMSE
MU, *see* MU
DFD, 73, 114
parallel, 7, 114, 138
successive, 114
DFE, 72, 139
diagonal loading, 171
dimension-flexible block conjugate gradient, *see* BCG
dimension-flexible block Lanczos, *see* BLR
direct sequence, *see* DS
direct sequence code division multiple-access, *see* DS-CDMA
direction matrix, 47, 52, 53, 67, 68, 97
$\boldsymbol{A}$-conjugate, 50, 51, 67
derivation, 47–50
DS, 113
CDMA, *see* DS-CDMA
DS-CDMA, 6, 7, 71–73, 121, 139, 141, 152
receiver, *see* Turbo receiver
spreader, *see* spreader
transmitter, 115

EDGE, 72
effective free distance criterion, 121
eigen decomposition, 14, 29, 37, 60, 62, 64, 182, 184
definition, 29, 56
eigensubspace, 3, 5, 14, 29, 71, 106, 132
method, *see* PC *or* CS
eigenvalue, 3, 29, 30, 59, 60, 63, 65, 66, 182–184, 198
computation, *see* eigenvalue problems
definition, 29, 56, 184
eigenvalue problems, 5, 35, 72
band Hessenberg structure, 37
solving via Arnoldi, *see* Arnoldi
solving via block Arnoldi, *see* block Arnoldi
solving via block Lanczos, *see* block Lanczos
solving via Lanczos, *see* Lanczos
eigenvector, 3, 29–31, 132, 184
computation, 30
definition, 29, 56
encoder, 6, 115, 118
convolutional, *see* convolutional encoder
FEC, *see* FEC
non-systematic, 116
systematic, 116
enhanced data rates for GSM evolution, *see* EDGE
equalizer
decision feedback, *see* DFE
MAP, *see* MAP
MU, *see* MU
Turbo, *see* Turbo receiver
estimator
conditional mean, *see* CME
MMSE, *see* MMSE
Euclidean norm, 4, 8, 179
Euler's formula, 180
EXIT
characteristic, 7, 141, 153
of CG, *see* MSWF
of CS, *see* CS
of MSWF, *see* MSWF
of PC, *see* PC
of WF, *see* WF
semianalytical calculation, 146–152
chart, 7, 141
derivation, 142–145
MSE analysis, *see* MSE
exponential power delay profile, *see* power delay profile
extrapolator, 13
extrinsic, *see* LLR
extrinsic information transfer, *see* EXIT

fast Fourier transform, *see* FFT
FEC, 115, 143
FFT, 14, 18, 60, 115
field, 185
Galois, *see* Galois field
filter factors
definition, 182
of BCC, *see* BCG
of CG, *see* CG
Tikhonov, *see* Tikhonov
finite impulse response, *see* FIR
FIR, 118
floating point operation, *see* FLOP
FLOP, 4, 6
definition, 17–18

forward error correction, *see* FEC
forward substitution, *see* FS
Fredholm integral equation, 181
Frobenius norm, 8
  definition, 178–179
FS, 18, 21
  algorithm, 21
  computational complexity, 21

Galois field, 185
  binary, 116, 185
Gauss–Jordan, 1–2
Gauss–Seidel, 2, 34
Gaussian elimination, 1–2, 34
generalized minimal residual, *see* GMRES
generalized sidelobe canceler, *see* GSC
generator matrix, *see* convolutional encoder
generator sequence, *see* convolutional encoder
global positioning system, *see* GPS
global systems for mobile communications, *see* GSM
GMRES, 2, 34
  block, *see* BGMRES
GPS, 71
Gram–Schmidt
  classical, 38, 39
  conjugate, 49–51
  modified, 39, 42, 191, 192
Grassmann manifold, 26, 196
Gray mapping, *see* mapper
GSC, 71, 77–78
  AV, 72
  MSMWF, 78
  MSWF, 3, 71
GSM, 72

hard decision, 4, 127
Hermitian inversion, *see* HI
Hessenberg matrix, 37, 39, 42, 43
  band, *see* band Hessenberg matrix
  definition, 37
HI, 25, 104, 106
  computational complexity, 189
  derivation, 187–189
Hilbert space, 59
Hilbert–Schmidt norm, *see* Frobenius norm
Householder reflection, 39
hypothesis testing, 3

i. i. d., 142
ill-conditioned problem, 5, 60, 182
  definition, 182
ill-posed problem, 5, 60
  definition, 182
independent and identically distributed, *see* i. i. d.
induced matrix norm, *see* 2-norm
interference
  intersymbol, *see* ISI
  multiple-access, *see* MAI
interference cancellation, 139
  iterative soft, *see* iterative soft interference cancellation
interleaver, 6, 114, 126, 130, 134, 137, 142, 148
  bitwise, 118
  block, 119
  derivation, 118–119
  random, 119
  S-random, 119
  symbolwise, 118
interpolator, 13
intersymbol interference, *see* ISI
inverse problem, 180–182
IS-95, 113
ISI, 113, 114, 122, 128, 154
iterative receiver, *see* Turbo receiver
iterative soft interference cancellation, 7, 114, 138, 139, 141, 152

Jacobi, 2

Karhunen–Loève transform, *see* KLT
KLT, 14, 29, 30
Kronecker delta, 9, 29, 122, 184, 187
Kronecker product, 8, 122, 183
  definition, 183
  properties, 184
Krylov subspace
  BCG, *see* BCG
  definition, 36
  MSMWF, *see* MSMWF
  MSWF, *see* MSWF

polynomial representation, 36–37

Lagrangian function, 197, 199
Lanczos, 2, 33
bi-orthogonalization, 2, 33
block, *see* block Lanczos
eigenvalue problems, 30, 33, 37
MSWF implementation, *see* MSWF
regularization, 60
latency time, 9, 129, 145, 151, 153, 159, 165
least mean squares, *see* LMS
least squares, *see* LS
LLR, 128–130, 137, 139, 141, 142
a posteriori, 128, 137
definition, 134, 137
a priori, 128–130, 139
definition, 135, 137
distribution, 142
joint distribution, 147
random variable, 137
AWGN channel, 142
definition, 9, 127
extrinsic, 134, 136, 150
definition, 135, 137
distribution, 150
meaning, 142
LMS, 71
log-likelihood ratio, *see* LLR
lower triangular inversion, *see* LTI
lower triangular product, *see* LTP
LS, 7, 33, 34, 72, 164, 182
channel estimation, 170–171
recursive, *see* RLS
LSQR, 34
LTI, 24, 187
algorithm, 188
computational complexity, 188
LTP, 187
algorithm, 189
computational complexity, 189
LU factorization, 2

MA, 6, 139
SIMO channel, *see* MA-SIMO
MA-SIMO, 73, 124, 128
MAI, 113, 114, 122, 128, 139, 146, 168
manifold, *see* Grassmann manifold
MAP
decoder, 114, 127, 138
derivation, 137–138
soft-input soft-output, 137, 138
equalizer, 114
mapper, 6, 130, 144, 145
definition, 119
derivation, 119–121
Gray, 120
Turbo, 120–121
QPSK, *see* QPSK
matched filter, 113, 156
matrix, *see* matrix matched filter
matrix correlator, 5, 6
definition, 77–79, 195–196
matrix inversion, 133
lower triangular, *see* LTI
partitioned matrix, 177–178
positive definite, *see* HI
Schur complement, 177
matrix matched filter, 195–196
matrix norm, 59
2-norm, *see* 2-norm
$\boldsymbol{A}$-norm, *see* $\boldsymbol{A}$-norm
definition, 178
Frobenius, *see* Frobenius norm
Hilbert–Schmidt, *see* Frobenius norm
induced, *see* 2-norm
matrix Wiener filter, *see* MWF
maximum a posteriori, *see* MAP
maximum likelihood, *see* ML
mean square error, *see* MSE
memory, *see* convolutional encoder
MI, 141, 145, 153
definition, 142–143
of CS, *see* CS
of MSWF, *see* MSWF
of PC, *see* PC
of WF, *see* WF
relationship to BER, *see* BER
MIMO, 73
minimum mean square error, *see* MMSE
minimum variance distortionless response, *see* MVDR
ML, 113
decoder, 138
detector, 113
MMSE
channel estimation, 171
detector, 113, 114, 135

estimator, 1, 3, 13–15
of CS, *see* CS
of MSMWF, *see* MSMWF
of MWF, *see* MWF
of PC, *see* PC
of reduced-rank MWF, *see* reduced-rank MWF
optimal MSMWF, 90
Monte Carlo
sequential sampling techniques, 115
simulations, 7, 141, 146, 149, 151, 156, 160, 165
Moore–Penrose, *see* pseudoinverse
MSE, 1, 4, 13, 14, 71
BCG connection, 97
definition, 15
EXIT chart analysis, 141, 145, 149, 156–161
minimum, *see* MMSE
optimization, 15
reduced-rank optimization, 27
MSMWF, 3, 5, 73, 132
algorithm, 83
backward recursion, 82
BCG connection, 97–98
BCG implementation, 103, 106, 151
BER performance, 165–171
block Arnoldi connection, 95–96, 107
block Lanczos connection, 95–96
BLR implementation, 105
algorithm, 103
derivation, 98–103
computational complexity, *see* BCG *or* BLR
dependency on blocking matrix, 87–90
derivation, 73–80
fundamental properties, 85–90
general version, 77
GSC, *see* GSC
invariance, 80
Krylov subspace, 91–95
MMSE of, 84
rank-flexible, 83
reduced-rank, 81–85
MSVWF, 159
BER performance, 168–170
computational complexity, *see* CG
MSE performance, 159–161
VCG implementation, 159
MSWF, 3, 71
AV, 72
CG implementation, 151
computational complexity, 163
correlation-subtraction architecture, 71
EXIT characteristic, 153–155
GSC, *see* GSC
Krylov subspace, 3, 71, 93
Lanczos implementation, 72
matrix, *see* MSMWF
MI performance, 156
MSE performance, 156–158
MU detector, 6, 72, 73
iterative, *see* Turbo receiver
MU equalizer, 127
derivation, 128–133
multiple signal classification, *see* MUSIC
multiple-access, *see* MA
multiple-access interference, *see* MAI
multiple-input multiple-output, *see* MIMO
multistage decomposition, *see* MSMWF
multistage matrix Wiener filter, *see* MSMWF
multistage vector Wiener filter, *see* MSVWF
multistage Wiener filter, *see* MSWF
multiuser detector, *see* MU detector
multiuser equalizer, *see* MU equalizer
MUSIC, 14
mutual information, *see* MI
MVDR, 71, 72, 78
MWF, 1, 4–5, 72, 113, 128
adjusted, 131
BER performance, 165–171
Cholesky based implementation, 18–22
computational complexity
comparison, 107–108
formula, 22
derivation, 14–17
for non-stationary random sequences, 17, 130
for non-zero mean random sequences, 16–17, 130
formula, 16

formula for equalization, 129
formula of adjusted, 131
formula of reduced-dimension, 27
matrix inversion lemma, 133
MMSE of, 16
multistage, *see* MSMWF
optimization, 15
reduced-complexity, *see* RCMWF
reduced-dimension, 26, 27
reduced-rank, *see* reduced-rank MWF

non-stationary random sequences, *see* MWF
non-zero mean random sequences, *see* MWF
normal distribution
complex, 89
multivariate, 125
real, 142
normalized matrix correlator, *see* matrix correlator

orthogonal variable spreading factor, *see* OVSF
outer product version, *see* CF
overmodeled problem, 26
OVSF, 122, 139, 165

parallel DFD, *see* DFD
parallel-to-serial conversion, 117, 118
PC, 3, 5, 14, 71, 132
computational complexity, 30, 106
derivation, 29–30
eigensubspace, 3, 30
EXIT characteristic, 153–155
MI performance, 156
MMSE of, 29
MSE performance, 156–158
prefilter formula, 29
vector, *see* VPC
permutation matrix, 115, 118, 119
phase shift keying, *see* PSK
polynomial expansion, 33, 72
Porat, 154, 158
definition, 153
power delay profile
exponential, 126, 151, 158, 163, 165
uniform, 126, 163
preconditioning, *see* BCG
predictor, 13
Proakis
b, 144, 151–154
c, 152, 154, 157
definition, 153
pseudoinverse, 8, 33, 47
definition, 179
PSK, 120

QAM, 120
QPSK, 120, 141, 145–151, 153, 158, 165
mapper, 120
soft demapper, 136, 150
statistics, 130
QR factorization
algorithm, 192
computational complexity, 192
derivation, 191–192
quadrature amplitude modulation, *see* QAM
quadriphase shift keying, *see* QPSK

Rake receiver, 113
random interleaver, *see* interleaver
random matrix theory, 72
random sequences, *see* MWF
rate of convergence
of BCG, *see* BCG
of CG, *see* CG
RC, 4, 6, 14, 18, 23, 26, 106, 132, 133, 166
MWF, *see* RCMWF
RCMWF
algorithm, 24
computational complexity
comparison, 107–108
formula, 25
derivation, 22–26
real normal distribution, *see* normal distribution
recursive least squares, *see* RLS
reduced-complexity, *see* RC
reduced-complexity matrix Wiener filter, *see* RCMWF
reduced-rank MSMWF, *see* MSMWF
reduced-rank MWF, 5
basis invariance, 28
computational efficiency, 28
CS based, *see* CS

derivation, 26–29
formula, 27
formula of adjusted, 132
MMSE of, 27
optimization, 27
PC based, *see* PC
prefilter matrix, 26
regularization
definition, 182
derivation, 180–183
diagonal loading, *see* diagonal loading
of BCG, *see* BCG
of CG, *see* CG
of CS, *see* CS
of Lanczos, *see* Lanczos
Tikhonov, *see* Tikhonov
reorthogonalization, 39, 43
residuum matrix, 47, 49–53, 68, 108
definition, 46
RLS, 71
Ruhe, *see* BLR

S-random interleaver, *see* interleaver
sampling techniques, *see* Monte Carlo
Schur complement, 24, 76, 100, 177, 178
definition, 177
inversion, *see* matrix inversion
scrambling, 133
sequential Monte Carlo sampling techniques, *see* Monte Carlo
set partitioning, 121
signal-to-noise ratio, *see* SNR
SIMO, *see* MA-SIMO
single-input multiple-output, *see* MA-SIMO
SNR, 153, 156
definition, 144
threshold, 121, 156
soft demapper, 121, 127, 128, 137, 141, 147, 149
derivation, 134–136
for QPSK, *see* QPSK
soft interference cancellation, *see* iterative soft interference cancellation
soft symbol, 139
spread spectrum techniques, *see* CDMA
spreader
derivation, 121–124
OVSF, *see* OVSF
spreading factor, 122, 151
orthogonal variable, *see* OVSF
square root matrix
general, 8, 197, 198
definition, 184
positive semidefinite, 8, 198
definition, 184
step size matrix, 67, 97
formula, 47, 53
optimization, 47
successive DFD, *see* DFD
superlinear convergence
of BCG, *see* BCG
of CG, *see* CG
symbol mapper, *see* mapper
symbolwise interleaving, *see* interleaver

Tikhonov, 182–183
factor, 182
filter factors, 182–183
time-invariant, 133, 148, 151, 156, 159, 165, 166
time-variant, *see* TV
Toeplitz, 18, 60
block, *see* block Toeplitz
Trench, 18
tridiagonal matrix, 30, 42, 43
block, *see* block tridiagonal matrix
Turbo equalizer, *see* Turbo receiver
Turbo receiver, 4, 6, 114, 141, 145, 146
BER performance, 165–171
complexity analysis, 163–164
derivation, 126–139
EXIT chart analysis, 152
MSE, *see* MSE
relationship to interference cancellation, 138–139
widely linear processing, 131
TV, 151, 165, 166

uniform power delay profile, *see* power delay profile
unit impulse, 9, 116

VCG
computational complexity, *see* CG
MSVWF implementation, *see* MSVWF

VCS, 31, 159
computational complexity, *see* CS
MSE performance, 159–161
vector 2-norm, *see* Euclidean norm
vector conjugate gradient, *see* VCG
vector cross-spectral, *see* VCS
vector norm, *see* Euclidean norm
vector principal component, *see* VPC
vector Wiener filter, *see* VWF
Viterbi, 138
VPC, 30, 159
computational complexity, *see* PC
MSE performance, 159–161
VWF, 17, 139, 144, 146, 159, 176
BER performance, 168–170
CG, *see* VCG
computational complexity, 160
CS, *see* VCS
EXIT characteristic, 151
MSE performance, 159–161
multistage, *see* MSVWF
PC, *see* VPC

WCDMA, 113
WF, 1, 14, 71, 144, 151
computational complexity, 163
continuous, 13
EXIT characteristic, 144, 151, 153–155
matrix, *see* MWF
MI performance, 156
MSE performance, 156–158
multistage, *see* MSWF
multistage matrix, *see* MSMWF
multistage vector, *see* MSVWF
vector, *see* VWF
wideband code division multiple-access, *see* WCDMA
widely linear processing, *see* Turbo receiver
Wiener filter, *see* WF
Wiener–Hopf equation, 1, 4, 17, 72
continuous, 13
formula, 16
solving via BCG, 97

zero padding, 123