

Modifications of Prony's Method for the Reconstruction of Structured Functions

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

„Doctor rerum naturalium“

der Georg-August-Universität Göttingen

im Promotionsstudiengang *Mathematical Sciences*

der Georg-August University School of Science (GAUSS)

vorgelegt von

INGEBORG MARLEN KELLER

aus Korbach

Göttingen, 01.03.2021

Betreuungsausschuss:

Prof. Dr. Gerlind Plonka-Hoch
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. Dr. Russell Luke
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. Dr. Thorsten Hohage
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Mitglieder der Prüfungskommission:

Referentin:

Prof. Dr. Gerlind Plonka-Hoch
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Korreferent:

Prof. Dr. Jürgen Prestin
Institut für Mathematik
Universität zu Lübeck

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Russell Luke
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. Dr. Thorsten Hohage
Institut für Numerische und Angewandte Mathematik
Georg-August-Universität Göttingen

Prof. Dr. Stephan Huckemann
Institut für Mathematische Stochastik
Georg-August-Universität Göttingen

Prof. Dr. Preda Mihailescu
Mathematisches Institut
Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 12.04.2021

Lassen wir uns nicht schrecken
durch die Ungunst äußerer Umstände,
haben wir für alle Schwierigkeiten
nur eine Antwort:
Erst recht!

CLARA ZETKIN

Acknowledgments

In the last years I had the opportunity to work in the Institute for Numerical and Applied Mathematics at the University of Göttingen. This thesis is the result of this work. First and foremost I would like to thank my supervisor Gerlind Plonka-Hoch. Not only did she introduce me to the field of image and signal processing. Without her guidance and continuous support this thesis would not exist. Moreover, my sincere thanks goes to Jürgen Prestin, who did not hesitate to agree when asked if he would be a referee for this thesis. I would also like to thank my co-supervisors Russell Luke and Thorsten Hohage for always having a sympathetic ear during these years.

Furthermore, I would like to gratefully acknowledge the financial support of the German Research Foundation (DFG) in the framework of the Research Training Group 2088 (GRK 2088) *Discovering structure in complex data: Statistics meets Optimization and Inverse Problems*. Being a member of the RTG during my doctoral studies enabled me to attend many conferences and workshops and helped to broaden my scientific and professional horizon, but also gave me the opportunity to improve my soft skills.

Moreover, I would like to thank my amazing current and former colleagues and friends from the *Mathematical Signal and Image Processing* group for the incredible working environment and the enormous support. In particular, I would like to thank Sina Bittens, Wolfgang Keller, Hanna Knirsch, Markus Petz and Max Pentzlin for proof-reading this thesis.

My very special gratitude goes to my family for their unconditional support and in particular, to my brothers Wilfried and Wolfgang, who are and always will be a big source of inspiration for me. I am eternally grateful to my friends for their tremendous support and patience with me throughout the last years and especially during pandemic times. Last but not least I would like to thank Max for his unconditional support, love and for always helping me turn on the light even if times seemed dark.

Contents

Notation	ix
List of Figures	xi
List of Tables	xiii
List of Algorithms	xv
1 Introduction	1
2 The Prony Method	5
2.1 The Classical Prony Method	5
2.2 The Generalized Prony Method	9
3 The Reconstruction of Generalized Exponential Sums and Generalized Trigonometric Sums	15
3.1 Reconstruction of Trigonometric Expansions	16
3.2 Reconstruction of Generalized Exponential Sums and Generalized Trigonometric Sums	27
3.2.1 Generalized Shift Operators	28
3.2.2 Generalized Symmetric Shift Operators	40
3.2.3 Reconstruction Using Linear Differential Operators	50
3.3 Applications to Special Expansions	53
4 Reconstruction of Expansions into Orthogonal Polynomials	65
4.1 Reconstruction Method Using Generating Functions	66
4.2 Reconstruction Methods Using Linear Operators	76
4.2.1 Quadratic Sampling Schemes	80
4.3 Reconstruction Method Using q -analogs	83
4.4 Numerical Examples and Applications	88
4.4.1 Reconstruction of Sparse Chebyshev Expansions of First Kind	88
4.4.2 Reconstruction of Sparse Legendre Expansions	94
5 Numerical Behaviour of the Prony Method for Generalized Exponential Sums	101
5.1 Modification of the ESPRIT Algorithm for Generalized Exponential Sums	102
5.2 Partially Known Frequency Parameters	113

5.3	Sub-Sampling Based Algorithm for the Recovery of Generalized Exponential Sums	114
5.3.1	Collision Free Sub-Sampling	115
5.3.2	Colliding Frequencies	120
5.4	Modification of Prony's Method for Sparse Approximation	123
5.4.1	Gauß-Newton and Levenberg-Marquardt Iteration	129
5.5	Numerical Examples	130
6	Conclusion and Outlook	139
	Bibliography	141

Notation

- $\mathbb{N} := \{1, 2, 3, \dots\}$ and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$
- \mathbb{Z} : integers
- \mathbb{R} : real numbers and \mathbb{R}_+ : positive real numbers
- \mathbb{C} : complex numbers
- \mathbf{I}_N : N -by- N identity matrix
- $\mathbf{0}_{N,M}$: N -by- M zero matrix
- $\mathbf{0}_N := (0, \dots, 0)^T \in \mathbb{R}^N$
- $\mathbf{H}_{N,M} \in \mathbb{C}^{N \times M}$ resp. $\mathbf{H}_M \in \mathbb{C}^{M \times M}$: N -by- M Hankel matrix resp. M -by- M Hankel matrix as in (2.10)
- $\mathbf{V}_{N,M} \in \mathbb{C}^{N \times M}$ resp. $\mathbf{V}_M \in \mathbb{C}^{M \times M}$: N -by- M Vandermonde matrix resp. M -by- M Vandermonde matrix
- $\mathbf{H}_M^{\text{Toep}} \in \mathbb{C}^{M \times M}$: M -by- M Hankel-plus-Toeplitz matrix
- $\text{diag}(a_1, \dots, a_N) \in \mathbb{C}^{N \times N}$ resp. $\text{diag}((a_j)_{j=1}^N)$: Diagonal matrix with a_1, \dots, a_N as main diagonal entries
- $\mathbf{C}_M(\mathbf{p}) \in \mathbb{C}^{(M \times M)}$: M -by- M companion matrix of the monic polynomial $p(z) = p_0 + p_1 z + \dots + z^M$ as in (2.6)
- $\mathbf{T}_N \in \mathbb{R}^{N \times N}$: Chebyshev Vandermonde matrix of size $N \times N$ as in (3.6)
- $C(\mathbb{R}) := \{f : \mathbb{R} \rightarrow \mathbb{C} \mid f \text{ continuous}\}$ vector space of continuous, real valued functions
- $C^d([a, b])$: vector space of d -times continuously differentiable functions on $[a, b]$
- $C^\infty(\mathbb{R})$: space of smooth, real valued functions
- $L(V)$: vector space of all linear operators on a (normed) vector space V
- \hat{f} : Fourier transform of the function f with $\hat{f}(\omega) := F(f)(\omega) := \int_{-\infty}^{\infty} f(x) e^{ix\omega} dx$
- $\mathcal{L}(f)$: Laplace transform of the function f with $\mathcal{L}(f)(s) := \int_0^{\infty} e^{-st} f(t) dt$

- $\text{dom}(A)$: domain of the operator A
- Id : Identity operator
- S_h : shift operator for the shift parameter h as in (3.1)
- $S_{h,-h}$: symmetric shift operator with shift parameter h as in (3.3)
- $S_{G,H,h}$: generalized shift operator as in (3.22)
- $\arg(c)$: argument of a complex number $c = re^{i\varphi}$ with $\arg(c) := \varphi$
- $L^2([a,b], \omega) := \left\{ f: (a,b) \rightarrow \mathbb{R} \mid \left(\int_a^b \omega(x) f(x)^2 dx \right)^{1/2} < \infty \right\}$: Hilbert space with inner product $\langle f, g \rangle := \int_a^b f(x)g(x)\omega(x)dx$
- $\delta_{k,l}$: Kronecker delta with $\delta_{k,l} := \begin{cases} 1 & \text{if } k = l, \\ 0 & \text{otherwise,} \end{cases}$ for $k, l \in \mathbb{Z}$
- $\deg(f)$: degree of the polynomial f
- $\langle f, g \rangle_Q := \int_a^b f(x)g(x)\omega(x)dx$ inner product corresponding to the set of orthogonal polynomials $\{Q_k | k \in \mathbb{N}_0\}$
- $g(x, t)$: generating function as in (4.6)
- $L_{p,q,\omega}$: Sturm-Liouville operator as in Definition 4.16
- $V^* := \{\varphi: V \rightarrow \mathbb{K} \mid \varphi \text{ linear}\}$ dual space of the vector space V over the field \mathbb{K}
- A^* : adjoint operator as in Definition 4.21
- \mathbf{A}^+ : Moore-Penrose inverse as in Definition 5.6
- $\text{gcd}(p, u)$: greatest common denominator of two numbers $p, u \in \mathbb{N}$
- \mathbf{P}_z : projection matrix as in (5.40)

List of Figures

3.1	Real and imaginary part of the sparse expansion into Gaussian chirps f as given in Example 3.50 in blue. The black dots indicate the used signal values. The reconstructed signal was obtained by Algorithm 3 and is shown in red. It cannot be distinguished visually from the original signal f	56
3.2	Original signal f of the form (3.30) with the parameters given in Table 3.3 in blue. The red dots indicate the used signal values of f and the black dots indicate the used signal values of \tilde{f} . The reconstructed signal was obtained by application of Corollary 3.51 and Algorithm 3 and is shown in red. It cannot be distinguished visually from the original signal f	58
3.3	Original signal f of the form (3.32) with the parameters given in Table 3.4 in blue. The black dots indicate the used signal values of f . The reconstructed signal was obtained by application of Corollary 3.53 and Algorithm 4 and is shown in red. It cannot be distinguished visually from the original signal f	60
4.1	Different generating functions for the Chebyshev polynomials of first kind for the parameter $t = \frac{1}{\sqrt{2}}$. The red line shows the generating function for the weights $w_k = 1$ for $k \in \mathbb{N}_0$, the blue line the generating function for $w_k = \frac{1}{k!}$ for $k \in \mathbb{N}_0$ and the black line for the weights $w_k = \frac{1}{k}$ for $k \in \mathbb{N}$	89
4.2	Original signal f of the form (4.26). The reconstructed signal was obtained by Corollary 4.42 and Algorithm 5 and is shown in red. The original signal f is shown in blue. Both signals cannot be distinguished visually.	93
4.3	4-sparse Legendre expansion with parameters given in Table 4.5 and its reconstruction. The reconstructed signal was obtained by Algorithm 5 and is shown in red. The original signal is shown in blue. Both signals cannot be distinguished visually.	97
4.4	Recovery of 5-sparse Legendre expansion with parameters given in Table 4.6 with Algorithm 5 shown in red. The blue signal shows the 5-sparse Legendre expansion recovered by the Gaussian quadrature approach. Both signals cannot be distinguished visually.	99

5.1	Singular values of the matrices $\mathbf{H}_{22,13}$ (blue) and $\mathbf{H}_{29,6}$ (red) used in Algorithm 7 on the logarithmic y -axis and the corresponding index on the x -axis.	131
5.2	5-sparse exponential cosine expansion with parameters given in Table 5.1. The black dots indicate the sampled used for the reconstruction using Algorithm 7.	132
5.3	Parameter α_j (black circles) as well as the corresponding recovered parameters obtained by Algorithm 8 (red stars) and Algorithm 3 (blue squares). The x -axis shows the real part of the parameter α_j . The y -axis the imaginary part of $\alpha_j, j = 1, \dots, 5$	135
5.4	Recovery of the 5-sparse exponential sine expansion (5.49) with Algorithm 3 in red. Recovery obtain by Algorithm 8 shown in blue. Both reconstructions cannot be distinguished visually.	136
5.5	Parameter α_j (black circles) as well as the corresponding recovered parameters obtained by Algorithm 8 (red stars). The x -axis shows the real part of the parameter α_j . The y -axis shows the imaginary part of $\alpha_j, j = 1, \dots, 6$	138

List of Tables

3.1	Examples of generalized shift operators of the form $S_{G,H,h}$ as in Definition 3.26 with $G(x) = \int_a^x \frac{1}{g(t)} dt$, their corresponding eigenfunctions $H(\cdot)e^{G(\cdot)}$ and the sample values needed for the reconstruction of a generalized exponential sum as in Definition 3.12.	38
3.2	Parameters c_j and α_j for an expansion into Gaussian chirps as in (3.28) with $M = 10$, see Figure 3.1.	55
3.3	Parameters c_j, s_j and α_j for $j = 1, \dots, M$ for an expansion f as in (3.30) with $M = 8$	58
3.4	Parameters c_j, α_j and β_j for $j = 1, 2$ for a expansion f as in (3.32) with $M = 2$	59
4.1	Orthogonal polynomials with their corresponding generating function $g(x, t)$ and the corresponding weight coefficients $w_n, n \in \mathbb{N}_0$	69
4.2	Error in the numerical evaluation of indices of the “active” basis polynomials T_{n_j} of the sparse Chebyshev expansion (4.26) using Algorithm 5 and different generating functions.	92
4.3	Error in the numerical evaluation of coefficients c_{n_j} of the sparse Chebyshev expansion (4.26) using Algorithm 5 and different generating functions.	93
4.4	Performance of the reconstruction of 4-sparse Chebyshev expansions using Algorithm 5 and different generating functions.	94
4.5	Numerical evaluation of the indices of the “active” basis polynomials and corresponding coefficients of the sparse Legendre expansion (4.30) using Algorithm 5 and the ordinary generating function.	96
4.6	Numerical evaluation of the indices of the “active” basis polynomials and corresponding coefficients of the sparse Legendre expansion (4.34) using the Gaussian quadrature approach as well as Algorithm 5 and the ordinary generating function.	98
5.1	Parameters c_j and α_j for a sparse exponential cosine expansion as in (5.47) with $M = 5$, see Figure 5.2.	131
5.2	Maximal error in the frequency coefficients α_j for a 5-sparse exponential cosine expansion as in (5.47) obtain by Algorithm 3 and Algorithm 7.	132
5.3	Parameters c_j and α_j for a 5-sparse expansion into Gaussian chirps as in (5.48).	133

5.4	Performance results of Algorithm 7 using different noise levels.	133
5.5	Parameters c_j and α_j , $j = 1, \dots, 5$, for a 5-sparse exponential sine expansion as in (5.49).	134
5.6	Error in the numerical evaluations of the parameters c_j and α_j for the 5-sparse exponential sine expansion as in (5.49) obtained using Algorithm 8.	134
5.7	Error in the numerical evaluations of the parameters c_j and α_j , $j = 1, \dots, 5$, for the 5-sparse exponential sine expansion as in (5.49) obtained using Algorithm 3.	135
5.8	Error in the numerical evaluations of the parameters c_j and α_j , $j = 1, \dots, 5$ for the 5-sparse exponential sine expansion obtained using Algorithm 7.	136
5.9	Parameters c_j and α_j for a 6-sparse expansion into Gaussian chirps as in (5.48).	137
5.10	Error in the numerical evaluations of the parameters c_j and α_j , $j = 1, \dots, 6$ for the 6-sparse expansion into Gaussian chirps with clustered frequencies obtained using Algorithm 8.	137

List of Algorithms

1	The Classical Prony Method	7
2	The Generalized Prony Method	13
3	Prony's Method in Terms of Generalized Shift Operators	40
4	Prony's Method in Terms of Generalized Symmetric Shift Operators	49
5	Algorithm for the Reconstruction of Expansions into Orthogonal Polynomials Using a Generating Function	73
6	q -Analog Based Prony Algorithm for the Reconstruction of Orthogonal Polynomials	87
7	ESPRIT Algorithm for Generalized Exponential Sums	108
8	Sub-Sampling Algorithm for Generalized Exponential Sums	119

1 Introduction

The reconstruction and analysis of sparse signals is a common and widely studied problem in signal processing, for example in wireless telecommunication, see, e.g., [QI99], biomedical engineering, see, e.g., [BKRE09], or power system theory, see, e.g., [TSSP91]. Hereby, most recovery methods exploit structures or special properties of the functions which are to be reconstructed. Particularly interesting are methods which aim to recover functions which possess a sparse representation in a given basis and use only a small set of sampling values.

Over the years different approaches have been used to analyse and recover structured functions such as greedy methods, see, e.g., [CDS98], or non-deterministic methods in the field of compressed sensing, see, e.g., [CRT06, Don06]. Usually, the reconstruction algorithms used in the framework of compressed sensing are based on l_1 -minimization methods and only recover the exact signal with a certain probability.

One example of a deterministic algorithm for the recovery of structured functions was given in [PW13]. Here, the authors used a small number of Fourier samples for the recovery of structured functions. In particular, the authors studied how real spline functions of order m with non-uniform knots containing N terms can be uniquely reconstructed by $m + N$ Fourier samples. Furthermore, in [Bit17, BZI19] deterministic fast Fourier transform based algorithms for the recovery of 2π periodic functions with short frequency support were derived.

One of the best studied examples of structured functions, which for example naturally appear in decay processes, are sparse exponential expansions, i.e.,

$$f(x) = \sum_{j=1}^M c_j e^{\alpha_j x},$$

where $c_j \in \mathbb{C} \setminus \{0\}$ and $\alpha_j \in \mathbb{C}$ with $|\text{Im}(\alpha_j)| < \pi$ for $j = 1, \dots, M$. One way to identify the parameters α_j and c_j , $j = 1, \dots, M$ and therefore, recover such sparse exponential expansions is by Prony's method due to [dP95], which gained great popularity in the last decade. This method uses $2M$ equidistant sample values $f(x_0 + hk)$ for $k = 0, \dots, 2M - 1$, with sampling distance $h \in \mathbb{R} \setminus \{0\}$ and starting point $x_0 \in \mathbb{R}$. Even though Prony's method is a rather simple one, it is highly versatile and can also be used for the algebraic recovery of piecewise smooth functions, see [Bat15], or sparse phase retrieval in one dimension, see [BP17]. For system reduction, Prony's

method is related to the problem of low-rank approximation of structured matrices, in particular Hankel matrices and corresponding nonlinear least-squares problems, see [Mar12,UM14]. The problem of sparse approximation also connects Prony's method and the AKK-theory, see [AAK71], which is a theory used for best approximation problems associated with (infinite) Hankel matrices and operators, and is used by engineers for model reduction. In [Pot17] this connection is used for the approximation of signals using exponential sums with frequencies inside the unit circle.

Since Prony's method is known to be numerically unstable, different numerically stable methods have been derived. One possible stabilization method is the so called Approximate Prony method, see [PT10], which is based on [BM05]. Another possible stabilizing method was proposed in [FMP12], where instead of samples of the original signal, a windowed average of their autocorrelation sequence is used. This approach is based on the application of operators of the form $\sum_{k=0}^M g\left(\frac{k}{M}\right) f(k)e^{(ik)}$, where g is a suitable smooth, even function studied in [MP00,MP05], and, enables rigorous performance guarantees for Prony's method.

One of the most notable generalizations of Prony's method, the so-called generalized Prony method, was introduced in 2013 by Peter and Plonka, see [PP13]. Here the reconstruction algorithm behind Prony's method was extended to the recovery of sparse expansions into eigenfunctions of certain linear operators, i.e.,

$$f(x) = \sum_{j \in J} c_j v_j(x),$$

where J is a index set of size $M \in \mathbb{N}$, $c_j \in \mathbb{C} \setminus \{0\}$ and v_j being the eigenfunction of a linear operator A . The reconstruction can now be done using $2M$ values of the form $F(A^k(f))$ for $k = 0, \dots, 2M - 1$, where F denotes a linear functional, satisfying $F(v_j) \neq 0$ for all eigenfunctions v_j .

While some examples of suitable linear operators were given by Peter and Plonka, e.g., the shift operator as well as certain differential operators, the sample values needed for the reconstruction are not always accessible in practice. This leads to the following questions.

- (i) Can we find other suitable linear operators with meaningful structured functions as eigenfunctions and easily accessible sample values?
- (ii) Can we modify methods for the numerical stabilization of the Prony method such as the ESPRIT method, see, e.g., [PT13], for the generalized Prony method?

Based on the questions above, we investigate what functions can be recovered using only sample values that are easily accessible. Based on the theory of one-parameter semigroups we derive so called generalized shift operators and their eigenfunctions, so-called generalized exponential sums. The framework of generalized exponential

sums and generalized shift operators covers all previously derived examples for the application of the generalized Prony method. Moreover, we extend this to generalized trigonometric expansions. Furthermore, we elaborate on the connection between generalized shift operators and linear differential operators. We introduce a Prony based method for the reconstruction of orthogonal polynomials based on generating functions and connect this recovery method to our generalized shift operators.

Additionally, we can answer the second question above in the affirmative. We show that the numerical stabilizations of Prony's method are not limited to the case of exponential expansions but also can be used for the structured functions we considered throughout this thesis.

This dissertation is organized as follows. In Chapter 2 we give an overview of Prony's method and derive the algorithm for the reconstruction of sparse exponential expansions. Moreover, we present the generalized Prony method for the reconstruction of M -sparse expansions into eigenfunctions and highlight how the classical Prony method can be viewed as a special case of the generalized Prony method by using shift operators.

In Chapter 3 we extend the notion of shift operators. We introduce symmetric shift operators for the reconstruction of trigonometric functions and generalized shift operators for the reconstruction of generalized exponential sums. To this end, we introduce the concept of operator-valued exponential functions and prove, based on the work [DOTV97], that each generalized shift operator can be written as such an operator-valued exponential function with a suitable differential operator A . We show how non-stationary signal such as generalized exponential sums or generalized trigonometric expansion can be recovered from suitable function values using the generalized Prony method. Furthermore, we use the established connection between shift operators and differential operator to derive a Prony based reconstruction method for generalized exponential sums using linear differential operators. Moreover, we illustrate our finding with numerical examples.

In Chapter 4 we focus on the reconstruction of sparse expansions into orthogonal polynomials. We recall some of the most important properties of orthogonal polynomials and derive a new recovery method for the reconstruction of sparse expansions into orthogonal polynomials using Prony's method and generating functions. Additionally, we explain how this recovery fits into the framework of generalized exponential sums and generalized shift operators, which we developed in Chapter 3. Moreover, we highlight the differences between our newly developed algorithm and some already existing approaches based on Prony's method and illustrate our findings with numerical examples.

Chapter 5 is dedicated to the numerical analysis of the Prony method for generalized exponential sums. Based on [PT13] we derive a modified ESPRIT algorithm for the reconstruction of generalized exponential sums. Furthermore, we analyse the case of

partially known frequencies and modify the results on clustered frequencies obtained in [CL20a] for generalized exponential sums. Finally, we elaborate on the modification of Prony's method for sparse approximation. We illustrate our derived algorithms with different numerical examples.

All experiments in this thesis were implemented in MATLAB 2020a and conducted with a 2,7 GHz Intel Core i5 processor and 8 GB 1867 MHz DDR3. For the computation of polynomials roots, singular values and least square solutions internal MATLAB routines have been used.

2 The Prony Method

In 1795 Gaspard Riche de Prony developed a method for the reconstruction of low order exponential functions, see [dP95]. Even though this method dates back to the 18th century, it has preserved up to this date and has inspired a variety of applications in the field of electrical engineering, see [VMB02], for the approximation of Green's function in quantum chemistry, see [YFG⁺04], and even in the medical field, see [JHLC16].

As the Prony method lies at the core of this thesis, we will begin by introducing the classical Prony method as well as one possible generalization in the form of the generalized Prony method by Peter and Plonka, see [PP13].

2.1 The Classical Prony Method

Prony's method is a procedure for the identification of parameters in **exponential sums**, that is, the Prony method can be used in order to determine frequencies α_j and corresponding coefficients c_j of a function $f: \mathbb{R} \rightarrow \mathbb{C}$ given as

$$f(x) = \sum_{j=1}^M c_j e^{\alpha_j x} \quad (2.1)$$

for $c_j \in \mathbb{C} \setminus \{0\}$ and $\alpha_j \in \mathbb{C}$ with $-\pi < \text{Im}(\alpha_1) \leq \text{Im}(\alpha_2) \leq \dots \text{Im}(\alpha_M) < \pi$ for $j = 1, \dots, M$. Here it is always assumed that M is minimal, i.e., that $\alpha_j \neq \alpha_i$ for $i \neq j$. One big advantage of Prony's method is that it only uses the minimal number of $2M$ values $f(k)$ for $k = 0, \dots, 2M - 1$, combined with methods from linear algebra to reconstruct f .

For each signal of the form (2.1), we can define the associated **Prony polynomial** as

$$P(z) := \prod_{j=1}^M (z - \lambda_j) = \sum_{k=0}^M p_k z^k \quad (2.2)$$

with $\lambda_j := e^{\alpha_j}$ for $j = 1, \dots, M$. We note that by definition the Prony polynomial is monic, i.e., $p_M = 1$.

We can interpret the signal f in (2.1) as the solution to a homogeneous linear difference equation and, therefore, understand the Prony polynomial P as the characteristic polynomial corresponding to the linear difference equation. We verify this by the

following computation for $m \in \mathbb{N}_0$

$$\begin{aligned} \sum_{k=0}^M p_k f(k+m) &= \sum_{k=0}^M p_k \sum_{j=1}^M c_j e^{\alpha_j(k+m)} = \sum_{j=1}^M c_j e^{\alpha_j m} \sum_{k=0}^M p_k e^{\alpha_j k} \\ &= \sum_{j=1}^M c_j e^{\alpha_j m} \underbrace{P(e^{\alpha_j})}_{=0} = 0. \end{aligned} \quad (2.3)$$

Employing the fact that the Prony polynomial is monic, we can rewrite (2.3) as

$$\sum_{k=0}^{M-1} p_k f(k+m) = -f(M+m) \quad \forall m \in \mathbb{N}_0. \quad (2.4)$$

Based on this we can form the minimal linear system for the computation of the coefficients $p_k, k = 0, \dots, M-1$, of the Prony polynomial P given in (2.2)

$$\underbrace{\begin{pmatrix} f(0) & f(1) & \dots & f(M-1) \\ f(1) & f(2) & \dots & f(M) \\ \vdots & \vdots & & \vdots \\ f(M-1) & f(M) & \dots & f(2M-2) \end{pmatrix}}_{=: \mathbf{H}_M} \begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_{M-1} \end{pmatrix} = - \begin{pmatrix} f(M) \\ f(M+1) \\ \vdots \\ f(2M-1) \end{pmatrix}. \quad (2.5)$$

The matrix $\mathbf{H}_M \in \mathbb{C}^{M \times M}$, which entries are constant along the diagonals parallel to the antidiagonal, is called a **Hankel matrix** and is indeed invertible since it can be factorized in the following way

$$\begin{aligned} \mathbf{H}_M &= (f(k+m))_{k,m=0}^{M-1} = \left(\sum_{j=1}^M c_j e^{\alpha_j(k+m)} \right)_{k,m=0}^{M-1} = \left(\sum_{j=1}^M c_j e^{\alpha_j k} e^{\alpha_j m} \right)_{k,m=0}^{M-1} \\ &= \left(c_j e^{\alpha_j k} \right)_{k=0,j=1}^{M-1,M} \cdot \left(e^{\alpha_j m} \right)_{j=1,m=0}^{M,M-1} \\ &= \left(e^{\alpha_j k} \right)_{k=0,j=1}^{M-1,M} \cdot \mathbf{diag}(c_1, c_2, \dots, c_M) \cdot \left(e^{\alpha_j m} \right)_{j=1,m=0}^{M,M-1} \\ &= \mathbf{V}_M(e^{\alpha_1}, \dots, e^{\alpha_M}) \cdot \mathbf{D} \cdot \mathbf{V}_M^T(e^{\alpha_1}, \dots, e^{\alpha_M}), \end{aligned}$$

where

$$\mathbf{V}_M(e^{\alpha_1}, \dots, e^{\alpha_M}) := \left(e^{\alpha_j k} \right)_{k=0,j=1}^{M-1,M} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ e^{\alpha_1} & e^{\alpha_2} & \dots & e^{\alpha_M} \\ \vdots & \vdots & \dots & \vdots \\ e^{\alpha_1(M-1)} & e^{\alpha_2(M-1)} & \dots & e^{\alpha_M(M-1)} \end{pmatrix} \in \mathbb{C}^{M \times M}$$

denotes the **Vandermonde matrix** to the nodes e^{α_j} and $\mathbf{D} := \mathbf{diag}(c_1, c_2, \dots, c_M)$ is the **diagonal matrix** containing the parameter c_j for $j = 1, \dots, M$. In order not to

overload the notation we will omit the specification of the nodes and only write \mathbf{V}_M up to Chapter 5.4. The matrix \mathbf{D} has only nonzero diagonal entries and is, therefore, invertible and the Vandermonde matrix \mathbf{V}_M is invertible because of $\alpha_j \neq \alpha_k$ for $j \neq k$, see [Sch02], Section 3.1.2, Lemma 3.1.2.

Once the coefficients of the Prony polynomial have been computed, we can calculate its roots $\lambda_j = e^{\alpha_j}$ for $j = 1, \dots, M$. One way to do this, is by solving the eigenvalue problem of the **companion matrix** $\mathbf{C}_M(\mathbf{p})$ of the Prony polynomial P with coefficients p_k , $k = 0, \dots, M-1$, and $\mathbf{p} = (p_0, \dots, p_{M-1})^T$, see [HJ13] pages 194 – 195, with

$$\mathbf{C}_M(\mathbf{p}) := \begin{pmatrix} 0 & 0 & \dots & 0 & p_0 \\ 1 & 0 & \dots & 0 & p_1 \\ 0 & 1 & \dots & 0 & p_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & p_{M-1} \end{pmatrix}. \quad (2.6)$$

The values α_j , $j = 1, \dots, M$, can then be easily computed by taking the principal branch of the logarithm.

Therefore, as a last step, we compute the coefficients c_j in (2.1) for $j = 1, \dots, M$. This can be done by solving the following Vandermonde system

$$f(k) = \sum_{j=1}^M c_j e^{\alpha_j k} \quad \text{for } k = 0, \dots, 2M-1. \quad (2.7)$$

We summarize the above results in the following algorithm.

Algorithm 1 The Classical Prony Method

Input: $M \in \mathbb{N}$ and $f(k)$ for $k = 0, \dots, 2M-1$, as in (2.1).

- 1: Form the Hankel matrix $\mathbf{H}_M := (f(k+\ell))_{k,\ell=0}^{M-1}$ as well as $\mathbf{f} := (f(M+\ell))_{\ell=0}^{M-1}$ and solve $\mathbf{H}_M \cdot \mathbf{p} = -\mathbf{f}$ for $\mathbf{p} := (p_\ell)_{\ell=0}^{M-1}$ as in (2.5).
- 2: Define the Prony polynomial $P(z) := \sum_{k=0}^M p_k z^k$ with $p_M = 1$ as in (2.2), find all roots $\lambda_j = e^{\alpha_j}$ and compute $\alpha_j = \log(\lambda_j)$ for $j = 1, \dots, M$.
- 3: Determine the coefficients c_j for $j = 1, \dots, M$ as the solution to the Vandermonde system $\mathbf{V}_{2M,M} \cdot \mathbf{c} = \mathbf{f}_{2M}$ as in (2.7) with $\mathbf{V}_{2M,M} := (e^{\alpha_j k})_{k=0,j=1}^{2M-1,M}$, $\mathbf{c} := (c_j)_{j=1}^M$ and $\mathbf{f}_{2M} := (f(k))_{k=0}^{2M-1}$.

Output: c_j and α_j for $j = 1, \dots, M$ as in (2.1).

Remark 2.1 The classical Prony method assumes M to be a known parameter. As this is often not the case in applications, see, e.g., [Hac05], there also exists a variant of Prony's method which computes the parameter M using oversampling and an upper bound $N > M$ on the number of terms in the exponential sum (2.1). In the noiseless case, we construct Hankel matrices $\mathbf{H}_k \in \mathbb{C}^{k \times k}$ similarly to the one in (2.5) but of

different size $k = 1, \dots, N$, and note that $\det(\mathbf{H}_k) \neq 0$ for $k = M$ and $\det(\mathbf{H}_k) = 0$ for $k > M$, see [KL03], Theorem 4. However, due to the nature of the signal f , we may get $\det(\mathbf{H}_k) = 0$ for some $k < M$. Therefore, it is not sufficient to determine $\det(\mathbf{H}_k)$ for growing k in order to find M . One possible approach to obtain the parameter M is to analyze the singular values of the matrix \mathbf{H}_N for some upper bound N for M and to determine M as the (numerical) rank of \mathbf{H}_N .

Since the matrices \mathbf{H}_M and $\mathbf{V}_{2M,M}$ occurring in Algorithm 1 are usually ill-conditioned, see [Tyr94] and [Pan16], there exist many methods aiming at higher numerical stability such as the approximate Prony method, see [PT10], or the ESPRIT algorithm, see [RK89], and Chapter 5. \diamond

Remark 2.2 (i) The classical Prony method is also known as the annihilating filter method and is often used in the field of electrical engineering especially, with regard to finite rate of innovation methods, see, e.g., [VMB02, DVB07]. In this framework the Prony polynomial P is the z -transform of a filter with coefficients p_k for $k = 0, \dots, M$, i.e., it is of the form $P(z) = \prod_{j=1}^M (1 - z^{-1}\lambda_j) = \sum_{k=0}^M p_k z^{-k}$ with $p_0 = 1$. The sampling scheme is shifted by $M + 1$ and instead of the Hankel system in (2.5), a so-called **Toeplitz** system of the form

$$\begin{aligned}
 & \begin{pmatrix} f(0) & f(-1) & \dots & f(-M+1) \\ f(1) & f(0) & \dots & f(-M+2) \\ \vdots & \vdots & & \vdots \\ f(M-1) & f(M-2) & \dots & f(0) \end{pmatrix} \\
 = & \begin{pmatrix} f(-M+1) & f(-M+2) & \dots & f(0) \\ f(-M+2) & f(-M+3) & \dots & f(1) \\ \vdots & \vdots & & \vdots \\ f(0) & f(1) & \dots & f(M-1) \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & \dots & 1 \\ 0 & \dots & 1 & 0 \\ 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 \end{pmatrix}
 \end{aligned}$$

is used, see [VMB02], page 1420. As seen above the entries of the Toeplitz matrix are constant along the extra-diagonals. Thus, the Toeplitz matrix is connected to the Hankel matrix via multiplication with the counter identity matrix.

(ii) The model f in (2.1) also has close relations to the expansion

$$g(x) = \sum_{j=1}^M c_j \delta(x - s_j)$$

for $c_j \in \mathbb{C} \setminus \{0\}$, $s_j \in \mathbb{R}$, and delta-distribution δ . Using the fact that the Fourier transform $\hat{g}(\omega) := F(g)(\omega) := \int_{-\infty}^{\infty} g(t) e^{i\omega t} dt$ can be extended to the space of tempered distributions, we can interpret the exponential sum f in (2.1) as a Fourier transform of g using $\alpha_j = is_j$ for $j = 1, \dots, M$. Therefore, g can also be recovered using Prony's method and equidistant Fourier samples. Moreover,

this idea was extended to finite linear combination of arbitrary shifts of a given function $\phi \in L^1(\mathbb{R})$ with non vanishing Fourier transform $\hat{\phi}$, i.e.,

$$g(x) = \sum_{j=1}^M c_j \phi(x - \alpha_j)$$

with $\alpha_j \in \mathbb{R}$, for $j = 1, \dots, M$, see [PW13].

(iii) If we assume $|e^{\alpha_j}| < 1$ for $j = 1, \dots, M$, considering the data samples $f(k)$ as Taylor series coefficients

$$\sum_{k=0}^{\infty} f(k)z^k = \sum_{k=0}^{\infty} \sum_{j=1}^M c_j e^{\alpha_j k} z^k = \sum_{j=1}^M \frac{c_j}{1 - e^{\alpha_j} z}$$

relates Prony's method to Padé approximation as shown in [PT14a,Cuy20,WM63].

◇

It is well known that Prony's method is highly instable in the presence of noise, see, e.g., [LC56], pages 276-280, due to the ill-conditioned Vandermonde matrices, see [Pan16]. Different models and algorithms were developed to tackle this problem. In 1973 Pisarenko used the covariance structure of the signal in order to introduce a set of noise reduced samples, see [Pis73]. Moreover, modified versions of Prony's method based on non-linear least square problems and iterative maximum likelihood estimation were introduced by Osborne et al., see [OS91,Os75].

Furthermore, if the minimal separation distance

$$q := \min_{\substack{i,j=1,\dots,M \\ i \neq j}} |\alpha_j - \alpha_i|$$

between the frequencies α_j , $j = 1, \dots, M$, is small, Prony's method is not able to distinguish between the frequencies. In [CL20a] a method for retrieving high resolution information from coarse-scale measurements, using uniform downsampling and exploiting aliasing in order to increase the resolution is introduced. The effect of clustered frequencies and on the stability on Vandermonde matrices have also been studied in [KN21].

2.2 The Generalized Prony Method

In recent years a variety of applications and modifications of Prony's method have been derived for the recovery of structured functions such as sparse expansions into Legendre polynomials, see [PPR13,PT16], or sparse expansions into Chebyshev polynomials, see [PT14b].

Furthermore, multivariate models of Prony's method have been established, see, e.g., [KPRv16, PV20, CL18].

Most notably are the generalized Prony method derived by Peter and Plonka in 2013, see [PP13], as well as the generalized operator based Prony method developed by Plonka and Stampfer in 2020, see [Sta18, SP20]. We will give a brief summary of the generalized Prony method as described in [PP13], and begin with reformulating the classical Prony method.

We note that we can interpret the exponential functions in (2.1) as expansions into eigenfunctions of the shift operator S_1 with

$$\begin{aligned} S_1: C(\mathbb{R}) &\rightarrow C(\mathbb{R}) \\ f &\mapsto f(\cdot + 1). \end{aligned}$$

Using

$$S_1 e^{\alpha_j x} = e^{\alpha_j(x+1)} = e^{\alpha_j} e^{\alpha_j x},$$

we can reinterpret the roots of the Prony polynomial P in (2.2) as the eigenvalues e^{α_j} corresponding to the eigenfunction $e^{\alpha_j x}$. Introducing the point evaluation functional $F_0: C(\mathbb{R}) \rightarrow \mathbb{C}$ with $F_0(f) := f(0)$, we are able to rewrite the samples used in the Prony method in the following way,

$$F_0(S_1^k f) = F_0(f(\cdot + k)) = f(k) \quad \text{for } k = 0, \dots, 2M - 1.$$

The key ingredients are the iterative application of the shift operator S_1 , its eigenvalues and the fact that S_1 is a linear operator. Thus, we have the basis for the generalized Prony method, whose major contribution is that it enables us not only to reconstruct exponential sums as in (2.1) but rather sparse expansions of eigenfunctions of a linear operator A .

We, therefore, consider a normed (complex) vector space V and a linear operator

$$\begin{aligned} A: V &\rightarrow V \\ v &\mapsto Av. \end{aligned}$$

We assume that A possesses eigenvalues λ_j and corresponding eigenfunctions v_j such that

$$Av_j = \lambda_j v_j,$$

where the eigenvalues λ_j are pairwise distinct. Furthermore, we assume that the eigenvalues are simple, i.e., there exists a unique correspondence between the eigenfunction

v_j and its corresponding eigenvalue λ_j .

Definition 2.3 (M-Sparse Expansion) Let $M \in \mathbb{N}$, V be a normed (complex) vector space and $A: V \rightarrow V$ be a linear operator with eigenfunctions v_j and corresponding unique eigenvalues λ_j . Then we call f an **M-sparse expansion into eigenfunctions of A** if f is of the form

$$f(x) = \sum_{j=1}^M c_j v_j(x) \quad (2.8)$$

with $c_j \in \mathbb{C} \setminus \{0\}$ for $j = 1, \dots, M$. We assume M to be minimal.

Definition 2.4 (Prony Polynomial) Let f be an M -sparse expansion into eigenfunctions of a linear operator A as in (2.8). The associated **Prony polynomial** is defined as

$$P(z) := \prod_{j=1}^M (z - \lambda_j) = \sum_{k=0}^M p_k z^k \quad (2.9)$$

with the roots of the Prony polynomial λ_j being the eigenvalues corresponding to the “active” eigenfunctions of f .

We want to proceed analogously to (2.3). Thus, we employ a linear functional $F: V \rightarrow \mathbb{C}$, which is non-vanishing on the eigenfunctions v_j of the linear operator A , i.e., $F(v_j) \neq 0$, and consider the following equation for $m \in \mathbb{N}_0$

$$\begin{aligned} \sum_{k=0}^M p_k F(A^{k+m} f) &= \sum_{k=0}^M p_k F\left(A^{k+m} \sum_{j=1}^M c_j v_j\right) = \sum_{k=0}^M p_k F\left(\sum_{j=1}^M c_j \underbrace{A^{k+m} v_j}_{=\lambda_j^{k+m} v_j}\right) \\ &= \sum_{k=0}^M p_k \sum_{j=1}^M c_j \lambda_j^k \lambda_j^m F(v_j) = \sum_{j=1}^M c_j F(v_j) \lambda_j^m \underbrace{\sum_{k=0}^M p_k \lambda_j^k}_{=P(\lambda_j)=0} = 0. \end{aligned}$$

Again, we can employ the fact that the Prony polynomial is monic and thus we obtain

$$\sum_{k=0}^{M-1} p_k F(A^{k+m} f) = -F(A^{M+m} f) \quad \forall m \in \mathbb{N}_0.$$

We formulate this as a Hankel system for $m = 0, \dots, M-1$ and get

$$\underbrace{\begin{pmatrix} F(f) & F(Af) & \dots & F(A^{M-1}f) \\ F(Af) & F(A^2f) & \dots & F(A^Mf) \\ \vdots & \vdots & & \vdots \\ F(A^{M-1}f) & F(A^Mf) & \dots & F(A^{2M-2}f) \end{pmatrix}}_{=: \mathbf{H}_M} \cdot \begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_{M-1} \end{pmatrix} = - \begin{pmatrix} F(A^Mf) \\ F(A^{M+1}f) \\ \vdots \\ F(A^{2M-1}f) \end{pmatrix}. \quad (2.10)$$

We can again diagonalize the Hankel matrix using Vandermonde matrices in the following way

$$\underbrace{\begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_M \\ \vdots & \vdots & & \vdots \\ \lambda_1^{M-1} & \lambda_2^{M-1} & \dots & \lambda_M^{M-1} \end{pmatrix}}_{=: \mathbf{V}_M} \cdot \underbrace{\begin{pmatrix} c_1F(v_1) & 0 & \dots & 0 \\ 0 & c_2F(v_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_MF(v_M) \end{pmatrix}}_{=: \mathbf{D}} \cdot \underbrace{\begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{M-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{M-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_M & \dots & \lambda_M^{M-1} \end{pmatrix}}_{=: \mathbf{V}_M^T}. \quad (2.11)$$

Since the eigenvalues λ_j for $j = 1, \dots, M$ are assumed to be pairwise distinct, the Vandermonde matrix \mathbf{V}_M is invertible. The diagonal matrix \mathbf{D} is invertible as it contains only nonzero entries on its diagonal, since $c_j \neq 0 \neq F(v_j)$ for $j = 1, \dots, M$.

Once we have computed the coefficients p_k , $k = 0, \dots, M-1$, of the Prony polynomial (2.9), we can calculate its roots λ_j for $j = 1, \dots, M$, which are eigenvalues of the linear operator A . Using the fact that there exists a one-to-one correspondence between the eigenvalues λ_j and eigenfunctions v_j , we can derive the ‘‘active’’ eigenfunctions in the M -sparse expansion f .

At last, we again need to compute the parameters c_j for $j = 1, \dots, M$. This is done by solving the linear system

$$F(A^k f) = \sum_{j=1}^M c_j \lambda_j^k F(v_j) \quad \text{for } k = 0, \dots, 2M-1. \quad (2.12)$$

We summarize these results in the following theorem.

Theorem 2.5 *Let f be an M -sparse expansion into eigenfunctions of the linear operator $A: V \rightarrow V$ as in (2.8). Let $F: V \rightarrow \mathbb{C}$ be a linear functional such that $F(v_j) \neq 0$ for all eigenfunctions v_j of the operator A . Then f can be uniquely reconstructed using only the $2M$ values $F(A^k f)$ for $k = 0, \dots, 2M-1$, i.e., we can uniquely determine the parameters c_j and the ‘‘active’’ eigenfunctions v_j of f .*

This yields the following algorithm.

Algorithm 2 The Generalized Prony Method

Input: $M \in \mathbb{N}$ and $F(A^k f)$ for $k = 0, \dots, 2M - 1$

- 1: Form the Hankel matrix $\mathbf{H}_M := \left(F(A^{(k+\ell)} f) \right)_{k,\ell=0}^{M-1}$ as well as $\mathbf{f} := \left(F(A^{(M+\ell)} f) \right)_{\ell=0}^{M-1}$ and solve $\mathbf{H}_M \cdot \mathbf{p} = -\mathbf{f}$ for $\mathbf{p} := (p_\ell)_{\ell=0}^{M-1}$ as in (2.10).
- 2: Define the Prony polynomial $P(z) := \sum_{k=0}^M p_k z^k$ with $p_M = 1$ as in (2.9), find all roots $\lambda_j, j = 1, \dots, M$, of P and determine the corresponding eigenfunctions v_j for $j = 1, \dots, M$.
- 3: Solve the Vandermonde-like system $\mathbf{V}_{2M,M} \cdot \tilde{\mathbf{c}} = \mathbf{f}_{2M}$ with $\mathbf{V}_{2M,M} := \left(\lambda_j^k \right)_{k=0,j=1}^{2M-1,M}$, $\tilde{\mathbf{c}} := (c_j F(v_j))_{j=1}^M$ and $\mathbf{f}_{2M} := (F(A^k f))_{k=0}^{2M-1}$ as in (2.12) and compute the coefficients $c_j = \frac{\tilde{c}_j}{F(v_j)}$ for $j = 1, \dots, M$.

Output: c_j and v_j for $j = 1, \dots, M$ as in (2.8).

Remark 2.6 The generalized operator based Prony method as derived in [Sta18,SP20] - connected to [PSK19] - is a further theoretical framework in order to systematically obtain new reconstruction schemes for sparse expansions into eigenfunctions of special linear operators with two essential goals.

First off, this method aims to derive simple reconstruction schemes for more general expansions. The second goal is to obtain a variety of different sample sets that admit the recovery of these expansions. In particular, this method introduced a generalization of the linear sampling function F .

Furthermore, in 2020 an algebraic framework has been proposed which generalizes several variants of Prony's method and explains their relationships towards one another, see [KRv20]. ◇

3 The Reconstruction of Generalized Exponential Sums and Generalized Trigonometric Sums

In the last chapter we have seen that the classical Prony method is only a special case of the generalized Prony method using the shift operator $S_1: C(\mathbb{R}) \rightarrow C(\mathbb{R})$ with $S_1(f) := f(\cdot + 1)$. Indeed, there is no need to restrict ourselves to the shift operator with the shift parameter 1. In fact, we can choose a more general approach and define the **shift operator with the shift parameter h** for $h \in \mathbb{R} \setminus \{0\}$ as

$$\begin{aligned} S_h: C(\mathbb{R}) &\rightarrow C(\mathbb{R}) \\ f &\mapsto f(\cdot + h). \end{aligned} \tag{3.1}$$

Thus, we can take a first step towards the generalization of shift operators and choose a starting point $x_0 \in \mathbb{R}$, and a sampling distance $h \in \mathbb{R} \setminus \{0\}$, and according to Theorem 2.5, we can recover the exponential sum f in (2.1) using the samples $f(x_0 + hk)$ for $k = 0, \dots, 2M - 1$. Here, we employ the Hankel matrix

$$\mathbf{H}_M = (f(x_0 + h(k + m)))_{k,m=0}^{M-1}$$

and assume that $|\operatorname{Im}(\alpha_j)| < |\frac{\pi}{h}|$ holds for all $j = 1, \dots, M$.

In this chapter we want to explore generalizations of the classical shift operator S_h as in (3.1) and derive connections to exponential operators. Furthermore, we want to identify which functions can be reconstructed using Prony's method and generalized shift operators.

Thus, a first generalization of the shift operator is studied in Section 3.1. This will lead to a Prony based recovery method for sparse trigonometric expansions.

In Section 3.2 we further generalize the shift operator based on the theory of exponential operators and semigroups. In particular, the so called generalized exponential sums are introduced in Section 3.2.1 as eigenfunction of generalized shift operators. Moreover, generalized trigonometric sums and their recovery are studied in Section 3.2.2. In Section 3.2.3 we use the fact that exponential functions $e^{\alpha_j x}$ are not only eigenfunctions of the shift operator S_h but also of the ordinary differential operator $\frac{d}{dx}$ in order to derive linear differential operators such that generalized exponential sums can be recovered using Prony's method and these differential operators. Addition-

ally, we will explain the connection between the linear differential operators and the generalized shift operators.

Section 3.3 is devoted to illustrating our results numerically. This chapter is based on our papers [PSK19], [KPS19] and parts of our survey [KP21].

We start by taking a closer look at the reconstruction of expansions into trigonometric functions, i.e., functions of the form

$$f(x) = \sum_{j=1}^M c_j \cos(\alpha_j x) \quad \text{or} \quad f(x) = \sum_{j=1}^M c_j \sin(\alpha_j x) \quad (3.2)$$

for $c_j \in \mathbb{C} \setminus \{0\}$ and $\alpha_j \in \mathbb{R}$ for $j = 1, \dots, M$, and construct a suitable shift operator for the recovery of such functions.

3.1 Reconstruction of Trigonometric Expansions

In the following we will focus on a sparse cosine expansion f , as analogue arguments can be made for sparse sine expansions as well as sparse hyperbolic expansions. Using the fact that $\cos(\alpha x) = \frac{1}{2}(e^{i\alpha x} + e^{-i\alpha x})$ holds, we can always rewrite f as an exponential sum of length $2M$ and use the classical Prony method and $4M$ functional values in order to recover the parameters c_j as well as α_j for $j = 1, \dots, M$. Moreover, this also motivates a first idea for a generalized shift operator, namely a linear combination of the classical shift operator S_h and S_{-h} in order to recover trigonometric functions, see also [PSK19].

Definition 3.1 (Symmetric Shift Operator) Let $h \in \mathbb{R} \setminus \{0\}$ be a shift parameter. Then the **symmetric shift operator** $S_{h,-h}: C(\mathbb{R}) \rightarrow C(\mathbb{R})$ is defined as

$$S_{h,-h}f(x) := \frac{1}{2}(S_h f(x) + S_{-h} f(x)) = \frac{1}{2}(f(x+h) + f(x-h)). \quad (3.3)$$

As the iterative application of the operator is a key ingredient in the generalized Prony method, we are interested in the properties of the symmetric shift operator.

Proposition 3.2 Let $h_1, h_2 \in \mathbb{R} \setminus \{0\}$ be shift parameters and $f \in C(\mathbb{R})$. Then the following is true

$$S_{h_1,-h_1} S_{h_2,-h_2} f = S_{h_2,-h_2} S_{h_1,-h_1} f = \frac{1}{2} \left(S_{h_1+h_2,-(h_1+h_2)} + S_{h_1-h_2,-(h_1-h_2)} \right) f.$$

In particular, we have

$$S_{h,-h}^k f = \frac{1}{2^{k-1}} \sum_{\ell=0}^{\lfloor (k-1)/2 \rfloor} \binom{k}{\ell} S_{(k-2\ell)h,-(k-2\ell)hf} + \delta_{k/2, \lfloor k/2 \rfloor} \frac{1}{2^k} \binom{k}{\lfloor k/2 \rfloor} f \quad (3.4)$$

for $k \in \mathbb{N}_0$.

Proof. Let $f \in C(\mathbb{R})$ and let $h_1, h_2 \in \mathbb{R} \setminus \{0\}$ be shift parameters. Then we find for $x \in \mathbb{R}$

$$\begin{aligned} S_{h_1, -h_1} S_{h_2, -h_2} f(x) &= S_{h_1, -h_1} \frac{1}{2} (f(x + h_2) + f(x - h_2)) \\ &= \frac{1}{2} (S_{h_1, -h_1} f(x + h_2) + S_{h_1, -h_1} f(x - h_2)) \\ &= \frac{1}{4} (f(x + h_2 + h_1) + f(x + h_2 - h_1) + f(x - h_2 + h_1) + f(x - h_2 - h_1)) \\ &= \frac{1}{2} (S_{h_1+h_2, -(h_1+h_2)} f(x) + S_{h_1-h_2, -(h_1-h_2)} f(x)). \end{aligned}$$

Analogously, we can conclude

$$S_{h_2, -h_2} S_{h_1, -h_1} f(x) = S_{h_1, -h_1} S_{h_2, -h_2} f(x).$$

Now let $k \in \mathbb{N}_0$. Then the k -th iteration is given by

$$\begin{aligned} S_{h, -h}^k f &= \left(\frac{1}{2} (S_h + S_{-h}) \right)^k f = \frac{1}{2^k} \sum_{\ell=0}^k \binom{k}{\ell} S_h^{k-\ell} S_{-h}^\ell f \\ &= \frac{1}{2^k} \sum_{\ell=0}^k \binom{k}{\ell} S_{h(k-2\ell)} f \\ &= \frac{1}{2^k} \left[\sum_{\ell=0}^{\lfloor (k-1)/2 \rfloor} \binom{k}{\ell} S_{h(k-2\ell)} f + \sum_{\ell=\lfloor (k-1)/2 \rfloor + 1}^k \binom{k}{\ell} S_{h(k-2\ell)} f + \binom{k}{\lfloor k/2 \rfloor} \delta_{k/2, \lfloor k/2 \rfloor} f \right] \\ &= \frac{1}{2^k} \left[\sum_{\ell=0}^{\lfloor (k-1)/2 \rfloor} \binom{k}{\ell} (S_{h(k-2\ell)} + S_{-h(k-2\ell)}) f + \binom{k}{\lfloor k/2 \rfloor} \delta_{k/2, \lfloor k/2 \rfloor} f \right] \\ &= \frac{1}{2^{k-1}} \sum_{\ell=0}^{\lfloor (k-1)/2 \rfloor} \binom{k}{\ell} S_{(k-2\ell)h, -(k-2\ell)h} f + \frac{1}{2^k} \delta_{k/2, \lfloor k/2 \rfloor} \binom{k}{\lfloor k/2 \rfloor} f. \end{aligned}$$

□

This iteration property yields a connection between the reconstruction of trigonometric functions and Chebyshev polynomials. In the following we will, therefore, introduce some of the most important properties.

Definition 3.3 (Chebyshev Polynomials of First Kind) Let $n \in \mathbb{N}_0$ and $x \in \mathbb{R}$. Then the **Chebyshev polynomial of first kind** of degree $n + 1$ is defined as

$$T_{n+1}(x) := 2^n \prod_{k=0}^n \left(x - \cos \left(\frac{(2k+1)\pi}{2(n+1)} \right) \right) =: \sum_{\ell=0}^{n+1} a_{n+1, \ell} x^\ell$$

with $T_0(x) = 1$.

We summarize some of the most important properties of the Chebyshev polynomials of first kind in the following lemma.

Lemma 3.4 Let $n \in \mathbb{N}_0$ and $x \in \mathbb{R}$.

(i) T_n is a polynomial of degree n .

(ii) The leading coefficient of the Chebyshev polynomial T_n is given by

$$a_{n,n} = \begin{cases} 2^{n-1}, & \text{if } n \geq 1, \\ 1, & \text{if } n = 0. \end{cases}$$

(iii) The n zeros of T_n are given by

$$t_{n,\ell} = \cos\left(\frac{(2\ell+1)\pi}{2n}\right) \text{ for } \ell = 0, \dots, n-1,$$

and are called the **Chebyshev nodes**.

(iv) For $x \in [-1, 1]$, the Chebyshev polynomial T_n can be written as

$$T_n(x) = \cos(n \arccos(x)).$$

(v) The Chebyshev polynomials satisfy the recurrence relation

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

with $T_0(x) = 1$ and $T_1(x) = x$.

(vi) The n -monomial can be expressed in terms of the Chebyshev polynomials in the following way

$$x^n = \frac{1}{2^{n-1}} \sum_{\ell=0}^{\lfloor (n-1/2) \rfloor} \binom{n}{\ell} T_{n-2\ell}(x) + \frac{1}{2^n} \delta_{n/2, \lfloor n/2 \rfloor} \binom{n}{\lfloor n/2 \rfloor} T_0(x). \quad (3.5)$$

Proof. The claims (i), (ii) and (iii) follow directly from Definition 3.3. Proofs of (iv) and (v) can be found in [PPST19], Section 6.1. The proof for (vi) can be found in [FP68], Section 3.6. \square

Definition 3.5 (Chebyshev Vandermonde Matrix) Let $n \in \mathbb{N}$ and $(x_k)_{k=1}^n \in \mathbb{R}^n$. The matrix

$$\mathbf{T}_n = (T_\ell(x_k))_{\ell=0, k=1}^{n-1, n} \in \mathbb{R}^{n \times n} \quad (3.6)$$

is called the **Chebyshev Vandermonde matrix**.

Lemma 3.6 Let $n \in \mathbb{N}$ and $(x_k)_{k=1}^n \in \mathbb{R}^n$ such that $x_j \neq x_k$ for all $j \neq k$ with $j, k = 1, \dots, n$. Then the Chebyshev Vandermonde matrix \mathbf{T}_n is invertible with

$$\det(\mathbf{T}_n) = 2^{(n-1)(n-2)/2} \prod_{1 \leq j < k \leq n} (x_j - x_k).$$

Proof. By Definition 3.3 and Lemma 3.4, the $j - 1$ -th Chebyshev polynomial can be written as $T_{j-1}(x) = 2^{j-2}x^{j-1} + \sum_{\ell=0}^{j-2} a_{j-1,\ell}x^\ell$ and is polynomials of degree $j - 1$ with $j = 1, \dots, n$. Therefore, there exist coefficients $b_k \in \mathbb{R}$ for $k = 0, \dots, n - 2$ such that

$$\sum_{k=0}^{j-2} b_k T_k(x) = \sum_{\ell=0}^{j-2} a_{n-1,\ell} x^\ell =: r_{j-2}(x) \quad (3.7)$$

holds for all $j > 1$. Hence, for each row of \mathbf{T}_n we can subtract the linear combination of the previous rows given in (3.7) and use the multilinearity of the determinant to get

$$\begin{aligned} \det(\mathbf{T}_n) &= \det \begin{pmatrix} T_0(x_1) & T_0(x_2) & \dots & T_0(x_n) \\ T_1(x_1) & T_1(x_2) & \dots & T_1(x_n) \\ T_2(x_1) & T_2(x_2) & \dots & T_3(x_n) \\ \vdots & \vdots & \dots & \vdots \\ T_{n-1}(x_1) & T_{n-1}(x_2) & \dots & T_{n-1}(x_n) \end{pmatrix} \\ &= \det \begin{pmatrix} 1 & & & 1 & \dots & 1 \\ x_1 & & & x_2 & \dots & x_n \\ 2x_1^2 + r_0(x_1) & & & 2x_2^2 + r_0(x_2) & \dots & 2x_n^2 + r_0(x_n) \\ \vdots & & & \vdots & \dots & \vdots \\ 2^{n-2}x_1^{n-1} + r_{n-2}(x_1) & & & 2^{n-2}x_2^{n-1} + r_{n-2}(x_2) & \dots & 2^{n-2}x_n^{n-1} + r_{n-2}(x_n) \end{pmatrix} \\ &= \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ 2x_1^2 & 2x_2^2 & \dots & 2x_n^2 \\ \vdots & \vdots & \dots & \vdots \\ 2^{n-2}x_1^{n-1} & 2^{n-2}x_2^{n-1} & \dots & 2^{n-2}x_n^{n-1} \end{pmatrix} \\ &= 2^{(n-2)(n-1)/2} \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \dots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{pmatrix} \\ &= 2^{(n-2)(n-1)/2} \prod_{1 \leq j < k \leq n} (x_j - x_k). \end{aligned}$$

The factor $2^{(n-2)(n-1)/2}$ appears since the j -th row contains the factors 2^{j-2} for $j > 1$. The last equality holds due to the fact that the Vandermonde determinate $(x_k^j)_{j=0,k=1}^{n-1,n}$ is of the given form, see [Sch02], Section 3.1.2, Lemma 3.1.2. \square

Remark 3.7 We recall that the Vandermonde matrix has a connection to polynomial interpolation in the monomial basis, i.e., the problem of, given $(n + 1)$ data tuples $(x_0, y_0), \dots, (x_n, y_n) \in \mathbb{R}^2$, finding a polynomial P of degree at most n which satisfies

$$P(x_k) = y_k \text{ for } k = 0, \dots, n.$$

If we rewrite this as a linear system, the corresponding matrix is a transposed Vandermonde matrix of size $(n + 1) \times (n + 1)$, which is invertible if the data points x_0, \dots, x_n are pairwise distinct. Consequently, if we consider the problem of polynomial interpolation in the Chebyshev basis, the corresponding linear system involves Chebyshev Vandermonde matrices. Lemma 3.6 ensures that this interpolation problem is, indeed, solvable if the given points x_0, \dots, x_n are pairwise distinct.

Furthermore, this idea can be generalized to the problem of polynomial interpolation for different polynomial basis such as different orthogonal polynomials. \diamond

Proposition 3.8 Let $h \in \mathbb{R} \setminus \{0\}$ be a shift parameter and $S_{h,-h}$ be the symmetric shift operator as in Definition 3.1. Then $S_{h,-h}$ possesses eigenfunctions of the form $\cos(\alpha x)$, $\sin(\alpha x)$ as well as $\cosh(\alpha x)$ and $\sinh(\alpha x)$ for $\alpha \in \mathbb{R}$.

Proof. We show the claim for the trigonometric functions $\cos(\alpha x)$ and $\sin(\alpha x)$. To this end, we use the trigonometric identities $\cos(x \pm y) = \cos(x)\cos(y) \mp \sin(x)\sin(y)$ and $\sin(x \pm y) = \sin(x)\cos(y) \pm \cos(x)\sin(y)$, respectively. This yields

$$\begin{aligned} S_{h,-h}(\cos(\alpha x)) &= \frac{1}{2} (\cos(\alpha(x+h)) + \cos(\alpha(x-h))) \\ &= \frac{1}{2} (\cos(\alpha x)\cos(\alpha h) - \sin(\alpha x)\sin(\alpha h) + \cos(\alpha x)\cos(\alpha h) + \sin(\alpha x)\sin(\alpha h)) \\ &= \cos(\alpha x)\cos(\alpha h) \end{aligned}$$

and

$$S_{h,-h}(\sin(\alpha x)) = \sin(\alpha x)\cos(\alpha h),$$

respectively. Since analogue trigonometric identities hold for $\cosh(\alpha x)$ and $\sinh(\alpha x)$ the claim follows. \square

Theorem 3.9 Let $M \in \mathbb{N}$ and f be an M -sparse cosine expansion, i.e.,

$$f(x) = \sum_{j=1}^M c_j \cos(\alpha_j x) \tag{3.8}$$

for $c_j \in \mathbb{C} \setminus \{0\}$ and $\alpha_j \in \mathbb{R}$ for $j = 1, \dots, M$. Furthermore, let $K \in \mathbb{R}$ such that $\alpha_j \in [0, K)$ for all $j = 1, \dots, M$ and $h \in \mathbb{R} \setminus \{0\}$ be a shift parameter such that $h = \frac{\pi}{K}$. Then f can be uniquely recovered using the $2M$ samples $f(kh)$ for $k = 0, \dots, 2M - 1$. Moreover, let $x_0 \in \mathbb{R}$

be an arbitrary sampling point satisfying $\alpha_j x_0 \neq \frac{\pi(2\ell+1)}{2}$ for all $\ell \in \mathbb{Z}$. Then f can be uniquely recovered using the $4M - 1$ samples $f(x_0 + hk)$ for $k = -2M + 1, \dots, 2M - 1$.

Proof. Equation (3.4) yields that the k -th iteration of the symmetric shift operator $S_{h,-h}$ can be written as a linear combination of the shift operators $S_{h\ell,-h\ell}$ for $\ell = 0, \dots, k$. Therefore, we will use the shifts $S_{h\ell,-h\ell}$ instead of $S_{h,-h}^k$. Furthermore, comparison of (3.5) and (3.4) yields that the linear combination used in both expression is, indeed, the same. Thus, it is advantageous to define the Prony polynomial in terms of the Chebyshev polynomials of first kind, i.e.,

$$P(z) := \prod_{j=1}^M (z - \cos(\alpha_j h)) = \sum_{k=0}^M p_k T_k(z).$$

We note that due to the fact that the leading coefficient of the k -th Chebyshev polynomial is 2^{k-1} , $p_M = \frac{1}{2^{M-1}}$ holds. For the first step we compute the coefficients p_k , $k = 0, \dots, M - 1$, of the Prony polynomial using the given samples. Therefore, we employ the definition of the Prony polynomial as well as Proposition 3.2 and get for $x_0 \in \mathbb{R}$ and the signal f in (3.8)

$$\begin{aligned} & \sum_{k=0}^M p_k (S_{hk,-hk} S_{mh} f(x_0)) \\ &= \frac{1}{2} \sum_{k=0}^M p_k (f(x_0 + (m+k)h) + f(x_0 + (m-k)h)) \\ &= \frac{1}{2} \sum_{k=0}^M p_k \sum_{j=1}^M c_j (\cos(\alpha_j(x_0 + h(m+k))) + \cos(\alpha_j(x_0 + h(m-k)))) \\ &= \sum_{k=0}^M p_k \sum_{j=1}^M c_j \cos(\alpha_j hk) \cos(\alpha_j(x_0 + hm)) \\ &= \sum_{j=1}^M c_j \cos(\alpha_j(x_0 + hm)) \sum_{k=0}^M p_k \cos(\alpha_j hk) \\ &= \sum_{j=1}^M c_j \cos(\alpha_j(x_0 + hm)) \underbrace{P(\cos(\alpha_j h))}_{=0} = 0 \end{aligned}$$

for all $m = 0, \dots, M - 1$. Analogously, we obtain

$$\sum_{k=0}^M p_k (S_{hk,-hk} S_{-mh} f(x_0)) = \sum_{j=1}^M c_j \cos(\alpha_j(x_0 - hm)) \sum_{k=0}^M p_k \cos(\alpha_j hk) = 0$$

for $m = 0, \dots, M - 1$. Now we distinguish between the case $x_0 \neq 0$ and $x_0 = 0$.

1. For $x_0 = 0$ we obtain the linear system

$$\sum_{k=0}^{M-1} p_k (f((m+k)h) + f((m-k)h)) = -\frac{1}{2^{M-1}} (f((m+M)h) + f((m-M)h))$$

for $m = 0, \dots, M-1$. Since f is an even function, it suffices to know the values $f(kh)$ for $k = 0, \dots, 2M-1$ to build this system. The corresponding matrix has Hankel-plus-Toeplitz structure

$$\begin{aligned} \mathbf{H}_M^{\text{Toep}} &= (f((m+k)h) + f((m-k)h))_{m,k=0}^{M-1} \\ &= 2 \left(\sum_{j=1}^M c_j \cos(\alpha_j mh) \cos(\alpha_j hk) \right)_{m,k=0}^{M-1} \\ &= 2 (\cos(\alpha_j mh))_{m=0,j=1}^{M-1,M} \cdot \text{diag}(c_1, \dots, c_M) \cdot (\cos(\alpha_j hk))_{j=1,k=0}^{M,M-1} \\ &= 2\mathbf{T}_M \cdot \text{diag}(c_1, \dots, c_M) \cdot \mathbf{T}_M^T \end{aligned} \quad (3.9)$$

with the Chebyshev Vandermonde matrix \mathbf{T}_M as in Definition 3.5 and nodes $\cos(\alpha_j h)$, $j = 1, \dots, M$. The terms $\cos(\alpha_j h)$, $j = 1, \dots, M$, are pairwise distinct and non-zero by assumption on h . Thus, Lemma 3.6 yields the invertibility of \mathbf{T}_M . The diagonal matrix $\text{diag}(c_1, \dots, c_M)$ is invertible since $c_j \neq 0$ for $j = 1, \dots, M$, by assumption and, therefore, $\mathbf{H}_M^{\text{Toep}}$ is invertible.

2. For $x_0 \neq 0$ we need to take all samples $S_{kh, -kh} S_{mh, -mh} f(x_0)$ into account. We consider

$$\begin{aligned} &\sum_{k=0}^{M-1} p_k (f(x_0 + (m+k)h) + f(x_0 - (m+k)h) + f(x_0 + (m-k)h) + f(x_0 - (m-k)h)) \\ &= -\frac{1}{2^{M-1}} (f(x_0 + (m+M)h) + f(x_0 - (m+M)h) + f(x_0 + (m-M)h) + f(x_0 - (m-M)h)). \end{aligned} \quad (3.10)$$

Similar to (3.9), the factorization of the occurring Hankel-plus-Toeplitz matrix is given by

$$\begin{aligned} \mathbf{H}_M^{\text{Toep}} &= (f(x_0 + (m+k)h) + f(x_0 - (m+k)h) + f(x_0 + (m-k)h) + f(x_0 - (m-k)h))_{k,m=0}^{M-1} \\ &= 4 \left(\sum_{j=1}^M c_j \cos(\alpha_j x_0) \cos(\alpha_j mh) \cos(\alpha_j kh) \right)_{m,k=0}^{M-1} \\ &= 4\mathbf{T}_M \cdot \text{diag}(c_1 \cos(\alpha_1 x_0), \dots, c_M \cos(\alpha_M x_0)) \cdot \mathbf{T}_M^T. \end{aligned}$$

The diagonal matrix $\text{diag}((c_j \cos(\alpha_j x_0))_{j=1}^M)$ is invertible if $c_j \neq 0$ for $j = 1, \dots, M$ and $\cos(\alpha_j x_0) \neq 0$. This is true if $\alpha_j x_0 \neq \frac{(2k+1)\pi}{2}$ for all $k \in \mathbb{Z}$ and $j = 1, \dots, M$, and, thus, satisfied by assumption.

Once the coefficients p_k , $k = 0, \dots, M-1$ of the Prony polynomial have been computed, we can derive its zeros $\cos(\alpha_j h)$, and, thus, α_j for $j = 1, \dots, M$. In the last step

we need to solve the linear system

$$f(x_0 + kh) = \sum_{j=1}^M c_j \cos(\alpha_j(x_0 + kh)), \text{ for } k = 0, \dots, 2M - 1,$$

in order to compute the coefficients c_j for $j = 1, \dots, M$. \square

Remark 3.10 We can also use Theorem 2.5 in order to recover an M -sparse cosine expansion f . In this case, we use the Hankel matrix $\mathbf{H}_M := \left(S_{-h,h}^{k+\ell} f(x_0) \right)_{k,\ell=0}^{M-1,M-1}$ instead of the Hankel-plus-Toeplitz matrices $\mathbf{H}_M^{\text{Toep}}$ used in (3.10). The corresponding Prony polynomial is considered in monomial basis.

One advantage of the approach given in Theorem 3.9 is that we can omit the linear combination of Shift operators given in (3.4). Furthermore, the computation of the roots of a polynomial in monomial basis is sensitive to noise in the coefficients of the polynomial, see [Wil59]. A more stable approach is to consider the Prony polynomial in Chebyshev basis and, therefore, the approach given in the proof of Theorem 3.9 is advantageous. \diamond

We have already noted that the symmetric shift operator possesses both cosine and sine as eigenfunctions. Thus, we can also apply it for the reconstruction for M -sparse sine expansions of the form

$$f(x) = \sum_{j=1}^M c_j \sin(\alpha_j x) \tag{3.11}$$

with $c_j \in \mathbb{C} \setminus \{0\}$ and $\alpha_j \in \mathbb{R} \setminus \{0\}$ for all $j = 1, \dots, M$. Indeed, the reconstruction is similar to the one in (3.8). The major difference is that since sine is an odd function the point $f(0)$ does not yield information for the recovery of (3.11).

Theorem 3.11 *Let $M \in \mathbb{N}$ and f be an M -sparse sine expansion as in (3.11). Let $K \in \mathbb{R}$ such that $\alpha_j \in (0, K)$ for all $j = 1, \dots, M$ and $h \in \mathbb{R} \setminus \{0\}$ be a shift parameter such that $h = \frac{\pi}{K}$. Let $x_0 \in \mathbb{R}$ be an arbitrary sampling point satisfying $\alpha_j x_0 \neq \pi \ell$ for all $\ell \in \mathbb{Z}$. Then f can be uniquely recovered using the $4M - 1$ samples $f(x_0 + hk)$ for $k = -2M + 1, \dots, 2M - 1$. In particular, f can be uniquely recovered using the $2M$ samples $f(kh)$ for $k = 1, \dots, 2M$.*

Proof. The proof is similar to the proof of Theorem 3.9. Theorem 3.8 implies that the eigenvalues corresponding to $\sin(\alpha x)$ are also $\cos(\alpha h)$. Hence, the Prony polynomial is defined analogously as in the proof of Theorem 3.9, i.e.,

$$P(z) := \prod_{j=1}^M (z - \cos(\alpha_j x)) = \sum_{k=0}^M p_k T_k(z)$$

with $p_M = 2^{-M+1}$. This yields the same linear system as in (3.10) with the correspond-

ing Hankel-plus-Toeplitz matrix

$$\begin{aligned} \mathbf{H}_M^{\text{Toep}} &= (f(x_0 + (m+k)h) + f(x_0 - (m+k)h) + f(x_0 + (m-k)h) + f(x_0 - (m-k)h))_{k,m=0}^{M-1} \\ &= 4 \left(\sum_{j=1}^M c_j \sin(\alpha_j x_0) \cos(\alpha_j m h) \cos(\alpha_j k h) \right)_{m,k=0}^{M-1} \\ &= 4 \mathbf{T}_M \cdot \text{diag}((c_j \sin(\alpha_j x_0))_{j=1}^M) \cdot \mathbf{T}_M^T, \end{aligned}$$

where \mathbf{T}_M denotes the Chebyshev Vandermonde matrix as in Definition 3.5. The invertibility follows by Lemma 3.6 and $\sin(\alpha_j x_0) \neq 0$ by assumption on $x_0 \alpha_j \neq \pi k$ for all $k \in \mathbb{Z}$. This is for example satisfied for $x_0 = \frac{\pi}{K} = h$. Thus, the function values $f(x_0 + hk) = f(h(k+1))$ for $k = 0, \dots, 2M-1$, are already sufficient for the reconstruction, since we have

$$f(x_0 - hk) = f(h(1-k)) = \begin{cases} 0, & \text{if } k = 1, \\ -f(h(1-k)), & \text{if } k \geq 2. \end{cases}$$

After computing the coefficients p_k , $k = 0, \dots, M-1$ of the Prony polynomial we derive its roots $\cos(h\alpha_j)$, and, thus, α_j for $j = 1, \dots, M$. Then we can compute the coefficients c_j for $j = 1, \dots, M$ as in (3.11) by solving a linear system using the sample values

$$f(x_0 + kh) = \sum_{j=1}^M c_j \sin(\alpha_j(x_0 + kh)), \quad \text{for } k = 0, \dots, 2M-1.$$

□

We have seen that in Proposition 3.8 the hyperbolic functions $\sinh(\alpha x)$ as well as $\cosh(\alpha x)$, $\alpha \in \mathbb{R}$, are eigenfunctions of the symmetric shift operator. Therefore, we can also prove similar results for sparse hyperbolic expansions.

Theorem 3.12 *Let $M \in \mathbb{N}$, $c_j \in \mathbb{C} \setminus \{0\}$ and $\alpha_j \in \mathbb{R}$ for all $j = 1, \dots, M$. Furthermore, let $K \in \mathbb{R}$ such that $\alpha_j \in [0, K)$ for all $j = 1, \dots, M$ and $h \in \mathbb{R} \setminus \{0\}$ be a shift parameter satisfying $h = \frac{\pi}{K}$.*

(i) *Let f be an M -sparse hyperbolic cosine expansion, i.e.,*

$$f(x) = \sum_{j=1}^M c_j \cosh(\alpha_j x).$$

Then f can be uniquely reconstructed using the samples $f(kh)$ for $k = 0, \dots, 2M-1$. Furthermore, let $x_0 \in \mathbb{R}$ satisfy $\cosh(x_0 \alpha_j) \neq 0$ for all $j = 1, \dots, M$. Then f can be uniquely recovered using the $4M-1$ samples $f(x_0 + kh)$ for $k = -2M+1, \dots, 2M-1$.

(ii) Let f be an M -sparse hyperbolic sine expansion, i.e.,

$$f(x) = \sum_{j=1}^M c_j \sinh(\alpha_j x)$$

with $\alpha_j \neq 0$ for all $j = 1, \dots, M$. Then f can be uniquely reconstructed using the samples $f(kh)$ for $k = 1, \dots, 2M$. Furthermore, let $x_0 \in \mathbb{R}$ satisfy $\sinh(x_0 \alpha_j) \neq 0$ for all $j = 1, \dots, M$. Then f can be uniquely recovered using the $4M - 1$ samples $f(x_0 + kh)$ for $k = -2M + 1, \dots, 2M - 1$.

Remark 3.13 We note that the trigonometric functions cosine and sine satisfy a differential equation of the form

$$y'' + y = 0.$$

We can define the differential operator $Df := -f''(x)$ as well as an evaluation functional F such that $F(\cos(\alpha_j x)) \neq 0$ or $F(\sin(\alpha_j x)) \neq 0$ for $j = 1, \dots, M$ holds and apply Theorem 2.5 in order to reconstruct sparse trigonometric expansion. Similar arguments can be applied for the hyperbolic functions. \diamond

Using Remark 2.2 (ii), we can also apply suitable transforms such as the Laplace transform and respectively its inverse in order to recover certain functions.

Corollary 3.14 Let $M \in \mathbb{N}$, $c_j \in \mathbb{C} \setminus \{0\}$ and $\alpha_j \in \mathbb{R}$ for all $j = 1, \dots, M$. Let f be a signal of the form

$$f(x) = \sum_{j=1}^M c_j \frac{x}{x^2 + \alpha_j^2}.$$

Furthermore, let $K \in \mathbb{R}$ such that $\alpha_j \in [0, K)$ for all $j = 1, \dots, M$, $x_0 \in \mathbb{R}$, and let $h \in \mathbb{R} \setminus \{0\}$ be a shift parameter satisfying $h = \frac{\pi}{K}$. Then f can be uniquely reconstructed using the $4M - 1$ samples $\mathcal{L}^{-1}(f)(x_0 + kh)$ for $k = -2M + 1, \dots, 2M - 1$, where \mathcal{L}^{-1} denotes the inverse Laplace transform given by

$$\mathcal{L}^{-1}(g)(s) := \frac{1}{2\pi i} \lim_{T \rightarrow \infty} \int_{\gamma - iT}^{\gamma + iT} e^{ts} g(t) dt$$

with $\gamma \in \mathbb{R}$ being a vertical contour in the complex plane such that all singularities of g are to the left and g is bounded on the line given by γ .

Proof. We consider the Laplace transform $\mathcal{L}f(s) := \int_0^\infty f(x)e^{-sx}dx$ and apply it to the signal of the form $\sum_{j=1}^M c_j \cos(\alpha_j x)$. This yields

$$\begin{aligned} \mathcal{L}\left(\sum_{j=1}^M c_j \cos(\alpha_j \cdot)\right)(s) &= \int_0^\infty \sum_{j=1}^M c_j \cos(\alpha_j x) e^{-sx} dx \\ &= \sum_{j=1}^M \frac{c_j}{2} \int_0^\infty \left(e^{\alpha_j i x} + e^{-\alpha_j i x}\right) e^{-sx} dx \\ &= \sum_{j=1}^M \frac{c_j}{2} \int_0^\infty e^{x(\alpha_j i - s)} + e^{-x(\alpha_j i - s)} dx \\ &= \sum_{j=1}^M \frac{c_j}{2} \left(\frac{1}{s - \alpha_j i} + \frac{1}{s + \alpha_j i}\right) = \sum_{j=1}^M c_j \frac{s}{s^2 + \alpha_j^2}. \end{aligned}$$

Since the coefficients c_j and α_j , $j = 1, \dots, M$ of f are the same as the coefficients of $\mathcal{L}^{-1}(f)$, we can apply Theorem 3.9 in order to reconstruct $\mathcal{L}^{-1}(f)$ and consequently, the given signal. \square

Analogously, models arising for the Laplace transform of sine, hyperbolic cosine and hyperbolic sine can be recovered using Prony's method.

3.2 Reconstruction of Generalized Exponential Sums and Generalized Trigonometric Sums

In Chapter 2 we have seen that we are able to recover M -sparse expansions of a linear operator A for $M \in \mathbb{N}$ using a generalized version of Prony's method. While some examples of suitable linear operators were given by Peter and Plonka, see [PP13], the sample values needed for the reconstruction are in practice not always accessible. This leads to the following questions:

1. Can we find further meaningful examples, i.e., can we find a class of functions with a corresponding linear operator for which we can apply the generalized Prony method?
2. What kind of information is needed in order to recover these expansions?

In order to tackle the first question we will introduce the notion of generalized exponential functions and generalized exponential sums.

Definition 3.15 (Generalized Exponential Sum) Let $G: \mathbb{R} \rightarrow \mathbb{R}$ and $H: \mathbb{R} \rightarrow \mathbb{C}$ with $G, H \in C(\mathbb{R})$ such that for a given interval $[a, b] \subset \mathbb{R}$ the function G is strictly monotone and the function H is non-vanishing. Then we call f a **generalized exponential**

sum or an M -sparse **generalized exponential** if it is of the form

$$f(x) = \sum_{j=1}^M c_j H(x) e^{\alpha_j G(x)} \quad x \in [a, b] \quad (3.12)$$

with $c_j \in \mathbb{C} \setminus \{0\}$, pairwise different $\alpha_j \in \mathbb{C}$ for $j = 1, \dots, M$ and $M \in \mathbb{N}$. The atoms $e^{\alpha_j G(x)} H(x)$, $j = 1, \dots, M$, of f are called **generalized exponential functions**. The function G is called **phase function**. The function H is called the **instantaneous amplitude function**. Furthermore, the signal f is called **non-stationary** if $H \not\equiv c$ or $G(x) \neq mx + d$ for $c, d, m \in \mathbb{R}$ for all $x \in \mathbb{R}$.

It is easy to see that the classical exponential sum as in (2.1) is a special case of the generalized exponential sum with $H \equiv 1$ and $G(x) = x$.

The following section is devoted to the construction of linear operators for which the generalized exponential sums are eigenfunctions.

3.2.1 Generalized Shift Operators

The fact that the exponential functions are eigenfunction of the classical shift operator S_h as given in (3.1) is the starting point for the construction of linear operators that admit generalized exponentials as eigenfunctions. Therefore, this section elaborates on the connection between shift operators and exponential operators. This will lead us to generalized shift operators, which we have studied in our papers [PSK19, KP21], and which were also investigated in [Sta18, SP20].

First of we begin by taking a closer look at the classical shift operator S_h as in (3.1) and we will recall some notations and theory of operator theory and one-parameter semigroups. See [EN00] and [Paz83] for more information on this topic.

Definition 3.16 (Bounded Operator) Let V be a complex Banach space with corresponding norm $\|\cdot\|$ and $A: V \rightarrow V$ be a linear operator. Then A is called **bounded** if

$$\|A\| := \sup_{\|v\|=1, v \in V} \|Av\| < \infty \quad \text{with } v \in V.$$

Definition 3.17 (Operator Valued Exponential Function) Let V be a complex Banach space with corresponding norm $\|\cdot\|$ and $A: V \rightarrow V$ be bounded linear operator. Then we define the **operator valued exponential function** as

$$e^{tA} = \sum_{k=0}^{\infty} \frac{(tA)^k}{k!} \quad (3.13)$$

for $t \in \mathbb{R}$. In the field of quantum physics this operator is also sometimes called (time) evolution operator.

Remark 3.18 The operator valued exponential function is, indeed, well-defined since A is bounded and

$$\|e^{tA}\| = \left\| \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} \right\| \leq \sum_{k=0}^{\infty} \frac{t^k}{k!} \|A\|^k = e^{t\|A\|} < \infty$$

holds. Furthermore, let $\sigma(A)$ be the **spectrum** of the operator A and $\rho(A) := \mathbb{C} \setminus \sigma(A)$ the **resolvent set** of the operator A . Then we can define the **resolvent**

$$R(\cdot, A): \rho(A) \rightarrow L(V)$$

as

$$R(\lambda, A) := (\lambda \text{Id} - A)^{-1},$$

where $\text{Id}: V \rightarrow V$ denotes the identity operator. This is an analytic map, see [EN00], Chapter IV, Section 1, Proposition 1.3. Using this we get an equivalent definition of the operator valued exponential function using Cauchy's integral formula, see [EN00], Chapter 1, Section 3, Definition 3.4, i.e.,

$$e^{tA} := \frac{1}{2\pi i} \int_{\delta U} e^{t\lambda} R(\lambda, A) d\lambda$$

for all $t \geq 0$ and U being an open neighbourhood of the spectrum $\sigma(A)$ with smooth, positively oriented boundary $+\delta U$. We note that the operator valued exponential function does not depend of the choice of the neighbourhood U . \diamond

These operator valued exponential functions have also been studied in the context of quantum mechanics and in particular as solutions of generalized difference equations, see [DOTV97, Wil67]. Moreover, these operators have a connection to non-linear differential equations, see [Ste84]. Special kinds of differential operators and their corresponding operator valued exponentials have been studied in [DOTV97, DL00].

As for the exponential function, it is easy to see that the operator valued exponential functions fulfil a linearity condition since usage of the Cauchy product yields

$$\begin{aligned} e^{(t+s)A} &= \sum_{k=0}^{\infty} \frac{(t+s)^k A^k}{k!} = \sum_{k=0}^{\infty} \sum_{\ell=0}^k \frac{k!}{(k-\ell)!\ell!} \frac{(t^\ell s^{k-\ell}) A^k}{k!} \\ &= \sum_{k=0}^{\infty} \sum_{\ell=0}^k \frac{t^\ell A^\ell s^{k-\ell} A^{k-\ell}}{\ell! (k-\ell)!} = e^{tA} e^{sA} \end{aligned}$$

for $t, s \in \mathbb{R}$. This motivates the following definition and theorem.

Definition 3.19 (One-Parameter Semigroup) Let V be a complex Banach space with corresponding norm $\|\cdot\|$. We call a family of bounded linear operators $(T(t))_{t \geq 0}$ a **(one-parameter) semigroup** on V or sometimes linear dynamical system if the following is satisfied

- (i) $T(t+s) = T(t)T(s)$ for $t, s \in \mathbb{R}_+$,
- (ii) $T(0) = \text{Id}$.

We call a linear operator A given as

$$Av := \lim_{t \rightarrow 0} \frac{T(t)v - v}{t} \quad (3.14)$$

for all $v \in \text{dom}(A) := \left\{ v \in V \mid \lim_{t \rightarrow 0} \frac{T(t)v - v}{t} \text{ exists} \right\}$ the **infinitesimal generator** of the semigroup $(T(t))_{t \geq 0}$.

Furthermore, we call semigroup $(T(t))_{t \geq 0}$ **strongly continuous** if

$$\lim_{t \rightarrow 0} T(t)v = v$$

holds for all $v \in V$.

As we have seen above it is easy to see, that $(e^{tA})_{t \geq 0}$ is, indeed, a semigroup. In particular, the following theorem derives a correspondence between operator valued exponential functions and semigroups.

Theorem 3.20 *Let V be a normed, complex Banach space with corresponding norm $\|\cdot\|$. The semigroup $(T(t))_{t \geq 0}$ is a uniformly continuous, i.e.,*

$$\lim_{h \rightarrow 0} \|T(t+h) - T(t)\| = 0$$

holds, if and only if its infinitesimal generator A is bounded and we have $T(t) = e^{tA}$.

Proof. The proof is a combination of Theorem 1.2. and Corollary 1.4 in [Paz83]. □

Remark 3.21 Theorem 3.20 implies that the map $e^{\cdot A}: \mathbb{R}_+ \rightarrow L(V)$ with $t \rightarrow T(t)$ is continuous and differentiable. Therefore, we can conclude that the solution to the homogeneous abstract Cauchy problem

$$\begin{cases} \frac{du}{dt} = Au(t) & t > 0 \\ u(0) = v & v \in V \text{ fixed} \end{cases}$$

is given by $u(t) = T(t)v$. The connection between the abstract Cauchy problem and continuous semigroups is elaborated on in [Paz83], Chapter 4.

◇

Unfortunately, not every semigroup is uniformly continuous. In particular, the operators we want to consider in the following are not bounded. Therefore, we do not necessarily have convergence of e^{tA} . Nevertheless, if we restrict ourselves to strongly continuous semigroups, we can still find an interpretation such that $T(t)$ "equals" e^{tA} .

Theorem 3.22 *Let V be a complex Banach space, $(T(t))_{t \geq 0}$ be a strongly continuous semigroup. Then we have*

$$T(t)v = \lim_{\tau \rightarrow 0} e^{tA(\tau)}v$$

for all $v \in V$ with

$$A(\tau)v := \frac{T(\tau)v - v}{\tau}. \quad (3.15)$$

Furthermore, the limit is uniform in t on any bounded interval $[0, K]$.

Proof. For a proof we refer to [Paz83], Theorem 8.1. □

We now want to connect the classical shift operator as in Definition 3.1 with the theory of one-parameter semigroups.

Proposition 3.23 *Let V be the space of all bounded, uniformly continuous functions on \mathbb{R} with the norm $\|f\|_\infty := \sup_{x \in \mathbb{R}} |f(x)|$. Furthermore, let $D: C^1(\mathbb{R}) \rightarrow C(\mathbb{R})$ be the differential operator defined as $Df(x) := \frac{d}{dx}f(x) =: f'(x)$, where f' denotes the first derivative of f for $f \in C^1(\mathbb{R})$. Let S_t be the shift operator as in (3.1) with $t \in \mathbb{R}_+$. Then $(S_t)_{t \geq 0}$ is a strongly continuous semigroup with infinitesimal operator D and $\text{dom}(D) = \{f \in V \mid f \in C^1(\mathbb{R})\}$. Furthermore, the following holds*

$$S_t f = e^{tD} f$$

for all analytic functions f .

Proof. Firstly we note, that V equipped with the supremum norm is a Banach space and begin by proving that $(S_t)_{t \geq 0}$ is a strongly continuous semigroup. The definition of the shift operator in (3.1) yields for all $f \in V$ and $x \in \mathbb{R}$

$$\begin{aligned} S_0 f(x) &= f(x), \\ S_{t+s} f(x) &= f(x+t+s) = f((x+s)+t) = S_t S_s f(x), \\ \lim_{t \rightarrow 0} S_t f(x) &= \lim_{t \rightarrow 0} f(x+t) = f(x), \end{aligned}$$

where the last equality holds since f is continuous. Furthermore, we have

$$\|S_t\| = \sup_{\|f\|=1, f \in V} \|f(x+t)\|_\infty < \infty$$

since f is bounded. Equation (3.14) implies, that the infinitesimal generator is given by

$$\lim_{t \rightarrow 0} \frac{S_t f(x) - f(x)}{t} = \lim_{t \rightarrow 0} \frac{f(x+t) - f(x)}{t} = f'(x) = Df(x)$$

for all $x \in \mathbb{R}$. Hence, we can conclude that $(S_t)_{t \geq 0}$ is a strongly continuous semigroup with infinitesimal generator D . We consider for all $f \in V$ and $x \in \mathbb{R}$

$$A(\tau)f(x) = \frac{S_\tau f(x) - f(x)}{\tau} = \frac{f(x+\tau) - f(x)}{\tau},$$

i.e., $A(\tau)$ is the difference quotient. Together with Theorem 3.22 we obtain

$$S_t f(x) = \lim_{\tau \rightarrow 0} e^{tA(\tau)} f(x) = \lim_{\tau \rightarrow 0} \sum_{k=0}^{\infty} \frac{(tA(\tau))^k}{k!} f(x) = \lim_{\tau \rightarrow 0} \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k(\tau) f(x),$$

where the limit exists uniformly for all $x \in \mathbb{R}$ and $t \in [0, K]$ with $K \in \mathbb{R}_+$. If f is analytic, we have $\lim_{\tau \rightarrow 0} A^k(\tau) f(x) = f^{(k)}(x)$ for all $k \in \mathbb{N}$ and hence, the claim follows. \square

Even though the differential operator D is unbounded, the action of e^{tD} on analytic functions is well-defined and yields the shift operator. In fact, this connection is well-known and even dates back to Lagrange.

Moreover, from Proposition 3.23 the question arises if a more general differential operator of first order may yield a generalized version of the shift operator. This question is answered by the following theorem. The idea of its proof can be found in [DOTV97], Section I.2.

Theorem 3.24 *Let $f: [a, b] \rightarrow \mathbb{R}$ be an analytic function and $D: C^1(\mathbb{R}) \rightarrow C(\mathbb{R})$ be the differential operator defined as $Df(x) = \frac{d}{dx} f(x) = f'(x)$, where f' denotes the first derivative of f for $f \in C^1(\mathbb{R})$ and let $h \in \mathbb{R}_+$. Furthermore, let $g: [a, b] \rightarrow \mathbb{R}$ be an analytic and non-vanishing function and $v: [a, b] \rightarrow \mathbb{R}$ be an analytic function. Then the following holds*

$$e^{hg(x)D} f(x) = f(G^{-1}(G(x) + h)), \quad (3.16)$$

where $G(x) := \int_a^x \frac{1}{g(t)} dt$. Furthermore, we have

$$e^{h(g(x)D + v(x)\text{Id})} f(x) = f(G^{-1}(G(x) + h)) \frac{H(x)}{H(G^{-1}(G(x) + h))} \quad (3.17)$$

with $H(x) := \exp\left(-\int_a^x \frac{v(t)}{g(t)} dt\right)$.

Proof. We note that in order to improve the readability of this proof, we will directly consider $e^{tg(x)D}$ and $e^{tg(x)D + v(x)}$ instead of $\lim_{\tau \rightarrow 0} e^{tA(\tau)}$ and $\lim_{\tau \rightarrow 0} e^{tA(\tau) + v(x)}$ with $A(\tau)f(x) = \frac{f(x+\tau) - f(x)}{\tau}$. This is in fact possible, since all considered functions are

analytic and we can apply the same arguments as in the proof of Proposition 3.23.

We consider $t \in \mathbb{R}_+$ and $A: C^1(\mathbb{R}) \rightarrow C(\mathbb{R})$ be a linear operator such that the action of e^{tA} is well defined on analytic functions f . Then

$$e^{tA}f(x) = f(e^{tA}xe^{-tA})e^{tA} \quad (3.18)$$

holds true since

$$\begin{aligned} e^{tA}f(x) &= e^{tA} \left(\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \right) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} e^{tA} x^k \\ &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \underbrace{e^{tA} x e^{-tA} e^{tA} x e^{-tA} \dots x e^{-tA}}_{k\text{-times}} e^{tA} \\ &= f(e^{tA} x e^{-tA}) e^{tA}. \end{aligned}$$

We now set $A := g(x)D$ and use the fact that $e^{hg(x)D}e^{-hg(x)D} = \text{Id}$ holds. Therefore, we can write

$$e^{hg(x)D}x = e^{hg(x)D}x e^{-hg(x)D} e^{hg(x)D} = \chi_1(h) e^{hg(x)D}$$

with

$$\chi_1(h) := e^{hg(x)D}x e^{-hg(x)D}$$

and $\chi_1(0) = x$. Differentiating χ_1 with respect to h yields

$$\frac{d}{dh}\chi_1(h) = e^{hg(x)D}[g(x)D, x]e^{-hg(x)D}, \quad (3.19)$$

where $[g(x)D, x] := g(x)Dx - xg(x)D$ denote the commutator brackets. Computation of the commutator brackets yields

$$\begin{aligned} [g(x)D, x]f(x) &= g(x) \underbrace{\frac{d}{dx}(xf(x))}_{f(x)+xf'(x)} - xg(x) \underbrace{\frac{d}{dx}f(x)}_{f'(x)} \\ &= g(x)f(x) + g(x)xf'(x) - xg(x)f'(x) = g(x)f(x). \end{aligned}$$

Substituting this in (3.19) and applying (3.18) we get

$$\frac{d}{dh}\chi_1(h) = e^{hg(x)D}g(x)e^{-hg(x)D} = g(\chi_1(h)). \quad (3.20)$$

This is a first order non-linear differential equation with the initial condition $\chi_1(0) = x$.

The solution is given by

$$\chi_1(h) = G^{-1}(G(x) + h) \quad \text{with} \quad G(x) := \int_a^x \frac{1}{g(t)} dt$$

since

$$\frac{d}{dh} G^{-1}(G(x) + h) = (G^{-1})'(G(x) + h) = \frac{1}{G'(G^{-1}(G(x) + h))} = g(G^{-1}(G(x) + h)).$$

Equation (3.18) now implies the first part of the claim.

We use the same technique in order to prove the second part of the theorem and begin by considering

$$e^{h(g(x)D+v(x)\text{Id})} x = e^{h(g(x)D+v(x)\text{Id})} x e^{-h(g(x)D+v(x)\text{Id})} e^{h(g(x)D+v(x)\text{Id})}.$$

We define

$$\chi_2(h) := e^{h(g(x)D+v(x)\text{Id})} x e^{-h(g(x)D+v(x)\text{Id})} \quad \text{with} \quad \chi_2(0) = x$$

and

$$u(h) := e^{h(g(x)D+v(x)\text{Id})} \quad \text{with} \quad u(0) = 1.$$

Taking the derivatives with respect to h and application of (3.18) yields

$$\begin{aligned} \frac{d}{dh} \chi_2(h) &= e^{h(g(x)D+v(x)\text{Id})} [g(x)D + v(x)\text{Id}, x] e^{-h(g(x)D+v(x)\text{Id})} \\ &= e^{h(g(x)D+v(x)\text{Id})} g(x) e^{-h(g(x)D+v(x)\text{Id})} = g(\chi_2(h)), \end{aligned}$$

since the commutator of $g(x)D + v(x)\text{Id}$ and x is given by

$$\begin{aligned} [g(x)D + v(x)\text{Id}, x]f(x) &= (g(x)D + v(x))xf(x) - x(g(x)D + v(x))f(x) \\ &= g(x)xf'(x) + g(x)f(x) + v(x)xf(x) \\ &\quad - xg(x)f'(x) - xv(x)f(x) \\ &= g(x)f(x). \end{aligned}$$

Thus, the derivative yields the same non-linear differential equation as in (3.20) with the same initial condition, and, therefore, the same solution.

Furthermore, we get for the derivative of u with respect to h

$$\begin{aligned} \frac{d}{dh}u(h) &= e^{h(g(x)D+v(x)\text{Id})} (g(x)D + v(x)\text{Id}) \\ &= e^{h(g(x)D+v(x)\text{Id})} g(x)D + e^{h(g(x)D+v(x)\text{Id})} v(x)\text{Id} \\ &= e^{h(g(x)D+v(x)\text{Id})} g(x)D + e^{h(g(x)D+v(x)\text{Id})} v(x) e^{-h(g(x)D+v(x)\text{Id})} e^{h(g(x)D+v(x)\text{Id})} \\ &= e^{h(g(x)D+v(x)\text{Id})} g(x)D + v(\chi_2(h))u(h). \end{aligned}$$

If we do not consider u as an operator but only as a function in h , we can simplify its derivative to

$$\frac{d}{dh}u(h) = v(\chi_2(h))u(h). \quad (3.21)$$

This is a ordinary first-order linear differential equation with the initial condition $u(0) = 1$ and solution

$$u(h) = \frac{H(x)}{H(G^{-1}(G(x) + h))}$$

with $H(x) := \exp\left(-\int_a^x \frac{v(t)}{g(t)} dt\right)$ since

$$\begin{aligned} &\frac{d}{dh} \frac{H(x)}{H(G^{-1}(G(x) + h))} \\ &= -\frac{H(x)}{(H(G^{-1}(G(x) + h)))^2} \cdot \underbrace{H'(G^{-1}(G(x) + h))}_{=H(G^{-1}(G(x)+h))\left(-\frac{v(G^{-1}(G(x)+h)}{g(G^{-1}(G(x)+h)}\right)} \cdot \underbrace{(G^{-1})'(G(x) + h)}_{=g(G^{-1}(G(x)+h))} \\ &= \frac{H(x)}{H(G^{-1}(G(x) + h))} v(G^{-1}(G(x) + h)) \end{aligned}$$

holds. Now let $n \in \mathbb{N}_0$ and we consider the action of the operator $e^{h(g(x)D+v(x)\text{Id})}$ on the monomials and get

$$\begin{aligned} u(h)x^n &= e^{h(g(x)D+v(x)\text{Id})} x^n \\ &= e^{h(g(x)D+v(x)\text{Id})} x e^{-h(g(x)D+v(x)\text{Id})} e^{h(g(x)D+v(x)\text{Id})} x \dots e^{h(g(x)D+v(x)\text{Id})} x e^{-h(g(x)D+v(x)\text{Id})} \\ &\quad \cdot e^{h(g(x)D+v(x)\text{Id})} \\ &= (e^{h(g(x)D+v(x)\text{Id})} x e^{-h(g(x)D+v(x)\text{Id})})^n e^{h(g(x)D+v(x)\text{Id})} = (\chi_2(h))^n u(h). \end{aligned}$$

The same argument as for (3.18), thus, yields

$$e^{h(g(x)D+v(x)\text{Id})} f(x) = f(\chi_2(h))u(h)$$

with χ_2 and u being the solutions to the differential equations (3.20) and (3.21), respectively. \square

Remark 3.25 If f is a signal of the form (3.12) with phase function $G(x) := \int_a^x \frac{1}{g(t)} dt$, then $\frac{1}{g}$ is called the **instantaneous phase function** as this function can be understood as the change in the phase G . \diamond

We want to use the shift action in (3.16) and (3.17) respectively, to motivate the definition for our generalized shift operator.

Definition 3.26 (Generalized Shift Operator) Let $G: \mathbb{R} \rightarrow \mathbb{R}$ and $H: \mathbb{R} \rightarrow \mathbb{C}$ be continuous functions such that for a given interval $[a, b] \subset \mathbb{R}$ the function G is strictly monotone and the function H is non-vanishing. Furthermore, let $h \in \mathbb{R} \setminus \{0\}$ be a shift parameter. Then the **generalized shift operator** $S_{G,H,h}: C([a, b]) \rightarrow C(\mathbb{R})$ is defined as

$$S_{G,H,h}f(x) := \frac{H(x)}{H(G^{-1}(G(x) + h))} f\left(G^{-1}(G(x) + h)\right). \quad (3.22)$$

In particular, we have for $H(x) \equiv c$ for some constant $c \in \mathbb{R} \setminus \{0\}$ and for all $x \in [a, b]$,

$$S_{G,h}f(x) := f\left(G^{-1}(G(x) + h)\right).$$

Similarly, for $G(x) = x$ for all $x \in [a, b]$,

$$S_{H,h}f(x) := \frac{H(x)}{H(x+h)} f(x+h).$$

Remark 3.27 The assumptions on G in Definition 3.26 ensure that G is invertible in the interval $[a, b]$. We can consider generalized shift operators for complex-valued functions $G: \mathbb{R} \rightarrow \mathbb{C}$ if we assume G to be invertible in $[a, b]$ instead of strictly monotone and continuous. \diamond

So far we have established how generalized shift operators can be constructed using operator valued exponential functions of special differential operators. In order to connect these operators to our generalized exponential sums given in (2.8), we take a closer look at their eigenfunctions.

Theorem 3.28 Let $G: \mathbb{R} \rightarrow \mathbb{R}$ and $H: \mathbb{R} \rightarrow \mathbb{C}$ be continuous functions such that for a given interval $[a, b] \subset \mathbb{R}$ the function G is strictly monotone and the function H is non-vanishing. Furthermore, let $h \in \mathbb{R} \setminus \{0\}$ be a shift parameter. Then the generalized shift operator $S_{G,H,h}$ as in Definition 3.26 possesses eigenfunctions of the form

$$e^{\alpha G(x)} H(x)$$

with corresponding eigenvalue $e^{\alpha h}$ for $\alpha \in \mathbb{C}$.

Proof. Let $\alpha \in \mathbb{C}$ and $x \in \mathbb{R}$. We employ the definition of the generalized shift operator $S_{G,H,h}$ and obtain

$$\begin{aligned} S_{G,H,h} \left(H(\cdot) e^{G(\cdot)\alpha} \right) (x) &= \frac{H(x)}{H(G^{-1}(G(x) + h))} H(G^{-1}(G(x) + h)) e^{\alpha G(G^{-1}(G(x) + h))} \\ &= H(x) e^{\alpha(G(x) + h)} = H(x) e^{\alpha G(x)} e^{\alpha h}. \end{aligned}$$

□

The above theorem thus, indeed, proves that the generalized exponential sums are expansions into eigenfunctions of the generalized shift operators given in Definition 3.26. Therefore, we have all tools at hand in order to recover such generalized exponential sums using the generalized Prony method.

As the iterative application of the linear operator is one of the key elements of this method, we want to analyze the action of the iterated generalized shift operator $S_{G,H,h}$.

Theorem 3.29 *Let $G: [a, b] \rightarrow \mathbb{R}$ be continuous and strictly monotone and $H: \mathbb{R} \rightarrow \mathbb{C}$ be continuous, such that $H(x) \neq 0$ for all $x \in [a, b]$. Furthermore, let $S_{G,H,h}$ denote the corresponding generalized shift operator as in Definition 3.26. Then the following holds for all $f \in C(\mathbb{R})$*

$$S_{G,H,h_1} S_{G,H,h_2} f = S_{G,H,h_1+h_2} f$$

for all shift parameters $h_1, h_2 \in \mathbb{R} \setminus \{0\}$ satisfying

$$\begin{aligned} G(x) + h_1, G(x) + h_2, G(x) + h_1 + h_2 &\in [G(a), G(b)] && \text{for } G(a) < G(b) && \text{or} \\ G(x) + h_1, G(x) + h_2, G(x) + h_1 + h_2 &\in [G(b), G(a)] && \text{for } G(b) < G(a). \end{aligned}$$

In particular, we have for $k \in \mathbb{N}_0$

$$S_{G,H,h}^k f = S_{G,H,kh} f$$

for

$$\begin{aligned} G(x) + kh &\in [G(a), G(b)] && \text{for } G(a) < G(b) && \text{or} \\ G(x) + kh &\in [G(b), G(a)] && \text{for } G(b) < G(a). \end{aligned}$$

Proof. We use the definition of the generalized shift operator. Then we have

$$\begin{aligned}
 S_{G,H,h_1}(S_{G,H,h_2})f(x) &= S_{G,H,h_1} \left(\frac{H(\cdot)}{H(G^{-1}(G(\cdot) + h_2))} f(G^{-1}(G(\cdot) + h_2)) \right) (x) \\
 &= \frac{H(x)}{H(G^{-1}(G(x) + h_1))} \frac{H(G^{-1}(G(x) + h_1))}{H(G^{-1}(G(x) + h_1 + h_2))} \\
 &\quad \cdot f(G^{-1}(G(x) + h_1 + h_2)) \\
 &= \frac{H(x)}{H(G^{-1}(G(x) + h_1 + h_2))} f(G^{-1}(G(x) + h_1 + h_2)) \\
 &= S_{G,H,h_1+h_2}f(x) = S_{G,H,h_2}S_{G,H,h_1}f(x).
 \end{aligned}$$

Consequently, we get

$$S_{G,H,h}^k f = S_{G,H,hk} f.$$

□

We list some examples for phase functions G , instantaneous amplitudes H , the eigenfunctions to the corresponding generalized shift operator $S_{G,H,h}$ as well as the required samples of a signal f as in Definition 3.12 in Table 3.1.

inst. phase g	phase G	inst. ampl. H	eigenfct.	sampling values
1	x	1	$e^{\alpha x}$	$f(x + hk)$
x	$\log(x)$	1	x^α	$f(e^{hk}x)$
$\frac{x^{p-1}}{p} (p > 1)$	x^p	1	$e^{\alpha x^p}$	$f(\sqrt[p]{x^p + hk})$
$-\sqrt{1-x^2}$	$\arccos(x)$	1	$e^{\alpha \arccos(x)}$	$f(\cos(\arccos(x) + kh))$
$\sqrt{1-x^2}$	$\arcsin(x)$	1	$e^{\alpha \arcsin(x)}$	$f(\sin(\arcsin(x) + kh))$
$\frac{1}{\cos(x)}$	$\sin(x)$	1	$e^{\alpha \sin(x)}$	$f(\arcsin(\sin(x) + kh))$
$-\frac{1}{\sin(x)}$	$\cos(x)$	1	$e^{\alpha \cos(x)}$	$f(\arccos(\cos(x) + kh))$
1	x	$x^r (r \in \mathbb{R})$	$x^r e^{\alpha x}$	$\frac{x^r}{(x+kh)^r} f(x + kh)$
1	x	$e^{-\beta x^2}$	$e^{-\beta x^2 + \alpha x}$	$e^{\beta h(2x+kh)} f(x + kh)$

Table 3.1: Examples of generalized shift operators of the form $S_{G,H,h}$ as in Definition 3.26 with $G(x) = \int_a^x \frac{1}{g(t)} dt$, their corresponding eigenfunctions $H(\cdot)e^{G(\cdot)}$ and the sample values needed for the reconstruction of a generalized exponential sum as in Definition 3.12.

Remark 3.30 Even though we have established a more general theory around generalized shift operators, in practice only the functions G and H are used. Therefore, it is also possible to consider special piecewise defined generalized exponential sums. We can also consider $G: [a, b] \rightarrow \mathbb{R}$ to be a piecewise continuous and piecewise strictly

monotone function, i.e.,

$$G(x) := \begin{cases} G_1(x) & x \in [a_1, a_2] \\ G_2(x) & x \in (a_2, a_3] \\ \vdots & \\ G_k(x) & x \in (a_k, a_{k+1}], \end{cases}$$

with $a_1 = a < a_2 < \dots < a_{k+1} = b$ and $G_j(x)$ continuous and strictly monotone on $(a_j, a_{j+1}]$ for $j = 1, \dots, k$. \diamond

Theorem 3.31 *Let $G: [a, b] \rightarrow \mathbb{R}$ and $H: [a, b] \rightarrow \mathbb{C}$ be continuous functions such that G is strictly monotone and H is non-vanishing. Furthermore, let f be a generalized exponential expansion as in (3.12), such that there exists a $K \in \mathbb{R}$ with $|\operatorname{Im}(\alpha_j)| < K$ for all $j = 1, \dots, M$. Then f can be uniquely reconstructed using the $2M$ samples $f(G^{-1}(G(x_0) + hk))$ for $k = 0, \dots, 2M - 1$ with $x_0 \in [a, b]$, $h \in \mathbb{R}$ such that $0 < |h| < \frac{\pi}{K}$ and*

$$\begin{aligned} G(x_0) + hk &\in [G(a), G(b)] && \text{for } G(a) < G(b) && \text{or} \\ G(x_0) + hk &\in [G(b), G(a)] && \text{for } G(a) > G(b) \end{aligned}$$

for all $k = 0, \dots, 2M - 1$.

Proof. Let f be a generalized exponential sum. Then Theorem 3.28 implies that f is, indeed, a linear combination of eigenfunctions of the generalized shift operator $S_{G,H,h}$. Applying Theorem 2.5 yields that f can be reconstructed using the functional values $F(S_{G,H,h}^k f)$ for a suitable functional $F: \mathbb{C}(\mathbb{R}) \rightarrow \mathbb{C}$ satisfying $F(H(\cdot)e^{\alpha G(\cdot)}) \neq 0$ for all $\alpha \in \mathbb{C}$. Theorem 3.29 yields $F(S_{G,H,h}^k f) = F(S_{G,H,kh} f)$ for a suitable functional $F: \mathbb{C}(\mathbb{R}) \rightarrow \mathbb{C}$ with $F(H(\cdot)e^{\alpha G(\cdot)}) \neq 0$ for $\alpha \in \mathbb{C}$. Choosing $Ff := f(x_0)$ for $x_0 \in [a, b]$ yields

$$F(S_{G,H,kh} f) = \frac{H(x_0)}{H(G^{-1}(G(x_0) + kh))} f(G^{-1}(G(x_0) + kh)).$$

Since the values $\frac{H(x_0)}{H(G^{-1}(G(x_0) + kh))}$, $k = 0, \dots, 2M - 1$ are independent of the parameters α_j and c_j for $j = 1, \dots, M$, they can be precomputed and consequently the claim follows. \square

Thus, we can reformulate the algorithm for the generalized Prony method for generalized shift operators.

Algorithm 3 Prony's Method in Terms of Generalized Shift Operators

Input: $M \in \mathbb{N}$, $h > 0$, sampled values $f(G^{-1}(h\ell + G(x_0)))$, $\ell = 0, \dots, 2M - 1$ and $G(x_0)$.

1: Form the Hankel matrix $\mathbf{H}_M := (a_{k+\ell} f(G^{-1}(h(\ell + k) + G(x_0))))_{k,\ell=0}^{M-1}$ as well as $\mathbf{f} := (a_{M+\ell} f(G^{-1}(G(x_0) + h(M + \ell))))_{\ell=0}^{M-1}$ with $a_\ell := \frac{H(x_0)}{H(G^{-1}(G(x_0)+h\ell))}$ for $\ell = 0, \dots, 2M - 1$, and compute $\mathbf{H}_M \cdot \mathbf{p} = -\mathbf{f}$ for $\mathbf{p} := (p_\ell)_{\ell=0}^{M-1}$.

2: Define the Prony polynomial $P(z) = \sum_{k=0}^M p_k z^k$ with $p_M = 1$, find all roots $\lambda_j = e^{\alpha_j h}$ and determine the parameters $\alpha_j := \frac{1}{h} \log(\lambda_j)$ for $j = 1, \dots, M$.

3: Solve the Vandermonde-like system $\mathbf{V}_{2M,M} \cdot \tilde{\mathbf{c}} = \mathbf{f}_{2M}$ with $\mathbf{V}_{2M,M} := (\lambda_j^k)_{k=0,j=1}^{2M-1,M}$, $\tilde{\mathbf{c}} := (c_j e^{\alpha_j G(x_0)})_{j=1}^M$ and $\mathbf{f}_{2M} := (\frac{f(G^{-1}(G(x_0)+hk))}{H(G^{-1}(G(x_0)+hk))})_{k=0}^{2M-1}$ and compute the coefficients $c_j = \frac{\tilde{c}_j}{e^{\alpha_j G(x_0)}}$ for $j = 1, \dots, M$.

Output: c_j and α_j for $j = 1, \dots, M$ as in (3.12).

3.2.2 Generalized Symmetric Shift Operators

The generalized shift operators $S_{G,H,h}$ introduced in Section 3.2.1 can be combined in a symmetric way as in Section 3.1. This motivates the following definitions.

Definition 3.32 (Generalized trigonometric expansions) Let $G: \mathbb{R} \rightarrow \mathbb{R}$ and $H: \mathbb{R} \rightarrow \mathbb{C}$ be continuous functions such that for a given interval $[a, b] \subset \mathbb{R}$ the function G is strictly monotone and the function H is non-vanishing. Then we call f an **M -sparse generalized trigonometric expansion** or **generalized trigonometric sum** if it is of the form

$$f(x) = \sum_{j=1}^M c_j H(x) \cos(\alpha_j G(x) + \beta_j) \quad \text{or} \quad f(x) = \sum_{j=1}^M c_j H(x) \sin(\alpha_j G(x) + \beta_j)$$

with $c_j \in \mathbb{C} \setminus \{0\}$, $\beta_j \in (-\pi, \pi)$ and pairwise different $\alpha_j \in \mathbb{C}$ for $j = 1, \dots, M$.

Remark 3.33 These generalized trigonometric sums can also be understood as a special case of a signal decomposition into so-called intrinsic mode functions, which are studied using empirical mode decomposition, see [HSL⁺98]. Furthermore, special types of generalized trigonometric sums, called non-harmonic Fourier sums, given by

$$f(x) = \sum_{j=1}^M c_j \cos(2\pi\alpha_j x + \beta_j)$$

with $c_j, \alpha_j \in \mathbb{R}_+$ and $\beta_j \in (0, 2\pi)$ and $\alpha_j \neq \alpha_k$ for $j, k = 1, \dots, M$ and $j \neq k$, are also analysed in [PPD21]. The authors show that such functions can be reconstructed using Fourier coefficients and rational approximation. \diamond

Definition 3.34 (Generalized Symmetric Shift Operator) Let $G: \mathbb{R} \rightarrow \mathbb{R}$ and $H: \mathbb{R} \rightarrow \mathbb{C}$ be continuous functions such that for a given interval $[a, b]$ the function G is strictly monotone and the function H is non-vanishing. Furthermore, let $h \in \mathbb{R} \setminus \{0\}$ be a shift parameter. Then the **generalized symmetric shift operator** $S_{G,H,h,-h}: C([a, b]) \rightarrow C(\mathbb{R})$ is defined as

$$S_{G,H,h,-h}f(x) := \frac{1}{2} (S_{G,H,h}f(x) + S_{G,H,-h}f(x))$$

where $S_{G,H,h}$ denotes the generalized shift operator as in Definition 3.26.

Analogously, to Proposition 3.2 we can show that the k -th iteration of a generalized symmetric shift operator $S_{G,H,h,-h}$ can be written as a linear combination of generalized shift operators.

Proposition 3.35 Let $f \in C(\mathbb{R})$ and let $G: \mathbb{R} \rightarrow \mathbb{R}$ and $H: \mathbb{R} \rightarrow \mathbb{C}$ be continuous functions such that for a given interval $[a, b]$ the function G is strictly monotone and the function H is non-vanishing. Moreover, let $S_{G,H,h,-h}$ denote the generalized symmetric shift operator. Then the following holds,

$$\begin{aligned} S_{G,H,h_1,-h_1}S_{G,H,h_2,-h_2}f &= S_{G,H,h_2,-h_2}S_{G,H,h_1,-h_1}f \\ &= \frac{1}{2} \left(S_{G,H,h_1+h_2,-(h_1+h_2)} + S_{G,H,h_1-h_2,-(h_1-h_2)} \right) f \end{aligned}$$

for $h_1, h_2 \in \mathbb{R} \setminus \{0\}$ satisfying

$$\begin{aligned} G(x) + h_1, G(x) + h_2, G(x) + h_1 + h_2 &\in [G(a), G(b)] && \text{for } G(a) < G(b) && \text{or} \\ G(x) + h_1, G(x) + h_2, G(x) + h_1 + h_2 &\in [G(b), G(a)] && \text{for } G(b) < G(a). \end{aligned}$$

In particular, we have for $k \in \mathbb{N}_0$

$$S_{G,H,h,-h}^k f = \frac{1}{2^{k-1}} \sum_{\ell=0}^{\lfloor (k-1)/2 \rfloor} \binom{k}{\ell} S_{G,H,(k-2\ell)h,-(k-2\ell)h} f + \delta_{k/2, \lfloor k/2 \rfloor} \frac{1}{2^k} \binom{k}{\lfloor k/2 \rfloor} f \quad (3.23)$$

for $h \in \mathbb{R} \setminus \{0\}$ satisfying

$$\begin{aligned} G(x) + kh &\in [G(a), G(b)] && \text{for } G(a) < G(b) && \text{or} \\ G(x) + kh &\in [G(b), G(a)] && \text{for } G(b) < G(a). \end{aligned}$$

Proof. Application of the definition of the generalized symmetric shift operator yields similarly as in the proof of Proposition 3.2

$$\begin{aligned} S_{G,H,h_1,-h_1}S_{G,H,h_2,-h_2}f(x) &= S_{G,H,h_1,-h_1} \frac{1}{2} (S_{G,H,h_2}f(x) + S_{G,H,-h_2}f(x)) \\ &= \frac{1}{2} \left(S_{G,H,h_1+h_2,-(h_1+h_2)}f(x) + S_{G,H,h_1-h_2,-(h_1-h_2)}f(x) \right). \end{aligned}$$

Now let $k \in \mathbb{N}$. Then the k -th iteration is given by

$$\begin{aligned}
 S_{G,H,h,-h}^k f &= \left(\frac{1}{2} (S_{G,H,h} + S_{G,H,-h}) \right)^k f(x) = \frac{1}{2^k} \sum_{\ell=0}^k \binom{k}{\ell} S_{G,H,h}^{k-\ell} S_{G,H,-h}^\ell f \\
 &= \frac{1}{2^k} \sum_{\ell=0}^k \binom{k}{\ell} S_{G,H,h(k-2\ell)} f \\
 &= \frac{1}{2^k} \left[\sum_{\ell=0}^{\lfloor (k-1/2) \rfloor} \binom{k}{\ell} (S_{G,H,h(k-2\ell)} + S_{G,H,-h(k-2\ell)}) f + \binom{k}{\lfloor k/2 \rfloor} \delta_{k/2, \lfloor k/2 \rfloor} f \right] \\
 &= \frac{1}{2^{k-1}} \sum_{\ell=0}^{\lfloor (k-1/2) \rfloor} \binom{k}{\ell} S_{G,H,(k-2\ell)h, -(k-2\ell)h} f + \frac{1}{2^k} \delta_{k/2, \lfloor k/2 \rfloor} \binom{k}{\lfloor k/2 \rfloor} f.
 \end{aligned}$$

□

In analogy to Proposition 3.8 we can conclude the eigenfunction for the generalized symmetric shift operators $S_{G,H,h,-h}$.

Proposition 3.36 *Let $h \in \mathbb{R} \setminus \{0\}$ be a shift parameter and $G: \mathbb{R} \rightarrow \mathbb{R}$ and $H: \mathbb{R} \rightarrow \mathbb{C}$ be continuous functions such that for a given interval $[a, b] \subset \mathbb{R}$ the function G is strictly monotone, the function H is non-vanishing and $S_{G,H,h,-h}$ denotes the generalized symmetric shift operator as in Definition 3.34. Then $S_{G,H,h,-h}$ possesses the eigenfunctions $H(x) \cos(\alpha G(x))$, $H(x) \sin(\alpha G(x))$ as well as $H(x) \cosh(\alpha G(x))$ and $H(x) \sinh(\alpha G(x))$ for $\alpha \in \mathbb{R}$.*

Proof. The claim is a direct implication of Theorem 3.28 and

$$\cos(G(x)\alpha) = \frac{1}{2} \left(e^{iG(x)\alpha} + e^{-iG(x)\alpha} \right)$$

as well as similar equations for $\sin(G(x)\alpha)$, $\cosh(\alpha G(x))$ $\sinh(\alpha G(x))$. We have

$$\begin{aligned}
 S_{G,H,h,-h} H(x) \cos(G(x)\alpha) &= \frac{1}{2} (S_{G,H,h} H(x) \cos(G(x)\alpha) + S_{G,H,-h} H(x) \cos(G(x)\alpha)) \\
 &= \frac{1}{4} \left(S_{G,H,h} \left(H(x) e^{iG(x)\alpha} + H(x) e^{-iG(x)\alpha} \right) \right. \\
 &\quad \left. + S_{G,H,-h} \left(H(x) e^{iG(x)\alpha} + H(x) e^{-iG(x)\alpha} \right) \right) \\
 &= \frac{1}{2} \left(H(x) e^{i\alpha G(x)} \cos(\alpha h) + H(x) e^{-i\alpha G(x)} \cos(\alpha h) \right) \\
 &= H(x) \cos(G(x)\alpha) \cos(\alpha h).
 \end{aligned}$$

Hence, the claim follows. □

Analogously, to Section 3.1 we can now consider generalized cosine expansion.

Theorem 3.37 Let $G: \mathbb{R} \rightarrow \mathbb{R}$ and $H: \mathbb{R} \rightarrow \mathbb{C}$ be continuous functions such that for a given interval $[a, b]$ the function G is strictly monotone and the function H is non-vanishing and $S_{G,H,h,-h}$ denotes the generalized symmetric shift operator as in Definition 3.34 for $h \in \mathbb{R} \setminus \{0\}$. Let $M \in \mathbb{N}$ and f be a M -sparse generalized cosine expansion, i.e.,

$$f(x) = \sum_{j=1}^M c_j H(x) \cos(G(x)\alpha_j + \beta_j) \quad (3.24)$$

for $c_j \in \mathbb{R} \setminus \{0\}$, $\beta_j \in (-\pi, \pi) \setminus \{\pm \frac{\pi}{2}\}$ and $\alpha_j \in \mathbb{R}$ for $j = 1, \dots, M$. Furthermore, let $\alpha_j \in [0, K]$ for all $j = 1, \dots, M$ and some $K > 0$. Then the signal f can be reconstructed using the samples $f(G^{-1}(G(x_0) + hk))$ for $k = -2M + 1, \dots, 2M + 1$ with $x_0 \in [a, b]$ satisfying $\cos(\alpha_j G(x_0) + \beta_j) \neq 0$ for $j = 1, \dots, M$ and $0 < h \leq \frac{\pi}{K}$ satisfying

$$\begin{aligned} G(x_0) + hk &\in [G(a), G(b)] && \text{for } G(a) < G(b) && \text{or} \\ G(x_0) + hk &\in [G(b), G(a)] && \text{for } G(a) > G(b) \end{aligned}$$

for $k = -2M + 1, \dots, 2M - 1$.

Moreover, if $x_0 \in [a, b]$ can be chosen such that $G(x_0) = 0$ is satisfied and $\beta_j = 0$ for all $j = 1, \dots, M$, the samples $f(G^{-1}(G(x_0) + hk))$ for $k = 0, \dots, 2M - 1$ are sufficient for the reconstruction of the signal f .

Proof. The proof is done analogously to the proof of Theorem 3.9. Equation (3.23) yields, that the k -th iteration of the generalized symmetric shift operator $S_{G,H,h,-h}$ can be written as a linear combination of the generalized symmetric shifts $S_{G,H,h\ell,-h\ell}$ for $\ell = 0, \dots, k$. Therefore, we will use the shifts $S_{G,H,h\ell,-h\ell} f(x_0)$ instead of $S_{G,H,h,-h}^\ell f(x_0)$. We define the Prony polynomial

$$P(z) := \prod_{j=1}^M (z - \cos(\alpha_j h)) = \sum_{k=0}^M p_k T_k(z).$$

We note that due to the fact that the leading coefficient of the k -th Chebyshev polynomial is 2^{k-1} , $p_M = \frac{1}{2^{M-1}}$ holds. First we compute the coefficients p_k , $k = 0, \dots, M - 1$, of the Prony polynomial using the given samples. Moreover, Proposition 3.36 implies

$$S_{G,H,h,-h} H(x) \cos(\alpha G(x) + \beta) = H(x) \cos(\alpha G(x) + \beta) \cos(\alpha h)$$

for $\alpha \in \mathbb{R}$ and $\beta \in (-\pi, \pi)$. Hence, we can conclude

$$S_{G,H,h,-h}^k f(x_0) = \sum_{j=1}^M c_j H(x_0) \cos(\alpha_j G(x_0) + \beta_j) \cos(\alpha_j k h)$$

for $k \in \mathbb{Z}$. Thus, we employ the definition of the Prony polynomial as well as Propo-

sition 3.35 and get for the signal f in (3.24) and for $m = 0, \dots, M-1$

$$\begin{aligned}
 & \sum_{k=0}^M p_k \left(S_{G,H,h(k+m),-h(k+m)} f(x_0) + S_{G,H,h(m-k),-h(m-k)} f(x_0) \right) \\
 &= \sum_{k=0}^M p_k \left(\sum_{j=1}^M c_j H(x_0) \cos(\alpha_j G(x_0) + \beta_j) \cos(\alpha_j h(k+m)) \right. \\
 & \quad \left. + \sum_{j=1}^M c_j H(x_0) \cos(\alpha_j G(x_0) + \beta_j) \cos(\alpha_j h(m-k)) \right) \\
 &= 2 \sum_{k=0}^M p_k \sum_{j=1}^M H(x_0) c_j \cos(\alpha_j G(x_0) + \beta_j) \cos(\alpha_j h k) \cos(\alpha_j m h) \\
 &= 2 \sum_{j=1}^M H(x_0) c_j \cos(\alpha_j G(x_0) + \beta_j) \cos(\alpha_j m h) \sum_{k=0}^M p_k \cos(\alpha_j h k) \\
 &= 2 \sum_{j=1}^M H(x_0) c_j \cos(\alpha_j G(x_0) + \beta_j) \cos(\alpha_j m h) \underbrace{P(\cos(\alpha_j h))}_{=0} = 0.
 \end{aligned}$$

Thus, we get the Hankel-plus-Toeplitz system

$$\begin{aligned}
 & \sum_{k=0}^{M-1} p_k \left(S_{G,H,h(m+k),-h(m+k)} f(x_0) + S_{G,H,h(m-k),-h(m-k)} f(x_0) \right) \\
 &= -\frac{1}{2^{M-1}} \left(S_{G,H,h(m+M),-h(m+M)} f(x_0) + S_{G,H,h(m-M),-h(m-M)} f(x_0) \right)
 \end{aligned}$$

for $m = 0, \dots, M-1$.

The factorization of the occurring Hankel-plus-Toeplitz matrix is given by

$$\begin{aligned}
 \mathbf{H}_M^{\text{Toep}} &= \left(S_{G,H,h(m+k),-h(m+k)} f(x_0) + S_{G,H,h(m-k),-h(m-k)} f(x_0) \right)_{k,m=0}^{M-1} \\
 &= 2 \left(\sum_{j=1}^M H(x_0) c_j \cos(\alpha_j G(x_0) + \beta_j) \cos(\alpha_j m h) \cos(\alpha_j k h) \right)_{m,k=0}^{M-1} \\
 &= 2 \mathbf{T}_M \cdot \text{diag}(H(x_0) c_j \cos(\alpha_j G(x_0) + \beta_j))_{j=1}^M \cdot \mathbf{T}_M^T,
 \end{aligned}$$

with the Chebyshev Vandermonde matrix \mathbf{T}_M as in Definition 3.5. The diagonal matrix $\text{diag} \left((H(x_0) c_j \cos(\alpha_j G(x_0) + \beta_j))_{j=1}^M \right)$ is invertible if $H(x_0) c_j \neq 0$ for $j = 1, \dots, M$ and $\cos(\alpha_j G(x_0) + \beta_j) \neq 0$, which is satisfied by assumption.

If the parameters β_j for $j = 1, \dots, M$ do not appear in the signal f and $x_0 \in [a, b]$ can be chosen such that $G(x_0) = 0$ is satisfied, we can use the fact that cosine is an even

function and only consider the system

$$\begin{aligned}
 & \sum_{k=0}^M p_k \left(\frac{H(x_0)}{H(G^{-1}(h(m+k)))} f(G^{-1}(h(m+k))) + \frac{H(x_0)}{H(G^{-1}(h(m-k)))} f(G^{-1}(h(m-k))) \right) \\
 &= \sum_{k=0}^M p_k \sum_{j=1}^M c_j H(x_0) (\cos(\alpha_j h(m+k)) + \cos(\alpha_j h(m-k))) \\
 &= 2 \sum_{k=0}^M p_k H(x_0) \sum_{j=1}^M c_j (\cos(\alpha_j h k) \cos(h m \alpha_j)) = 0.
 \end{aligned}$$

Thus, the corresponding Hankel-plus-Toeplitz matrix is of the form

$$\begin{aligned}
 \mathbf{H}_M^{\text{Toep}} &= \left(\frac{H(x_0)}{H(G^{-1}(h(m+k)))} f(G^{-1}(h(m+k))) + \frac{H(x_0)}{H(G^{-1}(h(m-k)))} f(G^{-1}(h(m-k))) \right)_{m,k=0}^{M-1} \\
 &= 2 \left(\sum_{j=1}^M c_j H(x_0) \cos(\alpha_j m h) \cos(\alpha_j h k) \right)_{m,k=0}^{M-1} \\
 &= 2 \mathbf{T}_M \cdot \text{diag}(c_j H(x_0)_{j=1}^M) \cdot \mathbf{T}_M^T.
 \end{aligned}$$

Once the coefficients p_k , $k = 0, \dots, M-1$ of the Prony polynomial have been computed, we can derive its zeros $\cos(\alpha_j h)$, and, thus, α_j for $j = 1, \dots, M$. Now we need to distinguish between the case $\beta_j = 0$ for all $j = 1, \dots, M$ and $\beta_j \neq 0$ for at least one $j \in \{1, \dots, M\}$.

1. If the parameters β_j do not appear in the signal f given in (3.24) we only need to solve the linear system for $k = 0, \dots, 2M-1$

$$\frac{H(x_0)}{H(G^{-1}(G(x_0) + kh))} f(G^{-1}(G(x_0) + kh)) = \sum_{j=1}^M c_j H(x_0) \cos(\alpha_j (G(x_0) + kh)),$$

in order to compute the coefficients c_j for $j = 1, \dots, M$.

2. If there exists a $j \in \{1, \dots, M\}$ such that $\beta_j \neq 0$, we first solve

$$\begin{aligned}
 & \frac{H(x_0)}{H(G^{-1}(G(x_0) + kh))} f(G^{-1}(G(x_0) + kh)) + \frac{H(x_0)}{H(G^{-1}(G(x_0) - kh))} f(G^{-1}(G(x_0) - kh)) \\
 &= 2 \sum_{j=1}^M c_j H(x_0) \cos(\alpha_j G(x_0) + \beta_j) \cos(\alpha_j h k)
 \end{aligned}$$

for $k = 0, \dots, M-1$ in order to compute $d_j := c_j H(x_0) \cos(\alpha_j G(x_0) + \beta_j)$ for $j = 1, \dots, M$. Moreover, we also solve

$$\begin{aligned}
 & \frac{H(x_0)}{H(G^{-1}(G(x_0) + kh))} f(G^{-1}(G(x_0) + kh)) - \frac{H(x_0)}{H(G^{-1}(G(x_0) - kh))} f(G^{-1}(G(x_0) - kh)) \\
 &= 2 \sum_{j=1}^M c_j H(x_0) \sin(\alpha_j G(x_0) + \beta_j) \sin(\alpha_j h k)
 \end{aligned}$$

for $k = 0, \dots, M-1$ in order to compute $\tilde{d}_j := H(x_0)c_j \sin(\alpha_j G(x_0) + \beta_j)$ for $j = 1, \dots, M$. Finally, we can conclude for the parameters c_j and β_j for $j = 1, \dots, M$

$$c_j = \frac{\sqrt{d_j^2 + \tilde{d}_j^2}}{H(x_0)} \quad \text{and} \quad \beta_j = \arg(d_j + i\tilde{d}_j) - \alpha_j G(x_0) \bmod 2\pi.$$

Analogously, to Section 3.1 we can also formulate a similar statement for generalized sine expansions. \square

Theorem 3.38 *Let $G: \mathbb{R} \rightarrow \mathbb{R}$ and $H: \mathbb{R} \rightarrow \mathbb{C}$ be continuous functions such that for a given interval $[a, b]$ the function G is strictly monotone and the function H is non-vanishing and $S_{G,H,h,-h}$ denote the generalized symmetric shift operator as in Definition 3.34 for $h \in \mathbb{R} \setminus \{0\}$. Let $M \in \mathbb{N}$ and f be an M -sparse generalized sine expansion, i.e.,*

$$f(x) = \sum_{j=1}^M c_j H(x) \sin(G(x)\alpha_j + \beta_j) \quad (3.25)$$

for $c_j \in \mathbb{R} \setminus \{0\}$, $\beta_j \in (-\pi, \pi) \setminus \{\pm\frac{\pi}{2}\}$ and $\alpha_j \in \mathbb{R}$ for $j = 1, \dots, M$ with $M \in \mathbb{N}$. Furthermore, let $\alpha_j \in [0, K)$ for all $j = 1, \dots, M$ and some $K > 0$. Then the signal f can be reconstructed using the samples $f(G^{-1}(G(x_0) + hk))$ for $k = -2M+1, \dots, 2M+1$ with $x_0 \in [a, b]$ satisfying $\sin(\alpha_j G(x_0) + \beta_j) \neq 0$ for $j = 1, \dots, M$ and $0 < h \leq \frac{\pi}{K}$ satisfying

$$\begin{aligned} G(x_0) + hk &\in [G(a), G(b)] && \text{for } G(a) < G(b) && \text{or} \\ G(x_0) + hk &\in [G(b), G(a)] && \text{for } G(a) > G(b) \end{aligned}$$

for $k = -2M+1, \dots, 2M-1$.

Proof. The proof is done analogously to the one of Theorem 3.37. We define the Prony polynomial

$$P(z) := \prod_{j=1}^M (z - \cos(\alpha_j h)) = \sum_{k=0}^M p_k T_k(z)$$

with $p_M = \frac{1}{2^{M-1}}$ holds. As a first step we compute the coefficients p_k , $k = 0, \dots, M-1$ of the Prony polynomial using the given samples. Proposition 3.36 implies

$$S_{G,H,h,-h} H(x) \sin(\alpha G(x) + \beta) = H(x) \sin(\alpha G(x) + \beta) \cos(\alpha h)$$

for $\alpha \in \mathbb{R}$ and $\beta \in (-\pi, \pi)$. Hence, we can conclude

$$S_{G,H,h,-h} f(x_0) = \sum_{j=1}^M c_j H(x_0) \sin(\alpha_j G(x_0) + \beta_j) \cos(\alpha_j h)$$

for $k \in \mathbb{Z}$. Thus, we employ the definition of the Prony polynomial as well as Propo-

sition 3.35 and get for the signal f in (3.24) and for $m = 0, \dots, M-1$

$$\begin{aligned} & \sum_{k=0}^M p_k \left(S_{G,H,h(k+m),-h(k+m)} f(x_0) + S_{G,H,h(m-k),-h(m-k)} f(x_0) \right) \\ &= 2 \sum_{j=1}^M H(x_0) c_j \sin(\alpha_j G(x_0) + \beta_j) \cos(\alpha_j m h) \sum_{k=0}^M p_k \cos(\alpha_j h k) \\ &= 2 \sum_{j=1}^M H(x_0) c_j \sin(\alpha_j G(x_0) + \beta_j) \cos(\alpha_j m h) \underbrace{P(\cos(\alpha_j h))}_{=0} = 0. \end{aligned}$$

We get the Hankel-plus-Toeplitz system

$$\begin{aligned} & \sum_{k=0}^{M-1} p_k \left(S_{G,H,h(m+k),-h(m+k)} f(x_0) + S_{G,H,h(m-k),-h(m-k)} f(x_0) \right) \\ &= \frac{1}{2^{M-1}} \left(S_{G,H,h(m+M),-h(m+M)} f(x_0) + S_{G,H,h(m-M),-h(m-M)} f(x_0) \right) \end{aligned}$$

for $m = 0, \dots, M-1$. The factorization of the occurring Hankel-plus-Toeplitz matrix is given by

$$\begin{aligned} \mathbf{H}_M^{\text{Toep}} &= \left(S_{G,H,h(m+k),-h(m+k)} f(x_0) + S_{G,H,h(m-k),-h(m-k)} f(x_0) \right)_{k,m=0}^{M-1} \\ &= 2 \mathbf{T}_M \cdot \text{diag}((c_j H(x_0) \sin(\alpha_j G(x_0) + \beta_j))_{j=1}^M) \cdot \mathbf{T}_M^T. \end{aligned}$$

The diagonal matrix $\text{diag}(H(x_0) c_j \sin(\alpha_j G(x_0) + \beta_j))_{j=1}^M$ is invertible if $H(x_0) c_j \neq 0$ for $j = 1, \dots, M$ and $\sin(\alpha_j G(x_0) + \beta_j) \neq 0$, which is satisfied by assumption.

Once the coefficients p_k , $k = 0, \dots, M-1$ of the Prony polynomial have been computed, we can derive its zeros $\cos(\alpha_j h)$, and, thus, α_j for $j = 1, \dots, M$. Now we need to distinguish between the case $\beta_j = 0$ for all $j = 1, \dots, M$ and $\beta_j \neq 0$ for at least one $j \in \{1, \dots, M\}$.

1. If the parameters β_j do not appear in the signal f given in (3.24), we only need to solve the linear system

$$\frac{H(x_0)}{H(G^{-1}(G(x_0) + kh))} f(G^{-1}(G(x_0) + kh)) = \sum_{j=1}^M c_j H(x_0) \sin(\alpha_j (G(x_0) + kh))$$

for $k = 0, \dots, 2M-1$ in order to compute the coefficients c_j for $j = 1, \dots, M$.

2. If there exists a $j \in \{1, \dots, M\}$ such that $\beta_j \neq 0$, we first solve

$$\begin{aligned} & \frac{H(x_0)}{H(G^{-1}(G(x_0) + kh))} f(G^{-1}(G(x_0) + kh)) + \frac{H(x_0)}{H(G^{-1}(G(x_0) - kh))} f(G^{-1}(G(x_0) - kh)) \\ &= 2 \sum_{j=1}^M c_j \sin(\alpha_j G(x_0) + \beta_j) \cos(\alpha_j h k) \end{aligned}$$

for $k = 0, \dots, M-1$ in order to compute $d_j := c_j \sin(\alpha_j G(x_0) + \beta_j)$ for $j = 1, \dots, M$.

Moreover, we also solve

$$\begin{aligned} & \frac{H(x_0)}{H(G^{-1}(G(x_0) + kh))} f(G^{-1}(G(x_0) + kh)) - \frac{H(x_0)}{H(G^{-1}(G(x_0) - kh))} f(G^{-1}(G(x_0) - kh)) \\ &= 2 \sum_{j=1}^M c_j \cos(\alpha_j G(x_0) + \beta_j) \sin(\alpha_j h k) \end{aligned}$$

for $k = 0, \dots, M-1$ in order to compute $\tilde{d}_j := H(x_0)c_j \cos(\alpha_j G(x_0) + \beta_j)$ for $j = 1, \dots, M$. Finally, we can conclude for the parameters c_j and β_j for $j = 1, \dots, M$ by

$$c_j = \frac{\sqrt{d_j^2 + \tilde{d}_j^2}}{H(x_0)} \quad \text{and} \quad \beta_j = \arg(\tilde{d}_j + id_j) - \alpha_j G(x_0) \bmod 2\pi.$$

□

Remark 3.39 (i) Analogously, to Remark 3.10 we can also use Theorem 2.5 for the recovery of M -sparse trigonometric expansions by considering the Prony polynomial in monomial basis instead of Chebyshev basis.

(ii) Analogously, to Theorem 3.12 we also use generalized symmetric shift operators $S_{G,H,h,-h}$ as in Definition 3.34 to recover M -sparse expansions into generalized hyperbolic functions such as

$$f(x) = \sum_{j=1}^M c_j H(x) \cosh(G(x)\alpha_j) \quad \text{and} \quad g(x) = \sum_{j=1}^M c_j H(x) \sinh(G(x)\alpha_j)$$

with $c_j \in \mathbb{C} \setminus \{0\}$ and $\alpha_j \in \mathbb{R}$ for $j = 1, \dots, M$.

◇

We summarize our results for the reconstruction of sparse generalized cosine expansion in the following algorithm.

Algorithm 4 Prony's Method in Terms of Generalized Symmetric Shift Operators

Input: $M \in \mathbb{N}$, $h > 0$ and sampled values $f(G^{-1}(h\ell + G(x_0)))$, $\ell = -2M + 1, \dots, 2M - 1$, $H(x_0)$, $G(x_0)$

1: Form the Hankel-plus-Toeplitz matrix

$$\mathbf{H}_M^{\text{Toep}} := \left(\begin{aligned} &a_{k+\ell} f(G^{-1}(h(\ell+k) + G(x_0))) + a_{k-\ell} f(G^{-1}(h(k-\ell) + G(x_0))) \\ &+ a_{-(k+\ell)} f(G^{-1}(-h(\ell+k) + G(x_0))) + a_{-(k-\ell)} f(G^{-1}(-h(k-\ell) + G(x_0))) \end{aligned} \right)_{k,\ell=0}^{M-1}$$

as well as

$$\mathbf{f} := \left(\begin{aligned} &a_{M+\ell} f(G^{-1}(G(x_0) + h(M+\ell))) + a_{\ell-M} f(G^{-1}(G(x_0) + h(\ell-M))) \\ &+ a_{-(M+\ell)} f(G^{-1}(G(x_0) - h(M+\ell))) + a_{-(\ell-M)} f(G^{-1}(G(x_0) - h(\ell-M))) \end{aligned} \right)_{\ell=0}^{M-1}$$

with $a_\ell := \frac{H(x_0)}{H(G^{-1}(G(x_0)+h\ell))}$, $\ell = -2M + 1, \dots, 2M - 1$ and solve $\mathbf{H}_M \cdot \mathbf{p} = -\mathbf{f}$ for $\mathbf{p} := (p_\ell)_{\ell=0}^{M-1}$.

2: Define the Prony polynomial $P(z) = \sum_{k=0}^M p_k T_k(z)$ with $p_M = 2^{-(M-1)}$, find all roots $\lambda_j = \cos(\alpha_j/h)$ and determine the parameters $\alpha_j := \frac{1}{h} \arccos(\lambda_j)$ for $j = 1, \dots, M$.

3: Solve the Vandermonde-like system $2\mathbf{V}_{2M,M} \cdot \mathbf{d} = \mathbf{f}_{2M}$ with

$$\begin{aligned} \mathbf{V}_{2M,M} &:= (\cos(\alpha_j k))_{k=0, j=1}^{2M-1, M}, \\ \mathbf{d} &:= (c_j \cos(\alpha_j G(x_0) + \beta_j))_{j=1}^M \text{ and} \\ \mathbf{f}_{2M} &:= \left(a_k f(G^{-1}(G(x_0) + hk)) + a_{-k} f(G^{-1}(G(x_0) - hk)) \right)_{k=0}^{2M-1}. \end{aligned}$$

4: Solve the Vandermonde-like system $2\tilde{\mathbf{V}}_{2M,M} \cdot \tilde{\mathbf{d}} = \tilde{\mathbf{f}}_{2M}$ with

$$\begin{aligned} \tilde{\mathbf{V}}_{2M,M} &:= (\sin(\alpha_j k))_{k=0, j=1}^{2M-1, M}, \\ \tilde{\mathbf{d}} &:= (c_j \sin(\alpha_j G(x_0) + \beta_j))_{j=1}^M \text{ and} \\ \tilde{\mathbf{f}}_{2M} &:= \left(a_k f(G^{-1}(G(x_0) + hk)) - a_{-k} f(G^{-1}(G(x_0) - hk)) \right)_{k=0}^{2M-1}. \end{aligned}$$

Compute the coefficients $c_j = \frac{\sqrt{d_j^2 + i\tilde{d}_j^2}}{H(x_0)}$ and $\beta_j = \arg(d_j + i\tilde{d}_j) - \alpha_j G(x_0) \bmod 2\pi$ for $j = 1, \dots, M$.

Output: c_j , α_j and β_j for $j = 1, \dots, M$ as in (3.24).

Remark 3.40 This algorithm can be adapted for the reconstruction of M -sparse generalized sine expansion as in (3.25). Therefore, in the third and fourth step, the vectors \mathbf{d} and $\tilde{\mathbf{d}}$ have to be exchanged. \diamond

3.2.3 Reconstruction Using Linear Differential Operators

In the previous section we have seen that every generalized exponential sum can be reconstructed using a corresponding generalized shift operator. However, the classical exponential sum $e^{\alpha_j x}$ is also an eigenfunction of the ordinary linear differential operator $\frac{d}{dx}$, i.e.,

$$\frac{d}{dx} e^{\alpha x} = \alpha e^{\alpha x} \quad \alpha \in \mathbb{C}.$$

We want to use this as a starting point to derive suitable linear differential operators such that the generalized exponentials as in Definition 3.12 are, indeed, eigenfunctions of these operators, since it is often simpler to find linear differential operators than to the generalized shift operators. Theorem 3.24 already gives rise to a possible linear differential operator and motivates the following definition.

Definition 3.41 Let $D: C^1(\mathbb{R}) \rightarrow C(\mathbb{R})$ be the differential operator defined as $Df := \frac{d}{dx}f$. Furthermore, let $g, v: [a, b] \rightarrow \mathbb{R}$ be smooth functions and g be non-vanishing. Then we define the **linear differential operator** $A: C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ by

$$Af(x) := g(x)(Df)(x) + v(x)f(x) \tag{3.26}$$

for all $f \in C^\infty(\mathbb{R})$.

Theorem 3.42 Let $g, v: [a, b] \rightarrow \mathbb{R}$ be smooth functions and g be non-vanishing. Then the linear differential operator $A: C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ given by (3.26) possesses eigenfunctions of the form

$$e^{\alpha G(x)} H(x)$$

with the eigenvalues $\alpha \in \mathbb{C}$, where $G, H: [a, b] \rightarrow \mathbb{R}$ are smooth functions given by $G(x) := \int_a^x \frac{1}{g(t)} dt$ and $H(x) := \exp\left(-\int_a^x \frac{v(t)}{g(t)} dt\right)$, respectively.

Proof. Let $G, H: [a, b] \rightarrow \mathbb{R}$ be defined as in Theorem 3.42, i.e., $G(x) := \int_a^x \frac{1}{g(t)} dt$ and $H(x) := \exp\left(-\int_a^x \frac{v(t)}{g(t)} dt\right)$. Then their derivatives are given by

$$\begin{aligned} G'(x) &= \frac{1}{g(x)} \quad \text{and} \\ H'(x) &= -H(x) \frac{v(x)}{g(x)}. \end{aligned} \tag{3.27}$$

We employ the definition of the linear differential operator A and (3.27) and get

$$\begin{aligned} A\left(e^{G(x)\alpha}H(x)\right) &= g(x)\left(e^{G(x)\alpha}H(x)\right)' + v(x)e^{G(x)\alpha}H(x) \\ &= g(x)\left(e^{G(x)\alpha}H'(x) + e^{G(x)\alpha}H(x)G'(x)\alpha\right) + v(x)e^{G(x)\alpha}H(x) \\ &= -e^{G(x)\alpha}H(x)v(x) + e^{G(x)\alpha}\alpha H(x) + e^{G(x)\alpha}H(x)v(x) \\ &= \alpha e^{G(x)\alpha}H(x). \end{aligned}$$

□

The above theorem tells us that we have found a suitable linear differential operator such that the generalized exponential sums are expansions into eigenfunctions of said operator. Therefore, we can use the generalized Prony method to reconstruct these sums.

Theorem 3.43 *Let $G, H \in C^\infty([a, b])$, such that G' and H are non-vanishing on $[a, b] \subset \mathbb{R}$ and let $x_0 \in [a, b]$ be fixed. Then the generalized exponential sum as in (3.12) can be viewed as a sparse expansion into eigenfunctions of the differential operator A in Definition 3.41, and f as in (3.12) can be uniquely reconstructed from the derivative samples $f^{(k)}(x_0)$ for $k = 0, \dots, 2M - 1$.*

Proof. Let f be a generalized exponential sum as in (3.12). Then Theorem 3.42 implies that f can be viewed as a sparse expansion into eigenfunctions of the linear differential operator A . Therefore, we can apply Theorem 2.5 for a suitable linear functional $F: C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ satisfying $F\left(H(x)e^{\alpha G(x)}\right) \neq 0$ for all $\alpha \in \mathbb{C}$. We choose F to be the point evaluation functional at x_0 with $x_0 \in [a, b]$, i.e., $F(f) := f(x_0)$, and have to consider the samples $F(A^k f)$ for $k = 0, \dots, M$.

Thus, in order to reconstruct the parameters α_j for $j = 1, \dots, M$, we first have to compute the required values $F(A^k f) = (A^k f)(x_0)$ for $k = 0, \dots, 2M - 1$.

For this purpose, we need to determine the lower triangular matrix

$$\mathbf{L} = (\lambda_{m,k})_{m,k=0}^{2M-1} \in \mathbb{R}^{2M \times 2M}$$

such that

$$\left(F(A^k f)\right)_{k=0}^{2M-1} = \left(A^k f(x_0)\right)_{k=0}^{2M-1} = \mathbf{L} \cdot \left(f^{(k)}(x_0)\right)_{k=0}^{2M-1}.$$

By definition of A we get $\lambda_{0,0} := 1$, $\lambda_{1,0} := g(x_0)$ and $\lambda_{1,1} := h(x_0)$. Generally we have to consider the elements $\lambda_{m,k}$ as functions in x in order to obtain the entries of \mathbf{L} , starting with $\lambda_{0,0} \equiv 1$. We use induction to show

$$A^k f(x) = \sum_{m=0}^k \lambda_{k,m}(x) f^{(m)}(x).$$

The base step is a direct consequence of equation (3.26). The induction step is given by

$$\begin{aligned} A^{k+1}f(x) &= A \left(\sum_{m=0}^k \lambda_{k,m}(x) f^{(m)}(x) \right) \\ &= g(x) \sum_{m=0}^k \left(\lambda'_{k,m}(x) f^{(m)}(x) + \lambda_{k,m}(x) f^{(m+1)}(x) \right) + v(x) \sum_{m=0}^k \lambda_{k,m}(x) f^{(m)}(x) \\ &= \sum_{m=0}^k \left(g(x) \lambda'_{k,m}(x) + v(x) \lambda_{k,m}(x) \right) f^{(m)}(x) + g(x) \lambda_{k,m}(x) f^{(m+1)}(x). \end{aligned}$$

We can conclude the recursion

$$\lambda_{k+1,m}(x) = \begin{cases} g(x) \lambda'_{k,m}(x) + v(x) \lambda_{k,m}(x), & m = 0 \\ g(x) \left(\lambda'_{k,m}(x) + \lambda_{k,m-1}(x) \right) + v(x) \lambda_{k,m}(x) & m = 1, \dots, k \\ g(x) \lambda_{k,m}(x) & m = k + 1, \end{cases}$$

and the induction claim follows. The matrix entries $\lambda_{m,k} := \lambda_{m,k}(x_0)$ are well-defined by assumption on H and G . Hence, application of Theorem 2.5 yields that the signal f can be uniquely reconstructed using the values

$$F(A^k f) = A^k f(x_0) = \sum_{m=0}^k \lambda_{k,m}(x_0) f^{(m)}(x_0),$$

and, thus, the claim follows. \square

Remark 3.44 One major difference between the reconstruction using linear differential operators and the reconstruction using generalized shift operators is the set of assumptions on G and H . In Section 3.2.1 it was sufficient to assume that G and H are continuous. In Section 3.2.3 we need to assume that G and H are smooth. \diamond

The connection between the linear differential operator A as in Definition 3.41 and the generalized shift operator given in Theorem 3.24 is given by the theory of one-parameter semigroups and operator valued exponential functions, as elaborated in Section 3.2.1.

Furthermore, this theory also fits into the framework of the generalized operator based Prony method, see [Sta18, SP20] and can be interpreted as a changing of operators, see [SP20], Theorem 3.4.

Theorem 3.45 *Let $A: V \rightarrow V$ be a linear operator, and let $\sigma(A) \neq 0$ be a subset of the point spectrum $\sigma_P(A)$ with pairwise different eigenvalues and with corresponding eigenfunctions v_α such that the map $\alpha \mapsto v_\alpha$ is injective for $\alpha \in \sigma(A)$. Let $\varphi: \sigma(A) \rightarrow \mathbb{C}$ be an analytic,*

injective function. Then $\varphi(A)$ is a well-defined linear operator on

$$M(A) := \left\{ \sum_{j=1}^M c_{\alpha_j} v_{\alpha_j} : M < \infty, c_{\alpha_j} \in \mathbb{C} \setminus \{0\} \text{ and } \alpha_j \in \sigma(A) \text{ pairwise distinct} \right\}.$$

Furthermore, if v_α is an eigenfunction of A corresponding to the eigenvalue α , then v_α is also an eigenfunction of $\varphi(A)$ corresponding to the eigenvalue $\varphi(\alpha)$.

3.3 Applications to Special Expansions

In this section we want to give some examples of special expansions, which can be reconstructed using generalized shift operators such as sparse expansions into Gaussian chirps and other non-stationary signals.

Definition 3.46 (Gaussian Chirp) Let $\beta \in \mathbb{C} \setminus \{0\}$ and $M \in \mathbb{N}$. We call f an M -sparse expansion into **Gaussian chirps** if f is of the form

$$f(x) = \sum_{j=1}^M c_j e^{-\beta(x-\alpha_j)^2} \quad (3.28)$$

for $c_j \in \mathbb{C} \setminus \{0\}$ and $\alpha_j \in \mathbb{C}$.

In order to reconstruct sparse expansions of the above form, we need to find a suitable generalized shift operator or linear differential operator such that the Gaussian chirps are eigenfunctions of said operator. First we observe that

$$e^{-\beta(x-\alpha)^2} = e^{-\beta(x^2-2x\alpha+\alpha^2)} = \underbrace{e^{-\beta\alpha^2} e^{-\beta x^2}}_{=:H(x)} e^{2\beta x\alpha}$$

implies that a Gaussian chirp, indeed, is a generalized exponential function, i.e., of the form $H(x)e^{G(x)\alpha}$ with

$$H(x) := e^{-\beta\alpha^2-\beta x^2} \quad \text{and} \quad G(x) := 2\beta x. \quad (3.29)$$

This is in fact a special case. In general it is not possible to choose the functions H or G depending on the parameter α , since the parameter is unknown. But since in this case

$$\frac{H(x)}{H(G^{-1}(G(x)+h))} = \frac{e^{-\beta(\alpha^2-x^2)}}{e^{-\beta(\alpha^2-((2\beta x+h)/(2\beta))^2)}} = e^{-h(x+h)/(4\beta)}$$

holds, the factor $e^{-\beta\alpha^2}$ cancels out and the operator $S_{G,H,h}$ is defined independently of α .

Therefore, the Gaussian chirps fit in the theoretical framework we have derived in

this chapter.

Theorem 3.47 *Let f be an M -sparse expansions into Gaussian chirps as in Definition 3.46 and $x_0 \in \mathbb{R}$. Then f can be uniquely reconstructed using the sample values $f^{(k)}(x_0)$ for $k = 0, \dots, 2M - 1$.*

Proof. Since the signal f is a generalized exponential sum with G and H given in (3.29), Theorem 3.43 implies that f is a sparse expansion into eigenfunctions of the differential operator $A := \frac{1}{G'(x)}D - \frac{1}{G'(x)}\frac{H'(x)}{H(x)}\text{Id} = \frac{1}{2\beta}D + x\text{Id}$, where D denotes the differential operator with $Df(x) := f'(x)$. Here, it is also important that the terms containing the parameter α cancel out due to

$$\frac{H'(x)}{H(x)} = \frac{e^{-\beta(a^2-x^2)}2\beta x}{e^{-\beta(a^2-x^2)}} = 2\beta x.$$

Thus, the operator A itself is independent of the parameter α .

Furthermore, Theorem 3.43 implies that f can be uniquely reconstructed using the values $f^{(k)}(x_0)$ for $k = 0, \dots, 2M - 1$. Moreover, we can choose any $x_0 \in \mathbb{R}$ since $G'(x_0) = 2\beta \neq 0$ and $H(x_0) \neq 0$ for all $x_0 \in \mathbb{R}$, which means that the interval $[a, b]$ in Theorem 3.43 can be chosen arbitrarily. \square

Theorem 3.48 *Let f be an M -sparse expansion into Gaussian chirps as in Definition 3.46 with known scale parameter $\beta \in \mathbb{R} \setminus \{0\}$ and $x_0 \in \mathbb{R}$. Furthermore, let the frequencies α_j of f be such that $|\text{Im}(\alpha_j)| < K$ for some $K \in \mathbb{R}$ and $h \in \mathbb{R}$ satisfying $0 < |h| < \frac{\pi}{K}$. Then f can be uniquely reconstructed using the $2M$ equidistant sample values $f(x_0 + \frac{hk}{2\beta})$ for $k = 0, \dots, 2M - 1$.*

Proof. As we have established above the signal f is a generalized exponential sum with G and H given in (3.29). Then we can apply Theorem 3.31 and can reconstruct f using the equidistant values $f(G^{-1}(G(x_0) + hk)) = f\left(\frac{hk+2\beta x_0}{2\beta}\right) = f\left(x_0 + \frac{hk}{2\beta}\right)$ for $k = 0, \dots, 2M - 1$. We note that we can choose any arbitrary interval $[a, b] \subset \mathbb{R}$. This implies that we can choose it such that

$$\frac{|G(b) - G(a)|}{2M} = \frac{2|\beta|(b-a)}{2M} > \frac{\pi}{K}$$

holds, and, therefore, no further condition on the choice of h is needed. \square

Remark 3.49 1. We note that the reconstruction of sparse expansions into Gaussian chirps in Theorem 3.48 can also be extended to $\beta \in \mathbb{C} \setminus \{0\}$. This is done by using the substitution $\tilde{\alpha}_j = \alpha_j 2\beta$ such that f in (3.28) can be rewritten as

$$f(x) = \sum_{j=1}^M \underbrace{c_j e^{-\beta \alpha_j^2}}_{=: \tilde{c}_j} e^{-\beta x^2} e^{\tilde{\alpha}_j x}.$$

Therefore, we can choose $G(x) := x$ and $H(x) := e^{-\beta x^2}$ and get the corresponding generalized shift operator $S_{G,H,h}$ with $S_{G,H,h}f(x) = e^{\beta h(2x+h)}f(x+h)$, see also [PSK19], Section 4.1.

2. In particular, the model in (3.28) also includes expansions into modulated shifted Gaussians

$$f(x) = \sum_{j=1}^M c_j e^{2\pi i s_j x} e^{-\beta(x-\alpha_j)^2}$$

with $s_j \in (0,1)$, $c_j \in \mathbb{R} \setminus \{0\}$ and $\alpha_j \in \mathbb{R}$ for $j = 1, \dots, M$, which we have considered in [PSK19]. Since

$$e^{2\pi i s_j x} e^{-\beta(x-\alpha_j)^2} = e^{-\beta\alpha_j^2} e^{-\beta x^2} e^{-x(2\beta\alpha_j + 2\pi i s_j)}$$

holds, we can choose $\tilde{\alpha}_j := 2\beta\alpha_j + 2\pi i s_j$ for $j = 1, \dots, M$ such that f can be rewritten as

$$f(x) = \sum_{j=1}^M \underbrace{c_j e^{-\beta\alpha_j^2}}_{=: \tilde{c}_j} e^{-\beta x^2} e^{-x\tilde{\alpha}_j}.$$

Then the reconstruction of the parameters $\tilde{\alpha}_j$ for $j = 1, \dots, M$ is sufficient to find the parameters α_j and s_j from the real and the imaginary part of $\tilde{\alpha}_j$ for $j = 1, \dots, M$, respectively.

◇

Example 3.50 We illustrate the recovery of expansions into Gaussian chirps and consider f of the form 3.28 with $M = 10$ and $\beta = i$. The original parameters in Table 3.2 have been drawn uniformly from the intervals $(-3,3) + i(-2,2)$ for c_j and from $(-2,2)$ for α_j .

param.	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$	$j = 9$	$j = 10$
Re c_j	2.357	1.212	0.334	-1.893	-1.728	-2.536	2.483	1.240	0.347	-1.119
Im c_j	-1.335	0.490	1.952	-1.318	-0.969	-0.413	-1.704	0.736	-0.390	1.931
α_j	-0.391	0.483	-1.356	-0.475	-1.355	1.032	1.484	-0.597	0.742	-0.823

Table 3.2: Parameters c_j and α_j for an expansion into Gaussian chirps as in (3.28) with $M = 10$, see Figure 3.1.

Since β is complex, we apply the substitution as in Remark 3.49. Moreover, the starting point is given by $x_0 = -1$ and the shift parameter by $h = 1$.

The reconstruction is done using Algorithm 3 and used 20 equidistant samples $f(k)$, for $k = -1, \dots, 18$, which are represented as black dots in Figure 3.1. The obtained

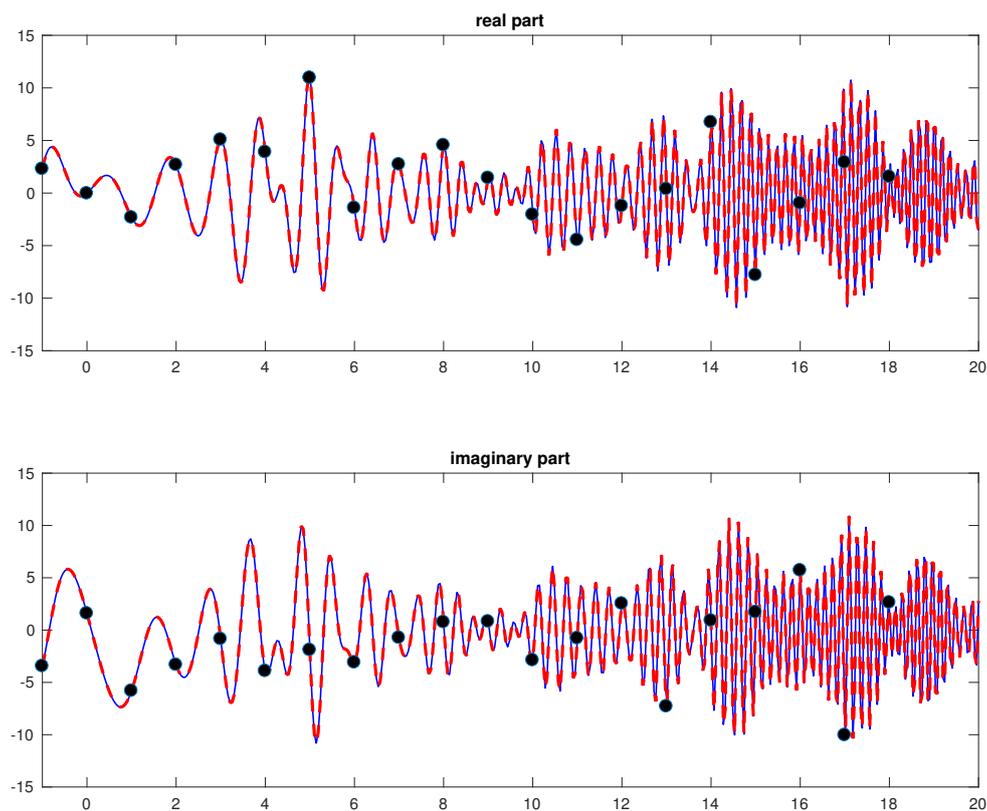


Figure 3.1: Real and imaginary part of the sparse expansion into Gaussian chirps f as given in Example 3.50 in blue. The black dots indicate the used signal values. The reconstructed signal was obtained by Algorithm 3 and is shown in red. It cannot be distinguished visually from the original signal f .

maximal reconstruction error for the parameters α_j and the parameters c_j are

$$\max_{j=1,\dots,10} |\alpha_j - \tilde{\alpha}_j| = 5.36 \cdot 10^{-12}, \quad \max_{j=1,\dots,10} |c_j - \tilde{c}_j| = 7.99 \cdot 10^{-10}$$

where \tilde{c}_j and $\tilde{\alpha}_j$, $j = 1, \dots, 10$, denote the parameters computed by Algorithm 3. ◇

Corollary 3.51 *Let f be a non-stationary signal with quadratic phase given by*

$$f(x) = \sum_{j=1}^M c_j \cos(x^2 + \alpha_j x + s_j) \quad (3.30)$$

with $M \in \mathbb{N}$, $c_j \in \mathbb{R} \setminus \{0\}$, $\alpha_j \in (-K, K)$ for some $K \in \mathbb{R}$ and $s_j \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ for $j = 1, \dots, M$. Furthermore, let $h \in \mathbb{R} \setminus \{0\}$ and $x_0 \in \mathbb{R}$. Then f in (3.30) can be reconstructed using the $2M$ sample values $f(x_0 + hk)$, $k = 0, \dots, 2M - 1$ and the $2M$ sample values $\tilde{f}(x_0 + hk)$,

where \tilde{f} is given by

$$\tilde{f}(x) = \sum_{j=1}^M c_j \sin(x^2 + \alpha_j x + s_j). \quad (3.31)$$

Proof. Using the substitutions $c_{j,1} := c_j \cos(s_j)$, $c_{j,2} := c_j \sin(s_j)$, $b_j := \left(\frac{c_{j,1} + ic_{j,2}}{2}\right)$, and $d_j := b_j e^{-ia_j^2/4}$, the signal given in (3.30) can be rewritten as

$$\begin{aligned} f(x) &= \sum_{j=1}^M (c_j \cos(s_j)) \cos(x^2 + \alpha_j x) + (c_j \sin(s_j)) \sin(x^2 + \alpha_j x) \\ &= \sum_{j=1}^M c_{j,1} \cos(x^2 + \alpha_j x) - c_{j,2} \sin(x^2 + \alpha_j x) \\ &= \sum_{j=1}^M \left(\frac{c_{j,1} + ic_{j,2}}{2} \right) e^{i(x^2 + \alpha_j x)} + \left(\frac{c_{j,1} - ic_{j,2}}{2} \right) e^{-i(x^2 + \alpha_j x)} \\ &= \sum_{j=1}^M b_j e^{i((x + \alpha_j/2)^2 - \alpha_j^2/4)} + \bar{b}_j e^{-i((x + \alpha_j/2)^2 - \alpha_j^2/4)} \\ &= 2 \operatorname{Re} \left(\sum_{j=1}^M b_j e^{-ia_j^2/4} e^{i(x + \alpha_j/2)^2} \right) \\ &= 2 \operatorname{Re} \left(\sum_{j=1}^M d_j e^{i(x + \alpha_j/2)^2} \right). \end{aligned}$$

Analogously, we observe that

$$\begin{aligned} \tilde{f}(x) &= \sum_{j=1}^M c_j \sin(x^2 + \alpha_j x + s_j) = \sum_{j=1}^M c_j \cos \left(x^2 + \alpha_j x + s_j - \frac{\pi}{2} \right) \\ &= 2 \operatorname{Im} \left(\sum_{j=1}^M d_j e^{i(x + \alpha_j/2)^2} \right) \end{aligned}$$

with $d_j = b_j e^{-ia_j^2/4}$. The model is, therefore, closely related to the model in (3.28). Hence, we can define the function $g(x) := f(x) + i\tilde{f}(x)$, which is an expansion into Gaussian chirps. Consequently, we can apply Theorem 3.48 in order to recover the parameters d_j and α_j for $j = 1, \dots, 2M$ using the $2M$ samples $g(x_0 + hk)$ for $k = 0, \dots, 2M - 1$. As a last step we compute the parameters c_j and s_j for $j = 1, \dots, M$ using

$$\begin{aligned} b_j &= d_j e^{ia_j^2/4}, & c_{j,1} &= 2 \operatorname{Re}(b_j), & c_{j,2} &= 2 \operatorname{Im}(b_j) \\ |c_j| &= 2|b_j|, & s_j &= \arg(b_j) \text{ and } \operatorname{sgn}(c_j) = \operatorname{sgn}(c_{j,1}). \end{aligned}$$

□

Example 3.52 We will illustrate the recovery of expansion f of the form (3.30) with $M = 8$. The original parameters are given in Table 3.3. The parameters have been drawn uniformly from the intervals $(-1, 5)$ for the coefficients c_j , from $(-\pi, \pi)$ for α_j and from $(-\frac{\pi}{2}, \frac{\pi}{2})$ for s_j for $j = 1, \dots, M$. We apply Corollary 3.51 with starting point $x_0 = 0$ and shift parameter $h = \frac{9}{20}$. Hence, we have used the signal values $f(\frac{9}{20}k)$ and $\tilde{f}(\frac{9}{20}k)$ for $k = 0, \dots, 15$ with \tilde{f} as in (3.31). The signal and its reconstruction are shown in Figure 3.2.

parameter	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
c_j	-0.2998	-0.285	4.639	2.873	1.877	2.836	2.268	2.883
α_j	0.276	1.389	0.141	3.102	-1.768	-2.477	-2.452	-2.742
s_j	-0.2997	-0.162	-0.422	0.828	0.402	0.854	1.360	1.485

Table 3.3: Parameters c_j , s_j and α_j for $j = 1, \dots, M$ for an expansion f as in (3.30) with $M = 8$.

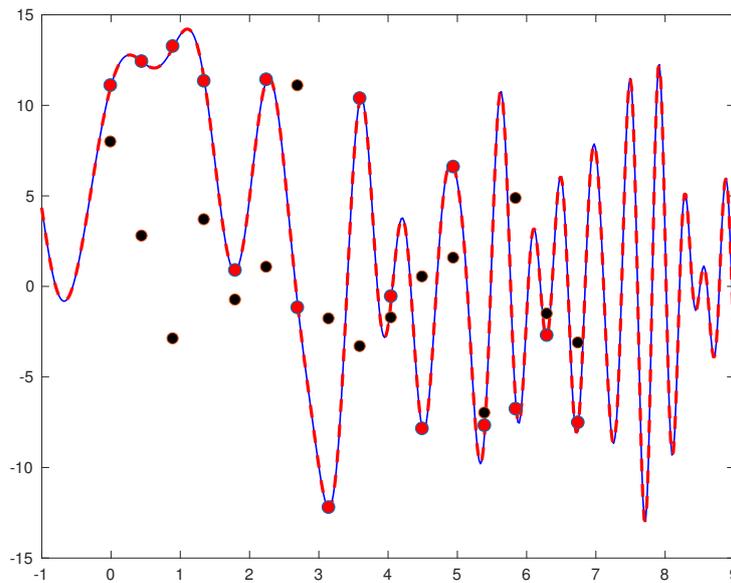


Figure 3.2: Original signal f of the form (3.30) with the parameters given in Table 3.3 in blue. The red dots indicate the used signal values of f and the black dots indicate the used signal values of \tilde{f} . The reconstructed signal was obtained by application of Corollary 3.51 and Algorithm 3 and is shown in red. It cannot be distinguished visually from the original signal f .

The obtained maximal reconstruction errors for the parameters α_j , c_j and s_j for $j = 1, \dots, 8$ are

$$\begin{aligned}\max_{j=1,\dots,8} |\alpha_j - \tilde{\alpha}_j| &= 1.7 \cdot 10^{-6}, \\ \max_{j=1,\dots,8} |c_j - \tilde{c}_j| &= 9.8 \cdot 10^{-5} \text{ and} \\ \max_{j=1,\dots,8} |s_j - \tilde{s}_j| &= 1.4 \cdot 10^{-4},\end{aligned}$$

where \tilde{c}_j , $\tilde{\alpha}_j$ and \tilde{s}_j for $j = 1, \dots, 8$ denote the recovered parameters. \diamond

Corollary 3.53 *Let $p \in \mathbb{N}$ odd and $M \in \mathbb{N}$. Let the signal f be given by*

$$f(x) = \sum_{j=1}^M c_j \cos(\alpha_j x^p + \beta_j) \quad (3.32)$$

with $c_j \in \mathbb{R} \setminus \{0\}$, $\alpha_j \in \mathbb{R}$, $\beta_j \in (-\pi, \pi) \setminus \{\pm \frac{\pi}{2}\}$ for $j = 1, \dots, M$. Moreover, let $\alpha_j \in [0, K)$ for some $K \in \mathbb{R}_+$ for all $j = 1, \dots, M$. Then f can be uniquely reconstructed using the samples $f\left(\text{sgn}(x_0 + hk) \sqrt[p]{|x_0 + kh|}\right)$ for $k = -2M + 1, \dots, 2M + 1$ with $x_0 \geq 0$ satisfying $\cos(\alpha_j x_0) \neq 0$ for $j = 1, \dots, M$ and $0 < h < \frac{\pi}{K}$. If we choose $x_0 = 0$, the functional values $f\left(\sqrt[p]{hk}\right)$ for $k = 0, \dots, 2M - 1$ are sufficient for the reconstruction.

Proof. We consider the intervall $[0, \infty)$ and set $G(x) := x^p$. Choosing the starting point as $\sqrt[p]{x_0}$, applying Theorem 3.37 and using the fact that cosine is an even function proves the claim. \square

Example 3.54 We illustrate the recovery method of a signal f of the form (3.32) with a small numerical example. Our signal f is of the form (3.32) with $M = 2$, $p = 3$ and the parameters T_j , c_j and α_j given in Table 3.4. The parameters have been drawn uniformly from the intervals $(0, 5)$ for the coefficients c_j , from $(0, \pi)$ for α_j and from $(-\pi, \pi)$ for β_j for $j = 1, 2$. We apply Corollary 3.53 with starting point $x_0 = 0$ and shift parameter $h = 1$. Hence, we have used the signal values $f(\pm \sqrt[3]{k})$ for $k = 0, \dots, 3$. The signal and its reconstruction are shown in Figure 3.3.

parameter	$j = 1$	$j = 2$
c_j	4.64157	2.90045
α_j	2.41598	1.82667
β_j	-3.03488	-2.38221

Table 3.4: Parameters c_j , α_j and β_j for $j = 1, 2$ for a expansion f as in (3.32) with $M = 2$.

The obtained maximal reconstruction error for the parameters α_j , c_j and β_j for $j =$

1,2 are

$$\begin{aligned} \max_{j=1,2} |\alpha_j - \tilde{\alpha}_j| &= 2.7 \cdot 10^{-15}, \\ \max_{j=1,2} |c_j - \tilde{c}_j| &= 1.9 \cdot 10^{-14} \text{ and} \\ \max_{j=1,2} |\beta_j - \tilde{\beta}_j| &= 1.8 \cdot 10^{-15}, \end{aligned}$$

where \tilde{c}_j , $\tilde{\beta}_j$ and $\tilde{\alpha}_j$ for $j = 1, 2$ denote the recovered parameters.

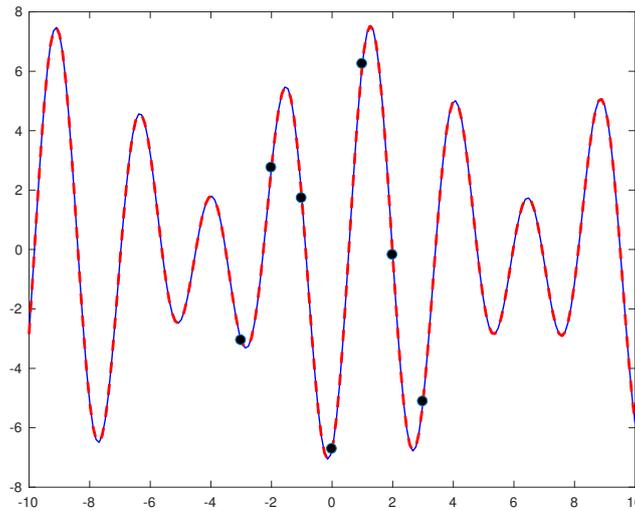


Figure 3.3: Original signal f of the form (3.32) with the parameters given in Table 3.4 in blue. The black dots indicate the used signal values of f . The reconstructed signal was obtained by application of Corollary 3.53 and Algorithm 4 and is shown in red. It cannot be distinguished visually from the original signal f .

◇

At last we want to consider an expansion into special orthogonal polynomials, namely Chebyshev polynomials. These expansions have also been studied in [PT14b, PP13] and in our paper [PSK19] and we will also study these expansion in Chapter 4.

The approach can also be transferred to Chebyshev polynomials of second, third and fourth kind, see [PT14b]. However, in [PT14b] the connection to shift operators with Chebyshev polynomials as eigenfunctions has not been explicitly used.

Definition 3.55 The following expansion

$$f(x) = \sum_{j=1}^M c_{n_j} T_{n_j}(x) \quad (3.33)$$

with $c_{n_j} \in \mathbb{C} \setminus \{0\}$, $n_j \in \mathbb{N}_0$, for $j = 1, \dots, M$ and T_{n_j} being the n_j -th Chebyshev polynomials of first kind as in Definition 3.3 is called an **M -sparse expansion into Chebyshev polynomials of first kind**. We assume the n_j to be ordered, i.e., $0 \leq n_1 < n_2 < \dots < n_M$.

We consider the interval $[-1, 1]$. Lemma 3.4 (iv) yields that for $x \in [-1, 1]$, a sparse expansion into Chebyshev polynomials f can be written as

$$f(x) = \sum_{j=1}^M c_{n_j} \cos(n_j \arccos(x)).$$

Thus, f is of the form (3.24) with $H \equiv 1$ and $G(x) := \arccos(x)$. Consequently, we can conclude the following corollary for the reconstruction of f .

Corollary 3.56 *Let $M \in \mathbb{N}$ and f be an M -sparse expansion into Chebyshev polynomials of first kind. Let K be an upper bound of the degree of f . Furthermore, let $0 < h \leq \frac{\pi}{K}$. Then f can be uniquely reconstructed using the $2M$ samples $f(\cos(hk))$ for $k = 0, \dots, 2M - 1$.*

Proof. This is a direct implication of Theorem 3.37 using the starting point $x_0 = 1$. Therefore, the required samples simplify to $f(\cos(\arccos(1) + kh)) = f(\cos(kh))$ for $k = -2M + 1, \dots, 2M + 1$. Using the fact that cosine is an even function the claim follows. □

Definition 3.57 (Chebyshev Polynomials of Second Kind) Let $n \in \mathbb{N}_0$ and $x \in \mathbb{R}$. Then the **Chebyshev polynomial of second kind** of degree $n + 1$ is defined by

$$U_{n+1} := 2^n \prod_{k=0}^n \left(x - \cos\left(\frac{(k+1)\pi}{n+2}\right) \right) =: \sum_{\ell=0}^{n+1} a_{n+1,\ell} x^\ell$$

with $U_0(x) = 1$.

We summarize some of the most common properties of the Chebyshev polynomials of second kind in the following lemma.

Lemma 3.58 Let $n \in \mathbb{N}_0$ and $x \in \mathbb{R}$.

(i) U_n is a polynomial of degree n .

(ii) The leading coefficient of the Chebyshev polynomial U_n is given by

$$a_{n,n} = \begin{cases} 2^{n-1}, & \text{if } n \geq 1, \\ 1, & \text{if } n = 0. \end{cases}$$

(iii) The n zeros of U_n are given by

$$t_{n,\ell} = \cos\left(\frac{\ell\pi}{n+1}\right) \text{ for } \ell = 1, \dots, n.$$

(iv) For $x \in [-1, 1]$ the Chebyshev polynomial U_n can be written as

$$U_n(x) = \frac{\sin((n+1)\arccos(x))}{\sin(\arccos(x))}.$$

(v) The Chebyshev polynomials satisfy the recurrence relation

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x)$$

with $U_0(x) = 1$ and $U_1(x) = x$.

(vi) The Chebyshev polynomials of second kind U_n can be expressed in terms of the derivative of the Chebyshev polynomials of first kind T_{n+1} , i.e.,

$$U_n(x) = \frac{1}{n+1} T'_{n+1}(x).$$

Proof. The claims (i), (ii) and (iii) follow directly from Definition 3.57. Proofs of (iv),(v) and (vi) can be found in [PPST19], Section 6.1. \square

Corollary 3.59 Let $M \in \mathbb{N}$ and f be an M -sparse expansion into Chebyshev polynomials of second kind, i.e.,

$$f(x) = \sum_{j=1}^M c_{n_j} U_{n_j}(x) \tag{3.34}$$

with $c_{n_j} \in \mathbb{C} \setminus \{0\}$ and $n_j \in \mathbb{N}_0$ for $j = 1, \dots, M$. We assume the n_j to be ordered, i.e., $0 \leq n_1 < n_2 < \dots < n_M$. Moreover, let K be an upper bound of the degree of f , i.e., $n_M < K$. Then f can be uniquely reconstructed using the samples $f(\cos(\arccos(x_0) + kh))$ for $k = -2M+1, \dots, 2M+1$ with $x_0 \in (-1, 1)$ and $0 < h < \frac{\pi}{K}$ satisfying $\arccos(x_0) + kh \in (0, \pi)$ for $k = -2M+1, \dots, 2M+1$.

Proof. Lemma 3.58 (iv) implies that f can be written as

$$f(x) = \sum_{j=1}^M c_{n_j} H(x) \sin((n_j + 1)G(x))$$

with $H(x) := \frac{1}{\sin(\arccos(x))}$ and $G(x) := \arccos(x)$. Both functions are non-vanishing on $(-1, 1)$ and G is monotone on $(-1, 1)$. Therefore, we can apply Theorem 3.38 in order to recover f . \square

Remark 3.60 The connection between Chebyshev polynomials of first kind T_n and Chebyshev polynomials of second kind given in Lemma 3.58 (vi) can also be seen in the operators used for the reconstruction. Let $G(x) := \arccos(x)$, $H \equiv 1$ and $S_{G,H,h,-h}$ denote the generalized symmetric shift operator as in Definition 3.34. We consider the action of this operator on a function f and its derivative

$$\begin{aligned} \frac{d}{dx} S_{G,H,h,-h} f(x) &= \frac{1}{2} \left(\frac{d}{dx} f(\cos(\arccos(x) + h)) + \frac{d}{dx} f(\cos(\arccos(x) - h)) \right) \\ &= \frac{1}{2} \left(f'(\cos(\arccos(x) + h)) \frac{\sin(\arccos(x) + h)}{\sin(\arccos(x))} \right. \\ &\quad \left. + f'(\cos(\arccos(x) - h)) \frac{\sin(\arccos(x) - h)}{\sin(\arccos(x))} \right). \end{aligned}$$

Viewing this as the action of an operator on f' , we see that this yields, indeed, the same generalized symmetric shift operator $S_{G,H,h,-h}$ as in the Corollary above. \diamond

4 Reconstruction of Expansions into Orthogonal Polynomials

In Section 3.3 we introduced operators for the reconstructions of M -sparse expansions into Chebyshev polynomials of first or second kind using only $2M$ functional values.

In recent years, there has been a number of research papers dedicated to the recovery of sparse expansions into special orthogonal polynomials such as generalized Chebyshev polynomials. In [IKY18], a substitution was used in order to reduce the sparse expansion into orthogonal polynomials into a sparse Laurent polynomial in power (standard) basis. Sparse expansions into Legendre or Chebyshev polynomials are analysed in [PT14b,PT16]. Furthermore, reconstruction methods for sparse expansions into non-standard basis such as the Pochhammer basis were given in [LS95].

This chapter focuses on the recovery of expansions into orthogonal polynomials and the general structure behind it. A Prony based reconstruction can be applied by using differential operators such as the Sturm-Liouville Operator and the generalized Prony method, see [PP13]. One major drawback is that the required samples are derivative values, which are not easily accessible. Only in some cases we can use shift operators, functional values and the generalized Prony method to recover expansions into orthogonal polynomials such as Chebyshev polynomials of first and second kind, see Chapter 3.

A further method based on inner products and dual operators as well as the generalized Prony method was introduced in [Sta18].

In this chapter we will introduce a new reconstruction method which is based on generating functions of the orthogonal polynomials. This new algorithm is presented in Section 4.1.

In Section 4.2 we expand the reconstruction approach based on linear differential operators and compare both methods.

Moreover, we will introduce a new application of the generalized Prony method for q -hypergeometric polynomials in Section 4.3.

In the last section of this chapter, Section 4.4, we will illustrate our newly developed algorithm using numerical examples.

4.1 Reconstruction Method Using Generating Functions

This section is devoted to the development of a new reconstruction method for sparse expansions into orthogonal polynomials based on their generating function. Thus, we will begin by introducing the definitions needed.

Definition 4.1 (Orthogonal Polynomials) Let $L^2([a, b], \omega)$ be the space of square integrable functions on the interval $[a, b] \subset \mathbb{R}$ with a weighted inner product

$$\langle f, g \rangle = \int_a^b f(x)g(x)\omega(x)dx \quad \text{for all } f, g \in L^2([a, b]) \quad (4.1)$$

with an integrable weight function $\omega : [a, b] \rightarrow \mathbb{R}$. Then the sequence $\{Q_k \mid k \in \mathbb{N}_0\}$ of polynomials is called **orthogonal** if

$$\deg(Q_k) = k \quad \text{and} \quad \langle Q_j, Q_k \rangle = 0 \quad \text{for } j \neq k.$$

In the following we denote the inner product corresponding to the set $\{Q_k \mid k \in \mathbb{N}_0\}$ of orthogonal polynomials with $\langle \cdot, \cdot \rangle_Q$ and

$$\langle Q_i, Q_j \rangle_Q = d_j \delta_{i,j} \quad (4.2)$$

with constants $d_j \in \mathbb{R}$ for $i, j \in \mathbb{N}_0$.

We will introduce a class of orthogonal polynomials, namely the Jacobi polynomials, and summarize some of their common properties. Many well studied and important orthogonal polynomials are special cases of Jacobi polynomials, such as Chebyshev polynomials, see Definition 3.3 and Definition 3.57, and Legendre polynomials.

Definition 4.2 (Jacobi Polynomials) Let $n \in \mathbb{N}_0$ and $\alpha, \beta > -1$, $x \in [-1, 1]$. Then the **Jacobi polynomial** of degree n is defined as

$$P_n^{(\alpha, \beta)}(x) := \frac{\Gamma(\alpha + n + 1)}{n! \Gamma(\alpha + \beta + n + 1)} \sum_{m=0}^n \binom{n}{m} \frac{\Gamma(\alpha + \beta + n + m + 1)}{\Gamma(\alpha + m + 1)} \left(\frac{x-1}{2}\right)^m$$

where $\Gamma(x) := \int_0^\infty t^{x-1} e^{-t} dt$ for $\text{Re}(x) > 0$ denotes the Gamma function.

Lemma 4.3 Let $n \in \mathbb{N}_0$ and $x \in [-1, 1]$. Then the following holds.

(i) $P_n^{(\alpha, \beta)}$ is a polynomial of degree n .

(ii) The Jacobi polynomials satisfy the Rodrigues' formula

$$P_n^{(\alpha, \beta)}(x) := \frac{(-1)^n}{2^n n!} (1-x)^{-\alpha} (1+x)^{-\beta} \frac{d^n}{dx^n} \left((1-x)^{\alpha+n} (1+x)^{\beta+n} \right).$$

(iii) The Jacobi polynomials are orthogonal on the interval $[-1, 1]$ with the weight function $\omega(x) := (1-x)^\alpha(1+x)^\beta$, i.e.,

$$\int_{-1}^1 P_j^{(\alpha, \beta)}(x) P_k^{(\alpha, \beta)}(x) (1-x)^\alpha (1+x)^\beta dx = \frac{2^{\alpha+\beta+1}}{2j+\alpha+\beta+1} \frac{\Gamma(j+\alpha+1)\Gamma(j+\beta+1)}{\Gamma(j+\alpha+\beta+1)j!} \delta_{jk} \quad (4.3)$$

for $j \neq 0$ and

$$\int_{-1}^1 P_j^{(\alpha, \beta)}(x) P_k^{(\alpha, \beta)}(x) (1-x)^\alpha (1+x)^\beta dx = \frac{2^{\alpha+\beta+1}}{\Gamma(\alpha+\beta+2)} \frac{\Gamma(j+\alpha+1)\Gamma(j+\beta+1)}{j!} \delta_{jk}$$

for $j = 0$.

(iv) The Jacobi polynomials $P_n(\alpha, \beta)$ satisfy the following linear homogeneous differential equation of second order

$$(1-x^2)y'' + [\beta - \alpha - (\alpha + \beta + 2)x]y' + n(n + \alpha + \beta + 1)y = 0.$$

Proof. The proof of (ii) as well as (iii) can be found in [Sze75], Section 4.3. The claim (i) is a consequence of (ii). The proof of (iv) can be found in [Sze75], Section 4.2. as the proof of Theorem 4.2.1. \square

The Chebyshev polynomials of first kind are a special case of the Jacobi polynomials for $\alpha = \beta = -\frac{1}{2}$, i.e., $P_n^{(-0.5, -0.5)}(x) = T_n(x)$. Moreover we have $U_n(x) = P_n^{(0.5, 0.5)}(x)$ and $P_n(x) = P_n^{(0, 0)}(x)$, where U_n denote the Chebyshev polynomials of second kind and P_n the Legendre polynomials, see [Sze75], Section 4.1.

Definition 4.4 (Generating Function) A function $g : I_1 \times I_2 \rightarrow \mathbb{R}$ with $I_1 \subset \mathbb{R}$ and $I_2 \subset \mathbb{C}$ is called a **generating function** for the orthogonal polynomials Q_k , $k = 0, 1, 2, \dots$, if it is of the following form

$$g(x, t) = \sum_{k=0}^{\infty} w_k Q_k(x) t^k \quad \text{with } x \in I_1, t \in I_2, \quad (4.4)$$

where $w_k \in \mathbb{R}$ and $w_k > 0$. Here, we use the convention $0^0 = 1$. The generating function is called **ordinary** if $w_k = 1$ for $k \in \mathbb{N}_0$ and **exponential** if $w_k = \frac{1}{k!}$ for $k \in \mathbb{N}_0$.

Remark 4.5 The above definition is a formal one. In order to see if the above series in (4.4) is, indeed, well-defined, we need to check its convergence. This has to be done with respect to the particular set $\{Q_k \mid k \in \mathbb{N}_0\}$ of orthogonal polynomials and the weight coefficients w_k , $k \in \mathbb{N}_0$ which are used. \diamond

Theorem 4.6 Let $P_n^{(\alpha,\beta)}$ be the Jacobi polynomials. Then the ordinary generating function of the Jacobi polynomials is given by

$$\sum_{n=0}^{\infty} P_n^{(\alpha,\beta)}(x)t^n = 2^{\alpha+\beta}R^{-1}(1-t+R)^{-\alpha}(1+t+R)^{-\beta},$$

with $R = R(x,t) = \sqrt{1-2xt+t^2}$.

Proof. The proof can be found in [Sze75], Section 4.4. □

Since the Chebyshev polynomials are special cases of the Jacobi polynomials, the above theorem yields the ordinary generating function for the Chebyshev polynomials. Indeed, for this special case a different and simple proof is possible.

Theorem 4.7 Let $\{T_n \mid n \in \mathbb{N}_0\}$ denote the set of Chebyshev polynomials of first kind and let $\{U_n \mid n \in \mathbb{N}_0\}$ be the set of Chebyshev polynomials of second kind. Then the ordinary generating functions are given by

$$\sum_{n=0}^{\infty} T_n(x)t^n = \frac{1-xt}{1-2xt+t^2}$$

for $x \in [-1,1]$ and $|t| < 1$ and

$$\sum_{n=0}^{\infty} U_n(x)t^n = \frac{1}{1-2xt+t^2}$$

for $x \in [-1,1]$ and $|t| < 1$, respectively.

Proof. We use the recurrence relation for the Chebyshev polynomials given in Lemma 3.4 (v) and consider

$$\begin{aligned} \sum_{n=0}^{\infty} T_{n+2}(x)t^{n+2} &= t2x \sum_{n=0}^{\infty} T_{n+1}(x)t^{n+1} - t^2 \sum_{n=0}^{\infty} T_n(x)t^n \\ \iff \sum_{n=0}^{\infty} T_n(x)t^n - 1 - xt &= 2xt \sum_{n=0}^{\infty} T_n(x)t^n - 2xt - t^2 \sum_{n=0}^{\infty} T_n(x)t^n \\ \iff \sum_{n=0}^{\infty} t^n T_n(x)(1-2xt+t^2) &= 1-xt \\ \implies \sum_{n=0}^{\infty} T_n(x)t^n &= \frac{1-xt}{1-2xt+t^2}. \end{aligned}$$

Lemma 3.4 (iv) yields, that $|T_n(x)| \leq 1$ for $x \in [-1,1]$ as well as $|T_n t^n| \leq |t|^n$. For $|t| < 1$, the series

$$\sum_{n=0}^{\infty} |T_n(x)t^n| \leq \sum_{n=0}^{\infty} |t|^n = \frac{1}{1-|t|}$$

converges and consequently, the Weierstrass M-test yields absolute and uniform convergence for all $x \in [-1, 1]$ and $|t| < 1$. Using the recurrence relation for the Chebyshev polynomials of second kind given in Lemma 3.58 (v) yields the generating function for U_n . \square

We list some of the most common polynomials with their corresponding generating functions $g(x, t)$ and the needed weight coefficients w_k in Table 4.1. The proof for the exponential generating function of the Chebyshev polynomials can be found in [Ces12]. For the other generating functions we refer to [AS64], pages 783-784. There exist different methods of obtaining generating function such as the Weisner method, which is based on partial differential equations and group transformation, or the Truesduell method based on differential-difference equations. A comprehensive overview on these methods can be found in [McB71].

name	weight-coeff. w_n	generating function
Chebyshev 1. kind T_n	1	$\frac{1-xt}{1-2xt+t^2}$ ($ t < 1$)
Chebyshev 1. kind T_n	$\frac{1}{n}$ ($n \geq 1$)	$\log\left(\frac{1}{\sqrt{1-2xt+t^2}}\right)$ ($ t < 1$)
Chebyshev 1. kind T_n	$\frac{1}{n!}$	$e^{tx} \cos(t\sqrt{1-x^2})$ ($t \in \mathbb{C}$)
Chebyshev 2. kind U_n	1	$\frac{1}{1-2xt+t^2}$ ($ t < 1$)
Chebyshev 2. kind U_n	$\frac{1}{n!}$	$e^{tx} \left(\cos(t\sqrt{x^2-1}) + \frac{x}{\sqrt{1-x^2}} \sin(t\sqrt{1-x^2}) \right)$ ($t \in \mathbb{C}$)
Legendre P_n	1	$\frac{1}{\sqrt{1-2xt+t^2}}$ ($ t < 1$)
Legendre P_n	$\frac{1}{n!}$	$e^{xt} J_0(t\sqrt{1-x^2})$ ($t \in \mathbb{R}$)
Laguerre L_n	1	$\frac{1}{1-t} e^{-tx/(1-t)}$ ($ t < 1$)
Hermite H_n	$\frac{1}{n!}$	e^{2xt-t^2} ($t \in \mathbb{C}$)

Table 4.1: Orthogonal polynomials with their corresponding generating function $g(x, t)$ and the corresponding weight coefficients $w_n, n \in \mathbb{N}_0$.

Definition 4.8 Let $M \in \mathbb{N}$. Then we call a signal f of the form

$$f(x) = \sum_{j=1}^M c_{n_j} Q_{n_j}(x) \tag{4.5}$$

M-sparse expansion into orthogonal polynomials Q_{n_j} with coefficients $c_{n_j} \in \mathbb{C} \setminus \{0\}$ and n_j are the indices of the “active” orthogonal polynomials for $j = 1, \dots, M$. We always assume that the sum representing f has minimal length, i.e., that $n_i \neq n_j$ for $i \neq j$ and the n_j to be ordered, i.e., $0 \leq n_1 < n_2 < \dots < n_M$.

Theorem 4.9 Let $\{Q_k \mid k \in \mathbb{N}_0\}$ with $Q_k \in L^2([a, b], \omega)$ be a set of orthogonal polynomials with corresponding inner product $\langle \cdot, \cdot \rangle_Q$ as in Definition 4.1 and generating function g as in (4.4). Let f be an M -sparse expansion into orthogonal polynomials as in (4.5) for $M \in \mathbb{N}$.

Furthermore, we suppose that there exists a $t \in \mathbb{C} \setminus \{0\}$ such that g is well-defined, i.e., uniformly convergent, for all $x \in [a, b]$ and t^k for $k \in \mathbb{N}$. Moreover, we assume that $t^{n_i} \neq t^{n_j}$ holds for $i \neq j$, where n_j denote the indices of the “active” orthogonal polynomials in f for $j = 1, \dots, M$. Then we can uniquely reconstruct the sparse expansion into orthogonal polynomials f as in (4.5) using the $2M$ samples $\langle f, g(\cdot, t^k) \rangle_Q$ for $k = 1, \dots, 2M$.

Proof. The proof is a constructive one. We use the definition of the signal f in (4.5) and the definition of the generating function given in (4.4) and apply the orthogonality

$$\begin{aligned} \langle f, g(\cdot, t) \rangle_Q &= \left\langle \sum_{j=1}^M c_{n_j} Q_{n_j}, g(\cdot, t) \right\rangle_Q = \sum_{j=1}^M c_{n_j} \left\langle Q_{n_j}, \sum_{k=0}^{\infty} w_k Q_k t^k \right\rangle_Q \\ &= \sum_{j=1}^M c_{n_j} \sum_{k=0}^{\infty} w_k t^k \underbrace{\langle Q_{n_j}, Q_k \rangle_Q}_{\delta_{n_j, k} d_{n_j}} = \sum_{j=1}^M c_{n_j} w_{n_j} d_{n_j} t^{n_j} \\ &= \sum_{j=1}^M \tilde{c}_{n_j} t^{n_j}, \end{aligned} \quad (4.6)$$

with $\tilde{c}_{n_j} := c_{n_j} w_{n_j} d_{n_j}$. Thus, this inner product “maps” the sparse expansion f into the orthogonal polynomials to a sparse monomial expansion. We note that the second step in the above equation is, indeed, possible for t as in the assumption. This monomial expansion can be recovered via the generalized Prony method. Therefore, we define the Prony polynomial

$$P(z) := \prod_{j=1}^M (z - t^{n_j}) = \sum_{k=0}^M p_k z^k.$$

We note that $p_M = 1$, employ (4.6) and consider the following equation for $\ell = 1, \dots, M$,

$$\begin{aligned} \sum_{k=0}^M p_k \langle f, g(\cdot, t^{(\ell+k)}) \rangle_Q &= \sum_{k=0}^M p_k \sum_{j=1}^M \tilde{c}_j t^{n_j(\ell+k)} \\ &= \sum_{j=1}^M \tilde{c}_j t^{n_j \ell} \underbrace{\sum_{k=0}^M p_k t^{n_j k}}_{=P(t^{n_j})=0} = 0. \end{aligned}$$

Exploiting the fact that $p_M = 1$ and the above equations one can compute the coefficients p_k , $k = 0, \dots, M-1$, of the Prony polynomial by considering the following Hankel system

$$\left(\langle f, g(\cdot, t^{(\ell+k)}) \rangle_Q \right)_{\ell=1, k=0}^{M, M-1} \cdot (p_0, \dots, p_{M-1})^T = - \left(\langle f, g(\cdot, t^{(\ell+M)}) \rangle_Q \right)_{\ell=1}^M$$

This matrix can be factorized in the following form

$$\begin{aligned}
 & \begin{pmatrix} \langle f, g(\cdot, t) \rangle_Q & \langle f, g(\cdot, t^2) \rangle_Q & \dots & \langle f, g(\cdot, t^M) \rangle_Q \\ \langle f, g(\cdot, t^2) \rangle_Q & \langle f, g(\cdot, t^3) \rangle_Q & \dots & \langle f, g(\cdot, t^{M+1}) \rangle_Q \\ \vdots & \vdots & \vdots & \vdots \\ \langle f, g(\cdot, t^M) \rangle_Q & \langle f, g(\cdot, t^{M+1}) \rangle_Q & \dots & \langle f, g(\cdot, t^{2M-1}) \rangle_Q \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 1 & \dots & 1 \\ t^{n_1} & t^{n_2} & \dots & t^{n_M} \\ \vdots & \vdots & \vdots & \vdots \\ t^{(M-1)n_1} & t^{(M-1)n_2} & \dots & t^{(M-1)n_M} \end{pmatrix} \cdot \begin{pmatrix} \tilde{c}_{n_1} & 0 & \dots & 0 \\ 0 & \tilde{c}_{n_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \tilde{c}_{n_M} \end{pmatrix} \cdot \begin{pmatrix} t^{n_1} & t^{2n_1} & \dots & t^{Mn_1} \\ t^{n_2} & t^{2n_2} & \dots & t^{Mn_2} \\ \vdots & \vdots & \vdots & \vdots \\ t^{n_M} & t^{2n_M} & \dots & t^{Mn_M} \end{pmatrix} \\
 &= \mathbf{V}_M \cdot \text{diag}(\tilde{c}_{n_1}, \dots, \tilde{c}_{n_M}) \cdot \text{diag}(t^{n_1}, \dots, t^{n_M}) \cdot \mathbf{V}_M^T,
 \end{aligned} \tag{4.7}$$

with $\mathbf{V}_M := \left(t^{n_j k} \right)_{k=0, j=1}^{M-1, M}$. Since we assumed that the sum determining f in (4.5) has minimal length we can conclude $\tilde{c}_{n_j} \neq 0$, $j = 1, \dots, M$ and $t^{n_j} \neq t^{n_i} \neq 0$ for $j \neq i$ by assumption. Therefore, the occurring Vandermonde matrices \mathbf{V}_M as well as the diagonal matrices are invertible, and, thus, the Hankel matrix itself is invertible.

Using the coefficients of the Prony polynomial we can compute its roots t^{n_j} , and, thus, the parameters n_j for $j = 1, \dots, M$, as well as the corresponding weights w_{n_j} .

As a last step, we now have to compute the coefficients c_{n_j} , $j = 1, \dots, M$, of f in (4.5). This can be done by solving the overdetermined linear system

$$\langle f, g(\cdot, t^k) \rangle_Q = \sum_{j=1}^M \tilde{c}_{n_j} t^{n_j k} \quad \text{for } k = 1, \dots, 2M, \tag{4.8}$$

and we obtain $c_{n_j} = \frac{\tilde{c}_{n_j}}{d_{n_j} w_{n_j}}$ for $j = 1, \dots, M$. Hence, the claim follows. \square

Remark 4.10 We omit the samples $\langle f(x), g(x, t^0) \rangle_Q = \langle f(x), g(x, 1) \rangle_Q$ since for a lot of generating functions $g(x, t)$ the corresponding series does not converge for $t = 1$ and therefore application of Theorem 4.9 would not be possible.

Furthermore, we want to emphasize that depending on the generating function the parameter t can be chosen as a primitive root of unity. The periodicity of the roots of unity imply that the condition $t^{n_j} \neq t^{n_k}$ for $j \neq k$ in Theorem 4.9 is, indeed, necessary. Since in practice the parameter n_j are unknown, we need $t^k \neq t^\ell$ for all $k, \ell \in \{0, \dots, 2MN\}$, where N is an upper bound on the degree of f . \diamond

We want to explore the connection between the generalized Prony method and the above mentioned reconstruction. Therefore, we consider the following operator.

Definition 4.11 (Dilation Operator) Let $h \in \mathbb{R} \setminus \{0\}$. Then the **dilation operator** $D_h: C(\mathbb{R}) \rightarrow C(\mathbb{R})$ is defined as

$$D_h f(x) := f(hx). \quad (4.9)$$

Remark 4.12 This operator is a special case of the generalized shift operator $S_{G,H,h}$ in (3.22) with $G(x) = \log(x)$ and $H \equiv 1$, i.e.,

$$S_{\log(\cdot),1,h} f(x) = f(e^{\log(x)+h}) = f(e^h x) = f(\tilde{h}x)$$

with $\tilde{h} = e^h$. In the previous section, we established that the eigenfunctions of this operator are monomials, see Table 3.1. Further information for the reconstruction using this operator can be found in [Pet14], Section 4.2. \diamond

The generating function reconstruction method does not use the linear operator for which the orthogonal polynomials are eigenfunctions but rather the dilation operator.

We recall (4.6) and consider the samples

$$\langle f, g(\cdot, t^k) \rangle_Q = \sum_{j=1}^M \tilde{c}_{n_j} t^{kn_j}$$

with $\tilde{c}_{n_j} = c_{n_j} w_{n_j} d_{n_j}$, $j = 1, \dots, M$, and $k \in \mathbb{N}$. We choose the point evaluation functional $F_1: C(\mathbb{R}) \rightarrow \mathbb{C}$ with $F_1(f) := f(1)$ and get

$$\sum_{j=1}^M \tilde{c}_{n_j} t^{kn_j} = \sum_{j=1}^M \tilde{c}_{n_j} F_1 \left(D_t^k x^{n_j} \right) = F_1 \left(D_t^k \left(\sum_{j=1}^M \tilde{c}_{n_j} x^{n_j} \right) \right). \quad (4.10)$$

Consequently, we can interpret (4.6) as "mapping" between the basis of orthogonal polynomials into a basis of monomials.

We summarize our results in the following algorithm.

Algorithm 5 Algorithm for the Reconstruction of Expansions into Orthogonal Polynomials Using a Generating Function

Input: $M \in \mathbb{N}$, $t \in \mathbb{C} \setminus \{0\}$ as in Theorem 4.9 and sampled values $\langle f, g(\cdot, t^\ell) \rangle_Q$, $\ell = 1, \dots, 2M$.

- 1: Form the Hankel matrix $\mathbf{H}_M := (\langle f, g(\cdot, t^{k+\ell}) \rangle_Q)_{\ell=0, k=1}^{M-1, M}$ as well as $\mathbf{f} := (\langle f, g(\cdot, t^{\ell+M}) \rangle_Q)_{\ell=1}^M$ and solve $\mathbf{H}_M \cdot \mathbf{p} = -\mathbf{f}$ for $\mathbf{p} := (p_\ell)_{\ell=0}^{M-1}$ as in (4.7).
- 2: Define the Prony polynomial $P(z) := \sum_{k=0}^M p_k z^k$ with $p_M = 1$, find all roots $\lambda_j = t^{n_j}$ of P , and determine the parameters $n_j := \log_t(\lambda_j)$ for $j = 1, \dots, M$. Determine the weights w_{n_j} corresponding to the n_j -orthogonal Polynomial Q_{n_j} in (4.4) as well as the parameters d_{n_j} in (4.2).
- 3: Solve the Vandermonde-like system $\mathbf{V}_{2M, M} \cdot \tilde{\mathbf{c}} = \mathbf{f}_{2M}$ with $\mathbf{V}_{2M, M} := (\lambda_j^k)_{k=1, j=1}^{2M, M}$, $\tilde{\mathbf{c}} = (\tilde{c}_{n_j})_{j=1}^M := (c_{n_j} d_{n_j} w_{n_j})_{j=1}^M$ and $\mathbf{f}_{2M} := (\langle f, g(\cdot, t^k) \rangle_Q)_{k=1}^{2M}$ and compute the coefficients $c_{n_j} = \frac{\tilde{c}_{n_j}}{d_{n_j} w_{n_j}}$ for $j = 1, \dots, M$, as in (4.8).

Output: c_{n_j} and n_j for $j = 1, \dots, M$ as in (4.5).

For application the required samples $\langle f, g(\cdot, t^\ell) \rangle_Q$ can be computed by numerical integration using Gaussian quadrature.

Theorem 4.13 Let $f \in C([a, b])$ and $n \in \mathbb{N}$. Furthermore, let $\{Q_k \mid k \in \mathbb{N}_0\}$ be a set of orthogonal polynomials in Definition 4.1 with corresponding weight function ω , and let x_i denote the i -th root of the the polynomial Q_n for $i = 1, \dots, n$. The **Gaussian quadrature** $I(f)$ of f is given by

$$I(f) = \sum_{i=1}^n \omega_i f(x_i) \quad (4.11)$$

with weights

$$\omega_i = \int_a^b \omega(x) \prod_{\substack{j=1 \\ i \neq j}}^n \frac{x - x_j}{x_i - x_j} dx \quad (4.12)$$

for $i = 1, \dots, n$. The Gaussian quadrature is exact for polynomials up to degree $2n - 1$, i.e.,

$$\int_a^b f(x) \omega(x) dx = I(f).$$

Proof. For the proof we refer to Theorem 3.6.12 in [SB02]. □

Theorem 4.14 Let $M \in \mathbb{N}$ and $\{Q_k \mid k \in \mathbb{N}_0\}$ with $Q_k \in L^2([a, b], \omega)$ be a set of orthogonal polynomials with corresponding inner product $\langle \cdot, \cdot \rangle_Q$ as in Definition 4.1 and generating function g in (4.4). Furthermore, let f be an M -sparse expansion into orthogonal polynomials as in (4.5) and $N \in \mathbb{N}$ be an upper bound on the degree of f , i.e., $\deg(f) < N$. We suppose

that there exists a $t \in \mathbb{C} \setminus \{0\}$ such that g is well-defined, i.e., uniformly convergent for all $x \in [a, b]$ and t^k for $k \in \mathbb{N}_0$. Moreover, we assume, that $t^\ell \neq t^k$ holds for all $k, \ell = 1, \dots, N$. Then the signal f can be recovered using only the N samples $f(x_k)$ with x_k denoting the k -th root of the polynomial Q_N for $k = 1, \dots, N$. In particular,

$$\langle f, g(\cdot, t^\ell) \rangle_Q = \sum_{i=1}^N \omega_i f(x_i) \sum_{n=0}^N t^{\ell n} Q_n(x_i) w_n \quad (4.13)$$

holds for $\ell = 1, \dots, 2M$ with ω_i as in (4.12) and w_n as in (4.4).

Proof. We introduce the partial sum

$$g_N(x, t) := \sum_{n=0}^N w_n t^n Q_n(x).$$

For fixed $t \in \mathbb{C}$ this is a polynomial of degree N , i.e., $\deg(g_N(\cdot, t)) = N$. We use (4.4) and consider for $\ell = 1, \dots, 2M$

$$\begin{aligned} \langle f, g(\cdot, t^\ell) \rangle_Q &= \int_a^b f(x) g(x, t^\ell) \omega(x) dx = \int_a^b f(x) \sum_{n=0}^{\infty} w_n (t^\ell)^n Q_n(x) \omega(x) dx \\ &= \int_a^b f(x) g_N(x, t^\ell) \omega(x) dx + \int_a^b f(x) \sum_{n=N+1}^{\infty} w_n t^{\ell n} Q_n(x) \omega(x) dx \\ &= \int_a^b f(x) g_N(x, t^\ell) \omega(x) dx + \sum_{n=N+1}^{\infty} w_n t^{\ell n} \sum_{j=1}^M c_{n_j} \int_a^b Q_{n_j}(x) Q_n(x) \omega(x) dx \end{aligned}$$

where the last equation follows by (4.5). Since N is an upper bound on the degree of f , we know that $\deg(f(\cdot) \cdot g_N(\cdot, t^\ell)) \leq \deg(f) + \deg(g_N(\cdot, t^\ell)) < 2N$ holds. Thus, $f g_N(\cdot, t^\ell)$ is a polynomial with degree of at most $2N - 1$ and, hence, Theorem 4.13 implies

$$\int_a^b f(x) g_N(x, t^\ell) \omega(x) dx = \sum_{i=1}^N \omega_i f(x_i) g_N(x_i, t^\ell),$$

with x_i the i -th root of the polynomial Q_N and ω_i as in (4.12), $i = 1, \dots, N$. Additionally, we know that $\deg(f) < N$ implies

$$\int_a^b Q_{n_j}(x) Q_n(x) \omega(x) dx = 0$$

for all $n \geq N + 1$ and Q_{n_j} the “active” orthogonal polynomials in f . Consequently, we can conclude

$$\langle f, g(\cdot, t^\ell) \rangle_Q = \sum_{i=1}^N \omega_i f(x_i) g_N(x_i, t^\ell).$$

Application of Theorem 4.9 yields the claim. □

In particular, we can conclude that the application of the Gaussian quadrature on the samples $\langle f, g(\cdot, t^\ell) \rangle_Q$ is sufficient for the reconstruction of a signal f .

Corollary 4.15 *Let $M \in \mathbb{N}$ and $\{Q_k \mid k \in \mathbb{N}_0\}$ with $Q_k \in L^2([a, b], \omega)$ be a set of orthogonal polynomials with corresponding inner product $\langle \cdot, \cdot \rangle_Q$ as in Definition 4.1 and generating function g in (4.4). Furthermore, let $M \in \mathbb{N}$ and f be an M -sparse expansion into orthogonal polynomials, i.e.,*

$$f(x) = \sum_{j=1}^M c_{n_j} Q_{n_j}(x)$$

and $N > \deg(f)$. We suppose that there exists a $t \in \mathbb{C} \setminus \{0\}$ such that g is uniformly convergent for all $x \in [a, b]$ and t^k for $k \in \mathbb{N}_0$. Moreover, we assume, that $t^\ell \neq t^k$ holds for all $k, \ell = 1, \dots, N$. Then we have

$$\left| \langle f, g(\cdot, t^\ell) \rangle_Q - \sum_{j=1}^N \omega_j f(x_j) g(x_j, t^\ell) \right| \leq \sum_{i=1}^N |\omega_i f(x_i)| \left\| g(x, t^\ell) - \sum_{n=0}^N w_n t^{\ell n} Q_n(x) \right\|_\infty$$

with ω_i as in (4.12), w_n as in (4.4), x_i the i -th root of the polynomial Q_N and $\ell = 1, \dots, 2M$, $i = 1, \dots, N$.

Proof. We want to apply Gaussian quadrature to the samples $\langle f, g(\cdot, t^\ell) \rangle_Q$ for $\ell = 1, \dots, 2M$ as in Theorem 4.13 and consider the corresponding error

$$\left| \langle f, g(\cdot, t^\ell) \rangle_Q - \sum_{i=1}^N \omega_i f(x_i) g(x_i, t^\ell) \right|$$

for ω_i as in (4.12) and x_i the i -th root of the orthogonal polynomial Q_N for $\ell = 1, \dots, 2M$. Application of Theorem 4.14 and, in particular, (4.13) and yields

$$\begin{aligned} \left| \langle f, g(\cdot, t^\ell) \rangle_Q - \sum_{i=1}^N \omega_i f(x_i) g(x_i, t^\ell) \right| &= \left| \sum_{i=1}^N \omega_i f(x_i) \sum_{n=0}^N w_n t^{\ell n} w_n Q_n(x_i) - \sum_{i=1}^N \omega_i f(x_i) g(x_i, t^\ell) \right| \\ &\leq \sum_{i=1}^N |\omega_i f(x_i)| \left| \sum_{n=0}^N w_n t^{\ell n} Q_n(x_i) - g(x_i, t^\ell) \right| \leq \sum_{i=1}^N |\omega_i f(x_i)| \left\| g(x, t^\ell) - \sum_{n=0}^N w_n t^{\ell n} Q_n(x) \right\|_\infty. \end{aligned}$$

□

In particular, if the generating function g converges uniformly and N is chosen large enough, the approximation obtained by the Gaussian quadrature is sufficient for the reconstruction since $\left\| g(x, t^\ell) - \sum_{n=0}^N w_n t^{\ell n} Q_n(x) \right\|_\infty \rightarrow 0$ for $N \rightarrow \infty$.

4.2 Reconstruction Methods Using Linear Operators

In the previous section we have developed an algorithm for the recovery of sparse expansions into orthogonal polynomials using a generating function and the generalized Prony method. In this section we want to take a look at two already existing methods for the recovery of such expansion using Prony's method and highlight the major differences. Both methods use a linear operator for which the orthogonal polynomials are eigenfunctions. This operator is also known as the Sturm-Liouville operator.

Definition 4.16 (Sturm-Liouville Operator) Let $[a, b] \subset \mathbb{R}$ a interval, then we can define the **Sturm-Liouville operator** $L_{p,q,\omega} : C^2([a, b]) \rightarrow C^1([a, b])$ by

$$L_{p,q,\omega}f(x) := \frac{1}{\omega} \left(-\frac{d}{dx} \left[p(x) \frac{d}{dx} f(x) \right] + q(x)f(x) \right) \quad x \in (a, b)$$

with $\omega, q \in C([a, b])$ and $p \in C^1([a, b])$ with $\omega(x) > 0$ and $p(x) > 0$ for all $x \in [a, b]$.

Remark 4.17 The problem of finding eigenfunctions and eigenvalues of the Sturm-Liouville operator $L_{p,q,\omega}$ on $C^2([a, b])$ together with the boundary conditions

- $\alpha_1 f(a) + \alpha_2 f'(a) = 0$ for $\alpha_1^2 + \alpha_2^2 > 0$,
- $\beta_1 f(b) + \beta_2 f'(b) = 0$ for $\beta_1 + \beta_2^2 > 0$

is called the **Sturm-Liouville problem**. Is it commonly studied in the field of non-linear partial differential evolution equations and has applications in quantum mechanics, for example the one-dimensional time-independent Schrödinger equation, see, e.g., [Mar86]. \diamond

Lemma 4.18 Let $[a, b] \subset \mathbb{R}$ an interval and $L_{p,q,\omega}$ be the Sturm-Liouville operator. Then the following holds.

- (i) If ω is the integrable weight function of $L^2([a, b], \omega)$, then the eigenfunctions of the Sturm-Liouville operator are orthogonal with respect to the inner product in (4.1).
- (ii) The eigenfunctions of the Sturm-Liouville operator are only polynomials if ω, p are chosen such that $-\frac{p}{\omega}$ and $-\frac{p'}{\omega}$ are polynomials of degree 2 and 1, respectively and $q \equiv 0$, i.e. the Sturm-Liouville operator can be simplified to

$$L_{\tilde{p},\tilde{q}}f := \tilde{p}f''(x) + \tilde{q}(x)f'(x) \tag{4.14}$$

with $\tilde{p} := -\frac{p}{\omega}$ and $\tilde{q} := -\frac{p'}{\omega}$.

Proof. For the proof of (i) we refer to [Pry93], Proposition 2.3 and for the proof of (ii) we refer to [Les62], Theorem 1. \square

Lemma 4.18 implies that we can use Theorem 2.5 in order to recover sparse expansions into orthogonal polynomials, see [PP13].

Theorem 4.19 Let $L_{p,q,\omega}$ be the Sturm-Liouville operator. Let p, ω and q be such that $-\frac{p}{\omega} =: \tilde{p}$ and $-\frac{p'}{\omega} =: \tilde{q}$ are polynomials of degree 2 and 1, respectively and $q \equiv 0$. Then the ℓ -th iteration of the Sturm-Liouville operator is given by

$$L_{p,q,\omega}^\ell f(x) = \sum_{k=1}^{2\ell} a_{k,\ell}(x) f^{(k)}(x) \quad f \in C^{2\ell}([a, b]) \quad (4.15)$$

for $\ell \in \mathbb{N}_0$ with $a_{1,1}(x) := \tilde{q}(x)$, $a_{2,1}(x) := \tilde{p}(x)$, and for $\ell \geq 2$, $a_{k,\ell}$ satisfies the recursion

$$\begin{aligned} a_{k,\ell}(x) = & k \left(\frac{k-1}{2} \tilde{p}''(x) + \tilde{q}'(x) \right) a_{k,\ell-1}(x) + ((k-1)\tilde{p}'(x) + \tilde{q}(x)) a_{k-1,\ell-1}(x) \\ & + \tilde{p}(x) a_{k-2,\ell-1}(x) \end{aligned} \quad (4.16)$$

with the convention $a_{k,\ell}(x) = 0$ for $\ell \geq 1$ and $k \notin \{1, \dots, 2\ell\}$. Moreover, let $\{Q_k \mid k \in \mathbb{N}_0\}$ be a set of orthogonal polynomials in Definition 4.1 with corresponding weight function ω and f be an M -sparse expansion into the orthogonal polynomials Q_k as in (4.5). Then f can be reconstructed using the derivative values $f^{(\ell)}(x_0)$ for $\ell = 0, \dots, 4M - 2$ and $x_0 \in [a, b]$.

Proof. For the proof of (4.15) with (4.16) we refer to [PP13], Theorem 4.1. Furthermore, Lemma 4.18 (i) and (ii) imply, that the orthogonal polynomials Q_k , $k \in \mathbb{N}_0$, are the eigenfunctions of the Sturm-Liouville operator. Therefore, we can apply Theorem 2.5 and the claim follows. \square

Although this theorem enables us to reconstruct an arbitrary finite expansion into orthogonal polynomials f , there is one major drawback of this approach.

In practice we usually want to reconstruct a signal or a function from sampled data, which come in form of functional values $f(x_j)$ for sampling points x_j with $j = 0, \dots, 2M - 1$ and not in form of derivative values $f^{(m)}(x_0)$ for $x_0 \in \mathbb{R}$ and $m = 0, \dots, 2M - 1$. Hence, the approach using the Sturm-Liouville operator is not practical.

A different approach for the reconstruction of orthogonal polynomials was introduced in [Sta18]. We will briefly discuss the method in order to compare it with the one we derived above. Therefore, we introduce the necessary definitions.

Definition 4.20 (Dual Space) Let V be a normed vector space over a field \mathbb{K} . Then the space $V^* := \{\varphi: V \rightarrow \mathbb{K} \mid \varphi \text{ linear}\}$ is called the **dual space** of V . Moreover, the mapping $(\cdot, \cdot): V \times V^* \rightarrow \mathbb{K}$ with $(v, \varphi) := \varphi(v)$ is called the **natural pairing** between V and its dual space V^* .

Definition 4.21 (Adjoint Operator) Let V be a Hilbert space, V^* be its dual space and (\cdot, \cdot) the corresponding natural pairing as in Definition 4.20. Moreover, let A be a linear operator defined on a dense subspace $\text{dom}(A) \subset V$ with $A: \text{dom}(A) \rightarrow V$. The

operator $A^* : \text{dom}(A^*) \rightarrow \text{dom}(A)$ is called the **(formal) adjoint operator** if it satisfies

$$(Av, \varphi) = (v, A^* \varphi) \quad (4.17)$$

for all $v \in \text{dom}(A)$ and $\varphi \in \text{dom}(A^*)$.

Remark 4.22 The above definition is usually given for unbounded operators A . If A is continuous and, hence, bounded, we have $\text{dom}(A) = V$. \diamond

Since this Chapter focusses on orthogonal polynomials and their expansion, we consider the Hilbert space $L^2([a, b], \omega)$.

Theorem 4.23 *Let $[a, b] \subset \mathbb{R}$ and let $L^2([a, b], \omega)$ be the vector space introduced in Definition 4.1. Moreover, let $L_{p,q,\omega}$ be the Sturm-Liouville operator given in Definition 4.16. The following holds.*

- (i) *The space $L^2([a, b], \omega)$ is self-dual and the natural pairing is the inner product $\langle \cdot, \cdot \rangle$ given in (4.1).*
- (ii) *If p, q are real functions, then the Sturm-Liouville operator is selfadjoint on all functions that satisfy the boundary conditions given in Remark 4.16.*

Proof. For the proof of (i) we refer to [Wer11], Theorem II.2.4. The proof for (ii) can be found in [Heu06], Section 3 pages 37–38. \square

Now, we can formulate the dual sampling approach in the context of orthogonal polynomials.

Theorem 4.24 *Let $\{Q_k \mid k \in \mathbb{N}_0\}$ be a set of orthogonal polynomials with $Q_k \in L^2([a, b], \omega)$ and inner product $\langle \cdot, \cdot \rangle_Q$ as in Definition 4.1. Moreover, let $M \in \mathbb{N}$ and f be an M -sparse expansion into the orthogonal polynomials Q_k as in (4.5) and $A : L^2([a, b], \omega) \rightarrow L^2([a, b], \omega)$ be a linear operator such that $AQ_k = \lambda_k Q_k$ holds for all $k \in \mathbb{N}_0$. Let A^* be the adjoint operator as in Definition 4.21 and $\phi \in \text{dom}(A^*)$ be fixed. Then f can be uniquely reconstructed using the samples $\langle f, (A^*)^\ell \phi \rangle_Q$ for $\ell = 0, \dots, 2M - 1$.*

Proof. Theorem 4.23 (i) implies, that the inner product $\langle \cdot, \cdot \rangle_Q$ coincides with the natural pairing on $L^2([a, b], \omega)$. Furthermore, since the orthogonal polynomials Q_k are eigenfunctions of A , we have $Q_k \in \text{dom}(A)$. Consequently, the claim is a direct application of Theorem 4.1.3. in [Sta18]. \square

Remark 4.25 The function $\phi \in \text{dom}(A^*)$ in the above theorem is called a **sampling kernel** in [Sta18]. Furthermore, we can interpret the above theorem as a special case of the generalized Prony method. Therefore, we define the linear functional

$F_\phi: L^2([a, b], \omega) \rightarrow \mathbb{R}$ with $F_\phi := \langle \cdot, \phi \rangle_Q$. Application of (4.17) yields

$$\langle f, (A^*)^\ell \phi \rangle_Q = \langle A^\ell f, \phi \rangle_Q = F_\phi \left(A^\ell f \right).$$

Hence, Theorem 2.5 can be applied in order to reconstruct the signal f . \diamond

While some first ideas of sampling kernels, adjoint operators and orthogonal polynomials were briefly discussed in [Sta18], the dual sampling approach remained rather theoretical. In [Sta18], pages 74–75 the choice for a sampling ϕ with

$$\phi(x) = \sum_{k=0}^{\infty} \beta_k Q_k(x) \quad (4.18)$$

for $|\beta_k| > 0$ and $\sum_{k=0}^{\infty} |\beta_k|^2 < \infty$ was briefly introduced but not further elaborated on. We take a closer look at these sampling kernels. Using the definition of the adjoint operator, we can conclude that the needed samples in Theorem 4.24 are of the form

$$\begin{aligned} \langle f, (A^*)^\ell \phi \rangle_Q &= \left\langle \sum_{j=1}^M c_{n_j} Q_{n_j}, \sum_{k=0}^{\infty} \beta_k (A^*)^\ell Q_k \right\rangle_Q = \sum_{k=0}^{\infty} \beta_k \sum_{j=1}^M c_{n_j} \langle Q_{n_j}, (A^*)^\ell Q_k \rangle_Q \\ &= \sum_{k=0}^{\infty} \beta_k \sum_{j=1}^M c_{n_j} \langle A^\ell Q_{n_j}, Q_k \rangle_Q = \sum_{k=0}^{\infty} \sum_{j=1}^M \beta_k c_{n_j} \lambda_{n_j}^\ell \langle Q_{n_j}, Q_k \rangle_Q \\ &= \sum_{j=1}^M c_{n_j} \beta_{n_j} \lambda_{n_j}^\ell d_{n_j}, \end{aligned}$$

where d_{n_j} denotes the constant given in (4.2).

We choose $\beta_k = t^k$ for $|t| < 1$ and $k \in \mathbb{N}_0$. We want to emphasize that this choice of sampling kernel is indeed possible, if we have uniform convergence of the series given in (4.18). This yields

$$\langle f, (A^*)^\ell \phi \rangle_Q = \sum_{j=1}^M c_{n_j} \lambda_{n_j}^\ell t^{n_j} d_{n_j}.$$

Comparing these samples with the ones used in Theorem 4.9, we observe that the major difference is the occurrence of eigenvalues λ_{n_j} in the dual sampling approach and therefore our above described method is more advantageous as it omits those parameters.

Furthermore, this highlights that we do not need the linear operator which has the orthogonal polynomials Q_k , $k \in \mathbb{N}_0$, as eigenfunctions for the reconstruction method based on the generating function.

We can specify the operator used in Theorem 4.24.

Theorem 4.26 *Let $\{Q_k \mid k \in \mathbb{N}_0\}$ be a set of orthogonal polynomials with corresponding inner product $\langle \cdot, \cdot \rangle_Q$ and generating function g as in (4.4) and let $t \in \mathbb{C} \setminus \{0\}$ such that g is*

well-defined for all $x \in [a, b]$. Let f be an M -sparse expansion into orthogonal polynomials as in (4.5) for $M \in \mathbb{N}$. Furthermore, let $L_{p,q,\omega}$ be the Sturm-Liouville operator given in Remark 4.16 with p, ω such that $-\frac{p}{\omega} =: \tilde{p}$ and $-\frac{p'}{\omega} =: \tilde{q}$ are polynomials of degree 2 and 1, respectively, and $q \equiv 0$. Then f can be reconstructed using the samples $\langle f, g(\cdot, t) \rangle_Q$ and

$$\sum_{k=1}^{2\ell} \langle f, a_{k,\ell}(\cdot) g^{(k)}(\cdot, t) \rangle_Q$$

for $\ell = 1, \dots, 2M - 1$ with $a_{1,1}(x) := \tilde{q}(x)$, $a_{2,1}(x) := \tilde{p}(x)$, and for $\ell \geq 2$, $a_{k,\ell}$ satisfies the recursion

$$\begin{aligned} a_{k,\ell}(x) = & k \left(\frac{k-1}{2} \tilde{p}''(x) + \tilde{q}'(x) \right) a_{k,\ell-1}(x) + ((k-1)\tilde{p}'(x) + \tilde{q}(x)) a_{k-1,\ell-1}(x) \\ & + \tilde{p}(x) a_{k-2,\ell-1}(x) \end{aligned}$$

with the convention $a_{k,\ell} = 0$ for $\ell \geq 1$ and $k \notin \{1, \dots, 2\ell\}$.

Proof. Theorem 4.18 (ii) implies that the Sturm-Liouville operator $L_{p,q,\omega}$ is self-adjoint with the orthogonal eigenfunctions Q_k for $k \in \mathbb{N}_0$. Therefore, application of Theorem 4.24 yields that the samples $\langle f, L_{p,q,\omega}^\ell g(\cdot, t) \rangle_Q$, $\ell = 0, \dots, 2M - 1$, are sufficient for the reconstruction of f . Theorem 4.19 yields that the values $L_{p,q,\omega}^\ell g(x, t)$ can be written as $L_{p,q,\omega}^\ell g(x, t) = \sum_{k=1}^{2\ell} a_{k,\ell}(x) g^{(k)}(x, t)$ with $a_{k,\ell}$ satisfying the recursion (4.16). \square

We note, that for the dual sampling method described above, we do not need derivative values of the M -sparse expansion f into orthogonal polynomials, but $4M - 2$ derivatives of the generating function g , which can be computed beforehand, since they are independent of the signal f .

Nevertheless, we need to compute $4M - 2$ inner products. Thus, in terms of needed samples our newly developed algorithm is favourable.

4.2.1 Quadratic Sampling Schemes

We also want to introduce a different possible choice for a sampling kernel ϕ in Theorem 4.24. Instead of ϕ as in (4.18), we can choose, a finite series as a sampling kernel, i.e.,

$$\phi(x) = \sum_{k=1}^N \beta_k Q_k(x),$$

with $\beta_k \in \mathbb{R} \setminus \{0\}$ and $N \in \mathbb{N}$. We want to apply Theorem 4.24 for the reconstruction of an M -sparse expansion into orthogonal polynomials. Thus, we take a closer look at

the samples, which are needed, i.e.,

$$\begin{aligned} \langle f, (A^*)^\ell \phi \rangle_Q &= \left\langle \sum_{j=1}^M A^\ell c_{n_j} Q_{n_j}, \sum_{k=1}^N \beta_k Q_k \right\rangle_Q \\ &= \sum_{j=1}^M \sum_{k=1}^N \lambda_{n_j}^\ell c_{n_j} \beta_k \underbrace{\langle Q_{n_j}, Q_k \rangle_Q}_{=\delta_{n_j, k} d_{n_j}} \quad \ell \in \mathbb{N}. \end{aligned}$$

In order to ensure that the above samples are not all vanishing, we need to assume that $N \geq n_j$ for all $j = 1, \dots, M$. Hence, we can simply choose

$$\phi(x) = \sum_{j=1}^M c_{n_j} Q_{n_j}(x) = f(x). \quad (4.19)$$

This yields a quadratic sampling scheme.

Theorem 4.27 *Let $\{Q_k \mid k \in \mathbb{N}\}$ be the set of orthogonal polynomials corresponding to the inner product $\langle \cdot, \cdot \rangle_Q$ on $L^2([a, b], \omega)$ as in Definition 4.1. Furthermore, let $A: L^2([a, b], \omega) \rightarrow L^2([a, b], \omega)$ be a linear operator such that $AQ_k = \lambda_k Q_k$ holds for all $k \in \mathbb{N}_0$. Moreover, let f be an M -sparse expansion of the form (4.5), i.e.,*

$$f(x) = \sum_{j=1}^M c_{n_j} Q_{n_j}(x)$$

with $c_{n_j} \in \mathbb{C} \setminus \{0\}$ for $j = 1, \dots, M$. Then the “active” indices n_j of the signal f can be uniquely reconstructed and the coefficients c_{n_j} can be reconstructed up to modulus using the samples $\langle A^\ell f, f \rangle_Q$ for $\ell = 0, \dots, 2M - 1$ and $j = 1, \dots, M$. If $c_{n_j} \in \mathbb{R}_+$ for $j = 1, \dots, M$ holds, then the reconstruction is unique.

Proof. The proof is a constructive one. We define the Prony polynomial

$$P(z) := \prod_{j=1}^M (z - \lambda_{n_j}) = \sum_{\ell=0}^M p_\ell z^\ell$$

with $p_M = 1$, which is determined by the (unknown) eigenvalues λ_{n_j} , corresponding to the “active” polynomials Q_{n_j} in (4.5). We consider $m \in \mathbb{Z}$, then

$$\begin{aligned} \sum_{\ell=0}^M p_\ell \langle A^{\ell+m} f, f \rangle_Q &= \sum_{\ell=0}^M p_\ell \left\langle A^{m+\ell} \left(\sum_{j=1}^M c_{n_j} Q_{n_j} \right), \sum_{k=1}^M c_{n_k} Q_{n_k} \right\rangle_Q \\ &= \sum_{\ell=0}^M p_\ell \sum_{j=1}^M \sum_{k=1}^M c_{n_j} c_{n_k} \langle A^{m+\ell} Q_{n_j}, Q_{n_k} \rangle_Q \end{aligned}$$

$$\begin{aligned}
 &= \sum_{\ell=0}^M p_\ell \sum_{j=1}^M \sum_{k=1}^M c_{n_j} c_{n_k} \lambda_{n_j}^{m+\ell} \underbrace{\langle Q_{n_j}, Q_{n_k} \rangle_Q}_{=\delta_{n_j, n_k} d_{n_j}} \\
 &= \sum_{j=1}^M c_{n_j}^2 d_{n_j} \lambda_{n_j}^m \underbrace{\sum_{\ell=0}^M p_\ell \lambda_{n_j}^\ell}_{=P(\lambda_{n_j})=0} = 0.
 \end{aligned}$$

Thus, we can solve the Hankel system

$$\left(\langle A^{m+\ell} f, f \rangle_Q \right)_{m, \ell=0}^{M-1} \cdot (p_0, p_1, \dots, p_{M-1})^T = - \left(\langle A^{M+\ell} f, f \rangle_Q \right)_{\ell=0}^{M-1}$$

for the coefficients of the Prony polynomial. Using the coefficients we can compute its roots λ_{n_j} for $j = 1, \dots, M$.

Finally, we solve the following linear system

$$\langle A^\ell f, f \rangle_Q = \sum_{j=1}^M c_{n_j}^2 d_{n_j} \lambda_{n_j}^\ell \quad \text{for } \ell = 0, \dots, 2M-1,$$

and compute $c_{n_j}^2 d_{n_j}$ and with this $|c_{n_j}|$ for $j = 1, \dots, M$. □

Corollary 4.28 Let $\{Q_k \mid k \in \mathbb{N}\}$ be the set of orthogonal polynomials corresponding to the inner product $\langle \cdot, \cdot \rangle_Q$ and $A: L^2([a, b], \omega) \rightarrow L^2([a, b], \omega)$ be a linear operator such that $AQ_k = \lambda_k Q_k$ for $k \in \mathbb{N}_0$. Furthermore, let f be an M -sparse expansion of the form (4.5), i.e.,

$$f(x) = \sum_{j=1}^M c_{n_j} Q_{n_j}(x)$$

with $c_{n_j} \in \mathbb{C} \setminus \{0\}$ for $j = 1, \dots, M$. Furthermore, let $x_k \in [a, b]$ be pairwise distinct sampling points for $k = 0, \dots, M$. Then the signal f can be uniquely reconstructed using the samples $\langle A^\ell f, f \rangle_Q$ and $f(x_k)$ for $\ell = 0, \dots, 2M-1$, $k = 1, \dots, M$, and $j = 1, \dots, M$.

Proof. We can apply Theorem 4.27 in order to recover the “active” indices n_j for $j = 1, \dots, M$, of the signal f . The coefficients c_{n_j} for $j = 1, \dots, M$ can be computed by solving the linear system

$$f(k) = \sum_{j=1}^M c_{n_j} Q_{n_j}(x_k) \quad \text{for } k = 1, \dots, M.$$

□

4.3 Reconstruction Method Using q -analogs

In this section we want to illustrate that the generalized Prony method can theoretically be employed for further and more general expansions into orthogonal polynomials if we can view them as eigenfunctions of suitable linear operators. Therefore, we use the Askey-Scheme and derive a approximation based approach for the reconstruction of orthogonal polynomials. The Askey-Scheme is a way of organizing orthogonal polynomials of hypergeometric or basic hypergeometric type into a hierarchy. Further information on this topic can be found in [KLS10]. We begin by introducing basic definitions of q -calculus. We want to mention that these results are of rather theoretical nature.

The following equation is the starting point for the theory of q -analogs,

$$\lim_{q \rightarrow 1} \frac{1 - q^n}{1 - q} = n \quad \text{for } n \in \mathbb{N}_0.$$

Using this, we define the q -analogs for factorials as well as for binomial coefficients.

Definition 4.29 (q -Bracket, q -Shifted Factorial) Let $n, k \in \mathbb{N}_0$ and $0 < q < 1$. Then we call the expression

$$[n]_q := \frac{1 - q^n}{1 - q} \tag{4.20}$$

the q -**bracket** of n . Furthermore, for $a \in \mathbb{R}$ we define the q -**shifted factorial** (or q -**analog of the Pochhammer symbol**) as

$$(a; q)_n := \prod_{k=0}^{n-1} (1 - aq^k) \tag{4.21}$$

with $(a; q)_0 := 1$ and $(a, q)_\infty := \prod_{k=0}^{\infty} (1 - aq^k)$.

Based on these definitions we can now construct a special differential operator.

Definition 4.30 (q -Differential Operator) Let $f: [0, \infty) \rightarrow \mathbb{C}$ be an arbitrary function and $0 < q < 1$. Then the q -**derivative** or **Jackson-derivative** of f at the point $x \in (0, \infty)$ is defined as

$$(\tilde{D}_q f)(x) := \begin{cases} \frac{f(qx) - f(x)}{x(q-1)} & x \neq 0 \\ f'(0) & x = 0. \end{cases}$$

We call \tilde{D}_q the q -**differential operator**. If the function f is differentiable, we have $\lim_{q \rightarrow 1} \tilde{D}_q f(x) = f'(x)$, see [KLS10], Section 1.15.

Remark 4.31 There also exists a corresponding q -integral. Let $f: [0, \infty) \rightarrow \mathbb{C}$ be an arbitrary real function and $0 < q < 1$. Then the q -integral or **Jackson-integral** of f on $(0, z)$ is defined as

$$\int_0^z f(t) d_q t := z(1-q) \sum_{n=0}^{\infty} f(q^n z) q^n.$$

The Jackson integral can also be defined on $(0, \infty)$ by

$$\int_0^{\infty} f(t) d_q(t) := (1-q) \sum_{n=-\infty}^{\infty} f(q^n) q^n.$$

For a continuous function f we have

$$\lim_{q \rightarrow 1} \int_0^z f(t) d_q(t) = \int_0^z f(t) dt.$$

◇

In the theory surrounding q -analogs one finds that for a lot of classical polynomials such as Laguerre and Legendre polynomials, there exists a q -analog version and for each Sturm-Liouville-equation there exists a q -difference equation. We will use this to approximate these classical polynomials using the q -approach.

Definition 4.32 (q -Hypergeometric Function) Let $0 < q < 1$, $n, j, k \in \mathbb{N}_0$ and

$a_1, a_2, \dots, a_j, b_1, b_2, \dots, b_k \in \mathbb{R}$. Then the basic hypergeometric or q -hypergeometric function for $|z| < 1$ is given by

$${}_j\phi_k \left[\begin{matrix} a_1 & a_2 & \dots & a_j \\ b_1 & b_2 & \dots & b_k \end{matrix}; q, z \right] = \sum_{n=0}^{\infty} \frac{(a_1, a_2, \dots, a_j; q)_n}{(b_1, b_2, \dots, b_k, q; q)_n} \left((-1)^n q^{\binom{n}{2}} \right)^{1+k-j} z^n$$

with $(a_1, a_2, \dots, a_j; q)_n := \prod_{m=1}^j (a_m; q)_n$ the q -shifted factorial as in (4.21) For $j = k + 1$ we get the special case

$${}_{k+1}\phi_k \left[\begin{matrix} a_1 & a_2 & \dots & a_{k+1} \\ b_1 & b_2 & \dots & b_k \end{matrix}; q, z \right] = \sum_{n=0}^{\infty} \frac{(a_1, a_2, \dots, a_{k+1}; q)_n}{(b_1, b_2, \dots, b_k, q; q)_n} z^n.$$

Definition 4.33 (Big q -Jacobi polynomials) Let $0 < q < 1$, $n \in \mathbb{N}_0$ and $a, b, c, x \in \mathbb{R}$. Then the **big q -Jacobi polynomials** P_n are defined as

$$P_n(x; a, b, c; q) := {}_3\phi_2 \left[\begin{matrix} q^{-n} & abq^{n+1} & x \\ aq & cq \end{matrix}; q, q \right].$$

The properties of the big q -Jacobi polynomials can be found in [KLS10], Chapter 14, Section 14.5. We summarize some in the following theorem.

Theorem 4.34 Let $0 < q < 1$, $a, b, c, x \in \mathbb{R}$ and $n, m \in \mathbb{N}_0$. Then the following holds.

(i) For $0 < aq < 1$, $0 \leq bq < 1$ and $c < 0$ the big q -Jacobi polynomials satisfy the following orthogonality relation

$$\int_{cq}^{aq} \frac{(a^{-1}x, c^{-1}x; q)_\infty}{(x, bc^{-1}x; q)_\infty} P_m(x; a, b, c; q) P_n(x; a, b, c; q) d_q x \\ = aq(1-q) \frac{(a, abq^2, a^{-1}c, ac^{-1}q; q)_\infty}{(aq, bq, cq, abc^{-1}q; q)_\infty} \frac{(1-abq)}{(1-abq^{2n+1})} \frac{(q, bq, abc^{-1}q; q)_n}{(aq, abq, cq; q)_n} (-acq^2)^n q^{\binom{n}{2}} \delta_{m,n}.$$

(ii) The big q -Jacobi polynomials obey the limiting behaviour

$$\lim_{q \rightarrow 1} P_n(x; q^\alpha, q^\beta, 0; q) = \frac{P_n^{(\alpha, \beta)}(2x-1)}{P_n^{(\alpha, \beta)}(1)},$$

and for arbitrary $\gamma \in \mathbb{R}_{\geq 0}$

$$\lim_{q \rightarrow 1} P_n(x; q^\alpha, q^\beta, -q^\gamma; q) = \frac{P_n^{(\alpha, \beta)}(x)}{P_n^{(\alpha, \beta)}(1)}.$$

Here, $P_n^{(\alpha, \beta)}(x)$ denotes the classical Jacobi polynomials as in Definition 4.2.

(iii) The big q -Jacobi polynomials satisfy the q -difference equation

$$q^{-n}(1-q^n)(1-abq^{n+1})x^2y(x) = B(x)y(qx) - [B(x) + D(x)]y(x) + D(x)y(q^{-1}x)$$

with $B(x) := aq(x-1)(bx-c)$ and $D(x) := (x-aq)(x-cq)$.

Therefore, we can reconstruct M -sparse expansions into big q -Jacobi polynomials, using the generalized Prony method and Theorem 2.5.

We use the difference equation given in 4.34(iii) and define the following linear operator.

Definition 4.35 (big q -Jacobi Sturm-Liouville operator) Let $0 < 1 < q$ and $\alpha, \beta \in \mathbb{R}$. Then the big q -Jacobi Sturm-Liouville operator $L_q^{(\alpha, \beta)}$ is given by

$$L_q^{(\alpha, \beta)} := \left(q^{\alpha+\beta+1} D_q - (q^{\alpha+\beta+1} + 1) \text{Id} + D_{\frac{1}{q}} \right) - \frac{1}{x} \left(q^{\alpha+\beta+1} D_q - q^{\alpha+1} (1+q^\beta) \text{Id} + q^{\alpha+1} D_{\frac{1}{q}} \right), \quad (4.22)$$

where $(D_q f)(x) = f(qx)$ is the dilation operator as in (4.9) and $\text{Id}f(x) = f(x)$ the identity operator.

The big q -Jacobi polynomials $P_n(x; q^\alpha, q^\beta, 0; q)$ as in Definition 4.33 are eigenfunctions of this operator with eigenvalues

$$\lambda_n^{(\alpha, \beta)} = q^{-n}(1-q^n)(1-q^{\alpha+\beta}q^{n+1}).$$

Therefore, we can use the generalized Prony method in order to uniquely reconstruct a sparse expansion into big q -Jacobi polynomials such as

$$f(x) = \sum_{j=1}^M c_{n_j} P_{n_j}(x; q^\alpha, q^\beta, 0; q) \quad (4.23)$$

using only the samples $(L_q^{(\alpha, \beta)})^k f(x_0)$ for $k = 0, \dots, 2M - 1$ and $x_0 \in \mathbb{R}$ with $x_0 \neq 0$. Hence, we can conclude the following theorem.

Theorem 4.36 *Let $M \in \mathbb{N}$, $0 < q < 1$ and $x_0 \in \mathbb{R}$ and f be an M -sparse expansion into big q -Jacobi polynomials, i.e., f is of the form (4.23). Then f can be reconstructed using the $2M$ samples $(L_q^{(\alpha, \beta)})^k f(x_0)$, $k = 0, \dots, 2M - 1$, with $L_q^{(\alpha, \beta)}$ as in (4.22).*

Remark 4.37 There also exists a continuous version for q -Jacobi polynomials. Let $0 < q < 1$, $n \in \mathbb{N}_0$ and $\alpha, \beta, x \in \mathbb{R}$ with $x = \cos(\theta)$. Then the **continuous q -Jacobi polynomials** $P_n^{(\alpha, \beta)}$ are defined as

$$P_n^{(\alpha, \beta)}(x|q) := \frac{(q^{\alpha+1}; q)_n}{(q; q)_n} {}_4\phi_3 \left[\begin{matrix} q^{-n} & q^{\alpha+\beta+n+1} & q^{0.5\alpha+0.25} e^{i\theta} & q^{0.5\alpha+0.25} e^{-i\theta} \\ q^{\alpha+1} & -q^{0.5(\alpha+\beta+1)} & -q^{0.5(\alpha+\beta+2)} & \end{matrix} ; q, q \right]$$

and they possess the following limiting behaviour

$$\lim_{q \rightarrow 1} P_n^{(\alpha, \beta)}(x|q) = P_n^{(\alpha, \beta)}(x),$$

where $P_n^{(\alpha, \beta)}$ denote the original Jacobi polynomials.

These continuous q -Jacobi polynomials also satisfy a q -difference equation, which is given in [KLS10], Chapter 14, Section 14.10, Equation 14.10.5, as follows,

$$(1 - q)^2 \tilde{D}_q[\tilde{w}(x; q^{\alpha+1}, q^{\beta+1} | q) \tilde{D}_q y(x)] + \lambda_n \tilde{w}(x; q^\alpha, q^\beta | q) y(x) = 0,$$

where \tilde{D}_q denotes the q -differential operator as in Definition 4.30. with

$$\begin{aligned} y(x) &= P_n^{(\alpha, \beta)}(x|q), \\ \lambda_n &= 4q^{-n-1}(1 - q^n)(1 - q^{n+\alpha+\beta+1}) \quad \text{and} \\ \tilde{w}(x; q^\alpha, q^\beta | q) &= \frac{\left| \frac{(e^{2i\theta}; q)_\infty}{(q^{0.5\alpha+0.25} e^{i\theta}, q^{0.5\alpha+0.75} e^{i\theta}, -q^{0.5\beta+0.25} e^{i\theta}, q^{0.5} e^{i\theta})_\infty} \right|^2}{\sqrt{1 - x^2}} \end{aligned}$$

for $x = \cos \theta$. Therefore, we can also construct an operator, such that the continuous q -Jacobi polynomials are eigenfunctions of this operator. Since such an operator is based on the q -differential operator and the function \tilde{w} , this is impractical. ◇

We want to use the limiting behaviour of the Jacobi polynomials to derive an approximation based Prony algorithm, i.e., we approximate the Sturm-Liouville operator given in (4.16) and its iterations by the big q -Jacobi Sturm-Liouville operator and its powers. This gives rise to the following Algorithm.

Algorithm 6 q -Analog Based Prony Algorithm for the Reconstruction of Orthogonal Polynomials

- Input:** $M \in \mathbb{N}$, $0 < q < 1$, $x_0 \neq 0$, $\alpha, \beta \in \mathbb{R}$ and $L_q^{(\alpha, \beta)} f(2x_0 - 1)$ for $k = 0, \dots, 2M - 1$
- 1: Set $f_k = (L_q^{(\alpha, \beta)} f)^k(2x_0 - 1)$ for $k = 0, \dots, 2M - 1$ and form the Hankel-matrix $\mathbf{H}_M := (f_{k+\ell})_{k, \ell=0}^{M-1}$ and solve the system $\mathbf{H}_M \cdot \mathbf{p} = -\mathbf{f}$ with $\mathbf{p} := (p_k)_{k=0}^{M-1}$ and $\mathbf{f} := (f_{M+k})_{k=0}^{M-1}$.
 - 2: Define $p_M = 1$ and find all roots λ_{n_j} , $j = 1, \dots, M$ of the polynomial $P(z) := \sum_{k=0}^M p_k z^k$ and compute n_j using $\lambda_{n_j} = q^{-n_j} - q^{\alpha+\beta+1} - 1 + q^{\alpha+\beta+1+n_j}$. Set \tilde{n}_j as the integer closest to n_j .
 - 3: Determine the unknowns \tilde{c}_{n_j} , $j = 1 \dots, M$ as the solution of the Vandermonde-like-system $\mathbf{V}_{2M, M} \cdot \tilde{\mathbf{c}} = \mathbf{f}_{2M}$ with $\mathbf{V}_{2M, M} := (\lambda_{n_j}^k)_{k=0, j=1}^{2M-1, M}$, $\tilde{\mathbf{c}} = (\tilde{c}_{n_j})_{j=1}^M := (c_{n_j} P_{\tilde{n}_j}(2x_0 - 1))_{j=1}^M$ and $\mathbf{f}_{2M} = (f_j)_{j=0}^{2M-1}$. Compute $c_{n_j} = \frac{\tilde{c}_{n_j}}{P_{\tilde{n}_j}(2x_0 - 1)}$ for $j = 1, \dots, M$.
- Output:** $c_{n_j} \in \mathbb{C} \setminus \{0\}$ and $\tilde{n}_j \in \mathbb{N}_0$ for $j = 1, \dots, M$ as approximations of the parameters in (4.5)
-

We remark that the iterated action $(L_q^{(\alpha, \beta)})^k$ is non trivial as the terms $(q^{\alpha+\beta+1} D_q - (q^{\alpha+\beta+1} + 1) \text{Id} + D_{\frac{1}{q}})$ and $\frac{1}{x} (q^{\alpha+\beta+1} D_q - q^{\alpha+1} (1 + q^\beta) \text{Id} + q^{\alpha+1} D_{\frac{1}{q}})$ do not commute since

$$\begin{aligned}
 & \left(q^{\alpha+\beta+1} D_q - (q^{\alpha+\beta+1} + 1) \text{Id} + D_{\frac{1}{q}} \right) \frac{1}{x} \left(q^{\alpha+\beta+1} D_q - q^{\alpha+1} (1 + q^\beta) \text{Id} + q^{\alpha+1} D_{\frac{1}{q}} \right) \\
 &= q^{2(\alpha+\beta)+1} \frac{1}{x} D_q^2 + \frac{1}{x} q^{\alpha+2} D_{\frac{1}{q}}^2 - [(q^{\alpha+\beta+1} (q^{\alpha+\beta+1} + q^{\alpha+\beta} + q^\alpha + 1))] \frac{1}{x} D_q \\
 & \quad + \frac{1}{x} [q^{2\alpha+\beta+1} + (q^{\alpha+\beta+1} + 1)(q^{\alpha+1} + q^{\alpha+\beta+1}) + q^{\alpha+\beta+2}] \text{Id} \\
 & \quad - \frac{1}{x} [(q^{\alpha+\beta+1} + 1)q^{\alpha+1} + (q^{\alpha+2} + q^{\alpha+\beta+2})] D_{\frac{1}{q}} \\
 & \neq q^{2(\alpha+\beta)+1} \frac{1}{x} D_q^2 + \frac{1}{x} q^{\alpha+1} D_{\frac{1}{q}}^2 - \frac{1}{x} [q^{\alpha+\beta+1} (2q^{\alpha+\beta+1} + q^{\alpha+1} + 1)] D_q \\
 & \quad + \frac{1}{x} [q^{\alpha+\beta+1} + q^{\alpha+1} (1 + q^\beta) (q^{\alpha+\beta+1} + 1) + q^{2\alpha+\beta+2}] \text{Id} \\
 & \quad - \frac{1}{x} [q^{\alpha+1} (q^{\alpha+\beta+1} + q^\beta + 2)] D_{\frac{1}{q}} \\
 &= \frac{1}{x} \left(q^{\alpha+\beta+1} D_q - q^{\alpha+1} (1 + q^\beta) \text{Id} + q^{\alpha+1} D_{\frac{1}{q}} \right) \left(q^{\alpha+\beta+1} D_q - (q^{\alpha+\beta+1} + 1) \text{Id} + D_{\frac{1}{q}} \right).
 \end{aligned}$$

Consequently, the above derived algorithm is of a rather theoretical nature.

Example 4.38 (A Toy Example) The Chebyshev polynomials of first kind $T_n(x)$ are given as $P_n^{(-0.5,-0.5)}$, i.e., $\alpha = \beta = -0.5$. We now consider $f(x) = cT_n(x)$ with parameters $c = 1$ and $n = 3$ that need to be recovered.

We want to sample $(L_q^{(\alpha,\beta)})^k f(2x-1)(x_0)$ for $k = 0, 1$ for $x_0 = 1$ and choose $q = 0.99$. Therefore, we have $q^{\alpha+\beta+1} = q^0 = 1$ and

$$L_q^{(-0.5,-0.5)} = \left(D_q - 2\text{Id} + D_{\frac{1}{q}}\right) - \frac{1}{x} \left(D_q - (1 + q^{0.5})\text{Id} + q^{0.5}D_{\frac{1}{q}}\right)$$

with the eigenvalues

$$\lambda_n^{(-0.5,-0.5)} = q^{-n}(1 - q^n)(1 - q^{-1}q^{n+1}).$$

Thus, we need the samples

$$\begin{aligned} f(2x-1)(1) &= 1, \\ (L_q^{(-0.5,-0.5)})f(2x-1)(1) &= (0.99^{0.5} - 1)f(2x-1)(1) + (1 - 0.99^{0.5}) \cdot f(2x-1)(q^{-1}) \\ &= (0.99^{0.5} - 1) + (1 - 0.99^{0.5}) \cdot 1.18675 = 0.000936096. \end{aligned}$$

This yields $p_0 = -0.000936096$ and $p_1 = 1$. Therefore, the Prony polynomial is given as $P(z) = -0.000936096 + z$. The root is $\lambda_n = 0.000936096 = 0.99^{-n}(1 - 0.99^n)^2$, and approximate this for $n \in \mathbb{N}_0$. The solution is given as $n \approx 3.04413$, and, thus, we get $n = 3$. Now we need to solve the linear system

$$\begin{aligned} 1 &= cT_3(1), \\ 0.000936096 &= cT_3(1)\lambda_3, \end{aligned}$$

which yields $c = 1$.

The reconstruction of the parameter c is exact and the error for the recovery of n is 0.04413. Further stability analysis lies beyond the scope of this thesis, but this result implies that the derived Algorithm is of instable and as of now not of practical use. \diamond

4.4 Numerical Examples and Applications

In this section we want to illustrate Algorithm 5 using different examples.

4.4.1 Reconstruction of Sparse Chebyshev Expansions of First Kind

We will begin by considering sparse expansions into Chebyshev polynomials, which we have also studied in Section 3.3.

We recall that the Chebyshev polynomials of first kind are a special case of the Jacobi polynomials for $\alpha = \beta = -\frac{1}{2}$. Therefore, Lemma 4.3 implies that they are orthogonal

with respect to the scalar product

$$\langle f, g \rangle_T := \int_{-1}^1 f(x)g(x) \frac{1}{\sqrt{(1-x^2)}} dx$$

with

$$\langle T_n, T_m \rangle_T = \int_{-1}^1 T_n(x)T_m(x) \frac{1}{\sqrt{(1-x^2)}} dx = \begin{cases} \pi & n = m = 0, \\ \frac{\pi}{2} & n = m \neq 0, \\ 0 & n \neq m. \end{cases}$$

We have seen the generating functions for Chebyshev polynomials of first kind in Table 4.1 and illustrate them in Figure 4.1.

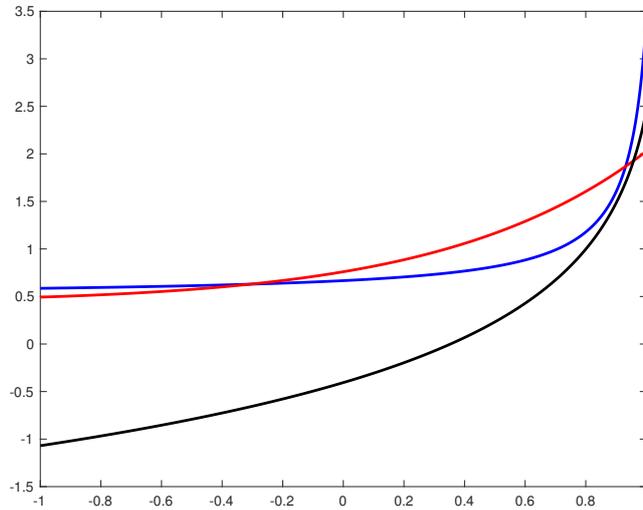


Figure 4.1: Different generating functions for the Chebyshev polynomials of first kind for the parameter $t = \frac{1}{\sqrt{2}}$. The red line shows the generating function for the weights $w_k = 1$ for $k \in \mathbb{N}_0$, the blue line the generating function for $w_k = \frac{1}{k!}$ for $k \in \mathbb{N}_0$ and the black line for the weights $w_k = \frac{1}{k}$ for $k \in \mathbb{N}$.

Thus, we can formulate the following theorem for the reconstruction of Chebyshev expansions using the generating function $g(x, t) = \frac{1-xt}{1-2xt+t^2}$.

Theorem 4.39 Let $M \in \mathbb{N}$, $|t| < 1$ and f be an M -sparse expansion into Chebyshev polynomials of first kind i.e.,

$$f(x) = \sum_{j=1}^M c_{n_j} T_{n_j}(x).$$

Then f can be reconstructed using the samples

$$\int_{-1}^1 f(x) \frac{1 - xt^\ell}{1 - 2xt^\ell + t^{2\ell}} \frac{1}{\sqrt{(1 - x^2)}} dx$$

for $\ell = 1, \dots, 2M$.

Proof. We employ the fact that $\frac{1-xt}{1-2xt+t^2}$ is a generating function for the Chebyshev polynomials. Thus, Theorem 4.9 yields the claim. \square

Corollary 4.40 *Let f be a Chebyshev expansion of the form (3.33) and $N \in \mathbb{N}$ an upper bound for the degree of the polynomial f . Furthermore, let $t \in \mathbb{R} \setminus \{0\}$ such that $|t| < 1$. The signal f can be reconstructed using the samples $f\left(\cos\left(\frac{(2k-1)\pi}{2N}\right)\right)$ for $k = 1, \dots, N$. Moreover, the absolute value of the error obtained from approximating the integral*

$$\int_{-1}^1 f(x) \frac{1 - xt^\ell}{1 - 2xt^\ell + t^{2\ell}} \frac{1}{\sqrt{1 - x^2}} dx \quad (4.24)$$

via Gauss quadrature can be bounded by

$$\pi \sum_{j=1}^M |c_{n_j}| \left\| \frac{1 - xt^\ell}{1 - 2xt^\ell + t^{2\ell}} - \sum_{n=0}^N t^{\ell n} T_n(x) \right\|_{\infty}$$

for $\ell = 1, \dots, M$ and $c_{n_j}, j = 1, \dots, M$, the coefficients of the sparse Chebyshev expansion f .

Proof. Theorem 4.7 implies that the ordinary generating function $g(x, t) = \frac{1-xt}{1-2xt+t^2}$ converges uniformly for all $x \in [-1, 1]$ and $|t| < 1$. We apply Theorem 4.14 and can reconstruct f using only the N samples $f(x_k), k = 1, \dots, N$, where x_k denotes the k -th root of the polynomial Q_N . Moreover, Lemma 3.4 (iii) yields $x_k = \cos\left(\frac{(2k-1)\pi}{2N}\right)$.

Lemma 3.4 (iv) and Corollary 4.15 imply that the absolute value of the error given by applying Gaussian quadrature to the integral (4.24) is bounded by

$$\begin{aligned} & \left| \langle f, g(\cdot, t^\ell) \rangle_Q - \sum_{j=1}^N \omega_j f(x_j) g(x_j, t^\ell) \right| \leq \sum_{i=1}^N |\omega_i f(x_i)| \left\| g(x, t^\ell) - \sum_{n=0}^N t^{\ell n} T_n(x) \right\|_{\infty} \\ & \leq \sum_{i=1}^N |\omega_i| \sum_{j=1}^M |c_{n_j}| \cdot |\cos(n_j \arccos(x_i))| \left\| g(x, t^\ell) - \sum_{n=0}^N t^{\ell n} T_n(x) \right\|_{\infty} \\ & \leq \sum_{i=1}^N |\omega_i| \sum_{j=1}^M |c_{n_j}| \left\| g(x, t^\ell) - \sum_{n=0}^N t^{\ell n} T_n(x) \right\|_{\infty}. \end{aligned}$$

The weights $\omega_i, i = 1, \dots, N$, given in (4.12) are $\omega_i = \frac{\pi}{N}$, see [AS64], Equation 25.4.38, and, hence, the claim follows. \square

As seen in Table 4.1 there is more than one generating function for the Chebyshev polynomials of first kind.

In the following we will now take a closer look at the reconstruction using the exponential generating function.

Theorem 4.41 *Let $M \in \mathbb{N}$ and f be an M -sparse Chebyshev expansion as in (3.33) and $t \in \mathbb{C}$ be a primitive N -th root of unity with $N \gg 2M$. Then f can be reconstructed using the samples*

$$\int_{-1}^1 f(x) e^{t^\ell x} \cos(t^\ell \sqrt{1-x^2}) dx$$

for $\ell = 1, \dots, 2M$.

Proof. The series $\sum_{k=0}^{\infty} \frac{1}{k!} t^k T_k(x)$ converges uniformly for all $t \in \mathbb{C}$ due to the Weierstrass M -test, since $|T_k(x)| \leq 1$ for $x \in [-1, 1]$ and $\sum_{k=0}^{\infty} \frac{1}{k!} t^k = e^t$ hold. Since $e^{tx} \cosh(t\sqrt{1-x^2})$ is a generating function for f , we can apply Theorem 4.9 and the claim follows. \square

Analogously, to the case of the ordinary generating function we approximate the needed samples using Gaussian quadrature.

Corollary 4.42 *Let $M \in \mathbb{N}$ and f be an M -sparse Chebyshev expansion of the form (3.33) and $N \in \mathbb{N}$ an upper bound on the degree of the polynomial f . Furthermore, let $t \in \mathbb{C}$ be a primitive L -th root of unity with $L \gg 2M$. Then the signal f can be reconstructed using the samples $f\left(\cos\left(\frac{(2k-1)\pi}{2N}\right)\right)$ for $k = 1, \dots, N$. Moreover, the absolute value of the error given obtained by approximating the integral*

$$\int_{-1}^1 f(x) e^{t^\ell x} \cos(t^\ell \sqrt{1-x^2}) \frac{1}{\sqrt{1-x^2}} dx \quad (4.25)$$

via Gauss quadrature can be bounded by

$$\pi \sum_{j=1}^M |c_{n_j}| \left\| e^{t^\ell x} \cos(t^\ell \sqrt{1-x^2}) - \sum_{n=0}^N \frac{t^{\ell n}}{n!} T_n(x) \right\|_{\infty}$$

for $\ell = 1, \dots, M$.

Proof. We consider the exponential generating function $g(x, t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} Q_n(x)$. The Weierstrass M -test implies that the right hand side is uniformly convergent for $x \in [-1, 1]$ and t as assumed. Theorem 4.14 yields that f can be reconstructed using the samples $f(x_k)$ with $x_k = \cos\left(\frac{(2k-1)\pi}{2N}\right)$, $k = 1, \dots, N$, denoting the roots of the N -th Chebyshev polynomial as in Lemma 3.4 (iii). Corollary 4.15 and Lemma 3.4 (iv) imply that the absolute value of the error of the Gaussian quadrature of the integral (4.25)

can be bounded by

$$\begin{aligned}
 & \left| \langle f, g(\cdot, t^\ell) \rangle_Q - \sum_{j=1}^N \omega_j f(x_j) g(x_j, t^\ell) \right| \leq \sum_{i=1}^N |\omega_i f(x_i)| \left\| g(x, t^\ell) - \sum_{n=0}^N \frac{t^{\ell n}}{n!} T_n(x) \right\|_\infty \\
 & \leq \sum_{i=1}^N |\omega_i| \sum_{j=1}^M |c_{n_j}| \cdot |\cos(n_j \arccos(x_i))| \left\| g(x, t^\ell) - \sum_{n=0}^N \frac{t^{\ell n}}{n!} T_n(x) \right\|_\infty \\
 & \leq \sum_{i=1}^N |\omega_i| \sum_{j=1}^M |c_{n_j}| \left\| g(x, t^\ell) - \sum_{n=0}^N \frac{t^{\ell n}}{n!} T_n(x) \right\|_\infty.
 \end{aligned}$$

Since $\omega_i = \frac{\pi}{N}$, see [AS64], Equation 25.4.38, the claim follows. \square

Example 4.43 We consider the following signal:

$$f(x) = -T_2(x) + 2.5T_4(x) + 0.7T_9(x) + 0.9T_{11}(x) - 1.2T_{14}(x). \quad (4.26)$$

This is a 5-sparse Chebyshev expansion. We use Algorithm 5, and compare the reconstruction using the different generating functions given in Table 4.1 to calculate approximations \tilde{c}_{n_j} and \tilde{n}_j of the original parameters c_{n_j} and n_j for $j = 1, \dots, M$ as shown in Table 4.2 and Table 4.3, respectively. For the ordinary generating function and the generating function with weights $\frac{1}{n}$, we choose $t = \frac{1}{\sqrt{2}}$ and for the exponential generating function we choose $t = \exp\left(\frac{\pi i}{15}\right)$. Furthermore, we use Gaussian quadrature with the samples $f\left(\cos\left(\frac{2k-1}{200}\right)\right)$ for $k = 1, \dots, 100$. The error in the reconstruction of the indices of the “active” basis polynomials T_{n_j} and the corresponding coefficients c_{n_j} for $j = 1, \dots, 5$ are shown in Table 4.2 and Table 4.3, respectively.

j	n_j	$ n_j - \tilde{n}_j (w_n = 1)$	$ n_j - \tilde{n}_j (w_n = \frac{1}{n})$	$ n_j - \tilde{n}_j (w_n = \frac{1}{n!})$
$j = 1$	2	$3.01248 \cdot 10^{-12}$	$2.14895 \cdot 10^{-12}$	$7.99312 \cdot 10^{-14}$
$j = 2$	4	$4.94558 \cdot 10^{-10}$	$7.18446 \cdot 10^{-10}$	$1.02602 \cdot 10^{-12}$
$j = 3$	9	$8.80341 \cdot 10^{-5}$	$2.94123 \cdot 10^{-4}$	$1.60405 \cdot 10^{-7}$
$j = 4$	11	0.00133	0.00547	$1.48308 \cdot 10^{-5}$
$j = 5$	14	0.00247	0.01291	0.00228

Table 4.2: Error in the numerical evaluation of indices of the “active” basis polynomials T_{n_j} of the sparse Chebyshev expansion (4.26) using Algorithm 5 and different generating functions.

j	c_{n_j}	$ c_{n_j} - \tilde{c}_{n_j} (w_n = 1)$	$ c_{n_j} - \tilde{c}_{n_j} (w_n = \frac{1}{n})$	$ c_{n_j} - \tilde{c}_{n_j} (w_n = \frac{1}{n!})$
$j = 1$	-1	$1.21369 \cdot 10^{-11}$	$8.66718 \cdot 10^{-12}$	$1.40923 \cdot 10^{-13}$
$j = 2$	2.5	$2.49116 \cdot 10^{-9}$	$3.63626 \cdot 10^{-09}$	$3.10319 \cdot 10^{-12}$
$j = 3$	0.7	0.00015	0.00049	$1.57102 \cdot 10^{-7}$
$j = 4$	0.9	0.00076	0.00312	$1.79510 \cdot 10^{-5}$
$j = 5$	-1.2	0.00063	0.00335	0.00437

Table 4.3: Error in the numerical evaluation of coefficients c_{n_j} of the sparse Chebyshev expansion (4.26) using Algorithm 5 and different generating functions.

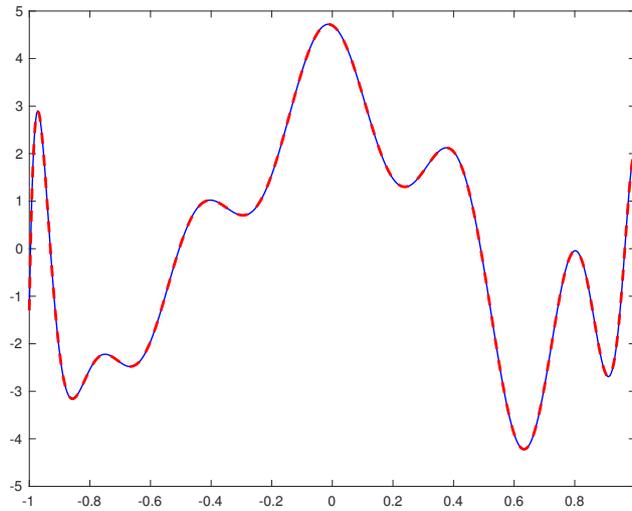


Figure 4.2: Original signal f of the form (4.26). The reconstructed signal was obtained by Corollary 4.42 and Algorithm 5 and is shown in red. The original signal f is shown in blue. Both signals cannot be distinguished visually.

Here, since we know that the degrees n_j of the polynomials are integers, we have rounded the reconstructed values \tilde{n}_j to the next integer before proceeding with the last step of Algorithm 5. The signal f and its reconstruction based on the exponential generating function are shown in Figure 4.2.

We observe, that the reconstruction method using the exponential generating function is slightly better than the reconstruction method using the other generating functions. In particular, we perceive that the reconstruction error is larger for the higher indices, which can be explained by the approximation error for the Gaussian quadrature derived in Corollary 4.42 and Corollary 4.40. \diamond

Example 4.44 We want to compare the reconstructions methods based on the different generating functions given in Table 4.1 in a more detailed way. Therefore, we consider 100 different randomly generated M -sparse Chebyshev expansions as in (3.33) with $M = 4$ and compare the average of the maximal reconstruction errors.

The “active” indices n_j have been drawn uniformly from $\{1, \dots, 17\}$ and the parameters c_{n_j} from $\{-20, \dots, 20\} \setminus \{0\}$ for $j = 1, \dots, 4$. Moreover, we compute the largest difference between the “active” indices n_j and the corresponding coefficients c_{n_j} respectively, and the parameters \tilde{n}_j and \tilde{c}_{n_j} recovered by Algorithm 5, for $j = 1, \dots, 4$.

Additionally, we average these values over the 100 calculations. We set $\mathbf{n} := (n_j)_{j=1}^4$, $\mathbf{c} = (c_{n_j})_{j=1}^4$, $\tilde{\mathbf{n}} := (\tilde{n}_j)_{j=1}^4$ and $\tilde{\mathbf{c}} = (\tilde{c}_{n_j})_{j=1}^4$, then, for each generating function we define the reconstruction failure err respectively by

$$\begin{aligned} err_n &:= \frac{1}{100} \sum_{j=1}^{100} \|\mathbf{n} - \tilde{\mathbf{n}}\|_{\infty}, \\ err_c &:= \frac{1}{100} \sum_{j=1}^{100} \|\mathbf{c} - \tilde{\mathbf{c}}\|_{\infty}. \end{aligned}$$

The result of this computation is presented in Table 4.4. We observe that all generating functions given in Table 4.1 give similar results.

One possible explanation for the obtained errors is that the absolute value of the coefficients c_{n_j} , $j = 1, \dots, 4$ can be relatively big, i.e., $1 < |c_{n_j}| \leq 20$. Furthermore, it is possible that the indices n_j , $j = 1, \dots, 4$, can all be relatively big and close to 17. Therefore, they may negatively impact the error of the Gaussian quadrature.

err	$w_n = 1$	$w_n = \frac{1}{n}$	$w_n = \frac{1}{n!}$
err_n	0.78169	0.85064	0.65250
err_c	0.22439	0.67256	0.32215

Table 4.4: Performance of the reconstruction of 4-sparse Chebyshev expansions using Algorithm 5 and different generating functions.

◇

4.4.2 Reconstruction of Sparse Legendre Expansions

In the following section we will illustrate the reconstruction method described in this chapter using Legendre polynomials. Therefore, we recall the definition of the Legendre polynomials and their most important properties.

Definition 4.45 (Legendre Polynomial) Let $n \in \mathbb{N}_0$ and $x \in [-1, 1]$. Then the **Legendre polynomial** P_n of degree n is defined as the solution of the differential equation

$$\frac{d}{dx} \left((1 - x^2) \frac{d}{dx} y(x) \right) + n(n + 1)y(x) = 0 \quad (4.27)$$

with $P_n(1) = 1$.

The Legendre polynomials are a special case of the Jacobi polynomials given in

Defintion 4.2 with $\alpha = \beta = 0$, see [AS64], Section 4.1. Therefore, Lemma 4.3 implies, that the Legendre polynomials are orthogonal with respect to the weight function $\omega \equiv 1$, i.e.,

$$\int_{-1}^1 P_n(x)P_m(x)dx = \frac{2}{2n+1}\delta_{n,m}. \quad (4.28)$$

Furthermore, according to Theorem 4.6, the ordinary generating function for the Legendre polynomials is given by

$$\sum_{n=0}^{\infty} P_n t^n = \frac{1}{\sqrt{1-2xt+t^2}}. \quad (4.29)$$

The convergence of the above series holds for $|t| < 1$, since

$$|P_n(x)t^n| \leq \underbrace{|P_n(x)|}_{\leq 1} |t|^n \leq |t|^n.$$

Thus, the Weierstrass M -test yields uniform convergence. Therefore, we arrive at the following theorem.

Theorem 4.46 *Let $M \in \mathbb{N}$ and f be an M -sparse expansion into Legendre polynomials, i.e.,*

$$f(x) = \sum_{j=1}^M c_{n_j} P_{n_j}(x) \quad (4.30)$$

with $c_{n_j} \in \mathbb{C} \setminus \{0\}$ and n_j being the indices of the “active” Legendre polynomials for $j = 1, \dots, M$ with $0 \leq n_1 < n_2 < \dots < n_M$. Then the signal f can be reconstructed using the samples

$$\int_{-1}^1 f(x) \frac{1}{\sqrt{1-2xt^\ell+t^{2\ell}}} dx \quad (4.31)$$

for $\ell = 1, \dots, 2M$.

Proof. Equation (4.29) yields, that $\frac{1}{\sqrt{1-2xt+t^2}}$ is the ordinary generating function for the Legendre polynomials. Therefore, we can apply Theorem 4.9 and the claim follows. \square

Corollary 4.47 *Let $M \in \mathbb{N}$ and f be an M -sparse expansion into Legendre polynomials as in (4.30) and $N \in \mathbb{N}$ an upper bound on the degree of the polynomial f . Furthermore, let $t \in \mathbb{R} \setminus \{0\}$ such that $|t| < 1$. Let x_k , $k = 1, \dots, N$, be the k -th root of the N -th Legendre polynomial P_N . Then the signal f can be reconstructed using the samples $f(x_k)$ for $k = 1, \dots, N$. Moreover, the absolute value of the error obtained from approximating the*

integral

$$\int_{-1}^1 f(x) \frac{1}{1 - 2xt^\ell + t^{2\ell}} dx \tag{4.32}$$

via Gauss quadrature can be bounded by

$$\sum_{i=1}^N \frac{2}{|(1 - x_i)|(P'_N(x_i))^2} \left| \sum_{j=1}^M |c_{n_j}| \left\| \frac{1}{1 - 2xt^\ell + t^{2\ell}} - \sum_{n=0}^N t^{\ell n} P_n(x) \right\|_\infty \right|$$

for $\ell = 1, \dots, M$ and $c_{n_j}, j = 1, \dots, M$, the coefficients of the sparse Legendre expansion f .

Proof. The generating function $g(x, t) = \frac{1}{1 - 2xt + t^2}$ converges uniformly. Theorem 4.14 yields that f can be reconstructed from the values $f(x_i)$, where x_i denotes the i -th root of the Legendre polynomial P_N . The weights ω_i of the Gaussian quadrature are given by

$$\omega_i = \frac{2}{(1 - x_i)(P'_N(x_i))^2}$$

for $i = 1, \dots, N$, see [AS64], Equation 23.4.32. Hence, $|P_n(x)| \leq 1$ for all $n \in \mathbb{N}_0$ and application of Corollary 4.15 yield the claim. \square

Example 4.48 We want to illustrate the reconstruction method with an example. Therefore, we consider a 4-sparse Legendre expansion. We use Algorithm 5 to calculate approximations \tilde{n}_j and \tilde{c}_{n_j} of the original parameters n_j and c_{n_j} for $j = 1, \dots, 4$ as shown in Table 4.5. The needed samples were approximated using Gaussian quadrature of order $N = 50$. The roots of the Legendre polynomials were computed using the Newton method with initial guess $y_k = \cos\left(\frac{(2k+1)\pi}{(2N)}\right) + \left(\frac{0.27}{N}\right) \sin\left(\pi\left(-1 + \frac{2k}{N}\right)\frac{N-1}{N+1}\right)$ for the k -th root with $k = 0, \dots, N - 1$. Furthermore, we choose $t = \frac{1}{\sqrt{2}}$. Since we know that the orders n_j of the polynomials are integers, we have rounded the values of the reconstructed values $\tilde{n}_j, j = 1, \dots, 4$, to the next integer before proceeding with the last step of Algorithm 5. Even though we obtain small errors in the recovered parameters, the reconstruction is accurate, as depicted in Figure 4.3.

j	n_j	c_{n_j}	\tilde{n}_j	\tilde{c}_{n_j}
$j = 1$	1	-1	0.999999999999999	-0.999999999999998
$j = 2$	7	1	6.99999999404930	0.99999998854422
$j = 3$	11	2	10.9999989830414	1.99999839505289
$j = 4$	25	-2	25.0193267666628	-2.01316759005214

Table 4.5: Numerical evaluation of the indices of the “active” basis polynomials and corresponding coefficients of the sparse Legendre expansion (4.30) using Algorithm 5 and the ordinary generating function.

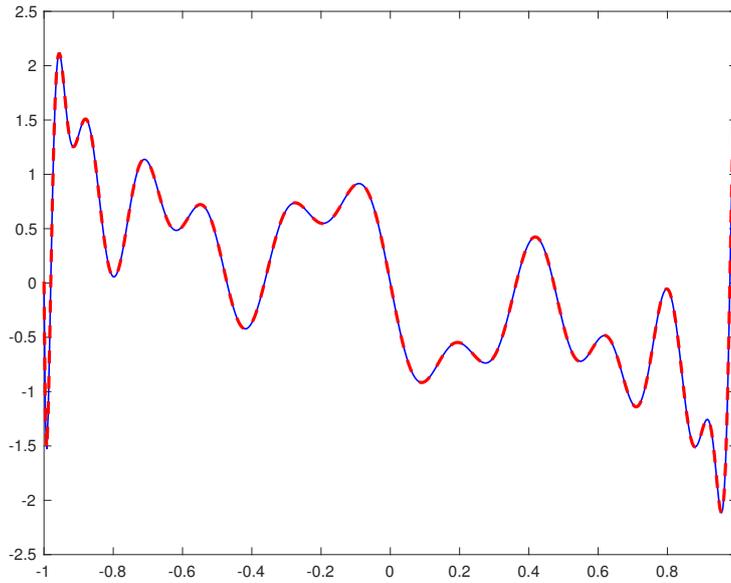


Figure 4.3: 4-sparse Legendre expansion with parameters given in Table 4.5 and its reconstruction. The reconstructed signal was obtained by Algorithm 5 and is shown in red. The original signal is shown in blue. Both signals cannot be distinguished visually.

◇

Example 4.49 Finally, we also want to compare our Algorithm with the (classical) “brute force” Gaussian quadrature approach. Therefore, we consider an M -sparse expansion f into Legendre polynomials as in (4.30) and N an upper bound on the degree of f . In order to determine the “active” Legendre polynomials P_{n_j} as well as the corresponding coefficients c_{n_j} , $j = 1, \dots, M$ we compute the inner products

$$\langle f, P_\ell \rangle_P = \left\langle \sum_{j=1}^M c_{n_j} P_{n_j}, P_\ell \right\rangle_P = \sum_{j=1}^M c_{n_j} \langle P_{n_j}, P_\ell \rangle_P = \begin{cases} c_{n_j} \frac{2}{2n_j+1} & \ell = n_j, \\ 0 & \ell \neq n_j \end{cases} \quad (4.33)$$

for $\ell = 1, \dots, N$. Hence, the “active” Legendre polynomials P_{n_j} are obtained by the non-vanishing inner products and the corresponding coefficients c_{n_j} can be computed by

$$c_{n_j} = \frac{\langle f, P_{n_j} \rangle_P (2n_j + 1)}{2}$$

for $j = 1, \dots, M$. If the parameter M is known, $\deg(f) + 1$ inner products are sufficient for the reconstruction. If M is unknown, we need to compute $N + 1$ inner products for the recovery of the M -sparse expansion. We use Gaussian quadrature of order $2N$ and note that Theorem 4.13 implies, that the Gaussian quadrature is exact, since we have

$\deg(f \cdot P_j) \leq \deg(f) + \deg(P_j) < 2N$. In particular, we need N samples for each inner product. In fact, this method is not only applicable for Legendre polynomials but for all types of orthogonal polynomials considered in Table 4.1.

We consider the following 5-sparse expansion into Legendre polynomials:

$$f(x) = 0.5P_0(x) - 4P_5(x) + 0.2P_8(x) + 8P_{11}(x) - 3P_{21}(x). \quad (4.34)$$

We use the classical Gaussian quadrature approach as described above and compare the results with the recovery obtained by Algorithm 5 using the ordinary generating function. In both cases we use the upper bound $N = 30$. Furthermore, we note that even though the Gaussian quadrature is exact for the inner products considered in (4.33), we may obtain small errors due to machine precision. Therefore, we consider a inner product to be vanishing if it is smaller than a predefined threshold parameter $\varepsilon > 0$. For our numerical example we use $\varepsilon = 10^{-14}$. The recovered values \tilde{n}_j and \tilde{c}_j , for $j = 1, \dots, 5$, for the approach by Gaussian quadrature and Algorithm 5 are shown in Table 4.6.

j	n_j	c_{n_j}	\tilde{n}_j (Algo. 5)	\tilde{c}_{n_j} (Algo. 5)	\tilde{n}_j (Quadrature)	\tilde{c}_j (Quadrature)
$j = 1$	0	0.5	$4.1644 \cdot 10^{-15}$	0.5000000000000006	0	0.5
$j = 2$	5	-4	5.0000	-3.999999951507359	5	-4.0000000000000007
$j = 3$	8	0.2	8.0001	0.200024082438281	8	0.1999999999999998
$j = 4$	11	8	11.0001	8.000295666331050	11	8.0000000000000005
$j = 5$	21	-3	20.8985	2.902549580175646	21	-3.0000000000000013

Table 4.6: Numerical evaluation of the indices of the “active” basis polynomials and corresponding coefficients of the sparse Legendre expansion (4.34) using the Gaussian quadrature approach as well as Algorithm 5 and the ordinary generating function.

We note, that the identification of the indices of the “active” Legendre polynomials is indeed exact by design of the algorithm. Due to the exactness of the Gaussian quadrature, we only obtain errors due to machine epsilon and hence, the recovery of the sparse Legendre-expansion using the “brute force”-Gaussian quadrature approach yields better results. Moreover, we also compare the run-time of both reconstruction methods. For Algorithm 5 we have an elapsed run-time of 67.46 seconds and for the Gaussian quadrature approach we have an elapsed runtime of 410.12 seconds . Thus, in terms of runtime as well as number of needed functional values Algorithm 5 is advantageous. Furthermore, Figure 4.4 shows that both recovery methods are highly accurate.

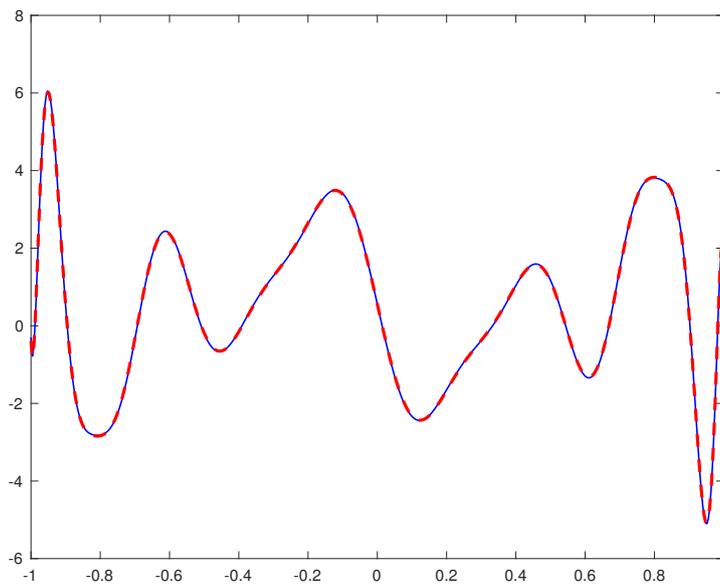


Figure 4.4: Recovery of 5-sparse Legendre expansion with parameters given in Table 4.6 with Algorithm 5 shown in red. The blue signal shows the 5-sparse Legendre expansion recovered by the Gaussian quadrature approach. Both signals cannot be distinguished visually.

◇

5 Numerical Behaviour of the Prony Method for Generalized Exponential Sums

Even though Prony's method is rather simple and relies mostly on linear algebra, one major drawback is its high instability with regard to noise, see [LC56], pages 276-280. Here, the author aims to recover the function

$$f(x) = 0.0951e^x + 0.8607e^{3x} + 1.5576e^{5x}$$

using 24 rounded samples of the form $f(0.05k)$ for $k = 0, \dots, 23$. Even though the reconstructed signal \tilde{f} fits the given data, i.e., $f(0.05k) = \tilde{f}(0.05k)$ for $k = 0, \dots, 23$, instead of the $M = 3$ exponents Prony's method only yields two frequencies $\alpha_1 = 1.58$ and $\alpha_2 = 4.45$. The corresponding recovered weights c_1 and c_2 are distorted by a factor of 7. This small example gives an insight of the effect of noisy data on the parameter estimation using Prony's method.

Hence, in recent years considerable research has been dedicated to the development of numerically stable Prony methods such as the Approximate Prony method, see [PT11], the ESPRIT method, see [RK89], and the Matrix Pencil method, see [HS90].

Thus, this chapter is focused on the modification of such stabilizing algorithms for generalized exponential sums.

In Section 5.1 we will introduce a modification of a subspace method known as the ESPRIT algorithm for generalized exponential sums.

Furthermore, we will elaborate on how this approach can be modified if one or more more frequencies are known beforehand in Section 5.2.

In the Section 5.3 we focus on the problem of clustered frequencies and derive a sub-sampling based recovery method for an improved recovery in this case.

Finally, in Section 5.4 we will survey a modification of Prony's method in order to solve the problem of best approximation of a given data vector by a vector of equidistant samples of a generalized exponential sum in the 2-norm.

In Section 5.5 we provided some numerical results for the derived algorithms. Partial results of this chapter have been published in our survey [KP21].

5.1 Modification of the ESPRIT Algorithm for Generalized Exponential Sums

One of the most used methods for the stable recovery of sparse exponential sums is the estimation of signal parameters via rotational techniques, a subspace method better known as the ESPRIT method, see [RK89]. In [PT13], it was shown that the ESPRIT method can be seen as a variant of Prony's method. We will generalize and modify this approach and apply it in the special case of generalized exponential sums.

Let us recall the some important notation and definitions given in the previous chapters.

We will again consider generalized exponential sums (3.12) as in Chapter 3, i.e.,

$$f(x) = \sum_{j=1}^M c_j H(x) e^{\alpha_j G(x)} \quad x \in [a, b], \quad (5.1)$$

with $c_j \in \mathbb{C} \setminus \{0\}$, $\alpha_j \in \mathbb{C}$ with $|\operatorname{Im}(\alpha_j)| < K$, $K > 0$, for $j = 1, \dots, M$, with G, H satisfying the restrictions given in Definition 3.15, i.e., $G, H \in C([a, b])$, G strictly monotone and H non-vanishing.

We assume that only an upper bound L on the number of terms M in (5.1) is known and oversample our signal f with $N \geq L \geq M$. Therefore, we need to choose $0 < |h| < \frac{\pi}{K}$ and $x_0 \in [a, b]$ such that the assumption of Theorem 3.29 are satisfied, i.e.,

$$\begin{aligned} G(x_0) + hk &\in [G(a), G(a)] & \text{for } G(a) < G(b) & \text{or} \\ G(x_0) + hk &\in [G(b), G(a)] & \text{for } G(a) > G(b) \end{aligned}$$

for all $k = 0, \dots, 2N - 1$. Furthermore, we assume the $2N$ data samples $f(G^{-1}(G(x_0) + kh))$, $k = 0, \dots, 2N - 1$, $h \in \mathbb{R} \setminus \{0\}$ with $N \geq L \geq M$ are given.

In order to simplify notation we define

$$f_k := S_{G,H,h}^k f(x_0) = \frac{H(x_0) f(G^{-1}(hk + G(x_0)))}{H(G^{-1}(hk + G(x_0)))} \quad k = 0, \dots, 2N - 1, \quad (5.2)$$

where $S_{G,H,h} f(x) := \frac{H(x)}{H(G^{-1}(G(x)+h))} f(G^{-1}(G(x)+h))$ for $h \in \mathbb{R} \setminus \{0\}$ denotes the generalized shift operator as in Definition 3.26 and f is a generalized exponential sum as in (5.1). Furthermore, we recall the definition of the corresponding Prony polynomial P given in (2.9), i.e.,

$$P(z) = \prod_{j=1}^M (z - \lambda_j) = \sum_{k=0}^M p_k z^k \quad (5.3)$$

with $\lambda_j := e^{\alpha_j h}$ for $h \in \mathbb{C} \setminus \{0\}$ for $j = 1, \dots, M$. Here, the parameters λ_j are the eigenvalues of the generalized shift operator $S_{G,H,h}$ corresponding to the eigenfunctions $e^{\alpha_j G(x)} H(x)$ for $j = 1, \dots, M$.

Using this notation we recall, that due to Theorem 2.5 and Theorem 3.31, the following equation holds true

$$\begin{aligned}
 \sum_{k=0}^M p_k f_{m+k} &= \sum_{k=0}^M p_k S_{G,H,h}^{k+m} f(x_0) = \sum_{k=0}^M p_k S_{G,H,h}^{k+m} \sum_{j=1}^M c_j H(x_0) e^{\alpha_j G(x_0)} \\
 &= \sum_{j=1}^M c_j \sum_{k=0}^M p_k S_{G,H,h}^{k+m} \left(H(x) e^{\alpha_j G(x)} \right) (x_0) \\
 &= \sum_{j=1}^M c_j \sum_{k=0}^M p_k e^{\alpha_j h(k+m)} H(x_0) e^{\alpha_j G(x_0)} \\
 &= \sum_{j=1}^M c_j e^{\alpha_j h m} H(x_0) e^{\alpha_j G(x_0)} \sum_{k=0}^M p_k e^{\alpha_j h k} = 0.
 \end{aligned} \tag{5.4}$$

Definition 5.1 Let $N, L \in \mathbb{N}$ with $N \geq L$ and let f_k be the noiseless data samples given in (5.2) for $k = 0, \dots, 2N - 1$. Then we define the **rectangular Hankel matrix** $\mathbf{H}_{2N-L, L+1}$ as

$$\mathbf{H}_{2N-L, L+1} := (f_{\ell+m})_{\ell, m=0}^{2N-L-1, L} \in \mathbb{C}^{(2N-L) \times (L+1)}. \tag{5.5}$$

Furthermore, we define the **Hankel sub-matrices** $\mathbf{H}_{2N-L, L}(0) \in \mathbb{C}^{2N-L, L}$ and $\mathbf{H}_{2N-L, L}(1) \in \mathbb{C}^{2N-L, L}$ given by

$$\mathbf{H}_{2N-L, L+1} = \begin{pmatrix} \mathbf{H}_{2N-L, L}(0) & (f_{\ell+L})_{\ell=0}^{2N-L-1} \end{pmatrix} = \begin{pmatrix} (f_{\ell})_{\ell=0}^{2N-L-1} & \mathbf{H}_{2N-L, L}(1) \end{pmatrix},$$

i.e., we obtain $\mathbf{H}_{2N-L, L}(0)$ by removing the last column of $\mathbf{H}_{2N-L, L+1}$ and $\mathbf{H}_{2N-L, L}(1)$ by removing the first column of $\mathbf{H}_{2N-L, L+1}$. Introducing the sub-matrix notation $\mathbf{A}(a : b, c : d)$ to select the rows a to b and the columns c to d of \mathbf{A} , we can write

$$\mathbf{H}_{2N-L, L}(0) := \mathbf{H}_{2N-L, L+1}(1 : 2N - L, 1 : L) \tag{5.6}$$

$$\mathbf{H}_{2N-L, L}(1) := \mathbf{H}_{2N-L, L+1}(1 : 2N - L, 2 : L + 1). \tag{5.7}$$

Using equation (5.4) we can determine the rank of the above matrices.

Lemma 5.2 Let $M, L, N \in \mathbb{N}$ with $M \leq L \leq N$ be given. Furthermore, let f_k be the noiseless data samples given in (5.2) for $k = 0, \dots, 2N - 1$ and $\mathbf{H}_{2N-L, L+1}$ and $\mathbf{H}_{2N-L, L}(s)$ as in Definition 5.1 for $s = 0, 1$. Then the following holds:

(i)

$$\text{rank}(\mathbf{H}_{2N-L, L+1}) = \text{rank}(\mathbf{H}_{2N-L, L}(s)) = M \text{ for } s = 0, 1.$$

(ii) If $L = M$, then $\ker(\mathbf{H}_{2N-M, M+1}) = \text{span}(\mathbf{p})$ and $\ker(\mathbf{H}_{2N-M, M+1}(s)) = \text{span}(\mathbf{0}_{M+1})$

with the singular values

$$\sigma_1(\mathbf{H}_{2N-M,M+1}) \geq \sigma_2(\mathbf{H}_{2N-M,M+1}) \geq \cdots \geq \sigma_M(\mathbf{H}_{2N-M,M+1}) > 0.$$

Additionally, we recall the definition of the companion matrix from (2.6).

Definition 5.3 (Companion Matrix) Let $n \in \mathbb{N}$ and P be a monic polynomial of degree n , i.e., $P(z) := z^n + \sum_{k=0}^{n-1} p_k z^k$ for some $p_k \in \mathbb{C}$, $k = 0, \dots, n-1$. Then the corresponding companion matrix $\mathbf{C}_n(\mathbf{p}) \in \mathbb{C}^{n \times n}$ is given by

$$\mathbf{C}_n(\mathbf{p}) := \begin{pmatrix} 0 & 0 & \cdots & 0 & p_0 \\ 1 & 0 & \cdots & 0 & p_1 \\ 0 & 1 & \cdots & 0 & p_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & p_{n-1} \end{pmatrix}, \quad (5.11)$$

where $\mathbf{p} := (p_0, \dots, p_{n-1})^T$ denotes the vector of the coefficients up to the degree $n-1$.

Companion matrices are used for the numerical computation of the roots of polynomials, since the eigenvalues of the companion matrix $\mathbf{C}_n(\mathbf{p})$ are the roots of the corresponding monic polynomial $P(z) = z^n + \sum_{k=0}^{n-1} p_k z^k$, see [HJ13], pages 194–195.

Lemma 5.4 Let $N, M \in \mathbb{N}$ with $M \leq N$ and $\mathbf{H}_{2N-M,M}(s)$ be the Hankel sub-matrices as in Definition 5.1 for $s = 0, 1$ and $L = M$. Furthermore, let P be the Prony polynomial given in (5.3) of degree M and $\mathbf{C}_M(\mathbf{p})$ be the corresponding companion matrix defined in (5.11). Then the following holds

$$\mathbf{H}_{2N-M,M}(0) \cdot \mathbf{C}_M(\mathbf{p}) = \mathbf{H}_{2N-M,M}(1). \quad (5.12)$$

Proof. Using the definition of the Hankel sub-matrix $\mathbf{H}_{2N-M,M}(0)$ in (5.6), Lemma 5.2 (ii) and $p_M = 1$ yields

$$\mathbf{H}_{2N-M,M}(0) \cdot \begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_{M-1} \end{pmatrix} = - (f_{\ell+M})_{\ell=0}^{2N-M-1}.$$

Therefore, we can use (5.6) and get

$$\mathbf{H}_{2N-M,M}(0) \cdot \mathbf{C}_M(\mathbf{p}) = \mathbf{H}_{2N-M,M}(1).$$

□

Remark 5.5 The above lemma implies that the eigenvalues $\lambda_j, j = 1, \dots, M$, of the companion matrix $\mathbf{C}_M(\mathbf{p})$ can be computed by solving a generalized eigenvalue problem. If λ is an eigenvalue of the companion matrix $\mathbf{C}_M(\mathbf{p})$ with corresponding eigenvector \mathbf{v} , then we have

$$\begin{aligned} \mathbf{H}_{2N-M,M}(0) \cdot \underbrace{\mathbf{C}_M(\mathbf{p}) \cdot \mathbf{v}}_{=\lambda \mathbf{v}} &= \mathbf{H}_{2N-M,M}(1) \cdot \mathbf{v} \\ \Leftrightarrow \lambda \mathbf{H}_{2N-M,M}(0) \cdot \mathbf{v} &= \mathbf{H}_{2N-M,M}(1) \cdot \mathbf{v}. \end{aligned} \quad (5.13)$$

Thus, the eigenvalues λ of the **matrix pencil** corresponding to the matrices $\mathbf{H}_{2N-M,M}(0)$ and $\mathbf{H}_{2N-M,M}(1)$, i.e., all $\lambda \in \mathbb{C}$ satisfying

$$\det(\mathbf{H}_{2N-M,M}(1) - \lambda \mathbf{H}_{2N-M,M}(0)) = 0,$$

are the eigenvalues of the companion matrix $\mathbf{C}_M(\mathbf{p})$. \diamond

Definition 5.6 (Moore-Penrose Inverse) Let $\mathbf{A} \in \mathbb{C}^{n \times m}$ be a matrix. Then a matrix $\mathbf{A}^+ \in \mathbb{C}^{m \times n}$ is called **Moore-Penrose inverse** or **generalized inverse** if it satisfies the Moore-Penrose conditions:

- (i) $\mathbf{A} \cdot \mathbf{A}^+ \cdot \mathbf{A} = \mathbf{A}$,
- (ii) $\mathbf{A}^+ \cdot \mathbf{A} \cdot \mathbf{A}^+ = \mathbf{A}^+$,
- (iii) $(\mathbf{A}^+ \cdot \mathbf{A})^* = \mathbf{A}^+ \cdot \mathbf{A}$,
- (iv) $(\mathbf{A} \cdot \mathbf{A}^+)^* = \mathbf{A} \cdot \mathbf{A}^+$.

We note some of the properties of the generalized inverse, see [SS90], Chapter III, Theorem 1.2 and Theorem 1.3.

Lemma 5.7 Let $\mathbf{A} \in \mathbb{C}^{n \times m}$, $\mathbf{A}^+ \in \mathbb{C}^{m \times n}$ be its Moore – Penrose inverse and $\mathbf{A}^* \in \mathbb{C}^{m \times n}$ be the conjugate transpose of \mathbf{A} . Then the following holds.

- (i) If the columns of \mathbf{A} are linearly independent, then

$$\mathbf{A}^+ := (\mathbf{A}^* \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^*. \quad (5.14)$$

- (ii) If the rows of \mathbf{A} are linearly independent, then

$$\mathbf{A}^+ := \mathbf{A}^* \cdot (\mathbf{A}^* \cdot \mathbf{A})^{-1}.$$

- (iii) If $\mathbf{A} = \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$, with $\mathbf{D} = \text{diag}(d_1, \dots, d_r)$ and $d_i \neq 0$ for $i = 1, \dots, r$, then the

Moore-Penrose inverse is given by

$$A^+ = \begin{pmatrix} \mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

(iv) The matrix $\mathbf{P}_A = \mathbf{A} \cdot \mathbf{A}^+ \in \mathbb{C}^{n \times n}$ is the orthogonal projector onto the column space of \mathbf{A} .

Theorem 5.8 Let $N, M \in \mathbb{N}$, f_k be the noiseless data samples given in (5.2) for $k = 0, \dots, 2N - 1$, and $\mathbf{H}_{2N-M, M+1}$ as in (5.9) with the singular value decomposition given in (5.10) using the matrices \mathbf{U}_{2N-M} , $\mathbf{D}_{2N-M, M+1}$ and \mathbf{W}_{M+1} . Furthermore, we define $\mathbf{W}_M(0)$ by removing the last column and the last row of \mathbf{W}_{M+1} and $\mathbf{W}_M(1)$ by removing the first column and last row of \mathbf{W}_{M+1} , i.e.,

$$\mathbf{W}_M(s) := \mathbf{W}_{M+1}(1 : M, 1 + s : M + s) \quad \text{for } s = 0, 1.$$

Moreover, let $\mathbf{C}_M(\mathbf{p})$ be the companion matrix defined in (5.11), where $\mathbf{p} = (p_0, p_1, \dots, p_{M-1})^T$ denotes the vector containing the (unknown) coefficients of the Prony polynomial P given in (5.3). Then the eigenvalues of $\mathbf{W}_M(0)^{-1} \cdot \mathbf{W}_M(1)$ are equal to the eigenvalues of $\mathbf{C}_M(\mathbf{p})$.

Proof. Definition (5.6) implies that the singular value decompositions of the Hankel sub-matrices are given by

$$\begin{aligned} \mathbf{H}_{2N-M, M}(0) &= \mathbf{U}_{2N-M} \cdot \mathbf{D}_{2N-M, M+1} \cdot \mathbf{W}_{M+1, M}(0), \\ \mathbf{H}_{2N-M, M}(1) &= \mathbf{U}_{2N-M} \cdot \mathbf{D}_{2N-M, M+1} \cdot \mathbf{W}_{M+1, M}(1), \end{aligned} \quad (5.15)$$

where $\mathbf{W}_{M+1, M}(0)$ is obtained by removing the last column of \mathbf{W}_{M+1} and $\mathbf{W}_{M+1, M}(1)$ by removing the first column, respectively. Hence, Lemma 5.4 implies

$$\mathbf{D}_{2N-M, M+1} \cdot \mathbf{W}_{M+1}(0) \cdot \mathbf{C}_M(\mathbf{p}) = \mathbf{D}_{2N-M, M+1} \cdot \mathbf{W}_{M+1}(1).$$

It follows from Lemma 5.7 (iii) that the generalized inverse of $\mathbf{D}_{2N-M, M+1}$ is given by

$$\mathbf{D}_{2N-M, M+1}^+ = \left(\text{diag} \left(\frac{1}{\sigma_1(\mathbf{H}_{2N-M, M+1})}, \dots, \frac{1}{\sigma_M(\mathbf{H}_{2N-M, M+1})}, 0 \right), \mathbf{0}_{2N-2M+1} \right)$$

with $\mathbf{D}_{2N-M, M+1}^+ \in \mathbb{R}^{M+1, 2N-M}$. Therefore, multiplication with $\mathbf{D}_{2N-M, M+1}^+$ yields

$$\mathbf{W}_M(0) \cdot \mathbf{C}_M(\mathbf{p}) = \mathbf{W}_M(1).$$

Since Lemma 5.2 yields that the matrices $\mathbf{H}_{2N-M, M}(0)$ and $\mathbf{H}_{2N-M, M}(1)$ have full rank M , it follows that $\mathbf{D}_{2N-M, M+1} \cdot \mathbf{W}_{M+1, M}(0)$ and $\mathbf{D}_{2N-M, M+1} \cdot \mathbf{W}_{M+1, M}(1)$ have rank M . Therefore, we obtain invertibility of the matrices $\mathbf{W}_M(1)$ and $\mathbf{W}_M(0)$. Hence, the claim follows. \square

The above theorem implies that it is sufficient to compute the eigenvalues of

$$\mathbf{W}_M(0)^{-1} \cdot \mathbf{W}_M(1) \quad (5.16)$$

in order to compute the eigenvalues $e^{\alpha_j h}$, $j = 1, \dots, M$ corresponding to Prony polynomial corresponding the generalized exponential sum in (5.1). At last we can compute the parameter c_j for $j = 1, \dots, M$ by solving

$$f_k = \sum_{j=1}^M c_j e^{\alpha_j G(x_0)} e^{\alpha_j h k} H(x_0)$$

for $k = 0, \dots, 2N - 1$.

We summarize our results in the following algorithm.

Algorithm 7 ESPRIT Algorithm for Generalized Exponential Sums

Input: $L, N \in \mathbb{N}$, $L \leq N$, L upper bound on the number of terms M in (5.1), $h > 0$, sampled values $f(G^{-1}(h\ell + G(x_0)))$, $\ell = 0, \dots, 2N - 1$, $G(x_0)$

1: Form the Hankel matrix $\mathbf{H}_{2N-L, L+1} := (a_{k+\ell} f(G^{-1}(h(\ell+k) + G(x_0))))_{k, \ell=0}^{2N-L-1, L+1}$ as in (5.5) with $a_\ell := \frac{H(x_0)}{H(G^{-1}(G(x_0)+h\ell))}$ for $\ell = 0, \dots, 2N - 1$. Determine the rank M of $\mathbf{H}_{2N-L, L+1}$.

2: Form the matrix $\mathbf{H}_{2N-M, M+1} := (a_{k+\ell} f(G^{-1}(h(\ell+k) + G(x_0))))_{k, \ell=0}^{2N-M-1, M+1}$ as in (5.9) and compute its singular value decomposition $\mathbf{H}_{2N-M, M+1} = \mathbf{U}_{2N-M} \cdot \mathbf{D}_{2N-M, M+1} \cdot \mathbf{W}_{M+1}$.

3: Build the restricted matrix $\mathbf{W}_M(0)$ by removing the last column and the last row of \mathbf{W}_{M+1} and $\mathbf{W}_M(1)$ by removing the first column and the last row of \mathbf{W}_{M+1} . Compute the eigenvalues λ_j , $j = 0, \dots, M$ of $\mathbf{W}_M(0)^{-1} \cdot \mathbf{W}_M(1)$.

4: Compute the frequencies $\alpha_j := \frac{1}{h} \log(\lambda_j)$ for $j = 1, \dots, M$.

5: Solve the Vandermonde-like system $\mathbf{V}_{2N, M} \cdot \tilde{\mathbf{c}} = \mathbf{f}_{2N}$ with $\mathbf{V}_{2N, M} := (\lambda_j^k)_{k=0, j=1}^{2N-1, M}$, $\tilde{\mathbf{c}} := (c_j \lambda_j^{G(x_0)/h})_{j=1}^M$ and $\mathbf{f}_{2N} := (f(G^{-1}(h(\ell+k) + G(x_0))))_{k=0}^{2N-1}$ and compute the coefficients $c_j = \frac{\tilde{c}_j}{e^{\alpha_j G(x_0)/h}}$ for $j = 1, \dots, M$.

Output: M , α_j and c_j for $j = 1, \dots, M$ as in (5.1).

Remark 5.9 (i) In practice, the determination of the rank of the Hankel matrix $\mathbf{H}_{2N-L, L+1}$ in step 1 is done by calculating the singular values $\sigma_1 \geq \dots \geq \sigma_{L+1}$ and defining M as the number of eigenvalues greater than a small predefined threshold bound $\varepsilon > 0$. In particular, in the case of noisy data, i.e., $\tilde{f}_k := f_k + \epsilon_k$ where ϵ_k is additive noise, the parameter ε needs to be chosen appropriately. We define the **error matrix** as $\mathbf{E}_{2N-L, L+1} := (\epsilon_{k+\ell})_{k, \ell=0}^{2N-L-1, L} \in \mathbb{C}^{2N-L, L+1}$ and obtain

the noisy Hankel matrix $\tilde{\mathbf{H}}_{2N-L,L+1} \in \mathbb{C}^{2N-L,L+1}$ given by

$$\tilde{\mathbf{H}}_{2N-L,L+1} := \mathbf{H}_{2N-L,L+1} + \mathbf{E}_{2N-L,L+1}.$$

If $2 \|\mathbf{E}_{2N-L,L+1}\|_2 \ll \sigma_M(\mathbf{H}_{2N-L,L+1})$, we can chose

$$\varepsilon \approx 2 \|\mathbf{E}_{2N-L,L+1}\|_2 / \sigma_1(\tilde{\mathbf{H}}_{2N-L,L+1}),$$

see [PT17], page 630.

- (ii) One possible way to compute the singular values of the Hankel matrices $\mathbf{H}_{2N-L,L}$ and $\mathbf{H}_{2N-M,M+1}$ with high accuracy is given in [Drm15]. The main idea behind this approach is to determine a so-called **rank revealing decomposition** (RRD) based on a modified one-sided Jacobian method prior to computing the singular value decomposition of the Hankel matrices. Similar approaches can be used for Vandermonde matrices. For further information on this topic we refer to [DGE⁺99, Dem00].
- (iii) The above algorithm can also be applied for the recovery of real generalized trigonometric functions studied in Chapter 3.2.2. In order to so do, we use the trigonometric identities

$$\begin{aligned} \cos(\alpha G(x) + \beta) &= \frac{1}{2} \left(e^{i(\alpha G(x) + \beta)} + e^{-i(\alpha G(x) + \beta)} \right), \\ \sin(\alpha G(x) + \beta) &= \frac{1}{2i} \left(e^{i(\alpha G(x) + \beta)} - e^{-i(\alpha G(x) + \beta)} \right). \end{aligned}$$

Thus, we obtain for the generalized cosine expansion

$$\begin{aligned} \sum_{j=1}^M c_j \cos(\alpha_j G(x) + \beta_j) &= \sum_{j=1}^M \frac{1}{2} c_j \left(e^{i\beta_j} e^{i\alpha_j G(x)} + e^{-i\beta_j} e^{-i\alpha_j G(x)} \right) \\ &= \sum_{j=1}^{2M} \tilde{c}_j e^{\tilde{\alpha}_j G(x)} \end{aligned}$$

with $\tilde{c}_j := \frac{1}{2} c_j e^{i\beta_j}$ and $\tilde{\alpha}_j := i\alpha_j$ for $j = 1, \dots, M$ as well as $\tilde{c}_j := \frac{1}{2} c_j e^{-i\beta_j}$ and $\tilde{\alpha}_j := -i\alpha_j$ for $j = M+1, \dots, 2M$. For real parameters c_j , it is sufficient to recover the parameter β_j and c_j from the real and imaginary part of \tilde{c}_j for $j = 1, \dots, M$. We can use analogous trigonometric identities for generalized sine expansions as well as generalized hyperbolic expansions.

- (iv) We can also combine the above ESPRIT algorithm with the algorithm for the reconstruction of sparse expansions into orthogonal polynomials based on generating functions obtained in Chapter 4. This is possible, since the algorithm derived in Chapter 4 is based on "mapping" the sparse expansion into orthogonal polynomials to the corresponding sparse monomial expansion and a special gen-

eralized shift operator, the dilation operator given in Definition 4.11, is used for the reconstruction. In this case, the samples f_k are replaced by the $\langle f, g(\cdot, t^k) \rangle_Q$ as in (4.8) for $k = 0, \dots, 2N - 1$.

◇

Based on Lemma 3.1. in [PT17], we can find a relationship between the singular values of $\mathbf{H}_{2N-L, L+1}$ and $\mathbf{H}_{2N-M, M+1}$. For this purpose, we recall some properties of eigenvalues of hermitian matrices.

Theorem 5.10 *Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ be a hermitian matrix and let $\lambda_1(\mathbf{A}) \leq \lambda_2(\mathbf{A}) \leq \dots \leq \lambda_n(\mathbf{A})$ be the eigenvalues of \mathbf{A} . Then the following holds.*

- (i) *(Cauchy's Interlacing Theorem) Let $\mathbf{B} \in \mathbb{C}^{(n-1) \times (n-1)}$ be a principal sub matrix, i.e., obtained by deleting the i -th row as well as the i -th column of \mathbf{A} for some $i \in \{1, \dots, n\}$. Moreover, let $\lambda_1(\mathbf{B}) \leq \lambda_2(\mathbf{B}) \leq \dots \leq \lambda_{n-1}(\mathbf{B})$ be the eigenvalues of \mathbf{B} . Then the following holds*

$$\lambda_1(\mathbf{A}) \leq \lambda_1(\mathbf{B}) \leq \lambda_2(\mathbf{A}) \leq \dots \leq \lambda_{n-1}(\mathbf{B}) \leq \lambda_n(\mathbf{A}).$$

- (ii) *(Weyl's Theorem) Let $\mathbf{C} \in \mathbb{C}^{n \times n}$ be a hermitian matrix and $\lambda_1(\mathbf{C}) \leq \lambda_2(\mathbf{C}) \leq \dots \leq \lambda_n(\mathbf{C})$ be the eigenvalues of \mathbf{C} . Furthermore, let $\lambda_1(\mathbf{A} + \mathbf{C}) \leq \dots \leq \lambda_n(\mathbf{A} + \mathbf{C})$ be the eigenvalues of $\mathbf{A} + \mathbf{C}$. Then we have*

$$\lambda_k(\mathbf{A}) + \lambda_1(\mathbf{C}) \leq \lambda_k(\mathbf{A} + \mathbf{C}) \leq \lambda_k(\mathbf{A}) + \lambda_n(\mathbf{C})$$

for $k = 1, \dots, n$.

Proof. For the proof of (i) we refer to [SS90], Chapter IV, Theorem 4.2. For the proof of (ii) we refer to [SS90], Chapter IV, Corollary 4.9. □

Lemma 5.11 *Let $L, M, N \in \mathbb{N}$ with $M \leq L \leq N$, f_k be the data samples given in (5.2) and $\mathbf{H}_{2N-M, M+1}$ be the rectangular Hankel matrix given in (5.5). Then the singular values of the $\mathbf{H}_{2N-M, M+1}$ satisfy the inequality*

$$\sigma_k(\mathbf{H}_{2N-L, L+1})^2 \leq \sigma_k(\mathbf{H}_{2N-M, M+1})^2 + \sum_{j=M+1}^L \|\mathbf{f}_j\|_2^2 \quad k = 1, \dots, M, \quad (5.17)$$

where $\mathbf{f}_j = (f_k)_{k=j}^{2N-1}$ denotes the last column of the matrix $\mathbf{H}_{2N-j, j+1}$ for $j = M + 1, \dots, L$.

Proof. This proof is based on the proof Lemma 3.1 in [PT17]. Therefore, we will recapitulate the first steps of its proof. Analogously to (5.6), we can represent the matrices

$\mathbf{H}_{2N-L+1,L}$ and $\mathbf{H}_{2N-L,L+1}$ using a block matrix structure

$$\begin{aligned}\mathbf{H}_{2N-L,L+1} &= \left(\mathbf{H}_{2N-L,L} \ (f_{\ell+L})_{\ell=0}^{2N-L-1} \right), \\ \mathbf{H}_{2N-L+1,L} &= \begin{pmatrix} \mathbf{H}_{2N-L,L} \\ \left((f_{\ell+2N-L})_{\ell=0}^{L-1} \right)^T \end{pmatrix}.\end{aligned}$$

Note that we write $\mathbf{H}_{2N-L,L}$ instead of $\mathbf{H}_{2N-L,L}(0)$ in this proof in order to simplify the notation and improve the readability of the proof. Furthermore, we set

$$\begin{aligned}\mathbf{f}_L &:= (f_{\ell+L})_{\ell=0}^{2N-L-1}, \\ \mathbf{h}_{2N-L} &:= (f_{\ell+2N-L})_{\ell=0}^{L-1}.\end{aligned}$$

Using this notation we obtain

$$\mathbf{A}_{2N-L} := \mathbf{H}_{2N-L,L+1} \cdot \mathbf{H}_{2N-L,L+1}^* = \mathbf{H}_{2N-L,L} \cdot \mathbf{H}_{2N-L,L}^* + \mathbf{f}_L \mathbf{f}_L^*, \quad (5.18)$$

$$\begin{aligned}\mathbf{A}_{2N-L+1} &:= \mathbf{H}_{2N-L+1,L} \cdot \mathbf{H}_{2N-L+1,L}^* \\ &= \begin{pmatrix} \mathbf{H}_{2N-L,L} \cdot \mathbf{H}_{2N-L,L}^* & \mathbf{H}_{2N-L,L} \bar{\mathbf{h}}_{2N-L} \\ \mathbf{h}_{2N-L}^T \bar{\mathbf{H}}_{2N-L,L}^* & \|\mathbf{h}_{2N-L}\|_2^2 \end{pmatrix}\end{aligned} \quad (5.19)$$

Equation (5.18) implies that \mathbf{A}_{2N-L} is a rank-one perturbation of $\mathbf{H}_{2N-L,L} \mathbf{H}_{2N-L,L}^*$. Moreover, equation (5.19) implies that the matrix $\mathbf{H}_{2N-L,L} \mathbf{H}_{2N-L,L}^*$ is a principal sub-matrix of \mathbf{A}_{2N-L+1} . Therefore, we can apply Cauchy interlacing Theorem (Theorem 5.10 (i)) and obtain

$$\lambda_k(\mathbf{H}_{2N-L,L} \cdot \mathbf{H}_{2N-L,L}^*) \leq \lambda_{k+1}(\mathbf{A}_{2N-L+1}) \quad (5.20)$$

for $k = 1, \dots, L$. Theorem 5.10 (ii) now yields

$$\lambda_1(\mathbf{f}_L \mathbf{f}_L^*) + \lambda_k(\mathbf{H}_{2N-L,L} \cdot \mathbf{H}_{2N-L,L}^*) \leq \lambda_k(\mathbf{A}_{2N-L})$$

for $k = 1, \dots, L$. We combine this with (5.20) and obtain

$$\lambda_1(-\mathbf{f}_L \mathbf{f}_L^*) + \lambda_k(\mathbf{A}_{2N-L}) \leq \lambda_k(\mathbf{H}_{2N-L,L} \cdot \mathbf{H}_{2N-L,L}^*) \leq \lambda_{k+1}(\mathbf{A})_{2N-L+1}$$

for $k = 1, \dots, L$. Using the fact that $-\mathbf{f}_L \mathbf{f}_L^*$ is a rank-one matrix, we can conclude that its first eigenvalue $\lambda_1(-\mathbf{f}_L \mathbf{f}_L^*)$ is equal to the norm $-\|\mathbf{f}_L\|_2^2$. Hence, we can obtain that the singular values of the matrix $\mathbf{H}_{2N-L,L+1}$ increase almost monotonously with respect to L , i.e., they satisfy the following inequality

$$\sigma_k(\mathbf{H}_{2N-L,L+1})^2 \leq \sigma_k(\mathbf{H}_{2N-L+1,L})^2 + \|\mathbf{f}_L\|_2^2 \quad k = 1, \dots, M,$$

where \mathbf{f}_L denotes the last column of the matrix $\mathbf{H}_{2N-L,L+1}$. Iterative application yields

the claim. □

Remark 5.12 The above lemma implies that the choice of the upper bound L has an effect on the lowest singular value $\sigma_M(\mathbf{H}_{2N-L,L+1}) > 0$ of $\mathbf{H}_{2N-L,L+1}$. Since the singular values are used to estimate the parameter M , the singular value σ_M should not be too small. Furthermore, we can conclude that a good choice of the upper bound L is close to the original parameter M , i.e., $L \approx M$ and $L \approx N$. That is if the matrices $\mathbf{H}_{2N-L,L+1}$ as well as $\mathbf{H}_{2N-M,M+1}$ are almost square. ◇

The difference between our modified version of the ESPRIT Algorithm 7 and the classical ESPRIT algorithm described in [PT13] is the second step. In the classical ESPRIT method, the construction of the matrix $\mathbf{H}_{2N-M,M+1}$ is omitted and instead the matrix $\mathbf{H}_{2N-L,L+1}$ and its singular value decomposition given as in (5.8) are used. Instead of almost square or square matrices only sub-matrices of the orthogonal matrices in the singular value decomposition are employed.

Using the established rank M of the Hankel matrix $\mathbf{H}_{2N-L,L+1}$, we can construct the matrices $\mathbf{U}_{2N-L,M} := \mathbf{U}_{2N-L}(1 : 2N - L, 1 : M)$ and $\mathbf{W}_{M,L+1} := \mathbf{W}_{L+1}(1 : M, 1 : L + 1)$ with orthonormal columns, as well as the diagonal matrix $\mathbf{D}_M := \text{diag}(\sigma_1, \dots, \sigma_M)$ and obtain a partial singular value decomposition of the matrix (5.5). Additionally, we set $\mathbf{W}_{M,L}(s) := \mathbf{W}_{M,L+1}(1 : M, 1 + s : L + s)$ for $s = 0, 1$. This implies that both Hankel sub-matrices in (5.6) can be simultaneously factorized, i.e.,

$$\mathbf{H}_{2N-L,L}(s) = \mathbf{U}_{2N-L,M} \cdot \mathbf{D}_M \cdot \mathbf{W}_{M,L}(s) \quad \text{for } s = 0, 1.$$

Hence, rather than the square matrices $\mathbf{W}_M(0)$ and $\mathbf{W}_M(1)$ the rectangular matrices $\mathbf{W}_{M,L}(0)$ and $\mathbf{W}_{M,L}(1)$ are considered. Therefore, instead of (5.16) we need to compute the eigenvalues of $(\mathbf{W}_{M,L}(0))^+ \cdot \mathbf{W}_{M,L}(1)$.

In recent years, different performance results for the ESPRIT method have been established. Most of the existing performance results are of statistical nature and focus on asymptotic - either in the sample size or the SNR - statements, see, e.g., [SS91, RH89, LVT91]. A deterministic result in the special case of exponential sums on the unit circle was given in [Aub16], Chapter 5, and is based on the application of the Bauer-Fike-Theorem, see [SS90], Chapter IV, Theorem 3.3, as well as an upper bound on the condition of Vandermonde matrices with nodes on the unit circle. In particular, these results imply that performance of the ESPRIT method is good, if the noise present in the used data samples is small. Since these results only depend on the frequency parameters e^{α_j} , $j = 1, \dots, M$, we can assume that similar results can be obtained for generalized exponential sums if $|e^{\alpha_j}| \leq 1$. Further results for generalized exponential sums with frequencies outside the unit circle as well as generalized trigonometric expansions have yet to be obtained.

5.2 Partially Known Frequency Parameters

In some applications, one or more of the frequency parameters α_j for some $j = 1, \dots, M$, in the model f given in (5.1) are known beforehand, while the corresponding weight coefficients c_j are unknown. Therefore, we cannot simply eliminate the term $c_j H(x) e^{\alpha_j G(x)}$ from the sum (5.1) in order to get a new set of measurements and use Prony's method. This section is dedicated to the recovery of generalized exponential sums in the case of partially known frequencies.

Theorem 5.13 *Let $M, N \in \mathbb{N}$, $M \leq N$ and f be a generalized exponential sum of the form (5.1). Furthermore, let f_k be the data samples given in (5.2) for $k = 0, \dots, 2N - 1$ and one frequency α_ℓ for $\ell = 1, \dots, M$, be known beforehand. Then the function f can be recovered using the samples $\tilde{f}_k = f_{k+1} - e^{\alpha_1 h} f_k$ for $k = 0, \dots, 2N - 2$.*

Proof. Without loss of generality we can assume that the frequency α_1 is known beforehand and we define $\lambda_1 := e^{\alpha_1 h}$. We recall the definition of the Prony polynomial P given in (5.3), i.e.,

$$P(z) = \sum_{k=0}^M p_k z^k = \prod_{j=1}^M (z - e^{\alpha_j h}) = (z - \lambda_1) \underbrace{\prod_{j=2}^M (z - e^{\alpha_j h})}_{=: Q(z)}, \quad (5.21)$$

where $\mathbf{p} := (p_0, \dots, p_M)^T$ the coefficient vector of the Prony polynomial. Furthermore, we have

$$Q(z) := \prod_{j=2}^M (z - \lambda_j) = \sum_{k=0}^{M-1} q_k z^k$$

with the coefficient vector $\mathbf{q} := (q_0, \dots, q_{M-1})^T$ and $\lambda_j := e^{\alpha_j h}$ for $j = 2, \dots, M$. Then (5.21) yields

$$\mathbf{p} = \begin{pmatrix} 0 \\ q_0 \\ \vdots \\ q_{M-1} \end{pmatrix} - \lambda_1 \begin{pmatrix} q_0 \\ \vdots \\ q_{M-1} \\ 0 \end{pmatrix}.$$

Thus, it follows from Lemma 5.2 (ii)

$$\mathbf{0}_{2N-M} = \mathbf{H}_{2N-M, M+1} \cdot \mathbf{p} = \mathbf{H}_{2N-M, M+1} \cdot \left(\begin{pmatrix} 0 \\ q_0 \\ \vdots \\ q_{M-1} \end{pmatrix} - \lambda_1 \begin{pmatrix} q_0 \\ \vdots \\ q_{M-1} \\ 0 \end{pmatrix} \right)$$

and, therefore, (5.6) yields

$$\mathbf{0}_{2N-M} = \mathbf{H}_{2N-M,M+1} \cdot \mathbf{p} = (\mathbf{H}_{2N-M,M}(1) - \lambda_1 \mathbf{H}_{2N-M,M}(0)) \mathbf{q},$$

where $\mathbf{H}_{2N-M,M}(s)$ denotes the Hankel sub-matrix given in (5.6) for $s = 0, 1$. Hence, we define a new Hankel matrix for the reduced problem as

$$\tilde{\mathbf{H}}_{2N-M,M} := \mathbf{H}_{2N-M,M}(1) - \lambda_1 \mathbf{H}_{2N-M,M}(0).$$

We observe that the components \tilde{f}_k for $k = 0, \dots, 2N - 2$ of this new Hankel matrix are of the form

$$\begin{aligned} \tilde{f}_k &= f_{k+1} - \lambda_1 f_k = \sum_{j=1}^M H(x_0) c_j e^{\alpha_j(G(x_0)+h(k+1))} - e^{\alpha_1 h} H(x_0) \sum_{j=1}^M c_j e^{\alpha_j(G(x_0)+hk)} \\ &= \sum_{j=2}^M c_j H(x_0) \left(e^{\alpha_j h} - e^{\alpha_1 h} \right) e^{\alpha_j(hk+G(x_0))}, \end{aligned}$$

i.e., the coefficients c_j for $j = 2, \dots, M$ are changed to $\tilde{c}_j = c_j(e^{\alpha_j h} - e^{\alpha_1 h})$. Consequently, we can use the samples \tilde{f}_k for $k = 0, \dots, 2N - 2$ to recover the shorter sum

$$\sum_{j=2}^M \tilde{c}_j H(x) e^{\alpha_j G(x)}.$$

Once we have computed the remaining α_j , $j = 2, \dots, M$, we obtain the coefficients c_j , $j = 1, \dots, M$, by solving the linear system

$$f(G^{-1}(G(x_0) + hk)) = \sum_{j=1}^M c_j e^{\alpha_j(G(x_0)+hk)} H(G^{-1}(G(x_0) + hk)).$$

□

5.3 Sub-Sampling Based Algorithm for the Recovery of Generalized Exponential Sums

As we have previously mentioned, the accuracy of the Prony method does not only depend on the accuracy of the input data, but also on the distribution of the frequencies, i.e., on the **minimal separation distance**

$$q := \min_{\substack{i,j=1,\dots,M \\ i \neq j}} |\alpha_j - \alpha_i|.$$

If the minimal separation distance decreases, the nodes $e^{\alpha_j h}$ of the Vandermonde matrix \mathbf{V}_M grow closer together and hence the conditions of the Vandermonde matrix

as well as the Hankel matrix \mathbf{H}_M grow. In [CL20a], the authors used sub-Nyquist sampling in order to improve the frequencies with a high resolution. In the following section, we modify this approach for generalized exponential sums.

We recall the generalized exponential sum as in (5.1),

$$f(x) = \sum_{j=1}^M c_j H(x) e^{\alpha_j G(x)},$$

with $c_j \in \mathbb{C} \setminus \{0\}$, $\alpha_j \in \mathbb{C}$, $|\operatorname{Im}(\alpha_j)| < T$, $T > 0$ and G and H satisfying the assumptions in Definition 3.15.

Up until this point, we have always chosen the sampling distance $h \in \mathbb{R} \setminus \{0\}$ such that $0 < |h| < \frac{\pi}{T}$ holds. This ensures that the computation of the frequencies α_j from the values $e^{\alpha_j h}$ yields a unique solution. In this section, we will deliberately sub-sample in order to increase the resolution and reconstruct the clustered frequencies more accurately.

Instead of the consecutive samples $f_k = S_{G,H,h}^k f(x_0)$, $k = 0, \dots, 2M - 1$, only every u -th sample with $u \in \mathbb{N}$, $u > 1$ is used, i.e.,

$$f_{uk} = S_{G,H,h}^{\ell u} f(x_0) = S_{G,H,h}^k f(x_0) \quad (5.22)$$

for $k = 0, \dots, 2M - 1$. Since sub-sampling may cause the aliasing, i.e., the collision of frequencies, we need to distinguish between the collision free and the colliding case.

5.3.1 Collision Free Sub-Sampling

In the following we will assume $e^{\alpha_j u h} \neq e^{\alpha_k u h}$ for $j \neq k$ and $j, k = 1, \dots, M$. Since we are purposely sub-sampling, we cannot uniquely recover the parameter $e^{\alpha h}$, and, thus, the frequencies α_j , respectively, from the values $e^{\alpha_j u h}$, $j = 1, \dots, M$, but obtain a set of possible solutions

$$U_j := \left\{ e^{\alpha_j h + \frac{2\pi i}{u} \ell} \mid \ell = 0, \dots, u - 1 \right\}. \quad (5.23)$$

In order to determine the correct frequency α_j , $j = 1, \dots, M$, we use the following lemma.

Lemma 5.14 *Let $p, u \in \mathbb{N}$ with $\gcd(u, p) = 1$. Furthermore, let $\alpha \in \mathbb{C}$ with $|\operatorname{Im}(\alpha)| < K$ for some $K > 0$ for $j = 1, \dots, M$ and $h = \frac{\pi}{K}$. Moreover, we define*

$$P := \left\{ e^{\alpha h + \frac{2\pi i}{p} \ell} \mid \ell = 0, \dots, p - 1 \right\} \quad \text{and}$$

$$U := \left\{ e^{\alpha h + \frac{2\pi i}{u} \ell} \mid \ell = 0, \dots, u - 1 \right\}.$$

Then the following holds.

(i) The sets U and P have a unique intersection $e^{\alpha h}$, i.e., $P \cap U = \{e^{\alpha h}\}$.

(ii) There exist $r_1, r_2 \in \mathbb{Z}$ such that $r_1 u + r_2 p = 1$, and we have

$$\left(e^{\alpha u h}\right)^{r_1} \left(e^{\alpha p h}\right)^{r_2} = e^{\alpha h}.$$

Proof. For the proof of (i) we refer to [CL20a], Lemma 1. Furthermore, Theorem 15, Chapter 2.4 in [Bos06], yields the existence of $r_1, r_2 \in \mathbb{Z}$ satisfying $r_1 u + r_2 p = \gcd(u, p) = 1$. Hence, we obtain

$$\left(e^{\alpha u h}\right)^{r_1} \left(e^{\alpha p h}\right)^{r_2} = e^{\alpha h}.$$

Consequently, the claim follows. \square

Remark 5.15 The integers $r_1, r_2 \in \mathbb{Z}$ given in the above lemma can be computed by the Euclidean algorithm, see [Bos06], Chapter 2.4, Theorem 15 and are called **Bézout coefficients**. \diamond

Theorem 5.16 Let $G: \mathbb{R} \rightarrow \mathbb{R}$ and $H: \mathbb{R} \rightarrow \mathbb{C}$ be continuous functions such that G is strictly monotone and H is non-vanishing on an interval $[a, b] \subset \mathbb{R}$. Furthermore, let f be a generalized exponential expansion of length M as in (5.1) such that there exists a $K \in \mathbb{R}$ with $|\operatorname{Im}(\alpha_j)| < K$ for all $j = 1, \dots, M$. Moreover, let $h = \frac{\pi}{K}$, $u, p \in \mathbb{N}$ satisfy $\gcd(u, p) = 1$ and $x_0 \in [a, b]$ such that

$$\begin{aligned} G(x_0) + huk &\in [G(a), G(b)] && \text{for } G(a) < G(b) && \text{or} \\ G(x_0) + huk &\in [G(b), G(a)] && \text{for } G(a) > G(b) \end{aligned}$$

for all $k = 0, \dots, 2M - 1$ as well as

$$\begin{aligned} G(x_0) + h(uk + p) &\in [G(a), G(b)] && \text{for } G(a) < G(b) && \text{or} \\ G(x_0) + h(uk + p) &\in [G(b), G(a)] && \text{for } G(a) > G(b) \end{aligned}$$

for all $k = 0, \dots, M - 1$ holds. If $e^{\alpha_j u h} \neq e^{\alpha_k u h}$ for $j \neq k$ holds, then f can be uniquely reconstructed using the samples $f(G^{-1}(G(x_0) + uhk))$ for $k = 0, \dots, 2M - 1$ and the samples $f(G^{-1}(G(x_0) + h(uk + p)))$ for $k = 0, \dots, M - 1$.

Proof. The proof is based on the Prony method for generalized exponential sums. We

define

$$f_{u\ell} := S_{G,H,hu}^\ell f(x_0) \frac{H(x_0)}{H(G^{-1}(G(x_0) + u\ell h))} f(G^{-1}(G(x_0) + u\ell h)) \quad \text{and} \quad (5.24)$$

$$f_{uk+p} := S_{G,H,hu}^k S_{G,H,hp} f(x_0) \frac{H(x_0)}{H(G^{-1}(G(x_0) + h(uk+p)))} f(G^{-1}(G(x_0) + h(uk+p))) \quad (5.25)$$

for $\ell = 0, \dots, 2M-1$ and $k = 0, \dots, M-1$. We note that the values

$$\frac{H(x_0)}{H(G^{-1}(G(x_0) + u\ell h))} \quad \text{and} \quad \frac{H(x_0)}{H(G^{-1}(G(x_0) + h(uk+p)))}$$

for $\ell = 0, \dots, 2M-1$, $k = 0, \dots, M-1$, are independent of the signal f and can be precomputed.

We follow the first steps of the Prony method for generalized exponential sums and define the Prony polynomial

$$P(z) = \prod_{j=1}^M (z - e^{\alpha_j u h}) = \sum_{k=0}^M p_k z^k. \quad (5.26)$$

Furthermore, we define the Hankel matrix $\mathbf{H}_M := \left(f_{u(\ell+m)} \right)_{\ell,m=0}^{M-1}$. Lemma 5.2 (ii) as well as $e^{\alpha_j u h} \neq e^{\alpha_k u h}$, $j \neq k$, and $p_M = 1$ yield

$$\mathbf{H}_M \cdot \mathbf{p} = -\mathbf{f}$$

with $\mathbf{p} := (p_k)_{k=0}^{M-1}$ and $\mathbf{f} := \left(f_{u(\ell+M)} \right)_{\ell=0}^{M-1}$. Once we have computed the coefficients p_k , $k = 0, \dots, M$, we compute the roots $e^{\alpha_j u h}$ of the Prony polynomial P . Moreover, we solve

$$f_{u\ell} = \sum_{j=1}^M \tilde{c}_j e^{\alpha_j u h} \quad \text{for } \ell = 0, \dots, 2M-1 \quad (5.27)$$

with

$$\tilde{c}_j := c_j H(x_0) e^{\alpha_j G(x_0)} \quad (5.28)$$

for $j = 1, \dots, M$. In order to uniquely recover the frequencies α_j for $j = 1, \dots, M$ from the set U_j as in (5.23), we consider the M additional samples f_{uk+p} as in (5.25) for $k = 0, \dots, M-1$. Employing Theorem 3.28 yields

$$f_{uk+p} = \sum_{j=1}^M c_j H(x_0) e^{G(x_0)\alpha_j} e^{\alpha_j p h} e^{\alpha_j u k h} = \sum_{j=1}^M \tilde{c}_j e^{\alpha_j p h} \left(e^{\alpha_j u h} \right)^k$$

with \tilde{c}_j as in (5.28) for $j = 1, \dots, M$. Hence, we obtain the linear system

$$\underbrace{\begin{pmatrix} 1 & \dots & 1 \\ e^{\alpha_1 u h} & \dots & e^{\alpha_M u h} \\ \vdots & \dots & \vdots \\ e^{\alpha_1 (M-1) u h} & \dots & e^{\alpha_M (M-1) u h} \end{pmatrix}}_{=: \mathbf{V}_M} \cdot \underbrace{\begin{pmatrix} \tilde{c}_1 & 0 & \dots & 0 \\ 0 & \tilde{c}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \tilde{c}_M \end{pmatrix}}_{=: \mathbf{D}_M} \cdot \begin{pmatrix} e^{\alpha_1 p h} \\ \vdots \\ e^{\alpha_M p h} \end{pmatrix} = \begin{pmatrix} f_{p+0u} \\ \vdots \\ f_{p+(M-1)u} \end{pmatrix}.$$

Since the diagonal entries of the matrix \mathbf{D}_M are nonzero, it is invertible. Moreover, the matrix \mathbf{V}_M is a Vandermonde matrix with distinct nodes and is, therefore, invertible. Consequently, this system can be uniquely solved for the values $e^{\alpha_j p h}$ for $j = 1, \dots, M$. Therefore, we can employ Lemma 5.14 and uniquely recover the values $e^{\alpha_j h}$, and, thus, the frequencies α_j , $j = 1, \dots, M$.

As a last step we use (5.28) in order to compute the parameters c_j , $j = 1, \dots, M$, which proves the claim. \square

Remark 5.17 The assumption of non colliding frequencies in the above theorem, i.e., that $e^{\alpha_j u h} \neq e^{\alpha_k u h}$ for $j \neq k$ is satisfied if one of the following conditions holds:

- (i) $\operatorname{Re}(\alpha_j) \neq \operatorname{Re}(\alpha_k)$ for $j \neq k$ or
- (ii) $u \frac{\operatorname{Im}(\alpha_j)}{2\pi} \neq u \frac{\operatorname{Im}(\alpha_k)}{2\pi} + \frac{\ell}{h}$ for $j \neq k$ and for all $\ell \in \mathbb{Z}$.

\diamond

We summarize our results in the following algorithm.

Algorithm 8 Sub-Sampling Algorithm for Generalized Exponential Sums

Input: $M \in \mathbb{N}$, $u, p \in \mathbb{N}$ with $\gcd(u, p) = 1$, $h > 0$, sampled values $f(G^{-1}(h(u\ell + kp) + G(x_0)))$, $\ell = 0, \dots, 2M - 1$, $k = 0, 1$, $G(x_0)$, $H(x_0)$

- 1: Form the Hankel matrix $\mathbf{H}_M := (a_{k+\ell} f(G^{-1}(h(u\ell + k) + G(x_0))))_{k, \ell=0}^{M-1}$ with $a_\ell := \frac{H(x_0)}{H(G^{-1}(G(x_0) + uh\ell))}$ for $\ell = 0, \dots, 2M - 1$ as well as $\mathbf{f} := (a_{\ell+M} f(G^{-1}(h(u(\ell + M) + G(x_0))))_{\ell=0}^{M-1}$ and compute $\mathbf{H}_M \cdot \mathbf{p} = -\mathbf{f}$ for $\mathbf{p} := (p_\ell)_{\ell=0}^{M-1}$.
- 2: Define the Prony polynomial $P(z) := \sum_{k=0}^M p_k z^k$ with $p_M = 1$ as in (5.26) and find all of its roots $\lambda_j = e^{\alpha_j u h}$, $j = 1, \dots, M$.
- 3: Form the Vandermonde-like system $\mathbf{V}_{2M, M} := (e^{\alpha_j u k h})_{k=0, j=1}^{2M-1, M}$ and solve the linear system $\mathbf{V}_{2M, M} \cdot \tilde{\mathbf{c}} = \mathbf{f}_{2M}$ with $\tilde{\mathbf{c}} := (\tilde{c}_j H(x_0) e^{\alpha_j G(x_0)})$ and $\mathbf{f}_{2M} := (a_{k+\ell} f(G^{-1}(G(x_0) + hu(\ell + k))))_{k=0}^{2M-1}$.
- 4: Form the Vandermonde matrix $\mathbf{V}_M := (e^{\alpha_j k u h})_{k=0, j=1}^{M-1, M}$ and the diagonal matrix $\mathbf{D}_M := \text{diag}(\tilde{c}_1, \dots, \tilde{c}_M)$ and solve the system $\mathbf{V}_M \cdot \mathbf{D}_M \cdot \mathbf{h} = \tilde{\mathbf{f}}$ for \mathbf{h} with $\mathbf{h} := (e^{\alpha_j p h})_{j=1}^M$ and $\tilde{\mathbf{f}} := (b_\ell f(G^{-1}(h(u\ell + p) + G(x_0))))_{\ell=0}^{M-1}$ with $b_\ell = \frac{H(x_0)}{H(G^{-1}(G(x_0) + h(u\ell + p)))}$.
- 5: Compute r_1, r_2 such that $r_1 u + r_2 p = 1$ and $(e^{\alpha_j u h})^{r_1} (e^{\alpha_j p h})^{r_2} = e^{\alpha_j h}$ for $j = 1, \dots, M$.
- 6: Compute the frequencies $\alpha_j := \frac{1}{h} \log(e^{\alpha_j h})$ and $c_j := \frac{\tilde{c}_j}{H(x_0) e^{\alpha_j G(x_0)}}$ for $j = 1, \dots, M$.

Output: α_j and c_j for $j = 1, \dots, M$ as in (5.1).

Remark 5.18 (i) The above algorithm can also be applied in the case of unknown sparsity number M . Similarly to the ESPRIT algorithm, we use an upper bound $L \geq M$ and $2N$ samples with $N \geq L \geq M$. Moreover, we also can omit the first two steps of the algorithm and compute the values $e^{\alpha_j u h}$, $j = 1, \dots, M$ by solving the generalized eigenvalue problem (5.13) as in Remark 5.5.

(ii) Analogously to Remark 5.9 (iii) we can recover the generalized trigonometric expansion discussed in Chapter 3.2.2 using Algorithm 8 and $4M$ samples instead of $2M$.

◇

The above describe algorithm can also be combined with the reconstruction method obtained in Chapter 4.

Corollary 5.19 Let $M \in \mathbb{N}$ and $\{Q_k \mid k \in \mathbb{N}_0\}$ with $Q_k \in L^2([a, b], \omega)$ be a set of orthogonal polynomials with corresponding inner product $\langle \cdot, \cdot \rangle_Q$ as in Definition 4.1 and generating function g in (4.4). Furthermore, let f be an M -sparse expansion into orthogonal polynomials

as in (4.5) i.e.,

$$f(x) = \sum_{j=1}^M c_{n_j} Q_{n_j}(x)$$

with $c_j \in \mathbb{C} \setminus \{0\}$ for $j = 1, \dots, M$. We suppose that there exists a $t \in \mathbb{C}$ such that g is uniformly convergent for all $x \in [a, b]$ and t^k for $k \in \mathbb{N}$. Furthermore, let $p, u \in \mathbb{N}$ satisfy $\gcd(u, p) = 1$. Then f can be uniquely reconstructed using the values $\langle f, g(\cdot, t^{uk}) \rangle_Q$ for $k = 1, \dots, 2M$ and the values $\langle f, g(\cdot, t^{u\ell+p}) \rangle_Q$ for $\ell = 1, \dots, M$.

Proof. We recall that (4.10) the inner product $\langle f, g(\cdot, t) \rangle_Q$ yields that a mapping between the sparse expansion into orthogonal polynomials and the corresponding sparse polynomials in monomial basis. This polynomial can be interpreted as a generalized exponential sum with $G(x) := \log(x)$ and $H \equiv 1$. In particular, the samples $\langle f, g(\cdot, t^{uk+\ell p}) \rangle_Q$ for $k = 1, \dots, 2M - 1$ and $\ell = 0, 1$ correspond to the samples needed for the recovery of the sparse polynomial using the shift operator $S_{G, H, t}$ and $x_0 = 1$. Since we have $n_j \in \mathbb{N}_0$ for $j = 1, \dots, M$, we obtain $e^{n_j t u} \neq e^{n_k t u}$ for $j \neq k$. Therefore, we can apply Theorem 5.16 and the claim follows. \square

5.3.2 Colliding Frequencies

One problem that may occur when the sub-sampling causes aliasing, i.e., that $e^{\alpha_j u h} = e^{\alpha_k u h}$ for $j \neq k$. Without loss of generality we can assume that the collisions appear in successive terms, i.e., there exists a partition of M such that

$$M = \sum_{i=1}^{\tilde{M}} (n_{i+1} - n_i) \quad (5.29)$$

with $n_1 = 1$, $n_i \leq n_{i+1}$, $i = 1, \dots, \tilde{M}$, $n_{\tilde{M}+1} = \tilde{M} + 1$ and $e^{\alpha_j u h} = e^{\alpha_k u h}$ with $n_i \leq j, k \leq n_{i+1} - 1$ and $j \neq k$. Consequently, we obtain for the samples in (5.22)

$$f_{ul} = \sum_{j=1}^{\tilde{M}} \tilde{c}_j^{(0)} e^{\alpha_j^{(0)} u \ell h} \quad (5.30)$$

with

$$\tilde{c}_j^{(0)} := \sum_{k=n_j}^{n_{j+1}-1} \tilde{c}_k = \sum_{k=n_j}^{n_{j+1}-1} c_k H(x_0) e^{\alpha_k G(x_0)} \quad (5.31)$$

and $e^{\alpha_j^{(0)} u h} := e^{\alpha_{n_j} u h} = \dots = e^{\alpha_{n_{j+1}-1} u h}$ for $j = 1, \dots, \tilde{M}$. However, we note that, depending on the parameters c_j , $j = 1, \dots, M$, x_0 and the function G it may happen that $\tilde{c}_k^{(0)} = 0$ for some $k = 1, \dots, \tilde{M}$. First, we consider the case $\tilde{c}_k^{(0)} \neq 0$ for all $k = 1, \dots, \tilde{M}$. Since we do not know which frequencies collide, we also need to compute \tilde{M} .

Theorem 5.20 Let $G, : \mathbb{R} \rightarrow \mathbb{R}$ and $H: \mathbb{R} \rightarrow \mathbb{C}$ be continuous functions such that G is strictly monotone and H is non-vanishing on an interval $[a, b] \subset \mathbb{R}$. Furthermore, let f be a generalized exponential expansion of length M as in (5.1), such that there exists a $K \in \mathbb{R}$ with $|\operatorname{Im}(\alpha_j)| < K$ for all $j = 1, \dots, M$. Moreover, let $h = \frac{\pi}{K}$, $u, p, \in \mathbb{N}$ satisfying $\gcd(u, p) = 1$ and $x_0 \in [a, b]$ such that

$$\begin{aligned} G(x_0) + h(uk + \ell p) &\in [G(a), G(b)] && \text{for } G(a) < G(b) && \text{or} \\ G(x_0) + h(uk + \ell p) &\in [G(b), G(a)] && \text{for } G(a) > G(b) \end{aligned}$$

for all $k, \ell = 0, \dots, 2M - 1$. We suppose $e^{\alpha_k uh} = e^{\alpha_j uh}$ for some $j \neq k$ holds and that $\tilde{c}_j^{(0)}$ as in (5.31) is non-vanishing for all $j = 1, \dots, \tilde{M}$. Then f can be uniquely reconstructed using the samples $f(G^{-1}(G(x_0) + h(uk + \ell p)))$ for $k, \ell = 0, \dots, 2M - 1$.

Proof. The proof is done similarly to the proof of Theorem 5.16. For simplicity we write

$$f_{uk+\ell p} := S_{G,H,h(uk+p\ell)} f(x_0) = \frac{H(x_0)}{H(G^{-1}(G(x_0)+h(uk+p\ell)))} f(G^{-1}(G(x_0) + h(uk + p\ell)))$$

for $k, \ell = 0, \dots, 2M - 1$.

We note that Lemma 5.2 implies that we can compute the parameter \tilde{M} as the rank of the Hankel matrix $(f_{u(k+m)})_{k,m=0}^{M-1}$. Furthermore, we can use the samples f_{uk+0p} , $k = 1, \dots, \tilde{M}$, and follow the first steps in the proof of Theorem 5.16 in order to recover the values $e^{\alpha_j^{(0)} uh}$ for $j = 1, \dots, \tilde{M}$. Additionally, equation (5.30) yields

$$\underbrace{\begin{pmatrix} 1 & \dots & 1 \\ e^{\alpha_1^{(0)} uh} & \dots & e^{\alpha_M^{(0)} uh} \\ \vdots & \dots & \vdots \\ e^{\alpha_1^{(0)} (\tilde{M}-1)uh} & \dots & e^{\alpha_M^{(0)} (\tilde{M}-1)uh} \end{pmatrix}}_{=: \mathbf{V}_{\tilde{M}}} \cdot \begin{pmatrix} \tilde{c}_1^{(0)} \\ \vdots \\ \tilde{c}_{\tilde{M}}^{(0)} \end{pmatrix} = \begin{pmatrix} f_0 \\ \vdots \\ f_{u(\tilde{M}-1)} \end{pmatrix}. \quad (5.32)$$

Theorem 3.28 and equation (5.30) imply for $k \geq 1$

$$\begin{aligned} f_{uk+p\ell} &= S_{G,H,h(uk+p\ell)} f(x_0) = S_{G,H,ph}^\ell S_{G,H,uh}^k f(x_0) \\ &= S_{G,H,ph}^\ell \left(\sum_{j=1}^{\tilde{M}} \sum_{i=n_j}^{n_{j+1}-1} c_i H(\cdot) e^{\alpha_i G(\cdot)} e^{\alpha_j^{(0)} kuh} \right) (x_0) \\ &= \sum_{j=1}^{\tilde{M}} \sum_{i=n_j}^{n_{j+1}-1} c_i e^{\alpha_j^{(0)} ukh} S_{G,H,ph}^\ell \left(H(\cdot) e^{\alpha_i G(\cdot)} \right) (x_0) \\ &= \sum_{j=1}^{\tilde{M}} \sum_{i=n_j}^{n_{j+1}-1} c_i e^{\alpha_j^{(0)} ukh} H(x_0) e^{\alpha_i ph\ell} e^{\alpha_i G(x_0)} \end{aligned}$$

$$= \sum_{j=1}^{\tilde{M}} \tilde{c}_j^{(\ell)} e^{\alpha_j^{(0)} u k h}$$

with

$$\tilde{c}_j^{(\ell)} := \sum_{i=n_j}^{n_{j+1}-1} \tilde{c}_i e^{\alpha_i p \ell h} = \sum_{i=n_j}^{n_{j+1}-1} c_i H(x_0) e^{\alpha_i p \ell h} e^{\alpha_i G(x_0)} \quad (5.33)$$

for $\ell = 0, \dots, 2M - 1$ and $j = 1, \dots, \tilde{M}$. Using the Vandermonde matrix $\mathbf{V}_{\tilde{M}}$ given in (5.32), we can rewrite this as

$$\mathbf{V}_{\tilde{M}} \cdot \left(\tilde{c}_j^{(\ell)} \right)_{j=1}^{\tilde{M}} = \left(f_{uk+p\ell} \right)_{k=0}^{\tilde{M}}$$

and compute $\tilde{c}_j^{(\ell)} = S_{G,H,ph}^{\ell} \left(\sum_{i=n_j}^{n_{j+1}-1} c_i H(\cdot) e^{\alpha_i G(\cdot)} \right) (x_0)$ for $j = 1, \dots, \tilde{M}$ and $\ell = 1, \dots, 2M - 1$.

The values $\tilde{c}_j^{(\ell)}$, $\ell = 1, \dots, 2M - 1$ are sample values of a generalized exponential sum of length $n_{j+1} - n_j$. In particular, we can compute the values $n_{j+1} - n_j$ by computing the rank of the Hankel matrix $\left(\tilde{c}_j^{(k+m)} \right)_{k,m=0}^{2M-1}$. Therefore, we can apply Theorem 5.16 and reconstruct the frequencies α_i and the corresponding coefficients c_i , $i = n_j, \dots, n_{j+1} - 1$ for $j = 1, \dots, \tilde{M}$. Consequently, the claim follows. \square

At last, we need to consider the case when aliasing causes a cancellation, i.e., when $\tilde{c}_k^{(0)} = 0$ for some $k = 1, \dots, M$.

Lemma 5.21 *Let $p, u \in \mathbb{N}$ with $\gcd(p, u) = 1$. Let $\alpha_1, \alpha_2 \in \mathbb{C}$, $\alpha_1 \neq \alpha_2$ with $|\operatorname{Im}(\alpha_j)| < K$, $j = 1, 2$ and $h = \frac{\pi}{K}$. If $e^{\alpha_1 u h} = e^{\alpha_2 u h}$, then $e^{\alpha_1 p h} \neq e^{\alpha_2 p h}$ holds.*

Proof. For a proof we refer to [CL20a], Lemma 3. \square

The above lemma implies that even in the case of colliding frequencies and vanishing coefficients we can recover the signal f . Therefore, we recall equation (5.33), i.e.,

$$\tilde{c}_j^{(\ell)} := \sum_{i=n_j}^{n_{j+1}-1} \tilde{c}_i e^{\alpha_i p \ell h} = \sum_{i=n_j}^{n_{j+1}-1} c_i H(x_0) e^{\alpha_i p \ell h} e^{\alpha_i G(x_0)} \quad (5.34)$$

for $j = 1, \dots, \tilde{M}$ and $\ell = 0, \dots, 2M - 1$. For fixed $j \in \mathbb{N}$ this implies the following Vandermonde system

$$\underbrace{\left(e^{\alpha_i p h \ell} \right)_{i=n_j, \ell=0}^{n_{j+1}-1, (n_{j+1}-n_j)}}_{=: \mathbf{V}_{n_{j+1}-n_j}} \cdot \left(\tilde{c}_i \right)_{i=n_j}^{n_{j+1}-1} = \left(\tilde{c}_j^{(\ell)} \right)_{\ell=0}^{n_{j+1}-n_j}.$$

Lemma 5.21 implies that the occurring Vandermonde matrix $\mathbf{V}_{n_{j+1}-n_j}$ is invertible.

Hence, we can conclude that not all $\tilde{c}_j^{(\ell)}$ are vanishing for $\ell = 0, \dots, n_{j+1} - n_j$ and $j = 1, \dots, \tilde{M}$. Consequently, we arrive at the following theorem.

Theorem 5.22 *Let $G: \mathbb{R} \rightarrow \mathbb{R}$ and $H: \mathbb{R} \rightarrow \mathbb{C}$ be continuous functions such that G is strictly monotone and H is non-vanishing on an interval $[a, b] \subset \mathbb{R}$. Furthermore, let f be a generalized exponential expansion of length M as in (5.1), such that there exists a $K \in \mathbb{R}$ with $|\text{Im}(\alpha_j)| < K$ for all $j = 1, \dots, M$. Moreover, let $h = \frac{\pi}{K}$, $u, p, \in \mathbb{N}$ satisfying $\text{gcd}(u, p) = 1$ and $x_0 \in [a, b]$ such that*

$$\begin{aligned} G(x_0) + h(uk + \ell p) \in [G(a), G(b)] & \quad \text{for } G(a) < G(b) & \quad \text{or} \\ G(x_0) + h(uk + \ell p) \in [G(b), G(a)] & \quad \text{for } G(a) > G(b) \end{aligned}$$

for all $k, \ell = 0, \dots, 2M - 1$. Then f can be uniquely reconstructed using the samples $f(G^{-1}(G(x_0) + h(uk + \ell p)))$ for $k, \ell = 0, \dots, 2M - 1$.

5.4 Modification of Prony's Method for Sparse Approximation

One of the most common applications of Prony's method is the sparse approximation of structured signals and in particular exponential sums, i.e., for signals that are either exactly or approximated by an exponential sum. Therefore, we want to study the question of approximating a given data vector $\mathbf{y} = (y_k)_{k=0}^N \in \mathbb{C}^{N+1}$ by a new vector $\mathbf{f} = (f_k)_{k=0}^N \in \mathbb{C}^{N+1}$ whose elements are of the form

$$f_k := \sum_{j=1}^M c_j z_j^k \quad (5.35)$$

with $2M \leq N$, $c_j, z_j \in \mathbb{C} \setminus \{0\}$, and z_j pairwise distinct for $j = 1, \dots, M$. We note that the length M of the sum in (5.35) is the rank of the Hankel matrix $\mathbf{H}_{N,N} = (f_{j+k})_{j,k=0}^N$ due to Kronecker's Theorem, see, e.g., Theorem 8.19, in [Fuh12], and, therefore, also has a close connection to structured low-rank Hankel approximation, see, e.g., [Mar12], which has been recently studied in the special case of rank-one Hankel approximation for the spectral norm as well as the Frobenius norm, see [KPP21].

The sparse approximation problem can be rewritten in terms of a **non-linear least squares problem**

$$\underset{\mathbf{c}, \mathbf{z} \in \mathbb{C}^M}{\text{argmin}} \left\| \mathbf{y} - \left(\sum_{j=1}^M c_j z_j^k \right)_{k=0}^N \right\|_2^2. \quad (5.36)$$

This formulation can be used to solve the problem of parameter estimation in the case

of noisy data, i.e., when the data samples y_k are of the form

$$y_k = \frac{H(x_0)}{H(G^{-1}(G(x_0) + hk))} f(G^{-1}(G(x_0) + hk)) + \epsilon_k \quad k = 0, \dots, N,$$

where $\epsilon_k, k = 0, \dots, N$, is the error term and f is a generalized exponential sum as in (5.1), i.e., of the form $f(x) = \sum_{j=1}^M c_j H(x) e^{\alpha_j G(x)}$, since we have

$$\frac{H(x_0)}{H(G^{-1}(G(x_0) + hk))} f(G^{-1}(G(x_0) + hk)) = \sum_{j=1}^M c_j H(x_0) e^{\alpha_j G(x_0)} e^{\alpha_j hk} = \sum_{j=1}^M \tilde{c}_j z_j^k$$

with $\tilde{c}_j := c_j H(x_0) e^{\alpha_j G(x_0)}$ and $z_j := e^{\alpha_j h}$. If the error terms $\epsilon_k, k = 0, \dots, N$, are independently and identically distributed random variables with $\epsilon_k \in N(0, \sigma^2)$, the minimization problem can also be interpreted as a maximum likelihood method, see [SM97], Appendix B, and [Os75]. Furthermore, we can also understand this as a non-linear approximation problem of finding a generalized exponential sum f as in (5.1) such that

$$\sum_{k=0}^N \left| y_k - \frac{H(x_0)}{H(G^{-1}(G(x_0) + hk))} f(G^{-1}(G(x_0) + hk)) \right|^2$$

is minimized.

We recall the Vandermonde matrix

$$\mathbf{V}_{N,M}(z_1, \dots, z_M) := \mathbf{V}_z := \begin{pmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_M \\ z_1^2 & z_2^2 & \dots & z_M^2 \\ \vdots & \vdots & & \vdots \\ z_1^N & z_2^N & \dots & z_M^N \end{pmatrix} \in \mathbb{C}^{(N+1) \times M} \quad (5.37)$$

with corresponding node vector $\mathbf{z} := (z_1, z_2, \dots, z_M)$. In the following our Vandermonde matrix always will be of size $(N+1) \times M$. Since we focus on the recovery of the parameters $z_j, j = 1, \dots, M$, we use the notation \mathbf{V}_z throughout this section. Using this notation we can write $\mathbf{f} = \mathbf{V}_z \cdot \mathbf{c}$, with $\mathbf{f} = (f_0, \dots, f_N)$ and $\mathbf{c} = (c_1, \dots, c_M)^T \in \mathbb{C}^M$ denoting the coefficient vector. Thus, we can reformulate (5.36) as

$$\operatorname{argmin}_{\mathbf{c}, \mathbf{z} \in \mathbb{C}^M} \|\mathbf{y} - \mathbf{V}_z \mathbf{c}\|_2. \quad (5.38)$$

We note that the non-linear least square problem in (5.36) is NP-hard to solve, see [Nat95]. We want to use the so-called variable projection method in order to solve (5.38) and transfer the problem to a minimization problem with respect to \mathbf{c} .

Lemma 5.23 *Let $N, M \in \mathbb{N}$ with $2M \leq N$ and let $\mathbf{y} \in \mathbb{C}^{N+1}$ be the vector of given data samples. Furthermore, let \mathbf{V}_z denote the Vandermonde matrix given in (5.37) and $\mathbf{c} \in$*

$\mathbb{C}^M \setminus \{0\}$ denote the vector of (unknown) coefficients. If the parameter $\mathbf{z} \in \mathbb{C}^M$ is fixed, the non-linear optimization problem given in (5.38) is equivalent to

$$\operatorname{argmax}_{\mathbf{z} \in \mathbb{C}^M} \|\mathbf{r}(\mathbf{z})\|_2^2 = \operatorname{argmax}_{\mathbf{z} \in \mathbb{C}^M} \|\mathbf{P}_z \mathbf{y}\|_2^2, \quad (5.39)$$

with $\mathbf{P}_z = \mathbf{V}_z \cdot \mathbf{V}_z^+$, $\mathbf{r}(\mathbf{z}) := \mathbf{P}_z \cdot \mathbf{y} \in \mathbb{C}^{N+1}$ and $\mathbf{c} = \mathbf{V}_z^+ \cdot \mathbf{y}$.

Proof. For fixed $\mathbf{z} \in \mathbb{C}^M$, the problem in (5.38) reduces to a linear least squares problem

$$\operatorname{argmin}_{\mathbf{c} \in \mathbb{C}^M} \|\mathbf{y} - \mathbf{V}_z \cdot \mathbf{c}\|_2^2.$$

Penrose's Theorem, see [SS90], Section III, Theorem 1.4, implies that the solution is given by $\mathbf{c} = \mathbf{V}_z^+ \mathbf{y}$, where \mathbf{V}_z^+ denotes the Moore-Penrose inverse given in Definition 5.6. Since the Vandermonde matrix \mathbf{V}_z has full rank M , Lemma 5.7 (i) yields

$$\mathbf{c} = \mathbf{V}_z^+ \cdot \mathbf{y} = (\mathbf{V}_z^* \cdot \mathbf{V}_z)^{-1} \cdot \mathbf{V}_z^* \cdot \mathbf{y},$$

where \mathbf{V}_z^* denotes the conjugate transpose of \mathbf{V}_z . Additionally, we introduce the projection matrix $\mathbf{P}_z \in \mathbb{C}^{N+1 \times N+1}$ by

$$\mathbf{P}_z := \mathbf{V}_z \cdot \mathbf{V}_z^+ \quad (5.40)$$

and satisfying

$$\mathbf{P}_z = \mathbf{P}_z^*, \mathbf{P}_z = \mathbf{P}_z^2, \mathbf{P}_z \cdot \mathbf{V}_z = \mathbf{V}_z \text{ and } \mathbf{V}_z^+ \cdot \mathbf{P}_z = \mathbf{V}_z^+. \quad (5.41)$$

Hence, problem (5.38) can be simplified to

$$\begin{aligned} \operatorname{argmin}_{\mathbf{z} \in \mathbb{C}^M} \|\mathbf{y} - \mathbf{V}_z \cdot \mathbf{V}_z^+ \cdot \mathbf{y}\|_2^2 &= \operatorname{argmin}_{\mathbf{z} \in \mathbb{C}^M} \|(\mathbf{I}_{N+1} - \mathbf{P}_z) \cdot \mathbf{y}\|_2^2 \\ &= \operatorname{argmin}_{\mathbf{z} \in \mathbb{C}^M} (\mathbf{y}^* \cdot (\mathbf{I}_{N+1} - \mathbf{P}_z)^* \cdot (\mathbf{I}_{N+1} - \mathbf{P}_z) \cdot \mathbf{y}) \\ &= \operatorname{argmin}_{\mathbf{z} \in \mathbb{C}^M} (\mathbf{y}^* \mathbf{y} - \mathbf{y}^* \cdot \mathbf{P}_z \cdot \mathbf{y}). \end{aligned}$$

Therefore, we need to solve

$$\operatorname{argmax}_{\mathbf{z} \in \mathbb{C}^M} \mathbf{y}^* \cdot \mathbf{P}_z \cdot \mathbf{y}.$$

Now we define $\mathbf{r}(\mathbf{z}) := \mathbf{P}_z \cdot \mathbf{y} \in \mathbb{C}^{N+1}$. Then the optimization problem is equivalent to

$$\operatorname{argmax}_{\mathbf{z} \in \mathbb{C}^M} \|\mathbf{r}(\mathbf{z})\|_2^2 = \operatorname{argmax}_{\mathbf{z} \in \mathbb{C}^M} \|\mathbf{P}_z \cdot \mathbf{y}\|_2^2.$$

□

In order to solve the derived maximization problem in (5.39) we consider the Jacobian matrix of the vector $\mathbf{r}(\mathbf{z})$.

Theorem 5.24 *Let $N, M \in \mathbb{N}$ with $N \geq 2M$ and $\mathbf{y} \in \mathbb{C}^{N+1}$ be a given data vector. Furthermore, let \mathbf{V}_z be the Vandermonde matrix given in (5.37), $\mathbf{P}_z = \mathbf{V}_z \cdot \mathbf{V}_z^+$ be the corresponding projection matrix and $\mathbf{P}_z \cdot \mathbf{y} =: \mathbf{r}(\mathbf{z}) = (r_\ell(\mathbf{z}))_{\ell=0}^N$. Then the Jacobian matrix $\mathbf{J}_z \in \mathbb{C}^{(N+1) \times M}$ of $\mathbf{r}(\mathbf{z})$ is given by*

$$\begin{aligned} \mathbf{J}_z &:= \left(\frac{\partial r_\ell(\mathbf{z})}{\partial z_j} \right)_{\ell=0, j=1}^{N, M} \\ &= (\mathbf{I}_{N+1} - \mathbf{P}_z) \cdot \mathbf{V}_z^+ \cdot \text{diag}(\mathbf{V}_z^+ \mathbf{y}) + (\mathbf{V}_z^+)^* \cdot \text{diag}((\mathbf{V}_z^+)^* \cdot (\mathbf{I}_{N+1} - \mathbf{P}_z) \cdot \mathbf{y}), \end{aligned} \quad (5.44)$$

where \mathbf{I}_{N+1} denotes the identity matrix of size $N+1$ and

$$\mathbf{V}_z' = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 1 & 1 & \dots & 1 \\ 2z_1 & 2z_2 & \dots & 2z_M \\ \vdots & \vdots & & \vdots \\ Nz_1^{N-1} & Nz_2^{N-1} & \dots & Nz_M^{N-1} \end{pmatrix} \in \mathbb{C}^{(N+1) \times M}.$$

In particular,

$$\nabla \|\mathbf{r}(\mathbf{z})\|_2^2 = 2\mathbf{J}_z^* \cdot \mathbf{r}(\mathbf{z}) = 2 \text{diag} \left((\mathbf{V}_z')^T \cdot (\mathbf{I}_{N+1} - \bar{\mathbf{P}}_z) \cdot \bar{\mathbf{y}} \right) \cdot \mathbf{V}_z^+ \cdot \mathbf{y}. \quad (5.45)$$

Proof. First, we observe that the partial derivative $\frac{\partial}{\partial z_j} \mathbf{V}_z$ is a rank-1 matrix of the form

$$\frac{\partial}{\partial z_j} \mathbf{V}_z = \mathbf{z}'_j \mathbf{e}_j^* \in \mathbb{C}^{(N+1) \times M}, \quad j = 1, \dots, M, \quad (5.46)$$

where $\mathbf{z}'_j := (0, 1, 2z_j, 3z_j^2, \dots, Nz_j^{N-1})^T$, and \mathbf{e}_j is the j -th unit vector of length M . Applying the general rule

$$\frac{\partial \mathbf{Y}(\mathbf{x})^{-1}}{\partial x_j} = -\mathbf{Y}(\mathbf{x})^{-1} \frac{\partial \mathbf{Y}(\mathbf{x})}{\partial x_j} \mathbf{Y}(\mathbf{x})^{-1}, \quad \mathbf{Y}(\mathbf{x}) \in \mathbb{C}^{n \times m}, \mathbf{x} = (x_1, \dots, x_m)^T \text{ for } n, m \in \mathbb{N},$$

to $\mathbf{P}_z = \mathbf{V}_z \cdot \mathbf{V}_z^+ = \mathbf{V}_z \cdot (\mathbf{V}_z^* \cdot \mathbf{V}_z)^{-1} \cdot \mathbf{V}_z^*$ and using (5.46) we obtain

$$\begin{aligned} \frac{\partial}{\partial z_j} \mathbf{r}(\mathbf{z}) &= \frac{\partial}{\partial z_j} (\mathbf{P}_z \cdot \mathbf{y}) = \frac{\partial}{\partial z_j} (\mathbf{V}_z \cdot \mathbf{V}_z^+ \cdot \mathbf{y}) \\ &= \left(\frac{\partial}{\partial z_j} \mathbf{V}_z \right) \cdot \mathbf{V}_z^+ \mathbf{y} - (\mathbf{V}_z^+)^* \cdot \left[\frac{\partial}{\partial z_j} (\mathbf{V}_z^* \cdot \mathbf{V}_z) \right] \cdot \mathbf{V}_z^+ \mathbf{y} + (\mathbf{V}_z^+)^* \cdot \left(\frac{\partial}{\partial z_j} \mathbf{V}_z \right)^* \cdot \mathbf{y} \\ &= (\mathbf{z}'_j \mathbf{e}_j^*) \mathbf{V}_z^+ \mathbf{y} - (\mathbf{V}_z^+)^* \cdot \left[(\mathbf{z}'_j \mathbf{e}_j^*)^* \mathbf{V}_z + \mathbf{V}_z^* (\mathbf{z}'_j \mathbf{e}_j^*) \right] \cdot \mathbf{V}_z^+ \mathbf{y} + (\mathbf{V}_z^+)^* \cdot (\mathbf{z}'_j \mathbf{e}_j^*)^* \cdot \mathbf{y} \\ &= (\mathbf{V}_z^+ \mathbf{y})_j \mathbf{z}'_j - \left((\mathbf{z}'_j)^* \mathbf{P}_z \mathbf{y} \right) \cdot (\mathbf{V}_z^+)^* \mathbf{e}_j - (\mathbf{V}_z^+ \mathbf{y})_j \mathbf{P}_z \mathbf{z}'_j + \left((\mathbf{z}'_j)^* \mathbf{y} \right) \cdot (\mathbf{V}_z^+)^* \mathbf{e}_j \end{aligned}$$

$$\begin{aligned}
 &= (\mathbf{V}_z^+ \cdot \mathbf{y})_j (\mathbf{I}_{N+1} - \mathbf{P}_z) \cdot \mathbf{z}'_j + ((\mathbf{z}'_j)^* \cdot (\mathbf{I}_{N+1} - \mathbf{P}_z) \cdot \mathbf{y}) \cdot (\mathbf{V}_z^+)^* \cdot \mathbf{e}_j \\
 &= (\mathbf{V}_z^+ \mathbf{y})_j (\mathbf{I}_{N+1} - \mathbf{P}_z) \cdot \mathbf{V}'_z \cdot \mathbf{e}_j + ((\mathbf{z}'_j)^* \cdot (\mathbf{I}_{N+1} - \mathbf{P}_z) \cdot \mathbf{y}) \cdot (\mathbf{V}_z^+)^* \cdot \mathbf{e}_j,
 \end{aligned}$$

where $(\mathbf{V}_z^+ \mathbf{y})_j$ denotes the j -th component of the vector $\mathbf{V}_z^+ \mathbf{y}$. This implies

$$\mathbf{J}_z = (\mathbf{I}_{N+1} - \mathbf{P}_z) \cdot \mathbf{V}'_z \cdot \text{diag}(\mathbf{V}_z^+ \mathbf{y}) + (\mathbf{V}_z^+)^* \cdot \text{diag} \left(\left((\mathbf{z}'_j)^* \cdot (\mathbf{I}_{N+1} - \mathbf{P}_z) \cdot \mathbf{y} \right)_{j=1}^M \right).$$

Consequently, we obtain that the Jacobian matrix \mathbf{J}_z of $\mathbf{r}(\mathbf{z})$ is of the form (5.44). Furthermore, using (5.41) we get

$$\begin{aligned}
 \mathbf{J}_z^* \mathbf{r}(\mathbf{z}) &= \left(\text{diag} \overline{\mathbf{V}_z^+ \mathbf{y}} \right) \cdot (\mathbf{V}'_z)^* \cdot (\mathbf{I}_{N+1} - \mathbf{P}_z)^* \cdot \mathbf{P}_z \cdot \mathbf{y} \\
 &\quad + \left(\text{diag} \left((\mathbf{V}'_z)^* \cdot (\mathbf{I}_{N+1} - \mathbf{P}_z) \cdot \mathbf{y} \right) \right)^* \cdot \mathbf{V}_z^+ \cdot \mathbf{P}_z \cdot \mathbf{y} \\
 &= \left(\text{diag} \left((\mathbf{V}'_z)^T \cdot (\mathbf{I}_{N+1} - \bar{\mathbf{P}}_z) \cdot \bar{\mathbf{y}} \right) \right) \cdot \mathbf{V}_z^+ \cdot \mathbf{y}.
 \end{aligned}$$

Hence, the gradient of $\|\mathbf{r}(\mathbf{z})\|_2^2$ given by $\nabla \|\mathbf{r}(\mathbf{z})\|_2^2 = 2\mathbf{J}_z \cdot \mathbf{r}(\mathbf{z})$ has the desired form (5.39). \square

Corollary 5.25 *Let $N, M \in \mathbb{N}$ with $N \geq 2M$ and $\mathbf{y} \in \mathbb{C}^{N+1}$ be the given data vector. Assume that the corresponding Hankel matrix $\mathbf{H}_{N-M+2, M} = (y_{k+\ell})_{k, \ell=0}^{N-M+1, M-1}$ has full rank M and $\mathbf{V}_z^T \mathbf{y}$ has no vanishing components. Then the solution to the optimization problem given in (5.39) necessarily satisfies*

$$(\mathbf{V}'_z)^T \cdot (\mathbf{I}_{N+1} - \bar{\mathbf{P}}_z) \cdot \bar{\mathbf{y}} = \mathbf{0}_{N+1}.$$

Proof. We assume \mathbf{z} solves the optimization problem in (5.39). This implies $\nabla \|\mathbf{r}(\mathbf{z})\|_2^2 = \mathbf{0}_{N+1}$. Hence, Theorem 5.24 yields

$$\nabla \|\mathbf{r}(\mathbf{z})\|_2^2 = 2 \text{diag} \left((\mathbf{V}'_z)^T \cdot (\mathbf{I}_{N+1} - \bar{\mathbf{P}}_z) \cdot \bar{\mathbf{y}} \right) \cdot \mathbf{V}_z^+ \cdot \mathbf{y} = \mathbf{0}_{N+1}.$$

Since $\mathbf{V}_z^+ \mathbf{y}$ has no vanishing components the claim follows. \square

Remark 5.26 Based on this analysis we proposed the idea for an iterative algorithm for the approximation on the vector \mathbf{z} in [KP21]. For the sake of scientific accuracy and validity, we want to explain, why this proposed algorithm is not feasible. Therefore, we will briefly present the suggested method: We start with initial guess $\mathbf{z}^{(0)}$ obtained by Algorithm 7. Then the vector $\mathbf{z}^{(j+1)}$, $j \in \mathbb{N}_0$ is updated by solving

$$(\mathbf{V}'_{z^{(j+1)}})^T \cdot (\mathbf{I}_{N+1} - \bar{\mathbf{P}}_{z^{(j)}}) \cdot \bar{\mathbf{y}} = \mathbf{0}_{N+1} \quad j = 0, \dots,$$

i.e., by computing the zeros of the polynomial $P_j(x) = \sum_{k=0}^N p_k^{(j)} x^k$ with the coefficient

vector $\mathbf{p}^{(j)} = (p_k^{(j)})_{k=0}^N := \text{diag}(0, 1, \dots, N) \cdot (\mathbf{I}_{N+1} - \overline{\mathbf{P}}_{\mathbf{z}^{(j)}}) \cdot \overline{\mathbf{y}}$. This results in N zeros. In order to get rid of the superfluous zeros we choose the subset of size M which is closest to the subset containing the coefficients of $\mathbf{z}^{(j)}$.

This algorithm is not feasible, since Lemma 5.7(iv) implies that if the vector $\mathbf{z}^{(j)}$ is the solution of the optimization problem given in (5.39), it lies in the column space of $\mathbf{V}_{\mathbf{z}^{(j)}}$. Therefore, we have $(\mathbf{I}_{N+1} - \mathbf{P}_{\mathbf{z}^{(j)}})\mathbf{y} = \mathbf{0}_{N+1}$ and hence $\mathbf{p}^{(j)} = \mathbf{0}_{N+1}$. Consequently, the corresponding polynomial $P_{(j)}$ is the zero polynomial. If the vector $\mathbf{z}^{(j)}$ is close to the optimal solution machine precision may yield only $\mathbf{p}^{(j)} \approx \mathbf{0}_{N+1}$, we may obtain $P_{(j)} \not\equiv 0$ but still $P_{(j)} \approx 0$. In this case, the roots of $P_{(j)}$ are in general not close to the optimal solution $\mathbf{z}^{(j)}$. \diamond

5.4.1 Gauß-Newton and Levenberg-Marquardt Iteration

In this section we will discuss a different approach to solve the non-linear least squares problem (5.39) based on the Newton method. In each iteration step, the parameter vector $\mathbf{r}(\mathbf{z}) = \mathbf{P}_{\mathbf{z}} \cdot \mathbf{y}$ is replaced by a new estimate $\mathbf{r}(\mathbf{z} + \delta)$ for $\delta \in \mathbb{C}^M$. This new vector is now approximated by its first order Taylor expansion $\mathbf{r}(\mathbf{z}) + \mathbf{J}_{\mathbf{z}} \cdot \delta$. Hence, instead of maximizing $\|\mathbf{r}(\mathbf{z} + \delta)\|_2^2$ we consider

$$\begin{aligned} \operatorname{argmax}_{\delta \in \mathbb{C}^M} \|\mathbf{r}(\mathbf{z}) + \mathbf{J}_{\mathbf{z}} \cdot \delta\|_2^2 &= \operatorname{argmax}_{\delta \in \mathbb{C}^M} (\mathbf{r}(\mathbf{z}) + \mathbf{J}_{\mathbf{z}} \cdot \delta)^* \cdot (\mathbf{r}(\mathbf{z}) + \mathbf{J}_{\mathbf{z}} \cdot \delta) \\ &= \operatorname{argmax}_{\delta \in \mathbb{C}^M} \left(\|\mathbf{r}(\mathbf{z})\|_2^2 + \mathbf{r}(\mathbf{z}) \cdot \mathbf{J}_{\mathbf{z}} \cdot \delta + \delta^* \cdot \mathbf{J}_{\mathbf{z}}^* \cdot \mathbf{r}(\mathbf{z}) + \delta^* \mathbf{J}_{\mathbf{z}}^* \cdot \mathbf{J}_{\mathbf{z}} \delta \right), \end{aligned}$$

which yields

$$2\operatorname{Re}(\mathbf{J}_{\mathbf{z}}^* \mathbf{r}(\mathbf{z})) + 2\mathbf{J}_{\mathbf{z}}^* \cdot \mathbf{J}_{\mathbf{z}} \delta = \mathbf{0}_{N+1}.$$

Hence, the corresponding **Gauss-Newton iteration** at the j -th step is of the form

$$(\mathbf{J}_{\mathbf{z}^{(j)}}^* \cdot \mathbf{J}_{\mathbf{z}^{(j)}}) \delta^{(j)} = -\operatorname{Re} \left(\mathbf{J}_{\mathbf{z}^{(j)}} \mathbf{r} \left(\mathbf{z}^{(j)} \right) \right),$$

in order to obtain the improved vector $\mathbf{z}^{(j+1)} = \mathbf{z}^{(j)} + \delta^{(j)}$ with $\mathbf{J}_{\mathbf{z}^{(j)}}$ as in Theorem 5.24.

Since the vector $(\mathbf{I}_{N+1} - \overline{\mathbf{P}}_{\mathbf{z}^{(j)}}) \cdot \mathbf{y}$ may already be close to the zero vector, the matrix $(\mathbf{J}_{\mathbf{z}^{(j)}}^* \mathbf{J}_{\mathbf{z}^{(j)}})$ is usually ill-conditioned. The **Levenberg-Marquardt algorithm**, see, e.g., [Mor78], introduces a regularization changing the coefficient matrix at each iteration step. The matrix is changed to

$$(\mathbf{J}_{\mathbf{z}^{(j)}}^* \cdot \mathbf{J}_{\mathbf{z}^{(j)}}) + \mu_j \mathbf{I}_M$$

for $\mu_j > 0$. Then the iteration reads as follows

$$(\mathbf{J}_{\mathbf{z}^{(j)}}^* \cdot \mathbf{J}_{\mathbf{z}^{(j)}} + \mu_j \mathbf{I}_M) \delta^{(j)} = -\operatorname{Re} \left(\mathbf{J}_{\mathbf{z}^{(j)}} \mathbf{r} \left(\mathbf{z}^{(j)} \right) \right).$$

If we arrive at a (local) maximum of $\|\mathbf{r}(\mathbf{z})\|_2^2$, the Levenberg-Marquardt iteration vanishes, and we obtain $\delta^{(j)} = \mathbf{0}$. This optimization algorithm is in general very fast. Since this algorithm may only lead to a local maximum, the solution highly depends on the input vector. One possible option for an input vector $\mathbf{z}^{(0)}$ is the solution obtained by the ESPRIT method described in Section 5.1. Nevertheless, one drawback is that the sparsity number M is not used and the matrices are of size $N + 1$. Therefore, it is beneficial to use the classical approach described in e.g., [ZP19] and above. Finally, we want to conclude that the numerical experiments in [ZP19] indicate, that the variable projection method based on the Levenberg-Marguardt Iteration as well as the Prony method or the ESPRIT method seem to perform well for the approximation problem considered in this section.

5.5 Numerical Examples

In this section we want to illustrate our derived algorithms using different numerical examples.

To this end, we begin by considering the following generalized exponential sum, which is sometime studied in the field of electrical engineering, see, e.g., [Mil15].

Definition 5.27 (Exponential Cosine Expansion) Let $M \in \mathbb{N}$. Then we call f an M -sparse exponential cosine expansion if f is of the form

$$f(x) = \sum_{j=1}^M c_j e^{\alpha_j \cos(x)} \quad (5.47)$$

with $c_j \in \mathbb{C} \setminus \{0\}$ and $\alpha_j \in \mathbb{C}$ for $j = 1, \dots, M$.

This expansion is, indeed, a generalized exponential as in (3.12) with $G(x) := \cos(x)$ and $H \equiv 1$. In order to ensure that G is strictly monotone, we chose the sampling interval $(\pi, 2\pi)$. Then the corresponding generalized shift operator $S_{G,H,h}$ is of the form

$$S_{G,H,h}f(x) = f(G^{-1}(G(x) + h)) = f(\arccos(h + \cos(x))).$$

Example 5.28 In this example we compare the performance of the classical Prony method in Algorithm 3 with the ESPRIT method in Algorithm 7 and focus on the reconstruction of the frequency parameters. We consider a sparse exponential cosine expansion f as in (5.47) and can use Theorem 3.31 for the reconstruction of f . We remark that the sampling distance h strongly depends on the length of the interval, in which G is strictly monotone, as well as on the slope of G^{-1} . Hence, in order to be able to apply Theorem 3.31, we need to choose a sampling distance $h \in \mathbb{R} \setminus \{0\}$ and a starting point $x_0 \in \mathbb{R}$ such that $\cos(x_0) + hk \in (-1, 1)$ for $k = 0, \dots, 2M - 1$.

This is satisfied for $x_0 = \pi + h/2$ for $0 < h \leq \frac{\pi}{2M+1}$. Since we also want to apply the modified ESPRIT method to determine the parameter M , we oversample this function and consider $2N$ samples with $N \geq M$. Therefore, we choose $0 < h \leq \frac{\pi}{2N+1}$ and $x_0 = \pi + h/2$, satisfying the conditions for Theorem 3.31 and the ESPRIT method.

For our numerical example we choose $M = 5$. The original real parameters in Table 5.1 have been drawn uniformly from the intervals $(-1, 1)$ for c_j and from $(-\pi, \pi)$ for α_j , $j = 1, \dots, M$.

parameters	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
c_j	0.7171	0.8221	0.3993	0.4504	-0.5402
α_j	-1.1251	0.0717	-2.7608	1.4180	0.3554

Table 5.1: Parameters c_j and α_j for a sparse exponential cosine expansion as in (5.47) with $M = 5$, see Figure 5.2.

For the reconstruction using Algorithm 7 we use $N = 17$ and, therefore, the samples $f(\arccos(hk + \cos(x_0)))$, $k = 0, \dots, 33$, with $h = \frac{1}{35}$ and $x_0 = \pi + \frac{h}{2} = \pi + \frac{1}{70}$. We assume an upper bound of $L = 12$. For the rank approximation in step 1 of Algorithm 7 we use the tolerance $\varepsilon = 10^{-10}$. The singular values of the matrices $\mathbf{H}_{22,13}$ as well as $\mathbf{H}_{29,6}$ are shown in Figure 5.1. In both cases a clear gap after the fifth singular value is visible. Furthermore, this visualisation coincides with Lemma 5.11 since the singular values of the matrix $\mathbf{H}_{22,13}$ are slightly larger than the singular values of $\mathbf{H}_{29,6}$.

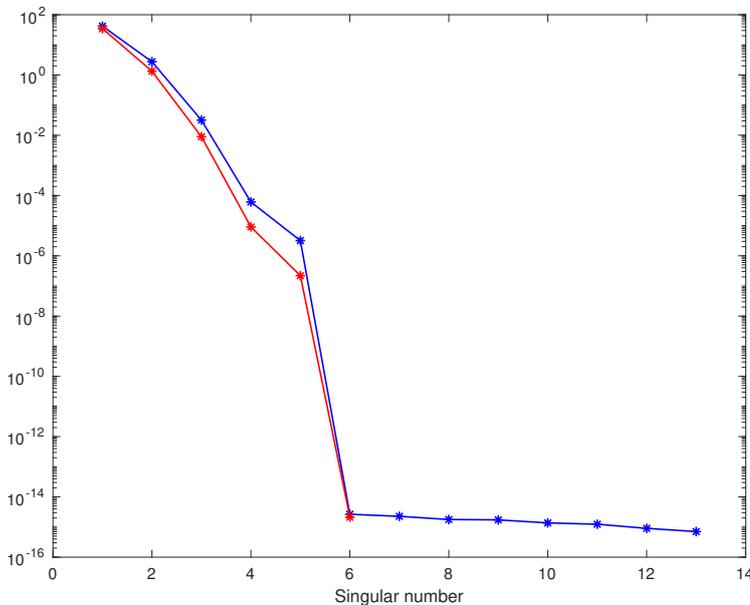


Figure 5.1: Singular values of the matrices $\mathbf{H}_{22,13}$ (blue) and $\mathbf{H}_{29,6}$ (red) used in Algorithm 7 on the logarithmic y-axis and the corresponding index on the x-axis.

For the reconstruction using the classical Prony Algorithm 3 we used the 10 samples $f(\arccos(hk + \cos(x_0)))$ for $k = 0, \dots, 9$.

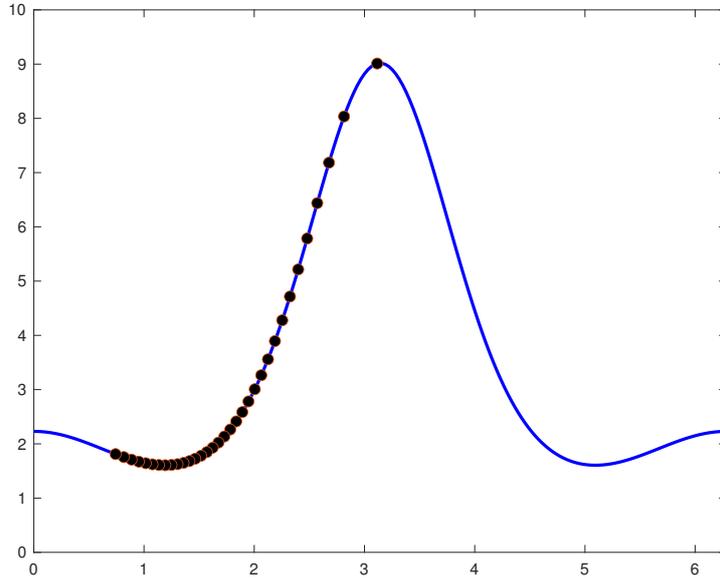


Figure 5.2: 5-sparse exponential cosine expansion with parameters given in Table 5.1. The black dots indicate the sampled used for the reconstruction using Algorithm 7.

Table 5.2 shows the maximal reconstruction error for the frequency parameters α_j , $j = 1, \dots, M$ given by

$$\text{err} := \max_{j=1, \dots, 5} |\alpha_j - \tilde{\alpha}_j|,$$

where $\tilde{\alpha}_j$ denote the recovered parameter for $j = 1, \dots, 5$.

	Algorithm 4	Algorithm 7 ($L = 12$)
err	0.0904	$3.1028 \cdot 10^{-6}$

Table 5.2: Maximal error in the frequency coefficients α_j for a 5-sparse exponential cosine expansion as in (5.47) obtain by Algorithm 3 and Algorithm 7.

This shows that the estimation obtained by the ESPRIT Algorithm yields a much more accurate result than Algorithm 3. One reason for this is the fact that our sampling distance h is rather small and the samples therefore close, as depicted in Figure 5.2.

◇

Example 5.29 We also want to test our modified ESPRIT method on noisy data. Therefore, we consider an M -sparse expansion into Gaussian chirps as in (3.28), i.e.,

$$f(x) = \sum_{j=1}^M c_j e^{-\beta(x-\alpha_j)^2} \quad (5.48)$$

with $\beta = i$, $c_j \in \mathbb{C} \setminus \{0\}$, $\alpha_j \in \mathbb{C}$, $j = 1, \dots, M$, as in Table 5.3 and $M = 5$.

parameters	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
c_j	1	-3	0.4	0.01	2
$\operatorname{Re}(\alpha_j)$	0	$-\frac{\pi}{2}$	-0.25	-1	0.2
$\operatorname{Im}(\alpha_j)$	1	0	0.4	0.01	0

Table 5.3: Parameters c_j and α_j for a 5-sparse expansion into Gaussian chirps as in (5.48).

Remark 3.49 implies, that this a generalized exponential sum with $G(x) = 0$ and $H(x) = e^{-ix^2}$ and can be reconstructed using the samples $f(x_0 + hk)$, $k = 0, \dots, 2M - 1$. We consider the noisy data samples with

$$\tilde{f}(x_0 + hk) := f(x_0 + hk) + \epsilon_k$$

where ϵ_k additive white Gaussian noise with mean 0 and variance σ^2 . We perform tests for 7 different variances $\sigma^2 = 10^{-14}, 10^{-13}, \dots, 10^{-8}$ using Algorithm 7. We use the starting point $x_0 = 0$ and the sampling distance $h = 0.1$. We choose an the upper bound of $L = 8$, $2N = 2 \cdot 12 = 24$ samples and a tolerance of $\epsilon = 10^{-6}$. The maximal reconstruction errors for the frequencies are shown in Table 5.4. Here, we have $\alpha := (\alpha_j)_{j=1}^5$ and $\tilde{\alpha} := (\tilde{\alpha}_j)_{j=1}^5$ with $\tilde{\alpha}_j$, $j = 1, \dots, 5$, the parameter obtained by Algorithm 7.

variance σ^2	$\ \alpha - \tilde{\alpha}\ _\infty$
10^{-14}	$5.8723 \cdot 10^{-7}$
10^{-13}	$1.8487 \cdot 10^{-6}$
10^{-12}	$3.5068 \cdot 10^{-6}$
10^{-11}	$1.4566 \cdot 10^{-5}$
10^{-10}	$5.5265 \cdot 10^{-5}$
10^{-09}	$8.8796 \cdot 10^{-5}$
10^{-08}	$5.1598 \cdot 10^{-4}$

Table 5.4: Performance results of Algorithm 7 using different noise levels.

We observe that for growing noise variance the accuracy of the reconstructed values decreases. In particular, this may result to the recovery of an additional frequency parameter α_6 , if the tolerance ϵ is not chosen appropriately.

◇

Example 5.30 In this example we want to test Algorithm 8 on a sparse exponential sine expansion. Therefore, let f be a 5-sparse exponential sine expansion, i.e.,

$$f(x) = \sum_{j=1}^5 c_j e^{\alpha_j \sin(x)} \quad (5.49)$$

with parameters c_j and α_j , $j = 1, \dots, 5$, as given in Table 5.5. The minimal separation distance in our example is given by $q = 0.01$ and the parameters α_1 and α_2 are clustered. Moreover, the signal f is, indeed, a generalized exponential sum with $H \equiv 1$ and $G(x) := \sin(x)$.

parameters	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
c_j	1	-0.5	-0.02	1.7	-0.84
Re α_j	-1	-0.99	-2	1.7	0.3
Im α_j	-1.12	-1.1199	-5	-5	-1.1189

Table 5.5: Parameters c_j and α_j , $j = 1, \dots, 5$, for a 5-sparse exponential sine expansion as in (5.49).

In order to ensure that G is strictly monotone we consider the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$. Similarly to Example 5.28 the sampling distance strongly depends on the length of our considered interval as well as on the slope of G^{-1} . We choose the sampling distance $h = \frac{1}{20}$ and the starting point $x_0 = -\frac{\pi}{2} + 0.01$. Furthermore, we set $p = 3$, $u = 11$ and choose the samples as in Theorem 5.16. For the application of Algorithm 8 we use 10 samples of the form $f(\arcsin(\sin(x_0) + uhl))$ for $\ell = 0, \dots, 9$ as well as 5 samples of the form $f(\arcsin(\sin(x_0) + h(u\ell + p)))$ for $\ell = 0, \dots, 5$. The parameters obtained by Algorithm 8 are denoted by \tilde{c}_j and $\tilde{\alpha}_j$, $j = 1, \dots, 5$. The reconstruction errors are shown in Table 5.6.

j	$ \alpha_j - \tilde{\alpha}_j $	$ c_j - \tilde{c}_j $
$j = 1$	$1.9183 \cdot 10^{-07}$	$3.8487 \cdot 10^{-05}$
$j = 2$	$3.3868 \cdot 10^{-07}$	$1.5666 \cdot 10^{-04}$
$j = 3$	$1.1341 \cdot 10^{-08}$	$4.9707 \cdot 10^{-12}$
$j = 4$	$6.9634 \cdot 10^{-13}$	$5.7275 \cdot 10^{-14}$
$j = 5$	$4.9968 \cdot 10^{-10}$	$7.7323 \cdot 10^{-10}$

Table 5.6: Error in the numerical evaluations of the parameters c_j and α_j for the 5-sparse exponential sine expansion as in (5.49) obtained using Algorithm 8.

We compare this result with the outcome of Algorithm 3. In this case, we only need 10 samples of the form $f(\arcsin(\sin(x_0) + h\ell))$ for $\ell = 0, \dots, 9$. The reconstruction errors obtained by Algorithm 3 are shown in Table 5.7. For better visualization we

present the frequencies α_j , $j = 1, \dots, 5$, as well as the parameters obtained by Algorithm 8 and Algorithm 3 in Figure 5.3.

j	$ \alpha_j - \tilde{\alpha}_j $	$ c_j - \tilde{c}_j $
$j = 1$	0.0096	0.4997
$j = 2$	1.7923	0.5000
$j = 3$	$1.2890 \cdot 10^{-6}$	$6.1338 \cdot 10^{-8}$
$j = 4$	$9.7038 \cdot 10^{-7}$	$2.7015 \cdot 10^{-6}$
$j = 5$	$9.7251 \cdot 10^{-5}$	$4.3585 \cdot 10^{-4}$

Table 5.7: Error in the numerical evaluations of the parameters c_j and α_j , $j = 1, \dots, 5$, for the 5-sparse exponential sine expansion as in (5.49) obtained using Algorithm 3.

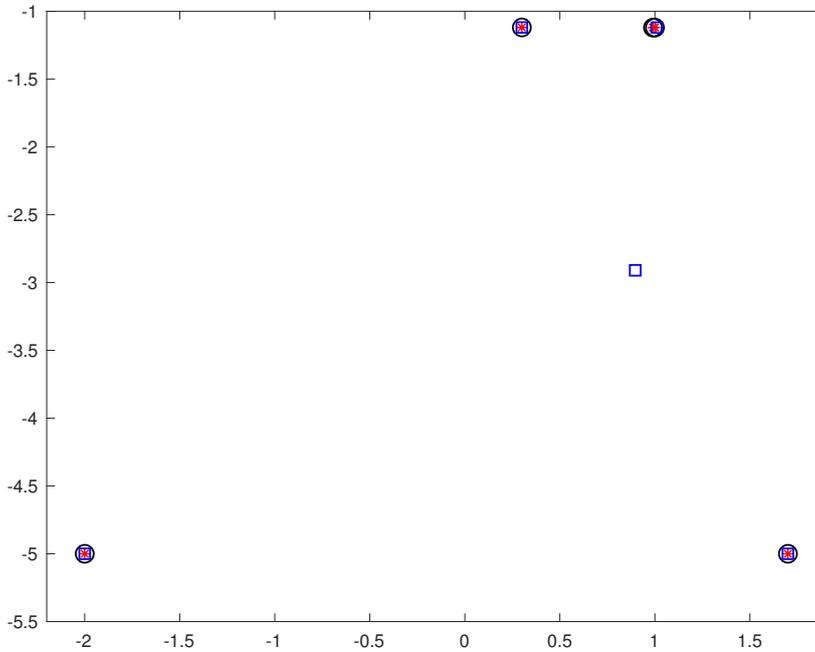


Figure 5.3: Parameter α_j (black circles) as well as the corresponding recovered parameters obtained by Algorithm 8 (red stars) and Algorithm 3 (blue squares). The x -axis shows the real part of the parameter α_j . The y -axis the imaginary part of α_j , $j = 1, \dots, 5$.

Algorithm 3 is not able to distinguish between the first two frequencies. We obtain $\tilde{\alpha}_2 = 1.0096 - 1.1198i$ and $\tilde{\alpha}_2 = 0.8967 - 2.90981i$. The corresponding parameters c_j , $j = 1, 2$ are adjusted accordingly and we obtain $\tilde{c}_1 = 0.5005 - 0.0003i$ and $\tilde{c}_2 = 2.8412 \cdot 10^{-06} + 2.1577i \cdot 10^{-05}$. This explains the large reconstruction errors for the values c_1 and c_2 . This is due to the fact, that the roots of the Prony polynomial $e^{\alpha_j h}$ for $j = 1, 2$ are close together. In comparison, the values $e^{\alpha_j u h}$ are much better separated

than the values $e^{\alpha_j h}$, $j = 1, \dots, 5$. Therefore, Algorithm 8 is able to recover the clustered frequencies with higher accuracy. Nevertheless, the performance of the recovery for both Algorithms is accurate as shown in Figure 5.4.

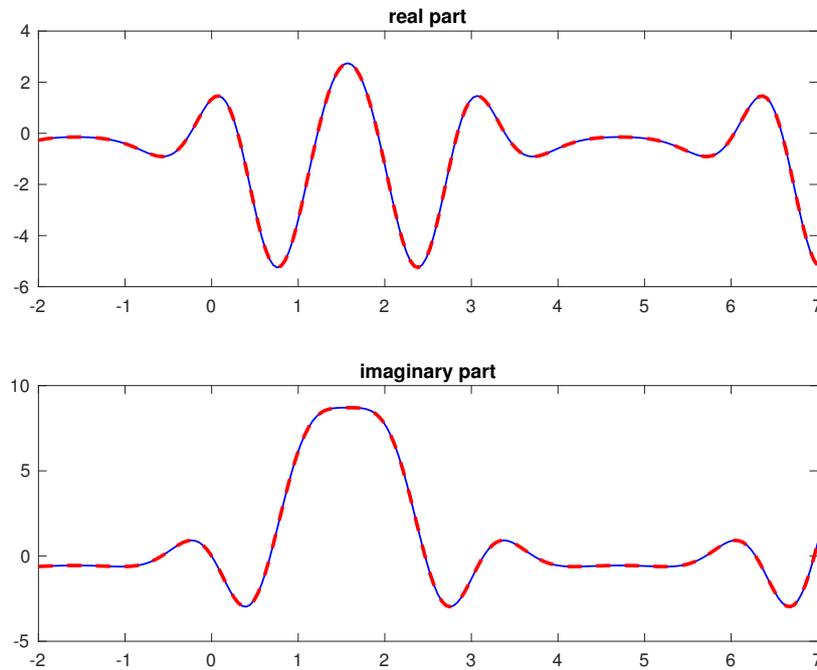


Figure 5.4: Recovery of the 5-sparse exponential sine expansion (5.49) with Algorithm 3 in red. Recovery obtain by Algorithm 8 shown in blue. Both reconstructions cannot be distinguished visually.

Finally, we also want to highlight that in this case our modified ESPRIT Algorithm 7 also yields high accuracy. Therefore, we oversample with $N = 14$. We choose $L = 8$ for the upper bound on the sparsity number $M = 5$ and tolerance $\varepsilon = 10^{-12}$. The reconstruction errors obtained by Algorithm 7 are shown in Table 5.8. We can see that the results obtained by Algorithm 8 are slightly better.

j	$ \alpha_j - \tilde{\alpha}_j $	$ c_j - \tilde{c}_j $
$j = 1$	0.0020	0.2965
$j = 2$	0.0049	0.2965
$j = 3$	$1.2158 \cdot 10^{-07}$	$2.8488 \cdot 10^{-09}$
$j = 4$	$3.5058 \cdot 10^{-10}$	$1.2134 \cdot 10^{-09}$
$j = 5$	$1.1696 \cdot 10^{-06}$	$5.0034 \cdot 10^{-06}$

Table 5.8: Error in the numerical evaluations of the parameters c_j and α_j , $j = 1, \dots, 5$ for the 5-sparse exponential sine expansion obtained using Algorithm 7.

◇

Example 5.31 We consider a M -sparse expansion into Gaussian chirps as in (5.48) with $M = 6$ and $\beta = 0.5$, i.e.,

$$f(x) = \sum_{j=1}^6 c_j e^{-0.5(x-\alpha_j)^2}$$

with the parameters c_j and α_j , $j = 1, \dots, 6$, as in Table 5.9.

parameters	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
c_j	-0.5	1	7	0.2	-3	8
Re α_j	0	0.0010	0.0022	0.00097	0	-0.00003
Im α_j	1.88π	1.8π	-1.5119	-1.5	-1.4889	1.7899π

Table 5.9: Parameters c_j and α_j for a 6-sparse expansion into Gaussian chirps as in (5.48).

We have 2 clusters, i.e., the frequencies α_1 , α_2 and α_6 as well as the frequencies α_3 , α_4 and α_5 are clustered together. The minimal separation distance is $q = 0.01$. We choose the starting point $h = \frac{1}{10}$ and the starting point $x_0 = 0$. For the sub-sampling parameters we choose $u = 5$ and $p = 3$. Then Theorem 5.16 implies that f can be recovered using the samples $f(x_0 + h(ku + \ell p))$ for $k = 0, \dots, 11$ and $\ell = 0, 1$.

The reconstruction errors obtained by Algorithm 8 are shown in Table 5.10. For better visualization we present the original frequencies α_j , $j = 1, \dots, 6$ as well as the parameters obtained by Algorithm 8 in Figure 5.5. We obtain a slightly greater error in the recovered parameters $\tilde{\alpha}_j$ for $j = 1, 2, 6$ and the corresponding parameters \tilde{c}_j , $j = 1, 2, 6$ are adjusted accordingly. Therefore, the recovered signal obtained by Algorithm 8 is still highly accurate. This shows that even in the case of multiple clusters containing more than 2 frequencies, the performance of Algorithm 8 is highly accurate.

j	$ \alpha_j - \tilde{\alpha}_j $	$ c_j - \tilde{c}_j $
$j = 1$	$1.2750 \cdot 10^{-4}$	0.0016
$j = 2$	0.0025	0.1282
$j = 3$	$2.4730 \cdot 10^{-7}$	$4.3621 \cdot 10^{-4}$
$j = 4$	$3.6358 \cdot 10^{-5}$	$8.2152 \cdot 10^{-5}$
$j = 5$	$6.3586 \cdot 10^{-7}$	$5.1094 \cdot 10^{-4}$
$j = 6$	$1.9650 \cdot 10^{-4}$	0.1294

Table 5.10: Error in the numerical evaluations of the parameters c_j and α_j , $j = 1, \dots, 6$ for the 6-sparse expansion into Gaussian chirps with clustered frequencies obtained using Algorithm 8.

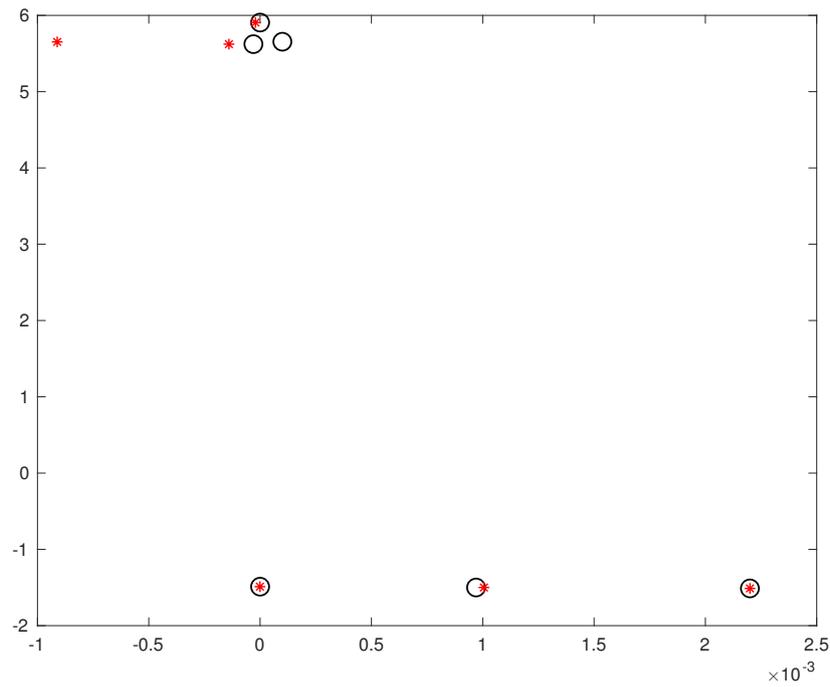


Figure 5.5: Parameter α_j (black circles) as well as the corresponding recovered parameters obtained by Algorithm 8 (red stars). The x -axis shows the real part of the parameter α_j . The y -axis shows the imaginary part of α_j , $j = 1, \dots, 6$.

◇

6 Conclusion and Outlook

In this thesis we have looked at the question of how we can recover structured functions using only a small set of functional values. Answering this question, we have deduced Prony based algorithms for the unique reconstruction of generalized exponential sums, generalized trigonometric sums and finite, sparse expansions into orthogonal polynomials based on the theory of the generalized Prony method derived in [PP13].

Based on the theory of one-parameter semigroups we derived a Prony method for the recovery of generalized exponential sums such as sparse cosine exponentials and generalized trigonometric functions. We showed that these structured functions can be understood as eigenfunctions of generalized shift operators as well as eigenfunctions of linear differential operators.

Additionally, we derived a Prony based method for the recovery of sparse expansions into orthogonal polynomials using generating functions and have embedded this reconstruction method into our theoretical concept of the generalized Prony method.

Furthermore, we showed that (stabilizing) numerical methods for the Prony method can be modified for the recovery of structured functions. In particular, we modified the ESPRIT method, introduced a sub-sampling Prony based recovery method for generalized exponential sums (with clustered frequencies) and analysed modifications of Prony's method for sparse approximation.

However, there exist further open problems for the recovery of structured functions. All of the derived methods have been studied in the one-dimensional case. Hence, an interesting object of investigation would be the two-dimensional or multivariate case. As we mentioned in Chapter 2, a lot of effort has been made to obtain a Prony-like method in several variables, see [PV20, KPRv16, CL18]. Therefore, it would be interesting to see if the results of this thesis can be extended to the multivariate case.

The numerical results in Section 5.5 show that the reconstruction obtained by Algorithm 8 is highly accurate in the case of clustered frequencies. The results in [CL20b] indicate that similar algorithms can be obtained for the reconstruction of generalized trigonometric expansions as well as generalized hyperbolic expansions. While there exists a variety of stability analysis for algorithms such as the ESPRIT method, see

e.g. [Aub16,PT13,LLF20] no performance results for the reconstruction method in 8 and similar algorithms have been derived. Hence, it would be of interest to obtain error analysis results for these algorithms.

Furthermore, the non-linear least squares problem considered in Section 5.4 is also closely related to structured low-rank approximation, see [Mar12,UM14]. Also, instead of the Euclidean norm, one can consider the maximum norm, see [Hac05] or the 1-norm, see [Skr17].

Additionally, one open question is, how functions in special smoothness spaces can be approximated by exponential sums (or generalized exponential sums), if the input data is known to be a sampling sequence of a function in a given smoothness space. Furthermore, it would be of interest to derive a convergence rate with respect to the number of terms M . The numerical results in [Hac05] show that the function $\frac{1}{x}$ can be approximated by an M -term exponential sum with an error of $O(\exp(c\sqrt{M}))$.

Finally, the results in [Pot17,PP19] indicate that we can hope for an exponential decay of the approximation error for a larger class of functions.

Bibliography

- [AAK71] V. M. Adamjan, D. Z. Arov, and M. G. Kreĭn. Analytic properties of the Schmidt pairs of a Hankel operator and the generalized Schur-Takagi problem. *Mat. Sbornik (N.S.)*, 86(128):34—75, 1971.
- [AS64] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 1964.
- [Aub16] C. Aubel. *Performance of super-resolution methods in parameter estimation and system identification*. PhD thesis, ETH Zürich, 2016. <https://www.research-collection.ethz.ch/handle/20.500.11850/156296>.
- [Bat15] D. Batenkov. Complete algebraic reconstruction of piecewise-smooth functions from Fourier data. *Math. Comput.*, 84(295):2329–2350, 2015.
- [Bit17] S. Bittens. Sparse FFT for Functions with Short Frequency Support. *Dolomites Res. Notes Approx.*, 10:43–55, 2017.
- [BKRE09] M. A. Bani-Hasan, Y. M. Kadah, M. E. M. Rasmy, and F. M. El-Hefnawi. Electrocardiogram signals identification for cardiac arrhythmias using Prony’s method and neural network. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1893–1896, 2009.
- [BM05] G. Beylkin and L. Monzón. On approximation of functions by exponential sums. *Applied and Computational Harmonic Analysis*, 19(1):17–48, 2005.
- [Bos06] S. Bosch. *Algebra (6. Aufl.)*. Springer, Berlin, 2006.
- [BP17] R. Beinert and G. Plonka. Sparse Phase Retrieval of One-Dimensional Signals by Prony’s Method. *Frontiers in Applied Mathematics and Statistics*, 3:5, 2017. <https://www.frontiersin.org/article/10.3389/fams.2017.00005>.
- [BZI19] S. Bittens, R. Zhang, and M. A. Iwen. A deterministic sparse FFT for functions with structured Fourier sparsity. *Adv. Comput. Math.*, 45(2):519–561, 2019.
- [CDS98] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998.

- [Ces12] C. Cesarano. Identities and generating functions on Chebyshev polynomials. *Georgian Mathematical Journal*, 19(3):427–440, 2012.
- [CL18] A. Cuyt and W.-s. Lee. Multivariate Exponential Analysis from the Minimal Number of Samples. *Advances in Computational Mathematics*, 44(4):987 — 1002, 2018.
- [CL20a] A. Cuyt and W.-s. Lee. How to get high resolution results from sparse and coarsely sampled data. *Applied and Computational Harmonic Analysis*, 48(3):1066–1087, 2020.
- [CL20b] A. Cuyt and W.-s. Lee. Parametric spectral analysis: scale and shift, 2020. arXiv:2008.02125.
- [CRT06] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489 – 509, 2006.
- [Cuy20] A. Cuyt. What Do Sparse Interpolation, Padé Approximation, Gaussian Quadrature and Tensor Decomposition Have in Common? In *Proceedings of the 45th International Symposium on Symbolic and Algebraic Computation, ISSAC '20*, page 12, New York, NY, USA, 2020. Association for Computing Machinery.
- [Dem00] J. Demmel. Accurate Singular Value Decompositions of Structured Matrices. *SIAM Journal on Matrix Analysis and Applications*, 21(2):562–580, 2000.
- [DGE⁺99] J. Demmel, M. Gu, S. Eisenstat, I. Slapničar, K. Veselić, and Z. Drmač. Computing the singular value decomposition with high relative accuracy. *Linear Algebra and its Applications*, 299(1):21–80, 1999.
- [DL00] G. Dattoli and D. Levi. Exponential operators and generalized difference equations. *Il Nuovo Cimento B*, 115(b)(6):653–662, 2000.
- [Don06] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.
- [DOTV97] G. Dattoli, P. Ottaviani, A. Torre, and L. Vázquez. Evolution operator equations: Integration with algebraic and finite difference methods. Applications to physical problems in classical and quantum mechanics and quantum field theory. *La Rivista del Nuovo Cimento(1978-1999)*, 20(2):3–133, 1997.
- [dP95] Gaspard Riche de Prony. Essai experimental et analytique sur les lois de la delatabilite des fluides et sur celles de la force expansive del la vapeur de l'eau el de l'alkool differentes temperatures. *l'Ecole Polytech.*, 1(2):24–76, 1795.

- [Drm15] Z. Drmač. SVD of Hankel matrices in Vandermonde-Cauchy product form. *Electronic transactions on numerical analysis*, 44:593–623, 2015.
- [DVB07] P. L. Dragotti, M. Vetterli, and T. Blu. Sampling Moments and Reconstructing Signals of Finite Rate of Innovation: Shannon Meets Strang–Fix. *IEEE Transactions on Signal Processing*, 55(5):1741–1757, 2007.
- [EN00] K.J. Engel and R Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Graduate Texts in Mathematics. Springer New York, 2000.
- [FMP12] F. Filbir, H.N. Mhaskar, and J. Prestin. On the Problem of Parameter Estimation in Exponential Sums. *Constructive Approximation*, 35:323–343, 2012.
- [FP68] L. Fox and I.B. Parker. *Chebyshev Polynomials in Numerical Analysis*. Oxford mathematical handbooks. Oxford University Press, London-New York-Toronto, Ont., 1968.
- [Fuh12] P.A. Fuhrmann. *A Polynomial Approach to Linear Algebra*. Springer, New York, second edition, 2012.
- [Hac05] W. Hackbusch. Entwicklungen nach Exponentialsummen. Technical report, MPI für Mathematik in den Naturwissenschaften, 2005. <https://www.mis.mpg.de/de/publications/andere-reihen/tr/report-0405.html>.
- [Heu06] H. Heuser. *Funktionalanalysis: Theorie und Anwendung*. Mathematische Leitfäden. Vieweg+Teubner Verlag, Wiesbaden, 2006.
- [HJ13] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, second edition, 2013.
- [HS90] Y. Hua and T. K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(5):814 – 824, 1990.
- [HSL⁺98] N. Huang, Z. Shen, S. Long, M. L. C. Wu, H. Shih, Q. Zheng, N.-C. Yen, C.-C. Tung, and H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454:903–995, 1998.
- [IKY18] E. Imamoglu, E. L. Kaltofen, and Z. Yang. Sparse Polynomial Interpolation With Arbitrary Orthogonal Polynomial Bases. In *Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation, ISSAC '18*, pages 223 — 230, New York, NY, USA, 2018. Association for Computing Machinery.

- [JHLC16] J.-W. Jung, J.-W. Hong, H.-K. Lee, and K. Choi. Estimation of viscoelastic parameters in Prony series from shear wave propagation. *Journal of Applied Physics*, 119:234701, 2016. <https://doi.org/10.1063/1.4953674>.
- [KL03] E. Kaltofen and W.-s. Lee. Early Termination in Sparse Interpolation Algorithms. *J. Symb. Comput.*, 36(3):365 – 400, 2003.
- [KLS10] R. Koekoek, P. Lesky, and R. Swarttouw. *Hypergeometric Orthogonal Polynomials and Their q -Analogues*. Springer-Verlag, Berlin Heidelberg, 2010.
- [KN21] S. Kunis and D. Nagel. On the condition number of Vandermonde matrices with pairs of nearly-colliding nodes. *Numerical Algorithms*, 87(1):473 – 496, 2021. <https://doi.org/10.1007/s11075-020-00974-x>.
- [KP21] I. Keller and G. Plonka. Modifications of Prony’s Method for the Recovery and Sparse Approximation with Generalized Exponential Sums. In Gregory E. Fasshauer, Marian Neamtu, and Larry L. Schumaker, editors, *Approximation Theory XVI*, pages 123 –152, Cham, 2021. Springer International Publishing.
- [KPP21] H. Knirsch, M. Petz, and G. Plonka. Optimal rank-1 Hankel approximation of matrices: Frobenius norm and spectral norm and Cadzow’s algorithm. *Linear Algebra and its Applications*, 629:1– 39, 2021.
- [KPRv16] S. Kunis, T. Peter, T. Römer, and U. von der Ohe. A multivariate generalization of Prony’s method. *Linear Algebra and its Applications*, 490:31 – 47, 2016.
- [KPS19] I. Keller, G. Plonka, and K. Stampfer. Reconstruction of Non-Stationary Signals by the Generalized Prony Method. *Proc. Appl. Math. Mech.*, 19:358 – 359, 2019.
- [KRv20] S. Kunis, T. Römer, and U. von der Ohe. Learning algebraic decompositions using Prony structures. *Advances in Applied Mathematics*, 118:102044, 2020.
- [LC56] C. Lanczos and Karreman Mathematics Research Collection. *Applied Analysis*. Mathematics series. Prentice-Hall, Inc., Englewood Cliffs, N. J., 1956.
- [Les62] P. Lesky. Die Charakterisierung der klassischen orthogonalen Polynome durch Sturm-Liouvillesche Differentialgleichungen. *Archive for Rational Mechanics and Analysis*, 10(1):341 – 351, 1962.
- [LLF20] W. Li, W. Liao, and A. Fannjiang. Super-Resolution Limit of the ESPRIT Algorithm. *IEEE Transactions on Information Theory*, 66(7):4593–4608, 2020.
- [LS95] Y. N. Lakshman and B. D. Saunders. Sparse Polynomial Interpolation in Nonstandard Bases. *SIAM Journal on Computing*, 24(2):387 – 397, 1995.

- [LVT91] F. Li, R. J. Vaccaro, and D. W. Tufts. Performance analysis of the state-space realization (TAM) and ESPRIT algorithms for DOA estimation. *IEEE Transactions on Antennas and Propagation*, 39(3):418 – 423, 1991.
- [Mar86] V. A. Marchenko. *Sturm-Liouville operators and applications*. Operator Theory: Advances and Applications. Birkhäuser Basel, 1986.
- [Mar12] I. Markovsky. *Low Rank Approximation: Algorithms, Implementation, Applications*. Springer, London, 2012.
- [McB71] E. B. McBride. *Obtaining Generating Functions*. Springer Tracts in Natural Philosophy. Springer-Verlag, Berlin Heidelberg, 1971.
- [Mil15] E. K. Miller. Using Prony’s Method To Synthesize Discrete Arrays for Prescribed Source Distributions and Exponentiated Patterns. *IEEE Antennas and Propagation Magazine*, 57(1):147–163, 2015.
- [Mor78] J. Moré. The Levenberg-Marquardt algorithm: Implementation and theory. In G.A. Watson, editor, *Numerical Analysis*, volume 630 of *Lecture Notes in Mathematics*, pages 105–116. Springer, Berlin, Heidelberg, 1978.
- [MP00] H. Mhaskar and J. Prestin. On the detection of singularities of a periodic function. *Adv. Comput. Math.*, 12(2-3):95–131, 2000.
- [MP05] H. Mhaskar and J. Prestin. On Local Smoothness Classes of Periodic Functions. *Journal of Fourier Analysis and Applications*, 11(3):353–373, 2005.
- [Nat95] B. K. Natarajan. Sparse Approximate Solutions to Linear Systems. *SIAM J. Comput.*, 24(2):227–234, 1995.
- [OS91] M. R. Osborne and G. Smyth. A Modified Prony Algorithm for Fitting Functions Defined by Difference Equations. *SIAM J. Sci. Comput.*, 12:362–382, 1991.
- [Osb75] M. R. Osborne. Some Special Nonlinear Least Squares Problems. *SIAM Journal on Numerical Analysis*, 12(4):571–592, 1975.
- [Pan16] V. Y. Pan. How Bad Are Vandermonde Matrices? *SIAM Journal on Matrix Analysis and Applications*, 37(2):676–694, 2016.
- [Paz83] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Number 44 in Applied Mathematical Sciences. Springer-Verlag, New York, 1983.
- [Pet14] T. Peter. *Generalized Prony Method*. PhD thesis, University of Göttingen, 2014. Der Andere Verlag.

- [Pis73] V. F. Pisarenko. The Retrieval of Harmonics from a Covariance Function. *Geophysical Journal International*, 33(3):347–366, 1973.
- [Pot17] V. Pototskaia. *Application of AAK Theory for Sparse Approximation*. PhD thesis, University of Göttingen, 2017. <http://hdl.handle.net/11858/00-1735-0000-0023-3F4B-1>.
- [PP13] T. Peter and G. Plonka. A generalized Prony method for reconstruction of sparse sums of eigenfunctions of linear operators. *Inverse Problems*, 29(2):025001, 2013.
- [PP19] G. Plonka and V. Pototskaia. Computation of Adaptive Fourier Series by Sparse Approximation of Exponential Sums. *Journal of Fourier Analysis and Applications*, 25(4):1580–1608, 2019.
- [PPD21] M. Petz, G. Plonka, and N. Derevianko. Exact Reconstruction of Sparse Non-Harmonic Signals from Fourier Coefficients. *Sampling Theory, Signal Processing, and Data Analysis*, 19(1):7, 2021. <https://doi.org/10.1007/s43670-021-00007-1>.
- [PPR13] T. Peter, G. Plonka, and D. Roşca. Representation of sparse Legendre Expansions. *Journal of Symbolic Computation*, 50:159–169, 2013.
- [PPST19] G. Plonka, D. Potts, G. Steidl, and M. Tasche. *Numerical Fourier Analysis*. Birkhäuser Basel, 2019.
- [Pry93] J. D. Pryce. *Numerical solution of Sturm-Liouville problems*. Monographs on Numerical Analysis. The Clarendon Press, Oxford University Press, New York, 1993. Oxford Science Publications.
- [PSK19] G. Plonka, K. Stampfer, and I. Keller. Reconstruction of stationary and non-stationary signals by the generalized Prony method. *Analysis and Applications*, 17(02):179 – 210, 2019.
- [PT10] D. Potts and M. Tasche. Parameter estimation for exponential sums by approximate Prony method. *Signal Processing*, 90(5):1631–1642, 2010.
- [PT11] D. Potts and M. Tasche. Nonlinear approximation by sums of nonincreasing exponentials. *Applicable Analysis*, 90(3-4):609–626, 2011.
- [PT13] D. Potts and M. Tasche. Parameter estimation for nonincreasing exponential sums by Prony-like methods. *Linear Algebra and its Applications*, 439(4):1024 – 1039, 2013. 17th Conference of the International Linear Algebra Society, Braunschweig, Germany, August 2011.
- [PT14a] G. Plonka and M. Tasche. Prony methods for recovery of structured functions. *GAMM-Mitteilungen*, 37(2):239–258, 2014.

- [PT14b] D. Potts and M. Tasche. Sparse polynomial interpolation in Chebyshev bases. *Linear Algebra and its Applications*, 441:61–87, 2014.
- [PT16] D. Potts and M. Tasche. Reconstruction of sparse Legendre and Gegenbauer expansions. *BIT Numerical Mathematics*, 56(56):1019–1043, 2016.
- [PT17] D. Potts and M. Tasche. Error Estimates for the ESPRIT Algorithm. In *Large Truncated Toeplitz Matrices, Toeplitz Operators, and Related Topics: The Albrecht Böttcher Anniversary Volume*, pages 621–648. Springer International Publishing, Cham, 2017.
- [PV20] J. Prestin and H. Veselovska. Prony-Type Polynomials and Their Common Zeros. *Frontiers in Applied Mathematics and Statistics*, 6:16, 2020. <https://www.frontiersin.org/article/10.3389/fams.2020.00016>.
- [PW13] G. Plonka and M. Wischerhoff. How many Fourier samples are needed for real function reconstruction? *Journal of Applied Mathematics and Computing*, 42:117 – 137, 2013.
- [QI99] R. C. Qiu and I-Tai Lu. Multipath resolving with frequency dependence for wide-band wireless channel modeling. *IEEE Transactions on Vehicular Technology*, 48(1):273–285, 1999.
- [RH89] B. D. Rao and K. V. S. Hari. Performance analysis of ESPRIT and TAM in determining the direction of arrival of plane waves in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(12):1990–1995, 1989.
- [RK89] R. Roy and T. Kailath. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(7):984–995, 1989.
- [SB02] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. Texts in Applied Mathematics. Springer, New York, 2002.
- [Sch02] M. Schatzman. *Numerical Analysis: A Mathematical Introduction*. Clarendon Press, 2002.
- [Skr17] M.R. Skrzipek. Signal recovery by discrete approximation and a Prony-like method. *Journal of Computational and Applied Mathematics*, 326:193–203, 2017.
- [SM97] P. Stoica and R. Moses. *Introduction to Spectral Analysis*. Prentice Hall, Upper Saddle River, 1997.
- [SP20] K. Stampfer and G. Plonka. The Generalized Operator Based Prony Method. *Constr. Approx.*, 52:247–282, 2020.

- [SS90] G. W. Stewart and J. Sun. *Matrix Perturbation Theory*. Academic Press, 1990.
- [SS91] P. Stoica and T. Soderstrom. Statistical analysis of MUSIC and subspace rotation estimates of sinusoidal frequencies. *IEEE Transactions on Signal Processing*, 39(8):1836–1847, 1991.
- [Sta18] K. Stampfer. *The Generalized Operator Based Prony Method*. PhD thesis, University of Göttingen, 2018. <http://hdl.handle.net/11858/00-1735-0000-002E-E631-3>.
- [Ste84] S. Steinberg. Lie series and nonlinear ordinary differential equations. *Journal of Mathematical Analysis and Applications*, 101(1):39 – 63, 1984.
- [Sze75] G. Szegő. *Orthogonal Polynomials*. American Math. Soc: Colloquium publ. American Mathematical Society, 4 edition, 1975.
- [TSSP91] D. J. Trudnowski, J. R. Smith, T. A. Short, and D. A. Pierre. An application of Prony methods in PSS design for multimachine systems. *IEEE Transactions on Power Systems*, 6(1):118–126, 1991.
- [Tyr94] E. E. Tyrtyshnikov. How bad are Hankel matrices? *Numerische Mathematik*, 67(2):261–269, 1994.
- [UM14] K. Usevich and I. Markovsky. Variable projection for affinely structured low-rank approximation in weighted 2-norms. *Journal of Computational and Applied Mathematics*, 272:430–448, 2014.
- [VMB02] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, 50(6):1417–1428, 2002.
- [Wer11] D. Werner. *Funktionalanalysis*. Springer-Lehrbuch. Springer-Verlag Berlin Heidelberg, 7 edition, 2011.
- [Wil59] J.H. Wilkinson. The evaluation of the zeros of ill-conditioned polynomials. Part I. *Numerische Mathematik*, 1:150–160, 1959.
- [Wil67] R. M. Wilcox. Exponential Operators and Parameter Differentiation in Quantum Physics. *Journal of Mathematical Physics*, 8(4):962–982, 1967.
- [WM63] L. Weiss and R. N. McDonough. Prony’s Method, Z-Transforms, and Padé Approximation. *SIAM Review*, 5(2):145–149, 1963.
- [YFG⁺04] T. Yanai, G. Fann, Z. Gan, R. Harrison, and G. Beylkin. Multiresolution quantum chemistry in multiwavelet bases: Analytic derivatives for Hartree-Fock and density functional theory. *The Journal of chemical physics*, 121:2866–76, 2004.

- [ZP19] R. Zhang and G. Plonka. Optimal approximation with exponential sums by maximum likelihood modification of Prony's method. *Adv. Comput Math*, 45(3):1657 – 1687, 2019.