The Generalized Operator Based Prony Method

Kilian Stampfer*   Gerlind Plonka†

16th January 2020

Abstract. The generalized Prony method introduced in [19] is a reconstruction technique for a large variety of sparse signal models that can be represented as sparse expansions into eigenfunctions of a linear operator A. However, this procedure requires the evaluation of higher powers of the linear operator A that are often expensive to provide.

In this paper we propose two important extensions of the generalized Prony method that simplify the acquisition of the needed samples essentially and at the same time can improve the numerical stability of the method. The first extension regards the change of operators from A to $\varphi(A)$, where φ is a suitable operator valued mapping, such that A and $\varphi(A)$ possess the same set of eigenfunctions. The goal is now to choose φ such that the powers of $\varphi(A)$ are much simpler to evaluate than the powers of A. The second extension concerns the choice of the sampling functionals. We show, how new sets of different sampling functionals F_k can be applied with the goal to reduce the needed number of powers of the operator A (resp. $\varphi(A)$) in the sampling scheme and to simplify the acquisition process for the recovery method.

Key words: Generalized Prony method, exponential operators, sparse expansions into eigenfunctions of linear operators, parameter identification, generalized sampling.

Mathematics Subject Classification: 41A30, 37M99, 65F15.

1 Introduction

The recovery of signals which can be represented or approximated by finite expansions into signal atoms is a task regularly encountered in a variety of fields such as signal processing, biology, and engineering. These "signal atoms" have a fixed structure and can be identified by a small number of real or complex parameters. Therefore, sparse expansions into these signal atoms often permit an arbitrarily high resolution in contrast to classical sampling schemes based on Hilbert space techniques. At the same time these signal models frequently allow a better physical interpretation. The most prominent and well-studied signal model of this kind is a

---

*Institute for Mathematical Stochastics, Göttingen University, Lotzestr. 16-18, 37083 Göttingen, Germany. Email: k.stampfer@math.uni-goettingen.de

†Institute for Numerical and Applied Mathematics, Göttingen University, Lotzestr. 16-18, 37083 Göttingen, Germany. Email: plonka@math.uni-goettingen.de

sparse expansion into complex exponentials, i.e.,

$$f(x) := \sum_{j=1}^{M} c_j \exp(T_j x) = \sum_{j=1}^{M} c_j z_j^x, \tag{1.1}$$

with pairwise different $z_j := \exp(T_j)$ and with parameters $c_j \in \mathbb{C} \setminus \{0\}$ and $T_j \in \mathbb{C}$. Using the classical Prony method, the parameters c_j and z_j can be computed from the $2M$ equidistant samples $f(\ell)$, $\ell = 0, \dots, 2M - 1$, see e.g. [27] or [25], Chapter 10, and the references therein. Observe that, in order to extract T_j from z_j in a unique way, we need to restrict $\operatorname{Im} T_j$ to an interval of length 2π.

In practical applications, we have to take special care of the numerical instabilities that can occur using Prony's method. There have been many attempts to provide improved numerical algorithms, including the Pisarenko method [22], MUSIC [31], ESPRIT [15], Matrix Pencil Methods [14] and the approximate Prony method [29]. Furthermore, to ensure the consistency in case of noisy measurements, modifications of Prony's method have been proposed, see e.g. [8,16,18,35]. The interest in Prony-like methods has been strongly increased during the last years, also because of their utilization for the recovery of signals of finite rate of innovation, see e.g. [7,12,33,34]. In particular, the close connection between the exponential sum in (1.1) and the expansion into shifted Diracs

$$s(t) = \sum_{j=1}^{M} c_j \, \delta(t - t_j) \tag{1.2}$$

with $c_j \in \mathbb{C} \setminus \{0\}$ and $t_j \in \mathbb{R}$ is extensively used. Indeed the Fourier transform of $s(t)$ is of the form (1.1), where $T_j = \mathrm{i} t_j$, and thus $s(t)$ can be reconstructed from only $2M$ of its Fourier samples, see also [21,28]. Moreover, Prony's method and its generalizations provide new approaches for nonlinear sparse approximation of smooth functions, and there are close relations to optimal approximation of functions in Hardy spaces [1,6,23,24].

An essential extension of the classical Prony method has been proposed in [19], where the recovery of expansions into exponentials has been generalized to the recovery of expansions into eigenfunctions of linear operators.

Let us assume that $A : V \to V$ is a linear operator on a normed vector space V, and let $\sigma(A)$ be a subset of the point spectrum of A that contains pairwise different eigenvalues. Further, we consider the corresponding set of eigenfunctions v_λ of A such that v_λ can be uniquely identified by $\lambda \in \sigma(A)$. In other words, the eigenspace to λ is fixed as a one-dimensional space. Then, the generalized Prony method in [19] allows the reconstruction of expansions f of the form

$$f = \sum_{j=1}^{M} c_j \, v_{\lambda_j} \tag{1.3}$$

with $c_j \in \mathbb{C} \setminus \{0\}$ and with pairwise distinct $\lambda_j \in \sigma(A)$. According to [19], the eigenvalues λ_j belonging to the "active" eigenfunctions v_{λ_j} and the coefficients c_j, $j = 1, \dots, M$ can be uniquely recovered from the (complex) values $F(A^\ell f)$, $\ell = 0, \dots, 2M - 1$, where $F : V \to \mathbb{C}$ is a functional that can be chosen arbitrarily up to the condition $Fv_\lambda \neq 0$ for all $\lambda \in \sigma(A)$. The expansion into exponentials in (1.1) can be seen as a special case of (1.3) if we take $V = C(\mathbb{R})$, $A = S_1$ with the shift

operator given by $S_1 f := f(\cdot + 1)$, and the point evaluation functional $Ff := f(0)$. Indeed, the exponentials $\exp(T_j \cdot)$ are eigenfunctions of S_1 to the eigenvalues $\exp(T_j)$ which are pairwise different for $T_j \in \mathbb{R} + \mathrm{i}[-\pi, \pi)$. The needed samples $F(A^\ell f)$ are in this case of the form $F(A^\ell f) = F(S_1^\ell f) = F(f(\cdot + \ell)) = f(\ell)$.

There have been other attempts to generalize the idea of Prony's method to different expansions, including sparse polynomials [4], piecewise sinusoidal signals [5], sparse expansions into Legendre polynomials [20] or Chebyshev polynomials [30] and into Lorentzians [2]. All these expansions can be also recovered directly using the approach in [19]. An extension of the generalized Prony method to the multivariate case based on Artinian Gorenstein algebras and the flat extension principle has been given by Mourrain [17].

However, the generalized Prony method is not always simple to apply since it requires the computation of higher powers of the operator A in order to achieve the needed sample values $F(A^\ell f)$ for the reconstruction procedure. While for shift operators these samples are easy to acquire, the problem is much more delicate for differential or integral operators of higher order. Indeed, the shift operator S_τ, with $S_\tau f := f(\cdot + \tau)$, and its generalizations play a special role, since the power S_τ^ℓ is equivalent to $S_{\ell\tau}$, i.e., to a simple shift operator with shift length $\ell\tau$. Expansions into eigenfunctions of generalized shift operators are therefore of special interest, since they can be recovered just by suitable function samples, see [26].

In this paper, we reconsider the generalized Prony method in [19] in more detail and particularly study two extensions that provide us more freedom in data acquisition for the recovery of expansions of the form (1.3).

The first extension is based on the observation that for a given linear operator A there is often a different linear operator B that possesses the same eigenfunctions to different eigenvalues. For example, the exponential function $\exp(Tx)$ is an eigenfunction of the shift operator S_τ to the eigenvalue $\mathrm{e}^{\tau T}$, but at the same time also an eigenfunction of the differential operator $\frac{\mathrm{d}}{\mathrm{d}x}$ to the eigenvalue T. Thus, we need to understand, how this observation can help us to solve the signal recovery problem, and in particular, for a given linear operator A, how to find a different linear operator B with the same eigenfunctions that may be easier to apply.

The second extension directly aims at generalizing the sampling functional F. While it is appealing that the $2M$ parameters of the signal model in (1.1) and (1.3) can be theoretically obtained from only $2M$ samples, in many applications we are faced with a parameter identification problem, where a large number of noisy samples is given, and we need to identify the parameters in a stable manner. Therefore, we go away from sampling schemes that use a minimal number of sampling values being ordered in matrices with Hankel structure. We will show that there is much more freedom to choose a set of different sampling functionals F_k, where each sampling functional leads to a linear equation providing one condition for the vector of coefficients of the Prony polynomial. Our approach also covers previous ideas to identify the frequency parameters T_j of the exponential sum in (1.1) using equispaced sampling sequences with different sampling sizes simultaneously, see [9].

Our ideas to provide simple acquisition schemes to recover expansions into eigenfunctions of linear operators also open the way for new approaches for sparse nonlinear approximation of (non-stationary) signals and images.

The paper is organized as follows. In Section 2 we reconsider the Prony method for exponential sums. We first show, how it can be understood as a method to recover a sparse expansion into eigenfunctions of the shift operator S_τ on the one hand and

of the differential operator $\frac{\mathrm{d}}{\mathrm{d}x}$ on the other hand. In Section 2.3 we employ an exponential operator notation to show how the two operators S_τ and $\frac{\mathrm{d}}{\mathrm{d}x}$ are related to each other. Further, we introduce the idea, how the sampling scheme can be generalized using a set of different sampling functionals F_k instead of $F(A^k)$.

Section 3 is devoted to the new generalized operator based Prony method (GOP). We start with recalling the generalized Prony method from [19] and transfer it into our new notation. Sections 3.2 and 3.3 are concerned with the two new extensions, first the change of operators from A to $\varphi(A)$, where φ is an analytic function, and second the generalization of the sampling scheme. In particular, we introduce admissible sets of sampling functionals F_k that allow a unique reconstruction of expansions of the form (1.3). In Section 3.4 we give a detailed example, where GOP is applied to sparse cosine expansions.

In Section 4, we discuss the application of GOP for the recovery of eigenfunctions of differential operators. We show that special linear differential operators of first and second order lead by a transfer from the operator A to $\varphi(A)$ (with an exponential map φ) to generalized shift operators whose powers can be simply evaluated in sampling schemes.

Section 5 is devoted to a further investigation of the second extension, the generalized sampling. We embed the functions f in (1.3) into a suitable Hilbert space and employ a dual approach for the sampling scheme. Then, our sampling functionals $F_k : V \to \mathbb{C}$ can be written as inner products with special kernels ϕ_k as Riesz representers, i.e., $F_k(f) = \langle f, \phi_k \rangle$. Therefore the application of F_k to powers $A^\ell f$ or $(\varphi(A))^\ell f$ to obtain the required sampling values can be rewritten by applying powers of the adjoint operator A^* to the kernel ϕ_k. In this way, we are able to find admissible sampling schemes for the recovery of expansions into eigenfunctions of differential operators in terms of moments. We demonstrate the principle for the recovery of exponential sums and for the recovery of sparse Legendre expansions using only moments of f.

The considerations in this paper provide the starting point for further studies that focus on the improvement of the numerical stability of the generalized Prony method. But this problem is beyond the scope of this paper and will be the further investigated.

2 An introductory example: Revisiting Prony's method using shift and differential operator

2.1 Prony's method based on the shift operator

The classical Prony method is a way to reconstruct the parameters $c_j \in \mathbb{C} \setminus \{0\}$, $T_j \in \mathbb{C}$, $j = 1, \ldots, M$, of the weighted sum of exponentials

$$f(x) = \sum_{j=1}^{M} c_j \exp(T_j x). \tag{2.1}$$

Using equidistant sample values $f(k)$, $k = 0, \ldots, 2M - 1$, exact recovery is possible if $T_j \in \mathbb{R} + \mathrm{i}[-\pi, \pi)$, see e.g. [27]. Usually, we assume that there is an a priori known bound C such that $\operatorname{Im} T_j \in [-C\pi, C\pi)$, and the parameters T_j can still be recovered using a rescaling argument and taking sampling values $f(kh)$ with $h \leq 1/C$ instead

of $h = 1$. With

$$\mathcal{M} := \left\{ \sum_{j=1}^{M} c_j \, e^{T_j x} \, : \, M < \infty, c_j \in \mathbb{C} \setminus \{0\}, T_j \in \mathbb{R} + \mathrm{i}[-C\pi, \, C\pi), \forall j \neq i : T_j \neq T_i, \right\}$$

we denote the model class of all finite linear combinations of complex exponentials that can be recovered by Prony's method.

Recalling the ideas in [19, 26], we can reinterpret and generalize the method using a shift operator. The exponential sum in (2.1) can be understood as an expansion into M eigenfunctions of the shift operator $S_\tau : C(\mathbb{R}) \to C(\mathbb{R})$ for some $\tau \neq 0$ with $S_\tau f(x) := f(x + \tau)$. More precisely, we observe that

$$(S_\tau \exp(T_j \cdot))(x) = \exp(T_j(x + \tau)) = \exp(T_j \tau) \, \exp(T_j x),$$

i.e., the exponentials $\exp(T_j x)$ occurring in (2.1) are eigenfunctions of S_τ to the eigenvalues $\exp(T_j \tau)$. This implies

$$(S_\tau - \exp(T_j \tau) I) \exp(T_j \cdot) = 0,$$

where I denotes the identity operator. We define the Prony polynomial

$$P(z) = P_\tau(z) := \prod_{j=1}^{M} (z - \exp(T_j \tau))$$

with the monomial representation

$$P(z) = \sum_{\ell=0}^{M} p_\ell z^\ell = z^M + \sum_{\ell=0}^{M-1} p_\ell \, z^\ell$$

and observe for f in (2.1) that

$$\begin{aligned}
P(S_\tau)f &= \sum_{\ell=0}^{M} p_\ell S_\tau^\ell f = \sum_{\ell=0}^{M} p_\ell S_\tau^\ell \sum_{j=1}^{M} c_j \exp(T_j \cdot) \\
&= \sum_{j=1}^{M} c_j \, \exp(T_j \cdot) \Big(\sum_{\ell=0}^{M} p_\ell \, \exp(T_j \tau \ell) \Big) \\
&= \sum_{j=1}^{M} c_j \, \exp(T_j \cdot) \, P(\exp(T_j \tau)) = 0.
\end{aligned}$$

Thus, f solves the difference equation $P(S_\tau)f = 0$. In particular, we also have

$$S_\tau^k P(S_\tau)f = P(S_\tau) \, S_\tau^k \, f = \sum_{\ell=0}^{M} p_\ell \, S_\tau^{\ell+k} f = 0, \qquad k \in \mathbb{Z}.$$

We fix an arbitrary value $x_0 \in \mathbb{R}$ and employ the point evaluation functional F_{x_0} with $F_{x_0} f := f(x_0)$ to compute the samples $F_{x_0} S_\tau^k f = f(x_0 + \tau k)$, $k = 0, \ldots, 2M - 1$. Then we obtain the homogeneous equation system

$$F_{x_0}(S_\tau^k P(S_\tau)f) = \sum_{\ell=0}^{M} p_\ell f(x_0 + \tau(k + \ell)) = 0, \quad k = 0, \ldots, M - 1, \tag{2.2}$$

for the vector $\mathbf{p} = (p_0, \ldots, p_M)^T$ of coefficients of $P(z)$. For $f \in \mathcal{M}$ and fixed $\tau < C^{-1}$ the arising coefficient matrix $(f(x_0 + \tau(k + \ell)))_{k=0,\ell=0}^{M-1,M} \in \mathbb{C}^{M \times M+1}$ is of Hankel structure and has full rank M, see [19, 27]. Thus, $\mathbf{p}$ is uniquely defined with $p_M = 1$, and we can extract the zeros $\exp(T_j\tau)$ of the polynomial $P(z)$ and compute T_j, $j = 1, \ldots, M$. Finally, the vector of coefficients $\mathbf{c} = (c_j)_{j=1}^M$ in (2.1) can be computed as a least squares solution of the Vandermonde system

$$\mathbf{V}_{2M,M}\,\mathbf{c} = (S_\tau^k f(x_0))_{k=0}^{2M-1} = (f(x_0 + \tau k))_{k=0}^{2M-1}$$

with $\mathbf{V}_{2M,M} := (\exp(T_j(x_0 + \tau k)))_{k=0,j=1}^{2M-1,M}$.

2.2 Prony's method based on the differential operator

We now present a different viewpoint and interpret $f(x)$ in (2.1) as the solution of a linear ordinary differential equation of order M. In fact, the functions $\exp(T_j x)$ are also eigenfunctions of the first derivative operator $\frac{\mathrm{d}}{\mathrm{d}x} : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$, i.e.,

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}\, \exp(T_j\cdot) \right)(x) = T_j\, \exp(T_j x),$$

and thus

$$\left(\frac{\mathrm{d}}{\mathrm{d}x} - T_j I \right) \exp(T_j\cdot) = 0$$

for all $T_j \in \mathbb{C}$, where I denotes the identity operator. We can now proceed similarly as before just by replacing the shift operator with the differential operator. Employing the eigenvalues T_j, we define the characteristic polynomial

$$\tilde{P}(z) := \prod_{j=1}^M (z - T_j) = \sum_{\ell=0}^M \tilde{p}_\ell\, z^\ell = z^M + \sum_{\ell=0}^{M-1} \tilde{p}_\ell\, z^\ell.$$

We apply the corresponding linear differential operator $\tilde{P}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)$ of order M to the function f in (2.1) and find

$$
\begin{aligned}
\tilde{P}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) f &= \prod_{j=1}^M \left(\frac{\mathrm{d}}{\mathrm{d}x} - T_j I\right) f = \left(\sum_{\ell=0}^M \tilde{p}_\ell \frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}\right) f \\
&= \sum_{\ell=0}^M \tilde{p}_\ell \sum_{j=1}^M c_j\, T_j^\ell\, \exp(T_j\cdot) = \sum_{j=1}^M c_j\, \exp(T_j\cdot)\left(\sum_{\ell=0}^M \tilde{p}_\ell\, T_j^\ell\right) \\
&= \sum_{j=1}^M c_j\, \exp(T_j\cdot)\, \tilde{P}(T_j) = 0,
\end{aligned}
$$

i.e., f in (2.1) solves the homogeneous differential equation $\tilde{P}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) f = 0$. We particularly observe that

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k} \tilde{P}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) f = \tilde{P}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) f^{(k)} = 0$$

for all $k \in \mathbb{N}$, where $f^{(k)}$ denotes the k-th derivative of f. As before, we can exploit this observation in order to reconstruct the parameters c_j and T_j, $j = 1, \ldots, M$, that

identify f. We fix a value $x_0 \in \mathbb{R}$ and apply the point evaluation functional F_{x_0} with $F_{x_0} f = f(x_0)$ to obtain the equations

$$F_{x_0}\left(\frac{\mathrm{d}^k}{\mathrm{d}x^k}\tilde{P}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)f\right) = F_{x_0}\left(\tilde{P}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)f^{(k)}\right) = \sum_{\ell=0}^{M}\tilde{p}_\ell f^{(k+\ell)}(x_0) = 0, \qquad (2.3)$$

for $k = 0, \ldots, M-1$. This homogeneous linear equation system yields the vector $\tilde{\mathbf{p}} = (\tilde{p}_0, \ldots, \tilde{p}_M)$ of coefficients of the Prony polynomial $\tilde{P}(z)$. Also here, the arising Hankel matrix $(f^{(k+\ell)}(x_0))_{k=0,\ell=0}^{M-1,M}$ has full rank M, such that $\tilde{\mathbf{p}}$ is uniquely defined with $\tilde{p}_M = 1$, see [19]. In turn we find the zeros T_j, $j = 1, \ldots, M$, of $\tilde{P}(z)$. Now the coefficients c_j can be obtained by solving the overdetermined linear system

$$\sum_{j=1}^{M} c_j T_j^k \exp(T_j x_0) = f^{(k)}(x_0), \qquad k = 0, \ldots, 2M-1.$$

2.3 Generalization 1: Switch between operators with the same eigenfunctions

An essential difference between the two approaches is that the required input values have completely different structure. Instead of the derivative values $f^{(k)}(x_0)$ for some $x_0 \in \mathbb{R}$ and $k = 0, \ldots, 2M-1$ for $\frac{\mathrm{d}}{\mathrm{d}x}$, we just need to provide the function values $f(x_0 + k\tau)$, $k = 0, \ldots, 2M-1$ for S_τ.

The second essential difference regards the condition of the matrices involved into the method. For $\frac{\mathrm{d}}{\mathrm{d}x}$ we have to find the zero eigenvector of the Hankel matrix $\tilde{\mathbf{H}} = (f^{(k+\ell)}(x_0))_{k=0,\ell=0}^{M-1,M} \in \mathbb{C}^{M \times M+1}$. Using the structure of $f(x)$ in (2.1), $\tilde{\mathbf{H}}$ has the factorization

$$\tilde{\mathbf{H}} = \tilde{\mathbf{V}}_{M,M} \operatorname{diag}(c_1, \ldots, c_M) \operatorname{diag}(\exp(T_1 x_0), \ldots, \exp(T_M x_0)) \tilde{\mathbf{V}}_{M+1,M}^T,$$

with the Vandermonde matrices $\tilde{\mathbf{V}}_{M,M} = (T_j^\ell)_{\ell=0,j=1}^{M-1,M}$ and $\tilde{\mathbf{V}}_{M+1,M} = (T_j^\ell)_{\ell=0,j=1}^{M,M}$. In contrast, for S_τ we have instead to solve the eigenvalue problem with the Hankel matrix $\mathbf{H} = (f(x_0 + \tau(k+\ell)))_{k=0,\ell=0}^{M-1,M}$ with the factorization

$$\mathbf{H} = \mathbf{V}_{M,M} \operatorname{diag}(c_1, \ldots, c_M) \operatorname{diag}(\exp(T_1 x_0), \ldots, \exp(T_M x_0)) \mathbf{V}_{M+1,M}^T,$$

where $\mathbf{V}_{M,M} = (\exp(T_j \tau \ell))_{\ell=0,j=1}^{M-1,M}$ and $\mathbf{V}_{M+1,M} = (\exp(T_j \tau \ell))_{\ell=0,j=1}^{M,M}$. Depending on the range of the parameters T_j the occurring Vandermonde matrices can have completely different condition number. If e.g. $T_j = \mathrm{i}\operatorname{Im} T_j$, then the knots $\exp(T_j \tau)$ determining $\mathbf{V}_{M,M}$ lie on the unit circle while the T_j determining $\tilde{\mathbf{V}}_{M,M}$ lie on the imaginary axis.

We are therefore interested in understanding the connection between the two methods to recover (2.1). Both approaches work, since the exponentials $\exp(T_j x)$ are eigenfunctions to the two different operators S_τ and $\frac{\mathrm{d}}{\mathrm{d}x}$. But the corresponding spectra are different. While the eigenvalues with regard to the differential operator $\frac{\mathrm{d}}{\mathrm{d}x}$ are of the form T_j, for the shift operator S_τ the eigenvalues are $\exp(T_j \tau)$. Obviously, the spectra are connected by the map $\exp(\tau \cdot) : \lambda \to \exp(\lambda \tau)$. With $\exp z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$ we indeed have

$$\exp\left(\tau\frac{\mathrm{d}}{\mathrm{d}x}\right)\exp(Tx) = \sum_{k=0}^{\infty}\frac{\tau^k}{k!}\frac{\mathrm{d}^k}{\mathrm{d}x^k}\exp(Tx) = \sum_{k=0}^{m}\frac{\tau^k}{k!}T^k\exp(Tx)$$

$$= \exp(\tau T) \exp(Tx) = S_\tau \exp(Tx) \tag{2.4}$$

for all $T \in \mathbb{C}$, and in turn for any analytic function $f \in \mathcal{M}$

$$\exp\left(\tau \frac{\mathrm{d}}{\mathrm{d}x}\right) f(x) = f(\tau + x) = (S_\tau f)(x),$$

see [11]. Thus, using the analytic function $\exp(\tau \cdot)$, we can map from the differential operator $\frac{\mathrm{d}}{\mathrm{d}x}$ to the shift operator S_τ, thereby staying with the same eigenfunctions but changing the eigenvalues. This observation is summarized in the following Theorem.

Theorem 2.1. *Let $\frac{\mathrm{d}}{\mathrm{d}x} : C^1(\mathbb{R}) \to C(\mathbb{R})$ be the first derivative operator. Then, each $T \in \mathbb{C}$ is an eigenvalue of $\frac{\mathrm{d}}{\mathrm{d}x}$. For some $C > 0$ let $\Lambda_C := \mathbb{R} + \mathrm{i}\,[-C\pi,\, C\pi)$ be a given subset of $\mathbb{C}$, and let $\varphi_\tau(x) := \exp(\tau x)$ with $\tau \le C^{-1}$. Then φ_τ is well-defined on $\mathbb{C}$ and*

$$\left(\frac{\mathrm{d}}{\mathrm{d}x} - T\,I\right) \exp(Tx) = 0$$

implies

$$\varphi_\tau\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) \exp(Tx) - \varphi_\tau(T\,I) \exp(Tx) = (S_\tau - \exp(\tau\,T)\,I)\,\exp(Tx) = 0,$$

where S_τ is the shift operator as before. Furthermore, the map $\varphi_\tau : T \to \exp(\tau\,T)$ is injective on Λ_C.

Proof. Obviously, $\frac{\mathrm{d}}{\mathrm{d}x} \exp(Tx) = T \exp(Tx)$ for all $T \in \mathbb{C}$. For all $T \in \Lambda_C$ the value $\varphi_\tau(T) = \exp(\tau\,T)$ is well-defined, and $\varphi_\tau(T_1) = \varphi_\tau(T_2)$ yields $T_1 = T_2 + \frac{2\pi k \mathrm{i}}{\tau}$, $k \in \mathbb{Z}$, i.e., $T_1 = T_2$ for $T_1, T_2 \in \Lambda_C$. The remaining assertions follow from (2.4). $\qquad\square$

Theorem 2.1 has strong implications on the reconstruction of $f(x)$ in (2.1) using Prony's method. We can replace the operator $\frac{\mathrm{d}}{\mathrm{d}x}$ by the operator S_τ in order to reconstruct f in (2.1), as we have seen in the previous two subsections.

2.4 Generalization 2: Changing the sampling scheme

In the two previous examples in Subsections 2.1 and 2.2 we have applied the point evaluation functional F_{x_0} with some $x_0 \in \mathbb{R}$ and used the samples

$$F_{x_0}(S_\tau^k f) = f(x_0 + k\tau) \quad \text{and} \quad F_{x_0}\left(\frac{\mathrm{d}^k}{\mathrm{d}x^k} f\right) = f^{(k)}(x_0), \quad k = 0, \ldots, 2M - 1,$$

respectively, to recover $f \in \mathcal{M}$. According to [19], we can however use any other linear functional $F : C^\infty \to \mathbb{C}$ with the only restriction that F applied to the eigenfunctions $\exp(Tx)$ should be well-defined and nonzero for all T in the parameter range we are interested in. We can for example take

$$Ff = \int_\Omega f(x)\,K(x)\,\mathrm{d}x$$

with some $\Omega \subset \mathbb{R}$ and some rather arbitrary kernel function $K(x)$ such that Ff is well defined and $\int_\Omega \exp(Tx)\,K(x)\,\mathrm{d}x \ne 0$ for all $T \in \mathbb{C}$. Thus, the choice of F gives us already some freedom to choose the sampling scheme. Taking e.g. $K(x) =$

$\sum_{r=-L}^{L} w_r \delta(x - r\tau)$ with the delta distribution δ and some positive weights w_r or just $K(x) := \chi_{[-1/2,1/2)}(x)$ we arrive at smoothed sampling values

$$F(S_\tau^k f) = \sum_{r=-L}^{L} w_r\, f((k + r)\tau) \quad \text{or} \quad F(S_\tau^k f) = \int_{-1/2}^{1/2} f(x + \tau k)\, \mathrm{d}x$$

instead of $f(x_0 + \tau k)$ for $k = 0, \ldots, 2M - 1$.

We can now generalize the sampling scheme even further if we allow ourselves to employ more than the minimal number of $2M$ input data. We inspect again the equations

$$F_{x_0}(S_\tau^k\, P(S_\tau)f) = 0, \qquad k = 0, \ldots, M - 1,$$

that lead in (2.2) to the Hankel system determining the coefficient vector $\mathbf{p}$ of the Prony polynomial $P(z)$. We already have $P(S_\tau)f = 0$, and the application of S_τ^k does not change the right-hand side of the equation. Therefore, for each $k = 0, \ldots, M-1$, we can replace $F_{x_0} S_\tau^k$ by a new linear functional F_k to obtain the M equations to recover $\mathbf{p}$. We only need to pay attention that the obtained M equations are linearly independent.

For example, we could take $F_k = F_{x_0} S_\theta^k$ with a parameter $\theta \notin \{0,\, \tau\}$ and obtain an equation system

$$F_{x_0}(S_\theta^k\, P(S_\tau)f) = \sum_{\ell=0}^{M} p_\ell f(x_0 + k\theta + \ell\tau) = 0, \quad k = 0, \ldots, M - 1.$$

The arising coefficient matrix $(f(x_0 + k\theta + \ell\tau))_{k=0,\ell=0}^{M-1,M}$ does not longer have Hankel structure but may possess a better condition than $(f(x_0 + (k + \ell)\tau))_{k=0,\ell=0}^{M-1,M}$. Taking e.g. $\theta = 2\tau$ we need the $3M - 1$ sample values $f(x_0 + \tau(2k + \ell))$ to recover f in (2.1).

Considering the method in Section 2.2, we can also replace the functionals $F_{x_0} \frac{\mathrm{d}^k}{\mathrm{d}x^k}$ in (2.3) by other linear functionals F_k. Taking for example $F_k = F_{x_0} S_\tau^k$ then we obtain the system

$$F_{x_0}\left(S_\tau^k \tilde{P}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)f\right) = \sum_{\ell=0}^{M} p_\ell\, f^{(\ell)}(x_0 + \tau k) = 0, \qquad k = 0, \ldots, M - 1.$$

Here, we need now the input data $f^{(\ell)}(x_0 + k\tau)$, $k = 0, \ldots, M - 1$, $\ell = 0, \ldots, M$, using only derivatives up to order M and its equidistant shifts. In Section 3.3 and in Section 5 we will investigate such generalized sampling schemes in more detail and particularly show that the examples above provide sampling matrices of full rank M, such that f in (2.1) can be uniquely reconstructed.

Remark 2.2. 1. Special generalized sampling schemes for the shift operator and the differential operator have also been proposed by Seelamantula [32], but without considering the relations between these operators. However, a rigorous investigation of rank properties of the involved matrices has not been given in [32]. The representation of Prony's method as an approach to reconstruct expansions into eigenfunctions of linear operators has been given already in [19].

2. For the special case of recovery of expansions into shifted Diracs in (1.2), it has been extensively studied how to retrieve the needed Fourier samples from low-pass projections with suitable sampling kernels, see e.g. [2, 5, 7, 12, 33, 34].

3 Generalized operator based Prony method

We want to study the two new observations considered for the special operators $\frac{\mathrm{d}}{\mathrm{d}x}$ and S_τ in Subsections 2.3 and 2.4 in a more general setting. We will call the new method Generalized Operator based Prony Method (GOP). For that purpose, we start with recalling the generalized Prony method from [19].

3.1 Generalized Prony method

Let V be a normed vector space over $\mathbb{C}$ and let $A : V \to V$ be a linear operator. Assume that A possesses a non-empty point spectrum $\sigma_P(A)$ and let $\sigma(A) \subset \sigma_P(A)$ be a (sub)set with pairwise different eigenvalues of A. We assume further that there is a corresponding set of eigenfunctions, i.e., for each $\lambda \in \sigma(A)$ we have a $v_\lambda \in V$ with $A v_\lambda = \lambda v_\lambda$, and the mapping $\lambda \mapsto v_\lambda$ is injective. In other words, the eigenspace to λ is one-dimensional, or, if this is not the case, we have to determine one relevant eigenfunction v_λ corresponding to λ in advance, which may occur in the expansion that we want to recover. Throughout the paper, we will assume that the considered eigenfunctions v_λ are normalized, i.e., $\|v_\lambda\|_V = 1$.

We want to reconstruct M-sparse expansions into eigenfunctions of A of the form

$$f = \sum_{j=1}^{M} c_j \, v_{\lambda_j} \tag{3.1}$$

where $\lambda_j \in \sigma(A)$ and where we always assume $c_j \in \mathbb{C} \setminus \{0\}$ for $j = 1, \dots, M$. The considered set of possible expansions is given as

$$\mathcal{M}(A) := \left\{ f = \sum_{j=1}^{M} c_j \, v_{\lambda_j} : \; M < \infty, \, c_j \in \mathbb{C} \setminus \{0\}, \, \lambda_j \in \sigma(A), \lambda_j \neq \lambda_k \text{ for } j \neq k \right\}. \tag{3.2}$$

The generalized Prony method in [19] provides an algorithm to recover f using only $2M$ complex measurements. For that purpose, a linear functional $F : V \to \mathbb{C}$ is introduced that satisfies $F(v_\lambda) \neq 0$ for all $\lambda \in \sigma(A)$.

Theorem 3.1 (Generalized Prony method [19]). *With the assumptions above, the expansion* (3.1) *of eigenfunctions v_{λ_j} of the linear operator A can be uniquely reconstructed from the values $F(A^k f)$, $k = 0, \dots, 2M - 1$.*

Proof. We give an outline of the proof in [19] with our notation. Observe that f is completely reconstructed if we recover the subset $\Lambda_f := \{\lambda_1, \dots, \lambda_M\} \subset \sigma(A)$ of "active eigenvalues" and the complex coefficients c_j, $j = 1, \dots, M$. The eigenfunctions v_{λ_j} are then uniquely determined by λ_j.

Let $P(z) = \prod_{j=1}^{M}(z - \lambda_j) = \sum_{\ell=0}^{M} p_\ell z^\ell$ be the Prony polynomial determined by the set of M pairwise different (unknown) active eigenvalues $\lambda_j \in \Lambda_f$, and $\mathbf{p} = (p_0, \dots, p_{M-1}, p_M)^T$ with $p_M = 1$ denotes the vector of its monomial coefficients. Then we obtain by (3.1)

$$P(Af) = \prod_{k=1}^{M}(A - \lambda_k I)\, f = \sum_{j=1}^{M} c_j \prod_{k=1}^{M}(A - \lambda_k I)\, v_{\lambda_j} = 0, \tag{3.3}$$

and therefore

$$F(A^k P(A) f) = F\Big(A^k \Big(\sum_{\ell=0}^{M} p_\ell A^\ell f\Big)\Big) = \sum_{\ell=0}^{M} p_\ell F(A^{\ell+k} f) = 0 \qquad (3.4)$$

for all $k \in \mathbb{N}$. Taking M equations for $k = 0, \ldots, M-1$, is already sufficient to recover the coefficient vector $\mathbf{p}$, since the matrix

$$\Big(F(A^{\ell+k} f)\Big)_{k=0,\ell=0}^{M-1,M}$$

has full rank M. This can be seen from the factorization

$$
\begin{aligned}
\Big(F(A^{\ell+k} f)\Big)_{k=0,\ell=0}^{M-1,M} &= \Big(F(A^{\ell+k} \sum_{j=1}^{M} c_j v_{\lambda_j})\Big)_{k=0,\ell=0}^{M-1,M} = \Big(\sum_{j=1}^{M} c_j F(A^{\ell+k} v_{\lambda_j})\Big)_{k=0,\ell=0}^{M-1,M} \\
&= \Big(\sum_{j=1}^{M} c_j F(v_{\lambda_j}) \lambda_j^{\ell+k}\Big)_{k=0,\ell=0}^{M-1,M} \\
&= V_{\Lambda_f,M,M} \operatorname{diag}\big(c_j F(v_{\lambda_j})\big)_{j=1}^{M} V_{\Lambda_f,M+1,M}^T \qquad (3.5)
\end{aligned}
$$

with the Vandermonde matrices

$$V_{\Lambda_f,M,M} := (\lambda_j^k)_{k=0,j=1}^{M-1,M}, \qquad V_{\Lambda_f,M+1,M} := (\lambda_j^k)_{k=0,j=1}^{M,M}$$

having full rank M. Thus, we can first compute $\mathbf{p}$ as the right eigenvector of $(F(A^{\ell+k} f))_{k=0,\ell=0}^{M-1,M}$ to the eigenvalue 0 with normalization $p_M = 1$, determine $P(z)$, then extract the zeros λ of $P(z)$ to recover λ_j, $j = 1, \ldots, M$, and finally compute the coefficients c_j, $j = 1, \ldots, M$, by solving an overdetermined linear system of the form

$$F(A^k f) = \sum_{j=1}^{M} c_j \lambda_j^k F(v_{\lambda_j}), \qquad k = 0, \ldots, 2M-1.$$

$\square$

Remark 3.2. As shown in [19] and [26], many expansions fit into the scheme of Theorem 3.1. In Section 2 we have used A to be the shift operator or the differential operator. Other examples in [19] and [26] include the dilation operator, generalized shift operators as well as the Sturm-Liouville differential operator of second order.

3.2 Generalization 1: Change of operators

The actions $A^k f$ needed for the generalized Prony method to recover $f \in \mathcal{M}(A)$ in (3.2) may be very expensive to acquire. Therefore we can try to replace the operator A by a different operator with the same eigenfunctions v_λ such that the powers of this new operator are simpler to realize. We start with the following definition.

Definition 3.3 (Iteration Operator)**.** Let $A : V \to V$ be a linear operator, and let $\sigma(A) \neq \emptyset$ be a subset of the point spectrum $\sigma_P(A)$ with pairwise different eigenvalues and with corresponding normalized eigenfunctions v_λ such that the map $\lambda \mapsto v_\lambda$ is injective for $\lambda \in \sigma(A)$. Further, let $\varphi : \sigma(A) \to \mathbb{C}$ be an injective function. We call $\Phi = \Phi_\varphi$ an *iteration operator to A* if $\Phi : \mathcal{M}(A) \to \mathcal{M}(A)$ is a well-defined linear operator and $\Phi v_\lambda = \varphi(\lambda) v_\lambda$ for all $\lambda \in \sigma(A)$.

The injectivity of φ in Definition 3.3 implies that the values $\varphi(\lambda)$ are pairwise different for all $\lambda \in \sigma(A)$. In particular, we can show that for analytic functions φ the operator $\Phi = \varphi(A)$ is an iteration operator.

Theorem 3.4. *Let $A : V \to V$ be a linear operator, and let $\sigma(A) \neq \emptyset$ be a subset of the point spectrum $\sigma_P(A)$ with pairwise different eigenvalues and with corresponding eigenfunctions v_λ such that the map $\lambda \mapsto v_\lambda$ is injective for $\lambda \in \sigma(A)$. Let $\varphi : \sigma(A) \to \mathbb{C}$ be an analytic, injective function. Then $\varphi(A)$ is an iteration operator, i.e., it is a well-defined linear operator on $\mathcal{M}(A)$ and*

$$(A - I\lambda)\, v_\lambda = 0$$

implies

$$(\varphi(A) - \varphi(\lambda)I)\, v_\lambda = 0.$$

This means, if v_λ is an eigenfunction of A corresponding to the eigenvalue λ, then v_λ is also an eigenfunction of $\varphi(A)$ corresponding to the eigenvalue $\varphi(\lambda)$.

Proof. Since φ is assumed to be analytic on $\sigma(A)$, it follows that its power series $\varphi(z) = \sum_{n=0}^\infty a_n z^n$ converges for $z \in \sigma(A)$. Thus, $A\, v_\lambda = \lambda\, v_\lambda$ implies for all $\lambda \in \sigma(A)$

$$\varphi(A)\, v_\lambda = \sum_{n=0}^\infty a_n\, A^n v_\lambda = \lim_{N \to \infty} \sum_{n=0}^N a_n \lambda^n v_\lambda = \varphi(\lambda)\, v_\lambda.$$

Further, the injectivity of φ implies that the eigenvalues $\varphi(\lambda)$, $\lambda \in \sigma(A)$, are pairwise distinct. Thus, $\varphi(A)$ is well-defined on $\mathcal{M}(A)$ and satisfies all assumptions of an iteration operator. $\qquad\square$

Example 3.5. 1. One example has been already seen in Section 2. We can take $V = C^\infty(\mathbb{R})$, $A = \frac{\mathrm{d}}{\mathrm{d}x}$ with $\sigma_P(A) = \mathbb{C}$ according to Theorem 2.1. Further, let $\sigma(A) = \mathbb{R} + \mathrm{i}\,[-C\pi,\, C\pi) \subset \sigma_P(A)$. Then, $\varphi(z) := \exp(\tau z)$ with $0 < \tau \leq 1/C$ is injective on $\sigma(A)$, and we obtain the iteration operator $\varphi(A) = S_\tau$ on $\mathcal{M}(A)$.

2. We take $\varphi(z) = z^{-1}$ and $\sigma(A) \in \sigma_P(A) \setminus \{0\}$. Then $\varphi(A) = A^{-1}$ is well-defined on $\mathcal{M}(A)$ and

$$A\, v_\lambda = \lambda\, v_\lambda \qquad \Leftrightarrow \qquad A^{-1}\, v_\lambda = \frac{1}{\lambda}\, v_\lambda.$$

For example, $A = S_\tau$ with $\tau \neq 0$ yields $A^{-1} = S_{-\tau}$. The dilation operator $D_a : C(\mathbb{R}) \to C(\mathbb{R})$ with $D_a f(x) := f(ax)$, $a \neq 0$ and $|a| \neq 1$, yields $D_a^{-1} f(x) = f(\frac{1}{a}x)$.

3. Consider the operator A on $C^\infty(\mathbb{R})$ given by

$$A f(x) := x\, \frac{\mathrm{d}f}{\mathrm{d}x}(x) = x\, f'(x)$$

with eigenfunctions x^p for $p \in \mathbb{R}$ to the eigenvalues $p \in \mathbb{R}$. We use $\varphi(z) = \exp(\tau z)$ with $\tau \in \mathbb{R} \setminus \{0\}$ and obtain for each polynomial x^m that

$$\exp(\tau x\, \frac{\mathrm{d}}{\mathrm{d}x})\, x^m = \sum_{\ell=0}^\infty \frac{\tau^\ell}{\ell!} \left(x\, \frac{\mathrm{d}}{\mathrm{d}x} \right)^\ell x^m = \sum_{\ell=0}^\infty \frac{\tau^\ell}{\ell!}\, m^\ell\, x^m = \mathrm{e}^{\tau m}\, x^m = (\mathrm{e}^\tau x)^m,$$

see also [10]. Thus, $\varphi(A)$ is here the dilation operator $D_{\exp(\tau)}$. The injectivity condition for $\varphi(z)$ is satisfied since $\exp(\tau p)$ is strictly monotone as a function in p. $\square$

What does a change from A to $\varphi(A)$ mean for the reconstruction scheme to recover an expansion f in (3.1)? Using the operator A and a functional F, Theorem 3.1 implies that we need (at least) the sample values $F(A^k f)$, $k = 0, \ldots, 2M - 1$ for the recovery of f. Changing from A to $\varphi(A)$, we observe that all assumptions required in Theorem 3.1 also hold for $\varphi(A)$, and we can now reconstruct f in (3.1) from samples $F(\varphi(A)^k f)$, $k = 0, \ldots, 2M - 1$, thereby employing the new Prony polynomial

$$P_\varphi(z) := \prod_{j=1}^{M} (z - \varphi(\lambda_j)) := \sum_{\ell=0}^{M} p_\ell \, z^\ell.$$

Taking a suitable φ may have two advantages. First, the samples $F(\varphi(A)^k f)$, $k = 0, \ldots, 2M - 1$, may be much simpler to acquire. In Section 3.5 and Section 4, we will present many examples, where a change from linear differential operators A to generalized shift operators $\varphi(A)$ leads to new recovery schemes for the expansions in (3.1) employing just function values of f instead of high order derivative values.

Second, the numerical scheme to recover f can be essentially stabilized. The main reason for that is the change of eigenvalues from $\lambda \in \Lambda_f$ to $\varphi(\lambda) \in \varphi(\Lambda_f)$. The eigenvalues play an important role for the matrices being involved in the Prony algorithms. Compared with the generalized Prony method, we get now instead of (3.5) the Hankel matrix factorization

$$\left(F(\varphi(A)^{\ell+k} f) \right)_{k=0,\ell=0}^{M-1,M} = V_{\varphi(\Lambda_f),M,M} \operatorname{diag} \left(c_j \, F(v_{\lambda_j}) \right)_{j=1}^{M} V_{\varphi(\Lambda_f),M+1,M}^{T}$$

with the Vandermonde matrices

$$V_{\varphi(\Lambda_f),M,M} := \left(\varphi(\lambda_j)^k \right)_{k=0,j=1}^{M-1,M}, \qquad V_{\varphi(\Lambda_f),M+1,M} := \left(\varphi(\lambda_j)^k \right)_{k=0,j=1}^{M,M}$$

to recover the coefficient vector $\mathbf{p} = (p_0, \ldots, p_M)^T$ of the Prony polynomial P_φ.

3.3 Generalization 2: Change the sampling scheme

As we have seen in Theorem 3.1 and Theorem 3.4, the expansion $f = \sum_{j=1}^{M} c_j \, v_{\lambda_j}$ into eigenfunctions of the operator A can be recovered using either the samples $F(A^k f)$ or the samples $F(\varphi(A)^k f)$ for $k = 0, \ldots, 2M - 1$, where $F : V \to \mathbb{C}$ is a linear functional satisfying $F(v_\lambda) \neq 0$ for all $\lambda \in \sigma(A)$. Having a closer look at the equations (3.3) and (3.4) we observe however that already $P(A)f = 0$, such that $F A^k$ can be replaced by different functionals.

Definition 3.6 (Sampling Functionals)**.** *Let $A : V \to V$ be a linear operator and let $\sigma(A)$ be a fixed subset of pairwise different eigenvalues of A. Further, let*

$$V_{\sigma(A)} := \{ v_\lambda : \; A \, v_\lambda = \lambda \, v_\lambda, \; \lambda \in \sigma(A), \; \|v_\lambda\|_V = 1 \}$$

be the corresponding set of eigenfunctions such that the mapping $\lambda \to v_\lambda$ is injective on $\sigma(A)$. Then $\{F_k\}_{k=0}^{M-1}$ with

$$F_k : V \to \mathbb{C}, \qquad k = 0, \ldots, M - 1,$$

forms an admissible set of sampling functionals for A if for all finite subsets $\Lambda_M \subset \sigma(A)$ with cardinality $M < \infty$ the matrix

$$\left(F_k(v_\lambda) \right)_{k=0,\lambda \in \Lambda_M}^{M-1}$$

has full rank M.

If the set of functionals $\{F_k\}_{k=0}^{M-1}$ is admissible for a linear operator A, then it is also admissible for any iteration operator $\varphi(A)$, since the eigenvectors v_λ do not change. Then we obtain

Theorem 3.7. *Assume that $\{F_k\}_{k=0}^{M-1}$ forms an admissible set of sampling functionals for the linear operator $A : V \to V$ according to Definition 3.6. Let $f \in \mathcal{M}(A)$ be a linear expansion into eigenfunctions of A as in (3.1). Then the sampling matrix*

$$\left(F_k(A^\ell f)\right)_{k=0,\ell=0}^{M-1,M} \in \mathbb{C}^{M \times (M+1)}$$

possesses rank M and is called admissible sampling matrix for f. Further, if $\Phi = \varphi(A)$ is an iteration operator of A as given in Theorem 3.4, then also

$$\left(F_k(\Phi^\ell f)\right)_{k=0,\ell=0}^{M-1,M} \in \mathbb{C}^{M \times (M+1)}$$

possesses rank M and is therefore an admissible sampling matrix.

Proof. We show the second equation for $\Phi = \varphi(A)$, where φ is an injective analytic function on $\sigma(A)$. Then the first equation follows by taking $\varphi(z) = z$. We find

$$\left(F_k(\varphi(A)^\ell f)\right)_{k=0,\ell=0}^{M-1,M} = \left(F_k(\varphi(A)^\ell \sum_{j=1}^{M} c_j v_{\lambda_j})\right)_{k=0,\ell=0}^{M-1,M} = \left(\sum_{j=1}^{M} c_j \varphi(\lambda_j)^\ell F_k(v_{\lambda_j})\right)_{k=0,\ell=0}^{M-1,M}$$

$$= \left(F_k(v_{\lambda_j})\right)_{k=0,j=1}^{M-1,M} \operatorname{diag}(c_j)_{j=1}^{M} \left(\varphi(\lambda_j)^\ell\right)_{j=1,\ell=0}^{M,M}.$$

All three matrices in this factorization have full rank M by assumption, and the assertion follows. In particular, the last matrix is a Vandermonde matrix generated by M pairwise distinct values $\varphi(\lambda_j)$, $j = 1, \ldots, M$. $\qquad\square$

Example 3.8. Comparison with formula (3.4) yields that $F_k = FA^k$, $k = 0, \ldots, M-1$, is always an admissible set of sampling functionals, since the proof of Theorem 3.1 shows that $(F(A^{k+\ell} f))_{k=0,\ell=0}^{M-1,M}$ has full rank M for each f in $\mathcal{M}(A)$. $\square$

Further we have

Lemma 3.9. *Let $A : V \to V$ be a linear operator, and let $\sigma(A) \neq \emptyset$ be a subset of the point spectrum $\sigma_P(A)$ with pairwise different eigenvalues and with corresponding eigenfunctions v_λ such that the map $\lambda \mapsto v_\lambda$ is injective for $\lambda \in \sigma(A)$. Let ψ be an analytic injective function on $\sigma(A)$. Assume that $F : V \to \mathbb{C}$ is a linear functional with $Fv_\lambda \neq 0$ for all $\lambda \in \sigma(A)$. Then $\{F_k\}_{k=0}^{M-1} := \{F(\psi(A)^k)\}_{k=0}^{M-1}$ is an admissible set of sampling functionals and the matrix*

$$(F_k(A^\ell f))_{k=0,\ell=0}^{M-1,M} = (F(\psi(A)^k A^\ell f))_{k=0,\ell=0}^{M-1,M} \in \mathbb{C}^{M \times M+1}$$

is an admissible sampling matrix for each $f \in \mathcal{M}(A)$.

Proof. From $\psi(A)^k v_\lambda = \psi(\lambda)^k v_\lambda$ it follows that

$$F_k(v_\lambda) = F(\psi(A)^k v_\lambda) = \psi(\lambda)^k F(v_\lambda)$$

is bounded and nonzero by assumption. Further, for $f \in \mathcal{M}(A)$,

$$\left(F(\psi(A)^k A^\ell f)\right)_{k=0,\ell=0}^{M-1,M} = \left(F\left(\psi(A)^k A^\ell \sum_{j=1}^{M} c_j v_{\lambda_j}\right)\right)_{k=0,\ell=0}^{M-1,M}$$

$$= \left(F\Big(\sum_{j=1}^{M} c_j \, \psi(\lambda_j)^k \, \lambda_j^{\ell} \, v_{\lambda_j} \Big) \right)_{k=0,\ell=0}^{M-1,M}$$

$$= \mathbf{V}_{\psi(\Lambda_f),M,M} \, \mathrm{diag}\, \big((c_j \, F(v_{\lambda_j})) \big)_{j=1}^{M} \, \mathbf{V}_{\Lambda_f,M+1,M}^{T}$$

with , with $\Lambda_f = \{\lambda_1, \ldots, \lambda_M\}$, $\mathbf{V}_{\psi(\Lambda_f),M,M} := ((\psi(\lambda_j))^k)_{k=0,j=1}^{M-1,M}$ and $\mathbf{V}_{\Lambda_f,M+1,M} := (\lambda_j^{\ell})_{\ell=0,j=1}^{M,M}$. These two Vandermonde matrices have full rank M since the $\lambda_j \in \Lambda_f$ are pairwise different and ψ is injective on Λ_f with $\psi(\lambda_j) \neq 0$ for $\lambda_j \in \Lambda_f$. $\qquad\square$

3.4 Generalized operator based Prony method (GOP)

The following theorem summarizes the central statement of the generalized operator-based Prony method (GOP) and the corresponding proof results in an algorithm to solve the reconstruction problem for $f \in \mathcal{M}(A)$ in (3.2).

Theorem 3.10 (Generalized Operator based Prony Method).
Let $A : V \to V$ be a linear operator on the normed vector space V over $\mathbb{C}$, and let $\sigma(A)$ be a subset of pairwise different eigenvalues of A. Let $\Phi = \varphi(A)$ be an iteration operator of A as given in Definition 3.3. Assume that the set $\{F_k\}_{k=0}^{M-1}$ is an admissible set of sampling functionals according to Definition 3.6. Then each $f \in \mathcal{M}(A)$ can be completely recovered from the complex samples $F_k(\varphi(A)^{\ell} f)$, $k = 0, \ldots, M-1$, $\ell = 0, \ldots, M$.

Proof. To recover $f = \sum_{j=1}^{M} c_j \, v_{\lambda_j} \in \mathcal{M}(A)$, we only have to determine the set $\Lambda_f = \{\lambda_1, \ldots, \lambda_M\}$ of "active eigenvalues" and the corresponding coefficients $c_j \in \mathbb{C} \setminus \{0\}$, $j = 1, \ldots, M$, since the map $\lambda \to v_\lambda$ is assumed to be injective. Further, since φ is also injective on $\sigma(A)$, we can determine the set $\varphi(\Lambda_f) = \{\varphi(\lambda_j) : j = 1, \ldots, M\}$ instead of Λ_f by Theorem 3.4.

Let now

$$P_\varphi(z) := \prod_{j=1}^{M} (z - \varphi(\lambda_j)) = \sum_{\ell=0}^{M} p_\ell \, z^{\ell}$$

be the Prony polynomial determined be the unknown pairwise different active eigenvalues $\varphi(\lambda_j)$ of $\varphi(A)$ for $\lambda_j \in \Lambda_f$, where $\mathbf{p} = (p_0, \ldots, p_{M-1}, p_M)^T$ with $p_M = 1$ denotes the vector of coefficients in the monomial representation of $P_\varphi(z)$. Then

$$\begin{aligned} P_\varphi(\varphi(A))f &= \prod_{k=1}^{M} (\varphi(A) - \varphi(\lambda_k)I)\, f \\ &= \sum_{j=1}^{M} c_j \prod_{k=1}^{M} (\varphi(A) - \lambda_k I) v_{\lambda_j} = 0, \end{aligned}$$

and therefore

$$F_k(P_\varphi(\varphi(A))f) = F_k\Big(\sum_{\ell=0}^{M} p_\ell \, \varphi(A)^{\ell} f \Big) = \sum_{\ell=0}^{M} p_\ell F_k(\varphi(A)^{\ell} f) = 0, \ k = 0, \ldots, M-1.$$

Thus, we obtain a homogeneous linear system to compute $\mathbf{p}$, where by Theorem 3.7 (with A replaced by $\varphi(A)$) the coefficient matrix is the admissible sampling matrix $(F_k(\varphi(A)^{\ell} f))_{k=0,\ell=0}^{M-1,M} \in \mathbb{C}^{M \times M+1}$ with full rank M. Hence, $\mathbf{p}$ is uniquely determined

by this system using the normalization $p_M = 1$. We can now extract the zeros $\varphi(\lambda_j)$, $j = 1, \ldots, M$, and thus $\Lambda_f = \{\lambda_1, \ldots, \lambda_M\}$. Finally, we compute the coefficients c_λ as solutions of the linear system

$$F_k(\varphi(A)^\ell f) = \sum_{j=1}^M c_j \varphi(\lambda_j)^\ell F_k(v_{\lambda_j}), \quad \ell = 0, \ldots, M, \tag{3.6}$$

where the coefficient matrix is of full rank, since $F_k(v_{\lambda_j}) \neq 0$ and the arising Vandermonde matrix $((\varphi(\lambda_j))^\ell)_{\ell=0,j=1}^{M,M}$ has full rank M since the values $\varphi(\lambda_j)$, $j = 1, \ldots, M$, are pairwise different. $\qquad\square$

The proof of Theorem 3.10 is constructive and leads to the following algorithm for the recovery of $f \in \mathcal{M}(A)$. We assume here that we have an iteration operator $\varphi(A)$ and a given set of admissible sampling functionals F_k such that the sampling matrix $(F_k(\varphi(A)^\ell f))_{k=0,\ell=0}^{M-1,M} \in \mathbb{C}^{M \times M+1}$ for the operator $\varphi(A)$ has full rank M.

Algorithm 3.11 (GOP).
Input: $F_k\big(\varphi(A)^\ell f\big)$, $\ell = 0, \ldots, M$, $k = 0, \ldots, M-1$, where $f \in \mathcal{M}(A)$.

- Compute the kernel vector $\mathbf{p} = (p_0, \ldots, p_{M-1}, p_M)^T$ with $p_M = 1$ of the matrix $(F_k(\varphi(A)^\ell f))_{k=0,\ell=0}^{M-1,M} \in \mathbb{C}^{M \times M+1}$.
- Compute the M zeros $\varphi(\lambda_j)$, $j = 1, \ldots, M$, of the Prony polynomial $P_\varphi(z) = \sum_{\ell=0}^M p_\ell z^\ell$ and identify the active eigenfunctions v_{λ_j} by $\varphi(A)\,v_{\lambda_j} = \varphi(\lambda_j)\,v_{\lambda_j}$. Compute λ_j from $\varphi(\lambda_j)$ to obtain $\Lambda_f = \{\lambda_1, \ldots, \lambda_M\}$.
- Compute c_j by solving the system in (3.6).

Output: Parameters λ_j and c_j, $j = 1, \ldots, M$ such that $f = \sum_{j=1}^M c_j\, v_{\lambda_j}$.

Remark 3.12. 1. The generalized Prony method in [19] is a special case of GOP if we take $\varphi(z) = z$ and $F_k = F\,A^k$ for some suitable functional F. In this case the sampling matrix has Hankel structure and we need only $2M$ input values.

2. If we choose $F_k = F(\psi(A)^k \cdot)$ for some analytic function ψ as in Lemma 3.9, then the sampling matrix can be taken in the form $(F(\psi(A)^k\,\varphi(A)^\ell f))_{k=0,\ell=0}^{M-1,M} \in \mathbb{C}^{M \times M+1}$, where compared to Lemma 3.9, we have replaced the powers of A by powers of $\varphi(A)$. This sampling matrix is also admissible, and the proof can be performed as for Lemma 3.9.

3. GOP can be also generalized to operators with eigenvalues of higher geometric multiplicity, similarly as the generalized Prony method, [19]. This approach leads to a Prony polynomial with zeros of higher multiplicity. We also refer to [3, 17]. In this paper we restrict ourselves to the case where the correspondence between λ resp. $\varphi(\lambda)$ and v_λ is bijective.

3.5 Application of GOP to cosine expansions

In this section, we want to explain the ideas of GOP in a simple example.

Consider the expansion

$$f(x) := \sum_{j=1}^M c_j\, \cos(\alpha_j x), \tag{3.7}$$

where we want to recover the $2M$ parameters $\alpha_j \in [0, C) \subset \mathbb{R}$ and $c_j \in \mathbb{C} \setminus \{0\}$, $j = 1, \ldots, M$. We observe that $A := -\frac{\mathrm{d}^2}{\mathrm{d}x^2}$ is an operator on $C^\infty(\mathbb{R})$ such that all functions $\cos(\alpha x)$ are eigenfunctions of A with

$$A \cos(\alpha \cdot) = \alpha^2 \cos(\alpha \cdot).$$

Using the generalized Prony method in Theorem 3.1, we can therefore reconstruct f in (3.7) using the samples $F(A^k f) = (-1)^k F(f^{(2k)})$, $k = 0, \ldots, 2M - 1$, where $f^{(2k)}$ denotes the $2k$-th derivative of f. Here, the sampling functional $F : C^\infty(\mathbb{R}) \to \mathbb{C}$ needs to satisfy $F(\cos(\alpha \cdot)) \neq 0$ for all all $\alpha \in [0, C)$.

Taking e.g. the point evaluation functional $Ff = f(0)$, we need the measurements $f^{(2k)}(0)$, $k = 0, \ldots, 2M - 1$. These measurements are usually difficult to provide, it would be much better to use just function values of f.

We want to apply now GOP in Theorem 3.10 to reconstruct f in (3.7) in a different way. We employ the analytic function $\varphi(z)$ of the form

$$\varphi(z) = \sum_{n=0}^{\infty} (-1)^n \frac{\tau^{2n} z^n}{(2n)!},$$

i.e., $\varphi(z^2) = \cos(\tau z)$, and observe that the application of $\varphi(A)$ to monomial functions x^m gives

$$
\begin{aligned}
\varphi(A) \, x^m &= \sum_{n=0}^{\infty} (-1)^n \frac{\tau^{2n}}{(2n)!} \left(-\frac{\mathrm{d}^2}{\mathrm{d}x^2} \right)^n x^m \\
&= \sum_{0 \leq 2n \leq m} \binom{m}{2n} \tau^{2n} \, x^{m-2n} \\
&= \frac{1}{2} \left(\sum_{0 \leq n' \leq m} \binom{m}{n'} \tau^{n'} x^{m-n'} + \sum_{0 \leq n' \leq m} \binom{m}{n'} (-\tau)^{n'} x^{m-n'} \right) \\
&= \frac{1}{2} \left((x+\tau)^m + (x-\tau)^m \right) = \frac{1}{2} (S_\tau + S_{-\tau}) \, x^m
\end{aligned}
$$

with the shift operator S_τ given by $S_\tau f = f(\cdot + \tau)$. Thus we have

$$\varphi(A) = \frac{1}{2} (S_\tau + S_{-\tau})$$

and by Theorem 3.4 it follows that

$$
\begin{aligned}
\varphi(A) \, \cos(\alpha \cdot) &= \frac{1}{2}(S_\tau + S_{-\tau}) \cos(\alpha \cdot) = \frac{1}{2} (\cos(\alpha(\cdot + \tau)) + \cos(\alpha(\cdot - \tau))) \\
&= \cos(\alpha \tau) \, \cos(\alpha \cdot),
\end{aligned}
$$

i.e., the eigenvalues α^2 of $A = -\frac{\mathrm{d}^2}{\mathrm{d}x^2}$ are transferred to $\cos(\tau \alpha)$. We can still identify $\alpha \in [0, C)$ uniquely from $\cos(\tau \alpha)$ if $\tau \leq \frac{\pi}{C}$.

In order to apply GOP, we also need to fix an admissible sampling matrix. According to Lemma 3.9, we can use an admissible set of sampling functionals

$$F_k = F(\varphi(A)^k) = F\left(\frac{1}{2}(S_\tau + S_{-\tau})\right)^k = F\left(\frac{1}{2^k} \sum_{r=0}^{k} \binom{k}{r} S_{(k-2r)\tau} \right) \qquad (3.8)$$

and arrive with the point evaluation functional $Ff := f(0)$ at the sampling matrix $\left(F_k(\varphi(A)^\ell f)\right)_{k=0,\ell=0}^{M-1,M}$ with entries

$$F_k(\varphi(A)^\ell f) = F(\varphi(A)^{k+\ell} f) = \frac{1}{2^{k+\ell}} \sum_{r=0}^{k+\ell} \binom{k+\ell}{r} f((k+\ell-2r)\tau).$$

This matrix involves the function samples $f(k\tau)$, $-2M+1 \le k \le 2M-1$. Since f in (3.7) is symmetric, it is sufficient to provide $f(k\tau)$, $k = 0, \ldots, 2M-1$. Indeed,

$$
\begin{aligned}
F(\varphi(A)^{k+\ell} f) &= \sum_{j=1}^{M} c_j \, F(\varphi(A)^{k+\ell} \cos(\alpha_j \cdot)) \\
&= \sum_{j=1}^{M} c_j \, (\cos(\alpha_j \tau))^{\ell+k} \, F(\cos(\alpha_j \cdot)) = \sum_{j=1}^{M} c_j \, (\cos(\alpha_j \tau))^{\ell+k}
\end{aligned}
$$

yields that the sampling matrix can be simply factorized, and all matrix factors have full rank M.

We can employ a different sampling matrix by taking

$$F_k(f) = ((S_{k\tau} + S_{-k\tau})f)(0)$$

instead of (3.8) and get the matrix entries

$$((S_{k\tau} + S_{-k\tau})(\varphi(A)^\ell f))(0) = \frac{1}{2^\ell} \sum_{r=0}^{\ell} \binom{\ell}{r} [f((\ell+k-2r)\tau) + f(\ell-k-2r)\tau)]. \quad (3.9)$$

For f of the form (3.7) this sampling matrix is also admissible since we obtain with the Chebyshev polynomial $T_k(z) := \cos(k(\arccos z))$ that

$$
\begin{aligned}
((S_{k\tau} &+ S_{-k\tau})\varphi(A)^\ell f)(0) \\
&= \frac{1}{2^\ell} \sum_{r=0}^{\ell} \binom{\ell}{r} \sum_{j=1}^{M} c_j [\cos(\alpha_j(\ell+k-2r)\tau) + \cos(\alpha_j(\ell-k-2r)\tau)] \\
&= \sum_{j=1}^{M} c_j \left(\frac{2}{2^\ell} \sum_{r=0}^{\ell} \binom{\ell}{r} \cos(\alpha_j(\ell-2r)\tau) \right) \cos(\alpha_j k\tau) \\
&= \sum_{j=1}^{M} c_j \left(\frac{2}{2^\ell} \sum_{r=0}^{\ell} \binom{\ell}{r} T_{|\ell-2r|}(\cos(\alpha_j \tau)) \right) \cos(\alpha_j k\tau) \\
&= 2 \sum_{j=1}^{M} c_j \, (\cos(\alpha_j \tau))^\ell \, \cos(\alpha_j k\tau),
\end{aligned}
$$

where we have used the identity $x^\ell = \frac{1}{2^\ell} \sum_{r=0}^{\ell} \binom{\ell}{r} T_{|\ell-2r|}(x)$. Thus

$$\left(((S_{k\tau} + S_{-k\tau})\varphi(A)^\ell f)(0)\right)_{k=0,\ell=0}^{M-1,M} = (\cos(\alpha_j k\tau))_{k=0,j=1}^{M-1,M} \operatorname{diag}(2c_j)_{j=1}^{M} ((\cos(\alpha_j \tau))^\ell)_{j=1,\ell=0}^{M,M},$$

where all matrix factors have full rank M. The sampling matrix in (3.9) applies the idea that instead of $F_k(f) = F(\varphi(A)^k f)$, $k = 0, \ldots, M-1$, we can also use

$$F_k(f) = F(p_k(\varphi(A)f)), \qquad k = 0, \ldots, M-1,$$

with a basis $\{p_k\}_{k=0}^{M-1}$ of the space of algebraic polynomials up to degree $M-1$. Here, (3.9) is obtained by using the basis of Chebyshev polynomials $p_k = T_k$, $k = 0, \ldots, M-1$.

Remark 3.13. A slightly different sampling scheme was applied in [30] and in [26], where the Prony polynomial has been written using a Chebyshev polynomial basis instead of the monomial basis.

4 GOP for special linear differential operators of first and second order

In this section we discuss the application of GOP for the recovery of expansions into eigenfunctions of linear differential operators. In this case, we will mainly apply iteration operators that are constructed using $\varphi(z) = \exp(\tau z)$ and $\varphi(z) = \cos(\tau z^{1/2})$. We will show that the obtained iteration operators are generalized shift operators that enable us to recover the considered expansions using only function values instead of derivative values. We will consider sampling functionals $F_k : \mathcal{M} \to \mathbb{C}$ of the form

$$F_k(f) = F(\varphi(A)^k f).$$

With this sampling, GOP is equivalent with the generalized Prony method for $\varphi(A)$ (instead of A) and a fixed functional F that only needs to satisfy the assumptions of Theorem 3.1. Then, the corresponding sampling matrix is always admissible for all $f \in \mathcal{M}(A)$ in (3.2), and we need the values $F((\varphi(A)^k f)$, $k = 0, \ldots, 2M - 1$ to reconstruct f in (3.1).

4.1 Differential operators of first order and generalized shifts

Assume that $G : I \to J \subset \mathbb{R}$ is in $C^\infty(I)$ and that its first derivative $G'(x)$ has no zero on I. This means in particular that $g(x) = 1/G'(x)$ is well-defined on I. Moreover, $G(x)$ is strictly monotone on I such that $G^{-1}(x)$ is also well-defined on I. Further, let $H \in C^\infty(I)$.

We want to reconstruct functions of the form

$$f(x) = \sum_{j=1}^{M} c_j \, \mathrm{e}^{H(x) + \lambda_j G(x)}, \tag{4.1}$$

i.e., we want to recover the parameters $c_j \in \mathbb{C} \setminus \{0\}$ and $\lambda_j \in \mathbb{R} + \mathrm{i}[-C, C)$. We define the functions

$$g(x) := \frac{1}{G'(x)}, \qquad h(x) := -\frac{H'(x)}{G'(x)}. \tag{4.2}$$

Then $v_{\lambda_j}(x) := \mathrm{e}^{H(x) + \lambda_j G(x)}$ are eigenfunctions of

$$A = g(\cdot) \frac{\mathrm{d}}{\mathrm{d}x} + h(\cdot), \tag{4.3}$$

since we have for all $\lambda \in \mathbb{C}$,

$$
\begin{aligned}
A \, v_\lambda(x) &= \left(g(x) \frac{\mathrm{d}}{\mathrm{d}x} + h(x) \right) \mathrm{e}^{H(x) + \lambda G(x)} \\
&= g(x) \, \mathrm{e}^{H(x) + \lambda G(x)} \frac{(-h(x) + \lambda)}{g(x)} + h(x) \, \mathrm{e}^{H(x) + \lambda G(x)} = \lambda \, v_\lambda(x). \quad (4.4)
\end{aligned}
$$

We can therefore apply the generalized Prony method to recover (4.1), and with the operator A in (4.3) this leads to a recovery scheme that involves the samples

$$F\Big(\Big(g(\cdot)\frac{\mathrm{d}}{\mathrm{d}x}+h(\cdot)\Big)^k f\Big), \qquad k=0,\ldots,2M-1.$$

However, these samples may be difficult to provide.

We therefore apply the GOP approach with $\varphi(z)=\exp(\tau z)$. For f of the form (4.1) it follows that

$$
\begin{aligned}
\mathrm{e}^{\tau A}f(x) &= \mathrm{e}^{\tau(g(\cdot)\frac{\mathrm{d}}{\mathrm{d}x}+h(\cdot))}f(x) = \sum_{\ell=0}^{\infty}\frac{\tau^{\ell}}{\ell!}\Big(g(\cdot)\frac{\mathrm{d}}{\mathrm{d}x}+h(\cdot)\Big)^{\ell}\Big(\sum_{j=1}^{M}c_j\,\mathrm{e}^{H(x)+\lambda_j G(x)}\Big) \\
&= \sum_{j=1}^{M}c_j\Big(\sum_{\ell=0}^{\infty}\frac{\tau^{\ell}}{\ell!}\lambda_j^{\ell}\Big)\mathrm{e}^{H(x)+\lambda_j G(x)} = \sum_{j=1}^{M}c_j\,\mathrm{e}^{\lambda_j\tau}\,\mathrm{e}^{H(x)+\lambda_j G(x)} \\
&= \mathrm{e}^{H(x)-H(G^{-1}(\tau+G(x)))}\sum_{j=1}^{M}c_j\,\mathrm{e}^{H(G^{-1}(\tau+G(x)))+\lambda_j G(G^{-1}(\tau+G(x)))} \\
&= \mathrm{e}^{H(x)-H(G^{-1}(\tau+G(x)))}\,f(G^{-1}(\tau+G(x))). \qquad (4.5)
\end{aligned}
$$

Thus, the iteration operator $\varphi(A)$ of A is the generalized shift operator $S_{G,H,\tau}:C(\mathbb{R})\to C(\mathbb{R})$ with

$$S_{G,H,\tau}f(x) := \varphi(A)f(x) = \mathrm{e}^{\tau A}f(x) = \mathrm{e}^{H(x)-H(G^{-1}(\tau+G(x)))}\,f(G^{-1}(\tau+G(x))). \quad (4.6)$$

This observation enables us to reconstruct f in (4.1) using function values instead of derivative values.

Theorem 4.1. *Let $G:I\to J\subset\mathbb{R}$ be in $C^{\infty}(I)$ with $|G'(x)|>0$ for all $x\in I$, and $H\in C^{\infty}(I)$. Further, for some fixed $x_0\in I$ and $0<|\tau|\le\pi/C$ let $\tau k+G(x_0)\in G(I)$ for $k=0,\ldots,2M-1$, where $G(I):=\{g(x):x\in I\}$ denotes the image of G. Then f in (4.1) with $|Im\,\lambda_j|\le C$ can be uniquely reconstructed from the function samples $f(G^{-1}(\tau k+G(x_0)))$, $k=0,\ldots,2M-1$.*

Proof. Taking the differential operator A in (4.3) with g and h as in (4.2), it follows from (4.4) that $\mathrm{e}^{H(x)+\lambda_j G(x)}$ are eigenfunctions of A to the pairwise distinct eigenvalues λ_j. As shown in (4.5), we can apply $\varphi(z)=\exp(\tau z)$ and obtain the generalized shift operator $\varphi(A)=S_{G,H,\tau}$ in (4.6). One important consequence of the computations in (4.5) is the observation that also

$$\varphi(A)^k f = \mathrm{e}^{\tau k A}f = \exp\Big(\tau k\Big(g(\cdot)\frac{\mathrm{d}}{\mathrm{d}x}+h(\cdot)\Big)\Big)f = S_{G,H,k\tau}f$$

holds. Therefore, we have $S_{G,H,\tau}^k = S_{G,H,k\tau}$, see also [26] for a different proof. We apply now Theorem 3.10 to f in (4.1) with the operator $\varphi(A)=S_{G,H,\tau}$, the point evaluation functional $F(f)=f(x_0)$, and with $F_k(f):=F(\varphi(A)^k f)$. By Theorem 3.4, the eigenfunctions $\mathrm{e}^{H(x)+\lambda_j G(x)}$ of $A=g(\cdot)\frac{\mathrm{d}}{\mathrm{d}x}+h(\cdot)$ to the eigenvalues λ_j are also eigenfunctions of $S_{G,H,\tau}$, now to the eigenvalues $\mathrm{e}^{\lambda_j\tau}$. We only need to pay attention that these new eigenvalues are pairwise distinct. Since $\lambda_j\in\mathbb{R}+\mathrm{i}[-C,C]$, this is satisfied if $0<\tau\le\frac{\pi}{C}$. Therefore the mapping from $\mathrm{e}^{\lambda_j\tau}$ to $v_{\lambda_j}=\mathrm{e}^{H(\cdot)+\lambda_j G(\cdot)}$ is bijective. Finally, $F(v_{\lambda_j})=v_{\lambda_j}(x_0)=\mathrm{e}^{H(x_0)+\lambda_j G(x_0)}\neq 0$. Hence, the sampling matrix

$$(F(\varphi(A)^{k+\ell}f))_{k,\ell=0}^{M-1,M} = ((S_{G,H,\tau(k+\ell)}f)(x_0))_{k,\ell=0}^{M-1,M}$$

$$= \left(\mathrm{e}^{H(x)-H(G^{-1}(\tau(k+\ell)+G(x_0)))} f(G^{-1}(\tau(k+\ell)+G(x_0))) \right)_{k,\ell=0}^{M-1,M}$$

is admissible by Lemma 3.9 and is already determined by the well-defined sampling values $f(G^{-1}(\tau k + G(x_0)))$, $k = 0, \ldots, 2M - 1$. Thus, Theorem 3.10 can be applied and the assertion follows. $\qquad\square$

Remark 4.2. If the generalized shift operator $S_{G,H,\tau}$ is used to recover the expansion f in (4.1), then the assumptions on G and H can be relaxed. It is sufficient to have continuous functions G and H, where G is monotone on I.

Example 4.3. We want to recover an expansion of the form

$$f(x) = \sum_{j=1}^{M} c_j \, \mathrm{e}^{\lambda_j \cos(x)} \tag{4.7}$$

and have to find the parameters $c_j \in \mathbb{C} \setminus \{0\}$ and $\lambda_j \in \mathbb{R} + \mathrm{i}[-\pi, \pi)$ by employing Theorem 4.1. We take $G(x) := \cos(x)$ which is monotone on $[0, \pi]$, i.e., we can choose $I = [0, \pi]$ and $G(I) = [-1, 1]$. Then, $G : I \to G(I)$ is bijective, and $G^{-1}(x) = \arccos(x)$ is well-defined as a function from $G(I)$ onto I. Further, let $H(x) := 0$. Taking $g(x) := \frac{1}{G'(x)} = \frac{-1}{\sin x}$ and $h(x) := 0$, we conclude that the functions $\mathrm{e}^{\lambda_j \cos(x)}$ in the expansion (4.7) are eigenfunctions of the differential operator $A = -\frac{1}{\sin(x)} \frac{\mathrm{d}}{\mathrm{d}x}$. We apply $\varphi(z) = \exp(\tau z)$ and obtain the generalized shift operator of the form

$$\varphi(A)f(x) = S_{\cos,0,\tau}f(x) = f(\arccos(\tau + \cos(x))).$$

We choose $x_0 = 0$, i.e., $G(x_0) = 1$, and $\tau = -\frac{1}{M}$ such that the values $\cos(x_0) + k\tau = 1 - k/M \in G(I)$ for $0 \ldots, 2M - 1$. Thus

$$S_{\cos,0,\tau}^k f(x_0) = S_{\cos,0,k\tau}f(0) = f(\arccos(k\tau + 1)), \qquad k = 0, \ldots, 2M - 1,$$

are well-defined. According to Theorem 4.1, $f(x)$ in (4.7) is already completely described by these values. In this case, $\mathrm{e}^{\lambda_j \cos(x)}$ are eigenfunctions to $S_{\cos,0,\tau}$ corresponding to the eigenvalues $\mathrm{e}^{\lambda_j \tau}$. Therefore, defining the Prony polynomial

$$P_{\cos}(z) = \prod_{j=1}^{M}(z - \mathrm{e}^{\lambda_j \tau}) = \sum_{\ell=0}^{M} p_\ell \, z^\ell$$

we find with (4.7)

$$\sum_{\ell=0}^{M} p_\ell \, f(\arccos(1 + (m+\ell)\tau)) = \sum_{\ell=0}^{M} p_\ell \sum_{j=1}^{M} c_j \, \mathrm{e}^{\lambda_j(\cos(\arccos(1+(m+\ell)\tau)))}$$

$$= \sum_{j=1}^{M} c_j \mathrm{e}^{\lambda_j(1+m\tau)} \sum_{\ell=0}^{M} p_\ell \, \mathrm{e}^{\lambda_j \ell \tau} = 0$$

for $m = 0, \ldots, M - 1$. This homogeneous linear system provides the coefficients p_0, ..., p_{M-1}, and $p_M = 1$ of $P_{\cos}(z)$. Having found $P_{\cos}(z)$, we can extract its zeros $\mathrm{e}^{\lambda_j \tau}$, recover λ_j and finally find c_j by solving a linear system for the given function values. $\square$

Example 4.4. We want to recover an expansion into shifted Gaussians of the form

$$f(x) = \sum_{j=1}^{M} c_j \, e^{-\alpha(x-\lambda_j)^2}, \tag{4.8}$$

where we assume that $\alpha \in \mathbb{R} \setminus \{0\}$ is given beforehand, and we need to reconstruct $c_j \in \mathbb{C} \setminus \{0\}$ and $\lambda_j \in \mathbb{R}$, $j = 1, \ldots, M$. By direct comparison we have $e^{-\alpha(x-\lambda_j)^2} = e^{\lambda_j^2} \, e^{H(x)+\lambda_j G(x)}$ with

$$H(x) = -\alpha x^2, \qquad G(x) = 2\alpha x,$$

and with the linear factor $e^{\lambda_j^2}$. Thus, taking $g(x) := 1/G'(x) = 1/(2\alpha)$ and $h(x) := H'(x)/G'(x) = -x$, it follows that $v_{\lambda_j}(x) = e^{-\alpha(x-\lambda_j)^2}$ satisfies the differential equation

$$\left(\frac{1}{2\alpha} \frac{\mathrm{d}}{\mathrm{d}x} - x \right) v_{\lambda_j}(x) = \lambda \, v_{\lambda_j}(x),$$

i.e., $e^{\alpha(x-\lambda_j)^2}$ are eigenfunctions of the operator A in (4.3) with g and h as above. According to Theorem 4.1 we can therefore recover the expansion into shifted Gaussians in (4.8) using the function samples

$$f(G^{-1}(k\tau + G(0)) = f\left(\frac{\tau}{2\alpha} k \right), \qquad k = 0, \ldots, 2M-1,$$

where we have taken $x_0 = 0$ and arbitrary real step size $\tau \neq 0$, since $G(x)$ is monotone on $\mathbb{R}$ and the eigenvalues $e^{\lambda_j \tau}$ are real, see also [26], Section 4.1. $\square$

Remark 4.5. We mention that there are other approaches to recover expansions into shifted Gaussians, see e.g. [34]. When one is interested in approximation of functions by sparse sums of the form (4.8), the question occurs, whether arbitrarily narrow Gauss pulses be constructed by linearly combining arbitrarily wider Gauss pulses. This question has been recently discussed in [13].

The approach to consider eigenfunctions of the form $v_\lambda(x) = e^{H(x)+\lambda G(x)}$ for differentiable functions $G(x)$ and $H(x)$, where $G(x)$ is strictly monotone on some interval I opens the way to recover many different expansions of the form (4.1) using only special function values of f. In Table 1, we summarize some examples for $g(x)$, $G(x)$, and arbitrary $H(x)$ (resp. $h(x)$), the corresponding eigenfunctions v_λ as well as the needed function samples for GOP.

4.2 Second order differential operators and generalized symmetric shifts

We consider now the reconstruction problem to find all parameters $c_j \in \mathbb{C} \setminus \{0\}$ and $\lambda_j \in [0, C)$ of

$$f(x) = \sum_{j=1}^{M} c_j \, \cos(\lambda_j \, G(x)). \tag{4.9}$$

As before, we assume that $G \in C^\infty(I)$ for some interval $I = [a, b] \subset \mathbb{R}$ and that G' is strictly positive (or strictly negative) on I. Let $g(x) := 1/G'(x)$. We consider now the special differential operator of second order acting on $f(x)$ as follows

$$Bf(x) := A^2 f(x) = \left(\left(g(\cdot) \frac{\mathrm{d}}{\mathrm{d}x} \right)^2 f \right)(x) = (g(x))^2 f''(x) + g(x) \, g'(x) f'(x). \tag{4.10}$$

$g(x)$	$G(x)$	eigenfunctions v_λ	sampling values
$1/x$	$-\frac{1}{2}x^2$	$\exp(H(x) - \frac{\lambda}{2}x^2)$	$f\left(\sqrt{-k\tau + x_0}\right)$
1	x	$\exp(H(x) + \lambda x)$	$f(k\tau + x_0)$
x	$\log(x)$	$\mathrm{e}^{H(x)}\, x^\lambda$	$f\left(\mathrm{e}^{k\tau} x_0\right)$
$x^p\ (p \neq 1)$	$\frac{x^{1-p}}{1-p}$	$\exp(H(x) + \lambda x^{1-p}/(1-p))$	$f((1-p)\tau k + x_0^{1-p})^{1/1-p})$
$-\sqrt{1-x^2}$	$\arccos(x)$	$\exp(H(x) + \lambda \arccos(x))$	$f(\cos(k\tau + \arccos(x_0)))$
$\sqrt{1-x^2}$	$\arcsin(x)$	$\exp(H(x) + \lambda \arcsin(x))$	$f\left(\sin(k\tau + \arcsin(x_0))\right)$
$\sqrt{x^2-1}$	$\mathrm{arcosh}\,(x)$	$\exp(H(x) + \lambda \,\mathrm{arcosh}\,(x))$	$f\left(\cosh(k\tau + \mathrm{arcosh}(x_0))\right)$
$\sqrt{x^2+1}$	$\mathrm{arsinh}\,(x)$	$\exp(H(x) + \lambda \,\mathrm{arsinh}\,(x))$	$f\left(\sinh(k\tau + \mathrm{arsinh}(x_0))\right)$
$\frac{1}{\cos(x)}$	$\sin(x)$	$\exp(H(x) + \lambda \sin(x))$	$f\left(\arcsin(k\tau + \sin(x_0))\right)$
$-\frac{1}{\sin(x)}$	$\cos(x)$	$\exp(H(x) + \lambda \cos(x))$	$f\left(\arccos(k\tau + \cos(x_0))\right)$
$-\frac{1}{\cosh(x)}$	$\sinh(x)$	$\exp(H(x) + \lambda \sinh(x))$	$f\left(\mathrm{arsinh}(k\tau + \sinh(x_0))\right)$
$-\frac{1}{\sinh(x)}$	$\cosh(x)$	$\exp(H(x) + \lambda \cosh(x))$	$f\left(\mathrm{arcosh}(k\tau + \cosh(x_0))\right)$

Table 1 *Examples of operators $A = g(\cdot)\frac{\mathrm{d}}{\mathrm{d}x} + h(\cdot)$, corresponding eigenfunctions $v_\lambda = \exp(H(\cdot) + \lambda\, G(\cdot))$ and sampling values for $k = 0, \ldots, 2M-1$ with sampling parameter τ to recover expansions f in (4.1).*

Similarly as in (4.4), we observe that the functions $\mathrm{e}^{\mathrm{i}\lambda G(x)}$ and $\mathrm{e}^{-\mathrm{i}\lambda G(x)}$ are the two eigenfunctions of B to the eigenvalue $-\lambda^2$. Therefore, also $\cos(\lambda\, G(x))$ and $\sin(\lambda\, G(x))$ are eigenfunctions of B to $-\lambda^2$.

In order to ensure that the map from eigenvalues to eigenfunctions $-\lambda^2 \to v_\lambda$ is bijective, we restrict ourselves to the eigenfunctions $\cos(\lambda\, G(x))$ with $\lambda \geq 0$.

Then, the function f in (4.9) can be understood as an expansion into eigenfunctions $\cos(\lambda_j\, G(x))$ of the operator B in (4.10), and according to the generalized Prony method in Theorem 3.1, we can reconstruct f using the values $F\left((g(\cdot)\frac{\mathrm{d}}{\mathrm{d}x})^{2k} f\right)$, $k = 0, \ldots, 2M-1$ with some suitable functional $F : C^\infty(I) \to \mathbb{C}$.

We want to apply GOP to derive a simpler reconstruction scheme. We take the analytic function $\varphi(z) = \cos(\tau z^{1/2})$ and obtain for f in (4.9) according to (4.5)

$$
\begin{aligned}
\varphi(B)f(x) &= \varphi(A^2)f(x) = \cos(\tau A)f(x) \\
&= \frac{1}{2}\left[\exp\left(\tau g(\cdot)\frac{\mathrm{d}}{\mathrm{d}x}\right) + \exp\left(-\tau g(\cdot)\frac{\mathrm{d}}{\mathrm{d}x}\right)\right]f(x) \\
&= \frac{1}{2}\left[f(G^{-1}(\tau + G(x))) + f(G^{-1}(-\tau + G(x)))\right].
\end{aligned}
$$

Thus, we find here a symmetric generalized shift operator

$$
S_{G,\tau}^{sym} f := \frac{1}{2}\left[f(G^{-1}(\tau + G(\cdot))) + f(G^{-1}(-\tau + G(\cdot)))\right]
$$

as an iteration operator of B, and f in (4.9) can also be understood as a sparse expansion into eigenfunctions of the operator $S_{G,\tau}^{sym}$ to the eigenvalues $\varphi(-\lambda_j^2) = \cos(\tau\lambda_j)$. This observation enables us to reconstruct f in (4.9) using only function values of f instead of derivative values.

Theorem 4.6. *Let $G : I \to J \subset \mathbb{R}$ be in $C^\infty(I)$ with $|G'(x)| > 0$ for all $x \in I$. Assume further, that for some fixed $x_0 \in I$ we have $\cos(\lambda G(x_0)) \neq 0$ for all $\lambda \in [0, C)$, and for a fixed τ with $0 < |\tau| \leq \pi/C$ we have $\tau k + G(x_0) \in G(I)$ for $k = -2M+1, \ldots, 2M-1$. Then the parameters $c_j \in \mathbb{C}\backslash\{0\}$ and $\lambda_j \in [0, C)$, $j = 1, \ldots, M$, of f in (4.9) can be uniquely reconstructed from the samples $f(G^{-1}(\tau k + G(x_0)))$, $k = -2M + 1, \ldots, 2M - 1$.*

Proof. We apply Theorem 3.10, where we use the operator $\varphi(B) = \cos(\tau A) = S_{G,\tau}^{sym}$, the point evaluation functional $Ff = f(x_0)$, and the set of sampling functionals $F_k = F(\varphi(B)^k)$, $k = 0, \ldots, M - 1$. From Theorem 3.4 it follows that the eigenfunctions $\cos(\lambda_j G(x))$ of B in (4.10) are also eigenfunctions of $S_{G,\tau}^{sym}$. Indeed, we find by direct computation

$$
\begin{aligned}
S_{G,\tau}^{sym} \cos(\lambda_j G(x)) &= \frac{1}{2}\left[\cos(\lambda_j G(G^{-1}(\tau + G(x)))) + \cos(\lambda_j G(G^{-1}(-\tau + G(x))))\right] \\
&= \frac{1}{2}\left[\cos(\lambda_j(\tau + G(x))) + \cos(\lambda_j(-\tau + G(x)))\right] \\
&= \cos(\lambda_j\tau)\,\cos(\lambda_j G(x)).
\end{aligned}
$$

Therefore, the eigenvalues have here the form $\cos(\lambda_j\tau)$ and are pairwise different for $\lambda_j \in [0, C)$ if $0 < \tau < \frac{\pi}{C}$. Further, the sampling matrix $(F_k(\varphi(B)^\ell f))_{k,\ell=0}^{M-1,M}$ is admissible by Lemma 3.9. This sampling matrix has Hankel structure and is determined by

$$
F_k(f) = F((S_{G,\tau}^{sym})^k f) = ((S_{G,\tau}^{sym})^k f)(x_0) = \frac{1}{2^k}\sum_{r=0}^k \binom{k}{r} f(G^{-1}(G(x_0) + (k - 2r)\tau))
$$

for $k = 0, \ldots, 2M - 1$. Thus the assertion follows. $\square$

Example 4.7. We want to reconstruct expansions of the form

$$
f(x) = \sum_{j=1}^M c_j \, \cos(\lambda_j \, \arccos(x)) \tag{4.11}
$$

with $c_j \in \mathbb{C} \setminus \{0\}$ and $\lambda_j \in [0, C)$. Therefore, we choose $G(x) := \arccos(x)$ on the interval $[-1, 1]$, and $g(x) := 1/G'(x) = -(1 - x^2)^{1/2}$. According to our observations we take $Af(x) = g(x)f'(x) = -\sqrt{1 - x^2}\,f'(x)$ and

$$
Bf(x) = A^2 f(x) = \left(\sqrt{1 - (\cdot)^2}\,\frac{\mathrm{d}}{\mathrm{d}x}\right)^2 f(x) = (1 - x^2)\,f''(x) - xf(x)
$$

on $I = [-1, 1]$. Then, B possesses the eigenfunctions $\cos(\lambda \arccos x)$ for $\lambda \geq 0$. Taking the non-negative integers $\lambda = n \in \mathbb{N}_0$, we particularly obtain the Chebyshev polynomials $T_n(x) = \cos(n \arccos x)$. According to Theorem 4.6 we can now reconstruct the expansion (4.11) using only the samples $((S_{\arccos,\tau}^{sym})^k f)(x_0)$, $k = 0, \ldots, 2M - 1$, which can be computed from the values

$$
f(\cos(k\tau + \arccos(x_0))), \qquad k = -2M + 1, \ldots, 2M - 1.
$$

We can choose $x_0 = 1$ to ensure that $\cos(\lambda\,G(x_0)) = \cos(\lambda \arccos(1)) = 1 \neq 0$ for all $\lambda \in [0, C)$. Further, we take $\tau \in (0, \min\{\frac{\pi}{C}, \frac{\pi}{2M}\})$ such that $k\tau + \arccos x_0 = k\tau \in [0, \pi)$ for $k = 0, \ldots, 2M - 1$. In this special case the values $f(\cos(k\tau))$, $k = 0, \ldots, 2M - 1$, are sufficient for full recovery since the cosine function is symmetric. Different approaches to recover expansions into Chebyshev polynomials are taken in [30] and [26]. $\square$

5 Generalized sampling for the Prony method

In this section we study admissible sampling schemes in GOP in more detail and want to give some special applications.

Let us assume that the normed vector space V is a subspace of $L^2([a, b])$ and fix the linear operator $A : V \to V$. We denote with $\sigma(A)$ a fixed set of pairwise different eigenvalues of A and consider the set V_σ of corresponding eigenvectors such that the map $\lambda \to v_\lambda$ is a bijective map from $\sigma(A)$ onto V_σ. By Theorem 3.10 we know that A can be replaced by an iteration operator $\varphi(A)$.

In this section we will focus on finding an admissible set $\{F_k\}_{k=0}^{M-1}$ of sampling functionals according to Definition 3.6 such that entries of the sampling matrix $(F_k(A^\ell f))_{k,\ell=0}^{M-1,M}$ can be simply computed. We recall that a set of sampling functionals $F_k : V \to \mathbb{C}$ is admissible if $(F_k(v_\lambda))_{k=0,\lambda \in \Lambda_M}^{M-1}$ has full rank M for all subsets $\Lambda_M \subset \sigma(A)$ with cardinality M. Then it follows by Theorem 3.7 that the sampling matrix $(F_k(A^\ell f))_{k,\ell=0}^{M-1,M}$ has full rank M for each $f \in \mathcal{M}(A)$ such that f can be uniquely recovered.

We consider functionals $F_k : \mathcal{M}(A) \to \mathbb{C}$ which can be written as

$$F_k(f) := \langle f, \phi_k \rangle = \int_a^b f(x)\, \phi_k(x)\, \mathrm{d}x, \tag{5.1}$$

where $(a, b) \subseteq \mathbb{R}$ is a suitable interval and ϕ_k is some kernel function or distribution, such that the integral in (5.1) is well-defined in a distribution sense. For example, we can take ϕ_k to be the δ-distribution,

$$F_k(f) := \langle f, \delta(\cdot - x_0) \rangle = \int_a^b f(x)\, \delta(\cdot - x_0)\, \mathrm{d}x = f(x_0), \qquad x_0 \in [a, b].$$

Using the adjoint operator, the entries of the sampling matrix can be written as

$$F_k(A^\ell f) = \langle A^\ell f, \phi_k \rangle = \langle f, (A^*)^\ell \phi_k \rangle = \int_a^b f(x)\, (A^*)^\ell\, \phi_k(x)\, \mathrm{d}x. \tag{5.2}$$

If A is a linear differential operator, the consideration of powers of the adjoint operator A^* applied to ϕ_k is particularly useful, if we cannot acquire derivative samples of f but special moments instead. In this case, we need to assume that the kernel functions ϕ_k are sufficiently smooth on $[a, b]$, such that $(A^*)^\ell \phi_k \in L^2([a, b])$. For admissibility we need now to ensure that $(\langle v_\lambda, \phi_k \rangle)_{k=0,\lambda \in \Lambda_M}^{M-1}$ has full rank M.

Example 5.1. We consider again the example of exponential sums to present the variety of possible sampling matrices that can be used. Let

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

with $c_j \in \mathbb{C} \setminus \{0\}$, $T_j \in \mathbb{R} + \mathrm{i}[-\pi, \pi)$, where $\mathrm{e}^{T_j x}$ are eigenfunctions of $A = \frac{\mathrm{d}}{\mathrm{d}x}$ to the eigenvalue T_j. Here $\mathcal{M}(A)$ is a subset of the Schwartz space, and thus obviously a subspace of $L^2([a, b])$ for each interval $[a, b]$ and also of $L^2(\mathbb{R})$. We present a variety of sampling schemes which are all admissible and of the form (5.2).

a) Let $F(f) := \int_{-\infty}^\infty f(x)\, \delta(x - x_0)\, \mathrm{d}x = f(x_0)$ be the point evaluation functional with $x_0 \in \mathbb{R}$ and let $F_k(f) := F(A^k f)$. Then the entries of the sampling matrix are of the form

$$F_k(A^\ell f) = \int_{-\infty}^\infty A^\ell f(x)\, A^k \delta(x - x_0)\, \mathrm{d}x = \int_{-\infty}^\infty f(x)\, \delta^{(k+\ell)}(x - x_0)\mathrm{d}x = f^{(k+\ell)}(x_0)$$

used in Section 2.2, where we need derivative values $f^{(k)}(x_0)$, $k = 0, \ldots, 2M-1$. The used kernel functions are in this case the distributions $\phi_k = A^k \delta(\cdot - x_0) = \delta^{(k)}(\cdot - x_0)$, i.e., derivatives of the Delta distribution. Admissibility is ensured since for any $T_j \in \mathbb{R} + \mathrm{i}[-\pi, \pi)$,

$$(\langle \mathrm{e}^{T_j \cdot}, \phi_k \rangle)_{k=0, j=1}^{M-1, M} = (T_j^k \mathrm{e}^{T_j x_0})_{k=0, j=1}^{M-1, M} = (T_j^k)_{k=0, j=1}^{M-1, M} \operatorname{diag}(\mathrm{e}^{T_j x_0})_{j=1}^{M}$$

has full rank M.

b) By Lemma 3.9 we can also take $F_k(f) = F(\psi(A)^k f)$ for some iteration operator $\psi(A)$ with F as in a). With $\psi(A) = \exp(\tau A) = S_\tau$, $\tau \neq 0$, see Example 3.5, we obtain the admissible sampling matrix with entries

$$
\begin{aligned}
F_k(A^\ell f) &= \int_{-\infty}^{\infty} (S_\tau^k A^\ell f)(x)\, \delta(x - x_0)\, \mathrm{d}x \\
&= \int_{-\infty}^{\infty} f(x)\, ((A^\ell)^* (S_\tau^k)^* \delta)(x - x_0)\, \mathrm{d}x \\
&= \int_{-\infty}^{\infty} f(x)\, \delta^{(\ell)}(x - \tau k - x_0)\, \mathrm{d}x = f^{(\ell)}(x_0 + \tau k),
\end{aligned}
$$

where we need the values $f^{(\ell)}(x_0 + k\tau)$, $\ell = 0, \ldots, M$, $k = 0, \ldots, M-1$, see Section 2.4. We have here $\phi_k = (S_\tau^k)^* \delta(\cdot - x_0) = \delta(\cdot - \tau k - x_0)$, $k = 0, \ldots, M-1$.

c) Consider now the functional

$$F(f) := \int_0^1 f(x)\, \phi(x)\, \mathrm{d}x \tag{5.3}$$

with $\phi(x) := x^{2M}(1 - x)^{2M}$. Then

$$F(\mathrm{e}^{T \cdot}) = \int_0^1 \mathrm{e}^{Tx}\, \phi(x)\mathrm{d}x \neq 0$$

for all $T \in \mathbb{R} + \mathrm{i}[-\pi, \pi)$ since $\phi(x) > 0$ for $x \in (0, 1)$. Thus, with $F_k := F(A^k)$ we obtain

$$
\begin{aligned}
F_k(A^\ell f) &= F(A^{k+\ell} f) = \int_0^1 f^{(k+\ell)}(x)\, x^{2M}(1 - x)^{2M}\, \mathrm{d}x \\
&= (-1)^{k+\ell} \int_0^1 f(x)\, [x^{2M}(1 - x)^{2M}]^{(k+\ell)}\, \mathrm{d}x
\end{aligned}
$$

is admissible. These values can be computed from the moments $\int_0^1 f(x) x^s\, \mathrm{d}x$ for $s = 0, \ldots, 4M$. The functions ϕ_k are here defined as $\phi_k := \phi^{(k)}$, $k = 0, \ldots, M-1$.

d) Let us now take the functional F as in (5.3), but with $\phi(x) := x^M(1 - x)^M$ and let $F_k(f) := F(\exp(kA)f) = F(S_1^k f)$ according to Lemma 3.9. Then we get the entries of the admissible sampling matrix in the form

$$
\begin{aligned}
F_k(A^\ell f) &= \int_0^1 f^{(\ell)}(x + k)\, x^M(1 - x)^M\, \mathrm{d}x = (-1)^\ell \int_0^1 f(x + k)\, [x^M(1 - x)^M]^{(\ell)}\, \mathrm{d}x \\
&= (-1)^\ell \int_k^{k+1} f(x)\, [(x - k)^M(k + 1 - x)^M]^{(\ell)}\, \mathrm{d}x.
\end{aligned}
$$

These entries can be computed from the moments $\int_0^1 f(x+k)\, x^s \mathrm{d}x$ for $k = 0, \ldots, M-1$ and $s = 0, \ldots, 2M$. The functions ϕ_k are of the form $\phi_k(x) = (x - k)^M(k + 1 - x)^M$, $k = 0, \ldots, M-1$.

e) Besides all the sampling schemes above, we know from Section 2.1 that f can be reconstructed using the $2M$ samples $f(x_0 + k\tau)$, $k = 0, \ldots, 2M - 1$, with $x_0 \in \mathbb{R}$, $\tau \neq 0$. This sampling scheme also follows from Theorem 3.10 by replacing A by the iteration operator $\exp(\tau A) = S_\tau$. The simple equidistant sampling is obtained by taking $F_k = F(S_\tau^k)$ and the kernel function $\phi(x) = \delta(x - x_0)$ as in a), such that

$$F_k((\exp(\tau A))^\ell f)) = F(S^{k+\ell} f) = f(x_0 + (k + \ell)\tau).$$

The kernel functions ϕ_k are here $\phi_k = \phi(\cdot - \tau k)$, $k = 0, \ldots, M - 1$. Taking instead $F_k = F(S_{2\tau}^k)$ we arrive at

$$F_k(S_\tau^\ell f) = F(S_{2\tau}^k S_\tau^\ell f) = f(x_0 + \tau(2k + \ell)), \quad k = 0, \ldots, M - 1, \ell = 0, \ldots, M,$$

and also this sampling matrix is admissible by Lemma 3.9. Here we have now $\phi_k = \phi(\cdot - 2\tau k)$, $k = 0, \ldots, M - 1$. $\square$

Besides the well-known example of exponential sums, we can also find new sampling schemes for expansions into eigenfunctions of differential operators of higher order, where we need to acquire moments instead of derivative values. This can be always achieved by employing suitable kernels ϕ_k and the adjoint operator representation in (5.2).

Let us consider the linear differential operator

$$A := \sum_{n=0}^{d} g_n(\cdot) \frac{\mathrm{d}^n}{\mathrm{d}x^n} \tag{5.4}$$

of order d with sufficiently smooth functions g_n. Further, let $\sigma(A)$ be a subset of pairwise distinct eigenvalues λ of A with corresponding eigenfunctions $v_\lambda \in L^2([a, b])$ such that we have a bijection $\lambda \to v_\lambda$.

Lemma 5.2. *Let A be an operator in (5.4) with $g_n \in C^d([a, b])$ for $n = 0, \ldots, d$, and let $F : L^2([a, b]) \to \mathbb{C}$ be a functional given by $Ff = \langle f, \phi \rangle$, where $\phi \in C^d([a, b])$ and*

$$\lim_{x \to a} \phi^{(\ell)}(x) = \lim_{x \to b} \phi^{(\ell)}(x) = 0, \qquad \ell = 0, \ldots, d.$$

Then

$$F(Af) = \langle Af, \phi \rangle = \left\langle f, \sum_{n=0}^{d} (-1)^r \sum_{\ell=0}^{r} \binom{r}{\ell} g_n^{(\ell)} \phi^{(r-\ell)} \right\rangle,$$

where $g_n^{(\ell)}$ and $\phi^{(\ell)}$ denote the ℓ-th derivative of g_n and ϕ, respectively.

Proof. The proof follows simply by partial integration, where the boundary terms vanish because of the assumption on ϕ. $\square$

Thus, we can apply the sampling scheme arising from (5.2) where we need to compute with derivatives of the kernel functions instead of derivatives of f.

Example 5.3 (Sparse Legendre Expansions)**.** We want to recover a sparse expansion into Legendre polynomials of the form

$$f(x) := \sum_{j=1}^{M} c_j P_{n_j}(x)$$

where $c_j \in \mathbb{C} \setminus \{0\}$, and $n_j \in \mathbb{N}_0$ with $0 \leq n_1 < n_2 < \ldots < n_M$. The Legendre polynomials P_n, $n \in \mathbb{N}_0$ are eigenfunctions of the differential operator of second order

$$A f(x) := (x^2 - 1) f''(x) + 2x \, f'(x),$$

and we have

$$A \, P_n = n(n+1) \, P_n.$$

Employing a functional of the form

$$F(f) := \int_a^b f(x) \phi_P(x) \, dx,$$

with a smooth kernel ϕ_P satisfying $\phi_P(a) = \phi_P(b) = 0$ and $\phi_P'(a) = \phi_P'(b) = 0$, it follows that

$$\int_a^b A f(x) \, \phi_P(x) \, dx = \int_a^b f(x) \, A \phi_P(x) \, dx.$$

We choose the kernel

$$\phi_P(x) := \begin{cases} (x-a)^{4M}(x-b)^{4M} \exp\left(-\alpha(x-\beta_0)^2(x-\beta_1)^2\right) & x \in [a, \, b], \\ 0 & x \notin [a, \, b]. \end{cases} \quad (5.5)$$

Here, the parameters β_0 and β_1 are chosen to be outside of the interval $[a, \, b]$, and $\alpha \geq 0$. For $\alpha = 0$, ϕ_P is a polynomial of degree $8M$.

Taking for example $[a, \, b] = [-1/2, \, 3/4]$, it follows that the functional F satisfies the admissibility condition $F(P_n) \neq 0$ for all $n \in \mathbb{N}_0$. Therefore, the expansion f can be recovered from the $2M$ samples

$$F(A^k f) = \int_{-1/2}^{3/4} f(x) \, A^k \phi_P(x) dx, \quad k = 0, \ldots, 2M - 1.$$

We consider a small computational example. We want to recover the parameters c_j and n_j of the expansion

$$f(x) = \sum_{j=1}^{3} c_j \, P_{n_j}(x)$$

from the 6 samples $F(A^k f)$, $k = 0, \ldots, 5$. The true parameters are given in Table 2.

n_j	1	4	9
c_j	1.703	3.193	3.710

Table 2 *Active degrees n_j and the corresponding linear coefficients c_j of f with parameters in Table 2.*

The signal with this parameters is presented in Figure 1.

We choose now the sampling kernel ϕ_P in (5.5) with $a = -1/2$, $b = 3/4$, $\alpha = 0.1$, and $-\beta_0 = \beta_1 = 2$. The kernels $A^k \phi_P$, $k = 0, \ldots, 5$, are depicted in Figure 2. These kernels can now be used for any $3-$sparse linear combination of arbitrary Legendre polynomials. For our example, the sampling matrix has the form

$$\begin{bmatrix} F(f) & F(Af)) & F(A^2 f) & F(A^3 f) \\ F(Af) & F(A^2 f)) & F(A^3 f) & F(A^4 f) \\ F(A^2 f) & F(A^3 f)) & F(A^4 f) & F(A^5 f) \end{bmatrix}.$$
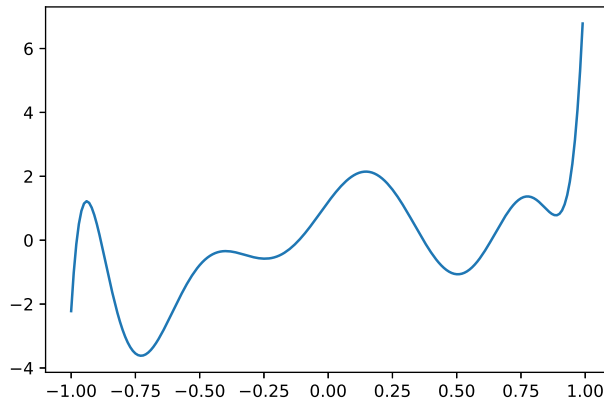
Figure 1 *3-sparse Legendre expansion f with parameters in Table 2.*
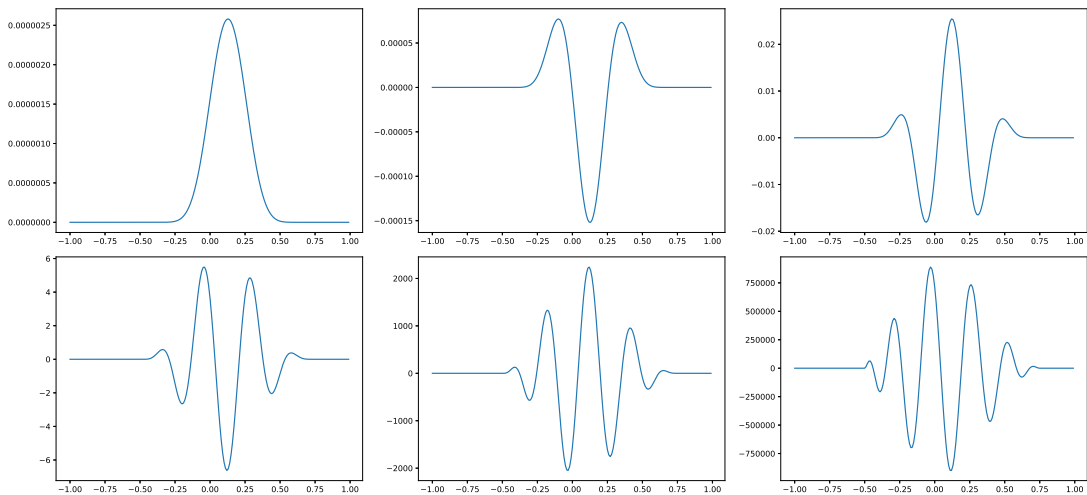

Figure 2 *Sampling kernels $A^k\phi_P$, $k = 0, 1, 2$ (first row) $k = 3, 4, 5$ (second row) for a 3-sparse Legendre expansion.*

The reconstructed parameters can be seen in Table 3. The polynomial degrees are correctly recovered up to small rounding errors. We round to the closest integer and get the exact values n_j. The coefficients c_j are found using a 3×3 Vandermonde system. Alternatively, to recover the coefficients, we can use the orthogonality of Legendre polynomials and obtain

$$c_j = \frac{2n_j + 1}{2} \int_{-1}^{1} f(x) P_{n_j}(x) dx.$$

The numerical instabilities due to the exponentially growing functions $A^k\phi_P$ are an issue in this approach. A clever choice of the parameters of ϕ can help to control the amplitudes of $A^k\phi_P$. Another way is to apply a set of different functionals F_k as proposed in Section 3.3.

n_j	1.00008823	4.00001099	9.00000026
c_j	1.703	3.193	3.710

Table 3 *Computed parameters n_j and c_j for f.*

References

[1] F. Andersson, M. Carlsson, and M.V de Hoop. Sparse approximation of functions using sums of exponentials and AAK theory. *J. Approx. Theory*, 163:213–248, 2011.

[2] G. Baechler, A. Scholefield, L. Baboulaz, and M. Vetterli. Sampling and exact reconstruction of pulses with variable width. *IEEE Trans. Signal Process.*, 65(10):2629–2644, 2017.

[3] D. Batenkov and Y. Yomdin. On the accuracy of solving confluent Prony systems. *SIAM J. Appl. Math.*, 73(1):134–154, 2013.

[4] M. Ben-Or and P. Tiwari. A deterministic algorithm for sparse multivariate polynomial interpolation. Proc. Twentieth Annual ACM Symp. Theory Comput., pages 301–309. ACM Press, New York, 1988.

[5] J. Berent, P.L. Dragotti, and T. Blu. Sampling piecewise sinusoidal signals with finite rate of innovation methods. *IEEE Trans. Signal Process.*, 58(2):613–625, 2010.

[6] G. Beylkin and L. Monzón. On approximation of functions by exponential sums. *Appl. Comput. Harmon. Anal.*, 19:17–48, 2005.

[7] A. Bhandari and Y. C. Eldar. Sampling and super resolution of sparse signals beyond the fourier domain. *IEEE Trans. Signal Process.*, 67(6):1508–1521, 2019.

[8] Y. Bresler and A. Macovski. Exact maximum likelihood parameter estimation of superimposed exponential signals in noise. *IEEE Trans. Acoust., Speech, Signal Process.*, 34(5):1081–1089, 1986.

[9] A. Cuyt and W.-s. Lee. How to get high resolution results from sparse and coarsely sampled data. *Appl. Comput. Harmon. Anal., online first*, 2018.

[10] G. Dattoli and D. Levi. Exponential operators and generalized difference equations. *Nuovo Cimento Soc. Ital. Fis. B (12)*, 115(6):653–662, 2000.

[11] G. Dattoli, P.L. Ottaviani, A. Torre, and L. Vázquez. Evolution operators equations: Integration with algebraic and finite difference methods. applications to physical problems in classical and quantum mechanics and quantum field theory. *Riv. Nuovo Cimento Soc. Ital. Fis. (4)*, 20(2):1–133, 1997.

[12] P.L. Dragotti, M. Vetterli, and T. Blu. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang–Fix. *IEEE Trans Signal Process.*, 55(5):1741–1757, 2007.

[13] P. J. S. G. Ferreira and A. J. Pinho. The natural scale of signals: Pulse duration and superoscillations. International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 4176–4179. IEEE, 2014.

[14] Y. Hua and T.K. Sarkar. On SVD for estimating generalized eigenvalues of singular matrix pencil in noise. *IEEE Trans. Signal Process.*, 39(4):892–900, 1991.

[15] T. Kailath. ESPRIT–estimation of signal parameters via rotational invariance techniques. *Optical Engineering*, 29(4):296, 1990.

[16] L. Li and T. P. Speed. Parametric deconvolution of positive spike trains. *The Annals of Statistics*, 28(5):1279–1301, 2000.

[17] B. Mourrain. Polynomial–exponential decomposition from moments. *Found. Comput. Math.*, 82(3):339, 2017.

[18] M.R. Osborne and G.K. Smyth. A modified Prony algorithm for exponential function fitting. *SIAM J. Sci. Comput.*, 16(1):119–138, 1995.

[19] T. Peter and G. Plonka. A generalized Prony method for reconstruction of sparse sums of eigenfunctions of linear operators. *Inverse Problems*, 29(2), 2013.

[20] T. Peter, G. Plonka, and D. Roşca. Representation of sparse Legendre expansions. *J. Symbolic Comput.*, 50:159–169, 2013.

[21] T. Peter, D. Potts, and M. Tasche. Nonlinear approximation by sums of exponentials and translates. *SIAM J. Sci. Comput.*, 33(4):1920–1947, 2011.

[22] V. F. Pisarenko. The retrieval of harmonics from a covariance function. *Geophys. J. Int.*, 33(3):347–366, 1973.

[23] G. Plonka and V. Pototskaia. Application of the AAK theory for sparse approximation of exponential sums. http://arxiv.org/pdf/1609.09603, 2016.

[24] G. Plonka and V. Pototskaia. Computation of adaptive fourier series by sparse approximation of exponential sums. *J. Fourier Anal. Appl.*, 25(4):1580–1608, 2019.

[25] G. Plonka, D. Potts, G. Steidl, and M. Tasche. *Numerical Fourier Analysis*. Birkhäuser, 2018.

[26] G. Plonka, K. Stampfer, and I. Keller. Reconstruction of stationary and non-stationary signals by the generalized Prony method. *Anal. and Appl.*, 17(2):179–210, 2019.

[27] G. Plonka and M. Tasche. Prony methods for recovery of structured functions. *GAMM Mitt.*, 37(2):239–258, 2014.

[28] G. Plonka and M. Wischerhoff. How many Fourier samples are needed for real function reconstruction? *J. Appl. Math. and Comput.*, 42(1-2):117–137, 2013.

[29] D. Potts and M. Tasche. Parameter estimation for exponential sums by approximate Prony method. *Signal Process.*, 90(5):1631–1642, 2010.

[30] D. Potts and M. Tasche. Sparse polynomial interpolation in Chebyshev bases. *Linear Algebra Appl.*, 441:61–87, 2014.

[31] R.O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Trans. On Antennas and Propagation*, 34(3):276–280, 1986.

[32] C.S. Seelamantula. Opera: Operator-based annihilation for finite-rate-of-innovation signal sampling. In G. Pfander, editor, *Sampling Theory, a Renaissance*, Applied and Numerical Harmonic Analysis, pages 461–484. Birkhäuser, Cham, 2015.

[33] J.A. Urigen, T. Blu, and P.L. Dragotti. FRI sampling with arbitrary kernels. *IEEE Trans. Signal Process.*, 61(21):5310–5323, 2013.

[34] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Trans. Signal Process.*, 50(6):1417–1428, 2002.

[35] R. Zhang and G. Plonka. Optimal approximation with exponential sums by a maximum likelihood modification of Prony's method. *Adv. Comput. Math.*, 45(3):1657–1687, 2019.