Phase Retrieval and System Identification in Dynamical Sampling via Prony's Method

Robert Beinert[1] and Marzieh Hasannasab[1]

[1]Institut für Mathematik
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin, Germany

Correspondence
R. Beinert: beinert@math.tu.berlin.de
M. Hasannasab: hasannas@math.tu-berlin.de

Abstract. Phase retrieval in dynamical sampling is a novel research direction, where an unknown signal has to be recovered from the phaseless measurements with respect to a dynamical frame, i.e. a sequence of sampling vectors constructed by the repeated action of an operator. The loss of the phase here turns the well-posed dynamical sampling into a severe ill-posed inverse problem. In the existing literature, the involved operator is usually completely known. In this paper, we combine phase retrieval in dynamical sampling with the identification of the system. For instance, if the dynamical frame is based on a repeated convolution, then we want to recover the unknown convolution kernel in advance. Using Prony's method, we establish several recovery guarantees for signal and system, whose proofs are constructive and yield analytic recovery methods. The required assumptions are satisfied by almost all signals, operators, and sampling vectors. Moreover, these guarantees not only hold for the finite-dimensional setting but also carry over to infinite-dimensional spaces. Studying the sensitivity of the analytic recovery procedures, we also establish error bounds for the applied approximate Prony method with respect to complex exponential sums.

Keywords. Phase retrieval, dynamical sampling, system identification, Prony's method, Vandermonde matrix.

AMS subject classification. 42A05, 94A12, 15A29, 94A20

1 Introduction

Phase retrieval is an ill-posed inverse problem consisting in the recovery of signals or images from phaseless measurements like the magnitude of the Fourier transform or the absolute values of inner products with respect to given sampling vectors. Phaseless reconstructions appear naturally in many applications like X-ray crystallography [31, 36, 44], astronomy [17, 22], laser optics [53, 54] and audio processing [23, 27, 39]. The mathematical analysis of this ill-posed problem has been studied intensively during the last decades, see for instance [2, 3, 12–14, 17, 30, 37, 38, 55, 59] and references therein.

In this paper, we consider phase retrieval in the context of dynamical sampling. Dynamical sampling is a novel research direction motivated by the work of Vetterli et al. [41, 51] and was introduced in [4, 6, 7, 10]. The topic instantly attracted attention in the scientific community, see for instance [1, 5, 18, 20, 21, 43, 43, 47, 56, 58] for further studies. Formulated in the setting of finite-dimensional spaces, the main question in dynamical sampling is to find conditions on the system $A \in \mathbb{C}^{d \times d}$ and the sampling vectors $\{\phi_i\}_{i=0}^{J-1} \subset \mathbb{C}^d$ such that each signal $x \in \mathbb{C}^d$ can be stably recovered from the spatiotemporal samples

$$\left\{ \langle x, A^\ell \phi_i \rangle \right\}_{\ell,i=0}^{L-1,J-1}$$

or such that $\{A^\ell \phi_i\}_{\ell,i=0}^{L-1,J-1}$ forms a frame for some $L, J \in \mathbb{N}$. Note that many structured measurements like the discrete Gabor transform may be interpreted as dynamical samples. For the Gabor transform, A would be a diagonal matrix corresponding to the modulation operator, and ϕ_i would be shifts of a window function. We refer to [4, 7] for motivations about this particular question.

Different from the classical finite-dimensional dynamical sampling, we consider the phaseless measurements

$$\left\{ |\langle x, A^\ell \phi_i \rangle|^2 \right\}_{\ell,i=0}^{L-1,J-1}$$

for some $L, J \in \mathbb{N}$. The main question is again: under which conditions on A and ϕ_i can x be recovered from the given measurements. Due to the loss of the phase, this problem becomes far more challenging since the recovery is now severely ill posed in advance.

Relation to existing literature Phase retrieval in dynamical sampling has already been studied. In [8, 9], the authors pose conditions on the operator A defined on a real Hilbert space and on the sampling vectors ϕ_i to ensure that the dynamical phase retrieval problem has a unique solution. The main strategy is here to ensure that the sequence $\{A^\ell \phi_i\}_{\ell=i=0}^{L_i-1,J-1}$ has the complementary property meaning that each subset or its complement spans the entire space. The restriction to the real-valued problem is crucial since the complementary property is not sufficient to allow phase retrieval in the complex case. Further, the results are of a theoretical nature, and the question how to recover the signal numerical remains open.

An approach for a numerical recovery procedure based on polarization identities has been considered in [11], where the measurement vectors ϕ_i have been designed to allow phase retrieval. The key idea has been to consider interfering measurement vectors that allow the recovery of the missing phase by polarization such that we obtain a classical dynamical sampling problem, which can be solved in a second step. The presented reconstruction technique works for almost all real or complex signals.

Contributions Besides the recovery of the real or complex signal x, we want to recover the unknown operator A from a certain class in advance. For instance, if the operator $A \coloneqq \operatorname{circ} a$ corresponds to the convolution with a, we want to recover the signal a

or the spectrum $\hat{\boldsymbol{a}}$ of $\boldsymbol{A}$, where $\hat{\cdot}$ denotes the discrete Fourier transform. The theoretical requirements to allow phase retrieval besides system identification is our main contribution and focus of this paper. The combination of phase retrieval, dynamical sampling, and system identification is to our knowledge a new research topic. Our work horse to establish the recovery guarantees for phase and system is Prony's method, which allow us to recover the wanted entities from the given measurements. As a result, all our proofs contain analytic recovery methods. The required assumptions are satisfied by almost all signals, spectra, and sampling vectors. Using several sampling vectors, phase retrieval and system identification is possible from only linearly many samples. The basic idea here generalizes to the infinite-dimensional setting. Moreover, we study the sensitivity of the applied Prony method resulting in error bounds that are interesting by their own outside the context of dynamical sampling. On this basis, we moreover study the sensitivity of the proposed analytic recovery procedures.

Roadmap This paper is organized as follows. In Section 2, we introduce the required notations. In Section 3, we recall Prony's method, and we explain how this method enables us to recover the missing information. In Section 4, we provide conditions to retrieve an unknown signal when the underlying dynamical frame is known. Section 5 is devoted to the system identification in case that the signal $\boldsymbol{x}$ is already known. In Section 6, we suppose that both the signal and the spectrum of $\boldsymbol{A}$ are unknown. In particular, we establish recovery guarantees when the operator $\boldsymbol{A}$ corresponds to a convolution with a low-pass filter as kernel. In Section 7, we consider multiple sampling vectors, which finally allow us to recover both – signal and operator. In Section 8, we adapt our results to the infinite-dimensional setting. The sensitivity of the analytic reconstructions is investigated in Section 9. In Section 10, we provide numerical examples to accompany our theoretical results. Section 11 concludes the paper with a number of final remarks.

2 Preliminary notes

In this section, we introduce the notations and definitions that are needed throughout this paper. All finite-dimensional vectors and matrices are stated in bold print. The zero matrix of dimension $L \times K$ is denoted by $0 := 0_{L,K}$ and the $(d \times d)$-dimensional identity by $\boldsymbol{I} := \boldsymbol{I}_d$. If the dimension is clear within the context, we usually skip the indices.

A matrix $\boldsymbol{A} \in \mathbb{C}^{d \times d}$ is called diagonalizable if there exist an invertible matrix $\boldsymbol{S}$, whose columns consists of eigenvectors of $\boldsymbol{A}$, and a diagonal matrix $\boldsymbol{\Lambda}$ with the eigenvalues of $\boldsymbol{A}$ on its diagonal, such that $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$. Throughout the paper, we always use this eigenvalue decomposition, where $\boldsymbol{S}$ does not have to be orthogonal implying that the columns of $\boldsymbol{S}$ only form a (maybe non-orthogonal) basis. Further, if the eigenvalues are pairwise distinct, we say that a given vector $\boldsymbol{\phi} \in \mathbb{C}^d$ *depends on all eigenspaces* of $\boldsymbol{A}$ if $\boldsymbol{S}^{-1}\boldsymbol{\phi}$ does not vanish anywhere, i.e. if all coordinate to the basis in $\boldsymbol{S}$ are non-zero. Note that in this case $\boldsymbol{S}$ is unique up to permutation and global phase of the columns.

For $\boldsymbol{a} \in \mathbb{C}^d$, we denote by $\operatorname{circ}(\boldsymbol{a})$ the circulant matrix whose first column is $\boldsymbol{a}$. Note that the multiplication with $\operatorname{circ}(\boldsymbol{a})$ results in the convolution with $\boldsymbol{a}$, i.e. $\operatorname{circ}(\boldsymbol{a})\,\boldsymbol{x} = \boldsymbol{a}*\boldsymbol{x}$. All circulant matrices are diagonalizable with respect to the discrete Fourier transform. More precisely, we have $\operatorname{circ}(\boldsymbol{a}) = {}^1\!/_d\,\boldsymbol{F}\operatorname{diag}(\hat{\boldsymbol{a}})\,\boldsymbol{F}^{-1}$, where $\boldsymbol{F} = (\mathrm{e}^{-2\pi \mathrm{i} jk/d})_{j,k=0}^{d-1}$ denotes the Fourier matrix and $\hat{\boldsymbol{a}} \coloneqq \boldsymbol{F}\boldsymbol{a}$ the discrete Fourier transform.

Given a vector $\boldsymbol{\beta} \in \mathbb{C}^K$ and $L \in \mathbb{N}$, we define the rectangular Vandermonde matrix $\boldsymbol{V}_L \in \mathbb{C}^{L\times K}$ by

$$\boldsymbol{V}_L \coloneqq \boldsymbol{V}_L(\boldsymbol{\beta}) \coloneqq (\beta_k^\ell)_{\ell,k=0}^{L-1,K-1}.$$

For $L = K$, we drop the subscript and denote the Vandermonde matrix by $\boldsymbol{V}$ or $\boldsymbol{V}(\boldsymbol{\beta})$.

Recall that the finite-dimensional p-norm is defined as

$$\|\boldsymbol{x}\|_p = \left(\sum_{k=0}^{d-1} |x_k|^p\right)^{1/p} \qquad \text{for} \qquad \boldsymbol{x} \in \mathbb{C}^d \quad \text{and} \quad p \in [1,\infty).$$

Moreover, the maximum norm is defined by $\|\boldsymbol{x}\|_\infty = \max_k |x_k|$. Against this background and for notational convenience, we define the minimum norm $\|\boldsymbol{x}\|_{-\infty} = \min_k |x_k|$ although this expression is clearly no norm.

The non-zero complex numbers are denote by $\mathbb{C}_*$. Without loss of generality, we always choose the phase $\arg(\cdot)$ of a complex number within the interval $[-\pi,\pi)$. Especially for calculations with phases, we denote by $\cdot \bmod 2\pi$ the remainder within $[-\pi,\pi)$, i.e. we add or subtract a multiple of 2π to obtain an number in the considered interval.

For a given vector $\boldsymbol{x} = (x_0,\dots,x_{d-1})$, we call the set of relative phases $\arg(x_j\bar{x}_k)$ the *winding direction* of $\boldsymbol{x}$. Figuratively, the winding direction describes how the phase is changing by traveling through the components of $\boldsymbol{x}$. We say that a vector $\boldsymbol{x}$ can be uniquely recovered up to the winding direction if the relatives phases are only reconstructable up to a global sign. If x_0 is real, a vector with the opposite winding direction can be computed by conjugating all components of $\boldsymbol{x}$, i.e. changing the sign of all relative phases.

Finally, we denote by $\#[\cdot]$ the cardinality of a set.

3 The approximate Prony method

In a nutshell, Prony's method [50] allows us to recover the non-zero coefficients $\eta_k \in \mathbb{C}_*$ and the pairwise distinct bases $\beta_k \in \mathbb{C}_*$ of an exponential sum

$$f(t) \coloneqq \sum_{k=0}^{K-1} \eta_k\,\beta_k^t \tag{1}$$

from the equispaced sampled data $h_\ell \coloneqq f(\ell)$ with $\ell = 0,\dots,2K-1$. The so-called Prony polynomial $P\colon \mathbb{C} \to \mathbb{C}$ is the monic polynomial whose zeros are the unknown bases, i.e.

$P(z) := \sum_{k=0}^{K} \gamma_k z^k = \prod_{k=0}^{K-1}(z - \beta_k)$ with $\gamma_K = 1$. Considering the linear equations

$$\sum_{k=0}^{K} \gamma_k \, h_{\ell+k} = \sum_{j=0}^{K-1} \eta_j \beta_j^{\ell} \, P(\beta_j) = 0, \qquad \ell = 0, \dots, K-1, \tag{2}$$

one may calculate the coefficients γ_k of the Prony polynomial by solving a linear equation system. Knowing the Prony polynomial, we may extract the unknown bases β_k via its roots. The coefficients η_k of the exponential sum are determined by an overdetermined linear equation system. To improve the numerical performance, the number of measurements may be increased [15, 46, 49]. On the basis of the rectangular Hankel matrix

$$\boldsymbol{H} := \left(h_{\ell+k}\right)_{\ell,k=0}^{L-K-1,K} \qquad \text{with} \qquad L \geq 2K, \tag{3}$$

the coefficients of the Prony polynomial are determined by the kernel of $\boldsymbol{H}$.

LEMMA 3.1. *For the exact samples h_ℓ with $\ell = 0, \dots, L-1$, the rectangular Hankel matrix (3) is of rank K, and the following assertions are equivalent:*
(i) *the polynomial $P(z) := \sum_{\ell=0}^{K} \gamma_\ell z^\ell$ has the K distinct roots $\beta_0, \dots, \beta_{K-1}$,*
(ii) *the vector $\boldsymbol{\gamma} := (\gamma_\ell)_{\ell=0}^{K}$ spans $\ker(\boldsymbol{H})$, i.e. $\boldsymbol{H}\boldsymbol{\gamma} = 0$.*

Proof. With $\boldsymbol{\eta} := (\eta_k)_{k=0}^{K-1}$ and $\boldsymbol{\beta} := (\beta_k)_{k=0}^{K-1}$, we may factorize the Hankel matrix (3) into

$$\boldsymbol{H} = \boldsymbol{V}_{L-K}(\boldsymbol{\beta}) \, \mathrm{diag}(\boldsymbol{\eta}) \, \boldsymbol{V}_{K+1}^{\mathrm{T}}(\boldsymbol{\beta}).$$

Since the occurring Vandermonde and diagonal matrices have full rank, we have $\operatorname{rank} \boldsymbol{H} = K$ meaning $\dim(\ker(\boldsymbol{H})) = 1$. Thus, $\boldsymbol{H}$ possesses the simple singular value zero. Considering (2) for $\ell = 0, \dots, L-K-1$, we obtain

$$\boldsymbol{H}\boldsymbol{\gamma} = \boldsymbol{V}_{L-K}(\boldsymbol{\beta}) \left(\eta_j \, P(\beta_j)\right)_{j=0}^{K-1}.$$

Since the Vandermonde matrix $\boldsymbol{V}_{L-K}$ has full rank due to the assumptions on (1), the equivalence follows immediately. □

Lemma 3.1 is the theoretical justification why Prony's method always yields the parameters of (1) for exact measurements h_ℓ. In practice, the measurements $\tilde{h}_\ell := h_\ell + e_\ell$ are disturbed by some small error e_ℓ; so we have only access to the disturbed rectangular Hankel matrix

$$\tilde{\boldsymbol{H}} := \boldsymbol{H} + \boldsymbol{E} = \left(h_{\ell+k} + e_{\ell+k}\right)_{\ell,k=0}^{L-K-1,K} \qquad \text{with} \qquad L \geq 2K, \tag{4}$$

where $\boldsymbol{E} := (e_{\ell+k})_{\ell,k=0}^{L-K-1,K}$ is the rectangular error Hankel matrix. If $L > 2K$, the kernel of the perturbed Hankel matrix $\tilde{\boldsymbol{H}}$ will be trivial almost surely. For this reason, Potts

& Tasche [49] suppose to approximate the kernel using the singular value decomposition. This approach is supported by the Lidskii–Weyl perturbation theorem for singular values, see [16, Prob III.6.13] or [40], yielding

$$\max_{k=0,\dots,K} |\sigma_k(\tilde{\boldsymbol{H}}) - \sigma_k(\boldsymbol{H})| \le \|\tilde{\boldsymbol{H}} - \boldsymbol{H}\|_2 \le \|\boldsymbol{E}\|_2. \tag{5}$$

If the non-zero singular values of $\boldsymbol{H}$ are greater than $2\|\boldsymbol{E}\|$, the singular vector to the smallest singular value of $\tilde{\boldsymbol{H}}$ seems to be a valid approximation for $\boldsymbol{\gamma}$. Summarized, we obtain the so-called approximate Prony method [49, Alg 3.3] here written down for complex exponential sums.

ALGORITHM 3.2 (APPROXIMATE PRONY METHOD).
Input: $\tilde{\boldsymbol{h}} := (\tilde{h}_\ell)_{\ell=0}^{L-1} \in \mathbb{C}^L$ with $L > 2K$.

(i) Compute the right singular vector $\tilde{\boldsymbol{\gamma}}$ to the smallest singular value σ_K of $\tilde{\boldsymbol{H}}$.
(ii) Determine the roots $\tilde{\boldsymbol{\beta}} := (\tilde{\beta}_k)_{k=0}^{K-1}$ of $\tilde{P}(z) = \sum_{k=0}^{K} \tilde{\gamma}_k z^k$.
(iii) Compute the least-squares solution of $\boldsymbol{V}_L(\tilde{\boldsymbol{\beta}})\, \tilde{\boldsymbol{\eta}} = \tilde{\boldsymbol{h}}$.

Output: $\tilde{\boldsymbol{\eta}} \in \mathbb{C}^K$, $\tilde{\boldsymbol{\beta}} \in \mathbb{C}^K$.

Finally, we would like to note that alternative methods to obtain unknown bases from the exponential sum in (1) can be employed, for instance matrix pencil methods [33, 34], ESPRIT estimation methods [52], and Cadzow denoising method [19].

4 EXCLUSIVE PHASE RETRIEVAL

In the following, we assume that $\boldsymbol{A} \in \mathbb{C}^{d\times d}$ is diagonalizable, i.e. $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$. For a fixed signal $\boldsymbol{x} \in \mathbb{C}^d$ and a fixed sampling vector $\boldsymbol{\phi} \in \mathbb{C}^d$, the given phaseless measurements are then of the form

$$|\langle \boldsymbol{x}, \boldsymbol{A}^\ell \boldsymbol{\phi}\rangle|^2 = |\langle \boldsymbol{y}, \boldsymbol{\Lambda}^\ell \boldsymbol{\psi}\rangle|^2 = \left|\sum_{k=0}^{d-1} \lambda_k^\ell \underbrace{\bar{y}_k \psi_k}_{=:c_k}\right|^2 = \sum_{j,k=0}^{d-1} c_j \bar{c}_k\, (\lambda_j \bar{\lambda}_k)^\ell, \tag{6}$$

where $\boldsymbol{y} := \boldsymbol{S}^* \boldsymbol{x}$ and $\boldsymbol{\psi} := \boldsymbol{S}^{-1}\boldsymbol{\phi}$. Notice that (6) is an exponential sum with coefficients $c_j \bar{c}_k$ and bases $\lambda_j \bar{\lambda}_k$. In the following, we require that the exponential sum has exactly d^2 unique bases. Therefore, we call $M := \{\mu_0, \dots, \mu_{d-1}\} \subset \mathbb{C}$,

- *collision-free* if the products $\mu_j \bar{\mu}_k$ are pairwise distinct for $j, k \in \{0, \dots, d-1\}$.
- *absolutely collision-free* if M is collision-free and if the products $|\mu_j||\mu_k|$ are pairwise distinct for $j > k$.

Note that a matrix with collision-free eigenvalues is always invertible, and that the eigenvalue decomposition becomes unique up to permutations and global phases of the columns of $\boldsymbol{S}$. If the system or the matrix $\boldsymbol{A}$ is known, we can usually recover the signal $\boldsymbol{x}$ using one sampling vector $\boldsymbol{\phi}$.

Theorem 4.1. *Let $A \in \mathbb{C}^{d \times d}$ be known and diagonalizable with collision-free eigenvalues, and let $\phi \in \mathbb{C}^d$ depend on all eigenspaces of A. Then every $x \in \mathbb{C}^d$ can be recovered from the samples $\{|\langle x, A^\ell \phi\rangle|\}_{\ell=0}^{d^2-1}$ up to global phase.*

Proof. Assume that A has the eigenvalue decomposition $A = S \Lambda S^{-1}$, and denote the coordinates of ϕ with respect to S by $\psi := S^{-1}\phi$. The given measurements have the form

$$|\langle x, A^\ell \phi\rangle|^2 = \sum_{j,k=0}^{d-1} c_j \bar{c}_k \, (\lambda_j \bar{\lambda}_k)^\ell$$

with $c_k = \bar{y}_k \psi_k$ as shown in (6). Due to the distinctness of the products $\lambda_j \bar{\lambda}_k$, the coefficients $c_j \bar{c}_k$ may be calculated by solving a linear equation system based on an invertible Vandermonde matrix. The products $c_j \bar{c}_k$ contain the absolute values $|c_k|$ and the relative phases $\arg(c_j \bar{c}_k)$; so the factors c_k are determined up to global phase. Since the components of ψ are non-zero, and since S is invertible, we finally obtain x up to global phase. □

Corollary 4.2. *For almost all $a \in \mathbb{C}^d$ and almost all $\phi \in \mathbb{C}^d$, the signal $x \in \mathbb{C}^d$ can be recovered from the samples $\{|\langle x, (\operatorname{circ} a)^\ell \phi\rangle|\}_{\ell=0}^{d^2-1}$ up to global phase.*

Proof. The eigenvalues of $A := \operatorname{circ} a$ are just given by the discrete Fourier transform $\hat{a}$, and for almost all vectors $a \in \mathbb{C}^d$ or, equivalently, $\hat{a} \in \mathbb{C}^d$, the products $\hat{a}_j \bar{\hat{a}}_k$ are pairwise distinct. Further, the vectors ϕ that are orthogonal to one column of the Fourier matrix form a hyperplane. □

We would like to note that phase retrieval from the sample $\{|\langle x, (\operatorname{circ} a)^\ell \phi_i\rangle|\}_{\ell,i=0}^{L-1,J-1}$ is possible with much less than d^2 temporal measurement if more spatial measurement vectors ϕ_i and polarization techniques are employed [11].

5 Exclusive system identification

The other way round, if the signal x is known, then we can usually identity the eigenvalues of the system $A = S\Lambda S^{-1}$, i.e. we assume that the eigenvectors S of the system are known. For a convolutional systems $A = \operatorname{circ} a$, the eigenvectors are just the columns of the Fourier matrix for instance.

Theorem 5.1. *Let $A = S\Lambda S^{-1}$ be diagonalizable by a known eigenvector basis S and assume that the eigenvalues are collision-free. Let $\phi \in \mathbb{C}^d$ depend on all eigenspaces of A, and let $x \in \mathbb{C}^d$ be given. If the coefficients c_k defined in (6) are collision-free too, then the eigenvalues $\lambda_0, \dots, \lambda_{d-1}$ of A are defined by the samples $\{|\langle x, A^\ell \phi\rangle|\}_{\ell=0}^{2d^2-1}$ up to global phase.*

Proof. The measurements again have the form

$$|\langle \boldsymbol{x}, \boldsymbol{A}^\ell \boldsymbol{\phi}\rangle|^2 = \sum_{j,k=0}^{d-1} c_j \bar{c}_k \, (\lambda_j \bar{\lambda}_k)^\ell$$

as shown in (6). By assumption, the bases $\lambda_j \bar{\lambda}_k$ of this exponential sum are pairwise distinct and the coefficients c_k are non-zero. Thus the products $\lambda_j \bar{\lambda}_k$ and $c_j \bar{c}_k$ are determinable by Prony's method. Note that Prony's method gives only the values but not the corresponding indices j and k. Exploiting that the products $c_j \bar{c}_k$ are known – $\boldsymbol{x}$, $\boldsymbol{\phi}$, $\boldsymbol{S}$ are known, we can however deduce these indices. Similarly to the proof of Theorem 4.1, the products $\lambda_j \bar{\lambda}_k$ contain the absolute values $|\lambda_k|$ and the relative phases $\arg(\lambda_j \bar{\lambda}_k)$; so the eigenvalues λ_k are determined up to global phase. □

COROLLARY 5.2. *For almost all $\boldsymbol{x} \in \mathbb{C}^d$ and almost all $\boldsymbol{\phi} \in \mathbb{C}^d$, almost all kernels $\boldsymbol{a} \in \mathbb{C}^d$ can be recovered from the samples $\{|\langle \boldsymbol{x}, (\operatorname{circ} \boldsymbol{a})^\ell \boldsymbol{\phi}\rangle|\}_{\ell=0}^{2d^2-1}$ up to global phase.*

Proof. Again, the vectors $\boldsymbol{\phi}$ that are orthogonal to one column of the Fourier matrix form a hyperplane. Further, for almost all $\boldsymbol{\phi}$ and $\boldsymbol{x}$, the products $c_j \bar{c}_k$ in (6) are pairwise distinct. As discussed in the proof of Corollary 4.2 almost all vectors $\boldsymbol{a} \in \mathbb{C}^d$ satisfy the assumption of Theorem 5.1. □

6 SIMULTANEOUS PHASE & SYSTEM IDENTIFICATION

If either the signal $\boldsymbol{x}$ or the spectrum of $\boldsymbol{A}$ are known, we can recover the respective unknown information from the temporal samples of only one sampling point. To a certain degree, we may even determine some information if both – the signal and the spectrum – are unknown. Using one sampling point, we however lose the order of the components. So we only obtain the unordered spectrum of $\boldsymbol{A}$.

THEOREM 6.1. *Let $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$ be diagonalizable by a known eigenvector basis $\boldsymbol{S}$ and assume that the eigenvalues are absolutely collision-free. Let $\boldsymbol{\phi} \in \mathbb{C}^d$ depend on all eigenspaces of $\boldsymbol{A}$, and let $\boldsymbol{y} := \boldsymbol{S}^* \boldsymbol{x}$ be elementwise non-zero for unknown $\boldsymbol{x} \in \mathbb{C}^d$. Then the spectrum of $\boldsymbol{A}$ is determined by the samples $\{|\langle \boldsymbol{x}, \boldsymbol{A}^\ell \boldsymbol{\phi}\rangle|\}_{\ell=0}^{2d^2-1}$ up to global phase and winding direction.*

Proof. Since the coefficients $c_k = \bar{y}_k \psi_k$ with $\boldsymbol{y} := \boldsymbol{S}^* \boldsymbol{x}$ and $\boldsymbol{\psi} := \boldsymbol{S}^{-1} \boldsymbol{\phi}$ are non-zero, and since the eigenvalues are absolutely collision-free, the measurements have the form

$$|\langle \boldsymbol{x}, \boldsymbol{A}^\ell \boldsymbol{\phi}\rangle|^2 = \sum_{j,k=0}^{d-1} c_j \bar{c}_k \, (\lambda_j \bar{\lambda}_k)^\ell = \sum_{k=0}^{d^2-1} \eta_k \beta_k^\ell$$

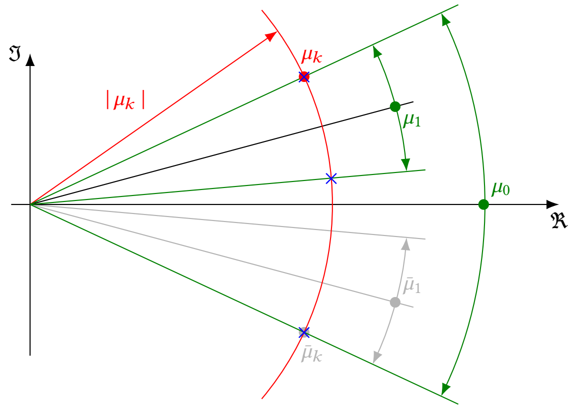


Figure 1: Propagating the phase in the proof of Theorem 6.1. The points μ_0 and μ_1 are already known. Using the relatives phases $\pm \arg(\mu_k \mu_0)$ and $\pm \arg(\mu_k \bar{\mu}_1)$, starting from μ_0 and μ_1, we obtain two possible candidates ($\times$) for μ_k respectively since $|\mu_k|$ is known too. Further, since μ_1 cannot also be real by assumption, exactly two candidates coincide yielding μ_k. For the other winding direction, i.e. choosing $\bar{\mu}_1$ instead of μ_1, we obtain $\bar{\mu}_k$.

as shown in (6), where β_k denotes the d^2 unique, unknown bases and η_k the corresponding coefficients. Applying Prony's method, we now recover the set $B := \{\beta_k\}_{k=0}^{d^2-1}$. Note that the relation between the elements of B and $\{\lambda_j \bar{\lambda}_k : k, j = 0, \ldots, d-1\}$ is still unrevealed.

In the following, we denote the recovered eigenvalues of $\boldsymbol{A}$ in absolutely decreasing order by μ_k, i.e. $|\mu_0| > \cdots > |\mu_{d-1}|$, and recover the permuted eigenvalues step by step. Our assumption guarantees that $\mu_j \bar{\mu}_k$ differs from $\mu_k \bar{\mu}_j$, i.e. the imaginary part cannot vanish; so the real values in B correspond to the magnitudes $|\mu_k|$. The absolute collision freedom now allow us to recover the products $\mu_j \bar{\mu}_k$ and $\mu_k \bar{\mu}_j$ in B corresponding to $|\mu_j|$ and $|\mu_k|$. We now assume that μ_0 is real and positive because the global phase cannot be recovered. Considering $\mu_0 \bar{\mu}_1$ and $\mu_1 \bar{\mu}_0$, we obtain the relative phase $\arg(\mu_0) - \arg(\mu_1) \bmod 2\pi$ up to sign. At this point, we have to chose one winding direction for the phase. For $k = 2, \ldots, d-1$, we may consider the relative phases between μ_k and the recovered μ_0 and μ_1, see Figure 1, which uniquely determines the remaining phases. □

Remark 6.2. Note that the spectrum retrieved in Theorem 6.1 is an unordered set, i.e. the relation to the known eigenvectors in $\boldsymbol{S}$ is not revealed. Applying the recovered relations between the bases, we may also recover the coefficients c_k in (6) up to global phase and winding direction. However, without knowing the actual order of the eigenvalues/coefficients, the recovery of the unknown signal is forlorn. ○

Supposing that the unknown complex eigenvalues of the operator $\boldsymbol{A}$ have a clearly recognizable structure like increasing/decreasing absolute values leads to highly artificial side condition. A nevertheless interesting special case are real-valued convolutional systems with symmetrically decreasing kernels in the frequency domain. For the following theorem, we therefore restrict the setup to real-valued signals $\boldsymbol{x} \in \mathbb{R}^d$, real-valued convolution operators $\operatorname{circ} \boldsymbol{a}$ with $\boldsymbol{a} \in \mathbb{R}^d$, and real-valued sampling vectors $\boldsymbol{\phi} \in \mathbb{R}^d$. We call a kernel $\boldsymbol{a}$ *strictly, symmetrically decreasing* when

$$\hat{\boldsymbol{a}} \in \mathbb{R}^d_{++}, \qquad \hat{a}_k = \hat{a}_{-k}, \qquad \text{and} \qquad \hat{a}_k > \hat{a}_j$$

for $k, j \in \{0, \dots, \lfloor d/2 \rfloor\}$ with $k < j$. The negative indices are here considered modulo d, and $\mathbb{R}_{++}$ denotes the real and positive half axis. Strictly, symmetrically decreasing kernels correspond to low-pass filters, whose identification in dynamical sampling has been studied in [56]. Note that the signal $\boldsymbol{a}$ is real and symmetric too. We call the kernel *collision-free in frequency* if the products $\hat{a}_j \hat{a}_k$ are unique for $k, j \in \{0, \dots, \lfloor d/2 \rfloor\}$ with $j \geq k$. This definition differs from the collision-free complex sets. In order to recover both – signal and kernel, we employ two sampling vectors $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$. We call $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$ *pointwise independent (in the frequency domain)* when $\hat{\phi}_{1,k}$ and $\hat{\phi}_{2,k}$ interpreted as two-dimensional real vectors are independent for $k = 1, \dots, \lfloor d/2 \rfloor$. For this specific setting, the identification of the system and the signal is usually possible.

THEOREM 6.3. *Let $\boldsymbol{a} \in \mathbb{R}^d$ be strictly, symmetrically decreasing and collision-free in frequency, let $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2 \in \mathbb{R}^d$ be pointwise independent, and let $\boldsymbol{x} \in \mathbb{R}^d$ satisfy $\Re[\bar{\hat{x}}_k \hat{\phi}_{i,k}] \neq 0$ for $k = -\lfloor (d-1)/2 \rfloor, \dots, \lfloor d/2 \rfloor$, $i = 1, 2$. Then $\boldsymbol{a}$ and $\boldsymbol{x}$ can be recovered from the samples*

$$\left\{ |\langle \boldsymbol{x}, (\operatorname{circ} \boldsymbol{a})^\ell \boldsymbol{\phi}_1 \rangle|, |\langle \boldsymbol{x}, (\operatorname{circ} \boldsymbol{a})^\ell \boldsymbol{\phi}_2 \rangle| \right\}_{\ell=0}^{L-1} \qquad \text{with} \qquad L := \left(\left\lfloor \tfrac{d}{2} \right\rfloor + 1\right)\left(\left\lfloor \tfrac{d}{2} \right\rfloor + 2\right)$$

up to global sign.

Proof. To simplify the notation, we first study the temporal samples with respect to an arbitrary sampling vector $\boldsymbol{\phi}$. Exploiting the symmetry of $\hat{\boldsymbol{a}}$ and the conjugated symmetry of $\boldsymbol{c} := (\bar{\hat{x}}_k \hat{\phi}_k)_{k=-\lfloor (d-1)/2 \rfloor}^{\lfloor d/2 \rfloor}$ caused by the Fourier transform, we combine the several times appearing bases in (6) to obtain

$$\begin{aligned} |\langle \boldsymbol{x}, (\operatorname{circ} \boldsymbol{a})^\ell \boldsymbol{\phi} \rangle|^2 &= \left| \sum_{k=-\lfloor (d-1)/2 \rfloor}^{\lfloor d/2 \rfloor} c_k \hat{a}_k^\ell \right|^2 = \left| \sum_{k=0}^{\lfloor d/2 \rfloor} \gamma_k \Re[c_k] \, \hat{a}_k^\ell \right|^2 \\ &= \sum_{k=0}^{\lfloor d/2 \rfloor} \sum_{j=0}^{\lfloor d/2 \rfloor} \gamma_k \gamma_j \Re[c_k] \Re[c_j] \, (\hat{a}_k \hat{a}_j)^\ell \\ &= \sum_{k=0}^{\lfloor d/2 \rfloor} \sum_{j=k}^{\lfloor d/2 \rfloor} \underbrace{\gamma_{k,j} \Re[c_k] \Re[c_j]}_{\eta_k} (\hat{a}_k \hat{a}_j)^\ell = \sum_{k=0}^{L/2-1} \eta_k \beta_k^\ell, \end{aligned}$$

with bases β_k related to $\hat{a}_k \hat{a}_j$ and coefficients η_k where the multipliers are given by

$$\gamma_k := \begin{cases} 1 & \text{if } k = 0, \\ 2 & \text{if } k = 1, \dots, \lfloor (d-1)/2 \rfloor, \\ 1 & \text{if } k = d/2 \text{ and } d \text{ is even,} \end{cases} \qquad \text{and} \qquad \gamma_{k,j} := \begin{cases} 2\gamma_k\gamma_j & \text{if } k \neq j, \\ \gamma_k^2 & \text{if } k = j. \end{cases}$$

This exponential sum has exactly $1/2(\lfloor d/2 \rfloor + 1)(\lfloor d/2 \rfloor + 2)$ distinct bases since $\boldsymbol{a}$ is collision-free in frequency.

Applying Prony's method, we compute the bases β_k and coefficients η_k. Because the bases β_k are all real and non-negative, we need a different procedure than before to reveal the relation to the factors $\hat{a}_k \hat{a}_j$. Let B be the set of recovered bases, where we assume $\beta_0 > \cdots > \beta_{L/2-1}$.

(i) The strict, symmetric decrease of $\boldsymbol{a}$ ensures $\beta_0 = \hat{a}_0^2$. Now, remove β_0 from B.

(ii) The next largest basis β_1 corresponds to $\hat{a}_0 \hat{a}_1$ allowing the recovery of $\hat{a}_1$. Remove $\beta_1 = \hat{a}_0 \hat{a}_1$ and $\hat{a}_1^2$ from B.

(iii) The largest remaining bases correspond to $\hat{a}_0 \hat{a}_2$, which gives us $\hat{a}_2$. Remove all products $\hat{a}_0 \hat{a}_2$, $\hat{a}_1 \hat{a}_2$, $\hat{a}_2 \hat{a}_2$ of $\hat{a}_2$ with the recovered components from B.

(iv) Repeating this procedure, we obtain $\hat{a}_0, \dots, \hat{a}_{\lfloor d/2 \rfloor}$ and, due to symmetry, the remaining half of $\hat{\boldsymbol{a}}$.

Alongside of the kernel, we also obtain the relation between η_k and $\gamma_{k,j} \Re[c_k] \Re[c_j]$ for each sampling vector $\boldsymbol{\phi}_1$, $\boldsymbol{\phi}_2$. Assuming $\Re[\bar{\hat{x}}_0 \hat{\phi}_{1,0}] = \bar{\hat{x}}_0 \hat{\phi}_{1,0} > 0$, we compute the real parts $\Re[\bar{\hat{x}}_k \hat{\phi}_{1,k}]$ for $k = 1, \dots, \lfloor d/2 \rfloor$ by exploiting the revealed relative phases (sign changes), transfer the sign from $\bar{\hat{x}}_0 \hat{\phi}_{1,0}$ to $\bar{\hat{x}}_0 \hat{\phi}_{2,0} = \Re[\bar{\hat{x}}_0 \hat{\phi}_{2,0}]$ since $\hat{\phi}_{1,0}$ and $\hat{\phi}_{2,0}$ are known, and determine $\Re[\bar{\hat{x}}_k \hat{\phi}_{2,k}]$ for $k = 1, \dots, \lfloor d/2 \rfloor$ analogously. Due to the pointwise linear independence, the equation systems

$$\Re[\bar{\hat{x}}_k \hat{\phi}_{1,k}] = \Re\hat{\phi}_{1,k} \Re\hat{x}_k + \Im\hat{\phi}_{1,k} \Im\hat{x}_k$$

$$\Re[\bar{\hat{x}}_k \hat{\phi}_{2,k}] = \Re\hat{\phi}_{2,k} \Re\hat{x}_k + \Im\hat{\phi}_{2,k} \Im\hat{x}_k$$

gives us $\hat{x}_k$ for $k = 1, \dots, \lfloor d/2 \rfloor$. With the conjugated symmetry of $\hat{\boldsymbol{x}}$, the inverse Fourier transform yields $\boldsymbol{x}$ up to the sign. □

Remark 6.4. Note that the assumption $\Re[\bar{\hat{x}}_k \hat{\phi}_{i,k}] \neq 0$ for $k = 0, \dots, d-1$ may be weakened to only hold for one sampling vector $\boldsymbol{\phi}_1$ or $\boldsymbol{\phi}_2$ as long as $\Re[\bar{\hat{x}}_0 \hat{\phi}_{i,0}] \neq 0$ for both. In this case, the exponential sum corresponding to the temporal samples of the other sampling vector may consist of less than $1/2\,(\lfloor d/2 \rfloor + 1)(\lfloor d/2 \rfloor + 2)$ bases. Exploiting that the coefficients η_k of the missing bases are zero, and spreading the sign between the non-zero coefficients, we can nevertheless recover $\boldsymbol{x}$. ○

It is also possible to identify the strictly, symmetrically decreasing kernel alongside a complex signal and to allow complex sampling vectors. In this case, the temporal samples corresponding to one sampling vector $\boldsymbol{\phi}$ possesses the form

$$|\langle \boldsymbol{x}, (\operatorname{circ} \boldsymbol{a})^\ell \boldsymbol{\phi} \rangle|^2 = \sum_{k=0}^{\lfloor d/2 \rfloor} \sum_{j=k}^{\lfloor d/2 \rfloor} \tfrac{\gamma_{k,j}}{4} \, \Re\left[(c_k + c_{-k})(\bar{c}_j + \bar{c}_{-j})\right] (\hat{a}_k \hat{a}_j)^\ell.$$

Similarly to the proof of Theorem 6.3, we may recover the kernel $\boldsymbol{a}$ from the temporal samples of one sampling vector if

$$\Re\left[(\bar{\hat{x}}_k \hat{\phi}_k + \bar{\hat{x}}_{-k} \hat{\phi}_{-k}) \overline{(\bar{\hat{x}}_j \hat{\phi}_j + \bar{\hat{x}}_{-j} \hat{\phi}_{-j})}\right] \neq 0 \tag{7}$$

for $k, j = 0, \dots, \lfloor d/2 \rfloor$. Additionally, the signal $\boldsymbol{x}$ may be recovered if four sampling vectors are employed. In this case, the coefficient of $\hat{a}_0^2$ is just $|\bar{\hat{x}}_0 \hat{\phi}_{i,0}|$; so fixing the phase for $c_{1,0}$, we may spread the phase to $c_{i,0}$, $i = 2, 3, 4$, where the first index stands for the related sampling vector, i.e. all coefficients $c_{i,0}$ are known. If the equation system

$$\begin{aligned}\Re[c_{i,0}(\bar{c}_{i,k} + \bar{c}_{i,-k})] &= \Re[\bar{c}_{i,0} \hat{\phi}_{i,k}] \, \Re \hat{x}_k + \Im[\bar{c}_{i,0} \hat{\phi}_{i,k}] \, \Im \hat{x}_k \\ &\quad + \Re[\bar{c}_{i,0} \hat{\phi}_{i,-k}] \, \Re \hat{x}_{-k} + \Im[\bar{c}_{i,0} \hat{\phi}_{i,-k}] \, \Im \hat{x}_{-k}\end{aligned}$$

with $i = 1, \dots, 4$ is solvable, we obtain $\hat{\boldsymbol{x}}$ and thus $\boldsymbol{x}$. Notice that the recovery of $\hat{\boldsymbol{a}}$ here is not a special case of Theorem 6.1 since $\hat{\boldsymbol{a}}$ is not collision-free as a complex set. In sum, the following statement can be established.

THEOREM 6.5. *Let $\boldsymbol{a} \in \mathbb{R}^d$ be strictly, symmetrically decreasing and collision-free in frequency, let $\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_4 \in \mathbb{C}^d$ and $\boldsymbol{x} \in \mathbb{C}^d$ satisfy (7). If the real-valued vectors*

$$(\Re[\bar{c}_{i,0} \hat{\phi}_{i,k}], \Im[\bar{c}_{i,0} \hat{\phi}_{i,k}], \Re[\bar{c}_{i,0} \hat{\phi}_{i,-k}], \Im[\bar{c}_{i,0} \hat{\phi}_{i,-k}])^{\mathrm{T}}, \qquad i = 1, \dots, 4,$$

are independent for each $k = 1, \dots, \lfloor (d-1)/2 \rfloor$, then $\boldsymbol{a}$ and $\boldsymbol{x}$ can be recovered from the samples

$$\left\{ |\langle \boldsymbol{x}, (\operatorname{circ} \boldsymbol{a})^\ell \boldsymbol{\phi}_i \rangle| \right\}_{\ell=0, i=1}^{L-1, 4} \qquad \text{with} \qquad L \coloneqq \left(\left\lfloor \tfrac{d}{2} \right\rfloor + 1 \right) \left(\left\lfloor \tfrac{d}{2} \right\rfloor + 2 \right)$$

up to global phase.

Remark 6.6. The strictly, symmetrically decreasing kernels form a $(\lfloor d/2 \rfloor + 1)$-dimensional manifold. Further, the not collision-free kernels live on the union of submanifolds with strictly smaller dimension; so almost all strictly, symmetrically decreasing kernels are collision-free. Moreover, almost all vectors $\boldsymbol{x}$ and $\boldsymbol{\phi}_i$ satisfy the posed conditions in the real as well as in the complex setting. ○

7 Multiple sampling vectors

Let us return to the parameter identification of arbitrary systems after that brief digression to strictly, symmetrically decreasing convolution kernels. Revisiting the statement in Theorem 6.1, our main problem has been that we cannot recover the order of the spectrum from merely one sampling vector if both – signal and eigenvalues – are unknown. Since our analysis is based on Prony's method, we have always relied on a squared number of measurements. To surmount these shortcomings, we suppose specifically constructed sets of sampling vectors.

Instead of assuming that the sampling vectors $\boldsymbol{\phi}_i$ depend on all eigenspaces of the system matrix $\boldsymbol{A}$, we now assume that $\boldsymbol{\phi}_i$ might only depends on a small set of eigenspaces. Considering the temporal samples for such a sampling vector, in analogy to (6), we have

$$|\langle \boldsymbol{x}, \boldsymbol{A}^\ell \boldsymbol{\phi}_i \rangle|^2 = |\langle \boldsymbol{y}, \boldsymbol{\Lambda}^\ell \boldsymbol{\psi}_i \rangle|^2 = \left| \sum_{k \in \mathcal{I}_i} \lambda_k^\ell \underbrace{\bar{y}_k \psi_{i,k}}_{=: c_{i,k}} \right|^2 = \sum_{j,k \in \mathcal{I}_i} c_{i,j} \bar{c}_{i,k} \, (\lambda_j \bar{\lambda}_k)^\ell,$$

where $\boldsymbol{y} := \boldsymbol{S}^* \boldsymbol{x}$, $\boldsymbol{\psi}_i := \boldsymbol{S}^{-1} \boldsymbol{\phi}_i$, and $\mathcal{I}_i := \operatorname{supp} \boldsymbol{\psi}_i$. Since $\boldsymbol{\phi}_i$ only captures a small part of the spectrum, the last sum only consists of $|\mathcal{I}_i|$ exponentials instead of d^2 and allows the recovery of a specific part of the spectrum. To combine these partial information and to overcome the mentioned issues, the sampling vectors $\{\boldsymbol{\phi}_i\}_{i=0}^{J-1}$ with $\boldsymbol{\psi}_i := \boldsymbol{S}^{-1} \boldsymbol{\phi}_i$ should allow

(i) *index separation:* the supports of $\{\boldsymbol{\psi}_i\}_{i=0}^{J-1}$ form a full cover meaning $\bigcup_{i=0}^{J-1} \operatorname{supp} \boldsymbol{\psi}_j = \{0, \dots, d-1\}$, and for every $k \in \{0, \dots, d-1\}$ there exist two index sets $\mathcal{F}_k$ and $\mathcal{G}_k$ such that

$$\{k\} = \bigcap_{i \in \mathcal{F}_k} \operatorname{supp} \boldsymbol{\psi}_i \Big\backslash \bigcup_{i \in \mathcal{G}_k} \operatorname{supp} \boldsymbol{\psi}_i, \tag{8}$$

(ii) *phase propagation:* the set $\{\boldsymbol{\phi}_i\}_{i=0}^{J-1}$ is ordered such that

$$\# \left[\operatorname{supp} \boldsymbol{\psi}_k \cap \bigcup_{i=0}^{k-1} \operatorname{supp} \boldsymbol{\psi}_i \right] = 2 \tag{9}$$

for $k = 1, \dots, J-1$, i.e. there is an overlap of two elements at least,

(iii) *winding direction determination:* there are indices i_1, i_2, k_1, k_2 such that

$$\arg(\psi_{i_1,k_1} \bar{\psi}_{i_1,k_2}) \not\equiv \arg(\psi_{i_2,k_1} \bar{\psi}_{i_2,k_2}) \bmod \pi, \tag{10}$$

where $\psi_{i_1,k_1} \bar{\psi}_{i_1,k_2}$ and $\psi_{i_2,k_1} \bar{\psi}_{i_2,k_2}$ are non-zero.

If the sampling vectors $\boldsymbol{\phi}_i$ fulfill all three assumptions, we say that the sampling set allows *parameter identification and phase retrieval* (up to global phase).

THEOREM 7.1. *Let $A = S\Lambda S^{-1}$ be diagonalizable by a known eigenvector basis S and assume that the eigenvalues are absolutely collision-free. Let $\{\phi_j\}_{j=0}^{J-1} \subset \mathbb{C}^d$ allow parameter identification and phase retrieval, and let $y := S^* x$ be elementwise non-zero for unknown $x \in \mathbb{C}^d$. Then the eigenvalues $\lambda_0, \dots, \lambda_{d-1}$ of A and the signal x are determined by the spatiotemporal samples*

$$\left\{ |\langle x, A^\ell \phi_i \rangle| \right\}_{\ell,i=0}^{L_i^2-1,J-1} \qquad \text{with} \qquad L_i := \#[\operatorname{supp}(S^{-1}\phi_i)]$$

up to global phase.

Proof. Using the procedure in the proof of Theorem 6.1, we recover the unblocked part $\Lambda_i := \{\lambda_k : k \in \mathcal{I}_i\}$ of the spectrum of A for each $i = 0, \dots, J-1$ up to global phase and winding direction. Note that we do not know which value in Λ_i corresponds to which index. However, since the eigenvalues are absolutely collision-free, and since the sampling set allows index separation, we have

$$\bigcap_{j \in \mathcal{F}_k} |\Lambda_i| \setminus \bigcup_{i \in \mathcal{G}_k} |\Lambda_i| = |\lambda_k|,$$

where the absolute value is applied element by element. Thus the true index of the eigenvalues is revealed.

Using that the sampling set allows phase propagation, we align the global phase and winding direction of the sets Λ_i as follows. First, we fix the global phase and winding direction of Λ_0. There are at least two eigenvalues λ_{k_1} and λ_{k_2} that are contained in Λ_0 and Λ_1. The collision-freedom ensures $\arg(\lambda_{k_1}\bar{\lambda}_{k_2}) \not\equiv 0 \bmod \pi$. Using λ_{k_1} and λ_{k_2}, which can be identified by their absolute values, the global phase and winding direction are uniquely transferable form Λ_0 to Λ_1, i.e. we obtain the eigenvalues in $\Lambda_0 \cup \Lambda_1$ up to global phase and winding direction. Repeating this argument, we propagate the phase information to the remaining subsets Λ_i, which results in the recovery of all eigenvalues $\lambda_0, \dots, \lambda_{d-1}$ up to global phase and winding direction.

The ambiguity with respect to the winding direction occurs since we have not been able to determine whether the true relative phase between λ_j and λ_k corresponds to $\arg(\lambda_j \bar{\lambda}_k)$ or to $\arg(\lambda_k \bar{\lambda}_j)$. Let us now consider the indices i_1, i_2, k_1, k_2 in the winding direction property (10) of $\{\phi_i\}_{i=0}^{J-1}$. Notice that both λ_{k_1} and λ_{k_2} are captured by the sampling vectors ϕ_{i_1}, ϕ_{i_2}. Due to the missing winding direction, the coefficients $c_{i_1,k_1}\bar{c}_{i_1,k_2}$ and $c_{i_2,k_1}\bar{c}_{i_2,k_2}$ can only be identified up to the conjugation; so we merely obtain $\Re[c_{i_1,k_1}\bar{c}_{i_1,k_2}]$ and $\Re[c_{i_2,k_1}\bar{c}_{i_2,k_2}]$, which however are given by

$$\Re[c_{i_1,k_1}\bar{c}_{i_1,k_2}] = \Re[y_{k_1}\bar{y}_{k_2}]\,\Re[\psi_{i_1,k_1}\bar{\psi}_{i_1,k_2}] + \Im[y_{k_1}\bar{y}_{k_2}]\,\Im[\psi_{i_1,k_1}\bar{\psi}_{i_1,k_2}],$$

$$\Re[c_{i_2,k_1}\bar{c}_{i_2,k_2}] = \Re[y_{k_1}\bar{y}_{k_2}]\,\Re[\psi_{i_2,k_1}\bar{\psi}_{i_2,k_2}] + \Im[y_{k_1}\bar{y}_{k_2}]\,\Im[\psi_{i_2,k_1}\bar{\psi}_{i_2,k_2}].$$

Our assumptions guarantees that this equation system has the unique answer $y_{k_1}\bar{y}_{k_2}$, which yields $c_{i_1,k_1}\bar{c}_{i_1,k_2}$ and $c_{i_2,k_1}\bar{c}_{i_2,k_2}$ without conjugation ambiguity. Further, at least one of the products $c_{i_1,k_1}\bar{c}_{i_1,k_2}$ and $c_{i_2,k_1}\bar{c}_{i_2,k_2}$ has a non-vanishing imaginary part again due to (10). The corresponding basis $\lambda_{k_1}\bar{\lambda}_{k_2}$ reveals the true winding direction resulting in the recovery of $\lambda_0, \dots, \lambda_{d-1}$ up to global phase.

Considering the coefficient of the temporal samples for each $\boldsymbol{\phi}_i$, we determine y_k with $k \in \operatorname{supp} \boldsymbol{\psi}_i$ up to global phase. The recovered components of $\boldsymbol{y}$ may now be aligned due to the overlap between the supports in (9) yielding $\boldsymbol{y}$ up to global phase. Applying the inverse of $\boldsymbol{S}^*$, we finally obtain the wanted signal $\boldsymbol{x}$ up to global phase. □

Remark 7.2. The absolute collision-freedom of the eigenvalues can be weakened. More precisely, we only require the absolute collision-freedom on the non-blocked parts of the spectrum with respect to $\{\boldsymbol{\phi}_i\}_{i=0}^{J-1}$, i.e. we only require that the sets Λ_i are absolutely collision-free. In order to propagate the phase, there have to be to at least two indices

$$k_1, k_2 \in \operatorname{supp} \boldsymbol{\psi}_k \cap \bigcup_{i=0}^{k-1} \operatorname{supp} \boldsymbol{\psi}_i$$

for $k = 1, \dots, J-1$, cf. (9), satisfying $\arg(\lambda_{k_1}\bar{\lambda}_{k_2}) \not\equiv 0 \bmod \pi$. ○

Theorem 7.1 not only allow us to recover the signal and the system's eigenvalues simultaneously but also to reduce the required number of samples. In the statements before, the number of measurements to apply Prony's method is always a multiple of the squared dimension, i.e. we require $\mathcal{O}(d^2)$ samples. In Theorem 7.1 the number of spatiotemporal samples mainly correlate with the support sparsity $L_i := \#[\operatorname{supp}(\boldsymbol{S}^{-1}\boldsymbol{\phi}_i)]$. With $L := \max\{L_i : i = 0, \dots J-1\}$, the number of samples is thus bounded by $2L^2J$. Notice that we need d vectors at the most to build a sampling set allow parameter identification and phase retrieval. For instance the sampling vectors may be constructed such that $\operatorname{supp} \boldsymbol{\psi}_i := \{i, \dots, i+L-1\}$ for $i = 0, \dots, d-L$ and $L \geq 3$. We then employ only $\mathcal{O}(dL^2)$ measurement. For a fixed sparsity L, we only need linearly many spatiotemporal samples.

COROLLARY 7.3. *Under the assumption of Theorem 7.1, the eigenvalues of $\boldsymbol{A} \in \mathbb{C}^{d\times d}$ and the unknown signal $\boldsymbol{x} \in \mathbb{C}^d$ are identifiable with $\mathcal{O}(d)$ spatiotemporal samples.*

The idea of blocking a part of the spectrum to reduce the number of required spatiotemporal samples clearly transfers to Theorem 6.3 and 6.5. The indices of the recovered eigenvalues is then determined by the strict, symmetrical decay; so the index separation, phase propagation, and winding direction determination is not required, although the supports of $\{\hat{\boldsymbol{\phi}}_i\}_{i=0}^{J-1}$ should still form a full cover. Considering Theorem 6.3 exemplarily, we instead need that, for every $k \in \{0, \dots, d-1\}$, there exists at least one index $j \in \{0, \dots, J-1\}$ such that $\Re[\bar{\hat{x}}_k \hat{\phi}_{i,k}] \neq 0$ to recover all components of $\hat{\boldsymbol{a}}$ and two indices $i_1, i_2 \in \{0, \dots, J-1\}$ such that $\hat{\phi}_{i_1,k}$ and $\hat{\phi}_{i_2,k}$ are linearly independent interpreted as two-dimensional real vectors to recover all components of $\hat{\boldsymbol{x}}$.

8 Phase & system identification in infinite dimensions

Up to this point, we only considered the finite-dimensional setting. The central ideas to apply Prony's method to identify the eigenvalues of the system and the unknown signal simultaneously is however extendable to the infinite-dimensional setting too. In the following, we consider an infinite-dimensional, complex Hilbert space $\mathcal{H}$ and call an invertible, bounded, linear operator $\mathcal{A} : \mathcal{H} \to \mathcal{H}$ *diagonalizable* if $\mathcal{A}$ can be factorized into $\mathcal{A} = \mathcal{S}\Lambda\mathcal{S}^{-1}$, where $\mathcal{S} \colon \ell^2(\mathcal{Z}) \to \mathcal{H}$ is an invertible, bounded, linear operator, $\Lambda \colon \ell^2(\mathcal{Z}) \to \ell^2(\mathcal{Z})$ is a multiplication operator, and $\mathcal{Z}$ is an infinite countable set like $\mathbb{N}$ or $\mathbb{Z}$. The elementwise *multiplication operator* $\Lambda \colon \ell^2(\mathcal{Z}) \to \ell^2(\mathcal{Z})$ is defined by

$$\Lambda(y) \coloneqq \big(\lambda_k \, y_k\big)_{k \in \mathcal{Z}}$$

with bounded eigenvalues $\lambda_k \in \mathbb{C}_*$, i.e. $\sup_{k\in\mathcal{Z}} |\lambda_k| < \infty$.

Similarly to the finite-dimensional setting, the temporal samples for one sampling vector $\boldsymbol{\phi}_i$ are given by

$$|\langle x, \mathcal{A}^\ell \phi_i \rangle_{\mathcal{H}}|^2 = |\langle y, \Lambda^\ell \psi_i \rangle_{\ell^2(\mathcal{Z})}|^2 = \Bigg| \sum_{k \in \mathcal{I}_i} \lambda_k^\ell \underbrace{\bar{y}_k \psi_{i,k}}_{=: c_{i,k}} \Bigg|^2 = \sum_{j,k \in \mathcal{I}_i} c_{i,j} \bar{c}_{i,k} \, (\lambda_j \bar{\lambda}_k)^\ell,$$

where $y \coloneqq \mathcal{S}^* x$, $\psi_i \coloneqq \mathcal{S}^{-1}\phi_i$, and $\mathcal{I}_i \coloneqq \operatorname{supp} \psi_i \subset \mathcal{Z}$. If $\operatorname{supp} \psi_i$ is finite, the sum on the right-hand side becomes finite such that Prony's method may be applied to recover the present eigenvalues (without indices). In order to determine the complete spectrum, the finite supports of ϕ_i have to form a full cover of $\mathcal{Z}$, which is only possible for infinitely many sampling vectors, i.e. $J = \infty$. To align the recovered subsets, we rely again on the parameter identification and phase retrieval properties in (8–10). In sum, we obtain the following recovery guarantee for infinite-dimensional Hilbert spaces.

Theorem 8.1. *Let $\mathcal{A} \colon \mathcal{H} \to \mathcal{H}$ with absolutely collision-free eigenvalues be diagonalizable by a known $\mathcal{S} \colon \ell^2(\mathcal{Z}) \to \mathcal{H}$, where $\mathcal{H}$ is an infinite-dimensional Hilbert space and $\mathcal{Z}$ an infinite countable set. Let $\{\phi_j\}_{j=0}^\infty \subset \mathcal{H}$ allows parameter identification and phase retrieval with finitely supported $\mathcal{S}^{-1}\phi_i$, and let $y \coloneqq \mathcal{S}^* x$ be elementwise non-zero for unknown $x \in \mathcal{H}$. Then the eigenvalues λ_k with $k \in \mathcal{Z}$ of $\mathcal{A}$ and the signal x are defined by the spatiotemporal samples*

$$\big\{ |\langle x, \mathcal{A}^\ell \phi_i \rangle| \big\}_{\ell,i=0}^{L_i^2 - 1, \infty} \qquad \text{with} \qquad L_i \coloneqq \#[\operatorname{supp}(\mathcal{S}^{-1}\phi_i)]$$

up to a global phase.

Since the statement can be established with the construction in the proof of Theorem 7.1, we omit the proof. Furthermore, Remark 7.2 carries over to the infinite-dimensional setting as well. Note that the non-zero assumption on $y \coloneqq \mathcal{S}^* x$ is crucial since otherwise a part of the spectrum is blocked in all spatiotemporal measurements and thus cannot be recovered.

An example for the infinite-dimensional Hilbert space setting is the repeated convolution of periodic function. For this, let $\mathcal{H}$ be the Hilbert space $L^2(\mathbb{T})$ of all square-integrable, one-periodic functions on the torus $\mathbb{T}$. The convolution operator with respect to an absolutely integrable function $a \in L^1(\mathbb{T})$ is defined by

$$\mathrm{conv}_a[\phi](t) := (a * \phi)(t) = \int_{\mathbb{T}} a(t-s)\,\phi(s)\,\mathrm{d}s$$

for $t \in \mathbb{T}$. The convolution operator conv_a is here an isomorphism on $L^2(\mathbb{T})$ due to Young's convolution inequality, see e.g., [48], and is diagonalized by the finite Fourier transform $\mathcal{F}\colon L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$ given by

$$\mathcal{F}[\phi](k) := \hat{\phi}(k) := \int_{\mathbb{T}} \phi(t)\,\mathrm{e}^{-2\pi \mathrm{i} k t}\,\mathrm{d}t.$$

More precisely, we have $\mathcal{S}^{-1} = \mathcal{F}$, $\mathcal{Z} = \mathbb{Z}$, and $\Lambda\colon \psi \mapsto \hat{a} \odot \psi$, where $\odot$ denotes the elementwise multiplication. Due to the support constraints on the Fourier coefficients, the sampling vectors $\{\phi_i\}_{i=0}^{\infty}$ are trigonometric polynomials.

COROLLARY 8.2. *Let $a \in L^1(\mathbb{T})$ with absolutely collision-free Fourier coefficients $\hat{a}$ be unknown, let $\{\phi_j\}_{j=0}^{\infty}$ be a set of trigonometric polynomials allowing parameter identification and phase retrieval, and let $\hat{f}$ be elementwise non-zero for unknown $f \in L^2(\mathbb{T})$. Then a and f are defined by the spatiotemporal samples*

$$\bigl\{|\langle f, \mathrm{conv}_a^{\ell}[\phi_i]\rangle|\bigr\}_{\ell,i=0}^{L_i^2-1,\infty} \qquad \text{with} \qquad L_i := \#[\mathrm{supp}(\hat{\phi}_i)]$$

up to global phase.

The proposed eigenvalue and signal identification can be generalized to arbitrary Banach spaces $\mathcal{X}$ that are isomorphic to a sequence space like $\ell^p(\mathcal{Z})$. In this case, the inner products have to be replaced by appropriate dual pairings.

9 SENSITIVITY ANALYSIS

In the previous sections, we have shown that the dynamical phase retrieval and system identification problem is solvable under certain assumptions from exact measurements. In the following, we study the situation for disturbed measurements. Since our constructive proofs have been heavily based on Prony's method, the sensitivity also mainly depends on it. On the bases of Potts & Tasche [49], initially, the sensitivity of the approximate Prony method is considered; hereby, we follow the proofs of [49] for real-valued exponential sums and generalize to the complex setting. In a second step, we analyse the error propagation in dynamical phase retrieval.

9.1 SENSITIVITY OF PRONY'S METHOD Essentially, the (approximate) Prony method is a two step approach to determine the parameters of the exponential sum (1). In the first step, the unknown bases $\boldsymbol{\beta}$ are recovered using a singular value decomposition and determining the roots of the Prony polynomial. In the second, the unknown coefficients $\boldsymbol{\eta}$ are computed by solving a linear least-square problem. To analyse the sensitivity of the first step, we require the following lemma estimating the norm of a rectangular Vandermonde matrix by the maximal radius of the bases

$$\rho_{\boldsymbol{\beta}} \coloneqq \max\{1, \|\boldsymbol{\beta}\|_\infty\}.$$

LEMMA 9.1. *For $\boldsymbol{\beta} \in \mathbb{C}^K$, the Vandermonde matrix $\boldsymbol{V}_L(\boldsymbol{\beta})$ satisfies*

$$\|\boldsymbol{V}_L(\boldsymbol{\beta})\|_\infty \le K\rho_{\boldsymbol{\beta}}^{L-1}, \qquad \|\boldsymbol{V}_L(\boldsymbol{\beta})\|_1 \le L\rho_{\boldsymbol{\beta}}^{L-1},$$

and thus

$$\|\boldsymbol{V}_L(\boldsymbol{\beta})\|_2 \le \sqrt{KL}\rho_{\boldsymbol{\beta}}^{L-1}.$$

Proof. The assertion immediately follows from

$$\|\boldsymbol{V}_L(\boldsymbol{\beta})\|_\infty \le \max_{0\le\ell<L} \sum_{k=0}^{K-1} |\beta_k|^\ell \le K \max_{0\le\ell<L} \|\boldsymbol{\beta}\|_\infty^\ell \le K \max\{1, \|\boldsymbol{\beta}\|_\infty^{L-1}\},$$

$$\|\boldsymbol{V}_L(\boldsymbol{\beta})\|_1 \le \max_{0\le k<K} \sum_{\ell=0}^{L-1} |\beta_k|^\ell \le L \max_{0\le k<K} \Big(\max\{1, \beta_k^{L-1}\}\Big) \le L \max\{1, \|\boldsymbol{\beta}\|_\infty^{L-1}\},$$

$$\|\boldsymbol{V}_L(\boldsymbol{\beta})\|_2 \le \sqrt{\|\boldsymbol{V}_L(\boldsymbol{\beta})\|_1 \|\boldsymbol{V}_L(\boldsymbol{\beta})\|_\infty}. \qquad \square$$

Further, we need a left inverse of the rectangular Vandermonde matrix. The inverse of a quadratic Vandermonde matrix has been well studied in the literature [24, 25, 28, 29, 32, 42, 45, 57] and is given by

$$\boldsymbol{V}^{-1}(\boldsymbol{\beta}) = \Big((-1)^{K-k-1} S_{K-k-1}^{(\ell)}(\boldsymbol{\beta}) \Big/ \Pi_\ell(\boldsymbol{\beta})\Big)_{\ell,k=0}^{K-1}, \tag{11}$$

where $S_k^{(\ell)}$ denotes the kth elementary symmetric polynomial without the ℓth variable, which is more precisely defined by

$$S_k^{(\ell)}(\boldsymbol{\beta}) = \sum_{\substack{0\le j_1<\dots<j_k\le K-1 \\ j_1,\dots,j_k\ne\ell}} \beta_{j_1}\dots\beta_{j_k} \qquad \text{and} \qquad S_0^{(\ell)}(\boldsymbol{\beta}) = 1,$$

and where Π_ℓ is the product of differences

$$\Pi_\ell(\boldsymbol{\beta}) \coloneqq \prod_{\substack{k=0 \\ k\ne\ell}}^{K-1} (\beta_\ell - \beta_k).$$

The classical elementary symmetric polynomials are based on all elements of $\boldsymbol{\beta}$, i.e. without the condition $j_1, \dots, j_k \neq \ell$, and are denoted by $S_k(\boldsymbol{\beta})$.

Lemma 9.2 (Gautschi [28]). *The elementary symmetric polynomial are bounded by*

$$\sum_{k=0}^{K-1} |S_k(\boldsymbol{\beta})| \leq \prod_{k=0}^{K-1} (1 + |\beta_k|).$$

Proof. For convenience, we give the brief proof from [28]. On the bases of Vieta's formula, the elementary symmetric polynomials are related to the polynomial

$$z \mapsto \sum_{k=0}^{K-1} (-1)^k \, S_k(\boldsymbol{\beta}) \, z^{K-k-1} = \prod_{k=0}^{K-1} (z - \beta_k).$$

Choosing $z = -1$, we obtain the assertion for real and positive β_k, $k = 0, \dots, K-1$. The general assertion then follows from $|S_k(\boldsymbol{\beta})| \leq S_k(|\boldsymbol{\beta}|)$, where $|\cdot|$ is applied elementwise. □

Defining the product radius $\pi_{\boldsymbol{\beta}}$ and the minimal separation $\sigma_{\boldsymbol{\beta}}$ of the bases in $\boldsymbol{\beta}$ as

$$\pi_{\boldsymbol{\beta}} \coloneqq \prod_{k=0}^{K-1} (1 + |\beta_k|) \qquad \text{and} \qquad \sigma_{\boldsymbol{\beta}} \coloneqq \min\{|\beta_\ell - \beta_k| : 0 \leq \ell < k \leq K-1\},$$

the norm of the inverse Vandermonde matrix is bounded as follows.

Proposition 9.3. *For $\boldsymbol{\beta} \in \mathbb{C}_*^K$ with distinct elements, the inverse of the quadratic Vandermonde matrix $\boldsymbol{V}(\boldsymbol{\beta})$ satisfies*

$$\| \boldsymbol{V}^{-1}(\boldsymbol{\beta}) \|_\infty \leq \frac{\pi_{\boldsymbol{\beta}}}{\sigma_{\boldsymbol{\beta}}^{K-1}}.$$

Proof. The bound follows immediately from the inversion formula (11) and from applying Lemma 9.2 to the sum over the elementary symmetric polynomials $S_k^{(\ell)}$ with fixed ℓ as well as multiplying the estimated for the row sums by the missing factor $(1 + |\beta_\ell|) > 1$. □

The norm estimates regarding the Vandermonde matrix allow us to study the quality of the Prony polynomial for perturbed measurements. If the error is small, the true bases are nearly roots; so we may hope that the first two steps of Algorithm 3.2 approximate the bases well. Recall that the approximate Prony method is based on the assumption that the measurement error ϵ with $|h_\ell + e_\ell| \leq \epsilon$ is small enough such that the singular values of the unperturbed Hankel matrix fulfil $\sigma_k(\boldsymbol{H}) \geq 2\|\boldsymbol{E}\|_2$. The spectral norm is here bounded by

$$\|\boldsymbol{E}\|_2 \leq \sqrt{\|\boldsymbol{E}\|_1 \|\boldsymbol{E}\|_\infty} \leq \sqrt{(L-K)(K+1)}\, \epsilon \leq (L+1)\, \epsilon/2.$$

THEOREM 9.4. *Let $L > 2K$, and let $\tilde{\boldsymbol{\gamma}}$ be a normalized right singular vector to the smallest singular value $\tilde{\sigma}_K$ of the perturbed Hankel matrix (4) with respect to the exponential sum (1). Then the corresponding polynomial $\tilde{P}(z) = \sum_{k=0}^{K} \tilde{\gamma}_k z^k$ satisfies*

$$\sum_{k=0}^{K-1} |\eta_k|^2 |\tilde{P}(\beta_k)|^2 \le L \left(\frac{\pi_{\boldsymbol{\beta}}}{\sigma_{\boldsymbol{\beta}}^{K-1}} \right)^2 \left(\tilde{\sigma}_K + \| \boldsymbol{E} \|_2 \right)^2.$$

Proof. Let $\tilde{\boldsymbol{v}}$ be the corresponding left singular vector, i.e. $\tilde{\boldsymbol{H}} \tilde{\boldsymbol{\gamma}} = \tilde{\sigma}_K \tilde{\boldsymbol{v}}$. Incorporating (4) and (1) into this equation, we obtain

$$\tilde{\sigma}_K \tilde{v}_\ell = \sum_{k=0}^{K} \tilde{h}_{\ell+k} \tilde{\gamma}_k = \sum_{k=0}^{K} (h_{\ell+k} + e_{\ell+k})\, \tilde{\gamma}_k = \sum_{j=0}^{K-1} \eta_j \beta_j^\ell \, \tilde{P}(\beta_j) + \sum_{k=0}^{K} e_{\ell+k} \tilde{\gamma}_k$$

for $\ell = 0, \dots, L-K-1$. In matrix-vector form, these equations are given by

$$\boldsymbol{V}_{L-K}(\boldsymbol{\beta}) \left(\eta_j \tilde{P}(\beta_j) \right)_{j=0}^{K-1} = \tilde{\sigma}_K \tilde{\boldsymbol{v}} - \boldsymbol{E} \tilde{\boldsymbol{\gamma}}.$$

Multiplying with the left inverse $\boldsymbol{V}_{L-K}^{+}(\boldsymbol{\beta}) := \left(\begin{smallmatrix} \boldsymbol{V}^{-1}(\boldsymbol{\beta}) \\ 0_{L-2K,K} \end{smallmatrix} \right)$, we obtain

$$\left(\eta_j \tilde{P}(\beta_j) \right)_{j=0}^{K-1} = \boldsymbol{V}_{L-K}^{+}(\boldsymbol{\beta}) \left(\tilde{\sigma}_K \tilde{\boldsymbol{v}} - \boldsymbol{E} \tilde{\boldsymbol{\gamma}} \right).$$

Taking the squared Euclidean norm, bounding the spectral norm by the row-sum norm, and applying Proposition 9.3 yields the assertion. □

THEOREM 9.5. *Let $L > 2K$, let $\tilde{\boldsymbol{\gamma}}$ be a normalized right singular vector to the smallest singular value $\tilde{\sigma}_K$ of the perturbed Hankel matrix (4) with $|h_\ell - \tilde{h}_\ell| \le \epsilon$, and let σ_{K-1} be the smallest non-zero singular value of the unperturbed Hankel matrix (3). Then the corresponding polynomial $\tilde{P}(z) = \sum_{k=0}^{K} \tilde{\gamma}_k z^k$ satisfies*

$$\sum_{k=0}^{K-1} |\tilde{P}(\beta_k)|^2 \le K L \, \rho_{\boldsymbol{\beta}}^{2L-2} \, \frac{\left(\tilde{\sigma}_K + \| \boldsymbol{E} \|_2 \right)^2}{\sigma_{K-1}^2}$$

Proof. First assume $\tilde{\boldsymbol{\gamma}} \notin \ker \boldsymbol{H}$. Letting $\boldsymbol{\gamma} := \operatorname{proj}_{\ker \boldsymbol{H}} \tilde{\boldsymbol{\gamma}}$, the projection $\boldsymbol{\gamma}$ is a maybe not normalized right singular vector for the singular value zero. Lemma 3.1 implies that the polynomial $P(z) := \sum_{k=0}^{K} \gamma_k z^k$ has the roots $\beta_0, \dots, \beta_{K-1}$. Therefore, we can write

$$\sum_{k=0}^{K-1} |\tilde{P}(\beta_k)|^2 = \sum_{k=0}^{K-1} |\tilde{P}(\beta_k) - P(\beta_k)|^2 = \| \boldsymbol{V}^{\mathrm{T}}(\boldsymbol{\beta})\, \tilde{\boldsymbol{\gamma}} - \boldsymbol{V}^{\mathrm{T}}(\boldsymbol{\beta})\, \boldsymbol{\gamma} \|_2^2 \le \| \boldsymbol{V}(\boldsymbol{\beta}) \|_2^2 \, \| \tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \|_2^2.$$

Now since $(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) \perp \ker \boldsymbol{H}$, we obtain

$$\sigma_{K-1}^2 \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}\|_2^2 \le \|\boldsymbol{H}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma})\|_2^2 = \|(\tilde{\boldsymbol{H}} - \boldsymbol{E})\tilde{\boldsymbol{\gamma}}\|_2^2 \le \left(\tilde{\sigma}_K + \|\boldsymbol{E}\|_2\right)^2.$$

Combining the above inequalities, and applying Lemma 9.1, we establish the assertion. For the remaining case $\tilde{\boldsymbol{\gamma}} \in \ker \boldsymbol{H}$, the bases β_k are roots of $\tilde{P}$ by Lemma 3.1. □

Remark 9.6. The above Theorems 9.4 and 9.5 essentially state that the true bases are nearly roots of the perturbed Prony polynomial. Therefore, we nurture the hope that the perturbed roots are close. Although this seems plausible for generic polynomials, we can construct pathological cases of very sensitive polynomials, where already slight disturbances of the coefficients have tremendous effects on the roots. In [56], the author tries to establishes an explicit bound on the reconstruction error regarding the roots of the Prony polynomial, which we initially wanted to adapt to our setting. Unfortunately, the key theorem studying a linear perturbation of the coefficient of a polynomial cannot be applied to our setting since here the perturbations e_ℓ in the measurements $\tilde{h}_\ell = h_\ell + e_\ell$ lead to non-linear perturbations of the coefficients in the Prony polynomial. ○

In the third step of Prony's method, the coefficients $\boldsymbol{\eta}$ of the exponential sum (1) are determined by solving $\boldsymbol{V}_L(\boldsymbol{\beta})\,\boldsymbol{\eta} = \tilde{\boldsymbol{h}}$ in the least-square sense, i.e. we have to determine the minimizer of $\|\boldsymbol{V}_L(\boldsymbol{\beta})\,\boldsymbol{\eta} - \tilde{\boldsymbol{h}}\|_2$. The minimizer is given by $\boldsymbol{V}_L^\dagger(\boldsymbol{\beta})\,\tilde{\boldsymbol{h}}$, where

$$\boldsymbol{V}_L^\dagger(\boldsymbol{\beta}) = (\boldsymbol{V}_L^*(\boldsymbol{\beta})\,\boldsymbol{V}_L(\boldsymbol{\beta}))^{-1}\,\boldsymbol{V}_L^*(\boldsymbol{\beta})$$

is the Moore–Penrose inverse. To estimate the reconstruction error with respect to $\boldsymbol{\eta}$, we need to estimate the norm of the Moore–Penrose inverse. For this, we exploit that the Moore–Penrose inverse is the zero continuation of the inverse with respect to the range of the orthogonal complement of the kernel. For an arbitrary full-rank matrix, the Moore–Penrose inverse is therefore the left inverse with the smallest norm.

Proposition 9.7. *Let $\boldsymbol{A} \in \mathbb{C}^{L \times K}$ with $L \ge K$ be a full-rank matrix, and let $\boldsymbol{A}^+$ be an arbitrary left inverse. For every $1 \le p \le \infty$, the Moore–Penrose inverse then satisfies*

$$\|\boldsymbol{A}^\dagger\|_p \le \|\boldsymbol{A}^+\|_p.$$

Proof. Since every left inverse $\boldsymbol{A}^+$ fulfils $\boldsymbol{A}^+\boldsymbol{A} = \boldsymbol{I}$, all left inverses coincide on the range of $\boldsymbol{A}$. The Moore–Penrose inverse is now the unique zero continuation from the range to the whole space $\mathbb{C}^L$, which geometrically means that the Moore–Penrose inverse is the projection onto $\operatorname{ran} \boldsymbol{A}$ composed with the unique inverse on the range. For the induced matrix norm, this means

$$\|\boldsymbol{A}^+\|_p = \sup_{\|\boldsymbol{x}\|_p = 1} \|\boldsymbol{A}^+\boldsymbol{x}\|_p \ge \sup_{\substack{\|\boldsymbol{x}\|_p = 1\\ \boldsymbol{x} \in \operatorname{ran} \boldsymbol{A}}} \|\boldsymbol{A}^+\boldsymbol{x}\|_p = \sup_{\|\boldsymbol{x}\|_p = 1} \|\boldsymbol{A}^\dagger\boldsymbol{x}\|_p = \|\boldsymbol{A}^\dagger\|_p$$

because $(\operatorname{ran} \boldsymbol{A})^\perp = \ker \boldsymbol{A}^\dagger$. This argumentation holds for all induced matrix norms and not only for the p-norm. □

Using this property of the Moore–Penrose inverse, we may immediately estimate the condition number $\kappa(\boldsymbol{V}_L(\boldsymbol{\beta})) := \|\boldsymbol{V}_L^\dagger(\boldsymbol{\beta})\|_2 \|\boldsymbol{V}_L(\boldsymbol{\beta})\|_2$ of the Vandermonde matrix $\boldsymbol{V}_L(\boldsymbol{\beta})$ if the bases $\boldsymbol{\beta}$ are known.

PROPOSITION 9.8. *The condition number of the Vandermonde matrix $\boldsymbol{V}_L(\boldsymbol{\beta})$ is bounded by*

$$\kappa(\boldsymbol{V}_L(\boldsymbol{\beta})) \le \sqrt{K}\, L\, \frac{\pi_{\boldsymbol{\beta}}\, \rho_{\boldsymbol{\beta}}^{L-1}}{\sigma_{\boldsymbol{\beta}}^{K-1}},$$

Proof. The bound follows from Lemma 9.1 and from Proposition 9.7 and 9.3 with the left inverse $\boldsymbol{V}_{L-K}^{+}(\boldsymbol{\beta}) := \begin{pmatrix} V^{-1}(\boldsymbol{\beta}) \\ 0_{L-2K,K} \end{pmatrix}$. □

PROPOSITION 9.9. *Let $\boldsymbol{\eta}$ and $\boldsymbol{\beta}$ be the parameters of the exponential sum* (1). *The least-squares solution $\tilde{\boldsymbol{\eta}}$ of the perturbed equation system $\boldsymbol{V}_L(\boldsymbol{\beta})\,\tilde{\boldsymbol{\eta}} = \tilde{\boldsymbol{h}}$ with $\|\boldsymbol{h} - \tilde{\boldsymbol{h}}\|_\infty \le \epsilon$ satisfies*

$$\|\boldsymbol{\eta} - \tilde{\boldsymbol{\eta}}\|_\infty \le \frac{\pi_{\boldsymbol{\beta}}}{\sigma_{\boldsymbol{\beta}}^{K-1}}\,\epsilon.$$

Proof. The inequality follows immediately from $\|\boldsymbol{\eta} - \tilde{\boldsymbol{\eta}}\|_\infty \le \|\boldsymbol{V}_L^\dagger(\boldsymbol{\beta})\|_\infty \|\boldsymbol{h} - \tilde{\boldsymbol{h}}\|_\infty$ and from applying Proposition 9.7 and 9.3 with the left inverse $\boldsymbol{V}_{L-K}^{+}(\boldsymbol{\beta}) := \begin{pmatrix} V^{-1}(\boldsymbol{\beta}) \\ 0_{L-2K,K} \end{pmatrix}$. □

Certainly, the computed bases $\tilde{\boldsymbol{\beta}}$ are themselves only approximations of $\boldsymbol{\beta}$ in practice. Therefore, besides the right-hand side $\tilde{\boldsymbol{h}}$, the Vandermonde matrix $\boldsymbol{V}_L(\tilde{\boldsymbol{\beta}})$ is perturbed too. For studying the effect to the recovered coefficients, we need the following lemmata.

LEMMA 9.10. *For $\boldsymbol{\beta} \in \mathbb{C}^K$, and for $\tilde{\boldsymbol{\beta}} \in \mathbb{C}^K$ with $\|\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}}\|_\infty \le \delta$, it holds*

$$\pi_{\tilde{\boldsymbol{\beta}}} \le \pi_{|\boldsymbol{\beta}|+\delta \mathbb{1}}.$$

Proof. The lemma is established by

$$\pi_{\tilde{\boldsymbol{\beta}}} = \prod_{k=0}^{K-1} (1 + |\tilde{\beta}_k|) \le \prod_{k=0}^{K-1} (1 + |\beta_k| + \delta) = \pi_{|\boldsymbol{\beta}|+\delta \mathbb{1}}.$$ □

LEMMA 9.11. *For $\boldsymbol{\beta} \in \mathbb{C}^K$, and for $\tilde{\boldsymbol{\beta}} \in \mathbb{C}^K$ with $\|\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}}\|_\infty \le \delta$, it holds*

$$\sigma_{\tilde{\boldsymbol{\beta}}} \ge \sigma_{\boldsymbol{\beta}} - 2\delta.$$

Proof. Using the triangle inequality, we may estimate the minimal separation by

$$|\tilde{\beta}_\ell - \tilde{\beta}_k| \ge |\beta_\ell - \beta_k| - |\beta_\ell - \tilde{\beta}_\ell| - |\beta_k - \tilde{\beta}_k| \ge |\beta_\ell - \beta_k| - 2\delta.$$ □

Lemma 9.12. *For $\boldsymbol{\beta} \in \mathbb{C}^K$, and for $\tilde{\boldsymbol{\beta}} \in \mathbb{C}^K$ with $\|\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}}\|_\infty \le \delta$, it holds*

$$\|\boldsymbol{V}_L(\tilde{\boldsymbol{\beta}}) - \boldsymbol{V}_L(\boldsymbol{\beta})\|_\infty \le \sqrt{2}\, KL\, \rho^{L-1}_{|\boldsymbol{\beta}|+\delta 1}\, \delta.$$

Proof. We use the following complex mean value theorem [26, Thm 2.2]: Let f be a holomorphic function defined on an open convex set $D \subset \mathbb{C}$, and let a and b be two distinct points in D. Then there exist $z_1, z_2 \in (a, b)$ such that

$$\Re\big(f'(z_1)\big) = \Re\left(\frac{f(b) - f(a)}{b-a}\right) \qquad \text{and} \qquad \Im\big(f'(z_2)\big) = \Im\left(\frac{f(b) - f(a)}{b-a}\right),$$

where (a, b) denotes the open line segment

$$(a, b) \coloneqq \{a + t(b-a) : t \in (0,1)\}.$$

On the basis of this complex mean value theorem, we obtain

$$\|\boldsymbol{V}_L(\boldsymbol{\beta}) - \boldsymbol{V}_L(\tilde{\boldsymbol{\beta}})\|_\infty = \max_{0\le \ell < L} \sum_{k=0}^{K-1} |\beta_k^\ell - \tilde{\beta}_k^\ell| = \max_{0\le \ell < L} \sum_{k=0}^{K-1} \ell |\beta_k - \tilde{\beta}_k| \,|\Re(\xi_{\ell,k}^{\ell-1}) + \mathrm{i}\Im(\zeta_{\ell,k}^{\ell-1})|$$

with intermediate points $\xi_{\ell,k}, \zeta_{\ell,k} \in (\beta_k, \tilde{\beta}_k)$. Since $|\xi_{\ell,k}| \le |\beta_k| + \delta$ as well as $|\zeta_{\ell,k}| \le |\beta_k| + \delta$, we finally have

$$\|\boldsymbol{V}_L(\boldsymbol{\beta}) - \boldsymbol{V}_L(\tilde{\boldsymbol{\beta}})\|_\infty \le \max_{\substack{0\le \ell < L\\ 0 \le k < K}} \sqrt{2}\, \ell K \,(|\beta_k| + \delta)^{\ell-1}\, \delta \le \sqrt{2}\, KL\, \rho^{L-1}_{|\boldsymbol{\beta}|+\delta 1}\, \delta. \qquad \square$$

Theorem 9.13. *Let $\boldsymbol{\eta}$ and $\boldsymbol{\beta}$ be the parameters of the exponential sum* (1). *The least-squares solution $\tilde{\boldsymbol{\eta}}$ of the perturbed equation system $\boldsymbol{V}_L(\tilde{\boldsymbol{\beta}})\, \tilde{\boldsymbol{\eta}} = \tilde{\boldsymbol{h}}$ with $\|\boldsymbol{h} - \tilde{\boldsymbol{h}}\|_\infty \le \epsilon$, $\|\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}}\|_\infty \le \delta$, and $\delta < \sigma_{\boldsymbol{\beta}}/2$ satisfies*

$$\|\boldsymbol{\eta} - \tilde{\boldsymbol{\eta}}\|_\infty \le \frac{\pi_{|\boldsymbol{\beta}|+\delta 1}}{(\sigma_{\boldsymbol{\beta}} - 2\delta)^{K-1}} \left(\sqrt{2}\, KL\, \frac{\pi_{\boldsymbol{\beta}}\, \rho^{L-1}_{|\boldsymbol{\beta}|+\delta 1}}{\sigma_{\boldsymbol{\beta}}^{K-1}} \|\boldsymbol{h}\|_\infty\, \delta + \epsilon\right).$$

Proof. Due to $\delta < \sigma_{\boldsymbol{\beta}}/2$, the perturbed Vandermonde matrix $\boldsymbol{V}_L(\tilde{\boldsymbol{\beta}})$ has full rank. Further, the reconstruction error may be estimated by

$$\begin{aligned}
\|\boldsymbol{\eta} - \tilde{\boldsymbol{\eta}}\|_\infty &= \|\boldsymbol{\eta} - \boldsymbol{V}_L^\dagger(\tilde{\boldsymbol{\beta}})\, \tilde{\boldsymbol{h}}\|_\infty \\
&= \|\boldsymbol{V}_L^\dagger(\tilde{\boldsymbol{\beta}})\, \boldsymbol{V}_L(\tilde{\boldsymbol{\beta}})\boldsymbol{\eta} - \boldsymbol{V}_L^\dagger(\tilde{\boldsymbol{\beta}})\, \boldsymbol{V}_L(\boldsymbol{\beta})\, \boldsymbol{\eta} + \boldsymbol{V}_L^\dagger(\tilde{\boldsymbol{\beta}})(\boldsymbol{h} - \tilde{\boldsymbol{h}})\|_\infty \\
&\le \|\boldsymbol{V}_L^\dagger(\tilde{\boldsymbol{\beta}})\|_\infty \big(\|\boldsymbol{V}_L(\tilde{\boldsymbol{\beta}}) - \boldsymbol{V}_L(\boldsymbol{\beta})\|_\infty \|\boldsymbol{\eta}\|_\infty + \|\boldsymbol{h} - \tilde{\boldsymbol{h}}\|_\infty\big)
\end{aligned}$$

The first factor may be estimated by applying Proposition 9.7 with perturbed left inverse $\boldsymbol{V}^{+}_{L-K}(\tilde{\boldsymbol{\beta}}) \coloneqq \binom{V^{-1}(\tilde{\boldsymbol{\beta}})}{0_{L-2K,K}}$ followed by Proposition 9.3, Lemma 9.10, and Lemma 9.11 yielding

$$\|\boldsymbol{V}_L^{\dagger}(\tilde{\boldsymbol{\beta}})\|_\infty \le \frac{\pi_{\tilde{\boldsymbol{\beta}}}}{\sigma_{\tilde{\boldsymbol{\beta}}}^{K-1}} \le \frac{\pi_{|\boldsymbol{\beta}|+\delta 1}}{(\sigma_{\boldsymbol{\beta}} - 2\delta)^{K-1}}.$$

Using Lemma 9.12 and that $\|\boldsymbol{\eta}\|_\infty \le \|\boldsymbol{V}_L^{\dagger}(\boldsymbol{\beta})\|_\infty \|\boldsymbol{h}\|_\infty$ together with Proposition 9.7 and Proposition 9.3, we finally arrive at

$$\|\boldsymbol{\eta} - \tilde{\boldsymbol{\eta}}\|_\infty \le \frac{\pi_{|\boldsymbol{\beta}|+\delta 1}}{(\sigma_{\boldsymbol{\beta}} - 2\delta)^{K-1}} \left(\sqrt{2}\, KL\, \rho_{|\boldsymbol{\beta}|+\delta 1}^{L-1}\, \delta\, \frac{\pi_{\boldsymbol{\beta}}}{\sigma_{\boldsymbol{\beta}}^{K-1}} \|\boldsymbol{h}\|_\infty + \epsilon \right). \qquad \square$$

9.2 Sensitivity of phase & system identification On the basis of the sensitivity analysis of Prony's method, we analyse the error propagation in dynamical phase retrieval. For this, we assume that the unknown bases $\lambda_j \bar{\lambda}_k$ and coefficients $c_j \bar{c}_k$ of the exponential sum describing the measurements (6) have been approximately computed. In the following, we denote the true bases and coefficients by

$$\beta_{\tau(j,k)} = \lambda_j \bar{\lambda}_k \qquad \text{and} \qquad \eta_{\tau(j,k)} = c_j \bar{c}_k, \tag{12}$$

where the bijective map

$$\tau\colon \{0,\dots,d-1\} \times \{0,\dots,d-1\} \to \{0,\dots,d^2-1\}$$

describes the relation between the indices. Assuming that the recovered bases $\tilde{\boldsymbol{\beta}}$ and coefficients $\tilde{\boldsymbol{\eta}}$ satisfy $\|\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_\infty \le \delta$ and $\|\tilde{\boldsymbol{\eta}} - \boldsymbol{\eta}\|_\infty \le \varepsilon$, where δ should be small enough such that the mapping τ can be recovered up to the winding direction by the above constructive proofs, i.e. the error is small enough such that the order of the absolute values $|\beta_k|$ remains unchanged, we want to estimate the errors in the recovered spectrum $\tilde{\boldsymbol{\lambda}}$ and signal $\tilde{\boldsymbol{x}}$. Note that $\tilde{\beta}_{\tau(j,k)}$ and $\tilde{\eta}_{\tau(j,k)}$ are simply conjugated for the opposite winding direction.

In line with the above procedures, where firstly the magnitudes of the unknown variables are determined, and secondly the phase is propagated between the elements, we decouple the sensitivity analysis of absolute value and phase. Further, we first discuss the sensitivity of the unknown operator spectrum, followed by the analysis of the unknown signal, and finally the error propagation for multiple sampling vectors.

Sensitivity of the spectrum The recovered bases $\tilde{\boldsymbol{\beta}}$ already contain estimates of the squared modulus of the spectrum $\boldsymbol{\lambda}$. After recovering the relation τ (up to winding direction), the magnitude of the spectrum is easily obtained by taking the square root, i.e.

$$|\tilde{\lambda}_j| \coloneqq \sqrt{|\tilde{\beta}_{\tau(j,j)}|}. \tag{13}$$

The sensitivity of the magnitude computation may be easily estimated via the mean value theorem.

LEMMA 9.14. *Assume* $|\tilde{\beta}_{\tau(j,j)} - \beta_{\tau(j,j)}| \le \delta$*, and estimate the magnitude* $|\lambda_j|$ *by* (13)*. If* $\delta < |\lambda_j|^2$*, then we have*

$$\big||\tilde{\lambda}_j| - |\lambda_j|\big| \le \frac{\delta}{2\sqrt{|\lambda_j|^2 - \delta}}$$

and, for $\delta < |\lambda_j|^2/2$*, in particular*

$$\big||\tilde{\lambda}_j| - |\lambda_j|\big| \le \frac{\sqrt{2}\,\delta}{2\sqrt{|\lambda_j|^2}}.$$

Proof. The statement immediately follows from applying the mean value theorem and the reversed triangle inequality by

$$\big||\tilde{\lambda}_j| - |\lambda_j|\big| = \big||\tilde{\beta}_{\tau(j,j)}|^{1/2} - |\beta_{\tau(j,j)}|^{1/2}\big| \le \frac{\delta}{2\sqrt{|\beta_{\tau(j,j)}| - \delta}}.$$

The second one is a trivial consequence. □

Recall that for computing the phase of $\tilde{\lambda}_j$, we first find the element with the largest magnitude, say $\tilde{\lambda}_k$, then set the phase of $\tilde{\lambda}_k$ to be zero due to the global phase ambiguity, and finally propagate the phase to $\tilde{\lambda}_j$ using the relative phase encoded in $\beta_{\tau(j,k)}$. More precisely, exploiting $\tilde{\beta}_{\tau(j,k)} \approx \lambda_j \bar{\lambda}_k$, we retrieve the phase of λ_j by

$$\tilde{\lambda}_j := \frac{\tilde{\beta}_{\tau(j,k)}}{|\tilde{\beta}_{\tau(j,k)}|}\,|\tilde{\lambda}_j|, \tag{14}$$

where $|\tilde{\lambda}_j|$ has been computed by (13) in the first step. Note that this phase propagation is a very simple method, which however allow to analyse the propagation error. For doing this, we assume that the map τ given in (12) has been identified with respect to the true winding direction. Otherwise, we consider the conjugated recovered spectrum $\bar{\tilde{\boldsymbol{\lambda}}}$ without loss of generality. For simplicity, we first consider the phase propagation only between two elements. The idea of the proof was motivated by [35].

LEMMA 9.15. *Assume* $|\tilde{\beta}_{\tau(j,k)} - \beta_{\tau(j,k)}| \le \delta$*, suppose that* λ_k *is real and positive, and estimate the phase* $\arg(\lambda_j)$ *by* (14)*. If* $\delta < |\lambda_j||\lambda_k|$*, then we have*

$$|\arg(\tilde{\lambda}_j) - \arg(\lambda_j) \bmod 2\pi| \le \frac{2\delta}{|\lambda_k||\lambda_j|}.$$

Proof. Since λ_k is supposed to be real and positive, the phase of λ_j is directly encoded in the basis $\beta_{\tau(j,k)}$ by

$$\arg(\beta_{\tau(j,k)}) = \arg(\lambda_j) - \arg(\lambda_k) \bmod 2\pi = \arg(\lambda_j).$$

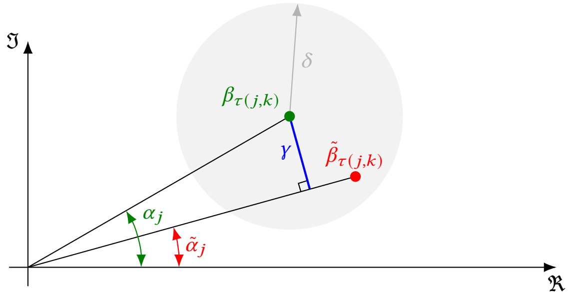


Figure 2: Geometrical relation between $\beta_{\tau(j,k)}$ and $\beta_{\tau(j,k)}$. In the proof of Lemma 9.15, we exploit the right-angled triangle between the rays with angle α_j and $\tilde{\alpha}_j$. Note that the point $\tilde{\beta}_{\tau(j,k)}$ may lay on the adjacent. The opposite γ is of length δ at the most.

During the proof, we denote the phases of $\beta_{\tau(j,k)}$ and $\tilde{\beta}_{\tau(j,k)}$ or λ_j and $\tilde{\lambda}_j$ by α_j and $\tilde{\alpha}_j$ respectively. Because of $|\tilde{\beta}_{\tau(j,k)} - \beta_{\tau(j,k)}| \le \delta < |\beta_{\tau(j,k)}|$, the phase difference $|\tilde{\alpha}_j - \alpha_j \bmod 2\pi|$ is always smaller than $\pi/2$. Thus we have

$$|\tilde{\alpha}_j - \alpha_j \bmod 2\pi| \le 2\sin(|\tilde{\alpha}_j - \alpha_j \bmod 2\pi|).$$

To estimate the sine of the phase difference, we exploit the geometrical relation between $\beta_{\tau(j,k)}$ and $\tilde{\beta}_{\tau(j,k)}$ schematically presented in Figure 2. Using the best-known sine relation of the right-angled triangle, we have

$$|\tilde{\alpha}_j - \alpha_j \bmod 2\pi| \le \frac{2\gamma}{|\beta_{\tau(j,k)}|} \le \frac{2\delta}{|\beta_{\tau(j,k)}|}. \qquad \square$$

Coupling the recovery of absolute values and the phase, we may estimate the total recovery error for the spectrum $\boldsymbol{\lambda}$, which mainly depends on $\|\boldsymbol{\lambda}\|_{-\infty}$.

PROPOSITION 9.16. *Assume $\|\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_\infty \le \delta$, and estimate $\boldsymbol{\lambda}$ by* (13) *and* (14)*, where the true winding direction is used without loss of generality, and where the phase is propagated from the element largest in magnitude. If $\delta < \|\boldsymbol{\lambda}\|^2_{-\infty}$, then we have*

$$\|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \le \left(\frac{2\sqrt{2}}{\|\boldsymbol{\lambda}\|_{-\infty}} + \frac{1}{2\sqrt{\|\boldsymbol{\lambda}\|^2_{-\infty} - \delta}}\right)\delta$$

and, for $\delta \le \|\boldsymbol{\lambda}\|^2_{-\infty}/2$, in particular

$$\|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \le \frac{5\sqrt{2}\,\delta}{2\|\boldsymbol{\lambda}\|_{-\infty}}.$$

Proof. Let $\tilde{\alpha}_j, \alpha_j$ be the phases of $\tilde{\lambda}_j, \lambda_j$ respectively. We decouple the phase and magnitude error by

$$|\tilde{\lambda}_j - \lambda_j| = \big||\tilde{\lambda}_j|\,\mathrm{e}^{\mathrm{i}\tilde{\alpha}_j} \pm |\tilde{\lambda}_j|\,\mathrm{e}^{\mathrm{i}\alpha_j} - |\lambda_j|\mathrm{e}^{\mathrm{i}\alpha_j}\big| \le |\tilde{\lambda}_j|\,|\mathrm{e}^{\mathrm{i}\tilde{\alpha}_j} - \mathrm{e}^{\mathrm{i}\alpha_j}| + \big||\tilde{\lambda}_j| - |\lambda_j|\big|.$$

The magnitude error may be simply estimated using Lemma 9.14 via

$$\big| |\tilde{\lambda}_j| - |\lambda_j| \big| \leq \frac{\delta}{2\sqrt{\|\boldsymbol{\lambda}\|_{-\infty}^2 - \delta}}.$$

For the phase error, assume that λ_k is the eigenvalue with largest magnitude, set $\arg(\lambda_k) = 0$, and propagate the phase from λ_k to the remaining λ_j by (14). The difference between the unimodular exponentials is now

$$\begin{aligned} |\mathrm{e}^{\mathrm{i}\tilde{\alpha}_j} - \mathrm{e}^{\mathrm{i}\alpha_j}| &= |\mathrm{e}^{\mathrm{i}(\tilde{\alpha}_j - \alpha_j)/2} - \mathrm{e}^{-\mathrm{i}(\tilde{\alpha}_j - \alpha_j)/2}| = 2|\sin(^{(\tilde{\alpha}_j - \alpha_j)}/_2)| \\ &\leq |\tilde{\alpha}_j - \alpha_j \bmod 2\pi| \leq \frac{2\delta}{\|\boldsymbol{\lambda}\|_\infty \|\boldsymbol{\lambda}\|_{-\infty}}, \end{aligned}$$

where the last inequality holds by Lemma 9.15. Using $|\tilde{\lambda}_j| \leq \sqrt{\|\boldsymbol{\lambda}\|_\infty^2 + \delta} \leq \sqrt{2}\|\boldsymbol{\lambda}\|_\infty$, we finally arrive at

$$\|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \frac{2\delta\sqrt{\|\boldsymbol{\lambda}\|_\infty^2 + \delta}}{\|\boldsymbol{\lambda}\|_\infty \|\boldsymbol{\lambda}\|_{-\infty}} + \frac{\delta}{2\sqrt{\|\boldsymbol{\lambda}\|_{-\infty}^2 - \delta}} \leq \left(\frac{2\sqrt{2}}{\|\boldsymbol{\lambda}\|_{-\infty}} + \frac{1}{2\sqrt{\|\boldsymbol{\lambda}\|_{-\infty}^2 - \delta}} \right) \delta.$$

If $\delta < {}^{\|\boldsymbol{\lambda}\|_{-\infty}}/_2$, we obtain

$$\|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \frac{2\sqrt{2}\,\delta}{\|\boldsymbol{\lambda}\|_{-\infty}} + \frac{\sqrt{2}\,\delta}{2\|\boldsymbol{\lambda}\|_{-\infty}} \leq \frac{5\sqrt{2}\delta}{2\|\boldsymbol{\lambda}\|_{-\infty}}.$$ □

Sensitivity of the signal As discussed in the previous sections, the components of $\tilde{\boldsymbol{\eta}}$ are in line with the structure of (12) meaning

$$\tilde{\eta}_{\tau(j,k)} \approx c_j \bar{c}_k \qquad \text{with} \qquad c_j = \bar{y}_j \psi_j = (\overline{S^* \boldsymbol{x}})_j \, (S^{-1}\boldsymbol{\phi})_j.$$

With respect to the above proofs, we recover the transformed signal $\boldsymbol{y} = \boldsymbol{S}^*\boldsymbol{x}$ similar to the spectrum $\boldsymbol{\lambda}$. Thus, we first recover the magnitudes via the real and positive values $\tilde{\eta}_{\tau(j,j)}$, then assume that $\tilde{c}_k$ largest in magnitude is real and positive, and spread the phase from $\tilde{c}_k$ to every other $\tilde{c}_j$ using the relative phase encoded in $\tilde{\eta}_{\tau(j,k)}$. Because of $y_j = c_j \psi_j^{-1}$ resulting in $|y_j| = |\psi_j^{-1}| \sqrt{\eta_{\tau(j,j)}}$ and $\arg(y_j) = \arg(\eta_{\tau(j,k)}) - \arg(\psi_j)$, we compute the tranformed components via

$$|\tilde{y}_j| := \frac{\sqrt{|\tilde{\eta}_{\tau(j,j)}|}}{|\psi_j|} \qquad \text{and} \qquad y_j := \frac{\eta_{\tau(j,k)}}{|\eta_{\tau(j,k)}|} \frac{\bar{\psi}_j}{|\psi_j|} |y_j|. \tag{15}$$

Adapting the considerations in the previous paragraph for the spectrum, we obtain the following sensitivities.

LEMMA 9.17. *Assume* $|\tilde{\eta}_{\tau(j,j)} - \eta_{\tau(j,j)}| \le \epsilon$, *and estimate the magnitude* $|y_j|$ *by* (15). *If* $\epsilon < |y_j|^2|\psi_j|^2$, *then we have*

$$\big||\tilde{y}_j| - |y_j|\big| \le \frac{\epsilon}{2|\psi_j|\sqrt{|y_j|^2|\psi_j|^2 - \epsilon}}.$$

Proof. Consider $||\tilde{y}_j| - |y_j|| = |\psi_j^{-1}|\,||\tilde{c}_j| - |c_j||$ and use the arguments in Lemma 9.14. □

LEMMA 9.18. *Assume* $|\tilde{\eta}_{\tau(j,k)} - \eta_{\tau(j,k)}| \le \epsilon$, *suppose that* y_k *is real and positive, and estimate the phase* $\arg(y_j)$ *by* (15). *If* $\epsilon < |y_j||y_k||\psi_j||\psi_k|$, *then we have*

$$|\arg(\tilde{y}_j) - \arg(y_j) \bmod 2\pi| \le \frac{2\epsilon}{|y_k||y_j||\psi_k||\psi_j|}.$$

Proof. Note that the phase difference may be written as

$$|\arg(\tilde{y}_j) - \arg(y_j)| = |\arg(\tilde{c}_j) - \arg(\psi_j) - \arg(c_j) + \arg(\psi_j)| = |\arg(\tilde{c}_j) - \arg(c_j)|,$$

and use the arguments of Lemma 9.18. □

PROPOSITION 9.19. *Assume* $\|\tilde{\boldsymbol{\eta}} - \boldsymbol{\eta}\|_\infty \le \epsilon$, *and estimate* $\boldsymbol{y}$ *by* (15), *where the true winding direction is used without loss of generality, and where the phase is propagated form the element largest in magnitude. If* $\epsilon < \|\boldsymbol{y}\|_{-\infty}^2 \|\boldsymbol{\psi}\|_{-\infty}^2$, *then we have*

$$\|\tilde{\boldsymbol{y}} - \boldsymbol{y}\|_\infty \le \left(\frac{2\sqrt{2}\,\|\boldsymbol{y}\|_\infty\,\|\boldsymbol{\psi}\|_\infty}{\|\boldsymbol{y}\|_{-\infty}^2\,\|\boldsymbol{\psi}\|_{-\infty}^2} + \frac{1}{2\,\|\boldsymbol{\psi}\|_{-\infty}\sqrt{\|\boldsymbol{y}\|_{-\infty}^2\,\|\boldsymbol{\psi}\|_{-\infty}^2 - \epsilon}}\right)\epsilon$$

and thus

$$\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|_\infty \le \left(\frac{2\sqrt{2}\,\|\boldsymbol{y}\|_\infty\,\|\boldsymbol{\psi}\|_\infty}{\|\boldsymbol{y}\|_{-\infty}^2\,\|\boldsymbol{\psi}\|_{-\infty}^2} + \frac{1}{2\,\|\boldsymbol{\psi}\|_{-\infty}\sqrt{\|\boldsymbol{y}\|_{-\infty}^2\,\|\boldsymbol{\psi}\|_{-\infty}^2 - \epsilon}}\right)\|S^{-1}\|_1\,\epsilon.$$

Proof. The statement follows using the same technique as for Proposition 9.16. Notice however that in the last estimate $|y_k|$ and $|\psi_k|$ would not have to correspond to $\|\boldsymbol{y}\|_\infty$ and $\|\boldsymbol{\psi}\|_\infty$ respectively since the phase is propagated from the coefficient $\tilde{c}_k \approx \bar{y}_k\psi_k$ largest in magnitude. Therefore the maximum norms do not cancel out. For the second part, exploit $\boldsymbol{x} = (S^*)^{-1}\boldsymbol{y}$ and $\|(S^{-1})^*\|_\infty = \|S^{-1}\|_1$. □

MULTIPLE SAMPLING VECTORS Finally, we would like to discuss the sensitivity of the phase propagation in the setting of Theorem 7.1, where we exploit spatiotemporal measurements with respect to several sampling vectors $\boldsymbol{\phi}_i$. Here we first recover the partial spectra $\tilde{\Lambda}_i = \{\tilde{\lambda}_k : k \in \operatorname{supp}\psi_i\}$ up to global phase and winding direction, then identify the order within the partial spectra, and afterwards align these to find the complete spectrum of $\boldsymbol{A}$ with one unified global phase and winding direction. In this process an extra

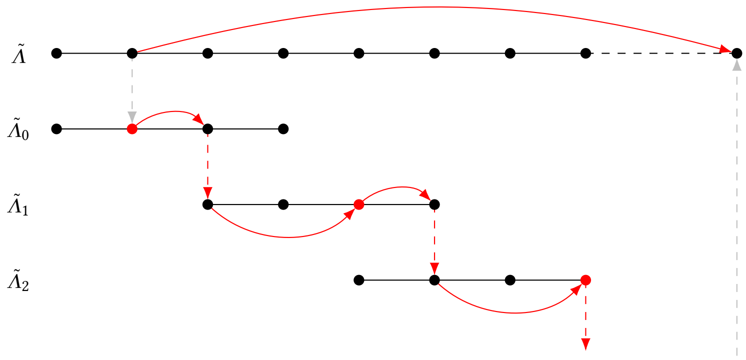


Figure 3: Schematic example for the propagation of the phase from some starting element over some path to another element. The elements of $\tilde{\Lambda}_i$ with the largest magnitude are marked in red. In each partial spectra, the phase error get worse by 2ρ at the most.

error will appear in the phase of eigenvalues because of the phase propagation between the partial spectra. Fortunately, the amplitude of the eigenvalues is not affected.

To demonstrate the issue in more detail, let us – for the moment – consider two partial spectra $\tilde{\Lambda}_0$ and $\tilde{\Lambda}_1$ and assume

$$|\arg(\tilde{\lambda}_{i,k}) - \arg(\lambda_k) \bmod 2\pi| \le \rho$$

if λ_k is covered by $\tilde{\Lambda}_i$. For simplicity, we assume that the winding directions are already aligned. If we now propagate the phase from $\tilde{\Lambda}_0$ over $\tilde{\lambda}_{0,k}$ and $\tilde{\lambda}_{1,k}$ to $\tilde{\Lambda}_1$, then the phases in $\tilde{\Lambda}_1$ have to be shifted by $\arg(\tilde{\lambda}_{0,k}) - \arg(\tilde{\lambda}_{1,k})$. Since the phase of $\tilde{\lambda}_{1,k}$ is already defective, the error within $\tilde{\Lambda}_1$ may accumulate at most to 2ρ. If we want to align the global phase of the entire spectrum, we may take the element with the largest magnitude in $\tilde{\Lambda}_0$, look for the shortest path over the partial spectra $\tilde{\Lambda}_i$ to $\tilde{\lambda}_j$, and propagate the phase along this path. The error of $\arg(\tilde{\lambda}_j)$ may then accumulate at most to $[1 + 2(M - 1)]\rho$, where M is the number of the employed spectra Λ_i. A schematic example of this procedure is shown in Figure 3. For the phase of the transformed signal $\boldsymbol{y}$, we may apply the same procedure.

10 Numerical examples

The constructive proofs of the uniqueness guarantees for phase retrieval and system identification can immediately be implemented to obtain numerical algorithms. Because of the sensitivity of Prony's method as corner stone of the proofs, these methods will however be vulnerable to noise. Nevertheless, we provide some small numerical examples to accompany the theoretical results and to show that simultaneous identification of system and signal is possible in principle. All numerical experiments have been implemented in Julia[1].

Example 10.1 (Prony's method). First, we apply the approximated Prony method in Algorithm 3.2 to the complex setting. For this, we generate exponential sums (1) by choosing the coefficients and bases from a ring in the complex plane. More precisely, the absolute values are drawn with respect to the uniform distributions $|\eta_k| \sim \mathcal{U}([1/8, 1])$ and $|\beta_k| \sim \mathcal{U}([1/2, 1])$ and the phases form $\mathcal{U}((-\pi, \pi])$ independently. The mean maximal reconstruction errors for different numbers of addends K and numbers of samples L. The results over 5 000 reconstructions are recorded in Table 1 and 2. For a small number of addends, the parameter are identified fairly well. Increasing the number of addends however leads to a significant loss of accuracy. To some degree, this may be compensated by employing more samples. We repeat this experiment with small additive noise $|e_k| \sim U([0, 10^{-10}])$ and $\arg(e_k) \sim \mathcal{U}((-\pi, \pi])$, see Table 3 and 4. ○

Example 10.2 (Simultaneous signal & system identification). In this numerical example, we consider the recovery of real-valued signals and convolution kernels as discussed in Section 6. The true, unknown kernel $\boldsymbol{a} \in \mathbb{R}^6$ is here chosen as

$$\hat{\boldsymbol{a}} := \big(\cos(2k)\big)_{k=-3}^{2},$$

where the indices are considered modulo 6. Besides the strictly, symmetrically decreasing kernel, the unknown signal $\boldsymbol{x} \in \mathbb{R}^6$ and the known measurement vectors $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2 \in \mathbb{R}^6$ have been randomly generated such that the requirements for the reconstruction are fulfilled, i.e. $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$ are pointwise independent in the frequency domain, and the assumption $\Re[\bar{\hat{x}}_k \hat{\phi}_{i,k}] \neq 0$ is satisfied for $k = 0, \dots, 5$, $i = 1, 2$. For reproducibility, the employed signals are shown in Table 5. Choosing $L := 4d^2 + 1 = 145$ to encounter the numerical sensitivity of Prony's method, we now apply the procedure in the constructive proof of Theorem 6.3. The reconstructions $\tilde{\boldsymbol{a}}$ and $\tilde{\boldsymbol{x}}$ of the true signals $\boldsymbol{a}$ and $\boldsymbol{x}$ are shown in Figure 4. Aligning the overall sign of $\boldsymbol{x}$ and $\tilde{\boldsymbol{x}}$, we are able to recover the unknown signals up to an error of $\|\hat{\boldsymbol{a}} - \hat{\tilde{\boldsymbol{a}}}\|_\infty = 8.650 \cdot 10^{-5}$ and $\|\boldsymbol{x} - \tilde{\boldsymbol{x}}\|_\infty = 1.141 \cdot 10^{-3}$. The theoretical procedure behind Theorem 6.3 thus allows the simultaneous recovery of signal and kernel numerically at least for small instances. ○

Example 10.3 (Multiple sampling vectors). Finally, we consider the identification of complex-valued signals and convolution kernels, i.e. $\boldsymbol{A} := \operatorname{circ} \boldsymbol{a}$, using multiple sampling vectors. For the experiment, the true but unknown signal $\boldsymbol{x} \in \mathbb{C}^{50}$ and kernel $\boldsymbol{a} \in \mathbb{C}^{50}$ have been randomly generated such that $\boldsymbol{x}$ has a non-vanishing Fourier transform and $\boldsymbol{a}$ is absolutely collision-free, see Figure 5. Further, we generate 47 sampling vectors $\boldsymbol{\phi}_i \in \mathbb{C}^{50}$ such that $\operatorname{supp} \hat{\boldsymbol{\phi}}_i = \{i, \dots, i+3\}$. Since the support of two consecutive sampling vectors is shifted by one, the generated sampling vectors allow index separation (8) and phase propagation (9). Additionally, we ensure that the winding direction

[1]The Julia Programming Language – Version 1.4.2 (`https://docs.julialang.org`)

	Number of samples L					
K	$2K+1$	$3K+1$	$4K+1$	$5K+1$	$8K+1$	$10K+1$
5	$7.380 \cdot 10^{-12}$	$1.317 \cdot 10^{-12}$	$1.446 \cdot 10^{-12}$	$9.931 \cdot 10^{-13}$	$2.162 \cdot 10^{-13}$	$4.886 \cdot 10^{-13}$
10	$1.212 \cdot 10^{-7}$	$1.822 \cdot 10^{-7}$	$4.340 \cdot 10^{-8}$	$1.526 \cdot 10^{-8}$	$4.081 \cdot 10^{-9}$	$5.598 \cdot 10^{-9}$
15	$1.286 \cdot 10^{-3}$	$2.475 \cdot 10^{-4}$	$1.646 \cdot 10^{-6}$	$3.558 \cdot 10^{-5}$	$2.163 \cdot 10^{-6}$	$2.460 \cdot 10^{-6}$
20	$1.503 \cdot 10^{-2}$	$5.406 \cdot 10^{-4}$	$7.727 \cdot 10^{-4}$	$3.951 \cdot 10^{-4}$	$2.998 \cdot 10^{-4}$	$2.325 \cdot 10^{-4}$

Table 1: The mean of the reconstruction error $\|\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}}\|_\infty$ over 5 000 experiments for different numbers of addends K and samples L in the noise-free setting, see Example 10.1.

	Number of samples L					
K	$2K+1$	$3K+1$	$4K+1$	$5K+1$	$8K+1$	$10K+1$
5	$1.298 \cdot 10^{-10}$	$5.154 \cdot 10^{-11}$	$4.864 \cdot 10^{-11}$	$3.644 \cdot 10^{-11}$	$6.745 \cdot 10^{-12}$	$1.988 \cdot 10^{-11}$
10	$3.517 \cdot 10^{-6}$	$6.538 \cdot 10^{-6}$	$6.285 \cdot 10^{-6}$	$5.281 \cdot 10^{-7}$	$1.483 \cdot 10^{-7}$	$2.763 \cdot 10^{-7}$
15	$2.194 \cdot 10^{-3}$	$1.403 \cdot 10^{-4}$	$1.193 \cdot 10^{-4}$	$2.193 \cdot 10^{-4}$	$6.814 \cdot 10^{-5}$	$6.406 \cdot 10^{-5}$
20	$1.860 \cdot 10^{-2}$	$2.040 \cdot 10^{-3}$	$2.445 \cdot 10^{-3}$	$1.503 \cdot 10^{-3}$	$2.021 \cdot 10^{-3}$	$1.405 \cdot 10^{-3}$

Table 2: The mean of the reconstruction error $\|\boldsymbol{\eta}-\tilde{\boldsymbol{\eta}}\|_\infty$ over 5 000 experiments for different numbers of addends K and samples L in the noise-free setting, see Example 10.1.

	Number of samples L					
K	$2K+1$	$3K+1$	$4K+1$	$5K+1$	$8K+1$	$10K+1$
5	$1.481 \cdot 10^{-5}$	$7.640 \cdot 10^{-6}$	$5.188 \cdot 10^{-7}$	$1.942 \cdot 10^{-7}$	$3.555 \cdot 10^{-7}$	$3.335 \cdot 10^{-7}$
10	$1.580 \cdot 10^{-2}$	$4.646 \cdot 10^{-3}$	$3.571 \cdot 10^{-3}$	$3.210 \cdot 10^{-3}$	$3.442 \cdot 10^{-3}$	$3.413 \cdot 10^{-3}$
15	$9.528 \cdot 10^{-2}$	$2.016 \cdot 10^{-2}$	$1.719 \cdot 10^{-2}$	$1.570 \cdot 10^{-2}$	$1.685 \cdot 10^{-2}$	$1.290 \cdot 10^{-2}$
20	$2.741 \cdot 10^{-1}$	$9.357 \cdot 10^{-2}$	$8.451 \cdot 10^{-2}$	$7.909 \cdot 10^{-2}$	$8.243 \cdot 10^{-2}$	$8.477 \cdot 10^{-2}$

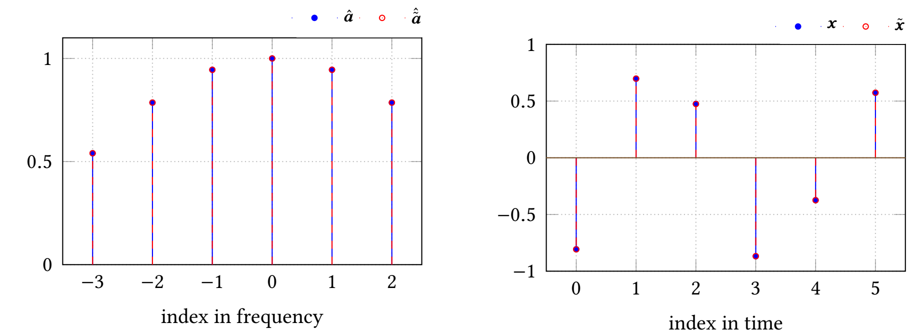
Table 3: The mean of the reconstruction error $\|\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}}\|_\infty$ over 5 000 experiments for different numbers of addends K and samples L in the noisy setting $|e_k| \sim U([0, 10^{-10}])$ and $\arg(e_k) \sim \mathcal{U}((-\pi, \pi])$, see Example 10.1.

	Number of samples L					
K	$2K+1$	$3K+1$	$4K+1$	$5K+1$	$8K+1$	$10K+1$
5	$2.680 \cdot 10^{-4}$	$2.120 \cdot 10^{-4}$	$1.215 \cdot 10^{-5}$	$3.500 \cdot 10^{-6}$	$1.419 \cdot 10^{-5}$	$6.576 \cdot 10^{-6}$
10	$1.580 \cdot 10^{-2}$	$4.646 \cdot 10^{-3}$	$3.571 \cdot 10^{-3}$	$3.210 \cdot 10^{-3}$	$3.442 \cdot 10^{-3}$	$3.413 \cdot 10^{-3}$
15	$9.304 \cdot 10^{-2}$	$3.209 \cdot 10^{-2}$	$3.271 \cdot 10^{-2}$	$2.968 \cdot 10^{-2}$	$3.093 \cdot 10^{-2}$	$2.804 \cdot 10^{-2}$
20	$2.256 \cdot 10^{-1}$	$1.224 \cdot 10^{-1}$	$1.178 \cdot 10^{-1}$	$1.135 \cdot 10^{-1}$	$1.184 \cdot 10^{-1}$	$1.230 \cdot 10^{-1}$

Table 4: The mean of the reconstruction error $\|\boldsymbol{\eta} - \tilde{\boldsymbol{\eta}}\|_\infty$ over 5 000 experiments for different numbers of addends K and samples L in the noisy setting $|e_k| \sim U([0, 10^{-10}])$ and $\arg(e_k) \sim \mathcal{U}((-\pi, \pi])$, see Example 10.1.

	Index k in time domain					
	0	1	2	3	4	5
x	-0.806 494 570 188	0.697 047 937 358	0.475 340 169 748	-0.868 496 176 947	-0.373 776 219 367	0.573 125 494 692
ϕ_1	0.299 100 737 288	-0.067 652 854 127	0.223 548 074 051	-0.419 039 372 471	0.398 336 559 020	0.439 827 094 742
ϕ_2	-0.222 947 251 005	0.185 111 331 800	0.508 076 580 285	-0.024 006 689 074	0.491 191 477 978	-0.360 304 943 116

Table 5: The randomly generated unknown signal $\boldsymbol{x}$ and the known measurement vectors $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2$ in Example 10.2 satisfying the assumptions of Theorem 6.3.



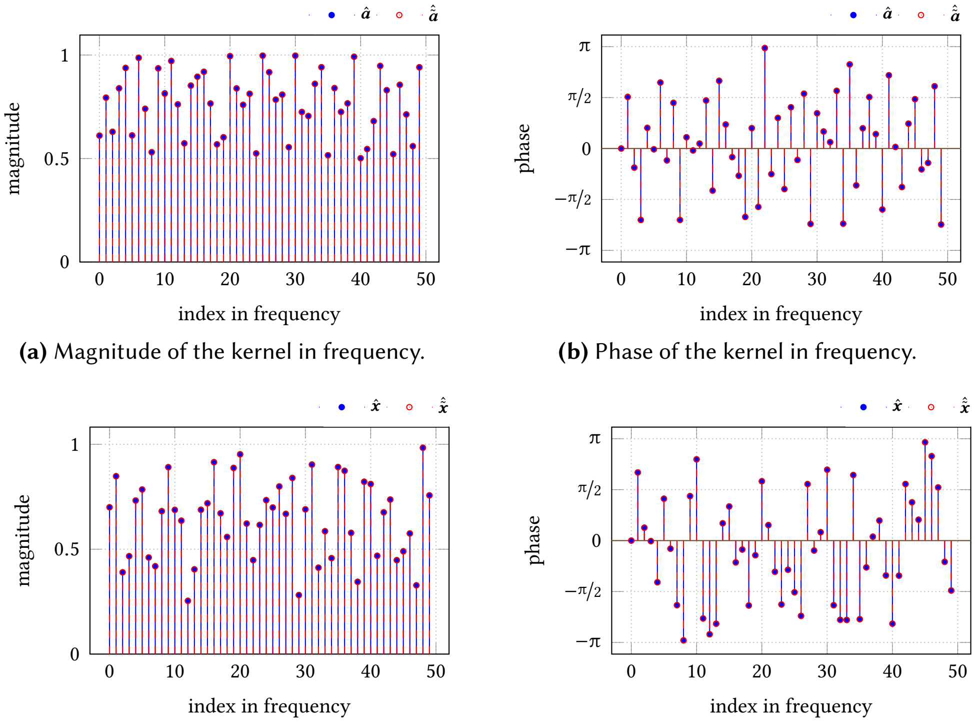
(a) Convolution kernel $\hat{\boldsymbol{a}}$. **(b)** Unknown signal $\boldsymbol{x}$.

Figure 4: The true and reconstructed signal and kernel in Example 10.2 by applying the procedure provided in Theorem 6.3.

determination property (10) is satisfied for $i_1 = 0, i_1 = 1, k_1 = 1, k_2 = 2$. Further, we employ for each sampling vector 65 samples, which is around twice the minimal required number to apply Prony's method. Next, we apply the construction behind the proof of Theorem 7.1 line by line, where the procedure in the proof of Theorem 6.1 is used to identify the partial spectrum of $\boldsymbol{a}$ with respect to $\boldsymbol{\phi}_i$. The recovered signal $\tilde{\boldsymbol{x}}$ and kernel $\tilde{\boldsymbol{a}}$ are shown in Figure 5. Aligning the phase of the true and recovered vectors at the first component, we here observe the reconstruction errors $\|\hat{\boldsymbol{a}} - \hat{\tilde{\boldsymbol{a}}}\|_\infty = 1.897 \cdot 10^{-3}$ and $\|\hat{\boldsymbol{x}} - \hat{\tilde{\boldsymbol{x}}}\|_\infty = 1.563 \cdot 10^{-4}$. As shown in this example, the techniques behind the theoretical proofs may be applied to recover signal and kernel from noise-free samples. ○

11 Conclusion

Phase retrieval in dynamical sampling is a novel research direction occurring a few years ago. As for most phase retrieval problems, the main issue is the ill-posedness especially emerging in the non-uniqueness of the solution. Besides the phase retrieval of the unknown signal, we additionally identify the unknown involved operator from a certain operator class. We have shown that both – phase retrieval and system identification – is in principle simultaneously possible if the spectrum of the operator is (absolutely) collision-free. The employed conditions to ensure the uniqueness of the combined phase and system identification hold for almost all signals, spectra, and measurement vectors. Our work horse has been the approximate Prony method for complex exponential sums. As a consequence, all proofs are constructive and give explicit analytic reconstruction methods. Unfortunately, Prony's method is notorious for its instability. We have studied the sensitivity in more details yielding error bounds that are interesting by themselves outside the context of dynamical sampling. The recovery error of phase and system here centrally depends on the well-separation of the pairwise products of the spectrum and



(a) Magnitude of the kernel in frequency.
(b) Phase of the kernel in frequency.
(c) Magnitude of the signal in frequency.
(d) Phase of the signal in frequency.

Figure 5: The true and reconstructed signal and kernel in Example 10.3 by applying the procedure behind Theorem 7.1.

how far the involved entities are away from zero. Especially for high-dimensional instances the well-separation gets worse and worse since the pairwise products start to cluster; so the analytic reconstructions can only be applied to small instances or a series of specially constructed sampling vectors numerically. The main contributions of this paper are the theoretical uniqueness guarantees, where the question of a practical recovery methods remains open for further research. In particular for phase retrieval, it would be interesting to adapt Prony's method to the occurring quadratic structure or to replace it by a more suitable method.

References

[1] R. Aceska and Y. H. Kim. Scalability of frames generated by dynamical operators. *Front Appl Math Stat*, 3:22, 2017.

[2] R. Alaifari, I. Daubechies, P. Grohs, and R. Yin. Stable phase retrieval in infinite dimensions. *Found Comput Math*, 19(4):869–900, 2019.

[3] R. Alaifari and M. Wellershoff. Stability estimates for phase retrieval from discrete Gabor measurements. *J Fourier Anal Appl*, 27(2):1–31, 2021.

[4] A. Aldroubi, C. Cabrelli, U. Molter, and S. Tang. Dynamical sampling. *Appl Comput Harmon Anal*, 42(3):378–401, 2017.

[5] A. Aldroubi, K. Gröchenig, L. Huang, P. Jaming, I. Krishtal, and J. L. Romero. Sampling the flow of a bandlimited function. *J Geom Anal*, pages 1–35, 2021.

[6] A. Aldroubi, L. Huang, and A. Petrosyan. Frames induced by the action of continuous powers of an operator. *J Math Anal Appl*, 478(2):1059–1084, 2019.

[7] A. Aldroubi and I. Krishtal. Krylov subspace methods in dynamical sampling. *Sampl Theory Signal Image Process*, 15:9–20, 2016.

[8] A. Aldroubi, I. Krishtal, and S. Tang. Phase retrieval of evolving signals from space-time samples. In *Proceedings of the SampTA 2017*, pages 46–49, 2017.

[9] A. Aldroubi, I. Krishtal, and S. Tang. Phaseless reconstruction from space–time samples. *Appl Comput Harmon Anal*, 48(1):395–414, 2020.

[10] A. Aldroubi and A. Petrosyan. Dynamical sampling and systems from iterative actions of operators. In *Frames and Other Bases in Abstract and Function Spaces*, chapter 2, pages 15–26. Birkhäuser, Cham, 2017.

[11] R. Beinert and M. Hasannasab. Phase retrieval via polarization in dynamical sampling. arXiv:2102.00536, 2021.

[12] R. Beinert and G. Plonka. Ambiguities in one-dimensional discrete phase retrieval from Fourier magnitudes. *J Fourier Anal Appl*, 21(6):169–1198, 2015.

[13] R. Beinert and G. Plonka. One-dimensional discrete-time phase retrieval. In *Nanoscale Photonic Imaging*, Nanoscale Photonic Imaging, chapter 24, pages 603–627. Springer, Cham, 2020.

[14] T. Bendory, R. Beinert, and Y. C. Eldar. Fourier phase retrieval: uniqueness and algorithms. In *Compressed Sensing and its Applications*, Applied and Numerical Harmonic Analysis, chapter 2, pages 55–91. Birkhäser, Cham, 2017.

[15] G. Beylkin and L. Monzón. On approximation of functions by exponential sums. *Appl Comput Harmon Anal*, 19(1):17–48, 2005.

[16] R. Bhatia. *Matrix Analysis.* Number 169 in Graduate Texts in Mathematics. Springer, New York, 1997.

[17] Y. M. Bruck and L. G. Sodin. On the ambiguity of the image reconstruction problem. *Opt Commun*, 30(3):304–308, September 1979.

[18] C. Cabrelli, U. Molter, V. Paternostro, and F. Philipp. Dynamical sampling on finite index sets. *J Anal Math*, 140(2):637–667, 2020.

[19] J. A. Cadzow. Signal enhancement-a composite property mapping algorithm. *IEEE Trans Acoust Speech Signal Process*, 36(1):49–62, 1988.

[20] O. Christensen and M. Hasannasab. Frame properties of systems arising via iterated actions of operators. *Appl Comput Harmon Anal*, 46(3):664–673, 2019.

[21] O. Christensen, M. Hasannasab, and F. Philipp. Frame properties of operator orbits. *Math Nachr*, 293(1):52–66, 2020.

[22] J. C. Dainty and J. R. Fienup. Phase retrieval and image reconstruction for astronomy. In *Image Recovery : Theory and Application*, chapter 7, pages 231–275. Academic Press, Orlando (Florida), 1987.

[23] J. R. Deller, J. H. L. Hansen, and J. G. Proakis. *Discrete-time processing of speech signals.* Institute of Electrical and Electronics Engineers, New York, 2000. Originally published: New York : Macmillan, 1993.

[24] A. Eisinberg and C. Picardi. On the inversion of Vandermonde matrix. In *Control science and technology for the progress of society*, volume 14, pages 507–511, Laxenburg, 1981. IFAC. 8th IFAC World Congress on Control Science and Technology for the Progress of Society, Kyoto, Japan, 24-28 August 1981.

[25] M. E. A. El-Mikkawy. Explicit inverse of a generalized Vandermonde matrix. *Appl Math Comput*, 146(2-3):643–651, 2003.

[26] J.-C. Evard and F. Jafari. A complex Rolle's theorem. *Am Math Mon*, 99(9):858–861, 1992.

[27] J. L. Flanagan and R. Golden. Phase vocoder. *Bell System Technical Journal*, 45(9):1493–1509, 1966.

[28] W. Gautschi. On inverses of Vandermonde and confluent Vandermonde matrices. *Numer Math*, 4:117–123, 1962.

[29] W. Gautschi. Norm estimates for inverses of Vandermonde matrices. *Numer Math*, 23:337–347, 1975.

[30] P. Grohs, S. Koppensteiner, and M. Rathmair. The mathematics of phase retrieval. arXiv:1901.07911, 2019.

[31] H. A. Hauptman. The phase problem of x-ray crystallography. *Rep Prog Phys*, 54(11):1427–1454, November 1991.

[32] M. S. Hosseini, A. Chen, and K. N. Plataniotis. On the closed form expression of elementary symmetric polynomials and the inverse of Vandermonde matrix. arXiv:1909.08155, 2019.

[33] Y. Hua and T. K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Trans Acoust Speech Signal Process*, 38(5):814–824, 1990.

[34] Y. Hua and T. K. Sarkar. On svd for estimating generalized eigenvalues of singular matrix pencil in noise. In *IEEE International Sympoisum on Circuits and Systems*, pages 2780–2783. IEEE, 1991.

[35] M. A. Iwen, A. Viswanathan, and Y. Wang. Fast phase retrieval from local correlation measurements. *SIAM J Imaging Sci*, 9(4):1655–1688, 2016.

[36] W. Kim and M. H. Hayes. The phase retrieval problem in x-ray crystallography. In *Proceedings of the ICASSP 91*, volume 3, pages 1765–1768, 1991.

[37] M. V. Klibanov and V. G. Kamburg. Uniqueness of a one-dimensional phase retrieval problem. *Inverse Probl*, 30(7):075004(10), July 2014.

[38] M. V. Klibanov, P. E. Sacks, and A. V. Tikhonravov. The phase retrieval problem. *Inverse Probl*, 11(1):1–28, 1995.

[39] J. Laroche and M. Dolson. Improved phase vocoder time-scale modification of audio. *IEEE Trans Audio Speech Lang Process*, 7(3):323–332, 1999.

[40] C.-K. Li and R. Mathias. The Lidskii-Mirsky-Wielandt theorem – additive and multiplicative versions. *Numer Math*, 81:377–413, 1999.

[41] Y. M. Lu and M. Vetterli. Spatial super-resolution of a diffusion field by temporal oversampling in sensor networks. In *Proceedings of the ICASSP 2009*, pages 2249–2252, 2009.

[42] N. Macon and A. Spitzbart. Inverses of Vandermonde matrices. *Amer Math Monthly*, 65:95–100, 1958.

[43] R. D. Martín, I. Medri, and U. Molter. Continuous and discrete dynamical sampling. *J Math Anal Appl*, page 125060, 2021.

[44] R. P. Millane. Phase retrieval in crystallography and optics. *J Opt Soc Amer A*, 7(3):394–411, 1990.

[45] V. Y. Pan. How bad are Vandermonde matrices? *SIAM J. Matrix Anal. Appl.*, 37(2):676–694, 2016.

[46] J. M. Papy, L. De Lathauwer, and S. Van Huffel. Exponential data fitting using multilinear algebra: the single-channel and multi-channel case. *Numer Linear Algebra Appl*, 12(8):809–826, 2005.

[47] F. Philipp. Bessel orbits of normal operators. *J Math Anal Appl*, 448(2):767–785, 2017.

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

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

[50] R. Prony. Essai expérimental et analytique sur les lois de la dilatabilité des fluides élastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, á différentes températures. *Journal de l'École polytechnique*, 2:24–76, 1795.

[51] J. Ranieri, A. Chebira, Y. M. Lu, and M. Vetterli. Sampling and reconstructing diffusion fields with localized sources. In *Proceedings of the ICASSP 2011*, pages 4016–4019, 2011.

[52] R. Roy, A. Paulraj, and T. Kailath. Esprit–a subspace rotation approach to estimation of parameters of cisoids in noise. *IEEE Trans Acoust Speech Signal Process*, 34(5):1340–1342, 1986.

[53] B. Seifert, H. Stolz, M. Donatelli, D. Langemann, and M. Tasche. Multilevel Gauss-Newton methods for phase retrieval problems. *J Phys A: Math Gen*, 39(16):4191–4206, 2006.

[54] B. Seifert, H. Stolz, and M. Tasche. Nontrivial ambiguities for blind frequency-resolved optical gating and the problem of uniqueness. *J Opt Soc Am B*, 21(5):1089–1097, May 2004.

[55] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev. Phase retrieval with application to optical imaging: A contemporary overview. *IEEE Signal Process Mag*, 32(3):87–109, 2015.

[56] S. Tang. System identification in dynamical sampling. *Adv Comput Math*, 43(3):555–580, 2017.

[57] R. Turner. Inverse of the Vandermonde matrix with applications. NASA Technical Note NASA TN D-3547, National Aeronautics and Space Administration, Washington, D.C., 1966.

[58] A. Ulanovskii and I. Zlotnikov. Reconstruction of bandlimited functions from space–time samples. *J. Funct. Anal.*, page 108962, 2021.

[59] P. van Hove, M. H. Hayes, J. S. Lim, and A. V. Oppenheim. Signal reconstruction from signed Fourier transform magnitude. *IEEE Trans Acoust Speech Signal Process*, ASSP-31(5):1286–1293, October 1983.