Journal of Symbolic Computation 125 (2024) 102313

Contents lists available at ScienceDirect

Journal of Symbolic Computation

journal homepage: www.elsevier.com/locate/jsc

Stabilized recovery and model reduction for multivariate exponential polynomials

Juan Manuel Peña [a], Tomas Sauer [b]

[a] Depto. Matemática Aplicada, Fac. Ciencias, Universidad de Zaragoza, E–50009 Zaragoza, Spain
[b] Lehrstuhl für Mathematik mit Schwerpunkt Digitale Bildverarbeitung & FORWISS, University of Passau, Fraunhofer IIS, Development Center X-Ray Technology, Innstr. 43, D–94032 Passau, Germany

ARTICLE INFO

Article history:
Received 7 October 2022
Received in revised form 5 November 2023
Accepted 27 February 2024
Available online 1 March 2024

MSC:
13P10
65T40
41D05

Keywords:
Prony's method
Gröbner basis
Backwards error
Model reduction

ABSTRACT

Recovery of multivariate exponential polynomials, i.e., the multivariate version of Prony's problem, can be stabilized by using more than the minimally needed multiinteger samples of the function. We present an algorithm that takes into account this extra information and prove a backward error estimate for the algebraic recovery method SMILE. In addition, we give a method to approximate data by an exponential polynomial sequence of a given structure as a step in the direction of multivariate model reduction.

1. Introduction

In this paper, we investigate the recovery of multivariate sparse exponential polynomials from discrete samples and approximation of discrete data by exponential polynomials of a given structure. That is, we consider functions of the form

$$f(x) = \sum_{\omega\in\Omega} f_\omega(x)\, e^{\omega^T x}, \qquad \Omega \subset (\mathbb{R} + i\mathbb{T})^s, \quad \#\Omega < \infty, \quad 0 \neq f_\omega \in \mathbb{C}[x] := \mathbb{C}[x_1, \dots, x_s]. \tag{1}$$

E-mail addresses: jmpena@unizar.es (J.M. Peña), Tomas.Sauer@uni-passau.de (T. Sauer).

https://doi.org/10.1016/j.jsc.2024.102313

The *complexity* N of such a model is the number $\#\Omega$ of *frequencies* in Ω, counted with multiplicities that depend on the respective polynomial f_ω. *Sparsity* of the representation (1) means that all coefficients are nonzero and of minimal degree if they are polynomials, hence the finite cardinality N of Ω is minimal for representing f. Usually it is also assumed that N is of modest size, whatever that means in detail. The task is to reconstruct the *frequencies* $\omega \in \Omega$ and the coefficients $f_\omega \in \mathbb{C}[x] \setminus \{0\}$ from measurements of f on a grid $\Lambda \subset \mathbb{Z}^s$, where typical examples are sets of the form

$$\Lambda = A \pm B = \{\alpha + \beta : \alpha \in A, \beta \in B\}, \qquad A, B \subset \mathbb{Z}^s,$$

and will have to satisfy additional properties. This recovery problem is known as the multivariate version of *Prony's problem*, (Prony, 1795), see (Harmouch and Mourrain, 2018; Mourrain, 2018; Sauer, 2017a,b). The a priory knowledge needed for such reconstructions is usually an upper estimate $N' \geq N$ for the number of frequencies (with multiplicities). Numerical aspects of Prony's problem in one and several variables have been an active field of research recently, ranging from the question of sparse reconstruction and minimal sampling (Cuyt and Lee, 2018; Kaltofen and Yang, 2016) and the conditioning of the underlying Vandermonde matrices (Batenkov et al., 2021) to new methods based on rational approximation (Derevianko and Plonka, 2023); this list is in no way complete. Here, we consider a different problem, namely some backwards error of the reconstruction and the effect of systematic oversampling to it.

Like in the univariate case, there are certain requirements on A and B in order to be able to reconstruct f from the restriction $f(\Lambda)$. Numerical experiments show, on the other hand, that it often helps to *oversample*, i.e., to use a set Λ that is larger than necessary. For example, in the univariate case it is well-known that the minimal number of samples is $2N$, more than $2N$ samples would then be oversampling. Having such extended data available, one can define a numerically stabilized version of the reconstruction algorithm by taking into account the (ideal) structure behind the frequency set Ω, more precisely, behind e^Ω. We describe this method, consider in Theorem 2 the backward error of the recovery and derive a quantitative estimate that also illustrates how oversampling can improve the accuracy of the process.

The second question we consider in this paper is in the direction of *model reduction* and gives a method to compute, for given input data, a function of the form (1) in a prescribed simple structure. In one variable this would mean to just bound the number of frequencies, in several variables however it corresponds to prescribing a certain ideal structure.

A different perspective to the problem is provided by the point of view of filters. Recall that a *filter*, more precisely an *FIR filter* (Hamming, 1989), is an operator on $\ell(\mathbb{Z}^s) := \{c : \mathbb{Z}^s \to \mathbb{C}\}$, defined as

$$Hf := h * f := \sum_{\alpha \in \mathbb{Z}^s} h(\cdot - \alpha)\, f(\alpha), \qquad f \in \ell(\mathbb{Z}^s), \tag{2}$$

with a finitely supported *impulse response* $h \in \ell_0(\mathbb{Z}^s)$, where ℓ_0 is the space of sequences with finite pseudo norm

$$\|h\|_0 := \#\left\{\alpha \in \mathbb{Z}^s : h(\alpha) \neq 0\right\}.$$

Using the concepts of the *shift operator* τ, defined as $\tau^\alpha f := f(\cdot + \alpha)$, and of the *z transform*

$$h^*(z) := \sum_{\alpha \in \mathbb{Z}^s} h(\alpha)\, z^{-\alpha}, \qquad z \in \mathbb{C}_\times^s := (\mathbb{C} \setminus \{0\})^s,$$

we can write the filter as

$$h * f = \sum_{\alpha \in \mathbb{Z}^s} h(\cdot - \alpha)\, f(\alpha) = \sum_{\alpha \in \mathbb{Z}^s} h(\alpha)\, f(\cdot - \alpha) = \sum_{\alpha \in \mathbb{Z}^s} h(\alpha) \tau^{-\alpha} f = h^*(\tau)\, f.$$

Recall that h^* is a *Laurent polynomial* if and only if $h \in \ell_0(\mathbb{Z}^s)$.

Definition 1. By $\Pi = \mathbb{C}[z]$ we denote the ring of all polynomials and use

$$\Pi_A := \operatorname{span}_{\mathbb{C}} \left\{ z^\alpha : \alpha \in A \right\}$$

for all polynomials with *support* in A. Moreover, we denote by $\Lambda = \mathbb{C}\left[z, z^{-1}\right]$ the ring of all *Laurent polynomials* with complex coefficients which are defined on $\mathbb{C}_\times^s$, where, as usual, $\mathbb{C}_\times := \mathbb{C} \setminus \{0\}$.

Considering the bilinear form

$$(\cdot, \cdot) : \ell_0(\mathbb{Z}^s) \times \ell(\mathbb{Z}^s) \to \ell(\mathbb{Z}^s),$$

defined as

$$(h, f) := h * f := h^*(\tau)\, f,$$

and the *homogeneous partial difference equation* $(h, f) = 0$, we find that this implies for fixed h that

$$0 = \tau^\alpha 0 = \tau^\alpha (h * f) = h * (\tau^\alpha f)$$

and for fixed f that

$$0 = \tau^\alpha 0 = \tau^\alpha h^*(\tau)\, f = \left((\cdot)^\alpha h^*\right)(\tau)\, f,$$

so that we can state the following simple but important structural observation.

Lemma 1. *Given $h \in \ell_0(\mathbb{Z}^s)$ the set $\left\{ f \in \ell(\mathbb{Z}^s) : (h, f) = 0 \right\}$ is a shift invariant linear subspace of $\ell(\mathbb{Z}^s)$ while for $f \in \ell(\mathbb{Z}^s)$ the set $\{h^* : (h, f) = 0\}$ is a* Laurent ideal*.*

Remark 1. A *Laurent ideal* is defined as an ideal in the ring of Laurent polynomials, i.e., a subset closed under addition and multiplication with arbitrary monomials with positive and negative exponents. Since, in contrast to polynomials, the monomials are *units* in the ring of Laurent polynomials, the whole ring is spanned as a vector space by units which makes the concept of a graded ring inapplicable. Moreover, Laurent polynomials are not defined for values z where $z_j = 0$ which makes basis computations for the ideals more intricate, cf. Möller and Sauer (2004).

In view of Lemma 1, it is convenient to consider the bilinear form $(\cdot, \cdot)$ as a mapping on $\Lambda \times \ell(\mathbb{Z}^s)$ with values in $\ell(\mathbb{Z}^s)$. We then have that

$$(\cdot, \cdot) : \Lambda \times \ell(\mathbb{Z}^s) \to \ell(\mathbb{Z}^s), \qquad (p, f) \mapsto p(\tau)\, f \tag{3}$$

and know that the homogeneous relations $(\cdot, f) = 0$ and $(p, \cdot) = 0$ define an ideal and a shift invariant subspace, respectively.

Remark 2. It is important to notice that the bilinear form in (3) is *not* an inner product since it is not scalar valued. It can be understood, however, as *partial difference equation* with constant coefficients, hence any f that solves $(p, f) = 0$ must be an exponential polynomial function, cf. Jordan (1965) for the univariate case. The multivariate analogy has been considered for example in Mourrain (2018) as well as in Sauer (2016), taking into account that partial difference equations are convolutions.

Alternatively, one can also consider filters based on the *correlation*

$$\widetilde{H} f := h \star f := \sum_{\alpha \in \mathbb{Z}^s} h(\alpha)\, f(\cdot + \alpha) = (h(-\cdot) * f), \qquad f \in \ell(\mathbb{Z}^s), \tag{4}$$

in which case we have that

$$h \star f = h(\tau)\, f, \qquad h(z) := h^*(z^{-1}) = \sum_{\alpha \in \mathbb{Z}^s} h(\alpha)\, z^\alpha, \qquad z \in \mathbb{C}^s_\times,$$

in terms of the *symbol* of h.

The paper is organized as follows. In Section 2 we recall some facts about Hankel and Toeplitz operators of finite rank which will be the mail tools for our investigations; the equivalence between finite dimensional shift invariant sequence spaces, Hankel or Toeplitz operators of finite rank and exponential polynomials revised there will turn out to be useful. Section 3 introduces the stabilized Prony method and gives a backward error estimate for it, shedding some light on the role of oversampling. Model reduction and algorithm for that purpose based on alternating minimization are finally stated and investigated in Section 4.

2. Finite rank operators

The convolution of $c \in \ell(\mathbb{Z}^s)$ with any $f \in \ell(\mathbb{Z}^s)$, regardless of its support, can be written as

$$f * c = \sum_{\alpha \in \mathbb{Z}^s} f(\cdot - \alpha) c(\alpha) = Fc,$$

where F is the biinfinite matrix with entries $F_{\alpha,\beta} := f(\alpha - \beta)$, $\alpha, \beta \in \mathbb{Z}^s$, called a *Toeplitz operator*. This can be formalized more generally as follows.

Definition 2. Given $A, B \subset \mathbb{Z}^s$, not necessarily finite, and $f \in \ell(\mathbb{Z}^s)$, the associated *Hankel operator* is defined as

$$H_{A,B}(f) := \left[f(\alpha + \beta) : \begin{matrix} \alpha \in A \\ \beta \in B \end{matrix} \right] \in \mathbb{C}^{A \times B} \tag{5}$$

and the associated *Toeplitz operator* as

$$T_{A,B}(f) := \left[f(\alpha - \beta) : \begin{matrix} \alpha \in A \\ \beta \in B \end{matrix} \right] \in \mathbb{C}^{A \times B}, \tag{6}$$

respectively. If A and B are finite, we speak of (generalized) Hankel and Toeplitz matrices, if $A = B = \mathbb{N}_0^s$, we simply write $H(f) := H_{\mathbb{N}_0^s, \mathbb{N}_0^s}(f)$ and for $A = B = \mathbb{Z}^s$ we use $T(f) := T_{\mathbb{Z}^s, \mathbb{Z}^s}(f)$.

Remark 3. Even when A, B are finite sets, arranging the above matrices from Definition 2 into a matrix in $\mathbb{C}^{\#A \times \#B}$ by means of some ordering of A and B clearly does **not** lead to a matrix with Hankel or Toeplitz structure. This is the reason why (5) and (6) are often called *generalized* Hankel/Toeplitz matrices. However, the multivariate structure becomes much clearer and more evident in the above form.

For any $c \in \mathbb{C}^B$ and $\alpha \in A$ we have that

$$\big(H_{A,B}(f)\, c\big)(\alpha) = \sum_{\beta \in B} c(\beta)\, f(\alpha + \beta) = \sum_{\beta \in \mathbb{N}_0^s} c(\beta)\, f(\alpha + \beta) = (c \star f)(\alpha), \qquad \alpha \in A, \tag{7}$$

and similarly

$$\big(T_{A,B}(f)\, c\big)(\alpha) = (c * f)(\alpha), \qquad \alpha \in A. \tag{8}$$

Here and in the rest of this paper, we identify an finitely supported c with its canonical embedding into $\ell(\mathbb{Z}^s)$ by means of zero padding, setting $c(\alpha) = 0$, $\alpha \in \mathbb{Z}^s \setminus B$. It is easy to see that whenever f is a simple *multi-exponential function*. i.e., of the form (1) with $0 \neq f_\omega \in \mathbb{C}$ instead of $f_\omega \in \Pi$, the above matrices factorize as

$$H_{A,B}(f) = V(e^\Omega, A)^T\, F_\Omega\, V(e^\Omega, B), \qquad T_{A,B}(f) = V(e^\Omega, A)^T\, F_\Omega\, V(e^{-\Omega}, B), \tag{9}$$

where the diagonal matrix $F_\Omega = \text{diag}\,[f_\omega : \omega \in \Omega]$ has rank $N = \#\Omega$ and where

$$V(X,\Gamma) := \left[x^\gamma : \begin{array}{c} x \in X \\ \gamma \in \Gamma \end{array}\right], \qquad \Gamma \subseteq \mathbb{Z}^s, \tag{10}$$

denotes the *Vandermonde matrix* for the space Π_Γ, with respect to $X \subset \mathbb{C}^s_\times$.

In the case that $f_\omega \in \Pi$, a factorization similar to (9) holds, cf. (Mourrain, 2018; Sauer, 2017b). This factorization, however, features a *block diagonal* matrix F_Ω and Vandermonde matrices with respect to a certain Hermite interpolation problem, where not only function values but also derivatives have to be interpolated. To keep the presentation simple here, we will avoid the intricacies and focus on simple multi-exponentials.

Remark 4. Negative powers in the Vandermonde matrix (10) are permitted in this approach since $e^\omega \in \mathbb{C}^s_\times$, $\omega \in \Omega$. Moreover, any finite $\Gamma \subset \mathbb{Z}^s$ can always be written as $\Gamma = \gamma^0 + \Gamma'$ with $\gamma^0 \in \mathbb{Z}^s$ and $\Gamma' \subset \mathbb{N}^s_0$, so that

$$x^\gamma = x^{\gamma^0} x^{\gamma-\gamma^0} =: x^{\gamma^0} x^{\gamma'}, \qquad \gamma' \in \Gamma',$$

hence

$$V(X,\Gamma) = \text{diag}\left[x^{\gamma^0} : x \in X\right] V(X,\Gamma') =: D(X,\gamma^0)\, V(X,\Gamma')$$

and the diagonal matrix is nonsingular as long as $X \subset \mathbb{C}^s_\times$. This implies the modified factorization

$$H_{A,B}(f) = V(e^\Omega, A')^T\, D(e^\Omega,\alpha^0)\, F_\Omega\, D(e^\Omega,\beta^0)\, V(e^\Omega, B'), \tag{11}$$

where now the Vandermonde matrices are *polynomial* ones and the diagonal matrix

$$F'_\Omega := D(e^\Omega,\alpha^0)\, F_\Omega\, D(e^\Omega,\beta^0)$$

still is nonsingular. In the case of the Toeplitz matrix $T_{A,B}(f)$ we get a similar expansion, just with a complex conjugate:

$$F'_\Omega = \overline{D(e^\Omega,\alpha^0)}\, F_\Omega\, D(e^\Omega,\beta^0).$$

The following observation will become a fundamental tool later.

Proposition 2. *If f is an exponential polynomial of the form (1) and A is any finite subset of $\mathbb{N}^s_0$ such that* $\text{rank}\, V(e^\Omega, A) = N = \#\Omega$, *then* $\ker H_{A,B}(f)$ *or* $\ker T_{A,B}(f)$ *consists of all filters h with impulse response supported on B such that*

$$h \star f = 0 \qquad \text{or} \qquad h * f = 0, \tag{12}$$

respectively. In other words, the kernels give all annihilators supported on B.

Proof. We prove the result for the Hankel matrix $H_{A,B}(f)$, the proof for the Toeplitz matrix $T_{A,B}(f)$ is identical. If h is supported on B, then $h \in \ker H_{A,B}(f)$ is equivalent to $(h \star f)(\alpha) = 0$ for any $\alpha \in A$. Since $\text{rank}\, V(e^\Omega, A) = N$, there exists a matrix $M \in \mathbb{C}^{A\times\Omega}$, for example the pseudo-inverse of $V(e^\Omega, A)$, such that $MV(e^\Omega, A) = I \in \mathbb{C}^{A\times A}$. Now we get for any $A' \subset \mathbb{Z}^s$ that

$$\begin{aligned}(h \star f)|_{A'} &= H_{A',B}(f)\, h = V(e^\Omega, A')^T\, F_\Omega\, V(e^\Omega, B)\, h \\ &= V(e^\Omega, A')^T\, M\, V(e^\Omega, A)\, F_\Omega\, V(e^\Omega, B)\, h = V(e^\Omega, A')^T\, M\, H_{A,B}(f)\, h = 0,\end{aligned}$$

hence $h \star f$ vanishes on arbitrary subsets of $\mathbb{Z}^s$. □

In several variables, the question of determining A such that $\operatorname{rank} V(e^\Omega, A) = N$ is more intricate than in one variable since now also the geometry of Ω matters. A choice for A proposed in (Sauer, 2018) is the positive part of the *hyperbolic cross*,

$$\Upsilon_N := \left\{ \alpha \in \mathbb{N}_0^s : \prod_{j=1}^{s} (\alpha_j + 1) \le N \right\}. \tag{13}$$

It can be shown that $\operatorname{rank} V(X, \Upsilon_N) = N$ for all $X \subset \mathbb{C}^s$ with $\#X \le N$ and that Υ_N is the union of all *lower sets* of cardinality at most N. Recall that a lower set of multiindices is defined by the property that it contains, together with some α, all multiindices that are bounded by α componentwise, see Definition 5 for a formal definition.

Now we consider the rank of Hankel and Toeplitz operators. To that end, we start with $f \in \ell(\mathbb{Z}^s)$ being an arbitrary sequence. Since increasing A or B means adding rows or columns, it readily follows that

$$\operatorname{rank} H_{A,B}(f) \le \operatorname{rank} H_{A',B'}(f), \qquad A \subseteq A', \quad B \subseteq B', \qquad \#A', \#B' \le \infty, \tag{14}$$

with the same holding true for Toeplitz operators as well. Since small Hankel matrices, i.e., matrices associated to small sets A, B, may not cover the full information on f, we make the following definition.

Definition 3. The rank of the Hankel and Toeplitz operators are defined as the limits

$$\operatorname{rank} H(f) := \sup_{A,B \subset \mathbb{N}_0^s} \operatorname{rank} H_{A,B}(f), \qquad \operatorname{rank} T(f) := \sup_{A,B \subset \mathbb{Z}^s} \operatorname{rank} T_{A,B}(f), \tag{15}$$

respectively. Moreover, we use the abbreviations

$$r_H(f) := \operatorname{rank} H(f), \qquad r_T(f) := \operatorname{rank} T(f),$$

for the (nonnegative) *Hankel rank* and *Toeplitz rank* of a given sequence.

In the remainder of this section, we recall a few well-known properties of Hankel operators of finite rank, cf. (Harmouch and Mourrain, 2018).

Proposition 3. *For any $f \in \ell(\mathbb{Z}^s)$ we have*

1. $r_H(f) = r_T(f)$,
2. $r_H(f) < \infty$ *if and only if f is an exponential polynomial, i.e.,*

$$f = \sum_{\omega \in \Omega} f_\omega(\cdot)\, e^{\omega^T \cdot}, \qquad f_\omega \in \Pi, \tag{16}$$

where Ω is a finite set of frequencies.

Since $T_{A,B}(f) = H_{A,-B}(f)$, $A, B \in \mathbb{Z}^s$, and since also

$$\operatorname{rank} H_{A,B}(f) = \sup_{A,B \in \mathbb{Z}^s} H_{A,B}(f),$$

the two suprema in (15) coincide and statement 1) of Proposition 3 is trivially true.

Definition 4. The shift invariant space generated by $f \in \ell(\mathbb{Z}^s)$ will be denoted by

$$\mathcal{S}(f) := \operatorname{span} \left\{ \tau^\alpha f : \alpha \in \mathbb{Z}^s \right\}. \tag{17}$$

Lemma 4. $r_H(f) = \dim \mathcal{S}(f)$.

If $r_H(f) < \infty$ then there exist finitely supported sequences $c \in \ell(\mathbb{N}_0^s)$ such that $H(f)c = 0$ and the set of polynomials

$$\left\{ \sum_{\alpha \in \mathbb{N}_0^s} c(\alpha) z^\alpha : H(f)c = 0,\ \|c\|_0 < \infty \right\}$$

forms a zero dimensional ideal $\mathcal{I}$, the so-called *Prony ideal* associated to f from (16), cf. Harmouch and Mourrain (2018); Mourrain (2018). The common zeros of $\mathcal{I}$, counted with proper multiplicity, are the $e^\omega \in \mathbb{C}_\times^s$. Concrete constructions for H-bases and Gröbner bases of this ideal can be found in Sauer (2017a, 2018). It has been shown in Sauer (2018) that, for $N = \dim \Pi/\mathcal{I}$, all information on $H(f)$ is already obtained by choosing $A = B = \Upsilon_N$. Hence, we have the following result.

Theorem 1. *If $r_H(f) < \infty$ then there exists some $N > 0$ such that*

$$r_H(f) = \operatorname{rank} H_{\Upsilon_N, \Upsilon_N}(f). \tag{18}$$

3. Stabilized computation of frequencies

The main effort in solving Prony's problem is made in reconstructing the frequencies Ω. Once they are found, determining the coefficients is only a linear problem that leads immediately to a linear system which is easy to solve, at least in principle (Sauer, 2017a).

To compute the frequencies, one starts with a sufficiently rich set $A \subset \mathbb{N}_0^s$, for example $A = \Upsilon_N$ for some $N \geq \#\Omega$, and builds the Hankel matrices $H_{A,B_j}(f)$ based on a nested sequence $B_0 \subset B_1 \subset \cdots \subset \mathbb{N}_0^s$. Then one successively determines the kernels $\ker H_{A,B_j}(f)$ until the ranks of the sequence $H_{A,B_j}(f)$ stabilize. The simplest way is to choose the B_j in such a way that a subsequence satisfies $B_{j(n)} = \{\alpha : |\alpha| = n\}$ since $\operatorname{rank} H_{A,B_{j(n+1)}}(f) = \operatorname{rank} H_{A,B_{j(n)}}(f)$ implies that $r_H(f) = \operatorname{rank} H_{A,B_{j(n)}}(f)$. This is the concept behind the SMILE (Sauer, 2018) and DNSIN (Sauer, 2017a) algorithms that can be seen as multivariate extensions of the well–known univariate Prony methods ESPRIT and MUSIC, (Roy and Kailath, 1989; Schmidt, 1986).

Definition 5. A subset $A \subset \mathbb{N}_0^s$ is called a *lower set* or *order closed ideal* if

$$\left\{\beta \in \mathbb{N}_0^s : \beta \leq \alpha\right\} \subseteq A, \qquad \alpha \in A,$$

where $\beta \leq \alpha$ iff $\alpha - \beta \in \mathbb{N}_0^s$. For $A \subset \mathbb{N}_0^s$ we denote by

$$L(A) = \bigcap_{\alpha \in A} \left(\alpha + \mathbb{N}_0^s\right)$$

the lower set generated by the (finite or infinite) set A of multiindices.

In the remainder of this section, we always assume the following setting: the available measurements of f are taken on $\Lambda := A + B \subset \mathbb{Z}^s$ where

1. $0 \in A$ and A is an *interpolation set* for e^Ω, i.e., the interpolation problem $p(e^\omega) = y_\omega$, $\omega \in \Omega$, has a solution in Π_A for any $y \in C^\Omega$ which, however, need not be unique; we can choose for example $A = \Upsilon_N$, $N \geq \#\Omega$,
2. $B \subset \mathbb{N}_0^s$ is a *lower set*, hence also $0 \in B$.

Note that $0 \in A \cap B$ implies that $A, B \subset \Lambda$. Under the above assumption on A, the SMILE algorithm determines a *minimal* lower set $B^* \subset B$ such that $\operatorname{rank} H_{A,B^*}(f) = \operatorname{rank} H(f)$ by means of computing a Gröbner basis for the associated Prony ideal.

Our additional assumption is that B is larger than B^* and thus provides some *oversampling* of f that can and will be used for stabilized version of the SMILE algorithm from (Sauer, 2018).

To that end, let us revisit the algorithm and mark the necessary changes. We order the elements of B as $\beta^0 \prec \beta^1 \prec \cdots$ with respect to the well–known *graded lexicographic* order "$\prec$", so that $\alpha \prec \beta$ if either $|\alpha| < |\beta|$ or $|\alpha| = |\beta|$, and there exists $1 \le j \le s$ such that $\alpha_k = \beta_k$, $k < j$, and $\alpha_j < \beta_j$. Then

$$B_j := \left\{\beta^0, \ldots, \beta^j\right\}, \qquad j = 0, \ldots, \#B - 1,$$

yields a nested sequence of subsets $B_0 \subset B_1 \subset \cdots \subseteq B$. We can rewrite the condition $\operatorname{rank} H_{A,B_j}(f) = \operatorname{rank} H_{A,B_{j+1}}(f)$ as the *rank deficiency condition* that there exists some vector $h \in \mathbb{C}^{B_{j+1}}$ such that

$$0 = H_{A,B_{j+1}}(f)\, h = H_{A,B_j}(f)\, h' + H_{A,\beta^{j+1}}(f), \qquad h' \in \mathbb{C}^{B_j}, \quad h = \begin{bmatrix} h' \\ 1 \end{bmatrix} \in \mathbb{C}^{B_{j+1}}. \tag{19}$$

The vector h corresponds to the polynomial

$$h(z) := \sum_{\alpha \in B_{j+1}} h(\alpha)\, z^\alpha = z^{\beta^{j+1}} + \sum_{\alpha \in B_j} h(\alpha)\, z^\alpha,$$

which is *monic* with respect to the term order "$\prec$" since its *leading term* $(\cdot)^{\beta^j}$ is a monomial, cf. (Cox et al., 1996).

Remark 5. To avoid further notation, we use the same symbol h for the polynomial and its coefficient vector since they describe the same object and the role of the symbol should be clear from the context.

The *leading terms* of the polynomials $(\cdot)^\alpha\, h$ generating the intersection of the ideal generated by h with Π_B, are the monomials $x^{\alpha+\beta^{j+1}}$, $\alpha \in (B - B_{j+1})_+$. Here, the nonnegative difference $(\Gamma - \Gamma')_+$ of two sets $\Gamma, \Gamma \subset \mathbb{N}_0^s$ is defined as

$$(\Gamma - \Gamma')_+ := (\Gamma - \Gamma') \cap \mathbb{N}_0^s := \left\{\gamma - \gamma' : \gamma \in \Gamma, \gamma' \in \Gamma'\right\} \cap \mathbb{N}_0^s,$$

i.e., the restriction of the difference to nonnegative integer vectors. Observe, however, that the case $(B - B_{j+1})_+ = \{0\}$ is not excluded from the considerations. This would reduce to the standard case of the SMILE algorithm.

The additional conditions due to the fact that the kernel elements of the Hankel matrix, interpreted as polynomials, form an ideal, are $H_{A,B}(f) h_\alpha = 0$, $h_\alpha := (\cdot)^\alpha\, h$, $\alpha \in (B - B_j)_+$, i.e.,

$$\begin{aligned} 0 = H_{A,B_{j+1}+\alpha}(f)\, h_\alpha &= \sum_{\beta \in B_{j+1}+\alpha} f(A+\beta) h_\alpha(\beta) = \sum_{\beta \in B_{j+1}+\alpha} f(A+\beta) h(\beta - \alpha) \\ &= \sum_{\beta \in B_{j+1}} f(A+\alpha+\beta)\, h(\beta) = H_{A+\alpha,B_{j+1}}(f)\, h, \end{aligned}$$

and lead to the algebraic *consistency conditions*

$$0 = H_{\alpha+A,B_j}(f)\, h' + H_{\alpha+A,\beta^{j+1}}(f), \qquad \alpha \in (B - B_{j+1})_+. \tag{20}$$

Remark 6. Note that (20) makes use of the fact that we computed a Gröbner basis for the Prony ideal $\mathcal{I}$ in the SMILE algorithm, since then (20) means that the solution considers $\mathcal{I} \cap \Pi_B$.

Thus, h is a member of the *Prony ideal* of all polynomials in the kernel of the infinite Hankel operator, cf. (Sauer, 2017a, 2018) if and only if it is the solution to the optimization problem

$$\min_{y\in\mathbb{C}^{B_j}} \Phi_{\beta^{j+1},f}(y), \qquad \Phi_{\beta^{j+1},f}(y) := \sum_{\alpha\in(B-B_{j+1})_+} \left\| H_{\alpha+A,B_j}(f)\, y + H_{\alpha+A,\beta^{j+1}}(f) \right\|_2^2, \tag{21}$$

that assumes the globally minimal value $\Phi_{\beta^{j+1},f}(h') = 0$.

Remark 7. The optimization problem in (21) tries to find a shortest vector y or a simple basis of the ideal that is, however, as consistent as possible over all the available data. This approach is different from the Prony problem mentioned in Osborne and Smyth (1991), see also Osborne and Smyth (1995); Zhang and Plonka (2019). There the norm $\|H_{A,B\setminus\{\beta\}}(f)y + H_{A,\beta}(f)\|_2$ with a y of maximal length is minimized, while our approach considers a shorter y and all its shift simultaneously.

Remark 8. The optimization problem in (20) is quite easy to solve. Since

$$\begin{aligned}\Phi_{\beta^{j+1},f}(y) = \sum_{\alpha\in(B-B_{j+1})_+} \Big\{ & y^H H_{\alpha+A,B_j}(f)^H H_{\alpha+A,B_j}(f) y \\ & + 2\mathrm{Re}\left(y^H H_{\alpha+A,B_j}(f)^H H_{\alpha+A,\beta^{j+1}}(f) \right) \\ & + H_{\alpha+A,\beta^{j+1}}(f)^H H_{\alpha+A,\beta^{j+1}}(f) \Big\},\end{aligned}$$

the minimizer is simply the solution of the normal equations

$$\left(\sum_{\alpha\in(B-B_{j+1})_+} H_{\alpha+A,B_j}(f)^H H_{\alpha+A,B_j}(f) \right) y = - \sum_{\alpha\in B-B_{j+1}} H_{\alpha+A,B_j}(f)^H H_{\alpha+A,\beta^{j+1}}(f). \tag{22}$$

Note that $H_{\alpha+A,B_j}(f)$ depends only on the samples of f on

$$\left\{ \alpha + A + B_j : \alpha \in (B - B_{j+1})_+ \right\} \subseteq A + B = \Lambda,$$

the latter of which is the input data for the algorithm.

The minimization (21) is the basis of the numerically stabilized variant of the SMILE algorithm (Sauer, 2018) and decides the ideal containment of some monic polynomial $h = (\cdot)^{\beta^{j+1}} + h'$ in the following way: instead of only checking whether $H_{A,B_j}(f)\, h' + H_{A,\beta_{j+1}}(f) = 0$ or its numerical counterpiece

$$\left\| H_{A,B_j}(f)\, h' + H_{A,\beta_{j+1}}(f) \right\|_2 < t \tag{23}$$

for some given *tolerance* $t > 0$, we use *all* the available information on f, namely $f(B)$, and choose $h = (\cdot)^{\beta_{j+1}} + h'$ as an ideal member if

$$h' = \operatorname{argmin}_{y\in\mathbb{C}^{\beta_j}} \Phi_{\beta^{j+1},f}(y) \qquad \text{and} \qquad \Phi_{\beta^{j+1},f}(h') < t. \tag{24}$$

That is, we first compute the solution of the optimization problem and then verify if the value of the target function is sufficiently small with respect to all consistency conditions, requesting not only an optimal but even a good solution. Clearly, this depends on the chosen tolerance $t > 0$.

Once a monic polynomial $h = (\cdot)^{\beta^{j+1}} + h'$ satisfies (24), we proceed as in (Sauer, 2018), make h a member of the ideal basis H, replace B by $B \setminus (\beta^{j+1} + \mathbb{N}_0^s)$ and proceed with the next multiindex in the modified B according to $\prec$.

It has been shown in (Sauer, 2018) that this process terminates and determines monic polynomials $h_j = (\cdot)^{\beta_j} + h'_j$, $j = 1, \ldots, n$, where the β^j now are the exponents identified as leading terms of the ideal basis in the above process and each h_j has coefficients in $B^* = L\left(\{\beta^j : j = 1, \ldots, n\}\right)$; In the case of an exact symbolic solution, the h_j are a Gröbner basis for the zero dimensional ideal $\mathcal{I} = \langle h_j : j = 1, \ldots, n \rangle$ they generate and

$$\Pi/\mathcal{I} = \Pi_{B^*} = \bigcap_{j=1}^{n} \ker D^{\beta_j}.$$

After computing the ideal basis, the frequencies can be obtained as joint eigenvalues of multiplication tables $M_1, \ldots, M_s \in \mathbb{C}^{B^* \times B^*}$ on $\Pi/\mathcal{I}$; recall that M_k represents the multiplication modulo the basis $h_1, \ldots, h_n$, i.e., its rows are

$$(M_k)_{:,\beta} = \begin{cases} -h'_j & \beta + \epsilon_k = \beta^j, \\ e_{\beta+\epsilon_k} & \text{otherwise}, \end{cases} \qquad \beta \in B^*. \tag{25}$$

These matrices represent multiplication modulo $\mathcal{I}$ if the h_j form an H-basis or a Gröbner basis and being such a basis is equivalent to the commuting of the multiplication tables, i.e., $M_j M_k = M_k M_j$, cf. Mourrain (2003). In one variable, (25) is exactly the Frobenius companion matrix.

Once the common zeros are determined via the joint eigenvalue problem, the coefficients are solutions of a linear system, cf. (Sauer, 2018). Note that the exponential polynomial f that has been reconstructed by determining its frequencies and their multiplicities in the above way, is the dual of the h_j in the sense that $(h_j, f) = 0$, $j = 1, \ldots, n$. With this extension we obtain a *stabilized SMILE* algorithm, for which we can state a backward error estimate on the difference between the given $\hat{f}$ and the f found by the algorithm. This of course only works if the algorithm detected the structure behind the data correctly, see Remark 10.

Theorem 2. *Let sampled data $\hat{f} \in \mathbb{C}^{\Lambda}$ be given and let*

$$G = \left\{ g_j = (\cdot)^{\beta_j} + g'_j : j = 1, \ldots, n \right\}, \qquad \operatorname{supp} g'_j \subseteq B^* \subseteq B,$$

be a reduced monic Gröbner basis for a zero dimensional ideal which satisfies

$$g'_j = \operatorname{argmin}_{\mathbb{C}^{B^*}} \Phi_{\beta^j, \hat{f}}(\cdot), \qquad j = 1, \ldots, n, \tag{26}$$

and

$$\Phi_{\beta^j, \hat{f}}(g'_j) < t, \qquad j = 1, \ldots, n. \tag{27}$$

Then there exists $f \in \mathbb{C}^{\Lambda}$ such that $\Phi_{\beta^j, f}(g_j) = 0$ and

$$\left\| f - \hat{f} \right\|_2 \leq C_G \sqrt{t}, \tag{28}$$

where the constant C_G depends only on the set G.

Remark 9. Theorem 2 is a backward stability result. Since usually the sampled data will be corrupted by noise, it will not be of exponential polynomial form any more, hence there exists no *exact* dual ideal for $\hat{f}$. The theorem says that if the solution of the optimization problem is an almost dual filter, then there exists an exponential sequence that approximates the samples with good accuracy. The quantitative relationship is expressed in (27) and (28).

Remark 10. The two assumptions of Theorem 2 are not easy to satisfy in general as a solution of the minimization problem (26) does not have to be a Gröbner basis. While the *exact symbolic solution* of the SMILE algorithm in (Sauer, 2018) yields a Gröbner basis when applied to *exact samples* of a exponential polynomial, this does not need to hold for the numerical version described above, even if its result satisfies (26) and (27). In other words, the h_j from the above numerical SMILE algorithm do *not* have to be a Gröbner basis for the ideal they generate. This means that the algorithm did not properly detect the algebraic structure of the exponential polynomial "hidden" in the data, whichs can be due to poor quality of the samples or a too large tolerance value t. A similar situation arises in the detection of *approximate varieties*, cf. Heldt et al. (2009); Sauer (2007). If the computed approximate

basis G is not a Gröbner basis of the ideal it generates, the algorithm can be considered to have failed; fortunately, this situation can be detected by means of Buchberger's algorithm: if a syzygy of leading terms does not reduce to zero, the g_j do not form a Gröbner basis.

The proof of the theorem uses a relationship between Hankel and Toeplitz matrices which is of independent interest. Adapted to our situation, it takes the following form.

Lemma 5. *For $\alpha \in \mathbb{Z}^s$, $f \in \mathbb{C}^{\alpha+A+B}$ and $y \in \mathbb{C}^B$ we have that*

$$\tau^\alpha H_{A,B}(f)\, y = T_{A,B}(y_-)\, \tau^\alpha f, \tag{29}$$

where $y_- \in \mathbb{C}^{A-B}$ is defined as the reverse of the zero padded vector y, i.e.,

$$y_-(\beta) := \begin{cases} y(-\beta), & \beta \in -B, \\ 0, & \textit{otherwise}, \end{cases} \qquad \beta \in \mathbb{Z}^s. \tag{30}$$

Proof. We compute

$$\begin{aligned} \tau^\alpha H_{A,B}(f)\, y = H_{\alpha+A,B}(f)\, y &= \left[\sum_{\beta\in B} f(\gamma+\beta)\, y(\beta) : \gamma \in \alpha + A \right] \\ &= \left[\sum_{\beta\in B} (\tau^\alpha f)(\gamma+\beta)\, y(\beta) : \gamma \in A \right] = \left[\sum_{\beta\in B+\gamma} (\tau^\alpha f)(\beta)\, y(\beta-\gamma) : \gamma \in A \right] \\ &= \left[\sum_{\beta\in B+\gamma} (\tau^\alpha f)(\beta)\, y_-(\gamma-\beta) : \gamma \in A \right] = \left(y_- * \tau^\alpha f\right)(A) = T_{A,B}(y_-)\, \tau^\alpha f \end{aligned}$$

to obtain (29). □

Remark 11. Setting $\alpha = 0$ in (29) yields a way to interchange the roles of filter and signal by passing from Hankel to Toeplitz operators and vice versa.

Definition 6. Given $f \in \ell(\mathbb{Z}^s)$ and a finite $A \subset \mathbb{Z}^s$, we define

$$f(A) = [f(\alpha) : \alpha \in A] \in \mathbb{C}^A.$$

Lemma 6. *Given the sampling vector $f(\Lambda) \in \mathbb{C}^\Lambda$ and $y' \in \mathbb{C}^{B_j}$, there exists a symmetric and positive semidefinite matrix $M(y') \in \mathbb{C}^{B\times B}$ such that*

$$\Phi_{\beta^{j+1},f}(y') = f(B)^H\, M(y')\, f(B), \qquad y' \in \mathbb{C}^{B_j}. \tag{31}$$

Proof. For $y' \in \mathbb{C}^{B_j}$ we set $y = \begin{bmatrix} y' \\ 1 \end{bmatrix} \in \mathbb{C}^{B_{j+1}}$. Since we assumed that the samples $f(\Lambda)$ and hence $f(B)$ are known, we can choose any $\alpha \in (B - B_{j+1})_+$, hence $\alpha + B_{j+1} \subseteq B$, and obtain by means of (29) that

$$\begin{aligned} H_{\alpha+A,B_{j+1}}(f)\, y &= T_{A,B_{j+1}}(y_-)\, \left(\tau^\alpha\right) f(B_j) = T_{A,B_{j+1}}(y_-) f(\alpha + B_j) \\ &= T_{A,B_{j+1}}(y_-)\, E_\alpha f(B) =: M_\alpha(y')\, f(B), \end{aligned}$$

where

$$E_\alpha := \sum_{\beta\in B_j} e_\beta e_{\alpha+\beta} \in \mathbb{R}^{B_j\times B}.$$

Now we conclude that

$$\begin{aligned}\sum_{\alpha\in(B-B_{j+1})_+} \left\|H_{\alpha+A,B_{j+1}}(f)\,y\right\|_2^2 &= \sum_{\alpha\in(B-B_{j+1})_+} \left\|M_\alpha(y')\,f(B)\right\|_2^2\\ &= f(B)^H\left(\sum_{\alpha\in(B-B_{j+1})_+} E_\alpha^T T_{A,B_{j+1}}(y_-)^H T_{A,B_{j+1}}(y_-)\,E_\alpha\right) f(B)\\ &=: f(B)^H\,M(y')\,f(B),\end{aligned}$$

where $M(y')$ is a symmetric positive semidefinite matrix, hence (31) follows; also note that

$$\ker M(y') = \bigcap_{\alpha\in(B-B_{j+1})_+} \ker M_\alpha(y'). \quad \square \tag{32}$$

Before we begin with the proof of Theorem 2, let us briefly revise the main idea behind solving the difference equation $(G,\cdot)=0$. For simplicity, we assume that the finitely many common zeros $Z\subset\mathbb{C}^s$ of G are simple and have no zero components. Then any exponential function of the form

$$f(x) = \sum_{\zeta\in Z} f_\zeta \zeta^x, \qquad f_\zeta\in\mathbb{C},$$

yields a Hankel operator with the factorization

$$H(f) = V(Z,\mathbb{N}_0^s)^T \operatorname{diag}\left[f_\zeta : \zeta\in Z\right] V(Z,\mathbb{N}_0^s)$$

and for any $g\in G$, we have that

$$(g,f) = H(f)g = V(Z,\mathbb{N}_0^s)^T \operatorname{diag}\left[f_\zeta : \zeta\in Z\right][g(\zeta) : \zeta\in Z] = 0,$$

since $g(Z)=0$. This is also the main idea in Prony's reconstruction method and the motivation for the Prony ideal.

Proof of Theorem 2. Since $\langle G\rangle$ is zero dimensional, there exists, as explained above, an exponential polynomial sequence f such that

$$0 = \left(g_j, f\right), \qquad j=1,\dots,n, \tag{33}$$

which implies, since $B_{j+1}\subseteq B^*\subseteq B$ for $j=1,\dots,n$, that

$$0 = H_{\alpha+A,B_{j+1}}(f)\,g_j = H_{\alpha+A,B^*}(f)\,g_j = H_{\alpha+A,B}(f)\,g_j$$

since the support of g_j is B_j. Hence,

$$0 = T_{A,B^*}\left((g_j)_-\right) E_\alpha\, f(B) =: M_\alpha(g_j)\,f(B)$$

holds for $\alpha\in(B-B^*)_+$ and $j=1,\dots,n$. All biinfinite sequences f that satisfy (33) form a shift invariant subspace $\mathcal{F}$ of $\ell(\mathbb{Z}^s)$ whose restriction $\mathcal{F}(B)$ to B is a subspace of $\mathbb{C}^B$ and coincides with the intersection of the kernels of $M_\alpha(g_j)$, $j=1,\dots,n$. Indeed, the finitely many common zeros of the g_j, together with their multiplicities, define a finite dimensional space of exponential polynomials as their kernels, and this space is shift invariant. This allows us to conclude that

$$\bigcap_{j=1}^n \ker M_\alpha(g_j) = \bigcap_{j=1}^n \ker M_0(g_j) = \mathcal{F}(B), \qquad \alpha\in B-B^*,$$

since $\mathcal{F}(B)$ is shift invariant. Let $f(B) \in \mathbb{C}^B$ be the orthogonal projection of $\hat{f}(B)$ on $\mathcal{F}(B)$. If $f = \hat{f}$ then the theorem is trivially true. In addition, let λ denote the smallest nonzero eigenvalue of

$$M = \sum_{j=1}^{n} M_0(g_j)^H M_0(g_j) = \sum_{j=1}^{n} T_{A,B^*}\left((g_j)_-\right)^H T_{A,B^*}\left((g_j)_-\right) \in \mathbb{C}^{B^* \times B^*}, \tag{34}$$

and note that

$$\ker M = \mathcal{F}(B^*).$$

Then,

$$\begin{aligned}
nt &\geq \sum_{j=1}^{n} \sum_{\alpha \in (B - B_{j+1})_+} \left\| H_{\alpha + A, B_{j+1}}(\hat{f})\, g_j \right\|_2^2 \\
&= \sum_{j=1}^{n} \hat{f}(B)^H \left(\sum_{\alpha \in (B - B_{j+1})_+} M_\alpha(g_j)^H M_\alpha(g_j) \right) \hat{f}(B) \\
&= \sum_{j=1}^{n} \left(f(B) - \hat{f}(B) \right)^H \left(\sum_{\alpha \in (B - B_{j+1})_+} M_\alpha(g_j)^H M_\alpha(g_j) \right) \left(f(B) - \hat{f}(B) \right) \\
&\geq \sum_{j=1}^{n} \left(f(B) - \hat{f}(B) \right)^H \left(\sum_{\alpha \in (B - B^*)_+} M_\alpha(g_j)^H M_\alpha(g_j) \right) \left(f(B) - \hat{f}(B) \right) \\
&\geq \sum_{\alpha \in (B - B^*)_+} \left(E_\alpha \left(f(B) - \hat{f}(B) \right) \right)^H M \left(E_\alpha \left(f(B) - \hat{f}(B) \right) \right).
\end{aligned}$$

Each summand $E_\alpha(f(B) - \hat{f}(B))$ is either zero or perpendicular to the kernel of M by our choice of f as an orthogonal projection, hence

$$\begin{aligned}
&\sum_{\alpha \in (B - B^*)_+} \left(E_\alpha \left(f(B) - \hat{f}(B) \right) \right)^H M \left(E_\alpha \left(f(B) - \hat{f}(B) \right) \right) \\
&\quad \geq \lambda \sum_{\alpha \in (B - B^*)_+} \left\| E_\alpha \left(f(B) - \hat{f}(B) \right) \right\|_2^2 = \lambda \sum_{\beta \in B} \left(f(\beta) - \hat{f}(\beta) \right)^2 c_\beta \\
&\quad \geq \lambda \left\| f(B) - \hat{f}(B) \right\|_2^2,
\end{aligned} \tag{35}$$

where

$$c_\beta := \#\left\{ \gamma \in (B - B^*)_+ : \beta \in \gamma + B^* \right\} \geq 1 \tag{36}$$

since $(0 \in B - B^*)_+$. Hence

$$\left\| f(B) - \hat{f}(B) \right\|_2^2 \leq \tau \, \frac{n}{\lambda},$$

as claimed. □

Remark 12. The two constants λ and n that define the constant C_G in (28) depend only on the Gröbner basis G. n is just the cardinality of the basis and also the matrix $M(y')$ depends only on this basis. Though these numbers depend on the accuracy τ, they are changing discontinuously with τ and at least remain constant for small perturbations of f due to (24).

Theorem 2 shows no effect or benefit of oversampling, however. In fact, the effect of the oversampling shows in the coefficients c_β from (36), because of which we call them *structural oversampling factors*. It is important to note that some of these coefficients at the corners of B, especially $\beta = 0$, will always have $c_\beta = 1$, hence we cannot get a worst case improvement by oversampling since it cannot be excluded that $f(\beta) \neq \hat{f}(\beta)$ only if $c_\beta = 1$.

Nevertheless, improvement has to be expected when looking at the average situation. To illustrate this, let us assume for simplicity that $\hat{f}(\beta)$ are independent identically distributed (iid) random values with mean $f(\beta)$ and variance σ^2, $\beta \in B$. Then the expectation for the 2-norm is

$$\mathbb{E}\left(\left\| f(B) - \hat{f}(B)\right\|_2^2\right) = \#B\,\sigma^2,$$

while

$$\mathbb{E}\left(\sum_{\beta\in B}\left(f(\beta) - \hat{f}(\beta)\right)^2 c_\beta\right) = \sigma^2 \sum_{\beta\in B} c_\beta = \left(\sum_{\beta\in B}\frac{c_\beta}{\#B}\right)\mathbb{E}\left(\left\| f(B) - \hat{f}(B)\right\|_2^2\right),$$

so that (28) can be replaced by

$$\mathbb{E}\left(\|f - \hat{f}\|_2^2\right)^{\frac{1}{2}} \le C_G \left(\sum_{\beta\in B}\frac{c_\beta}{\#B}\right)^{-1/2}\sqrt{\tau}. \tag{37}$$

Example 1. Let us begin with $s = 2$ and the situation of a tensor product configuration with $B^* = \{0,\dots,n\}^2$ and $B = \{0,\dots,n+1\}^2$, i.e., we oversample by one bi-degree. Then $(B - B^*)_+ = \{(0,0),(0,1),(1,0),(1,1)\}$ and we have that

$$c_\beta = \begin{cases} 1 & \beta \in \{(0,0),(n,0),(0,n+1),(n+1,n+1)\}, \\ 4 & \beta \in \{1,\dots,n\}^2 \\ 2 & \text{otherwise} \end{cases}$$

and therefore

$$\sum_{\beta\in B}\frac{c_\beta^{-1}}{\#B} = \frac{4 + 4n^2 + 2\left((n+2)^2 - n^2 - 4\right)}{(n+2)^2} = 4 - \frac{8n+8}{(n+2)^2}$$

so that this simple oversampling improves the backwards error by a factor of 2 in expectation.

Example 2. The standard situation for total degree bases algorithms like the one we present here is that $B^* = \Gamma_n := \{\alpha : |\alpha| \le n\}$ for some n; though normally it will only be subset with some multi-indices of length n missing, we assume the above for simplicity. Now we assume that the oversampled set will be $B = \Gamma_n + \Gamma_k = \Gamma_{n+k}$ and therefore $(B - B^*) = \Gamma_k$. To give a lower estimate for $\sum c_\beta$ we note that $c_\beta = \binom{k+s}{s}$ whenever $\beta \ge \epsilon := (k,\dots,k)$ and $|\beta| \le n$ which means that $\beta \in \epsilon + \Gamma_{n-sk} =: B_+$. This already restricts k to $k \le n/s$ in order to fully explore the effect of the oversampling. Then,

$$\begin{aligned}\sum_{\beta\in B_+}\frac{c_\beta}{\#B} &= \binom{k+s}{s}\frac{\#B_+}{\#B} = \binom{k+s}{s}\frac{\binom{n-(k-1)s}{s}}{\binom{n+k+s}{s}} = \binom{k+s}{s}\frac{(n-(k-1)s\cdots(n-ks+1)}{(n+k+s)\cdots(n+k+1)} \\ &= \binom{k+s}{s}\left(1 - \frac{k(s+1)}{n+k+s}\right)\cdots\left(1 - \frac{k(s+1)}{n+k+1}\right) \ge \binom{k+s}{s}\left(1 - \frac{k(s+1)}{n+k+1}\right)^s\end{aligned}$$

and using for c_β, $\beta \in B \setminus B_+$, the estimate $c_\beta \ge 1$, we get for any $0 < \lambda < 1$ that

$$\sum_{\beta\in B}\frac{c_\beta}{\#B}\geq\sum_{\beta\in B_+}\frac{c_\beta}{\#B}+\frac{\#B-\#B_+}{\#B}=1+\left(\binom{k+s}{s}-1\right)\frac{\#B_+}{\#B}$$
$$\geq 1+\left(\binom{k+s}{s}-1\right)\left(1-\frac{k(s+1)}{n+k+1}\right)^s\geq 1+\left(\binom{k+s}{s}-1\right)\lambda^s$$

as long as $k\leq\frac{(n+1)(1-\lambda)}{s+\lambda}$. Hence, oversampling *always* helps since the constant is >1 and the constant improves with an order of $k^{-s/2}$, provided that k is still relatively small relative to n.

Example 3. To show that the oversampling must respect the structure of B^*, we consider the extremal case when $B^*=\{0,\epsilon_1,\ldots,n\epsilon_1\}$ and we oversample to $B=[0,n]^s$. Then it is easily seen that $c_\beta=1$, $\beta\in B$, and even in expectation we get no improvement through oversampling. This coincides with the intuition since the oversampling was performed orthogonal to the geometry of B^*.

Let us summarize the results of this section: The augmented Prony algorithm based on least squares optimization admits a backwards analysis and is stable up to a problem depended constant C_G that depends on the algebraic structure of the problem found. It can be applied to arbitrary unstructured data and finds the closest exponential sequence. Moreover,

1. oversampling the data gives no improvement in the worst case, but this requires a special distribution of the error.
2. In expectation for the simple case of iid perturbations with zero mean it can be shown that oversampling is beneficial and will improve the stability of reconstruction of the frequencies.
3. The oversampling has to be adapted to the geometric and algebraic structure of the problem, more precisely to the support set of the Gröbner basis.

We assume that a reasonable way of oversampling might be to consider $B=B^*+\Gamma_k$, but even that would require a priori knowledge of B^* or at least a superset of B^*.

4. Model reduction

In our approach to model reduction we aim to approximately representing a sequence $\hat{f}$, again available on a finite set $\Lambda\subset\mathbb{N}_0^s$, with a structurally simpler exponential polynomial sequence, and to do this by determining the *difference equation* that defines the exponential polynomial. In other words, the goal is to compute an ideal basis H that eventually defines an exponential polynomial that is close to the original $\hat{f}$.

In this context, we do not even have to assume that $\hat{f}$ is of exponential polynomial form and simply look for a best approximation. In particular, $\hat{f}$ can be a noisy exponential polynomial sequence. The task is now to determine an approximating f and its ideal of dual filters simultaneously. Since we formulate the approximation in an ℓ_2 sense, it requires that $\hat{f}\in\ell_2(\mathbb{N}_0)$ is decaying and therefore the frequencies according to the optimal approximant will need to have negative real part and the domain of summation will only be chosen as $\mathbb{N}_0^s$, cf. Gragg and Reichel (1989) for the univariate case.

Simplicity will be defined by a *lower set* $B\subset\mathbb{N}_0^s$ with the property that there exists an index set A such that $\Lambda=A+B$. Of course, for any B there exist lots of sampling sets Λ with this property. In one variable the lower set B would necessarily be of the form $\mathbb{Z}_n:=\{0,\ldots,n-1\}$.

Any lower set $B'\subset B$ can be written as the intersection

$$B'=\bigcap_{j=1}^{n}\left(\beta^j+\mathbb{N}_0^s\right) \tag{38}$$

of upper sets, where we assume that $\beta^j\in B$, $j=1,\ldots,n$. Then any choice of polynomials $h_j\in(\cdot)^{\beta^j}+\Pi_{B'}$, $j=1,\ldots,s$, generates a zero dimensional ideal $\mathcal{I}$ with the property that $\Pi/\mathcal{I}\subseteq\Pi_{B'}$ with quality if and only if $H:=\left\{h_j:j=1,\ldots,n\right\}$ is an H-basis or a Gröbner basis or a border basis for $\mathcal{I}$, cf.

Mourrain (2003); in the case of these monic polynomials, all three notions coincide. Being such a good basis is in turn characterized by the commuting of the multiplication tables M_j from (25) which defines a quadratic constraint for H as long as the β^j are fixed. If this condition is not enforced, H can be replaced by an H-basis for $\langle H\rangle$ which is equivalent to replacing B' by a lower subset of B'.

Based on this concept, we can define the infinite dimensional model reduction problem in several variables as follows:

$$\min_{f\in\ell(\mathbb{N}_0^s),H} \|f-\hat{f}\|_2^2, \qquad \text{subject to } (H,f)=0, \qquad H=\left[h_j=(\cdot)^{\beta^j}+\Pi_{B'} : j=1,\dots,n\right], \tag{39}$$

where

$$(H,f) := \left[(h_j,f) : j=1,\dots,n\right] \in \ell^n(\mathbb{Z}^s).$$

Note that the constraint $(H,f)=0$ ensures that f is the solution of a homogeneous system of difference equations with constant coefficients and therefore an exponential polynomial function. Nevertheless, the optimization problem is non-convex and therefore hard to solve; in addition, it uses infinite information on $\hat{f}$.

The heuristic finite version of (39) then makes use of the sampled data only and reads as

$$\min_{f\in\mathbb{C}^B,H} \left\|f-\hat{f}\right\|_{B,2}^2, \qquad \text{subject to} \quad 0=\left[H_{A,B}(f)\,h_j : j=1,\dots,n\right] \in \left(\mathbb{C}^A\right)^n. \tag{40}$$

The constraint in (40) is equivalent to $(H,f)(A)=0$, i.e., it localizes the duality constraint while B encodes the domain of summation. If $A\supseteq\Upsilon_{\#B'}$, then H must be a basis for the associated Prony ideal whenever f is an exponential polynomial sequence with at most $\#B'$ terms. This means that the solution of (39) is admissible for (40) and the associated H is the one that belongs to this solution. In other words, the problem (40) is consistent for $A,B\to\mathbb{N}_0^s$.

Nevertheless, the finite problem is still non-convex which will be addressed by switching to minimizing the convex relaxation

$$\Psi_\mu(f,H) := \left\|f-\hat{f}\right\|_{B,2}^2 + \frac{1}{\mu}\sum_{j=1}^n \left\|H_{A,B}(f)\,h_j\right\|_2^2, \qquad \mu>0, \tag{41}$$

with respect to f and H.

Since the solution f^* of (40) satisfies $H_{A,B}(f^*)h_j=0$, $j=1,\dots,n$, it immediately follows that

$$\left\|f^*-\hat{f}\right\|_{B,2}^2 = \left\|f^*-\hat{f}\right\|_{B,2}^2 + \frac{1}{\mu}\sum_{j=1}^n \left\|H_{A,B}(f^*)\,h_j\right\|_2^2, \qquad \mu>0, \tag{42}$$

holds *independently of* μ, which yields the following observation.

Lemma 7. *For any H we have that*

$$\Psi_\mu(f^*,H) = \lim_{\mu\to 0}\min_f \Psi_\mu(f,H). \tag{43}$$

Proof. From the definition of f^* and (42), we know that $\Psi_\mu(f^*)$ is independent of μ and minimizes each of these functionals among all choices f such that $H_{A,B}(f)\,h_j=0$, $j=1,\dots,n$. If, on the other hand, some f does not satisfy the constraint, then

$$\sum_{j=1}^n \left\|H_{A,B}(f)\,h_j\right\|_2^2 > 0$$

and $\Psi_\mu(f) > \Psi_\mu(f^*)$ whenever

$$\mu\left(\left\|f^*-\hat{f}\right\|_{B,2}^2-\left\|f-\hat{f}\right\|_{B,2}^2\right)<\sum_{j=1}^n\left\|H_{A,B}(f)\,h_j\right\|_2^2$$

so that for

$$\mu<\frac{1}{\left|\left\|f^*-\hat{f}\right\|_{B,2}^2-\left\|f-\hat{f}\right\|_{B,2}^2\right|}\sum_{j=1}^n\left\|H_{A,B}(f)\,h_j\right\|_2^2$$

we have $\Psi_\mu(f^*) < \Psi_\mu(f)$ and f^* must indeed be the minimizer as claimed in (43). □

The function (41) can be optimized by alternating minimization. For that purpose, we first note that by Lemma 5 we can rewrite Ψ_μ as

$$\Psi_\mu(f,H)=\left\|f-\hat{f}\right\|_{B,2}^2+\frac{1}{\mu}\left\|T_{A,B}(H_-)\,f\right\|_2^2=\left\|f-\hat{f}\right\|_{B,2}^2+\frac{1}{\mu}\sum_{j=1}^n\left\|T_{A,B}((h_j)_-)\,f\right\|_2^2$$

so that

$$\nabla_f\Psi_\mu(f,H)=2\left(f-\hat{f}\right)+\frac{2}{\mu}T_{A,B}(H_-)^HT_{A,B}(H_-)\,f \tag{44}$$

is now linear in f. On the other hand, taking into account that $h_j \in \mathbb{C}^B$ can be written as $h_j = e_{\beta^j} + h'_j$, $h'_j \in \mathbb{C}^{B'}$, we get that

$$\Psi_\mu(f,H)=\|f-\hat{f}\|_{B,2}^2+\frac{1}{\mu}\sum_{j=1}^n\left\|H_{A,B}(f)\,\left(e_{\beta^j}+h'_j\right)\right\|_2^2$$

so that

$$\nabla_{h'_j}\Psi_\mu(f,H)=\frac{2}{\mu}\left(H_{A,B}(f)^Te_{\beta^j}+H_{A,B}(f)^TH_{A,B}(f)\,h'_j\right),\qquad j=1,\ldots,n, \tag{45}$$

hence

$$\begin{aligned}\frac{\mu}{2}&\nabla_{H'}\Psi_\mu(f,H)\\&=\begin{bmatrix}H_{A,B}^Te_{\beta^1}\\\vdots\\H_{A,B}^Te_{\beta^n}\end{bmatrix}+\begin{bmatrix}H_{A,B}(f)^TH_{A,B}(f)&&\\&\ddots&\\&&H_{A,B}(f)^TH_{A,B}(f)\end{bmatrix}\begin{bmatrix}h'_1\\\vdots\\h'_n\end{bmatrix}\\&=\left(1\otimes H_{\beta^1,\ldots,\beta^n,:}\right)+\left(I_n\otimes H_{A,B}(f)^TH_{A,B}(f)\right)H'\end{aligned}$$

is linear in $H' := \left[h'_j : j = 1,\ldots,n\right]$ as well. The complete gradient of Φ_μ can then be given as

$$\begin{aligned}\frac{\mu}{2}\nabla\Psi_\mu(f,H)&:=\frac{\mu}{2}\begin{bmatrix}\nabla_f\Psi_\mu(f,H)\\\nabla_{H'}\Psi_\mu(f,H)\end{bmatrix}\\&=\begin{bmatrix}\mu(f-\hat{f})\\1\otimes H_{\beta^1,\ldots,\beta^n,:}\end{bmatrix}+\begin{bmatrix}T_{A,B}(H_-)^HT_{A,B}(H_-)&\\&I_n\otimes H_{A,B}(f)^TH_{A,B}(f)\end{bmatrix}\begin{bmatrix}f\\H\end{bmatrix}.\end{aligned} \tag{46}$$

These expressions indeed suggest to minimize the convex (but non quadratic) functional by means of *alternating optimization*, starting with some f_0, H^0 and then solving, in the nth iteration

$$\min_f \Psi_\mu(f, H)$$

to obtain f_{n+1} as the solution of the linear system

$$\hat{f} = \left(I - \frac{1}{\mu} T_{A,B}(H_-^n)^T T_{A,B}(H_-^n) \right) f; \tag{47}$$

this is uniquely solvable for μ sufficiently small. By (45) the minimization problem

$$\min_{H'} \Psi_\mu(f, H)$$

turns into solving the equations

$$-H_{A,B}(f_{n+1})^T e_{\beta_j} = H_{A,B'}(f_{n+1})^T H_{A,B'}(f_{n+1}) h'_{n+1,j}, \qquad j = 1, \ldots, n, \tag{48}$$

whose solutions define the next basis candidate H^{n+1}. Note that the matrix in these linear systems is *independent* of j and thus has to be computed and factorized only once.

Remark 13. Moreover, the basis H that is computed in (48) need *not* be an H-basis for the ideal it generates. This can be interpreted in the sense that $\hat{f}$ can be approximated by a *simpler* model; in fact, computing an H-basis of $\langle H_{n+1} \rangle$ corresponds to adding polynomials with leading terms in $\Pi_{B'}$ to the computed H_{n+1} which does not add requirements on A or B.

Since each step (47) and (48) reduces the target function if the respective gradient is nonzero, the alternating minimization converges in that case. In that case, the resulting basis H_* is the *implicit* reduced model and its common zeros define a finite dimensional space of exponential polynomials from which the best approximation to f can be determined by standard means.

Remark 14. The choice of μ can be made adaptively. Starting with $\mu_0 = 1$ and arriving at an f_{n+1} such that

$$\|f_{n+1} - \hat{f}\|_{B,2}^2 + \frac{1}{\mu_n} \sum_{j=1}^n \|H_{A,B}(f_{n+1}) h_j\|_2^2 < \|f_n - \hat{f}\|_{B,2}^2 + \frac{1}{\mu_n} \sum_{j=1}^n \|H_{A,B}(f_n) h_j\|_2^2, \tag{49}$$

any choice of

$$\mu_n > \mu_{n+1} > \frac{\sum_{j=1}^n \|H_{A,B}(f_{n+1}) h_j\|_2^2}{\|f_n - \hat{f}\|_{B,2}^2 - \|f_{n+1} - \hat{f}\|_{B,2}^2 + \frac{1}{\mu_n} \sum_{j=1}^n \|H_{A,B}(f_n) h_j\|_2^2}$$

still leads to an improvement of the target function in the optimization step with respect to f, but increases the emphasis on the side condition of duality.

This choice ensures that

$$\Psi_{\mu_0}\left(f_0, H^0\right) > \Psi_{\mu_1}\left(f_1, H^0\right) \geq \Psi_{\mu_1}\left(f_1, H^1\right) > \Psi_{\mu_2}\left(f_2, H^1\right) \geq \cdots \geq 0$$

and since this sequence is bounded from below, it will eventually converge. If the resulting μ is too large, the process can be restarted with $\frac{\mu}{2}$ since (42) insures that $\min_f \Psi_\mu(f, H) \leq \|f^* - \hat{f}\|_{B,2}^2$ for any $\mu > 0$ and any H.

Remark 15. Since (46) gives an explicit expression for the complete gradient of Ψ_μ with respect to both parameters that can be evaluated efficiently, the optimization can also be performed by a gradient descent algorithm, cf. (Nocedal and Wright, 1999).

5. Summary

Using consistency conditions, we state a numerically stabilized version of Prony's method and derive a first backward error analysis result in the sense that whenever the algorithm to solve Prony's problem yields approximately dual filters h_j for the given data $\hat{f}$, then there exists an exponential polynomial sequence f that is close to $\hat{f}$ and is an exact dual of the filters. In other words, the data is only a small perturbation of the exact function which is reconstructed by the algorithm. We give a quantitative statement for the error and show that *properly chosen* oversampling can improve the stability of the reconstruction at least in the average, but examples show that the oversampling has to be chosen in a way that is compatible with the problem structure.

Moreover, given any data, an optimal approximant with a certain a priori prescribed "simpler" ideal structure can be computed by alternating optimization algorithms that correspond only to solving moderately sized linear systems. We formulate it here as a signal approximation problem, but it could also be seen as structured approximation of matrices.

Even if the understanding of the methods presented here require a certain amount of algebra, especially ideal bases, the computations themselves rely entirely on methods from linear algebra and can thus be performed numerically and with high efficiency.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgement

This work was partially supported through the Spanish research grant PGC2018-096321-B-I00 (MCIU/AEI), by Gobierno de Aragón (E41-17R,E41-20R) and Feder 2014-2020 "Construyendo Europa desde Aragón".

We also want to thank Benedikt Diederichs for a lot of helpful comments on an earlier version of this manuscript and the referees for corrections and valuable suggestions.

References

Batenkov, D., Diederichs, D., Goldman, G., Yomdin, Y., 2021. The spectral properties of Vandermonde matrices with clustered nodes. Linear Algebra Appl. 609, 37–72.

Cox, D., Little, J., O'Shea, D., 1996. Ideals, Varieties and Algorithms, 2. ed. Undergraduate Texts in Mathematics. Springer–Verlag.

Cuyt, A., Lee, W., 2018. Multivariate exponential analysis from the minimal number of samples. Adv. Comput. Math. 44, 987–1002.

Derevianko, N., Plonka, G., Petz, M., 2023. From ESPRIT to ESPIRA: estimation of signal parameters by iterative rational approximation. IMA J. Numer. Anal. 43, 789–827. https://doi.org/10.1093/imanum/drab108.

Gragg, W.B., Reichel, L., 1989. On singular values of Hankel operators of finite rank. Linear Algebra Appl. 121, 53–70.

Hamming, R.W., 1989. Digital Filters. Prentice–Hall. Republished by Dover Publications. 1998.

Harmouch, J., Mourrain, B., 2018. Structured low rank decomposition of multivariate Hankel matrices. Linear Algebra Appl. 542, 162–185.

Heldt, D., Kreuzer, M., Pokutta, S., Poulisse, H., 2009. Approximate computation of zero-dimensional polynomial ideals. J. Symb. Comput. 44, 1566–1591.

Jordan, C., 1965. Calculus of Finite Differences, 3rd ed. Chelsea.

Kaltofen, E.L., Yang, Z., 2016. Sparse multivariate function recovery with a small number of evaluations. J. Symb. Comput. 75, 209–218.

Möller, H.M., Sauer, T., 2004. Multivariate refinable functions of high approximation order via quotient ideals of Laurent polynomials. Adv. Comput. Math. 20, 205–228.

Mourrain, B., 2003. A new criterion for normal form algorithms. In: Fossorier, M., Imai, H., Lin, S., Poli, A. (Eds.), Applied Algebra, Algebraic Algorithms and Error-Correcting Codes. AAECC 1999. Springer, pp. 430–442.

Mourrain, B., 2018. Polynomial-exponential decomposition from moments. Found. Comput. Math. 18, 1435–1492.
Nocedal, J., Wright, S.J., 1999. Numerical Optimization. Springer Series in Operations Research. Springer.
Osborne, M.R., Smyth, G.K., 1991. A modified Prony algorithm for fitting functions defined by difference equations. SIAM J. Sci. Stat. Comput. 12, 362–382.
Osborne, M.R., Smyth, G.K., 1995. A modified Prony algorithm for exponential function fitting. SIAM J. Sci. Comput. 16, 119–138.
Prony, C., 1795. Essai expérimental et analytique sur les lois de la dilabilité des fluides élastiques, et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes températures. J. Éc. Polytech. 2, 24–77.
Roy, R., Kailath, T., 1989. ESPRIT – estimation of signal parameters via rotational invariance techniques. IEEE Trans. Acoust. Speech Signal Process. 37, 984–995.
Sauer, T., 2007. Approximate varieties, approximate ideals and dimension reduction. Numer. Algorithms 45, 295–313.
Sauer, T., 2016. Kernels of discrete convolutions and subdivision operators. Acta Appl. Math. 145, 115–131. https://doi.org/10.1007/s10440-016-0051-8. arXiv:1403.7724.
Sauer, T., 2017a. Prony's method in several variables. Numer. Math. 136, 411–438. https://doi.org/10.1007/s00211-016-0844-8. arXiv:1602.02352.
Sauer, T., 2017b. Reconstructing sparse exponential polynomials from samples: difference operators, Stirling numbers and Hermite interpolation. In: Floater, M., Lyche, T., Mazure, M.L., Moerken, K., Schumaker, L.L. (Eds.), Mathematical Methods for Curves and Surfaces. 9th International Conference. MMCS 2016, Tønsberg, Norway. Revised Selected Papers. Springer, pp. 233–251. arXiv:1610.02780.
Sauer, T., 2018. Prony's method in several variables: symbolic solutions by universal interpolation. J. Symb. Comput. 84, 95–112. https://doi.org/10.1016/j.jsc.2017.03.006. arXiv:1603.03944.
Schmidt, R., 1986. Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propag. 34, 276–280.
Zhang, R., Plonka, G., 2019. Optimal approximation with exponential sums by a maximum likelihood modification of Prony's method. Adv. Comput. Math. 45, 1657–1687.