Journal of Symbolic Computation 84 (2018) 95–112

Contents lists available at ScienceDirect

Journal of Symbolic Computation

www.elsevier.com/locate/jsc

Prony's method in several variables: Symbolic solutions by universal interpolation

Tomas Sauer

Lehrstuhl für Mathematik mit Schwerpunkt Digitale Bildverarbeitung & FORWISS, Universität Passau, Innstr. 43, D-94053 Passau, Germany

ARTICLE INFO

Article history:
Received 30 March 2016
Accepted 7 March 2017
Available online 24 March 2017

MSC:
13P10
65T40
41D05

Keywords:
Prony's method
Gröbner basis
Degree reducing interpolation
Hyperbolic cross

ABSTRACT

The paper considers a symbolic approach to Prony's method in several variables and its close connection to multivariate polynomial interpolation. Based on the concept of universal interpolation that can be seen as a weak generalization of univariate Chebychev systems, we can give estimates on the minimal number of evaluations needed to solve Prony's problem.

1. Introduction

Formulated in several variables, *Prony's problem* consists of reconstructing a function $f : \mathbb{R}^s \to \mathbb{C}$ of the form

$$f(x) = \sum_{\omega\in\Omega} f_\omega e^{\omega^T x}, \qquad f_\omega \in \mathbb{C} \setminus \{0\}, \quad \omega \in \Omega \subset (\mathbb{R} + i\mathbb{T})^s, \tag{1}$$

from discrete samples, where, as usual, $\mathbb{T}$ stands for the *torus* $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$. The restriction of the imaginary part of the frequencies is needed to avoid ambiguities in the solution. In one variable, this problem and its solution date back to Prony (Prony, 1795) in 1795 and since then various numerical methods have been devised to solve the problem, in particular the ESPRIT and MUSIC algorithms

E-mail address: Tomas.Sauer@uni-passau.de.

http://dx.doi.org/10.1016/j.jsc.2017.03.006
0747-7171/

(Roy and Kailath, 1989; Schmidt, 1986) from multi source radar detection with extensions to higher dimensions on grids in Rouquette and Najim (2001), Yilmazer et al. (2006). One should also consider Potts and Tasche (2015) for recent improvements and Plonka and Tasche (2014) for a survey on Prony's method and its extensions and generalizations. The *matrix pencil* approach for the Hankel matrices has also been considered in Hua and Sarkar (1990), Hua (1992) in one and two variables.

The closely related problem of reconstruction of sparse polynomials in several variables has also been considered by methods other than Prony's approach. Probabilistic methods with a small expected number of evaluations can be found in Zippel (1979), see also Zippel (1990) and, more recently, in Kaltofen and Yang (2016).

The multivariate version of Prony's method has gained some popularity recently and was approached by projection methods as in Diederichs and Iske (2015), Potts and Tasche (2013), as well as by more or less direct multivariate attempts Kunis et al. (2016). This paper is an extension of Sauer (2017), where the algebraic nature of the multivariate problem has been pointed out, resulting in a fast numerical method based on orthogonal H-bases.

Let us begin with a slightly informal presentation of the algebraic structure underlying Prony's problem in several variables: the approach consist of considering finite parts of the infinite *Hankel matrix*

$$F := \left[f(\alpha+\beta) : \alpha, \beta \in \mathbb{N}_0^s \right], \tag{2}$$

for example

$$F_n := [f(\alpha+\beta) : \alpha, \beta \in \Gamma_n], \qquad \Gamma_n := \left\{ \alpha \in \mathbb{N}_0^s : |\alpha| \le n \right\}, \tag{3}$$

with the standard *length* $|\alpha| = \alpha_1 + \cdots + \alpha_s$ of a *multiindex* $\alpha \in \mathbb{N}_0^s$. The crucial observation is that the *Prony ideal* I_Ω, the set of all polynomials vanishing on

$$X_\Omega = e^\Omega = \left\{ x_\omega = e^\omega = \left(e^{\omega_1}, \dots, e^{\omega_s} \right) : \omega \in \Omega \right\},$$

is in one-to-one correspondence with the kernels of the matrices F_n. More precisely, if we denote by Π_n the vector space of all polynomials of *total degree* at most n, and identify an polynomial

$$\Pi_n \ni p(x) = \sum_{|\alpha| \le n} p_\alpha \, x^\alpha$$

with its coefficient vector $p = [p_\alpha : \alpha \in \Gamma_n] \in \mathbb{C}^{\Gamma_n}$, then the following result from Sauer (2017) holds true.

Theorem 1. *For sufficiently large n we have that $p \in I_\Omega \cap \Pi_n$ if and only if $F_n p = 0$.*

The "sufficiently large" in Theorem 1 can be made concrete: it is the degree of some degree reducing interpolation space for X_Ω and since all degree reducing interpolation spaces have the same degree, this is indeed a *geometric* quantity depending on X_Ω only. The algorithm developed in Sauer (2017) relies on the a priori knowledge of such a sufficiently large n and then successively builds the matrices

$$F_{n,k} := \left[f(\alpha+\beta) : \begin{array}{c} \alpha \in \Gamma_n \\ \beta \in \Gamma_k \end{array} \right], \qquad k = 0, \dots, n,$$

for which the following observation has been made in Sauer (2017).

Theorem 2. *For sufficiently large n, we have that*

1. $\ker F_{n,k} \simeq I_\Omega \cap \Pi_k$, $k = 0, \dots, n$,
2. $k \mapsto \operatorname{rank} F_{n,k}$ *computes the* affine Hilbert function *for the ideal* I_Ω,

3. *there exists a number* $0 \le m \le n$ *such that*

$$\operatorname{rank} F_{n,0} < \cdots < \operatorname{rank} F_{n,m-1} = \operatorname{rank} F_{n,m} = \cdots = \operatorname{rank} F_{n,n} = \operatorname{rank} F_n,$$

and this m is the minimal choice for n.

Based on these structural observations, an algorithm can be derived that computes an orthogonal H-basis in the sense of Sauer (2001) as well as a graded homogeneous basis for the interpolation space Π/I_Ω entirely by application of standard techniques from Numerical Linear Algebra, namely singular valued decompositions and QR factorizations. This allows for fast and accurate solutions of even high dimensional problems in floating point arithmetic.

Once the H-basis and the interpolation space are determined, it is a fairly standard approach, see Gonzales-Vega et al. (1999), Stetter (1995) to determine the *multiplication tables*, a set of s commuting matrices of size $\#\Omega \times \#\Omega$ whose eigenvalues are the components of the x_ω that can be related by the respective eigenvectors, see Möller and Tenberg (2001).

This paper takes a somewhat different approach to Prony's problem by considering interpolation spaces spanned by a minimal number of monomials, see Boor (2007), Sauer (1997). While orthogonal H-bases are more favorable from a numerical point of view and work well in a numerical environment, the underlying methods from Numerical Linear Algebra, in particular orthogonal factorizations like the QR decomposition of matrices, cause difficulties in a symbolic framework due to the occurrence of square roots. In contrast to those numerical methods, this paper studies a *symbolic* approach that will provide us with a *minimal* sampling set for Prony's method and tell us what an (asymptotically) minimal number of evaluations of f needed for the reconstruction. In doing so, we will also gain some further inside into the algebraic structure of Prony's problem.

The paper is organized as follows. In Section 2 the notation will be fixed and Prony's problem will be expressed in terms of degree reducing interpolation. In Section 3 we study the fundamental algebraic tool, namely *uniform interpolation*. This means the identification of spaces that permit interpolation at any subset of $\mathbb{C}^s$ of a given cardinality. Based on this concept, Section 4 points out how to solve Prony's problem with a minimal number of evaluations. Detailed symbolic algorithms for that purpose are developed in Section 5. Finally, Section 6 briefly points out the connection to sparse polynomials and how those can be determined symbolically and Appendix A provides two valuable tools from computational ideal theory together with proofs.

All results presented in this paper are of algebraic nature. Of course, the numerical stability of the methods and of the reconstruction in general depends on the conditioning of Vandermonde matrices. This important and valuable question, however, is not in the scope of this paper.

2. Prony's problem revisited

Let $\Pi = \mathbb{C}[x_1, \dots, x_s]$ denote the algebra of polynomials in s variables with complex coefficients. We consider the *nonnegative integer grid* $\Gamma = \mathbb{N}_0^s$. For a finite $A \subset \Gamma$ we define

$$\Pi_A := \left\{ p(x) = \sum_{\alpha \in A} p_\alpha x^\alpha : p_\alpha \in \mathbb{C} \right\}$$

as the space of polynomials *supported* on A, a finite dimensional subspace of Π of dimension $\#A$. Recall that the *total degree* of a polynomial $p \in \Pi$ is defined as

$$\deg p := \max\{|\alpha| : p_\alpha \neq 0\}.$$

In the important case $A = \Gamma_n = \{\alpha : |\alpha| \le n\}$, we use the common abbreviation $\Pi_n := \Pi_{\Gamma_n}$ for the vector space of all polynomials of total degree (at most) n. In an analogous way, we define $\Gamma_n^0 = \{\alpha : |\alpha| = n\}$ as the set of *homogeneous* multiindices of length n and $\Pi_n^0 := \Pi_{\Gamma_n^0}$ as the *homogeneous* polynomials of degree (exactly) n.

The *coefficients* p_α of $p \in \Pi_A$ can be conveniently arranged into a vector $p = (p_\alpha : \alpha \in A) \in \mathbb{C}^A$; we use the same symbol for the polynomial and the vector, the respective meaning will be clear from the context. The next notion is standard in (polynomial) interpolation theory.

Definition 1. For finite sets $A \subset \Gamma$ and $X \subset \mathbb{C}^s$ we define the *Vandermonde matrix* $V(X, A)$ as

$$V(X, A) := \left[x^\alpha : \begin{array}{c} x \in X \\ \alpha \in A \end{array} \right].$$

Vandermonde matrices play a fundamental role in Prony's method as the following simple computation shows: for $A, B \subset \Gamma$ we define the Hankel matrix

$$F_{A,B} := \left[f(\alpha + \beta) : \begin{array}{c} \alpha \in A \\ \beta \in B \end{array} \right]. \tag{4}$$

Using the unit vectors $e_\alpha := (\delta_{\alpha,\alpha'} : \alpha' \in A) \in \mathbb{C}^A$, we find that

$$\begin{aligned}(F_{A,B})_{\alpha,\beta} = e_\alpha^T F_{A,B} e_\beta &= \sum_{\omega\in\Omega} f_\omega e^{\omega^T(\alpha+\beta)} = \sum_{\omega\in\Omega} f_\omega e^{\omega^T\alpha} e^{\omega^T\beta} \\ &= \sum_{\omega\in\Omega} e_\alpha^T V(X_\Omega, A)^T e_\omega f_\omega e_\omega^T V(X_\Omega, B) e_\beta,\end{aligned}$$

which yields the well-known factorization

$$F_{A,B} = V(X_\Omega, A)^T F_\Omega V(X_\Omega, B), \qquad F_\Omega := \operatorname{diag}(f_\omega : \omega \in \Omega), \tag{5}$$

already used in the univariate ESPRIT method (Roy and Kailath, 1989). Since, by assumption (1), $f_\omega \neq 0$, $\omega \in \Omega$, the rank of $F_{A,B}$ is at most $\#\Omega$ with equality if and only if

$$\operatorname{rank} V(X_\Omega, A) = \operatorname{rank} V(X_\Omega, B) = \#\Omega. \tag{6}$$

The meaning of (6) is well-known: Π_A and Π_B have to be interpolation spaces for X_Ω.

Definition 2. A subspace $\mathcal{P}$ of Π is called an *interpolation space* for X if for any $y \in \mathbb{C}^X$ there exists (at least one) $p \in \mathcal{P}$ such that $p(X) = y$.

Despite of its simple derivation, (6) has an immediate important consequence for Prony's problem and the reconstruction of Ω from $F_{A,B}$

Theorem 3. *The coefficients f_ω can be reconstructed from $F_{A,B}$ if and only if Π_A and Π_B are interpolation spaces with respect to X_Ω.*

Proof. Suppose that Π_A and Π_B are interpolation spaces with respect to X_Ω, then $\#A \geq \#\Omega$, and there exist coefficient vectors $p_\omega = (p_{\omega,\alpha} : \alpha \in A)$ such that for $\omega' \in \Omega$

$$\delta_{\omega,\omega'} = p_\omega(x_{\omega'}) = \sum_{\alpha\in A} p_{\omega,\alpha} x_{\omega'}^\alpha,$$

hence

$$V(X_\Omega, A) \left[p_{\omega,\alpha} : \begin{array}{c} \alpha \in A \\ \omega \in \Omega \end{array} \right] = I_{\#\Omega}.$$

In other words, this matrix is a right inverse of $V(X_\Omega, A)$ which we call $V(X_\Omega, A)^{-1}$ and since the same holds for $V(X_\Omega, B)$, it follows that

$$V(X_\Omega, A)^{-T} F_{A,B} V(X_\Omega, B)^{-1} = F_\Omega,$$

which reconstructs the coefficients under the assumption that Π_A and Π_B are interpolation spaces for X_Ω.

Conversely, if $\operatorname{rank} V(X_\Omega, B) < \#\Omega$, then there exists a nonzero diagonal matrix $F \in \mathbb{C}^{\Omega\times\Omega}$ such that $F\, V(X_\Omega, B) = 0$ and therefore

$$F_{A,B} = V(X_\Omega, A)^T\, (F_\Omega + F)\, V(X_\Omega, B)$$

so that F_Ω cannot be reconstructed from $F_{A,B}$. An analogous argument can also be used in the case that $\operatorname{rank} V(X_\Omega, A) < \#\Omega$. $\square$

Corollary 4. *Any sampling sets A, B for Prony's method must be chosen such that Π_A and Π_B are interpolation spaces for X_Ω.*

Definition 3. A subspace $\mathcal{P}$ of Π is called a *degree reducing interpolation space* for a finite set $X \subset \mathbb{C}^s$ if for any $q \in \Pi$ there exists a *unique* polynomial $p \in \mathcal{P}$ such that $p(X) = q(X)$ and $\deg p(X) \le \deg q(X)$.

Degree reducing interpolation spaces for some set $X \subset \mathbb{C}^s$ of *sites* have the advantage that they give the ideal $I_X = \{p \in \Pi : p(X) = 0\}$ almost for free. To be more concrete, let $A \subset \Gamma$ be such that Π_A is a degree reducing interpolation space for X and let $L_A : \Pi \to \Pi_A$ denote the *interpolation operator* defined by

$$(L_A p)(X) = p(X) \tag{7}$$

which is well defined because degree reducing interpolation is unique by definition. It can then be shown that the polynomials

$$h_\alpha := (\cdot)^\alpha - L_A(\cdot)^\alpha, \qquad \alpha \in A^c := \Gamma \setminus A, \tag{8}$$

form an *H-basis* of $I_X = \{q : q(X) = 0\}$, that is, any polynomial $q \in I_X$ can be written as

$$q = \sum_{\alpha \in A^c} q_\alpha\, h_\alpha, \qquad \deg q_\alpha \le \deg q - \deg h_\alpha, \tag{9}$$

where the sum in (9) is finite and uses the convention that $\deg p < 0$ iff $p = 0$. This fact, that is some folklore in ideal interpolation will be (re-)proved in even stronger form in Lemma 18 of the abstract.

By Corollary 4 we can reconstruct F_Ω from the symmetric sampling matrix $F_{A,A}$ if and only if Π_A is an interpolation space. A can be chosen minimally or at least of minimal cardinality by requesting Π_A to be a minimal degree interpolation space. In other words, computing a solution to Prony's problem with a minimal number of evaluations leads to determining such a space for the *unknown* point set X_Ω from $F_{A,A}$. The following two examples illustrate what can happen.

Example 1 *(Generic case).* Suppose for simplicity that

$$N := \#\Omega = r_n = \dim \Pi_n = \binom{n+s}{s}.$$

In this situation the set of all point configurations X such that Π_n is *the* degree reducing interpolation space is open and dense in $(\mathbb{C}^s)^{\#\Omega}$, hence,

$$\det V(X_\Omega, \Gamma_n) \ne 0$$

with probability 1. Hence, F_Ω can be reconstructed from F_n and the kernel of

$$F_{n,n+1} = \left[f(\alpha+\beta) : \begin{matrix} \alpha \in \Gamma_n \\ \beta \in \Gamma_{n+1} \end{matrix} \right]$$

determines an H-basis of I_Ω by Theorem 1. Hence, Prony's problem can be solved based on the knowledge of f on the grid

$$\{\alpha+\beta : \alpha \in \Gamma_n,\ \beta \in \Gamma_{n+1}\} = \Gamma_{2n+1}$$

since any multiindex of length $2n+1$ can be written as the sum of two multiindices, one of length n, one of length $n+1$. Hence the number of samples is r_{2n+1} and since

$$\frac{r_{2n+1}}{N} = \frac{\binom{2n+1+s}{s}}{\binom{n+s}{s}} = \frac{(2n+1+s)\cdots(2n+2)}{(n+s)\cdots(n+1)} = \prod_{j=1}^{s}\left(1+\frac{n+1}{n+j}\right) \le 2^s,$$

with 2^s being the smallest bound independent of n, it follows that the generic case needs $\approx 2^s \# \Omega$ samples of f without any *curse of dimensionality* in $\#\Omega$.

Unfortunately, not every configuration of the frequencies Ω and therefore of the points X_Ω is generic and even if any configuration could be made generic by an arbitrarily small perturbation, relying on the generic situation leads to numerical and structural problems. The second example shows that linear complexity cannot be expected in general.

Example 2 *(Hyperbola).* Let $\Omega \subset \mathbb{C}^2$ consist of $2N+1$ distinct frequencies of the form $(\omega_j, -\omega_j)$ with $\omega_j \in \mathbb{R}$, $j = 0, \dots, 2N+1$. Then the points $x_j = (e^{\omega_j}, e^{-\omega_j}) \in \mathbb{R}^2$ all lie on the hyperbola $x_1 x_2 = 1$, hence $q(x) = x_1 x_2 - 1 \in I_\Omega$, and the (unique) degree reducing interpolation space is spanned by $1, x_1, \dots, x_1^N, x_2, \dots, x_2^N$ which is a subset of Π_N. Hence, $A = \{0, \epsilon_1, \dots, N\epsilon_1, \epsilon_2, \dots, N\epsilon_2\}$ where ϵ_j stands for the unit multiindex. With $A' := A \cup \{(N+1)\epsilon_1, (N+1)\epsilon_2\}$ the minimal sampling set consists of

$$A + A' = \{\alpha + \beta : \alpha \in A, \beta \in A'\} \supset \{\alpha \in \Gamma : \|\alpha\|_\infty \le N\},$$

which has $> (N+1)^2$ elements and therefore at least $O(N^2)$ sampling points have to be used in this case.

Nevertheless, since $\#(A+A) \le (\#A)^2$ we could always have a chance to reconstruct f from $O(N^2)$ samples, independent of the dimension, provided we could find a set $A \subset \Gamma$ such that Π_A is a degree reducing interpolant. Why "degree reducing" is so important will become clear in Section 5 where we use this property to identify the ideal.

3. Universal interpolation spaces

The observations from the preceding section naturally suggest the following question.

Problem 5. *For any $N \ge 0$ determine a* minimal *set $\Upsilon_N \subset \Gamma$ such that for any set X with $\#X = N$ we find a subset $A \subset \Upsilon_N$ such that Π_A is a degree reducing interpolation space for X.*

In one variable, we know that $\Upsilon_N = \{0, \dots, N-1\}$ solves Problem 5 since the polynomials Π_{N-1} of degree at most $N-1$ form a *Chebychev system* of order N or span a *Haar space* of dimension N. Both means the same: any interpolation problem at N sites has a unique solution. Since Haar spaces only exist in one variable according to Mairhuber's theorem, cf. Lorentz (1966), a solution of Problem 5 must contain more than N elements.

Remark 1. Problem 5 is a simpler version of the classical problem of finding for any N a polynomial space of *minimal dimension* that allows for interpolation at arbitrary N points. To my knowledge, these spaces are only known for some very special cases.

In the following definition of universal interpolation spaces we define the data to be interpolated by means of polynomials. Since Π_N is always a universal interpolation space of order $N+1$, cf. Sauer (1998), this is no restriction, but more consistent when degree reducing interpolation is concerned.

Definition 4. A subspace $\mathcal{P}$ is called a *universal interpolation space* or *generalized Haar space* of order N if for any $X \subset \mathbb{C}^s$ with $\#X \le N$ and any $q \in \Pi$ there exists $p \in \mathcal{P}$ such that $p(X) = q(X)$. $\mathcal{P}$ is called a *degree reducing universal interpolation space* if the interpolant $p \in \mathcal{P}$ can be chosen such that $\deg p \le \deg q$.

Definition 5. A degree reducing universal interpolation space $\mathcal{P}$ of order N is called *redundant* if there exists a proper subspace $\mathcal{Q} \subset \mathcal{P}$ that is also a degree reducing universal interpolation space.

Lemma 6. *Any non-redundant degree reducing universal interpolation space $\mathcal{P}$ of order $N+1$ is a subspace of Π_N.*

Proof. If $\mathcal{P}$ is not a subspace of Π_N and not redundant, there must be a configuration $X \subset \mathbb{C}^s$ of N sites such that at least one of the polynomials ℓ_x defined by $\ell_x(x') = \delta_{x,x'}$, $x, x' \in X$, has degree $> N$ as otherwise $\mathcal{P}$ could be chosen as a subspace of Π_N. Let ℓ_x denote this polynomial and x the respective element of X. On the other hand, since Π_N is also a universal interpolation space there exists $\tilde{\ell}_x \in \Pi_N$ with the same interpolation property $\tilde{\ell}_x(x') = \delta_{x,x'}$, $x' \in X$, so that ℓ_x is the interpolant to $\tilde{\ell}_x$. But then $\deg \ell_x > N \ge \deg \tilde{\ell}_x$ contradicts the assumption that $\mathcal{P}$ is degree reducing. □

Returning to Problem 5, we now give an explicit non-redundant and therefore minimal monomial degree reducing interpolation space that is spanned by monomials, hence is of the form Π_A for some set $A \subset \Gamma$. This can be formalized as follows.

Definition 6. A set $A \subset \Gamma$ is called a *monomial degree reducing universal interpolation set* of order N if for any $X \subset \mathbb{C}^s$ with $\#X \le N$ the polynomial space Π_A is a degree reducing universal interpolation space.

To identify monomial degree reducing universal interpolation spaces, we need another fundamental concept.

Definition 7. For two multiindices $\alpha, \beta \in \Gamma$ we write $\alpha \le \beta$ if $\alpha_j \le \beta_j$, $j = 1, \dots, s$. We call $A \subset \Gamma$ a *lower set* if $\alpha \in A$ implies $\{\beta \in \Gamma : \beta \le \alpha\} \subset A$. By $L(\Gamma)$ we denote the set of all lower sets in Γ and

$$L_j(A) = \{B \in L(\Gamma) : \#B = j\}, \qquad j \in \mathbb{N},$$

stands for all lower sets of cardinality j.

Definition 8. For a finite set $A \subset \Gamma$ we define its *border* as

$$\partial A := \left(\bigcup_{j=1}^{s} \left(A + \epsilon_j \right) \right) \setminus A \tag{10}$$

and its *corona* as

$$\lceil A \rceil := A \cup \partial A. \tag{11}$$

We can now describe degree reducing sets of monomials in terms of lower sets.

Theorem 7. *If $A \subset \Gamma$ is a monomial degree reducing universal interpolation set of order N, then*

$$A \supseteq \bigcup_{j=1}^{N} \bigcup_{B \in L_j(\Gamma)} B. \tag{12}$$

Proof. Let B be any lower set with $\#B \le N$ and consider interpolation on the *grid*

$$X_B = \{\beta : \beta \in B\}.$$

Since A is a monomial degree reducing set, there exists a $B' \subseteq A$ such that $\Pi_{B'}$ is a degree reducing interpolation space for X_B where uniqueness of interpolation implies that $\#B' = \#B$. Since $\Pi_{B'}$ is degree reducing, the polynomials

$$(\cdot)^\beta - L_{B'}(\cdot)^\beta, \qquad \beta \in \partial B',$$

form an H-Basis of $I_{X'_B}$, see Lemma 18, and $\Pi_{B'}$ is the *normal form* or *reduced* space for X_B with respect to this H-basis and an appropriate inner product, see Lemma 19. The same holds true for B and the standard inner product, hence Π_B is also a normal form space. It has been proved in Sauer (2004) that for interpolation grids based on lower sets the normal form interpolation space is unique independently of grading and inner product, hence $\Pi_B = \Pi_{B'}$ and since both spaces are spanned by $\#B$ monomials, it follows that $B = B'$. $\square$

Proposition 8. *Any set A that satisfies (12) is a monomial degree reducing universal interpolation set of order N.*

Proof. Given $X \subset \mathbb{C}^s$, $\#X \le N$, let $\mathcal{G}$ be any reduced graded Gröbner basis for I_X and B the index set for the quotient space $\Pi_B \simeq \Pi / I_X$. Since the complement of B consists of the *upper set* leading terms of the ideal, cf. Cox et al. (1996), p. 230, it follows that B is a lower set and unique interpolation requests that $\#B = \#X$. Therefore, any interpolation problem with $\le N$ sites can be solved by an appropriate lower set B of the same cardinality. $\square$

Combining Theorem 7 with Proposition 8, we immediately get the following result.

Corollary 9. *The index set*

$$A_N^* := \bigcup_{j=1}^{N} \bigcup_{B \in L_j(\Gamma)} B \tag{13}$$

consisting of the union of all lower sets of cardinality at most N is a non-redundant, hence minimal, monomial degree reducing universal interpolation set.

The index set A_* defined in (13) can easily be described in a different way.

Lemma 10. *For any $N \ge 1$ we have*

$$\alpha \in A_N^* \qquad \Leftrightarrow \qquad \pi(\alpha) := \prod_{j=1}^{s} (\alpha_j + 1) \le N. \tag{14}$$

Proof. Since the set $\{\beta \in \Gamma : \beta \le \alpha\}$ has cardinality $(\alpha_1 + 1) \cdots (\alpha_s + 1)$, any multiindex satisfying the right hand side of (14) determines a lower subset of A_N^* of cardinality $\le N$, and therefore must belong to A_N^*. If, on the other hand $\pi(\alpha) > N$ then any lower set containing α contains $> N$ elements and cannot be a subset of A_N^*, hence $\alpha \notin A_N^*$. $\square$

Definition 9. The set $\Upsilon_N := \{\alpha \in \Gamma : \pi(\alpha) \le N\}$ is called the (positive octant of the) *hyperbolic cross* of order N.

In the sequel we will use "hyperbolic cross" for the positive octant since we only consider subsets of Γ. Hyperbolic crosses play an important role in the context of FFT methods (Döhler et al., 2009), but also in general multivariate Approximation Theory, cf. the recent survey (Dũng et al., 2015). In terms of interpolation, we can summarize the above results as follows.

Corollary 11. Υ_N *forms a minimal monomial degree reducing universal interpolation set of order N.*

Since $\#\Upsilon_N \le N \log^{s-1} N$, cf. Lubich (2008), Lemma 1.4, p. 71, the universal interpolation space based on Π_{Υ_N} allows for degree reducing interpolation at arbitrary N sites in $\mathbb{C}^s$ which can be seen as some “Haar space without uniqueness” where the number of elements only exceeds the number of points by a very moderate logarithmic factor.

Remark 2. The hyperbolic crosses Υ_N are only minimal *monomial* interpolation spaces, but not general ones. This can be easily seen in the case $N = 3$ in $s = 2$ where $\{1, x, y, x^2 + y^2\}$ forms a universal interpolation of minimal dimension 4 while $\Upsilon_3 = \{(0,0), (1,0), (2,0), (0,1), (0,2)\}$ already consists of 5 elements and therefore has dimension 5.

4. Minimal recovery

By the results of the preceding section, a simple method can be devised to solve Prony's problem, provided that $N = \#\Omega$ is given:

1. Set up the matrix
$$F = \left[f(\alpha + \beta) : \begin{array}{c} \alpha \in \Upsilon_N \\ \beta \in \lceil \Upsilon_N \rceil \end{array} \right].$$
2. Compute the kernel of this matrix and therefore the *Prony ideal* as the ideal generated by the kernel vectors, interpreted as polynomial coefficients.
3. Compute a graded Gröbner basis or an H-basis as in Möller and Sauer (2000b) and the normal form space, cf. Sauer (2006), for this ideal.
4. Determine the multiplication tables and their eigenvalues and therefore X_Ω as described in Sauer (2017).
5. Determine the coefficients by solving the Vandermonde system.

This procedure is an evaluation efficient way of solving Prony's problem where the number of function evaluations needed depends only in a very mild way on the dimension s.

Theorem 12. *Prony's problem can be solved for N frequencies in $\mathbb{C}^s$ on the basis of at most*

$$(s+1)N^2 \log^{2s-2} N$$

point evaluations on the grid Γ.

Proof. The theorem relies on two simple facts: any kernel element of F belongs the ideal I_Ω by Theorem 13, and by Lemma 18 from the appendix, the corona of Υ_N contains even an H-basis of the ideal I_Ω, hence the ideal and the quotient space can be determined from F. This allows for the reconstruction of frequencies by eigenvalues methods and coefficients by solving a simple linear system. Since the corona of a set contains at most $s + 1$ times as many elements as the set, the estimate $\#\Upsilon_N \le N \log^{s-1} N$ leads to the claim. $\square$

Remark 3. The order N^2 is optimal up to constants and logarithmic factors for reconstruction from Hankel matrices of samples from Γ. Indeed, we saw in Theorem 3 that even the coefficients can only be reconstructed from a matrix $F_{A,B}$ provided that A, B are interpolation sets for X_Ω. Without a priori information on Ω, these sets have to be universal and Υ_N is the minimal universal set, at least when degree reduction and monomiality are requested. Moreover, Example 2 showed that a complexity of N^2 is unavoidable for Hankel matrices already for $s = 2$.

Remark 4. In the generic case that happens with probability one, the complexity is even lower, namely $\sim 2^s N$, as pointed out in Example 1. Consequently, with probability one the number of variables even enters only as a constant.

Theorem 13. *Suppose that $N = \#\Omega$ and $A \subset \Gamma$. Then a vector $p \in \mathbb{C}^A$ belongs to the kernel of $F_{\Upsilon_N,A}$ if and only if the associated polynomial fulfills $p \in I_\Omega \cap \Pi_A$.*

Proof. The standard "Prony trick" yields that for any $\beta \in \Upsilon_N$ we have

$$\left(F_{\Upsilon_N,A}\, p\right)_\beta = \sum_{\alpha\in A} f(\alpha+\beta) p_\alpha = \sum_{\alpha\in A}\sum_{\omega\in\Omega} f_\omega e^{\omega^T(\alpha+\beta)}\, p_\alpha = \sum_{\omega\in\Omega} f_\omega\, e^{\omega^T\beta}\, p(x_\omega) \tag{15}$$

which we can rewrite in vector form as

$$F_{\Upsilon_N,A}\, p = V(X_\Omega, \Upsilon_N)^T F_\Omega \left[p(x_\omega) : \omega \in \Omega\right],$$

and since $\operatorname{rank} V(X_\Omega,\Upsilon_N)^T = \operatorname{rank} F_\Omega = \#\Omega$, this vector is zero if and only if p vanishes on X_Ω. □

A slightly closer inspection of the proof shows that we can reformulate Theorem 13 in even stronger form.

Corollary 14. *The equivalence*

$$p \in \ker F_{A,B} \qquad \Leftrightarrow \qquad p \in I_\Omega \cap \Pi_B$$

holds if and only if Π_A is an interpolation space for X_Ω.

By Theorem 13, the algorithm from Sauer (2017) could immediately be restated for the matrices

$$F_{n,k} := \left[f(\alpha+\beta) : \begin{matrix} \alpha \in \Upsilon_N \\ |\beta| \le k \end{matrix} \right],$$

but since this approach is based on orthogonal decompositions it requires square roots which makes it inappropriate for a symbolic environment. Therefore, the next section provides an algorithm that works in a symbolic and more "monomial" environment.

5. Symbolic algorithms

Now we are in position to turn the observations obtained so far into detailed symbolic algorithms for the reconstruction of f. The first one will be called Sparse Monomial Interpolation with Least Elements (SMILE), in contrast to the Sparse Homogeneous Interpolation Technique introduced in Sauer (2017). Both methods have in common that for $k = 0, 1, \ldots$ they *successively* compute $\Pi_A \cap \Pi_k$ and an H-basis of $I_\Omega \cap \Pi_k$ at the same time by appropriate update rules. This is more efficient than first determining some basis of the ideal, then a "good" basis (Gröbner or H-basis) and afterwards the quotient space.

We return to the function of the form (1) and assume that $N := \#\Omega$ is known. During the process, we will consider matrices of the form

$$F_k := \left[f(\alpha+\beta) : \begin{matrix} \alpha \in \Upsilon_N \\ \beta \in A_k \end{matrix} \right], \qquad A_k \subseteq \Gamma_k, \tag{16}$$

with nested sets $A_0 \subseteq A_1 \subseteq \cdots$ to be determined during the reconstruction process which will eventually terminate for some n with A_n being a monomial degree reducing set for interpolation at X_Ω. The goal is to decompose Γ_k into three sets A_k, B_k and I_k where A_k contains the exponents from $A \cap \Gamma_k$, where A is the interpolation space to be constructed eventually. I_k contains the leading powers of an H-basis of the ideal and B_k multiindices from $I_k + (\mathbb{N}_0^s \setminus \{0\})$. Since the kernel of F_k consists of the coefficient vectors from an ideal, polynomials with a leading term from B_k can be ignored in the inductive step which improves the performance of the algorithm.

The initialization is $A_0 = \{0\}$ which leaves us with

$$F_0 = \left[f(\alpha) : \alpha \in \Upsilon_N\right] \in \mathbb{C}^{\#\Upsilon_N \times 1},$$

which is $\neq 0$ since $F_0 = V(X_\Omega, \Upsilon_N)^T F_\Omega 1_N$ with $\operatorname{rank} V(X_\Omega, \Upsilon_N) = \#\Omega = N$ and $F_\Omega \neq 0$. Hence, $\operatorname{rank} F_0 = 1 = \#A_0$. Moreover, we define $I_0 = B_0 := \emptyset$ as a subset of Γ_0 and note that $\Gamma_0 = A_0 \cup I_0 \cup B_0$.

To advance from $k \to k+1$ we assume that $\operatorname{rank} F_k = \#A_k$ and $\Gamma_k = A_k \cup B_k \cup I_k$ and define the sets

$$B := \bigcup_{j=1}^{s} \left(A_k^c \cap \Gamma_k^0 \right) + \epsilon_j \subseteq \Gamma_{k+1}^0, \qquad \widetilde{A}_{k+1} := A_k \cup (\Gamma_{k+1}^0 \setminus B)$$

and extend F_k into

$$\widetilde{F}_{k+1} := \left[f(\alpha+\beta) : \begin{array}{c} \alpha \in \Upsilon_N \\ \beta \in \widetilde{A}_{k+1} \end{array} \right] = [F_k \,|\, G], \qquad G \in \mathbb{C}^{\Upsilon_N \times (\widetilde{A}_{k+1} \setminus A_k)},$$

with additional columns. Next, we compute a basis of

$$\ker \widetilde{F}_{k+1} = \left\{ y \in \mathbb{C}^{\widetilde{A}_{k+1}} \setminus \{0\} : \widetilde{F}_{k+1} y = 0 \right\}$$

and arrange it into a matrix $Y \in \mathbb{C}^{\widetilde{A}_{k+1} \times d}$ with $d := \dim \ker \widetilde{F}_{k+1} \le \#(\widetilde{A}_{k+1} \setminus A_k)$. We write

$$Y = \begin{bmatrix} Y_k \\ Y' \end{bmatrix}, \qquad Y_k \in \mathbb{C}^{A_k \times d}, \quad Y' \in \mathbb{C}^{\widetilde{A}_{k+1} \setminus A_k \times d},$$

and obtain that

$$0 = \widetilde{F}_{k+1} Y = [F_k \,|\, G] \begin{bmatrix} Y_k \\ Y' \end{bmatrix} = F_k Y_k + G Y'. \tag{17}$$

Since F_k is of maximal rank by assumption, it has a left inverse, for example the pseudoinverse F_k^+, which can be computed symbolically, cf. Springer (1987), giving $Y_k = -F_k^+ G Y'$ by (17). This implies the Schur complement relation

$$0 = \widetilde{F}_{k+1} \begin{bmatrix} -F_k^+ G \\ I \end{bmatrix} Y' = [F_k \,|\, G] \begin{bmatrix} -F_k^+ G \\ I \end{bmatrix} Y' = (I - F_k F_k^+) G Y'. \tag{18}$$

Still, $\operatorname{rank} Y' = d$, hence there exist d linear independent rows of Y' or a permutation P such that

$$Z := [I_{d\times d} \,|\, 0]\, P Y' \in \mathbb{R}^{d \times d}$$

is invertible, so that

$$P Y' Z^{-1} = \begin{bmatrix} I \\ * \end{bmatrix}.$$

After replacing Y' by $Y' Z^{-1}$ and ordering the elements of $\widetilde{A}_{k+1} \setminus A_k$ according to the permutation P, we can thus assume that $Y' = \begin{bmatrix} I_{d\times d} \\ * \end{bmatrix}$ and determine $Y_k = -F_k^* G Y'$ which of course also requires a compatible ordering of the columns of G. Now we set

$$A_{k+1} := A_k \cup \left(\widetilde{A}_{k+1} \setminus A_k \right) \left(d+1 : \#(\widetilde{A}_{k+1} \setminus A_k) \right), \tag{19}$$

$$I_{k+1} := I_k \cup \left(\widetilde{A}_{k+1} \setminus A_k \right) (1 : d), \tag{20}$$

$$B_{k+1} := B_k \cup B, \tag{21}$$

where the components of the vectors are indexed in a Matlab-like way. It follows directly from (19), (20), (21) and the assumption on A_k and I_k that $A_{k+1} \cup I_{k+1} = \Gamma_{k+1}$. For $\alpha \in \left(\widetilde{A}_{k+1} \setminus A_k \right) (1:d)$ we define polynomials $q_\alpha \in \Pi_{\widetilde{A}_k}$ whose coefficients are the respective columns of Y.

Having determined A_{k+1}, we can build the matrix F_{k+1} according to (16). If A_{k+1} is a proper superset of A_k, the matrix enlarges F_k by adding further columns which immediately yields that $\operatorname{rank} A_{k+1} \ge \operatorname{rank} A_k$. This construction also advances the rank hypothesis from k to $k+1$.

Lemma 15. *The matrix*

$$F_{k+1} = F_{\Upsilon_N, A_{k+1}} = V(\Upsilon_N, X_\Omega)^T F_\Omega V(A_{k+1}, X_\Omega) \tag{22}$$

has maximal rank $\# A_{k+1}$.

Proof. The claim is the induction hypothesis if $A_{k+1} = A_k$ and therefore trivial in this case. If $\# A_{k+1} > \# A_k$ we first note that, since $\operatorname{rank} F_k = \# A_k$ there exists a matrix

$$Z = \begin{bmatrix} * \\ 0_{\#(\widetilde{A}_{k+1} \setminus A_k) \times \# A_k} \end{bmatrix} \in \mathbb{C}^{\#\widetilde{A}_{k+1} \times \# A_k}$$

such that $\operatorname{rank} \widetilde{F}_{k+1} Z = \# A_k$. If we extend the matrix Y from above as

$$Z' = [Y \mid \hat{Y}] = \begin{bmatrix} * & * \\ I_{d \times d} & 0 \\ * & * \end{bmatrix} \in \mathbb{R}^{\#(\widetilde{A}_{k+1} \setminus A_k) \times (\widetilde{A}_{k+1} \setminus A_k)}$$

into a matrix of rank $\widetilde{A}_{k+1} \setminus A_k$, the fact that Y exactly contains the kernel of $\widetilde{F}_{k+1}$ implies that

$$\operatorname{rank} F_{k+1} [Z \mid \widehat{Y}] = \operatorname{rank} \widetilde{F}_{k+1} [Z \mid Y \mid \widehat{Y}] = \operatorname{rank} \widetilde{F}_{k+1} [Z \mid \widehat{Y}] = \# \widetilde{A}_{k+1} - d = \# A_{k+1}$$

and therefore $\operatorname{rank} F_{k+1} = \# A_{k+1}$. $\square$

This algorithm is repeated until $A_{k+1} = A_k$ and it solves Prony's problem at termination. Let us summarize the algorithm formally.

Algorithm 1 *(Prony's method, symbolic decomposition).*
Given: function $f : \Gamma \to \mathbb{C}$ and $N \geq 0$.

1. Initialization: $A_0 := \{0\}$, $I_0 := \emptyset$, $B_0 := \emptyset$ and $F_0 := [f(\alpha) : \alpha \in \Upsilon_N]$.
2. For $k = 0, 1, \dots$ repeat
   (a) Compute

   $$B := \bigcup_{j=1}^{s} \left(A_k^c \cap \Gamma_k^0 \right) + \epsilon_j$$

   set $b = \#(\Gamma_{k+1}^0 \setminus B)$ and compute

   $$G := \left[f(\alpha + \beta) : \begin{array}{c} \alpha \in \Upsilon_N \\ \beta \in \Gamma_{k+1}^0 \setminus B \end{array} \right].$$

   (b) Determine the kernel of $(I - F_k F_k^+) G$ and write it as a matrix $Y = \begin{bmatrix} I_{d \times d} \\ * \end{bmatrix}$, where d is the dimension of the kernel. This defines an ordering of $\Gamma_{k+1}^0 \setminus B$.
   (c) Set

   $$\begin{aligned} A_{k+1} &= A_k \cup (\Gamma_{k+1}^0 \setminus B)(d+1 : b), \\ I_{k+1} &= I_k \cup (\Gamma_{k+1}^0 \setminus B)(1 : d), \\ B_{k+1} &= B_k \cup B. \end{aligned}$$

   (d) Set

   $$F_{k+1} := \left[f(\alpha + \beta) : \begin{array}{c} \alpha \in \Upsilon_N \\ \beta \in A_{k+1} \end{array} \right] = \left[F_k \,\middle|\, f(\alpha + \beta) : \begin{array}{c} \alpha \in \Upsilon_N \\ \beta \in A_{k+1} \setminus A_k \end{array} \right].$$

(e) Define polynomials q_α, $\alpha \in I_{k+1} \setminus I_k$, by taking the αth column of the matrix $\begin{bmatrix} -F_k^+ GY \\ Y \end{bmatrix}$ as coefficient vectors.

until $A_{k+1} = A_k$.

Results: Monomial degree reducing interpolation space Π_{A_k} for X_Ω and H-basis $H = \left\{h_\alpha : \alpha \in I_{k+1}\right\}$ for I_Ω.

Remark 5. The intuitive meaning of (19), (20) and (21) is to split the multiindices from Γ_{k+1}^0 into three groups: A_{k+1} collects those which are used for the interpolation space, I_{k+1} those which are needed for the H-basis and B_{k+1} those which also belong to the ideal due to an H-basis element of lower degree.

Theorem 16. *If $N \geq \#\Omega$ Algorithm 1*

1. *terminates at some level $k = n \leq \#\Omega$,*
2. *determines a degree reducing interpolation set A_n for X_Ω,*
3. *determines an H-basis $H := \{h_\alpha : \alpha \in I_{n+1}\}$ for I_Ω.*

Proof. We will first verify that the polynomial set

$$H_k := \left\{(\cdot)^\beta h_\alpha : \beta \in \Pi_{k-|\alpha|}, \alpha \in I_k\right\}, \qquad k \in \mathbb{N}_0, \tag{23}$$

forms a vector space basis for $I_\Omega \cap \Pi_k$. This is trivially true as long as $I_k = \emptyset$. Since, for any $\alpha \in I_k$, we have $\deg h_\alpha = |\alpha|$ and since the coefficient vector of any h_α belongs to $\ker F_{|\alpha|-1}$, Theorem 13 ensures that $h_\alpha \in I_\Omega \cap \Pi_{|\alpha|}$. Thus, $H_k \subseteq I_\Omega \cap \Pi_k$. For the converse inclusion, we assume that $q \in \Pi_k \cap I_\Omega$ with $\deg q = k$. Again by Theorem 13, the coefficient vector of q has to belong to $\ker F_k$. By subtracting a proper element of H_k we can, like in the standard Gröbner basis division algorithm, eliminate all monomials from $\Gamma_k \setminus A_k$ from q and obtain another $q' \in I_\Omega \cap \Pi_k$ since we only subtracted ideal elements. Moreover, $q' \in \ker F_k$, but since, according to Lemma 15, F_k has rank $\#A_k$, the only polynomial $q' \in \Pi_{A_k} \cap I_\Omega$ is $q' = 0$. Hence, $q \in H_k$ and therefore $\operatorname{span} H_k \supseteq I_\Omega \cap \Pi_k$ as well, yielding $\operatorname{span} H_k = I_\Omega \cap \Pi_k$.

Since Π is a Noetherian ring, cf. Cox et al. (1996), there exists some $n \in \mathbb{N}_0$ such that the increasing chain $\langle H_k \rangle$, $k \in \mathbb{N}_0$, of ideals from (23) stabilizes, and since the strict inclusion $I_{k+1} \supset I_k$ implies $H_{k+1} \supset H_k$ in the strict sense, it follows that also $I_n = I_{n+1} = \cdots$. By (20) this means that either $d = 0$, i.e., all columns of Y correspond to ideal elements, or $B = \Gamma_{n+1}^0$. In both cases we have that $\Lambda(H_{n+1}) \cap \Pi_{n+1}^0 = \Pi_{n+1}^0$ and $A_n = A_{n+1} = \cdots$. Since $\Pi_n = A_n \oplus H_n$, it follows that Π_{A_n} is a degree reducing interpolation space, see Sauer (2006) and that

$$H := \{h_\alpha : \alpha \in I_n\}$$

is an H-basis for I_Ω. $\square$

The algorithm can also be formulated in a "term-by-term" way which gives an implicit version of the Möller–Buchberger algorithm from Möller and Buchberger (1982). To that end, we recall the classical *graded lexicographic ordering* "$\preceq$" where $\alpha \prec \beta$ provided that $|\alpha| < |\beta|$ or $|\alpha| = |\beta|$ and there exists some $k \in \{1, \ldots, s\}$ such that $\alpha_j = \beta_j$, $j = 1, \ldots, k-1$, and $\alpha_k < \beta_k$, cf. Cox et al. (1996). "$\preceq$" is a total ordering on Γ and therefore also induces a total order on the monomials or *terms* $(\cdot)^\alpha$, $\alpha \in \Gamma$. The algorithm, based on Sparse Monomial Interpolation with Least Elements (SMILE), now proceeds as follows.

Algorithm 2 *(Prony's method, SMILE).*
Given: function $f : \Gamma \to \mathbb{C}$ and $N \geq \#\Omega$.

1. Initialization: $B = \Gamma_{N+1}$, $I := \emptyset$, $A = \emptyset$, $F = [] \in \mathbb{C}^{\#\Upsilon_N \times 0}$.

2. While $B \neq \emptyset$
   (a) $\beta := \min_{\preceq} B$.
   (b) Expand the matrix by one column:
   $$\widetilde{F} = [F \mid f(\alpha+\beta) : \alpha \in \Upsilon_N].$$
   (c) If $\operatorname{rank} \widetilde{F} = \operatorname{rank} F$ then determine $0 \neq q_\beta \in \ker \widetilde{F}$ and set
   $$B := B \setminus \left(\beta + \Gamma_{N+1-|\beta|}\right),$$
   $$I := I \cup \{\beta\}.$$
   If $\operatorname{rank} \widetilde{F} > \operatorname{rank} F$ then set $F := \widetilde{F}$ and
   $$B := B \setminus \{\beta\},$$
   $$A := A \cup \{\beta\}.$$

Results: Gröbner basis $\{q_\alpha : \alpha \in I\}$ and monomial quotient space $\Pi_A \simeq \Pi / I_\Omega$.

The proof of the validity of this algorithm works like the proof of the preceding theorem, one only has to keep in mind that whenever the rank increases, a new term for the quotient space has been found which guarantees that always $\operatorname{rank} F = \#A$. If, on the other hand, the rank does not increase after adding the column, there must be a nontrivial kernel element, unique up to normalization with nonzero value in its βth component which becomes a member of the ideal basis.

Theorem 17. *Algorithm 2 computes the decomposition using at most*

$$s\,N^2 \log^{s-1} N$$

evaluations of f if $N = \#\Omega$.

Proof. Each column added needs at most $\#\Upsilon_N \leq N \log^{s-1} N$ evaluations of f. The number of columns added during the algorithm is

$$\#A + \#I \leq \#A + \#\partial A \leq \#A + (s-1)\#A = s\,\#A$$

since $I \subseteq \partial A$. $\square$

Once the set A_n and the H-basis H are determined, the points X_Ω can be determined by means of *multiplication tables* as described in Auzinger and Stetter (1988), Möller and Stetter (1995) and efficiently determined by the methods from Möller and Tenberg (2001). For reduction we can again use the inner product from Lemma 19. Once X_Ω and thus Ω are determined, the coefficients f_ω, $\omega \in \Omega$, are determined by solving a linear system, for details see Sauer (2017).

6. Sparse polynomials

A problem, closely related to Prony's problem is the reconstruction of *sparse polynomials*, i.e., of polynomials of the form

$$f(x) = \sum_{\kappa \in K} f_\kappa\, x^\kappa, \qquad f_\kappa \in \mathbb{C} \setminus \{0\}, \quad \kappa \in K,$$

where *sparsity* means that $\#K$ is (very) small relative to $\binom{\deg K + s}{s} = \Pi_{\deg K}$. This requirement is quite easy to achieve in several variables.

The "classical" method to reconstruct f from samples on Γ is the one from Ben-Or and Tiwari (1988) and uses a univariate Prony method together with divisibility aspects of relatively prime numbers. A variant with unit roots and the Chinese remainder theorem can be found in Giesbrecht et al. (2009).

As shown in Sauer (2017), it is easy to reduce this problem to Prony's problem: let $\Theta \in \mathbb{Z}^{s\times s}$ be any nonsingular matrix, then

$$f\left(2^{\Theta\alpha}\right) = \sum_{\kappa\in K} f_\kappa \, e^{\log 2\,(\Theta\kappa)^T\alpha} = \sum_{\kappa\in K} f_\kappa \, e^{\omega_\kappa^T\alpha}, \qquad \omega_\kappa := \log 2\,\Theta\kappa,$$

which is Prony's problem with $\Omega = \{\omega_\kappa : \kappa \in A\}$ which can be solved by considering the Hankel matrices

$$F_{A,B} := \left[f\left(2^{\Theta(\alpha+\beta)}\right) : \begin{array}{c} \alpha \in A \\ \beta \in B \end{array} \right].$$

If the coefficients f_κ of f belong to the *Gaussian integers* $\mathbb{Z} + i\mathbb{Z}$, which is the normal assumption in symbolic computations, the evaluations in $F_{A,B}$ are rational numbers and therefore also the ideal basis computed in the preceding section consists of *symbolic polynomials* with coefficients in $\mathbb{Q} + i\mathbb{Q}$. The same holds true for the multiplication tables and only the joint eigenvalues have to be computed in numerical precision giving the frequencies ω_κ from which the exponents can be computed as

$$\kappa = \mathrm{rd}\left(\frac{1}{\log 2}\Theta^{-1}\omega_\kappa\right)$$

by rounding to the next integer.

Acknowledgements

I want to thank the referees and the editor for their careful reading and helpful suggestions and references.

Appendix A. Two facts on interpolation spaces

This section gives a detailed exposition of some of the algebraic results used in the preceding ones. We begin by pointing out that any monomial degree reducing interpolation space automatically defines a natural H-basis.

Lemma 18. *If $A \subset \Gamma$ is a finite set such that Π_A is a degree reducing interpolation space for X then the polynomials $q_\alpha := (\cdot)^\alpha - L_A(\cdot)^\alpha$, $\alpha \in \partial A$, form an H-basis of I_X.*

Proof. Define

$$q_\alpha := (\cdot)^\alpha - L_A(\cdot)^\alpha, \qquad \alpha \in \Gamma, \tag{24}$$

and note that $q_\alpha = 0$ for $\alpha \in A$ and $\deg q_\alpha = |\alpha|$ since A is degree reducing. Moreover, $\frac{\partial^\beta q_\alpha}{\partial x^\beta}(0) = \alpha!\delta_{\alpha,\beta}$, $\alpha, \beta \in A^c := \Gamma \setminus A$. Therefore the polynomials q_α, $\alpha \in \Gamma_n \setminus A$, and $(\cdot)^\alpha$, $\alpha \in A_n := A \cap \Gamma_n$ form a basis of Π_n for any $n \in \mathbb{N}$. Consequently, any polynomial $p = \sum p_\alpha (\cdot)^\alpha \in \Pi$ can be written as

$$p(x) = L_A p(x) + p(x) - L_A p(x) = L_A p(x) + \sum_{\alpha\in A^c} p_\alpha \, q_\alpha(x)$$

and

$$p \in I_X \qquad \Leftrightarrow \qquad p(x) = \sum_{\alpha\in A^c} p_\alpha \, q_\alpha(x). \tag{25}$$

The representation on the right hand side of (25) is an H-representation (Möller and Sauer, 2000a), hence the polynomials $\{q_\alpha : \alpha \in A^c\}$, form an infinite H-basis, and $\left\{q_\alpha : \alpha \in A^c_{n+1}\right\}$, $A^c_{n+1} := A^c \cap \Gamma_{n+1}$, where $n := \deg A = \max\{|\alpha| : \alpha \in A\}$, is a finite H-basis of I_X.

Next, we fix $j \in \{1,\dots,s\}$ and $\alpha \in A^c$, and write $p(x) := L_A(\cdot)^\alpha(x) \in \Pi_A$ as $p(x) = \sum p_\beta x^\beta$. Then,

$$\begin{aligned}
q_{\alpha+\epsilon_j}(x) - x_j q_\alpha(x) &= x^{\alpha+\epsilon_j} - L_A(\cdot)^{\alpha+\epsilon_j}(x) - x^{\alpha+\epsilon_j} + x_j L_A(\cdot)^\alpha(x) \\
&= x_j L_A(\cdot)^\alpha(x) - L_A(\cdot)^{\alpha+\epsilon_j}(x) = \sum_{\beta\in A} p_\beta x^{\beta+\epsilon_j} - L_A(\cdot)^{\alpha+\epsilon_j}(x) \\
&= \sum_{\beta\in\partial A} p_{\beta-\epsilon_j} x^\beta + \sum_{\beta\in A\cap(A+\epsilon_j)} p_{\beta-\epsilon_j} x^\beta - L_A(\cdot)^{\alpha+\epsilon_j}(x) \\
&= \sum_{\beta\in\partial A} p_{\beta-\epsilon_j} q_\beta(x) + \sum_{\beta\in\partial A} p_{\beta-\epsilon_j} L_A(\cdot)^\beta(x) + \sum_{\beta\in A\cap(A+\epsilon_j)} p_{\beta-\epsilon_j} x^\beta - L_A(\cdot)^{\alpha+\epsilon_j}(x) \\
&= \sum_{\beta\in\partial A} p_{\beta-\epsilon_j} q_\beta(x) + \tilde p(x)
\end{aligned}$$

with some $\tilde p \in \Pi_A$. The polynomial on the left hand side belongs to the ideal and vanishes on X as do the q_β in the sum on the right hand side, hence $\tilde p(X) = 0$ and therefore, taking account on the lengths of the β appearing in the above decomposition,

$$q_{\alpha+\epsilon_j}(x) - x_j q_\alpha(x) = \sum_{\beta\in\partial A\cap\Gamma_{n+1}} c_{\alpha,\beta}\, q_\beta(x), \qquad c_{\alpha,\beta} \in \mathbb{C}. \tag{26}$$

Therefore, any q_α with $\alpha \in A^c$ such that $\alpha - \epsilon_j \in A^c$ has a H-representation by $(\cdot)_j q_{\alpha-\epsilon_j}$ and q_β, $\alpha \in \partial A \cap \Gamma_n$. If, on the other hand, $\alpha - \epsilon_j \notin A^c$, $j = 1,\dots,s$, and $\alpha \neq 0$, then there must be some j such that $\alpha - \epsilon_j \in A$ and therefore $\alpha \in \partial A$. Since any nontrivial degree reducing set must contain 0, it follows that $0 \notin A^c$ and therefore an inductive application of the above process shows that any q_α must eventually be written as a linear combination of q_β, $\beta \in \partial A \cap \Gamma_{|\alpha|}$. $\square$

We recall the notion of a reduced polynomial. Given an inner product $(\cdot,\cdot)$ on Π, we call a polynomial p *reduced* if each homogeneous term

$$p_j(x) := \sum_{|\alpha|=j} p_\alpha x^\alpha, \qquad j = 0,\dots,\deg p,$$

of p is perpendicular to the homogeneous leading forms in $\Lambda(I_X) \cap \Pi_j^0$, where $\Lambda(p) := p_{\deg p} \in \Pi^0_{\deg p}$. As shown in Sauer (2001) that whenever H is an H-basis for I_X there exists, for any polynomial $p \in \Pi$, a decomposition

$$p = \sum_{h\in H} q_h\, h + r, \qquad \deg q_h + \deg h \le \deg p, \tag{27}$$

such that r is reduced and depends only on I_X and $(\cdot,\cdot)$ and is zero if and only if $p \in I_X$. Therefore, r can be seen as a well defined mapping $r : \Pi \to \Pi$. Also note that (27) is the multivariate analog of *euclidean division* or *division with remainder* and that r is the natural interpolant of p.

Lemma 19. *If $A \subset \Gamma$ is a finite set such that Π_A is a degree reducing interpolation space for X then there exists an inner product $(\cdot,\cdot)$ such that $\Pi_A = r(\Pi)$.*

Proof. We use the H-basis q_α, $\alpha \in A^c$, defined in (24) and define the inner product separately on $\Pi_n^0 \times \Pi_n^0$, $n \in \{0,1,\dots\}$. If $n < \min\{|\alpha| : \alpha \in A^c\}$ and $n > \deg A$, we simply use the inner product of the coefficients,

$$(p,p')_n := \sum_{|\alpha|=n} \overline{p_\alpha}\, p'_\alpha, \qquad p, p' \in \Pi_n^0.$$

For other values of n we first observe that the polynomials x^α, $\alpha \in A \cap \Gamma_n^0$ and the leading forms $\Lambda(q_\alpha)$, $\alpha \in A^c \cap \Gamma_n^0$, span Π_n^0. We arrange the coefficient vectors into a nonsingular matrix $Y \in \mathbb{C}^{r_n^0 \times r_n^0}$ where $r_n^0 := \dim \Pi_n^0 = \binom{n+s-1}{s-1}$ and note that the Gramian $G := YY^H$ is hermitian and positive definite. Defining

$$(p, p')_n = p^H G^{-1} p = \sum_{|\alpha|=|\beta|=n} (G^{-1})_{\alpha,\beta} p_\alpha p'_\beta, \qquad p, p' \in \Pi_n^0,$$

we get that

$$(Y, Y)_n = Y^H G^{-1} Y = Y^H (YY^H)^{-1} Y = I$$

which means that the vectors e_α, $\alpha \in A \cap \Gamma_n^0$ are perpendicular to the coefficient vectors of $\Lambda(q_\alpha)$, $\alpha \in A^c \cap \Gamma_n^0$. Consequently, the inner product

$$(p, p') = \sum_{j \in \mathbb{N}_0} (p_j, p'_j)_j, \qquad p, p' \in \Pi,$$

has the property that a polynomial is reduced if and only if it belongs to Π_A, that is, $\Pi_A = r(\Pi)$ as claimed. $\square$

References

Auzinger, W., Stetter, H.J., 1988. An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations. In: Numerical Mathematics. Singapore, 1988. In: Internat. Schriftenreihe Numer. Math., vol. 86. Birkhäuser, Basel, pp. 11–30.

Ben-Or, M., Tiwari, P., 1988. A deterministic algorithm for sparse multivariate polynomial interpolation. In: Proc. Twentieth Annual ACM Symp. Theory Comput. ACM Press, New York, pp. 301–309.

Boor, C.d., 2007. Interpolation from spaces spanned by monomials. Adv. Comput. Math. 26, 63–70.

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

Diederichs, B., Iske, A., 2015. Parameter estimation for bivariate exponential sums. In: IEEE International Conference Sampling Theory and Applications, pp. 493–497.

Döhler, M., Kunis, S., Potts, D., 2009. Nonequispaced hyperbolic cross fast Fourier transform. SIAM J. Numer. Anal. 47, 4415–4428.

Dũng, D., Temlyakov, V.N., Ullrich, T., 2015. arXiv:1601.03978.

Giesbrecht, M., Labahn, G., Lee, W., 2009. Symbolic–numeric sparse interpolation of multivariate polynomials. J. Symb. Comput. 44, 943–959.

Gonzales-Vega, L., Rouillier, F., Roy, M.-F., Trujillo, G., 1999. Symbolic recipes for real solution. In: Cohen, A.M., Cuypers, H., Sterk, M. (Eds.), Some Tapas in Computer Algebra. In: Algorithms and Computations in Mathematics, vol. 4. Springer, pp. 121–162. Ch. 2.

Hua, Y., 1992. Estimating two-dimensional frequencies by matrix enhancement and matrix pencil. IEEE Trans. Signal Process. 40, 2267–2279.

Hua, Y., Sarkar, T.K., 1990. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. IEEE Trans. Acoust. Speech Signal Process. 38, 814–824.

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

Kunis, S., Peter, T., Römer, T., von der Ohe, U., 2016. A multivariate generalization of Prony's method. Linear Algebra Appl. 490, 31–47.

Lorentz, G.G., 1966. Approximation of Functions. Chelsea Publishing Company.

Lubich, C., 2008. From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis. European Mathematical Society.

Möller, H.M., Buchberger, B., 1982. The construction of multivariate polynomials with preassigned zeros. In: Goos, G., Hartmanis, J. (Eds.), Computer Algebra, EUROCAM '82, European Computer Algebra Conference. In: Lecture Notes in Computer Science, vol. 144. Springer-Verlag, pp. 24–31.

Möller, H.M., Sauer, T., 2000a. H-bases for polynomial interpolation and system solving. Adv. Comput. Math. 12 (4), 335–362.

Möller, H.M., Sauer, T., 2000b. H-bases I: the foundation. In: Cohen, A., Rabut, C., Schumaker, L.L. (Eds.), Curve and Surface fitting. Saint-Malo, 1999. Vanderbilt University Press, pp. 325–332.

Möller, H.M., Stetter, H.J., 1995. Multivariate polynomial equations with multiple zeros solved by matrix eigenproblems. Numer. Math. 70, 311–329.

Möller, H.M., Tenberg, R., 2001. Multivariate polynomial system solving using intersections of eigenspaces. J. Symb. Comput. 32, 513–531.

Plonka, G., Tasche, M., 2014. Prony methods for recovery of structured functions. GAMM-Mitt. 37, 239–258.

Potts, D., Tasche, M., 2013. Parameter estimation for multivariate exponential sums. Electron. Trans. Numer. Anal. 40, 204–224.

Potts, D., Tasche, M., 2015. Fast ESPRIT algorithms based on partial singular value decompositions. Appl. Numer. Math. 88, 31–45.

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.

Rouquette, S., Najim, M., 2001. Estimation of frequencies and damping factors by two-dimensional esprit type methods. IEEE Trans. Signal Process. 49, 237–245.

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., 1997. Polynomial interpolation of minimal degree. Numer. Math. 78 (1), 59–85.

Sauer, T., 1998. Polynomial interpolation of minimal degree and Gröbner bases. In: Buchberger, B., Winkler, F. (Eds.), Groebner Bases and Applications, Proc. of the Conf. 33 Years of Groebner Bases. In: London Math. Soc. Lecture Notes, vol. 251. Cambridge University Press, pp. 483–494.

Sauer, T., 2001. Gröbner bases, H-bases and interpolation. Trans. Am. Math. Soc. 353, 2293–2308.

Sauer, T., 2004. Lagrange interpolation on subgrids of tensor product grids. Math. Comput. 73, 181–190.

Sauer, T., 2006. Polynomial interpolation in several variables: lattices, differences, and ideals. In: Buhmann, M., Hausmann, W., Jetter, K., Schaback, W., Stöckler, J. (Eds.), Multivariate Approximation and Interpolation. Elsevier, pp. 189–228.

Sauer, T., 2017. Prony's method in several variables. Numer. Math. http://dx.doi.org/10.1007/s00211-016-0844-8. In press. arXiv:1602.02352.

Schmidt, R., 1986. Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propag. 34, 276–280.

Springer, J., 1987. Generalized inverses of integer matrix. Z. Angew. Math. Mech. 67, 503–506.

Stetter, H.J., 1995. Matrix eigenproblems at the heart of polynomial system solving. SIGSAM Bull. 30 (4), 22–25.

Yilmazer, N., Fernandez-Recio, R., Sarkar, T.K., 2006. Matrix pencil method for simultaneously estimating azimuth and elevation angles of arrival along with the frequency of the incoming signals. Digit. Signal Process. 16, 796–816.

Zippel, R., 1979. Probabilistic algorithms for sparse polynomials. PhD thesis. MIT.

Zippel, R., 1990. Interpolating polynomials from their values. J. Symb. Comput. 9, 375–403.