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.

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.

Keywords: Prony’s method, Gröbner basis, degree reducing interpolation, hyperbolic cross

2000 MSC: 13P10, 65T40, 41D05

1. Introduction

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

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

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 (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).

Email address: Tomas.Sauer@uni-passau.de (Tomas Sauer)

Preprint submitted to Elsevier

March 14, 2017
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 consists of considering finite parts of the infinite Han-
kel matrix

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

(2)

for example

\[ F_n := \left[ f(\alpha + \beta) : \alpha, \beta \in \Gamma_n \right], \quad \Gamma_n := \left\{ \alpha \in \mathbb{N}_0^s : \|\alpha\| \leq n \right\}, \]

(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 \( I_\Omega \), the set of all polynomials vanishing on

\[ X_\Omega = e^\Omega = \left\{ x_\omega = e^\omega = (e^{\omega_1}, \ldots, e^{\omega_s}) : \omega \in \Omega \right\}, \]

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

\[ \Pi_n \ni p(x) = \sum_{\|\alpha\|\leq n} p_\alpha x^\alpha \]

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 \( X_\Omega \) and since all degree reducing interpolation spaces have the
same degree, this is indeed a geometric quantity depending on \( X_\Omega \) 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

\[ F_{n,k} := \left[ f(\alpha + \beta) : \alpha \in \Gamma_n, \beta \in \Gamma_k \right], \quad k = 0, \ldots, n, \]

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

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

1. \( \ker F_{n,k} \cong I_\Omega \cap \Pi_k, \ k = 0, \ldots, n \),
2. \( k \mapsto \text{rank } F_{n,k} \) computes the affine Hilbert function for the ideal \( I_\Omega \),
3. there exists a number \( 0 \leq m \leq n \) such that

\[ \text{rank } F_{n,0} < \cdots < \text{rank } F_{n,m-1} = \text{rank } F_{n,m} = \cdots = \text{rank } F_{n,m} = \text{rank } F_n, \]

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 \( \Pi/I_\Omega \) 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 \( x_\omega \) 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 insight 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 an appendix in Section 7 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 \( \Pi = \mathbb{C}[x_1, \ldots, x_s] \) 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 \( \Pi \) 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 \( A = \Gamma_n = \{ \alpha : |\alpha| \leq n \} \), 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 \( \Pi_n^0 := \Pi_n \) as the homogeneous polynomials of degree (exactly) \( n \).

The coefficients \( p_\omega \) 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 \( X \subset \mathbb{C}^\ast \) we define the Vandermonde matrix \( V(X, A) \) as

\[
V(X, A) := \begin{bmatrix} x^\alpha : & x \in X \\
\alpha \in A & \end{bmatrix}
\]

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

\[
F_{A, B} := \begin{bmatrix} f(\alpha + \beta) : & \alpha \in A \\
\beta \in B & \end{bmatrix}.
\]

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

\[
(F_{A, B})_{\alpha, \beta} = e^T_\alpha F_{A, B} e_\beta = \sum_{\omega \in \Omega} f_\omega e^{i\omega(\alpha+\beta)} = \sum_{\omega \in \Omega} f_\omega e^{i\omega\alpha} e^{i\omega\beta}
\]

\[
= \sum_{\omega \in \Omega} e^T_\omega V(X_\Omega, A) e_\omega f_\omega e^T_\omega V(X_\Omega, B) e_\beta,
\]

which yields the well-known factorization

\[
F_{A, B} = V(X_\Omega, A)^T F_\Omega V(X_\Omega, B), \quad F_\Omega := \text{diag}(f_\omega : \omega \in \Omega),
\]

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

\[
\text{rank } V(X_\Omega, A) = \text{rank } V(X_\Omega, B) = \#\Omega.
\]

The meaning of (6) is well-known: \( \Pi_A \) and \( \Pi_B \) have to be interpolation spaces for \( X_\Omega \).

**Definition 2.** A subspace \( \mathcal{P} \) of \( \Pi \) 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 \( \Omega \) from \( F_{A, B} \)

**Theorem 3.** The coefficients \( f_\omega \) can be reconstructed from \( F_{A, B} \) if and only if \( \Pi_A \) and \( \Pi_B \) are interpolation spaces with respect to \( X_\Omega \).

**Proof.** Suppose that \( \Pi_A \) and \( \Pi_B \) are interpolation spaces with respect to \( X_\Omega \), 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
\[ V(\Omega, \mathcal{A}) \begin{bmatrix} p_{\alpha, \omega} : \alpha \in \mathcal{A} \\ \omega \in \Omega \end{bmatrix} = I_{\#\Omega}. \]

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

which reconstructs the coefficients under the assumption that \( \Pi_{\mathcal{A}} \) and \( \Pi_{\mathcal{B}} \) are interpolation spaces for \( X_{\Omega} \).

Conversely, if \( \operatorname{rank} V(\Omega, \mathcal{B}) < \#\Omega \), then there exists a nonzero diagonal matrix \( F \in \mathbb{C}^{\#\Omega \times \#\Omega} \) such that
\[ F V(\Omega, \mathcal{B}) = 0 \]
and therefore
\[ F_{\mathcal{A}, \mathcal{B}} = V(\Omega, \mathcal{A})^{T} (F_{\Omega} + F) V(\Omega, \mathcal{B}) \]
so that \( F_{\Omega} \) cannot be reconstructed from \( F_{\mathcal{A}, \mathcal{B}} \). An analogous argument can also be used in the case that \( \operatorname{rank} V(\Omega, \mathcal{A}) < \#\Omega \).

**Corollary 4.** Any sampling sets \( \mathcal{A}, \mathcal{B} \) for Prony’s method must be chosen such that \( \Pi_{\mathcal{A}} \) and \( \Pi_{\mathcal{B}} \) are interpolation spaces for \( X_{\Omega} \).

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

Degree reducing interpolation spaces for some set \( \Gamma \subset \mathbb{C}^s \) of sites have the advantage that they give the ideal \( I_{\Gamma} = \{ p \in \Pi : p(\Gamma) = 0 \} \) almost for free. To be more concrete, let \( \mathcal{A} \subset \Gamma \) be such that \( \Pi_{\mathcal{A}} \) is a degree reducing interpolation space for \( X \) and let \( L_{\mathcal{A}} : \Pi \to \Pi_{\mathcal{A}} \) denote the interpolation operator defined by
\[ (L_{\mathcal{A}} p)(X) = p(X) \quad (7) \]
which is well defined because degree reducing interpolation is unique by definition. It can then be shown that the polynomials
\[ h_{\alpha} := (\cdot)^{\alpha} - L_{\mathcal{A}}(\cdot)^{\alpha}, \quad \alpha \in \mathcal{A}^c := \Gamma \setminus \mathcal{A}, \quad (8) \]
form an \( H \)-basis of \( I_{\Gamma} = \{ q : q(\Gamma) = 0 \} \), that is, any polynomial \( q \in I_{\Gamma} \) can be written as
\[ q = \sum_{\alpha \in \mathcal{A}^c} q_{\alpha} h_{\alpha}, \quad \deg q_{\alpha} \leq \deg q - \deg h_{\alpha}, \quad (9) \]
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 \( F_{\Omega} \) from the symmetric sampling matrix \( F_{\mathcal{A}, \mathcal{A}} \) if and only if \( \Pi_{\mathcal{A}} \) is an interpolation space. \( \mathcal{A} \) can be chosen minimally or at least of minimal cardinality by requesting \( \Pi_{\mathcal{A}} \) 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 \( X_{\Omega} \) from \( F_{\mathcal{A}, \mathcal{A}} \). 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 \( \Pi_n \) is the degree reducing interpolation space is open and dense in \( (C^\infty)^\#\Omega \), hence,

\[ \det V(X_{\Omega}, \Gamma_n) \neq 0 \]

with probability 1. Hence, \( F_\Omega \) can be reconstructed from \( F_n \) and the kernel of

\[ F_{n,n+1} = \begin{bmatrix} f(\alpha + \beta) : \alpha \in \Gamma_n & \beta \in \Gamma_{n+1} \end{bmatrix} \]

determines an \( \mathbb{H}\)-basis of \( I_\Omega \) by Theorem \[\text{[1]}\]. 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

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

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 \( \Omega \) and therefore of the points \( X_{\Omega} \) 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 \( \omega_j \in \mathbb{R}, j = 0, \ldots, 2N+1 \). 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 \( g(x) = x_1 x_2 - 1 \in \mathbb{I}_\Omega \), and the (unique) degree reducing interpolation space is spanned by \( 1, x_1, \ldots, x^N_1, x_2, \ldots, x^N_2 \) which is a subset of \( \Pi_N \). Hence, \( A = \{0, \epsilon_1, \ldots, N \epsilon_1, \epsilon_2, \ldots, N \epsilon_2\} \) where \( \epsilon_j \) stands for the unit multiindex. With \( A' := A \cup (N+1) \epsilon_1, (N+1) \epsilon_2 \) the minimal sampling set consists of

\[ A + A' = \{ \alpha + \beta : \alpha \in A, \beta \in A' \} \supset \{ \alpha \in \Gamma : \|\alpha\|_\infty \leq N \}, \]

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

Nevertheless, since \( \#(A + A) \leq (\#A)^2 \) 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 \( \Pi_A \) is a degree reducing interpolant. Why “degree reducing” is so important will become clear in Section \[\text{[5]}\] 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 \( N \geq 0 \) 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 \( \Pi_A \) is a degree reducing interpolation space for \( X \).

In one variable, we know that \( \Upsilon_N = \{0, \ldots, N-1\} \) solves Problem 5 since the polynomials \( \Pi_{N-1} \) 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 \( \Pi_N \) 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 \( \#X \leq N \) 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 \( \deg p \leq \deg q \).

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

Proof. If \( \mathcal{P} \) is not a subspace of \( \Pi_N \) and not redundant, there must be a configuration \( X \subset \mathbb{C}^s \) of \( N \) sites such that at least one of the polynomials \( \ell_x \) 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 \( \Pi_N \). Let \( \ell_x \) denote this polynomial and \( x \) the respective element of \( X \). On the other hand, since \( \Pi_N \) 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 \( \ell_x \) is the interpolant to \( \tilde{\ell}_x \). But then \( \deg \ell_x > N \geq \deg \tilde{\ell}_x \) 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 \( \Pi_A \) 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 \( \#X \leq N \) the polynomial space \( \Pi_A \) 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 \( \alpha \leq \beta \) if \( \alpha_j \leq \beta_j \), \( j = 1, \ldots, s \). We call \( A \subset \Gamma \) a lower set if \( \alpha \in A \) implies \( \{ \beta \in \Gamma : \beta \leq \alpha \} \subset A \). By \( L(\Gamma) \) we denote the set of all lower sets in \( \Gamma \) and

\[
L_j(A) = \{ B \in L(\Gamma) : \#B = j \}, \quad j \in \mathbb{N},
\]

stands for all lower sets of cardinality \( j \).

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

\[
\partial A := \left( \bigcup_{j=1}^{s} (A + \epsilon_j) \right) \setminus A
\]

and its corona as

\[
[A] := A \cup \partial A.
\]

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

\[
A \supseteq \bigcup_{j=1}^{N} \bigcup_{B \in L_j(\Gamma)} B.
\]

**Proof.** Let \( B \) be any lower set with \( \#B \leq N \) 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

\[
(\cdot)^\beta - \mathcal{L}B(\cdot)^\beta, \quad \beta \in \partial B',
\]

form an H–Basis of \( I_X \), 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 \( \Pi_B \) is also a normal form space. It has been proved in \( \text{[Sauer, 2004]} \) 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 \( B = B' \).

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

**Proof.** Given \( X \subset \mathbb{C}^s \), \( \#X \leq N \), let \( \mathcal{G} \) be any reduced graded Gröbner basis for \( I_X \) and \( B \) the index set for the quotient space \( \Pi_B \cong \Pi/I_X \). Since the complement of \( B \) consists of the upper set leading terms of the ideal, cf. \( \text{[Cox et al., 1996]} \) p. 230), it follows that \( B \) is a lower set and unique interpolation requests that \( \#B = \#X \). Therefore, any interpolation problem with \( \leq N \) sites can be solved by an appropriate lower set \( B \) of the same cardinality.
Combining Theorem 7 with Proposition 8, we immediately get the following result.

**Corollary 9.** The index set

\[ A_N^* := \bigcup_{j=1}^{N} \bigcup_{B \in L_j(\Gamma)} B \]  

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 \( N \geq 1 \) we have

\[ \alpha \in A_N^* \iff \pi(\alpha) := \prod_{j=1}^{r}(\alpha_j + 1) \leq N. \]  

**Proof.** Since the set \( \{\beta \in \Gamma : \beta \leq \alpha\} \) 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 \( \leq N \), and therefore must belong to \( A_N^* \). If, on the other hand \( \pi(\alpha) > N \) then any lower set containing \( \alpha \) contains \( > N \) elements and cannot be a subset of \( A_N^* \), hence \( \alpha \notin A_N^* \).

**Definition 9.** The set \( \Upsilon_N := \{\alpha \in \Gamma : \pi(\alpha) \leq N\} \) 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 \( \Gamma \). Hyperbolic crosses play an important role in the context of FFT methods (Döhrer et al., 2009), but also in general multivariate Approximation Theory, cf. the recent survey (Dürr et al., 2015). In terms of interpolation, we can summarize the above results as follows.

**Corollary 11.** \( \Upsilon_N \) forms a minimal monomial degree reducing universal interpolation set of order \( N \).

Since \( \#\Upsilon_N \leq N \log^{s-1} N \), cf. (Lubich, 2008, Lemma 1.4, p. 71), the universal interpolation space based on \( \Pi_{\Upsilon_N} \) 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 \( \Upsilon_N \) 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 \( \Upsilon_2 = \{(0, 0), (1, 0), (2, 0), (0, 1), (0, 2)\} \) 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

\[ F = \left[ f(\alpha + \beta) : \frac{\alpha \in \Upsilon_N}{\beta \in \lceil \Upsilon_N \rceil} \right]. \]
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 $X_\Omega$ 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 $(s + 1)N^2 \log^{2r-2} N$ point evaluations on the grid $\Gamma$.

**Proof.** The theorem relies on two simple facts: any kernel element of $F$ belongs the ideal $I_\Omega$ by Theorem [13], and by Lemma [18] from the appendix, the corona of $\gamma_N$ contains even an H–basis of the ideal $I_\Omega$, 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 $\#\gamma_N \leq N \log^{s-1} N$ leads to the claim.

**Remark 3.** The order $N^2$ is optimal up to constants and logarithmic factors for reconstruction from Hankel matrices of samples from $\Gamma$. 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 $X_\Omega$. Without a priori information on $\Omega$, these sets have to be universal and $\gamma_N$ 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 $F_{\gamma_N,A}$ if and only if the associated polynomial fulfills $p \in I_\Omega \cap \Pi_A$.

**Proof.** The standard “Prony trick” yields that for any $\beta \in \gamma_N$ we have

\[
(F_{\gamma_N,A} p)_\beta = \sum_{\alpha \in A} f(\alpha + \beta) p_\alpha = \sum_{\alpha \in A} \sum_{\omega \in \Omega} f_\omega e^{i\omega(\alpha + \beta)} p_\alpha = \sum_{\omega \in \Omega} f_\omega e^{i\omega\beta} p(x_\omega) \tag{15}
\]

which we can rewrite in vector form as

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

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

A slightly closer inspection of the proof shows that we can reformulate Theorem 13 in even stronger form.
Corollary 14. The equivalence

\[ p \in \ker F_{A,B} \iff p \in I_\Omega \cap \Pi_B \]

holds if and only if \( \Pi_A \) is an interpolation space for \( X_\Omega \).

By Theorem 13, the algorithm from [Sauer, 2017] could immediately be restated for the matrices

\[ F_{n,k} := \left\{ f(\alpha + \beta) : \alpha \in \Upsilon_N, \beta \in A_k, |\beta| \leq k \right\}, \]

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

\[ F_k := \left\{ f(\alpha + \beta) : \alpha \in \Upsilon_N, \beta \in A_k, |\beta| \leq k \right\}, \quad A_k \subseteq \Gamma_k, \]

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 \( X_\Omega \). The goal is to decompose \( \Gamma_k \) 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 \( I_k + (\mathbb{N}_0^n \setminus \{0\}) \). 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

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

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

To advance from \( k \rightarrow k + 1 \) we assume that rank \( F_k = \#A_k \) and \( \Gamma_k = A_k \cup B_k \cup I_k \) and define the sets

\[ B := \bigcup_{j=1}^{k} (A_j^+ \cap \Gamma_j^0), \quad \epsilon_j \subseteq \Gamma_{k+1}^0, \quad \tilde{A}_{k+1} := A_k \cup (\Gamma_{k+1}^0 \setminus B) \]
and extend $F_k$ into

$$F_{k+1} := \left[ f(\alpha + \beta) : \begin{array}{c} \alpha \in Y \\\ \beta \in \tilde{A}_{k+1} \end{array} \right] = [F_k | G], \quad G \in \mathbb{C}^{Y_k \times (\tilde{A}_{k+1} \setminus \tilde{A}_k)},$$

with additional columns. Next, we compute a basis of

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

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

$$Y = \left[ \begin{array}{c} Y_k \\ Y' \end{array} \right], \quad Y_k \in \mathbb{C}^{\tilde{A}_{k+1} \times d}, \quad Y' \in \mathbb{C}^{\tilde{A}_{k+1} \setminus A_k \times d},$$

and obtain that

$$0 = \tilde{F}_{k+1} Y = [F_k | G] \left[ \begin{array}{c} Y_k \\ Y' \end{array} \right] = F_k Y_k + G Y'.$$  \hspace{1cm} (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

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

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

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

Still, 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

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

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

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

$$I_{k+1} := I_k \cup (\tilde{A}_{k+1} \setminus A_k) (1 : d),$$  \hspace{1cm} (20)

$$B_{k+1} := B_k \cup B,$$  \hspace{1cm} (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 $\alpha \in (\tilde{A}_{k+1} \setminus A_k) (1 : d)$ we define polynomials $q_\alpha \in \Pi_{A_k}$ 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 rank $A_{k+1} \geq$ rank $A_k$. This construction also advances the rank hypothesis from $k$ to $k + 1$. 

12
Lemma 15. The matrix

\[ F_{k+1} = F_{Y_k,A_k+1} = V(\gamma_N, X_\Omega)^T F_\Omega V(A_{k+1}, X_\Omega) \]  

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 rank \( F_k = \#A_k \) there exists a matrix

\[ Z = \begin{bmatrix} \ast & \ast & \ast \\ 0_{\#(\tilde{A}_{k+1}\setminus A_k) \times \#A_k} & \ast & \ast \end{bmatrix} \in \mathbb{C}^{\#\tilde{A}_{k+1} \times \#A_k} \]

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

\[ Z' = [Y \ | \ \tilde{Y}] = \begin{bmatrix} \ast & \ast & \ast \\ I_{\#d} & 0 & \ast \\ \ast & \ast & \ast \end{bmatrix} \in \mathbb{R}^{\#(\tilde{A}_{k+1}\setminus A_k) \times (\tilde{A}_{k+1}\setminus A_k)} \]

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

\[ \text{rank } F_{k+1}[Z \ | \ \tilde{Y}] = \text{rank } \tilde{F}_{k+1}[Z \ | \ \tilde{Y}] = \text{rank } \tilde{F}_{k+1}[Z \ | \ \tilde{Y}] = \#\tilde{A}_{k+1} - d = \#A_{k+1} \]

and therefore rank \( F_{k+1} = \#A_{k+1} \).

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 \( F_0 := [f(\alpha) : \alpha \in \gamma_N] \).

2. For \( k = 0, 1, \ldots \) repeat

   (a) Compute

   \[ B := \bigcup_{j=1}^s (A_j^c \cap \Gamma_0^j) + \epsilon_j \]

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

   \[ G := \left[ f(\alpha + \beta) : \alpha \in \gamma_N, \beta \in \bigcap_{k=1}^0 \Gamma_{k+1} \setminus B \right]. \]

   (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 \( \bigcap_{k=1}^0 \Gamma_{k+1} \setminus B \).

   (c) Set

   \[
   
   \begin{align*}
   \Gamma_{k+1} &= A_k \cup (\Gamma_0^k \setminus B)(d+1:b), \\
   I_{k+1} &= I_k \cup (\Gamma_0^k \setminus B)(1:d), \\
   B_{k+1} &= B_k \cup B.
   \end{align*}
   \]

13
(d) Set
\[ F_{k+1} := \left[ f(\alpha + \beta) : \alpha \in \mathcal{T}_N^k, \beta \in A_{k+1} \right] \]

(e) Define polynomials \( q_\alpha, \alpha \in I_{k+1} \backslash I_k \), by taking the \( \alpha \)th column of the matrix
\[
\begin{bmatrix}
-F^*GY
\end{bmatrix}
\]
as coefficient vectors.

\[ \text{until } A_{k+1} = A_k. \]

**Results:** Monomial degree reducing interpolation space \( \Pi_{\Lambda_k} \) for \( X_\Omega \) and \( \Pi \)-basis \( H = \{ h_\alpha : \alpha \in I_{k+1} \} \) for \( I_\Omega \).

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

**Theorem 16.** If \( N \geq \#\Omega \), Algorithm 7

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

**Proof.** We will first verify that the polynomial set
\[ H_k := \left\{ \gamma^h h_\alpha : \beta \in \Pi_k \backslash \pi_0, \alpha \in I_k \right\}, \quad k \in \mathbb{N}_0, \]

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 \( h_\alpha \) belongs to \( \ker F_{|\alpha|} \), Theorem 13 ensures that \( h_\alpha \in I_\Omega \cap \Pi_k \). 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 \( H_0 \) we can, like in the standard Gröbner basis division algorithm, eliminate all monomials from \( \Gamma_k \backslash \Pi_k \) 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 \( \#\Lambda_k \), the only polynomial \( q' \in \Pi_k \cap I_\Omega \) is \( q' = 0 \). Hence, \( q \in H_k \) and therefore \( \text{span } H_k = I_\Omega \cap \Pi_k \).

Since \( \Pi \) is a Noetherian ring, cf. (Cox et al., 1996), there exists some \( n \in \mathbb{N}_0 \) such that the increasing chain \( (H_k), k \in \mathbb{N}_0 \), 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 \( \Lambda(H_{n+1}) \cap \Pi_n^0 = \Pi_{n+1}^0 \) and \( A_n = A_{n+1} = \cdots \). Since \( \Pi_n = A_n \cup H_n \), it follows that \( \Pi_n \) is a degree reducing interpolation space, see (Sauer, 2006), and that

\[ H := \{ h_\alpha : \alpha \in I_n \} \]
is an \( \Pi \)-basis for \( I_\Omega \).
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 “≤” where \( \alpha < \beta \) 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). “≤” is a total ordering on \( \Gamma \) 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: \( B = \Gamma_{N+1}, I := \emptyset, A = \emptyset, F = [ ] \in \mathbb{C}^{\#\Omega \times 0} \).
2. While \( B \neq \emptyset \) (a) \( \beta := \min_{\leq} B \).
   (b) Expand the matrix by one column: \( \tilde{F} = [F | f(\alpha + \beta) : \alpha \in \Gamma_N] \).
   (c) If \( \text{rank} \tilde{F} = \text{rank} F \) then determine \( 0 \neq q_\beta \in \ker \tilde{F} \) and set \( B := B \setminus (\beta + \Gamma_{N+1-|\beta|}) \), \( I := I \cup \{\beta\} \).
   If \( \text{rank} \tilde{F} > \text{rank} F \) then set \( F := \tilde{F} \) and \( B := B \setminus \{\beta\} \), \( A := A \cup \{\beta\} \).

**Results:** Gröbner basis \( \{q_\alpha : \alpha \in I\} \) and monomial quotient space \( \Pi_A \cong \Pi/I_\Omega \).

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 \( \text{rank} F = \#A \). 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 \( \beta \)th 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 \( \#\Gamma_N \leq N \log^{s-1} N \) 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 \( I \subseteq \partial A \).

Once the set \( A \), and the \( \mathrm{H} \)-basis \( H \) are determined, the points \( X_\Omega \) 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 \( X_\Omega \) and thus \( \Omega \) are determined, the coefficients \( f_\omega, \omega \in \Omega \), 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

\[ f(x) = \sum_{\kappa \in K} f_\kappa x^\kappa, \quad f_\kappa \in \mathbb{C}\{0\}, \quad \kappa \in K, \]

where sparsity means that \( |K| \) is (very) small relative to \( \deg K + s \). This requirement is quite easy to achieve in several variables.

The “classical” method to reconstruct \( f \) from samples on \( \Gamma \) 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

\[ f(2^{\Theta \alpha}) = \sum_{\kappa \in K} f_\kappa e^{\log 2 (\Theta \kappa) T \alpha} = \sum_{\kappa \in K} f_\kappa e^{\omega_\kappa T \alpha}, \quad \omega_\kappa := \log 2 \Theta \kappa, \]

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

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

If the coefficients \( f_\kappa \) 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 \( \omega_\kappa \) from which the exponents can be computed as

\[ \kappa = \text{rd} \left( \frac{1}{\log 2} \Theta^{-1} \omega_\kappa \right) \]

by rounding to the next integer.

7. Appendix: 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 \( \Pi_A \) 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

\[ q_\alpha := (\cdot)^\alpha - L_A(\cdot)^\alpha, \quad \alpha \in \Gamma, \]

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

$$p(x) = L_\alpha p(x) + p(x) - L_\alpha p(x) = L_\alpha p(x) + \sum_{\alpha \in A} p_\alpha q_\alpha(x)$$

and

$$p \in I_X \iff p(x) = \sum_{\alpha \in A} p_\alpha q_\alpha(x). \quad (25)$$

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 $\{q_\alpha : \alpha \in A^c\}$, $A^c_{n+1} := A^c \cap \Gamma_{n+1}$, where $n := \deg A = \max\{|\alpha| : \alpha \in A\}$, is a finite H–basis of $I_X$.

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

$$q_{\alpha+\epsilon}(x) - x_j q_\alpha(x) = x^{\alpha+\epsilon} - L_\alpha (x)^{\alpha+\epsilon}(x) - x_j L_\alpha (x)^\alpha(x) = x_j L_\alpha (x)^\alpha(x) - L_\alpha (x)^{\alpha+\epsilon}(x) = \sum_{\beta \in \partial A} p_{\beta-\epsilon} x^\beta + \sum_{\beta \in \partial A} p_{\beta-\epsilon} x^\beta - L_\alpha (x)^{\alpha+\epsilon}(x)$$

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

$$q_{\alpha+\epsilon}(x) - x_j q_\alpha(x) = \sum_{\beta \in \partial A} c_{\alpha, \beta} q_\beta(x), \quad c_{\alpha, \beta} \in \mathbb{C}. \quad (26)$$

Therefore, any $q_\alpha$ with $\alpha \in A^c$ such that $\alpha - \epsilon_j \in A^c$ has a H–representation by $(.) q_{\alpha-\epsilon_j}$ and $q_{\beta}$, $\alpha \in \partial A \cap \Gamma_n$. If, on the other hand, $\alpha - \epsilon_j \notin A^c$, $j = 1, \ldots, s$, and $\alpha \neq 0$, then there must be some 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 $q_\alpha$ must eventually be written as a linear combination of $q_\beta$, $\beta \in \partial A \cap \Gamma_n$.

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

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

of $p$ is perpendicular to the homogeneous leading forms in $\Lambda(I_X) \cap \Pi_0^p$, where $\Lambda(p) := \deg p \in \Pi_0^\deg p$. 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

$$p = \sum_{h \in H} q_h h + r, \quad \deg q_h + \deg h \leq \deg p, \quad (27)$$
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 $\Pi_A$ 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 $q_\alpha$, $\alpha \in A^c$, defined in (24) and define the inner product separately on $\Pi_0^0 \times \Pi_0^0$, $n \in \{0, 1, \ldots\}$. If $n < \min\{ |\alpha| : \alpha \in A \}$ and $n > \deg A$, we simply use the inner product of the coefficients,

$$(p, p')_n := \sum_{|\alpha|=n} \overline{p_\alpha} p'_\alpha, \quad p, p' \in \Pi_0^n.$$ 

For other values of $n$ we first observe that the polynomials $x^\alpha$, $\alpha \in A \cap \Gamma_0^0$ and the leading forms $\Lambda(q_\alpha)$, $\alpha \in A^c \cap \Gamma_0^0$, span $\Pi_0^0$. We arrange the coefficient vectors into a nonsingular matrix $Y \in \mathbb{C}^{r_0^0 \times r_0^0}$, where $r_0^0 := \dim \Pi_0^0 = \left(\begin{smallmatrix} s+n-1 \\ s-1 \end{smallmatrix}\right)$ and note that the Gramian $G := YY^H$ is hermitian and positive definite. Defining

$$(p, p')_n = p^H G^{-1} p = \sum_{|\alpha|=|\beta|=n} (G^{-1})_{\alpha,\beta} p_\alpha p'_\beta, \quad p, p' \in \Pi_0^n,$$

we get that

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

which means that the vectors $e_\alpha$, $\alpha \in A \cap \Gamma_0^0$ are perpendicular to the coefficient vectors of $\Lambda(q_\alpha)$, $\alpha \in A^c \cap \Gamma_0^0$. Consequently, the inner product

$$(p, p') = \sum_{\alpha \in A} (p, p')_\alpha, \quad p, p' \in \Pi,$$

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

**Acknowledgement**

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

**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. Vol. 86 of Internat. Schriftenreihe Numer. Math. 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. Advances 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.
and elevation angles of arrival along with the frequency of the incoming signals. Digital Signal Processing: A Review
Journal 16, 796–816.
Zippel, R., 1979. Probabilistic algorithms for sparse polynomials. Ph.D. thesis, MIT.
Zippel, R., 1990. Interpolating polynomials from their values. J. Symbolic Comput 9, 375–403.