Methods and Applications of Analysis.
Vol. 25, No. 3, pp. 257–276, September 2018

005

# GEOMETRY AND SINGULARITIES OF PRONY VARIETIES*

GIL GOLDMAN†, YEHONATAN SALMAN†, AND YOSEF YOMDIN†

*To the Memory of John Mather*

**Abstract.** We start a systematic study of the topology, geometry and singularities of the Prony varieties $S_q(\mu)$, defined by the first $q+1$ equations of the classical Prony system $\sum_{j=1}^{d} a_j x_j^k = \mu_k,\ k=0,1,\dots$ .

Prony varieties, being a generalization of the Vandermonde varieties, introduced in [5,22], present a significant independent mathematical interest (compare [5, 20, 22]). The importance of Prony varieties in the study of the error amplification patterns in solving Prony system was shown in [1–4, 20]. In [20] a survey of these results was given, from the point of view of Singularity Theory.

In the present paper we show that for $q \geq d$ the variety $S_q(\mu)$ is diffeomerphic to the intersection of a certain affine subspace in the space $\mathcal{V}_d$ of polynomials of degree $d$, with the hyperbolic set $H_d$.

In case of the Prony curves $S_{2d-2}$ we study the behavior of the amplitudes $a_j$ as the nodes $x_j$ collide, and the nodes escape to infinity.

We discuss the behavior of the Prony varieties as the right hand side $\mu$ varies, and possible connections of this problem with J. Mather's result in [24] on smoothness of solutions in families of linear systems.

**Key words.** Prony system, spike-train signals, Prony and Vandermonde varieties, nodal collision singularities.

**Mathematics Subject Classification.** 65H10, 94A12, 32S05, 58K20.

## 1. Introduction.

This paper is devoted to a detailed study of "Prony varieties", which play important role in some problems of Signal Processing (in particular, in Fourier reconstruction of "spike-train signals" - see Section 1.1 below). We believe that Prony varieties present a significant independent mathematical interest, especially from the point of view of Singularity Theory (compare [20]).

In particular, in the course of our study we provide proofs of most of the results announced in [20]. (However, we keep the present paper independent of [20], and give all the necessary definitions).

### 1.1. Prony system.

We consider the classical *Prony system* of algebraic equations, of the form

$$\sum_{j=1}^{d} a_j x_j^k = \mu_k,\ k=0,1,\dots,2d-1. \tag{1.1}$$

Here $d$ and the right hand side $\mu=(\mu_0,\dots,\mu_{2d-1})$ are assumed to be known, while $a=(a_1,\dots,a_d)\in\mathbb{R}^d$ and $x=(x_1,\dots,x_d)\in\mathbb{R}^d$ are the unknowns to be found.

Prony system appears, in particular, in the problem of moment reconstruction of spike-trains, that is, of one-dimensional signals $F$ which are linear combinations of $d$ shifted $\delta$-functions:

$$F(x)=\sum_{j=1}^{d} a_j \delta\left(x-x_j\right). \tag{1.2}$$

*Received June 4, 2018; accepted for publication October 12, 2018.
†Department of Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel (gilgoldm @gmail.com; salman.yehonatan@gmail.com; yosef.yomdin@weizmann.ac.il).

We assume that the form (1.2) of signals $F$ is a priori known, but the specific parameters - the amplitudes $a_j$ and the nodes $x_j$ are unknown. Our goal is to reconstruct them from $2d$ moments $m_k(F) = \int_{-\infty}^{\infty} x^k F(x)dx,\ k = 0, \ldots, 2d-1$, which are known with a possible error bounded by $\epsilon > 0$.

An immediate computation shows that the moments $m_k(F)$ are expressed as $m_k(F) = \sum_{j=1}^{d} a_j x_j^k$. Hence our reconstruction problem is equivalent to solving the Prony system (1.1), with $\mu_k = m_k(F)$.

We will identify the signal $F$ with the pair $(a, x)$ of the amplitudes and the nodes of $F$. We will assume that the nodes $x$ are pairwise distinct and ordered: $x_1 < x_2 < \ldots < x_d$, and denote the space of the nodes by

$$\mathcal{P}_d^x \cong \Delta_d = \{(x_1, \ldots, x_d) \in \mathbb{R}^d,\ x_1 < x_2 < \ldots < x_d\}.$$

Denote the space of the amplitudes by $\mathcal{P}_d^a$. Finally denote by $\mathcal{P}_d = \mathcal{P}_d^a \times \mathcal{P}_d^x$, the parameter space of signals $F$ with $d$ nodes.

The space of the moments $m_0, \ldots, m_{2d-1}$ (or of the right-hand sides $\mu = (\mu_0, \ldots, \mu_{2d-1})$ of the Prony system (1.1)) will be denoted by $\mathcal{M}_d \cong \mathbb{R}^{2d}$.

Prony system appears in many theoretical and applied mathematical problems. There exists a vast literature on Prony and similar systems - see [6, 8–10, 12, 26–30] and references therein, as a very small sample.

Some applications of Prony system are of major practical importance, and, in case when some of the nodes $x_j$ nearly collide, it is well known that the Prony system presents major mathematical difficulties, in particular, in the context of "super-resolution problem" (see [1–4, 7, 9, 13–17, 19, 25, 26] as a small sample).

Recent papers [1–4] deal with the problem of "error amplification" in solving a Prony system in the case when some of the nodes $x_1, \ldots, x_d$ nearly collide. Our approach is independent of a specific method of inversion and deals with a possible amplification of the measurements errors, in the reconstruction process, caused by the geometric nature of the Prony system.

The main observations and results in [1–4] can be shortly summarized as follows: *the incorrect reconstructions, caused by the measurements noise, are spread along certain algebraic subvarieties $S_q$ in the parameter space $\mathcal{P}_d$, which we call the "Prony varieties" (see the next section).*

This important fact allows us to better understand the geometry of the error amplification in solving the Prony system. On one hand, we produce on this basis rather accurate upper and lower bounds for the worst case reconstruction error. On the other hand, we show that in some cases this geometric information can be used to improve the expected reconstruction accuracy.

A survey of the results on the geometry of error amplification, obtained in [1–4], is given in [20]. This survey stresses the role of Singularity Theory in the study of Prony inversion, and, in particular, in the study of the Prony varieties, and announces some new results in this direction. However, these results on the geometry and singularities of the Prony varieties are stated in [20] without proofs.

In the present paper we start a systematic study of the topology, geometry and singularities of the Prony varieties, providing, in particular, proofs for most of the results announced in [20].

### 1.2. Prony varieties.

Definition 1.1. *For $\mu \in \mathcal{M}_d$ and for $q = 0, \ldots, 2d-1$, the Prony variety $S_q = S_q(\mu)$ is an algebraic variety in the parameter space $\mathcal{P}_d$, defined by the first $q+1$*

*equations of the Prony system (1.1):*

$$\sum_{j=1}^{d} a_j x_j^k = \mu_k, \ k = 0, 1, \ldots, q. \tag{1.3}$$

Thus the variety $S_q(\mu)$ is completely determined by the first $q+1$ moments $\mu_k, \ k = 0, 1, \ldots, q$, which are preserved along $S_q(\mu)$. The dimension of the variety $S_q(\mu)$ is $2d - q - 1$. The chain of inclusions

$$S_0 \supset S_1 \supset \ldots \supset S_{2d-2} \supset S_{2d-1}$$

can be explicitly computed (in principle), from the known measurements $\mu = (\mu_0, \ldots, \mu_{2d-1}) \in \mathcal{M}_d$. Notice that $S_{2d-1}(\mu)$ coincides with the set of solutions of the "full" Prony system (1.1).

If in equations (1.3) above we fix the amplitudes $a_j$, we obtain the "Vandermonde varieties" in $\mathcal{P}_d^x$, as introduced in [5, 22]. We expect that Prony varieties, being, essentially, the "fibered spaces", with the Vandermonde varieties as the fibers, share important properties of the latter, described in [5, 22].

In our approach to solving Prony systems, the Prony varieties $S_q$ serve as approximations to the set of possible "noisy solutions" of (1.1) which appear for a noisy right-hand side $\mu$. The Prony curve $S_{2d-2}$ is especially essential in the presentation below.

An important fact discovered in [1–4] is that *if the nodes $x_1, \ldots, x_d$ form a cluster of a size $h \ll 1$, while the measurements error is of order $\epsilon$, then the worst case error in reconstruction of $S_q$ is of order $\epsilon h^{-q}$. Thus, for smaller $q$, the varieties $S_q$ are bigger, but the accuracy of their reconstruction is better. The same is true for the accuracy with which $S_q$ approximates noisy solutions of (1.1).* That is, the "real", as well as the noisy solutions to the Prony system (1.1) lie inside an $\epsilon h^{-q}$-neighborhood of the Prony variety $S_q(\mu)$ calculated from the noisy moment measurements $\mu$.

In particular, the worst case error in the reconstruction of the solution $S_{2d-1}$ of (1.1) is $\sim \epsilon h^{-2d+1}$, while the worst case error in reconstruction of the Prony curve $S_{2d-2}$ is of order $\epsilon h^{-2d+2}$. That is, the reconstruction of the Prony curve $S_{2d-2}$ can be done $h$ times better than the reconstruction of the solutions themselves.

Consequently, we can split the process of solution of (1.1) into two steps: first finding, with an improved accuracy, the Prony curve $S_{2d-2}(\mu)$, and then localizing the solution to (1.1) *on this curve.* In particular, in the presence of a certain additional a priori information about the expected solutions of the Prony system (for example, upper and/or lower bounds on the amplitudes), it was shown in [1–4, 20] that the Prony curves can be used to significantly improve the overall reconstruction accuracy.

We believe that the results of [1–4, 20, 21], as well as connections of the Prony varieties with Vandermonde varieties, justify a detailed algebraic-geometric study of the Prony varieties, including their singularities. (We refer the reader to [20] and references therein for a survey of these results from the point of view of Singularity Theory.)

The paper is organized as follows: in Section 2 our main results and their proofs are presented. They include a global algebraic-geometric description of the Prony varieties $S_q(\mu)$. It is convenient to consider separately the cases $q \leq d-1$ and $q \geq d$. The proofs in the first case are given in Section 2.1 and in the second case in Section 2.2.

In Section 2.3, we informally discuss the behavior of the Prony varieties $S_q(\mu)$ as functions of $\mu$. This leads, via the previous results, to solving parametric linear systems, and to possible connections of this problem with J. Mather's result on smoothness of solutions in families of linear systems (see [24]).

In Section 2.4 we consider the case of Prony curves $S_{2d-2}(\mu)$ and describe the behavior of the amplitudes at the singularities of nodes collision. We also investigate a situation when the nodes escape to infinity.

In Section 3, as an illustration, a complete description of the Prony varieties in the case of two nodes is given.

The authors would like to thank the referee for suggesting improvements in the presentation, and for bringing to our attention paper [18].

**2. Global description of Prony varieties.** In this section we start a global algebraic-geometric and topological investigation of the Prony varieties. Our main results are as follows:

THEOREM 2.1. *For each $\mu \in \mathcal{M}_d$ and for each $q = 0, \ldots, 2d-1$, the Prony variety $S_q(\mu)$ is a smooth submanifold of $\mathcal{P}_d$ (for $q \geq d$, possibly empty), and its node projection $S_q^x(\mu)$ is a smooth submanifold of $\mathcal{P}_d^x$.*

It is convenient to separate the cases $q \leq d-1$ and $q \geq d$. In the first case, we have the following result:

THEOREM 2.2. *For each $q \leq d-1$ and for each $\mu \in \mathcal{M}_d$, the variety $S_q(\mu)$ satisfies the following conditions:*

*1. $S_q(\mu)$ is non-empty, its dimension is equal to $2d-q-1$, and the equations* (1.3) *are regular at each point $P$ of $S_q(\mu)$, i.e., the rank of their Jacobian at $P$ is $q+1$.*

*2. The nodes $x_1, \ldots, x_d$ and any $d-q-1$ amplitudes (in particular, $a_{q+2}, \ldots, a_d$) form a global regular coordinate system on $S_q(\mu)$.*

*3. The projection $p_x : S_q(\mu) \to \mathcal{P}_d^x$ is onto. It defines $S_q(\mu)$ as a trivial affine bundle, naturally isomorphic to $\mathbb{R}^{d-q-1} \times \mathcal{P}_d^x$. In particular, $S_q(\mu)$ is topologically trivial.*

The case $q \geq d$ is somewhat more complicated. To state the result we need some definitions and notations. Let

$$
\begin{aligned}
\sigma_0(x_1, \ldots, x_d) &= 1 \\
\sigma_1(x_1, \ldots, x_d) &= -(x_1 + \ldots + x_d) \\
&\cdots \\
\sigma_k(x_1, \ldots, x_d) &= (-1)^k \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq d} x_{i_1} x_{i_2} \cdots x_{i_k} \\
&\cdots \\
\sigma_d(x_1, \ldots, x_d) &= (-1)^d x_1 x_2 \cdot \ldots \cdot x_{d-1} x_d,
\end{aligned}
\tag{2.1}
$$

be the elementary symmetric (Vieta) polynomials in $x_1, \ldots, x_d,$. Thus $\sigma_j$ are the coefficients of the univariate polynomial

$$
Q(z) = \prod_{j=1}^{d} (z - x_j) = z^d + \sigma_1 z^{d-1} + \ldots + \sigma_d = \sum_{i=0}^{d} \sigma_{d-i} z^i,
$$

whose roots are the nodes $x_1, \ldots, x_d$.

Let $\mathcal{V}_d \cong \mathbb{R}^d$ be the space of the coefficients $\sigma = (\sigma_1, \dots, \sigma_d)$ of the polynomials $Q(z)$ (which we identify with the space of the monic polynomials $Q$).

Consider a subset $H_d \subset \mathcal{V}_d$, consisting of *hyperbolic* polynomials $Q$, i.e. of polynomials $Q(z) = z^d + \sigma_1 z^{d-1} + \ldots + \sigma_d$ with all their roots real and distinct. Thus we consider *the open* hypebolic set, excluding the boundary.

The set $H_d$ is important in many problems, and it was intensively studied (see, as a small sample, [5, 22, 23] and references therein).

DEFINITION 2.1. *The "root mapping" $RM_d : H_d \to \mathcal{P}^x_d$ is defined by*

$$RM_d(Q) = x = (x_1, \dots, x_d) \in \mathcal{P}^x_d,$$

*where $x_1 < x_2 < \ldots < x_d$ are the ordered roots of the hyperbolic polynomial $Q(z) \in H_d$.*

*The "Vieta mapping" $VM_d : \mathcal{P}^x_d \to H_d$ is defined by*

$$VM_d(x_1, \dots, x_d) = (\sigma_1(x_1, \dots, x_d), \dots, \sigma_d(x_1, \dots, x_d)),$$

*where $\sigma_i = \sigma_i(x_1, \dots, x_d)$, $i = 1, \dots, d$, are the Vieta symmetric polynomials in $x_1, \dots, x_d$,.*

Clearly, on $H_d$ the root mapping $RM_d$ is regular, and $RM_d = VM_d^{-1}$. Therefore both the mapping $RM_d : H_d \to \mathcal{P}^x_d$ and its inverse $VM_d : \mathcal{P}^x_d \to H_d$ provide a regular algebraic diffeomorphism between $H_d$ and $\mathcal{P}^x_d$.

Let $\mu \in \mathcal{M}_d$ be given. For $q \geq d$, consider the following system of $q - d + 1$ linear equations for $\sigma_1, \dots, \sigma_d$:

$$\begin{aligned} \mu_{d-1}\sigma_1 + \mu_{d-2}\sigma_2 + \ldots + \mu_0\sigma_d &= -\mu_d \\ \mu_d\sigma_1 + \mu_{d-1}\sigma_2 + \ldots + \mu_1\sigma_d &= -\mu_{d+1} \\ \ldots\ldots\ldots \\ \mu_{q-1}\sigma_1 + \mu_{q-2}\sigma_2 + \ldots + \mu_{q-d}\sigma_d &= -\mu_q \end{aligned} \tag{2.2}$$

Taking into account that $\sigma_0 = 1$, this system can be rewritten as

$$\sum_{i=0}^{d} \mu_{l-i}\sigma_i = 0, \ l = d, \dots, q.$$

For a signal $F$ with nodes $x_1, \dots, x_d$ and moments $\mu = (\mu_0, \dots, \mu_q)$, system (2.2) forms a part of the standard (and classical) linear system for the coefficients of the polynomial $Q$ (see, for instance, [26, 29, 30]). For $q = 2d - 1$, the complete system is obtained.

Equations (2.2) define an affine subspace $L_q(\mu) \subset \mathcal{V}_d$, which is generically of dimension $2d - q - 1$ (but, depending on $\mu$, $L_q(\mu)$ may be empty, or of any dimension not smaller than $2d - q - 1$). We denote by $L^h_q(\mu)$ the intersection of $L_q(\mu)$ and the set $H_d$ of hyperbolic polynomials.

Finally, we notice that system (2.2), being a linear system in variables $\sigma_1, \dots, \sigma_d$, is a nonlinear system in $x_1, \dots, x_d$, if we consider $\sigma_j$ as the Vieta elementary symmetric polynomials in $x_1, \dots, x_d$.

Now we have all the tools required to describe the Prony varieties $S_q(\mu)$ for $q \geq d$:

THEOREM 2.3. *For each $q \geq d$ and for any $\mu \in \mathcal{M}_d$, the variety $S_q(\mu)$ satisfies the following conditions:*

*1. Either $S_q(\mu)$ is empty, or it is smooth and its dimension is greater than or equal to $2d-q-1$.*

*2. $S_q^x(\mu)$ is defined in $\mathcal{P}_d^x$ by system (2.2), considered as a non-linear system in $x_1,\dots,x_d$. The Vieta mapping $VM$ and its inverse root mapping $RM$ provide a diffeomorphism between $S_q^x(\mu)$ and $L_q^h(\mu)$.*

*3. The projection $p_x : S_q(\mu) \to S_q^x(\mu)$, as well as its inverse, are $1-1$, and provide a diffeomorphism between $S_q(\mu)$ and $S_q^x(\mu)$.*

We prove Theorems 2.1 - 2.3 in the following order: first Theorem 2.2, then Theorem 2.3. Theorem 2.1 then follows directly.

**2.1. The case $q \le d-1$. Proof of Theorem 2.2.** A direct computation shows that the upper left $d\times d$ minor of the Jacobian matrix $JP_{2d-1}$ of the Prony system of equations (1.1) is the Vandermonde matrix on the nodes $x_1,\dots,x_d$. By construction, for each signal $F$ in $\mathcal{P}_d$ these nodes are pairwise distinct, and hence the first $d$ rows of $JP_{2d-1}(F)$ are linearly independent. Therefore, for each $q \le d-1$, all the $q+1$ rows of the Jacobian matrix $JP_q(F)$ of the partial system (1.3) are linearly independent. This proves, via Implicit function theorem, that the variety $S_q(\mu)$ is smooth, of dimension $2d-q-1$, which is statement 1 of Theorem 2.2.

In order to prove statement 2, we rewrite equations (1.3) as

$$
\begin{aligned}
a_1 + a_2 + \ldots + a_{q+1} &= \mu_0 - a_{q+2} - \ldots - a_d\\
a_1x_1 + a_2x_2 + \ldots + a_{q+1}x_{q+1} &= \mu_1 - a_{q+2}x_{q+2} - \ldots - a_dx_d\\
a_1x_1^2 + a_2x_2^2 + \ldots + a_{q+1}x_{q+1}^2 &= \mu_2 - a_{q+2}x_{q+2}^2 - \ldots - a_dx_d^2\\
&\cdots\cdots\cdots\\
a_1x_1^q + a_2x_2^q + \ldots + a_{q+1}x_{q+1}^q &= \mu_q - a_{q+2}x_{q+2}^q - \ldots - a_dx_d^q
\end{aligned}
\tag{2.3}
$$

The left - hand side of (2.3) is the Vandermonde linear system on the pairwise different nodes $x_1,\dots,x_{q+1}$ with respect to $a_1,\dots,a_{q+1}$. Hence we can uniquely express from (2.3) the amplitudes $a_1,\dots,a_{q+1}$ using the Cramer rule. The resulting expressions will be linear in $\mu$ and in $a_{q+2},\dots,a_d$, with the coefficients being rational functions in the nodes. Their denominator is the Vandermonde determinant $V_q(x_1,\dots,x_{q+1}) = \prod_{1\le i<j\le q+1}(x_j-x_i)$, which does not vanish at the points $x=(x_1,\dots,x_{q+1},\dots,x_d)\in \mathcal{P}_d^x$. Thus (2.3) regularly expresses $a_1,\dots,a_{q+1}$ through $x_1,\dots,x_d,a_{q+2},\dots,a_d$. We conclude that these last parameters can be considered as a global regular coordinate system on $S_q(\mu)$. This proves statement 2 of Theorem 2.2.

In order to prove statement 3, we notice that by (2.3), for any fixed $x = (x_1,\dots,x_d)$, the fiber over $x$ of the projection $p_x : S_q(\mu) \to \mathcal{P}_d^x$ is an affine subspace in $\mathcal{P}_d$ of dimension $d-q-1$, regularly parametrized by $a_{q+2},\dots,a_d$. Therefore (2.3) provides an isomorphism of the affine fiber bundles $p_x : S_q(\mu)\to \mathcal{P}_d^x$ and $\mathbb{R}^{d-q-1}\times\mathcal{P}_d^x$. This completes the proof of statement 3 and of Theorem 2.2. □

Let us look at a special case $q=d-1$. In this case the Prony variety $S_{d-1}(\mu)$ has dimension $d$, and, according to Theorem 2.2, the nodes $x_1<x_2<\ldots<x_d$ can be taken as the coordinates on $S_{d-1}(\mu)$. The Cramer rule applied to (2.3) gives

$$
a_j = \frac{1}{V_d(x_1,\dots,x_d)}\sum_{l=0}^{q} A_l^j(x_1,\dots,x_d)\mu_l,\ j=1,\dots,d, \tag{2.4}
$$

with $A_l^j$ the corresponding minors of the Vandermonde matrix of (2.3). In fact, the coefficients in (2.4) can be written in a much simpler form.

By a certain abuse of notations, let us denote by $\varrho_k(u_1,\dots,u_{d-1}),\ k=1,\dots,d-1$, the Vieta symmetric polynomials in $d-1$ variables.

For $x=(x_1,\dots,x_d)$, we put $\pi_i(x)=(x_1,\dots,x_{i-1},x_{i+1},\dots,x_d),\ i=1,\dots,d$, and consider the "partial" symmetric polynomials

$$\varrho_k(\pi_i(x)),\ k=1,\dots,d-1,\ i=1,\dots,d.$$

Let $L_i(x)=\prod_{1\le l\le d,\ l\ne i}(x_i-x_l)$ be the denominators in the summands of the Lagrange interpolation polynomial on the nodes $x_1,\dots,x_d$.

Proposition 2.1. *The amplitudes $a_i$ of the points on the Prony variety $S_{d-1}(\mu)$ are given in terms of the nodes $x=(x_1,\dots,x_d)$ via the expressions*

$$a_i=\frac{P(\pi_i(x))}{L_i(x)}, \tag{2.5}$$

*where for $u=(u_1,\dots,u_{d-1})$ the polynomial $P(u)$ of $d-1$ variables is defined by*

$$P(u)=\mu_0\varrho_{d-1}(u)+\ldots+\mu_{d-2}\varrho_1(u)+\mu_{d-1}.$$

*Proof.* Observe that for $q=d-1$, system (2.3) can be rewritten as follows:

$$\begin{pmatrix}1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_d\\ & & \cdots\cdots & \\ x_1^{d-1} & x_2^{d-1} & \cdots & x_d^{d-1}\end{pmatrix}\begin{pmatrix}a_1\\ a_2\\ \cdots\\ a_d\end{pmatrix}=\begin{pmatrix}\mu_0\\ \mu_1\\ \cdots\\ \mu_{d-1}\end{pmatrix}.$$

Using the well-known formula for the inverse of the Vandermonde matrix (see, for example [31]) we obtain from the last matrix equation that

$$\begin{pmatrix}a_1\\ a_2\\ \cdots\\ a_d\end{pmatrix}=\begin{pmatrix}\frac{\varrho_{d-1}(\pi_1(x))}{L_1(x)} & \frac{\varrho_{d-2}(\pi_1(x))}{L_1(x)} & \cdots & \frac{\varrho_1(\pi_1(x))}{L_1(x)} & \frac{1}{L_1(x)}\\ \frac{\varrho_{d-1}(\pi_2(x))}{L_2(x)} & \frac{\varrho_{d-2}(\pi_2(x))}{L_2(x)} & \cdots & \frac{\varrho_1(\pi_2(x))}{L_2(x)} & \frac{1}{L_2(x)}\\ \cdots\cdots & & & & \\ \frac{\varrho_{d-1}(\pi_d(x))}{L_d(x)} & \frac{\varrho_{d-2}(\pi_d(x))}{L_d(x)} & \cdots & \frac{\varrho_1(\pi_d(x))}{L_d(x)} & \frac{1}{L_d(x)}\end{pmatrix}\begin{pmatrix}\mu_0\\ \mu_1\\ \cdots\\ \mu_{d-1}\end{pmatrix}. \tag{2.6}$$

Therefore for $i=1,\dots,d$, we get $a_i=\frac{P(\pi_i(x))}{L_i(x)}$, with

$$P(u)=\mu_0\varrho_{d-1}(u)+\ldots+\mu_{d-2}\varrho_1(u)+\mu_{d-1}$$

defined as above. This completes the proof of Proposition 2.1. □

Corollary 2.1. *For each $q\ge d-1$, equations (2.5), expressing the amplitudes $a_1,\dots,a_d$ through the nodes $x_1,\dots,x_d$, remain valid on the Prony varieties $S_q(\mu)$.*

*Proof.* By definition, for $q\ge d-1$, we have $S_q(\mu)\subset S_{d-1}(\mu)$. □

**2.2. The case $q \geq d$. Proof of Theorem 2.3.** For each $q \geq d$, the dimension $2d - q - 1$ of the Prony varieties $S_q(\mu)$ is strictly smaller than $d$. Consequently, the projections $S_q^x(\mu)$ of $S_q(\mu)$ onto the nodes subspace $\mathcal{P}_d^x$ are proper subvarieties in $\mathcal{P}_d^x$.

By Corollary 2.1 we conclude that for each $x = (x_1, \ldots, x_d) \in S_q^x(\mu)$, the amplitudes $a = (a_1, \ldots, a_d)$ are uniquely defined, and given by regular expressions (2.5). Thus the projection of $S_q(\mu)$ to $S_q^x(\mu)$ is one-to-one and regular, as well as its inverse. This proves statement 3 of Theorem 2.3.

Next we concentrate on the Prony varieties $S_q^x(\mu)$, and show that they are defined in $\mathcal{P}_d^x$ by system (2.2). We have to eliminate the amplitudes $a_1, \ldots, a_d$ from the equations (1.3). For this purpose we use a modification of the classical solution method of the Prony system.

First we show that for $q \geq d$, system (1.3) implies system (2.2). Indeed, for each $l = d, \ldots, q$, we obtain, using (1.3), that

$$\sum_{i=0}^{d} \mu_{l-i}\sigma_i = \sum_{i=0}^{d} \sigma_i \sum_{j=1}^{d} a_j x_j^{l-i} = \sum_{j=1}^{d} a_j \sum_{i=0}^{d} \sigma_i x_j^{l-i} = \sum_{j=1}^{d} a_j x_j^{l-d} Q(x_j) = 0,$$

since each node $x_j$ is a root of $Q(x)$. In other words, for each $(a, x) \in \mathcal{P}_d$ satisfying system (1.3), the component $x$ satisfies system (2.2). We conclude that the projection $S_q^x(\mu)$ of $S_q(\mu)$ onto the nodes subspace $\mathcal{P}_d^x$ is contained in the zero set of system (2.2).

To prove the opposite inclusion, let us assume that $x = (x_1, \ldots, x_d)$ satisfies system (2.2). We uniquely define the amplitudes $a = (a_1, \ldots, a_d)$ from the Vandermonde linear system, formed by the first $d$ equations of system (1.3), according to expressions (2.5). Now we form a signal

$$F(x) = \sum_{j=1}^{d} a_j \delta(x - x_j) = (a, x) \in \mathcal{P}_d,$$

which by construction satisfies the first $d$ equations of system (1.3). It remains to show that the last $q - d + 1$ equations of (1.3) are satisfied for $F(x)$.

Consider the rational function $R(z) = \sum_{j=1}^{d} \frac{a_j}{z - x_j}$. We have $R(z) = \frac{P(z)}{Q(z)}$ for a certain polynomial $P(z)$ of degree $d - 1$ and for

$$Q(z) = \prod_{j=1}^{d} (z - x_j) = z^d + \sigma_1 z^{d-1} + \ldots + \sigma_d,$$

where $\sigma_i = \sigma_i(x_1, \ldots, x_d)$, $i = 1, \ldots, d$, are, as above, the Vieta elementary symmetric polynomials in $x_1, \ldots, x_d,$.

Developing the elementary fractions in $R(z)$ into geometric progressions, we get

$$R(z) = \sum_{k=0}^{\infty} \frac{m_k}{z^{k+1}}, \quad m_k = m_k(F) = \sum_{j=1}^{d} a_j x_j^k. \tag{2.7}$$

Therefore, the moments $m_k = m_k(F)$, $k = 0, 1, \ldots,$ given by the left-hand side $\sum_{j=1}^{d} a_j x_j^k$ of system (1.3), are the Taylor coefficients of the rational function $R(z) = \frac{P(z)}{Q(z)}$, with $P(z)$ of degree $d - 1$, and $Q(z)$ of degree $d$. Starting with $k = d$ these Taylor coefficients $m_k$ of $R$ are known to satisfy the recurrence relation

$$m_k = -\sum_{s=1}^{d} \sigma_s m_{k-s}, \tag{2.8}$$

$\sigma_l$ being the coefficients of the denominator $Q(z)$ of $R(z)$. Since by the choice of the amplitudes $a_j$, the first $d$ equations of system (1.3) are satisfied, we conclude that $m_k = \mu_k,\ k = 0, \ldots, d-1$.

Now we use the assumption that system (2.2) is satisfied. Its equations show that $\mu_k$ satisfy exactly the same recurrence relation till $k = q$. Since the first $d$ terms are the same, we conclude that in fact $m_k = \mu_k,\ k = 0, \ldots, q$.

This means that the entire system (1.3) is satisfied. Consequently, $F \in S_q(\mu)$, and therefore $x = (x_1, \ldots, x_d) \in S_q^x(\mu) \subset \mathcal{P}_d^x$. We conclude that the zero set of system (2.2) is contained in $S_q^x(\mu)$. This completes the proof of the fact that $S_q^x(\mu)$ is defined in $\mathcal{P}_d^x$ by system (2.2).

We conclude that the Vieta mapping $VM$ transforms the points $x = (x_1, \ldots, x_d)$ of $S_q^x(\mu)$ into the hyperbolic polynomials $Q = VM(x)$ belonging to $L_q^h(\mu) \subset \mathcal{V}_d$. Conversely, for each $Q \in L_q^h(\mu)$ its image $RM(Q)$ under the root map belongs to $S_q^x(\mu)$. Therefore the Vieta mapping $VM$ and its inverse root mapping $RM$ provide a diffeomorphism between $S_q^x(\mu)$ and $L_q^h(\mu)$. This completes the proof of statement 2 of Theorem 2.3.

In order to prove statement 1 of Theorem 2.3, we notice that $S_q(\mu) \subset \mathcal{P}_d$ is a diffeomorphic image under $p_x^{-1} \circ RM$ of $L_q^h(\mu) \subset \mathcal{V}_d$, the latter being a finite union of open domains in an affine subspace $L_q(\mu)$ of $\mathcal{V}_d$. Since $L_q(\mu)$ is defined by system (2.2) of $q - d + 1$ linear equations, $L_q^h(\mu)$ is always smooth and either empty or of dimension not smaller than $2d - q - 1$. The same is true for the diffeomorphic images $S_q^x(\mu)$ and $S_q(\mu)$ of $L_q^h(\mu)$.

This completes the proof of statement 1 of Theorem 2.3, and of the entire Theorem 2.3. □

**2.3.** $S_q(\mu)$ **and** $L_q(\mu)$ **as functions of** $\mu$**.** In Theorem 2.3 we do not make any assumption on the rank of linear system (2.2). It is easy to give examples of a right-hand side $\mu = (\mu_0, \ldots, \mu_q)$ of (2.2) for which the solutions of this system form an empty set, or an affine subspace $L_q(\mu)$ of any dimension not smaller than $2d - q - 1$. Theorem 2.3 remains true in each of these cases. Compare a detailed discussion of the situation for two nodes ($d = 2$) in Section 3 below.

Possible degenerations of system (2.2) are closely related to the conditions of solvability of the Prony system (see, for example, Theorem 3.6 of [11], and the discussion thereafter). Both these questions are very important in the robustness analysis of the Prony inversion, but we do not discuss them here. In Section 3 we illustrate the above results providing a complete description of the Prony curves in the case of two nodes.

The above observations lead to a very important question: what can be said about the behavior of the affine subspace $L_q(\mu)$ as a function of $\mu$? Via the results above, answering this question will also describe (up to intersection with $H_d$) the behavior of the Prony varieties $S_q(\mu),\ q \geq d$, as a function of $\mu$. A very important special case is $q = 2d - 1$ where $S_{2d-1}(\mu)$ is the set of solutions $F_\mu$ of the original Prony system, while $L_{2d-1}(\mu)$ is the set of polynomials $Q_\mu$ whose roots are the nodes of the solutions $F_\mu$.

Linear system (2.2) essentially presents a family, parametrized by the moments $\mu = (\mu_0, \mu_1, \ldots)$. As it was mentioned above, the rank of this system typically changes with $\mu$, and the behavior of the affine subspaces $L_q(\mu)$, as a function of $\mu$, may be rather complicated.

We expect that J. Mather's theorem on smoothness of solutions of parametric families of linear systems in [24] will be important in analysis of this problem. Notice, however, that system (2.2) is rigidly structured: its matrices $M(\mu)$ are of Hankel type. Consequently, the transversality of the family $\mu \to M(\mu)$ to the rank stratification of the space of Hankel matrices, required in J. Mather's theorem, is not self-evident. The results on linear sections of determinantal varieties obtained in ( [18], Section 4) provide important information in this direction.

Also the second condition of J. Mather's theorem, the existence of a solution for each $\mu$, is a delicate question, closely related to solvability conditions for the Prony system.

We believe that approaching these problems with the tools used in the proof of J. Mather's theorem, may be very productive.

**2.4. The case $q = 2d - 2$: Prony curves.** The case $q = 2d - 2$ is especially important in the study of error amplification (see [1–4]). For generic $\mu \in \mathcal{M}_d$, the dimension of the Prony variety $S(\mu) = S_{2d-2}(\mu)$ is one, and by Theorem 2.3 the variety $S(\mu)$ is a smooth curve consisting of a finite number of open intervals. These intervals are parametrized, via the root mapping $RM$, by the intervals of the intersection $L^h_{2d-2}(\mu)$ of the straight line $L_{2d-2}(\mu)$ with the hyperbolic set $H_d$ in the polynomial space $\mathcal{V}_d$.

In turn, a convenient explicit parametrization of the straight line $L_{2d-2}(\mu)$ can be obtain as follows: consider system (2.2), with $q = 2d - 2$, whose equations define the affine space $L_{2d-2}(\mu)$ in the polynomial space $\mathcal{V}_d$. We add to this system the last equation, corresponding to $q = 2d - 1$. We obtain the new system

$$\begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_{d-2} & \mu_{d-1} \\ \mu_1 & \mu_2 & \cdots & \mu_{d-1} & \mu_d \\ & \cdots\cdots\cdots & & & \\ \mu_{d-1} & \mu_d & \cdots & \mu_{2d-3} & \mu_{2d-2} \end{pmatrix} \begin{pmatrix} \sigma_d \\ \sigma_{d-1} \\ \cdots\cdots \\ \sigma_1 \end{pmatrix} = - \begin{pmatrix} \mu_d \\ \mu_{d+1} \\ \cdots \\ \mu_{2d-1} \end{pmatrix}, \tag{2.9}$$

which is system (2.2) with $q = 2d - 1$. The matrix in the left-hand side of (2.9) is called the moment Hankel matrix $M_d(\mu)$. This matrix plays the central role in the study of Prony systems.

We denote the determinant of $M_d(\mu)$ by $\delta(\mu)$, and its minimal singular value by $\eta(\mu)$. Notice that, in fact, the last moment $\mu_{2d-1}$ does not enter $M_d(\mu)$, and therefore this matrix is constant along the Prony curve $S(\mu)$. We conclude that for $\delta(\mu) \neq 0$, the rank of the first $d-1$ equations of (2.9) is $d-1$, and hence $L_{2d-2}(\mu)$ is one-dimensional, i.e. a straight line in $\mathcal{V}_d$.

All the entries $\mu_j$ in (2.9) besides the bottom moment $\mu_{2d-1}$ on the right-hand side are fixed on the Prony curve $S(\mu)$. Accordingly, to obtain a parametrization of $L_{2d-2}(\mu)$ we put $t = \mu_{2d-1}$ and take it as a free parameter. For each given $\mu_{2d-1} = t$, we solve (2.9) and obtain the corresponding coordinates $\sigma_1(t), \ldots, \sigma_d(t)$ of the point $Q_t(z) \in L_{2d-2}(\mu)$. Using Cramer's rule and Theorem 2.3 we get

PROPOSITION 2.2. *Let $\mu \in \mathcal{M}_d$ be such that $\delta(\mu) \neq 0$. Then the line $L_{2d-2}(\mu) \subset \mathcal{V}_d$ possesses a parametrization $Q_t(z) = z^d + \sigma_1(t) z^{d-1} + \ldots + \sigma_d(t)$ with $t = \mu_{2d-1}$ and*

$$\sigma_k(t) = \alpha_k t + \beta_k, \ k = 1, ..., d, \quad \textit{with} \ \ \alpha_k = \frac{(-1)^l M_{d,l}(\mu)}{\delta(\mu)},$$

$$\beta_k = \frac{(-1)^l}{\delta(\mu)} \left(\mu_d \cdot M_{1,l}(\mu) - \mu_{d+1} \cdot M_{2,l}(\mu) + ... + (-1)^{d-2}\mu_{2d-2} \cdot M_{d-1,l}(\mu)\right),$$

*where $l = d - k + 1$, and $M_{i,j}(\mu)$ denotes the minor obtained by omitting the $i$-th row and $j$-th column of $M_d(\mu)$.*

Notice that by construction, the roots of $Q_t(z)$ are always the nodes $x_1, ..., x_d$ in the solution $F(t)$ of the original Prony system (1.1), with the right hand side $(\mu_0, \ldots, \mu_{2d-2}, t)$.

To get real and distinct nodes $x_1, ..., x_d$ we should take only those values of $t$ for which $Q_t(z) \in H_d$. Thus, we define $A_\mu \subseteq \mathbb{R}$ as the set of all $t \in \mathbb{R}$ for which $Q_t(z) \in H_d$. $A_\mu$ is a finite union of open intervals in $\mathbb{R}$.

It is important to accurately describe the behavior of the nodes and of the amplitudes of $F(t)$ along the Prony curve $S(\mu)$, in terms of the *known "measurements"* $\mu$. In [20] it is explained how this information can help us to improve the reconstruction accuracy. The following two results in this direction were announced in [20] without proof:

1. Assume that $\delta(\mu) \neq 0$. Then if two nodes $x_i(t), x_{i+1}(t)$ of $F(t)$ collide as $t \to t_0$, then both the amplitudes $a_i(t), a_{i+1}(t)$ tend to infinity when $t \to t_0$.
2. Assume that $\delta(\mu) \neq 0$, and that the upper left $(d-1) \times (d-1)$ minor of $M_d(\mu)$ is non-degenerate. Then for $t \to \pm\infty$ at most one node of $F(t)$ can tend to infinity.

These results will follow from significantly more accurate results of Theorem 2.4, Theorem 2.5, and Proposition 2.3 which we prove below.

**2.4.1. Behavior of the amplitudes as the nodes near-collide.** In the study of the nodes collisions it is convenient to slightly change the initial setting of the problem, and to consider *unordered, and possibly colliding* nodes $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$.

Denote by $\Sigma$ the "diagonal" in $\mathbb{R}^d$ consisting of all $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ with at least two coordinates equal. For $s = 2, \ldots, d$, let $J_s$ be the set of all the $s$-tuples

$$J = \{1 \leq r_1 < \ldots < r_s \leq d\}.$$

We denote by $J_s^i$ the subset of $J_s$ consisting of the $s$-tuples $J$ with $i \in J$.

For $J \in J_s$, let

$$\Sigma(J) = \{x = (x_1, \ldots, x_d) \in \mathbb{R}^d, \ x_{r_1} = x_{r_2} = \ldots = x_{r_s}\}$$

be the subspace in $\mathbb{R}^d$ where the corresponding $s$ coordinates coincide. Then we have

$$\Sigma = \cup_{s, J \in J_s} \Sigma(J). \tag{2.10}$$

We also define the set of all $s$-collisions as $\Sigma_s = \cup_{J \in J_s} \Sigma(J)$, and the set of all $s$-collisions, which include the $i$-th coordinate, as $\Sigma_s^i = \cup_{J \in J_s^i} \Sigma(J)$.

Previously we considered the Prony varieties $S_q^x(\mu)$ only inside $\Delta_d$, which is one of the components of $\mathbb{R}^d \setminus \Sigma$. But equations (2.2) define $S_q^x(\mu)$ in the entire space $\mathbb{R}^d$, and by Theorem 2.2 and Theorem 2.3 we conclude that all the singularities of $S_q^x(\mu)$ in $\mathbb{R}^d$ are contained in the diagonal $\Sigma$.

Notice that the points of $S_q^x(\mu) \cap \Sigma$ may be non-singular points of $S_q^x(\mu)$: compare Section 3 below. Another remark is that permutations of the nodes preserve the Prony varieties, and hence the study of their "out-of-collisions" part can be restricted to $\Delta_d$.

Now we consider Prony curves $S(\mu) = S_{2d-2}(\mu)$ and for $F = (a, x) \in S(\mu)$ describe the behavior of the amplitudes $a = (a_1(x), \ldots, a_d(x))$ as the nodes $x = (x_1, \ldots, x_d) \in S^x(\mu)$ approach the diagonal $\Sigma_d$.

For given $i$ and $s$ with $1 \le i, s \le d$, and for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, let $dist_s^i(x)$ denote the distance of $x$ to the stratum $\Sigma_s^i$ of the diagonal $\Sigma$ (we work with the Euclidean norm in $\mathbb{R}^d$).

Essentially, $dist_s^i(x)$ estimates the minimal size of a cluster of $s$ nodes $x_j$, containing the node $x_i$. Indeed, we have the following result:

LEMMA 2.1. *For given $i$ and $s$ with $1 \le i, s \le d$ and for any $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ there exist pairwise different, and different from $i$, indices $l_1, \ldots, l_{s-1}$ such that*

$$|x_{l_j} - x_i| \le 2\ dist_s^i(x),\ j = 1, \ldots, s-1.$$

*Proof.* By definition, there is a point $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_d) \in \Sigma_s^i$ such that $||x - \bar{x}|| = dist_s^i(x)$. In particular, we have $|x_j - \bar{x}_j| \le dist_s^i(x),\ j = 1, \ldots, d$. By definition, there are indices $J = \{1 \le r_1 < \ldots < r_s \le d\}$, one of them is equal to $i$, such that $\bar{x}_{r_1} = \bar{x}_{r_2} = \ldots = \bar{x}_{r_s}$. Re-denote by $l_j,\ j = 1, \ldots, s-1$, the indices $r_m$, different from $i$. Then for each $l_j$, we have that

$$|x_{l_j} - x_i| \le |x_{l_j} - \bar{x}_{l_j}| + |x_i - \bar{x}_i| \le 2\ dist_s^i(x),$$

since $\bar{x}_{l_j} = \bar{x}_i$. This completes the proof. □

Let us recall that by Proposition 2.1 the amplitudes $a_i$ on $S^x(\mu)$ are uniquely expressed through $x$ as

$$a_i(x) = \frac{P(\pi_i(x))}{L_i(x)}, \tag{2.11}$$

where for $u = (u_1, \ldots, u_{d-1})$ the polynomial $P$ is defined by

$$P(u) = \mu_0 \varrho_{d-1}(u) + \ldots + \mu_{d-2}\varrho_1(u) + \mu_{d-1}, \tag{2.12}$$

and $L_i(x) = \prod_{1 \le l \le d,\ l \ne i}(x_i - x_l)$.

Finally, for $d, \mu$ as above, and for a given $D > 1$, put

$$C_1 = \frac{1}{2^d D^{2d-s-1}\sqrt{d}},\ C_2 = \frac{1}{D^{d-1}\sqrt{d}},\ C_3 = \nu_{d-1}(1+D)^{d-1},$$

with $\nu_{d-1} = \max_{k=0,\ldots,d-1}|\mu_k|$.

Now we have all the definitions and preliminary facts required in order to state and prove our result on the behavior of the amplitudes near the nodes collision point on the Prony curve.

THEOREM 2.4. *Let $\mu \in \mathcal{M}_d$ be such that $\delta(\mu) \ne 0$. Then for each $x = (x_1, \ldots, x_d) \in S^x(\mu)$, with $|x_j| \le D,\ j = 1, \ldots, d$, for each $i = 1, \ldots, d$, and for each $s = 2, \ldots, d$, we have*

$$\frac{C_1 \eta(\mu)}{(dist_s^i(x))^s} \le \frac{C_2 \eta(\mu)}{L_i(x)} \le |a_i(x)| \le \frac{C_3}{L_i(x)}. \tag{2.13}$$

*In particular, if $x \in S^x(\mu)$ tends to $\bar{x} \in \Sigma(J)$ with $J \in J_s$, then for each $i \in J$, the amplitude $a_i(x)$ tends to infinity at least as fast as $\frac{C_1\eta(\mu)}{||x-\bar{x}||^s}$.*

*Proof.* Starting with (2.11) we bound $a_i(x)$ estimating the numerator and the denominator in this expression. For fixed $i$ and $s$, and for any $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, $|x_j| \le D$, (not only for $x \in S^x(\mu)$), we have the following inequality:

$$|L_i(x)| \le (2D)^{d-s}[2\ dist_s^i(x)]^s = 2^d D^{d-s}(dist_s^i(x))^s. \tag{2.14}$$

Indeed, by Lemma 2.1, some $s-1$ factors in $L_i(x)$ do not exceed $2\ dist_s^i(x)$, while the remaining factors are bounded by $2D$. Accordingly, it remains only, for any $x \in S^x(\mu)$, to bound the polynomial $P(\pi_i(x))$ from above and from below.

From (2.12) and from the assumption that $|x_j| \le D$ we get, denoting, as above, by $\nu_{d-1}$ the maximum $\max_{0 \le k \le d-1} |\mu_k|$,

$$|P(\pi_i(x))| \le \nu_{d-1} \sum_{k=0}^{d} \binom{d-1}{k} D^k = \nu_{d-1}(1+D)^{d-1}. \tag{2.15}$$

To prove the lower bound for $P(\pi_i(x))$ on the Prony curve $S^x(\mu)$, we add $P$ to the system of equations (2.2), defining $S(\mu)$, and transform the resulting system into a convenient form.

Let us start with a simple identity for the symmetric polynomials: we fix $i$ and denote $\pi_i(x)$ by $x^*$. Then

$$\sigma_k(x) = -\varrho_{k-1}(x^*)x_i + \varrho_k(x^*),\ k = 1, \dots, d.$$

Recall that $\varrho_0 \equiv 1,\ \varrho_d \equiv 0$. We use the shorthand $\varrho_k$ for $\varrho_k(x^*)$. System (2.2), defining, for $q = 2d-2$, the Prony curve $S^x(\mu) = S^x_{2d-2}(\mu)$, can be now rewritten as (*):

$$\mu_0(\varrho_{d-1}x_i - \varrho_d) + \mu_1(\varrho_{d-2}x_i - \varrho_{d-1}) + \dots + \mu_{d-1}(\varrho_0 x_i - \varrho_1) - \mu_d = 0$$
$$\mu_1(\varrho_{d-1}x_i - \varrho_d) + \mu_2(\varrho_{d-2}x_i - \varrho_{d-1}) + \dots + \mu_d(\varrho_0 x_i - \varrho_1) - \mu_{d+1} = 0$$
$$\dots\dots\dots\dots$$
$$\mu_{d-2}(\varrho_{d-1}x_i - \varrho_d) + \mu_{d-1}(\varrho_{d-2}x_i - \varrho_{d-1}) + \dots + \mu_{2d-3}(\varrho_0 x_i - \varrho_1) - \mu_{2d-2} = 0$$

Let us now assume that $P(x^*) = \gamma$, or, equivalently,

$$\mu_0\varrho_{d-1} + \mu_1\varrho_{d-2} + \dots + \mu_{d-2}\varrho_1 + \mu_{d-1}\varrho_0 = \gamma$$

Multiplying this equation by $x_i$ and subtracting it from the first equation in the system (*), we get all the terms containing the product with $x_i$ cancelled, and obtain a new equation

$$\mu_1\varrho_{d-1} + \mu_2\varrho_{d-2} + \dots + \mu_{d-1}\varrho_1 + \mu_d\varrho_0 = \gamma x_i.$$

Multiplying this new equation by $x_i$ and subtracting it from the second equation in the system (*), we get all the terms containing the product with $x_i$ cancelled, and obtain the next equation

$$\mu_2\varrho_{d-1} + \mu_3\varrho_{d-2} + \dots + \mu_d\varrho_1 + \mu_{d+1}\varrho_0 = \gamma x_i^2.$$

Continuing in this way, we obtain the following system:

$$\begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_{d-2} & \mu_{d-1} \\ \mu_1 & \mu_2 & \cdots & \mu_{d-1} & \mu_d \\ & \cdots\cdots\cdots & & & \\ \mu_{d-1} & \mu_d & \cdots & \mu_{2d-3} & \mu_{2d-2} \end{pmatrix} \begin{pmatrix} \varrho_{d-1} \\ \varrho_{d-2} \\ \cdots\cdots \\ \varrho_1 \\ 1 \end{pmatrix} = \begin{pmatrix} \gamma \\ \gamma x_i \\ \cdots \\ \gamma x_i^{d-2} \\ \gamma x_i^{d-1} \end{pmatrix}.$$

Since, by our assumption, the minimal singular value of the matrix $M_d(\mu)$ in the left-hand side of the last system is $\eta(\mu)$, we conclude that the norm of the right-hand side is at least $\eta(\mu)$:

$$|\gamma|(\sum_{j=0}^{d-1} x_i^{2j})^{\frac{1}{2}} \geq \eta(\mu), \text{ or } |\gamma| \geq \frac{\eta(\mu)}{(\sum_{j=0}^{d-1} x_i^{2j})^{\frac{1}{2}}} \geq \frac{\eta(\mu)}{D^{d-1}\sqrt{d}}. \tag{2.16}$$

The last inequality follows since $|x_i| \leq D$ by assumptions. Thus, by inequalities (2.14), (2.15) and (2.16), we have

$$|L_i(x)| \leq 2^d D^{d-s}(dist_s^i(x))^s, \quad \frac{\eta(\mu)}{D^{d-1}\sqrt{d}} \leq |P(\pi_i(x))| \leq \nu_{d-1}(1+D)^{d-1},$$

which implies, via (2.11)

$$\frac{\eta(\mu)}{2^d D^{2d-s-1}\sqrt{d}(dist_s^i(x))^s} \leq \frac{\eta(\mu)}{D^{d-1}\sqrt{d}L_i(x)} \leq |a_i(x)| \leq \frac{\nu_{d-1}(1+D)^{d-1}}{L_i(x)}.$$

This completes the proof of Theorem 2.4. ☐

**2.4.2. Escape of nodes to infinity.** In Proposition 2.2 above we describe a natural parametrization of the Prony curves $S^x(\mu)$ with non-degenerate $M_d(\mu)$. It goes via the root mapping $RM$ restricted to the intersection $L^h(\mu)$ of the line $L(\mu)$ with the hyperbolic set $H_d \subset \mathcal{V}_d$. In turn, the polynomials $Q_t(z) = z^d + \sigma_1(t)z^{d-1} + \ldots + \sigma_d(t)$ on the line $L(\mu)$ are given by

$$\sigma_k(t) = \alpha_k t + \beta_k, \ k = 1, ..., d, \tag{2.17}$$

with $t = \mu_{2d-1}$. The explicit expressions through $\mu_0, \ldots, \mu_{2d-2}$ for $\alpha_k, \beta_k$ are given in Proposition 2.2.

The set $A_\mu \subseteq \mathbb{R}$ was defined as the set of all $t \in \mathbb{R}$ for which $Q_t(z) \in H_d$. $A_\mu$ is a finite union of open intervals in $\mathbb{R}$. The expression $(x_1(t), \ldots, x_d(t)) = RM(Q_t)$ provides a diffeomorphic parametrization of the Prony curve $S^x(\mu)$, for $t \in A_\mu$. (See Section 3 for examples in the case of two nodes.)

THEOREM 2.5. *Let $\mu \in \mathcal{M}_d$ be such that the moment Hankel matrix $M_d(\mu)$ is non-degenerate, as well as its top-left $(d-1) \times (d-1)$ minor. Then for $t \to \pm\infty$ inside $A_\mu$, at most one node among $x_1(t), \ldots, x_d(t)$ can tend to infinity.*

*Proof.* By our assumptions, the parametrization (2.17) above is applicable, and, by Proposition 2.2, we have $\alpha_1 \neq 0$ in (2.17). Hence the required result is implied directly by the following statement (where we do not insist on all the roots of $Q_t(z)$ being real):

Proposition 2.3. *Let* $Q_t(z) = z^d + \sigma_1(t)z^{d-1} + \ldots + \sigma_d(t)$, $\sigma_k(t) = \alpha_k t + \beta_k$, *be a polynomial pencil, with* $a_1 \neq 0$. *There are positive constants* $t_0, t_1, \lambda_0, A_1 < A_2$ *(defined in the proof below) such that for* $\alpha_1 > 0$ *and for* $t \geq t_0$, *the polynomial* $Q_t(z)$ *has no real roots in the interval* $[\lambda_0, \infty)$.

*For* $\alpha_1 < 0$ *and for* $t \geq t_0$, *the polynomial* $Q_t(z)$ *has no real roots in the intervals* $[\lambda_0, A_1 t]$ *and* $[A_2 t, \infty)$ *and exactly one real root in the interval* $[A_1 t, A_2 t]$.

*Proof.* We use a version of the Descartes rule of signs, usually called the Budan-Fourier Theorem. Let $\nu_Q(\lambda)$ denote the number of sign changes in the sequence $(Q(\lambda), Q'(\lambda), ..., Q^{(n)}(\lambda))$ (the number of sign changes in a sequence of real numbers is counted with zeroes omitted).

Lemma 2.2. *For any polynomial* $Q(z)$ *of degree* $d$ *and* $a < b$, *the number of zeros of* $Q$ *(counted with their multiplicities) in the interval* $(a, b]$ *is less than or equal to* $\nu_Q(a) - \nu_Q(b)$, *and differs from it by an even number.*

For the derivatives $Q_t^{(r)}(\lambda)$, $r = 0, 1, \ldots, d$, of $Q_t(\lambda)$, denoting for $n \geq m$ by $A(n, m)$ the product $(n - m + 1)(n - m + 2) \ldots (n - 1)n$, we have

$$Q_t^{(r)}(\lambda) = A(d, r)\lambda^{d-r} + \sum_{k=1}^{d-r} A(d - k, r)(\alpha_k t + \beta_k)\lambda^{d-k-r}.$$

To make the computations more transparent, we normalize the derivatives, and estimate the signs of $\hat{Q}_t^{(r)}(\lambda) = \frac{Q_t^{(r)}(\lambda)}{A(d,r)\lambda^{d-r}}$, for which we obtain

$$\hat{Q}_t^{(r)}(\lambda) = 1 + \frac{C(r)t\alpha_1}{\lambda}\left(1 + \frac{\beta_1}{\alpha_1 t} + \sum_{k=2}^{d-r} \frac{B(k, r)}{\lambda^{k-1}}\left(\frac{\alpha_k}{\alpha_1} + \frac{\beta_k}{\alpha_1 t}\right)\right), \tag{2.18}$$

where $C(r) = \frac{A(d-1,r)}{A(d,r)} = \frac{d-r}{d}$, $B(k, r) = \frac{A(d-k,r)}{A(d-1,r)}$, for $r = 0, \ldots, d - 1$. Notice that $C(d) = 0$, and hence $\hat{Q}_t^{(d)}(\lambda) = 1$. Recall also that by assumptions $\alpha_1 \neq 0$.

Next we assume that

$$t \geq t_0 := 20 \max{}_{k=1,\ldots,d} \left|\frac{\beta_k}{\alpha_1}\right|, \quad \lambda \geq \lambda_0 := 20d \max{}_{k=2,\ldots,d,\ r=0,\ldots,d} B(k, r)\left(\left|\frac{\alpha_k}{\alpha_1}\right| + \frac{1}{20}\right).$$

Lemma 2.3. *Under these assumptions we have*

$$\hat{Q}_t^{(r)}(\lambda) = 1 + \frac{C(r)t\alpha_1}{\lambda}(1 + \kappa(r)), \ \textit{with}\ |\kappa(r)| \leq \frac{1}{10}, \ r = 0, \ldots d. \tag{2.19}$$

*Proof.* The required expression follows from (2.18) with

$$\kappa = \frac{\beta_1}{\alpha_1 t} + \sum_{k=2}^{d-r} \frac{B(k, r)}{\lambda^{k-1}}\left(\frac{\alpha_k}{\alpha_1} + \frac{\beta_k}{\alpha_1 t}\right).$$

Hence, by the assumptions on $t$ and $\lambda$ we have

$$|\kappa| \leq \frac{1}{20} + \frac{1}{\lambda}\sum_{k=2}^{d-r} B(k, r)\left(\left|\frac{\alpha_k}{\alpha_1}\right| + \frac{1}{20}\right) \leq \frac{1}{10}. \quad \square$$

Thus, we have to count the number of sign changes in sequence (2.19), with $r = 0, \ldots, d$, for different $\lambda$.

Assume first that $\alpha_1 > 0$, in which case there are no sign changes in (2.19), and we conclude for each $\lambda \geq \lambda_0$ and $t \geq t_0$ that we have $\nu_{Q_t}(\lambda) = 0$. Therefore $Q_t(z)$ does not have real roots on $[\lambda_1, \infty)$.

Now we consider the case $\alpha_1 < 0$. Put $c_1(d) = \min_{r=0,\ldots,d-1} C(r)$, $c_2(d) = \max_{r=0,\ldots,d-1} C(r)$, and define $t_1 := \frac{2\lambda_0}{c(d)}$, $A_1 := \frac{1}{2}c_1(d)|\alpha_1|$, $A_2 := 2c_2(d)|\alpha_1|$.

We get immediately that for each $t \geq t_1$,

$$\frac{C(r)t\alpha_1}{\lambda_0}(1 + \kappa(r)) < -1, \; r = 0, \ldots d - 1,$$

and for $\lambda_1(t) = A_1 t, \; \lambda_2(t) = A_2 t$,

$$\frac{C(r)t\alpha_1}{\lambda_1(t)}(1 + \kappa(r)) < -1, \; \frac{C(r)t\alpha_1}{\lambda_2(t)}(1 + \kappa(r)) > -\frac{1}{2}, \; r = 0, \ldots d - 1.$$

By Lemma 2.3 we get $Q_t^{(r)}(\lambda_0), Q_t^{(r)}(\lambda_1(t)) < 0$, $r = 0, \ldots, d-1$, while we have $Q_t^{(d)} \equiv 1$, and hence $\nu_{Q_t}(\lambda_1(t)) = 1$. For $\lambda_2(t)$ we obtain all $Q_t^{(r)}(\lambda_2(t))$, $r = 0, \ldots, d$, positive, i.e. $\nu_{Q_t}(\lambda_2(t)) = 0 = \nu_{Q_t}(\infty)$. Therefore there are no real roots of $Q_t(z)$ between $\lambda_0$ and $\lambda_1(t)$, there is exactly one real root of $Q_t(z)$ between $\lambda_1(t)$ and $\lambda_2(t)$, and no real roots in $[\lambda_2(t), \infty)$. This completes the proof of Proposition 2.3 and of Theorem 2.5. □

Remark. The cases $t \to -\infty$ and/or $z \to -\infty$ are reduced to the above case by the substitutions $\tau = -t$ and $w = -z$, respectively.

**3. Prony varieties for two nodes.** In this section we illustrate some of the above results, providing a complete description of the Prony varieties in the case of two nodes, i.e., for $d = 2$. Set $\mu = (\mu_0, \mu_1, \mu_2, \mu_3)$.

For $q = 0$, the varieties $S_0(\mu)$ are three-dimensional hyperplanes in $\mathcal{P}_2 \cong \mathbb{R}^4$, defined by the equation $a_1 + a_2 = \mu_0$.

For $q = 1 = d - 1$, the varieties $S_1(\mu)$ are two-dimensional subvarieties in $\mathcal{P}_2$, defined by the equations

$$a_1 + a_2 = \mu_0, \quad a_1 x_1 + a_2 x_2 = \mu_1. \tag{3.1}$$

This gives

$$a_1 = \frac{\mu_0 x_2 - \mu_1}{x_2 - x_1}, \; a_2 = \frac{-\mu_0 x_1 + \mu_1}{x_2 - x_1}, \tag{3.2}$$

which is a special case of expressions (2.5) for $d = 2$.

Consider now the case $q = 2 = d = 2d - 2$. Here the varieties $S_2(\mu)$ are (generically) algebraic curves in $\mathcal{P}_2$, defined by the equations

$$a_1 + a_2 = \mu_0, \quad a_1 x_1 + a_2 x_2 = \mu_1, \quad a_1 x_1^2 + a_2 x_2^2 = \mu_2, \tag{3.3}$$

For the corresponding curve $S_2^x(\mu)$, in the node space $\mathcal{P}_2^x \cong \Delta_2$ we obtain from Theorem 2.3 the equation $\mu_1\sigma_1 + \mu_0\sigma_2 = -\mu_2$, or, equivalently,

$$\mu_0 x_1 x_2 - \mu_1(x_1 + x_2) + \mu_2 = 0. \tag{3.4}$$

This equation leads to three different possibilities:

1. If $\mu_0 \neq 0$, then the curve $S_2^x(\mu)$ is a hyperbola

$$(x_1 - \frac{\mu_1}{\mu_0})(x_2 - \frac{\mu_1}{\mu_0}) + \frac{\mu_0\mu_2 - \mu_1^2}{\mu_0^2} = 0, \tag{3.5}$$

which is non-singular for $\mu_0\mu_2 - \mu_1^2 \neq 0$, and degenerates into two orthogonal coordinate lines, crossing at the diagonal $\{x_1 = x_2\}$, for $\mu_0\mu_2 - \mu_1^2 = 0$.

2. If $\mu_0 = 0$, but $\mu_1 \neq 0$, then the curve $S_2^x(\mu)$ is a straight line

$$x_1 + x_2 = \mu_2/\mu_1. \tag{3.6}$$

3. Finally, if $\mu_0 = \mu_1 = 0$, but $\mu_2 \neq 0$, then the curve $S_2^x(\mu)$ is empty, and for $\mu_0 = \mu_1 = \mu_2 = 0$, it coincides with the entire plane $\mathcal{P}_2^x$.

It is instructive to interpret the cases (1-3) above in terms of the relative position of the straight line $L_2(\mu)$, with respect to the set $H_2$ of hyperbolic polynomials $Q$. This line is defined in the space $\mathcal{V}_2$ of polynomials $Q(z) = z^2 + \sigma_1 z + \sigma_2$ by system (2.2), i.e., by the equation $\mu_1\sigma_1 + \mu_0\sigma_2 = -\mu_2$. Figure 1 illustrates possible positions of the line $L_2(\mu)$ with respect to the set $H_2$ of hyperbolic polynomials.

The discriminant $\Delta(\sigma_1, \sigma_2) = \sigma_1^2 - 4\sigma_2$ of $Q(z) = z^2 + \sigma_1 z + \sigma_2$ is positive for $Q \in H_2$. Therefore $H_2$ is the portion under the parabola $P = \{\sigma_2 = \frac{1}{4}\sigma_1^2\}$ in $\mathcal{V}_2$. (Compare Figure 1). The case $\mu_0 \neq 0$ corresponds to the lines $L_2(\mu)$, nonparallel to the $\sigma_2$-axis of $\mathcal{V}_2$. These lines may cross the parabola $P$ at two points (line $l_1$ on Figure 1), at one point, if tangent to $P$ (line $l_2$ on Figure 1), or they may not cross $P$ at all, and then they are entirely contained in $H_2$ (line $l_3$ on Figure 1). These cases correspond to $\mu_0\mu_2 - \mu_1^2 < 0$, $\mu_0\mu_2 - \mu_1^2 = 0$ and $\mu_0\mu_2 - \mu_1^2 > 0$, respectively.

The corresponding Prony curves $S^x(\mu)$, which are the images of the lines $L_2(\mu)$ intersected with $H_2$, under the root map $RM$, are shown on the bottom part of Figure 1.

For the line $L_2(\mu)$ crossing the parabola $P$ at two points (like $l_1$ on Figure 1), the corresponding hyperbola $S_2^x(\mu)$ crosses the diagonal in the plane $\mathcal{P}_2^x$, i.e., it contains a collision of the nodes $x_1, x_2$.

Notice that the Prony curve $S^x(\mu)$ remains non-singular at the crossing point. This fact holds also for the general case of "double collisions" on the Prony curves, and we plan to present it separately.

For the line $L_2(\mu)$ tangent to the parabola $P$ (like $l_2$ on Figure 1), the corresponding hyperbola $S_2^x(\mu)$ degenerates into two orthogonal coordinate lines, crossing at a certain point on the diagonal $\{x_1 = x_2\}$.

For the line $L_2(\mu)$ entirely contained in $H_2$ (like $l_3$ on Figure 1), the corresponding hyperbola $S_2^x(\mu)$ does not cross the diagonal $\{x_1 = x_2\}$, and so it does not lead to the collision of the nodes.

For $\mu_0 = 0$, but $\mu_1 \neq 0$, the lines $L_2(\mu)$ are parallel to the $\sigma_2$-axis of $V_2$ (like $l_4$ on Figure 1). They cross the parabola $P$ at exactly one point. The corresponding curve $S_2^x(\mu)$ for $l_4$ is a straight line $x_1 + x_2 = -\frac{\mu_2}{\mu_1}$.

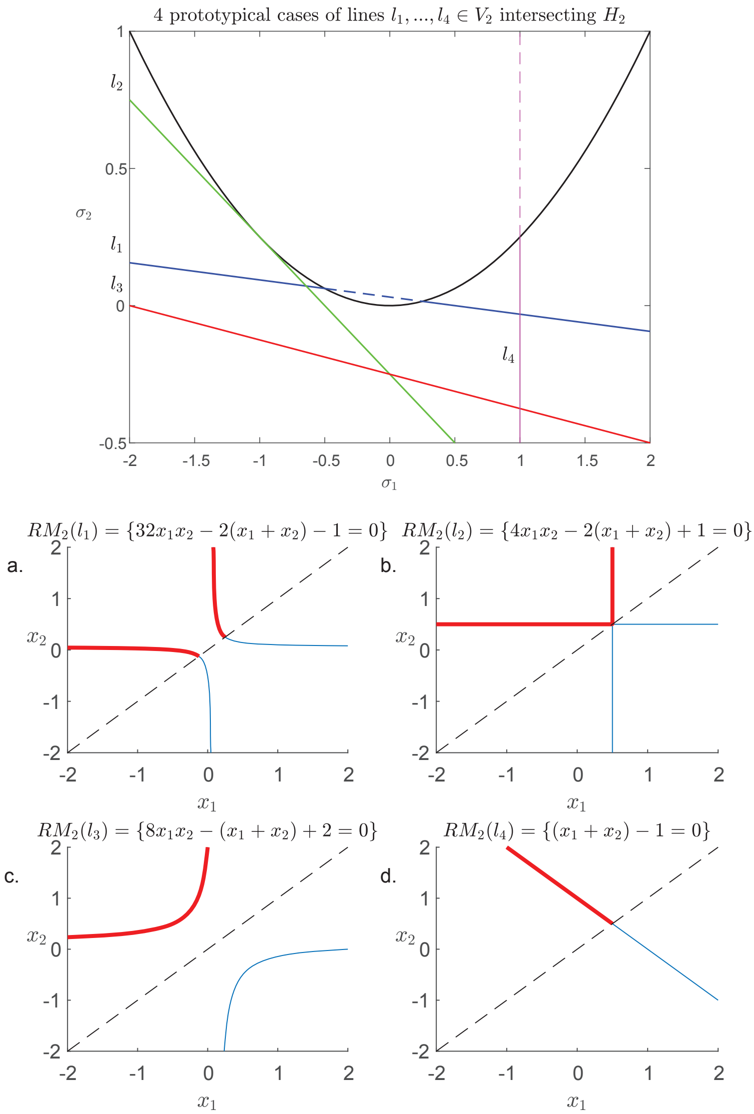


FIG. 1. *Visualized is the diffeomorphism* $RM : H_2 \to \mathcal{P}_2^x$ *acting on 4 prototypical lines* $l_1,\ l_2\ ,l_3\ ,l_4 \in \mathcal{V}_2$ *intersected with* $H_2$ *(the open set outside the parabola on the upper figure). In the bottom figure, the highlighted parts in the subplots a,b,c,d are the images, under* $RM_2$*, of* $l_1,\ l_2\ ,l_3\ ,l_4$ *intersected with* $H_2$*, respectively.*

## REFERENCES


[1] A. Akinshin, D. Batenkov, and Y. Yomdin, *Accuracy of spike-train Fourier reconstruction for colliding nodes*, in "2015 International Conference on Sampling Theory and Applications (SampTA)", pp. 617–621. IEEE, 2015.

[2] A. Akinshin, G. Goldman, V. Golubyatnikov, and Y. Yomdin, *Accuracy of reconstruction of spike-trains with two near-colliding nodes*, in "Proc. Complex Analysis and Dynamical Systems VII", volume 699, pp. 1–17. The AMS and Bar-Ilan University, 2015.

[3] A. Akinshin, G. Goldman, and Y. Yomdin, *Geometry of error amplification in solving Prony system with near-colliding nodes*, arXiv preprint arXiv:1701.04058, 2017.

[4] A. Akinshin, V. Golubyatnikov, and Y. Yomdin, *Low-dimensional Prony systems*, in Proc. International Conference "Lomonosov readings in Altai: fundamental problems of science and education", pp. 443–450, 20–24 October 2015.

[5] V. I. Arnol'd, *Hyperbolic polynomials and Vandermonde mappings*, Functional Analysis and Its Applications, 20:2 (1986), pp. 125–127.

[6] J. R. Auton, M. L. Van Blaricum, et al, *Investigation of procedures for automatic resonance extraction from noisy transient electromagnetics data*, AFWL Math. Note 7, General Research Corp Santa Barbara, Calif, 1981.

[7] J.-M. Azais, Y. De Castro, and F. Gamboa, *Spike detection from inaccurate samplings*, Applied and Computational Harmonic Analysis, 38:2 (2015), pp. 177–195.

[8] D. Batenkov, *Stability and super-resolution of generalized spike recovery*, Applied and Computational Harmonic Analysis, 2016.

[9] D. Batenkov, *Accurate solution of near-colliding Prony systems via decimation and homotopy continuation*, Theoretical Computer Science, 2017.

[10] D. Batenkov and Y. Yomdin, *On the accuracy of solving confluent Prony systems*, SIAM Journal on Applied Mathematics, 73:1 (2013), pp. 134–154.

[11] D. Batenkov and Y. Yomdin, *Geometry and singularities of the Prony mapping*, Journal of Singularities, 10 (2014), pp. 1–25.

[12] G. Beylkin and L. Monzón, *Approximation by exponential sums revisited*, Applied and Computational Harmonic Analysis, 28:2 (2010), pp. 131–149,.

[13] E. J. Candès and C. Fernandez-Granda, *Super-resolution from noisy data*, Journal of Fourier Analysis and Applications, 19:6 (2013), pp. 1229–1254.

[14] E. J. Candès and C. Fernandez-Granda, *Towards a mathematical theory of super-resolution*, Communications on Pure and Applied Mathematics, 67:6 (2014), pp. 906–956.

[15] L. Demanet, D. Needell, and N. Nguyen, *Super-resolution via superset selection and pruning*, arXiv preprint arXiv:1302.6288, 2013.

[16] L. Demanet and N. Nguyen, *The recoverability limit for superresolution via sparsity*, arXiv preprint arXiv:1502.01385, 2015.

[17] D. L. Donoho, *Superresolution via sparsity constraints*, SIAM journal on mathematical analysis, 23:5 (1992), pp. 1309–1331.

[18] D. Eisenbud, *Linear sections of determinantal varieties*, American Journal of Mathematics, 110 (1988), pp. 541–575.

[19] C. Fernandez-Granda, *Super-resolution of point sources via convex programming*, Information and Inference: A Journal of the IMA, 5:3 (2016), pp. 251–303.

[20] G. Goldman, Y. Salman, and Y. Yomdin, *Accuracy of noisy spike-train reconstruction: a singularity theory point of view*, Journal of Singularities to appear, arXiv preprint arXiv:1801.02177.

[21] G. Goldman and Y. Yomdin, *On algebraic properties of low rank approximations of Prony systems*, arXiv preprint arXiv:1803.09243, 2018.

[22] V. P. Kostov, *On the geometric properties of Vandermonde's mapping and on the problem of moments*, Proceedings of the Royal Society of Edinburgh Section A: Mathematics, 112:3-4 (1989), pp. 203–211.

[23] V. P. Kostov, *Topics on hyperbolic polynomials in one variable*, Panoramas et Synthèses-Société Mathématique de France, 33 (2011).

[24] J. N. Mather, *Solutions of generic linear equations*, Dynamical Systems, pp. 185–193. Elsevier, 1973.

[25] V. I. Morgenshtern and E. J. Candès, *Super-resolution of positive sources: the discrete setup*, SIAM Journal on Imaging Sciences, 9:1 (2016), pp. 412–444.

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

[27] T. Peter, D. Potts, and M. Tasche, *Nonlinear approximation by sums of exponentials and translates*, SIAM Journal on Scientific Computing, 33:4 (2011), pp. 1920–1947.

[28] G. Plonka and M. Wischerhoff, *How many Fourier samples are needed for real function reconstruction?*, Journal of Applied Mathematics and Computing, 42:1-2 (2013), pp. 117–137.

[29] D. Potts and M. Tasche, *Fast ESPRIT algorithms based on partial singular value decompositions*, Applied Numerical Mathematics, 88 (2015), pp. 31–45.

[30] R. Prony, *Essai experimental et analytique etc*, J. de l'Ecole Polytechnique, 1 (1975), pp. 24–76.

[31] L. R. Turner, *Inverse of the Vandermonde matrix with applications*, 1966.