The final version of this paper appears in [M.C. Beltrametti, J.R. Sendra, J. Sendra and M. Torrente. Applied Mathematics and Computation 375 (2020) 125083] and it is available at https://doi.org/10.1016/j.amc.2020.125083.

This article is dedicated to the memory of our friend and colleague Professor Mauro C. Beltrametti who passed away unexpectedly on November, 2019; he introduced us in the exciting field of the Hough Transform.

Moore–Penrose approach in the Hough transform framework

Abstract

Let $F(\boldsymbol{x}, \boldsymbol{a})$ be a real polynomial in two sets of variables, $\boldsymbol{x}$, $\boldsymbol{a}$. Assume that the polynomial is linear with respect to one set, say $\boldsymbol{a}$, of variables. Given a data set of points $\mathcal{D}$ in the real affine space $\mathbb{A}_{\boldsymbol{x}}$, we use the Moore–Penrose pseudo inverse real matrices approach to address the question of finding a point $\boldsymbol{a}^\star$ in the affine set $\mathbb{A}_{\boldsymbol{a}}$ of parameters such that the algebraic locus of equation $F(\boldsymbol{x}, \boldsymbol{a}^\star) = 0$ provides the best fit of $\mathcal{D}$. This allows us to deal with the questions of finding a region in the parameter space and bounding the sampling distance of the discretization of the region, which in turn consititute main issues in the Hough transform framework, a standard pattern recognition technique to detect loci in images.

keywords: Mutivariate polynomial, pesudo-inverse matrix, perturbed system, Hough transform, parameter region detection, parameter region discretization
MSC 2010: 15A09,14Q10, 68W30, 68T10

Introduction

Let $F(\boldsymbol{x}, \boldsymbol{a}) \in \mathbb{R}[\boldsymbol{x}, \boldsymbol{a}]$ be a real polynomial in two sets of variables $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{a} = (a_1, \ldots, a_t)$, let $\mathbb{A}^n_{\boldsymbol{x}}$ (the *image space*) and $\mathbb{A}^t_{\boldsymbol{a}}$ (the *parameter space*) be real affine spaces of coordinates $\boldsymbol{x}$ and $\boldsymbol{a}$, respectively. In this paper, we address the following question.

> Assume that we are given a data set $\mathcal{D}$ of N points in the image space $\mathbb{A}^n_{\boldsymbol{x}}$, with $N \gg t$, such that $F(\boldsymbol{x}_0, \boldsymbol{a})$ is not constant for every $\boldsymbol{x}_0 \in \mathcal{D}$. We then want to determine a point $\boldsymbol{a}^\star \in \mathbb{A}^t_{\boldsymbol{a}}$ such that the algebraic locus of equation $F(\boldsymbol{x}, \boldsymbol{a}^\star) = 0$ provides the best fit of $\mathcal{D}$.

A strong motivation comes from the Hough transform framework, a standard pattern recognition technique to detect curves in images. In this regard, specializing $\boldsymbol{x}$ and $\boldsymbol{a}$ in given points $p \in \mathbb{A}^n_{\boldsymbol{x}}$ and $\boldsymbol{\lambda} \in \mathbb{A}^t_{\boldsymbol{a}}$, respectively, the polynomial $F := F(\boldsymbol{x}, \boldsymbol{a})$ gives

rise to polynomials $f_p(\boldsymbol{a}) := F(p, \boldsymbol{a}) \in \mathbb{R}[\boldsymbol{a}]$ and $f_{\boldsymbol{\lambda}}(\boldsymbol{x}) := F(\boldsymbol{x}, \boldsymbol{\lambda}) \in \mathbb{R}[\boldsymbol{x}]$, whence one has two families of zero loci

$$\mathcal{F} := \{\Gamma_{\boldsymbol{\lambda}} := f_{\boldsymbol{\lambda}}(\boldsymbol{x}) = 0\}_{\boldsymbol{\lambda} \in \mathbb{A}^t_{\boldsymbol{a}}} \subset \mathbb{A}^n_{\boldsymbol{x}} \quad \text{and} \quad \mathcal{H} := \{\Gamma_p := f_p(\boldsymbol{a}) = 0\}_{p \in \mathbb{A}^n_{\boldsymbol{x}}} \subset \mathbb{A}^t_{\boldsymbol{a}}.$$

For sufficiently general points $p \in \mathbb{A}^n_{\boldsymbol{x}}$ and $\boldsymbol{\lambda} \in \mathbb{A}^t_{\boldsymbol{a}}$, it turns out that $\Gamma_{\boldsymbol{\lambda}}$ and Γ_p are hypersurfaces in $\mathbb{A}^n_{\boldsymbol{x}}$ and $\mathbb{A}^t_{\boldsymbol{a}}$, respectively. Moreover, according to the Hough transform terminology, Γ_p is called *Hough transform of p w.r.t. the family $\mathcal{F}$*.

Let's highlight the Hough transform technique. Given a profile of interest $\mathcal{P}$ in the image space (typically, either a curvilinear profile in $\mathbb{A}^2_{\boldsymbol{x}}$ or a superficial profile in $\mathbb{A}^3_{\boldsymbol{x}}$, pointed by a data set $\mathcal{D}$ of points), the aim is to detect a hypersurface from the family $\mathcal{F}$ best fitting the profile $\mathcal{P}$. A distinguished peculiarity of the Hough transform technique consists of being quite robust against the noise that, in general, affects the data set $\mathcal{D}$. Two crucial steps of the method are made up of finding a region T in the parameter space (where looking for the point $\boldsymbol{a}^\star \in \mathbb{A}^t_{\boldsymbol{a}}$ we aim to determine), and a discretization $\boldsymbol{\delta}$ of T (since, in most cases, it happens that $\cap_{p_j \in \mathcal{D}} \Gamma_{p_j} = \emptyset$). A voting procedure then provides the cell, or cells, of the discretization achieving the maximum number of crossing among the Hough transforms Γ_{p_j} of the points p_j belonging to $\mathcal{D}$. Depending on the goodness of the choice of both the region T and the discretization $\boldsymbol{\delta}$ (and, of course, of the family $\mathcal{F}$, that is, of the polynomial $F(\boldsymbol{x}, \boldsymbol{a})$), the center of the cell $\boldsymbol{a}^\star = (a_1^\star, \ldots, a_t^\star) \in \mathbb{A}^t_{\boldsymbol{a}}$ gives rise to the hypersurface $\Gamma_{\boldsymbol{a}^\star}$ from the family $\mathcal{F}$ best approximating the profile $\mathcal{P}$.

Indeed, under the crucial assumption we made all over throughout the paper that the polynomial $F(\boldsymbol{x}, \boldsymbol{a})$ is linear in the set of variables $\boldsymbol{a}$, if the points of the data set $\mathcal{D}$ we start from are either exact (that is, the data set $\mathcal{D}$ is given with absence of noise) or only affected by small perturbations, then the standard Moore–Penrose approach applies to answer the question posed above. However, in general, specially in applications, one cannot expect to receive the data set exactly. Thus, in this case, the Hough transform technique plays a crucial role.

The paper is organized as follows. In Section 1 we recall some basic results on Moore–Penrose pseudo-inverse real matrices. In Section 2 we address the question posed above assuming that the points of $\mathcal{D}$ are either exact or non-exact with no background noise. In Section 3, we consider the general and abstract setting made up of a multivariate polynomial $F := F(\boldsymbol{x}, \boldsymbol{a}) \in \mathbb{R}[\boldsymbol{x}, \boldsymbol{a}]$, $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{a} = (a_1, \ldots, a_t)$, linear in the variables $\boldsymbol{a}$, together with a given data set of exact points $\mathcal{D} \subset \mathbb{A}^n_{\boldsymbol{x}}$. Via the Moore–Penrose method we detect a region T in $\mathbb{A}^t_{\boldsymbol{a}}$ corresponding (w.r.t. the variety $V(F)$) to the data set $\mathcal{D}$. We also provide a goodness criterion for the region T (see Subsection 3.1). Section 4 is devoted to the more general non-exact case in presence of background noise. In such a generality the Moore–Penrose method is not anymore suitable to address the question of optimally fitting the data set $\mathcal{D}$, so one needs to employ the Hough transform technique. However, a specialization of our previous results discussed in Section 3 applies to detect a suitable region T in the

parameter space, offering this way an answer to one of the typical problems in the Hough transform setting. In Subsection 4.2 we validate the result providing as well illustrative examples (e.g., see Example 4.2). Finally, in Subsection 4.3, we consider a further task in the Hough transform framework, by suggesting an upper bound for the discretization of the region T needed to perform the voting procedure, as previously highlighted. Throughout the paper some illustrative examples are provided.

An up to date list of references covers the considered topics. The computations in this paper were performed using Maple 2018 and CoCoA [1].

1 Background material

In this section we fix some terminology and we recall general facts on the Moore–Penrose solution of a linear system. As far as the Hough transform framework is concerned we refer for instance to [21, Section 5], [11] and [17].

1.1 General setting

Let K be either the field of complex numbers $\mathbb{C}$ or the field of real numbers $\mathbb{R}$ (in the applications, $K = \mathbb{R}$). Let $F := F(\boldsymbol{x}, \boldsymbol{a}) \in K[\boldsymbol{x}, \boldsymbol{a}]$ be a polynomial in two set of variables $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{a} = (a_1, \ldots, a_t)$, let $\mathbb{A}^n_{\boldsymbol{x}}(K)$ and $\mathbb{A}^t_{\boldsymbol{a}}(K)$ be affine spaces over K of coordinates $\boldsymbol{x}$ and $\boldsymbol{a}$, and let $V(F)$ be the algebraic locus (the *incidence variety*) defined in $\mathbb{A}^n_{\boldsymbol{x}}(K) \times \mathbb{A}^t_{\boldsymbol{a}}(K)$ by the equation $F = 0$.

From now on, we will call $\mathbb{A}^n_{\boldsymbol{x}}(K)$ and $\mathbb{A}^t_{\boldsymbol{a}}(K)$ the *image space* and the *parameter space*, respectively. We will denote by $p = (x_1(p), \ldots, x_n(p)) \in \mathbb{A}^n_{\boldsymbol{x}}(K)$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_t) \in \mathbb{A}^t_{\boldsymbol{a}}(K)$ given points in $\mathbb{A}^n_{\boldsymbol{x}}(K)$ and $\mathbb{A}^t_{\boldsymbol{A}}(K)$, respectively. We put

$$f_p(a_1, \ldots, a_t) := F(x_1(p), \ldots, x_n(p)); a_1, \ldots, a_t) \in K[\boldsymbol{a}],$$

and

$$f_{\boldsymbol{\lambda}}(x_1, \ldots, x_n) := F(x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_t) \in K[\boldsymbol{x}].$$

We will assume that both $f_p(a_1, \ldots, a_t)$, $f_{\boldsymbol{\lambda}}(x_1, \ldots, x_n)$ are non-constant polynomials.

According to the nowadays well known terminology in the Hough transform setting (see [3], [4]), we say that the *Hough transform of p (w.r.t. $V(F)$)* is the locus

$$\Gamma_p := \{f_{\boldsymbol{p}}(a_1, \ldots, a_t) = 0\} \subset \mathbb{A}^t_{\boldsymbol{a}}(K). \tag{1}$$

We also set

$$\Gamma_{\boldsymbol{\lambda}} := \{f_{\boldsymbol{\lambda}}(x_1, \ldots, x_n) = 0\} \subset \mathbb{A}^n_{\boldsymbol{x}}(K),$$

and we consider the familly

$$\mathcal{F} := \{\Gamma_{\boldsymbol{\lambda}}\}_{\boldsymbol{\lambda} \in \mathbb{A}^t_{\boldsymbol{a}}(K)} \subset \mathbb{A}^n_{\boldsymbol{x}}(K). \tag{2}$$

If $K = \mathbb{C}$, the algebraic loci Γ_p and $\Gamma_{\boldsymbol{\lambda}}$ are of course hypersurfaces in $\mathbb{A}^t_{\boldsymbol{a}}(K)$ and $\mathbb{A}^n_{\boldsymbol{x}}(K)$, respectively. In the real case $K = \mathbb{R}$, the same is true for sufficiently general points in $\mathbb{A}^t_{\boldsymbol{a}}(K)$ and $\mathbb{A}^n_{\boldsymbol{x}}(K)$ (see [22, Prop. 2.25]).

Clearly, the following *duality condition* holds true

$$p \in \Gamma_{\boldsymbol{\lambda}} \iff f(p; \boldsymbol{\lambda}) = 0 \iff \boldsymbol{\lambda} \in \Gamma_p. \tag{3}$$

Corollary 1.1 *Let $\mathcal{F}$ be the familiy of hypersurfaces as above. Then for any fixed point $p \in \mathbb{A}^n_{\boldsymbol{x}}(K)$, one has*

$$\bigcap_{p \in \Gamma_{\boldsymbol{\lambda}}} \Gamma_p \ni \boldsymbol{\lambda}.$$

Proof. It immediately follows from condition (3).

From now on through the paper, unless otherwise specified, we assume that $K = \mathbb{R}$.

1.2 Moore–Penrose pseudo-inverse

Moore–Penrose pseudo-inverses were introduced simultaneously by Moore in [16], and by Penrose in [19]. Since then, this type of generalized inverses has been extensively studied and applied in many different contexts (see e.g. [5], [6], [14]). In this paper, we will use Moore–Penrose inverses of real matrices (for Moore–Penrose pseudo-inverses over other fields we refer to [23]), and we are specially interested in the Moore–Penrose inverse property of providing the best approximate solution of linear matrix equations (see [20] and [14, pp. 257–258]).

Let $A \in \mathcal{M}(m, n; \mathbb{R})$ be an $m \times n$ matrix with real entries. The *Moore–Penrose pseudo-inverse* $A^\dagger$ is the $n \times m$ matrix defined by the conditions

1. $AA^\dagger A = A$
2. $A^\dagger AA^\dagger = A^\dagger$
3. $(AA^\dagger)^T = AA^\dagger$
4. $(A^\dagger A)^T = A^\dagger A$

Moreover, in the case when $m \geq n$ and the matrix A is of full rank n (that is, A has linearly independent columns, the situation we will find in the applications), the matrix $A^\dagger$ can be expressed by the algebraic formula:

$$A^\dagger = (A^A)^{-1} A^T. \tag{4}$$

The pseudo-inverse matrix $A^\dagger$ provides the least squares solution to a system of linear equations

$$A\boldsymbol{x} = b$$

as follows (here $\boldsymbol{x}$, b denote column vectors $(x_1, \ldots, x_n)^T$, $(b_1, \ldots, b_m)^T$, respectively):

$$z = A^\dagger b. \tag{5}$$

Under the assumptions made, the system may have no solutions or, if it does, the solution is unique. The following inequality of the Euclidean distances holds true:

$$\|A\boldsymbol{x} - b\|_2 \geq \|Az - b\|_2, \tag{6}$$

for each $\boldsymbol{x} \in \mathbb{R}^n$ and $z = A^\dagger b$. In the sequel, we also refer to $z = A^\dagger b$ as the "Moore–Penrose solution" of $A\boldsymbol{x} = b$.

1.3 Perturbed linear system

In the sequel we use the following general fact. With the notation and assumptions as in Subsections 1.2, let $A\boldsymbol{x} = b$ be a linear system with $A \in \mathcal{M}(n \times n; \mathbb{R})$ of rank n and $b \in \mathcal{M}(n \times 1; \mathbb{R})$. Let

$$(A + \Delta A)(\boldsymbol{x} + \Delta\boldsymbol{x}) = b + \Delta b \tag{7}$$

be a perturbed linear system, where $\Delta A \in \mathcal{M}(n \times n; \mathbb{R})$ and $\Delta b \in \mathcal{M}(n \times 1; \mathbb{R})$ are perturbations of A and b, respectively, and $\boldsymbol{x} + \Delta\boldsymbol{x}$ denotes the solution of (7). Then, the following fact holds true (see [13, p. 383]).

Theorem 1.2 *let $A\boldsymbol{x} = b$ be a full rank square linear system, and let $(A + \Delta A)(\widetilde{\boldsymbol{x}}) = b + \Delta b$ be a perturbed system, where $\widetilde{\boldsymbol{x}} = \boldsymbol{x} + \Delta\boldsymbol{x}$. Let $p > 1$ be a real number. If $\|A^{-1}\|_p \, \|\Delta A\|_p < 1$, then*

$$\frac{\|\Delta\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p} \leq \frac{c_p(A)}{1 - c_p(A)\frac{\|\Delta A\|_p}{\|A\|_p}} \left(\frac{\|\Delta A\|_p}{\|A\|_p} + \frac{\|\Delta b\|_p}{\|b\|_p} \right),$$

where $c_p(A) := \|A\|_p \|A^{-1}\|_p$.

2 A fitting problem

With the notation as in Subsection 1.1, let $F(\boldsymbol{x}, \boldsymbol{a}) \in \mathbb{R}[\boldsymbol{x}, \boldsymbol{a}]$ be a polynomial in the two series of variables $\boldsymbol{x} = (x_1, \ldots, x_n)$ (coordinates in the image space $\mathbb{A}^n_{\boldsymbol{x}}$), $\boldsymbol{a} = (a_1, \ldots, a_t)$ (coordinates in the parameter space $\mathbb{A}^t_{\boldsymbol{a}}$). From now on, we need the polynomial $F(\boldsymbol{x}, \boldsymbol{a})$ to be linear w.r.t. the set of variables $\boldsymbol{a}$. Furthermore, we assume that we are given a data set $\mathcal{D}$ of N points in the image space, with $N \gg t$, such that for every $\boldsymbol{x}_0 \in \mathcal{D}$ it holds that $F(\boldsymbol{x}_0, \boldsymbol{a})$ is not constant.

Question 2.1 *The problem essentially consists in determining a point $\boldsymbol{a}^\star \in \mathbb{A}^t_{\boldsymbol{a}}$ such that the hypersurface defined by $F(\boldsymbol{x}, \boldsymbol{a}^\star)$ from the family $\mathcal{F}$ as in (2) provides the best fit of $\mathcal{D}$.*

In general, in practice, one cannot expect to receive the data set $\mathcal{D}$ exactly. More precisely, we will consider two different types of *noises* in the data set:

(a) *perturbation noise*, consisting in random perturbations of the points' location of the elements in $\mathcal{D}$. That is, perturb each coordinate of each point (x, y) in this database by means of a Gaussian distribution $\mathcal{N}(0, \sigma^2)$ with zero mean and standard deviation σ (see Figure 1, left panel).

(b) *background noise*, consisting in points on the image space that appear in $\mathcal{D}$, according to a uniform distribution, and, or, points in $\mathcal{D}$ that do not follow, in the sense of not being closed to, the shape sought (see Figure 1, right panel).

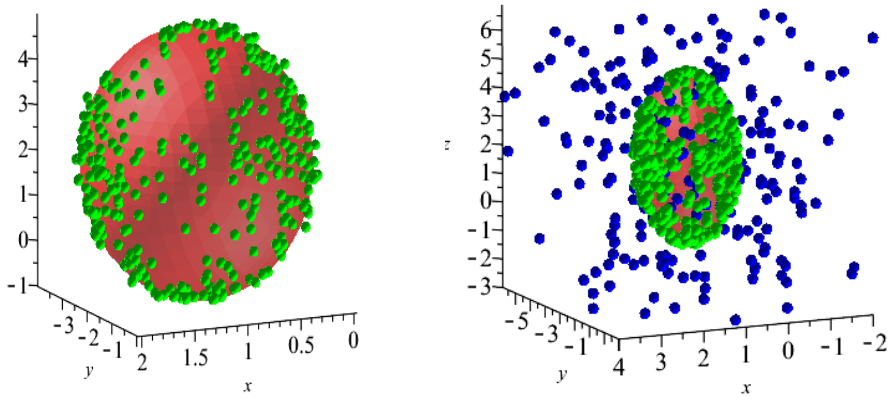


Figure 1: *Ellipsoid* $\mathcal{E}$ *of equation* $f(x,y,z) = x^2 - 2x + \frac{13}{9} + \frac{1}{4}y^2 + y + \frac{1}{9}z^2 - \frac{4}{9}z = 0$ *with data set consisting of small perturbation noise (green points, left panel), and with background noise (blue points, right panel).*

We will distinguish among the exact case (none noise, neither perturbed nor background) and the the non-exact case. Within the non-exact case, we still may distinguish between two different situations depending on the existence, or not, of background noise in the data set. We will say that $\mathcal{D}$ *has background noise* if there exists a significant (w.r.t. some threshold depending on the context) subset of points in $\mathcal{D}$ of type (b) as above, and that $\mathcal{D}$ *has no background noise* otherwise (see also Section 4). Both the exact case and the non-exact case without background noise can be treated via the Moore–Penrose approach.

2.1 The exact case

In this context, we assume the ideal situation where there exists a particular tuple $\boldsymbol{a}_0 \in \mathbb{A}_{\boldsymbol{a}}^t$ such that $F(\boldsymbol{x}, \boldsymbol{a}_0)$ is not constant and such that for all $\boldsymbol{x}_0 \in \mathcal{D}$ it holds that $F(\boldsymbol{x}_0, \boldsymbol{a}_0) = 0$; that is, $\mathcal{D}$ *is included in the variety defined by* $F(\boldsymbol{x}, \boldsymbol{a}_0)$. Let $\Gamma_{\boldsymbol{x}_0}$ denote the Hough transform of $\boldsymbol{x}_0 \in \mathbb{R}^n$ (w.r.t. $V(F)$), i.e., $\Gamma_{\boldsymbol{x}_0}$ is a hyperplane in the affine parameter space $\mathbb{A}_{\boldsymbol{a}}^t$ defined by $F(\boldsymbol{x}_0, \boldsymbol{a})$ (see Subsection 1.1). In this situation, the following regularity property holds true (see [4], [3]):

$$\bigcap_{\boldsymbol{x}_0 \in \mathcal{D}} \Gamma_{\boldsymbol{x}_0} = \{\boldsymbol{a}_0\}. \tag{8}$$

Therefore, in this ideal case, to compute the tuple $\boldsymbol{a}_0$ providing the hypersurface in the image space, described by $\mathcal{D}$, one simply needs to compute the intersection of the hyperplanes in (8), that is, to solve the linear system

$$\{F(\boldsymbol{x}_0, \boldsymbol{a}) = 0\}_{\boldsymbol{x}_0 \in \mathcal{D}}. \tag{9}$$

Let us illustrate this claim by the following toy example.

Example 2.2 We consider the ellipsoid $\mathcal{E}$ given by

$$f(x, y, z) = x^2 - 2x + \frac{13}{9} + \frac{1}{4}y^2 + y + \frac{1}{9}z^2 - \frac{4}{9}z.$$

Using a rational parametrization of $\mathcal{E}$ we randomly generate a data set $\mathcal{D}$ of 100 exact points on $\mathcal{E}$ (see Figure 2). Now, we consider the polynomial

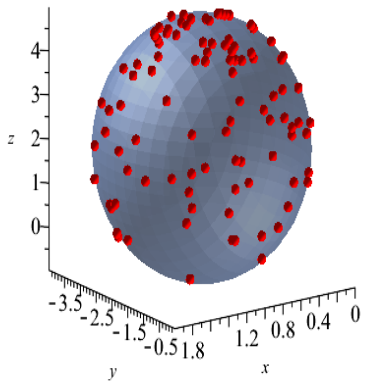


Figure 2: *Data set $\mathcal{D}$ (red points) and ellipsoid $\mathcal{E}$ in Example 2.2.*

.

$$F(\boldsymbol{x}, \boldsymbol{a}) = x^2 + a_1 y^2 + a_2 z_2^2 + a_3 xy + a_4 xz + a_5 yz + a_6 x + a_7 y + a_8 z + a_9 \tag{10}$$

representing all quadrics normalized in x^2. Of course, in general, one may lose information if the correct answer is not of the form (10). Nevertheless, this is not really a difficulty since one can either work projectively in $\boldsymbol{a}$ or use two different normalizations. Then $\{\Gamma_{\boldsymbol{x}_0}\}_{\boldsymbol{x}_0 \in \mathcal{D}}$ provides a linear system of 100 equations in the 9 unknowns $\{a_1, \ldots, a_9\}$. The system however has a unique solution, namely

$$\boldsymbol{a}_0 = \left(1, \frac{1}{4}, \frac{1}{9}, 0, 0, 0, -2, 1, -\frac{4}{9}, \frac{13}{9}\right),$$

and hence $f(x, y, z) = F(x, y, z, \boldsymbol{a}_0)$ is the solution. □

2.2 The non-exact case without background noise

Let $\mathcal{D}$ be a non-exact data set without background noise in the sense previously stated. Then, as in the exact case, the Hough transform $\{\Gamma_{\boldsymbol{x}_0}\}_{\boldsymbol{x}_0\in\mathcal{D}}$ provides the linear system (9) of N equations in t unknowns. Let us write (9) as:

$$A\boldsymbol{a}^T = A(a_1,\ldots,a_t)^T = b, \quad \text{where } A\in\mathcal{M}(N,t;\mathbb{R}),\ b\in\mathcal{M}(N,1;\mathbb{R}). \tag{11}$$

In general, the data set $\mathcal{D}$ is big enough to ensure that $N \gg t$ and that the points in $\mathcal{D}$ are "sufficiently general", that is, the coefficient matrix A associated to the system (11) is of full rank t. We then compute a "solution" via the Moore–Penrose pseudo-inverse, say

$$\boldsymbol{a}^\dagger = A^\dagger b. \tag{12}$$

In this situation, as pointed out in Section 1.2, the method returns $F\big(\boldsymbol{x},(a^\dagger)^T\big)$ as the best fitting variety for $\mathcal{D}$ within the family $F(\boldsymbol{x},\boldsymbol{a})$. Let us illustrate by two examples the case of data sets without background noise.

Example 2.3 We consider the same ellipsoid $\mathcal{E}$ as in Example 2.2 and we take a non-exact data set $\mathcal{D}$, without background noise, with 200 points. Indeed, it is the one shown in Figure 1 (left). Applying the Moore–Penrose approach w.r.t. to the

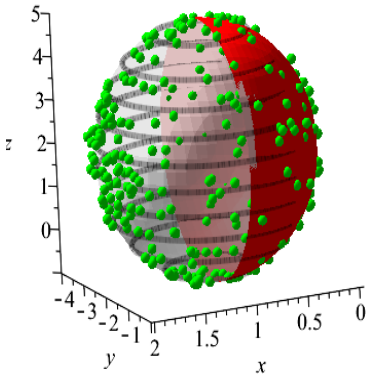


Figure 3: *Ellipsoid $\mathcal{E}$ as in Example 2.2 and ellipsoid $\mathcal{E}^\dagger$ of equation (13), respectively. Data set as $\mathcal{D}$ in Example 2.3 (green points).*

polynomial $F(\boldsymbol{x},\boldsymbol{a})$ as in (10) we get the ellipsoid $\mathcal{E}^\dagger$ defined by

$$x^2 - \frac{4305\,xy}{407698} - \frac{1539\,xz}{1088650} + \frac{11121\,y^2}{45770} - \frac{473\,yz}{1237796} + \frac{8261\,z^2}{75728} - \frac{65424\,x}{32437} + \frac{8033\,y}{8148} - \frac{10868\,z}{24879} + \frac{19951}{13780}.$$

The usual Euclidean distance in $\mathbb{R}^9$ between the coefficients of the defining polynomials of $\mathcal{E}$ and $\mathcal{E}^\dagger$ is ≈ 0.026 (see Figure 3). □

In the previous examples, the data set was generated artificially from a particular ellipsoid. In the next example, the data set is taken from the AIM@SHAPE repository [24].

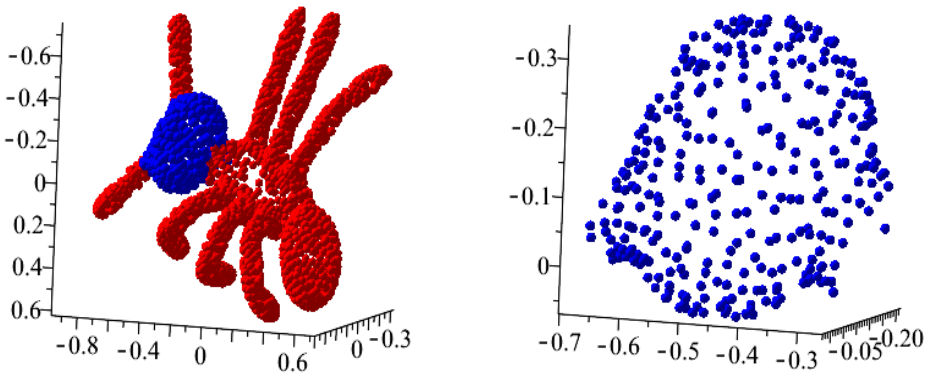


Figure 4: *Left panel: data set (blue and red points) of the ant. Right panel: data set of the ant head.*

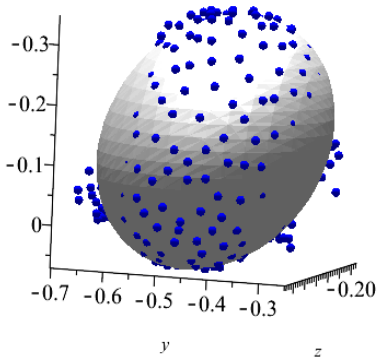


Figure 5: *Ellipsoid $\mathcal{E}^{\dagger}$ and data set $\mathcal{D}$ in Example 2.4 (blue points).*

Example 2.4 We consider the data set taken from the AIM@SHAPE repository, and representing an ant (see Figure 4). Our goal here is to detect the best ellipsoid fitting the head of the ant. Moreover, we consider $\mathcal{D}$ as the corresponding subset of points defining the head (see Figure 4). We note that $\#(\mathcal{D}) = 308$. Since there is no background noise we apply the Moore–Penrose method described above to get the ellipsoid

$\mathcal{E}^\dagger$ defined by the polynomial (see Figure 5):

$$x^2 + \frac{20327xy}{124120} - \frac{66581xz}{75703} + \frac{27881y^2}{33037} + \frac{2380yz}{28481} + \frac{66837z^2}{36527} + \frac{12472x}{43717} + \frac{28343y}{35136} + \frac{11666z}{65967} + \frac{12459}{71767}.$$

□

To treat the non-exact case in presence of background noise we need some more results. We discuss this in the next section.

3 Detecting a region in the parameter space

In this section, we use the same notation and assumptions as in Section 2. Here we address the following question.

> Assume that we are given a data set $\mathcal{D}$ of N points in the image space $\mathbb{A}^n_{\boldsymbol{x}}$, with $N \gg t$, such that for every $\boldsymbol{x}_0 \in \mathcal{D}$ it holds that $F(\boldsymbol{x}_0, \boldsymbol{a})$ is not constant. The goal is to find a region in $\mathbb{A}^t_{\boldsymbol{a}}$ corresponding (via the variety $V(F)$) to the cloud of points $\mathcal{D} \subset \mathbb{A}^n_{\boldsymbol{x}}$, that contains a good fitting parameter value $\boldsymbol{a}^\star_0$.

To this end we consider the following method based on the Moore–Penrose pseudo-inverse notion.

(1) Randomly choose ν subsets $\mathcal{D}_k$ of $\mathcal{D}$ of cardinality $M_k \geq t$, $k = 1, \ldots, \nu$.

(2) For each subset $\mathcal{D}_k$ consider the linear system of size $M_k \times t$ in the variables $\boldsymbol{a} = (a_1, \ldots, a_t)$ defined by

$$S_k := \{F(\boldsymbol{x}_0, \boldsymbol{a}) = 0\}_{\boldsymbol{x}_0 \in \mathcal{D}_k},\ k = 1, \ldots, \nu. \tag{13}$$

Since the points have been randomly chosen, we can assume the associated coefficients matrix A_k is of full rank t.

(3) Apply the Moore–Penrose approach described in Section 1.2 to each data set $\mathcal{D}_k$, to compute the corresponding Moore–Penrose solution z_k as in (5) of the linear system S_k, $k = 1, \ldots, \nu$. This way, we find ν points $z_1, \ldots, z_\nu$ in the parameter space $\mathbb{A}^t_{\boldsymbol{a}}$.

Let's consider the following "extreme" example.

Example 3.1 (Data set of points on a plane curve) Consider the polynomial $F(x, y; a_1, a_2)$ in the variables $\boldsymbol{x} = (x, y)$, $\boldsymbol{a} = (a_1, a_2)$, linear in the variables $\boldsymbol{a}$. Fix a curve $\mathcal{C}_{(a_1,a_2)} : f_{(a_1,a_2)}(x, y) = 0$ (for instance, a line of equation $x + a_1 y + a_2 = 0$), and consider a data set $\mathcal{D}$ of (distinct) points belonging to $\mathcal{C}_{(a_1,a_2)}$. Thus, the region in the parameter space corresponding to $\mathcal{D}$ reduces to the single point $(a_1, a_2) \in \mathbb{A}^2_{\boldsymbol{a}}$, since all the solutions z_k as in (13) coincide with that point. □

We note that property (6) certifies the goodness of the choice of the points z_k we found, implying that

- $F(\boldsymbol{x}, z_k)$ would provide the best fitting of $\mathcal{D}_k$, $k = 1, \ldots, \nu$.

Then we can conclude that, given a data set of points $\mathcal{D}$ in the image space, the Moore–Penrose approach provides the optimal solution z_k via the polynomial F w.r.t. each single subset $\mathcal{D}_k \subset \mathcal{D}$ in the parameter space. Thus, it is natural to ask how good the Moore–Penrose approach is w.r.t the whole data set of points $\mathcal{D}$. To this purpose, we suggest a measure to certify the goodness of the choice of the points as follows.

3.1 Convex hull goodness criterion

We require here the further assumption that $\{\mathcal{D}_1, \ldots, \mathcal{D}_\nu\}$ is a partition of $\mathcal{D}$, that is, $\mathcal{D} = \cup_{k=1}^{\nu} \mathcal{D}_k$ and $\mathcal{D}_i \cap \mathcal{D}_j = \emptyset$ for $i \neq j$.

- Letting $\mathcal{D}_k = \{p_1, \ldots, p_{M_k}\}$, we consider the linear system

$$F(p_1, \boldsymbol{a}) = \cdots = F(p_{M_k}, \boldsymbol{a}) = 0,$$

  that we rewrite in the form $A_k \boldsymbol{a}^T - b_k = \boldsymbol{0}$, where $A_k \in \mathcal{M}_{M_k \times t}(\mathbb{R})$, $k = 1, \ldots, \nu$.

- For $k = 1, \ldots, \nu$, we compute the Moore–Penrose solution $z_k^T = A_k^\dagger b_k$ of $A_k \boldsymbol{a}^T - b_k = \boldsymbol{0}$. Inequality (6) yields

$$\|A_k z_k^T - b_k\|_2 \leq \|A_k \boldsymbol{a}^T - b_k\|_2, \quad \text{for each } \boldsymbol{a} \in \mathbb{A}^t_{\boldsymbol{a}}.$$

- Consider the convex hull $\boldsymbol{\Delta} = \langle z_1, \ldots, z_\nu \rangle$ defined by the generated set of points $z_1, \ldots, z_\nu$. Each point $\boldsymbol{a}_0$ belonging to $\boldsymbol{\Delta}$ can be expressed as linear combination of the points $z_1, \ldots, z_\nu$, that is

$$\boldsymbol{a}_0 = \alpha_1 z_1 + \cdots + \alpha_\nu z_\nu$$

  for non-negative real numbers (depending on $\boldsymbol{a}_0$) $\alpha_1, \ldots, \alpha_\nu$ such that $\sum_{k=1}^{\nu} \alpha_k = 1$.

- Letting $\mathcal{D} = \{q_1, \dots, q_N\}$, we consider the linear system
$$F(q_1, \boldsymbol{a}) = \cdots = F(q_N, \boldsymbol{a}) = 0,$$
that we rewrite in the form $A\boldsymbol{a}^T - b = \mathbf{0}$, where $A \in \mathcal{M}_{N\times t}(\mathbb{R})$, $\boldsymbol{a}^T \in \mathcal{M}_{t\times 1}(\mathbb{R})$ and $b \in \mathcal{M}_{N\times 1}(\mathbb{R})$.

- Up to reordering the data set $\mathcal{D}$, we can assume that:
$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_\nu \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_\nu \end{pmatrix}$$

- For any given point $\boldsymbol{a}_0 \in \boldsymbol{\Delta}$ let us compute
$$\begin{aligned}
\|A\boldsymbol{a}_0^T - b\|_2 &\leq \sum_{k=1}^{\nu} \|A_k \boldsymbol{a}_0^T - b_k\|_2 \\
&= \sum_{k=1}^{\nu} \|A_k \ (\alpha_1 z_1^T + \cdots + \alpha_\nu z_\nu^T) - b_k\|_2 \\
&= \sum_{k=1}^{\nu} \left\| \sum_{j=1}^{\nu} \alpha_j A_k z_j^T - \sum_{j=1}^{\nu} \alpha_j b_k \right\|_2 \\
&= \sum_{k=1}^{\nu} \left\| \sum_{j=1}^{\nu} \alpha_j (A_k z_j^T - b_k) \right\|_2 \\
&\leq \sum_{k=1}^{\nu} \left(\sum_{j=1,\ j\neq k}^{\nu} \alpha_j \|A_k z_j^T - b_k\|_2 + \alpha_k \|A_k z_k^T - b_k\|_2 \right) \\
&\leq \sum_{k=1}^{\nu} \left(\sum_{j=1,\ j\neq k}^{\nu} \alpha_j \max_{j=1\dots\nu, j\neq k} \|A_k z_j^T - b_k\|_2 \right) + \sum_{k=1}^{\nu} \alpha_k \|A_k z_k^T - b_k\|_2 \\
&\leq \sum_{k=1}^{\nu} \max_{j=1,\dots,\nu,\ j\neq k} \|A_k z_j^T - b_k\|_2 + \sum_{k=1}^{\nu} \alpha_k \max_{k=1,\dots,\nu} \|A_k z_k^T - b_k\|_2 = \\
&= \sum_{k=1}^{\nu} \max_{j=1,\dots,\nu,\ j\neq k} \|A_k z_j^T - b_k\|_2 + \max_{k=1,\dots,\nu} \|A_k z_k^T - b_k\|_2 =: B \qquad (14)
\end{aligned}$$

- We then conclude that the evaluation $F(\boldsymbol{x}, \boldsymbol{a}_0)$ of the polynomial $F(\boldsymbol{x}, \boldsymbol{a})$ at each given point $\boldsymbol{a}_0 \in \Delta$ is bounded by
$$\left\| \begin{pmatrix} F(q_1, \boldsymbol{a}_0) \\ \vdots \\ F(q_N, \boldsymbol{a}_0) \end{pmatrix} \right\|_2 \leq B, \qquad (15)$$
where B is the quantity, only depending on $z_1, \dots, z_k$ and F, defined as in (14).

Definition 3.2 (Convex hull goodness criterion) In the situation as above, let $\varepsilon > 0$ be some threshold depending on the context. We say that the generated set of points $z_1, \ldots, z_\nu$ is optimal w.r.t. the whole data set of points $\mathcal{D}$, or that the region $\boldsymbol{\Delta}$ is *good*, if the inequality (15) holds true with $B < \varepsilon$.

We illustrate the convex hull criterion in the following example.

Example 3.3 We consider the data set $\mathcal{D}$ of Example 2.4 and we fix $\varepsilon = 0.75$. Following the steps outlined above, we compute the convex hull $\boldsymbol{\Delta}$ in the parameter space and we apply to it the convex hull goodness criterion. We work with $\nu \in \{5, 10, 15, 20, 25\}$ subsets $\mathcal{D}_k$ of $\mathcal{D}$ of cardinality $\#(\mathcal{D}_k) = M_k \geq \lfloor \frac{\#\mathcal{D}}{\nu} \rfloor$, where $\lfloor \cdot \rfloor$ denotes the floor function. In particular, we choose $M_k = \lfloor \frac{\#\mathcal{D}}{\nu} \rfloor$, for $k = 1, \ldots, \nu - 1$, and $M_\nu = \#(\mathcal{D}) - (\nu - 1)\lfloor \frac{\#\mathcal{D}}{\nu} \rfloor$ so that $\sum_{k=1}^{\nu} M_k = \#(\mathcal{D})$. For each data set, we repeat the computation of the convex hull $\boldsymbol{\Delta}$ and the quantity B defined in (14) for $N = 1000$ different choices of the partition $\{\mathcal{D}_1, \ldots, \mathcal{D}_\nu\}$ of $\mathcal{D}$. We gather the results in Table 1, where B_{av} denotes the arithmetic average of the quantity B computed for the N random partitions of the set $\mathcal{D}$. According to Definition 3.2, one sees that the regions $\boldsymbol{\Delta}$ are good for $\nu = 5, 10, 15$.

ν	5	10	15	20	25
B_{av}	0.2508	0.3508	0.5980	1.2400	1.8202

Table 1: *The arithmetic average B_{av} of the quantity B computed for $N = 1000$ random partitions of cardinality ν of the data set $\mathcal{D}$, in the cases $\nu = 5, 10, 15, 20, 25$.*

4 The non-exact case in presence of background noise

In this section, we use the same notation as in Section 2. So, we are given a polynomial $F(\boldsymbol{x}, \boldsymbol{a})$, in two different sets of variables $\boldsymbol{x}$ and $\boldsymbol{a}$, describing different shapes for each particular tuple $\boldsymbol{a} \in \mathbb{R}^t$, as well as a data set $\mathcal{D} \subset \mathbb{A}^n_{\boldsymbol{x}}$ that contains both perturbation noise and background noise. Let us express $\mathcal{D}$ as

$$\mathcal{D} = \mathcal{D}_P \cup \mathcal{D}_B, \tag{16}$$

where $\mathcal{D}_B$ contains the background noise points, and $\mathcal{D}_P$ the perturbation noise. Also, let us denote by $N_P = \#(\mathcal{D}_P)$ and $N_B = \#(\mathcal{D}_B)$ the cardinalities of $\mathcal{D}_P$ and $\mathcal{D}_B$, respectively. If

$$N_B/(N_P + N_B) \leq \varepsilon \tag{17}$$

for a certain tolerance ε depending on the context, we may consider the data set without background noise and hence the direct Moore–Penrose approach, described in Section 1, can be applied. In the sequel, we assume that the presence of background noise

is relevant, so, from now on, we need a different method, especially robust against background noise, based on the notion of the Hough transform of a point of the image space w.r.t. the variety $V(F)$.

Clearly, in presence of background noise, the Moore–Penrose approach is not suitable to solve Question 2.1. In this case, the method based on the Hough transform, a standard pattern recognition technique for detection of profiles in images, turns to be a quite accurate and robust method (we refer to [10], [3, Section 6], [15, Section 4], [21, Section 5], [27, Section 6] and [11] for full details and further references with special regard to a 2-dimensional image plane, and to [4] for a more general view).

4.1 Highlighting the Hough transform technique

To give an idea of the technique, let us consider the well established case of detection of curves in 2D images (e.g. see [3], [15]). Then, let $\mathcal{P}$ be a profile of interest in the image plane $\mathbb{A}^2_{\boldsymbol{x}}$, $\boldsymbol{x} = (x, y)$, pointed by a data set $\mathcal{D} \subset \mathbb{A}^2_{\boldsymbol{x}}$ of points p_j's. The Hough transform technique detects a curve from a family $\mathcal{F} = \{f_{\boldsymbol{a}}(x, y)\}$ best approximating the profile. A pre-processing step of the method consists of a standard edge detection technique on the image (see [8] for full details). This step reduces the number of points of which one has to compute the Hough transform (see Subsection 1.1). Because of the presence of noise and approximations (due to the floating point numbers representation encoding the real coordinates) of the points p_j's extracted from a digital image, and consequently on their Hough transforms $\Gamma_{p_j}(\mathcal{F})$, in practice in most cases it happens that $\cap_j \Gamma_{p_j}(\mathcal{F}) = \emptyset$; though we notice that there are regions in the parameter space with high density of Hough transform crossings. Then a discretization of a bounded region of the parameter space $\mathbb{A}^t_{\boldsymbol{a}}$ is required, which possibly exploit bounds on the parameters values $\boldsymbol{a}$. To find such bounds detecting the region of the parameter space to be considered is indeed one on the more critical issues of the method; in Subsection 4.2 we propose how to choose a suitable region. A last step is usually performed using the so called "voting procedure" which allows us to constructs the accumulator function in the discretized parameter space, such that the value of the accumulator in a cell of such a space corresponds to the number of times the Hough transforms of the points in the data set reach that cell. As a final outcome of the algorithm, the parameter values $\boldsymbol{a}^\star$ characterizing the hypersurface from the family $\mathcal{F} := \{\Gamma_{\boldsymbol{\lambda}}\}_{\boldsymbol{\lambda} \in \mathbb{A}^t_{\boldsymbol{a}}}$ (see equation (2)) to be detected in the image space is given by the parameter values identifying the cell where the accumulator function reaches its maximum. A further critical issue is then to find a suitable sampling distance for the discretization; an upper bound for it is discussed in Subsection 4.3.

4.2 Certifying points in the parameter space

Notation and assumptions as in Section 3. We aim to find a t-dimensional hyperrectangle in $\mathbb{A}^t_{\boldsymbol{a}}$ corresponding (via the variety $V(F)$) to the given data set of points

$\mathcal{D} = \{p_1, \ldots, p_N\} \subset \mathbb{A}^n_{\boldsymbol{x}}$. Note that the minimal number of points in general position in $\mathbb{A}^t_{\boldsymbol{a}}$ whose convex hull is a t-dimensional hyperrectangle is $t+1$, while the number of its vertices is 2^t. We specialize our generating points procedure as in Section 3 to the case $t+1 \leq \nu \leq 2^t$, as follows.

(1) Randomly choose ν subsets $\mathcal{D}_k$ of $\mathcal{D}$ of cardinality $M_k \geq t$, $k = 1, \ldots, \nu$, with $t+1 \leq \nu \leq 2^t$. Then apply the Moore–Penrose approach described in Section 1.2 to each data set $\mathcal{D}_k$ to compute the corresponding Moore–Penrose solution z_k of the linear system S_k as in (5), $k = 1, \ldots, \nu$. In this way, we find ν points $z_1, \ldots, z_\nu$ in the parameter space $\mathbb{A}^t_{\boldsymbol{a}}$.

(2) Consider the *axis-aligned minimum bounding box* $\boldsymbol{H}$ of the points $z_1, \ldots, z_\nu$, that is, the minimal t-hyperrectangle containing $z_1, \ldots, z_\nu$ subject to the constraint that its edges are parallel to the (Cartesian) coordinate axes. Let $\boldsymbol{h} = (h_1, \ldots, h_t)$ be the vector of the edge distances of $\boldsymbol{H}$ along the $a_1, \ldots, a_t$ axes, and let $\boldsymbol{a}_0$ be the center of $\boldsymbol{H}$. Set

$$\boldsymbol{h}_{\min} := \min\{h_1, \ldots, h_t\} \quad \text{and} \quad \boldsymbol{h}_{\max} := \max\{h_1, \ldots, h_t\}.$$

(3) We proceed by certifying the quality of a point by means of a voting process. Several voting procedures are nowadyas present in the literature. A main reference is a Matlab-based software freely available at the following URL: http://mida.dima.unige.it/g_software_htbone.html, described in [7].

Our voting process is different and it is based on the results in [25], [28, Section 5] (see also [27, Section 6]. Let's us briefly recall our procedure in the present context.

Consider the following steps.

- Let $\varepsilon > 0$ be fixed, let $\boldsymbol{a}_0 \in \mathbb{A}^t_{\boldsymbol{a}}$ and let $\Gamma_{\boldsymbol{x}_0}$ be the hyperplane defined by $f(\boldsymbol{a}) := F(\boldsymbol{x}_0, \boldsymbol{a})$ for a given $\boldsymbol{x}_0 \in \mathcal{D}$. Then:
  - (a) Theorem 3.2 and Remark 3.3 in [28] provide a bound, B_1, that depends on ε, $\boldsymbol{a}_0$ and $f(\boldsymbol{a})$, such that if $|f(\boldsymbol{a}_0)| > B_1$ then $\Gamma_{\boldsymbol{x}_0}$ does not pass through the ∞-ball centered at $\boldsymbol{a}_0$ and of radius ε.
  - (b) Theorem 4.6 and Remark 4.7 in [28] provide a bound, B_2, that depends on ε, $\boldsymbol{a}_0$ and $f(\boldsymbol{a})$, such that if $|f(\boldsymbol{a}_0)| < B_2$ then $\Gamma_{\boldsymbol{x}_0}$ does pass through the ∞-ball centered at $\boldsymbol{a}_0$ and of radius ε.
- Set $f_i(\boldsymbol{a}) := F(p_i, \boldsymbol{a})$, and let the hyperplanes $\Gamma_1, \ldots, \Gamma_N$ of $\mathbb{A}^t_{\boldsymbol{a}}$ be the voters, where $\Gamma_i := \{f_i(\boldsymbol{a}) = 0\}$, $i = 1, \ldots, N$.
- For $i = 1, \ldots, N$, apply item (a) with $\varepsilon := \dfrac{\boldsymbol{h}_{\max}}{2}$ and item (b) with $\varepsilon := \dfrac{\boldsymbol{h}_{\min}}{2}$. Precisely,
  - (a) If $|f_i(\boldsymbol{a}_0)| > B_1$, then Γ_i votes *no*, and its vote value is $v_i = 0$.

(b) If $|f_i(\boldsymbol{a}_0)| < B_2$, then Γ_i votes *yes*, and its vote value is

$$v_i = \frac{|B_2 - |f_i(\boldsymbol{a}_0)||}{B_2}.$$

Note that the vote increases when the difference between B_2 and $|f_i(\boldsymbol{a}_0)|$ increases, being the maximum value of the vote 1, which is achieved when $f_i(\boldsymbol{a}_0) = 0$, i.e., when $\boldsymbol{a}_0 \in \Gamma_i$.

(c) If none of the above conditions are satisfied, the vote is $v_i = 0$.

Let us denote by $\mathrm{Vote}(\boldsymbol{a}_0)$ the final scrutiny of votes:

$$\mathrm{Vote}(\boldsymbol{a}_0) := \sum_{i=1}^{N} v_i. \tag{18}$$

In other words, $\mathrm{Vote}(\boldsymbol{a}_0)$ is a weighted average of the number of hyperplanes crossing the ∞-ball centered at $\boldsymbol{a}_0$ and of radius $\frac{\boldsymbol{h}_{\min}}{2}$.

Following [15, Subsection 5.1], where a different voting procedure was used (see also [27, Section 6]), we measure the goodness of the point $\boldsymbol{a}_0$ as follows. Fix a percentage, say $\boldsymbol{p}\%$.

Definition 4.1 (Points goodness criterion) In the situation as above, we say that the point $\boldsymbol{a}_0$ is *good* if $\mathrm{Vote}(\boldsymbol{a}_0) > \dfrac{\boldsymbol{p}N}{100}$.

In the following we illustrate the method by some further examples.

Example 4.2 In Example 2.4 we have worked with a data set $\mathcal{D}_P$ of 308 points with only perturbation noise representing the head of an ant. Now, we take this set $\mathcal{D}_P$, and we introduce a new data set $\mathcal{D}_B$ with background noise. Then we apply our method to the new data set $\mathcal{D} := \mathcal{D}_P \cup \mathcal{D}_B$ to determine a region of interest $\boldsymbol{H}$ in the parameter space. We choose ν subsets D_k of $\mathcal{D}$ of cardinality M_k, with $\nu = \lceil 0.4 \times 2^9 \rceil = 205$ and we fix $M_k = 15$, $k = 1, \ldots, \nu$. In order to verify whether the output hyperrectangle $\boldsymbol{H}$ is valid, we will check whether the coefficients tuple $\mathcal{E}^\dagger$ of the solution ellipsoid as in Example 2.4 of equation

$$x^2 + \frac{20327xy}{124120} - \frac{66581xz}{75703} + \frac{27881y^2}{33037} + \frac{2380yz}{28481} + \frac{66837z^2}{36527} + \frac{12472x}{43717} + \frac{28343y}{35136} + \frac{11666z}{65967} + \frac{12459}{71767}$$

belongs to $\boldsymbol{H}$. Furthermore, we repeat the experiment for different cardinalities of $\mathcal{D}_B$ satisfying condition (17) with $\varepsilon = 0.8$, in the first three cases of Table 2, and with $\varepsilon = 0.9$ in the fourth case of the table (see also Fig. 6). The region $\boldsymbol{H}$ generated for each of the cases treated in these examples are shown below. For this purpose, instead of giving the 2^9 vertices of T we provide the center $\boldsymbol{a}_0$ of $\boldsymbol{H}$ and the vector $\boldsymbol{h}$ of its edge distances, as follows.

$N_P + N_B$	$N_P = \#(\mathcal{D}_P)$	perturbation noise	$N_B = \#(\mathcal{D}_B)$	background noise
514	308	60%	206	40%
770	308	40%	462	60%
1540	308	20%	1232	80%
3080	308	10%	2772	90%

Table 2: *Percentages of perturbation and background noises treated in Example* 4.2.

- Background noise = 40% (Figure 6, panel (a)):
  $\boldsymbol{a}_0 = (-0.440, -1.21, -0.292, -0.567, -1.01, -0.0250, -0.250, -0.871, -0.0739)$
  $\boldsymbol{h} = (4.55, 8.71, 6.08, 6.47, 8.78, 3.55, 5.17, 5.46, 1.60)$
  It holds that $\mathcal{E}^\dagger \in \boldsymbol{H}$ and $\|\mathcal{E}^\dagger - \boldsymbol{a}_0\|_2 = 3.04$

- Background noise = 60% (Figure 6, panel (b)):
  $\boldsymbol{a}_0 = (-0.854, 0.283, -0.194, 1.18, -0.0333, 0.0667, -0.743, 0.260, -0.121)$
  $\boldsymbol{h} = (4.96, 6.43, 6.61, 7.51, 6.07, 3.47, 5.09, 4.34, 1.21)$
  It holds that $\mathcal{E}^\dagger \in \boldsymbol{H}$ and $\|\mathcal{E}^\dagger - \boldsymbol{a}_0\|_2 = 2.06$

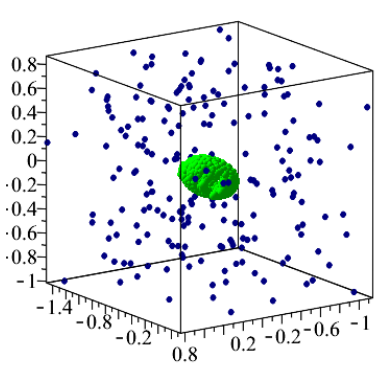
- Background noise = 80% (Figure 6, panel (c)):
  $\boldsymbol{a}_0 = (-0.103, 0.659, 0.333, -0.896, 0.0757, 0.384, -0.149, -0.0649, 0.0444)$
  $\boldsymbol{h} = (2.24, 3.54, 3.62, 5.54, 2.90, 2.48, 2.45, 2.42, 1.02)$
  It holds that $\mathcal{E}^\dagger \in \boldsymbol{H}$ and $\|\mathcal{E}^\dagger - \boldsymbol{a}_0\|_2 = 1.17$

- Background noise = 90% (Figure 6, panel (d)):
  $\boldsymbol{a}_0 = (-0.139, 0.164, -0.134, -0.0476, -0.0521, 0.286, -0.125, 0.0417, -0.135)$
  $\boldsymbol{h} = (1.94, 2.53, 3.23, 4.76, 2.23, 2.29, 2.25, 1.92, .930)$
  It holds that $\mathcal{E}^\dagger \notin \boldsymbol{H}$, but the distance $\|\mathcal{E}^\dagger - \boldsymbol{a}_0\|_2 = 1.67$ is not too big with respect to the context. □

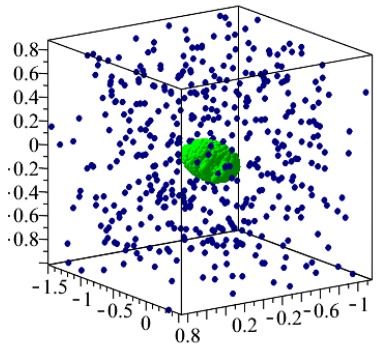
Remark 4.3 (Goodness of fit) Related to the general Question 2.1, a crucial task in detection of curves (hypersurfaces) in images in presence of noise is, of course, to certify the goodness of fit in the image space $\mathbb{A}^n_{\boldsymbol{x}}$. Let's add a comment for the general case $n \geq 3$.

Following a nowadays standard procedure on which the Hough transform technique is founded, one first constructs the bounding box $\boldsymbol{H}$ in the parameter space $\mathbb{A}^t_{\boldsymbol{a}}$, for instance as described in Subsection 4.2. Next, one considers a discretization $\boldsymbol{\delta}$ (a regular grid of points samples) of the region of interest $\boldsymbol{H}$, where $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_t)$ is the vector of the sampling distances along the $a_1, \ldots, a_t$ axes. Let $\boldsymbol{C}(\boldsymbol{\lambda})$ be a cell of the discretization δ, i.e., $\boldsymbol{C}(\boldsymbol{\lambda})$ is the ∞-ball of center a point $\boldsymbol{\lambda} \in \boldsymbol{H}$ and of radii $\frac{\delta_1}{2}, \ldots, \frac{\delta_t}{2}$ along the $a_1, \ldots, a_t$ axes. Then the region $\boldsymbol{H}$ is partitioned into a finite set of cells $\mathfrak{C} = \{\boldsymbol{C}(\boldsymbol{\lambda})\}$. Now, with the notation as in definition (18), set

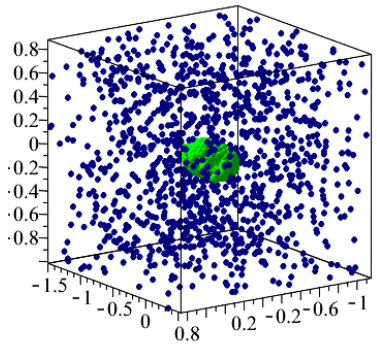
$$\mu(\boldsymbol{C}(\boldsymbol{\lambda})) := \#\{i \in \{1, \ldots, N\} \mid \Gamma_i \cap \boldsymbol{C}(\boldsymbol{\lambda}) \neq \emptyset\}.$$



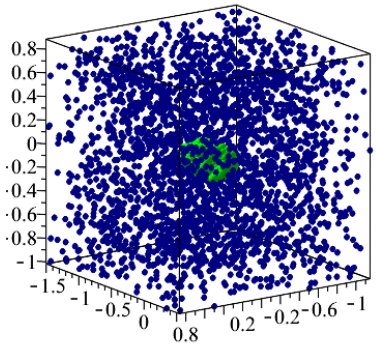
(a) Background noise = 40%.



(b) Background noise = 60%.



(c) Background noise = 80%.



(d) Background noise = 90%.

Figure 6: *Data set $\mathcal{D} = \mathcal{D}_P \cup \mathcal{D}_B$ with different percentages of background noise as in Example 4.2, Table 2: $\mathcal{D}_P$, perturbation noise (green points); $\mathcal{D}_B$, background noise, (blue points).*

We then choose $\boldsymbol{C}(\boldsymbol{a}^\star) \in \mathfrak{C}$ such that $\mu(\boldsymbol{C}(\boldsymbol{a}^\star))$ is maximum, and we output as optimal fitting of the data set $\mathcal{D}$ the hypersurface $\Gamma_{\boldsymbol{a}^\star}$ from the family $\mathcal{F}$ defined by $F(\boldsymbol{x}, \boldsymbol{a}^\star)$ as in (2).

In the case $n = 2$, a measure of the goodness of fit, based on a notion of "good point" analogous to Definition 4.1, was proposed in [15, Subsection 5.1] and [27, Section 6]. In particular, the goodness of fit criterion from [15] verbatim extends to $n \geq 3$ as follows.

Consider the subregion $\boldsymbol{H}' \subset \boldsymbol{H}$ consisting of the union of the cells $\boldsymbol{C}(\boldsymbol{\lambda})$, of the given discretization $\boldsymbol{\lambda}$ of $\boldsymbol{H}$, whose centers $\boldsymbol{\lambda}$ are good in the sense of Definition 4.1 (w.r.t. a fixed percentage $\boldsymbol{p}$). First, extract from the original data set $\mathcal{D}$ the points

whose Hough transforms cross $\boldsymbol{H}'$, that is, set

$$\mathcal{D}' := \{p_{i'} \in \mathcal{D}, i' \in \{1, \dots, N'\} \mid \Gamma_p \cap \boldsymbol{H}' \neq \emptyset\}.$$

Next, for each index i', compute the minimum Euclidean distance $d_{i'}$ between the point $p_{i'} \in \mathcal{D}'$ and the recognized hypersurface $\Gamma_{\boldsymbol{a}^\star}$

$$d_{i'} := \min_{\boldsymbol{x} \in \Gamma_{\boldsymbol{a}^\star}} \{\|p_{i'} - \boldsymbol{x}\|_2\}.$$

Finally, define the GoF measure

$$\text{GoF} := \frac{1}{N'} \sum_{i'=1}^{N'} d_{i'}.$$

For a fixed familly $\mathcal{F} = \{\Gamma_{\boldsymbol{a}}\}$ of hypersurfaces as in (2) and for the detected hypersurface $\Gamma_{\boldsymbol{a}^\star}$ from the family, the GoF measure provides a quantitative evaluation of the detection error (see [15, Subsection 5.1] for more comments and illustrative examples in the case $n = 2$). □

4.3 An upper bound for the sampling distance of the discretization

We provide here an upper bound for the discretization step of the region $\boldsymbol{H}$ of the parameter space $\mathbb{A}^t_{\boldsymbol{a}}(\mathbb{R})$ constructed as in Subsection 4.2 (we also refer to [27, Section 5] for a different approach based on the results of [29]). We proceed as follows.

(1) Since the polynomial $F(\boldsymbol{x}, \boldsymbol{a}) \in \mathbb{R}[\boldsymbol{x}, \boldsymbol{a}]$ is linear in the parameters $\boldsymbol{a} = (a_1, \dots, a_t)$, it can be written in the form

$$F(\boldsymbol{x}, a_1, \dots, a_t) = f_1(\boldsymbol{x})a_1 + \cdots + f_t(\boldsymbol{x})a_t - f_{t+1}(\boldsymbol{x}),$$

for some $f_r(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}]$, $r = 1, \dots, t+1$.

(2) Let $\mathbb{X} = \{p_1, \dots, p_t\}$ be a set of t points from the data set $\mathcal{D}$ in the image space $\mathbb{A}^n_{\boldsymbol{x}}(\mathbb{R})$ as in (16) of Section 4 such that the linear system

$$\begin{cases} f_1(p_1)a_1 + \cdots + f_t(p_1)a_t = f_{t+1}(p_1) \\ \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ f_1(p_t)a_1 + \cdots + f_t(p_t)a_t = f_{t+1}(p_t) \end{cases} \tag{19}$$

is of full rank t.

(3) Write the system (19) in the form $A(a_1,\dots,a_t)^T = b$, where

$$A = \begin{pmatrix} f_1(p_1) & \dots & f_t(p_1) \\ \vdots & & \vdots \\ f_1(p_t) & \dots & f_t(p_t) \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} f_{t+1}(p_1) \\ \vdots \\ f_{t+1}(p_t) \end{pmatrix},$$

so that the solution is

$$\overline{\boldsymbol{a}} = A^{-1}b.$$

(4) Consider perturbations $\Delta p_1,\dots,\Delta p_t$ of the points of $\mathbb{X}$ such that, for some positive real value $\epsilon < 1$,

$$\|\Delta p_j\|_\infty \le \epsilon,\ j = 1,\dots,t. \tag{20}$$

(5) We search for the solution of the perturbed linear system in the variables $a_r + \Delta a_r$, $r = 1,\dots,t$,

$$\begin{cases} f_1(p_1+\Delta p_1)(a_1+\Delta a_1) + \dots + f_t(p_1+\Delta p_1)(a_t+\Delta a_t) = f_{t+1}(p_1+\Delta p_1) \\ \quad\vdots \qquad\qquad\qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots \\ f_1(p_t+\Delta p_t)(a_1+\Delta a_1) + \dots + f_t(p_t+\Delta p_t)(a_t+\Delta a_t) = f_{t+1}(p_t+\Delta p_t). \end{cases} \tag{21}$$

For $r = 1,\dots,t+1$, $j = 1,\dots,t$, consider the formal Taylor expansion of the polynomial $f_r(\boldsymbol{x})$ at $p_j + \Delta p_j$,

$$f_r(p_j+\Delta p_j) = f_r(p_j) + \mathrm{grad}(f_r)(p_j)\Delta p_j + \mathrm{O}(\epsilon^2),$$

that is, $f_r(p_j+\Delta p_j) = f_r(p_j) + \mathrm{grad}(f_r)(p_j)\Delta p_j$ neglecting contributions of order $\mathrm{O}(\epsilon^2)$. Thus, by a first-order error analysis, the system (21) can be written in the form

$$(A+\Delta A)\begin{pmatrix} a_1+\Delta a_1 \\ \vdots \\ a_t+\Delta a_t \end{pmatrix} = b + \Delta b,$$

where

$$\Delta A = \begin{pmatrix} \mathrm{grad}(f_1)(p_1)\Delta p_1 & \dots & \mathrm{grad}(f_t)(p_1)\Delta p_1 \\ \vdots & & \vdots \\ \mathrm{grad}(f_1)(p_t)\Delta p_t & \dots & \mathrm{grad}(f_t)(p_t)\Delta p_t \end{pmatrix}$$

and

$$\Delta b = \begin{pmatrix} \mathrm{grad}(f_{t+1})(p_1)\Delta p_1 \\ \vdots \\ \mathrm{grad}(f_{t+1})(p_t)\Delta p_t \end{pmatrix}.$$

(6) By using standard relations between norms and the sub-multiplicative property (see [14, §2.3.1]), compute

$$\begin{aligned}\|\Delta A\|_\infty &= \max_{j=1,\dots,t}\Big\{\sum_{r'=1}^{t}\big|\mathrm{grad}(f_{r'})(p_j)\Delta p_j\big|\Big\}\\ &\le \max_{j=1,\dots,t}\Big\{\sum_{r'=1}^{t}\|\mathrm{grad}(f_{r'})(p_j)\|_1\ \|\Delta p_j\|_\infty\Big\}.\end{aligned}$$

Whence, recalling bound (20),

$$\|\Delta A\|_\infty \le \epsilon\Big(\max_{j=1,\dots,t}\Big\{\sum_{r'=1}^{t}\|\mathrm{grad}(f_{r'})(p_j)\|_1\Big\}\Big) =: \epsilon\ \mathcal{B}'. \tag{22}$$

Similarly, compute

$$\begin{aligned}\|\Delta b\|_\infty &= \max_{j=1,\dots,t}\Big\{\big|\mathrm{grad}(f_{t+1})(p_j)\Delta p_j\big|\Big\}\\ &\le \max_{j=1,\dots,t}\Big\{\|\mathrm{grad}(f_{t+1})(p_j)\|_1\ \|\Delta p_j\|_\infty\Big\}\\ &\le \epsilon\Big(\max_{j=1,\dots,t}\Big\{\|\mathrm{grad}(f_{t+1})(p_j)\|_1\Big\}\Big) =: \epsilon\ \mathcal{B}''.\end{aligned} \tag{23}$$

From now on, we need to assume

$$\epsilon < \min\left\{\frac{1}{\mathcal{B}'\|A^{-1}\|_\infty},\ 1\right\}. \tag{24}$$

Hence, in particular,

$$\|A^{-1}\|_\infty < \frac{1}{\epsilon\ \mathcal{B}'}.$$

Then, from bound (22), we get

$$\|A^{-1}\|_\infty\ \|\Delta A\|_\infty < 1.$$

Therefore, Theorem 1.2 applies to give, for the solution $\overline{\boldsymbol{a}}$ of the linear system (19),

$$\frac{\|\Delta\overline{\boldsymbol{a}}\|_\infty}{\|\overline{\boldsymbol{a}}\|_\infty} \le \frac{c_\infty(A)}{1-c_\infty(A)\frac{\|\Delta A\|_\infty}{\|A\|_\infty}}\left(\frac{\|\Delta A\|_\infty}{\|A\|_\infty}+\frac{\|\Delta b\|_\infty}{\|b\|_\infty}\right), \tag{25}$$

where $c_\infty(A) := \|A\|_\infty\|A^{-1}\|_\infty$. Thus, by combining (25) with bounds (22), (23), and recalling that $\epsilon < 1$, we find

$$\begin{aligned}\frac{\|\Delta\overline{\boldsymbol{a}}\|_\infty}{\|\overline{\boldsymbol{a}}\|_\infty} &\leq \frac{c_\infty(A)}{1-c_\infty(A)\dfrac{\epsilon\ \mathcal{B}'}{\|A\|_\infty}}\left(\frac{\epsilon\ \mathcal{B}'}{\|A\|_\infty}+\frac{\epsilon\ \mathcal{B}''}{\|b\|_\infty}\right)\\ &\leq \frac{c_\infty(A)}{1-c_\infty(A)\dfrac{\mathcal{B}'}{\|A\|_\infty}}\left(\frac{\epsilon\ \mathcal{B}'}{\|A\|_\infty}+\frac{\epsilon\ \mathcal{B}''}{\|b\|_\infty}\right)\\ &= \epsilon\left(\frac{c_\infty(A)}{1-c_\infty(A)\dfrac{\mathcal{B}'}{\|A\|_\infty}}\left(\frac{\mathcal{B}'}{\|A\|_\infty}+\frac{\mathcal{B}''}{\|b\|_\infty}\right)\right)=:\epsilon\ \mathcal{B}. \qquad (26)\end{aligned}$$

Finally, we then obtain the bound

$$\|\Delta\overline{\boldsymbol{a}}\|_\infty \leq \epsilon\big(\|\overline{\boldsymbol{a}}\|_\infty\ \mathcal{B}\big), \qquad (27)$$

with the quantity $\mathcal{B}$ defined as in (26).

To conclude, as final step, we randomly choose N sets $\mathbb{X}_\alpha = \{{p_1}^\alpha, \ldots, {p_t}^\alpha\}$, $\alpha = 1, \ldots, N$, of t points from the data set $\mathcal{D}$ in the image space $\mathbb{A}^n_{\boldsymbol{x}}(\mathbb{R})$, as done in step (2) of the procedure described above. Therefore, with clear meaning of the symbols, our argument provides N bounds

$$\|\Delta\overline{\boldsymbol{a}}_\alpha\|_\infty \leq \epsilon\big(\|\overline{\boldsymbol{a}}_\alpha\|_\infty\ \mathcal{B}_\alpha\big),\quad \alpha = 1, \ldots, N.$$

One then may consider as candidates for the discretization $\boldsymbol{\delta} = (\delta, \ldots, \delta)$ in the parameter space, with the same sampling distance δ along the $a_1, \ldots, a_t$ axes, one the following:

1. Up to renaming the indices α's, we can assume that

$$\|\overline{\boldsymbol{a}}_1\|_\infty\ \mathcal{B}_1 \leq \cdots \leq \|\overline{\boldsymbol{a}}_\alpha\|_\infty\ \mathcal{B}_\alpha \leq \cdots \leq \|\overline{\boldsymbol{a}}_N\|_\infty\ \mathcal{B}_N.$$

Then take the discretization $\boldsymbol{\delta}$ of sampling distance

$$\delta_{\text{median}} := \epsilon\big(\|\overline{\boldsymbol{a}}_s\|_\infty\ \mathcal{B}_s\big) \quad \text{if } N = 2s+1 \text{ is odd},$$

and

$$\delta_{\text{median}} := \epsilon\Big(\frac{\|\overline{\boldsymbol{a}}_s\|_\infty\ \mathcal{B}_s + \|\overline{\boldsymbol{a}}_{s+1}\|_\infty\ \mathcal{B}_{s+1}}{2}\Big) \quad \text{if } N = 2s \text{ is even}.$$

2. Consider the discretization $\boldsymbol{\delta}$ of sampling distance $\delta_{\text{mean}} := \epsilon\Big(\dfrac{\sum_{\alpha=1}^N \|\overline{\boldsymbol{a}}_\alpha\|_\infty\ \mathcal{B}_\alpha}{N}\Big)$.

Practical examples show that δ_{median} is a more suitable discretization step (clearly enough, other possible values as either $\delta_{\min} := \epsilon\big(\min_{\alpha=1,\ldots,N}\{\|\overline{\boldsymbol{a}}_\alpha\|_\infty\ \mathcal{B}_\alpha\}\big)$ or $\delta = \boldsymbol{\delta}_{\max} : \epsilon\big(\max_{\alpha=1,\ldots,N}\{\|\overline{\boldsymbol{a}}_\alpha\|_\infty\ \mathcal{B}_\alpha\}\big)$ does not look appropriate).

To conclude this section, we provide some illustrative examples.

Example 4.4 In the affine plane $\mathbb{A}^2_{(x,y)}(\mathbb{R})$ we consider the family $\mathcal{F} = \{\mathcal{C}_{\boldsymbol{a}}\}$, $\boldsymbol{a} = (a_1, a_2, a_3, a_4)$, of cubic curves of equation

$$\mathcal{C}_{\boldsymbol{a}} : x^2 = -a_3 y^3 + a_4 y^2 - a_1 y + a_2,$$

for $t = 4$ real parameters a_1, a_2, a_3, a_4, with a_3 positive. The general curve of the family is non-singular, so that it is an elliptic curve.

In the image plane $\mathbb{A}^2_{(x,y)}(\mathbb{R})$ we consider a set $\mathcal{D}$ of 636 points represented in Figure 7. The data set $\mathcal{D}$ has been taken from [25, Example 6.4] and from [15] (see Figure 3, upper panels, in [15]), after an edge detection processing and extracting the set of points lying in the box $[-3, 3] \times [-3, 3]$). We note that the points of the set

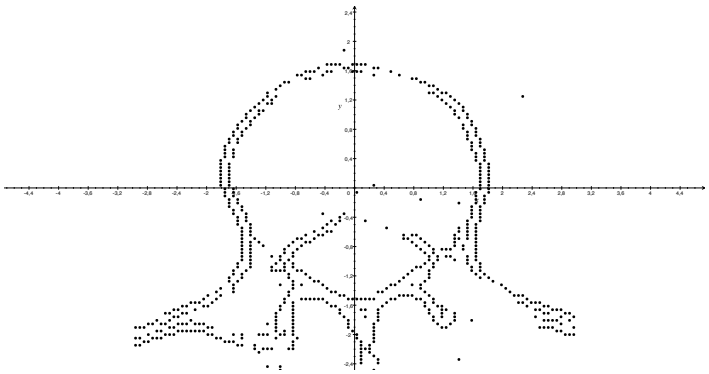


Figure 7: *Data set $\mathcal{D}$ in Example 4.4.*

$\mathcal{D}$ are given with four exact decimals, so that the component-wise difference between two points of $\mathcal{D}$ has to be (if not zero) greater than 10^{-4}. This suggests that the quantity $\|\Delta p_j\|_\infty$ as in formula (20) has to be of the order $\|\Delta p_j\|_\infty \approx \frac{10^{-4}}{2} = 5 \times 10^{-5}$, $j = 1, \ldots, t$.

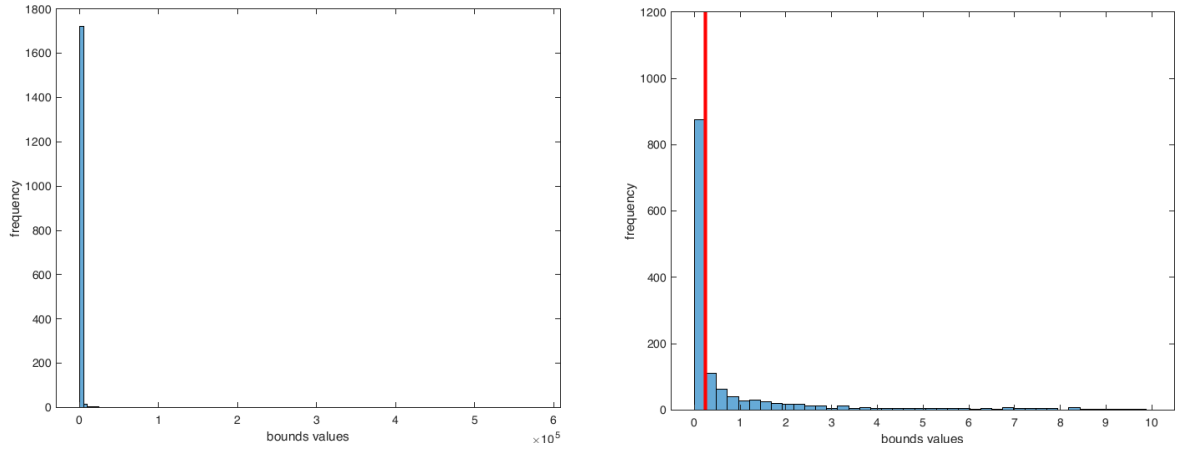
We use the procedure described in Subsection 4.3, where the polynomial $F(\boldsymbol{x}, \boldsymbol{a}) \in \mathbb{R}[\boldsymbol{x}, \boldsymbol{a}]$, $\boldsymbol{x} = (x, y)$, is now

$$F(x, y, a_1, a_2, a_3, a_4) = x^2 + a_3 y^3 - a_4 y^2 + a_1 y - a_2.$$

We then randomly choose $N = 2000$ subsets $\mathbb{X}_\alpha = \{{p_1}^\alpha, {p_2}^\alpha, {p_3}^\alpha, {p_4}^\alpha\}$, $\alpha = 1, \ldots, N$, of $t = 4$ points from the data set $\mathcal{D}$ in the image space $\mathbb{A}^2_{(x,y)}(\mathbb{R})$ and we compute the corresponding bound $\epsilon\big(\|\overline{\boldsymbol{a}}_\alpha\|_\infty \ \mathcal{B}_\alpha\big)$, $\alpha = 1, \ldots, N$, distributed as in the histograms represented in Figure 8. We observe that more than 88% of the subsets satisfy the assumption (24).

The value of δ_{median} is shown in the Table 3, where for completeness and comparison, we include the candidates δ_{min}, δ_{mean}, δ_{max} for the sampling distance of the discretization in the parameter space.

In Figure 9 we consider a profile of interest $\mathcal{P}$ highlighted by the data set $\mathcal{D}$ as in Figure 7, and the curve (in red) from the family $\mathcal{F}$ best approximating $\mathcal{P}$ and detected



(a) (b)

Figure 8: (a) *Histogram of the distribution of the bounds* $\epsilon\big(\|\overline{\boldsymbol{a}}_\alpha\|_\infty\, \mathcal{B}_\alpha\big)$, $\alpha = 1, \ldots, N = 2000$, *computed in Example* 4.4. (b) *Plot of histogram* (a) *obtained by using the bounds values that fall in the interval* $[0, 10]$. *The red line shows where the discretization step candidate* δ_{median} *is located.*

ϵ value	$\delta_{\min}$	δ_{median}	δ_{mean}	$\delta_{\max}$
$\epsilon = 5 \times 10^{-5}$	0.0026	0.2337	247.2576	9.0291×10^4

Table 3: Values of $\delta_{\min}$, δ_{median}, δ_{mean}, $\delta_{\max}$ for the sampling distance of the discretization in the parameter space.

via the Hough transform technique by using the RECOGNITION algorithm as in [25], with discretization $\boldsymbol{\delta} = (0.1, 0.1, 0.1, 0.1)$ (see Section 6, Figure 6.4). We then compare such a curve with the elliptic curve from the same family (in green) obtained in the same way by using the discretization $\boldsymbol{\delta}$ of sampling distance $\delta_{\text{median}} = 0.2337$ as in Table 3.

Example 4.5 In the affine space $\mathbb{A}^3_{(x,y,z)}(\mathbb{R})$ we consider the family $\mathcal{F} = \{\mathcal{E}_{\boldsymbol{a}}\}$, $\boldsymbol{a} = (a_1, \ldots, a_9)$ of ellipsoids of equations

$$\mathcal{E}_{\boldsymbol{a}} : x^2 + a_1 y^2 + a_2 z_2^2 + a_3 xy + a_4 xz + a_5 yz + a_6 x + a_7 y + a_8 z + a_9 = 0$$

for $t = 9$ parameters $a_1, \ldots, a_9$, as in Example 2.2. In the image space $\mathbb{A}^3_{(x,y,z)}(\mathbb{R})$ we consider a set $\mathcal{D}$ of 514 points represented in Figure 6, high-left panel. We note that the points of the set $\mathcal{D}$ are given with ten exact decimals, so that the component-wise difference between two points of $\mathcal{D}$ has to be (if not zero) greater then 10^{-10}. This would suggest that the quantity $\|\Delta p_j\|_\infty$ as in formula (20) has to be of the order $\|\Delta p_j\|_\infty \approx \frac{10^{-10}}{2} = 5 \times 10^{-11}$, $j = 1, \ldots, t$.

Again, we use the procedure described in Subsection 4.3, where the polynomial

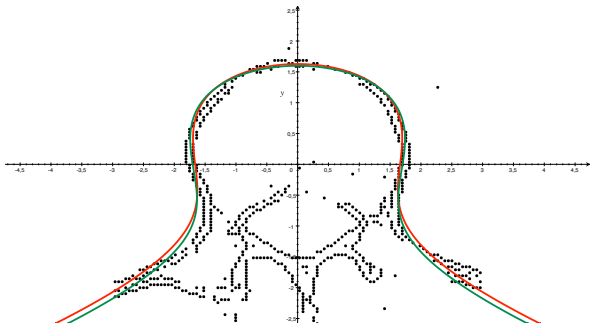


Figure 9: *Red curve of equation* $x^2 = -\frac{4}{5}y^3 + \frac{21}{50}y + \frac{69}{25}$*; green curve of equation* $x^2 = -\frac{4}{5}y^3 - \frac{1}{5}y^2 + \frac{2763}{5000}y + \frac{7237}{2500}$.

$F(\boldsymbol{x}, \boldsymbol{a}) \in \mathbb{R}[\boldsymbol{x}, \boldsymbol{a}]$, $\boldsymbol{x} = (x, y, z)$, is now

$$F(x, y, z, \boldsymbol{a}) = x^2 + a_1 y^2 + a_2 z_2^2 + a_3 xy + a_4 xz + a_5 yz + a_6 x + a_7 y + a_8 z + a_9.$$

We then randomly choose $N = 2000$ subsets $\mathbb{X}_\alpha = \{{p_1}^\alpha, {p_2}^\alpha, \ldots, {p_8}^\alpha, {p_9}^\alpha\}$, $\alpha = 1, \ldots, N$, of $t = 9$ points from the data set $\mathcal{D}$ in the image space $\mathbb{A}^3_{(x,y,z)}(\mathbb{R})$ and we compute the corresponding bound $\epsilon\big(\|\overline{\boldsymbol{a}}_\alpha\|_\infty \, \mathcal{B}_\alpha\big)$, $\alpha = 1, \ldots, N$. The value of δ_{median} is shown Table 4, where for completeness and comparison, we include the candidates δ_{min}, δ_{mean}, δ_{max} for the sampling distance of the discretization in the parameter space. For the threshold ϵ we consider bigger values than the "expected" order $\approx 10^{-11}$ mentioned above, since otherwise the value of δ_{median} looks probably too small and therefore producing a too big number of cells, with a consequent too high computational burden in the voting procedure of the Hough transform technique.

ϵ value	δ_{min}	δ_{median}	δ_{mean}	δ_{max}
10^{-10}	4.1761×10^{-8}	3.6263×10^{-6}	0.2500×10^{-3}	2.9658
10^{-8}	4.1973×10^{-6}	3.3565×10^{-4}	5719×10^{-1}	72725×10^{2}
10^{-6}	3.8398×10^{-4}	3.780×10^{-2}	3.0459×10	1.9627×10^{4}
10^{-4}	2.5700×10^{-2}	2.9408	1.2368×10^{2}	4.7188×10^{4}

Table 4: Values of δ_{min}, δ_{median}, δ_{mean}, δ_{max} for the sampling distance of the discretization in the parameter space.

Acknowledgements. J.R. Sendra and J. Sendra are partially supported by FEDER/Ministerio de Ciencia, Innovación y Universidades - Agencia Estatal de Investigación/MTM2017-88796-P (Symbolic Computation: new challenges in Algebra and Geometry together with its applications). Maria-Laura Torrente is a member of GNAMPA - Gruppo Nazionale per l´Analisi Matematica, la Probabilità e le loro Applicazioni of INDAM.

References

[1] J. Abbott, A.M. Bigatti and L. Robbiano, CoCoA: a system for doing Computations in Commutative Algebra. Available at `http://cocoa.dima.unige.it`

[2] M.C. Beltrametti, C. Campi, A.M. Massone and M. Torrente, Geometry of the Hough transform with applications to synthetic data, arXiv 1904.02587 (2019).

[3] M.C. Beltrametti, A.M. Massone and M. Piana, Hough transforms of special classes of curves, SIAM J. Imaging Sci., 6(1) (2013), 391–412.

[4] M.C. Beltrametti and L. Robbiano, An algebraic approach to Hough transforms, Journal of Algebra, 37 (2012), 669–681.

[5] A. Ben–Israel and T.N.E. Greville, Generalized Inverses: Theory and Applications, Second ed., Springer–Verlag, Berlin, 2003.

[6] S.L. Campbell and C.D. Meyer, Generalized Inverses of Linear Transformations. Series: Classics in Applied Mathematics, SIAM, 2009.

[7] C. Campi, A. Perasso, M. C. Beltrametti, G. Sambuceti, A.M. Massone and M. Piana, HT-BONE: A Graphical User Interface for the identification of bone profiles in CT images via extended Hough transform, Proc. SPIE 9784, Medical Imaging 2016: Image Processing, 978423.

[8] J. Canny, A Computational Approach to Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6) (1986), 679–698.

[9] C. Conti, L. Romani and D. Schenone, Semi-automathic spline fitting of planar curvlinear profiles in digital images using the Hough transform, Pattern Recognition, 74 (2018), 64–76.

[10] R.O. Duda and P.E. Hart, Use of the Hough transformation to detect lines and curves in pictures, *Comm. ACM*, **15**, vol. 1 (1972), 11–15.

[11] Hassanein A.S., Mohamed S., Sameer M. and Ragab M.E.: A survey on Hough transform, theory, techniques and applications, arXiv:1502.02160v1 [cs.CV] (2015).

[12] E. Hernandez and G. Weiss, *A First Course on Wavelets*, CRC-Press (1996).

[13] R.A. Horn and C.R. Johnson, *Matrix Analysis*, 2nd ed., Cambridge University Press, New York, 2013.

[14] G. Golub and C. F. Van Loan, *Matrix Computations*, 3td ed., Johns Hopkins, Baltimore, 1996.

[15] A.M. Massone, A. Perasso C. Campi and M.B. Beltrametti, Profile detection in medical and astronomical images by means of the Hough transform of special classes of curves, J. Math. Imaging Vis. 51(2) (2015), 296–310.

[16] E.H. Moore, On the reciprocal of the general algebraic matrix, Bulletin of the American Mathematical Society 26(3) (1920), 94–395.

[17] P. Mukhopadhyay and B.B. Chaudhuri, A survey of Hough Transform, Pattern Recognition 48(3) (2014), 993–1010.

[18] B.G. Osgood, *Lectures on the Fourier Transform and Its Applications*, Pure and Applied Undergraduate Texts, AMS (2019).

[19] R. Penrose, A generalized inverse for matrices, Proceedings of the Cambridge Philosophical Society, 51 (1955), 406–403.

[20] R. Penrose, On best approximate solution of linear matrix equations, Proceedings of the Cambridge Philosophical Society, 52 (1956), 17–19.

[21] G. Ricca, M.C. Beltrametti and A.M. Massone, Detecting curves of symmetry in images via Hough transform, Mathematics in Computer Sciences, Special Issue on Geometric Computation, 10(1) (2016), 179–205.

[22] L. Robbiano, Hyperplane sections, Gröbner bases, and Hough transforms, Journal of Pure and Applied Algebra, 219, 2434–2448 (2015).

[23] J.R. Sendra and J. Sendra, Computation of Moore–Penrose generalized inverses of matrices with meromorphic function entries, Applied Mathematics and Computation, 313C (2017), 355–366.

[24] The Shape Repository, 2011–2015, (http://visionair.ge.imati.cnr.it/ontologies/shapes/).

[25] M. Torrente and M.C. Beltrametti, Almost vanishing polynomials and an application to the Hough transform, J. Algebra Appl.,13(8) (2014), 1450057 [39 pages].

[26] M. Torrente and M.C. Beltrametti, Corrigendum for the paper "Almost vanishing polynomials and an application to the Hough transform", arXiv:1705.00158v1 [math.AG].

[27] M. Torrente, M.C. Beltrametti and J.R. Sendra, Perturbation of polynomials and applications to the Hough transform, J. Algebra, 486 (2017), 328–359.

[28] M. Torrente, M.C. Beltrametti and J.R. Sendra, r-norm bounds and metric properties for zero loci of real analytic functions Journal of Computational and Applied Mathematics, 336 (2018), 375–393.

[29] M. Torrente, M.C. Beltrametti and A.J. Sommese, Perturbation results on the zero-locus of a polynomial, J. Symbolic Comput., 80 (2017), 307–328.