Detecting Nonproperness of Likelihood Equations

Xiaoxian Tang*
School of Mathematical Sciences,
Beihang University
Beijing, China
xiaoxian@buaa.edu.cn

Bican Xia
School of Mathematical Sciences,
Peking University
Beijing, China
xbc@math.pku.edu.cn

Tianqi Zhao*
School of Mathematical Sciences,
Peking University
Beijing, China
zhaotq@pku.edu.cn

ABSTRACT

For an algebraic statistical model, the real critical points of the likelihood function are real solutions to a parametric polynomial system called a likelihood-equation system. Given a likelihood-equation system, a fundamental problem is the real root classification, i.e., classifying the data according to the number of real solutions. The nonproperness set, an essential component of the discriminant variety, plays a crucial role in the real root classification. We develop a novel method for computing nonproperness sets of likelihood equations. We show experimentally that it is far more efficient than the standard method in the literature.

CCS CONCEPTS

• **Computing methodologies** → **Symbolic and algebraic manipulation**; *Algebraic algorithms.*

KEYWORDS

Maximum likelihood estimation, Likelihood equation, Real root classification, Nonproperness

ACM Reference Format:
Xiaoxian Tang, Bican Xia, and Tianqi Zhao. 2021. Detecting Nonproperness of Likelihood Equations. In *International Symposium on Symbolic and Algebraic Computation (ISSAC '21), July 18–22, 2021, Saint Petersburg, Russia.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/...

1 INTRODUCTION

This work is motivated by the *real root classification* problem of *Lagrange likelihood equations* (1) [32, 33, 36]. Lagrange likelihood equations form a parametric polynomial system $\boldsymbol{f}$ in *probability variables* $p_0, \ldots, p_n$, *Lagrange multipliers* $\lambda_1, \ldots, \lambda_{s+1}$, and *parameters* $u_0, \ldots, u_n$ representing the data obtained from statistical experiments. It is well known that all critical points of the *maximum likelihood estimation* (MLE) problem are real solutions to the likelihood equations [1, 10, 11, 21–24]. Classifying the parameters (i.e., the data) according to the number of real solutions of likelihood equations (i.e., the real critical points of the MLE problem) is an important and a challenging problem since it is a specific *real quantifier elimination* problem [2–5, 7–9, 13, 14, 17, 20, 25, 29–31, 34, 35, 37]. In principle, the real root classification problem can be carried out automatically using some known software systems [12, 28, 39]. However, the expense of these tools is way beyond current computing capabilities since the likelihood-equation systems usually contain a lot of variables and parameters even for small statistical models.

One key step of real root classification is to compute the *discriminant variety*, which consists of two components: the discriminant locus and the nonproperness set, denoted by $\mathcal{V}(\boldsymbol{f})_J$ and $\mathcal{V}(\boldsymbol{f})_\infty$, respectively (see Definition 2.8). Both components are algebraically closed. The number of real solutions only changes when the parameters (data) pass the discriminant variety $\mathcal{V}(\boldsymbol{f})_J \cup \mathcal{V}(\boldsymbol{f})_\infty$. It is well known that when the parameters pass the discriminant locus, at least two simple real solutions might merge into a multiple one. When the parameters pass the nonproperness set, some real solutions might go to infinity. So, the number of real solutions might be "unusual" when the parameters are exactly located in the discriminant variety. For example, for the 3×3 symmetric matrix model [32, Example 6], the number of real solutions for the parameters from the complement of $\mathcal{V}(\boldsymbol{f})_J \cup \mathcal{V}(\boldsymbol{f})_\infty$ (an open region) is 2 or 6. However, the number of real solutions can be only 1 for some parameters from the nonproperness set. The standard methods for computing $\mathcal{V}(\boldsymbol{f})_J$ and $\mathcal{V}(\boldsymbol{f})_\infty$ are based on computing Gröbner bases [28]. However, these methods are not applicable for larger models. How to efficiently compute the generator of the discriminant locus $\mathcal{V}(\boldsymbol{f})_J$ for likelihood-equation systems was studied in [32, 33, 36].

The goal of this paper is to efficiently compute the generator of the nonproperness set for a given system of likelihood equations. In our setting, the likelihood-equation system is a general zero-dimensional system, and we assume the nonproperness set has codimension one. More formally, we have the following **problem statement**:
Input: Likelihood equations

$$f_0, \ldots, f_{n+s+1} \in \mathbb{Q}[u_0, \ldots, u_n, p_0, \ldots, p_n, \lambda_1, \ldots, \lambda_{s+1}].$$

Output: A generator of the nonproperness set, which is a polynomial in $\mathbb{Q}[u_0, \ldots, u_n]$.

Our **main contribution** is a probabilistic algorithm (see Algorithm 2) for computing nonproperness sets of Lagrange likelihood equations. We implemented the algorithm with `Maple2019`. Our experiments show that Algorithm 2 is significantly more efficient than the method which directly uses the standard approach in [28] for computing nonproperness sets (Algorithm 1). We summarize the computational timings for both methods in Table 1. See Section 5 for further details. The crucial idea to save computational time is

*Corresponding author

ISSAC '21, July 18–22, 2021, Saint Petersburg, Russia

ACM ISBN 978-1-4503-6084-5/19/07...$15.00
https://doi.org/10.1145/...

Models	$\#p_i$	MLD	Timings	
			Standard	Algorithm 2
1	4	3	0.465 s	**0.272** s
2	6	2	1.813 s	**0.** s
3	6	4	17.992 s	**0.** s
4	6	6	∞	**70.482** s
5	5	12	∞	**10649.391** s
6	9	10	∞	**1234.218** s
7	5	23	∞	*> 11 d*
8	8	14	∞	**6432.788** s
9	8	9	∞	**1251.420** s

Table 1: Timings for computing nonproperness sets (s: seconds; d: days). The models are available online (see Section 5.1). The column "Timings" compares the runtimes of Algorithm 1 (via a regular FGb [19] Gröbner basis computation) and Algorithm 2 (see the difference between "∞" and "*> 11 d*" in Section 5.2).

an observation that the generators of the nonproperness sets have lots of linear factors (for instance, see (20)).

The rest of this article is organized as follows. In Section 2, we recall some basic notions and definitions from algebraic statistics, and we review the notion of discriminant variety. In Section 3, we present the main theorem (Theorem 3.10) and the proofs. In Section 4, we review the standard method (Algorithm 1) for any general zero-dimensional system, and introduce Algorithm 2 and a list of sub-algorithms for computing the nonproperness sets of likelihood equations. The correctness of Algorithm 2 is guaranteed by Theorem 3.10. In Section 5, we explain the implementation details and compare the efficiency of Algorithm 2 with Algorithm 1. We end this work with a summary, see Section 6.

2 LIKELIHOOD EQUATIONS

We assume that the readers are familiar with the fundamental concepts of computational algebraic geometry. For a general overview, we refer the readers to [15]. Below, we recall the basic concepts from algebraic statistics. Also, we review general zero-dimensional system and discriminant variety.

Throughout the rest of the paper, we use bold letters for vectors, e.g., $\boldsymbol{z} = (z_1, \ldots, z_n)$. For $h \in \mathbb{Q}[\boldsymbol{z}]$, we denote the *total degree* of h by $\deg(h)$ and the degree of h with respect to (*w.r.t.*) a particular variable z_j by $\deg(h, z_j)$. We denote by $\mathrm{coeff}(h, z_j^i)$ the *coefficient* of h with respect to the monomial z_j^i. For $H \subseteq \mathbb{Q}[\boldsymbol{z}]$, we denote by $\langle H \rangle$ the *ideal* generated by H in $\mathbb{Q}[\boldsymbol{z}]$, and by $\mathcal{V}(H)$ the *affine variety* $\{\boldsymbol{z} \in \mathbb{C}^n \mid h(\boldsymbol{z}) = 0 \text{ for all } h \in H\}$. For any ideal $\mathcal{I} \subseteq \mathbb{Q}[\boldsymbol{z}]$, we denote by $\sqrt{\mathcal{I}}$ the *radical* of $\mathcal{I}$. For any polynomial $h \in \mathbb{Q}[\boldsymbol{z}]$, we denote by $\sqrt{h}$ the *square-free part* of h.

Definition 2.1 (**Probability Simplex**). The *n-dimensional probability simplex* is defined as $\Delta_n := \{(p_0, \ldots, p_n) \in \mathbb{R}^{n+1} \mid p_0 > 0, \ldots, p_n > 0,\ p_0 + \cdots + p_n = 1\}$.

Definition 2.2 (**Algebraic Statistical Model**). Given homogenous polynomials $g_1, \ldots, g_s$ in $\mathbb{Q}[p_0, \ldots, p_n]$ such that (*s.t.*)

$$\mathcal{V} := \mathcal{V}(\{g_1, \ldots, g_s\}) \ (\subsetneq \mathbb{C}^{n+1})$$

is irreducible and generically reduced, we define an *algebraic statistical model* as $\mathcal{M} := \mathcal{V} \cap \Delta_n$. If $g_1, \ldots, g_s$ are algebraically independent, then we say $\{g_1, \ldots, g_s\}$ is a set of *independent model invariants*.

Given an algebraic statistical model $\mathcal{M}$ and a *data vector* $\boldsymbol{u} := (u_0, \ldots, u_n) \in \mathbb{R}^{n+1}_{\geq 0}$, a fundamental problem in statistics is the *maximum likelihood estimation* (MLE) problem:

$$\max\ \Pi_{k=0}^{n} p_k^{u_k} \ \text{ subject to } (p_0, \ldots, p_n) \in \mathcal{M}.$$

One way to solve the above optimization problem is to solve likelihood equations formulated by the Lagrange multiplier method.

Definition 2.3 (**Lagrange Likelihood Equations**). Given an algebraic statistical model $\mathcal{M}$ with independent model invariants $g_1, \ldots, g_s \in \mathbb{Q}[p_0, \ldots, p_n]$, the polynomial set $\boldsymbol{f} = \{f_0, \ldots, f_{n+s+1}\}$ defined in (1) is said to be the system of *Lagrange likelihood equations* (or, simply *likelihood equations* in this paper) of $\mathcal{M}$ when set equal to zeros:

$$\begin{array}{rcl} f_0(\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{\lambda}) & := & p_0 \cdot \left(\frac{\partial g_1}{\partial p_0} \lambda_1 + \cdots + \frac{\partial g_s}{\partial p_0} \lambda_s + \lambda_{s+1} \right) - u_0, \\ & \vdots & \\ f_n(\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{\lambda}) & := & p_n \cdot \left(\frac{\partial g_1}{\partial p_n} \lambda_1 + \cdots + \frac{\partial g_s}{\partial p_n} \lambda_s + \lambda_{s+1} \right) - u_n, \\ f_{n+1}(\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{\lambda}) & := & g_1(p_0, \ldots, p_n), \\ & \vdots & \\ f_{n+s}(\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{\lambda}) & := & g_s(p_0, \ldots, p_n), \\ f_{n+s+1}(\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{\lambda}) & := & g_{s+1}(p_0, \ldots, p_n) = p_0 + \cdots + p_n - 1. \end{array} \tag{1}$$

Herein, $\boldsymbol{u} = (u_0, \ldots, u_n)$, $\boldsymbol{p} := (p_0, \ldots, p_n)$, and $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_{s+1})$.

We remark that by the notion $\mathcal{V}(\boldsymbol{f})$, we mean the variety

$$\{(\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{\lambda}) \in \mathbb{C}^{n+1} \times \mathbb{C}^{n+1} \times \mathbb{C}^{s+1} \mid f_0 = 0, \ldots, f_{n+s+1} = 0\}.$$

In Theorem 2.4 and later in Definition 2.8, we define π as the *canonical projection*: $\mathbb{C}^{n+1} \times \mathbb{C}^{n+s+2} \rightarrow \mathbb{C}^{n+1}$ such that $\pi(\boldsymbol{u}, \boldsymbol{p}, \boldsymbol{\lambda}) = \boldsymbol{u}$.

Theorem 2.4. *[24] Given a system of Lagrange likelihood equations $\boldsymbol{f}$ defined as in* (1)*, there exists an affine variety $V \subsetneq \mathbb{C}^{n+1}$, and a non-negative integer N such that for any $\boldsymbol{b}^* \in \mathbb{C}^{n+1} \backslash V$,*

$$\left|\pi^{-1}(\boldsymbol{b}^*) \cap \mathcal{V}(\boldsymbol{f})\right| = N.$$

Definition 2.5 (**Maximum-Likelihood-Degree**). [24] Given an algebraic statistical model $\mathcal{M}$ with a system of Lagrange likelihood equations $\boldsymbol{f}$ defined as in (1), the non-negative integer N stated in Theorem 2.4 is called the *ML-degree* of $\mathcal{M}$.

Definition 2.6 (**General Zero-Dimensional System**). A polynomial set $\boldsymbol{H} = \{h_1, \ldots, h_m\} \subseteq \mathbb{Q}[a_1, \ldots, a_k, x_1, \ldots, x_m]$ (here, $a_1, \ldots, a_k$ are parameters and $x_1, \ldots, x_m$ are variables) is called a *general zero-dimensional system* if there exists an affine variety $V \subsetneq \mathbb{C}^k$ such that for any $\boldsymbol{b} = (b_1, \ldots, b_k) \in \mathbb{C}^k \backslash V$, the equations $h_1(\boldsymbol{b}) = \cdots = h_m(\boldsymbol{b}) = 0$ satisfy:

(1) the number of complex solutions is a positive constant;
(2) all complex solutions are distinct;
(3) $x_1^{(1)} \neq x_1^{(2)}$ for every pair of distinct complex solutions $\boldsymbol{x}^{(1)} = (x_1^{(1)}, \ldots, x_m^{(1)})$ and $\boldsymbol{x}^{(2)} = (x_1^{(2)}, \ldots, x_m^{(2)})$.

Assumption 2.7. In this article, we assume any likelihood-equation system $\boldsymbol{f}$ is a general zero-dimensional system. Here, we view $u_0, \ldots, u_n$ as the parameters $a_1, \ldots, a_k$ stated in Definition 2.6, and we view $p_0, \ldots, p_n, \lambda_1, \ldots, \lambda_{s+1}$ as the variables $x_1, \ldots, x_m$.

Definition 2.8. Given a system of Lagrange likelihood equations $\boldsymbol{f}$ defined as in (1), we have the following:

• $\mathcal{V}(\boldsymbol{f})_\infty$ denotes the *set of nonproperness* of $\boldsymbol{f}$, i.e., the set of the $\boldsymbol{u} \in \overline{\pi(\mathcal{V}(\boldsymbol{f}))}$ such that there does not exist a compact neighborhood U of $\boldsymbol{u}$ where $\pi^{-1}(U) \cap \mathcal{V}(\boldsymbol{f})$ is compact;

• $\mathcal{V}(\boldsymbol{f})_J$ denotes $\overline{\pi(\mathcal{V}(\boldsymbol{f}) \cap \mathcal{V}(J))}$ where J denotes the determinant below

$$\det \begin{bmatrix} \frac{\partial f_0}{\partial p_0} & \cdots & \frac{\partial f_0}{\partial p_n} & \frac{\partial f_0}{\partial \lambda_1} & \cdots & \frac{\partial f_0}{\partial \lambda_{s+1}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{n+s+1}}{\partial p_0} & \cdots & \frac{\partial f_{n+s+1}}{\partial p_n} & \frac{\partial f_{n+s+1}}{\partial \lambda_1} & \cdots & \frac{\partial f_{n+s+1}}{\partial \lambda_{s+1}} \end{bmatrix}.$$

Geometrically, $\mathcal{V}(\boldsymbol{f})_\infty$ is the set of the data $\boldsymbol{u}$ such that the Lagrange likelihood equations have some solutions $(\boldsymbol{p}, \boldsymbol{\lambda})$ at infinity; this is the closure of the set of "nonproperness" as defined in [26, page 1] and [35, page 3]. Geometrically, $\mathcal{V}(\boldsymbol{f})_J$ is the closure of the union of the projection of the singular locus of $\mathcal{V}(\boldsymbol{f})$ and the set of critical values of the restriction of π to the regular locus of $\mathcal{V}(\boldsymbol{f})$ [28, Definition 2]. Both $\mathcal{V}(\boldsymbol{f})_\infty$ and $\mathcal{V}(\boldsymbol{f})_J$ play key roles in real root classification of likelihood equations because the complement of $\mathcal{V}(\boldsymbol{f})_\infty \cup \mathcal{V}(\boldsymbol{f})_J$ defines open connected components such that the number of real solutions is uniform over each open connected component (e.g, see [33, Theorem 2]).

3 THEOREMS

In order to make our notation simpler, for any i $(0 \le i \le n)$, let $x_i := p_i$, and for any i $(1 \le i \le s+1)$, let $x_{n+i} := \lambda_i$. Let $\boldsymbol{x} := (x_0, \ldots, x_{n+s+1})$. Then for a likelihood-equation system $\boldsymbol{f}$ defined as in (1), we have $\boldsymbol{f} \subseteq \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]$.

Assume $u_n < \cdots < u_0$ and $x_0 < \cdots < x_{n+s+1}$. We denote by $\prec_{\boldsymbol{u}}$ (or, by $\prec_{\boldsymbol{x}}$) the graded reverse lex order on the monomials only depending on $\boldsymbol{u}$ (or, only depending on $\boldsymbol{x}$). We denote by $\prec_{\boldsymbol{u}}^{lex}$ the lexicographic order on the monomials depending on $\boldsymbol{u}$.

For any admissible monomial order on $\boldsymbol{u}$ (e.g., $\prec_{\boldsymbol{u}}$), and for any admissible monomial order on $\boldsymbol{x}$ (e.g., $\prec_{\boldsymbol{x}}$), we denote by $(\prec_{\boldsymbol{u}}, \prec_{\boldsymbol{x}})$ an admissible *block monomial order* such that $\boldsymbol{u} \ll \boldsymbol{x}$ for the monomials in $\mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]$.

For any polynomial $g \in \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]$, and for any block monomial order (e.g., $(\prec_{\boldsymbol{u}}, \prec_{\boldsymbol{x}})$), we denote by $\mathrm{lm}_{\prec_{\boldsymbol{u}}, \prec_{\boldsymbol{x}}}(g)$ and $\mathrm{lc}_{\prec_{\boldsymbol{u}}, \prec_{\boldsymbol{x}}}(g)$ the leading monomial and the leading coefficient of g *w.r.t.* $(\prec_{\boldsymbol{u}}, \prec_{\boldsymbol{x}})$, respectively. And, for the monomial order $\prec_{\boldsymbol{x}}$, we denote by $\mathrm{lm}_{\prec_{\boldsymbol{x}}}(g)$ and $\mathrm{lc}_{\prec_{\boldsymbol{x}}}(g)$ the leading monomial and the leading coefficient of g *w.r.t.* $\prec_{\boldsymbol{x}}$, respectively. For any set $H \subseteq \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]$, define $\mathrm{lc}_{\prec_{\boldsymbol{x}}}(H) := \{\mathrm{lc}_{\prec_{\boldsymbol{x}}}(h) \mid h \in H\}$ $(\subseteq \mathbb{Q}[\boldsymbol{u}])$.

Suppose $\boldsymbol{f}$ is a likelihood-equation system defined as in (1). Let $G^{(r)}$ be the reduced Gröbner basis of the ideal $\langle \boldsymbol{f} \rangle \subseteq \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]$ *w.r.t.* $(\prec_{\boldsymbol{u}}, \prec_{\boldsymbol{x}})$. Let $G^{(\ell)}$ be the reduced Gröbner basis of the ideal $\langle \boldsymbol{f} \rangle \subseteq \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]$ *w.r.t.* $(\prec_{\boldsymbol{u}}^{lex}, \prec_{\boldsymbol{x}})$.

For any $j \in \{r, \ell\}$, and for any $i \in \{0, \ldots, n+s+1\}$, we define

$$G_i^{(j)} := \{g \in G^{(j)} \mid \exists\, k \ge 1,\ s.t.\ \mathrm{lm}_{\prec_{\boldsymbol{x}}}(g) = x_i^k\}\ (\subseteq \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]),$$

and

$$C_i^{(j)} := \mathrm{lc}_{\prec_{\boldsymbol{x}}}(G_i^{(j)})\ (\subseteq \mathbb{Q}[\boldsymbol{u}]).$$

Assumption 3.1. We assume that the ideals $\langle C_i^{(r)} \rangle$ and $\langle C_i^{(\ell)} \rangle$ are principal, for any i $(0 \le i \le n+s+1)$.

Under Assumption 3.1, we can assume that $\langle C_i^{(j)} \rangle = \langle h_i^{(j)} \rangle$, where $h_i^{(j)}$ is a monic polynomial in $\mathbb{Q}[\boldsymbol{u}]$ (i.e., $\mathrm{lc}_{\prec_{\boldsymbol{u}}}(h_i^{(r)}) = 1$, and $\mathrm{lc}_{\prec_{\boldsymbol{u}}^{lex}}(h_i^{(\ell)}) = 1$).

Lemma 3.2. *Under Assumption 3.1, for any i $(0 \le i \le n+s+1)$, we have* $\sqrt{h_i^{(r)}} = \sqrt{h_i^{(\ell)}}$.

Proof. By [28, Lemma 5] (see [28, line 3, page 649]), we have $\mathcal{V}(C_i^{(r)}) = \mathcal{V}(C_i^{(\ell)})$, and so, $\mathcal{I}(\mathcal{V}(C_i^{(r)})) = \mathcal{I}(\mathcal{V}(C_i^{(\ell)}))$. Recall that $\langle C_i^{(j)} \rangle = \langle h_i^{(j)} \rangle$ $(j \in \{r, \ell\})$. So, we have $\sqrt{h_i^{(r)}} = \sqrt{h_i^{(\ell)}}$. □

We remark that we can view $u_1, \ldots, u_n$ as parameters, and view $u_0, \boldsymbol{x}$ as variables. For any $g \in \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]$, we denote by $\mathrm{lm}_{\prec_{u_0}, \prec_{\boldsymbol{x}}}(g)$ and $\mathrm{lc}_{\prec_{u_0}, \prec_{\boldsymbol{x}}}(g)$ the leading monomial and the leading coefficient of g *w.r.t.* $(\prec_{u_0}, \prec_{\boldsymbol{x}})$, respectively.

For any generic point $\boldsymbol{a} := (a_1, \ldots, a_n) \in \mathbb{Q}^n$, define $\sigma_{\boldsymbol{a}}$ as a homomorphism $\sigma_{\boldsymbol{a}}:\ \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}] \to \mathbb{Q}[u_0, \boldsymbol{x}]$ such that

$$\sigma_{\boldsymbol{a}}(g) = g(u_0, a_1, \ldots, a_n, \boldsymbol{x}). \tag{2}$$

When the context is clear, we simply denote $\sigma_{\boldsymbol{a}}$ by σ. Define

$$\overline{\sigma}(g) := \frac{\sigma(g)}{\mathrm{lc}_{\prec_{u_0}, \prec_{\boldsymbol{x}}}(\sigma(g))},$$

for any $g \in \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]$.

Let $G^{(\sigma)}$ be the reduced Gröbner basis of the ideal $\langle \sigma(\boldsymbol{f}) \rangle \subseteq \mathbb{Q}[u_0, \boldsymbol{x}]$ *w.r.t.* $(\prec_{u_0}, \prec_{\boldsymbol{x}})$. For any i $(0 \le i \le n+s+1)$, define

$$G_i^{(\sigma)} := \{g \in G^{(\sigma)} \mid \exists\, k \ge 1,\ s.t.\ \mathrm{lm}_{\prec_{\boldsymbol{x}}}(g) = x_i^k\}\ (\subseteq \mathbb{Q}[u_0, \boldsymbol{x}]),$$

and

$$C_i^{(\sigma)} := \mathrm{lc}_{\prec_{\boldsymbol{x}}}(G_i^{(\sigma)})\ (\subseteq \mathbb{Q}[u_0]).$$

Note that the ideal $\langle C_i^{(\sigma)} \rangle \subseteq \mathbb{Q}[u_0]$ is principal for any i. Suppose that $\langle C_i^{(\sigma)} \rangle = \langle h_i^{(\sigma)} \rangle$, where $h_i^{(\sigma)}$ $(\in \mathbb{Q}[u_0])$ is monic.

Definition 3.3. [27, Definition 4.1] Given $H \subseteq \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]$, a subset H_m of H is a *noncomparable subset* of H *w.r.t.* $(\prec_{u_0}, \prec_{\boldsymbol{x}})$ if

(1) for every $p \in H$, there exists a polynomial $g \in H_m$ such that $\mathrm{lm}_{\prec_{u_0}, \prec_{\boldsymbol{x}}}(p)$ is a multiple of $\mathrm{lm}_{\prec_{u_0}, \prec_{\boldsymbol{x}}}(g)$, and

(2) for any $g_1, g_2 \in H_m$, with $g_1 \ne g_2$, $\mathrm{lm}_{\prec_{u_0}, \prec_{\boldsymbol{x}}}(g_1)$ is not a multiple of $\mathrm{lm}_{\prec_{u_0}, \prec_{\boldsymbol{x}}}(g_2)$, and $\mathrm{lm}_{\prec_{u_0}, \prec_{\boldsymbol{x}}}(g_2)$ is not a multiple of $\mathrm{lm}_{\prec_{u_0}, \prec_{\boldsymbol{x}}}(g_1)$.

For the given likelihood-equation system $\boldsymbol{f}$ and the corresponding Gröbner basis $G^{(\ell)}$, denote by G_m a noncomparable set of $G^{(\ell)}$.

Assumption 3.4. We assume the homomorphism σ satisfies

$$\sigma\left(\prod_{g \in G_m} \mathrm{lc}_{\prec_{u_0}, \prec_{\boldsymbol{x}}}(g)\right) \ne 0.$$

Lemma 3.5. *Under Assumption 3.4, we have*

$$C_i^{(\sigma)} = \mathrm{lc}_{\prec_{\boldsymbol{x}}}\left(\{g \in \overline{\sigma}(G_m) \mid \exists\, k \ge 1,\ s.t.\ \mathrm{lm}_{\prec_{\boldsymbol{x}}}(g) = x_i^k\}\right),$$

for any i $(0 \le i \le n+s+1)$.

Proof. First, by Assumption 2.7 and [16, page 1578, Lemma C.5], we have

$$G^{(\ell)} \cap \mathbb{Q}[\boldsymbol{u}] \;=\; \emptyset. \tag{3}$$

Then by [27, Theorem 4.3], $\overline{\sigma}(G_m)$ is a Gröbner basis of the ideal $\langle\sigma(\boldsymbol{f})\rangle$ (also, see [38, Lemma 2.3]). By the algorithm (e.g., [6, page 216, Algorithm REDGRÖBNER]) for computing the reduced Gröbner basis from a Gröbner basis, it is clear that $\mathrm{lm}_{<_{u_0},<_{\boldsymbol{x}}}(G^{(\sigma)}) = \mathrm{lm}_{<_{u_0},<_{\boldsymbol{x}}}(\overline{\sigma}(G_m))$. The conclusion follows immediately. □

Lemma 3.6. *Under Assumption 3.4, for any $g \in G_m$, we have*

$$\mathrm{lc}_{<_{\boldsymbol{x}}}(\overline{\sigma}(g)) \;=\; \overline{\sigma}\left(\mathrm{lc}_{<_{\boldsymbol{x}}}(g)\right).$$

Proof. We write $\mathrm{lc}_{<_{\boldsymbol{x}}}(g) = \mathrm{lc}_{<_{u_0},<_{\boldsymbol{x}}}(g)\cdot u_0^i + c_0$, where $c_0 \in \mathbb{Q}[\boldsymbol{u}]$, and $i = \deg\left(\mathrm{lc}_{<_{\boldsymbol{x}}}(g), u_0\right) \geq 0$. Then, it is straightforward to check that

$$\mathrm{lc}_{<_{\boldsymbol{x}}}(\overline{\sigma}(g)) \;=\; \overline{\sigma}\left(\mathrm{lc}_{<_{\boldsymbol{x}}}(g)\right) \;=\; u_0^i + \frac{\sigma(c_0)}{\sigma\left(\mathrm{lc}_{<_{u_0},<_{\boldsymbol{x}}}(g)\right)}.$$

□

Assumption 3.7. We assume that for any $g \in G_i^{(\ell)}$ ($i \in \{0,\ldots,n+s+1\}$), $\mathrm{lm}_{<_{\boldsymbol{u}},<_{\boldsymbol{x}}}(g) = u_0^t x_i^k$ for some $t \in \mathbb{Z}_{\geq 0}$ and $k \in \mathbb{Z}_{\geq 1}$.

Lemma 3.8. *Under Assumption 3.7, for any i ($0 \leq i \leq n+s+1$), we have $G_i^{(\ell)} \subseteq G_m$.*

Proof. If there exists $g \in G_i^{(\ell)} \backslash G_m$, by Assumption 3.7, we have $\mathrm{lm}_{<_{\boldsymbol{u}},<_{\boldsymbol{x}}}(g) = u_0^t x_i^k$, where $t \geq 0$ and $k \geq 1$. Then,

$$\mathrm{lm}_{<_{u_0},<_{\boldsymbol{x}}}(g) \;=\; u_0^t x_i^k. \tag{4}$$

Because G_m is a noncomparable set, so there exists $p \in G_m$ such that $\mathrm{lm}_{<_{u_0},<_{\boldsymbol{x}}}(p) \mid \mathrm{lm}_{<_{u_0},<_{\boldsymbol{x}}}(g)$. By (4), we have $\mathrm{lm}_{<_{u_0},<_{\boldsymbol{x}}}(p) = u_0^{t_1} x_i^{k_1}$ (which implies that $\mathrm{lm}_{<_{\boldsymbol{x}}}(p) = x_i^{k_1}$), where $t_1 \leq t$ and $k_1 \leq k$. Note also by (3), $k_1 \geq 1$. So, $p \in G_i^{(\ell)}$. Thus, by Assumption 3.7, we have $\mathrm{lm}_{<_{\boldsymbol{u}},<_{\boldsymbol{x}}}(p) = u_0^{t_1} x_i^{k_1}$, which means $\mathrm{lm}_{<_{\boldsymbol{u}},<_{\boldsymbol{x}}}(p) \mid \mathrm{lm}_{<_{\boldsymbol{u}},<_{\boldsymbol{x}}}(g)$. This is a contradiction to the fact that $G^{(\ell)}$ is the reduced Gröbner basis. So, the case $g \in G_i^{(\ell)} \backslash G_m$ will not happen. □

Lemma 3.9. *Under Assumptions 3.1, 3.4 and 3.7, for any i ($0 \leq i \leq n+s+1$), we have $h_i^{(\sigma)} \;=\; \overline{\sigma}(h_i^{(\ell)})$.*

Proof. We will first prove that

$$C_i^{(\sigma)} \;=\; \overline{\sigma}(C_i^{(\ell)}).$$

For any $q(u_0) \in C_i^{(\sigma)}$, by Lemma 3.5, there exists $p \in G_m$ such that $\mathrm{lm}_{<_{\boldsymbol{x}}}(p) = x_i^k$ (which implies $\mathrm{lc}_{<_{\boldsymbol{x}}}(p) \in C_i^{(\ell)}$), and $q(u_0) = \mathrm{lc}_{<_{\boldsymbol{x}}}(\overline{\sigma}(p))$. So, by Lemma 3.6, we have $q(u_0) = \overline{\sigma}\left(\mathrm{lc}_{<_{\boldsymbol{x}}}(p)\right)$. Thus, we have $q(u_0) \in \overline{\sigma}(C_i^{(\ell)})$.

On the other hand, for any $q(u_0) \in \overline{\sigma}(C_i^{(\ell)})$, by the fact that $C_i^{(\ell)} = \mathrm{lc}_{<_{\boldsymbol{x}}}(G_i^{(\ell)})$ (recall (3)), there exists $g \in G_i^{(\ell)}$ such that

$$q(u_0) \;=\; \overline{\sigma}\left(\mathrm{lc}_{<_{\boldsymbol{x}}}(g)\right). \tag{5}$$

(Note here, we have $\mathrm{lm}_{<_{\boldsymbol{x}}}(g) = x_i^k$ since $g \in G_i^{(\ell)}$, where $k \geq 1$.) By Lemma 3.8, we have $g \in G_m$. Then by (5) and Lemma 3.6, we have $q(u_0) = \mathrm{lc}_{<_{\boldsymbol{x}}}(\overline{\sigma}(g))$. So, by Lemma 3.5, we have $q(u_0) \in C_i^{(\sigma)}$. We complete the proof of $C_i^{(\sigma)} = \overline{\sigma}(C_i^{(\ell)})$.

By Assumption 3.1 and (3), we have $\langle h_i^{(\sigma)}\rangle \;=\; \langle\overline{\sigma}(h_i^{(\ell)})\rangle$. Recall that the polynomials $h_i^{(\sigma)}$ and $\overline{\sigma}(h_i^{(\ell)})$ are monic. So, $h_i^{(\sigma)} = \overline{\sigma}(h_i^{(\ell)})$. □

Define $\mathcal{I}^{(r)} := \bigcap_{i=0}^{n+s+1}\langle h_i^{(r)}\rangle$, and $\mathcal{I}^{(\sigma)} := \bigcap_{i=0}^{n+s+1}\langle h_i^{(\sigma)}\rangle$. Then, obviously

$$\sqrt{\mathcal{I}^{(r)}} \;=\; \langle\sqrt{\Pi_{i=0}^{n+s+1} h_i^{(r)}}\rangle \;\subseteq \mathbb{Q}[\boldsymbol{u}], \tag{6}$$

and

$$\sqrt{\mathcal{I}^{(\sigma)}} \;=\; \langle\sqrt{\Pi_{i=0}^{n+s+1} h_i^{(\sigma)}}\rangle \;\subseteq \mathbb{Q}[u_0].$$

Theorem 3.10. *Under Assumptions 3.1, 3.4 and 3.7, we have*

$$\sqrt{\Pi_{i=0}^{n+s+1} h_i^{(\sigma)}} \;=\; \overline{\sigma}(\sqrt{\Pi_{i=0}^{n+s+1} h_i^{(\ell)}}) \;=\; \overline{\sigma}(\sqrt{\Pi_{i=0}^{n+s+1} h_i^{(r)}}).$$

Proof. By Lemma 3.9, we have

$$\sqrt{\Pi_{i=0}^{n+s+1} h_i^{(\sigma)}} \;=\; \sqrt{\Pi_{i=0}^{n+s+1}\overline{\sigma}(h_i^{(\ell)})}.$$

And obviously,

$$\sqrt{\Pi_{i=0}^{n+s+1}\overline{\sigma}(h_i^{(\ell)})} \;=\; \sqrt{\overline{\sigma}(\Pi_{i=0}^{n+s+1} h_i^{(\ell)})}.$$

By [36, Proposition 3.2], we get

$$\sqrt{\overline{\sigma}(\Pi_{i=0}^{n+s+1} h_i^{(\ell)})} \;=\; \overline{\sigma}(\sqrt{\Pi_{i=0}^{n+s+1} h_i^{(\ell)}}).$$

By Lemma 3.2, we have

$$\overline{\sigma}(\sqrt{\Pi_{i=0}^{n+s+1} h_i^{(\ell)}}) \;=\; \overline{\sigma}(\sqrt{\Pi_{i=0}^{n+s+1} h_i^{(r)}}).$$

Then, we complete the proof. □

Remark 3.11. Fix any $i \in \{1,\ldots,n\}$, if we define the homomorphism in (2) as $\sigma(g) = g(a_0,\ldots,a_{i-1},v_i,a_{i+1},\ldots,a_n,\boldsymbol{x})$, then Theorem 3.10 is also correct.

4 ALGORITHMS

For a likelihood-equation system $\boldsymbol{f} = \{f_0,\ldots,f_{n+s+1}\} \subseteq \mathbb{Q}[\boldsymbol{u},\boldsymbol{x}]$, suppose Assumption 3.1 holds (our experiments show that this assumption is general enough in practice). Algorithm 1 is the standard method for computing $\mathcal{I}(\mathcal{V}(\boldsymbol{f})_\infty)$. From Algorithm 1, we see that under Assumption 3.1, $\mathcal{I}(\mathcal{V}(\boldsymbol{f})_\infty)$ is principal, and it is generated by $\sqrt{\Pi_{i=0}^{n+s+1} h_i^{(r)}}$ (recall (6)). Let $w := \sqrt{\Pi_{i=0}^{n+s+1} h_i^{(r)}}$. Remark that w is homogeneous [33, page 347]. Based on Theorem 3.10, we propose Algorithm 2 (a specialization/interpolation method inspired by [33]) for computing w. Without loss of generality, we assume $w \in \mathbb{Z}[\boldsymbol{u}]$ and the greatest common divisor (GCD) of all coefficients of w is 1. First, we write w as

$$w \;=\; LFactor^{(1)}(\boldsymbol{u}) \cdot nFactor(\boldsymbol{u}) \cdot LFactor^{(2)}(\boldsymbol{u}), \tag{7}$$

where $LFactor^{(1)}(\boldsymbol{u}) \in \mathbb{Z}[\boldsymbol{u}]$ is the product of all linear factors in which all coefficients are 1 (for instance, in (20), $LFactor^{(1)}(\boldsymbol{u})$ is the product of the first four factors), $nFactor(\boldsymbol{u}) \in \mathbb{Z}[\boldsymbol{u}]$ is the product of all nonlinear factors, and $LFactor^{(2)}(\boldsymbol{u}) \in \mathbb{Z}[\boldsymbol{u}]$ is the product of all linear factors in which not all coefficients are 1. The basic idea is to interpolate these factors one by one. The main algorithm has five steps; see Algorithm 2:

Algorithm 1: StandardMethod [28, Algorithm PROPERNESSDEFECTS]

Input : Lagrange likelihood equations $\boldsymbol{f} \subseteq \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]$, $\boldsymbol{u}$, $\boldsymbol{x}$
Output: w, the generator polynomial of $\mathcal{I}(\mathcal{V}(\boldsymbol{f})_\infty)$

$G^{(r)} \leftarrow$ the reduced Gröbner basis of the ideal $\langle \boldsymbol{f} \rangle \subseteq \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]$ *w.r.t.* $(<_{\boldsymbol{u}}, <_{\boldsymbol{x}})$
for i **from** 0 **to** $n+s+1$ **do**
    $C_i^{(r)} \leftarrow G^{(r)} \cap \mathbb{Q}[\boldsymbol{u}]$
    for g in $G^{(r)}$ **do**
        if there exists k $(k \geq 1)$ such that $\mathrm{lm}_{<_{\boldsymbol{x}}}(g) = x_i^k$ **then**
            $C_i^{(r)} \leftarrow C_i^{(r)} \cup \{\mathrm{lc}_{<_{\boldsymbol{x}}}(g)\}$
$w \leftarrow$ the generator polynomial of the ideal $\sqrt{\cap_{i=0}^{n+s+1} \langle C_i^{(r)} \rangle}$
return w

Algorithm 2: NewMethod (Main Algorithm)

Input : Lagrange likelihood equations $\boldsymbol{f} \subseteq \mathbb{Q}[\boldsymbol{u}, \boldsymbol{x}]$, $\boldsymbol{u}$, $\boldsymbol{x}$
Output: w, the generator polynomial of $\mathcal{I}(\mathcal{V}(\boldsymbol{f})_\infty)$

$b_0 \leftarrow$ a generic (nonzero) rational number
$\boldsymbol{f}(\boldsymbol{v}, \boldsymbol{x}) \leftarrow$ replace $u_0, u_1, \ldots, u_n$ with $v_0 + b_0, v_1 + v_0 + b_0, \ldots, v_n + v_0 + b_0$ in $\boldsymbol{f}$
Compute $LFactor^{(1)}(\boldsymbol{v})$ in (9) by **AllOne**$(\boldsymbol{f}(\boldsymbol{v}, \boldsymbol{x}), \boldsymbol{v}, \boldsymbol{x}, b_0)$
Record $\deg(LFactor^{(1)}(\boldsymbol{v}), v_i)$ in the list $LFactorDegree^{(1)}$
$wDegree$, $nFactorDegreeBound \leftarrow$ **Degrees**$(\boldsymbol{f}(\boldsymbol{v}, \boldsymbol{x}), \boldsymbol{v}, \boldsymbol{x})$
$LFactorDegree^{(2)} \leftarrow [\,]$
if $nFactorDegreeBound[1] > 1$ **then**
    Compute $nFactor(\boldsymbol{v})$ in (9) by **Nonlinear**$(\boldsymbol{f}(\boldsymbol{v}, \boldsymbol{x}), \boldsymbol{v}, \boldsymbol{x}, nFactorDegreeBound)$
    #Here, Algorithm **Nonlinear** *is [33, page 354, Algorithm 2].*
    Record $\deg(nFactor, v_i)$ in the list $nFactorDegree$
    for i **from** 0 **to** n **do**
        $LFactorDegree^{(2)}[i+1] \leftarrow$
          $wDegree[i+1] - nFactorDegree[i+1] - LFactorDegree^{(1)}[i+1]$
else
    for i **from** 0 **to** n **do**
        $LFactorDegree^{(2)}[i+1] \leftarrow wDegree[i+1] - LFactorDegree^{(1)}[i+1]$
if $LFactorDegree^{(2)}[1] > 0$ **then**
    Compute $LFactor^{(2)}(\boldsymbol{v})$ in (9) by
      NotAllOne$(\boldsymbol{f}(\boldsymbol{v}, \boldsymbol{x}), \boldsymbol{v}, \boldsymbol{x}, b_0, LFactor^{(1)}(\boldsymbol{v}), LFactorDegree^{(2)})$
$w(\boldsymbol{v}) \leftarrow LFactor^{(1)}(\boldsymbol{v}) \cdot nFactor(\boldsymbol{v}) \cdot LFactor^{(2)}(\boldsymbol{v})$
$w \leftarrow$ apply the inverse linear transformation to $w(\boldsymbol{v})$
return w

Step 1 In order to satisfy Assumption 3.7, we apply the following invertible linear transformation to $\boldsymbol{f}$:

$$u_0 = v_0 + b_0, \quad \text{and} \quad u_j = v_j + v_0 + b_0 \quad \text{for} \quad j = 1, \ldots, n, \qquad (8)$$

where b_0 is a generic (nonzero) rational number. See Example 4.1. Note here, after the linear transformation, the equality (7) becomes

$$w(\boldsymbol{v}) \;=\; LFactor^{(1)}(\boldsymbol{v}) \cdot nFactor(\boldsymbol{v}) \cdot LFactor^{(2)}(\boldsymbol{v}). \qquad (9)$$

Step 2 Compute the polynomial $LFactor^{(1)}(\boldsymbol{v})$ by applying a homomorphism $\sigma_{\boldsymbol{b}}$ to $\boldsymbol{f}(\boldsymbol{v}, \boldsymbol{x})$, where $\boldsymbol{b} := (b_1, \ldots, b_n)$ is a generic rational vector. See Algorithm 3 and Example 4.2.

Step 3 For any i $(0 \leq i \leq n)$, compute $\deg(w(\boldsymbol{v}), v_i)$ and an upper bound for $\deg(nFactor(\boldsymbol{v}), v_i)$. See Algorithm 4 and Example 4.3.

Step 4 Interpolate the nonlinear part $nFactor(\boldsymbol{v})$. The interpolation algorithm is [33, page 354, Algorithm 2] (Here, we call it Algorithm **Nonlinear**). See Example 4.4.

Step 5 Interpolate the polynomial $LFactor^{(2)}(\boldsymbol{v})$. See Algorithm 5 and Example 4.5. Finally, we only need to apply the inverse transformation of (8) to (9), and we will get $w(\boldsymbol{u})$.

Algorithm 3: AllOne (Sub-Algorithm of Algorithm 2)

Input : Lagrange likelihood equations after the linear transformation $\boldsymbol{f}(\boldsymbol{v}, \boldsymbol{x}) \subseteq \mathbb{Q}[\boldsymbol{v}, \boldsymbol{x}]$, $\boldsymbol{v}$, $\boldsymbol{x}$ and a generic (nonzero) rational number b_0
Output: $LFactor^{(1)}$, the product of all linear factors of $w(\boldsymbol{v})$ in which all coefficients are 1

$\frac{\alpha_1}{\beta_1}, \ldots, \frac{\alpha_n}{\beta_n} \leftarrow$ some generic rational numbers
#Here, for any $\boldsymbol{i} \in \{1, \ldots, \boldsymbol{n}\}$, $\alpha_{\boldsymbol{i}}, \beta_{\boldsymbol{i}} \in \mathbb{Z}\backslash\{0\}$, *and* $\mathrm{GCD}(\alpha_{\boldsymbol{i}}, \beta_{\boldsymbol{i}}) = 1$.
$\boldsymbol{f}^* \leftarrow$ replace $v_1, \ldots, v_n$ with $\frac{\alpha_1}{\beta_1}, \ldots, \frac{\alpha_n}{\beta_n}$ in $\boldsymbol{f}(\boldsymbol{v}, \boldsymbol{x})$
$w^* \leftarrow$ **StandardMethod**$(\boldsymbol{f}^*, v_0, \boldsymbol{x})$
for every factor q of w^* **do**
    if $\deg(q, v_0) = 1$ **then**
        $c_0 \leftarrow \mathrm{coeff}(q, v_0^0)$
        $c_1 \leftarrow \mathrm{coeff}(q, v_0)$
        #Here, $\boldsymbol{c}_0, \boldsymbol{c}_1 \in \mathbb{Z}\backslash\{0\}$, *and* $\mathrm{GCD}(\boldsymbol{c}_0, \boldsymbol{c}_1) = 1$.
        if there exist K $(0 < K \leq |H_1|)$ and $R \subseteq \{0, \ldots, \boldsymbol{n}\}$ such that $H_1 = Kc_1$ and $H_2 = Kc_0$ **then**
            $LFactor^{(1)} \leftarrow LFactor^{(1)} \cdot \left(c_1 v_0 + c_1 b_0 + \sum_{j \in R\backslash\{0\}} v_j\right)$
        #Here, $H_1 := \beta_0 \cdot |R| \cdot \left(\Pi_{i \in T\backslash\{0\}} \beta_i\right)$,
        $H_2 := \alpha_0 \cdot |R| \cdot \left(\Pi_{i \in T\backslash\{0\}} \beta_i\right) + \beta_0 \cdot \sum_{i \in T\backslash\{0\}} \left(\alpha_i \cdot \Pi_{j \in T\backslash\{0,i\}} \beta_j\right)$,
        where α_0 and β_0 are two integers such that $b_0 = \frac{\alpha_0}{\beta_0}$ $(\mathrm{GCD}(\alpha_0, \beta_0) = 1)$.
return $LFactor^{(1)}$

Algorithm 4: Degrees (Sub-Algorithm of Algorithm 2)

Input : Lagrange likelihood equations after the linear transformation $\boldsymbol{f}(\boldsymbol{v}, \boldsymbol{x}) \subseteq \mathbb{Q}[\boldsymbol{v}, \boldsymbol{x}]$, $\boldsymbol{v}$, $\boldsymbol{x}$
Output: $wDegree$ and $nFactorDegreeBound$ where

- $wDegree$ is a list, whose $(j+1)$-th entry is $\deg(w(\boldsymbol{v}), v_j)$ for $j = 0, \ldots, n$,
- $nFactorDegreeBound$ is a list, whose $(j+1)$-th entry is an upper bound for $\deg(nFactor(\boldsymbol{v}), v_j)$ for $j = 0, \ldots, n$

for i **from** 0 **to** n **do**
    $\boldsymbol{f}^* \leftarrow$ replace $v_0, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$ with some generic rational numbers in $\boldsymbol{f}(\boldsymbol{v}, \boldsymbol{x})$
    $w^* \leftarrow$ **StandardMethod**$(\boldsymbol{f}^*, v_i, \boldsymbol{x})$
    $d_i \leftarrow \deg(w^*, v_i)$
    $wfactor_1 \leftarrow \{q \mid q \text{ is a factor of } w^*,\ \deg(q, v_i) > 1\}$
    $wfactor_2 \leftarrow \{q \mid q \text{ is a factor of } w^*,\ \deg(q, v_i) \leq 1\}$
    $k_i \leftarrow |wfactor_1|$
    #Here, k_0 *is equal to the number of all nonlinear factors of* $w(\boldsymbol{v})$.
    $t_i \leftarrow (k_0 - k_i) + \sum_{q \in wfactor_1} \deg(q, v_i)$
$wDegree \leftarrow [d_0, \ldots, d_n]$
$nFactorDegreeBound \leftarrow [t_0, \ldots, t_n]$
return $wDegree$, $nFactorDegreeBound$

4.1 Running Example

Let $\boldsymbol{f} = \{f_0, \ldots, f_7\}$, where

$$\begin{array}{lcl}
f_0 & = & p_0(\lambda_1 + (-4p_0p_5 + 2p_2p_4)\lambda_2) - u_0, \\
f_1 & = & p_1(\lambda_1 + (8p_3p_5 - 2p_4^2)\lambda_2) - u_1, \\
f_2 & = & p_2(\lambda_1 + (2p_0p_4 - 4p_2p_3)\lambda_2) - u_2, \\
f_3 & = & p_3(\lambda_1 + (8p_1p_5 - 2p_2^2)\lambda_2) - u_3, \\
f_4 & = & p_4(\lambda_1 + (2p_0p_2 - 4p_1p_4)\lambda_2) - u_4, \\
f_5 & = & p_5(\lambda_1 + (-2p_0^2 + 8p_1p_3)\lambda_2) - u_5, \\
f_6 & = & 8p_1p_3p_5 - 2p_1p_4^2 - 2p_0^2p_5 + 2p_0p_2p_4 - 2p_2^2p_3, \\
f_7 & = & p_0 + p_1 + p_2 + p_3 + p_4 + p_5 - 1.
\end{array} \qquad (10)$$

There are 6 parameters $u_0, \ldots, u_5$ and 8 variables $p_0, \ldots, p_5, \lambda_1, \lambda_2$. In the following examples, we will show step by step how to compute a polynomial $w(\boldsymbol{u})$ by Algorithm 2 such that $\mathcal{I}(\mathcal{V}(\boldsymbol{f}(\boldsymbol{u}, \boldsymbol{x}))_\infty) = \langle w(\boldsymbol{u}) \rangle$. Recall that we assume w can be factorized as in (7).

Example 4.1 (**LinearTransformation**). We apply the linear transformation (8) (by taking $b_0 = 8$) to the system $\boldsymbol{f}$ (10):

$$\begin{array}{lll}
u_0 = v_0 + 8, & u_1 = v_1 + v_0 + 8, & u_2 = v_2 + v_0 + 8, \\
u_3 = v_3 + v_0 + 8, & u_4 = v_4 + v_0 + 8, & u_5 = v_5 + v_0 + 8.
\end{array} \qquad (11)$$

Algorithm 5: NotAllOne (Sub-Algorithm of Algorithm 2)

Input : Lagrange likelihood equations after the linear transformation $\boldsymbol{f}(\boldsymbol{v},\boldsymbol{x}) \subseteq \mathbb{Q}[\boldsymbol{v},\boldsymbol{x}]$, $\boldsymbol{v}$, $\boldsymbol{x}$, a generic (nonzero) rational number b_0, $LFactor^{(1)}(\boldsymbol{v})$ (see (9)), and $LFactorDegree^{(2)}$

Output: $LFactor^{(2)}$, the product of all linear factors of $w(\boldsymbol{v})$ in which not all coefficients are 1

$d \leftarrow LFactorDegree^{(2)}[1]$

$Q \leftarrow p \times d$ array, whose (i,j)-entry is a polynomial $Q_{i,j} = 1 + \sum_{k \in R} v_k$, where $R \subseteq \{0,\ldots,n\}$, $|R| \geq 2$, the polynomials $\Pi_{j=1}^{d} Q_{i,j}$ are different, and $LFactorDegree^{(2)}[k+1] = \deg\left(\Pi_{j=1}^{d} Q_{i,j}, v_k\right)$ for any k $(0 \leq k \leq n)$ and any i

for i **from** 1 **to** p **do**
- $Q_i \leftarrow \Pi_{j=1}^{d} Q_{i,j}$
- **for** j **from** 0 **to** $d-1$ **do**
  - $S_{i,j} \leftarrow$ the set of all monomials of coeff $\left(Q_i, v_0^j\right)$

if $p = 1$ **then**
- **for** j **from** 0 **to** $d-1$ **do**
  - $M_j \leftarrow$ **Coefficients**$(\boldsymbol{f}(\boldsymbol{v},\boldsymbol{x}), \boldsymbol{v}, \boldsymbol{x}, LFactor^{(1)}(\boldsymbol{v}), j, S_{1,j})$
- $LFactor^{(2)} \leftarrow v_0^d + \sum_{j=0}^{d-1} M_j v_0^j$
- **return** $LFactor^{(2)}$

else
- **for** j **from** 1 **to** $d-1$ **do**
  - $S_j \leftarrow \cup_{i=1}^{p} S_{i,j}$
- **for** j **from** 1 **to** $d-1$ **do**
  - $M_j \leftarrow$ **Coefficients**$(\boldsymbol{f}(\boldsymbol{v},\boldsymbol{x}), \boldsymbol{v}, \boldsymbol{x}, LFactor^{(1)}(\boldsymbol{v}), j, S_j)$
  - $B_j \leftarrow$ the set of all monomials of M_j
- $L \leftarrow \{\}$
- **for** i **from** 1 **to** p **do**
  - **if** $[S_{i,1},\ldots,S_{i,d-1}] = [B_1,\ldots,B_{d-1}]$ **then**
    - $L \leftarrow$ add Q_i to L
- **for** every Q_i in L **do**
  - $M_0 \leftarrow$ **Coefficients** $(\boldsymbol{f}(\boldsymbol{v},\boldsymbol{x}), \boldsymbol{v}, \boldsymbol{x}, LFactor^{(1)}(\boldsymbol{v}), 0, S_{i,0})$
  - $LFactor^{(2)} \leftarrow v_0^d + \sum_{j=0}^{d-1} M_j v_0^j$
  - **if** $LFactor^{(2)}$ can be factorized into d linear factors **then**
    - **return** $LFactor^{(2)}$
    - **break**

Algorithm 6: Coefficients (Sub-Algorithm of Algorithm 5)

Input : Lagrange likelihood equations after the linear transformation $\boldsymbol{f}(\boldsymbol{v},\boldsymbol{x}) \subseteq \mathbb{Q}[\boldsymbol{v},\boldsymbol{x}]$, $\boldsymbol{v}$, $\boldsymbol{x}$, $LFactor^{(1)}(\boldsymbol{v})$ (see (9)), an integer t, and a set of monomials S

Output: M_t, which is equal to coeff$(LFactor^{(2)}(\boldsymbol{v}), v_0^t)$

Rename all monomials of the set S as $V_1,\ldots,V_{|S|}$

for i **from** 1 **to** $|S|$ **do**
- $b_{i,1},\ldots,b_{i,n} \leftarrow$ some generic rational numbers
- $l_2^* \leftarrow$ **Intersect**$(\boldsymbol{f}(\boldsymbol{v},\boldsymbol{x}), \boldsymbol{v}, \boldsymbol{x}, LFactor^{(1)}(\boldsymbol{v}), b_{i,1},\ldots,b_{i,n})$
- $c_i^* \leftarrow$ coeff(l_2^*, v_0^t)

$\mathcal{R} \leftarrow |S| \times |S|$ matrix whose (i,j)-entry is $V_j(b_{i,1},\ldots,b_{i,n})$

$M_t \leftarrow (V_1,\ldots,V_{|S|})\mathcal{R}^{-1}(c_1^*,\ldots,c_{|S|}^*)^T$

return M_t

*Example 4.2 (***AllOne***).* We apply Algorithm 3 to compute the factor $LFactor^{(1)}(\boldsymbol{u})$ in (7). It is sufficient to compute $LFactor^{(1)}(\boldsymbol{v})$ in (9), the polynomial after the linear transformation (11). We randomly choose $b_1 = 29$, $b_2 = 43$, $b_3 = 89$, $b_4 = 149$, and $b_5 = 247$. We substitute $v_i = b_i$ $(i \in \{1,\ldots,5\})$ into $\boldsymbol{f}(\boldsymbol{v},\boldsymbol{x})$, and we call the resulting system $\boldsymbol{f}^*$. By Algorithm 1, we compute $w^* \in \mathbb{Q}[v_0]$ such that $\mathcal{I}(\mathcal{V}(\boldsymbol{f}^*)_\infty) = \langle w^* \rangle$. We list all linear factors of w^*:

$$[3v_0+142, 3v_0+343, 3v_0+509, 6v_0+605,$$
$$2v_0+359, 4v_0+133, 4v_0+359].$$

Algorithm 7: Intersect (Sub-Algorithm of Algorithm 6)

Input : Lagrange likelihood equations after the linear transformation $\boldsymbol{f}(\boldsymbol{v},\boldsymbol{x}) \subseteq \mathbb{Q}[\boldsymbol{v},\boldsymbol{x}]$, $\boldsymbol{v}$, $\boldsymbol{x}$, $LFactor^{(1)}(\boldsymbol{v})$ (see (9)), and some generic rational numbers $b_1,\ldots,b_n$

Output: $\mathcal{L}^*$, which is equal to $LFactor^{(2)}(v_0, b_1,\ldots,b_n)$

$f^* \leftarrow$ replace $v_1,\ldots,v_n$ with $b_1,\ldots,b_n$ in $\boldsymbol{f}(\boldsymbol{v},\boldsymbol{x})$

$l_1^* \leftarrow LFactor^{(1)}(v_0, b_1,\ldots,b_n)$

$w^* \leftarrow$ **StandardMethod**$(f^*, v_0, \boldsymbol{x})$

$\mathcal{L}^* \leftarrow 1$

for every factor q of w^* **do**
- **if** $\deg(q, v_0) = 1$, and q is not a factor of l_1^* **then**
  - $\mathcal{L}^* \leftarrow \mathcal{L}^* \cdot q$

return $\mathcal{L}^*$

Note that $3b_0+b_1+b_3 = 142$, $3b_0+b_1+b_2+b_5 = 343$, $3b_0+b_3+b_4+b_5 = 509$, and $6b_0+b_1+b_2+b_3+b_4+b_5 = 605$, where $b_0 = 8$. And, the last three factors in the above list do not satisfy the condition in Algorithm 3-Line 10. So, by Theorem 3.10, we have

$$\begin{aligned} LFactor^{(1)}(\boldsymbol{v}) \;=\; & (3v_0+v_1+v_3+24) \\ & \cdot (3v_0+v_1+v_2+v_5+24) \cdot (3v_0+v_3+v_4+v_5+24) \\ & \cdot (6v_0+v_1+v_2+v_3+v_4+v_5+48). \end{aligned} \tag{12}$$

*Example 4.3 (***Degrees***).* By Algorithm 4, we compute $\deg(w(\boldsymbol{v}), v_i)$ (recorded in the list $wDegree$), and for every $i \in \{0,\ldots,5\}$, we compute an upper bound for $\deg(nFactor(\boldsymbol{v}), v_i)$ (recorded in the list $nFactorDegreeBound$). First, we compute $\deg(w(\boldsymbol{v}), v_0)$ and $\deg(nFactor(\boldsymbol{v}), v_0)$. First, we choose some generic rational numbers: $v_1 = 4$, $v_2 = 3$, $v_3 = 7$, $v_4 = 8$, and $v_5 = 9$. By Algorithm 1 and Theorem 3.10, we can compute $w(v_0, 4, 3, 7, 8, 9)$ by calling

$$\textbf{StandardMethod}(\boldsymbol{f}(v_0,4,3,7,8,9,\boldsymbol{x}), v_0, \boldsymbol{x}),$$

which has the list of factors:

$$[v_0^3+\frac{97}{2}v_0^2+\frac{1521}{2}v_0+\frac{7673}{2}, v_0+\frac{35}{3}, v_0+\frac{61}{4}, v_0+\frac{79}{6},$$
$$v_0+\frac{27}{2}, v_0+\frac{40}{3}, v_0+16, v_0+\frac{43}{4}].$$

So, $\deg(w(\boldsymbol{v}), v_0) = 10$. By [36, Proposition 3.3] (Hilbert's irreducibility theorem), we have $\deg(nFactor(\boldsymbol{v}), v_0) = 3$.

Similarly, we can compute $w(2, v_1, 4, 7, 8, 1)$, which has the list of factors:

$$[v_1+35, v_1+37, v_1+22, v_1+\frac{291}{53}, v_1+80].$$

So, we have $\deg(w(\boldsymbol{v}), v_1) = 5$. Recall that $nFactor$ is a factor of w (see (9)). Notice that if $\deg(nFactor(\boldsymbol{v}), v_1) = 0$ (i.e., v_1 does not appear in the nonlinear factor of w), then $nFactor(2, v_1, 4, 7, 8, 1)$ is a constant, and this constant will not appear in the above list of factors returned by `Maple`. So, from this computation, we conclude $\deg(nFactor(\boldsymbol{v}), v_1) \leq 1$. Similarly, for any i $(2 \leq i \leq 5)$, we can compute $\deg(w(\boldsymbol{v}), v_i)$ and an upper bound for $\deg(nFactor(\boldsymbol{v}), v_i)$. Then, we obtain $wDegree = [10, 5, 6, 5, 6, 5]$, and

$$nFactorDegreeBound = [3, 1, 2, 1, 2, 1].$$

*Example 4.4 (***Nonlinear***).* For the system $\boldsymbol{f}(\boldsymbol{v},\boldsymbol{x})$ in Example 4.1, we apply Algorithm **Nonlinear** [33, page 354, Algorithm 2] to compute $nFactor(\boldsymbol{u})$ in (7). It is sufficient to compute $nFactor(\boldsymbol{v})$ in (9), the polynomial after the linear transformation (11).

By Example 4.3, we know $\deg(nFactor(\boldsymbol{v}), v_0) = 3$. Without loss of generality, we can assume $nFactor(\boldsymbol{v})$ (after scaling) is

$$v_0^3 + P_2 v_0^2 + P_1 v_0 + P_0,$$

where $P_i \in \mathbb{Q}[v_1, \ldots, v_5]$ has total degree $3 - i$ $(i = 0, 1, 2)$. So, we can assume $P_2 = P_{21}v_1 + P_{22}v_2 + P_{23}v_3 + P_{24}v_4 + P_{25}v_5 + 24$, where $P_{2i} \in \mathbb{Q}$ for $i = 1, \ldots, 5$.

We first compute $nFactor(v_0, 27, 17, 8, 5, 26)$ which is the only nonlinear factor of $w(v_0, 27, 17, 8, 5, 26)$:

$$nFactor(v_0, 27, 17, 8, 5, 26) \;=\; v_0^3 + \frac{209}{2}v_0^2 + \frac{6693}{2}v_0 + \frac{60669}{2} \quad (13)$$

Note here, $w(v_0, 27, 17, 8, 5, 26)$ is the output of

$$\textbf{StandardMethod}(\boldsymbol{f}(v_0, 27, 17, 8, 5, 26, \boldsymbol{x}), v_0, \boldsymbol{x}).$$

Similarly, we can compute

$$\begin{aligned}
nFactor(v_0, 9, 2, 18, 20, 24) &= v_0^3 + \frac{179}{2}v_0^2 + 2462v_0 + 20420, && (14)\\
nFactor(v_0, 11, 19, 25, 13, 29) &= v_0^3 + \frac{211}{2}v_0^2 + \frac{6749}{2}v_0 + 31264, && (15)\\
nFactor(v_0, 28, 3, 14, 6, 30) &= v_0^3 + \frac{255}{2}v_0^2 + \frac{9857}{2}v_0 + 54733, && (16)\\
nFactor(v_0, 15, 4, 10, 12, 16) &= v_0^3 + \frac{155}{2}v_0^2 + 1872v_0 + 14168. && (17)
\end{aligned}$$

Comparing $\mathrm{coeff}(nFactor(v_0), v_0^2)$ in equations (13)–(17), we have the linear equations below

$$\begin{aligned}
27P_{21} + 17P_{22} + 8P_{23} + 5P_{24} + 26P_{25} + 24 &= \frac{209}{2},\\
9P_{21} + 2P_{22} + 18P_{23} + 20P_{24} + 24P_{25} + 24 &= \frac{179}{2},\\
11P_{21} + 19P_{22} + 25P_{23} + 13P_{24} + 29P_{25} + 24 &= \frac{211}{2},\\
28P_{21} + 3P_{22} + 14P_{23} + 6P_{24} + 30P_{25} + 24 &= \frac{255}{2},\\
15P_{21} + 4P_{22} + 10P_{23} + 12P_{24} + 16P_{25} + 24 &= \frac{155}{2}.
\end{aligned}$$

Solving the linear system yields $P_{21} = \frac{3}{2}, P_{22} = -\frac{1}{2}, P_{23} = \frac{3}{2}, P_{24} = -\frac{1}{2}$, and $P_{25} = \frac{3}{2}$. We can similarly compute P_1 and P_0, and we get

$$\begin{aligned}
nFactor(\boldsymbol{v}) &= v_0^3 + \frac{3}{2}v_0^2 v_1 - \frac{1}{2}v_0^2 v_2 + \frac{3}{2}v_0^2 v_3 - \frac{1}{2}v_0^2 v_4 + \frac{3}{2}v_0^2 v_5 + 2v_0 v_1 v_3\\
&\quad - v_0 v_1 v_4 + 2v_0 v_1 v_5 - \frac{1}{2}v_0 v_2^2 - v_0 v_2 v_3 + \frac{1}{2}v_0 v_2 v_4 + 2v_0 v_3 v_5\\
&\quad - \frac{1}{2}v_0 v_4^2 + 2v_1 v_3 v_5 - \frac{1}{2}v_1 v_4^2 - \frac{1}{2}v_2^2 v_3 + \cdots \qquad (18)
\end{aligned}$$

Example 4.5 (**NotAllOne**). We apply Algorithm 5 to compute $LFactor^{(2)}(\boldsymbol{v})$, the polynomial after applying (11) to $LFactor^{(2)}(\boldsymbol{u})$ in (7). By (12), (18), and $\deg(w(\boldsymbol{v}), v_i)$ ($wDegree$) computed in Example 4.3, we get $\deg(LFactor^{(2)}(\boldsymbol{v}), v_i)$ for any i $(0 \le i \le 5)$, and we record them as $LFactorDegree^{(2)} = [3, 1, 2, 1, 2, 1]$, where

$$LFactorDegree^{(2)}[i+1] = \deg(LFactor^{(2)}(\boldsymbol{v}), v_i).$$

We can write $LFactor^{(2)}(\boldsymbol{v})$ as:

$$LFactor^{(2)}(\boldsymbol{v}) \;=\; M_3 v_0^3 + M_2 v_0^2 + M_1 v_0 + M_0,$$

where $M_3 \in \mathbb{Z}_{>0}$ and $M_i \in \mathbb{Q}[v_1, \ldots, v_5]$ for $i = 0, 1, 2$.

By (11), $\deg(LFactor^{(2)}(\boldsymbol{u})) = \deg(LFactor^{(2)}(\boldsymbol{v}), v_0) = 3$, so $LFactor^{(2)}(\boldsymbol{u})$ has three linear factors. According to $LFactorDegree^{(2)}$ and (11), $LFactor^{(2)}(\boldsymbol{u})$ can be factorized as only one of the following 37 possible products:

$$\begin{aligned}
&(1)\ (\Box u_0 + \Box u_2 + \Box u_4)(\Box u_0 + \Box u_2 + \Box u_4)(\Box u_0 + \Box u_1 + \Box u_3 + \Box u_5),\\
&(2)\ (\Box u_0 + \Box u_1)(\Box u_0 + \Box u_2 + \Box u_4)(\Box u_0 + \Box u_2 + \Box u_3 + \Box u_4 + \Box u_5)\\
&\qquad\vdots\\
&(37)\ (\Box u_0 + \Box u_2 + \Box u_5)(\Box u_0 + \Box u_3 + \Box u_4)(\Box u_0 + \Box u_1 + \Box u_2 + \Box u_4),
\end{aligned}$$

where $\Box$ denotes an unknown coefficient (integer) we need to compute. Expanding the above 37 polynomials and applying the linear transformation (11), we obtain 37 polynomials in $\mathbb{Q}[\boldsymbol{v}]$, say Q_i $(1 \le i \le 37)$. Assume

$$Q_i \;=\; Q_{(i,3)}v_0^3 + Q_{(i,2)}v_0^2 + Q_{(i,1)}v_0 + Q_{(i,0)},$$

where $Q_{(i,3)} \in \mathbb{Z}_{>0}$, and $Q_{(i,j)} \in \mathbb{Q}[v_1, \ldots, v_5]$ for $j = 0, 1, 2$. Let $S_{i,j}$ be the set of all possible monomials of $Q_{(i,j)}$ for any i, j. For instance,

$$\begin{aligned}
S_{1,2} &= \{v_1, v_2, v_3, v_4, v_5, 1\},\\
S_{1,1} &= \{v_1v_2, v_1v_4, v_2^2, v_2v_3, v_2v_4, v_2v_5, v_3v_4, v_4^2, v_4v_5, v_1, v_2, v_3, v_4, v_5, 1\},\\
S_{1,0} &= \{v_1v_2^2, v_1v_2v_4, v_1v_4^2, v_2^2v_3, v_2^2v_5, v_2v_3v_4, v_2v_4v_5, v_3v_4^2, v_4^2v_5, v_1v_2, v_1v_4, v_2^2, v_2v_3,\\
&\quad v_2v_4, v_2v_5, v_3v_4, v_4^2, v_4v_5, v_1, v_2, v_3, v_4, v_5, 1\}.
\end{aligned}$$

Let $S_j := \cup_{i=1}^{37} S_{i,j}$ for $j = 1, 2$. Then, S_j is the set of all possible monomials those may appear in M_j. First, we apply Algorithm **Nonlinear**, and interpolate M_1 and M_2. The computation results are

$$\begin{aligned}
M_2 &= 32v_1 + 32v_2 + 32v_3 + 32v_4 + 32v_5 + 1536,\\
M_1 &= 8v_1v_2 + 16v_1v_3 + 16v_1v_4 + 16v_1v_5 + 4v_2^2 + 16v_2v_3 + 12v_2v_4\\
&\quad + 8v_2v_5 + 8v_3v_4 + 16v_3v_5 + 4v_4^2 + 8v_4v_5 + 512v_1 + 512v_2\\
&\quad + 512v_3 + 512v_4 + 512v_5 + 12288.
\end{aligned}$$

Let B_j be the set of all monomials of M_j, where $1 \le j \le 2$. Note that $B_j \subseteq S_j$. Define $L := \{Q_i \mid S_{i,j} = B_j \text{ for } j = 1, 2\}$. Here, we have $L = \{Q_3, Q_{13}, Q_{22}, Q_{24}, Q_{25}, Q_{27}, Q_{29}, Q_{32}, Q_{37}\}$. Next, for every Q_i in L, we compute $Q_{(i,3)}$ and $Q_{(i,0)}$ until $Q_{(i,3)}v_0^3 + M_2v_0^2 + M_1v_0 + Q_{(i,0)}$ can be factorized into three linear factors. Here, the only polynomial in L satisfying the condition is Q_{24}. And, the corresponding results are $M_3 = Q_{(24,3)} = 64$, and

$$\begin{aligned}
M_0 = Q_{(24,0)} &= 4v_1v_2v_3 + 2v_1v_2v_4 + 4v_1v_3v_4 + 8v_1v_3v_5 + 2v_1v_4^2\\
&\quad + 4v_1v_4v_5 + 2v_2^2v_3 + v_2^2v_4 + 2v_2v_3v_4 + 4v_2v_3v_5 + v_2v_4^2\\
&\quad + 2v_2v_4v_5 + 64v_1v_2 + 128v_1v_3 + 128v_1v_4 + 128v_1v_5\\
&\quad + 32v_2^2 + 128v_2v_3 + 96v_2v_4 + 64v_2v_5 + 64v_3v_4\\
&\quad + 128v_3v_5 + 32v_4^2 + 64v_4v_5 + 2048v_1 + 2048v_2\\
&\quad + 2048v_3 + 2048v_4 + 2048v_5 + 32768.
\end{aligned}$$

By Theorem 3.10, we obtain

$$\begin{aligned}
LFactor^{(2)}(\boldsymbol{v}) &= (4v_0 + 2v_1 + v_2 + 32)\cdot(4v_0 + 2v_3 + v_4 + 32)\\
&\quad \cdot (4v_0 + v_2 + v_4 + 2v_5 + 32). \qquad (19)
\end{aligned}$$

Finally, we apply the inverse of the linear transformation (11) to $LFactor^{(1)}(\boldsymbol{v})$ in (12), $nFactor(\boldsymbol{v})$ in (18) and $LFactor^{(2)}(\boldsymbol{v})$ in (19), and we get the generator polynomial w of the ideal $\mathcal{I}(\mathcal{V}(\boldsymbol{f})_\infty)$:

$$\begin{aligned}
w &= LFactor^{(1)}(\boldsymbol{u})\cdot nFactor(\boldsymbol{u})\cdot LFactor^{(2)}(\boldsymbol{u})\\
&= (u_0 + u_1 + u_3)\cdot(u_1 + u_2 + u_5)\\
&\quad \cdot (u_3 + u_4 + u_5)\cdot(u_0 + u_1 + u_2 + u_3 + u_4 + u_5)\\
&\quad \cdot (u_0^2u_5 - u_0u_2u_4 - 4u_1u_3u_5 + u_1u_4^2 + u_2^2u_3)\\
&\quad \cdot (u_0 + 2u_1 + u_2)\cdot(u_0 + 2u_3 + u_4)\cdot(u_2 + u_4 + 2u_5). \qquad (20)
\end{aligned}$$

5 COMPUTATIONAL RESULTS

5.1 Implementation

Testing models, Maple code and computational results are available online via: https://github.com/zhao-tq/nonproperness.

Software We implemented Algorithm 2 with Maple2019, where we use the FGb command fgb_gbasis for computing reduced Gröbner bases in Algorithm 1-Line 1.

Hardware and System We used a 2.3 GHz Intel Core i7 processor (16 GB of RAM) under Ubuntu 18.04.5.

Testing Models Testing Models are chosen from the literatures [18, 24] and have been tested by both standard method (i.e. Algorithm 1) and Algorithm 2.

5.2 Computing nonproperness sets

Table 1 (in Section 1) compares the timings of Algorithm 1 and Algorithm 2.

Instruction for Table 1:

(1) For each testing model, the columns "#p_i" and "MLD" give the number of probability variables n and ML-degree N, respectively.

(2) In the column "Standard" we record the time to compute the polynomial generating $\mathcal{I}(\mathcal{V}(f)_\infty)$. When FGb returned no output until we run out the memory, we record "∞".

(3) We record timings of Algorithm 2 in the columns "Algorithm 2". The timing in italics font means the computation did not finish within a week, but we estimate a lower bound based on the sampling timing. For instance, in Model 7, its timing of Algorithm 7 is 27.7 minutes. To compute its nonproperness set, we need call Algorithm 7 592 times and solve the linear equations, so a lower bound of the total time is $592 \times 27.7\ m \doteq 11\ d$, where m means minutes, and d means days.

Conclusion from Table 1: Algorithm 2 indeed improves the efficiency significantly. For smaller models with ML-degree less than 5, standard method (Algorithm 1) finishes successfully within 20 seconds, while Algorithm 2 almost takes no time. For larger models with ML-degree greater than 5, only Algorithm 2 can get results.

6 SUMMARY

In this work, we propose an efficient method for computing nonproperness sets of likelihood-equation systems. From the computational results, we conjecture that the nonproperness set according to a probability variable p_i is a finite union of hyperplanes.

ACKNOWLEDGMENTS

This work was supported by the NSFC under grants 61732001 and 12001029.

REFERENCES

[1] C. Amendola, N. Bliss, I. Burke, C. R. Gibbons, M. Helmer, S. Hoşten, E. D. Nash, J. I. Rodriguez, and D. Smolkin. 2019. The maximum likelihood degree of toric varieties. *Journal of Symbolic Computation* 92, 222–242.

[2] D. S. Arnon and S. Dennis. 1988. A cluster-based cylindrical algebraic decompoisition algorithm. *J. Symb. Comput.* 5, 1, 189–212.

[3] S. Basu, R. Pollack, and M. F. Roy. 1996. On the combinatorial and algebraic complexity of quantifier elimination. *Journal of ACM* 43, 6, 1002–1045.

[4] S. Basu, R. Pollack, and M. F. Roy. 1999. Computing roadmaps of semi-algebraic sets on a variety. *Journal of the AMS* 3, 1, 55–82.

[5] S. Basu, R. Pollack, and M. F. Roy. 2006. *Algorithms in Real Algebraic Geometry.* Springer-Verlag.

[6] T. Becker and V. Weispfenning. 1998. *Groebner Bases.* Springer.

[7] C. W. Brown. 2001. Improved projection for cylindrical algebraic decomposition. *J. Symb. Comput.* 32, 5, 447–465.

[8] C. W. Brown. 2001. Simple CAD construction and its applications. *J. Symb. Comput.* 31, 5, 521–547.

[9] C. W. Brown. 2003. QEPCAD B: a program for computing with semi-algebraic sets using CADs. *ACM SIGSAM Bulletin* 37, 4, 97–108.

[10] M.-L. G. Buot, S. Hoşten, and D. Richards. 2007. Counting and locating the solutions of polynomial systems of maximum likelihood equations, II: The Behrens-Fisher problem. *Statistica Sinica* 17, 1343–1354.

[11] F. Catanese, S. Hoşten, A. Khetan, and B. Sturmfels. 2006. The maximum likelihood degree. *Amer. J. Math.* 128, 3, 671–697.

[12] C. Chen, J. H. Davemport, J. P. May, M. M. Maza, B. Xia, and R. Xiao. 2010. Triangular decomposition of semi-algebraic systems.. In *ISSAC'10* . ACM New York, 187–194.

[13] G. E. Collins. 1975. Quantifier Elimination for the Elementary Theory of Real Closed Fields by Cylindrical Algebraic Decomposition. In *Lecture Notes In Computer Science*, Vol. 33. Springer-Verlag, Berlin, 134–183.

[14] G. E. Collins and H. Hong. 1991. Cylindrical algebraic decomposition for quntifier elimination. *J. Symb. Comput.* 12, 3, 299–328.

[15] D. A. Cox, J. Little, and D. Oshea. 2015. *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra.* Springer.

[16] A. Dickenstein, M. P. Millan, A. Shiu, and X. Tang. 2019. Multistationarity in structured reaction networks. *Bulletin of mathematical biology* 81, 5, 1527–1581.

[17] A. Dolzmann, T. Sturm, and V. Weispfenning. 1998. A new approach for automatic theorem proving in real geometry. *Journal of Automated Reasoning* 21, 3, 357–380.

[18] M. Drton, B. Sturmfels, and S. Sullivant. 2009. *Lectures on algebraic statistics.* Springer.

[19] J. C. Faugère. 2010. FGb: A Library for Computing Gröbner Bases. In *Mathematical Software - ICMS 2010, Lecture Notes in Computer Science*, Vol. 6327. Springer Berlin / Heidelberg, 84–87.

[20] D. Grigoriev. 1988. Complexity of deciding Tarski algebra. *J. Symb. Comput.* 5, 1–2, 65–108.

[21] E. Gross, M. Drton, and S. Petrović. 2012. Maximum likelihood degree of variance component models. *Electronic Journal of Statistics* 6, 993–1016.

[22] E. Gross and J. I. Rodriguez. 2014. Maximum likelihood geometry in the presence of data zeros. In *ISSAC'14* . ACM New York, 232–239.

[23] J. Hauenstein, J. I. Rodriguez, and B. Sturmfels. 2014. Maximum likelihood for matrices with rank constraints. *Journal of Algebraic Statistics* 5, 1, 18–38.

[24] S. Hoşten, A. Khetan, and B. Sturmfels. 2005. Solving the likelihood equations. *Foundations of Computational Mathematics* 5, 4, 389–407.

[25] H. Hong and M. Safey EI Din. 2012. Variant Quantifier Elimination. *J. Symb. Comput.* 47, 7, 883–901.

[26] Z. Jelonek. 1999. Testing sets for properness of polynoimal mappings. *Math. Ann.* 315, 1–35.

[27] D. Kapur, Y. Sun, and D. K. Wang. 2010. A new algorithm for computing comprehensive Gröbner systems. In *ISSAC'10* . 29–36.

[28] D. Lazard and F. Rouillier. 2005. Solving parametric polynomial systems. *J. Symb. Comput.* 42, 6, 636–667.

[29] S. McCallum. 1988. An improved projection operation for cylindrical algebraic decomposition of three-dimensional space. *J. Symb. Comput.* 5, 1, 141–161.

[30] S. McCallum. 1999. On projection in CAD-Based quantifier elimination with equational constrants. In *ISSAC'99.* ACM, 145–149.

[31] J. Renegar. 1992. On the computational comlexity and geometry of the first-order theory of the reals. *J. Symb. Comput.* 13, 3, 255–352.

[32] J. I. Rodriguez and X. Tang. 2015. Data-Discriminants of Likelihood Equations. In *ISSAC'15* . ACM New York, 307–314.

[33] J. I. Rodriguez and X. Tang. 2017. A Probabilistic Algorithm for Computing Data-Discriminants of Likelihood Equations. *J. Symb. Comput.* 83, 342–364.

[34] M. Safey EI Din and E. Schost. 2003. Polar varieties and computation of one point in each connected component of a smooth real algebraic set. In *ISSAC'03* . 224–231.

[35] M. Safey EI Din and E. Schost. 2004. Properness defects of projections and computaion of in each connected component of a real algebraic set. *Discrete and Computational Geometry* 32, 3, 417–430.

[36] X. Tang, T. Wolff and R. Zhao. 2019. A New Method for Computing Elimination Ideals of Likelihood Equations. In *ISSAC'19* . 339–346.

[37] A. Tarski. 1951. *A decision method for elementary algebra and geometry.* University of California Press. University of California Press. University of California Press.

[38] V. Weispfenning. 1992. Comprehensive gröbner bases. *Journal of Symbolic Computation* 14, 1, 1–29.

[39] L. Yang and B. Xia. 2005. Real solution classifications of a class of parametric semi-algebraic systems. In *Proceedings of the A3L 2005 on Algorithmic Algebra and Logic.* 281–289.