Exact Algorithms for Polynomial Optimisation

Andrew Ferguson

THÈSE DE DOCTORAT DE
SORBONNE UNIVERSITÉ

Spécialité

Informatique

École Doctorale Informatique, Télécommunications et Électronique (Paris)

Présentée par

Andrew Ferguson

Pour obtenir le grade de

DOCTEUR de SORBONNE UNIVERSITÉ

Exact algorithms for polynomial optimisation

Thèse dirigée par Mohab Safey El Din et Jérémy Berthomieu

soutenue le lundi 24 octobre 2022

après avis des **rapporteurs** :

M. Victor Magron	Researcher, CNRS
M. Cordian Riener	Professor, University of Tromsø

devant le **jury** composé de :

M. Jérémy Berthomieu	Assistant Professor, Sorbonne Université
M. Alin Bostan	Director of Research, INRIA Saclay Île-de-France
M. Bruno Escoffier	Professor, Sorbonne Université, président du jury
M. Hamza Fawzi	Assistant Professor, University of Cambridge
M. Victor Magron	Researcher, CNRS
Mme Fatemeh Mohammadi	Professor, KU Leuven
M. Cordian Riener	Professor, University of Tromsø
M. Mohab Safey El Din	Professor, Sorbonne Université

Acknowledgements

The writing of this thesis and the work contained within it was made possible thanks to the help and support of many friends and colleagues.

Firstly, I would like to thank my advisors Mohab Safey El Din and Jérémy Berthomieu for their supervision and the helpful advice they have given me over the last three years. Through their guidance, I have learned the fundamental skills necessary for a successful academic career.

Secondly, I greatly appreciate the time and effort that Cordian Riener and Victor Magron spent reading and reviewing my thesis. In particular, I am grateful for their kind words and the many useful questions they had for me during the defence. Thanks also to Alin Bostan, Bruno Escoffier, Hamza Fawzi and Fatemeh Mohammadi for being part of the jury of my defence. Thank you as well to Stef Graillat and Bruno Salvy for being my mid-term evaluation committee and ensuring I was properly prepared for the defence a year later.

I would like to thank Fatemeh again for introducing me to research and the beautiful interplay of combinatorial commutative algebra. Additionally, thank you to Cordian for introducing me to the POEMA project and for your personal help prior to starting work on my PhD.

I am very glad to have had the opportunity to work with Alin Bostan, Huu Phuoc Le, Giorgio Ottaviani and Ettore Turatti on some very interesting problems. From these collaborations, I learned much more about the theoretical side of my work, from combinatorics to algebraic geometry.

I would also like to thank the members of the PolSys team for their support, scientific and otherwise. Thank you to Dimitri, Georgy, Hadrien, Hieu, Jorge, Olive, Rafael, Rémi, Sriram and Vincent for your friendship and for readily answering the many questions I asked over these years. Thank you also to the members of POEMA for your companionship at the many events, both online as well as in person.

Thank you to all my teachers who inspired in me a love of mathematics. Thank you to Carl Savage, Mr. West and Mr. Neil for pushing me to do my best in the early stages of my development. Thank you to Alexandra Kjuchukova and Jeremy Rickard for guiding me towards the beauty of pure mathematics.

Thank you to my friends Jack and Josh for all our time together full of laughs and fun games, I couldn't have built my confidence without your help. Thank you to Katy, Clara and Elle for never failing to make me laugh.

I am eternally grateful to my mum and to my sister for your dedicated love, care and support and endless Christmas time joy. Thank you to all my family for your support in me becoming who I am today.

Finally, thank you to Eshi, my partner. Thank you for always being by my side these last three years. Thank you for all you help reading over what I write and spending long hours listening to me ramble on about things you have no interest in understanding. I am certain that I couldn't have done this without you.

Contents

Chapter 1

Introduction en français

1.1 Motivation principale

Soit $f \in \mathbb{R}[x_1, \ldots, x_n]$ un polynôme de degré d. Nous désignerons $x_1, \ldots, x_n$ par $\boldsymbol{x}$. Nous considérons la classe suivante de problèmes d'optimisation polynomiale (POP). Notre objectif est de calculer l'infimum d'un polynôme f restreint à un ensemble semi-algébrique défini par des polynômes $g_1, \ldots, g_m \in \mathbb{R}[\boldsymbol{x}]$ de degrés $d_1, \ldots, d_m$ respectivement,

$$\mathfrak{S} := \{\boldsymbol{x} \in \mathbb{R}^n \mid g_1(\boldsymbol{x}) \geq 0, \ldots g_m(\boldsymbol{x}) \geq 0\}.$$

Ceci est formulé dans le problème d'optimisation suivant :

$$\begin{aligned} f^* &:= \inf_{\boldsymbol{x} \in \mathfrak{S}} f(\boldsymbol{x}) \\ &= \sup_{\lambda \in \mathbb{R}} \lambda \quad \text{s.t.} \quad f - \lambda \geq 0 \quad \text{over } \mathfrak{S}. \end{aligned}$$

La résolution des POP est d'une importance capitale dans de nombreux domaines de l'ingénierie et des statistiques (notamment la théorie du contrôle [50, 55], la vision par ordinateur [1, 88] et la conception optimale [25], etc.) En outre, l'optimisation polynomiale apparaît dans de nombreuses applications pratiques. Par exemple, dans les problèmes de flux de puissance optimale, soit en optimisation, où l'on optimise la puissance à travers un réseau, soit pour une simulation [40, 61]. Il est important de trouver l'optimum global exact pour les petits systèmes, par opposition aux solutions approximatives, quel que soit le temps de calcul nécessaire, car ces petits systèmes peuvent être intégrés dans des problèmes plus importants. La robotique est un autre domaine d'application important qui est actuellement très actif. Par exemple, voir [106] où les méthodes symboliques sont utilisées comme prétraitement pour les techniques numériques afin d'obtenir une solution globalement optimale pour le célèbre problème de cinématique inverse pour une certaine série de robots.

Nous allons décrire brièvement plusieurs méthodes qui ont été développées pour résoudre le POP. Tout d'abord, nous mentionnons les sommes de carrés (SOS) et les relaxations de moments au POP qui aboutissent à des solutions approximatives qui convergent vers le véritable infimum avant de souligner les pièges potentiels de ces techniques non exactes. Ensuite, nous présentons le cadre algorithmique de l'optimisation polynomiale exacte à l'aide de la méthode des *valeurs critiques généralisées* où se trouve notre contribution.

1.1.1 Les relaxations SOS et l'approche des moments

Un polynôme est SOS s'il peut être exprimé comme une somme de carrés dans $\mathbb{R}[\boldsymbol{x}]$. En restreignant la région réalisable, on obtient un nouveau problème d'optimisation, plus facile à

résoudre, dont la solution est une borne inférieure à la solution originale. Le *module quadratique* $\mathscr{M}(g)$ est défini par

$$\mathscr{M}(g) := \{s_0 g_0 + s_1 g_1 + \cdots + s_m g_m \mid g_0 = 1, s_i \text{ est SOS pour } 0 \leq i \leq m\}.$$

Ensuite, définissons pour $t \in \mathbb{N}$ le *module quadratique tronqué* :

$$\mathscr{M}(g)_{2t} := \{s_0 g_0 + s_1 g_1 + \cdots + s_m g_m \mid g_0 = 1, s_i \text{ est SOS et } \deg(s_i g_i) \leq 2t \text{ pour } 0 \leq i \leq m\}.$$

On peut alors définir la relaxation SOS de notre POP :

$$f^*_{\text{sos},t} := \sup_{\lambda \in \mathbb{R}} \lambda \quad \text{s.t.} \quad f - \lambda \in \mathscr{M}(g)_{2t}.$$

Puisque $\mathscr{M}(g)_{2t} \subset \mathscr{M}(g)_{2(t+1)} \subset \mathscr{M}(g)$, il est facile de voir que $f^*_{\text{sos},t} \leq f^*_{\text{sos},t+1} \leq f^*$. Cependant, dans [67], Lasserre a montré que si $\mathfrak{S}$ est *compact* et *archimédien*, alors $\lim_{t\to\infty} f^*_{\text{sos},t} = f^*$. De plus, on peut regarder le problème dual des moments concernant les fonctionnelles de $\mathbb{R}[\boldsymbol{x}]$ dans $\mathbb{R}$. De la même manière que précédemment, on peut relâcher le problème en utilisant le module quadratique tronqué. On arrive à une suite de programmes semi-définis (SDP), connue sous le nom de hiérarchie de Lasserre, pouvant maintenant être résolus à l'aide de solveurs SDP. Nous renvoyons vers les logiciels GloptiPoly [51], SOSTOOLS [85], SparsePOP [108], TSSOS [109] ou YALMIP [72]. Notons que SparsePOP et TSSOS résolvent des variantes creuses de POP. De plus, nous mentionnons RAGLib [94] et RealCertify [75] qui peuvent être utilisés pour obtenir des solutions certifiées. Ce cadre de dualité SOS/moments s'est avéré très efficace pour traiter un large éventail de POP, y compris ceux issus de la pratique. Cependant, comme les problèmes d'optimisation sont résolus numériquement, des problèmes peuvent survenir et entraîner des solutions inexactes. De plus, certains problèmes ne sont pas adaptés à cette approche, notamment ceux qui ne satisfont pas l'hypothèse de compacité. C'est pourquoi nous présentons maintenant un cadre alternatif qui permet de résoudre un plus grand nombre de problèmes, tout en fournissant une représentation exacte de la solution.

1.1.2 Valeurs critiques généralisées

Dans le cadre d'un ensemble semi-algébrique compact $\mathfrak{S}$, les extrema d'une application polynomiale sont contenus dans l'ensemble des valeurs critiques de l'application, noté $K_0(f)$. Cependant, si l'on considère une application polynomiale restreinte à un ensemble algébrique non compact, comme c'est le cas pour POP non contraint, les valeurs critiques de l'application peuvent ne pas suffire.

Dans [89], Rabier introduit l'ensemble des *valeurs critiques asymptotiques*, noté $K_\infty(f)$, dont l'union avec les valeurs critiques est appelée l'ensemble des *valeurs critiques généralisées*, noté $K(f)$. Ces valeurs fournissent une généralisation du théorème de la fibration d'Ehresmann à des paramètres non propres. Soit $\boldsymbol{f}$ une application polynomiale définie par

$$\boldsymbol{f} : \boldsymbol{x} \in X \mapsto (f_1(\boldsymbol{x}), \ldots, f_p(\boldsymbol{x})) \in \mathbb{K}^p,$$

où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et X est une variété lisse définie par une suite régulière réduite $\boldsymbol{g} = (g_1, \ldots, g_m)$. Ainsi, la restriction de $\boldsymbol{f}$ à $X \setminus \boldsymbol{f}^{-1}(K(\boldsymbol{f}))$ est une fibration localement triviale. Cela signifie que pour tout ensemble ouvert connexe $U \subset X \setminus K(\boldsymbol{f})$, pour tout $y \in U$ il existe un difféomorphisme φ tel que le diagramme suivant commute

$$\begin{array}{ccc} \boldsymbol{f}^{-1}(y) \times U & \xrightarrow{\varphi} & \boldsymbol{f}^{-1}(U) \\ & \searrow^{\pi} & \downarrow \boldsymbol{f} \\ & & U \end{array}$$

où π est la projection sur U [58, Théorème 3.1].

On désigne le jacobien de $\boldsymbol{f}$ et $\boldsymbol{g}$ par $\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})$, l'ensemble des valeurs critiques d'une application polynomiale est défini de la manière habituelle :

$$K_0(\boldsymbol{f}) = \{c \in \mathbb{C}^p \mid \exists \boldsymbol{x} \in X \text{ s.t. } \boldsymbol{f}(\boldsymbol{x}) = c \text{ and } \operatorname{rank}(\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})(\boldsymbol{x})) < m + p\}\,.$$

Ensuite, l'ensemble des *valeurs critiques asymptotiques* de l'application $\boldsymbol{f}$ est défini comme étant l'ensemble :

$$K_\infty(\boldsymbol{f}) = \{c \in \mathbb{C}^p \mid \exists (\boldsymbol{x}_t)_{t\in\mathbb{N}} \subset X \text{ s.t. } \|\boldsymbol{x}_t\| \to \infty, \boldsymbol{f}(\boldsymbol{x}_t) \to c \text{ and } \|\boldsymbol{x}_t\|\nu(\mathrm{d}\boldsymbol{f}(\boldsymbol{x}_t)) \to 0\}\,,$$

où $\mathrm{d}\boldsymbol{f}$ est la différentielle de l'application $\boldsymbol{f}$ et ν est la distance à l'ensemble des opérateurs singuliers. Défini de cette manière, Kurdyka, Orro et Simon montrent dans [62] que $K_\infty(\boldsymbol{f})$ satisfait un théorème de Sard généralisé. Cela signifie que la codimension de $K_\infty(\boldsymbol{f})$ est supérieure ou égale à un. Combiné à la propriété de fibration des valeurs critiques généralisées, le calcul de $K(\boldsymbol{f})$ permet de résoudre de nombreux problèmes de géométrie algébrique réelle, comme le calcul de points échantillons pour chaque composante connexe d'un ensemble semi-algèbrique défini par une inégalité unique, et dans le cas où $p = 1$, la résolution de POP. Pour un polynôme $f \in \mathbb{Q}[\boldsymbol{x}]$ et son infimum f^*, il existe trois cas :

- L'infimum f^* est atteint. Alors, f^* est une valeur critique de f ;
- f^* n'est atteint qu'à l'infini, ce qui signifie qu'il n'existe pas de minimiseur $\boldsymbol{x} \in \mathbb{R}^n$ mais un chemin $\boldsymbol{x}_t \in \mathbb{R}^n$ qui s'approche de l'infimum quand $\|\boldsymbol{x}_t\| \to \infty$. Alors, f^* est une valeur critique asymptotique de f ;
- $f^* = -\infty$.

Nous donnons un exemple d'un polynôme dont l'infimum tombe dans le second de ces trois cas.

Example 1.1. *Considérons le polynôme $f = x^2 + (xy-1)^2 \in \mathbb{R}[x, y]$ et soit f^* son infimum sur $\mathbb{R}^2$. Premièrement, puisque le gradient de f est égal à $(2x + 2y(xy-1), 2x(xy-1))$, nous voyons qu'il existe exactement un point critique $(x, y) = (0, 0)$. Ainsi, la seule valeur critique de f est 1. Cependant, remarquons que si l'on prend un chemin $\gamma(t) = (t, 1/t)$, alors quand $t \to 0$ nous avons $\|\gamma(t)\| \to \infty$, $f(\gamma(t)) \to 0$. Donc, $f^* \leq 0$. Puisque f est une somme de carrés, nous savons que $f^* \geq 0$ et donc $f^* = 0$. Par conséquent, 0 est une valeur critique asymptotique de f.*

Ainsi, comme première étape d'une stratégie d'optimisation polynomiale exacte, on calcule les représentations exactes de toutes les valeurs critiques généralisées de f. On peut le faire en calculant un polynôme dont les racines contiennent ces valeurs, puis en utilisant un algorithme d'isolation des racines réelles, par exemple celui décrit dans [91], pour calculer des intervalles d'isolation avec des extrémités rationnelles pour toutes les racines réelles.

On commence par les valeurs critiques, soit I l'idéal défini par les polynômes $f - c$, $\boldsymbol{g}$ et les mineurs maximaux de la jacobienne de f et $\boldsymbol{g}$, où $\boldsymbol{g}$ est une suite régulière réduite définissant un ensemble algébrique lisse $\mathbf{V}(\boldsymbol{g})$. Par le critère jacobien [27, Corollaire 16.20], pour obtenir une représentation polynomiale des valeurs critiques de f limité à $\mathbf{V}(\boldsymbol{g})$, on peut calculer une résolution géométrique de I ce qui donne une représentation triangulaire analogue à l'élimination gaussienne dans le cadre linéaire. Ensuite, on doit calculer l'ensemble des valeurs critiques asymptotiques de f restreintes à $\mathbf{V}(\boldsymbol{g})$ en utilisant, par exemple, l'algorithme présenté dans l'article [58].

Avec $C = \{c_1, \ldots, c_k\} \subset \mathbb{R}$ l'ensemble fini des valeurs critiques généralisées, on peut calculer des intervalles isolants avec des extrémités rationnelles pour chaque point dans C et donc on peut choisir des nombres rationnels $r_1, \ldots, r_k$ de sorte que

$$r_1 < c_1 < r_2 < \cdots < r_k < c_k.$$

Enfin, on peut utiliser la propriété de fibration de $K(f)$ pour décider laquelle, le cas échéant, des valeurs critiques généralisées de f est l'infimum f^* en décidant du caractère vide des fibres de f en $r_1, \ldots, r_k$ en utilisant, par exemple, l'algorithme proposé dans [95].

Par conséquent, on peut utiliser cette méthode des valeur critiques généralisées pour résoudre des POP dans des situations non compactes, sans approximations numériques ni erreurs numériques potentielles. De plus, le résultat est une représentation exacte de l'infimum du polynôme d'entrée, sous la forme d'un polynôme et d'un intervalle isolant.

Pour le calcul des valeurs critiques généralisées, la résolution du système polynomial est essentielle. Il existe de nombreuses approches pour résoudre les systèmes polynomiaux, telles que la déformation par homotopie, les bases de Gröbner ou les résolutions géométriques. Dans cette thèse, nous nous concentrons principalement sur les bases de Gröbner et utilisons également les résolutions géométriques et nous expliquons maintenant brièvement les raisons de ceci.

1.1.3 Déformation homotopique

Soit $S_1 \in \mathbb{K}[x_1, \ldots, x_n]^p$ un système polynomial de dimension zéro, ce qui signifie que son ensemble de solutions est fini. Les méthodes d'homotopie consistent à définir une déformation entre le système S_1 que l'on veut résoudre et un second système de dimension zéro S_0 de même degré dont les solutions sont plus faciles à décrire. Pour cela, laissons t être un paramètre et définissons le système S_t par

$$S_t = (1-t)S_0 + tS_1 \in \mathbb{K}[t, x_1, \ldots, x_n]^p,$$

de sorte que lorsque $t = 0$, S_t est égal au système que nous pouvons résoudre facilement, et lorsque $t = 1$, S_t est égal au système cible. Ainsi, les solutions du système S_t définissent un chemin dans $\mathbb{K}^n$ entre les racines de S_0 et S_1. L'idée fondatrice est alors de résoudre le système S_0 et pour chaque solution de parcourir ce chemin en augmentant la valeur de t pas à pas, en calculant les solutions des systèmes intermédiaires par itération de Newton, jusqu'à atteindre les solutions correspondantes du système S_1.

Il existe de nombreux algorithmes pour la continuation homotopique, ce qui est exposé ci-dessus, qui peuvent être classés en algorithmes numériques et symboliques, pour plus d'informations voir [2, 49, 71]. Cependant, des problèmes peuvent survenir lorsqu'un chemin passe par un système *mal conditionné*. Par exemple, supposons que pour $t \in]0, 1[$ le système intermédiaire S_t ait des solutions avec multiplicité. Cela signifierait que deux solutions de S_0 convergeraient en une seule solution. En fait, en raison de la perte de précision de l'arithmétique à virgule flottante ou du chevauchement des intervalles dans l'arithmétique des intervalles, les solutions n'ont pas besoin d'être multiples pour que ce problème se produise, mais simplement d'être trop proches les unes des autres. C'est pour cette raison que les algorithmes d'itération de type Newton qui calculent les solutions des systèmes intermédiaires étape par étape ont des difficultés à proximité d'un groupe de racines.

Des problèmes similaires peuvent se produire avec des chemins de solutions qui tendent vers l'infini. Bien que le nombre de solutions reste constant dans le cadre projectif, cela reviendrait à perdre une solution du système. Par conséquent, dans cette thèse, nous nous concentrons sur les méthodes symboliques qui évitent ces problèmes de manque de précision et de perte de racines, comme les bases de Gröbner et les résolutions géométriques.

1.1.4 Bases de Gröbner

Le calcul des bases de Gröbner sera l'outil principal de cette thèse pour la résolution de systèmes polynomiaux. Étant donné un idéal de dimension zéro $I \subset \mathbb{C}[x_1, \ldots, x_n]$, nous souhaitons calculer une représentation polynomiale de l'ensemble $\mathbf{V}(I)$. Nous disons que l'idéal I est dans en position *générique* si sa base de Gröbner lexicographique (LEX) a la forme

$$\{x_1 - f_1(x_n), \ldots, x_{n-1} - f_{n-1}(x_n), f_n(x_n)\},$$

où le degré de f_n est le degré de l'idéal I. De manière générale, le calcul direct d'un base de Gröbner LEX est coûteux. Une méthode rapide couramment utilisée en pratique consiste plutôt à calculer d'abord une base de Gröbner pour l'ordre du degré lexicographique inverse (DRL) de I, en utilisant par exemple l'algorithme F_5 de Faugère [29], puis à utiliser un algorithme de changement d'ordre, tel que FGLM [31], pour récupérer une base de Gröbner LEX.

De cette façon, nous pouvons éviter tout problème de perte de précision et nous avons la garantie de récupérer toutes les solutions du système polynomial donné en entrée satisfaisant l'hypothèse de position générique. Par conséquent, cette méthode s'inscrit bien dans le cadre de l'optimisation polynomiale exacte exposée ci-dessus, car nous devons calculer un polynôme dont les racines contiennent toutes les valeurs critiques généralisées du polynôme cible. Cependant, si le cadre est clair, il reste de nombreuses questions à soulever concernant la complexité d'une telle procédure.

1.2 Exposés des problèmes

Problème 1. Une méthode populaire de résolution de systèmes polynomiaux consiste à calculer d'abord une base de Gröbner DRL du système, puis à utiliser un algorithme de changement d'ordre pour obtenir une base LEX. Alors que le calcul d'une base de Gröbner DRL pour les systèmes déterminantiels dérivant de mineurs maximaux est bien compris, une étude de l'étape de changement d'ordre pour cette classe de systèmes fait défaut. Quelle est la structure d'une base de Gröbner DRL générique dans ce cadre de déterminants de mineurs maximaux ? De plus, les estimations de complexité pour le calcul des valeurs critiques peuvent-elles être améliorées en tirant parti de cette structure déterminantielle ?

Problème 2. Les idéaux déterminantiels structurés apparaissent fréquemment dans les applications. Par exemple, la résolution de programmes semi-définis pour trouver des décompositions de sommes de carrés, une méthode populaire pour l'optimisation polynomiale nous amène à étudier des matrices de moments qui sont symétriques. Comment cette structure supplémentaire affecte-t-elle les calculs de base de Gröbner sur les idéaux déterminantiels dérivés de défauts de rang de ces matrices symétriques ?

Problème 3. Le calcul des valeurs critiques asymptotiques est le goulot d'étranglement de la méthode globale des valeurs critiques généralisées pour résoudre un POP. Existe-t-il un algorithme plus efficace pour calculer l'ensemble des valeurs critiques asymptotiques qui ramène la complexité du calcul de l'infimum d'une application polynomiale $f : X \to \mathbb{R}$, où X est un ensemble algébrique lisse défini par une suite réduite et radicale $g_1, \ldots, g_m \in \mathbb{K}[x_1, \ldots, x_n]$ et où $f, g_1, \ldots, g_m$ sont de degré d, à exactement $d^{O(n)}$ opérations dans le corps $\mathbb{K}$, où $\mathbb{K}$ est $\mathbb{C}$ ou $\mathbb{R}$?

Problème 4. On cherche à comprendre la structure algébrique sous-jacente des décompositions en sommes de carrés. Dans un premier temps, en fixant un certain nombre de variables homogènes, quel est le degré de la variété de toutes les sommes de deux carrés ? De plus, étant donné un polynôme générique qui est une somme de carrés, quelle est la structure de toutes ses décompositions possibles ?

1.3 Travaux antérieurs et contributions

1.3.1 Problèmes 1 et 2

Travaux antérieurs. Les idéaux déterminantiels constituent un domaine d'étude actif en algèbre commutative. Une technique populaire dans ce domaine consiste à utiliser la théorie des bases de Gröbner pour relier ces idéaux à des objets combinatoires, afin d'utiliser les propriétés des anneaux de Stanley-Reisner de complexes simpliciaux, voir par exemple [16, 18, 19, 21, 22, 104].

Dans [22], les auteurs ont donné une formule explicite pour la série de Hilbert des idéaux déterminantiels définis par des matrices à coefficients des variables. En spécialisant ce résultat aux matrices déterminantieelles dérivées des mineurs maximaux, les auteurs de [33] trouvent la série de Hilbert des idéaux définissant l'ensemble des valeurs/points critiques d'un polynôme restreint à un ensemble algébrique sous certaines hypothèses de régularité. En utilisant cela, les auteurs donnent une limite supérieure sur le nombre d'opérations arithmétiques nécessaires pour calculer une base de Gröbner LEX d'un tel idéal dans le cadre de DRL à LEX dans le même article [33, Theorem 3].

Tout d'abord, en se basant sur [5, Theorem 7], les auteurs de [33, Theorem 3] utilisent la série de Hilbert des idéaux déterminantiels génériques pour analyser la complexité de l'étape DRL en utilisant l'algorithme F_5 de Faugère [29]. Ici, et dans tout le texte, les estimations de complexité sont données en termes d'opérations arithmétiques dans le corps de base $\mathbb{K}$. Ensuite, pour obtenir une base de Gröbner LEX, puisque nous sommes dans le cas zéro-dimensionnel, ils utilisent l'algorithme FGLM pour effectuer le changement d'ordre [31]. La complexité de FGLM est de $O(nD^3)$, où n est le nombre de variables et D est le degré de l'idéal. Par exemple, considérons la projection ϕ de $\mathbb{K}^n$ sur la première coordonnée restreinte à un ensemble algébrique défini par une suite régulière réduite $g_1, \ldots, g_m$ où $\deg g_i = d$ pour $1 \leq i \leq m$. Dans [84, Théorème 2.2], les auteurs utilisent la formule de Thom-Porteous-Giambelli pour prouver que le degré de l'idéal définissant les points critiques de la projection ϕ est

$$D = d^p(d-1)^{n-m}\binom{n-1}{m-1}.$$

Sous certaines hypothèses de stabilité, les auteurs de [30] et de [83] améliorent l'algorithme FGLM en appliquant des techniques d'algèbre linéaire rapide. Leurs algorithmes ont une complexité $O^{\sim}(D^\omega)$ et $O(nD^\omega \log D)$ respectivement, où ω est l'exposant de la multiplication matricielle. La meilleure borne théorique connue pour ω est de $2,37286$ donnée dans [3].

D'autres algorithmes ont été introduits pour tirer parti du caractère creux de la matrice de multiplication T_{x_n}, comme l'algorithme Sparse-FGLM dans [32]. Sous les mêmes hypothèses de stabilité et lorsque l'idéal d'intérêt est en position générique, c'est-à-dire que les monômes principaux de la base LEX sont $x_1, \ldots, x_{n-1}, x_n^D$, l'algorithme Sparse-FGLM s'appuie principalement sur le caractère creux de la matrice T_{x_n} associée à l'application linéaire de multiplication par x_n dans l'algèbre quotient de dimension finie $\mathbb{K}[x_1, \ldots, x_n]/I$. Elle a une complexité de $O(qD^2 + nD\log^2 D)$, où q est le nombre de colonnes non triviales de la matrice T_{x_n}.

Dans certains cas, $q \in O(D)$ et donc l'algorithme Sparse-FGLM n'est pas toujours plus rapide que les algorithmes de [30, 83] asymptotiquement. Cependant, très récemment, sous les mêmes hypothèses de position générique et de stabilité, les auteurs de [12] ont conçu un algorithme qui améliore celui de [30, 32, 83] en se concentrant sur la structure de la matrice T_{x_n}, au lieu juste de son caractère creux, avec une complexité de $O^{\sim}(q^{\omega-1}D)$.

Afin de comparer précisément les algorithmes de [12, 32] à ceux de [30, 83], on doit d'abord estimer le paramètre q. De plus, comme q est un paramètre fondamental des algorithmes de type FGLM, borner q est utile pour tout algorithme qui s'appuie sur la matrice de multiplication T_{x_n}.

En utilisant les résultats de [79] sur la structure de l'escalier DRL des intersections complètes génériques, le nombre q est étudié dans [32] pour cette classe de systèmes. De plus, en utilisant cette structure, les auteurs de [32] ont prouvé que la matrice T_{x_n} est telle qu'elle peut être calculée sans opérations arithmétiques. Cependant, des résultats similaires étaient auparavant inconnus pour d'autres classes d'idéaux, tels que les idéaux déterminantiels génériques. Cela signifie que les améliorations de la complexité de [12, 32] n'étaient pas entièrement comprises pour de nombreux problèmes importants, par exemple le calcul de la valeur critique.

Contributions. Les résultats concernant le problème 1 sont le fruit d'un travail conjoint avec Jérémy Berthomieu, Alin Bostan et Mohab Safey El Din et ont été publiés dans le Journal of

Algebra. Nous commençons par définir précisément la classe d'idéaux que nous allons considérer pour le problème 1, ce que nous appelons les idéaux déterminantiels-sommes génériques.

Definition 1.2. *Soit $\mathbb{K}$ un corps infini, soit $I \subset \mathbb{K}[x_1, \ldots, x_n]$ un idéal qui est la somme de m polynômes de degré au plus d et des mineurs maximaux d'une matrice avec des entrées polynomiales également de degré au plus d. Nous disons que I est un idéal déterminantiel-somme générique si les trois conditions suivantes sont réunies :*

- *l'idéal I est en position générique, ce qui signifie que la base de Gröbner réduite LEX avec $x_1 \succ \cdots \succ x_n$ admet les monomes de tête $x_1, \ldots, x_{n-1}, x_n^D$ où D est le degré de I,*
- *la série de Hilbert H de $\mathbb{K}[x_1, \ldots, x_n]/I$ est égale à*

$$H = \frac{\det(M(t^{d-1}))}{t^{(d-1)\binom{m-1}{2}}} \frac{(1-t^d)^m (1-t^{d-1})^{n-m}}{(1-t)^n}$$

  où $M(t)$ est la matrice $(m-1) \times (m-1)$ dont la (i,j)-ième entrée est $\sum_k \binom{m-i}{k}\binom{n-1-j}{k} t^k$,
- *pour tout $e \geq 1$, la série de Hilbert de $(\mathbb{K}[x_1, \ldots, x_n]/I)/\langle x_n^e \rangle$ est égale à la série $(1-t)H$ tronquée au premier coefficient négatif.*

Dans le chapitre 4, nous prouvons que, sous certaines hypothèses de régularité, l'idéal définissant l'ensemble des valeurs critiques d'une application polynomiale générique restreinte à un ensemble algébrique lisse tombe dans cette classe.

Notre première contribution principale est un résultat de structure sur la base de Gröbner DRL de tels idéaux, donnant le théorème suivant.

Theorem 1.3. *Soit I un idéal déterminantiel-somme générique de sorte que les conditions de la définition ?? soient réunies. Supposons que l'on connaisse une base de Gröbner réduite et minimale de I pour l'ordre DRL. Alors la matrice de multiplication T_{x_n} peut être construite sans effectuer d'opérations arithmétiques.*

En tenant compte de cette structure, nous donnons ensuite des formules pour le nombre de colonnes non triviales de T_{x_n}, que nous notons q. Nous donnons une formule exacte dans le cas $d = 2$ alors que pour $d \geq 3$ nous donnons une formule asymptotique.

Theorem 1.4. *Soit I un idéal déterminantiel-somme générique tel que les conditions de la définition ?? soient vérifiées, et soit T_{x_n} la matrice associée à l'application linéaire de multiplication par x_n. Dénotons par q le nombre de colonnes non triviales de T_{x_n}. Alors, pour $d = 2$ et $n \gg m$,*

$$q = \sum_{k=0}^{m-1} \binom{n-m-1+k}{k} \binom{m}{\lfloor 3m/2 \rfloor - 1 - j}. \tag{1.1}$$

De plus, pour $d \geq 3$ et $n \to \infty$,

$$q \approx \frac{1}{\sqrt{(n-m)\pi}} \sqrt{\frac{6}{(d-1)^2-1}} d^m (d-1)^{n-m} \binom{n-2}{m-1}. \tag{1.2}$$

Enfin, nous utilisons ces résultats et l'algorithme Sparse-FGLM [32, Theorem 3.2] pour donner un résultat de complexité pour le changement d'ordre de DRL à LEX pour les idéaux déterminantiels-sommes génériques.

Theorem 1.5. *Soit I un idéal déterminantiel-somme générique de sorte que les conditions de Definition ?? soient vérifiées. Supposons que l'on connaisse une base de Gröbner DRL réduite et*

minimale de I. Alors, pour $d \geq 3$, la complexité arithmétique du calcul d'une base de Gröbner LEX de I est bornée supérieurement par

$$O\left(\frac{d^{3m}(d-1)^{3(n-m)}}{\sqrt{(n-m)d\pi}}\binom{n-2}{m-1}\binom{n-1}{m-1}^2\right).$$

Par conséquent, le gain de complexité de Sparse-FGLM par rapport à FGLM pour les systèmes déterminantiels-sommes génériques est approximativement de

$$O\left(\frac{q}{nD}\right) \approx O\left(\frac{\sqrt{n-m}}{n^2(d-1)}\right).$$

De plus, pour le problème 2, nous considérons la classe des idéaux déterminantiels symétriques génériques. Ce sont les idéaux déterminantiels qui sont dérivés d'une matrice symétrique avec des entrées polynomiales génériques. Plus précisément, pour la matrice symétrique $S = (s_{i,j})_{1\leq i,j\leq \ell}$ avec des entrées les variables $\boldsymbol{s} = (s_{1,1}, s_{2,1}, s_{2,2}, \ldots, s_{\ell,1}, \ldots, s_{\ell,\ell})$ et un certain $r \in \mathbb{N}$, $\mathcal{S}_r$ est l'idéal homogène engendré par tous les mineurs de taille $r+1$ de S. Pour des valeurs fixes de $n, d \in \mathbb{N}$, soit $S^{n,d} = (f_{i,j})_{1\leq i,j\leq \ell}$ une matrice symétrique de taille $\ell \times \ell$ dont les entrées sont dans $\mathbb{K}[x_1, \ldots, x_n]_{\leq d}$. Alors, $\mathcal{S}_r^{n,d}$ est l'idéal défini par les mineurs de $S^{k,d}$ de taille $r+1$. Soit $\mathcal{H}_r$ et $\mathcal{H}_r^{n,d}$ les numérateurs réduits de la série de Hilbert des idéaux $\mathcal{S}_r$ et $\mathcal{S}_r^{n,d}$ respectivement. Soit maintenant $\mathcal{S}_r^{n,d,\mathrm{h}}$ l'idéal homogénéisé de $\mathcal{S}_r^{n,d}$. Les résultats suivants, réalisés en collaboration avec Huu Phuoc Le, ont été communiqués à la conférence ISSAC 2022 sous la forme d'un article.

Conjecture 1.6. *1. Étant donné $r \in \mathbb{N}$, le numérateur réduit $\mathcal{H}_r(t)$ de la série de Hilbert de l'idéal symétrique déterminantiel $\mathcal{S}_r$ est unimodal.*

2. Pour $e \geq 1$, soit $\mathcal{Q}_r^{n,d,e}$ la série de Hilbert de $\mathbb{K}[x_0, \ldots, x_n]/\left(\mathcal{S}_r^{n,d,\mathrm{h}} + \langle x_0, x_n^e\rangle\right)$. Nous supposons que $\mathcal{Q}_r^{n,d,e} = \left[(1-t^e)\mathcal{H}_r^{n,d}(t)\right]_+$, qui est la série $(1-t^e)\mathcal{H}_r^{n,d}(t)$ tronquée à son premier coefficient négatif.

Ensuite, sous ces hypothèses de régularité, nous prouvons le résultat de structure suivant pour les idéaux déterminantiels symétriques génériques.

Theorem 1.7. *Étant donné $r, \ell, d \in \mathbb{N}$ et $k = \binom{\ell-r+1}{2}$, il existe un sous-ensemble ouvert de Zariski non vide $\mathscr{F}_r$ de $\mathbb{K}[x_1, \ldots, x_k]_{\leq d}^{\ell(\ell+1)/2}$ tel que, lorsque les entrées de $S^{n,d}$ sont prises dans $\mathscr{F}_r$, alors*

*L'idéal $\mathcal{S}_r^{n,d}$ est de dimension nulle et radical. Si la conjecture **??** est vérifiée et qu'une base de Gröbner réduite de $\mathcal{S}_r^{n,d}$ par rapport à $\prec_{\mathrm{DRL}}$ est connue, la matrice T_{x_n} de la multiplication par x_n peut être construite sans aucune opération arithmétique. De plus, le nombre de colonnes denses de T_{x_n} est égal au plus grand coefficient de la série de Hilbert $\mathcal{H}_r^{n,d}$.*

En utilisant l'algorithme Sparse-FGLM, nous fournissons ensuite un résultat de complexité dédié au changement d'ordre de DRL à LEX pour les idéaux déterminantiels symétriques.

Theorem 1.8. *Étant donné $r, \ell, d \in \mathbb{N}$ et $k = \binom{\ell-r+1}{2}$, nous considérons la matrice $S^{n,d}$ dont les entrées sont prises dans l'ensemble ouvert de Zariski $\mathscr{F}_r$ défini dans le théorème **??**. Supposons que la conjecture **??** est vérifiée et que la base de Gröbner réduite de $\mathcal{S}_r^{n,d}$ pour $\prec_{\mathrm{DRL}}$ est connue. Alors quand $d \to \infty$, l'algorithme Sparse-FGLM calcule une base de Gröbner pour $\prec_{\mathrm{LEX}}$ de $\mathcal{S}_r^{n,d}$ en*

$$O\left(q\mathcal{H}_r^{n,d}(1)^2\right) = O\left(qd^{2n}\mathcal{H}_r(1)^2\right) = O\left(qd^{2n}\left(\prod_{i=0}^{\ell-r-1}\frac{\binom{\ell+i}{2i+r}}{\binom{2i+1}{i}}\right)^2\right)$$

opérations arithmétiques dans $\mathbb{K}$ où q est le nombre de colonnes denses de la matrice de multiplication T_{x_n}. De plus, quand $d \to \infty$, q est borné supérieurement par

$$d^{n-1}\mathcal{H}_r(1) = \sqrt{\frac{6}{n\pi}}d^{n-1}\prod_{i=0}^{\ell-r-1}\frac{\binom{n+i}{2i+r}}{\binom{2i+1}{i}}.$$

Dans le chapitre 5, nous étudierons trois autres cas particuliers et obtiendrons des résultats de complexité plus fins pour chacun d'eux.

1.3.2 Problème 3

Travaux antérieurs. Le calcul de l'ensemble des valeurs critiques d'une application polynomiale $\boldsymbol{f} = (f_1, \ldots, f_p)$ restreinte à un ensemble algébrique $X = \mathbf{V}(\boldsymbol{g}) = \mathbf{V}(g_1, \ldots, g_m)$ est classique. Sous certaines hypothèses de régularité, l'ensemble algébrique défini par l'intersection de X avec la variété définie par les mineurs maximaux de $\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})$ est égal à l'ensemble des points critiques de $\boldsymbol{f}$ [27, Corollaire 16.20].

Le premier travail vers le calcul des valeurs critiques asymptotiques d'une application polynomiale, dans le cadre non restreint, a été donné dans [62]. Dans cet article, les auteurs donnent une caractérisation géométrique de $K_\infty(\boldsymbol{f})$ qui permet de construire un ensemble algébrique de codimension au moins égale à un dans $\mathbb{C}^p$ qui contient les valeurs critiques asymptotiques en utilisant des algorithmes pour effectuer des opérations théoriques d'idéaux, tels que les algorithmes basés sur les bases de Gröbner. Ensuite, dans [58], les auteurs abordent le problème du calcul des valeurs critiques généralisées d'une application polynomiale restreinte à un ensemble algébrique. L'algorithme donné dans cet article suit un cadre similaire consistant à définir des ensembles algébriques, à considérer leurs intersections avec des hyperplans linéaires et à projeter sur l'espace cible. Cependant, cet algorithme nécessite la construction de $(p(m+p))^{\binom{n}{m+p}}$ ensembles localement fermés dans $\mathbb{C}^{(n+1)\binom{n}{m+p}+p+n}$ avant de projeter chacun d'eux sur $\mathbb{C}^p$, ce qui rend l'algorithme peu pratique, surtout si on le compare au calcul des valeurs critiques de $\boldsymbol{f}$. De plus, aucun résultat expérimental ni aucune analyse de complexité n'est donné pour cet algorithme.

En revenant au cadre sans restriction, et avec l'hypothèse supplémentaire que l'application polynomiale n'a qu'une seule composante, il y a eu plusieurs tentatives d'améliorer ce modèle algorithmique. Par exemple, l'auteur de [92] fait le lien entre les valeurs critiques généralisées et les propriétés des variétés polaires. Cette connexion a ensuite été exploitée dans [59] pour construire des arcs rationnels qui atteignent toutes les valeurs critiques généralisées d'un polynôme. De plus, dans [60], les auteurs font une distinction entre les valeurs critiques asymptotiques, en détectant celles qui sont trouvées de manière non triviale, c'est-à-dire loin du lieu critique du polynôme. Cependant, une telle distinction n'est pas nécessaire pour le problème de l'optimisation polynomiale et nous n'en tenons donc pas compte dans nos contributions.

Contributions. Nous supposons que $\boldsymbol{f}$ satisfait l'hypothèse de régularité suivante (R): "La clôture de Zariski de $X \setminus \mathrm{crit}(\boldsymbol{f}, X)$ est X", où

$$\mathrm{crit}(\boldsymbol{f}, X) = \{\boldsymbol{x} \in X \mid \mathrm{rank}(\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})(\boldsymbol{x})) < m + p\}$$

est le lieu critique de $\boldsymbol{f}$ sur X. Par conséquent, l'hypothèse (R) est équivalente à exiger que pour un point générique $\boldsymbol{x} \in X$, $\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})(\boldsymbol{x})$ est de rang plein.

Sous cette hypothèse de régularité, mes co-auteurs, Jérémy Berthomieu et Mohab Safey El Din, et moi-même résolvons le problème 3 en concevant des algorithmes qui ramènent le calcul des valeurs critiques asymptotiques dans le cas $p = 1$, le cas crucial pour l'optimisation polynomiale, à une complexité de $d^{O(n)}$. Ces résultats ont été soumis au Journal of Symbolic Computation. Pour prouver cela, nous donnons d'abord un résultat de degré.

Theorem 1.9. *Soit X un ensemble algébrique lisse défini par une suite régulière réduite $\boldsymbol{g} = (g_1, \ldots, g_m)$. Soit $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ une application polynomiale de X dans $\mathbb{K}^p$ satisfaisant l'hypothèse (R). Soit $d = \max(\deg f_1, \ldots, \deg f_p, \deg g_1, \ldots, \deg g_m)$. Alors, les valeurs critiques asymptotiques de $\boldsymbol{f}$ sont contenues dans une hypersurface de degré au plus*

$$p d^{n-p-1} \sum_{i=0}^{p+1} \binom{n+p-1}{m+2p-i} d^i.$$

Bien que nos algorithmes permettent le calcul des applications polynomiales, nous donnons un résultat de complexité dédié dans le cas particulier $p = 1$ qui est le cas d'intérêt pour l'optimisation polynomiale.

Theorem 1.10. *Soit $\boldsymbol{g} = (g_1, \ldots, g_m)$ une suite régulière réduite définissant un ensemble algébrique lisse X. Soit $f \in \mathbb{K}[\mathbf{z}]$ une application polynomials de X dans $\mathbb{K}$ satisfaisant l'hypothèse (R). Soient $d = \max(\deg f, \deg g_1, \ldots, \deg g_m)$ et $D = d^{n-2} \sum_{i=0}^{2} \binom{n}{m+2-i} d^i$. Alors, il existe un algorithme qui, pour une entrée $f, \boldsymbol{g}$, produit un polynôme non nul $H \in \mathbb{K}[c]$ tel que $K_\infty(\boldsymbol{f}) \subset \mathbf{V}(H)$ en utilisant au plus*

$$O^{\sim}(n^2 d^{n+2} D^5)$$

opérations arithmétiques dans $\mathbb{K}$.

Cependant, dans le cadre plus général des applications polynomiales, nous obtenons la complexité suivante.

Theorem 1.11. *Soit $\boldsymbol{g} = (g_1, \ldots, g_m)$ une suite régulière réduite définissant un ensemble algébrique lisse X. Soient $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ une application polynomiale de X dans $\mathbb{K}^p$ satisfaisant l'hypothèse (R). Soit $d = \max(\deg f_1, \ldots, \deg f_p, \deg g_1, \ldots, \deg g_m)$ et $D = d^{n-p-1} \sum_{i=0}^{p+1} \binom{n+p-1}{m+2p-i} d^i$. Alors, il existe un algorithme qui, sur des entrées de $\boldsymbol{f}$ et $\boldsymbol{g}$, produit des listes finies de p listes finies de polynômes non nuls $G_i \subset \mathbb{K}[\mathbf{c}]$ tels que $K_\infty(\boldsymbol{f}) \subset (\mathbf{V}(G_1) \cup \cdots \cup \mathbf{V}(G_p)) \subsetneq \mathbb{C}^p$ en utilisant au plus*

$$O^{\sim}\left(p^2 D^{p+5} + n^2 d^{n+2} D^{p+4}\right)$$

opérations arithmétiques dans $\mathbb{K}$.

1.3.3 Problème 4

Travaux antérieurs. L'étude algébrique des décompositions de polynômes homogènes a une longue histoire. Depuis les travaux classiques de Sylvester [105], l'étude des décompositions de polynômes homogènes par puissances de formes linéaires, est toujours un domaine de recherche actif. Dans [38] il a été prouvé que tout polynôme homogène général de degré $2d$ en $n+1$ variables est une somme d'au plus 2^n carrés. Pour n fixe, cette borne est atteinte pour tout d suffisamment grand.

Le nombre minimal de carrés requis dans la décomposition d'un polynôme est connu sous le nom de rang SOS. Les auteurs de [73] étudient ce rang pour des polynômes génériques en deux variables. Puis, dans [37], les auteurs donnent une conjecture sur le rang SOS générique des polynômes, en termes de nombre de variables et de degré.

Contributions. Nous considérons les décompositions SOS de polynômes de degré $2d$. Soit V un espace vectoriel complexe de dimension $n+1$. Alors, l'espace des polynômes homogènes de degré $2d$ en $n+1$ variables sera dénoté par $\mathrm{Sym}^{2d} V$.

Definition 1.12. *Soit $f \in \mathrm{Sym}^{2d} V$. Le polynôme f a un rang SOS de k si k est le nombre minimal pour qu'il existe $f_i \in \mathrm{Sym}^{d} V$ tel que*

$$f = \sum_{i=1}^{k} f_i^2.$$

Pour répondre à ce problème, nous définissons et étudions deux variétés liées aux décompositions SOS exactes. La première est définie par tous les polynômes de rang SOS inférieur ou égal à k.

Definition 1.13. *Soit* SOS_k *la sous-variété dans* $\mathrm{Sym}^{2d} V$ *obtenue à partir de la clôture de Zariski de l'ensemble de tous les polynômes de rang SOS* k*.*

$$\mathrm{SOS}_k = \overline{\{f_1^2 + \cdots + f_k^2 \mid f_i \text{ dans } \mathrm{Sym}^d V\}}.$$

Le rang SOS générique est le plus petit nombre k *tel que* SOS_k *couvre l'espace ambiant.*

Un autre objet qui peut être étudié est l'ensemble de toutes les différentes décompositions d'un polynôme f générique dans SOS_k pour $k \in \mathbb{N}$ fixe.

Definition 1.14. *Soit* $f \in \mathrm{SOS}_k$ *un polynôme générique. Nous définissons la variété de toutes les décompositions rang SOS* k *de* f *comme*

$$\mathrm{SOS}_k(f) = \left\{ (f_1, \ldots, f_k) \in \prod_{i=1}^{k} \mathrm{Sym}^d V \,\middle|\, \sum_{i=1}^{k} f_i^2 = f \right\}.$$

Pour le premier objet d'intérêt : $\mathrm{SOS}_k(f)$, nous donnons sa structure exacte dans le cas $k = 2$.

Theorem 1.15. *Soit* $f \in \mathrm{SOS}_2$ *un polynôme générique de rang SOS deux. Alors,* $\mathrm{SOS}_2(f)$ *a deux composantes irréductibles isomorphes à* $\mathrm{SO}(2)$*. Par conséquent,* $\mathrm{SOS}_2(f)$ *est isomorphe à* $\mathrm{O}(2)$*.*

Notez que le degré du polynôme f n'est pas important ici. De manière plus générale, nous calculons la dimension de cet objet.

Theorem 1.16. *Soit* $f \in \mathrm{SOS}_k$ *générique avec* $k \leq n$*. Alors ,*

$$\dim \mathrm{SOS}_k(f) = \binom{k}{2}.$$

De plus, nous conjecturons que $\mathrm{SOS}_k(f)$ est également isomorphe à $\mathrm{O}(k)$ et nous donnons quelques comptes de dimension et expériences pour soutenir cette conjecture. Par ailleurs, nous donnons le degré de SOS_1 et de SOS_2.

Theorem 1.17. *Soit* $N = \dim \mathrm{Sym}^d V = \binom{n+d}{d}$*. Les degrés des variétés des carrés et de la somme de deux carrés dans* $\mathbb{P}(\mathrm{Sym}^{2d} V)$ *sont donnés par*

$$\deg(\mathrm{SOS}_1) = 2^{N-1}, \ \deg(\mathrm{SOS}_2) = \prod_{i=0}^{N-3} \frac{\binom{N+i}{N-2-i}}{\binom{2i+1}{i}}.$$

Mes collaborateurs, Giorgio Ottaviani, Mohab Safey El Din et Ettore Turatti, et moi-même avons soumis ces résultats au Journal of Pure and Applied Algebra.

1.4 Structure de la thèse

Nous commençons par les préliminaires nécessaires à la présentation de nos résultats dans le chapitre 3. Les sections 3.1 et 3.2 donnent les définitions et propositions de base en algèbre et en géométrie algébrique qui seront fréquemment utilisées tout au long de cette thèse. La section 3.3 est consacrée aux bases de Gröbner, aux définitions de base et, en particulier, à l'introduction d'algorithmes de changement d'ordre qui sera central dans l'étude des problèmes 1 et 2. Enfin, la section 3.4 donnera quelques informations sur les décompositions de sommes de carrés et l'optimisation polynomiale exacte.

Nous présentons ensuite nos contributions qui sont organisées dans les quatre chapitres suivants.

- Le chapitre 4 détaille nos contributions au problème 1, la dérivation de formules asymptotiques pour un paramètre fondamental dans le changement d'ordre des bases de Gröbner pour les systèmes de valeurs critiques, améliorant les estimations de complexité connues précédemment.

  Ce travail a été publié sous forme d'article dans le Journal of Algebra : "Gröbner bases and critical values: The asymptotic combinatorics of determinantal systems" (Jérémy Berthomieu, Alin Bostan, Andrew Ferguson et Mohab Safey El Din) [10].

- Le chapitre 5 détaille nos contributions au problème 2, en analysant l'étape de changement d'ordre pour des classes plus générales de systèmes déterminantiels et en donnant les premières estimations de complexité dédiées dans ces cas.

  Le contenu de ce chapitre a été présenté comme un article de conférence publié dans les actes d'ISSAC 2022, Lille, France : "Finer Complexity Estimates for the Change of Ordering of Gröbner Bases for Generic Symmetric Determinantal Ideals" (Andrew Ferguson et Huu Phuoc Le) [35].

- Le chapitre 6 détaille nos contributions au problème 3, en introduisant de nouveaux algorithmes efficaces pour calculer l'ensemble des valeurs critiques asymptotiques des applications polynomiales à partir d'ensembles algébriques satisfaisant une hypothèse de régularité de base.

  Ces algorithmes et les résultats qui en découlent forment un article qui a été soumis au Journal of Symbolic computation : "Computing the set of asymptotic critical values of polynomial mappings from smooth algebraic sets" (Jérémy Berthomieu, Andrew Ferguson et Mohab Safey El Din).

- Le chapitre 7 détaille nos contributions au problème 4, en donnant le degré de la variété des sommes de deux carrés et en travaillant à la compréhension de la structure algébrique de toutes les décompositions possibles en sommes de carrés d'une somme de carrés générique.

  Ce travail a été soumis au Journal of Pure and Applied Algebra sous la forme d'un article : "On the degree of varieties of sum of squares" (Andrew Ferguson, Giorgio Ottaviani, Mohab Safey El Din et Ettore Turatti).

Enfin, nous concluons au chapitre 8 en résumant nos résultats et en analysant les prochaines étapes de chacun des quatre problèmes identifiés dans la section 1.2.

Chapter 2

Introduction

2.1 Main motivation

Let $f \in \mathbb{R}[x_1, \ldots, x_n]$ be a polynomial of degree d. We shall denote $x_1, \ldots, x_n$ by $\boldsymbol{x}$. We consider the following class of polynomial optimisation problems (POP). We aim to compute the infimum of a polynomial f restricted to a closed semi-algebraic set defined by polynomials $g_1, \ldots, g_m \in \mathbb{R}[\boldsymbol{x}]$ of degrees $d_1, \ldots, d_m$ respectively,

$$\mathfrak{S} := \{\boldsymbol{x} \in \mathbb{R}^n \mid g_1(\boldsymbol{x}) \geq 0, \ldots, g_m(\boldsymbol{x}) \geq 0\}.$$

This is formulated in the following optimisation problem:

$$\begin{aligned} f^* &:= \inf_{\boldsymbol{x} \in \mathfrak{S}} f(\boldsymbol{x}) \\ &= \sup_{\lambda \in \mathbb{R}} \lambda \quad \text{s.t.} \quad f - \lambda \geq 0 \quad \text{over } \mathfrak{S}. \end{aligned}$$

Solving POP is of principal importance in many areas of engineering and statistics, including control theory [50, 55], computer vision [1, 88] and optimal design [25] among others. Additionally, polynomial optimisation problems appear in many practical applications. For instance, in optimal power flow problems either in optimisation, where one optimises the power across a network, or for a simulation [40, 61]. Finding the exact global optimal for small systems, as opposed to approximate solutions, no matter how long it takes to compute is important as such small systems can be embedded in larger problems. Another important application domain that is currently very active is the analysis and design of robots. For example, see [106] where symbolic methods are used as pre-processing for numerical techniques to obtain a globally optimal solution for the famous Inverse Kinematics problem for a certain series of robots.

We shall briefly describe several methods that have been developed to solve POP. Firstly, we mention sums of squares (SOS) and moment relaxations to POP that result in approximate solutions that converge to the true infimum before highlighting the potential pitfalls of such non-exact techniques. Then, we present the algorithmic framework for exact polynomial optimisation using the *generalised critical value* method wherein lies our contributions.

2.1.1 SOS relaxations and the moment approach

A polynomial is SOS if it can be expressed as a sum of squares in $\mathbb{R}[\boldsymbol{x}]$. By restricting the feasible region, one obtains a new, easier to solve, optimisation problem whose solution is a lower bound to the original solution. The *quadratic module* $\mathscr{M}(g)$ is defined by

$$\mathscr{M}(g) := \{s_0 g_0 + s_1 g_1 + \cdots + s_m g_m \mid g_0 = 1, s_i \text{ is SOS for } 0 \leq i \leq m\}.$$

Then, define for $t \in \mathbb{N}$ the *truncated quadratic module* by

$$\mathscr{M}(g)_{2t} := \{s_0 g_0 + s_1 g_1 + \cdots + s_m g_m \mid g_0 = 1, s_i \text{ is SOS and } \deg(s_i g_i) \leq 2t \text{ for } 0 \leq i \leq m\}.$$

One can define the SOS relaxation of our POP:

$$f^*_{sos,t} := \sup_{\lambda \in \mathbb{R}} \lambda \quad \text{s.t.} \quad f - \lambda \in \mathscr{M}(g)_{2t}.$$

Since $\mathscr{M}(g)_{2t} \subset \mathscr{M}(g)_{2(t+1)} \subset \mathscr{M}(g)$, it is easy to see that $f^*_{sos,t} \leq f^*_{sos,t+1} \leq f^*$. However, in [67], Lasserre showed that if $\mathfrak{S}$ is *compact* and *Archimedean*, then $\lim_{t\to\infty} f^*_{sos,t} = f^*$. Moreover, one can look at the dual moment problem concerning functionals from $\mathbb{R}[\boldsymbol{x}]$ to $\mathbb{R}$. In the same way as before, we can relax the problem using the truncated quadratic module. One arrives at a sequence of semi-definite programs (SDP), known as Lasserre's hierarchy, that can now be solved using SDP solvers. For example, see the software packages GloptiPoly [51], SOSTOOLS [85], SparsePOP [108], TSSOS [109] or YALMIP [72]. Note that SparsePOP and TSSOS solve sparse variants of POP. Furthermore, we mention RAGlib [94] and RealCertify [75] that can be used to obtain certified solutions. This SOS/moment duality framework has been very successful in tackling a wide range of POPs, including those coming from practice. However, as the optimisation problems are solved numerically, problems can arise that result in inaccurate solutions. Moreover, certain problems are not suited to this approach, including those that do not satisfy the compactness assumption. Thus, we now present an alternative framework that can tackle a wider range of problems, as well as returning an exact representation of the solution.

2.1.2 Generalised critical values

In the setting of a compact semi-algebraic set $\mathfrak{S}$, the extrema of a polynomial mapping are contained within the set of critical values of the map, denoted $K_0(f)$. However, if we consider a polynomial mapping restricted to a non-compact algebraic set, as is the case for unconstrained POP, the critical values of the map may not suffice.

In [89], Rabier introduces the set of *asymptotic critical values*, denoted $K_\infty(f)$, the union of which with the critical values is called the set of *generalised critical values*, denoted $K(f)$. These values provide a generalisation of Ehresmann's fibration theorem to non-proper settings. Let $\boldsymbol{f}$ be a polynomial mapping defined by

$$\boldsymbol{f} : \boldsymbol{x} \in X \mapsto (f_1(\boldsymbol{x}), \dots, f_p(\boldsymbol{x})) \in \mathbb{K}^p,$$

where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and X is a smooth variety defined by a reduced, regular sequence $\boldsymbol{g} = (g_1, \dots, g_m)$. Them, the restriction of $\boldsymbol{f}$ to $X \setminus \boldsymbol{f}^{-1}(K(\boldsymbol{f}))$ is a locally trivial fibration. This means that for all connected open sets $U \subset X \setminus K(\boldsymbol{f})$, for all $y \in U$ there exists a diffeomorphism φ such that the following diagram commutes

$$\begin{array}{ccc} \boldsymbol{f}^{-1}(y) \times U & \xrightarrow{\varphi} & \boldsymbol{f}^{-1}(U) \\ & \searrow^{\pi} & \downarrow \boldsymbol{f} \\ & & U \end{array}$$

where π is the projection map onto U [58, Theorem 3.1].

Denoting the Jacobian of $\boldsymbol{f}$ and $\boldsymbol{g}$ by $\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})$, the set of critical values of a polynomial mapping is defined in the usual way:

$$K_0(\boldsymbol{f}) = \{c \in \mathbb{C}^p \mid \exists \boldsymbol{x} \in X \text{ s.t. } \boldsymbol{f}(\boldsymbol{x}) = c \text{ and } \operatorname{rank}(\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})(\boldsymbol{x})) < m + p\}.$$

Then, the set of *asymptotic critical values* of the mapping $\boldsymbol{f}$ is defined to be the set:

$$K_\infty(\boldsymbol{f}) = \{c \in \mathbb{C}^p \mid \exists (\boldsymbol{x}_t)_{t\in\mathbb{N}} \subset X \text{ s.t. } \|\boldsymbol{x}_t\| \to \infty, \boldsymbol{f}(\boldsymbol{x}_t) \to c \text{ and } \|\boldsymbol{x}_t\| \nu(\mathrm{d}\boldsymbol{f}(\boldsymbol{x}_t)) \to 0\},$$

where $\mathrm{d}\boldsymbol{f}$ is the differential of the mapping $\boldsymbol{f}$ and ν is the distance to the set of singular operators. Defined in this way, Kurdyka, Orro and Simon show in [62] that $K_\infty(\boldsymbol{f})$ satisfies a generalised Sard's theorem. This means that the codimension of $K_\infty(\boldsymbol{f})$ is greater than or equal to one.

Combined with the fibration property of the generalised critical values, computing $K(\boldsymbol{f})$ allows one to solve many problems in real algebraic geometry, such as computing sample points for each connected component of a semi-algebraic set defined by a single inequality, and in the case where $p = 1$, solving POP. For a polynomial $f \in \mathbb{Q}[\boldsymbol{x}]$ and its infimum f^*, there are three cases:

- f^* is reached. Then, f^* is a critical value of f;
- f^* is reached only at infinity, meaning that there is no minimiser $\boldsymbol{x} \in X$ but instead a path $(\boldsymbol{x}_t)_{t\in\mathbb{N}} \subset X$ that approaches the infimum as $\|\boldsymbol{x}_t\| \to \infty$. Then, f^* is an asymptotic critical value of f;
- $f^* = -\infty$.

We give an example of a polynomial whose infimum falls into the second of these three cases.

Example 2.1. *Consider the polynomial $f = x^2 + (xy-1)^2 \in \mathbb{R}[x, y]$ and let f^* be its infimum over $\mathbb{R}^2$. Firstly, since the gradient of f is equal to $(2x + 2y(xy-1), 2x(xy-1))$ we see that there is exactly one critical point $(x, y) = (0, 0)$. Thus, the only critical value of f is 1. However, notice that if one takes a path $\gamma(t) = (t, 1/t)$, then as $t \to 0$ we have $\|\gamma(t)\| \to \infty$ and $f(\gamma(t)) \to 0$. Hence, $f^* \leq 0$. Since f is a sum of squares, we know that $f^* \geq 0$ and so $f^* = 0$. Therefore, 0 is an asymptotic critical value of f.*

Thus, as the first step in a strategy for exact polynomial optimisation, one computes exact representations of all the generalised critical values of f. We can do this by computing a polynomial whose roots contain these values and then using a real root isolation algorithm, such as the one in [91], to compute isolating intervals with rational endpoints for all the real roots.

Beginning with the critical values, let I be the ideal defined by the polynomials $f - c$, $\boldsymbol{g}$ and the maximal minors of the Jacobian of f and $\boldsymbol{g}$, where $\boldsymbol{g}$ is a reduced regular sequence defining a smooth algebraic set $\mathbf{V}(\boldsymbol{g})$. By the Jacobian criterion [27, Corollary 16.20], to obtain a polynomial representation of the critical values of f restricted to $\mathbf{V}(\boldsymbol{g})$, one can compute a geometric resolution of I, giving a triangular representation analogously to Gaussian elimination in the linear setting. Then, one needs to compute the set of asymptotic critical values of f restricted to $\mathbf{V}(\boldsymbol{g})$ using, for instance, the algorithm presented in the paper [58].

With $C = \{c_1, \ldots, c_k\} \subset \mathbb{R}$ the finite set of generalised critical values, one can compute isolating intervals with rational endpoints for each point in C and so one can choose rational numbers $r_1, \ldots, r_k$ so that

$$r_1 < c_1 < r_2 < \cdots < r_k < c_k.$$

Finally, one can use the fibration property of $K(f)$ to decide which, if any, of the generalised critical values of f is the infimum f^* by deciding the emptiness of the fibres of f at $r_1, \ldots, r_k$ using, for example, the algorithm proposed in [95].

Therefore, one may use this generalised critical value method to solve POPs in non-compact situations, free of numeric approximations and potential numeric errors. Moreover, the output is an exact representation of the infimum of the input polynomial, in the form of a polynomial and an isolating interval.

In order to execute the above strategy for exact polynomial optimisation, we need to solve polynomial systems. In particular, for the computation of the generalised critical values, polynomial system solving is key. There are many approaches to solving polynomial systems such as homotopy deformation, Gröbner bases or geometric resolutions. In this thesis, we focus primarily on Gröbner bases and also make use of geometric resolutions and we now briefly explain the reasons for this.

2.1.3 Homotopy deformation

Let $S_1 \in \mathbb{C}[x_1, \ldots, x_n]^p$ be a zero-dimensional polynomial system, meaning that its set of solutions is finite. Homotopy methods involve defining a deformation between the system S_1 we want to solve and a second zero-dimensional system S_0 of the same degree whose solutions are easier described. To do so, let t be a parameter and define the system S_t by

$$S_t = (1-t)S_0 + tS_1 \in \mathbb{C}[t, x_1, \ldots, x_n]^p,$$

so that when $t = 0$, $S_t = S_0$, the system we expect to solve easily, and when $t = 1$, $S_t = S_1$, the target system. Hence, the solutions of the system S_t define a path in $\mathbb{C}^n$ between the roots of S_0 and S_1. The founding idea is then to solve the system S_0 and for each solution traverse this path by increasing the value of t step-by-step, computing the solutions of intermediate systems through Newton iteration, until you reach the corresponding solutions of the system S_1.

There are many algorithms for homotopy continuation, what is laid out above, which can be categorised into numeric and symbolic algorithms, for more information see [2, 49, 71]. However, issues can arise when a path goes through an *ill-conditioned* system. For example, suppose that for some $t \in (0, 1)$, the intermediary system S_t had solutions with multiplicity. This would mean that two solutions of S_0 would converge into one solution. In fact, due to loss of precision with floating point arithmetic or by overlapping intervals in interval arithmetic, the solutions need not have multiplicity for this issue to occur but simply be too close together. For this reason, the Newton-like iteration algorithms that compute the solutions of the intermediate systems step-by-step struggle near a cluster of roots.

Similar issues can occur with paths of solutions that tend to infinity. While the number of solutions remains constant in the projective setting, this would amount to losing a solution of the system. Therefore, in this thesis we focus on symbolic methods that avoid such issues with lack of precision and loss of roots such as Gröbner bases and geometric resolutions.

2.1.4 Gröbner bases

Computing Gröbner bases will be the main tool in this thesis for polynomial system solving. Given a zero-dimensional ideal $I \subset \mathbb{C}[x_1, \ldots, x_n]$, we wish to compute a polynomial encoding of the set $\mathbf{V}(I)$. We say that the ideal I is in *shape position* if its LEX Gröbner basis has the form

$$\{x_1 - f_1(x_n), \ldots, x_{n-1} - f_{n-1}(x_n), f_n(x_n)\},$$

where the degree of f_n is the degree of the ideal I. In broad terms, computing a LEX Gröbner basis directly is timely. A fast method commonly used in practice is instead to first compute a DRL Gröbner basis of I, using for example Faugère's $\mathsf{F_5}$ algorithm [29], and then use a change of ordering algorithm, such as FGLM [31], to recover a LEX Gröbner basis.

In this way, we can avoid any issues with loss of precision and we are guaranteed to recover all the solutions to an input polynomial system satisfying the shape assumption. Therefore, this method fits well into the framework for exact polynomial optimisation laid out above as we must compute a polynomial whose roots contain all the generalised critical values of the target polynomial. However, while the framework is clear, there are still many questions to raise concerning the complexity of such a procedure.

2.2 Problem statements

Problem 1. A popular method for polynomial system solving is to first compute a DRL Gröbner basis of the system and then use a change of ordering algorithm to obtain a LEX basis. While the computation of a DRL Gröbner basis for determinantal systems deriving from maximal minors is well understood, a study of the input DRL Gröbner basis and the change of ordering step for this class of systems is lacking. What is the structure of the generic DRL Gröbner basis in this

maximal minor determinantal setting? Moreover, can the complexity estimates for the critical value computation be improved by taking advantage of this determinantal structure?

Problem 2. Structured determinantal ideals appear frequently in applications. For example, solving semi-definite programs to find sums of squares decompositions, a popular method for polynomial optimisation, leads one to investigate moment matrices which are symmetric. How does this additional structure affect LEX Gröbner basis computations on the determinantal ideals derived from the rank defects of these symmetric matrices?

Problem 3. The computation of the asymptotic critical values is the bottleneck of the overall generalised critical value method for solving POP. Does there exist a more efficient algorithm for computing the set of asymptotic critical values that brings the complexity of computing the infimum of a polynomial map $f : X \to \mathbb{R}$, with domain X a smooth algebraic set defined by a reduced, radical sequence $g_1, \ldots, g_m \in \mathbb{K}[x_1, \ldots, x_n]$, where $f, g_1, \ldots, g_m$ have degree d, exactly to within $d^{O(n)}$ operations in the ground field $\mathbb{K}$, either $\mathbb{C}$ or $\mathbb{R}$?

Problem 4. One aims to understand the underlying algebraic structure of sums of squares decompositions. As first steps, fixing a number of homogeneous variables, what is the degree of the variety of all sums of two squares? Moreover, given a generic polynomial that is a sum of squares, what is the structure of all its possible decompositions?

2.3 Prior works and Contributions

2.3.1 Problems 1 and 2

Prior works. Determinantal ideals are an active area of study in commutative algebra. A popular technique in this subject is to use the theory of Gröbner bases to connect such ideals with combinatorial objects to utilise the properties of Stanley-Reisner rings of simplicial complexes, see for example [16, 18, 19, 21, 22, 104].

In [22], the authors gave an explicit formula for the Hilbert series of the determinantal ideals defined by variable matrices. Specialising this result to determinantal matrices derived from maximal minors, the authors of [33] find the Hilbert series of ideals defining the set of critical values/points of a polynomial restricted to an algebraic set under some regularity assumptions. Using this, the authors give an upper bound on the number of arithmetic operations necessary for computing a LEX Gröbner basis of such an ideal within the DRL to LEX framework in the same paper [33, Theorem 3].

Firstly, based on [5, Theorem 7], the authors of [33, Theorem 3] use the Hilbert series of generic determinantal ideals to analyse the complexity of the DRL step using Faugère's $\mathsf{F_5}$ algorithm [29]. Here, and in the whole text, complexity estimates are given in terms of arithmetic operations in the ground field $\mathbb{K}$. Then, to obtain a LEX Gröbner basis, since we are in the zero-dimensional case, they use the FGLM algorithm to perform the change of ordering [31]. The complexity of FGLM is $O(nD^3)$, where n is the number of variables and D is the degree of the ideal. For example, consider the projection map ϕ from $\mathbb{K}^n$ onto the first coordinate restricted to an algebraic set defined by a reduced regular sequence $g_1, \ldots, g_m$ where $\deg g_i = d$ for $1 \leq i \leq m$. In [84, Theorem 2.2], the authors use the Thom-Porteous-Giambelli formula to prove that the degree of the ideal defining the critical points of the projection map ϕ is

$$D = d^p(d-1)^{n-m}\binom{n-1}{m-1}.$$

Under some stability assumptions, the authors of [30] and [83] improve upon the FGLM algorithm by applying fast linear algebra techniques. Their algorithms have complexity $O^{\sim}(D^{\omega})$

and $O(nD^\omega \log(D))$ respectively, where ω is the exponent of matrix multiplication. The best known theoretical bound for ω is 2.37286 given in [3].

Other algorithms have been introduced to take advantage of the sparsity of the multiplication matrix T_{x_n}, such as the Sparse-FGLM algorithm in [32]. Under the same stability assumptions and when the ideal of interest is in shape position, meaning that the leading monomials of the LEX basis are $x_1, \ldots, x_{n-1}, x_n^D$, the Sparse-FGLM algorithm relies primarily on the sparsity of the matrix T_{x_n} associated to the linear map of multiplication by x_n in the finite dimensional quotient algebra $\mathbb{K}[x_1, \ldots, x_n]/I$. It has complexity $O(qD^2 + nD\log^2 D)$, where q is the number of non-trivial columns of the matrix T_{x_n}.

In some cases, $q \in O(D)$ and so the Sparse-FGLM algorithm is not always faster than the algorithms of [30, 83] asymptotically. However, very recently, under the same shape and stability assumptions, the authors of [12] designed an algorithm which improves upon [30, 32, 83] by focusing on the structure of the matrix T_{x_n}, instead of just its sparsity, with a complexity of $O^\sim(q^{\omega-1}D)$.

In order to accurately compare the algorithms of [12, 32] to those of [30, 83], one must first estimate the parameter q. Moreover, as q is a fundamental parameter in FGLM-like algorithms, a bound on q is useful for any algorithm that relies on the multiplication matrix T_{x_n}.

Using the results of [79] on the structure of the DRL staircase of generic complete intersections, the number q is studied in [32] for this class of systems. Moreover, using this structure, the authors of [32] proved that the matrix T_{x_n} is such that it can be computed free of arithmetic operations. However, similar results were previously unknown for other classes of ideals, such as generic determinantal ideals. This meant that the complexity improvements of [12, 32] were not fully understood for many important problems, for example critical value computation.

Contributions. The results towards Problem 1 were a joint a work with Jérémy Berthomieu, Alin Bostan and Mohab Safey El Din and have been published in the Journal of Algebra. We begin by precisely defining the class of ideals we will consider for Problem 1, what we call generic determinantal sum ideals.

Definition 2.2. *With $\mathbb{K}$ an infinite field, let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal which is the sum of m polynomials of degree at most d and the maximal minors of a matrix with polynomial entries also of degree at most d. We say that I is a generic determinantal sum ideal if the following three conditions hold:*

- *the ideal I is in shape position, meaning that the reduced LEX Gröbner basis with $x_1 \succ \cdots \succ x_n$ has leading monomials $x_1, \ldots, x_{n-1}, x_n^D$ where D is the degree of I,*
- *the Hilbert series H of $\mathbb{K}[x_1, \ldots, x_n]/I$ is equal to*

$$H = \frac{\det(M(t^{d-1}))}{t^{(d-1)\binom{m-1}{2}}} \frac{(1-t^d)^m (1-t^{d-1})^{n-m}}{(1-t)^n}$$

  where $M(t)$ is the $(m-1) \times (m-1)$ matrix whose (i,j)th entry is $\sum_k \binom{m-i}{k}\binom{n-1-j}{k} t^k$,
- *for all $e \geq 1$, the Hilbert series of $(\mathbb{K}[x_1, \ldots, x_n]/I)/\langle x_n^e \rangle$ is equal to the series $(1-t)H$ truncated at the first non-positive coefficient.*

In Chapter 4, we prove that, under some regularity assumptions, the ideal defining the set of critical values of a generic polynomial map restricted to a smooth algebraic set falls in this class.

Our first main contribution is a structure result on the DRL Gröbner basis of such ideals, giving the following theorem.

Theorem 2.3. *Let I be a generic determinantal sum ideal so that the conditions of Definition 2.2 hold. Assume that a reduced Gröbner basis of I with respect to a DRL ordering is known. Then the multiplication matrix T_{x_n} can be constructed without performing any arithmetic operations.*

Taking this structure into account, we then give formulae for the number of non-trivial columns of T_{x_n}, which we denote q. We give an exact formula in the case $d = 2$ while for $d \geq 3$ we give an asymptotic formula.

Theorem 2.4. *Let I be a generic determinantal sum ideal so that the conditions of Definition 2.2 hold, and let T_{x_n} be the matrix associated to the linear map of multiplication by x_n. Denote by q the number of non-trivial columns of T_{x_n}. Then, for $d = 2$ and $n \gg m$,*

$$q = \sum_{k=0}^{m-1} \binom{n-m-1+k}{k} \binom{m}{\lfloor 3m/2 \rfloor - 1 - j}. \tag{2.1}$$

Moreover, for $d \geq 3$ and $n \to \infty$,

$$q \approx \frac{1}{\sqrt{(n-m)\pi}} \sqrt{\frac{6}{(d-1)^2-1}} d^m (d-1)^{n-m} \binom{n-2}{m-1}. \tag{2.2}$$

Finally, we use these results and the Sparse-FGLM algorithm [32, Theorem 3.2] to give a complexity result for the change of ordering from DRL to LEX for generic determinantal sum ideals.

Theorem 2.5. *Let I be a generic determinantal sum ideal so that the conditions of Definition 2.2 hold. Assume that a reduced DRL Gröbner basis of I is known. Then, for $d \geq 3$, the arithmetic complexity of computing a LEX Gröbner basis of I is upper bounded by*

$$O\left(\frac{d^{3m}(d-1)^{3(n-m)}}{\sqrt{(n-m)d\pi}} \binom{n-2}{m-1} \binom{n-1}{m-1}^2 \right).$$

Hence, the complexity gain of Sparse-FGLM over FGLM for generic determinantal sum systems is approximately

$$O\left(\frac{q}{nD}\right) \approx O\left(\frac{\sqrt{n-m}}{n^2(d-1)}\right).$$

Furthermore, for Problem 2 we consider the class of generic symmetric determinantal ideals. These are the determinantal ideals that are derived from a symmetric matrix with generic polynomial entries. Specifically, for the symmetric matrix $S = (s_{i,j})_{1 \leq i,j \leq \ell}$ with entries the variables $\boldsymbol{s} = (s_{1,1}, s_{2,1}, s_{2,2}, \ldots, s_{n,1}, \ldots, s_{\ell,\ell})$ and some $r \in \mathbb{N}$, $\mathcal{S}_r$ is the homogeneous ideal generated by all the $(r+1)$-minors of S. For fixed $n, d \in \mathbb{N}$, let $S^{k,d} = (f_{i,j})_{1 \leq i,j \leq \ell}$ be an $\ell \times \ell$ symmetric matrix with entries in $\mathbb{K}[x_1, \ldots, x_n]_{\leq d}$. Then, $\mathcal{S}_r^{n,d}$ is the ideal defined by the $(r+1)$-minors of $S^{n,d}$. Let $\mathcal{H}_r$ and $\mathcal{H}_r^{n,d}$ be the reduced numerators of the Hilbert series of the ideals $\mathcal{S}_r$ and $\mathcal{S}_r^{n,d}$ respectively. Now, let $\mathcal{S}_r^{n,d,\mathrm{h}}$ be the homogenised ideal of $\mathcal{S}_r^{n,d}$. The following results, made in collaboration with Huu Phuoc Le, were communicated at the ISSAC 2022 conference in the form of a paper.

Conjecture 2.6.

1. *Given $r \in \mathbb{N}$, the reduced numerator $\mathcal{H}_r(t)$ of the Hilbert series of the symmetric determinantal ideal $\mathcal{S}_r$ is unimodal.*
2. *For $e \geq 1$, let $\mathcal{Q}_r^{n,d,e}$ be the Hilbert series of $\mathbb{K}[x_0, \ldots, x_n] / \left(\mathcal{S}_r^{n,d,\mathrm{h}} + \langle x_0, x_n^e \rangle \right)$. We conjecture that $\mathcal{Q}_r^{n,d,e} = \left[(1-t^e)\mathcal{H}_r^{n,d}(t) \right]_+$, which is the series $(1-t^e)\mathcal{H}_r^{n,d}(t)$ truncated at its first non-positive coefficient.*

Then, under these regularity assumptions, we prove the following structure result for generic symmetric determinantal ideals.

Theorem 2.7. *Given $r, \ell, d \in \mathbb{N}$ and $n = \binom{\ell-r+1}{2}$, there exists a non-empty Zariski-open subset $\mathscr{F}_r$ of $\mathbb{K}[x_1, \ldots, x_n]_{\leq d}^{\ell(\ell+1)/2}$ such that, when the entries of $S^{n,d}$ are taken in $\mathscr{F}_r$, the following holds:*

The ideal $\mathcal{S}_r^{n,d}$ is zero-dimensional and radical. When Conjecture 2.6 holds and a reduced Gröbner basis of $\mathcal{S}_r^{n,d}$ w.r.t. $\prec_{\text{DRL}}$ is known, the matrix T_{x_n} of multiplication by x_n can be constructed without any arithmetic operations. Moreover, the number of dense columns of T_{x_n} equals the largest coefficient of the Hilbert series $\mathcal{H}_r^{n,d}$.

Using the Sparse-FGLM algorithm we then provide a dedicated complexity result for the change of ordering from DRL to LEX Gröbner bases for symmetric determinantal ideals.

Theorem 2.8. *Given $r, \ell, d \in \mathbb{N}$ and $n = \binom{\ell-r+1}{2}$, we consider the matrix $S^{n,d}$ with entries taken in the Zariski-open set $\mathscr{F}_r$ defined in Theorem 2.7. Assume that Conjecture 2.6 holds and the reduced Gröbner basis of $\mathcal{S}_r^{n,d}$ w.r.t. $\prec_{\text{DRL}}$ is known. Then as $d \to \infty$, the Sparse-FGLM algorithm computes a $\prec_{\text{LEX}}$ Gröbner basis of $\mathcal{S}_r^{n,d}$ within*

$$O\left(q\mathcal{H}_r^{n,d}(1)^2\right) = O\left(qd^{2n}\mathcal{H}_r(1)^2\right) = O\left(qd^{2n}\left(\prod_{i=0}^{\ell-r-1}\frac{\binom{\ell+i}{2i+r}}{\binom{2i+1}{i}}\right)^2\right)$$

arithmetic operations in $\mathbb{K}$ where q is the number of dense columns of the multiplication matrix T_{x_n}. Moreover, as $d \to \infty$, q is bounded above by

$$d^{n-1}\mathcal{H}_r(1) = \sqrt{\frac{6}{n\pi}}d^{n-1}\prod_{i=0}^{n-r-1}\frac{\binom{\ell+i}{2i+r}}{\binom{2i+1}{i}}.$$

In Chapter 5, we will investigate further three special cases and obtain finer complexity results for each.

2.3.2 Problem 3

Prior works. Computing the set of critical values of a polynomial mapping $\boldsymbol{f} = (f_1, \ldots, f_p)$ restricted to an algebraic set $X = \mathbf{V}(\boldsymbol{g}) = \mathbf{V}(g_1, \ldots, g_m)$ is classical. Under some regularity assumptions, the algebraic set defined by the intersection of X with the variety defined by the maximal minors of $\text{jac}(\boldsymbol{f}, \boldsymbol{g})$ is equal to set the critical points of $\boldsymbol{f}$ [27, Corollary 16.20].

The first work towards the computation of the asymptotic critical values of a polynomial mapping, in the unrestricted setting, was given in [62]. In this paper, the authors give a geometric characterisation of $K_\infty(\boldsymbol{f})$ that allows one to construct an algebraic set of codimension at least one in $\mathbb{C}^p$ that contains the asymptotic critical values by using algorithms to perform ideal-theoretic operations, such as Gröbner basis based algorithms. Then, in [58], the authors tackle the problem of computing the generalised critical values of a polynomial mapping restricted to an algebraic set. The algorithm given in this paper follows a similar framework of defining algebraic sets, considering their intersections with linear hyperspaces and projecting onto the target space. However, this algorithm requires the construction of $(p(m+p))^{\binom{n}{m+p}}$ locally closed sets in $\mathbb{C}^{(n+1)\binom{n}{m+p}+p+n}$ before projecting each onto $\mathbb{C}^p$, making the algorithm impractical, especially when compared to the computation of the critical values of $\boldsymbol{f}$. Furthermore, there are no experimental results or a complexity analysis given for this algorithm.

Returning to the unrestricted setting, and with the additional assumption that the polynomial mapping has only one component, there have been several attempts to improve this algorithmic pattern. For example, the author of [92] makes the connection between generalised critical values and properties of polar varieties. This connection was later exploited in [59] to build rational arcs that reach all the generalised critical values of a polynomial. Moreover, in [60], the authors make a distinction between asymptotic critical values, detecting those that are found non-trivially, meaning away from the critical locus of the polynomial. However, such a distinction is not necessary for the problem of polynomial optimisation and so we do not consider this in our contributions.

Contributions. We assume that $\boldsymbol{f}$ satisfies the following regularity assumption (R): "The Zariski closure of $X \setminus \mathrm{crit}(\boldsymbol{f}, X)$ is X", where

$$\mathrm{crit}(\boldsymbol{f}, X) = \{\boldsymbol{x} \in X \mid \mathrm{rank}(\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})(\boldsymbol{x})) < m + p\}$$

is the critical locus of $\boldsymbol{f}$ on X. Hence, Assumption (R) is equivalent to requiring that for a generic point $\boldsymbol{x} \in X$, $\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})(\boldsymbol{x})$ has full rank.

Under this regularity assumption, my co-authors Jérémy Berthomieu and Mohab Safey El Din and I solve Problem 3 by designing algorithms that bring the computation of the asymptotic critical values in the case $p = 1$, the crucial case for polynomial optimisation, within a complexity of $d^{O(n)}$. These results have been submitted to the Journal of Symbolic Computation. To prove this, we first give a degree result.

Theorem 2.9. *Let X be a smooth algebraic set defined by a reduced regular sequence $\boldsymbol{g} = (g_1, \ldots, g_m)$. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping from X to $\mathbb{K}^p$ satisfying Assumption (R). Let $d = \max(\deg f_1, \ldots, \deg f_p, \deg g_1, \ldots, \deg g_m)$. Then, the asymptotic critical values of $\boldsymbol{f}$ are contained in a hypersurface of degree at most*

$$p d^{n-p-1} \sum_{i=0}^{p+1} \binom{n+p-1}{m+2p-i} d^i.$$

While our algorithms allow the computation of polynomial mappings, we give a dedicated complexity result in the special case $p = 1$ which is the case of interest for polynomial optimisation.

Theorem 2.10. *Let $\boldsymbol{g} = (g_1, \ldots, g_m)$ be a reduced regular sequence defining a smooth algebraic set X. Let $f \in \mathbb{K}[\mathbf{z}]$ be a polynomial mapping from X to $\mathbb{K}$ satisfying Assumption (R). Let $d = \max(\deg f, \deg g_1, \ldots, \deg g_m)$ and $D = d^{n-2} \sum_{i=0}^{2} \binom{n}{m+2-i} d^i$. Then, there exists an algorithm which, on input $f, \boldsymbol{g}$, outputs a non-zero polynomial $H \in \mathbb{K}[c]$ such that $K_\infty(\boldsymbol{f}) \subset \mathbf{V}(H)$ using at most*

$$O^{\sim}(n^2 d^{n+2} D^5)$$

arithmetic operations in $\mathbb{K}$.

However, in the more general setting of polynomial mappings, we achieve the following complexity.

Theorem 2.11. *Let $\boldsymbol{g} = (g_1, \ldots, g_m)$ be a reduced regular sequence defining a smooth algebraic set X. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping from X to $\mathbb{K}^p$ satisfying Assumption (R). Let $d = \max(\deg f_1, \ldots, \deg f_p, \deg g_1, \ldots, \deg g_m)$ and $D = d^{n-p-1} \sum_{i=0}^{p+1} \binom{n+p-1}{m+2p-i} d^i$. Then, there exists an algorithm which, on input $\boldsymbol{f}$ and $\boldsymbol{g}$, outputs p finite lists of non-zero polynomials $G_i \subset \mathbb{K}[\mathbf{c}]$ such that $K_\infty(\boldsymbol{f}) \subset (\mathbf{V}(G_1) \cup \cdots \cup \mathbf{V}(G_p)) \subsetneq \mathbb{C}^p$ using at most*

$$O^{\sim}\left(p^2 D^{p+5} + n^2 d^{n+2} D^{p+4}\right)$$

arithmetic operations in $\mathbb{K}$.

2.3.3 Problem 4

Prior works. The study of the decompositions of homogeneous polynomials has a long history. From the classical works of Sylvester [105], the study of decompositions of homogeneous polynomials by powers of linear forms, is still an active area of research. In [38] it was proved that general homogeneous polynomials of degree $2d$ in $n+1$ variables are sums of at most 2^n squares. For fixed n, this bound is sharp for all sufficiently large d.

The minimal number of squares required in a decomposition of a polynomial is known as the SOS-rank. The authors of [73] investigate this rank for generic polynomials in two variables. Then, in [37], the authors give a conjecture on the generic SOS-rank of polynomials, in terms of number of variables and degree.

Contributions. We consider SOS decompositions of polynomials of degree $2d$. Let V be a complex vector space of dimension $n+1$. Then, the space of homogeneous polynomials of degree $2d$ in $n+1$ variables will be denoted by $\mathrm{Sym}^{2d} V$.

Definition 2.12. *Let $f \in \mathrm{Sym}^{2d} V$. The polynomial f has SOS-rank k if k is the minimum number such that there exist $f_1, \ldots, f_k \in \mathrm{Sym}^d V$ such that*

$$f = \sum_{i=1}^{k} f_i^2.$$

Towards answering this problem, we define and study two varieties related to exact SOS decompositions. The first is defined by all polynomials of SOS-rank less than or equal to k.

Definition 2.13. *Let SOS_k be the subvariety in $\mathrm{Sym}^{2d} V$ obtained from the Zariski closure of the set of all SOS-rank k polynomials.*

$$\mathrm{SOS}_k = \overline{\{f_1^2 + \cdots + f_k^2 \mid f_i \in \mathrm{Sym}^d V\}}.$$

The generic SOS-rank is the smallest number k such that SOS_k covers the ambient space.

Another object that can be investigated is the set of all different decompositions of a polynomial f generic in SOS_k for fixed $k \in \mathbb{N}$.

Definition 2.14. *Let $f \in \mathrm{SOS}_k$ be a generic polynomial. We define the variety of all the SOS-rank k decompositions of f as*

$$\mathrm{SOS}_k(f) = \left\{ (f_1, \ldots, f_k) \in \prod_{i=1}^{k} \mathrm{Sym}^d V \,\middle|\, \sum_{i=1}^{k} f_i^2 = f \right\}.$$

For the first object of interest: $\mathrm{SOS}_k(f)$, we give its exact structure in the case $k = 2$.

Theorem 2.15. *Let $f \in \mathrm{SOS}_2$ be a generic polynomial of SOS-rank two. Then, $\mathrm{SOS}_2(f)$ has two irreducible components isomorphic to $\mathrm{SO}(2)$. Hence, $\mathrm{SOS}_2(f)$ is isomorphic to $\mathrm{O}(2)$.*

Note that the degree of the polynomial f is not important here. In more generality, we calculate the dimension of this object.

Theorem 2.16. *Let $f \in \mathrm{SOS}_k$ be generic with $k \leq n$. Then,*

$$\dim \mathrm{SOS}_k(f) = \binom{k}{2}.$$

Furthermore, we conjecture that $\mathrm{SOS}_k(f)$ is also isomorphic to $\mathrm{O}(k)$ and we give some dimension counts and experiments to support this conjecture. On the other hand, we give the degree of SOS_1 and SOS_2.

Theorem 2.17. *Let $N = \dim \mathrm{Sym}^d V = \binom{n+d}{d}$. The degrees of the varieties of squares and of sum of two squares in $\mathbb{P}(\mathrm{Sym}^{2d} V)$ are given by*

$$\deg(\mathrm{SOS}_1) = 2^{N-1}, \ \deg(\mathrm{SOS}_2) = \prod_{i=0}^{N-3} \frac{\binom{N+i}{N-2-i}}{\binom{2i+1}{i}}.$$

My collaborators, Giorgio Ottaviani, Mohab Safey El Din and Ettore Turatti, and I have submitted these results to the Journal of Pure and Applied Algebra.

2.4 Structure of the thesis

We begin with the preliminaries that are required to present our results in Chapter 3. Sections 3.1 and 3.2 give the basic definitions and propositions in algebra and algebraic geometry that will be used frequently throughout this thesis. Section 3.3 is dedicated to Gröbner bases, the basic definitions and, in particular, introducing the change of ordering algorithms that will be central to the study of Problems 1 and 2. Finally, Section 3.4 will give some background on sums of squares decompositions and exact polynomial optimisation.

We then give our contributions which are organised into the following four chapters.

- Chapter 4 details our contributions towards Problem 1, the derivation of asymptotic formulae for a fundamental parameter in the change of ordering of Gröbner bases for critical value systems, improving upon the previously known complexity estimates.

  This work has been published as an article in the Journal of Algebra: "Gröbner bases and critical values: The asymptotic combinatorics of determinantal systems" (Jérémy Berthomieu, Alin Bostan, Andrew Ferguson and Mohab Safey El Din) [10].

- Chapter 5 details our contributions towards Problem 2, analysing the change of ordering step for more general classes of determinantal systems and giving the first dedicated complexity estimates in those cases.

  The contents of this chapter have been presented as a conference paper published in the proceedings of ISSAC 2022, Lille, France: "Finer Complexity Estimates for the Change of Ordering of Gröbner Bases for Generic Symmetric Determinantal Ideals" (Andrew Ferguson and Huu Phuoc Le) [35].

- Chapter 6 details our contributions towards Problem 3, introducing new efficient algorithms for computing the set of asymptotic critical values of polynomial mappings from algebraic sets satisfying some basic regularity assumption.

  These algorithms and consequent results form an article that has been submitted to the Journal of Symbolic computation: "Computing the set of asymptotic critical values of polynomial mappings from smooth algebraic sets" (Jérémy Berthomieu, Andrew Ferguson and Mohab Safey El Din).

- Chapter 7 details our contributions towards Problem 4, giving the degree of the variety of sums of two squares and working towards understanding the algebraic structure of all possible sums of squares decompositions of a generic sum of squares.

  This work has been submitted to the Journal of Pure and Applied Algebra as the paper: "On the degree of varieties of sum of squares" (Andrew Ferguson, Giorgio Ottaviani, Mohab Safey El Din and Ettore Turatti).

Finally, in Chapter 8, we conclude by summarising our results and by analysing the next steps of each of the four problems identified in Section 2.2.

Chapter 3

Preliminaries

We begin by recalling some basic notions of algebra and algebraic geometry that will be used frequently throughout this thesis. Then, we introduce Gröbner bases and in particular the FGLM and Sparse-FGLM algorithms, given in [31] and [32] respectively. A key parameter in the latter of which will be the focus of our study of Problems 1 and 2. Finally, we give some background on exact polynomial optimisation.

3.1 Algebra

We start with preliminaries on commutative algebra and algebraic geometry. The contents of these two sections can be found in more detail in the standard texts on algebra by Cox, Little and O'Shea [24], Eisenbud [27], Hartshorne [47] and Lang [65].

Definition 3.1 [24, Definition 1.4.1, Lemma 1.4.3]**.** *A non-empty subset I of a ring R is called an* ideal *if the following conditions hold:*

1. *If f and g are elements of I then $(f+g) \in I$.*
2. *For all $f \in I$ and $g \in R$ we have $fg \in I$.*

For a subset $\boldsymbol{f} = \{f_1, \ldots, f_p\} \subset R$, we write $\langle \boldsymbol{f} \rangle$ for the ideal generated by $\boldsymbol{f}$, defined by

$$\langle \boldsymbol{f} \rangle = \{r_1 f_1 + \cdots + r_p f_p \mid r_1, \ldots, r_p \in R\}.$$

Lemma 3.2 [24, Section 4.3, Proposition 4.4.9]**.** *Let $I, J \subset R$ be ideals. Then, the following are ideals of R:*

1. *the* sum $I + J = \{f + g \mid \forall f \in I, g \in J\}$,
2. *the* intersection $I \cap J$,
3. *the* product $IJ = \{f_1 g_1 + \cdots + f_s g_s \mid \forall s \in \mathbb{N}, f_i \in I, g_i \in J \text{ for } 1 \leq i \leq s\}$,
4. *the* saturation $I : J^\infty = \{f \in R \mid \forall g \in J, \exists\, s \in \mathbb{N} \text{ such that } fg^s \in I\}$.

Definition 3.3 [24, Definition 3.1.1]**.** *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal. The ith* elimination ideal *of I is defined as $I \cap \mathbb{K}[x_{i+1}, \ldots, x_n]$.*

Definition 3.4 [24, Definition 4.2.2]**.** *Let $I \subset R$ be an ideal. The* radical *of I, denoted $\sqrt{I}$, is an ideal of R and is defined by*

$$\sqrt{I} = \{f \in R \mid \exists\, r \in \mathbb{N} \text{ such that } f^r \in I\}.$$

If $\sqrt{I} = I$, then we say that I is a radical *ideal.*

Example 3.5. *Consider the ring $\mathbb{Z}$. The ideal $\langle 12\rangle$ is not radical since $6^2 \in \langle 12\rangle$ and $6 \notin \langle 12\rangle$. We have that $\sqrt{\langle 12\rangle} = \langle 6\rangle$.*

Definition 3.6 [24, Definition 4.5.2, Definition 4.5.7]**.** *Let $I \subsetneq R$ be an ideal not equal to R. We say that I is a* prime *ideal if for all $f, g \in R$, $fg \in I$ implies that $f \in I$ or $g \in I$. We say that I is a* maximal *ideal if for all proper ideals $J \subsetneq R$, $I \subset J$ implies that $I = J$.*

Definition 3.7. *Let R be a commutative ring. A non-zero element $r \in R$ is a* zero-divisor *if there exists some non-zero $s \in R$ such that $sr = 0$. If R has no zero-divisors, then it is called an* integral domain.

Example 3.8. *Consider the ring $\mathbb{Z}$. Its prime ideals are $\langle \mathfrak{p}\rangle$, for all prime numbers $\mathfrak{p}$, and $\langle 0\rangle$. Indeed, for a ring R the ideal $\langle 0\rangle \subset R$ is prime if and only if R is an integral domain. Similarly, the ideal $\langle 0\rangle \subset R$ is maximal if and only if R is a field.*

Example 3.9. *Note that if an ideal is prime then it is radical. Thus, let $R = \mathbb{K}[x, y]$ be a ring with field $\mathbb{K}$. The ideal $I = \langle x^2 - 2xy + y^2\rangle$ is not radical, and so not prime, since $(x - y)^2 \in I$ but we have that $(x - y) \notin I$.*

Definition 3.10 [24, Section 5.2]**.** *Let $I \subset R$ be an ideal. Then,* congruence modulo I *defined by*

$$f \cong g \ mod\ I \qquad if\ (f - g) \in I$$

is an equivalence relation on R. The set of equivalence classes of R forms a ring and is called the quotient *of R by I and is denoted R/I.*

Example 3.11. *Let $\mathbb{K}$ be a field and $R = \mathbb{K}[x_1, \ldots, x_n]$ be a polynomial ring with an ideal I. Then, R is a $\mathbb{K}$-algebra and the quotient R/I inherits this structure and is called the* quotient algebra.

Definition 3.12 [24, Definition 7.1.6]**.** *A polynomial is* homogeneous *if all its non-zero terms have the same degree. An ideal is* homogeneous *if it can be generated by homogeneous polynomials.*

Definition 3.13 [24, Proposition 8.2.7]**.** *Let $f \in \mathbb{K}[x_1, \ldots, x_n]$ be a polynomial of degree d. Consider the expansion of f into homogeneous components, $f = \sum_{i=0}^{d} f_i$ where f_i has degree i. The* homogenisation *of f by a variable x_0 is the polynomial $f^{\mathrm{h}} \in \mathbb{K}[x_0, x_1, \ldots, x_n]$ defined by $f^{\mathrm{h}} = \sum_{i=1}^{d} x_0^{d-i} f_i$.*

Definition 3.14 [24, Definition 8.4.1]**.** *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal. The* homogenisation *of I by the variable x_0 is the homogeneous ideal $I^{\mathrm{h}} \subset \mathbb{K}[x_0, x_1, \ldots, x_n]$ defined by*

$$I^{\mathrm{h}} = \{f^{\mathrm{h}} \mid f \in I\}.$$

Example 3.15 [24, Example 8.4.3]**.** *Consider the ideal $I = \langle x - z^3, y - z^2\rangle \subset \mathbb{K}[x, y, z]$. Let I' be the ideal defined by the homogenisation of the polynomials defining I with respect to a new variable w, $\langle w^2x - z^3, wy - z^2\rangle$. Then, we claim that $I' \neq I^{\mathrm{h}}$. To see this, note that the polynomial*

$$x - z^3 - z(y - z^2) = (x - yz) \in I.$$

Therefore, $(wx - yz) \in I^{\mathrm{h}}$. However, since $I' = \langle w^2x - z^3, wy - z^2\rangle$ is homogeneous, the only degree two polynomials in I' are multiples of $wy - z^2$. Hence, $(wx - yz) \notin I'$.

Definition 3.16 [24, Proposition 9.3.3]**.** *Let $R = \mathbb{K}[x_0, x_1, \ldots, x_n]$ be a polynomial ring with a homogeneous ideal I. Let R_d denote the union of 0 with the set of all homogeneous polynomials of degree d in R. The* Hilbert series *of the quotient algebra R/I (equivalently of the ideal I) is defined as*

$$H_{R/I}(t) = \sum_{d=0}^{\infty} \left(\dim_{\mathbb{K}} R_d/(I \cap R_d)\right) t^d$$

where $\dim_{\mathbb{K}}$ *means dimension as a* $\mathbb{K}$*-vector space. The* d*th coefficient of the Hilbert series is the number of monomials of degree* d *not in* I*. We now define the Hilbert series of a non-homogeneous ideal* $I \subset \mathbb{K}[x_1, \ldots, x_n]$ *to be the Hilbert series of* $\mathbb{K}[x_0, x_1, \ldots, x_n]/(I^{\mathrm{h}} + \langle x_0 \rangle)$*, where* I^{h} *is the homogenisation of* I *by* x_0*.*

Definition 3.17 [27, Section 10.3]**.** *Let* $\boldsymbol{f} = (f_1, \ldots, f_p) \subset \mathbb{K}[x_0, x_1, \ldots, x_n]$ *be a sequence of homogeneous polynomials. We say that* $\boldsymbol{f}$ *is a* regular *sequence if for any* $1 \leq i \leq p$*,* f_i *is not a zero-divisor in* $\mathbb{K}[x_0, x_1, \ldots, x_n]/\langle f_1, \ldots, f_{i-1} \rangle$*. We say that any polynomial sequence is* regular *if the homogeneous parts of highest degree forms a regular sequence.*

Example 3.18. *The Hilbert series* H *of a regular sequence* $\boldsymbol{f} = (f_1, \ldots, f_p) \subset \mathbb{K}[x_1, \ldots, x_n]$ *with degrees* $\deg f_i = d_i$ *can be expressed as*

$$H(t) = \frac{\prod_{i=1}^{p}(1 - t^{d_i})}{(1-t)^n}.$$

Definition 3.19 [27, Page 425]**.** *Let* $I \subset \mathbb{K}[x_1, \ldots, x_n]$ *be an ideal. The* depth *of* I *is the maximal length of a regular sequence in* I*.*

Definition 3.20 [47, Page 6]**.** *Let* R *be a ring and let* $\mathfrak{p} \subset R$ *be a prime ideal. The* height *of* $\mathfrak{p}$ *is the supremum of all integers* s *such that there exists a chain of distinct prime ideals*

$$\mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \cdots \subsetneq \mathfrak{p}_s = \mathfrak{p}.$$

The Krull dimension *of* R*, denoted* $\dim R$*, is the supremum of all heights of all prime ideals of* R*.*

Example 3.21. *Consider the ideal* $I = \langle x - z^3, y - z^2 \rangle$*. Firstly, note that* I *is a prime ideal since the quotient algebra* $\mathbb{K}[x, y, z]/I \simeq \mathbb{K}[z]$ *is an integral domain. Therefore, the height of* I *is two, given by the chain of ideals:*

$$\langle 0 \rangle \subsetneq \langle x - z^3 \rangle \subsetneq \langle x - z^3, y - z^2 \rangle.$$

Consider the ideal $J = \langle (x-y)^2, x - z^3 \rangle$*. The generators of this ideal form a regular sequence, hence the depth of* J *is two.*

Example 3.22. *Consider the ring* $\mathbb{Z}$*. The prime ideals of* $\mathbb{Z}$ *are* $\langle 0 \rangle$ *and* $\langle \mathfrak{p} \rangle$ *for all prime numbers* $\mathfrak{p}$*. Clearly,* $\langle 0 \rangle$ *has zero height, while the chain* $\langle 0 \rangle \subsetneq \langle \mathfrak{p} \rangle$ *implies that* $\langle \mathfrak{p} \rangle$ *has height one for all primes* $\mathfrak{p}$*. Thus,* $\mathbb{Z}$ *has Krull dimension one. A field* $\mathbb{K}$ *only has one prime ideal,* $\langle 0 \rangle$*, and so has Krull dimension zero. The ring* $\mathbb{K}[x_1, \ldots, x_n]$ *has Krull dimension* n*, given by the maximal chain of prime ideals* $\langle 0 \rangle \subsetneq \langle x_1 \rangle \subsetneq \langle x_1, x_2 \rangle \subsetneq \cdots \subsetneq \langle x_1, \ldots, x_n \rangle$*.*

Definition 3.23 [27, Section 18.2]**.** *A ring* R *is called* Cohen-Macaulay *if for every maximal ideal* $\mathfrak{m} \subset R$ *the depth of* $\mathfrak{m}$ *equals the height of* $\mathfrak{m}$*.*

Definition 3.24 [65, Page 244]**.** *Let* $\mathbb{L}$ *be a finite extension of a field* $\mathbb{K}$*. If* $\mathbb{L} = \mathbb{K}(\alpha)$ *for some* $\alpha \in \mathbb{L}$ *then we say that* α *is a* primitive element *of* $\mathbb{L}$ *over* $\mathbb{K}$*.*

Example 3.25. *Consider the field extension* $\mathbb{Q}(\sqrt{2}, \sqrt{3})$*. Let* $\alpha = \sqrt{2} + \sqrt{3}$*. Clearly,* $\mathbb{Q}(\alpha) \subset \mathbb{Q}(\sqrt{2}, \sqrt{3})$*. Note that* $\frac{\alpha^2 - 5}{2} = \sqrt{6} \in \mathbb{Q}(\alpha)$ *implies that* $\sqrt{6}\alpha - 2\alpha = \sqrt{2} \in \mathbb{Q}(\alpha)$*. Hence,* $\sqrt{3} \in \mathbb{Q}(\alpha)$ *and so* $\mathbb{Q}(\sqrt{2}, \sqrt{3}) \subset \mathbb{Q}(\alpha)$*. Therefore,* α *is a primitive element of* $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ *over* $\mathbb{Q}$*.*

Example 3.26. *We remark that not every field extension has a primitive element. Consider the field* $\mathbb{K} = \mathbb{F}_p(x, y)$ *where* $\mathbb{F}_p$ *is the finite field with* p *elements. Let* $\mathbb{L}$ *be the field extension of* $\mathbb{K}$ *obtained by adjoining the* p*th roots of* x *and* y*,* $\mathbb{K}[w, z]/\langle x - w^p, y - z^p \rangle$*. Then,* $\mathbb{L}$ *is a finite extension of* $\mathbb{K}$ *of degree* p^2 *with basis* $(w^i z^j)_{0 \leq i,j \leq p-1}$*. Observe that for all* $\beta \in \mathbb{L}$*,* $\beta^p \in \mathbb{K}$*. Therefore, all elements of* $\mathbb{L}$ *have degree at most* p *and so* $\mathbb{L}$ *has no primitive element.*

3.2 Algebraic Geometry

In this section and for the remaining preliminaries, we consider the polynomial ring $\mathbb{K}[x_1, \ldots, x_n]$ where $\mathbb{K}$ is an algebraically closed field.

Definition 3.27 [24, Definition 1.2.1]**.** *Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[x_1, \ldots, x_n]$ be polynomials. Then, the* affine variety *defined by $\boldsymbol{f}$ is*

$$\mathbf{V}(\boldsymbol{f}) = \{(x_1, \ldots, x_n) \in \mathbb{K}^n \mid f_i(x_1, \ldots, x_n) = 0 \text{ for all } 1 \leq i \leq p\}.$$

Given a variety $V \subset \mathbb{K}^n$, the set $\mathbf{I}(V)$ defined by

$$\mathbf{I}(V) = \{f \in \mathbb{K}[x_1, \ldots, x_n] \mid f(x) = 0 \text{ for all } x \in V\}$$

is an ideal in $\mathbb{K}[x_1, \ldots, x_n]$.

Example 3.28. *Recall the ideal $\langle x - z^3, y - z^2 \rangle \subset \mathbb{K}[x, y, z]$ defined in Example 3.15. The variety $\mathfrak{C} = \mathbf{V}(x - z^3, y - z^2)$ is the twisted cubic in $\mathbb{K}^3$, parametrised by (t^3, t^2, t).*

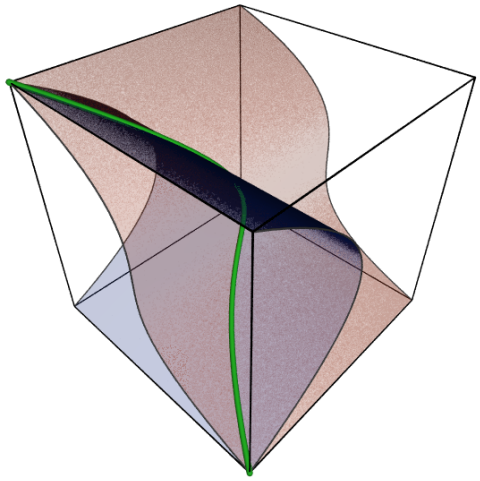


Figure 3.1 – The twisted cubic in the box [-1, 1], from [90].

Lemma 3.29 [24, Lemma 1.2.2, Section 4.3]**.** *Let $U = \mathbf{V}(f_1, \ldots, f_p)$ and $V = \mathbf{V}(g_1, \ldots, g_q)$ be varieties. Then,*

1. *their* intersection *is a variety with $U \cap V = \mathbf{V}(f_1, \ldots, f_p, g_1, \ldots, g_q)$,*
2. *their* union *is a variety with $U \cup V = \mathbf{V}(f_i g_j \mid 1 \leq i \leq p, 1 \leq j \leq q)$.*

Definition 3.30 [47, Page 10]**.** *The* Zariski topology *is defined by taking the closed sets to be varieties. The properties of a topology are easily verified by Lemma 3.29. It follows that finite unions and arbitrary intersections of varieties are also varieties. In this topology, a variety is called* Zariski-closed *and its complement is called* Zariski-open.

Definition 3.31 [24, Definition 4.4.2]**.** *Let $S \subset \mathbb{K}^n$. The* Zariski closure $\overline{S}$ *of S is $\mathbf{V}(\mathbf{I}(S))$, the smallest affine variety that contains S.*

Theorem 3.32 [24, Theorem 4.4.4]**.** *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal with variety $V = \mathbf{V}(I)$. Let $\pi_i : \mathbb{K}^n \to \mathbb{K}^{n-i}$ be the projection map onto the last $n - i$ coordinates. Then, for all $1 \leq i \leq n$, the variety defined by the ith elimination ideal $\mathbf{V}(I_i) = \overline{\pi_i(V)}$.*

Example 3.33. *Consider the subset of $\mathbb{C}$ without 0, $\mathbb{C}^* \subset \mathbb{C}$. The Zariski closure $\overline{\mathbb{C}^*}$ of $\mathbb{C}^*$ is $\mathbb{C}$. Note that $\mathbb{C}^*$ is the image of $\mathbf{V}(xy - 1) \subset \mathbb{C}^2$ under the projection map onto either coordinate. Thus, the projection of a variety is not necessarily a variety.*

Lemma 3.34 [24, Theorem 4.4.10]. *Let I, J be ideals of $\mathbb{K}[x_1, \ldots, x_n]$. Then,*

$$\overline{\mathbf{V}(I) \setminus \mathbf{V}(J)} = \mathbf{V}(I : J^\infty).$$

Example 3.35. *Let I be the ideal $\langle xy \rangle \subset \mathbb{K}[x, y]$. Then, the variety $\mathbf{V}(I)$ is the union of the x and y axes. Consider the saturation by the ideal $J = \langle x \rangle$, $\langle xy \rangle : \langle x \rangle^\infty = \langle y \rangle$ which defines the x-axis. Indeed we have that $\overline{\mathbf{V}(I) \setminus \mathbf{V}(J)} = \mathbf{V}(I : J^\infty)$.*

Theorem 3.36 [24, Theorem 4.1.2]. *Let V be a variety and let I be an ideal. Then,*

$$\mathbf{V}(\mathbf{I}(V)) = V \quad \text{and} \quad \mathbf{I}(\mathbf{V}(I)) = \sqrt{I}.$$

In other words, there is a one-to-one correspondence between varieties and radical ideals.

Definition 3.37. *Let V be a variety. A property P of V is said to hold* generically *if there exists a non-empty Zariski-open subset U of V such that P holds for U.*

Example 3.38. *Let $V = \mathbb{K}^{2\times 2}$ be the set of 2×2 matrices with entries in $\mathbb{K}$. An LU decomposition of a matrix $A = (a_{i,j})_{1\leq i,j\leq 2} \in V$ is a lower triangular matrix L and an upper triangular matrix U such that $A = LU$. Such a decomposition exists, which we call property P, if and only if A is invertible and $a_{1,1} \neq 0$. Hence, P holds on the non-empty Zariski-open complement U of $\mathbf{V}(a_{1,1}) \cup \mathbf{V}(a_{1,1}a_{2,2} - a_{1,2}a_{2,1}) = \mathbf{V}(a_{1,1}(a_{1,1}a_{2,2} - a_{1,2}a_{2,1}))$ and so P holds generically.*

Definition 3.39 [47, Proposition 1.7]. *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal with quotient algebra A. Then, the affine variety $V = \mathbf{V}(I)$ has* dimension *equal to the Krull dimension of the algebra A. This dimension is equal to the dimension of V as a topological space in the Zariski topology.*

Example 3.40. *The dimension of a variety V is equal to the minimal number m of* generic hyperplanes *H_i needed so that $V \cap H_1 \cap \cdots \cap H_m$ is a finite, non-zero number of points. For a trivial example, take the ideal $\langle 0 \rangle \in \mathbb{K}[x_1, \ldots, x_n]$. Clearly, $\mathbf{V}(0) = \mathbb{K}^n$. Thus, n hyperplanes are required for the intersection to yield a finite, non-zero number of points and so we expect the dimension to be n. Indeed, as in Example 3.22, the quotient algebra $\mathbb{K}[x_1, \ldots, x_n]/\langle 0 \rangle = \mathbb{K}[x_1, \ldots, x_n]$ has Krull dimension n.*

Definition 3.41 [24, Definition 8.2.1]. *Let V be a vector space over a field $\mathbb{K}$. Let $\sim$ be the equivalence relation on $V \setminus \{0\}$ defined by $v_1 \sim v_2$ if there exists some non-zero $\lambda \in \mathbb{K}$ such that $v_1 = \lambda v_2$. The* projective space *$\mathbb{P}(V)$ is the set of equivalence classes of $\sim$ on $V \setminus \{0\}$. When $V = \mathbb{K}^{n+1}$ for some $n \in \mathbb{N}$, we denote $\mathbb{P}(V)$ by $\mathbb{P}^n$.*

Definition 3.42. *Let $V \subset \mathbb{K}^n$ be a variety of dimension m. The* degree *of the variety V is equal to the number of points in the intersection of V with m generic hyperplanes.*

Example 3.43. *Let $f \in \mathbb{K}[x_1, \ldots, x_n]$ be a polynomial. Consider the hypersurface defined by f, $V = \mathbf{V}(f)$. Then, since V has dimension $n-1$, we intersect with $n-1$ general hyperplanes.*

$$V_H = V \cap H_1 \cap \cdots \cap H_{n-1} = \mathbf{V}(f, a_{1,0} + a_{1,1}x_1 + \cdots + a_{1,n}x_n, \ldots, a_{n-1,0} + a_{1,1}x_1 + \cdots + a_{n-1,n}x_n).$$

The coefficients of these hyperplanes defines an $(n-1) \times (n+1)$ matrix with coefficients in $\mathbb{K}$.

$$\begin{bmatrix} a_{1,0} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n-1,0} & \cdots & a_{n-1,n} \end{bmatrix}.$$

Performing Gaussian elimination on this matrix allows one to write V_H as

$$\mathbf{V}(f, x_1 - b_1x_n - c_1, \ldots, x_{n-1} - b_{n-1}x_n - c_{n-1}),$$

where $b_i, c_i \in \mathbb{K}$ for $1 \leq i \leq n-1$. Thus, one can now rewrite f as a univariate polynomial in x_n with the same degree. The roots of the resulting univariate polynomial then give the last coordinate of the points in the intersection from which one can recover the remaining $n-1$ coordinates. Therefore, the degree of V is equal to the degree of f.

Example 3.44. *Let $\mathfrak{C}$ be the twisted cubic $\mathbf{V}(x-z^3, y-z^2) \subset \mathbb{K}^3$. Note that $\mathfrak{C}$ has dimension one so we intersect $\mathfrak{C}$ with a sufficiently generic hyperplane of $\mathbb{K}^3$ to define a zero-dimensional variety. Thus, for the hyperplane $H = \mathbf{V}(x-3y+6z-8)$,*

$$\mathfrak{C}_H = \mathfrak{C} \cap H = \mathbf{V}(x-z^3, y-z^2, x-3y-6z+8) = \mathbf{V}(x-z^3, y-z^2, z^3-3z^2-6z+8).$$

Solving the cubic equation $z^3-3z^2-6z+8=0$ and substituting the resulting values to find x and y gives three points in the intersection, $(1,1,1), (-8,4,-2), (64,16,4)$, and therefore the twisted cubic $\mathfrak{C}$ has degree 3.

Definition 3.45 [24, Definition 8.2.5]**.** *Let $I \subset \mathbb{K}[x_0, \ldots, x_n]$ be a homogeneous ideal. Then the variety $\mathbf{V}(I)$ is a subset of a projective space $\mathbb{P}^n$ over $\mathbb{K}$ and is called a* projective *variety.*

Definition 3.46 [24, Definition 8.4.6]**.** *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal with corresponding variety $V = \mathbf{V}(I)$. The* projectivisation *of V is a projective variety V^{h} given by $V^{\mathrm{h}} = \mathbf{V}(I^{\mathrm{h}})$.*

Proposition 3.47. *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal with variety $V = \mathbf{V}(I)$. Consider the Hilbert series H of I,*

$$H(t) = \frac{h(t)}{(1-t)^m},$$

where $h(t)$ is a univariate polynomial with coefficients in $\mathbb{Z}$ that is not divisible by $1-t$, in other words $H(t)$ is a reduced fraction. Then, the dimension of the variety V is equal to m and the degree of V is equal to $h(1)$.

Example 3.48. *Recall that in Example 3.18, the Hilbert series H of a regular sequence $\boldsymbol{f} = (f_1, \ldots, f_p) \subset \mathbb{K}[x_1, \ldots, x_n]$ with degrees $\deg f_i = d_i$ is expressed as*

$$H(t) = \frac{\prod_{i=1}^{p}(1-t^{d_i})}{(1-t)^n}.$$

Then, reducing the fraction gives the following form of H:

$$H(t) = \frac{\prod_{i=1}^{p}(1+t+\cdots+t^{d_i-1})}{(1-t)^{n-p}}.$$

Hence, the variety $\mathbf{V}(\boldsymbol{f})$ has dimension $n-p$ and has degree $d_1 \cdots d_p$. Indeed, this agrees with the special case $p=1$, the hypersurface case, as in Example 3.43.

Definition 3.49 [24, Definition 4.5.1]**.** *A variety $V \subset \mathbb{K}^n$ is* irreducible *if whenever V is expressed as the union of two varieties $V = V_1 \cup V_2$, either $V = V_1$ or $V = V_2$.*

Theorem 3.50 [24, Theorem 4.6.2]**.** *Let $V \subset \mathbb{K}^n$ be a variety. Then, V can be expressed as a finite union $V = V_1 \cup \cdots \cup V_r$ where $V_1, \ldots, V_r$ are irreducible varieties in $\mathbb{K}^n$.*

Definition 3.51 [42, Definition 1]**.** *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal with finite variety $V = \mathbf{V}(I)$ of degree δ. A linear form $u = \lambda_1 x_1 + \cdots + \lambda_n x_n$ where $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ is called a* primitive element *if the powers $1, u, \ldots, u^{\delta-1}$ form a basis of the quotient algebra $\mathbb{K}[x_1, \ldots, x_n]/I$.*

Example 3.52. *Let $\pi_i : \mathbb{K}^n \to \mathbb{K}$ be the projection map onto the ith coordinate. Let $V = \{v_1, \ldots, v_D\} \in \mathbb{K}^n$ be a set of points such that $\pi_n(v_i) \neq \pi_n(v_j)$ for all $i \neq j$. Define the degree D polynomial $g_n = \prod_{i=1}^{n}(x_n - \pi_n(v_i))$ so that g_n vanishes on V. Furthermore, by interpolation, for each $1 \leq i \leq n-1$, the set $\{(\pi_n(v_1), \pi_i(v_1)), \ldots, (\pi_n(v_D), \pi_i(v_D))\}$ defines a unique polynomial $g_i \in \mathbb{K}[x_n]$ of degree at most $D-1$ such that $g_i(\pi_n(v_j)) = \pi_i(v_j)$. Hence, $x_i - g_i(x_n)$ vanishes on V as well and we have that*

$$\{x_1 - g_1(x_n), \ldots, x_{n-1} - g_{n-1}(x_n), g_n(x_n)\} \subseteq \mathbf{I}(V)$$

3.3 Gröbner bases

As we saw in Subsection 2.1.2, polynomial system solving is key to solving polynomial optimisation problems exactly using the generalised critical value method. There are many techniques available to solve polynomial systems. However, homotopy deformation techniques can lose solutions if a path passes through an ill-conditioned system. This can lead to an incorrect infimum when applied to POP if a generalised critical value is missed. Therefore, for their practical use, we focus on Gröbner bases in this thesis. In particular, we study Gröbner basis computation with the aim of solving Problems 1 and 2.

3.3.1 Using Gröbner bases

Definition 3.53 [24, Definition 2.2.1]**.** *A* monomial ordering $\prec$ *on* $\mathbb{K}[x_1, \ldots, x_n]$ *is a relation on* $\mathbb{Z}^n_{\geq 0}$ *satisfying the following:*

1. *For all* $\alpha, \beta \in \mathbb{Z}^n_{\geq 0}$*, exactly one of the following holds:*
$$\alpha \prec \beta, \quad \beta \prec \alpha \text{ or } \alpha = \beta.$$
2. *For all* $\alpha, \beta, \gamma \in \mathbb{Z}^n_{\geq 0}$*, if* $\alpha \prec \beta$*, then* $\alpha + \gamma \prec \beta + \gamma$*.*
3. *For every non-empty subset* $S \subset \mathbb{Z}^n_{\geq 0}$ *there exists some* $\alpha \in S$ *such that* $\alpha \prec \beta$ *for all* $\beta \in S \setminus \{\alpha\}$*.*

Definition 3.54 [24, Definitions 2.2.3 and 2.2.5]**.** *With the convention that* $x_n \prec \cdots \prec x_1$ *and for all* $\alpha, \beta \in \mathbb{Z}^n_{\geq 0}$*, define the following orders called the* lexicographic *(LEX) and* degree reverse lexicographic *(DRL) orderings respectively:*

- $\alpha \prec_{LEX} \beta$ *if and only if there exists* $1 \leq j \leq n$ *such that for all* $i < j$*,* $\alpha_i = \beta_i$ *and* $\alpha_j < \beta_j$*.*
- $\alpha \prec_{DRL} \beta$ *if and only if* $\alpha_1 + \cdots + \alpha_n < \beta_1 + \cdots + \beta_n$ *or if* $\alpha_1 + \cdots + \alpha_n = \beta_1 + \cdots + \beta_n$ *and there exists* $2 \leq j \leq n$ *such that for all* $i > j$*,* $\alpha_i = \beta_i$ *and* $\alpha_j > \beta_j$*.*

Definition 3.55 [24, Definition 2.2.7]**.** *Fix a monomial ordering* $\prec$ *on* $\mathbb{K}[x_1, \ldots, x_n]$*. The* leading monomial *of a polynomial* $f \in \mathbb{K}[x_1, \ldots, x_n]$*, denoted* $\mathrm{LM}_\prec(f)$*, is the largest monomial in* f *with respect to* $\prec$*. Similarly,* $\mathrm{LC}_\prec(f)$ *denotes the* leading coefficient *of* f*, the coefficient of* $\mathrm{LM}_\prec(f)$*, and* $\mathrm{LT}_\prec(f)$ *denotes the* leading term *of* f*, that is* $\mathrm{LT}_\prec(f) = \mathrm{LC}_\prec(f)\,\mathrm{LM}_\prec(f)$*.*

Definition 3.56 [24, Definition 2.5.1]**.** *Let* I *be an ideal of* $\mathbb{K}[x_1, \ldots, x_n]$ *with monomial ordering* $\prec$*. The* initial ideal *of* I *with respect to* $\prec$*, denoted* $\mathrm{LM}_\prec(I)$*, is defined by*
$$\mathrm{LM}_\prec(I) = \langle \mathrm{LM}_\prec(f) \mid f \in I \rangle.$$

Definition 3.57 [24, Definition 2.5.5]**.** *Let* $I \subset \mathbb{K}[x_1, \ldots, x_n]$ *be a non-zero ideal and let* $\prec$ *be a monomial ordering. A finite set* $G = \{g_1, \ldots, g_r\} \subset I$ *is a* Gröbner basis *of* I *with respect to* $\prec$ *if* $\langle \mathrm{LM}_\prec(g_1), \ldots, \mathrm{LM}_\prec(g_r) \rangle = \mathrm{LM}_\prec(I)$*.*

Example 3.58. *Consider the ideal* $I = \langle x - z^3, y - z^2, x - 3y - 6z + 8 \rangle$ *defining* $\mathfrak{C}_H$*, the twisted cubic intersected with a plane, as in Example 3.44. With* $x \succ y \succ z$*, let* $\prec_{DRL}$ *and* $\prec_{LEX}$ *be the LEX and DRL orders respectively. Then, the Gröbner basis* $\mathcal{G}_{DRL}$ *of* I *with respect to* $\prec_{DRL}$ *is*
$$\mathcal{G}_{DRL} = \{x - 3y - 6z + 8, z^2 - y, yz - 3y - 6z + 8, y^2 - 15y - 10z + 24\}.$$
Moreover, the Gröbner basis $\mathcal{G}_{LEX}$ *of* I *with respect to* $\prec_{LEX}$ *is*
$$\mathcal{G}_{LEX} = \{x - 3z^2 - 6z + 8, y - z^2, z^3 - 3z^2 - 6z + 8\}.$$
Note that z *is a primitive element of* $\mathbb{K}[x, y, z]/I$*.*

Proposition 3.59 [24, Proposition 2.6.1]. *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal with a Gröbner basis $G = \{g_1, \ldots, g_m\}$ with respect to a monomial ordering $\prec$. Let $f \in \mathbb{K}[x_1, \ldots, x_n]$ be a polynomial. Then, there exists a unique polynomial $r \in \mathbb{K}[x_1, \ldots, x_n]$, called the* normal form *of f with respect to G and denoted by $\mathrm{NF}(f, G, \prec)$, such that no term of r is divisible by any of $\mathrm{LT}_\prec(g_1), \ldots, \mathrm{LT}_\prec(g_m)$ and there exists $h \in I$ such that $f = h + r$.*

Definition 3.60 [24, Definition 2.7.4]. *Let G be a Gröbner basis of an ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$. We say that G is* reduced *if for all $f \in G$, $\mathrm{LC}_\prec(f) = 1$ and for all $g \in G \setminus \{f\}$, every monomial of g is not divisible by $\mathrm{LM}_\prec(f)$.*

Theorem 3.61 [24, Theorem 2.7.5]. *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be a non-zero ideal. Then, given a monomial ordering $\prec$, I has a unique reduced Gröbner basis with respect to $\prec$.*

Lemma 3.62 [24, Theorem 3.1.2]. *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal with a Gröbner basis G with respect to a lexicographic term ordering $\prec_{LEX}$ where $x_n \prec \cdots \prec x_1$. Then, for all $0 \leq i \leq n$, the set $G_i = G \cap \mathbb{K}[x_{i+1}, \ldots, x_n]$ forms a Gröbner basis of I_i, the ith elimination ideal of I.*

Lemma 3.63 [24, Theorem 4.3.11]. *Let $I = \langle f_1, \ldots, f_p \rangle, J = \langle g_1, \ldots, g_q \rangle \subset \mathbb{K}[x_1, \ldots, x_n]$ be ideals. Then, for an indeterminate y, the ideal $I \cap J = (yI + (1-y)J) \cap \mathbb{K}[x_1, \ldots, x_n]$. Therefore, if G is a Gröbner basis of $\langle yf_1, \ldots, yf_p, (1-y)g_1, \ldots, (1-y)g_q \rangle$ with respect to a lexicographic ordering in which y is the greatest variable, then $G \cap \mathbb{K}[x_1, \ldots, x_n]$ is a Gröbner basis of $I \cap J$.*

Lemma 3.64 [24, Theorem 4.4.14]. *Let $I = \langle f_1, \ldots, f_p \rangle \subset \mathbb{K}[x_1, \ldots, x_n]$ be an ideal and let $g \in \mathbb{K}[x_1, \ldots, x_n]$ be a polynomial. Then, for a new variable y,*

$$I : \langle g \rangle^\infty = \langle f_1, \ldots, f_p, 1 - yg \rangle \cap \mathbb{K}[x_1, \ldots, x_n].$$

Therefore, if G is a Gröbner basis of $\langle f_1, \ldots, f_p, 1 - yg \rangle$ with respect to a lexicographic ordering in which y is the greatest variable, then $G \cap \mathbb{K}[x_1, \ldots, x_n]$ is a Gröbner basis of $I : \langle g \rangle^\infty$.

Definition 3.65. *A set of monomials S is a* staircase *if for all monomials m_1, m_2, if $m_1 m_2 \in S$ then $m_1 \in S$ and $m_2 \in S$.*

Example 3.66. *Consider an ideal $I \in \mathbb{K}[x_1, \ldots, x_n]$ with Gröbner basis G with respect to an ordering $\prec$. Let S be the set of monomials such that $m \in S$ if and only if $\mathrm{LM}_\prec(g) \nmid m$ for all $g \in G$. Suppose that $m_1 m_2 \in S$ for some monomials m_1, m_2 so that $\mathrm{LM}_\prec(g) \nmid m_1 m_2$ for all $g \in G$. Then, $\mathrm{LM}_\prec(g) \nmid m_1$ and $\mathrm{LM}_\prec(g) \nmid m_2$ for all $g \in G$ and so S is a staircase. Note that S does not depend on the choice of G. Hence, we call S the staircase associated to $\prec$.*

Example 3.67 [24, Exercise 9.4.11]. *Let I be an ideal with Gröbner basis G for a monomial ordering $\prec$. Let S be the staircase associated to $\prec$. The dimension and degree of the ideal I are closely linked to the staircase S. In particular, the staircase S forms a natural basis of the quotient algebra $\mathbb{K}[x_1, \ldots, x_n]/I$. When I is zero-dimensional, that is when $\mathbf{V}(I)$ is finite, the dimension of this quotient algebra is equal to the degree of I. Hence, $|S|$ is finite and is equal to the degree of I, that is the number of points in $\mathbf{V}(I)$. Alternatively, when I has positive dimension, the staircase S is infinite.*

3.3.2 Computing Gröbner bases

A major focus of this thesis is the complexity estimates of polynomial optimisation. Thus, we introduce a few standard definitions in complexity theory.

Definition 3.68 [107, Definition 25.7]. *For two univariate functions $f : \mathbb{N} \to \mathbb{R}$ and $g : \mathbb{N} \to \mathbb{R}$, we say that $f = O(g)$ if there exists $N, M \in \mathbb{N}$ such that for all $n \geq N$ we have $\|f(n)\| \leq M\|g(n)\|$.*

Lemma 3.69 [17]. *Denote by $\mathsf{M}(d)$ the number of arithmetic operations in the base field required to multiply two univariate polynomials of degree at most d. Through the Cantor-Kaltofen algorithm, $\mathsf{M}(d)$ is at most $O(d \log d \log \log d)$.*

Algorithm 1: FGLM

Input: $\mathcal{G}_{\prec_1}$ a Gröbner basis of a *zero-dimensional* ideal $I \in \mathbb{K}[x_1, \ldots, x_n]$ with respect to the monomial ordering $\prec_1$ and $\prec_2$, the desired term order.

Output: $\mathcal{G}_{\prec_1}$, a Gröbner basis of I with respect to $\prec_2$.

$S_2 \leftarrow \{1\}$.
$\mathcal{G}_{\prec_2} \leftarrow \emptyset$.
$l \leftarrow \{x_n, \ldots, x_1\}$.
While $l \neq \emptyset$ **do**
  $m \leftarrow$ first entry of the list l.
  If *there exists a linear combination:* $\mathrm{NF}(m + \sum_{\mu \in S_2} c_\mu \mu, \mathcal{G}_{\prec_1}, \prec_1) = 0$ **then**
    $\mathcal{G}_{\prec_2} \leftarrow \mathcal{G}_{\prec_2} \cup \{m + \sum_{\mu \in S_2} c_\mu \mu\}$.
    Remove from l all monomials that are divisible by m.
  Else
    $S_2 \leftarrow S_2 \cup \{m\}$.
    Add to l the monomials $x_1 \cdot m, \ldots, x_n \cdot m$ except those that are divisible by a leading monomial of $\mathcal{G}_{\prec_2}$.
    Sort l with respect to $\prec_2$. Remove m from l.
Return $\mathcal{G}_{\prec_2}$.

Remark 3.70. *The FGLM algorithm, presented as Algorithm 1, necessarily terminates as the ideal I is zero-dimensional and therefore the constructed staircase S_2 is finite. It is shown in [31] that for an ideal $I \in \mathbb{K}[x_1, \ldots, x_n]$ of degree D, the FGLM algorithm has complexity $O(nD^3)$.*

Example 3.71. *Recall the ideal $I = \langle x - z^3, y - z^2, x - 3y - 6z + 8 \rangle$ in Example 3.58. We consider its DRL Gröbner basis with $x \succ y \succ z$,*

$$\mathcal{G}_{DRL} = \{x - 3y - 6z + 8, z^2 - y, yz - 3y - 6z + 8, y^2 - 15y - 10z + 24\}.$$

We will use the FGLM algorithm to compute the LEX Gröbner basis. The leading monomials of $\mathcal{G}_{DRL}$ are $\{x, z^2, yz, y^2\}$. Thus, the staircase S_{DRL} corresponding to $\prec_{DRL}$ is $\{1, y, z\}$. We will find the staircase S_{LEX} corresponding to $\prec_{LEX}$ and in doing so will discover the linear combinations of monomials that lie inside I. We begin by initialising our LEX staircase, $S_2 = \{1\}$, and our monomial list, $l = \{z, y, x\}$.

1. $\mathrm{NF}(z, \mathcal{G}_{DRL}, \prec_{DRL}) = z \implies z \in S_{LEX}$ *and* $l = \{z^2, y, yz, x, xz\}$.
2. $\mathrm{NF}(z^2, \mathcal{G}_{DRL}, \prec_{DRL}) = y \implies z^2 \in S_{LEX}$ *and* $l = \{z^3, y, yz, yz^2, x, xz, xz^2\}$.
3. $\mathrm{NF}(z^3 - 3z^2 + 6z - 8, \mathcal{G}_{DRL}, \prec_{DRL}) = 0 \implies z^3 - 3z^2 - 6z + 8 \in \mathcal{G}_{LEX}$ *and* $l = \{y, yz, yz^2, x, xz, xz^2\}$.
4. $\mathrm{NF}(y - z^2, \mathcal{G}_{DRL}, \prec_{DRL}) = 0 \implies y - z^2 \in \mathcal{G}_{LEX}$ *and* $l = \{x, xz, xz^2\}$.
5. $\mathrm{NF}(x - 3z^2 + 6z - 8, \mathcal{G}_{DRL}, \prec_{DRL}) = 0 \implies x - 3z^2 + 6z - 8 \in \mathcal{G}_{LEX}$ *and* $l = \{\emptyset\}$.

Hence, $S_{LEX} = \{1, z, z^2\}$ and $\mathcal{G}_{LEX}$ is as given in Example 3.58.

Definition 3.72. *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be a zero-dimensional ideal of degree D and let $\prec_{LEX}$ be a lexicographic term ordering with x_n as the least variable. We say that I is in* shape position *if the reduced Gröbner basis of I with respect to the ordering $\prec_{LEX}$ has the form*

$$\{x_1 - g_1(x_n), \ldots, x_{n-1} - g_{n-1}(x_n), g_n(x_n)\},$$

where $g_1, \ldots, g_{n-1}$ have degree at most $D - 1$ and g_n has degree D.

Lemma 3.73 [8, Proposition 5]**.** *Let $I \subset \mathbb{K}[x_1, \ldots, x_n]$ be a zero-dimensional radical ideal of degree D with and let $\mathbf{V}(I) = \{v_1, \ldots, v_D\} \subset \mathbb{K}^n$. Let π_n be the projection map from $\mathbb{K}^n$ onto the last coordinate. Then, I is in shape position if and only if $\pi_n(v_i) \neq \pi_n(v_j)$ for all $i \neq j$. Moreover, when $\mathbb{K}$ has characteristic zero, I is in shape position after applying a generic linear change of coordinates.*

Remark 3.74. *Suppose that the base field $\mathbb{K}$ has characteristic zero. Given a zero-dimensional radical ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$, one can define an ideal $J \subset \mathbb{K}[x_1, \ldots, x_n, u]$ by $J = I + \langle u - \sum_{i=1}^n \lambda_i x_i \rangle$, where $\lambda_i \in \mathbb{K}$ are generic, such that J is in shape position. This amounts to defining a new variable u to act as a primitive element. Thus, one can interpret the Shape Lemma 3.73 as saying that after a sufficiently general linear change of coordinates, x_n is a primitive element.*

When an ideal I is in shape position we have significant knowledge on the LEX basis of I before we begin any computations. Hence, it is natural to try to use this information to speed up the change of ordering from DRL to LEX. Indeed, this structure implies that the polynomial g_n is exactly the characteristic polynomial of the matrix T_{x_n} associated to multiplication by x_n in the quotient algebra $\mathbb{K}[x_1, \ldots, x_n]/I$. Thus, the polynomial g_n can be constructed efficiently using the Wiedemann algorithm. Then, for each $1 \leq i \leq n-1$, $x_i = g_i$ in the quotient algebra $\mathbb{K}[x_1, \ldots, x_n]/I$. Hence, $x_i \cdot x_n = g_i \cdot x_n$ can be expressed as x_n^D plus some smaller degree terms in x_n. By repeated multiplications by x_n, one constructs a Hankel system and can solve it to find g_i. As we will see, this methodology allows us to do better than the FGLM complexity of $O(nD^3)$ in the shape position case.

Definition 3.75. *Consider a zero-dimensional ideal $I \in \mathbb{K}[x_1, \ldots, x_n]$ of degree D. Let $\mathcal{G}$ be a Gröbner basis of I with respect to an ordering $\prec$ with associated staircase S. For each $1 \leq i \leq n$, the map of multiplication by x_i,*

$$\mathbb{K}[x_1, \ldots, x_n]/I \to \mathbb{K}[x_1, \ldots, x_n]/I \tag{3.1}$$

$$m \mapsto \mathrm{NF}(x_n \cdot m, \mathcal{G}, \prec) \tag{3.2}$$

can be represented by a matrix $T_{x_i} \in \mathbb{K}^{D \times D}$. For a monomial $m \in S$, the corresponding column of T_{x_i} falls into one of three cases:

1. *If $x_i \cdot m \in S$, then the column is a column of the identity matrix.*
2. *If $x_i \cdot m = \mathrm{LM}(g)$ for some $g \in \mathcal{G}$, then the column is given by $x_i m - g$.*
3. *Otherwise, the normal form $\mathrm{NF}(x_i \cdot m, \mathcal{G}, \prec)$ must be computed.*

When $i = n$, the paramount case when I is in shape position, the number of columns that fall into the latter two cases is denoted q.

Remark 3.76. *The polynomial g_n is the minimal polynomial of the matrix T_{x_n}. This is computed using the Wiedemann algorithm which computes the linear recurrence relation of the scalar sequence $r^t T_{x_n}^i s$, where r and s are generic. In this setting, we take $s = (1, 0, \ldots, 0)^t$ for the computation of g_n and for the computation of $g_1, \ldots, g_{n-1}$ we take $s = \mathbf{c}_1$ to $s = \mathbf{c}_{n-1}$, where $\mathbf{c}_j$ is the coefficient vector of $\mathrm{NF}(x_j, \mathcal{G}_{DRL}, \prec_{DRL})$, to construct a Hankel system. Note that in many cases $\mathbf{c}_1, \ldots, \mathbf{c}_{n-1}$ will be columns of the identity matrix. Since I is in shape position, g_n has degree D and so we need $2D$ terms to recover the recurrence using the Berlekamp–Massey algorithm [111]. By the structure imposed on T_{x_n} by Definition 3.75, each multiplication requires $O(qD)$ operations where q is the number of dense columns. The other polynomials, $g_1, \ldots, g_{n-1}$ can be computed through Hankel system solving in $O(\mathsf{M}(D)(n + \log D))$ base field operations, see [15] for more details. Therefore, the overall complexity of the shape position case of Sparse-FGLM is in the class $O(qD^2 + \mathsf{M}(D)(n + \log D))$.*

Algorithm 2: Sparse-FGLM

Input: $\mathcal{G}_{DRL}$ the DRL Gröbner basis of a *zero-dimensional* ideal $I \in \mathbb{K}[x_1, \dots, x_n]$ of degree D in shape position with respect to the monomial ordering $\prec_{DRL}$ and $\prec_{LEX}$ with $x_1 \succ \cdots \succ x_n$.

Output: $\mathcal{G}_{LEX}$, the Gröbner basis of I with respect to $\prec_{LEX}$.

$T_{x_n} \leftarrow$ built as in Definition 3.75.
$r \leftarrow$ random vector in $\mathbb{K}^D$.
$\mathbf{1} \leftarrow (1, 0, \dots, 0)^t$.
For i **from** 1 **to** $2D-1$ **do**
    Compute $r^t T_{x_n}^i \mathbf{1}$.
For j **from** 1 **to** $n-1$ **do**
    $\mathbf{c}_j \leftarrow$ coefficient vector of $\mathrm{NF}(x_j, \mathcal{G}_{DRL}, \prec_{DRL})$.
    For i **from** 1 **to** $D-1$ **do**
        Compute $r^t T_{x_n}^i \mathbf{c}_j$.
$g_n \leftarrow$ recovered by Wiedemann algorithm.
$g_1, \dots, g_{n-1} \leftarrow$ recovered through Hankel system solving.
Return $\{x_1 - g_1, \dots, x_{n-1} - g_{n-1}, g_n\}$.

Example 3.77. *Consider the ideal* $I = \langle x^2 + 3xy + 2x, y^4 + x^2y^2 + x^2 + 2x + 1\rangle \subset \mathbb{K}[x, y]$. *We start with the reduced Gröbner basis* $\mathcal{G}_{DRL}$ *of* I *with respect to the ordering* $\prec_{DRL}$ *and aim to compute a Gröbner basis* $\mathcal{G}_{LEX}$ *with respect to* $\prec_{LEX}$, *both with* $x \succ y$. *We have that*

$$\mathcal{G}_{DRL} = \Big\{ x^2 + 3yx + 2x, \\ xy^3 - \frac{1}{3}y^4 + \frac{2}{3}xy^2 + yx - \frac{1}{3}, \\ y^5 + \frac{8}{15}y^4 - \frac{1}{6}xy^2 + \frac{1}{5}yx + \frac{3}{10}x + y + \frac{8}{15} \Big\},$$

and begin by computing the multiplication matrix T_y. *Since the leading monomials of* $\mathcal{G}_{DRL}$ *are* $\{x^2, xy^3, y^5\}$, *the staircase* S_1 *associated to* $\prec_{DRL}$ *is* $\{1, y, x, y^2, xy, y^3, xy^2, y^4\}$. *Thus,* $T_y \in \mathbb{K}^{8\times 8}$ *and through the procedure laid out in Definition 3.75, we have that*

$$T_y = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & -\frac{8}{15} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{10} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & -\frac{1}{5} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -\frac{2}{3} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} & -\frac{8}{15} \end{bmatrix}.$$

Note that T_y *is sparse and that no normal forms needed to be computed as all of its columns fell into the first two cases of Definition 3.75. In this case, the number of dense columns* q *is two. Then, we compute the polynomial* $g_2 = y^8 + \frac{6}{5}y^7 + \frac{13}{10}y^6 + \frac{3}{5}y^5 + \frac{11}{10}y^4 + \frac{6}{5}y^3 + \frac{13}{10}y^2 + \frac{3}{5}y + \frac{1}{10}$ *using the Wiedemann algorithm. Since* g_2 *has degree* 8, *we confirm that* I *is in shape position. Next, for this system we find that* c_1 *is* $(0, 0, 1, 0, 0, 0, 0, 0)$. *Then, we compute the polynomial*

$g_1 = -\frac{15}{7}y^7 - \frac{8}{7}y^6 - \frac{45}{14}y^5 - 2y^4 - \frac{15}{7}y^3 - \frac{8}{7}y^2 - \frac{45}{14}y - 2$. *Finally, we output the LEX Gröbner basis*

$$\mathcal{G}_{LEX} = \left\{ x - \left(-\frac{15}{7}y^7 - \frac{8}{7}y^6 - \frac{45}{14}y^5 - 2y^4 - \frac{15}{7}y^3 - \frac{8}{7}y^2 - \frac{45}{14}y - 2 \right), \right.$$
$$\left. y^8 + \frac{6}{5}y^7 + \frac{13}{10}y^6 + \frac{3}{5}y^5 + \frac{11}{10}y^4 + \frac{6}{5}y^3 + \frac{13}{10}y^2 + \frac{3}{5}y + \frac{1}{10} \right\}.$$

3.4 Polynomial Optimisation

Definition 3.78. *A polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ is positive (resp. non-negative) on a set $S \subseteq \mathbb{R}^n$ if for all $x \in S$ we have $f(x) > 0$ (resp. $f(x) \geq 0$).*

Definition 3.79. *Let $(f_1, \ldots, f_s) \subset \mathbb{K}[x_1, \ldots, x_n]$. The* Jacobian matrix *of $(f_1, \ldots, f_s)$ is*

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_s}{\partial x_1} & \cdots & \frac{\partial f_s}{\partial x_n} \end{bmatrix}.$$

Definition 3.80 [6, Definition 5.55]**.** *Let $n, m, p \in \mathbb{Z}_{\geq 0}$ with $n \geq m + p$. Let $X = \mathbf{V}(g_1, \ldots, g_m)$ be a smooth variety where $g_1, \ldots, g_m \in \mathbb{K}[x_1, \ldots, x_n]$ is a reduced regular sequence. Consider the polynomial mapping $\boldsymbol{f}$ defined by*

$$\boldsymbol{f} : x \in X \mapsto (f_1(x), \ldots, f_p(x)),$$

where $f_1, \ldots, f_p \in \mathbb{K}[x_1, \ldots, x_n]$. Let J be the Jacobian matrix of $(f_1, \ldots, f_p, g_1, \ldots, g_m)$. We say that $x \in X$ is a critical point *if the evaluation of the Jacobian at x, $J(x)$, has rank less than $m + p$. A* critical value *of $\boldsymbol{f}$ is the image of a critical point x under $\boldsymbol{f}$.*

Theorem 3.81 [27, Theorem 16.19]**.** *Let $f, g_1, \ldots, g_m \in \mathbb{K}[x_1, \ldots, x_n]$ where $(g_1, \ldots, g_m)$ is a reduced regular sequence defining a smooth variety $X = \mathbf{V}(g_1, \ldots, g_m)$ of dimension $n - m$. Let J be the ideal defined by the $m \times m$ minors of the Jacobian matrix of $(f, g_1, \ldots, g_m)$. Then, $\mathbf{V}(J)$ is the set of critical points of the polynomial map*

$$f : x \in X \to \mathbb{K}.$$

Theorem 3.82 [96, Proposition B2]**.** *Let $X \subset \mathbb{K}^n$ be a variety and let $\boldsymbol{f} : X \to \mathbb{R}^m$ be a polynomial mapping. Then, the critical values of $\boldsymbol{f}$ are contained in an algebraic subset of $\mathbb{R}^m$ of dimension at most $m - 1$.*

Definition 3.83. *A polynomial $f \in \mathbb{K}[x_1, \ldots, x_n]$ of degree $2d$ is a sum of squares (SOS) if there exist polynomials $g_1, \ldots, g_m \in \mathbb{K}[x_1, \ldots, x_n]$ such that $f = \sum_{i=1}^m g_i^2$.*

Theorem 3.84 [56]**.** *Let $f \in \mathbb{R}[x_1, \ldots, x_n]$ be a polynomial of degree $2d$. If f is SOS then it is non-negative on $\mathbb{R}^n$. The converse holds for any such f if and only if $n = 1$ or $d = 1$ or $n = d = 2$.*

Example 3.85. *We demonstrate proofs that f is SOS if and only if f is non-negative on $\mathbb{R}^n$ in the simple cases $n = 1$ and $d = 1$.*

On the one hand, let $f \in \mathbb{R}[x]$ be a polynomial of degree $2d$ such that $f(x) \geq 0$ for all $x \in \mathbb{R}$. Let α be a root of f. If $\alpha \in \mathbb{R}$, then it must have even multiplicity else for some small ϵ, either $f(\alpha) + \epsilon < 0$ or $f(\alpha) - \epsilon < 0$. Hence, $(x - \alpha)^2 \mid f$. Otherwise, $\alpha = a + bi$ is complex and so its conjugate $\overline{\alpha} = a - bi$ is also a root of f. Note that $(x - a - bi)(x - a + bi) = (x - a)^2 + b^2$. Thus, f is a product of sums of two squares. By the identity $(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2$,

all products of sums of two squares are themselves sums of two squares. Therefore, f is a sum of two squares.

On the other hand, let $f \in \mathbb{R}[x_1, \ldots, x_n]$ be a non-negative polynomial of degree 2 and let $\boldsymbol{x} = (x_1, \ldots, x_n)$. Let $S \in \mathbb{R}^{n \times n}$ be the symmetric matrix such that $f = \boldsymbol{x} S \boldsymbol{x}^t$. For example, with

$$f = a_{1,1}x_1^2 + a_{1,2}x_1x_2 + \cdots + a_{2,2}x_2^2 + a_{2,3}x_2x_3 + \cdots + a_{n,n}x_n^2,$$

the matrix S is,

$$S = \begin{bmatrix} a_{1,1} & \frac{a_{1,2}}{2} & \cdots & \frac{a_{1,n}}{2} \\ \frac{a_{1,2}}{2} & a_{2,2} & \cdots & \frac{a_{2,n}}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{1,n}}{2} & \frac{a_{2,n}}{2} & \cdots & a_{n,n} \end{bmatrix}.$$

Since S is symmetric, it has an eigendecomposition, meaning that there exists a diagonal matrix $D \in \mathbb{R}^{n \times n}$ and an orthogonal matrix $O \in \mathbb{R}^{n \times n}$ such that $S = ODO^t$. Thus, $f = \boldsymbol{x}ODO^t\boldsymbol{x}^t = (\boldsymbol{x}O)D(\boldsymbol{x}O)^t$. Hence, f is a sum of squares.

Example 3.86 [80]**.** *A famous example of a non-negative polynomial that is* not *a sum of squares is the Motzkin polynomial $f = x^4y^2 + x^2y^4 - 3x^2y^2 + 1 \in \mathbb{R}[x, y]$. First, we show that f is non-negative on $\mathbb{R}^2$. Recall the inequality of arithmetic and geometric means: For all $i \in \mathbb{N}$ and for $z_i \geq 0$,*

$$\frac{z_1 + \cdots + z_n}{n} \geq \sqrt[n]{z_1 \cdots z_n}.$$

Hence, taking $z_1 = x^4y^2, z_2 = x^2y^4$ and $z_3 = 1$ we have that

$$\frac{x^4y^2 + x^2y^4 + 1}{3} \geq x^2y^2,$$

and so $f \geq 0$. Now, given the support of f, if it were to be a sum of squares it necessarily would be a sum of polynomials of the form $(ax^2y + bxy^2 + cxy + d)^2$, for $a, b, c, d \in \mathbb{R}$. However, no values of the coefficients a, b, c, d can give a negative x^2y^2 coefficient. Hence, f is not a sum of squares.

Chapter 4

Critical points, Determinantal ideals and Gröbner bases

Abstract. Determinantal polynomial systems are those involving the minors of some given matrix. An important situation where these arise is the computation of the critical values of a polynomial map restricted to an algebraic set. This leads directly to a strategy for, among other problems, polynomial optimisation.

Computing Gröbner bases is a classical method for solving polynomial systems in general. For practical computations, this consists of two main stages. First, a Gröbner basis is computed with respect to a DRL (degree reverse lexicographic) ordering. Then, a change of ordering algorithm, such as Sparse-FGLM, designed by Faugère and Mou, is used to find a Gröbner basis of the same system but with respect to a lexicographic ordering. The complexity of this latter step, in terms of the number of arithmetic operations in the ground field, is $O(qD^2)$, where D is the degree of the ideal generated by the input and q is the number of non-trivial columns of a certain $D \times D$ matrix.

While asymptotic estimates are known for q in the case of *generic* polynomial systems, thus far, the complexity of Sparse-FGLM was unknown for the class of determinantal systems.

By assuming Fröberg's conjecture, thus ensuring that the Hilbert series of generic determinantal ideals have the necessary structure, we expand the work of Moreno-Socías by detailing the structure of the DRL staircase in the determinantal setting. Then we study the asymptotics of the quantity q by relating it to the coefficients of these Hilbert series. Consequently, we arrive at a new bound on the complexity of the Sparse-FGLM algorithm for a certain class of generic determinantal systems and, in particular, for generic critical point systems.

We consider the ideal inside the polynomial ring $\mathbb{K}[x_1, \ldots, x_n]$, where $\mathbb{K}$ is some infinite field, generated by m generic polynomials of degree d and the maximal minors of an $m \times (n-1)$ polynomial matrix with generic entries of degree $d-1$. Then, in this setting, for the case $d = 2$ and for $n \gg m$ we establish an exact formula for q in terms of n and m. Moreover, for $d \geq 3$, we give a tight asymptotic formula, as $n \to \infty$, for q in terms of n, m and d.

This chapter contains joint work with J. Berthomieu, A. Bostan and M. Safey El Din and led to the publication [10].

4.1 Introduction

Motivation By the Lagrange multiplier theorem, the local extrema of a polynomial mapping restricted to a real algebraic set are contained in the set of critical values of the map. Thus, computing these values, and the corresponding minimum/critical points where these extrema are reached, leads to a strategy for polynomial optimisation under some regularity assumptions.

Polynomial optimisation is of principal importance in many areas of engineering and social sciences (including control theory [50, 55], computer vision [1, 88] and optimal design [25], etc.).

Critical point computations are also a fundamental task in the algorithms of effective real algebraic geometry. For example, the problems of deciding the emptiness of the set of real solutions of a polynomial system, counting the number of connected components of such sets and one block quantifier elimination can all be accomplished, under some regularity assumptions, by the so-called critical point method [6, Ch. 7], see also [57, 95].

With $\mathbb{K}$ an infinite field, let $\boldsymbol{g} = (g_1, \ldots, g_m) \in \mathbb{K}[x_1, \ldots, x_n]$ be a sequence of polynomials of degree d and let $\mathbf{V}(\boldsymbol{g}) \subset \mathbb{K}^n$ be their simultaneous vanishing set. Define φ_1 to be the projection map onto the first coordinate. We denote by $\mathcal{J}$ the Jacobian of $(\varphi_1, \boldsymbol{g})$,

$$\mathcal{J} := \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}.$$

An example of the ideals we consider in this chapter is the ideal I defined by $\boldsymbol{g}$ and the maximal minors of $\mathcal{J}$. By a corollary of the Jacobian criterion [27, Corollary 16.20], when $\boldsymbol{g}$ is a reduced regular sequence and $\mathbf{V}(\boldsymbol{g})$ is smooth, the algebraic set $\mathbf{V}(I)$ is exactly the set of critical points of the projection map φ_1 restricted to the algebraic set $\mathbf{V}(\boldsymbol{g})$.

Throughout this chapter, we shall consider what we call *generic determinantal sum* systems, that is the sum of a generic ideal with an ideal defined by the maximal minors of a polynomial matrix of a given size. Essentially, for the example of critical point systems, we choose the coefficients of the polynomials $g_1, \ldots, g_m$ so that they lie inside a non-empty Zariski open subset of $\mathbb{K}^{\binom{n+d}{d}}$ where the results of [33] hold. In particular, the generic systems we consider satisfy the conditions of the Jacobian criterion so that I encodes the critical points of φ_1 restricted to $\mathbf{V}(\boldsymbol{g})$ [96, Lemma A.2]. Moreover, by [33, Lemma 2] and [64, Proposition 4.2], I is a zero-dimensional, radical ideal. So, the quotient algebra $\mathbb{K}[x_1, \ldots, x_n]/I$ is a finite dimensional vector space over $\mathbb{K}$.

For the many applications of the critical point method previously discussed, one wishes to compute a rational parametrisation of this set of critical points. By our genericity conditions, we shall assume that the ideal I is in *shape position*, meaning that for a lexicographic (LEX) ordering with x_n as the least variable, the LEX Gröbner basis has the following structure:

$$\{x_1 - f_1(x_n), \ldots, x_{n-1} - f_{n-1}(x_n), f_n(x_n)\},$$

where the degree of f_n is the degree of the ideal I [8]. A fast method commonly used in practice, and the one which we shall use, to compute a LEX Gröbner basis is to first compute a Gröbner basis of I with respect to a degree reverse lexicographic ordering (DRL). Then, one uses a change of ordering algorithm to compute another Gröbner basis of I but with respect to a LEX ordering.

Previous works In [33, Theorem 3], Faugère, Safey El Din and Spaenlehauer give an upper bound on the number of arithmetic operations necessary for computing a LEX Gröbner basis of a generic determinantal sum system within the DRL to LEX framework. They do so by deriving the Hilbert series of such a system, using results by Conca and Herzog [22, Corollary 1].

Then, based on a result of Bardet, Faugère and Salvy [5, Theorem 7], the authors of [33, Theorem 3] analyse the complexity of the DRL step using Faugère's F_5 algorithm [29]. Here, and in the whole text, complexity estimates are given in terms of arithmetic operations in the ground field $\mathbb{K}$. Next, to obtain a LEX Gröbner basis, since we are in the zero-dimensional case, they use the FGLM algorithm to perform the change of ordering [31]. The complexity of FGLM is $O(nD^3)$, where D is the degree of the determinantal sum ideal. In [84, Theorem 2.2], Nie and Ranestad use the Thom-Porteous-Giambelli formula to prove that this degree is

$$D = d^p (d-1)^{n-m} \binom{n-1}{m-1}.$$

In [32], Faugère and Mou proposed another algorithm that solves the change of ordering step, the Sparse-FGLM algorithm. Under some stability assumptions, Sparse-FGLM relies primarily on the structure of the matrix T_{x_n} associated to the linear map of multiplication by x_n in the finite dimensional quotient algebra $\mathbb{K}[x_1, \ldots, x_n]/I$. When the ideal in question is in shape position, meaning that the leading monomials of its reduced LEX Gröbner basis are $x_1, \ldots, x_{n-1}, x_n^D$, its complexity is $O(qD^2 + nD\log^2 D)$, where q is the number of non-trivial columns of the matrix T_{x_n}. This number is studied in the same paper for generic complete intersections using the results of Moreno-Socías [79]. By deriving the asymptotics of the number of non-trivial columns, as well as by proving that the structure of the matrix T_{x_n} is such that it can be computed free of arithmetic operations, Faugère and Mou demonstrate in [32] that the complexity of Sparse-FGLM is indeed an improvement of that of FGLM.

Very recently, under the same shape and stability assumptions, Berthomieu, Neiger and Safey El Din, the authors of [12], designed an algorithm with a complexity of $O^{\sim}(q^{\omega-1}D)$. This algorithm improves upon [32] as well as the algorithms designed in [30, 83] which improve upon FGLM using fast linear algebra techniques to achieve complexities of $O^{\sim}(D^{\omega})$ and $O(nD^{\omega}\log(D))$ respectively.

Main results In this chapter, under similar genericity assumptions and by assuming a variant of Fröberg's conjecture [36], we extend the results of [32, 79] to generic determinantal sum ideals. We emphasise here that our results hold not only for critical point systems but indeed for any sufficiently generic determinantal sum system. This is made precise in Definition 2.2.

Firstly, we prove a result on the structure of the DRL staircase, which implies that the only non-trivial columns of T_{x_n} correspond one-to-one with monomials which, once multiplied by x_n, give a leading monomial in the reduced DRL Gröbner basis. Furthermore, for each such monomial, one can read the entries of the corresponding non-trivial column from the polynomial in the Gröbner basis with that leading monomial. This implies the following theorem.

Theorem 2.3. *Let I be a generic determinantal sum ideal so that the conditions of Definition 2.2 hold. Assume that a reduced Gröbner basis of I with respect to a DRL ordering is known. Then the multiplication matrix T_{x_n} can be constructed without performing any arithmetic operations.*

Continuing further, we prove an explicit formula for the number of non-trivial columns of T_{x_n}, which we denote q, in the case of quadratic polynomials with a large number of variables n compared to the number of polynomials m. Then, for any choice of degree $d \geq 3$ and for $n \to \infty$, we prove asymptotic formulae for q.

Theorem 2.4. *Let I be a generic determinantal sum ideal so that the conditions of Definition 2.2 hold, and let T_{x_n} be the matrix associated to the linear map of multiplication by x_n. Denote by q the number of non-trivial columns of T_{x_n}. Then, for $d = 2$ and $n \gg m$,*

$$q = \sum_{k=0}^{m-1} \binom{n-m-1+k}{k} \binom{m}{\lfloor 3m/2 \rfloor - 1 - j}. \tag{2.1}$$

Moreover, for $d \geq 3$ and $n \to \infty$,

$$q \approx \frac{1}{\sqrt{(n-m)\pi}} \sqrt{\frac{6}{(d-1)^2 - 1}} d^m (d-1)^{n-m} \binom{n-2}{m-1}. \tag{2.2}$$

By [32, Theorem 3.2], and since the ideals we consider are in shape position, Theorem 2.4 leads directly to a complexity result for the Sparse-FGLM algorithm. Therefore, we arrive at an improved upper bound on the complexity of the change of ordering step for generic determinantal sum systems.

Theorem 2.5. *Let I be a generic determinantal sum ideal so that the conditions of Definition 2.2 hold. Assume that a reduced DRL Gröbner basis of I is known. Then, for $d \geq 3$, the arithmetic complexity of computing a LEX Gröbner basis of I is upper bounded by*

$$O\left(\frac{d^{3m}(d-1)^{3(n-m)}}{\sqrt{(n-m)d\pi}}\binom{n-2}{m-1}\binom{n-1}{m-1}^2\right).$$

Hence, the complexity gain of Sparse-FGLM over FGLM for generic determinantal sum systems is approximately

$$O\left(\frac{q}{nD}\right) \approx O\left(\frac{\sqrt{n-m}}{n^2(d-1)}\right).$$

Organisation of the chapter The remainder of the chapter consists of: Section 7.2, where we define the class of ideals for which our results hold; Section 4.3, where we prove our main results; and Section 5.5, where we test our formula for the number of non-trivial columns of the matrix T_{x_n} for various parameters.

4.2 Preliminaries

4.2.1 Shape position

Let $g_1, \ldots, g_m \in \mathbb{K}[x_1, \ldots, x_n]$ be polynomials of degree d. Similarly, let $h_{1,2}, \ldots, h_{m,n} \in \mathbb{K}[x_1, \ldots, x_n]$ be polynomials of degree $d-1$. Let I be the determinantal sum ideal generated by $\langle g_1, \ldots, g_m \rangle$ and the maximal minors of the following matrix:

$$\begin{bmatrix} h_{1,2} & \cdots & h_{1,n} \\ \vdots & \ddots & \\ h_{m,2} & \cdots & h_{m,n} \end{bmatrix}.$$

The authors of [64, Proposition 4.2] show that if the coefficients of $g_1, \ldots, g_m$ and $h_{1,2}, \ldots, h_{m,n}$ are chosen in some non-empty Zariski open subsets of $\mathbb{K}^{\binom{n+d}{d}}$ and $\mathbb{K}^{\binom{n+d-1}{d-1}}$ respectively, then the ideal I defined above is radical and zero-dimensional.

In order to apply the results of [32] to our determinantal sum ideals, we require they be in shape position. To ensure this, we add a new indeterminate that acts as a primitive element of the quotient algebra. For any $\lambda \in \mathbb{K}^n$, define the ideal

$$J = I + \langle y - \sum_{j=1}^{n} \lambda_j x_j \rangle \quad \subset \quad \mathbb{K}[x_1, \ldots, x_n, y].$$

The idea of the following lemma is similar to that of applying a generic linear change of variables to the ideal I. However, introducing a new variable to be the least variable in the monomial ordering also allows one to avoid some degenerate cases that will be discussed in Remark 4.5.

Lemma 4.1. *Let $\mathbb{K}$ be an infinite field. Then, there exists a non-empty Zariski open subset $\mathcal{O}$ of $\mathbb{K}^n$ such that for all $\lambda \in \mathcal{O}$ and with y as the least variable in the LEX ordering, the ideal J is in shape position.*

Proof. By [64, Proposition 4.2], the ideal I is radical and zero-dimensional. Thus, for all $\lambda \in \mathbb{K}^n$, the ideal J is also zero-dimensional and radical. By [41, Proposition 1.6], [8, Proposition 5] and the genericity of the polynomials defining I, we have that J is in shape position if and only if each of the finitely many points in the algebraic set $\mathbf{V}(J)$ has a unique y-coordinate. As $\mathbb{K}$ is infinite, the finitely many linear equations that give equality of the y coordinate of any two points in $\mathbf{V}(J)$ define a proper Zariski closed subset of $\mathbb{K}^n$. Therefore, there exists a non-empty Zariski open subset $\mathcal{O}$ of $\mathbb{K}^n$ such that for all $\lambda \in \mathcal{O}$ the y coordinate of each point in the algebraic set $\mathbf{V}(J)$ is unique. Hence, for $\lambda \in \mathcal{O}$, the ideal J is in shape position. □

4.2.2 Fröberg's conjecture

As a direct consequence of [22], the authors of [33] further show that under the same genericity assumptions, the following proposition holds:

Proposition 4.2 [33, Proposition 1]. *The Hilbert series of $\mathbb{K}[x_1,\dots,x_n]/I$ is*

$$H = \frac{\det(M(t^{d-1}))}{t^{(d-1)\binom{m-1}{2}}}\frac{(1-t^d)^m(1-t^{d-1})^{n-m}}{(1-t)^n}$$

where $M(t)$ is the $(m-1)\times(m-1)$ matrix whose (i,j)th entry is $\sum_k \binom{m-i}{k}\binom{n-1-j}{k}t^k$.

We shall consider the quotients of the algebra $\mathbb{K}[x_1,\dots,x_n]/I$ by powers of generic linear forms. By the genericity introduced above, it suffices to consider the quotients of $A=\mathbb{K}[x_1,\dots,x_n,y]/J$ by powers of y. Thus, denote by HQ_e the Hilbert series of $A/\langle y^e\rangle$, for $e\geq 1$. In order to control the shape of this Hilbert series, we rely on a variant of Fröberg's conjecture given in [34, Lemma 14]. First, however, a definition.

Definition 4.3. *For a series $S=\sum_k a_k t^k$, we define*

$$\left[\sum_k a_k t^k\right]_+$$

to be the series S truncated at the first non-positive coefficient.

Lemma 4.4 [34, Lemma 14]. *If Fröberg's conjecture is true, then for all $e\geq 1$*

$$HQ_e = [(1-t^e)H]_+ .$$

We remark that in [86], Pardue showed that Moreno-Socías' conjecture [79, Conjecture 4.2] implies Fröberg's conjecture, as well as a number of other interesting conjectures. Moreover, while these conjectures are usually given in a homogeneous setting, we shall assume that Lemma 4.4 holds also in the affine case.

4.2.3 Generic determinantal sum ideals

With the assumption of Fröberg's conjecture, we recall the precise definition of the class of ideals we consider in this chapter.

Definition 2.2. *With $\mathbb{K}$ an infinite field, let $I\subset\mathbb{K}[x_1,\dots,x_n]$ be an ideal which is the sum of m polynomials of degree at most d and the maximal minors of a matrix with polynomial entries also of degree at most d. We say that I is a generic determinantal sum ideal if the following three conditions hold:*

- *the ideal I is in shape position, meaning that the reduced LEX Gröbner basis with $x_1\succ\cdots\succ x_n$ has leading monomials $x_1,\dots,x_{n-1},x_n^D$ where D is the degree of I,*
- *the Hilbert series H of $\mathbb{K}[x_1,\dots,x_n]/I$ is equal to*

  $$H = \frac{\det(M(t^{d-1}))}{t^{(d-1)\binom{m-1}{2}}}\frac{(1-t^d)^m(1-t^{d-1})^{n-m}}{(1-t)^n}$$

  where $M(t)$ is the $(m-1)\times(m-1)$ matrix whose (i,j)th entry is $\sum_k \binom{m-i}{k}\binom{n-1-j}{k}t^k$,
- *for all $e\geq 1$, the Hilbert series of $(\mathbb{K}[x_1,\dots,x_n]/I)/\langle x_n^e\rangle$ is equal to the series $(1-t)H$ truncated at the first non-positive coefficient.*

By our genericity assumptions, $g_1, \ldots, g_m$ is a reduced, regular sequence defining a smooth algebraic set $\mathbf{V}(g_1, \ldots, g_m)$. By [33, Lemma 6] and [24, Ch. 9, Sec. 3, Prop. 9], the determinantal sum ideal defining the critical points of the projection map onto the first coordinate restricted to $\mathbf{V}(g_1, \ldots, g_m)$ satisfies Proposition 4.2. Moreover, using the same addition of a new indeterminate as in Lemma 4.1, one may assume that such an ideal is in shape position. Thus, by assuming Fröberg's conjecture, these generic critical point systems are an important example of the generic determinantal sum ideals we consider.

Remark 4.5. *We note that without the addition of a new indeterminate, generic critical point systems may not satisfy the conditions of Definition 2.2. In particular, if one considers a DRL ordering with x_n as the least variable for the determinantal sum system defining the critical values of the projection map onto the x_n-axis, then Lemma 4.4 no longer holds. The consequence of this is that the results of this chapter cannot then be applied to this special case. However, introducing a new indeterminate to be the least variable in the DRL ordering, as in Lemma 4.1, rectifies this problem. Therefore, we may assume that all generic critical point systems satisfy the conditions of Definition 2.2.*

Furthermore, note that the Hilbert series of $\mathbb{K}[x_1, \ldots, x_n]/I$ is equal to the Hilbert series of $\mathbb{K}[x_1, \ldots, x_n, y]/J$. Therefore, for ease of notation, we shall assume that the determinantal sum ideals considered in this chapter satisfy Definition 2.2 without introducing the new indeterminate y.

4.3 Proofs

Roadmap Firstly, as in the papers [32, 33, 79], to prove our results we rely on manipulations of the Hilbert series H from Proposition 4.2. However, for our purposes, the form involving the determinant of the matrix M makes this difficult. Thus, our first step is to express H in a simpler form in Section 4.3.1. Then we show that this Hilbert series is always *unimodal* in Section 4.3.2. This property, along with the assumption of Fröberg's conjecture, allows us to prove in Section 4.3.3 a structure theorem on the generic DRL staircase. This leads to our first main result, that the multiplication matrix T_{x_n} can be constructed for free. Combining this result with the unimodality property, we show that the number of non-trivial columns of this matrix, a key parameter of the Sparse-FGLM algorithm, is equal to the largest coefficient of the series H. In Section 5.4, we conclude the proof of our main results by studying the asymptotics of the largest coefficient of H.

4.3.1 Simplification of the Hilbert series

As in the works we wish to generalise [32, 33, 79], our results rely heavily on the Hilbert series of the generic determinantal sum ideals we consider. Thus, the first stage we take is to simplify the form given in Proposition 4.2. We do so by expressing the determinant of the binomial matrix in this Hilbert series as a binomial sum. We start with some general results involving binomial matrices that will lead to the simplification we want as a special case.

Let $\mathcal{A} = (a_{ij})_{i,j\geq 0}$ be the infinite Pascal matrix defined by $a_{ij} = \binom{i}{j}$ for $j \leq i$ and $a_{ij} = 0$ for $j > i$. The minor of this matrix corresponding to rows $0 \leq a_1 < \cdots < a_n$ and columns $0 \leq b_1 < \cdots < b_n$ will be denoted by

$$\binom{a_1, \ldots, a_n}{b_1, \ldots, b_n} = \begin{vmatrix} \binom{a_1}{b_1} & \cdots & \binom{a_1}{b_n} \\ \vdots & \ddots & \vdots \\ \binom{a_n}{b_1} & \cdots & \binom{a_n}{b_n} \end{vmatrix}.$$

We recall the following two lemmas from [39].

Lemma 4.6 [39, Lemma 8]. *If $b_1 \neq 0$, then*

$$\binom{a_1,\ldots,a_k}{b_1,\ldots,b_k} = \frac{a_1\cdots a_k}{b_1\cdots b_k}\binom{a_1-1,\ldots,a_k-1}{b_1-1,\ldots,b_k-1}.$$

Lemma 4.7 [39, Lemma 9]. *The following holds*

$$\binom{a,a+1,\ldots,a+k-1}{0,b_2,\ldots,b_k} = \binom{a,a+1,\ldots,a+k-2}{b_2-1,b_3-1,\ldots,b_k-1}.$$

We can now prove the following identity.

Lemma 4.8. *Let S be the $k\times(k+1)$ submatrix corresponding to rows $a+1,a+2,\ldots,a+k$ and columns $0,1,\ldots,k$. Then, for $0\leq l\leq k$, the minors of this submatrix are equal to*

$$\binom{a+1,a+2,\ldots,a+k}{0,1,\ldots\ell-1,\ell+1,\ldots,k} = \binom{a+k-\ell}{k-\ell}.$$

Proof. Apply Lemma 4.7 ℓ times to the minor

$$\binom{a+1,a+2,\ldots,a+k}{0,1,\ldots\ell-1,\ell+1,\ldots,k}.$$

The result is the minor

$$\binom{a+1,a+2,\ldots,a+k-\ell}{1,\ldots,k-\ell}.$$

Next, apply Lemma 4.6 to obtain the minor

$$\frac{(a+1)\cdots(a+k-\ell)}{1\cdots(k-\ell)}\binom{a,\ldots,a-1+k-\ell}{0,1,\ldots,k-\ell-1} = \binom{a+k-\ell}{k-\ell}\binom{a,\ldots,a-1+k-\ell}{0,1,\ldots,k-\ell-1}.$$

Finally, apply Lemma 4.7 another $k-\ell-1$ times until the minor is reduced to a single entry

$$\binom{a+k-\ell}{k-\ell}\binom{a}{0} = \binom{a+k-\ell}{k-\ell}.$$ □

Lemma 4.9. *Let P be the $p\times p$ matrix with entries in $\mathbb{K}[x,y,t]$ defined by $P_{i,j}=\sum_{k=0}^{p}\binom{x-i}{k}\binom{y-j}{k}t^k$. Then*

$$\frac{\det(P)}{t^{\binom{p}{2}}} = \sum_{k=0}^{p}\binom{x-p-1+k}{k}\binom{y-p-1+k}{k}t^k.$$

Proof. Let A be the $p\times(p+1)$ matrix with entries $a_{ik}=\binom{x-i}{k-1}$. Let B be the $(p+1)\times p$ matrix with entries $B_{kj}=\binom{y-j}{k-1}t^{k-1}$. Observe that $P=AB$.

We shall write $A^{[\ell]}$ (resp. $B^{[\ell]}$) for the matrix A (resp. B) with its ℓth column (resp. row) removed. By the Cauchy-Binet formula

$$\det(P) = \sum_{\ell=1}^{p+1}\det(A^{[\ell]})\det(B^{[\ell]}).$$

We begin with the matrix A. Notice that by making $\frac{1}{2}p(p+1)$ column transpositions, one can rearrange A so that it is a submatrix of the Pascal matrix $\mathcal{A}$. Specifically, one can rearrange the columns of A so that it has rows $x-p,\ldots,x-1$ and columns $0,\ldots,p$ of $\mathcal{A}$. Then, by Lemma 4.8, the determinant of the minors of A equals, up to the sign difference from the transpositions,

$$\det(A^{[\ell]}) = \pm\binom{x-\ell}{p-\ell+1}.$$

Now, let C be the matrix B with $t = 1$. Then note that

$$\det(B^{[\ell]}) = \det(C^{[\ell]})t^{\binom{m}{2}+p-\ell+1}.$$

In the same way as for the matrix A, by taking the transpose of C and making $\frac{1}{2}p(p+1)$ column transpositions, one can rearrange C so that is has the form of a submatrix of $\mathcal{A}$. We find that

$$\det(C^{[\ell]}) = \pm\binom{y-\ell}{p-\ell+1}, \quad \text{and thus } \det(B^{[\ell]}) = \pm\binom{y-\ell}{p-\ell+1}t^{\binom{p}{2}+p-\ell+1}.$$

Returning to the Cauchy-Binet formula,

$$\det(P) = \sum_{\ell=1}^{p+1}\binom{x-\ell}{p-\ell+1}\binom{y-\ell}{p-\ell+1}t^{\binom{p}{2}+p-\ell+1}.$$

By a change of coordinates, substituting $k = p-\ell+1$, we arrive at

$$\det(P) = \sum_{k=0}^{p}\binom{x-p-1+k}{k}\binom{y-p-1+k}{k}t^{\binom{p}{2}+k}. \qquad \square$$

Corollary 4.10. *The Hilbert series H from Proposition 4.2 can be expressed as*

$$H = \left(\sum_{k=0}^{m-1}\binom{n-m-1+k}{k}t^{k(d-1)}\right)\frac{(1-t^d)^m(1-t^{d-1})^{n-m}}{(1-t)^n}.$$

Proof. By Proposition 4.2,

$$H = \frac{\det(M(t^{d-1}))}{t^{(d-1)\binom{m-1}{2}}}\frac{(1-t^d)^m(1-t^{d-1})^{n-m}}{(1-t)^n}$$

where $M(t)$ is the $(m-1)\times(m-1)$ matrix whose (i,j)th entry is $\sum_k\binom{m-i}{k}\binom{n-1-j}{k}t^k$. Thus, as the special case of Lemma 4.9 with $p = m-1, x = m$ and $y = n-1$,

$$\frac{\det(M(t))}{t^{\binom{m-1}{2}}} = \sum_{k=0}^{m-1}\binom{n-m-1+k}{k}t^k. \qquad \square$$

4.3.2 Unimodality

The Hilbert series of the systems we study are highly structured. In particular, it was shown in [79, Proposition 2.2] that the Hilbert series of generic complete intersections are symmetric and so-called unimodal polynomials. As we transition to more general determinantal sum ideals, we may lose some of this structure for certain choices of parameters. However, we show in this section that our series are always unimodal. This property will then be exploited in the remaining two parts of Section 4.3. We begin with the definition of unimodality.

Definition 4.11. *A polynomial $\sum_{k=0}^{n}a_kt^k$ with non-negative coefficients and $a_n > 0$ is* unimodal *if there exists $N \in \mathbb{N}$, $N \leq n$ such that*

$$\begin{array}{ll} a_{k-1} \leq a_k \leq a_N & \text{for} \quad 1 \leq k \leq N, \\ a_k \geq a_k \geq a_{k+1} & \text{for} \quad N \leq k \leq n-1. \end{array}$$

Unimodality is not necessarily preserved by multiplication. For example, the polynomial $f = 3+t+t^2$ is unimodal, while $f^2 = 9+6t+7t^2+2t^3+t^4$ is not.

Definition 4.12. *A polynomial f with non-negative coefficients is strongly unimodal if, for all unimodal polynomials g, the product fg is unimodal.*

Note that a strongly unimodal polynomial is also unimodal. A classical example of a strongly unimodal polynomial is as follows.

Lemma 4.13. *For any $d \in \mathbb{N}$, the polynomial $f = 1 + t + \cdots + t^d$ is strongly unimodal.*

Proof. Let $g = \sum_{k=0}^{n} a_k t^k$ be a unimodal polynomial with integer N such that

$$\begin{aligned} a_{k-1} \leq a_k \leq a_N \quad &\text{for} \quad 1 \leq k \leq N, \\ a_k \geq a_k \geq a_{k+1} \quad &\text{for} \quad N \leq k \leq n-1. \end{aligned}$$

For ease of notation, let $a_k = 0$ if $k < 0$ or $k > n$. Let $fg = \sum_{k=0}^{n+d} b_k t^k$ so that $b_k = a_{k-d} + \cdots + a_k$. Suppose that there does not exist an integer σ such that $b_{\sigma+1} < b_\sigma$, then fg is trivially unimodal. On the other hand, suppose such an index exists and let M be the least integer such that $b_{M+1} < b_M$. Clearly, $M \geq N$, since the coefficients of g are non-decreasing up to index N. Assume that for some k, for all ℓ such that $M \leq \ell < k$ we have that $b_{\ell+1} \leq b_\ell$. Then $a_k - a_{k-d-1} \leq 0$. Since $k + 1 \geq M + 1 > N$, by the unimodality of g, $a_{k+1} \leq a_k$. Similarly, if $k - d \leq N$ we have $a_{k-d-1} \leq a_{k-d}$. Hence, by the inductive assumption, $b_{k+1} - b_k = a_{k+1} - a_{k-d} \leq a_k - a_{k-d-1} \leq 0$. Alternatively, if $k - d > N$, then by unimodality of g we have $a_{k+1} - a_{k-d} \leq 0$. Hence, by induction, $b_{k+1} \leq b_k$ for all $k > M$. Thus, fg is a unimodal polynomial and we conclude that f is a strongly unimodal polynomial. □

Unlike unimodality, strong unimodality is preserved by multiplication.

Lemma 4.14. *Let f, g be strongly unimodal polynomials. Then, fg is a strongly unimodal polynomial.*

Proof. Let h be a unimodal polynomial. Then, since g is strongly unimodal, gh is a unimodal polynomial. Hence, since f is strongly unimodal, fgh is unimodal and so fg is strongly unimodal. □

We shall prove that the Hilbert series of a generic determinantal sum ideal is unimodal by showing that it is the product of a strongly unimodal polynomial and a unimodal polynomial.

Lemma 4.15. *Let H be the Hilbert series from Proposition 4.2, with parameters $n, m, d \in \mathbb{N}$ where $n > m$. Then H is a unimodal polynomial.*

Proof. Firstly, by Corollary 4.10,

$$H = \left(\sum_{k=0}^{m-1} \binom{n-m-1+k}{k} t^{k(d-1)} \right) \frac{(1-t^d)^m (1-t^{d-1})^{n-m}}{(1-t)^n}.$$

Our strategy is to show that we can write this polynomial as the product of a unimodal polynomial and a strongly unimodal polynomial. The polynomial H would then be unimodal by Definition 5.2.

For $d > 2$, the binomial sum factor

$$\sum_{k=0}^{m-1} \binom{n-m-1+k}{k} t^{k(d-1)}$$

is not unimodal. However, since $n \geq m - 1$, the remaining factor of H always has the following polynomial as a factor:

$$\frac{1-t^{d-1}}{1-t} = 1 + t + \cdots + t^{d-2}.$$

Therefore, we can always multiply this factor into the binomial sum above

$$\sum_{k=0}^{m-1} \binom{n-m-1+k}{k} t^{k(d-1)} (1 + t + \cdots + t^{d-2}) = \sum_{k=0}^{m-1} \sum_{i=0}^{d-2} \binom{n-m-1+k}{k} t^{k(d-1)+i}.$$

The resulting polynomial is unimodal as its coefficients are non-decreasing with no internal zeroes.

Consider the remaining quotient

$$\frac{(1-t^d)^m(1-t^{d-1})^{n-m-1}}{(1-t)^{n-1}}.$$

This polynomial is the product of $n-1$ polynomials of the form $1+t+\cdots+t^p$ for some $p \in \mathbb{N}$. By Lemma 4.13, each of these polynomials is strongly unimodal. Thus, by Lemma 4.14, the remaining quotient,

$$\frac{(1-t^d)^m(1-t^{d-1})^{n-m-1}}{(1-t)^{n-1}},$$

is strongly unimodal. Therefore, since H is the product of a strongly unimodal polynomial and a unimodal polynomial, H is unimodal. □

Remark 4.16. *In the context of this chapter, by the unimodality of the Hilbert series H, Definition 4.3 is equivalent to the definition given in [79, Section 1].*

4.3.3 Staircase structure

In this section, we prove a structure theorem on the DRL staircase for generic determinantal sum ideals. Let $(f_1, \ldots, f_k)$ be a reduced and minimal Gröbner basis of I with respect to a DRL ordering with x_n as the least variable. For $1 \leq i \leq k$, let $r_i \in \mathscr{M}$ be the leading monomial of f_i, where $\mathscr{M}$ is set of monomials of $\mathbb{K}[x_1, \ldots, x_n]$. Then we shall denote the DRL staircase by

$$E = \bigcap_{i=1}^{k} \{r \in \mathscr{M} \mid r_i \nmid r\}.$$

The elements of the staircase give a natural basis for the quotient algebra $\mathbb{K}[x_1, \ldots, x_n]/I$. For each $b \in E$, the columns of the matrix T_{x_n} are the normal forms of $x_n b$ with respect to the DRL Gröbner basis expressed in terms of the basis E. Thus, the construction of the column of T_{x_n} corresponding to $x_n b$ falls into exactly one of following three cases:

1. $x_n b \in E$: Then the corresponding column is sparse, consisting of all zeroes except one entry with a value of 1 in the row corresponding to $x_n b$.

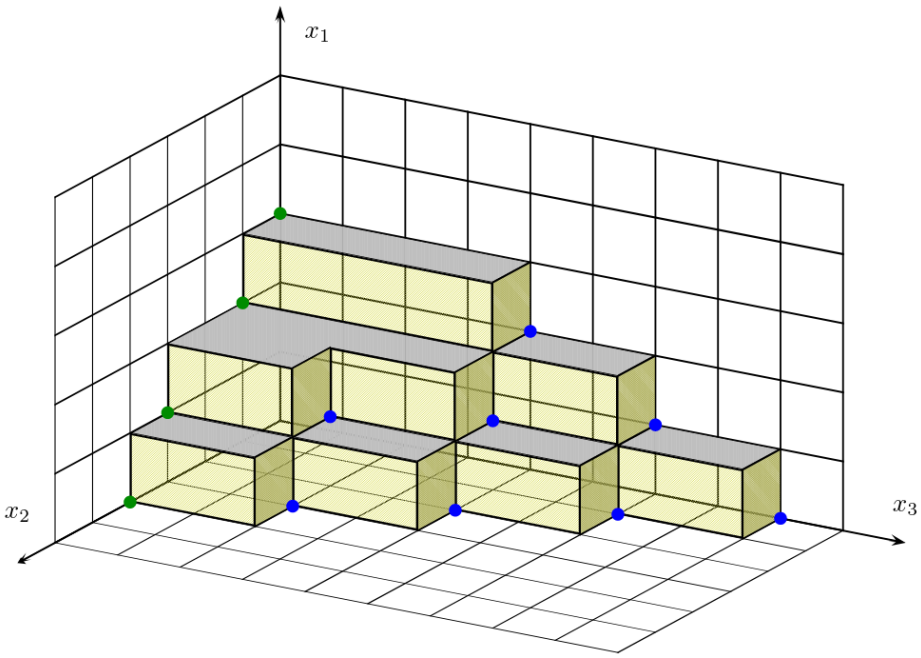
2. $x_n b$ is a leading term of the reduced DRL Gröbner basis: Then the normal form is obtained from the polynomial f in the Gröbner basis whose leading term is $x_n b$.

3. Otherwise, the normal form must be computed.

In the first case, the corresponding column is trivial. In the latter two cases, the corresponding columns are non-trivial. Usually, and in the case we consider with generic polynomials, these non-trivial columns are dense. Moreover, constructing columns that fall into the first two cases do not require any arithmetic operations.

We establish in this subsection that, for generic determinantal sum ideals, only the first two cases occur. This implies that the number of non-trivial columns of the matrix T_{x_n} is equal to the number of leading monomials of elements of the reduced DRL Gröbner basis that have positive degree in x_n.

To prove this result, we consider the Hilbert series H, its simplified form from Corollary 4.10 as well as the unimodal property of Lemma 4.15. Here, we illustrate an example of the DRL staircase in the case $(n, m, d) = (3, 2, 3)$.



Here, the cubes represent elements of the staircase and the dots are the leading monomials of the reduced DRL Gröbner basis. We can see that in this instance, the number of non-trivial columns is equal to the number of blue dots, the number of leading monomials of elements of the reduced Gröbner basis that have positive degree in x_n.

We recall the definition of HQ_e, the Hilbert series of $(\mathbb{K}[x_1,\ldots,x_n]/I)/\langle x_n^e\rangle$, for $e \geq 1$. Also, recall that we assume that Fröberg's conjecture is true and so the conclusion of Lemma 4.4 holds. In particular, this implies that the degree of the polynomial HQ_1 is equal to the degree of the term of largest coefficient of H, or the least such degree if there are multiple terms with equal largest coefficient. We shall refer to this degree by Σ. Moreover, for ease of notation, we shall denote

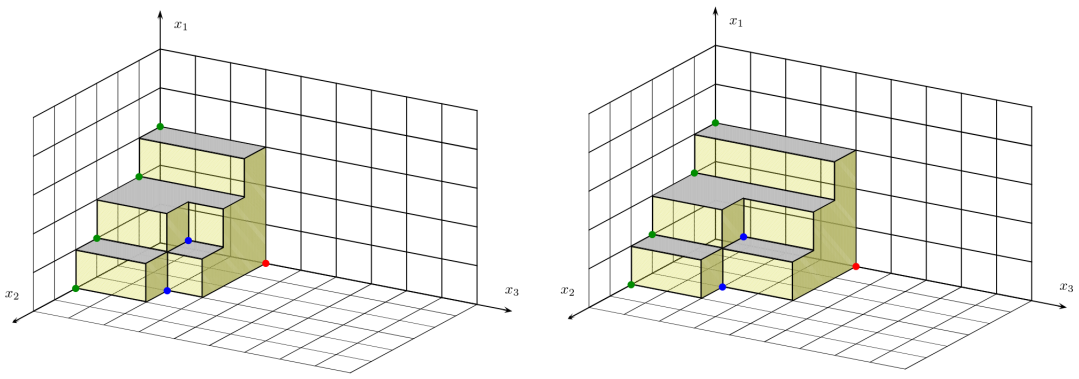
$$\Delta = (m-1)(d-1) + m(d-1) + (n-m)(d-2),$$

so that Δ equals the degree of H.

Note that the DRL ordering with x_n as the least variable is compatible with these quotients. We recall the following property that can be easily verified:

Lemma 4.17 [79, Lemma 1.9]**.** *Let $I \in \mathbb{K}[x_1,\ldots,x_n]$ be a polynomial ideal and let $\{f_1,\ldots,f_k\}$ be a Gröbner basis of I with respect to a DRL ordering with x_n as the least variable. Then $\{f_1,\ldots,f_k,x_n^e\}$ is a Gröbner basis of $I+\langle x_n^e\rangle$. Moreover, if $\{f_1,\ldots,f_k\}$ is additionally a reduced Gröbner basis, then removing from $\{f_1,\ldots,f_k,x_n^e\}$ all f_i such that x_n^e divides the leading term of f_i gives a reduced Gröbner basis of $I+\langle x_n^e\rangle$.*

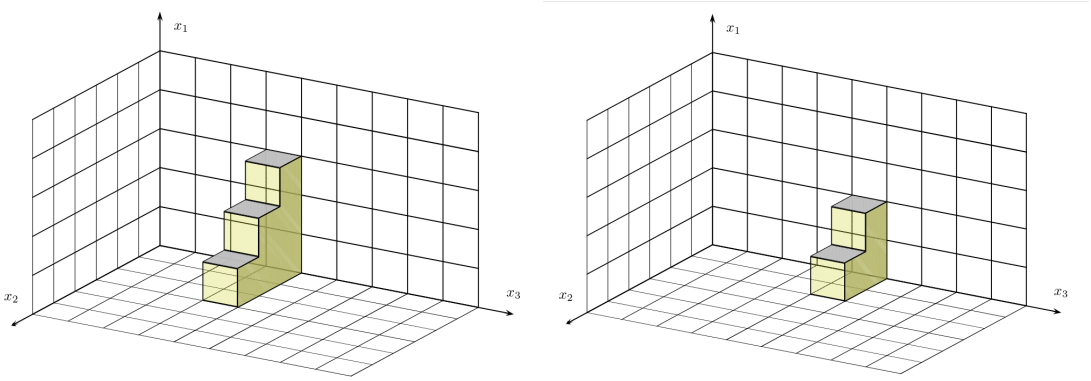
This compatibility can be easily seen from the corresponding staircases:



Here we see the quotients by x_3^3 and x_3^4. Adding these monomials to the Gröbner basis is indicated by the red dots. As in [79], for $e \geq 1$ we consider the eth section

$$H_e = \frac{HQ_{e+1} - HQ_e}{t^e}.$$

Effectively, we consider the Hilbert series of a cross section of the DRL staircase.



Here we illustrate $t^3 H_3$ and $t^4 H_4$. From this example, it is clear that by scaling these polynomials, by dividing by t^3 and t^4 respectively, the difference of these polynomials tells us about how the stairs change as we increase the degree of x_n. To study these sections, we first prove a result restricting the degree they can have.

Lemma 4.18. *For all $e \geq 1$, $\deg HQ_{e+1} - \deg HQ_e \in \{0, 1\}$.*

Proof. Let $H = \sum_{k=0}^{\Delta} a_k t^k$. For a given e, let σ be the degree of HQ_e. By Lemma 4.15, H is unimodal. Therefore, by Lemma 4.4,

$$HQ_e = a_0 + \cdots + a_{e-1}t^{e-1} + (a_e - a_0)t^e + \cdots + (a_\sigma - a_{\sigma-e})t^\sigma$$

Moreover, since H is unimodal, the degree of HQ_{e+1} is at least the degree of HQ_e and $\sigma \geq \Sigma$, where Σ is the degree of HQ_1. For the purpose of contradiction, suppose that the degree of HQ_{e+1} is $\sigma + 2$. Then

$$HQ_e = \left[a_0 + \cdots + (a_\sigma - a_{\sigma-e})t^\sigma + (a_{\sigma+1} - a_{\sigma+1-e})t^{\sigma+1}\right]_+$$

and

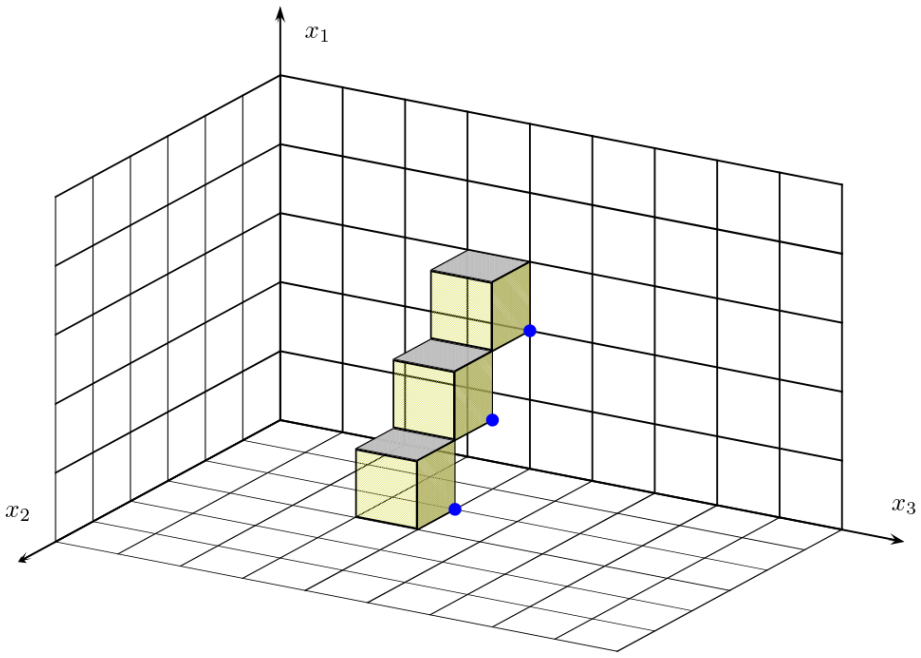
$$HQ_{e+1} = a_0 + \cdots + (a_{\sigma+2} - a_{\sigma+1-e})t^{\sigma+2}.$$

This implies that

$$a_{\sigma+1} \leq a_{\sigma+1-e},$$
$$a_{\sigma+2} > a_{\sigma+1-e}.$$

Therefore, $a_{\sigma+1} < a_{\sigma+2}$. This is a contradiction, as a_Σ is the largest coefficient of H and so $a_{\sigma+1} \geq a_{\sigma+2}$ by unimodality. Clearly, the same argument holds if the degree of HQ_{e+1} is greater than $\sigma + 2$. Therefore, the degree of HQ_{e+1} is either σ or $\sigma + 1$. □

With Lemma 4.18, we greatly restrict the possible degrees these sections can have. This allows us to prove a result on the differences of these sections.



We see here that the difference of sections tells us when there are drops in the staircase as we increase the degree of x_n. Note that the three monomials in the illustration of the difference $t^3(H_4 - H_3)$ correspond to the three leading monomials in the reduced Gröbner basis that have degree 4 in x_3. With the following lemma and proposition, we show that this correspondence always occurs.

Lemma 4.19. *For all $e \geq 1$, the difference $H_{e+1} - H_e$ is either 0 or a monomial.*

Proof. For a fixed e we need to consider the three quotients HQ_e, HQ_{e+1} and HQ_{e+2}. Let $\sigma \geq \Sigma$ be the degree of HQ_e. Then, by Lemma 4.18, the degree of HQ_{e+1} is either σ or $\sigma + 1$ and the degree of HQ_{e+2} is between σ and $\sigma + 2$. We consider the following four cases and show that the result holds in each:

- $\deg HQ_{e+1} - \deg HQ_e = 0$ and $\deg HQ_{e+2} - \deg HQ_{e+1} = 0$. Then we have the quotients:

$$\begin{aligned} HQ_e &= a_0 + \cdots + a_{e-1}t^{e-1} + (a_e - a_0)t^e + \cdots + (a_\sigma - a_{\sigma-e})t^\sigma, \\ HQ_{e+1} &= a_0 + \cdots + a_e t^e + (a_{e+1} - a_0)t^{e+1} + \cdots + (a_\sigma - a_{\sigma-e-1})t^\sigma, \\ HQ_{e+2} &= a_0 + \cdots + a_{e+1}t^{e+1} + (a_{e+2} - a_0)t^{e+2} + \cdots + (a_\sigma - a_{\sigma-e-2})t^\sigma. \end{aligned}$$

  This gives the sections:

$$\begin{aligned} H_e &= a_0 + (a_1 - a_0)t + \cdots + (a_{\sigma-e-1} - a_{\sigma-e-2})t^{\sigma-e-1} \\ &\quad + (a_{\sigma-e} - a_{\sigma-e-1})t^{\sigma-e}, \\ H_{e+1} &= a_0 + (a_1 - a_0)t + \cdots + (a_{\sigma-e-1} - a_{\sigma-e-2})t^{\sigma-e-1}. \end{aligned}$$

  Therefore, the difference is:

$$H_{e+1} - H_e = (a_{\sigma-e-1} - a_{\sigma-e})t^{\sigma-e}.$$

- $\deg HQ_{e+1} - \deg HQ_e = 1$ and $\deg HQ_{e+2} - \deg HQ_{e+1} = 0$. Then we have the quotients:

$$\begin{aligned} HQ_e &= a_0 + \cdots + a_{e-1}t^{e-1} + (a_e - a_0)t^e + \cdots + (a_\sigma - a_{\sigma-e})t^\sigma, \\ HQ_{e+1} &= a_0 + \cdots + a_e t^e + (a_{e+1} - a_0)t^{e+1} + \cdots + (a_\sigma - a_{\sigma-e-1})t^\sigma \\ &\quad + (a_{\sigma+1} - a_{\sigma-e})t^{\sigma+1}, \\ HQ_{e+2} &= a_0 + \cdots + a_{e+1}t^{e+1} + (a_{e+2} - a_0)t^{e+2} + \cdots + (a_\sigma - a_{\sigma-e-2})t^\sigma \\ &\quad + (a_{\sigma+1} - a_{\sigma-e-1})t^{\sigma+1}. \end{aligned}$$

This gives the sections:

$$\begin{aligned} H_e &= a_0 + (a_1 - a_0)t + \cdots + (a_{\sigma-e} - a_{\sigma-e-1})t^{\sigma-e} \\ &\quad + (a_{\sigma+1} - a_{\sigma-e})t^{\sigma-e+1}, \\ H_{e+1} &= a_0 + (a_1 - a_0)t + \cdots + (a_{\sigma-e} - a_{\sigma-e-1})t^{\sigma-e}. \end{aligned}$$

Therefore, the difference is:

$$H_{e+1} - H_e = (a_{\sigma-e} - a_{\sigma+1})t^{\sigma-e+1}.$$

- $\deg HQ_{e+1} - \deg HQ_e = 0$ and $\deg HQ_{e+2} - \deg HQ_{e+1} = 1$. Then we have the quotients:

$$\begin{aligned} HQ_e &= a_0 + \cdots + a_{e-1}t^{e-1} + (a_e - a_0)t^e + \cdots + (a_\sigma - a_{\sigma-e})t^\sigma, \\ HQ_{e+1} &= a_0 + \cdots + a_e t^e + (a_{e+1} - a_0)t^{e+1} + \cdots + (a_\sigma - a_{\sigma-e-1})t^\sigma, \\ HQ_{e+2} &= a_0 + \cdots + a_{e+1}t^{e+1} + (a_{e+2} - a_0)t^{e+2} + \cdots + (a_\sigma - a_{\sigma-e-2})t^\sigma \\ &\quad + (a_{\sigma+1} - a_{\sigma-e-1})t^{\sigma+1}. \end{aligned}$$

This gives the sections:

$$\begin{aligned} H_e &= a_0 + (a_1 - a_0)t + \cdots + (a_{\sigma-e-1} - a_{\sigma-e-2})t^{\sigma-e-1} \\ &\quad + (a_{\sigma-e} - a_{\sigma-e-1})t^{\sigma-e}, \\ H_{e+1} &= a_0 + (a_1 - a_0)t + \cdots + (a_{\sigma-e-1} - a_{\sigma-e-2})t^{\sigma-e-1} \\ &\quad + (a_{\sigma+1} - a_{\sigma-e-1})t^{\sigma-e}. \end{aligned}$$

Therefore, the difference is:

$$H_{e+1} - H_e = (a_{\sigma+1} - a_{\sigma-e})t^{\sigma-e}.$$

- $\deg HQ_{e+1} - \deg HQ_e = 1$ and $\deg HQ_{e+2} - \deg HQ_{e+1} = 1$. Then we have the quotients:

$$\begin{aligned} HQ_e &= a_0 + \cdots + a_{e-1}t^{e-1} + (a_e - a_0)t^e + \cdots + (a_\sigma - a_{\sigma-e})t^\sigma, \\ HQ_{e+1} &= a_0 + \cdots + a_e t^e + (a_{e+1} - a_0)t^{e+1} + \cdots + (a_\sigma - a_{\sigma-e-1})t^\sigma \\ &\quad + (a_{\sigma+1} - a_{\sigma-e})t^{\sigma+1}, \\ HQ_{e+2} &= a_0 + \cdots + a_{e+1}t^{e+1} + (a_{e+2} - a_0)t^{e+2} + \ldots \\ &\quad + (a_{\sigma+1} - a_{\sigma-e-1})t^{\sigma+1} + (a_{\sigma+2} - a_{\sigma-e})t^{\sigma+2}. \end{aligned}$$

This gives the sections:

$$\begin{aligned} H_e &= a_0 + (a_1 - a_0)t + \cdots + (a_{\sigma-e} - a_{\sigma-e-1})t^{\sigma-e} \\ &\quad + (a_{\sigma+1} - a_{\sigma-e})t^{\sigma-e+1}, \\ H_{e+1} &= a_0 + (a_1 - a_0)t + \cdots + (a_{\sigma-e} - a_{\sigma-e-1})t^{\sigma-e}, \\ &\quad + (a_{\sigma+2} - a_{\sigma-e})t^{\sigma-e+1}. \end{aligned}$$

Therefore, the difference is:

$$H_{e+1} - H_e = (a_{\sigma+2} - a_{\sigma+1})t^{\sigma-e+1}. \qquad \square$$

We now can translate these results to describe the DRL staircase. For all $e \geq 0$, the sections of the staircase will be denoted by

$$E^e = \{x_1^{i_1} \cdots x_{n-1}^{i_{n-1}} \mid x_1^{i_1} \cdots x_{n-1}^{i_{n-1}} x_n^e \in E\}.$$

We can now state and prove our structure result.

Proposition 4.20. *For all $b \in E$, either $x_n b \in E$ or $x_n b$ is a leading monomial in the reduced DRL Gröbner basis of I.*

Proof. Let $b \in E$ be a monomial of degree δ. Assume that $x_n b \notin E$. Let $b' \in E^e$ so that $b = b' x_n^e$. The coefficient of the δth term of a Hilbert series is the number of monomials of degree δ under the staircase. Thus, b is accounted for in the δth term of HQ_{e+1}. Furthermore, since $x_n^e \mid b$, b is not accounted for in HQ_e and so in the section H_e, b' is accounted for in the $(\delta - e)$th coefficient. However, since $x_n^{e+1} \nmid b$, b is still accounted for in the δth term of HQ_{e+2}. Therefore, these parts cancel in the section H_{e+1} and so in the difference $H_{e+1} - H_e$, b' is accounted for in the $(\delta - e)$th term. The absolute value of the sum of the coefficients of this difference gives the number of monomials that are in E^e that are not in E^{e+1}. By Lemma 4.19, $H_{e+1} - H_e$ is a monomial. Therefore, all monomials that are in E^e and are not in E^{e+1} are of the same degree and so are independent. The monomial b' is accounted for in the coefficient of $H_{e+1} - H_e$ and so $x_n b$ is a leading monomial in the reduced DRL Gröbner basis of I. □

Theorem 2.3. *Let I be a generic determinantal sum ideal so that the conditions of Definition 2.2 hold. Assume that a reduced Gröbner basis of I with respect to a DRL ordering is known. Then the multiplication matrix T_{x_n} can be constructed without performing any arithmetic operations.*

Proof. Each column of the matrix T_{x_n} is the normal form of a monomial $x_n b$ such that $b \in E$. By Lemma 4.20, either $x_n b \in E$, in which case the column is all zeroes except one entry with a value of 1 in the row corresponding to $x_n b$, or $x_n b$ is a leading term in the reduced DRL Gröbner basis of I. In the latter case, the normal form is obtained from the DRL Gröbner basis without cost. Therefore, the multiplication matrix T_{x_n} can be constructed for free. □

With this structure theorem in tow, we aim to count the number of non-trivial columns. The following lemma gives a useful classification of this number.

Lemma 4.21. *If Fröberg's conjecture is true, then the number of non-trivial columns of T_{x_n} is equal to the largest coefficient of H.*

Proof. By Theorem 2.3, we can count the number of non-trivial columns of T_{x_n} by counting the number of polynomials in the reduced and minimal DRL Gröbner basis whose leading terms have positive degree in x_n. Lemma 4.20 implies that this number is equal to the number of monomials $b \in E$ such that $x_n b \notin E$. Note that this number is also equal to the number of monomials in the section E^0. The monomials in this section form a monomial basis of the quotient algebra $(\mathbb{K}[\boldsymbol{x}]/I)/\langle x_n \rangle$. Thus, the number of non-trivial columns of T_{x_n} is equal to the sum of the coefficients of the Hilbert series HQ_1 of this algebra. By Lemma 4.4, we can express HQ_1 in terms of the coefficients of H:

$$HQ_1 = a_0 + (a_1 - a_0)t + \cdots + (a_\Sigma - a_{\Sigma-1})t^\Sigma.$$

Therefore, the sum of the coefficients of HQ_1, and so the number of non-trivial columns of T_{x_n}, equals a_Σ, the largest coefficient of H. □

4.3.4 Asymptotics

By [32], the complexity of the Sparse-FGLM algorithm depends linearly on the number of non-trivial columns of the multiplication matrix T_{x_n}, denoted q. In the previous section, we proved Lemma 4.21, meaning that we can determine this number by finding the largest coefficient of the Hilbert series H from Proposition 4.2. We consider two cases. Firstly, we suppose that $d = 2$. This assumption leads to a simplification of the Hilbert series so that, by Corollary 4.10 and a trivial identity, it can be written as

$$H = \left(\sum_{k=0}^{m-1} \binom{n-m-1+k}{k} t^k \right) (1+t)^m.$$

On the other hand, for any $d \geq 2$, to find an asymptotic formula for the largest coefficient of H we will consider the central coefficients of polynomials of the form $(1+t+\cdots+t^r)^s$ for some r, s. Therefore, we recall an abridged version of the following result from [103].

Proposition 4.22 [103, Theorem 2]**.** *Let $r, s \geq 1$ and choose $0 \leq k \leq s^{1/2}$. Then the $\frac{1}{2}(sr+k)$th coefficient of the polynomial $(1+t+\cdots+t^r)^s$ is asymptotically equal to*

$$\frac{1}{\sqrt{s\pi}}\sqrt{\frac{6}{r^2-1}}r^s\left(1+O\left(\frac{k}{s}\right)\right).$$

We can now restate and prove our main result.

Theorem 2.4. *Let I be a generic determinantal sum ideal so that the conditions of Definition 2.2 hold, and let T_{x_n} be the matrix associated to the linear map of multiplication by x_n. Denote by q the number of non-trivial columns of T_{x_n}. Then, for $d = 2$ and $n \gg m$,*

$$q = \sum_{k=0}^{m-1} \binom{n-m-1+k}{k}\binom{m}{\lfloor 3m/2 \rfloor - 1 - j}. \tag{2.1}$$

Moreover, for $d \geq 3$ and $n \to \infty$,

$$q \approx \frac{1}{\sqrt{(n-m)\pi}}\sqrt{\frac{6}{(d-1)^2-1}}d^m(d-1)^{n-m}\binom{n-2}{m-1}. \tag{2.2}$$

Proof. By Lemma 4.21, q is equal to the largest coefficient of the Hilbert series H. First, assume that $d = 2$. Then the Hilbert series can be written as

$$H = \left(\sum_{k=0}^{m-1}\binom{n-m-1+k}{k}t^k\right)(1+t)^m = \sum_{k=0}^{m(m-1)} h_k t^k.$$

In this setting, we consider the binomial coefficients:

$$(1+t)^m = \sum_{k=0}^{m}\binom{m}{k}t^k = \sum_{k=0}^{m} a_k t^k.$$

We shall prove our first result by finding the degree of the term of H with the largest coefficient. The number q can then be found by a convolution formula.

Firstly, note that $(1+t)^m$ is a symmetric unimodal polynomial. Therefore, its largest coefficient is at the term of degree $\lfloor \frac{p}{2} \rfloor$. Since this polynomial is unimodal,

$$h_{\lfloor \frac{3m}{2} \rfloor} = \sum_{k=0}^{m-1}\binom{n-2-k}{m-1-k}a_{\lfloor \frac{m}{2} \rfloor+1+k} \leq \sum_{k=0}^{m-1}\binom{n-2-k}{m-1-k}a_{\lfloor \frac{m}{2} \rfloor+k} = h_{\lfloor \frac{3m}{2} \rfloor-1}.$$

By Lemma 4.15, H is unimodal and so the largest coefficient of H is at least $h_{\lfloor \frac{3m}{2} \rfloor-1}$. We now show that the previous coefficient of H is also no more than $h_{\lfloor \frac{3m}{2} \rfloor-1}$. By unimodality, this shows that $h_{\lfloor \frac{3m}{2} \rfloor-1}$ is the largest coefficient. Hence,

$$h_{\lfloor \frac{3m}{2} \rfloor-1} = \sum_{k=0}^{m-1}\binom{n-m-1+k}{k}\binom{m}{\lfloor \frac{3m}{2} \rfloor - 1 - k}$$

and

$$h_{\lfloor \frac{3m}{2} \rfloor-2} = \sum_{k=0}^{m-1}\binom{n-m-1+k}{k}\binom{m}{\lfloor \frac{3m}{2} \rfloor - 2 - k}.$$

As $n \to \infty$ we can write this as:

$$h_{\lfloor \frac{3m}{2} \rfloor - 1} = \binom{n-2}{m-1}\binom{m}{\lfloor \frac{m}{2} \rfloor} + O(n^{m-2})$$

and

$$h_{\lfloor \frac{3m}{2} \rfloor - 2} = \binom{n-2}{m-1}\binom{m}{\lfloor \frac{m}{2} \rfloor - 1} + O(n^{m-2}).$$

Therefore,

$$h_{\lfloor \frac{3m}{2} \rfloor - 1} - h_{\lfloor \frac{3m}{2} \rfloor - 2} = \binom{n-2}{m-1}\left(\binom{m}{\lfloor \frac{m}{2} \rfloor} - \binom{m}{\lfloor \frac{m}{2} \rfloor - 1}\right) + O(n^{m-2})$$

If $m = 1$, then $H = 1 + t$, and so the largest coefficient is indeed $h_0 = 1$. Otherwise, $\binom{m}{\lfloor \frac{m}{2} \rfloor} > \binom{m}{\lfloor \frac{m}{2} \rfloor - 1}$ and so this difference tends to mositive infinity as $n \to \infty$.

Therefore, for sufficiently large n, the largest coefficient is $h_{\lfloor \frac{3m}{2} \rfloor - 1}$.

Suppose now that $d > 2$. We return to the Hilbert series form given in Corollary 4.10 along with a trivial identity:

$$H = \left(\sum_{k=0}^{m-1} \binom{n-m-1+k}{k} t^{k(d-1)}\right)(1 + t + \cdots + t^{d-1})^m (1 + t + \cdots + t^{d-2})^{n-m}.$$

Firstly, consider the binomial sum factor. Note that as $n \to \infty$, the dominant term is the term of highest degree. Specifically, we may write

$$\sum_{k=0}^{m-1} \binom{n-m-1+k}{k} t^{k(d-1)} = \binom{n-2}{m-1} t^{(m-1)(d-1)} + O(n^{m-2} t^{(m-2)(d-1)}).$$

Therefore, since we only consider the largest coefficient of H as $n \to \infty$, we see that this is equal to the largest coefficient of the polynomial

$$h = \binom{n-2}{m-1}(1 + t + \cdots + t^{d-1})^m (1 + t + \cdots + t^{d-2})^{n-m}.$$

Thus, we can replace the binomial sum in the expression we consider with just a binomial coefficient.

For ease of notation, denote the other factors of h by $f_1 = (1 + t + \cdots + t^{d-1})^m$ and $f_2 = (1 + t + \cdots + t^{d-2})^{n-m}$. By [79, Proposition 2.2], these polynomials are symmetric. In particular, this means that $a_i = a_{(d-2)(n-m)-i}$, where

$$f_2 = \sum_{i=0}^{(d-2)(n-m)} a_i t^i.$$

Then, by Lemma 4.13 the polynomial f_2 is unimodal and so its largest coefficient is the central one. Therefore, by Proposition 4.22, the largest coefficient of f_2 is asymptotically equal to

$$\frac{1}{\sqrt{(n-m)\pi}} \sqrt{\frac{6}{(d-1)^2 - 1}} (d-1)^{n-m}.$$

Also by Proposition 4.22, since $m(d-1)+1$ is fixed as $n \to \infty$, the central $m(d-1)+1$ coefficients of f_2 tend to its largest coefficient. Note that for sufficiently large n, the largest coefficient of the product $f_1 f_2$ depends only on the central $m(d-1)+1$ coefficients of f_2, since f_1 does not

depend on n. Therefore, since the sum of the coefficients of f_1 equals d^m, as $n \to \infty$, the largest coefficient of H is asymptotically equal to

$$\frac{1}{\sqrt{(n-m)\pi}}\sqrt{\frac{6}{(d-1)^2-1}}d^m(d-1)^{n-m}\binom{n-2}{m-1}.$$

We conclude that, for $d \geq 3$ and $n \to \infty$, the number of non-trivial columns of T_{x_n} is asymptotically equal to

$$q \approx \frac{1}{\sqrt{(n-m)\pi}}\sqrt{\frac{6}{(d-1)^2-1}}d^p(d-1)^{n-m}\binom{n-2}{m-1}. \qquad \square$$

Theorem 2.5. *Let I be a generic determinantal sum ideal so that the conditions of Definition 2.2 hold. Assume that a reduced DRL Gröbner basis of I is known. Then, for $d \geq 3$, the arithmetic complexity of computing a LEX Gröbner basis of I is upper bounded by*

$$O\left(\frac{d^{3m}(d-1)^{3(n-m)}}{\sqrt{(n-m)d\pi}}\binom{n-2}{m-1}\binom{n-1}{m-1}^2\right).$$

Hence, the complexity gain of Sparse-FGLM over FGLM for generic determinantal sum systems is approximately

$$O\left(\frac{q}{nD}\right) \approx O\left(\frac{\sqrt{n-m}}{n^2(d-1)}\right).$$

Proof. Firstly, by Definition 2.2, we may apply the shape position variant of the Sparse-FGLM algorithm. Assuming the multiplication matrix T_{x_n} is constructed, its complexity is $O(qD^2 + nD\log^2(D))$, where q is the number of non-trivial columns of the multiplication matrix T_{x_n} and D is the degree of the ideal I [32, Theorem 3.2]. By Theorem 2.3, the construction of the matrix T_{x_n} requires no arithmetic operations. Recall that the degree of the ideal I is equal to

$$D = d^m(d-1)^{n-m}\binom{n-1}{m-1}.$$

Then, for $d \geq 3$, by Theorem 2.4, as $n \to \infty$,

$$q \approx \frac{1}{\sqrt{(n-m)\pi}}\sqrt{\frac{6}{(d-1)^2-1}}d^m(d-1)^{n-m}\binom{n-2}{m-1}.$$

Since the dominant term of the complexity is $O(qD^2)$, substituting the formula for D and the asymptotics of q gives the complexity result.

The complexity gain is then

$$O\left(\frac{qD^2}{nD^3}\right) = O\left(\frac{q}{nD}\right) \approx O\left(\frac{\sqrt{n-m}}{n^2(d-1)}\right). \qquad \square$$

4.4 Experiments

In this section, we test the practical accuracy of our formulae in Theorem 2.4, for the number of dense columns of the multiplication matrix T_{x_n}. For $d = 2$ we use our exact formula (2.1), while for $d \geq 3$ we use the asymptotic formula (2.2). The matrix density refers to the number of non-zero entries of T_{x_n} divided by its total number of entries. As seen in Theorem 2.11, the matrix density gives an idea of the complexity gain of using Sparse-FGLM over FGLM for the change of ordering.

Table 4.1 originates as a cropped version of [32, Table 2]. There, the authors give the values in the "Actual" column, obtained by computing the multiplication matrix and calculating exactly

Parameters	Degree	Matrix Density		
(d, m, n)	D	Actual	Theoretical	Asymptotic
$(2, 4, 9)$	896	30.17%	30.80%	30.80%
$(2, 4, 10)$	1344	31.13%	31.77%	31.77%
$(2, 4, 11)$	1920	31.86%	32.50%	32.50%
$(3, 3, 6)$	2160	17.52%	18.52%	27.73%
$(3, 3, 7)$	6480	17.39%	18.31%	26.62%
$(3, 3, 8)$	18144	17.63%	18.72%	25.50%
$(4, 2, 5)$	1728	14.46%	15.45%	21.24%
$(4, 2, 6)$	6480	14.11%	15.13%	19.56%
$(5, 2, 5)$	6400	11.00%	11.94%	15.47%
$(6, 2, 5)$	18000	8.80%	9.63%	12.22%

Table 4.1 – Density of multiplication matrix T_{x_n} for generic critical point systems

the number of non-zero entries, but the entries in the theoretical and asymptotic columns were blank. Now, with Theorem 2.4 we can complete this table, and we put the new entries in blue. The entries of the theoretical and asymptotic columns are the values of q/D, approximately the density of non-zero entries, for the varying parameters. In the theoretical column, the value of q is taken to be the largest coefficient of the Hilbert series. Then for the asymptotic column we take q as in Theorem 2.4.

Exceptionally, in Figure 4.1, we consider the generic determinantal sum ideals defined by two quartics, and also the generic determinantal sum ideals defined by four polynomials of degree 8, with an increasing number of variables n.

Note that the number of dense columns increases exponentially with n in about the same exponent for either the theoretical or the asymptotic in both examples. On the other hand, the matrix density can have different behaviours as the number of variables n increases for different degrees d and number of polynomials m. However, in both examples we see that the asymptotic approximation of the matrix density is rather inaccurate for small n. But, for moderate n, the approximation becomes good.

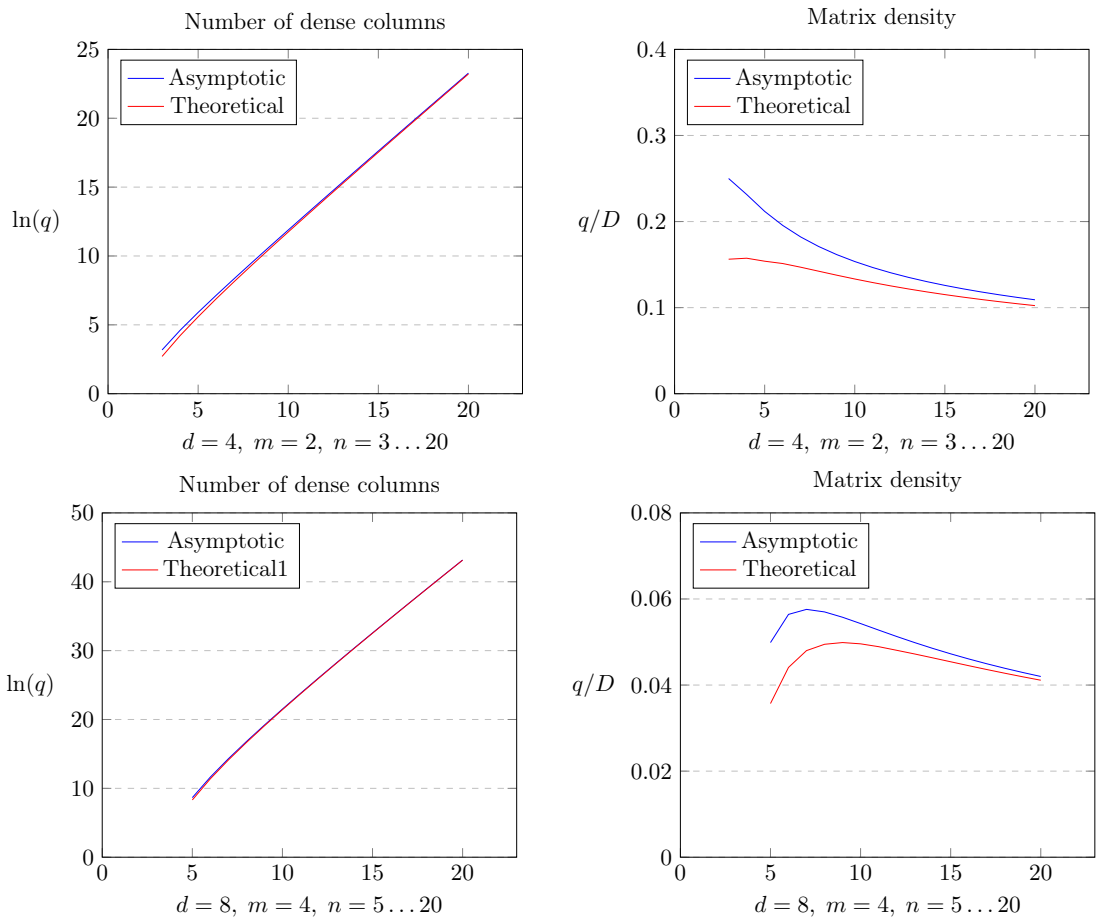


Figure 4.1 – Comparison of our asymptotic formulae against the theoretical number of dense columns and matrix density for generic critical point systems with parameters (d, m, n).

Chapter 5

Symmetric Determinantal ideals and Gröbner bases

Abstract. Polynomial matrices and ideals generated by their minors appear in various domains such as cryptography, polynomial optimization and effective algebraic geometry. When the given matrix is symmetric, this additional structure on top of the determinantal structure, affects computations on the derived ideals. Thus, understanding the complexity of these computations is important. Moreover, this study serves as a stepping stone towards further understanding the effects of structure in determinantal systems, such as those coming from moment matrices. In this chapter, we focus on the Sparse-FGLM algorithm, the state-of-the-art for changing ordering of Gröbner bases of zero-dimensional ideals. Under a variant of Fröberg's conjecture, we study its complexity for symmetric determinantal ideals and identify the gain of exploiting sparsity in the Sparse-FGLM algorithm compared with the classical FGLM algorithm. For an $\ell \times \ell$ symmetric matrix with polynomial entries of degree d, we show that the complexity of Sparse-FGLM for zero-dimensional determinantal ideals obtained from this matrix over that of the FGLM algorithm is at least $O(1/d)$. Moreover, for some specific sizes of minors, we prove finer results of at least $O(1/\ell d)$ and $O(1/\ell^3 d)$.

This chapter contains joint work with H. P. Le and led to the publication [35].

5.1 Introduction

Let $\mathbb{K}$ be a field of characteristic 0 and $\overline{\mathbb{K}}$ denote its algebraic closure. We consider a set of variables $\boldsymbol{x} = (x_1, \ldots, x_n)$ and an $\ell \times \ell$ symmetric matrix $S = (f_{i,j})_{1 \leq i,j \leq \ell}$ where $f_{i,j} \in \mathbb{K}[x_1, \ldots, x_n]$ and $f_{i,j} = f_{j,i}$. Given $r \in \mathbb{N}$, the ideal generated by all $(r+1)$-minors of S defines an algebraic subset of $\overline{\mathbb{K}}^n$ at which S has rank at most r. We call such an ideal a *symmetric determinantal ideal*.

Polynomial matrices with special structures such as those above appear frequently in computer algebra. For example, determinantal ideals arise in cryptography especially through the Min-Rank problem (see e.g. [34]). Additionally, critical point methods in effective algebraic geometry often lead to polynomial systems defined by minors of Jacobian matrices. Symbolic computation based methods for semi-definite programming, such as in [52, 53, 54, 82], lead to the study of rank defects of polynomial matrices, including symmetric and Hankel ones. In [69, 70], an algorithm for solving parametric polynomial systems is developed based on parametric Hermite matrices which are symmetric matrices that encode the numbers of real/complex solutions to zero-dimensional parametric systems. Determinantal ideals obtained from those Hermite matrices define algebraic sets such that the parametric system under study has at most a given number of distinct complex solutions.

Thus, a task of great importance in the aforementioned works is to handle computations involving determinantal ideals efficiently and to understand the complexity of those computations. The Gröbner basis method for computing with ideals is commonly used. The most efficient

Gröbner basis algorithms include the $\mathsf{F_4}/\mathsf{F_5}$ [28, 29], FGLM [31] and Sparse-FGLM [32] algorithms. In this chapter, we study the complexity of the Sparse-FGLM algorithm [32] on zero-dimensional ideals generated by minors of symmetric polynomial matrices. Our main objective is to provide finer complexity estimates for these algorithms on special determinantal ideals compared to already known general complexity results.

Related works. Ideals generated by minors of a matrix whose entries are variables are studied intensively in commutative algebra. A popular technique in this subject is to use the theory of Gröbner bases to associate initial ideals of determinantal ideals (w.r.t. a suitable ordering) to simplicial complexes. This allows one to make a connection between determinantal ideals with combinatorial objects and establish many results using the Stanley-Reisner rings of those simplicial complexes (see e.g [16, 18, 19, 21, 22, 104]).

In this chapter, we are more interested in the computational aspects that arise when one considers matrices whose entries are multivariate polynomials. Computing with determinantal ideals generated by minors of these matrices gives rise to the question of estimating the complexity of Gröbner basis algorithms, e.g., $\mathsf{F_4}/\mathsf{F_5}$ [28, 29] and FGLM-like [31, 32] algorithms, to this class of ideals.

Previous works on the complexity of these algorithms depend on some regularity properties as well as some quantities of the given ideal that can be read from its Hilbert series. It is well-known that the practical behavior of Gröbner basis computation depends on the choice of monomial ordering. While Gröbner bases of lexicographical orderings provides many information on the solutions to a given system, algorithms like $\mathsf{F_4}/\mathsf{F_5}$ operate more efficiently for computing Gröbner bases w.r.t. graded reversed lexicographic (grevlex) orderings. Hence, a popular strategy for computing lexicographic Gröbner bases is to start with an easy ordering such as grevlex and then to apply a change of ordering algorithm. For this second step, the FGLM algorithm [31] can be used in the zero-dimensional case. Given a zero-dimensional ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ of degree D, the classical FGLM algorithm is based on linear algebra operations in $\mathbb{K}[x_1, \ldots, x_n]/I$ which has the structure of a $\mathbb{K}$-vector space of dimension D. This leads to a complexity of $O(nD^3)$.

However, the matrices representing linear maps of multiplication in the quotient ring used by the FGLM algorithm are sparse. In particular, the majority of the columns of the multiplication matrix T_{x_n} associated to the least variable x_n contain only one entry while the rest are dense. An improved variant of the FGLM algorithm that exploits this sparsity pattern was designed in [32] to obtain a more efficient change of ordering algorithm with better complexity results. With N the number of non-zero entries of T_{x_n}, the authors of [32] prove, under some genericity assumptions, the complexity $O(ND + nD\log(D)^2)$. Due to the structure of this multiplication matrix, one can bound N by qD, where q is its number of dense columns. When the input zero-dimensional system is generic, an asymptotic bound for q is given using the knowledge of the Hilbert series of the given system. Inspired by [32], there have been attempts to study the complexity of the Sparse-FGLM algorithm for systems with special structures, the main task being to estimate the sparsity of the multiplication matrices involved. Research in this direction was undertaken in [10]. Focusing on zero-dimensional ideals defining critical loci of polynomial maps restricted to algebraic sets, [10] introduces an explicit formula of the Hilbert series of those given ideals which significantly simplifies the formula given in [22]. This allows one to derive a sharp asymptotic bound for the number of non-zero entries of the multiplication matrix T_{x_n}, when the number of variables n tends to infinity. Applying this to the complexity result of the Sparse-FGLM algorithm allows one to improve the change-of-ordering complexity estimate for critical loci computation compared to [33], which relies on the classical FGLM algorithm. Computational experiments are also provided to support that theoretical bound. We continue in this direction by considering determinantal ideals obtained from symmetric matrices.

Besides the Sparse-FGLM algorithm which exploits the sparsity of multiplication matrices, other algorithms are also developed using fast linear algebra techniques to improve the classical FGLM. In particular, under certain assumptions, [30] and [83] present two algorithms of complexity

$O^{\sim}(D^{\omega})$ and $O(nD^{\omega}\log(D))$ respectively, where ω is the exponent of matrix multiplication, for changing ordering of Gröbner bases. The best known theoretical bound for ω is 2.37286 given in [3]. Comparing the Sparse-FGLM algorithm with these algorithms requires estimating the parameter q. Moreover, a bound on q serves independently as an indicator for the sparsity of T_{x_n} and could be useful for any algorithm that relies on this sparsity (e.g., the algorithm of [12] that improves [30, 83]).

Main results. Our main result is a refined complexity of the Sparse-FGLM algorithm for zero-dimensional symmetric determinantal systems by bounding the aforementioned parameter q.

We consider a symmetric matrix $S = (s_{i,j})_{1\leq i,j\leq \ell}$ where $\boldsymbol{s} = (s_{1,1}, s_{2,1}, s_{2,2}, \ldots, s_{\ell,1}, \ldots, s_{\ell,\ell})$ are variables. Let $\mathbb{K}[\boldsymbol{s}]_d$ denote the set of homogeneous polynomials of degree d in $\mathbb{K}[\boldsymbol{s}]$ and 0.

Given $r \in \mathbb{N}$, $\mathcal{S}_r$ denotes the ideal generated by all $(r+1)$-minors of S and $A_r = \mathbb{K}[\boldsymbol{s}]/\mathcal{S}_r$. The Hilbert series of A_r is defined as

$$\mathrm{HS}_{A_r}(t) = \sum_{d=0}^{\infty} \dim_{\mathbb{K}} \mathbb{K}[\boldsymbol{s}]_d / \left(\mathcal{S}_r \cap \mathbb{K}[\boldsymbol{s}]_d\right) \cdot t^d$$

where $\dim_{\mathbb{K}}$ means the dimension as a $\mathbb{K}$-vector space. It is well-known that $\mathrm{HS}_{A_r}(t)$ can be written in the form

$$\mathrm{HS}_{A_r}(t) = \frac{\mathcal{H}_r(t)}{(1-t)^{\delta}}$$

where δ is the Krull dimension of A_r and $\mathcal{H}_r(t) \in \mathbb{Z}[t]$ such that $\mathcal{H}_r(1) \neq 0$ [24, Theorem 10.2.4] [27, Ch. 8]. We call this polynomial $\mathcal{H}_r(t)$ the reduced numerator of $\mathrm{HS}_{A_r}(t)$.

For $d \in \mathbb{N}$ and $k = \binom{\ell-r+1}{2}$, $\mathbb{K}[x_1, \ldots, x_n]_{\leq d}$ denotes the set of polynomials in $\mathbb{K}[x_1, \ldots, x_n]$ of degree at most d and $S^{n,d}$ be the symmetric matrix where $s_{i,j}$ are replaced by $f_{i,j} \in \mathbb{K}[x_1, \ldots, x_n]_{\leq d}$. For sufficiently generic $f_{i,j}$, we will prove that the ideal $\mathcal{S}_r^{n,d}$ generated by the $(r+1)$-minors of $S^{n,d}$ is zero-dimensional.

For any ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ (not necessarily homogeneous), let I^{h} be the homogenized ideal of I with a new variable x_0. The Hilbert series of $\mathbb{K}[x_1, \ldots, x_n]/I$ is defined as the Hilbert series of $\mathbb{K}[x_0, \ldots, x_n]/(I^{\mathrm{h}} + \langle x_0 \rangle)$ in the homogeneous setting. Let $\mathcal{S}_r^{n,d,\mathrm{h}}$ be the homogenization of $\mathcal{S}_r^{n,d}$ and $\mathcal{H}_r^{n,d}$ be the Hilbert series of $\mathbb{K}[x_1, \ldots, x_n]/\mathcal{S}_r^{n,d}$. Our main results rely on some conditions on the Hilbert series associated to $\mathcal{S}_r$ and $\mathcal{S}_r^{n,d}$ below.

Definition 5.1. *A polynomial $\sum_{i=0}^{n} a_i t^i$ with non-negative coefficients and $a_n > 0$ is* unimodal *if there exists $N \in \mathbb{N}$, $N \leq n$ such that*

$$\begin{array}{lll} a_{i-1} \leq a_i \leq a_N & \text{for} & 1 \leq i \leq N, \\ a_N \geq a_i \geq a_{i+1} & \text{for} & N \leq i \leq n-1. \end{array}$$

Additionally, we require a condition on the cross-sections of the Hilbert series of $\mathcal{S}_r^{n,d}$. This conjecture is a determinantal variant of Fröberg's well-known conjecture [36] on the shape of the Hilbert series of ideals generated by generic polynomial sequences.

Conjecture 2.6. 1. *Given $r \in \mathbb{N}$, the reduced numerator $\mathcal{H}_r(t)$ of the Hilbert series of the symmetric determinantal ideal $\mathcal{S}_r$ is unimodal.*

2. *For $e \geq 1$, let $\mathcal{Q}_r^{n,d,e}$ be the Hilbert series of $\mathbb{K}[x_0, \ldots, x_n]/\left(\mathcal{S}_r^{n,d,\mathrm{h}} + \langle x_0, x_n^e \rangle\right)$. We conjecture that $\mathcal{Q}_r^{n,d,e} = \left[(1-t^e)\mathcal{H}_r^{n,d}(t)\right]_+$, which is the series $(1-t^e)\mathcal{H}_r^{n,d}(t)$ truncated at its first non-positive coefficient.*

Fröberg's conjecture and the second part of Conjecture 2.6 also relate to the strong Lefschetz property in homogeneous setting. A graded Artinian algebra A has the strong Lefschetz property if there exists a linear form u such that the Hilbert series of the quotient $A/\langle u^e \rangle$ is equal to

$[(1-t^e)\mathrm{HS}_A(t)]_+$ for any $e \geq 1$. We refer the interested readers to [77] for a survey on this subject.

To support Conjecture 2.6 , we refer in Section 5.5 to a computational database for testing the two conditions in our conjecture.

Throughout this chapter, the notations $\prec_{\mathrm{DRL}}$ and $\prec_{\mathrm{LEX}}$ always denote the grevlex and lexicographic orderings in $\mathbb{K}[x_1, \ldots, x_n]$ with $x_1 \succ \cdots \succ x_n$. We can now state our main results.

Theorem 2.7. *Given $r, \ell, d \in \mathbb{N}$ and $n = \binom{\ell-r+1}{2}$, there exists a non-empty Zariski-open subset $\mathscr{F}_r$ of $\mathbb{K}[x_1, \ldots, x_n]_{\leq d}^{\ell(\ell+1)/2}$ such that, when the entries of $S^{n,d}$ are taken in $\mathscr{F}_r$, the following holds:*

The ideal $\mathcal{S}_r^{n,d}$ is zero-dimensional and radical. When Conjecture 2.6 holds and a reduced Gröbner basis of $\mathcal{S}_r^{n,d}$ w.r.t. $\prec_{\mathrm{DRL}}$ is known, the matrix T_{x_n} of multiplication by x_n can be constructed without any arithmetic operations. Moreover, the number of dense columns of T_{x_n} equals the largest coefficient of the Hilbert series $\mathcal{H}_r^{n,d}$.

Through the Sparse-FGLM algorithm [32], Theorem 2.7 leads directly to a complexity result for the change-of-ordering to a $\prec_{\mathrm{LEX}}$ Gröbner basis for symmetric determinantal ideals.

Theorem 2.8. *Given $r, \ell, d \in \mathbb{N}$ and $n = \binom{\ell-r+1}{2}$, we consider the matrix $S^{n,d}$ with entries taken in the Zariski-open set $\mathscr{F}_r$ defined in Theorem 2.7. Assume that Conjecture 2.6 holds and the reduced Gröbner basis of $\mathcal{S}_r^{n,d}$ w.r.t. $\prec_{\mathrm{DRL}}$ is known. Then as $d \to \infty$, the Sparse-FGLM algorithm computes a $\prec_{\mathrm{LEX}}$ Gröbner basis of $\mathcal{S}_r^{n,d}$ within*

$$O\left(q\mathcal{H}_r^{n,d}(1)^2\right) = O\left(qd^{2n}\mathcal{H}_r(1)^2\right) = O\left(qd^{2n}\left(\prod_{i=0}^{\ell-r-1}\frac{\binom{\ell+i}{2i+r}}{\binom{2i+1}{i}}\right)^2\right)$$

arithmetic operations in $\mathbb{K}$ where q is the number of dense columns of the multiplication matrix T_{x_n}. Moreover, as $d \to \infty$, q is bounded above by

$$d^{n-1}\mathcal{H}_r(1) = \sqrt{\frac{6}{n\pi}}d^{n-1}\prod_{i=0}^{n-r-1}\frac{\binom{\ell+i}{2i+r}}{\binom{2i+1}{i}}.$$

Our results provide dedicated estimates of the complexity of the Sparse-FGLM algorithm for symmetric determinantal ideals. This new complexity result is finer than previous results that do not take the specific structure into account. Moreover, we focus on three special cases in particular, $n = \ell - 2, n = \ell - 3$ and $r = 1$. In these cases, the Hilbert series is known [19, 21]. This allows us to provide sharper complexity results by analyzing the largest coefficients of these Hilbert series. To illustrate this result, we provide some numerical results to compare this theoretical bound with the actual number of dense columns that is observed in practice.

Organization of the chapter. In Section 5.2, we recall some basic notions and known results for determinantal ideals that will be used further. The transition from variable matrices to polynomial matrices is described in Section 5.3. There, we prove some properties that relate the largest coefficient of the Hilbert series to the complexity of the Sparse-FGLM algorithm applied to symmetric determinantal ideals. Using these properties, in Section 5.4 we asymptotically bound said complexity, with sharper estimates in some special cases. Based on our findings, we touch on topics for further study, including triangular and moment matrices, in Section 5.6. Finally, in Section 5.5, experiments are provided to support our asymptotic bounds.

5.2 Preliminaries

In this section, we recall some properties of determinantal systems associated to symmetric matrices. In Section 5.3, we show that these properties can be transferred to determinantal ideals

generated by polynomial matrices. Under certain hypotheses, these properties serve as main ingredients for our complexity estimate of the Sparse-FGLM algorithm for symmetric determinantal ideals in Section 5.4.

We start with variable matrices before transitioning to the zero-dimensional setting. As in Section 7.1, consider a symmetric matrix $S = (s_{i,j})_{1\leq i,j\leq \ell}$ with entries the variables $\boldsymbol{s} = (s_{1,1}, s_{2,1}, s_{2,2}, \ldots, s_{\ell,1}, \ldots, s_{\ell,\ell})$. For $r \in \mathbb{N}$, $\mathcal{S}_r$ is the homogeneous ideal generated by all the $(r+1)$-minors of S and $A_r = \mathbb{K}[\boldsymbol{s}]/\mathcal{S}_r$. The reduced numerator of the Hilbert series of A_r is denoted by $\mathcal{H}_r(t)$.

By [63], the quotient ring $\mathbb{K}[\boldsymbol{s}]/\mathcal{S}_r$ is a Cohen-Macaulay normal domain. Moreover, we have the following properties:

- The Krull dimension of A_r is
$$\dim A_r = \binom{\ell+1}{2} - \binom{\ell-r+1}{2} = \frac{(2\ell+1-r)r}{2}.$$
- The degree of A_r, i.e. $\mathcal{H}_r(1)$, equals
$$\mathcal{H}_r(1) = \prod_{i=0}^{\ell-r-1} \frac{\binom{\ell+i}{2i+r}}{\binom{2i+1}{i}} \leq \frac{n^{\binom{\ell-r+1}{2}}}{2^{\binom{\ell-r}{2}} \prod_{i=1}^{\ell-r-1} i!}.$$

Now we discuss some particular cases when the numerator of the Hilbert series is unimodal (Definition 5.1). Note that unimodality is not necessarily preserved by multiplication, for example $f = 3 + t + t^2$ is unimodal (for $N = 0$) while $f^2 = 9 + 6t + 7t^2 + 2t^3 + t^4$ is not. This motivates the following definition.

Definition 5.2. *A polynomial f with non-negative coefficients is* strongly unimodal *if, for any unimodal polynomial g, the product fg is unimodal.*

For an $n \times m$, with $n \leq m$, general variable matrix, the authors of [10] simplify a formula given in [22] for the Hilbert series of the ideal generated by its maximal minors. The reduced numerator in this simplified formula of the Hilbert series,
$$\sum_{i=0}^{n-1} \binom{m-n+i}{i} t^i,$$
is easily seen to be unimodal. This allows one to derive the Hilbert series of ideals generated by the maximal minors of matrices whose entries are generic homogeneous polynomials of *the same degree d*. Using the strong unimodality of $1 + \cdots + t^{d-1}$, it is also proved in [10] that the corresponding reduced numerator is also unimodal.

In the case of symmetric matrices, we focus on the following special cases for which the Hilbert series are known [19, 21]:

- When $r = \ell - 2$, the Hilbert series of $\mathcal{S}_{\ell-2}$ is
$$\frac{1}{(1-t)^{\ell(\ell+1)/2-3}} \sum_{i=0}^{\ell-2} \binom{i+2}{2} t^i.$$
- When $r = \ell - 3$, the Hilbert series of $\mathcal{S}_{n-3}$ is symmetric
$$\frac{1}{(1-t)^{\ell(\ell+1)/2-6}} \left(\sum_{i=0}^{\ell-3} \binom{i+5}{5} t^i + \sum_{i=0}^{\ell-4} \binom{i+5}{5} t^{2\ell-6-i} \right).$$

- When $r = 1$, the Hilbert series of $\mathcal{S}_1$ is

$$\frac{1}{(1-t)^\ell} \sum_{i=0}^{\lfloor \frac{\ell}{2} \rfloor} \binom{\ell}{2i} t^i.$$

One can see that the reduced numerators of these Hilbert series are unimodal. However, except these cases, closed forms of the Hilbert series are unknown. Although whether all the reduced numerators are unimodal remains open, an affirmative answer can be observed experimentally for generic determinantal systems (see Section 5.5).

5.3 The zero-dimensional setting

As in [10, 33, 34], we are interested in studying the behavior of Gröbner basis computations for zero-dimensional systems. In this section, some properties of zero-dimensional ideals generated by minors of a symmetric polynomial matrix are established.

We denote by $\mathbb{K}[x_1, \ldots, x_n]_{\leq d}$ the subset of $\mathbb{K}[x_1, \ldots, x_n]$ of polynomials of degree *at most* d. Let $S^{k,d} = (f_{i,j})_{1 \leq i,j \leq \ell}$ be an $\ell \times \ell$ symmetric matrix with entries in $\mathbb{K}[x_1, \ldots, x_n]_{\leq d}$. Then, for $r \in \mathbb{N}$, $\mathcal{S}_r^{n,d}$ denotes the ideal generated by the $(r+1)$-minors of $S^{n,d}$. It is expected that when the entries of $S^{n,d}$ are sufficiently generic, the ideal $\mathcal{S}_r^{n,d}$ retains some of the structure of $\mathcal{S}_r$ defined in Section 5.2.

In order to apply the reasoning of [32] to generic symmetric determinantal ideals we require them to be in shape position. This means that for a $\prec_{\text{LEX}}$ ordering with x_n as the least variable, the $\prec_{\text{LEX}}$ Gröbner basis has the structure

$$\{x_1 - g_1(x_n), \ldots, x_{n-1} - g_{n-1}(x_n), g_n(x_n)\},$$

where for $1 \leq i \leq n-1$, $\deg g_i < \deg g_n = D$, the degree of I.

Proposition 5.3. *Let $r, d \in \mathbb{N}$, $\mathcal{H}_r(t)$ be the reduced numerator of the Hilbert series of the ideal $\mathcal{S}_r$ and $n = \binom{\ell-r+1}{2}$, the codimension of $\mathcal{S}_r$. There exists a non-empty Zariski-open subset $\mathscr{F}_r$ of $\mathbb{K}[x_1, \ldots, x_n]_{\leq d}^{\ell(\ell+1)/2}$ such that if the entries of the matrix $S^{n,d}$ are taken in $\mathscr{F}_r$, then the ideal $\mathcal{S}_r^{n,d}$ is radical and zero-dimensional and its Hilbert series $\mathcal{H}_r^{n,d}$ is equal to*

$$\left(1 + t + \cdots + t^{d-1}\right)^n \mathcal{H}_r\left(t^d\right).$$

Moreover, there exists a non-empty Zariski-open subset $\mathscr{O}$ of the set $\mathrm{GL}(n, \mathbb{K})$ of invertible $n \times n$ matrices such that, after applying any linear change of coordinates $A \in \mathscr{O}$, the ideal $\mathcal{S}_r^{n,d}$ is in shape position.

Proof. We start in a homogeneous setting with $\mathbb{K}[x_0, x_1, \ldots, x_n]_d$ denoting the subset of homogeneous polynomials of degree d in $\mathbb{K}[x_0, x_1, \ldots, x_n]$ together with 0. Let $S = (s_{i,j})_{1 \leq i,j \leq \ell}$ be an $\ell \times \ell$ symmetric matrix. Throughout this proof, $\mathcal{S}_r$ denotes the ideal of $\mathbb{K}[\boldsymbol{s}, x_0, \ldots, x_n]$ generated by the $(r+1)$-minors of S. By [63], $\mathbb{K}[\boldsymbol{s}, x_0, \ldots, x_n]/\mathcal{S}_r$ is a Cohen-Macaulay ring.

By giving the weighted degrees d and 1 for the variables $\boldsymbol{s}$ and $x_0, \ldots, x_n$ respectively, the Hilbert series of $\mathbb{K}[\boldsymbol{s}, x_0, \ldots, x_n]/\mathcal{S}_r$ is

$$\tilde{\mathcal{H}_r}^{\mathrm{h}}(t) = \frac{\mathcal{H}_r(t^d)}{(1-t)^{k+1}\left(1-t^d\right)^{\ell(\ell+1)/2-k}}.$$

Assume that $f_{i,j} \in \mathbb{K}[x_1, \ldots, x_n]_d$. Let $f_{i,j}^{\mathrm{h}}$ be the homogenization of $f_{i,j}$ in $\mathbb{K}[x_0, \ldots, x_n]$. We consider the quasi-homogeneous ideal

$$J = \mathcal{S}_r + \langle s_{i,j} - f_{i,j}^{\mathrm{h}} | 1 \leq i \leq j \leq \ell \rangle.$$

Through similar techniques as in [34, Sec. 3 and 4], there exists a non-empty Zariski-open subset $\mathscr{Z}$ of $\mathbb{K}[x_0,\dots,x_n]_d^{\ell(\ell+1)/2}$ such that when the polynomials $f_{i,j}^{\mathrm{h}}$ lie in $\mathscr{Z}$, the ideals J and $J+\langle x_0\rangle$ have dimension one and zero respectively. Since $\mathcal{S}_r^{n,d}$ is the dehomogenized ideal of J, it has dimension zero. Moreover, by the unmixedness theorem [27, Cor. 18.14], the $\ell(\ell+1)/2+1$ polynomials

$$s_{i,j}-f_{i,j}^{\mathrm{h}} \text{ for } 1\le i\le j\le n \text{ and } x_0$$

forms a regular sequence over $\mathbb{K}[\boldsymbol{s},x_0,\dots,x_n]/\mathcal{S}_r$. Therefore, the Hilbert series of $J+\langle x_0\rangle$ is equal to

$$(1-t^d)^{\frac{\ell(\ell+1)}{2}}(1-t)\tilde{\mathcal{H}_r}^{\mathrm{h}}(t)=\left(1+\cdots+t^{d-1}\right)^n\mathcal{H}_r\left(t^d\right).$$

Next, we prove that there exists a non-empty Zariski-open subset $\mathscr{J}\subset\mathbb{K}[x_0,\dots,x_n]_d^{\ell(\ell+1)/2}$ such that, for $f_{i,j}^{\mathrm{h}}\in\mathscr{J}$, J is radical. By [18, Theorem 2.9], there exists a monomial ordering $\prec$ such that the corresponding initial ideal $in_\prec(\mathcal{S}_r)$ is generated by square-free monomials and so, is radical. Thus, $\mathcal{S}_r$ is a radical ideal of codimension $\binom{\ell-r+1}{2}$. Fixing an r-minor $\mathfrak{m}$ of S, we consider the set $\mathfrak{M}$ of the $\binom{n-r+1}{2}$ $(r+1)$-minors that contain $\mathfrak{m}$ as a submatrix. As the ideal $\mathcal{S}_r$ is radical, so is the ideal generated by the minors $\mathfrak{M}$. By the exchange lemma [4, Lemma 4], these minors, together with $\mathfrak{m}\neq 0$, define the locally closed algebraic set $V(\mathcal{S}_r)\setminus V(\mathfrak{m})$, which has codimension $\binom{\ell-r+1}{2}$.

We now consider the coefficients of $f_{i,j}^{\mathrm{h}}$ as new variables $\boldsymbol{c}$ in the space $\mathscr{C}=\overline{\mathbb{K}}[x_0,\dots,x_n]_d^{\ell(\ell+1)/2}$. Define the map φ by

$$\begin{array}{rccc}\varphi: & \overline{\mathbb{K}}^{\binom{\ell+1}{2}+k}\times\mathscr{C} & \to & \overline{\mathbb{K}}^{\binom{\ell-r+1}{2}}\times\overline{\mathbb{K}}^{\binom{\ell+1}{2}}\\ & (\boldsymbol{s},\boldsymbol{x},\boldsymbol{c}) & \mapsto & (\mathfrak{M},s_{1,1}-f_{1,1}^{\mathrm{h}},\dots,s_{\ell,\ell}-f_{\ell,\ell}^{\mathrm{h}})\end{array}$$

and $\varphi_{\boldsymbol{c}}$ denotes the restriction of the map φ to a given $\boldsymbol{c}\in\mathscr{C}$. Let $\mathrm{jac}_{\boldsymbol{s}}(\mathfrak{M})$ be the Jacobian matrix of $\mathfrak{M}$ w.r.t. $\boldsymbol{s}$. Note that the Jacobian matrix of φ has the following structure

$$\mathrm{jac}(\varphi)\coloneqq\begin{bmatrix}\mathrm{jac}_{\boldsymbol{s}}(\mathfrak{M}) & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ * & x_0^d\mathrm{Id} & \cdots & x_k^d\mathrm{Id} & \cdots\end{bmatrix},$$

where the blocks $\boldsymbol{x}_i^d\mathrm{Id}$ come from the derivatives of $s_{i,j}-f_{i,j}^{\mathrm{h}}$ w.r.t. the coefficients of x_i^d of $f_{i,j}^h$.

For any $\boldsymbol{s}$ such that $\mathfrak{m}(\boldsymbol{s})\neq 0$, $\mathrm{jac}_{\boldsymbol{s}}(\mathfrak{M})$, and therefore $\mathrm{jac}(\varphi)$, has maximal rank over the projective space of $(x_0,\dots,x_n)$. Thus, the Jacobian criterion [27, Theorem 16.19] implies that $\mathbf{0}$ is a regular value of φ. By Thom's weak transversality theorem [96, Proposition B.3], there exists a Zariski-open dense subset $\mathscr{C}_{\mathfrak{m}}$ of $\mathscr{C}$ such that for any $\boldsymbol{c}\in\mathscr{C}_{\mathfrak{m}}$, $\mathbf{0}$ is a regular value of $\varphi_{\boldsymbol{c}}$ and the Jacobian matrix of $\varphi_{\boldsymbol{c}}$ has maximal rank when $\mathfrak{m}(\boldsymbol{s})\neq 0$ and $(x_0,\dots,x_n)\neq\mathbf{0}$, which means J is radical. By dehomogenizing J, the ideal $\mathcal{S}_r^{n,d}$ is radical.

Now, let $\mathcal{S}_r^{n,d,\mathrm{h}}$ be the homogenized ideal of $\mathcal{S}_r^{n,d}$. The radicality of $\mathcal{S}_r^{n,d}$ implies that $\mathcal{S}_r^{n,d,\mathrm{h}}$ is also radical. As $J+\langle x_0\rangle$ has dimension zero, the projective varieties $V(J)$ and $V(\mathcal{S}_r^{n,d,\mathrm{h}})$ in $\mathbb{P}(\overline{\mathbb{K}})^n$ coincide. Since J is radical, the homogeneous Hilbert's Nullstellensatz [6, Corollary 4.80] gives $J=I(V(J))=I(V(\mathcal{S}_r^{n,d,\mathrm{h}}))=\mathcal{S}_r^{n,d,\mathrm{h}}$. Thus, the Hilbert series of $\mathcal{S}_r^{n,d}$ equals the Hilbert series of $J+\langle x_0\rangle$ whose explicit form is already proven above.

Finally, let $\mathscr{F}_r$ be the intersection of $\mathscr{Z}$ and $\mathscr{J}$, identified as a Zariski-open dense subset of $\mathbb{K}[x_1,\dots,x_n]_{\le d}^{\ell(\ell+1)/2}$ by specializing x_0 to one, with the sets $\mathscr{C}_{\mathfrak{m}}$ for all r-minors $\mathfrak{m}$ of S. For any $\boldsymbol{c}\in\mathscr{F}_r$, the ideal $\mathcal{S}_r^{n,d}$ is zero-dimensional and radical as the Jacobian matrix associated to its defining equations has rank $\binom{\ell+1}{2}+\binom{\ell-r+1}{2}$. Therefore, we may apply the shape lemma [8, Proposition 5]. There exists a Zariski-open dense subset $\mathscr{O}$ of $\mathrm{GL}(k,\mathbb{K})$ such that for all $A\in\mathscr{O}$, after applying A, the points of the variety $\mathbf{V}(\mathcal{S}_r^{n,d})$ have distinct x_n coordinates. Thus, the ideal $\mathcal{S}_r^{n,d}$ is in shape position. □

5.4 Asymptotic complexity

5.4.1 The general case

Given a Gröbner basis of a zero-dimensional ideal in $\mathbb{K}[x_1, \ldots, x_n]$ w.r.t. an ordering $\prec_1$, the Sparse-FGLM algorithm [32] computes a Gröbner basis of the same ideal but w.r.t. a target ordering $\prec_2$. A common change of ordering for practical uses is from a grevlex ordering to a lexicographic one [23, 33, 34]. In this section, we prove an asymptotic upper bound on the complexity of this computation for zero-dimensional symmetric determinantal ideals.

We keep the same setting as in Section 5.3. Given $\ell, r \in \mathbb{N}$ and $n = \binom{\ell-r+1}{2}$, we consider an $\ell \times \ell$ symmetric matrix $S^{n,d}$ whose entries are taken in $\mathscr{F}_r \subset \mathbb{K}[x_1, \ldots, x_n]_{\leq d}^{\ell(\ell+1)/2}$ defined by Proposition 5.3. Then, the ideal $\mathcal{S}_r^{n,d}$ is zero-dimensional and in shape position.

Given a zero-dimensional ideal $I \subset \mathbb{K}[x_1, \ldots, x_n]$ of degree D, let $\mathcal{G}$ be its reduced Gröbner basis w.r.t. the ordering $\prec_{\text{DRL}}$. It is well known that $\mathbb{K}[x_1, \ldots, x_n]/I$ is a finite-dimensional vector space for which the set $\mathcal{B}$ of monomials irreducible by $\mathcal{G}$ forms a basis. The multiplications by $x_1, \ldots, x_n$ are linear maps of $\mathbb{K}[x_1, \ldots, x_n]/I$, whose matrix representations $T_{x_1}, \ldots, T_{x_n}$ in $\mathcal{B}$ appear with sparsity. The Sparse-FGLM algorithm [32] improves upon the classical FGLM algorithm [31], whose arithmetic complexity is $O(nD^3)$, by taking advantage of this sparsity. In [32], the authors also provide a careful complexity analysis of their algorithm. By assuming the widely accepted Moreno-Socías conjecture [79, Conjecture 4.1], they show that the matrix T_{x_n} can be obtained from $\mathcal{G}$ without additional cost. With q as the number of dense columns of T_{x_n}, when I is in shape position they bound the complexity of this algorithm by

$$O\left(qD^2 + nD\log^2 D\right).$$

This complexity analysis relies on the observation that there are three possible cases when one multiplies a monomial $b \in \mathcal{B}$ by x_n:

- $x_n \cdot b \in \mathcal{B}$: in this case, the associated column in T_{x_n} is $(0, \ldots, 0, 1, 0, \ldots, 0)$ where the row of 1 corresponds to $x_n \cdot b$.
- $x_n \cdot b$ is the leading monomial of some $g \in \mathcal{G}$: in this case, the column is easily obtained from the coefficients of $x_n \cdot b - g$.
- Otherwise, the column is non-trivial and requires a normal form reduction of $x_n \cdot b$ by $\mathcal{G}$ to compute its canonical representation in $\mathcal{B}$, i.e. the corresponding column in T_{x_n}.

The most dense columns of the matrix T_{x_n} correspond to the second and the third cases. Only the third case requires extra computation. If Moreno-Socías' conjecture holds, then the third case does not occur for generic polynomial systems [32]. Thus, the multiplication matrix T_{x_n} can be obtained without further computation. In [10], it is shown that under similar genericity assumptions, the third case does not occur for critical point systems either. We shall now prove that the same holds for generic symmetric determinantal ideals.

Theorem 2.7. *Given $r, \ell, d \in \mathbb{N}$ and $n = \binom{\ell-r+1}{2}$, there exists a non-empty Zariski-open subset $\mathscr{F}_r$ of $\mathbb{K}[x_1, \ldots, x_n]_{\leq d}^{\ell(\ell+1)/2}$ such that, when the entries of $S^{n,d}$ are taken in $\mathscr{F}_r$, the following holds:*

The ideal $\mathcal{S}_r^{n,d}$ is zero-dimensional and radical. When Conjecture 2.6 holds and a reduced Gröbner basis of $\mathcal{S}_r^{n,d}$ w.r.t. $\prec_{\text{DRL}}$ is known, the matrix T_{x_n} of multiplication by x_n can be constructed without any arithmetic operations. Moreover, the number of dense columns of T_{x_n} equals the largest coefficient of the Hilbert series $\mathcal{H}_r^{n,d}$.

Proof. The existence of the set $\mathscr{F}_r$ such that $\mathcal{S}_r^{n,d}$ is zero-dimensional and radical is given by Proposition 5.3. By the first item of Conjecture 2.6, $\mathcal{H}_r$ is unimodal, which then implies that

$(1+\cdots+t^{d-1})\mathcal{H}_r(t^d)$ is unimodal. By [10, Lemma 17], $1+\cdots+t^{d-1}$ is a strongly unimodal polynomial. Hence, the Hilbert series $\mathcal{H}_r^{n,d}$ given in Proposition 5.3 is also unimodal.

Let $A_r^{k,d}=\mathbb{K}[x_1,\ldots,x_n]/\mathcal{S}_r^{n,d}$ and $\mathcal{S}_r^{n,d,\mathrm{h}}$ be the homogenization of $\mathcal{S}_r^{n,d}$. We shall construct the matrix T_{x_n} column by column. As in [32], the columns are indexed by elements in the basis $\mathcal{B}$ of $A_r^{n,d}$, given by the ordering $\prec_{\mathrm{DRL}}$. For any $b\in\mathcal{B}$, the entries in its corresponding column are the coefficients of the normal form of $x_n\cdot b$ expressed in terms of the basis $\mathcal{B}$. By [10, Theorem 1], under Conjecture 2.6, $x_n\cdot b$ is either an element of $\mathcal{B}$ or a leading monomial of the known grevlex Gröbner basis $\mathcal{G}$. In the first case, the column corresponding to b is a column of the identity matrix and requires no computation. In the second case, the column corresponding to b can be read from the coefficients of the polynomial in $\mathcal{G}$ for which $x_n\cdot b$ is the leading monomial. Thus, there are bijections between the dense columns of T_n, the polynomials in $\mathcal{G}$ whose leading terms are divisible by x_n and then the elements in $\mathcal{B}$ which are not divisible by x_n. The cardinal of the last set equals the dimension of $\mathbb{K}[x_0,\ldots,x_n]/\left(\mathcal{S}_r^{n,d,\mathrm{h}}+\langle x_0,x_n\rangle\right)$ which can be read by evaluating $\mathcal{Q}_r^{n,d,1}(1)$. When Conjecture 2.6 holds, similar to [10, Lemma 25], we deduce that the largest coefficient of $\mathcal{H}_r^{n,d}$ equals $\mathcal{Q}_r^{n,d,1}(1)$. □

Hence, assuming that $n=\binom{\ell-r+1}{2}$ and that the entries of $S^{n,d}$ are taken from $\mathscr{F}_r$ described in Proposition 5.3, we study the asymptotic behavior of the largest coefficient of the Hilbert series of the zero-dimensional ideal $\mathcal{S}_r^{n,d}$ as d tends to infinity.

Lemma 5.4. *Let $n=\binom{\ell-r+1}{2}$. The largest coefficient of*

$$\mathcal{H}_r^{n,d}(t)=(1+t+\cdots+t^{d-1})^n\mathcal{H}_r(t)$$

as $d\to\infty$ is bounded above by

$$\sqrt{\frac{6}{n\pi}}d^{n-1}\mathcal{H}_r(1)=\sqrt{\frac{6}{n\pi}}d^{n-1}\prod_{i=0}^{\ell-r-1}\frac{\binom{\ell+i}{2i+r}}{\binom{2i+1}{i}}.$$

Proof. By [32, Corollary 5.10], as $d\to\infty$, all the coefficients of $\left(1+\cdots+t^{d-1}\right)^n$ are bounded by $\sqrt{\frac{6}{n\pi}}d^{n-1}$. Substituting this asymptotic formula into the convolution formula for the largest coefficient gives the first result. By [63], we conclude using the equation

$$\mathcal{H}_r(1)=\prod_{i=0}^{\ell-r-1}\frac{\binom{\ell+i}{2i+r}}{\binom{2i+1}{i}}.$$ □

We now apply Lemma 5.4 to prove Theorem 2.8 which provides an asymptotic complexity estimate for the Sparse-FGLM algorithm on generic symmetric determinantal systems.

Theorem 2.8. *Given $r,\ell,d\in\mathbb{N}$ and $n=\binom{\ell-r+1}{2}$, we consider the matrix $S^{n,d}$ with entries taken in the Zariski-open set $\mathscr{F}_r$ defined in Theorem 2.7. Assume that Conjecture 2.6 holds and the reduced Gröbner basis of $\mathcal{S}_r^{n,d}$ w.r.t. $\prec_{\mathrm{DRL}}$ is known. Then as $d\to\infty$, the Sparse-FGLM algorithm computes a $\prec_{\mathrm{LEX}}$ Gröbner basis of $\mathcal{S}_r^{n,d}$ within*

$$O\left(q\mathcal{H}_r^{n,d}(1)^2\right)=O\left(qd^{2n}\mathcal{H}_r(1)^2\right)=O\left(qd^{2n}\left(\prod_{i=0}^{\ell-r-1}\frac{\binom{\ell+i}{2i+r}}{\binom{2i+1}{i}}\right)^2\right)$$

arithmetic operations in $\mathbb{K}$ where q is the number of dense columns of the multiplication matrix T_{x_n}. Moreover, as $d\to\infty$, q is bounded above by

$$d^{n-1}\mathcal{H}_r(1)=\sqrt{\frac{6}{n\pi}}d^{n-1}\prod_{i=0}^{n-r-1}\frac{\binom{\ell+i}{2i+r}}{\binom{2i+1}{i}}.$$

Proof. By Proposition 5.3, we apply the shape position variant of the Sparse-FGLM algorithm. Then, by Theorem 2.7, the multiplication matrix T_{x_n} can be constructed without any additional arithmetic operations and the number of dense columns q equals the largest coefficient of the Hilbert series of $\mathcal{S}_r^{n,d}$. The dominant term in the complexity is $O(qD^2)$, where D is the degree of $\mathcal{S}_r^{n,d}$. This degree is given by the evaluation of the Hilbert series

$$\mathcal{H}_r^{n,d}(t) = \left(1 + t + \cdots + t^{d-1}\right)^n \mathcal{H}_r\left(t^d\right)$$

of $\mathcal{S}_r^{n,d}$ at one. By [63], the degree of $S_r^{n,d}$ is equal to

$$D = \mathcal{H}_r^{n,d}(1) = d^n \mathcal{H}_r(1) = d^n \prod_{i=0}^{\ell-r-1} \frac{\binom{\ell+i}{2i+r}}{\binom{2i+1}{i}}.$$

Finally, Lemma 5.4 implies the bound on q as $d \to \infty$. □

Corollary 5.5. *The complexity of the Sparse-FGLM algorithm over that of the FGLM algorithm for generic symmetric determinantal ideals as $d \to \infty$ is at least $O(1/d)$.*

5.4.2 Cases $r = n-2, r = n-3$ and $r = 1$

In this subsection, we treat the cases of $r = \ell - 2$, $r = \ell - 3$ and $r = 1$ separately. By taking into account the knowledge on the corresponding Hilbert series, the first item of Conjecture 2.6 holds in these cases. Furthermore, one can arrive at finer asymptotic estimates on the largest coefficient. Recall that the codimension of $\mathcal{S}_r$, and hence the number of variables we consider in the zero-dimensional setting, equals 3, 6 and $\binom{\ell}{2}$ for these cases respectively.

We start by identifying the largest coefficient of $\mathcal{H}_{\ell-2}^{3,d}$ exactly.

Proposition 5.6. *The largest coefficient of*

$$\mathcal{H}_{\ell-2}^{3,d}(t) = \left(1 + t + \cdots + t^{d-1}\right)^3 \sum_{i=0}^{\ell-2} \binom{i+2}{2} t^{id}$$

is the value of

$$\binom{\ell-1}{2}\binom{j+1}{2} + \binom{\ell}{2}\left(\binom{d+1}{2} + j(d-j-1)\right).$$

when j is any integer that minimises $\left|\frac{2\ell d - \ell - 2}{2(\ell+2)} - j\right|$.

Proof. Note that

$$\left(1 + t + \cdots + t^{d-1}\right) \sum_{i=0}^{\ell-2} \binom{i+2}{2} t^{id} = \sum_{i=0}^{\ell-2} \sum_{j=0}^{d-1} \binom{i+2}{2} t^{id+j}.$$

We write these coefficients in the following $d \times ((\ell-2)d - 1)$ grid:

t^0	$\cdots$	t^{d-1}	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$t^{(\ell-1)d-1}$	$\cdots$	$t^{(\ell-2)d-2}$
		1	$\cdots$	1	$\cdots$	$\cdots$	$\binom{\ell}{2}$	$\cdots$	$\binom{\ell}{2}$
	$\iddots$	$\vdots$	$\iddots$	$\iddots$	$\iddots$	$\iddots$	$\vdots$	$\iddots$	
1	$\cdots$	1	$\cdots$	$\cdots$	$\binom{\ell}{2}$	$\cdots$	$\binom{\ell}{2}$		

The coefficients of $\left(1 + t + \cdots + t^{d-1}\right)^2 \mathcal{H}_{\ell-2}(t)$ are the sums of columns of this grid, which are

$$\binom{i+2}{2}(j+1) + \binom{i+1}{2}(d-j-1).$$

Thus, the coefficients of $\left(1+t+\cdots+t^{d-1}\right)^3 \mathcal{H}_{\ell-2}(t)$ can be computed by summing all d consecutive columns of the above grid.

As $\binom{i+2}{2}$ is increasing as a sequence in i, the largest coefficient of $\mathcal{H}_{\ell-2}^{3,d}$ must be the coefficient of $t^{\ell d-j-2}$ for some $0 \le j \le d-1$. By a simple calculation, this coefficient can be expressed as

$$\begin{aligned}&\binom{\ell-1}{2}\binom{j+1}{2}+\binom{\ell}{2}\left(\binom{d+1}{2}+j(d-j-1)\right)\\&=C-\frac{(\ell-1)(\ell+2)}{16}\left(\frac{2\ell d-\ell-2}{\ell+2}-2j\right)^2\end{aligned}$$

where

$$C=\binom{\ell}{2}\binom{d+1}{2}+\frac{(\ell-1)(2\ell d-\ell-2)^2}{16(\ell+2)}$$

does not depend on j. Hence, to identify j, we minimize

$$\min_{j\in\mathbb{N},0\le j\le d-1}\left|\frac{2\ell d-\ell-2}{2(\ell+2)}-j\right|.$$

Let $\alpha=\frac{2\ell d-\ell-2}{2(\ell+2)}$, which lies in $[0,d-1/2)$ if $\ell\ge 2$. Then, to conclude the proof, we take j to be the nearest integer to α. □

Recall that D denotes the degree of the ideal under study. When $r=\ell-2$ we have that $D=\binom{d+1}{3}$. Since the complexity of the Sparse-FGLM algorithm over that of the FGLM algorithm is $O\left(\frac{qD^2}{nD^3}\right)=O\left(\frac{q}{nD}\right)$, Proposition 5.6 immediately implies the following corollary.

Corollary 5.7. *By the proof of Proposition 5.6, we can bound*

$$q\le C=\binom{\ell}{2}\binom{d+1}{2}+\frac{(\ell-1)(2\ell d-\ell-2)^2}{16(\ell+2)}.$$

Hence, the complexity of the Sparse-FGLM algorithm over that of the FGLM algorithm when $r=\ell-2$ is at least $O\left(\frac{1}{\ell d}\right)$.

Next, we consider $r=\ell-3$. Notice that the reduced numerator $\mathcal{H}_{\ell-3}$ is symmetric, i.e. $\mathcal{H}_{\ell-3}(t)=t^{\deg(h)}\mathcal{H}_{\ell-3}(1/t)$. The lemma below will be useful for proving a finer complexity in this case.

Lemma 5.8. *Let $f(t)$ be a unimodal symmetric polynomial. Then*

$$g(t)=(1+t+\cdots+t^{d-1})f(t)$$

is also unimodal and symmetric. Moreover, the c largest coefficients of $g(t)$ are combinations of the $d+c-1$ largest coefficients of $f(t)$. As a point of notation, if $f(t)$ has fewer than $d+c-1$ coefficients then we consider all other coefficients to be zero.

Proof. First, the unimodality of g comes from the strong unimodality of $1+t+\cdots+t^{d-1}$. The symmetry can be deduced from the equality

$$t^{\deg(g)}g(1/t)=\left(1+\cdots+t^{d-1}\right)t^{\deg(g)}f(1/t)=g(t).$$

Note that the coefficient of t^i in g is the sum of the coefficients of $t^{i-d+1},\ldots,t^i$ in f. As f is unimodal and symmetric, the largest coefficient of g is the sum of the d central coefficients of f. Since g is unimodal and symmetric, the c largest coefficients of g are consecutive and any of them is at most $\left\lceil\frac{c-1}{2}\right\rceil$ elements away from the central and thus largest coefficient. Hence, the c largest coefficients of g involve only the central $d+c-1$ coefficients of f. □

Proposition 5.9. *The largest coefficient of the Hilbert series*

$$\mathcal{H}_{\ell-3}^{6,d}(t) = \left(1 + t + \cdots + t^{d-1}\right)^6 \mathcal{H}_{\ell-3}\left(t^d\right)$$

as $d \to \infty$ is bounded above by

$$\left(\binom{\ell+2}{5} + 2\binom{\ell+1}{5} + 2\binom{\ell}{5}\right)\sqrt{\frac{1}{\pi}}d^5 \leq 5\binom{\ell+2}{5}\sqrt{\frac{1}{\pi}}d^5 \in O\left(\ell^5 d^5\right).$$

Proof. By Lemma 5.8, $(1+\cdots+t^{d-1})^6$ is unimodal and symmetric. And so is $\mathcal{H}_{\ell-3}$ from its explicit formula. Thus, by Lemma 5.8, the largest coefficient m of the Hilbert series $\mathcal{H}_{\ell-3}^{6,d}$, which is actually a polynomial, depends only on the central $5(d-1)+1$ coefficients of $(1+\cdots+t^{d-1})\mathcal{H}_{\ell-3}(t^d)$. This number then depends on at most the central 5 coefficients of the polynomial $\mathcal{H}_{\ell-3}(t)$.

By [32, Corollary 5.10], all coefficients of $(1+\cdots+t^{d-1})^6$ are at most $\sqrt{\frac{1}{\pi}}d^5$. Therefore, by the definition of $\mathcal{H}_{\ell-3}$ and its symmetry, we have that, as $d \to \infty$,

$$q \leq \left(\binom{\ell+2}{5} + 2\binom{\ell+1}{5} + 2\binom{\ell}{5}\right)\sqrt{\frac{1}{\pi}}d^5 \leq 5\binom{\ell+2}{5}\sqrt{\frac{1}{\pi}}d^5. \qquad \square$$

When $r = \ell - 3$, the ideal $\mathcal{S}_{\ell-3}^{k,d}$ has degree

$$D = \left(\binom{\ell+2}{6} + \binom{\ell+3}{6}\right)d^6 \in O(\ell^6 d^6).$$

By Proposition 5.9, the number of dense columns q lies in $O(\ell^5 d^5)$ as $d \to \infty$, which implies Corollary 5.10.

Corollary 5.10. *Let $r = \ell - 3$. As $d \to \infty$, the complexity improvement of the Sparse-FGLM algorithm over that of the FGLM algorithm for the generic symmetric determinantal ideal $S_{\ell-3}^{6,d}$ is at least $O\left(\frac{1}{\ell d}\right)$.*

Finally, in the case $r = 1$, the number of variables n is chosen to be equal to $\binom{\ell}{2}$. Since this depends on ℓ, we consider the asymptotic complexity as $\ell \to \infty$.

Proposition 5.11. *The largest coefficient of*

$$\mathcal{H}_1^{\binom{\ell}{2},d}(t) = \left(1 + t + \cdots + t^{d-1}\right)^{\binom{\ell}{2}} \sum_{i=0}^{\lfloor \frac{\ell}{2} \rfloor} \binom{n}{2i} t^{id}$$

as $\ell \to \infty$ is at most

$$\sqrt{\frac{6}{\binom{\ell}{2}\pi(d^2-1)}}d^{\binom{\ell}{2}}2^{\ell-1} \in O\left(\frac{2^{\ell-1}}{\ell}d^{\binom{\ell}{2}-1}\right).$$

Proof. As $(1+\cdots+t^{d-1})^{\binom{\ell}{2}}$ is symmetric and unimodal, its largest coefficient is central. By an abridged version of [103, Theorem 2], this largest coefficient is asymptotically equal to

$$\sqrt{\frac{6}{\binom{\ell}{2}\pi(d^2-1)}}d^{\binom{\ell}{2}}$$

as $\ell \to \infty$. Then, the largest coefficient of $\mathcal{H}_1^{\binom{\ell}{2},d}$ is at most

$$\sqrt{\frac{6}{\binom{\ell}{2}\pi(d^2-1)}}d^{\binom{\ell}{2}}\sum_{i=0}^{\lfloor \frac{\ell}{2} \rfloor}\binom{\ell}{2i}.$$

The following equality gives the result

$$\sum_{i=0}^{\lfloor \frac{\ell}{2} \rfloor}\binom{\ell}{2i} = \sum_{i=0}^{\lfloor \frac{\ell}{2} \rfloor}\left(\binom{\ell-1}{2i-1} + \binom{\ell-1}{2i}\right) = \sum_{i=0}^{\ell-1}\binom{\ell-1}{i} = 2^{\ell-1}. \qquad \square$$

As the degree of $\mathcal{S}_1^{\binom{\ell}{2},d}$ is $d^{\binom{\ell}{2}}2^{\ell-1}$, by applying Proposition 5.11 to Theorem 2.8 we arrive at the following corollary.

Corollary 5.12. *When $r = 1$ the degree of $\mathcal{S}_1^{\binom{\ell}{2},d}$ is $d^{\binom{\ell}{2}}2^{\ell-1}$.*

Therefore, the complexity of the Sparse-FGLM algorithm over that of the FGLM algorithm as $\ell \to \infty$ is at least

$$O\left(\frac{1}{n\ell d}\right) = O\left(\frac{1}{\binom{\ell}{2}\ell d}\right) = O\left(\frac{1}{\ell^3 d}\right).$$

Moreover, the bound on m in Theorem 2.8 implies that the complexity gain as $d \to \infty$ is also at least

$$O\left(\frac{1}{n^{3/2}d}\right) = O\left(\frac{1}{\binom{\ell}{2}^{3/2}d}\right) = O\left(\frac{1}{\ell^3 d}\right).$$

5.5 Experiments

5.5.1 Supporting Conjecture 2.6

This subsection reports on our testing of Conjecture 2.6 upon which our main results rely. Firstly, except for the cases $r \in \{1, \ell-2, \ell-3\}$ considered in Subsection 5.4.2, the unimodality of the Hilbert polynomials of generic symmetric determinantal ideals remains open in general. Moreover, for non-symmetric determinantal ideals, while a formula for the Hilbert series is known in the generic case [34], it is not proven to be unimodal.

Secondly, the second item of Conjecture 2.6 is not proven in any of the cases we consider. We test this conjecture by computing the leading monomials of the reduced Gröbner basis of a generic symmetric determinantal system I with Hilbert series P. Homogenizing this Gröbner basis, we obtain a Gröbner basis of the homogenized ideal I^{h} w.r.t. the $\prec_{\mathrm{DRL}}$ ordering where $x_1 \succ \cdots x_n \succ x_0$. Finally, adding $\langle x_0, x_n\rangle$ gives a Gröbner basis of $I^{\mathrm{h}} + \langle x_0, x_n\rangle$ w.r.t. the $\prec_{\mathrm{DRL}}$ ordering with $x_1 \succ \cdots \succ x_n \succ x_0$ [79, Lemma 1.9]. Then, we can compute the Hilbert series and compare this to the formula $[(1-t^e)P]_+$ to test the second item. The current status of testing this conjecture can be found at the following website: https://www-polsys.lip6.fr/~ferguson/conjecture_testing.html.

5.5.2 Asymptotics in practice

In this subsection, we compare the true density of the multiplication matrix T_{x_n} (Actual) against the percentage of dense columns (Theoretical) and the asymptotic bounds established in Section 5.4 (Asymptotic), following the notation of [32, Table 2].

We begin with $\ell \times \ell$ symmetric matrices with rank at most $r = \ell - 2$. We consider 3 variables and vary the size of the matrix and the degree of its entries. When the entries are sufficiently generic, this construction yields symmetric determinantal ideals of dimension zero. Figure 5.1 reports on the exact numbers of dense columns in the matrices T_{x_3} using Proposition 5.6.

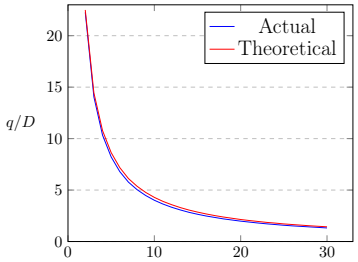


Figure 5.1 – Density of T_{x_3} for $\mathcal{S}_{\ell-2}^{3,d}$ for $d \in \{2, \ldots, 50\}$

In Table 5.1, we analyze the ideal $S_{\ell-3}^{6,d}$, where we also compare the matrix density and number of dense columns against the asymptotic bound obtained in Proposition 5.9 (Asymptotic). Additionally, Figure 5.2 illustrates how the asymptotic result approaches the true number of dense columns as the degree d increases.

Parameters	Degree	Matrix Density		
(d, ℓ)	D	Actual	Theoretical	Asymptotic
$(2, 5)$	2240	20.23%	21.96%	28.21%
$(3, 5)$	25515	12.58%	13.96%	18.81%
$(2, 6)$	7168	17.40%	19.14%	27.71%
$(3, 6)$	81648	10.89%	12.26%	18.47%
$(2, 7)$	18816	15.20%	16.96%	26.87%

Table 5.1 – Density of T_{x_6} for $\mathcal{S}_3^{6,d}$

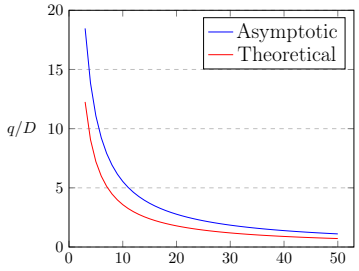


Figure 5.2 – Density of T_{x_6} for $\mathcal{S}_{\ell-3}^{6,d}$ for $d \in \{3, \ldots, 50\}$

Finally, Figure 5.3 reports on the case $r = 1$ in where we fix $d = 4$ and increase the size of the matrix ℓ. Here, the Asymptotic curve comes from Proposition 5.11.

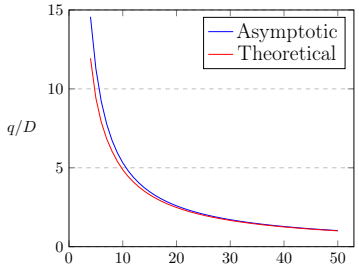


Figure 5.3 – Density of $T_{x_{\binom{\ell}{2}}}$ for $\mathcal{S}_1^{\binom{\ell}{2},4}$ for $\ell \in \{4, \ldots, 50\}$

5.6 Perspectives

Our results describe the fundamental parameter q, the number of dense columns of T_{x_n}. Therefore, while the complexity results in this article focus on the application to the Sparse-FGLM algorithm, we can also apply the propositions of Section 5.4 to the new change-of-ordering algorithm of [12]. There, the authors prove a complexity result, excluding logarithmic factors, of $O^{\sim}(q^{\omega-1}D)$, where ω is the exponent of the complexity of matrix multiplication. Applying our estimates for m leads to even finer complexity results for symmetric determinantal systems. Our bound on q enables more precise comparison of this new algorithm in [12] with the existing algorithms based on fast linear algebra [30, 83] whose complexities lie in $O^{\sim}(D^{\omega})$.

The finer complexity results of Section 5.4 rely primarily on the knowledge of the Hilbert series of the special cases $r = 1, \ell - 2$ and $\ell - 3$. Should further cases be explored, we could expect to obtain stronger results for those cases as well. We would also like to study more types of matrix structure such as moment matrices. For instance, we discuss the case of Hankel variable matrices and derive an alternative derivation of the Hilbert series of $\mathcal{S}_{\ell-2}$.

Let $\ell \in \mathbb{N}$, $c_0, \ldots, c_{2\ell-2}$ be new variables and C be the Hankel matrix

$$C = \begin{bmatrix} c_0 & \cdots & c_{\ell-1} \\ \vdots & \iddots & \vdots \\ c_{\ell-1} & \cdots & c_{2\ell-2} \end{bmatrix}.$$

We denote by $\mathcal{C}_r$ the ideal generated by all the $(r+1)$-minors of C.

Lemma 5.13. *Given $r \in \mathbb{N}$, the Hilbert series of $\mathcal{C}_r$ is equal to*

$$\frac{1}{(1-t)^{2r}} \sum_{i=0}^{r} \binom{2\ell - 2r - 2 + i}{i} t^i.$$

Proof. By [20, Corollary 2.2], $\mathcal{C}_r$ coincides with the ideal generated by $(r+1)$-minors of the $(r+1) \times (2\ell - r - 1)$ Hankel matrix

$$\overline{C} = (c_{i+j})_{0 \leq i \leq r, 0 \leq j \leq 2\ell - r - 2}$$

and the codimension of $\mathcal{C}_r$ is $2\ell - 2r - 1$.

Let $M = (m_{i,j})_{0 \leq i \leq r, 0 \leq j \leq 2\ell - r - 2}$ be a general variable matrix of the same size of $\overline{C}$ and I be the ideal generated by all the $(r+1)$-minors of M. Hence, the ideal $\mathcal{C}_r$ can be identified with

$$I + \langle c_i - m_{j,i-j}, \mid 0 \leq i \leq 2\ell - 2, 0 \leq j \leq i \rangle.$$

Since $\mathbb{K}[m_{0,0}, \ldots, m_{r,2\ell-r-2}]/I$ is a Cohen-Macaulay ring of the same codimension $2\ell - 2r - 1$ as $\mathbb{K}[c_0, \ldots, c_{2\ell-2}]/\mathcal{C}_r$, the unmixedness theorem [27, Cor. 18.14] and [10] give the result. □

The above lemma allows one to study similar problems on Hankel matrices. Furthermore, using the same technique as in Lemma 5.13 and noting that both $\mathcal{C}_{\ell-2}$ and $\mathcal{S}_{\ell-2}$ have codimension three, one can obtain a different derivation of the Hilbert series of $\mathcal{S}_{\ell-2}$.

Additionally, we make the following conjecture for triangular matrices that, as far as we are aware, is new.

Conjecture 5.14. *Let T be an $\ell \times \ell$ triangular variable matrix and $\mathcal{T}_r$ be the ideal generated by its $(r+1)$-minors. Then the Hilbert series associated to $\mathcal{T}_r$ equals the Hilbert series associated to the ideal $\mathcal{S}_r$.*

As the proofs in this chapter rely solely on the Hilbert series of the ideal we consider, if Conjecture 5.14 holds then our results also hold for ideals generated by minors of triangular matrices.

Chapter 6

Computing the set of asymptotic critical values of polynomial mappings from smooth algebraic sets

Abstract. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{Q}[z_1, \ldots, z_n]$ be a polynomial tuple. Define the polynomial mapping $\boldsymbol{f} : X \to \mathbb{C}^p$, where X is a smooth algebraic set defined by the simultaneous vanishing of the reduced regular sequence $g_1, \ldots, g_m$, with $m + p \leq n$. Let $d = \max(\deg f_1, \ldots, \deg f_p, \deg g_1, \ldots, \deg g_m)$, $\mathrm{d}\boldsymbol{f}$ be the differential of $\boldsymbol{f}$ and κ be a continuous function measuring the distance of a linear operator to the set of singular linear operators from $\mathbb{C}^n$ to $\mathbb{C}^p$. We consider the problem of computing the set of asymptotic critical values of $\boldsymbol{f}$. This is the set of values c in the target space of $\boldsymbol{f}$ such that there exists a sequence of points $(\mathbf{x}_i)_{i\in\mathbb{N}}$ tending to ∞ for which $\boldsymbol{f}(\mathbf{x}_i)$ tends to c and $\|\mathbf{x}_i\|\kappa(\mathrm{d}\boldsymbol{f}(\mathbf{x}_i))$ tends to 0 when i tends to infinity.

The union of the classical and asymptotic critical values contains the so-called bifurcation set of a polynomial mapping. Thus, by computing both the critical values and the asymptotic critical values, one can utilise generalisations of Ehresmann's fibration theorem in non-proper settings for applications in polynomial optimisation and computational real algebraic geometry.

We design new efficient algorithms for computing the set of asymptotic critical values of a polynomial mapping restricted to a smooth algebraic set. We give the first bound on the degree of these values, showing that they are contained in a hypersurface of degree at most pD, where $D = d^{n-p-1}\sum_{i=0}^{p+1}\binom{n+p-1}{m+2p+i}d^i$. We also give the first complexity analysis of this problem, showing that it requires at most $O^{\sim}(p^2D^{p+5} + n^2d^{n+2}D^{p+4})$ operations in the base field. Moreover, in the special case $p = 1$, we give a sharper complexity estimate of $O^{\sim}(n^2d^{n+2}D^5)$ arithmetic operations.

Additionally, we show how to apply these algorithms to polynomial optimisation problems and the problem of computing sample points per connected component of a semi-algebraic set defined by a single inequality/inequation.

We provide implementations of our algorithms and use them to test their practical capabilities. We show that our algorithms significantly outperform the current state-of-the-art algorithms by tackling previously out of reach benchmark examples.

This chapter contains joint work with J. Berthomieu and M. Safey El Din and led to the submission of an article.

6.1 Introduction

Definition of asymptotic critical values Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$ and let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[z_1, \ldots, z_n]^p$ be a polynomial mapping. Let $\boldsymbol{g} = (g_1, \ldots, g_m)$ be a reduced regular sequence such that the variety $X = \mathbf{V}(g_1, \ldots, g_m) \subset \mathbb{C}^n$ is smooth. We consider the polynomial mapping

$$\boldsymbol{f} : \boldsymbol{x} = (x_1, \ldots, x_n) \in X \mapsto (f_1(x_1, \ldots, x_n), \ldots, f_p(x_1, \ldots, x_n)) \in \mathbb{K}^p.$$

We assume that $n \geq m+p$ and that this mapping is dominant, so that the image of $\boldsymbol{f}$ is dense in $\mathbb{C}^p$. For ease of notation, we shall denote $z_1, \ldots, z_n$ by $\mathbf{z}$ and $c_1, \ldots, c_p$ by $\mathbf{c}$, for the value of the polynomial mapping $\boldsymbol{f}$. Denote by $\mathrm{d}\boldsymbol{f}$ the differential of the mapping $\boldsymbol{f}$ and, for a given point $\boldsymbol{x} \in X$, $\mathrm{d}\boldsymbol{f}(\boldsymbol{x})$ the differential of $\boldsymbol{f}$ at $\boldsymbol{x}$, a linear map from the tangent space $T_{\boldsymbol{x}}X$ of X at $\boldsymbol{x}$ to the tangent space $T_{\boldsymbol{f}(\boldsymbol{x})}\mathbb{K}^p$ of $\mathbb{K}^p$ at $\boldsymbol{f}(\boldsymbol{x})$.

Then, the set of critical values of $\boldsymbol{f}$ are defined as

$$K_0(\boldsymbol{f}) = \{\mathbf{c} \in \mathbb{C}^p \mid \exists \boldsymbol{x} \in X \text{ s.t. } \boldsymbol{f}(\boldsymbol{x}) = c \text{ and } \operatorname{rank}(\operatorname{jac}(\boldsymbol{f}, \boldsymbol{g})(\boldsymbol{x})) < m+p\}\,.$$

Denote by $L(\mathbb{K}^n, \mathbb{K}^p)$ the space of linear mappings from $\mathbb{K}^n$ to $\mathbb{K}^p$ and by Σ the singular set of $L(\mathbb{K}^n, \mathbb{K}^p)$. First defined in [89], denote by ν the distance of an operator $A \in L(\mathbb{K}^n, \mathbb{K}^p)$ to the set of singular operators: [62, Proposition 2.2]

$$\nu(A) = \operatorname{dist}(A, \Sigma) = \inf_{B \in \Sigma} \|A - B\|.$$

Then, the set of asymptotic critical values of the polynomial mapping $\boldsymbol{f}$ restricted to the algebraic set X is defined as follows:

$$K_\infty(\boldsymbol{f}) = \{\mathbf{c} \in \mathbb{C}^p \mid \exists (\mathbf{z}_t)_{t\in\mathbb{N}} \subset X \text{ s.t.} \|\mathbf{z}_t\| \to \infty, \boldsymbol{f}(\mathbf{z}_t) \to c \text{ and } \|\mathbf{z}_t\| \nu(\mathrm{d}\boldsymbol{f}(\mathbf{z}_t)) \to 0\}\,.$$

Motivation The set of generalised critical values is defined to be the union of the classical critical values and the asymptotic critical values, $K(\boldsymbol{f}) = K_0(\boldsymbol{f}) \cup K_\infty(\boldsymbol{f})$. In [89], the author proved that this set contains the so-called bifurcation set of $\boldsymbol{f}$. Essentially, this provides a generalisation of Ehresmann's fibration theorem to non-proper settings. Thus,

$$\boldsymbol{f} : X \setminus f^{-1}(K(\boldsymbol{f})) \to \mathbb{K}^p \setminus K(\boldsymbol{f})$$

is a locally trivial fibration which by definition, means that for all connected open sets $U \subset \mathbb{K}^p \setminus K(\boldsymbol{f})$, for all $y \in U$ there exists a diffeomorphism φ such that the following diagram commutes:

$$\begin{array}{ccc} \boldsymbol{f}^{-1}(y) \times U & \xrightarrow{\varphi} & \boldsymbol{f}^{-1}(U) \\ & \searrow^{\pi} & \downarrow \boldsymbol{f} \\ & & U \end{array}$$

where π is the projection map onto U [58, Theorem 3.1]. However, for this to be computationally meaningful, we require the set $K(\boldsymbol{f})$ not to be dense in $\mathbb{K}^p$. It is well known that by Bertini's algebraic version of Sard's theorem, the set $K_0(\boldsymbol{f})$ has codimension at least one in $\mathbb{C}^p$. Crucially, it has also been shown that the set of asymptotic critical values satisfies a generalised Sard's theorem [58, Theorem 3.3].

Therefore, the computation of the generalised critical values for effective uses in real algebraic geometry is appealing. Their fibration property has been capitalised upon in [44, 92] to design algorithms for

- exact polynomial optimisation (i.e. computing the minimal polynomial of the infimum of the map $x \to f(x)$ restricted to $X \cap \mathbb{R}^n$ and an isolating interval for this infimum),
- computing sample points for each connected component of a semi-algebraic set defined by a single inequality.

Prior works Computing the set of critical values of a polynomial mapping restricted to an algebraic set is classical. By the Jacobian criterion under the assumption that X is smooth and $\boldsymbol{g}$ is a reduced regular sequence, one may consider the algebraic set defined by the intersection of

X with the variety defined by the maximal minors of $\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})$ to find the critical points of $\boldsymbol{f}$. Then, the set $K_0(\boldsymbol{f})$ is equal to the set of values of $\boldsymbol{f}$ at these points [27, Corollary 16.20].

As far as we are aware, the first work towards the computation of the asymptotic critical values of a polynomial mapping was given in [62]. This is based on a geometric characterisation of $K_\infty(\boldsymbol{f})$ that allows one to construct an algebraic set of codimension at least one in $\mathbb{C}^p$ that contains the asymptotic critical values. Then, one can construct polynomials defining this algebraic set by using algorithms that compute elimination ideals in polynomial rings, such as Gröbner basis based algorithms. Note that the authors of this paper only consider polynomial mappings with an unrestricted domain. Later, the authors of [58] proposed an algorithm for computing the generalised critical values of a polynomial mapping restricted to an algebraic set. This follows a similar schematic of defining algebraic sets, considering their intersections with linear hyperspaces and projecting onto the target space. However, this algorithm constructs $(p(m+p))^{\binom{n}{m+p}}$ locally closed sets in $\mathbb{C}^{(n+1)\binom{n}{m+p}+p+n}$ before projecting onto $\mathbb{C}^p$ making the algorithm impractical. Furthermore, a complexity analysis for this algorithm is lacking.

Several attempts to improve this algorithmic pattern have been made in the global case with $p = 1$. We mention [92] in which the author makes the connection between generalised critical values and properties of polar varieties. This connection is exploited in [59] where the authors build rational arcs that reach all the generalised critical values of a polynomial. Moreover, in [60], the authors make a distinction between asymptotic critical values, detecting those that are found non-trivially, meaning away from the critical locus of the polynomial, something not covered in this paper.

Main results We assume that $\boldsymbol{f}$ satisfies the following regularity assumption (R): "The Zariski closure of $X \setminus \mathrm{crit}(\boldsymbol{f}, X)$ is X", where

$$\mathrm{crit}(\boldsymbol{f}, X) = \{\boldsymbol{x} \in X \mid \mathrm{rank}(\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})(\boldsymbol{x})) < m + p\}$$

is the critical locus of $\boldsymbol{f}$ on X. Hence, Assumption (R) is equivalent to requiring that for a generic point $\boldsymbol{x} \in X$, $\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})(\boldsymbol{x})$ has full rank.

By adapting the results of [62, Section 4], building a geometric characterisation of $K_\infty(\boldsymbol{f})$ using Lagrange multipliers, we develop efficient algorithms for computing asymptotic critical values under the restriction to a smooth algebraic set. We introduce an element of randomisation to avoid some combinatorial steps in the algorithm designed in [58]. Next, with a geometric result, we reduce the computation of $K_\infty(\boldsymbol{f})$ to intersecting the Zariski closure of some locally closed subset of $\mathbb{C}^{n+m+2p}$ with a linear affine subspace of codimension 2 such that the projection onto the target space of $\boldsymbol{f}$ of this intersection contains $K_\infty(\boldsymbol{f})$. Then, by taking advantage of the multi-homogeneous structure of the objects defined in this algorithm, we give a bound on the degree of the asymptotic critical values.

Theorem 2.9. *Let X be a smooth algebraic set defined by a reduced regular sequence $\boldsymbol{g} = (g_1, \ldots, g_m)$. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping from X to $\mathbb{K}^p$ satisfying Assumption (R). Let $d = \max(\deg f_1, \ldots, \deg f_p, \deg g_1, \ldots, \deg g_m)$. Then, the asymptotic critical values of $\boldsymbol{f}$ are contained in a hypersurface of degree at most*

$$p d^{n-p-1} \sum_{i=0}^{p+1} \binom{n+p-1}{m+2p-i} d^i.$$

We note that in many cases, the bound given in Theorem 2.9, combined with the bound on the degree of the critical values in [33, Corollary 2] in the $p = 1$ case, is less than the bound given on the degree of the generalised critical values in [58, Theorem 4.1]. However, for certain values of the parameters m, p and n, the latter bound is actually smaller. This is discussed in Section 6.9.

While in practice, and in our experiments, Gröbner bases are the tool of choice for performing the algebraic elimination routines necessary in our algorithms, we study their complexity by

utilising the geometric resolution algorithm given in [42]. We recall the "soft-Oh" notation: $f(n) \in O^{\sim}(g(n))$ means that $f(n) \in g(n)\log^{O(1)}(3+g(n))$, see also [107, Chapter 25, Section 7].

We now give our first complexity result. The following is for the special case $p = 1$, which is of particular importance for many applications such as polynomial optimisation.

Theorem 2.10. *Let $\boldsymbol{g} = (g_1, \ldots, g_m)$ be a reduced regular sequence defining a smooth algebraic set X. Let $f \in \mathbb{K}[\mathbf{z}]$ be a polynomial mapping from X to $\mathbb{K}$ satisfying Assumption (R). Let $d = \max(\deg f, \deg g_1, \ldots, \deg g_m)$ and $D = d^{n-2}\sum_{i=0}^{2}\binom{n}{m+2-i}d^i$. Then, there exists an algorithm which, on input $f, \boldsymbol{g}$, outputs a non-zero polynomial $H \in \mathbb{K}[c]$ such that $K_\infty(\boldsymbol{f}) \subset \mathbf{V}(H)$ using at most*

$$O^{\sim}(n^2 d^{n+2} D^5)$$

arithmetic operations in $\mathbb{K}$.

In the $p = 1$ case, we can perform some necessary eliminations through the computation of resultants. This leads to a sharper complexity bound. However, in the $p > 1$ case, we must change our methodology for technical reasons. We use the FGLM algorithm [31] which has dominant complexity in our algorithm, to arrive at the following result.

Theorem 2.11. *Let $\boldsymbol{g} = (g_1, \ldots, g_m)$ be a reduced regular sequence defining a smooth algebraic set X. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping from X to $\mathbb{K}^p$ satisfying Assumption (R). Let $d = \max(\deg f_1, \ldots, \deg f_p, \deg g_1, \ldots, \deg g_m)$ and $D = d^{n-p-1}\sum_{i=0}^{p+1}\binom{n+p-1}{m+2p-i}d^i$. Then, there exists an algorithm which, on input $\boldsymbol{f}$ and $\boldsymbol{g}$, outputs p finite lists of non-zero polynomials $G_i \subset \mathbb{K}[\mathbf{c}]$ such that $K_\infty(\boldsymbol{f}) \subset (\mathbf{V}(G_1) \cup \cdots \cup \mathbf{V}(G_p)) \subsetneq \mathbb{C}^p$ using at most*

$$O^{\sim}\left(p^2 D^{p+5} + n^2 d^{n+2} D^{p+4}\right)$$

arithmetic operations in $\mathbb{K}$.

Furthermore, we have implemented all the algorithms given in this paper in the MAPLE [76] computer algebra system. For the Gröbner basis computations, we rely on the Gröbner package in MAPLE. Testing these implementations for a wide range of benchmark examples, we illustrate that our algorithms significantly outperform the state-of-the-art.

Structure of the paper In Section 7.2, we develop the geometric characterisation of the asymptotic critical values given in [62] to the setting of restrictions to smooth algebraic sets. Then, we explore an interpretation of this characterisation in terms of Lagrange multipliers that leads directly to an algorithm for computing the set of asymptotic critical values. In Section 6.3, we prove our main geometric result, upon which the efficiency of our algorithms relies. Then, in Section 6.4, we apply the results of the previous two sections to introduce two elements of randomisation in order to design new algorithms more efficient than the state-of-the-art. In Sections 6.5 and 6.6, we prove our main results by analysing the degree of the objects computed in, and the complexity of, our new algorithms. An additional algorithm, deriving from a different interpretation of the geometric characterisation of the asymptotic critical values is presented in Section 6.7. We illustrate how our algorithms can be applied to solve polynomial optimisation problems and other problems in real algebraic geometry in Section 6.8. Finally, in Section 6.9, we compare all the algorithms given in this paper in terms of time. Furthermore, we compare our degree result to the bound given in [58, Theorem 4.1] and to the true number of asymptotic critical values for a set of benchmark examples.

6.2 Preliminaries

We begin with a lemma in linear algebra.

Lemma 6.1. *With $n \geq m$, consider the linear maps $F : \mathbb{C}^n \to \mathbb{C}^m$ and $P : \mathbb{C}^n \to \mathbb{C}$, defined by $F(x) = (F_1 \cdot x, \ldots, F_m \cdot x)$ and $P(x) = \ell \cdot x$ respectively where $F_1, \ldots, F_m, \ell \in \mathbb{C}^n$. Then,*

$$\ell \in \operatorname{span}(F_1, \ldots, F_m) \iff \ker F \subset \ker P.$$

Proof. We shall prove this by double inclusion. Firstly, assume that $P \in \operatorname{span}(F_1, \ldots, F_m)$ so that $\ell = \sum_{i=1}^m y_i F_i$ for some $y \in \mathbb{C}^m$. Then, for all $x \in \ker F$,

$$P(x) = \left(\sum_{i=1}^m y_i F_i \right) \cdot x = \sum_{i=1}^m y_i (F_i \cdot x) = 0.$$

Hence, $x \in \ker P$ and $\ker F \subset \ker P$.

Now, assume that $\ker F \subset \ker P$. Consider the map $G : \mathbb{C}^n \to \mathbb{C}^{m+1}$ defined by $G(x) = (F_1 \cdot x, \ldots, F_m \cdot x, \ell \cdot x)$. For $x \in \ker F \subset \ker P$, we have $G(x) = 0$, hence $x \in \ker G$. Conversely, if $x \in \ker G$, then $(F_1 \cdot x, \ldots, F_m \cdot x, P \cdot x) = (0, \ldots, 0, 0)$ and $x \in \ker F \subset \ker P$. Hence $\ker G = \ker F$.

By the rank-nullity theorem, we have $\dim \operatorname{im} F = n - \dim \ker F = n - \dim \ker G = \dim \operatorname{im} G$. Hence $\operatorname{im} F$ is isomorphic to $\operatorname{im} G$ and $\operatorname{im} P$ can be identified with a vector subspace of $\operatorname{im} F$. In other words, $\ell \in \operatorname{span}(F_1, \ldots, F_m)$. □

Let $\boldsymbol{g}$ be a reduced regular sequence defining a smooth algebraic set X. Let $\boldsymbol{f} : X \to \mathbb{C}^p$ be a polynomial mapping satisfying Assumption (R). By [58, Theorem 3.3] the set of asymptotic critical values of $\boldsymbol{f}$ has codimension at least one in $\mathbb{C}^p$. The aim of this section is to define an algebraic set containing $K_\infty(\boldsymbol{f})$ that also has codimension at least one in $\mathbb{C}^p$.

Let $\operatorname{jac}(\boldsymbol{f}, \boldsymbol{g})$ be the Jacobian matrix associated to the mapping $(f_1, \ldots, f_p, g_1, \ldots, g_m)$,

$$\operatorname{jac}(\boldsymbol{f}, \boldsymbol{g}) = \begin{bmatrix} \frac{\partial f_1}{\partial z_1} & \cdots & \frac{\partial f_1}{\partial z_n} \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial z_1} & \cdots & \frac{\partial f_p}{\partial z_n} \\ \frac{\partial g_1}{\partial z_1} & \cdots & \frac{\partial g_1}{\partial z_n} \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial z_1} & \cdots & \frac{\partial g_m}{\partial z_n} \end{bmatrix}.$$

For $1 \leq j \leq p$, denote by $\operatorname{jac}(\boldsymbol{f}, \boldsymbol{g})^{[j]}$ the submatrix of $\operatorname{jac}(\boldsymbol{f}, \boldsymbol{g})$ obtained by removing the jth row. Note that we only ever remove one of the first p rows. Denote by N_j the right kernel of the matrix $\operatorname{jac}(\boldsymbol{f}, \boldsymbol{g})^{[j]}$. In the special case $(p, m) = (1, 0)$, $j = 1$ and the resulting matrix has no entries. So, by convention, we say its kernel N_1 is $\mathbb{K}^n$. The differential $\mathrm{d}f_j(z)$ of the map f_j at z induces a linear map from $\mathbb{C}^n$ to $\mathbb{C}$. Since N_j is a vector subspace of $\mathbb{C}^n$, we denote by $w_j(z)$ the restriction of this linear map to N_j

Following [58, Proposition 2.3], for a linear subspace $H \subset \mathbb{K}^n$ defined by vectors $B_1, \ldots, B_m$, let $F \in L(H, \mathbb{K}^p)$ be a linear map represented by a matrix with rows $(A_1, \ldots, A_p) \subset \mathbb{K}^n$. We consider the so-called Kuo distance defined by

$$\kappa(F) = \min_{1 \leq j \leq p} \operatorname{dist}\left(A_i, \operatorname{span}((A_j)_{i \neq j}, (B_k)_{1 \leq k \leq m})\right).$$

In particular, for $z \in X$ we have that

$$\kappa(\mathrm{d}\boldsymbol{f}(z)) = \min_{1 \leq j \leq p} ||w_j(z)||.$$

By [58, Corollary 2.1], the function ν is equivalent to the Kuo distance. Hence, an equivalent definition of the set of asymptotic critical values, the one that we shall primarily use, is the following:

$$K_\infty(\boldsymbol{f}) = \{\mathbf{c} \in \mathbb{C}^p \mid \exists (\mathbf{z}_t)_{t \in \mathbb{N}} \subset X \text{ s.t.} \|\mathbf{z}_t\| \to \infty, \boldsymbol{f}(\mathbf{z}_t) \to c \text{ and } \|\mathbf{z}_t\| \kappa(\mathrm{d}\boldsymbol{f}(\mathbf{z}_t)) \to 0\}.$$

Restriction to a proper algebraic subset of $\mathbb{C}^n$ can affect the asymptotic critical values of a polynomial mapping in subtle ways. For example, a path that leads to an asymptotic critical value in the unrestricted setting may not satisfy the Jacobian condition in the definition of $K_\infty(\boldsymbol{f})$. However, restricting $\boldsymbol{f}$ to an algebraic set that contains this path and thereby adding rows to said Jacobian, can result in this path now satisfying all the above conditions.

Example 6.2. *Let $f = z_1^2 + (z_1z_2 - 1)^2$. First we consider the global case, $f : \mathbb{C}^2 \to \mathbb{C}$. We shall show that $0 \in K_\infty(f)$. The gradient is equal to*

$$\mathrm{d}f = (2z_1 + 2z_2(z_1z_2 - 1), 2z_1(z_1z_2 - 1)).$$

Then, consider the path $z(t) = (t, (1/t) - t)$ as $t \to 0$. We see that $\|z(t)\| \to \infty$ and $f(z(t)) = t^2 + t^4 \to 0$. Furthermore, we have that $\mathrm{d}f(z(t)) = (2t^3, -2t^3)$. Since $p = 1$ and $m = 0$, the Kuo distance κ can be simply replaced by the 2-norm. Hence,

$$\|z(t)\|^2\|\mathrm{d}f(z(t))\|^2 = 8t^6\left(t^2 + \left(\frac{1}{t} - t\right)^2\right) \to 0,$$

and so 0 is an asymptotic critical value of f.

Note that the path $y(t) = (t, 1/t)$ satisfies the first two conditions for a path towards the asymptotic critical value 0, $\|y(t)\| \to \infty$ and $f(y(t)) \to 0$ as $t \to 0$. However, $\mathrm{d}f(y(t)) = (2t, 0)$ and so

$$\|y(t)\|^2\|\mathrm{d}f(y(t))\|^2 = 4t^2\left(t^2 + \frac{1}{t^2}\right) = 4t^4 + 4 \to 4.$$

Now, consider the algebraic set $X = \mathbf{V}(g) = \mathbf{V}(z_1z_2 - 1)$ and the restricted polynomial map $f|_X : X \to \mathbb{C}$ defined by $f|_X(z) = f(z)$. Then, consider the Jacobian

$$\mathrm{jac}(f, g) = \begin{bmatrix} 2z_1 + 2z_2(z_1z_2 - 1) & 2z_1(z_1z_2 - 1) \\ z_2 & z_1 \end{bmatrix}.$$

Let N_1 be the right kernel of $\mathrm{jac}(f, g)^{[1]} = \mathrm{d}g$, then $w_1 = \mathrm{d}f|_{N_j}$ and $\kappa(\mathrm{d}f) = \|w_1\|$. Choose a basis for N_1, say $(-z_1, z_2)$. Then,

$$w_1 : (z_1, z_2) \mapsto -2z_1(2z_1 + 2z_2(z_1z_2 - 1)) + z_2(2z_1(z_1z_2 - 1)) = 2z_1z_2 - 2z_1^2z_2^2 - 4z_1^2.$$

Clearly, the path $y(t)$ is in the set $\mathbf{V}(z_1z_2 - 1)$ for all $t > 0$, so we have $\|y(t)\| \to \infty$ and $f(y(t)) \to 0$ as $t \to 0$ but now we also have

$$\|y(t)\|^2\kappa(\mathrm{d}f(y(t)))^2 = \|y(t)\|^2\|w_1(y(t))\|^2 = 16t^4\left(t^2 + \frac{1}{t^2}\right) \to 0.$$

Hence, the path $y(t) = (t, 1/t)$ does not allow us to conclude that 0 is an asymptotic critical value of $f : \mathbb{C}^2 \to \mathbb{C}$ but it does for its restricted polynomial mapping $f|_X; X \to \mathbb{C}$.

To access the asymptotic behaviour algebraically, we utilise the following transformation that sends $z_s = 0$ to ∞:

$$\tau_s(z) = \left(\frac{z_1}{z_s}, \ldots, \frac{z_{s-1}}{z_s}, \frac{1}{z_s}, \frac{z_{s+1}}{z_s}, \ldots, \frac{z_n}{z_s}\right).$$

For each choice of $s = 1, \ldots, n$, $j = 1, \ldots, p$ and point $\boldsymbol{x} \in X$, let $W_s^j(\boldsymbol{x})$ be the graph of $x_s w_j(\boldsymbol{x})$, a point in the Grassmannian of linear subspaces of $\mathbb{C}^n \times \mathbb{C}$ that are of dimension $n - p - m + 1$, denoted by $\mathbb{G}_{n-p-m+1}(\mathbb{C}^n \times \mathbb{C})$. Recall that this Grassmannian is a compact smooth manifold that parameterises all $(n - p - m + 1)$-dimensional linear subspaces of $\mathbb{C}^n \times \mathbb{C}$. Since X is a smooth affine variety and the mapping $\boldsymbol{f}$ satisfies Assumption (R), there exists a non-empty Zariski-open subset $\mathcal{O}_X \subset X$ such that for all $\boldsymbol{x} \in \mathcal{O}_X$, $W_s^j(\boldsymbol{x})$ is well-defined, that is when the right kernel of $\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})^{[j]}$ has dimension $n - p - m + 1$.

Then, define the rational mapping

$$M_s^j(\boldsymbol{f}) : X \setminus \{z_s = 0\} \to \mathbb{C}^p \times \mathbb{G}_{n-p-m+1}(\mathbb{C}^n \times \mathbb{C}),$$
$$z \mapsto (\boldsymbol{f}(\tau_s(z)), W_s^j(\tau_s(z))).$$

Let $\Lambda = \mathbb{G}_{n-p-m+1}(\mathbb{C}^n \times 0)$. This is the set of $(n-p-m+1)$-dimensional graphs of linear maps from $\mathbb{C}^n$ to $\mathbb{C}$ that are identically the zero map.

$$L_s^j(\boldsymbol{f}) = \overline{\operatorname{graph} M_s^j(\boldsymbol{f})} \cap (\{z \in X | z_s = 0\} \times \mathbb{C}^p \times \Lambda). \tag{6.1}$$

Define $\pi : X \times \mathbb{C}^p \times \mathbb{G}_{n-k+1}(\mathbb{C}^n \times \mathbb{C}) \to \mathbb{C}^p$ to be the projection map and take $K_s^j(\boldsymbol{f}) = \pi(L_s^j(\boldsymbol{f}))$. We shall prove that $L_s^j(\boldsymbol{f})$ is an algebraic set.

To this end, for a reduced rational function, φ/θ, we define the function numer by $\operatorname{numer}(\varphi/\theta) = \varphi$. Likewise, for a vector of reduced rational functions, $(\varphi_1/\theta_1, \ldots, \varphi_m/\theta_m)$, we extend the function numer so that $\operatorname{numer}(\varphi_1/\theta_1, \ldots, \varphi_m/\theta_m) = (\varphi_1, \ldots, \varphi_m)$. The advantage of the transformation τ_s is that it allows us to give an algebraic description of the sets $L_s^j(\boldsymbol{f})$. By elimination of variables, we are then able to compute the Zariski-closure of the projection of $L_s^j(\boldsymbol{f})$ on the $\mathbf{c}$-space. This gives, in general, a superset of the asymptotic critical values of codimension at least 1 in $\mathbb{C}^p$. Moreover, in the special case $p = 1$, of particular interest for many applications such as polynomial optimisation, this inclusion becomes equality.

First, we give a lemma that will allow a Lagrange multiplier interpretation of the Kuo distance.

In the algorithms presented in this paper, we shall derive polynomials whose simultaneous vanishing set is the Zariski-closure of the graph of the map $M_s^j(\boldsymbol{f}^A)$. For this purpose, we introduce $m + p - 1$ new variables $(\lambda_1, \ldots, \lambda_{m+p-1}) = \boldsymbol{\lambda}$, that will be Lagrange multipliers. Additionally, we recall that $(c_1, \ldots, c_p) = \mathbf{c}$ are indeterminates representing the values of $\boldsymbol{f}^A$ and thus the p first coordinates of the values of $M_s^j(\boldsymbol{f}^A)$. We also introduce n new variables $(u_1, \ldots, u_n) = \mathbf{u}$ for the last n coordinates of the values of $M_s^j(\boldsymbol{f}^A)$.

Lemma 6.3. *Let X be a smooth algebraic set defined by a reduced regular sequence $\boldsymbol{g} = (g_1, \ldots, g_m)$. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping from X to $\mathbb{K}^p$ satisfying Assumption (R). Then, there exist indeterminates $\mathbf{c} = (c_1, \ldots, c_p), \mathbf{u} = (u_1, \ldots, u_n)$, Lagrange multipliers $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_{m+p+1})$ and polynomials $h_1, \ldots, h_{n+m+p}$ in $\mathbb{K}[\mathbf{z}, \mathbf{c}, \mathbf{u}, \boldsymbol{\lambda}]$ such that*

$$h_i = \operatorname{numer}(f_i(\tau_s(z)) - c_1),\ 1 \le i \le p,$$
$$h_{p+i} = \operatorname{numer}(g_i(\tau_s(z))),\ 1 \le i \le m,$$
$$h_{p+m+i} = \operatorname{numer}\left(z_s \operatorname{jac}(f_j)_i - \sum_{k=1}^{m+p-1} \lambda_k \operatorname{jac}(\boldsymbol{f}, \boldsymbol{g})_{k,i}^{[j]} - u_i\right),\ 1 \le i \le n,$$
$$\overline{\operatorname{graph} M_s^j(\boldsymbol{f})} = \overline{\mathbf{V}(h_1, \ldots, h_{n+m+p}) \setminus \mathbf{V}(z_s)},$$
$$L_s^j(\boldsymbol{f}) = \overline{\operatorname{graph} M_s^j(\boldsymbol{f})} \cap \mathbf{V}(z_s, u_1, \ldots, u_n),$$

where $\operatorname{jac}(f_j)_i$ is the ith coefficient of $\operatorname{jac}(f_j)$ and $\operatorname{jac}(\boldsymbol{f}, \boldsymbol{g})_{k,i}^{[j]}$ is the coefficient on the kth row and ith column of $\operatorname{jac}(\boldsymbol{f}, \boldsymbol{g})^{[j]}$.

Proof. From the first p components of the map $M_s^j(\boldsymbol{f})$, we take $h_1, \ldots, h_p$ to be $\operatorname{numer}(\boldsymbol{f}(\tau_s(z)) - \mathbf{c})$, where $\mathbf{c}$ are new indeterminates for the value of $\boldsymbol{f}$ at $\tau_s(z)$ and we take the numerators of these rational functions to get polynomials. We shall handle the denominators of these polynomials later by removing the algebraic set they define, thus ensuring these rational functions are always well-defined.

Then, restricting to the algebraic set $X = \mathbf{V}(g_1, \ldots, g_m)$, we set $h_{p+1}, \ldots, h_{p+m}$ to be $\operatorname{numer}(\boldsymbol{g}(\tau_s(z)))$. Now, we need an algebraic interpretation of $W_s^j(\tau_s(z))$ and $\Lambda = \mathbb{G}_{n-p-m+1}(\mathbb{C}^n \times 0)$.

Let us recall that $W_s^j(\tau_s(z))$ is an element of the Grassmannian $\mathbb{G}_{n-p+1}(\mathbb{C}^n \times \mathbb{C})$, since the map $M_s^j(\boldsymbol{f})$ is well-defined outside of a nowhere dense algebraic set, and that $w_j(z)$ is the restriction of $z_s \,\mathrm{d}f_j$ to the right kernel of the Jacobian matrix of $\boldsymbol{f}$ with the jth row removed. Recall that the construction of the set $L_s^j(\boldsymbol{f})$, as in equation (6.1), involves the intersection with $\Lambda = \mathbb{G}_{n-p-m+1}(\mathbb{C}^n \times 0)$. This means that we must find some path towards an asymptotic critical value such that $W_s^j(\tau_s(z)) \to W$, for some $W \in \Lambda$. This implies that the right kernel of $\operatorname{jac}(\boldsymbol{f}, \boldsymbol{g})^{[j]}$ tends to a subset of the right kernel of $z_s \,\mathrm{d}f_j$. By Lemma 6.1, this is equivalent to the evaluation of $z_s \,\mathrm{d}f_j$ tending to a vector in the span of the evaluation of $\operatorname{jac}(\boldsymbol{f}, \boldsymbol{g})^{[j]}$ at $\tau_s(z)$. Thus, we may use Lagrange multipliers to represent $W_s^j(\tau_s(z))$ and its limit in Λ. Hence, we set $h_{p+m+1}, \ldots, h_{p+m+n}$ to be the numerators of the following polynomials at $\tau_s(z)$,

$$z_s \operatorname{jac}(f_j) - \sum_{i=1}^{m+p-1} \lambda_i \operatorname{jac}(\boldsymbol{f}, \boldsymbol{g})_i^{[j]} - \mathbf{u}.$$

Now, note that all the denominators of the rational functions we have defined are all powers of z_s. Thus, according to the definition of the map $M_s^j(\boldsymbol{f})$, by removing the algebraic set $\mathbf{V}(z_s)$ from $\mathbf{V}(h_1, \ldots, h_{n+m+p})$, we get exactly the graph of $M_s^j(\boldsymbol{f})$. Therefore, the algebraic closures give us the first equality

$$\overline{\operatorname{graph} M_s^j(\boldsymbol{f})} = \overline{\mathbf{V}(h_1, \ldots, h_{n+m+p}) \setminus \mathbf{V}(z_s)}.$$

Secondly, to compute $L_s^j(\boldsymbol{f})$ we intersect with the space $(\{z \in X | z_s = 0\} \times \mathbb{C}^p \times \Lambda)$. As discussed above, by Lemma 6.1, the intersection with Λ is achieved by setting the introduced $\mathbf{u}$ variables to 0. Then, the second equality is clear

$$L_s^j(\boldsymbol{f}) = \overline{\operatorname{graph} M_s^j(\boldsymbol{f})} \cap \mathbf{V}(z_s, u_1, \ldots, u_n). \qquad \square$$

This framework suggests looking at each coordinate tending to infinity separately. Instead, we shall introduce a probabilistic element that allows one to investigate every coordinate tending to infinity at once.

Definition 6.4. *Let $A \in \mathrm{GL}_n(\mathbb{K})$ be an invertible matrix and $\boldsymbol{g} = (g_1, \ldots, g_m) : \mathbb{C}^n \to \mathbb{C}^m$ be a polynomial mapping. We let $\boldsymbol{g}^A : z \in \mathbb{C}^n \to \boldsymbol{g}(Az) = (g_1(Az), \ldots, g_m(Az)) \in \mathbb{C}$.*

For an algebraic set $X = \mathbf{V}(\boldsymbol{g}) \subset \mathbb{C}^n$, we let $X^A = \mathbf{V}(\boldsymbol{g}^A)$.

For a polynomial mapping $\boldsymbol{f} = (f_1, \ldots, f_p) : X \to \mathbb{C}^p$, we let

$$\boldsymbol{f}^A : z \in X^A \to \boldsymbol{f}(Az) = (f_1(Az), \ldots, f_p(Az)) = (f_1^A(z), \ldots, f_p^A(z)) \in \mathbb{C}^p.$$

Lemma 6.5. *Let $\boldsymbol{f} : X \to \mathbb{C}^p$ be a polynomial mapping from an algebraic set X. Let $A \in \mathrm{GL}_n(\mathbb{K})$ be an invertible matrix and $\boldsymbol{f}^A : X^A \to \mathbb{C}^p$ be defined as in Definition 6.4. Then, $K_\infty(\boldsymbol{f}) = K_\infty(\boldsymbol{f}^A)$.*

Proof. Let $\mathbf{c} \in K_\infty(\boldsymbol{f})$ be an asymptotic critical value with a path $z(t) \subset X$ such that $\|z(t)\| \to \infty, \boldsymbol{f}(z(t)) \to \mathbf{c}$ and $\|z(t)\| \nu(\mathrm{d}\boldsymbol{f}(z(t))) \to 0$ as $t \to \infty$. Then, for a given invertible matrix $A \in \mathrm{GL}_n(\mathbb{K})$, define the path $y(t) = A^{-1} z(t) \subset X^A$. Clearly, as $t \to \infty$, $\|y(t)\| \to \infty$ and $f^A(y(t)) \to \mathbf{c}$. Then, to prove that $\mathbf{c} \in K_\infty(\boldsymbol{f}^A)$, it remains to show that $\|y(t)\| \nu(\mathrm{d}\boldsymbol{f}^A(y(t))) \to 0$. Firstly, by [62, Proposition 2.1], $\|y(t)\| \leq \nu(A^{-1}) \|z(t)\|$. Moreover, by the chain rule we have

$$\mathrm{d}\boldsymbol{f}^A(y(t)) = \mathrm{d}\boldsymbol{f}^A(A^{-1} z(t)) = \mathrm{d}\boldsymbol{f}(z(t)) A.$$

Then, since A is an invertible matrix and since the Rabier distance is the distance to the set of singular operators [62, Proposition 2.2], we have that $\|z(t)\| \nu(\mathrm{d}\boldsymbol{f}(z(t))A) \to 0$ and hence $\mathbf{c} \in K_\infty(\boldsymbol{f}^A)$. The reverse direction holds with the same argument. $\square$

Lemma 6.6. *Let X be a smooth algebraic set defined by a reduced regular sequence $\boldsymbol{g} = (g_1, \ldots, g_m)$. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping from X to $\mathbb{K}^p$ satisfying Assumption (R) and let $A \in \mathrm{GL}_n(\mathbb{K})$. Then, there exist indeterminates $\mathbf{c} = (c_1, \ldots, c_p), \mathbf{u} = (u_1, \ldots, u_n)$, Lagrange multipliers $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_{m+p+1})$ and polynomials $h_1, \ldots, h_{n+m+p} \in \mathbb{K}[\mathbf{z}, \mathbf{c}, \mathbf{u}, \boldsymbol{\lambda}]$ such that*

$$\begin{aligned}
h_i &= \mathrm{numer}(f_i^A(\tau_s(z)) - c_1), \ 1 \le i \le p,\\
h_{p+i} &= \mathrm{numer}(g_i^A(\tau_s(z))), \ 1 \le i \le m,\\
h_{p+m+i} &= \mathrm{numer}\left(z_s \,\mathrm{jac}(f_j^A)_i - \sum_{k=1}^{m+p-1} \lambda_k \,\mathrm{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)_{k,i}^{[j]} - u_i\right), \ 1 \le i \le n,\\
\overline{\mathrm{graph}\, M_s^j(\boldsymbol{f}^A)} &= \overline{\mathbf{V}(h_1, \ldots, h_{n+m+p}) \setminus \mathbf{V}(z_s)},\\
L_s^j(\boldsymbol{f}^A) &= \overline{\mathrm{graph}\, M_s^j(\boldsymbol{f}^A)} \cap \mathbf{V}(z_s, u_1, \ldots, u_n),
\end{aligned}$$

where $\mathrm{jac}(f_j^A)_i$ is the ith coefficient of $\mathrm{jac}(f_j^A)$ and $\mathrm{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)_{k,i}^{[j]}$ is the coefficient on the kth row and ith column of $\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})^{[j]}$.

Proof. Firstly, since $A \in \mathrm{GL}_n(\mathbb{K})$, if $\boldsymbol{f}$ satisfies Assumption (R), then so does $\boldsymbol{f}^A$. Thus, $M_s^j(\boldsymbol{f}^A)$ is well-defined outside of a nowhere dense algebraic set. Then, it suffices to apply Lemma 6.3 on $\boldsymbol{f}^A$ and X^A defined by $\boldsymbol{g}^A$ to prove the existence of polynomials $h_1, \ldots, h_{n+m+p}$. □

Lemma 6.7. *Let $\boldsymbol{f} \in \mathbb{K}[\mathbf{z}]^p$ be a dominant polynomial mapping with domain a smooth algebraic set X defined by a reduced, regular sequence $(g_1, \ldots, g_m)$ and let $A \in \mathrm{GL}_n(\mathbb{K})$. Then, there exist polynomials $h_1, \ldots, h_{n+m+p} \in \mathbb{K}[\mathbf{z}, \mathbf{c}, \mathbf{u}, \boldsymbol{\lambda}]$ such that*

$$\begin{aligned}
\overline{\mathrm{graph}\, M_s^j(\boldsymbol{f}^A)} &= \overline{\mathbf{V}(h_1, \ldots, h_{n+m+p}) \setminus \mathbf{V}(z_s)},\\
L_s^j(\boldsymbol{f}^A) &= \overline{\mathrm{graph}\, M_s^j(\boldsymbol{f}^A)} \cap \mathbf{V}(z_s, u_1, \ldots, u_n).
\end{aligned}$$

Proof. Firstly, since $A \in \mathrm{GL}_n(\mathbb{K})$, $\boldsymbol{f}$ being dominant implies that $\boldsymbol{f}^A$ is also dominant. Thus, $M_s^j(\boldsymbol{f}^A)$ is well-defined outside of a nowhere dense algebraic set.

By the first p components of the map $M_s^j(\boldsymbol{f}^A)$, we take $h_1, \ldots, h_p$ to be $\mathrm{numer}(\boldsymbol{f}^A(\tau_s(z)) - \mathbf{c})$, where $\mathbf{c}$ are new indeterminates for the value of $\boldsymbol{f}^A$ at $\tau_s(z)$ and we take the numerators of these rational functions to get polynomials. We shall handle the denominators of these polynomials later by removing the algebraic set they define, thus ensuring these rational functions are always well-defined.

Then, restricting to the algebraic set $X^A = \mathbf{V}(g_1^A, \ldots, g_m^A)$, we set $h_{p+1}, \ldots, h_{p+m}$ to be $\mathrm{numer}(\boldsymbol{g}^A(\tau_s(z)))$. Now, we need an algebraic interpretation of $W_s^j(\tau_s(z))$ and $\Lambda = \mathbb{G}_{n-p-m+1}(\mathbb{C}^n \times 0)$.

Recall that $W_s^j(\tau_s(z))$ is an element of the Grassmannian $\mathbb{G}_{n-p+1}(\mathbb{C}^n \times \mathbb{C})$, since the map $M_s^j(\boldsymbol{f}^A)$ is well-defined outside of a nowhere dense algebraic set, and that $w_j(z)$ is the restriction of $z_s \,\mathrm{d}f_j^A$ to the right kernel of the Jacobian matrix of $\boldsymbol{f}$ with the jth row removed. Recall that the construction of the set $L_s^j(\boldsymbol{f}^A)$, as in equation (6.1), involves the intersection with $\Lambda = \mathbb{G}_{n-p-m+1}(\mathbb{C}^n \times 0)$. This means that we must find some path towards an asymptotic critical value such that $W_s^j(\tau_s(z)) \to W$, for some $W \in \Lambda$. This implies that the right kernel of $\mathrm{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)^{[j]}$ tends to a subset of the right kernel of $z_s \,\mathrm{d}f_j^A$. By Lemma 6.1, this is equivalent to the evaluation of $z_s \,\mathrm{d}f_j^A$ tending to a vector in the span of the evaluation of $\mathrm{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)^{[j]}$ at $\tau_s(z)$. Thus, we may use Lagrange multipliers to represent $W_s^j(\tau_s(z))$ and its limit in Λ. Hence, we set $h_{p+m+1}, \ldots, h_{p+m+n}$ to be the numerators of the following polynomials at $\tau_s(z)$,

$$z_s \,\mathrm{d}f_j^A - \sum_{i=1}^{m+p-1} \lambda_i \,\mathrm{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)_i^{[j]} - \mathbf{u},$$

where $\mathbf{u}$ are n new indeterminates to represent the value of this Lagrangian function, $\boldsymbol{\lambda}$ are Lagrange multipliers and $\operatorname{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)_i^{[j]}$ is the ith row-vector of $\operatorname{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)^{[j]}$.

Now, note that all the denominators of the rational functions we have defined are all powers of z_s. Thus, according to the definition of the map $M_s^j(\boldsymbol{f}^A)$, by removing the algebraic set $\mathbf{V}(z_s)$ from $\mathbf{V}(h_1, \ldots, h_{n+m+p})$, we get exactly the graph of $M_s^j(\boldsymbol{f}^A)$. Therefore, the algebraic closures give us the first equality

$$\overline{\operatorname{graph} M_s^j(\boldsymbol{f}^A)} = \overline{\mathbf{V}(h_1, \ldots, h_{n+m+p}) \setminus \mathbf{V}(z_s)}.$$

Secondly, to compute $L_s^j(\boldsymbol{f}^A)$ we intersect with the space $(\{z \in X | z_s = 0\} \times \mathbb{C}^p \times \Lambda)$. As discussed above, by Lemma 6.1, the intersection with Λ is achieved by setting the introduced $\mathbf{u}$ variables to 0. Then, the second equality is clear

$$L_s^j(\boldsymbol{f}^A) = \overline{\operatorname{graph} M_s^j(\boldsymbol{f}^A)} \cap \mathbf{V}(z_s, u_1, \ldots, u_n). \qquad \square$$

We now have an algebraic description of the np sets $L_s^j(\boldsymbol{f}^A)$ and hence of their projections $K_s^j(\boldsymbol{f}^A)$. However, we shall see that by choosing a sufficiently generic A, it suffices to consider only p of these sets, for instance the sets $K_1^1(\boldsymbol{f}^A), \ldots, K_1^p(\boldsymbol{f}^A)$.

Proposition 6.8. *Let X be a smooth algebraic set. Let $\boldsymbol{f} \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping with domain X satisfying Assumption (R). Then, there exists a non-empty Zariski-open subset $\mathcal{O}_{\mathrm{GL}}$ of $\mathrm{GL}_n(\mathbb{K})$ such that for $A \in \mathcal{O}_{\mathrm{GL}}$ the following equality holds:*

$$K_\infty(\boldsymbol{f}) \subseteq \overline{\bigcup_{j=1}^{p} K_1^j(\boldsymbol{f}^A)}.$$

Let $\boldsymbol{f} \in \mathbb{K}[\mathbf{z}]^p$ be a dominant polynomial mapping with domain a smooth algebraic set X. There exists a non-empty Zariski-open subset $\mathcal{O}_{\mathrm{GL}}$ of $\mathrm{GL}_n(\mathbb{K})$ such that for $A \in \mathcal{O}_{\mathrm{GL}}$ the following equality holds:

$$K_\infty(\boldsymbol{f}) \subseteq \overline{\bigcup_{j=1}^{p} K_1^j(\boldsymbol{f}^A)}.$$

Proof. By assumption on X and $\boldsymbol{f}$ and since the matrix $A \in \mathrm{GL}_n(\mathbb{K})$, we may apply Lemma 6.5 so that

$$K_\infty(\boldsymbol{f}) = K_\infty(\boldsymbol{f}^A) \subseteq \overline{\bigcup_{s=1}^{n} \bigcup_{j=1}^{p} K_s^j(\boldsymbol{f}^A)}.$$

It remains to show, that we can restrict this union to sets K_1^j for $1 \leq j \leq p$.

Consider an irreducible component $C \subset \overline{\mathbb{K}_\infty(\boldsymbol{f})}$ of dimension β and degree δ. Consider β generic hyperplanes,

$$\begin{cases} H_1 & = \ell_{1,0} + \ell_{1,1}x_1 + \cdots + \ell_{1,p}x_p \\ & \vdots \\ H_\beta & = \ell_{\beta,0} + \ell_{\beta,1}x_1 + \cdots + \ell_{\beta,p}x_p \end{cases}$$

and their intersection with C,

$$C \cap H_1 \cap \cdots \cap H_\beta = \{c_1, \ldots, c_\delta\}.$$

Then, $c_1, \ldots, c_\delta$ lie in the algebraic closure $\mathbb{L}$ of the field $\mathbb{K}(\ell_{1,0}, \ldots, \ell_{\beta,p})$. Consider, without loss of generality, the asymptotic critical value $c = c_1$. Then, there exists some sequence $(\boldsymbol{x}_i)_{i \in \mathbb{N}} \subset X$ such that as $i \to \infty$,

$$\|\boldsymbol{x}_i\| \to \infty, \boldsymbol{f}(\boldsymbol{x}_i) \to c \text{ and } \|\boldsymbol{x}_i\| \kappa(\mathrm{d}\boldsymbol{f}(\boldsymbol{x}_i)) \to 0.$$

By the isomorphism between $\mathbb{C}^n$ and $\mathbb{R}^{2n}$, we consider a hyperball $\mathcal{B}$ in $\mathbb{R}^{2n}$ such that $\boldsymbol{f}(\mathcal{B})$ contains an open set around c and $0 \in \{\|\boldsymbol{x}\|\kappa(\mathrm{d}\boldsymbol{f}(\boldsymbol{x})) \,|\, \boldsymbol{x} \in \mathcal{B}\}$. Then, one can apply the curve selection lemma at infinity [62, Lemma 3.3], an extension of the classical curve selection lemma [13, Theorem 2.5.5] obtained by considering a semi-algebraic compactification of $\mathbb{R}^{2n}$. Recall that such semi-algebraic curves may be chosen to be Nash curves [13, Proposition 8.1.12].

Therefore, there exists a path $\gamma : (0,1) \to X$ such that

$$\boldsymbol{f}(\gamma(t)) \to c, \|\gamma(t)\| \to \infty \text{ and } \|\gamma(t)\|\kappa(\mathrm{d}\boldsymbol{f}(\gamma(t))) \to 0 \text{ as } t \to 0 \tag{6.2}$$

where each component of γ is a Puiseux series in t with coefficients in $\mathbb{L}$, that is $\gamma \in \overline{\mathbb{L}(t)}^n$ and depends on $\ell_{1,0}, \ldots, \ell_{\beta,p}$. Then, by the definition of Puiseux series, each component of $\gamma(t)$ has finitely many terms with negative exponents. Let r be the least rational number such that t^r has a non-zero coefficient for some component of $\gamma(t)$, or in other words, the exponent of the term that tends to infinity fastest as $t \to 0$. Thus, we can write

$$\gamma(t) = \left(\sum_{k \geq r} \gamma_{1,k} t^k, \ldots, \sum_{k \geq r} \gamma_{n,k} t^k \right) \in \overline{\mathbb{L}(t)}^n.$$

Consider the group of $n \times n$ invertible matrices $\mathrm{GL}_n(\mathbb{L})$ with entries in $\mathbb{L}$. For $B = (b_{i,k})_{1 \leq i,k \leq n} \in \mathrm{GL}_n(\mathbb{L})$, let $y(t) = B\gamma(t)$ and set

$$y_1 = \sum_{k \geq r} y_{1,k} t^k.$$

Consider the coefficient

$$y_{1,r} = \sum_{k=1}^{n} b_{1,k} \gamma_{k,r}.$$

Then, $y_{1,r} = 0$ defines the Zariski-closed subset $\mathscr{C}$ of $\mathrm{GL}_n(\mathbb{L})$ such that $B \in \mathscr{C}$ implies that the first component of $B\gamma(t)$ is such that r is not the least exponent. By definition, some $y_{i,r}$ is non-zero and so $\mathscr{C}$ is a proper subset. Therefore, there exists a non-empty Zariski-open subset $\mathcal{O}_{\mathbb{L}}^{-1}$ of $\mathrm{GL}_n(\mathbb{L})$ such that for $B \in \mathcal{O}_{\mathbb{L}}^{-1}$, $\|(B\gamma)_1(t)\|$ tends to infinity at the same speed as $\|\gamma(t)\|$ as $t \to 0$. Let $\mathcal{O}_{\mathbb{L}}$ be the non-empty Zariski closed subset of $\mathrm{GL}_n(\mathbb{L})$ defined by $A \in \mathcal{O}_{\mathbb{L}} \iff A^{-1} \in \mathcal{O}_{\mathbb{L}}^{-1}$.

Choose some $A \in \mathcal{O}_{\mathbb{L}}$ and consider the polynomial mapping $\boldsymbol{f}^A = \boldsymbol{f}(Az)$ restricted to the algebraic set defined by $X^A = \mathbf{V}(\boldsymbol{g}^A) = \mathbf{V}(\boldsymbol{g}(Az))$ and the path $\Gamma(t) = A^{-1}\gamma(t)$. As $t \to 0$, $\|\Gamma(t)\| \to \infty$ and $\boldsymbol{f}^A(\Gamma(t)) \to c$. Furthermore, by the construction of $\mathcal{O}_{\mathbb{L}}$, the first coordinate Γ_1 of the path Γ is such that $\|\Gamma_1(t)\| \to \infty$ as $t \to 0$. Recall that κ is equivalent to ν. Thus, since $A \in \mathrm{GL}_n(\mathbb{L})$, by [62, Corollary 2.1], we have $\|y(t)\|\nu(\mathrm{d}\boldsymbol{f}^A(y(t))) \to 0$ which implies that $\|\Gamma(t)\|\kappa(\mathrm{d}\boldsymbol{f}^A(\Gamma(t))) \to 0$. Choose j such that $\kappa(\mathrm{d}\boldsymbol{f}^A(\Gamma(t))) = \|w_j(\Gamma(t))\|$. Then, since the Grassmannian $\mathbb{G}_{n-k+1}(\mathbb{L}^n \times \mathbb{L})$ is compact, there is a limit W_1^j of graphs $\Gamma_1(t)w_j(\Gamma(t))$ where $W_1^j \in \Lambda$ by [78, Lemma 5.1]. Therefore, we have in the limit $(0, c, W_1^j) \in L_1^j(\boldsymbol{f})$ and so $c \in K_1^j(\boldsymbol{f}^A)$.

We now demonstrate that there is a Zariski-open subset of specialisations ℓ, i.e. specialisations of $\ell_{1,0}, \ldots, \ell_{\beta,p}$ in elements of $\mathbb{K}$, such that the specialised path satisfies the conditions of the definition of the asymptotic critical values in equation (6.2). Firstly, the denominator of the coefficient corresponding to the rth exponent of γ_1 is a polynomial in $\mathbb{K}[\ell_{1,0}, \ldots, \ell_{\beta,p}]$. Additionally, the coefficients of $f(\gamma)$ have finitely many algebraically independent denominators. Similarly, the differential $\mathrm{d}\boldsymbol{f}$ can only introduce finitely many algebraically independent denominators in the coefficients of $\mathrm{d}\boldsymbol{f}(\gamma)$. Hence, there exists a Zariski-open subset $\mathfrak{L}$ of $\mathbb{K}^{\beta(p+1)}$ such that for all $\ell \in \mathfrak{L}$, the specialisation γ_ℓ of the path γ behaves well, meaning that the conditions in equation (6.2) are satisfied for some asymptotic critical value $c_\ell \in C$.

Consider the variable matrix $A = (a_{i,k})_{1 \leq i,k \leq n}$. Since $\mathcal{O}_{\mathbb{L}}$ is non-empty, there exists some non-zero $\Delta \in \mathbb{L}[a_{1,1}, \ldots, a_{n,n}]$ that defines the Zariski-closed complement of $\mathcal{O}_{\mathbb{L}}$. Choose some a such that $\Delta(a) \neq 0$. Then, $\Delta(a)$ is a rational fraction of the parameters $\ell_{1,0}, \ldots, \ell_{\beta,p}$. Hence, there exists a Zariski-open subset $\mathcal{L} \subset \mathfrak{L} \subset \mathbb{K}^{\beta(p+1)}$ such that any specialisation $\ell \in \mathcal{L}$ is such

that $\Delta(a) \neq 0$. Then, for such a specialisation ℓ we obtain a path γ_ℓ such that $f(\gamma_\ell(t)) \to c_\ell$ as $t \to 0$ for some $c_\ell \in C$. Therefore, we can define a non-empty Zariski-open subset $\mathcal{O}_{\mathbb{K}}$ of $\mathrm{GL}_n(\mathbb{K})$ by evaluating all matrices in $\mathcal{O}_{\mathbb{L}}$ at a given $\ell \in \mathcal{L}$.

Let $C_1, \ldots, C_k$ be the irreducible components of $\overline{K_\infty(\boldsymbol{f})}$. For each C_i, we define a non-empty Zariski-open subset $\mathcal{O}_{\mathbb{K},i}$ of $\mathrm{GL}_n(\mathbb{K})$ as above. Thus, $\mathcal{O}_{\mathrm{GL}} = \bigcap_{i=1}^{k} \mathcal{O}_{\mathbb{K},i}$ is a non-empty Zariski-open subset of $\mathrm{GL}_n(\mathbb{K})$ such that for all $A \in \mathcal{O}_{\mathrm{GL}}$, $\bigcup_{j=1}^{p} K_1^j(\boldsymbol{f}^A)$ contains a Zariski dense subset of $\overline{K_\infty(\boldsymbol{f})}$. Thus, the following equality holds

$$K_\infty(\boldsymbol{f}) \subseteq \overline{\bigcup_{j=1}^{p} K_1^j(\boldsymbol{f}^A)}. \qquad \square$$

6.3 Geometric result

In this section, we state our main geometric result that will form the basis of the proof of correctness of the probabilistic algorithms we give in Sections 6.4 and 6.7.

Proposition 6.9. *Let $W \subset \mathbb{C}^N$ be an algebraic set and let $n \leq s < N$. Let Z be a hyperplane of $\mathbb{C}^N$ and let $\overline{W \setminus Z} = V_1 \cup \cdots \cup V_k$ be an irreducible decomposition. Let π be the canonical projection map from W onto $\mathbb{C}^n$ and let $\mathbb{G}_2(\mathbb{C}^n)$ be the Grassmannian of planes through the origin in $\mathbb{C}^n$. Suppose that the following hold:*

- *$V_1, \ldots, V_k$ have dimension s,*
- *π restricted to V_i is dominant for all i.*

Then, there exists a dense Zariski-open dense subset $\mathcal{O}_\mathcal{E}$ of $\mathbb{G}_2(\mathbb{C}^n)$ such that for all $E \in \mathcal{O}_\mathcal{E}$,

$$\overline{\pi^{-1}(E) \setminus Z} = \overline{W \setminus Z} \cap \pi^{-1}(E), \quad \dim \overline{\pi^{-1}(E) \setminus Z} = s - n + 2.$$

Proof. Observe that we can restrict ourselves to the case $k = 1$. Indeed, if $k > 1$, then we can first build the dense Zariski-open subset $\mathcal{O}_{\mathcal{E},i}$ of $\mathbb{G}_2(\mathbb{C}^n)$ for $W_i = \overline{V_i \setminus Z}$, for each i, and then take their intersection $\mathcal{O}_\mathcal{E}$, which is still a dense Zariski-open subset of $\mathbb{G}_2(\mathbb{C}^n)$.

We now assume that $\overline{W \setminus Z} = V$ is irreducible of dimension m. Let $V_H \in \mathbb{P}^N$ be the projectivisation of V. Then, the map π naturally extends to a projection map $\pi_H : V_H \to \mathbb{P}^n$. Note that π_H is a morphism of varieties since $\dim V \geq n$ and π is dominant. Hence $\dim \pi_H(V_H) + 1 > n$ and by Bertini's theorem, or an extension thereof [68, Theorem 3.3.1], the preimage of every line $L \in \mathbb{P}^n$, $\pi_H^{-1}(L)$, is irreducible in the Zariski topology of V_H. This implies that there exists a Zariski-open subset $\mathcal{O}_C$ of affine lines in $\mathbb{C}^n$ such that for all $L \in \mathcal{O}_C$, the preimage $\pi^{-1}(L)$ is irreducible. Let $\mathbb{C}[u_1, \ldots, u_n]$ be a coordinate ring of $\mathbb{C}^n$. Then, each line in $\mathcal{O}_C$ may be parametrised by the equations

$$u_1 = a_1 e_1 + b_1, \quad \ldots, \quad u_n = a_n e_1 + b_n,$$

where e_1 is a parameter and $\mathbf{a} = (a_1, \ldots, a_n), \mathbf{b} = (b_1, \ldots, b_n)$ are vectors of $\mathbb{C}^n$ outside of some proper Zariski-closed subset defined by $\mathcal{O}_C$. From each line in $\mathcal{O}_C$ we get a plane defined by the two parameter equations

$$u_1 = a_1 e_1 + b_1 e_2, \quad \ldots, \quad u_n = a_n e_1 + b_n e_2. \tag{6.3}$$

Thus, by Bertini's theorem, there exists a dense Zariski-open subset $\mathcal{O}_1$ of $\mathbb{G}_2(\mathbb{C}^n)$ so that for all $E \in \mathcal{O}_1$ the preimage $\pi^{-1}(E)$ is an irreducible section of V.

Consider $E \in \mathcal{O}_1$ and the parametrisation given by equation (6.3). Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$ be parameters and consider the ideal $I(\pi^{-1}(E))$. Since Z is a hyperplane of $\mathbb{C}^N$, there exists a linear form F such that $Z = \mathbf{V}(F)$. Then, the subset of E such that $\pi^{-1}(E) \subset Z$ is

given by the normal form of F with respect to a Gröbner basis of $I(\pi^{-1}(E))$. Either the normal form is identically zero, or we obtain a polynomial whose coefficients are polynomials in the parameters $\boldsymbol{\alpha}, \boldsymbol{\beta}$. Thus, this subset of planes that we must avoid is a Zariski-closed subset of $\mathbb{G}_2(\mathbb{C}^n)$ which we now show is not $\mathbb{G}_2(\mathbb{C}^n)$. To do so, take some $x \in V \setminus Z$. Since π is dominant, there exists some $E \in \mathbb{G}_2(\mathbb{C}^n)$ such that $x \in \pi^{-1}(E)$. Recall that $V \setminus Z$ is a dense Zariski-open subset of V. Hence, there exists a Zariski-open dense subset $\mathcal{O}_2$ of $\mathbb{G}_2(\mathbb{C}^n)$ such that for all $E \in \mathcal{O}_2$, $\pi^{-1}(E) \nsubseteq Z$.

Let $\mathcal{O}_{\mathcal{E}} = \mathcal{O}_1 \cup \mathcal{O}_2$. Then, $\mathcal{O}_{\mathcal{E}}$ is a Zariski-open dense subset of $\mathbb{G}_2(\mathbb{C}^n)$. Fix some $E \in \mathcal{O}_{\mathcal{E}}$. Then, since $\pi^{-1}(E)$ is irreducible and is not contained in Z we have that

$$\overline{\pi^{-1}(E) \setminus Z} = \overline{W \setminus Z} \cap \pi^{-1}(E).$$

Note that E has codimension $n-2$. Hence, by the genericity of E and by the theorem on the dimension of fibres [101, Theorem 1.25],

$$\dim \overline{\pi^{-1}(E) \setminus Z} = \dim \overline{W \setminus Z} \cap \pi^{-1}(E) = m - (n-2) = s - n + 2. \qquad \square$$

We aim to apply the results of Proposition 6.9 to reduce the dimension of the algebraic sets we consider in our algorithms. First, however, we give an algebraic condition that is sufficient to prove the required dominance of the projection from the graph of $M_1^j(\boldsymbol{f}^A)$ onto the $\mathbf{u}$-space. For that purpose, for given $\boldsymbol{f}, \boldsymbol{g}$ and $A \in \mathrm{GL}_n(\mathbb{K})$, define for each $1 \leq j \leq p$ the sequence of polynomials $H_j = (\boldsymbol{g}^A, z_1 \operatorname{jac}(f_j^A) - \sum_{i=1}^{m+p-1} \lambda_i \operatorname{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)_i^{[j]})$.

Lemma 6.10. *Let X be a smooth algebraic set defined by a reduced regular sequence $\boldsymbol{g} = (g_1, \ldots, g_m)$ and let $\boldsymbol{f} \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping with domain X satifying Assumption (R). Let $A \in \mathcal{O}_{\mathrm{GL}}$, where $\mathcal{O}_{\mathrm{GL}}$ is the Zariski-open subset of $\mathrm{GL}_n(\mathbb{K})$ defined in Proposition 6.8. Let π be the projection map from $\overline{\operatorname{graph} M_s^j(\boldsymbol{f}^A)}$ onto the $\mathbf{u}$-space. If the Jacobian matrix $\operatorname{jac}(H_j)$ has full rank for all j, then π is a dominant map.*

Proof. Fix some $1 \leq j \leq p$. We aim to show that the set of points in $\mathbb{C}^n$ that are not in the image of an irreducible component of $\overline{\operatorname{graph} M_s^j(\boldsymbol{f}^A)}$ by π is a proper Zariski-closed subset which would imply that the image of π is Zariski-dense and that π is dominant. By assumption on X, $\boldsymbol{g}$, $\boldsymbol{f}$ and A, we can apply Lemma 6.7, so that

$$\overline{\operatorname{graph} M_1^j(\boldsymbol{f}^A)} = \overline{\mathbf{V}(h_1, \ldots, h_{n+m+p}) \setminus \mathbf{V}(z_1)},$$

where

$$\begin{aligned}
h_i &= \operatorname{numer}(\boldsymbol{f}_i^A(\tau_1(z)) - c_i) \quad \text{for } 1 \leq i \leq p,\\
h_{p+i} &= \operatorname{numer}(\boldsymbol{g}_i^A(\tau_1(z))) \quad \text{for } 1 \leq i \leq m,\\
h_{m+p+1} &= \operatorname{jac}(f_j^A)(\tau_1(z))_i - \sum_{k=1}^{m+p-1} \lambda_k \operatorname{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)_{k,i}^{[j]}(\tau_1(z)) - z_1 u_i) \quad \text{for } 1 \leq i \leq n.
\end{aligned}$$

Let $\mathbf{a} = (a_1, \ldots, a_n)$ be a generic point of $\mathbb{C}^n$ and C be an irreducible component of $\overline{\operatorname{graph} M_1^j(\boldsymbol{f}^A)}$. We shall show the existence of a point $(z_1, \ldots, z_n, c_1, \ldots, c_p, u_1, \ldots, u_n, \lambda_1, \ldots, \lambda_{m+p-1}) \in C$ where $(u_1, \ldots, u_n) = (a_1, \ldots, a_n)$. Then, a generic point of C is one such that $z_1 \neq 0$ and so where τ_1 is invertible. Hence, at a generic point we have that $(h_{p+1}, \ldots, h_{n+m+p}) = H_j$. Consider the system of equations

$$\begin{cases}
v_1 = g_1^A, \ldots, v_m = g_m^A \\
w_1 = z_1 \frac{\partial f_j^A}{\partial z_1} - \sum_{i=1}^{m+p-1} \lambda_i \operatorname{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)_{i,1}^{[j]}, \ldots, w_n = z_1 \frac{\partial f_j^A}{\partial z_n} - \sum_{i=1}^{m+p-1} \lambda_i \operatorname{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)_{i,n}^{[j]} \\
x_1 = \sum_{i=1}^{n} b_{1,1} z_i + \sum_{i=1}^{m+p-1} b_{1,n+i} \lambda_i, \ldots, x_{p-1} = \sum_{i=1}^{n} b_{p-1,1} z_i + \sum_{i=1}^{m+p-1} b_{p-1,n+i} \lambda_i
\end{cases}$$

for generic $\mathbf{b} = (b_1, \ldots, b_{p-1}) \in \mathbb{C}^{p-1}$. Since the Jacobian of H_j has full rank, by the genericity of $\mathbf{b}$ we have that the Jacobian of the polynomials on the right-hand side of these equations has full rank. Thus, by applying the inverse function theorem to the above system, there exist equations, defined for $z_1 \neq 0$, $(\mathbf{z}, \boldsymbol{\lambda}) = (\phi_1(\mathbf{v}, \mathbf{w}, \boldsymbol{x}), \ldots, \phi_{n+m+p-1}(\mathbf{v}, \mathbf{w}, \boldsymbol{x}))$. Therefore, substituting $\mathbf{v}$ for 0, $\mathbf{w}$ for $\mathbf{a}$ and $\boldsymbol{x}$ for $\mathbf{b} \cdot (\mathbf{z}, \boldsymbol{\lambda})$, we have constructed a point $(\mathbf{z}, \boldsymbol{f}^A(\mathbf{z}), \mathbf{a}, \boldsymbol{\lambda}) \in C$. Hence, the image C of π is a Zariski-dense subset of $\mathbb{C}^n$ and so π is dominant. □

6.4 Algorithms

6.4.1 Subroutines

The algorithms in this paper rely primarily on algebraic geometric operations. By the ideal-variety correspondence, these shall be performed using ideal theoretic operations. We specify three such subroutines that will be used in our algorithms and proofs.

Eliminate$(P, \mathbf{v}, \mathbf{w})$**:**

Input: P, a finite basis of an ideal, I, of a polynomial ring (with base field $\mathbb{K}$ and two lists of indeterminates, $\mathbf{v}$ and $\mathbf{w}$) which we denote $\mathbb{K}[\mathbf{v}, \mathbf{w}]$.

Output: E, a finite basis of the ideal $I \cap \mathbb{K}[\mathbf{w}]$.

Intersect$(P_1, \ldots, P_k)$**:**

Input: $P_1, \ldots, P_k$, finite bases of ideals, $I_1, \ldots, I_k$, of a polynomial ring.

Output: P, a finite basis of the ideal $\bigcap_{i=1}^k I_i$.

Saturate(P_1, P_2)**:**

Input: P_1, P_2, finite bases of ideals, I_1, I_2, of a polynomial ring R.

Output: S, a finite basis of the ideal

$$I_1 : I_2^\infty = \{f \in R \mid \forall g \in I_2,\ \exists\, s \in \mathbb{N} \text{ such that } fg^s \in I_1\}.$$

Remark 6.11. *These ideal theoretic operations can be performed using, for example, Gröbner bases. We refer to [24, Chapter 3, Section 1, Theorem 2], [9, Proposition 6.19] and [7, 27] for algorithms for computing a finite basis for respectively elimination ideals, intersection of ideals and the saturation of ideals.*

6.4.2 Computing asymptotic critical values

To demonstrate how Algorithm 3 works, we give a simple example.

Example 6.12. *Following Example 6.2, the polynomial $f = z_1^2 + (z_1 z_2 - 1)^2$ has the asymptotic critical value 0. We will compute this value using Algorithm 3 and show that this is the only one.*

First, we generate a sufficiently random matrix A, for example

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix},$$

and we apply this change of coordinates to f to obtain $f^A = z_1^4 + 10z_1^3 z_2 + 37z_1^2 z_2^2 + 60z_1 z_2^3 + 36z_2^4 - z_1^2 - 6z_1 z_2 - 8z_2^2 + 1$. Then, we generate random vectors $\mathbf{a} = (1, 2), \mathbf{b} = (2, 3)$. From the gradient of f^A, we then construct the vector $\mathbf{v}(z)$ so that $N(z)$ is given by

$$\begin{aligned} N = \{&z_1^4 + 10z_1^3 z_2 + 37z_1^2 z_2^2 + 60z_1 z_2^3 + 36z_2^4 - z_1^2 - 6z_1 z_2 - 8z_2^2 + 1 - c_1, \\ &4z_1^4 + 30z_1^3 z_2 + 74z_1^2 z_2^2 + 60z_1 z_2^3 - 2z_1^2 - 6z_1 z_2 - e_1 - 2e_2, \\ &10z_1^4 + 74z_1^3 z_2 + 180z_1^2 z_2^2 + 144z_1 z_2^3 - 6z_1^2 - 16z_1 z_2 - 2e_1 - 3e_2\}. \end{aligned}$$

Algorithm 3: acv1

Input: $\boldsymbol{g}$ a reduced regular sequence defining a smooth algebraic set X, $\boldsymbol{f} : X \to \mathbb{K}^p$ a polynomial mapping with components in the ring $\mathbb{K}[\mathbf{z}]$ satisfying Assumption (R) and the list $\mathbf{z}$.

Output: R, a finite list of polynomials whose zero set has codimension at least 1 in $\mathbb{C}^p$ and contains the set of asymptotic critical values of $\boldsymbol{f}$.

1 Generate a random matrix $A \in \mathbb{K}^{n\times n}$ and set $\boldsymbol{f}^A \leftarrow \boldsymbol{f}(Az)$, $\boldsymbol{g}^A \leftarrow \boldsymbol{g}(Az)$.

2 Generate random vectors $\mathbf{a}, \mathbf{b} \in \mathbb{K}^n$.

3 **For** j **from** 1 **to** p **do**

4 $\quad \mathbf{v}(z) \leftarrow z_1 \operatorname{jac}(f_j^A) - \lambda_1 \operatorname{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)_1^{[j]} - \cdots - \lambda_{m+p-1} \operatorname{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)_{m+p-1}^{[j]} - \mathbf{a}e_1$.

5 $\quad N(z) \leftarrow \{f_1^A - c_1, \ldots, f_p^A - c_p, g_1^A, \ldots, g_m^A, v_1 - b_1 e_2, \ldots, v_n - b_n e_2\}$.

6 $\quad G \leftarrow \operatorname{numer}(N(\tau_1(z)))$.

7 $\quad G_s \leftarrow \operatorname{Saturate}(G, z_1)$.

8 $\quad L \leftarrow G_s \cup \{z_1, e_1, e_2\}$.

9 $\quad \mathrm{M}_j \leftarrow \operatorname{Eliminate}(L, \{\mathbf{z}, e_1, e_2, \lambda_1, \ldots, \lambda_{m+p-1}\}, \{\mathbf{c}\})$.

10 $\mathrm{R} \leftarrow \operatorname{Intersect}(\mathrm{M}_1, \ldots, \mathrm{M}_p)$.

11 **Return** R.

Next, we apply τ_1 to N and take the numerators to form G:

$$\begin{aligned} G = \{&z_1^4 - 8z_2^2 z_1^2 + 36z_2^4 - 6z_1^2 z_2 + 60z_2^3 - z_1^2 + 37z_2^2 + 10z_2 + 1 - c_1 z_1^4, \\ &- 6z_1^2 z_2 + 60z_2^3 - 2z_1^2 + 74z_2^2 + 30z_2 + 4 - e_1 z_1^4 - 2e_2 z_1^4, \\ &- 16z_1^2 z_2 + 144z_2^3 - 6z_1^2 + 180z_2^2 + 74z_2 + 10 - 2e_1 z_1^4 - 3e_2 z_1^4\}. \end{aligned}$$

We now perform the first ideal-theoretic operation. Using, for example, Gröbner bases we compute a finite basis G_s of the ideal $\langle G\rangle : \langle z_1\rangle^\infty$. Then, we form the list L by adding z_1, e_1 and e_2 to G and by similar methods as before we compute a finite basis M_j of the ideal $L \cap \mathbb{K}[c_1]$. We find that $\mathrm{M}_j = \{c_1\}$ and hence $K_\infty(f) \subseteq \{0\}$.

Following the Proof of Proposition 6.9, for two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$, we let $\mathcal{P}_{\mathbf{a},\mathbf{b}}$ be the plane spanned by these two vectors.

Theorem 6.13. *Let X be a smooth algebraic set defined by a reduced regular sequence $\boldsymbol{g} = (g_1, \ldots, g_m)$. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping with domain X satisfying Assumption (R). Suppose that $A \in \mathcal{O}_{\mathrm{GL}}$, where $\mathcal{O}_{\mathrm{GL}}$ is the Zariski-open subset of $\mathrm{GL}_n(\mathbb{K})$ defined in Proposition 6.8, and suppose that $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$ are such that $\mathcal{P}_{\mathbf{a},\mathbf{b}} \in \mathcal{O}_{\mathcal{E}}$, where $\mathcal{O}_{\mathcal{E}}$ is the Zariski-open subset of $\mathbb{G}_2(\mathbb{C}^n)$ defined in Proposition 6.9. Suppose that $\operatorname{jac}(H_j)$ has full rank for all j. Then, Algorithm 3 terminates and returns as output a finite basis whose zero set has codimension at least 1 in $\mathbb{C}^p$ and contains the set of asymptotic critical values of $\boldsymbol{f}$.*

Proof. Firstly, Algorithm 3 relies on multivariate polynomial routines that are correct and terminate, see Remark 6.11. Hence, Algorithm 3 terminates in finitely many steps. As the matrix A is taken at random in $\mathbb{K}^{n\times n}$ and $\mathcal{O}_{\mathrm{GL}}$ is a Zariski-open subset of $\mathrm{GL}_n(\mathbb{K})$ which is a Zariski-open subset of $\mathbb{K}^{n\times n}$, then, with probability 1, A lies in $\mathcal{O}_{\mathrm{GL}}$. Thus, by assumption on X, $\boldsymbol{f}$ and A, we can apply Proposition 6.8. Therefore, we aim to compute the Zariski closures of the sets $K_1^j(\boldsymbol{f}^A)$ for $1 \le j \le p$. We shall show that the algebraic sets defined by the list of polynomials M_j computed in step 9 contains $\overline{K_1^j(\boldsymbol{f}^A)}$ and has codimension at least 1 in $\mathbb{C}^p$. Then, the union of these algebraic sets, $\mathbf{V}(\mathrm{R})$ as computed in step 10, contains the asymptotic critical values of $\boldsymbol{f}$ and has codimension at least 1 in $\mathbb{C}^p$ by [24, Chapter 9, Section 4, Theorem 8].

Thus, fix some $1 \leq j \leq p$. Since A lies in $\mathrm{GL}_n(\mathbb{K})$ with probability 1, by assumption on X and $\boldsymbol{f}$, we can apply Lemma 6.7, i.e. there exists polynomials $h_1, \ldots, h_{n+m+p} \in \mathbb{K}[\mathbf{z}, \mathbf{c}, \mathbf{u}, \boldsymbol{\lambda}]$ such that

$$\overline{\text{graph } M_1^j(\boldsymbol{f}^A)} = \overline{\mathbf{V}(h_1, \ldots, h_{n+m+p}) \setminus \mathbf{V}(z_1)},$$
$$L_1^j(\boldsymbol{f}^A) = \overline{\text{graph } M_1^j(\boldsymbol{f}^A)} \cap \mathbf{V}(z_1, u_1, \ldots, u_n).$$

Let $\mathcal{P}_{\mathbf{a},\mathbf{b}}$ be the plane in the $\mathbf{u}$-space parametrised by the equations $u_i = a_i e_1 + b_i e_2$. Let $W = \mathbf{V}(h_1, \ldots, h_{n+m+p})$. Then, $\overline{\text{graph } M_1^j(\boldsymbol{f}^A)}$ is the union of the irreducible components of W that do not vanish on $\mathbf{V}(z_1)$. Then, $\overline{\text{graph } M_1^j(\boldsymbol{f}^A)}$ is equidimensional of dimension $n+p-1$ and, by Lemma 6.10, the projection map π from $\overline{\text{graph } M_1^j(\boldsymbol{f}^A)}$ onto the $\mathbf{u}$-space is dominant. Then, by Proposition 6.9, by the choice of $\mathcal{P}_{\mathbf{a},\mathbf{b}}$ we have that

$$\overline{\pi^{-1}(\mathcal{P}_{\mathbf{a},\mathbf{b}}) \setminus \mathbf{V}(z_1)} = \overline{W \setminus \mathbf{V}(z_1)} \cap \pi^{-1}(\mathcal{P}_{\mathbf{a},\mathbf{b}}).$$

Therefore, since $\mathcal{P}_{\mathbf{a},\mathbf{b}}$ contains the origin of the $\mathbf{u}$-space,

$$\begin{aligned} L_1^j(\boldsymbol{f}^A) &= \overline{\text{graph } M_1^j(\boldsymbol{f}^A)} \cap \mathbf{V}(z_1, u_1, \ldots, u_n) \\ &= \overline{W \setminus \mathbf{V}(z_1)} \cap \mathbf{V}(z_1, u_1, \ldots, u_n) \\ &= \overline{\pi^{-1}(\mathcal{P}_{\mathbf{a},\mathbf{b}}) \setminus \mathbf{V}(z_1)} \cap \mathbf{V}(z_1, u_1, \ldots, u_n) \\ &= \overline{\pi^{-1}(\mathcal{P}_{\mathbf{a},\mathbf{b}}) \setminus \mathbf{V}(z_1)} \cap \mathbf{V}(z_1, e_1, e_2). \end{aligned}$$

Thus, we may replace u_i by $a_i e_1 + b_i e_2$, its value in the parametrisation of $\mathcal{P}_{\mathbf{a},\mathbf{b}}$. By the definition of $h_1, \ldots, h_{n+m+p}$, the resulting polynomials are exactly those defined in step 6. Therefore, the algebraic set defined by L as defined in step 8 is $L_1^j(\boldsymbol{f}^A)$. Then, by [24, Chapter 4, Section 4, Theorem 4], eliminating all variables except $\mathbf{c}$ computes the closure of the projection onto the $\mathbf{c}$-space. The resulting algebraic set is exactly $\overline{K_1^j(\boldsymbol{f}^A)}$. Moreover, by Proposition 6.9, the hyperspace section $\pi^{-1}(\mathcal{P}_{\mathbf{a},\mathbf{b}})$ has dimension $(n+p-1)-(n-2) = p+1$. By the dominance of the projection onto the $\mathbf{u}$-space, e_1, e_2 are not identically zero on $\pi^{-1}(\mathcal{P}_{\mathbf{a},\mathbf{b}})$, hence $L_1^j(\boldsymbol{f}^A)$ has dimension at most $p-1$. Therefore, the projection $K_1^j(\boldsymbol{f}^A)$ onto the $\mathbf{c}$-space has dimension at most $p-1$ and so has codimension at least one in $\mathbb{C}^p$. □

Note that by step 8 of Algorithm 3, we find equations defining an algebraic set of dimension at most $p+1$. However, we then intersect with 3 hyperplanes but we only require the dimension to drop by 2. We take advantage of this behaviour in the following algorithm, which reduces the number of equations and variables by one each. This algorithm will subsequently allow us to obtain sharper degree bounds on the set of asymptotic critical values.

Example 6.14. *By again considering the polynomial $f = z_1^2 + (z_1 z_2 - 1)^2$, as in Example 6.12, we will compare Algorithms 3 and 4 by computing the asymptotic critical value* 0.

As before, we generate a random matrix A,

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix},$$

and vectors $\mathbf{a} = (1,2), \mathbf{b} = (2,3)$ and proceed by computing the gradient of $f^A = z_1^4 + 10z_1^3 z_2 + 37z_1^2 z_2^2 + 60 z_1 z_2^3 + 36 z_2^4 - z_1^2 - 6 z_1 z_2 - 8 z_2^2 + 1$. The first difference from Algorithm 3 occurs in Step 5 where we instead define the set

$$\begin{aligned} N' = \{&z_1^4 + 10z_1^3 z_2 + 37z_1^2 z_2^2 + 60 z_1 z_2^3 + 36 z_2^4 - z_1^2 - 6 z_1 z_2 - 8 z_2^2 + 1 - c_1, \\ &8z_1^4 + 58z_1^3 z_2 + 138 z_1^2 z_2^2 + 108 z_1 z_2^3 - 6 z_1^2 - 14 z_1 z_2 - e_1\}. \end{aligned}$$

Algorithm 4: acv2

Input: $\boldsymbol{g}$ a reduced regular sequence defining a smooth algebraic set X, $\boldsymbol{f} : X \to \mathbb{K}^p$ a polynomial mapping with components in the ring $\mathbb{K}[\mathbf{z}]$ satisfying Assumption (R) and the list $\mathbf{z}$.

Output: R, a finite list of polynomials whose zero set has codimension at least 1 in $\mathbb{C}^p$ and contains the set of asymptotic critical values of $\boldsymbol{f}$.

1 Generate a random matrix $A \in \mathbb{K}^{n\times n}$ and set $\boldsymbol{f}^A \leftarrow \boldsymbol{f}(Az)$, $\boldsymbol{g}^A \leftarrow \boldsymbol{g}(Az)$.

2 Generate random vectors $\mathbf{a}, \mathbf{b} \in \mathbb{K}^n$.

3 **For** j **from** 1 **to** p **do**

4 $\quad \mathbf{v}(z) \leftarrow z_1 \operatorname{jac}(f_j^A) - \lambda_1 \operatorname{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)_1^{[j]} - \cdots - \lambda_{m+p-1} \operatorname{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)_{m+p-1}^{[j]} - \mathbf{a}e_1$.

5 $\quad N'(z) \leftarrow \{f_1^A - c_1, \ldots, f_p^A - c_p, g_1^A, \ldots, g_m^A, b_2v_1 - b_1v_2, \ldots, b_nv_1 - b_1v_n\}$.

6 $\quad G' \leftarrow \operatorname{numer}(N'(\tau_1(z)))$.

7 $\quad G'_s \leftarrow \operatorname{Saturate}(G', z_1)$.

8 $\quad L' \leftarrow G'_s \cup \{z_1, e_1\}$.

9 $\quad \mathrm{M}'_j \leftarrow \operatorname{Eliminate}(L', \{\mathbf{z}, e_1, \boldsymbol{\lambda}\}, \{\mathbf{c}\})$.

10 $\mathrm{R}' \leftarrow \operatorname{Intersect}(\mathrm{M}'_1, \ldots, \mathrm{M}'_p)$.

11 **Return** R'.

Note that we now have one fewer polynomial and variable compared to Algorithm 3. Then, the set G' is defined by applying τ_1 to N' and taking the numerators,

$$G' = \{z_1^4 - 8z_2^2z_1^2 + 36z_2^4 - 6z_1^2z_2 + 60z_2^3 - z_1^2 + 37z_2^2 + 10z_2 + 1 - c_1z_1^4,\\ - 14z_1^2z_2 + 108z_2^3 - 6z_1^2 + 138z_2^2 + 58z_2 + 8 - e_1z_1^4\}.$$

As in Algorithm 3, we compute a finite basis G'_s of the ideal $\langle G'\rangle : \langle z_1\rangle^\infty$. However, since we no longer have the variable e_2, we form the list L' by adding just z_1 and e_1 to G'. Then, we compute a finite basis M'_j of the ideal $L' \cap \mathbb{K}[c_1]$ and we find that $\mathrm{M}'_j = \{c_1\}$ and hence $K_\infty(f) \subseteq \{0\}$.

To prove the correctness of this algorithm, we need the following lemma.

Lemma 6.15. *Fix some $1 \le j \le p$ and let G and G' be the list of polynomials computed at step 6 of Algorithm 3 and at step 6 of Algorithm 4 respectively for the same sufficiently generic choice of $A, \mathbf{a}$ and $\mathbf{b}$. Then,*

$$\langle G'\rangle = \langle G\rangle \cap \mathbb{C}[\mathbf{z}, e_1, \boldsymbol{\lambda}, \mathbf{c}].$$

Proof. Firstly, the polynomials $f_1^A - c_1, \ldots, f_p^A - c_p, g_1^A, \ldots, g_m^A$ at $\tau_1(z)$ are elements of both lists G and G' which are contained in the polynomial ring $\mathbb{C}[\mathbf{z}, \mathbf{c}]$. Hence, we need only consider the remaining polynomials that are in the ring $\mathbb{C}[\mathbf{z}, e_1, e_2, \boldsymbol{\lambda}]$.

We shall prove this by double inclusion. To simplify notation, we shall write v'_i for $v_i \circ \tau_1$. Firstly, take some $\operatorname{numer}(b_iv'_1 - b_1v'_i) \in G'$. We have that $b_i \operatorname{numer}(v'_1 - b_1e_2) - b_1 \operatorname{numer}(v'_i - b_ie_2) \in \langle G\rangle \cap \mathbb{C}[\mathbf{z}, e_1, \boldsymbol{\lambda}, \mathbf{c}]$. However, since v_1, v_i have the same degree in $\mathbf{z}$, $b_i \operatorname{numer}(v'_1 - b_1e_2) - b_1 \operatorname{numer}(v'_i - b_ie_2) = b_i \operatorname{numer}(v'_1) - b_1 \operatorname{numer}(v'_i) = \operatorname{numer}(b_iv'_1 - b_1v'_i) \in \langle G\rangle \cap \mathbb{C}[\mathbf{z}, e_1, \boldsymbol{\lambda}, \mathbf{c}]$.

On the other hand, let $h \in \langle G\rangle \cap \mathbb{C}[\mathbf{z}, e_1, \boldsymbol{\lambda}, \mathbf{c}]$. Let $G = \{h_1, \ldots, h_{n+m+p}\}$, then $h \in \langle G\rangle$ equals $\sum_{i=1}^{n+m+p} y_ih_i$ such that $y_i \in \mathbb{C}[\mathbf{z}, e_1, e_2, \boldsymbol{\lambda}, \mathbf{c}]$ and all e_2-terms are cancelled. Considering a monomial ordering such that e_2 is the largest monomial, $(y_1, \ldots, y_{n+m+p})$ is a syzygy on the leading terms of $h_1, \ldots, h_{n+m+p}$ that involve e_2. The S-polynomials, which are elements of $\langle G'\rangle$, generate the set of syzygies [24, Chapter 2, Section 10, Proposition 5]. Hence, $h \in \langle G'\rangle$. □

Thus, Algorithm 4 is the same as Algorithm 3 except that we eliminate the variable e_2 before computing the saturation in step 7.

Theorem 6.16. *Let X be a smooth algebraic set defined by a reduced regular sequence $\boldsymbol{g} = (g_1, \ldots, g_m)$. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping with domain X satisfying Assumption (R). Suppose that $A \in \mathcal{O}_{\mathrm{GL}}$, where $\mathcal{O}_{\mathrm{GL}}$ is the Zariski-open subset of $\mathrm{GL}_n(\mathbb{K})$ defined in Proposition 6.8, and suppose that $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$ are such that $\mathcal{P}_{\mathbf{a},\mathbf{b}} \in \mathcal{O}_{\mathcal{E}}$, where $\mathcal{O}_{\mathcal{E}}$ is the Zariski-open subset of $\mathbb{G}_2(\mathbb{C}^n)$ defined in Proposition 6.9. Suppose that $\mathrm{jac}(H_j)$ has full rank for all j. Then, Algorithm 4 terminates and returns as output a finite basis whose zero set has codimension at least 1 in $\mathbb{C}^p$ and contains the set of asymptotic critical values of $\boldsymbol{f}$.*

Proof. As in Algorithm 3, Algorithm 4 relies on multivariate polynomial routines that are correct and terminate, see Remark 6.11. Hence, Algorithm 4 terminates in finitely many steps.

Fix some $1 \leq j \leq p$ and let G and G' be the list of polynomials computed at step 6 of Algorithm 3 and at step 6 of Algorithm 4 respectively for the same sufficiently generic choice of $A, \mathbf{a}$ and $\mathbf{b}$. By assumption on X, $\boldsymbol{g}$, $\boldsymbol{f}$ and A, applying Lemma 6.10 and by assumption on $\mathbf{a}$ and $\mathbf{b}$, applying Proposition 6.9, we have that the projection from $\mathbf{V}(G)$ onto the (e_1, e_2)-space is dominant. By Lemma 6.15 and [24, Chapter 4, Section 4, Theorem 4], $\mathbf{V}(G')$ is the Zariski closure of the projection π_{e_2} of $\mathbf{V}(G)$ that eliminates e_2. Thus, $\mathbf{V}(G')$ remains two-dimensional and by Proposition 6.9, we have that

$$\overline{\mathbf{V}(G') \setminus \mathbf{V}(z_1)} = \overline{\overline{\pi_{e_2}(\mathbf{V}(G))} \setminus \mathbf{V}(z_1)} = \overline{\pi_{e_2}\left(\overline{\mathbf{V}(G) \setminus \mathbf{V}(z_1)}\right)}.$$

By [24, Chapter 4, Section 4, Theorem 10], this is equal to $\mathbf{V}(G'_s)$, where G'_s is the list computed at step 7. Therefore, there exist embeddings of the algebraic sets defined in Algorithm 4 in their counterparts defined in Algorithm 3. Thus, $\mathbf{V}(\mathrm{R}')$ contains $K_1^j(\boldsymbol{f}^A)$. It remains to show that $\mathbf{V}(\mathrm{R}')$ is contained in a proper Zariski-closed subset of $\mathbb{C}^p$.

By the dominance of the projection onto the e_1-axis, we have that e_1 is not identically zero over $\mathbf{V}(G'_s)$. Furthermore, by the saturation in step 7, z_1 is not identically zero either. Hence, $\mathbf{V}(L')$ has dimension at most $p-1$. Thus, $\mathbf{V}(\mathrm{M}'_j)$ has codimension at least 1 in $\mathbb{C}^p$ and so does $\mathbf{V}(\mathrm{R}')$. □

6.5 Degree result

Theorem 2.9. *Let X be a smooth algebraic set defined by a reduced regular sequence $\boldsymbol{g} = (g_1, \ldots, g_m)$. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping from X to $\mathbb{K}^p$ satisfying Assumption (R). Let $d = \max(\deg f_1, \ldots, \deg f_p, \deg g_1, \ldots, \deg g_m)$. Then, the asymptotic critical values of $\boldsymbol{f}$ are contained in a hypersurface of degree at most*

$$pd^{n-p-1} \sum_{i=0}^{p+1} \binom{n+p-1}{m+2p-i} d^i.$$

Proof. Let $A \in \mathrm{GL}_n(\mathbb{K})$ and $\mathbf{a}, \mathbf{b} \in \mathbb{K}^n$ be such that the genericity assumptions of Theorem 6.16 hold. That is $A \in \mathcal{O}_{\mathrm{GL}}$, the Zariski-open subset of $\mathrm{GL}_n(\mathbb{K})$ defined in Proposition 6.8 and $\mathbf{a}, \mathbf{b} \in \mathbb{K}^n$ are such that the plane $\mathcal{P}_{\mathbf{a},\mathbf{b}}$ they span is $\mathcal{O}_{\mathcal{E}}$, the Zariski-open subset of $\mathbb{G}_2(\mathbb{C}^n)$ defined in Proposition 6.9.

By assumption on X, $\boldsymbol{g}$, $\boldsymbol{f}$ and A and application of Proposition 6.8, we have that

$$K_\infty(\boldsymbol{f}) \subseteq \bigcup_{j=1}^{p} K_1^j(\boldsymbol{f}^A).$$

By Theorem 6.16, the sets $K_1^1, \ldots, K_1^p$ are contained in the algebraic sets $\mathrm{M}_1, \ldots, \mathrm{M}_p$ returned by each pass of step 9 of Algorithm 4. Thus, we shall bound the degree of $K_\infty(\boldsymbol{f})$ by the sum of the degrees of the M_i's. In order to do so, we shall compute a uniform bound on these degrees that depends on n, d, p and m so that the degree of $K_\infty(\boldsymbol{f})$ is bounded by p times this bound.

Hence, without loss of generality, we consider M_1. Then, let G, G_s, L and M_1 be the finite lists of polynomials as defined in the $j = 1$ pass of Algorithm 4 in steps 6, 7, 8 and 9 respectively.

By [24, Chapter 4, Section 4, Theorem 10], the saturation

$$\langle G_s \rangle = \langle G \rangle : \langle z_1 \rangle^\infty$$

corresponds to the variety

$$\mathbf{V}(G_s) = \overline{\mathbf{V}(G) \setminus \mathbf{V}(z_1)}.$$

Thus, $\mathbf{V}(G_s)$ is the union of a subset of the irreducible components of $\mathbf{V}(G)$. Clearly, the degree of $\mathbf{V}(L)$ is then also bounded by the degree of $\mathbf{V}(G)$. Since projection cannot increase the degree either [48, Lemma 2], we have that the degree of $\mathbf{V}(\mathrm{M}_1)$ is bounded by the degree of $\mathbf{V}(G)$.

Now, we shall bound the degree of $\mathbf{V}(G)$ by taking advantage of the multi-homogeneous structure of its defining system. Firstly, we note that G consists of $n + m + p - 1$ polynomials in $n + m + 2p$ variables. We shall split the variables into two blocks: on the one hand, $\mathbf{z}$ and, on the other hand, $\mathbf{c}, e_1, \boldsymbol{\lambda}$. Note that G consists of m multi-homogeneous polynomials of multi-degree at most $(d, 0)$ and $n + p - 1$ multi-homogeneous polynomials that have multi-degree at most $(d, 1)$. Then, by the multi-homogeneous Bézout bound [97, Proposition 3], the $(p + 1)$-equidimensional component of $\mathbf{V}(G)$ has degree at most the sum of the coefficients of the normal form of the polynomial $(dv_1 + v_2)^{n+p-1} d^m v_1^m$ with respect to the ideal $\langle v_1^{n+1}, v_2^{m+2p+1} \rangle$. Therefore, by binomial expansion, the degree of $\mathbf{V}(\mathrm{M}_1)$ is at most

$$\begin{aligned} \deg \mathbf{V}(\mathrm{M}_1) &\leq d^m \sum_{k=n-m-p-1}^{n-m} \binom{n+p-1}{k} d^k \\ &= d^{n-p-1} \sum_{i=0}^{p+1} \binom{n+p-1}{m+2p-i} d^i. \end{aligned}$$

Multiplication by p completes the proof of the bound in the statement. □

Note that for m, p, d fixed, the degree of the set of asymptotic critical values is in $O(n^{2p+m} d^n)$.

6.6 Complexity result

In this subsection, we analyse the worst-case complexity of Algorithm 4. We focus on this algorithm as it allows us to obtain the lowest degree bound. This is accomplished in Section 6.5 through the multi-homogeneous Bézout bound. Additionally, Algorithm 4 handles the fewest variables out of the algorithms given in this paper. This is important as the dimension of the ambient space is an important parameter in the complexity of these algorithms.

Firstly, let $\mathsf{M}(d)$ be the number of base field operations required for multiplying two univariate polynomials of degree at most d. For example, using the Cantor–Kaltofen algorithm, we would have that $\mathsf{M}(d) = O(d \log d \log \log d)$ [17].

Algorithm 4 takes as input a polynomial mapping $\boldsymbol{f} : X \to \mathbb{C}^p$, $\boldsymbol{f} = (f_1(z), \ldots, f_p) \in \mathbb{K}[\mathbf{z}]$, where X is a smooth algebraic set defined by a reduced regular sequence $\boldsymbol{g} = (g_1, \ldots, g_m)$. Let $d = \max(\deg \boldsymbol{f}, \deg \boldsymbol{g})$. The first steps of this algorithm, for each $1 \leq j \leq p$, is to construct a list of polynomials $h_1, \ldots, h_{n+m+p-1}$. In Section 6.5 it is proven that these polynomials define an algebraic set of degree at most

$$d^{n-p-1} \sum_{i=0}^{p+1} \binom{n+p-1}{m+2p-i} d^i.$$

This is a key quantity in our complexity result and so we denote this degree by D.

The remaining steps of Algorithm 4 involve applying algebraic elimination subroutines with the list of polynomials $h_1, \ldots, h_{n+m+p-1}$ as the initial input. The best complexity estimates are

through the use of the geometric resolution algorithm of [42]. However, since these polynomials define an algebraic set of dimension $p+1$, we must first specialise our system to obtain a zero-dimensional input for the geometric resolution algorithm. Then, we can apply the lifting algorithm of [98] to obtain a parametric representation of the $(p+1)$-dimensional system. Performing the final necessary intersections and projections of varieties is then done by computing resultants. In the end, we will obtain a polynomial whose solution set contains the set of asymptotic critical values.

Firstly, recall the representation given as the output of the geometric resolution algorithm. Consider polynomials $\varphi_1, \ldots, \varphi_\ell, \psi \in \mathbb{K}[x_1, \ldots, x_\ell]$. Suppose that $\varphi_1, \ldots, \varphi_\ell$ are a regular sequence so that the system S defined by $\varphi_1 = \cdots = \varphi_\ell = 0, \psi \neq 0$, is zero-dimensional of degree $\mathcal{D}$. Let T be a linear form in the variables $x_1, \ldots, x_\ell$. Then, with the system S as input, the geometric resolution algorithm returns a representation of the solution set of S as follows:

$$\begin{cases} Q(T) & = 0 \\ \frac{\mathrm{d}Q}{\mathrm{d}T}(T)\, x_1 & = P_1(T) \\ & \vdots \\ \frac{\mathrm{d}Q}{\mathrm{d}T}(T)\, x_\ell & = P_\ell(T), \end{cases}$$

where $Q, P_1, \ldots, P_\ell \in \mathbb{Q}[T]$ are univariate polynomials such that $\deg Q = \mathcal{D}, \deg P_i < \mathcal{D}$. Note that this representation is well-defined outside of the Zariski-closed subset $\mathbf{V}(\frac{\mathrm{d}Q}{\mathrm{d}T})$ of $\mathbb{C}^\ell$. We can now restate and prove our main complexity result.

We recall the complexity of the geometric resolution algorithm in the specialised context in which we shall use it [42, Theorem 1].

Lemma 6.17. *Let $\boldsymbol{g} = (g_1, \ldots, g_m)$ be a reduced regular sequence defining a smooth algebraic set X. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping from X to $\mathbb{K}^p$ satisfying Assumption (R). Let $d = \max(\deg f_1, \ldots, \deg f_p, \deg g_1, \ldots, \deg g_m)$ and $D = d^{n-p-1} \sum_{i=0}^{p+1} \binom{n+p-1}{m+2p-i} d^i$. Fix some $1 \leq j \leq p$ and define $(h_1, \ldots, h_{n+m+p-1})$ to be the output of step 6 of Algorithm 4. Let $L_1, \ldots, L_{p+1}$ be generic linear forms of the variables of the h_i. Let $y_1, \ldots, y_{p+1} \in \mathbb{K}$ be such that the following system S is zero-dimensional,*

$$h_1 = \cdots = h_{n+m+p-1} = 0,\ L_1 = y_1, \ldots, L_{p+1} = y_{p+1},\ z_1 \neq 0.$$

Then, a geometric resolution of this system can be computed within

$$O^{\sim}(n^2 d^{n+2} D^2)$$

arithmetic operations in the base field $\mathbb{K}$.

Proof. Let δ be the degree of the system Zariski closure of the variety defined by the system of equations

$$h_1 = \cdots = h_{n+m+p-1} = 0,\ L_1 = y_1, \ldots, L_p = y_p,\ z_1 \neq 0.$$

Since $L_1, \ldots, L_p$ are generic linear forms, we have that $\delta \leq D$. Then, since the final equation we include is $L_{p+1} = y_{p+1}$ which has degree 1, by [42, Theorem 1], computing a geometric resolution of the zero-dimensional system S requires at most

$$O((n+m+2p)((n+m+2p)\mathsf{L} + (n+m+2p)^{\Omega})\mathsf{M}(d\delta)^2)$$

arithmetic operations in $\mathbb{K}$, where L is the evaluation complexity and Ω is the exponent of the complexity of matrix multiplication. Moreover, by iterate Horner scheme, a multivariate polynomial of degree d in n variable can be evaluated in $\mathsf{L} = O(d^n)$ operations. Thus, excluding logarithmic factors, this step has complexity

$$O^{\sim}((n+m+p)^2 d^2 D^2((n+m+p)^{\Omega-1} + d^n)).$$

Finally, since $m+p \leq n$ and $d \geq 2$, the dominant term of the rightmost factor is d^n. This yields the simpler complexity estimate

$$O^{\sim}(n^2 d^{n+2} D^2) \qquad \square.$$

A particular case of interest is $p = 1$. Indeed, the study of this case allows one to tackle applications such as exact polynomial optimisation and other problems in computational real algebraic geometry. Hence, we first give a complexity result in this special case.

Theorem 2.10. *Let $\boldsymbol{g} = (g_1, \ldots, g_m)$ be a reduced regular sequence defining a smooth algebraic set X. Let $f \in \mathbb{K}[\mathbf{z}]$ be a polynomial mapping from X to $\mathbb{K}$ satisfying Assumption (R). Let $d = \max(\deg f, \deg g_1, \ldots, \deg g_m)$ and $D = d^{n-2} \sum_{i=0}^{2} \binom{n}{m+2-i} d^i$. Then, there exists an algorithm which, on input $f, \boldsymbol{g}$, outputs a non-zero polynomial $H \in \mathbb{K}[c]$ such that $K_\infty(\boldsymbol{f}) \subset \mathbf{V}(H)$ using at most*

$$O^{\sim}(n^2 d^{n+2} D^5)$$

arithmetic operations in $\mathbb{K}$.

Proof. To prove this result, we shall analyse the complexity of Algorithm 4. We aim to construct a non-zero polynomial $H \in \mathbb{K}[c]$ such that $K_\infty(\boldsymbol{f}) \subset \mathbf{V}(H)$. Thus, we begin choosing a sufficiently generic linear change of coordinates A, so that the result of Proposition 6.8 holds, and vectors $\mathbf{a}, \mathbf{b} \in \mathbb{K}^n$ so that the plane E they span is in the Zariski-open set $\mathcal{O}_{\mathcal{E}}$ as defined in Proposition 6.9. Then, with $j = 1$, let G' be the result of step 6 of Algorithm 4 so that

$$G' = (h_1, \ldots, h_{n+m}) \subset \mathbb{K}[\mathbf{z}, c, e_1, \boldsymbol{\lambda}].$$

We are interested in the irreducible components of the algebraic set defined by G' that are not contained in $\mathbf{V}(z_1)$. By Proposition 6.9 applied to $W = \operatorname{graph} M_1^1$, $Z = \mathbf{V}(z_1)$, $s = n$, $N = n+m+2$ and $\overline{\pi^{-1}(E) \setminus Z} = \mathbf{V}(G_s)$, we have that these components have dimension at most 2. However, the geometric resolution algorithm that we will rely upon requires the input system to be zero-dimensional. Thus, knowing that we can lift the result of a specialised computation, we introduce two generic linear forms of the variables $(\mathbf{z}, c, e_1, \boldsymbol{\lambda})$ that when specialised will reduce $\mathbf{V}(G')$ to a zero-dimensional algebraic set [42, 98]. Let L_1, L_2 be these linear forms. Then, with $y_1, y_2 \in \mathbb{K}$ generic, consider the zero-dimensional system:

$$h_1 = \cdots = h_{n+m} = 0, \; L_1 = y_1, L_2 = y_2, \; z_1 \neq 0.$$

Using the geometric resolution algorithm of [42], with T another linear form, we compute a representation

$$\begin{cases} q(T) & = 0 \\ \frac{\mathrm{d}q}{\mathrm{d}T}(T) c & = v_1(T) \\ \frac{\mathrm{d}q}{\mathrm{d}T}(T) z_1 & = v_2(T) \\ & \vdots \\ \frac{\mathrm{d}q}{\mathrm{d}T}(T) z_n & = v_{n+1}(T) \\ \frac{\mathrm{d}q}{\mathrm{d}T}(T) e_1 & = v_{n+2}(T) \\ \frac{\mathrm{d}q}{\mathrm{d}T}(T) \lambda_1 & = v_{n+3}(T) \\ & \vdots \\ \frac{\mathrm{d}q}{\mathrm{d}T}(T) \lambda_m & = v_{n+m+2}(T) \end{cases}$$

of this system where $q, v_1, \ldots, v_{n+m+2} \in \mathbb{K}[T]$ have degree at most D, the degree bound given in Theorem 2.9. By Lemma 6.17, this requires at most

$$O^{\sim}(n^2 d^{n+2} D^2)$$

arithmetic operations in $\mathbb{K}$.

We can then consider a lifted representation with polynomials of degree at most D in L_1, L_2, T using the algorithm given in [98].

$$\begin{cases} Q(L_1, L_2, T) & = 0 \\ \frac{\mathrm{d}Q}{\mathrm{d}T}(L_1, L_2, T)c & = P_1(L_1, L_2, T) \\ \frac{\mathrm{d}Q}{\mathrm{d}T}(L_1, L_2, T)z_1 & = P_2(L_1, L_2, T) \\ & \vdots \\ \frac{\mathrm{d}Q}{\mathrm{d}T}(L_1, L_2, T)z_n & = P_{n+1}(L_1, L_2, T) \\ \frac{\mathrm{d}Q}{\mathrm{d}T}(L_1, L_2, T)e_1 & = P_{n+2}(L_1, L_2, T) \\ \frac{\mathrm{d}Q}{\mathrm{d}T}(L_1, L_2, T)\lambda_1 & = P_{n+3}(L_1, L_2, T) \\ & \vdots \\ \frac{\mathrm{d}Q}{\mathrm{d}T}(L_1, L_2, T)\lambda_m & = P_{n+m+2}(L_1, L_2, T). \end{cases}$$

Note that P_i, Q are indeed polynomials as L_1, L_2 are generic linear forms. Hence the system is in Noether position and the number of solutions is constant for all specialisations, counted with multiplicities. We aim to compute the intersection of the variety defined by this system of equations with $\mathbf{V}(z_1, e_1)$, as in step 8 of Algorithm 4. To do so, we compute the projection of the variety defined by this system onto the (c, z_1, e_1)-space, and then will set z_1 and e_1 to zero. We accomplish this using evaluation-interpolation techniques. Specialising the L_i variables, eliminating T, and then interpolating the result. Therefore, we can in fact skip the lifting step and instead consider many different geometric resolutions by choosing different generic y_1, y_2. However, the existence of the lifted system will inform us on the degree of the polynomials we must interpolate.

Consider the first 2 equations of the specialised system, and eliminate the variable T by computing the resultant in T, $W = \mathrm{Res}_T(q, \frac{\mathrm{d}q}{\mathrm{d}T}c - v_1)$, a univariate polynomial in c. By [107, Corollary 11.21], we can compute this bivariate resultant within $O^{\sim}(D^2)$ arithmetic operations in $\mathbb{K}$. On the other hand, in the lifted system, we may express the polynomials P_1, P_2, P_{n+2}, Q as univariate polynomials in T by a Kronecker substitution, see [107, Chapter 8, Section 4]. Since L_1, L_2, T appear with degree at most D, the Kronecker substituted polynomials will have degree in the order of $O(D^3)$ in T.

Therefore, we must specialise the system in $O(D^3)$ points in y_1, y_2 and compute the same number of geometric resolutions and resultants. We then interpolate the resulting polynomials to find a polynomials $F \in \mathbb{K}[c, z_1, e_1]$. By [107, Chapter 10.2], this can be accomplished within $O^{\sim}(D^3)$ operations.

Then, define $H(c) = F(c, 0, 0)$. We have that $\mathbf{V}(H)$ contains the algebraic set defined by the result of step 9 in Algorithm 4. By assumption on $\boldsymbol{g}$, $\mathbf{V}(\boldsymbol{g})$, $\boldsymbol{f}$, A, $\mathbf{a}$ and $\mathbf{b}$ and application of Theorem 6.16, the algebraic set has codimension at least 1 in $\mathbb{C}$ and so H is non-zero. Hence, the overall complexity is dominated by computing $O(D^3)$ geometric resolutions and is in the class

$$O^{\sim}(n^2 d^{n+2} D^5). \qquad \square$$

In the case $p > 1$, we are no longer able to use a single resultant to eliminate the linear form T from the $p+1$ equations in the parametric representation we obtain from the geometric resolution algorithm. Thus, we opt for the FGLM algorithm to compute a representation where T is the greatest variable and so can be eliminated [31].

Theorem 2.11. *Let $\boldsymbol{g} = (g_1, \ldots, g_m)$ be a reduced regular sequence defining a smooth algebraic set X. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping from X to $\mathbb{K}^p$ satisfying Assumption (R). Let $d = \max(\deg f_1, \ldots, \deg f_p, \deg g_1, \ldots, \deg g_m)$ and $D = d^{n-p-1} \sum_{i=0}^{p+1} \binom{n+p-1}{m+2p-i} d^i$. Then, there exists an algorithm which, on input $\boldsymbol{f}$ and $\boldsymbol{g}$, outputs p finite lists of non-zero polynomials $G_i \subset \mathbb{K}[\mathbf{c}]$ such that $K_\infty(\boldsymbol{f}) \subset (\mathbf{V}(G_1) \cup \cdots \cup \mathbf{V}(G_p)) \subsetneq \mathbb{C}^p$ using at most*

$$O^{\sim}\left(p^2 D^{p+5} + n^2 d^{n+2} D^{p+4}\right)$$

arithmetic operations in $\mathbb{K}$.

Proof. As in the proof of Theorem 2.10, we shall analyse the complexity of Algorithm 4 and so for each $1 \leq j \leq p$, we begin with the list of polynomials

$$G' = (h_1, \ldots, h_{n+m+p-1}) \subset \mathbb{K}[\mathbf{z}, \mathbf{c}, e_1, \boldsymbol{\lambda}].$$

By the proof of Theorem 2.9, the degree of $\mathbf{V}(G')$ is at most D. Moreover, by Proposition 6.9 applied to $W = \operatorname{graph} M_1^1$, $Z = \mathbf{V}(z_1)$, $s = n+p-1$, $N = n+m+2p$ and $\overline{\pi^{-1}(E) \setminus Z} = \mathbf{V}(G_s)$, we have that the system

$$h_1 = \cdots = h_{n+m+p-1} = 0,\ z_1 \neq 0$$

defines a variety of dimension $p+1$. Hence, we introduce $p+1$ generic linear forms, $L_1, \ldots, L_{p+1}$, of the variables $(\mathbf{z}, \mathbf{c}, e_1, \boldsymbol{\lambda})$ and specialise them to generic $y_1, \ldots, y_{p+1} \in \mathbb{K}$ respectively to reduce to a zero-dimensional algebraic set. Consider a parametric representation of the system

$$h_1 = \cdots = h_{n+m+p-1} = 0,\ L_1 = y_1, \ldots, L_{p+1} = y_{p+1},\ z_1 \neq 0,$$

with T another linear form,

$$\begin{cases} q(T) & = 0 \\ \frac{\mathrm{d}q}{\mathrm{d}T}(T)c_i & = v_i(T), \quad 1 \leq i \leq p \\ \frac{\mathrm{d}q}{\mathrm{d}T}(T)z_i & = v_{p+i}(T), \quad 1 \leq i \leq n \\ \frac{\mathrm{d}q}{\mathrm{d}T}(T)e_1 & = v_{p+n+1}(T) \\ \frac{\mathrm{d}q}{\mathrm{d}T}(T)\lambda_i & = v_{p+n+1+i}(T), \quad 1 \leq i \leq m+p-1 \end{cases}$$

where $q, v_1, \ldots, v_{n+m+2p} \in \mathbb{K}[T]$ have degree at most D, the degree bound given in Theorem 2.9. By Lemma 6.17, such a representation can be computed using the geometric resolution algorithm within

$$O^{\sim}(n^2 d^{n+2} D^2)$$

arithmetic operations in the base field $\mathbb{K}$.

Consider the first $p+1$ equations of this rational parametrisation. Note that these form a Gröbner basis with respect to a lexicographic ordering with T as the least variable. Using the FGLM algorithm [31], we can compute a Gröbner basis defining the same ideal but with respect to a term ordering where T is the greatest variable, thereby eliminating T. Thus, let $\prec$ be the lexicographic monomial ordering $T > c_p > \cdots > c_1$. Let G_2 be the Gröbner basis with respect to the ordering $\prec$ returned by the FGLM algorithm with input basis $(q, \frac{\mathrm{d}q}{\mathrm{d}T}c_1 - v_1, \ldots, \frac{\mathrm{d}q}{\mathrm{d}T}c_p - v_p)$. Since the FGLM algorithm returns a reduced Gröbner basis and since the input polynomial system has degree D, we have that G_2 contains at most $(p+1)D$ polynomials. By [31, Theorem 5.1], this requires at most $O(pD^3)$ arithmetic operations in $\mathbb{K}$.

We aim to compute the intersection of the system obtained from the FGLM algorithm with $\mathbf{V}(z_1, e_1)$. To do so, we shall interpolate polynomials in $\mathbf{c}, z_1, e_1$ from different systems obtained by many specialisations. As in Theorem 2.10, by [98], there exists a lifted representation with polynomials, since we have Noether position, of degree at most D in $L_1, \ldots, L_{p+1}, T$.

$$\begin{cases} Q(L_1, \ldots, L_{p+1}, T) & = 0 \\ \frac{\mathrm{d}Q}{\mathrm{d}T}(L_1, \ldots, L_{p+1}, T)c_i & = P_i(L_1, \ldots, L_{p+1}, T), \quad 1 \leq i \leq p \\ \frac{\mathrm{d}Q}{\mathrm{d}T}(L_1, \ldots, L_{p+1}, T)z_i & = P_{p+i}(L_1, \ldots, L_{p+1}, T), \quad 1 \leq i \leq n \\ \frac{\mathrm{d}Q}{\mathrm{d}T}(L_1, \ldots, L_{p+1}, T)e_1 & = P_{p+n+1}(L_1, \ldots, L_{p+1}, T) \\ \frac{\mathrm{d}Q}{\mathrm{d}T}(L_1, \ldots, L_{p+1}, T)\lambda_i & = P_{p+n+1+i}(L_1, \ldots, L_{p+1}, T), \quad 1 \leq i \leq m+p-1. \end{cases}$$

We may express the corresponding polynomials $P_1, \ldots, P_{p+1}, P_{n+p+1}, Q$ as univariate polynomials in T by a Kronecker substitution. Since $L_1, \ldots, L_{p+1}, T$ appear with degree at most

D, the Kronecker substituted polynomials will have degree in the order of $O(D^{p+2})$ in T [107, Chapter 8.4]. Therefore, the polynomials we wish to interpolate have at most the same degree and so we must specialise the system in $O(D^{p+2})$ points in $y_1, \ldots, y_{p+1}$. We then interpolate the resulting polynomials to find polynomials $F_1, \ldots, F_{(p+1)D} \in \mathbb{K}[\mathbf{c}, z_1, e_1]$. By [107, Chapter 10.2], this can be accomplished within $O^{\sim}(D^{p+2})$ operations. Define $G_i(\mathbf{c}) = F_i(c_1, \ldots, c_p, 0, 0)$. Then, the polynomials $G_1, \ldots, G_{(p+1)D}$ define an algebraic set containing the one defined by the result of step 9 in Algorithm 4.

Therefore, for each $1 \leq j \leq p$, we output a separate list of these G_i. Hence, the overall complexity is given by calling the geometric resolution and FGLM algorithms $O(D^{p+2})$ times and so is in the class

$$O^{\sim}\left(p^2 D^{p+5} + n^2 d^{n+2} D^{p+4}\right). \qquad \square$$

6.7 Alternate description of the Jacobian condition

In this section, we develop a different interpretation of the geometric characterisation of the asymptotic critical values given in Section 7.2. Instead of a Lagrange multiplier based approach, we construct a basis of the right kernel of $\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})^{[j]}$ by introducing a matrix of new variables. Thus, define the set of variables $\mathbf{u} = (u_{i,k})_{1 \leq i \leq n, 1 \leq k \leq n-m-p+1}$ and the variable matrix

$$M_U = \begin{bmatrix} u_{1,1} & \cdots & u_{1,n-m-p+1} \\ \vdots & \ddots & \vdots \\ u_{n,1} & \cdots & u_{n,n-m-p+1} \end{bmatrix}.$$

Firstly, we introduce the equations

$$\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})^{[j]} \cdot M_U = 0,$$

so that the columns of the matrix M_U are elements of the right kernel of $\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})^{[j]}$. Then, we ensure that the matrix M_U has full rank by introducing a matrix of sufficiently generic scalars $T_U \in \mathbb{K}^{(n-m-p+1)\times n}$ and the equations

$$T_U M_U = \mathrm{Id}_{n-m-p+1},$$

where $\mathrm{Id}_{n-m-p+1}$ is the identity matrix of size $n - m - p + 1$.

Lemma 6.18. *There exists a proper Zariski open subset $\mathcal{O}_M$ of $\mathbb{K}^{(n-m-p+1)\times n}$ such that if $T_U \in \mathcal{O}_M$ then $T_U M_U = \mathrm{Id}_{n-m-p+1}$ implies that* $\mathrm{rank}(M_U) = n - m - p + 1$.

Proof. Consider an $(n - m - p + 1) \times n$ variable matrix. Then, the list of maximal minors of this matrix defines a proper Zariski closed subset of $\mathbb{K}^{(n-m-p+1)\times n}$ where the specialisations of T_U do not have full rank. Let $\mathcal{O}_M$ be the complement of this Zariski closed subset. Suppose $T_U \in \mathcal{O}_M$, then T_U has full rank and so $T_U M_U = \mathrm{Id}_{n-m-p+1}$ implies that $\mathrm{rank}(M_U) = n - m - p + 1$. $\square$

With the equations defined by Lemma 6.18, the columns of the matrix M_U are defined to be linearly independent. Therefore, since the image of the evaluation of $\mathrm{jac}(\boldsymbol{f}, \boldsymbol{g})^{[j]}$ has dimension $n - m - p + 1$ outside of a proper Zariski-closed subset of $\mathbb{C}^n$, they form a basis of the right kernel. Thus, we give Algorithm 5, an alternative version to Algorithm 4 which, as we shall prove, terminates and returns the same output.

Theorem 6.19. *Let X be a smooth algebraic set defined by a reduced regular sequence $\boldsymbol{g} = (g_1, \ldots, g_m)$. Let $\boldsymbol{f} = (f_1, \ldots, f_p) \in \mathbb{K}[\mathbf{z}]^p$ be a polynomial mapping with domain X satisfying Assumption (R). Suppose that $A \in \mathcal{O}_{\mathrm{GL}}$, where $\mathcal{O}_{\mathrm{GL}}$ is the Zariski-open subset of $\mathrm{GL}_n(\mathbb{K})$ defined in Proposition 6.8, and suppose that $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$ define a plane $E \in \mathcal{O}_{\mathcal{E}}$, where $\mathcal{O}_{\mathcal{E}}$ is the Zariski-open subset of $\mathbb{G}_2(\mathbb{C}^n)$ defined in Proposition 6.9. Suppose that $\mathrm{jac}(H_j)$ has full rank for all j. Suppose that the projection map π from $\overline{\mathrm{graph}\, M_s^j(\boldsymbol{f}^A)}$ to $\mathbb{C}^n$ is dominant. Then, Algorithm 5 terminates and returns as output a finite basis whose zero set has codimension at least 1 in $\mathbb{C}^p$ and contains the set of asymptotic critical values of $\boldsymbol{f}$.*

Algorithm 5: acv3

Input: $\boldsymbol{g}$ a reduced regular sequence defining a smooth algebraic set X, $\boldsymbol{f} : X \to \mathbb{K}^p$ a polynomial mapping with components in the ring $\mathbb{K}[\mathbf{z}]$ satisfying Assumption (R), a variable matrix M_U of size $n \times (n-m-p+1)$ with entries in the variable list $\mathbf{u}$ and the list $\mathbf{z}$.

Output: R, a finite list of polynomials whose zero set has codimension at least 1 in $\mathbb{C}^p$ and contains the set of asymptotic critical values of $\boldsymbol{f}$.

1 Generate a random matrix $T_U \in \mathbb{K}^{(n-m-p+1)\times n}$.
2 Generate a random matrix $A \in \mathbb{K}^{n\times n}$ and set $\boldsymbol{f}^A \leftarrow \boldsymbol{f}(Az)$, $\boldsymbol{g}^A \leftarrow \boldsymbol{g}(Az)$.
3 Generate random vectors $\mathbf{a}, \mathbf{b} \in \mathbb{K}^{n-m-p+1}$ and set
4 **For** j **from** 1 **to** p **do**
5 $\quad R_U \leftarrow$ List of polynomials $T_U M_U - \mathrm{Id}_{n-m-p+1}$.
6 $\quad J_U \leftarrow$ List of equations of $\mathrm{jac}(\boldsymbol{f}^A, \boldsymbol{g}^A)^{[j]} M_U$.
7 $\quad (v_1(z), \ldots, v_{n-p-m+1}(z)) \leftarrow \mathrm{jac}(f_j^A) M_U - \mathbf{a}e_1$.
8 $\quad N'(z) \leftarrow$
$\quad \{f_1^A - c_1, \ldots, f_p^A - c_p, g_1^A, \ldots, g_m^A, b_2 v_1 - b_1 v_2, \ldots, b_{n-m-p+1} v_1 - b_1 v_{n-m-p+1}\} \cup J_U$.
9 $\quad G' \leftarrow \mathrm{numer}(N'(\tau_1(z))) \cup R_U$.
10 $\quad G'_s \leftarrow \mathrm{Saturate}(G', z_1)$.
11 $\quad L' \leftarrow G'_s \cup \{z_1, e_1\}$.
12 $\quad \mathrm{M}'_j \leftarrow \mathrm{Eliminate}(L', \{\mathbf{z}, e_1, \mathbf{u}\}, \{\mathbf{c}\})$.
13 $\mathrm{R}' \leftarrow \mathrm{Intersect}(\mathrm{M}'_1, \ldots, \mathrm{M}'_p)$.
14 **Return** R'.

Proof. Since the genericity condition of Lemma 6.18 holds, the equations $R_U \cup J_U$ give conditions for the columns of the matrix M_U to define a basis for the right kernel of $\mathrm{jac}((\boldsymbol{f}^A, \boldsymbol{g}^A)^{[j]})$. Thus, the equations $f_j^A M_U$ define w_j, the restriction of the differential $\mathrm{d}f_j$ to the right kernel.

The remainder of the algorithm follows exactly Algorithm 4, with only one exception. The projection onto the $\mathbf{c}$-space now requires the elimination of the newly introduced variables $\mathbf{u}$ instead of $\boldsymbol{\lambda}$. Thus, by assumption on X, $\boldsymbol{g}$, $\boldsymbol{f}$, A, $\mathbf{a}$ and $\mathbf{b}$ and Theorem 6.16, Algorithm 5 terminates and returns as output a finite basis whose zero set has codimension at least 1 in $\mathbb{C}^p$ and contains the set of asymptotic critical values of $\boldsymbol{f}$. □

Remark 6.20. *One can perform a similar analysis of the degree of the objects computed in, and the complexity of, Algorithm 5 as is done for Algorithm 4 in Sections 6.5 and 6.6. Indeed, one can take advantage of both the multi-homogeneous structure of the polynomials constructed in step 9 as well as the number of linear forms from Lemma 6.18. Thus, using the multi-homogeneous Bézout bound, one can arrive at the following formula [97, Proposition 3].*

$$pd^m \sum_{i=0}^{n-m} \sum_{j=n-m-p-1-i}^{n-m-i} \binom{n-m}{i} \binom{(m+p-1)(n-m-p+1)}{j} d^i (d-1)^j.$$

Moreover, we saw that the number of variables is an important factor in the complexity of the geometric resolution algorithm used in Section 6.6 to perform the ideal theoretic operations [42]. Algorithm 5 works within a polynomial ring with $n(n-m-p+1)-m-p$ more variables than Algorithm 4. Hence, this leads to a worse bound on the degree and the worst-case complexity as $n \to \infty$ whenever $m+p < n$. However, as we will see in Section 6.9, there are some problems for which Algorithm 5 is faster.

6.8 Applications

6.8.1 Solving Polynomial Optimisation Problems

In this subsection we present how to use the algorithms detailed in this paper to solve polynomial optimisation problems without inequalities. Firstly, we review the problem we wish to solve. Consider a polynomial $f \in \mathbb{Q}[\mathbf{z}]$. We aim to compute the infimum of this polynomial over a smooth algebraic set X defined by a reduced regular sequence $\boldsymbol{g}$, $\inf_{\boldsymbol{x}\in X} f(\boldsymbol{x}) = f^* \in \mathbb{R} \cup \{-\infty\}$. We can solve this problem exactly by computing the generalised critical values of f restricted to X. There are three cases:

- f^* is reached. Then, f^* is a critical value of f;
- f^* is reached only at infinity, meaning that there is no minimiser $\boldsymbol{x} \in X$ but instead a path $\boldsymbol{x}_t \in \mathbb{R}^n$ that approaches the infimum as $\|\boldsymbol{x}_t\| \to \infty$. Then, f^* is an asymptotic critical value of f;
- $f^* = -\infty$.

Note that this methodology allows for the consideration of a non-compact domain X. The procedure is as follows: We first compute an algebraic representation of the generalised critical values of f restricted to X. One method to accomplish this is to compute asymptotic critical values and classical critical values separately. Firstly, we compute a polynomial whose roots contain the asymptotic critical values by using, for example Algorithm 4. Then, by the Jacobian criterion [27, Corollary 16.20], one can compute a geometric resolution of the system comprised of the polynomials $f - c$, $\boldsymbol{g}$ and the maximal minors of the Jacobian of f and $\boldsymbol{g}$, to obtain a polynomial representation of the critical values of f.

There are algebraic elimination algorithms that compute such polynomials with rational coefficients, for example Gröbner bases [24, Chapter 2] or the geometric resolution algorithm designed in [42], since we assumed that $\boldsymbol{f} \in \mathbb{Q}[\mathbf{z}]$. See Section 6.6 for a discussion on implementing Algorithm 4 using the geometric resolution algorithm. Thus, after finding a common denominator, we may assume these polynomials have integer coefficients. Then, we may use a real root isolation algorithm such as in [91], based on Descartes' rule of sign [6, Theorem 2.44], to compute isolating intervals with rational endpoints for all real roots of these polynomials.

Let $C = \{c_1, \ldots, c_k\} \subset \mathbb{R}$ be the finite set of real algebraic numbers that are the real roots of the above polynomials. Then, the set C contains the generalised critical values of f. By [58, Theorem 3.1], the polynomial f with restricted domain $f : X \setminus f^{-1}(K(f)) \to \mathbb{R} \setminus K(f)$ is a locally trivial fibration over each connected component of $\mathbb{R} \setminus K(f)$. Therefore, since C is finite, the restriction $f : X \setminus f^{-1}(C) \to \mathbb{R} \setminus C$ is also a locally trivial fibration. Hence, to decide the emptiness of each connected component of $\mathbb{R} \setminus C$, it is sufficient to decide the emptiness of one fibre for each connected component.

After computing the isolating intervals for the elements of C, we may now choose rational numbers $r_1, \ldots, r_k$ so that

$$r_1 < c_1 < r_2 < \cdots < r_k < c_k.$$

We must assess the emptiness of the fibres of these values. We do so using the algorithm designed in [95]. We consider, for $1 \leq i \leq k$, the ideal $\langle f - r_i \rangle$. This algorithm requires a radical ideal such that $\mathbf{V}(f - r_i)$ is smooth and equidimensional. Since r_i is outside of these isolating intervals, we have that $\mathbf{V}(f - r_i)$ is smooth and equidimensional. Furthermore, while $\langle f - r_i \rangle$ may not be radical, we have that $\mathbf{V}(\sqrt{\langle f - r_i \rangle}) = \mathbf{V}(f - r_i)$ and so we consider $\sqrt{\langle f - r_i \rangle}$ to decide the emptiness of $\mathbf{V}_{\mathbb{R}}(f - r_i) = \mathbf{V}(f - r_i) \cap \mathbb{R}^n$.

Firstly, if $\mathbf{V}_{\mathbb{R}}(f - r_1)$ is non-empty then we must be in the third case and so $f^* = -\infty$. For the remaining two cases, let c_j be the least critical value and let i be the least index such that $\mathbf{V}_{\mathbb{R}}(f - r_i)$ is non-empty, if such an index or critical value exist. If $r_i > c_j$, which one may decide from the isolating intervals, then c_j is the minimum of f. Otherwise, $r_i \leq c_j$ and c_{i-1} is an

asymptotic critical value and is the infimum of f. Finally, if such an index does not exist, then c_j is the minimum of f and if f also does not have any critical values, then the infimum is c_k.

We consider the complexity of the algorithm for polynomial optimisation described above. For a polynomial $f \in \mathbb{Q}[\mathbf{z}]$ and reduced regular sequence $\boldsymbol{g}$ of degree at most d, we first compute a polynomial representation of $K(f)$. With

$$D = \binom{n}{m+2} d^{n-2} + \binom{n}{m+1} d^{n-1} + \binom{n}{m} d^n,$$

by Theorem 2.10, one can compute a polynomial representation of the asymptotic critical values with complexity

$$O^{\sim}\left(n^2 d^{n+2} D^5\right).$$

By [33, Corollary 2], the set of critical values has degree at most $d^m(d-1)^{n-m}\binom{n}{m}$. Hence, with the geometric resolution algorithm, one can compute a polynomial representation of the critical values within

$$O^{\sim}\left(n^2 d^{2m+n+2}(d-1)^{2(n-m)}\binom{n}{m}^2\right)$$

arithmetic operations in $\mathbb{Q}$ [42]. By Theorem 2.9 and [33, Corollary 2], the product of these polynomials has at most

$$\Delta = \left(\binom{n}{m+2} d^{n-2} + \binom{n}{m+1} d^{n-1} + \binom{n}{m} d^n\right) + d^m(d-1)^{n-m}\binom{n}{m}$$

roots and hence f has at most this many generalised critical values. With β bounding the bit-size of the input polynomial, isolating the real roots using the algorithm designed in [91] requires at most $O(\beta\Delta^4)$ operations. We must then choose at most $d^n + 1$ points in $\mathbb{Q}$, the $r_1, \ldots, r_\Delta$ as above, and decide the emptiness of each of the real varieties $\mathbf{V}_{\mathbb{R}}(f - r_i)$. This requires the use of the algorithm designed in [95] at most Δ times with each computation requiring $O(n^7\Delta^3)$ operations. Thus, one can compute an isolating interval for the infimum of a polynomial $f \in \mathbb{Q}[\mathbf{z}]$ restricted to an algebraic set defined by a reduced regular sequence with degrees at most d in approximately $O^{\sim}(n^7\Delta^4 + n^2 d^n D^5)$ arithmetic operations in $\mathbb{Q}$.

Example 6.21. *Consider the polynomial $f = z_1^2 z_2^2 + 2z_1 z_2^3 + z_2^4 + z_1^2 + 3z_1 z_2 + 2z_2^2$. First, we compute the set of generalised critical values. Note that in this simple example it is possible to find exactly the real algebraic numbers that contain the generalised critical values because the degrees of the polynomials we compute in our algorithms are small. We find that $K_0(f) = \{0\}$ and using Algorithm 4 we find $K_\infty(f) \subset \{-\frac{1}{4}\}$. Now, to show that $f^* = -\frac{1}{4}$ one must first show that f is bounded from below. To do so, decide the emptiness of the real variety $\mathbf{V}_{\mathbb{R}}(f - r)$ for some rational number $r < -\frac{1}{4}$. For example, we can choose $r = -1$ and find that this variety is indeed empty. Finally, one must show that $-\frac{1}{4}$ truly is an asymptotic critical value as Algorithm 4 computes a superset of the asymptotic critical values. Thus, one shows that f takes values between $-\frac{1}{4}$ and 0 by deciding the emptiness of a fibre. Since the variety $\mathbf{V}_{\mathbb{R}}(f + \frac{1}{8})$ is non-empty and since*

$$f|_{\mathbb{R}^2 \setminus f^{-1}\{0,-\frac{1}{4}\}} \to \mathbb{R} \setminus \{0, -\frac{1}{4}\}$$

is a locally trivial fibration [58, Theorem 3.1], we conclude that the infimum of f is $-\frac{1}{4}$.

Example 6.22. *Consider the polynomial $f = z_1^3 + z_1^2 z_2^2 - 2z_1 z_2 + 1$. We find that $K_0(f) = \{1\}$ and $K_\infty(f) \subset \{0\}$. We first test the third case. Take a value less than 0, for example -1, and decide the emptiness of $\mathbf{V}_{\mathbb{R}}(f + 1)$. We find that this fibre is not empty and so by [58, Theorem 3.1], we conclude that $f^* = -\infty$.*

For more information on solving polynomial optimisation problems, we refer to [44, 93, 99].

6.8.2 Deciding the emptiness of semi-algebraic sets defined by a single inequality

In this subsection, we continue to explore the applications of algorithms computing generalised critical values. Let $f \in \mathbb{Q}[\mathbf{z}]$ be a polynomial with degree d and consider the semi-algebraic set S defined by the single inequality $f > 0$. The goal is to test the emptiness of the set S and in the case that S is not empty to compute at least one point in each connected component. There exists $e \in \mathbb{Q}^+$ small enough such that the problem is reduced to computing at least one point in each connected component of the real algebraic set $\mathbf{V}_{\mathbb{R}}(f - e)$. Such an e is small enough in this sense if it is less than the least positive generalised critical value of the map $z \in \mathbb{R}^n \to f(z) \in \mathbb{R}$, we refer to [92, Theorem 5.1]. To decide when this is the case, one computes isolating intervals for the generalised critical values by [6, Algorithm 10.63]. Once an appropriate e has been chosen, it remains to compute at least one point in each connected component of $\mathbf{V}_{\mathbb{R}}(f - e)$. This may be accomplished using the algorithm designed in [95]. To apply this algorithm, we require that $\langle f - e \rangle$ is radical and $\mathbf{V}(f - e)$ is equidimensional and smooth. Since e is away from any generalised critical values we have that $\mathbf{V}(f - e)$ is equidimensional and smooth. Moreover, if $\langle f - e \rangle$ is not radical, we may simply take the square-free part of $f - e$ instead as $\mathbf{V}(\sqrt{\langle f - e \rangle}) = \mathbf{V}(f - e)$.

As in the previous application, the complexity of computing isolating intervals for all real generalised critical values is in the class $O^{\sim}(n^7 d^{4n})$. After choosing an appropriate rational number e, it remains to apply the algorithm designed in [95]. This requires $O(n^7 d^3 n)$ operations. Therefore, the overall complexity of deciding the emptiness of the semi-algebraic set defined by $f > 0$ is in the class $O^{\sim}(n^7 d^{4n})$. Moreover, in the case where this set is not empty, at least one point in each connected component is computed.

Example 6.23. *Consider the polynomial $f = z_1^2(1 - z_2) - (z_1 z_2^2 - 1)^2$. Again, in this simple example we obtain polynomials of degree at most 2 from our algorithms and so we can give explicitly the set containing the generalised critical values. The polynomial giving the asymptotic critical values is c while for the critical values it is $229c^2 - 202c - 27$. Hence, we find that $K(f) \subset \{0, 1, \frac{-27}{229}\}$. We note that the value 1 is a critical value, hence we may decide immediately that the semi-algebraic set defined by $f > 0$ is non-empty. Now, to compute at least one sample point in each connected component of this set, we must choose a suitable fibre to investigate. Thus, we choose a rational value greater than 0 and less than the least critical value, such as $\frac{1}{2}$, and use the algorithm in [95] to compute sample points for each connected component of $\mathbf{V}_{\mathbb{R}}(f - \frac{1}{2})$. We may do so because $\langle f - \frac{1}{2} \rangle$ is a radical ideal. Let α be a real root of $x^4 + x - 1$. Then,*

$$(z_1, z_2) = \left(\frac{3}{4}(\alpha^3 + \alpha^2 + 1), \alpha\right)$$

is a sample point.

6.9 Experiments

The three algorithms given in this paper have been implemented in the MAPLE computer algebra system [76]. For our timing results, we use the `Groebner` package implemented in MAPLE to perform the algebraic eliminations. Alternatively, to obtain our degree results, we use MSOLVE [11], implemented in C, for the Gröbner basis computations. We present the experimental results of these implementations with computations performed on a computing server with 1536 GB of memory and an Intel Xeon E7-4820 v4 2GHz processor. To closer analyse our algebraic complexity result, all computations were performed over finite fields so as to avoid additional computation time due to coefficient growth. We choose the finite field $\mathbb{F}_{2147483647} = \mathbb{F}_{2^{31}-1}$ so that the probability of choosing bad random values in our algorithms is low. All computations that could not be completed within two days have been given the entry ∞ in Tables 6.1 and 6.2 and the entry N/A in Tables 6.3 and 6.4.

With the intention of comparison, we have attempted to implement the algorithm by Kurdyka and Jelonek, given in [58, Section 5.1], that computes the set of generalised critical values of a polynomial mapping whose domain is restricted to an algebraic set. However, this algorithm fails for some examples, such as $f = x^2 + (xy-1)^2$ restricted to $\mathbf{V}(xy-1)$. In Example 6.2, we saw that $0 \in \mathbb{K}_\infty(\boldsymbol{f})$, found along the path $(x, y) = (t, 1/t)$ as $t \to 0$. However, our implementation finds no values. Moreover, we understand that there may be some typos in the presentation of the algorithm. Based on our reading of this paper and the results obtained, we attempted another implementation fixing these mistakes. However, for the same example, we still fail to find the asymptotic critical value. On the other hand, in the global setting, one can infer an algorithm from the results of [62, Section 4] that is similar to a version of Algorithm 3 where we do not apply Proposition 6.9. This means we consider the polynomials directly as given in Lemma 6.1. Hence, an implementation of this algorithm, under the name acv0, will be compared to the algorithms designed in this paper. As we will see, this will illustrate the efficiency that Algorithms 3, 4 and 5 get from applying Proposition 6.9.

To further aid comparison, we implement versions of all these algorithms with and without a generic linear change of coordinates. This means not applying Proposition 6.8 and so we must compute np sets instead of p. However, while our complexity and degree results rely on a generic linear change of coordinates, for some problems this change can have a negative effect on the efficiency of the algorithm. This is to be expected for some sparse problems as such a generic change of coordinates destroys all structure in the input and means we perform operations on polynomials with dense support.

For our implementations of the algorithms given in this paper, and of the algorithm presented in [58, Section 5.1], see the webpage https://www-polsys.lip6.fr/~ferguson/acv_algorithms.html.

For the purpose of comparing the algorithms we develop, we introduce a number of families of polynomial mappings that have asymptotic critical values. Firstly, in the global setting, we give three families of polynomials. For $n \geq 2$, let

$$f_n = z_1^2 + \sum_{i=2}^{n} (z_1 z_i - 1)^2, \quad g_n = \sum_{i=1}^{n} \frac{\prod_{j=1}^{n} z_j^2}{z_i^2}, \quad h_n = \sum_{i=1}^{n} \prod_{j=1}^{i} z_j^{2^{i-j}}.$$

For $n \geq 2$, each of these polynomials has an asymptotic critical value at 0. For $n \geq 3$, f_n also has an asymptotic critical value at n. Additionally, we consider two families of polynomial mappings restricted to algebraic sets. For $n \geq 2$, let

$$\alpha_n : \mathbf{V}(z_1 z_2 - 1, \ldots, z_1 z_n - 1) \to \mathbb{C}, \quad \alpha_n(z) = z_1^2 + (z_1 z_2 - 1)^2 + \cdots + (z_1 z_n - 1)^2,$$
$$\beta_n : \mathbf{V}(z_1^3 z_2 \cdots z_n - 1) \to \mathbb{C}, \quad \beta_n(z) = z_1 \cdots z_n.$$

For $n \geq 3$, the map β_n has an asymptotic critical value at 0. The polynomial mapping α_n, extended from Example 6.2, also has an asymptotic critical value at 0 for all $n \geq 2$. We note that the critical locus of α_n, β_n is empty, so these asymptotic critical values are non-trivial. The system α_n has a fixed degree of 4 for all n is restricted to an algebraic set defined by $n-1$ polynomials each of degree 2. This allows us to test how our algorithms behave as we greatly increase the number of variables and the number of constraints. On the other hand, β_n has linear degree in n and has one restraint of degree $n+2$. Additionally, we compare these algorithms with random dense polynomials in both the global setting and under the restriction to a hypersurface defined by a random dense polynomial of the same degree. We denote this type of system, with degrees s in k variables, by $d_s n_k$.

6.9.1 Timing experiments

In Table 6.1 and 6.2, we see that for structured systems like α_n and β_n, the generic linear change of coordinates increases the computation time. This can be explained by two factors: Firstly,

	with A				without A			
	acv0	acv3	acv4	acv5	acv0	acv3	acv4	acv5
System	time (s)							
f_{20}	∞	3.3	2.4	220	650	3.0	1.5	230
f_{40}	∞	150	130	∞	∞	29	18	∞
f_{60}	∞	2300	1600	∞	∞	120	84	∞
g_4	∞	8.4	0.028	0.3	2700	6.7	0.044	0.86
g_6	∞	∞	19	1300	∞	∞	5.1	21000
g_8	∞	∞	83000	∞	∞	∞	1500	∞
h_3	0.46	0.21	0.020	0.070	0.068	0.20	0.017	0.20
h_4	∞	230	0.47	21000	∞	∞	0.59	∞
h_5	∞	∞	120	∞	∞	∞	4200	∞
d_2n_{20}	21	0.10	0.15	2.9	450	0.35	0.35	83
d_2n_{100}	∞	160	160	∞	∞	20	26	∞
d_3n_5	∞	13000	0.075	0.14	∞	63000	0.21	0.33
d_3n_7	∞	∞	0.42	1.6	∞	∞	1.1	8.3
d_4n_4	∞	0.13	0.38	1.2	∞	∞	1.1	0.71
d_4n_6	∞	∞	3.7	22	∞	∞	18	120

Table 6.1 – Timings for global systems given to 2 significant figures.

	with A				without A			
	acv0	acv3	acv4	acv5	acv0	acv3	acv4	acv5
System	time (s)							
α_{10}	0.82	0.15	0.075	0.039	0.66	0.21	0.20	0.092
α_{20}	53	1.3	1.2	0.61	23	2.1	2.3	1.0
α_{30}	720	9.5	10	6.8	240	9.3	9.6	4.8
α_{40}	5200	42	39	36	1600	28	29	16
α_{50}	∞	110	110	86	5300	73	75	46
α_{60}	∞	280	280	210	∞	160	150	110
β_4	5.1	0.33	0.25	0.26	0.19	0.25	0.18	0.34
β_5	300	2.0	0.67	4.0	0.75	0.97	0.67	2.6
β_6	∞	7.5	3.1	7.2	3.9	2.7	2.1	5.5
β_7	∞	41	9.9	120	21	7.2	4.2	41
β_8	∞	190	38	420	130	14	13	55
β_9	∞	1000	240	∞	1100	35	25	370
d_2n_4	18	0.37	0.026	0.072	69	1.4	0.079	0.29
d_2n_6	∞	7.2	0.10	0.35	∞	41	0.30	2.0
d_3n_3	21000	220	0.21	280	59000	670	0.59	820
d_4n_2	2.1	2.2	0.19	0.013	3.9	5.1	0.33	0.020
d_4n_4	∞	∞	5300	∞	∞	∞	22000	∞
d_6n_2	660	770	1.2	0.050	1400	1500	2.3	0.082

Table 6.2 – Timings for restricted systems given to 2 significant figures.

	Degree bound				True degree	
Sys.	acv4	Theorem 2.9	Crit. Values	[58, Theorem 4]	$K_\infty(\boldsymbol{f})$	$K(\boldsymbol{f})$
f_{20}	4	1.97×10^{13}	3.49×10^{9}	1.10×10^{12}	3	3
f_{40}	4	7.22×10^{25}	1.21×10^{19}	1.21×10^{24}	3	3
f_{60}	4	1.68×10^{38}	4.24×10^{28}	1.33×10^{36}	3	3
g_4	42	2 376	625	1 296	1	1
g_6	162	1 750 000	531 441	1 000 000	1	1
g_8	420	2 529 924 096	815 730 721	1 475 789 056	1	1
h_4	124	65 475	38 416	50 625	1	1
h_5	N/A	33 544 666	24 300 000	28 629 151	1	1
h_6	N/A	68 714 415 882	56 800 235 584	62 523 502 209	1	1
d_2n_{20}	3	61 341 696	1	1 048 576	0	1
d_2n_{100}	3	1.63×10^{33}	1	1.27×10^{30}	0	1
d_3n_5	64	918	32	243	0	32
d_3n_7	256	12 393	128	2 187	0	128
d_4n_4	135	608	81	256	0	81
d_4n_6	1215	14 080	729	4 096	0	729

Table 6.3 – Degree of asymptotic/generalised critical values in the unrestricted case.

the change of coordinates destroying the sparsity in the polynomials. Secondly, when there are many variables, the application of the linear change of variables A becomes more time consuming. For example, to solve the examples d_2n_{20} and d_2n_{100} applying A takes almost all computation time at around 0.1 and 160 seconds respectively. Similarly for the families f_n and α_n, applying the linear change of variables takes around half the time due to the large number of variables. Moreover, for generic systems, the change of coordinates effectively does not change the system. Hence, excluding the time spent applying A, the change of coordinates decreases computation time by approximately a factor of n, the number of variables, due to the algorithm computing one set instead of n sets.

Note that by considering the symmetry in the problem, one could improve the efficiency of our algorithms further. For example, for α_n and β_n, there is only one special variable, z_1. All other variables are symmetric and so the asymptotic critical values computed without a generic linear change of variables in the second to the nth set are the same. Therefore, one only needs to compute two sets, instead of n. Such symmetry reductions resulting in more efficient algorithms are a topic of future study.

From the timings presented in Table 6.2, the benefit of applying Proposition 6.9 is clear. Algorithms 3 and 4, which rely on this geometric result, are in general significantly faster than acv0. We note the special case $n = 2$, where Algorithm 3 can be slower than acv0. This is because in this setting we do not decrease the dimension of the algebraic sets we consider when we apply Proposition 6.9. However, we find that Algorithm 4 is in general faster than both acv0 and Algorithm 3. We also observe that the different formulations of this result, Algorithms 4 and 5, can have different behaviours depending on the problem. For example, Algorithm 4 computes the asymptotic critical values of β_n faster but Algorithm 5 is better at handling α_n as we increase the number of variables.

6.9.2 Degree experiments

We consider the degree of the algebraic set defined by the list of polynomials constructed in step 6 which is the basis of Theorem 2.9. Then, we give the bound of Theorem 2.9 as well as a bound on the number of critical values given in [33, Corollary 2] and compare this to the bound on the generalised critical values given in [58, Theorem 4].

In Table 6.3, we see that for unrestricted systems, the bound of [58, Theorem 4] is better.

	Degree bound				True degree	
Sys.	acv4	Theorem 2.9	Crit. Values	[58, Theorem 4]	$K_\infty(\boldsymbol{f})$	$K(\boldsymbol{f})$
α_{10}	2	30 146 560	787 320	161 414 428	1	1
α_{20}	2	1.53×10^{14}	9.30×10^{10}	4.56×10^{16}	1	1
α_{30}	2	4.53×10^{20}	8.23×10^{15}	1.29×10^{25}	1	1
α_{40}	2	1.03×10^{27}	6.48×10^{20}	3.64×10^{33}	1	1
α_{50}	2	2.00×10^{33}	4.79×10^{25}	1.03×10^{42}	1	1
α_{60}	2	3.51×10^{39}	3.39×10^{30}	2.90×10^{50}	1	1
β_4	21	6 624	3 000	7 986	1	1
β_5	25	111 475	45 360	199 927	1	1
β_6	29	2 146 304	806 736	6 075 000	1	1
β_7	33	46 707 759	16 515 072	217 238 121	1	1
β_8	37	1 136 000 000	382 637 520	8 938 717 390	1	1
β_9	41	30 575 371 299	9 900 000 000	416 051 452 971	1	1
d_2n_4	36	128	8	54	0	8
d_2n_6	64	1 184	12	486	0	12
d_3n_3	75	111	36	75	0	36
d_4n_2	32	36	24	28	0	24
d_4n_4	792	1 472	432	1 372	0	432
d_6n_2	72	78	60	66	0	60
d_6n_6	N/A	422 496	112 500	966 306	0	112 500

Table 6.4 – Degree of asymptotic/generalised critical values in the restricted case.

However, in Table 6.4 our degree bound is significantly smaller outside of a few cases where the parameters n and d are small. We note that the polynomial systems we compute in Algorithm 4 do not reach the bound of Theorem 2.9. Moreover, we are unaware of any examples of polynomial systems with a large number of asymptotic critical values, since generic systems contain no such values.

Chapter 7

On the degree of varieties of sum of squares

Abstract. We study the problem of how many different sum of squares decompositions a general polynomial f with SOS-rank k admits. We show that there is a link between the variety $\mathrm{SOS}_k(f)$ of all SOS-decompositions of f and the orthogonal group $\mathrm{O}(k)$. We exploit this connection to obtain the dimension of $\mathrm{SOS}_k(f)$ and show that its degree is bounded from below by the degree of $\mathrm{O}(k)$. In particular, for $k = 2$ we show that $\mathrm{SOS}_2(f)$ is isomorphic to $\mathrm{O}(2)$ and hence the degree bound becomes an equality. Moreover, we compute the dimension of the space of polynomials of SOS-rank k and obtain the degree in the special case $k = 2$.

This chapter contains joint work with G. Ottaviani, M. Safey El Din, E. Turatti and led to the submission of an article.

7.1 Introduction

Motivation. Let V be a complex vector space of dimension $n+1$ with basis $\{x_0 \dots, x_n\}$ and let $d \geq 0$ be an integer. Let $f \in \mathbb{C}[x_0, \dots, x_n]$ be a homogeneous polynomial of degree $2d$, that is $f \in \mathrm{Sym}^{2d} V$. A starting case, when f is real, is the problem of computing the global infimum of f, $f^* = \inf_{z \in \mathbb{R}^n} f(z)$. This problem is of principal importance in many areas of engineering and social sciences (including control theory [50, 55], computer vision [88, 1] and optimal design [25], etc.). However, even for $\deg f \geq 4$ this is an NP-hard problem [81]. As such, many methods have been developed to approximate f^*. A popular method is to relax the optimisation problem:

$$\max_{\lambda \in \mathbb{R}} \lambda \text{ s.t. } f - \lambda = \sum_{i=1}^{k} g_i^2, \quad g_i \in \mathrm{Sym}^d V.$$

Clearly, being a sum of squares implies non-negativity. It is well-known that these notions are equivalent in two homogeneous variables. However, due to the counter example by Motzkin this is not true in general [80].

In [66], using the duality between moments and sums of squares, Lasserre constructed a hierarchy of semi-definite programs whose solutions converge to the true infimum f^*. However, in general, the decompositions obtained from semi-definite programming are *approximate* certificates of non-negativity. In recent years there has been an increased study on computing *exact* certificates [87, 74]. Hence, one wants to understand the algebraic structure of SOS decompositions and the related semi-definite programs.

Prior works. Following the classical works of Sylvester [105], the study of so-called Waring decompositions, decompositions of homogeneous polynomials by powers of linear forms, is an

active area of research. In [38] it was proved that any general $f \in \mathrm{Sym}^{2d} V$ is a sum of at most 2^n squares. For fixed n, this bound is sharp for all sufficiently large d. The authors of [73] investigate the minimal numbers of squares in a decomposition of a generic polynomial in two variables. Then, in [37], the authors give a conjecture on the generic SOS-rank of polynomials, see Definition 2.13, in terms of number of variables and degree. On the other hand, in this chapter we will study generic polynomials of a given SOS-rank.

In this chapter, one aim is to analyse the degree of SOS decompositions directly from an algebraic geometry point of view. Another aim is to better understand the structure of the SOS decompositions of a given polynomial.

Main results. Recall the definitions of the two objects of interest in this chapter:

Definition 2.13. *Let* SOS_k *be the subvariety in* $\mathrm{Sym}^{2d} V$ *obtained from the Zariski closure of the set of all SOS-rank* k *polynomials.*

$$\mathrm{SOS}_k = \overline{\{f_1^2 + \cdots + f_k^2 \mid f_i \in \mathrm{Sym}^d V\}}.$$

The generic SOS-rank is the smallest number k *such that* SOS_k *covers the ambient space.*

Definition 2.14. *Let* $f \in \mathrm{SOS}_k$ *be a generic polynomial. We define the variety of all the SOS-rank* k *decompositions of* f *as*

$$\mathrm{SOS}_k(f) = \left\{ (f_1, \ldots, f_k) \in \prod_{i=1}^{k} \mathrm{Sym}^d V \,\middle|\, \sum_{i=1}^{k} f_i^2 = f \right\}.$$

While we investigate the $\mathrm{SOS}(f)$ variety for all ranks k, in particular we give a complete description of the $k = 2$ case.

Theorem 2.15. *Let* $f \in \mathrm{SOS}_2$ *be a generic polynomial of SOS-rank two. Then,* $\mathrm{SOS}_2(f)$ *has two irreducible components isomorphic to* $\mathrm{SO}(2)$*. Hence,* $\mathrm{SOS}_2(f)$ *is isomorphic to* $\mathrm{O}(2)$*.*

Since $\mathrm{SO}(k)$ acts on any decomposition using k squares, we have the inequality

$$\dim \mathrm{SOS}_k(f) \geq \dim \mathrm{SO}(k) = \binom{k}{2}.$$

In Corollary 7.17 we prove a statement which implies the following result.

Theorem 2.16. *Let* $f \in \mathrm{SOS}_k$ *be generic with* $k \leq n$*. Then,*

$$\dim \mathrm{SOS}_k(f) = \binom{k}{2}.$$

By analysing the general polynomial in SOS_2, we prove a formula for the degree of this variety.

Theorem 2.17. *Let* $N = \dim \mathrm{Sym}^d V = \binom{n+d}{d}$*. The degrees of the varieties of squares and of sum of two squares in* $\mathbb{P}(\mathrm{Sym}^{2d} V)$ *are given by*

$$\deg(\mathrm{SOS}_1) = 2^{N-1}, \ \deg(\mathrm{SOS}_2) = \prod_{i=0}^{N-3} \frac{\binom{N+i}{N-2-i}}{\binom{2i+1}{i}}.$$

Moreover, the dominant map

$$\begin{array}{rccc} \pi\colon & \prod_{i=1}^{k} \mathrm{Sym}^d V & \to & \mathrm{SOS}_k \\ & (f_1, \ldots, f_k) & \mapsto & \sum_{i=1}^{k} f_i^2 \end{array}$$

has fibers $\pi^{-1}(f) = \mathrm{SOS}_k(f)$, so that Theorem 2.16 implies the following.

Corollary 7.1.

$$\dim \mathrm{SOS}_k \leq k \binom{n+d}{n} - \binom{k}{2}$$

and equality holds for $k \leq n$ *and a general* $f \in \mathrm{SOS}_k$*.*

Structure of the chapter. In Section 7.2, we begin by recalling some definitions in sums of squares decompositions, algebraic geometry and commutative algebra. Then, in Section 7.3 we investigate the variety of all possible sums of k-squares decompositions of a given polynomial. We describe the action of the orthogonal group of size k on this variety and conjecture that there is an isomorphism between these two objects. We provide experimental and theoretical support for this conjecture and conclude by showing that it holds for $k = 2$. Finally, in Section 7.4 we use the results of Section 7.3 to prove a formula for the degree of the variety of all SOS decompositions of two squares in addition to a upper bound on this degree for $k \geq 3$.

7.2 Preliminaries

Let V be a complex vector space of dimension $n+1$. We will denote the n-dimensional projective space associated to V by $\mathbb{P}V$.

Definition 7.2. *We define the d-Veronese embedding as the map*

$$\nu_d : \mathbb{P}V \to \mathbb{P}\operatorname{Sym}^d(V),\ \ell \mapsto \ell^d.$$

Notice that the map ν_d is closed [101]. Therefore, we define the d-Veronese variety in $\mathbb{P}\operatorname{Sym}^d(V)$ as the the image of $\mathbb{P}V$ under the Veronese embedding ν_d.

Definition 7.3. *A polynomial $f \in \operatorname{Sym}^d V$ has rank one, or is decomposable, if $f = v^d$. The rank of a polynomial f is defined as the minimum number $r \in \mathbb{N}$ such that*

$$f = \sum_{i=1}^{r} v_i^d.$$

In other words, f is the sum of r decomposable polynomials.

Observe that the Veronese variety $\nu_d(V) \subset \mathbb{P}\operatorname{Sym}^d V$ consists exactly of the rank one polynomials.

Definition 7.4. *Let X be a subvariety of V. The k-th secant variety of X, denoted $\Sigma_k(X)$, is defined as the Zariski closure of the union of all the k linear subspaces spanned by points in X. That is*

$$\Sigma_k(X) = \overline{\bigcup_{x_1,\ldots,x_k \in X} \operatorname{span}\{x_1, \ldots, x_k\}}.$$

If $X = \nu_d(\mathbb{P}V) \subset \mathbb{P}\operatorname{Sym}^d V$, then the generic elements in the k-th secant variety of the Veronese variety consist exactly of polynomials of rank k as long the inclusion $\Sigma_k(\nu_d(\mathbb{P}V)) \subset \mathbb{P}\operatorname{Sym}^d V$ is strict.

Let U denote $\operatorname{Sym}^d V$. We can decompose $\operatorname{Sym}^{2d} U$ follows:

$$\operatorname{Sym}^2 U = \operatorname{Sym}^{2d} V \oplus C,$$

where C is obtained by plethysm, see [110] for more details. The space C corresponds to the quadrics on U that vanish on $\nu_d(\mathbb{P}V)$. Moreover, $\operatorname{Sym}^{2d} V$ is the degree two piece of the coordinate ring of $\nu_d(\mathbb{P}V)$.

Let $\{x_0, \ldots, x_n\}$ be a basis of V. Consider a basis $w_1 = x_0^d, w_2 = x_0^{d-1}x_1, \ldots, w_N = x_n^d$, with $N = \binom{n+d}{d}$. A rank one quadric q in $\operatorname{Sym}^2 U$ has an expression $q = (\alpha_1 w_1 + \cdots + \alpha_N w_N)^2$, with $\alpha_1, \ldots, \alpha_N \in \mathbb{C}$. Switching to the coordinates given by V we have

$$q = (\alpha_1 x_0^d + \cdots + \alpha_N x_n^d)^2.$$

This means that rank one quadrics in $\operatorname{Sym}^2 U$ correspond to square powers in $\operatorname{Sym}^{2d} V$. Furthermore, applying the same argument for a rank k quadric $f \in \operatorname{Sym}^2 U$, we see that f corresponds to a sum of k squares in $\operatorname{Sym}^{2d} V$.

Notice that if $(f_1, \dots, f_k) \in \mathrm{SOS}_k(f)$, as defined in Definition 2.14, then for any permutation $\sigma \in S_k$, where S_k is the symmetric group of order k, we have that

$$\left(f_{\sigma(1)}, \dots, f_{\sigma(k)}\right) \in \mathrm{SOS}_k(f).$$

One could desire to remove such "overlapping" points by taking the quotient by S_k. However, there is another important group, containing such permutations, that acts on $\mathrm{SOS}(f)$.

Let $\mathrm{O}(k)$ be the orthogonal group of order k. Fix a point $(f_1, \dots, f_k) \in \mathrm{SOS}_{\mathrm{k}}(f)$ and fix the ordering of the basis $\{w_1, \dots, w_N\}$ of $\mathrm{Sym}^d V$. Define A to be the $k \times N$ matrix whose i-th row is the coefficients of the polynomial f_i. Then,

$$xA^tAx^t = f.$$

Let $O \in \mathrm{O}(k)$, then the action on the left by A preserves the polynomial f. That is,

$$x(OA)^tOAx^t = f.$$

Essentially, such an action leads to a different decomposition $(f'_1, \dots, f'_k)$ of f, where f'_i is the ith row of the matrix OA.

Definition 7.5. *Let $f \in \mathrm{Sym}^d V$, let $\{x_0, \dots, x_n\}$ be a basis of V and let $\partial_0, \dots, \partial_n$ be the dual basis of $V^\vee$. For each $m < d$, we define the linear map*

$$\begin{aligned} W_f^m : &\mathrm{Sym}^m V^\vee \to \mathrm{Sym}^{d-m} V, \\ &\partial_{i_1} \cdots \partial_{i_m} \mapsto \frac{\partial f}{\partial x_{i_1} \cdots \partial x_{i_m}}. \end{aligned}$$

The matrix corresponding to this linear map is called the catalecticant *matrix of f.*

We give some cohomological definitions that are going to be used later on. Let $S = \bigoplus_q \mathrm{Sym}^q(V)$ be the symmetric algebra of V.

Definition 7.6. *Let R be a ring and F a free module of rank r over R. Given an R-linear map $k : F \to R$, the complex*

$$0 \to \bigwedge^r F \xrightarrow{\varphi_r} \bigwedge^{r-1} F \xrightarrow{\varphi_{r-1}} \dots \xrightarrow{\varphi_2} F \xrightarrow{\varphi_1} R \to 0$$

is called the Koszul complex associated to k. The maps φ_l are defined as

$$\varphi_l(e_1 \wedge \cdots \wedge e_\ell) = \sum_{i=1}^{\ell} (-1)^{i+1} k(e_i) e_1 \wedge \cdots \wedge \widehat{e_i} \wedge \cdots \wedge e_\ell,$$

where the notation $\widehat{e_i}$ means that this element is omitted from the product.

Definition 7.7. *Let M be a finitely generated graded S-module and let $F_0, \dots, F_m$ be the free S-modules that give a minimal free resolution of M. That is, there is an exact sequence*

$$0 \to F_m \to F_{m-1} \to \cdots \to F_1 \to F_0 \to M \to 0,$$

and the matrices of the maps $\phi_i : F_{i+1} \to F_i$ have no non-zero constant entry, see [27]. The Betti number $\beta_{i,j}$ is the number of generators of degree j needed to describe F_i. That is, $F_i = \bigoplus_j S(-j)^{\beta_{i,j}}$, where $S(-j)$ is the j-graded part of S.

Definition 7.8. *Let M, N be two graded S-modules and let $F_\bullet$ be a free resolution of N. Consider the complex $F_\bullet \otimes M$. The Tor groups are defined by*

$$\mathrm{Tor}_p^S(M, N) = H^p(F_\bullet \otimes M).$$

The next result shows the relation between the Tor groups of M and the Betti numbers of a free resolution of M.

Proposition 7.9. *[43, Section 1] Let $\mathfrak{m} \subset S$ be the maximal ideal $\mathfrak{m} = \bigoplus_{q\geq 1} \operatorname{Sym}^q(V)$ and let $\underline{k} = S/\mathfrak{m}$ be the residual field. Then, $\operatorname{Tor}_p^S(M, \underline{k})_q$ has rank equal to $\beta_{p,q}$.*

This connection between the Betti numbers and the Tor groups is important because it correlates the Betti numbers with cohomology. This allows us to use semi-continuity on the Betti numbers, as explained in the next theorem.

Theorem 7.10. *[47, Theorem 12.8] Let $f : X \to Y$ be a projective morphism of noetherian schemes. Let $\mathcal{F}$ be a coherent sheaf on X and flat over Y, in other words, $\mathcal{F}$ is a finitely presented $\mathcal{O}_X$-module and the functor $- \otimes \mathcal{O}_{Y,f(x)} : \operatorname{Mod}_{\mathcal{F}_x} \to \operatorname{Mod}_{\mathcal{F}_x}$ is exact for every $x \in X$. Then for each $i \geq 0$, the function*

$$y \mapsto \dim H^i(X_y, \mathcal{F}_y)$$

is upper semi-continuous on Y.

7.3 The degree of the variety of all SOS decompositions

Let $f = \sum_{i=1}^k f_i^2 \in \mathbb{C}[x_0, \ldots, x_n]$ be a sum of squares with degree $2d$. We consider the variety in the ambient space $\prod_{i=1}^k Sym^d(V)$ of all possible SOS decompositions of the given polynomial f.

$$\operatorname{SOS}_k(f) = \{(f_1, \ldots, f_k) \in \prod_{i=1}^k Sym^d(V) | \sum_{i=1}^k f_i^2 = f\}$$

We conjecture the degree of this variety, when $n \geq k$, to be the degree of the orthogonal group $\mathrm{O}(k)$. In [14] the authors give the degree of $\mathrm{SO}(k)$, and thus $\mathrm{O}(k)$, to be the determinant of the following binomial matrix

$$\deg \mathrm{O}(k) = 2^k \det\left(\left(\binom{2k-2i-2j}{k-2i}\right)_{1\leq i,j\leq \lfloor \frac{k}{2} \rfloor}\right). \tag{7.1}$$

For the case $d = 1$, the argument simplifies and so we give the following lemma.

Lemma 7.11. *Let $f \in \operatorname{Sym}^2 V$ be a quadric of* SOS*-rank $k \leq n$. Then, in the affine setting, the degree of $\operatorname{SOS}_k(f)$ is equal to the degree of $\mathrm{O}(k)$.*

Proof. With $f = \sum_{i=1}^k f_i^2$, $n \geq k$ implies that we can encode f in a $k \times (n+1)$ matrix, A, whose rows give the coefficients of the linear forms f_i. Then, with $\boldsymbol{x} = (x_0, \ldots, x_n)$ we have that $\|Ax^t\|^2 = f$. Thus, for any orthogonal matrix $O \in \mathrm{O}(k)$ we have that

$$\|OA\boldsymbol{x}^t\|^2 = (OA\boldsymbol{x}^t)^t(OA\boldsymbol{x}^t) = \boldsymbol{x}A^tO^tOA\boldsymbol{x}^t = \boldsymbol{x}A^tA\boldsymbol{x}^t = \|A\boldsymbol{x}^t\|^2 = f$$

Hence, there is an action on the $\operatorname{SOS}_k(f)$ variety by $\mathrm{O}(k)$. Additionally, there are at least two identical irreducible components that correspond to $\det O = \pm 1$.

We now show that up to a change of coordinates and multiplication by an orthogonal matrix, this SOS expression is unique. Let A and B be $k \times (n+1)$ matrices encoding SOS decomposition of f. Then, up to a change of coordinates, we can ensure that the first k columns are linearly independent and so QR decompositions can be found. Thus, let $A = Q_1R_1$ and $B = Q_2R_2$ where Q_1, Q_2 are $k \times k$ orthogonal matrices and R_1, R_2 are $k \times (n+1)$ upper triangular matrices. Then, R_1 and R_2 also encode SOS decompositions of f. By the equation $\|R_1x^t\|^2 = f$ we can identify exactly the entries of R_1, up to multiplication by ± 1 in the rows, or in other words, up to multiplication by an orthogonal matrix. The same holds for R_2 and so the decompositions encoded by A and B must be in the same orbit of the action of $\mathrm{O}(k)$ on $\operatorname{SOS}_k(f)$. Therefore, there is only one orbit and so the degree of $\operatorname{SOS}_k(f)$ is equal to the degree of $\mathrm{O}(k)$. □

The argument above also works partially for the case $d \geq 2$. Once a basis is chosen for $\mathrm{Sym}^d V$, we can construct the matrix in the same way with k rows but $\binom{n+d}{d}$ columns. Then, the group $\mathrm{O}(k)$ acts on the left to give new decompositions. However, the QR decomposition no longer implies uniqueness of the orbit. This is because there exist relations between the monomials described by the columns of f. In other words, the Gram matrix associated to f is not only symmetric but also has a moment structure. Thus, it is no longer easy to see that the non-linear equations given by the norm of Ax^t squared, $\|Ax^t\|^2$, have a unique solution.

Experimentally, up to $k \leq 6$, we observe a stabilisation of the degree of the variety $\mathrm{SOS}(f)$ as the degree of f increases. The following table derives from [14, Table 1]. Since the degree of $\mathrm{SOS}_7(f)$ is at least $233,232$ for a generic $f \in \mathrm{SOS}_7$, $k \leq 6$ is the currently the limit for our experimental methods.

k	Symbolic	Formula ($\mathrm{O}(k)$)	Formula ($\mathrm{SO}(k)$)
2	4	4	2
3	16	16	8
4	80	80	40
5	768	768	384
6	9356	9356	4768
7	-	233232	111616
8	-	6867200	3433600
9	-	393936896	196968448

Table 7.1 – Degree of $\mathrm{SOS}_k(f)$ for $n \geq k$. See formula 7.1 for the degree of $\mathrm{O}(k)$.

The next example shows that the condition $n \geq k$ is sharp.

Example 7.12. *The general plane quartic can be expressed as $g_1^2 + g_2^2 + g_3^2$ in 63 ways, where $g_i \in \mathrm{Sym}^2 \mathbb{C}^3$. A proof of such result is presented in [26, Theorem 6.2.3]. The idea is to consider the quartic form as the determinant of a 2×2 matrix whose entries are quadric forms.*

The next lemma gives an indication of the connection between k-SOS decompositions and the orthogonal group $\mathrm{O}(k)$. Indeed, fixing a matrix $A_0 \in \mathcal{M}_{k \times N}$ is equivalent to fixing a sum of squares decomposition of rank k of $f = x^t A_0^T A_0 x$.

Lemma 7.13. *Let $N \geq k \geq 1$ be integers and $A, A_0 \in \mathcal{M}_{k \times N}$ be matrices, A_0 of maximal rank and consider the entries of A as variables x_{ij}. Then the variety Y defined by the equation*

$$A^t A = A_0^t A_0$$

is isomorphic to $\mathrm{O}(k)$.

Proof. Up to an action of the group of $N \times N$ invertible matrices, $\mathrm{GL}(N)$, on the left and $\mathrm{O}(k)$ on the right of A_0, we may suppose without loss of generality that $A_0 = \begin{bmatrix} I_k & 0 \end{bmatrix}$, with I_k the $k \times k$ identity matrix and 0 a null matrix of size $k \times (N-k)$. Let $A = \begin{bmatrix} X_0 & X_1 \end{bmatrix}$, again with X_0 a $k \times k$ matrix and X_1 a $k \times (N-k)$ matrix.

In those coordinates, the variety is determined by

$$\begin{bmatrix} X_0^t X_0 & X_0^t X_1 \\ X_1^t X_0 & X_1^t X_1 \end{bmatrix} = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}.$$

The first block implies that $X_0^t X_0 \in \mathrm{O}(k)$. Moreover $X_0^t X_1 = 0$ implies that $X_1 = 0$ since X_0 is invertible. □

Let $f \in \mathrm{Sym}^{2d} V$ be a sum of k squares $f = \sum_{i=1}^k f_i^2$. Then, $f = xA^t A x^t$, where $A \in \mathcal{M}_{k \times N}$ has the coefficients of f_i as its i-th row. This gives a natural isomorphism

$$\mathrm{SOS}_k(f) \cong \{B \in \mathcal{M}_{k \times N} | xB^t B x^t = f\}. \tag{7.2}$$

Denote the Gram matrix $W_A = A^tA$ and note that $\operatorname{rk} W_A = k$ when the above decomposition is minimal. The previous lemma implies that $\mathrm{O}(k) \cong \{B \in \mathcal{M}_{k\times N} | W_B = W_A\} \subset \mathrm{SOS}_k(f)$. This, together with the isomorphism (7.2) implies that $\mathrm{SOS}_k(f)$ can be described by as many copies of $\mathrm{O}(k)$ as the number of distinct symmetric matrices W_A of rank k such that $x^tW_Ax = f$.

Let $f \in \operatorname{Sym}^{2d} V$ and $N = \binom{n+d}{d}$. Notice that the following diagram commutes.

$$\begin{array}{ccccc} & & A \mapsto xA^tAx^t & & \\ & \overset{\curvearrowright}{} & & \overset{\curvearrowright}{} & \\ \mathbb{C}^k \otimes \mathbb{C}^N & \xrightarrow[A \longmapsto A^tA]{\varphi} & \operatorname{Sym}^2(\mathbb{C}^N) & \xrightarrow[B \longmapsto xBx^t]{\pi} & \operatorname{Sym}^{2d} V \end{array} \tag{7.3}$$

We have that $\mathrm{SOS}_k(f) = \{A | \pi(A^tA) = f\}$. Moreover, if $B \in \operatorname{im}\varphi$ then $\operatorname{rk} B \leq k$.

The fiber $\pi^{-1}(f) = W_0 + C$, where W_0 is the rank N catalecticant matrix of f such that $xWx^t = f$ and C is the variety

$$C = \{C_0 \in S^N \mid x^TC_0x = 0\}.$$

This means that the problem can be reformulated in terms of the intersection $\varphi(\mathbb{C}^k \otimes \mathbb{C}^N) \cap (C + W_0)$: when this intersection is just a single point, as is the case for $k \leq 6$ shown in Table 7.1, this implies that there exists only one $C_0 \in C$ such that $W_0 + C_0$ has rank k. This is equivalent to saying that $\mathrm{SOS}_k(f)$ consists of a single copy of $\mathrm{O}(k)$. Thus, we arrive at the following conjecture.

Conjecture 7.14. *Let $f \in \operatorname{Sym}^{2d} V$ be generic of SOS-rank $k \leq n$ and let $N = \binom{n+d}{d}$. Then,*

$$\mathrm{SOS}_k(f) \cong \{A \in \mathbb{C}^k \otimes \mathbb{C}^N \mid A^tA = W,\ xWx^t = f\} = \{A \in \mathbb{C}^k \otimes \mathbb{C}^N \mid A^tA = W_0 + C_0\} = \mathrm{O}(k).$$

Of course, if this intersection consists of more than a single point, one would arrive at exactly the number of copies of $\mathrm{O}(k)$ such that $\mathrm{SOS}_k(f)$ is isomorphic.

Consider a tuple $(f_1, \ldots, f_k) \in \mathrm{SOS}_k(f)$, we denote the tangent space of $\mathrm{SOS}_k(f)$ at this point by $\mathrm{TSOS}_k(f)_{(f_1,\ldots,f_k)}$. Recall that if we consider an orthogonal matrix $O \in \mathrm{O}(k)$ and $A_f \in \mathcal{M}_{k\times N}$, then the rows of A_fO are polynomials giving a k-SOS decomposition of f.

We are interested in understanding the local behavior of this variety. More specifically, we want to show that the tangent space $\mathrm{TSOS}_k(f)_{(f_1,\ldots,f_k)}$ has dimension equal to the dimension of $\mathrm{O}(k)$. This means that locally, the variety $\mathrm{SOS}_k(f)$ is exactly equal to $\mathrm{O}(k)$. In order to do that, we can show that the only syzygies of a vector $(f_1, \ldots, f_k) \in \mathrm{SOS}_k(f)$ are given by the Koszul syzygies. In the next paragraphs, we further explain the concept of Koszul syzygies and how they are related to the tangent space of $\mathrm{SOS}_k(f)$.

Let A_f be the matrix whose rows are the coefficients of $f_1, \ldots, f_k$. Observe that the map

$$\phi : A \mapsto xA^tAx^t - f$$

gives $\mathrm{SOS}_k(f)$ as the fiber at zero. Therefore, the tangent space $\mathrm{TSOS}_k(f)_{(f_1,\ldots,f_k)}$ is the space generated by the nullity of the derivative of ϕ at the point $(f_1, \ldots, f_k)$. This equivalent to saying that

$$x(A_f^tV + V^tA_f)x^t = 0 \tag{7.4}$$

where $V \in \mathcal{M}_{k\times N}$. Notice that equation (7.4) is trivially satisfied when A_f^tV is a skew-symmetric matrix. A syzygy satisfying this equation is a Koszul syzygy of $(f_1, \ldots, f_k)$. If we have that the Koszul syzygies are the only syzygies of the point $(f_1, \ldots, f_k)$, we obtain that they span the tangent space at this point. In such case, the tangent space has dimension equal to the dimension of $\mathrm{O}(k)$.

A more geometric and intuitive explanation can be described by looking at the usual set of coordinates instead of matrices. We may see $\mathrm{SOS}_k(f)$ as the nullity of the map

$$\varphi : (h_1, \ldots, h_k) \mapsto \sum_{i=1}^{k} h_i^2 - f.$$

The tangent space $\mathrm{TSOS}_k(f)_{(f_1,\dots,f_k)}$ is computed once again as the space generated by the nullity of the derivative of the expression $\sum_{i=1}^k h_i^2 - f$ at the point $(f_1,\dots,f_k)$, that is $\varphi'(f_1,\dots,f_k)=0$. This means that the tangent space is generated by

$$\sum_{i=1}^{k} f_i g_i = 0.$$

The vanishing of this expression by considering tuples $(g_1,\dots,g_k)$ such that we have pairs $i \neq j$ with $g_i = f_j$ and $g_j = -f_i$ is a Koszul syzygy of the vector $(f_1,\dots,f_k)$. Observe that this corresponds exactly to the matrix $A_f^t V$ being skew-symmetric, where V is the matrix that has g_i as the ith-row.

Proposition 7.15. *The only syzygies of the vector $(x_0^d,\dots,x_k^d)$ are the Koszul syzygies.*

Proof. Let $A = \begin{bmatrix} I & 0 \end{bmatrix}$ be a matrix as in equation (7.2) giving a SOS decomposition of $f = x_0^{2d}+\cdots+x_k^{2d}$ in a basis $\{x_0^d,\dots,x_n^d,\dots\}$, where I is the $k\times k$-identity matrix. Consider $V = \begin{bmatrix} v_{ij} \end{bmatrix}$ a $k \times N$-matrix, then $\partial\varphi(A) = V^tA + A^tV$ is the derivative of φ. The statement is equivalent to show that $x\partial\varphi(A)x^t = 0$ if and only if V^tA is skew-symmetric.

In this basis, $W = V^tA + A^tV = \begin{bmatrix} v_{ij}+v_{ji} \end{bmatrix}$, and

$$xWx^t = \sum_{i=1}^{k}\left(\sum_{j=1}^{k}(v_{ij}+v_{ji})x_j^d\right)x_i^d = 0,$$

since each monomial coefficient is equal to zero we obtain $2(v_{ij}+v_{ji})=0$ as desired. □

The importance of this result is that it guarantees that at the point $(x_0^d,\dots,x_k^d)$ the tangent space to $\mathrm{SOS}_k(x_0^{2d}+\cdots+x_k^{2d})$ has dimension equal to the number of Koszul syzygies, since they span the null space of $\varphi'(x_0^d,\dots,x_k^d)$. Moreover, this dimension is exactly equal to the dimension of the tangent space of $\mathrm{O}(k)$. This implies that locally at the point $(x_0^d,\dots,x_k^d)$, the variety $\mathrm{SOS}_k(f)$ is equal to the subvariety $\mathrm{O}(k)\subset\mathrm{SOS}_k(f)$. We wish to extend this result to every point $(f_1,\dots,f_k)\in\mathrm{SOS}_k(f)$. We obtain that this can be extended to a vector $(f_1,\dots,f_k)$ by means of semi-continuity. Indeed, consider the kernel K of the map

$$\mathcal{O}_{\mathbb{P}V}(-d)^k \xrightarrow{[f_1\cdots f_k]} \mathcal{O}_{\mathbb{P}V}$$

defined by the vector $(f_1,\dots,f_k)$, where $\mathcal{O}_{\mathbb{P}V}$ is the sheaf defining $\mathbb{P}V$ as a scheme $(\mathbb{P}V,\mathcal{O}_{\mathbb{P}_V})$. The minimal resolution of the kernel, when there are only Koszul syzygies, start with

$$\dots\to\mathcal{O}_{\mathbb{P}V}(-2d)^{\binom{k}{2}}\to K\to 0$$

By Proposition 7.9, the Betti numbers $\beta_{p,p+q}$ of the minimal resolution of K correspond to the rank of $\mathrm{Tor}_p^S(K,\underline{k})_{p+q}$, this is the component of degree $p+q$ of $Tor_p^S(K,\underline{k})$. Since we can correlate the Betti numbers with cohomology dimensions using Proposition 7.9, we have by Theorem 7.10 that for a local deformation of K, the Betti numbers satisfy semi-continuity. Moreover, since we know that for any other point $(f_1,\dots,f_k)$ will have at least the Koszul syzygies, this implies that it will have only them.

Corollary 7.16. *Suppose that $k\le n$ and $f\in\mathrm{SOS}_k$ is general. Let $(f_1,\dots,f_k)$ be a vector in $(\mathrm{Sym}^d V)^{\times k}$ giving the decomposition as k sum of squares of a polynomial f. Then the only syzygies of $(f_1,\dots,f_k)$ are the Koszul ones.*

Corollary 7.17. *Suppose that $k\le n$ and $f\in\mathrm{SOS}_k$ is general. We have an isomorphism $\mathrm{SOS}_k(f)\cong\mathrm{O}(k)^p$, for some $p\in\mathbb{Z}_+$. Note that this does not depend on the degree of f. In particular $\deg\mathrm{SOS}_k(f)\ge\deg\mathrm{O}(k)$ which is computed in Table 7.1, in fact $\deg\mathrm{SOS}_k(f)\equiv 0 \mod \deg O(k)$.*

We notice that the diagram 7.3 can have its conclusion interpreted in a different manner. Instead of considering W_0 a maximal rank matrix, one may consider a fixed matrix A_0 defining f, and let

$$\mathrm{SOS}_k(f) = \{B^TB + C_0 \mid \mathrm{rank}(B^TB + C_0) = k,\ B^TB = A_0^TA_0 \text{ and } C_0 \in C\}.$$

Notice that such interpretation means that adding $C_0 \neq 0$ is equivalent to changing the $\mathrm{O}(k)$ component of $\mathrm{SOS}_k(f)$. Thus, if there exists no other matrix C_0 besides 0 such that $\mathrm{rank}(A_0^TA_0 + C_0) = k$, it implies that there exists only one component.

In the next pages we explore this equivalent problem and compare the dimensions of symmetric matrices of rank k and C. Although a proof that the only translation by C preserving the rank is 0 is not obtained, by a comparison of dimensions we get a clear indicator that we should not expect other solutions.

Let S_k^N be the variety of symmetric matrices of size $N = \binom{n+d}{d}$ of rank at most k. Then, for some fixed $W \in S_k^N$, consider the variety

$$(S + W)_k^N = \{B \mid B + W \in S_k^N\}.$$

Note that this is indeed a variety as it is defined by the minors of the matrix $B+W$ and moreover, for all $M \in (S+W)_k^N$, we have that $M - W \in S_k^N$. Hence, we can consider this variety a translation of S_k^N by the matrix W.

$$(S + W)_k^N = S_k^N - W.$$

Recall the variety $C = \{C_0 \in S^N \mid x^TC_0x = 0\}$ and note that the following statement holds:

$$\text{For a generic } W,\ (S+W)_k^N \cap C = 0 \iff \text{For a generic } f,\ \deg \mathrm{SOS}_k(f) = \deg \mathrm{O}(k).$$

Firstly, note that since W is symmetric of rank k, there exists a decomposition of the form $W = A^TA$ where $A \in \mathcal{M}_{k\times N}$. Then, since every symmetric matrix of size N gives a polynomial, through a moment vector x, we obtain a decomposition of x^TWx as a sum of k squares as x^TA^TAx. Then, as is discussed above, we would obtain equality for Corollary 7.17.

From the translation argument above, we obtain the following equivalences,

$$(S+W)_k^N \cap C = 0 \iff (S_k^N - W) \cap C = 0 \iff S_k^N \cap (C + W) = W.$$

The equations defining C are not general. Each equation specifies that a particular coefficient in the expansion of x^tBx be zero. Hence, no coefficients of a general f are zero, we have that a generic W is not contained in the hyperplanes defined by any of the $\binom{n+2d}{2d}$ equations defining C.

Let $N = \binom{n+d}{d}$ and let S be the polynomial ring $\mathbb{C}[x_{ij} | 1 \leq i, j \leq N]$. We set $x_{ij} = x_{ji}$ and consider $X = (x_{ij})_{1\leq i,j\leq N}$ to be an $N \times N$ variable symmetric matrix. For $1 \leq k \leq N-1$, we denote by I_k the ideal generated by the $k+1$ minors of X. It is known that S/I_k is a Cohen-Macaulay normal domain with dimension

$$\dim S/I_k = \frac{(2N + 1 - k)k}{2}.$$

Then, recall that

$$\mathrm{Sym}^2(\mathrm{Sym}^d V) = \mathrm{Sym}^{2d} V \oplus C.$$

Thus,

$$\mathrm{codim}\, C = \dim \mathrm{Sym}^{2d} V = \binom{n+2d}{2d}.$$

The following lemma, through a dimension count, gives further support for Conjecture 7.14.

Lemma 7.18. *Let $k \leq n$. Then, for all $n, d \geq 1$, $\dim S/I_k < \mathrm{codim}\, C$.*

Proof. Firstly, note that $\dim S/I_k$ is maximal when $k = N = \binom{n+d}{d}$ and that the dimension decreases monotonically as k decreases. However, since we restrict to $k \leq n$, it suffices to show that $\dim S/I_n < \operatorname{codim} C$. Now, suppose that $d = 1$. Then,

$$\begin{aligned}\dim S/I_n - \operatorname{codim} C =& \frac{(2(n+1)+1-n)n}{2} - \frac{(n+2)(n+1)}{2} \\ =& \frac{n(n+3)-(n+1)(n+2)}{2} \\ =& -1.\end{aligned}$$

Next, consider $d \geq 2$. Note that for all $n \geq 1$,

$$\frac{(2\binom{n+d}{d}+1-n)n}{2} \leq n\binom{n+d}{d}.$$

Hence, it suffices to prove that for all $d \geq 2$,

$$\binom{n+2d}{2d} > n\binom{n+d}{d}.$$

We proceed by induction on d. In the base case $d = 2$ we have,

$$\begin{aligned}\binom{n+4}{4} - n\binom{n+2}{2} &= \frac{(n+4)(n+3)(n+2)(n+1)}{4!} - \frac{(n+2)(n+1)n}{2!} \\ &= \frac{(n+1)(n+2)}{4!}(n^2-5n+12).\end{aligned}$$

It is easy to see that the polynomial $n^2 - 5n + 12$ is positive for all n and so the base case holds. Now, assume for some fixed $d \geq 2$ that $\binom{n+2d}{2d} > n\binom{n+d}{d}$ and consider

$$\begin{aligned}\binom{n+2d+2}{2d+2} - n\binom{n+d+1}{d+1} &= \frac{(n+2d+2)(n+2d+1)}{(2d+2)(2d+1)}\binom{n+2d}{2d} - \frac{n+d+1}{d+1}n\binom{n+d}{d} \\ &= n\binom{n+d}{d}\left(\frac{(n+2d+2)(n+2d+1)}{(2d+2)(2d+1)}\frac{\binom{n+2d}{2d}}{n\binom{n+d}{d}} - \frac{n+d+1}{d+1}\right) \\ &> n\binom{n+d}{d}\left(\frac{(n+2d+2)(n+2d+1)}{(2d+2)(2d+1)} - \frac{n+d+1}{d+1}\right) \\ &= n\binom{n+d}{d}\left(\left(1+\frac{n}{2d+2}\right)\left(1+\frac{n}{2d+1}\right) - \left(1+\frac{n}{d+1}\right)\right) \\ &> n\binom{n+d}{d}\left(\left(1+\frac{n}{2d+2}\right)^2 - \left(1+\frac{n}{d+1}\right)\right) \\ &> n\binom{n+d}{d}\left(\left(1+\frac{2n}{2d+2}\right) - \left(1+\frac{n}{d+1}\right)\right) = 0.\end{aligned}$$

Thus, by induction, $\operatorname{codim} C - \dim S/I_k > 0$. □

We finish this section by proving that Conjecture 7.14 holds for $k = 2$.

Theorem 2.15. *Let $f \in \mathrm{SOS}_2$ be a generic polynomial of SOS-rank two. Then, $\mathrm{SOS}_2(f)$ has two irreducible components isomorphic to $\mathrm{SO}(2)$. Hence, $\mathrm{SOS}_2(f)$ is isomorphic to $\mathrm{O}(2)$.*

Proof. Let $f \in \operatorname{Sym}^{2d} V$ be a general polynomial such that $f = g^2 + h^2 = (g+ih)(g-ih)$. Since this factorization is unique (UFD), we have for any other g', h' such that $f = g'^2 + h'^2$, then $\lambda(g'+ih') = g+ih$ and $\lambda^{-1}(g'-ih') = g-ih$, or $\lambda(g'+ih') = g-ih$ and $\lambda^{-1}(g'-ih') = g+ih$.

Consider the first set of conditions, then $g = g'\frac{\lambda+\lambda^{-1}}{2} + h'\frac{i(\lambda-\lambda^{-1})}{2}$ and $h = g'\frac{\lambda-\lambda^{-1}}{2i} + h'\frac{\lambda+\lambda^{-1}}{2}$. Thus

$$\begin{bmatrix} g \\ h \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{\lambda+\lambda^{-1}}{2} & \frac{i(\lambda-\lambda^{-1})}{2} \\ \frac{\lambda-\lambda^{-1}}{2i} & \frac{\lambda+\lambda^{-1}}{2} \end{bmatrix}}_{A} \begin{bmatrix} g' \\ h' \end{bmatrix}.$$

Since $\det(A) = 1$ and $AA^t = I$, this corresponds to one component of $\mathrm{O}(2)$. Then, the last copy of $\mathrm{SO}(2)$ is obtained from the other two conditions. □

7.4 The degree of the variety of the sum of two squares

Let V be a complex vector space of dimension $n+1$ and let $d \geq 0$ be an integer. Let $U = \mathrm{Sym}^d V$ and π_C be the projection of $\mathrm{Sym}^2 U = \mathrm{Sym}^{2d} V \oplus C$ centered at C, that is

$$\pi_C : \mathrm{Sym}^{2d} V \oplus C \to \mathrm{Sym}^{2d} V.$$

Notice that whenever $\Sigma_k(\nu_2(U)) \cap C = 0$, $\pi_C|_{\Sigma_k(\nu_2(U))}$ is a well-defined morphism. In such a case, assuming $\pi_C|_{\Sigma_k(\nu_2(U))}$ is an isomorphism, this means that

$$\deg(\mathrm{SOS}_k) = \deg\big(\Sigma_k(\nu_2(U))\big).$$

Theorem 7.19. *Following the previous notation we have that*

$$\Sigma_1(\nu_2(\mathbb{P}U)) \cap C = \emptyset \quad \text{and} \quad \Sigma_2(\nu_2(\mathbb{P}U)) \cap C = \emptyset.$$

Proof. It is known from the Borel-Weil Theorem, see [110], that whenever $X = G/P \subset \mathbb{P}(V_\lambda)$, where G is an algebraic group and $P \subset G$ a parabolic group, then $H^0(X, \mathbb{P}(V_\lambda(k))) = V_{k\lambda}$ [100]. If we consider $X = \nu_d(\mathbb{P}V) \subset \mathbb{P}U$ and $\mathrm{Sym}^{2d} V = V_{2d}$, we have the short exact sequence

$$0 \to \mathcal{I}_X \to \mathcal{O}_{\mathbb{P}U} \to \mathcal{O}_X \to 0.$$

Twisting it by $\mathcal{O}_{\mathbb{P}U}(2)$ and taking the long exact sequence of cohomologies we obtain

$$0 \to H^0(\mathcal{I}_X(2)) \to H^0(\mathcal{O}_{\mathbb{P}U}(2)) \to H^0(X, \mathcal{O}_X(2)) \to 0.$$

Notice that the last map is a surjection since $H^0(X, \mathcal{O}_X(2)) = V_{2d}$ that is irreducible.

We remark the follow identifications: $H^0(\mathcal{I}_X(2))$ is given by the quadric forms on the ideal sheaf of X, that is, the quadric forms that belong to C. $H^0(\mathcal{O}_{\mathbb{P}U}(2)) = \mathrm{Sym}^2(U) = \mathrm{Sym}^2(\mathrm{Sym}^d V)$ and $H^0(X, \mathcal{O}_X(2)) = \mathrm{Sym}^{2d} V$.

Assume for the purpose of contradiction that $\Sigma_1(\nu_2(\mathbb{P}U)) \cap C \neq \emptyset$ and $\Sigma_2(\nu_2(\mathbb{P}U)) \cap C \neq \emptyset$, this means that there exists polynomials $f, g \in C$ of respective ranks 1 and 2 in $\mathbb{P}\,\mathrm{Sym}^2 U$ such that $f, g \in H^0(\mathcal{I}_X(2))$. This implies that X is contained in the hyperplane determined by $f = l^2$ and in the union of hyperplanes determined by $g = l_0^2 + l_1^2 = (l_0 + il_1)(l_0 - il_1)$, where l, l_0, l_1 are linear forms in $\mathbb{P}U$. However $X = \nu_2(\mathbb{P}U)$ is not contained in any hyperplane, thus the intersection must be empty. □

Lemma 7.20. *The map π_C is injective in $\nu_2(\mathbb{P}U)$.*

Proof. Let $x,\ y$ be elements both in $\nu_2(\mathbb{P}U)$. The map is given by

$$x \mapsto \overline{x, C} \cap \mathrm{Sym}^{2d} V,$$

Thus, the equation $\pi_C(x) = \pi_C(y)$ implies that $\overline{x, C} = \overline{y, C}$. This means that there exists $\lambda \in \mathbb{C}$ and $c \in C$ such that $x = \lambda y + c$. Therefore, $x - \lambda y = c \in C$ which is a contradiction since $x - \lambda y \in \Sigma_2(\nu_2(U))$. □

Lemma 7.21. *The projection $\pi_C : \mathrm{Sym}^2(\mathrm{Sym}^d V) \to \mathrm{Sym}^{2d} V$ restricted to the second secant variety of the Veronese variety $\Sigma_2(\nu_2(\mathrm{Sym}^d V))$ is injective.*

Proof. Consider the projection

$$\mathrm{Sym}^d V \times \mathrm{Sym}^d V \times \mathrm{Sym}^2(\mathrm{Sym}^d V) \to \mathrm{Sym}^2(\mathrm{Sym}^d V).$$

Let $\mathrm{Ab}_2(\nu_2(\mathrm{Sym}^d V)) = \{(\alpha, \beta, g) | \alpha^2 + \beta^2 = g\}$ be the abstract Veronese variety that under the projection is mapped to $\Sigma_2(\nu_2(\mathrm{Sym}^d V))$. Notice that the fiber of this projection on a point g is $\mathrm{O}(2)$ by Lemma 7.20.

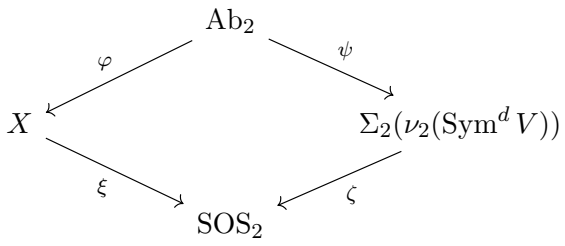
We may consider a similar projection

$$\mathrm{Sym}^d V \times \mathrm{Sym}^d V \times \mathrm{Sym}^{2d} V \to \mathrm{Sym}^{2d} V.$$

We may define $X = \{(\alpha, \beta, f) | \alpha^2 + \beta^2 = f\}$ in the same fashion as before. Under this projection we have that X is mapped to SOS_2 and the fiber on a point f is $\mathrm{SOS}_2(f) = \mathrm{O}(2)$ by Lemma 2.15.

Notice that the map $\mathrm{Sym}^2(\mathrm{Sym}^d V) \to \mathrm{Sym}^{2d} V$ that corresponds to the change of coordinates $w_1 = x_0^d, \ldots, w_N = x_n^d$ is injective when restricted to $\Sigma_2(\nu_2(\mathrm{Sym}^d V))$ and so is the induced linear map from Ab_2 to X.

Joining those maps into a diagram we obtain:



From the previous remarks, φ is an one-to-one map and the fibers of ψ and ξ are both equal to $\mathrm{O}(2)$. Since the diagram commutes, we also obtain that ζ is a one-to-one map. □

Theorem 2.17. *Let $N = \dim \mathrm{Sym}^d V = \binom{n+d}{d}$. The degrees of the varieties of squares and of sum of two squares in $\mathbb{P}(\mathrm{Sym}^{2d} V)$ are given by*

$$\deg(\mathrm{SOS}_1) = 2^{N-1}, \ \deg(\mathrm{SOS}_2) = \prod_{i=0}^{N-3} \frac{\binom{N+i}{N-2-i}}{\binom{2i+1}{i}}.$$

Proof. Since $\Sigma_k(\nu_2(\mathbb{P}U)) \cap C = \emptyset$ and $\pi_C|_{\Sigma_k(\nu_2(\mathbb{P}U))}$ is injective for $k = 1, 2$, it follows $\deg(\mathrm{SOS}_j) = \deg(\Sigma_j(\nu_2(\mathbb{P}U)))$. A classical result by Segre [46] states that for any $j \leq N$

$$\deg(\Sigma_j(\nu_2(\mathbb{P}U))) = \prod_{i=0}^{N-1-j} \frac{\binom{N+i}{N-j-i}}{\binom{2i+1}{i}}.$$ □

We notice that in the case of $n = 2$ and $d = 2$ Theorem 7.19 is sharp in the sense that for the 3-secant variety of $\nu_2(\mathbb{P}U)$ the intersection with C is non-empty. Indeed, one can find by computation that the intersection of $\Sigma_1(\nu_2(\mathbb{P}U))$ and $\Sigma_2(\nu_2(\mathbb{P}U))$ with C are empty. Thus, $\deg(\mathrm{SOS}_1) = 32$ and $\deg(\mathrm{SOS}_2) = 126$ as expected. However, for SOS_3 the intersection has codimension 3 in $\mathbb{P}^5 = \mathbb{P}U$. When the intersection is non-empty, the degree of $\Sigma_k(\nu_2(\mathbb{P}U))$ is still an upper bound for the degree of SOS_k.

Chapter 8

Conclusion and Perspectives

8.1 Problem 1

Conclusion. Under a variant of Fröberg's conjecture, we gave a complete description of the structure of the multiplication matrix T_{x_n} for generic critical point ideals. In particular, we gave an asymptotic formula for the number of dense columns q, a fundamental parameter in the change of ordering step of Gröbner bases. Moreover, we showed that this matrix can be computed solely from the DRL Gröbner basis, without any arithmetic operations. By proving a combinatorial result on the Hilbert series and combining this with the structure of the DRL staircase, we showed that this number q is equal to the largest coefficient of the Hilbert series. This allows for a finer complexity estimate for the Sparse-FGLM algorithm as well as for any other algorithms depending on the structure of T_{x_n}. Furthermore, we laid out the foundation for similar result for other classes of determinantal ideals, a thread which is explored in more detail in Chapter 5.

Perspectives. The next step would be to handle the case of mixed degrees. That is, investigate the change of ordering step for the critical points of a polynomial restricted to $\mathbf{V}(g_1, \ldots, g_m)$ where the degrees of $g_1, \ldots, g_m$ are allowed to differ. The computation of the reduced DRL Gröbner basis has been investigated in this setting [102]. In this paper, under some regularity assumptions, the author constructs the Hilbert series of generic determinantal sum ideals in this mixed degree setting using the Eagon-Northcott complex to obtain a graded free resolution from which the Hilbert series can be deduced. However, the numerator of the series is given as an alternating, combinatorial sum which is difficult to work with. In order to derive similar results as we obtained in Chapter 4, we would first need to provide a simplified formula of this Hilbert series. This would allow for a more refined complexity estimate for critical value computation when applied to polynomial optimisation.

8.2 Problem 2

Conclusion. In this thesis, we laid out the framework for studying the change of ordering from DRL to LEX Gröbner bases for generic determinantal ideals derived from structured matrices. For generic determianantal systems, assuming a variant of Fröberg's conjecture and a combinatorial property of the Hilbert series, we derive an approximation of a key parameter in the complexity of sparsity based FGLM-like algorithms such as Sparse-FGLM. For symmetric matrices in particular, for certain sizes of minors for which the Hilbert series of the resulting determinantal ideal is known, we give a finer approximation of this parameter and so a finer complexity analysis.

Perspectives. Firstly, Conjecture 2.6 could be lessened for generic determinantal ideals. One can apply the Cauchy-Binet formula, as in Chapter 4, to write the Hilbert series as a binomial sum. However, since the sum is over the set of minors of a given rank which, outside of the

maximal minor case studied in Chapter 4, it is still unclear how to prove unimodality. The study of the LEX Gröbner basis computation for determinantal ideals can also be extended to matrices with other structures. In particular, Hankel systems appear in many applications. We do not yet know if results such as Lemma 5.13 exist for other sizes of Hankel systems. Another structure of interest is that of moment matrices. The determinantal ideals derived from these matrices arise in the study of SOS decompositions as we saw in Chapter 7. Additionally, we recall Conjecture 5.14. Indeed, while the determinantal ideals defined from the minors of symmetric and triangular matrices are distinct they appear to have the same Hilbert series. Intuitively, this could be due to the correspondence between the minors in the symmetric case which are equal with multiple distinct indices and the minors in the triangular case which are zero due to the zeroes in the matrix. However, we are unsure how to prove this conjecture concretely. Furthermore, as with Problem 1, it could be interesting the tackle the case where the entries of the matrices we consider have varying degrees as this would leave to even more refined complexity estimates for the DRL to LEX framework.

8.3 Problem 3

Conclusion. In this thesis we have solved Problem 3 by designing algorithms that compute a superset of the asymptotic critical values for a polynomial map $f : X \to \mathbb{R}$, where X is a smooth algebraic set defined by a reduced, regular sequence $g_1, \ldots, g_m$ and $f, g_1, \ldots, g_m$ have degree d, with complexity singly exponential in n, the dimension of the ambient space. This brings this complexity in line with that of computing the critical values. Moreover, we control the size of the output to be essentially the Bézout bound of d^n so that the remaining root isolation and fibre emptiness test steps do not dominate. Furthermore, we extended our algorithms, under the same regularity assumptions, to compute the asymptotic critical values of polynomial mappings to $\mathbb{R}^p$ for $p > 1$ and obtain similar a complexity result and a bound on the size of the output.

Perspectives. An interesting area to investigate further is the study of *non-trivial* asymptotic critical values, denoted $N_\infty(\boldsymbol{f})$. These are the asymptotic critical values that do not lie along the critical locus, such values being called *trivial* asymptotic critical values. The algorithm of [60] manages to compute only $N_\infty(f)$ where f is polynomial. For the application of polynomial optimisation, we would like to extend this algorithm to be able to compute asymptotic critical values of polynomial mappings restricted to smooth algebraic sets. Moreover, we emphasise that a complexity analysis of the algorithm given in [60] is lacking and so we must ensure that such an extension be as, or more, efficient than the ones we propose in this thesis. The main advantage of computing only $N_\infty(\boldsymbol{f})$ is that it means there is no overlap in the computation of the critical values. This would result in a smaller output and could result in a speed-up by computing fewer values and also by having to perform fewer real root isolation.

8.4 Problem 4

Conclusion. We made progress towards Problem 4 by relating the variety of all possible k-SOS decompositions of a generic polynomial of SOS-rank k to the group of orthogonal matrices $O(k)$. Specifically, we show that these objects are have the same dimension and are in fact isomorphic when $k = 2$. Using this result, we proved a formula for the degree of the variety of all sums of two squares in any number of homogeneous variables. For $k > 2$, we conjectured that this isomorphism with $O(k)$ still holds and we gave some theoretical results and experimental findings to support this conjecture.

Perspectives. Our main goal would be to prove Conjecture 7.14 in its full generality. With Lemma 7.18 in tow, it would suffice to prove a certain transversality result. A standard result in

differential topology is that transversality is a generic property [45, Exercise 2.3.7]. In our setting, this means that if we take a variety $V \subset \mathbb{K}^n$ and a linear subspace $L \subset \mathbb{K}^n$, then for a generic point $x \in \mathbb{K}^n$, $L + x$ intersects V transversely. However, to prove Conjecture 7.14 we require taking x to be a generic point in V. While this appears to be true in our experimental testing, we were unable to find this result in the literature and we are unsure how to prove it ourselves.

Chapter 9

Bibliography

[1] C. Aholt, B. Sturmfels, and R. Thomas. A Hilbert Scheme in Computer Vision. *Canadian Journal of Mathematics*, 65(5):961–988, Oct. 2013.

[2] E. L. Allgower and K. Georg. *Numerical Continuation Methods: an Introduction*, volume 13. Springer Science & Business Media, 2012.

[3] J. Alman and V. V. Williams. A refined laser method and faster matrix multiplication. In *Proceedings of the Thirty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '21, pages 522–539, USA, 2021. Society for Industrial and Applied Mathematics.

[4] B. Bank, M. Giusti, J. Heintz, and G. M. Mbakop. Polar varieties and efficient real elimination. *Mathematische Zeitschrift*, 238(1):115–144, 2001.

[5] M. Bardet, J.-Ch. Faugère, and B. Salvy. On the complexity of Gröbner basis computation of semi-regular overdetermined algebraic equations. In *Proceedings of the International Conference on Polynomial System Solving*, pages 71–74, 2004.

[6] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in Real Algebraic Geometry (Algorithms and Computation in Mathematics)*. Springer-Verlag, Berlin, Heidelberg, 2006.

[7] D. A. Bayer. *The division algorithm and the Hilbert scheme*. Harvard University, 1982.

[8] E. Becker, T. Mora, M. G. Marinari, and C. Traverso. The Shape of the Shape Lemma. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation*, ISSAC '94, page 129–133, New York, NY, USA, 1994. Association for Computing Machinery.

[9] T. Becker and V. Weispfenning. *Gröbner Bases*. Springer New York, New York, NY, 1993.

[10] J. Berthomieu, A. Bostan, A. Ferguson, and M. Safey El Din. Gröbner bases and critical values: The asymptotic combinatorics of determinantal systems. *Journal of Algebra*, 602:154–180, 2022.

[11] J. Berthomieu, C. Eder, and M. Safey El Din. msolve: A Library for Solving Polynomial Systems. In *2021 International Symposium on Symbolic and Algebraic Computation*, 46th International Symposium on Symbolic and Algebraic Computation, Saint Petersburg, Russia, Jul. 2021.

[12] J. Berthomieu, V. Neiger, and M. Safey El Din. Faster Change of Order Algorithm for Gröbner Bases Under Shape and Stability Assumptions. In *Proceedings of the 2022 International Symposium on Symbolic and Algebraic Computation*, ISSAC '22, New York, NY, USA, 2022. Association for Computing Machinery.

[13] J. Bochnak, M. Coste, and M.-F. Roy. *Real Algebraic Geometry*, volume 36. Springer Science & Business Media, 2013.

[14] M. Brandt, J. Bruce, T. Brysiewicz, R. Krone, and E. Robeva. The Degree of $\mathrm{SO}(n, \mathbb{C})$. In *Combinatorial Algebraic Geometry*, pages 229–246. Springer, 2017.

[15] R. P. Brent, F. G. Gustavson, and D. Y. Yun. Fast solution of Toeplitz systems of equations and computation of Padé approximants. *Journal of Algorithms*, 1(3):259–295, 1980.

[16] W. Bruns and U. Vetter. *Determinantal Rings*, volume 1327. Springer, 2006.

[17] D. G. Cantor and E. Kaltofen. On fast multiplication of polynomials over arbitrary algebras. *Acta Informatica*, 28:693–701, 1991.

[18] A. Conca. Gröbner Bases of Ideals of Minors of a Symmetric Matrix. *Journal of Algebra*, 166:406–421, 1994.

[19] A. Conca. Symmetric ladders. *Nagoya Mathematical Journal*, 136:35–56, 1994.

[20] A. Conca. Straightening Law and Powers of Determinantal Ideals of Hankel Matrices. *Advances in Mathematics*, 138(2):263–292, 1998.

[21] A. Conca, E. De Negri, and V. Welker. A Gorenstein simplicial complex for symmetric minors. *Israel Journal of Mathematics*, 212:237–257, 2014.

[22] A. Conca and J. Herzog. On the Hilbert function of determinantal rings and their canonical module. *Proc. Amer. Math. Soc.*, 122(3):677–681, 1994.

[23] D. A. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry.* Graduate Texts in Mathematics. Springer New York, 2006.

[24] D. A. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 4/e (Undergraduate Texts in Mathematics).* Springer-Verlag, Berlin, Heidelberg, 2015.

[25] Y. De Castro, F. Gamboa, D. Henrion, R. Hess, and J.-B. Lasserre. Approximate optimal designs for multivariate polynomial regression. *Ann. Statist.*, 47(1):127–155, 2019.

[26] I. V. Dolgachev. *Classical Algebraic Geometry: A Modern View.* Cambridge University Press, 2012.

[27] D. Eisenbud. *Commutative Algebra: with a view toward algebraic geometry*, volume 150. Springer Science & Business Media, 2013.

[28] J.-C. Faugère. A new efficient algorithm for computing Gröbner bases (F4). *Journal of Pure and Applied Algebra*, 139(1-3):61–88, 1999.

[29] J.-C. Faugère. A new efficient algorithm for computing Gröbner bases without reduction to zero (F_5). In *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation*, pages 75–83. ACM, New York, 2002.

[30] J.-C. Faugère, P. Gaudry, L. Huot, and G. Renault. Sub-Cubic change of ordering for Gröbner basis: A probabilistic approach. In *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, ISSAC '14, pages 170–177, New York, NY, USA, 2014. Association for Computing Machinery.

[31] J.-C. Faugère, P. Gianni, D. Lazard, and T. Mora. Efficient computation of zero-dimensional Gröbner bases by change of ordering. *J. Symbolic Comput.*, 16(4):329–344, 1993.

[32] J.-C. Faugère and C. Mou. Sparse FGLM algorithms. *J. Symbolic Comput.*, 80(part 3):538–569, 2017.

[33] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer. Critical points and Gröbner bases: the unmixed case. In *ISSAC 2012—Proceedings of the 37th International Symposium on Symbolic and Algebraic Computation*, pages 162–169. ACM, New York, 2012.

[34] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer. On the complexity of the generalized MinRank problem. *J. Symbolic Comput.*, 55:30–58, 2013.

[35] A. Ferguson and H. P. Le. Finer Complexity Estimates for the Change of Ordering of Gröbner Bases for Generic Symmetric Determinantal Ideals. In *Proceedings of the 2022 International Symposium on Symbolic and Algebraic Computation*, ISSAC '22, New York, NY, USA, 2022. Association for Computing Machinery.

[36] R. Fröberg. An inequality for Hilbert series of graded algebras. *Math. Scand.*, 56(2):117–144, 1985.

[37] R. Fröberg, S. Lundqvist, A. Oneto, and B. Shapiro. Algebraic stories from one and from the other pockets. *Arnold Mathematical Journal*, 4(2):137–160, 2018.

[38] R. Fröberg, G. Ottaviani, and B. Shapiro. On the Waring problem for polynomial rings. *Proc. Natl. Acad. Sci. USA*, 109(15):5600–5602, 2012.

[39] I. Gessel and G. Viennot. Binomial determinants, paths, and hook length formulae. *Adv. in Math.*, 58(3):300–321, 1985.

[40] B. Ghaddar, J. Marecek, and M. Mevissen. Optimal power flow as a polynomial optimization problem. *IEEE Transactions on Power Systems*, 31(1):539–546, 2015.

[41] P. Gianni and T. Mora. Algebraic solution of systems of polynomial equations using Gröbner bases. In *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes (Menorca, 1987)*, volume 356 of *Lecture Notes in Comput. Sci.*, pages 247–257. Springer, Berlin, 1989.

[42] M. Giusti, G. Lecerf, and B. Salvy. A Gröbner free alternative for polynomial system solving. *Journal of Complexity*, 17(1):154–211, 2001.

[43] M. L. Green. Koszul Cohomology and Geometry. In *Lectures on Riemann Surfaces*, pages 177–200. World Scientific, 1989.

[44] A. Greuet and M. Safey El Din. Probabilistic Algorithm for Polynomial Optimization over a Real Algebraic Set. *SIAM Journal on Optimization*, 24(3):1313–1343, Aug. 2014.

[45] V. Guillemin and A. Pollack. *Differential topology*, volume 370. American Mathematical Soc., 2010.

[46] J. Harris and L. W. Tu. On symmetric and skew-symmetric determinantal varieties. *Topology*, 23(1):71–84, 1984.

[47] R. Hartshorne. *Algebraic Geometry*. Graduate Texts in Mathematics. Springer, 1977.

[48] J. Heintz. Definability and fast quantifier elimination in algebraically closed fields. *Theoretical Computer Science*, 24(3):239–277, 1983.

[49] J. Heintz, T. Krick, S. Puddu, J. Sabia, and A. Waissbein. Deformation techniques for efficient polynomial equation solving. *Journal of Complexity*, 16(1):70–109, 2000.

[50] D. Henrion and A. Garulli, editors. *Positive Polynomials in Control*, volume 312 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 2005.

[51] D. Henrion and J.-B. Lasserre. GloptiPoly: Global Optimization over Polynomials with Matlab and SeDuMi. *ACM Trans. Math. Softw.*, 29(2):165–194, Jun. 2003.

[52] D. Henrion, S. Naldi, and M. Safey El Din. Real Root Finding for Rank Defects in Linear Hankel Matrices. In *Proceedings of the 2015 ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC '15, page 221–228, New York, NY, USA, 2015. Association for Computing Machinery.

[53] D. Henrion, S. Naldi, and M. Safey El Din. Exact Algorithms for Linear Matrix Inequalities. *SIAM Journal on Optimization*, 26(4):2512–2539, 2016.

[54] D. Henrion, S. Naldi, and M. Safey El Din. Real root finding for determinants of linear matrices. *Journal of Symbolic Computation*, 74:205–238, 2016.

[55] D. Henrion, M. Šebek, and V. Kučera. Positive polynomials and robust stabilization with fixed-order controllers. *IEEE Trans. Automat. Control*, 48(7):1178–1186, 2003.

[56] D. Hilbert. Über die darstellung definiter formen als summe von formenquadraten. *Mathematische Annalen*, 32(3):342–350, 1888.

[57] H. Hong and M. Safey El Din. Variant quantifier elimination. *J. Symbolic Comput.*, 47(7):883–901, 2012.

[58] Z. Jelonek and K. Kurdyka. Quantitative Generalized Bertini-Sard Theorem for Smooth Affine Varieties. *Discrete and Computational Geometry, v.34, 659-678 (2005)*, 34, Nov. 2005.

[59] Z. Jelonek and K. Kurdyka. Reaching generalized critical values of a polynomial. *Mathematische Zeitschrift*, 276(1-2):557–570, 2014.

[60] Z. Jelonek and M. Tibăr. Detecting asymptotic non-regular values by polar curves. *International Mathematics Research Notices*, 2017(3):809–829, 2017.

[61] C. Josz and D. K. Molzahn. Lasserre hierarchy for large scale polynomial optimization in real and complex variables. *SIAM Journal on Optimization*, 28(2):1017–1048, 2018.

[62] K. Kurdyka, P. Orro, and S. Simon. Semialgebraic Sard Theorem for Generalized Critical Values. *J. Differential Geom.*, 56(1):67–92, Sep. 2000.

[63] R. E. Kutz. Cohen-Macaulay Rings and Ideal Theory in Rings of Invariants of Algebraic Groups. *Trans. Am. Math. Soc.*, 194:115–129, 1974.

[64] G. Labahn, M. Safey El Din, É. Schost, and T. X. Vu. Homotopy techniques for solving sparse column support determinantal polynomial systems. *Journal of Complexity*, 2021.

[65] S. Lang. *Algebra*, volume 211. Springer Science & Business Media, 2012.

[66] J.-B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. on Optimization*, 11(3):796–817, Mar. 2000.

[67] J.-B. Lasserre. *Moments, Positive Polynomials and Their Applications*. Imperial College Press, Oct. 2009.

[68] R. Lazarsfeld. *Positivity in Algebraic Geometry I: Classical Setting: Line Bundles and Linear Series*, volume 48. Springer, 2017.

[69] H. P. Le and M. Safey El Din. Faster One Block Quantifier Elimination for Regular Polynomial Systems of Equations. In *Proceedings of the 2021 on International Symposium on Symbolic and Algebraic Computation*, ISSAC '21, page 265–272, New York, NY, USA, 2021. Association for Computing Machinery.

[70] H. P. Le and M. Safey El Din. Solving parametric systems of polynomial equations over the reals through Hermite matrices. *Journal of Symbolic Computation*, 112:25–61, 2022.

[71] T.-Y. Li. Numerical solution of multivariate polynomial systems by homotopy continuation methods. *Acta Numerica*, 6:399–436, 1997.

[72] J. Lofberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *2004 IEEE international conference on robotics and automation (IEEE Cat. No. 04CH37508)*, pages 284–289. IEEE, 2004.

[73] S. Lundqvist, A. Oneto, B. Reznick, and B. Shapiro. On generic and maximal k-ranks of binary forms. *Journal of Pure and Applied Algebra*, 223(5):2062–2079, 2019.

[74] V. Magron and M. Safey El Din. On exact Polya and Putinar's representations. In *Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*, pages 279–286, 2018.

[75] V. Magron and M. Safey El Din. RealCertify: a Maple package for certifying non-negativity. *ACM Communications in Computer Algebra*, 52(2):34–37, 2018.

[76] Maplesoft, a division of Waterloo Maple Inc.. Maple, 2021.

[77] J. Migliore and U. Nagel. Survey Article: A tour of the weak and strong Lefschetz properties. *Journal of Commutative Algebra*, 5(3):329–358, 2013.

[78] J. W. Milnor and J. D. Stasheff. Characteristic classes. *Annals of Mathematics Studies*, 76, 1975.

[79] G. Moreno-Socías. Degrevlex Gröbner bases of generic complete intersections. *J. Pure Appl. Algebra*, 180(3):263–283, 2003.

[80] T. S. Motzkin. The arithmetic-geometric inequality. *Inequalities (Proc. Sympos. Wright-Patterson Air Force Base, Ohio, 1965)*, pages 205–224, 1967.

[81] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39:117–129, 1987.

[82] S. Naldi. Solving rank-constrained semidefinite programs in exact arithmetic. *Journal of Symbolic Computation*, 85:206–223, 2018. 41th International Symposium on Symbolic and Algebraic Computation (ISSAC'16).

[83] V. Neiger and É. Schost. Computing syzygies in finite dimension using fast linear algebra. *Journal of Complexity*, 60:101502, 2020.

[84] J. Nie and K. Ranestad. Algebraic degree of polynomial optimization. *SIAM J. Optim.*, 20(1):485–502, 2009.

[85] A. Papachristodoulou, J. Anderson, G. Valmorbida, S. Prajna, P. Seiler, and P. A. Parrilo. *SOSTOOLS: Sum of squares optimization toolbox for MATLAB*. `http://arxiv.org/abs/1310.4716`, 2013. Available from `http://www.eng.ox.ac.uk/control/sostools`, `http://www.cds.caltech.edu/sostools` and `http://www.mit.edu/~parrilo/sostools`.

[86] K. Pardue. Generic sequences of polynomials. *J. Algebra*, 324(4):579–590, 2010.

[87] H. Peyrl and P. Parrilo. Computing sum of squares decompositions with rational coefficients. *Theoretical Computer Science*, 409:269–281, Dec. 2008.

[88] T. Probst, D. P. Paudel, A. Chhatkuli, and L. Van Gool. Convex Relaxations for Consensus and Non-Minimal Problems in 3D Vision. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 10233–10242, 2019.

[89] P. J. Rabier. Ehresmann Fibrations and Palais-Smale Conditions for Morphisms of Finsler Manifolds. *Annals of Mathematics*, 146(3):647–691, 1997.

[90] C. Rocchini. Twisted cubic image. *Wikipedia, the free encyclopedia*, 2007.

[91] F. Rouillier and P. Zimmermann. Efficient isolation of polynomial's real roots. *Journal of Computational and Applied Mathematics*, 162(1):33–50, 2004.

[92] M. Safey El Din. Testing sign conditions on a multivariate polynomial and applications. *Mathematics in Computer Science*, 1(1):177–207, 2007.

[93] M. Safey El Din. Computing the global optimum of a multivariate polynomial over the reals. In *Proceedings of the 21st International Symposium on Symbolic and Algebraic Computation*, pages 71–78, 2008.

[94] M. Safey El Din. Real Alebraic Geometry library, RAGlib (version 3.4), 2017.

[95] M. Safey El Din and É. Schost. Polar varieties and computation of one point in each connected component of a smooth algebraic set. In *Proceedings of the 2003 International Symposium on Symbolic and Algebraic Computation*, pages 224–231. ACM, New York, 2003.

[96] M. Safey El Din and É. Schost. A nearly optimal algorithm for deciding connectivity queries in smooth and bounded real algebraic sets. *J. ACM*, 63(6):48:1–48:37, 2017.

[97] M. Safey El Din and É. Schost. Bit complexity for multi-homogeneous polynomial system solving—application to polynomial minimization. *Journal of Symbolic Computation*, 87:176–206, 2018.

[98] É. Schost. Computing parametric geometric resolutions. *Applicable Algebra in Engineering, Communication and Computing*, 13(5):349–393, 2003.

[99] M. Schweighofer. Global optimization of polynomials using gradient tentacles and sums of squares. *SIAM Journal on Optimization*, 17(3):920–942, 2006.

[100] J.-P. Serre. Représentations linéaires et espaces homogenes kählériens des groupes de Lie compacts. *Séminaire Bourbaki*, 2:447–454, 1954.

[101] I. R. Shafarevich and M. Reid. *Basic Algebraic Geometry 1: Varieties in Projective Space*. SpringerLink : Bücher. Springer Berlin Heidelberg, 2013.

[102] P.-J. Spaenlehauer. On the Complexity of Computing Critical Points with Gröbner Bases. *SIAM Journal on Optimization*, 24:1382–1401, Jul. 2014.

[103] Z. Star. An asymptotic formula in the theory of compositions. *Aequationes Math.*, 13(3):279–284, 1975.

[104] B. Sturmfels. Gröbner bases and Stanley decompositions of determinantal rings. *Mathematische Zeitschrift*, 205:137–144, 1990.

[105] J. J. Sylvester. On a remarkable discovery in the theory of canonical forms and of hyperdeterminants. *Philosophical Magazine*, I:265–283, 1851.

[106] P. Trutman, M. Safey El Din, D. Henrion, and T. Pajdla. Globally optimal solution to inverse kinematics of 7DOF serial manipulator. *IEEE Robotics and Automation Letters*, 7(3):6012–6019, 2022.

[107] J. von zur Gathen and J. Gerhard. *Modern Computer Algebra*. Cambridge University Press, third edition, 2013.

[108] H. Waki, S. Kim, M. Kojima, M. Muramatsu, and H. Sugimoto. Algorithm 883: SparsePOP—A Sparse Semidefinite Programming Relaxation of Polynomial Optimization Problems. *ACM Trans. Math. Softw.*, 35(2), Jul. 2008.

[109] J. Wang, V. Magron, and J.-B. Lasserre. TSSOS: A Moment-SOS hierarchy that exploits term sparsity. *SIAM Journal on Optimization*, 31(1):30–58, 2021.

[110] J. Weyman. *Cohomology of Vector Bundles and Syzygies*. Cambridge Tracts in Mathematics. Cambridge University Press, 2003.

[111] D. Wiedemann. Solving sparse linear equations over finite fields. *IEEE transactions on information theory*, 32(1):54–62, 1986.