On the expected number of zeros of polynomials and the real tau-conjecture

A dissertation submitted towards the degree Doctor of Natural Sciences of the Faculty of Mathematics and Computer Science of Saarland University

by Charilaos Zisopoulos

Saarbrücken / 2023

Day of Colloquium: 19. December 2023
Dean of the Faculty: Prof. Dr. Jürgen Steimle

Chair of the Committee: Prof. Dr. Karl Bringmann
Reporters
First reviewer: Prof. Dr. Markus Bläser
Second reviewer: Prof. Dr. Frank-Olaf Schreyer
Academic Assistant: Dr. Evangelos Kipouridis

Abstract

Abstract We study the expected number of real zeros of random sparse polynomials, motivated by the real τ-conjecture. Our main results focus on random k-sparse polynomials following the standard normal distribution, i.e., their coefficients are i.i.d. random variables following this distribution. We show a $\Theta(\sqrt{k})$ upper bound for the expected number of real zeros for such polynomials. In addition, we show that the real zeros of such polynomials concentrate in a small neighborhood around $|x| = 1$. Furthermore, we show a matching $\Theta(\sqrt{k})$ lower bound for the same quantity for a specific family of such k-sparse polynomials, thus proving that the upper bound is asymptotically optimal. We generalize the techniques used to random k-sparse polynomials following absolutely continuous distributions. In particular, we show that it suffices to study the expected number of real zeros in the interval $(-1, 1)$. If these distributions are in addition symmetric about 0, we show that it suffices to consider the interval $(0, 1)$. Moreover, we show that for random k-sparse polynomials following the Rademacher distribution, previous results can be adapted to the sparse case, obtaining an upper bound of $\Theta(\sqrt{k})$. We survey previously known results on the number of real zeros of fixed and random polynomials, while also presenting a detailed analysis of results due to Descartes and Laguerre.

Zusammenfassung

Zusammenfassung Wir untersuchen die erwartete Anzahl reellen Nullstellen von dünnbesetzten Polynomen, motiviert durch die reelle τ-Vermutung. Unsere Hauptergebnisse konzentrieren sich auf zufällige k-dünnbesetzte standardnormalverteilte Polynome, d. h., ihre Koeffizienten sind u.i.v. standardnormalverteilte Zufallsvariablen. Wir zeigen eine $\Theta(\sqrt{k})$ obere Schranke für die erwartete Anzahl reellen Nullstellen für solche Polynome. Außerdem zeigen wir, dass sich die reelle Nullstellen solcher Polynome in einer kleinen Umgebung um $|x| = 1$ konzentrieren. Darüber hinaus zeigen wir für eine bestimmte Familie solcher k-dünnbesetzten Polynome eine entsprechende $\Theta(\sqrt{k})$ untere Schranke und beweisen damit, dass die obere Schranke asymptotisch optimal ist. Wir verallgemeinern diese verwendeten Techniken auf absolut stetig verteilte Polynome. Insbesondere zeigen wir, dass es ausreicht, die erwartete Anzahl reeller Nullstellen im Intervall $(-1, 1)$ zu untersuchen. Wenn diese Verteilungen zusätzlich symmetrisch um 0 sind, zeigen wir, dass es ausreicht, das Intervall $(0, 1)$ zu berücksichtigen. Darüber hinaus zeigen wir anhand früherer Ergebnisse eine obere Schranke von $\Theta(\sqrt{k})$ für zufällige k-dünnbesetzte Rademacher-verteilte Polynome. Wir untersuchen bekannte Ergebnisse zur Anzahl reeller Nullstellen fester und zufälliger Polynome und präsentieren eine Analyse der Ergebnisse von Descartes und Laguerre.

Dedicated to my daughter Elpida. You give me hope that my accomplishments will allow yours to be even greater.

Acknowledgments

I would like to thank my advisor Prof. Dr. Markus Bläser for his continuing support throughout my doctoral studies. He has taught me invaluable lessons not only in technical matters relating to our field of research, but also on how to be a supportive educator and a valuable member of a research community that fosters collaboration. He has inspired me not only to apply these lessons for the rest of my life, but to make sure they are transferred to the next generations. He has shown tremendous understanding and patience throughout the many difficulties I have faced completing this work and I sincerely believe that he was indispensable in making it possible for me to achieve this goal.

I am grateful for the sustained efforts, insightful discussions and constructive feedback of my co-authors, Dr. Gorav Jindal, Dr. Anurag Pandey and Himanshu Shukla. The results contained in this work are the fruit of our wonderful collaboration and have been made possible through our combined contributions. In particular, I would like to thank Dr. Jindal and Dr. Pandey for introducing me to this line of research that has not only been the focus of my research but also my passion for the past few years. I also deeply appreciate the mathematical rigor and unique insights Himanshu Shukla brought along when he joined our team. I would like to thank all of them for their perseverance and hard work that made our results possible.

I would like to extend my gratitude to the many group members and members of the research community I interacted throughout the years. I am proud to be part both of our research group as well as the extended research community. In particular, I would like to thank Prof. Dr. Peter Bürgisser for his continued efforts that have renewed interest in the particular research topic this work is focused on, as well as for his helpful comments and insights. I also deeply appreciate the efforts of both faculty and administrative members of Saarland University that helped me tremendously throughout my studies. I would also like to thank Professor Kostas Tsichlas who during my bachelor studied was my mentor and introduced me to Computational Complexity and in addition helped me with finding a position abroad that was relevant to my research interests.

This work would not have been possible without the continuous support and love of my wife Maria. You have been with me every step of the way and if it was your kind words that kept me going through the worst of times. I thank you greatly for sharing these burdens with me. You are a wonderful wife to me and mother to my daughter and I cannot imagine my life without you. There is not enough space in this work or words to describe how much you mean for me, but I hope my actions everyday express my feelings better than anything I could ever write down.

I deeply value the care, love and support of my parents throughout the years. Without their many sacrifices and their support it would not be possible for me to even be able to start this work, let alone complete it. I hope that my achievements and my conduct make you proud for the person I have become. Thank you for believing in me every step of the way.

I'm beyond grateful for the support and encouragement of many family members

and friends throughout the years. In particular I would like to thank Dr. Evangelos Kipouridis for his friendship and for his unique perspective in life that has allowed me to view the world through new eyes. In addition, I would like to thank Dr. Pandey for his friendship, in particular the countless hours we spent discussing many topics from research to philosophy and everything in between, as well as for his support when it was needed.

Last but not least, I would like to honor the memory of those that are unfortunately not here to see this work to completion. I want to thank Athanasios Mpasogiannis for being my first mentor in Mathematics that taught me many things, including how to approach mathematical problems in a formal way. I also would like to thank the Honorable Professor Moses Elisaf not only for diagnosing me and assisting me with dealing with my chronic illness, but also for offering profound life advice. Finally, I would like to thank my beloved grandparents Kalliopi and Charilaos for their unconditional love and being a large part of my upbringing that helped me become who I am today.

Contents

CHAPTER 1

Introduction

In this work we study the number of real zeros a polynomial may have. We focus on univariate random polynomials, where by random we mean that the coefficients of a polynomial follow a distribution. Since the polynomials are random, so is the number of roots, we thus more precisely study the expected number of roots.

The question of how many roots a polynomial has, either in total or of a particular kind, such as integer roots or real roots, has been a question that has in one form or another been posed throughout human history. In particular, the study of the solutions of polynomials has been of interest for most of recorded history. The archaeological record suggests that in practice such interest arose naturally from everyday activities of the first civilized societies such as trade and would resemble the simple linear equations in one unknown that every student encounters in primary education. For example, a Sumerian trader that knew the price of a bundle (or as they were called, gur) of barley, say 2 coins of silver (called shekels) and also knew that he needs 200 silver coins to sustain himself and carry out his next business venture, could solve the simple equation $2x = 200$ to figure out that he must purchase at least 100 bundles of wheat to reach his goal. Of course, the method he would proceed to perform this calculation would look much more foreign than the equation above, but in principle, the result would be the same. Interestingly, these ancient methods resemble more the modern notion of computation rather than mathematical formulas, although they lacked the abstraction that is central in modern Computer Science. Rather, those interested in these methods such as traders or state officials learned through repetition, trial and error, not unsimilar to how supervised learning works in the context of Machine Learning.

As these societies thrived and began engaging in ever more complex activities, such as the building of large monuments and the necessary advances in mathematics and architecture, so did naturally the related problems arising from practice. General methods for solving such problems would be the progenitor of modern theoretical approaches, albeit lacking the level of formalization we are familiar with nowadays. Archaeological evidence for the above include the Moscow Mathematical Papyrus [71] originating in Egypt around 1850 BC, containing simple problems that we would classify as algebraic equations today, as well as in tablets from the Babylonian period bearing close resemblance to the well-known formula to solve quadratic equations or even a list of integer solutions for the equation $x^2 + y^2 = z^2$, i.e. Pythagorean triples, as witnessed for example in the so-called Plimpton 322 tablet [74] dating from 1800 BC. Of particular interest is that the latest problem is a first attempt to solve the problem of counting roots of a polynomial, and in fact the rather difficult problem to tackle of counting integer roots of a multivariate polynomial. As we have the benefit of knowing today, the above equation has an infinite number of solutions, but this was definitely out of the possibilities of mathematics at the time, with most numerical systems only having a finite amount of numbers.

Thankfully nowadays we are equipped with a much more rigorous theory that allows us to study problems such as polynomial root counting with a high level of formalization. In Chapter 2 we introduce the necessary notions, theory and notation that form the foundation upon both our work as well as similar previous results build upon. We start Section 2.1 by introducing the basic notions regarding polynomials, including two different yet complementary ways of approaching polynomials: as objects of symbolic computation, i.e., formal polynomials, but also as polynomial functions. Both notions are useful and their interplay is at the core of this work, since we use properties of a polynomial as a function, namely its number of roots, to derive statements about its computational complexity, a (symbolic) computation notion, in a way that is made precise later in this work. This groundwork allows us to state bounds on the number of roots of a polynomial with respect to its degree, an endeavour that was one of the first problems to be stated in a somewhat formal way that would seem familiar to us today in the late 17th century, although proper formal proofs of such statements would only be given two centuries later, which we present, using modern notation.

In Section 2.1.2 we introduce the notion of *sparsity*, which is simply the number of terms of a polynomial that have a non-zero coefficient. Sparsity as a notion is attractive to computer scientists and in particular complexity theorists, since it captures the complexity of representing the polynomial much more accurately than the degree. For example, a polynomial of very high degree may only have very few terms with nonzero coefficients, that is, very small sparsity. We call such polynomials sparse and they are a central object of study in our work. Typically we assume that all elements from the underlying field to which the values of the coefficients belong have the same complexity in our work, but in other fields of study each element may be assigned a different value of a measure of complexity, such as for example bit-length. Then, combined with sparsity, we may perform a much more fine-tuned analysis of the related measures, a notion useful for fields of study such as numerical analysis or computer algebra.

In addition, it turns out that sparsity plays a significant role in bounding the number of *real* roots of a polynomial. Over a closed field, such as the complex numbers, the Fundamental Theorem of Algebra guarantees that a polynomial has as many roots as its degree. However, it leaves open the question of how many of these roots are real numbers. Descartes' rule of signs, first stated in 1637 by Descartes [25], bounds the number of positive real roots of a real polynomial by the number of sign changes of its coefficient sequence. We provide a formal statement in Theorem 2.17 together with a formal proof due to Gauß. Note that the number of sign changes of the coefficient sequence is upper bounded by its sparsity. In particular, often the coefficients are not known beforehand, or at all, for example if we consider symbolic computation as we often do, therefore the sparsity of the polynomial is the best upper bound on the number of real roots we can offer in such cases.

Descartes' rule remains important even today and it is often used to bound the number of real roots of a univariate real polynomial. In particular, it can be shown that certain polynomials match the bound, therefore no better bound can be achieved in this context. In our work and relevant works in the field, it serves as a baseline since any bound obtained in special cases should at least be an improvement upon the general one given by Descartes' rule. Furthermore, it can serve as a building block for more complex bounds. For a simple example, consider the product of m polynomials of sparsity at most

k. The resulting polynomial will have sparsity at most k^m and by applying Descartes' rule directly that would be the bound on the number of real roots. However, by first applying the bound on each factor and considering that the roots of the product of polynomials is the union of the roots of its factors, we derive a much improved bound of mk.

Due to its importance, we examine Descartes' rule in detail and state many variants of it. Theorem 2.21 is the most complete versions for real roots, although often simpler versions of the rule are used. We also provide an in-depth analysis that demonstrates precisely why Descartes' rule holds and goes beyond simple proofs of the statement, by examining the fundamental mechanics that it depends on. Furthermore, we provide a variant of Descartes' rule that uses only the sparsity of a polynomial in Corollary 2.23. We also mention subsequent similar work and generalizations of the rule, such as for example the rule due to Laguerre given in Theorem 2.28.

In Section 2.2 we introduce the necessary background for another cornerstone of our work, the notion of random polynomials. This notion allows us to go beyond the worst-case bounds given by Descartes' and similar rules and investigate the number of real zeros of classes of polynomials by considering the expected number of real roots. Quite often this expectation is much lower than the worst-case bound, since a single polynomial with a high number of real roots suffices to restrict the bounds possible. This also reveals much more information for the roots of the polynomials in the class than simply considering the worst case. For example, by a simple application of Markov's inequality we know that at most half the polynomials in a class may have more than twice the expected number of real roots. In Section 2.2.1 we introduce the fundamental notions from probability theory necessary to provide a solid foundation for our work. We then define random polynomials and the expected number of zeros of a random polynomial formally in Section 2.2.2.

Having established these fundamental notions for random polynomials, we proceed in Section 2.2.3 to illustrate the most important developments in this area of research, which occurred during the mid 20th century. These results can be viewed as a natural continuation of the work of Laguerre in the end of the 19th century which provided the best known results for bounding the number of real roots of a univariate polynomial. The first three decades of the century very few results were produced in the area, with most of them attempting to utilize the already known results to other domains of mathematics, with a focus in number theory. To the best of our knowledge, the first work to consider the expected number of random polynomials is due to Bloch and Pòlya [10]. They consider random polynomials of degree d which have coefficients equal to either $1, -1$ or 0 with equal probability and independently of one another and show that the expected number of real zeros in this case is $O(\sqrt{d})$. While their focus was for dense polynomials, their arguments easily adapt to the sparse case. In fact, in this work we showcase a version that allows us to say that k-sparse random polynomials following the same distribution have $O(\sqrt{k})$ real roots in expectation. It is also easy to show that the same arguments hold if we instead consider the Rademacher distribution, that is, each coefficient is either 1 or -1 with equal probability.

Less than a decade later, Littlewood and Offord in [58] and [59] continued this line of work. They consider random polynomials whose coefficients are independent and identically distributed and all follow the same distribution, considering three separate

cases: the standard normal distribution, the uniform continuous distribution and the aforementioned Rademacher distribution. In all three cases, they show that the expected number of real roots is $O(\log^2 d)$. The next noteworthy contribution was due to Kac in 1943 [45]. It is noteworthy not only for the results produced, but most notably for the techniques involved. In particular, Kac developed a method that expresses the number of real roots of a polynomial by a simple to state integral. We state this result in Theorem 2.40. We note that this result generalizes to differentiable functions, as well as the multivariate case, thus it is applicable to a wide range of functions used both in practice and theory. Even more pertinent to our work, Kac extends the result to random polynomials, a result that we state in Theorem 2.41, thus enabling us to express the expected number of roots of a random polynomial as an integral.

Kac used this method to improve the results of Littlewood and Offord for the two cases of continuous distributions they investigated, the standard normal distribution and the uniform continuous distribution. In both cases he shows that the expected number of real roots is $\Theta(\log d)$, where d is the degree of the random polynomial. On the other hand, he notes that dealing with discrete distributions is much more difficult, which seems to be confirmed by the time required for the next improvement along the line, despite significant efforts. It was only in 1956 that Erdős and Offord [29] proved the same bound of $\Theta(\log d)$ for the Rademacher distribution. The proof is highly complex with a large number of parameters and does not generalize to other discrete distributions. In subsequent decades, more results were achieved that generalized these results in various ways, by extending them to related families of distributions, proving results about other moments of the distribution of real zeros such as its variance, or by closely examining the previous proofs and improving the constants involved. We highlight the work of Edelman and Kostlan [28], who in the mid 1990s re-examined the results of Kac for the standard normal distribution, giving a different formulation of the integral involved which is fundamental to our work.

Having briefly reviewed the history of this research area by highlighting the most notable results, it would be remiss not to mention the modern applications of the theory as well. Apart from its importance as a purely mathematical question, root counting has many practical applications as well. To begin with, root counting serves as a natural way to find a stopping condition for the closely related problem of root finding. Root finding is the problem of finding one or more roots of a function, in our case, a polynomial. Quite often, it may not be possible to determine the root in all cases with finite precision, for example if we are searching for real roots. Then, we solve the related problem of root isolation, where we wish to output a small enough interval for each root that contains it. In both cases, knowing the total number of roots serves as a stopping criterion. For example, we may be able to determine by simple root counting methods, such as the Descartes' rule, that in fact no roots exist. Or if we wish to find multiple roots, we may use a root counting method to provide the maximum number of required iterations. Even simply knowing the expected number of roots can be beneficial, since it allows algorithm designers to make an informed decision about the number of iterations a randomized algorithm should perform.

Of course, polynomials are widely used in a diverse array of real-world applications, from finance to engineering and from natural to medical sciences. For example, linear system solving, perhaps one of the most common problems encountered in real-world

applications, is a special case of polynomial system solving. Many functions that either model the parameters of a problem or describe a solution are often polynomial functions, such as objective functions in optimization. Even when the function themselves are not polynomial, they are typically continuous and thus can be approximated arbitrarily close by polynomials in a closed interval by the Weierstrass approximation theorem, which also can be generalized for the multivariate case. One such example is designing and analyzing the trajectory of vehicles. These can be fixed track vehicles, such as trains or even roller coasters, or vehicles with many degrees of freedom of movement, such as spacecraft. Indeed, consider the simplified example where we are given the trajectory of a spacecraft as a univariate polynomial that measures the distance $d(t)$ from an object of interest, such as a planet, in terms of the time since launch t. Often mission goals require a specific action at a given distance from the object of interest, such as releasing satellites that contain science instrument or performing a maneuver to alter the trajectory. If the desired height is say, h, solving $d(t) - h = 0$ can identify a point in time for performing this mission. Similarly, computing the number of roots of this polynomial gives us the number of possible attempts or "mission windows" for performing the necessary action, since additional restrictions other than the distance may apply. In a more realistic setting, the same problems must be solved for a system of polynomials, since multiple conditions must be satisfied such as velocity, orientation, relative position and so on.

Another practical application comes from medicine, where simple univariate polynomials can be used to model the administering of medication to a patient. Typically there are polynomials that are defined for $t \geq 0$, where $t = 0$ is the time of the drug administration. These polynomials are usually defined empirically via clinical trials and similar experiments. For example, consider the following polynomial $f(t) = -2t^3 + 4t + 8$. The value $f(t)$ for $0 \leq t \leq 2$ represents the amount of drug active in the body of the patient at the given time. At $t = 2$, that is after 2 hours, the effect of the drug is eliminated completely, thus after this point the polynomial does not model the real world scenario anymore. By derivating it is easy to see that the maximum amount of drug in the body is achieved at $t = \sqrt{\frac{2}{3}}$, while by integration we can see that the total amount of drug active in the body throughout the two hour period is 16 units, for example milligrams (mg). Similar to the previous case example, root counting can be used to identify windows of opportunity, where the drug concentration is sufficient for a medical operation to take place. Furthermore, more complicated examples are possible either in the univariate or multivariate case, with the multivariate case allowing for the modeling of further parameters.

We conclude this diversion into real world applications by mentioning another practical use scenario from electrical engineering. We have already mentioned the work of Kac and noted that it can be extended for differentiable functions. We may then use that method to count the number of zeros in various generalizations of polynomials, such as trigonometric polynomials. Recall that a trigonometric polynomial is a function of the following form

$$f(x) = a_0 + \sum_{k=0}^{n-1} a_k \cos kx + \sum_{k=0}^{n-1} b_k \sin kx \,,$$

where in our case we will assume the coefficients a_i and b_i are real numbers, but in

general can also be complex. Such trigonometric polynomials can be used to describe electrical signals in systems, especially ones that are mixtures of simpler signals, which are modeled by the individual sine and cosine terms. Roots of trigonometric polynomials are of interest to engineers since they represent points where the amplitude of the signal is zero, for example the voltage in an alternating current system. At these points, it is possible to perform alterations to the signal without generating unwanted electrical interference, among other applications. Counting the number of zeros is also useful in determining properties of the signal, such as its period.

Apart from real world applications, the problem of root counting has many applications in other fields of Mathematics and related sciences. In fact, the motivation of this work stems from Computer Science or more precisely from Algebraic Complexity Theory. In Section 2.3 we provide a short introduction to this field, along with highlighting the necessary notions from the area that motivate our present work. In Section 2.3.1 we present the models of computation in this area of Complexity Theory, with a focus on algebraic circuits (also commonly referred to as arithmetic circuits). We also present the most important complexity measures, in particular circuit size which plays a crucial role in stating our motivation. We close that section by mentioning some fundamental structural results in the area.

Having defined the basics of the computational model of interest, we proceed in Section 2.3.2 to define the objects whose complexity we wish to study. In particular, the problem we wish to solve in this case is computing polynomials via algebraic circuits. More precisely, we wish to compute families of polynomials, such as for example the determinant for each $n \in \mathbb{N}$, with families of circuits, one for each value of the family parameter. We present the main complexity classes in Algebraic Complexity Theory and highlight the most important problems in them, as well as present a notion of reduction. This allows us to state the major open problem in the area, namely Valiant's conjecture, which is in some ways similar to the $\mathsf{P} \neq \mathsf{NP}$ question. Very briefly, we consider VP to be the class of "easy" to compute families of polynomials and VNP the one containing families of polynomials we suspect are hard to compute. Then Valiant's conjecture states that $\mathsf{VP} \neq \mathsf{VNP}$. We also state some related questions and related complexity classes and problems.

With the basics of Algebraic Complexity Theory in place, we proceed in Section 2.3.3 to rigorously define the link connecting Valiant's conjecture to the problem of studying the real zeros of univariate polynomials. That is achieved via the real τ-conjecture, which we state in that Section, while we also provide a detailed overview of its proof. This conjecture is about the number of real zeros of a specific set of univariate polynomials. Namely, consider the following expression of sum of products of k-sparse univariate polynomials:

$$\sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j}(x)\,.$$

The conjecture states that if every polynomial of the above form has a number of real roots that is polynomial in the parameters k, the number of summands, m, the number of factors in each product, and t, the sparsity of the $f_{i,j}$ polynomials, then Valiant's conjecture is true, thus $\mathsf{VP} \neq \mathsf{VNP}$. We are thus able to answer a complexity theoretic

question simply by an algebraic one. This result is similar to other results in Complexity Theory in the sense that it proves a trade-off between hardness and derandomization. Essentially, if the number of real roots of all such polynomials is sufficiently small, then it is possible to build a deterministic polynomial identity testing (PIT) algorithm for this set of polynomials in the form of a hitting set. A hitting set is a set of values such that for all polynomials in the set, with the exception of the zero polynomial, there is at least one value so that the polynomial is nonzero at this point.

Clearly, using a hitting set we can easily solve the PIT testing problem, which requires us to determine whether a given polynomial (as a circuit) is the zero polynomial, by evaluating at every point of the hitting set. The theorem that enables the real τ-conjecture then essentially proves the following trade-off: either PIT for this class of polynomials is easy to derandomize or the permanent is easy to compute, for notions of "easy" that are specific in the theorem's statement. However, both statements cannot be true simultaneously. Thus, if we show that the number of real roots for this class of polynomials is small, we obtain a sufficiently small hitting set and thus the permanent must be hard to compute. The real τ-conjecture then is essentially about the number of real roots of such polynomials. We state several forms of the conjecture, in fact the simplified version stated above is not the one given in the original statement of the conjecture.

With the motivation for the study of the number of real roots of univariate polynomials made clear, we proceed in Section 2.3.4 to state previous results in the case of sparse polynomials. Unlike the well-studied dense case which as we've previously mentioned we examine in Section 2.2.3, much less is known in the sparse case. Naturally, the aforementioned Descartes' rule of signs completely solves the problem for fixed polynomials when only the sparsity is concerned and although it cannot be further improved, it is often far from optimal. Further results have been achieved in the multivariate case, especially the classical work on "fewnomials" due to Khovanskii [49]. More recently work on random sparse polynomials has been presented, with our own work being part of this body of work. We briefly present these previous recent works before concluding this section.

Having introduced the necessary notions and preliminaries in Chapter 2, we are now able to present our results in Chapter 3. Most of these results were published in [43], a work by the author of the current work, together with colleagues and co-authors Dr. Gorav Jindal, Dr. Anurag Pandey and Himanshu Shukla, in which all authors contributed equally. This work explores the related notions in much greater detail and provides detailed proofs that were not possible to include in their entirety in the original work due to length constraints. Also, the proofs of both lower and upper bounds in this work have been scrutinized to achieve the best constants possible in the confines of the proof followed. While this does not alter the result asymptotically, constants in similar results in the dense case have been considered important enough even for standalone publications, especially since the upper and lower bounds match asymptotically. Furthermore, not only this allows us to state our results as accurately as possible, it may also serve as intuition for further work either into similar problems or for those wishing to improve the constants obtained.

We begin in Section 3.1 by presenting the reformulation of Kac's integral by Edelman and Kostlan [28]. We explore their approach in great detail, elaborating further on the points made on their own work. We do so in a manner that is general and could be

applied to any distribution and differentiable functions, although in this work we only consider univariate polynomials. Furthermore, we prove and state the property that makes the standard normal distribution one of the easiest distributions to tackle for the problem of computing the expected number of real zeros of random polynomials using this integral. We then proceed to state the version of the formula in this particular case for any family of univariate polynomials having the same support vector.

In Section 3.2 we present our upper and lower bounds for the expected number of real zeros for random k-sparse polynomials whose coefficients are i.i.d. random variables following the standard normal (Gaussian) distribution. We begin by rewriting the Edelman-Kostlan integral in a form that is more suitable to the calculations we wish to perform, while also being able to describe the integrand as the square root of this function. We then proceed to prove some basic results about this function that also translate into statements for the expected real zeros of random polynomials, when performing certain operations on their support vector. Using these statements and considering the support vector S as a set, for each expression of S as a disjoint union $S = S_1 \uplus S_2$, we are able to upper bound the expected number of real zeros for random Gaussian sparse polynomials with support vector S in terms of the same quantity for polynomials with support vectors S_1 and S_2. Together with simple proofs for the case of 1-sparse and 2-sparse polynomials, we can use this inequality to prove via induction an upper bound of $\frac{8}{\pi}\sqrt{k-1}$ for k-sparse polynomials. Furthermore, we are also able to show that similar to the dense case, most zeros concentrate around $x = 1$, providing important information for the distribution of the roots.

The above results are complemented by a lower bound that has the same asymptotic growth. While the above results were valid for all k dimensional support vectors, and thus k-sparse random polynomials, we focus on a specifically crafted support to achieve this bound. It should be already be clear from the dense case, which can be considered k-sparse for $k = d + 1$, that this lower bound is not achieved for every support, since the classic results of Kac gives a $\Theta(\log d) = \Theta(\log k)$ bound. The proof of the lower bound also proceeds iteratively by carefully considering the modified version of the Edelman-Kostlan integral. Once again, considering the support vector as a set that can be written as a disjoint union $S = S_1 \uplus S_2$ is key to the proof as it makes the iteration possible. The largest part of the lower bound consists of carefully bounding the added term that is obtained at every step of the iteration. Similar to the upper bound, we obtain an inequality for each step of the iteration, which together with the exact answer for 2-sparse random Gaussian polynomials allows us to derive a lower bound of $0.014\sqrt{k} - 0.007$.

Section 3.3 attempts to generalize key notions from the standard Gaussian case to arbitrary absolutely continuous distributions. As we've already mentioned, although the expected number of real zeros can be easily stated as an integral, solving the integral itself is highly non-trivial. Thus, it's highly unlikely that a general method exists to solve the integral for any distribution. We restrict ourselves to absolutely continuous integrals since they are guaranteed to have a probability density function. We then use a generalized version of the integral, already known to Edelman and Kostlan, to state the expected number of zeros given such a density function. Our contributions, which have not been published before, stem by carefully considering the proof strategy employed for the standard Gaussian case. In that case one can show that counting the number of zeros

is essentially the same as determining the arclength of the support curve in the interval $(0, 1)$. While the relation to the arclength cannot be derived for every distribution, since it depends on the specific properties of the standard normal distribution, we are still able to generalize certain steps of the methodology.

In fact, we are able to show that when studying the expected zeros of k-sparse random polynomials, we can always restrict our attention to the interval $(-1, 1)$. Note that for this to hold, we must only parameterize the polynomials by their sparsity k and no further restrictions are allowed, i.e., we must consider all k-sparse polynomials, with each contributing to the expectation according to the measure given by the distribution at hand. This is beneficial since it allows us to study the expected number of real zeros only in a bounded interval rather than the entirety of $\mathbb{R}$, which is unbounded. Secondly, we can show that if the distribution is symmetrical (about 0), then it only suffices to study the expected number of zeros in the interval $(0, 1)$. Since in this interval we can assume the variable x of the polynomial to be always positive, this either simplifies certain algebraic manipulations or allows them in the first place, which in our experience working with the standard normal distribution can be crucial to solving the integral.

We conclude this work in Chapter 4, in which we briefly summarize our main results, as well as the best known previous results on each case. In addition, we state avenues for future research, important open questions and hurdles that we consider important towards both proving similar results for more distributions as well as proving the real τ-conjecture, which as we've stated has been the main motivation for this work.

After this short introduction, we continue with presenting the rest of our work, as detailed above, starting with the basic notions necessary in Chapter 2.

CHAPTER 2

Notation and Preliminaries

2.1 Polynomials and Root Counting

2.1.1 Basic definitions

In this chapter, we introduce the notions, terminology, notations, as well as any assumptions and conventions, both basic and specific to the topic at hand, that we will adhere to throughout this work. Furthermore, we expand upon the ground notions and establish the necessary prerequisites that we build upon on later chapters. We attempt to follow the standard notation used in the relevant underlying fields of study to the degree that they are compatible and also do not interfere with the legibility and aims of this work. In particular, readers familiar with the field of Algebraic Geometry in Mathematics and Algebraic Complexity Theory in Computer Science should find the various notations and assumptions familiar. Naturally, if no conflict arises, we utilize the standard notation in mathematics, for example the natural numbers are denoted by $\mathbb{N} = \{0, 1, 2, \dots\}$, the field of real numbers by $\mathbb{R}$, the field of complex numbers by $\mathbb{C}$ and an arbitrary field by $\mathbb{K}$. Variables are labeled by $x_1, \dots, x_n$ and particularly when the number of variables n is greater than 1, we refer to the n-dimensional variable vector, namely $X = (x_1, \dots, x_n)$. In the case of a single variable, we simply use x to refer to it and omit references to a variable vector. Thus, $f(X) = f(x_1, \dots, x_n)$ refers to a function of the n variables, $x_1, \dots, x_n$, also called a n-variate function, and $f(x)$ to a function of a single variable x, i.e., a univariate function. Most of the functions encountered in this work will be associated with a polynomial and thus will be polynomial functions.

As should be evident already from the title of this work, polynomials play a central role in it and thus we prioritize their formal definition in great detail. As we already discussed in Chapter 1, polynomials have been studied extensively for thousands of years and thus should be to no surprise that depending on the field or application the specifics of the definition can differ to some degree, however all of them essentially attempt to capture the same intuitive notion. Furthermore, rather unsurprisingly given their fundamental role in mathematics, polynomials can be examined via a number of different formalizations, each focusing on the aspect of interest at hand and with a specific goal in mind. Even in the narrow context of this work, we require two different approaches and the corresponding definitions to the notion of the polynomial, which complement each other and are both necessary and fundamental in its scope. In addition, later in the chapter we present a further interpretation of a polynomial which is instrumental in constructing a geometric interpretation of our results and is thus invaluable with regard to the goal of making the results easy to understand and intuitive.

We begin with what is likely the most fundamental and general definition of the notion of polynomials, that is, the *formal polynomial*. This definition stems from abstract algebra

and regards polynomials as elements of a ring that has a special structure. Specifically, consider some arbitrary field $\mathbb{K}$ and the indeterminate x. In all generality, $\mathbb{K}$ can even be considered to be a ring, not necessarily commutative, however in this work it will always be a field and that field furthermore will almost always be that of the real numbers, $\mathbb{R}$. Now consider the ring obtained by adjoining x to $\mathbb{K}$. Since this new structure must be again a ring, it must be closed under the operations of addition $(+)$ and multiplication $(\cdot)$.[1] With regards to multiplication, apart from x, all its powers $x^d = \underbrace{x \cdot x \cdots x}_{d \text{ times}}$ for $d \in \mathbb{N}$ (under the usual assumption that x^0 equals 1, i.e., the identity element for multiplication) must also belong in this ring. Furthermore, any product of an arbitrary element $c \in \mathbb{K}$ with these powers, $c \cdot x^d$ must also be in the ring. It is assumed that the indeterminate x and elements of $\mathbb{K}$ always commute with respect to multiplication, regardless whether $\mathbb{K}$ itself is a commutative ring or not. Since in this work $\mathbb{K}$ is always a field and thus commutative, this distinction will be of no further concern. Turning to addition, any sum of the aforementioned forms is also in the ring, therefore we conclude that any expression of the form $\sum_{i=0}^{d} c_i x^i$ for $c_i \in \mathbb{K}$, $i \in \{0, \cdots, d\}$ and $d \in \mathbb{N}$ should be included in the new ring.

From the above, it should be clear that his newly constructed ring is nothing else than the polynomial ring $\mathbb{K}[x]$ of all polynomials in the variable x with coefficients from the field $\mathbb{K}$. In fact, the polynomial ring $\mathbb{K}[x]$ is the smallest field extension containing both $\mathbb{K}$ and the adjoined element x, since any field extension must contain both $\mathbb{K}$, the powers of x and their linear combinations, which describes $\mathbb{K}[x]$. Similarly, we can define polynomials on more than one variable as elements of the polynomial ring $\mathbb{K}[x_1, \ldots, x_n]$, by adjoining the elements $x_1, \ldots, x_n$ to $\mathbb{K}$. Recall that we demand that coefficients from $\mathbb{K}$ and variables commute. However, the same is not necessarily true for the variables themselves. If we do not require them to commute and thus we have that, e.g., $x_1 x_2 \neq x_2 x_1$, the resulting structure is called a free algebra and its elements noncommutative polynomials [Section 6.2][23]. However, in this work we will always assume the variables commute and thus we obtain the usual aforementioned polynomial rings in many variables. With the core notions established, we now present the definition of formal polynomials and the associated notation.

Definition 2.1 (Formal polynomial)**.** *Let $\mathbb{K}$ be a field and $X = (x_1, \ldots, x_n)$ be a n-dimensional vector called the variable vector, where $\{x_1, \ldots, x_n\}$ is the set of indeterminates and n is a positive integer, $n \in \mathbb{Z}^+$. Furthermore, let $E \subset \mathbb{N}^n$ be a finite set of n-dimensional vectors $e = (e_1, \ldots, e_n)$. We call the elements of the set E exponent vectors or simply exponents and E the set of exponents. A* formal polynomial *is an expression of the following form*

$$\sum_{e \in E} c_e X^e = \sum_{e \in E} c_e x_1^{e_1} \cdots x_n^{e_n} . \tag{2.1}$$

We denote a formal polynomial by $f(X)$ or $f(x_1, \ldots, x_n)$. All such polynomials form a ring which we call the polynomial ring on n variables over $\mathbb{K}$ and denote by $\mathbb{K}[X]$ or $\mathbb{K}[x_1, \ldots, x_n]$.

[1] As is common practice, we will often omit $\cdot$ and simply append the quantities to be multiplied

With regards to notation, when dealing with multiple polynomials we may employ other letters such as $h(X), g(X)$ etc., use indices, namely $f_1, f_2, \ldots, f_m$ or utilize both variations at once.

With the definition established, we also now state the related terminology. To begin with, note that in the definition we refer to the components of the *variable* vector as *indeterminates*. The difference between the two notions is subtle, in that when we refer to the indeterminate x, it refers to some element that as the name implies, is not determined further. This is useful when writing expressions and performing symbolic computations. For example, with the above definition of the formal polynomial, we can proceed to define more complex notions, such for example the complexity of computing a polynomial, without worrying which values x may obtain, which is the essence of symbolic computation. On the other hand, the notion of a variable x implies that x obtains values from some algebraic structure that should be determined. A relevant situation arises naturally in this work, since it concerns itself with the real zeros of univariate polynomials, therefore we care for values of the variable that are real numbers, that is, $x \in \mathbb{R}$.

Throughout this work, we will almost always use the term *variable* even if we use x symbolically, as this will improve the clarity of the text and the notion employed can be effortlessly derived from context. We only use the term indeterminate when we wish to emphasize that x should be thought of symbolically, like in the above definition, or where we deem there is enough ambiguity to affect the meaning of the text.

A product of the variables $x_1, \ldots, x_n$, each raised to the power denoted by the component of some exponent vector $e = (e_1, \ldots, e_n)$, that is $X^e = x_1^{e_1} \cdots x_n^{e_n}$ is called a *monomial* of exponent e. Note that if the corresponding component of the exponent vector is equal to zero, say $e_i = 0$, then we have $x_i^{e_i} = 1$ and we may simply omit that variable. Likewise, we may call X^e as a monomial on the variables that appear with a nonzero component on the exponent.

Further examining Definition 2.1, we see that each monomial X^e is multiplied by an element c_e from $\mathbb{K}$. In this context, we call $\mathbb{K}$ the field (or if appropriate, ring) of coefficients. The coefficients of a polynomial f are indexed by the set of exponents E and we say that the monomial X_e appears with coefficient c_e in f. By inducing a total order on the set of exponents E, we may present the coefficients of a polynomial as an $|E|$-dimensional vector which is appropriately named the *coefficient vector*.

The product of the coefficient and the corresponding monomial $c_e X^e$ is called a *term* of the polynomial. If a monomial has coefficient 0, we regard it as not appearing in the polynomial, in particular when we fix a set of exponents E and examine several polynomials over E. Therefore, one can regard a polynomial also as a sum of terms with nonzero coefficients. An exception to the latter condition is the *zero polynomial*, that is the polynomial where each monomial appears with coefficient 0. We denote the zero polynomial by 0, however it should not be confused with the constant $0 \in \mathbb{K}$. Unless explicitly clear from context, we make sure to stress this distinction whenever the zero polynomial appears in the text.

Unfortunately the terminology for monomials and terms is not standard; in fact in several areas of study the two definitions are swapped. Therefore, for the sake of both clarity and convenience, we will seldom refer to terms of a polynomial and rather refer to monomials and coefficients of that monomial. In certain instances this may lead to abuse of terminology where we will refer by monomial to the product of the coefficient

and the monomial itself, however it should be clear from context whether the coefficient is included or not.

We can easily define the operations of addition and multiplications for formal polynomials. Let $f(X), g(X) \in \mathbb{K}[X]$ be two formal polynomials. We define their addition as follows

$$(f+g)(X) = f(X)+g(X) = \sum_{e\in E_f} c_e X^e + \sum_{i\in E_g} c_i X^i = \sum_{j\in(E_f\cup E_g)} \left(\mathbf{1}_{E_f}(j)c_j + \mathbf{1}_{E_g}(j)d_j\right) X^j, \tag{2.2}$$

where $\mathbf{1}_{E_f}(j)$ is the indicator function of the set E_f, that is it is 1 if and only if $j \in E_f$. If $\mathbf{1}_{E_f}(j) = 0$ we may safely assume that $c_j = 0$ I. These notions apply similarly to the set E_g. Note that the result of addition is again a polynomial with exponent set $E_f \cup E_g$ and coefficients defined as above.

Multiplication is also easy to define. We have that

$$(fg)(X) = f(X)g(X) = \left(\sum_{e\in E_f} c_e X^e\right)\left(\sum_{i\in E_g} c_i X^i\right) = \sum_{e\in E_f}\sum_{i\in E_g} (c_e c_i)\, X^{e+i} = \sum_{j\in(E_f+E_g)} c_j X^j\,, \tag{2.3}$$

where $E_f + E_g = \{e + i \,|\, e \in E_f, i \in E_g\}$ is the Minkowski sum of E_f and E_g. The result is again a polynomial has exponent vector $E_f + E_g$ and coefficient as described above. Note that in the last equation different summands may result in the same exponent. From the above, as expected from a ring structure, it is easy to see that polynomials are closed as a set under both operations.

Having defined the above, we may also define when we regard two formal polynomials $f, g, \in \mathbb{K}[X]$, say $f = \sum_{e\in E_f} b_e X^e$ and $g = \sum_{e\in E_g} b_e X^e$, to be equal. f and g may only be equal if the respective sets of exponents are equal, $E_f = E_g = E$. Furthermore, it must be that for each $e \in E$, the coefficients of f and g corresponding to the monomial indexed by e must be equal, i.e., $\forall e \in E\,,\, b_e = c_e$.

We call e_i the degree of the variable x_i in the monomial $X^e = x_1^{e_1}\cdots x_n^{e_n}$ and write $\deg_{x_i}(X^e) = \deg_{x_i}(x_1^{e_1}\cdots x_n^{e_n}) = \deg(x_i^{e_i}) = e_i$. In addition, we define the *degree of the monomial* X^e as the sum of the degrees of its variables, that is

$$\deg(X^e) = \deg(x_1^{e_1}\cdots x_n^{e_n}) = \sum_{i=1}^{n} \deg(x_i^{e_i}) = \sum_{i=1}^{n} e_i\,. \tag{2.4}$$

For example, we call $x_1^{14}x_3^5$ a monomial of degree 19 in the variables x_1 and x_3. In this monomial, x_1 appears with degree 14 and x_3 with degree 5. We may say that x_2 (and any variables with higher index if relevant in the context) do not appear in the monomial or equivalently that they appear with degree 0.

In turn, we may define the degree of the polynomial $f(X)$ as the maximum over the degrees of all its monomials with nonzero coefficient. We denote the degree of the

polynomial by $\deg(f)$, $\deg(f(X))$ or $\deg(f(x_1, \ldots, x_n))$, employing the latter options when we wish to make clear the variables involved. Therefore, we have that

$$\deg(f) = \max\{\deg(X^e) \mid e \in E \wedge c_e \neq 0\}\,. \tag{2.5}$$

Furthermore, we may wish to refer to the degree of the polynomial f with respect to the variable x_i, which we denote by $\deg_{x_i}(f)$. The same definitions as above hold, however we apply them only to the variable of interest. Similarly, we may generalize this concept to any subset of the variables.

Thus, the polynomial $f(x_1, x_2) = \pi x_1^2 x_2^5 + 25x_1^{500} + \sqrt{2}$ has degree $\deg(f) = 500$. Furthermore, it has degree $\deg_{x_1}(f) = 500$ in the variable x_1 and $\deg_{x_2}(f) = 5$ in the variable x_2. Some notable cases of polynomials with respect to their degree are the constant polynomials and the zero polynomial. *Constant polynomials* are precisely the polynomials of degree 0, that is those polynomials whose only term that appears with nonzero coefficient is the one corresponding to the 0-vector exponent. Note that those stand in bijection with the elements of the field of coefficients $\mathbb{K}$, hence the name, since they involve no variables. With regards to the degree of the zero polynomial, there is no widely accepted definition. Some texts define the degree of the zero polynomial to be also 0, similar to the constant polynomials, others to be $-\infty$ or some arbitrary negative integer, often -1, while others choose to simply have the degree of the zero polynomial to be undefined. Since in the definition of degree of a polynomial in Equation (2.4) the set in the right hand side would be empty, it would have no maximum and thus for the sake of consistency we choose the degree of the zero polynomial to be undefined as well. In general, the zero polynomial will often appear in this work as a corner case and this choice emphasizes this fact and will allow the statement of certain results in all generality.

If all monomials of a polynomial have the same degree, we call that polynomial *homogeneous*. We denote the set of all homogeneous polynomials of degree d on the variable vector X over $\mathbb{K}$ by $\mathbb{K}[X]_d$. Note that this set together with the operation of addition forms a group. Note that thus the degree induces a grading of the associated polynomial ring. Often, we also wish to refer to polynomials of degree at most d, which we denote by $\mathbb{K}[X]_{\leq d}$. Note that any polynomial $f \in \mathbb{K}[X]_{\leq d}$ can be written as a sum of homogeneous polynomials from $\mathbb{K}[X]_i$ for $0 \leq i \leq d$. That is, for any polynomial $f(X)$ of degree d, there exist unique polynomials $f_i \in \mathbb{K}[X]_i$ such that

$$f(X) = \sum_{i=0}^{d} f_i(X)\,.$$

The above summation is unique and we refer to it as the *representation* of f into homogeneous parts.

Most of this work focuses on univariate polynomials in the variable x. All definitions stated above naturally apply in the univariate case as well, while some can be simplified to a high degree. For example, we can simply refer to the exponent of a monomial as a positive integer rather than a (1-dimensional) exponent vector. Similarly, the degree of a univariate polynomial $f(x)$ is always equal to its degree with respect to the variable x and by definition that is the maximum degree over all monomials that appear with a nonzero coefficient, see Equation (2.5). It follows that homogeneous univariate polynomials must

have a single monomial and that the representation of a univariate polynomial into its homogeneous parts is equivalent to considering the set of monomials that appear in the polynomial.

Definition 2.1 is a very general notion of what could be referred to as a polynomial. As its name betrays, it is also the most formal one and the technical correct one when referring to a "polynomial", especially in the context of abstract algebra. As we've already mentioned, regarding the variables X as indeterminates allows for great flexibility, since we do not have to specify the product of a coefficient and a monomial as an element of $\mathbb{K}$, other than its symbolic expression. The same holds for the sum of two or more monomials. Thus, this definition allows us to consider polynomial purely as a collection of coefficients and monomials, to which we may apply further structure if it is desired. For example, the coefficients of a polynomial could be numbers, e.g. elements from the field of reals giving rise to the ring $\mathbb{R}[X]$ or they could be more complex objects such as $n \times n$ matrices with elements from $\mathbb{K}$, giving rise to the ring $\mathbb{K}^{n\times n}[X]$. Although in Definition 2.1 we insisted on $\mathbb{K}$ being a field since this is the focus of this work, recall that the coefficients can instead belong to some ring, such as the ring of $\mathbb{K}^{n\times n}$, the ring of $n \times n$ matrices with entries from $\mathbb{K}$. In general, any ring can be used to give rise to a polynomial ring, including another polynomial ring $K[X]$ giving rise to the ring $(\mathbb{K}[X])[Y]$, where $X \cap Y = \emptyset$, i.e., we use "fresh" variables. In fact, this is often a construction that allows us to generalize polynomial rings in one variable to that of two or more variables. A polynomial in $\mathbb{R}[X]$ and one in $\mathbb{K}^{n\times n}[X]$ can be given wildly different additional structure that simply makes no sense in the context of the other, for example in the latter case we may refer to the determinant of a polynomial. Nevertheless, the notion of formal polynomial can be utilized to describe both cases.

This great generality makes this definition highly desirable in the context of algebraic computation and thus also for Algebraic Complexity Theory. Namely, we may assume that the coefficients of a polynomial come from some arbitrary, and most importantly, unspecified ring $\mathbb{K}$. This allows us to make generalized statements about polynomials and their computational complexity, as it will be defined in the following, without worrying about the specifics of $\mathbb{K}$. Furthermore, it allows such results and findings to be seamlessly specialized to interesting cases. Even when we concern ourselves with a specific field, such in our case where we are mostly interested in $\mathbb{R}[X]$, we need not concern ourselves with issues such as the representation of the coefficients, the precision of the operations and so on, unless we wish to do so, usually in the context of specific applications or numerical analysis. The strength of this definition will therefore become abundantly clear in the subsequent Section 2.3, where we will define the computational complexity of a polynomial.

On the other hand, quite often it becomes desirable to introduce additional structure to the notion of a polynomial that transcends the limits of Definition 2.1. In particular, in this work we concern ourselves with the zeros of a polynomial, that is, given a polynomial $f \in \mathbb{K}[X]$, we wish to find the values of X for which $f(X)$ evaluates to $0 \in \mathbb{K}$ with respect to the operations of addition and multiplication in $\mathbb{K}$ or some field extension $\mathbb{L} \supseteq \mathbb{K}$, which we may also denote by $\mathbb{L}/\mathbb{K}$. Note that this notion does not make sense in the context of formal polynomials. To begin with, since we must evaluate the polynomial for different values of X, we may no longer treat X as indeterminates but rather as variables obtaining values in $\mathbb{L}/\mathbb{K}$ or a restricted substructure of it, if so desired. Furthermore, to

derive the result of the evaluation itself, we must specify the operations of addition and multiplication in $\mathbb{L}/\mathbb{K}$ and by restriction, also in $\mathbb{K}$ or other appropriate substructure.

Note that it is still possible to maintain great generality by using an abstract field $\mathbb{K}$ and an appropriate, also abstract, field extension $\mathbb{L}/\mathbb{K}$, thus relinquishing only the use of indeterminates compared to formal polynomials. Essentially this is equivalent to using indeterminates for the coefficients of a polynomial as well and then carrying out any computations in a completely symbolic manner, where the assignment of some indeterminates as coefficients is only semantic and thus can be considered as arbitrary. Not only such an approach may be useful at times, quite often it is desirable due to the goal at hand of being much more specific. For instance, in this work we specifically wish to examine polynomials with coefficients over the reals $\mathbb{R}$, thus the field of coefficients is explicitly specified. Quite often in the context of Algebraic Complexity Theory, the field is explicitly specified or at the very least certain properties are fixed, such as the characteristic of the field or whether the field is algebraically closed, a notion that we will explicitly introduce below.

Once we have specified $\mathbb{K}$ and $\mathbb{L}$, we may then consider for each polynomial $f \in \mathbb{K}[X]$ and point $a \in \mathbb{L}^n$ the evaluation at the point $f(a)$. The set of ordered pairs $(a, f(a))$ is what we are interested in and the reader should have identified that this set precisely describes the notion of a function. Unsurprisingly, we call that function a *polynomial function* and define it precisely as follows:

Definition 2.2 (Polynomial function)**.** *Let $\mathbb{K}$ be a field and $\mathbb{L}/\mathbb{K}$ a field extension of $\mathbb{K}$. Given a formal polynomial $f(X) = \sum_{e \in E} c_e X^e$ in $\mathbb{K}[X]$, we associate with it a function $f : \mathbb{L}^n \to \mathbb{L}$ that we call the* polynomial function *f. This function maps each element of $\mathbb{L}^n$ to the evaluation of f when the variable vector X is equal to a, according to the operations of addition and multiplication in $\mathbb{L}$. That is, the function f is defined as*

$$f : a \mapsto f(a)\,.$$

We call the elements of $\mathbb{L}^n$ for which f evaluates to 0, $f(a) = 0$, the roots *of the polynomial f. For univariate polynomials, $n = 1$, we say that a root a of the polynomial $f(x)$ has* multiplicity *$m \in \mathbb{N}^+$ if m is maximal such that $(x-a)^m$ divides $f(x)$ perfectly, i.e., there exists $g(x) \in \mathbb{K}[x]$ such that $f(x) = (x-a)^m g(x)$ but no $h(x) \in \mathbb{K}[x]$ such that $f(x) = (x-a)^{m+1} h(x)$.*

We now make certain necessary clarifications. To begin with, note that we essentially use the same notation for polynomial functions as for formal polynomials. While at first this might seem like a choice that could lead to significant ambiguity and thus confusion, it is a very deliberate choice that we now justify. We start with the more menial reasons, building up to the intentions behind this decision. To begin with, note that polynomial functions and formal polynomials, while both related to the intuitive notion of a polynomial, are very different objects mathematically. The view of the formal polynomial is most useful when we view polynomials as a list of coefficients, with the indeterminates and exponents essentially providing a way with which to introduce an arbitrary order to this list. It also relates to polynomials as elements of a highly-structured ring, the polynomial ring, and how these structures behave in the context of abstract algebra. On the other hand, polynomial functions refer to concepts more familiar to classical algebra, such as solving algebraic, i.e., polynomial equations. In addition,

assuming the underlying field permits it, polynomial functions may be associated with notions such as the graph of a function or even concepts from integral calculus. It is thus clear that in most, if not all, cases, it should be abundantly clear from the context whether the notion $f(X)$ refers to a formal polynomial or a polynomial function.

On the other hand, both definitions have properties that we desire to associate with the intuitive notion of a polynomial. Since relating a polynomial function and a formal polynomial is rather straightforward and as noted above poses little risk for confusion, it seems sensible to associate both with the intuitive notion of a polynomial. In fact, that is often the approach taken when introducing polynomials for the first time and only much later the distinction between the two notions is made. We will follow the same convention and thus refer to both terms as "polynomials", with the context used to distinguish between formal polynomials and polynomial functions. In almost all cases, this distinction is immediate, in the rare cases it is not, we will explicitly fully specify the appropriate term. This allows us great flexibility, since we may use notions related to both definitions in our study of polynomials.

While in most cases this approach allows great flexibility and follows intuitively, there are cases where one must exercise caution. To begin with, note that although both definitions specify a field $\mathbb{K}$, polynomial functions do require that the operations of $\mathbb{K}$ are explicitly specified, while formal polynomials only require that the coefficients specified exist in $\mathbb{K}$. For example, consider the formal polynomial $f(x) = 2x^2+x+1$. For all number fields excluding $\mathbb{F}_2$, since 2 is not an element of that field, the above formal polynomial can be interpreted correctly. While in all formality we should consider the above $f(x)$ over say $\mathbb{F}_3$ and $\mathbb{C}$ as distinct formal polynomials, since we perform no operations on the polynomial itself, the distinction is only formal and may be disregarded with little potential for ambiguity. On the other hand, it is clear that while over $\mathbb{F}_3$ we may specify the polynomial function by simply listing the set of pairs of all possible evaluations $\{(0,1)\,(1,1)\,(2,2)\}$, over $\mathbb{C}$ clearly the polynomial function contains an infinite number of evaluations and thus clearly the two functions are not equal, even if we disregard they are defined over different fields with different operations. In summary, polynomial functions depend heavily on the field they are defined over, while formal polynomials do less so.

Although at first glance the above point seems to be pedantic, recall that eventually we wish to associate formal polynomials and polynomial functions in a way that corresponds to the intuitive notion of a polynomial. Thus, it is clear from the above that explicitly specifying the base field is crucial. The natural question that arises is that once the field is specified, what is the mapping from formal polynomial to polynomial functions. Ideally, we would like this to be a bijective mapping, since this would guarantee us that we could treat both the formal polynomial and the polynomial function as representatives of the same element and there would be no possibility of error. Unfortunately, this is not always the case. Consider for example the finite field of characteristic 2, $\mathbb{F}_2$. Then the formal polynomials 0, i.e., the zero polynomial, and $x^2 + x$ correspond to the same function that maps both elements of the field to 0. In fact, it is easily seen that the same is true for any polynomial of the form $x^d + x$, $n \in \mathbb{N}$, $n \geq 2$, since $\forall x \in \mathbb{F}_2$, $x^2 = x \pmod 2$. Thus clearly a bijective mapping between the two notions is impossible in this field.

In fact, a simple counting argument can show that this must be the case for any finite field: Let $\mathbb{K}$ be a finite field of order p, i.e., it contains p elements. For each $d \in \mathbb{N}$, there

exist p^{d+1} distinct formal polynomials and more importantly d may take any of an infinite number of values. It is thus clear that the number of polynomials in $\mathbb{K}[x]$ is countably infinite. A similar counting argument can be used in the multivariate case. However, since a polynomial function can be uniquely identified by the set of evaluations for each of the p elements of $\mathbb{K}$, it follows that p^p distinct (polynomial) functions exist.[2] Again, in the multivariate case it is easily seen this argument generalizes and the number of distinct functions is p^{p^n} where n is the number of variables. In both cases, it is clear that only finitely many functions exist, for a fixed number of variables n in the multivariate case. We thus reach the same conclusion that over finite fields, such a desired bijective mapping is impossible.

The above observation would seem to contradict our previous assurances that the two notions can be easily distinguished. Indeed, the above shows that over finite fields, an infinite number of formal polynomials correspond to the zero function, including naturally the zero polynomial, which is often a point of significant confusion for new students of polynomials over finite fields. Recall however that in this work we are mostly concerned with polynomials over $\mathbb{R}[x]$ and their roots in $\mathbb{R}$, thus we are not concerned by these peculiarities of finite fields. As we will now show, for many fields, including $\mathbb{R}$ and $\mathbb{C}$, we may in fact uniquely associate a polynomial function with a formal polynomial.

Proposition 2.3 (Zero function and zero polynomial coincide over infinite fields [94])**.** *Let $\mathbb{K}$ be an infinite field, $f_{formal}(X) \in K[X]$ a formal polynomial over $\mathbb{K}$ in n variables and $f_{function}(X) : \mathbb{L}^n \to \mathbb{L}$ the polynomial function associated with $f_{formal}(X)$ for an appropriate field extension $\mathbb{L}/\mathbb{K}$. Then, $f_{formal}(X)$ is the zero polynomial if and only if $f_{function}(X)$ is the zero function.*

The above can be proven by induction on the number of variables n. We treat the base case $n = 1$ separately, since it it will be later of independent interest to us. To prove the base case, we require the following well-known lemma.

Lemma 2.4 (Degree as an upper bound to the number of roots)**.** *Let $\mathbb{K}$ be a field and $f(x)$ be a polynomial function that corresponds to a formal polynomial with coefficients over $\mathbb{K}$ of degree d. Then, the number of roots of f in $\mathbb{K}$, counted with multiplicity, is at most d.*

Proof. We also proceed inductively, this time on the degree d of the polynomial $f(x)$. If f is of degree 0, then it is a constant polynomial $f(x) = c \in \mathbb{K}$. If $c \neq 0$, then clearly the polynomial has no roots and the above holds. If $c = 0$, by the deliberate definition given above, its degree is undefined, so that the statement holds vacuously. Alternatively, since it is clear constant polynomials have no roots unless they are the zero polynomial, we can exclude them and have polynomials of degree $d = 1$ as the base case. Then f will be of the form $f(x) = c_0 + c_1 x$ with $c_1 \neq 0$ and since $\mathbb{K}$ is a field it has exactly a single root given by $x = -\frac{c_0}{c_1}$, satisfying the Lemma.

For the induction step, assume the above holds for degree strictly less than d and suppose as in the statement that $f(x)$ has degree $\deg(f) = d$. Recall that all fields are also Euclidean domains and therefore so is $\mathbb{K}$. By their defining property, if $\mathbb{A}$ is

[2] As an additional point, note that any function over a finite field can be considered to be a polynomial function.

a Euclidean domain, there exists a function, called the *degree function*, $d : \mathbb{A} \setminus 0 \to \mathbb{N}$, such that for each $f, g \in \mathbb{A}$ with $g \neq 0$, there exist $q, r \in \mathbb{A}$ that satisfy $f = gq + r$ and either $r = 0$ or $d(r) < d(g)$. This generalizes the notion of Euclidean division over the integers and naturally we call q the quotient and r the remainder. Furthermore, if $\mathbb{K}$ is a field, then $\mathbb{K}[x]$ is a Euclidean domain [23, Section 10.2] (If X has $n \geq 2$ variables, $K[X]$ is a unique factorization domain but not a principal ideal domain and thus neither a Euclidean domain as well). It is clear from the above that the function that maps each nonzero polynomial $f \in \mathbb{K}[x]$ to its degree $\deg(f)$ fulfills the definition of the degree function.

Suppose now that x_0 is a root of f in $\mathbb{K}$. If such a x_0 does not exist, then f has no roots and the statement is trivially true. Since $x - x_0$ is of smaller degree than $f(x)$, there exist $q(x), r(x) \in \mathbb{K}[X]$ such that $f(x) = q(x)(x - x_0) + r(x)$. Now since $f(x_0) = 0$, we have that $r(x_0) = f(x_0) - q(x_0)(x_0 - x_0) = f(x_0) = 0$ as well. Furthermore, since by definition $\deg(r) < \deg(x - x_0) = 1$, r must be a constant polynomial and thus we conclude that $r(x)$ is the zero polynomial. We may thus write $f(x) = q(x)(x - x_0)$. From this expression, it is immediate that any zero of $f(x)$ other than x_0 must also be a zero of $q(x)$ and vice versa.

Moreover, note that $q(x)$ can be of degree at most $d - 1$. Suppose not and let $c_i x^i$ with $c_i \neq 0$ and $i \geq d$ be the term of highest degree of q. Then by expanding the above expression for $f(x)$, we would obtain the term $c_i x^{i+1}$ which cannot be cancelled out by any other term in the expansion. Thus the right hand side would be a polynomial of degree $i + 1 \geq d + 1$, while f is of degree only d, thus reaching a contradiction. Finally, since $\deg(q) \leq d-1$, by the induction hypothesis it can have at most $d-1$ zeros. Including x_0 and by the above observation for all other roots of $f(x)$ and $q(x)$, we conclude that $f(x)$ can have at most d zeros, as required. □

We have thus obtained a first, albeit rather trivial, bound on the number of roots of a polynomial f, in terms of its degree. Furthermore, note that the same argument as above can also be applied to any field extension $\mathbb{L}/\mathbb{K}$, thus obtaining the same bound. In the following, we will consider $\mathbb{L}/\mathbb{K}$ to be the minimal such field extension so that it contains all roots of $f(x)$ and obtain a relevant lower bound. Before that however, we must first present the proof of Proposition 2.3.

Proof of Proposition 2.3. Clearly, if $f_{\text{formal}}(X)$ is the zero polynomial, then $f_{\text{function}}(X)$ must be the zero function. In the other direction, assume that $f_{\text{function}}(X)$ is the zero function. We will show that if $\mathbb{K}$ is an infinite field, then f_{formal} must be the zero polynomial.

As already noted, we proceed inductively on the number of variables. For $n = 1$, since $f_{\text{function}}(x)$ is the zero function, it follows that every $a \in \mathbb{L}$ is a root of f_{function}. Since $\mathbb{L}$ is a field extension of $\mathbb{K}$, it is also infinite. Therefore, by Lemma 2.4, the degree of f_{formal} cannot be bounded and thus cannot take any value $d \in \mathbb{N}$. Since all nonzero polynomials have such a degree, we conclude that f_{formal} must necessarily be the zero polynomial.

For the inductive hypothesis, assume the statement holds for strictly less than n variables. We can consider $f_{\text{formal}}(x_1, \ldots, x_n)$ as an element of $\mathbb{K}[x_1, \ldots, x_{n-1}][x_n]$, that is, a univariate polynomial in x_n with coefficients from $\mathbb{K}[x_1, \ldots, x_{n-1}]$. If f_{formal} is the zero polynomial in that polynomial ring, it must also be in the original, by the obvious ring isomorphism. Furthermore, although $\mathbb{K}[x_1, \ldots, x_{n-1}]$ is only a ring and our theorem

requires a field, we may simply consider its elements to be elements from the appropriate field of fractions, that is the field of rational functions $\mathbb{K}(x_1, \ldots, x_{n-1})$. We may thus write

$$g(x_n) = f(x_1, \ldots, x_n) = \sum_{i=0}^{deg_{x_n}(f)} f_i(x_1, \ldots, x_{n-1}) x_n^i \,,$$

where $f_i(x_1, \ldots, x_{n-1})$ are appropriate polynomials in $\mathbb{K}[x_1, \ldots, x_{n-1}]$. Now note that for each $a \in \mathbb{L}$, $g(a) = 0$ and that g is zero as a function. Thus, by the induction hypothesis, it follows that $g(x_n)$ is the zero polynomial in $\mathbb{K}[x_1, \ldots, x_{n-1}][x_n]$. Therefore, by the aforementioned ring isomorphism, it must be that $f_{\text{formal}}(X)$ is the zero polynomial in $\mathbb{K}[X]$.

□

The following corollary follows immediately from Proposition 2.3.

Corollary 2.5 (Equality of polynomials via their functions)**.** *Let $\mathbb{K}$ be an infinite field and $f(X), g(X) \in \mathbb{K}[X]$ be two (formal) polynomials. Then $f(X) = g(X)$ as elements of $\mathbb{K}[X]$ iff $h(X) = f(X) - g(X)$ is the zero function.*

Since in this work our focus is on infinite fields, specifically the reals $\mathbb{R}$ and in extension, the complex numbers $\mathbb{C}$, Proposition 2.3 formalizes the aforementioned flexibility in referring to both formal polynomial and polynomial functions to simply as "polynomials" without the need to be precise in the distinction. This property is significant in the context of Algebraic Complexity Theory. Note for example, that over an infinite field, Proposition 2.3 allows us to determine whether a polynomial is the zero polynomial by evaluating the associated polynomial function. In turn, solving this problem has numerous applications and connections to many important problems in the associated theory, to which already Corollary 2.5 hints to.

Note that so far, we have referred to some "appropriately chosen" extension $\mathbb{L}/\mathbb{K}$, so that $\mathbb{L}$ contains all roots of polynomials in $\mathbb{K}[X]$, without further specification. It is clear that this notion is relevant to us, since we wish to study the number of zeros that belong to a specific field and it is possible that a polynomial may have roots that belong in some field extension, which would affect the final answer. The common example one encounters of this is by considering the polynomial $f(x) = x^2 + 1$ over $\mathbb{R}[x]$. From Lemma 2.4, f can potentially have 2 roots, but it is clear that for all $x \in \mathbb{R}$, the function is strictly positive and thus the polynomial has no real roots. This paves the way for the introduction of the complex numbers $\mathbb{C}$, since by extending $\mathbb{R}$ by the imaginary unit i that is defined so that $i^2 + 1 = 0$ (and thus also $(-i)^2 + 1 = 0$), we obtain two roots for the polynomial. In fact, the above can be formally shown by proving that the quotient ring $\mathbb{R}[x]/(x^2 + 1)$ is isomorphic to $\mathbb{C}$ [23, Section 7.2]. Therefore, by selecting $\mathbb{C}$ as the field extension, not only we find additional roots of the above polynomial that were not in $\mathbb{R}$, but in fact we obtain 2 of them, thus matching the upper bound by the degree of the polynomial. Two questions arise naturally: First, does $\mathbb{C}$ contain all roots for a polynomial in $\mathbb{R}[X]$ or do we require to extend further? And secondly, if $\mathbb{C}$ does indeed contains all roots of a polynomial, is their number always equal to the degree of the polynomial, as in the example?

As we've already mentioned in Chapter 1, these questions have arisen naturally through thousands of years of practical applications and in particular as some of the earliest results of modern mathematics. Not only these results were fundamental in the early studies of classical algebra, i.e., algebraic equations, but also do happen to provide the very first lower bound regarding the number of zeros of a univariate polynomial. As the example above showcases, it is possible that in the original field $\mathbb{K}$ a polynomial might even have no zeros, which nevertheless can be used to extend the field. Thus, to derive a lower bound, we must demand a property from this extension, for which then we can also answer the two questions posed. The property is that we wish that the field $\mathbb{K}$ is such that all roots of all polynomials in $\mathbb{K}[x]$ lie in $\mathbb{K}$ itself. If that is not the case, we wish to extend $\mathbb{K}$ to a way similar to how we extended $\mathbb{R}$ to $\mathbb{C}$ so that we obtain a field extension $\mathbb{L}/\mathbb{K}$, potentially repeating the process until we obtain a field that has the property. We now define the relevant notions and make several observations about them.

Definition 2.6 (Algebraic extension)**.** *Let $\mathbb{K}$ be a field and let $\mathbb{L}/\mathbb{K}$ be a field extension. We call an element $a \in \mathbb{L}$ an* algebraic element *over $\mathbb{K}$ if and only if there exists a nonzero polynomial $f \in \mathbb{K}[x]$ such that $f(a) = 0$. We call $\mathbb{L}$ an* algebraic extension *of $\mathbb{K}$ if all its elements are algebraic over $\mathbb{K}$.*

Having defined what an algebraic extension of a field $\mathbb{K}$ is, our aim is clear: We can construct an operation based on algebraic extensions, where we examine whether there exists a polynomial in $\mathbb{K}[x]$ that has a root not in $\mathbb{K}$. Suppose such a polynomial exists, say $f(x)$ and furthermore we choose f to be minimal, in the sense that it is irreducible over $K[x]$, i.e., it cannot be factorized further. This is possible since as we already mentioned, if $\mathbb{K}$ is a field, then $\mathbb{K}[x]$ is a unique factorization domain. Then we may obtain a field extension isomorphic to $\mathbb{K}[x]/(f(x))$. We continue this process until no such polynomial exists. Then by definition, for that algebraic extension $\mathbb{L}/\mathbb{K}$, all polynomials over $\mathbb{L}[X]$, and thus $\mathbb{K}[x]$ as well, have roots that are contained in $\mathbb{L}$. We call such an algebraic extension *algebraically closed*.

Before formally defining the property of being algebraically closed, two questions arise: First, whether an algebraic extension exists such that a polynomial $f(x) \in \mathbb{K}[x]$ that has a root not in $\mathbb{K}$ has a root in that field extension. As already noted, it is sufficient to choose $f(x)$ to be irreducible, which is guaranteed to exist by the unique factorization property, and then consider the field extension isomorphic to $\mathbb{K}[x]/(f(x))$. The second issue is that it is conceivable that the process never terminates: i.e., that there is an infinite chain of algebraic field extensions of a field $\mathbb{K}$. This would mean that no field extension of $\mathbb{K}$ that is algebraically closed exists. Thankfully, this is not the case and numerous proofs of the existence of such an algebraic extension exist. In particular, the proof in [42] is both short and elegant, while simultaneously demanding only minimal prior knowledge, since it only requires defining an appropriate order on the set of algebraic extensions of a field $\mathbb{K}$ and then utilizing Zorn's lemma to show that this set must have a maximal element, which by its maximality must be algebraically closed. It can also further shown that the algebraic closure of $\mathbb{K}$ must be unique, up to isomorphism.

Having guaranteed the uniqueness and existence of an extension of every field $\mathbb{K}$, we may now formally define when the field itself attains that property:

Definition 2.7 (Algebraic closure of a field and algebraically closed fields)**.** *Let* $\mathbb{K}$ *be a field. We call* $\overline{\mathbb{K}}$ *the* algebraic closure *of* $\mathbb{K}$ *if it is the maximal algebraic extension of* $\mathbb{K}$*, i.e., it is not a subfield of any algebraic extension of* $\mathbb{K}$ *other than itself. If* $\mathbb{K} = \overline{\mathbb{K}}$*, we say that* $\mathbb{K}$ *is* algebraically closed*.*

For example, as we've already seen $\mathbb{R}$ is not an algebraically closed since the polynomial $x^2 + 1$ has no real roots. A more general example comes from finite fields. Let $\mathbb{F}_q$ be the finite field of order q, where q is a prime power, i.e., $q = p^k$ for p prime and $k \in \mathbb{N}^+$. Note that the polynomial $x^q - x$ vanishes for all $a \in \mathbb{F}$, a fact we already used to show there exist nonzero polynomials that map to the zero function over finite fields. Now, consider the polynomial $f(x) = x^q - x + 1$. By the above, it follows that f has no roots in $\mathbb{F}_q$ and thus F_q is not algebraically closed. In fact, one can show that the algebraic closure of $\mathbb{F}_q$ is simply

$$\overline{\mathbb{F}}_q = \bigcup_{i \in \mathbb{N}^+} \mathbb{F}_{q^i} .$$

We are now in a position to answer the two questions posed above. In the terminology we have introduced, we want to know whether $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$ and if yes, is there a lower bound on the number of roots of a polynomial with respect to its degree? Both questions can be answered simultaneously in the following way: It is sufficient to show that every nonconstant polynomial $f(x) \in \mathbb{R}[x]$ has at least one complex root, say $a \in \mathbb{C}$. It can then be shown that this statement is equivalent to each polynomial $f(x)$ in $\mathbb{R}[x]$ having exactly $\deg(f)$ complex roots and equivalent to $\mathbb{C}$ being the algebraic closure of $\mathbb{R}$.

Although as we've noted in Chapter 1, the field of algebra has expanded to encompass a plethora of topics other than the study of algebraic equations, the theorem was and still is considered of significant importance to earn the moniker "Fundamental Theorem of Algebra", where Algebra is best interpreted in the narrow sense of algebraic equations. We now formally state the theorem, as well as a Corollary that contains the aforementioned equivalent formulations, before providing a proof.

Theorem 2.8 (Fundamental Theorem of Algebra)**.** *Let* $f(x)$ *be a nonconstant polynomial in* $\mathbb{R}[x]$*, i.e.,* $\deg(f) \geq 1$*. Then there exists* $a \in \mathbb{C}$*, such that* $f(a) = 0$*.*

Corollary 2.9 (Alternative formulations of the Fundamental Theorem of Algebra)**.** *The following statements are equivalent to Theorem 2.8*

(1) Every nonzero polynomial $f(x) \in \mathbb{R}[x]$ *of degree* $\deg(f)$ *has precisely* $\deg(f)$ *complex roots, counted with multiplicity.*

(2) $\mathbb{C}$ *is the algebraic closure of* $\mathbb{R}$*, i.e.,* $\mathbb{C} \cong \overline{\mathbb{R}}$*.*

Proof. The equivalence can be shown to follow through elementary arguments:

- Theorem 2.8 $\implies$ (1)

  We prove the implication via induction on the degree $\deg(f)$ of the polynomial f. For $\deg(f) = 1$, the statement is straightforward and it is clear that not only (1) holds but also that $a \in \mathbb{R}$.

Now for the inductive step, note that by the statement of the Fundamental Theorem of Algebra, for every nonconstant $f(x) \in \mathbb{R}[x]$ there exists $a \in \mathbb{C}$ such that $f(a) = 0$. Then, as we've already seen in the proof of Lemma 2.4, the polynomial $x - a$ must divide $f(x)$ perfectly, say $f(x) = q(x)(x - a)$. Following the same argument, it is clear from this expression that the quotient $q(x)$ must have degree exactly $\deg(q) = \deg(f) - 1$ and all its roots must also be roots of $f(x)$. However, by the induction hypothesis, q must have $\deg(q)$ roots, thus the statement follows immediately.

- (1) $\implies$ (2)

  Let $f(x)$ be a nonzero polynomial in $\mathbb{R}[x]$. By (1), it follows that f must have $\deg(f)$ complex roots. Furthermore, note that this number of roots matches the upper bound already proven in Lemma 2.4 and thus it follows that there can be no further roots. It follows that for every nonzero polynomial in $\mathbb{R}[x]$, all its roots lie in $\mathbb{C}$. Since $\mathbb{C}$ is an algebraic extension of $\mathbb{R}$, this is precisely the definition of the algebraic closure given in Definition 2.7.

- (2) $\implies$ Theorem 2.8

  Suppose the implication does not hold. Then, there must exist a nonconstant polynomial $f(x)$ that does not have a complex root. We may then form the algebraic extension $\mathbb{R}[x]/(f(x))$, which must necessarily contain some root of $f(x)$, say a. By the above assumption it follows that $a \notin \mathbb{C}$. However, since $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$ and unique up to isomorphism, it follows by definition that it is the maximal algebraic extension of $\mathbb{R}$. Therefore it must contain $\mathbb{R}[x]/(f(x))$ and in particular a, thus we have reached a contradiction.

□

Having showed that the statements are indeed equivalent, it remains to provide a proof for Theorem 2.8. Numerous proofs of the Fundamental Theorem of Algebra exist, varying in their length, simplicity and assumptions. We follow here an elegant proof by Professor David H. Bailey, which can be found online [4]. The proof is explicitly constructed in order to be both elementary and self-contained with a minimal set of prerequisites. We include the proof in the following both for the sake of this work to be self-contained, as well as for posterity. Furthermore, we also add some additional clarifications and notes to further clarify the proof.

Proof of Theorem 2.8. We will actually prove the stronger statement where we allow the coefficient of the polynomial to be complex numbers instead of just reals. Therefore, we have $f(x) \in \mathbb{C}[x]$, which naturally contains also all real polynomials as a subset. Similarly, the equivalent statements presented above hold for this generalization as well. We only require the following prerequisite for this proof, which is true for both the reals and the complex numbers due to their completeness as metric spaces.

Proposition 2.10. *Let $f(x)$ be a real-valued continuous function defined on a closed set $S \subseteq \mathbb{C}$. Then there exists a point for which f attains its minimum value in S, i.e.,* $\exists t \in S : \forall s \in S\,,\ f(t) \leq f(s)$

For complex functions, we can apply the above by considering the magnitude of the function, which is a real-valued function and continuous if the original function is, as a composition of continuous functions. Also note that while this property might seem intuitive for real or complex numbers, it is not straightforward for every field. For example, consider the field of rational numbers $\mathbb{Q}$ and the real-valued continuous function $f(x) = (x - \pi)^2$. We can easily construct sets of rational numbers that approximate the solution, for example consider the exact finite approximations of π up to the n-th digit, for $n \in \mathbb{N}$. This set will approach infinitesimally close to the actual minimum that is achieved for $x = \pi$, however that limit point is not a rational number and thus the function never achieves its minimum in $S \cap \mathbb{Q}$.

Continuing with the proof, consider a polynomial $f(z) \in \mathbb{C}[z]$ given by $f(z) = \sum_{i=0}^{n} c_i z^i$. Let us first examine polynomials of degree 1, which must necessarily have $c_1 \neq 0$, otherwise they would be constant. It is clear that such a polynomial $f(z) = c_0 + c_1 z$ has precisely a complex root at $z = -\frac{c_0}{c_1}$. For any polynomial of degree larger than 1, we can assume that $c_0 \neq 0$ and $c_n \neq 0$. This does not result in a loss of generality since if $c_0 = 0$, then clearly $z = 0$ is a complex solution of $f(z)$ and if $c_n = 0$, we may consider a polynomial of smaller degree, until we obtain a polynomial with non-zero leading coefficient, with polynomials of degree 1 forming the base case. Thus, in the following we may assume that $\deg(f) = n$.

We now choose a closed set S so that any potential root of $f(z)$ will always be contained in it. Consider the magnitude $|f(z)|$ of the complex polynomial $f(z)$. Note that for z of sufficiently large magnitude, the magnitude of the leading term $|c_n z^n|$ will dominate the magnitude of the sum of the remaining terms. Indeed, let $m \coloneqq \max_i \left|\frac{c_i}{c_n}\right|$ and note that it must be that $m \geq 1$. Then, assuming that $|z| > 2m \geq 2$, we have that $\left|\sum_{i=0}^{n-1} c_i z^i\right| < |c_n z^n|$. Indeed, it is

$$\begin{aligned} \frac{1}{|c_n|}\left|\sum_{i=0}^{n-1} c_i z^i\right| &\leq \sum_{i=0}^{n-1}\left|\frac{c_i}{c_n}\right||z|^i \leq m\sum_{i=0}^{n-1}|z|^i = m\frac{|z|^n - 1}{|z| - 1} && (|z| - 1 > 2m - 1 > 0) \\ &< \frac{m}{2m-1}\left(|z|^n - 1\right) \leq |z|^n - 1 \leq |z|^n . && (m \geq 1) \end{aligned}$$

Furthermore, note that $\lim_{|z|\to\infty} |f(z)| = \lim_{|z|\to\infty} |c_n z^n| = \infty$, i.e., the magnitude of the dominating leading term and thus, of the polynomial, tends to infinity. It follows that given any $B > 0$, there exists a sufficiently large s, such that $|f(z)| > B$ for any $|z| \geq s$. We choose $B = 2|f(0)| = 2|c_0|$ and consider the closed disk $|z| \leq s$. By the above, we can guarantee that the magnitude of the polynomial is strictly positive outside this disk, thus any possible root of the polynomial can only occur within this closed disk.

Now, note that the closed disk $S \coloneqq |z| \leq s$ is a closed set and that $|f(z)|$ restricted on this disk is a real-valued continuous function. Thus, Proposition 2.10 applies and $|f(z)|$ attains its minimum value on the disk. Furthermore, by the choice of s above, note that for $|z| = s$ we have $|f(z)| > B = 2|f(0)| \geq |f(0)|$ and $z = 0$ is in the interior of the disk. It follows that the minimum must necessarily be obtained in the interior of the disk S, say at some point t. Let this value be $|p(t)| = M$. If this minimum is obtained in more than one point, we may arbitrarily choose one without loss of generality.

If $|p(t)| = 0$, it follows that $p(t) = 0$ and thus t is a root of $f(z)$, which indeed is what the theorem implies. Thus, we will suppose that $|p(t)| = M > 0$, i.e., that the minimum value of the magnitude of the polynomial within the disk S (and in extension in $\mathbb{C}$ by our assumptions) is strictly positive and derive a contradiction.

Consider the polynomial $q(z) = f(z+t) = \sum_{i=0}^{n} q_i z^i$, where we translate the argument so that the minimum is obtained at $z = 0$. We will show that there exists a value z_0 such that $|q(z_0)| < |q(0)|$, thus contradicting the minimality of t. Note that if such a value exists, it must also be in the disk S, since we would have $|q(z_0)| < |q(0)| < B$. We now choose such a value. Consider $r \in \mathbb{R}$ be an arbitrarily small positive real value that will be defined precisely later. Also let q_j be the nonzero coefficient of $q(z)$ with the smallest positive index, i.e., excluding q_0. Since f is nonconstant, such a coefficient is guaranteed to exist. Then, we select the following value

$$z_0 = r\left(\frac{-q_0}{q_j}\right)^{1/j}.$$

We will now show that under the above assumptions, we have $|q(z_0)| < |q(0)|$. We have

$$q(z_0) = q_0 - q_j\frac{q_0}{q_j}r^j + q_{j+1}\left(\frac{-q0}{q_j}\right)^{j+1/j} r^{j+1} + \cdots + q_n\left(\frac{-q0}{q_j}\right)^{n/j} r^n = q_0 - q_0 r^j + E\,,$$

where E stands for all the terms with index strictly larger than j. We will now show that $|E|$ can be upper bounded. Recall that $|z_0| = r\left|\frac{q_0}{q_j}\right|^{1/j}$ and assume we choose r small enough such that $|z_0| < 1$, i.e., $r < \left|\frac{q_j}{q_0}\right|^{1/j}$. Also let $q_{max} \coloneqq \max_i |q_i|$. We then have

$$\begin{aligned}
|E| &= \left|q_{j+1}\left(\frac{-q_0}{q_j}\right)^{j+1/j} r^{j+1} + \cdots + q_n\left(\frac{-q_0}{q_j}\right)^{n/j} r^n\right| = \left|\sum_{i=j+1}^{n} q_i\left(\frac{-q_0}{q_j}\right)^{i/j} r^i\right| \\
&\leq \sum_{i=j+1}^{n} |q_i|\left(r\left|\frac{q_0}{q_j}\right|^{1/j}\right)^i \leq \sum_{i=j+1}^{n} q_{max}\,|z_0|^i = q_{max}\,|z_0|^{j+1}\sum_{i=0}^{n-j-1}|z_0|^i \qquad (|z_0| < 1) \\
&\leq q_{max}\frac{|z_0|^{j+1}}{1-|z_0|}\,. \qquad (2.6)
\end{aligned}$$

We now would like to choose r small enough such that we may bound $|E|$ by a small enough value, in particular we would like to have

$$\begin{aligned}
&|E| < |q_0|\frac{r^j}{2} \overset{eq.\ (2.6)}{\iff} q_{max}\frac{|z_0|^{j+1}}{1-|z_0|} < |q_0|\frac{r^j}{2} \iff \frac{q_{max}}{1-|z_0|}\frac{|q_0|^{1+1/j}}{|q_j|^{1+1/j}}r^{j+1} < |q_0|\frac{r^j}{2} \\
&\iff 2q_{max}\,|q_0|^{1/j}\,r < |q_j|^{1+1/j}\,(1-|z_0|) \iff 2q_{max}\,|q_0|^{1/j}\,r < |q_j|^{1+1/j} - |q_j|\,|q_0|^{1/j}\,r \\
&\iff r < \frac{|q_j|^{1+1/j}}{|q_0|^{1/j}\,(2q_{max}+|q_j|)}\,. \qquad (2.7)
\end{aligned}$$

In fact, the condition given by Equation (2.7) is the only one needed, since it supersedes the previous one given. It is straightforward to verify this as follows

$$\frac{|q_j|^{1+1/j}}{|q_0|^{1/j}\left(2q_{max}+|q_j|\right)} < \left|\frac{q_j}{q_0}\right|^{1/j} \iff |q_j| < 2q_{max}+|q_j| \iff q_{max} > 0\,,$$

with the last inequality being true by the definition of q_{max} and the fact that q is a non-constant polynomial. Since this is only an upper bound on r, we may choose it to be arbitrarily small, as we originally noted. Now with the above bound on $|E|$, namely that $|E| < |q_0|\,\frac{r^j}{2}$ and r arbitrarily close to 0 and certainly $0 < r < 1$, we have

$$|q(z_0)| = \left|q_0 - q_0 r^j + E\right| \leq \left|q_0 - q_0 r^j\right| + |E| < |q_0|\,(1-r^j) + |q_0|\,\frac{r^j}{2} < |q_0|\left(1-\frac{r^j}{2}\right) < |q_0|\,.$$

Now by the simple observation that $|q_0| = |q(0)|$, we see that we have shown that indeed there exists a value z_0 where q obtains a value smaller than its claimed minimum and thus we have reached a contradiction. We therefore conclude that the minimum magnitude of q (and thus also f) on the interior of the disk must be 0 and thus we can always find a complex number a that is a root of f, i.e., $f(a) = 0$. □

The above proof crucially uses certain properties of the complex numbers, as we've already noted in Proposition 2.10, namely that $\mathbb{C}$ together with the usual metric, given by the magnitude of the difference of two complex numbers, is a complete metric space. Nevertheless, similar statements can be proven for any algebraic closure: if $\mathbb{K}$ is algebraically closed, any polynomial f of degree $\deg(f)$ will have exactly $\deg(f)$ roots. To see this, note that the Fundamental Theorem of Algebra is crucial in proving the field of complex numbers $\mathbb{C}$ is indeed the algebraic closure of $\mathbb{R}$, not in proving that such a closure must exist, since we have already seen that the existence of a closure is guaranteed for any field regardless. Furthermore, given any field $\mathbb{K}$, it can be shown that any nonconstant polynomial over $\mathbb{K}$ must have at least one root in its algebraic closure $\overline{\mathbb{K}}$. Then it's straightforward to adapt Corollary 2.9 for any $\mathbb{K}$, since the proof only uses basic facts about algebraically closed fields not particular to $\mathbb{C}$, thus guaranteeing that the number of roots matches the degree.

It follows that for algebraically closed fields $\mathbb{K}$, counting the number of roots of any polynomial f over $\mathbb{K}$ has a straightforward answer that requires no further study: It must be precisely the degree $\deg(f)$. Nevertheless, an interesting question emerges effortlessly: Given a field $\mathbb{K}$ that is *not* algebraically closed, how many roots of a given polynomial f lie in $\mathbb{K}$ itself and not in some non-trivial algebraic extension? The example of the real numbers $\mathbb{R}$, which is also the object of our study, shows that this is an interesting question from both a theoretical and practical perspective: Theoretically, it is easy to construct real polynomials f whose number of real roots is $N \coloneqq \deg(f) - 2k$ for any $k \in \mathbb{N}$ such that the above quantity is nonnegative.[3] For example, given any N real numbers, say $x_1, \ldots, x_N$, the polynomial $g(x) = \prod_{i=1}^{N}(x - x_i)$ clearly has N real roots.

[3]That the number of real roots must have the same parity as the degree of the polynomial follows in a straightforward way from the easy to prove fact that non-real roots of a real polynomial must appear in conjugate pairs.

Then, we may multiply this with an even polynomial $h(x)$ of degree $2k$ with nonzero constant term, which guarantees that $h(x) > 0$ for all $x \in \mathbb{R}$. This product will indeed have exactly N real roots and the degree specified, i.e., $\deg(f)$.

Practically, given an equation that describes some physical phenomenon or problem we wish to examine, it is often the case that we are interested in real solutions to the equation. An easy example from school-level physics is determining when an object thrown at some initial speed upwards at an angle will reach a certain height. The object will follow a parabolic trajectory and thus its height at a given time t can be expressed as a quadratic equation $h(t)$. Naturally, since we will solve for time (e.g., in seconds), we only care for real solutions to this equation. This rather inconspicuous problem can be quickly be made rather complicated and useful in practice. For example, we may consider the object to be a space rocket and the trajectory to be a polynomial approximation of its trajectory. We may now wish to know not only at which times the rocket will be at a certain orbit height above Earth but also the count of such crossings of the given orbit, if for example we wish to deploy a certain payload such as a satellite at that height. This is an example where not only the real solutions, but also their number would be of practical interest.

Returning to our more down to earth theoretical concerns and since we established the necessary background in the above, we now may define some useful notation before stating a question that lies in the center of this work.

Definition 2.11 (Number of roots of a polynomial over some field)**.** *Let $f \in \mathbb{K}[x]$ be a polynomial and $S \subseteq \overline{\mathbb{K}}$ be a set. Then we denote the number of roots of f in S by $N_S(f)$ and say that f has $N_S(f)$ roots in S. In particular if $\mathbb{K} = S = \mathbb{R}$, we omit the subscript and simply write $N(f)$ and refer to real roots of f.*

While this definition is quite broad, in this work we will concern ourselves with only a handful of sets, namely the entirety of $\mathbb{R}$ and certain simple subsets that arise naturally in our study. We may now pose the following core question:

Question 2.12 (Real zeros of a real polynomial)**.** *Given a polynomial $f \in \mathbb{R}[x]$, what is the number of its real roots $N(f)$?*

As the reader would hopefully be convinced during the reading of this chapter, this is an interesting question on its own right and has been studied in various forms for centuries and up to this day. Before examining this question in further detail, we briefly mention some generalizations. It is obvious from Definition 2.11 that we may pose such a question for any field $\mathbb{K}$ and set S, often with $S = \mathbb{K}$. Apart from the case of real roots of real polynomials, investigating the existence or the number of integer or rational roots of polynomials is often a question that arises naturally in the study of various problems, with number theory being an obvious example in the case of integer roots. Many problems, e.g., in cryptography, also require working modulo some prime, in such case the same questions can be posed over the appropriate finite field.

Furthermore, we may pose the same question for multivariate polynomials. We restrict ourselves to $\mathbb{R}$ for simplicity and consider the simplest case of a bivariate polynomial, i.e. $f \in \mathbb{R}[x_1, x_2]$. If we were to consider the number of roots of a single polynomial, we quickly run into trouble. While there exist polynomials where the number of real roots is finite, e.g. ,$f(x_1, x_2) = x_1^2 + x_2^2$ clearly only has the real root $(0, 0)$, in all generality

the number of roots will be infinite. To see this, fix x_1 and note that for each such value we obtain a polynomial, say $g_{x_1}(x_2) \in \mathbb{R}[x_2]$ whose roots are the values of x_2 such that (x_1, x_2) is a root of f. Thus, we may describe the roots of f in the form $(x_1, g_{x_1}(x_2))$,[4] which for most cases will produce an infinite number of roots. The solution is to introduce an additional equation, which in the general case will result to only a finite number of solutions. We thus obtain *polynomial systems*, which we now define:

Definition 2.13 (Polynomial system)**.** *Let $f_1, \ldots, f_m \in K[X]$ be n-variate polynomials. We call a* polynomial system *$F(X)$ a set of m polynomial equations of which we require to be simultaneously satisfied:*

$$F(X) = \begin{cases} f_1(X) = 0\,, \\ \quad \vdots \\ f_m(X) = 0\,. \end{cases}$$

This is equivalent to determining the common roots of $f_1, \ldots, f_m$. We say that some $X \in \overline{\mathbb{K}}$ is a solution *of the system F if it satisfies all equations and is thus a common root.*

Expanding upon the above observations, we can see that in the general case if the number of equations m is less than the number of variables n, i.e., $m < n$, the number of solutions will be infinite. Such systems are called *underdetermined*. On the other hand, if $m > n$, in the general case no solutions exist and we call such systems *overdetermined*, since the number of constraints, i.e., equations, is greater than the number of variables. More interestingly, if $m = n$, then the number of solutions is either infinite or at most the product of the degrees of the polynomial equations $\deg(f_1) \cdots \deg(f_m)$. This is essentially the statement of Bézout's theorem, see Chapter 7 in [35], which can be thought of as the multivariate generalization of the Fundamental Theorem of Algebra. If $\mathbb{K}$ is algebraically closed, then in the case of finite roots those are precisely the product of the degrees for $m = n$ and underdetermined systems have either 0 or infinite solutions, although as we've already noted this case is not as interesting to us.

Furthermore, the relevant notation easily generalizes to systems of polynomial equations. In particular, we may refer to the number of solutions of the polynomial system in the set S by $N_S(F)$ or $N_S(f_1, \ldots, f_m)$. In the context of polynomial systems, we may thus pose a generalization of Question 2.12.

Question 2.14 (Number of roots of polynomial system)**.** *Given a polynomial system $F = \{f_1(X) = 0, \ldots, f_n(X) = 0\}$ of n-variate polynomials, what is the number of its real solutions $N(F)$?*

In this work we will only refer to certain results and recent advances regarding Question 2.14 and focus instead on the univariate case. Nevertheless it is a topic of active research that actively interfaces with the work presented herein and thus we will often refer to such work, such as the classical treatment [49] by Khovanskii. The reader already familiar with this work will note that we have not yet introduced a notion central to the study of root counting, an omission which we will amend in the following.

[4] Note that g depends on x_1. In general it is not possible to produce a closed expression for x_2 in terms of x_1, as it follows from the insolubility of the quintic.

2.1.2 Sparse polynomials

In the previous section we built the necessary background to examine the number of roots of a polynomial in a formal setting. We also concluded via the Fundamental Theorem of Algebra and its generalizations that over algebraically closed fields such as the complex numbers $\mathbb{C}$, the number of roots of a polynomial is precisely determined by its degree. On the other hand, we demonstrated that the number of real roots of a polynomial may vary from 0 to matching the degree itself. Given this close relationship between the number of complex roots and the degree, it is natural to seek such an elegant statement for the number of real roots as well. Namely, given a polynomial in $\mathbb{R}[x]$, is there a quantity associated with it that completely determines its number of real roots? It turns out that given a polynomial, there is such a quantity that is easy to compute, that provides an upper bound on the number of real zeros. Furthermore, this upper bound is optimal, with respect to the aforementioned quantity, since there exist polynomials whose number of real roots match the bound exactly. Unfortunately, there are also cases where the actual number of real roots deviates strongly from said quantity. To state the bound, we require to define a number of related notions, which we now proceed to.

We begin with the notion of *sparsity*, which plays a central role in this work. Consider a polynomial f of degree exactly $\deg(f) = d$, which can thus be written as $f(x) = \sum_{i=0}^{d} c_i x^i$. Note that this information reveals very little for the polynomial: we know that the coefficient c_d must be nonzero due to the degree requirement, but not much else. The polynomial f could be as simple as $f(x) = c_d x^d$, with all other coefficients being zero and thus be trivial to study, or as complex as having $n+1$ nonzero coefficients and exhibit a much more complicated behavior in many regards, including its number of real roots. It seems natural therefore to define a quantity, which captures precisely this notion.

Definition 2.15 (Sparsity of a polynomial)**.** *Let $f(x) \in \mathbb{R}[x]$ be a polynomial of degree d and write $f(x) = \sum_{i=0}^{d} c_i x^i$. We write $k(f)$ or simply k and define as the* sparsity *of the polynomial f the number of non-zero coefficients of f, that is*

$$k = k(f) \coloneqq |\{c_i \mid c_i \neq 0\,,\, 0 \leq i \leq d\}|\,.$$

Almost always the polynomial f will be referred to implicitly and be clear from context and thus we will simply use k to denote its sparsity. We will also say that f is k-sparse. By the definition it is clear that $k \in \mathbb{N}$ and if f is nonzero, then $k \geq 1$. It is also clear that the sparsity k of a polynomial f is upper bounded by a quantity almost equal to its degree, since there are $\deg(f) + 1$ coefficients and all of them could potentially be nonzero. Furthermore, while our focus is on univariate polynomials, the notion naturally extends to the multivariate case.

Note that as we've already hinted to, the notion of sparsity is a good indicator of how "complex" it is to describe a polynomial. Assume that we either demand arbitrary precision at constant cost (i.e., infinite precision), as will be the assumption in this work, or more realistically in applied cases, we decide on a certain fixed precision. Under either assumption we assume that there is a fixed cost for representing a single element from the base field $\mathbb{K}$. This cost is measured in the number of bits necessary for the chosen level of precision and we denote it by p. Then if f is k-sparse, we know that we can describe it by its list of coefficients in pk bits. This concept is referred in the literature as *descriptive*

complexity and as the name betrays, in many situations it is a much better measure of the complexity of performing computations on a polynomial than its degree. This scenario is particularly interesting when the sparsity k of a polynomial is much smaller than its degree d, $k << d$. We refer to such polynomials as sparse and contrast then with the dense case when the sparsity matches its upper bound $k = d + 1$ or more loosely, when the sparsity is asymptotically equal to the degree, $k = \Theta(d)$. We will explore these notions more closely in subsequent chapters, particularly in Section 2.3.

That sparsity closely relates to the complexity of describing polynomial does not only emerge in a computational context but can already emerge in terms of notation. Clearly, we may omit zero coefficients of a polynomial without the loss of any information. More precisely, it is sufficient to list all non-zero coefficients of a polynomial to identify it completely. Assuming a polynomial f of degree d is k-sparse, we may write it as a sum of k terms, obtaining thus such a short description. Recall Definition 2.1 and note that we may project the exponent vector to only those indices that correspond to a term with a nonzero coefficient without loss of information. We may thus instead of the original trivial exponent vector $(0, \dots, d)$, use $E = (i \mid c_i \neq 0\,,\, 0 \leq i \leq d) = (e_0, \dots, e_{k-1})$, where we sort the exponents in ascending order. Similarly, we may consider the projection of the coefficient vector according to E, that is $A = (c_i \mid i \in E) = (a_0, \dots, a_{k-1})$. For example, if $f(x) = 3 + 5x^3 - 25.834x^8 + ex^{10}$, f has sparsity 4, it has the exponent vector $E = (0, 3, 8, 10)$ and the coefficient vector $A = (3, 5, -25.834, e)$.

In all generality, we may use this compact representation to write

$$f(x) = \sum_{i=0}^{k-1} a_i x^{e_i} \quad , a_i \in A\,,\, e_i \in E\,.$$

A brief examination also reveals that this form can be derived from the (projected) coefficient and exponent vectors mentioned above and in addition that these vectors are necessary to fully determine the polynomial. Furthermore, it is often convenient to consider set of sparse polynomials that share the same exponent vector $E = (e_0, \dots, e_{k-1})$. In that case, we often refer to the vector $(x^{e_0}, \dots, x^{e_{k-1}})$ as the *support vector* of the polynomial or set of polynomials in question.

While sparsity can easily capture the complexity of describing the polynomial, it only determines the number of non-zero coefficients. This is still useful and even desirable in certain contexts, since often we deal with sets, or more precisely, families, of polynomials whose coefficients can vary wildly. However we lose information for each coefficient, specifically when talking about polynomials in $\mathbb{R}[x]$, the magnitude (i.e., absolute value) and sign of each coefficient, which are also sufficient to completely define any real number. While in the context of both root counting and identification the absolute magnitudes of the coefficients are irrelevant, since $f(x)$ and $cf(x)$ for any nonzero $c \in \mathbb{R}$ have exactly the same roots, that is not the case with the sign of the coefficients. For example, a trivial observation is that if all signs are the same, then the polynomial has no positive roots. Although the ratio of magnitudes of the coefficients together with their signs would reveal the whole picture, it is essentially equivalent to listing the coefficient themselves and thus equivalent to specifying the polynomial completely. Even listing just the sign of the coefficients already requires at least k bits in the worst case, which at least asymptotically is equivalent to specifying the polynomial, assuming fixed (or infinite)

precision. A compromise would be, as in the case of sparsity, to simply list the number of sign changes of the sequence of coefficients. Furthermore, since multiplying by -1 does not affect the roots, whether we start with a positive or negative coefficient is irrelevant. It turns out that indeed this number is really useful in bounding the number of real roots. Before proceeding to show how this is possible, we formally define this notion and relate it to sparsity.

Definition 2.16 (Number of sign changes of the coefficient sequence)**.** *Let $f \in \mathbb{R}[x]$ be a k-sparse polynomial, say $f(x) = \sum_{i=0}^{k-1} a_i x^{e_i}$. Consider the sequence obtained by listing the components of the coefficient vector $A = (a_0, \ldots, a_{k-1})$ in ascending order according to the index, i.e., the sequence $a_0, \ldots, a_{k-1}$. We call this sequence the* coefficient sequence *of the polynomial f. We say that a* sign change *occurs in the sequence when two subsequent terms have opposite signs, that is precisely when*

$$a_i a_{i+1} < 0\,.$$

We denote the number of sign changes *of the coefficient sequence of f by $\#s(f)$ and define it in the natural way by*

$$\#s(f) = |\{a_i \mid a_i a_{i+1} < 0\,,\, 0 \leq i \leq k-2\}|\,.$$

Furthermore, we say that the coefficients of f have $\#s(f)$ sign changes.

As we noted above, the number of sign changes of the coefficients can be used to provide an upper bound on the number of real roots of a polynomial in $\mathbb{R}[x]$. We now state this celebrated result due to Descartes, originally published in his work "La Géométrie" in 1637 and commonly referred to as Descartes' rule of signs.

Theorem 2.17 (Descartes' rule of signs, [25])**.** *Let $f(x) \in \mathbb{R}[x]$ be a polynomial with a coefficient sequence with $\#s(f)$ number of sign changes. Then the number of* positive *real roots of f counted with multiplicity is at most $\#s(f)$, that is*

$$N_{\mathbb{R}^+}(f) \leq \#s(f)\,.$$

Among the myriad of proofs of the above theorem, we present a simple proof originally attributed to Gauß that is commonly taught together with the theorem and is essentially folklore. A modern exposition of the proof in detail can be found in [57].

Proof. Let $f(x) = \sum_{i=0}^{d} c_i x^i$ be a real polynomial and let $A = (a_0, \ldots, a_{k-1})$ be its nonzero coefficient vector. We will prove the above by induction on the number of positive real roots. The base case of f having no roots is immediate, since the number of sign changes $\#s(f)$ is by definition a non-negative number.

Now assume the result is true for polynomials with strictly less than N roots. We will now show the result for f with exactly N positive roots, assuming the inductive hypothesis. Let $a \in \mathbb{R}^+$ be any root of f. By the unique factorization of polynomials, there exists a unique polynomial $g(x)$ of degree $d-1$ with exactly $N-1$ positive roots such that

$$f(x) = (x-a)g(x)\,.$$

Since g has $N-1$ positive roots, the inductive hypothesis applies. Thus, it follows that g must exhibit at least $N-1$ sign changes. We will use the above form to show that in this case, f has to exhibit at least N sign changes, which will conclude the proof. We will do this by comparing the sign changes of the coefficient sequences of f and g. Let $g(x) = \sum_{i=0}^{d-1} g_i x^i$. Expand the above form for f, obtaining

$$\begin{aligned} f(x) &= (x-a)\sum_{i=0}^{d-1} g_i x^i = \sum_{i=0}^{d-1} g_i x^{i+1} - \sum_{i=0}^{d-1} a g_i x^i = \sum_{i=1}^{d} g_{i-1} x^i - \sum_{i=0}^{d-1} a g_i x^i \\ &= -a g_0 + \sum_{i=1}^{d-1} \left(g_{i-1} - a g_i\right) x^i + g_{d-1} x^d\,. \end{aligned}$$

From the above expression we may thus express the coefficients of f by

$$\begin{cases} c_0 = -a g_0\,, \\ c_i = g_{i-1} - a g_i\,,\ 1 \leq i \leq d-1\,, \\ c_d = g_{d-1}\,. \end{cases} \tag{2.8}$$

Let us now compare the sign changes of the two coefficient sequences. To begin with, clearly c_d has the same sign as g_{d-1} since they are equal. Let M be the exact number of sign changes occurring in g. We clearly have $N-1 \leq M \leq d$. Now consider the subsequence $g_{i_0}, \ldots, g_{i_j}, \ldots, g_{i_{M-1}}$ of the coefficient sequence of g obtained by selecting only those coefficients where a sign change occurs, that is $g_{i_{j-1}} g_{i_j} < 0$ for $0 \leq j \leq M-1$. We now note that the signs of $g_{i_{j-1}}$ and c_{i_j} agree. This is easy to verify case by case, namely

$$\begin{cases} g_{i_{j-1}} < 0\,,\ g_{i_j} > 0\ :\ c_{i_j} = g_{i_{j-1}} - \overbrace{a g_{i_j}}^{>0} < 0\,, \\ g_{i_{j-1}} > 0\,,\ g_{i_j} < 0\ :\ c_{i_j} = g_{i_{j-1}} - \underbrace{a g_{i_j}}_{<0} > 0\,. \end{cases} \tag{2.9}$$

From the above we can conclude that the coefficient sequence of f must exhibit at least as many sign changes as g, that is at least $M \geq N-1$. To see this, consider the subsequence $c_{i_0+1}, \ldots, c_{i_j+1}, \ldots, c_{i_{M-1}+1}$ of the coefficient sequence of f. By construction this subsequence contains M terms. Furthermore, it is easy to see by Equation (2.9) and the definition of the sequence $g_{i_0}, \ldots, g_{i_j}, \ldots, g_{i_{M-1}}$ that the signs of this subsequence in fact alternate. It follows that this subsequence exhibits exactly M sign changes and thus the coefficient sequence of f exhibits at least as many, since the number of sign changes in this subsequence will clearly lower bound the number of sign changes in the original sequence.

Finally, note that we have not yet considered $c_0 = -a g_0$. By definition, g_0 and g_{i_0} must have the same sign. Furthermore since $a > 0$, c_0 has the opposite sign of g_{i_0} and by extension of c_{i_0+1}, thus at least one additional sign change occurs in the coefficient sequence of f compared to that of g. It follows that the coefficient sequence of f exhibits at

least $M+1 \geq N$ sign changes, where N is the number of positive roots of f, $N := N_{\mathbb{R}^+}(f)$, as required by the theorem. □

With some additional work, one can even prove that similar to the case of degree, the number of sign changes $\#s(f)$ and the number of positive real roots $N_{\mathbb{R}^+}(f)$ of a real polynomial f must have the same parity. We thus obtain the following corollary.

Corollary 2.18 (Parity of number of sign changes and number of positive roots)**.** *Let $f(x) \in \mathbb{R}[x]$ be a polynomial with $N_{\mathbb{R}^+}(f)$ positive roots and having a coefficient sequence with $\#s(f)$ number of sign changes. Then $\#s(f)$ has the same parity as $N_{\mathbb{R}^+}(f)$, i.e.,*

$$\#s(f) = N_{\mathbb{R}^+}(f) + 2l\,, \quad \textit{for some } l \in \mathbb{N}\,.$$

Proof. That $N_{\mathbb{R}^+}(f)$ is smaller clearly follows immediately from Theorem 2.17. To show that the two quantities share the same parity we employ a proof that requires only two facts:

First, note that given a coefficient sequence, the parity of $\#s(f)$ is the same as the number of sign changes in the subsequence c_0, c_n, i.e., whether it exhibits a sign change or not. To see this, consider that the signs of the leading and trailing coefficients are fixed and the rest have a positive sign. The parity of the number of sign changes clearly matches that the number of sign changes of the subsequence. In fact unless both c_0 and c_n have a negative sign, the numbers are equal. Now imagine we correct the signs as needed to match those of the coefficient sequence of f, one at a time. Changing a sign from positive to negative can have the following effect on the number of sign changes: If it is adjacent to another negative sign, clearly the number of sign changes will remain the same. Otherwise, if both signs bordering it are positive, then switching the sign to a negative one increases the number of sign changes by 2. In particular, the parity remains the same. We may repeat this process until we obtain the signs of the coefficient sequence of f.

Secondly, we may apply the same reasoning if the number of known signs is more than two, that is we know the signs of some subsequence, as was the case in the proof of Theorem 2.17. Therefore, we may conclude that if $f(x) = (x-a)g(x)$ as in the proof and g exhibits m sign changes, then f exhibits $m+1+2l$ sign changes for some $l \in \mathbb{N}$, where the subsequence obtained together with the trailing coefficient exhibits $m+1$ sign changes. We may now apply this argument recursively for every positive real root a_i of f. In each step, the parity of the number of sign changes will remain the same.

It remains to deal with the two possible base cases. If all the roots of f are positive real numbers, the last step will be a linear polynomial of the form $x - a_N$ for $a_N > 0$. This polynomial has exactly one root and one sign change, thus not only the parity but also the numbers match exactly. The other base case is if after repeating the process for all positive real roots of f, we obtain a polynomial g_M having no positive real roots. It should be the case that the parity of $\#s(g_M)$ must be even, same as 0. Suppose not. Then, $g_{M,0}$ and $g_{M,d}$, the trailing and leading coefficients of g_M, must be of opposite sign, since the subsequence $g_{M,0}, g_{M,d}$ must have the same parity as the number of sign changes as mentioned above. However $g_M(0) = g_{M,0}$ and as noted in the proof of Theorem 2.8, for large enough x the sign of the polynomial is the same as the sign of the leading coefficient, since it dominates. By the continuity of polynomials, this would imply that at least one positive root must exist, contradicting the assumption of no positive roots. □

Before continuing with analyzing Theorem 2.17 and gaining some much needed intuition, we offer a number of easy to derive corollaries that are nevertheless useful in giving a bound on the number of zeros over the entirety of $\mathbb{R}$ instead of only the positive real roots. We begin with negative real roots, where we obtain the same bound almost immediately.

Corollary 2.19 (Descartes' rule of signs for negative real roots)**.** *Let $f(x) \in \mathbb{R}[x]$ be a polynomial and define $g(x) \coloneqq f(-x)$. Also denote the negative real numbers by $\mathbb{R}_{<0} \coloneqq \mathbb{R} \setminus (\mathbb{R}^+ \cup \{0\})$. Let $\#g(x)$ be the number of sign changes in the coefficient sequence of g. Then the number of* negative *real roots of f counted with multiplicity is at most $\#s(g)$, and if smaller, it is so by an even number. Therefore we have*

$$\#s(g) = N_{\mathbb{R}_{<0}}(f) + 2l\,, \quad \text{for some } l \in \mathbb{N}\,.$$

Proof. Clearly for every negative real root x_0 of f, $f(x_0) = 0$, $-x_0$ is a positive real root of g, since $g(-x_0) = f(x_0) = 0$. Therefore by applying Theorem 2.17 and Corollary 2.18 on g, we immediately obtain the result. □

Note that since $c(-x)^{2l} = cx^{2l}$ and $c(-x)^{2l+1} = -cx^{2l+1}$, we obtain the same coefficient sequence as g by switching the sign of every coefficient corresponding to an odd exponent of f, which we mention since it is a common way of stating this result.

It only remains to determine the real zeros precisely at $x = 0$ to obtain the complete picture. It is straightforward to show that the exponent of the trailing coefficient not only provides a bound on the zeros at 0 counted with multiplicity, but in fact completely determines them.

Corollary 2.20 (Real zeros at zero)**.** *Let $f(x) = \sum_{i=0}^{k-1} c_i x^{e_i}$ be a real k-sparse polynomial. If $e_0 = 0$, then 0 is not a root of f, otherwise it is a root with multiplicity exactly e_0. That is, we have*

$$N_{\{0\}}(f) = e_0\,.$$

Proof. If $e_0 = 0$, then we have that $f(0) = c_0 \neq 0$, therefore f indeed has no root at $x = 0$ (we may also say in a slight abuse of terminology that it has a root of multiplicity 0). Otherwise if $e_0 > 0$, clearly $f_0 = 0$. Furthermore, for $0 < m < e_0$, the m-th order derivative has a trailing exponent of $e_0 - m > 0$ and thus we also have that $f^{(m)}(0) = 0$. On the other hand, the trailing exponent of the (e_0)-th derivative is precisely 0 and the trailing coefficient is equal to $c_0\,(e_0)! \neq 0$. Since $f(0) = \cdots = f^{(e_0-1)}(0) = 0$ and $f^{(e_0)} \neq 0$, we conclude that f has a zero at $x = 0$ of multiplicity exactly e_0, as required. □

Summing up the above, the next Theorem follows immediately by applying Theorem 2.17, Corollary 2.19 and Corollary 2.20, thus providing a bound on the number of real roots with multiplicity.

Theorem 2.21 (Bound on real roots based on Descartes)**.** *Let $f(x) = \sum_{i=0}^{k-1} c_i x^{e_i}$ be a real polynomial with a coefficient sequence with $\#s(f)$ number of sign changes. Also let $g(x) \coloneqq f(-x)$ be a polynomial with a coefficient sequence with $\#s(g)$ number of*

sign changes. Then the number of real roots of f counted with multiplicity is at most $\#s(f) + \#s(g) + e_0$, and if smaller, it is so by an even number. That is

$$N_{\mathbb{R}}(f) + 2l = \#s(f) + \#s(g) + e_0\,, \quad \textit{for some } l \in \mathbb{N}\,.$$

We now would like to elaborate further upon the Descartes' rule of signs, since not only it is still a useful tool in studying sparse polynomials, but it also forms a baseline with which to compare bounds obtained by other means, as the ones we present in this study. Before we form this baseline and examine some generalizations of Theorem 2.17, we would like to present some intuition on why the theorem is true. Compared with bounds obtained by the degree, the Descartes' rule of signs is not as straightforward to discern. Lemma 2.4 that presents the bound of the degree is straightforward to justify, especially when restricting our attention to real univariate polynomials: Clearly, a polynomial cannot have more roots than its degree, since then it could be written as a polynomial of strictly higher degree by expanding its unique factorization. Furthermore, such a justification is immediate by the proof of Theorem 2.8, since while a polynomial has degree at least 1, we may follow the proof of the Fundamental Theorem of Algebra to verify the existence of a root, say x_0, and then divide that polynomial by $x - x_0$ and repeat the process. On the other hand, the proof of the Descartes' rule of signs does not indicate why a real root should exist, since it assumes its existence and proceeds to show that it is reflected on the number of sign changes. We now present two arguments to rectify this discrepancy and in the process hopefully inform the intuition of the reader.

The first argument essentially foregoes the assumption required by the proof that indeed N roots exist and instead identifies intervals that may contain roots by examining the monotonicity of the function under certain assumptions. These assumptions are equivalent to assuming that the maximum possible number of roots occurs. We obtain this worst case by assuming that in each interval where the function is decreasing and has a positive value in the left boundary, a root will occur and by consequence, the right boundary will have a negative value. We follow the same reasoning for the "mirrored" scenario of an increasing function with a negative value at the left boundary of the interval. To obtain information on the monotonicity of the function, we rely on the signs of the coefficients and employ a special type of derivative. In particular, a sign change implies a change in monotonicity and thus allows for a potential root by the above reasoning.

We mention the use of a special kind of derivative, which are simple enough to be considered folklore, although we were not able to identify a standard terminology for them. We thus take the liberty to define them as *coefficient-sign* derivatives. They are formulated specifically with the study of positive real roots in mind. The idea behind them is rather simple: Given a polynomial $f(x) = \sum_{i=0}^{k-1} a_i x^{e_i}$, we may divide it by x^{e_0}, i.e., the trailing term without its coefficient, without affecting the number of positive real roots. Then, taking the derivative has the effect of eliminating the trailing coefficient and by dividing once more with the variable raised to the trailing exponent of the result, i.e., $x^{e_1 - e0 - 1}$ we ensure the next nonzero coefficient of the intermediate outcome is the constant term of the final result. Note that derivating does not alter the sign of the coefficients but if the trailing term is constant, then it may eliminate one, which by construction is guaranteed in this process. Therefore we essentially obtain a subsequence of the signs of the coefficient sequence, disregarding the trailing coefficient at every step.

We then repeat the process until we obtain a constant polynomial, i.e., essentially only examining the sign of the leading coefficient. The above process is equivalent to studying the monotonicity of f for $x > 0$ under the assumptions that allow for the worst-case scenario in terms of the number of positive roots $N_{\mathbb{R}^+}(f)$. We now give the formal definition of this type of derivative.

Definition 2.22 (Coefficient-sign derivative)**.** *Let $f(x) = \sum_{i=0}^{k-1} a_i x^{e_i}$ be a real univariate polynomial. We define its (first-order)* coefficient-sign *derivative by*

$$\frac{\partial_{cs} f}{\partial x}(x) = f'_{cs}(x) = \sum_{i=1}^{k-1} a_i \left(e_i - e_0\right) x^{e_i - e_1} .$$

By convention, the 0-th order coefficient-sign derivative of f is obtained by only dividing by x^{e_0}, i.e., without derivation. Higher-order coefficient-sign derivatives may be obtained by repeated application of the operation and the i-th order coefficient-sign derivative is denoted simply by $f_{cs}^{(i)}(x)$.

Note that evaluating the i-th coefficient-sign derivative at $x = 0$ is equal to the i-th nonzero coefficient of the polynomial multiplied by a positive factor, namely the product of the differences of the i-th exponent with all exponents of smaller index. That is, we have $f_{cs}^{(i)}(0) = \prod_{j=0}^{i-1} \left(e_i - e_j\right) a_i$. By convention for $i = 0$, we simply obtain the trailing coefficient with no multiplicative factor. In particular, the value of the i-th coefficient-sign derivative at $x = 0$ shares its sign with the i-th coefficient of the polynomial, hence the name we have chosen. We will use the sign function $\operatorname{sgn}(x)$ in the following, recall that

$$\operatorname{sgn}(x) = \begin{cases} 1 & , x > 0 \,, \\ 0 & , x = 0 \,, \\ -1 & , x < 0 \,. \end{cases}$$

It should now start to become clearer how we may use coefficient-sign derivatives to obtain some insight into Descartes' rule of signs. Assume without loss of generality that the leading coefficient is positive, $c_{k-1} > 0$, since if it is not, we may simply multiply the polynomial by -1 without changing its roots. As we've already seen in the proof of Theorem 2.8, this immediately implies that for large enough x, the leading term dominates and thus all coefficient-sign derivatives are strictly positive for some sufficiently large x. On the other hand, examining the value of the $(k-1)$-th order coefficient-sign derivative at $x = 0$, we have that $\operatorname{sgn}\left(f_{cs}^{(k-1)}(0)\right) = \operatorname{sgn}\left(c_{k-1}\right) > 0$. In fact since it is constant the same is true for any $x \in \mathbb{R}^+$. Therefore the $(k-1)$-th coefficient derivative is positive everywhere in $\mathbb{R}^+$ which implies that the $(k-2)$-th coefficient derivative is strictly increasing for $x > 0$, since the division operation may only remove zeros at $x = 0$ and does not otherwise affect the sign.

Now proceeding similar for the $(k-2)$-th coefficient-sign derivative, we have that $\operatorname{sgn}\left(f_{cs}^{(k-2)}(0)\right) = \operatorname{sgn}(c_{k-2})$. If $c_{k-2} > 0$, since the $(k-2)$-th coefficient derivative is already positive at 0 and strictly increasing for $x > 0$, it must be positive in the entirety of $\mathbb{R}^+$. On the other hand, if $c_{k-2} < 0$, then not only a sign change appears in the

coefficient sequence of the original polynomial f but also the $(k-2)$-th derivative must have some positive root, say x_0, by continuity. Since the first scenario is identical to the previous step and does not increase the number of roots, we proceed with the second one and now examine the $(k-3)$-th coefficient derivative in turn. Note that $f_{cs}^{(k-3)}$ must be strictly decreasing in $(0, x_0)$ and strictly increasing in (x_0, ∞). If $c_{k-3} < 0$, i.e., f exhibits no sign change, this reduces to the previous case and we can guarantee at least one root by continuity, as above. More interestingly if $c_{k-3} > 0$, we have two distinct scenarios: If $f_{cs}^{(k-3)}(x_0) > 0$, then there is no root, since this is the minimum of $f_{cs}^{(k-3)}$ in $\mathbb{R}^+$. If on the other hand $f_{cs}^{(k-3)}(x_0) < 0$, then we can be certain that two positive real roots appear for the $(k-3)$-th coefficient-sign derivative, say x_1 and x_2 with $0 < x_1 < x_0 < x_2$. The following table illustrates the above in a succinct manner.

Function \ Interval	$[0, x_1)$	(x_1, x_0)	(x_0, x_2)	(x_2, ∞)
$f_{cs}^{(k-1)}$	$+$	$+$	$+$	$+$
$f_{cs}^{(k-2)}$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$
$f_{cs}^{(k-2)}$	$-$	$-$	$+$	$+$
$f_{cs}^{(k-3)}$	$\searrow$	$\searrow$	$\nearrow$	$\nearrow$
$f_{cs}^{(k-3)}$	$+$	$-$	$-$	$+$
$f_{cs}^{(k-4)}$	$\nearrow$	$\searrow$	$\searrow$	$\nearrow$

We may continue in a similar manner until we reach f_{cs}^{0} which has the same positive real zeros as f. At each step we are faced with two possibilities, one that corresponds to not having a sign change, which does not change the potential outcome with respect to the previous step, and the more interesting one that does correspond to a sign change and may introduce an additional root under certain assumptions for the value of the corresponding coefficient-sign derivative at a critical point. Each sign change corresponds to a potential change in monotonicity. In the worst case scenario, we obtain as many roots as the number of sign changes, as the bound given by Descartes' rule of signs promises.

The scrupulous reader will wonder how the above corresponds to the traditional method of examining the classical derivatives of the polynomial. Naturally, the above method is not intended to replace the traditional method of employing derivatives to analyze the behavior of a polynomial. Rather, it is designed with the underlying assumptions in mind to achieve the maximum number of roots possible and thus with the purpose of providing an upper bound on the number of positive roots. In particular, it foregoes additional information on their position and exact number, that we could obtain by classical analytical methods. Furthermore, the following observation reconciles the two methods. Consider all derivatives of f and their respective coefficient sequences. Since we only care about their sign and not their magnitude, we may only examine the so-called sign sequences instead. Clearly these sign sequences are subsequences of the sign sequence of f obtained by not considering the appropriate trailing coefficients. The coefficient-sign derivatives correspond to the points where we "drop" an additional trailing coefficient and hence its sign from the sequence. Now assume that we have completely determined the (potential) points where the $(i+1)$-th derivative changes sign,

thus also the monotonicity of the i-th derivative. If the i-th derivative does not have a constant term, the sign sequence will remain the same and in particular the sign of the trailing coefficient will too. Since the two derivatives share the same sign for their trailing sequence, we can deduce that either at the first interval, say $(0, x_0)$, either the i-th derivative is positive and increasing or negative and decreasing. Suppose without loss of generality the first case occurs. In the next interval, assuming it occurs, we know that the monotonicity will change, thus the function will decrease. This will be a potential point of a root that corresponds to the first critical point. By the same reasoning, for each interval where the i-th derivative is monotone, the critical point that forms the left end of the interval corresponds to a potential root. Thus, the potential number of roots of the i-th derivative in the worst case will remain the same as for the $i+1$-th derivative, albeit changing the location of these roots on the real line. Coefficient-sign derivatives abstract this information, since we only care for a bound on the number of real positive roots.

We conclude from the above that the number of roots in the worst case will remain the same as long as the derivatives have the same coefficient sign sequence. This is the point that is essentially simplified in the proof of Theorem 2.17 by assuming the exact number of roots. Furthermore, the only point where additional sign changes and thus potential roots may occur is when we consider an additional coefficient. In summary, the number of sign changes limits the number of changes in monotonicity of the polynomial f and therefore also the number of potential roots.

An additional point is that this process also sheds some light on Corollary 2.18, namely that the number of positive real roots $N_{\mathbb{R}^+}(f)$, the number of sign changes $\#s(f)$, and the number of sign changes of the coefficient subsequence $c_0 c_d$ have the same parity. Assume per the usual assumption that the leading coefficient c_d is positive and the trailing coefficient c_0 is negative. This implies that the parity of the number of sign changes is also odd. Furthermore, since we start with the function being negative, any root must occur in an interval where the function is increasing. At least one such root must occur since the leading and trailing coefficients have opposite signs as we've mentioned in the proof of Corollary 2.18. If a subsequent root occurs, this means that the function becomes negative again and since for large enough x it is positive, one additional root must occur. That is, after the first root, any subsequent roots must occur in pairs. We thus conclude that the parity of the number of positive real roots is odd, same as the parity of the number of sign changes. The same arguments applies for the case of an even number of sign changes, by omitting the initial root, since the trailing coefficient and thus the function near 0 will also be positive.

The first argument above shows that the number of sign changes of the coefficient sequence of f upper bounds the number of monotonicity changes of the function and thus also its real roots. However it provides no information on which intervals these roots may be located, even if we make certain assumptions about the magnitude of the coefficients. Note in particular that by the manner coefficient-sign derivatives were defined, which exponents correspond to nonzero coefficients was irrelevant to the argument and thus the argument was more suited to the study of sparse polynomials in general, where exponents may differ significantly. A different approach would be the following: Assume that a sign change occurs for the term with exponent m. Descartes' rule of signs says that this sign change implies a potential root (there may be further sign changes that affect the parity).

Furthermore, we saw in the proof of Theorem 2.8 that a term with larger exponent will eventually dominate all other terms and determine whether the function is negative or positive. By making certain assumptions about the other terms of the polynomial and the magnitude of the coefficients, can we determine when this term dominates and what does this imply for the roots of the polynomial?

Let us first discuss what assumptions would be logical to make. Recall that as in Theorem 2.17, we only care for positive roots. This allows us to assume that $x > 0$ and greatly simplifies the argumentation. To begin with, consider the possible terms with exponent smaller than m, that is

$$\sum_{i=0}^{m-1} c_i x_i \,.$$

Since we would like to know when the m-term dominates this sum, it makes sense to choose the assumptions that maximize this quantity. Since $x > 0$, clearly any nonpositive coefficient would result in a smaller result. Thus we may assume that all the coefficients are positive, $c_i > 0$ for $0 \leq i \leq m-1$. Also let $c_{max}(j)$ be the absolute value of the largest in magnitude coefficient from the trailing to the j-th coefficient, $c_{max}(j) = \max_{0 \leq i \leq j} \{c_i\}$. We are interested in the largest in magnitude coefficient up to $m-1$, that is $c_{max}(m-1)$. Clearly assuming that $c_i = c_{max}(m-1)$ for all $i < m$ can only increase the value of this sum. Since we wish for a sign change to occur, we must assign to the m-th term a negative coefficient, say $-c_m$, for $c_m > 0$. Finally let r be the ratio of the magnitudes of the absolutely largest preceding coefficient to that of that of the m-th term, that is, $r \coloneqq c_{max}(m-1)/c_m$. Clearly as r becomes smaller, so does the value of x for which we expect the m-th term to dominate. Finally, we assume that terms of higher exponent of the polynomial have the same sign as the m-th term and thus can only assist in dominating the smaller terms. This is to isolate the effect of a single sign change. Since higher order terms can only be beneficial, assuming they do not occur can only delay the point where the m-th term dominates. Finally, note that by the above assumptions, the associated polynomial, say f, has value $f(0) = c_0 > 0$ at $x = 0$ and is negative for large enough x, since no further sign changes occur and thus the leading coefficient must also be of negative sign. A root is then guaranteed to occur and thus with the above assumptions, we wish to determine the value of x, or at least an interval, in which the function first becomes negative, that is

$$-c_m x^m + \sum_{i=0}^{m-1} c_i x_i \leq 0 \iff c_m x^m \geq \sum_{i=0}^{m-1} c_i x_i \iff c_m x^m \geq c_{max}(m-1) \sum_{i=0}^{m-1} x^i \iff$$
$$x^m \geq r \sum_{i=0}^{m-1} x^i \,. \tag{2.10}$$

For $x = 1$, the m-th term can only match the sum in the right hand side of Equation (2.10) if and only if $r \leq 1/m$. Furthermore, if $m = 1$, i.e., if the sum degenerates to the constant term, the degree 1 term dominates if and only if $x \geq r$. Thus for the remainder, we may assume that $x \neq 1$ and $m \geq 2$ to simplify the analysis. Therefore, if $x \in \mathbb{R}^+ \setminus \{1\}$, we can write Equation (2.10) as

$$x^m \geq r\frac{x^m-1}{x-1} \iff \begin{cases} x^{m+1}-(r+1)x^m+r \geq 0\,,\ x>1\,, \\ x^{m+1}-(r+1)x^m+r \leq 0\,,\ x<1\,. \end{cases} \tag{2.11}$$

Therefore it suffices to study the polynomial $g(x) = x^{m+1}-(r+1)x^m+r$ for $x \in \mathbb{R}^+$ and determine when it is negative for $x<1$ and positive for $x>1$. To begin with, we note that $g(0) = r > 0$, $g(1) = 0$ and $g(r) = r - r^m$. The value at $x = r$ changes sign depending on the value of r, namely

$$\begin{cases} g(r) > 0\,,\ r<1\,, \\ g(r) < 0\,,\ r>1\,, \\ g(r) = 0\,,\ r=1\,. \end{cases} \tag{2.12}$$

Furthermore we have $g'(x) = (m+1)x^m-(r+1)mx^{m-1} = x^{m-1}((m+1)x-(r+1)m)$. For $x \in \mathbb{R}^+$, we clearly have that

$$\begin{cases} g'(x) = 0\,,\ x=0 \vee x=\frac{(r+1)m}{m+1}\,, \\ g'(x) < 0\,,\ x<\frac{(r+1)m}{m+1}\,, \\ g'(x) > 0\,,\ x>\frac{(r+1)m}{m+1}\,. \end{cases} \tag{2.13}$$

From the above it follows that g is strictly decreasing in $\left(0, \frac{(r+1)m}{m+1}\right)$ and strictly increasing in $\left(\frac{(r+1)m}{m+1}, \infty\right)$. Therefore the point $x = \frac{(r+1)m}{m+1}$ is always where the minimum of g is achieved in $\mathbb{R}^+$. Since $g(0) > 0$, $g(1) = 0$, and the function is strictly monotone in the two intervals given above, it follows that $f(\frac{(r+1)m}{m+1}) < 0$ except for the corner case when $\frac{(r+1)m}{m+1} = 1 \iff r = \frac{1}{m}$. In that case, when $r = 1/m$, the critical point of g and $x = 1$ coincide and thus g exhibits a double root at $x = 1$. Since we have determined the sign of g for 1, r and $\frac{(r+1)m}{m+1}$, comparing these values is crucial to ensure the analysis is properly carried out. In particular, we have

$$\frac{(r+1)m}{m+1} \geq 1 \iff r \geq \frac{1}{m}\,,$$

$$\frac{(r+1)m}{m+1} \geq r \iff r \leq m\,.$$

Combined with Equation (2.12), that relates the sign of $g(r)$ with the value of r itself, we can distinguish four general cases depending on how the value of r compares to $1/m$, 1 and m and three special cases that correspond to r coinciding with one of these values. We now present tables that correspond to each case and summarize the analysis of the sign and zeros of g via its derivative, along with the necessary clarifications.

We begin with the case $0 < r < 1/m < 1 < m$ (recall that $m \geq 2$ by assumption as the case $m = 1$ was treated separately). We have

Function \ Interval	$[0, r]$	$\left(r, \frac{(r+1)m}{m+1}\right]$	$\left(\frac{(r+1)m}{m+1}, 1\right]$	$(1, \infty)$
$g'(x)$	$0 \quad -$	$- \quad 0$	$+$	
$g(x)$	$\searrow$		$\nearrow$	
$g(x)$	$+ \quad + \quad +$	$+ \quad 0 \quad - \quad -$	$- \quad 0$	$+$

Table 2.1: The analysis for $r < 1/m$

The first row describes the sign changes of $g'(x)$, the second the monotonicity of $g(x)$ and the third row the sign of $g(x)$. The signs (or zero) in superscript denote the sign of the function at the point indicated by the vertical line separator. For example, we see from above that $g'(0) = 0$ and $g(r) > 0$ in this case. A dashed vertical separator which is always accompanied by a zero in superscript indicates that a root occurs in this interval, however we cannot determine its precise location. For example, in the above g must have a root in the interval $\left(r, \frac{(r+1)m}{m+1}\right]$.

As we can see from the above table, for $x > 1$, the m-th term always dominates. This is clear since as we've seen above for the corner case of $x = 1$, the same is true for $r \leq 1/m$, a condition that is essentially the same with the exception of the limit point. Furthermore, for $x < 1$, we see that there exists some value $a \in (r, \frac{(r+1)m}{m+1})$ such that $g(a) = 0$ and in particular, the m-th term dominates in $[a, 1)$. However, for the corner case mentioned above when $r = 1/m$, the situation is somewhat different for $x < 1$, which we now summarize in the following table.

Function \ Interval	$[0, r]$	$\left(r, \frac{(r+1)m}{m+1} = 1\right]$	$(1, \infty)$
$g'(x)$	$0 \quad -$	$- \quad 0$	$+$
$g(x)$	$\searrow$		$\nearrow$
$g(x)$	$+ \quad + \quad +$	$+ \quad 0$	$+$

Table 2.2: The special case $r = 1/m$

It is clear that this case is similar to the one for $r < 1/m$, with the exception that since the minimum occurs for $x = 1$ and $g(1) = 0$, g never becomes negative and thus the m-th term never dominates for $x < 1$.

The remaining cases are very similar, with essentially the only difference being the position of r relative to the other values and thus the sign of $g(r)$, which does not however affect the dominance of the m-th term. Nevertheless, we provide all relevant tables for the benefit of the interested reader.

Function \ Interval	$[0, r]$	$(r, 1]$	$\left(1, \frac{(r+1)m}{m+1}\right]$	$\left(\frac{(r+1)m}{m+1}, \infty\right)$
$g'(x)$	$0 \quad -$	$- \quad 0$	$+$	
$g(x)$	$\searrow$		$\nearrow$	
$g(x)$	$+ \quad + \quad +$	$+ \quad 0 \quad - \quad -$	$- \quad 0 \quad +$	

Table 2.3: The case $1/m < r < 1 < m$

Function \ Interval	$[0, r=1]$	$\left(1, \frac{(r+1)m}{m+1}\right]$	$\left(\frac{(r+1)m}{m+1}, \infty\right)$
$g'(x)$	$0 \quad -$	$- \quad 0$	$+$
$g(x)$	$\searrow$		$\nearrow$
$g(x)$	$+ \quad + \quad 0$	$- \quad -$	$- \quad 0 \quad +$

Table 2.4: The case $r = 1$

Function \ Interval	$[0, 1]$	$(1, r]$	$\left(r, \frac{(r+1)m}{m+1}\right]$	$\left(\frac{(r+1)m}{m+1}, \infty\right)$
$g'(x)$	$0 \quad -$		$- \quad 0$	$+$
$g(x)$	$\searrow$			$\nearrow$
$g(x)$	$+ \quad + \quad 0$	$- \quad -$	$- \quad -$	$- \quad 0 \quad +$

Table 2.5: The case $1 < r < m$

Function \ Interval	$[0, 1]$	$\left(1, \frac{(r+1)m}{m+1} = r\right]$	(r, ∞)
$g'(x)$	$0 \quad -$	$- \quad 0$	$+$
$g(x)$	$\searrow$		$\nearrow$
$g(x)$	$+ \quad + \quad 0$	$- \quad -$	$- \quad 0 \quad +$

Table 2.6: The case $r = m$

Function \ Interval	$[0, 1]$	$\left(1, \frac{(r+1)m}{m+1}\right]$	$\left(\frac{(r+1)m}{m+1}, r\right]$	(r, ∞)
$g'(x)$	$0 \quad -$	$- \quad 0$	$+$	
$g(x)$	$\searrow$		$\nearrow$	
$g(x)$	$+ \quad + \quad 0$	$- \quad -$	$- \quad -$	$- \quad 0 \quad +$

Table 2.7: The case $r > m$

As it is clear from the tables above, for $r > 1/m$, the m-th term never dominates for $x < 1$. It is easy to see why this is the case, since for $x < 1$ a single term of the sum of terms up to $m - 1$ suffices to dominate the m-th term, that is $1 > x > \cdots > x^{m-1} > x^m$. By the condition on r, the coefficient of the m-th term is too small to overcome this. Turning our attention to $x > 1$ under the same condition, we observe that the minimum of g always occurs for $x > 1$, as always at $\frac{(r+1)m}{m+1}$. Therefore the point where the m-th term starts dominating must occur after this point and furthermore at the case $r > m$, when r is larger than the point of minimality and in addition we have $g(r) < 0$, this must thus occur for $x > r$. Finally, by the the monotonicity of g, we can ensure that no other interval may occur where the m-th term dominates, under the aforementioned assumptions.

To summarize, we have the following possibilities for satisfying Equation (2.10), i.e., for the m-th term to dominate: There are two special cases, the first being $m = 1$. The

m-th term then dominates for $x \geq r$. The second special case is for $x = 1$, in which case the m-th term then dominates as long as $r \leq 1/m$. In the general case, we may thus assume $x \neq 1$ and $m \geq 2$, which in turn allows us to derive Equation (2.11) and conclude the following: If $x < 1$, the m-th term dominates only when $r < 1/m$. The m-th term will dominate in an interval $[a, 1)$ for some a such that $r < a < \frac{(r+1)m}{m+1}$. If $x > 1$, we distinguish two cases. If $r \leq 1/m$, the m-th term dominates in the entirety of $(1, \infty)$. However, if $r > 1/m$, then the m-th term only dominates in the interval $[a, \infty)$, where $g(a) = 0$ and $a > \max\{\frac{(r+1)m}{m+1}, r\}$, with r being the maximum value only when $r \geq m$.

The above analysis identified not only where the point in which the m-th term starts to dominate and thus the original polynomial f becomes negative, but also shows that after this point the m-th term will continue to dominate, as we would expect from our remarks in the proof of Theorem 2.8. It remains to associate these findings with the bound given by Theorem 2.17. To begin with, note that we assumed that all terms with exponent smaller than m did occur and with an opposite sign to the m-th one, which may not be the case. Furthermore, we assumed that all coefficients of these terms have the same magnitude that is actually the maximum of their magnitudes. In reality, most polynomials would have varying magnitudes and signs for their coefficients. Nevertheless, as we noted already in the analysis, this would allow the m-th term to dominate even sooner, i.e., the above assumptions form the worst-case scenario. The main point is that even in this case, we can ensure that at most one real root may occur, since there was one sign change, as in the statement of Theorem 2.17. For subsequent sign changes, we can operate under the same "worst case" hypothesis and still guarantee at most one root. This essentially derives the bound given by the Descartes' rule of signs.

A further point is that under our assumptions, we could guarantee that a root would occur, and the analysis was only necessary to clarify the interval where this indeed happens. However we are aware that for many polynomials, a root may not occur despite one or more sign changes. This deviation is due to our final assumption, where we assumed that terms of exponent higher than m have the same sign as the m-th term and thus considering them would only result in the m-th term dominating earlier in the real line. It is clear however by the same analysis, that even a single term of opposite sign to the m-th one, say the $m + 1$-th one, together with a large enough coefficient, can completely negate the dominance of the m-th term. Therefore, this would have the effect of a root not occurring despite the sign change, thus explaining why the Descartes' rule of signs provides a rather weak upper bound, as we will also see below. Furthermore, this provides some justification for why the parity of real roots must match that of sign changes, albeit more obscure than that of the first argument: Since we require an additional sign change to negate the dominance of the m-th term, reducing the number of roots by 1 requires increasing the number of sign changes also by 1, thus they must have the same parity. The argument may then repeated for each actual sign change that occurs in the actual polynomial f.

The two arguments above were presented in great detail and hopefully provided some intuition on why the Descartes' rule of signs is indeed true and its inner workings, which will serve the reader well in the subsequent sections and understanding the much more complicated families of polynomials to be studied. Before we proceed with introducing further necessary terms for presenting our work, we will conclude this section by commenting on how Theorem 2.21 may be used for families of polynomials, comment on how

the rule may be further improved and also present a number of generalizations.

As we have already mentioned, the notion of sparsity k of a polynomial f is central not only for this present work but it is a notion very often used to study polynomials. Nevertheless, Theorem 2.21 provides a bound in terms of the number of sign changes of f, $\#s(f)$. Somewhat paradoxically, the notion of sparsity is both less restrictive and more descriptive compared to that of the number of sign changes. In terms of restriction, the sparsity k only identifies how many terms have non-zero coefficients, without further specifying which exponents they correspond to or what the sign and magnitude of those coefficients are. On the other hand, the number of sign changes immediately implies that at least $\#s(f)$ coefficients must have opposite sign of the remaining ones, while the parity of $\#s(f)$ even determines whether the trailing and leading coefficients will be sharing a sign or not, as we've already seen. In turn, the sign of the leading coefficient determines the behavior of the polynomial at the two points of infinity on the real line, while the sign of the trailing coefficient determines the sign of the polynomial at 0. On the other hand, note that as long as we do not alter the number of sign changes, we may add an arbitrary number of coefficients to a polynomial, potentially greatly altering how it behaves, naturally with some restrictions. Finally, it is difficult both in theory and practice to reason about polynomials that share the same number of sign changes, while the restriction on the sparsity is much more natural.

The above would seem to imply that the Descartes' rule of signs is rather useless when considering families of polynomials. However, one simple observation is sufficient to alter the rule so that it relies on the sparsity of a polynomial rather than the number of sign changes, thus overcoming the above shortcomings when dealing with families of polynomials. That is nothing else than simply noting that the number of sign changes of a polynomial $f, \#s(f)$, can clearly be at most 1 fewer than its sparsity k, i.e., $\#s(f) \leq k-1$. This is trivial by the definition of sign changes, since each sign change requires a pair of subsequent terms of opposite sign and there are $k-1$ such pairs. Therefore we may adapt Theorem 2.17 and in turn Theorem 2.21 accordingly, albeit at the cost of losing any information regarding the parity of the number of real roots as well as weakening the bound further.

Corollary 2.23 (Descartes-based bound on real roots using sparsity)**.** *Let* $f(x) = \sum_{i=0}^{k-1} c_i x^{e_i}$ *be a real polynomial of sparsity* k*. Then the number of positive real roots of* f *counted with multiplicity is at most* $k-1$ *and consequently the number of real roots is at most*

$$N_{\mathbb{R}}(f) \leq 2(k-1) + e_0\,,\ l \in \mathbb{N}\,.$$

Quite often since e_0 can vary significantly for families of polynomials, instead of using the maximal such value, we will omit the zeros at $x = 0$ and instead talk about a bound of $2k-2$ for nonzero real roots, with multiplicity. This allows us to state a bound on the number of real zeros that only depends on the sparsity and thus allows us great freedom in the families of polynomials it can apply to.

As it is clear from the above and we've already mentioned, using the sparsity instead of the sign changes may weaken the upper bound significantly in certain cases, e.g., when no sign changes occur. But how tight is the Descartes' rule of signs to begin with? Can

we hope to improve the bound further in its current form and if not, what are some generalizations that we can make to obtain better bounds?

We start with some observations on the Descartes' rule of signs itself. We start by exhibiting a polynomial with as many roots as sign changes. For example, by the binomial theorem we have

$$(x-1)^d = \sum_{i=0}^{d} \binom{d}{i} (-1)^i x^{d-i}\,.$$

Since binomial coefficients are positive, clearly the above polynomial has the maximum number of sign changes as expressed by the term $(-1)^i$, that is d. Furthermore, clearly, from the left hand side, it has a root at $x = 1$ of multiplicity d, matching the bound given by Theorem 2.17. Note that however this polynomial is dense and its sparsity is the maximum possible. By more careful consideration, one can obtain polynomials that not only match the bound, but also have smaller sparsity. Furthermore, it is possible by appropriate manipulation of the coefficients to obtain any possible number of roots, i.e. any number with the same parity and smaller or equal to the number of sign changes. Such constructions can be found in many works and in particular the ones we describe above can be found in [32].

On the other hand, we are aware that the Descartes' rule of signs is a rather weak bound and one can easily construct examples where despite any number of sign changes, no real roots occur. For example, assume that f is a polynomial of even degree d and that the leading coefficient is positive, $c_d > 0$, ensuring that at both points of infinity the polynomial tends to infinity and in particular is positive, $\lim_{x\to-\infty} f(x) = \lim_{x\to\infty} f(x) = \infty$. Furthermore assume that the trailing coefficient is also positive, $c_0 > 0$. Clearly then the polynomial only has roots if it attains a minimum nonpositive value at some point μ, which nevertheless must be finite. Assume that this is the case and note that the above conditions do not restrict the number of sign changes, expect in that they must be of even parity, since the leading and trailing coefficients share a sign. Now, let $g(x) = f(x) - f(\mu) + 1$. Clearly in comparison to the coefficient sequence of f, that of g only changes for the constant term. Furthermore, the sign of the constant term will remain positive, since $f(\mu) < 0$. Finally, clearly the minimum value of g is now achieved at μ and is $g(\mu) = 1 > 0$, i.e., g has no real roots. By the above, we can have any number of sign changes from 0 to $k-1$, where k is the sparsity of f, while the number of real roots remains zero.

The two statements above show that despite the Descartes' rule of signs being a weak bound, unfortunately any bound that only relies on the number of sign changes of the coefficient sequence cannot be further improved. Nevertheless, a number of generalizations exist, which clearly alter certain assumptions of Theorem 2.17 to walk around the above restriction. We will only focus on univariate polynomials, but nevertheless mention the work of Khovanskii [49], which expands the ideas behind Descartes' rule of signs to the multivariate setting. Returning to the univariate setting, these various "rules" were obtained by mathematicians who closely worked on the Descartes' rule of signs or even provided some of its earlier proofs, since Descartes only stated the rule with no formal proof. The work [5] not only includes these rules due to Budan, Fourier, de Gua and Laguerre, along with their proofs and how they relate to the Descartes' rule of signs, but

also a historical recount of the development of these rules.

We begin with the work of Budan and Fourier, that derived similar theorems that generalize Descartes' rule of signs to any interval $(a, b]$ of the real line. Both theorems are very similar to the point of certain historical controversy and arrive at the same conclusion, albeit using different coefficient sequences that nevertheless share the same number of sign changes. We first state the result due to Budan.

Theorem 2.24 (Budan's rule [14])**.** *Let $f(x) = \sum_{i=0}^{k-1} c_i x^{e_i}$ be a real univariate polynomial and $(a, b]$ be an interval on the real line, $a\,,\ b \in \mathbb{R}$, with $a < b$. Also let $g(x) \coloneqq f(x+a)$ and $h(x) \coloneqq f(x+b)$. Then the number of real roots of f in $(a, b]$ is at most $\#s(g) - \#s(h)$ or smaller by an even integer, that is*

$$N_{(a,b]}(f) + 2l = \#s(g) - \#s(h)\,, \quad \textit{for some } l \in N\,.$$

Budan's rule relates the number of real roots in the interval $(a, b]$ to the number of sign changes of the polynomial f shifted to the endpoints of the interval, $g(x) = f(x + a)$ and $h(x) = f(x + b)$. Without delving into a formal proof, it is easy to see how this stems almost directly from the Descartes' rule of signs. Shifting the polynomial to a point, say the left endpoint $g(x) = f(x + a)$ and applying Theorem 2.17 to g, we obtain $N_{(a,\infty]}(f)+2l = \#s(g)\,,\ l \in \mathbb{N}$. This is easy to see since real positive roots of g correspond to roots of f in $(a, \infty]$. With a similar argument for h we obtain that $N_{(b,\infty]} + 2m = \#s(h)\,,\ m \in \mathbb{N}$. Now we only have to consider that $N_{(a,b]}(f) = N_{(a,\infty]}(f) - N_{(b,\infty]}(f)$ and substitute to essentially obtain the above.

We can even see that Budan's rule is truly a generalization of Descartes' rule of signs, since we may obtain the latter by setting $a = 0$ and b to a large enough value that we now specify. Consider all the derivatives of f from order 0 (i.e., f itself), to the highest order derivative that is not identically zero, that is of order d. It is easy to see that viewed as polynomials, their leading coefficients have the same sign, positive without loss of generality, since they are equal to the leading coefficient of f, c_d, multiplied by some positive factor. Therefore, for large enough x, all of them must be positive. Thus, we may choose b to be such an x so that all the derivatives of f are strictly positive in (b, ∞). Note that $h'(x) = f'(x + b)$ so we have thus also obtained $h(x)$ and all its derivatives. If we now evaluate them all at 0, we obtain their trailing coefficients, which have the same sign as the coefficients of $h(x)$. Specifically, the trailing coefficient of the i-th order derivative has the same sign as the coefficient of exponent i, which may potentially be 0. Furthermore, recall that the sign of a polynomial close to 0 is that of its trailing coefficient and we know that all these derivatives are always positive. It follows that all coefficients of h are positive and thus it exhibits no sign change due to our choice of b, Thus $\#s(h) = 0$ and since $\#s(g) = \#s(f)$ if $a = 0$, we obtain precisely the statement of Descartes' rule of signs.

Fourier's theorem [31] is very similar, the only difference is that it does not use the coefficient sequence of the shifted polynomials at the endpoints, but rather the sequence of derivatives of all orders at these points, that is it counts the number of sign changes in the sequence $f(a), f^{(1)}(a), \ldots, f^d(a)$ and similarly for $f(b), f^{(1)}(b), \ldots, f^d(b)$. The argument we used before for deriving Descartes' rule of signs from Budan's rule already reveals the proof idea. Furthermore, note that the number of sign changes of the first sequence is exactly the same as $\#s(g)$, using the above notation, and similarly for the second

sequence it is equal to $\#s(h)$, see also [5]. This is easy to see by considering the Taylor series of the polynomial at the endpoints of the interval. For example, consider the Taylor sequence of f at $x = a$. We have

$$f(x) = \sum_{i=0}^{d} \frac{f^i(a)}{i!}(x-a)^i \,.$$

Now by shifting the polynomial by a we obtain

$$g(x) = f(x+a) = \sum_{i=0}^{d} \frac{f^i(a)}{i!}x^i \,,$$

from where it follows that the i-th coefficient of $g(x)$ has the same sign as the i-th order derivative of f at $x = a$.

De Gua played an important rule in formalizing the Descartes' rule of signs. Furthermore, his contributions include the following statement, known as De Gua's rule [24]. It indirectly provides an upper bound on the number of real roots, by way of a lower bound on the number of imaginary (i.e., complex non-real) roots. While as demonstrated in [5] this rule is weaker than Theorem 2.21, it can be useful in gaining insights into results, as we will see after its statement.

Lemma 2.25 (De Gua's rule [24])**.** *Let $f(x) = \sum_{i=0}^{k-1} c_i x^{e_i}$ be a real univariate polynomial and let $r := \max_{i=1,\ldots,k-1}\{e_i - e_{i-1} - 1\}$, i.e., the largest number of terms of consecutive exponents with zero coefficient in f. Let k and l, with $0 \leq k \leq l \leq d$ be the exponents delimiting this "gap" of consecutive terms, i.e. we have $c_k \neq 0$, $c_l \neq 0$ and $c_i = 0$, $i = k+1, \ldots, l-1$. Then the number of imaginary roots of f, $N_{\mathbb{C}\setminus\mathbb{R}}(f)$ is bounded from below as follows*

$$\begin{cases} N_{\mathbb{C}\setminus\mathbb{R}}(f) \geq r \,,\ r = 2m \,,\ m \in \mathbb{N} \,, \\ N_{\mathbb{C}\setminus\mathbb{R}}(f) \geq r+1 \,,\ (r = 2m+1) \wedge (c_k c_l > 0) \ ,\ m \in \mathbb{N} \,, \\ N_{\mathbb{C}\setminus\mathbb{R}}(f) \geq r-1 \,,\ (r = 2m+1) \wedge (c_k c_l < 0) \ ,\ m \in \mathbb{N} \,. \end{cases}$$

While the above rule is interesting enough on its own right to be mentioned, given the unique approach of providing a lower bound for roots in $\mathbb{C} \setminus \mathbb{R}$ rather than an upper bound, it is possible that it can be quite useful in furthering our understanding in various cases. In particular, the above rule is useful in the study of sparse polynomials that have the same sparsity but have different exponents corresponding to nonzero coefficients. Specifically, we may keep the sparsity k constant but modify r, the size of the largest gap mentioned in Lemma 2.25. We can thus limit the number of real zeros of a polynomial by utilizing the above lemma and the usual degree bound on the total number of roots by Theorem 2.8. Together with additional structure on the polynomial, this can also be useful in constructing examples whose number of roots can be manipulated to be quite precise.

For example, consider a dense polynomial that has degree m and all its roots are distinct, real and positive, say $f(x) = \prod_{i=0}^{m-1}(x - x_i)$. By the condition on the roots and Theorem 2.17, it follows that all coefficients must be nonzero and alternating in sign. Now we may obtain a new polynomial g by multiplying f with x^{r+1}, where r is

odd and corresponds to the desired gap size as in the Lemma above. Finally we add a new constant term of arbitrarily small magnitude. In particular, we choose the constant term to have the same sign as the trailing coefficient of f, say positive without loss of generality. We thus set $g_0 = \varepsilon > 0$ to be small enough so that there exists $\delta > 0$ so that in each interval $[x_i - \delta, x_i + \delta]$, $i = 0, \ldots, m-1$ there exists exactly one point x_i' so that $(x_i')^{r+1} f(x_i') = -\varepsilon$. For this to be true, by the continuity of polynomials, it is sufficient that in the above intervals for x, $f(x)$ contains no turning point. Now, by applying the above argument to $g(x) = x^{r+1} f(x) + \varepsilon$, we see that g also has at least m positive real roots. Furthermore, by Lemma 2.25, in particular the case where r is odd and the coefficients delimiting the gap share the same sign, g must have at least $r+1$ imaginary roots and since $\deg(g) = m + r + 1$, all roots have been accounted for. This technique can be used in greater generality for families of polynomials, by choosing ε appropriately.

We close this section with a number of generalizations of Descartes' rule of signs due to Laguerre. We follow the original work [56], simplified both by our own contributions as well as certain remarks found in [5]. The first statement generalizes the Descartes' rule of signs to so-called "extended" polynomials. These differ from run-of-the-mill polynomials by allowing the exponents to be any real number instead of only natural numbers. We have the following statement

Lemma 2.26 (Descartes' rule for extended polynomials)**.** *Let* $f(x) = \sum_{i=0}^{k-1} c_i x^{e_i}$*, where* $c_i, e_i \in \mathbb{R}$ *for* $0 \leq i \leq k-1$ *and* $e_{i-1} < e_i$ *for* $i = 1, \ldots, k-1$*. Also let* $\#s(f)$ *be the number of sign changes in the coefficient sequence of* f*. Then similar to Theorem 2.17, the number of positive real roots of* f *counted with multiplicity is at most* $\#s(f)$*, that is*

$$N_{\mathbb{R}^+}(f) \leq \#s(f)\,.$$

Proof. This proof clearly also applies in the special case of polynomials, thus is an alternative proof of Descartes' rule of signs, due to Laguerre. The proof is much shorter and although it obscures even more intuitive details, it is remarkably simple and short. The proof proceeds by induction on the number of sign changes $\#s(f)$. For the base case, $\#s(f) = 0$, note that this implies that all coefficients of f must have the same sign, without loss of generality positive. It is now easy to see that since $x^i > 0$ for $x > 0$, it follows immediately that $f(x) > 0$ for all $x \in \mathbb{R}^+$, thus the number of positive real roots is 0 and the statement is true.

Now let us consider the induction case and assume the statement is true for $\#s(f) \leq m-1$. Assume that f has $\#s(f) = m$ sign changes. Let $a \in \mathbb{R}$, whose value will be specified later, and note that $x^{-a} f(x)$ has the same positive real zeros as f. Furthermore, recall the statement of Rolle's theorem: if $h(x)$ is continuous in $[l, u]$, differentiable in (l, u) and $f(l) = f(u)$ with $l < u$, there must exist $t \in (l, u)$ such that $h'(t) = 0$. Clearly extended polynomials satisfy these conditions in $\mathbb{R}^+$, in fact in $\mathbb{R} \setminus \{0\}$, therefore we can derive the following statements:

- If h is an extended polynomial and has a root of multiplicity n_i at $x = x_i$, then h' has a root of multiplicity of $n_i - 1$ at the same point.

- By Rolle's theorem, for any two distinct roots x_i and x_j of h, there exists a root c of h'.

Now consider the derivative of $x^{-a}f(x)$. Recall that $x^{-a}f(x)$ has $N_{\mathbb{R}^+}(f)$ positive real zeros. We have

$$\left(x^{-a}f(x)\right)' = x^{-a}f'(x) - ax^{-a-1}f(x) = x^{-a-1}\left(xf'(x) - af(x)\right) .$$

By the above observations and since $x^{-a-1} > 0$ for $x > 0$, it suffices to examine the zeros of $g(x) \coloneqq xf'(x) - af(x)$. Assume that $x_1, \ldots, x_D$ are the distinct real positive roots of $f(x)$ with respective multiplicities $m_1, \ldots, m_D$. Clearly we have $N_{\mathbb{R}^+}(f) = N = \sum_{i=1}^{N} m_i$. Knowing the number of roots of $f(x)$ and by the above observations relating the zeros of an extended polynomial and its derivative, we can conclude the following for the number of zeros of g

$$N_{\mathbb{R}^+}(g) \geq \sum_{i=1}^{N} (m_i - 1) + N - 1 = \sum_{i=1}^{N} m_i - 1 = N - 1 .$$

Furthermore, we may use the above expression to obtain the coefficient sequence of g

$$g(x) = xf'(x) - af(x) = \sum_{i=0}^{k-1} e_i c_i x^{e_i} - \sum_{i=0}^{k-1} a c_i x^{e_i} = \sum_{i=0}^{k-1} c_i \left(e_i - a\right) x^{e_i} .$$

From the above expression, it is clear that the choice of a will affect the sign of the coefficients of g, namely for all $e_i < a$ the sign of the coefficient of g will be opposite to that of f and for $e_i > a$ it will be the same. Now, note that since we are in the inductive case at least one sign change must occur in the coefficient sequence of f, say for $i = l$, i.e, $c_l c_{l+1} < 0$. Now choose a so that $e_l < a < e_{l+1}$. This will have the following effect on the sign of the coefficients of g:

- For $i < l$, all coefficients of g will have the opposite sign of those of f. However, since all coefficients change sign, any sign change will remain, just with the order the signs appear reversed.
- For $i \geq l + 1$, all coefficients of g have the same sign as those of f. Thus clearly, any sign changes will once again persist.
- For $i = l$, a sign change occurs for f. But for the coefficient of g, we have that $c_l c_{l+1} < 0$ and $(e_l - a)(e_{l+1} - a) < 0$, therefore $\left(c_l(e_l - a)\right)\left(c_{l+1}(e_{l+1} - a)\right) > 0$ and thus no sign change occurs.

From the above, it follows that for this choice of a, g exhibits exactly $m - 1$ sign changes. Therefore, the induction hypothesis applies and since we know g has at least $N - 1$ real positive roots, we conclude that

$$N - 1 \leq N_{\mathbb{R}^+}(g) \leq m - 1 \Longrightarrow N \leq m .$$

But recall that $N = N_{\mathbb{R}^+}(f)$ is exactly the number of real zeros of f and $m = \#s(f)$ is the number of sign changes of f, therefore the statement holds in the inductive case for m sign changes and the proof is concluded. □

Note that we may generalize the above statement further in the following way. We may assume that f is an infinite series, which can be thought as having infinite sparsity, however having only a finite number of sign changes $\#s(f)$.

Corollary 2.27 (Descartes' rule for infinite series [56])**.** *Let* $f(x) = \sum_{i=0}^{\infty} c_i x^{e_i}$ *be an infinite series where* $c_i, e_i \in \mathbb{R}$ *for all* $i \in \mathbb{N}$*. Let* $\#s(f) < \infty$ *be the number of* finite *sign changes of the sequence* c_i *of coefficients of* f*. Also let* I *be the interval of convergence of* $f(x)$*. Then Lemma 2.26 applies and in particular the number of real positive roots of* f *in* I *is at most* $\#s(f)$*, that is*

$$N_{I\cap\mathbb{R}^+}(f) \leq \#s(f)\,.$$

Proof. We will first make some assumptions and show that they do not result in a loss of generality. To begin with, note that we may assume that the exponents e_i appear in ascending order, that is $e_i < e_{i+1}$. This does not result in a loss of generalization since we can always rearrange the terms of the series and rename the e_i appropriately. Furthermore, we may assume that there exists $l \in \mathbb{N}$ such that $e_l > 0$. Note that coupled with the previous assumption, this implies in fact an infinite number of exponents will be positive, i.e., $e_i > 0$ for all $i \geq l$. If that is not the case, we may always multiply with an appropriate power x^a for $a \in \mathbb{R}^+$ to ensure the assumption holds. Both transformations do not affect the number of sign changes of the coefficient sequence or the number of positive real zeros.

Under the above assumptions, we can now make two crucial observations about f. The first is that since the number of sign changes $\#s(f)$ is finite, there must exist $l \in \mathbb{N}$ such that all coefficients of index $i \geq l$ must share the same sign. Furthermore since $x \in \mathbb{R}^+$ and by our assumptions the leading infinite terms must eventually have positive exponents, by a similar argumentation as in the proof of Theorem 2.8, these leading terms will eventually dominate the partial sum of terms with index $0 \leq i < l$, therefore there exists some $x \in \mathbb{R}^+$ so that f diverges. Moreover, we notice the following fact about the existence of roots of f. If no roots exist, then the statement is satisfied always since $\#s(f) \geq 0$. On the other hands if roots do exist, clearly f must be convergent with limit 0 at these points. Furthermore, considering that f is continuous, that $x \in \mathbb{R}^+$ and that by our above assumptions the exponents e_i are increasing, it follows that there exists some $x_0 \in \mathbb{R}^+$ so that f is convergent in $(0, x_0)$. Finally, once the leading terms dominate, this will remain the case for all $x \geq x_0$, so f is divergent in $[x_0, \infty)$. It follows that it suffices to study the zeros of f in the interval $(0, x_0)$.

With the above assumptions we see that f satisfies the assumptions of Rolle's theorem in $(0, x_0)$ and more generally, every condition of the proof Lemma 2.26 is satisfied and thus we may apply the same proof to this statement. Informally, this is similar to regarding f as an extended polynomial by regarding the leading terms that share the same sign as a single term with that sign. □

While the above Corollary is interesting in its own right as a generalization of Descartes' rule of signs, it is also useful in proving a much broader generalization that applies to polynomials. This generalization is also due to Laguerre and has unfortunately fallen into obscurity. While not as simple to state as Descartes' rule of signs, it is

still rather straightforward and in the opinion of the writer it has the potential to be utilized in current research, completely supersedes Descartes' result and is also flexible in its application. We hope that providing the statement here in detail, translating and modernizing the arguments from the original 19th century French text will help this statement to gain the popularity it deserves.

The idea behind the result is rather simple: Instead of the sequence of coefficients, we derive another sequence from the polynomial, which in fact is an infinite series. Then, we may employ the aforementioned Corollary 2.27 to obtain a bound on the number of roots. A caveat is that the bound applies to all roots larger than some $a \in \mathbb{R}^+$ and for $a = 0$, the theorem is identical to Descartes' rule of signs. However, we will show how to circumvent this and obtain a bound on all positive roots. We begin with the statement of the general Theorem, which we will refer to as Laguerre's rule of signs. Note that a number of results are referred with the same name in several contexts, which are not relevant to us in this work. More generally, the identifier "for polynomials" may be appended to avoid any ambiguity.

Theorem 2.28 (Laguerre's rule of signs [56])**.** *Let $f(x) = \sum_{i=0}^{d} c_i x^i$ be a univariate real polynomial and let $a \in \mathbb{R}^+ \cup \{0\}$ be a nonnegative real number. Consider the following sequence of polynomials derived from f*

$$\begin{cases} f_0(x) &= f(x) = \sum_{i=0}^{d} c_i x^i \,, \\ f_1(x) &= \sum_{i=1}^{d} c_i x^{i-1} \,, \\ &\vdots \\ f_m(x) &= \sum_{i=m}^{d} c_i x^{i-m} \,, \\ &\vdots \\ f_d(x) &= c_d \,. \end{cases} \tag{2.14}$$

Note that in particular we may also obtain this sequence by the recursive relation $f_i(x) = x f_{i+1}(x) + c_i$ for $i = 0, \ldots, d-1$ and $f_d(x) = c_d$ as above. Consider the above sequence of polynomials evaluated at a, that is $f_d(a), \ldots, f_m(a), \ldots, f(a)$ and let $\#sl(f)$ be the number of sign changes exhibited by this sequence. Then the number of positive real roots of f greater than a are bounded by $\#sl(f)$, that is

$$N_{(a,\infty)}(f) \leq \#sl(f) \,.$$

Proof. Consider the polynomial division of $f(x)$ by the polynomial $(x - a)$. We have for $x \neq a$

$$\frac{f(x)}{x-a} = f_d(a)x^{d-1} + f_{d-1}(a)x^{d-2} + \cdots + f_2(a)x + f_1(a) + \frac{f(a)}{x-a} \,.$$

Now note that for $x > a \iff \frac{a}{x} < 1$ the trailing term of the above expression can be expressed as the following convergent series

$$f(a)\frac{1}{x-a} = f(a)x^{-1}\frac{1}{1-\frac{a}{x}} = f(a)x^{-1}\sum_{i=0}^{\infty}\left(\frac{a}{x}\right)^i = \sum_{i=0}^{\infty} f(a)a^i x^{-(i+1)}\,.$$

By substituting in the above expression, we now obtain the following infinite series

$$\frac{f(x)}{x-a} = f_d(a)x^{d-1} + f_{d-1}(a)x^{d-2} + \cdots + f_2(a)x + f_1(a) + f(a)x^{-1} + f(a)ax^{-2} + \ldots\,.$$

Note that all terms that correspond to negative powers of x have the same sign, that of $f(a)$, if $a > 0$. In the corner case $a = 0$, those are all 0. Therefore it is clear that the above series can only have a finite number of sign changes. Furthermore, its sequence of coefficients is precisely $f_d(a), \ldots, f_m(a), \ldots, f(a), \ldots$ and thus exhibits precisely $\#sl(f)$ sign changes. Finally, the series by construction is convergent for $x > a$. Clearly, the conditions of Corollary 2.27 are satisfied and by its application, the proof is concluded. □

We now comment on the result. To begin with, a hasty examination might leave the reader considering that in the worst case we may obtain bounds equal to the degree, since this is the number of coefficients examined. However, if one pays close attention to the recursive relation for the sequence, $f_i(a) = af_{i+1}(a) + c_i$, we note that if the i-th coefficient of f is zero, $c_i = 0$, then $f_i(a)$ will have the same sign as $f_{i+1}(a)$. Therefore, similar to Descartes' rule of signs, the number of sign changes is bounded by $k-1$, where k is the sparsity of the polynomial.

An additional point is that we may easily obtain a similar bound for the positive roots of f that are small than a, i.e., $N_{(0,a)}(f)$. The proof is almost identical to that of Theorem 2.28, except that we replace the trailing term with the following series, which converges for $0 < x < a$ (in fact for $x \leq |a|$, but as stated, Corollary 2.27 holds only in $\mathbb{R}^+$).

$$f(a)\frac{1}{x-a} = -f(a)a^{-1}\frac{1}{1-\frac{x}{a}} = -f(a)a^{-1}\sum_{i=0}^{\infty}\left(\frac{x}{a}\right)^i = -\sum_{i=0}^{\infty} f(a)x^i a^{-(i+1)}\,.$$

The only difference is that the very first coefficient is $-f(a)$ instead of $f(a)$, which may affect the number of sign changes in the sequence by at most 1. However, it should also be noted that for all the above generalizations, similar statements to Descartes' rule of signs can be made with regards to the parity of the number of roots, a fact we omitted for clarity, since it is of little use in our studies. That is, in all stated results in this section, the number of roots has the same parity as the number of sign changes. In that light, the two bounds, for positive roots smaller and greater than some a, are the same.

Furthermore, it is clear from the above that for $a = 0$, we obtain precisely Descartes' rule of signs. However, the benefits are quickly realized for other values of a, since we obtain a sequence of coefficients that for its i-th term depends on all coefficients except the $i-1$ trailing ones and also on the value of a. Compare this to applying Theorem 2.17 to the shifted polynomial $f(x+a)$, where by applying the Binomial Theorem we obtain

$$f(x+a) = \sum_{i=0}^{d} c_i \sum_{j=0}^{i}\binom{j}{i}a^{i-j}x^j\,.$$

Consider the coefficient sequence of the shifted polynomial compared to the sequence used by Theorem 2.28. While both sequences are very likely to be dense, in the sense that most terms of the sequence will be nonzero, note that the recursive relation mentioned in the statement of Laguerre's rule allows us to obtain the coefficients recursively using only two operations, one addition and one multiplication. On the other hand, the shifted polynomial involves the computation of many binomial coefficients and does not seem to have such an elegant recursive expression. Therefore, we can expect to be computationally slower. Furthermore, note that by the use of the aforementioned binomial coefficients, the shifted polynomial method gives more weight to coefficients equidistant from 0 and the degree d of the polynomial in index, since these would involve binomial coefficients of much larger value such as for example $\binom{d}{d/2}$. It would be interesting to investigate how the two rules compare in terms of providing an upper bound, however such precise calculations are outside the scope of this work, since we rarely know the coefficients precisely.

An interesting case arises when $a = 1$. Then $f_i(a) = f_i(1)$ is simply the sum of coefficients of exponent equal or greater to i. Furthermore, in this case, we may obtain a bound on the number of roots smaller than 1 not by the method detailed above but by examining the polynomial $g(x) \coloneqq f(1/x)$ instead. Clearly, for every root x_0 of f with $x_0 > 1$, g has $1/x_0$ as a root which is indeed greater than 1. Thus, by applying Theorem 2.28 to g for $a = 1$, we obtain a bound that relies on the sequence $g_i(1)$, where it is easy to see that $g_i(1)$ is simply the sum of the i trailing coefficients of f. Finally, it is easy to check whether $f(1) = 0$ and even determine the multiplicity of a root at $x = 1$ by identifying the largest m such that $(x-1)^m$ divides f. The two bounds combined allows us to obtain a bound on the number of positive zeros of f, similar to Descartes' rule of signs. At first glance this bound seems weaker since it could be as high as two times that of Descartes' rule of signs in the worst case. However, note that in these bounds, not only the sign of the coefficients plays a role but also their magnitude, since terms of the sequence concern partial sums of the coefficients. Therefore, coefficients of large magnitude might even result in no sign changes in this sequence, a fact that would never be taken into account when relying on Theorem 2.17. Such behavior is close to the behavior of the polynomial, where indeed the magnitude of the coefficients clearly affects its sign and roots, for example setting the trailing and leading coefficient to be of the same sign and large enough magnitude would lead to them dominating the sequence and would result in the polynomial having no real positive roots. It is clear that this technique is of great advantage when we can put significant restrictions on the coefficients. For example, if all coefficients share the same magnitude, a condition that might seem too restrictive but corresponds to an important case that appears in the literature and we will mention in subsequent sections, whether a sign change occurs in the sequence depends on the number of positive against the number of negative coefficients encountered thus far.

We hope that the above brief exploration of more obscure results on bounds of real roots have left the reader informed and curious to further explore the topic. Indeed, Laguerre in his work [56] has many more results of interest. It is possible that these contain results that although are useful for current research endeavors, they have so far been disregarded. This is due to a variety of reasons including that to the best of our knowledge his works have only been published in the original French, including reprints

about a century later. It is also the commitment of the author to closely examine this work for theorems that can be used as tools in the area of research that contains the present work.

In this section, we established in great detail the fundamentals of polynomials, especially regarding those relevant to the study of their roots. We presented the most important historical results in this subject, which naturally also allowed us to shift our attentions to sparse polynomials and how sparsity is a much better indicator of the number of real roots via results such as Descartes' rule of signs and the generalizations we presented. In the next section, we present another key pillar of our study, that of random polynomials. We will examine how their number of zeros may be studied and present upper bounds on their number of roots reminiscent of the Fundamental Theorem of Algebra, in the sense that these bounds rely on the degree d of the polynomial rather than its sparsity.

2.2 Random polynomials and root counting in expectation

2.2.1 Fundamental notions in probability theory

In the previous section, we introduced the basic terminology and notation for polynomials, especially in the setting of root counting. We also went into great detail into some of the most important and classical theorems of root counting for real univariate polynomials, specifically the Fundamental Theorem of Algebra we stated in Theorem 2.8 that shows the number of complex roots of a polynomial is exactly its degree d and thus also a bound on the real roots, as well as the Descartes' rule of signs stated in Theorem 2.17 that provides an upper bound on the number of real roots of a polynomial in terms of its number of variations $\#s(f)$ and thus also its sparsity k.

Quite often both in practice and in theory, we care more about a class of polynomials, whether those are classes defined by some property such as their sparsity k as in our study or for example the set of all possible inputs to an algorithm in the context of a specific application. While the above theorems are stated for a single polynomial, they can still be of use. Clearly, if we wish to identify the worst-case scenario, that is an upper (respectively, lower) bound that holds for each polynomial in the class, it is sufficient to provide a bound with the polynomial with the most (respectively, least) real roots in the class. We may then use the aforementioned theorems to derive the needed bound. Essentially, providing an upper bound for all polynomials in the class reduces to providing a bound for a single polynomial.

While in principle the above seems straightforward, in reality finding such a polynomial or even the additional conditions such a "worst-case" polynomial must satisfy is a complicated task if even possible. It is worth noting that more often than not, these classes of polynomials are of infinite cardinality and even if finite, the numbers involved are prohibitively large for any sort of exhaustive search. For example, imagine a toy example of polynomials of degree 10 with the leading coefficient fixed to 1 and the remaining 10 coefficients allowed to attain any of 10 arbitrary values. This small example still corresponds to a family of 10^{10}, i.e., 10 billion polynomials. Furthermore, even in theory, we impose additional conditions and structure on the polynomials, either to simplify their study or due to the question of interest, while in practice such conditions exist but

are often unknown and hard to specify. Therefore, the bounds obtained by theorems that only rely on general quantities such as the degree or the sparsity of the polynomial can deviate quite far from the actual answer. We have already seen in Section 2.1 that it is possible to construct polynomials with a fixed number of sign changes in their coefficient sequence $\#s(f)$ whose number of real roots can vary to any number from 0 to $\#s(f)$ of the same parity.

In addition, it can be that the worst-case scenario is far more restrictive than we require. For example, consider the class of all degree 1 real polynomials, that all have exactly 1 real root. Now consider extending the family by a single polynomial of the form $f(x) = \prod_{i=0}^{d-1}(x - x_i)$ for $x_i \in \mathbb{R}$ for some $d \in N$ that we consider to be large enough. Clearly this polynomial has d roots and thus in the worst-case, the number of real roots of polynomials in the class is bounded by d, a bound matched by Lemma 2.4. However, it is clear that this bound is a poor representation for the number of roots of the polynomials of this class, excluding the added polynomial f. This can be significant both in theoretical concerns, where we wish to understand the behavior of the class in its entirety rather than a few outliers in the class in terms of the number of roots, as well as in practical concerns, where it is very likely for example that such outliers are much less likely to appear during normal operations.

In such cases, it is useful to consider the *average-case* scenario instead. That is, we consider essentially the *weighted average* number of real roots for the polynomials in the class. Considering the above example, we would intuitively expect this number to be 1. Even this small example however, showcases that it is rather haphazard to consider such weighted averages over classes of infinite cardinality. It is clear that a robust theory is required to ensure the validity of the outcome, as well as specify several aspects that seem unclear: What types of weight functions should be allowed? How is the average to be computed? How to deal with classes of infinite cardinality and how their treatment compares to those of finite cardinality? Thankfully, such questions have already been answered in the well grounded Theory of Probability, specifically its rigorous treatment through Measure Theory. Investigating even the basic concepts and rigorous grounding of the theory occupies several chapters of any textbook on the subject and is thus quite outside the scope of our work. Rather, we introduce here the basic concepts this work relies upon and necessary to understand its scope. We do so in a synoptic manner and highly recommend the interested reader to examine a relevant book on the subject. Specifically, we rely on the rigorous treatment detailed in the book [1] by Ash and Doléans-Dade, which contains most definitions below, with certain alterations for style.

We begin by considering the intuitive treatment of discrete probability that is taught to new students of the subject and then expand the associated concepts accordingly. We consider some random experiment with a finite number of possible outcomes, such as rolling a die or tossing a coin. Call the set of all possible outcomes Ω. Now, we may consider any subset of outcomes, that is any subset $A \subseteq \Omega$. We call each such subset an *event*. The question is then how likely is the event A to occur or in the language of the subject, what is the probability of A occurring. In the simplest of cases such as the throw of a fair die (i.e., all sides have the same chance to be face-up), this is as simple as computing $\Pr(A) = |A|/|\Omega|$, with other weighted functions on the outcome not being particularly difficult to generalize to. Likewise, we typically normalize the weighted function assigned to all possible outcomes so that the probability of the entire space Ω, i.e.

any outcome occurring, is 1. Furthermore, supposing we know the probability of an event A, it should be straightforward to determine the probability of its complement A^{c}, which we would expect to be $1-\Pr(A)$. Similarly, it would be sane that an event A described as the union of *disjoint* events, say $A = \bigcup_{i=1}^{n} A_i$, should have a probability of occurring that is equal to the sum of probabilities of the constituent events, i.e., $\Pr(A) = \sum_{i=1}^{n} \Pr(A_i)$. These axioms are essentially the ones necessary to build up the theory of discrete probability.

We can generalize these concepts to include infinite sets, albeit with a much higher level of formality. Infinite discrete sets can be treated with small alterations to the above axioms, namely using countable many unions instead of finitely many. The situation is rather different for continuous sample spaces Ω. Throughout the following section, we may consider $\Omega = \mathbb{R}$ as the exemplar of such sets. Unlike the discrete case, in the continuous case the structure of subsets of Ω can be quite complicated. We can nevertheless agree that simple subsets of say $\mathbb{R}$, such as every finite interval $[a, b] \subset \mathbb{R}$ should be able to be assigned a probability as an event. For example, such a natural weight would be the length of the interval $b - a$. However, if we wish to replicate the axioms above, it is not sufficient to include just all finite intervals of $\mathbb{R}$. It is easy to see that the set of all finite intervals of $\mathbb{R}$ is not closed under complement or countable union, making it unsuitable to generalize the notions of discrete probability to. We thus first define appropriate sets of subsets, which by definition satisfy these closure conditions.

Definition 2.29 (σ-field)**.** *Let $\mathcal{F}$ be a collection of subsets of a set Ω. Then $\mathcal{F}$ is called a σ-field if and only if*

(1)
$$\Omega \in \mathcal{F}\,,$$

(2)
$$\forall A \in \mathcal{F}\,,\; A^{\mathsf{c}} \in \mathcal{F}\,,$$

(3)
$$\forall A_i \in \mathcal{F}\,,\; i = 1, 2, \cdots \Longrightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}\,.$$

That is, the collection $\mathcal{F}$ is non-empty and closed under complement and countable union.

We note that often σ-fields are also called σ-algebras, here we use the former name in line with the cited source. Returning to the example of $\Omega = \mathbb{R}$, if $\mathcal{F}$ is the collection of all intervals $[a, b] \subset \mathbb{R}$, the smallest σ-field containing $\mathcal{F}$ is called the *Borel* field of $\mathbb{R}$. The Borel σ-field of the reals is the standard structure probability measures are defined over, not only due to fulfilling the axioms we detailed above, but also due to a number of desirable properties, see [1] for more details. It can be assumed throughout this work that any event that we are concerned with is indeed an element of a *Borel* field of $\mathbb{R}$ or more generally R^n for some $n \in \mathbb{N}^+$. Such events correspond to sets that we naturally call Borel sets.

Having defined the collection of subsets we are interested in, the next step is to now define what functions we may use to consistently assign weights to these subsets.

Definition 2.30 (Probability measure)**.** *A measure on a σ-field is a function $\mu : \mathcal{F} \to \overline{\mathbb{R}}$ that satisfies the following conditions.*

(1)

$$\forall A \in \mathcal{F}\,,\ \mu(A) \geq 0\,,$$

(2)

$$\forall A_i \in \mathcal{F}\,,\ i = 1, 2, \ldots \text{ such that } A_i \cap A_j = \emptyset\,,\ \forall i \neq j \Longrightarrow \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu\left(A_i\right)\,.$$

If in addition we demand that $\mu(\Omega) = 1$, we call μ a probability measure. Furthermore, if $A \in \mathcal{F}$ and $\mu\left(A^{\mathsf{c}}\right) = 0$, we say that μ is concentrated on A.

It is also useful to state the following basic properties of measures. Let μ be any (probability) measure, then the following hold.

(1)

$$\mu(\emptyset) = 0\,,$$

(2)

$$\forall A, B \in \mathcal{F}\,,\ \mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)\,,$$

(3)

$$\forall A, B \in \mathcal{F} \text{ such that } B \subset A\,,\ \mu(A) = \mu(B) + \mu(B \setminus A)\,,$$

(4)

$$\forall A_i \in \mathcal{F}\,,\ i = 1, 2, \ldots\,,\ \mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu\left(A_i\right)\,. \tag{2.15}$$

The property given in Equation (2.15) is often referred to as the (countable) union bound.

Equipped with the above, we may define *probability spaces*, which are simply a triple $(\Omega, \mathcal{F}, \mu)$ such that Ω is a set, $\mathcal{F}$ a σ-field of Ω and μ a probability measure defined on $\mathcal{F}$. Similarly we may define measure spaces.

We may now define appropriate measures over the Borel σ-field of $\mathbb{R}$, which we denote by $\mathcal{B}(\mathbb{R})$, which will be the basis to construct probability distributions over $\mathbb{R}$ and in extension, $\mathbb{R}^n$ for $n \in \mathbb{N}^+$. To begin with, we define the first notion that should seem familiar to anyone familiar with continuous distributions.

Definition 2.31 (Distribution function)**.** *A Lebesgue-Stieltjes measure on $\mathbb{R}$ is a measure μ over $\mathcal{B}(\mathbb{R})$ such that $\mu(I) < \infty$ for each bounded interval $I = [a, b]$. A* distribution function *on $\mathbb{R}$ is a map $F : \mathbb{R} \to \mathbb{R}$ that is* increasing*, i.e., $a < b \Longrightarrow F(a) \leq F(b)$, and right-continuous, i.e., $\lim_{x \to x_0^+} F(x) = F(x_0)$. Furthermore, there is a one-to-one correspondence between Lebesgue-Stieltjes measures and distribution functions up to a constant difference, via the formula $\mu\left((a, b]\right) = F(b) - F(a)$.*

Similarly we may define distribution functions for R^n, see [1, Section 1.4] for details. The only part of the puzzle now missing is the use of integration, specifically Lebesgue integrals which allow us to deal with cases where the classical Riemann integral is not well-defined, see [1, Section 1.5] for a complete treatment. Now, we may employ the Radon-Nikodym theorem, to relate two measures via integration. We first recall the theorem's statement.

Theorem 2.32 (Radon-Nikodym, Theorem 2.2.1 in [1])**.** *Let μ be a σ-finite measure and λ a signed measure on the σ-field $\mathcal{F}$ of subsets of Ω. Assume that λ is absolutely continuous with respect to μ. Then there is a* Borel measurable *function $g : \Omega \to \overline{R}$ such that*

$$\lambda(A) = \int_A g \, d\mu \, , \forall A \in \mathcal{F} \, .$$

The statement of the theorem contains notions that we have not precisely defined but can be found on the reference. For the purposes of this work however, it is sufficient for the reader not familiar with measure theory to understand that the Radon-Nikodym theorem allows us to use some well-behaved with respect to Borel sets function g and the Lebesgue measure μ, which for example over $\mathbb{R}$ corresponds intuitively to length, to derive some probability measure λ, by integrating the function g with respect to the Lebesgue measure. This will allow us to properly define probability distributions, using the machinery we have introduced so far.

A *probability distribution* is simply a probability measure defined over $\mathbb{R}$ or more generally, $\mathbb{R}^n$ for $n \in \mathbb{N}^+$. We may distinguish between two kinds of distributions: *discrete distributions* are concentrated over a countably infinite or potentially finite Borel set A of $\mathbb{R}$, i.e., $\mu(A) = 1$. On the other hand, *continuous* distributions are concentrated in an uncountably infinite set $A \subseteq \mathbb{R}$, which could also be the entirety of $\mathbb{R}$. We may make a further distinction for continuous distributions: If the set A that the distribution is concentrated on has Lebesgue measure 0, intuitively meaning that its length is can be made arbitrarily small, then we say that the distribution is singular continuous or simply singular. Such distributions are quite exotic and rarely encountered in practice, an example is the Cantor distribution. In our work such distributions are of little interest, thus we will not refer to them any longer. However, they play an important role, as the other type of continuous distributions are the *absolutely continuous* distributions. This is a consequence of Lebesgue's decomposition theorem [1, Theorem 2.2.6], which states that Lebesgue-Stieltjes measures such as those that we consider in our work can be decomposed into a sum consisting of a discrete measure, a singular continuous measure and an absolutely continuous measure. In the following, we will only consider either discrete or absolutely continuous measures, excluding singular measures or any other combination, i.e., probability mixtures.

We have already described the axioms for discrete probability places. Of more interest is how we may define absolute continuous distributions, to which we will refer in the remaining simply as continuous distributions. Since the associated measure λ is absolutely continuous, we may apply Theorem 2.32, choosing μ as the standard Lebesgue measure over R. Thus, we are guaranteed the existence of a Borel measurable function g that assigns to each Borel set $A \in \mathcal{B}(\mathbb{R})$ a value in $\mathbb{R}$ which we interpret as the probability

of occurrence of the event A. We call g the probability density function (pdf) of the distribution λ. Clearly, if $A = \mathbb{R}$, then we obtain that $\Pr(\mathbb{R}) = \int_{\mathbb{R}} g \, d\mu = 1$. Even more interestingly, by Definition 2.31 and since λ is finite for each A by the conditions of the Radon-Nikodym theorem, we have that there exists a function F such that if $A = (a, b]$, there exists an increasing and right-continuous function F such that

$$\lambda((a, b]) = \int_a^b g \, d\mu = F(b) - F(a) \, .$$

Over $\mathbb{R}$, we may also simply write

$$F(x) = \lambda((-\infty, x]) = \int_{-\infty}^{x} f(t) \, dt \, .$$

We call F the cumulative distribution function (cdf) of the distribution λ. From the above, we may derive that F must be an absolutely continuous function. Furthermore, as with all continuous distributions (including singular ones), we have that for any $r \in \mathbb{R}$, $\lambda(\{r\}) = \lambda((a, a]) = 0$.

As in most applications, so in our work, we typically do not work directly with a distribution but we are rather interested in a variable that assumes values in $\mathbb{R}$ with respect to the probability distribution λ. Such variables are naturally called *random variables*, which we now formally define.

Definition 2.33 (Random variable)**.** *A random variable X on a probability space $(\Omega, \mathcal{F}, \lambda)$ is a Borel measurable function from Ω to $\mathbb{R}$. Formally, we have $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. We say that the random variable X follows the distribution λ and write*

$$\lambda_X(B) = \Pr(\omega : X(\omega) \in B) \, , \, B \in \mathcal{B}(\mathbb{R}) \, .$$

Naturally, we also associate the distribution function F of λ with the random variable X and write $F_X(x) = \Pr(\omega : X(\omega) \leq x)$ for $x \in \mathbb{R}$.

We now present some distributions that appear in the following and also are commonly studied with respect to random polynomials. All of them are well-known distributions with applications outside our line of work, hence we only state them with their name and associated probability density function. These are

- The *Gaussian distribution*, denoted as $\mathcal{N}(\mu, \sigma^2)$ also called normal distribution, with density

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right) \, ,$$

  where $\mu = \mathbb{E}[X]$ is the expected value of X and $\sigma^2 = \mathbb{E}[(X - \mu)^2]$ is its variance, $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$.

- Quite often we will refer to the standard Gaussian distribution, often called *standard normal* distribution and denoted $\mathcal{N}(0,1)$, which is the Gaussian distribution parametrized with mean $\mu = 0$ and variance $\sigma^2 = 1$, which then naturally has density

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) .$$

- The *uniform continuous* distribution, denoted $\mathcal{U}(a,b)$, which intuitively allows the variable X to take values from some bounded interval $[a,b]$ with equal probability. Thus, its density function is

$$f_X(x) = \begin{cases} \frac{1}{b-a} \, , \; x \in [a,b] \, , \\ 0 \, , \;\; \text{otherwise} \;\; . \end{cases}$$

  In our work and related literature, quite often the focus is on the $\mathcal{U}(-1,1)$ uniform distribution.

- The *Rademacher* distribution, denoted $\mathcal{R}(a)$, a *discrete* distribution that is concentrated only on two values a and $-a$ in $\mathbb{R}$. It has a probability mass function given by

$$Pr[X = x] = \begin{cases} \frac{1}{2} \, , \; x = a \, , \\ \frac{1}{2} \, , \; x = -a \, , \\ 0 \, , \;\; \text{otherwise} \;\; . \end{cases} \tag{2.16}$$

  Similar to the case of the uniform distribution, a common parametrization in the study of real roots is $\mathcal{R}(1)$.

As it will soon be apparent, most of this work concerns itself with vectors of random variables, each following a distribution. We may then talk of *random vectors* $\overline{X}$ and refer to their *joint distribution* as per usual in probability theory. The random vectors that we primarily use will always have independent random variables as coefficients, for reasons that we detail below. More details on random vectors, including rigorous definitions, can be found in [1, Section 4.7]. More generally, Chapter 4 of the above reference contains formalizations for the notions we require from probability theory.

2.2.2 Random polynomials

Having defined the basic notions we require from random theory, we may proceed to define random polynomials. Clearly our eventual goal is to sample some appropriate family of polynomials, say $\mathsf{S} \subseteq \mathbb{R}[x]$. More precisely, while it is possible to imagine various scenarios and restrictions on the family S, we will typically either restrict polynomials to be of a specific degree d or of specific sparsity k. As we will see in the following, the second restriction based on sparsity is much more general than that on the degree in most cases, although both are useful both in practice and theory. Furthermore, we will

be typically interested on the expected number of roots as either the degree d or sparsity k tends to infinity, which reflects our wish to both uncover an overarching behavior of the expected number of random zeros with respect to the parameter, as well as avoiding potential irregularities for small values of the parameter that are not interesting either in theory nor practice.

Having decided on our sample space, we may now consider how to sample polynomials from it, regardless of the distribution considered. That is, the distribution can be considered to remain the same while the sampling method changes. We examine three different methods, two of which are of particular interest. The first method is the trivial method of considering the primitives that we sample as the polynomials itself. This method has several weaknesses: To begin with, it is not clear how to carry it out for every family of polynomials. For example, assuming we are interested in polynomials in $\mathbb{R}[x]$ of degree d, it is not clear how to select one at random a priori. Furthermore, there is not a natural agreed upon total order on polynomials, even if we restrict ourselves to univariate polynomials. Thus, it is not clear how usual distributions such as the ones mentioned in the previous section can be applied.

A more sophisticated approach can be found by examining the definition of polynomials given in Section 2.1. In particular, we recall that we may associate a polynomial with its coefficient vector $A \subseteq \mathbb{R}^n$ for some $n \in \mathbb{N}^+$. This convention can be seen to remove the problems we noted above: it is straightforward to consider joint probability distributions over $\mathbb{R}^n$ and sample from them. Furthermore, for the two classes considered above, polynomials of degree d and polynomials of sparsity k, it is straightforward to consider distributions over R^d and R^k respectively. We then only need to decide only upon a monomial ordering, which in particular for univariate polynomials is straightforward. An added advantage is that quite often we are interested in coefficients that do not depend upon one another, as is the case for the two families we are concerned with. In other words, we may consider the coefficients of the polynomial as *mutually independent* random variables, which makes sampling that much easier since we only need to work over $\mathbb{R}$. Furthermore, it is trivial to obtain the joint distribution for the coefficient vector itself from the distributions of the coefficient themselves. Recall that if we have mutually independent random variables $X_1, \ldots, X_n$ with probability density functions $g_1, \ldots, g_n$ respectively, then the joint probability density of the vector $X = (X_1, \ldots, X_n)$ is simply the product of the density function of its components, that is

$$g(X) = \prod_{i=1}^{n} g_i(X_i)\,.$$

We have already seen in Section 2.1 how there is an intimate relationship between the coefficients of a polynomial and its roots. Furthermore, we established in Corollary 2.5 that two polynomials are equal if and only if their coefficient vectors are the same. In addition, there is a number of theorems, mostly concerning root isolation, the study of isolating the roots of polynomial by appropriate intervals or more generally, neighborhoods, a close relative of root counting in many ways, that relate the coefficients to the position of roots on the real line. In the previous approach, we consider the coefficient as free random variables that once determined, will completely determine the roots of the polynomial. The third approach to constructing polynomials, which was introduced less than a decade ago by [72], goes the opposite way, that is instead of considering the

coefficients of the polynomial as the free random variables, it does so for the zeros of the polynomial themselves and then studies the resulting polynomials. However this method does not seem particularly suited to our study, since we wish to study the number of zeros in the first place, thus we wish this to not be a free variable we determine ourselves. Nevertheless, it could be interesting in future questions, since we may thus restrict the number of real zeros, potentially with additional restrictions, then study how the coefficients of the polynomial behave. Note however that as will become apparent in subsequent Chapters, even when the coefficients of a random polynomial follow simple and well-studied distributions, the corresponding distribution of the roots of the polynomial is highly non-trivial, thus we can expect this to be the case when following the reverse direction.

Taking into account the above approaches, we choose to follow the approach most common in literature and treating random polynomials as essentially random coefficient vectors from R^n for some $n \in \mathbb{N}^+$. We now provide the necessary formal definitions.

Definition 2.34 (Random polynomial)**.** *Let $A = (a_1, \ldots, a_n) \in \mathbb{R}^n$, $n \in \mathbb{N}^+$ be a random vector such that the i-th component is a random variable following the distribution λ_i, $a_i \sim \lambda_i$. We call the family of polynomials $f(x)$ a* random polynomial *with exponent vector $E = (e_1, \ldots, e_n)$ for fixed $e_i \in \mathbb{N}$, $1 \leq i \leq n$, and coefficient vector A and write*

$$f(x) = \sum_{i=1}^{n} a_i x^{e_i} .$$

We denote the joint distribution of the random vector A by λ and by extension we also say that the polynomial f follows the distribution λ and write $f \sim \lambda$. In particular, if the random coefficients a_i are mutually independent, we have that $\lambda = \prod_{i=1}^{n} \lambda_i$. Finally, we refer to the family of polynomials obtained by considering all possible polynomials in $\mathbb{R}^n$ under the distribution λ by $\mathfrak{F}_\lambda(x)$.

By convention in the above definition, if the marginal distributions λ_i are identical and the corresponding random variables are mutually independent, we may also say that the polynomial follows the distribution λ_i instead of referring to the joint distribution λ explicitly. For example, we may say that a random polynomial f follows the standard normal distribution $\mathcal{N}(0, 1)$, meaning that all its coefficients are independent and identically distributed according to $\mathcal{N}(0, 1)$.

The classical results in the field focused over polynomials of degree d, which as we've already mentioned can be constructed by sampling from $\mathbb{R}^d$. Note that if the distribution λ_d of the leading coefficient is continuous, we have that $\Pr(a_d = 0) = 0$ and thus the polynomial can be assumed to be of degree exactly d with probability 1. Extending this reasoning to the remaining random coefficients, it follows that the random polynomial can be assumed to be also dense, in particular having the maximum possible sparsity of $d + 1$.

For discrete distributions, assuming that the value 0 is allowed for the random coefficients, we consider the case when the leading coefficient is zero separately. Suppose that the probability of the a_i coefficient being 0 is $\Pr(a_i = 0) = p_i \geq 0$. Then we can assume that with probability $1 - p_d$ we obtain a polynomial of degree d, which we then handle under this conditional assumption. Otherwise, we may assume that with

probability $p_d(1-p_{d-1})$ the polynomial is of degree $d-1$ and continuing in a similar manner, the random polynomial has degree i with probability $(1-p_i)\prod_{j=i+1}^{d} p_j$. Very often the related statements and theorems are proven inductively based on the degree, thus we may only consider the case where $a_d \neq 0$. If the case that all coefficients are zero occurs, with probability $\prod_{i=0}^{d} p_i$, we obviously obtain the zero polynomial, which we typically deal with separately or more commonly ignore, since it would have zeros along the entirety of $\mathbb{R}$ and thus their number would be infinite with non-zero probability.

Similar arguments can be made for the case that is the focus of this work, where the random polynomials are of sparsity k. Especially, we may assume that polynomials whose coefficients follow continuous distributions will have sparsity exactly k with probability 1.

We also make a convention with regards to notation. For the remaining of this work, we will be using $A = (a_1, \ldots, a_n)$ for the coefficients of a random polynomial $f = \sum_i a_i x^{e_i}$ with the understanding that its coefficients a_i are random variables. On the other hand, when we refer to the coefficients by $C = (c_1, \ldots, c_n)$ and write $f = \sum_i c_i x^{e_i}$, we will be referring to a fixed polynomial that in turn has fixed coefficients c_i. Thus it will be obviously immediately whether we refer to random polynomials or not.

Having defined random polynomials and some conventions about them, we may approach the original question that is crucial to answer in our work, namely how may we consider the average-case scenario with respect to the number of real roots of a polynomial in $\mathbb{R}[x]$. This is straightforward, given the foundation of Probability Theory we presented earlier in this section. Instead of speaking vaguely about some kind of weighted average, we can simply refer to the notion of expectation. Namely, given a random polynomial f, consider the family of polynomials induced by the distribution λ, $\mathfrak{F}_\lambda$. Every fixed polynomial in the family, say $f_i \in \mathfrak{F}_\lambda$, has a number of real roots $N_{\mathbb{R}}(f_i)$. To obtain the expected value of real roots over the entire family $\mathfrak{F}$, equivalently the expected number of real roots of the random polynomial f, it suffices to consider the expected value of $N_{\mathbb{R}}(f_i)$ with respect to the distribution λ. We now formally define the expected number of zeros of a random polynomial f with respect to a distribution λ.

Definition 2.35 (Expected number of zeros of a random polynomial)**.** *Let $f(x) = \sum_{i=1}^{n} a_i x^{e_i}$ be a random polynomial whose coefficient vector $A = (a_1, \ldots, a_n) \in \mathbb{R}^n$ follows the distribution λ. As the coefficient vector varies A in $\mathbb{R}^n$, let f_C be the polynomial corresponding to each particular value $C = (c_1, \ldots, c_n)$ that A assumes, $f_C = \sum_{i=1}^{n} c_i x^{e_i}$, and let $S \subseteq \mathbb{C}$ be a set. We define the* expected number of zeros *of the random polynomial f in S with respect to the distribution λ and write $Z_S(f, \lambda)$ as the expectation with respect to the random vector A over all values of $N_S(f_C)$, where $A \sim \lambda$, that is*

- *If λ is a continuous distribution with density g_λ and $d\mu$ is the Lebesgue measure over $\mathbb{R}^n$, we have*

$$Z_S(f, \lambda) = \int_{C \in \mathbb{R}^n} N_S(f_C) g_\lambda(C)\, d\mu\,. \tag{2.17}$$

- *If λ is a discrete distribution concentrated in the countable set $\Omega \subset \mathbb{R}$ with probability mass function $p_\lambda(C) = \Pr_\lambda(A = C)$, $C \in \Omega$, we have*

$$Z_S(f,\lambda) = \sum_{C\in\Omega} N_S(f_C)p_\lambda(C)\,. \tag{2.18}$$

In particular, if $S = \mathbb{R}$, we omit the subscript and write $Z(f,\lambda)$ and speak about the expected number of real zeros of the random polynomial f.

Having defined the necessary notions, we now proceed to highlight several classical results in the study of the expected real roots of random polynomials.

2.2.3 Classic results on the expected number of real roots

Following the work of Laguerre in the late 19th century that we have presented to some extent in Section 2.1 and which is covered in all detail in [56], it seems that the problem was sufficiently solved for the needs of the time, which we assume to be mostly theoretical. For reasons that would be hard to specify, perhaps due to the research trends of the time or the lack of practical application, the study of the problem fell out of favor in the early 20th century, with few exceptions. It was only in the 1930s that interest in the problem re-emerged, in the form of questions about the expected number of roots of random polynomials of degree d under various distributions, most of which we already described in Section 2.2. A string of results covering the period from the early 1930s to the mid 1950s primarily due to Bloch, Erdős, Kac, Littlewood, Offord, and Pólya initiated the study of this topic. Since the 1960s, the topic has remained active and there is a string of results, improving upon these classical results in a variety of ways, continuing to the day of writing this work. Furthermore, in recent years, there has been renewed interest on the number of roots of sparse polynomials, both towards the multivariate setting as well as due to connections with other fields, which is also the motivation of the present work. In this section, we restrict ourselves to the classical results that parametrize polynomials according to their degree, while we defer the study of the known results for sparse polynomials to the next section, after we have introduced the necessary notions to fully demonstrate their relevance to this body of work.

Between the late 19th century and 1930, few works can be found relating to the subject. Indeed, we can only identify a link between the work carried out in 1930 and that of Laguerre: Mihály Fekete in late 1912 in a letter to Pólya [30], both former students of Fejér, poses the following conjecture. Consider the following polynomial

$$g(x) = \left(\frac{1}{q}\right) + \left(\frac{2}{q}\right)x + \cdots + \left(\frac{i}{q}\right)x^{i-1} + \cdots + \left(\frac{q-1}{q}\right)x^{q-2}\,. \tag{2.19}$$

where $\left(\frac{i}{q}\right)$ is the Legendre symbol which is defined as follows for $q \geq 3$ a prime number

$$\left(\frac{i}{q}\right) = \begin{cases} 1\,,\ i = n^2 \mod q\,, \text{ for some } n \in \mathbb{N}^+\,, \\ 0\,,\ i = 0 \mod q\,, \\ -1\,,\ \text{otherwise}\ , \end{cases}$$

i.e., the Legendre symbol is equal to 1 when i is a perfect square modulo q, 0 when it is divisible by q and -1 otherwise. Fekete conjectured that this polynomial g has no

roots in $(0,1)$ and therefore must be positive in this interval, since g as a polynomial is continuous and clearly $g(0) = \left(\frac{1}{q}\right) = 1$. A few years later, Pólya would solve this conjecture in the negative [73, Section II]. In this series of letters, aptly named "Über ein Problem von Laguerre" (About a problem of Laguerre), they utilize a hypothesis originally posed by Laguerre [56]. Specifically, this method is essentially a limiting process that repeats the process from which Laguerre's rule of signs (Theorem 2.28) for $x = 1$ is derived from the coefficient sequence of a polynomial f. As we've already mentioned, for the choice $x = 1$, the coefficient sequence whose variations we examine for Laguerre's rule is $A_0, \ldots, A_n$ where A_i is the sum of all coefficients of the polynomial in question, say f, with index equal or greater than i. We may now apply this process once again, defining a sequence $A_i^{(1)}$ such that the i-th term is the sum of all terms A_i of index i or greater. We may continue in the same way to define sequences $A_i^{(j)}$. As with Laguerre's rule, this is equivalent to dividing by $(1 - x)$. The hypothesis now is that as j tends to infinity, the variations of the sequence $A_i^{(j)}$ would tend to the number of real roots of the polynomial f, that is

$$\lim_{j\to\infty} \left(\#s\left(A^{(j)}\right)\right) = N_{(0,1)}(f)\,,$$

where we extend the notation $\#s(A^{(j)})$ to denote the number of sign changes of the sequence $A_i^{(j)}$. This type of limiting process would be interesting both theoretically as well as practically, since assuming a reasonable rate of convergence one would obtain much better bounds in a rather simple manner. They were able to achieve limited success, for example Pólya notes that the statement is true if f only has simple zeros in $(0,1)$, that is, of multiplicity 1. To the best of our knowledge, the question remains open and resolving it would be interesting at the very least theoretically.

As it is already apparent from the involvement of Legendre symbols, a notion from number theory, as well as the title of Pólya's work "Verschiedene Bemerkungen zur Zahlentheorie" [73], which translates to "Various Remarks on Number Theory", there was a shift of attention from classical algebra, i.e., the study of polynomial equations, to number theoretic concerns. This shift was already apparent in the work of Laguerre [56], by using polynomials and their number of roots as approximations to questions about series, and thus should be of little surprise. It is also quite apparent in the doctoral work of Dr. Gottfried Grimm [33], himself a student of Pólya, which expanded upon the topic discussed above and its relation to Dirichlet L-functions. Indeed in more generality, the early modern number theory utilized many concepts relating to polynomials and their zeros, including among others Galois theory and in general polynomials with integer or rational coefficients in an attempt to study functions central to the field such as the Riemann zeta function and the Dirichlet L-function, see for example [41] for a classical treatment of number theory and [68] for a book focused on algebraic number theory. Even in terms of computer algebra, a topic much closer to this work, the connection between integers and polynomials is made clear early on, see for example the classical textbook [93].

This connection to number theory can also be seen in the first publication we consider to deal specifically with the topic of the expected number of zeros of a random polynomial, namely the work of Bloch and Pólya [10]. This work considers random polynomials of

degree d of the following type

$$f(x) = 1 + \sum_{i=1}^{d} a_i x^i \,, \tag{2.20}$$

where the coefficients a_i for $1 \leq i \leq n$ are random variables with the following distribution $\mathcal{H}$: a_i have values either 1, 0 or -1 with equal probability $1/3$, that is, $\Pr_{\mathcal{H}}[a_i = 1] = \Pr_{\mathcal{H}}[a_i = -1] = \Pr_{\mathcal{H}}[a_i = 0] = 1/3$. We can assume the same is true for the trailing coefficient and that we obtain a polynomial of the above form, that is with constant coefficient 1, by dividing with the (nonzero) trailing term, which does not affect the number of positive real roots. Although the writing style of the time is quite different than today, with introductions about motivation and conclusions relating to open questions either completely absent or interspersed in the text of the main proof itself, this work contains a few motivating examples, with three relating to number theory and the rest involving simple series.

The first such example is the aforementioned polynomial involving Legendre symbols which we already defined in Equation (2.19). In [10] its connection to the positive real roots of the related Dirichlet series is made explicit. The second example involves the Liouville function $\lambda : \mathbb{Z}^+ \to \{-1, 1\}$ which is defined for $n \in \mathbb{Z}^+$ with respect to the parity of the exponents of prime numbers in its unique factorization, that is

$$\lambda(n) = (-1)^{e_1 + \cdots + e_m} \,, \text{ where } n = p_1^{e_1} \dots p_m^{e_m} \,,\, p_i : \text{prime} \,.$$

In particular the following relation is true

$$\sum_{i=1}^{\infty} \lambda(i) i^{-s} = \frac{\zeta(2s)}{\zeta(s)} \,,\, s \in \mathbb{C} \,,\, Re(s) > 1 \,,$$

where ζ is the Riemann zeta function. Finally, the third such example involves the Möbius function, which can be defined with respect to the Liouville function as follows for $n = p_1^{e_1} \dots p_m^{e_m}$:

$$\mu(n) = \begin{cases} \lambda(n) \,,\, e_1, \dots, e_m \leq 1 \,, \\ 0 \,,\, \exists\, e_i \geq 2 \,, \end{cases}$$

that is, it is equal to the Liouville function for square-free positive integers and 0 otherwise. It also relates to the Riemann zeta function as follows

$$\sum_{i=1}^{\infty} \mu(i) i^{-s} = \frac{1}{\zeta(s)} \,.$$

We may then define the following polynomials for Liouville and Möbius functions respectively

$$f_\lambda = \sum_{i=1}^{n} \lambda(i) x^{i-1} \,,$$
$$f_\mu = \sum_{i=1}^{n} \mu(i) x^{i-1} \,.$$

Clearly all examples above are in the sample space $\mathfrak{F}_{\mathcal{H}}$ of the random polynomials of the above form in Equation (2.20), which among other reasons motivates their study. For the first two examples, they cite previous work determining the number of roots up to some degree n, while for the case of the Möbius function, they can determine that such a polynomial has more than $o(\log n)$ roots.

More interestingly to our investigation are two arguments presented therein. The first argument is easy to understand and allows us to measure the number of real zeros of a random polynomial in the entirety of $\mathbb{R}$ by examining only the number of zeros in the interval $(0, 1]$ assuming certain conditions are satisfied. These conditions are satisfied for all distributions that we examine. In our findings in subsequent sections, we will present in a formal way the necessary conditions a distribution must satisfy for this to hold. In particular, it is easy to find distributions for which the above does not hold, for example certain trivial distributions, i.e., those that are equivalent to selecting a fixed polynomial with probability 1, since we clearly may have polynomials with the same number of zeros in $[0, 1]$ that nevertheless have different number of real zeros.

The argument reducing the study of zeros to $(0, 1]$ is rather simple. Consider a fixed polynomial $f(x) = \sum_{i=0}^{k-1} c_i x^{e_i}$. For each negative real root $x_0 < 0$ of f, $f(x_0) = 0$, the polynomial $g(x) \coloneqq f(-x)$ has a positive real root at $-x_0 > 0$ since

$$g(-x_0) = f(x_0) = 0\,. \tag{2.21}$$

Thus we may study the number of positive zeros of $g(x)$ instead of the negative zeros of f. Similarly, now consider a root of f, say x_0, in $[1, \infty)$. For each such root, the following polynomial has the root $1/x_0$ in $(0, 1]$

$$g(x) \coloneqq x^d f(1/x) = \sum_{i=0}^{k-1} c_i x^{d-e_i}\,, \tag{2.22}$$

where $d = \deg(f)$. This is true since $g(1/x_0) = x_0^{-d} f(x_0) = 0$. Thus we may study the zeros of g in $(0, 1]$ instead of those of f in $[1, \infty)$. From the above it is now clear that for a random polynomial f, we may study the number of zeros restricted in the interval $(0, 1]$, provided that the associated family of polynomials $\mathfrak{F}$ is closed under the above operations.

We also note that the above argument does not deal with roots at $x = 0$, which however can be easily identified by the exponent of the trailing term. Since this is not feasible for random polynomials, this case is dealt with arguments according to the distribution at hand. In particular, for continuous distributions whether the interval includes its limit points or not is irrelevant. To see this, it is sufficient to consider that the roots of a polynomial are completely determined by its coefficients. Then, it is clear that the roots of a random polynomial are themselves random variables with some distribution determined by that of the coefficients and thus also continuous. It follows that the probability that a random polynomial has a specific root, say $x = 1$, is 0, i.e., $Pr(f(1) = 0) = 0$. This also means that whether we count zeros with multiplicity or not does not alter the expected number of zeros. We thus obtain the following:

Proposition 2.36. *Let f be a random polynomial following a continuous distribution $\mathcal{D}$. Let $[a, b]$ be some interval of $\mathbb{R}$ and let $Z_S^m(f, \mathcal{D})$ denote the expected number of zeros of f with respect to $\mathcal{D}$ counted with* multiplicity*. Then the following is true*

$$Z_{[a,b]}(f,\mathcal{D}) = Z_{(a,b]}(f,\mathcal{D}) = Z_{[a,b)}(f,\mathcal{D}) = Z_{(a,b)}(f,\mathcal{D})\,, \tag{2.23}$$

$$Z_S^m(f,\mathcal{D}) = Z_S(f,\mathcal{D})\,. \tag{2.24}$$

This observation simplifies many statements and is implicit in the relevant results in the literature, although never stated. We note that by Sard's theorem [78], the critical set of any continuously differentiable function has Lebesgue measure 0. That statement not only provides an alternative justification of the above statement about counting with multiplicity or not, but also allows us to generalize the same statement for multivariate polynomials.

In the case of the distribution $\mathcal{H}$, that Bloch and Pólya study for dense polynomials, we see that the above argument holds. Furthermore the constant term is always 1, as we mentioned above by dividing with the trailing coefficient if necessary. Thus $f(0) = 1$ always and there are no roots at $x = 0$.

The second argument given in their work [10] relates to the expected number of zeros of a random polynomial of the form in Equation (2.20), i.e., following the distribution $\mathcal{H}$. In fact, the argument also follows and is in fact easier to state for the related Rademacher distribution $\mathcal{R}(1)$, which as we've already defined in Equation (2.16), assigns either the value 1 or -1 to the coefficients with the same probability.

The argument uses the version of Laguerre's rule of signs which bounds the number of positive real zeros below a number, in this case $x = 1$. Alternatively we could use the argument given above to study the roots at $[1, \infty)$ and use the version stated in Theorem 2.28. We now shortly present the proof given in their work, written in a more modern way, since it will be useful to us in the following.

Theorem 2.37 (Expected number of real zeros with respect to $\mathcal{R}(1)$, dense case, [10])**.** *Let f be a random dense polynomial of degree d following the Rademacher distribution $\mathcal{R}(1)$. Then its expected number of real zeros is $Z_{\mathbb{R}}(f,\mathcal{R}(1)) = O(\sqrt{d})$.*

Proof. Let $f = \sum_{i=0}^{k-1} a_i x^{e_i}$. As we've already mentioned and given that $|f(0)| = 1$, it suffices to study the zeros in $(0, 1]$. We initially also exclude the point $x = 1$, which we can deal with as a special case later. We apply Laguerre's rule of signs for positive zeros smaller than $x = 1$, by which the number of real zeros of f in $[0, 1]$ is bounded by the number of sign changes in the sequence

$$a_0, a_0 + a_1, \ldots, \sum_{i=0}^{j} a_i, \ldots, \sum_{i=0}^{k-1} a_i\,.$$

Since $a_{2j} = 1$ or $a_{2j} = -1$, for the $(2j)$-th term in the above coefficient sequence, say A_{2j}, to precede a change of sign, in addition to $A_{2j-1}A_{2j+1} < 0$, we must also have that $A_{2j} = 0$. The probability of the latter happening is equal to precisely half of the a_i having the opposite value to the other half, that is with probability

$$\Pr\left(A_{2j} = 0\right) = \frac{1}{2^{2j}}\binom{2j}{j}\,.$$

Then that the A_{2j+1} term has the opposite sign of A_{2j-1} occurs if a_{2j+1} has the value opposite that of A_{2j-1}, which happens with probability $1/2$. So a sign change

happens with probability $\frac{1}{2^{2j+1}}\binom{2j}{j}$. Since a sign change may occur at most every other term, ignoring terms equal to 0 for the sake of simplicity, and by setting $m = \lceil \frac{k-1}{2} \rceil$ we have that the expected number of sign changes in $(0, 1]$ is

$$\sum_{i=1}^{m} \frac{1}{2^{2i+1}} \binom{2i}{i} \leq \sum_{i=1}^{m} \frac{1}{2\sqrt{i\pi}} \leq \sqrt{\frac{k+1}{2\pi}} = O(\sqrt{k}) \,.$$

Since the random polynomial following the distribution $\mathcal{R}(1)$ is closed with respect to the operations given in Equation (2.21) and Equation (2.22), the value in the entirety of $\mathbb{R}$ is at most 4 times the above and thus asymptotically the same.

It now only remains to deal with the case when $x = 1$ is a root. We note that $f(1) = \sum_{i=0}^{k-1} a_i$. Thus, for f to have a root at this point, it is equal to precisely half of the coefficients being fixed to a value opposite that of the others. By the above, this happens with a probability bounded by $\frac{1}{2\sqrt{k\pi}}$ and in this case, by Descartes' rule of signs, Theorem 2.17, the number of real roots of f in $\mathbb{R}$ is bounded by twice the number of sign changes, which are at most $\frac{k-1}{2}$. We thus have that the expected number of roots of f in $\mathbb{R}$ is at most

$$\frac{k-1}{2\sqrt{k\pi}} + \left(1 - \frac{1}{2\sqrt{k\pi}}\right) c\sqrt{k} \leq \left(c + \frac{1}{2\sqrt{\pi}}\right) \sqrt{k} - \frac{c}{2\sqrt{\pi}} = O(\sqrt{k}) \,,$$

for some $c \in \mathbb{R}^+$. The proof now concludes by taking f to be a dense polynomial as in Equation (2.20) and thus setting $k = d + 1$. □

Proving the same result for random polynomial under the slightly more complicated distribution $\mathcal{H}$ is not too difficult, we only need to condition on the number of random coefficients that potentially can be zero.

Corollary 2.38 (Expected number of real zeros with respect to $\mathcal{H}$, dense case)**.** *Let f be a random dense polynomial of degree d following the distribution $\mathcal{H}$, as given in Equation* (2.20)*. Then its expected number of real zeros is $Z_{\mathbb{R}}(f, \mathcal{H}) = O(\sqrt{d})$.*

Proof. Examining the proof of Theorem 2.37, we note that we make use only of Descartes' and Laguerre's rule of signs, as well as other arguments that do not depend on the polynomial being dense. In fact, the theorem as written is true for polynomials of any sparsity. Thus, it suffices to separate the cases with respect to the number of non-zero coefficients, to which we them may apply the previous result and finally take the expectation.

We begin by counting how many polynomials of degree d with respect to the distribution $\mathcal{H}$ exist in the collection of polynomials $\mathfrak{F}_{\mathcal{H}}$. Clearly this collection has 3^{d+1} polynomials of degree d. Of those, a simple counting argument reveals that $\binom{d+1}{i}2^{d+1-i}$ have exactly i zero coefficients, where the binomial coefficient signifies that we choose i coefficients to be zero and then the remaining can vary between the other two values. Then each such polynomial with i zero coefficients will have at most $O\left(\sqrt{d+1-i}\right)$ zeros in expectation by Theorem 2.37. An exception is the case $i = d + 1$, where all coefficients are 0 and thus we obtain the 0 polynomial. Since in this case the number of zeros would be infinite, we explicitly exclude it from the calculation but note that it occurs with probability $1/3^{d+1}$. Thus in total, the expected number of zeros will be

$$\sum_{i=0}^{d} \frac{\binom{d+1}{i} 2^{d+1-i}}{3^{d+1}} \sqrt{d+1-i} \leq \frac{\sqrt{d+1}}{3^{d+1}} \sum_{i=0}^{d} \binom{d+1}{i} 2^{d+1-i}$$
$$= \frac{3^{d+1}-1}{3^{d+1}} \sqrt{d+1} = O\left(\sqrt{d}\right),$$

as claimed.

□

A few years later, Littlewood and Offord substantially expanded the scope of the previous work to various other distributions, once again for the dense case. Their investigation lead to a series of papers, starting with [58], which references many of the same examples as Bloch and Pólya as motivation and state their results without proof. Shortly after in [59], they also published the proofs for the aforementioned results. We now state the most important results of their paper that also relate to the present work, before commenting on them briefly.

Theorem 2.39 (Expected number of real zeros in the dense case [59])**.** *Let $f(x)$ be a random polynomial of degree d that follows one of the following distributions:*

- *The standard normal distribution $\mathcal{N}(0,1)$.*
- *The uniform continuous distribution $\mathcal{U}(-1,1)$.*
- *The Rademacher distribution $\mathcal{R}(1)$.*

Then the expected number of real zeros of f in the entirety of $\mathbb{R}$ is at most $O(\log^2 d)$, that is

$$Z_{\mathbb{R}}(f, \mathcal{N}(0,1)) = O(\log^2 d)\,, \tag{2.25}$$
$$Z_{\mathbb{R}}(f, \mathcal{U}(0,1)) = O(\log^2 d)\,, \tag{2.26}$$
$$Z_{\mathbb{R}}(f, \mathcal{R}(1)) = O(\log^2 d)\,. \tag{2.27}$$

A significant difference that we have to note compared to the previous work is that the proof techniques used in the proof of the above heavily exploit the fact that the polynomial is dense, in fact many of them essentially rely on properties of the polynomial viewed as a geometric series. In particular this means that in contrast with the results of Bloch and Pólya, these results cannot readily be applied to sparse polynomials. Even more disappointingly, after close examination we have determined that it is highly unlikely that the results can be modified. Precisely, the proofs contained in [59] showcase that Lemma 2 cannot be satisfied by most sparse polynomials, where in the context of the proof a_i are the terms (without coefficient) corresponding to nonzero coefficients. Furthermore, a crucial part of the proof of Theorem 2.39 is combining Lemmas 3 and 4 while also setting appropriate values for several parameters contained therein; the appropriate choice of parameters is important in achieving the final result. Unfortunately, in the case of sparse polynomials it is impossible to satisfy both relations for any meaningful

choice of parametrization, at best we were able to show a bound that would match the one given by the Descartes' rule of signs in the worst case and thus of little interest. In subsequent sections, we hint towards methods that we believe hold promise for achieving the analogous results in the sparse case. Finally, it is worth mentioning that in their subsequent and final work in this series [60], they expand the above for the Rademacher distribution $\mathcal{R}(1)$ as follows: each random coefficient a_i of the random polynomial f is allowed to be multiplied by an arbitrary complex number c_i. Despite this case being much more general than the one presented above, they are still able to show a bound of the order $O(\log^6 d + \log d \log M)$, where $M = \sum_{i=0}^{d} |c_i|$, thus we obtain a bound $O(\log^6 d)$ if we ignore the magnitude of the arbitrary weights.

Thus far, the methods used to obtain the results we detailed above varied as the problem they were employed to solve, although already in the work of Littlewood and Offord the desire for a common approach is readily visible, although while the same result is obtained, it requires conversion between the different cases. That would change with the work of Kac in 1943 [45], which introduced a methodology to count the number of real zeros of any polynomial and careful examination of the proof reveals the same technique may be extended to any differentiable function. The importance of this work cannot be stressed out enough, as it would become fundamental in the field of study and remains so to this day. It is also essentially the start of the study of similar questions for any kind of stochastic process, which was formalized as the Rice formulas, thus also often called Kac-Rice formulas, depending on the context, see [3, Chapter 3] for a detailed treatment.

Here we will restrict ourselves to random univariate polynomials, since this is the scope we are mostly interested in in this work, matches the presentation in the original Kac paper and is ideal in showcasing the main idea. Consider $f(x)$ to be a univariate polynomial, which for now we consider to be fixed. Also consider some subinterval of $\mathbb{R}$, say (a, b), in which we will count the number of zeros of f. Let $\varepsilon > 0$ be sufficiently small enough and define the following indicator function

$$\psi_\varepsilon(y) = \begin{cases} 1\,, & -\varepsilon < y < \varepsilon\,, \\ 0\,, & \text{otherwise}\ . \end{cases}$$

Now consider $\psi_\varepsilon(f(x))$ and the set $\mathcal{E}$ that contains all points $x \in (a, b)$ such that $\psi_\varepsilon(f(x)) = 1$. Since f is continuous this is an open set in $\mathbb{R}$ and can therefore be written as a sum of disjoint open intervals, $\mathcal{E} = \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_r$, which must be finitely many since the interval is bounded. We now choose ε to be small enough so that the following conditions are satisfied. Note that such an ε always exists.

- For all $x \in (a, b)$ such that $-\varepsilon < f(x) < \varepsilon$, i.e., in a small strip around the x axis, x is not a turning point of f, with the exception when $f(x) = 0$, i.e., x is a root of f.
- Neither a nor b are contained in any of the $\mathcal{I}_i$.

Now since each $\mathcal{I}_i$ denotes an open interval that is delimited by two points, say $I_i = (a_i, b_i)$ such that $|f(a_i)| = |f(b_i)| = \varepsilon$, we have

$$\int_{I_i} |f'(x)|\,dx = 2\varepsilon\,,\ 1 \le i \le r\,.$$

Thus in total for the entirety of the interval (a, b) we can compute that

$$\frac{1}{2\varepsilon}\int_{(a,b)} \psi_\varepsilon(f(x))|f'(x)|\,dx = \frac{1}{2\varepsilon}\sum_{i=1}^{r}\int_{I_i} |f'(x)|\,dx = r\,. \tag{2.28}$$

We may now ponder what precisely the above process achieved. Recall that ε was chosen so that there are no turning points of f in the strip $-\varepsilon < f(x) < \varepsilon$ unless that turning point is also a root. Thus in each interval $I_i = (a_i, b_i)$ the behavior of f is easy to analyze and in fact falls in one of the following four cases:

(1) f contains no turning point and $f(a_i) = \varepsilon$, thus it follows that $f(b_i) = -\varepsilon$. Furthermore, from the mean value theorem, there must exist a point $c \in I_i$ such that $f(c) = 0$. Also by the absence of a turning point, this point is unique, thus the interval I_i contains precisely a single root.

(2) A mirror case of the first case, with the only difference being that $f(a_i) = -\varepsilon$ and thus $f(b_i) = \varepsilon$. We again conclude the existence of a single root in I_i.

(3) f contains a turning point and $f(a_i) = \varepsilon$. By the choice of ε, this turning point, say c, must also be a root, $f(c) = 0$. Also note that no further roots may occur in the interval I_i, since that would require an additional turning point that is not a root. We conclude thus that $f(b_i) = \varepsilon$ and that the interval I_i contains exactly one root.

(4) This case is a mirror of the previous case and is similar in its analysis, with the difference that $f(a_i) = f(b_i) = -\varepsilon$.

Thus, in every case the interval I_i contains exactly one root. It is also easy to see that every root must be contained in such an interval, since for every root x_0 we have that $-\varepsilon < f(x_0) = 0 < \varepsilon$. It follows that r, the number of intervals identified above is precisely equal to the number of real zeros of f in the interval (a, b). We only require a few final remarks. First, the choice of ε above clearly depends on the polynomial, however for each polynomial there is such an ε and in fact the statement remains true for any further $0 < \varepsilon' < \varepsilon$. Thus, we may take the limit $\varepsilon \to 0$ and satisfy the conditions for any polynomial. Furthermore, the above can be applied to the entirety of $\mathbb{R}$ in the case of polynomials, since in reality the limits of integration are finite. This is since, as we've mentioned several times by now, eventually the leading term of a polynomial dominates and the value of the polynomial tends to infinity. Thus there exists x_0 large enough such that for all $|x| > x_0$ we have that $|f(x)| > \varepsilon$ and thus it suffices to consider $(-x_0, x_0)$. Finally, we must note that as it is clear from the above, this method does not count the roots with multiplicity, but rather it only counts distinct roots. However, as we've already shown in Proposition 2.36, this does not make a difference for the cases we are interested in. We therefore obtain the following, due to Kac

Theorem 2.40 (Kac formula for the number of real zeros of a univariate polynomial[45])**.** *Let $f(x)$ be a univariate polynomial in $\mathbb{R}[x]$. Then its number of distinct real zeros, counted without multiplicity is equal to*

$$N_{\mathbb{R}}(f) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{\mathbb{R}} \psi_{\varepsilon}(f(x)) |f'(x)| \, dx \, .$$

Furthermore, the above can be extended to random polynomials with relatively little work. Let f be a random polynomial following the distribution $\mathcal{D}$, resulting in the sample space $\mathfrak{F}_{\mathcal{D}}$. By definition, the expected number of real zeros of the random polynomial f is then given by

$$Z_{\mathbb{R}}(f, \mathcal{D}) = \int_{\mathfrak{F}_{\mathcal{D}}} \lim_{\varepsilon \to 0} \left(\frac{1}{2\varepsilon} \int_{\mathbb{R}} \psi_{\varepsilon}(f(x)) |f'(x)| \, dx \right) \, d\mathcal{D} \, .$$

where $d\mathcal{D}$ is the distribution measure associated with the distribution $\mathcal{D}$. We now simply need to rewrite the above so it is easier to work with by utilizing some fundamental theorems, specifically we may use the Fubini-Tonelli Theorem [1, Theorem 2.6.6] to interchange the two integrals and in addition the Lebesgue's dominated convergence Theorem [1, Theorem 1.6.9] to interchange the outer integral and the limit process. We thus obtain the following

Theorem 2.41 (Kac formula for random univariate polynomials [45])**.** *Let f be a random univariate polynomial following the distribution $\mathcal{D}$. Then the expected number of real zeros of f with respect to the distribution $\mathcal{D}$ is given by*

$$Z_{\mathbb{R}}(f, \mathcal{D}) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{\mathbb{R}} \int_{\mathfrak{F}_{\mathcal{D}}} \psi_{\varepsilon}(f(x)) |f'(x)| \, d\mathcal{D} \, dx \, . \tag{2.29}$$

Assuming that $\mathcal{D}$ is an absolutely continuous distribution over $\mathbb{R}^k$ with probability density function $g_{\mathcal{D}}(x)$, which also implies that all polynomials in $\mathfrak{F}_{\mathcal{D}}$ have sparsity at most k, we may rewrite Equation (2.29) as

$$Z_{\mathbb{R}}(f, \mathcal{D}) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{\mathbb{R}^{k+1}} \psi_{\varepsilon}(f(x)) |f'(x)| g_{\mathcal{D}}(f) \, d\mu \, , \tag{2.30}$$

where $d\mu$ is the Lebesgue measure over $\mathbb{R}^{k+1}$. We often refer to both versions of the above as the *Kac integral* but caution the reader that this may also refer to the Feynman-Kac formula used in physics.

As it is evident from the above, the methodology of Kac is quite general and can be adapted to any sample space of polynomials, or more generally differentiable functions. In fact, most of the subsequent work in the field stems from carefully analysing Equation (2.29) for the distribution at hand, including our own work. The reader should be forewarned however that while this integral is deceptively simple, solving the integral even asymptotically can prove to be extremely difficulty if at all possible depending on the distribution, as Kac himself also concludes in his closing remarks. In the case of his work [45], he applies the above to random polynomials following the standard normal distribution $\mathcal{N}(0, 1)$. His proof is straightforward as he proceeds by considering $f'(x)$ and $\psi_{\varepsilon}(f(x))$ as random variables and computing the density function for their joint probability distribution, which in turn allows the calculation of the expected value above.

The technique only relies on facts from probability theory about the standard normal distribution and joint distributions and does not readily apply to other cases or offer any further insight, so we omit and refer the interested reader in the work [45] itself. Furthermore, in subsequent sections we will refer to a subsequent work which examines the Kac integral in a much more intuitive way and forms one of the pillars of our results. We finish the examination of Kac's work by stating his main result.

Theorem 2.42 (Expected number of real zeros of dense polynomials following $\mathcal{N}(0,1)$)**.** *Let $f(x)$ be a random polynomial of degree d that follows the standard normal distribution $\mathcal{N}(0,1)$. Then the expected number of real zeros of f is*

$$Z_{\mathbb{R}}(f, \mathcal{N}(0,1)) \approx \frac{2}{\pi} \log d + \frac{14}{\pi} = \Theta(\log d)\,.$$

In particular, we see that Kac's result improves upon the bound in Equation (2.25) by Littlewood and Offord. Furthermore, Kac remarks that the integral can be used to show that most of the roots occur near either $x = 1$ and $x = -1$. In fact, it can be shown that outside these neighborhoods, the number of expected real roots is $O(1)$, i.e., constant. These remarks were later investigated and improved upon by various works in the field, some of which we will review below.

The first investigation was due to Kac himself, who in his previous work noted that his result can be extended to any random polynomial whose coefficients a_i are independent and identically distributed according to a distribution $\mathcal{D}$ whose variance is equal to 1, i.e., $\sigma_{\mathcal{D}}^2 = 1$. It turns out that this original assessment was erroneous with regards to the methods used, in particular in his follow-up work [46] he notes the argument evoked, i.e., using the central limit theorem does not apply in all cases, in particular in the case of the Rademacher distribution $\mathcal{R}(1)$. He also notes many of the difficulties that arise when considering such discrete distributions, while in contrast he is able to improve upon Littlewood and Offord with respect to the continuous uniform distribution $\mathcal{U}(-1,1)$.

Theorem 2.43 (Expected number of real zeros of dense polynomials following $\mathcal{U}(-1,1)$, [46])**.** *Let $f(x)$ be a random polynomial of degree d that follows the continuous uniform distribution $\mathcal{U}(-1,1)$. Then the expected number of real zeros of f is*

$$Z_{\mathbb{R}}(f, \mathcal{U}(-1,1)) \approx \frac{2}{\pi} \log d = \Theta(\log d)\,.$$

We consider the last of the classical results to be the work of Erdős and Offord [29] published in 1956. In this work, they are able to use a number of arguments specific to the case of the Rademacher distribution $\mathcal{R}(1)$ to show a similar improvement. In particular, they consider carefully chosen intervals in $(\frac{1}{2}, 1)$ so that with high probability each contains a single root and then determine that their number rarely differs from the desired expectation. Note that a polynomial following the Rademacher distribution cannot have roots in $[0, \frac{1}{2}]$ since the constant term is equal to 1 and $\sum_{i=1}^{d} c_i x^i \leq \sum_{i=1}^{d} x^i < 1$ for $x \leq 1/2$. They thus obtain the following result:

Theorem 2.44 (Expected number of real zeros of dense polynomials following $\mathcal{R}(1)$, [29])**.** *Let $f(x)$ be a random polynomial of degree d that follows the Rademacher distribution $\mathcal{R}(1)$. Then the expected number of real zeros of f is*

$$Z_{\mathbb{R}}(f, \mathcal{R}(1)) \approx \frac{2}{\pi} \log d = \Theta(\log d)\,.$$

The avid reader will have already noticed that for dense polynomials following the distributions considered in these classical results, namely the standard normal $\mathcal{N}(0,1)$ via Theorem 2.42, the continuous uniform case $\mathcal{U}(-1,1)$ via Theorem 2.43, the Rademacher case via Theorem 2.44, and via extension on the latter using the same argument as in Corollary 2.38, for the distribution $\mathcal{H}$, which was defined in Equation (2.20), we obtain the same bound of approximately $\frac{2}{\pi} \log d$ and in particular $\Theta(\log d)$ for the expected number of real roots. As Kac's remark above reveals, it was already suspected that this bound applied to a much wider family of distributions than the standard normal, a conjecture that the above results strengthen. Naturally, as we've already remarked, it is easy to construct distributions, especially discrete or even trivial ones, whose expected number of roots is the maximum possible, that is for dense random polynomials, equal to d, matching the upper bound given by Lemma 2.4. Similar results are also demonstrated in the above classical works, although the class of polynomials given are typically trivial with respect to the randomness involved.

Subsequent results in this topic can be roughly categorized in three classes: The first explores the limits of the above conjecture by investigating which distributions adhere to a similar asymptotic result as given in Theorem 2.42 for the expected number of zeros of *dense* random polynomials following them. The second class of results attempts to either improve on the constants involved in the approximation of either the original result by Kac or subsequent generalizations for other distributions or otherwise improve the results already established, for example by deriving facts about the distribution of zeros in such cases. The third, which also includes the current work, attempts to generalize such results to settings other than the dense polynomials, for example sparse polynomials or multivariate ones. There's a plethora of such results spanning over half a century, we choose here a select few that are good representatives in our judgement but by no means encapsulate the totality of achievements of this research area.

Beginning in the late 1960s, a number of results extended the results above to a much wider class of distributions. To begin with, Stevens in his doctoral thesis [84], whose results are also summarized in the subsequent publication [85], shows that if the coefficients a_i of a polynomial follow a distribution $\mathcal{D}$, $a_i \sim \mathcal{D}$, such that the expected value of the coefficients is zero, $\mathbb{E}_{\mathcal{D}}(a_i) = 0$, and their variance is 1, $\sigma^2(a_i) = 1$, in addition to additional conditions that bound the fourth moment of a_i as well as the rate with which the tails of the distribution decay, then the result of Kac given in Theorem 2.42 also applies in this case. A few years later, Logan and Shepp extended the same bound of $\Theta(\log d)$ initially to the Cauchy distribution [61], which is of interest since all its finite moments, including the expected value and variance are undefined, and subsequently generalized it even further in [62] to any distribution with characteristic function given by $e^{-|z|^\alpha}$ for $0 < \alpha \leq 2$, which subsumes both the standard normal case which we obtain for $\alpha = 2$ and the Cauchy distribution which we obtain for $\alpha = 1$.

This string of results continued into the early 1970s, by a series of results due to Ibragimov and Maslova. The two authors published in 1971 their work [38] that extends the bound of $\Theta(\log d)$ to any distribution $\mathcal{D}$ with mean 0, i.e., $\mathbb{E}_{\mathcal{D}}(a_i) = 0$, that is additionally in the domain of attraction of the normal law. Recall that a distribution $\mathcal{D}$

is in the domain of attraction of the normal law if given m independent and identically distributed (i.i.d.) random variables $X_i \sim \mathcal{D}$, there exist constants A_m and B_m such that the distribution of the random variable X defined by

$$X = \frac{\sum_{i=1}^{m} X_i - A_m}{B_m},$$

tends to some normal distribution $\mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 \in \mathbb{R}^+$ as $m \to \infty$, see also [37]. This is essentially a generalization of the central limit theorem and can be informally be viewed as a distribution that asymptotically can be approximated by some normal distribution, up to the constants involved. In their subsequent work [39] they also improve upon this result by removing the condition of requiring a zero mean. This result, to the best of our knowledge, was the greatest generalization of the results of Kac with respect to the size and importance of the class of distributions involved, although further smaller improvements have followed.

We now examine some representatives from results that improve upon previous results. We begin with a work by Maslova which is in many a ways a follow-up to the aforementioned results in [38]. In the work [65], she shows that if a random dense polynomial has i.i.d. coefficients a_i that follow a distribution $\mathcal{D}$ with mean $\mathbb{E}_{\mathcal{D}}(a_i) = 0$, with a finite absolute moment of order $2 + \varepsilon$ for some $\varepsilon \in \mathbb{R}^+$, $\mathbb{E}(|a_i|^{2+s}) < \infty$, and the probability of the a_i being 0 is itself 0, i.e., $\Pr(a_i = 0) = 0$, we obtain that the variance of the number of zeros (which is a random variable) is also $\Theta(\log d)$. Note that the final condition is satisfied for all continuous distributions and by discrete distributions whose sample space does not include 0. Under the same conditions, in a subsequent publication [66] she shows that the distribution of zeros of such a random polynomial is asymptotically normal, i.e., the zeros of such a random polynomial when sampled n times for $n \to \infty$ tend to the distribution $\mathcal{N}(\log d, \log d)$, where we omit the constants in the parametrization of the distribution.

Other results include an improvement to the constants of the results given by Kac and others. These results are important especially from the scope of mathematics, since the precise constant can often reveal interconnections and implications that would otherwise remain obscured. On the other hand, from the scope of this current work, we care more about asymptotic behavior and as will become clear in the next Section, for the ultimate end goal even a polynomially large error would not affect the validity of the result. Nevertheless, such results remain important since such close scrutiny often derives many useful properties about the distribution of zeros. A perfect example of such a work would be the celebrated publication of Edelman and Kostlan [28], in which they improve the expectation given by Kac. In particular, they show that for a random dense polynomial of degree d following the standard normal $\mathcal{N}(0, 1)$ distribution, the expected number of zeros is given by

$$Z_{\mathbb{R}}(f, \mathcal{N}(0,1)) = \frac{2}{\pi} \log d + 0.6257358072\ldots + \frac{2}{d\pi} + O\left(\frac{1}{d^2}\right). \tag{2.31}$$

This work contains an examination of Kac's result for the standard normal distribution which is of intrinsic value to our own work, as well as several generalizations, thus we defer a more close examination to the next Chapter. Similar results have appeared more

recently, for example Nguyen,Nguyen and Vu in 2016 [69] show that any distribution $\mathcal{D}$ with mean $\mathbb{E}_{\mathcal{D}}(a_i) = 0$, variance $\sigma^2 = 1$ and bounded $(2+\varepsilon)$ moment has an expected number of zeros given by

$$Z_{\mathbb{R}}(f, \mathcal{D}) = \frac{2}{\pi} \log d + c_{\varepsilon,\mathcal{D}}\,,$$

where $c_{\varepsilon,\mathcal{D}}$ is a constant depending only on the distribution $\mathcal{D}$ and thus consequently also on ε. In a follow-up work Do, Nguyen and Vu [26] extend their result as follows: For continuous distributions, they substitute the condition on the variance with the more general requirement of the distribution having a p-integrable density function $g(x)$ for some $p \in \mathbb{N}^+$. This means that the following integral exists and is finite

$$\int_R |g(x)|^p \, dx < \infty\,.$$

They also show the same result for the family of discrete distributions described by the following probability mass function

$$\Pr\left(a_i = v\right) = \frac{1}{2N}\,,\; v \in \{-N, \ldots, -1, 1, \ldots, N\}\,,$$

which in particular includes the Rademacher distribution $\mathcal{R}(1)$ for $N = 1$. Finally, they exhibit that such random polynomials are unlikely to have zeros of multiplicity 2 or greater.

The third category of results expands into other models than dense random polynomials. In a seminal work that inspired many subsequent results in recent years, Tao and Vu [89] examine in detail the problem and show necessary conditions for two distributions to behave similarly with respect to the distribution of their zeros. In particular, they utilize the framework they develop to show results for a number of generalizations, including Weyl polynomials, whose random coefficient a_i are weighted by $\frac{1}{\sqrt{i!}}$, i.e., the i-th term has coefficient $\frac{a_i}{\sqrt{i!}}$ and elliptic polynomials, whose random coefficients a_i are weighted instead by $\sqrt{\binom{d}{i}}$. Among many interesting results, they show that for such dense polynomials of degree d and for distributions $\mathcal{D}$ with mean 0, variance 1 and bounded $(2+\varepsilon)$-moment we obtain the following bound on the expected number of zeros:

$$Z_{\mathbb{R}}(f, \mathcal{D}) = \Theta(\sqrt{d})\,.$$

Their work contains a treasure trove of results relating to the behavior of zeros of random polynomials, to which we strongly refer the interested reader. Another example of results into other models is the results on sparse polynomials, which also form the core of our contributions. We thus defer their examination for the next sections, once the necessary prerequisites have been introduced. This primarily include a short introduction into Algebraic Complexity Theory, the primary area of our interest and how it is connected to the study of the zeros of polynomials, which we present in the next section.

2.3 Algebraic Complexity Theory

2.3.1 Models of computation

In this section we provide a short introduction to the exciting field of Algebraic Complexity Theory. This research area provides the main motivation for this work and is the primary research field of the author. It is a rich and active area of research with connections both to classical Complexity Theory (i.e. using Turing machines as the model) and Computer Algebra, among others. While broadly defined it contains a variety of topics, we will focus narrowly on algebraic (also called arithmetic) circuits, the related complexity classes and the central question associated with them. Even within this narrow context, we will only provide the necessary introductory knowledge required to understand the questions that are relevant to this work, which leaves out a breadth of interesting theory and results. For the reader interested in further details, the classical textbook on the topic is [19], that also deals with topics in the broader sense of the area as defined above. The related publication [15] deals more specifically with the complexity classes, notions of reduction and complete problems we will present below in a much more detailed manner. A more modern survey that contains an introduction to the topic with an actively updated list of recent lower bound results in the area can be found in [77]. Alternatively, the slightly older survey [79] covers many overlapping topics but also contains some additional results.

Specifically we will study the model of arithmetic circuits, also called algebraic circuits. We will use the two terms interchangeably, preferring the latter due to our subject matter. We remark that the former term is more commonly used and established, thus perhaps more useful for the reader interesting in finding further material, especially older publications. On the other hand, recent work, including the present one, focuses on the model defined for the reals $\mathbb{R}$ or the complex numbers $\mathbb{C}$, justifying our use of the latter term. The typical variant of the model used captures precisely the computation of polynomials, as will become clear during its definition. It should make clear why we focused on polynomials in previous Sections of this Chapter.

Before defining the model formally, it is instructive to investigate its origins. In particular, we would like to answer the question on why studying the efficiency of computing polynomials under this model is of particular interest. In general, investigating the efficiency of different ways of performing certain algebraic computations was intuitively done long before the formalization of notions such as complexity, algorithm or even computation. It is likely that as soon as algebraic computations arose, which happened really early in history as we've seen in Chapter 1, the informal search for efficient methods also arose. More concentrated efforts arose much later as the result of the formalization of mathematics in the last few centuries, as the results presented in Section 2.1 betray. Nevertheless, the systematic investigation of the topic is much more recent. Many authors, including [15] cite Ostrowski's investigation on the efficiency of Horner's rule in [70], published in 1954, as the first formal foray into algebraic complexity. While in that work the notion of the first model we wish to examine appeared, Strassen much later in [87] offers a much more complete definition of the model we now refer to as *straight-line programs*. The idea behind this model is rather simple and it is not unique to algebraic calculations, but can also accommodate several types of non-uniform

computations, including Boolean ones, by changing the operators allowed.

To define the model in general, we only require to define the allowed input and operations. In the case of algebraic straight-line programs, the input originates from some field $\mathbb{K}$ and the allowed operations are the basic algebraic operations of addition $+$, subtraction $-$, multiplication $*$ and division $/$, defined in the usual way for the field in question. As has been the case so far, we are typically interested in the case where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. The semantics of a straight-line program are then as follows: A straight-line program consists of lines that are executed in order. Most of the lines correspond to operations, with the exception of a number of lines at the beginning which correspond to the input. The input can be either variables or constants from the underlying field $\mathbb{K}$. Each line is indexed by an appropriate symbol, which is also used to denote either the input element or the output of the operation of that line. For each line that corresponds to an operation, the operands then have to be previous lines, so that they have already been well-defined and computed. The output of the straight-line program is the same as the output of the last line.

The model is very intuitive, although a bit cumbersome to work with, as the following simple example that computes the discriminant of the generic quadratic polynomial $ax^2 + bx + c$ showcases:

$L1 \leftarrow a$;
$L2 \leftarrow b$;
$L3 \leftarrow c$;
$L4 \leftarrow 4$;
$L5 \leftarrow L2 * L2$; // b^2
$L6 \leftarrow L1 * L3$; // ac
$L7 \leftarrow L4 * L6$; // $4ac$
$L8 \leftarrow L5 - L7$; // $b^2 - 4ac$

Algorithm 1: A straight line program for the discriminant of quadratic polynomials

A few immediate shortcomings of the model are immediately apparent. To begin with one that is actually intended, straight-line programs as perhaps the name betrays do not allow for branching; the lines of the program will always be executed in the same order and in their entirety. It is also clear that this type of model is not particularly human-readable. Even in the small example above, it is difficult to determine the quantity being computed just by looking at the program. Another weakness of the model is that its lack of visual detail makes difficult to distinguish how the different elements of the input interact to compute the algebraic expression at the output. For example, if we were to consider questions such as whether a particular program requires multiplying two input elements, which may not appear in the final expression due to cancellations, or which parts of the program can be computed in parallel, the model itself is not particularly effective in showcasing this information.

On the other hand, the model of (algebraic) circuits, which we will now introduce, not only is computationally equivalent to straight-line programs, but much more appealing to work with in general and in particular in the context of this work. The model rather unsurprisingly is similar to Boolean circuits that one encounters both in practice and in theory and very intuitive. Its main strength is that it combines the simplicity of

straight-line programs with the aforementioned intuitiveness of circuits, as well as an object that is much more manageable mathematically. We now proceed with its formal definition.

Definition 2.45 (Algebraic circuit)**.** *An algebraic circuit is a directed acyclic graph (DAG). Semantically, the circuit computes a function f over some field $\mathbb{K}$, that is $f\colon \mathbb{K}^n \to \mathbb{K}^m$, where n is the number of leaf nodes in the DAG and m the number of roots. Thus, n also denotes the size of the input to the circuit, including constants from $\mathbb{K}$. Typically $m = 1$, that is there is a single root, and $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.*

The vertices of the DAG can be semantically divided into three distinct categories, as follows: Leaf nodes, denoted input nodes *have indegree* 0 *and arbitrary outdegree and they simply denote the input to the circuit, similar to input instructions in a straight-line program. They can thus be labeled by either a variable $X_1, \ldots, X_n$ or a constant from $\mathbb{K}$. All other nodes are called gates and are labeled by one of the allowed operations. For algebraic circuits those are addition,subtraction,multiplication and division denoted respectively by $\{+, -, *, /\}$ and have indegree* 2*. Semantically, the gate computes the operation denoted by its label with inputs given by its children, denoted as its input, and forwards the output to any outgoing edges. Thus gates correspond to operation lines in a straight-line program. The gates corresponding to the roots are distinct from all other nodes since they have outdegree* 0 *and are called* output gates*. Output gates are distinguished by an arrow which has no end node (thus note that this is not an edge in the graph). In our work, there is typically only one such gate. All other gates have outdegree at least* 1*. When order is important for the input of a gate or the outputs of a circuit, e.g. for the division operation, we order the inputs from left to right according to the respective edges in the drawing of the graph.*

As an example Figure 2.1 illustrates an algebraic circuit for the discriminant of the generic quadratic polynomial $ax^2 + bx + c$, corresponding to the straight line program in Algorithm 1.

Note how even in this small example, the circuit model graphically reveals a wealth of information that was not apparent in the equivalent straight-line program: We can immediately see that this circuit for the discriminant uses 3 multiplication and 1 subtraction operation, with one of the multiplication involving b twice as an argument. Furthermore, the circuit can easily be parallelized into two separate threads, which only merge at the root node. Since none of the subtrees share an input, except for the gate computing b^2, it's clear that no cancellations occur, as well.

Even more importantly, having a graph as the underlying model allows us to use related notions that would be hard to express in terms of straight line programs. To begin with, there exist a number of measures, some more obvious than not, that one could define for this model. Furthermore, there are a number of restrictions we could impose to this model that make sense when we visualize the algorithm as a DAG and would be much more opaque if we were thinking simply in terms of straight-line programs. We now introduce the most important measures and restrictions, as well as the core complexity classes that they give rise to.

The most important measure is *size*, which can be defined as either the number of nodes or the number of edges in the graph. When defining it as the number of nodes, typically we only include gates, that is we exclude input nodes. Thus, with this definition,

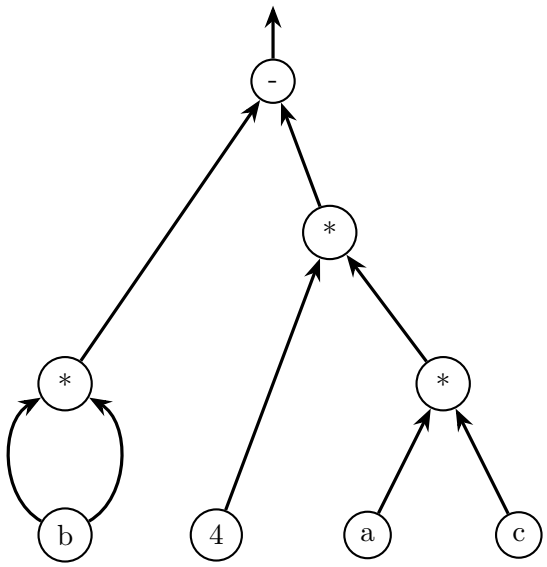

Figure 2.1: An algebraic circuit computing the discriminant of quadratic polynomials

the circuit in Figure 2.1 has size 4. Note that this will always be equivalent to the number of operations carried out. If we used the number of edges instead, the circuit size would instead be 8. This is precisely the total number of operands for all operations in the circuit. In this example, we have 4 operations all of arity 2, thus a circuit size of 8. Since all operations have constant arity, the two definitions are asymptotically equivalent. However, there are variants of the model where this is no longer the case, which we will soon discuss. In any case, the two definitions will differ at most by a polynomial factor. To see this, assume we have a circuit with s gates and n input nodes, including constants. Each one of the gates can thus have at most $s + n - 1$ operands in the worst case, thus the number of edges is at most $s(s + n - 1) = O(s^2)$, assuming $n < s$. In this study, polynomial factors are negligible, so the definition chosen is not of great importance. We will thus assume for the sake of simplicity and ease that the size of the circuit corresponds to its number of vertices.

A different measure often used is the *depth* of a circuit, which is defined simply as the length of the longest path from a root to an input node in the underlying graph. In the circuit for the discriminant in Figure 2.1, that would be the path from root to the input nodes labeled either with a or c, which would be 3. The depth can also interpreted as the length of the longest thread in a parallel computation of the circuit, which in turn would also be the running time. There are a number of results relating to depth and in recent years circuits of very small constant depth (3 or 4) have been studied extensively, as can be seen in [77].

The last statement can seem curious to someone not familiar with the area. After all, in the model described, a circuit of depth 4 could have at most 31 gates and 32 inputs, severely limiting the possible functions that it is possible to compute (in fact, there's only a finite number of different circuits, thus also functions). Instead, the results for

constant depth are possible by using a relaxation of the above model, which allows for much more flexibility. The relaxed model alters the definition in the following way:

(1) Multiplication gates have arbitrary arity.

(2) Addition gates have arbitrary arity and allow each operand to be multiplied by a constant from the underlying field $\mathbb{K}$, thus becoming essentially linear combination gates.

Note that under those conditions, every polynomial has a circuit of depth 2, since it can be expressed as a sum of monomials and each monomial can be computed by a single multiplication gate. The circuit for the discriminant in Figure 2.1 is then transformed as follows, in the relaxed model.

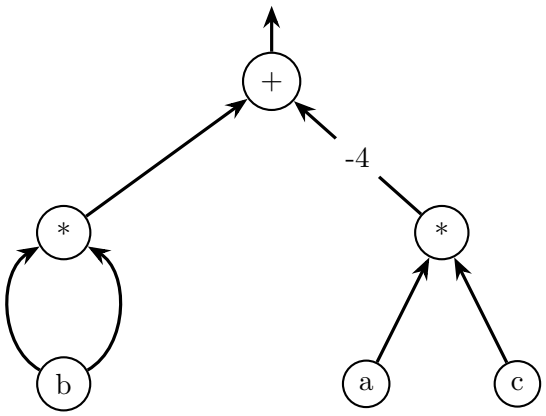

Figure 2.2: A relaxed version of the algebraic circuit in Figure 2.1

Note that both the size and depth of even this small circuit are affected by the relaxation. It is thus important to consider the context of the problem studied and choose the appropriate model definition. Furthermore there are additional ways in which the model may be changed, as for example allowing additional operations. Such an example that has been studied often enough is the addition of powering gates denoted by $\wedge$, see for example [34], which allow raising the input to any constant integer power. Note that this does not affect the size asymptotically, since the same operation could be done by a constant number of multiplication gates in the original model, but it does affect the depth. On a different example, the study of constant-free circuits has also been of interest, see [52], in particular Section 6. In such circuits only the basic constants of a field, namely 0 and 1, are allowed "for free", i.e. as an input node. All other constants must be constructed via gates. Therefore, large constants involved in the computation have an impact in the circuit size. Of particular cases are instances where large constants appear in intermediate steps, but not on the desired output, since such circuits come with added costs that would not be present in the original model.

In this work, we use the relaxed model, not only because it is often more convenient but also because sparse polynomials have an intimate connection with small depth

circuits. For example, note that since as we've seen above we can derive a simple depth 2 (relaxed) circuit for polynomials written as a sum of monomials, a polynomial having sparsity k implies it can be computed by a depth 2 circuit of $O(k)$ size. Furthermore, a conjecture that is central to this work can be related to circuits of depth 4, as we will soon demonstrate.

In addition, we may restrict our set of operations to only addition $+$ and multiplication $*$. The lack of subtraction is of little issue since it can be easily simulated by addition and using an appropriate constant, in our case -1 for $\mathbb{R}$ and $\mathbb{C}$. The lack of division becomes a problem however, since although we have access to the inverse of constants from the underlying field, there is no way to compute even simple rational functions such as $\frac{1}{x}$. If we instead restrict our attention to computing polynomials, then a theorem by Strassen [88] allows us to remove any division gates without a large cost. In particular we have

Theorem 2.46 (Division elimination [88])**.** *Let f be a polynomial of degree d over $\mathbb{K}[x_1, \ldots, x_n]$ that is computed by a s-size circuit $\mathcal{C}$ with addition, multiplication and division gates. Then f is computable by a circuit $\mathcal{C}'$ using only addition and multiplication gates of size polynomial in d and s, that is $O(s^{c_1} d^{c_2})$, for some $c_1, c_2 \in \mathbb{N}^+$.*

Proof. We only present the proof sketch of the theorem here and direct the reader interested in all the technical details to the cited work. Alternatively, a more general proof appears in Section 2.5 of [79]. Clearly, if $\mathcal{C}$ contains no division gates, we are done, so assume it contains at least one. The first step is regardless of the amount of division gates in the circuit, to reduce them to a single one that will in fact be the output gate. This is possible via simple algebraic manipulation using the following rules, as many times as it is necessary. In the following figures, the labeled triangles simply represent a subcircuit. The labels of the subcircuits will also stand for their output.

For addition, we employ the rule $A + \frac{B}{C} = \frac{AC+B}{C}$, that is we have

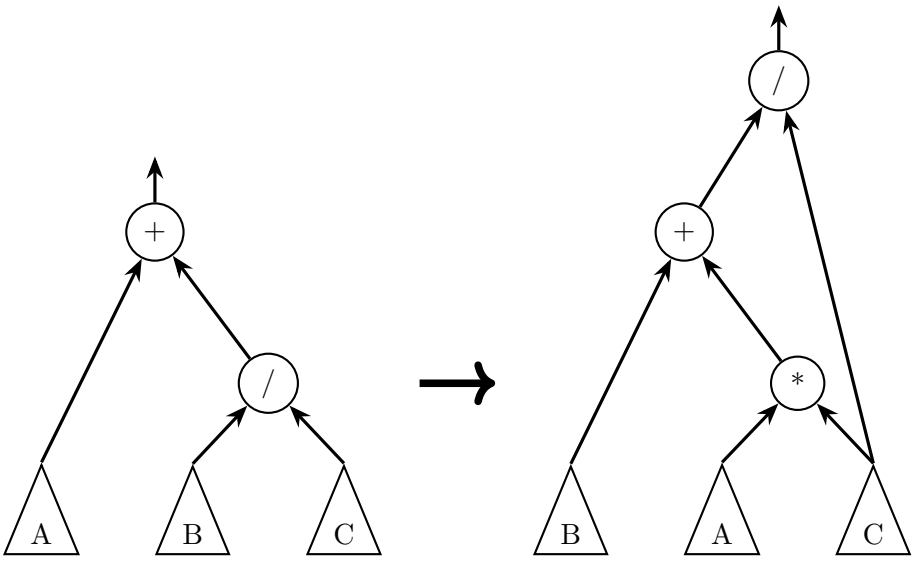

Figure 2.3: Changing order of division and addition

For multiplication, the rule is simply $A * (B/C) = (A * B)/C$, or graphically

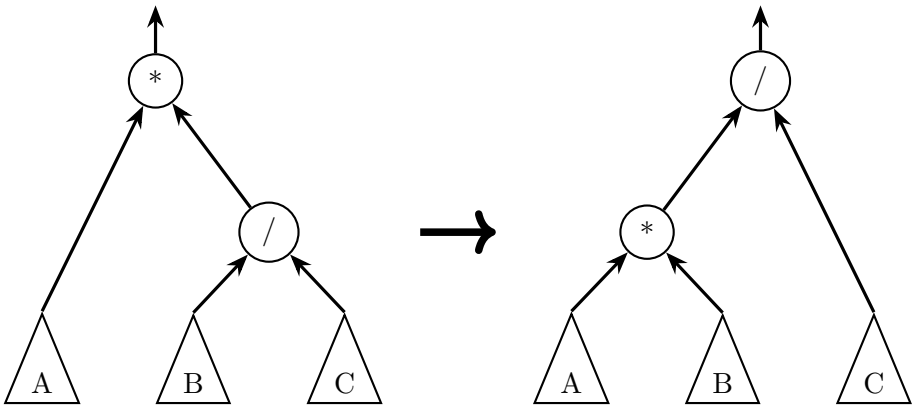

Figure 2.4: Changing order of division and multiplication

Finally, we may merge two divisions by the rule $\frac{A}{\frac{B}{C}} = \frac{A*C}{B}$, that is

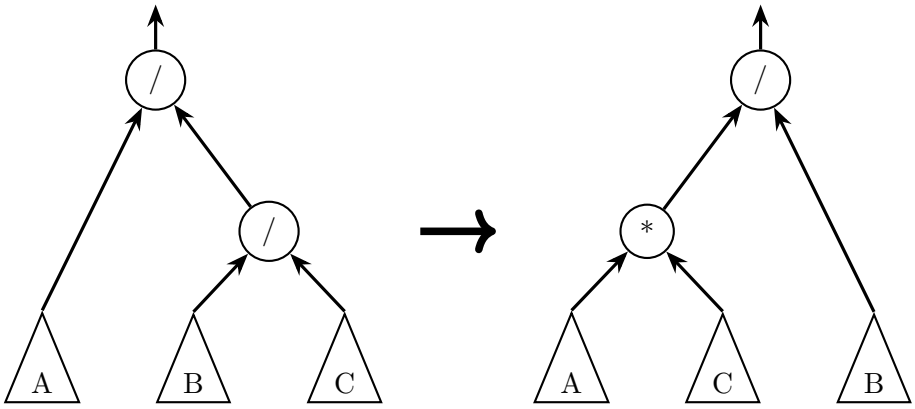

Figure 2.5: Merging two division operations

After applying the above operations repeatedly as needed, we can ensure that the resulting circuit for the polynomial f is of the form $\frac{G}{H}$ where both G and H are polynomials computed by the relevant subcircuits. Also note that the resulting circuit is well-defined, i.e. none of the denominators evaluates to 0, for at least the same values that the original circuit $\mathcal{C}$ was. What remains is to remove this final division operation that is the output gate of this intermediary circuit.

To do so, we rely on the Maclaurin series of the function $r(x) = \frac{1}{1-x}$. It is easy to verify that the i-th order derivative (with 0-th order being the function itself) of $r(x)$ is simply

$$r^{(i)}(x) = \frac{\partial^i}{\partial x}\left(\frac{1}{1-x}\right) = \frac{i!}{(1-x)^{i+1}} = i!\,(r(x))^{i+1}\ ,\ i \in \mathbb{N}^+\,. \tag{2.32}$$

Thus for the Maclaurin series, evaluating at $x = 0$, we obtain by Equation (2.32)

$$r^{(i)}(0) = \frac{i!}{1^{i+1}} = i!\,,\ i \in \mathbb{N}^+\,. \tag{2.33}$$

Thus the Maclaurin series of $r(x)$ is simply

$$r(x) = \sum_{i=0}^{\infty} \frac{f^{(i)}(0)}{i!} x^i = \sum_{i=0}^{\infty} \frac{i!}{i!} x^i = \sum_{i=0}^{\infty} x^i\,. \tag{2.34}$$

For the above we have to assume that $x \neq 1$, due to the domain of the function in question. If we were to consider x over $\mathbb{R}$ or $\mathbb{C}$ as is usually the case, we would also have to assume that x is in the radius of convergence of the right hand side, that is $|x| < 1$. However if we treat x as a formal variable, we no longer have to worry about convergence. In particular, we assume that x is a formal power series, that is taking values from $\mathbb{K}[[X]]$, for $X = (x_1, \ldots, x_n)$, similar to the original work [88].

Now we have to bring the polynomial f, which we have computed as $f(X) = \frac{G(X)}{H(X)}$, to a form resembling $r(x)$ while also respecting the necessary restrictions. Note that we can easily rewrite f as follows

$$f(X) = \frac{G(X)}{H(X)} = G(X)\frac{1}{1 - (1 - H(X))}\,. \tag{2.35}$$

We thus set $x \coloneqq 1 - H(X)$, which means that the following two conditions must be respected

$$x \neq 1 \iff 1 - H(X) \neq 1 \iff H(X) \neq 0\,, \tag{2.36}$$

$$r(0) = 1 \iff \frac{1}{1 - (1 - H(0, \ldots, 0))} = 1 \iff H(0, \ldots, 0) = 1\,. \tag{2.37}$$

The first condition in Equation (2.36) ensures that $r(x)$ is well-defined and is satisfied automatically since we are ensured that $H(X)$ is well-defined in the computation of $f(X)$. The second condition in Equation (2.37) ensures that the evaluation at $X = (0, \ldots, 0)$ for the Maclaurin series produces the correct value. While we cannot ensure that $H(0, \ldots, 0) = 1$, we can slightly modify H to ensure this is the case. From the first condition we know that $H(0, \ldots, 0) = a \neq 0$, since f is well-defined and thus $H(X) \neq 0$ for all x. We can thus define $N(X) \coloneqq \frac{H(X)}{a}$ and similarly $M(X) = \frac{G(X)}{a}$. We thus have $N(0, \ldots, 0) = \frac{H(0,\ldots,0)}{a} = 1$ and $N(X) \neq 0$ for all $X \in \mathbb{R}^n$. We then can write f as

$$f(x) = \frac{G(X)}{H(X)} = \frac{\frac{G(X)}{a}}{\frac{H(X)}{a}} = \frac{M(X)}{N(X)} = M(X)\frac{1}{1 - (1 - N(X))}\,. \tag{2.38}$$

We could now use Equation (2.34) to write

$$f(X) = M(X)\frac{1}{1 - (1 - N(X))} = \sum_{i=0}^{\infty} M(X)(1 - N(X))^i\,. \tag{2.39}$$

However, note that this leaves us with infinite terms which we cannot possibly compute. The final trick is to instead of $f(X)$, compute its homogeneous parts, i.e., of same

degree, separately. To do so, we also have to compute the homogeneous parts of the expression on the right. Note that to do that we require only polynomially many more gates. For each addition, let $O(X)$ be the polynomial computed by the gate and $A(X)$ and $B(X)$ the two inputs, that is $O(X) = A(X) + B(X)$. Let $\mathrm{HOM}_i\left(O(X)\right)$ stand for the homogeneous part of O of degree i for $0 \leq i \leq d$, where d is the degree of O. We employ the same notation for the other polynomials as well. We can now compute $\mathrm{HOM}_i\left(O(X)\right)$ simply by

$$\mathrm{HOM}_i\left(O(X)\right) = \mathrm{HOM}_i\left(A(X)\right) + \mathrm{HOM}_i\left(B(X)\right) , \tag{2.40}$$

for $0 \leq i \leq d$ thus we require d additional gates compared to the original circuit. For multiplication, suppose $O(X) = A(X) * B(X)$, For each homogeneous part of O, say of degree i, there are i possible ways to obtain a part of degree i, that is

$$\mathrm{HOM}_i\left(O(X)\right) = \sum_{j=0}^{i} \mathrm{HOM}_j\left(A(X)\right) * \mathrm{HOM}_{i-j}\left(B(X)\right) . \tag{2.41}$$

Thus we require $O(i)$ gates, $i+1$ multiplication gates and $O(i)$ addition gates to compute the end result (in the original model). In total, we thus require $\sum_{j=0}^{d} i = O(d^2)$ gates in total. In total, we require $O(d^2 s)$ gates in the worst case to compute the homogeneous part of the polynomial separately.

It remains to actually show how to compute the homogeneous parts of f from those of M and N. To do so, we need to establish certain facts. To begin with, since $N(0, \ldots, 0) = 1$, it follows that $1 - N(0, \ldots, 0) = 0$ therefore we have that $\mathrm{HOM}_0\left(1 - N(X)\right) = 0$, i.e., the constant term of $1 - N(X)$ is 0. It is easy to see that since the smallest possible degree of a monomial in $1 - N(X)$ is 1, the smallest possible degree of a monomial in $(1 - N(X))^i$ is also i, which can also verified by applying the rule in Equation (2.41). We thus have that

$$\mathrm{HOM}_j\left(\left(1 - N(X)\right)^i\right) = 0\,,\; j < i\,,\; i \in \mathbb{N}^+ . \tag{2.42}$$

Now to compute the j-th homogeneous part of f, we have that

$$
\begin{aligned}
\mathrm{HOM}_j\left(f(X)\right) &= \mathrm{HOM}_i\left(\sum_{i=0}^{\infty} M(X)\left(1-N(X)\right)^i\right) && \text{via Equation (2.39)}\\
&= \sum_{i=0}^{\infty} \mathrm{HOM}_j\left(M(X)\left(1-N(X)\right)^i\right) && \text{via Equation (2.40)}\\
&= \sum_{i=0}^{\infty}\sum_{\lambda=0}^{j} \mathrm{HOM}_{j-\lambda}\left(M(X)\right)\mathrm{HOM}_{\lambda}\left(\left(1-N(X)\right)^i\right) && \text{via Equation (2.41)}\\
&= \sum_{i=0}^{j}\sum_{\lambda=0}^{j} \mathrm{HOM}_{j-\lambda}\left(M(X)\right)\mathrm{HOM}_{\lambda}\left(\left(1-N(X)\right)^i\right) && \text{via Equation (2.42)}\\
&= \sum_{i=0}^{j} \mathrm{HOM}_j\left(M(X)\left(1-N(X)\right)^i\right) .
\end{aligned}
$$

In the penultimate line, we used the fact given by Equation (2.42) that for $\lambda < j < i$, the λ-th homogeneous part of $1 - N(X)$ will be 0 and that such terms do not contribute to the sum. Therefore, to compute the homogeneous part of f, we only require a multiplication, an addition and raising $1 - N(X)$ to the i-th power, which can be done in time $O(\log i) = O(\log d)$ since $0 \leq i \leq j \leq d$. We may then use $O(d)$ additions (or a single one in the relaxed model) to combine the homogeneous parts and compute f. Also note that we only need to compute the required homogeneous parts of the right hand side only once and reuse the results. In total, the size of the resulting circuit is polynomial in s and d, as required.

□

Another useful structural result is that of depth reduction, which relates the degree of the polynomial to be computed to the depth of the circuit computing it. In particular, this result due to [92] states that if a polynomial f of degree d is computed by a circuit $\mathcal{C}$ of size s, then it can be computed by a circuit $\mathcal{C}'$ of size polynomial in s and d and depth polylogarithmic in d and s. We now give the precise statement of the theorem, without proof.

Theorem 2.47 (Depth reduction)**.** *Let f be a degree d polynomial computed by a size s circuit. Then f can be computed by a circuit of depth $O\left(\log d\left(\log d + \log s\right)\right)$ and size $O(s^{c_1}d^{c_2})$, for some $c_1, c_2 \in \mathbb{N}^+$.*

A streamlined proof of the theorem can be found in Section 2.4 of [79]. This result is of particular interest, since the above implies that if a circuit has polynomial size and polynomially bounded degree, it has a circuit that also has polylogarithmic depth. This is in stark contrast with what is true for Boolean circuits, where a strict depth hierarchy exists where increasing the depth by just 1 allows circuits to compute more Boolean functions [81]. We also note for the interested reader that both [79] as well as the more recent [77] contain results for depth reduction in small constant depth, specifically depth 3 and 4, which formed the basis of some interesting results in the research area in the past decade.

Before continuing with the model at hand, we briefly mention other models that are also studied in algebraic complexity theory and deserve mention, although we will not employ them further. To begin with, we already mentioned that algebraic circuits are only a more convenient model to study straight-line programs but are computationally equivalent. In particular, they do not allow branching, no matter the input received the same path of execution will be followed, hence the name straight-line. If we also allow branching, we then obtain the model of computation trees. This model includes an additional type of node called a branching node. The branching node receives as input two operands and conditions on one of several operations, depending on the base field. We can always test for equality and in addition if the base field $\mathbb{K}$ is ordered, comparisons of the two operands are allowed. If the condition is satisfied, the flow of computation follows one of the branching node's children, say the left, otherwise, it follows the other one, say the right. Thus a branching node has in- and out-degree 2. We can use branching nodes to either modify the algebraic circuit model or more commonly, the straight-line program one, see e.g. Section 4.4 of [19] for a complete description of the model. It is clear that computation trees are strictly a stronger model, since we can construct simple functions that are not possible to be computed by an algebraic circuit. Computations trees are, often implicitly, the model of choice when studying computer algebra and the related algorithms.

So far all models we've mentioned are non-uniform. In their work [11], Blum, Shub and Smale introduced the model named after them, in short the BSS model. Briefly described, the BSS model resembles Turing or random access machines with a single yet important exception: Instead of simply a finite set of symbols, each memory cell (or register) may hold any element from $\mathbb{K}$, with arbitrary precision, with the original work having $\mathbb{K} = \mathbb{R}$. It is then possible to describe analogues of classical complexity classes in this model, for example for $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$ analogues of P and NP have been defined and studied.

Having defined the necessary models, we now proceed to define the rest of the significant notions of this theory, such as how we define "problems" to compute, complexity classes and so on.

2.3.2 Complexity classes, reductions and important problems

The next step in our study is to define the notion of a problem, which will in turn allow us to define complexity classes. We've already mentioned that we wish to compute polynomials. We could take any single polynomial and pose the question what is the smallest size (or minimal for any other measure) circuit computing it and indeed in certain cases this is a question posed for small examples, either to gain intuition into the problem or as a testbed for attacking the more general problem, see for example [95],[8]. This setup however would be far too restricting and in particular we would not be able to study how the complexity of the problem, as expressed by the various relevant measures, grows as a function of the input size.

To remedy this, we study instead *families* of polynomials. Formally, a family of polynomials is a set of polynomials $\mathfrak{F}_n = \{f_n(X)\,,\, n \in \mathbb{N}^+\}$, such that for each $n \in \mathbb{N}^+$, a unique member of the set corresponds to it. While one could construct various families in such a fashion, typically the members of a family follow a specific pattern. For example,

two families that are central to the field of study and we will define formally later are the determinants and permanents of $n \times n$ matrices. A yet different example that is given by a closed formula could be the sum of n-th powers of n variables, namely

$$f_n(X) = \sum_{i=1}^{n} X_i^n \,.$$

Note that although often closed formulas like the above exist that describe the polynomial in question, they do not necessarily describe the most efficient way to compute the polynomial. Perhaps the most famous example of this is the study of the complexity of matrix multiplication, see for example [9], for which the usual closed form was already shown by Strassen in [86] to not be optimal, which spawned a search for the precise complexity of matrix multiplication that has seen myriads of research papers and continues to this day.

Once we have a family of polynomials $\mathfrak{F}_n$ at hand, we can associate with it a family of circuits, say $\mathfrak{C}_n = \{\mathrm{C}_n \,,\, n \in \mathbb{N}^+\}$, so that C_n computes f_n. We can also associate with each circuit the relevant complexity measure we are interested in; in this study typically we are simply interested in its size, say s_{C_n}. Now, let the family of circuits $\mathfrak{C}_n$ for $\mathfrak{F}_n$ be chosen so that for each $n \in \mathbb{N}^+$, the circuit C_n has the smallest size among all circuits computing f_n. Then the following function is precisely the size complexity of $\mathfrak{F}_n$

$$s(n) = s_{\mathrm{C}_n} \,.$$

As is the case in classical complexity as well, we consider families of polynomials whose size complexity is polynomially bounded to be "easy". This includes many of the polynomials computed in practice, which typically also have a cubic or even quadratic bounded complexity. This class is named VP, where V stands for Valiant, who originally gave the definition [91] and is formally defined as follows

Definition 2.48 (VP)**.** *The class VP consists of all families of polynomials $\mathfrak{F}_n$ that satisfy the following conditions:*

(1) All polynomials in $\mathfrak{F}_n$ have polynomially many variables in n. That is, there exists a polynomial $v(n) : \mathbb{N}^+ \to \mathbb{N}^+$ such that for every $n \in \mathbb{N}^+$, the number of variables of f_n is bounded by $d(n)$, i.e., $f_n(X) = f_n(X_1, \ldots, X_{N(n)})$ such that $N(n) \leq v(n)$, $n \in \mathbb{N}^+$.

(2) All polynomials in $\mathfrak{F}_n$ have polynomially bounded degree in n. That is, there exists a polynomial $d(n) : \mathbb{N}^+ \to \mathbb{N}^+$ such that for every $n \in \mathbb{N}^+$, the degree of f_n is bounded by $v(n)$, i.e., $\deg(f_n) \leq d(n)$, $n \in \mathbb{N}^+$.

(3) All polynomials in $\mathfrak{F}_n$ have circuits of polynomially bounded size. That is, for each polynomial f_n in $\mathfrak{F}(n)$, there exists a circuit C_n that computes f_n of size $s(n)$ and a polynomial $b(n) : \mathbb{N}^+ \to \mathbb{R}$, such that the size $s(n)$ is bounded by $b(n)$, i.e., $s(n) \leq b(n)$, $n \in \mathbb{N}^+$.

While there exist polynomials of degree higher than polynomial that have polynomial sized circuits, they are not in VP due to the second condition. The reason for that, as well for the first condition, is that although we typically consider the polynomials as formal

polynomials, we would like for them to be able to be computed efficiently in practice. If we allowed for polynomials of superpolynomial degree, then the bit size of the output could also be superpolynomial with regards to the input bit size, which would make the computation impractical.

As we've already mentioned, a lot of polynomials computed in practice have algorithms that either directly or with a little work give polynomial algebraic circuits, thus placing them in VP. A large class of problems arises from treating the coefficients of one or more polynomials as the variables and then computing various expressions that are polynomial expressions of the coefficient themselves. Such examples include but are not limited to polynomial arithmetic, computing the resultant of two polynomials, computing the Taylor series up to a degree and the clearly related problem of computing the partial derivatives of a polynomial.

Such polynomials whose variables are actually the coefficients of different polynomials can also be used as so-called *test* polynomials, so named because they can be used to test for certain properties, that is their zero locus corresponds to polynomials for a specific property. A simple example is the discriminant of a quadratic polynomial already computed in Algorithm 1, which is zero precisely when the polynomial in question can be written as a perfect square, something that remains true for bivariate homogeneous polynomials of the form $ax^2 + bxy + cy^2$. Test polynomials play an important role in Geometric Complexity Theory, a research program that has a strong connection with Algebraic Complexity Theory and seems as the most promising approach for non-trivial lower bounds, see [21] for an overview of this exciting area of research.

Another class of polynomial arise from graph properties, such as whether a graph represents a tree or if all vertices have even degree, see Section 3.2 of [15] for more examples. Finally, a large class of polynomials originates from linear algebra and matrix multiplication, which although seem distinct at first, have a close connection, see Chapter 16 in [19] for a detailed analysis. In short, many problems in linear algebra boil down to computing the determinant, which in turn can be reduced to matrix multiplication.

It turns out that the determinant is of particular interest in terms of Algebraic Complexity Theory as well. The determinant of an $n \times n$ matrix A can be seen as a polynomial in the entries of the matrix, say $A_{i,j}$, $1 \leq i, j \leq n$. In particular, the determinant is given in closed form by the following so-called Leibniz formula

$$\det_n(A) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} A_{i,\sigma(i)}\,, \tag{2.43}$$

where $\mathcal{S}_n$ is the set of all permutations of $1, \ldots, n$ and $\operatorname{sgn}(\sigma)$ is the sign of the permutation σ (not to be confused with the sign of a real number).

From the above closed expression it is clear that the determinant has $n!$ monomials, degree n and n^2 variables. While the Leibniz formula does not give an efficient method to compute the determinant, it is possible to show that the determinant has polynomial sized circuits. While the common method of Gaussian elimination involves branching to find suitable pivots, it is possible to compute the determinant using a polynomial-sized circuit [6].

Not only is the determinant family $\det_n \in \mathsf{VP}$, it plays a very important role in Algebraic Complexity Theory. Before we are able to showcase that, we require to define

one additional complexity class, namely VNP. As the name might suggest, this class can be considered an analogue of NP, although one should be careful to not carry this analogy too far, as there are significant differences as we'll see in the following. It is sufficient to think of VNP as the class that it is likely to contain problems that we strongly believe to be hard to solve, i.e., that are not in VP. We now proceed with a precise definition of the class.

Definition 2.49 (VNP)**.** *The class VNP consists of all families of polynomials $\mathfrak{F}_n$ that satisfy the following condition:*

Let $N : \mathbb{N}^+ \to \mathbb{N}^+$ and $M : \mathbb{N}^+ \to \mathbb{N}^+$ be polynomials. Then $\mathfrak{F}_n \in$ VNP if there exists a family of polynomials $\mathfrak{G}_n$, such that $\mathfrak{G}_n \in$ VP, and for every $n \in \mathbb{N}^+$, $f_n \in \mathfrak{F}_n$ can be written as

$$f_n(X_1, \ldots, X_{N(n)}) = \sum_{W \in \{0,1\}^{M(n)}} g_n\left(X_1, \ldots, X_{N(n)}, W_1, \ldots, W_{M(n)}\right) , \tag{2.44}$$

where $g_n \in \mathfrak{G}_n$ and W_i refers to the i-th component of W, $1 \leq i \leq M(n)$.

Essentially the above says that if a family of polynomials $\mathfrak{F}_n$ is in VNP, then for each $n \in \mathbb{N}^+$, there exists a polynomial g_n with $M(n) + N(n)$ variables so that the family $\mathfrak{G}_n = \{g_n\,,\, n \in \mathbb{N}^+\}$ is in VP. Each summand corresponds to a vector in $\{0,1\}^{M(n)}$. The connection to NP becomes clear if we think W as playing a role similar to that of the certificate for NP problems. However, this also reveals where the analogy falls short: For NP, the existence of a single certificate is sufficient to guarantee membership in the class, whereas Equation (2.44) is more akin to summing over all possible certificates. It thus also bears some similarity to #P, although even this is not a perfect comparison, since not only that class is based on a uniform model, but also has a very precise definition in terms of counting accepting paths, specifically of nondeterministic Turing machines running in polynomial time.

There are many families of polynomials that belong in VNP. To begin with, since we like to think of families of polynomials in VP as being "easy" and those in VNP as "hard", it seems intuitive that we would like to have VP $\subseteq$ VNP, analogous to P $\subseteq$ NP. Indeed that is the case, which is not that difficult to prove. Given any VP family $\mathfrak{F}_n$, we can for each polynomial $f_n(X_1, \ldots, X_{N(n)}) \in \mathfrak{F}_n$ define the following polynomial

$$g_n\left(X_1, \ldots, X_{N(n)}, W_1\right) = W_1 f_n(X_1, \ldots, X_{N(n)})\,. \tag{2.45}$$

Clearly g_n requires just one additional multiplication gate compared to f_n, so the respective family $\mathfrak{G}_n$ is also a VP family. Furthermore, Equation (2.45) can be used to write f_n in the form of Equation (2.44). Specifically we have

$$\sum_{W_1=0}^{1} g_n\left(X_1,\ldots,X_{N(n)},W_1\right) = \sum_{W_1=0}^{1} W_1 f_n(X_1,\ldots,X_{N(n)}) = 0f_n(X_1,\ldots,X_{N(n)}) + 1f_n(X_1,\ldots,X_{N(n)}) = f_n(X_1,\ldots,X_{N(n)})\,.$$

Thus it is clear that $\mathfrak{F}_n$ belongs in VNP and since it was an arbitrary VP family of polynomials, it follows that VP $\subseteq$ VNP.

Of course, the definition of such a class would make little sense without problems that we consider hard to solve, i.e. not to be in VP. There are many families of polynomials that we know to belong to VNP but we believe not to be in VP, see Section 3.3 of [15] for examples. Perhaps the most important of those families is the *permanent*. The permanent is superficially similar to the determinant, in fact both are examples of a more general class of polynomials called immanants. Specifically, the permanent of a $n \times n$ matrix A with entries $A_{i,j}$ is given by the following formula

$$\mathrm{perm}_n(A) = \sum_{\sigma\in\mathcal{S}_n}\prod_{i=1}^{n} A_{i,\sigma(i)}\,, \tag{2.46}$$

where $\mathcal{S}_n$ is the set of all permutations of $1,\ldots,n$. A comparison with the determinant's definition in Equation (2.43) shows that the only difference is the sign of each permutation for each summand, which can be ± 1 for the determinant but is always 1 for the permanent. In fact, over fields of characteristic 2 the two are equivalent. This seemingly small difference makes for a world of change, in particular no polynomial algebraic circuit is known for the permanent. Not only that, but as we will see the permanent plays a central role in Algebraic Complexity Theory, for which also comparing it to the determinant is relevant.

From Equation (2.46), it is not clear that the permanent family is even a VNP family. The following formula, due to Ryser [76], not only shows the permanent to be in VNP but also is the most efficient known to date, apart from some non-asymptotic improvements.

$$\mathrm{perm}_n(A) = \sum_{T\subseteq\{1,\ldots,n\}} (-1)^{n-|T|}\prod_{i=1}^{n}\sum_{j\in T} A_{i,j} = \sum_{T\in\{0,1\}^n} (-1)^{n-\sum_{i=1}^{n} T_i}\prod_{i=1}^{n}\sum_{j=1}^{n} A_{i,j}T_j\,. \tag{2.47}$$

The first equality is the Ryser formula, while the second is the same formula written in the form of Equation (2.44). To represent every subset of $\{1,\ldots,n\}$, we simply have n indicator variables $T_1,\ldots,T_n$ which signify whether the corresponding element belongs in the set or not. Note that all possible variations of the variables correspond to the 2^n possible subsets. Then, we note that under that scheme we have that

$$|T| = \sum_{i=1}^{n} T_i \,, \tag{2.48}$$

$$\sum_{j \in T} A_{i,j} = \sum_{i=1}^{n} A_{i,j} T_j \,. \tag{2.49}$$

Substituting Equation (2.48) and Equation (2.49) allows us to then obtain the second equality of Equation (2.47). Now note that the family defined by

$$g_n(A,T) = (-1)^{n - \sum_{i=1}^{n} T_i} \prod_{i=1}^{n} \sum_{j=1}^{n} A_{i,j} T_j \,,$$

where $T = (T_1, \ldots, T_n)$, can be computed by an $O(n^2)$ circuit. Note that the expression is indeed a polynomial since the exponential dependence on T_i only serves as a shorthand to determine the value of the sign. In any case, since we are only interested in evaluations such that $T_i = 0$ or $T_i = 1$, for such values we have that $(-1)^{-\sum_{i=1}^{n} T_i} = \prod_{i=1}^{n} (-2T_i + 1)$, which we can substitute in the above expression. In any case, we conclude that g_n defines a VP family, as required.

We have already seen that $\mathsf{VP} \subseteq \mathsf{VNP}$. Similar to classical complexity theory, the central question of Algebraic Complexity Theory is whether the inclusion is strict. Also similar to the classical case, most experts in the field believe this to be the case. This conjecture was originally posed by Valiant and thus bears his name.

Conjecture 2.50 (Valiant's conjecture)**.** *$\mathsf{VP} \neq \mathsf{VNP}$.*

Note that in Definition 2.45, we noted that the underlying field is important in the definition of the model, thus also in that of the class. Technically VP and VNP are not a single class, but for each field we could potentially obtain a different class. For two fields $\mathbb{K}$ and $\mathbb{L}$, such that there exists a field element $a \in \mathbb{K}$ but $a \notin \mathbb{L}$, it is trivial that there exist families of polynomials in $\mathsf{VP}_{\mathbb{K}}$ but not $\mathsf{VP}_{\mathbb{L}}$, for example $\mathfrak{F}_n = \sum_{i=1}^{n} a x_i$. Nevertheless, when it comes to Valiant's conjecture, it can be shown that solving the question for one field can yield the same answer for other fields, under specific conditions. In particular, for algebraically closed fields, it turns out that only the characteristic of the field is important and not the particular field itself. The same is true for purely transcendental field extensions of infinite fields as well as for finite algebraic extensions. In particular note that $\mathbb{C}$ is an algebraic extension of $\mathbb{R}$ of degree 2. Section 4.1 in [15] investigates this dependence on the field, in particular Corollary 4.2 contains the above results. Furthermore, under standard assumptions, it has been shown that if Valiant's conjecture does not hold, then the polynomial hierarchy collapses [16], strengthening the belief in this hypothesis via implying an unlikely outcome in classical complexity.

We have already mentioned that most experts believe $\mathrm{perm}_n \notin \mathsf{VP}$. We would like to make this formal, elevating the permanent to a role similar to that the satisfiability problem plays for NP. Formally, we would like to introduce a notion of reduction that allows us to reduce the computation of one problem to that of another. This will also allow us to define complete problems for a class, similar to classical complexity, meaning problems to which all other problems in the class reduce to.

The notion of reduction in Algebraic Complexity Theory is rather simple and essentially amounts to substitution. We say that a polynomial f is a projection of another polynomial g, if it can be obtained from g by renaming variables or setting variables to any constant from the underlying field $\mathbb{K}$. We've already seen something that resembles this in the definition of VNP in Equation (2.44), where each summand is a projection of a VP polynomial. Note however that the entire sum is not considered a projection, since we only allow a single substitution to occur. One can contrast this with many-one and Turing reductions in classical complexity theory, with projections being more akin to the former. Finally, similar to polynomials, we can have that an entire family of polynomials $\mathfrak{F}_n$ is a projection of another family $\mathfrak{G}_n$, meaning that for each $n \in \mathbb{N}^+$, f_n is a projection of g_n. The formal definition follows

Definition 2.51 (Polynomial projections)**.** *A polynomial $f(X_1, \ldots, X_N)$ is called a* projection *of another polynomial $g(Y_1, \ldots, Y_M)$ if there exists $a \in (\{X_1, \ldots, X_N\} \cup \mathbb{K})^M$, where $\mathbb{K}$ is the underlying field, such that we have*

$$f(X_1, \ldots, X_N) = g(a_1, \ldots, a_M) .$$

We then write $f \leq_p g$.

Similarly, a family of polynomials $\mathfrak{F}_n$ is called a polynomial projection *of another family of polynomials $\mathfrak{F}_n$ if there exists a polynomially bounded $t : \mathbb{N}^+ \to \mathbb{N}^+$ such that for every $n \in \mathbb{N}^+$,$f_n \in \mathfrak{F}_n$ is a projection of $g_{t(n)} \in \mathfrak{G}_n$, that is there exists $a \in (\{X_1, \ldots, X_N\} \cup \mathbb{K})^{M(t(n))}$ such that*

$$f(X_1, \ldots, X_N) = g_{t(n)}\left(a_1, \ldots, a_{M(t(n))}\right) .$$

Similarly as above, we then write $\mathfrak{F}_n \leq_p \mathfrak{G}_{t(n)}$.

Having defined a notion of reduction, we can now introduce the fundamental notion of hardness and completeness. The definitions are similar to classical complexity theory.

Definition 2.52 (Hard and complete problems)**.** *Let C be an algebraic complexity class and let $\mathfrak{F}_n$ be a polynomial family. We say that $\mathfrak{F}_n$ is C-hard if every polynomial family $\mathfrak{G}_n \in \mathsf{C}$ is a projection of $\mathfrak{F}_n$, that is*

$$\forall \mathfrak{G}_n \in \mathsf{C},\ \mathfrak{G}_n \leq_p \mathfrak{F}_n .$$

If in addition $\mathfrak{F}_n \in \mathsf{C}$, then we say that $\mathfrak{F}_n$ is C-complete.

A central result due to Valiant states that the permanent is indeed complete for VNP. In particular we have

Theorem 2.53 (Completeness of the permanent, [91])**.** *The permanent family $perm_n$ is VNP-complete for fields of every characteristic except 2.*

A detailed proof of the above statement can also be found in Section 2.2. of [15]. The already mentioned Section 3.3 of the same work contains more examples of VNP-complete families. For a long time, generic VP-complete polynomials were known, but they were based on technical constructions built for this specific purpose. [27] proposed a number of "natural" complete problems for VP, in the sense that they stem from problems not

directly related to the definition of the class. In particular, these families of polynomials relate to graph homomorphisms. A subsequent work, [22], provides similar complete polynomials for other important classes in Algebraic Complexity Theory that we will mention briefly below, namely VF and VBP, as well as variants of such polynomials that are VNP-complete.

The avid reader will have noticed that despite contrasting the complexity of the determinant with the permanent in the text, we have not mentioned the determinant as a VP-complete problem. That is because the determinant family of polynomials is not known to be VP-complete, neither an impossibility result proven. Rather, it is known to be complete for a larger class, called VQP. This class is defined similarly to VP, with the exception that we allow the size of the circuit computing a VQP family to be quasi-polynomial bounded instead of simply polynomially bounded. Recall that quasi-polynomial means exponential with an exponent depending polylogarithmically in n. Thus, we may define VQP similar to Definition 2.48, with the exception of changing Condition (3). The formal definition follows.

Definition 2.54 (VQP)**.** *The class VQP consists of all families of polynomials $\mathfrak{F}_n$ that satisfy the following conditions:*

(1) All polynomials in $\mathfrak{F}_n$ have polynomially many variables in n. That is, there exists a polynomial $v(n) : \mathbb{N}^+ \to \mathbb{N}^+$ such that for every $n \in \mathbb{N}^+$, the number of variables of f_n is bounded by $v(n)$, i.e., $f_n(X) = f_n(X_1, \ldots, X_{N(n)})$ such that $N(n) \leq v(n)$, $n \in \mathbb{N}^+$.

(2) All polynomials in $\mathfrak{F}_n$ have polynomially bounded degree in n. That is, there exists a polynomial $d(n) : \mathbb{N}^+ \to \mathbb{N}^+$ such that for every $n \in \mathbb{N}^+$, the degree of f_n is bounded by $d(n)$, i.e., $\deg(f_n) \leq d(n)$, $n \in \mathbb{N}^+$.

(3) All polynomials in $\mathfrak{F}_n$ have circuits of quasi-polynomially bounded size. That is, for each polynomial f_n in $\mathfrak{F}(n)$, there exists a circuit C_n that computes f_n of size $s(n)$ and a constant $c \in \mathbb{N}^+$, such that $s(n) \leq 2^{\log^c n}$, $n \in \mathbb{N}^+$.

Similarly, we can define quasi-polynomial projections. The only difference to Definition 2.51 is that the number of variables of the polynomials we project can be quasi-polynomially large. Thus we have the following:

Definition 2.55 (Quasi-polynomial projections)**.** *We call a family of polynomials $\mathfrak{F}_n$ a* quasi-polynomial projection *of another family of polynomials $\mathfrak{F}_n$ and write $\mathfrak{F}_n \leq_{qp} \mathfrak{G}_{t(n)}$ if there exists a quasi-polynomial bounded function $t : \mathbb{N}^+ \to \mathbb{N}^+$, i.e., $t(n) \leq 2^{\log^c n}$ for some $c \in \mathbb{N}^+$, such that for every $n \in \mathbb{N}^+$,$f_n \in \mathfrak{F}_n$ is a projection of $g_{t(n)} \in \mathfrak{G}_n$, that is there exists $a \in (\{X_1, \ldots, X_N\} \cup \mathbb{K})^{M(t(n))}$ such that*

$$f\left(X_1, \ldots, X_N\right) = g_{t(n)}\left(a_1, \ldots, a_{M(t(n))}\right) .$$

That the determinant is VQP-complete was already known to Valiant, as attested in the same work that contains the proof of the VNP completeness of the permanent.

Theorem 2.56 (Completeness of the determinant, [91])**.** *The determinant family det_n is VQP-complete under quasi-polynomial projections.*

It is clear that $\mathsf{VP} \subseteq \mathsf{VQP}$ by definition, in fact it is possible to show that this separation is strict, that is, VQP is strictly larger. Thus, it would suffice to separate VQP from VNP to prove Valiant's conjecture, although the former statement is weaker than the latter. Since both classes have complete polynomials that not only are defined outside the context of Algebraic Complexity Theory but also bear close resemblance, they offer a way to state a weaker version of Valiant's conjecture.

Conjecture 2.57 (Valiant's weak conjecture)**.** *$\mathsf{VNP} \not\subseteq \mathsf{VQP}$. Due to Theorem 2.53, this is equivalent to $perm_n \notin \mathsf{VQP}$ for all fields, except those of characteristic 2. Furthermore, using Theorem 2.56 and under the same characteristic restriction, this is equivalent to the permanent not being a quasi-polynomial projection of the determinant that is*

$$perm_n \not\leq_{qp} det_n \,. \tag{2.50}$$

Especially the last formulation often leads to this version of the conjecture being referred to as the "*permanent versus determinant*" question. This formulation is undoubtedly the central question of Algebraic Complexity Theory and is central to the aforementioned Geometric Complexity Theory program. In fact, a central advantage of this formulation is being able to cast the question as an algebraic geometry problem involving the permanent and the determinant, with no mention to any complexity notions whatsoever, while also allowing the use of well-studied tools from several branches of mathematics.

Before closing our short introduction into Algebraic Complexity Theory, we draw the attention of the reader to two issues. The first is a surprising result that showcases the difference between Algebraic and Classical Complexity and should serve as a caution between drawing analogies too far. Note that the first formulation of Valiant's weak conjecture, Conjecture 2.57, is stated as an obstruction of an inclusion rather than an inequality. That is because not only are the classes known to be unequal, it is known that there exist VQP families of polynomials that are not in VNP, i.e., $\mathsf{VQP} \not\subseteq \mathsf{VNP}$. A proof of this statement, which also shows the aforementioned separation $\mathsf{VP} \subseteq \mathsf{VQP}$, can be found in Section 8.2 of [15]. In the classical world, the "analogous" classes, QP and NP are also known to not be equal, i.e., $\mathsf{QP} \neq \mathsf{NP}$, as it can be easily seen that the former class is closed under quasi-polynomial (many-one) reductions, while the latter isn't. Then, an analogous result of Conjecture 2.57, that is $\mathsf{NP} \not\subseteq \mathsf{QP}$ would be implied by the well-known exponential time-hypothesis [40]. However, proving that $\mathsf{QP} \not\subseteq \mathsf{NP}$ would also solve at least one major open question in Classical Complexity by the following argument.[5] Suppose a language L exists such that $L \in \mathsf{QP} \setminus \mathsf{NP}$. Then one of the following must be true:

(1) L is NP-hard, thus $\mathsf{NP} \subseteq \mathsf{QP}$, which would violate the exponential time hypothesis.

(2) $L \in \mathsf{SPACE}\left(\log^c n\right)$, the class of polylogarithmic space. Since $L \notin \mathsf{NP}$ and $\mathsf{P} \subseteq \mathsf{NP}$, it follows that $L \notin \mathsf{P}$. However, since $\mathsf{SPACE}\left(\log^c n\right) \subsetneq \mathsf{PSPACE}$, we have that $L \in \mathsf{PSPACE}$. Therefore L separates P from PSPACE, i.e., $\mathsf{P} \subsetneq \mathsf{PSPACE}$, which would be a significant breakthrough.

[5] The argument is folklore, see the online discussion in [36] for more details

(3) Similarly, if $L \in \mathsf{QP} \setminus \mathsf{SPACE}(\log^c n)$, this would imply that there exists L' such that $L' \in \mathsf{P} \setminus \mathsf{L}$ via a standard padding argument, where $\mathsf{L} = \mathsf{SPACE}(\log n)$, which would also answer a major open question.

The first condition could be true independently of the other two, while clearly the last two are mutually exclusive.

That we already have obtained such a result in the algebraic setting strengthens a belief that many researchers hold which is a major motivation for the study of Algebraic Complexity Theory. Namely, that since polynomials are well-studied and understood functions that behave in a more predictable manner compared to Boolean functions, the equivalent questions in Algebraic Complexity Theory should be easier to solve than their Boolean (i.e., classical) counterparts. We refer the interested reader to Chapter 4 of [15], which examines closely the relation between the two theories and how the relevant conjectures are connected. In particular, under standard assumptions, it can be seen that the non-uniform version of $\mathsf{P} \neq \mathsf{NP}$ implies the algebraic version, making the implication formal. That is we have

$$\mathsf{P/poly} \neq \mathsf{NP/poly} \Longrightarrow \mathsf{VP} \neq \mathsf{VNP}$$

for fields of characteristic 0 (including in particular $\mathbb{R}$ and $\mathbb{C}$) and finite fields.

Before closing this section, we briefly present some additional complexity classes from Algebraic Complexity Theory. These are classes obtained from imposing certain restrictions on the circuit. A large number of classes is obtained by imposing a restriction on the shape, i.e., the structure of the underlying graph of the circuit. For example, recall that all graphs representing circuits are directed acyclic graphs (DAG). Note that DAGs are acyclic with respect to the direction of the edges, but if we forget the direction of every edge, it is possible that in the induced undirected graph, cycles appear. For example, in Figure 2.2, the subcircuit computing b^2 induces an undirected cycle. If we specifically disallow undirected cycles, that is, demand that the underlying structure of the graph is a tree, we obtain the restricted class of circuits called *formulas*. It is easy to see that formulas correspond precisely to circuits that do not re-use any calculations, since if every gate has out-degree 1, by removing that edge we isolate the subcircuit rooted at the gate from the rest of the graph. Therefore no cycles could exist, since the existence of an undirected cycle would imply the existence of at least one alternate path. The same definition is applicable for straight-line programs, where formulas correspond to straight-line programs where each computed line only appears once after its computation. Note that each formula corresponds to a closed form for a polynomial that has as many operations as the formula has gates, hence the name. On the other hand, for circuits the number of operations would be more, since we have to replace more than one occurence of at least one gate.

By restricting ourselves to formulas in the definition of VP, Definition 2.48, we obtain the class VF of polynomially sized formulas (with the additional restrictions in the number of variables and the degree). This is clearly a subclass of VP, that is $\mathsf{VF} \subseteq \mathsf{VP}$. The inclusion is widely believed to be strict, however a proof eludes us and current lower bounds are far from being close to an answer. For example, the best lower bound for the formula size for the determinant is $\Omega\left(n^3\right)$ [47]. On the other hand, the best known formula for the determinant has size $n^{O(\log n)}$.

Yet another class that can be obtained by a less severe restriction on the shape of the circuit is VBP. Its name comes from the model of *algebraic branching programs* (ABPs), which is yet another model of algebraic computation. The model is rather simple to describe, which we now do by its formal definition.

Definition 2.58 (Algebraic branching program)**.** *Let $G = (V, E)$ be an acyclic graph with node set V and edge node E. Also let $X = (X_1, \ldots, X_n)$ be the set of indeterminates and $\mathbb{K}$ the underlying field. The graph has the following structure: There exist two special nodes $s, t \in V$, s being called the source and t the sink. Every edge $e \in E$ is labeled with a homogeneous linear polynomial in $\mathbb{K}[X]$, that is a label $L_e(X) = \sum_{i=1}^{n} c_{e,i} X_i$. Each path $p = (e_1, \ldots, e_d)$ from s to t in the graph computes a polynomial that is the product of the edges on the path, that is*

$$f_p(X) = \prod_{i=1}^{d} L_e(X)\,.$$

Finally, the polynomial f computed by the ABP is simply the sum of all polynomials f_p for each st-path on G, that is

$$f(X) = \sum_{p:s \rightsquigarrow t} f_p(X)\,.$$

By the addition of edges labeled simply by 1, we can assume that the ABP is "layered", with nodes at layer i being the i-th node in a st-path, thus s is at layer 0 and t at layer d. Furthermore, note also that since each label is a linear polynomial, $d \geq \deg(f)$.

One way to define VBP is as the class of families of polynomials that can be computed by a family of polynomially-sized ABPs, where the size is defined as the number of nodes in the underlying graph. Another measure of complexity for ABPs that we will not study further in this work but allows us to pose certain interesting questions is the width of the ABP, defined as the maximum number of nodes in any single layer of an ABP.

VBP also has a close connection to linear algebra problems. In fact, it can be shown that not only the determinant is in VBP via an inventive combinatorial characterization [63], but it is indeed complete for the class [91]. An alternative characterization of ABPs is in terms of iterated (i.e., repeated) matrix multiplication, which should not be surprising given the aforementioned connection between matrix multiplication and the determinant.

Finally, it is also possible to define VBP in terms of circuit shape. In particular, we can capture the class by restricting our circuits to the so-called *weakly-skewed* circuits. Weakly-skewed circuits satisfy the following property on the multiplication gates of the circuit: We assume that every multiplication gate M has in-degree 2 with children say L and R. Then a circuit is weakly-skewed if for every such gate M, L and R are disjoint and removing M results in the underlying graph splitting into at least two components. In other words, the subcircuits rooted at the children of every multiplication gate must be disjoint and results from at most one of them can be re-used by the rest of the circuit. Clearly, formulas are a special type of weakly-skewed circuits, since we allow no reuse of results at all, thus we can easily obtain that $\mathsf{VF} \subseteq \mathsf{VBP}$. Furthermore, clearly weakly-skewed circuits are a restriction of general algebraic circuits. All together, we have the following relation for the class of efficiently computable families of polynomials

$$\mathsf{VF} \subseteq \mathsf{VBP} \subseteq \mathsf{VP}\,. \tag{2.51}$$

All the above inclusions are suspected to be strict, however a proof of this statement eludes the community thus far.

It is worth mentioning one last type of restriction applied to classes. When defining algebraic circuits, Definition 2.45, we allow constants from the underlying field $\mathbb{K}$. This is a gross oversimplification compared to practical computation, since if say, $\mathbb{K} = \mathbb{R}$, this implies instant access to any real number with arbitrary precision at the cost of a single input gate. We can instead restrict circuits to be *constant-free*, where only the constants 0,1 and -1 are allowed. This is clearly much more restrictive than general circuits, since every other constant used in the circuit must be computed using gates. Perhaps one would doubt whether this is relevant in our model, since by definition all monomials of the permanent have coefficient equal to 1. However, note that this implies nothing for the intermediate results a circuit might use to compute the permanent (or any other polynomial), which could potentially be very large and simply cancel out before obtaining the final result.

Allowing only for constant-free circuits, we can then obtain the constant-free equivalent of any algebraic complexity class, in particular we obtain VP^0 from Definition 2.48 and similarly VNP^0 from Definition 2.49. The same argument as used in the original setting can show that $\mathsf{VP}^0 \subseteq \mathsf{VNP}^0$, although on the other hand the permanent is not proven to be VNP^0-complete since the original proof crucially uses the constant $\frac{1}{2}$. However, one can show the following: If the permanent has polynomial constant-free circuits, that is $\mathrm{perm}_n \in \mathsf{VP}^0$, then for every family $\mathfrak{F}_n \in \mathsf{VNP}^0$ there exists a polynomially bounded function $p(n)$ such that the family $2^{p(n)}\mathfrak{F}_n$ is in VP^0 [50]. Using the above, we can then pose the analogue of Valiant's weak conjecture in the constant-free setting. Furthermore, note that complete polynomials for VNP^0 exist, for example the Hamiltonian cycle polynomials HC_n, which are defined similarly to the permanent but instead of all permutations, the sum is over only cyclic permutations.

As we've already remarked, the computation of constants plays an important role for the above classes. Furthermore, note that the bound on the degree in Definition 2.48 also applies to constants in the constant-free circuits. Formally, we use the notion of "formal degree", which is essentially a black-box upper bound on the degree of a polynomial computed by a circuit, where we assume that each multiplication gate may double the degree. Since the bound is "black-box", this is still the case even if both arguments of the multiplication gate are constants, or if cancellations occur. Thus, it is conceivable that for the same characteristic, there exist polynomially bounded circuits for a family of polynomials, such as the permanent, however these utilize large-enough constants so that it is not possible to compute them, or if necessary, replace them, for a field of the same characteristic but smaller size. In other words, it is possible that we have polynomially bounded in size constant-free circuits for the permanent, but due to violating the formal degree condition, the permanent family is not in VP^0. However, as we've already noted, it is possible however to pose the same questions in the constant-free versions, with small alterations, by either multiplying the polynomials with a constant depending on n or using a different complete polynomial than the permanent.

2.3.3 The real tau conjecture and relevant results

Having introduced the basic notions in Algebraic Complexity Theory, we are now ready to present a relatively recent result in the area which is of particular interest to this work. This result associates Valiant's hypothesis with a seemingly unrelated hypothesis about the number of real roots of a class of univariate polynomials of a specific form, hence our focus on the roots of polynomials.

While Complexity Theory can be broadly classified as the study of computational classes and their relation in the model at hand, it quite often boils down to proving lower bounds, since algorithms are often the easiest way to prove upper bounds and thus prove membership in complexity classes. In turn, proving non-membership in a class seems to be much harder, as both mathematical intuition and the research record thus far show. The mathematical intuition comes from the simple fact that while a single algorithm suffices to prove membership in a class, proving non-membership via a lower bound requires proving that none of the infinitely many algorithms that compute the problem correctly violates the bound. Thus, while often an algorithm exploits a specific property to attack a problem, a lower bound requires showing that none of those "properties" are sufficient to violate it, therefore requiring a deep understanding of the problem at hand. Complexity theorists have devised a number of methods to cope with this difficulty. One such way is to alter the model in question appropriately. Such an alteration often restricts the power of the model, such as for example restricting ourselves to algebraic formulas or constant-free circuits, with the hope that the added restrictions limit the power of the model significantly enough to allow for a proof. Other times however, seemingly strengthening the model or changing it all together can also make the question easier, as it is (seemingly) the case for Valiant's conjecture stated in Conjecture 2.50, which we have already seen that is implied by the non-uniform version of $\mathsf{P} \neq \mathsf{NP}$, under reasonable assumptions.

A different method of coping with difficult research questions is to build a "web" of conditional results, starting from a central hypothesis or conjecture we wish to answer. By the implications between the different results, we can better estimate our belief on the statement of the conjecture, since it being true (or false) would imply a number of different results, for which we may have better intuition than for the original problem, or even be able to solve them. It can also be seen as a self-referential paradigm, since the hardest problem is reduced to a hopefully easier one. Such examples abound in classical Complexity Theory, with $\mathsf{P} \neq \mathsf{NP}$ playing the role of the central conjecture. For example, the simple proof of a large number of problems as NP-complete, many of which have been studied extensively but none have been found to have polynomial time algorithms, strengthens the belief that the two classes are indeed separate. Similarly, believing that $\mathsf{P} = \mathsf{NP}$ implies that P is actually equal to the polynomial-time hierarchy, i.e., the hierarchy collapses, which implies a vast amount of classes that seem to be even more difficult than NP, must also have polynomial time solutions, to name a few examples. The collapse of the polynomial hierarchy itself is widely used as a result implied by other results, due to the belief in the impossibility of such a collapse.

Perhaps one of the most ingenious ideas in this direction are results that connect upper bounds, i.e., the existence of an algorithm, with a lower bound. Such results can use the belief that an algorithm does not exist, such as the exponential time hypothesis,

to show a lower bound for some other, seemingly unrelated problem. Many such results have been shown in recent years in the emerging field of Fine-Grained Complexity Theory, see [12] for an overview.

On the opposite direction, we may link a lower bound with the existence of an algorithm. A classic such result is due to Kabanets and Impagliazzo [44], that links an algorithm for the Polynomial Identity Problem (PIT) with lower bounds for either nondeterministic Turing machines running in exponential time, namely that $\mathsf{NEXP} \not\subset \mathsf{P/poly}$ or that the permanent family requires superpolynomial algebraic circuits, which implies $\mathsf{VNP} \not\subseteq \mathsf{VP}$. This means that derandomizing PIT implies superpolynomial lower bounds for Boolean or algebraic circuits.

But what exactly is the Polynomial Identity Testing problem? While many variations exist, in its simplest form the PIT problem is defined as follows: Given a representation of a polynomial, in our case, an algebraic circuit, decide whether it computes the zero polynomial. As we've already remarked earlier in this section, see, e.g., Proposition 2.3 and the discussion before it, this might require working in larger fields than the original, however such technicalities do not overly concern us in our work since we work over $\mathbb{R}$ or $\mathbb{C}$.

While the result of Kabanets and Impagliazzo is perhaps the most well known, several results in that vein preceded it and followed it. One such result is the focus of our work and associates the existence of an efficient type of algorithm for a special case of the PIT problem with lower bounds for the permanent. This result is known as the *real τ-conjecture* and its statement is due to Koiran [51]. We will now present the most important details of the statement of the conjecture, its origins and the proof idea, while referring the reader to the original work for the complete proof.

The original τ-conjecture was posed by Shub and Smale in [80], see also Problem 4 in [82]. It states the following: Let $f \in \mathbb{Z}[x]$ be a univariate polynomial with integer coefficients. Let $N_{\mathbb{Z}}(f)$ denote the number of *distinct* integer roots of f and $s(f)$ its constant-free algebraic circuit complexity. Then if $N_{\mathbb{Z}}(f)$ is polynomially bounded by $s(f)$, the classes P and NP over $\mathbb{C}$ in the BSS model are distinct, that is we have

$$\exists c \in \mathbb{Z}^+ : N_{\mathbb{Z}}(f) \leq (s(f))^c \Longrightarrow \mathsf{P}_{\mathbb{C}} \neq \mathsf{NP}_{\mathbb{C}} \,. \tag{2.52}$$

Recall that the BSS model is essentially a Turing machine over a field, in this case $\mathbb{C}$, with each cell of the Turing machine being able to store an element from the field with arbitrary precision, thus it can be viewed as the uniform analog of algebraic circuits. In other words, the above says that if the constant-free complexity of an arbitrary integer polynomial is lower bounded by the c-th root of its number of distinct integer roots, for some $c \in \mathbb{Z}^+$, the two classes are separate. Bürgisser further showed that the t-conjecture implies super-polynomial lower bounds for the permanent.

Koiran modified the τ-conjecture in several related ways. First, the class of polynomials whose roots must be bounded are limited to polynomials of a specific form. Secondly, while the bound technically depends on the complexity of the polynomial, this can be replaced with parameters from the form of the polynomial, making the statement less convoluted. We now shortly give some necessary definitions to present the result.

We begin with the class of $\mathsf{SPS}_{s,e}$ polynomials. Those are $\mathbb{Z}[x]$ polynomials that satisfy certain conditions, namely:

Definition 2.59 (SPS polynomials)**.** *The class $\mathsf{SPS}_{s,e}$ is the class of polynomials from $\mathbb{Z}[x]$ which satisfy the following conditions:*

(1) They can be written in the form $\sum_i \prod_j f_{i,j}$, where $f_{i,j} \in \mathbb{Z}[x]$.

(2) The sum of monomials over all $f_{i,j}$ is bounded by s.

(3) Every coefficient of the polynomials $f_{i,j}$ can be written as the difference of two nonnegative integers that have at most s 1s in their binary representation.

(4) The degree of the polynomials $f_{i,j}$ is bounded by e.

(5) The coefficients are bounded in absolute value by 2^e.

We also recall the definition of the following complexity class:

Definition 2.60 (Counting hierarchy)**.** *Let A be a complexity class. We define the operator $\mathfrak{C}.$ as follows: The class $\mathfrak{C}.\mathsf{A}$ contains all languages L such that there exists a language $L' \in \mathsf{A}$ and a polynomial $p(n)$ satisfying*

$$x \in L \iff \#\{y \in \{0,1\}^{p(|x|)} : (x,y) \in L'\} \geq 2^{p(|x|)-1}\,.$$

That is, membership in a $\mathfrak{C}.\mathsf{A}$ language L means there is a L' verifier in the class A so that at least half the certificates verify membership in the language.

We can then use this operator to define the levels of the counting hierarchy CH recursively. The 0-th level is defined simply as $\mathsf{C}_0\mathsf{P} = \mathsf{P}$ and then we have recursively $\mathsf{C}_{i+1}\mathsf{P} = \mathfrak{C}.\mathsf{C}_i\mathsf{P}$.

The counting hierarchy is defined as the union over all levels i, that is $\mathsf{CH} = \cup_{i\geq 0}\mathsf{C}_i\mathsf{P}$. We may also define a non-uniform version of the class with polynomial advice, namely $\mathsf{CH/poly}$.

It is easy to see from the definition how CH compares with other known complexity classes. Namely we have

$$\mathsf{PH} \subseteq \mathsf{CH} \subseteq \mathsf{PSPACE}\,.$$

We may now define a different class of polynomials, which will play the role of efficiently computable polynomials, under a very generous complexity theoretic assumption. In particular, we have

Definition 2.61 (Algebraic number generators)**.** *An* algebraic number generator *is a sequence $\mathfrak{F}_i, i \geq 1$ of nonzero univariate polynomials $f_i(x) = \sum_a A(a,i)x^a$ such that for some integer constant $c \geq 1$:*

(1) The exponents a range from 0 to i^c, $0 \leq a \leq i^c$.

(2) $A(a,i)$ is a sequence of integers whose absolute value is bounded by 2^{i^c}, that is $|A(a,i)| \leq 2^{i^c}$.

(3) The coefficients of f_i can be computed bitwise in $\mathsf{CH/poly}$ in the following sense: The language $L(f) = \{(a,i,j,b) : \text{the } j\text{-th bit of } A(a,i) \text{ is equal to } b\}$ is in $\mathsf{CH/poly}$.

Our goal is to show that the existence of a sufficiently small hitting set for $\mathsf{SPS}_{s,e}$ polynomials implies that the permanent is hard to compute. We will make both the notion of sufficiently small and hard to compute precise in the following. We do not present every detail of the proof but only a comprehensive proof sketch. The interested reader can refer to [51] for all the details.

The entire proof is by contradiction, specifically we will assume that the permanent is easy to compute. In particular, we will assume that the permanent has constant-free polynomial size circuits, that is $\text{perm}_n \in \mathsf{VP}^0$. This assumption is powerful and allows us to use some strong claims, that enable proving the result. In particular, we have:

Lemma 2.62 (Lemma 2 in [51], Lemma 2.6 and 2.13 in [17])**.** *If $\text{perm}_n \in \mathsf{VP}^0$, then $\mathsf{CH/poly} = \mathsf{P/poly}$.*

The above lemma allows us under the assumption to compute bitwise the coefficients of an algebraic number generator in polynomial non-uniform time.

Under the same assumption that $\text{perm}_n \in \mathsf{VP}^0$, we can state a result about the complexity of every family of VNP^0 polynomials. Due to the use of constants, we stop short of proving they are in VP^0, but we can state something quite similar.

Theorem 2.63 (Theorem 4.3 in [50])**.** *Assume that $\text{perm}_n \in \mathsf{VP}^0$. For every family $\mathfrak{F}_n \in \mathsf{VNP}^0$, there exists a polynomially bounded function $p(n)$ such that the family $2^{p(n)}\mathfrak{F}_n$ is in VP^0.*

The factor $2^{p(n)}$ is due to the use of the constant $\frac{1}{2}$ in the completeness proof of the permanent for VNP. Also note that for each polynomial, the factor $2^{p(n)}$ multiplies all coefficients of the polynomial and only depends on n and not the variables of the polynomial X , so the function computed remains a polynomial in X.

We will also require the following useful criterion for membership in VNP^0. This is the constant-free version of Valiant's criterion, which states that if the coefficients of a polynomial are computable in $\#\mathsf{P/poly}$, then that polynomial is in VNP. Valiant's criterion was first introduced unsurprisingly by Valiant in [91]. A more modern statement can be found in Proposition 2.20 of [15]. We now state the constant-free version:

Proposition 2.64 (Valiant's criterion, constant-free version (Lemma 1, [51]))**.** *Suppose that $p(n)$ is a polynomially bounded function and that $g : \mathbb{N} \times \mathbb{N} \to \mathbb{Z}$ is such that the map $1^n 0 j \mapsto g(j, n)$ is in the complexity class $\mathsf{GapP/poly}$. Then the family $\mathfrak{F}_n$ of multilinear polynomials defined by*

$$\mathfrak{F}_n(X_1, \ldots, X_{p(n)}) = \sum_{j \in \{0,1\}^{p(n)}} g(j, n) X_1^{j_1} \cdots X_{p(n)}^{j_{p(n)}} \tag{2.53}$$

is in VNP^0, where j_k denotes the bit of weight 2^{k-1} of j.

As a reminder, the class GapP contains all functions that can be written as the difference of two $\#\mathsf{P}$ functions and here we are interested in its non-uniform version. Compared to the original Valiant's criterion, we note that instead of $\#\mathsf{P}/\operatorname{poly}$, the coefficient may be computed in the class $\mathsf{GapP/poly}$, which trivially contains $\#\mathsf{P}/\operatorname{poly}$, while in addition the polynomials may only be multilinear. The input to computing the coefficients $1^n 0 j$ is a technical condition, which means the input to g is (n, j), however

n is given in unary so the input is large enough and the 0 serves to separate the n ones from the rest of the input.

To achieve the contradiction, we will define a family of polynomials $\mathfrak{G}_n$ based on an algebraic number generator $\mathfrak{F}_i$, so that such for each n, the polynomial g_n vanishes on every root of the polynomials f_i for $1 \leq i \leq 2^n$. We will also define a hitting set $\mathcal{H}_m$ that comprises of all the roots of the polynomials f_i for $1 \leq i \leq m$. Then we show that the family of polynomials $\mathfrak{G}_n$ is in $\mathsf{SPS}_{s,e}$. Thus, if we assume that for a carefully selected value of m,s and e, $\mathcal{H}_m$ is a hitting set for $\mathsf{SPS}_{s,e}$, we have reached a contradiction, since by construction any polynomial in the family $\mathfrak{G}_n$ will vanish on all points of $\mathcal{H}_m$, while not being identically zero, contrary to the definition of a hitting set.

Most of the work needed to be done will be to show that the family of polynomials $\mathfrak{G}_n$ is in $\mathsf{SPS}_{s,e}$. We begin with a proposition that shows how to obtain a $\mathsf{SPS}_{s,e}$ family of polynomials from a VP^0 family of polynomials via substitution. Specifically, we have:

Proposition 2.65 (Proposition 1 of [51])**.** *Let $\mathfrak{F}_n(Y, Z)$ be a VP^0 family of multilinear polynomials with Y and Z two tuples of variables of length cn each, for some constant c. Let $f_n^s(x)$ be the univariate polynomial obtained by f_n by the following substitution*

$$f_n^s(x) = f_n(x^{2^0}, x^{2^1}, \ldots, x^{2^{cn-1}}, 2^{2^0}, 2^{2^1}, \ldots, 2^{2^{cn-1}}) \,. \tag{2.54}$$

Then f_n^s is a $\mathsf{SPS}_{s,e}$ polynomial for $s = n^{O(\sqrt{n}\log n)}$ and $e = 2^{O(n)}$.

This proposition relies on a simple characterization for the complexity of depth 4 constant-free circuits for VP^0 families of polynomials, which can be found in Corollary 1 of [51]. Furthermore, the structure of the polynomials f_n^s can be made more precise.

We saw above that the definition of algebraic number generators requires its coefficients to be bitwise computable in $\mathsf{CH/poly}$. We now introduce two different notions of computation in the above class, one also relating to the bitwise computation of coefficients and one to evaluation of polynomials.

The first notion we will examine relates to the computation of coefficients and its main difference to the notion we used to define algebraic number generators is a technical one, namely we allow one argument to be written in unary instead of binary. Furthermore, this notion is much more general. We thus define $\mathsf{CH/poly}$ definable coefficient sequences.

Definition 2.66 ($\mathsf{CH/poly}$ definable coefficient sequences [53])**.** *Let $\mathfrak{F}_n$ be a family of polynomials in $\mathbb{Z}[x]$ such that the degree of f_n and the bitsize of its coefficients are smaller than $2^{p(n)}$ for some polynomial p. The coefficient sequence of $\mathfrak{F}_n$ is the sequence of integers $g(n, a)$ defined by the relation*

$$f_n(x) = \sum_{a=0}^{2^{p(n)}-1} g(n, a) x^a \,. \tag{2.55}$$

The coefficient sequence $g(n, a)$ is said to be definable in $\mathsf{CH/poly}$ if the language $\mathsf{Bit}(a) = \{(1^n, a, j, b) : \text{the } j\text{-th bit of } a(n, a) \text{ is equal to } b\}$ is in $\mathsf{CH/poly}$.

Note that in contrast to algebraic number generators, that are of degree at most polynomial in the sequence argument i, the polynomials in Equation (2.55) can have degree exponential in n, which justifies the difference in the definition of the $\mathsf{CH/poly}$ language used.

We also give a different definition which relates to the evaluation of polynomials, rather than the computation of its coefficients. In particular, we proceed to define CH/poly evaluation of polynomials.

Definition 2.67 (CH/poly evaluation [53])**.** *Let $\mathfrak{F}_n$ be a family of polynomials given by Equation* (2.55)*. We say that the family of polynomials can be evaluated in CH/poly if the language $L(f) = \{(1^n, i, j, b) : 0 \leq i < 2^{p(n)}$ and the j-th bit of $f_n(i)$ is equal to $b\}$.*

We see that this definition is also a bitwise one, while since the polynomials in Equation (2.55) are in $\mathbb{Z}[x]$, the argument i is expected to be a positive integer as well, of magnitude at most $2^{p(n)} - 1$. Restricting the input in this particular way is intentional to allow for the following result:

Theorem 2.68 (Theorem 3.5 in [53])**.** *Let $\mathfrak{F}_n$ be a family of polynomials of the form given in Equation* (2.55)*. if $\mathfrak{F}_n$ can be evaluated in CH/poly at integer points, the coefficient sequence of $\mathfrak{F}_n$ is definable in CH/poly.*

We are now ready to define the family of polynomials that will play the central role in this result.

Theorem 2.69. *Let $\mathfrak{F}_i$ be an algebraic number generator. From this sequence we define a family $\mathfrak{G}_n$ of univariate polynomials as follows*

$$g_n(x) = \prod_{i=1}^{2^n} f_i(x)\,. \tag{2.56}$$

Furthermore, the coefficient sequence $b(n, a)$ of the family $\mathfrak{G}_n$, given by

$$g_n(x) = \sum_a b(n, a)x^a\,,$$

is definable in CH/poly.

As we see, the fact that the polynomials g_n vanish on all roots of the polynomials of the algebraic number generator f_i for $1 \leq i \leq 2^n$ is simply followed by taking their product. That the coefficient sequence of $\mathfrak{G}_n$ is definable in CH/poly follows from the fact that integer sequences definable in CH/poly are stable under products and summations, as shown in Theorem 10 of [17] and then by applying Theorem 2.68.

We now present a lemma that will allow us to show that the polynomials of the family $\mathfrak{G}_n$ defined in Equation (2.56) are in $\mathsf{SPS}_{s,e}$ for appropriate values of s and e.

Lemma 2.70 (Membership in $\mathsf{SPS}_{s,e}$)**.** *Let $g_n(x) = \sum_a b(n, a)x^a$ where the $a \in \mathbb{Z}$ ranges from 0 to $2^{cn} - 1$, $0 \leq a \leq 2^{cn} - 1$, $b(n, a)$ is a sequence of integers such that their absolute value satisfies $|b(n, a)| < 2^{2^{cn}}$ and are definable in CH/poly, and $c \in \mathbb{Z}^+$ is an integer constant independent of n. If $\mathrm{perm}_n \in \mathsf{VP}^0$, there is a polynomially bounded function $p(n)$ such that $2^{p(n)}g_n \in \mathsf{SPS}_{s,e}$, where $s = n^{O(\sqrt{n}\log n)}$ and $e = 2^{O(n)}$.*

Proof. The proof of this lemma is straightforward using the previously established results. Note that the definition of g_n above is similar to the definition of algebraic number generators, see Definition 2.61, with two differences: where in the definition of algebraic

number generators the coefficient sequence and exponents are defined with respect to i (or by renaming n), in the above definition they depend on 2^n. Furthermore, instead of demanding that the coefficients of g_n are bitwise computable in $\mathsf{CH/poly}$, we demand that they are $\mathsf{CH/poly}$ definable, as per Definition 2.66. By close examination of the two definitions, as we've already mentioned, the difference again is allowing degrees exponential in n.

With the above information, we only need to slightly rewrite the coefficients of g_n, in particular, write them in their binary expansion as integers:

$$b(n,a) = \sum_{i=0}^{2^{cn}-1} b_i(n,a)2^i \,. \tag{2.57}$$

We now define the following multilinear polynomial

$$h_n(x_1,\ldots,x_{cn},z_1,\ldots,z_{cn}) = \sum_{a=0}^{2^{cn}-1}\sum_{i=0}^{2^{cn}-1} z_1^{i_1}\cdots z_{cn}^{i_{cn}} x_1^{a_1}\cdots x_{cn}^{a_{cn}} \,, \tag{2.58}$$

where i_j and a_j are the binary digits of i and a respectively, that is $i = \sum_{j=1}^{\lfloor \log i \rfloor+1} i_j 2^{j-1}$ and $a = \sum_{j=1}^{\lfloor \log a \rfloor+1} a_j 2^{j-1}$. Note that since we have $0 \leq i, a \leq 2^{cn}-1$, $\lfloor \log(2^{cn}-1) \rfloor + 1 = \log 2^{cn-1} + 1 = cn$ bits are sufficient.

By now choosing the following appropriate evaluation, we may obtain g_n from Equation (2.58) as follows.

$$\begin{aligned} &h_n(x^{2^0}, x^{2^1}, \ldots, x^{2^{cn-1}}, 2^{2^0}, 2^{2^1}, \ldots, 2^{2^{cn-1}}) = \\ &\sum_{a=0}^{2^{cn}-1}\sum_{i=0}^{2^{cn}-1} a_i(n,a) 2^{i_1+2i_2+\cdots+2^{cn-1}i_{cn}} x^{a_1+2a_2+\ldots 2^{cn-1}a_{cn}} = \\ &\sum_{a=0}^{2^{cn}-1}\sum_{i=0}^{2^{cn}-1} a_i(n,a)2^i x^a = \sum_{a=0}^{2^{cn}-1} a(n,a)x^a = g_n(x) \,. \end{aligned} \tag{2.59}$$

Now by the assumption $\mathrm{perm}_n \in \mathsf{VP}^0$, by using Lemma 2.62, it follows that the coefficients of h_n, $a_i(n,a)$, which are equivalent to the i-th bit of $a(n,a)$, are computable in $\mathsf{P/poly} \subseteq \mathsf{GapP/poly}$. Therefore, using Valiant's criterion for constant-free circuits, i.e., Proposition 2.64, we conclude that the family of polynomials corresponding to h_n, $\mathfrak{H}_n \in \mathsf{VNP}^0$. In turn, by Theorem 2.63 there exists a polynomially bounded family $p(n)$ such that the family $\mathfrak{F}_n$ defined by $f_n = 2^{p(n)}h_n$ is in VP^0. Finally, by applying Proposition 2.65 to $\mathfrak{F}_n$, we conclude that the family defined by g_n, $\mathfrak{G}_n \in \mathsf{SPS}_{s,e}$ for $s = n^{O(\sqrt{n}\log n)}$ and $e = 2^{O(n)}$. □

With the above Lemma, we are now ready to show that the family of polynomials defined in Theorem 2.69 is in $\mathsf{SPS}_{s,e}$ for appropriate values of s and e. In particular, we have

Theorem 2.71. *Let $\mathfrak{F}_i$ be an algebraic number generator and $\mathfrak{G}_n$ be the family of polynomials defined by $g_n(x) = \sum_{i=1}^{2^n} f_i(x)$. If $\operatorname{perm}_n \in \mathsf{VP}^0$, there is a polynomially bounded function $p(n)$ such that $2^{p(n)} g_n \in \mathsf{SPS}_{s,e}$, where $s = n^{O(\sqrt{n}\log n)}$ and $e = 2^{O(n)}$.*

To prove the above Theorem, we need to check that the conditions of Lemma 2.70 apply in terms of bounds on the exponent and coefficient absolute value, which is straightforward. Furthermore, by applying Theorem 2.69 we also obtain that the coefficients of polynomials in the family $\mathfrak{G}_n$ are $\mathsf{CH/poly}$ definable and thus Lemma 2.70 is applicable.

We are now finally ready to prove the main theorem, which connects the existence of a hitting set for $\mathsf{SPS}_{s,e}$ with the hardness of the permanent for constant-free circuits.

Theorem 2.72 (Permanent lower bound from hitting sets, Theorem 7 in [51])**.** *Let $\mathfrak{F}_i$ be an algebraic number generator and let $\mathcal{H}_m$ be the set of all roots of the polynomials f_i for $i \leq m$. Let q and r be two functions such that $\mathcal{H}_{q(s)+r(e)}$ is a hitting set for $\mathsf{SPS}_{s,e}$. The permanent is not in VP^0 if $r(e) = e^{o(1)}$ and q satisfies the following condition: For some constant $0 < c < 1$ and s large enough, $q(s) \leq 2^{(\log s)^{1+c}} = s^{\log^c s}$.*

Proof. We begin with some commentary on the statement of the theorem. It may seem at first strange that an unrelated algebraic number generator may provide a hitting set for the $\mathsf{SPS}_{s,e}$ polynomials. Recall however that these are univariate polynomials, whose number of roots is bounded by their degree. By examining Definition 2.59 we see that each such polynomial can be of degree at most se, therefore any set of $se + 1$ points is a hitting set. Here instead we demand that the hitting set derived is of size at most subpolynomial in e and quasi-polynomial in s. Note that s mostly bounds the size of the expression (including how the coefficients may be described), while e bounds the size of the coefficients and the exponents and are thus independent, as the notation implies.

We will now prove the result by contradiction. For that, as we've already mentioned, we will use the family of polynomials defined in Theorem 2.69, that is, $g_n(x) = \prod_{i=1}^{2^n} f_i(x)$. We now assume by contradiction that both the hitting set in the statement exists and that $\operatorname{perm}_n \in \mathsf{VP}^0$.

Since $\operatorname{perm}_n \in \mathsf{VP}^0$ by Theorem 2.71 we have that $2^{p(n)} g_n \in \mathsf{SPS}_{s,e}$ for $s = n^{O(\sqrt{n}\log n)}$, $e = 2^{O(n)}$ and some polynomially bounded function $p(n)$. Note that for these values of s and e, we have

$$\begin{aligned} q(s) + r(e) &= 2^{(\log s)^{1+c}} + e^{o(1)} = 2^{\left(\log\left(n^{d\sqrt{n}\log n}\right)\right)^{1+c}} + 2^{o(n)} \\ &= 2^{d^{1+c} n^{\frac{1+c}{2}} \log^{2(1+c)} n} + 2^{o(n)} = 2^{o(n)}, \end{aligned} \tag{2.60}$$

where $d \in \mathbb{R}^+$ a constant. The last equality follows since $c < 1$ so $n^{\frac{1+c}{2}}$ is in $o(n)$, as the rest of the expression as well. Therefore by the assumption that $\mathcal{H}_{q(s)+r(e)}$ is a hitting set for $\mathsf{SPS}_{s,e}$, it follows that H_{2^n} is a hitting set for the same class. We have thus reached a contradiction, since by the above $2^{p(n)} g_n \in \mathsf{SPS}_{s,e}$ so there exists a point $t \in H_{2^n}$ such that $2^{p(n)} g_n(t) \neq 0$. However by definition, H_{2^n} consists of all roots of the polynomial f_i for $i \leq 2^n$ of the algebraic number generator $\mathfrak{F}_n$ and thus the polynomial $2^{p(n)} g(n) = 2^{p(n)} \prod_{i=1}^{2^n} f_i(x)$ vanishes on all values of H_{2^n}. □

We now interpret the above theorem. In essence, the theorem boils down to the implication that if the hitting set $\mathcal{H}$ exists (for the appropriate parametrization), then $\text{perm}_n \notin \mathsf{VP}^0$. This opens a number of possibilities. First of all, the most non-interesting case is that the hitting set does not exist. By the contrapositive of the above implication, it is sufficient that $\text{perm}_n \in \mathsf{VP}^0$ for this to occur, but not necessary. Thus, disproving the existence of such a hitting set simply eliminates this path towards a lower bound for the permanent. The most interesting case is that the hitting set for $\mathsf{SPS}_{s,e}$ does exist. Then, by the above theorem, a superpolynomial bound for constant-free circuits for the permanent exists. Thus, it suffices to answer a question about whether a specific class of polynomials has a hitting set to answer a complexity theoretic question. In fact, we can utilize the theorem to distance the question at hand from any notion of algebraic complexity further. Note that in the statement of Theorem 2.72, we are free to choose the algebraic number generator. Since the $\mathsf{SPS}_{s,e}$ polynomials are in $\mathbb{Z}[x]$, it makes sense to choose integer roots, thus we may select the algebraic number generator $f_i(x) = x - i$ for $i \geq 1$. This is essentially an improvement on the original τ-conjecture posed by Shub and Smale in [80], the difference being that instead of arbitrary algebraic circuits, we may only focus on circuits that conform to the form of the SPS polynomials, that is sum of products of sparse polynomials, i.e., depth-4 circuits. We thus arrive at the first form of a new τ-conjecture. This conjecture and all following ones can be found in [51], unless noted otherwise.

Conjecture 2.73 (τ-conjecture for $\mathsf{SPS}_{s,e}$ polynomials)**.** *If every non-zero polynomial in $\mathsf{SPS}_{s,e}$ has at most $(s+\log e)^c$ integer zeros for some constant $c \in \mathbb{N}^+$, then $\text{perm}_n \notin \mathsf{VP}^0$.*

The above conjecture is still rather restrictive. In particular, the conditions on the coefficients and the exponents in the definition of $\mathsf{SPS}_{s,e}$ polynomials are rather difficult to work with and might not be necessary to consider. We may therefore consider a stronger version of the conjecture, which targets a larger class of polynomials, which essentially shed those restrictions.

Conjecture 2.74 (Strong τ-conjecture for sums of products of sparse polynomials)**.** *Consider every non-zero polynomial of the following form*

$$f(x) = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j}(x)\,, \tag{2.61}$$

where each $f_{i,j} \in \mathbb{Z}[x]$ has sparsity at most t. If the number of integer zeros of every such f is at most $(kmt)^c$ for some constant $c \in \mathbb{N}^+$, then $\text{perm}_n \notin \mathsf{VP}^0$.

Note that in the form given in Equation (2.61), no restrictions are placed in the coefficients or the degree of the $f_{i,j}$. We can still derive a bound on the total number of monomials over all $f_{i,j}$ as in the definition of $\mathsf{SPS}_{s,e}$ polynomials, by setting $s = mt$. While we have simplified the polynomials considered at a significant level, we can still make a stronger conjecture that would imply both previous ones. Examining Conjecture 2.74 there is an obvious point of improvement. Each sparse polynomial $f_{i,j}$ is defined over $\mathbb{Z}[x]$, which strongly implies the use of number theoretic techniques would be required to resolve the conjecture. Instead, one can turn into the more convenient field of analysis, where polynomials are much better understood. To achieve this, one would require to

work in some field containing the integers. The obvious choice is the real numbers $\mathbb{R}$, since the field of rationals $\mathbb{Q}$ contains the integers and $\mathbb{R}$ is its usual completion, although not the only one. We can now take advantage of the established and understood notions from real analysis such as limits, integration and so on. We thus arrive at the first version of the real τ-conjecture.

Conjecture 2.75 (Real τ-conjecture, constant-free version)**.** *Consider every non-zero polynomial of the form given in Equation* (2.61)*, where each of the $f_{i,j} \in \mathbb{R}[x]$ has sparsity at most t. If the number of* real *roots of every such f is at most $(kmt)^c$ for some constant $c \in \mathbb{N}^+$, then $\text{perm}_n \notin \mathsf{VP}^0$.*

We can still improve the conjecture somewhat, as we will now demonstrate. Both improvements we will introduce are due to Tavenas, which he describes in his Ph.D. thesis [90], in French. For the reader not fluent in French, unfortunately these arguments have not been translated to English or any other language. Despite that, after a personal communication with the author, we were informed that the subsequent work [55] contains very similar arguments for the reader highly interested in their proof. We present here the improvements without proof.

The first improvement is to improve the lower bound obtained. Instead of proving a lower bound for constant-free circuits only, it is possible to show a lower bound for general algebraic circuits, that is, $\text{perm}_n \notin \mathsf{VP}$. This is achieved by finding a "hard" polynomial, i.e., one that can be defined as a projection of the permanent, whose coefficients are definable in P. Details of this proof can be found in pages 53 and 54, subsection "Avec la définissabilité dans P" (With definability in P) of [90]. Recall that the definition of an algebraic circuit, Definition 2.45, as well as the definitions of the classes VP and VNP by extension, Definition 2.48 and Definition 2.49 respectively, depend on the underlying field $\mathbb{K}$. The above proof works for fields of characteristic 0, in particular for $\mathbb{K} = \mathbb{C}$.

The second improvement can be achieved by a more careful examination of the proof. This allows us to demand a dependence exponential in the parameter m, which limits the size of each product in Equation (2.61), rather than polynomial. Details of this can be found in Conjecture 3.23 of [90].

Combining these two improvements, we can now present the main version of the real τ-conjecture which we wish to eventually resolve.

Conjecture 2.76 (Real τ-conjecture)**.** *Consider every non-zero polynomial of the form given in Equation* (2.61)*, that is*

$$f(x) = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j}(x)\,,$$

where each of the $f_{i,j} \in \mathbb{R}[x]$ has sparsity at most t. If the number of real *roots of every such f is at most $(k2^m t)^c$ for some constant $c \in \mathbb{N}^+$, then $\text{perm}_n \notin \mathsf{VP}$.*

We note that further forms of the conjecture have been developed, several of which can be found in Chapter 3 of [90], including a version for bivariate polynomials that has been published in [55]. We focus on Conjecture 2.76, which is also the central motivation of this work. Given its implication as well as current techniques, the conjecture does not seem trivial to solve and will require substantial effort. Nevertheless, it provides us with

a method to resolve Valiant's conjecture, see Conjecture 2.57, or at very least gauge our confidence in either outcome of resolving Valiant's conjecture. A first attempt to judge how difficult to resolve the conjecture is applying Descartes' rule of signs, in particular the version which uses the sparsity of the polynomial to obtain an upper bound on the number of real roots, Corollary 2.23. A simple examination of Equation (2.61) reveals that the polynomial f can have sparsity at most kt^m, thus we immediately obtain an upper bound of the same magnitude. Unfortunately, this bound is not polynomial in t. Furthermore, as we've already detailed in Section 2.1.2, it is easy to construct polynomials that match this upper bound, for example the Pochhammer-Wilkinson polynomials given by

$$f_n(x) = \prod_{i=1}^{n}(x - i)\,,$$

for $n = kt^m$, or in fact, any other polynomial of this form that is simply the product of linear polynomials vanishing at a single root, for the required sparsity. However, when brought in the form given in Equation (2.61), it is not clear if these polynomials would disprove the conjecture. As can be seen in this example, the Pochhammer-Wilkinson polynomials indeed obey the conjecture since their number of roots is linear in the size of the product.

While further attempts have provided less trivial lower bounds, some of them detailed in Chapters 4 and 5 of [90], which elaborate on the results of [55] and [54], those are still far from the goal and with no obvious way to proceed forward. A significant difficulty in proving the conjecture is that it is sufficient that a single polynomial of the form in Equation (2.61) has more roots than the bound posed to refute the conjecture. Therefore, as a first step, we may look at randomized version of the conjecture, where all polynomials of that form have coefficients following some distribution. We may then examine the *expected* number of zeros with respect to this distribution. This may lead to two outcomes: First, if the expected number surpasses the bound in the conjecture, via the probabilistic method a fixed polynomial exists with this many real roots and thus the conjecture is refuted. On the other hand, if the expected number obeys the bound, we may not resolve the conjecture one way or the other, since the possibility of a polynomial refuting the conjecture exists, yet we gain new understanding and methods to approach the problem, as well as allowing ourselves to adjust our confidence in the conjecture based on the result.

Furthermore, even the behavior of a single sparse random polynomial has not been well-studied, in contrast to the results for the dense case we presented in Section 2.2.3. As a first step, in this work we present results for exactly this case, with the focus being sparse polynomials with coefficients being i.i.d. random variables following the standard normal distribution. Before proceeding with our main results in Chapter 3, we present the few previous results for sparse polynomials, both with randomness and without in the next section.

2.3.4 Previous results for the number of roots of sparse polynomials

In Section 2.2.3, we detailed several results for the expected number of real roots for random polynomials with respect to several well-known distributions. In that case, there

was no real restriction on the polynomials studied and we parametrized them, as well as the results, in terms of the degree d, taking into account also the well-known bound in Lemma 2.4 that we've already examined in detail. In this section, we review similar results on the number of real roots of sparse polynomials, both in the fixed polynomial setting as well as for random polynomials.

The starting point is the fixed case for univariate polynomials. As we've already elaborated in great detail, Descartes' rule of sign given in Theorem 2.17 implies an upper bound of $2k-1$ for distinct real roots, as can be seen by Corollary 2.23. Furthermore, as we've already mentioned, there are polynomials that match this bound, see [32] for details. Therefore, no further improvement is possible in the fixed case.

Continuing with the fixed case, we turn our attention to multivariate polynomials instead. In this case, the existing results are not known to be optimal and it is suspected that the upper bound obtained can be significantly improved. We have already defined the notion of a polynomial system in Definition 2.13. In the same section, we've already mentioned that in order to count the number of roots of a polynomial system, we should demand that the number of equations matches the number of variables, n. Otherwise, either the system is undetermined and in general has an infinite number of roots in the general case, or it is overdetermined and in the general case has no roots at all. Suppose thus that we have the following system of n polynomial equations in n variables:

$$F(X) = \begin{cases} f_1(X) = \sum\limits_{i=0}^{k_1-1} c_{1,i} x^{e_{1,i}} = 0\,, \\ \quad\vdots \\ f_j(X) = \sum\limits_{i=0}^{k_j-1} c_{j,i} x^{e_{j,i}} = 0\,, \\ \quad\vdots \\ f_n(X) = \sum\limits_{i=0}^{k_n-1} c_{n,i} x^{e_{n,i}} = 0\,, \end{cases} \tag{2.62}$$

where $X = (x_1, \ldots, x_n)$ and the polynomial f_j has degree denoted by d_j, i.e., $\deg(f_j) = d_j = e_{j,k_j-1}$ for $1 \leq j \leq n$. We would like to have an analogue of Descartes' rule of signs for multivariate polynomials. Recall that Bézout's theorem, see Chapter 7 in [35], provides an upper bound in terms of the degrees $d_1, \ldots, d_n$ of the polynomials in the polynomial system. The theorem states that over an algebraically closed field, such as $\mathbb{C}$, the number of roots of a polynomial system such as the one in Equation (2.62) is either infinite or bounded by the product of the degrees of the polynomials involved in the polynomial system, that is $d_1 \cdots d_n$. It is an immediate corollary of the theorem that the number of real roots of the polynomial system F is either infinite or bounded by $d_1 \cdots d_n$. This can be viewed as a multivariate analogue of the Fundamental Theorem of Algebra, stated in Theorem 2.8.

Khovanskii in his seminal work "Fewnomials" [49] provides precisely an analogue of Descartes' rule of signs for multivariate polynomials that complements Bézout's theorem. It provides an upper bound on the number of real roots of a polynomial system in terms of its sparsity. We state Khovanskii's result without proof, but first we must define a few relevant notions.

To begin with, we must define what the sparsity of a polynomial system is. There is a number of different ways to define such a notion. The most obvious one and the one that is of interest to us, is the total number of non-zero terms, that is, the sum of the sparsities of the constituent polynomials of the system. We call this the *total sparsity*, denote it by k and following the notation of Equation (2.62), we define it as $k \coloneqq \sum_{i=1}^{n} k_i$.

We also need to define the *positive orthant* of a space, in particular $\mathbb{R}^n$. This is simply the subspace defined by allowing the components of the vectors in $\mathbb{R}^n$ to only obtain positive values. That is, if we use the same notation as the variables of the polynomial system $(x_1, \ldots, x_n)$, we obtain the positive orthant by demanding $x_1, \ldots, x_n > 0$. Khovanskii's theorem concerns the number of solutions in the positive orthant, analogous to how Descartes' rule concerns the number of positive solutions. We may now state the theorem.

Theorem 2.77 (Khovanskii's theorem [49],[48])**.** *Let $F(X)$ be a polynomial system of n equations with n variables, as in Equation* (2.62)*, with total sparsity k. Then the number of non-degenerate real roots in the positive orthant of $\mathbb{R}^n$ is bounded by*

$$2^{\binom{k}{2}} (n+1)^k \ .$$

A solution $X_0 = (x_{0,1}, \ldots, x_{0,n})$ is considered degenerate if the Jacobian determinant of the system vanishes at X_0. Recall that the Jacobian of the polynomial system F is defined as

$$J(F) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} .$$

Similar to Descartes' rule of signs, we may apply Theorem 2.77 to every other orthant as well. By the definition of the positive orthant, it is trivial to see that the total number of orthants for $\mathbb{R}^n$ is 2^n, e.g., the real plane $\mathbb{R}^2$ has 4 orthants, which are more commonly known also as quadrants in this case. We thus obtain a bound on the total number of real roots as follows.

Theorem 2.78 (Khovanskii's theorem extended)**.** *Let $F(X)$ be a polynomial system of n equations with n variables, as in Equation* (2.62)*, with total sparsity k. Then the number of non-degenerate real roots in $\mathbb{R}^n$ is bounded by*

$$2^n 2^{\binom{k}{2}} (n+1)^k \ .$$

Note that unlike the Descartes' rule of signs, Khovanskii's result only involves the (total) sparsity of the polynomial system, not the number of signs changes that occur in it. Whether it is possible to state a result that depends on some notion similar to the number of signs changes for polynomial systems is an open problem. Also note that the non-degenerate condition is necessary, in particular it avoids the issue of polynomials with an infinite number of real roots.

We also remark that the number of roots is exponential in the total sparsity k. It is widely believed that the bound is far from optimal and in fact, it should be possible to show that the number of roots is only polynomial in k. Unfortunately, the results already

obtained are far from achieving such a bound. In the general case, the only significant improvement was obtained by Bihan and Sottile [7], as follows

Theorem 2.79 (Bihan-Sottile bound [7])**.** *Let $F(X)$ be a polynomial system of n equations with n variables, as in Equation* (2.62)*, with total sparsity k, such that $k > n$. In particular, let $l \in \mathbb{Z}^+$ be defined so that $k = n+l+1$. Then the number of non-degenerate real roots in the positive orthant of $\mathbb{R}^n$ is bounded by*

$$\frac{e^2+3}{4} 2^{\binom{k-n-1}{2}} n^{k-n-1} = \frac{e^2+3}{4} 2^{\binom{l}{2}} n^l \,.$$

For the interested reader, the book by Sottile [83] provides the necessary details as well as further results on the topic.

Other extensions of Khovanskii's result exist, for example the result has been extended to other fields such as number fields or p-adic fields [75]. Other results study special cases, for example the number of real roots for a polynomial system comprising of a bivariate polynomial and a real line [2]. In particular, it is clear that this line of research has several open questions and improvements to be potentially obtained.

Given that improving the bound on the fixed case seems out of reach for current techniques, it is instructive to also examine the random case. Definition 2.34 naturally generalizes to random polynomials, the only difference being the monomials that the coefficients correspond to. We may thus consider random polynomial systems on n variables with n random polynomial equations and study their expected number of real roots. Until recently, very few results concerned themselves with this question, for example [64]. However, the real τ-conjecture we presented in Section 2.3.3 has sparked new interest in the topic. In particular, we focus on two recent related results.

The first result considers a special type of polynomial systems, called *unmixed systems*. These are polynomial systems, typically of sparse polynomials, such that all polynomials belonging in the system are defined over the same support vector, say $S = (x_1^{e_{1,1}} \cdots x_n^{e_{1,n}}, \ldots, x_i^{e_{i,1}} \cdots x_n^{e_{i,n}}, \ldots, x_n^{e_{k,1}} \cdots x_n^{e_{k,n}})$, where k is the shared sparsity of the polynomials in the system. In contrast, we call systems that do not share the same support vector *mixed*. For unmixed systems, it is possible to state improved bounds in terms of the sparsity k of each polynomial, rather than the total sparsity which would simply be nk. In particular, we have the following result due to [20].

Theorem 2.80 (Bound for random unmixed polynomial systems)**.** *Let $F(X)$ be an unmixed polynomial system of n random polynomials on n variables defined on the support vector $S = (s_1, \ldots, s_k)$. Also let the coefficients of each polynomial in the system be i.i.d. centered Gaussian random variables, with the coefficient corresponding to the monomial s_i having variance σ_i^2 and let $\sigma^2 = (\sigma_1^2, \ldots, \sigma_k^2)$. That is, we have $c_{j,i} \sim \mathcal{N}(0, \sigma_i)$ for $1 \leq j \leq n$ and $c_{j,i}$ is the coefficient of the monomial s_i in the polynomial f_j. Then the expected number of non-degenerate real zeros of $F(X)$ in the positive orthant of $\mathbb{R}^n$, denoted by $Z_{\mathbb{R}^n_+}\left(F, \mathcal{N}(0, \sigma^2)\right)$, is bounded by*

$$Z_{\mathbb{R}^n_+}\left(F, \mathcal{N}(0, \sigma^2)\right) \leq \frac{1}{2^{n-1}} \binom{k}{n} \,.$$

Note in particular that the bound does not depend exponentially on n, although it still does on k. Furthermore, while for $n = 1$ the bound is dominated by the Descartes' rule of signs, in the same work [20] the following bound for the univariate case is given.

Theorem 2.81 (Univariate random sparse case (previous result))**.** *Let $f(x)$ be a univariate random k-sparse polynomial following the standard normal distribution $\mathcal{N}(0,1)$. Then its expected number of real zeros of $f(x)$ is bounded by*

$$Z_{\mathbb{R}}\left(f, \mathcal{N}(0,1)\right) \leq \frac{2}{\pi}\sqrt{k}\log k\,.$$

Our work improves on that bound by removing the logarithmic factor and proving the new bound optimal by showing a family of polynomials exists whose expected number of real roots matches it asymptotically. Furthermore, we do so by a technique detailed in Section 3.2 which can be generalized for other distributions. A very recent result [18] extends Theorem 2.80 to the more general mixed case as well. We state only the main statement of this work and refer the interested reader to it for details and clarifications.

Theorem 2.82. *Let $F(X)$ be a polynomial system of n random polynomials on n variables, as in Equation* (2.62)*. Let the coefficients of each polynomial in the system be i.i.d. centered Gaussian random variables, with the coefficient corresponding to the monomial s_i having variance σ_i^2 and let $\sigma^2 = (\sigma_1^2, \ldots, \sigma_k^2)$. That is, we have $c_{j,i} \sim \mathcal{N}(0, \sigma_i)$ for $1 \leq j \leq n$ and $c_{j,i}$ is the coefficient of the monomial s_i in the polynomial f_j. Let the polynomial f_i have sparsity k_i and also let the exponents of its monomials, viewed as points in $\mathbb{Z}^n$, form a set $A_i \subset \mathbb{Z}^n$ with cardinality k_i. Also let P_i be the convex hull of the points in A_i and $P \coloneqq P_1 + \cdots + P_n$ be the Minkowski sum of the convex hulls P_i. Then the expected number of non-degenerate real zeros in the positive orthant of R^n, denoted by $Z_{\mathbb{R}^n_+}(F, A_1, \ldots, A_n)$, is bounded by*

$$Z_{\mathbb{R}^n_+}(F, A_1, \ldots, A_n) \leq (2\pi)^{-\frac{n}{2}}\, V_0(k_1 - 1)\cdots(k_n - 1)\,,$$

where V_0 denotes the number of vertices of the Minkowski sum P.

In this chapter, we defined the basic notions and notation necessary to present our main results, as well as provided a detailed view of results starting from Descartes' rule of signs in 1637 and stating several key results up to the present day. Using this groundwork, we have provided the interested reader not only with the necessary background to assess our result and comprehend it in its entirety, but also provided a springboard for a deep dive in this wonderful line of research. In the following chapter, we will state the main result of this work, as well as the necessary notions particular to it.

CHAPTER 3

Expected number of zeros of random sparse polynomials

In this chapter we present our main findings relating to the expected number of zeros of random polynomials of sparsity k. Initially in Section 3.1, we examine the work of Edelman and Kostlan [28] in great detail, specifically how their approach and reformulation of Kac's integral [45] provides us with a foundation for our investigation. Then in Section 3.2 we utilize this framework to obtain our main result, a $\Theta(\sqrt{k})$ bound for the expected number of zeros of k-sparse polynomials following the standard normal distribution. We also show, similar to the dense case, that the zeros of such a random sparse polynomial tend to concentrate in the neighborhood of $x = 1$ and $x = -1$.

Additionally, once again inspired by the work of Edelman and Kostlan, in Section 3.3 we present how their methodology for arbitrary distributions can be adapted to investigate sparse polynomials. Furthermore, we present some partial results towards generalizing the above results to other continuous distributions, similar in spirit to the classical results generalizing Kac's work that we examined in Section 2.2.3. In particular, we demonstrate technical contributions that generalize the techniques used to obtain our main results and can be used in the future to obtain similar results for other distributions. In that regard, we attempt to present such results in the greatest generalization possible and establish standard techniques that have been long used in the dense case.

3.1 The Edelman-Kostlan integral

In their work [28], Edelman and Kostlan derive the Kac integral given in Equation (2.29) through a simple geometric argument that greatly simplifies both the statement and its interpretation. In particular, certain properties of the Gaussian distribution that we will shortly examine simplify the argument even more, although it is possible to state the results for any continuous distribution, as we will see in Section 3.3. In the following we deviate slightly from the presentation in the original work and choose to use a generic support vector from the start, rather than restricting ourselves to the dense case, then generalizing.

In Section 2.1 we saw that we may describe any polynomial by two vectors, the coefficient vector $C = (c_0, \ldots, c_{k-1})$ and the exponent vector $E = (e_0, \ldots, e_{k-1})$. We also noted in Section 2.1.2 that especially for univariate polynomials, where we do not have to fix an order for the variables, we may simply use the support vector $S = (x^{e_0}, \ldots, x^{e_{k-1}})$, which is easily derived from the exponent vector. In fact, we use the term vector since we consider x to be fixed to some arbitrary value that we do not specify, while in reality as $x \in \mathbb{R}$ varies, we rather obtain a curve in $\mathbb{R}^k$, which we simply call the *support curve* and denote by $S(x)$. Often when it is clear from context, we refer to both the support

vector and the derived curve as simply the "support vector" for brevity.

Given the above, we may describe the value of a k-sparse polynomial at some x simply as the *inner product* of two vectors in $\mathbb{R}^k$, namely

$$f(x) = C \cdot S = \sum_{i=0}^{k-1} c_0 x^{e_i} \,. \tag{3.1}$$

Now x_0 being a zero of the polynomial f, that is $f(x_0) = 0$, is equivalent to the $C \cdot S(x_0) = 0$. It is a well-known fact that the inner product of two vectors being 0 implies that the two vectors, here C and $S(x_0)$, are perpendicular. Thus for the polynomial f, the distinct points on the real line that are roots have a one-to-one correspondence with the points x that satisfy the relation

$$f(x) = C \cdot S(x) = 0 \,.$$

Therefore, for a given polynomial, we may count the number of real zeros by simply counting all points where the support vector S is perpendicular to the (fixed) coefficient vector. In other words, it is

$$N_{\mathbb{R}}(f) = |\{\, x \in \mathbb{R} \mid C \cdot S(x) = 0\}| \,. \tag{3.2}$$

Clearly, $\mathbb{R}$ can be replaced by any other set in the above to obtain the equivalent to Definition 2.11.

Similar to our approach in Chapter 2, we may now generalize the above for families of polynomials. More specifically, given a random polynomial f following the distribution $\mathcal{D}$, we may examine the expected number of zeros for the family of polynomials $\mathfrak{F}_{\mathcal{D}}(x)$. Assuming that the exponent vector, and thus consequently also the support vector apart from the value of x, is fixed, we see that the sample space $\mathfrak{F}_{\mathcal{D}}(x)$ can be viewed as a set of vectors of coefficients A that is a subset of $\mathbb{R}^k$. In particular, we may make the assumption that the entirety of $\mathbb{R}^k$ is the sample space, with $\mathbb{R}^k \setminus \mathfrak{F}_{\mathcal{D}}(x)$ having probability 0 with respect to $\mathcal{D}$.

As above, assume that the exponents are fixed. Then, the expected number of zeros could be now computed as in Definition 2.35, which often is interpreted as performing the following for each coefficient vector A: Fix the coefficient vector A and then count the zeros for each $x \in \mathbb{R}$ as the support vector S follows the support curve. We may instead obtain the same result alternatively as follows. For each x, fix the associated support vector S and count the number of coefficient vectors A, that is polynomials, that have x as a root. Clearly the two ways of counting the expected number of roots of a random polynomial f following the support curve $S(x)$ are equivalent.

The important observation is that the second method has a significant advantage geometrically. Recall that due to Equation (3.2) the two vectors will be perpendicular if and only if the polynomial has a root at that point x. Given any fixed vector, say in $\mathbb{R}^k$ for $k \geq 2$, the vectors perpendicular to it form a space that is isomorphic to $\mathbb{R}^{k-1}$. For example, given any vector in $\mathbb{R}^2$, the vectors perpendicular to it form a line and similarly given any vector in $\mathbb{R}^3$, the vectors perpendicular to it form a plane. On the other hand, when fixing a coefficient vector $A \in \mathbb{R}^k$ we obtain a space of perpendicular vectors to it in $\mathbb{R}^{k-1}$. However, not all of these vectors will be on the support curve

$S(x)$. Therefore, we must intersect the two to obtain the points corresponding to real roots of the polynomial with coefficients given by A. In contrast, if we fix the support vector $S \in \mathbb{R}^k$, all the points of the space isomorphic to $\mathbb{R}^{k-1}$ correspond to some valid polynomial.

In close observation, this is similar to the Kac integral, given in Equation (2.29). The outer integral "enumerates" through all points in the real line corresponding to roots, while the inner one counts the expected number of zeros of the polynomials that have x as a root, with respect to the measure associated with the distribution $\mathcal{D}$. We will examine this relation closer in Section 3.3 and derive a general version of the Kac formula, under certain assumptions, however for now we focus our interest in the standard normal case, as was the case in Kac's work.

A particular property of the standard normal distribution $\mathcal{N}(0,1)$ that we rely upon is its rotational invariance. Often, this is used to show that linear combinations of random variables following normal distributions also follow some normal distribution themselves. We are more interested in the following more fundamental property.

Lemma 3.1. *Let $A = (a_0, \ldots, a_{k-1})$ be a vector in $\mathbb{R}^k$ whose components a_i are i.i.d. random variables following the standard normal distribution, $a_i \sim \mathcal{N}(0,1)$. Also let $\|A\| = \sqrt{a_0^2 + \cdots + a_{k-1}^2}$ be the length of the vector A, so that $\hat{A} = A/\|A\|$ stands for the normalization of the vector A. Then $\hat{A}$ is uniformly distributed on the unit sphere $\mathcal{S}^{k-1}$.*

Proof. Recall that for each a_i the density function is

$$g(a_i) = (2\pi)^{-\frac{1}{2}} e^{-\frac{a_i^2}{2}} .$$

Since the a_i are independent, the joint distribution for A can then be easily computed as the product of the marginal ones, that is

$$g(A) = \prod_{i=0}^{k-1} g(a_i) = \prod_{i=0}^{k-1} (2\pi)^{-\frac{1}{2}} e^{-\frac{a_i^2}{2}} = (2\pi)^{-\frac{k}{2}} e^{-\frac{1}{2}\sum_{i=0}^{k-1} a_i^2} = (2\pi)^{-\frac{k}{2}} e^{-\frac{1}{2}\|A\|^2} .$$

From the above it is clear that the density for a vector A only depends on its length $\|A\|$, therefore all vectors with the same length also have the same density, i.e., all vectors on a sphere of radius $r \coloneqq \|A\|$ have the same density. It remains to transform the underlying sample space from $\mathbb{R}^k$ to $\mathcal{S}^{k-1}$. This is easily accomplished via a transformation to hyperspherical coordinates. For k dimensions, these consist of one coordinate that represents the length r of the vector associated with a point in $\mathbb{R}^k$ and $k-1$ angular coordinates, say $\phi_1, \ldots, \phi_{k-1}$ that specify a single point in the k-sphere of radius r. The relation between Cartesian and hyperspherical coordinates in $\mathbb{R}^k$ is then given by

$$
\begin{aligned}
a_0 &= r \prod_{i=1}^{k-1} \sin\phi_i \,, \\
a_1 &= r \cos\phi_1 \prod_{i=2}^{k-1} \sin\phi_i \,, \\
a_2 &= r \cos\phi_2 \prod_{i=3}^{k-1} \sin\phi_i \,, \\
&\vdots \\
a_{k-2} &= r \cos\phi_{k-2} \sin\phi_{k-1} \,, \\
a_{k-1} &= r \cos\phi_{k-1} \,.
\end{aligned}
$$

From the above, one can derive the Jacobian that is necessary for the transformation, see [67], which is equal to

$$
\mathfrak{J}_k = (-1)^{k-1}\, r^{k-1} \prod_{i=2}^{k-1} \sin^{i-1}\phi_i \,.
$$

Let $d\phi = d\phi_1 \cdots d\phi_{k-1}$ and let Φ be the domain of $(\phi_1, \ldots, \phi_{k-1})$, specifically we have $\phi_1 \in [0, 2\pi)$ and $\phi_i \in [0, \pi]$ for $i = 2, \ldots, k-1$. We then have

$$
\begin{aligned}
\int_{\mathbb{R}^k} g(A)\, da_0 \cdots da_{k-1} &= \int_{\mathbb{R}^k} (2\pi)^{-\frac{k}{2}}\, e^{-\frac{1}{2}\|A\|^2}\, da_0 \cdots da_{k-1} \\
&= \int_0^\infty \int_\Phi (2\pi)^{-\frac{k}{2}}\, e^{-\frac{1}{2}r^2} |\mathfrak{J}_k|\, d\phi\, dr \\
&= \int_0^\infty (2\pi)^{-\frac{k}{2}}\, r^{k-1} e^{-\frac{1}{2}r^2} \left(\int_\Phi \prod_{i=2}^{k-1} \sin^{i-1}\phi_i \, d\phi \right) dr \,. \qquad (3.3)
\end{aligned}
$$

Now note that the inner integral, which runs over Φ with measure $d\phi$, is precisely the surface area of $\mathcal{S}^{k-1}$, in fact if we were to include r^{k-1} in the inner integral we would obtain the surface area of the $(k-1)$-sphere of radius r. Let $\mathcal{A}_{k-1}$ denote the surface area of S^{k-1}. We then have that Equation (3.3) is equal to

$$
\int_0^\infty (2\pi)^{-\frac{k}{2}}\, \mathcal{A}_{k-1} r^{k-1} e^{-\frac{1}{2}r^2}\, dr \,.
$$

Since $(2\pi)^{-\frac{k}{2}}\, \mathcal{A}_{k-1}$ is fixed for a given k, it follows that density of A only depends on the radius r. Therefore, since each point on the unit sphere corresponds to a collection of points with radius $r \in [0, \infty)$, it follows that the normalized vector $\hat{A}$ has the same density for any point on the unit sphere and is therefore uniformly distributed. □

Also note from the above that if we were to choose some set $I \subseteq \mathbb{R}^k$ as the domain of integration, we could in theory find the corresponding set $I_c \subseteq \mathcal{S}^{k-1}$ and potential ranges of r, which may also differ for each point. In practice however, for most sets it would be

a daunting proposition. Thankfully in our case the sets we are interested in are easy to describe. Recall that given a support curve $S(x)$ in R^k, we wish to find for each point on the curve S corresponding to some value of x, the vectors in $\mathbb{R}^k$ that are perpendicular to it, since they correspond to polynomials that have x as a root. To take advantage of the above, we transform both the coefficient and the support vector to vectors on the unit sphere S^{k-1}. We have

$$\hat{A} = \frac{A}{\|A\|} = \left(\frac{a_0}{\|A\|}, \ldots, \frac{a_{k-1}}{\|A\|} \right),$$
$$\hat{S} = \frac{S}{\|S\|} = \left(\frac{x_0}{\|S\|}, \ldots, \frac{x_{k-1}}{\|S\|} \right),$$

where $\|A\| = \sqrt{\sum_{i=0}^{k-1} a_i^2}$ and $\|S\| = \sqrt{\sum_{i=0}^{k-1} x_i^2}$. We also employ the notation $\gamma(x)$ for the normalized support vector $\hat{S}$ and support curve $\hat{S}(x)$, as this notation was also employed in many works, including [28].

Similar to $\mathbb{R}^k$, for a normalized support vector $\hat{S}$ in $\mathcal{S}^{k-1}$, the points perpendicular to it are easy to describe. In fact, in correspondence with the Euclidean space, these points form a space isomorphic to $\mathcal{S}^{k-2}$. Equivalently, this space can be obtained by intersecting the unit sphere $\mathcal{S}^{k-1}$ with the space of vectors perpendicular to $\hat{S}$ in $\mathbb{R}^k$. Following Edelman and Kostlan, we call these sets equators and denote them by $S_\perp$. The easiest to visualize example that also sheds some light to the origin of the terminology used is considering a sphere $\mathcal{S}^2$ and a point S on it. The points perpendicular to it on the surface of $\mathcal{S}^2$ then form a great circle $\mathcal{S}^1$. In particular, assuming that the planet Earth is a perfect sphere rather than an oblate spheroid and taking S as the vector corresponding to either of the poles, we then obtain what we also colloquially call the Equator. Higher dimensions, although harder to visualize, behave in a similar manner.

Now, consider the vector $\hat{S}$ as it follows the curve $\hat{S}(x)$. As S varies, so does its equator $\hat{S}_\perp$ and it is easy to see that it does so in a continuous manner. As it does, it prescribes an area on $\mathcal{S}^{k-1}$, which may also overlap itself. Returning to the familiar example, we may imagine the projection of the Sun's center point of mass to the Earth's surface and consider this as our support vector. As the sun moves along its orbit through space, it prescribes a curve on the Earth's surface, which roughly corresponds with the Earth's equator. For any given point on that curve, the points perpendicular to it are simply the terminator, i.e., the equator that divides the Earth into the part lit by sunlight and that in which it is currently nighttime. Clearly, if we follow this curve for a distance corresponding to the length of a day, each point on Earth will be swept twice by the equator: once corresponding to sunrise and one to sunset. If we continue following the curve, each point will be swept even more times.

We may now apply our intuition from the above to the problem of counting zeros. In this section, we do so only for the standard normal distribution and leave the generalization to more distributions for Section 3.3. The main advantage of insisting the random polynomial follows $\mathcal{N}(0,1)$ is that we may utilize Lemma 3.1, by which every point on $\mathcal{S}^{k-1}$ has the same density, thus we do not need to keep track of which points the equators "sweep" through, only their count with respect to the measure. If $\gamma(x)$ is

the support curve, let $\gamma(x)_\perp$ be the space obtained by the union of equators for every support vector $\hat{S}$ on the support curve $\hat{S}(x)$. Measuring the area of this space will be equivalent to computing the expected number of zeros. For the above simple example, if we select any great circle of $\mathcal{S}^2$ as our support curve, we will sweep each point twice and thus the expected number of zeros would be 2. Generalizing from that intuition, we may find this area only by the length of the support curve. Let $|\gamma(x)|$ be the arclength of the support curve and also let $|\gamma(x)_\perp|$ be the area of the area swept by the equators of $\gamma(x)$, counted with multiplicity.

Do note that this multiplicity only refers to points of the sphere that happen to be swept separate times by an equator and not to the multiplicity of the zeros themselves. In fact, the multiplicity of a single point in $\gamma(x)_\perp$ corresponds to the number of distinct real zeros for that polynomial. For example, for the support vector $\left(1/\sqrt{x^2+1}, x^2/\sqrt{x^2+1}\right)$, the coefficient vector $\left(-1/\sqrt{2}.1/\sqrt{2},\right)$ will be swept twice since it corresponds to the polynomial $(-1,1)\cdot(1,x^2) = x^2-1$, which has two roots at $x=1$ and $x=-1$. It is also easy to verify that indeed at those points the inner product of the two vectors (normalized or not) is indeed zero.

With that in mind, as also given in [28, Lemma 2.1], we have

$$Z_{\mathbb{R}}\left(f, \mathcal{N}(0,1)\right) = \frac{|\gamma(x)_\perp|}{\mathcal{A}_{k-1}} = \frac{1}{\pi}|\gamma(x)|\,. \tag{3.4}$$

Therefore, to determine the expected number of real zeros of a random polynomial f following $\mathcal{N}(0,1)$, it suffices to compute the arclength of the support curve $|\gamma(x)|$. In [28, Theorem 2.1], the full proof for *dense* random polynomials following $\mathcal{N}(0,1)$ is given, i.e., when the support curve is $S = (1, x, x^2, \ldots, x^d)$, which is essentially the moment curve. We do not substitute for the moment curve here and only derive the general expression, referring the interested reader to the above work. We repeat here only the information that will also be applied on our own work presented in Section 3.2.

The arclength of a curve can be computed by measuring the length of the tangent at every infinitesimal point, which in turn is given by the derivative at that point. We thus have

$$|\gamma(x)| = \int_{\mathbb{R}} \|\gamma'(x)\|\, dx\,. \tag{3.5}$$

We now require to compute $\|\gamma'(x)\|$, with the first step consisting of calculating the vector $\gamma'(x)$. By the following computation we have

$$\begin{aligned}
\gamma'(x) &= \left(\frac{S(x)}{\|S(x)\|}\right)' = \left(\frac{S(x)}{(S(x)\cdot S(x))^{1/2}}\right)' \\
&= \frac{S'(x)\,(S(x)\cdot S(x))^{1/2} - S(x)\left([S(x)\cdot S(x)]^{1/2}\right)'}{S(x)\cdot S(x)} \\
&= \frac{S'(x)\,(S(x)\cdot S(x))^{1/2} - \frac{1}{2}S(x)\,(S(x)\cdot S(x))^{-1/2}\,(S(x)\cdot S(x))'}{S(x)\cdot S(x)} \\
&= \frac{S'(x)\,(S(x)\cdot S(x))^{1/2} - S(x)\,(S(x)\cdot S(x))^{-1/2}\,(S'(x)\cdot S(x))}{S(x)\cdot S(x)} \\
&= \frac{S'(x)\,(S(x)\cdot S(x)) - S(x)\,(S'(x)\cdot S(x))}{(S(x)\cdot S(x))^{3/2}}\,. \qquad (3.6)
\end{aligned}$$

We may then compute the length of the vector $\gamma'(x)$ by

$$\begin{aligned}
\|\gamma'(x)\| &= \left(\gamma'(x)\cdot\gamma'(x)\right)^{1/2} \\
&\overset{(3.6)}{=} \left[\frac{(S(x)\cdot S(x))^2\,(S'(x)\cdot S'(x)) - (S(x)\cdot S(x))\,(S'(x)\cdot S(x))^2}{(S(x)\cdot S(x))^3}\right]^{1/2} \\
&= \left[\frac{(S(x)\cdot S(x))\,(S'(x)\cdot S'(x)) - (S'(x)\cdot S(x))^2}{(S(x)\cdot S(x))^2}\right]^{1/2}. \qquad (3.7)
\end{aligned}$$

We may now substitute the final expression in Equation (3.7) for $\|\gamma'(x)\|$ into Equation (3.5) to obtain the final expression. Also combined with Equation (3.4), we obtain

$$Z_{\mathbb{R}}\left(f,\mathcal{N}(0,1)\right) = \frac{1}{\pi}\int_{\mathbb{R}}\left[\frac{(S(x)\cdot S(x))\,(S'(x)\cdot S'(x)) - (S'(x)\cdot S(x))^2}{(S(x)\cdot S(x))^2}\right]^{1/2} dx\,. \qquad (3.8)$$

Now, by substituting the dense moment curve $S(x) = (1, x, \ldots, x^d)$ and some clever algebraic calculations which are detailed in [28], Edelman and Kostlan obtain their close approximation of the expected number of real zeros in the dense standard normal case which we already gave in Equation (2.31). Although the expression might seem taunting, in the dense case once one has obtained Equation (3.8), it is more a question of carefully carrying out the necessary algebraic operations and the integral is rather easy to solve. However, as we've already remarked in Section 2.2.3, this is not always the case for every curve. In particular, adapting the above to the sparse case was not straightforward and required careful handling of the integral. We expand on the necessary techniques and present our main result in the next section.

3.2 Expected number of zeros of random Gaussian sparse polynomials

In this section we present our main results for k-sparse random polynomials following the standard normal distribution $\mathcal{N}(0,1)$, which were originally published in [43]. Our results are similar in spirit to the work of Kac for the dense case and we are able to show asymptotically matching upper and lower bounds for the expected number of zeros of such random polynomials. In particular, we show a random k-sparse polynomial f following the standard normal distribution $\mathcal{N}(0,1)$, has $\Theta(\sqrt{k})$ expected number of real zeros. The theorems stated below also contain precise constants for both upper and lower bounds. Furthermore, we prove that similar to the dense case, most of the zeros of such a random polynomial are concentrated in a small neighborhood around $x = 1$ and $x = -1$.

Our results are made possible by careful consideration of the Edelman-Kostlan integral given in Equation (3.8), where we use an arbitrary support vector of k components, thus giving rise to k-sparse polynomials. In particular, we substitute in the above expression for $S = (x^{e_0}, \ldots, x^{e_{k-1}})$, with $e_i \in \mathbb{N}$ and $e_0 < e_1 < \cdots < e_{k-1} = \deg(f)$. The only restriction we impose is that $e_0 = 0$, which does not result in a loss of generalization since the standard normal distribution is absolutely continuous, therefore we have that $Pr(f(0) = 0) = Pr(a_0 = 0) = 0$, where a_0 is the constant term of the random polynomial f, thus the expected number of real zeros will not be affected.

Throughout this section f will be a random polynomial with a k-sparse arbitrary support vector S as above and following the standard normal distribution $\mathcal{N}(0,1)$. We begin by rewriting the Edelman-Kostlan integral so that it is more suited for our purposes. Let $g_S(x)$ be defined as follows

$$g_S(x) := \|S(x)\|^2 = \sum_{i=0}^{k-1} x^{2e_i} \,. \tag{3.9}$$

In the following lemma, we show that we can express $Z_I(f, \mathcal{N}(0,1))$ entirely in terms of $g_S(x)$ and its derivatives. In particular for any $g : \mathbb{R} \to \mathbb{R}^+$ which is a differentiable function such that $g^{-1}(0)$ is finite, we may define the function $\mathcal{I}(g) : \mathbb{R} \to \mathbb{R}$ as follows

$$\mathcal{I}(g) := \left(\frac{g'(x)}{g(x)}\right)' + \frac{g'(x)}{xg(x)} = (\log(g(x)))'' + \frac{(\log(g(x)))'}{x} \,. \tag{3.10}$$

Note that whenever the Edelman-Kostlan integral is well-defined, the conditions on g which make $\mathcal{I}(g)$ well-defined and non-negative are also satisfied. We now give our alternative formulation.

Lemma 3.2. *For f a random polynomial with support vector S and a set $I \subseteq \mathbb{R}$, we have the following equality for $Z_I(f, \mathcal{N}(0,1))$*

$$Z_I\left(f, \mathcal{N}(0,1)\right) = \frac{1}{2\pi} \int_I \sqrt{\mathcal{I}(g_S(x))}\, dx \,.$$

Proof. We can rewrite Equation (3.8) as

$$Z_I\left(f, \mathcal{N}(0,1)\right) = \frac{1}{\pi} \int_I \frac{\sqrt{g_S(x)\,\|S'(x)\|^2 - \left(S(x) \cdot S'(x)\right)^2}}{g_S(x)}\, dx \,. \tag{3.11}$$

We now examine some simple equalities that allow us to express everything in terms of $g_S(x)$ as in Equation (3.10). We begin with the following equality for $S(x) \cdot S'(x)$.

$$S(x) \cdot S'(x) = \sum_{i=0}^{k-1} e_i x^{2e_i - 1} = \frac{g'_S(x)}{2} \,. \tag{3.12}$$

Similarly, we also have the following equality for $(\|S'(x)\|_2)^2$.

$$\begin{aligned}
(\|S'(x)\|)^2 &= \sum_{i=0}^{k-1} e_i^2 x^{2e_i - 2} = \frac{1}{4}\left(\sum_{i=0}^{k-1} 4e_i^2 x^{2e_i-2}\right) \\
&= \frac{1}{4}\left(\sum_{i=0}^{k-1} ((2e_i(2e_i-1)) + 2e_i) x^{2e_i-2}\right) \\
&= \frac{1}{4}\left(\sum_{i=0}^{k-1} (2e_i(2e_i-1)) x^{2e_i-2}\right) + \frac{1}{4}\left(\sum_{i=0}^{k-1} 2e_i x^{2e_i-2}\right) \\
&= \frac{1}{4} g''_S(x) + \frac{1}{4x} g'_S(x) \,.
\end{aligned} \tag{3.13}$$

Substituting Equation (3.12) and Equation (3.13) into Equation (3.11), we obtain

$$\begin{aligned}
Z_I\left(f, \mathcal{N}(0,1)\right) &= \frac{1}{\pi} \int_I \sqrt{\frac{1}{4}\left(\frac{g_S(x)(g''_S(x) + \frac{1}{x} g'_S(x)) - \left(g'_S(x)\right)^2}{(g_S(x))^2}\right)} \, dx \\
&= \frac{1}{2\pi} \int_I \sqrt{\frac{g''_S(x)}{g_S(x)} - \left(\frac{g'_S(x)}{g_S(x)}\right)^2 + \frac{g'_S(x)}{x g_S(x)}} \, dx \\
&= \frac{1}{2\pi} \int_I \sqrt{\left(\frac{g'_S(x)}{g_S(x)}\right)' + \frac{g'_S(x)}{x g_S(x)}} \, dx \\
&= \frac{1}{2\pi} \int_I \sqrt{\mathcal{I}(g_S(x))} \, dx \,.
\end{aligned}$$

□

We could deduce that the above integrand is always well-formed, that is $\mathcal{I}(g_S(x)) > 0$ since it was derived from a well-formed integral that represents the length of a vector. This statement appears without proof in the original publication [43] due to constraints on the number of pages, but for the sake of formality and completeness, we proceed with a simple proof that this is indeed the case.

Proposition 3.3. *Let $\mathcal{I}(g_S(x))$ be defined as in Equation* (3.10)*. We then have that for all $x \in \mathbb{R} \setminus \{0\}$ that*

$$\mathcal{I}(g_S(x)) > 0 \,.$$

Proof. As we've already seen in the proof of Lemma 3.2, we can write $\mathcal{I}(g_S(x))$ as

$$\begin{aligned}\mathcal{I}(g_S(x)) &= \frac{1}{4}\left(\frac{g_S(x)(g''_S(x)+\frac{1}{x}g'_S(x))-(g'_S(x))^2}{(g_S(x))^2}\right)\\ &= \frac{1}{4}\frac{g_S(x)g''_S(x)-g'_S(x)g'_S(x)+\frac{1}{x}g_S(x)g'_S(x)}{g_S^2(x)}\,. \end{aligned} \tag{3.14}$$

Recall the definition of $g_S(x)$ in Equation (3.9) from which it follows that the denominator is always positive for $x \in \mathbb{R}\setminus\{0\}$. From the same definition we can also derive that

$$g'_S(x) = \sum_{i=0}^{k-1} 2e_i x^{2e_i-1}\,, \tag{3.15}$$

$$g''_S(x) = \sum_{i=0}^{k-1} 2e_i\,(2e_i-1)\,x^{2e_i-2}\,. \tag{3.16}$$

By substitution, we have that

$$g_S(x)g''_S(x) = \left(\sum_{j=0}^{k-1} x^{2e_j}\right)\left(\sum_{i=0}^{k-1} 2e_i\,(2e_i-1)\,x^{2e_i-2}\right) = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} 2e_i\,(2e_i-1)\,x^{2e_i+2e_j-2} \tag{3.17}$$

$$g'_S(x)g'_S(x) = \left(\sum_{i=0}^{k-1} 2e_i x^{2e_i-1}\right)\left(\sum_{j=0}^{k-1} 2e_j x^{2e_j-1}\right) = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} 4e_ie_j x^{2e_i+2e_j-2} \tag{3.18}$$

$$\frac{1}{x}g_S(x)g'_S(x) = \frac{1}{x}\left(\sum_{j=0}^{k-1} x^{2e_j}\right)\left(\sum_{i=0}^{k-1} 2e_i x^{2e_i-1}\right) = \sum_{i=0}^{k-1}\sum_{i=0}^{k-1} 2e_i x^{2e_i+2e_j-2}\,. \tag{3.19}$$

Thus the numerator of the expression in Equation (3.14), using Equation (3.17), Equation (3.18) and Equation (3.19) can be written as

$$\begin{aligned}&g_S(x)g''_S(x)-g'_S(x)g'_S(x)+\frac{1}{x}g_S(x)g'_S(x) =\\ &\sum_{i=0}^{k-1}\sum_{j=0}^{k-1}\left(4e_i^2-2e_i-4e_ie_j+2e_i\right)x^{2e_i+2e_j-2} = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} 4\left(e_i^2-e_ie_j\right)x^{2e_i+2e_j-2}\,. \end{aligned} \tag{3.20}$$

Now we need only to rewrite the sum in a convenient way. Consider all possible indices (i,j) with $0 \le i,j \le k-1$. For every index pair (i,j) there is a distinct index pair (j,i) except when $i=j$. However, note that the summand becomes 0 for $i=j$. If we group together the indices (i,j) and (j,i) for each i,j in the above range, we thus can write Equation (3.20) as

$$\sum_{i=0}^{k-1}\sum_{j=0}^{k-1} 4\left(e_i^2 - e_ie_j\right)x^{2e_i+2e_j-2} = \sum_{\substack{(i,j),i<j\\0\le i,j\le k-1}} 4\left(e_i^2 - e_ie_j + e_j^2 - e_ie_j\right)x^{2e_i+2e_j-2}$$
$$= \sum_{i=0}^{k-1}\sum_{j=i+1}^{k-1} 4\left(e_i^2 - 2e_ie_j + e_j^2\right)x^{2e_i+2e_j-2} = \sum_{i=0}^{k-1}\sum_{j=i+1}^{k-1} 4\left(e_i - e_j\right)^2 x^{2(e_i+2e_j-1)} > 0\,.$$

That the quantity is positive now follows easily. The only point to remark is the exponent of x above. Since $0 = e_0 < e_1 < \cdots < e_{k-1}$, all e_i and e_j are nonnegative and also by the sum conditions have so be unequal. At the worst case we have $i = 0$ and $j = 1$ which means $e_i + e_j = 0 + e_1 \ge 1$, therefore x in the above is always raised in an even power, thus must always be positive since $x \neq 0$.

Since both the numerator and denominator are always positive, so is the quantity, which is precisely what was claimed. □

The new formulation given in Lemma 3.2 allows us to derive several useful properties of the integrand. To begin with, we can show how the integrand behaves when the input is a product of functions g. Specifically, we have the following.

Lemma 3.4. *Given two functions $g_1, g_2 : \mathbb{R} \to \mathbb{R}^+$ defined as in Equation (3.9), we have that*

$$\sqrt{\mathcal{I}(g_1g_2)} \le \sqrt{\mathcal{I}(g_1)} + \sqrt{\mathcal{I}(g_2)}\,.$$

Proof. We show in fact that the property holds with equality for the $\mathcal{I}$ function. That is

$$\begin{aligned}\mathcal{I}(g_1g_2) &= \left(\log\left(g_1(x)g_2(x)\right)\right)'' + \frac{\left(\log\left(g_1(x)g_2(x)\right)\right)'}{x}\\ &= \left(\log(g_1(x))\right)'' + \frac{\left(\log(g_1(x))\right)'}{x} + \left(\log(g_2(x))\right)'' + \frac{\left(\log(g_2(x))\right)'}{x}\\ &= \mathcal{I}(g_1) + \mathcal{I}(g_2)\,,\end{aligned} \tag{3.21}$$

where in the second line we used the linearity of differentiation and the fact that $\log(g_1g_2) = \log(g_1) + log(g_2)$. Now the claim follows by the subadditivity of the square root, i.e., $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ for non-negative $a, b \in \mathbb{R}^+$. □

The above Lemma allows us to describe relations on the expected number of zeros with respect to set operations performed on the support vector, which essentially is the same as operations on its exponents. Recall that the exponents must be unique and have an inherent order, thus we can always view them as a set without loss of information. In particular, let f be a random polynomial with support vector S. If there exist support vectors S_1 and S_2 such that $S = S_1 \cup S_2$ or $S = S_1 + S_2$, where $+$ is the Minkowski sum, and we view all three support vectors as sets, we can bound the number of zeros of f by the number of zeros of the polynomials f_1 and f_2 with support vectors S_1 and S_2 respectively. Before proceeding further, we recall the definition of the Minkowski sum operation in our context.

Definition 3.5 (Minkowski sum and collission-free sets)**.** *For sets $A, B \subseteq \mathbb{N}$, we define the* Minkowski sum *of A and B as the following operation*

$$A + B \coloneqq \{a + b : a \in A, b \in B\}.$$

where the $a + b$ operation corresponds to the usual addition over $\mathbb{N}$. We say that A and B are collision-free *if $|A + B| = |A|\,|B| = |A \times B|$, i.e., when all the natural numbers $a + b : a \in A, b \in B$ produced by the above operation are mutually distinct.*

Since support vectors have a one-to-one correspondence with subsets of $\mathbb{N}$, we will slightly abuse the notation and write $S = S_1 + S_2$ while referring to the sets of exponents defined by the support vectors. With the Minkowski sum formally defined, we may now show that the expected number of zeros is subadditive with respect to it.

Lemma 3.6. *Let f_1 and f_2 be two random polynomials with support vectors $S_1 \in \mathbb{R}^{k_1}$ and $S_2 \in \mathbb{R}^{k_2}$ respectively, such that S_1 and S_2 are collision-free as sets. Also let f be a random polynomial with support vector $S = S_1 + S_2 \in \mathbb{R}^k$, with $k = k_1 k_2$. Then we have*

$$Z_I(f, \mathcal{D}) \leq Z_I(f_1, \mathcal{D}) + Z_I(f_2, \mathcal{D}).$$

Proof. Let $S = (e_0, \ldots, e_{k-1})$, $S_1 = (\varepsilon_0, \ldots, \varepsilon_{k_1-1})$ and $S_2 = (\delta_0, \ldots, \delta_{k_2-1})$. From the definition of g_S and given that S_1 and S_2 are collision-free, we have

$$\begin{aligned} g_{S_1+S_2}(x) &= \sum_{e_i \in S} x^{2e_i} = \sum_{\substack{e_i = \varepsilon_j + \delta_l \\ \forall \varepsilon_j \in S_1, \delta_l \in S_2}} x^{2(\varepsilon_j + \delta_l)} \\ &= \left(\sum_{\varepsilon_j \in S_1} x^{2\varepsilon_j} \right) \left(\sum_{\delta_l \in S_2} x^{2\delta_l} \right) = g_{S_1}(x)\, g_{S_2}(x). \end{aligned} \tag{3.22}$$

We may now use Lemma 3.2 and substitute for Equation (3.22) above to obtain

$$\begin{aligned} Z_I(f, \mathcal{D}) &= \frac{1}{2\pi} \int_I \sqrt{\mathcal{I}(g_{S_1+S_2}(x))}\, dx = \frac{1}{2\pi} \int_I \sqrt{\mathcal{I}(g_{S_1}(x)\, g_{S_2}(x))}\, dx \\ &\leq \frac{1}{2\pi} \int_I \sqrt{\mathcal{I}(g_{S_1}(x))}\, dx + \frac{1}{2\pi} \int_I \sqrt{\mathcal{I}(g_{S_2}(x))}\, dx \\ &= Z_I(f_1, \mathcal{D}) + Z_I(f_2, \mathcal{D}). \end{aligned}$$

where the inequality in the second line follows from Lemma 3.4. □

Note that the above can be used as the basis of an inductive argument, as long as we find appropriate collision-free subsets of a set. For example, suppose that $S = \left(1, x, \ldots, x^m, x^{d-m}, x^{d-m+1}, \ldots, x_d\right)$ where $d > 2m$. Note that for the exponents of S we have $\{0, 1, \ldots, m, d-m, d-m+1, \ldots, d\} = \{0, 1, \ldots, m\} + \{0, d-m\}$. Let $S_1 = (1, x, \ldots, x^m)$ and $S_2 = (1, x^{d-m})$. Now S_1 is a dense support vector of degree m, thus by Theorem 2.42 the expected number of real zeros of a random polynomial f_1 with support vector S_1 and following $\mathcal{N}(0, 1)$ will be $Z_{\mathbb{R}}(f_1, \mathcal{N}(0, 1)) = \Theta(\log m)$. Similarly a

random polynomial f_2 with support vector S_2 will have at most 2 distinct real zeros by Theorem 2.21 since its sparsity is 2. Therefore random polynomials with support vector S, which will be $2m$ sparse will have only $\Theta(\log m)$ expected real zeros.

We can also provide an inequality for the set operation of union on the support S viewed as a set. Specifically, we wish to establish an upper bound for $z_{S_1 \uplus S_2}$, where $S_1 \uplus S_2$ denotes the disjoint union of S_1 and S_2. First we state the following proposition which is easy to verify.

Proposition 3.7. *Let $S_1 \in \mathbb{R}^{k_1}$ and $S_2 \in \mathbb{R}^{k_2}$ be support vectors, such that S_1 and S_2 are disjoint as sets, i.e., $S_1 \cap S_2 = \emptyset$. Then for $S = S_1 \uplus S_2 \in \mathbb{R}^k$, with $k = k_1 + k_2$, we have that*

$$g_{S_1 \uplus S_2}(x) = g_{S_1}(x) + g_{S_2}(x)\,.$$

Proof. Let $S = (e_0, \ldots, e_{k-1})$, $S_1 = (\varepsilon_0, \ldots, \varepsilon_{k_1-1})$ and $S_2 = (\delta_0, \ldots, \delta_{k_2-1})$. From the definition of g_S and given that S_1 and S_2 are disjoint, we have

$$\begin{aligned} g_{S_1 \uplus S_2}(x) &= \sum_{e_i \in S} x^{2e_i} = \sum_{\substack{(e_i = \varepsilon_j \in S_1) \oplus \\ (e_i = \delta_l \in S_2)}} x^{2e_i} \\ &= \left(\sum_{\varepsilon_j \in S_1} x^{2\varepsilon_j} \right) + \left(\sum_{\delta_l \in S_2} x^{2\delta_l} \right) = g_{S_1}(x) + g_{S_2}(x)\,. \end{aligned}$$

□

We will require the following definitions in the following to simplify the presentation of the lemma for the expected number of zeros under the operation of disjoint union. While the above results can be adapted for any interval I, we focus on $I = (0, 1)$ which is sufficient for our results.

Definition 3.8. *Let S_1, S_2 be two support vectors that are disjoint as sets. Also suppose that $\left(\frac{g_{S_1}}{g_{S_2}}\right)' \geq 0$ at $x = 0$. Let $x_1, \ldots, x_m$ such that $x_i \leq x_{i+1}$ for $i \in \{1, \ldots, m-1\}$ be the critical points of odd multiplicity of $\frac{g_{S_1}}{g_{S_2}}$ in $(0, 1)$, i.e., $f'(x_i) = 0$ and the root at $x = x_i$ has odd multiplicity for $i \in \{1, \ldots, m\}$. We also define $x_0 := 0$ and $x_{m+1} := 1$. We then define the following quantities, for all $0 \leq i \leq m$ and $x \in (0, 1)$.*

$$\begin{aligned} \gamma_{S_1,S_2}(x) &:= \sqrt{\frac{g_{s_1}(x)}{g_{s_2}(x)}}\,, \\ T^i_{S_1,S_2} &:= (-1)^i \left(\arctan\left(\gamma_{S_1,S_2}(x_{i+1})\right) - \arctan\left(\gamma_{S_1,S_2}(x_i)\right) \right)\,, \\ R_{S_1,S_2} &:= \sum_{i=0}^{m} T^i_{S_1,S_2}\,. \end{aligned}$$

We also require the following technical proposition in the proof of the following Lemma.

Proposition 3.9. *The following identity is true for all $a, b, c, d \in \mathbb{R}$ such that $b \neq 0$,$d \neq 0$ and $b \neq -d$*

$$\left(\frac{a+c}{b+d}\right)^2 = \left(\frac{b}{b+d}\right)\left(\frac{a}{b}\right)^2 + \left(\frac{d}{b+d}\right)\left(\frac{c}{d}\right)^2 - \frac{1}{bd}\left(\frac{bc-ad}{b+d}\right)^2 .$$

Proof. The right hand side of the above can be written as

$$\begin{aligned}
&\left(\frac{b}{b+d}\right)\left(\frac{a}{b}\right)^2 + \left(\frac{d}{b+d}\right)\left(\frac{c}{d}\right)^2 - \frac{1}{bd}\left(\frac{bc-ad}{b+d}\right)^2 \\
&= \frac{a^2}{b(b+d)} + \frac{c^2}{d(b+d)} - \frac{b^2c^2 + a^2d^2 - 2abcd}{bd(b+d)^2} \\
&= \frac{a^2d(b+d) + bc^2(b+d) - b^2c^2 - a^2d^2 + 2abcd}{bd(b+d)^2} \\
&= \frac{a^2bd + a^2d^2 + b^2c^2 + bc^2d - b^2c^2 - a^2d^2 + 2abcd}{bd(b+d)^2} \\
&= \frac{bd\left(a^2 + 2ac + c^2\right)}{bd(b+d)^2} = \left(\frac{a+c}{b+d}\right)^2 .
\end{aligned}$$

□

We may now state the key result about the behavior of the expected number of zeros with respect to the disjoint union operation on the support vectors.

Lemma 3.10. *Let f_1 and f_2 be two random polynomials with support vectors $S_1 \in \mathbb{R}^{k_1}$ and $S_2 \in \mathbb{R}^{k_2}$ respectively, such that S_1 and S_2 are disjoint as sets, $S_1 \cap S_2 = \emptyset$. Also let f be a random polynomial with support vector $S = S_1 \uplus S_2 \in \mathbb{R}^k$, with $k = k_1 + k_2$. Assume that $\left(\frac{g_{S_1}(x)}{g_{S_2}(x)}\right)' \geq 0$ for $x = 0$. Then we have*

$$Z_{(0,1)}(f, \mathcal{N}(0,1)) \leq Z_{(0,1)}(f_1, \mathcal{N}(0,1)) + Z_{(0,1)}(f_2, \mathcal{N}(0,1)) + \frac{1}{\pi} R_{S_1,S_2} .$$

Proof. Note that the assumption that $\left(\frac{g_{S_1}(x)}{g_{S_2}(x)}\right)' \geq 0$ for $x = 0$ is always satisfied up to renaming S_1 and S_2. This is actually the case for any x, since we have

$$\left(\frac{g_{S_1}(x)}{g_{S_2}(x)}\right)' = \frac{g'_{S_1}(x)g_{S_2}(x) - g_{S_1}(x)g'_{S_2}(x)}{\left(g_{S_2}(x)\right)^2} \tag{3.23}$$

$$\left(\frac{g_{S_2}(x)}{g_{S_1}(x)}\right)' = \frac{g_{S_1}(x)g'_{S_2}(x) - g'_{S_1}(x)g_{S_2}(x)}{\left(g_{S_1}(x)\right)^2} = \frac{-\left(g'_{S_1}(x)g_{S_2}(x) - g_{S_1}(x)g'_{S_2}(x)\right)}{\left(g_{S_1}(x)\right)^2} . \tag{3.24}$$

Since the numerators of the two expressions is the same quantity with opposite sign and the denominators are positive, the two quantities are also of opposite sign thus one must always be non-negative. Even more so, in the case we are interested in, recall that we demand the support vector S to have the constant term $x^0 = 1$ as a component. Since S_1 and S_2 are disjoint and $S = S_1 \uplus S_2$, only one of them will also have the constant term

as a component, say S_2 without loss of generality. Thus $g_{S_2}(0) = 1$, while $g_{S_1}(0) = 0$, so we must choose Equation (3.23). By substitution we can then write it as

$$\frac{g'_{S_1}(x)1 - 0g'_{S_2}(x)}{1^2} = g'_{S_1}(x) = \sum_{\varepsilon_i \in S_1} 2\varepsilon_i x^{2\varepsilon_i - 1}.$$

Now note that since $0 \notin S_1$, we must have that $g'_{S_1}(0) = 0$ and thus the assumption is always satisfied.

We begin by using Proposition 3.7 to derive an expression for the $\mathcal{I}$ function when the support vector is expressed as a disjoint union. We have

$$\begin{aligned}
\mathcal{I}(g_{S_1 \uplus S_2}) &\overset{(3.7)}{=} \mathcal{I}(g_{S_1} + g_{S_2}) = \left(\frac{(g_{S_1} + g_{S_2})'}{g_{S_1} + g_{S_2}}\right)' + \frac{(g_{S_1} + g_{S_2})'}{x\,(g_{S_1} + g_{S_2})} \\
&= \frac{(g_{S_1} + g_{S_2})''\,(g_{S_1} + g_{S_2}) - \left((g_{S_1} + g_{S_2})'\right)^2}{(g_{S_1} + g_{S_2})^2} + \frac{(g_{S_1} + g_{S_2})'}{x\,(g_{S_1} + g_{S_2})} \\
&= \frac{g''_{S_1} + g''_{S_2}}{g_{S_1} + g_{S_2}} - \left(\frac{g'_{S_1} + g'_{S_2}}{g_{S_1} + g_{S_2}}\right)^2 + \frac{1}{x}\left(\frac{g'_{S_1} + g'_{S_2}}{g_{S_1} + g_{S_2}}\right). \qquad (3.25)
\end{aligned}$$

Note the middle term corresponds to the left hand side in Proposition 3.9 for $a = g'_{S_1}$, $b = g_{S_1}$, $c = g'_{S_2}$ and $d = g_{S_2}$. We thus have

$$\left(\frac{g'_{S_1} + g'_{S_2}}{g_{S_1} + g_{S_2}}\right)^2 = \frac{g_{S_1}}{g_{S_1} + g_{S_2}}\left(\frac{g'_{S_1}}{g_{S_1}}\right)^2 + \frac{g_{S_2}}{g_{S_1} + g_{S_2}}\left(\frac{g'_{S_2}}{g_{S_2}}\right)^2 - \frac{1}{g_{S_1} g_{S_2}}\left(\frac{g_{S_1} g'_{S_2} - g'_{S_1} g_{S_2}}{g_{S_1} + g_{S_2}}\right)^2. \qquad (3.26)$$

By substituting Equation (3.26) into Equation (3.25) we thus obtain

$$\begin{aligned}
\mathcal{I}(g_{S_1 \uplus S_2}) &= \frac{g_{S_1}}{g_{S_1} + g_{S_2}}\left(\frac{g''_{S_1}}{g_{S_1}} - \left(\frac{g'_{S_1}}{g_{S_1}}\right)^2 + \frac{g'_{S_1}}{x g_{S_1}}\right) + \frac{g_{S_2}}{g_{S_1} + g_{S_2}}\left(\frac{g''_{S_2}}{g_{S_2}} - \left(\frac{g'_{S_2}}{g_{S_2}}\right)^2 + \frac{g'_{S_2}}{x g_{S_2}}\right) \\
&\quad + \frac{1}{g_{S_1} g_{S_2}}\left(\frac{g_{S_1} g'_{S_2} - g'_{S_1} g_{S_2}}{g_{S_1} + g_{S_2}}\right)^2 \\
&= \frac{g_{S_1}}{g_{S_1} + g_{S_2}}\mathcal{I}(g_{S_1}) + \frac{g_{S_2}}{g_{S_1} + g_{S_2}}\mathcal{I}(g_{S_2}) + \frac{1}{g_{S_1} g_{S_2}}\left(\frac{g_{S_1} g'_{S_2} - g_{S_2} g'_{S_1}}{g_{S_1} + g_{S_2}}\right)^2. \qquad (3.27)
\end{aligned}$$

We now can compute a bound for the number of zeros of the random polynomial f with support vector $S = S_1 \uplus S_2$. Using Equation (3.27) we have

$$
\begin{aligned}
Z_{(0,1)}(f,\mathcal{N}(0,1)) &= \frac{1}{2\pi}\int_0^1 \sqrt{\mathcal{I}(g_{S_1\uplus S_2}(x))}\,dx \\
&= \frac{1}{2\pi}\int_0^1 \sqrt{\frac{g_{S_1}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_1})+\frac{g_{S_2}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_2})+\frac{1}{g_{S_1}g_{S_2}}\left(\frac{g_{S_1}g'_{S_2}-g_{S_2}g'_{S_1}}{g_{S_1}+g_{S_2}}\right)^2}\,dx \\
&= \frac{1}{2\pi}\int_0^1 \sqrt{\frac{g_{S_1}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_1})+\frac{g_{S_2}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_2})+\frac{1}{g_{S_1}g_{S_2}}\left(\frac{g_{S_2}g'_{S_1}-g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\right)^2}\,dx \\
&\qquad\qquad (x^2=(-x)^2) \\
&\le \frac{1}{2\pi}\left(\int_0^1 \sqrt{\mathcal{I}(g_{S_1})}\,dx+\int_0^1 \sqrt{\mathcal{I}(g_{S_2})}\,dx+\int_0^1 \left|\frac{1}{\sqrt{g_{S_1}g_{S_2}}}\left(\frac{g_{S_2}g'_{S_1}-g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\right)\right|\,dx\right) \\
&= Z_{(0,1)}(f_1,\mathcal{N}(0,1))+Z_{(0,1)}(f_2,\mathcal{N}(0,1))+\frac{1}{2\pi}\int_0^1 \left|\frac{1}{\sqrt{g_{S_1}g_{S_2}}}\left(\frac{g_{S_2}g'_{S_1}-g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\right)\right|\,dx.
\end{aligned}
\tag{3.28}
$$

In the third equality above we used that $x^2=(-x)^2$ to write the numerator so that it resembles the numerator of $\left(\frac{g_{S_1}}{g_{S_2}}\right)'$, to make the analysis below more aesthetically pleasing, although this step is not required.

Comparing the above to the desired result, it is clear that what remains is to bound the value of the last term. Note that the only expression in the absolute value that can change sign is the numerator $g_{S_2}g'_{S_1}-g_{S_1}g'_{S_2}$. As it can be seen in Equation (3.23), this numerator is positive if and only if $\left(\frac{g_{S_1}}{g_{S_2}}\right)'$. Therefore $g_{S_2}g'_{S_1}-g_{S_1}g'_{S_2}$ changes sign exactly on the critical points of odd multiplicity of $\frac{g_{S_1}}{g_{S_2}}$. As in Definition 3.8, let those points in $(0,1)$ be $x_1,\dots,x_m$ and in addition $x_0=0$ and $x_{m+1}=1$. We can thus partition $(0,1)$ into subintervals where the above expression is nonnegative or not. Let (x_i,x_{i+1}) be such a subinterval where the expression is nonnegative. We then can write the above integral as follows

$$
\begin{aligned}
&\int_{x_i}^{x_{i+1}} \frac{1}{\sqrt{g_{S_1}g_{S_2}}}\left(\frac{g_{S_2}g'_{S_1}-g_{S_1}g'_{S_2}}{g_{S_2}^2}\right)\left(\frac{g_{S_2}^2}{g_{S_1}+g_{S_2}}\right)\,dx \\
&= \int_{x_i}^{x_{i+1}} \sqrt{\frac{g_{S_2}}{g_{S_1}}}\left(\frac{g_{S_2}g'_{S_1}-g_{S_1}g'_{S_2}}{g_{S_2}^2}\right)\left(\frac{g_{S_2}}{g_{S_1}+g_{S_2}}\right)\,dx \\
&= 2\int_{x_i}^{x_{i+1}} \frac{1}{2}\frac{1}{\sqrt{\frac{g_{S_1}}{g_{S_2}}}}\left(\frac{g_{s_1}}{g_{s_2}}\right)'\left(\frac{1}{1+\frac{g_{s_1}}{g_{s_2}}}\right)\,dx \\
&= 2\int_{x_i}^{x_{i+1}} \left(\sqrt{\frac{g_{s_1}}{g_{s_2}}}\right)'\left(\frac{1}{1+\left(\sqrt{\frac{g_{s_1}}{g_{s_2}}}\right)^2}\right)\,dx\,.
\end{aligned}
\tag{3.29}
$$

We may now perform the change of variables $u=\sqrt{\frac{g_{s_1}}{g_{s_2}}}$. Let $\alpha=\gamma_{S_1,S_2}(x_i)=\sqrt{\frac{g_{s_1}(x_i)}{g_{s_2}(x_i)}}$

and $\beta = \gamma_{S_1,S_2}(x_{i+1})$, where γ_{S_1,S_2} matches the definition given in Definition 3.8. We thus can write Equation (3.29) as

$$\int_{x_i}^{x_{i+1}} \frac{1}{\sqrt{g_{S_1} g_{S_2}}} \left(\frac{g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}}{g_{S_2}^2} \right) \left(\frac{g_{S_2}^2}{g_{S_1} + g_{S_2}} \right) dx = 2 \int_{\alpha}^{\beta} \left(\frac{1}{1+u^2} \right) du$$
$$= 2(\arctan(\beta) - \arctan(\alpha)) = 2 \left(\arctan(\gamma_{S_1,S_2}(x_{i+1})) - \arctan(\gamma_{S_1,S_2}(x_i)) \right) . \quad (3.30)$$

For intervals where $g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}$ is negative instead, we obtain a similar expression with an additional minus sign, thus for such an interval (x_{i+1}, x_{i+2}) we have that

$$- \int_{x_{i+1}}^{x_{i+2}} \frac{1}{\sqrt{g_{S_1} g_{S_2}}} \left(\frac{g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}}{g_{S_2}^2} \right) \left(\frac{g_{S_2}^2}{g_{S_1} + g_{S_2}} \right) dx$$
$$= \int_{x_{i+1}}^{x_{i+2}} \frac{1}{\sqrt{g_{S_1} g_{S_2}}} \left(\frac{g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}}{g_{S_2}^2} \right) \left(\frac{g_{S_2}^2}{g_{S_1} + g_{S_2}} \right)$$
$$= 2(\arctan(\gamma_{S_1,S_2}(x_{i+1})) - \arctan(\gamma_{S_1,S_2}(x_{i+2}))) . \quad (3.31)$$

Now we may utilize the remaining definitions in Definition 3.8, Equation (3.30) and Equation (3.31) to express the integral for the entire interval $(0, 1)$. Furthermore from our assumption, for the interval $(0, x_1)$ we have $\left(\frac{g_{S_1}(x)}{g_{S_2}(x)} \right)' \geq 0$. We may thus write

$$\int_0^1 \left| \frac{1}{\sqrt{g_{S_1} g_{S_2}}} \left(\frac{g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}}{g_{S_1} + g_{S_2}} \right) \right| dx = \sum_{i=0}^{m} \int_{x_i}^{x_{i+1}} \left| \frac{1}{\sqrt{g_{S_1} g_{S_2}}} \left(\frac{g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}}{g_{S_1} + g_{S_2}} \right) \right| dx$$
$$= \sum_{i=0}^{m} 2(-1)^i \left(\arctan(\gamma_{S_1,S_2}(x_{i+1})) - \arctan(\gamma_{S_1,S_2}(x_i)) \right) = 2 \sum_{i=0}^{m} T^i_{S_1,S_2} = 2R_{S_1,S_2} . \quad (3.32)$$

Therefore, by finally substituting Equation (3.32) into Equation (3.28) we finally obtain

$$Z_{(0,1)}(f, \mathcal{N}(0,1)) =$$
$$Z_{(0,1)}(f_1, \mathcal{N}(0,1)) + Z_{(0,1)}(f_2, \mathcal{N}(0,1)) + \frac{1}{2\pi} \int_0^1 \left| \frac{1}{\sqrt{g_{S_1} g_{S_2}}} \left(\frac{g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}}{g_{S_1} + g_{S_2}} \right) \right| dx$$
$$\leq Z_{(0,1)}(f_1, \mathcal{N}(0,1)) + Z_{(0,1)}(f_2, \mathcal{N}(0,1)) + \frac{1}{\pi} R_{S_1,S_2} ,$$

as stated in the Lemma. □

Lemma 3.10 will be very useful in the proof of our main result as the basis of an iterative argument. The overarching proof idea is given a support vector S, which when viewed as a set can be expressed as two carefully chosen disjoint unions, say S_1 and S_2. These not only give rise to support vectors, but also since the union is disjoint, they are

of strictly smaller sparsity. In particular, for our strategy we elect to keep one vector, say S_1, rather large and obeying the restriction of $0 \in S_1$, thus satisfying our conditions for support vectors. In fact, if S has sparsity k, we choose S_1 to have sparsity $k-1$. This implies that S_2 will have sparsity 1 and thus will be easy to determine its number of zeros.

We first determine the expected number of zeros for the base cases of very low sparsity $k = 1$ and $k = 2$. Note that as the statements of the above Lemma might betray, it is sufficient to consider the interval $(0, 1)$, since the sample space $\mathfrak{F}$ with respect to the standard normal distribution $\mathcal{N}(0, 1)$ is closed with respect to the operations given in Equation (2.21) and *Equation* (2.22). We have that

Lemma 3.11. *Let f be a random polynomial with support vector $S = (x^e)$ for $e \in \mathbb{N}$. Then we have*

$$Z_{(0,1)}(f, \mathcal{N}(0, 1)) = 0\,.$$

Proof. The Lemma is easy to prove via a variety of ways, for example one may examine that any nonzero random polynomial ax^e only has zeros at $x = 0 \notin (0, 1)$ or by invoking Theorem 2.17, since a single term exhibits no variations and $(0, 1) \subset \mathbb{R}^+$. We verify this using our already established methodology, which also serves consistency. Observe that for $S = (x^e)$, it is $g_S(x) = x^{2e}$. Hence

$$\mathcal{I}(g_S) = (2e \log x)'' + \frac{(2e \log x)'}{x} = -\frac{2e}{x^2} + \frac{2e}{x^2} = 0\,. \tag{3.33}$$

It therefore follows that

$$Z_{(0,1)}(f, \mathcal{N}(0, 1)) = \int_0^1 0 \, dx = 0\,.$$

□

While we may use $k = 1$ as the base case, it turns out that the result for $k = 2$ is easy to derive exactly and is non-zero, thus we elect to use it instead. We may prove the following

Lemma 3.12. *Let f be a random polynomial with support vector $S = (1, x^e)$ for $e \in \mathbb{N}^+$. Then we have*

$$Z_{(0,1)}(f, \mathcal{N}(0, 1)) = \frac{1}{4}\,.$$

Proof. We can easily calculate that $g_S(x) = 1 + x^{2e}$. We can thus calculate the following:

$$\begin{aligned}\mathcal{I}(g_S(x)) &= \left(\frac{2ex^{2e-1}}{(1+x^{2e})}\right)' + \frac{2ex^{2e-1}}{x(1+x^{2e})} \\ &= \frac{2e(2e-1)x^{2e-2}(1+x^{2e}) - 4e^2x^{4e-2}}{(1+x^{2e})^2} + \frac{2ex^{2e-2}(1+x^{2e})}{(1+x^{2e})^2} \\ &= \frac{4e^2x^{2e-2} - 2ex^{2e-2} + 4e^2x^{4e-2} - 2ex^{4e-2} - 4e^2x^{4e-2} + 2ex^{2e-2} + 2ex^{4e-2}}{(1+x^{2e})^2} \\ &= \frac{4e^2x^{2e-2}}{(1+x^{2e})^2} = \left(\frac{2ex^{e-1}}{1+x^{2e}}\right)^2 . \end{aligned} \tag{3.34}$$

Substituting Equation (3.34) into the calculation for the expected number of zeros we obtain

$$Z_{(0,1)}(f,\mathcal{N}(0,1)) = \frac{1}{2\pi}\int_0^1 \sqrt{\left(\frac{2ex^{e-1}}{1+x^{2e}}\right)^2}\,dx = \frac{1}{\pi}\int_0^1 \frac{ex^{e-1}}{1+x^{2e}}\,dx = \frac{1}{\pi}\int_0^1 \frac{(x^e)'}{1+x^{2e}}\,dx\,. \tag{3.35}$$

We perform the change of variable $u = x^e$ for Equation (3.35) and obtain the now familiar integral

$$Z_{(0,1)}(f,\mathcal{N}(0,1)) = \frac{1}{\pi}\int_0^1 \frac{1}{1+u^2}\,du = \frac{1}{\pi}\left(\arctan(1) - \arctan(0)\right) = \frac{1}{\pi}\left(\frac{\pi}{4} - 0\right) = \frac{1}{4}\,.$$

□

We are now ready to prove what forms the iteration step in our proof. Together with Lemma 3.10 they form the main technical parts of the proof of our main theorem. The following lemma essentially upper bounds the "overhead" for reducing the sparsity by 1 in our proof.

Lemma 3.13. *Let f be a random polynomial with support vector S following $\mathcal{N}(0,1)$, such that $x^0 \in S$ and $|S| = k$. Let $e_{k-1} \in \mathbb{N}^+$ be the largest exponent in S. We let f_1 be a random polynomial with support vector $S_1 = (e_{k-1})$ and f_2 also a random polynomial with support vector $S_2 = (x^0, x^{e_1}, \ldots, x^{e_{k-2}})$, both following $\mathcal{N}(0,1)$. We then have*

$$Z_{(0,1)}(f,\mathcal{N}(0,1)) \leq Z_{(0,1)}(f_2,\mathcal{N}(0,1)) + \frac{1}{\pi}\arctan\left(\frac{1}{\sqrt{k}}\right).$$

Proof. Since $S = S_1 \uplus S_2$ we will make use of Lemma 3.10. Recall from Definition 3.8 that doing so requires identifying the number of critical points of odd multiplicity of $\frac{g_{S_1}}{g_{S_2}}$. Substituting $g_{S_1} = x^{2e_{k-1}}$ and $g_{S_2} = \sum_{i=0}^{k-2} x^{2e_i}$, we obtain

$$\left(\frac{g_{S_1}}{g_{S_2}}\right)' = \frac{1}{g_{S_2}^2}\left(2e_{k-1}x^{2e_{k-1}-1}\sum_{i=0}^{k-2}x^{2e_i} - x^{2e_{k-1}}\sum_{i=0}^{k-2}2e_i x^{2e_i-1}\right)$$
$$= \frac{2x^{e_{k-1}-1}}{g_{S_2}^2}\left(\sum_{i=0}^{k-2}e_{k-1}x^{2e_i} - \sum_{i=0}^{k-2}e_i x^{2e_i}\right) = \frac{2x^{e_{k-1}-1}}{g_{S_2}^2}\sum_{i=0}^{k-2}(e_{k-1}-e_i)\,x^{2e_i} > 0\,. \tag{3.36}$$

The last inequality follows since $x \in (0,1)$ and $e_{k-1} > e_i$ for all $i = 0, \ldots, k-2$. It is clear that $\frac{g_{S_1}}{g_{S_2}}$ has no critical points in the interval of interest and in fact in the entirety of $\mathbb{R}^+$. Therefore, by applying Definition 3.8, we have that $m = 0$, $x_0 = 0$ and $x_1 = 1$ and we obtain

$$R_{S_1,S_2} = T^0_{S_1,S_2} = \arctan\left(\sqrt{\frac{g_{S_1}(1)}{g_{S_2}(1)}}\right) - \arctan\left(\sqrt{\frac{g_{S_1}(0)}{g_{S_2}(0)}}\right)$$
$$= \arctan\left(\sqrt{\frac{1}{k}}\right) - \arctan(0) = \arctan\left(\sqrt{\frac{1}{k}}\right). \tag{3.37}$$

Since $S_1 = (x^{2e_{k-1}})$ we may utilize Lemma 3.11 to obtain

$$Z_{(0,1)}(f_1, \mathcal{N}(0,1)) = 0\,. \tag{3.38}$$

Plugging Equation (3.37) and Equation (3.38) into the expression of *Lemma* 3.10, we thus have

$$Z_{(0,1)}(f, \mathcal{N}(0,1)) \leq Z_{(0,1)}(f_1, \mathcal{N}(0,1)) + Z_{(0,1)}(f_2, \mathcal{N}(0,1)) + \frac{1}{\pi}R_{S_1,S_2}$$
$$= 0 + Z_{(0,1)}(f_2, \mathcal{N}(0,1)) + \frac{1}{\pi}\arctan\left(\sqrt{\frac{1}{k}}\right)$$
$$= Z_{(0,1)}(f_2, \mathcal{N}(0,1)) + \frac{1}{\pi}\arctan\left(\sqrt{\frac{1}{k}}\right).$$

□

The proof of the main theorem is now straightforward. We simply have to apply Lemma 3.13 iteratively until we reach the base case of $k = 2$.

Theorem 3.14 (Real zeros of k-sparse polynomials following $\mathcal{N}(0,1)$)**.** *Let f be a random polynomial with support vector $S = (1, x^{e_1}, \ldots, x^{e_{k-1}})$ following $\mathcal{N}(0,1)$ such that $|S| = k$. We then have*

$$Z_{\mathbb{R}}(f, \mathcal{N}(0,1)) \leq 1 + \frac{8}{\pi}(\sqrt{k-1} - 1) \leq \frac{8}{\pi}\sqrt{k-1}\,.$$

Proof. We obtain a similar result for $(0,1)$ before easily generalizing to the entirety of $\mathbb{R}$. For $k \leq 2$, the bound holds taken into consideration Lemma 3.11 and $0 \leq \frac{8}{\pi}\sqrt{0}$ for $k = 1$, and Lemma 3.12 and $\frac{1}{4} \leq \frac{8}{\pi}$ for $k = 2$, so we may assume $k > 2$. We will apply Lemma 3.13 $k-2$ times, one time for each iteration step. Since for each such step the sparsity decreases by one, after $k-2$ steps the sparsity will be 2 and we may apply Lemma 3.12 to this base case. Let f_i be the random polynomial after step i for $i = 1, \ldots, k-2$. f_i has support vector $S_i = (1, x^{e_1}, \ldots, x^{e_{k-1-i}})$ and coefficient vector $A = (a_0, \ldots, a_{k-1-i})$ and therefore $f(i) = \sum_{j=0}^{k-1-i} a_j x^{e_j}$. By convention let $f_0 = f$ with support vector $S_0 = S$. We obtain the following inequality

$$\begin{aligned} Z_{(0,1)}(f, \mathcal{N}(0,1)) &\leq Z_{(0,1)}(f_1, \mathcal{N}(0,1)) + \frac{1}{\pi} \arctan\left(\frac{1}{\sqrt{k}}\right) \leq \ldots \\ &\leq Z_{(0,1)}(f_i, \mathcal{N}(0,1)) + \frac{1}{\pi} \sum_{j=k-i}^{k-1} \arctan\left(\frac{1}{\sqrt{j}}\right) \leq \ldots \\ &\leq Z_{(0,1)}(f_{k-2}, \mathcal{N}(0,1)) + \frac{1}{\pi} \sum_{j=2}^{k-1} \arctan\left(\frac{1}{\sqrt{i}}\right) \\ &= \frac{1}{4} + \frac{1}{\pi} \sum_{j=2}^{k-1} \arctan\left(\frac{1}{\sqrt{i}}\right). \end{aligned} \tag{3.39}$$

Let $h(x) = x - \arctan(x)$. We have that $h'(x) = \frac{x^2}{x^2+1}$ so the function is monotone for $x \in \mathbb{R}^+$ increasing and in addition $h(0) = 0$. It follows that for $x \in \mathbb{R}^+$

$$\arctan(x) < x\,. \tag{3.40}$$

By using Equation (3.40) into Equation (3.39) we thus obtain

$$\begin{aligned} Z_{(0,1)}(f, \mathcal{N}(0,1)) &\leq \frac{1}{4} + \frac{1}{\pi} \sum_{j=2}^{k-1} \arctan\left(\frac{1}{\sqrt{j}}\right) \leq \frac{1}{4} + \frac{1}{\pi} \sum_{j=2}^{k-1} \frac{1}{\sqrt{j}} \\ &\leq \frac{1}{4} + \frac{1}{\pi} \int_1^{k-1} \frac{1}{\sqrt{i}}\, dx = \frac{1}{4} + \frac{2}{\pi}\left(\sqrt{k-1} - 1\right). \end{aligned} \tag{3.41}$$

Now since the sample space $\mathfrak{F}$ of random k-sparse polynomials with respect to $\mathcal{N}(0,1)$ is closed under Equation (2.21) and Equation (2.22), we have that

$$Z_{\mathbb{R}}\left(f, \mathcal{N}(0,1)\right) = 4Z_{(0,1)}\left(f, \mathcal{N}(0,1)\right) \leq 1 + \frac{8}{\pi}\left(\sqrt{k-1} - 1\right) \leq \frac{8}{\pi}\sqrt{k-1}\,,$$

which is precisely the claimed upper bound.

□

We are also able to show that in fact most of the real roots $(0,1)$ cluster near $x = 1$. By the symmetries given by Equation (2.21) and Equation (2.22), this implies that most of the real zeros of a random polynomial following $\mathcal{N}(0,1)$ lie in a small neighborhood around $x = -1$ and $x = 1$. We first require the following technical proposition.

Proposition 3.15. *For all $x \in (0,1)$, we have*

$$\sqrt{\sum_{e=0}^{\infty} e^2 x^{2e-2}} \leq \frac{1}{1-x^2} + \frac{2x}{(1-x^2)^{\frac{3}{2}}}\,.$$

Proof. It is well-known that we may write the following as a geometric series

$$\frac{1}{1-x^2} = \sum_{e=0}^{\infty} x^{2e}\,. \tag{3.42}$$

We may use Equation (3.42) to derive

$$\sum_{e=0}^{\infty} 2e(2e-1)x^{2e-2} = \left(\sum_{e=0}^{\infty} x^{2e}\right)'' = \left(\frac{1}{1-x^2}\right)'' = \left(\frac{2x}{(1-x^2)^2}\right)' = \frac{2(1+3x^2)}{(1-x^2)^3}\,. \tag{3.43}$$

Using the equality of the leftmost and rightmost expressions in Equation (3.43) and the fact that $e^2 \leq e(2e-1)$ for all $e \in \mathbb{N}$, we obtain

$$\begin{aligned}\sqrt{\sum_{e=0}^{\infty} e^2 x^{2e-2}} &\leq \sqrt{\sum_{e>0} e(2e-1)x^{2e-2}} \overset{(3.43)}{=} \sqrt{\frac{1+3x^2}{(1-x^2)^3}} = \sqrt{\frac{1-x^2}{(1-x^2)^3} + \frac{4x^2}{(1-x^2)^3}} \\ &= \sqrt{\frac{1}{(1-x^2)^2} + \frac{4x^2}{(1-x^2)^3}} \leq \frac{1}{1-x^2} + \frac{2x}{(1-x^2)^{\frac{3}{2}}}\,.\end{aligned}$$

□

We may now prove that most of the roots in $(0,1)$ are concentrated in $(1-\varepsilon,1)$ for $\varepsilon > 0$ arbitrary. We prove the statement that the number of roots in $(0,1-\varepsilon)$ can be bounded by a constant depending only on ε. Combined with Theorem 3.14, it then follows that most of the roots must be concentrated close to $x = 1$. We have

Theorem 3.16 (Concentration of roots of random k-sparse polynomials around $x = 1$)**.** *Let f be a random k-sparse polynomial following $\mathcal{N}(0,1)$ with support vector S. Also let $\varepsilon > 0$ be fixed. We then have*

$$Z_{(0,1-\varepsilon)}(f, \mathcal{N}(0,1)) \leq \frac{1}{2\pi}\left(\log\left(\frac{2}{\varepsilon}\right) + \frac{4}{\sqrt{\varepsilon}} - 4\right)\,.$$

Proof. As was the case before, we assume that the trailing term of S is the constant term, i.e., $e_0 = 0$, therefore we have

$$\|S(x)\| = \sqrt{\sum_{i=0}^{k-1} x^{2e_i}} \geq \sqrt{x^{2e_0}} = 1\,,\ x \in \mathbb{R}\,. \tag{3.44}$$

Recall Equation (3.11) that was used in the proof of Lemma 3.2, which we may slightly rewrite by noting that $(S(x) \cdot S'(x))^2 \geq 0$ as follows

$$
\begin{aligned}
Z_{(0,1-\varepsilon)}\left(f,\mathcal{N}(0,1)\right) &= \frac{1}{\pi}\int_0^{1-\varepsilon}\frac{\sqrt{g_S(x)\left\|S'(x)\right\|^2-\left(S(x)\cdot S'(x)\right)^2}}{g_S(x)}\,dx \\
&\leq \frac{1}{\pi}\int_0^{1-\varepsilon}\frac{\sqrt{g_S(x)\left\|S'(x)\right\|^2}}{g_S(x)}\,dx = \frac{1}{\pi}\int_0^{1-\varepsilon}\frac{\left\|S'(x)\right\|}{\sqrt{g_S(x)}}\,dx \\
&= \frac{1}{\pi}\int_0^{1-\varepsilon}\frac{\left\|S'(x)\right\|}{\left\|S(x)\right\|}\,dx \overset{(3.44)}{\leq} \frac{1}{\pi}\int_0^{1-\varepsilon}\left\|S'(x)\right\|\,dx\,. \qquad (3.45)
\end{aligned}
$$

Now by using Proposition 3.15 we have the following inequality for $\|S'(x)\|$

$$
\left\|S'(x)\right\| = \sqrt{\sum_{i=0}^{k-1} e_i^2 x^{2e_i-2}} \leq \sqrt{\sum_{e=0}^{\infty} e^2 x^{2e-2}} \leq \frac{1}{1-x^2}+\frac{2x}{(1-x^2)^{\frac{3}{2}}}\,. \qquad (3.46)
$$

Therefore by substituting Equation (3.46) into Equation (3.45), we have

$$
\begin{aligned}
Z_{(0,1-\varepsilon)}\left(f,\mathcal{N}(0,1)\right) &\leq \frac{1}{\pi}\int_0^{1-\varepsilon}\left\|S'(x)\right\|\,dx \leq \frac{1}{\pi}\int_0^{1-\varepsilon}\frac{1}{1-x^2}+\frac{2x}{(1-x^2)^{\frac{3}{2}}}\,dx \\
&= \frac{1}{\pi}\left(\int_0^{1-\varepsilon}\frac{1}{1-x^2}\,dx+\int_0^{1-\varepsilon}\frac{2x}{(1-x^2)^{\frac{3}{2}}}\,dx\right)\,. \qquad (3.47)
\end{aligned}
$$

We now remark that the above integrands can be written as the following derivatives

$$
\frac{1}{2}\left(\log\left(\frac{1+x}{1-x}\right)\right)' = \frac{1}{2}\frac{\frac{2}{(1-x)^2}}{\frac{1+x}{1-x}} = \frac{1}{2}\frac{2}{(1-x)(1+x)} = \frac{1}{1-x^2} \qquad (3.48)
$$

$$
\left(\frac{2}{\sqrt{1-x^2}}\right)' = \frac{-2\frac{1}{2}(1-x^2)^{-\frac{1}{2}}(-2x)}{1-x^2} = \frac{2x}{(1-x^2)^{\frac{3}{2}}}\,. \qquad (3.49)
$$

We can therefore write Equation (3.47) using Equation (3.48) and Equation (3.49) as follows

$$
\begin{aligned}
Z_{(0,1-\varepsilon)}\left(f,\mathcal{N}(0,1)\right) &\leq \frac{1}{\pi}\left(\left[\frac{1}{2}\log\left(\frac{1+x}{1-x}\right)\right]_0^{1-\varepsilon}+\left[\frac{2}{\sqrt{1-x^2}}\right]_0^{1-\varepsilon}\right) \\
&= \frac{1}{\pi}\left(\frac{1}{2}\log\left(\frac{2-\varepsilon}{\varepsilon}\right)+\frac{2}{\sqrt{\varepsilon(2-\varepsilon)}}-2\right) \qquad (0<\varepsilon<1) \\
&\leq \frac{1}{2\pi}\left(\log\left(\frac{2}{\varepsilon}\right)+\frac{4}{\sqrt{\varepsilon}}-4\right)\,.
\end{aligned}
$$

□

From the above it is easy to generalize to the entirety of $\mathbb{R}$. By the symmetry given by Equation (2.22), we may obtain a similar result for the interval $(1+\frac{\varepsilon}{1-\varepsilon},\infty)$. Note that $\frac{\varepsilon}{1-\varepsilon}$ tends to 0 as ε does. We thus may conclude that other than $\frac{1}{\pi}\left(\log\left(\frac{2}{\varepsilon}\right)+\frac{4}{\sqrt{\varepsilon}}-4\right)$ many roots, all positive real roots of a random polynomial f following $\mathcal{N}(0,1)$ are concentrated in the interval $(1-\varepsilon,1+\frac{\varepsilon}{1-\varepsilon})$. Similarly, by utilizing the symmetry given by Equation (2.21), we conclude that a similar statement is true for the negative real zeros of such a polynomial, that is, they are concentrated in the interval $(-1-\frac{\varepsilon}{1-\varepsilon},-1+\varepsilon)$.

We also note that the above statement works for any $0<\varepsilon<1$, even when ε depends on properties of the polynomial such as its sparsity k or its degree d. When ε is a constant with no such dependencies, Theorem 3.16 is always of interest and in particular it is easy to see that we may make ε arbitrarily small, as long as it remains a constant, which implies by the above that also the number of real roots in $(0,1-\varepsilon)$ will remain constant. On the other hand, taking into account the upper bound given by Theorem 3.14 that the number of real roots of a random polynomial following $\mathcal{N}(0,1)$ is bounded by $\frac{8}{\pi}\sqrt{k-1}$, as given by Theorem 3.14, it is easy to see that any $\varepsilon\geq\frac{1}{\sqrt{k-1}}$ results in a trivial bound that is superseded by the general bound for the entirety of $\mathbb{R}$. Furthermore, note that for a fixed sparsity k, the degree d of a polynomial can be arbitrarily large. Therefore, if the constant depends on the degree d, it may be that the bound obtained is also superseded by the bound given for $(0,1)$. Given the above, we prefer to state the above result so that ε is an arbitrarily small constant that does not depend on the polynomial in any way.

Thus far we have proven an upper bound of $O(\sqrt{k})$ for the real roots of a random polynomial following $\mathcal{N}(0,1)$ and also that its roots concentrate around $|x|=1$. We now complete the picture at this level of detail by stating an asymptotically matching lower bound. As we've already stated, this lower bound does not apply to *all* random polynomials. Rather, we will specify a specific support vector or more precisely, a family of support vectors depending on k. We begin with a technical lemma that complements Theorem 3.16 by providing an upper bound for the roots in the interval $(1-\varepsilon,1)$ for $\varepsilon>0$. We have the following

Lemma 3.17. *Let f be a random k-sparse polynomial following $\mathcal{N}(0,1)$ with support vector $S=(x^{e_0},\ldots,x^{e_{k-1}})$. Also let $\varepsilon>0$ be fixed. We then can bound the number of real zeros of f in $(1-\varepsilon,1)$ by*

$$Z_{(1-\varepsilon,1)}\left(f,\mathcal{N}(0,1)\right)\leq\varepsilon\frac{e_{k-1}}{\pi}\sqrt{k}\,.$$

Proof. As in the proof of Theorem 3.16 we use Equation (3.11) and the observation that $\left(S(x)\cdot S'(x)\right)^2\geq 0$ to derive an expression similar to Equation (3.45), with only the limits of the integral changing, which does not affect the calculation.

$$
\begin{aligned}
Z_{(1-\varepsilon,1)}\left(f,\mathcal{N}(0,1)\right) &= \frac{1}{\pi}\int_{1-\varepsilon}^{1}\frac{\sqrt{g_S(x)\left\|S'(x)\right\|^2-\left(S(x)\cdot S'(x)\right)^2}}{g_S(x)}\,dx \\
&\leq \frac{1}{\pi}\int_{1-\varepsilon}^{1}\frac{\sqrt{g_S(x)\left\|S'(x)\right\|^2}}{g_S(x)}\,dx = \frac{1}{\pi}\int_{1-\varepsilon}^{1}\frac{\left\|S'(x)\right\|}{\sqrt{g_S(x)}}\,dx \\
&= \frac{1}{\pi}\int_{1-\varepsilon}^{1}\frac{\left\|S'(x)\right\|}{\left\|S(x)\right\|}\,dx \overset{(3.44)}{\leq} \frac{1}{\pi}\int_{1-\varepsilon}^{1}\left\|S'(x)\right\|\,dx\,. \qquad (3.50)
\end{aligned}
$$

We now may use Equation (3.50) to derive a much simpler bound, by additionally observing that $e_{k-1} > e_{k-2} > \cdots > e_0 = 0$ and that $x > 0$ since $x \in (1-\varepsilon, 1)$. We have that

$$
\begin{aligned}
Z_{(1-\varepsilon,1)}\left(f,\mathcal{N}(0,1)\right) &\leq \frac{1}{\pi}\int_{1-\varepsilon}^{1}\left\|S'(x)\right\|\,dx = \frac{1}{\pi}\int_{1-\varepsilon}^{1}\sqrt{\sum_{i=0}^{k-1}e_i^2x^{2e_i-2}}\,dx \\
&\leq \frac{1}{\pi}\int_{1-\varepsilon}^{1}\sqrt{\sum_{i=0}^{k-1}e_{k-1}^2 1^{2e_i-2}}\,dx = \frac{1}{\pi}\int_{1-\varepsilon}^{1}\sqrt{ke_{k-1}^2}\,dx \\
&= \frac{e_{k-1}}{\pi}\sqrt{k}\int_{1-\varepsilon}^{1}1\,dx = \varepsilon\frac{e_{k-1}}{\pi}\sqrt{k}\,.
\end{aligned}
$$

□

We remark that this is a very weak upper bound on itself, since it depends on e_{k-1} which is the largest exponent appearing in f and therefore equal to its degree, $d = e_{k-1}$. Thus, while this only bounds the zeros in a very small interval, for ε constant it provides a bound even worse than the one given by Lemma 2.4 in terms of the degree. However, note that the bound also depends on ε and the smaller it becomes, so does the bound. For example for roughly $\varepsilon = e_{k-1}^{-1}$ it is already asymptotically equal to the bound given by Theorem 3.14, albeit for a much smaller interval. Therefore, as was the case with Theorem 3.16, we may choose ε accordingly to achieve our purpose. While in that case we chose ε to be constant, here it will be in our best interest to choose an ε that depends on e_{k-1} to remove that dependency from the bound, as well as allowing the interval $(1-\varepsilon, 1)$ to be as small as we require.

As a side note, by setting $\varepsilon \leq \frac{1}{e_{k-1}\sqrt{k}}$, we obtain a constant bound for the real roots of f in $(1-\varepsilon, 1)$. Combined with Theorem 3.16 for some $\varepsilon = c > 0$ an infinitesimally small constant, we can conclude that almost all the real roots with $0 \leq x \leq 1$ of a random, k-sparse, degree d polynomial following $\mathcal{N}(0,1)$ lie in $(1-c, 1-\frac{1}{d\sqrt{k}})$. By the usual symmetries given in Equation (2.22) and Equation (2.21), we conclude that for the entirety of $\mathbb{R}$, the real roots of such a polynomial lie accordingly in the intervals

$(1+\frac{1}{d\sqrt{k-1}}, 1+\frac{1}{1-c})$ for $x > 1$, $(-1+\frac{1}{d\sqrt{k}}, -1+c)$ for $-1 < x < 0$ and $(-1-\frac{1}{1-c}, -1-\frac{1}{d\sqrt{k-1}})$ for $x \leq -1$.

As we've already mentioned, the lower bound applies to a specific family of support vectors depending on k. These support vectors are defined for $k \geq 3$ as follows:

$$S = \left(x^0, x^1, x^{2^{2^1}}, \ldots, x^{2^{2^i}}, \ldots, x^{2^{2^{k-2}}}\right) = \left(1, x, x^4, \ldots, x^{2^{2^i}}, \ldots, x^{2^{2^{k-2}}}\right),\ 1 \leq i \leq k-2. \tag{3.51}$$

We will assume the support vector belongs to the above family from this point on and for the remainder of the proof of the lower bound. Our proof strategy is simple in light of our previous fundamental results. We will express the support vector (or more precisely, its set of exponents) as the disjoint union of two other support vectors. In particular, the second vector will be almost trivial consisting of a single term, which will correspond to the largest exponent of S. That is, we will obtain the support vector $S_1 = (x^{2^{2^{k-2}}})$ and consequently, the other support vector will be $S_2 = \left(1, x, x^4, \ldots, x^{2^{2^i}}, \ldots, x^{2^{2^{k-3}}}\right)$, $1 \leq i \leq k-3$. It is clear from this construction that S_1 and S_2 are disjoint as sets and furthermore $S = S_1 \uplus S_2$. We let f_k be a random polynomial with support vector S following $\mathcal{N}(0,1)$ and respectively g with support vector S_1 and f_{k-1} with support vector S_2 are also random polynomials following the same distribution. Recall from the proof of Lemma 3.10 and in particular the first three lines of Equation (3.28) that for disjoint support vectors as above, we may write

$$\begin{aligned} Z_{(0,1)}\left(f, \mathcal{N}(0,1)\right) &= \frac{1}{2\pi} \int_0^1 \sqrt{\mathcal{I}(g_{S_1 \uplus S_2}(x))}\, dx \\ &= \frac{1}{2\pi} \int_0^1 \sqrt{\frac{g_{S_1}}{g_{S_1}+g_{S_2}} \mathcal{I}(g_{S_1}) + \frac{g_{S_2}}{g_{S_1}+g_{S_2}} \mathcal{I}(g_{S_2}) + \frac{1}{g_{S_1} g_{S_2}} \left(\frac{g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}}{g_{S_1}+g_{S_2}}\right)^2}\, dx\,. \end{aligned} \tag{3.52}$$

The above Equation (3.52) hints on how to proceed further. Recall from the proof of Lemma 3.11 and in particular Equation (3.33) that for support vectors such as S_1 having a single term we have that $\mathcal{I}(g_{S_1}) = 0$. Furthermore, S_2 is a support vector of the family given in Equation (3.51) with sparsity strictly smaller than that of S. It is clear that we may utilize this to form the basis of a recursive relation up to a base case. To obtain a lower bound, we would also require to bound the contribution by the last term of in the above equation, that we would obtain at every step of the recursion. The below important lemma does precisely that, albeit for slight different integration bounds, for reasons that will become clear in the actual proof of the lower bound.

Lemma 3.18. *Let g be a random polynomial following $\mathcal{N}(0,1)$ with support vector $S_1 = (x^{2^{2^{k-2}}})$ and f_{k-1} be a random polynomial following $\mathcal{N}(0,1)$ with support vector $S_2 = \left(1, x, x^4, \ldots, x^{2^{2^i}}, \ldots, x^{2^{2^{k-3}}}\right)$, $1 \leq i \leq k-3$ and $k \geq 3$. Also let $\alpha = 2^{2^{k-2}}$. We then have*

$$\int_{1-\frac{1}{2\alpha}}^{1} \frac{1}{\sqrt{g_{S_1} g_{S_2}}} \left| \frac{g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}}{g_{S_1} + g_{S_2}} \right| dx \geq 2 \arctan\left(\frac{\sqrt{e} - \sqrt{2}}{1 + \sqrt{2e}} \frac{1}{\sqrt{k-2}} \right)$$
$$\geq 2 \arctan\left(\frac{0.07}{\sqrt{k-2}} \right).$$

Proof. We begin by noting that under the above assumptions, we have that $\left(\frac{g_{S_1}}{g_{S_2}}\right)' > 0$, similar to the proof of Lemma 3.13. We only need to verify that it holds for $k \geq 3$. Indeed it is

$$\left(\frac{g_{S_1}}{g_{S_2}}\right)' = \frac{g'_{S_1} g_{S_2} - g'_{S_2} g_{S_1}}{g_{S_2}^2}, \tag{3.53}$$

thus it is sufficient to prove the numerator is positive. Recall that

$$g_{S_1} = \left(x^{2^{2^{k-2}}}\right)^2 = x^{2^{2^{k-2}+1}}, \tag{3.54}$$
$$g_{S_2} = 1 + x^2 + \sum_{i=1}^{k-3} x^{2^{2^i+1}}, \tag{3.55}$$
$$g'_{S_1} = 2^{2^{k-2}+1} x^{2^{2^{k-2}+1}-1}, \tag{3.56}$$
$$g'_{S_2} = 2x + \sum_{i=1}^{k-3} 2^{2^i+1} x^{2^{2^i+1}-1}. \tag{3.57}$$

We thus may write the numerator of Equation (3.53) as

$$\begin{aligned}
&g'_{S_1} g_{S_2} - g'_{S_2} g_{S_1} = \\
&2^{2^{k-2}+1} x^{2^{2^{k-2}+1}-1} \left(1 + x^2 + \sum_{i=1}^{k-3} x^{2^{2^i+1}}\right) - x^{2^{2^{k-2}+1}} \left(2x + \sum_{i=1}^{k-3} 2^{2^i+1} x^{2^{2^i+1}-1}\right) \\
&2^{2^{k-2}+1} x^{2^{2^{k-2}+1}-1} + 2^{2^{k-2}+1} x^{2^{2^{k-2}+1}+1} + \sum_{i=1}^{k-3} 2^{2^{k-2}+1} x^{2^{2^{k-2}+1}+2^{2^i+1}-1} \\
&- 2x^{2^{2^{k-2}+1}+1} - \sum_{i=1}^{k-3} 2^{2^i+1} x^{2^{2^{k-2}+1}+2^{2^i+1}-1} \\
&= \sum_{i=1}^{k-3} \left(2^{2^{k-2}+1} - 2^{2^i+1}\right) x^{2^{2^{k-2}+1}+2^{2^i+1}-1} + \left(2^{2^{k-2}+1} - 2\right) x^{2^{2^{k-2}+1}+1} \\
&\quad + 2^{2^{k-2}+1} x^{2^{2^{k-2}+1}-1} > 0.
\end{aligned} \tag{3.58}$$

since $x > 0$, $2^{2^{k-2}+1} > 2^{2^i+1} > 2$ for $k \geq 3$, and $1 \leq i \leq k-3$. If $k = 3$, by standard convention we assume all the sums above to have no summands and thus be 0.

By Equation (3.58) and since in addition from Equation (3.54) and Equation (3.55) it is clear that $g_{S_1}+g_{S_2} > 0$, we may rewrite the integral as follows, similar to Equation (3.29)

$$\begin{aligned}
&\int_{1-\frac{1}{2\alpha}}^{1} \frac{1}{\sqrt{g_{S_1} g_{S_2}}} \left| \frac{g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}}{g_{S_1} + g_{S_2}} \right| dx = \int_{1-\frac{1}{2\alpha}}^{1} \frac{1}{\sqrt{g_{S_1} g_{S_2}}} \frac{g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}}{g_{S_1} + g_{S_2}} \, dx \\
&= \int_{1-\frac{1}{2\alpha}}^{1} \frac{g_{S_2}}{\sqrt{g_{S_1} g_{S_2}}} \frac{g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}}{g_{S_2}^2} \frac{g_{S_2}}{g_{S_1} + g_{S_2}} \, dx = \int_{1-\frac{1}{2\alpha}}^{1} \frac{1}{\sqrt{\frac{g_{S_1}}{g_{S_2}}}} \left(\frac{g_{S_1}}{g_{S_2}} \right)' \frac{1}{\left(\sqrt{\frac{g_{S_1}}{g_{S_2}}} \right)^2 + 1} \, dx \,.
\end{aligned} \tag{3.59}$$

Now by setting $u(x) = \sqrt{\frac{g_{S_1}}{g_{S_2}}}$, we have

$$du = \frac{1}{2} \frac{1}{\sqrt{\frac{g_{S_1}}{g_{S_2}}}} \left(\frac{g_{S_1}}{g_{S_2}} \right)' dx \,. \tag{3.60}$$

Using this change of variables and according to Equation (3.60), we may write Equation (3.59) as

$$\begin{aligned}
&\int_{1-\frac{1}{2\alpha}}^{1} \frac{1}{\sqrt{\frac{g_{S_1}}{g_{S_2}}}} \left(\frac{g_{S_1}}{g_{S_2}} \right)' \frac{1}{\left(\sqrt{\frac{g_{S_1}}{g_{S_2}}} \right)^2 + 1} \, dx = 2 \int_{u\left(1-\frac{1}{2\alpha}\right)}^{u(1)} \frac{1}{u^2+1} \, du \\
&= 2 \left(\arctan\left(u(1)\right) - \arctan\left(u\left(1 - \frac{1}{2\alpha} \right) \right) \right) \\
&= 2 \left(\arctan\left(\sqrt{\frac{g_{S_1}(1)}{g_{S_2}(1)}} \right) - \arctan\left(\sqrt{\frac{g_{S_1}\left(1-\frac{1}{2\alpha}\right)}{g_{S_2}\left(1-\frac{1}{2\alpha}\right)}} \right) \right) .
\end{aligned} \tag{3.61}$$

The argument of the first arctangent above is easy to compute, since by Equation (3.54) and Equation (3.55) we have

$$\begin{aligned}
g_{S_1}(1) &= 1^{2^{2^{k-2}}} = 1 \,, \\
g_{S_2}(1) &= 1 + 1^2 + \sum_{i=1}^{k-3} 1^{2^{2^i+1}} = k - 1 \,,
\end{aligned}$$

therefore it is easy to see that

$$\arctan\left(\sqrt{\frac{g_{S_1}(1)}{g_{S_2}(1)}} \right) = \arctan\left(\frac{1}{\sqrt{k-1}} \right) . \tag{3.62}$$

The arguments of the second arctangent are much harder to compute precisely, however we can bound them without much effort. Note that since the arctangent function is strictly increasing and a minus sign precedes it, an upper bound for its argument is

needed, which means an upper bound on the numerator and a lower bound on the denominator. Towards this goal, we require the following two inequalities

$$\left(1-\frac{1}{n}\right)^n \leq \frac{1}{e}\,,\; n \in \mathbb{N}^+\,, \tag{3.63}$$

$$(1+x)^r \geq 1+xr\,,\; r \in \mathbb{N}\,,\; x > -1\,. \tag{3.64}$$

Equation (3.63) is well-known and straightforward to prove. As a function of n it is strictly increasing as can be seen by its derivative

$$\left(\left(1-\frac{1}{n}\right)^n\right)' = n\left(1-\frac{1}{n}\right)^{n-1}\frac{1}{n^2} = \left(1-\frac{1}{n}\right)^{n-1}\frac{1}{n} > 0\,,$$

for $n \in \mathbb{N}^+$. Thus the largest value of the quantity is attained as n tends to infinity and by one of the definitions of e, we have that

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n = \frac{1}{e}\,.$$

The inequality now follows immediately from the monotonicity and the above limit.

The other inequality, Equation (3.64), can also be proven easily by induction on r. For the basis step of $r = 0$, we have that

$$(1+x)^0 = 1 \geq 1+0x\,.$$

Now we may prove the result inductively assuming it is true for $r-1$. We have that

$$\begin{aligned}
&(1+x)^{r-1} \geq 1+(r-1)x \iff (1+x)^{r-1} \geq 1+rx-x\\
&\overset{x>-1}{\iff} (1+x)^r \geq (1+rx-x)(1+x) \iff (1+x)^r \geq 1+x+rx+rx^2-x-x^2\\
&(1+x)^r \geq (r-1)x^2+rx+1 \overset{r\geq 1}{\geq} 1+rx\,.
\end{aligned}$$

We now are ready to prove the bounds for the argument of the second arctangent. The numerator is the easier of the two, since it involves a single term. Specifically using Equation (3.54) and Equation (3.63) we have

$$g_{S_1}\left(1-\frac{1}{2\alpha}\right) = \left(1-\frac{1}{2\,2^{2^{k-2}}}\right)^{2^{2^{k-2}+1}} = \left(1-\frac{1}{2^{2^{k-2}+1}}\right)^{2^{2^{k-2}+1}} \leq \frac{1}{e}\,. \tag{3.65}$$

Similarly using Equation (3.55) and Equation (3.64), since $2^{2^i} > 4 > 1$ we obtain for the denominator

$$g_{S_2}\left(1-\frac{1}{2\alpha}\right) = 1 + \left(1-\frac{1}{2\alpha}\right)^2 + \sum_{i=1}^{k-3}\left(1-\frac{1}{2\alpha}\right)^{2^{2^i+1}} \qquad \text{(Using Equation (3.64) twice)}$$
$$\geq 1 + 1 - \frac{2}{2a} + \sum_{i=1}^{k-3} 1 - \sum_{i=1}^{k-3}\frac{2^{2^i+1}}{2^{2^{k-2}+1}} \geq k - 1 - \frac{1}{a} - \sum_{i=1}^{k-3}\frac{2^{2^{k-3}+1}}{2^{2^{k-2}+1}}$$
$$= k - 1 - \frac{1}{a} - \frac{k-3}{2^{2^{k-3}}}\,. \tag{3.66}$$

For simplicity's sake it suffices if we transform the above Equation (3.66) to a bound depending only on k. A short examination of the two quantities that are not whole numbers in this inequality reveals that they can easily by lower bounded by $-\frac{1}{2}$ each, thus giving the closest possible lower bound that is a whole number. Specifically, since $a = 2^{2^{k-2}}$ for $k \geq 3$, we have that

$$a > 2^2 \iff \frac{1}{a} < \frac{1}{4} \iff -\frac{1}{a} > -\frac{1}{4} > -\frac{1}{2}\,. \tag{3.67}$$

For the other quantity we also want that

$$-\frac{k-3}{2^{2^{k-3}}} > -\frac{1}{2} \iff 2^{2^{k-3}-1} - k + 3 > 0\,. \tag{3.68}$$

Let $l(k) = 2^{2^{k-3}-1} - k + 3$, we will analyze the behavior of that function to obtain the bound. Recall that we care only for integer values of k and in particular $k \geq 3$. It is clear that $l(3) = 2^{2^0-1} - 3 + 3 = 1 > 0$ and $l(4) = 2^{2^1-1} - 4 + 3 = 1 > 0$. The derivative of the above function is

$$l'(k) = 2^{2^{k-3}-1}\ln(2)\left(2^{k-3}-1\right)' - 1 = 2^{2^{k-3}+k-4}\ln^2(2) - 1\,. \tag{3.69}$$

It is easy to see that this quantity should be strictly increasing as a composition of monotone functions. Indeed, by taking the derivative once more we have

$$l''(k) = 2^{2^{k-3}+k-4}\ln^3(2)\left(2^{k-3}+k-4\right)' = 2^{2^{k-3}+2k-7}\ln^4(2) + 2^{2^{k-3}+k-4}\ln^3(2) > 0\,. \tag{3.70}$$

All quantities in Equation (3.70) are positive for any k and in particular for $k \geq 3$. While the first derivative given in Equation (3.69) is increasing, we have that $l'(3) = 2^{1+3-4}\ln^2(2) - 1 = \ln^2(2) - 1 < 0$, since $\ln(2) < 1$. However, we have that $l'(4) = 2^{2+4-4}\ln^2(2) - 1 = 4\ln^2(2) - 1 = (2\ln(2))^2 - 1 = \ln^2(4) - 1 > 0$ since $\ln(4) > 1$. Since l' is strictly increasing, this implies that $l' > 0$ for $k \geq 4$. In turn, this means $l(k)$ is strictly increasing for $k \geq 4$ and since $l(4) = l(3) = 1 > 0$, it follows that the bound in Equation (3.68) holds for $k \geq 3$, since we only care for integer values.

Now by Equation (3.67) and Equation (3.68), we may write Equation (3.66) as

$$g_{S_2}\left(1-\frac{1}{2\alpha}\right) \geq k - 1 - \frac{1}{a} - \frac{k-3}{2^{2^{k-3}}} \geq k - 1 - \frac{1}{4} - \frac{1}{2} > k - 2\,. \tag{3.71}$$

We may now finally bound the argument of the second arctangent in Equation (3.61) using Equation (3.65) and Equation (3.71). We have

$$\arctan\left(\sqrt{\frac{g_{S_1}\left(1-\frac{1}{2\alpha}\right)}{g_{S_2}\left(1-\frac{1}{2\alpha}\right)}}\right) \leq \arctan\left(\sqrt{\frac{\frac{1}{e}}{k-2}}\right) = \arctan\left(\frac{1}{\sqrt{e(k-2)}}\right). \tag{3.72}$$

Having determined the argument of the first arctangent and bounded those of the second, we may now proceed with the remainder of the proof. The following well-known arctangent subtraction formula, that in turn can be proven using the equivalent subtraction formulas for sine and cosine, will be useful

$$\arctan(u) - \arctan(v) = \arctan\left(\frac{u-v}{1+uv}\right). \tag{3.73}$$

Now using Equation (3.62), Equation (3.72) and Equation (3.73) we can write Equation (3.61) as

$$\begin{aligned}
&2\left(\arctan\left(\sqrt{\frac{g_{S_1}(1)}{g_{S_2}(1)}}\right) - \arctan\left(\sqrt{\frac{g_{S_1}\left(1-\frac{1}{2\alpha}\right)}{g_{S_2}\left(1-\frac{1}{2\alpha}\right)}}\right)\right) \\
&\geq 2\left(\arctan\left(\frac{1}{\sqrt{k-1}}\right) - \arctan\left(\frac{1}{\sqrt{e(k-2)}}\right)\right) = 2\arctan\left(\frac{\frac{1}{\sqrt{k-1}} - \frac{1}{\sqrt{e(k-2)}}}{1+\frac{1}{\sqrt{e(k-1)(k-2)}}}\right) \\
&= 2\arctan\left(\frac{\sqrt{e}\sqrt{k-2}-\sqrt{k-1}}{\sqrt{e}\sqrt{k-1}\sqrt{k-2}+1}\right).
\end{aligned} \tag{3.74}$$

We now only need to simplify the argument of the single arctangent in the expression through a few basic inequalities. We have that

$$\sqrt{k-2} \geq \frac{\sqrt{k-1}}{\sqrt{2}} \iff k-2 \geq \frac{k-1}{2} \iff k \geq 3\,, \tag{3.75}$$

which by the equivalence clearly holds for $k \geq 3$. Similarly, we also have

$$\begin{aligned}
&\sqrt{e}\sqrt{k-1}\sqrt{k-2}+1 \leq \left(1+\frac{1}{\sqrt{2e}}\right)\sqrt{e}\sqrt{k-1}\sqrt{k-2} \iff \frac{\sqrt{k-1}\sqrt{k-2}}{\sqrt{2}} \geq 1 \\
&\iff \sqrt{k-1}\sqrt{k-2}-\sqrt{2} \geq 0\,.
\end{aligned} \tag{3.76}$$

Note that in the last inequality, the left hand side is a strictly increasing function for $k \geq 2$ and already for $k = 3$ we have equality since $\sqrt{2}\sqrt{1}-\sqrt{2} = 0$. Thus this inequality and its equivalent we are interested in holds for all $k \geq 3$.

Using Equation (3.75) and Equation (3.76) we can now write Equation (3.74) as

$$2\arctan\left(\frac{\sqrt{e}\sqrt{k-2}-\sqrt{k-1}}{\sqrt{e}\sqrt{k-1}\sqrt{k-2}+1}\right) \geq 2\arctan\left(\frac{\left(\frac{\sqrt{e}}{\sqrt{2}}-1\right)\sqrt{k-1}}{\left(1+\frac{1}{\sqrt{2e}}\right)\sqrt{e}\sqrt{k-1}\sqrt{k-2}}\right)$$
$$= 2\arctan\left(\frac{\sqrt{\frac{e}{2}}-1}{\sqrt{e}\left(1+\frac{1}{\sqrt{2e}}\right)}\frac{1}{\sqrt{k-2}}\right)$$
$$= 2\arctan\left(\frac{\sqrt{e}-\sqrt{2}}{1+\sqrt{2e}}\frac{1}{\sqrt{k-2}}\right) \geq 2\arctan\left(\frac{0.07}{\sqrt{k-2}}\right),$$

which is precisely what was claimed. □

We can now proceed with the proof of the lower bound, which consists of two steps. As we've already mentioned, the proof idea is to use Equation (3.52) as the basis for a recursive relation. In the first step, we will derive this recursive inequality, then proceed to solve it. We now state the lower bound and present its proof.

Theorem 3.19. *Let f_k be a random k-sparse polynomial following $\mathcal{N}(0,1)$ and with support vector $S(x) = \left(1, x, x^{2^2}, \ldots, x^{2^{2^i}}, \ldots, x^{2^{2^{k-2}}}\right)$, $1 \leq i \leq k-2$, for $k \geq 3$. Then its number of real roots in $(0,1)$, and by extension $\mathbb{R}$, is at least*

$$Z_{\mathbb{R}}\left(f_k, \mathcal{N}(0,1)\right) \geq Z_{(0,1)}\left(f_k, \mathcal{N}(0,1)\right) \geq 0.014\sqrt{k} - \frac{0.016}{\sqrt{k}} \geq 0.014\sqrt{k} - 0.007$$

for $k \geq 6$ and greater than 0.018 for $3 \leq k \leq 5$.

Proof. In line with the setup we have performed so far, we consider S as the disjoint union of two separate vectors $S = S_1 \uplus S_2$, with $S_1 = \left(x^{2^{2^{k-2}}}\right)$ and $S_2 = \left(1, x, x^{2^2}, \ldots, x^{2^{2^i}}, \ldots, x^{2^{2^{k-3}}}\right)$, $1 \leq i \leq k-3$ for the first step of the recursion. At subsequent steps the recursion will continue by considering the term involving the highest order exponent separately, until we reach the base case of $k = 2$ that has already been solved exactly in Lemma 3.12.

In particular, S_1 in each recursive step is a single term support vector, thus by Equation (3.33) we have that $\mathcal{I}(g_{S_1}) = 0$. We may thus write

$$Z_{(0,1)}\left(f, \mathcal{N}(0,1)\right) = \frac{1}{2\pi}\int_0^1 \sqrt{\mathcal{I}(g_{S_1 \uplus S_2}(x))}\, dx$$
$$= \frac{1}{2\pi}\int_0^1 \sqrt{\frac{g_{S_1}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_1}) + \frac{g_{S_2}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_2}) + \frac{1}{g_{S_1}g_{S_2}}\left(\frac{g_{S_2}g'_{S_1} - g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\right)^2}\, dx$$
$$= \frac{1}{2\pi}\int_0^1 \sqrt{\frac{g_{S_2}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_2}) + \frac{1}{g_{S_1}g_{S_2}}\left(\frac{g_{S_2}g'_{S_1} - g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\right)^2}\, dx\,. \tag{3.77}$$

A careful examination of Equation (3.77) is necessary to proceed. In particular, let us examine the two summands of the square root in the integrand. Both of them are positive, since g_{S_1} and g_{S_2} are, as well as $\mathcal{I}(g_{S_2})$, as we've already shown in Proposition 3.3. Thus, we can omit either one and obtain a lower bound to the expression. That alone however is not sufficient, since if we omit the first summand we will be foregoing the advantage of a recursive relation, while if we omit the second, the bound obtained will be too weak and essentially equivalent to bounding by the base case.

Instead, a further ingredient is required for our proof strategy. Specifically, we split the interval of integration as follows: $(0,1) = (0, 1-\frac{1}{2\alpha}) \uplus (1-\frac{1}{2\alpha}, 1)$ for $\alpha = 2^{2^{k-2}}$ as above. The idea is that we may obtain a lower bound by considering only the second summand for $(1-\frac{1}{2\alpha}, 1)$, which by Lemma 3.18 is lower bounded by a sufficiently large amount that also depends on k, while the remaining interval does not contain 1, thus it should only contain a small number of roots and can be used to form the recursion using the first summand without too much loss in precision.

A second trick in our arsenal is that while the above provides a recursive relation, the interval of integration would be $(0, 1-\frac{1}{2\alpha})$ for all subsequent steps, which we would expect to provide weak lower bounds since by design it does not contain 1. The trick is to rewrite this integral as a difference of an integral over $(0,1)$ and one over $(1-\frac{1}{2\alpha}, 1)$ and then utilize Lemma 3.17 to upper bound the second integral or equivalently, lower bound its negation. While the contribution by this bound is rather small, the benefit is that we may maintain the same interval of integration, $(0,1)$, for every recursive step.

By putting all the above ideas into action, we thus may write Equation (3.77) as

$$
\begin{aligned}
&\frac{1}{2\pi}\int_0^1 \sqrt{\frac{g_{S_2}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_2}) + \frac{1}{g_{S_1}g_{S_2}}\left(\frac{g_{S_2}g'_{S_1} - g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\right)^2}\, dx \\
&= \frac{1}{2\pi}\left(\int_0^{1-\frac{1}{2\alpha}} \sqrt{\frac{g_{S_2}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_2}) + \frac{1}{g_{S_1}g_{S_2}}\left(\frac{g_{S_2}g'_{S_1} - g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\right)^2}\, dx\right. \\
&\left. + \int_{1-\frac{1}{2\alpha}}^1 \sqrt{\frac{g_{S_2}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_2}) + \frac{1}{g_{S_1}g_{S_2}}\left(\frac{g_{S_2}g'_{S_1} - g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\right)^2}\, dx\right) \qquad (g_{S_1}, g_{S_2}, \mathcal{I}(g_{S_2}) \geq 0) \\
&\geq \frac{1}{2\pi}\left(\int_0^{1-\frac{1}{2\alpha}} \sqrt{\frac{g_{S_2}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_2})}\, dx + \int_{1-\frac{1}{2\alpha}}^1 \sqrt{\frac{1}{g_{S_1}g_{S_2}}\left(\frac{g_{S_2}g'_{S_1} - g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\right)^2}\, dx\right) \\
&= \frac{1}{2\pi}\left(\int_0^{1-\frac{1}{2\alpha}} \sqrt{\frac{g_{S_2}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_2})}\, dx + \int_{1-\frac{1}{2\alpha}}^1 \frac{1}{\sqrt{g_{S_1}g_{S_2}}}\left|\frac{g_{S_2}g'_{S_1} - g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\right|\, dx\right)
\end{aligned}
$$

$$= \frac{1}{2\pi} \int_0^1 \sqrt{\frac{g_{S_2}}{g_{S_1}+g_{S_2}} \mathcal{I}(g_{S_2})}\, dx - \frac{1}{2\pi} \int_{1-\frac{1}{2\alpha}}^1 \sqrt{\frac{g_{S_2}}{g_{S_1}+g_{S_2}} \mathcal{I}(g_{S_2})}\, dx$$
$$+ \frac{1}{2\pi} \int_{1-\frac{1}{2\alpha}}^1 \frac{1}{\sqrt{g_{S_1} g_{S_2}}} \left| \frac{g_{S_2} g'_{S_1} - g_{S_1} g'_{S_2}}{g_{S_1}+g_{S_2}} \right| dx\,. \tag{3.78}$$

Note that since $g_{S_2} \geq 1 > 0$ for $x \in \mathbb{R}$ the first factor in the first and second integral can be written as

$$\sqrt{\frac{g_{S_2}}{g_{S_1}+g_{S_2}}} = \frac{1}{\sqrt{1+\frac{g_{S_1}}{g_{S_2}}}}\,. \tag{3.79}$$

We may simplify all integrals involved in Equation (3.78) by taking into consideration that $\left(\frac{g_{S_1}}{g_{S_2}}\right)' > 0$ as already shown in Equation (3.58). For the third integral this means that we can simplify the absolute value, as in Lemma 3.18, while for the two other integrals we may bound the term in Equation (3.79) since

$$\frac{g_{S_1}(0)}{g_{S_2}(0)} < \frac{g_{S_1}(x)}{g_{S_2}(x)} < \frac{g_{S_1}(1)}{g_{S_2}(1)} \iff 0 < \frac{g_{S_1}(x)}{g_{S_2}(x)} < \frac{1}{k-1} \iff 1 < 1 + \frac{g_{S_1}(x)}{g_{S_2}(x)} < 1 + \frac{1}{k-1}$$
$$\iff 1 > \frac{1}{1+\frac{g_{S_1}(x)}{g_{S_2}(x)}} > \frac{k-1}{k}\,. \tag{3.80}$$

By the right inequality of Equation (3.80) and considering that the square root is a strictly increasing function, we directly obtain that

$$\frac{1}{\sqrt{1+\frac{g_{S_1}(x)}{g_{S_2}(x)}}} > \sqrt{\frac{k-1}{k}}\,, \tag{3.81}$$

while from the left inequality we similarly obtain

$$\frac{1}{\sqrt{1+\frac{g_{S_1}(x)}{g_{S_2}(x)}}} < 1 \iff -\frac{1}{\sqrt{1+\frac{g_{S_1}(x)}{g_{S_2}(x)}}} > -1\,. \tag{3.82}$$

We may thus rewrite Equation (3.78) using Equation (3.81),Equation (3.82) and simplifying the absolute value as

$$\begin{aligned}
&\frac{1}{2\pi}\int_0^1 \sqrt{\frac{g_{S_2}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_2})}\,dx - \frac{1}{2\pi}\int_{1-\frac{1}{2\alpha}}^1 \sqrt{\frac{g_{S_2}}{g_{S_1}+g_{S_2}}\mathcal{I}(g_{S_2})}\,dx \\
&\quad + \frac{1}{2\pi}\int_{1-\frac{1}{2\alpha}}^1 \frac{1}{\sqrt{g_{S_1}g_{S_2}}}\left|\frac{g_{S_2}g'_{S_1}-g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\right| dx \\
&> \frac{1}{2\pi}\sqrt{\frac{k-1}{k}}\int_0^1 \sqrt{\mathcal{I}(g_{S_2})}\,dx - \frac{1}{2\pi}\int_{1-\frac{1}{2\alpha}}^1 \sqrt{\mathcal{I}(g_{S_2})}\,dx \\
&\quad + \frac{1}{2\pi}\int_{1-\frac{1}{2\alpha}}^1 \frac{1}{\sqrt{g_{S_1}g_{S_2}}}\frac{g_{S_2}g'_{S_1}-g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\,dx\,. \qquad (3.83)
\end{aligned}$$

The first two integrals are precisely the formulation of root counting given by Lemma 3.2, thus we may use that substitution. The first of these integrals will remain as is and be used for the recursive relation. The second integral can be bounded using Lemma 3.17, while the last integral that is essentially the contribution per step of recursion can also be bounded by Lemma 3.18. This lemma was stated precisely for this case, thus no parametrization is needed.

For the second integral, using Lemma 3.17 for S_2, we have that $k \coloneqq k-1, e_{k-1} \coloneqq 2^{2^{k-3}}$ and we choose $\varepsilon = \frac{1}{2\alpha} = \frac{1}{2^{2^{k-2}+1}}$. Consistently with previous lemmas used, we call f_{k-1} the random polynomial following $\mathcal{N}(0,1)$ with support vector S_2. We thus obtain

$$Z_{(1-\varepsilon,1)}\left(f_{k-1},\mathcal{N}(0,1)\right) \leq \frac{1}{2^{2^{k-2}+1}}\frac{2^{2^{k-3}}}{\pi}\sqrt{k-1} = \frac{\sqrt{k-1}}{\pi 2^{2^{k-3}+1}}\,. \qquad (3.84)$$

Substituting Lemma 3.18 and Equation (3.84) into Equation (3.83), we thus obtain

$$\begin{aligned}
&\frac{1}{2\pi}\sqrt{\frac{k-1}{k}}\int_0^1 \sqrt{\mathcal{I}(g_{S_2})}\,dx - \frac{1}{2\pi}\int_{1-\frac{1}{2\alpha}}^1 \sqrt{\mathcal{I}(g_{S_2})}\,dx \\
&\quad + \frac{1}{2\pi}\int_{1-\frac{1}{2\alpha}}^1 \frac{1}{\sqrt{g_{S_1}g_{S_2}}}\frac{g_{S_2}g'_{S_1}-g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\,dx \\
&\geq \sqrt{\frac{k-1}{k}}Z_{(0,1)}\left(f_{k-1},\mathcal{N}(0,1)\right) - Z_{(1-\frac{1}{2\alpha},1)}\left(f_{k-1},\mathcal{N}(0,1)\right) \\
&\quad + \frac{1}{2\pi}\int_{1-\frac{1}{2\alpha}}^1 \frac{1}{\sqrt{g_{S_1}g_{S_2}}}\frac{g_{S_2}g'_{S_1}-g_{S_1}g'_{S_2}}{g_{S_1}+g_{S_2}}\,dx \\
&\geq \sqrt{\frac{k-1}{k}}Z_{(0,1)}\left(f_{k-1},\mathcal{N}(0,1)\right) - \frac{\sqrt{k-1}}{\pi 2^{2^{k-3}+1}} + \frac{1}{\pi}\arctan\left(\frac{\sqrt{e}-\sqrt{2}}{1+\sqrt{2e}}\frac{1}{\sqrt{k-2}}\right). \qquad (3.85)
\end{aligned}$$

The first term of the above expression will form the basis of our recursion. The right hand term will be the contribution per step, which with careful manipulation we may write as a somewhat more easy to handle expression. For that purpose we will require the following inequality, which is also tight.

Proposition 3.20. $\arctan(x) \geq \frac{\pi}{4}x\,,\ 0 \leq x \leq 1\,.$

Proof. Let $h(x) = \arctan(x) - \frac{\pi}{4}x$. It is easy to see that its derivative is

$$h'(x) = \frac{1}{1+x^2} - \frac{\pi}{4}\,.$$

We have that

$$h'(x) = 0 \iff \frac{1}{1+x^2} = \frac{\pi}{4} \iff x^2 = \frac{4}{\pi} - 1 \iff x = \pm\sqrt{\frac{4}{\pi} - 1} \approx \pm 0.52\,.$$

It is easy to see from the above that $h'(x) > 0$ for $|x| < \sqrt{\frac{4}{\pi} - 1}$. Thus for our interval of interest $[0, 1]$, we have that h is increasing up to this maximum point and then it is decreasing. Furthermore, we have that $h(0) = \arctan(0) - \frac{\pi}{4}\ 0 = 0$ and $h(1) = \arctan(1) - \frac{\pi}{4} = \frac{\pi}{4} - \frac{\pi}{4} = 0$. Combined with the monotonicity of the function, we have that $h(x) \geq 0$ for $0 \leq x \leq 1$ and the stated claim now follows.

□

Note that the argument of the arctangent in our case is a decreasing always positive function that has a value of approximately $0.07 < 1$ for $k = 3$. Thus we can now use Proposition 3.20 to write Equation (3.85) as follows

$$\begin{aligned}
&\sqrt{\frac{k-1}{k}} Z_{(0,1)}\left(f_{k-1}, \mathcal{N}(0,1)\right) - \frac{\sqrt{k-1}}{\pi 2^{2^{k-3}+1}} + \frac{1}{\pi}\arctan\left(\frac{\sqrt{e}-\sqrt{2}}{1+\sqrt{2e}}\frac{1}{\sqrt{k-2}}\right) \\
&\geq \sqrt{\frac{k-1}{k}} Z_{(0,1)}\left(f_{k-1}, \mathcal{N}(0,1)\right) - \frac{\sqrt{k-1}}{\pi 2^{2^{k-3}+1}} + \frac{1}{\pi}\frac{\pi}{4}\frac{\sqrt{e}-\sqrt{2}}{1+\sqrt{2e}}\frac{1}{\sqrt{k-2}} \\
&= \sqrt{\frac{k-1}{k}} Z_{(0,1)}\left(f_{k-1}, \mathcal{N}(0,1)\right) - \frac{\sqrt{k-1}}{\pi 2^{2^{k-3}+1}} + \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)}\frac{1}{\sqrt{k-2}} \\
&= \sqrt{\frac{k-1}{k}} Z_{(0,1)}\left(f_{k-1}, \mathcal{N}(0,1)\right) + \frac{1}{\sqrt{k-2}}\left(\frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{\sqrt{k-1}\sqrt{k-2}}{\pi 2^{2^{k-3}+1}}\right) \\
&= \sqrt{\frac{k-1}{k}} Z_{(0,1)}\left(f_{k-1}, \mathcal{N}(0,1)\right) + \frac{c(k)}{\sqrt{k-2}}\,, \qquad (3.86)
\end{aligned}$$

where

$$c(k) = \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{\sqrt{k-1}\sqrt{k-2}}{\pi 2^{2^{k-3}+1}}\,. \qquad (3.87)$$

The recursive inequality describing how the number of real roots relates to k has already been essentially given in Equation (3.86), which we can state more clearly as

$$Z_{(0,1)}\left(f_k, \mathcal{N}(0,1)\right) > \sqrt{\frac{k-1}{k}} Z_{(0,1)}\left(f_{k-1}, \mathcal{N}(0,1)\right) + \frac{c(k)}{\sqrt{k-2}}\,. \tag{3.88}$$

Recall that f_k has support vector $S = \left(1, x, x^{2^2}, \ldots, x^{2^{2^i}}, \ldots, x^{2^{2^{k-2}}}\right)$, $1 \leq i \leq k-2$ and sparsity k, while f_{k-1} has support vector $S_2 = \left(1, x, x^{2^2}, \ldots, x^{2^{2^i}}, \ldots, x^{2^{2^{k-3}}}\right)$, $1 \leq i \leq k-3$ and thus sparsity $k-1$. We will now solve this recursive inequality, which requires two steps. One is figuring out the general formula after j steps of recursion have been performed. The second is figuring out a bound on the value of $c(k)$ which will essentially be bounded by a constant.

Let us begin with the formula after j steps of recursion. Keeping with the notation above, let f_j be a random polynomial with sparsity j following $\mathcal{N}(0,1)$ with support vector $S_{k-j+1} = \left(1, x, x^{2^2}, \ldots, x^{2^{2^i}}, \ldots, x^{2^{2^{j-2}}}\right)$, $1 \leq i \leq j-2$ for $3 \leq j \leq k-2$. In addition, for $j = 2$, we have the special case of f_2 with support vector $S_{k-1} = (1, x)$ which will be the base case of the recursion. Obtaining one more step of the recursion should be sufficient to make the pattern apparent. To lower bound the average number of roots for sparsity $k-1$, or equivalently, $j = k-2$, we have that

$$Z_{(0,1)}\left(f_{k-1}, \mathcal{N}(0,1)\right) > \sqrt{\frac{k-2}{k-1}} Z_{(0,1)}\left(f_{k-2}, \mathcal{N}(0,1)\right) + \frac{c(k-1)}{\sqrt{k-3}}\,. \tag{3.89}$$

Substituting the above Equation (3.89) into Equation (3.88) we have that

$$\begin{aligned} Z_{(0,1)}\left(f_k, \mathcal{N}(0,1)\right) &> \sqrt{\frac{k-1}{k}} \left(\sqrt{\frac{k-2}{k-1}} Z_{(0,1)}\left(f_{k-2}, \mathcal{N}(0,1)\right) + \frac{c(k-1)}{\sqrt{k-3}} \right) + \frac{c(k)}{\sqrt{k-2}} \\ &= \sqrt{\frac{k-2}{k}} Z_{(0,1)}\left(f_{k-2}, \mathcal{N}(0,1)\right) + \frac{1}{\sqrt{k}} \frac{\sqrt{k-1}}{\sqrt{k-3}} c(k-1) + \frac{1}{\sqrt{k}} \frac{\sqrt{k}}{\sqrt{k-2}} c(k)\,. \end{aligned} \tag{3.90}$$

It is clear from the above, as well perhaps by further expansion, that the general expression for the $(k-j)$-th recursion step, $2 \leq j \leq k-1$ is

$$Z_{(0,1)}\left(f_k, \mathcal{N}(0,1)\right) > \sqrt{\frac{j}{k}} Z_{(0,1)}\left(f_j, \mathcal{N}(0,1)\right) + \frac{1}{\sqrt{k}} \sum_{i=j}^{k-1} \left(\sqrt{\frac{i+1}{i-1}}\, c(i+1) \right)\,. \tag{3.91}$$

Since for $j = 2$, f_2 has sparsity 2, we know from Lemma 3.12 that this random polynomial has precisely $1/4$ real roots on average. We could thus substitute this value and solve the above inequality precisely, giving an upper bound. The only obstacle to this is determining a value for the sum expression, which we now proceed to do.

The easier of the two expressions involved in finding a lower bound for $d(i) = \sqrt{\frac{i+1}{i-1}} = \sqrt{1 + \frac{2}{i-1}}$. A simple calculation of its derivative reveals that this function is strictly decreasing where it is defined and certainly for $i \geq 2$ since

$$d'(i) = -\frac{1}{\sqrt{i+1}(i-1)^{3/2}} < 0\,.$$

Since the function $d(i)$ is decreasing, we obtain by taking the limit at infinity that

$$d(i) = \sqrt{\frac{i+1}{i-1}} = \sqrt{1+\frac{2}{i-1}} > \lim_{i\to\infty}\sqrt{1+\frac{2}{i-1}} = 1\,. \tag{3.92}$$

This leaves the already defined $c(k)$, which we mentioned will be lower bounded by a constant, except for a few special cases. A closer observation of Equation (3.87) reveals that the first part of expression is truly constant and in fact except for the $1/4$ factor was already calculated in the proof of Lemma 3.18. We have that

$$\frac{1}{4}\frac{\sqrt{e}-\sqrt{2}}{\left(1+\sqrt{2e}\right)} > \frac{0.07}{4} > 0.017\,, \tag{3.93}$$

where the last number is also the best possible approximation from below of the quantity with three digits of precision.

The second part of $c(k)$ is a function of k that clearly will be increasing eventually since $\sqrt{k} \in o\left(2^{2^k}\right)$ and we are interested in the negation of such a ratio. It only remains to determine for which k this occurs and what is the appropriate constant to obtain a lower bound.

In particular we will study the following function

$$l(k) = \frac{\sqrt{k-1}\sqrt{k-2}}{\pi 2^{2^{k-3}+1}}\,. \tag{3.94}$$

We will once again compute its derivative to investigative its monotonicity behavior. We have that

$$\begin{aligned}
l'(k) &= \left(\frac{\sqrt{k-1}\sqrt{k-2}}{\pi 2^{2^{k-3}+1}}\right)' = \frac{\left(\sqrt{k-1}\sqrt{k-2}\right)'\pi 2^{2^{k-3}+1} - \sqrt{k-1}\sqrt{k-2}\left(\pi 2^{2^{k-3}+1}\right)'}{\pi^2 2^{2^{k-2}+2}} \\
&= \frac{\frac{2k-3}{2\sqrt{k-1}\sqrt{k-2}}\pi 2^{2^{k-3}+1} - \sqrt{k-1}\sqrt{k-2}\,\pi\ln^2(2)2^{2^{k-3}+k-2}}{\pi^2 2^{2^{k-2}+2}} \\
&= \frac{(2k-3)2^{2^{k-3}} - \ln^2(2)(k-1)(k-2)2^{2^{k-3}+k-2}}{\pi\sqrt{k-1}\sqrt{k-2}\;2^{2^{k-2}+2}} \\
&= \frac{2k-3-(k-1)(k-2)\ln^2(2)2^{k-2}}{\pi\sqrt{k-1}\sqrt{k-2}\;2^{2^{k-3}+2}}\,.
\end{aligned} \tag{3.95}$$

Clearly the denominator of the derivative is positive for $k \geq 3$, so it suffices to analyze the behavior of the numerator. Let $n(k) = 2k-3-(k-1)(k-2)\ln^2(2)2^{k-2}$. We have that

$$\begin{aligned}
n'(k) &= 2-(2k-3)\ln^2(2)2^{k-2} - (k-1)(k-2)\ln^3(2)2^{k-2} \\
&= 2-\ln^2(2)2^{k-2}\left(2k-3+\ln(2)(k-1)(k-2)\right)\,.
\end{aligned} \tag{3.96}$$

This expression is clearly now decreasing for $k \geq 3$ since the expression being negated consists of strictly increasing positive functions. Already for $k = 3$, we can see that

$n'(3) = 2\left(1 - ln^2(2)\left(3 + 2\ln(2)\right)\right) < 0$, therefore $n'(k) < 0$ for $k \geq 3$. In turn this implies that $n(k)$ is decreasing for $k \geq 3$. We have that $n(3) = 3 - 4\ln^2(2) > 0$, however $n(4) = 5 - 24\ln^2(2) < 0$. Thus for $k \geq 4$ we have that $n(k) < 0$ and since this is the numerator of the derivative of $l(k)$, it follows that $l(k)$ is decreasing for $k \geq 4$.

Since $c(k) = \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - l(k)$, with the first part of the expression being a constant, it follows that $c(k)$ will have the opposite monotonicity of $l(k)$ and thus will be increasing for $k \geq 4$. Since we only care for integer values, ideally we would like $c(3)$ and $c(4)$ to be positive, which would guarantee all the other values would also be and thus provide a useful lower bound. Unfortunately, that turns out not to be the case, but rather we only obtain a positive value for $k \geq 6$. Specifically, it is

$$
\begin{aligned}
c(3) &= \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{\sqrt{2}}{\pi 2^{1+1}} = \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{\sqrt{2}}{4\pi} \approx -0.10\,,\\
c(4) &= \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{\sqrt{3}\sqrt{2}}{\pi 2^{2+1}} = \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{\sqrt{6}}{8\pi} \approx -0.08\,,\\
c(5) &= \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{\sqrt{4}\sqrt{3}}{\pi 2^{2^2+1}} = \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{\sqrt{3}}{16\pi} \approx -0.02\,,\\
c(6) &= \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{\sqrt{5}\sqrt{4}}{2^{2^3+1}\pi} = \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{\sqrt{5}}{256\pi} \approx 0.014 > 0\,.
\end{aligned}
$$

All above approximations are lower bounds and accurate up to the digit included. Since $c(k)$ is an increasing function for $k \geq 4$, we can conclude from the above that

$$
c(k) \geq c(6) = \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{\sqrt{5}}{256\pi} > 0.014\,,\ k \geq 6\,. \tag{3.97}
$$

Before we state the lower bound for $k \geq 6$, we take care of the special cases for $k = 3, 4, 5$. Note that as we've already mentioned for $k = 2$ we may invoke Lemma 3.12 to determine the expected number of real roots precisely. The same bound can be used for any k for the base case $j = 2$, that is

$$
Z_{(0,1)}\left(f_2, \mathcal{N}(0,1)\right) = \frac{1}{4}\,. \tag{3.98}
$$

Note that for $k \geq 3$, the recursion in Equation (3.91) is valid, which we may use to derive the values for the special cases. Also note that the bounds differ for a polynomial that originally has sparsity 3 and a step on the recursion that has sparsity 3 but derived for an initial polynomial of sparsity $k > 3$. To avoid confusion, we will use the notation $g_{i,j}$ for the polynomials involved in the special cases, where $g_{i,j}$ has initial sparsity i and the sparsity at the current recursion step is j. With that notation, we have

$$
\begin{aligned}
Z_{(0,1)}\left(g_{3,3},\mathcal{N}(0,1)\right) &\geq \sqrt{\frac{2}{3}} Z_{(0,1)}\left(g_{3,2},\mathcal{N}(0,1)\right) + c(3) \\
&= \frac{\sqrt{2}}{\sqrt{3}}\frac{1}{4} + \frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{\sqrt{2}}{4\pi} > 0.1\,, \qquad (3.99) \\
Z_{(0,1)}\left(g_{4,4},\mathcal{N}(0,1)\right) &\geq \sqrt{\frac{2}{4}} Z_{(0,1)}\left(g_{4,2},\mathcal{N}(0,1)\right) + \frac{1}{2}\left(\sqrt{3}c(3) + \sqrt{2}c(4)\right) \\
&= \frac{1}{\sqrt{2}}\frac{1}{4} + \frac{\sqrt{3}}{2}c(3) + \frac{1}{\sqrt{2}}c(4) > 0.033\,, \qquad (3.100) \\
Z_{(0,1)}\left(g_{5,4},\mathcal{N}(0,1)\right) &\geq \sqrt{\frac{2}{5}} Z_{(0,1)}\left(g_{5,2},\mathcal{N}(0,1)\right) + \frac{1}{\sqrt{5}}\left(\sqrt{3}c(3) + \sqrt{2}c(4) + \frac{\sqrt{5}}{\sqrt{3}}c(5)\right) \\
&= \sqrt{\frac{2}{5}}\frac{1}{4} + \frac{\sqrt{3}}{\sqrt{5}}c(3) + \frac{\sqrt{2}}{\sqrt{5}}c(4) + \frac{1}{\sqrt{3}}c(5) > 0.018\,. \qquad (3.101)
\end{aligned}
$$

We are now ready to prove the lower bound. Our starting point is Equation (3.91) for $j = 2$ and general $k \geq 6$. Precisely, we have that

$$
\begin{aligned}
&Z_{(0,1)}\left(f_k,\mathcal{N}(0,1)\right) > \sqrt{\frac{2}{k}} Z_{(0,1)}\left(f_2,\mathcal{N}(0,1)\right) + \frac{1}{\sqrt{k}}\sum_{i=2}^{k-1}\left(\sqrt{\frac{i+1}{i-1}}\, c(i+1)\right) \\
&= \frac{1}{2^{\frac{3}{2}}\sqrt{k}} + \frac{1}{\sqrt{k}}\sum_{i=2}^{4}\left(\sqrt{\frac{i+1}{i-1}}\, c(i+1)\right) + \frac{1}{\sqrt{k}}\sum_{i=5}^{k-1}\left(\sqrt{\frac{i+1}{i-1}}\, c(i+1)\right), \qquad (3.102)
\end{aligned}
$$

where we used Lemma 3.12 to substitute $Z_{(0,1)}\left(f_2,\mathcal{N}(0,1)\right) = \frac{1}{4}$. We now only need to perform some substitutions that we have already derived, namely using Equation (3.92) and Equation (3.97) to write Equation (3.102) as

$$
\begin{aligned}
&Z_{(0,1)}\left(f_k,\mathcal{N}(0,1)\right) > \frac{1}{2^{\frac{3}{2}}\sqrt{k}} + \frac{1}{\sqrt{k}}\left(\sqrt{3}c(3) + \sqrt{2}c(4) + \frac{\sqrt{5}}{\sqrt{3}}c(5)\right) + \frac{1}{\sqrt{k}}\sum_{i=5}^{k-1} c(i+1) \\
&\geq \frac{1}{\sqrt{k}}\sum_{i=5}^{k-1} c(6) + \frac{1}{2^{\frac{3}{2}}\sqrt{k}} \\
&\quad + \frac{1}{\sqrt{k}}\left(\left(\sqrt{3}+\sqrt{2}+\frac{\sqrt{5}}{\sqrt{3}}\right)\frac{\sqrt{e}-\sqrt{2}}{4\left(1+\sqrt{2e}\right)} - \frac{1}{2\pi}\left(\frac{\sqrt{3}}{\sqrt{2}} + \frac{\sqrt{3}}{2} + \frac{\sqrt{5}}{8}\right)\right) \\
&= \frac{(k-5)c(6)}{\sqrt{k}} + \frac{0.054}{\sqrt{k}} \geq 0.014\sqrt{k} - \frac{0.016}{\sqrt{k}} \geq 0.014\sqrt{k} - \frac{0.016}{\sqrt{6}} > 0.014\sqrt{k} - 0.007\,. \qquad (3.103)
\end{aligned}
$$

□

Having presented all related results for random polynomials following $\mathcal{N}(0,1)$, we now examine the results in total. Since there is at least a family of support vectors

that matches the upper bound given in Theorem 3.14 asymptotically, we conclude that neither bound can be improved by more than a constant multiplicative factor.

On the other hand, while the upper bound applies to all support vectors of size k, the same is not true for the lower bound. In fact, this is already evident from the classical results of Kac [45] that give a $\Theta(\log k)$ bound for the dense case since for the dense support we have that $k = d + 1$. While the upper bound limits the average number of real zeros to $O(\sqrt{k})$ for random polynomials following $\mathcal{N}(0, 1)$, this could potentially vary greatly for different support vectors.

We also have shown that for all supports, real zeros for such random polynomials following $\mathcal{N}(0, 1)$ will cluster in a small neighborhood around $|x| = 1$, which expands upon similar remarks by Kac on the dense case. While our investigation has answered many important questions regarding the real zeros of random sparse polynomials, especially those following the standard normal distribution, there are still many further questions that point to future avenues of research to explore that we consider in the concluding Chapter 4. Before that, we continue on the next Section 3.3 by expanding our remarks to arbitrary distributions and investigating whether our results presented in this section can be extended to other cases.

3.3 Expected number of zeros of sparse polynomials following arbitrary distributions

In Section 3.1 we analyzed in detail how Edelman and Kostlan in [28] rewrote the Kac integral by first considering the support curve $\gamma(x)$ with k components on the unit sphere S_{k-1} and then examining the points perpendicular to each point of the support curve. For each such point S, we called the points on the sphere perpendicular to it the equator $S_\perp$. Then, as the support curve $\gamma(x)$ varies, these equators sweep out a space that we call $\gamma(x)_\perp$. The total volume of the space swept out, potentially with overlapping, is equivalent to the expected number of zeros, with each point weighted with respect to the distribution measure chosen. For the standard normal case, the great advantage is that on the unit sphere, every point has the same measure, thus we only need to consider the volume itself that as was shown relates directly to the arclength of the support curve in the $(0, 1)$ interval.

In this section we generalize this approach to any absolutely continuous distribution. We restrict ourselves to this class of distributions since they are the class of distributions that have a probability density function, which is necessary for this approach. As we've discussed in Section 2.2, this excludes singular and discrete distributions, as well as mixtures containing them. While singular distributions are rather exotic and rarely encountered in practice, discrete distributions are an important class of distributions that is not addressed by this method. We first present this method in general, then we adapt some of the arguments used in the standard normal case so that they apply to as many distributions as possible.

It was already known to Edelman and Kostlan that their integral could be adapted for any absolutely continuous distribution, as they briefly mention in Section 5.1 of [28]. Here, we closely follow their work, although we examine the result in more detail. We work over $\mathbb{R}^k$ instead of the hypersphere, since we cannot rely on the property of the

standard normal distribution. We assume that $f(x)$ is a random polynomial with support vector $S(x) = (x^{e_0}, \ldots, x^{e_{k-1}})$. Also similar as above let $\gamma(x)$ be the normalized support curve, that is

$$\gamma(x) = \frac{S(x)}{\|S(x)\|} .$$

The coefficients of the polynomial f can similarly be described by a vector, namely the coefficient vector $a = (a_0, \ldots, a_{k-1})$ and are i.i.d. random variables following an absolutely continuous distribution with probability density function $\sigma(a_i)$ for $0 \leq i \leq k-1$. Thus the joint distribution for a is given by $\sigma(a) = \prod_{i=0}^{k-1} \sigma(a_i)$. The polynomial $f(x)$ can then by written as

$$f(x) = a \cdot S(x) = \sum_{i=0}^{k-1} a_i x^{e_i} . \tag{3.104}$$

Now let us fix a particular x, say $x = x_0$. We examine when the random polynomial is equal to 0 for this particular value. By substituting into Equation (3.104), we have

$$f(x_0) = 0 \iff \sum_{i=0}^{k-1} a_i x_0^{e_i} = 0 \iff a \cdot S(x_0) = 0 . \tag{3.105}$$

The vector $S(x_0)$ corresponds to the unique point on the support curve for which $x = x_0$. All coefficients vector a in $\mathbb{R}^k$ that satisfy this equation correspond to coefficients of polynomials for which x_0 is a root. Furthermore, by the inner product given in the last equation in Equation (3.105), we see that this equation is satisfied by all coefficient vectors $a \in \mathbb{R}^k$ that are perpendicular to $S(x_0)$. Note that since the normalized support vector $\gamma(x_0)$ is colinear to $S(x_0)$, the same holds true. So far the setup is almost identical to the one we established in Section 3.1 for the hypersphere. Now we identify the space perpendicular to the vector $\gamma(x_0)$, which is the first point where the two approaches differ. Since this space consists of all vectors in $\mathbb{R}^k$ perpendicular to $S(x_0)$, it is clear that it forms a subspace of dimension $k-1$ that is defined by the equation $a \cdot \gamma(x_0) = 0$. In fact, this space is isomorphic to R^{k-1}. For a few easy to visualize examples, in $\mathbb{R}^2$ all vectors perpendicular to some fixed vector form a line that is isomorphic to $\mathbb{R}$ and similarly all vectors perpendicular to a fixed vector in $\mathbb{R}^3$ form a plane. To show that this is the case for any $k \in \mathbb{N}^+$ is not much harder: Consider a basis $B = (b_0, \ldots, b_{k-1})$ for $\mathbb{R}^k$ such that $b_0 = \gamma(x_0)$. Now consider any vector U with coefficients $U = (u_0, \ldots, u_{k-1})$ with respect to the basis B such that $u_0 = 0$. Since $\gamma(x_0) = b_0 = (1, 0, \ldots, 0)$ with respect to this basis, it follows that $U \cdot \gamma(x_0) = 0$ with $u_1, \ldots, u_{k-1}$ arbitrary over $\mathbb{R}$, thus clearly they span a subspace isomorphic to $\mathbb{R}^{k-1}$.

With the picture clear for a fixed support vector, we now have to consider how the hyperplane of vectors perpendicular to the support vector changes as the support vector varies along the normalized support curve $\gamma(x)$. It is beneficial to consider an infinitesimal change along the curve, from the normalized support vector $\gamma(x_0)$ to $\gamma(x_0 + dx)$. As the support vector changes, so will the hyperplane perpendicular to it. On the previous case of the unit sphere, that space was bounded by two equators. In $\mathbb{R}^{k-1}$, this area will be bounded by two hyperplanes. We call that area a "wedge". For an easy to visualize example that also motivates the name given, in $\mathbb{R}^2$ this area will be the area contained

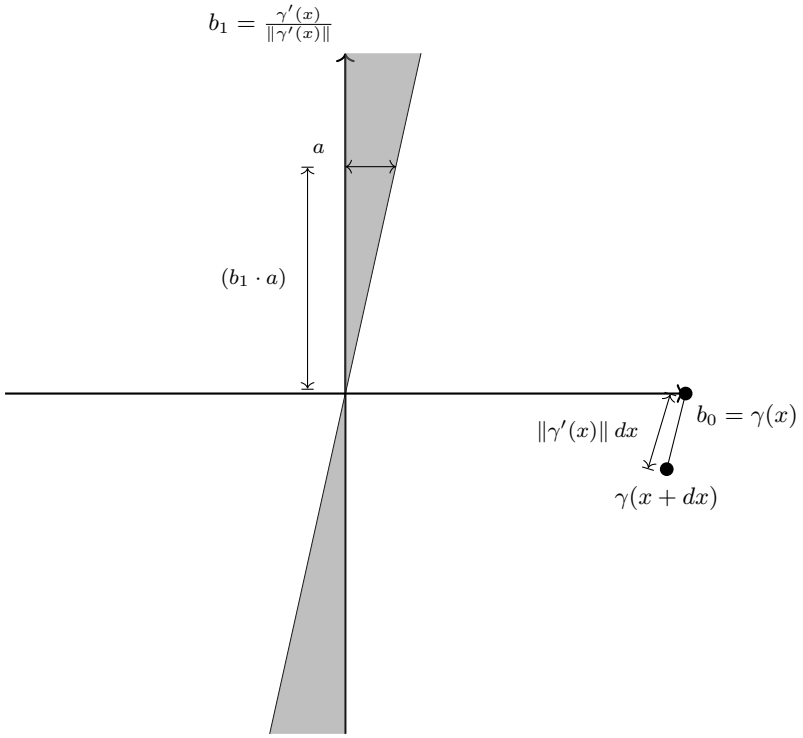

Figure 3.1: Geometric picture of the wedge area bounded by two hyperplanes

between two lines, see Figure 3.1. A similar figure appears as Figure 6 in Section 5.1 of [28], here we have made slight alterations to make the picture easier to understand.

Similar to above, it is easier to work over a different base than the standard one for $\mathbb{R}^{k-1}$, say $B = (b_0, \ldots, b_{k-1})$. Since the space we wish to investigate is perpendicular to $\gamma(x)$ for each x, it makes sense to have $b_0 = \gamma(x)$. Since this is a unit vector, we can construct the entire base as being orthonormal, meaning that the basis vectors are both pairwise perpendicular and unit vectors. In fact, it is not necessary to specify every basis vector, only two of them. We have already specified b_0, we now specify also b_1. We choose a vector that will both play a significant role in the calculation as well as be perpendicular to $\gamma(x)$ for all x. We thus set $b_1 = \frac{\gamma'(x)}{\|\gamma'(x)\|}$. Note that since $\gamma(x)$ is a unit vector, we have

$$b_0(x) \cdot b_1(x) = \gamma(x) \cdot \frac{\gamma'(x)}{\|\gamma'(x)\|} = \frac{1}{2\|\gamma'(x)\|} \frac{\partial}{dx} \Big(\gamma(x) \cdot \gamma(x) \Big) = \frac{1}{2\|\gamma'(x)\|} \frac{\partial}{dx} (1) = 0 \,,$$

therefore the two vectors are perpendicular. We now are ready to compute the volume of the wedge area that contains all polynomials with x as root. We begin by projecting onto the plane spanned by b_0 and b_1 and examining the area of the wedge in this space. Figure 3.1 clarifies the situation greatly. Since all vectors A that correspond to polynomial that have x as a root are perpendicular to $\gamma(x)$, when projected they are colinear with b_1, i.e. they appear on the y-axis in Figure 3.1. As $\gamma(x)$ moves to $\gamma(x+dx)$, the perpendicular space moves as well. In the plane $\mathsf{span}(b_0, b_1)$ the wedge is contained between the y axis

and the line perpendicular to $\gamma(x + dx)$ and is colored gray in Figure 3.1. Let us first consider the rate at which the plane is swept by the plane perpendicular to $\gamma(x)$. Since the line that sweeps the plane is always perpendicular to $\gamma(x)$, it will sweep the space at the same rate as $\gamma(x)$ moves throughout the space. That rate is equivalent to the arclength traversed, that is $\|\gamma'(x)\|\, dx$, as can be seen near the x-axis in Figure 3.1.

Next we need to measure the contribution of each $a \in \mathbb{R}^k$ that is perpendicular to $\gamma(x)$. As we've already mentioned, these vectors will be colinear with b_1. Their length along the line defined by b_1 is by definition equal to $|a \cdot b_1|$. Furthermore the contribution for every a must be weighted according to the density function $\sigma(a)$. To consider all a, we compute the Lebesgue integral for all a in the $\mathsf{span}(b_0, b_1)$ space, which is isomorphic to $\mathbb{R}$, that are also perpendicular to $\gamma(x)$, that is they satisfy $a \cdot \gamma(x) = 0$. This integral is with respect to db, the standard Lebesgue measure for that space. We thus have that the wedge area in Figure 3.1 is equal to

$$\|\gamma'(x)\|\, dx \int\limits_{\substack{\mathsf{span}(b_0,b_1),\\ \{a\cdot\gamma(x)=0\}}} |a \cdot b_1|\sigma(a)\, db\,. \tag{3.106}$$

Having computed the wedge area in $\mathsf{span}(b_0, b_1)$, we can now easily compute the volume of the entire wedge in $\mathbb{R}^k$. The wedge space can be obtained from the 2-dimensional wedge by taking the Cartesian product of the 2-dimensional wedge with $\mathbb{R}^{k-2}$, the space spanned by the remaining $k - 2$ basis vectors, since the wedge in the $\mathsf{span}(b_0, b_1)$ was a projection of the wedge in $\mathbb{R}^k$. Let db^{k-1} be the Lebesgue measure for the hyperplane defined by $a \cdot \gamma(x) = 0$. We then have that the volume of the wedge is

$$\|\gamma'(x)\|\, dx \int\limits_{\{a\cdot\gamma(x)=0\}} |a \cdot b_1|\sigma(a)\, db^{k-1}\,. \tag{3.107}$$

The quantity computed in Equation (3.107) is the density of real zeros for a single x. To compute the expected number of zeros we only need to integrate over all $x \in \mathbb{R}$. We thus obtain

$$\int\limits_{\mathbb{R}} \|\gamma'(x)\| \int\limits_{\{a\cdot\gamma(x)=0\}} |a \cdot b_1|\sigma(a)\, db^{k-1}\, dx = \int\limits_{\mathbb{R}} \int\limits_{\{a\cdot\gamma(x)=0\}} |a \cdot \gamma'(x)|\sigma(a)\, db^{k-1}\, dx\,, \tag{3.108}$$

where we substituted $b_1 = \frac{\gamma'(x)}{\|\gamma'(x)\|}$. We summarize the above into the following Lemma.

Lemma 3.21 (Expected number of zeros, arbitrary density (Theorem 5.1 in [28]))**.** *Let $f(x) = a \cdot S(x)$ be a random polynomial with $S(x)$ its support vector and $\gamma(x)$ its normalized support curve. The coefficient vector of $f(x)$ given by $a = (a_0, \cdots, a_{k-1})$ has i.i.d. random variables as components, each distributed according to some absolutely continuous distribution with density $\sigma(a_i)$ for $0 \le i \le k - 1$. The joint density function for a is then given by $\sigma(a) = \prod_{i=0}^{k-1} \sigma(a_i)$. Then the expected number of real zeros of $f(x)$ with respect to $\sigma(a)$ is*

$$\int\limits_{\mathbb{R}} \int\limits_{\{a\cdot\gamma(x)=0\}} |a \cdot \gamma'(x)|\sigma(a)\, db^{k-1}\, dx\,,$$

where db^{k-1} is the standard Lebesgue measure in the subspace perpendicular to $\gamma(x)$.

With an expression for the expected number of real zeros for any absolutely continuous distribution, we may now prove a few basic results that make solving the integral potentially easier. These results are generalizations of techniques already used in the standard normal case. We make the conditions for these results to hold as general as possible, so that they apply to as many distributions as possible. From the statements of these results, it is clear that these conditions are necessary. We state two such results, both aiming at reducing the domain of integration with respect to the variable x. The first of these results is the most general and perhaps the most useful. Not only it reduces the domain of integration to a finite interval, namely $(-1, 1)$, it also only requires that the random polynomial whose zeros we compute to simply be k-sparse. Further restrictions, especially with respect to the support vectors allowed, must be carefully selected, since the result requires that for each support vector $S(x)$, a related support vector $W(x)$ is also in the family of polynomials considered. The condition on the distribution of the coefficients of the polynomial is rather generous, it requires that they are independent but they need not be identically distributed. In fact, it is sufficient that specific pairs of coefficients follow the same distribution. Although this is the most general condition that still allows us to prove the result, the more restrictive condition of demanding that the coefficients are both independent and identically distributed is more likely to be used in practice. We now state this result, followed by its proof.

Theorem 3.22 (Reducing domain of integration to $(-1, 1)$)**.** *Let $f(x) = a \cdot S(x)$ be a random k-sparse polynomial with support vector $S(x) = (1, x^{e_1}, \ldots, x^{e_{k-1}})$, degree $d = e_{k-1}$, and coefficient vector $a = (a_0, \ldots, a_{k-1})$ such that the coefficients a_i are* independent *random variables following absolutely continuous distributions with density $\sigma(a_i)$ for $0 \leq i \leq k-1$, such that $\sigma(a_i) = \sigma(a_{k-1-i})$ for $0 \leq i \leq \lfloor \frac{k-1}{2} \rfloor$. Thus, the joint distribution $\mathcal{D}$ for the coefficient vector a has density $\sigma(a) = \prod_{i=0}^{k-1} \sigma(a_i)$.*

Also let $g(x) = a^R \cdot W(x)$ be a random k-sparse polynomial with support vector $W(x) = (1, x^{d-e_{k-2}}, \ldots, x^{d-e_{k-1-i}}, \ldots, x^d)$, degree $d = e_{k-1}$ and coefficient vector $a^R = (a_{k-1}, \ldots, a_{k-1-i}, \ldots, a_0)$ for $0 \leq i \leq k-1$. That is, the exponents of the support vector $W(x)$ are obtained from those of $S(x)$ by subtracting each exponent from the highest one, $d = e_{k-1}$, and its coefficient vector a^R is obtained by the coefficient vector a by "reversing" its order of coefficients from right to left, that is the first coefficient of a is the last of a^R and vice versa.

We then have that the expected number of real zeros of $f(x)$ over $\mathbb{R}$ with respect to the distribution $\mathcal{D}$ is equal to the sum of the expected number of real zeros of $f(x)$ and $g(x)$ over $(-1, 1)$, that is

$$Z_{\mathbb{R}}\left(f(x), \mathcal{D}\right) = Z_{(-1,1)}\left(f(x), \mathcal{D}\right) + Z_{(-1,1)}\left(g(x), \mathcal{D}\right) .$$

Proof. Let $\gamma_S(x)$ be the normalized support curve corresponding to $S(x)$ and similarly $\gamma_W(x)$ be the normalized support curve corresponding to $W(x)$. By Lemma 3.21 we can express $Z_{\mathbb{R}}\left(f(x), \mathcal{D}\right)$ by an integral, which we then decompose into three integrals with appropriately chosen domains of integration. We have

$$\int_{\mathbb{R}} \int_{\{a\cdot\gamma_S(x)=0\}} |a\cdot\gamma_S'(x)|\sigma(a)\, db^{k-1}\, dx = \int_{-1}^{1} \int_{\{a\cdot\gamma_S(x)=0\}} |a\cdot\gamma_S'(x)|\sigma(a)\, db^{k-1}\, dx + \int_{-\infty}^{-1} \int_{\{a\cdot\gamma_S(x)=0\}} |a\cdot\gamma_S'(x)|\sigma(a)\, db^{k-1}\, dx + \int_{1}^{\infty} \int_{\{a\cdot\gamma_S(x)=0\}} |a\cdot\gamma_S'(x)|\sigma(a)\, db^{k-1}\, dx\,. \tag{3.109}$$

The first integral is precisely $Z_{(-1,1)}\left(f(x), \mathcal{D}\right)$, which is in the final expression we wish to prove, we only need to rewrite the other two integrals. A quick look at the chosen intervals reveals the intent, which is to perform a change of variable $u \coloneqq \frac{1}{x}$. Then the sum of the last two integrals will be an integral over $(-1, 1)$ as well. To perform this variable change, we need to carefully rewrite the two integrals. We do so by examining their separate components and establishing some key equations. We begin with the random polynomials themselves, $f(x) = a \cdot S(x) = \sum_{i=0}^{k-1} a_i x^{e_i}$ and $g(x) = a^R \cdot W(x) = \sum_{i=0}^{k-1} a_{k-1-i} x^{d-e_{k-1-i}}$, for which we have

$$\begin{aligned} f(x) = a\cdot S(x) = \sum_{i=0}^{k-1} a_i x^{e_i} = \sum_{i=0}^{k-1} a_i \left(\frac{1}{x}\right)^{-e_i} = \left(\frac{1}{x}\right)^{-d} \sum_{i=0}^{k-1} a_i \left(\frac{1}{x}\right)^{d-e_i} \\ \text{(set } j \coloneqq k-1-i\text{)} \\ = x^d \sum_{j=0}^{k-1} a_{k-1-j} \left(\frac{1}{x}\right)^{d-e_{k-1-j}} = x^d g\left(\frac{1}{x}\right)\,. \end{aligned} \tag{3.110}$$

Next we establish a relation between the magnitudes of the two vectors. We first derive a related helpful equation. We have that

$$\begin{aligned} S(x)\cdot S(x) = \sum_{i=0}^{k-1} x^{2e_i} = \sum_{i=0}^{k-1} \left(\frac{1}{x}\right)^{-2e_i} = \left(\frac{1}{x}\right)^{-2d} \sum_{i=0}^{k-1} \left(\frac{1}{x}\right)^{2(d-e_i)} \quad \text{(set } j \coloneqq k-1-i\text{)} \\ = x^{2d} \sum_{j=0}^{k-1} \left(\frac{1}{x}\right)^{2(d-e_{k-1-j})} = x^{2d}\left(W\left(\frac{1}{x}\right)\cdot W\left(\frac{1}{x}\right)\right)\,. \end{aligned} \tag{3.111}$$

From Equation (3.111) we can directly establish the following relation for the magnitudes of the vectors

$$\|S(x)\| = x^d \left\| W\left(\frac{1}{x}\right)\right\|\,. \tag{3.112}$$

We can now show the following useful equation for the defining equation of the hyperplane containing vectors perpendicular to $\gamma_S(x)$. We have that

$$\begin{aligned} a \cdot \gamma_S(x) &= \frac{a \cdot S(x)}{\|S(x)\|} = \frac{f(x)}{\|S(x)\|} && \text{(using eq. (3.110), eq. (3.112))} \\ &= \frac{x^d g\left(\frac{1}{x}\right)}{x^d \left\|W\left(\frac{1}{x}\right)\right\|} = \frac{g\left(\frac{1}{x}\right)}{\left\|W\left(\frac{1}{x}\right)\right\|} = \frac{a^R \cdot W\left(\frac{1}{x}\right)}{\left\|W\left(\frac{1}{x}\right)\right\|} = a^R \cdot \gamma_W\left(\frac{1}{x}\right). && (3.113) \end{aligned}$$

By differentiating Equation (3.113), we obtain the following:

$$\frac{\partial}{\partial x}\left(a \cdot \gamma_S(x)\right) = \frac{\partial}{\partial x}\left(a^R \cdot \gamma_W\left(\frac{1}{x}\right)\right) \iff a \cdot \gamma_S'(x) = a^R \cdot \gamma_W'\left(\frac{1}{x}\right)\frac{\partial}{\partial x}\left(\frac{1}{x}\right). \quad (3.114)$$

Finally, the following holds for the joint densities of a and a^R:

$$\begin{aligned} \sigma(a) &= \prod_{i=0}^{k-1} \sigma(a_i) && \text{(using } \sigma(a_i) = \sigma(a_{k-1-i})\text{)} \\ &= \prod_{i=0}^{k-1} \sigma(a_{k-1-i}) = \sigma(a^R). && (3.115) \end{aligned}$$

We are now ready to write the second and third integral in Equation (3.109) in a more suitable form. We first carry out the calculations for the second integral, using Equation (3.113), Equation (3.114) and Equation (3.115). We have that

$$\begin{aligned} &\int_{-\infty}^{-1} \int_{\{a\cdot\gamma_S(x)=0\}} |a \cdot \gamma_S'(x)| \sigma(a)\, db^{k-1}\, dx \\ &= \int_{-\infty}^{-1} \int_{\{a^R\cdot\gamma_W\left(\frac{1}{x}\right)=0\}} \left|a^R \cdot \gamma_W'\left(\frac{1}{x}\right)\right| \left|\frac{\partial}{\partial x}\left(\frac{1}{x}\right)\right| \sigma(a^R)\, db^{k-1}\, dx && \left(\tfrac{\partial}{\partial x}\left(\tfrac{1}{x}\right) \le 0\right) \\ &= \int_{-\infty}^{-1} \int_{\{a^R\cdot\gamma_W\left(\frac{1}{x}\right)=0\}} \left|a^R \cdot \gamma_W'\left(\frac{1}{x}\right)\right| \left(-\frac{\partial}{\partial x}\left(\frac{1}{x}\right)\right) \sigma(a^R)\, db^{k-1}\, dx \\ &&& \text{(reverse limits of outer integral)} \\ &= \int_{-1}^{-\infty} \int_{\{a^R\cdot\gamma_W\left(\frac{1}{x}\right)=0\}} \left|a^R \cdot \gamma_W'\left(\frac{1}{x}\right)\right| \sigma(a^R)\, db^{k-1} \frac{\partial}{\partial x}\left(\frac{1}{x}\right) dx \\ &&& \text{(set } u \coloneqq \tfrac{1}{x}\text{ , } du = \tfrac{\partial}{\partial x}\left(\tfrac{1}{x}\right) dx\text{)} \\ &= \int_{-1}^{0} \int_{\{a^R\cdot\gamma_W(u)=0\}} \left|a^R \cdot \gamma_W'(u)\right| \sigma(a^R)\, db^{k-1}\, du. && (3.116) \end{aligned}$$

Similarly for the other integral we obtain that

$$\int_{1}^{\infty} \int_{\{a\cdot\gamma_S(x)=0\}} |a\cdot\gamma_S'(x)|\sigma(a)\,db^{k-1}\,dx = \int_{0}^{1} \int_{\{a^R\cdot\gamma_W(u)=0\}} \left|a^R\cdot\gamma_W'(u)\right|\sigma(a^R)\,db^{k-1}\,du\,. \tag{3.117}$$

Now by substituting Equation (3.116) and Equation (3.117) into Equation (3.109) we have

$$\begin{aligned}
&\int_{\mathbb{R}} \int_{\{a\cdot\gamma_S(x)=0\}} |a\cdot\gamma_S'(x)|\sigma(a)\,db^{k-1}\,dx = \\
&\int_{-1}^{1} \int_{\{a\cdot\gamma_S(x)=0\}} |a\cdot\gamma_S'(x)|\sigma(a)\,db^{k-1}\,dx \qquad + \int_{-1}^{0} \int_{\{a^R\cdot\gamma_W(u)=0\}} \left|a^R\cdot\gamma_W'(u)\right|\sigma(a^R)\,db^{k-1}\,du \\
&\qquad\qquad + \int_{0}^{1} \int_{\{a^R\cdot\gamma_W(u)=0\}} \left|a^R\cdot\gamma_W'(u)\right|\sigma(a^R)\,db^{k-1}\,du \\
&= \int_{-1}^{1} \int_{\{a\cdot\gamma_S(x)=0\}} |a\cdot\gamma_S'(x)|\sigma(a)\,db^{k-1}\,dx \quad + \int_{-1}^{1} \int_{\{a^R\cdot\gamma_W(u)=0\}} \left|a^R\cdot\gamma_W'(u)\right|\sigma(a^R)\,db^{k-1}\,du \\
&= Z_{(-1,1)}\left(f(x),\mathcal{D}\right) + Z_{(-1,1)}\left(g(x),\mathcal{D}\right)\,,
\end{aligned}$$

where in the last step we used Lemma 3.21 twice. □

The usefulness of the above lemma for proving bounds on the expected number of real zeros of k-sparse random polynomials should be clear: As long as no further restrictions are imposed, it suffices to bound the same quantity over the interval $(-1,1)$. Then by Theorem 3.22 the same bound applies for the expected number of real zeros over $\mathbb{R}$ with a loss of only a multiplicative factor of 2.

The next result is similar in nature and restricts further the domain of integration to $(0,1)$. This can often be useful as working over positive values of the variable x can either allow manipulations that are not possible otherwise or greatly simplify others. Unfortunately unlike the previous result, this requires a condition on the distribution other than being absolutely continuous. Namely, it requires that the distribution is symmetric about 0. This condition still allows for a large number of distributions including all normal distributions and continuous uniform distributions with mean 0. Furthermore, this result can apply to any support vector with additional arbitrary restrictions other than simply being k-sparse, since we do not require a second support vector as in the previous result. Nevertheless, we choose to state the result for the interval $(-1,1)$ to highlight how it can be used in conjunction with Theorem 3.22. The proof of the following Theorem is rather similar to the one above, although rather simpler. We now state the relevant result and give its proof.

Theorem 3.23 (Reducing domain of integration to $(0,1)$)**.** *Let $f(x) = a\cdot S(x)$ be a random k-sparse polynomial with support vector $S(x) = (1, x^{e_1}, \ldots, x^{e_{k-1}})$, degree $d =$*

e_{k-1}, and coefficient vector $a = (a_0, \ldots, a_{k-1})$ such that the coefficients a_i are independent random variables following absolutely continuous distributions that are symmetric, that is $\sigma(a_i) = \sigma(-a_i)$ for $0 \leq i \leq k-1$ and $a_i \in \mathbb{R}$. Thus, the joint distribution $\mathcal{D}$ for the coefficient vector a has density $\sigma(a) = \prod_{i=0}^{k-1} \sigma(a_i)$. Also let $g(x) = a^P \cdot S(x)$ be a random k-sparse polynomial with the same support vector $S(x)$, degree $d = e_{k-1}$ and coefficient vector a^P where $a_i^P = (-1)^{e_i} a_i$ for $0 \leq i \leq k-1$.

We then have that the expected number of real zeros of $f(x)$ over $(-1, 1)$ with respect to the distribution $\mathcal{D}$ is equal to the sum of the expected number of real zeros of $f(x)$ and $g(x)$ over $(0, 1)$, that is

$$Z_{(-1,1)}\left(f(x), \mathcal{D}\right) = Z_{(0,1)}\left(f(x), \mathcal{D}\right) + Z_{(0,1)}\left(g(x), \mathcal{D}\right) .$$

Proof. Similar to the previous proof, we let $\gamma(x)$ be the normalized support curve corresponding to $S(x)$. We also use Lemma 3.21 to express $Z_{(-1,1)}\left(f(x), \mathcal{D}\right)$ as an integral which we then decompose. We have

$$\begin{aligned}
&\int_{-1}^{1} \int_{\{a \cdot \gamma(x) = 0\}} |a \cdot \gamma'(x)| \sigma(a)\, db^{k-1}\, dx = \\
&\int_{0}^{1} \int_{\{a \cdot \gamma(x) = 0\}} |a \cdot \gamma'(x)| \sigma(a)\, db^{k-1}\, dx + \int_{-1}^{0} \int_{\{a \cdot \gamma(x) = 0\}} |a \cdot \gamma'(x)| \sigma(a)\, db^{k-1}\, dx \,. \qquad (3.118)
\end{aligned}$$

The first integral is equal to $Z_{(0,1)}\left(f(x), \mathcal{D}\right)$, so no further action is required. We now rewrite the second integral to match the other summand. As before, we will use an appropriate change of variable, this time $u \coloneqq -x$. We first examine the relationship between the two random polynomials $f(x) = a \cdot S(x) = \sum_{i=0}^{k-1} a_i x^{e_i}$ and $g(x) = a^P \cdot S(x) = \sum_{i=0}^{k-1} a_i^P x^{e_i} = \sum_{i=0}^{k-1} (-1)^{e_i} a_i x^{e_i}$, for which we have

$$\begin{aligned}
f(x) = a \cdot S(x) = \sum_{i=0}^{k-1} a_i x^{e_i} = \\
\sum_{i=0}^{k-1} (-1)^{e_i} a_i \left(-x\right)^{e_i} = \sum_{i=0}^{k-1} a_i^P \left(-x\right)^{e_i} = a^P \cdot S(-x) = g(-x) \,. \qquad (3.119)
\end{aligned}$$

It is easy to see that the following is true for the magnitude of $S(x)$

$$\|S(x)\| = \sum_{i=0}^{k-1} x^{2e_i} = \sum_{i=0}^{k-1} \left(-x\right)^{2e_i} = \|S(-x)\| \,. \qquad (3.120)$$

We can show the following relation for the defining equation of the hyperplane perpendicular to $\gamma(x)$. Namely, we have

$$a \cdot \gamma(x) = \frac{a \cdot S(x)}{\|S(x)\|} = \frac{f(x)}{\|S(x)\|} \qquad \text{(using eq. (3.119), eq. (3.120))}$$
$$= \frac{g(-x)}{\|S(-x)\|} = \frac{a^P \cdot S(-x)}{\|S(-x)\|} = a^P \cdot \gamma(-x)\,. \tag{3.121}$$

By differentiating Equation (3.121), we also obtain the following:

$$\frac{\partial}{\partial x}\left(a \cdot \gamma(x)\right) = \frac{\partial}{\partial x}\left(a^P \cdot \gamma(-x)\right) \iff a \cdot \gamma'(x) = -a^P \gamma'(-x)\,. \tag{3.122}$$

It is easy to see that the following is true for the joint densities of a and a^P

$$\sigma(a) = \prod_{i=0}^{k-1} \sigma(a_i) \qquad \text{(using } \sigma(a_i) = \sigma(-a_i)\text{)}$$
$$= \prod_{i=0}^{k-1} \sigma((-1)^{e_i} a_i) = \prod_{i=0}^{k-1} \sigma(a_i^P) = \sigma(a^P)\,. \tag{3.123}$$

Now by substituting Equation (3.121), Equation (3.122) and Equation (3.123) into the second integral of Equation (3.118) we have

$$\int_{-1}^{0} \int_{\{a\cdot\gamma(x)=0\}} |a \cdot \gamma'(x)| \sigma(a)\, db^{k-1}\, dx =$$
$$\int_{-1}^{0} \int_{\{a^P\cdot\gamma(-x)=0\}} |-a^P \cdot \gamma'(-x)| \sigma(a^P)\, db^{k-1}\, dx = \qquad \text{(reverse limits of outer integral)}$$
$$\int_{0}^{-1} \int_{\{a^P\cdot\gamma(-x)=0\}} |a^P \cdot \gamma'(-x)| \sigma(a^P)\, db^{k-1}\, \frac{\partial}{\partial x}(-x)\, dx\,. \tag{3.124}$$

Before proceeding we must address an important point with respect to the measure of the inner integral. In Theorem 3.22 only the order of the random variables a_i changed, thus it is clear that the measure remains the same. In this case however, for all a_i such that e_i is odd, a_i is replaced by $-a_i$ which could potentially alter the measure. We now show that the measure indeed remains the same. Let P be the Jacobian matrix that represents this change of variables, that is $a_i = Pa^P$. Clearly P is a diagonal matrix with entries $(-1)^{e_i}$, thus its determinant is equal to $\det(P) = (-1)^{\sum_{i=0}^{k-1} e_i}$ and in particular $|\det(P)| = 1$.

Before this change of variables must occur however, we must change from the basis $B = (b_0, b_1, \ldots, b_{k-1})$ of the space that db^{k-1} was defined over to the standard basis of R^{k-1}. Let da^{k-1} be the measure over R^{k-1} with respect to the standard basis, da_P^{k-1} the measure with respect to a^P and $d\beta^{k-1}$ the measure after the change of variables in

Equation (3.124) has occurred. Furthermore, let M be the Jacobian that performs the change of variable from the basis B to the standard normal basis, that is $U = M \cdot A$, where U are the coefficients of a vector over the base B. Naturally the reverse matrix M^{-1} performs the change of variables $A = M^{-1}U$. Thus to perform the change of variables in the integral, we can change from the base B to the standard normal base, perform the change of variables that alters the appropriate signs according to P and then perform the reverse transform M^{-1} to once again obtain the coordinates over the base B. These three change of variables can be performed all that once by taking the product of the three matrices, that is if U' are the coefficients we obtain after all three transforms we have

$$U' = (M^{-1}PM)U\,.$$

Thus the two measures and db^{k-1} and $d\beta^{k-1}$ are related by the following equation

$$\begin{aligned}
&db^{k-1} = \left|\det(M^{-1}PM)\right| d\beta^{k-1} \iff \\
&db^{k-1} = \left|\det(M^{-1})\det(P)\det(M))\right| d\beta^{k-1} \iff && (\det(M^{-1}) = (\det(M))^{-1}) \\
&db^{k-1} = |\det(P))|\, d\beta^{k-1} \iff db^{k-1} = d\beta^{k-1}\,. && (3.125)
\end{aligned}$$

Having shown that the two measures are equal, we can now proceed. For simplicity's sake, we keep the same notation for the measure. We can write Equation (3.124) by using Equation (3.125) as

$$\begin{aligned}
&\int_0^{-1} \int_{\{a^P\cdot\gamma(-x)=0\}} |a^P\cdot\gamma'(-x)|\sigma(a^P)\, db^{k-1}\, \frac{\partial}{\partial x}(-x)\, dx = \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (\text{set } u \coloneqq -x\ ,\ du = \tfrac{\partial}{\partial x}(-x)\, dx) \\
&\int_0^{1} \int_{\{a^P\cdot\gamma(u)=0\}} \left|a^P\cdot\gamma'(u)\right| \sigma(a^P)\, db^{k-1}\, du = Z_{(0,1)}\left(g(x), \mathcal{D}\right)\,. \qquad (3.126)
\end{aligned}$$

Now by simply substituting Equation (3.126) into Equation (3.118) we obtain the result stated.

□

Note that we stated the above result for $(-1, 1)$ in context with the previous result to emphasize that if the conditions of both theorems are satisfied, we can reduce the study of expected real zeros over the entirety of $\mathbb{R}$ to simply over the interval $(0, 1)$. However, the above result can be used more generally to replace any interval (a, b) over the negative reals to the corresponding interval over the positives, i.e. $(|b|, |a|)$, or in the case of $(-\infty, b)$ simply $(|b|, \infty)$. Similarly Theorem 3.22 can be used to reduce the domain from integration from any interval $(1, a)$ for $a > 1$ to $(\frac{1}{a}, 1)$ and similarly any interval $(-a, -1)$ to $(-1, -\frac{1}{a})$.

This concludes this section of basic results that can be used for arbitrary continuous distributions. Although they help with reducing the domain of integration for the integral,

the main difficulty remains to solve the integral for each case of interest. Nevertheless, we hope that these results will serve further efforts by both the author and his colleagues as well as the broader community interested in the question towards resolving certain cases.

This is also the end of the current chapter that presented the results of the present work. In Chapter 4 we conclude this work by summarizing our results, posing open questions that we consider interesting and expressing our thoughts regarding future results and directions of research.

CHAPTER 4

Conclusion

In this chapter we conclude this work by summarizing the best known results for the questions we discussed, as well as pose what we consider to be important open questions. We recall the state of the art results for each case, regardless if they are well-known results or new ones, with the objective that the reader can easily glimpse what is the state of the knowledge in the area, at least at the time of the writing of this work. When it comes to open problems and questions, there is a greater focus on our own work, since as authors we are the ones most familiar with the work and how it can be expanded and supplemented. Nevertheless, we also bring attention to important open questions that we wish the community to examine closely and hopefully resolve in the near future.

We begin with results regarding the real zeros of polynomials. We distinguished between the fixed case, where polynomials have specific coefficients and the random case, where the random polynomials have coefficients that are random variables that follow some distribution. We also distinguished between the univariate case, where the polynomials have only one variable x, and the typically harder multivariate case, where polynomials can have more than one variable, namely a variable vector $X = (x_1, \ldots, x_n)$.

The fixed univariate case is perhaps the easiest to handle from all the possible variations. Indeed, both in the dense and sparse case optimal bounds have been known for centuries, in terms of the degree d in the former and in terms of the sparsity k in the latter. For the dense case, a polynomial can have at most d real zeros, a well known upper bound which we stated and proven in Lemma 2.4. Furthermore, it is easy to construct polynomials that have d real zeros. Similarly, in the sparse case, a bound of $2k - 1$ real roots for k-sparse univariate polynomials follows from Descartes' rule of signs, as we analyzed in Theorem 2.17 and Theorem 2.21. Similarly, we noted in Section 2.1.2 that k-sparse polynomials exist that match this upper bound, see [32] for details. We conclude that the fixed univariate case has essentially been solved, although interesting questions can still be posed for restricted questions or settings different than the reals.

In this work we also provided a detailed analysis of Descartes' rule of signs, offering a deep understanding on its inner workings that to the best of our knowledge is not to be found in any modern text. That is not surprising since such a detailed discussion is out of scope for most textbooks on the subject while too lengthy for research work. We hope that our investigation, which includes an original approach via a specially defined type of derivative, provides a reference to those interested in exploring this important to this day Theorem further. In the same spirit, we briefly survey the most important results that followed Descartes' work and generalized it in several ways. In particular, we mention Budan's rule in Theorem 2.24 that generalizes Descartes' rule on any real interval. Fourier has proven the same result as Budan, albeit using a different approach. De Gua's rule stated in Lemma 2.25 provides a lower bound on the number of complex roots a univariate polynomial can have. Clearly this also can be made into an upper bound

for the number of real roots, although it is also interesting in how one can manipulate the number of complex roots by only altering the exponents of a polynomial, including random ones. We also show a number of extensions for Descartes' rule, including allowing for real exponents in Lemma 2.26, as well as infinite series in Corollary 2.27. One of the most refined results of this type is due to Laguerre that we stated in Theorem 2.28. Laguerre's work not only generalized all previous results but also potentially provides better upper bounds for real zeros in practice due to its formulation. We strongly believe that his work is relatively unknown even among experts in the field and deserves more recognition, since it can be potentially a powerful tool.

In this work we focused on the univariate case, but we also briefly mentioned some results for the multivariate case. We now briefly commentate on the results in the fixed case for multivariate polynomials. For dense polynomials, by the classical result due to Bézout, see Chapter 7 in [35] we can obtain a bound on the number of real roots of a polynomial system of n equations $f_1, \ldots, f_n$ on n unknowns, assuming that number is finite. Namely, if $d_i = \deg(f_i)$, we have that the number of real roots is either at most $d_1 \cdots d_n$ or infinite. This result naturally generalizes the bound on the univariate case. The sparse case is unfortunately not as well-studied, with Khovanskii's work on fewnomials, see Theorem 2.77 and [49], being the most systematic treatment to date and being a few decades old, relatively recent compared to Bézout's theorem that is over two centuries old. The best known result in the general case is due to Bihan and Sottile [7] and states that for a polynomial system of n equations with n variables and total sparsity k, the number of real roots in the positive orthant of $\mathbb{R}^n$ is bounded by $\frac{e^2+3}{4} 2^{\binom{k-n-1}{2}} n^{k-n-1}$. It is widely believed that this upper bound is far from optimal, see Theorem 2.79 for details. In both cases, there exist results that study restricted cases, although as we've already stated the literature is significant richer for the dense case.

Having presented the results for the fixed case, we proceed to discuss the results for random polynomials. The coefficients of such polynomials are random variables and in almost all results are also independent and identically distributed. We thus say that a random polynomial follows a distribution meaning each of its coefficient does. Typically these include well-known distributions that are also used often in practice. The dense case for random polynomials was intensively studied from the 1930s onwards, with optimal results obtained already in the 1940s and 1950s. The work due to Kac, see Theorem 2.41 and [45], is instrumental since it allows us to state the expected number of zeros with respect to some distribution as an integral. Kac used this method to show that the expected number of real zeros of a random degree d polynomial following the standard normal distribution is $\Theta(\log d)$, see Theorem 2.42. Using the same technique, he also showed that for the continuous uniform distribution $\mathcal{U}(-1, 1)$ the expected number of real zeros is also $\Theta(\log d)$, see Theorem 2.43. More than a decade later, Erdős and Offord managed to show that the same bound holds for the Rademacher distribution, see Theorem 2.44. In general, discrete distributions appear to be much harder to work with, a pattern that appears to hold also for the sparse case. Questions regarding the expected number of real zeros of random dense polynomials, as well as relevant questions regarding in general the distribution of zeros of such polynomials, continued constantly to be a topic of research interest and do so to this day.

We now turn our attention to sparse random polynomials. The results in this area are much more recent, with almost all of them being motivated by the real τ-conjecture,

including this work. We briefly recall the best results in the multivariate case, before focusing in the univariate case which also includes the results of this work. We distinguish between two kinds of sparse polynomials systems: unmixed systems where each polynomial has the same vector and mixed systems where each polynomial may have a different support vector. For unmixed systems it is known, see Theorem 2.80 and [20], that for polynomial systems of n random k-sparse polynomials on n variables where the i-th coefficient of each polynomial is an i.i.d. random variable following a normal distribution with mean 0, that the expected number of real zeros in the positive orthant of $\mathbb{R}^n$ is bounded by $\frac{1}{2^{n-1}}\binom{k}{n}$. Note that the coefficients of the same monomial must follow the same distribution in all polynomials. Now consider the same setup with the exception that each random polynomial is allowed to have a different support vector, that is the system is mixed, with the polynomial f_i having sparsity k_i for $1 \leq i \leq n$. Let the monomials of f_i viewed as a vector form a set A_i that has convex hull P_i and let $P = P_1 + \cdots + P_n$ be the Minkowski sum of the convex sets P_i. Then it is known, see Theorem 2.82 and [18], that the expected number of real zeros of the polynomial system in the positive orthant of R^n is bounded by $(2\pi)^{-\frac{n}{2}} V_0(k_1 - 1) \cdots (k_n - 1)$, where V_0 denotes the number of vertices of the Minkowski sum P.

Similarly, in the univariate case for sparse random polynomials, almost all results are recent. In fact, this work contains most of the recent results known. As before, the coefficients are i.i.d. random variables following some distribution. For random k-sparse polynomials following the standard normal distribution, we showed in detail the results of our previous work [43], which establish both lower and upper bounds. In particular, by Theorem 3.14 we show that the number of expected real zeros in this case is bounded by $\frac{8}{\pi}\sqrt{k-1}$. Similarly, we showed in Theorem 3.19 that for the support vector $S(x) = \left(1, x, x^{2^2}, \ldots, x^{2^{2^i}}, \ldots, x^{2^{2^{k-2}}}\right)$, $1 \leq i \leq k-2$, the expected number of real zeros is at least $0.014\sqrt{k} - 0.007$. Thus the upper and lower bound match asymptotically and we can say that for k-sparse random polynomials following the standard normal distribution, the expected number of zeros is $\Theta(\sqrt{k})$. In fact, we're able to show that most of the zeros of such random polynomials concentrate near $x = 1$. Specifically, we show in Theorem 3.16 that for any fixed $\varepsilon > 0$, the expected number of real zeros in the $(0, 1-\varepsilon)$ interval is at most $\frac{1}{2\pi}\left(\log\left(\frac{2}{\varepsilon}\right) + \frac{4}{\sqrt{\varepsilon}} - 4\right)$. In this work, we also show in Theorem 2.37 that with little effort the results due to Bloch and Pólya [10] for dense random polynomials following the Rademacher distribution can be adapted to the sparse case, showing an upper bound of $O(\sqrt{k})$. The same arguments as in Corollary 2.38 can be used to extend this upper bound for the distribution that assigns the values $1, -1$ and 0 to the random coefficients with equal probability.

We also presented some basic results that generalize techniques used in the standard normal case to arbitrary absolutely continuous distributions. These are intermediary results that have not yet been used to show bounds but we believe have the potential to simplify future results. In particular, we show in Theorem 3.22 that if we wish to bound the expected number of real zeros of a random k-sparse polynomial $f(x)$ following the absolutely continuous distribution $\mathcal{D}$ over the entirety of $\mathbb{R}$, it suffices to bound the expected number of real zeros in the interval $(-1, 1)$. Furthermore, if the distribution $\mathcal{D}$ is symmetric about 0, we show in Theorem 3.23 that it suffices to bound the expected number of real zeros in the interval $(0, 1)$. Assuming the conditions of both results hold, we

can thus obtain a bound for the expected number of real zeros in $(0, 1)$ and automatically have a bound for the entirety of $\mathbb{R}$ with only an added multiplicative factor of 4, which is in particular constant. These results are implicitly employed in the standard normal case and we believe that they can be instrumental in simplifying the calculations for other distributions too.

It would be remiss to end this summary of the best known results without mentioning a significant relatively recent development in the real τ-conjecture, which as we've detailed in Section 2.3.3 forms the main motivation for this work. In particular, in [13] it is shown that indeed the real τ-conjecture is true on average, with respect to the standard normal distribution. Furthermore, they show that the result can be generalized to more distributions that satisfy certain mild conditions. This is a significant development that has several immediate consequences. First of all, it strengthens our confidence that the real τ-conjecture may be true. Although it does not preclude the possibility that some fixed polynomial may violate the conjecture, it allows us to say with certain confidence that such a polynomial must be a rather special case and that most polynomials of the form given in the conjecture will have a polynomial number of real roots in the parameters. In fact, by a simple application of Markov's inequality, we have that a random polynomial following the standard normal distribution has more than $c\, Z_{\mathbb{R}}(f(x), \mathcal{N}(0, 1))$ zeros with probability at most $\frac{1}{c}$ for $c > 1$. Thus a very small amount of polynomials would not satisfy the conditions of the conjecture, although on the other hand it suffices that a single polynomial does to refute it. Furthermore, this results underlines the need to further study the subject of both the expected number of zeros of k-sparse polynomial under certain distributions, as well as how these behave when combined with various operations, in particular the form that appears in the statement of the real τ-conjecture.

Having summarized the most relevant best known results that concerned us in this work, we continue with posing open questions and problems that we believe merit further investigation by the broader research community interested in the subject. All of these questions relate to one or more of the results summarized above and we believe that solving any of them would further our understanding. Nevertheless, we feel that some are of greater significance than others, although the difficulty of resolving the open question should also be taken into account. We thus state these suggested research directions in the order of importance we consider appropriate, although naturally such a ranking has a certain degree of subjectivity.

We start with the real τ-conjecture itself, since it is our primary motivation for this work. Naturally, resolving the conjecture in the positive would allow us to prove Valiant's conjecture, see Conjecture 2.50, thus separating VP from VNP. We expect however both Valiant's conjecture as well as the real τ-conjecture that implies it to be highly non-trivial to solve and require considerable sustained effort from the research community. To highlight how rudimentary the current knowledge we possess is, consider the expression $fg+1$ where both f and g are fixed k-sparse univariate polynomials. It is easy to see that fg can have at most $2k$ roots, however for $fg + 1$ the only trivial bound is by Descartes' rule of signs. Sign this expression has $O(k^2)$ sparsity, that is also the magnitude of the bound obtained. In [55] an improved upper bound of $O(k^{\frac{4}{3}})$ was shown assuming f and g are bivariate k-sparse polynomials. It is suspected that both in the bivariate as well as the univariate case linear bounds should be possible, although that remains an open question. Clearly this is one of the simplest expressions that conform to the form given

in the real τ-conjecture and the inability to solve it highlights the hurdles that must be overcome. Improving the upper bound for this simple case to a linear one or proving it cannot be further improved would be a significant improvement and would likely require the development of techniques that would allow us to handle the zeros of sums of sparse polynomials. We thus have the first two open questions:

Question 4.1 (Real τ-conjecture)**.** *Is the real τ conjecture (see Conjecture 2.76) true?*

Question 4.2 ($fg+1$ problem)**.** *Let f and g be k-sparse univariate polynomials. Is it possible to show that the polynomial $fg+1$ has a number of real zeros bounded by $O(k)$?*

A promising intermediate step towards resolving the τ-conjecture and in general related questions about fixed polynomials is understanding the behavior of polynomials under random distributions. The intuition is that randomness allow us to avoid the peculiarities of certain difficult cases, while still getting a good overall picture of the behavior in average. Clearly, the choice of distribution severely affects this. The first step would be to target other absolutely continuous distributions, since for all continuous distributions no single "bad" case can affect the outcome, since every single point occurs with probability 0. By the same reasoning, an increase in difficulty would be resolving the real τ-conjecture on average with respect to discrete distributions, since a single case with a high number of zeros can severely affect the expected number of zeros. We expect that as the tools developed mature, handling discrete distributions is the natural way to proceed and gain useful insights that can then be used to handle the non-random case. We thus pose the following problem to the community:

Question 4.3 (Real τ-conjecture on average for other distributions)**.** *Can the results of [13] be generalized for more distributions? In particular, is it possible to show similar results, i.e., that the real τ-conjecture is true on average, with respect to discrete distributions?*

Related to the question of resolving the real τ-conjecture for other distributions is determining the expected number of real zeros of k-sparse random polynomials for distributions other than the one we examined in this work, namely the standard normal distribution. Essentially we can break down resolving the real τ conjecture into two parts: understanding the behavior of the number of zeros of each polynomial in the expression and then understanding how operations on the polynomials affects that number. Furthermore, as we've detailed in this work, the expected number of zeros of sparse polynomials with respect to various distributions is of independent interest. Once again absolutely continuous distributions make for an ideal candidate for a first challenge to tackle, especially given the tools developed in Section 3.3. The continuous uniform distribution with mean 0, which has already been studied in the dense case, seems like an obvious candidate for this sort of investigation. Once again, a step-up in difficulty comes with discrete distributions, for the same reasons we detailed above. We thus have the following set of problems that can be tackled.

Question 4.4 (Expected number of real zeros of k-sparse polynomials under other distributions)**.** *What is the expected number of real zeros of k-sparse univariate random polynomials under distributions other than the standard normal one? How significant*

the change in that number is for various absolutely continuous distributions? Can the same questions be answered for discrete distributions and how do discrete and continuous distributions compare with respect to the expected number of real zeros?

Another approach is to consider other settings for the same question of the expected number of real zeros. In particular, we have already mentioned that both in the random and fixed case in the multivariate setting the bounds obtained are believed to be highly suboptimal, see Section 2.3.4 for details. These results are of interest not only independently, but also in relation to the previous topics discussed, since there exists a formulation of the real τ-conjecture that concerns bivariate polynomials. Furthermore, as in the univariate case, distributions other than the normal one are poorly studied. All variations such as considering other distributions, continuous or discrete, are thus open, similar to Question 4.4. In addition, non-trivial lower bounds are not known for the multivariate case, unlike the univariate case where with Theorem 3.19 we showed that the bound proven in Theorem 3.14 cannot be further improved. The following open questions can thus be addressed.

Question 4.5 (Expected number of real zeros for the sparse multivariate case)**.** *Can the bounds given for the sparse multivariate case be improved, either for the fixed or random case? In particular, can the bounds given for the normal distributions be extended to other distributions, either absolutely continuous or discrete? What is the relation between the bound on the expected zeros and the various distributions? Can lower bounds, ideally matching the upper bounds given, be proven?*

Thus far the questions were general and although related to this work, could be also considered independent of it. The following open problems stem directly from this work and aim to improve certain aspects of it or investigate interesting questions that developed as a direct result of this work. The first such open question concerns the relation between the support vector chosen and the expected number of real zeros of a random k-sparse polynomial following the standard normal distribution. In this work we have proven an upper bound, see Theorem 3.14 that states that the expected number of real zeros of such polynomials is upper bounded by $O(\sqrt{k})$. Note that this applies to all k-dimensional support vectors. On the other hand, the lower bound presented in Theorem 3.19 applies to a specific support, namely $S(x) = \left(1, x, x^{2^2}, \ldots, x^{2^{2^i}}, \ldots, x^{2^{2^{k-2}}}\right)$, $1 \leq i \leq k-2$. For this specific support we show that the expected number of zeros is $\Theta(\sqrt{k})$, thus proving that the upper bound cannot be further improved. However, note that it is not necessary that all support vectors obey that lower bound. In fact, we know that this is not the case, by considering the well-studied dense case in terms of sparsity with $k = d + 1$. Then by Theorem 2.42 the expected number of zeros is $\Theta(\log d) = \Theta(\log k)$, which is significantly smaller. The question thus arises, how the choice of support vector affects the expected number of zeros? Are there specific attributes of a support vector that determine that number? It is clear that certain support vectors are related in this regard, for example by multiplying all components of a support vector $S(x) = (x_{e_0}, \ldots, x_{e_{k-1}})$ by a power of x, the expected number of real zeros does not change, since $x = 0$ depends on a coefficient being zero which has a probability of 0 under the standard normal distribution. It is clear that the relative distance between subsequent exponents $e_i - e_{i-1}$, plays an important role in this relation. Note that this intuition is verified by De Gua's rule, see Lemma 2.25,

which however only considers the largest gap of that type. Thus, further investigation is required. Towards this goal, we have performed a small number of experiments, by considering all possible support vectors up to a specific sparsity and degree, since for each k the number of possible support vectors is infinite. This limited experimental data hints at a dichotomy, where support vectors would either match the dense case with a number of expected zeros close to $\Theta(\log k)$ or they would match the sparse case of $\Theta(\sqrt{k})$. Rather surprisingly, no polynomials fell in between in our limited experiments, as one would expect from continuous objects. Note however that we consider a continuous distribution that must be sampled and that the number of possible support vectors increases exponentially, since for degree d and sparsity k there are $\binom{d+1}{k} = O(d^k)$ distinct support vectors, thus rather quickly the numbers involved become too large for any reasonable computer simulation. A theoretical approach depending asymptotically on k would thus be of much greater interest, which leads to the following conjecture.

Conjecture 4.6 (Support vector dichotomy conjecture)**.** *Let $f(x) = a \cdot S(x)$ be a random k-sparse polynomial following the standard normal distribution. Then depending on the support vector $S(x)$, its expected number of real zeros is either $\Theta(\log k)$ or $\Theta(\sqrt{k})$.*

Another interesting question once again relates to the support vector for which the lower bound is obtained. Recall that we show that for a k-sparse random polynomial following the standard normal distribution and with support vector is $S(x) = \left(1, x, x^{2^2}, \ldots, x^{2^{2^i}}, \ldots, x^{2^{2^{k-2}}}\right)$, $1 \leq i \leq k-2$, the expected number of real roots is $\Theta(\sqrt{k})$, as shown in Theorem 3.19. This support vector has exponents that increase doubly exponentially with i, as does the gap between subsequent exponents as well. Naturally, the same applies to the last exponent ,that is the degree, which depends doubly exponential in k, that is $d = 2^{2^{k-2}}$. The obvious question to ask is whether such large exponents are necessary to the result or whether they exist support vectors of smaller degree and smaller gaps between exponents that also have $\Theta(\sqrt{k})$ expected number of real zeros for such polynomials. In particular, it would be interesting to see if there exist support vectors whose degree is at most polynomial in k that exhibit a high expected number of real zeros. This question naturally could also be answered if the previously stated Conjecture 4.6 is.

Question 4.7 (Support vectors with a high expected number of real zeros)**.** *Let $f(x) = a \cdot S(x)$ be a random k-sparse polynomial following the standard normal distribution. Do they exist support vectors other than the one stated in Theorem 3.19 or ones easily derived from it, such that the expected number of real zeros is $\Theta(\sqrt{k})$. In particular, does such a support vector $S(x)$ exist so that the degree of $f(x)$ is polynomial in k?*

The last open question we would like to pose directly relates to the main results we presented in Theorem 3.14 and Theorem 3.19. Note that by the former the expected number of real zeros of a random k-sparse polynomial following the standard normal distribution is bounded by $\frac{8}{\pi}\sqrt{k-1}$, while in Theorem 3.19 we show that there exists a support vector for which the expected number of real zeros is at least $0.014\sqrt{k} - 0.007$. While asymptotically this bound match, there is still room for improvement in terms of constants. Similar work has been carried out for the results due to Kac, including the work due to Edelman and Kostlan [28], which among other results improves upon the

constants involved. While such questions are often not commonly studied in Computer Science where asymptotic results are the norm, they are often regarded as more important in Mathematics. Another area of possible improvement with respect to this particular result is to investigate higher moments relating to the distribution of real zeros of such a random polynomial. In particular, it would be interesting to know the standard deviation σ of the distribution of real zeros. Once again, similar work has been carried out for the dense case, see [65] for an example. We thus have the following open problem.

Question 4.8 (Refinement of the main results of this work)**.** *Can the constants involved in Theorem 3.14 and Theorem 3.19 be further improved? Furthermore, is it possible to ascertain additional information for the distribution of real zeros of a random k-sparse polynomial following the standard normal distribution, such as its variance or other higher moments?*

This concludes this work. We hope that the reader appreciated this effort as much as the author did and that it was beneficial to their undertakings. Furthermore, we hope that the questions raised in this section will be put to scientific dialogue and leads to new exciting results and frontiers to explore.

List of Figures

List of Tables

List of Acronyms

i.i.d. independent and identically distributed (referring to random variables)

Bibliography

[1] R. B. Ash and C. A. Doléans-Dade. *Probability and measure theory*. Harcourt/Academic Press, San Diego, 2nd edition, 2000.

[2] M. Avendaño. The number of roots of a lacunary bivariate polynomial on a line. *Journal of Symbolic Computation*, 44(9):1280–1284, 2009.

[3] J.-M. Azaïs and M. Wschebor. *Level sets and extrema of random processes and fields*. John Wiley & Sons, Inc., Hoboken, NJ, USA, Feb. 2009.

[4] D. H. Bailey. Simple proofs: The fundamental theorem of algebra. `https://mathscholar.org/2018/09/simple-proofs-the-fundamental-theorem-of-algebra/`.

[5] M. Bensimhoun. Historical account and ultra-simple proofs of Descartes's rule of signs, De Gua, Fourier, and Budan's rule. arXiv:1309.6664, 2013.

[6] S. J. Berkowitz. On computing the determinant in small parallel time using a small number of processors. *Information processing letters*, 18(3):147–150, 1984.

[7] F. Bihan and F. Sottile. New fewnomial upper bounds from Gale dual polynomial systems. *Moscow mathematical journal*, 7(3), 2007.

[8] M. Bläser. On the complexity of the multiplication of matrices of small formats. *Journal of Complexity*, 19(1):43–60, 2003.

[9] M. Bläser. Fast matrix multiplication. *Theory of Computing*, pages 1–60, 2013.

[10] A. Bloch and G. Pólya. On the roots of certain algebraic equations. *Proceedings of the London Mathematical Society*, 2(1):102–114, 1932.

[11] L. Blum, M. Shub, and S. Smale. On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines. *Bulletin of the American Mathematical Society*, 21(1):1–46, 1989.

[12] K. Bringmann. Fine-Grained Complexity Theory. In *36th International Symposium on Theoretical Aspects of Computer Science*, pages 4:1–4:7, 2019.

[13] I. Briquel and P. Bürgisser. The real tau-conjecture is true on average. *Random Structures & Algorithms*, 57(2):279–303, 2020.

[14] F.-D. Budan. *Nouvelle méthode pour la résolution des équations numériques d'un degré quelconque: d'après laquelle toute le calcul exigé pour cette résolution se réduit à l'emploi des deux premières règles de l'arithmétique*. Courcier, 1807.

[15] P. Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory*, volume 7 of *Algorithms and Computation in Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2000.

[16] P. Bürgisser. Cook's versus Valiant's hypothesis. *Theoretical Computer Science*, 235(1):71–88, 2000.

[17] P. Bürgisser. On defining integers and proving arithmetic circuit lower bounds. *Computational Complexity*, 18(1):81–103, 2009.

[18] P. Bürgisser. Real zeros of mixed random fewnomial systems. In *Proceedings of the 2023 International Symposium on Symbolic and Algebraic Computation*, pages 107–115. Association for Computing Machinery, 2023.

[19] P. Bürgisser, M. Clausen, and M. A. Shokrollahi. *Algebraic Complexity Theory*, volume 315 of *Grundlehren der mathematischen Wissenschaften*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1997.

[20] P. Bürgisser, A. A. Ergür, and J. Tonelli-Cueto. On the Number of Real Zeros of Random Fewnomials. *SIAM Journal on Applied Algebra and Geometry*, 3(4):721–732, Jan. 2019.

[21] P. Bürgisser, J. M. Landsberg, L. Manivel, and J. Weyman. An overview of mathematical issues arising in the Geometric Complexity Theory approach to VP≠VNP. *SIAM Journal on Computing*, 40(4):1179–1209, 2011.

[22] P. Chaugule, N. Limaye, and A. Varre. Variants of homomorphism polynomials complete for algebraic complexity classes. *ACM Transactions on Computation Theory (TOCT)*, 13(4):1–26, 2021.

[23] P. M. Cohn. *Basic algebra: Groups, Rings, and Fields*. Springer Science & Business Media, London, 2003.

[24] J. P. De Gua. Recherches du nombre des racines réelles ou imaginaires, réelles positives ou réelles négatives, qui peuvent se trouver dans les équations de tous les degrés. *Histoire de l'Academie royale des sciences (sec. mémoires)*, pages 435–494, 1741.

[25] R. Descartes. *La Géométrie*. Hermann, 1886.

[26] Y. Do, H. Nguyen, and V. Vu. Real roots of random polynomials: expectation and repulsion. *Proceedings of the London Mathematical Society*, 111(6):1231–1260, 2015.

[27] A. Durand, M. Mahajan, G. Malod, N. de Rugy-Altherre, and N. Saurabh. Homomorphism polynomials complete for VP. In *34th International Conference on Foundation of Software Technology and Theoretical Computer Science (FSTTCS 2014)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2014.

[28] A. Edelman and E. Kostlan. How many zeros of a random polynomial are real? *Bulletin of the American Mathematical Society*, 32(1):1–37, 1995.

[29] P. Erdös and A. C. Offord. On the number of real roots of a random algebraic equation. *Proceedings of the London Mathematical Society*, 3(1):139–160, 1956.

[30] M. Fekete and G. Pólya. Über ein problem von Laguerre. *Rendiconti del Circolo Matematico di Palermo*, 34(1):89–120, 1912.

[31] J. Fourier. Sur l'usage du théoreme de Descartes dans la recherche des limites des racines. *Bulletin des sciences par la Société philomatique de Paris*, 156(165):181–187, 1820.

[32] D. J. Grabiner. Descartes' rule of signs: Another construction. *The American Mathematical Monthly*, 106(9):854–856, 1999.

[33] G. Grimm. *Über die reellen Nullstellen Dirichlet'scher L-Reihen.* PhD thesis, ETH Zurich, 1932.

[34] A. Gupta, P. Kamath, N. Kayal, and R. Saptharishi. Arithmetic circuits: A chasm at depth 3. *SIAM Journal on Computing*, 45(3):1064–1079, 2016.

[35] R. Hartshorne. *Algebraic geometry.* Graduate Texts in Mathematics, No. 52. Springer-Verlag, New York, 1977.

[36] TStancek. Do we know some quasi-polynomial problem that is known to not be in NP? Theoretical Computer Science Stack Exchange. `https://cstheory.stackexchange.com/q/38828`.

[37] I. A. Ibragimov. Independent and stationary sequences of random variables. *Wolters, Noordhoff Pub.*, 1975.

[38] I. A. Ibragimov and N. B. Maslova. On the expected number of real zeros of random polynomials I. Coefficients with zero means. *Theory of Probability & Its Applications*, 16(2):228–248, 1971.

[39] I. A. Ibragimov and N. B. Maslova. On the expected number of real zeros of random polynomials. II. Coefficients with non-zero means. *Theory of Probability & Its Applications*, 16(3):485–493, 1971.

[40] R. Impagliazzo and R. Paturi. On the complexity of k-SAT. *Journal of Computer and System Sciences*, 62(2):367–375, 2001.

[41] K. Ireland and M. Rosen. *A Classical Introduction to Modern Number Theory*, volume 84 of *Graduate Texts in Mathematics*. Springer New York, New York, NY, 1990.

[42] Z. Jelonek. A simple proof of the existence of the algebraic closure of a field. *Universitatis Iagellonicae Acta Mathematica*, 30:131–132, 1993.

[43] G. Jindal, A. Pandey, H. Shukla, and C. Zisopoulos. How many zeros of a random sparse polynomial are real? In *Proceedings of the 45th International Symposium on Symbolic and Algebraic Computation*, pages 273–280, 2020.

[44] V. Kabanets and R. Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 355–364, 2003.

[45] M. Kac. On the average number of real roots of a random algebraic equation. *Bulletin of the American Mathematical Society*, 49(4):314–320, 1943.

[46] M. Kac. On the Average Number of Real Roots of a Random Algebraic Equation (II). *Proceedings of the London Mathematical Society*, s2-50(1):390–408, 1948.

[47] K. A. Kalorkoti. A lower bound for the formula size of rational functions. *SIAM Journal on Computing*, 14(3):678–687, 1985.

[48] A. G. Khovanskii. A class of systems of transcendental equations. *Doklady Akademii Nauk*, 255(4):804–807, 1980.

[49] A. G. Khovanskii. *Fewnomials*, volume 88. American Mathematical Soc., 1991.

[50] P. Koiran. Valiant's model and the cost of computing integers. *Computational Complexity*, 13(3-4):131–146, 2005.

[51] P. Koiran. Shallow circuits with high-powered inputs. arXiv:1004.4960, 2010.

[52] P. Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theoretical Computer Science*, 448:56–65, 2012.

[53] P. Koiran and S. Perifel. Interpolation in Valiant's theory. *Computational Complexity*, 20:1–20, 2011.

[54] P. Koiran, N. Portier, and S. Tavenas. A Wronskian approach to the real τ-conjecture. *Journal of Symbolic Computation*, 68:195–214, 2015.

[55] P. Koiran, N. Portier, S. Tavenas, and S. Thomassé. A τ-conjecture for Newton polygons. *Foundations of computational mathematics*, 15(1):185–197, 2015.

[56] E. N. Laguerre. *Théorie des équations numériques*. Gauthier-Villars, 1884.

[57] S. A. Levin. Descartes' rule of signs-How hard can it be. `https://sepwww.stanford.edu/oldsep/stew_save/descartes.pdf`, 2002.

[58] J. E. Littlewood and A. C. Offord. On the number of real roots of a random algebraic equation. *Journal of the London Mathematical Society*, s1-13(4):288–295, 1938.

[59] J. E. Littlewood and A. C. Offord. On the number of real roots of a random algebraic equation. II. *Mathematical Proceedings of the Cambridge Philosophical Society*, 35(2):133–148, 1939.

[60] J. E. Littlewood and A. C. Offord. On the number of real roots of a random algebraic equation (III). *Matematicheskii Sbornik*, 12(3):277–286, 1943.

[61] B. F. Logan and L. A. Shepp. Real zeros of random polynomials. *Proceedings of the London Mathematical Society*, s3-18(1):29–35, 1968.

[62] B. F. Logan and L. A. Shepp. Real zeros of random polynomials. II. *Proceedings of the London Mathematical Society*, s3-18(2):308–314, 1968.

[63] M. Mahajan and V. Vinay. A combinatorial algorithm for the determinant. In *Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '97, page 730–738. Society for Industrial and Applied Mathematics, 1997.

[64] G. Malajovich and J. M. Rojas. High probability analysis of the condition number of sparse polynomial systems. *Theoretical computer science*, 315(2-3):525–555, 2004.

[65] N. B. Maslova. On the variance of the number of real roots of random polynomials. *Theory of Probability & Its Applications*, 19(1):35–52, 1974.

[66] N. B. Maslova. On the distribution of the number of real roots of random polynomials. *Theory of Probability & Its Applications*, 19(3):461–473, 1975.

[67] A. Muleshkov and T. Nguyen. Easy proof of the Jacobian for the n-dimensional polar coordinates. *Pi Mu Epsilon Journal*, 14(4):269–273, 2016.

[68] J. Neukirch. *Algebraic Number Theory*, volume 322 of *Grundlehren der mathematischen Wissenschaften*. Springer Berlin Heidelberg, 1999.

[69] H. Nguyen, O. Nguyen, and V. Vu. On the number of real roots of random polynomials. *Communications in Contemporary Mathematics*, 18(04):1550052, 2016.

[70] A. M. Ostrowski. On two problems in abstract algebra connected with Horner's rule. In *Studies in Mathematics and Mechanics presented to Richard von Mises*, pages 40–48. Academic Press San Diego, 1954.

[71] T. E. Peet. Mathematischer Papyrus des staatlichen Museums der schönen Künste in Moskau. *The Journal of Egyptian Archaeology*, 17(1):154–160, 1931.

[72] R. Pemantle and I. Rivin. The distribution of zeros of the derivative of a random polynomial. In *Advances in Combinatorics*, pages 259–273. Springer Berlin Heidelberg, 2013.

[73] G. Pólya. Verschiedene Bemerkungen zur Zahlentheorie. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 28:31–40, 1919.

[74] E. Robson. Words and pictures: New light on Plimpton 322. *The American Mathematical Monthly*, 109(2):105–120, 2002.

[75] J. M. Rojas. Arithmetic multivariate Descartes' rule. *American Journal of Mathematics*, 126(1):1–30, 2004.

[76] H. J. Ryser. *Combinatorial mathematics*, volume 14. American Mathematical Soc., 1963.

[77] R. Saptharishi. A survey of lower bounds in arithmetic circuit complexity. `https://github.com/dasarpmar/lowerbounds-survey`.

[78] A. Sard. The measure of the critical values of differentiable maps. *Bulletin of the American Mathematical Society*, 48(12):883–890, 1942.

[79] A. Shpilka and A. Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends® in Theoretical Computer Science*, 5(3–4):207–388, 2010.

[80] M. Shub and S. Smale. On the intractability of Hilbert's Nullstellensatz and an algebraic version of "NP $\neq$ P?". *Duke Mathematical Journal*, 81(1):47–54, 1995.

[81] M. Sipser. Borel sets and circuit complexity. In *Proceedings of the fifteenth annual ACM symposium on Theory of computing*, pages 61–69, 1983.

[82] S. Smale. Mathematical problems for the next century. *Mathematical Intelligencer*, 20(2):7–15, 1998.

[83] F. Sottile. *Real solutions to equations from geometry*, volume 57. American Mathematical Soc., 2011.

[84] D. C. Stevens. *The average and variance of the number of real zeros of random functions*. PhD thesis, New York University, 1965.

[85] D. C. Stevens. The average number of real zeros of a random polynomial. *Communications on Pure and Applied Mathematics*, 22(4):457–477, 1969.

[86] V. Strassen. Gaussian elimination is not optimal. *Numerische mathematik*, 13(4):354–356, 1969.

[87] V. Strassen. Berechnung und Programm. I. *Acta Informatica*, 1(4):320–335, 1972.

[88] V. Strassen. Vermeidung von Divisionen. *Journal für die reine und angewandte Mathematik*, 264:184–202, 1973.

[89] T. Tao and V. Vu. Local universality of zeroes of random polynomials. *International Mathematics Research Notices*, 2015(13):5053–5139, 2015.

[90] S. Tavenas. *Bornes inferieures et superieures dans les circuits arithmetiques*. PhD thesis, Ecole normale supérieure de Lyon-ENS LYON, 2014.

[91] L. G. Valiant. Completeness classes in algebra. In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 249–261, 1979.

[92] L. G. Valiant, S. Skyum, S. Berkowitz, and C. Rackoff. Fast Parallel Computation of Polynomials Using Few Processors. *SIAM Journal on Computing*, 12(4):641–644, Nov. 1983.

[93] J. von zur Gathen and J. Gerhard. *Modern Computer Algebra*. Cambridge University Press, 3rd edition, 2013.

[94] H. Weyl. *The classical groups: their invariants and representations*. Princeton university press, 2nd edition, 1946.

[95] S. Winograd. On multiplication of 2×2 matrices. *Linear algebra and its applications*, 4(4):381–388, 1971.