On complex roots of the independence polynomial[*]

Ferenc Bencs[†] Péter Csikvári[‡] Piyush Srivastava[§] Jan Vondrák[¶]

Abstract

The independence polynomial of a graph is the generating polynomial of all its independent sets. Formally, given a graph G, its independence polynomial $Z_G(\lambda)$ is given by $\sum_I \lambda^{|I|}$, where the sum is over all independent sets I of G. The independence polynomial has been an important object of study in both combinatorics and computer science. In particular, the algorithmic problem of estimating $Z_G(\lambda)$ for a fixed positive λ on an input graph G is a natural generalization of the problem of counting independent sets, and its study has led to some of the most striking connections between computational complexity and the theory of phase transitions. More surprisingly, the independence polynomial for negative and complex values of λ also turns out to be related to problems in statistical physics and combinatorics. In particular, the locations of the complex roots of the independence polynomial of bounded degree graphs turn out to be very closely related to the Lovász local lemma, and also to the questions in the computational complexity of counting. Consequently, the locations of such zeros have been studied in many works. In this direction, it is known from the work of Shearer [29] and of Scott and Sokal [27] – inspired by the study of the Lovász local lemma – that the independence polynomial $Z_G(\lambda)$ of a graph G of maximum degree at most $d+1$ does not vanish provided that $|\lambda| \leq \frac{d^d}{(d+1)^{d+1}}$. Significant extensions of this result have recently been given in the case when λ is in the *right* half-plane (i.e., when $\Re\lambda \geq 0$) by Peters and Regts [26] and Bencs and Csikvári [9]. In this paper, our motivation is to further extend these results to find new zero free regions not only in the right half plane, but also in the *left* half-plane, that is, when $\Re\lambda \leq 0$.

We give new geometric criterions for establishing zero-free regions as well as for carrying out semi-rigorous numerical explorations. We then provide two examples of the (rigorous) use of these criterions, by establishing two new zero-free regions in the left-half plane. We also extend the results of Bencs and Csikvári [9] for the right half-plane using our framework. By a direct application of the interpolation method of Barvinok [5], combined with extensions due to Patel and Regts [25], our results also imply deterministic polynomial time approximation algorithms for the independence polynomial of bounded degree graphs in the new zero-free regions.

1 Introduction

The independence polynomial, also known as the partition function of the hard core lattice gas in the statistical physics literature, is the graph polynomial given by

$$Z_G(\lambda) := \sum_{I:\text{independent set in } G} \lambda^{|I|}.$$

An independent set in a graph G is subset of its vertices no two of which are adjacent in G. In statistical mechanics, the polynomial arises in the modeling of adsorption phenomena (usually with G being a lattice);

---

while in combinatorics, it is the natural generating function of independent sets of graphs, and offers a natural generalization to the problem of counting independent sets in a graph. These connections have led to the polynomial being studied extensively in the setting $\lambda > 0$, both in statistical physics and in computational complexity, and, in particular, has led to some very tight connections between the two fields [32, 35].

The setting $\lambda < 0$, and more generally, of complex λ is also of interest. In particular, the problem of understanding where the complex zeros of Z_G lie for graphs G in a given class is of special interest. In statistical mechanics, it relates to the Yang-Lee theory of phase transitions [36]. In the special case when G is a lattice, the work of Dobrushin and Shlosman [14, 15] also related the question to other, more probabilistic notions of phase transitions. In combinatorics, the behavior of Z_G at negative and complex λ plays an important role in the study of the Lovász local lemma; for a detailed discussion of this connection, we refer to the work of Shearer [29] as elucidated by Scott and Sokal [27]. For our purposes, we start with the following result proved in the above two papers. We denote by $\mathcal{G}_\Delta$ the set of finite graphs with vertex degrees at most Δ, for some fixed $\Delta \geq 3$.

THEOREM 1.1 ([**29**], **see also Corollary 5.7 and the discussion following it in [27]**). *Let $d \geq 2$ be an integer. If $\lambda \in \mathbb{C}$ is such that $|\lambda| \leq \lambda^*(d) := \frac{d^d}{(d+1)^{d+1}}$ then $Z_G(\lambda) \neq 0$ for all graphs $G \in \mathcal{G}_{d+1}$. Further, for any negative real $\lambda_1 < -\lambda^*(d)$, there exists a graph $G \in G_{d+1}$ and λ' satisfying $\lambda_1 < \lambda' < -\lambda^*(d)$ such that $Z_G(\lambda') = 0$.*

It is also known that the above theorem gives a full description of the zero-free region of Z_G, as G varies over $\mathcal{G}_{d+1}$, on the negative real line. The emphasis in the works leading to Theorem 1.1 was on obtaining zero-free regions shaped like disks (or like product of disks – *polydisks* – in the more general setting of the multivariate independence polynomial that we will not consider in this paper), and in the univariate setting, Theorem 1.1 essentially characterizes the radius of the largest such zero-free disk centered at the origin. Further, two different polynomial time approximation algorithms for $Z_G(\lambda)$ for $G \in \mathcal{G}_{d+1}$ and $|\lambda| < \lambda^*(d)$ were given by Patel and Regts [25] and Harvey, Srivastava and Vondrák [20].

Zero-free regions and algorithms. It is well known by now, however, that the actual zero-free region for Z_G as G varies over $\mathcal{G}_{d+1}$ is not described by a disk. It also turns out that the work towards characterizing this region is of importance for the algorithmic problem of approximating Z_G for an input graph $G \in \mathcal{G}_{d+1}$. In order to describe this connection, we first recall the work of Peters and Regts [26] on proving a conjecture of Sokal. There, they considered Z_G as G varies over spherically symmetric d-ary trees, and proved that it is non zero as long as $\lambda \in U_d$, where U_d is the open region (see fig. 1 for an example drawing of this curve) containing the origin bounded by the curve

$$(1.1) \qquad \partial U_d := \left\{ \kappa(\alpha) := \frac{-\alpha d^d}{(d+\alpha)^{d+1}} \,\middle|\, |\alpha| = 1 \right\}.$$

Using the results of Peters and Regts [26] on the existence of zeros near the boundary of ∂U_d, Bezáková, Galanis, Goldberg, and Štefankovič [11] showed that for every complex rational λ outside the closure of U_d that does *not* lie on the positive real line, the problem of approximating (up to any polynomial factor) $Z_G(\lambda)$ for graphs G in $\mathcal{G}_{d+1}$ is #P-hard.[1] On the other hand, due to the results of Barvinok [5] and Patel and Regts [25], the same problem admits a fully polynomial time approximation scheme (FPTAS) for any complex rational λ if for some $\epsilon > 0$, the ϵ-neighborhood of the line segment $[0, \lambda]$ is zero-free for the polynomials Z_G for all $G \in \mathcal{G}_{d+1}$.

Known results. In light of the above results, the problem of characterizing the location of the zeros of Z_G for general graph in $\mathcal{G}_{d+1}$ becomes of interest. Recall that U_d is the zero-free region for spherically symmetric d-ary trees. Perhaps the first natural question to ask is whether the region U_d is zero-free for Z_G even as G varies over *all* graphs in $\mathcal{G}_{d+1}$. The answer to this is no: Buys [12] showed that one can obtain a counterexample for $3 \leq d+1 \leq 9$ by considering spherically symmetric trees in which the arity of each vertex depends upon the distance from the root of the tree. Thus, the location of zeros of Z_G for G in $\mathcal{G}_{d+1}$ inside the region U_d needs to be studied more closely.

---

[1]In contrast, for positive real λ outside U_d, the same problem is NP-hard [18, 32], and is unlikely to be #P-hard for all such λ unless there is a collapse in the polynomial hierarchy. (The fact that approximate counting with positive weights cannot be #P-hard under standard complexity theoretic assumptions is a well-known direct consequence of Toda's theorem [34] and earlier results of Stockmeyer [33] and Sipser [30]; see, e.g., Ex. 17.5 in [1].)

In preparation for stating the contributions of this paper, we now turn to describing what is known about the zero-free region of Z_G for graphs in $\mathcal{G}_{d+1}$. Let $\lambda_c(d) := \frac{d^d}{(d-1)^{d+1}}$ be the unique point of intersection of the curve ∂U_d with the positive real line. (We note in passing that this quantity, known as the *uniqueness threshold* for the hard core model, has played a central role in the study of the algorithmic estimation of Z_G on the positive real line: in particular, this study led to some of the tightest known connections between statistical mechanics phase transitions and computational complexity [18, 31, 32, 35].) Peters and Regts [26] showed that for any positive $\lambda' < \lambda_c(d)$, there is an $\epsilon' = \epsilon'(\lambda') > 0$ such that for any z satisfying $|\Im z| \le \epsilon'$ and $\Re z = \lambda'$, $Z_G(z) \ne 0$ for all $G \in \mathcal{G}_{d+1}$. They also gave explicit lower bounds on $\epsilon'(\lambda')$ for $\lambda' \in (0, \tan(\pi/(2d)))$ (note that $\tan(\pi/(2d)) > \lambda^*(d)$ for $d \ge 2$, so these results are not implied by Theorem 1.1). Bencs and Csikvári [9], using different methods, improved on the latter lower bounds, and thereby significantly extended the known zero-free region inside U_d in the right half-plane. In Section 1.3 below, we describe some more recent papers that study phenomena such as the limit shape (after appropriate scaling) of the zero-free region as $d \uparrow \infty$, and that explore further connections between zero-freeness and other aspects of the independent set model. However, for specific finite d, none of these results seem to provide any new zero-free regions in the left half-plane beyond the half-disk implied by Theorem 1.1.

1.1 Contributions

In this paper, we give two geometric criterions (Theorems 4.5 and 4.6) which together give a framework for rigorously establishing (connected) zero-free regions as well as a way to carry out semi-rigorous numerical explorations.

We provide several examples of the (rigorous) use of these criterions. We establish two new zero-free regions in the left half plane: Theorem 6.1 gives a better result in the vicinity of the negative real line, while Theorem 7.1 gives a better result near the imaginary line. When restricted to the imaginary axis, the latter region agrees with the result of Bencs and Csikvári [9] for the right half-plane. We also extend the previous zero-freeness results of Bencs and Csikvári [9] for the right-half plane using the geometric criterions developed in this paper (Theorem 8.1, see also Remark 8.3). We also show that our framework gives a new proof of the Sokal conjecture, which was first proved by Peters and Regts via a potential function argument [26] (Theorem 5.1). See fig. 1 for a graphical illustration of these new zero-freeness results.

Algorithmic implications. Following the template provided by the results of Barvinok [5] and Patel and Regts [25], these new zero-freeness results also immediately lead to new polynomial-time algorithms for the approximation of $Z_G(\lambda)$ for λ lying in the interior of these regions on graphs $G \in \mathcal{G}_{d+1}$ of maximum degree at most $d + 1$. Given complex numbers Z and $\hat{Z}$ we say that $\hat{Z}$ is a multiplicative ε-approximation of Z if $e^{-\varepsilon} < \frac{|\hat{Z}|}{|Z|} < e^{-\varepsilon}$ and the angle between $\hat{Z}$ and Z considered as vectors in $\mathbb{C} = \mathbb{R}^2$ is at most ε. Then the algorithmic framework of Barvinok [5] and Patel and Regts [25] combined with our zero-free regions provides a deterministic algorithm of running time $\left(\frac{|V|}{\varepsilon}\right)^{O_{d,\lambda}(1)}$ for obtaining a multiplicative ε-approximation of $Z_G(\lambda)$ whenever $G = (V, E)$ has maximum degree $d + 1$ and λ is in the interior of the zero-free region provided by this paper.

Numerical explorations. We now comment briefly on the connections to numerical explorations – alluded to above – of our work. The naive method to numerically check whether a point λ is in the zero free region of Z_G for all graphs in $\mathcal{G}_{d+1}$ would be to evaluate $Z_G(\lambda)$ for all such graphs, and to check if it evaluates to 0. Known results allow one to restrict the set of graphs one has to explore to trees in $\mathcal{G}_{d+1}$ (see Theorem 3.2 below), but the resulting procedure is still computationally infeasible. In contrast, the geometric criterions in Theorems 4.5 and 4.6 allow one to do the following. Given d and λ, one tries to construct a curve in the complex plane with certain prescribed properties. The existence of such a curve then certifies that no graph in $\mathcal{G}_{d+1}$ has $Z_G(\lambda) = 0$. What curves would "work" for a given d and λ can then be explored numerically: in fact, many of our zero-freeness results listed above were obtained by first conjecturing the form of such a curve guided by numerical experiments, and then rigorously verifying – as done in the proofs of the theorems listed above – that the curve has the prescribed properties. We want to highlight, however, that although this method is "sound" – in the sense that producing such a curve as a certificate guarantees zero-freeness – it is not necessarily "complete" – one may not be able to construct such a curve as a certificate even though λ is in the zero-free region.

1.2 Organization of the paper

After a short section of preliminaries, we introduce in Section 3 various simple criterions to prove that a $\lambda \in \mathbb{C}$ is in the zero-free region of independence polynomial of graphs of bounded
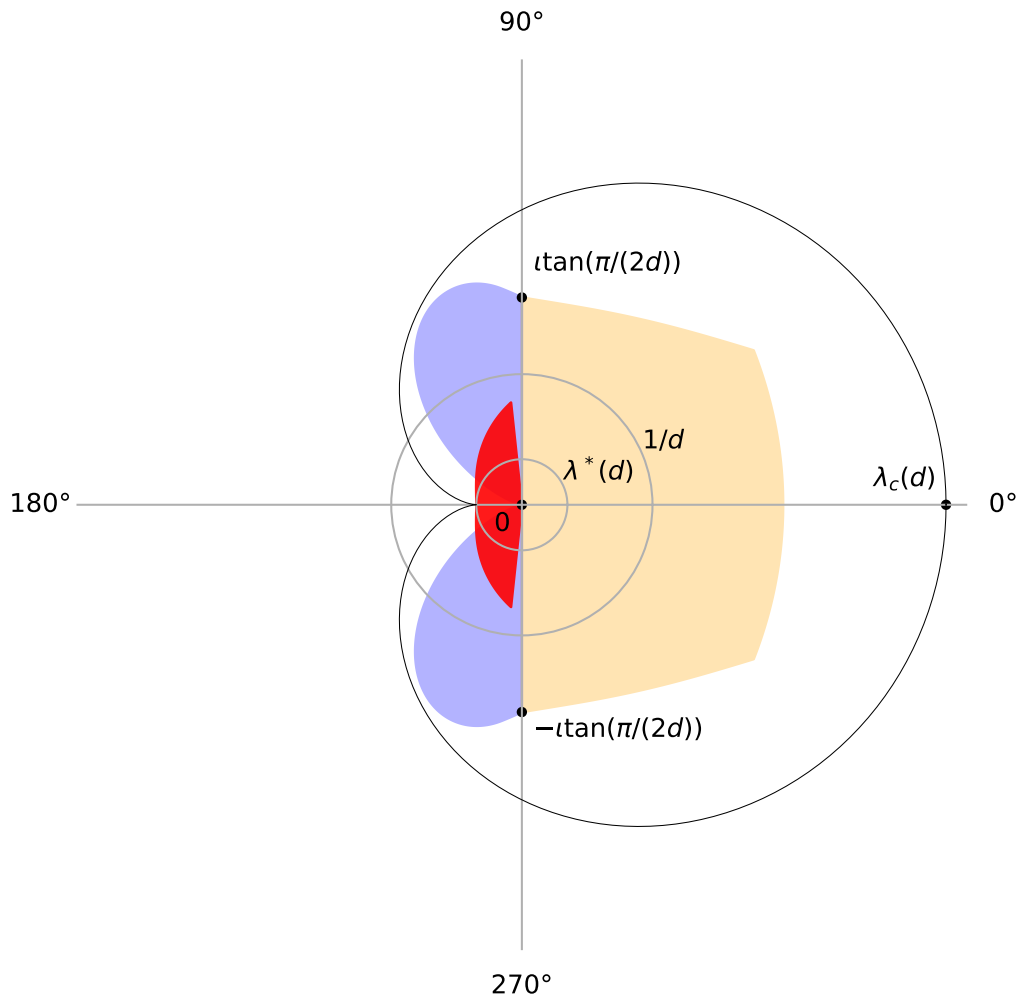
Figure 1: New zero free regions for $d = 9$ (graphs of degree at most 10). In the left half plane, the red region corresponds to Theorem 6.1, while the blue region corresponds to Theorem 7.1. In the right half plane, the yellow region corresponds to Theorem 8.1. The smaller grey circle around the origin has radius $\lambda^*(d)$ (the "Shearer radius" from Theorem 1.1), and the points $\pm \iota \tan(\pi/(2d))$ are marked on the imaginary axis. The outer black "cardioid-shaped" curve is the boundary ∂U_d as defined in eq. (1.1). A magnified version of the red region (corresponding to Theorem 6.1) is given in Figure 5.

degree. Building on this work, we introduce in Section 4 two new criterions (Theorems 4.5 and 4.6) which use constructions of certain curves in order to prove zero-free regions. The remaining sections are direct applications of these two criterions, and are independent of each other. Since some of the proofs are somewhat technical, these sections are arranged in increasing order of difficulty. In Section 5 we give a new proof of Sokal's conjecture originally proven by Peters and Regts. In Section 6 we give a new zero-free region in the vicinity of the critical point $\frac{d^d}{(d+1)^{d+1}}$. In Section 7 we provide a zero-free region close the imaginary axis. Finally, in Section 8 we prove a zero-free region in the right half plane.

For those readers who are interested in the ideas in general, but want to avoid technical difficulties we recommend reading the paper till the end of Section 5 and omitting Theorem 4.6 and its proof.

1.3 Related work The work of Barvinok [3] (see also [5]) pioneered the direct use of zero-free regions for designing algorithms for approximate counting. However, in most examples, a direct application of Barvinok's method gives a quasi-polynomial time algorithm: Patel and Regts [25] showed how to use various combinatorial tools in order to reduce this quasi-polynomial runtime to a polynomial runtime in various "bounded-degree"

settings. The method has since then been used to attack a wide variety of approximate counting problems: see, e.g., [3, 4, 6, 10, 16, 19, 23]. How this method relates to other methods of approximate counting, such as Markov chain Monte Carlo, or the method of reduction to tree-recurrences and "correlation decay" (first used by Bandyopadhyay and Gamarnik [2] and Weitz [35]), has also been explored in several papers, see, e.g., [21, 22, 24, 28]. In the statistical physics literature, very strong connections between Markov chain Monte Carlo and zero-freeness are known in the special case of integer lattices through the work of Dobrushin and Shlosman [14, 15].

As already discussed in detail in the previous subsections, the complex zeros of the independence polynomial have also been studied extensively in the context of its connections to the Lovász local lemma and also in the context of its computational complexity [11, 20, 25, 27, 29]. Here we describe a few more recent works in this direction. Recent work of de Boer, Buys, Guerini, Peters, and Regts [13] establishes strong formal connections between the computational complexity of the hard core model, complex dynamics, and zero-freeness of the partition function (see Main Theorem of [13]): in particular they prove that the zeros of Z_G for graphs in $\mathcal{G}_{d+1}$ are dense in the complement of U_d. The limit shapes of the zero-free regions have also been studied: Bencs, Buys, and Peters [8] show that in the $d \to \infty$ limit, a rescaled version of zero-free region tends to a bounded 0-star shaped region, whose boundary intersects $\lim_{d \to \infty} d \cdot \partial U_d$ only at real parameters. In particular, the results of [8] show that for large enough d the zero-free region is strictly contained in U_d except for the two real parameters. The problem of fully characterizing the zero-free region of Z_G for $G \in \mathcal{G}_{d+1}$ in the complex plane, however, still remains open.

2 Preliminaries

Branch cuts We adopt the following convention for defining fractional powers and complex logarithms. Given $z = re^{\iota\theta}$ with $r > 0$ and $\theta \in (-\pi, \pi]$, we define

$$(2.1) \qquad \log z := \log r + \iota\theta, \text{ and}$$
$$(2.2) \qquad z^\delta := r^\delta \exp(\iota\delta\theta), \text{ for any } \delta > 0.$$

We leave the functions undefined when $z = 0$ (except that we adopt the usual convention that $0^0 = 1$). Note that with the above definition, log and z^δ for non-integral δ are defined but discontinuous on the negative real line. However, we do have the following identity for all $z \neq 0$ and $\delta \geq 0$:

$$z^\delta = \exp(\delta \log z).$$

Further, for $z \neq 0$, we use the convention $\arg z = \Im(\log z)$.

Graphs and independence polynomials For the sake of providing a quick reference, we recollect here some basic notation and terminology about graphs and their independence polynomials that was introduced in the introduction above. We denote the set of all graphs of degree at most $d + 1$ by $\mathcal{G}_{d+1}$. The *independence polynomial* $Z_G(\lambda)$ of a graph G is given by

$$(2.3) \qquad Z_G(\lambda) := \sum_{I:\text{independent set in } G} \lambda^{|I|}.$$

Two quantities of interest with respect to the independence polynomial are the *Shearer radius* $\lambda^*(d) := \frac{d^d}{(d+1)^{d+1}}$, and the *uniqueness threshold* $\lambda_c(d) := \frac{d^d}{(d-1)^{d+1}}$. The former, $\lambda^*(d)$, is specially connected to the Lovász local lemma, and also the radius of the largest circular disk around the origin in which Z_G is zero-free for all graphs in $\mathcal{G}_{d+1}$ (see Theorem 1.1 above). The latter, $\lambda_c(d)$, is intimately connected to the complexity of approximating $Z_G(\lambda)$ for $G \in \mathcal{G}_{d+1}$, for λ on the positive real line: in particular, Weitz [35] gave a deterministic fully polynomial approximation scheme (FPTAS) for $Z_G(\lambda)$ for $G \in \mathcal{G}_{d+1}$, provided $\lambda < \lambda_c(d)$, while in a series of works [17, 18, 31, 32] starting with a paper of Sly, it was shown that a randomized fully polynomial approximation scheme (FPRAS) for the same problem in the regime $\lambda > \lambda_c(d)$ would imply NP = RP. Further, in terms of the curve ∂U_d of Peters and Regts [26] (see eq. (1.1) above), $\lambda_c(d)$ is the unique point of intersection of ∂U_d with the positive real line, while $-\lambda^*(d)$ is the unique point of intersection of ∂U_d with the negative real line.

3 A relaxed recurrence for the independence polynomial

As stated in the introduction, our goal is to study the location of the complex zeros of the (univariate) independence polynomial

$$Z_G(\lambda) := \sum_{I \subseteq V(G) \text{ independent}} \lambda^{|I|}.$$

Recall that we focus on the class of graphs $\mathcal{G}_\Delta$ with degrees at most Δ for some fixed $\Delta \geq 3$. It is more convenient, however, to work in terms of the notation $d := \Delta - 1$. We now proceed to describe a known characterization of zero-free regions for the independence polynomial of bounded degree graphs, in preparation for which we introduce the following definition.

DEFINITION 3.1 (**The set** $S_\lambda = S_\lambda(d)$). For $\lambda \in \mathbb{C}$, define $S_\lambda \subseteq \mathbb{C}$ as the set of points that can be generated by the following rules:

- $0 \in S_\lambda(d)$,

- If $z_1, \ldots, z_d \in S_\lambda(d)$ are such that $z_i \neq -1$ for $1 \leq i \leq d$, then

  $$(3.1) \qquad\qquad f(z_1, \ldots, z_d) = \frac{\lambda}{\prod_{i=1}^{d}(1 + z_i)}$$

  is also in $S_\lambda(d)$.

(Although the definition of S_λ depends on d, we will often omit this dependence from our notation when the value of d is clear from the context.)

The following theorem is well known [27,35], and has been used in previous work on the subject (e.g. in [9,26]). It can most directly be obtained from a result of Bencs [7], who showed that the independence polynomial of a graph divides (as a polynomial) the independence polynomial of the so-called "self-avoiding walk tree" of the graph. The zero-free regions of the independence polynomial of a tree can in turn be analyzed in terms of the "tree recurrences" described in eq. (3.1) [27,35].

THEOREM 3.2 (**see, e.g., Proposition 2.7 (1) of [7], and Lemma 2.1 of [9]**). *Fix $d \geq 2$. $Z_G(\lambda) = 0$ for some graph $G \in \mathcal{G}_{d+1}$ if and only if $-1 \in S_\lambda(d)$.*

A standard application of Theorem 3.2 is to define a "trapping region" T such that $0 \in T$, $-1 \notin T$ and f maps T to T. For instance, if $|\lambda| \leq \frac{d^d}{(d+1)^{d+1}}$, then $T = \{z \in \mathbb{C} \mid |z| \leq \frac{1}{d+1}\}$ is such a region:

$$\left| \frac{\lambda}{\prod_{i=1}^{d}(1+z_i)} \right| \leq \frac{d^d}{(d+1)^{d+1}} \prod_{i=1}^{d} \frac{1}{1 - \frac{1}{d+1}} = \frac{1}{d+1}$$

showing Shearer's result. In general, it is not easy to handle d variables at the same time. Therefore, in what follows we try to find sufficient conditions that only require understanding the behaviour of a univariate map.

In the following, we relax the recurrence $f(z_1, \ldots, z_d)$ to allow fractional powers and more than d arguments. As we will see, this in fact leads to a simplification of the problem.

DEFINITION 3.3 (**The set** $\tilde{S}_\lambda = \tilde{S}_\lambda(d)$). For $\lambda \in \mathbb{C}$, define $\tilde{S}_\lambda \subseteq \mathbb{C}$ as the set of points that can be generated by the following rules:

- $0 \in \tilde{S}_\lambda(d)$,

- If $z_1, \ldots, z_k \in \tilde{S}_\lambda(d)$ and $\delta_1, \ldots, \delta_k \geq 0$ are such that $\sum_{i=1}^{k} \delta_i \leq d$ and $z_i \neq -1$ for $1 \leq i \leq k$, then

  $$(3.2) \qquad\qquad f_{\delta_1, \ldots, \delta_k}(z_1, \ldots, z_k) = \frac{\lambda}{\prod_{i=1}^{k}(1 + z_i)^{\delta_i}}$$

  is also in $\tilde{S}_\lambda(d)$.

(As with S_λ, although the definition of S_λ depends on d, we will often omit this dependence from our notation when the value of d is clear from the context.)

Clearly, we have $S_\lambda(d) \subseteq \tilde{S}_\lambda(d)$, since the new generation rule subsumes $f(z_1, \ldots, z_d)$. Hence, from Theorem 3.2, we directly obtain the following.

LEMMA 3.4. *Fix $d \geq 2$. If $-1 \notin \tilde{S}_\lambda(d)$, then $Z_G(\lambda) \neq 0$ for every $G \in \mathcal{G}_{d+1}$.*

The main advantage of the relaxed recurrence is that it allows us to replace the multivariate recurrence by a univariate one. We do this as follows: Consider the set $\{\log(1+z) : z \in \tilde{S}_\lambda\}$. Note that this is well defined if $-1 \notin \tilde{S}_\lambda$. If we write $w_i = \log(1 + z_i)$, then the recurrence

$$f_{\delta_1,\ldots,\delta_k}(z_1, \ldots, z_k) = \frac{\lambda}{\prod_{i=1}^k (1+z_i)^{\delta_i}}$$

can be rewritten by substitution as

$$g_{\delta_1,\ldots,\delta_k}(w_1, \ldots, w_k) = \log(1 + f_{\delta_1,\ldots,\delta_k}(e^{w_1} - 1, \ldots, e^{w_k} - 1))$$

$$= \log\left(1 + \lambda \prod_{i=1}^k e^{-\delta_i w_i}\right) = \log\left(1 + \lambda e^{-\sum_{i=1}^k \delta_i w_i}\right).$$

Hence, a combination of fractional powers in $f_{\delta_1,\ldots,\delta_k}$ corresponds to a linear combination of the points $w_i = \log(1 + z_i)$. If we normalize the linear combination by $\frac{1}{d}$, and use the fact that 0 is always a possible choice for w_i, we obtain a convex linear combination of $w_1, \ldots, w_k$ in the exponent. This motivates the following characterization. Note that the characterization is in terms of the behavior of a function of only one complex variable.

THEOREM 3.5. *Fix $d \geq 2$. The number -1 is not contained in $\tilde{S}_\lambda(d)$ if and only if there is a convex set $T \subset \mathbb{C}$ containing 0 such that for every $w \in T$,*
$$g(w) = \log(1 + \lambda e^{-dw})$$
is well-defined and $g(w) \in T$.

Proof. Suppose first that $-1 \notin \tilde{S}_\lambda = \tilde{S}_\lambda(d)$. We define

$$T = \text{conv}\left\{\log(1+z) | z \in \tilde{S}_\lambda\right\}.$$

Note that since $-1 \notin \tilde{S}_\lambda$, T is well-defined, and further, is convex by definition. Also, $0 \in T$, since $0 \in \tilde{S}_\lambda$. Now consider $w \in T$. By Caratheodory's theorem, there exist $\delta_1, \delta_2, \delta_3 \geq 0$ summing up to d, and $z_1, z_2, z_3 \in \tilde{S}_\lambda$, such that $w = \frac{1}{d}\sum_{i=1}^3 \delta_i \log(1 + z_i)$. We thus have $\lambda \exp(-dw) = f_{\delta_1,\delta_2,\delta_3}(z_1, z_2, z_3) \in \tilde{S}_\lambda$. Thus, $\lambda \exp(-dw) \neq -1$ and hence $g(w) = \log(1 + \lambda \exp(-dw)) = \log(1 + f_{\delta_1,\delta_2,\delta_3}(z_1, z_2, z_3))$ is well-defined and lies in T.

Conversely, suppose that T is any arbitrary convex set containing 0, on which the map $g(w) = \log(1 + \lambda \exp(-dw))$ is well defined, and satisfies $g(w) \in T$ for all $w \in T$. We claim that if $-1 \in \tilde{S}_\lambda$, then there exists $w \in T$ such that $-1 = \lambda \exp(-dw)$.

To see this, define the *depth* of every $z \in \tilde{S}_\lambda$ as follows: $\text{depth}(0) = 0$, and for $z \neq 0$, $\text{depth}(z)$ is the smallest integer D such that z can be written as $f_{\delta_1,\delta_2,\ldots,\delta_k}(z_1, z_2, \ldots, z_k)$ where k is a positive integer, $\delta_i \geq 0$ sum to at most d, and $z_i \in \tilde{S}_\lambda$ have depth at most $D-1$. Note that $\text{depth}(z) \geq 1$ for $z \neq 0$. Now, if $-1 \in \tilde{S}(\lambda)$, let $D_{-1} = \text{depth}(-1)$.

We claim now that for all $z \in \tilde{S}_\lambda$ of depth at most $D_{-1} - 1$, $\log(1+z) \in T$. This is proved by induction on the depth of z: it is true in the base case $\text{depth}(z) = 0$ (so that $z = 0$), since $0 \in T$. Otherwise, from the definition of depth, we can find $z_1, z_2, \ldots z_k$ of depth strictly smaller than z, and $\delta_i \geq 0$ summing up to at most d, such that

$$(3.3) \qquad z = f_{\delta_1,\delta_2,\ldots,\delta_k}(z_1, z_2, \ldots, z_k) = \lambda \exp\left(-d\sum_{i=1}^k \frac{\delta_i}{d} \log(1 + z_i)\right).$$

Thus, we have $\log(1+z) = g(w)$ where w is a convex combination of 0 and the quantities $\log(1 + z_i)$. The latter quantities are all inductively in T, so that w is also in T (as T is convex). But since T is closed under applications

of g, this implies that $g(w) = \log(1+z)$ is also in T. This establishes the claim that for every $z \in \tilde{S}_\lambda$ of depth at most $D_1 - 1$, $\log(1+z)$ is an element of T.

Now, applying the argument leading to eq. (3.3) with $z = -1$ (which by assumption has depth D_{-1}), we conclude that there exists a $w \in T$ such that $-1 = \lambda \exp(-dw)$. But this contradicts the hypothesis that $g(w) = \log(1 + \lambda \exp(-dw))$ is well-defined on T. Thus, it cannot be the case that $-1 \in \tilde{S}_\lambda$. $\qquad\square$

For natural reasons, we call a T as in the statement of the above theorem a *trapping region* for λ.

Remark: Sometimes it is desirable to avoid -1 even in the closure of S_λ, or $\tilde{S}_\lambda$. By our transformation, this corresponds to the property that there is a convex set T containing 0 and closed under $g(w) = \log(1 + \lambda e^{-dw})$, such that $\Re(w) \geq -K$ for every $w \in T$ and some constant $K > 0$. This is equivalent to saying that every point $z \in \tilde{S}_\lambda$ satisfies $|1+z| \geq e^{-K}$.

4 Criterion in the original complex plane

The previous section shows that we get a rather clean picture when we study the behavior of the extended recurrence (eq. (3.2)) after a change of variable, $w = \log(1+z)$. However, we can also formulate a criterion using trapping regions in the original variable z. This criterion looks more intuitive, but it seems we lose a bit in the transition (in particular, we do not get an equivalence here).

THEOREM 4.1. *If there is a convex set $S \subset \mathbb{C}$ containing 0, not containing -1, such that $f(z) = \frac{\lambda}{(1+z)^d} \in S$ for every $z \in S$, then $Z_G(\lambda) \neq 0$ for every $G \in \mathcal{G}_\Delta$.*

In order to prove this statement, we need the following fact about the behavior of arithmetic vs. geometric averages in the complex plane. While we believe this fact to be standard, we are unable to find an exact reference, and hence provide a proof for completeness.[2]

LEMMA 4.2 ("geometric averages dominate arithmetic averages"). *For any two points $z_1, z_2 \in \mathbb{C} \setminus \{0\}$ satisfying $|\arg(z_1) - \arg(z_2)| \leq \pi$, and $\alpha \in [0,1]$, there exist $\beta \in [0,1]$ and $t \in [0,1]$ such that*

$$t z_1^\alpha z_2^{1-\alpha} = \beta z_1 + (1-\beta) z_2.$$

Proof. We reduce to the case where $z_2 = 1$, by dividing by z_2 and substituting $z = z_1/z_2$. Our goal then is to find $t, \beta \in [0,1]$ such that

$$t z^\alpha = \beta z + (1-\beta).$$

We can also assume that $\arg z \in [0,\pi]$, by complex conjugation if this is not the case.

Now, if $\arg z = 0$ the claim is trivially true because, then, if $z = \Re z \geq 1$, we can take $\beta = 0, t = z^{-\alpha}$, while when $0 < z = \Re z < 1$, we can take $\beta = 1, t = z^{1-\alpha}$. Similarly, when $\arg z = \pi$, z is a strictly negative real number, so that we can choose $t = 0$ and $\beta = \frac{1}{1-z} \in [0,1]$.

We can thus assume that $\theta := \arg(z) \in (0,\pi)$ and $r := |z| > 0$. Note that $\arg(z^\alpha) = \alpha\theta$. Let $y(\alpha)$ be the unique point with argument $\alpha\theta$ on the line segment joining 1 and z. From an elementary geometric argument, we then have

$$|y(\alpha)| = \frac{r \sin \theta}{r \sin((1-\alpha)\theta) + \sin \alpha\theta}.$$

We now define the function $f(x) : [0,1] \to \mathbb{R}$ as

$$f(x) := \log \frac{|y(x)|}{|z^x|} = \log \frac{r^{1-x} \sin \theta}{r \sin((1-x)\theta) + \sin x\theta}.$$

Note that the claim of the lemma is equivalent to showing that $f(x) \leq 0$ for all $x \in [0,1]$ (the quantity t can then be taken to $e^{f(\alpha)} \in [0,1]$ and $\beta \in [0,1]$ is such that $\beta z + 1 - \beta = y(\alpha)$).

---

[2]Note that, despite the title, the lemma does not contradict the usual inequality between the arithmetic and the geometric means of positive reals. The lemma is in fact a trivial statement for the case of positive reals.

To this end, we first note that $f(0) = f(1) = 0$, so the claim would follow if f is convex on $[0, 1]$. We verify this by directly computing the second derivative of f and checking that it is non-negative in $[0, 1]$:

$$f''(x) = \frac{\theta^2(1 + r^2 - 2r\cos\theta)}{(r\sin((1-x)\theta) + \sin x\theta)^2} \geq 0, \text{ when } x \in [0, 1]. \qquad \square$$

Proof of Theorem 4.1. Suppose that there is a convex set S containing 0, not containing -1, and closed under the map $f(z) = \frac{\lambda}{(1+z)^d}$. We will transform S into a convex set T satisfying the assumptions of Theorem 3.5. Define

$$T = \text{conv}\{\log(1 + z) : z \in S\}.$$

By construction, T is convex and it contains 0 (since $0 \in S$). We need to prove that $g(w) = \log(1 + \lambda e^{-dw})$ is well-defined on T and preserves membership in T. Consider $w \in T$, i.e., $w = \sum_{i=1}^{k} \alpha_i \log(1 + z_i)$, a convex combination of points $\log(1 + z_i)$ such that $z_i \in S$. We can set $\delta_i = d\alpha_i$, hence $w = \frac{1}{d}\sum_{i=1}^{k} \delta_i \log(1 + z_i)$. Then,

$$g(w) = \log(1 + \lambda e^{-dw}) = \log\left(1 + \frac{\lambda}{\prod_{i=1}^{k}(1 + z_i)^{\delta_i}}\right).$$

Now, we appeal to Lemma 4.2. Note first that $1 + S := \{1 + z | z \in S\}$ is a convex set containing 1 and not containing 0, so that by the separating hyperplane theorem, all of $1 + S$ lies in a halfplane defined by a line passing through 0, and therefore $|\arg(1 + u) - \arg(1 + v)| \leq \pi$ is true for all $u, v \in S$. Now, we take one pair z_i, z_j of points at a time, and consider $u_{ij} = (1 + z_i)^{\frac{\delta_i}{\delta_i + \delta_j}}(1 + z_j)^{\frac{\delta_j}{\delta_i + \delta_j}}$. By Lemma 4.2, there is $\beta \in [0, 1]$ such that $1 + z_{ij} = \beta(1 + z_i) + (1 - \beta)(1 + z_j)$ is a point of the same argument and smaller-or-equal modulus as u_{ij}. Hence we can replace both z_i and z_j by z_{ij} and continue. We maintain the property that the argument of $\prod(1 + z_i)^{\delta_i}$ remains preserved and the modulus can only decrease. Eventually, we obtain a point $\tilde{z} \in \text{conv}\{z_1, \ldots, z_k\} \subseteq S$ such that $\arg((1 + \tilde{z})^d) = \arg(\prod(1 + z_i))$ and $|(1 + \tilde{z})^d| \leq |\prod(1 + z_i)|$. Hence, we can write

$$g(w) = \log\left(1 + \frac{c \cdot \lambda}{(1 + \tilde{z})^d}\right),$$

where $c \leq 1$ is a non-negative real number. Now, note that since $\tilde{z} \in S$, we have $f(\tilde{z}) = \frac{\lambda}{(1+\tilde{z})^d} \in S$, as S is closed under applications of f. Then, since $0 \in S$, and S is convex, we get $y := \frac{c \cdot \lambda}{(1+\tilde{z})^d} \in S$, and further that $y \neq -1$, as $-1 \notin S$. Thus, by definition of T, $g(w) = \log(1 + y) \in T$, as $y \in S$. This proves that T satisfies the assumptions of Theorem 3.5 and hence -1 is not contained in $\tilde{S}_\lambda$, which implies that $Z_G(\lambda) \neq 0$. $\qquad\square$

Next, we present a more abstract extended version of this criterion, where we allow a "convex" initial segment $h(t), t \in [0, 1]$ rather than a line segment. First we define the following notion:

DEFINITION 4.3 (-1-**covered points**). *A point $z \in \mathbb{C}$ is -1-covered by $z' \in \mathbb{C}$ if $\arg(1 + z) = \arg(1 + z')$ and $|1 + z| \geq |1 + z'|$. More generally, a set T is -1 covered by a set S if for every $z \in T$, there is a point $z' \in S$ such that z is -1-covered by z'.*

Geometrically, the above notion captures z being "covered" by z' when "viewed" from the point -1. The utility of this definition for our purposes comes from the following simple observation.

OBSERVATION 4.4. *Fix an integer $d \geq 2$ and a $\lambda \in \mathbb{C}$, and consider $f(z) := \frac{\lambda}{(1+z)^d}$. If $z \in \mathbb{C}$ is -1-covered by $w \in \mathbb{C}$, then $f(z) = \alpha f(w)$ for some $\alpha \in [0, 1]$.*

Proof. If $\lambda = 0$, there is nothing to prove, so assume $\lambda \neq 0$. Since z is -1-covered by w, we have $|1 + z| \geq |1 + w|$ and $\arg(1 + w) = \arg(1 + z)$. It follows that $\arg(f(z)) = \arg(f(w))$, and $|f(z)| \leq |f(w)|$. Thus, $f(z)$ lies on the segment joining the origin to $f(w)$, and the claim follows. $\qquad\square$

Next we state and prove our first main geometric criterion for zero-freeness, which will be applied multiple times in the subsequent sections.

THEOREM 4.5. *For $\lambda \in \mathbb{C}$, assume that there is a curve $\{h(t) : t \in [a, b]\}$, where $a < b$ are real numbers, such that*

- $h(t) = 0$ *for some $t \in [a, b]$,*

- $\arg(1 + h(t))$ *is strictly increasing for $t \in [a, b)$,*

- $h(t)$ *is "convex" in the sense that for any $t_1, t_2 \in [a, b], \alpha \in [0, 1]$, $\alpha h(t_1) + (1 - \alpha)h(t_2)$ is -1-covered by $h(t)$ for some $t \in [t_1, t_2]$.*

- *for every $t \in [a, b]$, $f(h(t)) := \frac{\lambda}{(1+h(t))^d}$ is -1-covered by $h(t')$ for some $t' \in [a, b]$.*

Then $Z_G(\lambda) \neq 0$ for any $G \in \mathcal{G}_{d+1}$.

Proof. By an affine reparameterization of the curve h, if necessary, we assume that $a = 0$ and $b = 1$. Given the curve $h(t), t \in [0, 1]$, we define a trapping region, in the sense of Theorem 4.1, as a "shadow of the curve $h(t)$ when illuminated from the point -1":

$$S = \{z \in \mathbb{C} : \exists t \in [0, 1] \text{ such that } z \text{ is } -1\text{-covered by } h(t)\}.$$

This is a convex set, since for any $z_1, z_2 \in S$, z_1 and z_2 are -1-covered by $h(t_1), h(t_2)$ respectively, $z = \alpha z_1 + (1 - \alpha)z_2$ is -1-covered by $z' = \alpha' h(t_1) + (1 - \alpha')h(t_2)$ (for some $\alpha' \in [0, 1]$), and z' in turn is covered by $h(t)$ for some $t \in [t_1, t_2]$ by the convexity of $h(t)$.

Also, 0 is contained in S because $h(t) = 0$ for some $t \in [0, 1]$; -1 is not contained in S, since 0 is the only real value on the curve $h(t)$ (this follows since $\arg(1 + h(t))$ is assumed to be a strictly increasing function of t), so that S contains only non-negative real numbers.

To apply Theorem 4.1 in order to conclude the proof, it remains to prove that S is closed under the map $f(z) = \frac{\lambda}{(1+z)^d}$. For any $z \in S$, there is a $t \in [0, 1]$ such that z is -1-covered by $h(t)$. Observation 4.4 then implies that $f(z) = \alpha f(h(t))$ for some $\alpha \in [0, 1]$. By the assumptions of the theorem, we know that $f(h(t))$ is -1-covered by some point $h(t')$, $t' \in [0, 1]$, which implies that $f(h(t)) \in S$. Hence by convexity, since $0 \in S$, we also get $f(z) = \alpha f(h(t)) \in S$, as required. $\qquad\square$

Next, we formulate a more concrete sufficient condition which can be used in numerical experiments.[3] The proof we give here highlights the connections of this result with numerical exploration, even though alternative proofs may be possible (see Remarks 4.7 and 4.8 following the proof).

THEOREM 4.6. *For $\lambda \in \mathbb{C}$, $\Im(\lambda) > 0$, define a curve*

- $h(t) = t\lambda$ *for $t \in [0, 1]$,*

- $h(t) = \frac{\lambda}{(1+h(t-1))^d}$ *for $t > 1$.*

If $\Im(h(t)) \geq 0$ for all $t \geq 0$, then $Z_G(\lambda) \neq 0$ for all $G \in \mathcal{G}_{d+1}$. (See also Figure 2 for an example of the curve h in the statement of the theorem.)

Proof. Note that the curve $h(t)$ is continuous since it is continuous at $t \leq 1$: this is because at $t = 1$ we have $\lim_{t \to 1-} h(t) = \lim_{t \to 1+} h(t) = \lambda$, and by the recursion h is continuous at t if it is continuous at $t - 1$. By continuity and the assumption $\Im(h(t)) \geq 0$ for all $t \geq 0$ this implies that the continuous functions $\arg(h(t))$ and $\arg(1 + h(t))$ are also non-negative for $t \geq 0$. The identity $\arg(h(t)) = \arg(\lambda) - d\arg(1 + h(t-1))$ for $t \geq 1$ (valid whenever the right hand side lies in $(-\pi, \pi]$) then implies that $\arg(1 + h(t - 1))$ cannot exceed $\frac{1}{d}\arg(\lambda)$, because if t is the infimum of points for which $\arg(1 + h(t - 1)) > \frac{1}{d}\arg(\lambda)$ then the identity gives a contradiction to $\arg(h(t + \epsilon)) \geq 0$ for some small enough positive ϵ. Since $\arg(1 + h(t - 1)) \geq 0$, the identity then also implies that

$$(4.1) \qquad 0 \leq \arg(h(t)) = \arg(\lambda) - d\arg(1 + h(t-1)) \leq \arg(\lambda).$$

Hence the argument of any point of the curve is contained in $[0, \arg(\lambda)]$.

---

[3]This was the result stated in a talk at the Simons Institute for the Theory of Computing, UC Berkeley, on March 18, 2019.
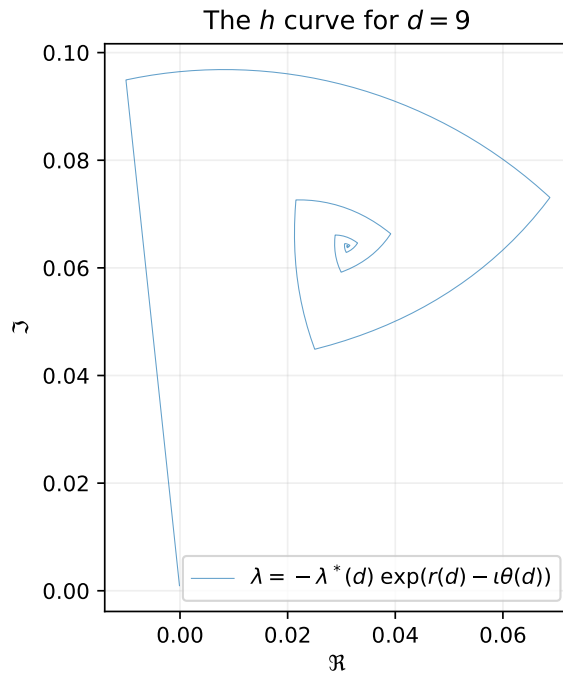
Figure 2: The h curve from Theorem 4.6. We set $d = 9$, and $\lambda^*(d)$ is the Shearer threshold described in Theorem 1.1. $\theta(d)$ is chosen to be $\cos^{-1} \frac{1}{d+0.5}$, while $r(d) := \min \left\{ d \log(1 + \frac{1}{d}), \frac{2d(d+1)\sin^2(\theta/2)}{d^2+4(d+1)\sin^2(\theta/2)} \right\}$. The choice of these parameters is based on their use in the later Theorem 6.1, which applies Theorem 4.6.

In order to define a trapping region, we start by defining the following quantity:

$$(4.2) \qquad \tau^\star := \sup\{t' : \arg(1 + h(t)) \text{ is non-decreasing for all } t \in [0, t']\}.$$

Note that $\tau^\star \geq 1$. We also allow $\tau^\star = \infty$, although this cannot really happen. We will now show that the region S in the upper half plane bounded by the line segments $[0, h(1)]$ and $[0, h(\tau^\star + 1)]$ and the curve $\{h(t+1) | 0 \leq t \leq \tau^\star\}$ is a trapping region in the sense of Theorem 4.1. Note that by definition $0 \in S$, while $-1 \notin S$ (since, from the observations above, $\arg z \in [0, \arg \lambda]$ for all $z \in S$). It remains to show that (1) S is convex, and (2) $f(z) = \lambda/(1+z)^d \in S$ for all z in S.

We start by proving that S is convex. Since the curve h lies in the upper half plane (see eq. (4.1)), this will follow if we establish the following two facts:

1. $\arg(h(t))$ is non-increasing for $t \in [1, \tau^\star + 1]$.

2. The curve $\{h(t) : t \in [1, \tau^\star + 1]\}$ is "turning to the right". More formally, for any $t \in [1, \tau^\star + 1)$, there is a small enough neighborhood N_t of t such that for $t_1 \leq t_2$ in N_t, $\arg(D^-(h)(t_1)/D^+(h)(t_2)) \geq 0$. Here D^+ and D^- denote the right and left one-sided derivatives.

We first prove item 1. This follows since in the interval $t \in [1, \tau^\star + 1]$, we have $\arg(h(t)) = \arg(\lambda) - d \arg(1 + h(t-1))$, which is non-increasing by the definition of $\tau^\star$.

We now consider item 2. Note that $h(t)$ is continuously differentiable in the neighborhood of any t which is not an integer. Further, for such a t, we have $h'(t) = \frac{-\lambda d}{(1+h(t-1))^{d+1}} h'(t-1)$. We now prove the claim for such t (i.e., non-integral t) using an induction on $\lceil t \rceil$. In the base case, when $\lceil t \rceil = 1$, we have $h'(t) = \lambda$, so $\arg(D^-(h)(t_1)/D^+(h)(t_2)) = \arg(1) = 0$ for t_1, t_2 in any small enough neighborhood of t. In the inductive case, we have, for $t_1 \leq t_2$ in a small enough neighborhood of t,

$$(4.3) \qquad \arg \frac{D^-(h)(t_1)}{D^+(h)(t_2)} = (d+1) \arg \frac{1 + h(t_2 - 1)}{1 + h(t_1 - 1)} + \arg \frac{D^-(h)(t_1 - 1)}{D^+(h)(t_2 - 1)}.$$

The claim now follows since the first term is non-negative due to the definition of $\tau^\star$, while the second is non-negative by the inductive hypothesis.

We now consider the case of integral t. Here, we find via a direct induction that

$$(4.4) \qquad \arg \frac{D^-(h)(t)}{D^+(h)(t)} = \arg \frac{D^-(h)(1)}{D^+(h)(1)} = \arg \frac{\lambda}{(-d\lambda^2)} = \pi - \arg \lambda \geq 0.$$

The proof for item 2 now follows from the already proved case of non-integral t and the fact that the derivative h' is a well-defined continuous function except at integral t. As noted earlier, this proves that S is convex. In fact, from the definition of $\tau^\star$, we also obtain that $0 = \arg(1 + h(0)) \leq \arg(1 + z) \leq \arg(1 + h(\tau^\star))$ for all $z \in S$. Since $\arg(1 + h(t))$ is non-decreasing for $t \in [0, \tau^\star]$ and S is convex, it follows that if a line is drawn from $z \in S$ in the direction of -1, it will intersect the boundary of S at some point $h(t)$ for $0 \leq t \leq \tau^\star$.

We can now prove that S satisfies the remaining requirement for being a trapping region, which is, that it is closed under application of f. Consider any point $z \in S$. As noted above, if a line is drawn from z towards -1, then it must intersect the boundary of S on a point $\tilde{z}$ of the form $h(t)$ for $t \in [0, \tau^\star]$. Hence, there is a point $\tilde{z} = h(t), t \in [0, \tau^\star]$ such that $\arg(1 + h(t)) = \arg(1 + z)$ and $|1 + h(t)| \leq |1 + z|$. By construction,

$$(4.5) \qquad f(h(t)) = \frac{\lambda}{(1 + h(t))^d} = h(t + 1)$$

which is still in S (since $t \in [0, \tau^\star]$). Finally, $f(z) = \frac{\lambda}{(1+z)^d}$ has the same argument as $h(t + 1)$, and possibly smaller modulus, hence $f(z) \in S$ by convexity (since $0 \in S$).

Thus, S as defined above is a trapping region, and this concludes the proof. $\qquad\square$

REMARK 4.7. *We note that the proof of Theorem 4.6 also indicates a numerical approach to check this criterion, see Figure 3. We do not have to track the curve for $t \to \infty$. It is sufficient to compute $h(t)$ for $t \in [0, \tau^\star + 1]$ as defined above. If $\arg(1 + h(\tau^\star)) \leq \frac{1}{d} \arg(\lambda)$ and $\arg(1 + h(t))$ is non-increasing for $t \in [\tau^\star, \tau^\star + 1]$, the argument above implies that S is a trapping region and the entire curve is contained in the upper half-plane.*

REMARK 4.8. *We also remark that it is possible to prove Theorem 4.6 from Theorem 4.5 by considering the curve from the proof of Theorem 4.6 on the interval $[0, \tau^\star]$. The curve in Theorem 4.5 represents the portion of the curve in Theorem 4.6 "visible from -1"; i.e., the points $h(t)$ that are not -1-covered by any other point $h(t')$.*

5 Derivation of the Sokal conjecture

Here we provide a short proof using Theorem 4.5 that there are no complex roots close to the positive real axis, up to the critical point $\lambda^* = \frac{d^d}{(d-1)^{d+1}}$. This was first proved by Peters and Regts [26].

THEOREM 5.1. *For every fixed $d \geq 2$, and every $\epsilon \in (0, 1)$, there is an $\epsilon' > 0$ such that $Z_G(\lambda) \neq 0$ for G of maximum degree at most $d + 1$ when $\lambda = (1 - \epsilon)\frac{(d-\epsilon)^d}{(d-1)^{d+1}} \exp(\iota\theta)$ with $|\theta| \leq \epsilon'$.*
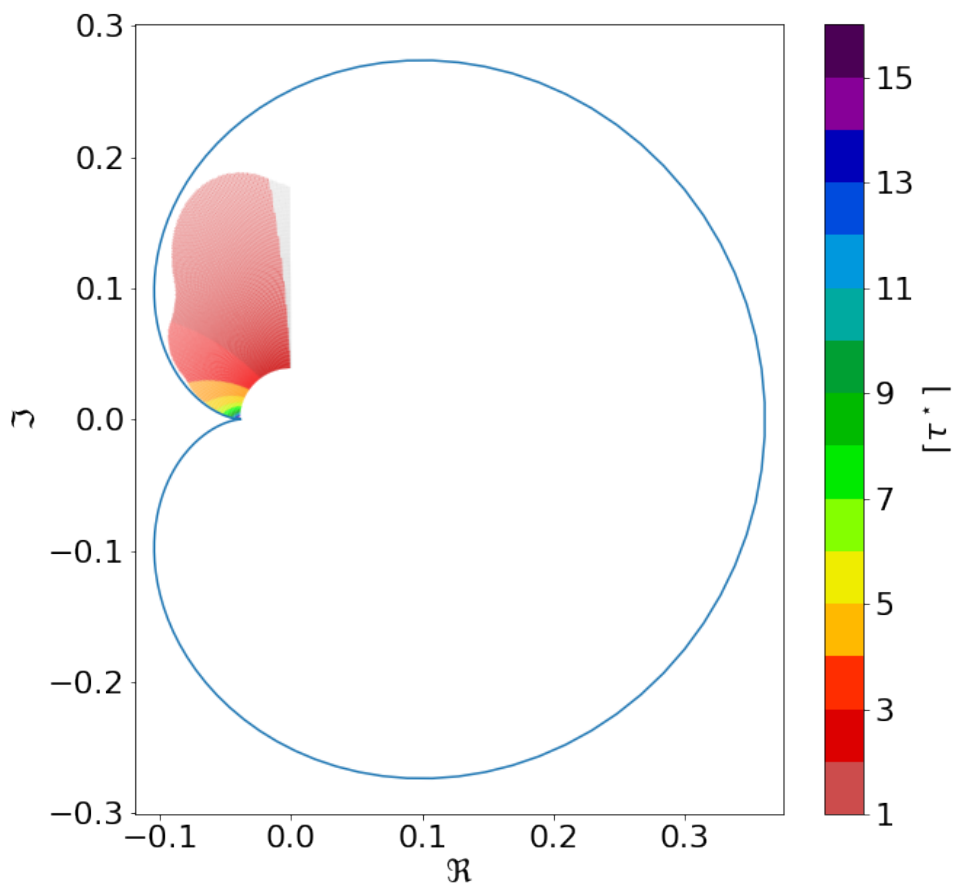
Proof. Note that $z_0 := \frac{1-\epsilon}{d-1}$ is a fixed point of the map $z \mapsto \frac{\lambda_0}{(1+z)^d}$ where $\lambda_0 = (1 - \epsilon)\frac{(d-\epsilon)^d}{(d-1)^{d+1}}$. Note also that $\lambda_0 \to \lambda^*$ as $\epsilon \to 0$. We consider two complex conjugate points $z_\pm = \frac{1-\epsilon\pm\iota\delta}{d-1}$ for some $\delta > 0$ to be fixed later. We have $\lambda = \lambda_0 \exp(\iota\theta)$ where $\theta > 0$ is small enough (as a function of ϵ and δ) to be fixed later. We will now use Theorem 4.5 for the curve defined by

$$(5.1) \qquad h(t) = \begin{cases} -tz_- & \text{if } t \in [-1, 0) \\ tz_+ & \text{if } t \in [0, 1] \end{cases}.$$

The first three conditions required of the curve h in Theorem 4.5 are satisfied by construction (see fig. 4 for an example sketch). We now proceed to verify the fourth condition. To start with, a direct computation reveals that

$$(5.2) \qquad f(z_-) = \frac{\lambda}{(1 + z_-)^d} = (1 - \epsilon)\frac{(d-\epsilon)^d}{(d-1)^{d+1}}\left(1 + \frac{1 - \epsilon - \iota\delta}{d - 1}\right)^{-d}\exp(\iota\theta)$$

$$(5.3) \qquad = \frac{1 - \epsilon}{d - 1}\left(1 - \frac{\iota\delta}{d - \epsilon}\right)^{-d}\exp(\iota\theta).$$

Figure 3: A numerical exploration of points that satisfy the condition of Theorem 4.6 for $d = 9$. Colors represent the value of $\lceil \tau^\star \rceil$.
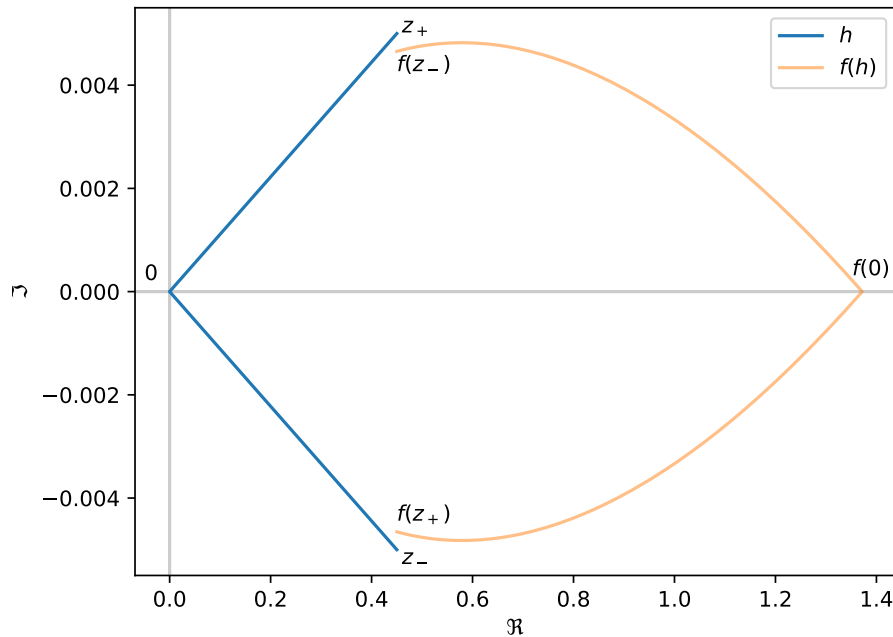
Figure 4: A sketch of the h curve in the proof of Theorem 5.1. In the notation of the theorem, the sketch corresponds to $d := 3$, $\epsilon := 0.1$ and $\delta := 0.01$. θ has been set to 0 for simplicity. As before, f is the map $z \mapsto \lambda/(1+z)^d$. The aspect ratio in the figure has been chosen to be different from 1 to accentuate features close to the real line.

We have $\arg(1 - \frac{\delta i}{d-\epsilon}) = -\tan^{-1}(\frac{\delta}{d-\epsilon}) = -\frac{\delta}{d-\epsilon} + O(\delta^2)$. Therefore,

$$(5.4) \qquad \arg(f(z_-)) = -d \arg\left(1 - \frac{\iota\delta}{d-\epsilon}\right) + \theta = \frac{d\delta}{d-\epsilon} + O(\delta^2) + \theta.$$

In comparison,

$$(5.5) \qquad \arg z_+ = \tan^{-1}\left(\frac{\delta}{1-\epsilon}\right) = \frac{\delta}{1-\epsilon} + O(\delta^2).$$

Thus, for all $\delta > 0$ small enough (depending on ϵ and d) and all $\theta \geq 0$ small enough (depending on $d \geq 2$, δ and ϵ), we have

$$(5.6) \qquad \arg z_+ > \arg f(z_-) > 0.$$

Observe also that $|f(z_-)| < |\frac{1-\epsilon}{d-1}| < |z_+|$. This implies that $f(z_-)$ is -1-covered by some point on the line segment from 0 to z_+, and in particular, $0 \leq \arg(1 + f(z_-)) \leq \arg(1 + z_+)$. An essentially symmetric argument shows that $f(z_+)$ is also -1-covered by some point on the line segment between 0 and z_-, and in particular, $0 \geq \arg(1 + f(z_+)) \geq \arg(1 + z_-)$. The inequality analogous to eq. (5.6) for $f(z_+)$ is (again, provided that θ has small enough magnitude)

$$(5.7) \qquad \arg z_- < \arg f(z_+) < 0.$$

For later use, we also record the following computation. For small positive θ and δ, we have, by a direct

computation,

$$\text{(5.8)} \qquad \arg \frac{f(z_+)}{1 + f(z_+)} = -\frac{d(d-1)}{(d-\epsilon)^2}\delta + O(\delta^2) + O(\theta), \text{ and}$$

$$\text{(5.9)} \qquad \arg \frac{z_+}{1 + z_+} = \frac{d-1}{(d-\epsilon)(1-\epsilon)}\delta + O(\delta^2).$$

In particular, since $d \geq 2$, we have

$$\text{(5.10)} \qquad \arg \frac{f(z_+)}{1 + f(z_+)} + \arg \frac{z_+}{1 + z_+} \geq 0$$

for all small enough $\delta > 0$ and $\theta > 0$. At this point, we specify our choice of θ and δ: we choose $\delta < 1$ and $\theta < \pi/10$ positive and small enough that (i) eq. (5.6), its analogue eq. (5.7) for $f(z_+)$, and eq. (5.10) are all valid, and (ii) $|\arg z_+| = |\arg z_-| \leq \pi/10$. In the following, we use these conditions imposed on δ and θ without comment.

We now claim that the curve $\{f(tz_+) : 0 \leq t \leq 1\}$ is also -1-covered by the curve h defined above. To prove this, we define $\gamma(t) := 1 + f(tz_+)$, and note that we have (for $s \in (0,1)$) $\frac{d}{ds}\arg\gamma(s)|_{s=t} = \Im\frac{\gamma'(t)}{\gamma(t)}$.

We begin by noting that as t increases from 0 to 1, $\arg f(tz_+)$ decreases from $\theta = \arg f(0)$ to $\arg f(z_+) > \arg z_-$ (since $\arg(1 + tz_+)$ increases as t increases), while $|f(tz_+)|$ decreases from λ_0 to $|f(z_+)| < |z_-|$ (again, since $|1 + tz_+|$ increases as t increases). We now compute

$$\text{(5.11)} \qquad \frac{d}{dt}\arg\gamma(t) = \Im\frac{\gamma'(t)}{\gamma(t)} = -d\Im\left(\frac{z_+}{1 + tz_+} \cdot \frac{f(tz_+)}{1 + f(tz_+)}\right).$$

We will now show that $\frac{d}{dt}\arg\gamma(t) \leq 0$ for all $t \in (0,1)$. For any particular t, if $\arg f(tz_+) \geq 0$, then eq. (5.11) immediately implies that $\frac{d}{dt}\arg\gamma(t) < 0$ (since $\arg z_+ > 0$). For any other $t \in (0,1)$, we must have $\arg f(z_+) \leq \arg f(tz_+) \leq 0$. We then get

$$\text{(5.12)} \qquad \arg\frac{f(tz_+)}{1 + f(tz_+)} = \tan^{-1}\frac{\sin\arg f(tz_+)}{|f(tz_+)| + \cos\arg f(tz_+)}$$

$$\text{(5.13)} \qquad\qquad\qquad \geq \tan^{-1}\frac{\sin\arg f(z_+)}{|f(tz_+)| + \cos\arg f(z_+)} \quad \text{since } \arg f(tz_+) \geq \arg f(z_+),$$

$$\text{(5.14)} \qquad\qquad\qquad \geq \tan^{-1}\frac{\sin\arg f(z_+)}{|f(z_+)| + \cos\arg f(z_+)} \quad \text{since } |f(tz_+)| \geq |f(z_+)| \text{ and } \arg f(z_+) \leq 0,$$

$$\text{(5.15)} \qquad\qquad\qquad = \arg\frac{f(z_+)}{1 + f(z_+)} \geq -\arg\frac{z_+}{1 + z_+}.$$

Here, the last inequality comes from eq. (5.10). Combining this with the observation that $\arg\frac{z_+}{1+tz_+}$ is strictly decreasing in t for $t \in (0,1)$, and substituting in eq. (5.11), we get the required claim that $\frac{d}{dt}\arg\gamma(t) \leq 0$ for all t in $(0,1)$. Thus, $\arg(1 + f(tz_+))$ *decreases* as t increases from 0 to 1. An essentially symmetrical argument shows that $\arg(1 + f(tz_-))$ *increases* as t increases from 0 to 1. Since we already established that $f(z_+)$ and $f(z_-)$ are -1-covered by h, and also that $\arg z_- < \arg f(h(t)) < \arg z_+$ for all $t \in [-1, 1]$, this establishes that the whole curve $\{f(h(t)) \mid t \in [-1,1]\}$ is -1-covered by h.

We thus see that the fourth condition of Theorem 4.5 is also satisfied for the curve h. We conclude therefore that $Z_G(\lambda) \neq 0$ for all graphs G of maximum degree at most $d + 1$. $\qquad\square$

6 A new zero-free region in the vicinity of the critical point

In this section, we use Theorem 4.6 to establish a new zero-free region for the independence polynomial in the vicinity of the negative real line. The result in this section applies more generally to points in the left half-plane away from the imaginary axis; we consider points close to the imaginary axis in Section 7.

We recall that $\lambda^* = \lambda^*(d) := \frac{d^d}{(d+1)^{d+1}}$ is the Shearer threshold. Consider the boundary ∂U_d of the "cardioid-shaped" region U_d (eq. (1.1)) of Peters and Regts [26]. Near the negative real line, one can calculate that the curve
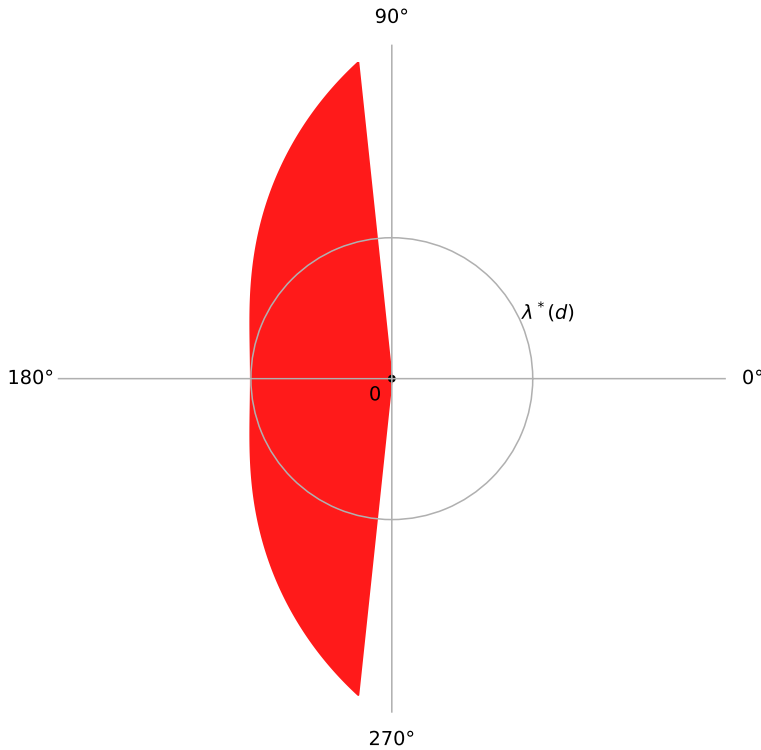
Figure 5: The red region is the zero-free region in Theorem 6.1, plotted here for $d = 9$ (i.e., for graphs of degree at most 10). The black circle around the origin has radius $\lambda^*(d)$, and the markings are according to polar coordinates.

∂U_d follows a power law of the following form. Let $R_U(\theta)$ denote the polar equation of ∂U_d, and let $(X_U(\theta), Y_U(\theta))$ denote the corresponding Cartesian coordinates $(R_U(\theta)\cos\theta, R_U(\theta)\sin\theta)$. In the vicinity of the point $-\lambda^*(d)$ on ∂U_d, a somewhat tedious but straightforward calculation shows that for small ϕ,

$$(6.1) \qquad X_U(\pi + \phi) = -\lambda^*(d) - c_d \cdot |\phi|^{2/3} + o(|\phi|)^{2/3}),$$

where c_d is a positive constant depending only on d. While we cannot prove that the true root-free region matches this exact power law, we have the following result which gives a weaker power law (see fig. 5 for a pictorial description).

THEOREM 6.1. *Fix an integer $d \geq 2$. If $\lambda = -\lambda^* \exp(r - \iota\theta)$, where $\theta \in (0, \cos^{-1}\frac{1}{d+0.5}]$ and $0 \leq r \leq \min\left\{d\log(1 + \frac{1}{d}), \frac{2d(d+1)\sin^2(\theta/2)}{d^2 + 4(d+1)\sin^2(\theta/2)}\right\}$, then $Z_G(\lambda) \neq 0$ for any graph G of degree at most $d + 1$.*

Before proving the theorem, we briefly describe the power law (analogous to the one stated above for ∂U_d) that the region described in the theorem follows. Again, we denote by $\tilde{R}(\theta)$ the polar equation of the boundary of the region described by the theorem, and let $(\tilde{X}(\theta), \tilde{Y}(\theta))$ denote the corresponding Cartesian coordinates $(\tilde{R}(\theta)\cos\theta, \tilde{R}(\theta)\sin\theta)$. In the vicinity of the point $-\lambda^*(d) = R_U(\pi) = \tilde{R}(\pi)$, a similar computation as above then shows that for small ϕ,

$$(6.2) \qquad \tilde{X}(\pi + \phi) = -\lambda^*(d) - \tilde{c}_d \cdot |\phi|^2 + o(|\phi|^2),$$

where $\tilde{c}_d = \frac{\lambda_*(d)}{2d}$ is a positive constant depending only on d.

Proof of Theorem 6.1. We use Theorem 4.6. In particular, we will show that the curve $h(t)$ defined there lies in the upper half plane $\{z | \Im z \geq 0\}$ for every $t \geq 0$. (See Figure 2 for an example of this curve for a particular setting of the parameters d, r and θ.) In fact, we will prove by an induction on $\lceil t \rceil$ that for all $t \geq 0$,

1. $|h(t)| \leq \tau := \frac{1}{d+1-\delta} \leq \frac{1}{d+0.5} \leq \frac{1}{2}$, and

2. $\arg h(t) \in [0, \pi - \theta]$,

where $\delta = \delta(d, r, \theta) < 1/2$ is a fixed non-negative constant. We first verify these for the base case $\lceil t \rceil = 1$. In this case, we have $|h(t)| = t\,|\lambda| \leq \exp(r)\frac{d^d}{(d+1)^{d+1}} \leq \frac{1}{d+1} \leq \tau$ since $r \leq d\log(1+1/d)$. Further, $\arg h(t) = \arg \lambda = \pi - \theta$.

We now proceed with the induction. For ease of notation, we denote $|h(t-1)|$ as ρ and $\arg h(t-1)$ as $\pi - \alpha$. From the induction hypothesis, we have $\rho \leq \tau$ and $\alpha \in [\theta, \pi]$. This gives

$$|1 + h(t-1)| = \sqrt{1 + \rho^2 - 2\rho\cos\alpha} \geq \sqrt{1 + \rho^2 - 2\rho\cos\theta}$$

$$\geq \sqrt{1 + \tau^2 - 2\tau\cos\theta} = (1-\tau)\sqrt{1 + 2\tau \cdot \frac{2\sin^2(\theta/2)}{(1-\tau)^2}}$$

$$(6.3) \qquad\qquad \geq \frac{d-\delta}{d+1-\delta} \cdot \sqrt{1 + \frac{4(d+1)\sin^2(\theta/2)}{d^2}}.$$

Here, for the second inequality we use the fact that the quantity inside the square-root is decreasing in ρ since $\rho \leq \tau \leq \frac{1}{d+0.5} \leq \cos\theta$, since $\theta \in [0, \cos^{-1}(1/(d+0.5))]$. Similarly, the last inequality uses $\tau \geq \frac{1}{d+1}$. Now, note that since $\arg h(t-1) \in [0, \pi]$, $|h(t-1)| \leq \tau \leq 1$ and $|1 + h(t-1)| > 0$, we have

$$(6.4) \qquad\qquad \arg(1 + h(t-1)) \geq 0,$$

and also

$$\arg(1 + h(t-1)) \leq \frac{\Im h(t-1)}{1 + \Re h(t-1)} = \frac{\rho \sin\alpha}{1 - \rho\cos\alpha} \leq \frac{\tau \sin\alpha}{1 - \tau\cos\alpha}$$

$$(6.5) \qquad\qquad = \frac{\sin\alpha}{d+1-\delta-\cos\alpha}.$$

From this, using the fact that $d\sin\alpha + \cos\alpha \leq \sqrt{d^2+1}$ for $\alpha \in [0, \pi]$, we deduce that

$$(6.6) \qquad\qquad \arg(1 + h(t-1)) \leq 1/d,$$

provided that $\delta \leq 1/2$. Now, we have

$$(6.7) \qquad\qquad \arg h(t) = \arg\lambda - d\arg(1 + h(t-1)) = \pi - \theta - d\arg(1 + h(t-1)),$$

so that eqs. (6.4) and (6.6) imply item 2 of the induction hypothesis (since $\theta \leq \pi/2$ so that $\pi - \theta - 1 \geq 0$). For item 1, we use eq. (6.3) to calculate
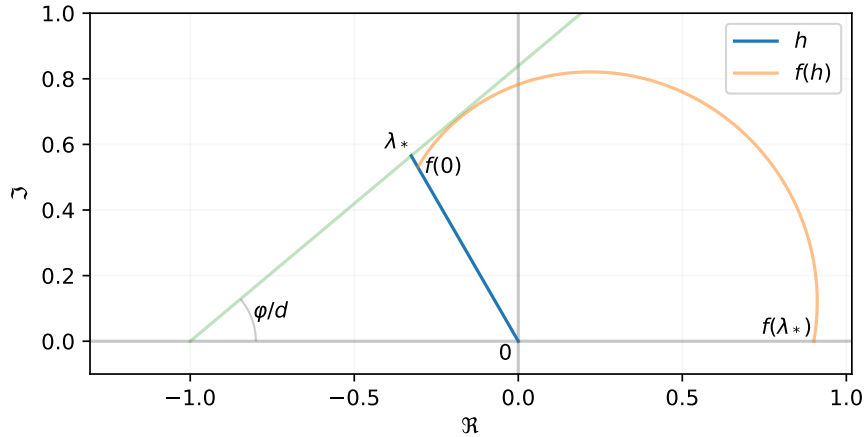
$$\log|h(t)| + \log(d+1-\delta) = r + d\log d - (d+1)\log(d+1) - d\log|1+h(t-1)| + \log(d+1-\delta)$$

$$(6.8) \qquad\qquad \leq r + (d+1)\log\left(1 - \frac{\delta}{d+1}\right) - d\log\left(1 - \frac{\delta}{d}\right)$$

$$(6.9) \qquad\qquad - \frac{d}{2}\log\left(1 + \frac{4(d+1)\sin^2(\theta/2)}{d^2}\right)$$

$$(6.10) \qquad\qquad < r - \delta + \frac{d\delta}{d-\delta} - \frac{2d(d+1)\sin^2(\theta/2)}{d^2 + 4(d+1)\sin^2(\theta/2)}$$

$$(6.11) \qquad\qquad = r - \frac{2d(d+1)\sin^2(\theta/2)}{d^2 + 4(d+1)\sin^2(\theta/2)} + \frac{\delta^2}{d-\delta} \leq 0,$$

provided $r \leq \frac{2d(d+1)\sin^2(\theta/2)}{d^2 + 4(d+1)\sin^2(\theta/2)}$ and δ is chosen to be a small enough non-negative constant depending only upon r, d and θ (note that the second inequality above is strict when δ is positive). $\qquad\square$

Figure 6: A sketch of the h curve in the proof of Theorem 7.1. In the notation of the theorem, the sketch corresponds to $d := 3$ and $\varphi := 2\pi/3$, and r has been chosen to be 0.99 times the value on the right hand side of eq. (7.1). As before, f is the map $z \mapsto \lambda/(1+z)^d$.

7   A zero-free region close to the imaginary axis

The analysis in the previous section was devoted to understanding the behavior of the zero-free region close to the negative real line. We now turn to understanding the behavior of the zero region close to the imaginary axis. The theorem below, while it covers all arguments in the third argument, is most interesting when the argument of the activity λ is closer to $\pi/2$ than to π.

THEOREM 7.1. *Let $d \geq 2$. Suppose that $\lambda = re^{\iota\varphi}$ where $\varphi \in [\frac{\pi}{2}, \pi)$ and*

$$(7.1) \qquad\qquad r < \frac{\sin(\varphi/d)\sin^d(\varphi)}{\sin((d-1)\varphi/d - d\psi^*)\sin^d(\varphi - \psi^*)},$$

where $\psi^* = \max\left(\frac{1}{d+1}\left((2-1/d)\varphi - \pi\right), 0\right)$. Then $Z_G(\lambda) \neq 0$ for any graph of degree at most $d+1$.

Proof. For a given φ, let us choose r_*, λ_* so that $\arg(1 + \lambda_*) = \varphi/d$ and $\arg(\lambda_*) = \arg(\lambda) = \varphi$. Note that for $t \in [0,1]$, $\varphi/d \geq \arg(1 + t\lambda_*) \geq 0$. As before denote $f(z) = \frac{\lambda}{(1+z)^d}$.

We claim that the function $h(t) = t\lambda_*$ satisfies the conditions of Theorem 4.5. The first three conditions of the theorem are satisfied trivially, thus we only have to show that the points $f(t\lambda_*)$, $t \in [0,1]$ are -1-covered by the segment

$$(7.2) \qquad\qquad \{t\lambda_* \ : \ t \in [0,1]\}.$$

Further, as $\arg(f(t\lambda_*)) = \varphi - d\arg(1 + t\lambda_*)$ decreases monotonically from φ to 0 as t goes from 0 to 1, it would be sufficient to prove that $\arg(1 + f(t\lambda_*))$ is at most φ/d for all $t \in [0,1)$. (See the example sketch in fig. 6.)

To prove this, we investigate the curve $\gamma(t) = 1 + f(t\lambda_*)$ for $t \in [0,1]$. Note first that we have $0 \leq \arg(\gamma(t)) \leq \arg(f(t\lambda_*)) \leq \varphi$ for all $t \in [0,1]$. Further, for all $t \in [0,1)$, we have (here, we denote by
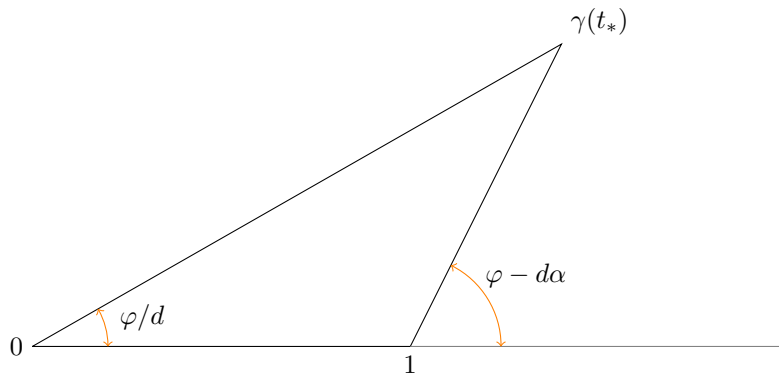
Figure 7: Angles appearing in eq. (7.6).

$\gamma'(t)$ the right one-sided derivative of γ at t)

$$(7.3) \qquad \gamma'(t) = -\frac{d\lambda\lambda_*}{(1 + t\lambda_*)^{d+1}}.$$

Since $\varphi/d \geq \arg(1 + t\lambda_*) \geq 0$ and $\arg(-\lambda\lambda_*) = 2\varphi - \pi$, we have that as t increases from 0 to 1,

$$(7.4) \qquad \arg\gamma'(t) = 2\varphi - \pi - (d+1)\arg(1 + t\lambda_*)$$

decreases monotonically from $2\varphi - \pi \in [0, \pi)$ to $(1 - 1/d)\varphi - \pi \geq -\pi + \varphi/d$.

Next, we compute that $\frac{d}{ds}\arg\gamma(s)|_{s=t} = \Im\frac{\gamma'(t)}{\gamma(t)}$ has the same sign as $\sin(\arg\gamma'(t) - \arg\gamma(t))$: note that the existence of this derivative follows since $\gamma(t)$ and $\gamma'(t)$ are non-zero, and since $\arg\gamma(t) \in [0, \pi)$ for $t \in [0, 1)$.

We now claim that $\arg(\gamma(t)) \leq \varphi/d$ for all $t \in [0, 1]$. For the sake of contradiction let us assume that $\arg\gamma(t)$ can be bigger than φ/d. As $\arg\gamma(1) = 0$, we then see that there must exist a $t_* \in [0, 1)$ such that $\arg\gamma(t_*) = \varphi/d$ and $\frac{d}{ds}\arg\gamma(s)|_{s=t_*} \leq 0$. Using the fact (noted just below eq. (7.4)) that $\pi > \arg\gamma'(t_*) \geq -\pi + \varphi/d$ and the expression for the sign of $\frac{d}{ds}\arg(\gamma(s))|_{s=t_*}$ noted above, these conditions can be written as

$$(7.5) \qquad \arg\gamma(t_*) = \varphi/d \qquad \text{and} \qquad \arg\gamma'(t_*) \leq \varphi/d.$$

Define $\alpha := \arg(1 + t_*\lambda_*) \geq 0$. Equation (7.5), along with the expression for $\arg\gamma'(t)$ in eq. (7.4), and the fact $\arg\gamma(t_*) \leq \arg f(t_*\lambda_*) = \phi - d\alpha$ noted above, gives

$$(7.6) \qquad \alpha \leq \frac{\varphi}{d}(1 - 1/d) \qquad \text{and} \qquad \alpha \geq \psi^*.$$

The standard sine rule applied to the triangle with vertices $0, 1, \gamma(t_*)$ gives us that (see fig. 7)

$$(7.7) \qquad \frac{|\gamma(t_*) - 1|}{\sin\arg\gamma(t_*)} = \frac{1}{\sin\left(\arg\left(\gamma(t_*) - 1\right) - \arg\gamma(t_*)\right)}.$$

Using the facts that (i) $|\gamma(t_*) - 1| = |f(t_*\lambda_*)| = r/|1 + t_*\lambda_*|^d$, (ii) $\arg\left(\gamma(t_*) - 1\right) = \arg f(t_*\lambda_*) = \varphi - d\alpha$, and (iii) $\arg\gamma(t_*) = \varphi/d$, we get

$$\sin(\varphi/d) = \frac{r\sin^d(\varphi - \alpha)}{\sin^d\varphi}\sin((1 - 1/d)\varphi - d\alpha)$$

$$(7.8) \qquad < \sin(\varphi/d)\frac{\sin^d(\varphi - \alpha)\sin((1 - 1/d)\varphi - d\alpha)}{\sin^d(\varphi - \psi^*)\sin((1 - 1/d)\varphi - d\psi^*)}.$$

But, in conjunction with eq. (7.6), this contradicts the fact that the function

$$(7.9) \qquad x \mapsto \sin^d(\varphi - x)\sin((1 - 1/d)\varphi - dx)$$

is a strictly decreasing function on $(\psi^*, (1 - 1/d)\varphi/d)$, since its derivative is

$$(7.10) \qquad -d\sin^{d-1}(\varphi - x) \cdot \sin\left((2 - 1/d)\varphi - (d+1)x\right) < 0.$$

Here, we use the condition $x \in (\psi^*, (1 - 1/d)\varphi/d)$ and the definition of ψ^* as $\max\left(\frac{1}{d+1}\left((2 - 1/d)\varphi - \pi\right), 0\right)$ to deduce the last inequality. $\qquad\square$

8 Zero free regions in the right half plane

In this section, we use the framework of Section 4 to establish a zero free region for the independence polynomial in the right half plane. The results here improve upon those in the manuscript [9] when λ is close to the real axis and match those results when λ is on the imaginary axis: see Remark 8.3 for a more detailed discussion.

We start with some notation. For any integer $d \geq 2$, let $\theta_d \in (\pi/(2(d+1)), \pi/2)$ be the unique solution of

$$(8.1) \qquad \tan(2x/d) = \frac{\tan((\pi/2 - x)/d)}{1 - \frac{\tan((\pi/2-x)/d)}{\tan(x)}}.$$

To see that θ_d exists and is unique, we first note that the left hand side of the above equation is monotone increasing while the right hand side is monotone decreasing, so that it has at most one solution in the given interval. To show existence, we note that as $x \downarrow \pi/(2(d+1))$, we have

$$(8.2) \qquad \lim_{x\downarrow\pi/(2(d+1))} \tan(2x/d) = \tan(\pi/(d(d+1))) < \lim_{x\downarrow\pi/(2(d+1))} \frac{\tan((\pi/2 - x)/d)}{1 - \frac{\tan((\pi/2-x)/d)}{\tan(x)}} = \infty,$$

while as $x \uparrow \pi/2$ we have

$$(8.3) \qquad \lim_{x\uparrow\pi/2} \tan(2x/d) = \tan(\pi/d) > \lim_{x\uparrow\pi/2} \frac{\tan((\pi/2 - x)/d)}{1 - \frac{\tan((\pi/2-x)/d)}{\tan(x)}} = 0.$$

Together, these show that there is a unique solution θ_d, such that for all x such that $\pi/(2(d+1)) < x < \theta_d$, we have

$$(8.4) \qquad \frac{\tan(2x/d)}{\sin x} < \frac{\tan((\pi/2 - x)/d)}{\sin(x) - \cos(x)\tan((\pi/2 - x)/d)},$$

while for $\theta_d < x < \pi/2$,

$$(8.5) \qquad \frac{\tan(2x/d)}{\sin x} > \frac{\tan((\pi/2 - x)/d)}{\sin(x) - \cos(x)\tan((\pi/2 - x)/d)},$$

For later comparison with results of [9], we also note that at $x = \pi/6$ we have (assuming $d \geq 3$)

$$(8.6) \qquad \frac{\tan(2x/d)}{\sin x} = \frac{\tan(\pi/(3d))}{\sin(\pi/6)} < \frac{\tan(\pi/(3d))}{\sin(\pi/6) - \cos(\pi/6)\tan(\pi/(3d))},$$

which implies that $\pi/6 < \theta_d$ for all $d \geq 2$ (this conclusion is trivially true for $d = 2$). We are now ready to state the main result describing zeros in the right half plane.

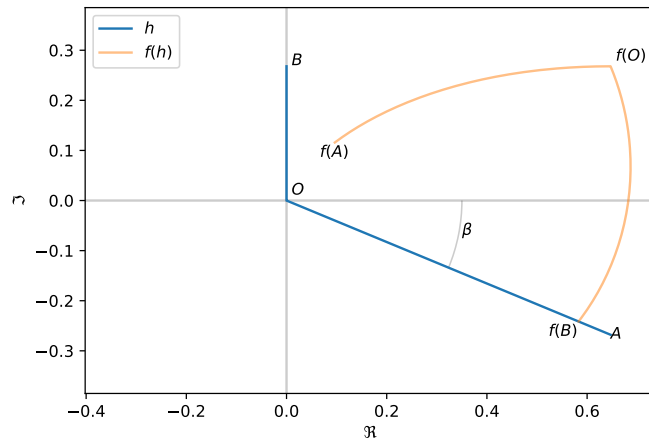THEOREM 8.1. *Let $\theta \in (0, \pi/2]$ and $0 \leq r \leq r_{1,d}(\theta)$, where*

$$(8.7) \qquad r_{1,d}(\theta) = \begin{cases} \frac{\tan(2\theta/d)}{\sin(\theta)} & \text{if } \theta \leq \theta_d \\ \frac{\tan((\theta+\beta^*)/d)}{\sin(\theta)} & \text{if } \theta > \theta_d \end{cases},$$

and where $\beta^* \in (0, \theta)$ is defined as the unique solution of

$$(8.8) \qquad \frac{\tan((\theta + x)/d)}{\sin(\theta)} = \frac{\tan((\pi/2 - \theta)/d)}{\sin(x) - \cos(x)\tan((\pi/2 - \theta)/d)},$$

when $\theta \in [\theta_d, \pi/2)$ and as $\beta^* := 0$ when $\theta = \pi/2$. If $\lambda = r\exp(\iota\theta)$, then $Z_G(\lambda) \neq 0$ for any graph G with degree at most $d + 1$.

Figure 8: An example sketch of the h curve from Lemma 8.2. In the notation of the lemma, the sketch corresponds to $d := 3$, $\beta := \theta := \pi/(2(d+1))$, $\psi := 2\theta/d$, $r_2 := r := \tan(2\theta/d)/\sin\theta$ and $\lambda := r\exp(\iota\theta)$. As before, f is the map $z \mapsto \lambda/(1+z)^d$.

The proof of this theorem is based on the following technical lemma, which employs the framework of Theorem 4.5.

LEMMA 8.2. *Let $\lambda = r\exp(\iota\theta)$ with $\theta \in (0, \pi/2]$ and $r > 0$. Suppose that there exist $r_2 \geq 0$ and $\beta, \psi \in [0, \pi/2]$ satisfying*

1. $\theta - d\psi \geq -\beta$,

2. $r_2 \geq r$,

3. $r\sin(\theta) \leq \tan\psi$,

4. $\theta + d\arg(1 + r_2\exp(\iota\beta)) \leq \pi/2$, *and*

5. $\theta \geq \beta$.

Then the curve

(8.9)
$$h(t) := \begin{cases} -t \cdot r_2\exp(-\iota\beta) & \text{if } t \in [-1, 0], \text{and} \\ t \cdot \tan(\psi)\iota & \text{if } t \in [0, 1] \end{cases}$$

satisfies the conditions of Theorem 4.5.

Proof. Let A denote the point $r_2\exp(-\iota\beta)$, B the point $\iota\tan\psi$, and O the origin. (See fig. 8 for an example sketch.) Since $h(0) = 0$, h satisfies the first condition of Theorem 4.5. Further, the curve $h(t)$ traverses the directed line segment AO as t varies from -1 to 0 and the line segment OB as t varies from 0 to 1, and this establishes the second condition of Theorem 4.5 (since $\beta \in [0, \pi/2]$ and $\psi \geq 0$).

A convex combination of any two points $h(t_1)$ and $h(t_2)$, where $t_1 < t_2$, lies either on the curve h (when $0 \notin (t_1, t_2)$), or on the boundary of the triangle with vertices $h(t_1) \in AO, h(0) = 0$ and $h(t_2) \in OB$ (when $0 \in (t_1, t_2)$). It is therefore -1 covered by $h(t)$ for some $t \in [t_1, t_2]$. This establishes the third condition of Theorem 4.5.

Note that the segment OB of the curve h -1-covers every point in the set

(8.10)
$$L_1 := \{z : \Re z \geq 0 \text{ and } 0 \leq \Im z \leq \Im B = \tan\psi \geq 0\},$$

while the segment AO -1-covers every point in the set

$$(8.11) \qquad L_2 := \{z : -\pi/2 < -\beta \leq \arg z \leq 0 \text{ and } \Im z \geq \Im A = -r_2 \sin \beta \leq 0\}.$$

It remains to verify the fourth condition of Theorem 4.5, which is that for every $t \in [0,1]$, $f(h(t))$ is -1-covered by some point on the curve h. We do so by proving that for all $t \in [-1,1]$, $f(h(t)) \in L_1 \cup L_2$.

Consider first a point $f(h(t))$ for $t \in [0,1]$. Note that as t increases from 0 to 1, $\arg(1 + h(t))$ increases from 0 to ψ. From item 1 in the statement of the lemma, we thus get that $\arg f(h(t)) \in [-\beta, \theta]$, while item 2 gives $|f(h(t))| \leq r \leq r_2$. Together with item 3, these imply that for $t \in [0,1]$,

1. $0 \leq \Im f(h(t)) \leq r \sin \theta \leq \tan \psi$ (when $\arg f(h(t)) \geq 0$), and

2. $0 \geq \Im f(h(t)) \geq -r \sin \beta \geq -r_2 \sin \beta$ (when $\arg f(h(t)) \leq 0$).

Thus, for all $t \in [0,1]$, $f(h(t)) \in L_1 \cup L_2$, and thus is -1-covered by the curve h.

Now, consider a point $f(h(t))$ for $t \in [-1,0]$. Define $g(t) := f(h(t))$. From item 4, we get that $\arg g(t) \in [\theta, \pi/2]$ for all $t \in [-1,0]$. We also have $g(0) = \lambda \in L_1$ (where the last inclusion follows from item 3). Thus, in order to establish that $g(t) \in L_1$ for all $t \in [-1,0]$, it suffices to prove that $k(t) := \Im g(t)$ has a non-negative right derivative at every $t \in [-1,0)$. The latter in turn would follow if we establish that $\arg g'(t) \in [0, \pi)$ for all $t \in [-1,0)$, where $g'(t)$ denotes the right derivative of g at t.

We now compute, for $t \in [-1,0)$,

$$(8.12) \qquad g'(t) = d \cdot r_2 \cdot \frac{\lambda}{(1 - t \cdot r_2 \exp(-\iota\beta))^d} \frac{\exp(-\iota\beta)}{(1 - t \cdot r_2 \exp(-\iota\beta))}$$

so that (after multiplying denominators with conjugates and ignoring positive real factors)

$$(8.13) \qquad \arg g'(t) = \arg \left(\underbrace{[\exp(\iota\theta) \cdot (1 - t \cdot r_2 \exp(\iota\beta))^d]}_{=:\mu} \cdot \underbrace{[\exp(-\iota\beta) - t \cdot r_2]}_{=:\nu} \right).$$

Since $t \in [-1,0)$, item 4 in the statement of the lemma then implies that $\arg \mu \in [\theta, \pi/2]$. Further, $\arg \nu \in [-\beta, 0]$. Together with $\theta \geq \beta$ (item 5 in the statement of the lemma), this implies that $\arg g'(t) = \arg(\mu \cdot \nu) \in [0, \pi/2] \subseteq [0, \pi)$. Given the above discussion about the relationship between the functions k and g, this completes the proof. $\qquad \square$

With the above lemma, we can now complete the proof of Theorem 8.1.

Proof of Theorem 8.1. We will prove that if $r \leq r_{1,d}(\theta)$, then we can find $\beta, \psi \in [0, \pi/2)$ and $r_2 > 0$, such that the conditions of Lemma 8.2 hold. By applying Theorem 4.5 for the curve h obtained from the lemma, we get the desired statement.

1. Consider first the case $\theta \leq \theta_d$. Then, let

$$(8.14) \qquad \beta = \theta, \qquad \psi = \frac{2\theta}{d}, \qquad r_2 = \arg \max_{t \geq 0} \left(\arg(1 + t \exp(\iota\theta)) \leq \frac{\pi/2 - \theta}{d} \right) \leq \infty.$$

Items 1 and 3 to 5 in Lemma 8.2 are satisfied by construction (as discussed below, we might have to redefine r_2 to make sure it is finite). We now show that item 2 holds:

- If $\theta \leq \pi/(2(d+1))$, then $r_2 = \infty \geq r$. In this case, we redefine $r_2 = r$, and all of the conditions continue to hold.

- Otherwise $\theta > \pi/(2(d+1))$. In this case, we have $r_2 = \frac{\tan((\pi/2-\theta)/d)}{\sin(\theta) - \cos(\theta)\tan((\pi/2-\theta)/d)} \geq \frac{\tan(2\theta/d)}{\sin(\theta)} = r_{1,d}(\theta) \geq r$, where the first inequality follows from eq. (8.4) since $\pi/(2(d+1)) < \theta \leq \theta_d$.

2. Consider now the case $\theta > \theta_d$. Then let β be $\beta^* \in [0, \theta]$ as described in the statement of the theorem. By definition, $\beta^* = 0$ when $\theta = \pi/2$, so we first show that even when $\theta \in (\theta_d, \pi/2)$, this β^* exists and is unique. To see this, note that

$$(8.15) \qquad \gamma_1(x) := \frac{\tan((\theta + x)/d)}{\sin(\theta)}$$

is continuous, monotone increasing and positive on $[0, \theta]$ (when $0 < \theta < \pi/2$). On the other hand,

$$(8.16) \qquad \gamma_2(x) := \frac{\tan((\pi/2 - \theta)/d)}{\sin(x) - \cos(x)\tan((\pi/2 - \theta)/d)}$$

is continuous in $[0, (\pi/2 - \theta)/d)) \cup ((\pi/2 - \theta)/d, \theta]$, negative in $[0, (\pi/2 - \theta)/d))$, and monotone decreasing in $((\pi/2 - \theta)/d, \theta]$. Further, in the interval $((\pi/2 - \theta)/d, \theta]$ we also have

$$(8.17) \qquad \gamma_1((\pi/2 - \theta)/d) < \infty, \qquad \text{and} \qquad \lim_{x \downarrow (\pi/2 - \theta)/d} \gamma_2(x) = \infty,$$

at the left endpoint, while at the right endpoint, $\theta > \theta_d$ implies $\gamma_1(\theta) > \gamma_2(\theta)$ (due to eq. (8.5)). The above observations imply that when $\theta \in (\theta_d, \pi/2)$, $\gamma_1(x) = \gamma_2(x)$ has exactly one solution $\beta^* \in [0, \theta]$, which lies in $((\pi/2 - \theta)/d, \theta]$.

Then, we define

$$\beta = \beta^*, \qquad \psi = (\theta + \beta)/d,$$

$$(8.18) \qquad r_2 = \begin{cases} \arg\max_{t \geq 0} \left(\arg(1 + t\exp(\iota\beta)) \leq \frac{\pi/2 - \theta}{d} \right) < \infty & \text{when } \theta \in (\theta_d, \pi/2), \\ \tan(\pi/(2d)) & \text{when } \theta = \pi/2. \end{cases}$$

Note that r_2 is finite when $\theta \in (\theta_d, \pi/2)$ since $\beta = \beta^* > (\pi/2 - \theta)/d$.

Again, items 1 and 3 to 5 in Lemma 8.2 are satisfied by construction. We now show that item 2 holds. To see this, we first note that when $\theta_d < \theta < \pi/2$, item 2 holds since in that case, eq. (8.18) and the definition of β^* give $r_2 = \frac{\tan((\pi/2 - \theta)/d)}{\sin(\beta^*) - \cos(\beta^*)\tan((\pi/2 - \theta)/d)} = \frac{\tan((\theta + \beta^*)/d)}{\sin(\theta)} = r_{1,d}(\theta) \geq r$. In the remaining case $\theta = \pi/2$, item 2 holds since in that case, eq. (8.18) gives again $r_2 = \tan(\pi/(2d)) = r_{1,d}(\pi/2) \geq r$. $\qquad\square$

REMARK 8.3. We remark that the zero-free region established in Theorem 8.1 contains the zero-free region described in the manuscript [9] when $\arg\lambda = \theta \leq \theta_d$ (recall also from the paragraph just before the statement of Theorem 8.1 that θ_d is always greater than $\pi/6$). For such θ, the above theorem gives zero-freeness for all λ with $|\lambda| < \tan(2\theta/d)/\sin(\theta)$ and $\arg(\lambda) = \theta$. On the other hand, the zero-free region in Theorem 1.4 of [9] requires at least that $|\lambda| \leq \tan(\pi/(2d))$. But when $d \geq 2$ and $\theta \in (0, \pi/2)$, elementary arguments involving the convexity of the function $\theta \mapsto \tan(2\theta/d) - \sin(\theta)\tan(\pi/(2d))$ in the interval $(0, \pi/2)$ imply that $\tan(\pi/(2d)) < \tan(2\theta/d)/\sin(\theta)$, showing that Theorem 8.1 gives a larger zero-free region. For the case $\theta = \pi/2$, we compute directly that $r_{1,d}(\pi/2) = \tan(\pi/(2d))$.

For the case $\theta_d < \arg\lambda < \pi/2$, the zero-free region in [9] has only an implicit description, and numerical calculations show that even in this case, the zero free region described in Theorem 8.1 is better than the one in [9], except possibly in the close vicinity of $\theta = \pi/2$.

References

[1] ARORA, S., AND BARAK, B. *Computational Complexity: A Modern Approach*. Cambridge University Press, 2009.

[2] BANDYOPADHYAY, A., AND GAMARNIK, D. Counting without sampling: Asymptotics of the log-partition function for certain statistical physics models. *Random Structures & Algorithms 33*, 4 (2008), 452–479. Extended abstract in SODA 2006.

[3] BARVINOK, A. Computing the Partition Function for Cliques in a Graph. *Theory Comput. 11* (Dec. 2015), 339–355.

[4] BARVINOK, A. Computing the Permanent of (Some) Complex Matrices. *Found. Comput. Math. 16*, 2 (Jan. 2015), 329–342.

[5] BARVINOK, A. *Combinatorics and Complexity of Partition Functions.* Algorithms and Combinatorics. Springer, 2017.

[6] BARVINOK, A., AND SOBERÓN, P. Computing the partition function for graph homomorphisms. *Combinatorica 37*, 4 (2017), 633–650.

[7] BENCS, F. On trees with real-rooted independence polynomial. *Disc. Math. 341*, 12 (Dec. 2018), 3321–3330.

[8] BENCS, F., BUYS, P., AND PETERS, H. The limit of the zero locus of the independence polynomial for bounded degree graphs, Nov. 2021. arXiv:2111.06451.

[9] BENCS, F., AND CSIKVÁRI, P. Note on the zero-free region of the hard-core model, Jul 2018. arXiv:1807.08963.

[10] BENCS, F., DAVIES, E., PATEL, V., AND REGTS, G. On zero-free regions for the anti-ferromagnetic Potts model on bounded-degree graphs. *Ann. Inst. Henri Poincaré D 8*, 3 (Sept. 2021), 459–489.

[11] BEZÁKOVÁ, I., GALANIS, A., GOLDBERG, L. A., AND ŠTEFANKOVIČ, D. Inapproximability of the Independent Set Polynomial in the Complex Plane. *SIAM J. Comput. 49*, 5 (2020), STOC18–395–STOC18–448. STOC 2018 Special Section.

[12] BUYS, P. Cayley Trees do Not Determine the Maximal Zero-Free Locus of the Independence Polynomial. *Michigan Math. J. 70*, 3 (Aug. 2021), 635–648.

[13] DE BOER, D., BUYS, P., GUERINI, L., PETERS, H., AND REGTS, G. Zeros, chaotic ratios and the computational complexity of approximating the independence polynomial, Apr. 2021. arXiv:2104.11615.

[14] DOBRUSHIN, R. L., AND SHLOSMAN, S. B. Completely Analytical Gibbs Fields. In *Statistical Physics and Dynamical Systems: Rigorous Results*, J. Fritz, A. Jaffe, and D. Szász, Eds., Progress in Physics. Birkhäuser, Boston, MA, 1985, pp. 371–403.

[15] DOBRUSHIN, R. L., AND SHLOSMAN, S. B. Completely Analytical Interactions: Constructive Description. *J. Stat. Phys. 46* (1987), 983–1014.

[16] ELDAR, L., AND MEHRABAN, S. Approximating the permanent of a random matrix with vanishing mean. In *Proceedings of the 59th Annual IEEE Symposium on Foundations of Computer Science (FOCS)* (2018), IEEE, pp. 23–34.

[17] GALANIS, A., GE, Q., ŠTEFANKOVIČ, D., VIGODA, E., AND YANG, L. Improved inapproximability results for counting independent sets in the hard-core model. *Random Struct. Algorithms 45*, 1 (2014), 78–110. Extended abstract in APPROX-RANDOM 2011.

[18] GALANIS, A., ŠTEFANKOVIČ, D., AND VIGODA, E. Inapproximability for antiferromagnetic spin systems in the tree nonuniqueness region. *J. ACM 62*, 6 (2015), 50:1–50:60. Extended abstract in STOC 2014.

[19] HARROW, A. W., MEHRABAN, S., AND SOLEIMANIFAR, M. Classical algorithms, correlation decay, and complex zeros of partition functions of Quantum many-body systems. In *Proceedings of the 52nd Annual ACM Symposium on Theory of Computing (STOC)*. ACM, June 2020, pp. 378–386.

[20] HARVEY, N. J. A., SRIVASTAVA, P., AND VONDRÁK, J. Computing the independence polynomial: From the tree threshold down to the roots. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* (2018), SIAM, pp. 1557–1576.

[21] LI, L., AND XIE, G. Complex contraction on trees without proof of correlation decay, Dec. 2021. arXiv:2112.15347.

[22] LIU, J. *Approximate Counting, Phase Transitions and Geometry of Polynomials.* PhD Thesis, UC Berkeley, 2019. Available at https://www2.eecs.berkeley.edu/Pubs/TechRpts/2019/EECS-2019-110.html.

[23] LIU, J., SINCLAIR, A., AND SRIVASTAVA, P. A deterministic algorithm for counting colorings with 2Δ colors. In *Proceedings of the 60th Annual IEEE Symposium on Foundations of Computer Science (FOCS)* (2019), IEEE, pp. 1380–1404.

[24] LIU, J., SINCLAIR, A., AND SRIVASTAVA, P. Fisher zeros and correlation decay in the Ising model. *J. Math. Phys. 60* (2019), 103304. Extended abstract in proceedings of ITCS 2019.

[25] PATEL, V., AND REGTS, G. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM J. Comput. 46*, 6 (2017), 1893–1919.

[26] PETERS, H., AND REGTS, G. On a Conjecture of Sokal Concerning Roots of the Independence Polynomial. *Michigan Math. J. 68*, 1 (2019), 33–55.

[27] SCOTT, A. D., AND SOKAL, A. D. The Repulsive Lattice Gas, the Independent-Set Polynomial, and the Lovász Local Lemma. *J. Stat. Phys. 118*, 5-6 (Mar. 2005), 1151–1261.

[28] SHAO, S., AND SUN, Y. Contraction: A Unified Perspective of Correlation Decay and Zero-Freeness of 2-Spin Systems. *J. Stat. Phys. 185*, 2 (Oct. 2021), 12. Extended abstract in proceedings of ICALP 2020.

[29] SHEARER, J. B. On a problem of Spencer. *Combinatorica 5*, 3 (Sept. 1985), 241–245.

[30] SIPSER, M. A complexity theoretic approach to randomness. In *Proceedings of the 15th Annual ACM Symposium on Theory of Computing (STOC)* (1983), ACM, pp. 330–335.

[31] SLY, A. Computational transition at the uniqueness threshold. In *Proceedings of the 51st Annual IEEE Symposium*

on Foundations of Computer Science (FOCS) (2010), IEEE, pp. 287–296.

[32] SLY, A., AND SUN, N. Counting in two-spin models on *d*-regular graphs. *Ann. Probab. 42*, 6 (2014), 2383–2416.

[33] STOCKMEYER, L. The complexity of approximate counting. In *Proceedings of the 15th Annual ACM Symposium on Theory of Computing (STOC)* (1983), ACM, pp. 118–126.

[34] TODA, S. PP is as Hard as the Polynomial-Time Hierarchy. *SIAM J. Comput. 20*, 5 (Oct. 1991), 865–877.

[35] WEITZ, D. Counting independent sets up to the tree threshold. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC)* (2006), ACM, pp. 140–149.

[36] YANG, C. N., AND LEE, T. D. Statistical theory of equations of state and phase transitions. I. Theory of condensation. *Phys. Rev. 87*, 3 (1952), 404–409.