Paving property for real stable polynomials and strongly Rayleigh processes

KASRA ALISHAHI1,*, MILAD BARZEGAR1,†

1Department of Mathematical Sciences, Sharif University of Technology, Tehran, Iran.
E-mail: *alishahi@sharif.edu; †milad.barzegar@sharif.edu

One of the equivalent formulations of the Kadison–Singer problem which was resolved in 2013 by Marcus, Spielman and Srivastava, is the “paving conjecture”. In this paper, we first extend this result to real stable polynomials. We prove that for every multi-affine real stable polynomial satisfying a simple condition, it is possible to partition its set of variables to a small number of subsets such that the “restriction” of the polynomial to each subset has small roots. Then, we derive a probabilistic interpretation of this result. We show that there exists a partition of the underlying space of every strongly Rayleigh process into a small number of sets such that the restriction of the point process to each set has “almost independent” points. This result implies that the dependence structure of strongly Rayleigh processes is constrained—a phenomenon that is to be expected in negatively dependent measures. To prove this result, we introduce and study the notion of kernel polynomial for strongly Rayleigh processes. This notion is a natural generalization of the kernel of determinantal processes. We also derive an entropy bound for strongly Rayleigh processes in terms of the roots of the kernel polynomial which is interesting on its own.

Keywords: Paving property; Real stable polynomials; Strongly Rayleigh processes

1. Introduction

In 1959, Richard V. Kadison and Igor M. Singer [21] raised the question whether every pure state on the algebra of bounded diagonal operators on $\ell^2(\mathbb{N})$ has a unique extension to a state on the algebra of all bounded operators on $\ell^2(\mathbb{N})$. This problem has come to be known as the Kadison–Singer problem. Over the next 54 years, this problem attracted a significant amount of research until it was resolved in the affirmative in 2013 by Adam Marcus, Daniel Spielman and Nikhil Srivastava [23].

One important aspect of the Kadison–Singer problem is that it has been shown to be equivalent to a large number of problems in various fields. One of these equivalent formulations, which will be our main focus, is as follows.

**Problem 1.1.** Let $\varepsilon \in (0, 1)$. Does there exist $r \in \mathbb{N}$ such that every Hermitian matrix $A$ whose diagonal entries are zero can be $(r, \varepsilon)$-paved, namely there are diagonal projections $P_1, \ldots, P_r$ such that $\sum_{i=1}^{r} P_i = I$ and

$$\forall i \in [r] : \|P_i A P_i\|_{op} \leq \varepsilon \|A\|_{op},$$

where $[r] = \{1, \ldots, r\}$ and $\|\cdot\|_{op}$ denotes the operator norm.

This formulation of the Kadison–Singer problem, which is known as the paving problem (or the paving conjecture for the assertion that the answer to the above question is “yes”), was discovered by Joel Anderson [6]. Anderson showed that the answer to the paving problem is positive if and only if the answer to the Kadison–Singer problem is positive. For a background on the Kadison–Singer problem and its equivalent formulations see [10] and the references therein.
Marcus et al. [23] proved a stronger version of “Weaver’s vector balancing formulation” of the Kadison–Singer problem (see [10]) using the “method of interlacing families”. This method was first introduced in [24] and provides a powerful technique for proving the existence of certain combinatorial objects. The application of this method results in an analysis of the locations of the roots of a “real stable polynomial”. The “multivariate barrier method” is introduced in [23] as a framework for such an analysis. Marcus et al. [23] obtained the following paving bound for positive semidefinite contractions with bounded diagonal entries.

Theorem 1.2. Let \( \alpha \) be a positive number and \( r \) be an integer such that \( r \geq 2 \). For every positive semidefinite contraction \( A \in \mathcal{M}_n(\mathbb{C}) \) with diagonal entries at most \( \alpha \), there are diagonal projections \( P_1, \ldots, P_r \in \mathcal{M}_n(\mathbb{C}) \) such that \( \sum_{i=1}^r P_i = I_n \) and

\[
\forall i \in [r] : \| P_i A P_i \|_{op} \leq \left( \frac{1}{r} + \sqrt{\alpha} \right)^2.
\]

Leake and Ravichandran [22] adapted the methods of [23] to directly prove the paving conjecture and they got sharper paving bounds. They used the method of interlacing families in conjunction with a modified version of the multivariate barrier method. Their result is as follows.

Theorem 1.3. Let \( r \in \mathbb{Z} \) and \( \alpha \in \mathbb{R} \) such that \( r \geq 2 \) and \( 0 < \alpha \leq (r-1)^2/r^2 \). For every positive semidefinite contraction \( A \in \mathcal{M}_n(\mathbb{C}) \) with diagonal entries at most \( \alpha \), there are diagonal projections \( P_1, \ldots, P_r \in \mathcal{M}_n(\mathbb{C}) \) such that \( \sum_{i=1}^r P_i = I_n \) and

\[
\forall i \in [r] : \| P_i A P_i \|_{op} \leq \left( \frac{1}{r} - \frac{\alpha}{r-1} + \sqrt{\alpha} \right)^2.
\]

In Section 3, we will prove a generalization of the above theorem to “real stable polynomials” using the methods of [22]. We are able to extend the main steps of Leake–Ravichandran’s proof, which are mostly linear algebraic in nature, by exploiting the “interlacing” properties of real stable polynomials. It is worthwhile to mention that two other generalizations of the Kadison–Singer problem appear in [2] and [12], both of which are through the main result of [23], namely Weaver’s formulation.

A polynomial \( p \in \mathbb{C}[z_1, \ldots, z_n] \) is stable if it has no roots in \( \mathbb{H}^n \), where \( \mathbb{H} \) is the open upper half-plane, and it is real stable if, in addition, its coefficients are real. These polynomials have received a lot of attention in recent years and they have found numerous applications in probability theory, operator theory, combinatorics, algorithms and sampling; see, for example, [11, 7, 8, 9, 28, 24, 23, 19, 4, 31]. We will review stable polynomials in Subsection 2.1.

Before stating our result, let us fix some notations. When \( n \) is specified in the context, we will use the notation \( z = (z_1, \ldots, z_n) \). Let \( \partial_i := \partial / \partial z_i \) and define \( z^I = \prod_{i \in I} z_i \) and \( \partial^I = \prod_{i \in I} \partial_i \), for every \( I \subset [n] \). The diagonalization of \( p \in \mathbb{C}[z_1, \ldots, z_n] \), denoted \( \overline{p} \), is defined by \( \overline{p}(x) = p(x, \ldots, x) \). For a real rooted polynomial \( p \), we denote its maximum root by \( \text{maxroot}(p) \). Our generalization of Theorem 1.3 is as follows.

Theorem 1.4. Let \( r \in \mathbb{Z} \) and \( \alpha \in \mathbb{R} \) such that \( r \geq 2 \) and \( 0 < \alpha \leq (r-1)^2/r^2 \). Assume that \( g \in \mathbb{R}[z_1, \ldots, z_n] \) is a multiaffine real stable polynomial and \( g(z) = \sum_{A \subseteq [n]} a_A z^A \). If all the roots of \( \overline{g} \) are in the interval \([0, 1]\), \( a_0 = 1 \) and \( |a_{\{i\}}| \leq \alpha \) for \( i = 1, \ldots, n \), then there exists a partition \( \{S_1, \ldots, S_r\} \) of \([n]\) such that

\[
\forall i \in [r] : \text{maxroot} \left( \partial^{S_i} g \right) \leq \left( \frac{1}{r} - \frac{\alpha}{r-1} + \sqrt{\alpha} \right)^2.
\]
As mentioned above, we will prove this theorem in Section 3. Here, we show that Theorem 1.3 is a special case of Theorem 1.4. The multivariate characteristic polynomial of a matrix $A \in \mathcal{M}_n(\mathbb{C})$, denoted $\chi[A]$, is defined by $\chi[A](z) = \det[Z - A]$, where $Z = \text{Diag}(z_1, \ldots, z_n)$. It is well-known that the multivariate characteristic polynomial of a Hermitian matrix is real stable (see the remarks following Proposition 2.3).

Let $A$ be as in Theorem 1.3. Note that the coefficient of the monomial $z_1 \ldots z_n$ in $\chi[A]$ is equal to 1. Also, for each $i \in [n]$, the coefficient of the monomial $z_1 \ldots z_{i-1}z_{i+1} \ldots z_n$ is equal to $A_{i,i}$ and thus its absolute value is less than $\alpha$. Since $\chi[A]$ is the characteristic polynomial of $A$ and $A$ is a positive semidefinite contraction, all its roots are in the interval $[0,1]$. Therefore, by Theorem 1.4, there exists a partition $\{S_1, \ldots, S_r\}$ of $[n]$ such that

$$\forall i \in [r]: \max\{\partial^{S_i}\chi[A]\} \leq \left(\sqrt{\frac{1}{r} - \frac{\alpha}{r-1}} + \sqrt{\alpha}\right)^2.$$ 

Let $P_i \in \mathcal{M}_n(\mathbb{C})$ be the diagonal matrix whose $k$-th diagonal entry is equal to 1 if $k \in S_i$ and is equal to 0 if $k \notin S_i$. Since $\{S_1, \ldots, S_r\}$ is a partition of $[n]$, we have $\sum_{i=1}^r P_i = I$. Now, Theorem 1.3 follows since

$$\|P_i AP_i\|_{op} = \max\{\chi[P_i AP_i]\} \quad \text{and} \quad \chi[P_i AP_i] = \partial^{S_i}\chi[A].$$

We will use Theorem 1.4 to prove a “paving property” for strongly Rayleigh point processes. A point process $\mathcal{X}$ on $[n]$, i.e. a random subset of $[n]$, is strongly Rayleigh if its probability generating polynomial, defined as

$$f_{\mathcal{X}}(z) = \sum_{A \subseteq [n]} \mathbb{P}(\mathcal{X} = A) z^{|A|},$$

is real stable.

Robin Pemantle in [27] emphasized the need for a theory of negative dependence which would take shape around an appropriate notion of negative dependence. Borcea, Brändén and Liggett [8] proposed the strong Rayleigh property as this appropriate notion. Strongly Rayleigh point processes satisfy many useful properties, such as negative association which is the strongest form of negative dependence, and they include several well-known examples of negatively dependent processes, most notably discrete “determinantal processes”. Since their introduction, strongly Rayleigh processes have been well studied and they have found many applications; see, for example, [26, 16, 13, 29, 2, 3, 17]. We will review strongly Rayleigh processes in Subsection 2.2.

We need the notion of entropy in order to state the paving property of strongly Rayleigh processes. Recall that the entropy of a random element $X$ from a finite set $S$, denoted $H(X)$, is defined by

$$H(X) = -\sum_{x \in S} \mathbb{P}(X = x) \log(\mathbb{P}(X = x)),$$

where the logarithms are taken in base 2. We use $h(p)$ to denote the entropy of a Bernoulli random variable $X$ with $\mathbb{P}(X = 1) = p$. The paving property for strongly Rayleigh processes is as follows.

**Theorem 1.5.** For each positive number $\delta$, there exists an integer $r$ such that for every strongly Rayleigh process $\mathcal{X}$ on any space $S$, it is possible to partition $S$ into $r$ subsets $S_1, \ldots, S_r$ such that

$$\forall i \in [r]: \frac{1}{|S_i|} H(\mathcal{X} \cap S_i) - \frac{1}{|S_i|} \sum_{j \in S_i} h(p_j) < \delta,$$

(1)
where $|S_i|$ denotes the size of $S_i$ and $p_j = \mathbb{P}(j \in \mathcal{X})$.

Note that in the above theorem, $r$ does not depend on the size of $S$. So Theorem 1.5 states that for every strongly Rayleigh process, there is a partition of the underlying space into a small number of sets such that the mean entropy per particle of the restriction of the process to each set is close to the mean entropy per particle of the unique product measure with the same marginals. In this sense, the points of each restriction are “almost independent”. In fact, if $\mu_i$ denotes the law of $\mathcal{X} \cap S_j$ and $\hat{\mu}_i$ denotes the product measure with the same marginals, then a straightforward calculation shows that (1) is equivalent to $D_{\text{KL}}(\mu_i \| \hat{\mu}_i) < |S_i| \delta$, where $D_{\text{KL}}$ denotes the Kullback–Leibler divergence.

This behavior hints at a fundamental difference between the notions of negative and positive dependence: while a large number of positively dependent random variables can be simultaneously strongly dependent, the dependence structure of negatively dependent random variables is more rigid. In a large set of negatively dependent random variables, there must be many that are weakly dependent.

This phenomenon is readily apparent in the level of pairwise correlations. Note that, for random variables $X_1, \ldots, X_n$, we have $0 \leq \text{var}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \text{var}(X_i) + 2 \sum_{i<j} \text{cov}(X_i, X_j)$. If $X_1, \ldots, X_n$ have pairwise negative correlations, then the above inequality implies that the “covariances” must typically be much smaller than the “variances”. There is no such restriction in the case of pairwise positive correlations. For instance, in the extreme case $X_1 = \cdots = X_n$ we have $\text{cov}(X_i, X_j) = \text{var}(X_1)$ for all $i, j \in [n]$.

Pairwise negative correlations is the weakest notion of negative dependence. It is natural to expect stronger manifestations of the rigidity of dependence structure in stronger notions of negative dependence. The paving property, Theorem 1.5, is a very strong manifestation of this phenomenon in the case of strongly Rayleigh measures. A consequence of this theorem is that in every strongly Rayleigh process, there is a large subset of the underlying space (with size of the same order as the size of the underlying space) whose points are almost independent.

We will prove Theorem 1.5 in Section 4. Essential to the proof, is the notion of “kernel polynomial” for strongly Rayleigh processes. We will introduce this notion in Subsection 2.2. The kernel polynomial is essentially a generalization of the kernel of determinantal processes and provides a unified framework for studying strongly Rayleigh and determinantal processes.

To prove Theorem 1.5, we will apply a slightly modified version of Theorem 1.4, presented in Proposition 4.2, to the kernel polynomial of $\mathcal{X}$ (the kernel polynomial is the appropriate object here because its partial derivatives correspond to the restrictions of the process; see Proposition 2.13). This proves the existence of a partition of the underlying space with the property that the roots of “centered versions” of the kernels of the restricted processes are small. We will translate this algebraic condition to an entropy inequality via the connection between stability and “hyperbolicity” and exploiting the majorization properties of hyperbolic polynomials (this is done in Subsection 4.2). We will also use an entropy bound in terms of the roots of the kernel polynomial, presented in Subsection 4.1.

This entropy bound is interesting on its own. As mentioned above, the dependence structure of a strongly Rayleigh process is constrained. In particular, its entropy cannot be too small. [5, Corollary 5.6] gives a lower bound for the entropy of strongly Rayleigh processes in terms of the entropy of its marginals. Our result provides a bound in terms of the roots of the diagonalization of the kernel polynomial (see Theorem 4.3).

In Subsection 2.2, we will see that the roots of the diagonalization of the kernel polynomial play a similar role to the eigenvalues of the kernel of determinantal processes. We know from the theory of determinantal processes that there is a “canonical probabilistic interpretation” for the eigenvalues of the kernel (see [20, Theorem 4.5.3]). It is natural to ask if there is a similar probabilistic interpretation for the roots of the diagonalization of the kernel polynomial. Even formalizing such an interpretation is challenging. In Subsection 4.1, we will propose a conjecture which can be regarded as a first step in this direction.
2. Preliminaries

The notation $A^c$ will be used to denote the complement of a set $A$, when the universal set is clear in the context. We will use $1$ to denote the vector of all 1’s, i.e. $1 = (1, \ldots, 1)$. Similarly, $0 := (0, \ldots, 0)$. We will denote the $i$-th entry of a vector $v$ by $v_i$. For $v, w \in \mathbb{R}^n$, we will use $v \geq w$ when $v_i \geq w_i$ for all $i \in [n]$. For $p \in \mathbb{C}[z_1, \ldots, z_n]$ let $\deg(p)$ denote the degree of $p$ and $\deg_j(p)$ denote the degree of $p$ in $z_j$. Also, for $v \in \mathbb{N}^n$, we will use $[z^v]_p$ to denote the coefficient of the monomial $z^v$ in $p$, where $z^v := \prod_i z_i^{v_i}$. For a real rooted polynomial $p \in \mathbb{R}[x]$, we will use $\lambda(p)$ to denote the nonincreasing vector of its roots and $\lambda_i(p)$ to denote its $i$-th largest root.

2.1. Stable Polynomials

Stable polynomials are a natural multivariate generalization of real rooted polynomials. These polynomials have many nice algebraic and geometric properties. In this subsection, we summarize some of these properties that we need for later use. See the surveys [28] and [33] for a thorough overview of this subject.

**Definition 2.1.** A polynomial $p \in \mathbb{C}[z_1, \ldots, z_n]$ is stable if

$$\text{Im}(z_1) > 0, \ldots, \text{Im}(z_n) > 0 \implies p(z_1, \ldots, z_n) \neq 0.$$ 

$p$ is real stable if, in addition, its coefficients are real. We use $\mathcal{H}_n(\mathbb{C})$ and $\mathcal{H}_n(\mathbb{R})$ to denote the set of $n$-variate stable and real stable polynomials, respectively.

Note that a univariate real polynomial is (real) stable if and only if it is real rooted. The following proposition is an immediate consequence of the definition.

**Proposition 2.2.** A polynomial $p \in \mathbb{C}[z_1, \ldots, z_n]$ is stable (real stable, respectively) if and only if for every $\alpha \in \mathbb{R}^n$ and $v \in \mathbb{R}_+^n$, the univariate polynomial $t \mapsto p(tv + \alpha)$ is stable (real stable, respectively).

Determinantal polynomials are the most important example of real stable polynomials.

**Proposition 2.3** (Proposition 1.12 of [9]). If $B \in \mathcal{M}_n(\mathbb{C})$ is Hermitian and $A_1, \ldots, A_m \in \mathcal{M}_n(\mathbb{C})$ are positive semidefinite, then the polynomial $\det(B + z_1 A_1 + \cdots + z_m A_m)$ is either identically zero or real stable.

It follows from the above proposition that for every Hermitian matrix $K \in \mathcal{M}_n(\mathbb{C})$, the polynomial $\det(Z - K)$ is real stable, where $Z := \text{Diag}(z_1, \ldots, z_n)$. This polynomial is the multivariate characteristic polynomial of $K$ and we denote it by $\chi_K(z)$.

The class of (real) stable polynomials is closed under several elementary operations. Some of these closure properties are summarized in the following proposition. See [9] and [7] for the proofs.

**Proposition 2.4.** If $p$ is a (real) stable polynomial in $n$ variables, then

1. $\partial_i p$ is identically zero or (real) stable for $i \in [n]$;
2. $p(z_1, \ldots, z_i-1, \beta, z_{i+1}, \ldots, z_n)$ is identically zero or (real) stable for $i \in [n]$ and $\beta \in \mathbb{R}$;
3. $p(z_1, \ldots, z_{i-1}, z_j, z_{i+1}, \ldots, z_n)$ is (real) stable for distinct $i, j \in [n]$. In particular $\overline{p}$ is (real) stable—recall that $\overline{p}(x) = p(x, \ldots, x)$. 


4. If \( p \) is real stable and \( \deg_i(p) = d_i \) for \( i \in [n] \), then \( z_1^{d_1} \ldots z_n^{d_n} p(\gamma_1z_1^{-1}, \ldots, \gamma_nz_n^{-1}) \) is real stable for \( \pm(\gamma_1, \ldots, \gamma_n) \in \mathbb{R}_+^n \).

As mentioned above, stability is the natural multivariate generalization of real rootedness. One important notion in the study of real rooted polynomials is “interlacing”. Polynomials \( p, q \in \mathcal{H}_1(\mathbb{R}) \) are interlacing if their roots alternate, namely

\[
\lambda_1(p) \geq \lambda_1(q) \geq \lambda_2(p) \geq \lambda_2(q) \geq \ldots \quad \text{or} \quad \lambda_1(q) \geq \lambda_1(p) \geq \lambda_2(q) \geq \lambda_2(p) \geq \ldots,
\]

in which case we must clearly have \( |\deg(p) - \deg(q)| \leq 1 \). In the case \( \deg(q) = \deg(p) - 1 \), we say that \( q \) interlaces \( p \).

It is a well-known fact that if \( p \) and \( q \) are interlacing, then the Wronskian, defined by \( W[p, q] = pq' - p'q \), is either nonnegative or nonpositive on the real line. We say that \( q \) is in proper position with respect to \( p \), denoted \( q \ll p \), if \( p \) and \( q \) are interlacing and \( W[p, q] \leq 0 \). If the leading coefficients of \( p \) and \( q \) have the same sign, then \( q \ll p \) if and only if

\[
\lambda_1(p) \geq \lambda_1(q) \geq \lambda_2(p) \geq \lambda_2(q) \geq \ldots.
\]

Also note that \( q \ll p \) if and only if \( -p \ll q \). The notion of proper position has been generalized by Borcea and Brändén to multivariate polynomials.

**Definition 2.5.** Let \( p, q \in \mathbb{R}[z_1, \ldots, z_n] \). We say that \( q \) is in proper position with respect to \( p \), denoted \( q \ll p \), if for all \( \alpha \in \mathbb{R}^n \) and \( v \in \mathbb{R}^n_+ \) the univariate polynomial \( q(tv + \alpha) \) is in proper position with respect to \( p(tv + \alpha) \).

It follows from **Proposition 2.2** and the Hermite–Biehler theorem (see [30]) that \( q \ll p \) if and only if \( p + iq \in \mathcal{H}_n(\mathbb{C}) \). It is a well-known fact that we have \( \partial_i p \ll p \) for every \( p \in \mathcal{H}_n(\mathbb{R}) \) and \( i \in [n] \) (see, e.g. [9]). An important consequence of the definition is that \( q \ll p \) implies \( \overline{q} \ll \overline{p} \). We will use this fact several times.

A polynomial \( p \in \mathbb{C}[z_1, \ldots, z_n] \) is multiaffine if \( \deg_i(p) \leq 1 \) for all \( i \in [n] \). The class of multiaffine real stable polynomials is of special importance because, as a consequence of the Grace–Walsh–Szegö coincidence theorem, the study of (real) stable polynomials can be reduced to the special case of multiaffine (real) stable polynomials (see, e.g. [14]). In this paper, we are mainly interested in these polynomials. In the rest of this subsection, we present some results concerning multiaffine real stable polynomials that will be useful for us later on.

**Proposition 2.6.** Let \( p \in \mathcal{H}_n(\mathbb{R}) \) be multiaffine and \( p = r + z_ns \), where \( r, s \in \mathbb{R}[z_1, \ldots, z_{n-1}] \). We have \( s \ll p \) and \( s \ll r \).

**Proof.** Since \( s = \partial_n p \) and \( \partial_n p \ll p \), we have \( s \ll p \). Since \( p|_{z_n=1} \) is stable and \( p|_{z_n=1} = r + is \), we have \( s \ll r \). \( \square \)

For a multiaffine polynomial \( p \in \mathbb{R}[z_1, \ldots, z_n] \) such that \( p(z) = \sum_{A \subseteq [n]} a_A z^A \), define its support, denoted by \( \text{supp}(p) \), to be the set \( \{ A \subseteq [n] : a_A \neq 0 \} \).

**Proposition 2.7** (Corollary 3.7 of [11]). Let \( p \) be a multiaffine real stable polynomial with nonnegative coefficients. If \( A \subseteq C \subseteq B \) and \( A, B \in \text{supp}(p) \), then \( C \in \text{supp}(p) \).
Lemma 2.8. Let $p \in \mathbb{R}[z_1, \ldots, z_n]$ with $p(z) = \sum_{A \subseteq [n]} a_A z^A$ be a multi-affine real stable polynomial and $a_{[n]} > 0$. All of the roots of $\overline{p}$ are nonnegative if and only if $(-1)^{n-|A|} a_A \geq 0$ for every $A \subseteq [n]$.

Proof. First we prove the “if” part. Note that the coefficient of $x^k$ in $p$ is $\sum_{|A| = k} a_A$. Therefore $(-1)^{n-k} [x^k] \overline{p} \geq 0$, for every $k \in \{0, \ldots, n\}$; namely, the signs of the coefficients of $\overline{p}$ alternate. Such a polynomial can only have nonnegative roots.

Now we prove the “only if” part. By the remarks following Definition 2.5, for every $A \subseteq [n]$ with $A = \{i_1, \ldots, i_k\}$ we have

$$
\overline{p} \gg \partial^{(i_1)} p \gg \partial^{(i_1, i_2)} p \gg \ldots \gg \partial^A p.
$$

Since $p$ is multi-affine and $a_{[n]} > 0$, the degree of each polynomial in the above sequence is one less than the degree of the polynomial preceding it. Therefore, each polynomial incorporates the polynomial preceding it. In particular, either $\partial^A p$ is identically zero or all its roots are nonnegative. In the case $\partial^A p \neq 0$, the polynomial $\partial^A p$ is of degree $n - |A|$ and its leading coefficient is positive. Therefore, we have $(-1)^{n-|A|} \partial^A p(0) \geq 0$. The lemma follows since $\partial^A p(0) = a_A$.

We will use the following corollary several times.

Corollary 2.9. Let $p$ be a multi-affine real stable polynomial and all the roots of $\overline{p}$ be nonnegative. If $A \subseteq C \subseteq B$ and $A, B \in \text{supp}(p)$, then $C \in \text{supp}(p)$.

Proof. Define $q(z) = (-1)^n p(-z)$. Note that $\text{supp}(p) = \text{supp}(q)$. By Lemma 2.8, $q$ has nonnegative coefficients. Now, the result follows from Proposition 2.7.

2.2. Strongly Rayleigh Processes

In this subsection we introduce the notion of kernel polynomial for strongly Rayleigh processes. This notion, in a sense, is a generalization of the kernel of determinantal processes, hence the name. We will show that the kernel polynomial satisfies most of the important properties of the kernel of determinantal processes.

A point process $\mathcal{X}$ on a finite set $S$ is a random subset of $S$. Note that the law of $\mathcal{X}$ is a probability measure on the lattice of all the subsets of $S$. Alternatively, $\mathcal{X}$ can be identified with its indicator (random) vector, namely $(X_i)_{i \in S}$, where $X_i$ is the indicator function of the event $\{i \in \mathcal{X}\}$. When $|S| = n$ we can replace $S$ with $[n]$ without loss of generality.

Definition 2.10. A point process $\mathcal{X}$ on $[n]$ is strongly Rayleigh if its probability generating polynomial, denoted $f_{\mathcal{X}}$ and defined by

$$
f_{\mathcal{X}}(z) = \sum_{A \subseteq [n]} \mathbb{P}(\mathcal{X} = A) z^A,
$$

is real stable.

Strongly Rayleigh processes have many nice properties, some of which are as follows:
1. Strongly Rayleigh processes have the negative association property. This is proved in [8].
2. The class of strongly Rayleigh processes is closed under many natural operations, including products, projections, external fields, conditioning and symmetric homogenization. These properties are proved in [8].
3. Strongly Rayleigh processes have strong concentration properties. For example, it is proved in [29] that Lipschitz functionals of strongly Rayleigh processes satisfy an Azuma-type concentration inequality.

For more information on strongly Rayleigh processes see [8]. The most important example of strongly Rayleigh processes is the class of (discrete) determinantal processes. A point process \( X \) on \([n]\) is determinantal if there exists a Hermitian matrix \( K \in \mathcal{M}_n(\mathbb{C}) \), called the kernel of \( X \), such that for every \( A \subseteq [n] \) we have

\[
P(A \subseteq X) = \det K_A,
\]

where \( K_A \) is the principal submatrix of \( K \) with rows and columns in \( A \). For a background on determinantal processes see [20]. It is shown in [8, Proposition 3.5] that determinantal processes are strongly Rayleigh and if \( Y \) is a determinantal process with kernel \( K \), then \( f_Y(z) = \det(KZ + I - K) \), where \( Z = \text{Diag}(z_1, \ldots, z_n) \).

**Definition 2.11.** The kernel of a point process \( X \) on \([n]\), denoted \( g_X \), is defined by

\[
g_X(z_1, \ldots, z_n) = z_1 \ldots z_n f_X(1 - \frac{1}{z_1}, \ldots, 1 - \frac{1}{z_n}),
\]

where \( f_X \) is the probability generating polynomial of \( X \).

By computing the coefficients of \( g_X \), we get

\[
g_X(z) = \sum_{A \subseteq [n]} (-1)^{|A|} P(A \subseteq X) z^A. \tag{2}
\]

By a straightforward calculation we see that for a determinantal process \( Y \) with kernel \( K \), we have \( g_Y(z) = \det(Z - K) \)—that is, the kernel polynomial of a determinantal process is the multivariate characteristic polynomial of its kernel. Note that, it is more natural to study a determinantal process through its kernel polynomial rather than the probability generating polynomial, whose connection to \( K \) is rather complicated.

In the following, we will show that in the case of strongly Rayleigh processes, the kernel polynomial retains many important features of the kernel matrix of determinantal processes. We expect that, at least in some settings, it is more appropriate to work with the kernel polynomial rather than the probability generating polynomial. As we will see in Section 4, the paving property is one such example, whose proof relies heavily on the notion of kernel polynomial.

**Proposition 2.12.** The kernel of a strongly Rayleigh point process is real stable.

**Proof.** Define operators \( \mathcal{T}, \mathcal{R} : \mathbb{C}[z_1, \ldots, z_n] \rightarrow \mathbb{C}[z_1, \ldots, z_n] \) as follows

\[
\mathcal{T}(p)(z_1, \ldots, z_n) = p(1 - z_1, \ldots, 1 - z_n),
\]

\[
\mathcal{R}(p)(z_1, \ldots, z_n) = z_1 \ldots z_n p(\frac{1}{z_1}, \ldots, \frac{1}{z_n}).
\]
Note that $\mathcal{T}$ is real stability preserving. Also, $\mathcal{R}$ is real stability preserving by part 4 of Proposition 2.4. The proposition follows since $g_X = \mathcal{R}(\mathcal{T}(f_X))$. □

For a determinantal process $\mathcal{Y}$ with kernel $K$ and every $B \subseteq [n]$, the matrix $K_B$ is the kernel of $\mathcal{Y} \cap B$, the restriction of $\mathcal{Y}$ to $B$. The following proposition is a generalization of this fact.

**Proposition 2.13.** Let $X$ be a strongly Rayleigh process on $[n]$ with kernel $g_X$. For every $B \subseteq [n]$, the polynomial $\partial^B g_X$ is the kernel polynomial of $X \cap B$.

**Proof.** By taking derivatives from both sides of (2), we get

$$\partial^B g_X(z) = \sum_{A \subseteq B} (-1)^{|A|} P(A \subseteq X) z^{B \setminus A}.$$ 

Since $P(A \subseteq X \cap B) = P(A \subseteq X)$ for every $A \subseteq B$, the above polynomial is the kernel polynomial of $X \cap B$. □

[20, Theorem 4.5.5] states that a Hermitian matrix $K$ is the kernel of a determinantal process if and only if it is a positive semidefinite contraction. The following proposition extends this result to strongly Rayleigh processes.

**Theorem 2.14.** Let $g \in \mathbb{R}[z_1, \ldots, z_n]$ be a multi-affine real stable polynomial. Then $g$ is the kernel of a strongly Rayleigh process if and only if $[z_1 \ldots z_n]_g = 1$ and all the roots of $\overline{g}$ are in the interval $[0, 1]$.

**Proof.** First we prove the “only if” part. Let $X$ be a strongly Rayleigh process with kernel $g_X$. It is clear from (2) that $[z_1 \ldots z_n]g_X = 1$. Assume that $\lambda_1, \ldots, \lambda_n$ are the roots of $\overline{g}_X$. We have

$$\overline{g}_X(x) = (x - \lambda_1) \ldots (x - \lambda_n).$$

It follows from the definition of kernel that

$$\overline{f}_X(x) = (1 - x)^n \overline{g}_X \left( \frac{1}{1 - x} \right) = (\lambda_1 x + 1 - \lambda_1) \ldots (\lambda_n x + 1 - \lambda_n). \quad (3)$$

Since the coefficients of $\overline{f}_X$ are nonnegative, we have $(1 - \lambda_i)/\lambda_i \geq 0$ for each nonzero $\lambda_i$. This implies that $\lambda_i \in [0, 1]$.

Now, consider the “if” part. Let $\mathcal{R}$ and $\mathcal{T}$ be as defined in the proof of Proposition 2.12 and $f := \mathcal{T}(\mathcal{R}(g))$. Assume $f(z) = \sum_{B \subseteq [n]} b_B z^B$. Note that $g$ is the kernel of a strongly Rayleigh process if and only if $f$ is a real stable probability generating polynomial. Since $\mathcal{R}$ and $\mathcal{T}$ are real stability preserving, $f$ is real stable. It remains to prove that $b_B \in [0, 1]$ for every $B \subseteq [n]$, and $\sum_{B \subseteq [n]} b_B = 1$.

Since $g$ is real stable and $[z_1 \ldots z_n] = 1$, the polynomial $\overline{g}$ is a monic real rooted polynomial of degree $n$. Let $\lambda_1, \ldots, \lambda_n$ denote the roots of this polynomial. By the definition of $f$,

$$\overline{f}(x) = (1 - x)^n \overline{g} \left( \frac{1}{1 - x} \right) = (\lambda_1 x + 1 - \lambda_1) \ldots (\lambda_n x + 1 - \lambda_n).$$

Therefore,

$$\sum_{B \subseteq [n]} b_B = f(1) = \overline{f}(1) = 1.$$
First, consider the case where \( \lambda_i \neq 0 \) for all \( i \in [n] \). Since \( f \) is multiaffine, we have \( b_{[n]} = \lambda_1 \cdots \lambda_n > 0 \). Also, since \( \lambda_i \in (0, 1] \), all of the roots of \( f \) are nonpositive. By applying Lemma 2.8 to the polynomial 
\[
(-1)^n f(-z),
\]
we conclude that \( b_{[n]} \geq 0 \) for all \( B \subseteq [n] \). In the general case, define the polynomial 
\[
g_{\varepsilon}(z) = (1 + \varepsilon)^{-n} g((1 + \varepsilon)z - \varepsilon 1)
\]
for every \( \varepsilon > 0 \). Note that \( g_{\varepsilon} \) is real stable, \( \prod_{i=1}^n g_{\varepsilon} = 1 \) and all of the roots of \( g_{\varepsilon} \) are in the interval \((0, 1]\). Therefore, \( T(R(g_{\varepsilon})) \) is a probability generating polynomial.

Since \( f = \lim_{\varepsilon \to 0} T(R(g_{\varepsilon})) \), it follows that \( f \) is a probability generating polynomial. This completes the proof.

We get the following result by comparing the coefficients of the two sides of (3).

**Proposition 2.15.** Let \( X \) be a strongly Rayleigh process on a set of size \( n \) and \( \lambda_1, \ldots, \lambda_n \) be the roots of \( \overline{g}_X \). By Theorem 2.14, \( \lambda_i \in [0, 1] \). Let \( I_1, \ldots, I_n \) be independent Bernoulli variables such that \( I_i \sim \text{Bernoulli}(\lambda_i) \). We have

\[
|X| \sim I_1 + \cdots + I_n.
\]

It is shown in [29, Lemma 4.1] that the size of a strongly Rayleigh process has the same distribution as a sum of independent Bernoulli variables. The above proposition describes this distribution in a canonical way and generalizes a similar result about determinantal processes. The special case of the above proposition for determinantal processes is proved in [20, Theorem 4.5.3]. In fact, [20, Theorem 4.5.3] is much stronger and provides a canonical probabilistic interpretation for the eigenvalues of the determinant of deterministic processes. A natural question is whether this result extends to strongly Rayleigh processes. We believe that such an extension requires a deep understanding of the structure of real stable polynomials. We will propose a conjecture in Subsection 4.1 which can be regarded as a first step in this direction.

3. Paving Property for Real Stable Polynomials

In this section we prove Theorem 1.4. As mentioned in Introduction, we will extend the main steps of Leake–Ravichandran’s proof [22] to real stable polynomials. To keep the paper concise, we will defer those arguments that work here with simple changes to the supplemental article [1].

Let \( \mathcal{P}_r(n) \) denote the set of all partitions of \([n]\) into \( r \), possibly empty, subsets. For a polynomial \( g \in R[z_1, \ldots, z_n] \) and \( S \in \mathcal{P}_r(n) \) with \( S = \{S_1, \ldots, S_r\} \), define \( g_S \in R[x] \) as

\[
g_S = \prod_{i=1}^r g_{S_i}
\]

and \( g_r \in R[x] \) as

\[
g_r = \sum_{S \in \mathcal{P}_r(n)} g_S.
\]

**Theorem 3.1.** Let \( g \) be as in Theorem 1.4. The polynomial \( g_r \) is real rooted and there exists a partition \( S \in \mathcal{P}_r(n) \) such that

\[
\maxroot(g_S) \leq \maxroot(g_r).
\]
The above theorem is a generalization of [22, Theorem 1.6] and the proof presented in [22, Section 2] works for this more general result with obvious changes. The idea of the proof is to construct an “interlacing family” of polynomials such that the set of “leaf-polynomials” is \( \{ g_S : S \in \mathcal{P}_r(n) \} \). The details of the proof can be found in Section A of the supplemental article [1].

Note that for \( S \in \mathcal{P}_r(n) \) with \( S = \{ S_1, \ldots, S_r \} \), we have

\[
\max \text{root}(g_S) = \max_{i \in [r]} \left( \max \text{root} \left( \frac{\partial S_i g}{\partial z_i} \right) \right).
\]

Therefore, for the partition \( \{ S_1, \ldots, S_r \} \in \mathcal{P}_r(n) \) given by Theorem 3.1, we have

\[
\forall i \in [r] : \max \text{root} \left( \frac{\partial S_i g}{\partial z_i} \right) \leq \max \text{root}(g_r).
\]

Thus, to prove Theorem 1.4, it is sufficient to show that

\[
\max \text{root}(g_r) \leq \left( \frac{1}{r} - \frac{\alpha}{r-1} + \sqrt{\alpha} \right)^2.
\] (4)

We will prove this inequality in Subsection 3.1 using Leake–Ravichandran’s version of “multivariate barrier method”. This version of the barrier method can be used to obtain (upper) bounds for the largest root of partial derivatives of a stable polynomial. The following proposition provides such an expression for \( g_r \).

**Proposition 3.2.** If \( g \) satisfies the assumptions of Theorem 1.4, then

\[
g_r(x) = \left( \frac{1}{(r-1)!} \right)^n \left[ \left( \prod_{i=1}^n \frac{\partial^{r-1} g(z)}{\partial z_i^{r-1}} \right) g(z) \right]_{z=x}^{r}.
\]

This proposition is a generalization of [22, Lemma 3.1] and the proof presented in [22] works here with obvious changes. For the sake of completeness, we provide a proof in Section B of the supplemental article [1].

### 3.1. The Barrier Method: Upper Bound for \( \max \text{root} (g_r) \)

In this subsection we prove inequality (4) using the multivariate barrier method. This method was introduced by Marcus, Spielman and Srivastava in [23] and provides a framework for analyzing the effect of differential operators on the “top” roots of real stable polynomials. We will employ Leake–Ravichandran’s version of the barrier method which adapts this machinery to the case of partial derivatives. We proceed as in [22, Section 4].

**Definition 3.3.** Given a real stable polynomial \( p \in \mathcal{H}_n(\mathbb{R}) \) and a point \( u \in \mathbb{R}^n \) with \( p(u) \neq 0 \), the barrier function in the direction \( i \) at \( u \), denoted \( \Phi_p^i(u) \), is defined by

\[
\Phi_p^i(u) = \frac{\partial_p}{p} (u).
\]

**Definition 3.4.** Let \( p \in \mathbb{R}[z_1, \ldots, z_n] \). A point \( u \in \mathbb{R}^n \) is above the roots of \( p \) if

\[
\forall w \geq u : p(w) \neq 0.
\]
We will use $\text{Ab}_p$ to denote the set of all the points above the roots of $p$.

In general, the set of the points above the roots of a polynomial can be empty, but this is not the case for real stable polynomials. This fact is stated in the following proposition. To avoid digression, we defer the proof of this proposition to Appendix A.

**Proposition 3.5.** If $p \in \mathbb{R}[z_1, \ldots, z_n]$ is real stable then $\text{Ab}_p$ is nonempty.

The main idea behind the barrier method is that the evolution of the points above the roots of a real stable polynomial under simple differential operators is governed by the barrier functions. For instance, it is known that if $i \in [n]$ and $\partial_i p \neq 0$, then $\Phi_i^p(u) > 0$ for every $u \in \text{Ab}_p$ (for a proof of this fact see [32, Lemmas 16 and 17]). In particular, we have $\partial_i p(u) \neq 0$, from which it follows that $\text{Ab}_p \subseteq \text{Ab}_{\partial_i p}$.

Let $g$ satisfy the assumptions of Theorem 1.4. Because of the expression given in Proposition 3.2 for $g_r$, to prove (4) it is sufficient to show that

$$\left( \sqrt{\frac{1}{r} - \alpha} + \sqrt{\alpha} \right)^2 1 \in \text{Ab}_{\partial_r^{-1} \ldots \partial_1^{-1} g^r}.$$ 

To prove this, we will start from a point above the roots of $g^r$ and follow its evolution under iterative applications of the operators $\partial_r^{-1}, \ldots, \partial_1^{-1}$ on $g^r$ until we obtain a point above the roots of $\partial_n^{-1} \ldots \partial_1^{-1} g^r$. In each iteration, the point is shifted along one of the axes and its displacements can be estimated in terms of the barrier functions. In order to apply this iterative argument, we need a control on the behavior of the barrier functions during the process. The following proposition is the key tool that we will use.

**Proposition 3.6.** Let $j \in [n]$ and $p \in \mathcal{H}_n(\mathbb{R})$ be a real stable polynomial of degree at most $r$ in $z_j$. If $u \in \text{Ab}_p$ and $\delta$ satisfies

$$0 \leq \delta \leq \frac{(r - 1)^2}{r} \left( \frac{1}{\Phi_j^p(u)} - \frac{1}{\lambda_r} \right),$$

where $\lambda_r$ is the smallest root of the univariate polynomial $p(u_1, \ldots, u_{j-1}, x, u_{j+1}, \ldots, u_n)$, then we have $u - \delta e_j \in \text{Ab}_{\partial_j^{-1} p}$ and

$$\Phi_{\partial_j^{-1} p}^i (u - \delta e_j) \leq \Phi_i^p(u)$$

for every $i \in [n]$.

The above proposition is an immediate consequence of [22, Propositions 4.2 and 4.3]. For the sake of completeness, we provide a complete proof in Section C of the supplemental article [1].

Proposition 3.6 relates the evolution of above the roots of a real stable polynomial to the barrier functions and the smallest roots of its one dimensional restrictions. Now, we estimate these quantities for a polynomial satisfying the assumptions of Theorem 1.4. We will derive a lower bound for $\lambda_r$ in Proposition 3.7 and an upper bound for $\Phi_j^p(u)$ in Proposition 3.10.
Proposition 3.7. Let \( g \) and \( r \) be as in Theorem 1.4 and \( p := g^r \). Also, let \( u \in \Lambda g \). Then, for every \( (i_1, \ldots, i_n) \in \mathbb{Z}^n \) such that \( 0 \leq i_k \leq r - 1 \) for all \( k \in [n] \), and every \( i \in [n] \), all the roots of the univariate polynomial

\[
q_i(x) := \left( \prod_{k=1}^{n} \sigma_{i_k}^{y_k} \right) (u_1, \ldots, u_{i-1}(x), u_{i+1}, \ldots, u_n)
\]

are nonnegative.

The above proposition is a generalization of [22, Proposition 4.6] and we take a similar route to [22] to prove it. We need the following two lemmas.

Lemma 3.8. Let \( p, q \in \mathbb{R}[z_1, \ldots, z_n] \setminus \{0\} \). If \( q \ll p \) and the leading coefficients of \( p \) and \( q \) have the same sign, then \( \Lambda p \subseteq \Lambda q \).

Proof. By Proposition 3.5, \( \Lambda p \) and \( \Lambda q \) are nonempty. Let \( a \in \Lambda p \) and \( b \in \Lambda q \). Note that for every \( \alpha \in \mathbb{R}^n \) and \( \eta \in \mathbb{R}^n \), we have \( t \eta + \alpha \geq a \) and \( t \eta + \alpha \geq b \) for large enough \( t \). Therefore, \( t \eta + \alpha \in \Lambda p \cap \Lambda q \) for large enough \( t \). Now, let \( a \in \Lambda p \) and \( b \in \mathbb{R}^n \) be such that \( u < w \). By the above observation, \( a \) is both in \( \Lambda p \cap \Lambda q \) for large enough \( t \). Since the sign of a polynomial does not change above its roots, we have \( \text{sgn}(p(t)) = \text{sgn}(p(t e + v)) \) and \( \text{sgn}(q(t)) = \text{sgn}(q(t e + v)) \) for large enough \( t \). Since the leading coefficients of \( p \) and \( q \) have the same sign, it follows that the leading coefficients of \( p(t e + v) \) and \( q(t e + v) \) have the same sign. On the other hand, since \( q \ll p \) and \( e \in \mathbb{R}^n_+ \), we have \( q(t e + v) \ll p(t e + v) \). Therefore, \( \lambda_1(q(t e + v)) \leq \lambda_1(p(t e + v)) \) (see the remarks following Proposition 2.4).

Since \( u \in \Lambda p \), we have \( \lambda_1(p(t e + v)) < 1 \) and hence \( \lambda_1(q(t e + v)) < 1 \). In particular, \( q(w) \neq 0 \). Overall, we have proved that \( q(w) \neq 0 \) for every \( w \in \mathbb{R}^n \) such that \( w \geq u \). This implies \( u \in \Lambda q \), which completes the proof.

Lemma 3.9. Let \( p \in \mathbb{R}[z_1, \ldots, z_n] \) with \( p(z) = \sum_{A \subseteq [n]} a_A z^A \) be a multiaffine real stable polynomial such that \( a_n > 0 \) and all of the roots of \( \overline{p} \) are nonnegative. If \( u \in \Lambda p \) then we have \( p(u_1, \ldots, u_{n-1}, x) = \alpha + \beta x \) where \( \alpha \leq 0 \) and \( \beta > 0 \).

Proof. Because \( p \) is multiaffine, we can write \( g = r + z_n s \), where \( r, s \in \mathbb{R}[z_1, \ldots, z_{n-1}] \). Thus \( g(u_1, \ldots, u_{n-1}, x) = r(u) + x s(u) \). From this point forth, we treat \( r \) and \( s \) as polynomials in \( \mathbb{R}[z_1, \ldots, z_{n-1}] \). As discussed in the remarks following Proposition 3.5, we have \( \Lambda p \subseteq \Lambda b p \) and hence \( u \in \Lambda b \). Note that the leading coefficient of \( \pi \) is \( a_{[n]} \), which is positive. Therefore, \( \lim_{r \to \infty} \pi(x) > 0 \). As highlighted in the proof of Lemma 3.8, we have \( x 1 \in \Lambda b \) for large enough \( x \). Since the sign of a real stable polynomial does not change above its roots, it follows that \( s(u) > 0 \).

If \( r = 0 \) then \( r(u) = 0 \) and there is nothing left to prove. Now assume \( r \neq 0 \). Since \( [n] \in \text{supp}(p) \) and \( r \neq 0 \), Corollary 2.9 implies that the maximum-degree monomial of \( r \) is \( a_{[n]} r z^{[n]} \). Hence, by Lemma 2.8, we have \( a_{[n]} r < 0 \) and so the leading coefficient of \( \pi \) is negative. Thus the leading coefficients of \( -\pi \) and \( \pi \) are both positive. On the other hand, by Proposition 2.6 we have \( s \ll r \); equivalently \( -r \ll s \). Therefore, by Lemma 3.8, we have \( \Lambda b \subseteq \Lambda r \). Hence \( u \in \Lambda r \) and by the same argument that we used above in the case of \( s \), it follows that \( r(u) < 0 \). This completes the proof.

Proof of Proposition 3.7. Since \( g \) is multiaffine,

\[
\left( \prod_{k=1}^{n} \sigma_{i_k}^{y_k} \right) p(z) = \sum_{(A_1, \ldots, A_r) \in A} \prod_{i=1}^{r} \sigma_{A_i} g(z),
\]
where $A$ is an appropriate subset of $(\mathbb{Z}^n)^r$ whose exact form is not important in this proof. Note that for every $B \subseteq [n]$ we have $\partial^B g(z) = \sum_{A \subseteq B} a_A z^{A \setminus B}$. Since all of the roots of $\bar{g}$ are nonnegative, the “only if” direction of Lemma 2.8 implies $(-1)^{|A|} a_A \geq 0$ for every $A \subseteq [n]$. Now, by applying the “if” direction of Lemma 2.8 to $\partial^B g(z)$, it follows that all of the roots of $\bar{g}$ are nonnegative. Therefore, by Lemma 3.9, each $\partial^A g$ is negative at every point $(u_1, \ldots, u_{i-1}, x, u_{i+1}, \ldots, u_n)$ with $x < 0$. Thus $q_i(x)$ is nonzero for all $x < 0$, which implies that all of the roots of $q_i$ are nonnegative.

The following proposition, which is a generalization of [22, Lemma 5.4], provides an upper bound for the barrier functions.

**Proposition 3.10.** Let $g$ and $r$ satisfy the assumptions of Theorem 1.4 and $p := g^r$. If $b \geq 1$, then

$$\forall i \in [n] : \Phi_i^p(b^1) \leq r \left( \frac{\alpha}{b - 1} + \frac{1 - \alpha}{b} \right).$$

To prove the above proposition, we need the following three lemmas.

**Lemma 3.11** (Lemma 9.B.3 of [25]). Given real numbers $c_1, \ldots, c_{n-1}$ and $\lambda_1, \ldots, \lambda_n$ satisfying the interlacing property

$$\lambda_1 \geq c_1 \geq \lambda_2 \geq \cdots \geq c_{n-1} \geq \lambda_n,$$

there exists a real symmetric $n \times n$ matrix of the form

$$W = \begin{bmatrix} D_c & v^t \\ v & v_n \end{bmatrix}$$

with eigenvalues $\lambda_1, \ldots, \lambda_n$.

**Lemma 3.12** (Lemma 5.1 of [22]). For any matrix $A \in \mathcal{M}_n(\mathbb{C})$ and any vector $v \in \mathbb{C}^n$, we have

$$\det(A_{v^\perp}) = (v^* A^{-1} v) \det(A),$$

where $A_{v^\perp} \in \mathcal{M}_{n-1}(\mathbb{C})$ is the compression of $A$ onto $v^\perp$.

**Lemma 3.13** (Lemma 5.3 of [22]). Let $A \in \mathcal{M}_n(\mathbb{C}$) be positive semidefinite contraction, $i \in [n]$ and $A_{i,i} \leq \alpha$. Then, for any $a \geq 1$,

$$e_i^a (aI - A)^{-1} e_i \leq \frac{\alpha}{a - 1} + \frac{1 - \alpha}{a}.$$

[22, Lemma 5.3] is slightly weaker than the above lemma but its proof only uses these weaker assumptions. Now, we are ready to prove Proposition 3.10.

**Proof of Proposition 3.10.** Let $\gamma_1 \geq \cdots \geq \gamma_n$ be the roots of $\bar{g}$ and $\delta_1 \geq \cdots \geq \delta_{n-1}$ be the roots of $\partial_i \bar{g}$. We have

$$
\Phi_i^p(b^1) = \frac{rg^{r-1} \partial_i g}{g^r}(b^1) = \frac{\partial_i \bar{g}(b)}{\bar{g}(b)} = \frac{\prod_{j=1}^{n-1}(b - \delta_j)}{\prod_{j=1}^{n-1}(b - \gamma_j)}.
$$

(6)
Since $\partial_ig \ll g$, we have $\gamma_1 \geq \delta_1 \geq \gamma_2 \geq \cdots \geq \delta_{n-1} \geq \gamma_n$. Thus, by Lemma 3.11, there is an $n \times n$ real symmetric matrix

$$A = \begin{bmatrix} D_\delta & v^t \\ v & v_n \end{bmatrix}$$

with $D_\delta = \text{Diag}(\delta_1, \ldots, \delta_{n-1})$, whose eigenvalues are $\gamma_1, \ldots, \gamma_n$. By Lemma 3.12,

$$\frac{\prod_{j=1}^{n-1} (b - \delta_j)}{\prod_{j=1}^{n} (b - \gamma_j)} = \frac{\det(bI_{n-1} - D_\delta)}{\det(bI_n - A)} = e_n^t (bI_n - A)^{-1} e_n,$$

(7)

where $I_k$ denotes the $k \times k$ identity matrix. We have

$$v_n = \sum_{j=1}^{n} \gamma_j - \sum_{j=1}^{n-1} \delta_j = [x^{n-1}] g - [x^{n-2}] \partial_ig = \sum_{j \in [n]} (-a_{\{j\}}) - \sum_{j \in [n] \setminus \{i\}} (-a_{\{j\}}) = -a_{\{i\}} \leq \alpha,$$

where in the second equality we have used the fact that $g$ and $\partial_ig$ are monic. Also by the assumption, $\gamma_j \in [0, 1]$ for all $j \in [n]$ and thus $A$ is a positive semidefinite contraction. Therefore, by Lemma 3.13,

$$e_n^t (bI_n - A)^{-1} e_n \leq \frac{\alpha}{b-1} + \frac{1 - \alpha}{b}. \quad (8)$$

The lemma follows from (6), (7) and (8).

The following lemma provides a subset of $Abg$ which will serve as the set of starting points for our iterative argument, over which we will optimize.

**Lemma 3.14.** Assume that $g \in \mathbb{R}[z_1, \ldots, z_n]$ is a multiaffine real stable polynomial such that $[z_1 \ldots z_n]g > 0$. Denote the largest root of $g$ by $\lambda$. Then, for every $b > \lambda$, the point $b1 \in \mathbb{R}^n$ lies above the roots of $g$.

**Proof.** Define $f(z_1, \ldots, z_n) = g(\lambda - z_1, \ldots, \lambda - z_n)$ and let $f(z) = \sum_{A \subseteq [n]} a_A z^A$. Note that $f$ is real stable. Since $\overline{f}(x) = g(\lambda - x)$, all of the roots of $\overline{f}$ are nonnegative. Thus by Lemma 2.8, we have $(-1)^{|A|} a_A \geq 0$ for every $A \subseteq [n]$. Therefore, $f(u) \neq 0$ for every $u \in (-\infty, 0)^n$; equivalently, $g(u) \neq 0$ for every $u \in (\lambda, \infty)^n$ and the lemma follows.

**Lemma 3.15** (Lemma 5.5 of [22]). For $\alpha, \beta \in [0, 1]$, we have

$$\inf_{a > 1} \left( a - \frac{\beta}{\alpha} \right) \left( \frac{1}{a-1} + \frac{1 - \alpha}{a} \right) = \begin{cases} \left( \sqrt{\alpha \beta} + \sqrt{(1-\alpha)(1-\beta)} \right)^2, & \alpha \leq \beta \\ \frac{1}{1}, & \alpha \geq \beta \end{cases}.$$
We have generalized all the tools that are used in [22] and so we can apply the proof of [22, Theorem 5.6] to deduce (4). For the sake of completeness we repeat the argument here.

**Theorem 3.16.** Let \( g_r \) be as in Proposition 3.2. We have

\[
\text{maxroot}(g_r) \leq \left( \sqrt{\frac{1}{r} - \frac{\alpha}{r-1}} + \sqrt{\alpha} \right)^2.
\]

**Proof.** Fix \( b > 1 \) and define \( w_0 \in \mathbb{R}^n \) as \( w_0 = b\mathbf{1} \). By Lemma 3.14, \( w_0 \) is above the roots of \( g^r \). Let \( p_0 = g^r \) and iteratively define

\[
p_k = \partial_r^{r-1} p_{k-1}, \quad k = 1, \ldots, n,
\]

and

\[
\delta_k = \frac{(r-1)^2}{r} \left( \frac{1}{\Phi_{p_{k-1}}(w_{k-1}) - \frac{1}{b}} \right) \quad \text{and} \quad w_k = w_{k-1} - \delta_k e_k.
\]

By Proposition 3.6 and Proposition 3.7, for every \( k \in [n] \) we have \( w_k \in \text{Ab}_{p_k} \) and

\[
\forall i \in [n] : \Phi^i_{p_k} (w_{k-1} - \delta_k e_k) \leq \Phi^i_{p_{k-1}} (w_{k-1}).
\]

It follows from Proposition 3.10 that

\[
\delta_k \geq \frac{(r-1)^2}{r} \left( \frac{1}{\Phi_{p_0}(b\mathbf{1}) - \frac{1}{b}} \right) \geq \frac{(r-1)^2}{r} \left( \frac{1}{r \left( \frac{\alpha}{b-1} + \frac{1-\alpha}{b} \right) - \frac{1}{b}} \right) =: \delta.
\]

Hence \( (b - \delta)\mathbf{1} \geq w_n \) and since \( w_n \in \text{Ab}_{p_n} \), we have \( (b - \delta)\mathbf{1} \in \text{Ab}_{p_n} \). Also, by Proposition 3.2, we have \( g_r = p_r \). Therefore,

\[
\text{maxroot}(g_r) \leq \inf_{b > 1} \left\{ b - \frac{(r-1)^2}{r} \left( \frac{1}{r \left( \frac{\alpha}{b-1} + \frac{1-\alpha}{b} \right) - \frac{1}{b}} \right) \right\}
\]

\[
= \inf_{b > 1} \left\{ b - \frac{r - 1}{r} \left( \frac{1}{r \alpha/(r-1) + \frac{1-r\alpha/(r-1)}{b-1}} \right) \right\}.
\]

By Lemma 3.15, if \((r - 1)^2/r^2 \geq \alpha\) then

\[
\text{maxroot}(g_r) \leq \left( \sqrt{\frac{1}{r} - \frac{\alpha}{r-1}} + \sqrt{\alpha} \right)^2.
\]

\[\square\]
4. Paving Property for Strongly Rayleigh Processes

Our main goal in this section is to prove Theorem 1.5. An important step in the proof is to use the polynomial paving property. The connection between these two results is more apparent in the case of determinantal processes. In this case, the polynomial paving property reduces to the classical paving conjecture. The following proposition is the paving property for zero-diagonal Hermitian matrices.

**Proposition 4.1** (Corollary 26 of [32]). Let $\Lambda$ be a positive number and $r$ be an integer such that $r \geq 2$. For every Hermitian matrix $A \in \mathcal{M}_n(\mathbb{C})$ with vanishing diagonal and $\|A\|_{op} \leq \Lambda$, there are diagonal projections $P_1, \ldots, P_r \in \mathcal{M}_n(\mathbb{C})$ such that \[ \sum_{i=1}^r P_i = I_n \] and \[ \forall i \in [r^2] : \|P_i A P_i\|_{op} \leq \left( \frac{2\sqrt{2}}{\sqrt{r}} + \frac{1}{r} \right) \Lambda. \]

Let $K$ be the kernel of a determinantal process $\mathcal{Y}$ and $D := \text{Diag}(K)$. By applying the above proposition to $K - D$, we conclude that for every positive $\varepsilon$, there is a positive integer $r$ such that it is possible to partition $[n]$ into $r$ subsets $S_1, \ldots, S_r$ such that \[ \|K_{S_i} - D_{S_i}\|_{op} \leq \varepsilon. \]

Note that for every $A \subseteq [n]$, the matrix $K_A$ is the kernel of $\mathcal{Y} \cap A$, and a determinantal process has independent points if and only if its kernel is a diagonal matrix. Hence we can interpret the above inequality as the restrictions of the determinantal process to each $S_i$ having “almost independent points”.

Motivated by the above argument, the first step in the proof of Theorem 1.5 will be to apply a generalization of Proposition 4.1 to the kernel polynomial of the given strongly Rayleigh process. The following proposition is the appropriate generalization of Proposition 4.1.

**Proposition 4.2.** Let $r \geq 4$ be an integer and $\Lambda \in \mathbb{R}_+$. Assume that $g \in \mathbb{R}[z_1, \ldots, z_n]$ is a multiaffine real stable polynomial and $g(z) = \sum_{A \subseteq [n]} a_A z^A$. If all the roots of $\overline{g}$ are in $[-\Lambda, \Lambda]$, $a_{\emptyset} = 1$ and $a_{\{i\}} = 0$ for $i = 1, \ldots, n$, then there exists a partition $S_1, \ldots, S_r$ of $[n]$ such that \[ \forall i \in [r^2] : \text{M} \left( \frac{\partial^{n_i} g}{\partial z_i^n} \right) \leq \left( \frac{r - 2}{r(r - 1)} + \frac{\sqrt{r - 2}}{r(r - 1)} \right) \Lambda, \]

where $\text{M}(\cdot)$ is the maximum absolute value of roots.

The proof of this result is similar to the proof presented in [32] for Proposition 4.1 and can be found in Section D of the supplemental article [1]. To complete the proof of Theorem 1.5, we will need a relationship between the entropy of a strongly Rayleigh process and the entropy of the roots of its kernel. This is presented in Subsection 4.1.

### 4.1. An Entropy Lower Bound

Let $\mathcal{X}$ be a strongly Rayleigh process on $[n]$. Recall the Bernoulli variables $I_1, \ldots, I_n$ from Proposition 2.15 for which $|\mathcal{X}| \sim I_1 + \cdots + I_n$. In this subsection we prove that the entropy of $\mathcal{X}$ is greater than or equal to the entropy of $(I_1, \ldots, I_n)$. Note that an obvious lower bound for the entropy of $\mathcal{X}$ is $H(|\mathcal{X}|)$. Our result provides a stronger lower bound.
Lemma 4.4. Between interlacing and majorization. For more information on the theory of majorization see [25].

Theorem 4.3. Let $X$ be a strongly Rayleigh process on $[n]$ with kernel $g_X$. We have

$$H(X) \geq \sum_{i=1}^{n} h(\lambda_i),$$

where $\lambda_1, \ldots, \lambda_n$ are the roots of $\overline{g}_X$.

In the proof of the above theorem, we will we will exploit the connection between interlacing and “majorization” in order to prove entropy inequalities. Given $u, v \in \mathbb{R}^n$, we say that $u$ majorizes $v$, denoted $u \succ v$, if $\sum_{i=1}^{n} u_i \geq \sum_{i=1}^{n} v_i$ and

$$\forall k \in [n-1]: \sum_{i=1}^{k} u_{(i)} \geq \sum_{i=1}^{k} v_{(i)},$$

where $u_{(i)}$ (respectively, $v_{(i)}$) is the the $i$-th largest component of $u$ (respectively, $v$). A function $\phi : \mathbb{R}^n \to \mathbb{R}$ is Schur-concave if $\phi(u) \leq \phi(v)$ for every $u, v \in \mathbb{R}^n$ such that $u \succ v$. A well-known example of Schur-concave functions is the entropy function. The following lemma establishes the relationship between interlacing and majorization. For more information on the theory of majorization see [25].

Lemma 4.4 (5.B.4 of [25]). If $b_1 \geq \cdots \geq b_{n-1}$ interlaces $a_1 \geq \cdots \geq a_n$, then

$$(a_1, \ldots, a_n) \succ (b_1, \ldots, b_{n-1}, b^*)$$

namely $(a_1, \ldots, a_n)$ majorizes $(b_1, \ldots, b_{n-1}, b^*)$, where $b^* = \sum_{i=1}^{n} a_n - \sum_{i=1}^{n-1} b_i$.

We will also need the following two lemmas in the proof of Theorem 4.3.

Lemma 4.5 (Lemma 3.2 of [9]). Let $p(z_1, z_2) = a_{11}z_1z_2 + a_{10}z_1 + a_{01}z_2 + a_{00} \in \mathbb{R}[z_1, z_2] \setminus \{0\}$. Then, $p \in \mathcal{H}_2(\mathbb{R})$ if and only if $\det[a_{ij}] \leq 0$.

Recall that the class of strongly Rayleigh processes is closed under conditioning and projection.

Lemma 4.6. Let $X$ be a strongly Rayleigh process on $[n]$ with kernel polynomial $g$. Denote the kernel polynomials of $(X \cap [n-1] \mid n \in X)$ and $(X \cap [n-1] \mid n \notin X)$ by $g_1$ and $g_0$, respectively. Then, $\overline{g}_1 \ll \overline{g}$ and $\overline{g}_0 \ll \overline{g}$.

The proof of this lemma is technical and to avoid digression, we defer it to Appendix B.

Proof of Theorem 4.3. We use induction on $n$. The base case is $n = 2$. Denote the probability generating polynomial of $X$ by $f_X$ and let $f_X(z_1, z_2) = a_{11}z_1z_2 + a_{10}z_1 + a_{01}z_2 + a_{00}$. Recall that $f_X(x) = (\lambda_1 x + 1 - \lambda_1)(\lambda_2 x + 1 - \lambda_2)$. Therefore,

$$a_{11}x^2 + (a_{10} + a_{01})x + a_{00} = \lambda_1 \lambda_2 x^2 + (\lambda_1(1 - \lambda_2) + \lambda_2(1 - \lambda_1))x + (1 - \lambda_1)(1 - \lambda_2).$$

By comparing the coefficients, we get

$$a_{11} = \lambda_1 \lambda_2,$$  \hspace{1cm} (9)

$$a_{10} + a_{01} = \lambda_1(1 - \lambda_2) + \lambda_2(1 - \lambda_1),$$  \hspace{1cm} (10)

$$a_{00} = (1 - \lambda_1)(1 - \lambda_2).$$  \hspace{1cm} (11)
By Lemma 4.5 we have \( a_{10}a_{01} \geq a_{11}a_{00} = ((1 - \lambda_1)\lambda_1)((1 - \lambda_2)\lambda_2) \). Because of (10), it follows that \((1 - \lambda_1)\lambda_1, (1 - \lambda_2)\lambda_2 \rangle \succ (a_{10}, a_{01}).\) Now, in view of (9) and (11), we can conclude that
\[
(\lambda_1\lambda_2, \lambda_1(1 - \lambda_2), \lambda_2(1 - \lambda_1), (1 - \lambda_1)(1 - \lambda_2)) \succ (a_{11}, a_{10}, a_{01}, a_{00}).
\]
Since entropy is a Schur-concave function, we have
\[
H(\mathcal{X}) = H(a_{11}, a_{10}, a_{01}, a_{00}) \geq H(\lambda_1\lambda_2, \lambda_1(1 - \lambda_2), \lambda_2(1 - \lambda_1), (1 - \lambda_1)(1 - \lambda_2)) = H(I_1, I_2).
\]

Now assume that the statement is true for \( n - 1 \). Let \( \mathcal{X}^\prime = \mathcal{X} \cap [n - 1] \) and denote the nonincreasing vectors of the roots of the kernel polynomials of \((\mathcal{X}^\prime \mid n \in \mathcal{X})\) and \((\mathcal{X}^\prime \mid n \notin \mathcal{X})\) by \( \gamma \) and \( \delta \), respectively. By Lemma 4.6, \( \gamma \) and \( \delta \) both interlace \( \lambda \), where \( \lambda \) is the nonincreasing vector of the roots of \( \mathcal{Y}_\mathcal{X} \).

By the induction hypothesis,
\[
H(\mathcal{X}^\prime \mid n \in \mathcal{X}) \geq \sum_{i=1}^{n-1} h(\gamma_i) \quad \text{and} \quad H(\mathcal{X}^\prime \mid n \notin \mathcal{X}) \geq \sum_{i=1}^{n-1} h(\delta_i).
\]
Therefore,
\[
H(\mathcal{X}) = p_n H(\mathcal{X}^\prime \mid n \in \mathcal{X}) + (1 - p_n) H(\mathcal{X}^\prime \mid n \notin \mathcal{X}) + h(p_n)
\geq p_n \left( \sum_{i=1}^{n-1} h(\gamma_i) \right) + (1 - p_n) \left( \sum_{i=1}^{n-1} h(\delta_i) \right) + h(p_n).
\]  
(12)

Let \( \alpha = \sum_{i=1}^{n} \lambda_i - \sum_{i=1}^{n-1} \gamma_i \) and \( \beta = \sum_{i=1}^{n} \lambda_i - \sum_{i=1}^{n-1} \delta_i \). Since \( \gamma \) and \( \delta \) both interlace \( \lambda \), we have \( \alpha \geq 0 \) and \( \beta \geq 0 \). Note that
\[
\sum_{i=1}^{n-1} \gamma_i = \mathbb{E}[|\mathcal{X}^\prime| \mid n \in \mathcal{X}] = \mathbb{E}[|\mathcal{X}| \mid n \in \mathcal{X}] - 1 \quad \text{and} \quad \sum_{i=1}^{n-1} \delta_i = \mathbb{E}[|\mathcal{X}^\prime| \mid n \notin \mathcal{X}] = \mathbb{E}[|\mathcal{X}| \mid n \notin \mathcal{X}].
\]

By conditioning on the inclusion of \( n \) in \( \mathcal{X} \) and using the fact that \( \sum_{i=1}^{n} \lambda_i = \mathbb{E}[|\mathcal{X}|] \), we get
\[
p_n \alpha + (1 - p_n) \beta = p_n.
\]  
(13)

Since \( \beta \geq 0 \), it follows from the above equation that \( \alpha \leq 1 \). Also, by negative dependence, \( \mathbb{E}[|\mathcal{X}^\prime| \mid n \notin \mathcal{X}] \geq \mathbb{E}[|\mathcal{X}^\prime|] \) which implies that
\[
\beta = \mathbb{E}[|\mathcal{X}|] - \mathbb{E}[|\mathcal{X}^\prime| \mid n \notin \mathcal{X}] \leq (\mathbb{E}[|\mathcal{X}^\prime|] + p_n) - \mathbb{E}[|\mathcal{X}^\prime|] \leq 1.
\]

Thus \( h(\alpha) \) and \( h(\beta) \) are well-defined.

Now, by Lemma 4.4, we have \( \lambda \succ (\gamma, \alpha) \) and \( \lambda \succ (\delta, \beta) \). Since the entropy function is concave and sum of concave functions is Schur-concave,
\[
\sum_{i=1}^{n-1} h(\gamma_i) + h(\alpha) \geq \sum_{i=1}^{n} h(\lambda_i),
\]
\[
\sum_{i=1}^{n-1} h(\delta_i) + h(\beta) \geq \sum_{i=1}^{n} h(\lambda_i).
\]
Therefore,
\[
p_n \left( \sum_{i=1}^{n-1} h(\gamma_i) + (1 - p_n) \sum_{i=1}^{n-1} h(\delta_i) \right) + \left( p_n h(\alpha) + (1 - p_n) h(\beta) \right) \geq \sum_{i=1}^{n} h(\lambda_i). \tag{14}
\]

Also, by (13) and concavity of \( h \) we have
\[
h(p_n) \geq p_n h(\alpha) + (1 - p_n) h(\beta). \tag{15}
\]

The result follows from (12), (14) and (15).

A stronger result holds for (discrete) determinantal processes.

**Proposition 4.7.** Let \( \mathcal{Y} \) be a determinantal process on \([n]\) with kernel \( K \) and \( \lambda_1, \ldots, \lambda_n \) be the eigenvalues of \( K \). If \( I_1, \ldots, I_n \) are independent Bernoulli variables such that \( I_i \sim \text{Bernoulli}(\lambda_i) \), then the probability vector corresponding to \( \mathcal{Y} \) is majorized by the probability vector corresponding to \((I_1, \ldots, I_n)\)—namely
\[
(\mathbb{P}(I = A) : A \subseteq [n]) \succ (\mathbb{P}(\mathcal{Y} = A) : A \subseteq [n]),
\]
where \( I = (I_1, \ldots, I_n) \).

Since entropy is Schur-concave, this result is stronger than Theorem 4.3 in the case of determinantal processes. Our proof for the above proposition relies on several facts from the theory of determinantal processes. To avoid digression, we only present a sketch of proof here and defer a full proof to Section E of the supplemental article [1]. By the spectral decomposition, we have \( K = \sum_{i=1}^{n} \lambda_i v_i v_i^* \), where \( v_1, \ldots, v_n \) are orthonormal. Define \( K_I = \sum_{i=1}^{n} I_i v_i v_i^* \) and let \( X_I \) be the (random) determinantal process with kernel \( K_I \). [20, Theorem 4.5.3] states that \( X_I \sim X \). This implies
\[
(\mathbb{P}(\mathcal{Y} = A) : A \subseteq [n]) = (\mathbb{P}(I = A) : A \subseteq [n]) M,
\]
where \( M \) is a \( 2^n \times 2^n \) matrix with \( M(A, B) = \mathbb{P}(X_I = B | I = A) \) for \( A, B \subseteq [n] \). We know from the theory of majorization that (16) holds if and only if \( M \) is doubly stochastic (see [25]). It is obvious that \( \sum_{B \subseteq [n]} M(A, B) = 1 \) for each \( A \subseteq [n] \). It remains to prove that \( \sum_{A \subseteq [n]} \mathbb{P}(X_I = B | I = A) = 1 \) for every \( B \subseteq [n] \). This is done using the geometric description of the law of “determinantal projection process”—and the observation that \( [X_I | I = A] \) is a from this class.

It is natural to expect that Proposition 4.7 extends to strongly Rayleigh processes.

**Conjecture 4.8.** Let \( X \) be a strongly Rayleigh process on \([n]\) with kernel \( g_X \) and \( \lambda_1, \ldots, \lambda_n \) be the roots of \( g_X \). Assuming \( \lambda = (\lambda_1, \ldots, \lambda_n) \), we have
\[
(\lambda^A (1 - \lambda)^{A^c} : A \subseteq [n]) \succ (\mathbb{P}(X = A) : A \subseteq [n])
\]
where \( \lambda = (\lambda_1, \ldots, \lambda_n) \) and \( \lambda^A (1 - \lambda)^{A^c} = \prod_{i \in A} \lambda_i \prod_{i \in A^c} (1 - \lambda_i) \).

The above conjecture can be regarded as a first step in generalizing [20, Theorem 4.5.3]. This conjecture is equivalent to the existence of a doubly stochastic \( 2^n \times 2^n \) matrix \( M \) such that
\[
(\mathbb{P}(X = A) : A \subseteq [n]) = (\lambda^A (1 - \lambda)^{A^c} : A \subseteq [n]) M.
\]
A full description of the entries of this matrix will lead to a generalization of [20, Theorem 4.5.3] to strongly Rayleigh processes.
4.2. Proof of the Paving Property for Strongly Rayleigh Processes

In this subsection we prove Theorem 1.5. The following is a corollary of Proposition 4.2.

Corollary 4.9. For every positive \( \varepsilon \), there is an integer \( r \) such that for every strongly Rayleigh process \( \mathcal{X} \) on any space \( S \), it is possible to partition \( S \) into \( r \) subsets \( S_1, \ldots, S_r \) such that

\[
\forall i \in [r] : M(\xi_i) \leq \varepsilon, \quad \xi_i(z) := g_{\mathcal{X}}(z + p),
\]

where \( M(.) \) denotes the maximum absolute value of roots and \( p = (p_j)_{j\in S} \) with \( p_j = P(j \in \mathcal{X}) \).

Proof. Without loss of generality, we can assume \( S = [n] \). Define \( \xi(z) = g_{\mathcal{X}}(z + p) \). We claim that \( \xi \) satisfies the assumptions of Proposition 4.2 with \( \Lambda = 1 \). It is straightforward to verify that \( \xi \) is multiaffine real stable, \( [z_1 \ldots z_n] \xi = 1 \) and \( [z_1 \ldots z_{i-1} z_{i+1} \ldots z_n] \xi = 0 \) for all \( i \in [n] \). Now, we prove \( M(\xi) \leq 1 \). By Theorem 2.14, we have \( \lambda_i(\overline{g}_{\mathcal{X}}) \in (0, 1] \), for \( i = 1, \ldots, n \). Since \( \lambda(\overline{g}_{\mathcal{X}}) \leq 1 \), Lemma 3.14 implies that \( b_1 \in \text{Ab}_{g_{\mathcal{X}}} \) for every \( b \in (1, \infty) \). Let \( b > 1 \) and \( u \geq b_1 \). Since \( p \geq 0 \), we have \( u + p \geq b_1 \) and hence \( \xi(u) = g_{\mathcal{X}}(u + p) \neq 0 \). Therefore, \( b_1 \in \text{Ab}_{\xi} \) for every \( b \in (1, \infty) \), which implies that \( \lambda_i(\xi) \leq 1 \) for \( i = 1, \ldots, n \).

Since \( \lambda(\overline{g}_{\mathcal{X}}) \geq 0 \), by applying Lemma 3.14 to \( (-1)^ng_{\mathcal{X}}(1 - z) \) we conclude that \( b_1 \in \text{Ab}_{g_{\mathcal{X}}}(-x) \) for every \( b \in (0, \infty) \). Let \( b > 1 \) and \( u \geq b_1 \). Since \( p \leq 1 \), we have \( u - p \geq 0 \) and hence \( \xi(u) = g_{\mathcal{X}}(u + p) = g_{\mathcal{X}}(u - p) \neq 0 \). Therefore, \( b_1 \in \text{Ab}_{\xi}(x) \) for every \( b \in (1, \infty) \), which implies that \( \lambda_i(\xi) \geq -1 \), for \( i = 1, \ldots, n \). Overall, we showed that all of the roots of \( \xi \) lie in \([-1, 1]\) and so \( \xi \) satisfies the assumptions of Proposition 4.2 with \( \Lambda = 1 \). The corollary follows by applying Proposition 4.2 to \( \xi \) and choosing a large enough \( r \).

\( \square \)

Theorem 1.5 follows by applying the following proposition to each \( \mathcal{X} \cap S_i \), where \( S_i \) are given by the above corollary.

Proposition 4.10. Assume that \( \mathcal{X} \) is a strongly Rayleigh process on \([n]\) with kernel \( g_{\mathcal{X}} \). Let \( p_i = P(i \in \mathcal{X}) \) and \( \xi(z_1, \ldots, z_n) = g_{\mathcal{X}}(z_1 + p_1, \ldots, z_n + p_n) \). For every positive \( \delta \), there exists a positive \( \varepsilon \) such that if all the roots of \( \xi \) have absolute value less than \( \varepsilon \), then

\[
\left| \frac{1}{n} H(\mathcal{X}) - \frac{1}{n} \sum_{i=1}^{n} b(p_i) \right| < \delta.
\]

We will use majorization properties of hyperbolic polynomials to prove the above proposition. A homogeneous polynomial \( q \in \mathbb{R}[z_1, \ldots, z_n] \) is hyperbolic with respect to a vector \( e \in \mathbb{R}^n \) if \( q(e) > 0 \) and \( q(te + \alpha) \in \mathbb{R}[t] \) is real rooted for all \( \alpha \in \mathbb{R}^n \). We use \( \lambda_q(q) \) to denote the vector of roots of the polynomial \( q(te + \alpha) \), in the nonincreasing order. The following theorem is proved in [18].

Theorem 4.11. Let \( q \in \mathbb{R}[z_1, \ldots, z_n] \) be hyperbolic with respect to \( e \). For \( v, u \in \mathbb{R}^n \) we have

\[
\lambda_{v+u}(q) \preceq \lambda_{v}(q) + \lambda_{u}(q).
\]

We will use the following lemma in the proof of Proposition 4.10.

Lemma 4.12. Let \( p_i \) be the vector of \( p_i \)’s in the nonincreasing order. We have

\[
\lambda(\overline{g}_{\mathcal{X}}) \prec \lambda(\xi) + p_i^\perp.
\]
Lemma 4.12 that since the entropy function is concave and sum of concave functions is Schur-concave, it follows from \( \lambda \).

This concludes the proof.

Proof. Suppose that \( \xi(z) = \sum_{A \subseteq [n]} b_A z^A e \). Since \( g_X(z_1, \ldots, z_n) = \xi(z_1 - p_1, \ldots, z_n - p_n) \), we have

\[
g_X(z) = \sum_{A \subseteq [n]} \left( \sum_{B \subseteq A} b_B (-p)_B A \right) z^A e.
\]

Define the polynomial \( F \in \mathbb{R}[z_1, \ldots, z_n, u_1, \ldots, u_n, w] \) as

\[
F(z, u, w) = \sum_{A \subseteq [n]} \left( \sum_{B \subseteq A} b_B u_{B|A} z^A e \right) w^A,
\]

where \( u = (u_1, \ldots, u_n) \) and \( z = (z_1, \ldots, z_n) \). We claim that \( F \) is hyperbolic with respect to \( e \in \mathbb{R}^{2n+1} \), where \( e_1 = \cdots = e_n = 1 \) and \( e_{n+1} = \cdots = e_{2n+1} = 0 \). Let \( z = (z, u, w) \in \mathbb{R}^{2n+1} \). If \( w \neq 0 \), then

\[
F(te + z) = w^n \xi \left( \frac{1}{w} (t + z_1 + u_1), \ldots, \frac{1}{w} (t + z_1 + u_1) \right).
\]

Since \( \xi \) is real stable and \( u_1, \ldots, u_n, w \in \mathbb{R} \), the above polynomial is real rooted. If \( w = 0 \), then

\[
F(te + z) = \sum_{A \subseteq [n]} u_A (t1 + z)^A e = \prod_{i=1}^n (t + z_i + u_i).
\]

The above polynomial is also real rooted. This completes the proof of our claim. Now, by Theorem 4.11,

\[
\lambda_{(0,-p,1)}(F) \times \lambda_{(0,0,1)}(F) + \lambda_{(0,-p,0)}(F),
\]

where \( 0 = (0, \ldots, 0) \in \mathbb{R}^n \). It is straightforward to verify that

\[
\lambda_{(0,-p,1)}(F) = \lambda(\overline{g}_X), \quad \lambda_{(0,0,1)}(F) = \lambda(\overline{X}), \quad \lambda_{(0,-p,0)}(F) = p_1.
\]

This concludes the proof.

Now we are ready to prove Proposition 4.10.

Proof of Proposition 4.10. Let \( \gamma_1, \ldots, \gamma_n \), indexed in nonincreasing order, be the roots of \( \overline{\xi} \) and \( \lambda_1, \ldots, \lambda_n \), indexed in nonincreasing order, be the roots of \( \overline{g}_X \). Also assume \( p_i = (p_1, \ldots, p(n)) \).

Since the entropy function is concave and sum of concave functions is Schur-concave, it follows from Lemma 4.12 that

\[
\sum_{i=1}^n h(\gamma_i + p_i) \leq \sum_{i=1}^n h(\lambda_i). \tag{17}
\]

Choose \( \varepsilon \) such that if \( |x - y| < \varepsilon \), then \( |h(x) - h(y)| < \delta \). Therefore, if \( |\gamma_i| < \varepsilon \) for all \( i \in [n] \), then by (17),

\[
\sum_{i=1}^n h(\lambda_i) \geq \sum_{i=1}^n h(\gamma_i + p_i) \geq \sum_{i=1}^n h(p_i) - n\delta. \tag{18}
\]
On the other hand, by Theorem 4.3,
\[
\sum_{i=1}^{n} h(\lambda_i) \leq H(\mathcal{X}) \leq \sum_{i=1}^{n} h(p_i). \tag{19}
\]
Combining (17), (18) and (19) we get
\[
\sum_{i=1}^{n} h(p_i) - n\delta \leq H(\mathcal{X}) \leq \sum_{i=1}^{n} h(p_i).
\]
This completes the proof. \qed

**Appendix A: Proof of Proposition 3.5**

**Proposition A.1.** If \( p \in \mathbb{R}[z_1, \ldots, z_n] \) is real stable then \( A_b p \) is nonempty.

This proposition is a direct consequence of the following observation of Tao [32].

**Proposition A.2.** Let \( p(z_1, z_2) \) be a real stable polynomial. Note that for every \( a \in \mathbb{R} \), the polynomial \( p(a, z) \) is either identically zero or real rooted and the former case occurs for finitely many \( a \). The largest root of \( p(a, \cdot) \) is a locally nonincreasing function of \( a \) on the set \( \{ a \in \mathbb{R} : p(a, \cdot) \neq 0 \} \).

For a proof of the above proposition see [32, proof of Lemma 17] or [22, Proposition 4.1].

**Proof of Proposition A.1.** We use induction on \( n \). The base case, \( n = 1 \), is trivial. Now assume that the statement is true for \( n - 1 \). Let \( p \in \mathcal{H}_n(\mathbb{R}) \) and \( \deg_n(p) \neq 0 \). We can write \( p = \sum_{i=0}^{r} z_i q_i \), where \( r = \deg_n(p) \) and \( q_0, \ldots, q_r \in \mathbb{R}[z_1, \ldots, z_{n-1}] \). Since \( q_r = (1/r!) \partial_{n}^{r} p \), by part 1 of Proposition 2.4, \( q_r \) is either identically zero or real stable. Since \( \deg_n(p) = r \), we have \( q_r \neq 0 \) and hence \( q_r \) is real stable. Thus, by the induction hypothesis, \( A_b q_r \) is nonempty. Let \( a \in A_b q_r \). For every \( b \in \mathbb{R}^{n-1} \) with \( b \geq a \), we have \( q_r(b) \neq 0 \) and hence \( p(b, \cdot) \neq 0 \). Therefore, by Proposition A.2, \( \lambda_1(p(b, \cdot)) \) is nonincreasing in \( b_i \) for every \( i \in [n-1] \), on the set \( \{ b \in \mathbb{R}^{n-1} : b \geq a \} \). Consequently, \( \lambda_1(p(b, \cdot)) \) is a nonincreasing function of \( b \) on the set \( \{ b \in \mathbb{R}^{n-1} : b \geq a \} \)—that is, if \( b_1 \geq b_2 \geq a \) then \( \lambda_1(p(b_1, \cdot)) \leq \lambda_1(p(b_2, \cdot)) \).

We claim that \( (a, \gamma) \in A_b p \) for every \( \gamma \in \mathbb{R} \) with \( \gamma > \lambda_1(p(a, \cdot)) \). Let \( b \in \mathbb{R}^{n-1} \) and \( \delta \in \mathbb{R} \) such that \( (b, \delta) \geq (a, \gamma) \); namely, \( b \geq a \) and \( \delta \geq \gamma \). We proved above that \( \lambda_1(p(b, \cdot)) \leq \lambda_1(p(a, \cdot)) \). Therefore, \( \delta \geq \gamma > \lambda_1(p(a, \cdot)) \geq \lambda_1(p(b, \cdot)) \). Thus \( p(b, \delta) \neq 0 \) and our claim follows. \qed

**Appendix B: Proof of Lemma 4.6**

**Lemma B.1.** Let \( \mathcal{X} \) be a strongly Rayleigh process on \([n]\) with kernel polynomial \( g \). Denote the kernel polynomials of \((\mathcal{X} \cap [n-1] \mid n \in \mathcal{X}) \) and \((\mathcal{X} \cap [n-1] \mid n \not\in \mathcal{X})\) by \( g_1 \) and \( g_0 \), respectively. Then, \( \mathcal{Y}_1 \ll \mathcal{Y} \) and \( \mathcal{Y}_0 \ll \mathcal{Y} \).

**Proof.** Let \( \lambda = \lambda(\mathcal{Y}) \), \( \gamma = \lambda(\mathcal{Y}_1) \) and \( \delta = \lambda(\mathcal{Y}_0) \). By Theorem 2.14, \( \lambda_i \in [0, 1] \) for all \( i \in [n] \). First, we consider the case where \( \lambda_i \in (0, 1) \) for all \( i \in [n] \). Let \( f \) be the probability generating polynomial of \( \mathcal{X} \). Denote the probability generating polynomials of \((\mathcal{X} \cap [n-1] \mid n \in \mathcal{X}) \) and \((\mathcal{X} \cap [n-1] \mid n \not\in \mathcal{X}) \) by \( f_1 \) and \( f_0 \), respectively. Note that \( f = z_n(p_n f_1) + (1 - p_n) f_0 \) and \( f_1, f_0 \in \mathbb{R}[z_1, \ldots, z_{n-1}] \).
By Proposition 2.6 we have $f_1 \ll f$ which implies $f_1 \ll f$. On the other hand,

$$
\tilde{f}(x) = (1 - x)^n \gamma \left( \frac{1}{1 - x} \right) = (\lambda_1 x + 1 - \lambda_1) \ldots (\lambda_n x + 1 - \lambda_n),
$$

$$
\tilde{f}_1(x) = (1 - x)^n \gamma_1 \left( \frac{1}{1 - x} \right) = (\gamma_1 x + 1 - \gamma_1) \ldots (\gamma_{n-1} x + 1 - \gamma_{n-1}).
$$

Since $\lambda_i \in (0, 1)$ for all $i \in [n]$, the polynomial $\tilde{f}$ is of degree $n$ and all its roots are negative. Since $\tilde{f}_1 \ll f$ and $\deg(f_1) < \deg(f)$, the polynomial $\tilde{f}_1$ is of degree $n - 1$ (consequently, $\gamma_i > 0$ for all $i \in [n - 1]$) and its roots interlace the roots of $\tilde{f}$. Note that the roots of $\tilde{f}$ are $(\lambda_i - 1)/\lambda_i$ for $i = 1, \ldots, n$, and the roots of $\tilde{f}_1$ are $(\gamma_i - 1)/\gamma_i$ for $i = 1, \ldots, n - 1$. Since the function $(x - 1)/x$ is strictly increasing on $\mathbb{R}_+$, it follows that $\gamma$ interlaces $\lambda$ which implies $\tilde{f}_1 \ll \tilde{f}$.

Now we prove $f_0 \ll f$. For every multiaffine polynomial $p \in \mathbb{R}[z_1, \ldots, z_n]$ define

$$
\mathcal{R}_n(p) = z_1 \ldots z_n p(z_1^{-1}, \ldots, z_n^{-1}).
$$

By part 4 of Proposition 2.4, $\mathcal{R}$ is a real stability preserving operator. Since $\mathcal{R}$ is linear, we have $\mathcal{R}_n(f) = p_n \mathcal{R}_{n-1}(f_1) + z_n ((1 - p_n) \mathcal{R}_{n-1}(f_0))$. Therefore, $\mathcal{R}_{n-1}(f_0) \ll \mathcal{R}_n(f)$ which implies $\mathcal{R}_n(f_0) \ll \mathcal{R}_n(f)$. On the other hand,

$$
\mathcal{R}_n(f)(x) = x^n \tilde{f} \left( \frac{1}{x} \right) = ((1 - \lambda_1)x + \lambda_1) \ldots ((1 - \lambda_n)x + \lambda_n),
$$

$$
\mathcal{R}_{n-1}(f_0)(x) = x^{n-1} \tilde{f}_0 \left( \frac{1}{x} \right) = ((1 - \delta_1)x + \delta_1) \ldots ((1 - \delta_{n-1})x + \delta_{n-1}).
$$

By an argument similar to the one used above to prove $\tilde{f}_1 \ll \tilde{f}$, we can deduce $f_0 \ll f$.

In general, every kernel polynomial can be approximated by kernel polynomials whose diagonalizations have roots in the interval $(0, 1)$. For instance, this can be achieved by the polynomials $(1 + 2\varepsilon)^{-n} g((1 + 2\varepsilon)z - \varepsilon 1)$, where $\varepsilon > 0$. The general case follows since interlacing is preserved under taking limits.

**Acknowledgements**

We would like to express our deepest appreciation to Amir Daneshgar and Mohmadsadegh Zamani for their valuable and instructive comments on an early draft of this paper. We also wish to thank Ziheng Zhu who pointed out a technical mistake in an earlier version of this paper.

**Supplementary Material**

Supplement to “Paving property for real stable polynomials and strongly Rayleigh processes”
In the supplemental article [1], we provide proofs that are omitted due to their similarity to arguments presented in [22] and [32]. Proofs of Theorem 3.1, Proposition 3.2, Proposition 3.6 and Proposition 4.2 are included. Also, a full proof of Proposition 4.7 is presented.
References


