ORTHANT PROBABILITIES AND THE ATTAINMENT OF MAXIMA ON A VERTEX OF A SIMPLEX

DAMIÁN PINASCO, EZEQUIEL SMUCLER, AND IGNACIO ZALDUENDO

Abstract. We calculate bounds for orthant probabilities for the equicorrelated multivariate normal distribution and use these bounds to show the following: for degree \( k \geq 4 \), the probability that a \( k \)-homogeneous polynomial in \( n \) variables attains a relative maximum on a vertex of the \( n \)-dimensional simplex tends to one as the dimension \( n \) grows. The bounds we obtain for the orthant probabilities are tight up to \( \log(n) \) factors.

Introduction

A number of papers have appeared addressing the question of the probability that an \( n \)-variable polynomial \( P : \mathbb{R}^n \to \mathbb{R} \) attain relative maxima on specific points ([3], [9], [10], [11]). It has been proved for instance, that for degree \( k \geq 2 \), the probability that a \( k \)-homogeneous polynomial in \( n \) variables attain a relative maxima on some vertex of the unit sphere of \( \ell_1^n \) tends to one as \( n \) increases [10], and that this is false for \( k = 2 \) [9].

To consider such problems a probability measure is defined on the space of \( k \)-homogeneous \( n \)-variable polynomials \( \mathcal{P}^k(\mathbb{R}^n) \) (we recall the necessary constructions in section 1). Functions \( X : \mathcal{P}^k(\mathbb{R}^n) \to \mathbb{R} \) are then random variables and the questions regarding relative maxima may be set in terms of inequalities involving some of these random variables.

In this paper we consider the case of relative maxima attained on a vertex of the zero-centered simplex. This translates into a classical problem in Statistics, that of computing orthant probabilities: let \( X_1, \ldots, X_n \) be jointly normally distributed random variables with mean \( \mu = 0 \), standard deviation \( \sigma = 1 \), and \( \mathbb{E}(X_iX_j) = \rho_{ij} \); then what is the probability that they be simultaneously positive? Sheppard [14] proved that \( P(X_1 > 0, X_2 > 0) = \frac{1}{4} + \frac{(2\pi)^{-1}}{4} \arcsin(\rho_{12}) \). David [4] develops a recursive formula that allows one to compute \( P(X_1 > 0, \ldots, X_n > 0) \) for \( n \) odd as a function of orthant probabilities of lower dimension. Applying this formula to \( n = 3 \) yields \( P(X_1 > 0, X_2 > 0, X_3 > 0) = 1/8+1/(4\pi) (\arcsin(\rho_{12}) + \arcsin(\rho_{13}) + \arcsin(\rho_{23})) \). No closed form formulas for \( n \geq 4 \) are known. Series expansions for the general case were derived in [13]. Steck [15] studies the special case in which \( \mathbb{E}(X_iX_j) = \rho \), which is the setting of interest in the present work, and lists a number of interesting properties of the function \( f(n, \rho) = P(X_1 > 0, \ldots, X_n > 0) \), including the identity \( f(n, \rho) = \mathbb{E} \left[ \Phi^n(Z\sqrt{\rho/(1-\rho)}) \right] \), where \( Z \) is a standard normal random variable and \( \Phi \) is its cumulative distribution function. In [11] there is a formula for \( f(n, \rho) \) in terms of the gaussian measure of a simplex. A more thorough review of

2010 Mathematics Subject Classification. 26D05, 46G25, 62H05.

Key words and phrases. Homogeneous Polynomial, Orthant Probabilities, Simplex.

Supported partially by PIP 112 201301 00422 CO (CONICET - Argentina).
results regarding orthant probabilities can be found in [8]. In this paper, we will show that for \( \rho > 0 \) and large \( n \), \( f(n, \rho) \) behaves essentially like \( n^{1-1/\rho} \), except for multiplicative \( \log(n) \) factors.

A central issue pertaining to the attainment of maxima at several different vertices is the independence or non-independence of the random variables involved. This accounts for the essential difference in the case of \( k > 2 \) or \( k = 2 \) mentioned above for the unit sphere of \( \ell_1^n \). It is also the main difficulty in the study of the analogous problem on the unit sphere of \( \ell_\infty^n \). As we will show in section 2, the particular geometry of the simplex gives rise to a situation where the random variables involved are not independent, but their dependence is weak. We use the Devroye-Mehrabian-Reddad total-variation bound [5] to approximate asymptotically the non-independent situation by independent ones.

This paper is organized as follows. In Section 1 we recall [11] the construction of the probability measure in \( P^k(\mathbb{R}^n) \), and present some important random variables and their correlations. In Section 2 we present the necessary geometry of the simplex, study the attainment of maxima on a given vertex and give bounds for orthant probabilities. Finally, in Section 3 we consider the problem of attaining maxima at any vertex and prove that for \( k > 4 \) the probability of attaining a relative maxima at some vertex of the simplex tends to one as the dimension grows.

1. Probability in spaces of polynomials

Given a \( k \)-homogeneous polynomial \( P : \mathbb{R}^n \rightarrow \mathbb{R} \) there exists a unique symmetric \( k \)-linear function \( \phi \) that satisfies \( P(x) = \phi(x, \ldots, x) \). We treat the space of \( k \)-homogeneous polynomials over \( \mathbb{R}^n \) as the dual of the symmetric tensor product \( \otimes_k \mathbb{R}^n \). The Hilbert space structure on the full tensor product \( \otimes_k \mathbb{R}^n \) is given by the inner product

\[
\langle v^1 \otimes \cdots \otimes v^k, w^1 \otimes \cdots \otimes w^k \rangle = \langle v^1, w^1 \rangle \cdots \langle v^k, w^k \rangle.
\]

See [6] for details.

Let the symmetrization operator \( S : \otimes_k \mathbb{R}^n \rightarrow \otimes_k \mathbb{R}^n \) be given by setting its values on a basis as

\[
S(e_{j_1} \otimes \cdots \otimes e_{j_k}) = \frac{1}{k!} \sum_\sigma e_{j_\sigma(1)} \otimes \cdots \otimes e_{j_\sigma(k)},
\]

where \( \sigma \) runs through all permutations of \( \{1, \ldots, k\} \). The image of \( S \)—the symmetric tensor product \( \otimes_k \mathbb{R}^n \)—is a predual of the space of polynomials over \( \mathbb{R}^n \), \( P^k(\mathbb{R}^n) \). We consider on \( \otimes_k \mathbb{R}^n \) the Hilbert space structure induced by the ambient space \( \otimes_k \mathbb{R}^n \), and on \( P^k(\mathbb{R}^n) \) the dual Hilbert space structure. The resulting norm on \( P^k(\mathbb{R}^n) \) is the Bombieri norm [1]

\[
\|P\| = \left( \sum_{|\alpha| = k} a_\alpha^2 \frac{\alpha!}{k!} \right)^{1/2},
\]

if \( P(x) = \sum_{|\alpha| = k} a_\alpha x^\alpha \) is the monomial-sum expression of \( P \).

All linear forms \( \varphi : P^k(\mathbb{R}^n) \rightarrow \mathbb{R} \) can be identified with an element of \( \otimes_k \mathbb{R}^n \). For instance, evaluation at \( x \): \( e_x(P) = P(x) \) is given by \( x \otimes \cdots \otimes x \). In the present
work, we will mainly deal with linear forms
\[ \frac{\partial}{\partial v}(a) : \mathcal{P}^k(\mathbb{R}^n) \rightarrow \mathbb{R} \text{ given by } P \mapsto \frac{\partial P}{\partial v}(a), \]
which can be identified with an element of \( \bigotimes_{k,s} \mathbb{R}^n \) as follows
\[ \frac{\partial}{\partial v}(a) = v \otimes a \otimes \cdots \otimes a + a \otimes v \otimes a \otimes \cdots \otimes a + \cdots + a \otimes \cdots \otimes a \otimes v. \]
The inner products between such linear forms are given by
\[ \langle \frac{\partial}{\partial v}(a), \frac{\partial}{\partial w}(b) \rangle = \langle v \otimes a \otimes \cdots \otimes a + \cdots + a \otimes \cdots \otimes a \otimes v, w \otimes b \otimes \cdots \otimes b + \cdots + b \otimes \cdots \otimes b \otimes w \rangle = k \langle v, w \rangle \langle a, b \rangle - \langle v, b \rangle \langle a, w \rangle \langle a, b \rangle. \]
If \( \langle a, b \rangle = 0 \), the display above is zero whenever \( k > 2 \). However, when \( k = 2 \) one has \( 2 \langle a, w \rangle \langle v, b \rangle \) which can be non-zero. In fact, if \( k = 2 \) and \( n = 2 \),
\[ \frac{\partial}{\partial (1,0)} (0,1) \text{ and } \frac{\partial}{\partial (0,1)} (1,0) \]
are the same linear form.

We consider on \( \mathcal{P}^k(\mathbb{R}^n) \) the standard Gaussian measure \( W \) corresponding to its Hilbert space structure, i.e., the measure
\[ W(B) = \frac{1}{(2\pi)^{d/2}} \int_B e^{-\|P\|^2/2} \, dP, \]
for any Borel set \( B \subset \mathcal{P}^k(\mathbb{R}^n) \),
where \( d = \binom{n+k-1}{k} \), is the dimension of \( \mathcal{P}^k(\mathbb{R}^n) \). \( W \) is rotation-invariant. Moreover, if \( T : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is an orthogonal transformation, then
\[ \tilde{T} : \mathcal{P}^k(\mathbb{R}^n) \rightarrow \mathcal{P}^k(\mathbb{R}^n) \text{ such that } \tilde{T}(P) = P \circ T \]
is a measure-preserving map.

Finally, note that if \( \varphi : \mathcal{P}^k(\mathbb{R}^n) \rightarrow \mathbb{R} \) is a linear form then \( \varphi \) is a normal random variable with zero mean and standard deviation given by \( \| \varphi \| \).

2. The simplex and orthant probabilities

We construct the zero-centered \( n \)-dimensional simplex in the following way. Take
\[ \Delta^+ = \left\{ x \in \mathbb{R}^{n+1} : x_i \geq 0 \text{ for } i = 1, \ldots, n \text{ and } \sum_{i=1}^{n+1} x_i = 1 \right\}. \]
Note that \( \Delta^+ \) is the convex hull of the canonical basis \( \{e_1, \ldots, e_{n+1}\} \) of \( \mathbb{R}^{n+1} \). The center of \( \Delta^+ \) is \( z = \left( \frac{1}{n+1}, \ldots, \frac{1}{n+1} \right) \). We now translate \( \Delta^+ \) by subtracting \( z \): \( \Delta = \Delta^+ - z \). \( \Delta \) and \( \Delta^+ \) are isometric copies of one another, and \( \Delta \) is contained in the subspace
\[ S = \left\{ x \in \mathbb{R}^{n+1} : \sum_{i=1}^{n+1} x_i = 0 \right\}. \]
We identify \( S \) isometrically with \( \mathbb{R}^n \) and call \( \Delta \) the zero-centered \( n \)-dimensional simplex. The identification of \( S \) with \( \mathbb{R}^n \) determines how \( \Delta \) is immersed in \( \mathbb{R}^n \), but as long as its center is at 0, its position will be irrelevant to us, for one can be taken to another by an isometry.
We will need to calculate the angles that two edges form at a vertex, and also the angles between vertices. For this we refer to $\Delta^+$. Let $a, a_1, \ldots, a_n$ be the vertices of $\Delta$, and set $v_i = \frac{a-a_i}{\|a-a_i\|}$ for $i = 1, \ldots, n$.

- Angles between edges at a vertex:
  \[
  \arccos \frac{\langle v_i, v_j \rangle}{\|v_i\|\|v_j\|} = \arccos \frac{\langle e_k - e_i, e_k - e_j \rangle}{\|e_k - e_i\||e_k - e_j\|}
  \]
  \[
  = \arccos \frac{|e_k|^2 - \langle e_k - e_i, e_k - e_j \rangle}{\|e_k - e_i\|^2}
  \]
  \[
  = \arccos \frac{1}{2} = 60^\circ.
  \]

- Angles (at zero) between vertices:
  \[
  \arccos \frac{\langle a_i, a_j \rangle}{\|a_i\||a_j\|} = \arccos \frac{\langle e_i - z, e_j - z \rangle}{\|e_i - z||e_j - z\|}
  \]
  \[
  = \arccos \frac{\langle e_i, e_j \rangle - \langle e_i, z \rangle - \langle z, e_j \rangle + \|z\|^2}{\|e_i - z\|^2}
  \]
  \[
  = \arccos \frac{-\frac{2}{n+1} + \frac{4}{n+1}}{n+1}
  \]
  \[
  = \arccos \frac{-1}{n}.
  \]

We will also need to calculate $\langle v_i, a \rangle$ and $\|a\|$:

\[
\langle v_i, a \rangle = \frac{\langle e_k - e_i, e_k - z \rangle}{\|e_k - e_i\|} = \frac{1}{\sqrt{2}} \left( 1 - \langle e_k, z \rangle - \langle e_i, e_k \rangle + \langle e_i, z \rangle \right) = \frac{1}{\sqrt{2}}.
\]

\[
\|a\|^2 = \langle e_k - z, e_k - z \rangle = \langle e_k, e_k \rangle - \langle e_k, z \rangle - \langle z, e_k \rangle + \|z\|^2 = 1 - \frac{2}{n+1} + \frac{1}{n+1} = \frac{n}{n+1}.
\]

Thus, the zero-centered $n$-dimensional simplex has $n+1$ vertices, in each of which incide $n$ edges at $60^\circ$ angles between each other. The angles formed (at zero) by different vertices are all equal to $\arccos \frac{-1}{n}$, and are therefore closer and closer to perpendicular as the dimension $n$ grows. This will be an important feature for our calculations.

A homogeneous polynomial $P : \mathbb{R}^n \rightarrow \mathbb{R}$ will have (relative to $\Delta$) a maximum at the vertex $a$ if its gradient is “outward pointing”,

that is, if $\frac{\partial P}{\partial v_i}(a) = \nabla P(a)(v_i) > 0$ for $i = 1, \ldots, n$. Thus our problem of relative maximum at $a$ translates into orthant probability: we want the measure of the subset of $\mathcal{P}_k(\mathbb{R}^n)$ on which

$X_1(a) = \frac{\partial}{\partial v_1}(a) = \frac{\partial}{\partial v_1}(a)\|\frac{\partial}{\partial v_1}(a)\|, \ldots, X_n(a) = \frac{\partial}{\partial v_n}(a) = \frac{\partial}{\partial v_n}(a)\|\frac{\partial}{\partial v_n}(a)\|

are simultaneously positive. In order to calculate the covariance matrix of the random vector $(X_1(a), \ldots, X_n(a))$ we need the following.

$$\left\| \frac{\partial}{\partial v_i}(a) \right\|^2 = \left\langle \frac{\partial}{\partial v_i}(a), \frac{\partial}{\partial v_i}(a) \right\rangle = k\langle v_i, v_i \rangle (a, a)^{k-1} + (k^2 - k)\langle v_i, a \rangle^2 (a, a)^{k-2}$$

$$= k\|a\|^2k^{k-2} + (k^2 - k)\frac{1}{2}\|a\|^{2k-4}$$

$$= k\|a\|^2k^{k-4} \left( \|a\|^2 + \frac{k-1}{2} \right).$$

$$\left\langle \frac{\partial}{\partial v_i}(a), \frac{\partial}{\partial v_j}(a) \right\rangle = k\langle v_i, v_j \rangle (a, a)^{k-1} + (k^2 - k)\langle v_i, a \rangle \langle a, v_j \rangle (a, a)^{k-2}$$

$$= \frac{1}{2}k\|a\|^2k^{k-2} + (k^2 - k)\frac{1}{2}\|a\|^{2k-4}$$

$$= k\|a\|^2k^{k-4} \left( \frac{1}{2}\|a\|^2 + \frac{k-1}{2} \right).$$

In section 3 we will consider maxima at different vertices, $a$ and $b$. For this, we need the following bound (here the $w_j$s define the edges inciding at $b$).
\[
\left\langle \frac{\partial}{\partial v_i}(a), \frac{\partial}{\partial w_j}(b) \right\rangle = |k(v_i, w_j)(a, b)^{k-1} + (k^2 - k)(v_i, b)(a, w_j)(a, b)^{k-2}|
\]
\[
= |k(v_i, w_j) \left( \|a\|^2 \frac{1}{n} \right)^{k-1} + (k^2 - k)(v_i, b)(a, w_j) \left( \|a\|^2 \frac{1}{n} \right)^{k-2}|
\]
\[
\leq k\|a\|^{2k-2} \left( \frac{1}{n} \right)^{k-1} + (k^2 - k)\|a\| \|b\|^{2k-4} \left( \frac{1}{n} \right)^{k-2}|
\]
\[
= k\|a\|^{2k-4} \left( \|a\|^2 \frac{1}{n^{k-1}} + (k-1)\|a\| \frac{1}{n^{k-2}} \right)
\]

Thus, after normalization, we have \( \|X_i(a)\| = 1 \), so \( X_i(a) \) are \( N(0, 1) \) random variables, with correlation

\[
\langle X_i(a), X_j(a) \rangle = \frac{k\|a\|^{2k-4} \left( \frac{1}{n} \right)^{k-1} + (k^2 - k)\|a\| \|b\|^{2k-4} \left( \frac{1}{n} \right)^{k-2}]{2\|a\|^2 + (k-1)}
\]
\[
\frac{n k + (k-1)}{n(k+1) + (k-1)}
\]
\[
= \rho_n.
\]

Note that \( \frac{1}{2} < \rho_n < 1 \) and \( \rho_n \to k/(k+1) \). We also have the bound

\[
|\langle X_i(a), X_j(b) \rangle| \leq \frac{\|a\|^2 \left( \frac{1}{n} \right)^{k-1} + (k-1)\|a\| \|b\|^2 \left( \frac{1}{n} \right)^{k-2}]{2\|a\|^2 + (k-1)}
\]
\[
= \frac{1}{n^{k-2} 2k-1} \left( \frac{1}{n} + (k-1) \right)
\]

Recall that for \( x > 0 \) the gamma function is given by

\[
\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt.
\]

For \( a, b > 0 \) the beta function is given by

\[
B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt.
\]

It is well known that

\[
B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}
\]

For \( Z \sim N(0, 1) \) we let \( \Phi \) be the cumulative distribution function of \( Z \). We will also use the notation \( \phi(x) = \Phi'(x) \). We now prove the following bounds for orthant probabilities.

**Theorem 2.1.** Let \( \rho \in (0, 1) \) and let \((X_1, \ldots, X_n)\) be a multivariate normal random vector with zero means, variances equal to one and correlations equal to \( \rho \). Let \( f(n, \rho) = P(X_1 > 0, \ldots, X_n > 0) \).
(1) There exists a universal constant \(c_1 > 0\) such that if \(\rho > 1/2\) and \(n \geq 2\), then

\[
\frac{c_1}{\sqrt{2 + (1/\rho - 1) \log(n)}} \leq \frac{f(n, \rho)}{n^{1-1/\rho}}.
\]

and

\[
f(n, \rho) \leq \frac{1}{2^{2-1/\rho}} \sqrt{\frac{1-\rho}{\rho}} \{1 + B(2, 1/\rho - 1)\}.
\]

(2) If \(\rho < 1/2\) and \(n\) satisfies \(n \geq (1/\rho - 1) \log(n)/\log(2)\), then

\[
f(n, \rho) \leq \frac{1}{2^{2-1/\rho}} \sqrt{\frac{1-\rho}{\rho}} \left\{ \Gamma(1/\rho - 1) - 1 \right\} \leq \frac{f(n, \rho)}{n^{1-1/\rho}}.
\]

Moreover, there exists \(n_0 = n_0(\rho)\) such that if \(n \geq n_0\) then

\[
f(n, \rho) \leq \frac{1}{2^{2-1/\rho}} \sqrt{\frac{1-\rho}{\rho}} \left\{ (1/\rho - 1) \log(n) \right\}^{1/\rho - 2}.
\]

Proof. Set \(s = \rho/(1 - \rho)\). By equation (2.5) in [15]

\[
f(n, \rho) = \mathbb{E} \left[ \Phi^n \left( Z \sqrt{s} \right) \right].
\]

We begin with the proof of part (1). Note that by Gordon’s inequality [7], for all positive \(x\),

\[1 - \Phi(x) < \frac{\phi(x)}{x}\]

and that by Birnbaum’s inequality [2]

\[
\frac{2\phi(x)}{\sqrt{4 + x^2} + x} < 1 - \Phi(x).
\]

Let \(\alpha_n = n^{1/\rho-1}\). By Markov’s inequality we have

\[
\mathbb{E} \left[ \Phi^n \left( Z \sqrt{s} \right) \right] \geq \Phi^n \left( \sqrt{2 \log(\alpha_n) s} \right) \mathbb{P} \left( \Phi^n \left( Z \sqrt{s} \right) \geq \Phi^n \left( \sqrt{2 \log(\alpha_n) s} \right) \right)
\]

\[
= \Phi^n \left( \sqrt{2 \log(\alpha_n) s} \right) \mathbb{P} \left( Z \geq \sqrt{2 \log(\alpha_n)} \right).
\]

Using Birnbaum’s lower bound we obtain

\[
\mathbb{P} \left( Z \geq \sqrt{2 \log(\alpha_n)} \right) \geq \frac{2\phi(\sqrt{2 \log(\alpha_n)})}{\sqrt{4 + 2 \log(\alpha_n) + 2 \log(\alpha_n)}}
\]

\[
= \frac{2-1}{\sqrt{2\pi} \alpha_n \sqrt{4 + 2 \log(\alpha_n) + 2 \log(\alpha_n)}}.
\]

Recall that \(\alpha_n = n^{1/\rho-1}\), so

\[
\mathbb{P} \left( Z \geq \sqrt{2 \log(\alpha_n)} \right) \geq n^{1-1/\rho} \frac{2}{\sqrt{2\pi} \sqrt{4 + 2(1/\rho - 1) \log(n) + 2(1/\rho - 1) \log(n)}}.
\]

On the other hand, using Gordon’s upper bound, we obtain

\[
\Phi^n \left( \sqrt{2 \log(\alpha_n) s} \right) \geq \left\{ 1 - \frac{\phi(\sqrt{2 \log(\alpha_n) s})}{\sqrt{2 \log(\alpha_n) s}} \right\}^n = \left\{ 1 - \frac{\alpha_n^{-s}}{\sqrt{4\pi \log(\alpha_n) s}} \right\}^n = \left\{ 1 - \frac{1}{n \sqrt{4\pi \log(n)}} \right\}^n.
\]
Therefore, for \( n \geq 2 \)
\[
\Phi^n \left( \sqrt{2 \log(n)} \right) \geq \left\{ 1 - \frac{1}{\sqrt{4\pi \log(n)}} \right\}^n \geq \left\{ 1 - \frac{1}{2\sqrt{4\pi \log(2)}} \right\}^2.
\]
We conclude that
\[
f(n, \rho) \geq n^{1-1/\rho} \frac{1}{\sqrt{4 + 2(1/\rho - 1) \log(n)}} \frac{2}{\sqrt{2\pi}} \left\{ 1 - \frac{1}{2\sqrt{4\pi \log(2)}} \right\}^2.
\]
Now (2.2) follows immediately from the last display.

We turn now to the proof of (2.3). First, some preliminaries. Let
\[ U = \Phi \left( Z \sqrt{s} \right). \]
Note that \( U \) takes values in \([0,1]\). We will calculate the density function of \( U \), \( f_U \).

Take \( u \in [0,1] \). Then
\[
P(U \leq u) = P(\Phi \left( Z \sqrt{s} \right) \leq u) = P\left( Z \leq \Phi^{-1}(u)/\sqrt{s} \right).
\]
Differentiating this last display with respect to \( u \) we get
\[
f_U(u) = \frac{1}{\sqrt{s}} \phi \left( \Phi^{-1}(u)/\sqrt{s} \right).
\]
Simplifying, we get
\[
f_U(u) = \frac{1}{\sqrt{s}} \exp \left\{ \frac{-1}{2s} (\Phi^{-1}(u))^2 + \frac{1}{2} (\Phi^{-1}(u))^2 \right\} = \frac{1}{\sqrt{s}} \exp \left\{ \frac{1}{2} (\Phi^{-1}(u))^2 \left( 1 - \frac{1}{s} \right) \right\}
\]
\[
= \left( \frac{\sqrt{2\pi}}{\sqrt{s}} \right)^{1/s-1} \left[ \phi \left( \Phi^{-1}(u) \right) \right]^{1/s-1}.
\]
Hence, by (2.6)
\[
(2.7) \quad f(n, \rho) = \frac{(\sqrt{2\pi})^{1/s-1}}{\sqrt{s}} \int_0^1 x^n \left[ \phi \left( \Phi^{-1}(x) \right) \right]^{1/s-1} dx.
\]
We claim that
\[
(2.8) \quad \min(x, 1-x) \sqrt{\frac{2}{\pi}} \leq \phi \left( \Phi^{-1}(x) \right)
\]
By symmetry it suffices to show that \( \phi \left( \Phi^{-1}(x) \right) \geq \sqrt{2/\pi} x \) for \( x \in [0,1/2] \). This follows from the fact that \( \phi \left( \Phi^{-1}(x) \right) \) is concave, and equal to \( \sqrt{2/\pi} x \) for \( x = 0 \) and \( x = 1/2 \).

We are now ready to prove (2.3). The assumption \( \rho > 1/2 \) implies \( s > 1 \) and hence \( 1/s - 1 < 0 \). Now write
\[
\int_0^1 x^n \left[ \phi \left( \Phi^{-1}(x) \right) \right]^{1/s-1} dx = \int_0^{1/2} x^n \left[ \phi \left( \Phi^{-1}(x) \right) \right]^{1/s-1} dx + \int_{1/2}^1 x^n \left[ \phi \left( \Phi^{-1}(x) \right) \right]^{1/s-1} dx.
\]
Using the lower bound in (2.8) we get

\[
\int_0^{1/2} x^n \left[ \phi \left( \Phi^{-1}(x) \right) \right]^{1/s-1} \, dx + \int_{1/2}^1 x^n \left[ \phi \left( \Phi^{-1}(x) \right) \right]^{1/s-1} \, dx \leq \\
\left( \sqrt{\frac{\pi}{2}} \right)^{1-1/s} \left\{ \int_0^{1/2} x^n x^{1/s-1} \, dx + \int_{1/2}^1 x^n (1-x)^{1/s-1} \, dx \right\} = \\
\left( \sqrt{\frac{\pi}{2}} \right)^{1-1/s} \left\{ \left( \frac{1}{2} \right)^{n+1/s} \frac{1}{n+1/s} + \int_{1/2}^1 x^n (1-x)^{1/s-1} \, dx \right\} \leq \\
\left( \sqrt{\frac{\pi}{2}} \right)^{1-1/s} \left\{ \left( \frac{1}{2} \right)^{n+1/s} \frac{1}{n+1/s} + B(n+1, 1/s) \right\}.
\]

Thus

\[
f(n, \rho) \leq \frac{1}{2^{1-1/s} \sqrt{s}} \left\{ \left( \frac{1}{2} \right)^{n+1/s} \frac{1}{n+1/s} + B(n+1, 1/s) \right\}.
\]

Replacing \( s = \rho/(1-\rho) \) yields

\[
f(n, \rho) \leq \frac{1}{2^{2-1/\rho}} \sqrt{\frac{1-\rho}{\rho}} \left\{ \left( \frac{1}{2} \right)^{n+1/\rho-1} \frac{1}{n+1/\rho-1} + B(n+1, 1/\rho - 1) \right\}.
\]

Now, for all \( n, 2^{-n-1/s} (n+1)^{-1} \leq n^{1-1/\rho} \). Recall that

\[
B(n+1, 1/s) = \frac{\Gamma(n+1)\Gamma(1/s)}{\Gamma(n+1+1/s)}.
\]

Using inequality (3.75) in [12] with \( x = n, a = 1, b = 1/s + 1, x_0 = 1 \), we have

\[
\frac{\Gamma(n+1)\Gamma(1/s)}{\Gamma(n+1+1/s)} \leq n^{1-1/s} \frac{\Gamma(2)\Gamma(1/s)}{\Gamma(2+1/s)} = n^{1-1/\rho} B(2, 1/s) = n^{1-1/\rho} B(2, 1/\rho - 1).
\]

Thus we have

\[
f(n, \rho) \leq n^{1-1/\rho} \frac{1}{2^{2-1/\rho}} \frac{1-\rho}{\rho} \left\{ 1 + B(2, 1/\rho - 1) \right\},
\]

which is precisely (2.3). This finishes the proof of part (1).
Next we prove part (2). Since \( \rho < 1/2 \), we have \( s < 1 \) and hence \( 1/s - 1 > 0 \).

Using the lower bound in (2.8) we get
\[
\int_{0}^{1/2} x^n \left[ \phi \left( \Phi^{-1}(x) \right) \right]^{1/s-1} dx + \int_{1/2}^{1} x^n \left[ \phi \left( \Phi^{-1}(x) \right) \right]^{1/s-1} dx \geq \left( \sqrt{\frac{\pi}{2}} \right)^{1-1/s} \int_{1/2}^{1} x^n (1-x)^{1/s-1} dx = \left( \sqrt{\frac{\pi}{2}} \right)^{1-1/s} \int_{1/2}^{1} x^n (1-x)^{1/s-1} dx + B(n+1,1/s) \geq \left( \sqrt{\frac{\pi}{2}} \right)^{1-1/s} \frac{1}{2n} \int_{0}^{1/2} (1-x)^{1/s-1} dx + B(n+1,1/s) = \left( \sqrt{\frac{\pi}{2}} \right)^{1-1/s} \left\{ -\frac{s}{2n} + \frac{s}{2n+1} + B(n+1,1/s) \right\} \geq \left( \sqrt{\frac{\pi}{2}} \right)^{1-1/s} \left\{ -\frac{s}{2n} + B(n+1,1/s) \right\}.
\]

For all \( n \) such that \( n \geq (1/\rho - 1) \log(n)/\log(2) \), \( s 2^{-n} \leq 2^{-n} \leq n^{1-1/\rho} \). Using inequality (3.75) in [12] with \( x = n, a = 1, b = 1/s + 1, x_0 = 1 \), we have
\[
B(n+1,1/s) = \frac{\Gamma(n+1)\Gamma(1/s)}{\Gamma(n+1+1/s)} \geq n^{-1/s} \Gamma(1/s) = n^{1-1/\rho} \Gamma(1/\rho - 1).
\]

Since \( \rho < 1/2 \), \( \Gamma(1/\rho - 1) > \Gamma(1) = 1 \). Thus
\[
\left( \sqrt{\frac{\pi}{2}} \right)^{1-1/s} \left\{ -\frac{s}{2n} + B(n+1,1/s) \right\} \geq n^{1-1/\rho} \left( \sqrt{\frac{\pi}{2}} \right)^{1-1/s} \{ \Gamma(1/\rho - 1) - 1 \}.
\]

We have shown that
\[
f(n,\rho) \geq n^{1-1/\rho} \frac{\sqrt{2\pi}}{\sqrt{s}} \left( \sqrt{\frac{\pi}{2}} \right)^{1-1/s} \{ \Gamma(1/\rho - 1) - 1 \}
= n^{1-1/\rho} \frac{1}{2^{2-1/\rho}} \sqrt{\frac{1-\rho}{\rho}} \{ \Gamma(1/\rho - 1) - 1 \}.
\]

This proves (2.4).

Finally, turn to the proof of (2.5). Let \( h_n = n/ \lceil \log(n) \log(n^{1/s}) \rceil \) and \( \gamma_n = \Phi(\sqrt{2 \log(h_n)}) \). Using (2.7) write (2.9)
\[
f(n,\rho) = \frac{(\sqrt{2\pi})^{1/s-1}}{\sqrt{s}} \left\{ \int_{\gamma_n}^{\gamma_n} x^n \left[ \phi \left( \Phi^{-1}(x) \right) \right]^{1/s-1} dx + \int_{\gamma_n}^{1} x^n \left[ \phi \left( \Phi^{-1}(x) \right) \right]^{1/s-1} dx \right\}.
\]

Since \( \rho < 1/2 \) we have \( s < 1 \) and hence \( 1/s - 1 > 0 \). Then \( \left[ \phi \left( \Phi^{-1}(x) \right) \right]^{1/s-1} \) is a decreasing function over \( (1/2,1) \). Since for sufficiently large \( n \) it holds that
\( \gamma_n > 1/2 \), we have that

\[
\int_{\gamma_n}^{1} x^n \left[ \phi \left( \Phi^{-1}(x) \right) \right]^{1/s-1} dx \leq \\
\left[ \phi \left( \Phi^{-1}(\gamma_n) \right) \right]^{1/s-1} \int_{\gamma_n}^{1} x^n dx = \\
\frac{1}{(\sqrt{2\pi})^{1/s-1}} \left[ \exp(-\log(n) + \log(\log(n) \log(n^{1/s})) \right]^{1/s-1} \frac{1 - \gamma_n^{n+1}}{n+1} \leq \\
\frac{1}{(\sqrt{2\pi})^{1/s-1}} n^{1-1/s} \left[ \log(n) \log(n^{1/s}) \right]^{1/s-1} \frac{1 - \gamma_n^{n+1}}{n+1} \leq \\
\frac{1}{(\sqrt{2\pi})^{1/s-1}} n^{-1/s} \left[ \log(n) \log(n^{1/s}) \right]^{1/s-1} = \\
(2.10) \quad \frac{1}{(\sqrt{2\pi})^{1/s-2}} n^{1-1/\rho} \left[ (1/\rho - 1)\log(n)^2 \right]^{1/\rho-2}.
\]

By Birnbaum’s inequality

\[
\gamma_n^n \leq \left( 1 - \frac{\phi(\sqrt{2 \log(h_n)})}{\sqrt{4 + 2 \log(h_n) + \sqrt{2 \log(h_n)}}} \right)^n = \left( 1 - \frac{\log(n) \log(n^{1/s})}{n\sqrt{2\pi}(\sqrt{4 + 2 \log(h_n) + \sqrt{2 \log(h_n)})} \right)^n.
\]

Now, since \( h_n = n / \left[ \log(n) \log(n^{1/s}) \right] \), for all sufficiently large \( n \)

\[
\frac{\log(n) \log(n^{1/s})}{n\sqrt{2\pi}(\sqrt{4 + 2 \log(h_n) + \sqrt{2 \log(h_n)})} \geq \frac{\log(n^{1/s})}{n}.
\]

Thus

\[
\gamma_n^n \leq \left( 1 - \frac{\log(n^{1/s})}{n} \right)^n.
\]

Also, for all sufficiently large \( n \)

\[
\left( 1 - \frac{\log(n^{1/s})}{n} \right)^{n/\log(n^{1/s})} \leq \exp(-1)
\]

and hence

\[
\gamma_n^n \leq \exp(-\log(n^{1/s})) = n^{-1/s} = n^{1-1/\rho}.
\]

This together with \( 2.9 \), \( 2.10 \) and \( 2.11 \) implies that, for all sufficiently large \( n \),

\[
f(n, \rho) \leq n^{1-1/\rho} \sqrt{\frac{1 - \rho}{\rho}} \left[ (1/\rho - 1)\log(n)^2 \right]^{1/\rho-2},
\]

This proves \( 2.5 \).
The case $\rho = 1/2$ is excluded from Theorem 2.1, since it is well known, see [8] for example, that $f(n,1/2) = 1/(n+1)$. Theorem 2.1 implies that, for $\rho > 0$ and large $n$, $f(n,\rho)$ behaves essentially like $n^{-1/\rho}$, except for logarithmic factors.

Recall that the probability that a $k$-homogeneous polynomial, $k \geq 2$, attains a relative maximum at a given vertex of the $n$-dimensional simplex is given by the probability that the standard normal random variables

$$X_1(a), \ldots, X_n(a)$$

are all positive. By (2.1), the covariance matrix of the multivariate normal vector $X(a) = (X_1(a), \ldots, X_n(a))$ is given

$$A = \begin{pmatrix}
1 & \rho_n & \rho_n & \ldots & \rho_n \\
\rho_n & 1 & \rho_n & \ldots & \rho_n \\
\vdots & \vdots & \ddots & \ldots & \rho_n \\
\rho_n & \rho_n & \rho_n & \ldots & 1
\end{pmatrix},$$

where $\rho_n = (nk + (k-1))/(n(k+1) + (k-1))$ satisfies $\rho_n > 1/2$ and $\rho_n \to k/(k+1)$. Thus, using the notation of Theorem 2.1 for $k \geq 2$, the probability that a $k$-homogeneous polynomial attain a relative maximum at a given vertex of the $n$-dimensional simplex is given by $f(n,\rho_n)$. By Theorem 2.1 for large enough $n$, we have

$$\frac{c}{\sqrt{\log(n)}} \leq \frac{f(n,\rho_n)}{n^{1/\rho_n - 1}} \leq c$$

where $c > 0$ is a universal constant.

3. Attaining maxima as $n$ grows

In this section we will consider the probability of a $k$-homogeneous ($k \geq 2$) polynomial attaining relative maxima at any vertex of the simplex. Denote by $A_i$ the set of polynomials attaining a relative maximum at the $i$th vertex, and by $W(A_i)$ the measure of this set. We know —because of rotation-invariance— that $W(A_i) = W(A_j) = f(n,\rho_n)$ for all $i, j$. Note that if the $A_i$’s were independent events (they are not), the probability of attaining a relative maximum at some vertex of the simplex would be

$$W\left(\bigcup_{i=1}^{n+1} A_i\right) = 1 - W\left(\bigcap_{i=1}^{n+1} A_i^c\right) = 1 - W(A_i)^{n+1} = 1 - (1 - f(n,\rho_n))^{n+1}$$

We know that for large enough $n$, $f(n,\rho_n) \geq \frac{\beta_n}{n+1}$, where

$$\beta_n = c \frac{n^{n+1} \log n}{\log(n+1)} (n+1).$$

But we have $\beta_n \to \infty$, and $\frac{\beta_n}{n+1} \to 0$, so

$$1 - (1 - f(n,\rho_n))^{n+1} \geq 1 - \left(1 - \frac{\beta_n}{n+1}\right)^{n+1} \to 1.$$
$(n+1)n$ matrices

$$\Sigma_0 = \begin{pmatrix} A & 0 & 0 & \cdots & 0 \\ 0 & A & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & A \end{pmatrix}$$

and

$$\Sigma_\varepsilon = \begin{pmatrix} A & M_\varepsilon & M_\varepsilon & \cdots & M_\varepsilon \\ M_\varepsilon & A & M_\varepsilon & \cdots & M_\varepsilon \\ \vdots & \vdots & \ddots & \cdots & M_\varepsilon \\ M_\varepsilon & M_\varepsilon & M_\varepsilon & \cdots & A \end{pmatrix}$$

where $A$ is the $n \times n$ covariance matrix of $\mathbf{X}(a)$, and $M_\varepsilon$ is the matrix whose $(i,j)$-th component is $(\langle X_i(a), X_j(b) \rangle)$, where $a$ and $b$ are different vertices of the simplex. Recall that each component of $M_\varepsilon$ is bounded in absolute value by $\varepsilon = \frac{1}{n^2(n-k+1)} \left( \frac{1}{n} + (k-1) \right)$. We need the following lemma.

**Lemma 3.1.** The inverse of $A$ is

$$B = \begin{pmatrix} \alpha & \beta & \beta & \cdots & \beta \\ \beta & \alpha & \beta & \cdots & \beta \\ \vdots & \vdots & \ddots & \cdots & \beta \\ \beta & \beta & \beta & \cdots & \alpha \end{pmatrix}$$

where $\alpha$ and $\beta$ are such that, as $n \to \infty$, $\alpha \to k+1$ and $(n-1)|\beta| \to k+1$.

**Proof.** For $AB = I$ we need

$$\alpha + (n-1)\rho_n \beta = 1$$

and

$$\rho_n \alpha + [(n-2)\rho_n + 1] \beta = 0.$$ 

Thus

$$\alpha = \frac{(n-2)\rho_n + 1}{(n-2)\rho_n + 1 - \rho_n^2(n-1)}$$

and

$$\beta = \frac{-\rho_n}{(n-2)\rho_n + 1 - \rho_n^2(n-1)},$$

which, for our value of $\rho_n$ is

$$\alpha = \frac{n^3(k^2 + k) - n^2(3k + 1) - n(k^2 - 2k + 3) + 2k - 2}{n^4k - n^2(k^2 + k + 1) + 2n(k+1) + k^2 - 1}$$

and

$$\beta = \frac{n^2(k^2 + k) + n(2k^2 - k - 1) + (k-1)^2}{-n^3k - n^2k}.$$ 

$\square$

We now have the following.
Theorem 3.2. The probability that a $k$-homogeneous polynomial $P : \mathbb{R}^n \to \mathbb{R}$ of degree $k > 4$ attains a relative maximum at a vertex of the $n$-dimensional simplex tends to one as the dimension $n$ tends to infinity.

Proof. Consider the $n(n+1)$ random variables

$$X_1(a_1), \ldots, X_n(a_1), X_1(a_2), \ldots, X_n(a_2), \ldots, X_1(a_{n+1}), \ldots, X_n(a_{n+1})$$

where $a_1, a_2, \ldots, a_{n+1}$ are the $n + 1$ vertices of the simplex $\Delta$. These are $N(0, 1)$ random variables with covariance matrix $\Sigma_\varepsilon$. If the events $A_1, \ldots, A_{n+1}$ were independent, the covariance matrix would be $\Sigma_0$. We denote by $G_\varepsilon$ the gaussian measure on $\mathbb{R}^{n(n+1)}$ with covariance matrix $\Sigma_\varepsilon$ and by $G_0$ the gaussian measure on $\mathbb{R}^{n(n+1)}$ whose covariance matrix is $\Sigma_0$. We are interested in the probability of the event $\bigcup_{i=1}^{n+1} A_i$: “for some vertex $a_1, a_2, \ldots, a_{n+1}$, all $X_1, X_2, \ldots, X_n$ (corresponding to that vertex) are positive”. We calculate the probability of the complement $\bigcap_{i=1}^{n+1} A_i^c$: “for all vertices $a_1, a_2, \ldots, a_{n+1}$, some $X_j$ (corresponding to each vertex) is negative”. We do this by integrating the gaussian measure $G_\varepsilon$ of covariance $\Sigma_\varepsilon$ over the subset $\Omega$ of

$$\mathbb{R}^{n(n+1)} = \{(z_1, \ldots, z_{n+1}) : z_i = (z_{i1}, \ldots, z_{in}) \in \mathbb{R}^n\}$$

given by

$$\Omega = \{(z_1, \ldots, z_{n+1}) \in \mathbb{R}^{n(n+1)} : \text{for all } i \leq n + 1 \text{ there is a } j \leq n : z_{ij} \leq 0\},$$

and analogously for $G_0$. Thus

$$1 - (1 - f(n, \rho_n))^{n+1} - W\left(\bigcup_{i=1}^{n+1} A_i\right) = \left|1 - (1 - f(n, \rho_n))^{n+1} - 1 + W\left(\bigcap_{i=1}^{n+1} A_i\right)\right|$$

$$= \left|W\left(\bigcap_{i=1}^{n+1} A_i\right) - (1 - f(n, \rho_n))^{n+1}\right|$$

$$= \int_{\Omega} dG_\varepsilon - dG_0.$$

This is bounded by the total variation distance between the measures $G_\varepsilon$ and $G_0$. By the Devroye-Mehrabian-Reddad theorem [5], this distance will in turn be bounded by

$$\frac{3}{2} \left\| \Sigma_\varepsilon \Sigma_0^{-1} - I \right\|_F,$$

where $\| \cdot \|_F$ is the Frobenius norm. Thus consider

$$\Sigma_\varepsilon \Sigma_0^{-1} - I = \begin{pmatrix}
0 & M_\varepsilon B & M_\varepsilon B & \ldots & M_\varepsilon B \\
M_\varepsilon B & 0 & M_\varepsilon B & \ldots & M_\varepsilon B \\
\vdots & \vdots & \ddots & \ldots & M_\varepsilon B \\
M_\varepsilon B & M_\varepsilon B & M_\varepsilon B & \ldots & M_\varepsilon B \\
M_\varepsilon B & M_\varepsilon B & M_\varepsilon B & \ldots & 0
\end{pmatrix}.$$

Each of the $(n + 1)^2 - (n + 1)$ non-zero blocks is an $n \times n$ matrix whose entries are inner products between a row of $M_\varepsilon$ and a column of $B$. Thus each of these entries is bounded by $\varepsilon \|((\alpha, \beta, \ldots, \beta))\|_1$. We then have

$$\left\| \Sigma_\varepsilon^{-1} \Sigma_0 - I \right\|_F \leq (n+1)^2 - (n+1) n^2 \varepsilon^2 \|((\alpha, \beta, \ldots, \beta))\|_1^2 < (n+1)^2 n^2 \varepsilon^2 (\alpha + (n-1)\beta)^2,$$
which for large $n$ is bounded by

$$C n^4 \varepsilon^2 (k + 1)^2 \leq C' \frac{(k + 1)^2}{n^{2k-8}},$$

where $C$ and $C'$ are positive constants. Thus if $k \geq 5$, for large $n$,

$$
\begin{aligned}
1 - W \left( \bigcup_{i=1}^{n+1} A_i \right) &\leq \left| (1 - f(n, \rho_n))^{n+1} \right| + \left| 1 - (1 - f(n, \rho_n))^{n+1} - W \left( \bigcup_{i=1}^{n+1} A_i \right) \right| \\
&\leq \left| (1 - f(n, \rho_n))^{n+1} \right| + C' \frac{(k + 1)^2}{n^{2k-8}} \to 0.
\end{aligned}
$$

This finishes the proof of the theorem.

\[ \square \]

References

[1] Bernard Beauzamy, Enrico Bombieri, Per Enflo, and Hugh L Montgomery. Products of polynomials in many variables. *Journal of Number Theory*, 36:219–245, 1990.
[2] Z. W. Birnbaum. An inequality for mill’s ratio. *Ann. Math. Statist.*, 13(2):245–246, 06 1942.
[3] Daniel Carando and Ignacio Zalduendo. Where do homogeneous polynomials attain their norm? *Publ. Math. Debrecen*, 63(1-2):19–28, 2003.
[4] FN David. A note on the evaluation of the multivariate normal integral. *Biometrika*, 40(3-4):458–459, 1953.
[5] Luc Devroye, Abbas Mehrabian, and Tommy Reddad. The total variation distance between high-dimensional gaussians. *arXiv preprint arXiv:1810.08693* 2018.
[6] Thomas AW Dwyer III. Partial differential equations in fischer-fock spaces for the hilbert-schmidt holomorphy type. *Bulletin of the American Mathematical Society*, 77(5):725–730, 1971.
[7] Robert D Gordon. Values of mills’ ratio of area to bounding ordinate and of the normal probability integral for large values of the argument. *The Annals of Mathematical Statistics*, 12(3):364–366, 1941.
[8] Donald B. Owen. Orthant probabilities. *Encyclopedia of Statistical Sciences*, 6:521–523, 1985.
[9] David Pérez-García and Ignacio Villanueva. Where do homogeneous polynomials on $\ell^1_1$ attain their norm? *Journal of Approximation Theory*, 127(1):124–133, 2004.
[10] Damián Pinasco and Ignacio Zalduendo. A probabilistic approach to polynomial inequalities. *Israel Journal of Mathematics*, 190(1):67–82, 2012.
[11] Damián Pinasco and Ignacio Zalduendo. On the measure of polynomials attaining maxima on a vertex. *Mathematical Inequalities & Applications*, 22(2):421–432, 2019.
[12] Feng Qi. Bounds for the ratio of two gamma functions. *Journal of Inequalities and Applications*, 2010:1–84, 2010.
[13] Harold Ruben. Probability content of regions under spherical normal distributions, iii: The bivariate normal integral. *The Annals of Mathematical Statistics*, pages 171–186, 1961.
[14] WF Sheppard. On the geometrical treatment of the normal curve of statistics, with especial reference to correlation and to the theory of error. *Proceedings of the Royal Society of London*, 62(379-387):170–173, 1898.
[15] George Powell Steck. Orthant probabilities for the equicorrelated multivariate normal distribution. *Biometrika*, 49(3/4):433–445, 1962.

 Universidad Torcuato Di Tella Av. Figueroa Alcorta 7350 (C1428BCW) Buenos Aires, Argentina and CONICET
E-mail address: dpinasco@utdt.edu

 Universidad Torcuato Di Tella Av. Figueroa Alcorta 7350 (C1428BCW) Buenos Aires, Argentina and CONICET
E-mail address: esmucler@utdt.edu

 Universidad Torcuato Di Tella Av. Figueroa Alcorta 7350 (C1428BCW) Buenos Aires, Argentina and CONICET
E-mail address: izalduendo@utdt.edu