

# Bounds on the determinant of an exponential matrix

Michael S. Floater\*

September 7, 2026

## Abstract

We derive upper and lower bounds on the determinant of an exponential matrix. They can be transformed into corresponding bounds for the determinant of a univariate Gaussian matrix.

*Keywords:* exponential matrix, determinant, total positivity, divided difference, Vandermonde determinant, Gaussian matrix.

*Math Subject Classification:* 15A15, 33B10, 65D05, 65F40.

## 1 Introduction

Consider the  $n \times n$  matrix

$$A := [e^{x_i y_j}]_{i,j=1,\dots,n} = \begin{bmatrix} e^{x_1 y_1} & \dots & e^{x_1 y_n} \\ \vdots & & \vdots \\ e^{x_n y_1} & \dots & e^{x_n y_n} \end{bmatrix}, \quad (1)$$

for real  $x_1 < x_2 < \dots < x_n$  and  $y_1 < y_2 < \dots < y_n$ . It includes the Vandermonde and generalized Vandermonde matrices as special cases, see, for example, Pinkus [23, Section 4.2]. In this paper, we study upper and lower bounds on the determinant of  $A$ . The motivation comes from Schumaker's

---

\*Department of Mathematics, University of Oslo, PO Box 1053, Blindern, 0316 Oslo, Norway, email: [michael.floater@math.uio.no](mailto:michael.floater@math.uio.no)

conjecture about partial interpolation by multivariate Bernstein polynomials [2], [19], [11]. The conjecture can be recast as the non-singularity of Hadamard (element-wise) products of matrices in the form of  $A$ , and the hope is that bounds on the determinant of  $A$ , or the technique used to find them, may help in the future towards a solution to the conjecture. Bounds on the determinant  $A$  have also been applied, either directly or indirectly, to the error analysis of interpolation by exponential functions by Yarotsky [27] and to finding optimal points for interpolation by exponentials and Gaussian radial basis functions by Bos and de Marchi [4].

Let us recall some basic facts about  $A$ . The determinant of  $A$  is positive, which was shown by Pólya in 1920 using a form of Descartes' Rule of Signs, see Pólya and Szegő [24, Part Five, Problem 76] and Gantmacher and Krein [13]. Since all square submatrices of  $A$  have the same form as  $A$  itself, they too have positive determinant, and thus  $A$  is totally positive in the sense that all its minors are positive. Every totally positive matrix satisfies the Hadamard inequality: its determinant is bounded above by the product of its diagonal elements (this follows from Sylvester's determinant identity; see [13, Théorème 1], [23, Section 1.4], and [8, Chapter 6]), and so

$$\det(A) \leq e^{x_1 y_1 + \dots + x_n y_n}.$$

Another consequence of total positivity is the spectral theory of Gantmacher and Krein [13] (which builds on Perron-Frobenius theory and properties of compound matrices). So the eigenvalues of  $A$  are real, positive, and distinct, and the eigenvectors satisfy sign-change properties; see [23, Theorem 5.3].

We can also express the determinant of  $A$  using the remarkable Harish-Chandra-Itzykson-Zuber formula [16, 18, 14], which involves an integral over unitary matrices (see Section 2), and is connected to Schur functions [21, 7, 1]. While this formula is far from elementary to derive, it was observed by Bos and de Marchi [4] that the integrand can be bounded above and below using the trace inequality of Coope and Renaud [5], and this leads to corresponding bounds on  $\det(A)$ , which are as follows. Let  $\mathbf{x} := (x_1, x_2, \dots, x_n)$  and  $\mathbf{y} := (y_1, y_2, \dots, y_n)$  and denote by  $V(\mathbf{x})$  the Vandermonde determinant

$$V(\mathbf{x}) = V(x_1, x_2, \dots, x_n) := \prod_{1 \leq i < j \leq n} (x_j - x_i)$$

(the determinant of the matrix  $[x_i^{j-1}]_{i,j=1,\dots,n}$ ). Let  $c_n$  be the ‘superfactorial’

$$c_n := \prod_{i=1}^n i!.$$

Then the bounds are

**Theorem 1**

$$e^{x_1 y_n + x_2 y_{n-1} + \dots + x_n y_1} \leq \frac{c_{n-1} \det(A)}{V(\mathbf{x})V(\mathbf{y})} \leq e^{x_1 y_1 + x_2 y_2 + \dots + x_n y_n}.$$

On the other hand, it would be preferable to have a more elementary proof of Theorem 1. This turns out to be the case, due to the fact that we can alternatively express  $\det(A)$  recursively, as an integral over a rectangle in  $\mathbb{R}^{n-1}$  in which the integrand is itself an exponential determinant, but of order  $n-1$  (Lemma 1). In this way we obtain Theorem 1 by proving it first in the case  $n=2$ , and using induction on  $n$  for  $n \geq 3$ . But not only that, using this recursive approach, we obtain better bounds, which are as follows. Define the sum,

$$s(\mathbf{x}) := x_1 + x_2 + \dots + x_n,$$

and denote the permanent of  $A$  by  $\text{per}(A)$ .

**Theorem 2**

$$\frac{1}{c_{n-1}} e^{s(\mathbf{x})s(\mathbf{y})/n} \leq \frac{\det(A)}{V(\mathbf{x})V(\mathbf{y})} \leq \frac{1}{c_n} \text{per}(A).$$

To see that this is an improvement, let  $S_n$  be the symmetric group of permutations of  $(1, 2, \dots, n)$ , and for each permutation  $\sigma \in S_n$ , let

$$\phi_\sigma := \sum_{i=1}^n x_i y_{\sigma(i)}. \tag{2}$$

Then the rearrangement inequality [15, Thm 368] bounds  $\phi_\sigma$ :

$$\sum_{i=1}^n x_i y_{n+1-i} \leq \phi_\sigma \leq \sum_{i=1}^n x_i y_i. \tag{3}$$

Let  $m_a$  and  $m_g$  be the arithmetic and geometric means of the exponentials  $e^{\phi_\sigma}$ ,  $\sigma \in S_n$ ,

$$m_a := \frac{1}{n!} \sum_{\sigma \in S_n} e^{\phi_\sigma}, \quad m_g := \left( \prod_{\sigma \in S_n} e^{\phi_\sigma} \right)^{1/n!}.$$

Then,

$$m_a = \frac{1}{n!} \text{per}(A),$$

and since

$$\sum_{\sigma \in S_n} \phi_\sigma = \sum_{i=1}^n x_i \sum_{\sigma \in S_n} y_{\sigma(i)} = (n-1)! s(\mathbf{x}) s(\mathbf{y}),$$

we find

$$m_g = e^{s(\mathbf{x})s(\mathbf{y})/n},$$

and so Theorem 2 can be rewritten as

$$m_g \leq \frac{c_{n-1} \det(A)}{V(\mathbf{x})V(\mathbf{y})} \leq m_a.$$

and by (3),

$$e^{x_1 y_n + x_2 y_{n-1} + \dots + x_n y_1} < m_g < m_a < e^{x_1 y_1 + x_2 y_2 + \dots + x_n y_n}.$$

While the recursive proofs of Theorems 1 and 2 are elementary, they are not simple either. When we apply the induction step, the difficult task is to compute the resulting integral, and the integrand is either a Vandermonde determinant or the product of a Vandermonde determinant and a function of the sum of the variables. It turns out that there is an identity relating this kind of integral to divided differences, which is derived in [12]. For the lower bound of Theorem 2, we also need an estimate on a divided difference from below, and for this we use an error formula from [10]. For a survey of divided differences, see de Boor [6], and for some basic theory, see Isaacson and Keller [17, Section 6.1] and Steffensen [26]. For the upper bound of Theorem 2, we need an auxiliary result concerning the maximization of a function that is both convex and symmetric.

The bounds of Theorem 2 can be transformed into corresponding bounds for the determinant of a univariate Gaussian matrix, and an application of the lower bound is a simple method for selecting the shape parameter. We treat this in Section 10.

## 2 The unitary matrix integral

The unitary matrix integral formula ((3.15) of Gross and Richards [14]) is

$$\frac{c_{n-1} \det(A)}{V(\mathbf{x})V(\mathbf{y})} = \int_{U(n)} e^{\text{tr}(XUYU^*)} dU, \quad (4)$$

where the integration is over the group of  $n \times n$  unitary matrices  $U$  with respect to the normalized Haar measure. Here,  $X = \text{diag}(x_1, x_2, \dots, x_n)$ ,  $Y = \text{diag}(y_1, y_2, \dots, y_n)$ ,  $U^*$  is the conjugate transpose of  $U$ , and  $\text{tr}$  denotes the trace. Multiplying out the matrices in the integrand, and taking the trace gives

$$\text{tr}(XUYU^*) = \mathbf{x}^T (U \circ \bar{U}) \mathbf{y},$$

with  $\mathbf{x}$  and  $\mathbf{y}$  treated as column vectors, and where  $\circ$  denotes the Hadamard product. It follows that if  $U = [u_{ij}]$ , then

$$\text{tr}(XUYU^*) = \mathbf{x}^T V \mathbf{y},$$

where  $V = [v_{ij}]$  is the real matrix with elements  $v_{ij} = |u_{ij}|^2$ . Since  $U$  is unitary,  $V$  is doubly stochastic:  $v_{ij} \geq 0$ , and

$$\sum_{j=1}^n v_{ij} = 1, \quad i = 1, \dots, n; \quad \sum_{i=1}^n v_{ij} = 1, \quad j = 1, \dots, n.$$

So by the Birkhoff-von Neumann theorem [3, p. 63], we can express  $V$  as a convex combination of the permutation matrices  $P_\sigma = [p_{ij}]$  defined by

$$p_{ij} = \begin{cases} 1 & \text{if } j = \sigma(i); \\ 0 & \text{otherwise.} \end{cases}$$

So there are weights  $\lambda_\sigma \geq 0$  with  $\sum_{\sigma \in S_n} \lambda_\sigma = 1$  such that

$$V = \sum_{\sigma \in S_n} \lambda_\sigma P_\sigma.$$

Therefore,

$$\mathbf{x}^T V \mathbf{y} = \sum_{\sigma \in S_n} \lambda_\sigma \mathbf{x}^T P_\sigma \mathbf{y} = \sum_{\sigma \in S_n} \lambda_\sigma \phi_\sigma,$$

with  $\phi_\sigma$  as in (2), and so by (3),

$$\sum_{i=1}^n x_i y_{n+1-i} \leq \mathbf{x}^T V \mathbf{y} \leq \sum_{i=1}^n x_i y_i.$$

Applying this to the integrand of (4), and since  $dU$  is normalized to give volume 1, we deduce Theorem 1.

### 3 The case $n = 2$

Next we give an elementary proof of both Theorems 1 and 2 when  $n = 2$ . In this case,

$$\det(A) = e^b - e^a,$$

where  $a := x_1 y_2 + x_2 y_1$  and  $b := x_1 y_1 + x_2 y_2$ , and

$$V(\mathbf{x})V(\mathbf{y}) = (x_2 - x_1)(y_2 - y_1) = b - a,$$

and so

$$\frac{\det(A)}{V(\mathbf{x})V(\mathbf{y})} = \frac{e^b - e^a}{b - a},$$

a first order divided difference of the exponential function. By the mean value theorem, there is some  $\xi \in (a, b)$  such that

$$\frac{e^b - e^a}{b - a} = e^\xi,$$

and so

$$e^a \leq \frac{e^b - e^a}{b - a} \leq e^b,$$

and

$$e^{x_1 y_2 + x_2 y_1} \leq \frac{\det(A)}{V(\mathbf{x})V(\mathbf{y})} \leq e^{x_1 y_1 + x_2 y_2},$$

which is Theorem 1 when  $n = 2$ . However, we can improve on these bounds by expressing the difference  $e^b - e^a$  in terms of the hyperbolic sine,

$$e^b - e^a = 2e^{(a+b)/2} \sinh((b-a)/2).$$

Using the well known estimates

$$t \leq \sinh(t) \leq t \cosh(t), \quad t \geq 0,$$

we then find that

$$e^{(a+b)/2} \leq \frac{e^b - e^a}{b - a} \leq \frac{e^a + e^b}{2},$$

and so

$$e^{(x_1+x_2)(y_1+y_2)/2} \leq \frac{\det(A)}{V(\mathbf{x})V(\mathbf{y})} \leq \frac{1}{2}\text{per}(A),$$

which is Theorem 2 when  $n = 2$ .

## 4 A multiple integral

To prove Theorems 1 and 2 for  $n \geq 3$  we express  $\det(A)$  as a multiple integral whose integrand is the determinant of an exponential matrix of order  $n - 1$ . Let  $R(\mathbf{x}) \subset \mathbb{R}^{n-1}$  denote the ‘sequential’ rectangle

$$R(\mathbf{x}) := [x_1, x_2] \times [x_2, x_3] \times \cdots \times [x_{n-1}, x_n], \quad (5)$$

and for a continuous function  $f : R(\mathbf{x}) \rightarrow \mathbb{R}$ , denote its integral in  $R(\mathbf{x})$  as

$$\int_{R(\mathbf{x})} f(\mathbf{t}) d\mathbf{t} := \int_{x_{n-1}}^{x_n} \cdots \int_{x_2}^{x_3} \int_{x_1}^{x_2} f(t_1, t_2, \dots, t_{n-1}) dt_1 dt_2 \cdots dt_{n-1},$$

where  $\mathbf{t} := (t_1, t_2, \dots, t_{n-1}) \in R(\mathbf{x})$ . Consider again  $A$  in (1) with  $n \geq 3$ . Choose any column index  $k \in \{1, 2, \dots, n\}$  and define

$$u_j := \begin{cases} y_j - y_k, & j = 1, \dots, k-1; \\ y_{j+1} - y_k & j = k, \dots, n-1, \end{cases} \quad (6)$$

and

$$\omega_k := (-1)^{k-1} u_1 \cdots u_{n-1} \quad (7)$$

Then

$$u_1 < \cdots < u_{k-1} < 0 < u_k < \cdots < u_{n-1}, \quad \text{and} \quad \omega_k > 0.$$

**Lemma 1**

$$\det(A) = e^{s(\mathbf{x})y_k} \omega_k \int_{R(\mathbf{x})} \det(B_k(\mathbf{t})) d\mathbf{t},$$

where

$$B_k(\mathbf{t}) := \begin{bmatrix} e^{t_1 u_1} & \cdots & e^{t_1 u_{n-1}} \\ \vdots & & \vdots \\ e^{t_{n-1} u_1} & \cdots & e^{t_{n-1} u_{n-1}} \end{bmatrix}.$$

*Proof.* For  $i = 1, \dots, n$ , we divide the  $i$ -th row of  $A$  by  $e^{x_i y_k}$  so that

$$\det(A) = e^{x_1 y_k} \dots e^{x_n y_k} \det(A') = e^{s(\mathbf{x}) y_k} \det(A'),$$

where

$$A' := \begin{bmatrix} e^{x_1 u_1} & \dots & e^{x_1 u_{k-1}} & 1 & e^{x_1 u_k} & \dots & e^{x_1 u_{n-1}} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ e^{x_n u_1} & \dots & e^{x_n u_{k-1}} & 1 & e^{x_n u_k} & \dots & e^{x_n u_{n-1}} \end{bmatrix}.$$

Next we subtract row  $n-1$  of  $A'$  from row  $n$  which does not change  $\det(A')$ , and write the differences as

$$e^{x_n u_j} - e^{x_{n-1} u_j} = u_j \int_{x_{n-1}}^{x_n} e^{t u_j} dt, \quad j = 1, \dots, n-1.$$

Then, with  $\int = \int_{x_{n-1}}^{x_n} dt$ , we can replace row  $n$  by

$$\left[ u_1 \int e^{t u_1}, \dots, u_{k-1} \int e^{t u_{k-1}}, 0, u_k \int e^{t u_k}, \dots, u_{n-1} \int e^{t u_{n-1}} \right]. \quad (8)$$

Next we subtract row  $n-2$  from row  $n-1$ , and then we can replace row  $n-1$  by (8) but now with  $\int = \int_{x_{n-2}}^{x_{n-1}} dt$ . We continue like this until we have modified the second row of  $A'$  to (8) with  $\int = \int_{x_1}^{x_2} dt$ . Now  $\det(A')$  reduces to a determinant of order  $n-1$ , which can be written as the product of  $\omega_k$  and the stated multiple integral of  $\det(B_k(\mathbf{t}))$ .  $\square$

## 5 The bounds of Theorem 1

We now prove Theorem 1 in general and suppose that  $n \geq 3$ . We can choose any  $k \in \{1, \dots, n\}$  and apply Lemma 1, and since  $t_1 < t_2 < \dots < t_{n-1}$  and  $u_1 < u_2 < \dots < u_{n-1}$  in the matrix  $B_k(\mathbf{t})$ , we can apply the induction hypothesis to deduce that

$$e^{t_1 u_{n-1} + \dots + t_{n-1} u_1} \leq \frac{c_{n-2} \det(B_k(\mathbf{t}))}{V(\mathbf{t})V(\mathbf{u})} \leq e^{t_1 u_1 + \dots + t_{n-1} u_{n-1}}.$$

Since  $V(\mathbf{u}) = V(y_1, \dots, y_{k-1}, y_{k+1}, \dots, y_n)$  and

$$\omega_k = (y_k - y_1) \dots (y_k - y_{k-1})(y_{k+1} - y_k) \dots (y_n - y_k),$$

we find

$$\omega_k V(\mathbf{u}) = V(\mathbf{y}), \quad (9)$$

and so

$$\int_{R(\mathbf{x})} V(\mathbf{t}) e^{t_1 u_{n-1} + \dots + t_{n-1} u_1} d\mathbf{t} \leq \frac{c_{n-2} \det(A)}{V(\mathbf{y}) e^{s(\mathbf{x}) y_k}} \leq \int_{R(\mathbf{x})} V(\mathbf{t}) e^{t_1 u_1 + \dots + t_{n-1} u_{n-1}} d\mathbf{t}.$$

Since  $u_1, \dots, u_{k-1}$  are negative and  $u_k, \dots, u_{n-1}$  are positive,

$$\begin{aligned} t_1 u_{n-1} + \dots + t_{n-k} u_k &\geq x_1 u_{n-1} + \dots + x_{n-k} u_k, \\ t_{n+1-k} u_{k-1} + \dots + t_{n-1} u_1 &\geq x_{n+2-k} u_{k-1} + \dots + x_n u_1, \end{aligned}$$

and so

$$s(\mathbf{x}) y_k + t_1 u_{n-1} + \dots + t_{n-1} u_1 \geq x_1 y_n + \dots + x_n y_1.$$

Similarly,

$$\begin{aligned} t_1 u_1 + \dots + t_{k-1} u_{k-1} &\leq x_1 u_1 + \dots + x_{k-1} u_{k-1}, \\ t_k u_k + \dots + t_{n-1} u_{n-1} &\leq x_{k+1} u_k + \dots + x_n u_{n-1}, \end{aligned}$$

and so

$$s(\mathbf{x}) y_k + t_1 u_1 + \dots + t_{n-1} u_{n-1} \leq x_1 y_1 + \dots + x_n y_n.$$

Therefore,

$$e^{x_1 y_n + \dots + x_n y_1} \int_{R(\mathbf{x})} V(\mathbf{t}) d\mathbf{t} \leq \frac{c_{n-2} \det(A)}{V(\mathbf{y})} \leq e^{x_1 y_1 + \dots + x_n y_n} \int_{R(\mathbf{x})} V(\mathbf{t}) d\mathbf{t}.$$

By Corollary 1 of [12],

$$\int_{R(\mathbf{x})} V(\mathbf{t}) d\mathbf{t} = \frac{1}{(n-1)!} V(\mathbf{x}), \quad (10)$$

and observing that  $c_{n-2}(n-1)! = c_{n-1}$  completes the proof.

## 6 Divided difference error formula

For proving the lower bound of Theorem 2, we use a formula for the error incurred when approximating a divided difference by a normalized derivative. Let  $v_1, v_2, \dots, v_n \in \mathbb{R}$ , let  $\bar{v}$  be their mean

$$\bar{v} := \frac{v_1 + v_2 + \dots + v_n}{n},$$

and let

$$S = \sum_{i=0}^n (v_i - \bar{v})^2.$$

**Lemma 2** *If  $f$  is  $C^{n+1}$  in the smallest interval containing  $v_1, \dots, v_n$  then there exists  $\xi$  in that interval such that*

$$[v_1, v_2, \dots, v_n]f = \frac{f^{(n-1)}(\bar{v})}{(n-1)!} + \frac{S}{2} \frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

*Proof.* This was shown in Theorem 2 and Lemma 3 of [10].  $\square$

## 7 The lower bound of Theorem 2

To prove the lower bound of Theorem 2, we suppose  $n \geq 3$ . We can apply Lemma 1 with any  $k \in \{1, \dots, n\}$ , and by the induction hypothesis,

$$\det(B_k(\mathbf{t})) \geq \frac{1}{c_{n-2}} V(\mathbf{t}) V(\mathbf{u}) e^{s(\mathbf{t})s(\mathbf{u})/(n-1)}.$$

and recalling (9),

$$\det(A) \geq \frac{1}{c_{n-2}} V(\mathbf{y}) e^{s(\mathbf{x})y_k} \int_{R(\mathbf{x})} V(\mathbf{t}) e^{s(\mathbf{t})s(\mathbf{u})/(n-1)} d\mathbf{t}. \quad (11)$$

If  $s(\mathbf{u}) = 0$ , then, by (10), the integral in (11) becomes  $V(\mathbf{x})/(n-1)!$ , and the lower bound follows from the fact that  $y_k = s(\mathbf{y})/n$ . Otherwise,  $s(\mathbf{u}) \neq 0$ , and we can write the integral in (11) as

$$\int_{R(\mathbf{x})} V(\mathbf{t}) f^{(n-1)}(s(\mathbf{t})) d\mathbf{t},$$

where

$$f(\alpha) := \frac{e^{\bar{u}\alpha}}{\bar{u}^{n-1}} \quad \text{and} \quad \bar{u} := \frac{s(\mathbf{u})}{n-1} \neq 0.$$

Then by Theorem 1 of [12],

$$\int_{R(\mathbf{x})} V(\mathbf{t}) f^{(n-1)}(s(\mathbf{t})) d\mathbf{t} = V(\mathbf{x}) [p_1, p_2, \dots, p_n] f,$$

where

$$p_i := x_1 + \cdots + x_{n-i} + x_{n+2-i} + \cdots + x_n = \left( \sum_{j=1}^n x_j \right) - x_{n+1-i}. \quad (12)$$

Since  $f^{(n+1)}(\alpha) = \bar{u}^2 e^{\bar{u}\alpha} \geq 0$  for all  $\alpha \in \mathbb{R}$ , Lemma 2 implies

$$[p_1, p_2, \dots, p_n]f \geq \frac{1}{(n-1)!} f^{(n-1)}\left(\frac{p_1 + \cdots + p_n}{n}\right) = \frac{1}{(n-1)!} e^{\bar{u}(p_1 + \cdots + p_n)/n}.$$

From (12),

$$p_1 + \cdots + p_n = n \left( \sum_{j=1}^n x_j \right) - \sum_{i=1}^n x_{n+1-i} = (n-1) \sum_{j=1}^n x_j,$$

and so

$$\int_{R(\mathbf{x})} V(\mathbf{t}) f^{(n-1)}(s(\mathbf{t})) d\mathbf{t} \geq \frac{1}{(n-1)!} V(\mathbf{x}) e^{s(\mathbf{u})s(\mathbf{x})/n},$$

and

$$\det(A) \geq \frac{1}{c_{n-1}} V(\mathbf{x}) V(\mathbf{y}) e^{s(\mathbf{x})(y_k + s(\mathbf{u})/n)}.$$

Since

$$y_k + \frac{s(\mathbf{u})}{n} = \frac{1}{n} \left( ny_k + \sum_{j=1}^n (y_j - y_k) \right) = \frac{s(\mathbf{y})}{n},$$

the proof is complete.

## 8 Maximization on $R(\mathbf{x})$

In order to prove the upper bound of Theorem 2, we will make use of the following auxiliary result. Observe that the sequential rectangle  $R(\mathbf{x})$  in (5) is contained in the  $(n-1)$ -dimensional cube  $[x_1, x_n]^{n-1}$ .

**Lemma 3** *For a convex function  $f : [x_1, x_n]^{n-1} \rightarrow \mathbb{R}$  that is symmetric in its variables,*

$$\max_{\mathbf{t} \in R(\mathbf{x})} f(\mathbf{t}) = \max_{i=1, \dots, n} f(\mathbf{v}_i),$$

where

$$\mathbf{v}_i := (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n), \quad i = 1, \dots, n. \quad (13)$$

So  $f$  attains its maximum in  $R(\mathbf{x})$  at one of  $n$  of the  $2^{n-1}$  vertices of  $R(\mathbf{x})$ .

*Proof.* Since  $f$  is convex on  $R(\mathbf{x})$ , it attains its maximum at a vertex of  $R(\mathbf{x})$ . Let  $\mathbf{t} = (t_1, \dots, t_{n-1})$  be a vertex of  $R(\mathbf{x})$ . Then

$$t_j \in \{x_j, x_{j+1}\}, \quad j = 1, \dots, n-1.$$

So  $x_1$  and  $x_n$  occur at most once in  $\mathbf{t}$  but for  $2 \leq j \leq n-1$ ,  $x_j$  may occur twice, in which case

$$t_{j-1} = x_j = t_j. \quad (14)$$

Let  $m(\mathbf{t})$  denote the number of repeated  $x_j$  in  $\mathbf{t}$ , i.e., the number of indices  $j \in \{2, \dots, n-1\}$  such that (14) holds. We will show that if  $m(\mathbf{t}) > 0$ , then there are vertices  $\mathbf{t}'$  and  $\mathbf{t}''$  of  $R(\mathbf{x})$  with  $m(\mathbf{t}') = m(\mathbf{t}'') = m(\mathbf{t}) - 1$  such that

$$f(\mathbf{t}) \leq \max\{f(\mathbf{t}'), f(\mathbf{t}'')\}. \quad (15)$$

By repeating this argument, this will imply that the maximum of  $f$  over the vertices of  $R(\mathbf{x})$  is attained at a vertex  $\mathbf{t}$  with  $m(\mathbf{t}) = 0$ , in which case  $\mathbf{t}$  is one of the  $\mathbf{v}_i$  in (13).

To prove (15), choose some  $j \in \{2, \dots, n-1\}$  such that (14) holds. Then there must be at least one index  $i \in \{1, \dots, j-1\}$  such that  $x_i \notin \{t_1, \dots, t_{j-2}\}$ . Let  $i$  be the largest such index. Similarly, there must be at least one index  $k \in \{j+1, \dots, n\}$  such that  $x_k \notin \{t_{j+1}, \dots, t_{n-1}\}$ . Let  $k$  be the smallest such index. We now have

$$t_r = \begin{cases} x_{r+1}, & r = i, \dots, j-1, \\ x_r, & r = j, \dots, k-1. \end{cases}$$

Since  $t_j = x_j$  and  $x_i < x_j < x_k$ , we can express  $\mathbf{t}$  as

$$\mathbf{t} = (1 - \lambda)\mathbf{a} + \lambda\mathbf{b},$$

where

$$\lambda := \frac{x_j - x_i}{x_k - x_i},$$

and

$$\mathbf{a} := (t_1, \dots, t_{j-1}, x_i, t_{j+1}, \dots, t_{n-1}),$$

$$\mathbf{b} := (t_1, \dots, t_{j-1}, x_k, t_{j+1}, \dots, t_{n-1}),$$

and by the convexity of  $f$  in  $[x_1, x_n]^{n-1}$ ,

$$f(\mathbf{t}) \leq \max\{f(\mathbf{a}), f(\mathbf{b})\}.$$

However,

$$\mathbf{a} = (t_1, \dots, t_{i-1}, x_{i+1}, \dots, x_j, x_i, t_{j+1}, \dots, t_{n-1}),$$

and its elements are not non-decreasing, and so  $\mathbf{a}$  is not a vertex of  $R(\mathbf{x})$ . But we can re-order the elements to be non-decreasing, defining

$$\mathbf{a}' := (t_1, \dots, t_{i-1}, x_i, x_{i+1}, \dots, x_j, t_{j+1}, \dots, t_{n-1}).$$

Then, by the choice of  $i$  and  $j$ ,  $\mathbf{a}'$  is a vertex of  $R(\mathbf{x})$  and  $m(\mathbf{a}') = m(\mathbf{t}) - 1$ , and  $f(\mathbf{a}') = f(\mathbf{a})$  by the symmetry of  $f$ .

The point  $\mathbf{b}$  may or may not be a vertex of  $R(\mathbf{x})$ , and there are two cases. If  $k = j + 1$  then  $\mathbf{b}$  is a vertex with  $m(\mathbf{b}) = m(\mathbf{t}) - 1$ . Otherwise,  $k \geq j + 2$  and then

$$\mathbf{b} = (t_1, \dots, t_{j-1}, x_k, x_{j+1}, \dots, x_{k-1}, t_k, \dots, t_{n-1}),$$

which is not a vertex. Reordering the elements gives

$$\mathbf{b}' := (t_1, \dots, t_{j-1}, x_{j+1}, \dots, x_{k-1}, x_k, t_k, \dots, t_{n-1}),$$

and then  $\mathbf{b}'$  is a vertex with  $m(\mathbf{b}') = m(\mathbf{t}) - 1$  and  $f(\mathbf{b}') = f(\mathbf{b})$ . Thus we have proved (15).  $\square$

## 9 The upper bound of Theorem 2

For the upper bound of Theorem 2, let  $n \geq 3$ . For any  $k \in \{1, \dots, n\}$ , we can apply Lemma 1, and by the induction hypothesis applied to  $B_k(\mathbf{t})$ , and recalling (9),

$$\det(A) \leq \frac{1}{c_{n-1}} e^{s(\mathbf{x})y_k} V(\mathbf{y}) \int_{R(\mathbf{x})} V(\mathbf{t}) \text{per}(B_k(\mathbf{t})) d\mathbf{t}.$$

Taking the average of this over  $k = 1, \dots, n$  implies

$$\det(A) \leq \frac{1}{nc_{n-1}} V(\mathbf{y}) \int_{R(\mathbf{x})} V(\mathbf{t}) f(\mathbf{t}) d\mathbf{t},$$

where

$$f(\mathbf{t}) := \sum_{k=1}^n e^{s(\mathbf{x})y_k} \text{per}(B_k(\mathbf{t})).$$

Then

$$\det(A) \leq \frac{1}{nc_{n-1}} \int_{R(\mathbf{x})} V(\mathbf{t}) d\mathbf{t} \max_{\mathbf{t} \in R(\mathbf{x})} f(\mathbf{t}),$$

and by (10),

$$\det(A) \leq \frac{1}{c_n} V(\mathbf{x})V(\mathbf{y}) \max_{\mathbf{t} \in R(\mathbf{x})} f(\mathbf{t}).$$

Next, observe that swapping rows of  $B_k(\mathbf{t})$  does not change its permanent, and so  $\text{per}(B_k(\mathbf{t}))$  is symmetric in  $t_1, \dots, t_{n-1}$ , and therefore so is  $f(\mathbf{t})$ . Further, for any real vector  $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_{n-1}]^T$ , the function

$$g(\mathbf{t}) := e^{\lambda_1 t_1 + \lambda_2 t_2 + \dots + \lambda_{n-1} t_{n-1}},$$

is convex because the Hessian of  $g$  is  $H(g) = g\boldsymbol{\lambda}\boldsymbol{\lambda}^T$ , and so for any  $\mathbf{v} \in \mathbb{R}^{n-1}$ ,

$$\mathbf{v}^T H(g) \mathbf{v} = (\mathbf{v}^T \boldsymbol{\lambda})^2 g \geq 0.$$

Therefore,  $\text{per}(B_k(\mathbf{t}))$ , being the sum of functions  $g(\mathbf{t})$ , is also convex, and so  $f(\mathbf{t})$  is convex as well. Therefore, by Lemma 3,  $f$  attains its maximum in  $R(\mathbf{x})$  at one of the vertices  $\mathbf{v}_i$  in (13). Now observe that

$$B_k(\mathbf{v}_i) = \begin{bmatrix} e^{x_1 u_1} & \dots & e^{x_1 u_{n-1}} \\ \vdots & & \vdots \\ e^{x_{i-1} u_1} & \dots & e^{x_{i-1} u_{n-1}} \\ e^{x_{i+1} u_1} & \dots & e^{x_{i+1} u_{n-1}} \\ \vdots & & \vdots \\ e^{x_{n-1} u_1} & \dots & e^{x_{n-1} u_{n-1}} \end{bmatrix}.$$

Multiplying its rows by the factors

$$e^{x_1 y_k}, \dots, e^{x_{i-1} y_k}, e^{x_{i+1} y_k}, \dots, e^{x_n y_k},$$

respectively, we deduce that

$$e^{s(\mathbf{x})y_k} \text{per}(B_k(\mathbf{v}_i)) = e^{x_i y_k} \text{per}(A_{ik}),$$

where  $A_{ik}$  is the submatrix of  $A$  formed by deleting row  $i$  and column  $k$ . Therefore, by the Laplace expansion of  $\text{per}(A)$  by its  $i$ -th row we find

$$f(\mathbf{v}_i) = \sum_{k=1}^n e^{x_i y_k} \text{per}(A_{ik}) = \text{per}(A), \quad \text{for all } i = 1, \dots, n.$$

Thus the maximum of  $f$  in  $R(\mathbf{x})$  is  $\text{per}(A)$ , which proves the upper bound of Theorem 2.

## 10 Gaussian matrices

A matrix related to  $A$  is the univariate Gaussian matrix

$$B := [e^{-\lambda(t_j - t_i)^2/2}]_{i,j=1,\dots,n}, \quad (16)$$

where  $t_1, \dots, t_n \in \mathbb{R}$  are distinct and  $\lambda > 0$  is a parameter. This is the matrix that arises when we interpolate data at  $t_1, \dots, t_n$  with a linear combination of the Gaussian radial basis functions

$$\phi_j(t) := e^{-\lambda(t - t_j)^2/2}, \quad j = 1, \dots, n, \quad (17)$$

see Micchelli [22], Fasshauer [9]. We obtain bounds on  $\det(B)$  from the bounds of Theorem 2. Let

$$\bar{t} := \frac{1}{n} \sum_{i=1}^n t_i, \quad S := \sum_{i=1}^n (t_i - \bar{t})^2, \quad N := \binom{n}{2}.$$

**Corollary 1** *For  $B$  in (16),*

$$\frac{e^{-\lambda S}}{c_{n-1}} \leq \frac{\det(B)}{\lambda^N V(\mathbf{t})^2} \leq \frac{1}{c_n} \text{per}(B). \quad (18)$$

*Proof.* We may assume that  $t_1 < \dots < t_n$ . We multiply both row  $i$  and column  $i$  of  $B$  by  $e^{\lambda t_i^2/2}$ , and then

$$\det(B) = e^{-\lambda(t_1^2 + \dots + t_n^2)} \det(B'),$$

where

$$B' := [e^{\lambda t_i t_j}]_{i,j=1,\dots,n}.$$

With the substitutions  $x_i := \sqrt{\lambda}t_i$  and  $y_j := \sqrt{\lambda}t_j$ , the matrix  $B'$  becomes  $A$  in (1). So the upper bound of Theorem 2 implies

$$\det(B') \leq \frac{1}{c_n} \lambda^N V(\mathbf{t})^2 \text{per}(B'),$$

and thus the upper bound for  $\det(B)$  follows.

The lower bound of Theorem 2 implies

$$\det(B') \geq \frac{1}{c_{n-1}} \lambda^N V(\mathbf{t})^2 e^{\lambda s(\mathbf{t})^2/n},$$

and so

$$\det(B) \geq \frac{1}{c_{n-1}} \lambda^N V(\mathbf{t})^2 e^{-\lambda((t_1^2 + \dots + t_n^2) - s(\mathbf{t})^2/n)}.$$

As is well known,

$$\sum_{i=1}^n t_i^2 - \frac{1}{n} \left( \sum_{i=1}^n t_i \right)^2 = S$$

(by replacing  $t_i$  by  $(t_i - \bar{t}) + \bar{t}$  in the first sum) and thus we obtain the lower bound of (18).  $\square$

Various methods have been proposed for optimizing the shape parameter in radial basis interpolation; see [25], [9, Section 5.2], [20]. As a general rule, for the Gaussians (17), it is usual to choose  $\lambda$  at least to be inversely proportional to the squares of the distances among the  $t_i$ . We can use the lower bound of Corollary 1 to make a choice of  $\lambda$  that agrees with that general rule. We can write the lower bound as

$$\det(B) \geq \frac{1}{c_{n-1}} V(\mathbf{t})^2 F(\lambda),$$

where

$$F(\lambda) := \lambda^N e^{-\lambda S}.$$

The function  $F(\lambda)$  approaches zero both as  $\lambda \rightarrow 0$  and as  $\lambda \rightarrow \infty$ , and this suggests choosing the parameter  $\lambda$  to maximize  $F$ . The maximum is attained when  $F'(\lambda) = 0$ , i.e., when

$$\lambda = \lambda_* := \frac{N}{S} = \frac{n(n-1)}{2} \bigg/ \sum_{i=1}^n (t_i - \bar{t})^2.$$

| $\lambda$               | Error, $f_1$         | Error, $f_2$ |
|-------------------------|----------------------|--------------|
| $0.01 \times \lambda_*$ | $2.8 \times 10^{-6}$ | 0.47         |
| $0.1 \times \lambda_*$  | $5.8 \times 10^{-8}$ | 0.030        |
| $\lambda_*$             | $5.4 \times 10^{-8}$ | 0.020        |
| $10 \times \lambda_*$   | 0.013                | 0.044        |
| $100 \times \lambda_*$  | 0.13                 | 0.10         |

Table 1: Effect of choice of  $\lambda$

The following numerical examples were used to test various values of  $\lambda$ . We took  $n = 101$  and  $t_i = (i - 1)/100$ ,  $i = 1, 2, \dots, 101$ , and

$$f_1(x) := x \cos(3\pi x), \quad f_2(x) := 2x\sqrt{1-x}, \quad 0 \leq x \leq 1.$$

We computed the interpolant for  $\lambda = 10^k \lambda_*$ , with  $k = -2, -1, 0, 1, 2$ . The max errors in  $[0, 1]$  for  $f_1$  and  $f_2$  are shown in Table 1, and we see that  $k = 0$  gives the smallest error for both  $f_1$  and  $f_2$ , suggesting that  $\lambda_*$  seems to offer a good choice of shape parameter. Figure 1 shows the Gauss interpolant to  $f_1$  with  $\lambda = \lambda_*$ .

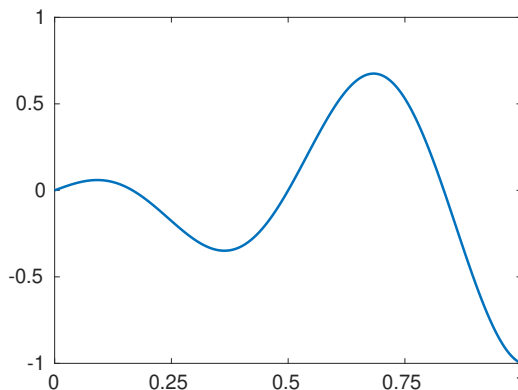


Figure 1: Gauss interpolant to  $f_1$  with  $\lambda = \lambda_*$ .

**Acknowledgement.** I wish to thank the referees for their suggestions for improvements, especially the importance of the unitary matrix integral, to Stefano de Marchi and Jonathan Novak for help with references, to Geir Dahl for the reminder about the elegant Birkhoff-von Neumann theorem, and to Espen Sande for the encouragement to persevere with this project.

## References

- [1] R. Ait-Haddou and M.-L. Mazure, *The fundamental blossoming inequality in Chebyshev spaces — I: applications to Schur functions*, Found. Comp. Math. **18** (2018), 135–158.
- [2] P. Alfeld and L. L. Schumaker, *A  $C^2$  trivariate macro-element based on the Clough-Tocher-split of a tetrahedron*, Comp. Aided Geom. Design. **22** (2005), 710–721.
- [3] R. B. Bapat and T. E. S. Raghavan, *Nonnegative matrices and applications*, Cambridge University Press, 1997.
- [4] L. Bos and S. de Marchi, *On optimal points for interpolation*, Dolomites Research Notes on Approximation **4** (2011), 8–12.
- [5] I. D. Coope and P. F. Renaud, *Trace inequalities with applications to orthogonal regression and matrix nearness problems*, J. Inequalities Pure Appl. Math. **10** (2009), Paper No. 92, pp. 1–7.
- [6] C. de Boor, *Divided differences*, Surveys in Approx. Theory **1** (2005), 46–69.
- [7] S. de Marchi, *Polynomials arising in factoring generalized Vandermonde determinants: an algorithm for computing their coefficients*, Math. Comp. Modelling **34** (2001), 271–281.
- [8] S. M. Fallat and C. R. Johnson, *Totally non-negative matrices*, Princeton University Press, 2011.
- [9] G. E. Fasshauer, *Positive definite kernels: past, present and future*, Dolomites Research Notes on Approximation **4** (2011), 21–63.
- [10] M. S. Floater, *Error formulas for divided difference expansions and numerical differentiation*, J. Approx. Theory **122** (2003), 1–9.
- [11] ———, *On a conjecture concerning interpolation by bivariate Bernstein polynomials*, J. Approx. Theory **293** (2023), 105920, pp. 1–14.
- [12] ———, *A divided difference identity for a class of multiple integrals*, preprint, 2026.

- [13] F. R. Gantmakher and M. G. Krein, *Sur les matrices complètement non négatives et oscillatoires*, Compositio Math. **4** (1937), 445–476.
- [14] K. Gross and D. St. P. Richards, *Total positivity, spherical series, and hypergeometric functions of matrix argument*, J. Approx. Theory **59** (1989), 224–246.
- [15] G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities, 2nd edition*, Cambridge University Press, 1952.
- [16] Harish-Chandra, *Differential operators on a semisimple Lie algebra*, Amer. J. Math. **79** (1957), 87–120.
- [17] E. Isaacson and H. B. Keller, *Analysis of numerical methods*, Dover, New York, 1994.
- [18] C. Itzykson and J. B. Zuber, *The planar approximation. II*, J. Math. Phys. **21** (1980), 411–421.
- [19] G. Jaklič and T. Kanduč, *On positivity of principal minors of bivariate Bezier collocation matrix*, J. Appl. Math. and Comp. **227** (2014), 320–328.
- [20] E. Larsson and B. Fornberg, *Theoretical and computational aspects of multivariate interpolation with increasingly flat radial basis functions*, Comp. Math. Appl. **49** (2005), 103–130.
- [21] I. G. Macdonald, *Symmetric functions and Hall polynomials*, Oxford University Press, 1995.
- [22] C. A. Micchelli, *Interpolation of scattered data: distance matrices and conditionally positive definite functions*, Const. Approx. **2** (1986), 11–22.
- [23] A. Pinkus, *Totally positive matrices*, Cambridge University Press, 2009.
- [24] G. Pólya and G. Szegő, *Problems and theorems in analysis II*, Springer-Verlag, New York, 1976.
- [25] S. Rippa, *An algorithm for selecting a good value for the parameter  $c$  in radial basis function interpolation*, Adv. Comp. Math. **11** (1999), 193–210.

- [26] J. F. Steffensen, *Interpolation*, Williams and Wilkins, Baltimore, 1927.
- [27] D. Yarotsky, *Univariate interpolation by exponential functions and Gaussian RBFs for generic sets of nodes*, J. Approx. Theory **166** (2013), 163–175.