Numer Algor (2015) 70:249–267
DOI 10.1007/s11075-014-9945-6

ORIGINAL PAPER

Reusing Chebyshev points for polynomial interpolation

Saman Ghili · Gianluca Iaccarino

Received: 18 February 2014 / Accepted: 27 November 2014 / Published online: 18 December 2014

Abstract Let X_l^C be the set of l Chebyshev points in the interval $[-1, 1]$. If n and n_0 are such that $n = 2^m n_0 - 1$ for some positive integer m, then $X_{n_0}^C \subset X_n^C$. This property can be utilized in order to reuse previous function values when one wants to increase the degree of the polynomial interpolation. For given n_0 and n, $n > n_0$, where $n \neq 2^m n_0 - 1$, we give a simple procedure to build a set of n points in the interval $[-1, 1]$ that include the set of n_0 Chebyshev points and have favorable interpolation properties. We show that the nodal polynomial for these points has a maximum norm that is at most $O(n)$ times larger than that of the Chebyshev points of the same size. We also present numerical evidence suggesting that the Lebesgue constant for these points grows at most linearly in n.

Keywords Polynomial interpolation · Chebyshev points · Lebesgue constant

1 Introduction

Chebyshev points of the second kind, defined as the extrema $X_n^C = \{\cos(\frac{(k-1)\pi}{n-1}) : 1 \leq k \leq n\}$ of the Chebyshev polynomials of the first kind, are a favorite choice for polynomial interpolation and integration [5, 12, 14] of functions on the interval

This work is supported by the PECASE Award, sponsored by the Lawrence Livermore National Laboratory under grant B597952 and the Office of Science of the U.S. Department of Energy under grant DE-SC0005384.

S. Ghili (✉)
Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA, USA
e-mail: samang@stanford.edu

G. Iaccarino
Department of Mechanical Engineering and Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA, USA
e-mail: jops@stanford.edu

$D = [-1, 1]$ (in the literature, they are also referred to as Clenshaw-Curtis points in the context of numerical integration). In this work, by Chebyshev points, we always mean the Chebyshev points of the second kind (as opposed to Chebyshev points of the first kind which are the roots of the same Chebyshev polynomials). Among many other reasons, their popularity is due to their very slowly growing Lebesgue constant. The Lebesgue constant for a set of n points $X_n = \{x_1, x_2, \cdots, x_n\} \subset D$ is defined as

$$\Lambda(X_n) = \sup_{f \in \mathrm{C}(D), f \neq 0} \frac{||p||_D}{||f||_D}, \tag{1.1}$$

where C(D) is the space of all continuous functions defined over the domain D, $||.||_D$ is the supremum norm over D, and p is the polynomial of degree at most $n-1$ that interpolates f at the points in X_n [16, Chapter 15].

The importance of $\Lambda(X_n)$ is due to the following error estimate. Let $P^*_{n-1}[f]$ be the best approximation to f in the supremum norm among all polynomials of degree at most $n-1$, and $P_{X_n}[f]$ be the polynomial of degree at most $n-1$ that interpolates f at the points in X_n. Then by a simple use of the triangular inequality it can be shown that

$$||f - P_{X_n}[f]||_D \leq (\Lambda(X_n) + 1)||f - P^*_{n-1}[f]||_D. \tag{1.2}$$

Therefore, the smaller $\Lambda(X_n)$, the closer our approximation to the true best polynomial approximation. It is well known that for any sequence of point-sets $\{X_n\}_{n=1}^{\infty}$, the Lebesgue constant grows at least logarithmically in n [2, 6, 16]. Chebyshev points are one system of points that have this optimal (logarithmically growing) Lebesgue constant [16].

Another feature of Chebyshev points is their nestedness property. For any n_0, the set of n_0 Chebyshev points $X^C_{n_0}$ is a subset of $n = 2n_0 - 1$ Chebyshev points X^C_n. If we have previously interpolated a function f at n_0 Chebyshev points, and now want to increase the degree of our polynomial interpolation to n, we only need to evaluate the function atthe$n_0 - 1$ new points. Other than saving computational cost when increasing the polynomial degree, this property is usful in other applications.An example is sparse grid quadratures for interpolation and integration of multivariate smoothfunctionsthatexploitthe nestednesspropertyof theClenshaw-Curtisnodes[11].

However, there might be situations where we have interpolated a function at n_0 Chebyshev points, and we want to increase the degree of our interpolating polynomial to an n that is smaller than $2n_0-1$. This can happen when the function evaluations are very expensive. Similarly, when we are interpolating a multivariate function at tensor-product Chebyshev points, doubling the number of points in each variable would increase the total number of points by a factor of 2^d (d is the number of independent variables), which depending on d and n_0, may become prohibitively expensive. If some variables are not as important as others, one could double the number of points only in the important variables. However, if the variable are equally important, one would have to settle for a smaller increase in the number of points i.e., an increase by a factor that is smaller than 2 in each direction. With this motivation, for a set of n_0 Chebyshev points, $X^C_{n_0}$ and an $n < 2n_0 - 1$, we try to find a set of n points $\tilde{X}_{n,n_0}$ that is a superset of $X^C_{n_0}$, and at the same time has a reasonably small Lebesgue constant.

It is worth mentioning that there are sets of interpolation points called Leja points [1, 9] that allow for reusing function values from smaller sets of points. Given an arbitrary point $x_1^L \in D$, the set of Leja points X_n^L is defined recursively by the equation

$$x_n^L = \underset{x \in D}{\operatorname{argmax}} \prod_{i=1}^{n-1} |x - x_i^L|. \tag{1.3}$$

Note that this optimization may have more than one solution. Therefore the Leja points are not unique. We can see from equation (1.3) that for any n_1 and n_2 with $n_1 < n_2$, we have $X_{n_1}^L \subset X_{n_2}^L$. Although Leja points are completely nested, they are not suited for our purpose. First, their Lebesgue constant has only been shown to grow sub-exponentially in n [13] (compare this with the logarithmic growth of the Lebesgue constant for Chebyshev points). Besides, here we are trying to add $n - n_0$ new points to an already fixed set of n_0 Chebyshev points, and not build a completely nested set of points. It is possible to start from the n_0 Chebyshev points and find the new points using (1.3), but this will result in a sub-optimal set of points because if we add the points one by one, we will not be taking full advantage of the freedom in choosing the $n - n_0$ extra points.

The rest of this paper is organized as follows. Section 2 presents a procedure for constructing a set of $n < 2n_0 - 1$ points that includes n_0 Chebyshev points, and analyze some of its properties. In particular, we show that the nodal polynomial for these nodes has a supremum norm that is at most $O(n)$ times larger than that of the set of Chebyshev points of the same size. Section 3 presents numerical results which show that the Lebesgue constant for these nodes grows at most linearly in n. We also study the norm of the inverse of the interpolation matrix (with Chebyshev polynomials as basis polynomials) for these points, and measure its performance in approximating two test functions. The concluding remarks are given in Section 4.

2 Description and analysis of the procedure

Our method for construction the set of nodes is somewhat similar to the idea of "mock-Chebyshev" points used by Boyd [3] for sampling a subset of uniformly distributed points that mimic the behavior of Chebyshev points. In other words, we want to find a set of n points in the interval $[-1, 1]$ that i) include our n_0 Chebyshev points, ii) are distributed like Chebyshev points, i.e., according to $\frac{1}{\pi\sqrt{1-x^2}}$, and iii) there are no two points that are too close to each other. Before we describe the algorithm for generating the points, we state and prove a very simple proposition that we will use later.

Proposition 1 *Let n_0 and n be two positive integers with $n_0 < n < 2n_0 - 1$, and let $h_0 = \frac{\pi}{2(n_0-1)}$ and $h = \frac{\pi}{2(n-1)}$. Let $\Theta_{n_0} = \{\theta_j^{(n_0)} = 2(j-1)h_0 : 1 \leq j \leq n_0\}$. Also let $I_1 = [0, h)$, $I_n = [(2n-3)h, \pi]$, and $I_k = [(2k-3)h, (2k-1)h)$ for $2 \leq k \leq n-1$. If we show the cardinality of a set by $|.|$, then the following statements are true:*

1. *$|I_k \cap \Theta_{n_0}| \leq 1$, $1 \leq k \leq n$.*
2. *For any $2 \leq k \leq n-1$, if $|I_k \cap \Theta_{n_0}| = 0$, then $|I_{k-1} \cap \Theta_{n_0}| = 1$ and $|I_{k+1} \cap \Theta_{n_0}| = 1$.*
3. *If $\theta_j^{(n_0)} \in I_{k-1}$ and $\theta_{j+1}^{(n_0)} \in I_{k+1}$, then $\frac{1}{2}(\theta_j^{(n_0)} + \theta_{j+1}^{(n_0)}) \in I_k$.*

Proof In the following proofs we use the the inequalities $h_0 < 2h < 2h_0$. To see this, subtract one from $n_0 < n < 2n_0 - 1$ to get $n_0 - 1 < n - 1 < 2(n_0 - 1)$. Now, inverting and multiplying by π gives us $h_0 < 2h < 2h_0$.

1. The length of each interval I_k is at most $2h$, while the distance between any two $\theta_j^{(n_0)}$ and $\theta_{j'}^{(n_0)}$ is at least $2h_0 > 2h$. Therefore there can not be two or more of them in any I_k.
2. If there is no $\theta_j^{(n_0)}$ in any two neighboring I_k's, then there is no $\theta_j^{(n_0)}$ in a sub-interval of length at least $3h_0$, which contradicts the fact that the distance between any two neighboring $\theta_j^{(n_0)}$'s is $2h_0$.
3. The distance between $\frac{1}{2}(\theta_j^{(n_0)} + \theta_{j+1}^{(n_0)})$ and any of $\theta_j^{(n_0)}$ or $\theta_{j+1}^{(n_0)}$ is h_0, while the distance between I_{k-1} and I_{k+1} is $2h > h_0$. Therefore $\frac{1}{2}(\theta_j^{(n_0)} + \theta_{j+1}^{(n_0)})$ must be in I_k. □

Notice that for the set of n Chebyshev points X_n^C, we have $x_k^{(n)} = \cos(\theta_k^{(n)})$, where $\theta_1^{(n)} = 0$, $\theta_n^{(n)} = \pi$, and $\theta_k^{(n)}$ is the middle point of I_k for all $2 \le k \le n-1$. Now if we look at all the n subintervals, n_0 of them have a $\theta_j^{(n_0)}$ in them (there are n_0 $\theta_j^{(n_0)}$'s, and each interval contains at most one of them). For the remaining $n - n_0$ intervals that do not contain any $\theta_j^{(n_0)}$'s, based on Proposition 2.1, their adjacent intervals contain a $\theta_j^{(n_0)}$ each, and their average is in I_k. So, if we set $\tilde{\Theta}_{n,n_0}$ to be the union of Θ_{n_0} with the set of $n - n_0$ averages that we compute for each I_k with $I_k \cap \Theta_{n_0} = \emptyset$, then the set $\tilde{\Theta}_{n,n_0}$ will have the following properties:

1. $|I_k \cap \tilde{\Theta}_{n,n_0}| = 1, 1 \le k \le n$.
2. The minimum distance between any two points in $\tilde{\Theta}_{n,n_0}$ is $h_0 = \frac{\pi}{2(n_0-1)}$.

These two properties mean that, roughly speaking, the points in $\tilde{\Theta}_{n,n_0}$ are uniformly distributed in $[0, \pi]$, and no two of them are too close to each other. The following summarizes this procedure:

Algorithm 1: The procedure for constructing $\tilde{\Theta}_{n,n_0}$.

$h_0 \leftarrow \frac{\pi}{2(n_0-1)}$
$h \leftarrow \frac{\pi}{2(n-1)}$
$\Theta_{n_0} \leftarrow \{\theta_j^{(n_0)} = 2(j-1)h_0 : 1 \le j \le n_0\}$
$\tilde{\Theta}_{n,n_0} \leftarrow \Theta_{n_0}$
Let $I_1 = [0, \pi h)$, $I_n = [(2n-3)h, \pi]$, and $I_k = [(2k-3)h, (2k-1)h)$ for $2 \le k \le n-1$
for $k = 1 : n$ **do**
  if $I_k \cap \Theta_{n_0} = \emptyset$ **then**
    $\tilde{\Theta}_{n,n_0} \leftarrow \tilde{\Theta}_{n,n_0} \cup \{\frac{1}{2}[(I_{k-1} \cap \Theta_{n_0}) + (I_{k+1} \cap \Theta_{n_0})]\}$ *
  end
end

*: Based on Proposition 2.1, if $I_k \cap \Theta_{n_0} = \emptyset$, then $I_{k-1} \cap \Theta_{n_0}$ and $I_{k+1} \cap \Theta_{n_0}$ will both have exactly one element. This is why we abuse the notation to take their average.

We finally build $\tilde{X}_{n,n_0}$ from $\tilde{\Theta}_{n,n_0}$ by taking the cosine of each of its members, $\tilde{X}_{n,n_0} = \{\tilde{x} = \cos(\tilde{\theta}) : \tilde{\theta} \in \tilde{\Theta}_{n,n_0}\}$. Figure 1 shows the nodes (both θ_j and x_j) for $n_0 = 20$, and four different values of n: 21, 26, 32, and 38.

Now we show that the nodal polynomial for these points will be at most $O(n)$ times larger than that of the Chebyshev points in the supremum norm. But first let us review the Lagrange interpolation theorem to see why this is an important quantity. The following is Theorem 3.2 in [7].

Theorem 1 *(Lagrange's Interpolation Theorem) Given a function f that is defined at a set of n distinct nodes $X_n = \{x_1, x_2, \cdots, x_n\}$ in $D = [-1, 1]$, there exists a unique polynomial of degree at most $n-1$, $P_{n-1}(x)$ such that*

$$P_{n-1}(x_i) = f(x_i), \qquad 1 \leq i \leq n.$$

This polynomial is given by

$$P_{n-1}(x) = \sum_{i=1}^{n} f(x_i) L_i(x),$$

where $L_i(x)$ is defined by

$$L_i(x) = \frac{w_{X_n}(x)}{(x - x_i) w'_{X_n}(x)} = \frac{\prod_{j=1, j \neq i}^{n} (x - x_j)}{\prod_{j=1, j \neq i}^{n} (x_i - x_j)}, \tag{2.1}$$

$w_{X_n}(x)$ being the nodal polynomial $w_{X_n}(x) = \prod_{j=1}^{n} (x - x_j)$.

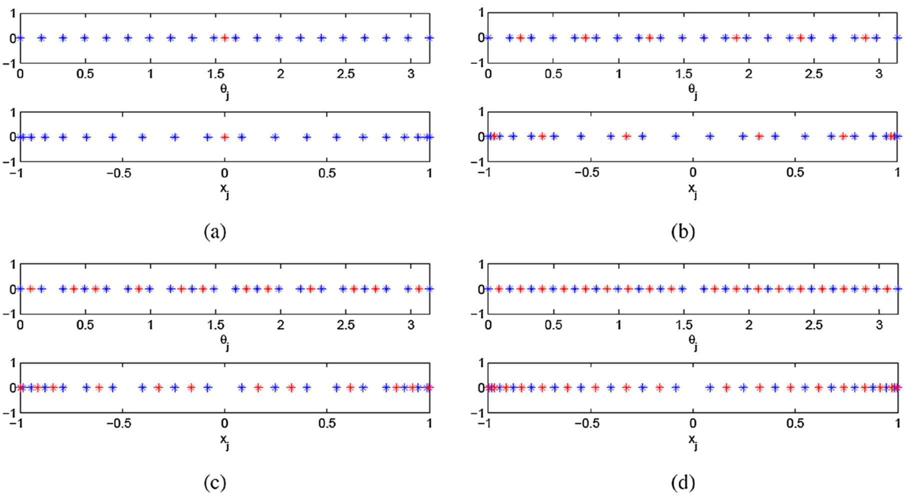

Fig. 1 $n_0 = 20$ Chebyshev points (*blue*) and $n - n_0$ added points (*red*), for **a** $n = 21$, **b** $n = 26$, **c** $n = 32$, **d** $n = 38$

Additionally, if $f \in C^n[-1, 1]$*, then for any* $x \in D$ *there exists a* $\zeta_x \in (-1, 1)$*, such that*

$$f(x) - P_{n-1}(x) = \frac{f^{(n)}(\zeta_x)}{n!} w_{X_n}(x). \tag{2.2}$$

The error expression in equation (2.2) suggests that a set of points with a smaller upper bound on $|w_{X_n}(x)|$ should in general give us a smaller interpolation error. In fact $|w_{X_n}(x)|$ is bounded by $\frac{1}{2^{n-1}}$ and $\frac{1}{2^{n-2}}$ for the roots and extrema of Chebyshev polynomials (Chebyshev points of the first and second kinds), respectively [12]. For a set of points constructed using our procedure we have the following theorem.

Theorem 2 *For a positive integer* n*, let the intervals* I_k *be as in Proposition 2.1. Let* $\Theta_n = \{\theta_1, \theta_2, \cdots, \theta_n\}$ *be such that* $\theta_1 = 0$*,* $\theta_n = \pi$*, and* $\theta_k \in I_k$ *for* $2 \leq k \leq n-1$*. If we define* $w_{\Theta_n}(\theta) = \prod_{j=1}^{n}(\cos(\theta) - \cos(\theta_j))$*, we have*

$$|w_{\Theta_n}(\theta)| \leq \frac{\pi(n-1)}{2^{n-2}}.$$

Trivially, the same bound holds for $|w_{X_n}(x)|$*, where* $x_j = \cos(\theta_j)$*,* $1 \leq j \leq n$*.*

Proof See Appendix A. □

The connection between this bound and the bound on the Lebesgue constant can be seen through equation (2.1). The Lebesgue constant for X_n can be computed by the following formula [16, Chapter 15]:

$$\Lambda(X_n) = \sup_{x \in D} \sum_{i=1}^{n} |L_i(x)|. \tag{2.3}$$

We can use Theorem 2 to bound each $|L_i(x)|$, which can in turn be used to bound $\Lambda(X_n)$. Also note that the conditions of Theorem 2 are less restrictive than those we imposed on our points. More specifically, it does not require them to be far enough from each other. This latter condition is needed to prevent the denominator of $|L_i(x)|$ from becoming too small. In fact, using this condition and Theorem 2, it is not very hard to prove an upper bound on $|L_i(x)|$ that is linear in n, which in turn gives us a quadratic upper bound on the Lebesgue constant. However, we do not include such a proof here since our numerical results suggest that the Lebesgue constant grows at most linearly in n, and therefore a quadratic bound would be very pessimistic.

3 Numerical results

In this section we test the performance of polynomial interpolation at the points obtained by the procedure given in the previous section. To do this, we look at both the Lebesgue constant and the norm of the inverse of the interpolation matrix (for the definition of the interpolation matrix see Section 3.2), and study their growth as a function of the number of interpolation points. We also look at the L_∞ and L_2 approximation errors when we interpolate a test function at these points for a fixed n

and different values of n_0, and finally, we look at the L_2 errors when we interpolate a function of three variables at a tensor-product grid constructed on the basis of these points.

3.1 The Lebesgue constant

Assume that we have n points in the interval $[-1, 1]$ obtained by adding $n - n_0$ points to n_0 Chebyshev points ($n_0 < n < 2n_0 - 1$) using our procedure. First, let us consider the rather extreme cases where n is either very close to n_0 or very close to $2n_0 - 1$. Figure 2 shows the growth of the Lebesgue constant with the number of interpolation points for six different cases, $n = n_0 + 1$, $n = n_0 + 3$ $n = n_0 + 5$, $n = 2n_0 - 6$, $n = 2n_0 - 4$, and $n = 2n_0 - 2$. The Lebesgue constant for Chebyshev points has also been plotted for reference. It can be seen that the Lebesgue constant grows very slowly when we add only one point to $n_0 = n - 1$ Chebyshev points, and that it is slightly faster when $n = n_0 + 3$ or $n = n_0 + 5$. In both cases the growth of the Lebesgue constant is very close to that of the Chebyshev points. On the other hand, when n is very close to $2n_0 - 1$, the Lebesgue constant grows much faster, but this fast growth rate drops quickly as we move away from $2n_0 - 1$. It is worth emphasizing that our points are not any closer to n Chebyshev points when $n = n_0 + 1$ compared to $n = 2n_0 - 2$, therefore the smaller Lebesgue constant is not the result of the points being closer to the actual Chebyshev points.

By looking at Fig. 2, one can see that for any n, the Lebesgue constant is the largest when $n = 2n_0 - 2$ compared to all the other values of n_0 in that plot. We experimented with various values of n and n_0, and for any fixed n, the Lebesgue constant was the largest when $n = 2n_0 - 2$, compared to any other n_0 with $\frac{n+2}{2} < n_0 < n$ (for which $n_0 < n < 2n_0 - 2$). Additionally, for $n = 2n_0 - 2$, the plots suggest that the Lebesgue constant grows linearly with n. In fact, in our experiments, the Lebesgue constant is

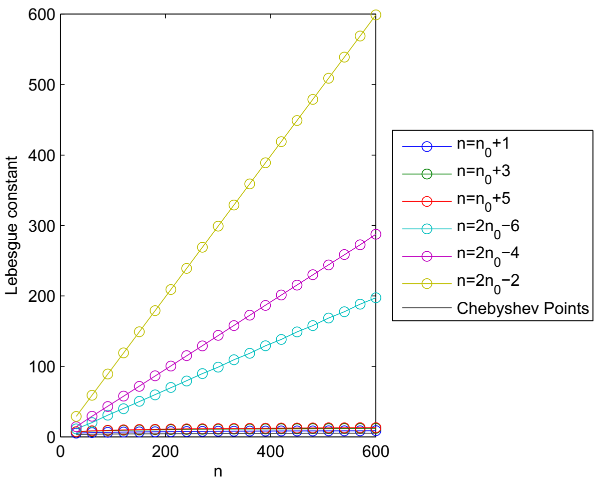

Fig. 2 The Lebesgue constant as a function of n when n is very close to either n_0 or $2n_0 - 1$

exactly $n - 1$. These observations suggest that the Lebesgue constant, for any set of n points constructed by this procedure, should be at most $n - 1$.

Figure 2 tells us about the growth of the Lebesgue constant when n is obtained by either adding a small number to n_0 or subtracting a small number from $2n_0 - 1$. However, from a practical standpoint, it might be better to study the behavior of the Lebesgue constant when n is proportional to n_0. Figure 3 shows the growth of the Lebesgue constant as a function of n, when $n = \lfloor \alpha n_0 \rfloor$, for four different values of α (note that α must be between 1 and 2), along with the Lebesgue constants for the fast Leja points of the same size [1]. Roughly the same pattern can be seen here: the Lebesgue constant tends to be larger for larger values of α.

We also see that for all four values of α, and particularly for large values of n, the Lebesgue constant for the points constructed by our procedure is significantly smaller than that of the fast Leja points.

The observation that the growth of the Lebesgue constant is slower for smaller α, i.e., when n is close to n_0, could be important from a practical point of view. In practice, the cases where n is close to $2n_0 - 1$ are much less common. For example, if it is not too expensive to evaluate the target function at $n = \lfloor 1.95 n_0 \rfloor$ points, then for a little extra cost we can evaluate the function at $n = 2n_0 - 1$ Chebyshev points.

All the Lebesgue constants in this section were computed using the "Chebfun" package [15].

3.2 Norm of the inverse of the interpolation matrix

Besides the Lebesgue constant, the condition number [4, 8] or the norm of the inverse (the smallest singular value) [10] of an interpolation matrix constructed using appropriate basis functions, can also be used in order to asses the quality of an interpolation

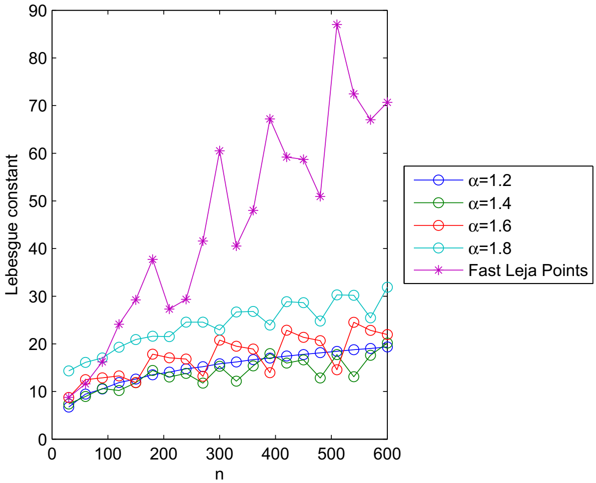

Fig. 3 The Lebesgue constant as a function of n when n is related to n_0 by $n = \lfloor \alpha n_0 \rfloor$

procedure. Here, we look at the norm of the inverse of the interpolation matrix constructed from orthonormal Chebyshev polynomials, $\{\hat{T}_0(x), \hat{T}_1(x), \cdots, \hat{T}_{n-1}(x)\}$. The orthonormality property satisfied by these polynomials is

$$\int_{-1}^{1} \hat{T}_i(x)\hat{T}_j(x)w(x)dx = \delta_{ij}, \qquad 0 \le i, j \le n-1, \tag{3.1}$$

where $w = \frac{1}{\pi\sqrt{1-x^2}}$ is the weight function and δ_{ij} is the Kronecker delta. Monic Chebyshev polynomials $T_j(x) = \cos(j \arccos(x))$ are orthogonal with respect to $w(x)$, and we have

$$\int_{-1}^{1} T_j^2(x)w(x)dx = \begin{cases} 1, & i = j = 0; \\ \frac{1}{2}, & j \ge 1. \end{cases} \tag{3.2}$$

Therefore, in order to obtain the orthonormal Chebyshev polynomials we only need to multiply the monic Chebyshev polynomials (except the one with degree zero) by $\sqrt{2}$, i.e.,

$$\hat{T}_j(x) = \begin{cases} T_j(x), & j = 0; \\ \sqrt{2}T_j(x), & 1 \le j \le n-1. \end{cases} \tag{3.3}$$

For a set of n points $X_n = \{x_1, x_2, \cdots x_n\} \subset D$, we define the interpolation matrix with orthonormal Chebyshev polynomials of degree up to $n-1$ as basis functions as follows:

$$P = \frac{1}{\sqrt{n}} \begin{bmatrix} \hat{T}_0(x_1) & \hat{T}_1(x_1) & \cdots & \hat{T}_{n-1}(x_1) \\ \hat{T}_0(x_2) & \hat{T}_1(x_2) & \cdots & \hat{T}_{n-1}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{T}_0(x_n) & \hat{T}_1(x_n) & \cdots & \hat{T}_{n-1}(x_n) \end{bmatrix}. \tag{3.4}$$

For a function f and a set of interpolation points X_n, let $p = \sum_{j=0}^{n-1} a_j \hat{T}_j(x)$ be the polynomial of degree at most $n-1$ that interpolates f at the nodes in X_n. If we let $b = [f(x_1), f(x_2), \cdots, f(x_n)]^T$, and $a = [a_0, a_1, \cdots, a_{n-1}]^T$, then one can easily verify that,

$$Pa = \frac{1}{\sqrt{n}} b. \tag{3.5}$$

The importance of $\|P^{-1}\|_2$ is due to the following Theorem.

Theorem 3 *Let $D = [-1, 1]$ and let $f : D \to \mathbb{R}$ have a convergent expansion $f = \sum_{j=0}^{\infty} a_j \hat{T}_j(x)$, where $\hat{T}_j(x)$ is the normalized Chebyshev polynomial of degree j. For a set of points, $X_n = \{x_1, x_2, \cdots x_n\} \subset D$, let P be the associated interpolation matrix and p be the polynomial of degree at most $n-1$ that interpolates f at the points in X_n. If there exist $C > 0$ and $s > 1$ such that for any $k \ge n$, $|a_k| \le C(k+1)^{-s}$, then*

$$\|f - p\|_{L_2(D,w)} \le C \left[\frac{1}{\sqrt{2s-1}} n^{-(s-\frac{1}{2})} + \frac{\sqrt{2}}{s-1} \|P^{-1}\|_2 n^{-(s-1)} \right], \tag{3.6}$$

or if there exist $C > 0$ *and* $\rho > 1$ *such that for any* $k \geq n$, $|a_k| \leq C\rho^{-n}$, *then*

$$\|f - p\|_{L_2(D,w)} \leq C\rho^{-n} \left[\frac{1}{\sqrt{1-\rho^{-2}}} + \frac{\sqrt{2}}{1-\rho^{-1}} \|P^{-1}\|_2 \right], \tag{3.7}$$

where $\|.\|_{L_2(D,w)}$ *is the weighted* L_2 *norm on* D *with* $w(x) = \frac{1}{\pi\sqrt{1-x^2}}$.

Proof See Appendix B. □

Equations (3.6) and (3.7) show how $\|P^{-1}\|_2$ can control the L_2 error of interpolation. It is worth mentioning that $\|P^{-1}\|_2$ is equal to 1 for Chebyshev points of the first kind, and about $\sqrt{2}$ for Chebyshev points of the second kind. Also note that inequalities of the type $|a_k| \leq C(k+1)^{-s}$ and $|a_k| \leq C\rho^{-n}$ hold when f has a number of continuous derivatives, or an analytic continuation in a neighborhood of D in the complex plane [14].

We now look at the norm of the inverse of P for the nodes constructed by our procedure. As before, we plot $\|P^{-1}\|_2$ as a function of the grid size n for four different cases: $n = n_0 + 1$, $n = n_0 + 3$, $n = 2n_0 - 4$, and $n = 2n_0 - 2$. Figure 4 shows the plot of $\|P^{-1}\|_2$ when n is very close to either n_0 or $2n_0 - 1$. The growth rate of $\|P^{-1}\|_2$ in this case is not necessarily faster when n is close to $2n_0 - 1$.

Like Fig. 3, Fig. 5 shows the plot of $\|P^{-1}\|_2$ as a function of n, when $n = \lfloor \alpha n_0 \rfloor$, for different values of α, along with that of the fast Leja points. Again, it can be seen that the norm of the inverse of the interpolation matrix is smaller for the points constructed by our procedure compared to the fast Leja points, especially for large n. These plots of $\|P^{-1}\|_2$ as a function of n (with different values for n_0) suggest that $\|P^{-1}\|_2$ grows sub-linearly with n.

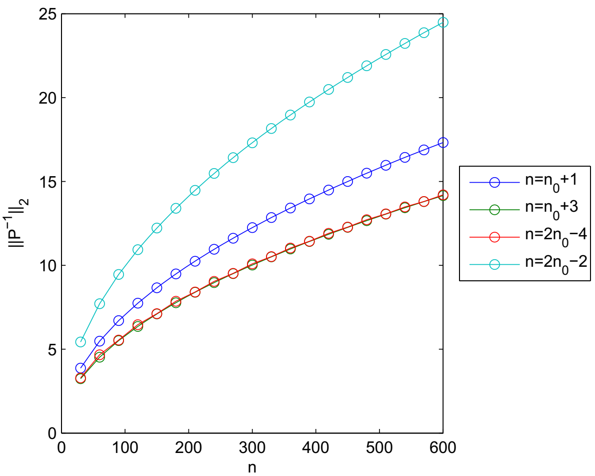

Fig. 4 $\|P^{-1}\|_2$ as a function of n when n is very close to either n_0 or $2n_0 - 1$

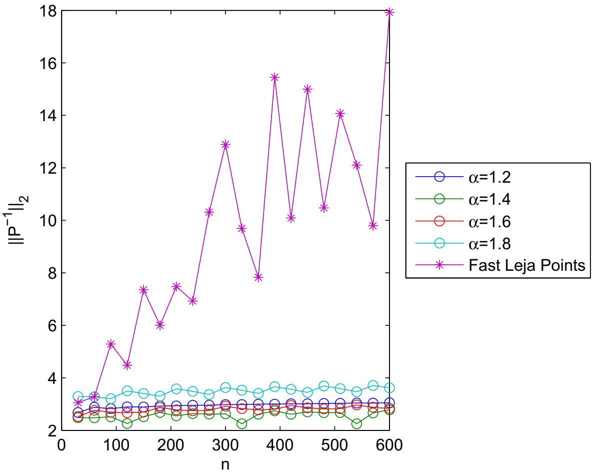

Fig. 5 $\|P^{-1}\|$ as a function of n when n is related to n_0 by $n = \lfloor \alpha n_0 \rfloor$

3.3 Interpolating two test functions

We finish this section by using the points constructed by our procedure to interpolate two test functions, and compare the results to the results obtained from the same number of Chebyshev points. First, we look at a function of one variable, and then we interpolate a function defined on the cube $[-1, 1]^3$ at a tensor-product grid constructed from our points.

3.3.1 One-dimensional example

Consider the function $f(x) = \frac{1}{1+25(x+0.4)^2}$ defined in $[-1, 1]$. Figure 6 shows the absolute error of interpolation at $n = 45$ points, for four different values of n_0. The absolute error from interpolation at $n = 45$ Chebyshev points has also been plotted for reference. Again, we see that the approximation is more accurate when n is closer to n_0.

Now let us look at the relative L_2 error $E = \frac{||f(x)-\tilde{p}(x)||_{L_2([-1,1],w)}}{||f(x)||_{L_2([-1,1],w)}}$ when we approximate f by a polynomial $\tilde{p}(x)$ that interpolates it at the same $n = 45$ interpolation nodes. Here, we approximate the integrals that appear in the expression for E using the Clenshaw-Curtis quadrature rule with 4000 quadrature points and treat them as the exact values for the L_2 norms. The relative errors for different values of n_0, and also $n = 45$ Chebyshev points are given in Table 1.

Figure 7 shows a comparison of the Chebyshev expansion coefficients of f (in absolute value) computed using different nodes.

3.3.2 Three-dimensional example

As we mentioned earlier, if we have a function in d dimensions that we are interpolating at tensor-product Chebyshev points, then doubling the number of points in

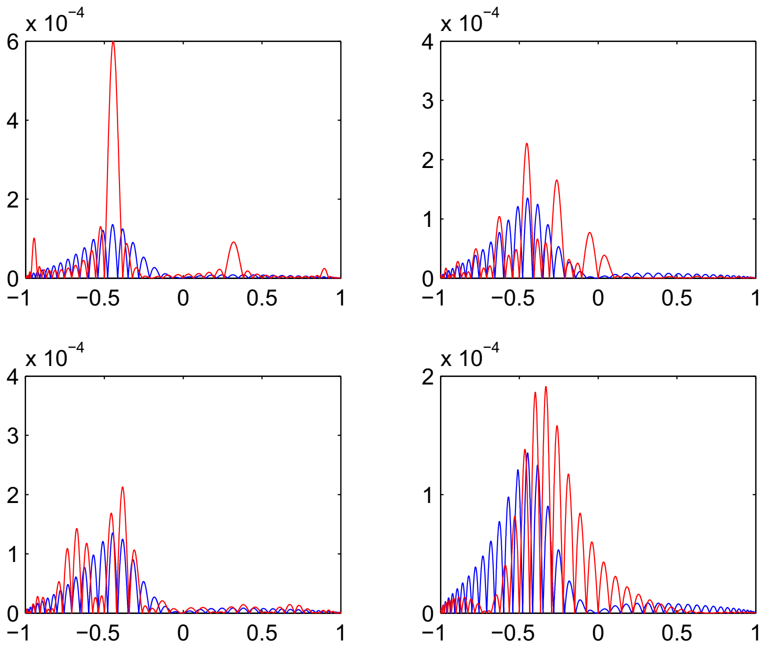

Fig. 6 *Horizontal axis*: x. *Vertical axis*: absolute error in interpolating f at $n = 45$ points. *Top left*: $n_0 = 25$. *Top right*: $n_0 = 31$. *Bottom left*: $n_0 = 37$. *Bottom right*: $n_0 = 43$. The absolute error for 45 Chebyshev points is plotted (*blue curves*) for comparison

each direction will increase the total number of points by a factor of 2^d, and in order to keep the final number of points in a manageable range, we may need to settle for a smaller increase in the number of points than a factor of 2. Here, we give an example of this in three dimensions and measure the quality of the resulting sets of interpolation points by interpolating a test function at those points, computing the relative L_2 error, and comparing the result to the relative L_2 error of interpola ting f at tensor-product Chebyshov points of the same size. Consider the function $f(x_1, x_2, x_3) = e^{2(x_1,+x_2+x_3)}$ defined in $[-1, 1]^3$, and assume that we have initially interpolated this function at a set of tensor-product Chebyshev points with $n_0 = 10$ points in each direction. The total number of initial points is $N_0 = n_0^3 = 1000$. We build three sets of tensor-product interpolation points based on one-dimensional

Table 1 The relative L_2 errors in approximating f by polynomial interpolation at $n = 45$ points, with four different values of n_0

$n_0 = 25$	$E = 2.5627e - 004$
$n_0 = 31$	$E = 1.2197e - 004$
$n_0 = 37$	$E = 1.2805e - 004$
$n_0 = 43$	$E = 1.2685e - 004$
Chebyshev points	$E = 9.3368e - 005$

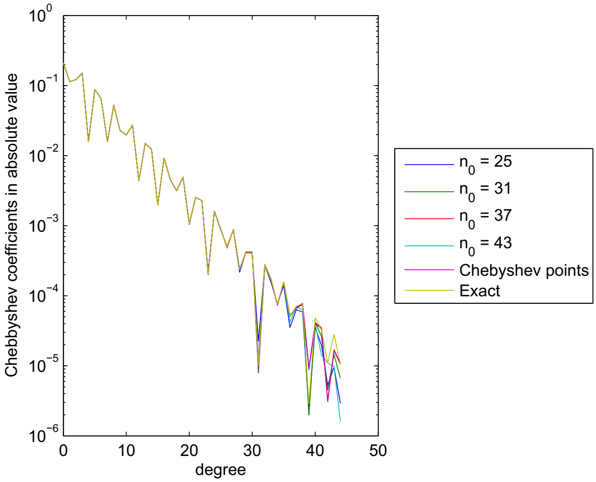

Fig. 7 Chebyshev coefficients of f in absolute value, for different values of n_0, and Chebyshev points

points obtained by adding one, two, and three points to the original 10 Chebyshev points. The total number of points $N = n^3$ will be 1331, 1728, and 2197, respectively. Figure 8 shows the original tensor-product grid and the grid based on adding one point in each direction. Figure 9 shows the tensor-product grids based on adding two and three points in each direction.

Let $\tilde{p}(x_1, x_2, x_3)$ be the polynomial that interpolates f at these points and whose degree in each variable is smaller than n. Here, we look at the relative L_2 error of interpolation defined as $E = \frac{||f(x_1,x_2,x_3)-\tilde{p}(x_1,x_2,x_3)||_{L_2([-1,1]^3,w)}}{||f(x_1,x_2,x_3)||_{L_2([-1,1]^3,w)}}$, where $w(x_1, x_2, x_3) = \prod_{k=1}^{3} \frac{1}{\pi\sqrt{1-x_k^2}}$. Table 2 summarizes the relative L_2 errors in approximating f by interpolation at the grid points shown in Figs. 8 and 9, and tensor-product Chebyshev points of the same size. It can be seen that the accuracy lost due to the fact that our points are not exactly Chebyshev points is very small, especially when n is close to n_0.

Table 2 The relative L_2 errors in approximating f at tensor-product grids based on various sets of points in one dimension

	Tensor-product grid based on $\tilde{X}_{n,n_0}$	Tensor-product grid based on X_n^C
$n_0 = 10, n = 10$	–	$3.1215e - 007$
$n_0 = 10, n = 11$	$2.8241e - 008$	$2.8147e - 008$
$n_0 = 10, n = 12$	$2.7717e - 009$	$2.3295e - 009$
$n_0 = 10, n = 13$	$2.5194e - 010$	$1.7815e - 010$

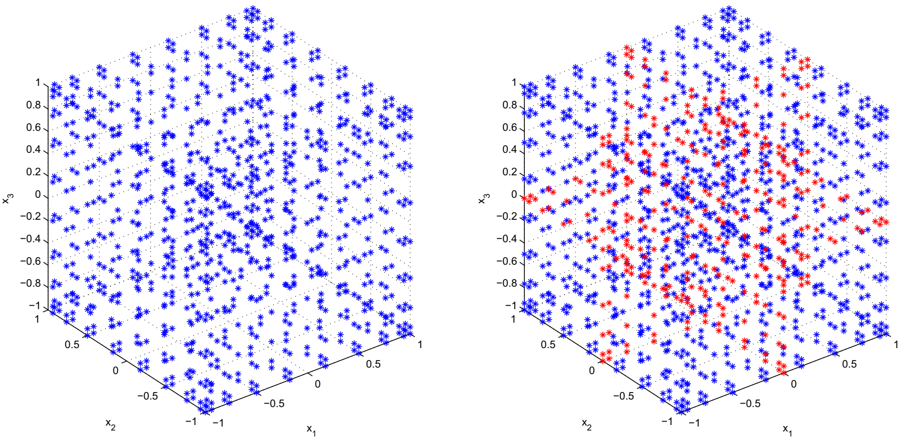

Fig. 8 The *blue stars* are the initial points and the *red stars* are the added points. *Left*: Tensor-product Chebyshev points with $n_0 = 10$ points in each direction. *Right*: Tensor-product of $n_0 = 10$ and one added point ($n = 11$) in each direction

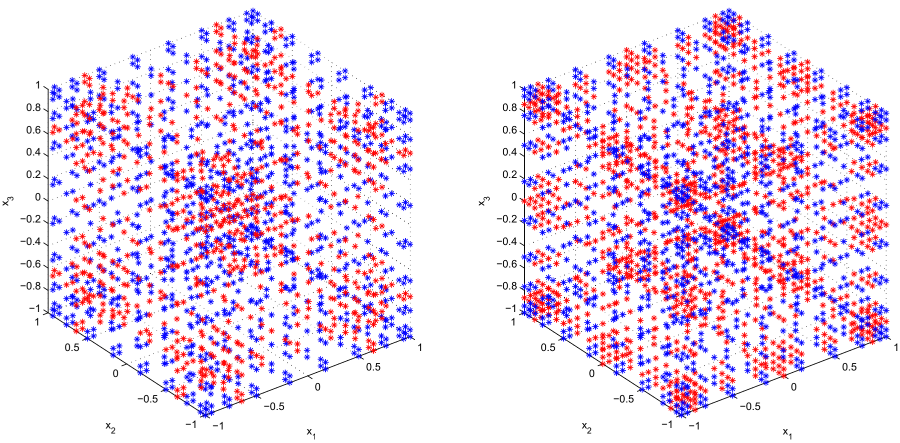

Fig. 9 *Blue stars* are the initial points and the *red stars* are the added points. *Left*: Tensor-product of $n_0 = 10$ and 2 added points ($n = 12$) in each direction. *Right*: Tensor-product of $n_0 = 10$ and 3 added point ($n = 13$) in each direction

4 Conclusion

We presented a procedure for constructing interpolation points of arbitrary size n, that include a set of n_0 Chebyshev points in the interval $D = [-1, 1]$. It was shown that the maximum norm of the monic nodal polynomial for these points is at most $O(n)$ times larger than that of Chebyshev points, and numerical evidence was presented suggesting that the Lebesgue constant and the norm of the inverse of the interpolation matrix associated with these points grow slowly with n.

Appendix A: Proof of Theorem 2

Proof First note that at the nodes $\theta_i, i = 1, 2, \cdots, n$, we have $w_{\Theta_n}(\theta_i) = 0$, which trivially satisfies the bound. So it remains to show that the bound holds for $\theta \neq \theta_i$. For an arbitrary i, $1 \leq i \leq n-1$, consider the interval $I = [\theta_i, \theta_{i+1}]$. Now set $h = \frac{\pi}{2(n-1)}$ as before, and let $\Phi = \{\phi_1, \phi_2, \cdots, \phi_n\}$ be another set of points in the interval $[0, \pi]$ with the following properties:

1. $\phi_1 = 0$ and $\phi_n = \pi$.
2. If $i > 1$, for all $1 < j \leq i$, $\phi_j = (2j-3)h$. This is the leftmost point in the interval $[(2j-3)h, (2j-1)h]$ which is the closure of the interval which θ_j belongs to.
3. If $i < n-1$, for all $i+1 \leq j < n$, $\phi_j = (2j-1)h$. Again, this is the rightmost point in the interval $[(2j-3)h, (2j-1)h]$ which is the closure of the interval which θ_j belongs to.

These properties imply that for any $1 \leq j \leq n$, and any $\theta \in I$, we have $|\theta - \theta_j| \leq |\theta - \phi_j|$, which in turn means that $|\cos(\theta) - \cos(\theta_j)| \leq |\cos(\theta) - \cos(\phi_j)|$. By multiplying these for all j we get

$$\prod_{j=1}^{n} |\cos(\theta) - \cos(\theta_j)| \leq \prod_{j=1}^{n} |\cos(\theta) - \cos(\phi_j)|, \qquad \theta \in I,$$

or

$$|w_{\Theta_n}(\theta)| \leq |w_\Phi(\theta)|, \qquad \theta \in I.$$

Since $|w_{\Theta_n}(\theta)| \leq |w_\Phi(\theta)|$ in I, an upper bound for $|w_\Phi(\theta)|$ will also bound $|w_{\Theta_n}(\theta)|$ in this interval. Also, we have $I = [\theta_i, \theta_{i+1}] \subseteq [\phi_i, \phi_{i+1}]$. Therefore, it is sufficient to prove the bound for $|w_\Phi(\theta)|$ in the interval $[\phi_i, \phi_{i+1}]$. But note that for $2 \leq j \leq n-1$, the ϕ_j's are $n-2$ (out of $n-1$) zeros of $\cos((n-1)\theta)$, and $\bar{\phi} = (2i-1)h$ is the only one that is missing. Using this, and the identity $\cos((n-1)\theta) = 2^{n-2} \prod_{j=1}^{n-1} (\cos(\theta) - \cos((2j-1)h))$, for $\theta \neq \bar{\phi}$ we can write

$$\begin{aligned} w_\Phi(\theta) &= \frac{(\cos(\theta)-1)(\cos(\theta)+1)}{\cos(\theta)-\cos(\bar{\phi})} \left(\cos(\theta) - \cos(\bar{\phi})\right) \prod_{j=2}^{n-1} (\cos(\theta) - \cos(\phi_j)) \\ &= \frac{1}{2^{n-2}} \frac{-\sin^2(\theta)\cos((n-1)\theta)}{\cos(\theta) - \cos(\bar{\phi})}. \end{aligned} \tag{A.1}$$

Since $w_\Phi(\theta)$ is continuous, its value at $\phi = \bar{\phi}$ can be obtained by computing the limit when θ goes to $\bar{\phi}$ in equation (A.1). Using the 'Hospital rule,

$$\begin{aligned} w_\Phi(\bar{\phi}) &= \frac{1}{2^{n-2}} \frac{-\sin(2\bar{\phi})\cos((n-1)\bar{\phi}) + (n-1)\sin^2(\bar{\phi})\sin((n-1)\bar{\phi})}{-\sin(\bar{\phi})} \\ &= -\sin(\bar{\phi})\sin((n-1)\bar{\phi})\frac{n-1}{2^{n-2}}. \end{aligned} \tag{A.2}$$

Taking the absolute value of equation (A.2) yields

$$|w_\Phi(\phi)| \leq \frac{n-1}{2^{n-2}}. \tag{A.3}$$

Now it remains to prove that the bound holds for $\theta \neq \bar{\phi}$. For this, we look at equation (A.1) again. Let us denote the numerator of the second fraction by A and its denominator by B. In the interval $[\phi_i, \phi_{i+1}]$ we have

$$|A| \leq (n-1)\sin^2(\theta)|\theta - \bar{\phi}|. \tag{A.4}$$

To see this, note that $|\cos((n-1)\theta)|$ is concave in both intervals $[\phi_i, \bar{\phi}]$ and $[\bar{\phi}, \phi_{i+1}]$, and its left and right derivatives at $\theta = \bar{\phi}$ are $-(n-1)$ and $(n-1)$, respectively. This will give us $|\cos((n-1)\theta)| \leq (n-1)|\theta - \bar{\phi}|$, and multiplying by $\sin^2(\theta)$ gives us equation (A.4).

For the denominator we have,

$$\cos(\theta) - \cos(\bar{\phi}) = -2\sin\left(\frac{\theta - \bar{\phi}}{2}\right)\sin\left(\frac{\theta + \bar{\phi}}{2}\right),$$

which after taking the absolute value becomes

$$|\cos(\theta) - \cos(\bar{\phi})| = 2\sin\left(\frac{|\theta - \bar{\phi}|}{2}\right)\sin\left(\frac{\theta + \bar{\phi}}{2}\right). \tag{A.5}$$

Now, noting that $\frac{\theta+\bar{\phi}}{2} \in [0, \pi]$ and $\frac{|\theta-\bar{\phi}|}{2} \in [0, \frac{\pi}{2}]$, and using the fact that $\sin(\theta)$ is concave and non-negative in $[0, \pi]$, we can write

$$\sin\left(\frac{\theta + \bar{\phi}}{2}\right) \geq \frac{1}{2}[\sin(\theta) + \sin(\bar{\phi})] \geq \frac{1}{2}\sin(\theta), \tag{A.6}$$

and

$$\sin\left(\frac{|\theta - \bar{\phi}|}{2}\right) \geq \frac{\frac{|\theta-\bar{\phi}|}{2}}{\frac{\pi}{2}} = \frac{|\theta - \bar{\phi}|}{\pi}. \tag{A.7}$$

If we plug (A.6) and (A.7) into equation (A.5), we will have

$$|B| \geq \frac{1}{\pi}\sin(\theta)|\theta - \bar{\phi}|. \tag{A.8}$$

Using (A.4) and (A.8) in (A.1) and taking the absolute value will give us for $\theta \in [\phi_i, \phi_{i+1}]$ and $\theta \neq \phi$:

$$|w_\Phi(\theta)| \leq \frac{\pi(n-1)\sin(\bar{\phi})}{2^{n-2}} \leq \frac{\pi(n-1)}{2^{n-2}}, \tag{A.9}$$

which together with (A.3) proves the bound for the interval $[\phi_i, \phi_{i+1}]$ and therefore $[\theta_i, \theta_{i+1}]$. Since i was arbitrary, the bound holds for all subintervals, and therefore the whole interval $[0, \pi]$. □

Appendix B: Proof of Theorem 3

We break the proof into two lemmas.

Lemma 1 *Let $D = [-1, 1]$ and let $f : D \to \mathbb{R}$ have a convergent expansion $f = \sum_{j=0}^{\infty} a_j \hat{T}_j(x)$, where $\hat{T}_j(x)$ is the normalized Chebyshev polynomial of degree j.*

For a set of nodes $X_n = \{x_1, x_2, \cdots x_n\} \subset D$, let P be the associated interpolation matrix and p be the polynomial of degree at most $n-1$ that interpolates f at the nodes in X_n. We have

$$\|f - p\|_{L_2(D,w)} \leq \left(\sum_{k=n}^{\infty} a_k^2\right)^{\frac{1}{2}} + \|P^{-1}\|_2 \sqrt{2} \sum_{k=n}^{\infty} |a_k|. \tag{B.1}$$

Proof Let p^* be the Chebyshev projection of f onto the space of polynomials of degree at most $n-1$, $\mathbb{P}_{n-1}$. Using the triangular inequality, we can write

$$\|f - p\|_{L_2(D,w)} \leq \|f - p^*\|_{L_2(D,w)} + \|p^* - p\|_{L_2(D,w)}. \tag{B.2}$$

Since $p^* = \sum_{k=0}^{n-1} a_k \hat{T}_k(x)$, the first term in the right hand side of Equation (B.2) is simply $\left(\sum_{k=n}^{\infty} a_k^2\right)^{\frac{1}{2}}$.

For the second term, if we denote the coefficients in the Chebyshev expansion of the interpolant p by $\tilde{a}_j$, $0 \leq j \leq n-1$, we have $\|p^* - p\|_{L_2(D,w)} = \|a - \tilde{a}\|_2$, where $a = [a_0, a_1, \cdots, a_{n-1}]^T$ and $\tilde{a} = [\tilde{a}_0, \tilde{a}_1, \cdots, \tilde{a}_{n-1}]^T$. Therefore we only need to show that $\|a - \tilde{a}\|_2 \leq \|P^{-1}\|_2 \sqrt{2} \sum_{k=n}^{\infty} |a_k|$. Let us define $b = [p^*(x_1), p^*(x_2), \cdots, p^*(x_n)]^T$ and $\tilde{b} = [f(x_1), f(x_2), \cdots, f(x_n)]^T$. Since p interpolates f and p^* interpolates itself at X_n, we have

$$Pa = \frac{1}{\sqrt{n}} b, \qquad P\tilde{a} = \frac{1}{\sqrt{n}} \tilde{b}.$$

By subtracting these equations we can write

$$a - \tilde{a} = \frac{1}{\sqrt{n}} P^{-1}(b - \tilde{b}).$$

Therefore,

$$\begin{aligned}
\|a - \tilde{a}\|_2 &\leq \|P^{-1}\|_2 \frac{1}{\sqrt{n}} \|b - \tilde{b}\|_2 \\
&= \|P^{-1}\|_2 \left(\frac{1}{n} \sum_{i=1}^{n} \left(f(x_i) - p^*(x_i)\right)^2\right)^{\frac{1}{2}} \\
&= \|P^{-1}\|_2 \left(\frac{1}{n} \sum_{i=1}^{n} \left(\sum_{k=n}^{\infty} a_k \hat{T}_k(x_i)\right)^2\right)^{\frac{1}{2}} \\
&\leq \|P^{-1}\|_2 \max_{1 \leq i \leq n} \left|\sum_{k=n}^{\infty} a_k \hat{T}_k(x_i)\right| \\
&\leq \|P^{-1}\|_2 \sum_{k=n}^{\infty} |a_k| \max_{1 \leq i \leq n} \left|\hat{T}_k(x_i)\right| \\
&\leq \|P^{-1}\|_2 \sqrt{2} \sum_{k=n}^{\infty} |a_k|,
\end{aligned}$$

where in the last line we have used equation (3.3). □

Lemma 2 *Let* $\{a_k\}_{k=0}^{\infty}$ *be a sequence of real numbers. For a positive integer n, if there are* $C > 0$ *and* $s > 1$ *such that for any* $k \geq n$, $|a_k| \leq C(k+1)^{-s}$, *then,*

$$\left(\sum_{k=n}^{\infty} a_k^2\right)^{\frac{1}{2}} \leq \frac{C}{\sqrt{2s-1}} n^{-\left(s-\frac{1}{2}\right)} \tag{B.3}$$

and

$$\sum_{k=n}^{\infty} |a_k| \leq \frac{C}{s-1} n^{-(s-1)}. \tag{B.4}$$

If there are $C > 0$ *and* $\rho > 1$ *such that for any* $k \geq n$, $|a_k| \leq C\rho^{-n}$, *then,*

$$\left(\sum_{k=n}^{\infty} a_k^2\right)^{\frac{1}{2}} \leq \frac{C}{\sqrt{1-\rho^{-2}}} \rho^{-n} \tag{B.5}$$

and

$$\sum_{k=n}^{\infty} |a_k| \leq \frac{C}{1-\rho^{-1}} \rho^{-n}. \tag{B.6}$$

Proof Inequality (B.4) can be proven using the common technique of bounding by an improper integral,

$$\sum_{k=n}^{\infty} \frac{C}{(k+1)^s} < \int_n^{\infty} \frac{C}{x^s} dx.$$

Inequality (B.3) can be obtained in a similar fashion. Inequalities (B.5) and (B.6) can be shown using geometric series. □

Proof of Theorem 3 The proof follows directly from Lemmas 1 and 2. □

References

1. Baglama, J., Calvetti, D., Reichel, L.: Fast Leja points. Electron. Trans. Numer. Anal. **7**, 124–140 (1998)
2. Bernstein, S.: Sur la limitation des valeurs d'un polynôme. Bull. Acad. Sci. de l'URSS **8**, 1025–1050 (1931)
3. Boyd, J.P., Xu, F.: Divergence (Runge phenomenon) for least-squares polynomial approximation on an equispaced grid and Mock-Chebyshev subset interpolation. Appl. Math. Comput. **210**, 158–168 (2009)
4. Boyd, J.P., Gildersleeve, K.W.: Numerical experiments on the condition number of the interpolation matrices for radial basis functions. Appl. Numer. Math. **61**, 443–459 (2011)
5. Clenshaw, C.W., Curtis, A.R.: A method for numerical integration on an automatic computer. Numer. Math. **2**, 197–205 (1960)
6. Erdös, P.: Problems and results on the theory of interpolation. II. Acta Math. Acad. Sci. Hung. **12**, 235–244 (1961)
7. Gil, A., Segura, J., Temme, N.M.: Numerical Methods for Special Functions. SIAM. 978-0-898716-34-4 (2007)

8. Jung, J., Stefan, W.: A simple regularization of the polynomial interpolation for the resolution of the Runge phenomenon. J. Sci. Comput. **46**, 225–242 (2010)
9. Leja, F.: Sur certaines suites liées aux ensembles plans et leur application à la représentation conforme. Ann. Polon. Math. **3**, 8–13 (1957)
10. Narcowich, F.J., Ward, J.D.: Norms of inverses and condition numbers for matrices associated with scattered data. J. Approx. Theory **64**, 69–94 (1991)
11. Nobile, F., Tempone, R., Webster, C.G.: A sparse grid stochastic collocation method for partial differential equations with random input data. SIAM J. Numer. Anal. **46**, 2309–2345 (2008)
12. Salzer, H.E.: Lagrangian interpolation at the Chebyshev points $x_{n,} = \cos(\nu\pi/n)$, $\nu = 0(1)n$; some unnoted advantages. Comput. J. **15**, 156–159 (1972)
13. Taylor, R.: Lagrange interpolation on Leja points, Graduate School Theses and Dissertations. http://scholarcommons.usf.edu/etd/530 (2008)
14. Trefethen, L.N.: Is Gauss quadrature better than Clenshaw-Curtis? SIAM Rev. **50**, 67–87 (2008)
15. Trefethen, L.N. et al.: Chebfun Version 4.2. The Chebfun Development Team. http://www.chebfun.org (2011)
16. Trefethen, L.N.: Approximation Theory and Approximation Practice. SIAM, Philadelphia (2012)