A MAPPED POLYNOMIAL METHOD FOR HIGH-ACCURACY APPROXIMATIONS ON ARBITRARY GRIDS

BEN ADCOCK* AND RODRIGO B. PLATTE†

Abstract. The focus of this paper is the approximation of analytic functions on compact intervals from their pointwise values on arbitrary grids. We introduce a new method for this problem based on mapped polynomial approximation. By careful selection of the mapping parameter, we ensure both high accuracy of the approximation and an asymptotically optimal scaling of the polynomial degree with the grid spacing. As we explain, efficient implementation of this method can be achieved using Nonuniform Fast Fourier Transforms (NFFTs). Numerical results demonstrate the efficiency and accuracy of this approach.

Key words. Equispaced nodes, scattered data, spectral methods, Runge phenomenon, analytic functions

AMS subject classifications. 65D15, 65D05, 65T50

1. Introduction. Let $f : [-1, 1] \to \mathbb{C}$ be an analytic function and $-1 \leq z_0 < \ldots < z_M \leq 1$ a fixed grid of $M$ points. In this paper, we consider the problem of approximating of $f$ from the grid values $f(z_m), \ m = 0, \ldots , M$.

A classical mean of doing this is to interpolate $f$ using a polynomial of degree $M$. However, the famous Runge phenomenon illustrates the pitfalls of such an approach. In the case of equispaced grids, for example, the corresponding interpolants diverge exponentially fast as $M \to \infty$ for any function with complex singularities lying sufficiently close to the interval $[-1, 1]$. Moreover, the approximation is ill-conditioned, and so one sees divergence of the interpolants in finite precision, even for entire functions. Neither is this phenomenon isolated to equispaced data. It is well known that to avoid a Runge-type phenomenon the data should cluster at the endpoints according to a Chebyshev distribution. In other words, polynomial interpolants are generally inadvisable for all but very special grids.

One possible way to overcome the Runge phenomenon is to reduce the polynomial degree, to, say, $N < M$, and perform an overdetermined (weighted) least-squares fit of the data. In the case of equispaced grids, provided $N$ is chosen sufficiently small in comparison to $M$, this leads to a stable and convergent approximation [7]. However, due to a result of Coppersmith & Rivlin [13] (see also [32]), it is known that $N$ can grow no faster than $\sqrt{M}$ as $M \to \infty$ to maintain stability and convergence. Thus, the effective convergence rate of the approximation, determined by the size of $N$, is greatly lessened. Although the best approximation of an analytic function in $P_N$ is exponentially-accurate in $N$, this scaling translates into only root-exponential in the number of data points $M$. More generally, for arbitrary grids with maximal separation $h$, we can expect only root-exponential convergence in $1/h$ as $h \to 0$.

The severity of this scaling is due to the behaviour of derivatives of polynomials, and specifically, the fact that $\|p'|_\infty$ grows maximally like $N^2 \|p\|_\infty$ for a polynomial $p \in P_N$ (this is commonly known as Markov’s inequality – see [6, Theorem 5.1.8],

*Department of Mathematics, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada (ben_adcock@sfu.ca, http://www.benadcock.org).
†School of Mathematical and Statistical Sciences, Arizona State University, P.O. Box 871804, Tempe, AZ 85287-1804, USA (rbp@asu.edu, https://math.la.asu.edu/~platte/).
for example). On the other hand, trigonometric polynomials possess derivatives that grow at most linearly in $N$. Hence, a trigonometric polynomial least squares approximation will permit a linear scaling of $N$ with $M$ (or, more generally, $1/h$) whilst maintaining stability. Unfortunately, trigonometric polynomials are a poor means of approximating analytic functions. Unless $f$ happens to also be periodic, there is no uniform convergence as $N \to \infty$ and one witnesses the undesirable Gibbs phenomenon near the interval endpoints.

The purpose of this paper is to introduce a weighted least-squares method for approximating analytic functions from their values on equispaced or scattered points that combines the good features of both algebraic and trigonometric polynomial approximations. For appropriate parameter choices, the method we introduce has a linear scaling of $N$ with $M$ (or, in general, $1/h$), much as with trigonometric polynomial approximation, but delivers high accuracy reminiscent of polynomial approximation. Moreover, the method is practical, simple and can be implemented efficiently in $O(M \log M)$ operations using nonuniform Fast Fourier Transforms (NFFTs) [17, 18, 20, 31, 35].

1.1. Mapped polynomial approximations. Our approximation method is based on mapped algebraic polynomials. The corresponding approximation space

\begin{equation}
P_N^\alpha = \{ p \circ m_\alpha : p \in P_N \},
\end{equation}

consists of algebraic polynomials in a mapped variable $y = m_\alpha(x)$, where $m_\alpha : [-1,1] \to [-1,1]$ is a particular one-parameter family of mappings indexed by a parameter $0 \leq \alpha \leq 1$. When $\alpha = 0$, $P_N^0$ coincides with the space $P_N$ of algebraic polynomials of degree $N$, and when $\alpha = 1$ it consists of functions closely related to trigonometric polynomials. By selecting $\alpha$ sufficiently close to one, it is therefore expected that one can retain the good approximation properties of the $\alpha = 0$ case, whilst also improving its severe scaling of $N$ with $M$ (respectively $h$).

The approximation space (1.1) is not new. Mapped polynomial methods have been used extensively in the context of numerical quadrature and spectral methods for PDEs. Here the mapping $m_\alpha$ is used to overcome the severe time-step requirements of standard Chebyshev spectral methods or to improve the poor resolution properties of Chebychev grids [8, 24, 26]. The most widely-used such map is due to Kosloff and Tal–Ezer [26]. However, other mapped have also been considered, including most recently in the work of Hale & Trefethen [24]. Note that the situation considered in such applications is roughly speaking the reverse of ours. Therein the mapping is used to distribute a Chebyshev grid more evenly over the domain, and hence improve the time-step restriction. Conversely, in our setting we consider a fixed, but arbitrary, grid of data points, which we map to a grid that is closer to a Chebyshev distribution. The motivation for doing this is to suppress the maximal polynomial derivatives, and correspondingly improve the scaling of $N$ with $M$ required for stability. Thus, an interesting conclusion of this paper is that mappings are not just useful in applications such as spectral methods and numerical quadrature, they are also useful in the reconstruction problem of approximating analytic functions to high accuracy from arbitrary grids.

Note that it is not our aim in this paper to compare different mappings. We shall use the mapping due to Kosloff and Tal–Ezer [26] throughout due to its simplicity. We note, however, that other mappings may provide some advantages. For a discussion on this issue in relation to spectral methods and numerical quadrature, see [24].
In the context of spectral methods, the choice of the mapping parameter $\alpha$ has also been the subject of an extensive debate. See [1, 8, 14, 15, 16, 24, 33, 34] and references therein. There are two standard approaches for doing this. First, $0 < \alpha < 1$ is fixed and close to one, and second, $\alpha = \alpha_N \to 1^-$ as $N \to \infty$. As we will discuss later, approximations from the space $P_N^\alpha$ converge geometrically in the first case. In the second case, the standard approach is to introduce a finite maximal accuracy $\epsilon$ (typically on the order of machine precision [26]), and choose $\alpha_N$ so that the error of the approximation is on the order of $\epsilon$ for large $N$. Approximations from the space $P_N^{\alpha_N}$ no longer converge classically (i.e. down to zero in exact arithmetic as $N \to \infty$), but in practice, high accuracy is expected by taking $\epsilon$ on the order of machine precision. From the point of view of spectral methods, the advantage of the second approach is that it delivers asymptotically optimal time-step and resolution properties, on the order of those of Fourier spectral methods.

1.2. Our contributions. After introducing the method in §2 and discussing its efficient implementation using NFFTs, we devote the remainder of the paper to the key issue of how to choose the parameter $N$ (the size of the approximation space) in relation to $M$ (or, in general, $h$) for various different choices of $\alpha$ (the mapping parameter). We first show that for fixed $\alpha$ one cannot improve the asymptotic scaling of $N$ with $M$ beyond that of the $\alpha = 0$ case, i.e. $N = \mathcal{O}(\sqrt{M})$. The only possible improvement is in the constant. Conversely, if $\alpha = \alpha_N \to 1^-$ in an appropriate way we show that stability is guaranteed with a linear scaling of $N$ with $M$ and with an explicit constant (Theorem 4.2). Whilst classical convergence is forfeited, high accuracy is guaranteed by an appropriate choice of $\epsilon$.

The proofs of these results follow from the derivation of a Markov inequality for $P_N^\alpha$ which is uniform in both $N$ and $\alpha$ (Theorem 4.1). Notably, the constant in this inequality is explicit and not overly large for practical choices of parameters. This is an interesting virtue of our analysis.

A summary of our main results is given in Table 1. Note that the results are stated for equispaced data only; the primary example we use throughout this paper. However, they can be easily recast in terms of general scattered grids by replacing $M$ with $1/h$. Table 1 also includes some terminology for convergence that will be used throughout this paper. In particular, we will say that a sequence $a_n$ converges geometrically if $a_n = \mathcal{O}(\rho^{-n})$ for large $n$ for some $\rho > 1$. We say the convergence is subgeometric with index $0 < r < 1$ if $a_n = \mathcal{O}(\rho^{-nr})$. When $r = 1/2$ we also refer to this convergence as root-exponential.

Finally, we say the sequence converges algebraically with index $k \geq 1$ if $a_n = \mathcal{O}\left(n^{-k}\right)$ as $n \to \infty$.

1.3. Methods for function approximation from equispaced data. Many methods have been developed for the approximation of analytic functions from equispaced data. For an extensive list, see [10, 30] and references therein. A recent result [30] states that no method for this problem can be both stable and exponentially convergent. In fact, the best possible convergence rate of a stable method is root exponential in the number of equispaced points $M$. Such a convergence rate is achieved by polynomial least-squares, for example. However in practice (as we shall see later in our numerical results) polynomial least-squares tends to give poor results. On the other hand, when $\alpha = \alpha_N$ is varied appropriately (see the fifth line in Table 1), the mapped polynomial method we introduce in this paper achieves high accuracy and numerical stability. Yet the aforementioned theorem is not avoided, since classical convergence is sacrificed for finite accuracy.
Summary of our main results for equispaced data, where $h = 1/M$. In the fourth and fifth row, the notation $\sim$ denotes the behaviour of $\alpha = \alpha_N$ as $N \to \infty$. In the fourth row $0 < \sigma < 1$ and $\alpha_0 > 0$ are fixed numbers. In the fifth row, $\epsilon > 0$ is a fixed number, chosen sufficiently close to machine epsilon to give high accuracy. In this case, the geometric rate of convergence is limited by the mapping $m_\alpha$ (see Theorem 3.3). Although not guaranteed to be convergent, for large $N$ we expect the error to be proportional to $\epsilon$.

Table 1

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>Conv. rate in $N$</th>
<th>Scaling with $M$</th>
<th>Conv. rate in $M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>geometric</td>
<td>$O(M^{\frac{1}{2}})$</td>
<td>root exp.</td>
</tr>
<tr>
<td>$0 &lt; \alpha &lt; 1$ fixed</td>
<td>geometric$^1$</td>
<td>$O(M^{\frac{1}{2}})$</td>
<td>root exp.</td>
</tr>
<tr>
<td>1</td>
<td>algebraic, index 1</td>
<td>$O(M)$</td>
<td>algebraic, index 1</td>
</tr>
<tr>
<td>$\sim 1 - \frac{\alpha_0}{\sqrt{M}}$</td>
<td>subgeo., index $1 - \sigma$</td>
<td>$O(M^{\frac{1}{\sigma}})$</td>
<td>subgeo., index $\frac{1}{\sigma}$</td>
</tr>
<tr>
<td>$\sim 1 - \frac{2\log \epsilon}{N\pi}$</td>
<td>not applicable$^2$</td>
<td>$O(M)$</td>
<td>not convergent$^2$</td>
</tr>
</tbody>
</table>

As discussed in [10, 30], a number of other methods for equispaced function approximation also offer high accuracy and stability in practice. In §7 of this paper, we compare mapped polynomial methods with two well known schemes for approximation on arbitrary grids: discrete polynomial least-squares and cubic splines. A thorough comparison with other methods for the approximation of analytic functions from equispaced data, including rational approximation [21], Fourier extension [5, 4, 9, 11], and windowed Fourier [28], will be presented in [29].

2. The mapped polynomial method.

2.1. Preliminaries. Throughout this paper, we denote the space of functions which are square integrable with respect to a weight function $w(x)$ by $L^2_w(-1, 1)$. The corresponding inner product is written as $\langle \cdot, \cdot \rangle_w$ and the norm denoted by $\| \cdot \|_w$. The space $L^\infty(-1, 1)$ consists of those functions that are bounded a.e. on $[-1, 1]$, equipped with the norm $\| \cdot \|_\infty$.

The mapping used in this paper, due to Kosloff and Tal–Ezer [26], is:

$$m_\alpha(x) = \frac{\sin(\alpha \pi x/2)}{\sin(\alpha \pi/2)}, \quad x \in [-1, 1], \quad \alpha \in (0, 1].$$

For completeness, we define

$$m_0(x) = \lim_{\alpha \to 0^+} m_\alpha(x) = x.$$

Observe that $m_\alpha(x)$ is a bijection of $[-1, 1]$, and in particular,

$$m_\alpha(x) \leq m_\alpha(x') \iff x \leq x'.$$

Throughout, we write $y = m_\alpha(x) \in [-1, 1]$ for the variable in the mapped domain. Note that

$$x = m_\alpha^{-1}(y) = \frac{2}{\alpha \pi} \sin^{-1}(\sin(\alpha \pi/2)y),$$

and also that

$$\frac{dy}{dx} = m'_\alpha(x) = \frac{\alpha \pi \cos(\alpha \pi x/2)}{2 \sin(\alpha \pi/2)} = \frac{\alpha \pi}{2 \beta} \sqrt{1 - \beta^2 y^2},$$

with $\beta = \frac{\pi}{2\sqrt{\alpha}}$. 


where

\[ \beta = \sin(\alpha \pi/2). \]

We shall also write

\[ f(x) = g^\alpha(y), \quad g^\alpha = f \circ m^{-1}_\alpha, \]

for the image of \( f \) under the mapping \( m_\alpha \).

We next define the approximation space we use in this paper. Let

\[ P^\alpha_N = \{ p \circ m_\alpha : p \in P_N \}, \]

be the space of mapped polynomials in the variable \( x \) of degree at most \( N \). Observe that \( P^0_N = P_N \) is the space of algebraic polynomials of degree at most \( N \). For computational purposes, it is also necessary to have a basis for \( P^\alpha_N \). Let

\[ \phi_n(x) = c_n T_n(m_\alpha(x)), \quad n = 0, 1, 2, \ldots, \]

where \( T_n(y) \in P_n \) is the \( n \)th Chebyshev polynomial of the first kind in \( y \) and the normalization factor \( c_n \) is given by \( \sqrt{1/\pi} \) if \( n = 0 \) and \( \sqrt{2/\pi} \) otherwise.

**Lemma 2.1.** The functions \( \{ \phi_n \}_{n=0}^\infty \) form an orthonormal basis for the weighted space \( L^2_{w_\alpha}(-1,1) \) with weight

\[ w_\alpha(x) = \frac{\alpha \pi}{2} \frac{\cos(\alpha \pi x/2)}{\sqrt{\sin^2(\alpha \pi/2) - \sin^2(\alpha \pi x/2)}}. \]

**Proof.** The functions \( c_n T_n(y) \) are orthonormal with respect to the Chebyshev weight function \( 1/\sqrt{1-y^2} \). Using the substitution \( y = m_\alpha(x) \), we have

\[ \int_{-1}^1 \phi_n(x) \phi_m(x) w_\alpha(x) \, dx = c_n c_m \int_{-1}^1 T_n(y) T_m(y) \frac{2\beta w_\alpha(m_\alpha^{-1}(y))}{\alpha \pi \sqrt{1 - \beta^2 y^2}} \, dy, \]

where \( \beta \) is as in (2.2). Thus, for orthonormality, we require that

\[ \frac{2\beta w_\alpha(m_\alpha^{-1}(y))}{\alpha \pi \sqrt{1 - \beta^2 y^2}} = \frac{1}{\sqrt{1-y^2}}. \]

Substituting \( y = m_\alpha(x) \) now gives the result. \( \Box \)

Note that one could in theory use any orthonormal polynomial basis for \( P_N \) in order to construct a basis of \( P^\alpha_N \). We use Chebyshev polynomials for their computational efficiency (see §2.3). But it is in theory possible to use any other polynomial basis, for example Gegenbauer polynomials, which have been used widely for other problems in scientific computing (see [19, 22, 23] and references therein).

**2.2. The method.** Having introduced the approximation space \( P^\alpha_N \), we next formulate the mapped polynomial method. To this end, let

\[ -1 \leq z_0 < z_1 < \ldots < z_M \leq 1, \]

be an ordered set of \( M + 1 \) data points, where \( M \geq N \), and write \( Z = \{ z_m \}_{m=0}^M \). Define the maximal spacing \( h > 0 \) by

\[ h = \max_{n=-1,\ldots,M} \{ z_{n+1} - z_n \}, \]

where

\[ \beta = \sin(\alpha \pi/2). \]
where \( z_{-1} = -1 \) and \( z_{M+1} = 1 \). We will use equispaced grids as our primary example throughout this paper. In this case, we set

\[
z_m = -1 + \frac{2m}{M}, \quad m = 0, \ldots, M,
\]

and therefore \( h = 2/M \).

Given the data \( \{f(z_m)\}_{m=1}^M \) and the approximation space \( P_N^\alpha \), we construct the approximation to \( f \) by a weighted least-squares fitting:

\[
F_N^\alpha Z(f) = \arg\min_{p^\alpha \in P_N^\alpha} \sum_{m=0}^M \mu_m |f(z_m) - p^\alpha(z_m)|^2.
\]

Here the weights \( \mu_n > 0 \) are trapezoidal quadrature weights corresponding to \( Z \):

\[
\mu_n = \frac{1}{2} \int_{m_\alpha(z_{n-1})}^{m_\alpha(z_{n+1})} \frac{1}{\sqrt{1 - y^2}} \, dy = \frac{1}{2} \left( \sin^{-1}(m_\alpha(z_{n+1})) - \sin^{-1}(m_\alpha(z_{n-1})) \right), \quad n = 0, \ldots, M.
\]

We make this choice over simpler strategies because it avoids conditioning issues if \( z_0 \) or \( z_M \) are close to their respective endpoints. Note that, given \( Z \) and the weights \( \mu_n \), the parameters \( \alpha \) and \( N \) are both chosen by the user. This is the key issue we consider in §3 and §4.

The approximation \( F_N^\alpha Z(f) \) is defined in the physical \( x \)-domain. Let

\[
F_N^\alpha Z(f)(x) = G_N^\alpha Z(g^\alpha)(y),
\]

be its image in the \( y \)-domain, where \( g^\alpha \) is given by (2.3). Note that \( G_N^\alpha Z(g^\alpha) \) can also be defined as a polynomial least-squares fit:

\[
G_N^\alpha Z(g^\alpha) = \arg\min_{p \in P_N} \sum_{m=0}^M \mu_m |g^\alpha(m_\alpha(z_m)) - p(m_\alpha(z_m))|^2.
\]

At this stage it is convenient to introduce the discrete inner product

\[
\langle f, g \rangle_Z = \sum_{m=0}^M \mu_m f(z_m)g(z_m), \quad f, g \in L^\infty(-1, 1).
\]

and write \( \| \cdot \|_Z \) for the corresponding discrete semi-norm. We also define the discrete uniform semi-norm:

\[
\|f\|_{Z, \infty} = \max_{m=0, \ldots, M} |f(z_m)|, \quad f \in L^\infty(-1, 1).
\]

Finally, since \( F_N^\alpha Z(f) \) is defined as a least-square fit of the data, it is the solution of the corresponding normal equations. Written in variational form, one sees that \( F_N^\alpha Z(f) \) is the solution to the problem

\[
\text{find } \hat{f} \in P_N^\alpha \text{ such that } \langle \hat{f}, p^\alpha \rangle_Z = \langle f, p^\alpha \rangle_Z, \forall p^\alpha \in P_N^\alpha.
\]

We shall use this observation later when analyzing the method.
2.3. Computation of the approximation. Let $\phi_n$ be the basis from (2.5). If

$$F_{N,Z}^\alpha(f) = \sum_{n=0}^{N} a_n \phi_n,$$

for unknown coefficients $a_n \in \mathbb{C}$ then the weighted least-squares (2.8) is equivalent to the algebraic least squares (2.11)

$$Aa \approx b,$$

where $A \in \mathbb{C}^{M \times N}$ has $(m,n)$th entry $\sqrt{\mu_m} \phi_n(z_m)$, $a = (a_0, \ldots, a_N)^\top$, $b = (b_0, \ldots, b_M)^\top$ and $b_m = \sqrt{\mu_m} f(z_m)$. Computation of the approximation $F_{N,Z}^\alpha(f)$ can be carried out using a standard solver such as conjugate gradients with the number of iterations being proportional to the square root of the condition number of $A$, $\kappa(A)$.

Let $\sigma_{\text{max}}$ and $\sigma_{\text{min}}$ be the maximal and minimal singular values of $A$ respectively, so that $\kappa(A) = \sigma_{\text{max}} / \sigma_{\text{min}}$. We note the following:

**Lemma 2.2.** Let $y_n = m_\alpha(z_n)$ for $n = 0, \ldots, M$. Then

$$(\sigma_{\text{max}})^2 = \max \left\{ \sum_{n=0}^{M} \mu_n |p(y_n)|^2 : p \in P_N, \|p\|_w = 1 \right\},$$

$$(\sigma_{\text{min}})^2 = \min \left\{ \sum_{n=0}^{M} \mu_n |p(y_n)|^2 : p \in P_N, \|p\|_w = 1 \right\},$$

where $w(y) = 1/\sqrt{1 - y^2}$ is the Chebyshev weight.

**Proof.** Let $a = (a_0, \ldots, a_N)^\top$, $f = \sum_{n=0}^{N} a_n \phi_n \in P_N^\alpha$ and $p = g^{\alpha} \in P_N$ be as in (2.3). Note that $p$ is a sum of orthonormal Chebyshev polynomials in $y$, and therefore

$$\sum_{n=0}^{N} |a_n|^2 = \|p\|^2_w.$$

On the other hand,

$$\|Aa\|^2 = \sum_{n=0}^{M} \mu_n |f(m_\alpha(z_n))|^2 = \sum_{n=0}^{M} \mu_n |p(y_n)|^2.$$

The result now follows immediately. [Q.E.D.]

In §5 we will use this lemma to analyze the conditioning of $A$ and show that $\kappa(A)$, in practice, remains bounded for appropriate choices of the parameters $\alpha$ and $N$. Using this, in §6 we will describe the fast implementation of the approximation in $O(M \log M)$ time using NFFTs.

2.4. Parameter choices. We now define the various parameter choices we consider in this paper:

(i) $\alpha = 0$,

(ii) $0 < \alpha < 1$ fixed,

(iii) $\alpha = 1$,

(iv) $\alpha = \alpha_N \sim 1 - \alpha_0/N^\sigma$ as $N \to \infty$, where $\alpha_0 > 0$ and $0 < \sigma < 1$,

(v) $\alpha = \alpha_N = \frac{4}{\pi} \arctan(\epsilon^{1/N})$, where $\epsilon > 0$ is small.
Note that in case (i), $P_N^\alpha = \mathbb{P}_N$ and therefore $F_{N,Z}^\alpha$ is just an algebraic polynomial least-squares fit. In case (iii) we shall see in §3.2.2 that $P_N^1$ is similar to the space of trigonometric polynomials, and has correspondingly poor approximation properties. Choices (iv) and (v) involve varying $\alpha$ with $N$. The particular choice of $\alpha_N$ in (v), originally due to Kosloff and Tal–Ezer [26], will be explained in §3.2.4.

3. Analysis of the mapped polynomial method. Having introduced the method, we will now analyze its stability and convergence properties and how they depend on the mapping parameter and the oversampling rate.

3.1. Stability and convergence. We first define the condition number of the approximation. Since $F_{N,Z}^\alpha$ is linear, its $L^\infty$ condition number is given by

$$
\kappa = \kappa_{N,Z}^\alpha = \sup_{\|f\|_{-1,1} \neq 0} \left\{ \frac{\|F_{N,Z}^\alpha(f)\|_\infty}{\|f\|_{-1,1}} \right\}.
$$

It transpires that $\kappa$ is a little difficult to analyze in practice. Thus we work with the smaller quantity

$$
\tilde{\kappa} = \tilde{\kappa}_{N,Z}^\alpha = \sup_{p^\alpha \in P_N^\alpha \atop \|p^\alpha\|_{Z,\infty} \neq 0} \left\{ \frac{\|p^\alpha\|_\infty}{\|p^\alpha\|_{Z,\infty}} \right\} = \sup_{p^\alpha \in P_N^\alpha \atop \|p^\alpha\|_{Z,\infty} \neq 0} \left\{ \frac{\|p^\alpha\|_\infty}{\|p^\alpha\|_{Z,\infty}} \right\}.
$$

Note that the second equality follows from $|Z| = M + 1 \geq N + 1$. Indeed, since $P_N^\alpha$ consists of mapped polynomials, $p^\alpha = 0$ if and only if $\|p^\alpha\|_{Z,\infty} = 0$. The following lemma relates $\tilde{\kappa}$ to the condition number $\kappa$:

**Lemma 3.1.** We have $\tilde{\kappa} \leq \kappa \leq \sigma \tilde{\kappa}$, where

$$
\sigma = \sqrt{\pi} \sqrt{\min_{n=0,\ldots,M} \{\mu_n\}}.
$$

**Proof.** Since $M \geq N$ by assumption, and since the points $z_0, \ldots, z_M$ are distinct, the matrix $A$ has full column rank. Hence $F_{N,Z}^\alpha(f)$ exists uniquely for any $f \in L^\infty(-1,1)$. The operator $D_{N,Z}^\alpha$ is also a projection onto $P_N^\alpha$. Therefore

$$
\kappa = \sup_{f \in L^\infty(-1,1) \atop \|f\|_{Z,\infty} \neq 0} \left\{ \frac{\|F_{N,Z}^\alpha(f)\|_\infty}{\|f\|_{Z,\infty}} \right\}
$$

$$
\geq \sup_{p^\alpha \in P_N^\alpha \atop \|p^\alpha\|_{Z,\infty} \neq 0} \left\{ \frac{\|F_{N,Z}^\alpha(p^\alpha)\|_\infty}{\|p^\alpha\|_{Z,\infty}} \right\}
$$

$$
= \sup_{p^\alpha \in P_N^\alpha \atop \|p^\alpha\|_{Z,\infty} \neq 0} \left\{ \frac{\|p^\alpha\|_\infty}{\|p^\alpha\|_{Z,\infty}} \right\} = \tilde{\kappa},
$$

which gives the lower bound. For the upper bound, we first note that

$$
\|F_{N,Z}^\alpha(f)\|_\infty \leq \tilde{\kappa} \|F_{N,Z}^\alpha(f)\|_{Z,\infty}.
$$

We now use the variational form (2.10). Setting $p^\alpha = \tilde{f} = F_{N,Z}^\alpha(f)$ and using the Cauchy–Schwarz inequality for the discrete inner product $(\cdot, \cdot)_Z$, we find that

$$
\|F_{N,Z}^\alpha(f)\|_Z \leq \|f\|_Z.
$$
A MAPPED POLYNOMIAL METHOD

Fig. 1. Numerically estimated condition numbers $\kappa$ for equispaced nodes with $M = 2N$ and two values of the mapping parameter, $\alpha = 0.5$ and 0.9. The solid lines present the values of $\kappa$, while the dashed lines are the bounds in Lemma 3.1.

Thus

$$\|f\|_{Z,\infty} \geq \frac{\|f\|_{Z}}{\sqrt{\sum_{n=0}^{M} \mu_n}} \geq \frac{\|F_{N,Z}^{\alpha}(f)\|_{Z}}{\sqrt{\sum_{n=0}^{M} \mu_n}} \geq \frac{\sqrt{\min_{n=0,\ldots,M}\{\mu_n\}}}{\sqrt{\sum_{n=0}^{M} \mu_n}} \|F_{N,Z}^{\alpha}(f)\|_{Z,\infty}.$$  

Observe that

$$\sum_{n=0}^{M} \mu_n = \frac{1}{2} (\pi + \sin^{-1}(m_\alpha(z_M)) - \sin^{-1}(m_\alpha(z_0))) \leq \int_{-1}^{1} \frac{1}{\sqrt{1 - y^2}} dy = \pi,$$

where the equality holds when $z_0 = -1$ and $z_M = 1$. Hence this gives

$$\|F_{N,Z}^{\alpha}(f)\|_{\infty} \leq \tilde{\kappa} \sigma \|f\|_{Z,\infty}.$$  

The first result now follows immediately. $\square$

Note that $\kappa$, and therefore $\tilde{\kappa}$, determines the condition number of the approximation $F_{N,Z}^{\alpha}$. Figure 1 shows how tight the bounds in Lemma 3.1 are for equispaced nodes. In this figure, $M = 2N$ was used for $\alpha = 0.5$ and $\alpha = 0.9$. Notice in particular that $\kappa$ is very close to $\tilde{\kappa}$.

We next consider the error of the approximation:

**Theorem 3.2.** We have

$$\|f - F_{N,Z}^{\alpha}(f)\|_{\infty} \leq (1 + \sigma \tilde{\kappa}) E_{N}^{\alpha}(f), \quad E_{N}^{\alpha}(f) = \inf_{p^\alpha \in P_{N}^{\alpha}} \|f - p^\alpha\|_{\infty}.$$  

**Proof.** For any $p^\alpha \in P_{N}^{\alpha}$,

$$\|f - F_{N,Z}^{\alpha}(f)\|_{\infty} \leq \|f - p^\alpha\|_{\infty} + \|F_{N,Z}^{\alpha}(f - p^\alpha)\|_{\infty} \leq \|f - p^\alpha\|_{\infty} + \kappa \|f - p^\alpha\|_{\infty}.$$  

We now use Lemma 3.1. $\square$

This result shows that the error of the mapped polynomial approximation decouples into a term $\tilde{\kappa}$ depending on the data and a term $E_{N}^{\alpha}(f)$ that is independent of the data and depends only on the parameters $\alpha$ and $N$. Clearly, our interest lies in
choosing $\alpha$ and $N$ such that $E_N^\alpha(f)$ is as small as possible. But this must be balanced with the fact that the best choices for minimizing $E_N^\alpha(f)$ may lead to a large condition number $\bar{\kappa}$. We dedicate §4 to the issue of balancing these parameters based on an estimate for $\bar{\kappa}$ which we derive. In order to do so, however, it is first necessary to consider the behaviour of the best approximation error $E_N^\alpha(f)$.

### 3.2. Behaviour of the best approximation error.

We now consider the decay rate of $E_N^\alpha(f)$ with respect to $N$ for the five choices of $\alpha$ introduced in §2.4.

#### 3.2.1. The case of fixed $0 \leq \alpha < 1$.

We shall focus on analytic functions. Let

$$B(\rho) = \left\{ \frac{1}{2} (\rho^{-1} e^{i\theta} + \rho e^{-i\theta}) : \theta \in [-\pi, \pi) \right\} \subseteq \mathbb{C},$$

be the usual Bernstein ellipse in the complex $y$-plane with index $\rho \geq 1$ and write $D_\alpha(\rho) \subseteq \mathbb{C}$ for the image of $B(\rho)$ in the complex $x$-plane under the inverse mapping $x = m_\alpha^{-1}(y)$. Then we have the following:

**Theorem 3.3.** Let $N \in \mathbb{N}$ and $0 \leq \alpha < 1$ be given and suppose that $f$ is analytic in $D_\alpha(\rho')$ for some $\rho' > 1$ and continuous on its boundary. Then

$$E_N^\alpha(f) \leq \frac{2e^\alpha(f)}{\rho - 1} \rho^{-N}, \quad \varepsilon^\alpha(f) = \max_{z \in D_\alpha(\rho)} |f(z)|,$$

where

$$\rho = \min \left\{ \cot \left( \frac{\alpha \pi}{4} \right), \rho' \right\}, \quad 0 < \alpha < 1, \quad \rho = \rho', \quad \alpha = 0.$$

**Proof.** Note that

$$E_N^\alpha(f) = \inf_{p^\alpha \in P_N^\alpha} \|f - p^\alpha\|_\infty = \inf_{p \in \mathbb{P}_N} \|g^\alpha - p\|_\infty,$$

where $g^\alpha = f \circ m_\alpha^{-1}$. The mapping $y = m_\alpha^{-1}(x)$ has singularities at $y = \pm 1/\sin(\alpha \pi/2)$. Note that, if $\rho \geq 1$ satisfies

$$\frac{1}{2} (\rho + \rho^{-1}) = \frac{1}{\sin(\alpha \pi/2)},$$

then $\rho = \cot(\alpha \pi/4)$. Since $f$ is analytic in $D_\alpha(\rho')$, the function $g^\alpha$ is therefore analytic in the Bernstein ellipse $B(\rho)$. Thus, the standard Bernstein estimate for best uniform approximation by polynomial (see [37, Chpt. 8] for example) now gives the result. $\Box$

Note that when $\alpha = 0$, i.e. when $P_N^\alpha = \mathbb{P}_N$, then $\rho = \rho'$ and we recover the usual result for polynomial approximation. For all other values of $\alpha$, the rate of geometric convergence of $E_N^\alpha(f)$ is limited to at most $\cot(\alpha \pi/4)$ by the singularity introduced by the inverse mapping $m_\alpha^{-1}$. Nevertheless, for any fixed $0 \leq \alpha < 1$ the convergence rate remains geometric, in contrast to the cases described next.

#### 3.2.2. The case $\alpha = 1$.

We first require the following result:

**Lemma 3.4.** For even $N$, the space $P_N^\alpha$ defined by (2.4) has equivalent expression

$$P_N^\alpha = \left\{ \sum_{n=0}^{N/2} a_n \cos(n \alpha \pi x) + \sum_{n=1}^{N/2} b_n \sin(n \alpha(n - 1/2)\pi x) : a_n, b_n \in \mathbb{C} \right\}.$$
The set \( \{ \cos(n\pi x) \}_{n=0}^{\infty} \cup \{ \sin(\alpha(n-1/2)\pi x) \}_{n=1}^{\infty} \) is precisely the orthogonal basis of eigenfunctions of the Laplace operator on \([-1/\alpha, 1/\alpha]\) subject to homogeneous Neumann boundary conditions.

Proof. By [36, Lem. 1], one has

\[
\cos(n\pi x) = (-1)^n T_{2n}(\sin(\frac{\alpha \pi}{2} x)) = (-1)^n T_{2n} \left( \sin(\frac{\alpha \pi}{2} m_{\alpha}(x) \right),
\]

and also

\[
\sin(\alpha(n-1/2)\pi x) = (-1)^{n-1} T_{2n-1}(\sin(\frac{\alpha \pi}{2} x)) = (-1)^{n-1} T_{2n-1} \left( \sin(\frac{\alpha \pi}{2} m_{\alpha}(x) \right).
\]

The functions \( \{ T_{2n} \left( \sin(\frac{\alpha \pi}{2} m_{\alpha}(x) \right) \}_{n=0}^{N} \cup \{ T_{2n-1} \left( \sin(\frac{\alpha \pi}{2} m_{\alpha}(x) \right) \}_{n=1}^{N} \) form a basis for \( P^N_{\alpha} \). Hence the result follows.

This lemma shows that \( P^N_{\alpha} \) is closely related to the space of trigonometric polynomials (the case of odd \( N \) is similar, and hence is omitted). Indeed, if the factor of \((n-1/2)\) were replaced by \(n\), then the space \( P^N_{\alpha} \) would be precisely the space of trigonometric polynomials on \([-1,1]\). For a comparison of trigonometric polynomial approximation and approximation with Laplace–Neumann eigenfunctions we refer to [2, 3]. One difference is that Laplace–Neumann approximations converge uniformly, whereas trigonometric polynomials do not. Specifically, in [2] it was shown that \( E_N(f) \to 0 \) as \( N \to \infty \) whenever \( f \in H^1(-1,1) \), where \( H^1(-1,1) \) denotes the standard Sobolev space of order 1. However, the convergence rate is limited to \( O(N^{-1}) \) unless \( f \) obeys specific endpoint conditions, analogous to periodicity conditions in trigonometric polynomial approximation. Hence, in general, the choice \( \alpha = 1 \) results in lower orders of approximation (specifically, algebraic with index one), even for analytic \( f \).

3.2.3. The case \( \alpha \sim 1 - \alpha_0/N^\sigma \) as \( N \to \infty \). Suppose now that

\[
\alpha = \alpha_N \sim 1 - \alpha_0/N^\sigma, \quad N \to \infty,
\]

where \( \alpha_0 > 0 \) and \( 0 < \sigma < 1 \) are fixed. To estimate the convergence rate in this case, we use Theorem 3.3. Observe that

\[
\cot(\alpha_N \pi /4) \sim 1 + \alpha_0 \pi/(2N^\sigma), \quad N \to \infty,
\]

and therefore

\[
\rho^{-N} \sim (1 + \alpha_0 \pi/(2N^\sigma))^{-N} \sim (\exp(\alpha_0 \pi/2))^{-N^{1-\sigma}}, \quad N \to \infty.
\]

Hence for \( 0 < \sigma < 1 \) we have subgeometric convergence with index \( 1 - \sigma \). Note that when \( \sigma = 1 \), there is no decay.

3.2.4. The case \( \alpha = 4/\pi \arctan(e^{1/N}) \). This final choice of \( \alpha \) is motivated by Theorem 3.3. Let \( \epsilon > 0 \) be a fixed, user-controlled tolerance. Since the error is proportional to \( (\cot(\alpha\pi/4))^{-N} \), the choice

\[
\alpha = \alpha_N = \frac{4}{\pi} \arctan \left( e^{1/N} \right),
\]

gives

\[
(\cot(\alpha_N \pi/4))^{-N} = \epsilon.
\]
Hence, provided $\epsilon$ is chosen sufficiently small (e.g. on the order of machine precision), we expect a small approximation error, even though rapid classical convergence of $E_N^\alpha(f)$ down to zero is no longer guaranteed. Observe that

$$\alpha_N = 1 - \frac{2|\log \epsilon|}{N\pi} + O(N^{-2}), \quad N \to \infty,$$

in this case. Hence, in the previous notation we have $\sigma = 1$ and $\alpha_0 = 2|\log \epsilon|/\pi$.

4. Parameter choices. As discussed in the previous section, given a set of data $Z$ with maximal spacing $h$, one must choose $N$ and $\alpha$ in such a way to keep the condition number $\tilde{\kappa}$ small, whilst at the same time minimizing the approximation error $E_N^\alpha(f)$. In this section we address this issue.

4.1. A Markov inequality for $P_N^\alpha$. We first require the following:

**Theorem 4.1.** We have

$$\|p^\alpha\|_\infty \leq C^\alpha_N \|p\|_\infty, \quad \forall p \in P_N^\alpha,$$

where

$$C^\alpha_N \leq \frac{\alpha\pi}{2} N \sqrt{1 + \cot^2(\alpha\pi/2)N^2}.$$

Recall that the classical Markov inequality for algebraic polynomials states that

$$\|p'\|_\infty \leq N^2 \|p\|_\infty, \quad \forall p \in \mathbb{P}_N,$$

where the constant $N^2$ is sharp; see [6, Theorem 5.1.8], for example. Conversely, for trigonometric polynomials, Bernstein’s inequality [6, Theorem 5.1.4] gives

$$\|p'\|_\infty \leq \frac{N\pi}{2} \|p\|_\infty, \quad \forall p \in \mathbb{T}_N,$$

where $\mathbb{T}_N = \left\{ \sum_{n=-N/2}^{N/2} a_n e^{in\pi x} : a_n \in \mathbb{C} \right\}$ is the space of trigonometric polynomials. Theorem 4.1 gives a Markov inequality for the spaces $P_N^\alpha$, $0 < \alpha < 1$. Note that the bound (4.1) reduces to (4.3) when $\alpha = 0$. Similarly, it reduces to (4.4) when $\alpha = 1$, except, of course, that $P_N^1$ is not the space $\mathbb{T}_N$ of trigonometric polynomials but the space of Laplace–Neumann eigenfunction on $[-1,1]$ (see Lemma 3.4). Nevertheless, one can show that $P_N^1$ satisfies the same Bernstein inequality (4.4) as $\mathbb{T}_N$ [3]. In other words, the general Markov inequality (4.1) is sharp in the extreme cases $\alpha = 0$ and $\alpha = 1$.

**Proof.** Recall that $p^\alpha(x) = p \circ m_\alpha(x) = p(y)$, where $p \in \mathbb{P}_N$. We have $\|p^\alpha\|_\infty = \|p\|_\infty$ and

$$(p^\alpha(x))' = p'(y)m_\alpha'(x) = \frac{\alpha\pi}{2\beta} p'(y) \sqrt{1 - \beta^2 y^2},$$

where $\beta$ is as in (2.2). Thus

$$\| (p^\alpha)' \|_\infty = \frac{\alpha\pi}{2\beta} \max_{-1 \leq y \leq 1} |p'(y)\sqrt{1 - \beta^2 y^2}|.$$

We now recall the following inequality for algebraic polynomials, due to Bernstein (see, for example, [6, Theorem 5.1.7]):

$$|p'(y)\sqrt{1 - y^2}| \leq N\|p\|_\infty, \quad p \in \mathbb{P}_N, \quad -1 \leq y \leq 1.$$
Now consider $|p'(y)\sqrt{1-\beta^2y^2}|$. Let $0 < \tau < 1$ and suppose that $|y| \leq \sqrt{1-\tau}$. Then
\[
|p'(y)\sqrt{1-\beta^2y^2}| = \sqrt{\frac{1-\beta^2y^2}{1-y^2}}|p'(y)\sqrt{1-y^2}| = \sqrt{\beta^2 + \frac{1-\beta^2}{1-y^2}|p'(y)\sqrt{1-y^2}|}.
\]
Thus, by (4.6) and the assumption on $y$,
\[
|p'(y)\sqrt{1-\beta^2y^2}| \leq \sqrt{\beta^2 + \frac{1-\beta^2}{\tau}N\|p\|_\infty}, \quad |y| < \sqrt{1-\tau}.
\]
Suppose that $\sqrt{1-\tau} \leq |y| \leq 1$. Then, by Markov’s inequality (4.3),
\[
|p'(y)\sqrt{1-\beta^2y^2}| \leq \sqrt{1-\beta^2(1-\tau)N^2\|p\|_\infty}, \quad \sqrt{1-\tau} \leq |y| \leq 1.
\]
Combining this with the previous estimate, we obtain
\[
|p'(y)\sqrt{1-\beta^2y^2}| \leq \max \left\{ \sqrt{\beta^2 + \frac{1-\beta^2}{\tau}N, \sqrt{\beta^2+1}} \right\} ||p||_\infty.
\]
We now set $\tau = 1/N^2$, substitute into (4.5) and use the definition of $\beta$ to deduce the result. \(\square\)

The Markov inequality (4.1) gives the following estimate for the condition number:

**Theorem 4.2.** The condition number
\[
\tilde{\kappa}_{N,M}^\alpha \leq \frac{1}{1-hC_N^\alpha/2},
\]
where $C_N^\alpha$ is as in (4.2). In particular, suppose that $c \geq 1$ is fixed. Then
\[
\tilde{\kappa}_{N,Z}^\alpha \leq c,
\]
whenever $N$ and $\alpha$ satisfy
\[
N\sqrt{1 + \cot^2(\alpha\pi/2)}N^2 \leq \frac{4(1-1/c)}{\alpha\pi h}.
\]

**Proof.** Let $x \in [-1,1]$. Then there exists an $m = -1, \ldots, M+1$ such that $|x-z_m| \leq h/2$. By the mean value theorem
\[
|p^\alpha(x)| \leq |p^\alpha(x_m)| + |x-z_m|\|p^\alpha\|_\infty \leq \|p^\alpha\|_{Z,\infty} + hC_N^\alpha/2\|p^\alpha\|_\infty.
\]
Taking the supremum over $x \in [-1,1]$ and rearranging now gives
\[
\|p^\alpha\|_\infty \leq 1/(1-hC_N^\alpha/2)\|p^\alpha\|_{Z,\infty},
\]
Since this holds for all $p^\alpha \in P_N^\alpha$ we obtain the first result. For (4.7), we note that $\tilde{\kappa}_{N,M}^\alpha \leq c$ provided
\[
C_N^\alpha \leq \frac{2(1-1/c)}{h}.
\]
Substituting the expression (4.2) for $C_N^\alpha$ now gives (4.7). \(\square\)
4.2. The choice of \( N \) and \( \alpha \). With Theorem 4.2 in hand, we may now consider how to select the parameter \( N \) for the choices of \( \alpha \) listed in §2.4.

4.2.1. The case \( \alpha = 0 \). Recall that \( P_N^\alpha \) is the space \( \mathbb{P}_N \) of polynomials of degree \( N \). Hence the sufficient condition (4.7) for a bounded condition number reduces to

\[
N \leq 2\sqrt{\frac{(1 - 1/c)}{\pi h}},
\]

i.e. \( N = \mathcal{O}(1/\sqrt{h}) \) as \( h \to 0 \). Since \( E_0^N(f) \) decays geometrically fast in \( N \), this translates into root-exponential convergence in \( 1/h \) as \( h \to 0 \). In other words, although algebraic polynomial approximations have good intrinsic approximation properties, they result in severe scalings of \( N \) with \( h \), leading to a less than desirable effective convergence rate in terms of \( h \).

4.2.2. The case \( \alpha = 1 \). At the other extreme, when \( \alpha = 1 \) condition (4.7) reads

\[
N \leq 4(1 - 1/c)\pi^2\alpha h.
\]

Hence a bounded condition number is ensured with a linear scaling \( N = \mathcal{O}(1/h) \). However, as discussed in §3.2, the best approximation error \( E_1^N(f) \) decays only very slowly in this case. Whilst setting \( \alpha = 1 \) overcomes the unpleasant scaling of the \( \alpha = 0 \) case, it destroys the beneficial approximation properties.

4.2.3. The case of fixed \( 0 < \alpha < 1 \). In this case, (4.7) results in the sufficient condition

\[
N = \mathcal{O}(1/\sqrt{h}) \text{ as } h \to 0,
\]

where \( 0 < \sigma < 1 \) and \( \alpha_0 > 0 \). Then for large \( N \) the left-hand side of (4.7) reads

\[
N\sqrt{1 + \cot^2(\alpha_N \pi/2)N^2} \sim \frac{\alpha_0\pi}{2}N^{2-\sigma},
\]

and therefore (4.7) results in the condition

\[
N \leq \left( \frac{8(1 - 1/c)}{\pi^2\alpha_0} \right)^{\frac{1}{2-\sigma}} h^{-\frac{1}{2-\sigma}} + o(1), \quad N \to \infty.
\]

In other words, we require \( N = \mathcal{O}(h^{-1/(2-\sigma)}) \) as \( h \to 0 \). Thus, by taking \( \sigma \) close to 1, we reduce the scaling to almost linear in \( 1/h \). But recall that the decay rate of \( E_\alpha^N(f) \) in this case is subgeometric with index \( 1 - \sigma \). This means that

\[
\|f - F_{N,Z}^\alpha(f)\|_\infty = \mathcal{O}\left( e^{-(1/h)^{1-\frac{1}{2-\sigma}}} \right), \quad h \to 0,
\]
for some $c > 1$ whenever $N = \mathcal{O}\left( h^{-\frac{1}{2\alpha}} \right)$ and $\alpha_N$ is as in (3.2) with $0 < \sigma < 1$. In other words, the effective convergence rate is subgeometric in $1/h$ with index $1 - \frac{1}{2 - \sigma}$, and this drops to zero as $\sigma$ approaches one.

4.2.5. The case $\alpha = 4/\pi \arctan(\epsilon^{1/N})$. Suppose now that $\alpha_N$ is given by (3.3) for some $\epsilon > 0$. Due to (3.4), we find that (4.7) gives

$$N \leq \frac{8(1 - 1/c)}{\pi \sqrt{1 + |\log \epsilon|^2/4}} h^{-1} + o(1), \quad N \to \infty.$$ 

Hence a linear scaling $N = \mathcal{O}(1/h)$ suffices in this case, much as in the case of $\alpha = 1$. However, unlike that case we expect high accuracy from this approach, provided $\epsilon$ is sufficiently small. Note that this does not contradict the aforementioned impossibility theorem of root-exponential convergence, since this choice of $\alpha_N$ does not lead to high-order classical convergence down to zero, but only down to approximately $\epsilon$.

5. The condition number of the matrix $A$. In the previous section, we demonstrated stability and accuracy, provided $\alpha$ and $N$ scale in the appropriate manner with $h$. Yet, as discussed in §2.3, it is important that the condition number of the matrix $A$ also remains bounded as $h \to 0$ for the same choices of $\alpha$ and $N$. If it does, the number of conjugate gradient iterations required to compute the approximation is $\mathcal{O}(1)$ irrespective of the problem size. We now show that this is indeed the case.

**Theorem 5.1.** Suppose that $h \leq 1/2$. Then the condition number of the matrix $A$ satisfies

$$\kappa(A) \leq \sqrt{\frac{1 + \Theta(\alpha, N, h)}{1 - \Theta(\alpha, N, h)}},$$

where, for any $0 < \delta < N^2$,

$$\Theta(\alpha, N, h) \leq c \left( Nh + N\sqrt{h} \sqrt{1 - \alpha + \delta^2} \left( \frac{h^2 N^2}{\delta^2} + \frac{h N}{\delta} + \frac{h N^2 (1 - \alpha)}{\delta^2} \right)^2 + \sqrt{\delta^2 + h^2 N^2 + h N^2 \sqrt{1 - \beta^2 (1 - \delta^2 / N^2)^2}} \right),$$

(5.1)

for some constant $c > 0$ independent of $\delta$, $\alpha$, $N$ and $h$, where $\beta = \sin(\alpha \pi/2)$.

The proof of this theorem is given in the appendix.

**Corollary 5.2.** Consider the following three cases:

(i) $0 \leq \alpha < 1$ fixed,

(ii) $\alpha = 1$,

(iii) $\alpha = \alpha_N \to 1^-$ as $N \to \infty$ with $\alpha_N \sim 1 - \alpha_0/N^\sigma$ for $0 < \sigma \leq 1$ and $\alpha_0 > 0$.

For each $\epsilon > 0$, there exists a $c_0(\epsilon) > 0$ such that

$$\kappa(A) \leq \sqrt{\frac{1 + \epsilon}{1 - \epsilon}},$$

provided $N \leq c_0(\epsilon) h^{-\gamma}$, where $\gamma$ satisfies

(i) $\gamma = 1/2$, 

(ii) $\gamma = 1$, 

(iii) $\gamma = \frac{1}{2 - \sigma}.$
Proof. By Theorem 5.1, it suffices to provide conditions under which \( \Theta(\alpha, N, h) \) is bounded away from 1. Suppose first that \( \alpha = 1 \). Then (5.1) gives

\[
\Theta(1, N, h) \leq c \left( Nh + \delta^2 \left( \frac{h^2 N^2}{\delta^2} + \frac{hN}{\delta} \right)^2 + \sqrt{\delta^2 + h^2 N^2 + hN\sqrt{2}} \right)
\]

Setting \( \delta = Nh \) gives

\[
\Theta(1, N, h) \leq c' \left( Nh + N^2 h^2 \right),
\]

for some constant \( c' \) independent of \( N \) and \( h \). Hence, taking \( Nh \) sufficiently small ensures \( \Theta(1, N, h) \) is bounded away from one.

Next suppose that \( 0 \leq \alpha < 1 \) is fixed. Then (5.1) reduces to

\[
\Theta(\alpha, N, h) \leq c \left( N\sqrt{h} + \frac{h^2 N^4}{\delta^2} + \delta \right).
\]

Setting \( \delta = hN^2 \) now gives \( \Theta(\alpha, N, h) \leq c \left( N\sqrt{h} + hN^2 \right) \) as required.

Finally, consider the case \( \alpha N \sim 1 - \alpha_0/N^\sigma \) for fixed \( \alpha_0 > 0 \) and \( 0 < \sigma \leq 1 \). Note that

\[
N\sqrt{h}\sqrt{1 - \sigma} \sim \sqrt{\alpha_0 N^{1-\sigma/2}} \sqrt{h},
\]

as \( N \to \infty \), uniformly in \( h \), and also that

\[
hN^2(1 - \alpha) \sim \alpha_0 hN^{2-\sigma},
\]

and

\[
hN^2 \sqrt{1 - \beta^2(1 - \delta^2/N^2)^2} \sim O(hN^{2-\sigma}), \quad N \to \infty, h \to 0.
\]

It follows that the right-hand side of (5.1) is small provided \( N = O(h^{-\frac{1}{2-\sigma}}) \), as required. \( \Box \)

Comparing this with §4.2, we note that the same scalings of \( N \) with \( h \) that ensure stability and accuracy of the approximation also ensure good conditioning of the linear system. Unfortunately, unlike in §4.2 we have no explicit values for the constant in this case. Although careful bookkeeping in the proof would give such a constant, it would likely be woefully pessimistic. However, we note that \( A \) is at least invertible for any choice of \( N, \alpha \) and \( h \), provided the number of points \( M \geq N \). Moreover, good conditioning of \( A \) can easily be checked numerically.

6. Fast implementation. The matrix \( A \) of the mapped polynomial method can be written as \( A = WFC \), where \( W = \text{diag}(\sqrt{\mu_0}, \ldots, \sqrt{\mu_M}) \in \mathbb{C}^{(M+1) \times (M+1)} \) and \( C = \text{diag}(\sqrt{1/\pi}, \sqrt{2/\pi}, \ldots, \sqrt{2/\pi}) \in \mathbb{C}^{(N+1) \times (N+1)} \) are diagonal, and

\[
F = \left\{ \cos(n \cos^{-1}(m_\alpha(z_m))) \right\}_{m=0, n=0}^{M, N} \in \mathbb{C}^{M \times N}.
\]

Our goal is to find the coefficient vector \( a \) that minimizes \( \|Aa - b\|_2 \), where the entries of the vector \( b \) are given by \( b_m = \sqrt{\mu_m} f(z_m) \). In our computations, the points \( z_m \) are either equispaced or scattered on the interval \([-1,1]\).
The operations \( a \mapsto Fa \) and \( b \mapsto F^*b \) can both be evaluated in \( \mathcal{O}(M \log M) \) operations using a NFFT package. In our implementation we use the NFFT library developed at the Mathematical Institute of the University of Lübeck [25, 31]. This library also includes the nonuniform fast cosine transform (NFCT), which the transformation performed by \( F \).

In order to find the expansion coefficients, we solve the least-squares problem with the LSQR algorithm [27]. In our experiments we use MATLAB’s LSQR function with tolerance \( 10^{-12} \) and the NFCT with the default parameters provided in the NFFT library. The number of iterations required by the solver is proportional to the square root of the condition number of \( A \). Hence, as long as \( A \) remains well-conditioned, the computational cost for finding the expansion coefficients is approximately \( \mathcal{O}(M \log M) \) operations. The performance of the NFFT implementation is demonstrated in the next section.

7. Numerical examples. In this section we present numerical results for approximations on \([-1, 1]\). Figure 2 shows the condition numbers (Lebesgue constants) \( \kappa_N^\alpha \) for equispaced nodes for several values of \( M \) and four choices of the least-squares aspect ratio. As expected, when \( M/N = 1 \), \( \kappa_N^\alpha \) is too large for practical computations. The condition number improves significantly as the oversampling rate is increased. The bottom-left panel of Fig. 2 indicates that \( \kappa_N^\alpha \) is approximately \( 10^3 \) if \( \alpha = 1 + \frac{2\log_{10} N}{N\pi} \), which is how we chose \( \alpha \) in all computations for the remainder of this paper.

Figure 3 confirms our results from §5 that the choice \( \alpha = 1 + \frac{2\log_{10} \epsilon}{N\pi} \) leads to stable computations when the least-squares process is computed using mapped Chebyshev polynomials as the approximation basis. Notice that \( \kappa(A) \) grows at a sub-algebraic rate and that for \( M/N = 2 \) it remains roughly \( 10^3 \) for practical values of \( M \). This indicates that, in double precision, oversampling by a factor of 2 is sufficient if the desired accuracy is roughly \( 10^{-12} \).

We present the error in the approximation of analytic functions in Figure 4. In all four cases the mapped polynomial approximations were obtained with \( M = 2N \). The norm of the error \( \|f - F_{N,Z}^\alpha(f)\|_\infty \) was estimated on a fine grid of 4000 uniformly distributed points. Notice that the method is particularly accurate for the function \( f_1(x) = 1/(1 + 100x^2) \). In fact geometric convergence can be observed on this plot. By contrast, polynomial interpolation of \( f_1 \) is known to diverge with the error growing exponentially fast near the ends of the interval. For reference, the errors for cubic splines (not-a-knot) and polynomial least-squares are also included. For stability the degree of the polynomial least-squares approximation must satisfy \( N = O(\sqrt{M}) \) [32]. In our numerical examples, \( N = 4\sqrt{M} \) was used. The superior convergence of the mapped approximations can also be observed for the functions \( f_2(x) = \frac{1}{1 + 16\sin^2(\pi x)} \) and \( f_3(x) = \sin(200x) \), which is entire but highly oscillatory. In the latter case, the convergence of the mapped polynomial scheme starts with a resolution of approximately 4.4 points per wavelength. The error then sharply drops several orders of magnitude and then slowly asymptotes to about \( 10^{-10} \). We point out that the condition number of \( f_3 \) is 200, and hence a couple of digits of accuracy are expected to be lost (in comparison to the other error plots) to rounding errors.

The function \( f_4(x) = \sqrt{1.01 + x} \) has a singularity near \( x = -1 \). In this case, the error decay for approximation on equispaced nodes is significantly slower. The effective rate seems to be sub-geometric for all values of \( M \) used in Fig. 4. Moreover, the polynomial least-squares is slightly more accurate than the mapped approximation. In contrast to the Runge function \( f_1 \), which has poles near \( x = 0 \), \( f_4 \) is most difficult
Fig. 2. Numerically estimated condition numbers $\kappa_{\alpha N}$ for several values of $\alpha$ and $N$. The colormap shows $\log_{10}(\kappa_{\alpha N})$. Four least-square aspect ratios (number of points / approximation degree) are considered: 1, 1.5, 2, and 2.5. The solid lines represent the curves $\alpha = 1 + \frac{2 \log \epsilon}{N \pi}$, with $\epsilon = 10^{-4}, 10^{-10},$ and $10^{-16}$.

Fig. 3. Condition number of the least-squares matrix $A$ for several values of $M$ (number of equispaced data points) and three least-squares aspect ratios $M/N$. The mapping parameter was chosen so that $\alpha = 1 + \frac{2 \log \epsilon}{M \pi}$. 
A MAPPED POLYNOMIAL METHOD

\[ f_1(x) = \frac{1}{1+100x^2} \]
\[ f_2(x) = \frac{1}{1+16\sin^2(7x)} \]
\[ f_3(x) = \sin(200x) \]
\[ f_4(x) = \sqrt{1.01 + x} \]

Fig. 4: Error in the approximation of four functions sampled at \( M \) equispaced points on \([-1, 1]\). Polynomial least-squares and cubic spline approximations are included for reference. The mapping parameter was chosen so that \( \alpha = 1 + \frac{2\log 10^{-12}}{N\pi} \) and \( M = 2N \).

Figure 5 presents the error in the approximation of \( f_1 \) and \( f_4 \) from scattered data. In this computation the data points were chosen as perturbation of equispaced nodes. More precisely,

\[ z_m = \delta_m + (-1 + 2m/M), \quad m = 1 \ldots M - 1, \quad z_0 = -1, \quad z_M = 1, \]

where the perturbations \( \delta_m \) were drawn uniformly from the open interval \((-1/M, 1/M)\). The error decay in this case is in good agreement with the equispaced case as expected.

The numerical results presented in Fig. 4 and Fig. 5 were computed using the NFFT implementation described in §6. Fig. 6 presents the elapsed time required to approximate \( f_5(x) = 1/(1 + 100\sin^2(30x)) \) on a 2010 MacBook Pro laptop (3.06 GHz Intel Core 2 Duo). The number of iterations used by LSQR is also reported in Fig. 6 (right panel). Notice that the number of iterations remains low even when the number of points is more than \( 10^4 \). For reference, elapsed time for the computation of the least-squares approximation using MATLAB’s backslash (which uses a Householder QR factorization) is also included. When \( M = 10^4 \), the LSQR iteration using NFFTs is roughly a thousand times faster than the matrix QR direct solver.
Fig. 5. Error in the approximation of two functions sampled at $M$ uniformly scattered points on $[-1, 1]$. Polynomial least-squares and cubic spline approximations are included for reference.

Fig. 6. Left: Elapsed time to approximate $f_5(x) = 1/(1 + 100 \sin^2(30x))$ using NFFTs and matrix computations. Dashed lines correspond to $O(M^3)$, $O(M^2)$ and $O(M)$. Right: Number of iterations used by LSQR.

8. Conclusions. The purpose of this paper was to introduce an efficient method for high-accuracy approximation from scattered grids based on mapped polynomials. Through a judicious choice of the parameter $\alpha$, the method has an asymptotically optimal scaling of the dimension of the approximation space $N$ with the maximal spacing $h$. While this parameter choice forfeits convergence down to zero, high accuracy is expected if the quantity $\epsilon$ is chosen close to machine epsilon. Efficient implementation of the method is achieved using NFFTs.

There are a number of topics we have not addressed in this paper. The first concerns the behaviour of the best approximation error $E_N^\alpha(f)$ for the parameter choice (3.3). Although this choice was derived so as to ensure the error bound of Theorem 3.3 is roughly $\epsilon$ for large $N$, this says nothing about how fast the error decreases in practice. The numerical experiments of the previous section suggest that the error decreases rapidly, at least initially, when it is orders of magnitude bigger than $\epsilon$. But how does one make precise mathematical statements to this effect when, after all, the approximation is not guaranteed to converge? It turns out that this can be done, but it is beyond the scope of this paper. We will report the details in a
future work.

Second, we have only presented our method in one dimension. For functions defined on hypercubes, an extension to higher dimensions using a tensor product of the one-dimensional mapping is conceptually straightforward. We expect that much of the analysis, in particular, the various scalings derived in §4, will also carry over to this setting. However, there may be more efficient, non-tensor product, approaches which warrant further investigations.

Other topics for investigation include the choice of mapping $m_\alpha$. We have used (2.1) throughout, however there are other possibilities. See [24] for an overview. We leave the question of the best choice of mapping for future work. Somewhat related to this is the optimal choice of $\alpha$. Here we have used the choice (3.3) due to Kosloff & Tal Ezer. However, other strategies may bring further benefits.

Acknowledgements. We thank Nick Trefethen and Nick Hale for fruitful discussions on using mapped methods for equispaced approximation. Much of this paper was developed during the authors’ visits to the American Institute of Mathematics (AIM). We thank AIM for providing a stimulating research environment. RP was supported in part by AFOSR FA9550-15-1-0152 and NSF-DMS1522639. BA was supported in part by the Alfred P. Sloan Foundation and NSERC 611675.

Appendix A. Proof of Theorem 5.1. To prove this theorem, we need several preliminary observations. First, let $y_n = m_\alpha(z_n)$, $n = -1, \ldots, M + 1$. Then

$$y_{n+1} - y_n \leq \|m'_\alpha\|_\infty h \leq \frac{\alpha \pi}{2\beta} h$$

Since

$$\frac{\alpha \pi}{2\beta} = \frac{\alpha \pi / 2}{\sin(\alpha \pi / 2)} \leq \frac{\pi}{2}, \quad 0 \leq \alpha \leq 1,$$

we find that

$$y_{n+1} - y_n \leq \frac{\pi h}{2}, \quad n = -1, \ldots, M.$$  \hspace{1cm} (A.2)

Unfortunately, this shall not be sufficient to prove the theorem, since it does not describe how the points $y_n$ cluster near the endpoints. To this end, we have the following:

**Lemma A.1.** Suppose that $y_n \geq 0$ and that $y \in [y_n, y_{n+1}]$. Then

$$|y - y_n| \leq \frac{\pi h}{2} \sqrt{1 - \beta^2(y_n)^2}, \quad |y - y_n| \leq \frac{\pi h}{2} \left(\pi h + \sqrt{1 - \beta^2(y_{n+1})^2}\right),$$

where $\beta$ is as in (2.2). Conversely, if $y_{n+1} \leq 0$ and $y \in [y_n, y_{n+1}]$ then we have

$$|y - y_n| \leq \frac{\pi h}{2} \sqrt{1 - \beta^2(y_{n+1})^2}, \quad |y - y_n| \leq \frac{\pi h}{2} \left(\pi h + \sqrt{1 - \beta^2(y_n)^2}\right),$$
Proof. Let \( y = m_\alpha(z) \) and \( y_n = m_\alpha(z_n) \). Then, for \( y_n \geq 0 \), or equivalently \( z_n \geq 0 \),

\[
|y - y_n| = \frac{1}{\sin(\alpha \pi/2)} |\sin(\alpha \pi z/2) - \sin(\alpha \pi z_n/2)|
\]

\[
\leq \frac{\alpha \pi/2}{\sin(\alpha \pi/2)} |z - z_n| |\cos(\alpha \pi \xi/2)|, \quad \xi \in [z_n, z]
\]

\[
\leq \frac{\alpha \pi/2}{\sin(\alpha \pi/2)} |z - z_n| |\cos(\alpha \pi z_n/2)|
\]

\[
\leq \frac{\alpha \pi/2}{\sin(\alpha \pi/2)} \beta h \sqrt{1 - \sin^2(\alpha \pi z_n/2)}
\]

\[
= \frac{\alpha \pi}{2 \beta} \beta h \sqrt{1 - \beta^2(y_n)^2}.
\]

To deduce the first inequality, we use (A.1). For the second, we first note that it suffices to take \( y = y_{n+1} \) and then let \( z = y_{n+1} - y_n \). Then by the first inequality

\[
z^2 \leq \frac{\pi^2 h^2}{4} (1 - \beta^2(y_{n+1})^2 + \beta^2 ((y_{n+1})^2 - (y_n)^2))
\]

\[
\leq \frac{\pi^2 h^2}{4} (1 - \beta^2(y_{n+1})^2 + 2\beta^2 y_n).
\]

A simple exercise gives that if \( z^2 \leq cz + d^2 \) with \( c, d \geq 0 \) then \( z \leq c + d \). Hence we obtain

\[
z \leq \frac{\pi^2 h^2}{2} + \frac{\pi h}{2} \sqrt{1 - \beta^2(y_{n+1})^2},
\]

as required. The case of \( y_{n+1} \leq 0 \) is similar. \( \square \)

We are now ready to prove Theorem 5.1. Throughout the proof, we shall use the notation \( a \lesssim b \) to mean that there exists a constant \( c > 0 \) independent of \( N, \alpha, Z \) and \( p \in \mathbb{P}_N \) such that \( a \leq cb \). Let \( y_n = m_\alpha(z_n) \). By Lemma 2.2, we wish to find constants \( c_1, c_2 > 0 \) such that

\[
\text{(A.3)} \quad c_1 \|p\|^2_w \leq \sum_{n=0}^{M} \mu_n \|p(y_n)\|^2 \leq c_2 \|p\|^2_w, \quad \forall p \in \mathbb{P}_N,
\]

where \( w(y) = 1/\sqrt{1 - y^2} \), in which case the condition number \( \kappa(A) \leq \sqrt{c_2/c_1} \). We now note the following. The weights

\[
\mu_n = \frac{1}{2} \left( \mu_n^l + \mu_n^r \right), \quad \mu_n^l = \int_{y_n}^{y_{n+1}} w(y) \, dy, \quad \mu_n^r = \int_{y_{n-1}}^{y_n} w(y) \, dy.
\]

Thus the theorem holds provided (A.3) holds with \( \mu_n \) replaced by \( \mu_n^l \) and \( \mu_n^r \). By symmetry, it suffices to the result for \( \mu_n^l \) only. That is, we need only show that

\[
c_1 \|p\|^2_w \leq \sum_{n=0}^{M} \mu_n^l \|p(y_n)\|^2 \leq c_2 \|p\|^2_w, \quad \forall p \in \mathbb{P}_N.
\]

Define the function \( \chi(y) = \sum_{n=0}^{N} p(y_n) \delta_{(y, y_{n+1})}(y) \). By definition of \( \mu_n^l \), we have that

\[
\sum_{n=0}^{M} \mu_n^l \|p(y_n)\|^2 = \int_{-1}^{1} |\chi(y)|^2 w(y) \, dy = \|\chi\|^2_w.
\]
Since

(A.4) \[ \|p\|_w - \|p - \chi\|_w \leq \|\chi\|_w \leq \|p\|_w + \|p - \chi\|_w, \quad \|p\|_w^2 = \int_{-1}^{1} \frac{|p(y)|^2}{\sqrt{1 - y^2}} \mathrm{d}y, \]

it suffices to estimate \(\|p - \chi\|_w\). We have

(A.5) \[ \|p - \chi\|_w^2 = \sum_{n=0}^{N} \int_{y_n}^{y_{n+1}} |p(y) - p(y_n)|^2 w(y) \mathrm{d}y + \int_{-1}^{y_0} |p(y)|^2 w(y) \mathrm{d}y + \sum_{n=0}^{M} J_n + I. \]

We now note the following inequality:

(A.6) \[ \|p\|_\infty \leq \sqrt{\frac{2N + 1}{\pi}} \|p\|_w, \quad \forall p \in P_N. \]

This follows immediately by expanding \(p\) in normalized Chebyshev polynomials \(c_n T_n\), where \(c_0 = \sqrt{1/\pi}\) and \(c_n = \sqrt{2/\pi}\) otherwise, and using the Cauchy–Schwarz inequality. In particular, this gives

\[ \int_{-1}^{y_0} |p(y)|^2 w(y) \mathrm{d}y \leq \|p\|_w^2 \frac{2N + 1}{\pi} \int_{-1}^{y_0} w(y) \mathrm{d}y \lesssim N \sqrt{1 + y_0 \|p\|_w^2}. \]

Note that \(1 + y_0 = y_0 - y_{-1}\). Hence, by Lemma A.1

\[ |1 + y_0| \lesssim h^2 + h \sqrt{1 - \beta^2} = h^2 + h \cos(\alpha \pi/2) \lesssim h^2 + h(1 - \alpha), \]

since \(\cos(\alpha \pi/2) \leq \pi(1 - \alpha)/2, 0 \leq \alpha \leq 1\). Thus

(A.7) \[ \int_{-1}^{y_0} |p(y)|^2 w(y) \mathrm{d}y \lesssim \left(Nh + N \sqrt{h \sqrt{1 - \alpha}}\right) \|p\|_w^2. \]

We now focus on the other terms of (A.5). Consider the integral \(J_n\):

\[ J_n = \int_{y_n}^{y_{n+1}} \left| \int_{y_n}^{y} p'(t) \mathrm{d}t \right|^2 \frac{1}{\sqrt{1 - y^2}} \mathrm{d}y \leq \int_{y_n}^{y_{n+1}} \left( \int_{y_n}^{y} \sqrt{1 - t^2} |p'(t)|^2 \mathrm{d}t \right) \left( \int_{y_n}^{y} \frac{1}{\sqrt{1 - t^2}} \mathrm{d}t \right) \frac{1}{\sqrt{1 - y^2}} \mathrm{d}y \]

(A.8) \[ \leq \left( \int_{y_n}^{y_{n+1}} \frac{1}{\sqrt{1 - y^2}} \mathrm{d}y \right)^2 \int_{y_n}^{y_{n+1}} \sqrt{1 - t^2} |p'(t)|^2 \mathrm{d}t \]

We wish to estimate the first integral. To do so, let \(0 < \epsilon < 1\) be a parameter (whose
value we choose later). Suppose first that $y_n \geq 0$. By Lemma A.1,

$$
\int_{y_n}^{y_{n+1}} \frac{1}{\sqrt{1-y^2}} dy \leq \frac{y_{n+1} - y_n}{\sqrt{1 - (y_{n+1})^2}}
$$

\[
\lesssim \frac{h^2}{\sqrt{1 - (y_{n+1})^2}} + h \sqrt{\frac{1 - \beta^2(y_{n+1})^2}{1 - (y_{n+1})^2}}
\]

\[
\lesssim \frac{h^2}{\sqrt{\epsilon}} + h \sqrt{1 + \frac{(1 - \beta^2)(y_{n+1})^2}{1 - (y_{n+1})^2}}
\]

\[
\lesssim \frac{h^2}{\sqrt{\epsilon}} + h \sqrt{1 + (1 - \alpha)^2/\epsilon}
\]

\[
\lesssim \sqrt{\epsilon} \left( \frac{h^2}{\epsilon} + \frac{h(1 - \alpha)}{\epsilon} \right).
\]

Hence

$$
\sum_{y_n \geq 0, y_{n+1} < 1 - \epsilon} J_n \lesssim \epsilon \left( \frac{h^2}{\epsilon} + \frac{h(1 - \alpha)}{\epsilon} \right)^2 \|p'\|_{1/w}^2
$$

(A.9)

\[
\lesssim \epsilon N^2 \left( \frac{h^2}{\epsilon} + \frac{h(1 - \alpha)}{\epsilon} \right)^2 \|p\|_w^2.
\]

Note that the second step is due to the inequality $\|p'\|_{1/w} \lesssim N\|p\|_w$ (see, for example, [12, (5.5.5)]). Near-identical arguments also give

$$
\sum_{y_n \leq 0, y_{n+1} > 1 - \epsilon} J_n \lesssim \epsilon N^2 \left( \frac{h^2}{\epsilon} + \frac{h(1 - \alpha)}{\epsilon} \right)^2 \|p\|_w^2.
$$

(A.10)

Now consider terms $J_n$ with $y_n \geq 0$ and $y_{n+1} \geq 1 - \epsilon$. Then

$$
J_n \leq \|p\|_w^2 \int_{y_n}^{y_{n+1}} w(y) dy,
$$

and therefore we get that

$$
\sum_{y_n \geq 0, y_{n+1} \geq 1 - \epsilon} J_n \leq \|p\|_w^2 \int_{y_n}^{y_{n+1}} w(y) dy \lesssim N\|p\|_w^2 \sqrt{1 - y_n}.
$$

By Lemma A.1,

$$
1 - y_n = 1 - y_{n+1} + y_{n+1} - y_n \leq \epsilon + y_{n+1} - y_n \lesssim \epsilon + h^2 + h \sqrt{1 - \beta^2(1 - \epsilon)^2}
$$

Therefore

$$
\sum_{y_n \geq 0, y_{n+1} \geq 1 - \epsilon} J_n \lesssim N \sqrt{\epsilon + h^2 + h \sqrt{1 - \beta^2(1 - \epsilon)^2}} \|p\|_w^2.
$$

(A.11)
A similar estimate holds for the case $y_{n+1} \leq 0, y_n \leq -1 + \epsilon$. Finally, let $n_0$ be such that $y_{n_0} \leq 0$ and $y_{n_0+1} > 0$. Without loss of generality, suppose that $|y_{n_0}| \geq |y_{n_0+1}|$ and therefore $|y_{n_0}| \leq h\pi/2$. Hence

\begin{equation}
(A.12) \quad \int_{y_{n_0}}^{y_{n_0+1}} |p(y)|^2 w(y) \, dy \leq \|p\|_w^2 oh w(y_{n_0}) \lesssim \frac{Nh}{\sqrt{1-h^2 \pi^2/4}} \|p\|_w^2 \lesssim Nh \|p\|_w^2,
\end{equation}

since $h < 1/2 < 2/\pi$. With this to hand, we now substitute (A.7), (A.9), (A.10), (A.11) and (A.12) into (A.5) to get

$$
\|x - p\|_2^2 \lesssim \left[ \left( Nh + N\sqrt{h(1-\alpha)} \right) + \epsilon N^2 \left( \frac{h^2}{\epsilon} + \frac{h}{\sqrt{\epsilon}} + \frac{h(1-\alpha)}{\epsilon} \right)^2 
+ N \sqrt{\epsilon} + h^2 + h \sqrt{1 - \beta^2 (1-\epsilon)^2} + Nh \right] \|p\|_w^2,
$$

The result now follows by setting $\epsilon = \delta^2/N^2$.

REFERENCES