LEAST SQUARES SPHERICAL HARMONICS APPROXIMATION ON THE CUBED SPHERE

JEAN-BAPTISTE BELLET AND JEAN-PIERRE CROISILLE

Abstract. The Cubed Sphere grid is an important tool to approximate functions or data on the sphere. We introduce an approximation framework on this grid based on least squares and on a suitably chosen subspace of spherical harmonics. The main claim is that for the equiangular Cubed Sphere with step size $\pi/(2N)$, a relevant spherical harmonics subspace is the one of all SH of degree less than $2N$. This choice, which matches the Nyquist's cutoff frequency along the equatorial great circle, provides an approximation problem both well-posed and well-conditioned. A series of theoretical and numerical results supporting this fact are presented.

1. Introduction

We consider the approximation of functions defined on the Cubed Sphere grid by mean of Spherical Harmonics. Suppose that $\varkappa \in \mathrm{CS}_N \mapsto y(\varkappa)$ is a gridfunction (a set of data) defined at the nodes $\varkappa$ of the Cubed Sphere CS_N. We approximate these data by a Spherical Harmonic (SH) $f \in \mathcal{Y}_D$ where $\mathcal{Y}_D = \mathbb{Y}_0 \oplus \cdots \oplus \mathbb{Y}_D$ is the space of all spherical harmonics with degree less than or equal to D. The standard least squares approximation problem is

$$\inf_{f \in \mathcal{Y}_D} \sum_{\varkappa \in \mathrm{CS}_N} |f(\varkappa) - y(\varkappa)|^2. \tag{LS}$$

Our main observation is that the choice $D = 2N - 1$ leads to a well posed and well conditioned problem. In addition, the resulting SH approximant possesses interesting properties for approximating a function given at the nodes of CS_N only. These facts are assessed theoretically and numerically hereafter.

In [4], we have introduced a particular SH subspace with good interpolating properties on the Cubed Sphere (Lagrange interpolation). This space consists of the direct sum $\mathcal{Y}_{2N-1} \oplus \mathcal{Y}'$. The second subspace $\mathcal{Y}'$ complements $\mathcal{Y}_{2N-1}$. It is such that $\mathcal{Y}' \subsetneq \mathbb{Y}_{2N} \oplus \cdots \oplus \mathbb{Y}_{3N}$. This interpolation framework has been used in [5] to define new spherical quadrature rules of accuracy comparable to optimal ones (Lebedev rules). Here, we show that the first subspace $\mathcal{Y}_{2N-1}$ is a suitable choice if one wants a least squares approximant instead of an interpolant.

Approximating and interpolating data on the sphere by spherical harmonics is an old and important topic. It is still widely used nowadays in many areas in physics such as quantum chemistry, numerical climatology, cosmology, gravitation, neutronic, etc. It is also central in harmonic analysis on spheres and balls since it is the three dimensional counterpart of trigonometric approximation. For fundamental and applied aspects of spherical harmonics analysis, refer to the two recent monographs [2,6] (theory and applications). Concerning applications in geomathematics, many chapters in the reference [7] are concerned with spherical harmonics. Regarding specifically least squares, recent works include [1, 8].

The outline is as follows. In Section 2 the setup of the problem is given. A general positive weight function is included to define the least squares functional. Theoretical results are given in Section 3. These results in particular concern estimates of the condition number of the collocation matrix. Section 4 is devoted to the structure of the matrix involved in the least squares problem (LS). In particular, the attractive block structure of this matrix is described in case of a rotationally invariant weight function. Finally, various numerical results are reported in Section 5.

Date: Mar 08 2023.

Key words and phrases. Cubed sphere, least squares, spherical harmonics.

2. Setup of the Least Squares problem

Our notation is as follows. For any $N \geq 1$, the equiangular Cubed Sphere CS_N is the set of $\bar{N} = 6N^2 + 2$ nodes $\varkappa \in \mathbb{S}^2$, with Cartesian coordinates in $\mathbb{R}^3$ given by

$$\mathrm{CS}_N \triangleq \Big\{ \tfrac{1}{r}(\pm 1, u, v),\ \tfrac{1}{r}(u, \pm 1, v),\ \tfrac{1}{r}(u, v, \pm 1); \\ r = (1+u^2+v^2)^{1/2},\ u = \tan\tfrac{i\pi}{2N},\ v = \tan\tfrac{j\pi}{2N},\ -\tfrac{N}{2} \leq i, j \leq \tfrac{N}{2} \Big\}.$$

Refer to [11–13]. For any function $g : x \in \mathbb{S}^2 \mapsto g(x) \in \mathbb{R}$, we denote g^* the gridfunction defined by $g^* : \varkappa \in \mathrm{CS}_N \mapsto g^*(\varkappa)$,

$$g^*(\varkappa) = g(\varkappa),\ \varkappa \in \mathrm{CS}_N.$$

We denote Y_n^m the Spherical Harmonic with index (n, m) in real form,

$$Y_n^m(x(\theta,\phi)) = q_n^m(\sin\theta) \cdot (\cos\theta)^{|m|} \cdot \begin{cases} \sin m\phi, & m < 0, \\ \cos m\phi, & m \geq 0. \end{cases} \tag{1}$$

where

$$x(\theta,\phi) = (\cos\theta\cos\phi, \cos\theta\sin\phi, \sin\theta), \quad \phi \in \mathbb{R},\ \theta \in [-\pi/2, \pi/2]. \tag{2}$$

In (1), q_n^m is the polynomial of degree $n - |m|$, with the parity of $n + |m|$, defined by

$$q_n^m(t) = \sqrt{\tfrac{(n+1/2)(n-|m|)!}{\pi(n+|m|)!}} \cdot \Big(\tfrac{\mathrm{d}^{|m|+n}}{\mathrm{d}t^{|m|+n}} \tfrac{1}{2^n n!} (t^2-1)^n \Big) \cdot \begin{cases} -1, & m < 0, \\ \frac{1}{\sqrt{2}}, & m = 0, \\ 1, & m > 0. \end{cases} \tag{3}$$

We note

$$\langle u, v \rangle_{L^2(\mathbb{S}^2)} = \int_{\mathbb{S}^2} u(x)v(x)\mathrm{d}\sigma, \qquad \|u\|_{L^2(\mathbb{S}^2)} = \langle u, u \rangle_{L^2(\mathbb{S}^2)}^{1/2}.$$

Any $f \in L^2(\mathbb{S}^2)$ is expressed in the Hilbert basis $(Y_n^m)_{|m|\leq n,\, n\in\mathbb{N}}$ as the Fourier like decomposition

$$f = \sum_{|m|\leq n} \hat{f}_n^m Y_n^m, \quad \text{with} \quad \hat{f}_n^m = \langle f, Y_n^m \rangle_{L^2(\mathbb{S}^2)}. \tag{4}$$

For any $D \geq 0$, the space $\mathcal{Y}_D = \mathrm{Span}(Y_n^m)_{0\leq|m|\leq n\leq D} = \oplus_{n=0}^D \mathbb{Y}_n$ is such that $\dim \mathcal{Y}_D = (D+1)^2$. Let $\omega(\varkappa) > 0, \varkappa \in \mathrm{CS}_N$ be a given positive weight function. Let $y(\varkappa),\ \varkappa \in \mathrm{CS}_N$ be a set of data given at the nodes of the CS_N. The functional $\mathcal{L}$ is defined as

$$f \mapsto \mathcal{L}(f) = \sum_{\varkappa\in\mathrm{CS}_N} \omega(\varkappa)|f(\varkappa) - y(\varkappa)|^2. \tag{5}$$

We consider the least squares problem: find $f \in \mathcal{Y}_D$ solution of

$$\inf_{f\in\mathcal{Y}_D} \mathcal{L}(f). \tag{WLS}$$

We also use the quadrature rule $\mathcal{Q}$ associated to ω. For $f : \mathbb{S}^2 \to \mathbb{R}$, we have by

$$\begin{aligned} \mathcal{Q}(f) &= \sum_{\varkappa\in\mathrm{CS}_N} \omega(\varkappa) f(\varkappa) \\ &= \int_{\mathbb{S}^2} f(x)\mathrm{d}\sigma - e_N(f), \end{aligned} \tag{6}$$

where e_N denotes the quadrature error. In the particular case where the data y are such that $y = g^*$ for a given function g, we have

$$\mathcal{L}(f) = \|f - g\|^2_{L^2(\mathbb{S}^2)} - e_N(|f-g|^2). \tag{7}$$

For fixed values of N and D, we call the *Vandermonde matrix* of the problem the rectangular matrix A_N^D defined by

$$A_N^D = \big[Y_n^m(\varkappa)\big]_{\substack{\varkappa\in\mathrm{CS}_N \\ |m|\leq n\leq D}} \in \mathbb{R}^{\bar{N}\times(D+1)^2}. \tag{8}$$

We define the diagonal matrix $\Omega_N \in \mathbb{R}^{\bar{N}\times\bar{N}}$ by

$$\Omega_N = \mathrm{diag}(\omega(\mathfrak{x}))_{\mathfrak{x}\in \mathrm{CS}_N} \in \mathbb{R}^{\bar{N}\times\bar{N}}. \tag{9}$$

In vector form, the problem (WLS) is expressed as

$$\inf_{\hat{f}\in\mathbb{R}^{(D+1)^2}} \|\Omega_N^{1/2}\big(A_N^D\hat{f} - y\big)\|^2, \tag{10}$$

where $y = [y(\mathfrak{x})]_{\mathfrak{x}\in\mathrm{CS}_N} \in \mathbb{R}^{\bar{N}}$. Uniqueness for (10), or equivalently for (LS) or (WLS), is equivalent to the injectivity of A_N^D. In this case, (10) is equivalent to the linear system

$$A_N^{D\intercal}\Omega_N A_N^D\,\hat{f} = A_N^{D\intercal}\Omega_N\,y. \tag{11}$$

A natural interpretation of (11) is as follows. Consider the following analog of the Discrete Fourier Transform (DFT) of the data $y = g^*$, located at the nodes of CS_N instead of at the $\theta_j = 2j\pi/N \in [0, 2\pi), j = 0, \ldots, N$, in the standard DFT. The Fourier-like coefficients are the components of the vector

$$\begin{aligned}\mathrm{DFT}(y) &= \Big[\sum_{\mathfrak{x}\in\mathrm{CS}_N} \omega(\mathfrak{x})Y_n^m(\mathfrak{x})y(\mathfrak{x})\Big]_{(n,m)}\\ &= A_N^{D\intercal}\Omega_N\,y.\end{aligned} \tag{12}$$

On the other hand, the analog of the Inverse Discrete Fourier Transform (IDFT) of a set of data $\hat{f} = [\hat{f}_n^m]_{0\le|m|\le n\le D}$ is the gridfunction

$$\begin{aligned}\mathrm{IDFT}[\hat{f}](\mathfrak{x}) &= \sum_{0\le|m|\le n\le D} \hat{f}_n^m Y_n^m(\mathfrak{x}), \quad \mathfrak{x}\in\mathrm{CS}_N,\\ &= \Big[(A_N^D)\hat{f}\Big](\mathfrak{x}).\end{aligned} \tag{13}$$

This means that in matrix form, A_N^D coincides with the IDFT operator. Therefore in terms of DFT/IDFT transforms, the solution $f \in \mathcal{Y}_D$ of (11) has coefficients $\hat{f} = [\hat{f}_n^m]$ solution of

$$\mathrm{DFT}\Big(\mathrm{IDFT}[\hat{f}] - y\Big) = 0. \tag{14}$$

For any N, there is a maximal degree D such that the matrix A_N^D is injective (full column rank), thus guaranteeing that (WLS) has a unique solution. The proof consists in observing that such degrees D form a nonempty set of integers. We have $\mathrm{rank}\,A_N^D \le \min\big((\bar{N}, (D+1)^2\big)$. Therefore assuming that A_N^D has full column rank implies that $(D+1)^2 \le \bar{N}$ which means

$$D \le \bar{N}^{1/2} - 1 \approx 2.45N - 1. \tag{15}$$

Fix a value N, and consider the Cubed Sphere CS_N. How to select $N \mapsto D_N$ in order for the two following conditions to hold?

$$(P)\quad \begin{cases} \text{(i) For every degree } D \le D_N\text{, the Vandermonde matrix } A_N^D \text{ is injective so that}\\ \text{the least squares problem (LS) has a unique solution.}\\ \text{(ii) The condition number } \mathrm{cond}(A_N^{D_N}) \text{ is bounded above for } N \to +\infty. \end{cases} \tag{16}$$

In other words, the matrix $A_N^{D_N}$ is required to satisfy both injectivity and asymptotic stability. In what follows, we assess that

$$D_N = 2N - 1 \tag{17}$$

is a natural candidate for *(P)* to be fullfilled.

Remark 1. Note that the value $D = 2N-1$ corresponds to the Nyquist cutoff angular frequency of a signal sampled with stepsize $\pi/(2N)$ (one dimensional problem). Note also that for a given integer N, it may exist $D > 2N-1$ such that A_N^D is injective. However, we are interested in a generic value of D, expressed in function of N such that Property *(P)* holds.

In Section 3, several theoretical results are proved, supporting (17). And Section 5 reports numerical results further supporting this claim.

3. Theoretical results

In this section we prove several facts supporting that $D = 2N - 1$ is a truncation value fullfilling Property *(P)*. Specifically we will show that

(1) For $1 \leq N \leq 4$, the condition $D \leq 2N - 1$ is equivalent to the injectivity of the matrix A_N^D, (Prop. 4).
(2) For $N \geq 5$, we show that $D \leq N + 2$ implies injectivity of the matrix A_N^D (Prop. 5). Combined with (15), this gives that a condition on the largest D for A_N^D to be injective is that
$$N + 2 \leq D \leq \sqrt{6N^2 + 2} - 1 \tag{18}$$
(3) Finally, Theorem 7 shows that $D = 2N$ gives that the condition number of A_N^{2N} is asymptotically unbounded. Thus $D \geq 2N$ does not satisfy *(ii)* in *(P)* and therefore one must select $D \leq 2N - 1$.

Remark 2. A full proof of the fact that the matrix A_N^{2N-1} is injective for all N is not yet available.

3.1. **The case $1 \leq N \leq 4$.** The case $1 \leq N \leq 4$ corresponds to a small Cubed Sphere grid ranging from $\bar{N} = 8, (N = 1)$ nodes to $\bar{N} = 98, (N = 4)$ nodes. Consider the spherical harmonic $Y_{2N}^{-2N} \in \mathbb{Y}_{2N}$, given by, see (1)
$$Y_{2N}^{-2N}(x(\theta, \phi)) = q_{2N}^{-2N}(\sin\theta) \cdot \cos^{2N}\theta \cdot \sin(2N\phi). \tag{19}$$
By shifting the angle ϕ by $\pi/4$ we obtain $f_N \in \mathbb{Y}_{2N}$ defined by
$$f_N(x(\theta, \phi)) = Y_{2N}^{-2N}(x(\theta, \phi - \tfrac{\pi}{4})). \tag{20}$$

Lemma 3 (Function f_N restricted to CS_N for $N \leq 4$)**.** *For $1 \leq N \leq 4$, the function $f_N \in \mathbb{Y}_{2N}$ vanishes at all nodes of CS_N ($f^* \equiv 0$). This implies that A_N^D has not full column rank if $N \leq 4$ and $D \geq 2N$.*

Proof. The spherical harmonic f_N is deduced from Y_{2N}^{-2N} by a rotation of $\pi/4$ around the pole axis. By invariance of $\mathbb{Y}_{2N}$ by rotation, we have $f_N \in \mathbb{Y}_{2N}$. In addition, for any $N \geq 4$, it turns out that CS_N is contained in the set M_N of meridians defined by
$$M_N \triangleq \Big\{ x(\theta, \phi) : \theta \in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}],\ \phi \equiv \tfrac{\pi}{4}\,(\tfrac{\pi}{2N}) \Big\}. \tag{21}$$
Along these meridians, the longitude angle ϕ is such that $2N(\phi - \frac{\pi}{4}) \equiv 0\,(\pi)$, hence
$$f_N(x(\theta, \phi)) = q_{2N}^{-2N}(\sin\theta) \cdot \cos^{2N}\theta \cdot \sin\big(2N(\phi - \tfrac{\pi}{4})\big) = 0. \tag{22}$$
This implies that $f(\varkappa) = 0$ for all $\varkappa \in \mathrm{CS}_N$. In particular, for any $D \geq 2N$, the linear map $f \in \mathcal{Y}_{2N} \mapsto f^*$ is not injective. Therefore for $D \geq 2N$, the matrix A_N^D is not injective. □

Proposition 4 (Full column rank in the case $1 \leq N \leq 4$)**.** *Let A_N^D be the Vandermonde matrix* (8)*. If $1 \leq N \leq 4$, then*
$$A_N^D \text{ has full column rank } \Leftrightarrow D \leq 2N - 1;$$
In particular, the largest degree D_N such that $A_N^{D_N}$ has full column rank is $D_N = 2N - 1$.

Proof. Equivalently, we prove that the linear map $f \in \mathcal{Y}_D \mapsto [f(\xi)]_{\xi \in \mathrm{CS}_N} \in \mathbb{R}^{6N^2+2}$ is injective. So we fix $f \in \mathcal{Y}_D$ such that $f(\xi) = 0$ for every $\xi \in \mathrm{CS}_N$ and we prove that $f = 0$. First, we introduce $2N$ meridian circles associated to the longitudes $\phi \equiv \frac{\pi}{4}\,(\frac{\pi}{2N})$,
$$\mathscr{C}(\psi) = \{x(\theta, \psi),\, \theta \in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}]\} \cup \{x(\theta, \psi + \pi),\, \theta \in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}]\}, \quad \psi \equiv \tfrac{\pi}{4}\,(\tfrac{\pi}{2N}),$$
and we prove that f is null on these great circles, *i.e.*
$$f|_{\mathscr{C}(\psi)} = 0, \quad \psi \equiv \tfrac{\pi}{4}\,(\tfrac{\pi}{2N}), \tag{23}$$
where $f|_{\mathscr{C}(\psi)}$ denotes the restriction of f to $\mathscr{C}(\psi)$. The assumption $N \leq 4$ implies that each great circle $\mathscr{C}(\psi)$ contains $4N$ points of CS_N. Since f vanishes on CS_N, these points give $4N$ zeros for $f|_{\mathscr{C}(\psi)}$. Since $f|_{\mathscr{C}(\psi)}$ represents a trigonometric polynomial with degree at most D, with $4N$ zeros, and $4N \geq 2D + 1$, we obtain $f|_{\mathscr{C}(\psi)} = 0$.

Second, any great circle $\mathscr{C}$ not containing the pole $(0,0,1)$, contains $4N$ points in the set $\cup_{\psi\equiv\frac{\pi}{4}\,(\frac{\pi}{2N})}\mathscr{C}(\psi)$. Therefore (23) implies that f has $4N$ zeros on $\mathscr{C}$. It results that $f|_{\mathscr{C}}$ represents a trigonometric polynomial with degree at most D and $4N$ zeros; since $4N \geq 2D+1$, we obtain

$$f|_{\mathscr{C}} = 0, \quad (0,0,1) \notin \mathscr{C}. \tag{24}$$

Since the great circles in (23) and (24) cover the sphere, we have $f = 0$ on $\mathbb{S}^2$. □

3.2. **The case $N \geq 5$.** Lemma 3 and Prop. 4 give a full answer to the injectivity of A_{2N}^{2N-1} in the case $1 \leq N \leq 4$. Consider now the case $N \geq 5$. We have

Proposition 5 (Case $N \geq 5$)**.** *Suppose $N \geq 5$. We have*

$$D \leq N+2 \Rightarrow A_N^D \text{ has full column rank},$$

Therefore, the largest degree D such that A_N^D has full column rank satisfies

$$N+2 \leq D \leq \bar{N}^{1/2} - 1\,(\approx 2.45N - 1).$$

Proof. Fix $N \geq 5$ and $D \leq N+2$. Consider $f \in \mathcal{Y}_D$ such that $f(\xi) = 0$ for every $\xi \in \mathrm{CS}_N$. For $\psi = -\frac{\pi}{4}, \frac{\pi}{4}$, $\mathscr{C}(\psi)$ contains $4N$ points from CS_N, which implies

$$f|_{\mathscr{C}(\psi)=0}, \quad \psi = \pm\tfrac{\pi}{4}.$$

The two considered circles intersect at the poles $(0,0,\pm1)$. Therefore any tangential derivative of f is zero at the poles, and each pole is a zero of order at least 2. Next, for any other angle $\psi \equiv \frac{\pi}{4}\,(\frac{\pi}{2N})$, $\mathscr{C}(\psi)$ contains at least $2N+2$ zeros of f on CS_N, and the poles as two additional zeros of order 2. Rolle's Theorem implies that the derivative of $f|_{\mathscr{C}(\psi)}$ (identified with a trigonometric polynomial) has $2N+6$ zeros. Since it is a trigonometric polynomial with degree at most D and $2N+6 \geq 2D+1$, we obtain (23). And we conclude as in the proof of Lemma 3. □

Remark 6. In the proof with $N \geq 5$, the bottleneck on the degree comes from the meridian circles that do not contain $4N$ points of CS_N. If there were $4N$ points per meridian circle, one would obtain the degree $2N-1$.

As in the proof of Lemma 3, the function f_N defined in (20) vanishes on the set M_N defined in (21). This implies that f_N vanishes at all nodes of the four equatorial panels (I), (II),(III), (IV) of CS_N. Regarding panels (V) and (VI), f_N satisfies the estimate

$$|f_N(x(\theta,\phi))| \leq \gamma_N \cdot \cos^{2N}\theta, \quad \theta \in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}],\ \phi \in \mathbb{R}. \tag{25}$$

The constant γ_N is

$$\gamma_N = \sqrt{\tfrac{4N+1}{2\pi}} \cdot \tfrac{\sqrt{(4N)!}}{2^{2N}(2N)!} \sim \tfrac{1}{\pi^{1/2}}\left(\tfrac{2N}{\pi}\right)^{1/4}\,(\approx 0.504\,N^{1/4}). \tag{26}$$

(The Stirling formula has been used). The behaviour of f_N on the north panel (V) and south panel (VI) is obtained by inspecting the nodes located outside the set M_N in (21), where the estimate (25) holds. Let $H_N \subset \mathrm{CS}_N$ be the set of nodes defined by

$$H_N \triangleq \big\{\tfrac{1}{r}(u,v,\pm1) : r = (1+u^2+v^2)^{1/2},\ u = \tan\tfrac{i\pi}{2N},\ v = \tan\tfrac{j\pi}{2N},\\ -\tfrac{N}{2} < i,j < \tfrac{N}{2},\ |i| \neq |j| \text{ and } i \neq 0 \text{ and } j \neq 0\big\}. \tag{27}$$

It turns out (see Fig. 1) that

$$\mathrm{CS}_N \setminus M_N \subset H_N. \tag{28}$$

Furthermore, the number of nodes in the set H_N is given by

$$|H_N| = \begin{cases} 2(N-1)(N-3), & \text{if } N \text{ is odd,} \\ 2(N-2)(N-4), & \text{if } N \text{ is even.} \end{cases} \tag{29}$$

In the next theorem it is proved that f_N^* "almost vanishes" at all the Cubed Sphere nodes $\mathfrak{x} \in \mathrm{CS}_N$. This will show that when taking $D = 2N$, the Vandermonde matrix A_N^{2N} cannot have full column rank (injective) while keeping a bounded condition number.

Theorem 7 (Asymptotics for the condition number of A_N^D)**.** *Fix $N \geq 1$ and $D \geq 2N$.*

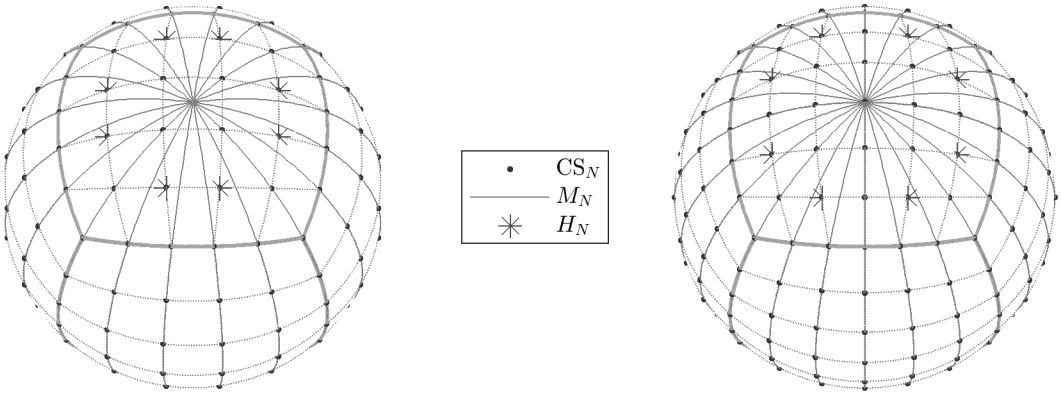

FIGURE 1. Equiangular Cubed Sphere and equiangular meridians. The Cubed Sphere CS_N (black dots) meshes $\mathbb{S}^2$ with equiangular arcs of great circles (dotted lines), including the radial projection of the edges of $[-1,1]^3$ (bold gray lines). The set M_N of equiangular meridians with longitude $\phi \equiv \frac{\pi}{4}\,(\frac{\pi}{2N})$ (gray lines) contains "many" points of CS_N; the remaining points of CS_N belong to the set H_N (indicated with star symbols) defined in (27). The size of H_N is given in (29), and is estimated by $|H_N| \sim \frac{1}{3}\bar{N}$. Left panel: N is odd ($N = 5$), right panel: N is even ($N = 6$).

(i) The smallest singular value of the matrix A_N^D, denoted by $\sigma_{\min}(A_N^D)$, satisfies

$$\sigma_{\min}(A_N^D)^2 \le \gamma_N^2 \cdot |H_N| \cdot \left(\frac{2}{3}\right)^{2N} \underset{N\to+\infty}{\sim} N\left(\frac{2N}{\pi}\right)^{3/2}\left(\frac{2}{3}\right)^{2N} \underset{N\to+\infty}{\to} 0,$$

where γ_N is given by (26)*, and $|H_N|$ is the estimation* (29) *of the size of $\mathrm{CS}_N \setminus M_N$. In particular*

$$\lim_{N\to+\infty} \sigma_{\min}\Big((A_N^{2N})^\intercal A_N^{2N}\Big) = 0. \tag{30}$$

(ii) In the case where A_N^D is injective, the condition number of A_N^D, denoted by $\mathrm{cond}(A_N^D)$, satisfies

$$\mathrm{cond}(A_N^D)^2 \ge \frac{\bar{N}}{|H_N|} \cdot \frac{1}{4\pi\gamma_N^2} \cdot \left(\frac{3}{2}\right)^{2N} \underset{N\to+\infty}{\sim} \frac{1}{4}\Big(\frac{\pi}{2N}\Big)^{1/2}\Big(\frac{3}{2}\Big)^{2N+1} \underset{N\to+\infty}{\to} +\infty,$$

where $\bar{N} = 6N^2 + 2$. In particular,

$$\lim_{N\to+\infty} \mathrm{cond}\left((A_N^{2N})^\intercal A_N^{2N}\right) = +\infty. \tag{31}$$

Proof. (i) Let $D \ge 2N$ be fixed. Consider first the case $\bar{N} < (D+1)^2$, the matrix A_N^D cannot have full column rank. In this case $\sigma_{\min}(A_N^D) = 0$ and the result is obvious. Next consider the case $\bar{N} \ge (D+1)^2$. Then $\sigma_{\min}(A_N^D)^2$ is the smallest eigenvalue of the symmetric matrix $A_N^{D\intercal}A_N^D$. It is expressed as the minimum Rayleigh quotient

$$\sigma_{\min}(A_N^D)^2 = \inf_{\substack{\hat{f}\in\mathbb{R}^{(D+1)^2} \\ \|\hat{f}\|=1}} \Big(\hat{f}^\intercal A_N^{D\intercal} A_N^D \hat{f}\Big).$$

With the Fourier-like expansion (13), we obtain $\hat{f}^\intercal A_N^{D\intercal} A_N^D \hat{f} = \|A_N^D \hat{f}\|^2 = \sum_{\varkappa \in \mathrm{CS}_N} f(\varkappa)^2$, so that

$$\sigma_{\min}(A_N^D)^2 = \inf_{\substack{f \in \mathcal{Y}_D \\ \|f\|_{L^2(\mathbb{S}^2)}=1}} \left(\sum_{\varkappa \in \mathrm{CS}_N} f(\varkappa)^2 \right).$$

Let f_N be the function defined in (20); then f_N is a rotation of the unitary function $Y_{2N}^{-2N} \in \mathcal{Y}_D$, so $f_N \in \mathcal{Y}_D$ with $\|f\| = 1$, which proves that

$$\sigma_{\min}(A_N^D)^2 \leq \sum_{\varkappa \in \mathrm{CS}_N} f(\varkappa)^2.$$

Using $\mathrm{CS}_N = (\mathrm{CS}_N \cap M_N) \sqcup (CS_N \cap H_N)$ and that $f_N \equiv 0$ on M_N, we deduce

$$\sigma_{\min}(A_N^D)^2 \leq \sum_{\varkappa \in \mathrm{CS}_N \cap H_N} f(\varkappa)^2.$$

If $H_N = \emptyset$, we have $\sigma_{\min}(A_N^D)^2 = 0$ and (i) is proved. Otherwise,

$$\sigma_{\min}(A_N^D)^2 \leq |H_N| \max_{\varkappa \in H_N} f(\varkappa)^2,$$

where $|H_N|$ is given by (29); Using (25), we have

$$\max_{\varkappa \in H_N} f(\varkappa)^2 \leq \gamma_N^2 c^{2N}, \quad \text{with} \quad c = \max_{\varkappa \in H_N} \cos^2 \theta(\varkappa).$$

For any $\varkappa = \frac{1}{(1+u^2+v^2)^{1/2}}(u, v, \pm 1) \in H_N$, with $|u|, |v| < 1$, the latitude angle $\theta(\varkappa)$ is such that $\cos^2 \theta(\varkappa) = 1 - \sin^2 \theta(\varkappa) = 1 - \frac{1}{1+u^2+v^2} < \frac{2}{3}$, which proves that $c < \frac{2}{3}$.

(ii) If A_N^D is injective, the condition number is the ratio

$$\mathrm{cond}(A_N^D) = \frac{\sigma_{\max}(A_N^D)}{\sigma_{\min}(A_N^D)},$$

where $\sigma_{\min}(A_N^D)$ has been bounded from below in (i), and $\sigma_{\max}(A_N^D)$ denotes the largest singular value of A_N^D. The square $\sigma_{\max}(A_N^D)^2$ is the largest eigenvalue of $A_N^{D\intercal} A_N^D$ and it is the maximum Rayleigh ratio

$$\sigma_{\max}(A_N^D)^2 = \sup_{\substack{\hat{f} \in \mathbb{R}^{(D+1)^2} \\ \|\hat{f}\|=1}} \left(\hat{f}^\intercal A_N^{D\intercal} A_N^D \hat{f} \right) = \sup_{\substack{f \in \mathcal{Y}_D \\ \|f\|_{L^2(\mathbb{S}^2)}=1}} \sum_{\varkappa \in \mathrm{CS}_N} f(\varkappa)^2.$$

With the particular choice $f(x) = Y_0^0(x) = \frac{1}{\sqrt{4\pi}}$ we obtain the lower bound $\sigma_{\max}(A_N^D)^2 \geq \frac{\bar{N}}{4\pi}$. □

4. Structure of the matrix $(A_N^D)^\intercal \Omega_N A_N^D$

In this section, we consider the problem (WLS) and the matching quadrature rule (6). Recall that the matrix attached to (WLS) is the matrix $A_N^{D\intercal} \Omega_N A_N^D$ in (11). We show in Theorem 8 below how close to an orthonormal system the set of functions (Y_n^m) is, for D a fixed integer. Next, Sec. 4.2 considers the particular case where the weight function $\omega(\varkappa)$ has the cubic symmetry. In this case, a suitable ordering of the indices n and m leads to a particular block diagonal structure of the matrix $A_N^{D\intercal} \Omega_N A_N^D$, which is fully specified.

4.1. Least squares and quadrature rule accuracy.

Suppose the integer D fixed and consider the least squares problem (WLS) in Section 2. Proving the well posedness of (WLS) amounts to establish bounds for the condition number of the matrix $A_N^{D\intercal} \Omega_N A_N^D$. We have

$$A_N^{D\intercal} \Omega_N A_N^D = \left[\sum_{\varkappa \in \mathrm{CS}_N} \omega(\varkappa) Y_n^m(\varkappa) Y_{n'}^{m'}(\varkappa) \right]_{\substack{|m| \leq n \leq D \\ |m'| \leq n' \leq D}} \in \mathbb{R}^{(D+1)^2 \times (D+1)^2}. \tag{32}$$

This matrix contains inner products involving the gridfunctions $(Y_n^m)^*$, for the discrete weighted inner product defined by

$$(y_1, y_2)_\omega \triangleq \sum_{\varkappa \in \mathrm{CS}_N} \omega_N(\varkappa) y_1(\varkappa) y_2(\varkappa), \quad y_1, y_2 : \mathrm{CS}_N \to \mathbb{R}. \tag{33}$$

The functions Y_n^m are orthonormal for the inner product $\langle \cdot, \cdot \rangle_{L^2(\mathbb{S}^2)}$. However, for the discrete product $(.,.)_\omega$, we only have $(Y_n^m, Y_{n'}^{m'})_\omega \approx 0$. Let E_N^D be the symmetric matrix defined by

$$E_N^D = \left[e_N(Y_n^m Y_{n'}^{m'}) \right]_{\substack{|m| \le n \le D \\ |m'| \le n' \le D}} \in \mathbb{R}^{(D+1)^2 \times (D+1)^2}. \tag{34}$$

The entries of the matrix E_N^D are the quadrature errors of the products $Y_n^m Y_{n'}^{m'}$.

Theorem 8. *Fix $N \ge 1$ and $D \ge 0$. Let A_N^D be the Vandermonde matrix with degree D on CS_N, defined in (8). Let $\omega : \mathrm{CS}_N \to (0, \infty)$ be the weight of a spherical quadrature rule on CS_N, with error e_N; let Ω_N be the associated diagonal matrix, defined in (9). Then*

$$A_N^{D\intercal} \Omega_N A_N^D = \mathrm{I}_{(D+1)^2} - E_N^D. \tag{35}$$

In particular, assume that $(\omega_N)_{N \ge 1}$ is a sequence of weight functions defining a convergent quadrature rule on $\mathcal{Y}_{2D}$, i.e. *$\forall f \in \mathcal{Y}_{2D}$, $e_N(f) \xrightarrow[N \to \infty]{} 0$, then*

$$A_N^{D\intercal} \Omega_N A_N^D \xrightarrow[N \to \infty]{} \mathrm{I}_{(D+1)^2}.$$

Moreover, if the rule converges with order $p > 0$, i.e. *$\forall f \in \mathcal{Y}_{2D}$, $\exists C_f \ge 0$, $\forall N \ge 1$, $|e_N(f)| \le C_f N^{-p}$, then*

$$A_N^{D\intercal} \Omega_N A_N^D = \mathrm{I}_{(D+1)^2} + \mathcal{O}\left(\frac{1}{N^p} \right).$$

The relation (35) expresses the fact that the matrix $A_N^{D\intercal} \Omega_N A_N^D$ is close to the identity, assuming that the error matrix entries are small. This is in particular the case when ω defines an accurate quadrature rule on the space $\mathcal{Y}_{2D}$. This will require that D is not too large compared to N. On the contrary, for large values of D, the entries in E_N^D are not a priori small.

Proof. In the matrix $A_N^{D\intercal} \Omega_N A_N^D$, the entry with row index (n, m) and column index (n', m') contains the discrete inner product $\left((Y_n^m)^*, (Y_{n'}^{m'})^* \right)_{\omega_N}$, as described in (32) and (33). Using the quadrature rule (6) with $g = Y_n^m Y_{n'}^{m'}$ shows that this element is expressed as

$$\left((Y_n^m)^*, (Y_{n'}^{m'})^* \right)_{\omega_N} = \int_{\mathbb{S}^2} Y_n^m(x) Y_{n'}^{m'}(x) \, \mathrm{d}\sigma - e_N(Y_n^m Y_{n'}^{m'}).$$

Since the family $(Y_n^m)_{0 \le |m| \le n \le D}$ is orthonormal in $L^2(\mathbb{S}^2)$, we have

$$\int_{\mathbb{S}^2} Y_n^m(x) Y_{n'}^{m'}(x) \, \mathrm{d}\sigma = \left\langle Y_n^m, Y_{n'}^{m'} \right\rangle_{L^2(\mathbb{S}^2)} = \begin{cases} 1, & \text{if } (n, m) = (n', m'), \\ 0, & \text{otherwise.} \end{cases}$$

This proves (35). The symmetry of E_N^D is obvious.

Finally for a convergent rule, for all $|m| \le n \le D$ and $|m'| \le n' \le D$, the entry of E_N^D with indices (n, m) and (n', m') is related to $f = Y_n^m Y_{n'}^{m'} \in \mathcal{Y}_{2D}$, so that by hypothesis $e_N(Y_n^m Y_{n'}^{m'}) \to 0$. For a convergence of order $p > 0$, there is furthermore a constant $C_{n,m}^{n',m'}$ such that $|e_N(Y_n^m Y_{n'}^{m'})| \le C_{n,m}^{n',m'} N^{-p}$. □

4.2. **Block structure of $(A_N^D)^\intercal \Omega_N A_N^D$ for a symmetric weight function.** The weight function $\varkappa \in \mathrm{CS}_N \mapsto \omega(\varkappa)$ plays the role of a parameter in the problem (WLS). Here we consider the particular case where $\omega(\varkappa)$ has the cubic symmetry. This property has been considered in [5, 10].

Theorem 9. *Assume that $\omega : \mathrm{CS}_N \to (0, \infty)$ is invariant under the symmetry group $\mathcal{G}$ of the cube $\{-1, 1\}^3$. Consider a nonzero entry $e_N(Y_n^m Y_{n'}^{m'})$ in the matrix E_N^D defined in (34), with row index (n, m), and column index (n', m'). Then the following conditions hold*

(i) $n \equiv n'\,(2)$ *(same parity for the degrees);*
(ii) $m, m' \geq 0$, *or* $m, m' < 0$ *(same sign for the orders);*
(iii) $m \equiv m'\,(2)$ *(same parity for the orders);*
(iv) if $m, m' \equiv 0\,(2)$, *then* $m \equiv m'\,(4)$.

Proof. The principle of the proof is close to the one of [5, Corollary 10]. The group of the Cubed Sphere coincides with the group of the Cube [3]. In matrix form, it is expressed as

$$\mathcal{G} = \left\{ \begin{bmatrix} \epsilon_1 e_{\sigma_1} & \epsilon_2 e_{\sigma_2} & \epsilon_3 e_{\sigma_3} \end{bmatrix}, \sigma \in \mathfrak{S}_3, \epsilon \in \{-1,1\}^3 \right\},\ e_1 = (1,0,0),\ e_2 = (0,1,0),\ e_3 = (0,0,1), \tag{36}$$

Therefore, the quadrature error defines a linear form

$$e_N : \mathcal{Y}_{2D} \to \mathbb{R}, \quad e_N(g) = \int_{\mathbb{S}^2} g(x)\,\mathrm{d}\sigma - \sum_{x \in \mathrm{CS}_N} \omega(x) g(x),$$

which is invariant under $\mathcal{G}$, *i.e.* $\forall Q \in \mathcal{G}$, $e_N(g(Q^\intercal \cdot)) = e_N(g)$. In the sequel, for all (n,m) and (n',m') violating at least one of the conditions *(i)-(iv)* in Theorem 9, we consider $g = Y_n^m Y_{n'}^{m'} \in \mathcal{Y}_{2D}$, and we exhibit a matrix $Q \in \mathcal{G}$ satisfying $g(Q^\intercal x) = -g(x)$. This is a sufficient condition to ensure that $e_N(g) = 0$, due to $e_N(g) = e_N(g(Q^\intercal \cdot)) = e_N(-g) = -e_N(g)$. The proof is a calculation in spherical coordinates, based on the expression

$$\begin{aligned} g(x(\theta,\phi)) = (q_n^m q_{n'}^{m'})(\sin\theta) \cdot \cos^{|m|}\theta \cos^{|m'|}\theta \\ \cdot (\sin(m\phi)\mathbf{1}_{m<0} + \cos(m\phi)\mathbf{1}_{m\geq 0})(\sin(m'\phi)\mathbf{1}_{m'<0} + \cos(m'\phi)\mathbf{1}_{m'\geq 0}). \end{aligned}$$

Case 1: (ii) is violated. Assume $m < 0$ and $m' \geq 0$ (without loss of generality), then

$$g(Q^\intercal x(\theta,\phi)) = g(x(\theta,-\phi)) = -g(x(\theta,\phi)), \quad \text{for} \quad Q := \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Case 2: (iii) is violated. Assume that $m \equiv 1\,(2)$ and $m' \equiv 0\,(2)$ (without loss of generality). Then $m(\phi+\pi) \equiv m\phi + \pi\,(2\pi)$, $m'(\phi+\pi) \equiv m'\phi\,(2\pi)$, and

$$g(Q^\intercal x(\theta,\phi)) = g(x(\theta,\phi+\pi)) = -g(x(\theta,\phi)), \quad \text{for} \quad Q := \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Case 3: (iii) is satisfied but (i) is violated. Assume that $n + |m| \equiv 1\,(2)$ and $n' + |m'| \equiv 0\,(2)$ (without loss of generality). Then $\theta \mapsto (q_n^m q_{n'}^{m'})(\sin\theta)$ is odd, hence

$$g(Q^\intercal x(\theta,\phi)) = g(x(-\theta,\phi)) = -g(x(\theta,\phi)), \quad \text{for} \quad Q := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Case 4: (iv) is violated. Assume that $m \equiv 2\,(4)$ and $m' \equiv 0\,(4)$ (without loss of generality). Then $m(\phi+\frac{\pi}{2}) \equiv m\phi + \pi\,(2\pi)$, $m'(\phi+\frac{\pi}{2}) \equiv m'\phi\,(2\pi)$, and

$$g(Q^\intercal x(\theta,\phi)) = g(x(\theta,\phi+\tfrac{\pi}{2})) = -g(x(\theta,\phi)), \quad \text{for} \quad Q := \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad \square$$

Roughly speaking, if the weight function ω is symmetric, then at most a percentage of

$$100 \times \frac{3}{32} (= 9.375\%) \tag{37}$$

of all the entries in E_N^D are nonzero. Indeed, Case *(i)* divides by 2 the number of possible nonzero entries. Then Case *(ii)* further divides by 2 this number. And finally Cases *(iii-iv)* multiply this number by $\frac{3}{8}$. At this point, two facts suggest the approximation

$$A_N^{D\intercal} \Omega_N A_N^D \approx \mathrm{I}_{(D+1)^2} :$$

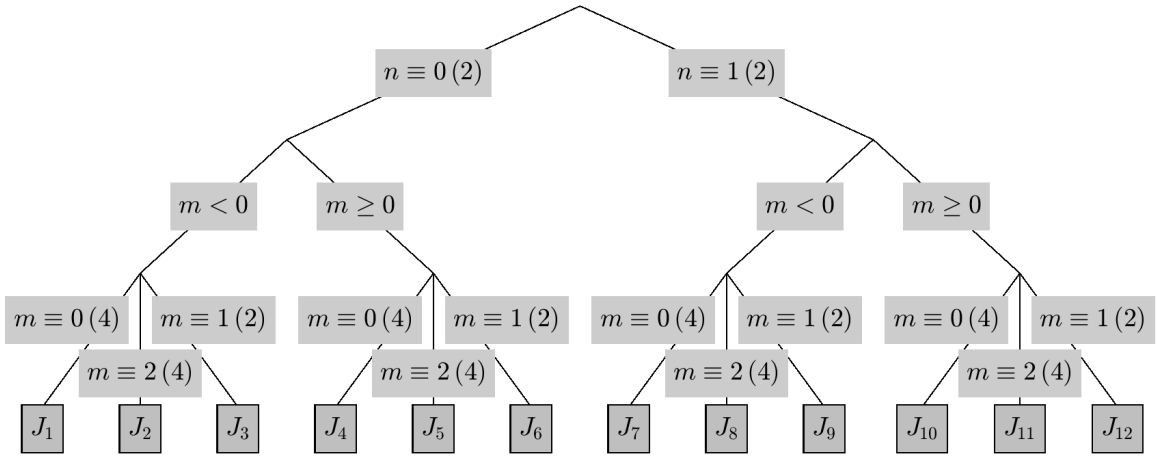

FIGURE 2. Classification tree for partioning the set of indices $\{(n,m) : |m| \leq n \leq D\}$ as a disjoint union $J_1 \sqcup \ldots \sqcup J_{12}$; for instance, $J_4 = \{|m| \leq n \leq D,\, n \equiv 0\,(2),\, m \geq 0,\, m \equiv 0\,(4)\}$.

- an approximate ratio of $\frac{29}{32}$ of all entries are zero if the weight function ω is assumed symmetric (Theorem 9);
- the remaining entries (approximate ratio of $\frac{3}{32}$) are small assuming that ω defines an accurate spherical quadrature rule in $\mathcal{Y}_{2D}$, (see Theorem 8).

In particular, the condition number of the matrix $A_N^{D\intercal}\Omega_N A_N^D$ is expected to be close to 1, so that (WLS) is expected to be well-posed. This point is further investigated numerically in Section 5.3.

Next, we go one step further in the analysis taking benefit from the orthogonality relations in Theorem 9. Indeed, Theorem 9 suggests to sort the indices (n,m) using the following criteria, ordered by decreasing priority:

- case $n \equiv 0\,(2)$ and case $n \equiv 1\,(2)$;
- case $m < 0$ and case $m \geq 0$;
- case $m \equiv 0\,(4)$, then case $m \equiv 2\,(4)$, and finally case $m \equiv 1\,(2)$.

This particular ordering expresses the set of indices as a disjoint union of the twelve sets J_k, $1 \leq k \leq 12$. Fig. 2 displays the resulting classification tree. It is an expression of the orthogonality relations in Theorem 9.

Corollary 10. *Fix $N \geq 1$, $D \geq 0$. Fix a weight function $\omega : \mathrm{CS}_N \to (0,\infty)$ invariant under the group $\mathcal{G}$ of $\{-1,1\}^3$ in (36). Let J_k, $1 \leq k \leq 12$, denotes a partioning of the set of indices $|m| \leq n \leq D$, displayed in Fig. 2.*

(i) Assume that the indices $(n,m) \in J_k$, $1 \leq k \leq 12$ in the Vandermonde matrix A_N^D are sorted along increasing k (for the rows and for the columns). Then $A_N^{D\intercal}\Omega_N A_N^D$ is block diagonal, as shown in Fig. 3.

(ii) The following orthogonal decomposition holds for the discrete inner product (33)*,*

$$\big\{f^*,\, f \in \mathcal{Y}_D\big\} = \bigoplus_{k=1}^{12} \mathrm{Span}\{(Y_n^m)^*,\, (n,m) \in J_k\}.$$

Assuming a symmetric weight function ω, Corollary 10 reveals that the matrix $A_N^{D\intercal}\Omega_N A_N^D$, associated to (WLS), is block diagonal for a particular ordering of the indices. This has the following consequence to solve the system (11). Instead of solving a linear system with $(D+1)^2$ unknowns, the system is solved by blocks. It consists in solving 8 square linear systems with approximately $\frac{1}{16}(D+1)^2$ unknowns, and 4 square linear systems with approximately $\frac{1}{8}(D+1)^2$ unknowns. These resolutions can be obviously performed in parallel.

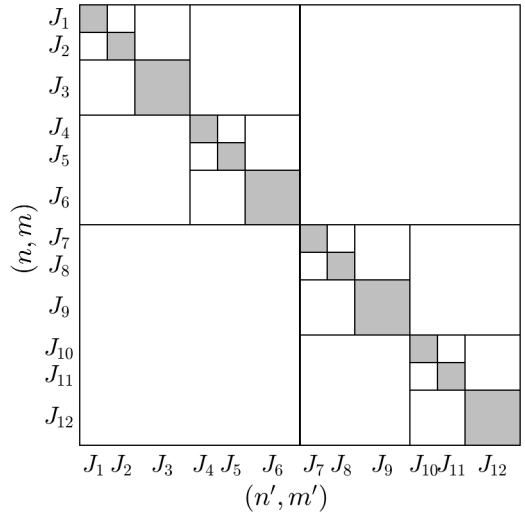

FIGURE 3. Block diagonal structure of the matrix $A_N^{D\intercal}\Omega_N A_N^D$, assuming that ω is invariant under $\mathcal{G}$; the sets of indices J_k are defined in Figure 2. The white cells contains only null coefficients; they represent an approximate ratio of $\frac{29}{32}$ of the entries.

Remark 11. In [3, Corollary 10], the "15/16" property is coined as meaning exactness of a symmetric quadrature rule for a certain proportion of 15/16 of all spherical harmonics. This property can be deduced from Corollary 10. Indeed, consider a symmetric weight function ω, a row index $(n, m) \notin J_4$, and the column index $(n', m') = (0, 0) \in J_4$; then $Y_{n'}^{m'} \equiv (4\pi)^{-1/2}$, and we deduce from Corollary 10.(i) that

$$\mathcal{Q}(Y_n^m) = \sum_{\chi \in \mathrm{CS}_N} \omega(\chi) Y_n^m(\chi) = 0 = \int_{\mathbb{S}^2} Y_n^m(x)\, \mathrm{d}\sigma.$$

Since J_4 contains about 1/16 of the indices, we see that the quadrature rule $\mathcal{Q}$ associated to ω exactly integrates an approximate proportion of 15/16 of all the Y_n^m.

Remark 12. In [5], a comparison of cubic symmetric rules of the form (6) with optimal Lebedev rules has been presented. Other rules with symmetries are the *level symmetric rules* in usage in computational neutron transport [9]. A comparison with this family of rules, where the nodes lie along circles, could be of interest.

5. Numerical results

5.1. **Condition number of the Vandermonde matrix.** In this section, we assess numerically that the problem (LS) is well-posed for the degree $D = 2N - 1$, but not for $D = 2N$. We proceed as follows. For all $1 \le N \le 32$, with $D = 2N - 1$ and $D = 2N$, we first compute a singular value decomposition of the Vandermonde matrix A_N^D in (8). Second, we extract the minimal singular value $\sigma_{\min}(A_N^D)$ and the maximal one $\sigma_{\max}(A_N^D)$. Then the condition number $\mathrm{cond}(A_N^D) = \sigma_{\max}(A_N^D)/\sigma_{\min}(A_N^D)$ is evaluated. The computation has been performed in double precision in `MATLAB`, using the `svd` function. The results in Fig. 4 are as follows

(1) For $D = 2N - 1$ (left panel in Fig. 4), we observe that the minimal singular value is "far" from 0, and that $\mathrm{cond}(A_N^D) \approx 1.19$ is close to 1; This is a numerical indication that the matrix A_N^{2N-1} is injective, and therefore that the critical degree D in (15) for injectivity of A_N^D is such that $D \ge 2N-1$. One also observes that $\mathrm{cond}(A_N^{D\intercal} A_N^D) = \mathrm{cond}(A_N^D)^2 \approx 1.41$. This suggests that the problem (LS) is well-posed for $D = 2N - 1$.
(2) For $D = 2N$ (right panel in Fig. 4), $\sigma_{\min}(A_N^D)$ is observed to be close to 0 for $N \in \{1, 2, 3, 4, 5, 7, 9\}$ (the machine epsilon is about $2.2 \cdot 10^{-16}$); for $N \ge 10$, it is positive and

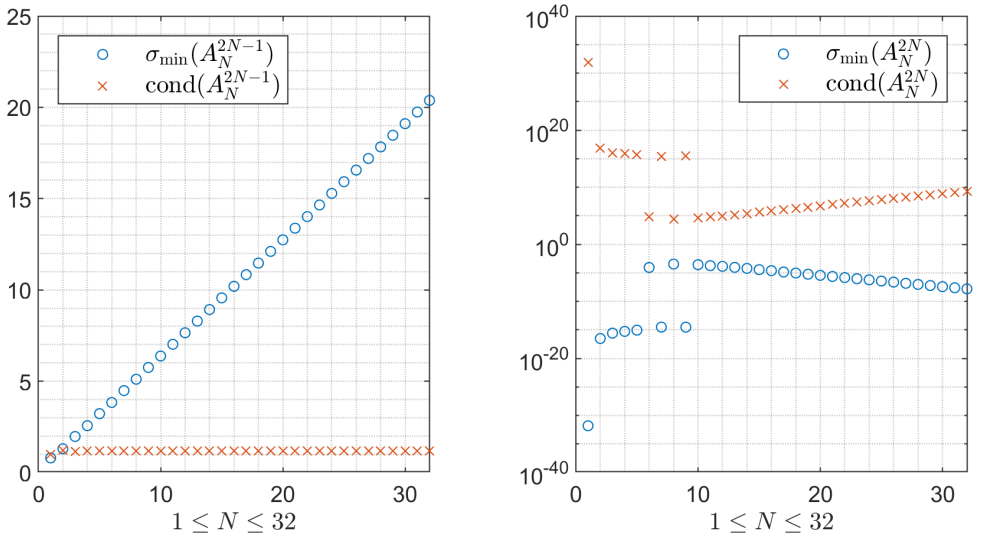

FIGURE 4. Smallest singular value ($\sigma_{\min}$) and condition number (cond) of the Vandermonde matrices A_N^D, $1 \le N \le 32$, with $D = 2N-1$ (left panel), and $D = 2N$ (right panel, in log-scale). Left panel: A_N^{2N-1} is injective ($\sigma_{\min} \gg 0$) and well-conditioned (cond ≈ 1.19). Right panel: A_N^{2N} is observed to be not numerically injective if N is small ($\sigma_{\min} \approx 0$), and is ill-conditioned otherwise (cond $> 10^4$).

decays to 0 when N increases. Hence, for $N \in \{1,2,3,4,5,7,9\}$, A_N^{2N} is not injective. This suggests $D_N \le 2N-1$. This is consistent with Prop. 4 which proves the result for $N \le 4$. This numerical observation, combined with the discussion above, supports the fact that

$$D_N = 2N-1, \quad N \in \{1,2,3,4,5,7,9\}.$$

For the other values of N, it is numerically apparent that A_N^{2N} is injective. Nevertheless, for these values of N, $\mathrm{cond}(A_N^{2N}) > 10^4$, and blows-up when N increases. This implies that $\mathrm{cond}(A_N^{2N\intercal} A_N^{2N}) > 10^8$ and blows-up as well. Therefore, for $D = 2N$, these numerical results are in agreement with the theoretical result in Theorem 7 and indicates that the ill-posedness of (LS) is true in all cases. In addition, the ill-posedness level increases with N.

The numerical study above suggests that for any $N \ge 1$, the value of D_N in Property *(P)* in (16) is $D_N = 2N-1$. This in particular means that any $f \in \mathcal{Y}_{2N-1}$ is correctly sampled on the Cubed Sphere CS_N with angular step $\frac{\pi}{2N}$, since (LS) can reconstruct f from f^* in a stable way. If $f \in \mathcal{Y}_D$ with $D \ge 2N$, this property is not guaranteed.

5.2. **Accuracy of the least squares approximation.** The results in Section 3 assess the fact that $\mathcal{Y}_{2N-1}$ is the largest SH subspace leading to well-posedness and well-conditonnining or the interpolation problem (LS). Here we further assess this property by evaluating the accuracy of least-squares approximations of a series of test functions.

First, we report in Table 1 five functions (see [3] and the references therein); this series of functions is representative of various regularity properties. For each $1 \le N \le 32$ and for each test function f_i, $1 \le i \le 5$, we compute the least-squares approximation $\tilde{f}_i \in \mathcal{Y}_{2N-1}$ of f_i from the grid function $(f_i)^*$: $\tilde{f}_i$ is evaluated as the unique solution to (LS), for $D = 2N-1$ and $y = (f_i)^*$. The accuracy is measured by the relative discrete error, on a fixed fine grid CS_M. We have chosen $M = 65$. The relative error is defined by

$$\epsilon(f_i) \triangleq \left(\frac{\sum_{x \in \mathrm{CS}_M} |f_i(x) - \tilde{f}_i(x)|^2}{\sum_{x \in \mathrm{CS}_M} |f_i(x)|^2} \right)^{1/2}, \quad M = 65. \tag{38}$$

i	$f_i(x,y,z)$	Comment
1	$\exp(x)$	Very smooth
2	$\frac{3}{4}\exp[-\frac{(9x-2)^2}{4}-\frac{(9y-2)^2}{4}-\frac{(9z-2)^2}{4}]$ $+\frac{3}{4}\exp[-\frac{(9x+1)^2}{49}-\frac{9y+1}{10}-\frac{9z+1}{10}]$ $+\frac{1}{2}\exp[-\frac{(9x-7)^2}{4}-\frac{(9y-3)^2}{4}-\frac{(9z-5)^2}{4}]$ $-\frac{1}{5}\exp[-(9x-4)^2-(9y-7)^2-(9z-5)^2]$	Smooth
3	$\frac{1}{10}\frac{\exp(x+2y+3z)}{(x^2+y^2+(z+1)^2)^{1/2}}\mathbf{1}(z>-1)$	Infinite spike at the south pole ($z=-1$)
4	$\cos(3\arccos z)\mathbf{1}(3\arccos z\le\frac{\pi}{2})$	Continuous, not differentiable ($z=\frac{\sqrt{3}}{2}$)
5	$\mathbf{1}(z\ge\frac{1}{2})$	Discontinuous spherical cap ($z=\frac{1}{2}$)

TABLE 1. A series of test functions representative of various regularity properties.

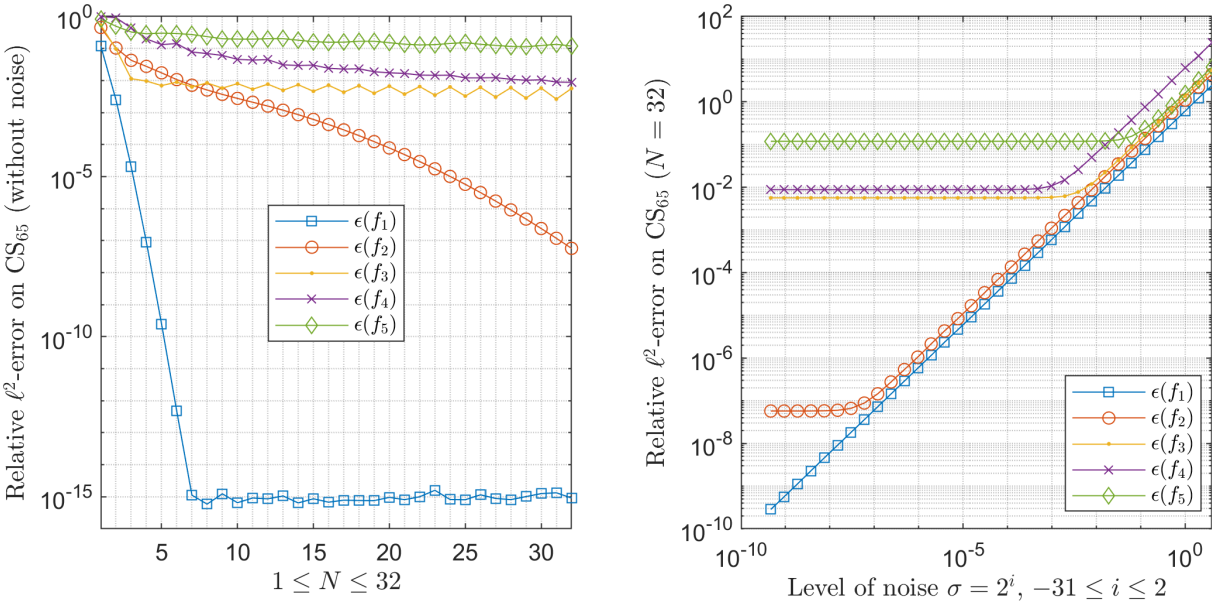

FIGURE 5. Least-squares approximation (LS) of the test functions f_i in Table 1. Left panel: for any $1\le N\le 32$, the approximation $\tilde{f}_i\in\mathcal{Y}_{2N-1}$ is computed from $(f_i)^*$, and the relative ℓ^2-error $\epsilon(f_i)$ defined in (38) is plotted. Right panel: for any level of noise $\sigma=2^j, -31\le j\le 2$, the approximation $\tilde{f}_i\in\mathcal{Y}_{63}$ is computed from a noisy dataset $(f_i)^*+\sigma\mathcal{N}(0,1)$ with $N=32$, and $\epsilon(f_i)$ is plotted.

The errors $\epsilon(f_i)$ are displayed in Fig. 5 (left panel). For the smooth functions f_1 and f_2, the error rapidly converges to 0 when N increases; this is especially true for f_1. For the continuous but not differentiable function f_4, the convergence is slow. For the spike function f_3 and the discontinuous function f_5, the convergence cannot be claimed from the plot. These observations are not surprising: it is expected that the convergence rate depends on the decay of the Fourier coefficients, which is related to the smoothness.

Second, fix the grid resolution to $N=32$. For each test function f_i, $1\le i\le 5$, for any $\sigma=2^j$, $-31\le j\le 2$, we corrupt the grid function $(f_i)^*$ with a gaussian noise with zero mean, and standard deviation σ. We compute an approximation $\tilde{f}_i\in\mathcal{Y}_{63}$ of f_i as the unique solution to (LS), for $D=2N-1$ and $y(\mathfrak{x})=f_i(\mathfrak{x})+\sigma u(\mathfrak{x})$, $\mathfrak{x}\in CS_{32}$, where $[u(\mathfrak{x})]$ contains independent realizations of the normal law $\mathcal{N}(0,1)$. Here again, we evaluate the accuracy of this approximation by the relative error (38); this error depends on σ (and on the experiment), and we denote it by $\epsilon(f_i)(\sigma)$. These errors are displayed in Figure 5 (right panel). One observes that

$$\epsilon(f_i)(\sigma)\approx\epsilon(f_i)(0)+\sigma,$$

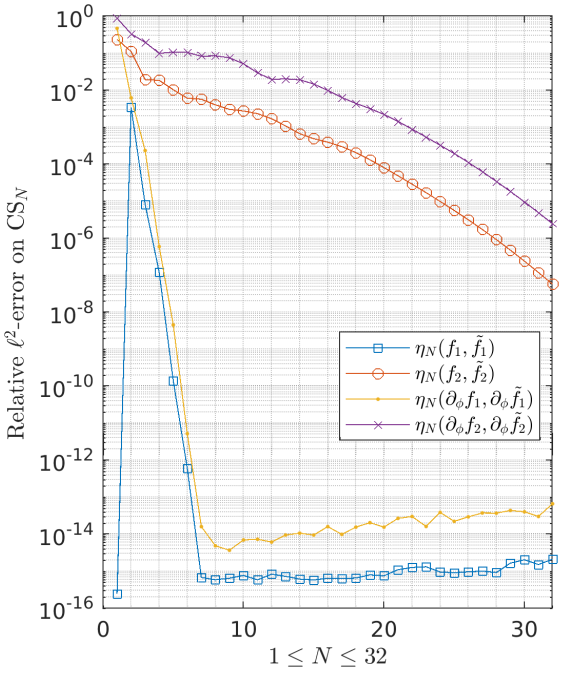

FIGURE 6. Spectral differentiation on CS_N with respect to the longitude angle ϕ. For $f = f_1,\ f_2$ from Table 1, for any $1 \leq N \leq 32$, the approximate derivative $\partial_\phi \tilde{f}$ is computed from the least-squares approximation $\tilde{f} \in \mathcal{Y}_{2N-1}$. Relative ℓ^2-errors defined in (41) are plotted.

where $\epsilon(f_i)(0)$ is the error without noise for $N = 32$ (displayed on the left panel). In other words, a level of noise σ in the dataset increases the error by σ. This reveals that approximating a function by least-squares on CS_N in the space $\mathcal{Y}_{2N-1}$ is very stable.

Third, we show numerically that differentiating the least-squares approximation (LS) in $\mathcal{Y}_{2N-1}$ ($D = 2N-1$) permits to approximate derivatives. Assume that f is a differentiable function on $\mathbb{S}^2$, known by the grid function f^*. The least squares approximation (LS) of f with $D = 2N-1$ is

$$\tilde{f} = \sum_{|m| \leq n \leq 2N-1} \tilde{f}_n^m Y_n^m \in \mathcal{Y}_{2N-1}, \tag{39}$$

and $y = f^*$. Consider for instance the derivative with respect to the longitude ϕ,

$$\partial_\phi \tilde{f}(x(\theta,\phi)) = \sum_{|m| \leq n \leq 2N-1} m \cdot \tilde{f}_n^{-m}\, Y_n^m(x(\theta,\phi)) \in \mathcal{Y}_{2N-1}. \tag{40}$$

We test this principle on the smooth functions defined in Table 1: $f = f_1,\ f_2$. For each value of $1 \leq N \leq 32$, we approximate $\partial_\phi f$ by $\partial_\phi \tilde{f}$ satisfying (40), and we calculate the relative ℓ^2-errors on the grid CS_N:

$$\eta_N(f, \tilde{f}) = \left(\frac{\sum_{x \in \mathrm{CS}_N} |f(x) - \tilde{f}(x)|^2}{\sum_{x \in \mathrm{CS}_N} |f(x)|^2} \right)^{1/2}, \ \eta_N(\partial_\phi f, \partial_\phi \tilde{f}) = \left(\frac{\sum_{x \in \mathrm{CS}_N} |\partial_\phi f(x) - \partial_\phi \tilde{f}(x)|^2}{\sum_{x \in \mathrm{CS}_N} |\partial_\phi f(x)|^2} \right)^{1/2}; \tag{41}$$

here, the exact derivative is given by

$$\partial_\phi f(x(\theta,\phi)) = -x_2(\theta,\phi)\partial_{x_1} f(x(\theta,\phi)) + x_1(\theta,\phi)\partial_{x_2} f(x(\theta,\phi),$$

where $x_1(\theta,\phi)$ and $x_2(\theta,\phi)$ denote the horizontal coordinates of $x(\theta,\phi)$. As can be observed in Figure 6, the error for the derivative and the error for the function itself have a similar behavior in function of N. The least squares approximation converges to the exact function and the spectral derivative converges to the exact derivative; the observed convergence rates are similar.

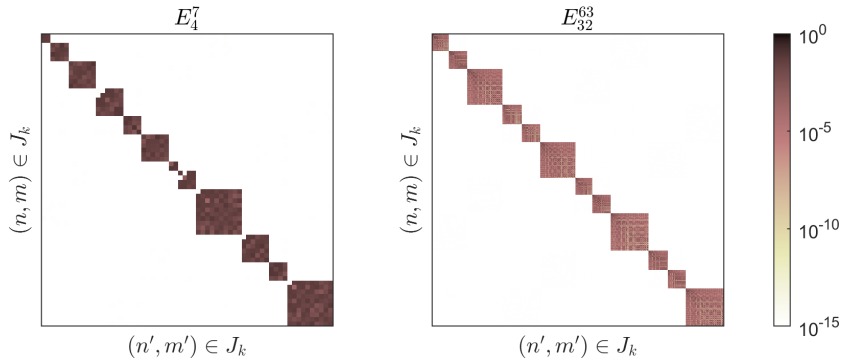

FIGURE 7. Matrix $E_N^D = [e_N(Y_n^m Y_{n'}^{m'})]$ from (35)-(34), with the uniform weight $\omega = 4\pi/\bar{N}$, $D = 2N-1$, $N = 4$ (left panel) and $N = 32$ (right panel). The indices are arranged by the classification tree of Figure 2. The displayed value is $10^{-15} + |e_N(Y_n^m Y_{n'}^{m'})|$, in logarithmic scale. The observed structure is the block diagonal structure predicted by Figure 3 (Corollary 10). The sparsity score is 9.961% (left panel), resp. 9.387% (right panel), which is close to a ratio of 3/32.

5.3. **Pseudo-orthogonality for the discrete inner product.** We evaluate numerically the relation (35); it represents some "pseudo-orthogonality" of the Legendre basis, for the discrete inner product (33).

First, we consider the uniform quadrature rule on CS_N, defined by $\omega(x) = 4\pi/\bar{N}$. In this case, the matrix $A_N^{D\intercal}\Omega_N A_N^D$ in (11) is expressed as

$$A_N^{D\intercal}\Omega_N A_N^D = \frac{4\pi}{\bar{N}} A_N^{D\intercal} A_N^D.$$

The uniform weight is invariant under $\mathcal{G}$. Thus Theorem 9 predicting a sparse structure of $E_N^D = \mathrm{I}_{(D+1)^2} - A_N^{D\intercal}\Omega_N A_N^D$ can be applied. More precisely, Corollary 10 predicts that E_N^D has the block diagonal structure in Figure 3, for a suitable ordering of the indices. Figure 7 reports this structure, where E_N^{2N-1} is displayed for $N = 4$ and $N = 32$. In these matrices, the percentage of coefficients above 10^{-14} is respectively 9.961% and 9.387%, which is close to the ratio (37). Furthermore, we compute the largest entry of E_N^{2N-1} for $1 \leq N \leq 32$. As displayed in Figure 8, this value is about 0.1 (except for $N = 1$, for which the observed value is the machine epsilon). Therefore, the matrix corresponding to (LS) (without weight) is close to be proportional to the identity matrix.

$$A_N^{2N-1\intercal} A_N^{2N-1} \approx \frac{\bar{N}}{4\pi}(\mathrm{I}_{4N^2} \pm 0.1).$$

Second, we consider the weight ω of the trapezoidal rule in the Definition 3.1 in [10]. It is invariant under $\mathcal{G}$, so E_N^D has a sparse structure as before. This rule is second order accurate; so it is more accurate than the uniform one, and the entries of E_N^D are expected to be smaller, due to Theorem 8. This is confirmed in Figure 8; the maximal entry of E_N^{2N-1} is below 0.1, and it decays to zero when N increases.

6. Conclusion

This paper considers weighted least-squares approximation by spherical harmonics on the equiangular Cubed Sphere CS_N. From a theoretical point of view, the symmetric positive semi-definite matrix of the normal equations is expected to be a pertubation of the identity matrix; the magnitude of the pertubation depends on the accuracy of the quadrature rule associated to the weight. This indicates that the Legendre spherical harmonics should be almost orthogonal for some discrete inner product on CS_N. In the case of a symmetrical weight, the matrix is block diagonal; this structure directly provides subspaces of spherical harmonics which are exactly orthogonal for the discrete inner product, disregarding the magnitude of the pertubation.

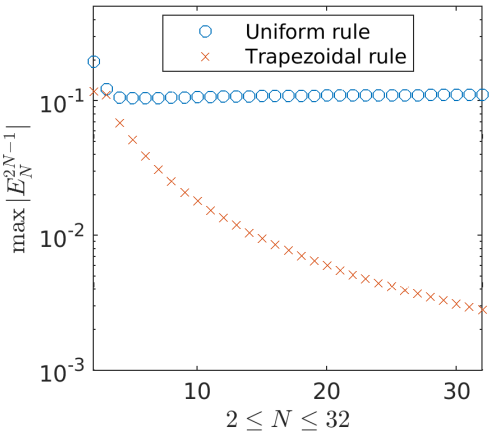

FIGURE 8. Maximal entry $\max |e_N(Y_n^m Y_{n'}^{m'})|$ of the matrix E_N^{2N-1}, $2 \leq N \leq 32$. The result depends on the accuracy of the quadrature rule ω. It is smaller for the trapezoidal weight than for the uniform weight.

From a numerical point of view, the matrix has a condition number close to 1 if the cutoff (angular) frequency is fixed to $2N-1$, whereas it is not anymore the case if higher frequencies are considered. Numerical results indicate that $\mathcal{Y}_{2N-1}$ is a suitable approximation space for fitting or differentiating a smooth function from values on CS_N.

Future work also includes further mathematical analysis on the one hand. On the other hand, the block structure and the well conditioning of the matrix shown in Section 4.2 opens the way to a parallel Conjugate Gradient solver. This is a preliminary step before to investigate a genuine fast solver.

REFERENCES

[1] C. An, X. Chen, I.H. Sloan, and R.S. Womersley. Regularized least squares approximations on the sphere using spherical designs. *SIAM J. Numer. Anal.*, 50(3):1513–1534, 2012.
[2] K. Atkinson and W. Han. *Spherical Harmonics and Approximations on the Unit Sphere: an introduction.* Number 2044 in Lect. Notes. Math. Springer-Verlag, 2012.
[3] J.-B. Bellet. Symmetry group of the equiangular cubed sphere. *Quart. of App. Math.*, 80:69–86, 2022.
[4] J.-B. Bellet, M. Brachet, and J.-P. Croisille. Interpolation on the Cubed Sphere with Spherical Harmonics. *Numer. Math.*, 2022. https://doi.org/10.1007/s00211-022-01340-w.
[5] J.-B. Bellet, M. Brachet, and J.-P. Croisille. Quadrature and symmetry on the Cubed Sphere. *J. Comp. App. Math.*, 409(114142), 2022.
[6] F. Dai and Y. Xu. *Approximation theory and harmonic analysis on spheres and balls.* Springer Mongraphs in Mathematics. Springer-Verlag, 2013.
[7] W. Freeden, M. Z. Nashed, and T. Sonar. *Handbook of Geomathematics.* Springer, 2010.
[8] K. Hesse and Q. T. L. Gia. L^2 error estimates for polynomial discrete penalized least-squares approximation on the sphere from noisy data. *Jour. Comp. App. Math.*, 408(114118), 2022.
[9] E.E. Lewis and W.F. Miller. *Computational methods of neutron transport.* John Wiley & Sons, 1984.
[10] B. Portelenelle and J.-P. Croisille. An efficient quadrature rule on the Cubed Sphere. *J. Comp. App. Math*, 328:59–74, 2018.
[11] M. Rančić, R.J. Purser, and F. Mesinger. A global shallow-water model using an expanded spherical cube: Gnomonic versus conformal coordinates. *Quart. J. Roy. Meteor. Soc.*, 122:959–982, 1996.
[12] C. Ronchi, R. Iacono, and P. S. Paolucci. The Cubed Sphere: A new method for the solution of partial differential equations in spherical geometry. *J. Comput. Phys.*, 124:93–114, 1996.
[13] R. Sadourny. Conservative finite-difference approximations of the primitive equations on quasi-uniform spherical grids. *Mon. Wea. Rev.*, 100:136–144, 1972.

UNIVERSITÉ DE LORRAINE, CNRS, IECL, F-57000 METZ, FRANCE
Email address: jean-baptiste.bellet@univ-lorraine.fr, jean-pierre.croisille@univ-lorraine.fr