Optimal polynomial meshes and Caratheodory-Tchakaloff submeshes

Paul Leopardi

Australian Bureau of Meteorology; University of Melbourne.

For presentation at Tight Frames and Approximation 2018.

Joint work with Alvise Sommariva and Marco Vianello.

Made possible by donations from 31 supporters via Pozible.

February 2018

Acknowledgements

Joint work with Alvise Sommariva and Marco Vianello.

Rob Womersley.

Program on "Minimal Energy Point Sets, Lattices, and Designs" at the Erwin Schrödinger International Institute for Mathematical Physics, 2014.

Australian National University.
University of Newcastle, Australia.
University of Melbourne.

Shayne Waldron.

Pozible supporters (2014)

$32+ SvA, AD, CF, OF, KM, JP, AHR, RR, ES, MT.

$50+ Yvonne Barrett, Angela M. Fearon, Sally Greenaway, Dennis Pritchard, Susan Shaw, Bronny Wright.

$64+ Naomi Cole.

$100+ Russell family, Jonno Zilber.

$128+ Jennifer Lanspeary, Vikram.

. . . and others, who did not want acknowledgement.

Outline of talk

- Polynomial meshes
- Dubiner distance
- The unit sphere
- EQ spherical codes
- Caratheodory-Tchakaloff subsampling
- Numerical results

Polynomials on a compact manifold

For a compact manifold $K \subset \mathbb{R}^d$,

$\mathbb{P}_n^d(K)$ denotes the subspace of d-variate polynomials of total degree not exceeding n restricted to K, and $N_n = N_n(K) = \dim(\mathbb{P}_n^d(K))$ denotes its dimension.

For example $N_n = (n+1)^2$ for the sphere $\mathbb{S}^2$.

Norming meshes

Definition 1

For compact $K \subset \mathbb{R}^d$, and $C(K)$ the space of continuous functions on K, given a sequence of finite dimensional subspaces $P_n(K) \subset C(K)$, a P_n*-norming mesh* is a sequence $(\mathcal{A}_n)$ of finite subsets of K such that

$$\|p\|_{L^\infty(K)} \leqslant c \sup_{z \in \mathcal{A}_n} |p(z)| \quad \text{for all } p \in P_n.$$

(Calvi and Levenberg 2008)

Polynomial meshes

A *polynomial mesh* on a compact manifold $K \subset \mathbb{R}^d$ is a sequence of finite norming subsets $\mathcal{A}_n \subset K$ such that

$$\|p\|_{L^\infty(K)} \leqslant C\,\|p\|_{\ell^\infty(\mathcal{A}_n)}\;,\;\forall p \in \mathbb{P}_n^d(K)\;, \qquad (1)$$

where $M_n = \mathrm{card}(\mathcal{A}_n) = \mathcal{O}(N_n^\beta)$, $\beta \geqslant 1$ and C is a constant independent of n.

Since $\mathcal{A}_n$ is automatically $\mathbb{P}_n^d(K)$-determining, then $M_n \geqslant N_n = \dim(\mathbb{P}_n^d(K)) = \dim(\mathbb{P}_n^d(\mathcal{A}_n))$.

Such a mesh is called *optimal* when $\beta = 1$.

Weakly admissible meshes

When C is substituted by a sequence C_n that increases subexponentially,

$$\|p\|_{L^\infty(K)} \leqslant C_n \, \|p\|_{\ell^\infty(\mathcal{A}_n)} \, , \; \forall p \in \mathbb{P}_n^d(K) \, , \tag{2}$$

in particular when $C_n = \mathcal{O}(n^s)$, $s \geqslant 0$,

we speak of a *weakly admissible polynomial mesh*.

(Calvi and Levenberg 2008)

Properties of polynomial meshes

Polynomial meshes

- contain computable near optimal interpolation sets, and
- are near optimal for uniform Least Squares approximation:

$$\Lambda(\mathcal{A}_n) = \|\mathcal{L}_{\mathcal{A}_n}\| := \sup_{f \in C(K), f \neq 0} \frac{\|\mathcal{L}_{\mathcal{A}_n} f\|_{L^\infty(K)}}{\|f\|_{L^\infty(K)}} \leqslant C\sqrt{M_n}\,, \tag{3}$$

where $\mathcal{L}_{\mathcal{A}_n}$ is the $\ell^2(\mathcal{A}_n)$-orthogonal projection operator $C(K) \to \mathbb{P}_n^d(K)$, from which follows

$$\|f - \mathcal{L}_{\mathcal{A}_n} f\|_{L^\infty(K)} \leqslant \left(1 + C\sqrt{M_n}\right) \min_{p \in \mathbb{P}_n^d(K)} \|f - p\|_{L^\infty(K)}\,. \tag{4}$$

(Bos, Calvi, Levenberg, Sommariva and Vianello 2011; Bos, De Marchi, Sommariva and Vianello 2010)

(Calvi and Levenberg 2008)

Dubiner distance (1)

The *Dubiner distance* in a compact set or manifold is

$$\mathrm{dist}_D(x,y) = \sup_{deg(p)\geqslant 1} \left\{ \frac{1}{deg(p)} \left| \cos^{-1}(p(x)) - \cos^{-1}(p(y)) \right| \right\} , \tag{5}$$

where the sup is taken over the polynomials $p \in \mathbb{P}_n^d(K)$ such that $\|p\|_{L^\infty(K)} \leqslant 1$.

(Dubiner 1995)

Dubiner distance (2)

The Dubiner distance is known analytically only in a very few cases: the interval, the cube, the simplex, the ball, and the sphere. In particular, it can be proved via the classical van der Corput-Schaake inequality that on the sphere it coincides with the usual *geodesic distance*,

$$\mathrm{dist}_D(x,y) = \gamma(x,y) = \cos^{-1}(\langle x,y \rangle)\ , \ \ \forall x,y \in \mathbb{S}^2\ , \tag{6}$$

where $\langle x,y \rangle$ denotes the Euclidean scalar product in $\mathbb{R}^3$.

(Bos, Levenberg and Waldron 2004)

Covering radius inequality

A simple connection of the Dubiner distance with the theory of polynomial meshes is given by the following

Proposition 1

Let $\mathcal{A}_n$ *be a compact subset of a compact set or manifold* $K \subset \mathbb{R}^d$ *whose covering radius with respect to the Dubiner distance does not exceed* θ/n, *where* $\theta \in (0,1)$ *and* $n \geqslant 1$, *i.e.*

$$\forall x \in K \; \exists y \in \mathcal{A}_n : \; dist_D(x,y) \leqslant \frac{\theta}{n} \,. \tag{7}$$

Then the following inequality holds,

$$\|p\|_{L^\infty(K)} \leqslant \frac{1}{1-\theta} \|p\|_{L^\infty(\mathcal{A}_n)} \,, \quad \forall p \in \mathbb{P}_n^d(K) \,. \tag{8}$$

Proof

In view of (7), the proof of Proposition 1 is an immediate consequence of the elementary inequality

$$|p(x)| \leqslant |p(y)| + |p(x) - p(y)| \leqslant |p(y)| + n\,\mathrm{dist}_D(x,y)\,\|p\|_{L^\infty(K)} \,. \tag{9}$$

Note that $\mathcal{A}_n$ need not be discrete.
In the case where (7) is satisfied by a sequence of finite subsets with $\mathrm{card}(\mathcal{A}_n) = \mathcal{O}(N_n^\beta)$, $\beta \geqslant 1$, then these subsets form a polynomial mesh like (1) for K, with $C = 1/(1-\theta)$.

The case of the sphere $K = \mathbb{S}^2$

A sequence of finite point configurations $X_M \subset \mathbb{S}^2$, with cardinality $M \geqslant 2$, is called a "good covering" of the sphere if its *covering radius*

$$\eta(X_M) = \max_{x\in\mathbb{S}^2} \min_{y\in X_M} |x - y| \tag{10}$$

satisfies the inequality

$$\eta(X_M) \leqslant \frac{\alpha}{\sqrt{M}}\ , \tag{11}$$

for some $\alpha > 0$

(Hardin, Michaels and Saff 2016)

Covering inequality for the sphere

The following result can also be obtained via a tangential Markov inequality on the sphere with exponent 1.

Proposition 2

Let $\{X_M\}$, $M \geqslant 1$, *be a good covering of* $\mathbb{S}^2$. *Then for every fixed* $\theta \in (0,1)$ *the sequence* $\mathcal{A}_n = X_{M_n}$, *with*

$$M_n = \lceil \sigma_n^2 n^2 \rceil \ , \quad \sigma_n = \frac{2\pi\alpha}{\theta(2\pi - \theta/n)} \sim \frac{\alpha}{\theta} \ , \quad n \to \infty \ . \qquad (12)$$

is an optimal polynomial mesh of $\mathbb{S}^2$ *with* $C = 1/(1-\theta)$.

Proof (1)

Proof. By (10)-(11) and simple geometric considerations, for every $x \in \mathbb{S}^2$ there exists $y \in X_M$ such that

$$\gamma(x, y) = 2 \sin^{-1} \left(\frac{|x - y|}{2} \right) \leqslant 2 \sin^{-1} \left(\frac{\alpha}{2\sqrt{M}} \right)$$

when $\sqrt{M} \geqslant \alpha/2$, where γ is the (geodesic) Dubiner distance.

Proof (2)

By Proposition 1, in order to determine M_n it is sufficient that

$$2\sin^{-1}\left(\frac{\alpha}{2\sqrt{M}}\right) \leqslant \frac{\theta}{n}$$

or equivalently

$$\frac{\alpha}{2\sqrt{M}} \leqslant \sin\left(\frac{\theta}{2n}\right) \quad . \tag{13}$$

Proof (3)

By the trigonometric inequality $\sin(t) \geqslant t(1 - t/\pi)$, valid for $0 \leqslant t \leqslant \pi$, we get

$$\sin\left(\frac{\theta}{2n}\right) \geqslant \frac{\theta}{2n}\left(1 - \frac{\theta}{2\pi n}\right) ,$$

and thus (13) is satisfied if

$$\frac{\alpha}{2\sqrt{M}} \leqslant \frac{\theta}{2n}\left(1 - \frac{\theta}{2\pi n}\right) ,$$

i.e. for $M \geqslant \sigma_n^2 n^2$. $\square$

Equal area partitions of the unit sphere

An *equal area partition* of $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ is a finite set $\mathcal{P}$ of Lebesgue measurable subsets of $\mathbb{S}^d$, such that

$$\bigcup_{R \in \mathcal{P}} R = \mathbb{S}^d,$$

and for each $R \in \mathcal{P}$,

$$\lambda_d(R) = \frac{\lambda_d(\mathbb{S}^d)}{|\mathcal{P}|},$$

where λ_d is the Lebesgue area measure on $\mathbb{S}^d$.

Diameter bounded sets of partitions

The *diameter* of a region $R \subset \mathbb{R}^{d+1}$ is defined by

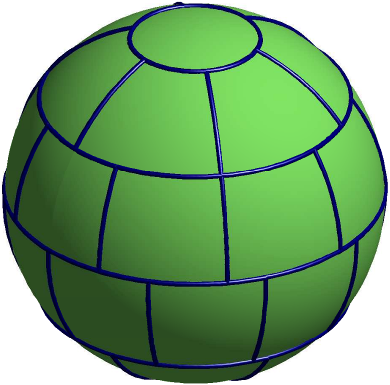
$$\operatorname{diam} R := \sup\{\|\mathbf{x} - \mathbf{y}\| \mid \mathbf{x}, \mathbf{y} \in R\}.$$

A set Ξ of partitions of $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ is *diameter-bounded* with *diameter bound* $K \in \mathbb{R}_+$ if for all $\mathcal{P} \in \Xi$, for each $R \in \mathcal{P}$,

$$\operatorname{diam} R \leqslant K \, |\mathcal{P}|^{-1/d}.$$

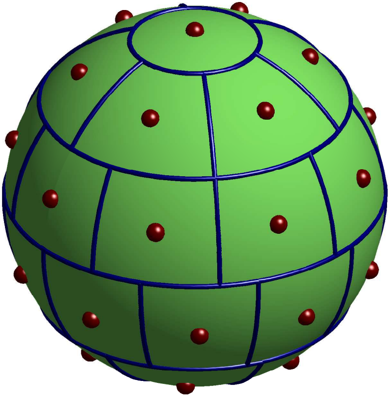
The partition EQ(2,33) on $\mathbb{S}^2 \subset \mathbb{R}^3$

EQ partitions: Recursive Zonal Equal Area partitions of $\mathbb{S}^d$,
$\bigcup \mathrm{EQ}(d, \mathcal{N}) = \mathbb{S}^d$, with $|\mathrm{EQ}(d, \mathcal{N})| = \mathcal{N}$.



The spherical code EQP(2,33) on $\mathbb{S}^2$

EQ codes: The Recursive Zonal Equal Area spherical codes, $\text{EQP}(d,\mathcal{N}) \subset \mathbb{S}^d$, with $|\text{EQP}(d,\mathcal{N})| = \mathcal{N}$.



Properties of the EQ partition of $\mathbb{S}^d$

$\mathrm{EQ}(d, \mathcal{N})$ is the *recursive zonal equal area* partition of $\mathbb{S}^d$ into $\mathcal{N}$ regions.

The set of partitions $\mathrm{EQ}(d) := \{\mathrm{EQ}(d, \mathcal{N}) \mid \mathcal{N} \in \mathbb{N}_+\}$.

The EQ partition satisfies:

Theorem 2

For $d \geqslant 1$, $\mathcal{N} \geqslant 1$, $\mathrm{EQ}(d, \mathcal{N})$ *is an equal-area partition.*

Theorem 3

For $d \geqslant 1$, $\mathrm{EQ}(d)$ *is diameter-bounded.*

Properties of the EQP code on $\mathbb{S}^d$

For $\mathrm{EQP}(d, \mathcal{N})$

Good:

- Centre points of regions of diameter $= \mathrm{O}(\mathcal{N}^{-1/d})$,
- Mesh norm (covering radius) $= \mathrm{O}(\mathcal{N}^{-1/d})$,
- Minimum distance and packing radius $= \Omega(\mathcal{N}^{-1/d})$.
- Normalized spherical cap discrepancy $= \mathrm{O}(\mathcal{N}^{-1/d})$,

Not so good:

- Mesh ratio $= \Omega(\sqrt{d})$,
- Packing density $\leqslant \frac{\pi^{d/2}}{2^d \; \Gamma(d/2+1)}$ as $\mathcal{N} \to \infty$.
- *Inefficient for polynomial interpolation*

The EQP codes form a polynomial mesh on $\mathbb{S}^2$

The EQP codes on the sphere $\mathbb{S}^2$ are a theoretically good covering with $\alpha = 3.5$, but numerical experiments suggest $\alpha = 2.5$.

For example, taking $\theta = 1/2$, by Proposition 2 we have

Corollary 4

The zonal equal area codes $\mathrm{EQP}(M_n, 2)$ *with* $M_n = \left\lceil 49n^2 \left(1 - \frac{1}{4\pi n}\right)^{-2} \right\rceil$ *points are an optimal polynomial mesh on the sphere, with* $C = 2$.

(Hardin, Michaels and Saff 2016; L 2006; L 2007)

The EQP codes form a polynomial mesh on $\mathbb{S}^d$

For general dimension d the following holds.

Theorem 5

The EQ codes form a $\mathbb{P}_\nu(\mathbb{S}^d)$-polynomial mesh .

Proof.

Any finite point set on the unit sphere $\mathbb{S}^d$ with mesh norm at most $(1-c)/\nu$ generates a norming set with constant c for $\mathbb{P}_\nu(\mathbb{S}^d)$. The EQ spherical codes have mesh norm at most $C_d\mathcal{N}^{-1/d}$. Thus if $\mathcal{N} \geqslant (C_d/(1-c))^d\ \nu^d$, then $\mathrm{EQP}(d,\mathcal{N})$ is a norming set with constant c for $\mathbb{P}_\nu(\mathbb{S}^d)$. □

(Jetter, Stöckler and Ward, 1998, 1999; L and Vianello 2014)

Caratheodory-Tchakaloff (CATCH) submeshes (1)

In order to reduce the cardinality of a polynomial mesh, we may try to relax the boundedness requirement for the ratio $\|p\|_{L^\infty(K)}/\|p\|_{\ell^\infty(\mathcal{A}_n)}$, seeking a *weakly admissible mesh* contained in the original one, where the ratio is allowed to increase subexponentially with respect to the degree.

A discrete approach that can be considered a sort of fully discrete hyperinterpolation, is the extraction of *Caratheodory-Tchakaloff submeshes*. These are computable by Linear or Quadratic Programming, and there are rigorous bounds for the corresponding constants C_n .

We recall a discrete version of the Tchakaloff theorem, whose proof is based on the Caratheodory theorem about combinations of finite dimensional cones.

(Sloan 1995; Borwein and Vanderwerff 2010)

Discrete Tchakaloff theorem

Theorem 6

Let μ *be a multivariate discrete measure supported at a finite set* $X = \{x_i\} \subset \mathbb{R}^d$ *, with correspondent positive weights* $\boldsymbol{\lambda} = \{\lambda_i\}$ *,* $i = 1, \ldots, M$ *.*
Then there exists a quadrature formula with nodes $T = \{t_j\} \subseteq X$ *and positive weights* $\mathbf{w} = \{w_j\}$ *,* $1 \leqslant j \leqslant m \leqslant N_\nu = \dim(\mathbb{P}_\nu^d(X))$ *, such that*

$$\int_X p(x)\, d\mu = \sum_{i=1}^{M} \lambda_i\, p(x_i) = \sum_{j=1}^{m} w_j\, p(t_j)\ , \quad \forall p \in \mathbb{P}_\nu^d(X)\ . \qquad (14)$$

We call $T = \{t_j\}$ a set of
Caratheodory-Tchakaloff (CATCH) quadrature points.

(Borwein and Vanderwerff 2010)

Caratheodory-Tchakaloff submeshes (2)

We apply the Tchakaloff theorem to the extraction of a weakly admissible submesh from a polynomial mesh.

Proposition 3

Let $\mathcal{A}_n \subset K$ be a polynomial mesh like (1) for K with cardinality $M_n > N_{2n} = dim(\mathbb{P}^d_{2n}(K))$, let μ be the discrete measure with unit weights supported at $\mathcal{A}_n$, and let $T_{2n} = \{t_j\}$ be the $m \leqslant N_{2n}$ CATCH quadrature points for degree $\nu = 2n$, extracted from $\mathcal{A}_n$, with corresponding weights $\mathbf{w} = \{w_j\}$, $1 \leqslant j \leqslant m$.

Then T_{2n} is a weakly-admissible CATCH submesh for K with $C_n = C\sqrt{M_n}$, and the following estimate holds for the corresponding weighted least squares approximation

$$\|f - \mathcal{L}^{w}_{T_{2n}} f\|_{L^\infty(K)} \leqslant (1 + C_n) \min_{p \in \mathbb{P}^d_n(K)} \|f - p\|_{L^\infty(K)} . \qquad (15)$$

Caratheodory-Tchakaloff submeshes (3)

The error estimate (15) follows from the inequality

$$\Lambda_w(T_{2n}) = \|\mathcal{L}^w_{T_{2n}}\| = \sup_{f \in C(K), f \neq 0} \frac{\|\mathcal{L}^w_{T_{2n}} f\|_{L^\infty(K)}}{\|f\|_{L^\infty(K)}} \leqslant C_n = C\sqrt{M_n}\ . \tag{16}$$

The error estimate (15) for weighted discrete Least Squares on the CATCH submesh turns out to coincide with the natural error estimate (4) for unweighted Least Squares on the original polynomial mesh.

CATCH submeshes on the sphere (1)

From Proposition 2 and 4 and Corollary 1 and 2 we get

Corollary 7

Let $\mathcal{A}_n$ be a good covering optimal polynomial mesh as in Proposition 2, and let T_{2n} be the extracted CATCH submesh (with corresponding weights).

Then, T_{2n} is a weakly admissible mesh for the sphere with cardinality $N_{2n} = \dim(\mathbb{P}^3_{2n}(\mathbb{S}^2)) = (2n+1)^2$, and (15) holds for the corresponding weighted Least Squares polynomial approximation $\mathcal{L}^w_{T_{2n}} f$ to $f \in C(\mathbb{S}^2)$, where

$$C_n = \frac{\sigma_n\, n}{1-\theta} \sim \frac{\alpha n}{\theta(1-\theta)}\,, \quad n \to \infty\,. \tag{17}$$

CATCH submeshes on the sphere (2)

In particular, for a CATCH submesh of the zonal equal area configurations of Corollary 1, we have

$$C_n = \frac{14n}{1-(4\pi n)^{-1}} \sim 14n \; , \quad n \to \infty \; . \tag{18}$$

By (3), (16) and (17) we get $\mathcal{O}(n)$ estimates for the least squares operator norms, whereas the best projection operators on $\mathbb{P}_n^3(\mathbb{S}^2)$ have a $\mathcal{O}(n^{1/2})$ norm. On the other hand, (16) turns out to be an overestimate of the actual norm, as we shall see in the numerical examples.

(Sloan and Womersley 2000)

Numerical approaches

In order to compute a sparse nonnegative solution to the underdermined system that exists by the Tchakaloff theorem, there are a number of different approaches available.

On the sphere, we use the classical spherical harmonics basis to define the Vandermonde-like matrix V.

Non-Negative Least Squares

A first approach uses Quadratic Programming, specifically the Non-Negative Least Squares problem

$$\text{QP}: \quad \begin{cases} \min \|V^t \mathbf{u} - \mathbf{b}\|_2 \\ \mathbf{u} \geqslant \mathbf{0} \end{cases} \tag{19}$$

which can be solved by the Lawson-Hanson active set method, which naturally seeks a sparse solution.

The nonzero components of $\mathbf{u}$ identify the weights $\mathbf{w} = \{w_j\}$ and the corresponding CATCH submesh T_{2n}.

(Piazzon, Sommariva and Vianello 2016; Sommariva and Vianello 2015)

Linear Programming

A second approach is based on Linear Programming.

$$\text{LP}: \quad \begin{cases} \min \mathbf{c}^t \mathbf{u} \\ V^t \mathbf{u} = \mathbf{b}\,, \quad \mathbf{u} \geqslant \mathbf{0} \end{cases} \tag{20}$$

where the constraints identify a polytope and the vector $\mathbf{c}$ is suitably chosen.

Solving the problem by the classical Simplex Method, we get a nonnegative sparse solution to the underdetermined system.

(Piazzon, Sommariva and Vianello 2016; Ryu and Boyd 2015; Tchernychova 2015)

Numerical results

In our Matlab codes for Caratheodory-Tchakaloff Least Squares we used both the QP approach (via both the *lsqnonneg* function and an optimized version of this function), and the LP approach (via the Simplex Method in the Matlab interface of the CPLEX package).

In Table 1, we report the numerical results corresponding to the extraction of CATCH submeshes from zonal equal area meshes of $\mathbb{S}^2$, for a sequence of degrees.

The cardinality of the CATCH submeshes is $\dim(\mathbb{P}^3_{2n}(\mathbb{S}^2)) = (2n+1)^2$, and the Compression Ratio, $C_{ratio} = card(\mathcal{A}_n)/card(T_{2n})$, increases, approaching the asymptotic value $49/4 = 12.25$.

(Piazzon, Sommariva and Vianello 2016)

Results of the QP approach (1)

Table: Results of lsqnonneg for different Weakly Admissible Meshes

deg	$\|\mathcal{A}_n\|$	$\|\mathcal{T}_{2n}\|$	C_{ratio}	$\frac{W_{max}}{W_{avg}}$	$\frac{W_{min}}{W_{avg}}$	$\Lambda_{\mathcal{A}_n}$	$1.5\sqrt{n}$	$\Lambda_{\mathcal{T}_{2n}}$
2	213	25	8.5	2.1	2.1×10^{-2}	2.2	2.1	2.5
5	1265	121	10.5	2.4	1.9×10^{-4}	3.3	3.4	3.7
8	3200	289	11.1	2.4	7.5×10^{-5}	4.2	4.2	4.7
11	6016	529	11.4	2.4	7.5×10^{-6}	4.9	5.0	5.3
14	9715	841	11.6	2.6	1.2×10^{-5}	5.6	5.6	5.9
17	14295	1225	11.7	2.6	1.3×10^{-6}	6.2	6.2	6.5
20	19757	1681	11.8	2.5	2.3×10^{-6}	6.7	6.7	7.1

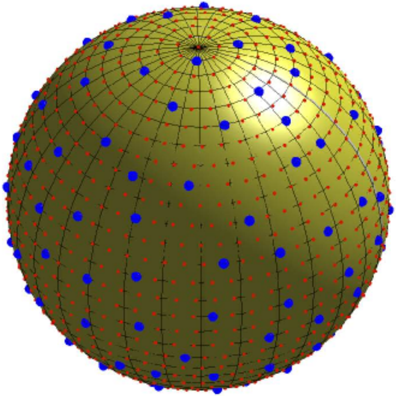
Results of the QP approach (2)

Since $\sum w_j = card(\mathcal{A}_n)$, the average CATCH weight turns out to coincide with the Compression Ratio.

The compressed least squares operator norms $\Lambda_w(T_{2n})$ are close to the norm $\Lambda(\mathcal{A}_n)$ of the least squares operator on the starting mesh.

On the other hand, all the norms are much lower than the theoretical overestimate $C_n \sim 14n$ in Corollary 2, having substantially a $\mathcal{O}(n^{1/2})$ increase.

Example of a CATCH submesh on the sphere



CATCH submesh (121 points, blue) extracted by NNLS from EQP(2, 1187) (red) for degree $n = 5$.