

Computing on Spheres

From spherical designs to scattered, random, and data-driven points

Hao-Ning Wu
University of Georgia

Old Dominion University
Jan 27, 2026

❑ Approximation theory and practice

- Point distribution and numerical integration
- Polynomials
- Splines
- Kernel-based approximation

❑ Numerical methods for PDEs and integral equations

- Spectral methods for PDEs
- quadrature-based methods for integral equations
- Spline-based methods

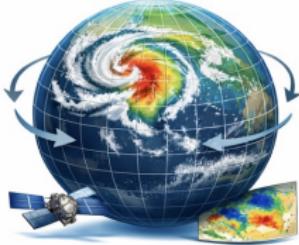
❑ Numerical methods for optimal control

- Integrating numerical PDE solvers with optimization techniques

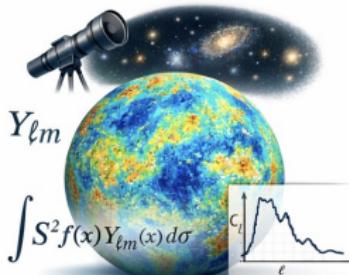
❑ Image/signal processing along with optimization

- Compressed sensing and image reconstruction
- First-order optimization algorithms

Spherical World



Atmospheric & Oceanic Sciences



Astronomy & Cosmology



Spherical Image & VR Processing

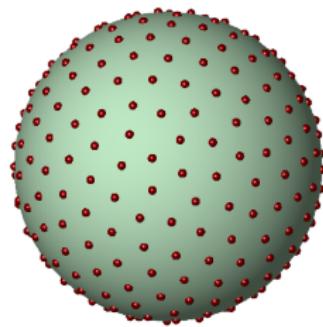
- ❑ Many core scientific problems live on the sphere
- ❑ Climate, astronomy, and immersive imaging are not edge cases
- ❑ Accuracy on the sphere is foundational, not cosmetic

Numerical Integration on Spheres

The Classical Assumption

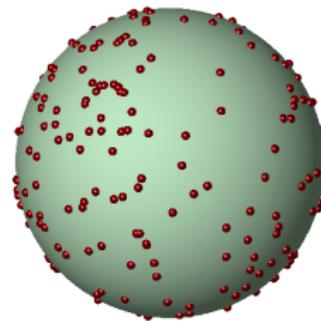
What numerical analysis assumes

- Exact quadrature rules
- Highly structured point sets
- Full access over sampling locations



What data gives us

- Scattered or random samples
- Data collected externally
- Limited or no control over point locations



Quadrature exactness is a mathematical luxury, not a practical assumption.

Spherical Designs

□ Consider spherical integrals

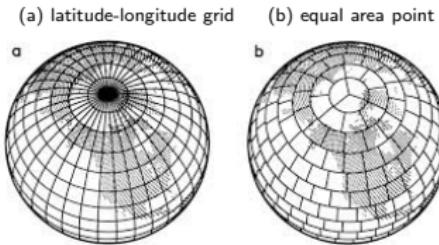
$$\int_{S^d} f d\omega = \int_0^{2\pi} \int_{-1}^1 f(\lambda, \mu) d\mu d\lambda,$$

where $S^d = \{x \in \mathbb{R}^{d+1} : \|x\|_2 = 1\}$.

latitude: Gauss–Legendre in μ

longitude: uniform grid in λ

✗ Severe point clustering near poles



□ **Spherical analogue of Gauss quadrature:** A spherical t -design $\{x_j\}_{j=1}^m$ is a carefully constructed set of points such that

$$\frac{1}{|S^d|} \int_{S^d} g d\omega = \frac{1}{m} \sum_{j=1}^m g(x_j) \quad \forall g \in \mathbb{P}_t$$

- High accuracy for approximation, PDEs, and integral equations
- High-accuracy designs are difficult to compute and incompatible with scattered, random, or data-driven sampling.

From Exactness to Marcinkiewicz–Zygmund inequalities

□ Classical View: Accuracy via Exactness

Accuracy comes from exact integration:

$$\sum_{j=1}^m w_j p(x_j) = \int_{\mathbb{S}^d} p d\omega \quad \forall p \in \mathbb{P}_{2n}$$

□ Our View: Accuracy via Stability

(Marcinkiewicz–Zygmund 1937) Stability comes from geometric sampling conditions: For $\eta \in [0, 1]$,

$$(1 - \eta) \int_{\mathbb{S}^d} p^2 d\omega \leq \sum_{j=1}^m w_j p(x_j)^2 \leq (1 + \eta) \int_{\mathbb{S}^d} p^2 d\omega \quad \forall p \in \mathbb{P}_n,$$

implying

$$\left| \sum_{j=1}^m w_j p(x_j)^2 - \int_{\mathbb{S}^d} p^2 d\omega \right| \leq \eta \int_{\mathbb{S}^d} p^2 d\omega$$

We don't need quadrature exactness but MZ.

Why MZ is the right language: Encodes geometry, not exactness

- ❑ Includes spherical t -designs as a special case
- ❑ Holds for scattered and random points

scattered data

(Filbir & Mhaskar, 2011)

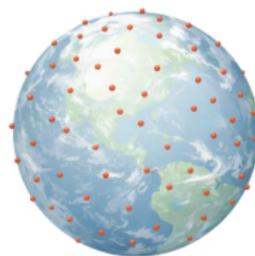
- ❑ MZ holds if $n \lesssim \eta / h_{\mathcal{X}_m}$



random data

(Le Gia & Mhaskar, 2009)

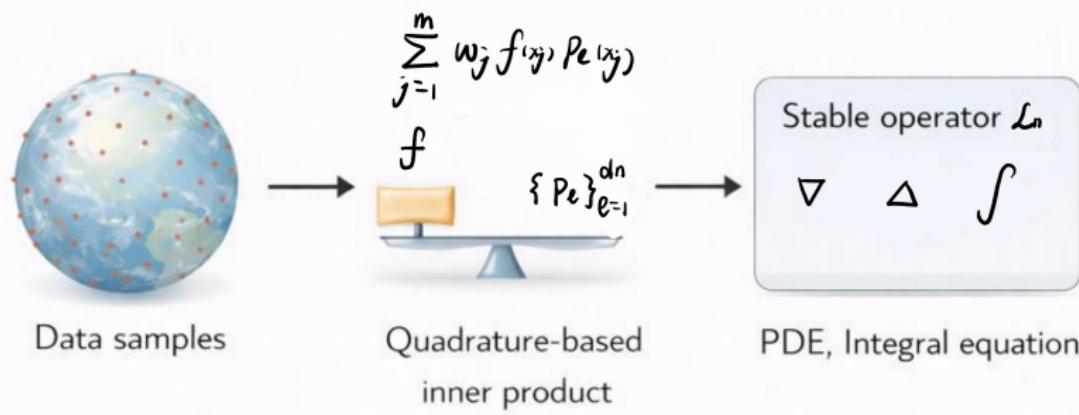
- ❑ MZ holds with probability at least $1 - \bar{c}n^{-\gamma}$ on the condition $m \geq \bar{c}n^d \log n / \eta^2$



Numerical analysis: from quadrature exactness-based to MZ-based.

From Numerical Integration to Function Approximation

Hyperinterpolation as an Approximation Scheme



- **Goal:** To build an approximation operator from **sampled data** at scattered locations on the sphere/manifold.
- **Scheme:** To replace continuous inner products by a **weighted quadrature rule**:

$$\mathcal{L}_n f := \sum_{\ell=1}^{d_n} \langle f, p_\ell \rangle_m p_\ell,$$

where $\langle f, g \rangle_m := \sum_{j=1}^m w_j f(x_j)g(x_j)$ with all $w_j > 0$.

What we gain and lose by relaxing exactness

Sloan (1995): quadrature exactness degree $2n$

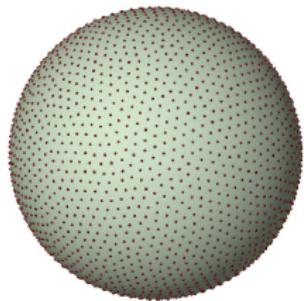
- $\mathcal{L}_n p = p$ for any $p \in \mathbb{P}_n$
- $\|\mathcal{L}_n f\|_2 \leq |\mathbb{S}^d|^{1/2} \|f\|_\infty$
- $\|\mathcal{L}_n f - f\|_2 \leq 2|\mathbb{S}^d|^{1/2} E_n(f)$

where $E_n(f) := \inf_{p \in \mathbb{P}_n} \|f - p\|_\infty$.

Ours (2022): quadrature exactness degree $n + k$ ($0 < k \leq n$) + MZ

- $\mathcal{L}_n p = p$ for any $p \in \mathbb{P}_{\textcolor{red}{k}}$
- $\|\mathcal{L}_n f\|_2 \leq \frac{|\mathbb{S}^d|^{1/2}}{\sqrt{1-\eta}} \|f\|_\infty;$
- $\|\mathcal{L}_n f - f\|_2 \leq \left(\frac{1}{\sqrt{1-\eta}} + 1 \right) |\mathbb{S}^d|^{1/2} E_{\textcolor{red}{k}}(f)$

spherical 50-design: 2601 pts



spherical 30-design: 961 pts

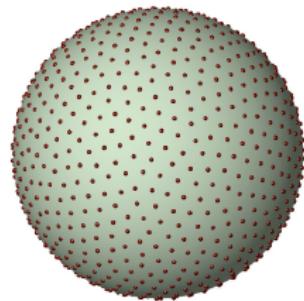
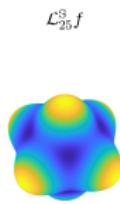
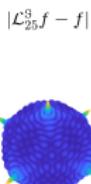


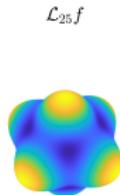
Figure: Spherical 50- and 30-designs, generated by the optimization method proposed by An, Chen, Sloan, and Womersley (2010).


 $\mathcal{L}_{25}^3 f$

1.56
1.54
1.52
1.50
1.48
1.46
1.44


 $|\mathcal{L}_{25}^3 f - f|$

$\times 10^{-6}$
2
1
0.5


 $\mathcal{L}_{25} f$

1.56
1.54
1.52
1.50
1.48
1.46
1.44


 $|\mathcal{L}_{25} f - f|$

$\times 10^{-4}$
2
1
0.5

$n = 25$
f: Wendland function (4.3)

$(k, n + k, m)$	$\ \mathcal{L}_n f - f\ _2$	$\ \mathcal{L}_n f - f\ _\infty$
(1,26,729)	1.4703e-04	1.1973e-02
(2,27,784)	1.0036e-04	7.2393e-03
(3,28,841)	7.7225e-05	5.6280e-03
(4,29,900)	3.6550e-06	2.2721e-04
(5,30,961)	2.7813e-06	2.1562e-04
(6,31,1024)	9.0144e-07	7.3522e-05
(7,32,1089)	6.3510e-07	5.4311e-05
(8,33,1156)	1.5667e-07	1.4221e-05
(9,34,1225)	1.2137e-07	1.0454e-05
(10,35,1296)	6.0979e-08	7.9442e-06
(11,36,1369)	5.3640e-08	5.4959e-06
(12,37,1444)	1.8896e-08	3.3341e-06
(13,38,1521)	1.9095e-08	3.7055e-06
(14,39,1600)	1.6651e-08	3.2061e-06
(15,40,1681)	1.4991e-08	2.6047e-06
(16,41,1764)	1.4137e-08	2.9486e-06
(17,42,1849)	1.3659e-08	2.5557e-06
(18,43,1936)	1.3509e-08	2.5579e-06
(19,44,2025)	1.3433e-08	2.5896e-06
(20,45,2116)	1.3354e-08	2.6336e-06
(21,46,2209)	1.3318e-08	2.5630e-06
(22,47,2304)	1.3309e-08	2.4906e-06
(23,48,2401)	1.3309e-08	2.5130e-06
(24,49,2500)	1.3294e-08	2.4568e-06
(25,50,2601)	1.3294e-08	2.4959e-06

Accuracy degrades gracefully, not catastrophically.

Assume the quadrature rule **satisfies the MZ property** only. Then for any $f \in C(\Omega)$:

$$\begin{aligned}\|\mathcal{L}_n f - f\|_{L^2} &\leq \left(\sqrt{1 + \eta} \left(\sum_{j=1}^m w_j \right)^{1/2} + |\mathbb{S}^d|^{1/2} \right) E_{\textcolor{red}{n}}(f) \\ &\quad + \sqrt{\eta^2 + 4\eta} \|p^*\|_{L^2},\end{aligned}$$

where $\|f - p^*\|_\infty = E_n(f)$.

- ❑ Error bounds depends on sampling geometry
- ❑ Applicable to fully data-driven settings

From Approximation to PDEs and Integral Equations

Application I: Semi-linear PDEs on Spheres

- To compute smooth solutions of **semi-linear PDEs** on $S^{d-1} \subset \mathbb{R}^d$ with dimension $d \geq 3$ of the form

$$u_t = \mathbf{L}u + \mathbf{N}(u), \quad u(0, x) = u_0(x),$$

where \mathbf{L} is a constant-coefficient linear differential operator, and \mathbf{N} is a constant-coefficient nonlinear differential (or non-differential) operator of lower order.

- Nonlinearity + sphere geometry
- High accuracy traditionally needs highly structured sampling
- Data-driven point sets break classical assumptions

- Example: **Allen–Cahn equation**

$$u_t = \nu^2 \Delta u - F'(u), \quad u(0, x) = u_0(x),$$

where $F'(u) = f(u) = u^3 - u$.

Our idea: Linearizing the nonlinear part $\mathcal{N}(u)$ by hyperinterpolation:

$$\begin{cases} \frac{u^{n+1} - u^n}{\tau} = \nu^2 \Delta u^{n+1} - \mathcal{L}_N ((u^n)^3 - u^n), & n \geq 0, \\ u^0 = \mathcal{L}_N u_0 \end{cases}$$

where $\tau > 0$ is the time step.

$$(n, t) = (0, 0)$$

$$(n, t) = (10, 5)$$

$$(n, t) = (20, 10)$$

$$(n, t) = (30, 15)$$

$$(n, t) = (140, 70)$$

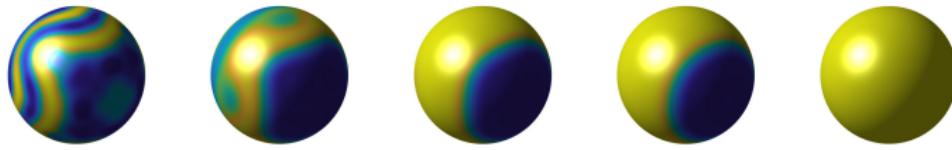


Figure: Numerical solution to the Allen–Cahn equation with $\nu = 0.1$ and initial condition $u(0, x, y, z) = \cos(\cosh(5xz) - 10y)$ using our scheme with $\tau = 0.5$, $N = 15$ and $m = (2N + 1)^2 = 961$ **equal area points**

Application II: Fredholm Integral Equations with Singular Kernels

- Consider the Fredholm integral equation of the second kind

$$\varphi(\mathbf{x}) - \int_{\mathbb{S}^2} h(|\mathbf{x} - \mathbf{y}|) K(\mathbf{x}, \mathbf{y}) \varphi(\mathbf{y}) d\omega(\mathbf{y}) = f(\mathbf{x})$$

on \mathbb{S}^2 , where $|\mathbf{x} - \mathbf{y}| := \sqrt{2(1 - \mathbf{x} \cdot \mathbf{y})}$, the kernel K is **continuous** but possibly **oscillatory**, and the weight function $h : (0, \infty) \rightarrow \mathbb{R}$ may be **weakly singular**, i.e., h is continuous for all $\mathbf{x}, \mathbf{y} \in \mathbb{S}^2$ with $\mathbf{x} \neq \mathbf{y}$, and there exists positive constants M and $\alpha \in (0, 2]$ such that

$$|h(|\mathbf{x} - \mathbf{y}|)| \leq M |\mathbf{x} - \mathbf{y}|^{\alpha-2};$$

- Potential issues:

- Singular kernels amplify point clustering
- Standard pointwise quadrature can explode
- Need a discretization that respects the singular structure

A semi-analytical approach for the singular kernel

- Let $\{Y_{\ell,k} : \ell = 0, \dots, n, k = 1, \dots, 2\ell + 1\}$ be the set of **spherical harmonics** of degree $\leq n$. They are orthogonal polynomials on spheres.
- Modified moments** (derived by Funk–Hecke formula) Computing the singular part analytically

$$\int_{S^2} h(|\mathbf{x} - \mathbf{y}|) Y_{\ell,k}(\mathbf{y}) d\omega(\mathbf{y}) = \mu_\ell Y_{\ell,k}(\mathbf{x}),$$

where

$$(|\mathbf{x} - \mathbf{y}| := \sqrt{2(1 - \mathbf{x} \cdot \mathbf{y})} \text{ and } h(|\mathbf{x} - \mathbf{y}|) = (2^{1/2}(1 - \mathbf{x} \cdot \mathbf{y})^{1/2}))$$

$$\mu_\ell := 2\pi \int_{-1}^1 h(2^{1/2}(1 - t)^{1/2}) P_\ell(t) dt.$$

☞ **Example 1:** $h(|\mathbf{x} - \mathbf{y}|) = |\mathbf{x} - \mathbf{y}|^\nu$ with $-2 < \nu < 0$. Then

$$\mu_\ell = 2^{\nu+2}\pi\left(-\frac{\nu}{2}\right)_\ell \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\ell + \frac{\nu}{2} + 2\right)}$$

with $(x)_n := \frac{\Gamma(x+n)}{\Gamma(x)}$ for $x > 0$.

☞ **Example 2:** $h(|\mathbf{x} - \mathbf{y}|) = \log|\mathbf{x} - \mathbf{y}|$. Then

$$\mu_\ell = 2\pi \int_{-1}^1 \log(2^{1/2}(1-t)^{1/2})P_\ell(t)dt = \pi \int_{-1}^1 \log(2(1-t))P_\ell(t)dt.$$

☞ **Example 3:** $h(|\mathbf{x} - \mathbf{y}|) = |\mathbf{x} - \mathbf{y}|^{\nu_1}|\mathbf{x} + \mathbf{y}|^{\nu_2}$ with $-2 \leq \nu_1, \nu_2 < 0$. Then

$$\mu_\ell = 2^{(\nu_1+\nu_2)/2}(2\pi) \int_{-1}^1 (1-t)^{\nu_1/2}(1+t)^{\nu_2/2}P_\ell(t)dt.$$

A new quadrature rule for the integral operator

For the integral operator $\int_{\mathbb{S}^2} h(|\mathbf{x} - \mathbf{y}|) \mathcal{K}(\mathbf{x}, \mathbf{y}) \varphi(\mathbf{y}) d\omega(\mathbf{y})$, we approximate it as

$$\begin{aligned} & \int_{\mathbb{S}^2} h(|\mathbf{x} - \mathbf{y}|) \mathcal{L}_n(K(\mathbf{x}, \cdot) \varphi(\cdot)) d\omega(\mathbf{y}) \\ &= \int_{\mathbb{S}^2} h(|\mathbf{x} - \mathbf{y}|) \left(\sum_{\ell=0}^n \sum_{k=1}^{2\ell+1} \langle K(\mathbf{x}, \cdot) \varphi, Y_{\ell,k} \rangle_m Y_{\ell,k}(\mathbf{y}) \right) d\omega(\mathbf{y}) \\ &= \sum_{\ell=0}^n \sum_{k=1}^{2\ell+1} \boxed{\left(\int_{\mathbb{S}^2} h(|\mathbf{x} - \mathbf{y}|) Y_{\ell,k}(\mathbf{y}) d\omega(\mathbf{y}) \right)} \langle K(\mathbf{x}, \cdot) \varphi, Y_{\ell,k} \rangle_m \\ &= \sum_{j=1}^m \textcolor{teal}{w_j} \left(\sum_{\ell=0}^n \sum_{k=1}^{2\ell+1} \boxed{\mu_\ell Y_{\ell,k}(\mathbf{x})} Y_{\ell,k}(\mathbf{x}_j) \right) K(\mathbf{x}, \mathbf{x}_j) \varphi(\mathbf{x}_j) \\ &=: \sum_{j=1}^m \textcolor{teal}{W_j}(\mathbf{x}) K(\mathbf{x}, \mathbf{x}_j) \varphi(\mathbf{x}_j), \end{aligned}$$

Two-stage numerical scheme for the integral equation

Let φ_γ denotes the numerical solution:

$$\varphi_\gamma(\mathbf{x}) - \sum_{j=1}^m W_j(\mathbf{x}) K(\mathbf{x}, \mathbf{x}_j) \varphi_\gamma(\mathbf{x}_j) = f(\mathbf{x})$$

☞ **First stage** We set $\mathbf{x} = \mathbf{x}_j$, $j = 1, \dots, m$, then numerically solves the obtained system of linear equations

$$\varphi_\gamma(\mathbf{x}_i) - \sum_{j=1}^m W_j(\mathbf{x}_i) K(\mathbf{x}_i, \mathbf{x}_j) \varphi_\gamma(\mathbf{x}_j) = f(\mathbf{x}_i), \quad i = 1, \dots, m$$

for the quantities $\varphi_\gamma(\mathbf{x}_j)$, $j = 1, \dots, m$.

☞ **Second stage** The value of $\varphi_\gamma(\mathbf{t})$ at any $\mathbf{t} \in \mathbb{S}^2$ can be evaluated by the direct usage of

$$\varphi_\gamma(\mathbf{t}) = f(\mathbf{t}) + \sum_{j=1}^m W_j(\mathbf{t}) K(\mathbf{t}, \mathbf{x}_j) \varphi_\gamma(\mathbf{x}_j).$$

Numerical experiments for the integral equation solver

Toy example setting: Let $\varphi \equiv 1$ and compare the computed solution with 1.

Let $h(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|^{-0.5}$ and $K(\mathbf{x}, \mathbf{y}) = \cos(10|\mathbf{x} - \mathbf{y}|)$, thus

$$f(\mathbf{x}, \mathbf{y}) \equiv 1 - 2\pi \int_{-1}^1 \left(\sqrt{2(1-t)} \right)^{-0.5} \cos \left(10\sqrt{2(1-t)} \right) dt$$

$$\approx 0.303738699125466.$$

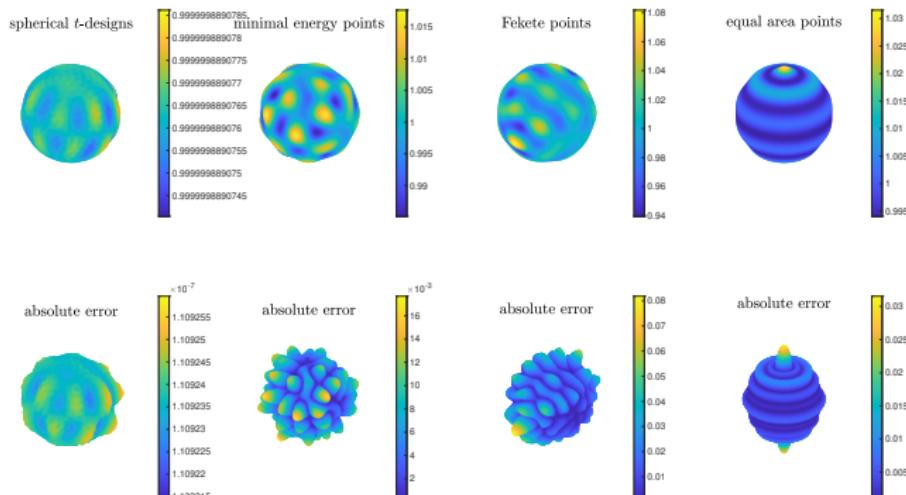


Figure: Numerical solutions with $n = 20$ and $m = (2n+1)^2$.

Research Vision

Computing on Manifolds from Imperfect Data

A long-term, cohesive research program.

- Sampling geometry as a unifying principle
- Approximation and learning from scattered data
- PDEs and control on manifolds driven by data
- Applications in atmospheric science

Direction A: Sampling Geometry Beyond Spheres

Core questions: Adaptive & data-aware sampling over

- other compact manifolds
- spherical triangles
- Euclidean domains

Where I'm well-positioned

- MZ-type conditions as a transferable language
- Existing results generalize naturally

This direction generalizes my current results, rather than starting over.

Core questions

- ❑ Learning operators from samples with stability guarantees
- ❑ Noise-robust approximation on manifolds
- ❑ Data-aware learning theory

My connection

- ☒ I bring: stability, geometry, guarantees
- ☒ ML brings: scalability, representation, data efficiency

My goal is to provide mathematical structure that makes learning reliable.

Core questions

- ❑ Data-driven discretization of PDEs on manifolds
- ❑ Reduced-order models under imperfect sampling
- ❑ Computation and control with learned or partially known dynamics

Potential funding interface

- Computational mathematics
- Scientific machine learning

Application of the Cohesive Program: Atmospheric Science

Scientific challenges

- ❑ Global climate and weather models are posed on the sphere
- ❑ Data are massive, heterogeneous, and imperfectly sampled
- ❑ Long-time integration demands numerical stability

Potential funding interface

- Computational mathematics for climate modeling
- Data-driven and uncertainty-aware numerical methods
- Scientific machine learning for geophysical systems

A natural bridge between numerical analysis, data, and atmospheric science.

Thanks for your attention.



Photo taken at the Castillo San Cristóbal, San Juan, Puerto Rico.