



GeomScale

Scalable Geometric Computing

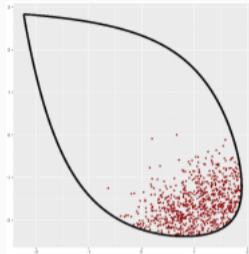
Vissarion Fysikopoulos

October 22, 2025

NumFOCUS Summit 2025

What is GeomScale?

- Open-source organization (since 2019)
- NumFOCUS affiliated project (since 2021)
- Google Summer of Code organization (6 years)



Central problem: sampling from a multi-dimensional distribution constrained within a constraint space

- Volume Computation: Estimating the size of such domains.
- Integration: Computing integrals of functions over these domains.
- Optimization: Optimize a convex function over these domains.
- Copula Computation: Estimating multivariate cumulative distribution functions.

Random walks, domains and distributions

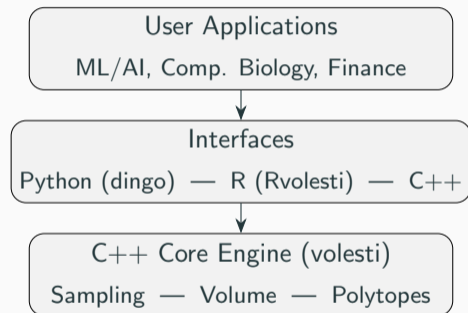
GeomScale's core technology integrates three key aspects:

- **Random Walks:** > 10 Monte Carlo-based (geometric) sampling algorithms tailored to different geometric and statistical problems (and constraints).
- **Constraint Domains:**
 - Polyhedra: Linear inequalities, convex hulls, and Minkowski sums (e.g., Zonotopes).
 - Non-linear convex sets: Spectrahedra (SDP feasibility regions), intersections of polyhedra with ellipsoids or spheres.
 - General Convex Bodies: Represented by oracles for membership, boundary, or reflection operations.
 - Non-Convex Bodies: Intersections of polyhedra boundary with an ellipsoid.
- **Distributions:** Uniform, Gaussian, exponential, and general log-concave distributions, with specialized/optimized algorithms for specific cases.

Three dimensions:

- **Variety:** Broad range of random walk algorithms and support for diverse constraint domains (e.g., only software that computes volumes of zonotopes in hundreds of dimensions).
- **Specialization:** Optimized methods for specific distributions and domain characteristics (e.g., fastest sampler for uniform distributions on polyhedra).
- **Efficiency with guarantees:** Computational scalability for high-dimensional problems. Quality of results is guaranteed by statistical tests.

Integrated Compute Engine



Applications and the NumFocus Ecosystem

- Possible interactions with NF projects
 - PyMC, Stan: support for sampling over general constrained domains
 - Equadratures: complement numerical integration with efficient sampling under convex constraints
- Applications
 - Metabolic networks (computational biology)
 - Computational finance: risk/uncertainty modeling
 - AI/ML: sampling in constrained models, approximate counting problems

Resources

- Source code:



- Documentation (C++):



- Webpage & blogs:

