Computational Geometry for Quantum Gravity

Will Cunningham

Perimeter Institute for Theoretical Physics

9 June 2020

RPI Cyberinfrastructure Training

Primary Research Questions

Two main approaches to discrete QG:

How can I discretize a particular geometric space?

- What structures do I use? What are the relevant symmetries?
- How much of the manifold do I keep? Metric/coordinates?
- What are "good" observables? Easy to measure, fast convergence, renormalizable.

Inverse Problem: How can I recover a particular geometric space?

- How do I construct statistical ensembles of discrete objects?
- How do I restrict focus to a particular dimension, manifold, topology, etc?
- Which subset of all discrete objects are also geometric?
- How does continuum Lorentzian geometry emerge in discrete quantum gravity?

Outline

Discrete Geometric Structures

- Sets, Graphs, and Simplicial Complexes
- Random Geometric Graphs
- Lorentzian Geometry from Order Theory

Observables and Convergence

- Geodesic Paths
- Cycles and Path Length Distribution
- Ricci Curvature

Statistical Physics of Random Graphs

- Energy/Action of Graphs
- Graph Ensembles and Monte Carlo Methods
- Manifold Reconstruction

Discrete Geometric Structures

Sets, Graphs, and Simplicial Complexes

The simplest discretization is to consider a point-cloud.

- Consider a compact region of a differential manifold with a Riemannian or Lorentzian metric $(\mathcal{M}, g_{\mu\nu})$.
- Sample coordinates within the region uniformly with respect to the metric tensor
- This stochastic process is called the *Poisson Point Process*:

$$P(N(V) = n) = \frac{(\rho V)^n}{n!} e^{-\rho V}$$

- The measure is invariant under symmetries of the metric, e.g., translations, rotations, boosts
- Discreteness scale: $\ell = \rho^{-1/d} = (N/V)^{-1/d}$
- Data structures: vector of $N \times d$ doubles (coordinates)
- Complexity: $O(N)$

Example: galaxy clusters/filaments

Sets, Graphs, and Simplicial Complexes

A manifold is a union of open balls (Riem.) or diamonds (Lor.)

- Let us define a characteristic length $r \gg \ell$. When $d(x_i, x_y) < r$, we say that the two neighborhoods overlap, and we add a *relation* between the pair of points.
- For Lorentzian metrics, $r = 0$. Why is that?
- After considering all pairs, the resulting discrete structure is a *random geometric graph*. This is the skeleton of a simplicial complex called the Vietoris-Rips complex, which is homotopy-equivalent to the manifold. Lorentzian RGGs are called *causal sets*.
- If $r/\ell \to \infty$ as $N \to \infty$, the graph is *dense*. If the mean number of connections remains constant, the graph is *sparse*.
- Data structures: adjacency matrix or edge list; $O(N^2)$

Sets, Graphs, and Simplicial Complexes

A simplicial complex is a graph with higher-order connections.

- When two neighborhoods overlap, we draw an edge between them.
- When three mutually overlap, we draw a *face*. A k-clique in the graph (skeleton) forms a $(k-1)$-dimensional facet.
- Simplicial complexes form the basis for two approaches to quantum gravity, Euclidean/Causal Dynamical Triangulations (EDT, CDT)
- Data structures: simplex tree

Uses: topological data analysis, persistent homology

Random Geometric Graphs

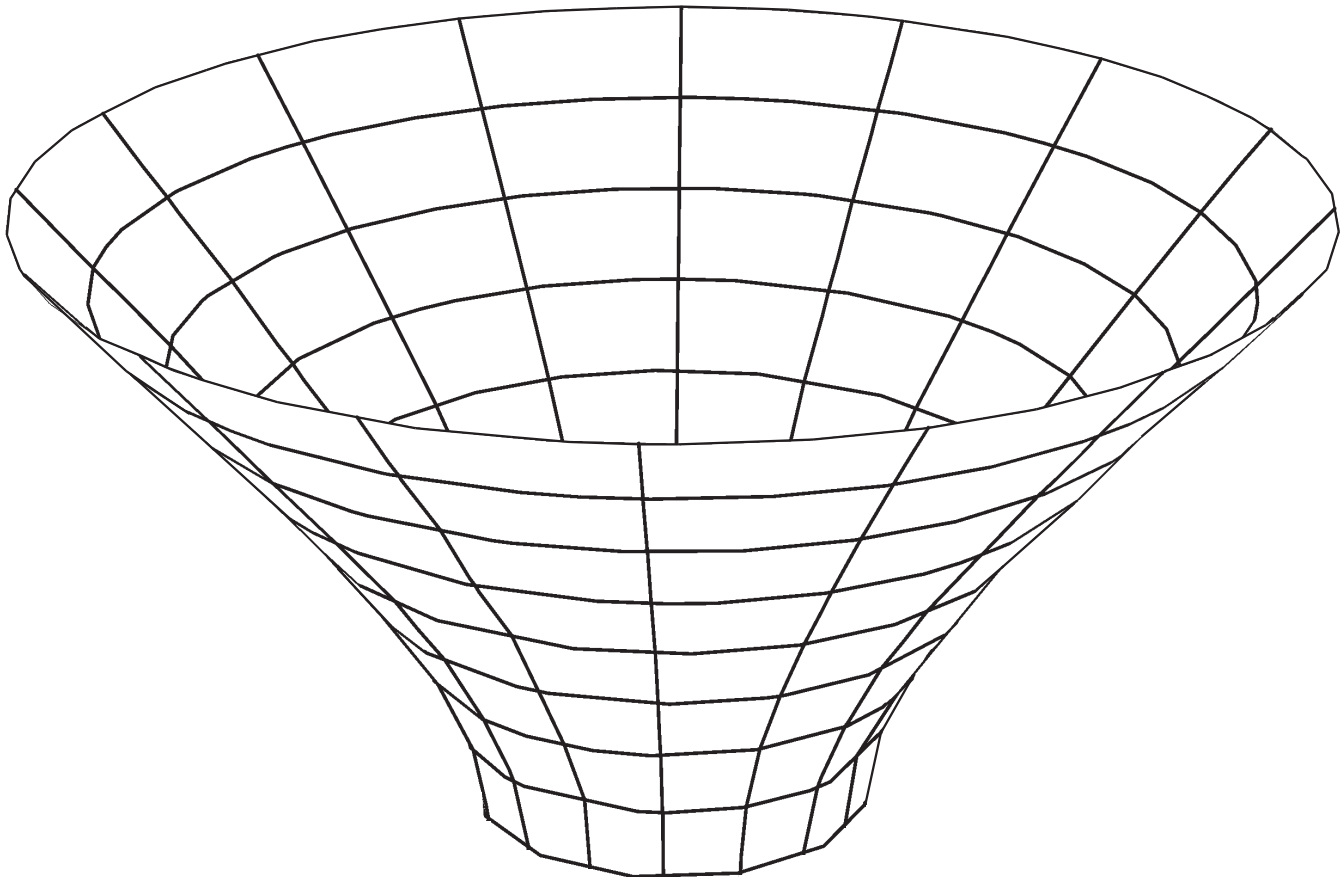
How do we discretize curved spacetime? Consider (1+1) de Sitter.

1. Write the line element:
   $ds^2 = g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + \cosh^2(t) d\theta^2$
2. The volume form gives a joint measure on the variables (t, θ)
   $dV = \sqrt{|g|}\, dt\, d\theta = \cosh(t)\, dt\, d\theta \qquad p(t, \theta) = \cosh(t)$
3. If the measure is a product $p(t)p(\theta)$, sample points from the normalized distributions w.r.t the boundaries
   $p(t) = \frac{\cosh(t)}{\sinh(t_0)} \qquad p(\theta) = \frac{1}{2\pi}$
4. When $ds^2 < 0$, points are timelike-separated (causally connected); otherwise, they are spacelike-separated. We check to see if light cones overlap.

Is this graph dense or sparse? How does it depend on the region?

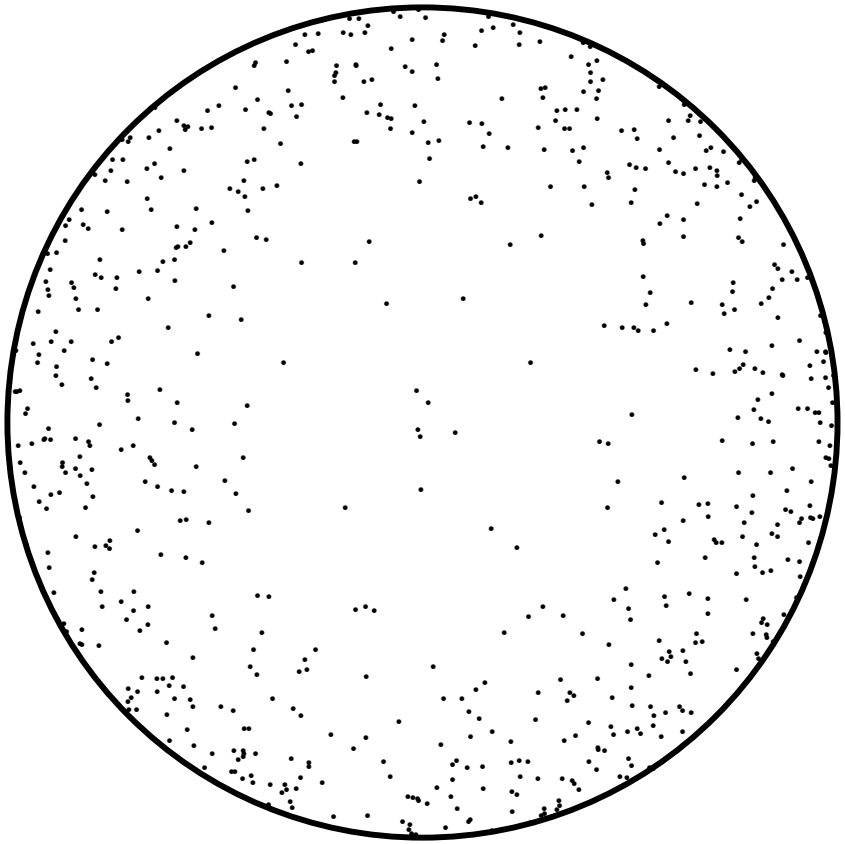
Random Geometric Graphs

Consider a slab of de Sitter spacetime: $t \in [0, t_0]$, $\theta \in [0, 2\pi)$



Random Geometric Graphs

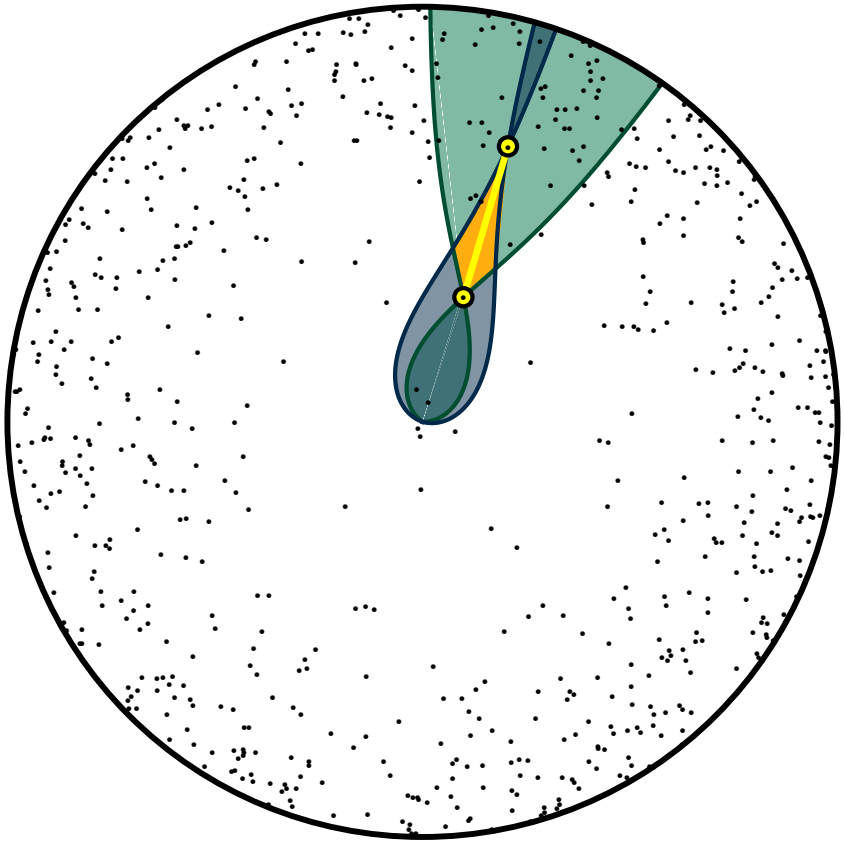
Perform a Poisson "sprinkling" of $\bar{N} = 700$ points using $p(t, \theta)$



This Poincare representation (top-down view) uses $t \to r$

Random Geometric Graphs

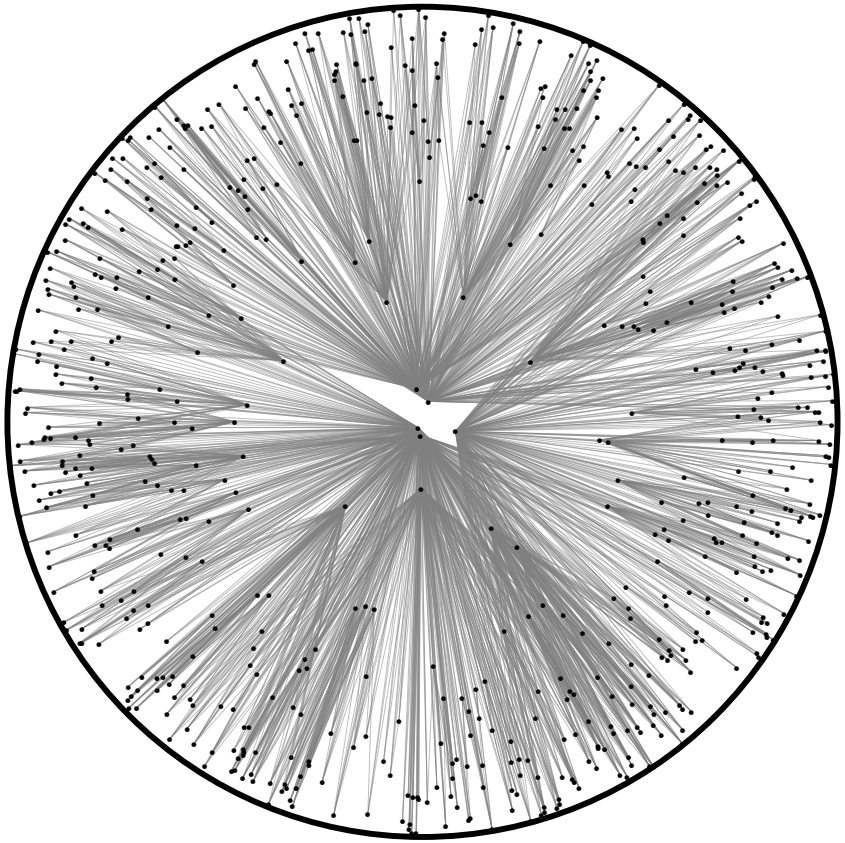
Draw the light cones for a pair of nodes



If they overlap, the pair is causally related

Random Geometric Graphs

Repeat for all pairs to get one graph sample.



What is the precise mathematical condition for linking?

Random Geometric Graphs

Use coordinates which make light cones simple:

- Use *conformal time*: $\cosh t = \sec \eta$
- $ds^2 = \sec^2 \eta(-d\eta^2 + d\theta^2)$
- When $\Delta\theta_{ij} < \Delta\eta_{ij}$, $ds^2 < 0$, add link $A_{ij} = A_{ji} = 1$
- $\Delta\theta_{ij} = \frac{\pi}{2} - \left|\frac{\pi}{2} - |\theta_i - \theta_j|\right|$ since boundaries are periodic
- $\Delta\eta_{ij} = |\sec^{-1}\cosh(t_i) - \sec^{-1}\cosh(t_j)|$

In general, conformally flat RGGs are much easier to generate.

Can you think of any methods to optimize this procedure?

Geometry from Order Theory

Order theory plays an important role in 2D.

- Lorentzian RGGs are directed and acyclic (partial orders)
- "Natural Labeling:" $i \prec j \Rightarrow \lambda(i) < \lambda(j)$
- The set of all natural labelings form the *linear extensions*

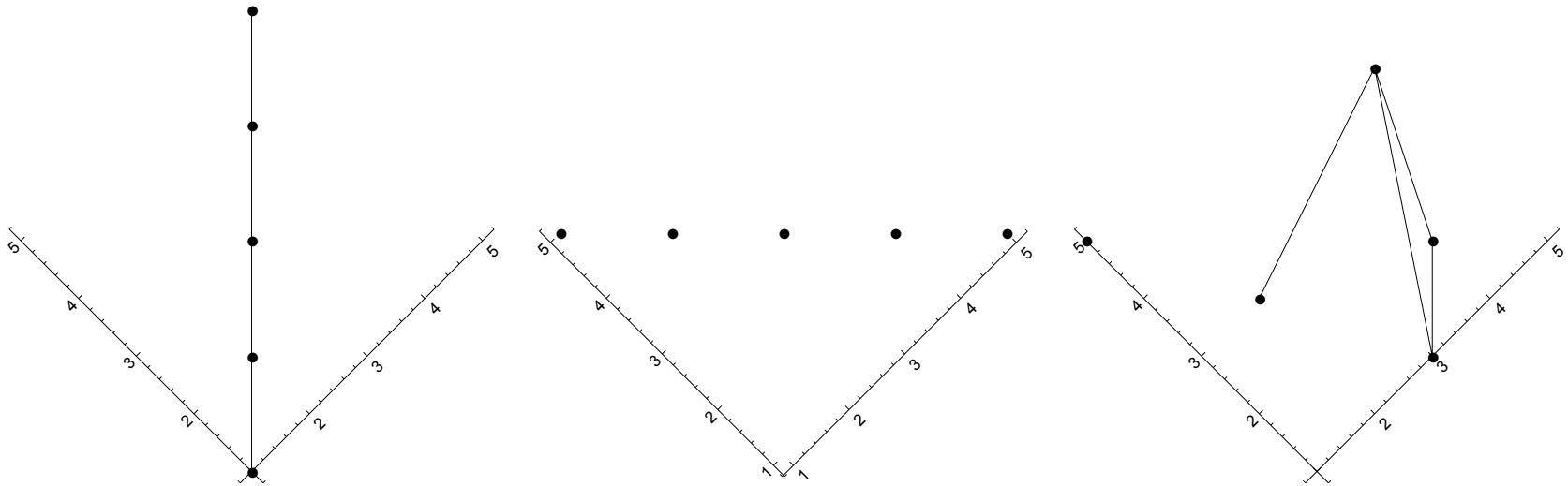
  Diamond: $\{1,2,3,4\}$ and $\{1,3,2,4\}$
- The 2d-orders are the graphs represented by intersections of at most two linear extensions λ, μ:

  $\lambda(i) < \lambda(j) \wedge \mu(i) < \mu(j) \Rightarrow i \prec j$
- A uniform measure over the 2d-orders is peaked around sprinklings into 2D Minkowski diamonds!

Unsolved: generalization to higher dimension.
Measuring order dimension for $d > 2$ is NP-complete

Geometry from Order Theory

Examples of 2D Orders:



The two linear extensions are the *light cone coordinates*.

Brightwell, Henson & Surya, Class. Quantum Grav 25 (2008).

Algorithms for Order Theory

What is the relationship between ensembles of Lorentzian sprinklings and ensembles of d-orders?

- Generate d_{max} vectors of length N, initialized to $\{1, \ldots, N\}$
- Fisher-Yates shuffle gives an $O(N)$ random permutation:

```
for (i = N - 1; i > 0; i--)
    swap(X[i], X[rng() * i])
```

- When an order relation agree among all d_{max} orders, add it to the graph
- Monte Carlo updates: swap an adjacent pair of entries in a random order. What about a "cluster" update?
- The height/width of the graph is the length of the longest ascending/descending sequence present in all orders

Exercise 1

1. Sprinkle $\bar{N} = 500$ points into a 3D Minkowski diamond using rejection sampling. First sprinkle uniformly (t, x, y) into a unit cube. The Minkowski diamond is a pair of cones $r = h = 0.5$ whose bases are touching. Reject points outside.
2. For each pair, link them if $|\Delta X| < |\Delta t|$. The former is the 2D Euclidean distance!
3. What is the mean number of relations per node? Is this value correlated with time? Why or why not? Perform an average over 20 realizations of the graph.
4. Now construct a 3D order with $N = 500$. Permute three vectors $\{1, \ldots, N\}$ using a F-Y shuffle. Construct an adjacency matrix for a graph, where $A_{ij} = 1$ if $\lambda_1(i) < \lambda_2(j)$ **and** $\lambda_2(i) < \lambda_2(j)$ **and** $\lambda_3(i) < \lambda_3(j)$. What is the mean number of relations? How does it compare to the RGG? Again average over 20 realizations.

Observables and Convergence

Geometric Observables

What geometric quantities can we measure?

- Dimension (Hausdorff, spectral, Myrheim-Meyer)
- Geodesic distance: $d(x_i, x_j) = d_{ij}$ (weighted) or $d(x_i, x_j) = 1$ (unweighted)

  Shortest/longest path in Riemannian/Lorentzian graph
- Volume (width, height, cardinality)
- Ricci curvature (Ollivier-Ricci, Benincasa-Dowker)
- Extrinsic curvature, boundary geometry

Pre-geometric statistical quantities:

- Ordering fraction $R/\binom{N}{2}$
- Degree distribution
- Clustering (fraction of triangles)
- Spectral geometry

Dimension

Dimension estimation is in general a hard problem.

1. *Hausdorff Dimension* (Riemannian)

   Study the scaling of ball volumes

   $N(r) = r^{d_h} \Rightarrow d_h = \lim_{r\to\infty} \frac{\ln N(r)}{\ln r}$

2. *Spectral Dimension* (Riemannian, Lorentzian)

   Study the return time of a random walk

   $d_s = -2 \lim_{t\to\infty} \frac{\ln p_t}{\ln t}$

3. *Myrheim-Meyer Dimension* (Lorentzian)

   Study the fraction of relations

   $\frac{R}{N^2} = \frac{\Gamma(d+1)\Gamma(d/2)}{4\Gamma(3d/2)}$ (Exact for Minkowski diamonds)

To be certain, use multiple estimators & take ensemble averages

Geodesic Distance

Riemannian:

- Shortest paths are geodesics
- Unweighted distance: breadth first search (BFS), $O(R + N)$
- Weighted distance: Dijkstra, $O(R + N \log N)$, or A*
- The "stretch" is the ratio between graph and manifold geodesic distance. $\lim_{N \to \infty} \bar{s} = 1^+$

Lorentzian:

- Longest paths are (timelike) geodesics; use DFS
- Graph should be *topologically sorted*
- Depth first search, $O(R + N)$
- Minkowski: $T = a_d L / \ell$

Dijkstra's Algorithm

Let's study an example in detail.

1. Divide nodes into visited/frontier/unvisited sets. Set $d_i = \infty$ for all nodes.
2. Let the source be in the frontier, set $d_s = 0$.
3. Pick the node n_i in the frontier set with the smallest d_i. If this is the target, return d_i.
4. For each neighbor n_j of n_i in the frontier set, let $d_j = \min(d_j, d_i + d_{ij})$
5. Move n_i from the frontier to the visited set; add all unvisited neighbors n_j to the frontier set
6. Go to step 3

How would you change this for multiple targets? Check out A*.

Ricci Curvature

We use the Ollivier-Ricci curvature in Riemannian graphs.

- The Wasserstein metric on discrete distributions is related to the graph curvature. It identifies the minimum distance and joint probability measure given two marginal distributions on graph neighborhoods.
- Consider uniform distributions on two δ-balls $\{x_i \in \mathcal{N}(x)\}$ and $\{y_i \in \mathcal{N}(y)\}$. $m_x(x_i) = 1/|\mathcal{N}(x)|$ and $m_y(y_i) = 1/|\mathcal{N}(y)|$.
- Construct a discrete distance matrix with entries $d_{ij} = d(x_i, y_j)$
- Minimize $W = \sum_{i,j} d_{ij}\rho_{ij}m_x(x_i)$ such that

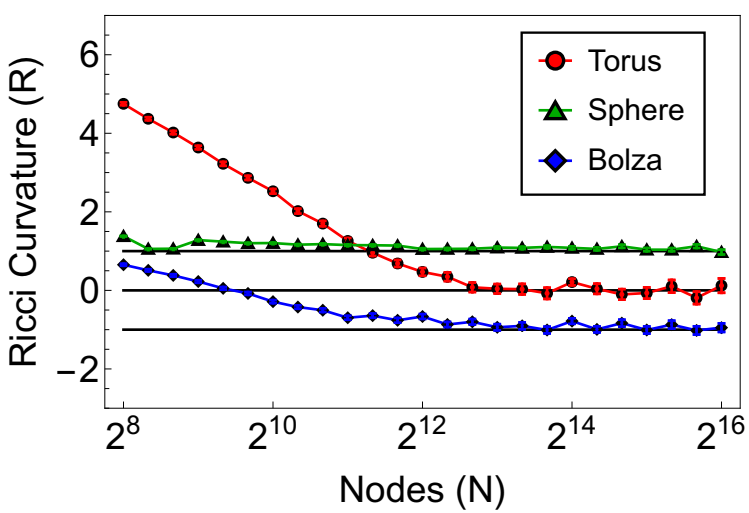
$$\sum_j \rho_{ij} = 1 \ \forall \ i \qquad 0 \leq \rho_{ij} \leq 1 \ \forall \ i,j$$

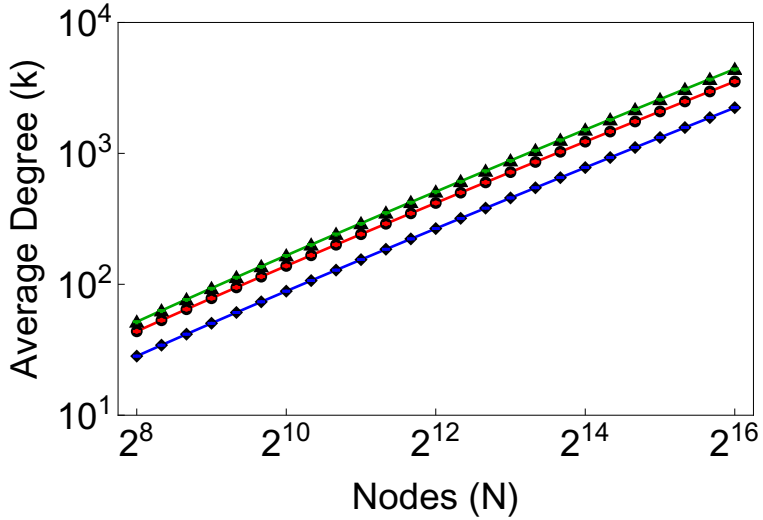
$$\sum_i \rho_{ij} m_x(x_i) = m_y(y_j) \ \forall \ j$$

Ricci Curvature

The Wasserstein distance W is related to the curvature R:

Ollivier: $\kappa = 1 - W/d$ Ricci: $R = \lim_{N\to\infty} \frac{2(2+d)}{\delta^2}\kappa$





Quasi-local curvature requires fairly large graphs.

Exercise 2

1. Construct a 2D Euclidean RGG. Sprinkle $N = 500$ points into a unit square and use the connection radius $r = 0.15$.
2. Estimate the Hausdorff dimension. First write a method to identify neighborhoods.
3. Estimate the clustering coefficient. This is the fractional number of triangles in the graph, $c = \frac{1}{2*3}\mathrm{Tr}A^3$. Can you guess how to measure the number of pentagons? What is the distribution of cycles of length n?
4. Estimate the length of the diagonal by measuring the longest path and multiplying it by the discreteness scale $\ell = N^{-1/2}$.
5. For each pair of nodes, divide the shortest weighted path in the graph by the Euclidean distance. Take an average over all pairs to measure the stretch.

Statistical Physics of Random Graphs

Energy/Action of Graphs

We can form a Gibbs measure on graphs:
$H(G) = aN + bR$ gives a weight a to nodes and b to relations

- Partition Function: $Z(\beta) = \sum_{G \in \mathcal{G}} e^{-\beta H(G)}$
- Analytic Continuation: $Z(\beta) \to \mathcal{Z}(\beta) = \int \mathcal{D}[g_{\mu\nu}] e^{iS(g_{\mu\nu})/\hbar}$

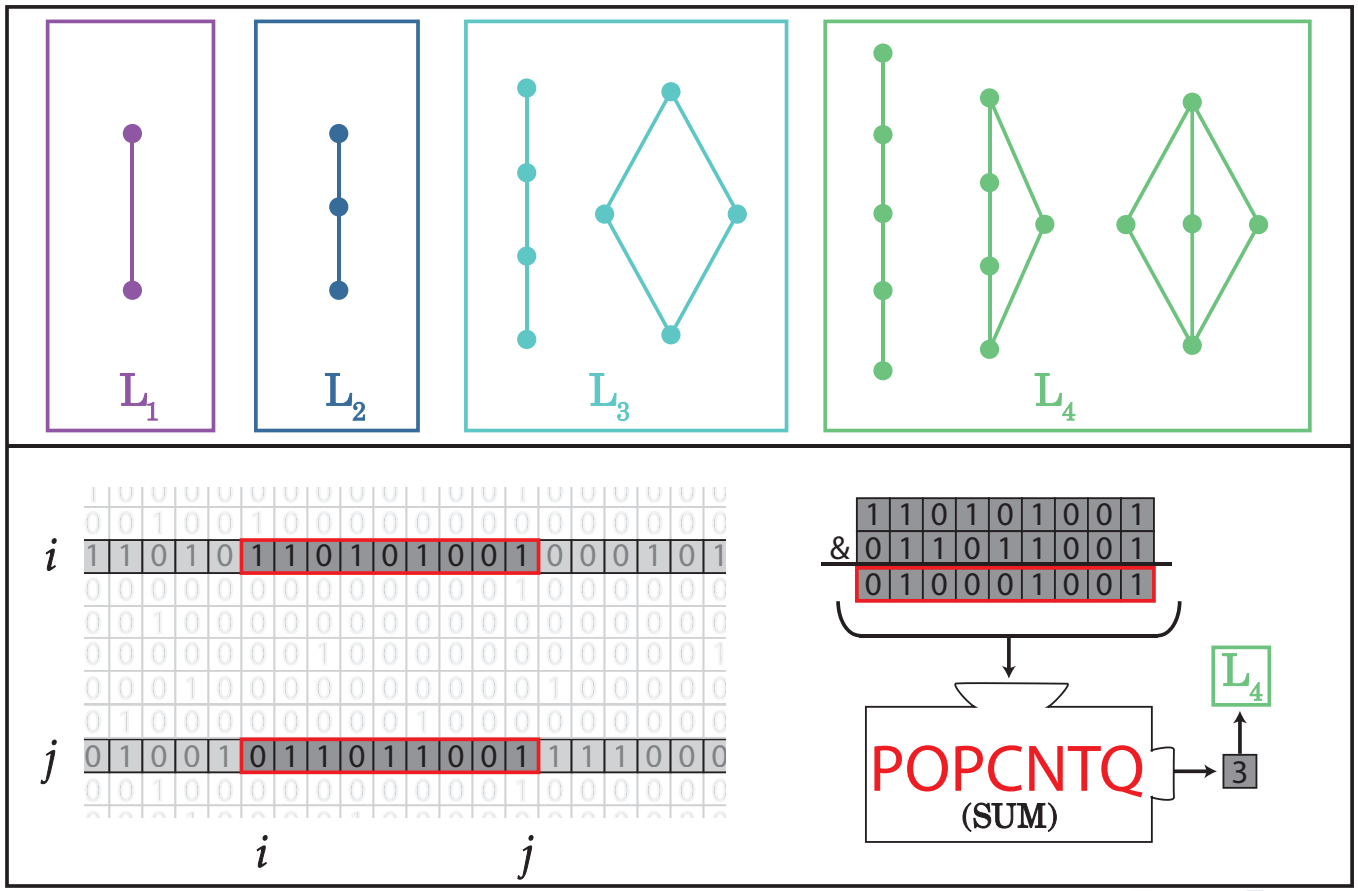
Unsolved: which choice of $H(G)$, $S(G)$ leads to the right measure?
Important features:

- Dynamics produces a second-order phase transition
- At least two phases: geometric, random
- Scale-free structure at large N at the critical point
- Renormalization flow of coefficients a, b
- Dimension is selected by the dynamics

Energy/Action of Graphs

For Lorentzian graphs, we use the Benincasa-Dowker action.

$$Z(N,d,\mathcal{T}) = \sum_{G\in\mathcal{G}} e^{-\beta S_d(G)/\hbar} \qquad S_2(G) = 2(N - 2n_1 + 4n_2 - 2n_3)$$



Monte Carlo Methods

Consider the Metropolis Monte Carlo method:

1. Pick an initial configuration, e.g., a random graph.
2. Identify your ergodic update. Link addition/removal is the simplest.
3. Choose a Hamiltonian. Pick a linear function of subgraphs with some coefficients.
4. Accept an update with probability $p = min(1, e^{-\beta \Delta E})$
5. Record your graph observables every sweep for 1000 sweeps.

   Unlike the Ising model, we need $O(N^2)$ updates per sweep. Why?

What are the properties of a typical configuration selected by your Hamiltonian? Do they change with β?

Monte Carlo Methods

How do we filter out correlated data?

- Demonstrate ergodicity using different initial configurations
- Measure the autocorrelations τ_{exp} and τ_{int}. Rule of thumb: discard first $20\tau_{\text{exp}}$ samples (thermalization). Independent samples are those $2\tau_{\text{int}}$ sweeps apart.
- Each observable has its own autocorrelations. When sampling/discarding, use the maximum.
- The simulation is complete when you have 10000 independent samples.
- Autocorrelation will be very large (1) near critical points and (2) at low temperatures. Scan β coarsely.
- Look to advanced methods like parallel tempering and population annealing to decrease autocorrelation.

Geometric Reconstruction

Is the typical graph from MCMC also a typical graph from a sprinkling? What manifold? What is the statistical certainty?

- In general, *inverse spectral problems* are unsolved. There is no set of discrete observables which is a "smoking gun."
- In practice, we measure as many observables as possible, study their distribution and convergence with N ("finite-size scaling")
- Good observables are distributions: subgraph distribution, degree distribution, degree-degree correlations, eigenvalue spectrum, Laplacian spectrum

Possible solution: large deviations theory. What observable is "rare" for all ensembles but one? Size-X clique/cycle?

Exercise 3

1. Generate a random graph with $N = 50$: sprinkling, random A_{ij}, d-order
2. Write a subroutine which picks a random linked pair and a random unlinked pair. You will use this to move links randomly.
3. Consider a Hamiltonian $H(G) = aR + bn_\Delta$. Choose $a, b, \beta \in [-1, 1]$. Measure and record the initial energy.
4. Write a loop which repeatedly moves a link and accepts the update according to the Metropolis condition.
5. Record the energy every $N(N-1)/2$ updates. Continue until you have at least 10 values recorded.
6. Record the final configuration. Measure some of the observables we discussed earlier. Is your graph geometric?