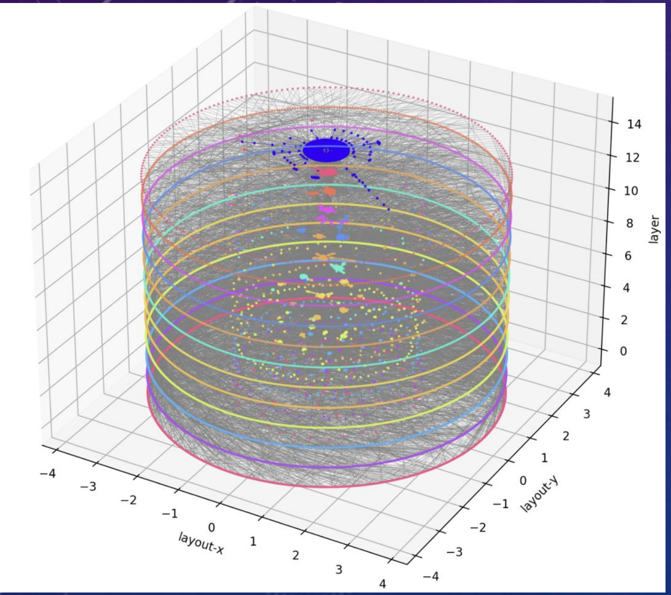
Mainly based on arXiv:2603.15128

第十八届粒子物理、核物理和宇宙学交叉学科前沿问题研讨会

基于暗晕图的星系形成路径似然表述
A Path-Likelihood Formulation of Galaxy Formation on Halo Graphs



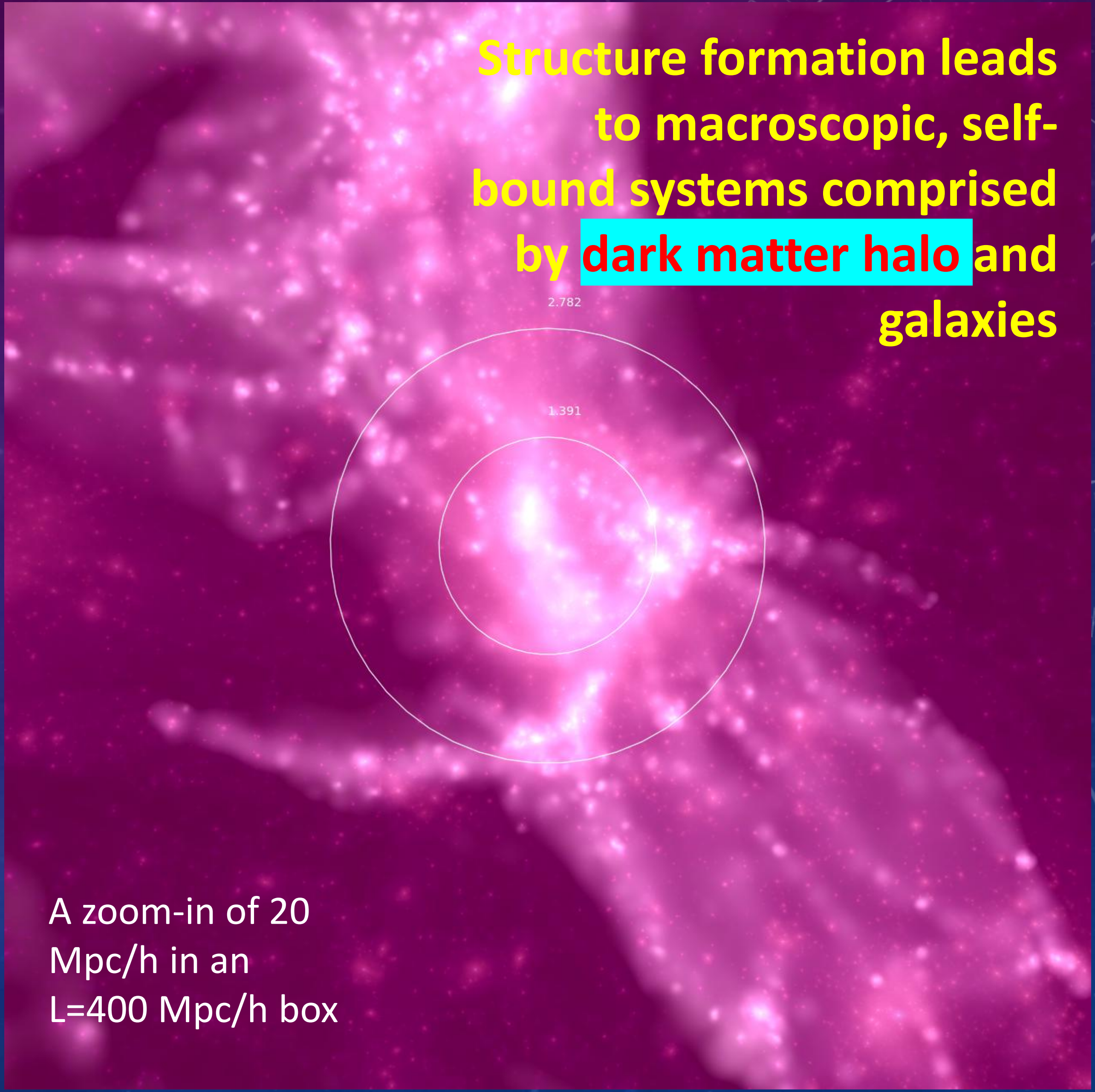
Daneng Yang (杨大能)
中国科学院紫金山天文台(PMO)

April 11, 2026
中国 桂林

https://github.com/DanengYang/GraphPathLikelihood

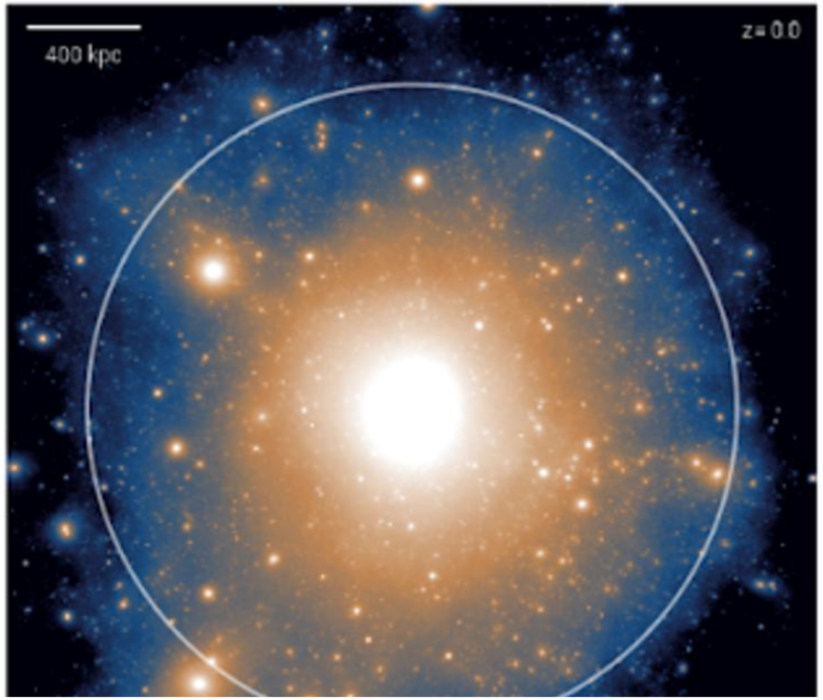
Outline

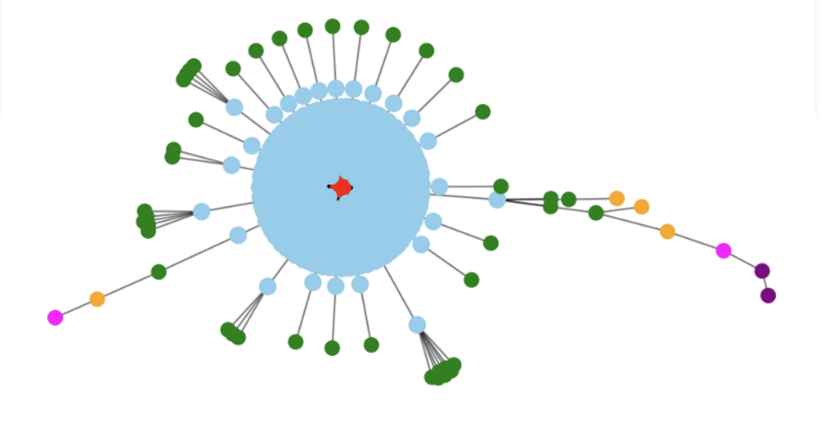
- Galaxy formation on Graphs
- A path measure model
- A machine learning realization
- Path-Space Tools
- Outlook



Can we write galaxy formation as a path probability?

The most massive host in TNG-50-1



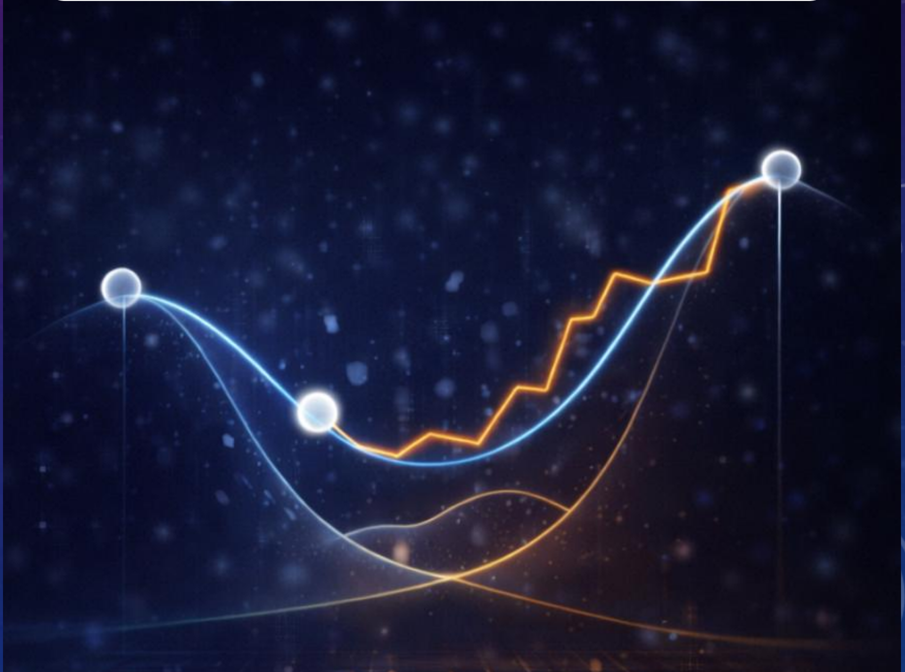


How rare is a given galaxy history?

How do we compare two different evolutionary paths?

What is the response of the system to controlled perturbations?

Forward models ≠ path measures



$$P(x|G) \propto p_{\text{attach}}(x|G)\, e^{-S(x;G)}$$

AI generated

Can we write galaxy formation as a path probability?

Galaxy formation = complex, high-dimensional system
Coarse-grained → stochastic effective dynamics
❗ But no path probability in standard models

Forward models ≠ path measures

Observables averaged over solutions of Langevin equation can be written as a path-integral

$$\dot{x}(t) = b(x(t)) + \eta(t)$$

The Langevin equation

The Martin-Siggia-Rose-Janssen-De Dominicis (MSRJD) formalism

$$Z = \int \mathcal{D}\eta\, \mathcal{D}x\, \mathcal{D}\hat{x}\; e^{\int dt\, \hat{x}^T(\dot{x}-b-\eta)}\; e^{-\frac{1}{2}\int dt\, \eta^T D^{-1}\eta}$$

Cosmological halo formation can be encoded by graphs

Yang & Yu 2022 PR Research:

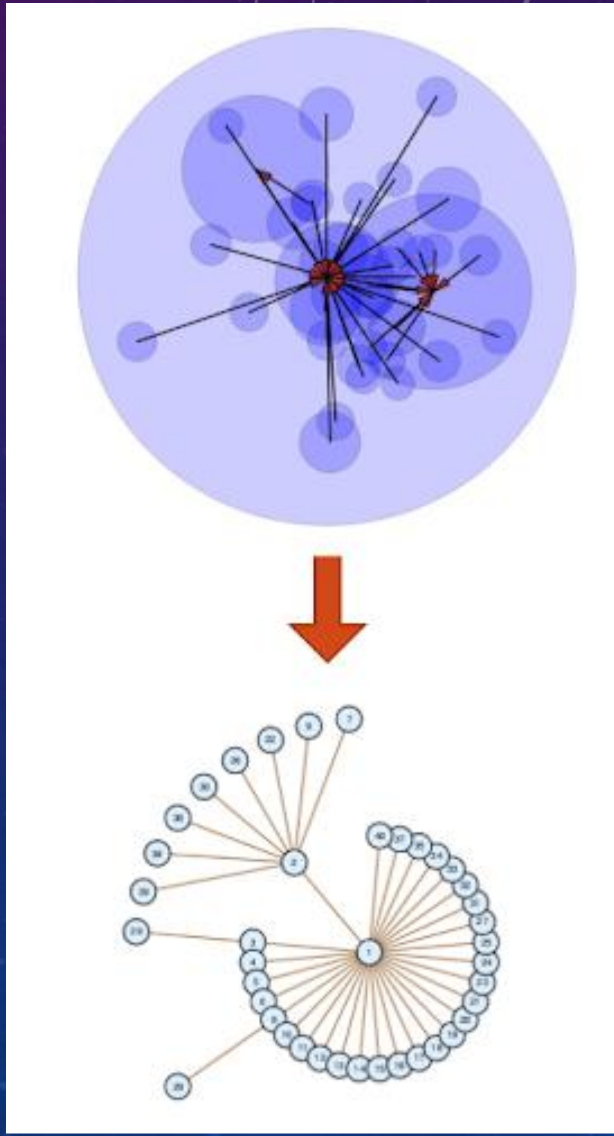
- Halo graphs can be constructed using **preferential attachment**
- They are power-law graphs

$$\text{Attachment Probability} = \frac{A_i}{\sum_j A_j}$$

Preferential attachment

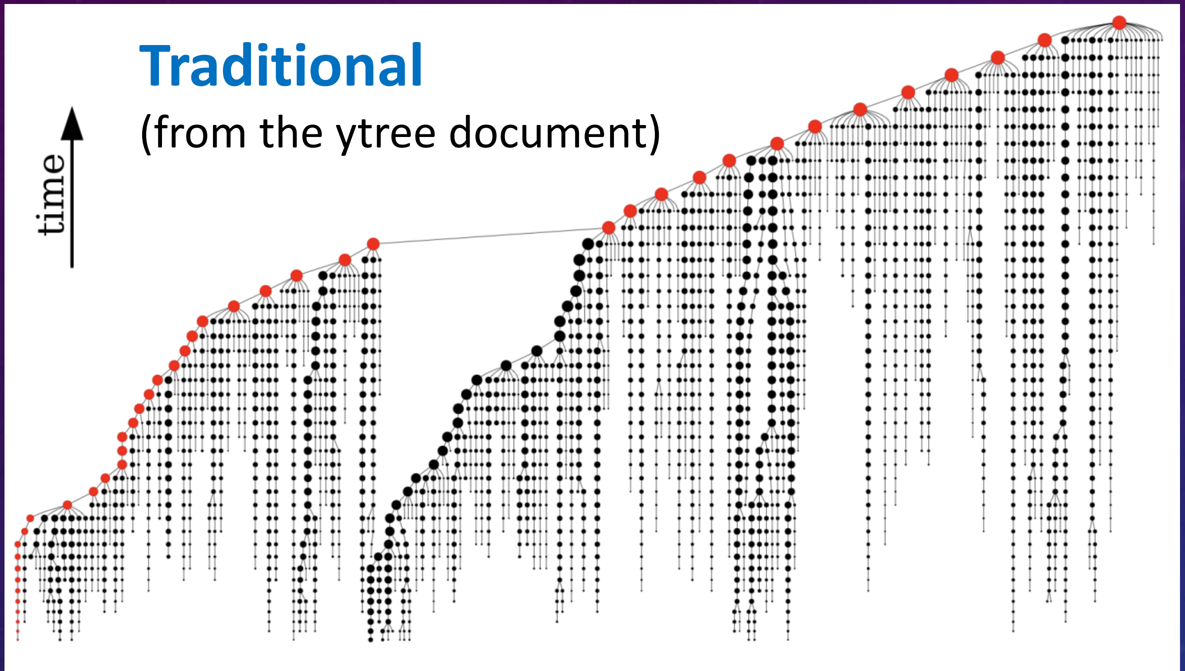
- Rich becomes richer
- Early attachment advantage

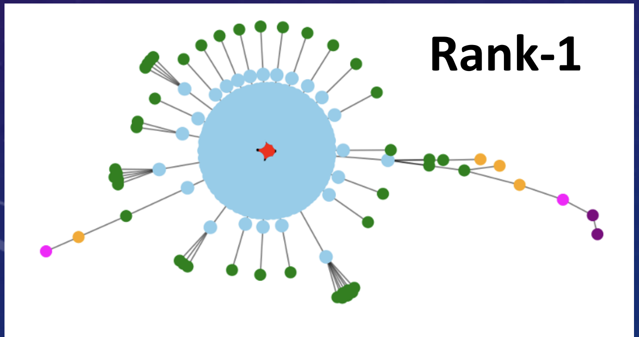
Self-similarity: Hierarchical structure naturally arise



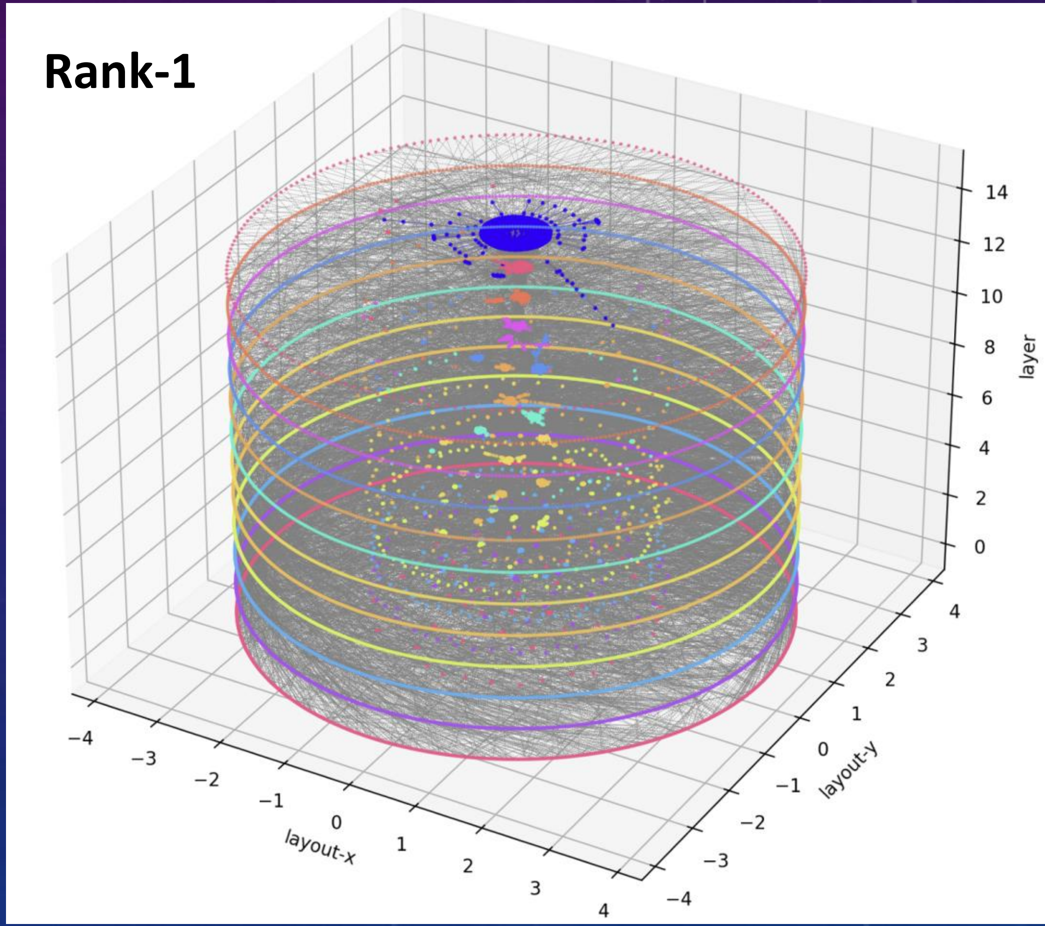
Recovering cosmic accretion history through detachment

One halo at z=0 has one merger tree





New: all halos belonging to an z=0 host halo has one layered halo graph



The probability-based construction naturally defines a graph measure P(G)

A Path Measure Model

Hierarchical structure formation

Graph-trajectory measure

$$P(\mathbf{x}, G) = P(\mathbf{x} \mid G) P(G)$$

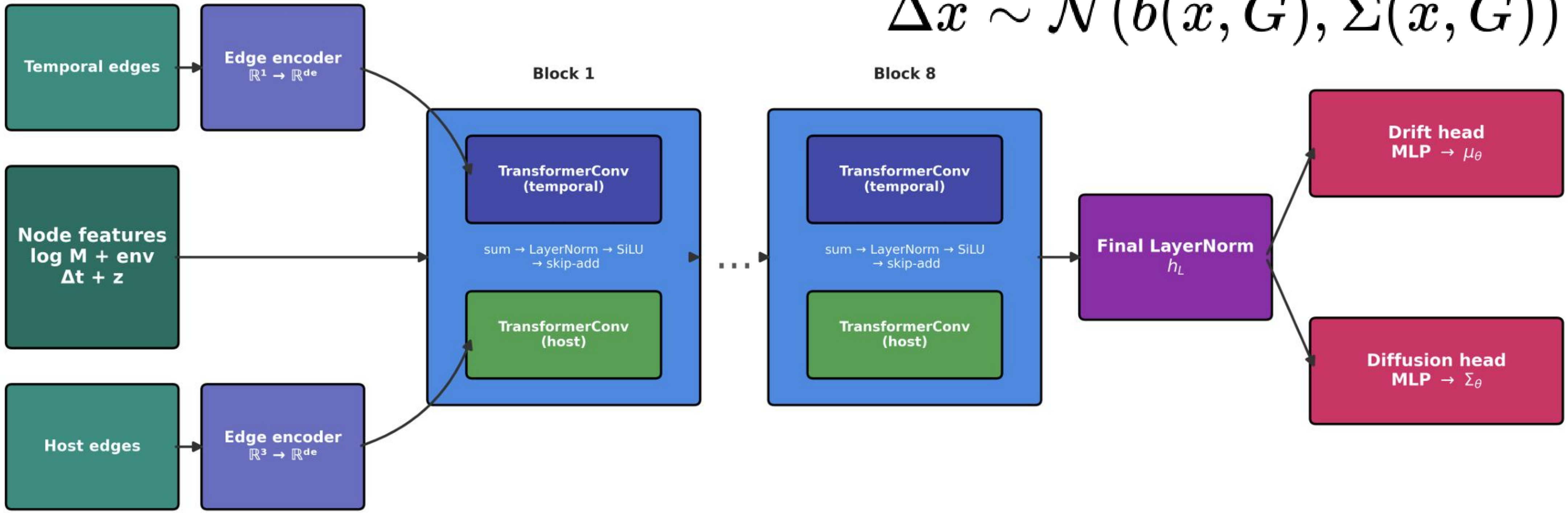
Graph-conditioned path measure

$$P(\mathbf{x} \mid G) \propto p_{\mathrm{attach}}(\mathbf{x} \mid G) \exp\left[-S(\mathbf{x}; G)\right]$$

Evolution in isolation

Evolution in environment

A Graph Path Likelihood Model (GPLM) from Machine Learning



$$\Delta x \sim \mathcal{N}(b(x,G), \Sigma(x,G))$$

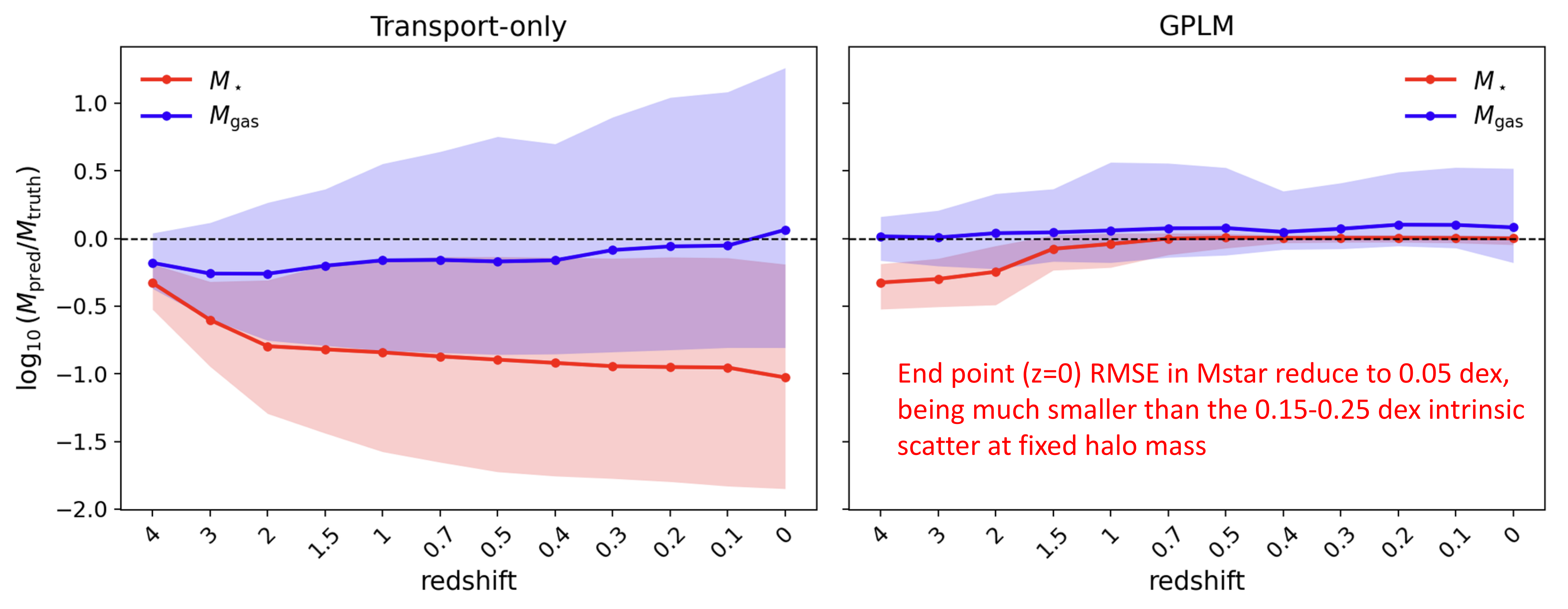
Graph encoding, conditioning

Attention-based Graph Neural Network

- b: learned drift
- D: learned diffusion

Model performance

GPLM substantially reduces both the median bias and the spread of the residual distribution.



MSRJD in GPLM: Discrete Realization of Stochastic Evolution

$$\Delta x_{i,k} \mid s_{i,k}, G \sim \mathcal{N}\left(\delta_{i,k}, \widehat{\Sigma}_{i,k}\right)$$

$$\delta_{i,k} \equiv b_{\theta,i,k} \Delta t_k \qquad \widehat{\Sigma}_{i,k} \equiv D_{\theta,i,k} \Delta t_k$$

$$\Delta x_{i,k} = b_{\theta,i,k}(s_{i,k}, G_k)\, \Delta t_k + \xi_{i,k}$$

$$\langle \xi_{i,k}\, \xi_{j,k'}^T \rangle = D_{\theta,i,k}\, \Delta t_k\, \delta_{ij}\, \delta_{kk'}$$

$$S_{\mathrm{MSRJD}}[x, \hat{x}; G] = \int dt \left[\hat{x}^T \left(\dot{x}_{\mathrm{res}} - b_\theta(x^{\mathrm{tr}}, G)\right) - \frac{1}{2}\hat{x}^T D_\theta(x^{\mathrm{tr}}, G)\hat{x}\right]$$

enforces the stochastic dynamics

Gaussian Onsager-Machlup (OM) action

$$S_{\mathrm{OM}}[\mathbf{x}; G] = \frac{1}{2} \sum_{(i,k)\in\mathcal{V}_{\mathrm{sup}}} \left[(\Delta x_{i,k} - \delta_{i,k})^{\mathsf{T}} \widehat{\Sigma}_{i,k}^{-1} (\Delta x_{i,k} - \delta_{i,k}) + \log\det \widehat{\Sigma}_{i,k}\right]$$

non-Gaussianity could be handled by higher-order terms

Path Space Tools

Operator average	Controlled likelihood deformation	Likelihood Ratio; forward-backward asymmetry

Applications to Dark Matter Microphysics

The effect of Self-Interacting Dark Matter (SIDM) could be understood as a controlled deformation of the path measure

We aim to improve/extend the parametric model for SIDM halos (Yang et al. 2305.16176)

See 项树城's talk

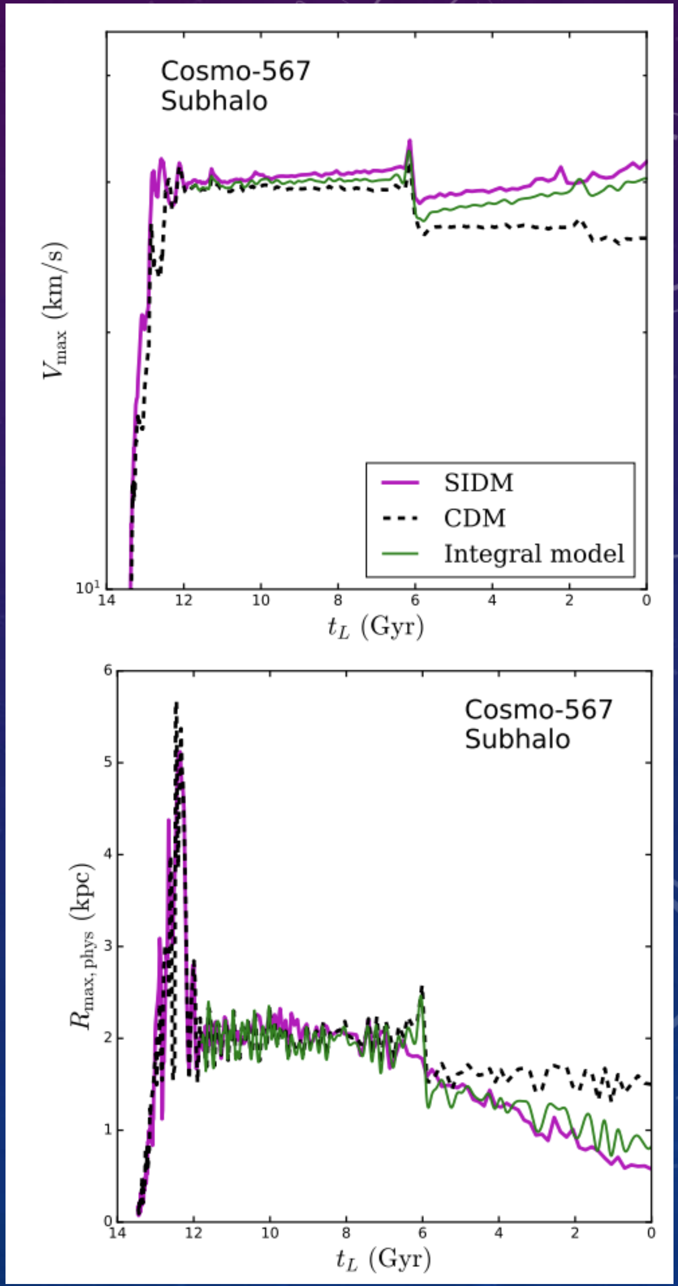
$$\mathbf{y}_k \equiv (V_{\max}, R_{\max})_k$$

$$\Delta \mathbf{y}_k = \Delta \mathbf{y}_k^{\mathrm{tr}} + \Delta \tau_k \mathbf{b}_{\mathrm{SIDM}}(s_k;\theta) + \boldsymbol{\epsilon}_k$$

(Non-linear) Transport

Residual Drift

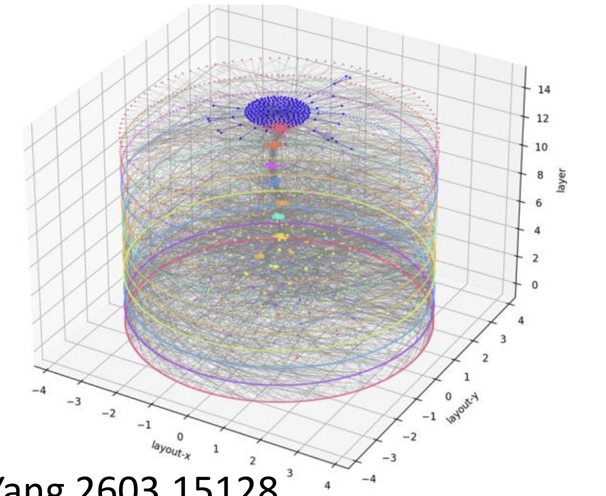
Residual Scatter



- ➢ A universal kernel approximates SIDM effects on (Vmax,Rmax) per time step
- ➢ Removing the time conditioning in bSIDM mirrors the universality

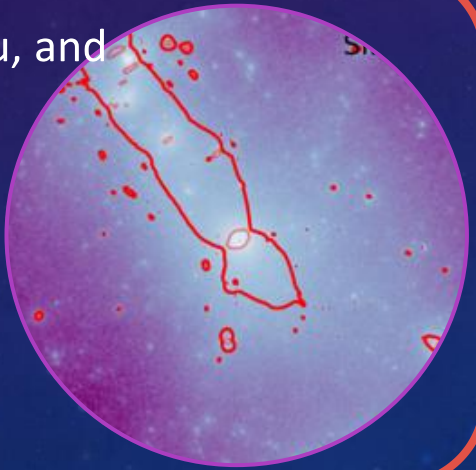
Our recent progress

A Graph Path Likelihood Model



Yang 2603.15128

Yang, Fan, Hou, and Tsai, **Science Bulletin 2026** & PRD **2025** on mass segregation



Elastic scatterings
- Gravothermal evolution

Topological and statistical correlations
- Correlation functions and Graphs

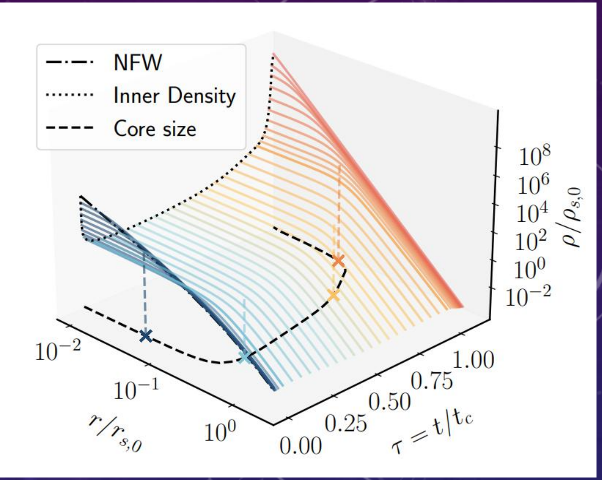
DM in a Path Action

Dissipative scatterings
- Boosted gravothermal evolution

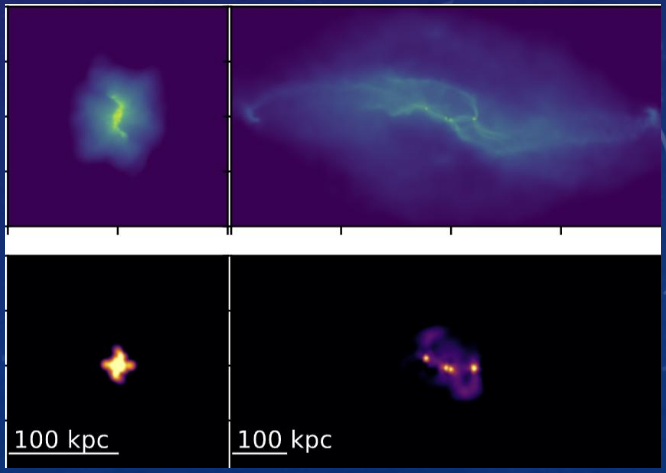
Multi-components
- Mass segregation

Self-boosting scatterings
- Gravitational binding

A parametric model for SIDM halos



https://github.com/DanengYang/parametricSIDM
Yang+2024 JCAP 02 (2024) 032
Yang 2024 PRD 110 (2024) 10, 103044
Yang+2025 PDU 47 (**2025**) 101807
Hou & Yang + JCAP08 (**2025**) 048



Collisional Formation of Baryon Dominated Dwarf Galaxies

Wang & Yang et al. **ApJL 2026**

Potential Theoretical Directions?

Non-Gaussian scatter

Optimal Transport

Geometric interpretation

The action can be rewrite as an energy functional of a curve in Riemannian geometry

$$S = \frac{1}{2}\int dt\, v^T g v$$

$$v \equiv \dot{x} - b(x)$$

$$\ddot{x}^i + \Gamma^i_{jk}\dot{x}^j\dot{x}^k = F^i(x, \dot{x})$$

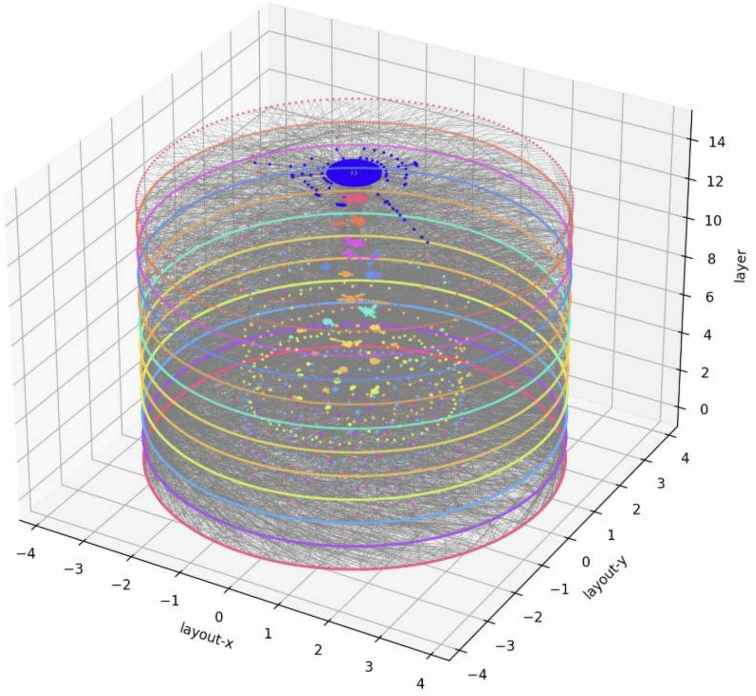
Most probable path=geodesic in g=D^{-1} under an effective potential

Galaxy evolution path=geodesic in environment-shaped metric+drift forcing

README

Graph Path Likelihood for Galaxy Formation on Layered Halo Graphs

- This repository presents the training and inference workflow used in https://arxiv.org/pdf/2603.15128
- The layered halo graph organizes the merger histories of a main halo and its subhalos into a single, connected layered object, explicitly encoding spatial connections via host edges and causal connections through temporal edges.
- The graph bundle under `dataGraphs/` contains the TNG-derived layered graphs used by the training and test lists. To save space, these graphs may be stored as individual `.json.gz` files. The convenience scripts below restore them automatically when needed. All graphs under `dataGraphs/` were constructed from the publicly available IllustrisTNG TNG50-1 data, which can be accessed at `www.tng-project.org/data`.



Rank-1 layered halo graph constructed from the TNG50-1 data.

For any questions, comments, and potential collaborations, please contact [yangdn_at_pmo.ac.cn].

Training & Inference codes available at https://github.com/DanengYang/GraphPathLikelihood

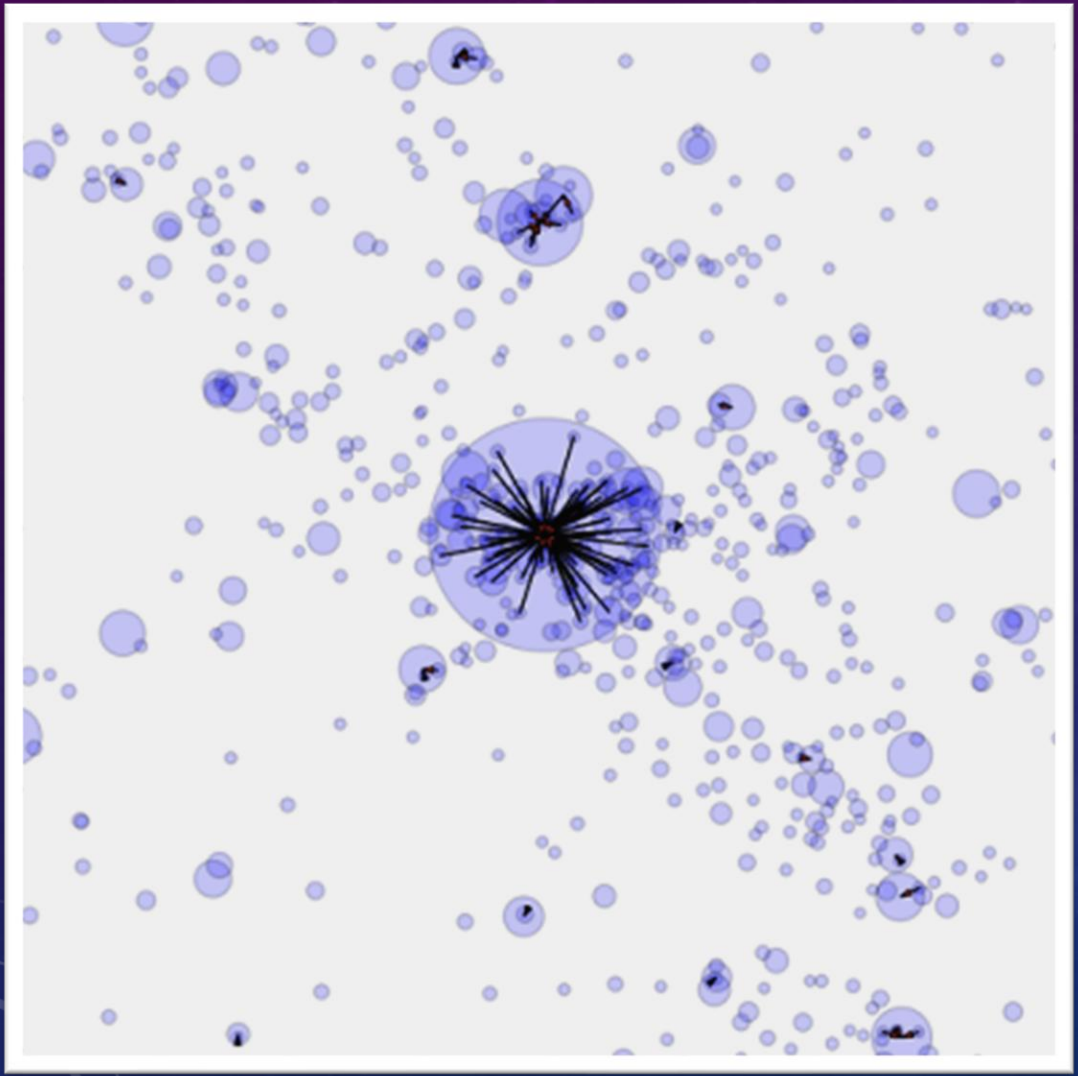
Training data converted from TNG-50-1 simulation data available under the *dataGraphs* folder

AI-drafted & AI-readable document facilitates further developments

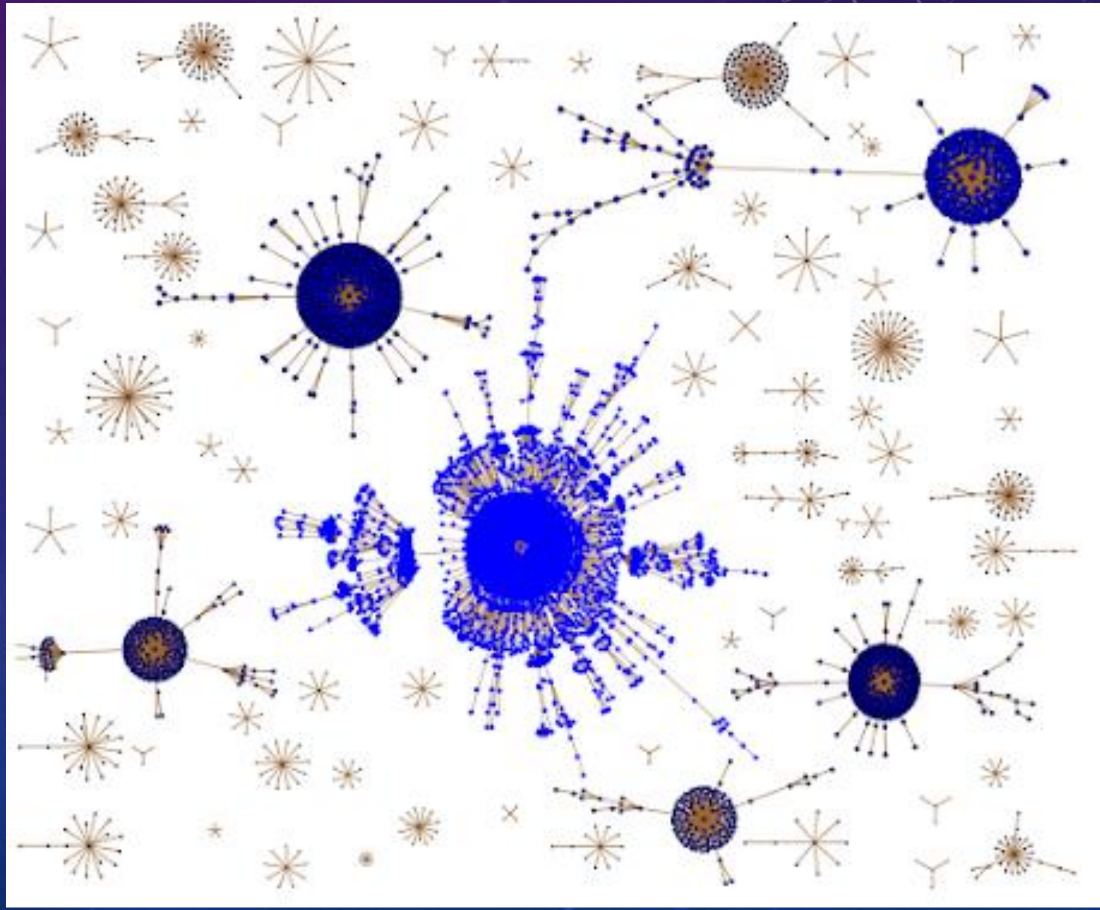
Thanks for your attention!

Back up

Cosmological halo formation can be encoded by graphs

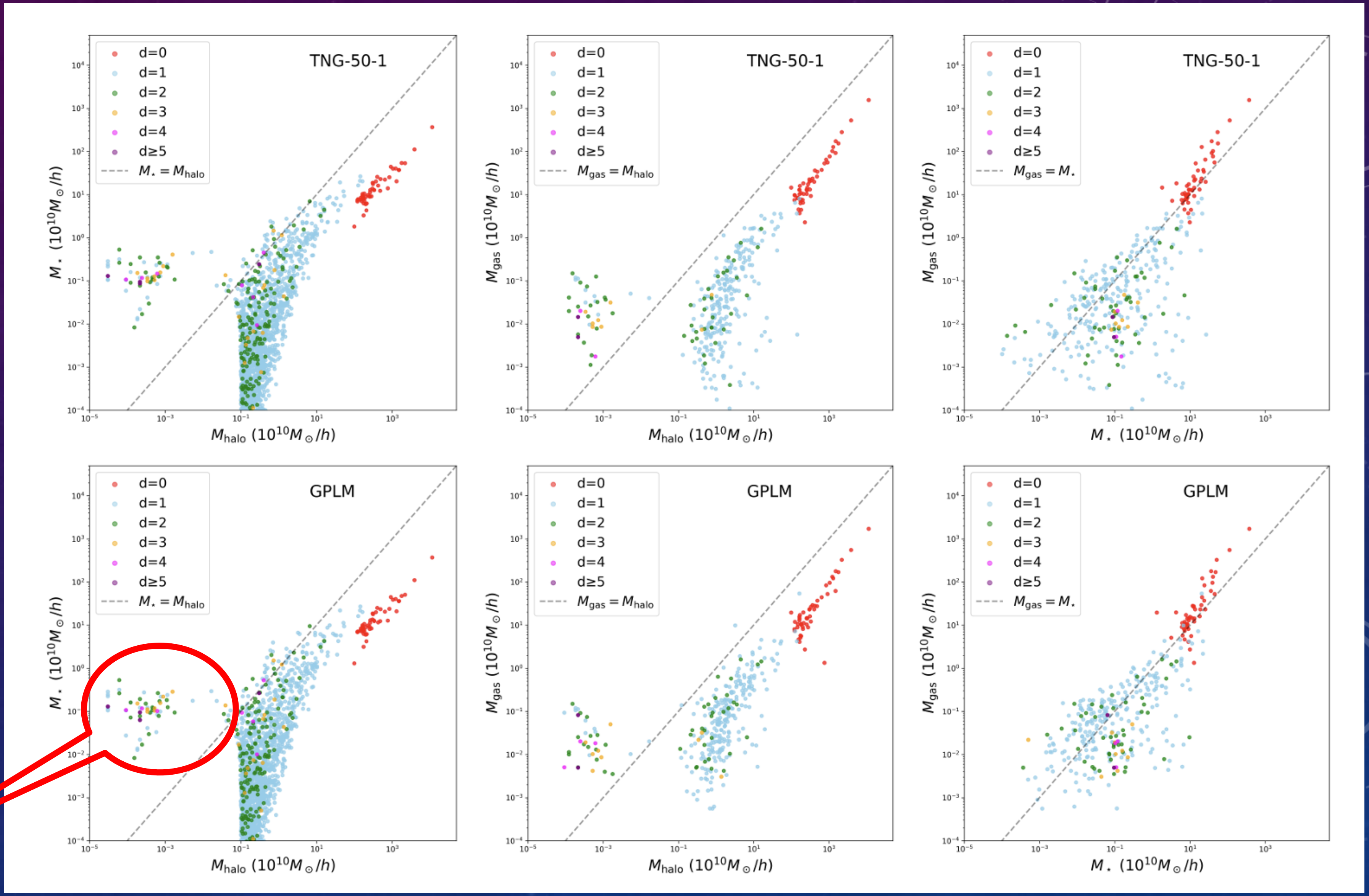


Yang & Yu, 2206.05578 [astro-ph.CO]
Phys. Rev. Research 5, 043187 (2023)



Model performance

The trained GPLM reliably reproduces the dG-split population structure observed in the simulation



Dark Matter Deficient Galaxies

Dark Matter Deficient Galaxy (DMDG) Probabilities

DMDG operator:

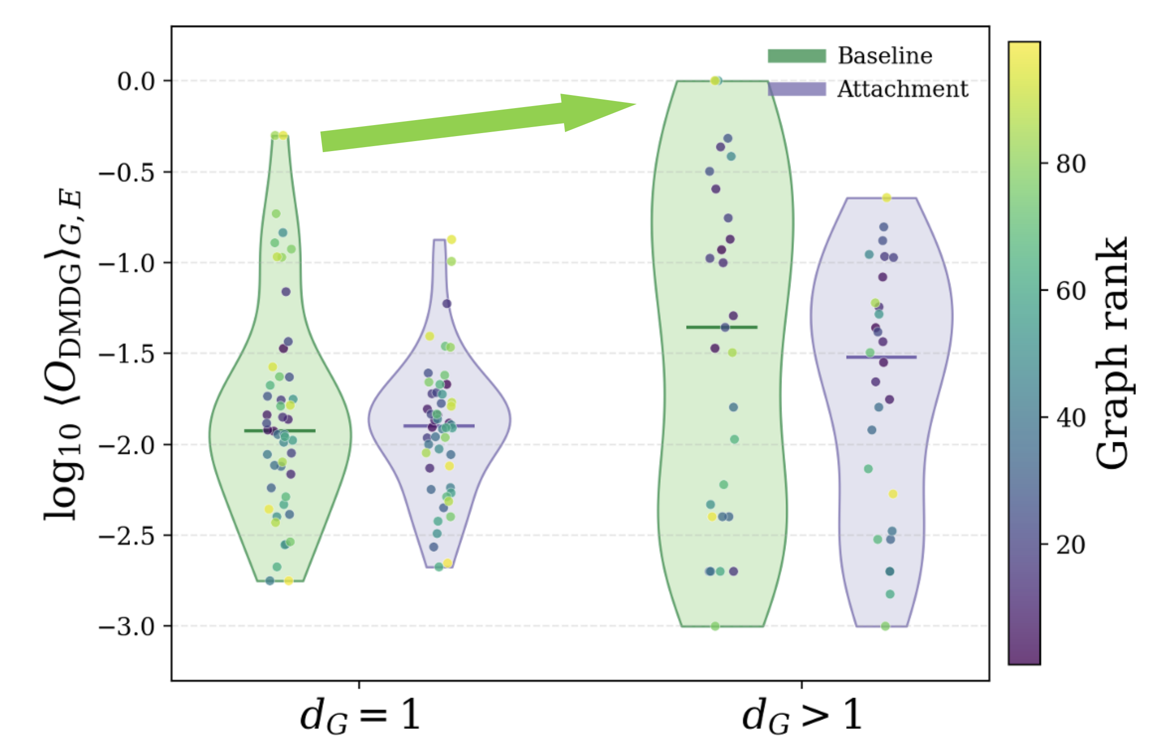
$$O^{(i)}_{\mathrm{DMDG}}(\mathbf{x}_i) = \Theta\left(\frac{M^{(i)}_\star(z=0) + M^{(i)}_{\mathrm{gas}}(z=0)}{M^{(i)}_{\mathrm{halo}}(z=0)} - 1\right)$$

dG=1: satellites
dG>1: higher order satellites

Higher tidal stripping in dG>1 satellites leads to more DMDGs but with greater scatter

Operator average

$$\langle O_{\mathrm{DMDG}} \rangle_{G,E} \equiv \int \mathcal{D}x \left[\frac{\sum_{i\in E} O^{(i)}_{\mathrm{DMDG}}[\mathbf{x}]}{N_E}\right] P_{\mathrm{tot}}(\mathbf{x} \mid G, x_0)$$



Gas-Rich response as a controlled likelihood deformation

Gas-Rich operator: $O_{\rm GR}(i) \equiv \Theta[M_{{\rm gas},i}(z{=}0) - M_{\star,i}(z{=}0)]$

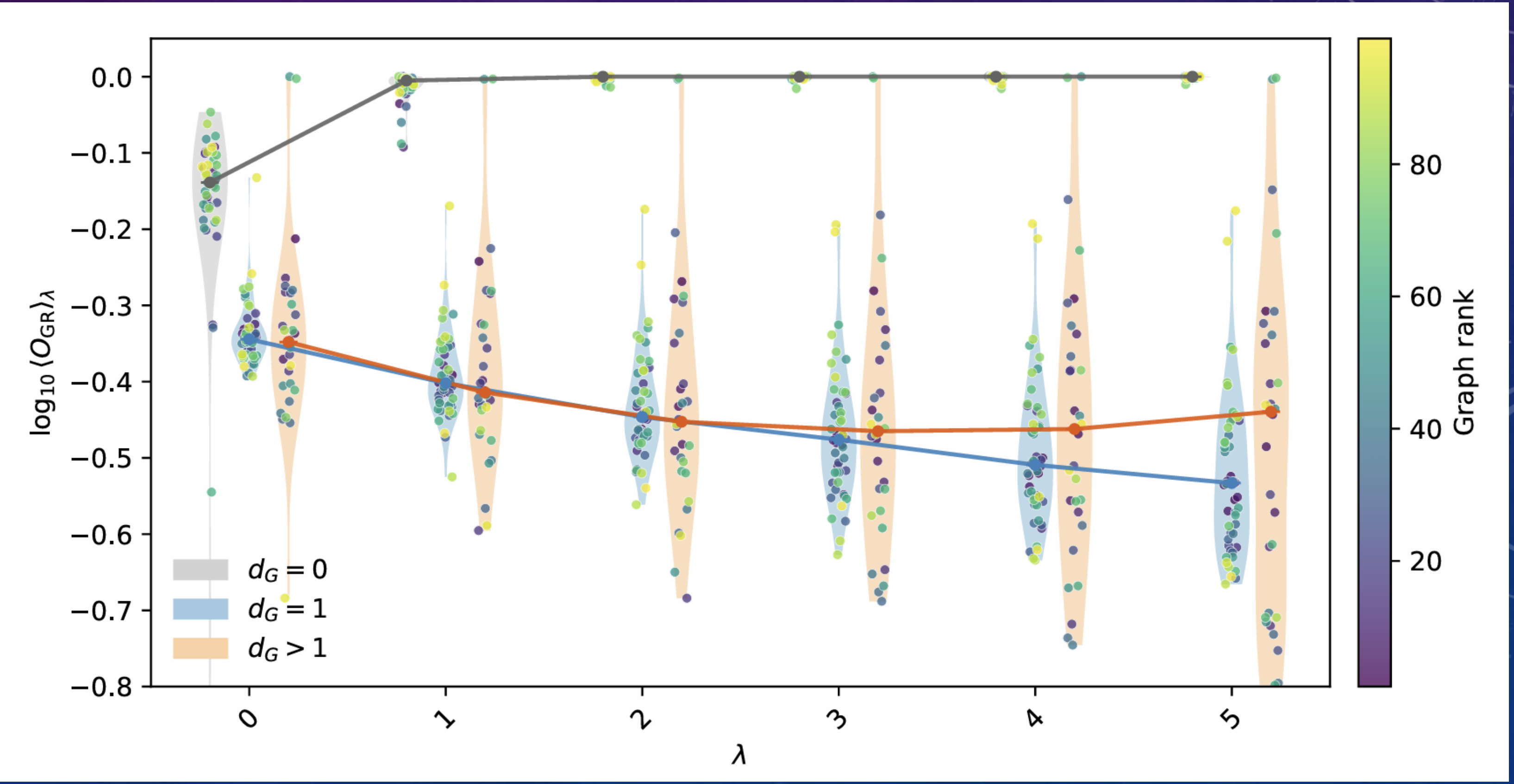
Controlled deformation: $b_\lambda(x_k, G) \equiv b_{\rm trans}(x_k, G) + \lambda \Pi_{\rm gas} b_{\rm host}(x_k, G).$

Ram-pressure stripping causes gas of satellites to feed into the host

dG=0: host

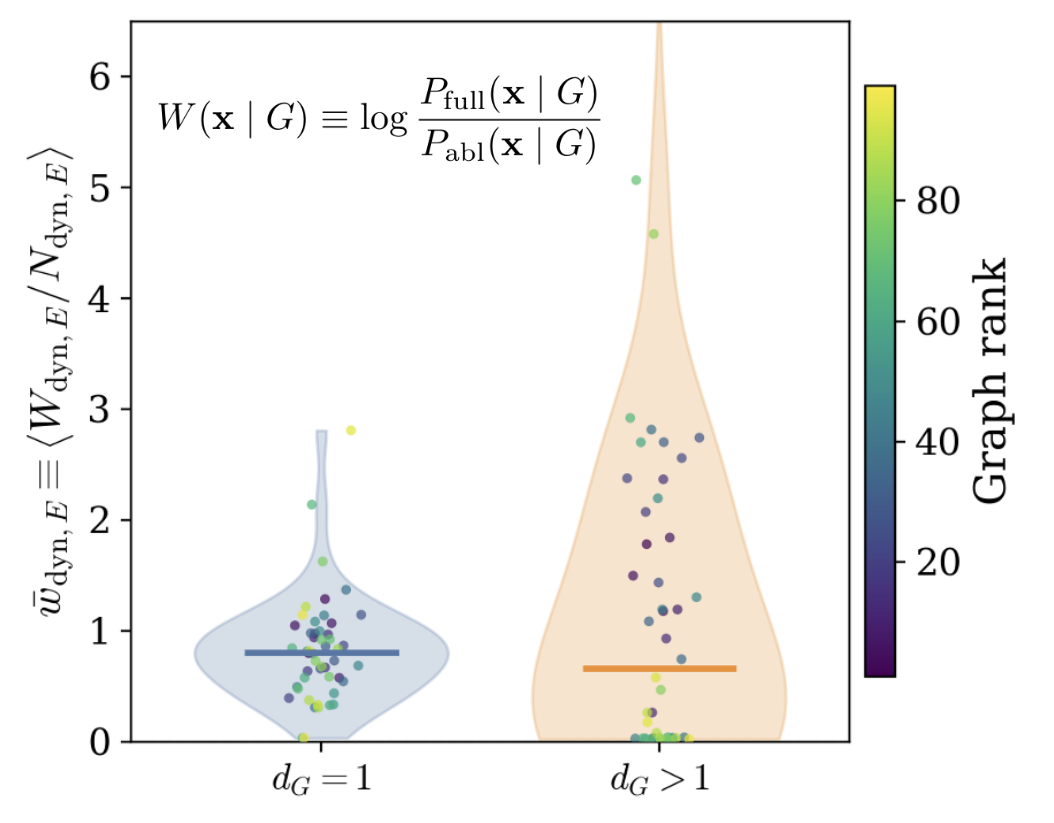
dG=1: satellites

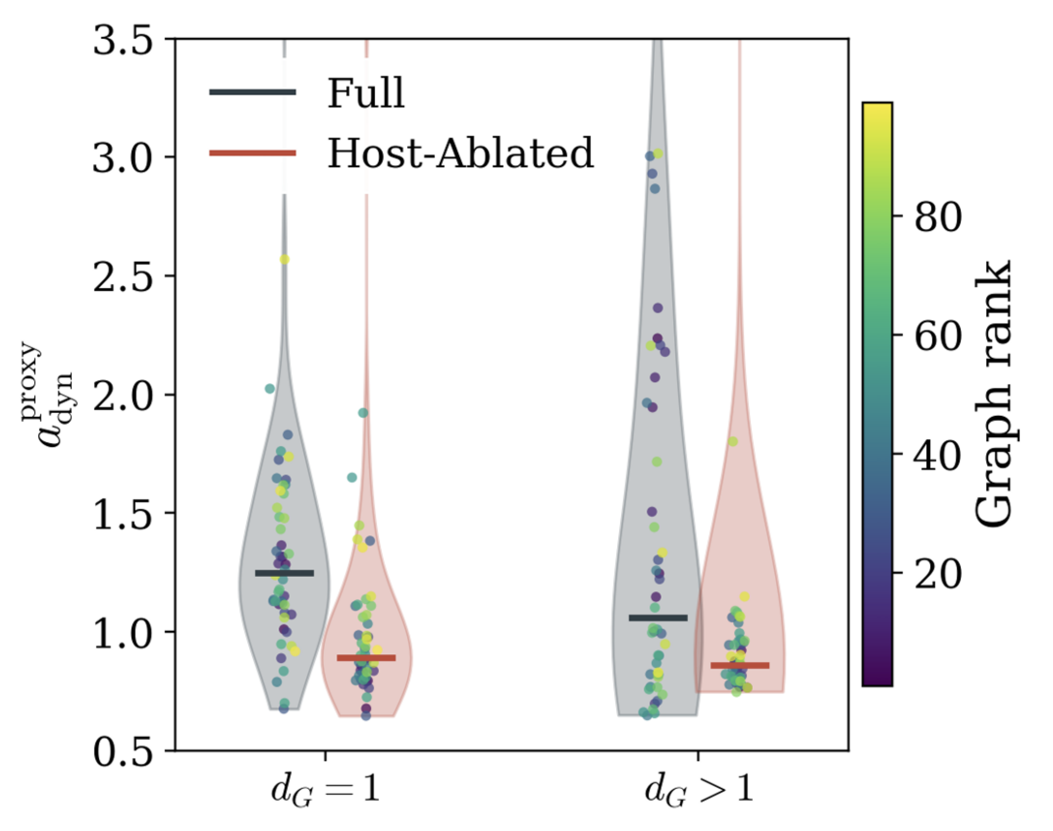
dG>1: higher order satellites



Path-Space likelihood diagnostics

$$\Delta s^{\text{proxy}}_{\text{dyn},i,k} = \log \frac{p^{\text{fwd}}_{i,k}}{p^{\text{rev,proxy}}_{i,k}}$$





Host-edge has a more diverse effect in those satellites

Host-edge conditioning weakens local forward–reverse alignment and increases graph-to-graph variability.

Stochastic field theory: noise-induced scatter

Aspect	QFT	SFT on layered graphs
Origin of scatter	Quantum fluctuations	Coarse-grained / environmental noise
Path integral weight	e^{iS}	e^{-S}
Nature	Coherent (phase)	Probabilistic
EOM	Deterministic + fluctuations	Intrinsically stochastic
Correlators	Propagators	Noise-driven correlations
Geometry	Spacetime	Noise structure
Mean of scatter	Quantum variance	Classical ensemble variance

MSRJD FROM NOISE MEASURE (1)

$$\dot{x}(t) = b(x(t)) + \eta(t)$$

$$Z = \int \mathcal{D}\eta \; P[\eta]$$

$$P[\eta] \propto \exp\left[-\frac{1}{2}\int dt\, \eta^T D^{-1} \eta\right]$$

$$1 = \int \mathcal{D}x \; \delta[\dot{x} - b(x) - \eta]$$

$$\delta[\dot{x} - b - \eta] = \int \mathcal{D}\hat{x} \; \exp\left[\int dt\, \hat{x}^T (\dot{x} - b - \eta)\right]$$

$$Z = \int \mathcal{D}\eta\, \mathcal{D}x\, \mathcal{D}\hat{x}\; e^{\int dt\, \hat{x}^T (\dot{x} - b - \eta)}\; e^{-\frac{1}{2}\int dt\, \eta^T D^{-1} \eta}$$

MSRJD FROM NOISE MEASURE (2)

$$\int \mathcal{D}\eta\, e^{-\frac{1}{2}\eta^T D^{-1}\eta - \hat{x}^T\eta} \;\propto\; \exp\left[-\frac{1}{2}\int dt\, \hat{x}^T D \hat{x}\right]$$

$$Z = \int \mathcal{D}x\, \mathcal{D}\hat{x}\; e^{-S_{\mathrm{MSRJD}}[x,\hat{x}]}$$

$$S_{\mathrm{MSRJD}} = \int dt\left[\hat{x}^T(\dot{x} - b(x)) - \frac{1}{2}\hat{x}^T D \hat{x}\right]$$

$$A \equiv (\dot{x}_{\mathrm{res}} - b_\theta)$$

$$-\frac{1}{2}\hat{x}^T D\hat{x} + \hat{x}^T A = -\frac{1}{2}(\hat{x} - D^{-1}A)^T D(\hat{x} - D^{-1}A) + \frac{1}{2}A^T D^{-1} A$$

$$\int \mathcal{D}\hat{x}\; \exp\left[-\frac{1}{2}(\hat{x}-\mu)^T D(\hat{x}-\mu)\right] \propto (\det D)^{-1/2}$$

$$Z \propto \int \mathcal{D}x\; \exp\left[-\frac{1}{2}\int dt\, A^T D^{-1} A\right] \times \exp\left[-\frac{1}{2}\int dt\, \log\det D_\theta\right]$$

MSRJD in GPLM

$$\dot{x}_{\mathrm{res}}(t) = b_\theta(x^{\mathrm{tr}}(t), G(t)) + \xi(t), \qquad \langle \xi(t)\xi(t')^T \rangle = D_\theta\, \delta(t - t').$$

$$S_{\mathrm{MSRJD}}[x, \hat{x}; G] = \int dt\, \left[\hat{x}^T (\dot{x}_{\mathrm{res}} - b_\theta(x^{\mathrm{tr}}, G)) - \frac{1}{2} \hat{x}^T D_\theta(x^{\mathrm{tr}}, G) \hat{x} \right]$$

$$S_{\mathrm{OM}}[x; G] = \frac{1}{2} \int dt\, (\dot{x}_{\mathrm{res}} - b_\theta)^T D_\theta^{-1} (\dot{x}_{\mathrm{res}} - b_\theta) + \frac{1}{2} \int dt\, \log\det D_\theta(x^{\mathrm{tr}}, G)$$

Discretization Matters: Same SDE, Different Path Measures

$$P[x] = \int \mathcal{D}\eta \, P[\eta] \, \delta[\dot{x} - b(x) - \eta]$$

$$P[x] = P[\eta = \dot{x} - b(x)] \times \det\left(\frac{\delta\eta}{\delta x}\right)$$

The definition of δ[ẋ−b(x)−η] is ambiguous without specifying discretization

$$x_{t+\Delta t} - x_t = b(x_\alpha)\,\Delta t + \sqrt{D(x_\alpha)}\,\Delta W$$

$$x_\alpha = (1-\alpha)x_t + \alpha x_{t+\Delta t}$$

- Itô: no Jacobian term
- Stratonovich: +(1/2)∇·b
- General α: continuous family

$$P[x] \propto \exp\left[-\tfrac{1}{2}(\dot{x}-b)^T D^{-1}(\dot{x}-b) - \tfrac{1}{2}\log\det D - \alpha \nabla \cdot b \right]$$

The most probable path

The action can be rewrite as an energy functional of a curve in Riemannian geometry

$$S = \frac{1}{2}\int dt\, v^T g v$$

$$v \equiv \dot{x} - b(x)$$

$$\ddot{x}^i + \Gamma^i_{jk}\dot{x}^j\dot{x}^k = F^i(x, \dot{x})$$

Most probable path=geodesic in g=D^{-1} under an effective potential

Galaxy evolution path=geodesic in environment-shaped metric+drift forcing

The most probable path equation is a diffusion equation

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0$$

$$\ddot{x} = (\nabla b)\dot{x} - D(\nabla b)^T D^{-1}(\dot{x} - b)$$

$$b(x) = -D\nabla\Phi(x)$$

Assuming constant D: $\ddot{x} = D^2\nabla\left[\frac{1}{2}|\nabla\Phi|^2\right]$

Hamiltonian structure

$$p \equiv D^{-1}(\dot{x} - b)$$

p encodes deviation from deterministic drift

$$\dot{x} = b(x, G, t) + \sqrt{D(x, G, t)}\,\eta$$

$$\dot{x} = b + Dp, \quad \dot{p} = -(\nabla_x b)^T p$$

Non-Gaussianities

$$\dot{x}_{\mathrm{res},i}(t) = b_i\big(x^{\mathrm{tr}}(t), G(t)\big) + \xi_i(t),$$

$$\langle \xi_i(t)\xi_j(t')\rangle = D_{ij}\big(x^{\mathrm{tr}}(t), G(t)\big)\,\delta(t-t').$$

$$S[x,\hat{x};G] = \int dt\,\left[\hat{x}_i\big(\dot{x}_{\mathrm{res},i} - b_i(x^{\mathrm{tr}},G)\big) - \frac{1}{2}\hat{x}_i D_{ij}(x^{\mathrm{tr}},G)\hat{x}_j\right] + S_{\mathrm{NG}},$$

$$D_{ij}(x,G) = L_{ik}(G)L_{jk}(G) + \mathcal{O}(x)$$

$$S_{\mathrm{NG}} = \int dt\,\left[\frac{1}{3!}C^{(3)}_{ijk}(G)\,\hat{x}_i\hat{x}_j\hat{x}_k + \frac{1}{4!}C^{(4)}_{ijkl}(G)\,\hat{x}_i\hat{x}_j\hat{x}_k\hat{x}_l + \cdots\right]$$

$$b_i(x,G) = A_{ij}(G)x_j + \frac{1}{2}B_{ijk}(G)x_jx_k + \frac{1}{3!}C^{(x)}_{ijkl}(G)x_jx_kx_l + \cdots$$

Non-Gaussian stochastic dynamics = interacting field theory in (x,p)

Connects to optimal transport / Wasserstein geometry

$\rho(x,t|G)$: the probability density (or empirical distribution) of galaxy states at layer t, conditioned on the graph/environment G

The same structure can be understood with optimal transport. In optimal transport (Benamou–Brenier):

$$\inf_{\rho,v} \int dt \int dx\, \rho(x,t)\, v^T M^{-1} v$$

Subject to (without drift):

$$\partial_t \rho + \nabla \cdot (\rho v) = 0$$

With drift:

$$\partial_t \rho = -\nabla \cdot (\rho b) - \nabla \cdot (\rho v)$$

Stochastic dynamics = gradient flow of free energy in Wasserstein metric

GPLM = graph-conditioned optimal transport