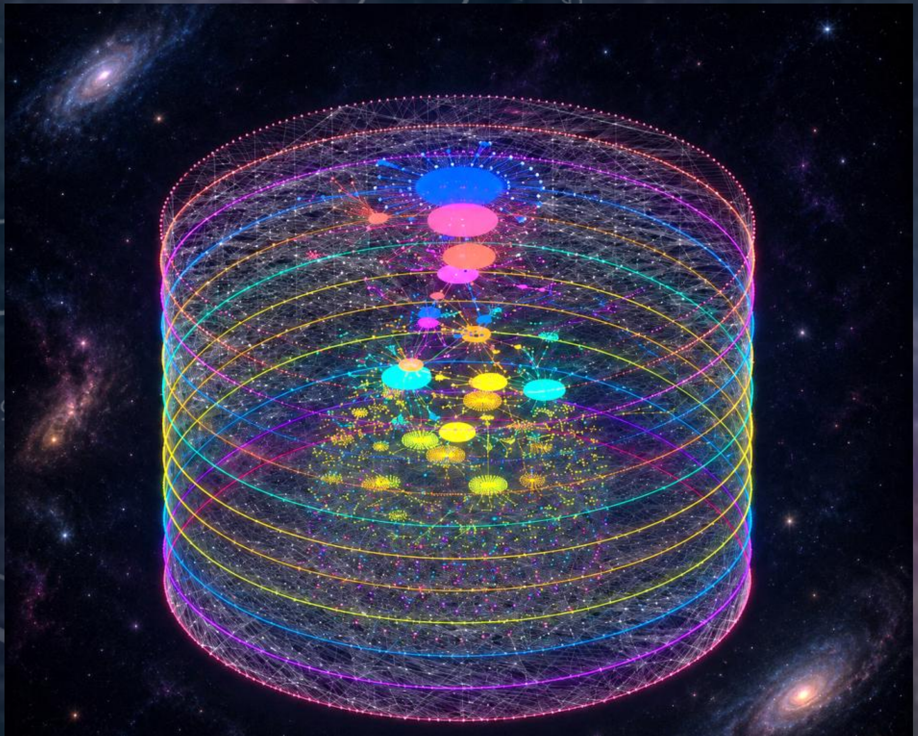
Mainly based on **arXiv:2603.15128**
arXiv:2206.05578 (PRR)
arXiv:2506.14898 (Sci. Bull.)
arXiv:2305.16176 (JCAP)

2026 紫金山暗物质研讨会

Path Measures for Stochastic Galaxy Formation on Layered Halo Graphs --- 星系形成的路径测度视角



https://github.com/DanengYang/GraphPathLikelihood

Daneng Yang (杨大能)
中国科学院紫金山天文台(PMO)

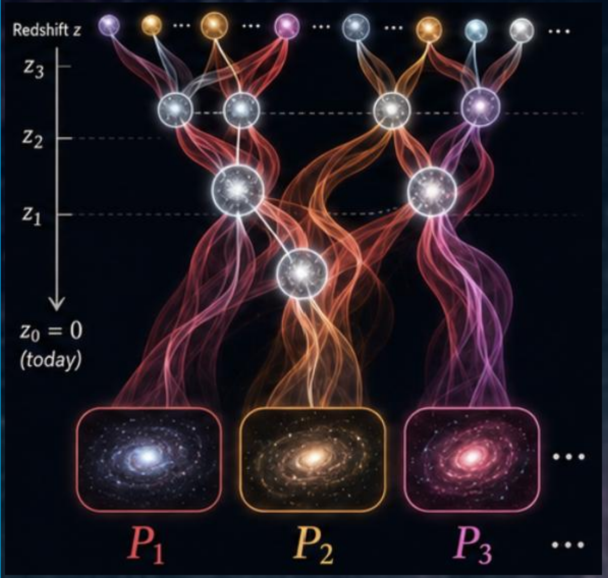
July 23, 2026
中国 南京

Outline

- Why do we need a path measure?
- Layered halo graphs and P(G)
- A path measure from training
- Example applications
- SIDM as controlled deformation on graph paths

A zoom-in of 20 Mpc/h in an L=400 Mpc/h box

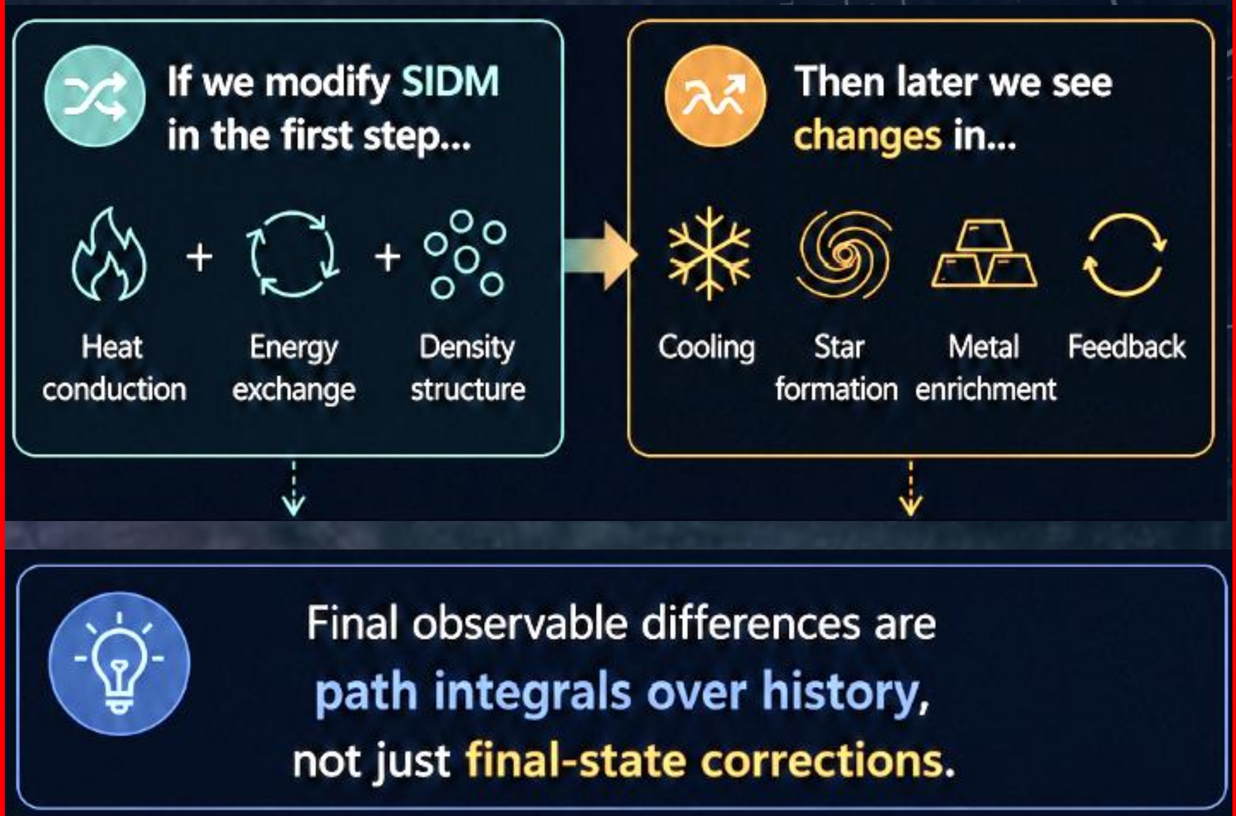
Why Think in Terms of Paths?



Given halo merger history, existing models can predict galaxies in halos

Forward models ≠ path measures

Similar galaxies can have different histories

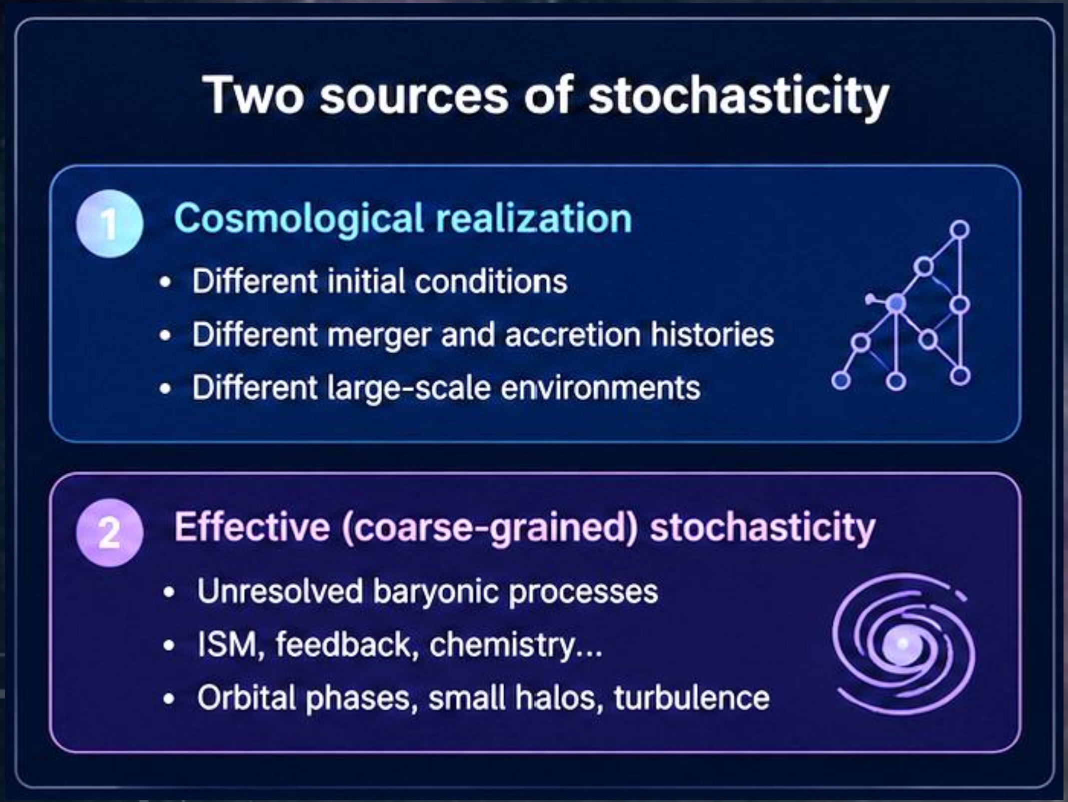


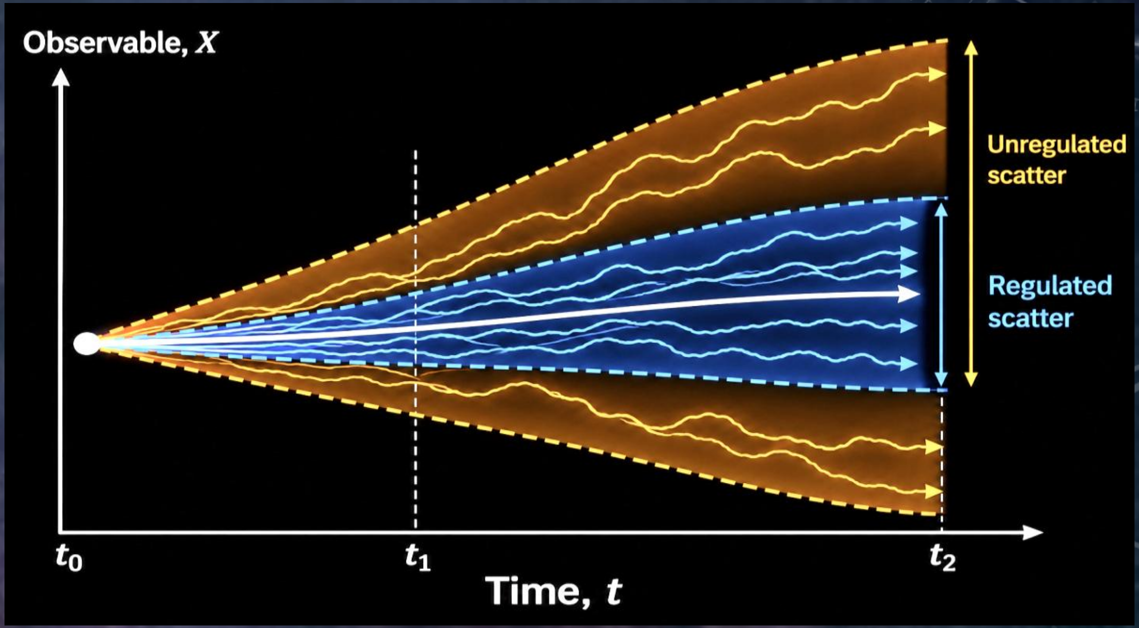
A self-consistent model for SIDM should evolve SIDM & galaxy formation simultaneously, in a way more close to doing path integrals

Galaxy formation along halo assembly has intrinsic scatters

Full simulation data → Snapshots → Merger trees & catalogs

Randomness is not put in by hand, it emerges after coarse-graining





Coarse-graining transforms deterministic microscopic dynamics into an effective stochastic dynamics

Model factorization

Graph: Hierarchical structure formation

Graph-trajectory measure

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

Graph-conditioned Path Measure

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

Evolution in isolation

Evolution in environment

Future extension:

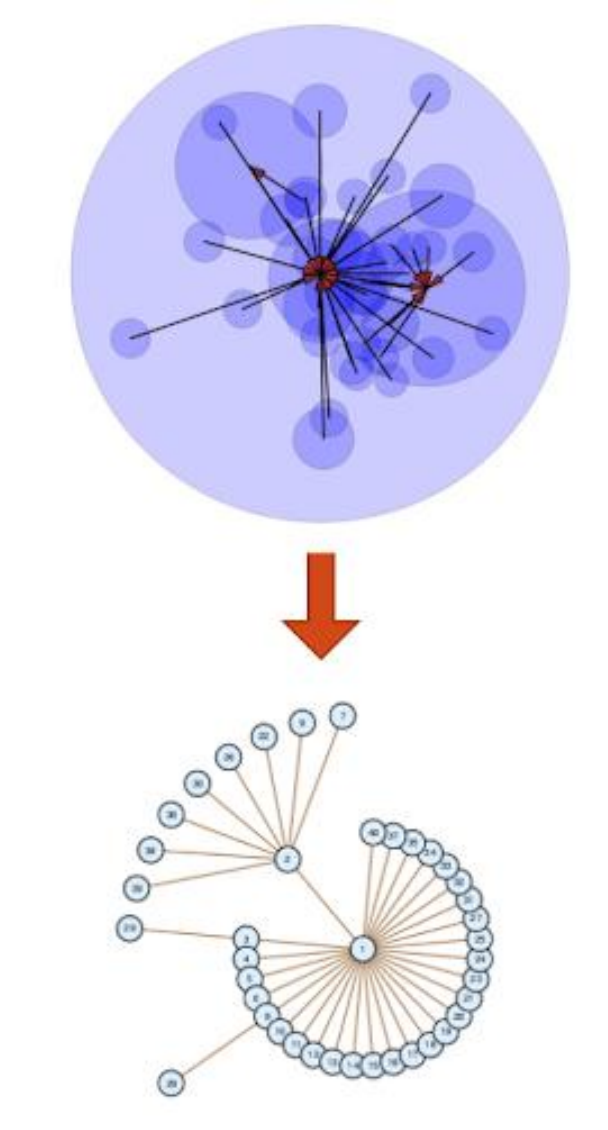
$$S_{\mathrm{tot}} = S_G[G_{0:K}] + S_x[\mathrm{x}_{0:K}; G_{0:K}] + S_{\mathrm{bnd}}[\mathrm{x}_{0:K}; G_{0:K}]$$

S_G: structure term (graph evolution)

S_x: internal term (formation physics)

S_{bnd}: boundary term (environment)

Probabilistic halo graph construction for clustering



Yang & Yu 2023 PR Research:

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

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

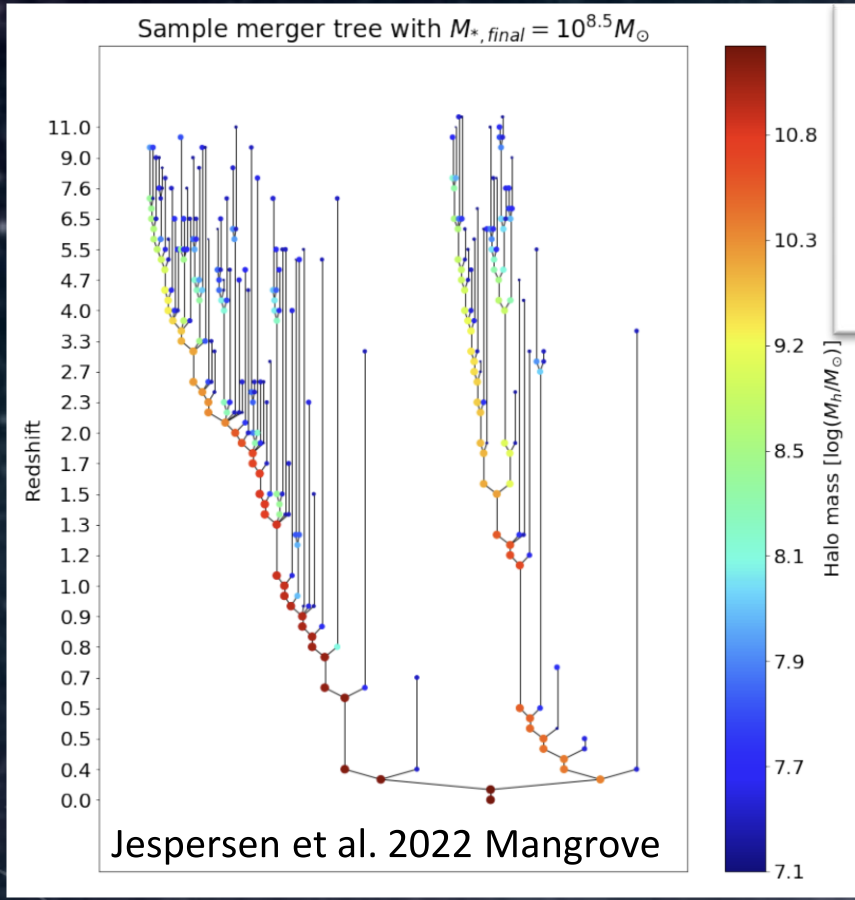
Preferential attachment

- Rich becomes richer
- Early attachment advantage

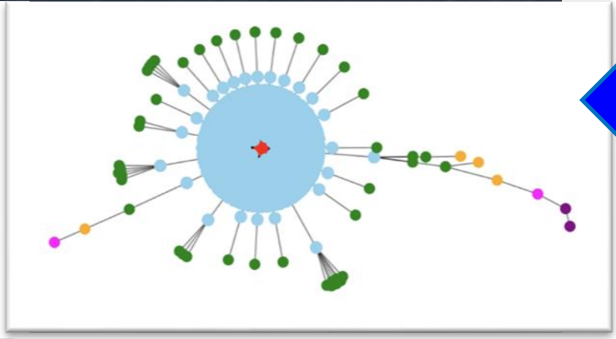
Self-similarity: Hierarchical structure naturally arise

A natural object to encode hierarchical structure formation and environmental features

One halo at z=0 has one merger tree

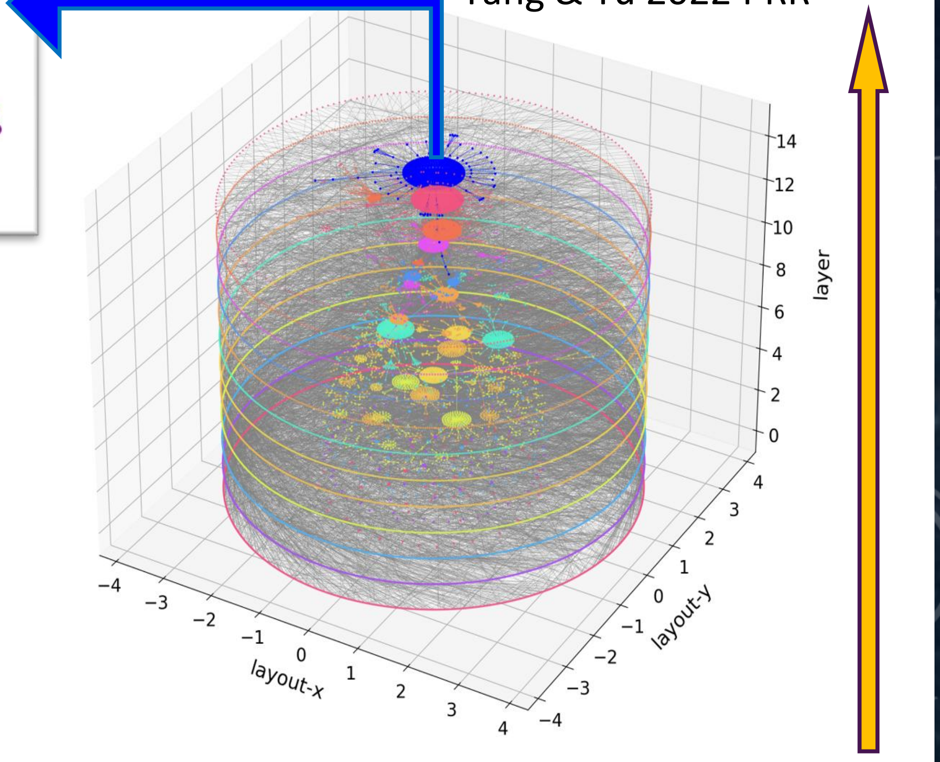


Jespersen et al. 2022 Mangrove



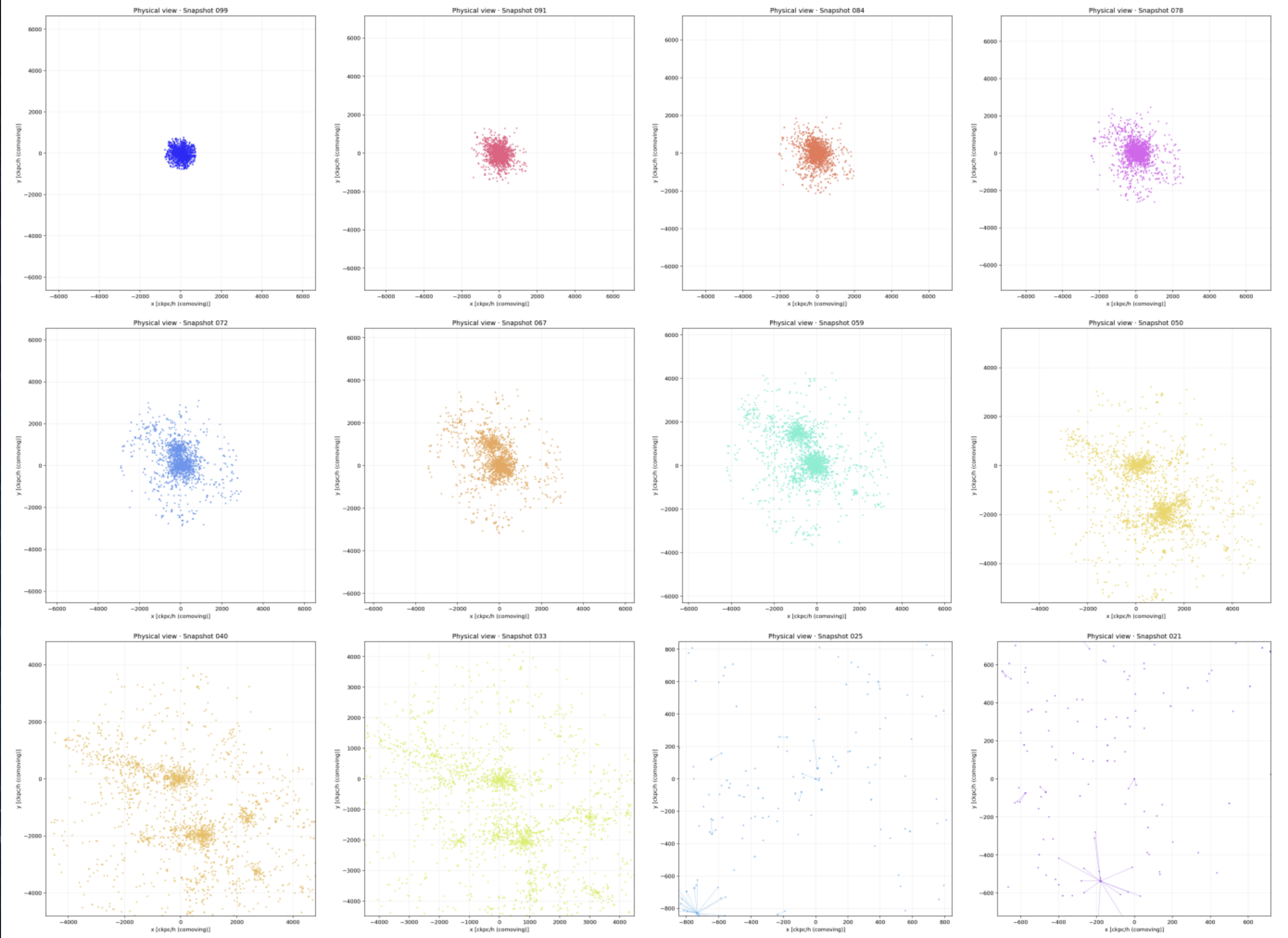
New: all halos belonging to an z=0 host halo has one Layered Halo Graph

Grouped histories for main and sub halos



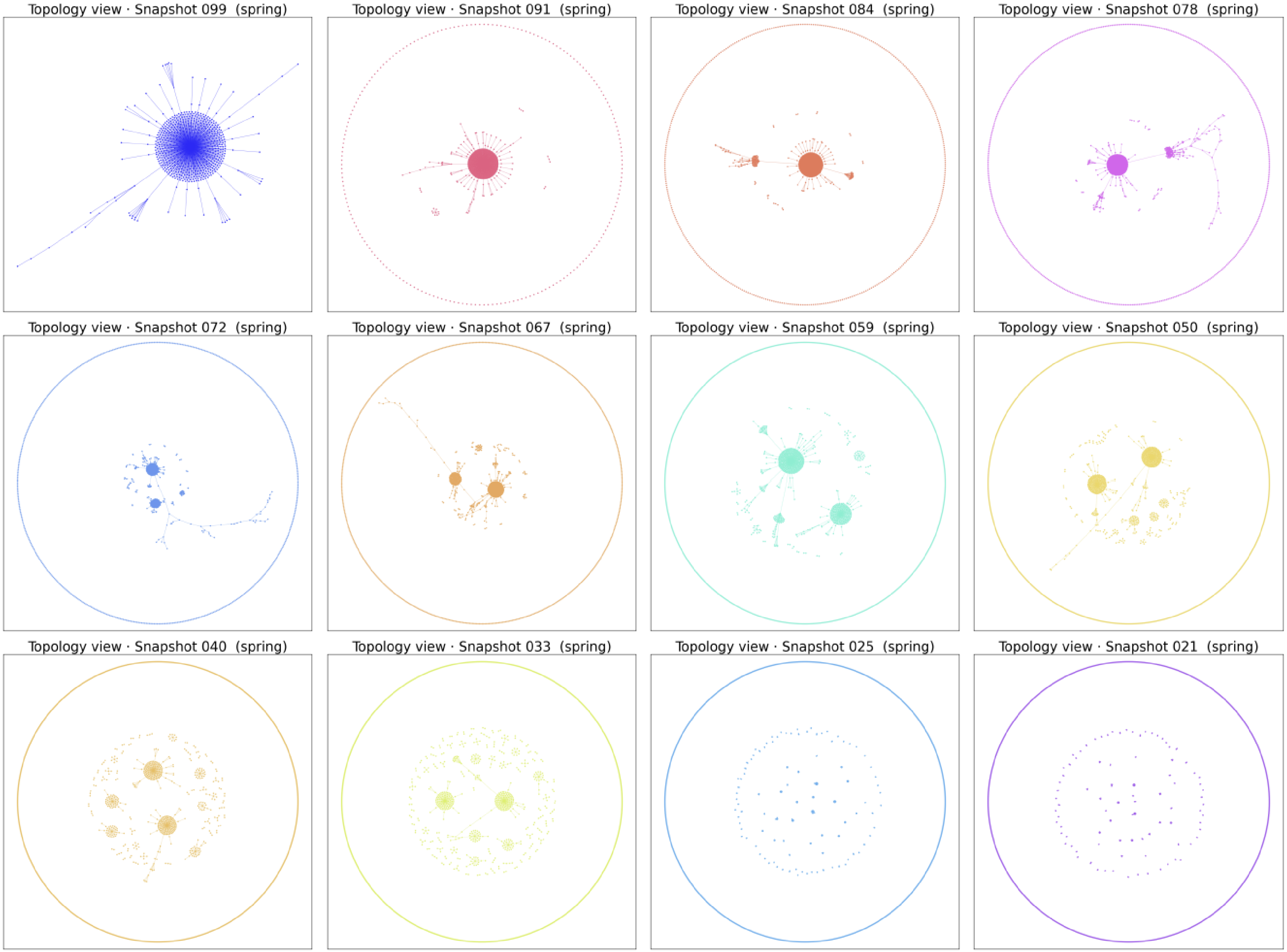
Yang 2603.15128, Yang & Yu 2022 PRR

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



Physical (comoving-coordinate) view

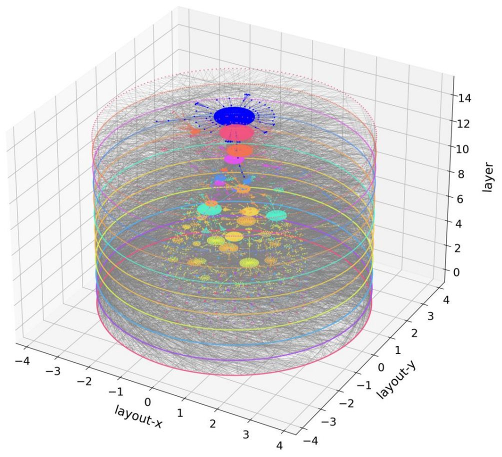
Hierarchical structures merge into a self-bound halo



Topological view

- Point has no size
- Edge has no length

Stack the layers



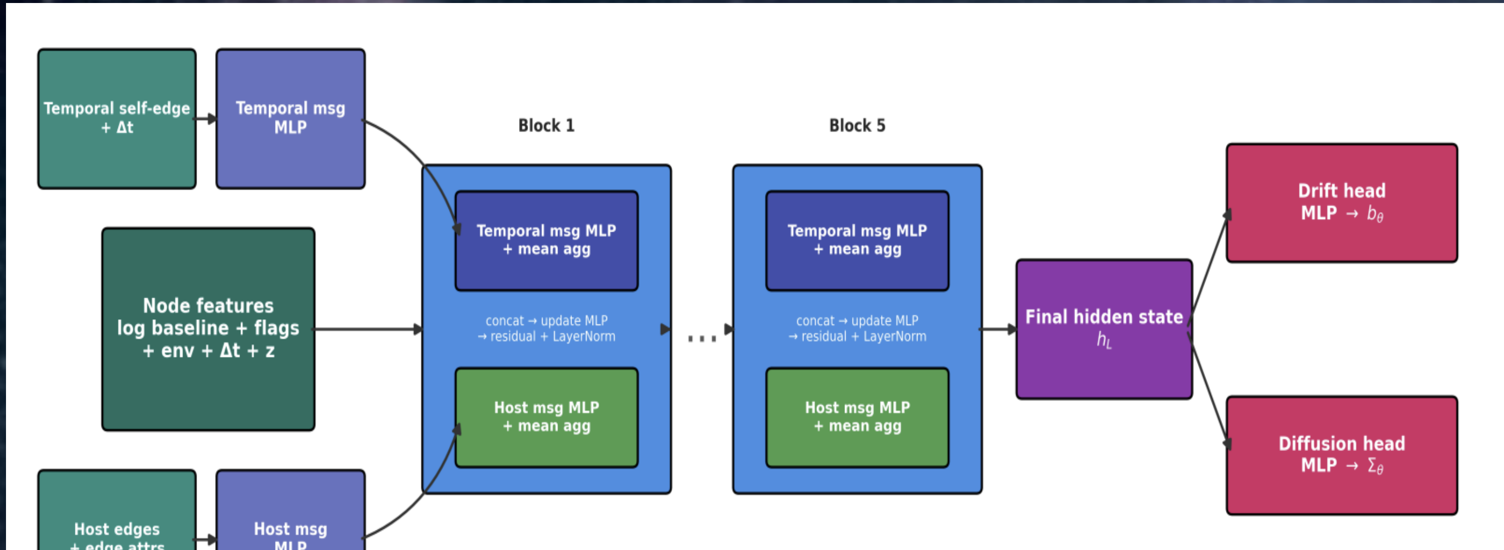
A Graph Path Likelihood Model (GPLM) from machine learning

Residual drift

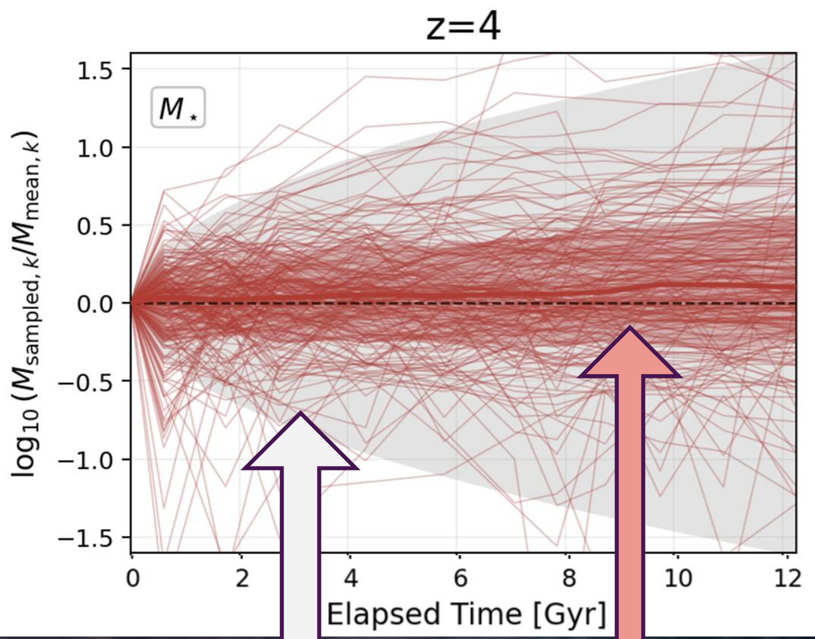
$$\Delta x \sim \mathcal{N}(b(x,G), \Sigma(x,G))$$

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

Training minimizes a discrete OM action build on data across histories： 6-layer (13 total) window



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



Unregulated scatter

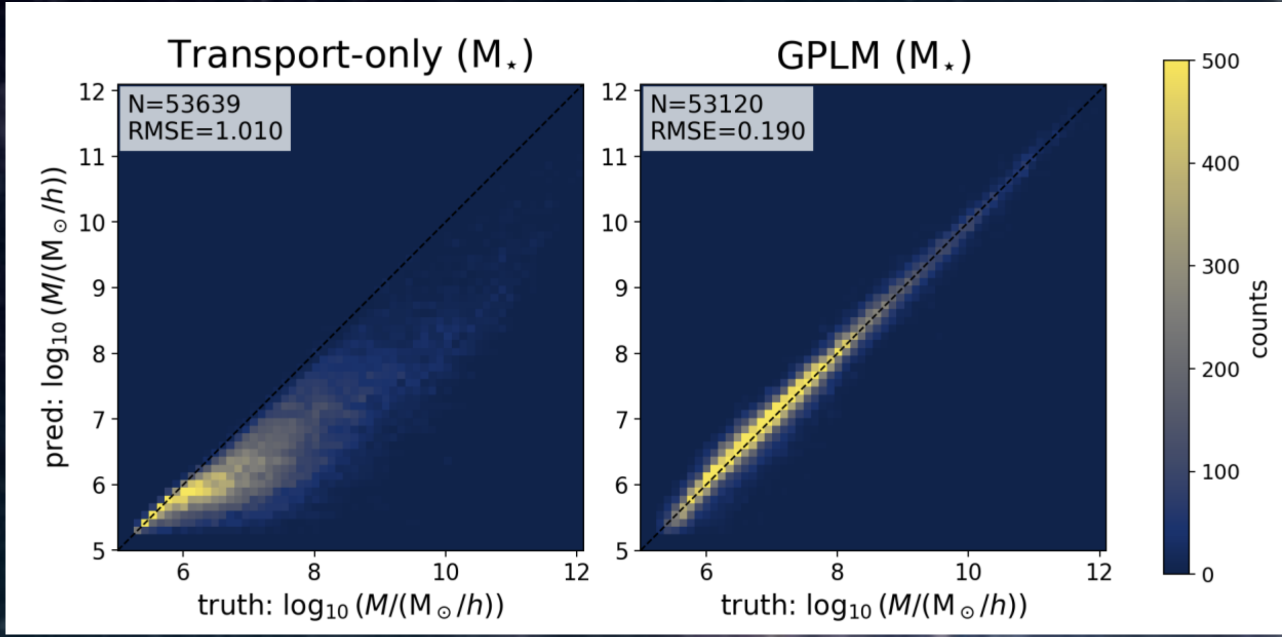
Regulated scatter

Graph encoding, conditioning

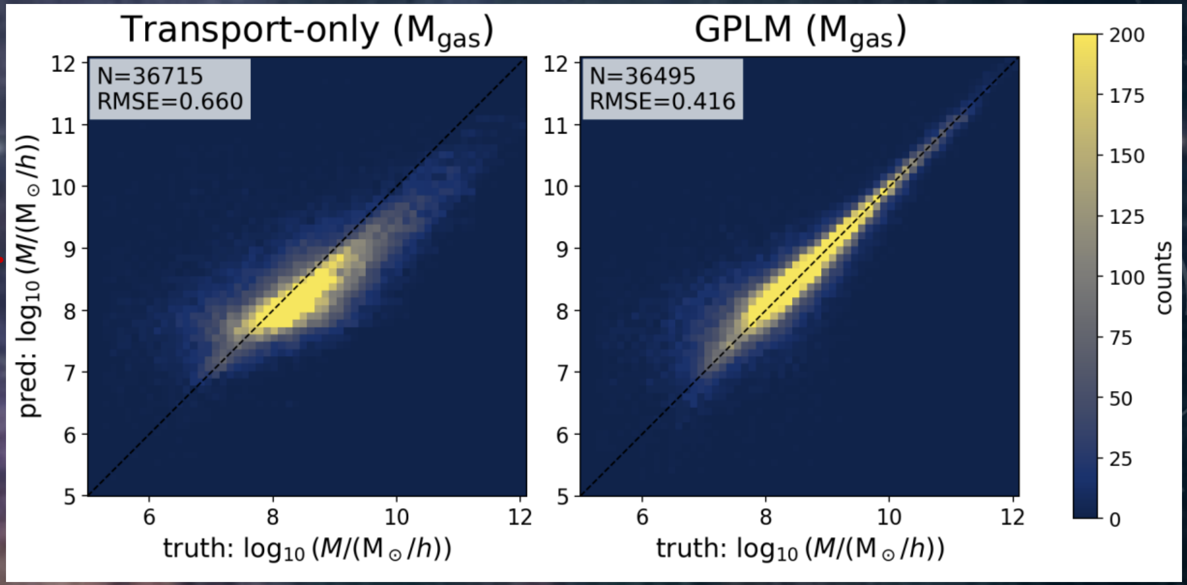
Graph Neural Network

- b: learned residual drift
- D: learned diffusion

Model performance



Prediction vs Truth in Stellar Mass



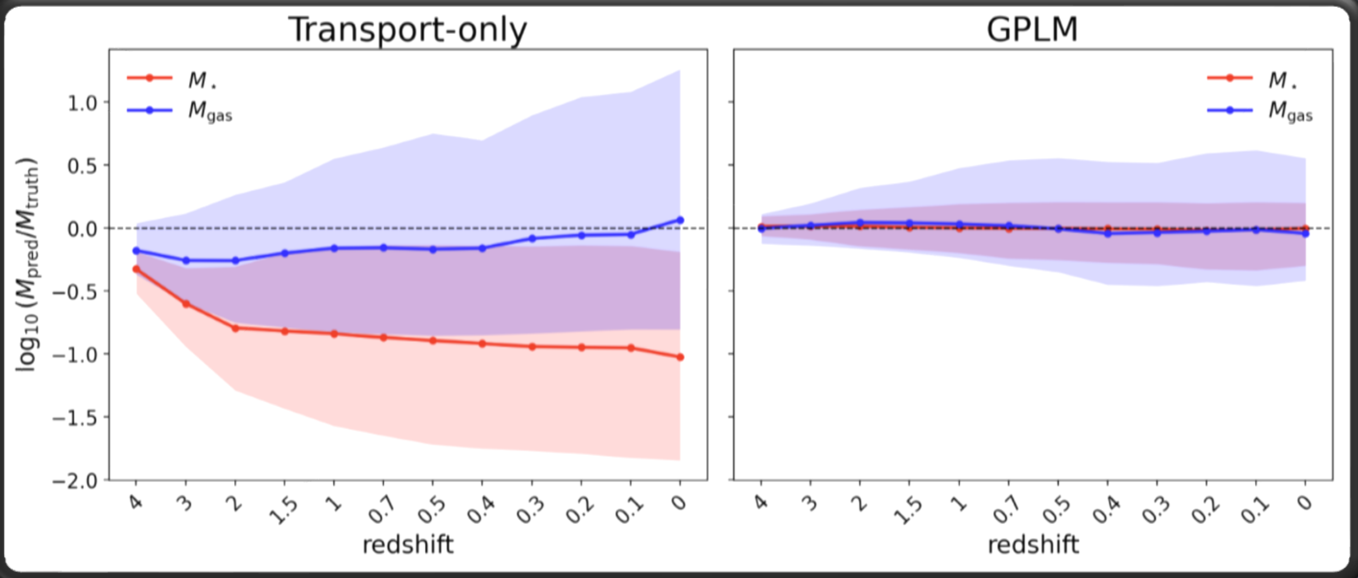
Prediction vs Truth in Gas Mass

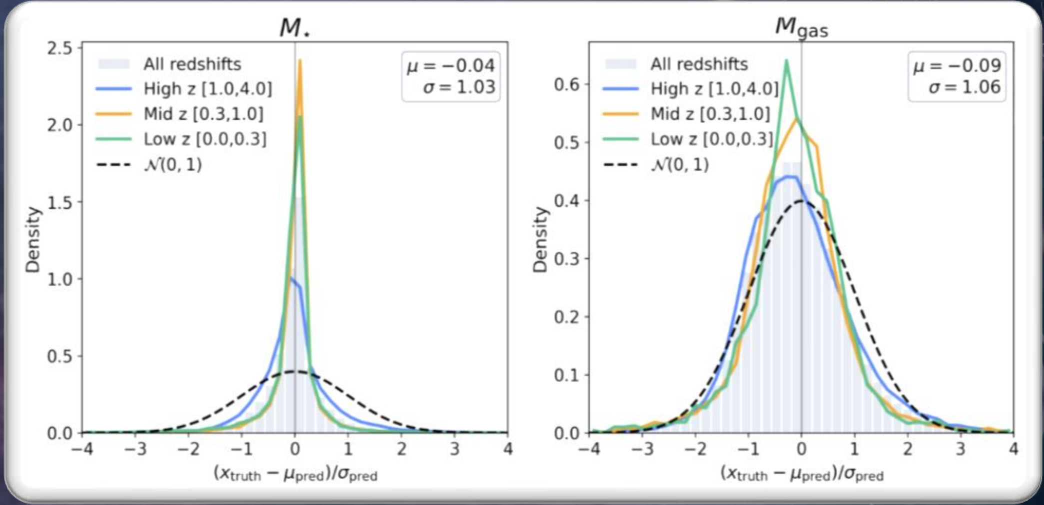
A Graph Path Likelihood Model (GPLM) from machine learning

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

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

- b: learned drift
- D: learned diffusion



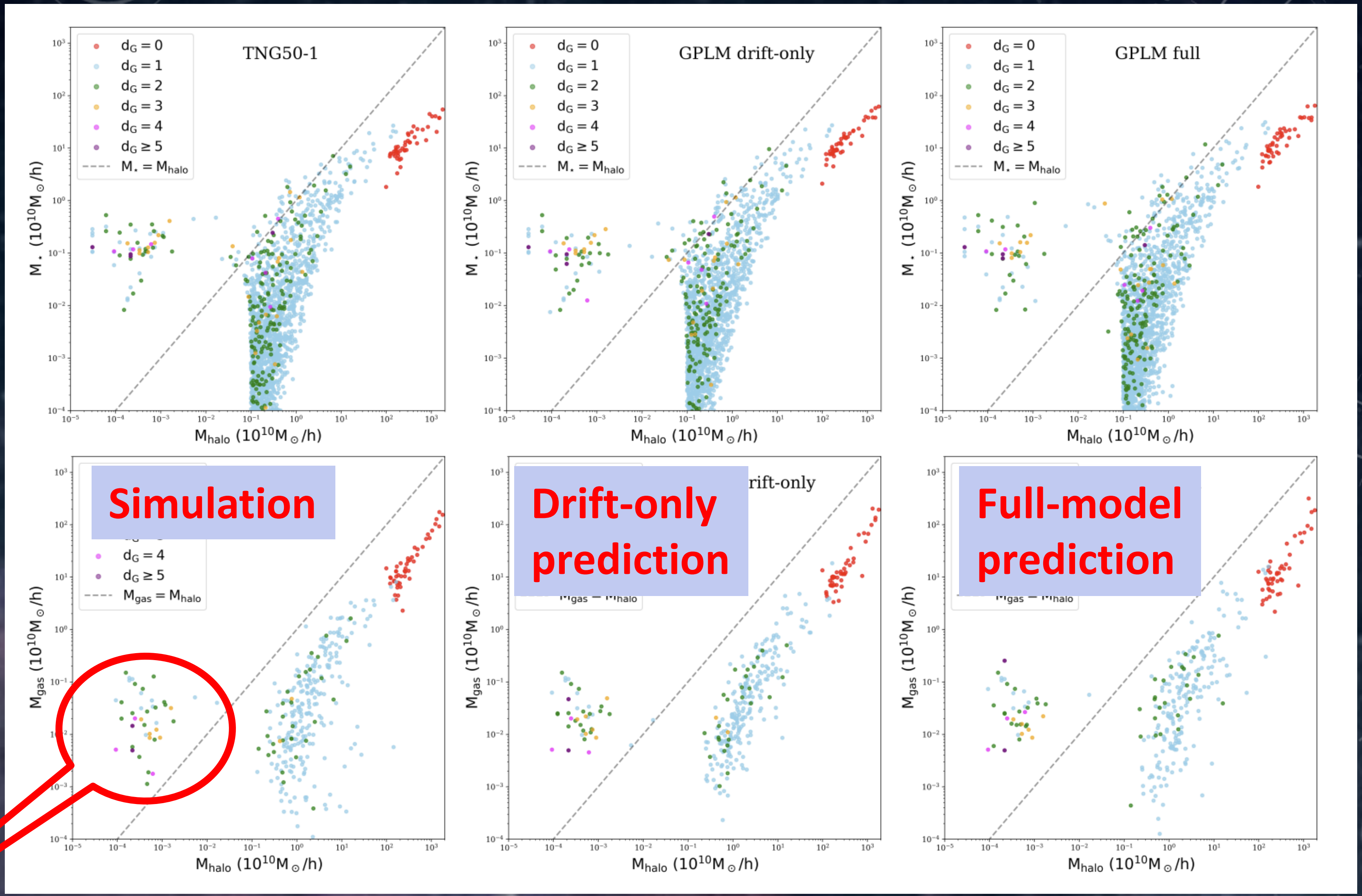


GPLM substantially reduces both the median bias and the spread of the residual distribution

The learned diffusions demonstrate **deviations from Gaussian**, which could be meaningful and be modeled by incorporating higher order terms

Model performance

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



Dark Matter Deficient Galaxies

- *Different sampled path batches represent* ***distinct stochastic realizations*** *of the same learned effective dynamics.*
- *The endpoint observables retain* ***analogous stochastic structure*** *across these different realizations.*

A Functional View under Effective Dynamical Scatter

Observables become functionals of paths where distortions can be introduced
O=F[x(t)]

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

The Langevin equation

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

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

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

Dark Matter Deficient Galaxy (DMDG) Probabilities

DMDG operator:

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

Operator average

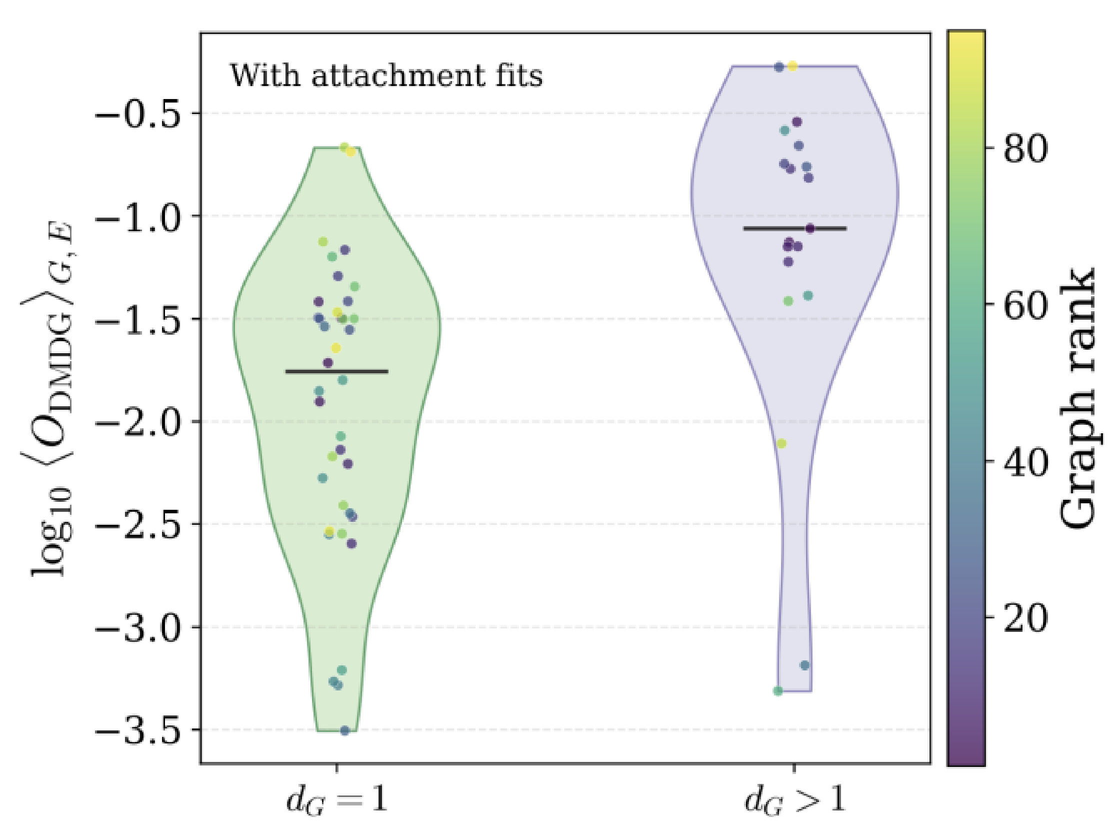
$$\langle O_{\mathrm{DMDG}}\rangle_{G,E} \equiv \int \mathcal{D}x \left[\frac{\sum_{i\in E} O^{(i)}_{\mathrm{DMDG}}[\mathbf{x}]}{N_E}\right] P_{\mathrm{tot}}(\mathbf{x} \mid G, x_0)$$

dG=1: satellites

dG>1: higher order satellites

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

➢ **Most probable path for DMDG formation?**



Gas-Rich response as a controlled likelihood deformation

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

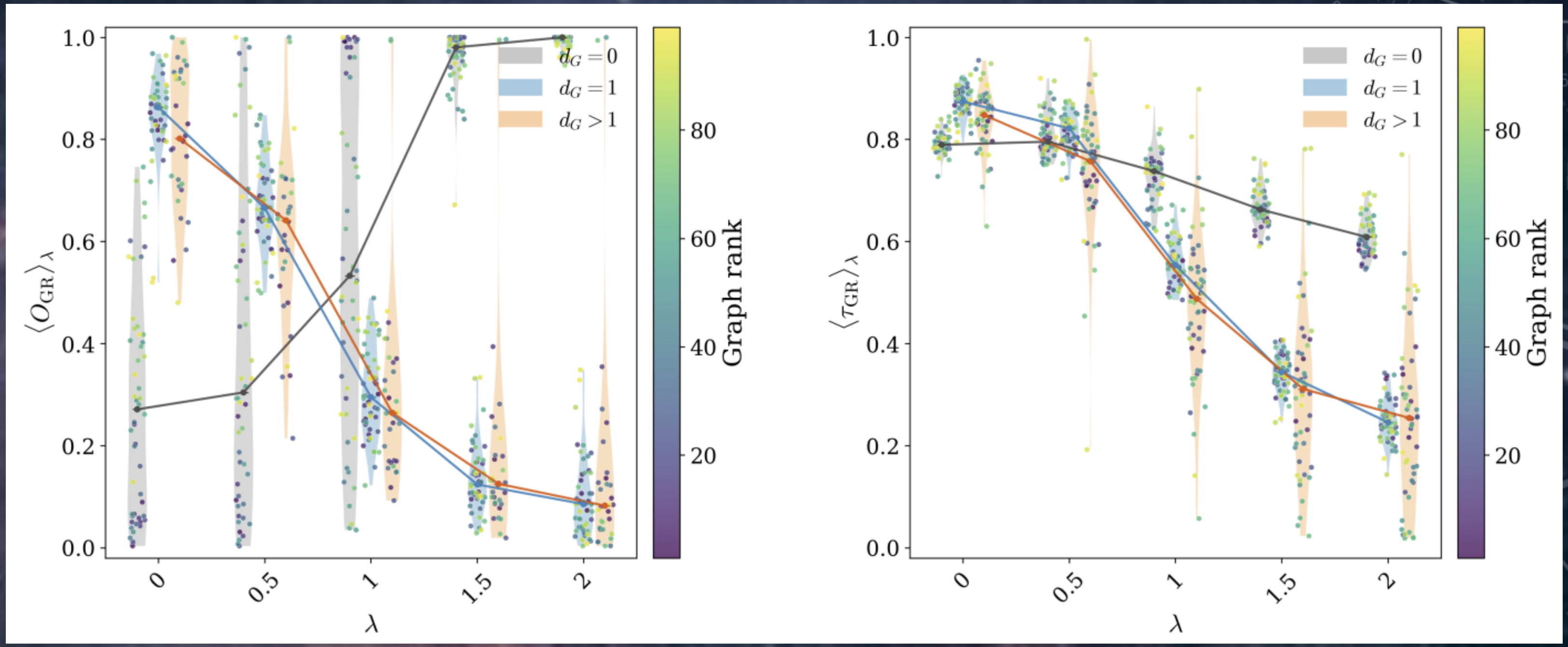
Controlled deformation: $b_{{\rm gas},\lambda}(x_k, G) = b_{\rm gas,off}(x_k, G) + \lambda[b_{\rm gas,full}(x_k, G) - b_{\rm gas,off}(x_k, G)]$

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

dG=0: host

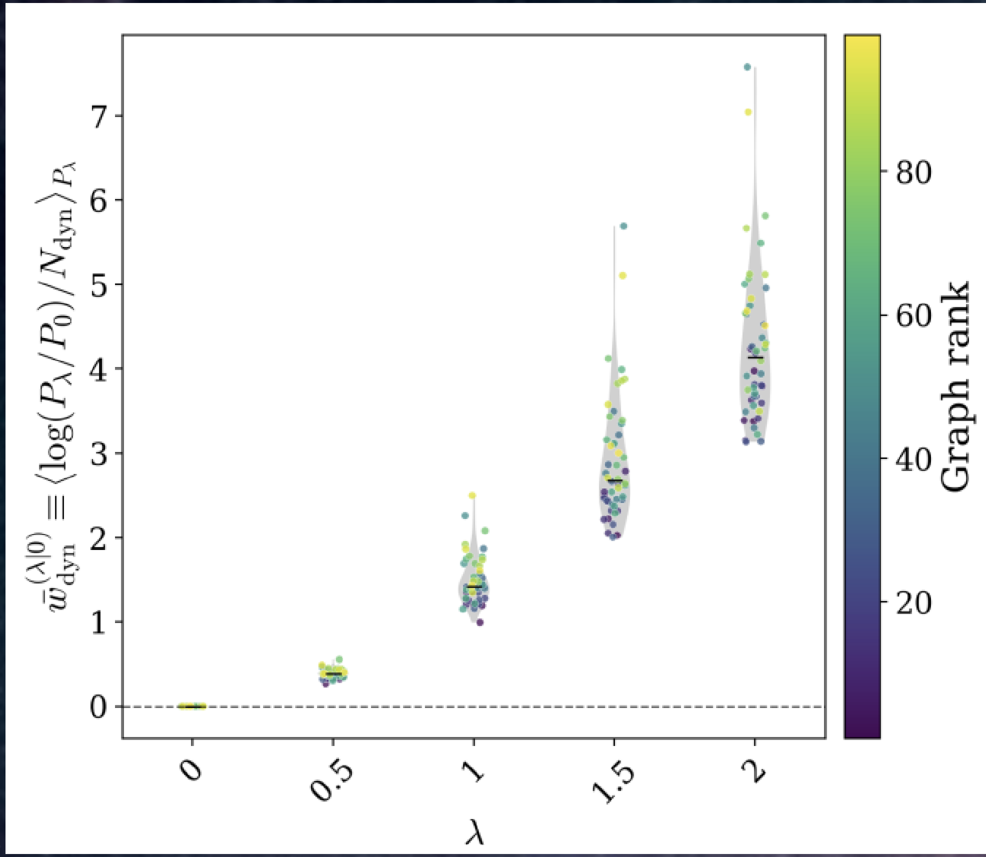
dG=1: satellites

dG>1: higher order satellites

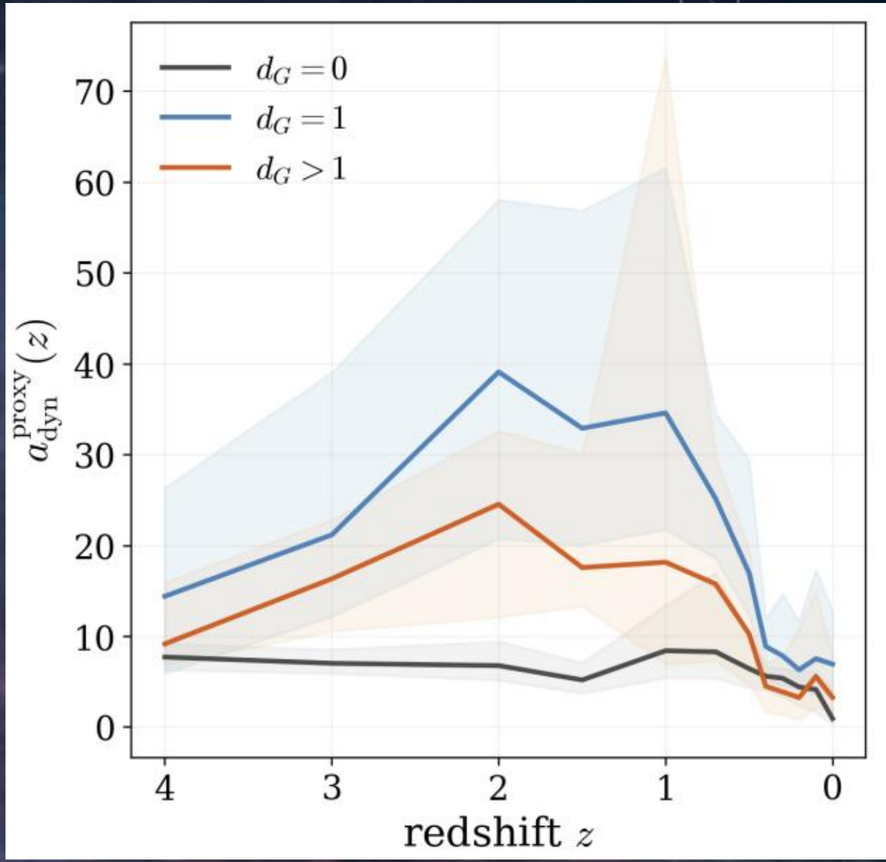


Path-Space likelihood diagnostics

A path-level KL-like distance



A local forward-reverse asymmetry



Larger λ drives a larger departure from λ=0 and broadens the graph-to-graph scatter

Peaks around cosmic noon, indicating stronger history-dependent evolution there

Theoretical Perspectives

Non-Gaussian scatter

Geometric interpretation

Optimal Transport

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

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

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

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

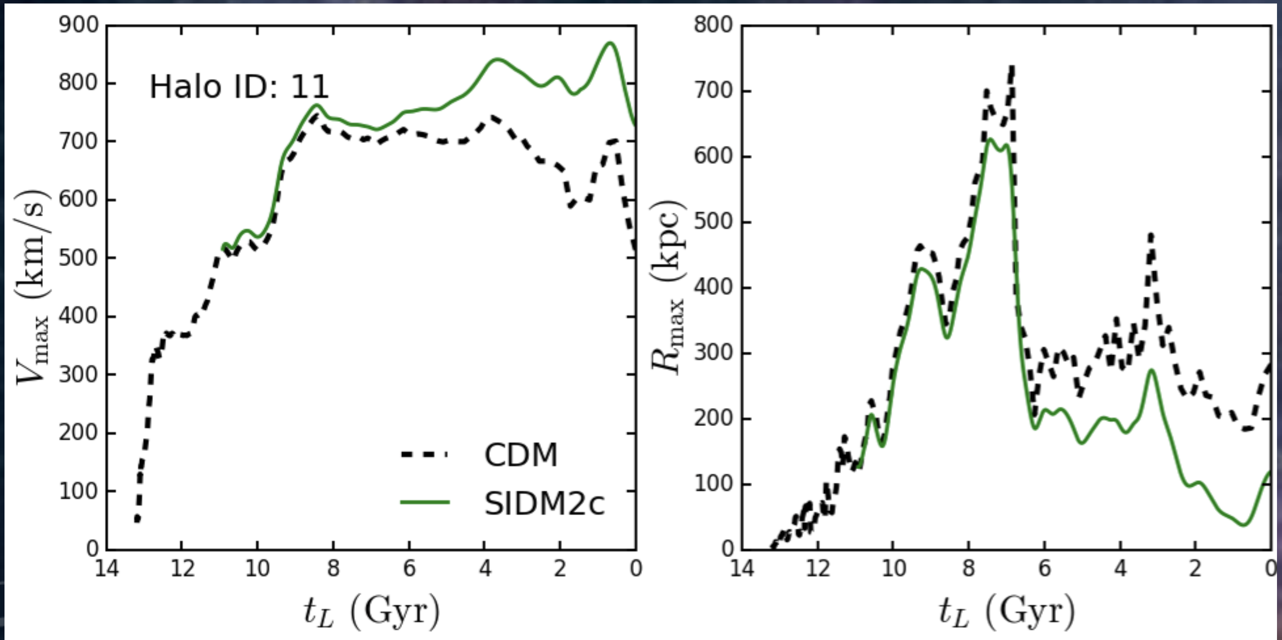
Most probable path=geodesic in g=D^{-1} under an effective potential

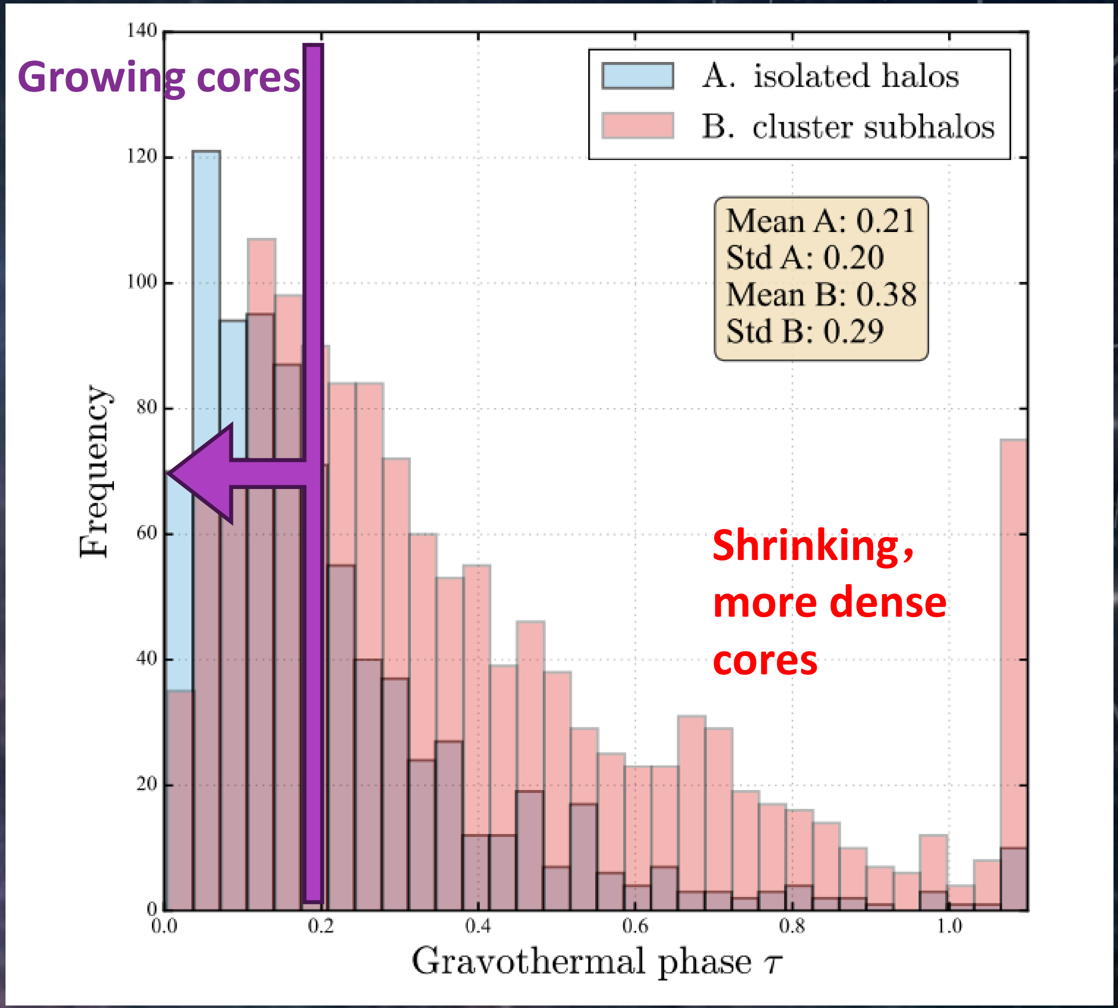
Galaxy evolution path=geodesic in environment-shaped metric+drift forcing

Path-conditioned predictions for 2-comp SIDM | Deterministic

Yang, Fan, Hou, Tsai, Sci. Bull. 71 (2026)

- Isolated dwarfs are mostly core forming (tau<0.2)
- Dwarfs in clusters are mostly core collapsing (tau>0.2)





SIDM as controlled deformation on graph paths

Analytic kernel → Neural Network Agent

- SIDM deforms (Vmax,Rmax) per time step
- Environmental Paths change endpoint observables

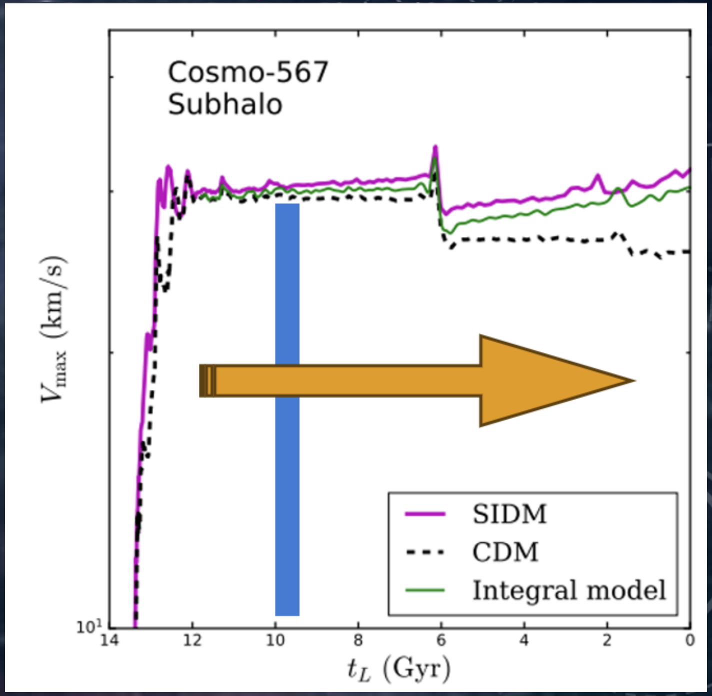
$$\mathbf{y}_k \equiv (V_{\max}, R_{\max})_k$$

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

(Non-linear) Transport

Residual Drift

Residual Scatter

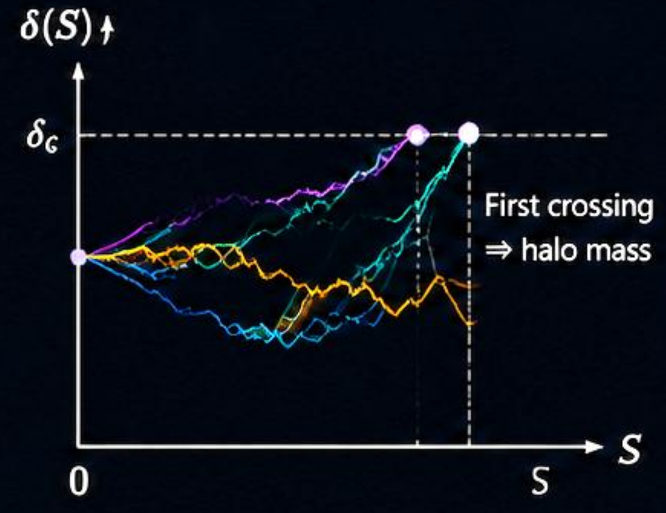


Yang et al. 2305.16176

From Stochastic Halo Growth to Stochastic Layered Graphs

Excursion-set theory: halo growth is a stochastic trajectory of density fluctuations across smoothing scales.

1. Excursion-set: random walks



Stochasticity comes from random initial fluctuations ⇒ random walks.

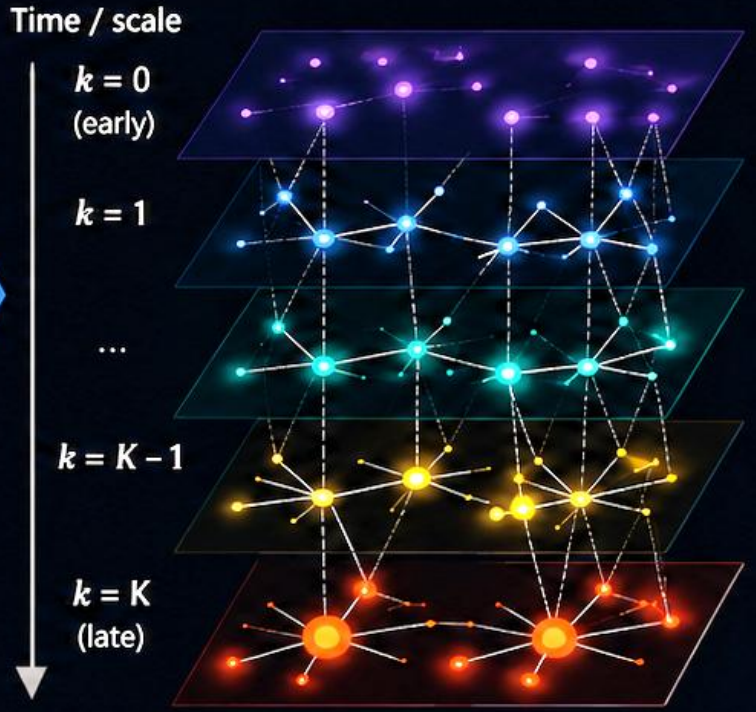
2. Stochastic growth (Langevin equation)

$$\frac{d\delta}{dS} = \mu(S) + \sqrt{D(S)}\,\eta(S)$$

- $\mu(S)$: drift (e.g., gravity)
- $D(S)$: diffusion (power spectrum)
- $\eta(S)$: Gaussian white noise $\langle \eta(S)\eta(S') \rangle = \delta_D(S - S')$

Drift + random forcing ⇒ stochastic trajectories

3. From trajectories to layered halo graphs



- Halos (nodes)
- Mergers (within layer)
- Progenitor links (across layers)

Joint probability

$$P(G_{0:K}) = P(G_0) \prod_{k=0}^{K-1} P(G_{k+1} \mid G_k)$$

4. Joint model: halo graph + galaxy fields

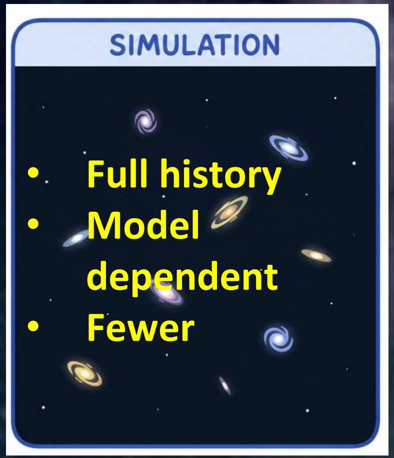
Halo graph action		Galaxy/field action		Total action
$S_G[G_{0:K}]$	$+$	$S_x[\mathbf{x}_{0:K}; G_{0:K}]$	$\Rightarrow$	$\boxed{S_{\text{tot}} = S_G + S_x}$

Thanks for your attention!

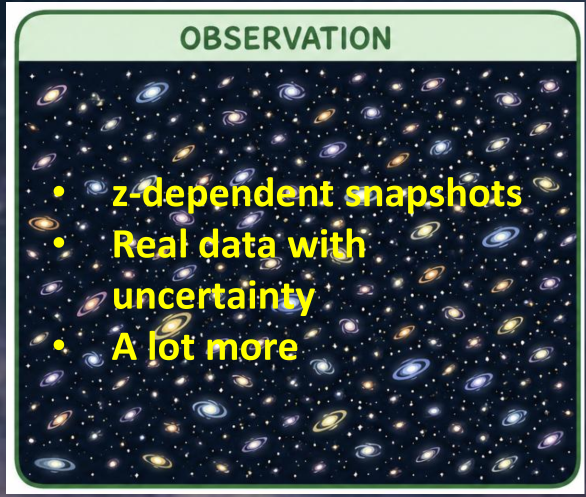
Back up

Which graph is most consistent with observations?

$$P(\mathcal{G} \mid d_{\text{end}}) \propto P(\mathcal{G}) \int \mathcal{D}\mathbf{x}\, \boxed{P(d_{\text{end}} \mid \mathbf{x}, \mathcal{G})}\, \boxed{P(\mathbf{x} \mid \mathcal{G})}$$

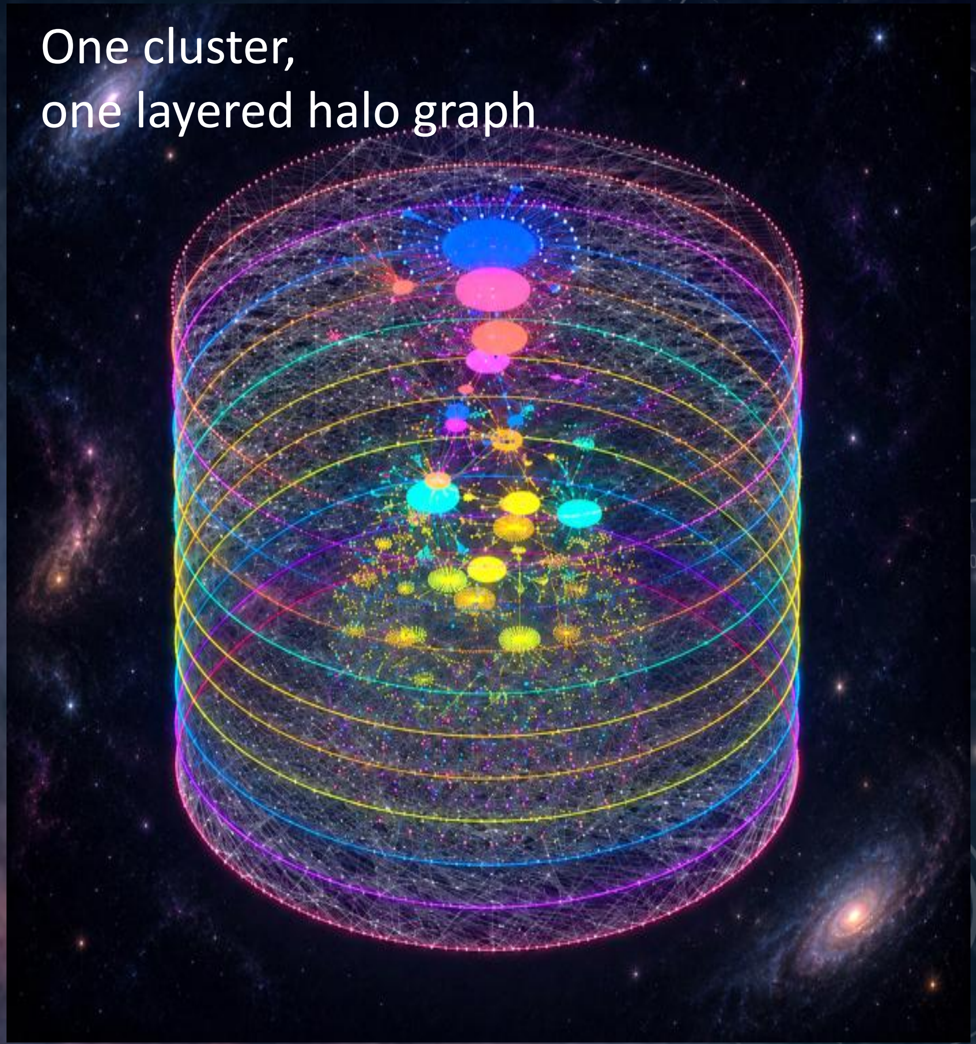


Can We Learn from Observational Data?

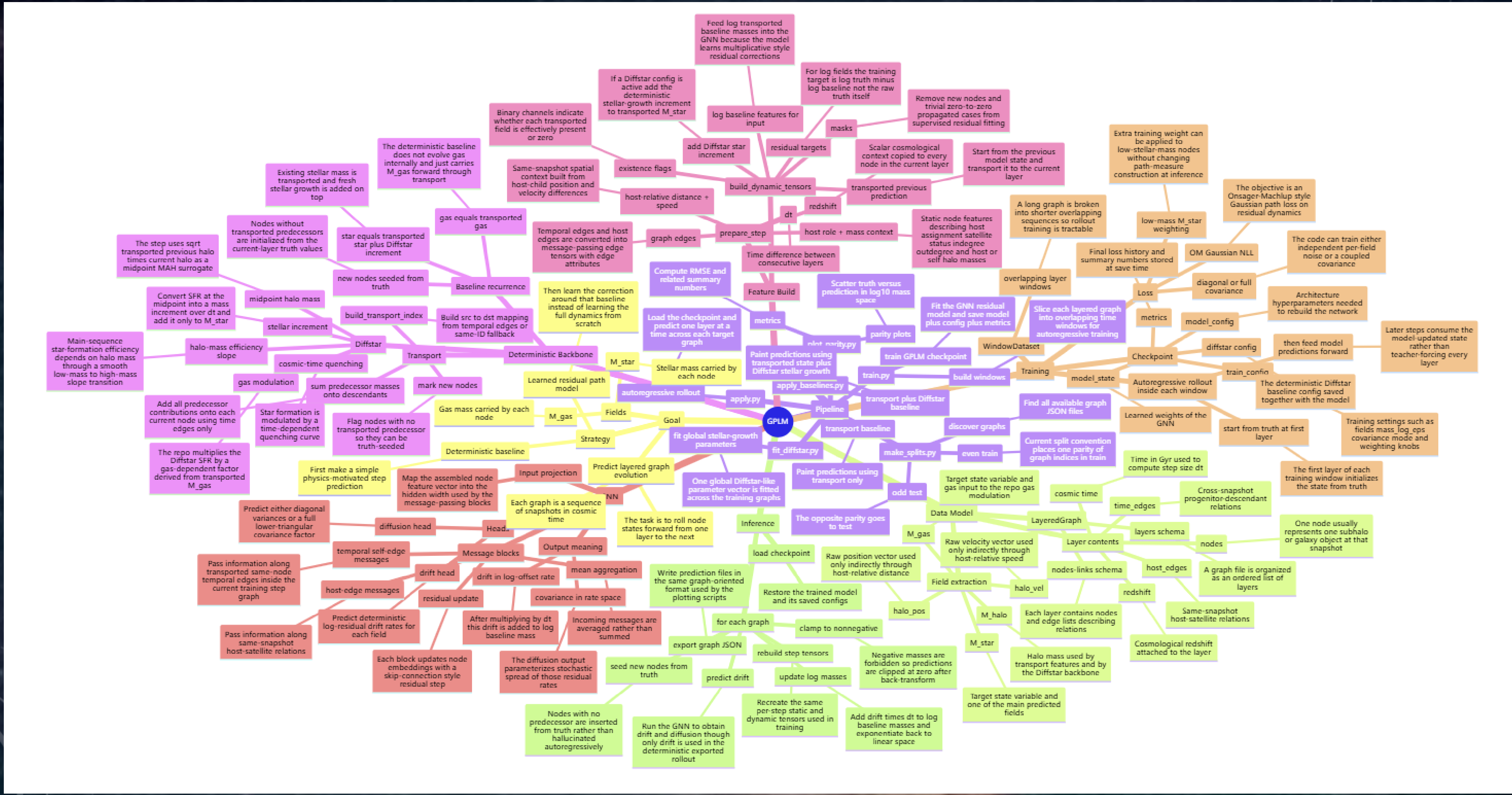


The vast amount of data z-dependent could, in principle, be used to fine-tune GPLM

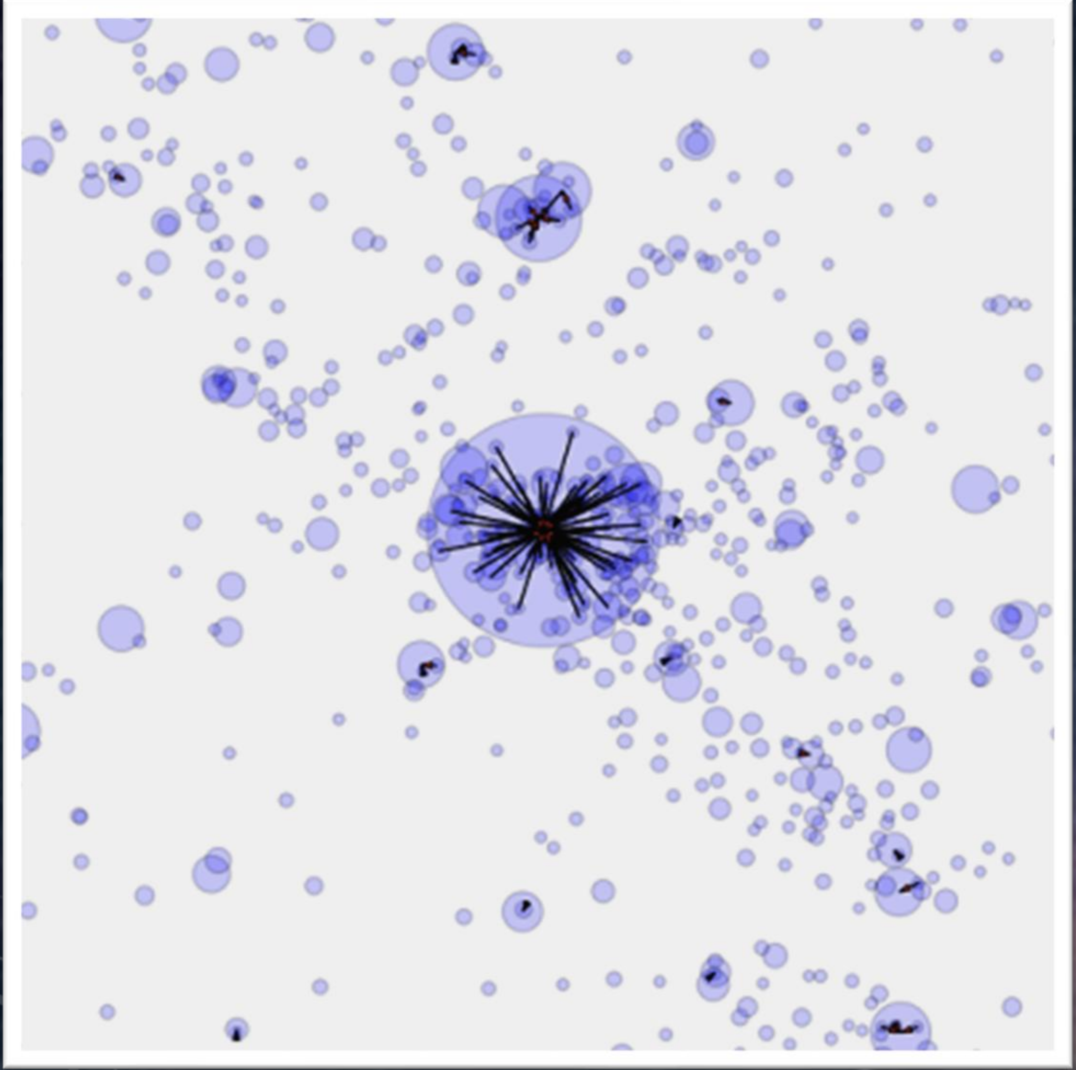
$$P(d_{\text{all}} \mid \vartheta) \sim \prod_{z} \int \mathcal{D}G\, \mathcal{D}\mathbf{x}\, P(d_z \mid \mathbf{x}_z, G)\, \boxed{P_{\vartheta}(\mathbf{x}, G)}$$



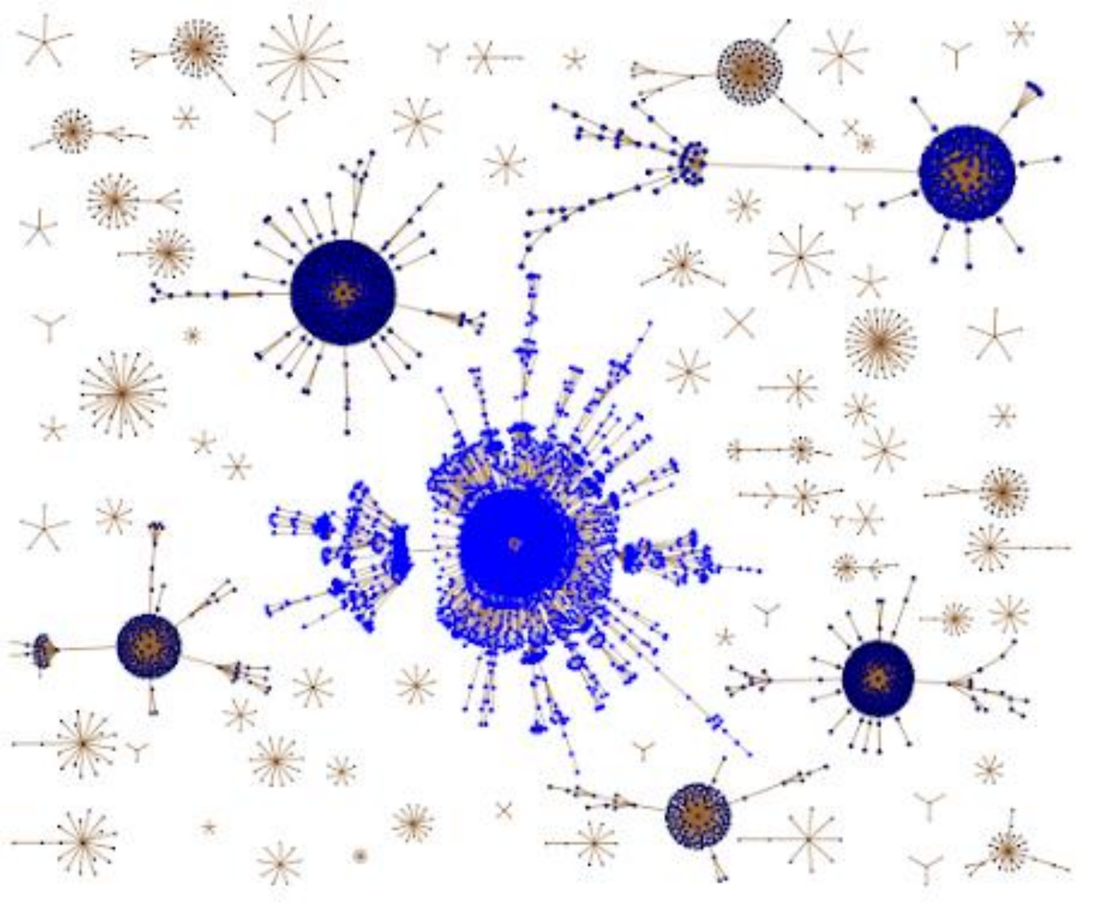
The GPLM code layout



Cosmological halo formation can be encoded by graphs



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



MSRJD in GPLM: Discrete Realization of Stochastic Evolution

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

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

$$\begin{cases} \Delta x_{i,k} = b_{\theta,i,k}(s_{i,k}, G_k)\, \Delta t_k + \xi_{i,k} \\ \langle \xi_{i,k}\, \xi_{j,k'}^T \rangle = D_{\theta,i,k}\, \Delta t_k\, \delta_{ij}\, \delta_{kk'} \end{cases}$$

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

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

non-Gaussianity could be handled by higher-order terms

Path Space Tools

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

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)

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

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

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

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

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

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

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