Cosmological Small-Scale Puzzles – tentative solutions within ΛCDM and in alternative models

Fangzhou Jiang

https://www.fzjiang.com

姜方周

KIAA Peking University

高能所 2025-06-25

Zixiang Jia (PKU student), Jinning Liang(PKU student), Guangze Sun (PKU student), Haonan Zheng (KIAA fellow)

Andrew Benson (Carnegie), Oren Slone (CCA/NYU), Manoj Kaplinghat (Irvine), Mariangela Lisanti (Princeton), Phil Hopkins (Caltech), Annika Peter (OSU), Xiaolong Du (UCLA), Shengqi Yang (Carnegie), Xuejian Shen (MIT), Zhichao Zeng (OSU), Dylan Folsom (Princeton), Igor Palubski (Irvine), Frank van den Bosch (Yale), Hai-Bo Yu (Riverside), Wei-Xiang Feng (Tsinghua), Daneng Yang (PMO), Avishai Dekel (Hebrew University), Jing Wang (KIAA), Zhaozhou Li (Hebrew University), Jonathan Freundlich (Strasbourg)

非线性结构形成 · 宇宙学数值模拟

$a \approx 0.05$

千禧年数值模拟 Millennium Simulation

非线性结构形成 · 宇宙学数值模拟

$a \approx 0.15$

非线性结构形成 · 宇宙学数值模拟

$a \approx 0.42$

非线性结构形成 · 宇宙学数值模拟

$a = 1$

- 暗物质晕：密度场涨落的峰，在引力作用下塌缩形成的结构

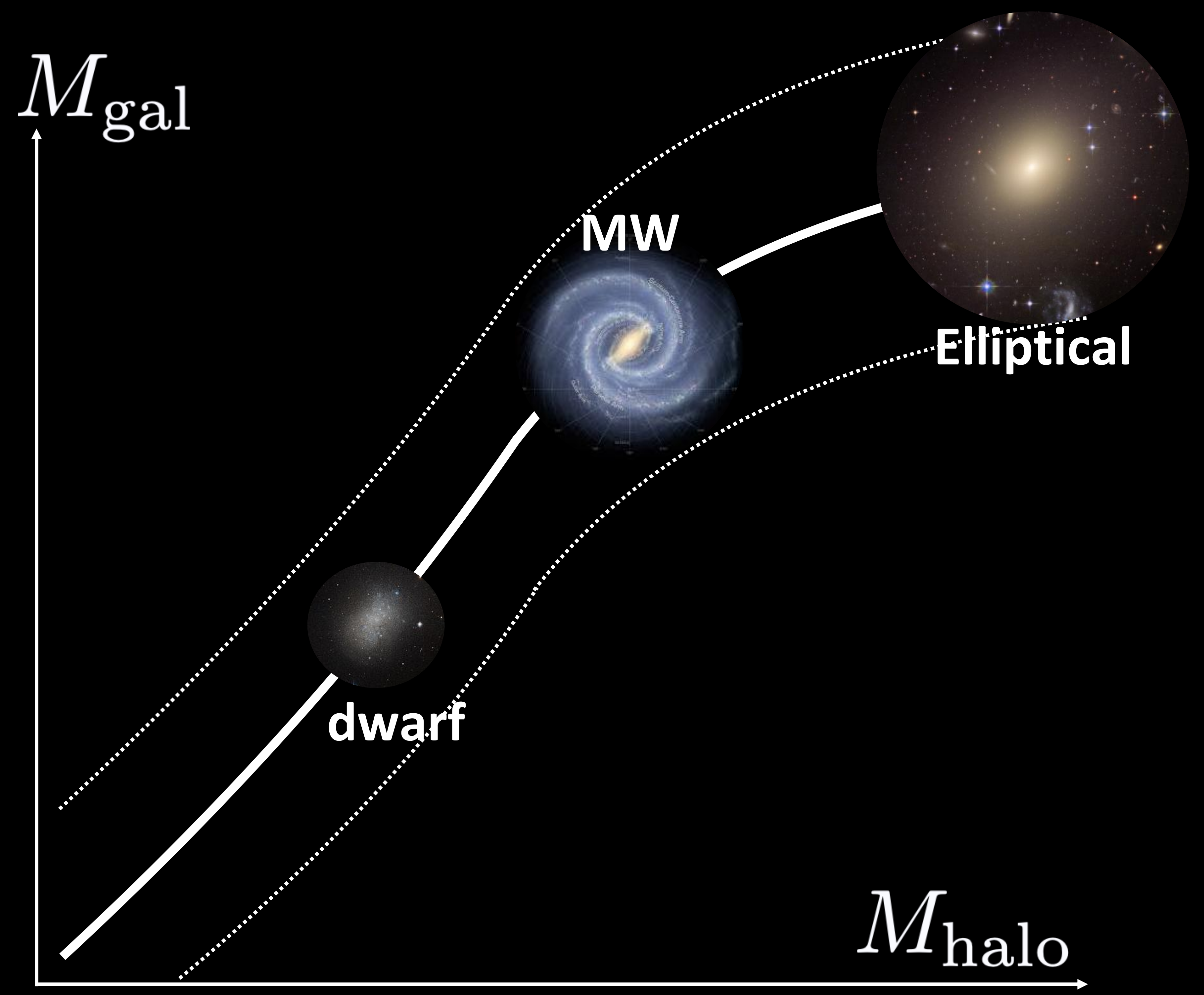
dark-matter halo
subhalo
simulation visualized by Ralf Kaehler, Carter Emmart, Tom Abel


halo
A few hundred kpc
galaxy
a few kpc

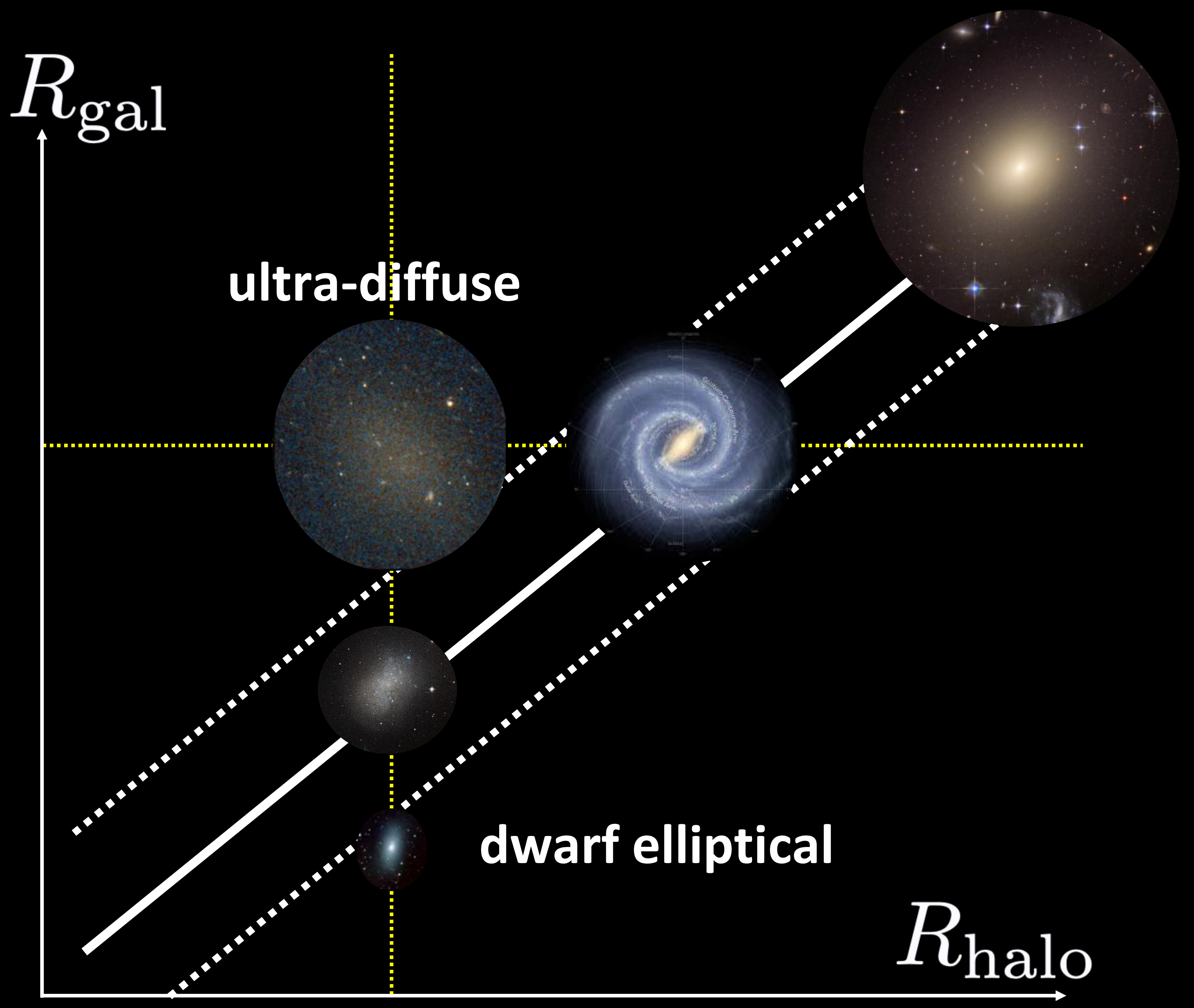
gas DM stars

IllustrisTNG100 宇宙学流体数值模拟

Galaxy-halo connection : What kind of galaxies live in what kind of DM halos ?



"**halo abundance matching**" e.g, Behroozi+13; Moster+13



e.g, Mo, Mao, White 98; Kravtsov 13

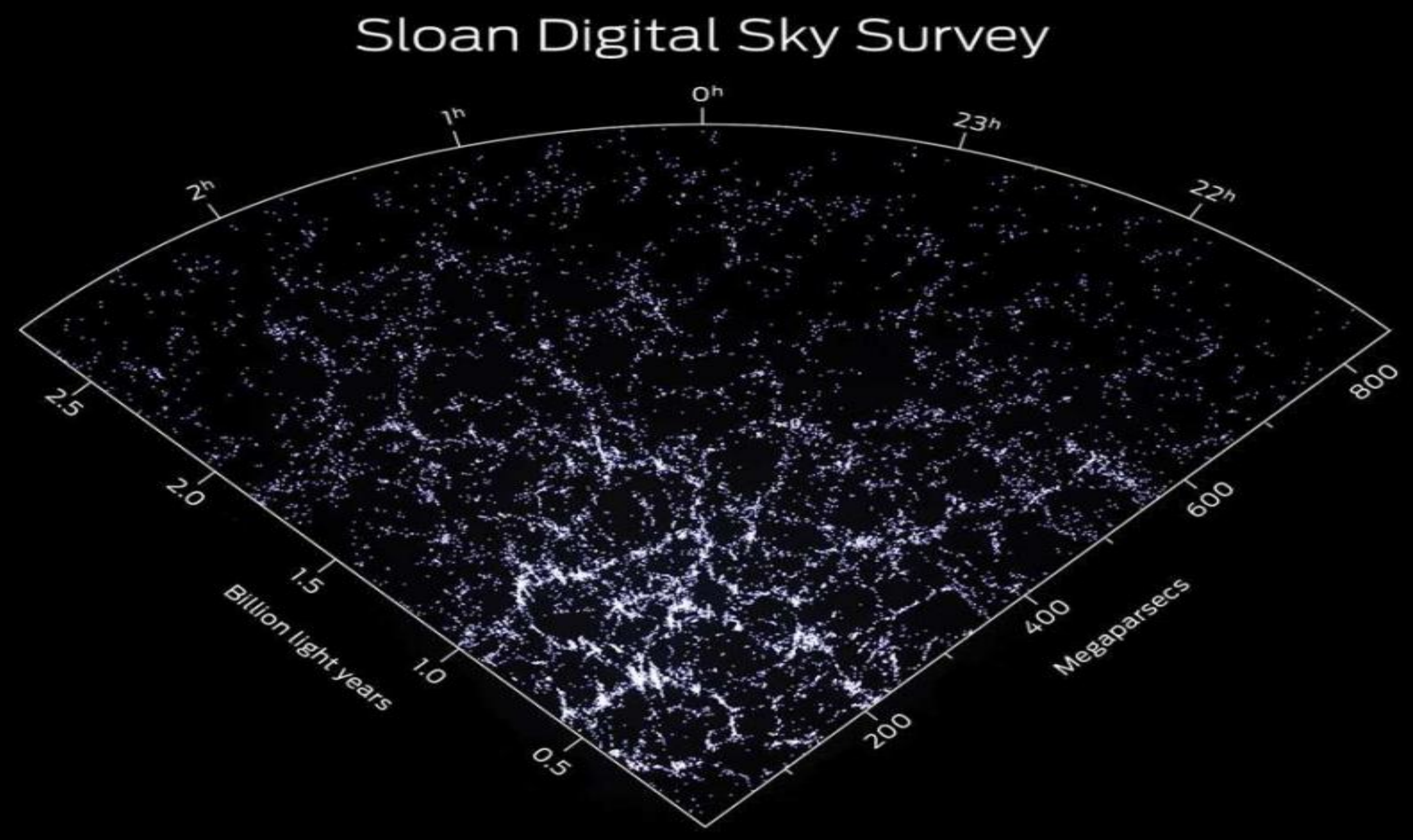
$$\mathcal{P}(Q_{\mathrm{gal}} | \underline{M_{\mathrm{halo}}}, p_1, p_2, \dots)$$

secondary halo parameters: structure, assembly history, large scale environment, …

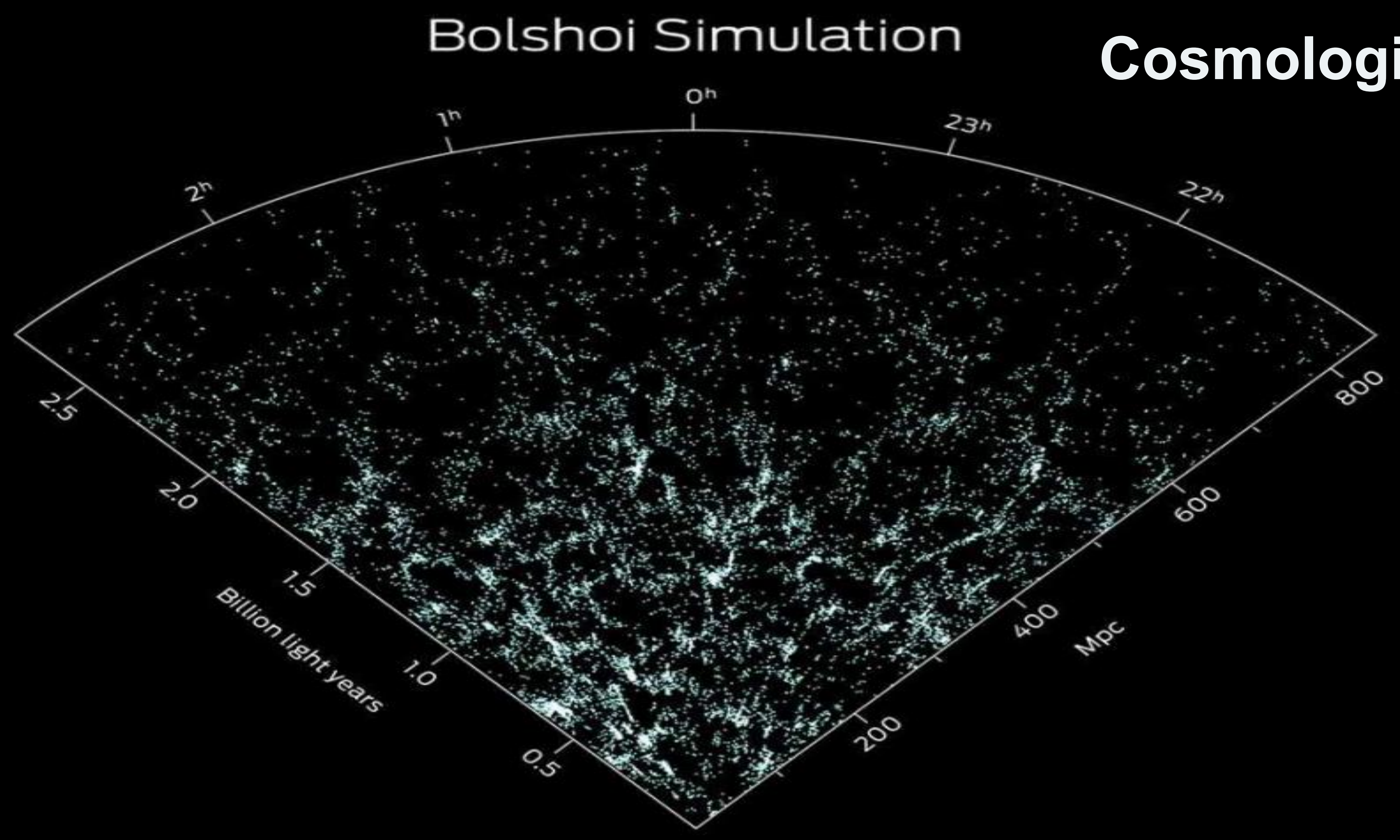
$$\Updownarrow \quad M_{\mathrm{halo}} = \frac{4}{3}\pi R_{\mathrm{halo}}^3 \Delta\rho_{\mathrm{crit}}$$

R_{halo}

ΛCDM is successful in reproducing large-scale structures



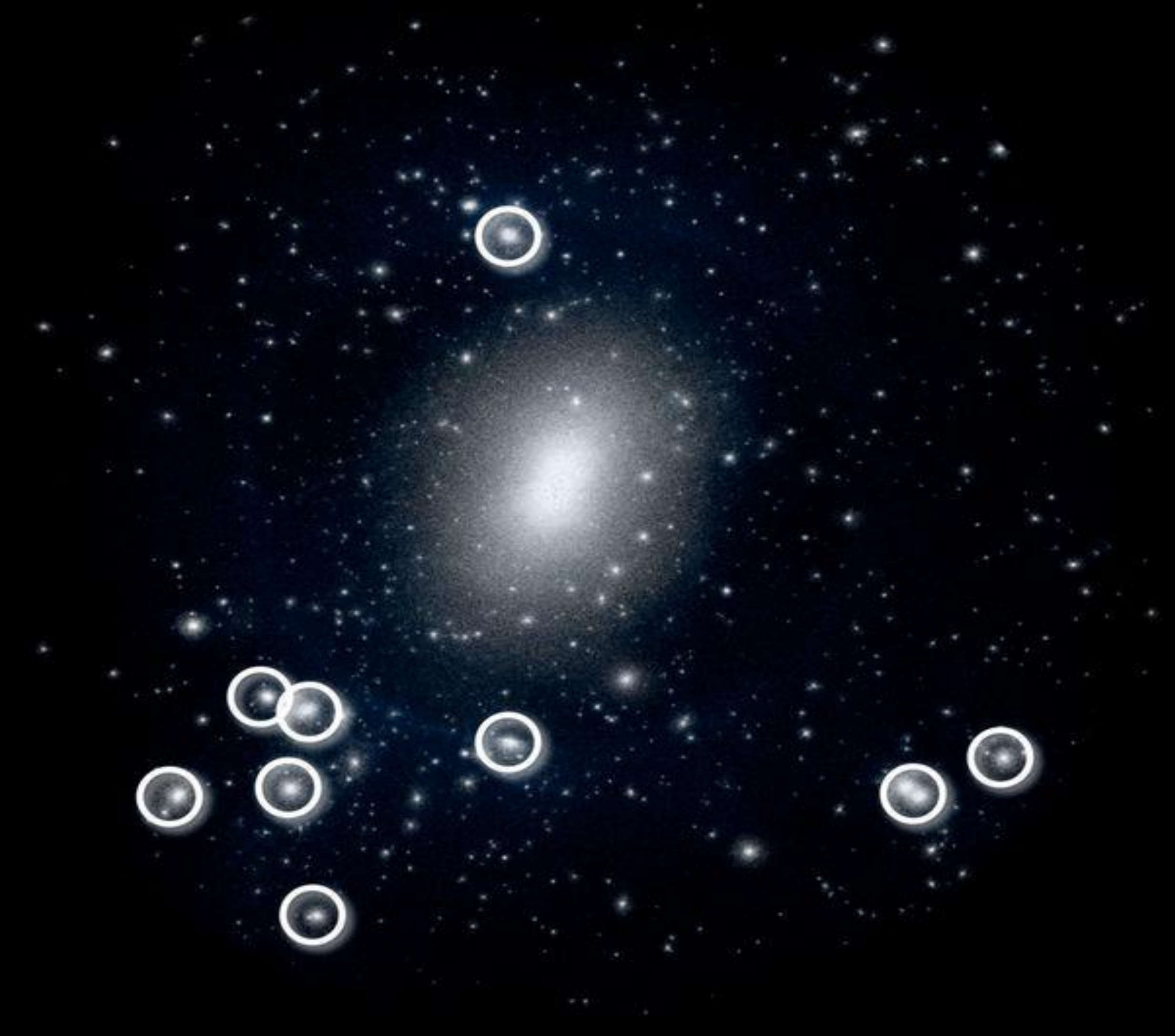
observation



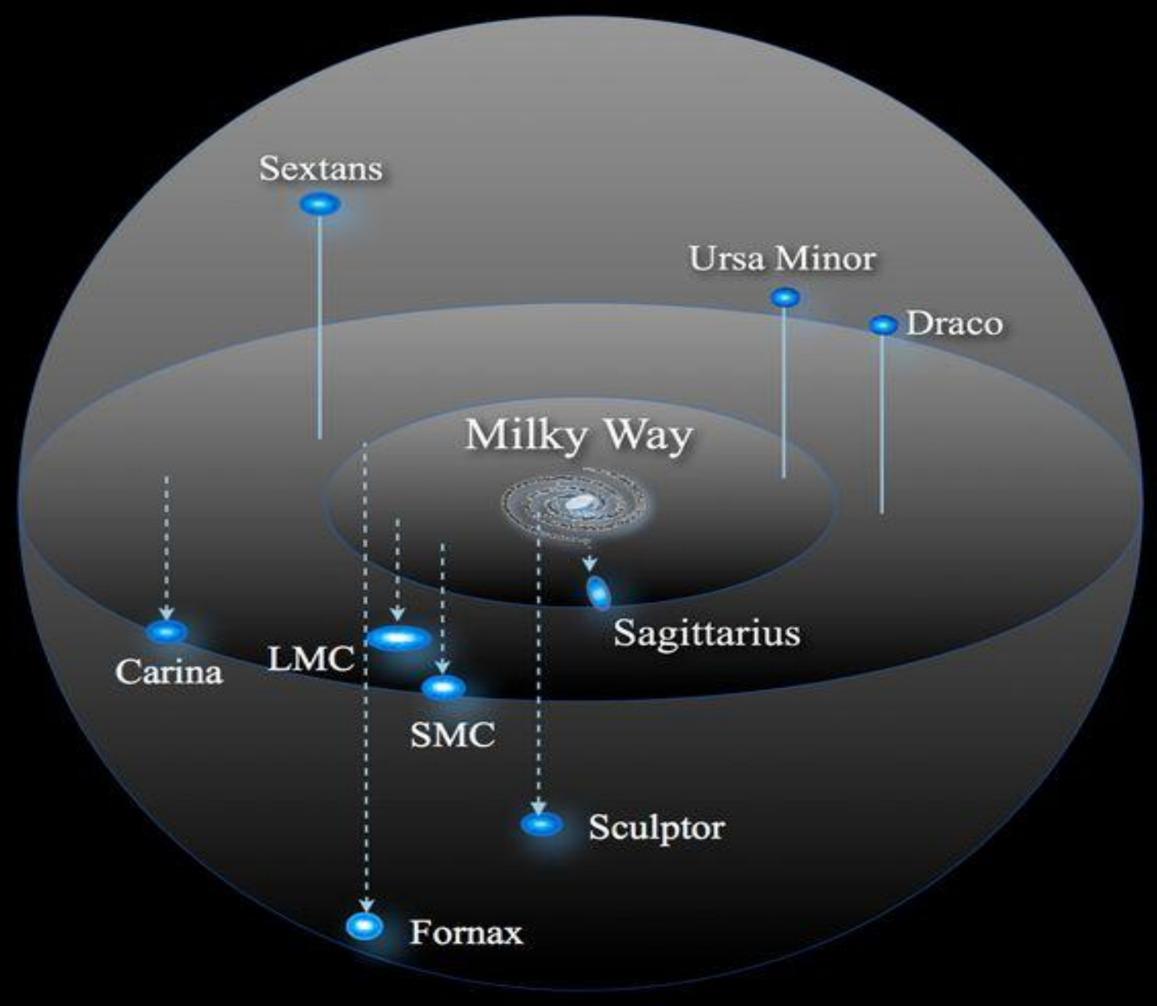
Cosmological simulation based on ΛCDM

Small-scale challenges to ΛCDM： satellite number, structure, and spatial distribution ALL WRONG!

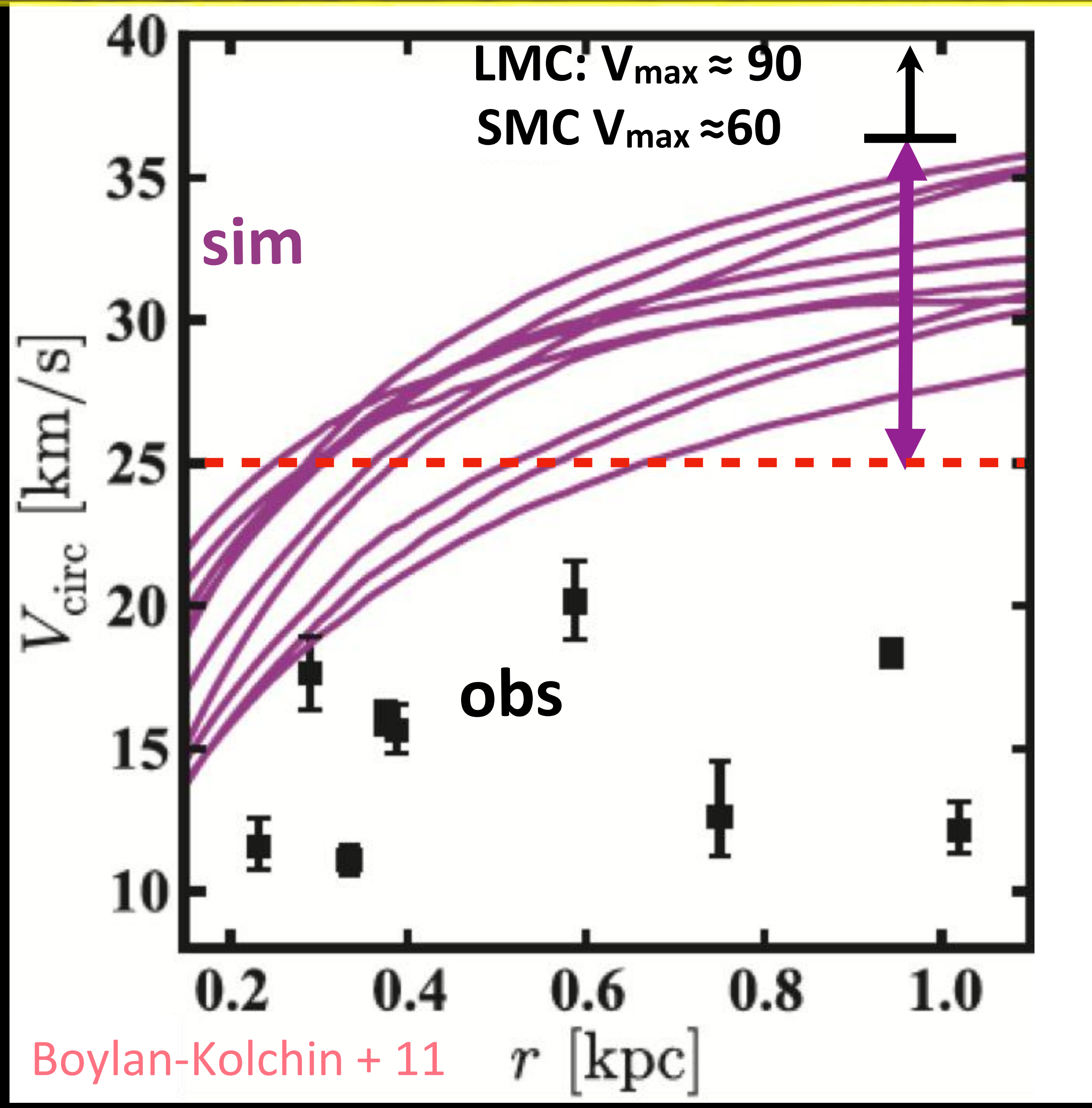
Simulated MW-mass halo



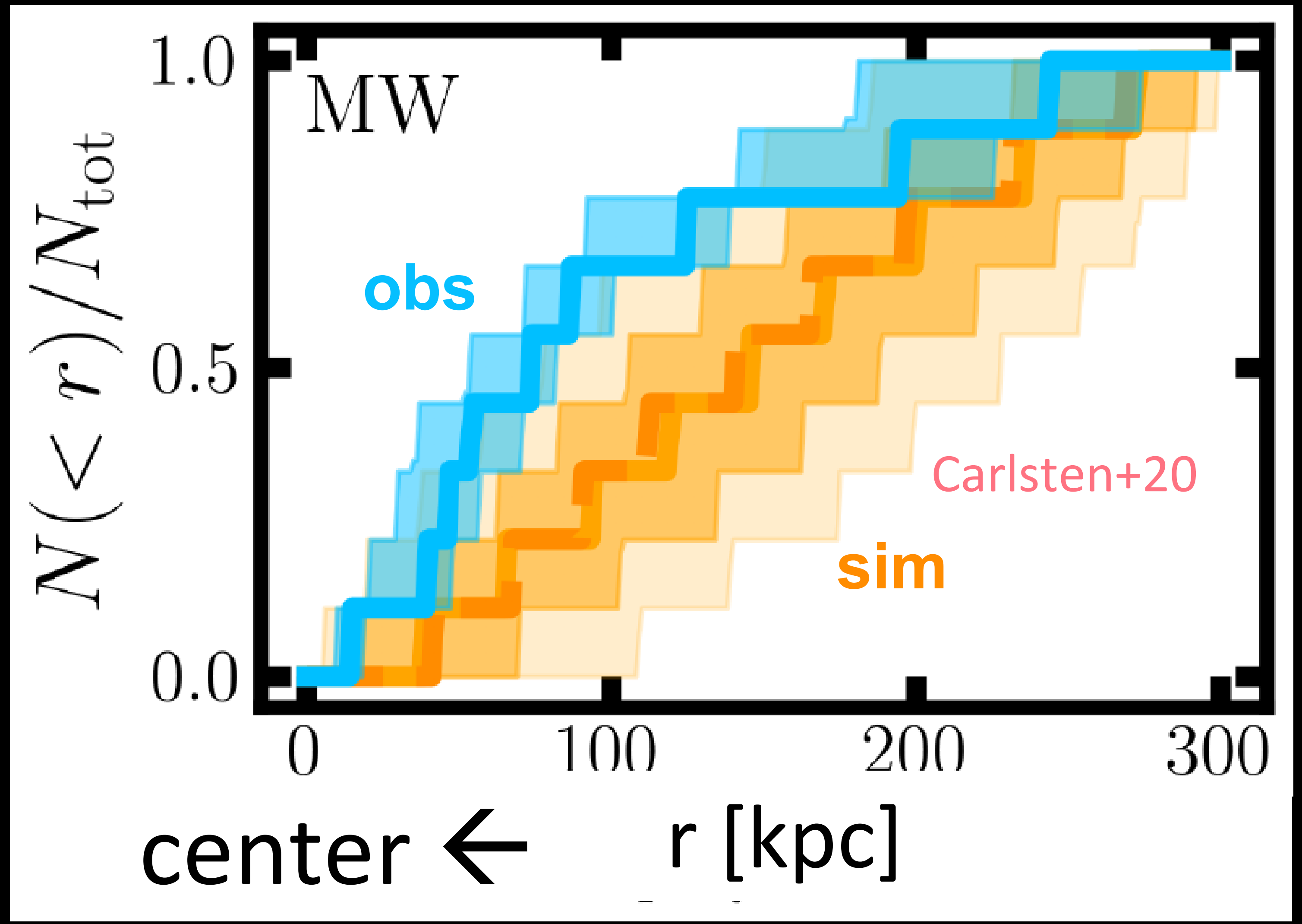
Real MW



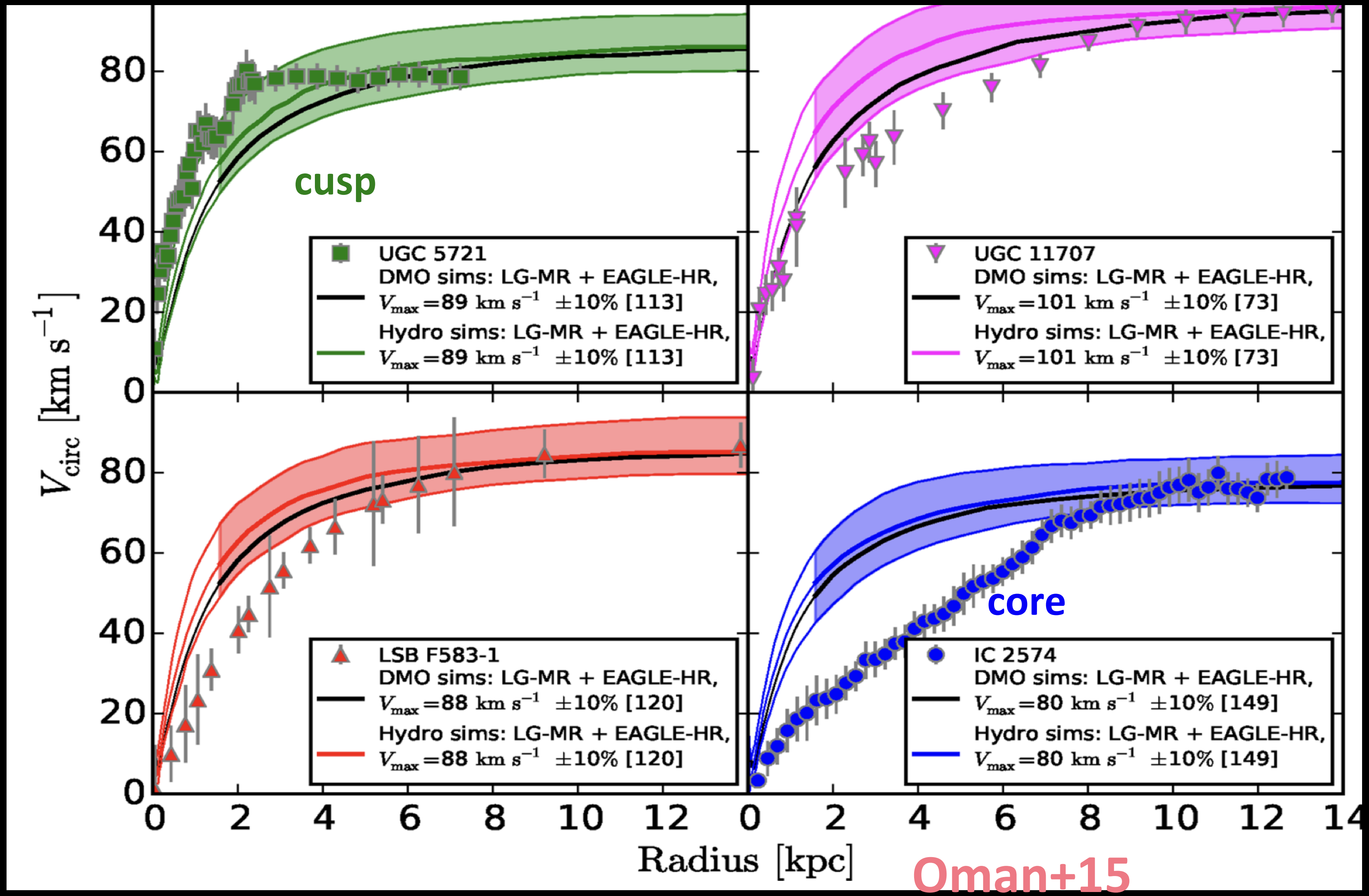
Weinberg + 15



- Missing satellite
- Spatial distribution

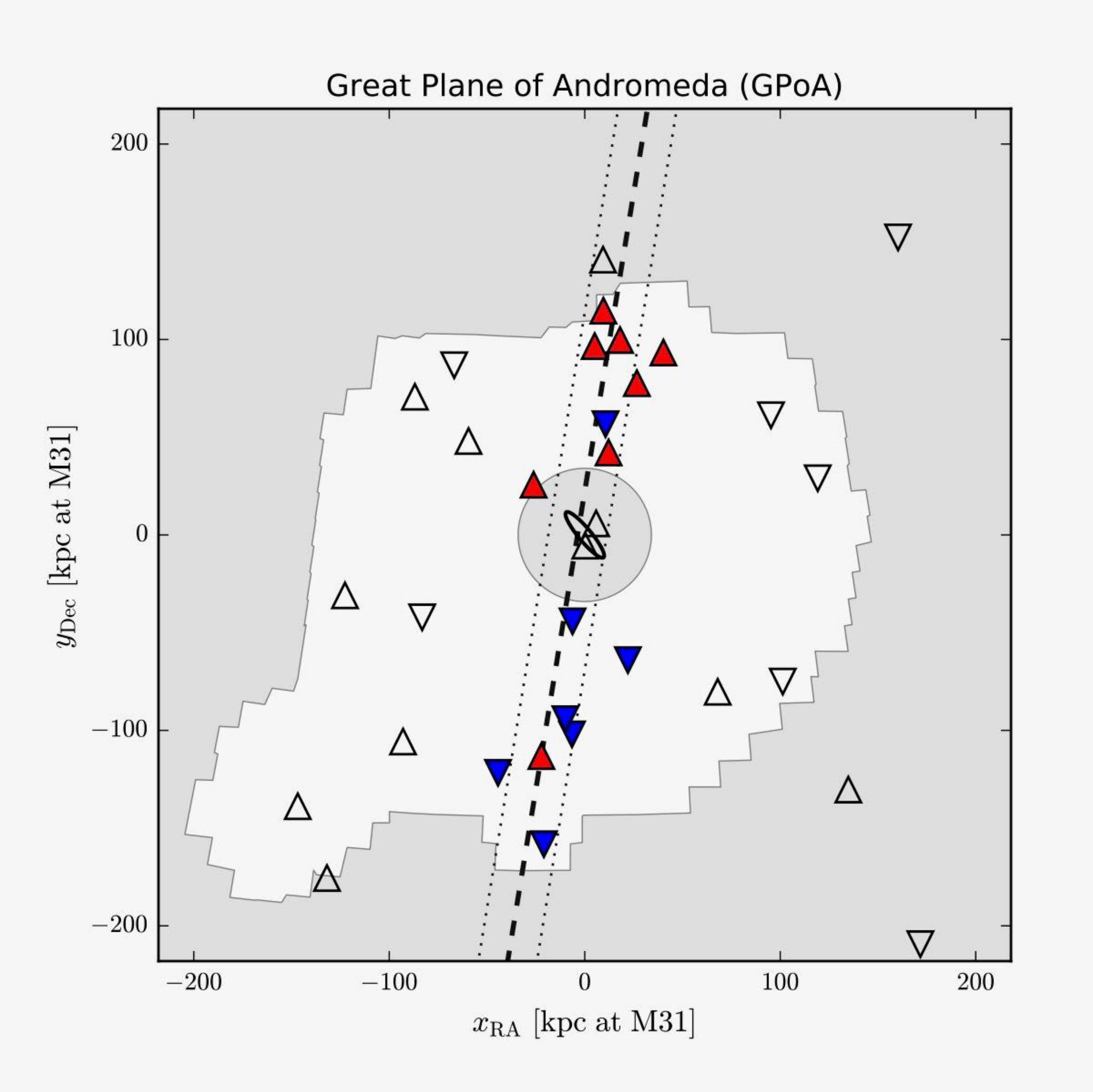


- "too big to fail"——massive satellites too dense
- Rotation curve diversity

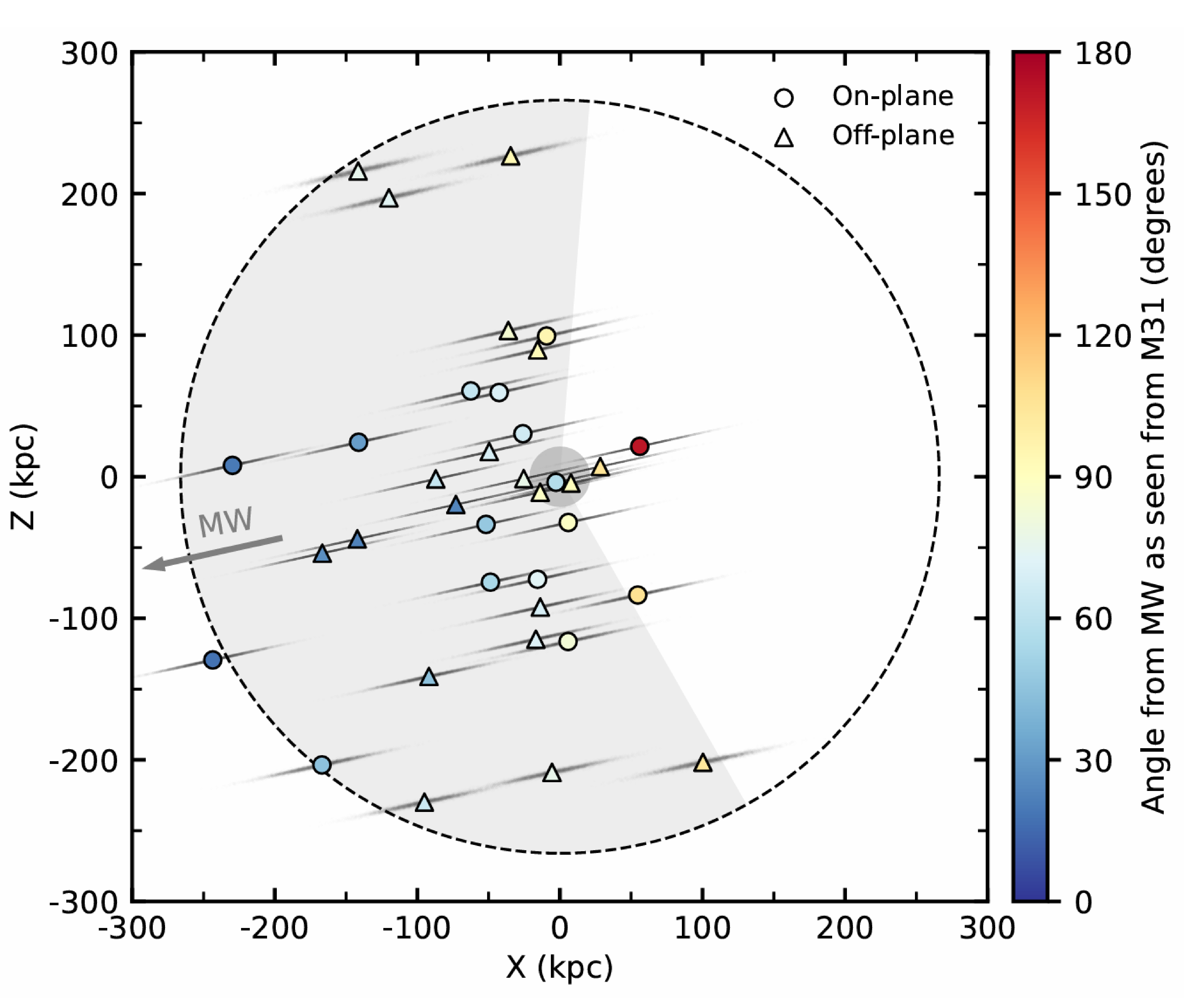


Small-scale challenges to ΛCDM： satellite number, structure, and spatial distribution ALL WRONG!

- **Plane of satellites**



- **Lopsidedness of satellites**

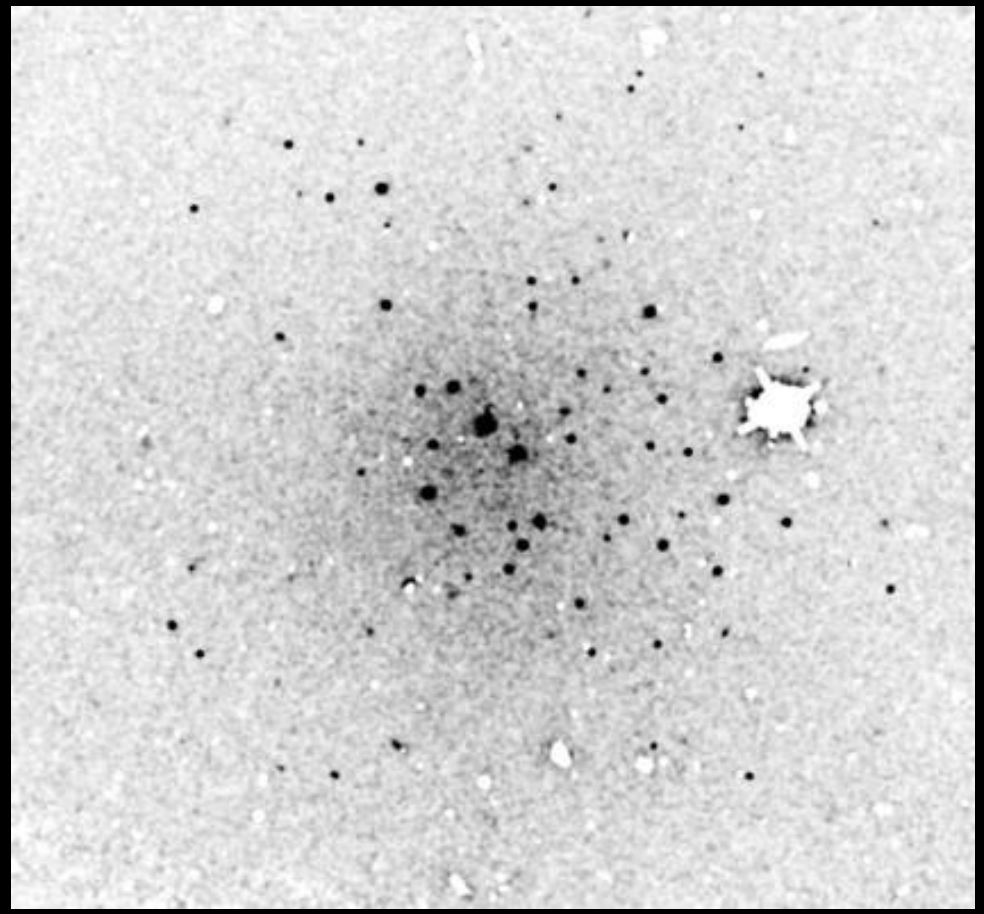


Small-scale challenges to ΛCDM： striking diversity of dwarf galaxies, hard to reproduced in cosmological ΛCDM simulations

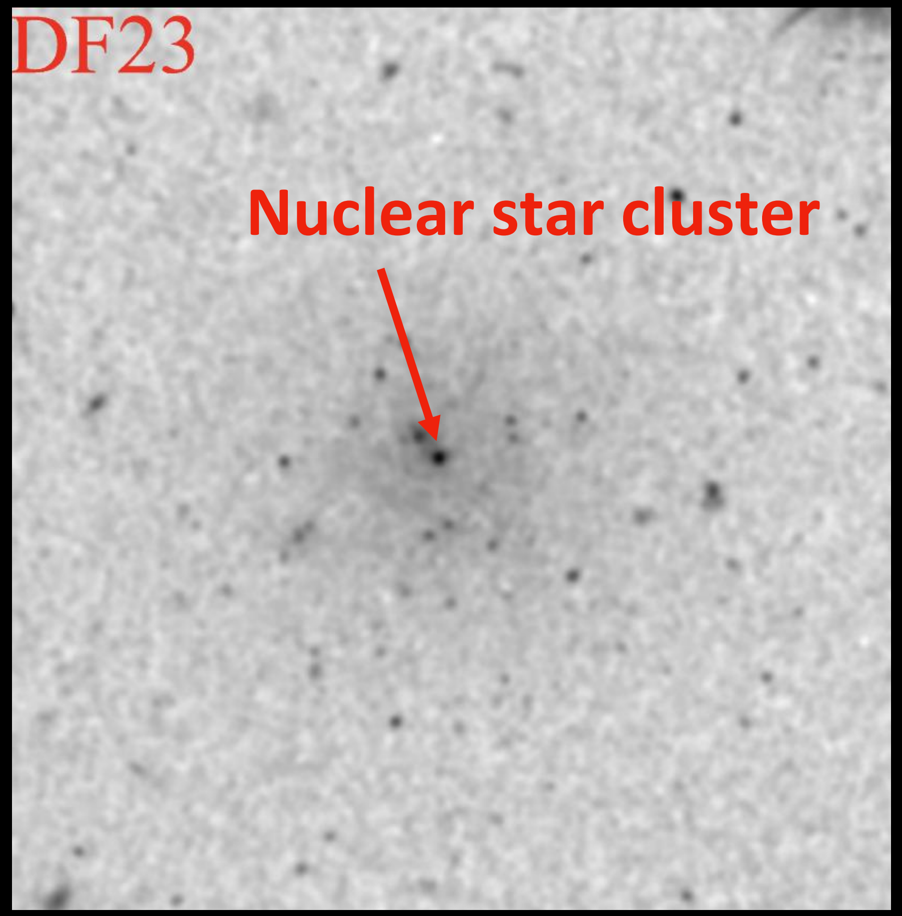
- **Co-existence of compact & diffuse dwarfs**

- **Dwarfs with a lot of globular clusters**
- **Nucleated dwarfs**

Liang, **Jiang*** + 24 ApJ
arXiv:2304.14431

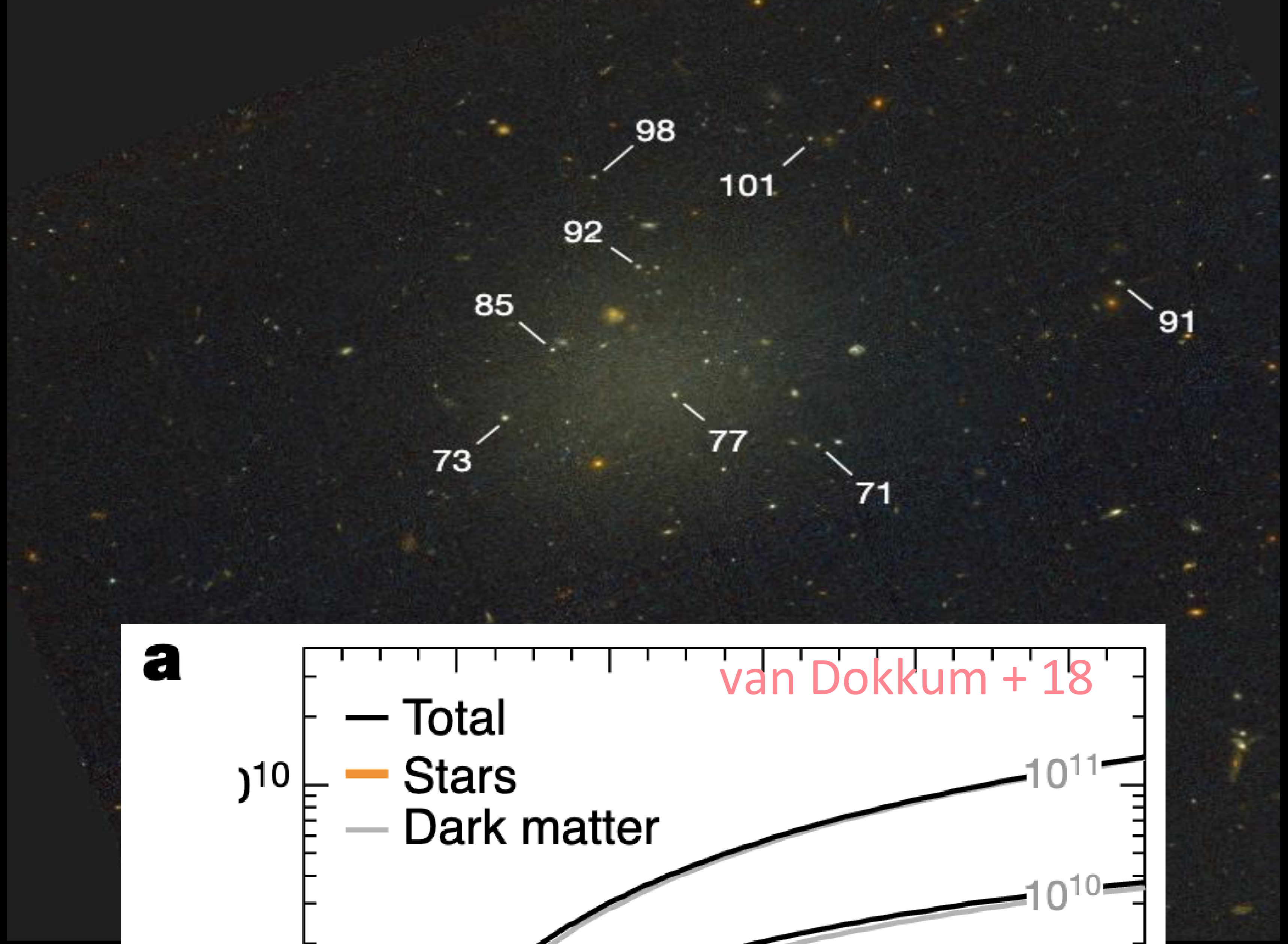


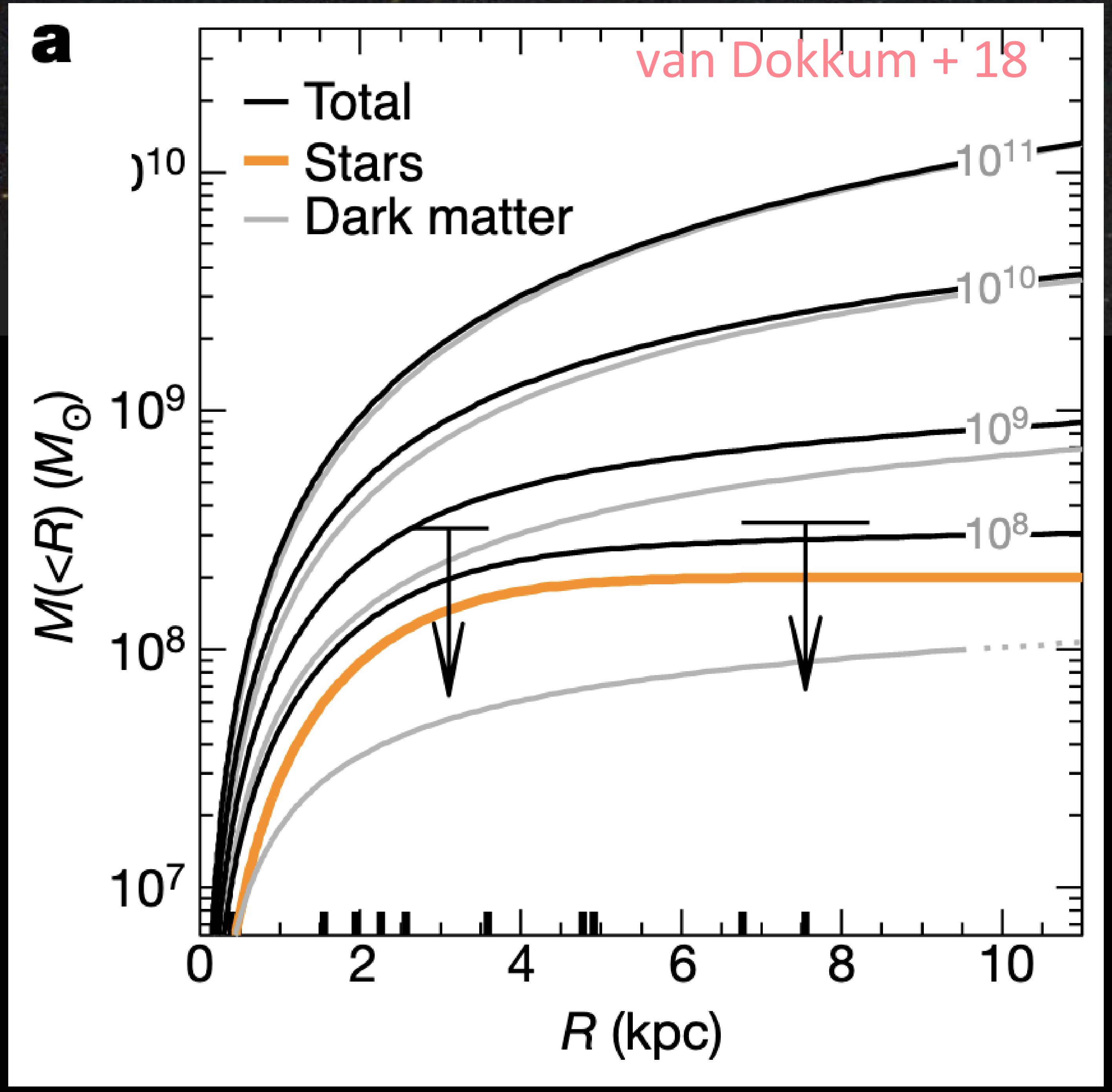
Danieli + 21



Lim + 18

- **Dark-matter deficient galaxies**





Small-scale puzzles (broadly speaking): Disk Galaxies – Baryonic Tully-Fisher Relation

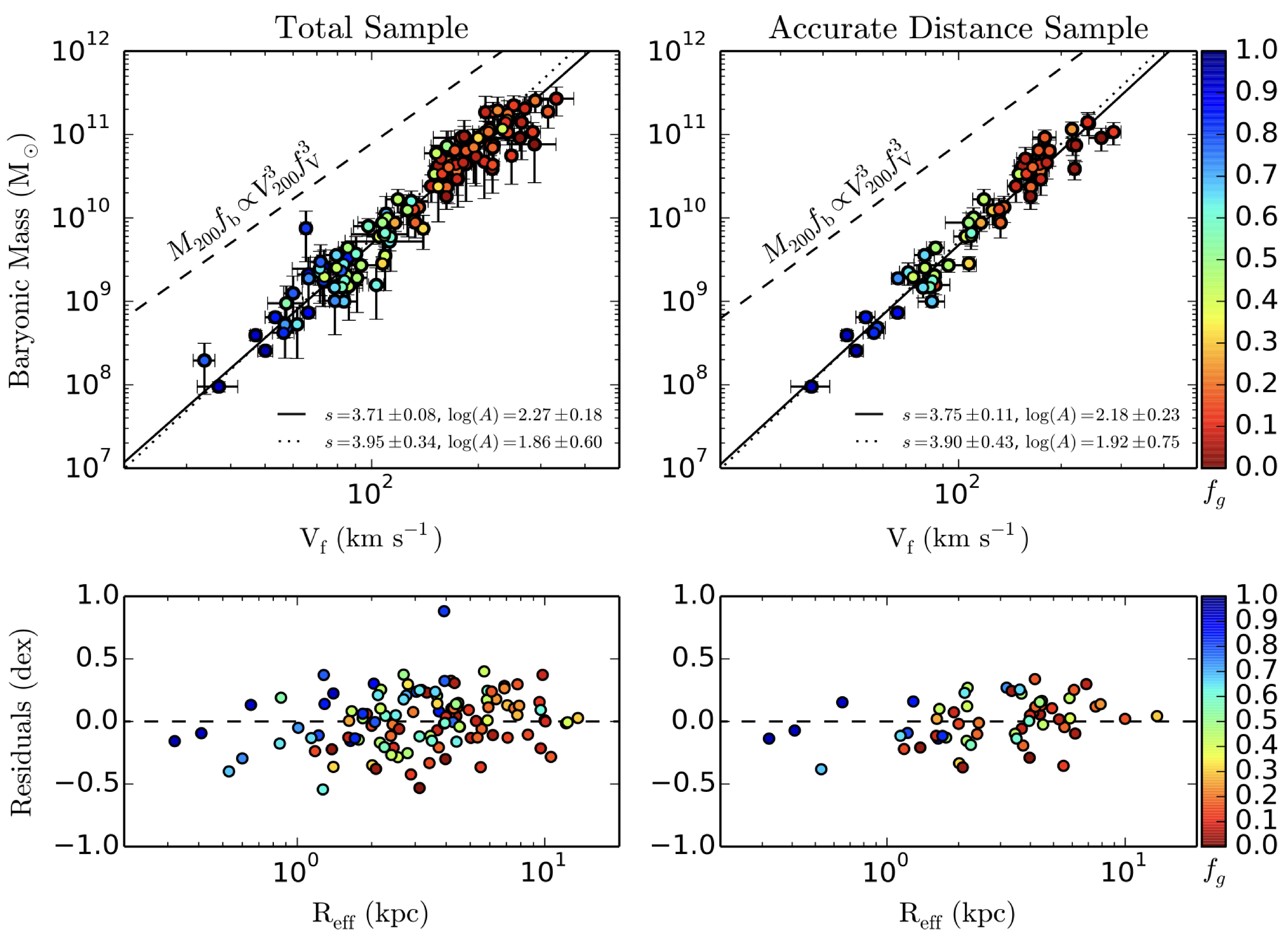
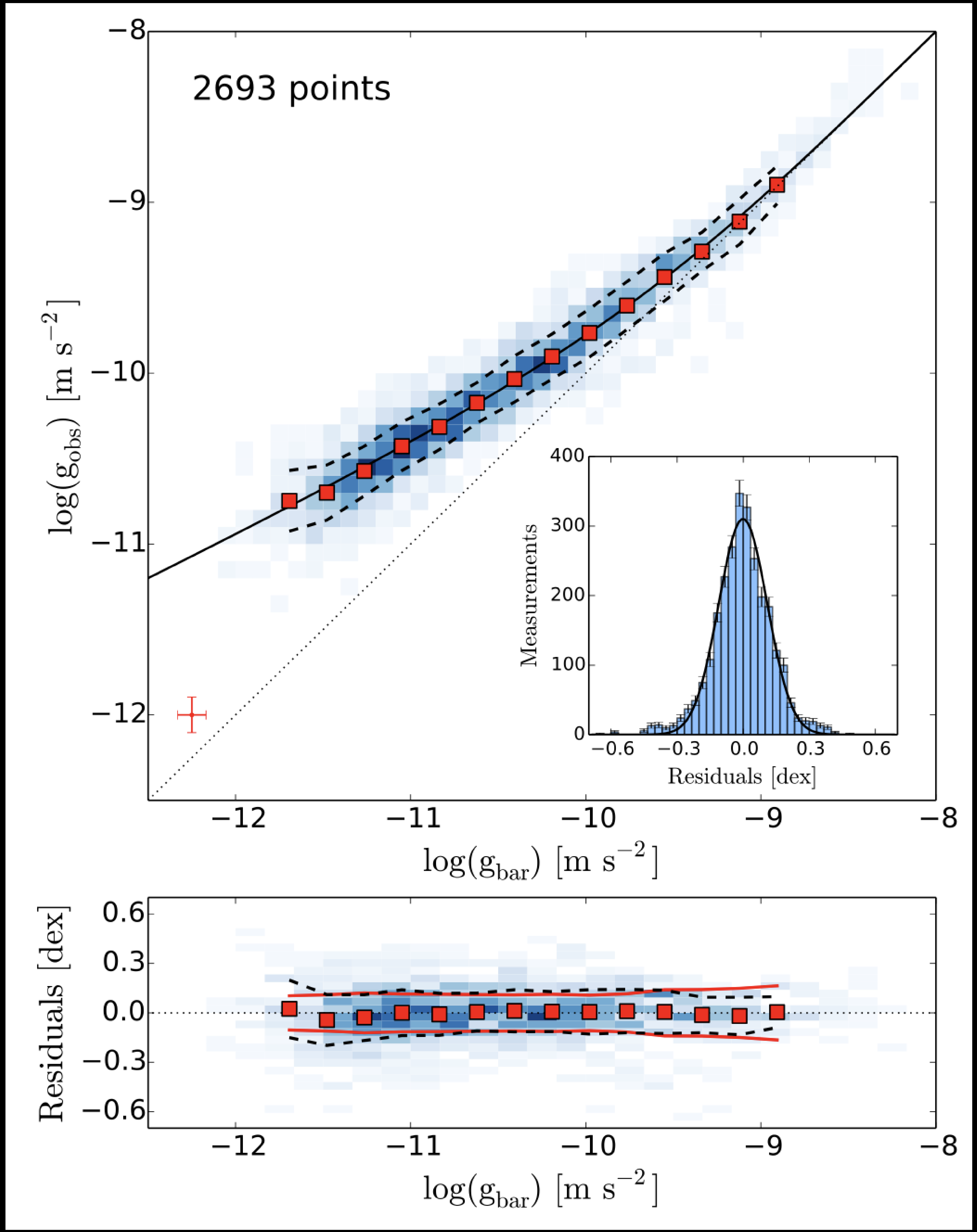


FIG. 2.— *Top panels*: BTFR adopting $\Upsilon_* = 0.5\ M_\odot/L_\odot$. Galaxies are color-coded by $f_g = M_g/M_b$. Solid lines show error-weighted fits. Dotted lines show fits weighted by f_g^2, increasing the importance of gas-dominated galaxies. The dashed line shows the ΛCDM initial condition with $f_V = 1$ and $f_b = 0.17$ (the cosmic value). *Bottom panels*: residuals from the error-weighted fits versus the galaxy effective radius. The outlier is UGC 7125, which has an unusually high correction for Virgocentric infall and lies near the region where the infall solution is triple-valued. If we consider only the correction for Local Group motion, UGC 7125 lies on the BTFR within the scatter.

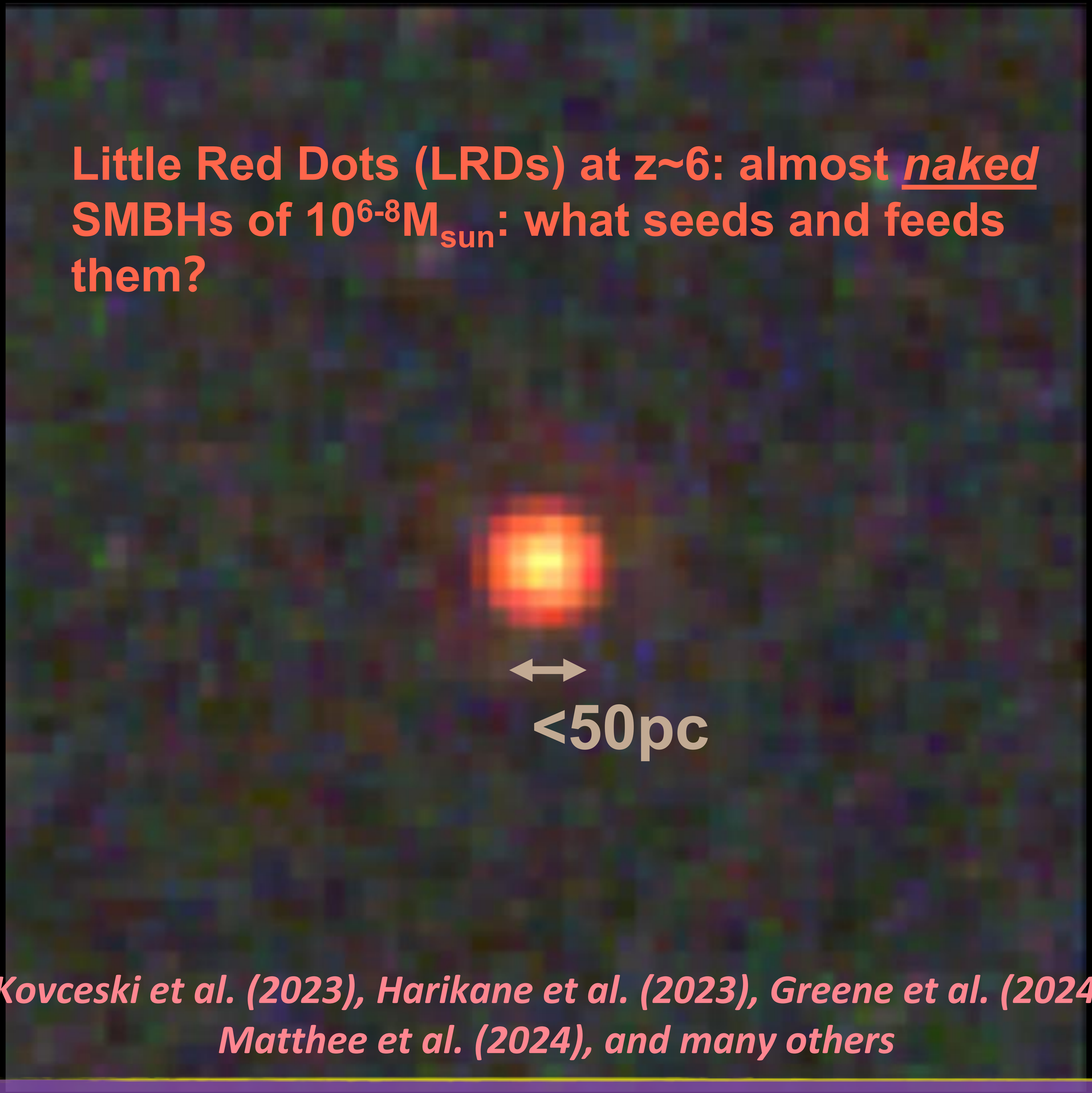
Lelli+2015 SPARC Survey

Small-scale puzzles: Disk Galaxies Radial Acceleration Relation (RAR)

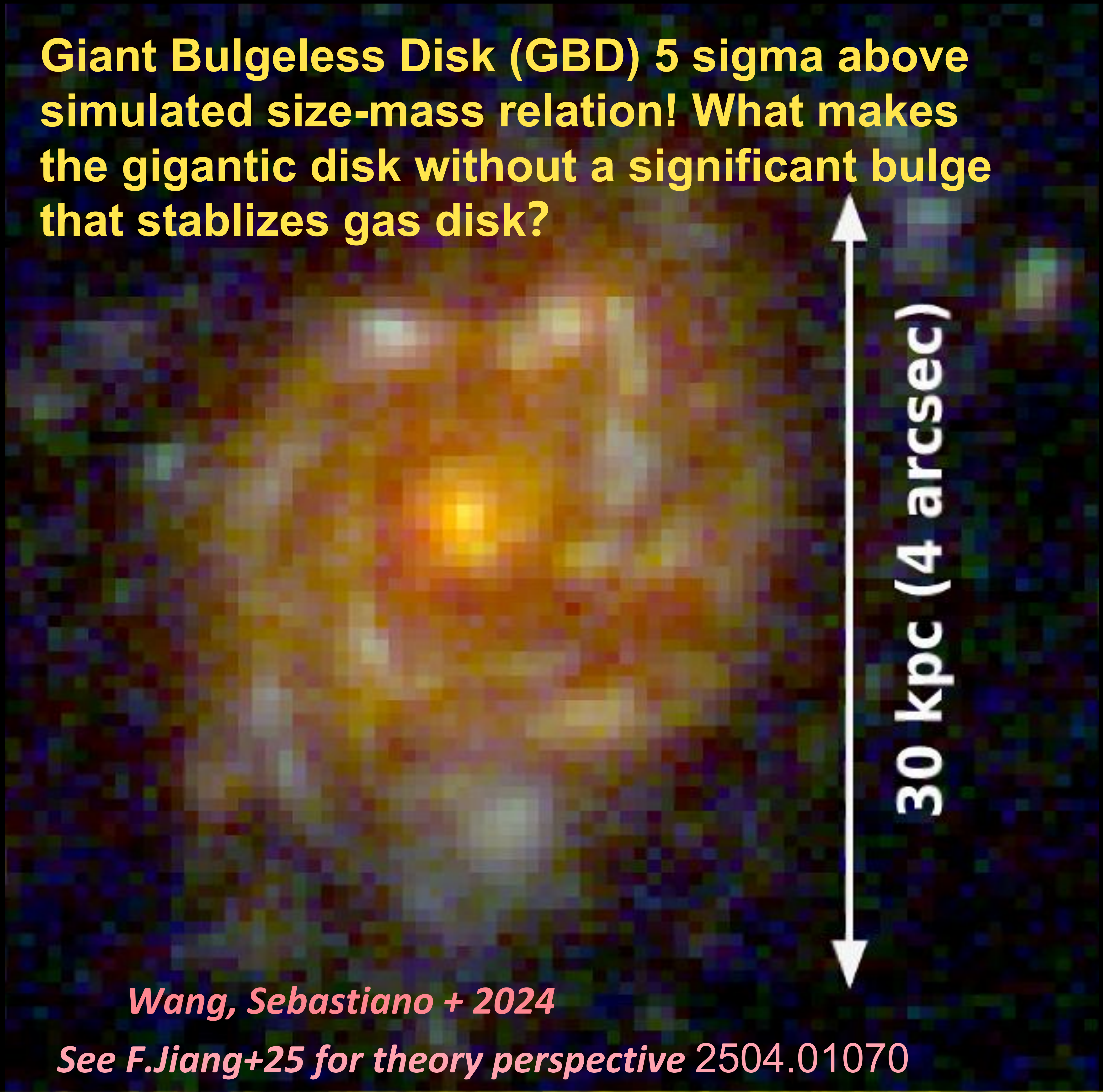


McGaugh, Lelli & Schombert
2016 SPARC Survey

Small-scale puzzles in the JWST era： extreme galaxies in the early Universe



Kovceski et al. (2023), Harikane et al. (2023), Greene et al. (2024), Matthee et al. (2024), and many others



Wang, Sebastiano + 2024

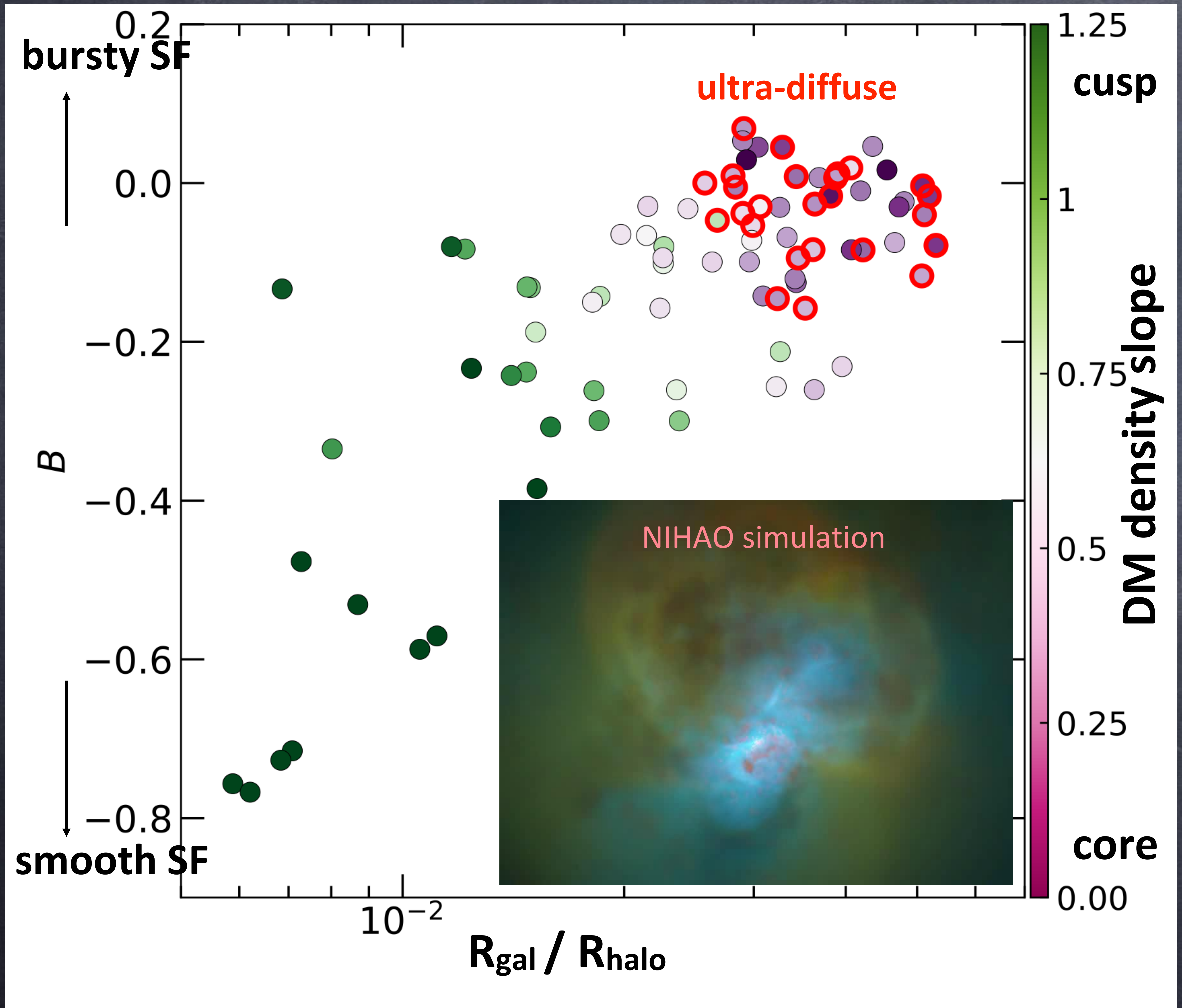
See F.Jiang+25 for theory perspective 2504.01070

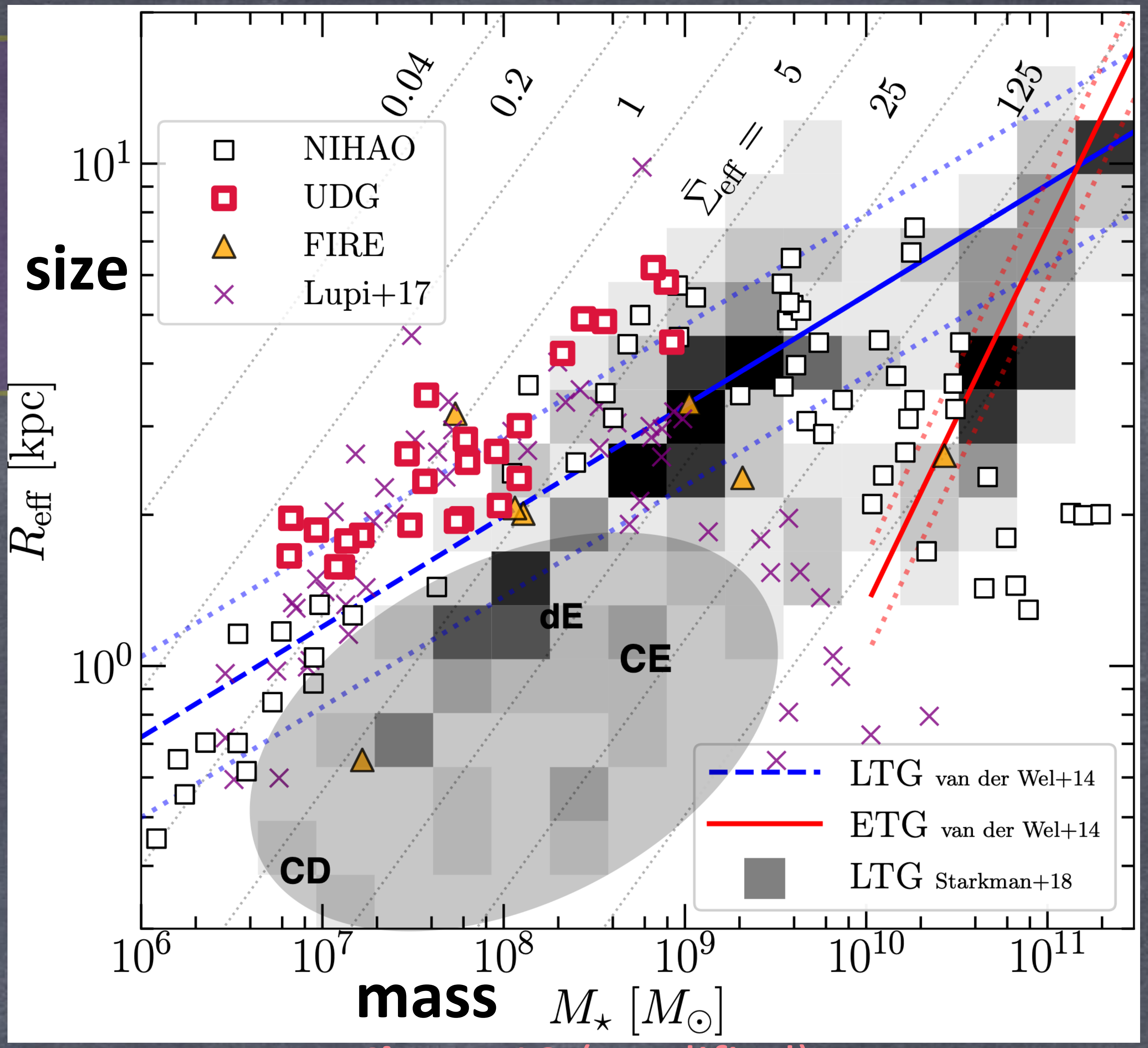
Extreme-morphology systems stress-test our understanding of gal-halo connections

I. Can we solve the small-scale puzzles within ΛCDM, accounting for baryonic and environmental processes ?

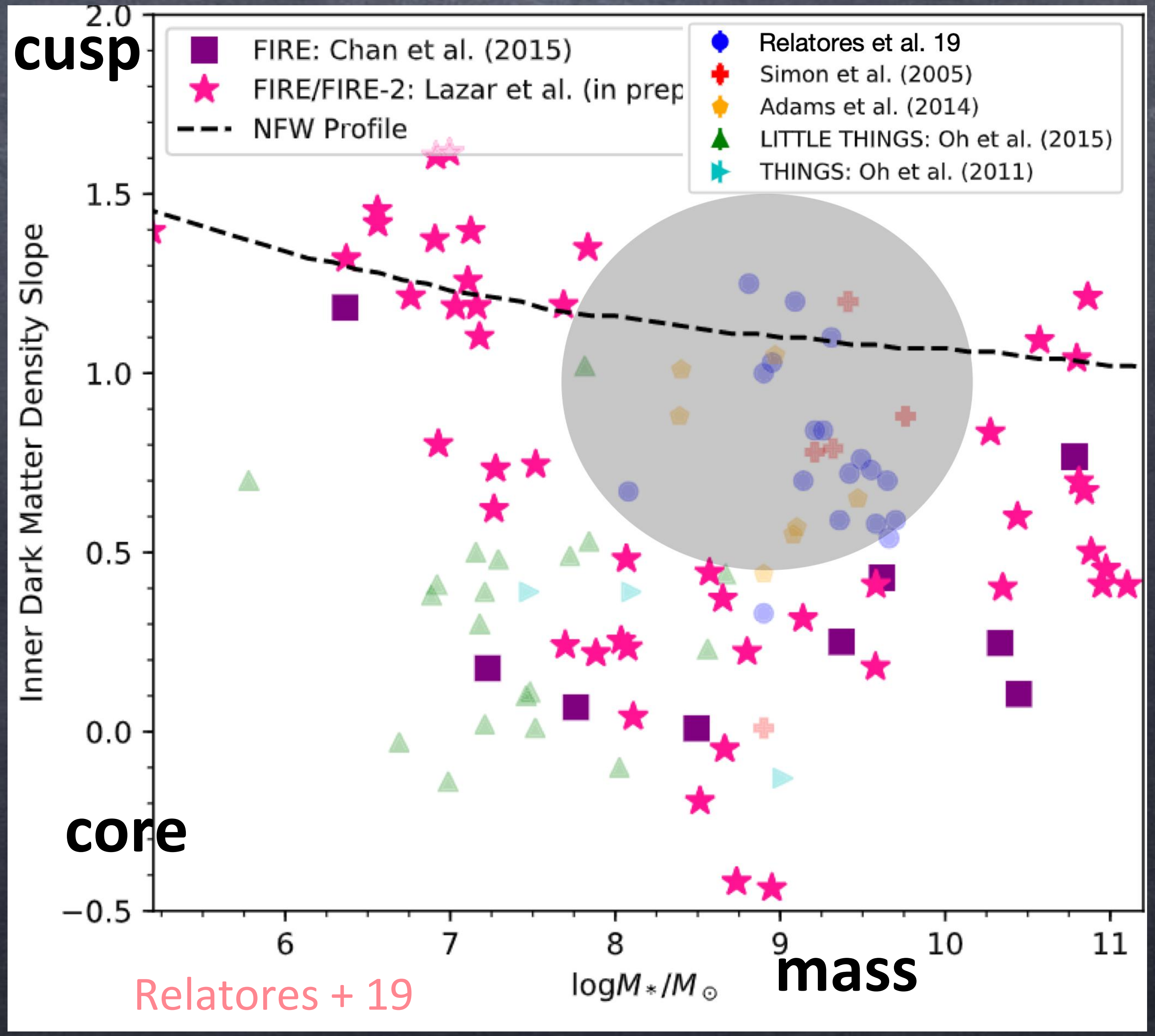
Modern simulations over-produce puffy galaxies (cored halos), but likely are short of compact (cuspy) systems, in bright-dwarf regime ($M_{star} = 10^{7.5-9.5}$ M_{sun})

- Star-formation / feedback too bursty?





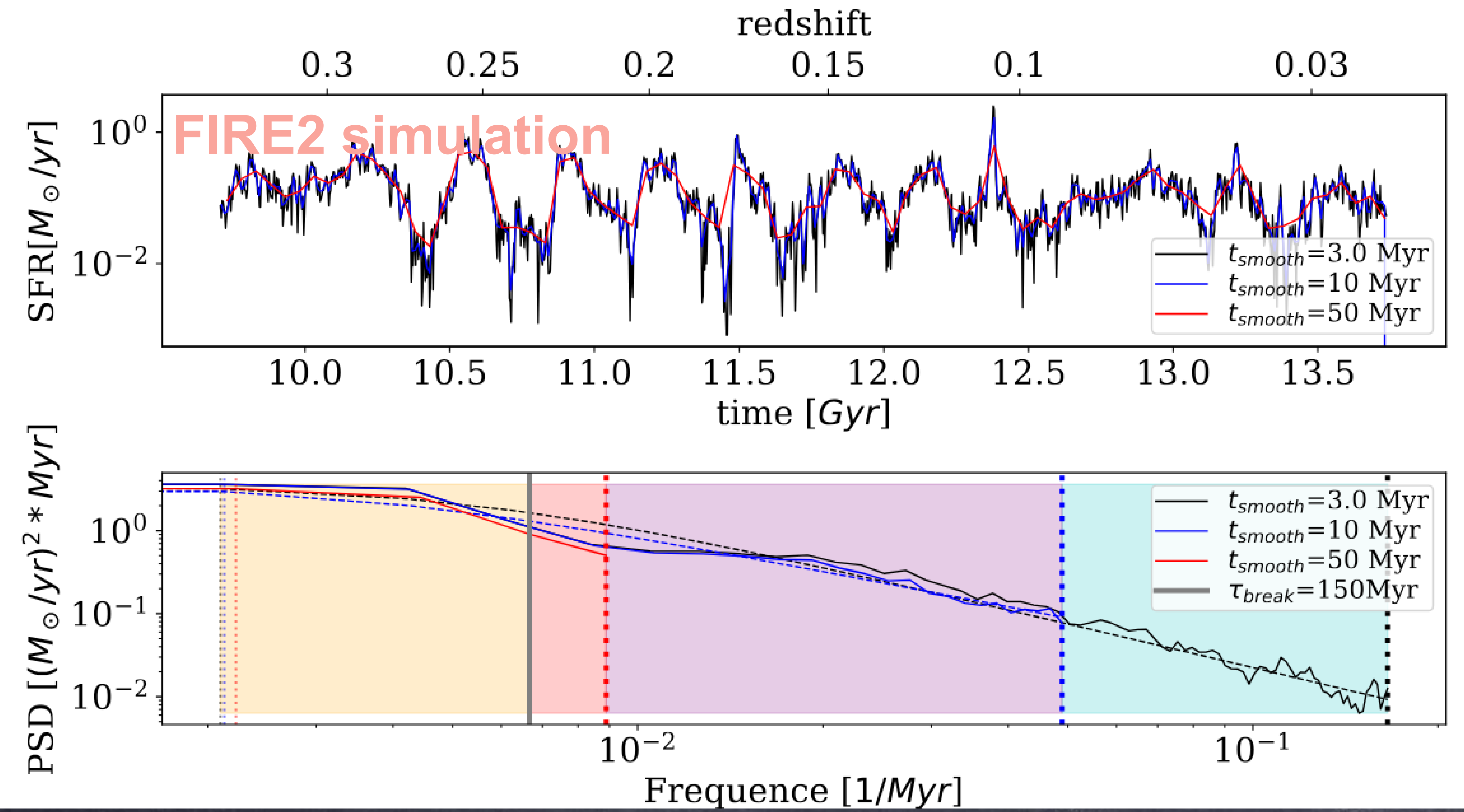
Jiang+19 (modified)

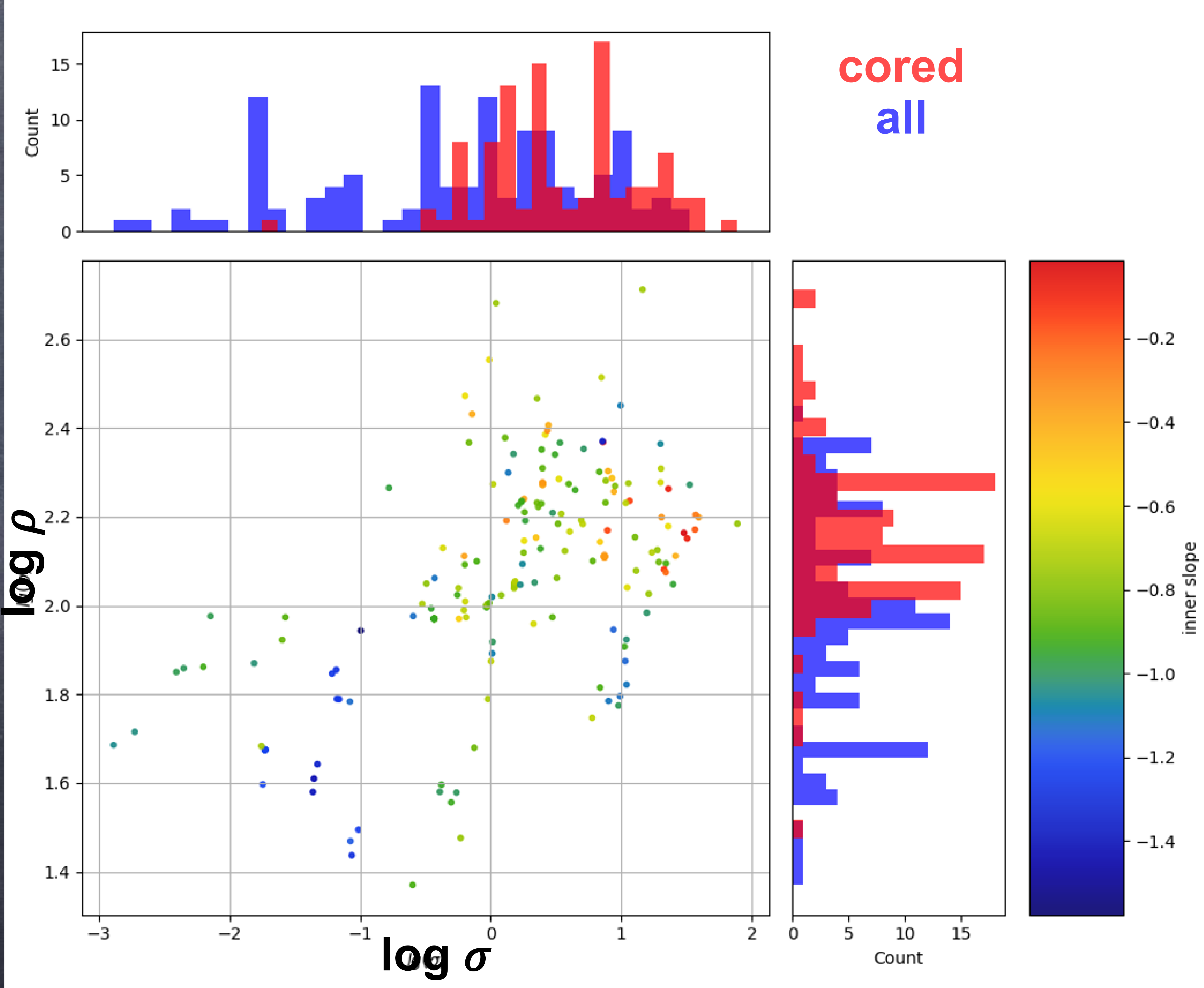


Core formation is strongly correlated with bursty star formation

$$\mathrm{PSD}(f) = \frac{\sigma^2}{1 + (\tau_{\mathrm{break}} f)^\alpha}$$

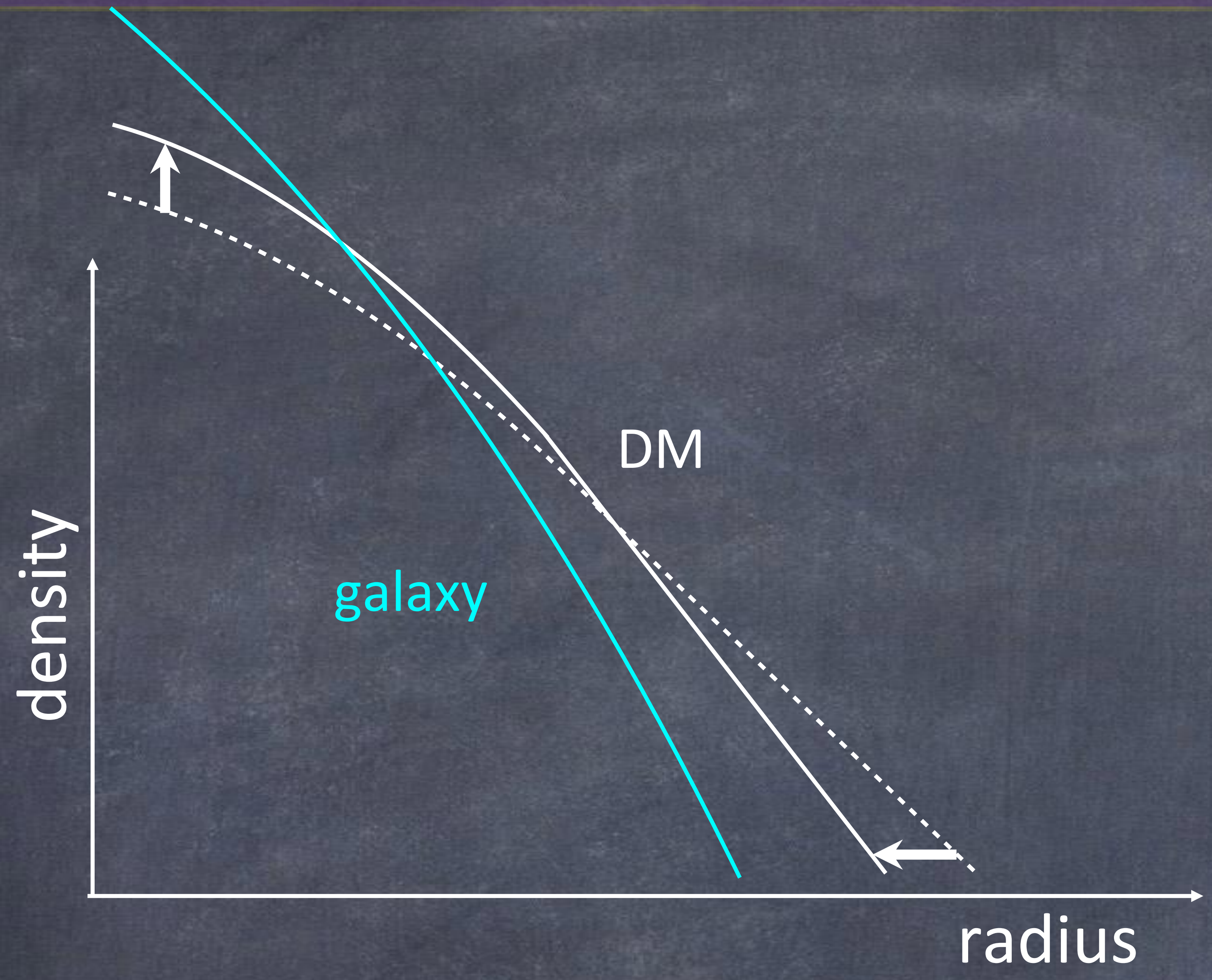
He, **Jiang***+25 in prep

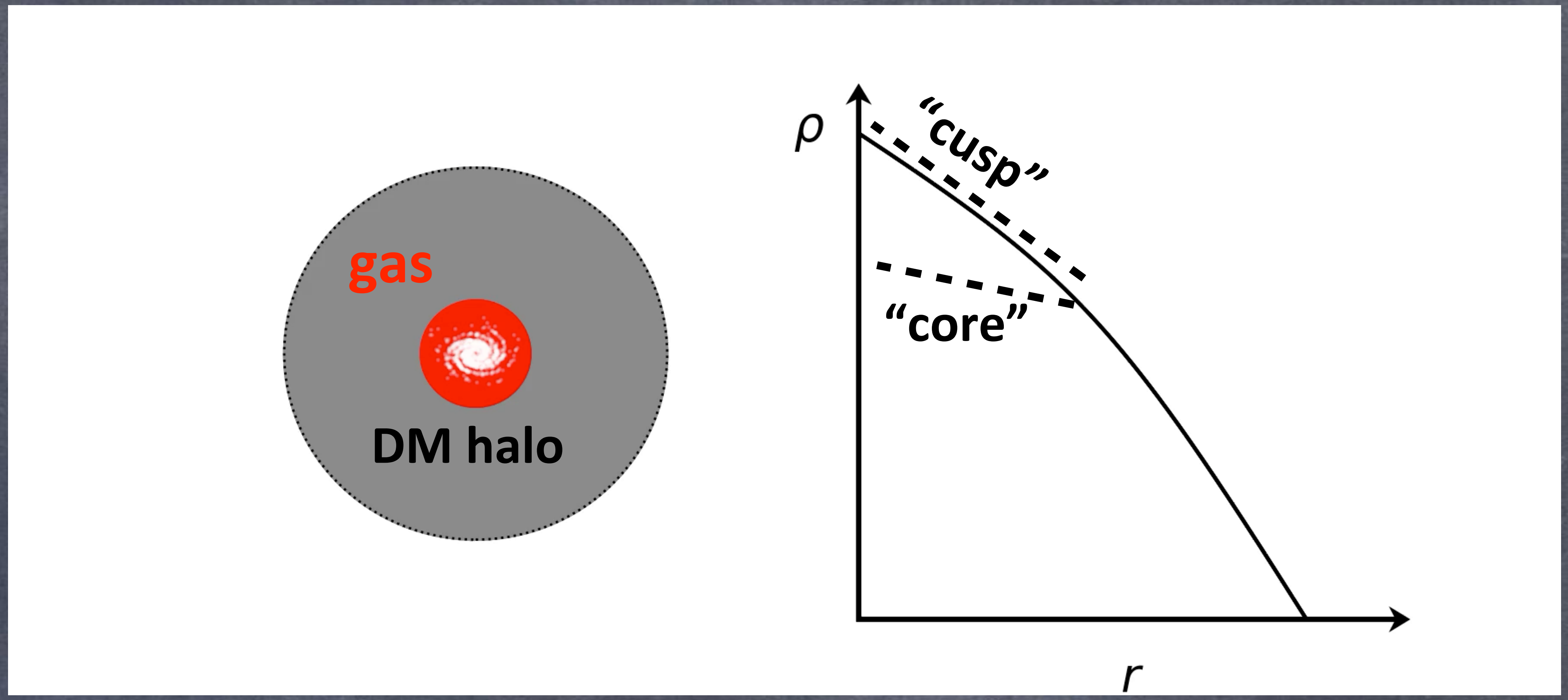




$$\rho = \frac{\tau_{\mathrm{break}}\sqrt{\alpha - 1}}{2\pi}$$

CDM halos respond to baryonic processes — contraction & expansion

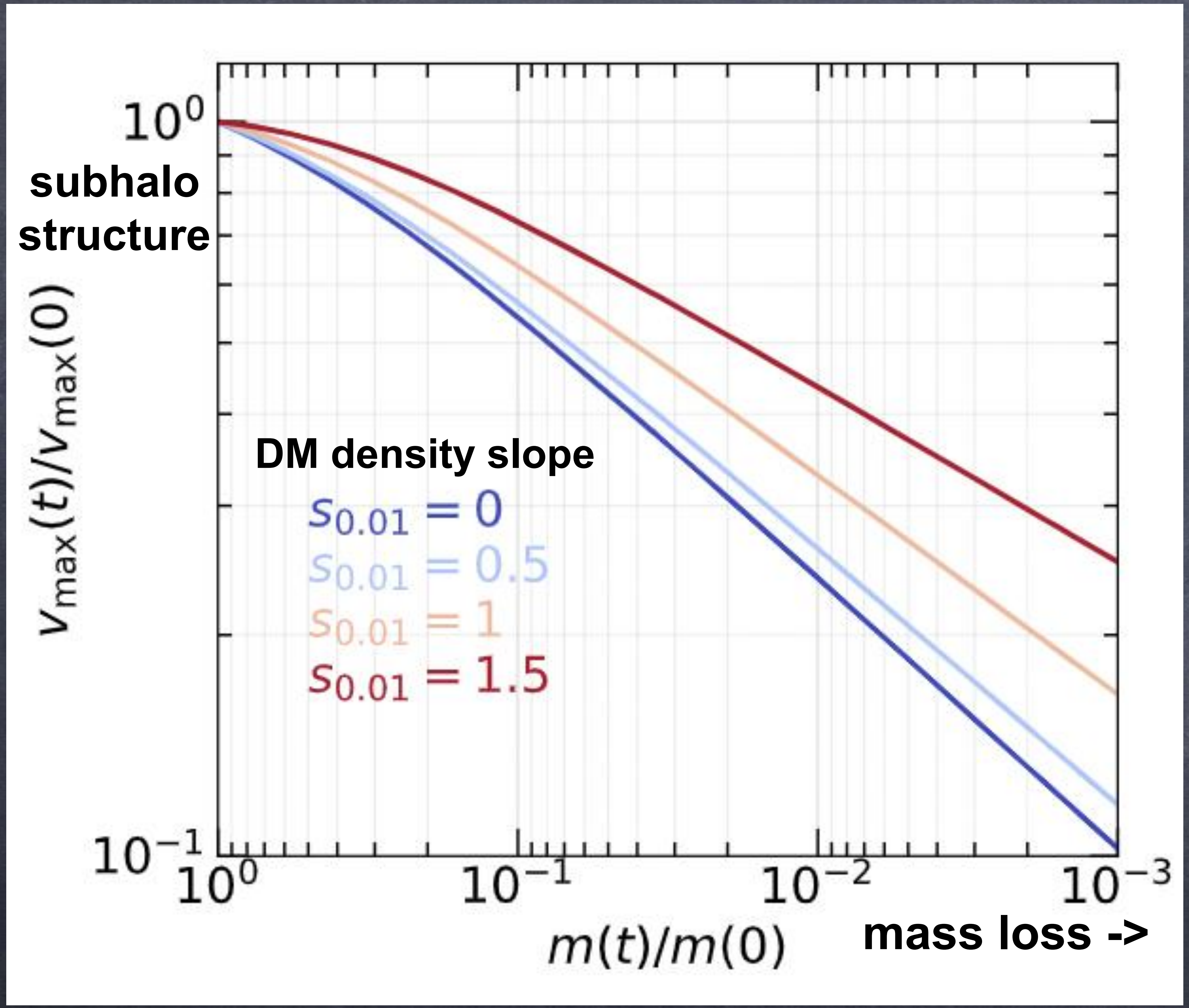


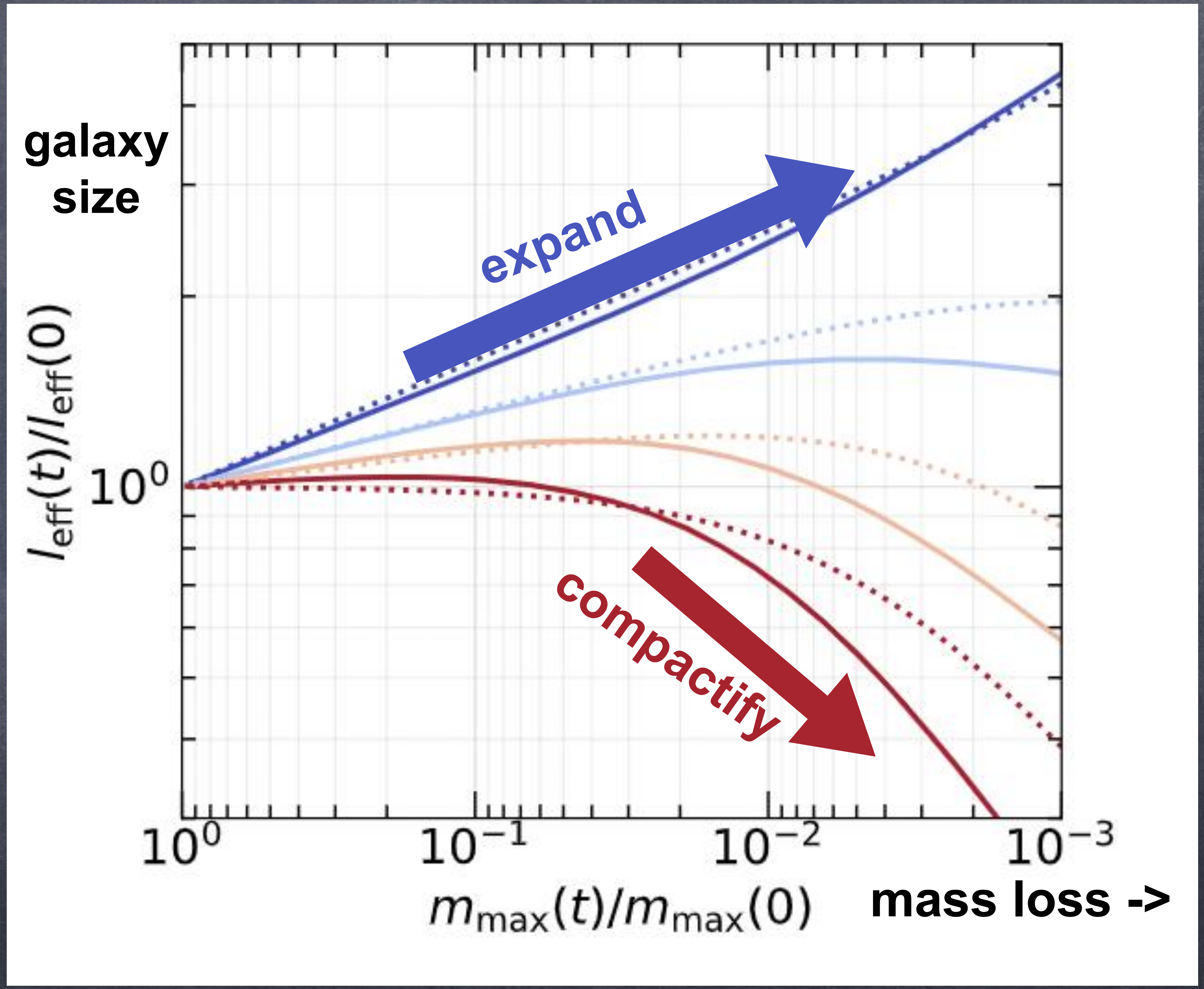


- Baryonic potential —> adiabatic contraction
- Feedback driven gas outflows —> expansion and cusp-core transformation

“Halo response” amplified in dense environments — enhancing diversity

- The “**tidal evolution tracks**” depend on the initial structures and **diverge** as the satellites evolve

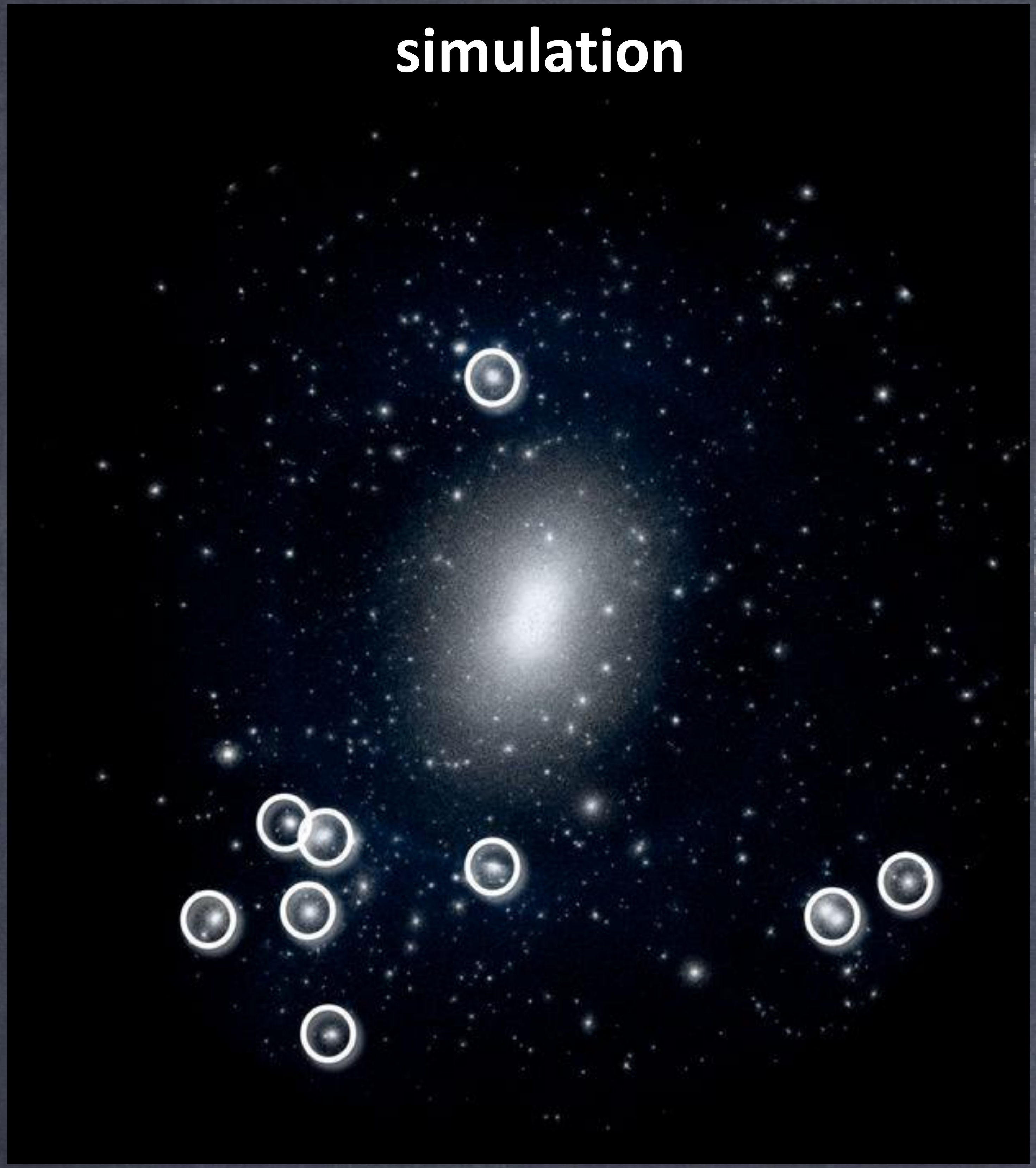




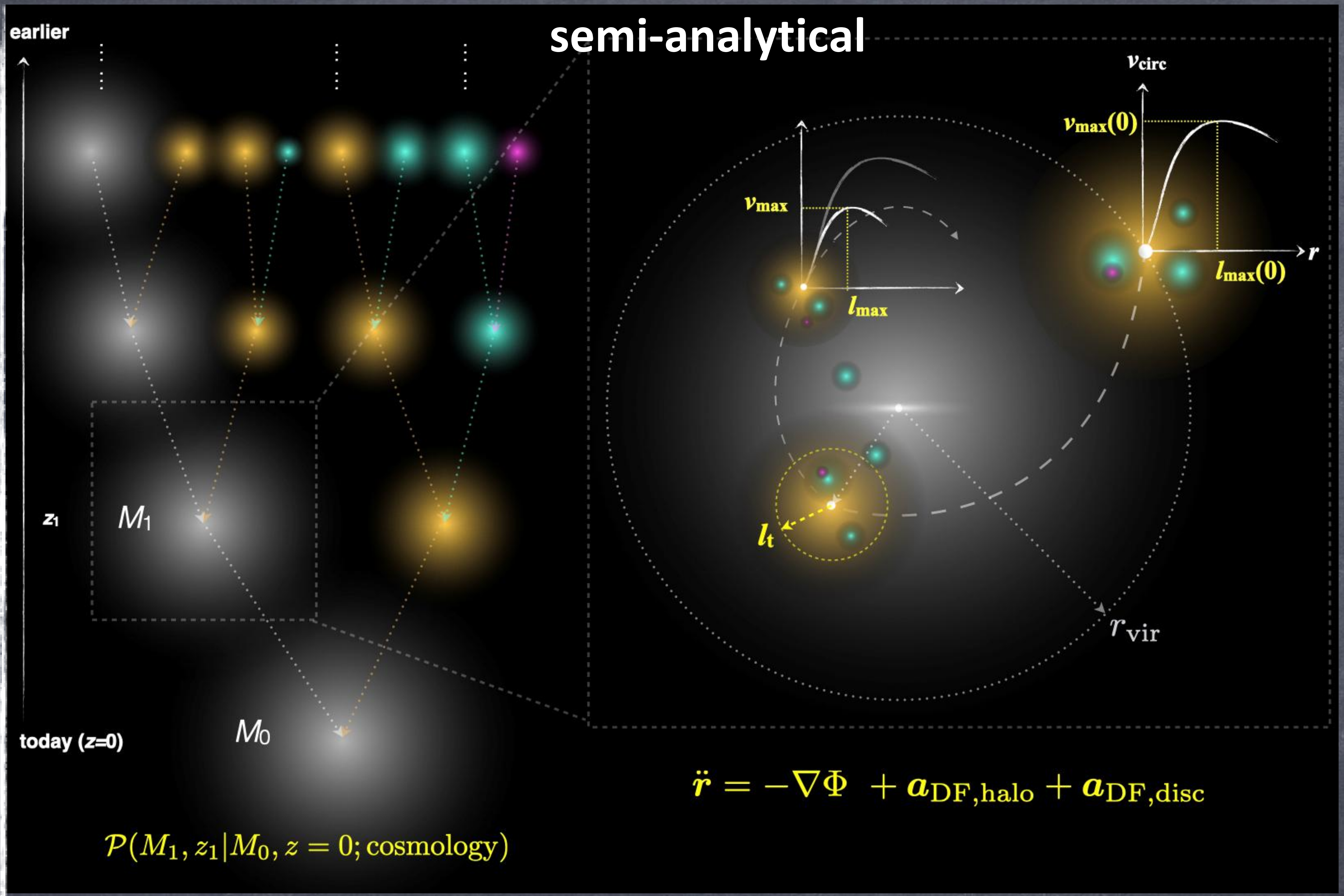
Hayashi+03; Penarrubia+08,10; Errani+15,18

Stucker+22; Benson & Du 22, Du + (including **Jiang**) 24, Yan, **Jiang***+ in prep

Comparison of methods: simulation versus semi-analytical model



ELVIS simulation (Garrison-Kimmel+14)



"SatGen" (Jiang & van den Bosch 14,15,16,17; Jiang+21)

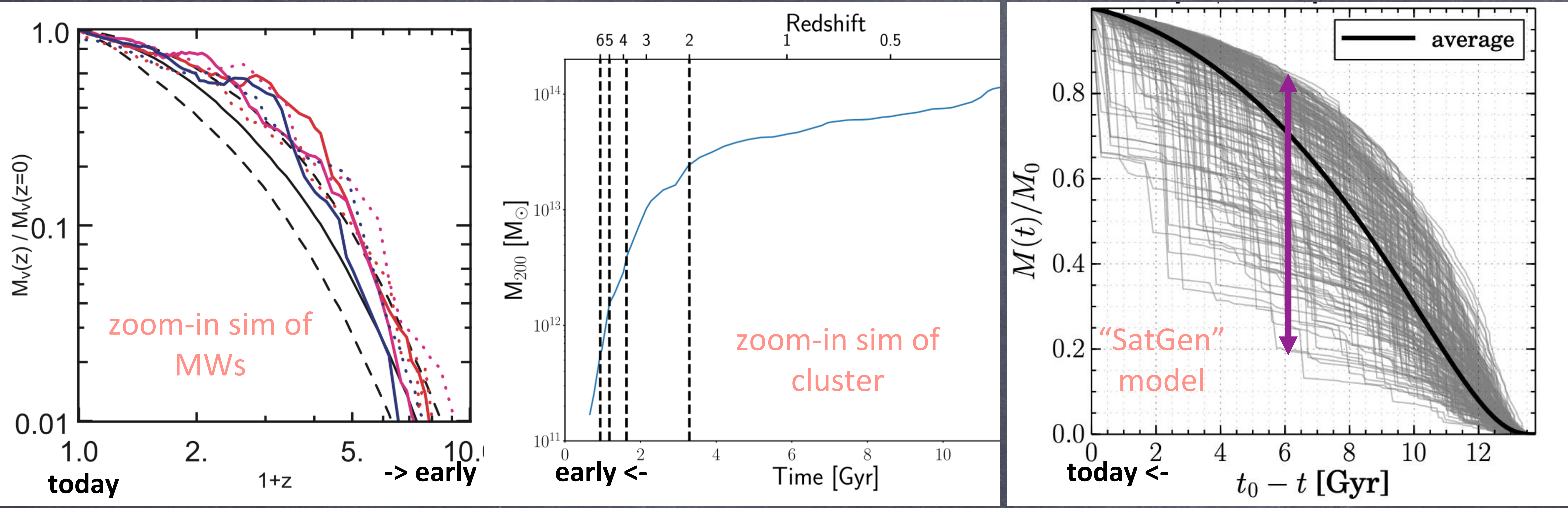
- operate on particles/cells: gravity + hydrodynamics + subgrid recipes

- operate on halo merger trees: assign galaxies to halos according to empirical/analytical galaxy-halo connections —> evolve satellites using empirical/analytical recipes for dense environments

https://github.com/JiangFangzhou/SatGen

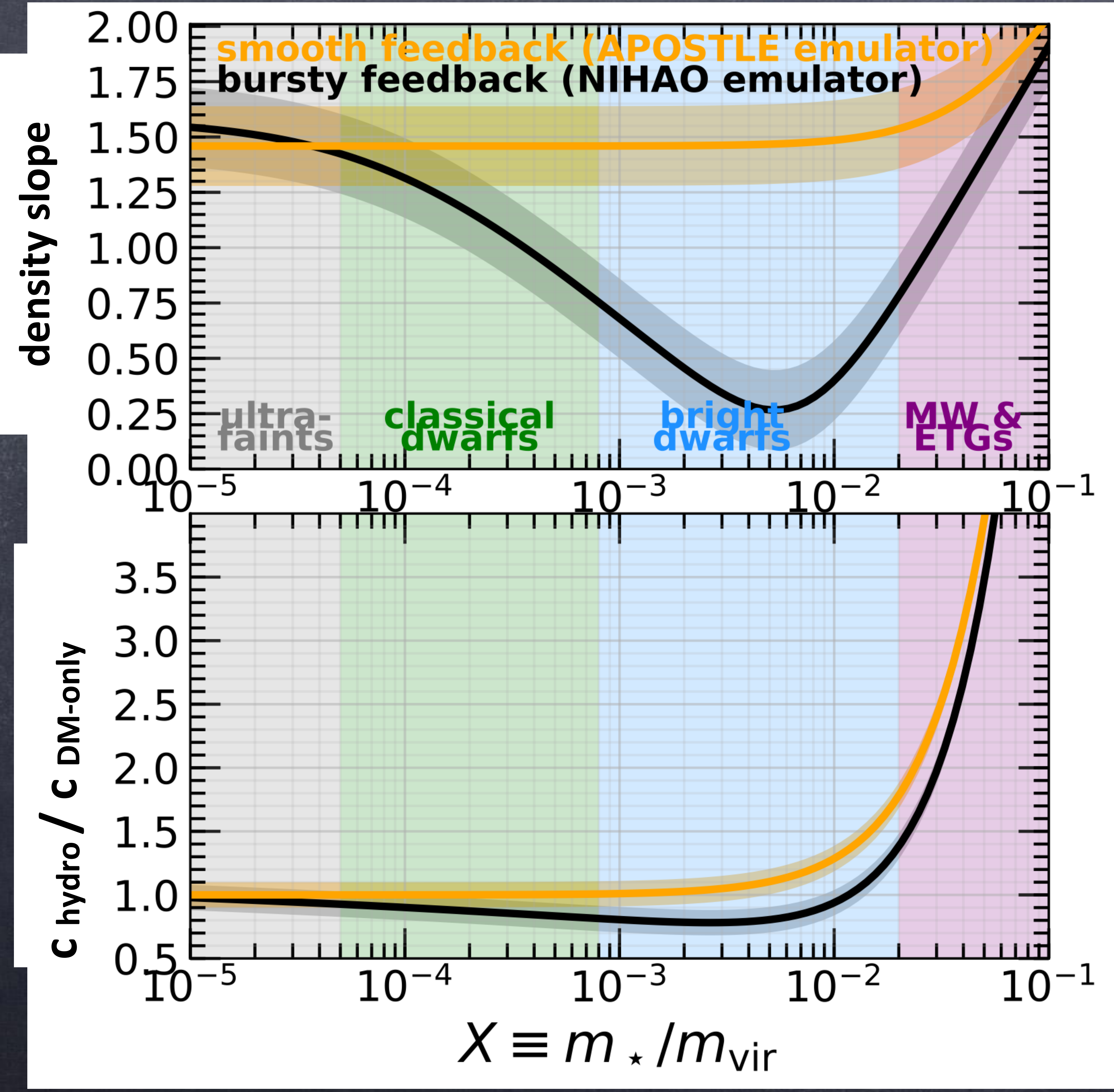
Comparison of methods: simulation versus semi-analytical model

Halo mass assembly histories



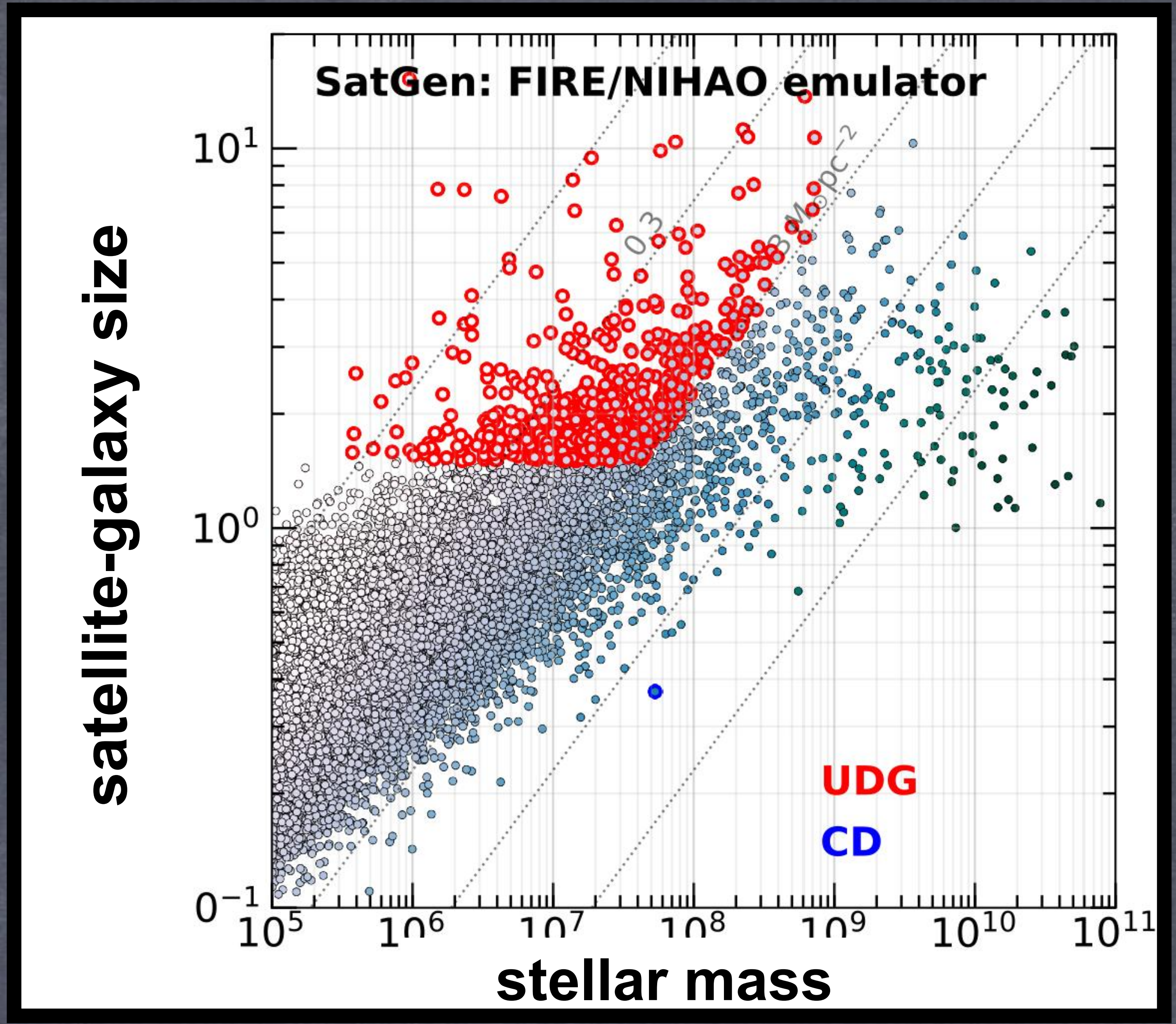
- A zoom-in simulation suite has ≈10 MWs or ≈1 cluster
- Adequate sampling of the cosmic halo-to-halo variance requires a sample of >100

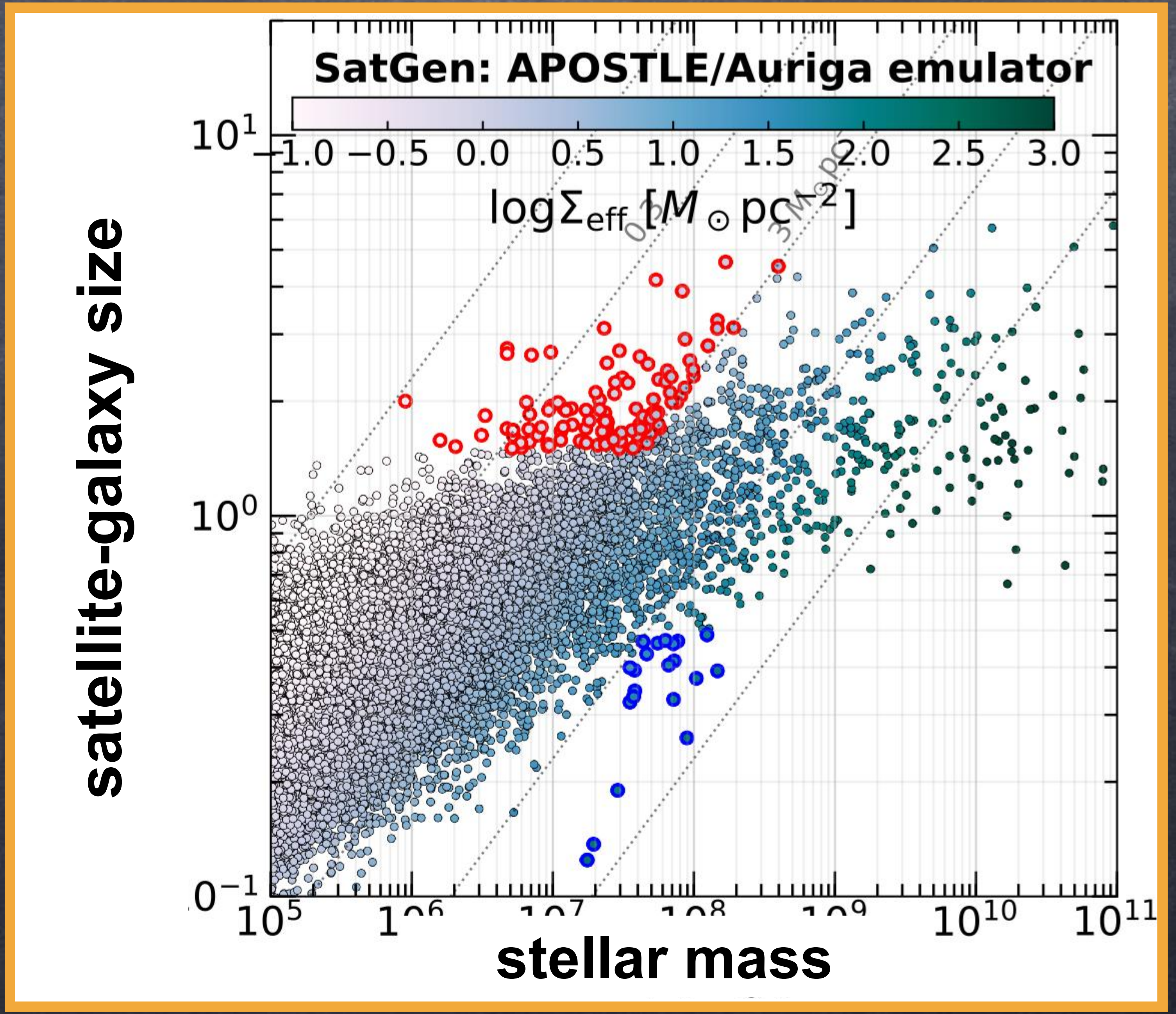
Environmental effects broadens structural diversity, but perhaps not enough …



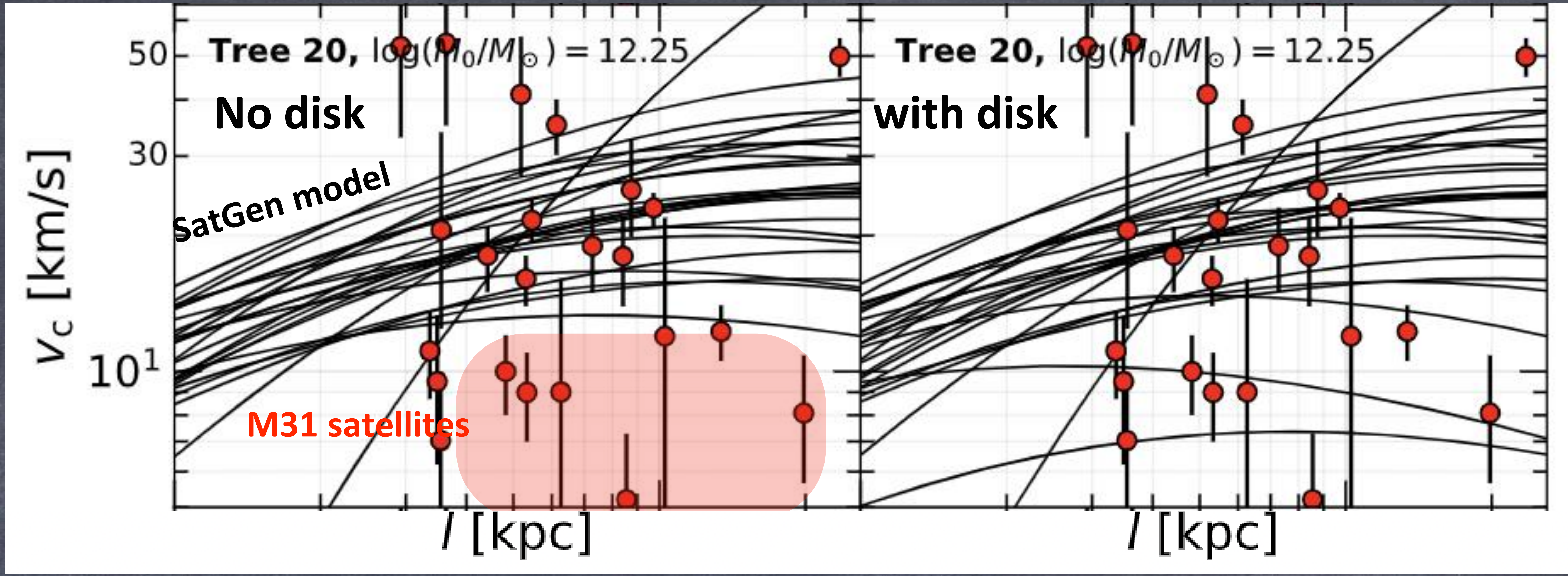
SatGen
Jiang+21

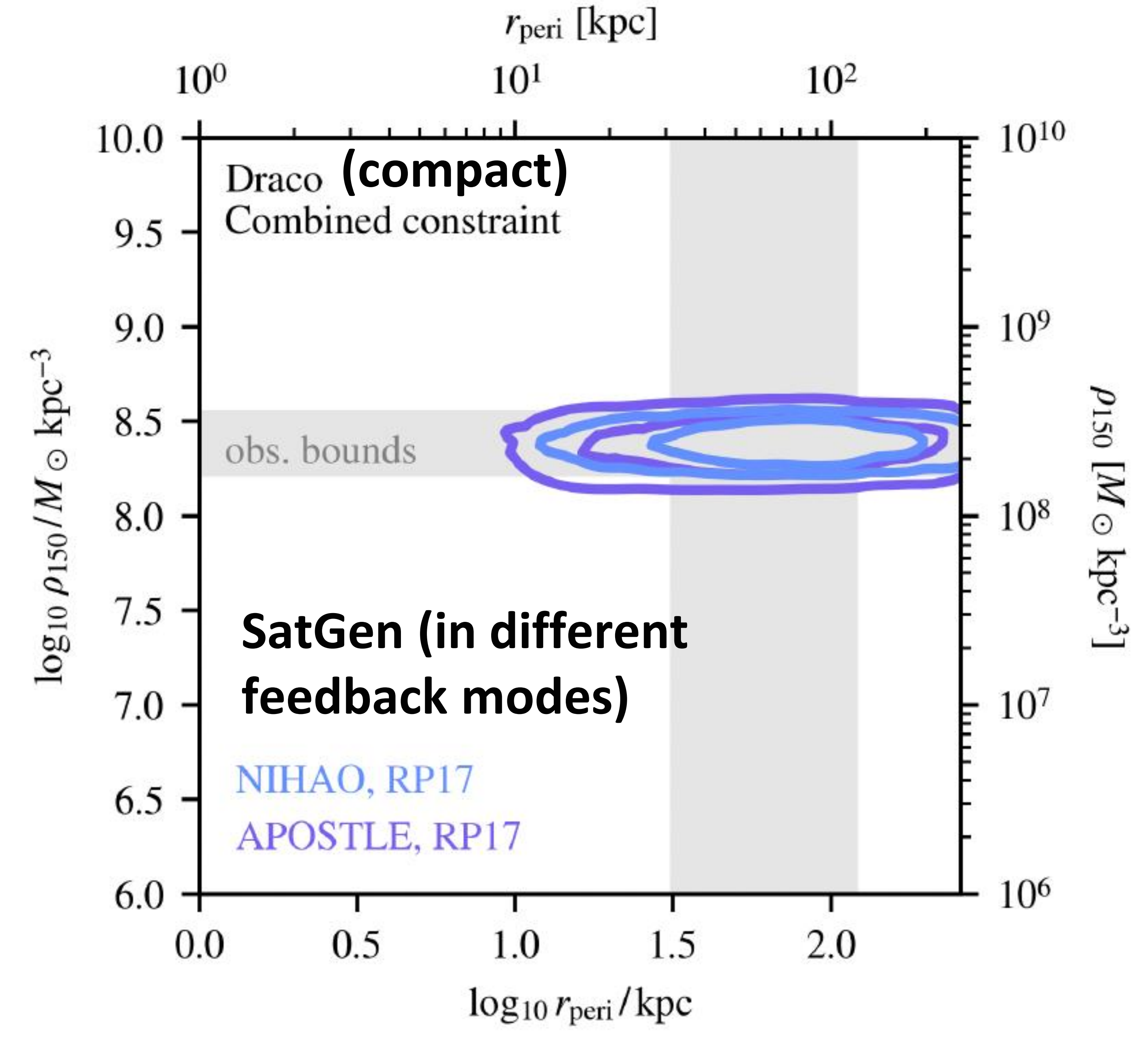
github.com/JiangFangzhou/SatGen

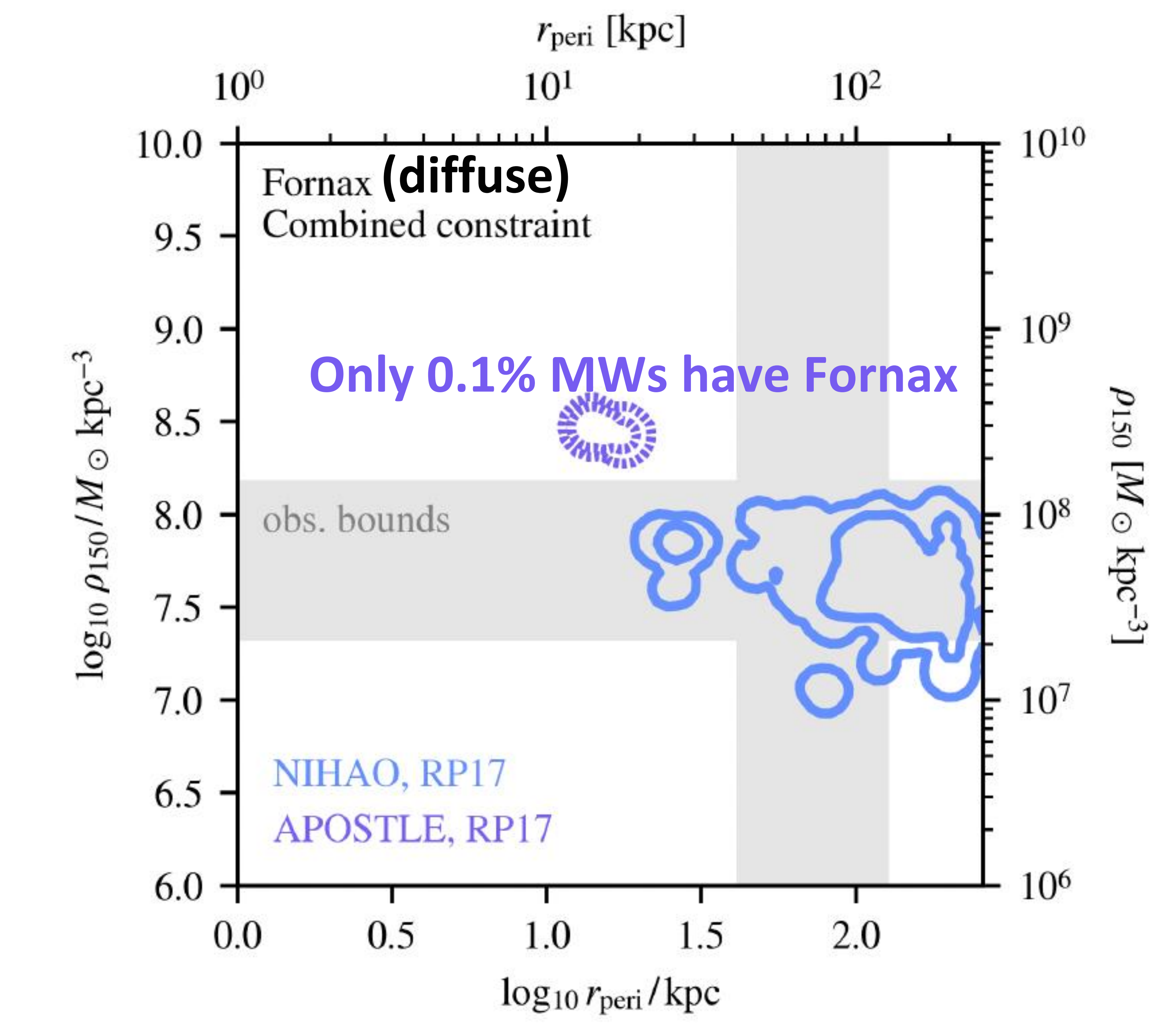


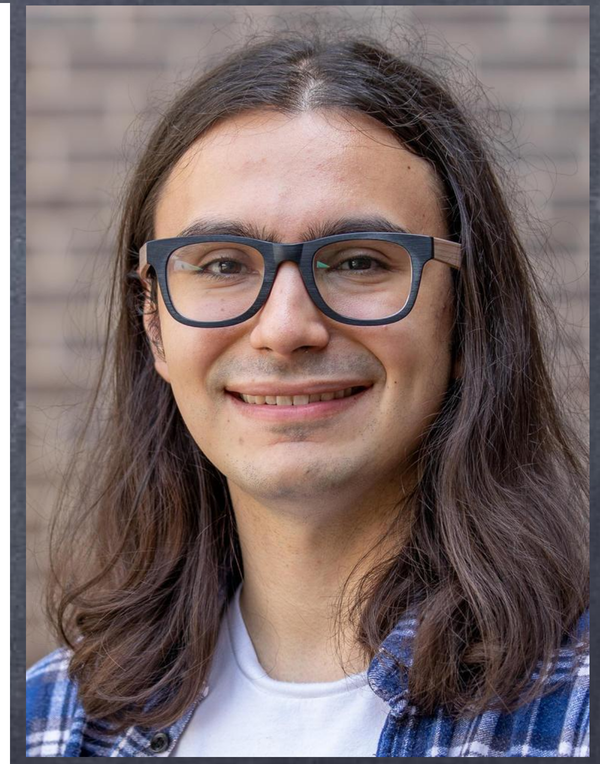


Environmental effects broadens structural diversity, but not enough …









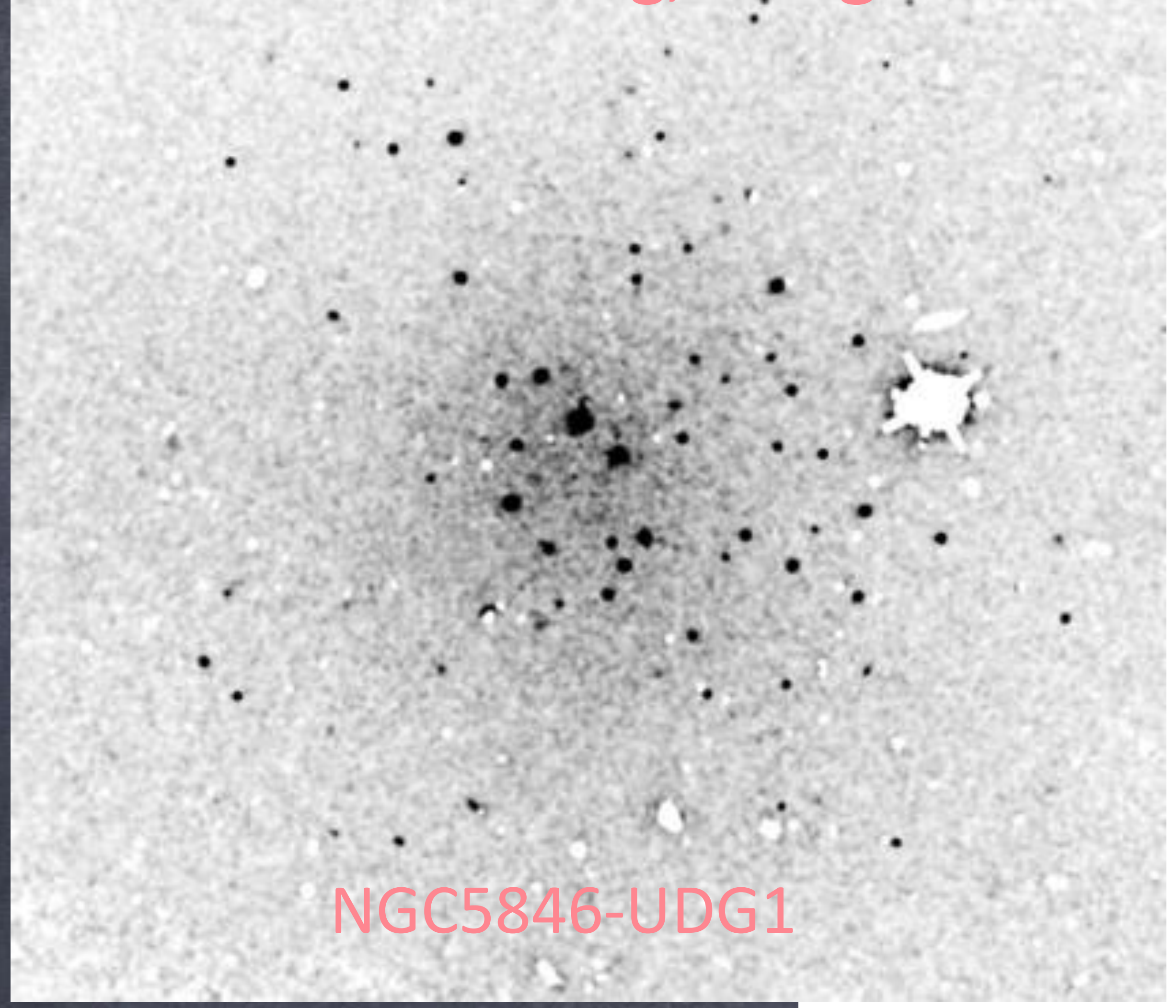
Dylan Folsom

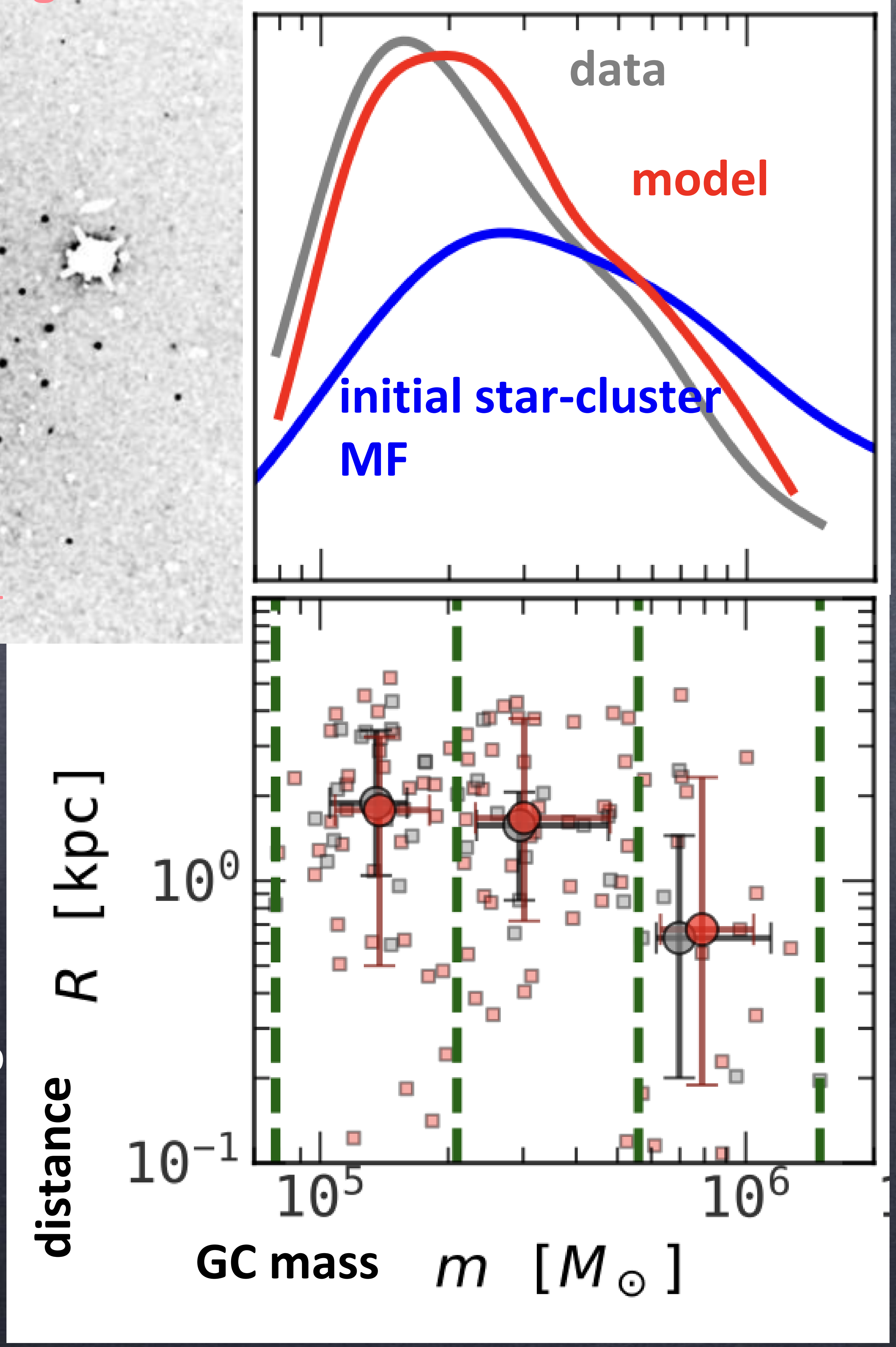
Folsom + (including Jiang) 2024, arXiv:2311.05676

Formation of nucleated dwarfs : dynamical friction + cuspy halo

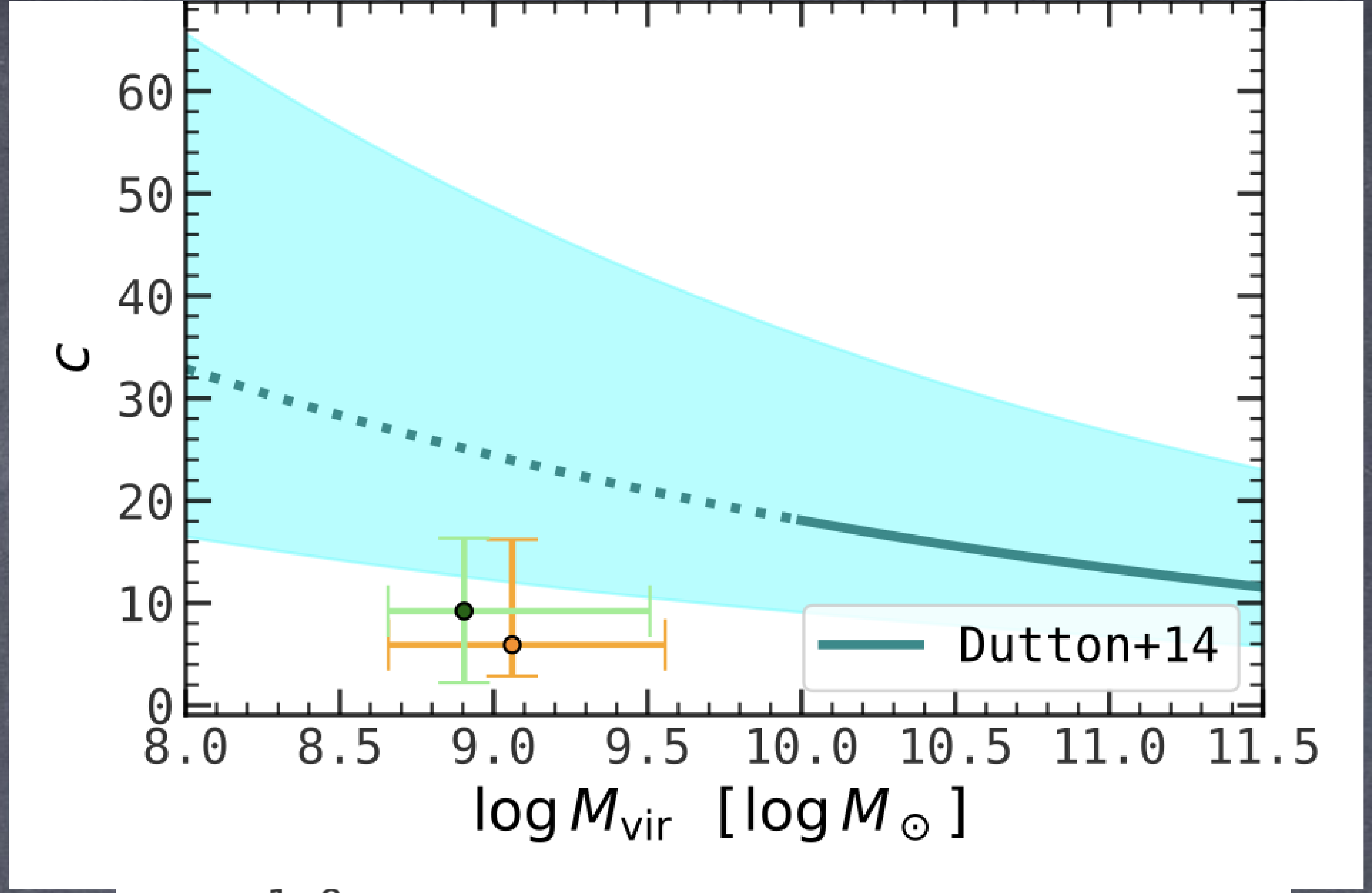
Jinning Liang

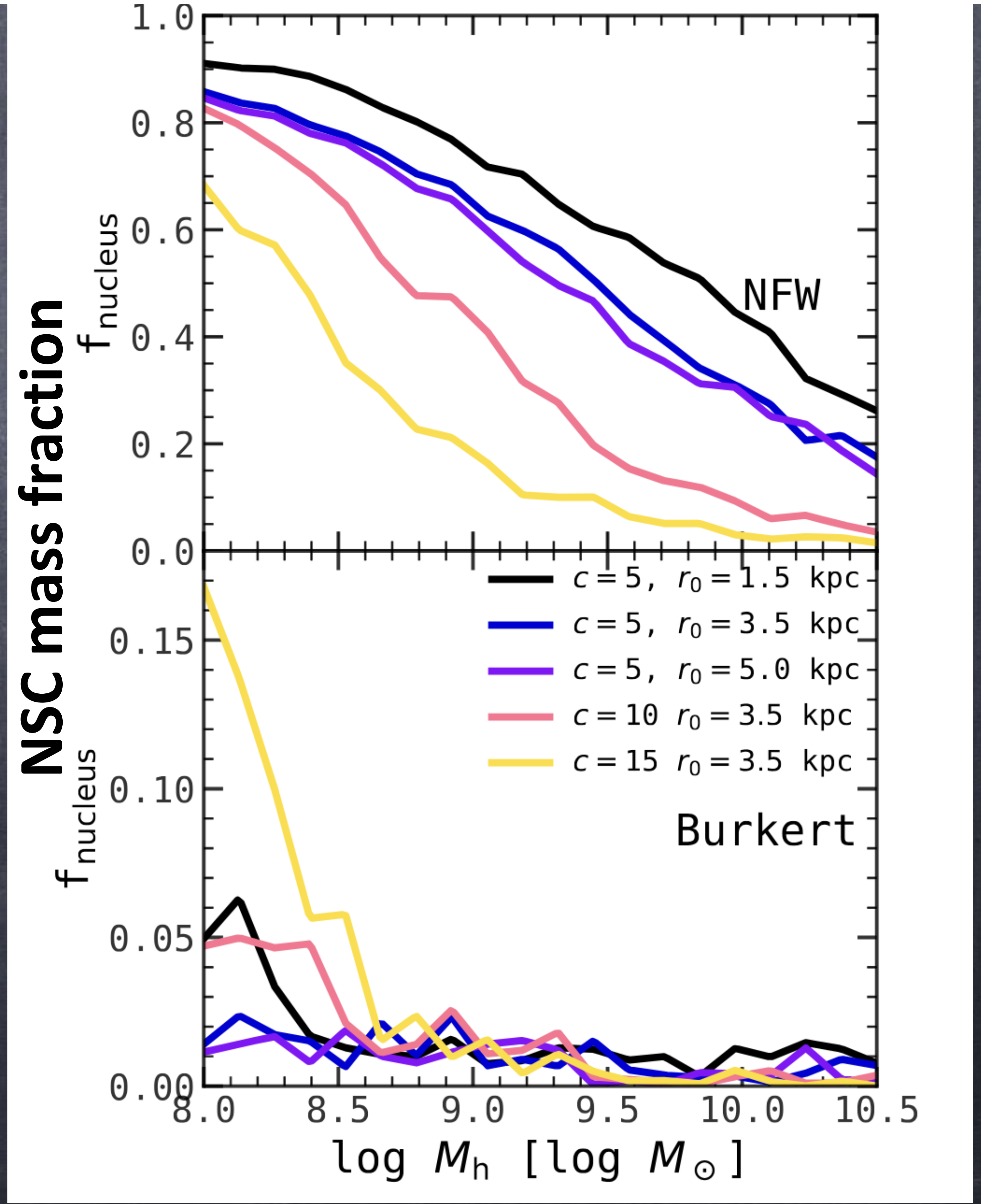
Liang, **Jiang*** + 2024 arXiv:2304.14431



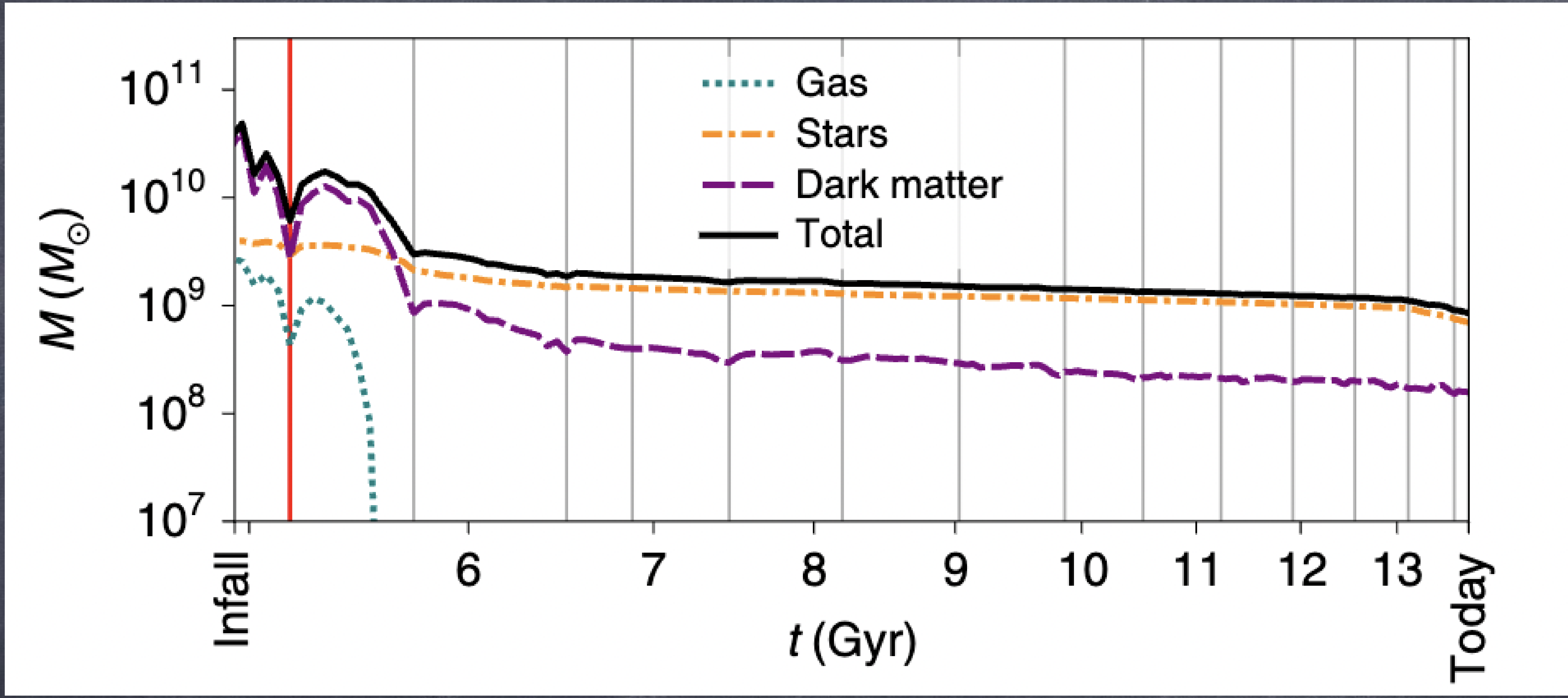


- Model of GC evolution + dynamical friction
- constrain DM halo (with imaging alone!)





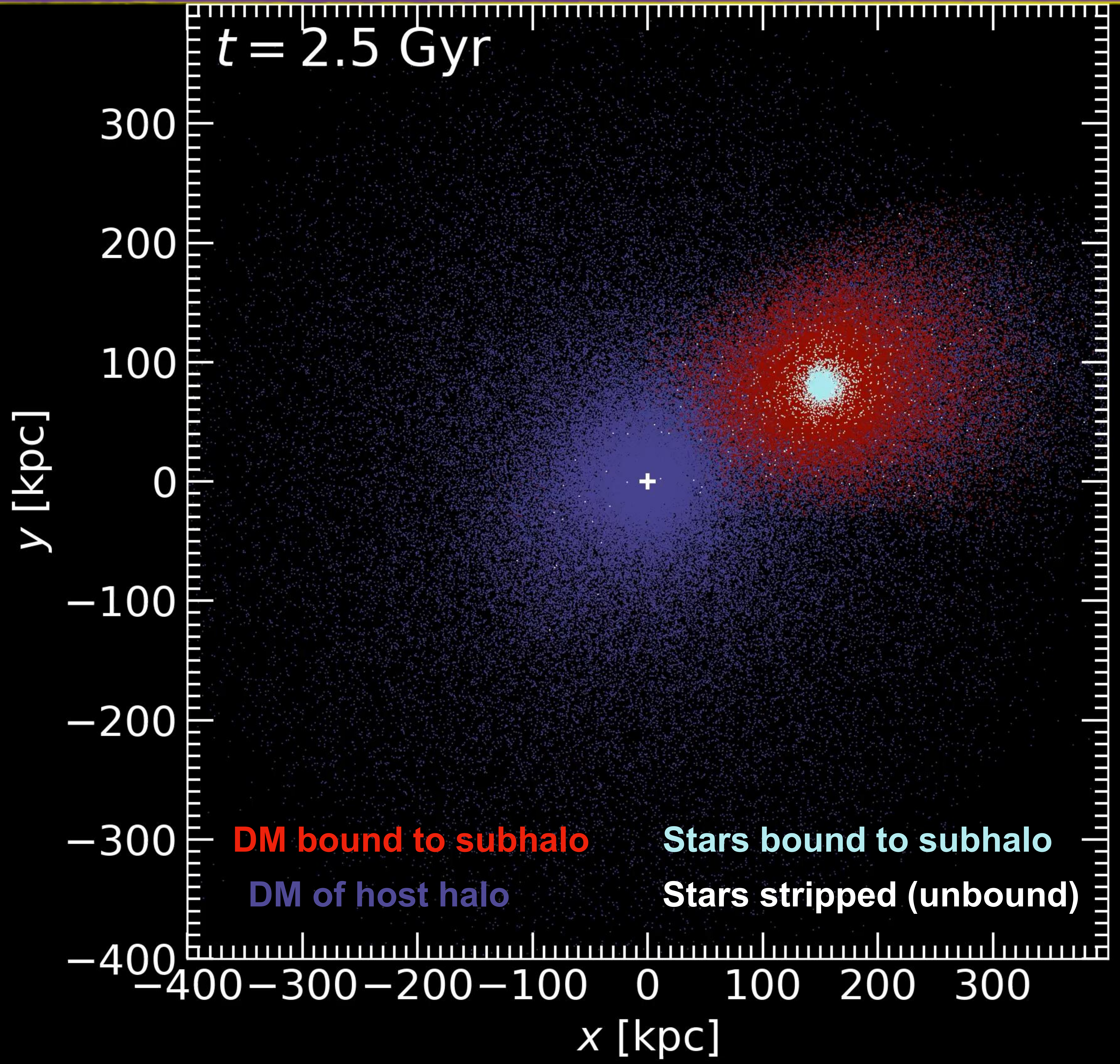
Formation of DM-deficient galaxies: differential stripping of DM and stars



Moreno + (including **Jiang**) 22 (FIRE-box simulation)

- caveat: numerical disruption of DM subhalos …

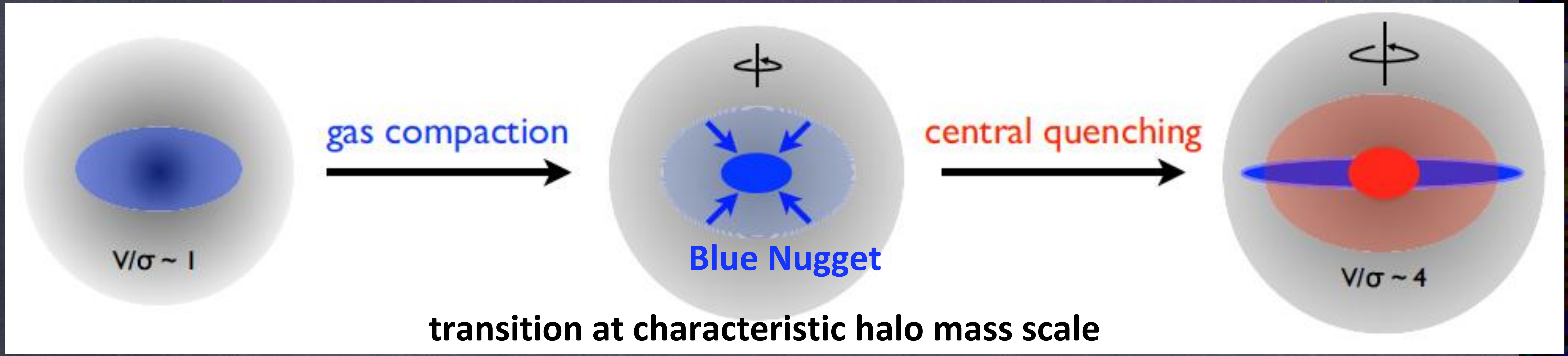
Differential Stripping → DM deficient dwarf?

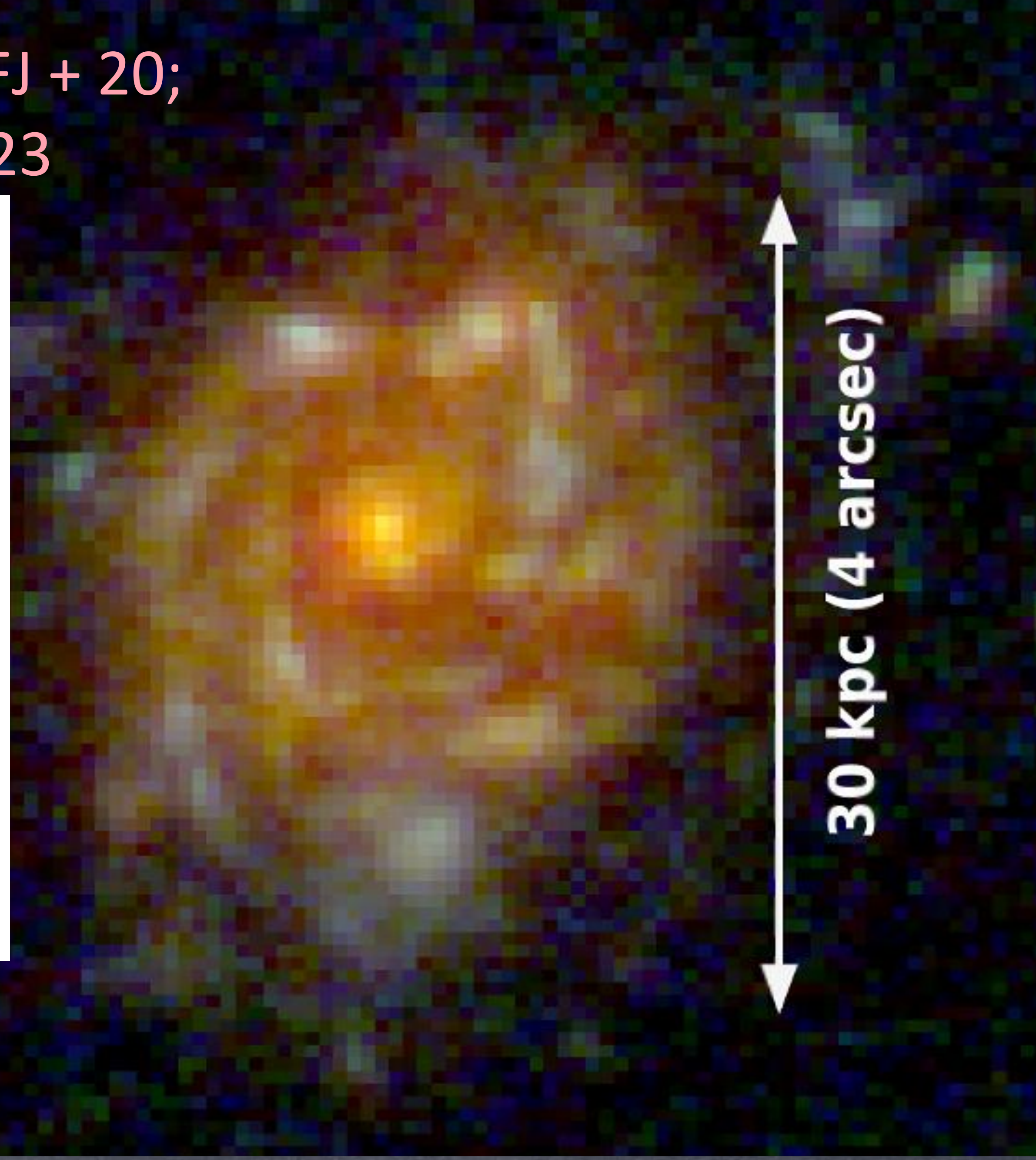


Fangzhou Jiang

DM-Halo conditions for disk formation/evolution

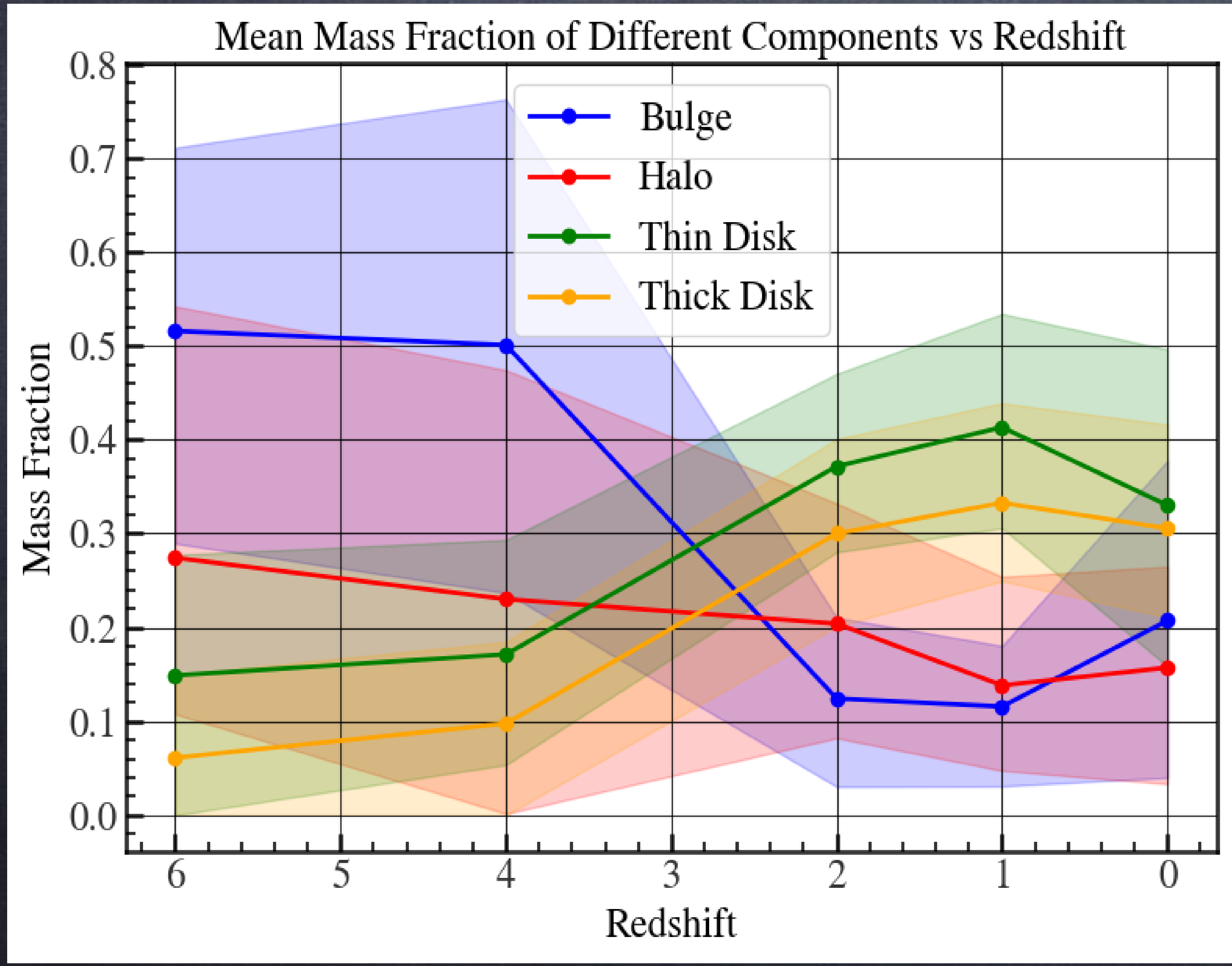
FJ+19; Dekel, Ginzburg, FJ + 20;
Hopkins + (including FJ) 23



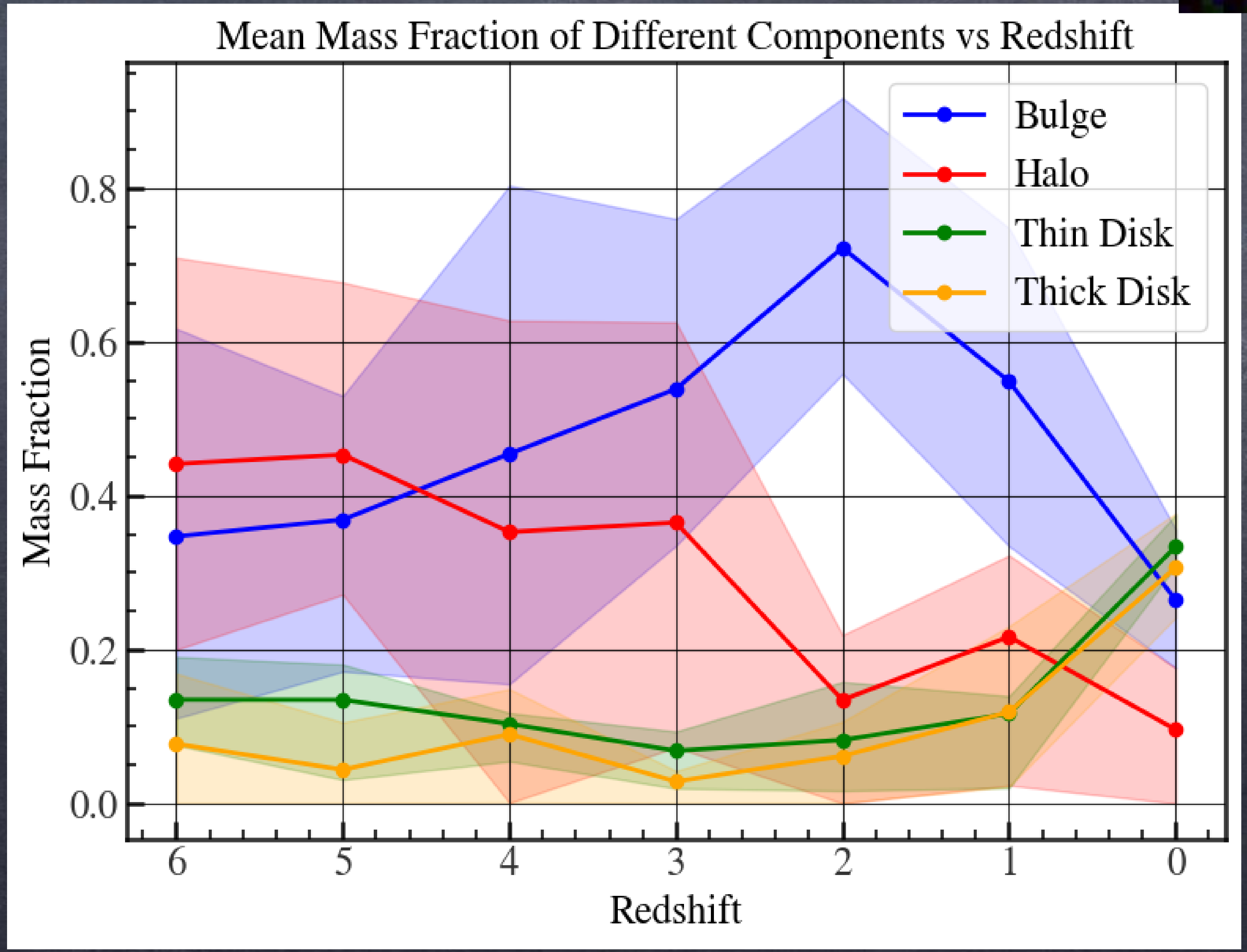


Disk formation in different simulations …

TNG50 MW-mass



FIRE2 m12's



Xing, FJ* + in prep

I. Can we solve-small scale puzzles within ΛCDM?

- For isolated dwarfs, puffy galaxies (cored halos) are over-produced, compact (cuspy) ones missing

  Jiang+ 19a, Relatores+19, Sales+21 review

- For satellites, environmental processes amplify subhalo response, giving rise to some diversity, but diversity is still limited

  Jiang & van den Bosch 2015; **Jiang**+21 (“SatGen”), Folsom+24 (using SatGen)

- Nuclear star clusters (NSCs) imply cuspy halos but ultra diffuse galaxies are expected to populate cored ones according to ΛCDM simulations…

  Liang, **Jiang***+24 arXiv:2304.14431

- Differential stripping of DM and stars can produce DM-poor satellites, but there may be numerical stripping

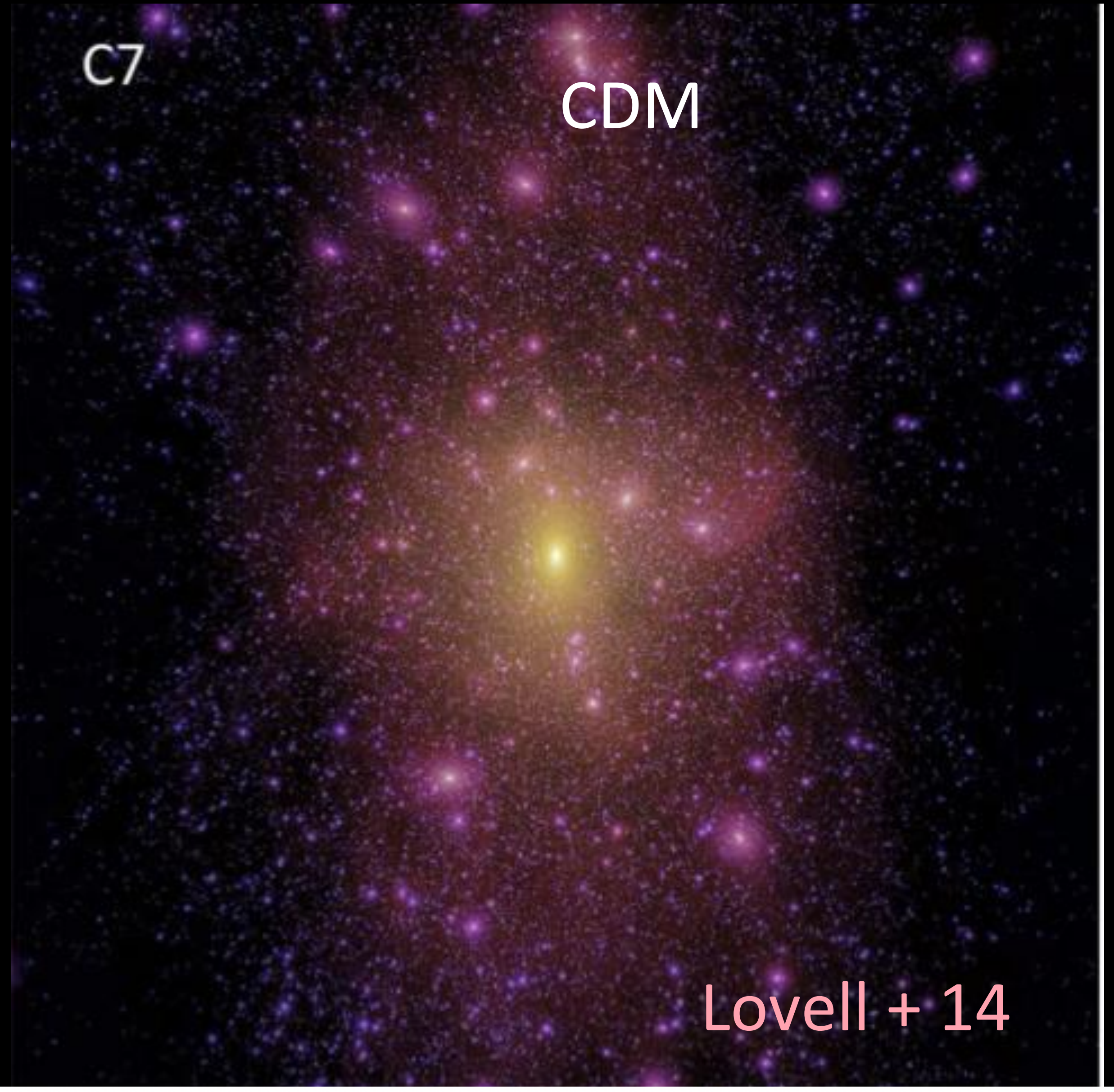
  Moreno + (including **Jiang**) 22

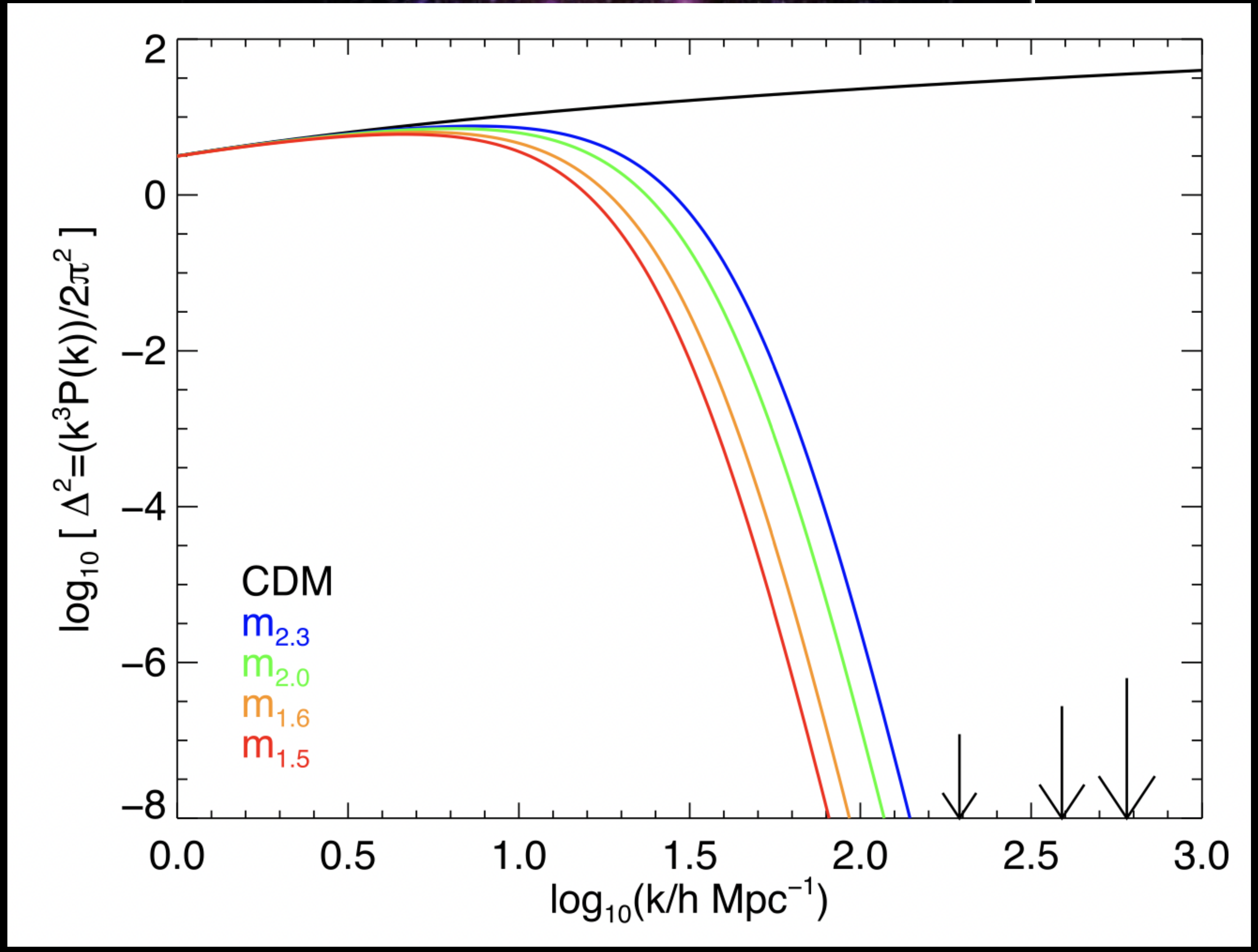
- ΛCDM simulations still struggle in reproducing galaxy morphological evolution (esp disks) is still problematic

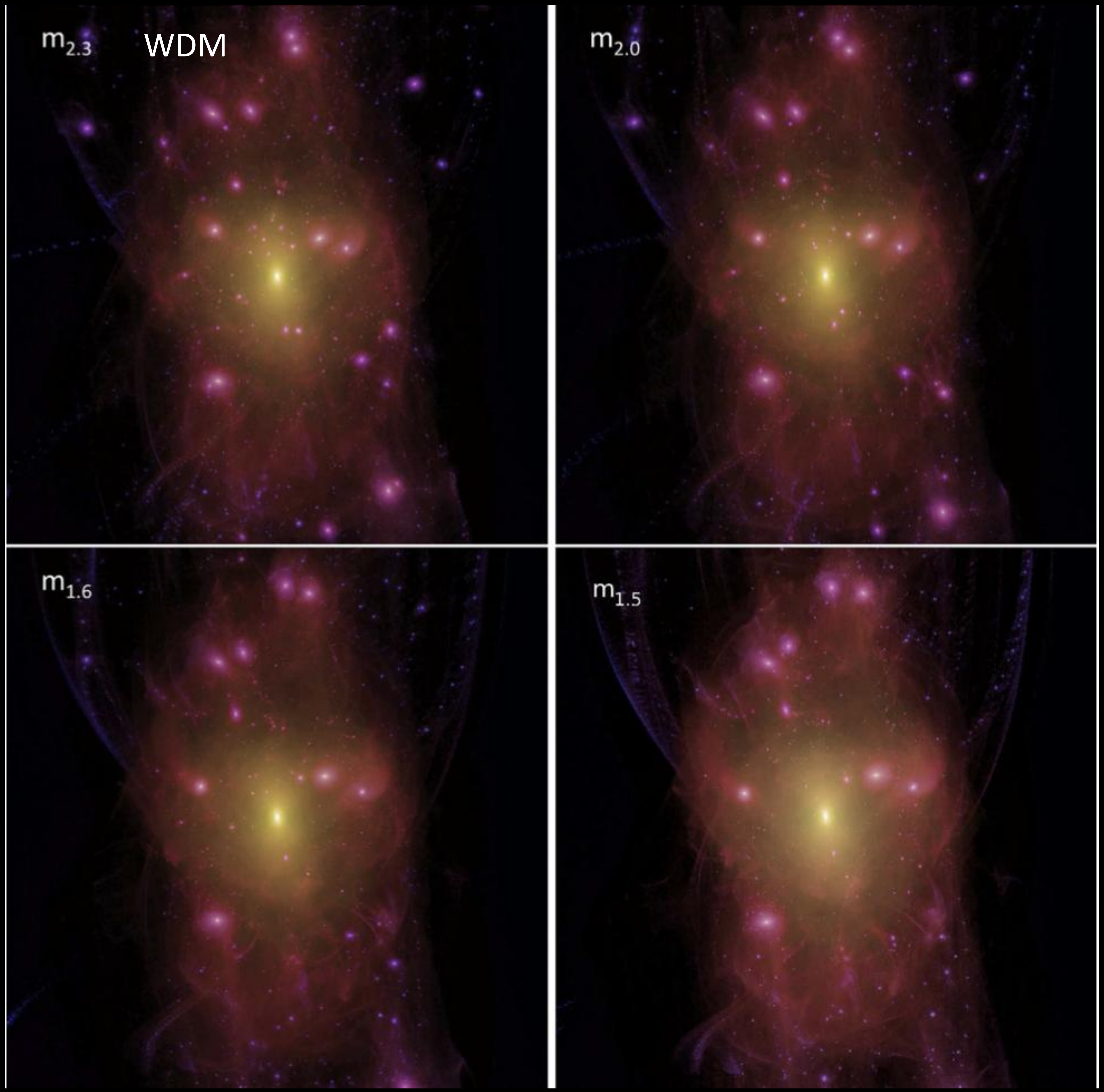
  Jiang+25 2504.01070 **;** Xing, **Jiang** + in prep

II. Alternative Dark Matter models (Warm, Fuzzy, Self-interacting)

标准宇宙学面临的小尺度结构问题与挑战——暗物质解决方案·温暗物质







FDM: Power Spectrum

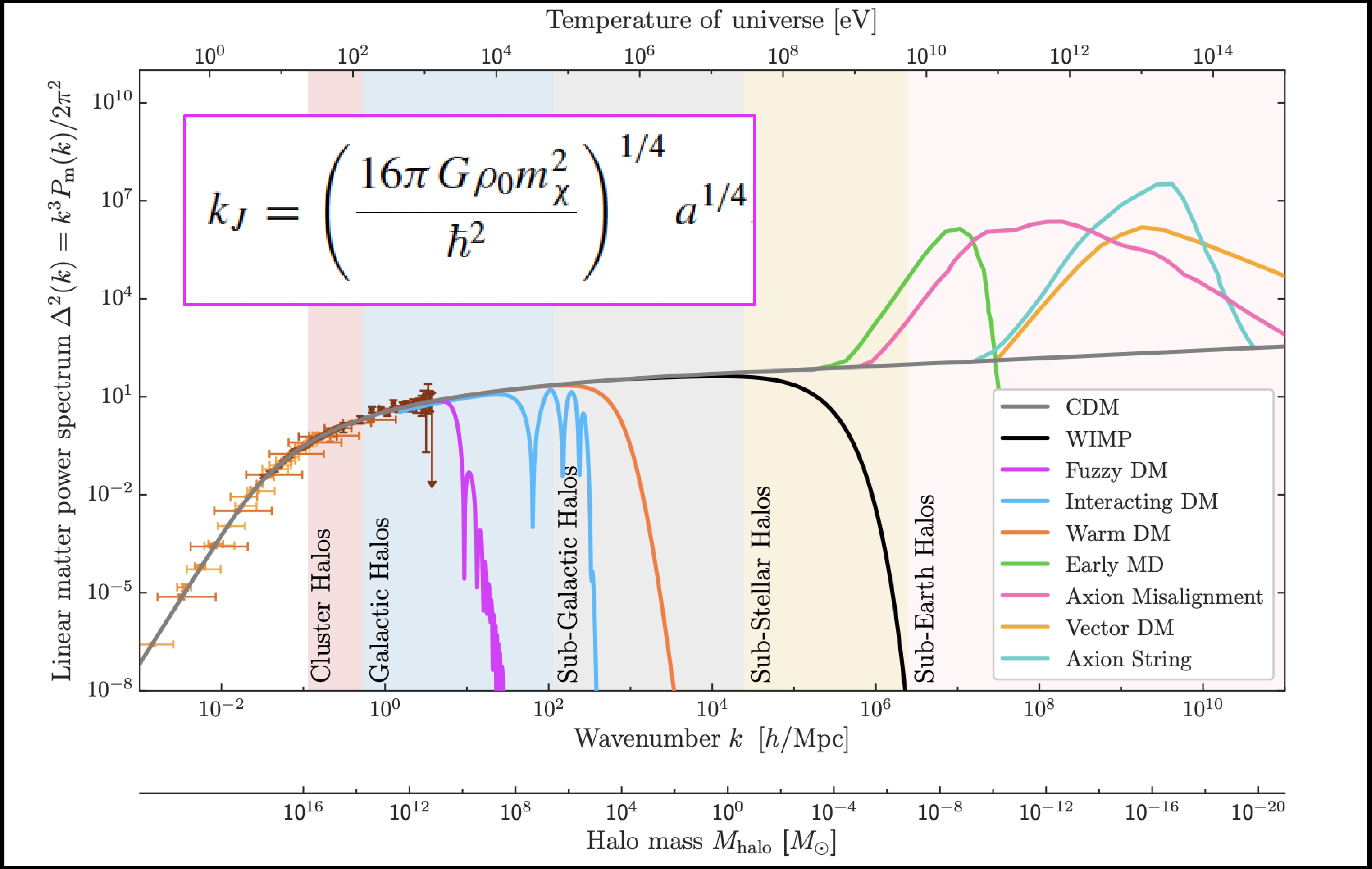
Bechtol+22

- Particle mass ~10^{-22} eV
- The Equation of Motion for FDM is a Schrodinger-Poisson equation:

$$i\hbar\left(\dot{\psi}+\frac{3}{2}H\psi\right)=\left(-\frac{\hbar^2}{2m_\chi a^2}\nabla^2+m_\chi\Phi\right)\psi,$$

- The linear growth equation has an extra term of quantum pressure

$$\ddot{\delta}_k+2H\dot{\delta}_k+\left(\frac{\hbar^2k^4}{4m_\chi^2a^4}-\frac{4\pi G\rho_0}{a^3}\right)\delta_k=0,$$



- quantum Jeans scale beyond which gravity is balanced by quantum pressure

FDM: HMF

Elgamal +24

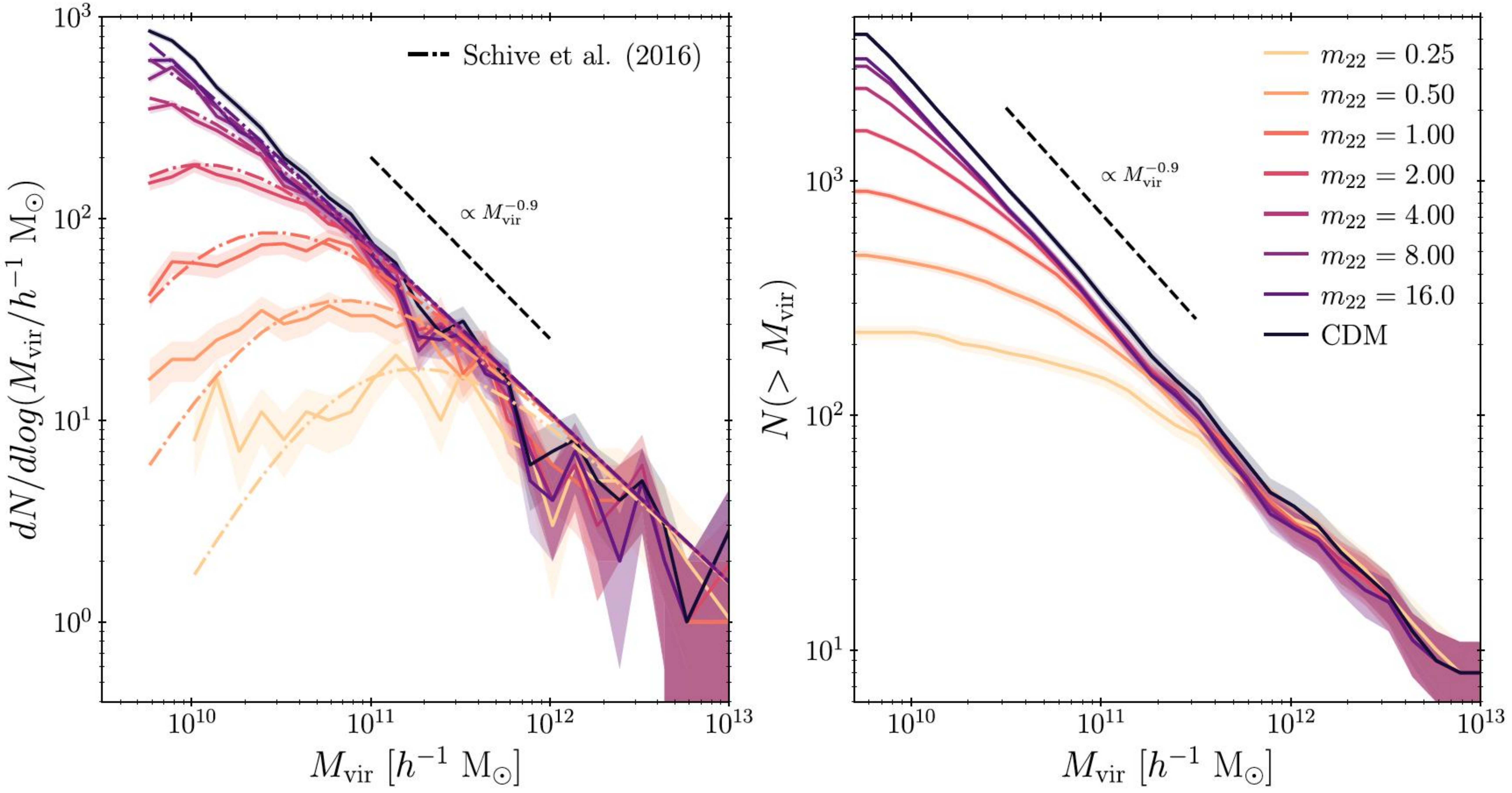
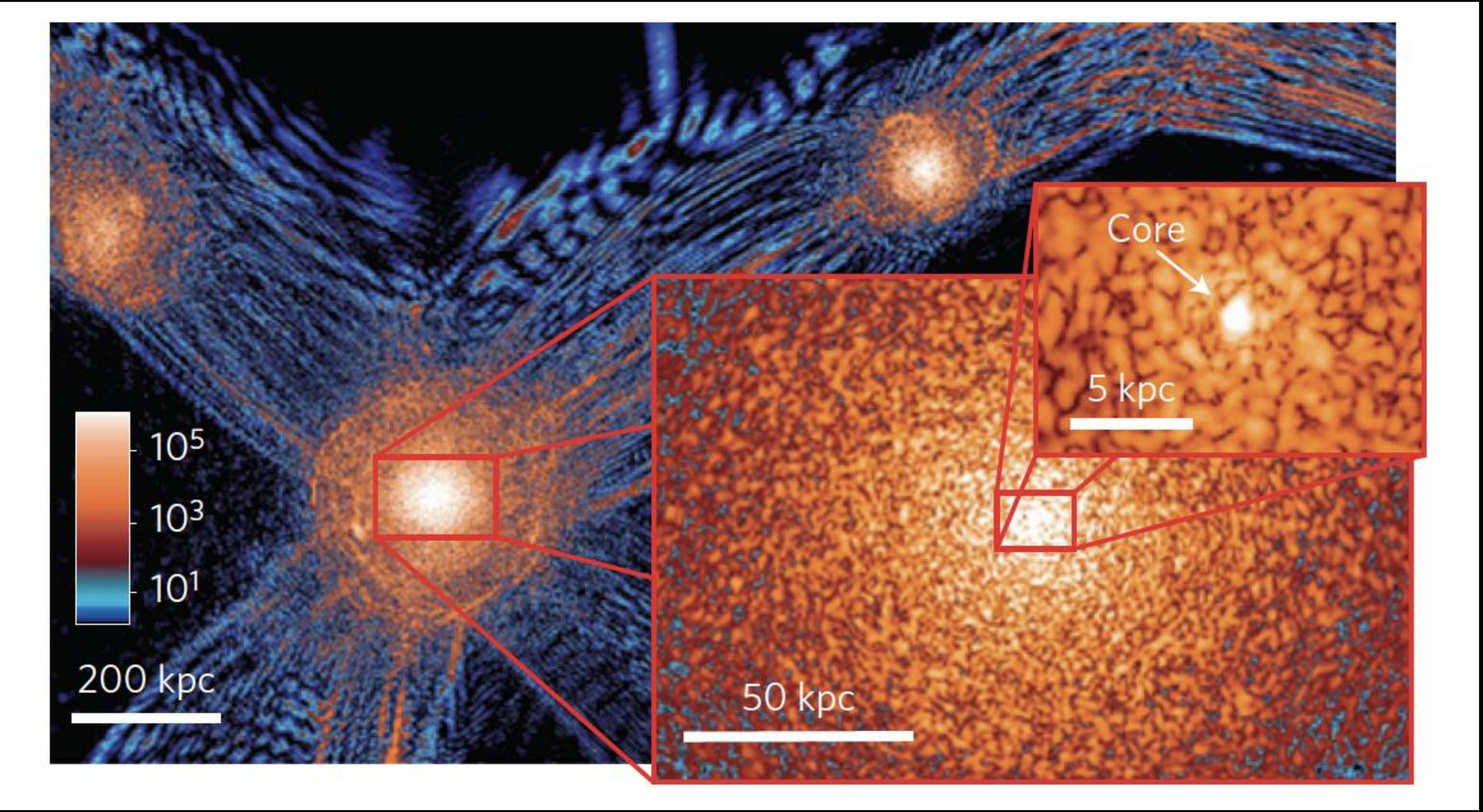


Figure 2. The CDM and FDM differential (left) and cumulative (right) HMF for different FDM masses. The dashed black line is the HMF scaling obtained in Springel et al. (2008) for CDM. The HMF given by Schive et al. (2016) fitting formula is shown by the dash–dotted curves on the left plot. The shaded regions correspond to the $\pm 1\sigma$ (i.e. standard deviation) of the number count distribution.

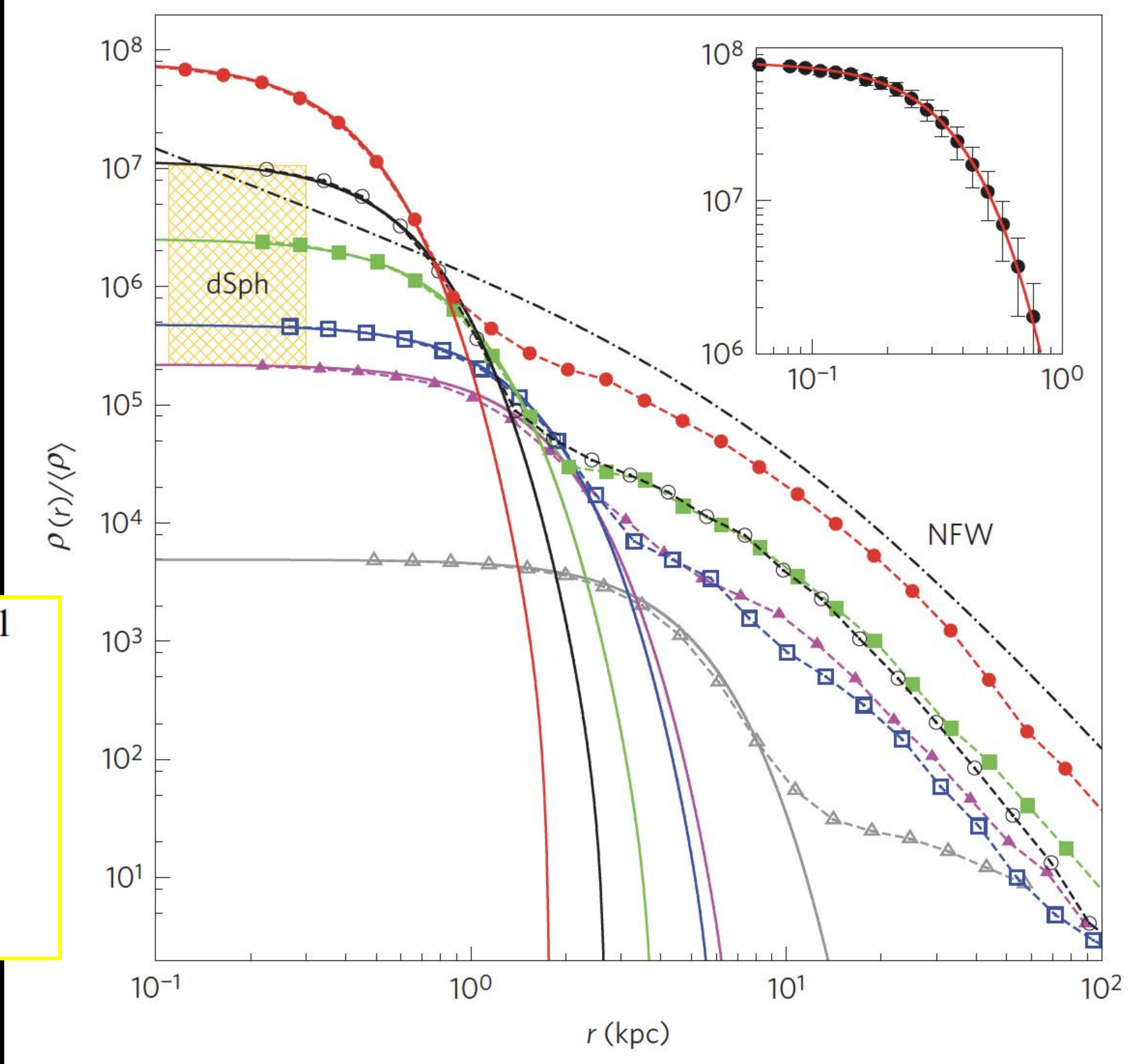
FDM: Soliton Core

Schive+14



FDM: Soliton Core radius – halo mass relation

Schive+14



$$r_0 = 1.6 \times q(z) \times (1+z)^{-0.5} \left(\frac{m}{10^{-22}\mathrm{eV}}\right)^{-1} \times \left(\frac{M_{\rm vir}}{10^9 M_\odot}\right)^{-1/3} \mathrm{kpc}$$

FDM: Soliton Core – in tension with dwarf galaxy densities

Burkert18

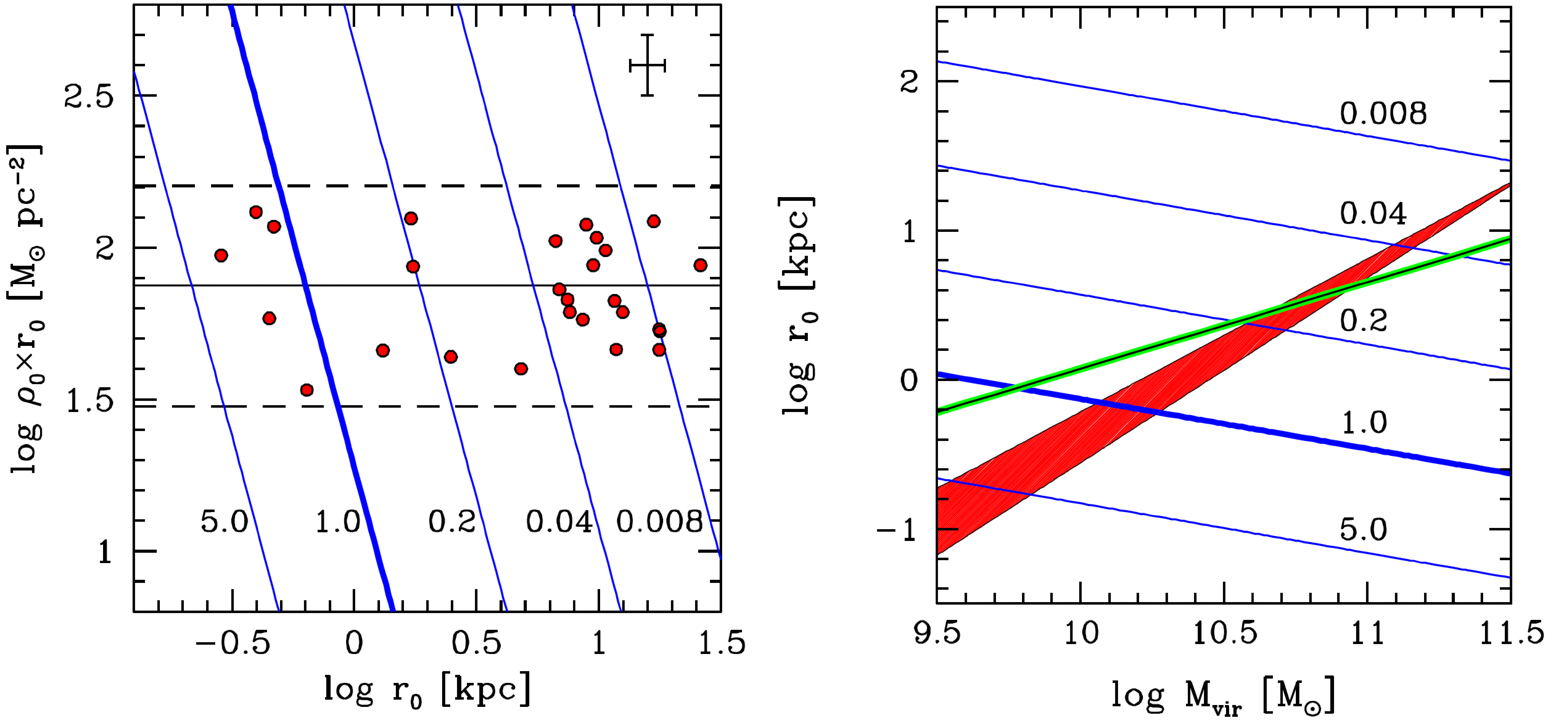
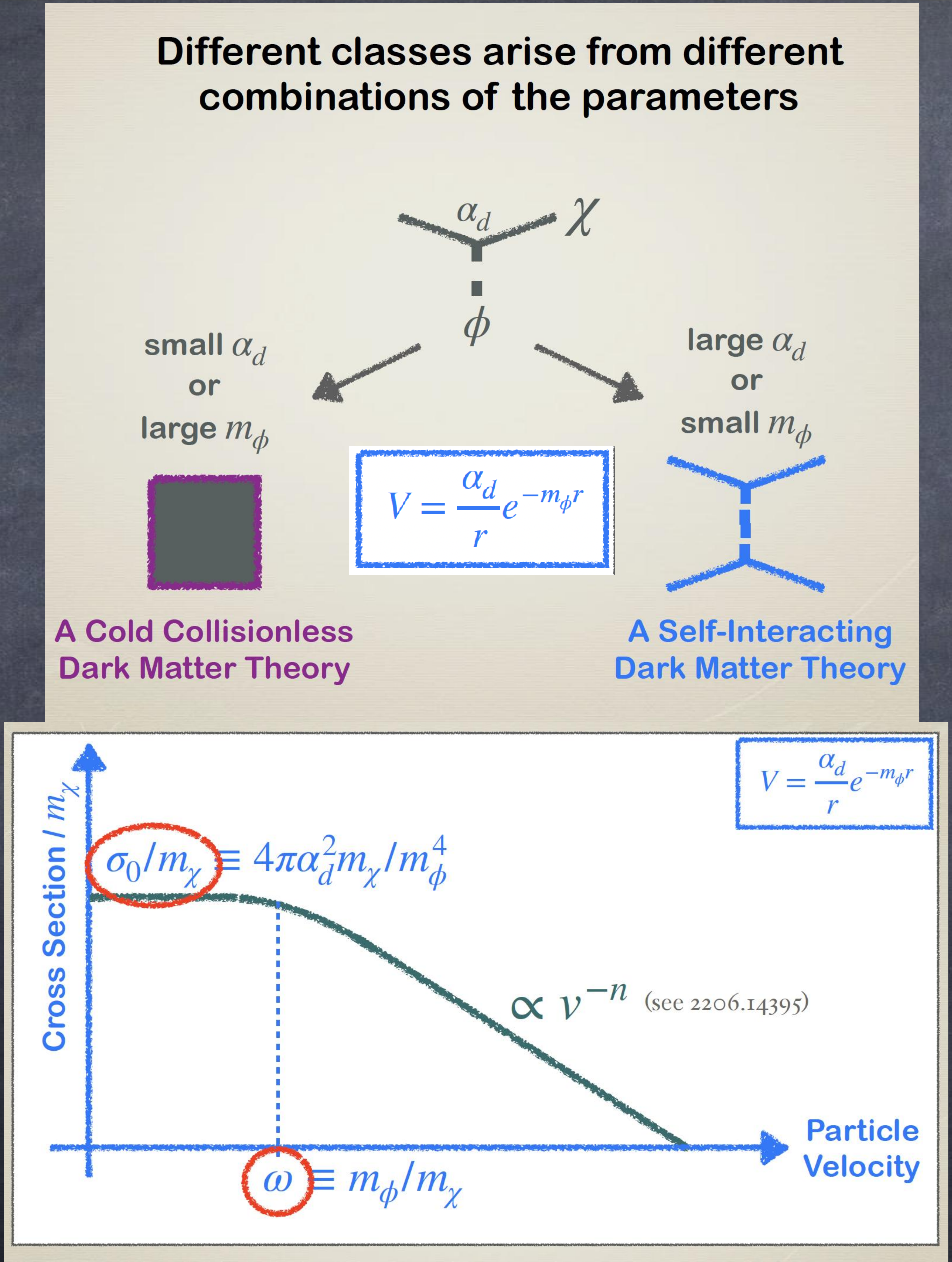


Figure 1. Observationally inferred dark halo core properties are compared with the predictions of the FDM model. The red points in the left panel show the observed core column densities vs. the corresponding core radii. The black solid horizontal line depicts the average value of 75 $M_{\odot}$ pc^{-2} and the horizontal dashed lines show the observed scatter. The error bar in the upper right corner shows the typical uncertainties. Blue lines denote the predictions for soliton cores. The FDM cores follow a completely different trend compared to the observations. The numbers associated with each blue line in the left and right panel depict the corresponding particle mass m in units of the standard value of 10^{-22} eV. The thick solid blue line corresponds to this standard value. The right panel shows the correlation between core radius and dark halo virial mass. The red area depicts the observationally inferred correlation, adopting a stellar-to-dark-matter mass conversion as predicted from abundance matching, combined with a constant stellar mass to light ratio and assuming a halo formation time in the range $z = 0$ (lower boundary) to $z = 8$ (upper boundary). Galaxies in this redshift range should populate this area. The green line shows the prediction from a universal rotation curve analysis of disk galaxies by Salucci et al. (2007). The redshift zero FDM predictions (blue lines) follow a completely different trend, with r_0 decreasing with increasing M_{vir}.

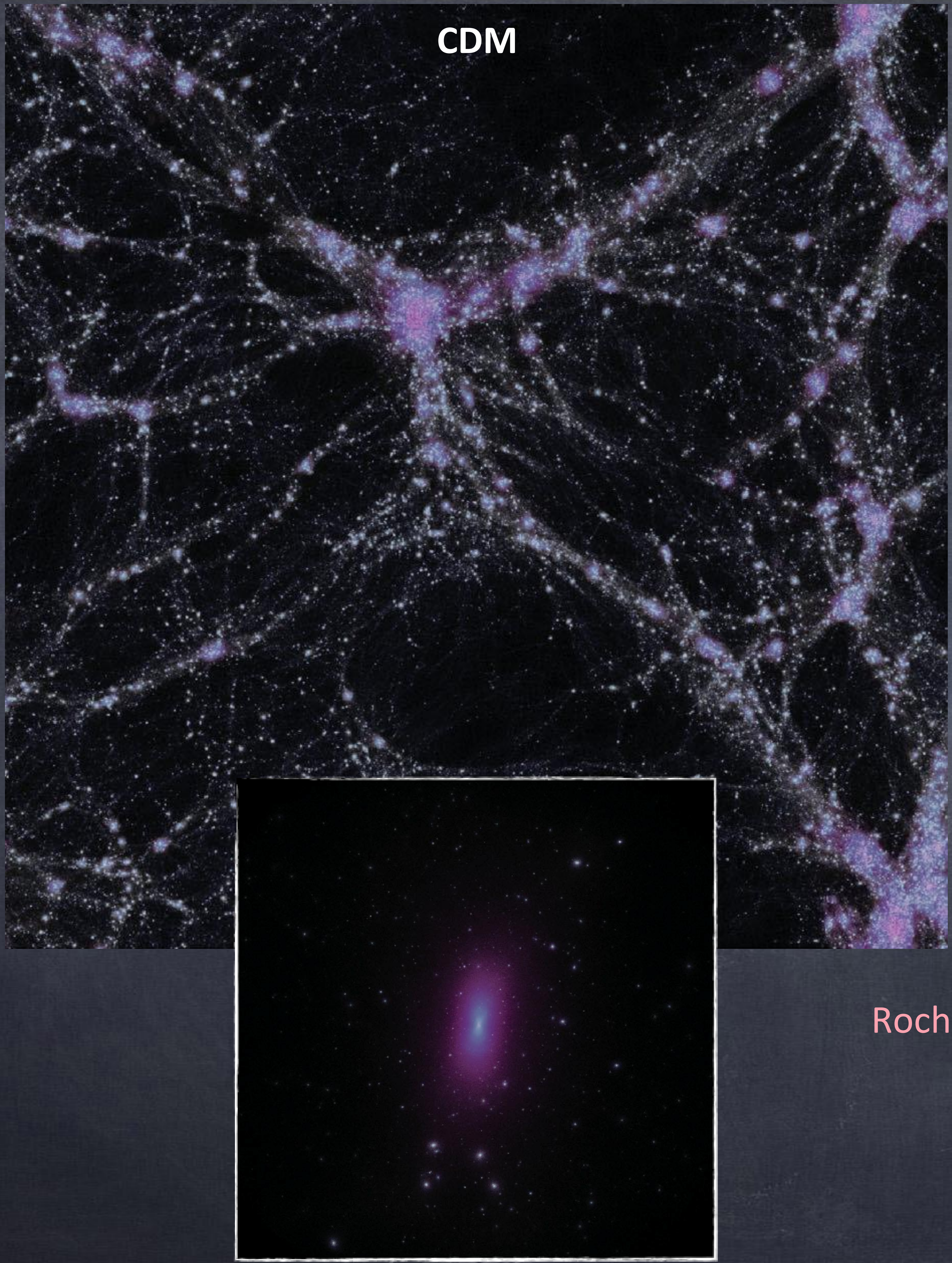
SIDM vs CDM: particle physics

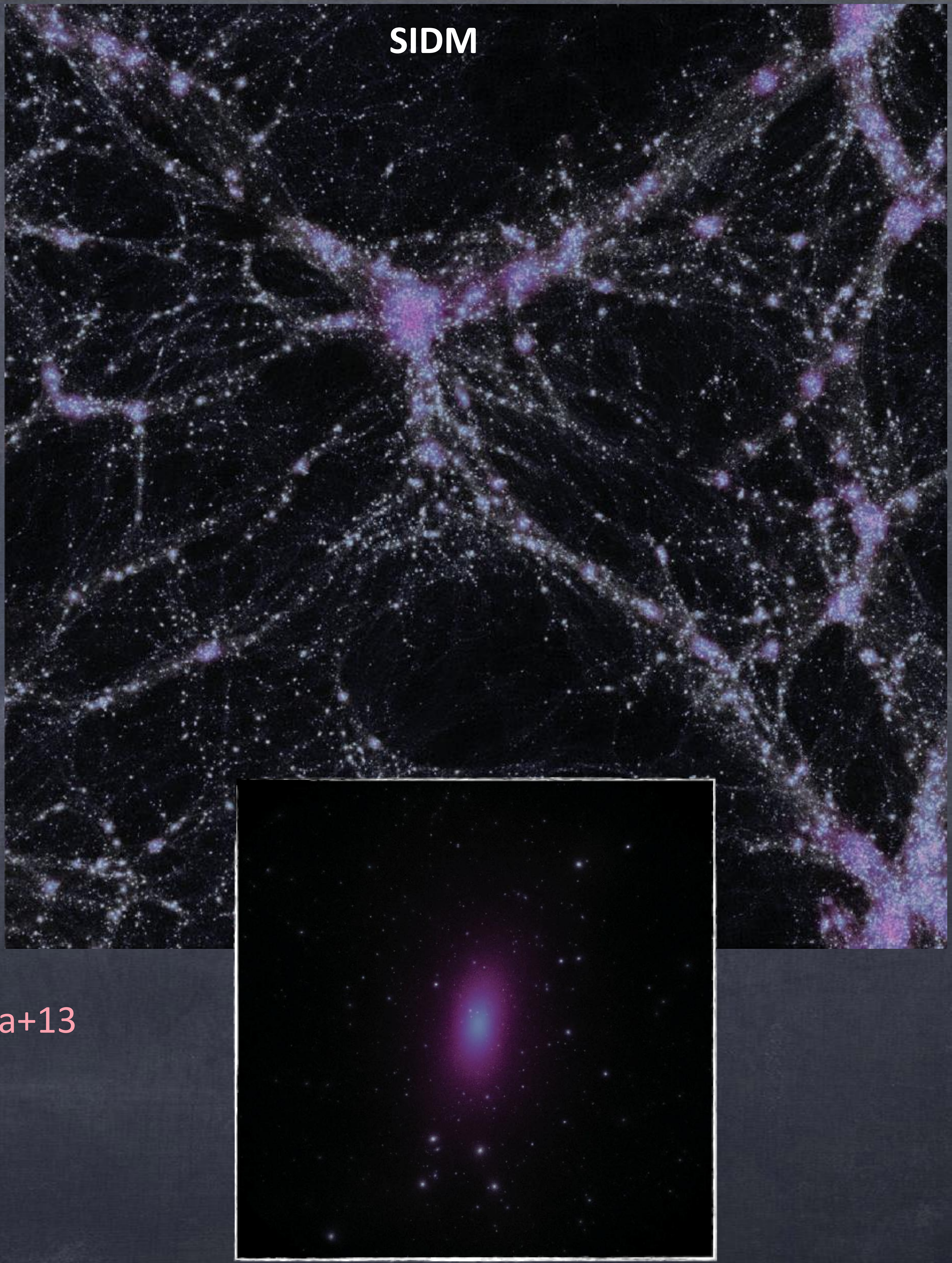


- SIDM and CDM are just two extreme cases of self-interactions obeying a Yukawa-like potential
- cross-section is velocity-dependent
- simulations & models usually consider effective constant cross-section for given halo mass scale of interest

Cartoon by Oren Slone

SIDM: same as CDM on large scale, only modifies halo center





Rocha+13

SIDM effects — isothermal core and core-collapse

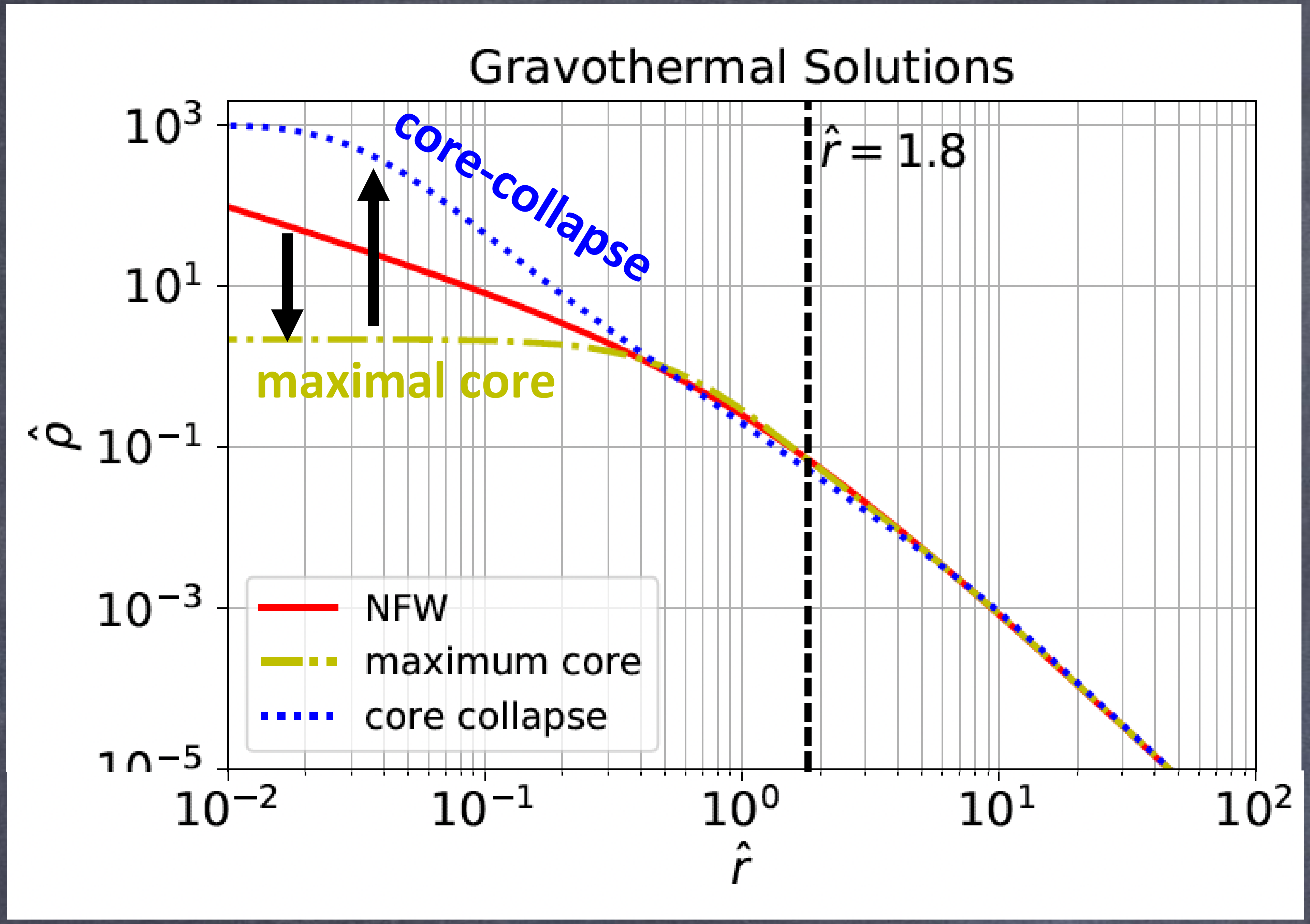
Yang S.*, **Jiang*** + 24 arXiv: 2305.05067

- Self-interaction thermalizes halo center
- Isothermal core hotter than outskirt -> self-gravitating systems have negative heat capacity -> core collapse
- core collapse -> 1% halo mass turns into seed BH, independent of halo mass or halo concentration

See Feng+21 and Feng+22 for detailed relativistic calculations

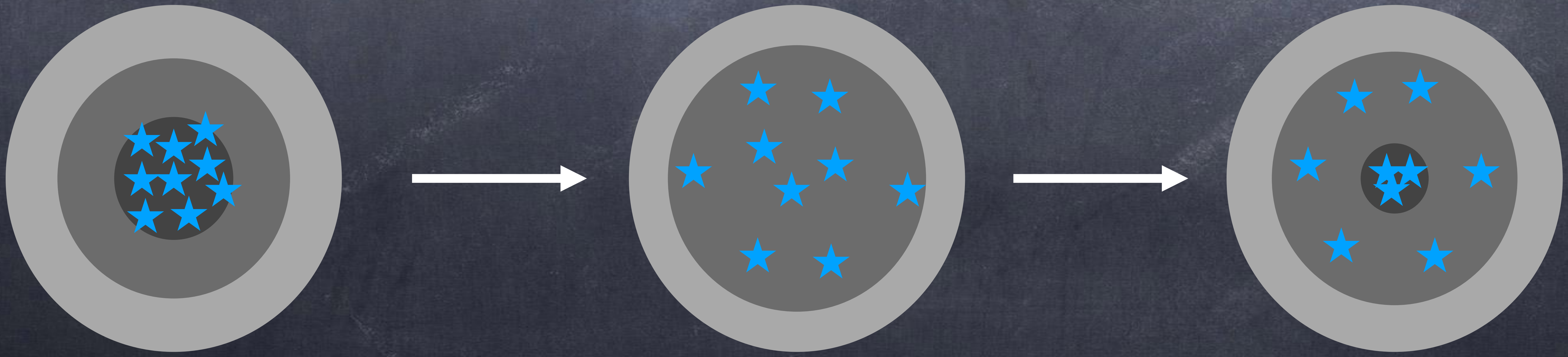
- If core-collapse occurs before re-ionization, "naked" BHs in halos without host galaxy can form -> LRDs ??

Structural diversity of dwarf galaxies can be explained by different evolution stages of SIDM halo



FJ+23 -- public code for computing SIDM profile:
https://github.com/JiangFangzhou/SIDM

Yang S.*, **Jiang***, Benson + 24 arXiv: 2305.05067

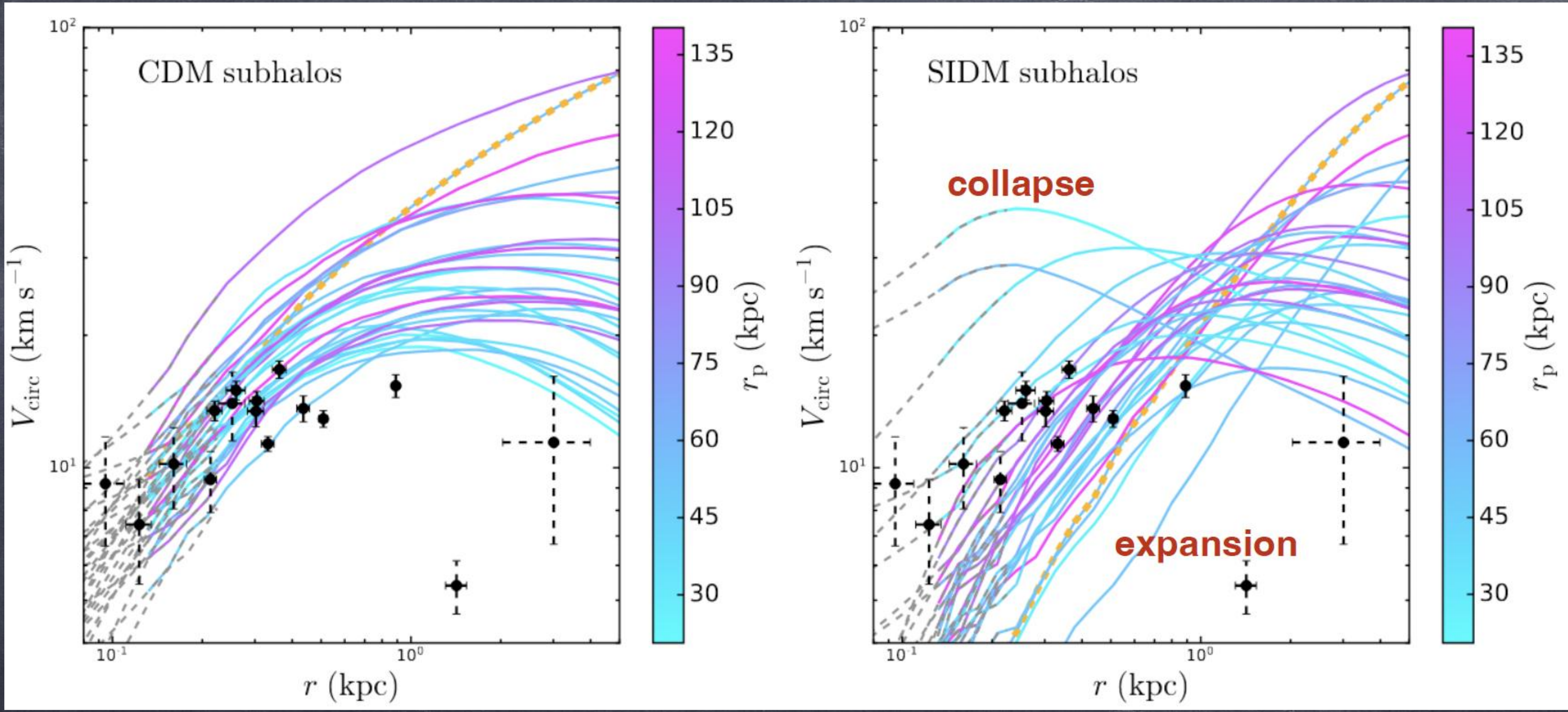


Normal dwarf in NFW halo

Ultra-diffuse galaxy in isothermal SIDM core

SIDM core-collapse -> nuclear star clusters & BH

SIDM can produce sufficient satellite structural diversity



Yang D., Nadler, Yu + 22

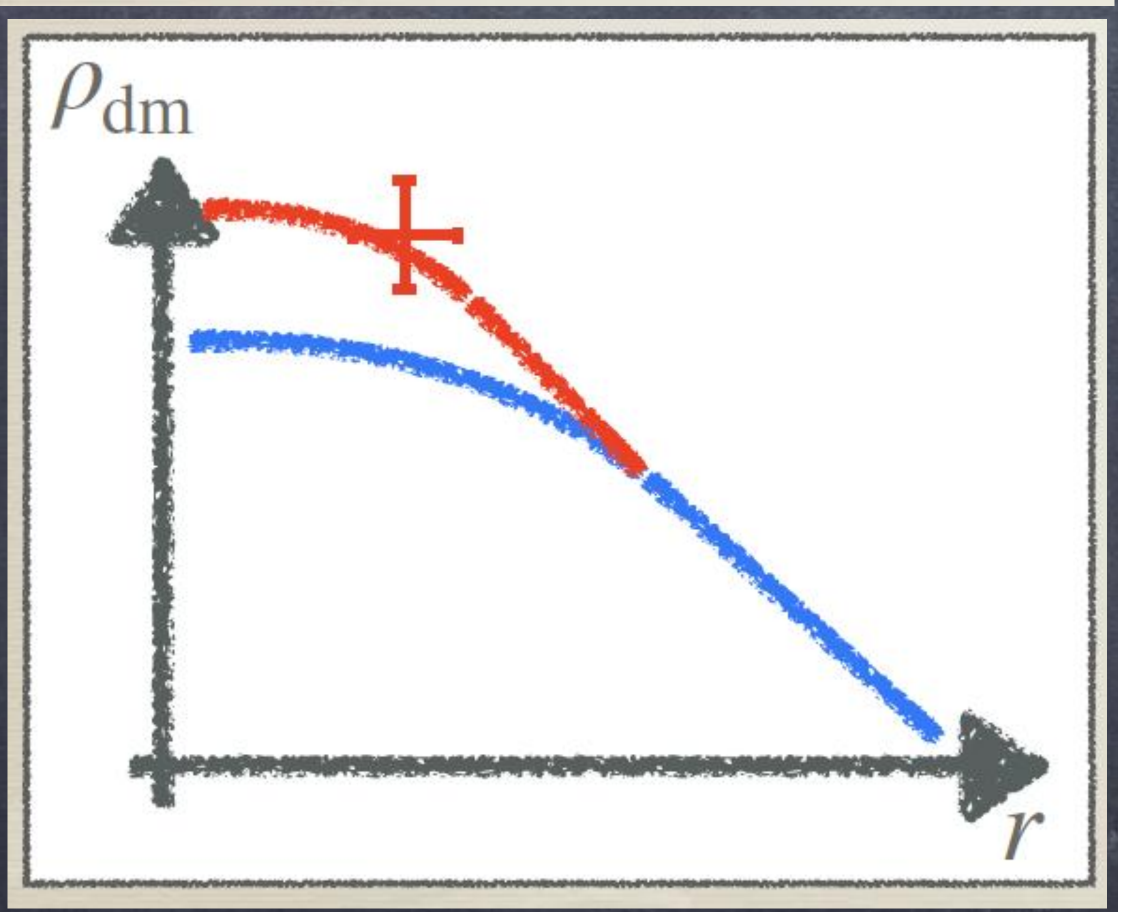
- SIDM with large cross sections produces a mixture of core-collapsing and isothermal-coring subhalos

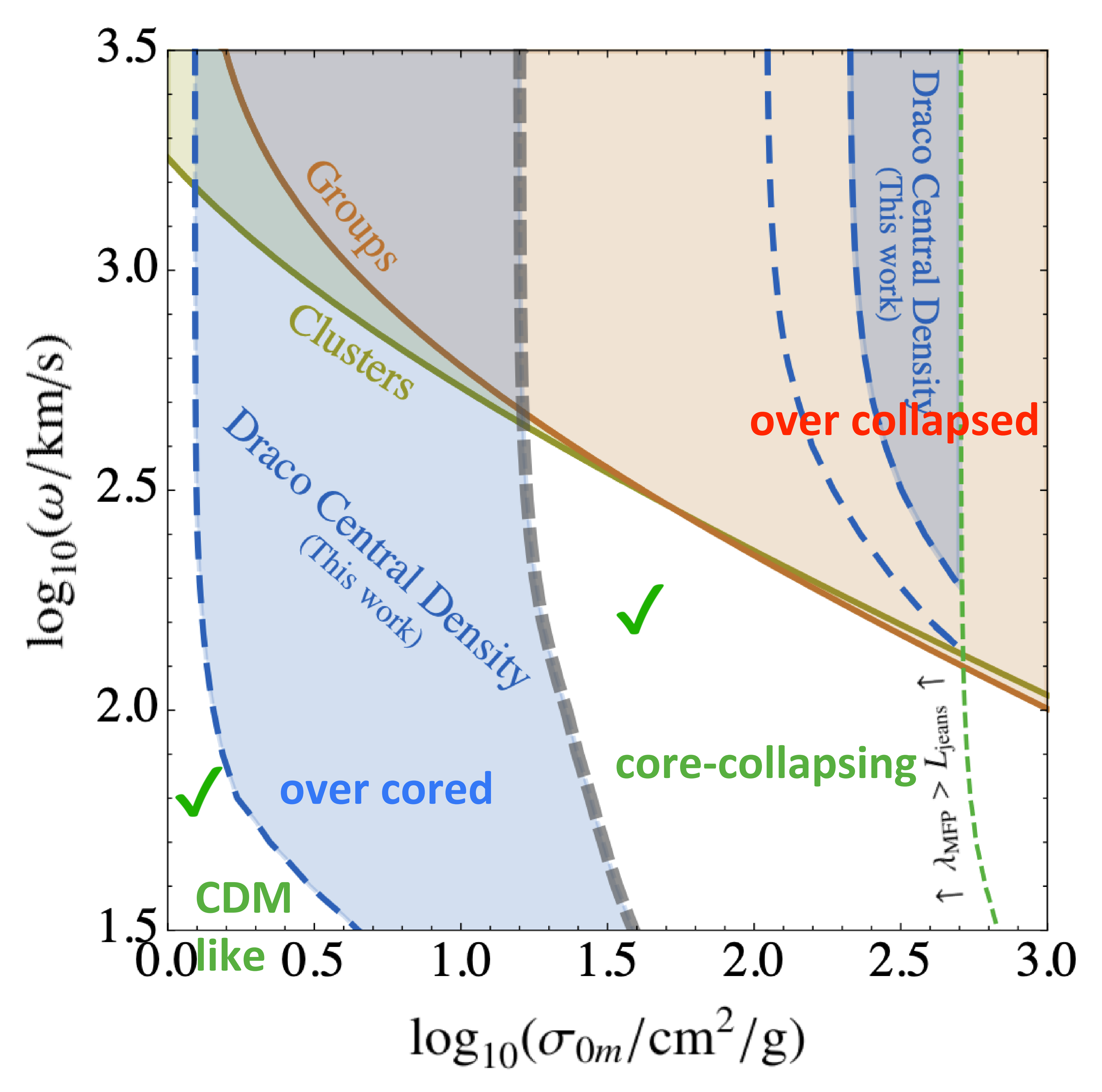
Constrain SIDM cross section using dwarf central density

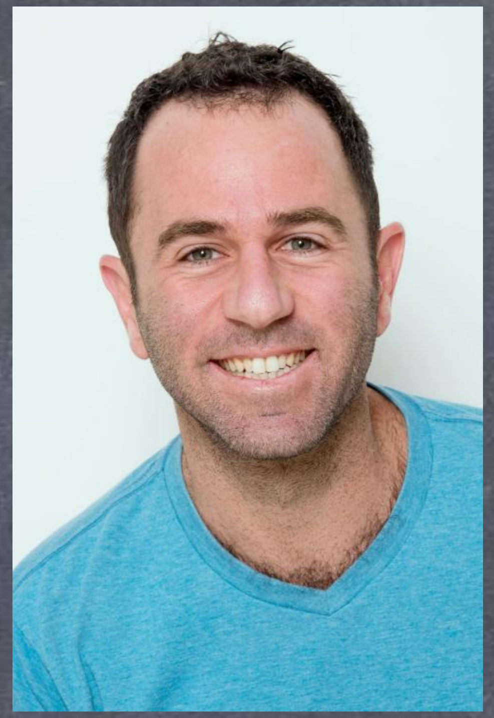
SatGen-SIDM proof-of-concept

Draco's measured central density:

$\rho_{150} = 0.212 \pm 0.045\, M_\odot/\mathrm{pc}^3$







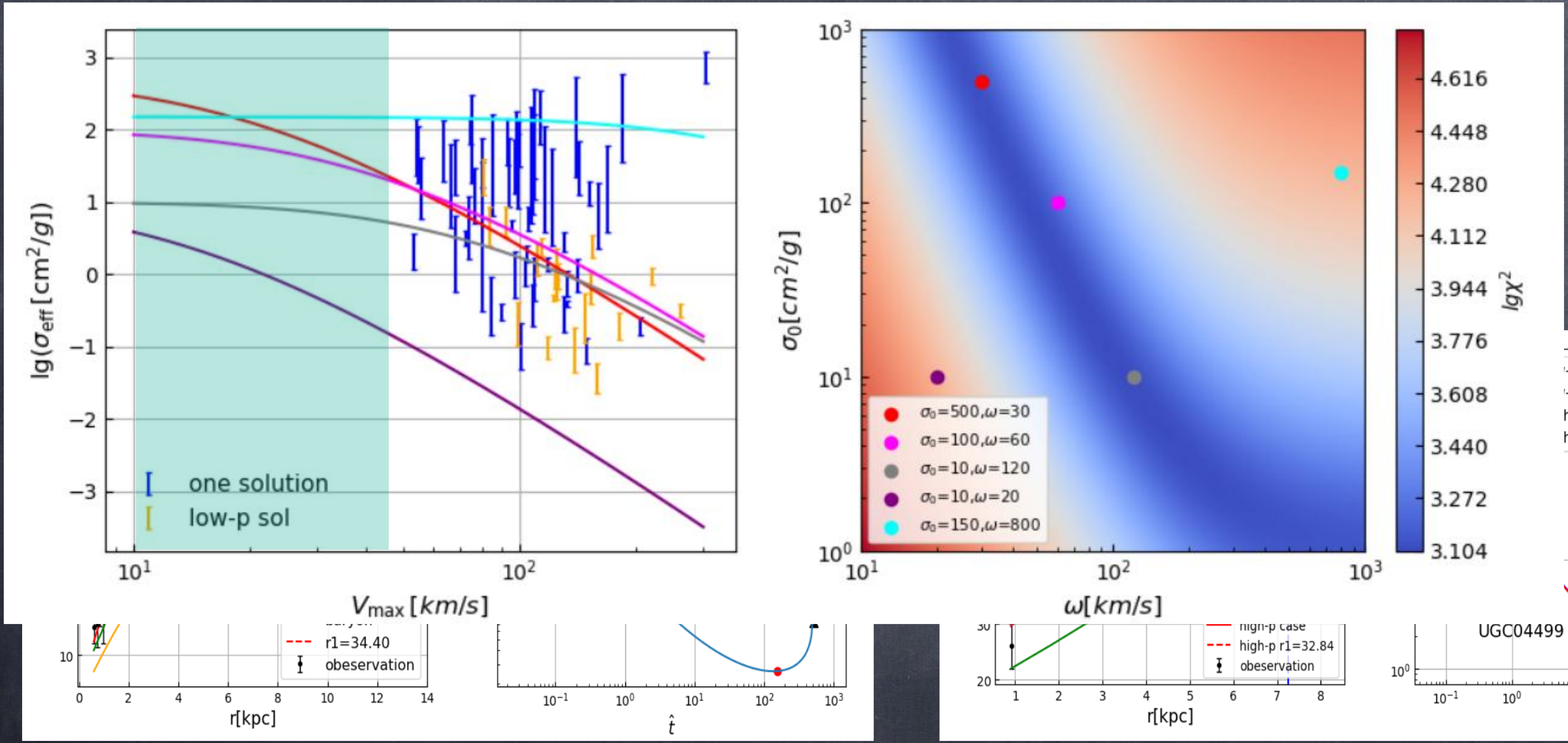
Oren Slone

Slone, **Jiang**, Lisanti, Kaplinghat, 23, PRD, 107, 4

- for SIDM to be interesting, the cross section needs to be large: σ_{m0} ~ 10-1000 cm^2 /g, ω ~ 10-100 km/s

Constrain SIDM cross section with galaxy kinematics

Jia, Jiang* + 25 in prep



Zixiang Jia 郑梓翔

- **with spatially resolved kinematics (IFU) + HI observations (e.g., from FAST core array) can break the degeneracy between σ_{m0} ~ 100-1000 cm^2 /g, ω ~ 10-100 km/s and σ_{m0} <10 cm^2 /g , ω > 100 km/s**

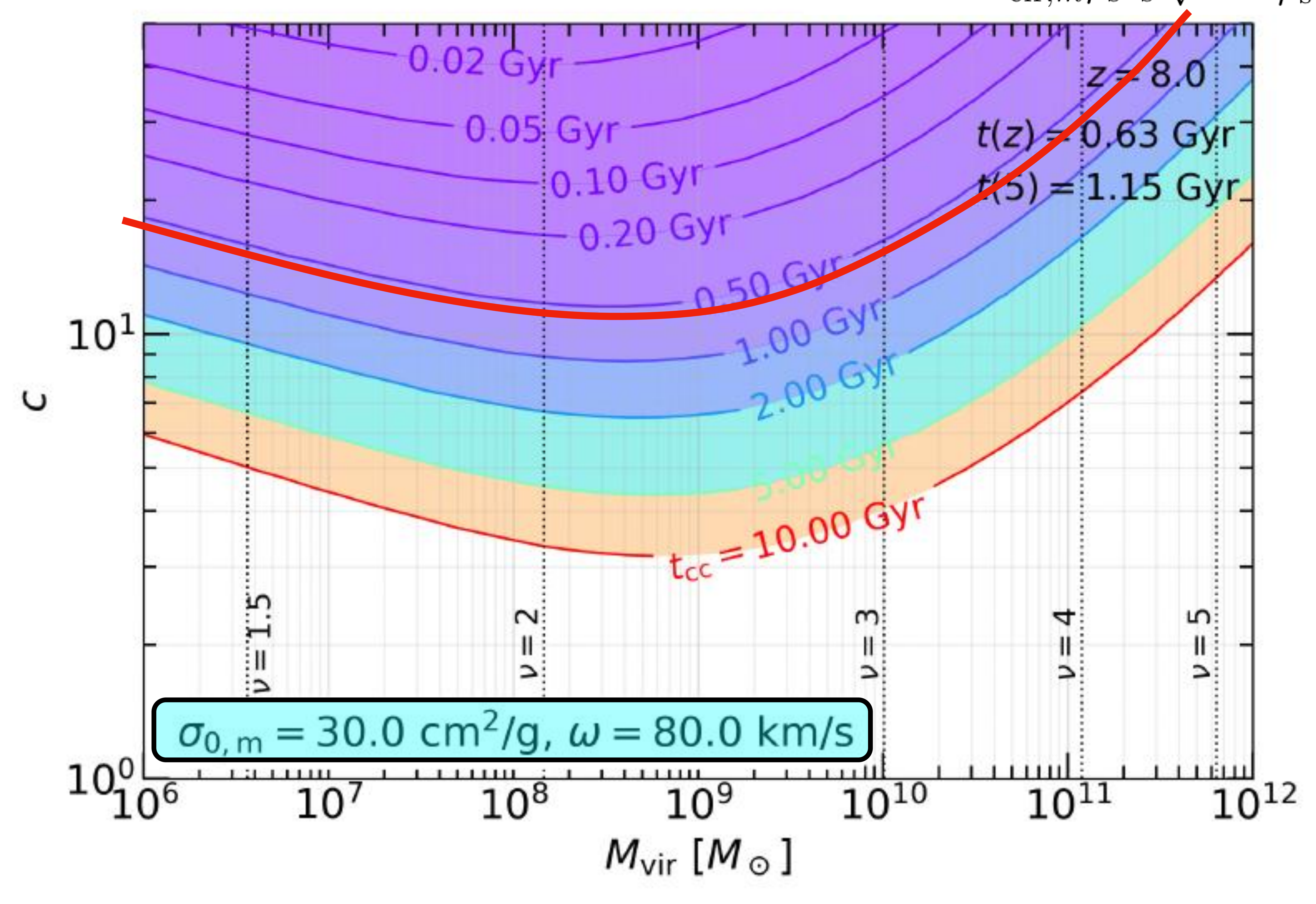
FJ + 2025, 2503.23710

Condition for BH seeding by SIDM core-collapse

t_{cc} (M_{vir}, c, σ_{0m}, ω | z)

halo mass — halo concentration

$$t_{\rm cc} \approx \frac{150}{C}\,\frac{1}{\sigma_{{\rm eff},m}\rho_{\rm s} r_{\rm s}}\,\frac{1}{\sqrt{4\pi G\rho_{\rm s}}}$$



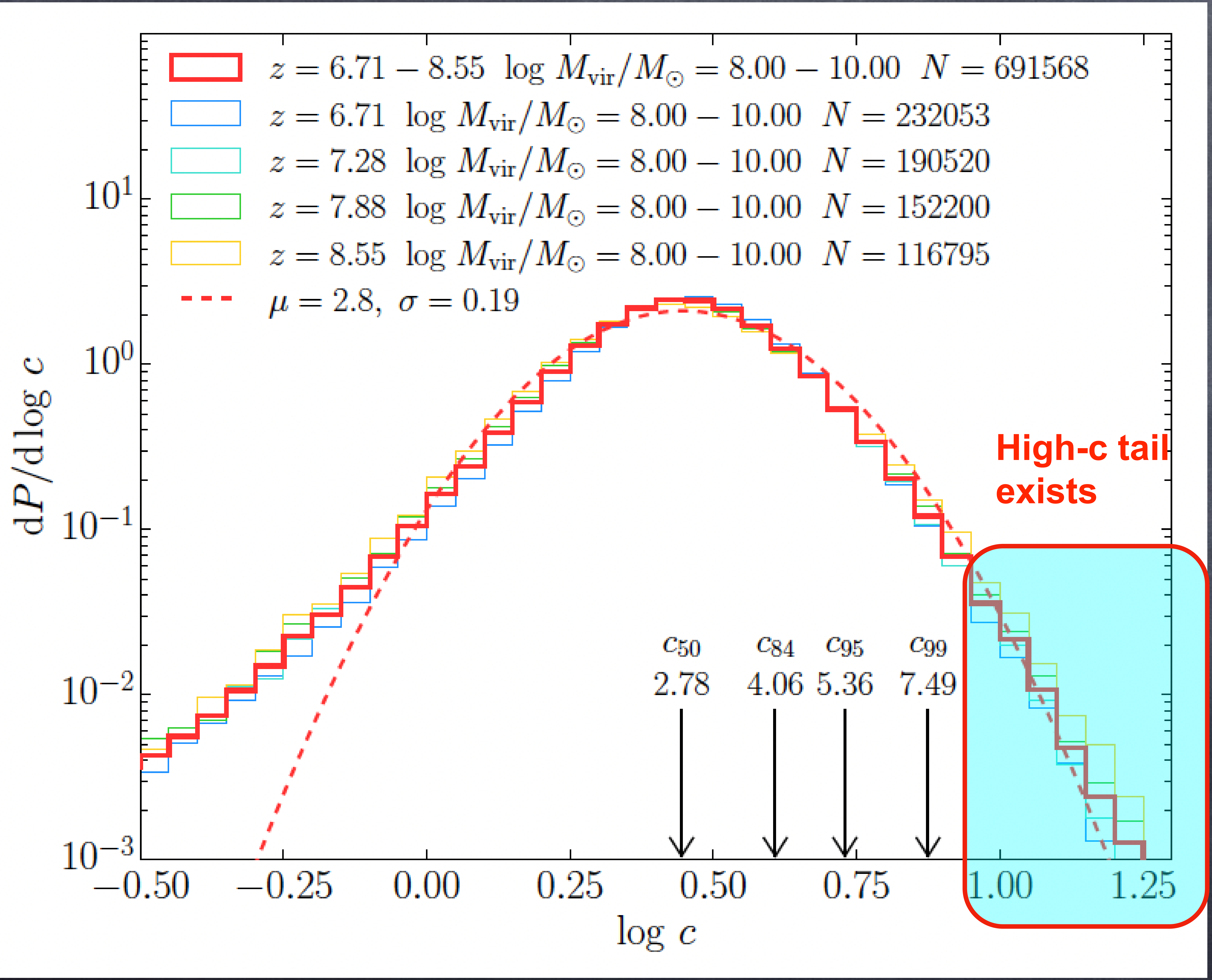
- Time to core-collapse and thus seed BH depends on halo mass, concentration, and cross section: higher c, larger cross section -> faster core-collapse

  e.g., Balberg & Shapiro 2002; Pollack+2015; Yang+2023 …

- Given cross-section & c, there is a characteristic halo mass where t_{cc} is shortest
- Mass trend: mass too low -> DM density too low -> inefficient self-interaction;
- mass too high -> cross section too small (because of velocity-dependence) -> inefficient self-interaction

Concentration at extremely high z from cosmological simulations

FJ + 2025, 2503.23710

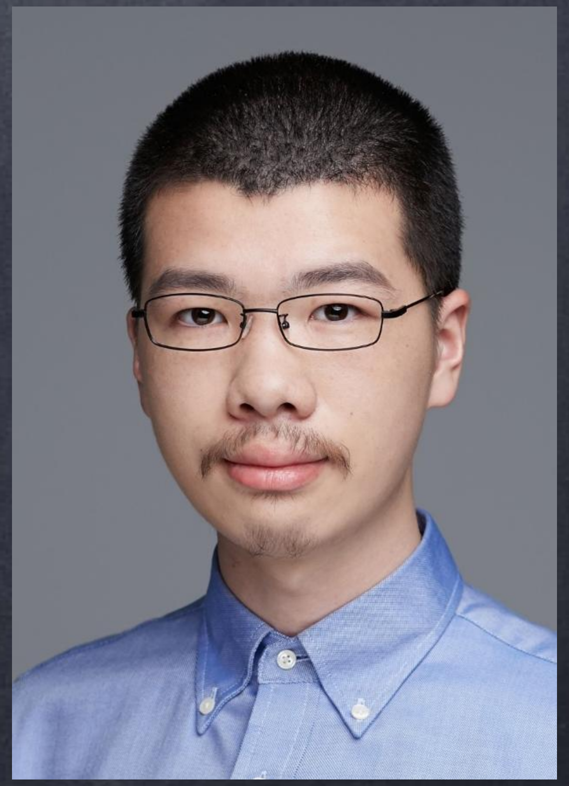


- Halo concentration exhibits marginal trend with mass or redshift at extremely high redshifts, and can reach values as high as 10-15, needed for SIDM core-collapse

$\log c \in N\,[\,\mu = \log(2.8),\ \sigma = 0.19]$

From “VVV” cosmological simulations optimized for high-z halos, measured by KIAA postdoc Haonan Zheng;

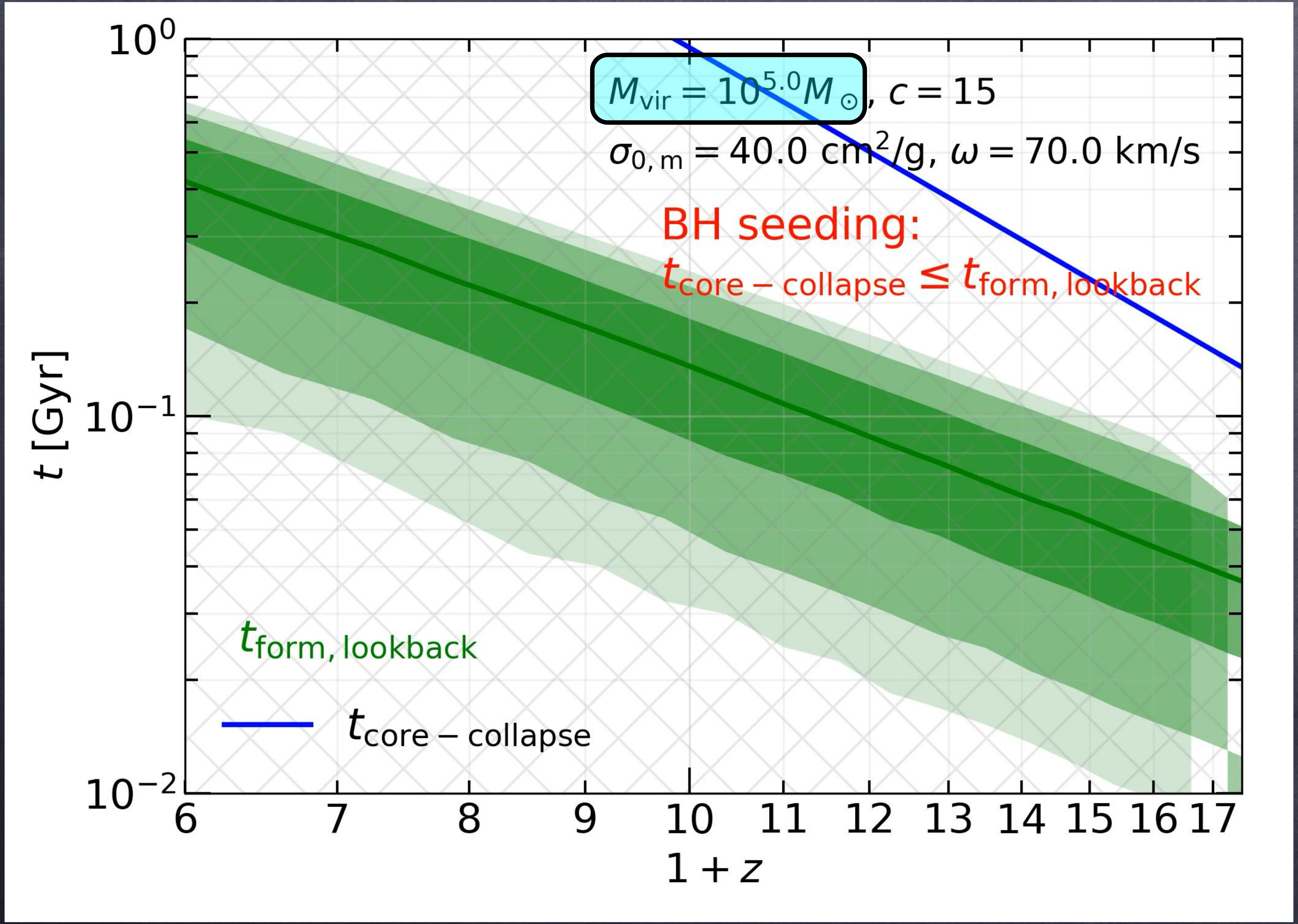
Also confirmed in Gureft, Thesan-HR simulations



Condition for BH seeding by SIDM: collapse time < time since halo formation

FJ + 2025, 2503.23710

$$t_{cc} < t_{form,lookback}(M_{vir}, c \mid z)$$



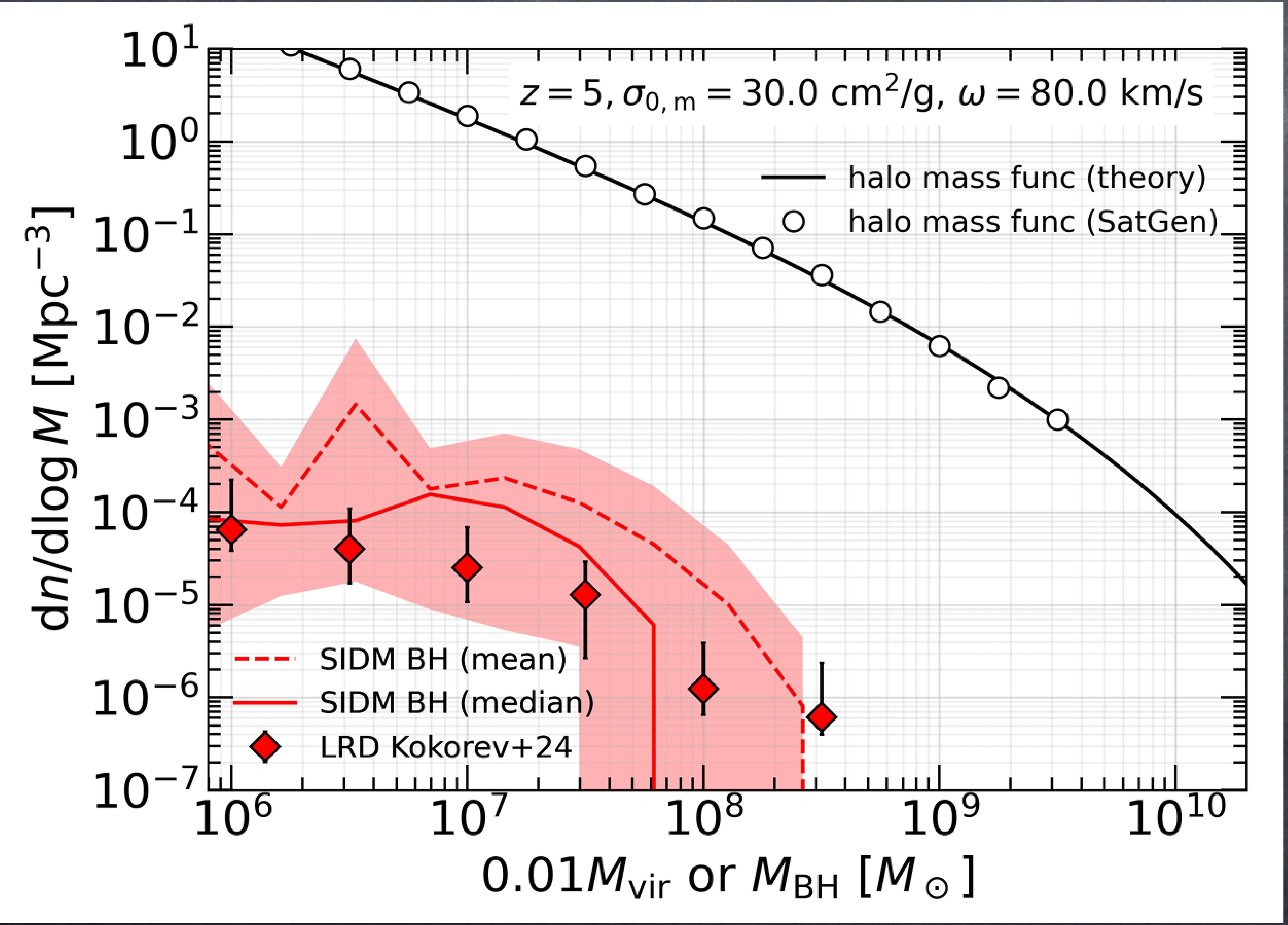
- In this movie, we keep *c* and cross section fixed, and scan halo mass → BH is seeded only for halos of a certain mass range $M_{vir} \approx 10^{6\text{-}8.5} M_{sun}$ and thus give BH seed mass of $M_{BH} \approx 10^{4\text{-}6.5} M_{sun}$

- Interestingly, this seeding mechanism only operates at very high redshift ($z>8.5$) — so there is no need to worry that all halos eventually collapse such that BHs are too abundant

- Technical note: green bands indicate halo formation-time distributions, (extended-Press-Schechter theory)

Explain LRDs with semi-analytic model of SIDM BH seeds + simple BH growth

FJ + 2025, 2503.23710

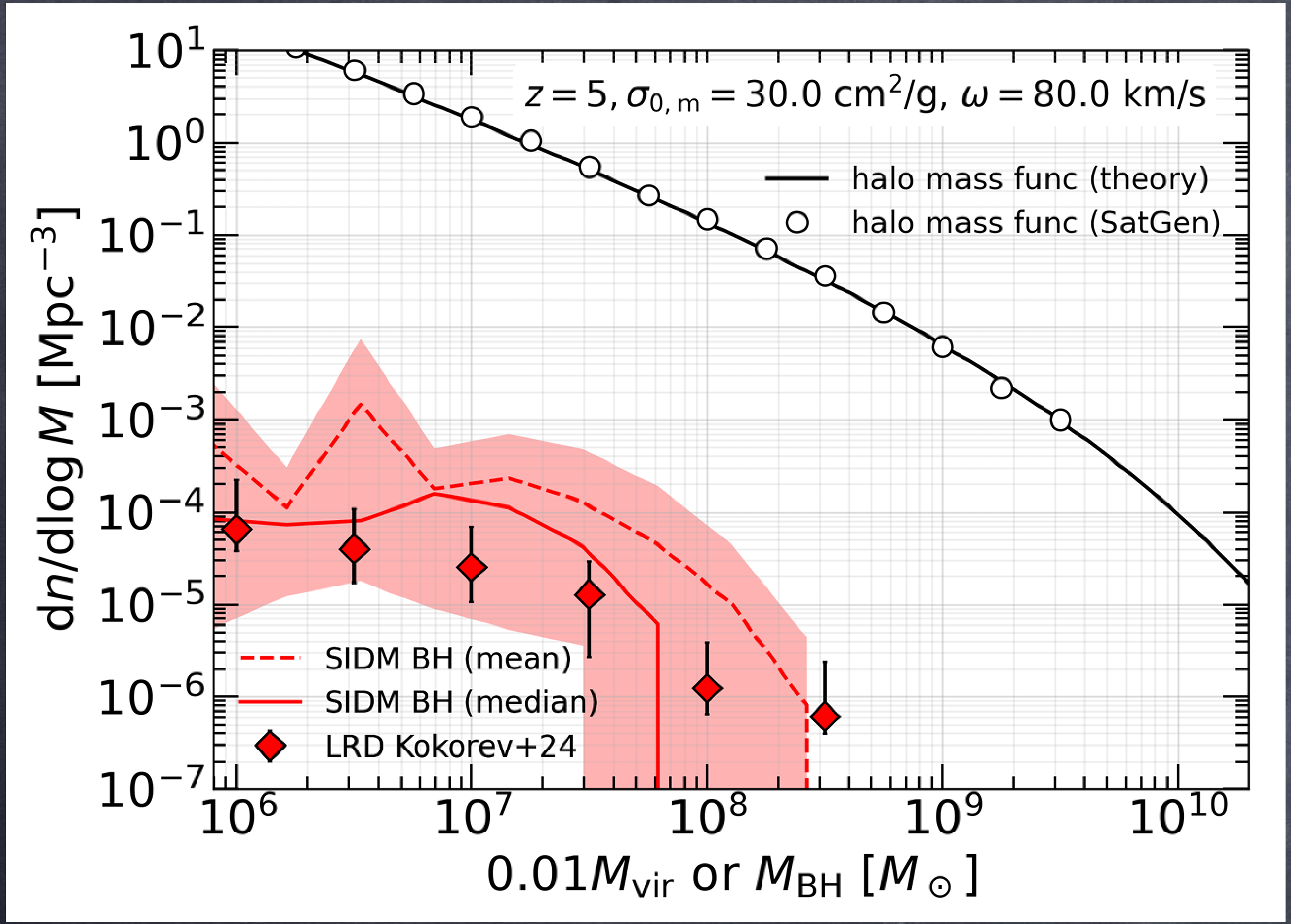


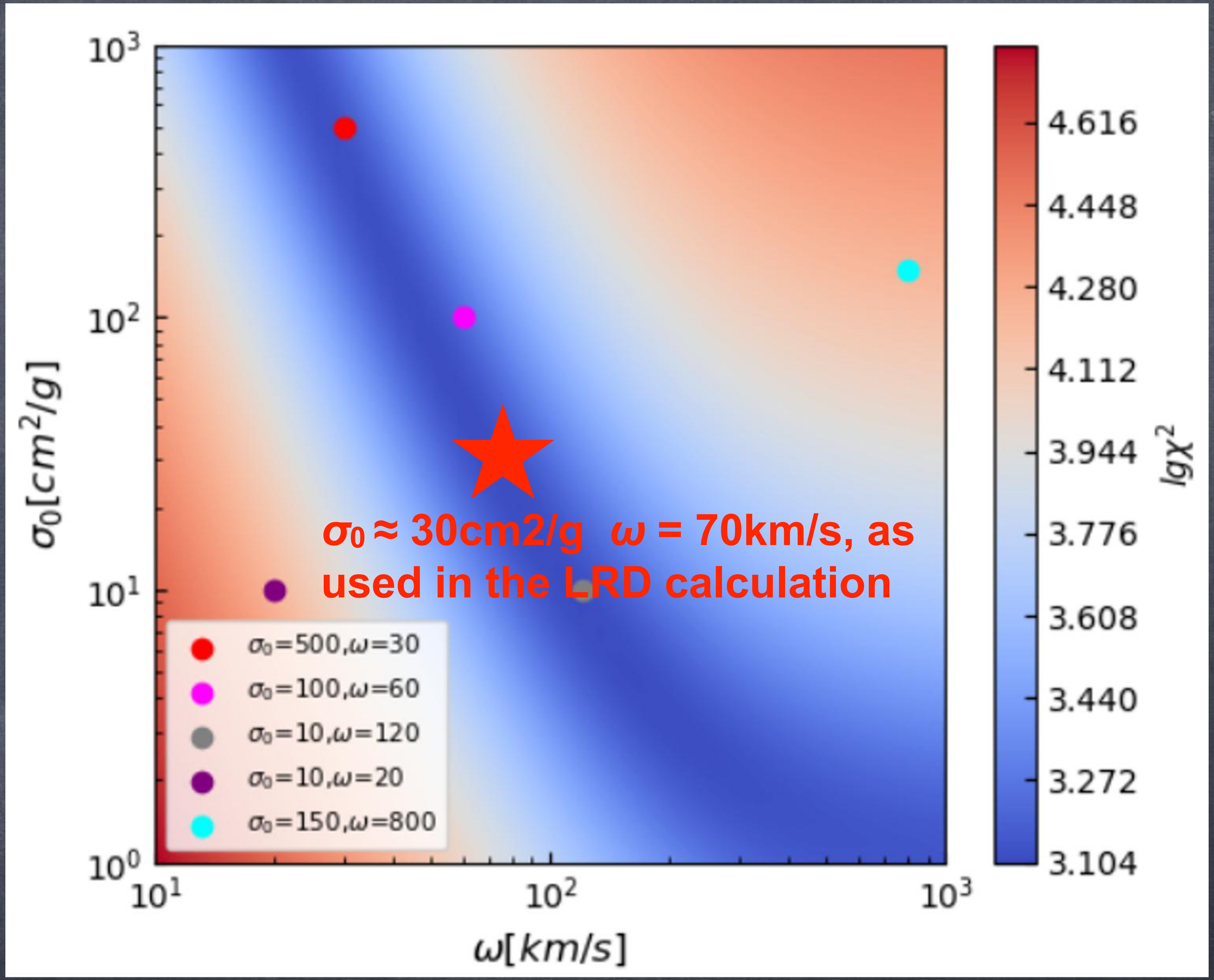
- Model ingredients:
  - halo merger trees
    Jiang+21 **SatGen** model
  - BH seeding by SIDM core-collapse if
    $t_{\mathrm{cc}} < t_{\mathrm{form,lookback}}$
  - BH mass growth w. Eddington ratio
    $\log \lambda \in N\,[\,\mu = \log(0.4),\ \sigma = 0.3]$
  - BH mergers by dynamical friction

- SIDM cross section adopted: $\sigma_0 \approx 30\ \mathrm{cm^2/g}$ $\omega = 70\mathrm{km/s}$ — interestingly consistent with local constraint (see next page) !

- (Not shown here) LRD mass function is sensitive to the SIDM cross section — promising way to constrain DM

LRDs may provide an orthogonal way of constraining SIDM cross section





- Blue band : local constraint of SIDM cross section from kinematics using ~100 rotation curves from SPARC and an improved SIDM model of Jiang+23, Kaplinghat+14,16
- Red star : the cross section used for the LRD mass function on the previous page — consistent with local constraint!

Take-away II. Alternative DM models

- FDM and WDM both suppress small-scale structures, with caveats (FDM: soliton core; WDM: lower mass limit); SIDM relates structural diversity to formation time
- Little Red Dots can be seeded from SIDM halos: on early-forming and high-concentration tail, ONLY at high z
- We use SatGen to present a statistical study of the BHs seeded by SIDM core-collapse, and reproduce observed LRD mass functions without fine-tuning
- Seeding BHs in the early universe is inevitable in the SIDM paradigm, which potentially provides a orthogonal way to constrain SIDM