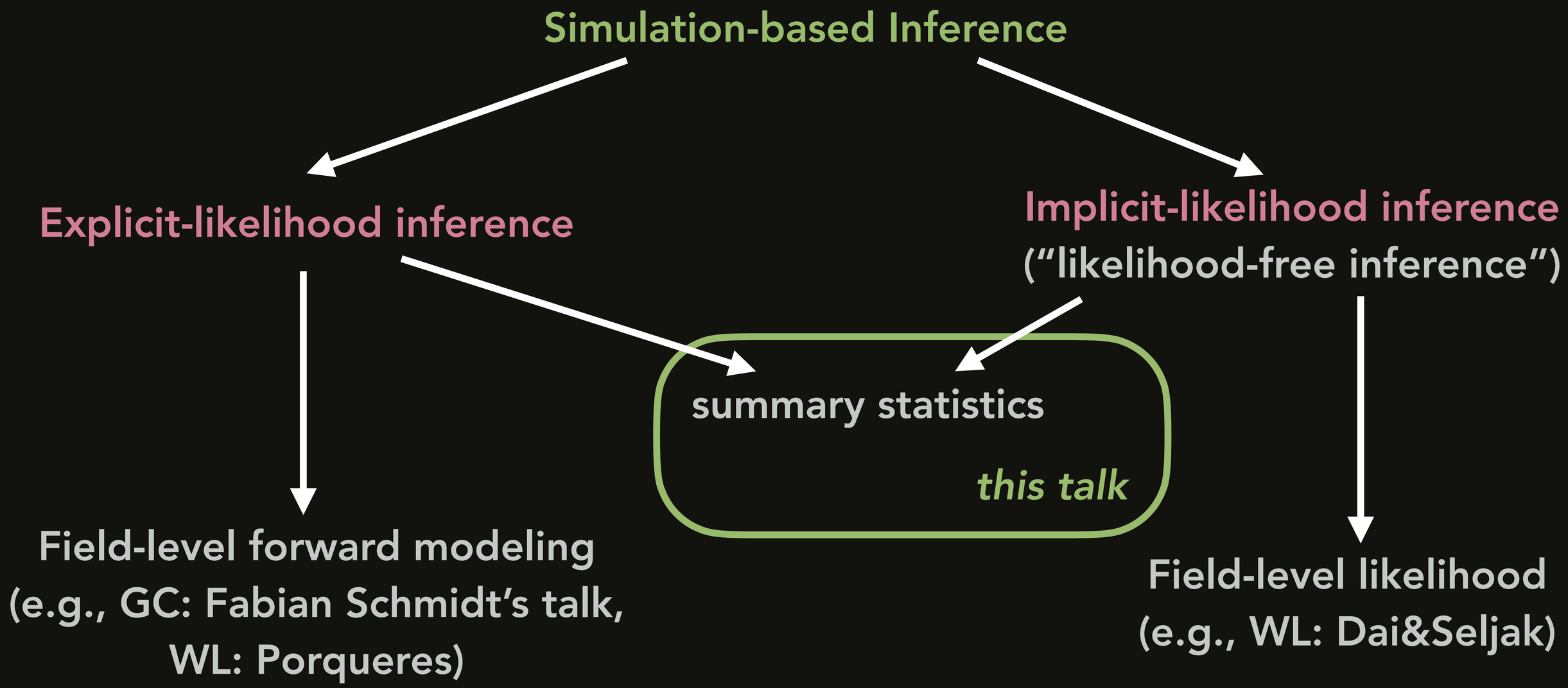
Simulation-based Inference

Leander Thiele (Kavli IPMU)

11th KIAS Workshop on Cosmology and Structure Formation, 10/28/2024

A word about words



Axiom:

We can simulate everything

… except black holes

Bayesian inference



Bayesian inference

More concretely:

θ=interesting parameters, η=nuisance parameters, ζ=initial conditions,

x=data, m=model:

$$P(x \mid \theta) = \int D\eta\, D\zeta\, \delta[\, x - m(\theta, \eta, \zeta)\,]$$



Bayesian inference

More concretely:

θ=interesting parameters, η=nuisance parameters, ζ=initial conditions,

x=data, m=model:

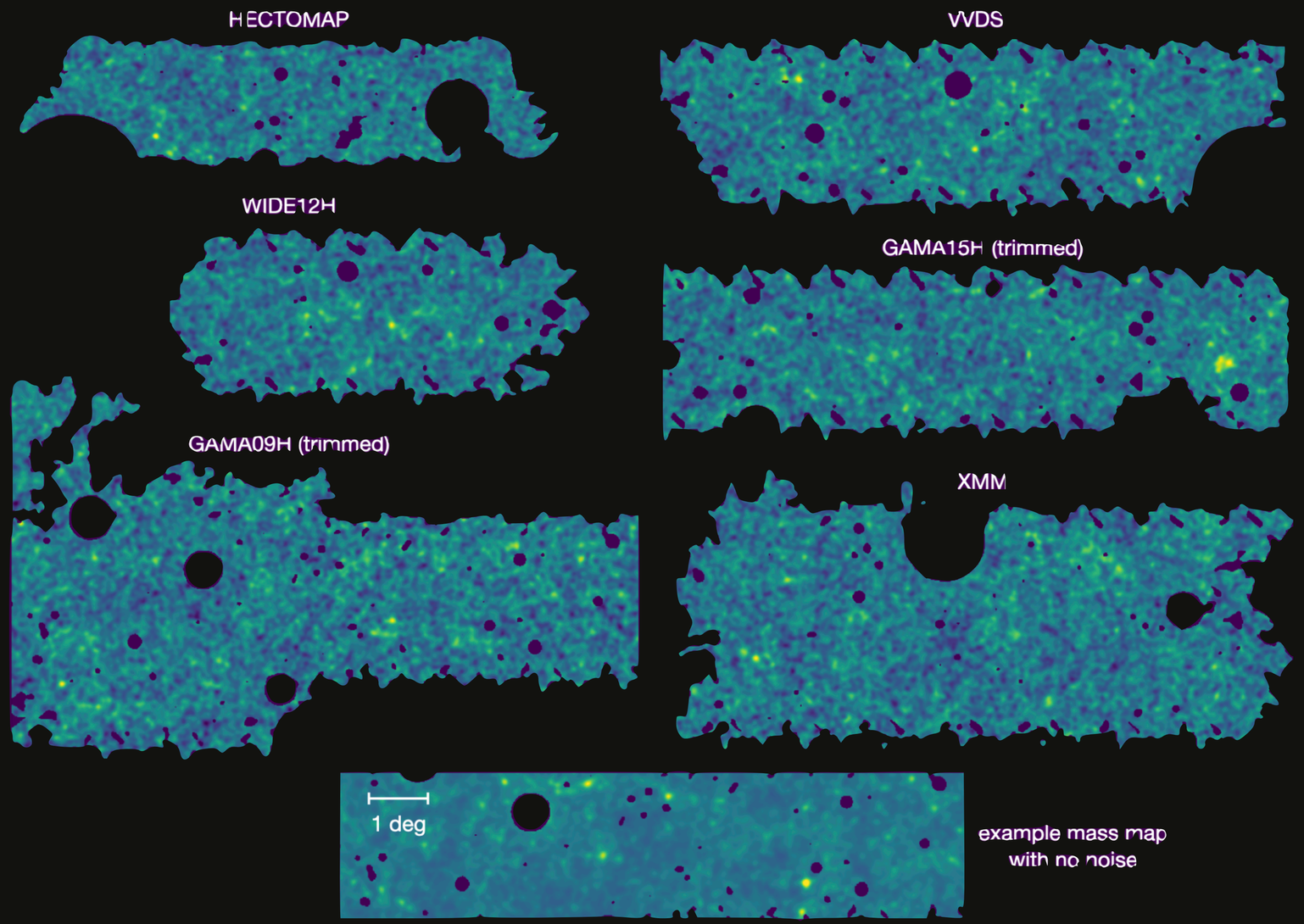
$$P(x \mid \theta) = \int D\eta\, D\zeta\, \delta[\, x - m(\theta, \eta, \zeta)\,]$$

Traditional case:

$$P(x \mid \theta) = \int D\eta\, \text{Gaussian}[\, x - \mu(\theta, \eta), \Sigma\,]$$

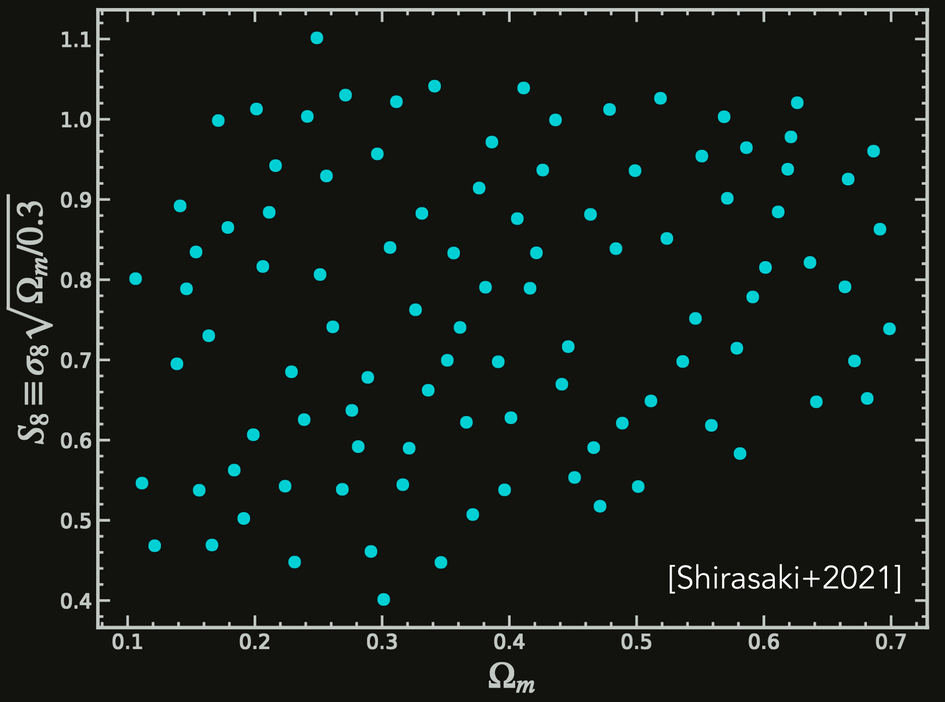
[do remaining low-dimensional η-integral with Monte Carlo]

Cosmology from HSC year-1 higher order statistics



example mass map
with no noise

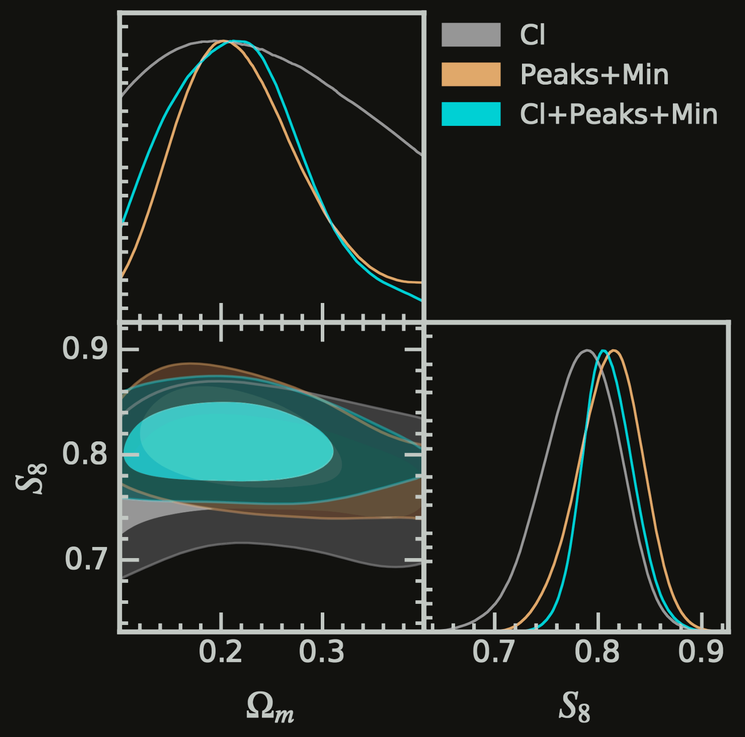
training simulations



with Jessica Cowell, Daniela Grandon, Joaquin Armijo, Camila Novaes, Sihao Cheng, Gabriela Marques, Masato Shirasaki, Jia Liu

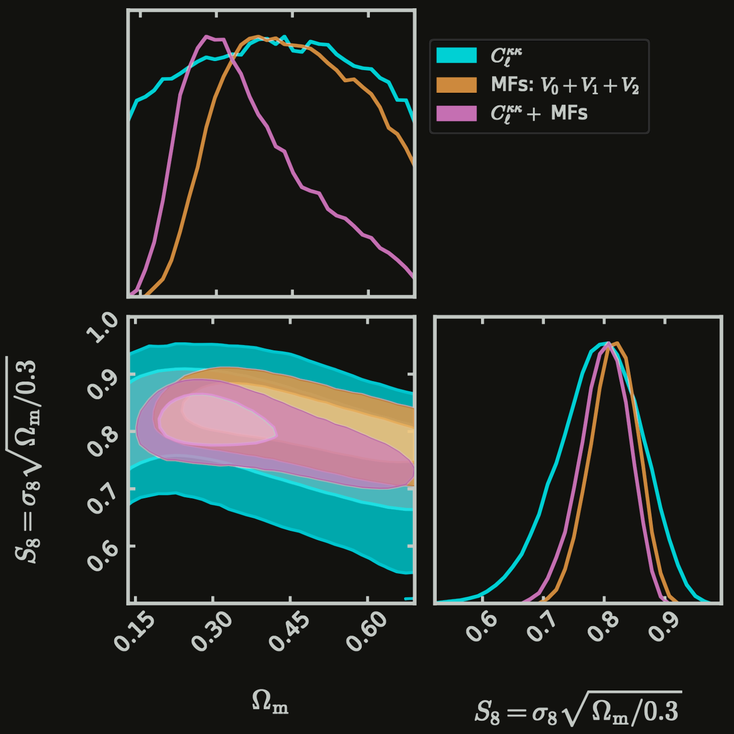
Cosmology from HSC year-1 higher order statistics

[Marques+2024]



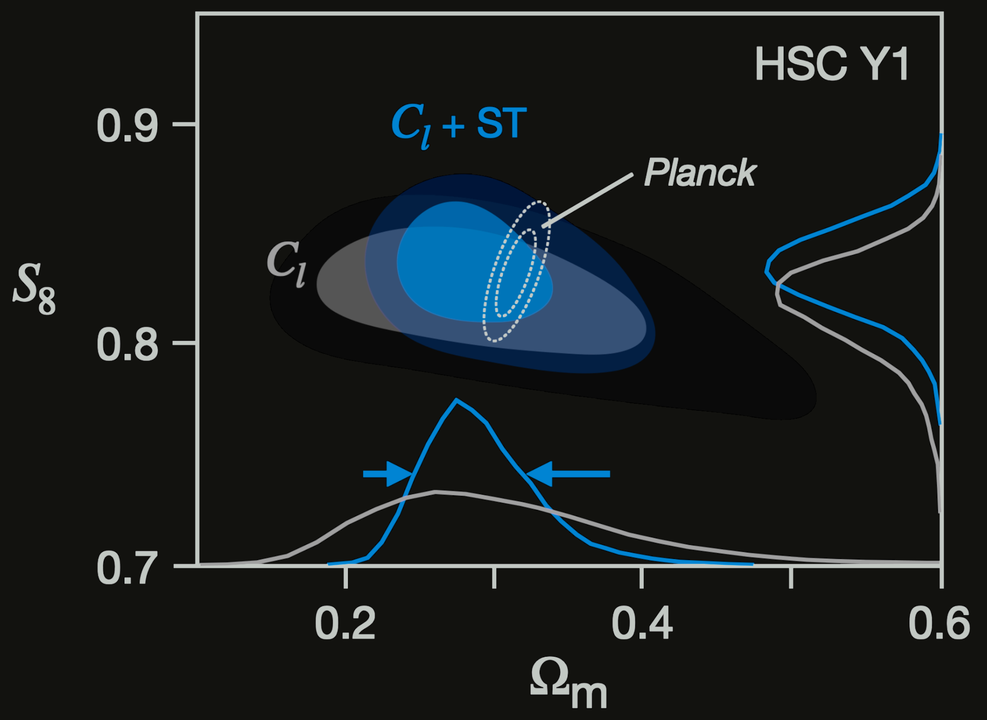
peaks and minima

[Armijo+2024]



Minkowski functionals

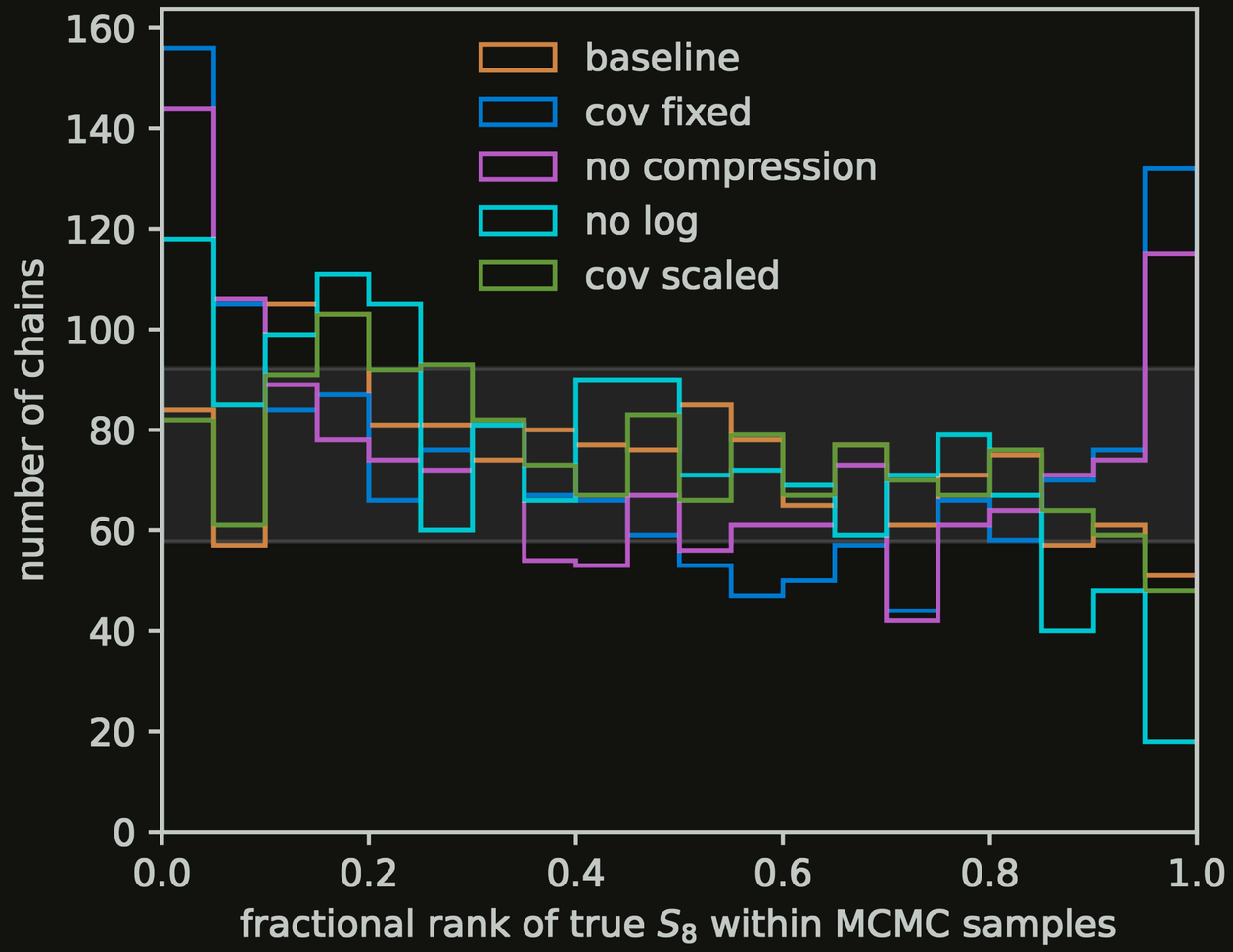
[Cheng+2024]

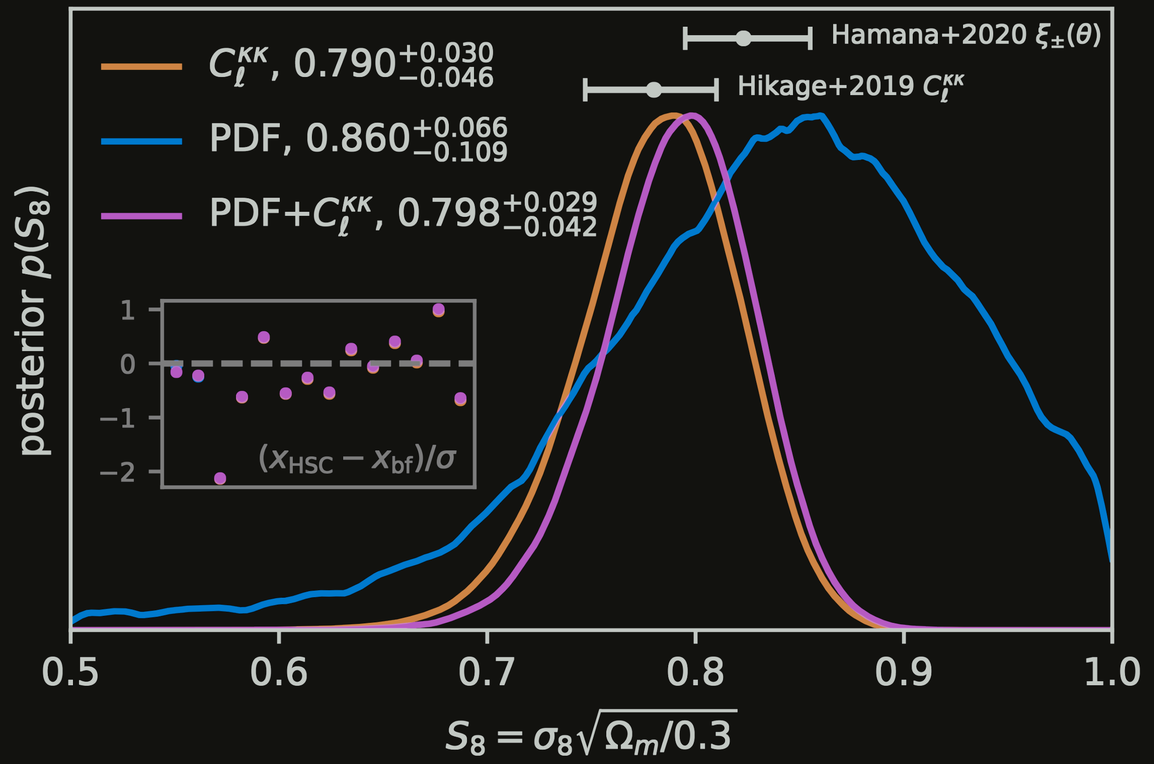


scattering transform

Cosmology from HSC year-1 higher order statistics

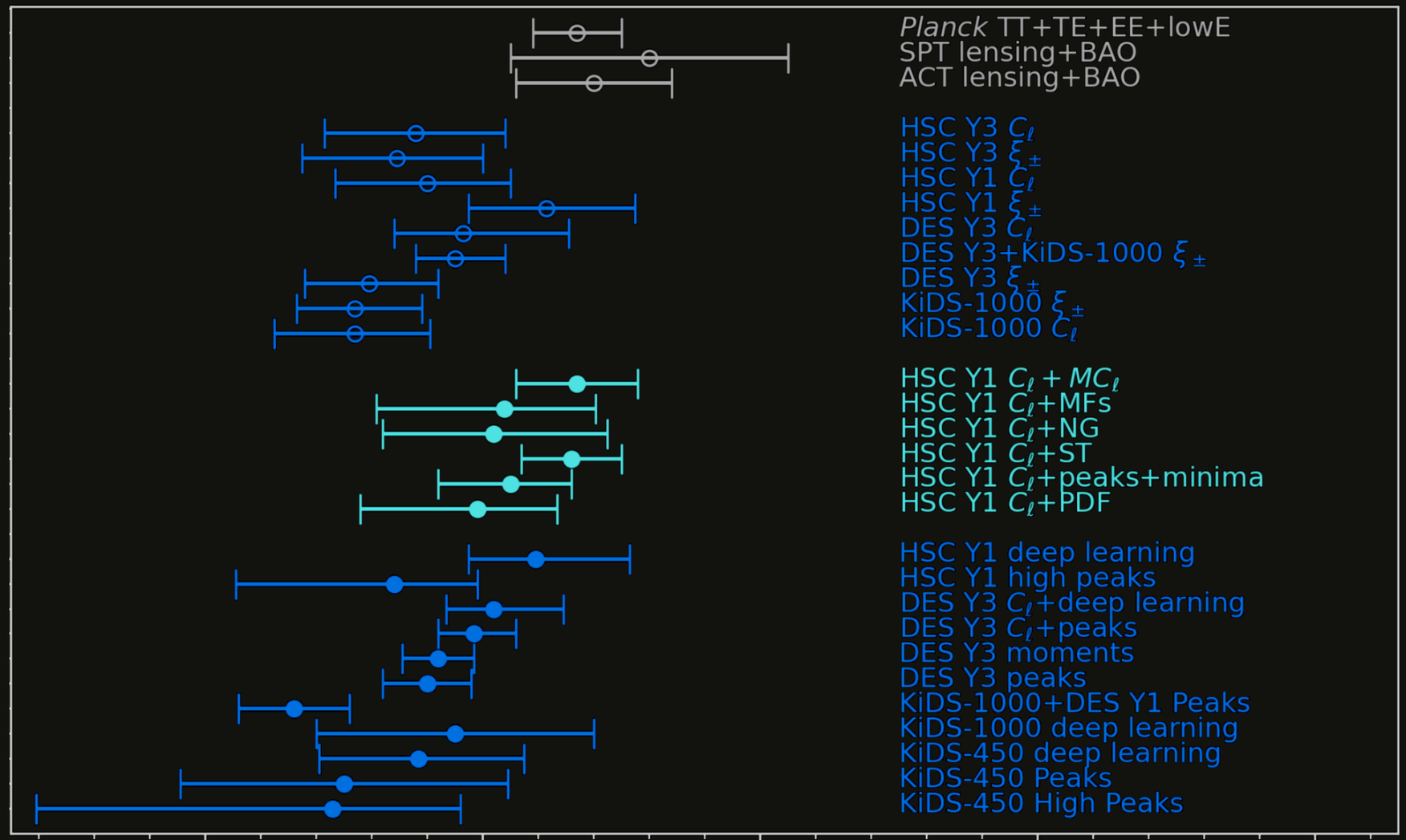
PDF: likelihood a bit tricky, guess an approximation using calibration diagnostics





[Thiele+2023]

Cosmology from HSC year-1 higher order statistics



Aghanim+20
Bianchini+20
MadhavacheriI+23

Dalal+23
Li+23
Hikage+19
Hamana+20
Doux+22
DES & KiDS 23
Secco+22 & Amon+22
Asgari+21
Loureiro+22

Cowell+24
Armijo+24
Novaes+24
Cheng+24
Marques+24
Thiele+23

Lu+23
Liu+23
Jeffrey+24
Zürcher+22
Gatti+22
Zürcher+22
Harnois-D'eraps+24
Fluri+22
Fluri+19
Martinet+18
Shan+18

Bayesian inference

More concretely:

θ=interesting parameters, η=nuisance parameters, ζ=initial conditions,

x=data, m=model:

$$P(x \mid \theta) = \int D\eta\, D\zeta\, \delta[\, x - m(\theta, \eta, \zeta)\,]$$

Traditional case:

$$P(x \mid \theta) = \int D\eta\, \text{Gaussian}[\, x - \mu(\theta, \eta), \Sigma\,]$$

But what do we do if the Gaussian approximation doesn't hold?

→ assume we have simulator that evaluates $m(\theta, \eta, \zeta)$ accurately

Neural Implicit Likelihood Inference (ILI)

More concretely:

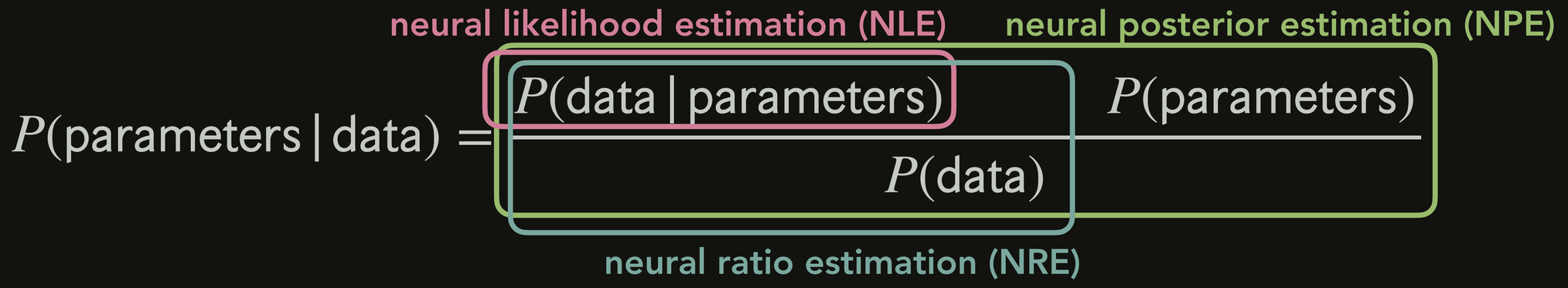
θ=interesting parameters, η=nuisance parameters, ζ=initial conditions,

x=data, m=model:

$$P(x \mid \theta) = \int D\eta\, D\zeta\, \delta[\, x - m(\theta, \eta, \zeta)\,]$$

But what do we do if the Gaussian approximation doesn't hold?

→ assume we have simulator that evaluates $m(\theta, \eta, \zeta)$ accurately



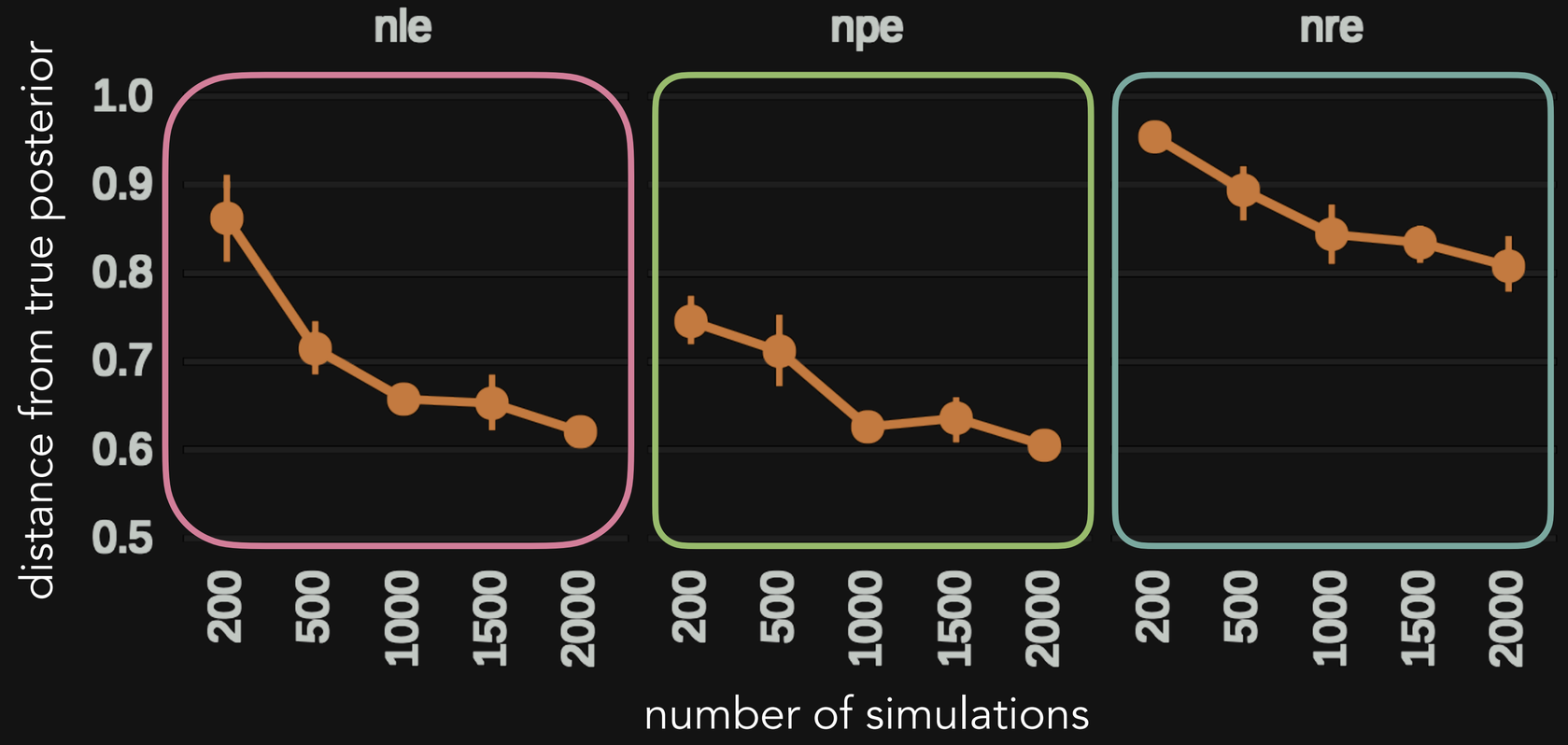
Neural Implicit Likelihood Inference (ILI)

neural likelihood estimation (NLE)

neural posterior estimation (NPE)

$$P(\text{parameters} \mid \text{data}) = \frac{P(\text{data} \mid \text{parameters})\, P(\text{parameters})}{P(\text{data})}$$

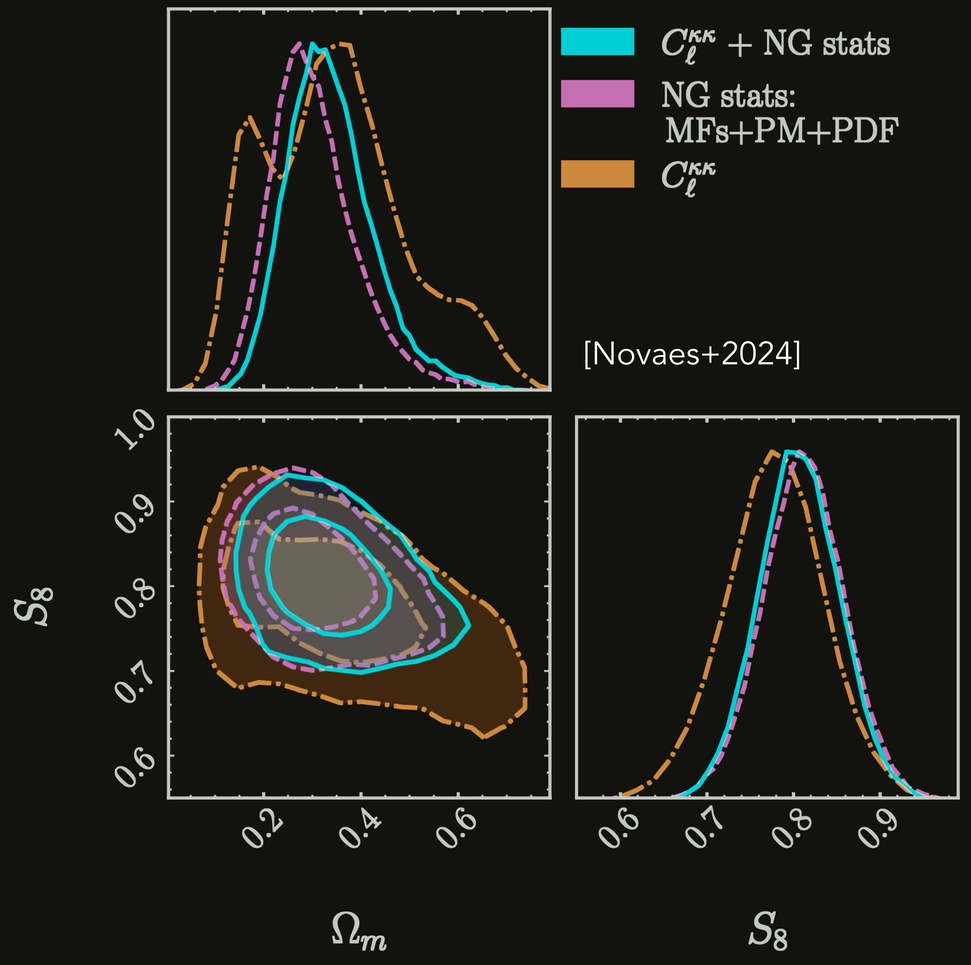
neural ratio estimation (NRE)



[Zeghal+2024]

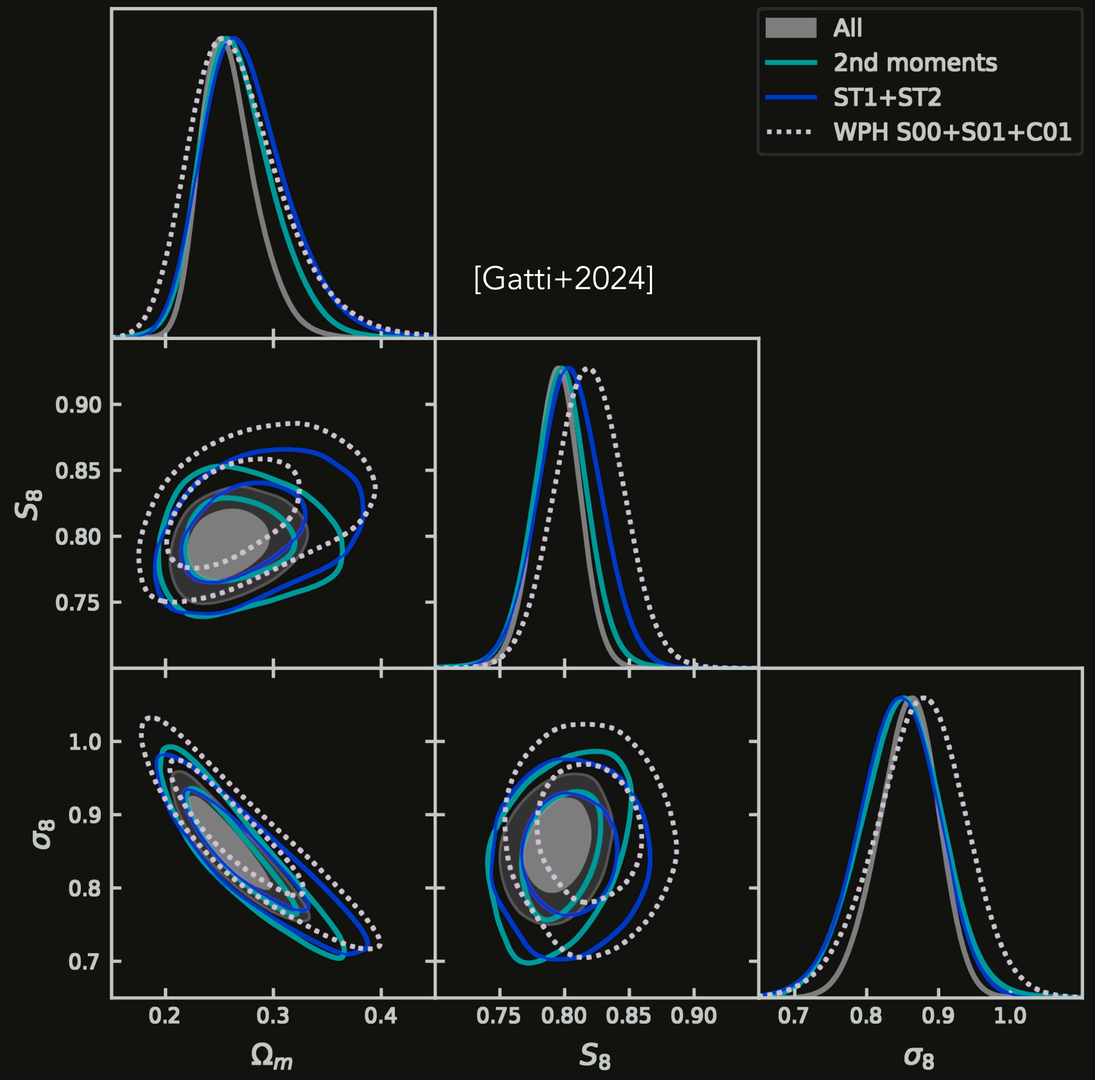
ILI applied in weak lensing

HSC Y1



[Novaes+2024]

DES Y3

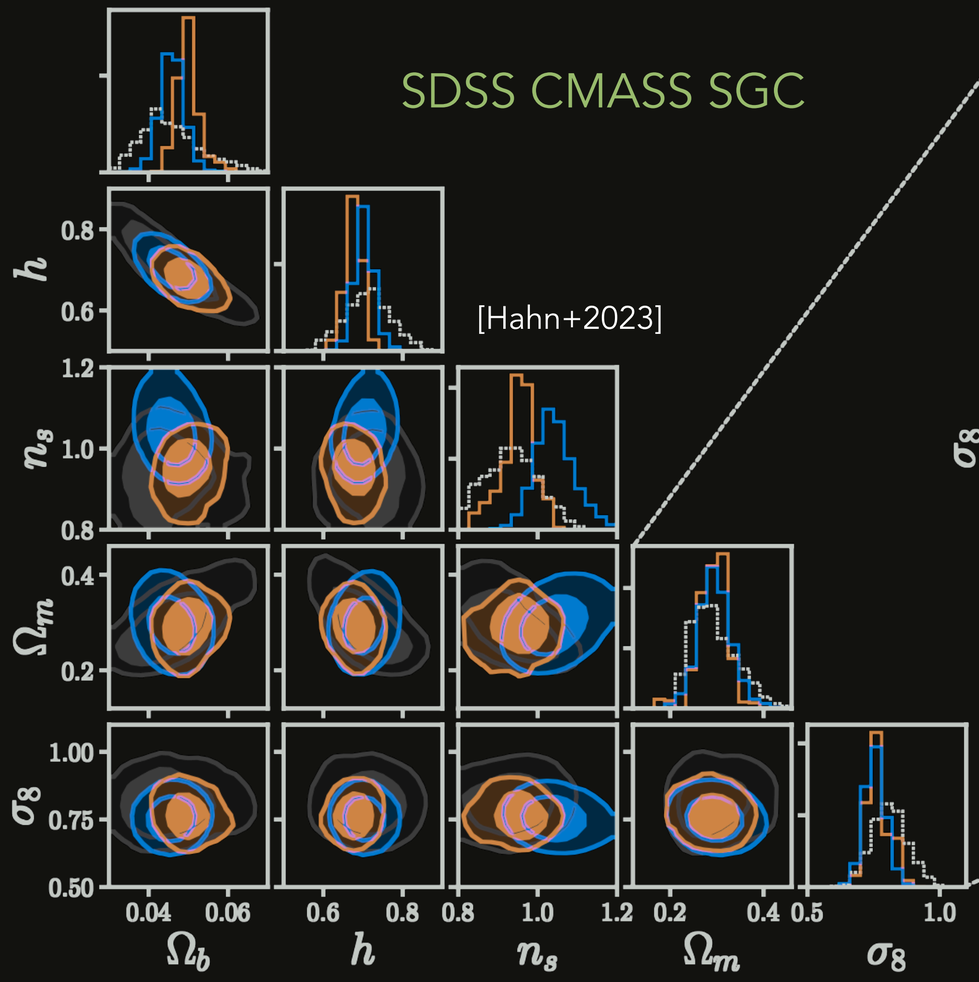


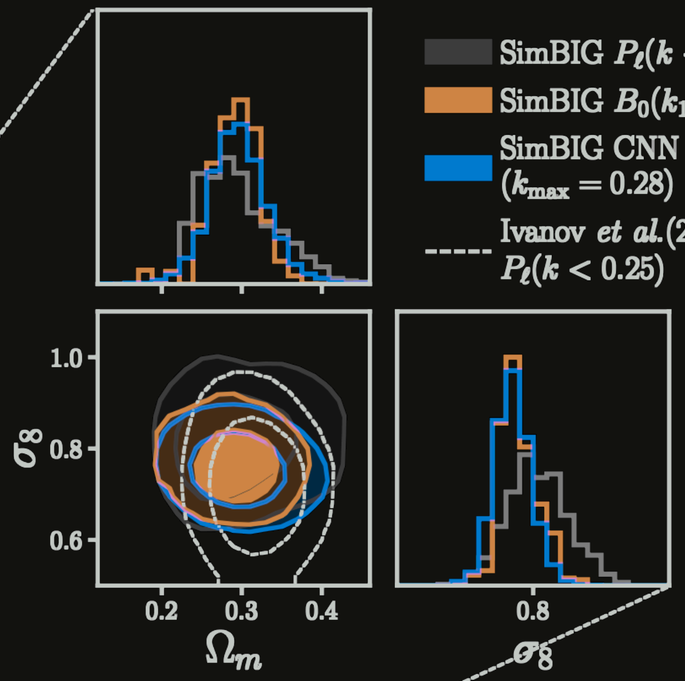
[Gatti+2024]

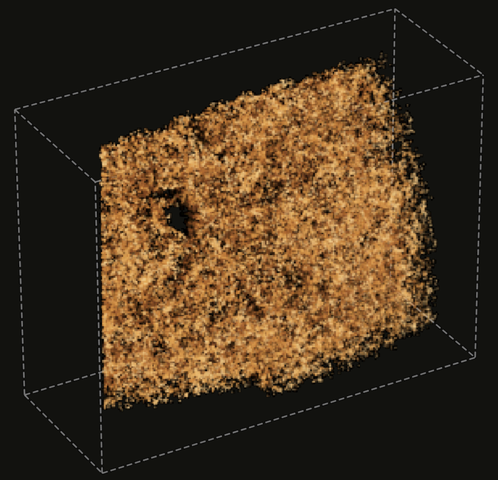
ILI applied in galaxy clustering

SDSS CMASS SGC

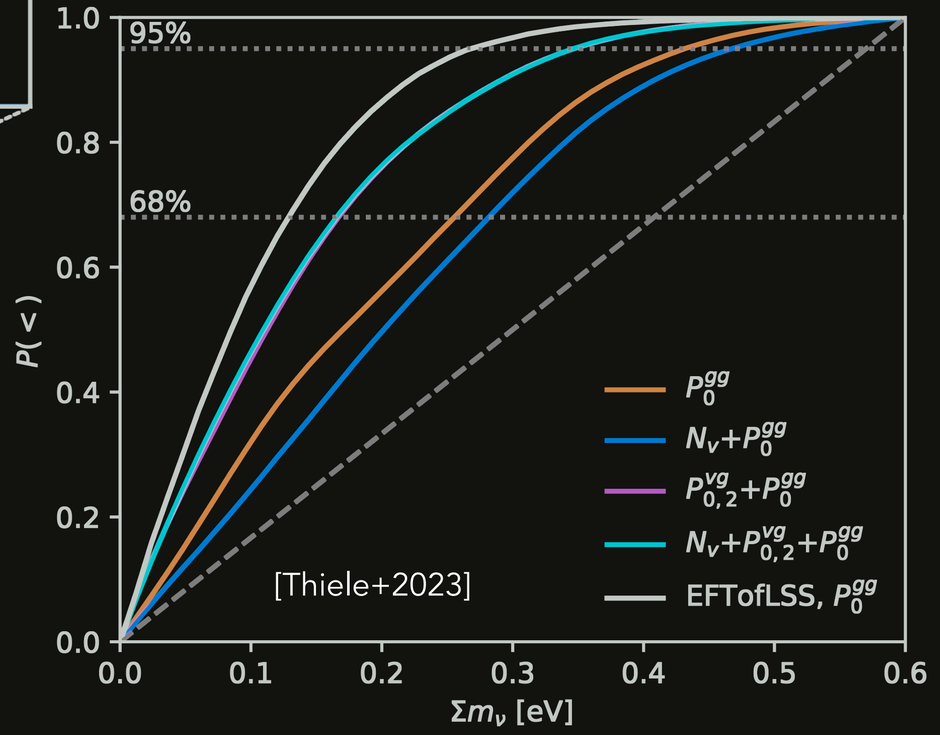
[Hahn+2023]







SDSS CMASS NGC voids



[Thiele+2023]

So far, more or less “toy examples”.

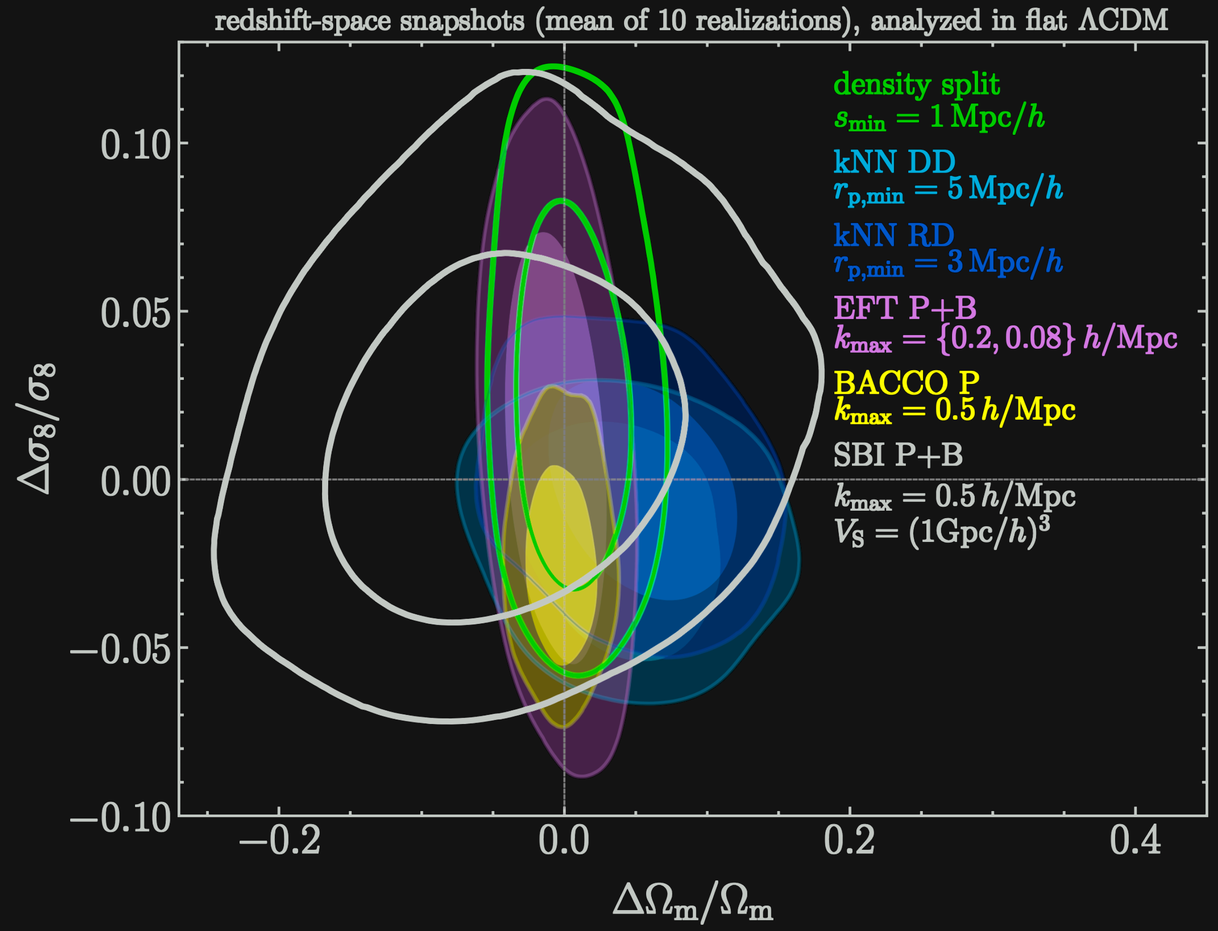
Not the same level of trust as traditional analyses.

A Parameter-Masked Mock Data Challenge for Beyond-Two-Point Galaxy Clustering Statistics*

The Beyond-2pt Collaboration

Elisabeth Krause,[1] Yosuke Kobayashi,[1,2] Andrés N. Salcedo,[1] Mikhail M. Ivanov,[3] Tom Abel,[4,5,6] Kazuyuki Akitsu,[7] Raul E. Angulo,[8,9] Giovanni Cabass,[10] Sofia Contarini,[11,12,13] Carolina Cuesta-Lazaro,[14,15,16] ChangHoon Hahn,[17] Nico Hamaus,[18,19] Donghui Jeong,[20,21] Chirag Modi,[22,23] Nhat-Minh Nguyen,[24,25] Takahiro Nishimichi,[2,26,27] Enrique Paillas,[28,29] Marcos Pellejero Ibañez,[30] Oliver H. E. Philcox,[31,32] Alice Pisani,[33,22,34,17] Fabian Schmidt,[35] Satoshi Tanaka,[26] Giovanni Verza,[36,22] Sihan Yuan,[4,6] Matteo Zennaro,[37]

building trust



So far, more or less “toy examples”.

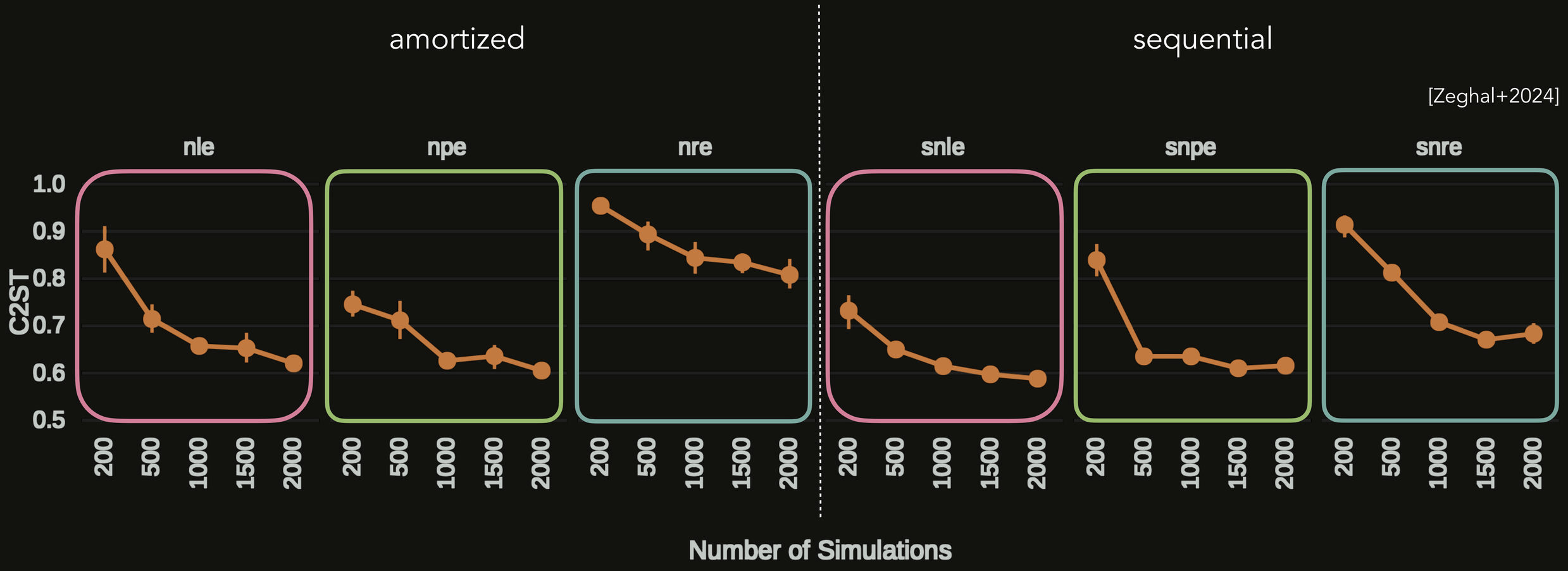
Not the same level of trust as traditional analyses.

Stage-IV deluge of data will make the problem much more challenging…

Need to develop clever ways to run better simulations for cheaper!

Sequential inference

representation learning only gets us so far…



Sequential inference

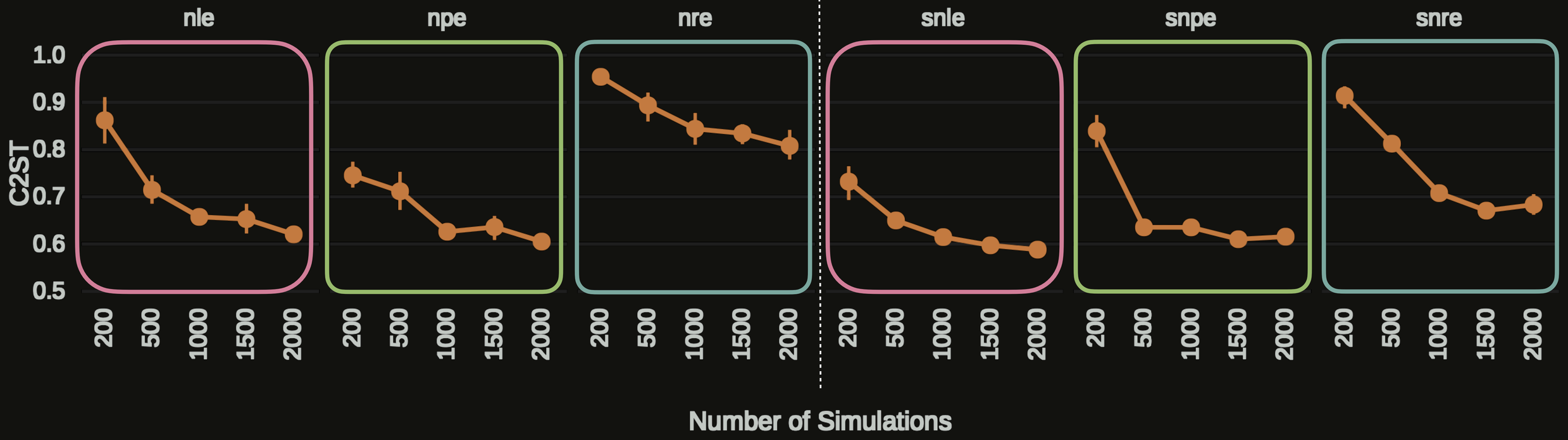
representation learning only gets us so far…

but what about coverage tests?

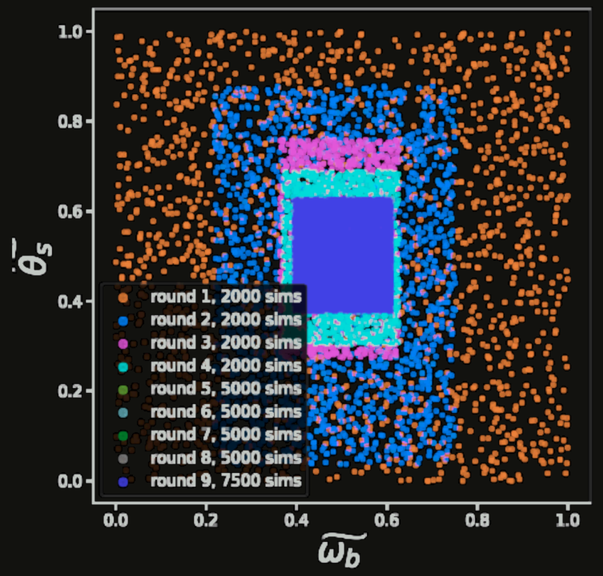
amortized

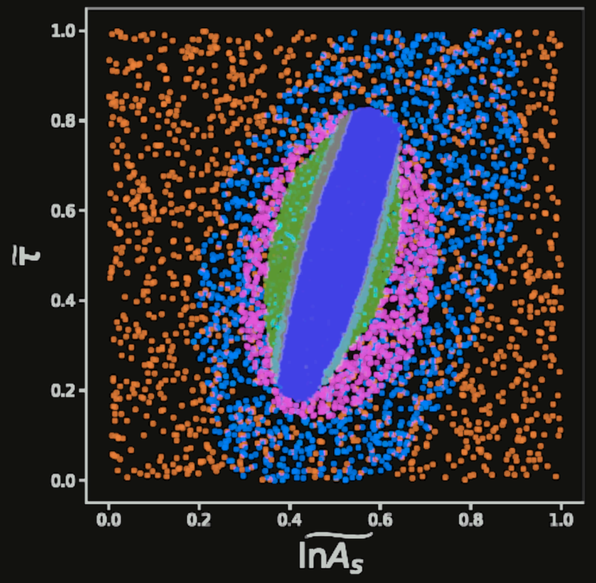
sequential

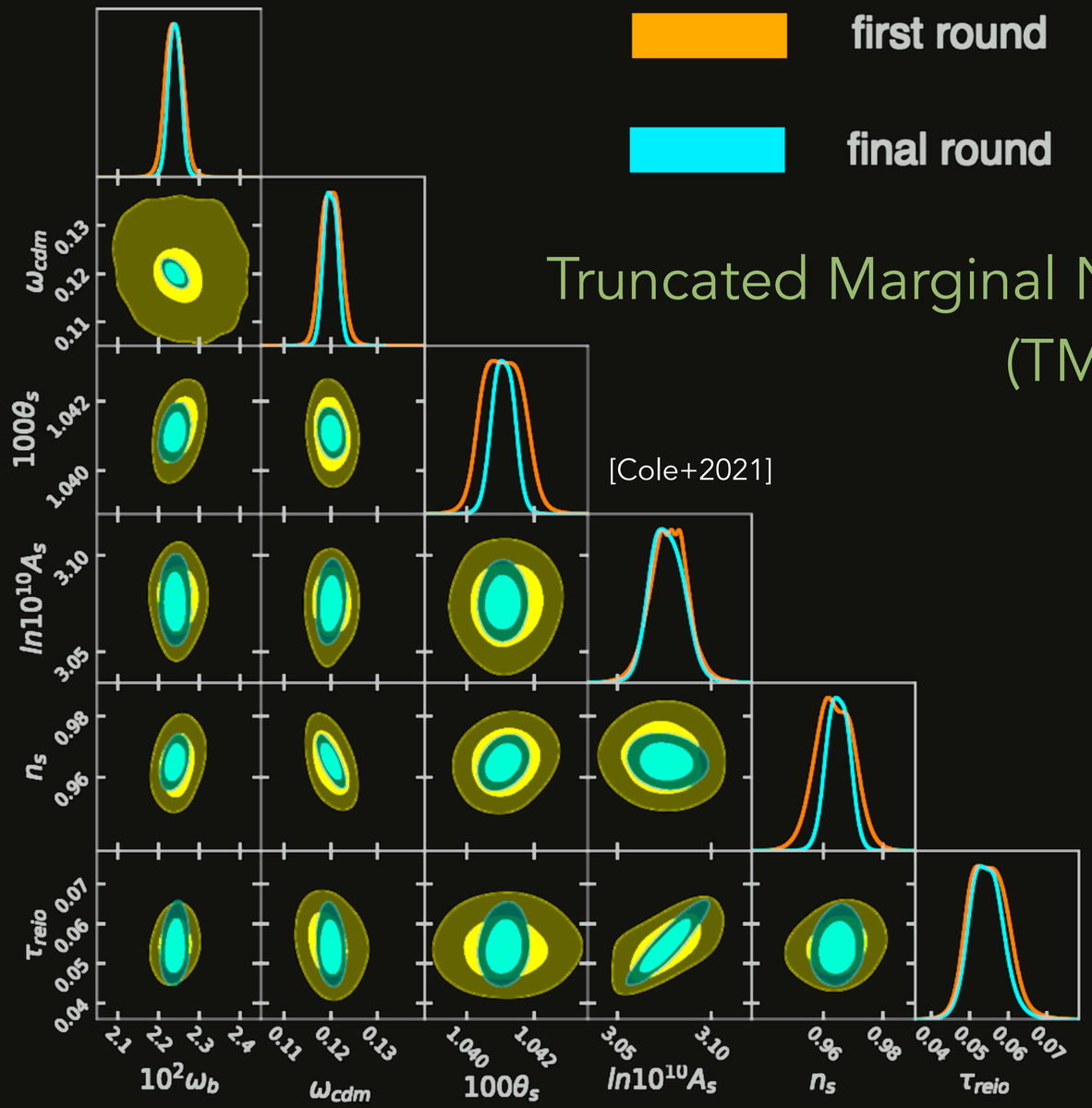
[Zeghal+2024]



Sequential inference



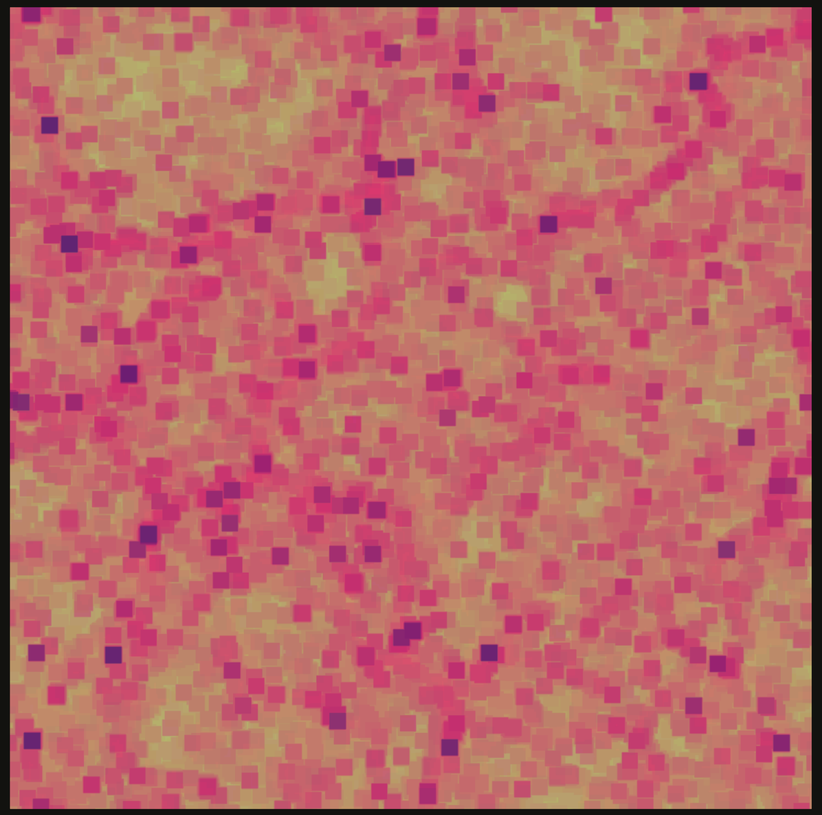




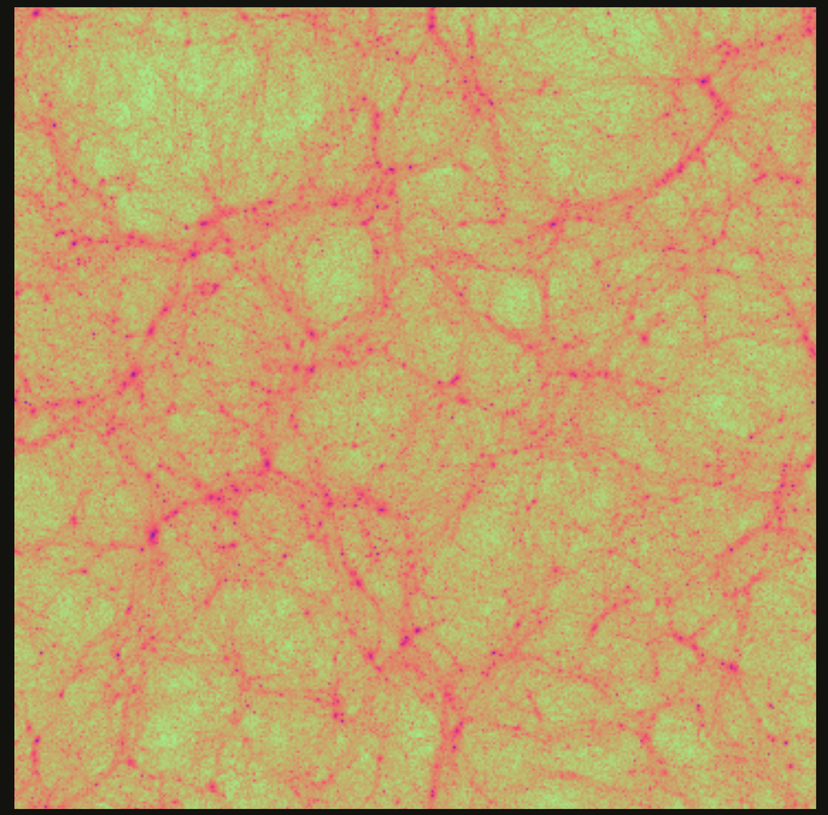
[Cole+2021]

Truncated Marginal Neural Ratio Estimation (TMNRE)

Learning corrections

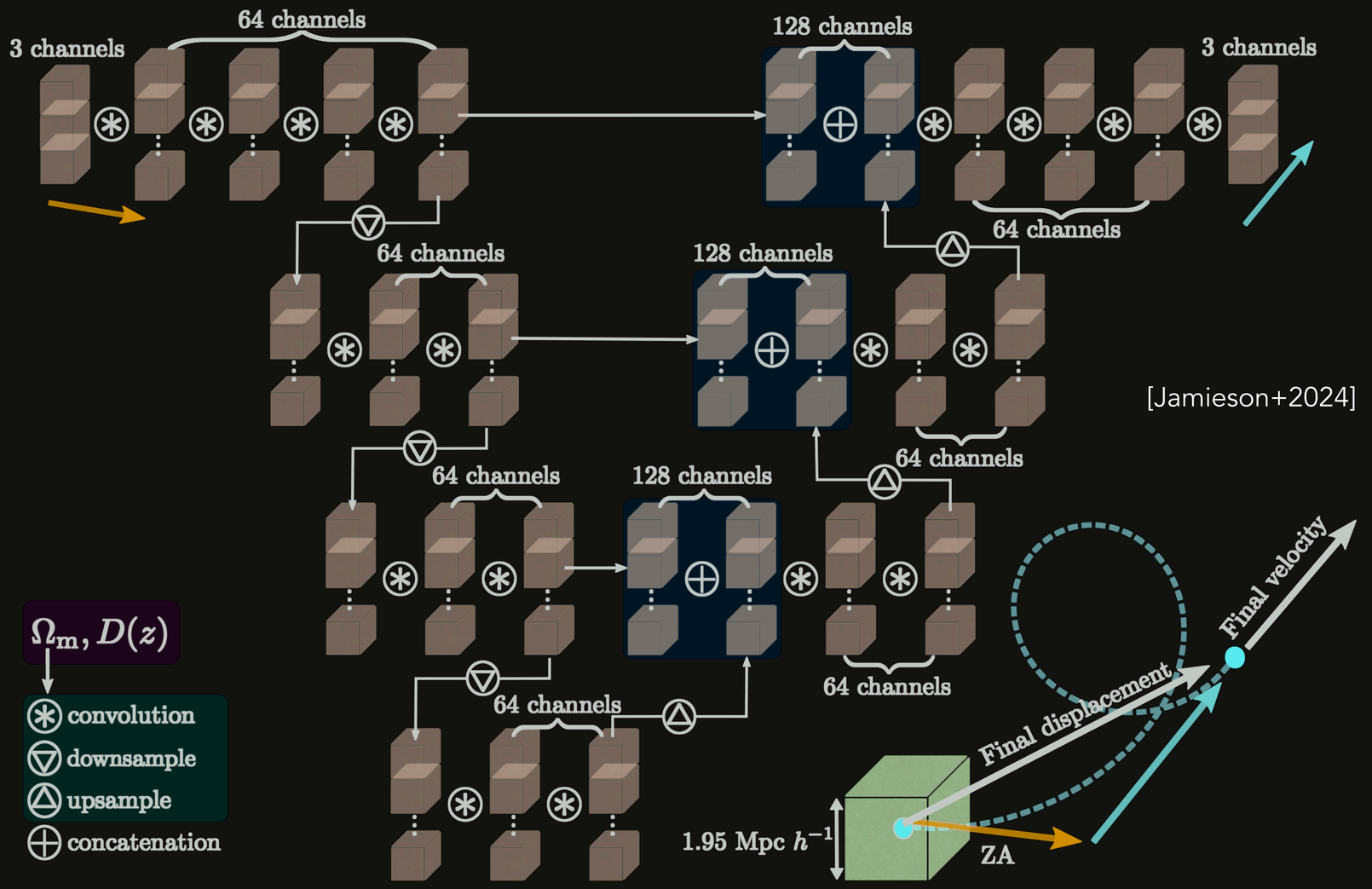


HODs, SAMs, …,
machine learning



Learning corrections

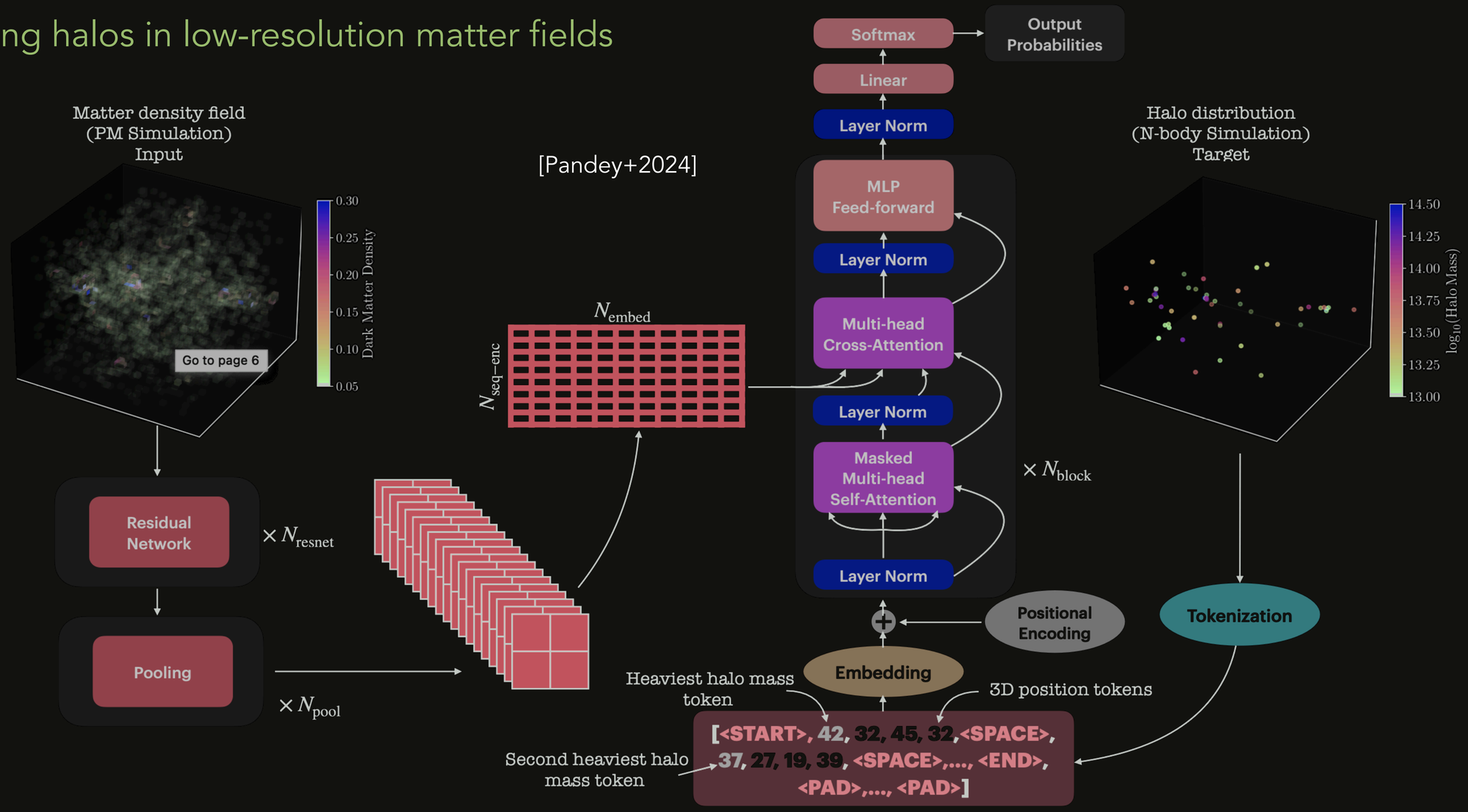
enhancing accuracy of gravity solvers



[Jamieson+2024]

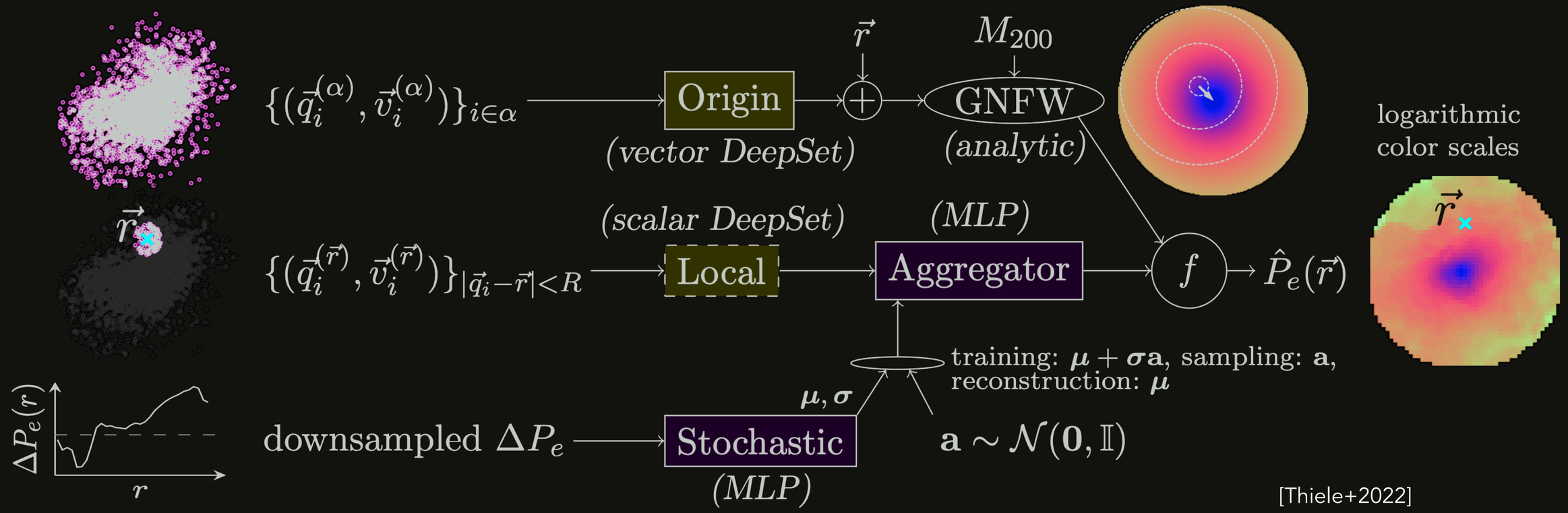
Learning corrections

finding halos in low-resolution matter fields



Learning corrections

painting tSZ on gravity-only halos

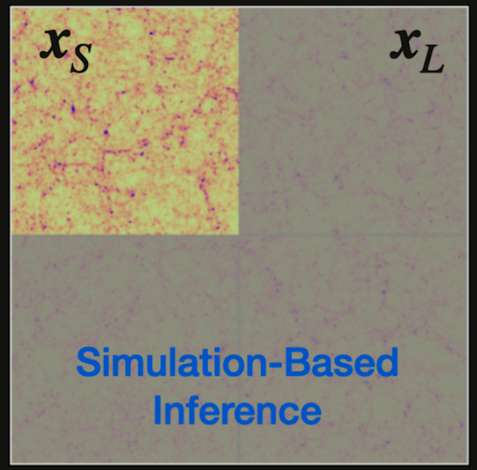


[Thiele+2022]

Hybrid ILI

$$P_L(k) + 2\int \frac{d\mathbf{p}}{(2\pi)^3} \left|F_2(\mathbf{p}, \mathbf{k}-\mathbf{p})\right|^2 P_L(p) P_L(|\mathbf{k}-\mathbf{p}|) + 6\int \frac{d\mathbf{p}}{(2\pi)^3} F_3(\mathbf{p}, -\mathbf{p}, \mathbf{k}) P_L(p) P_L(k) - 2c_s^2 k^2 P_L(k)$$

Perturbation Theory



Large scales are modeled analytically

$$x_L : P(k), B(\{k_i\}), \ldots \rightarrow P(x_L|\theta)$$

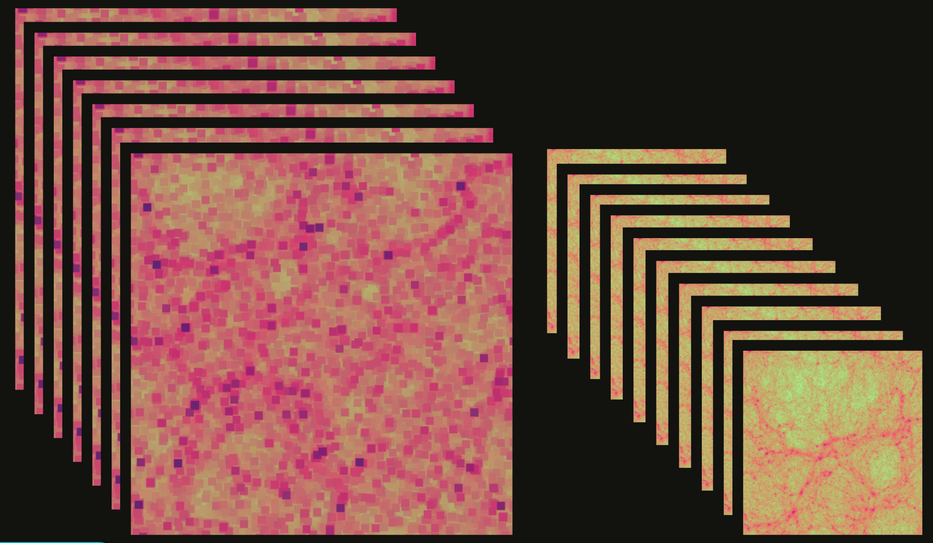
$$P(x_L, x_S|\theta)$$

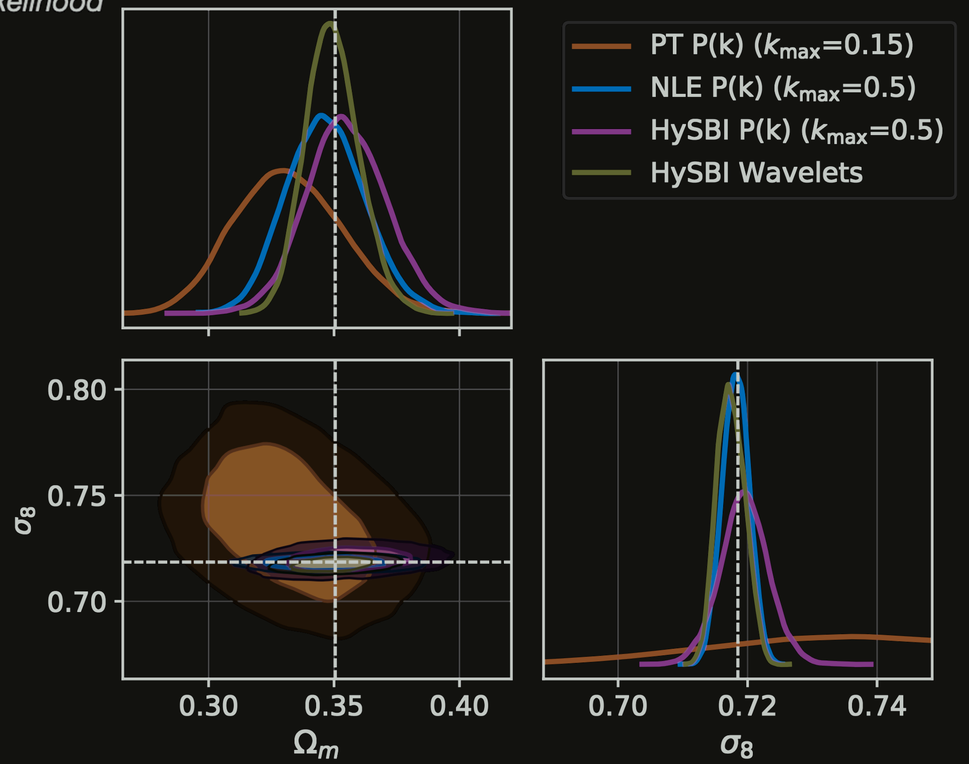
Joint likelihood

$$x_S : P(k), \mathrm{WST}, \ldots \rightarrow P(x_S|x_L, \theta)$$

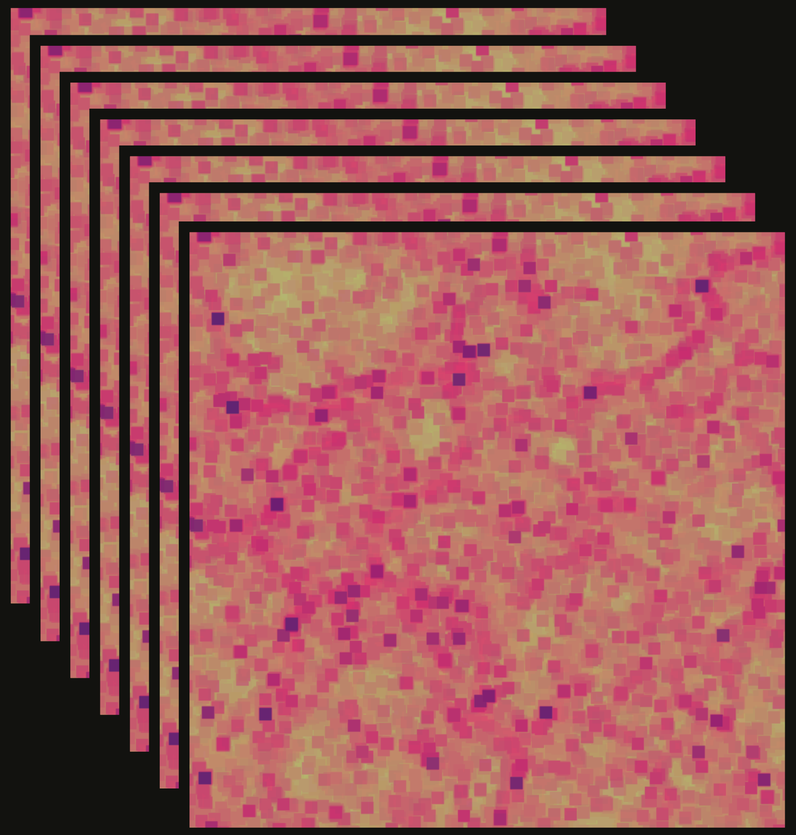
Small scales are modeled numerically

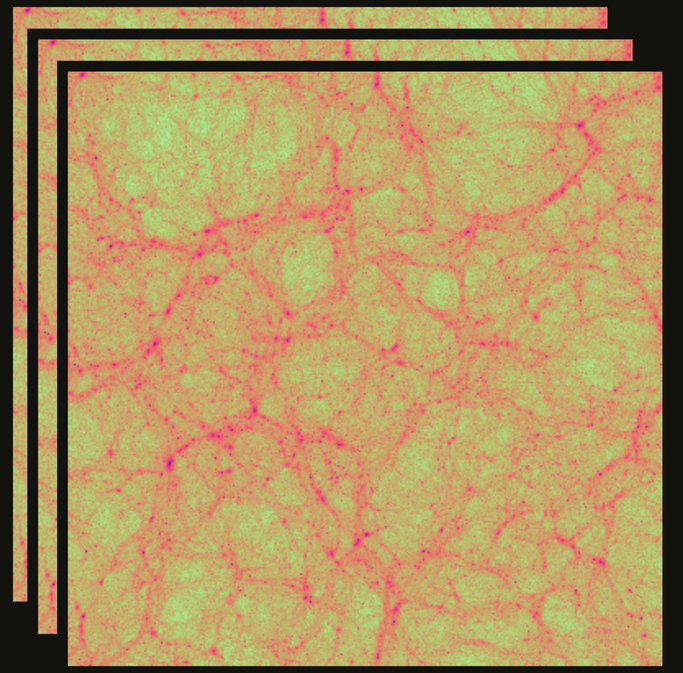
[Modi & Philcox 2023]





Multi-fidelity?



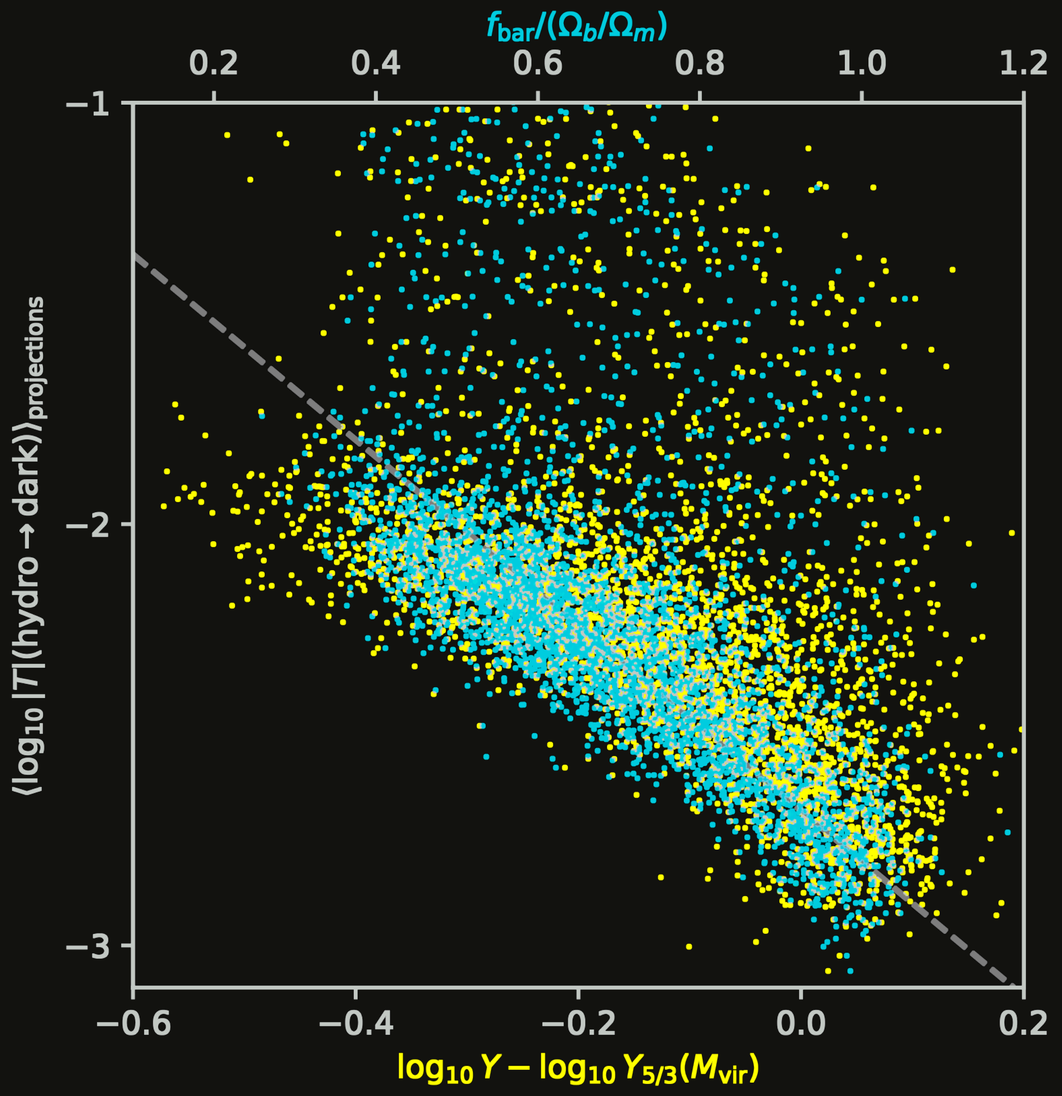


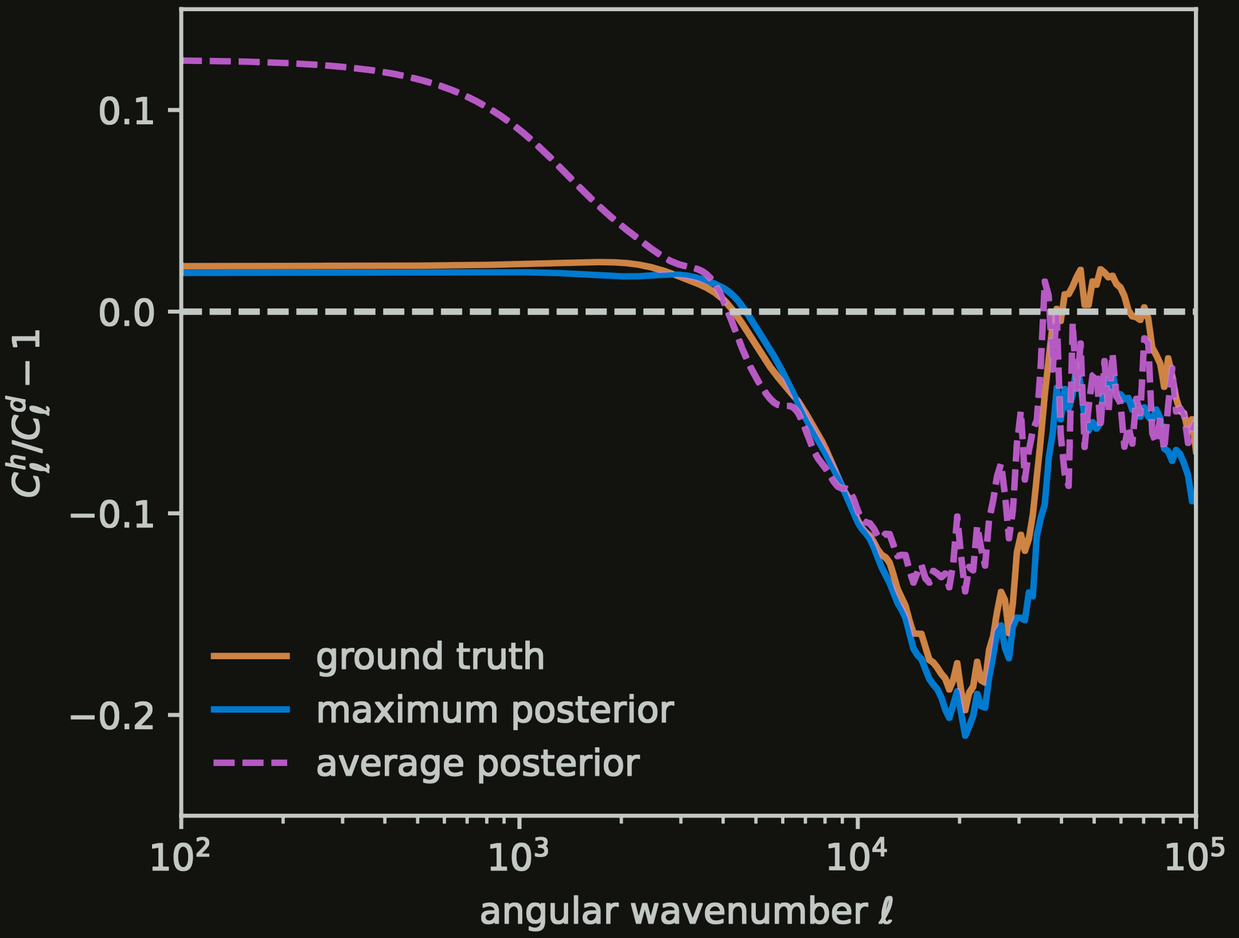
naive approach: transfer learning

maybe better: use correlated simulations
(similar to control variates aka “CARPOOL”)

work in progress

We can simulate everything … except black holes





Summary

- Simulation-based inference has proven viability in simple examples:
  - weak lensing (explicit-likelihood HSC Y1, implicit likelihood)
  - galaxy clustering (SDSS CMASS)
- In order to make it a standard tool, need to increase simulation quality while reducing cost
  - sequential methods
  - learning corrections
  - combine with traditional approaches for large scales
  - multi-fidelity training