Computational Cosmology
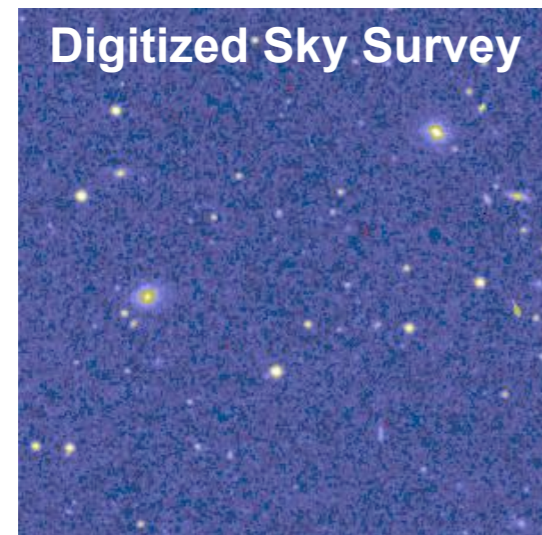
Katrin Heitmann

DOE HEP/ASCR Exascale Requirements Review

June 10, 2015
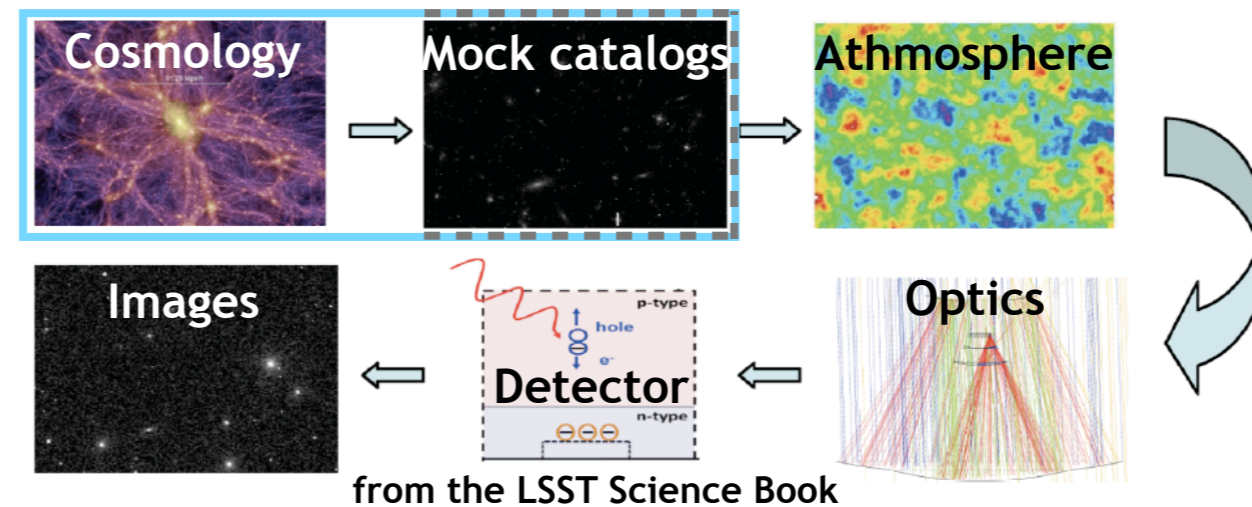
Roles of Cosmological Simulations in DE Survey Science
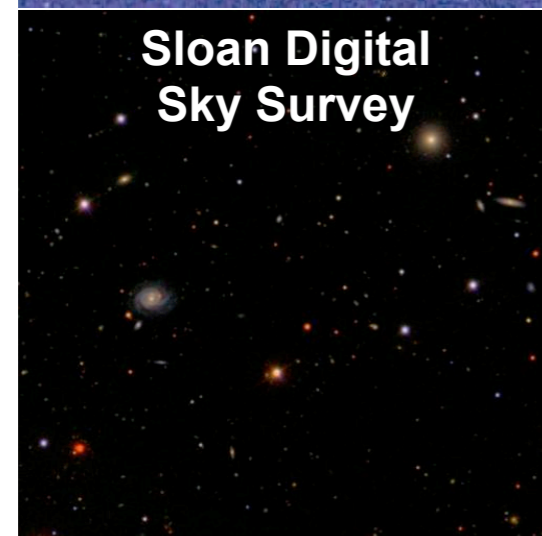
Past

Digitized Sky Survey

Present

Sloan Digital Sky Survey

Future

Deep Lens Survey

LSST

(1) Cosmology simulations and the survey

- First part of end-to-end simulation
- Control of systematics

Cosmology → Mock catalogs → Athmosphere

Images ← Detector ← Optics
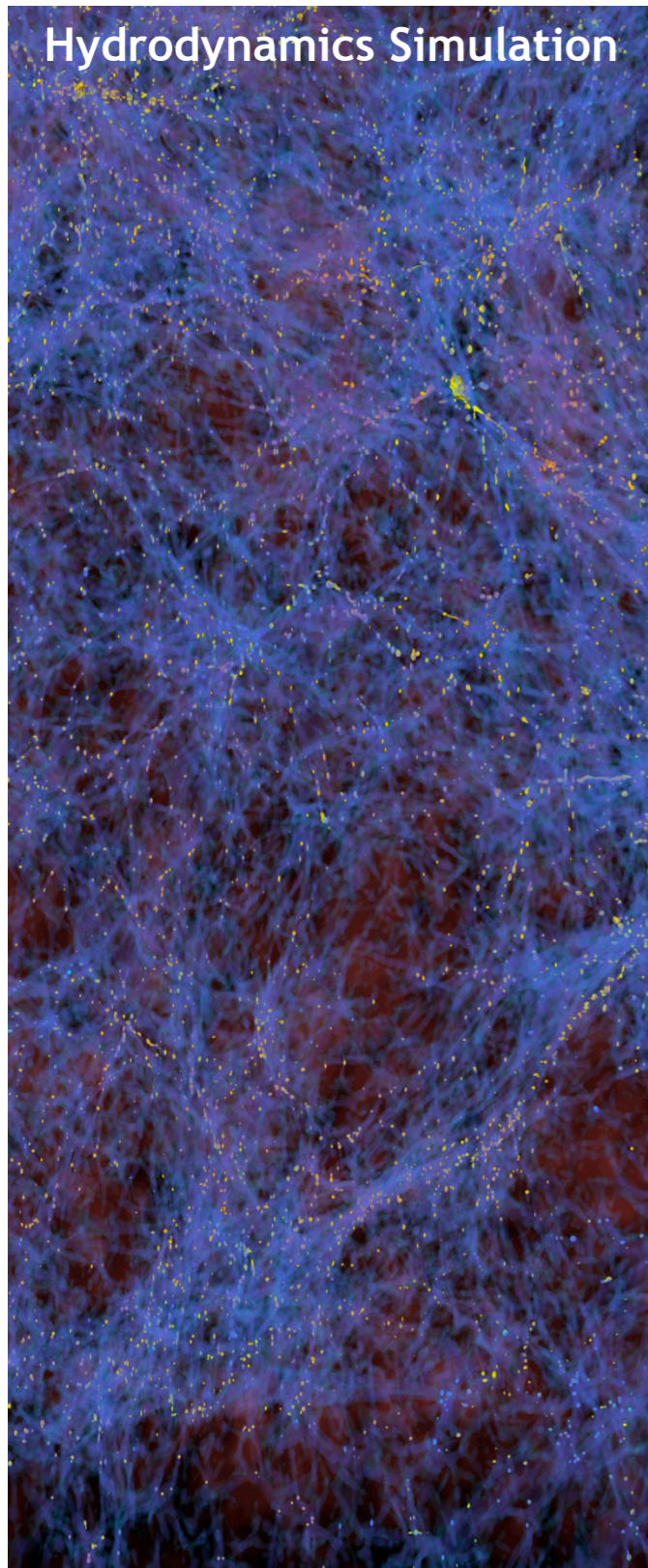
from the LSST Science Book
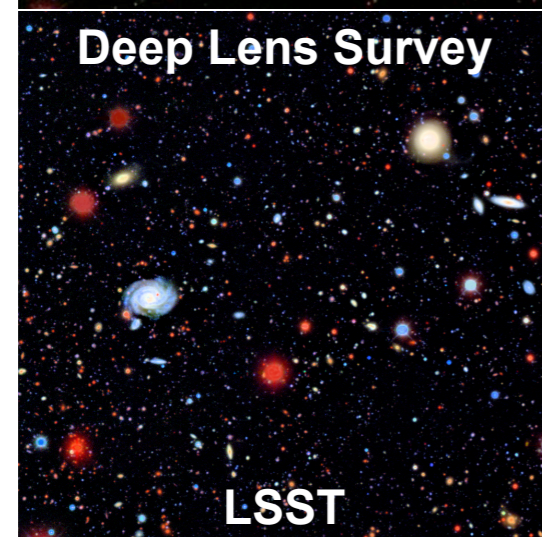
(2) Solving the Inverse Problem

- Exploring fundamental physics
- Fast, very accurate predictions tools (emulators) for physics and observables of interest
- Astrophysical systematics, e.g. baryonic effects
- Predictions for covariances

# Gravity-only and Hydrodynamics Simulations

Hydrodynamics Simulation

- **Gravity-only simulations:**
  - ▶ Gravity dominates on large scales, solve Vlasov-Poisson equation
  - ▶ VPE is 6-D, cannot be solved as PDE, therefore N-body methods, particles as tracers of the dark matter in the Universe
  - ▶ Different algorithms: tree, particle mesh, mixed implementations

- **Hydrodynamics simulations (include gravity solvers):**
  - ▶ Euler fluid equations, increase of cost at least a factor of 10
  - ▶ Subgrid modeling on the small scales, including SNe, AGN, star formation, radiative cooling
  - ▶ Different algorithms (Lagrangian vs. Eulerian): AMR, SPH, Moving mesh
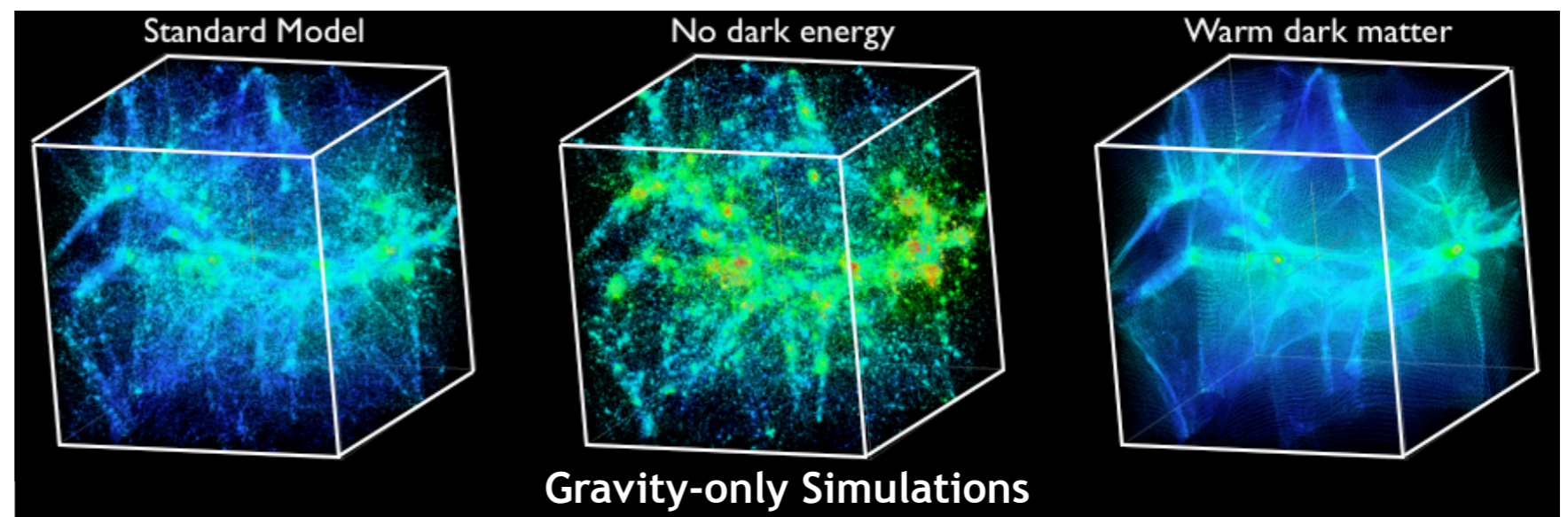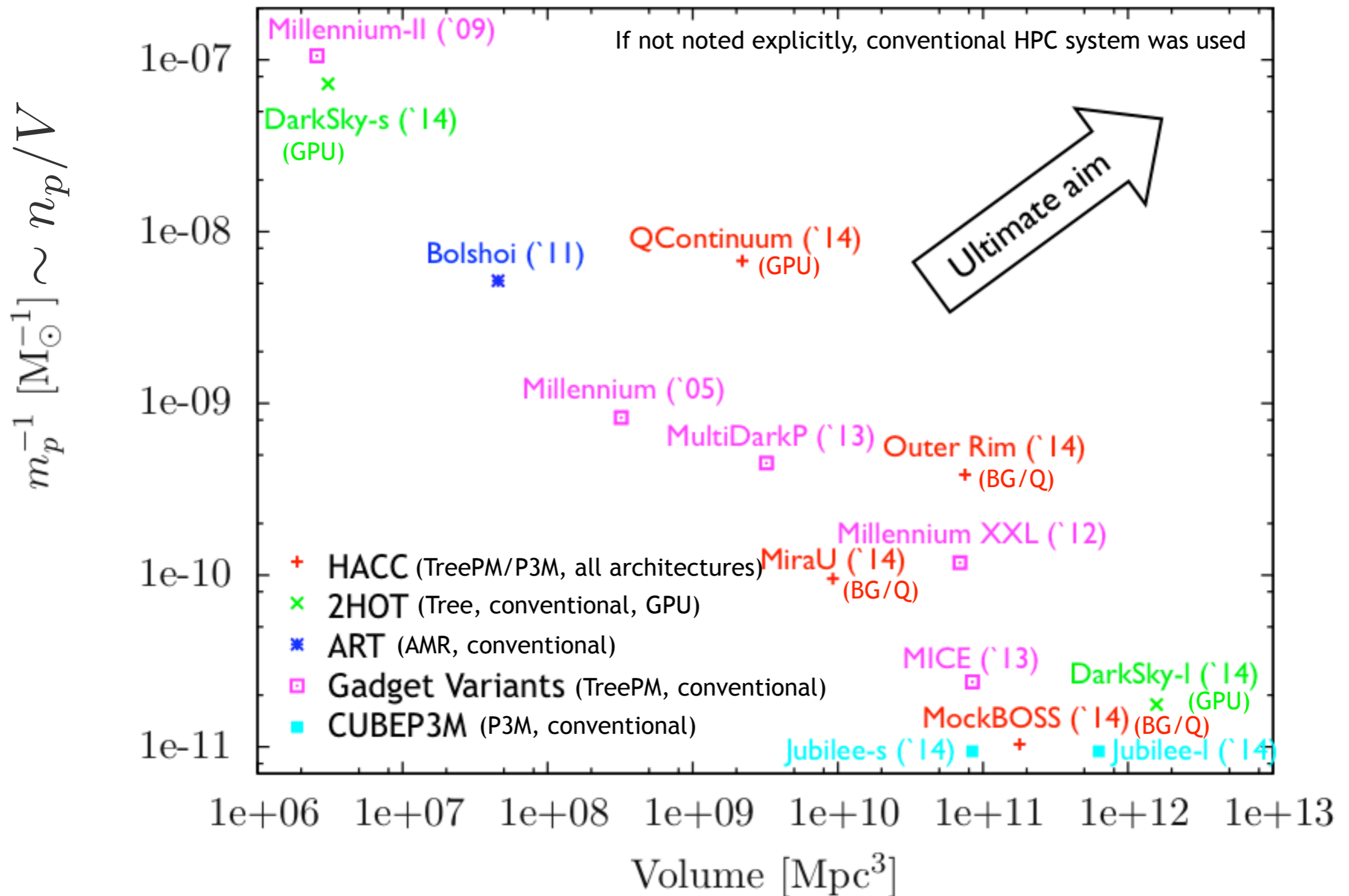
Standard Model   No dark energy   Warm dark matter
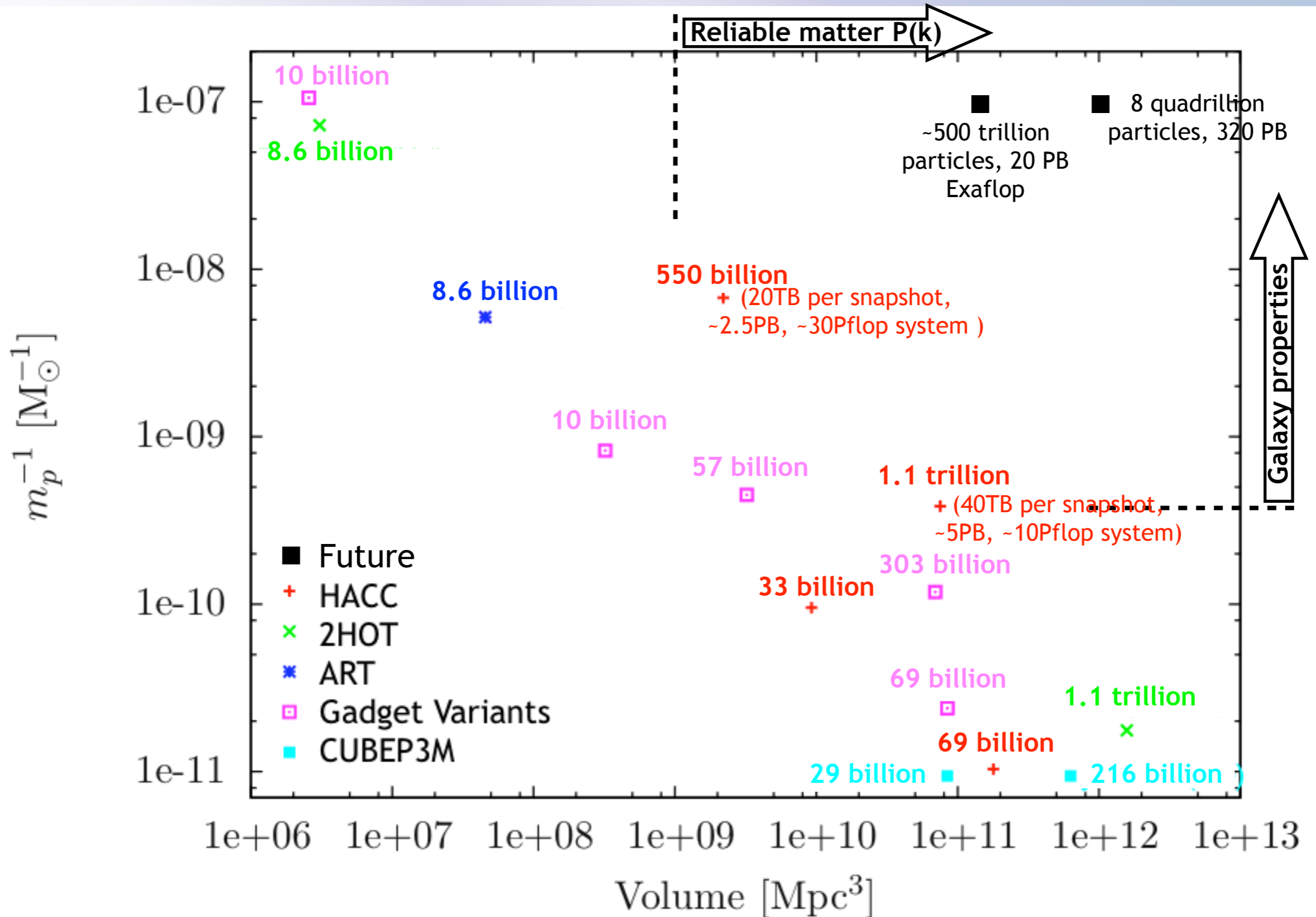
Gravity-only Simulations

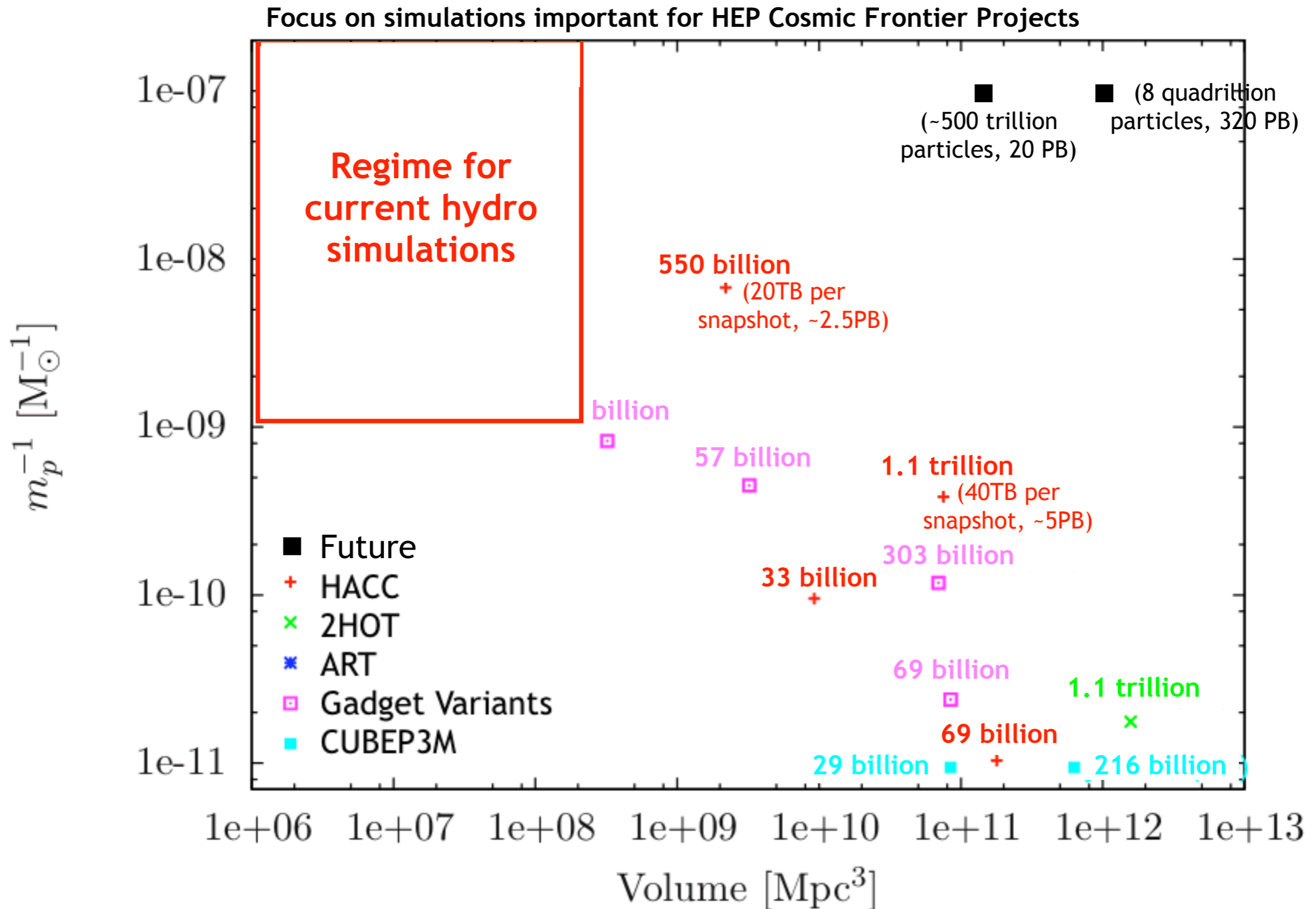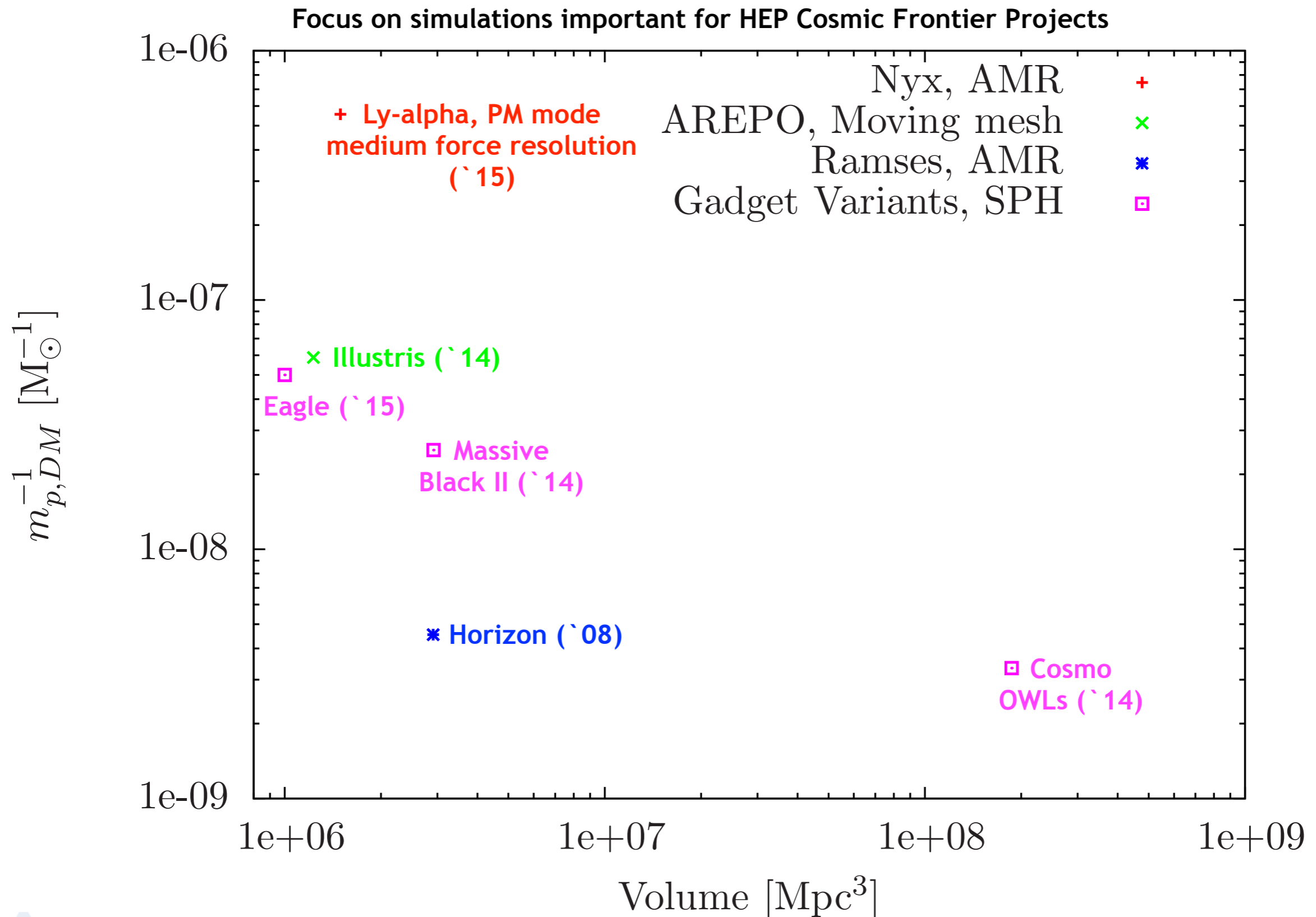# Current State-of-the-Art Simulations, Gravity Only

Current State-of-the-Art Simulations, Gravity Only
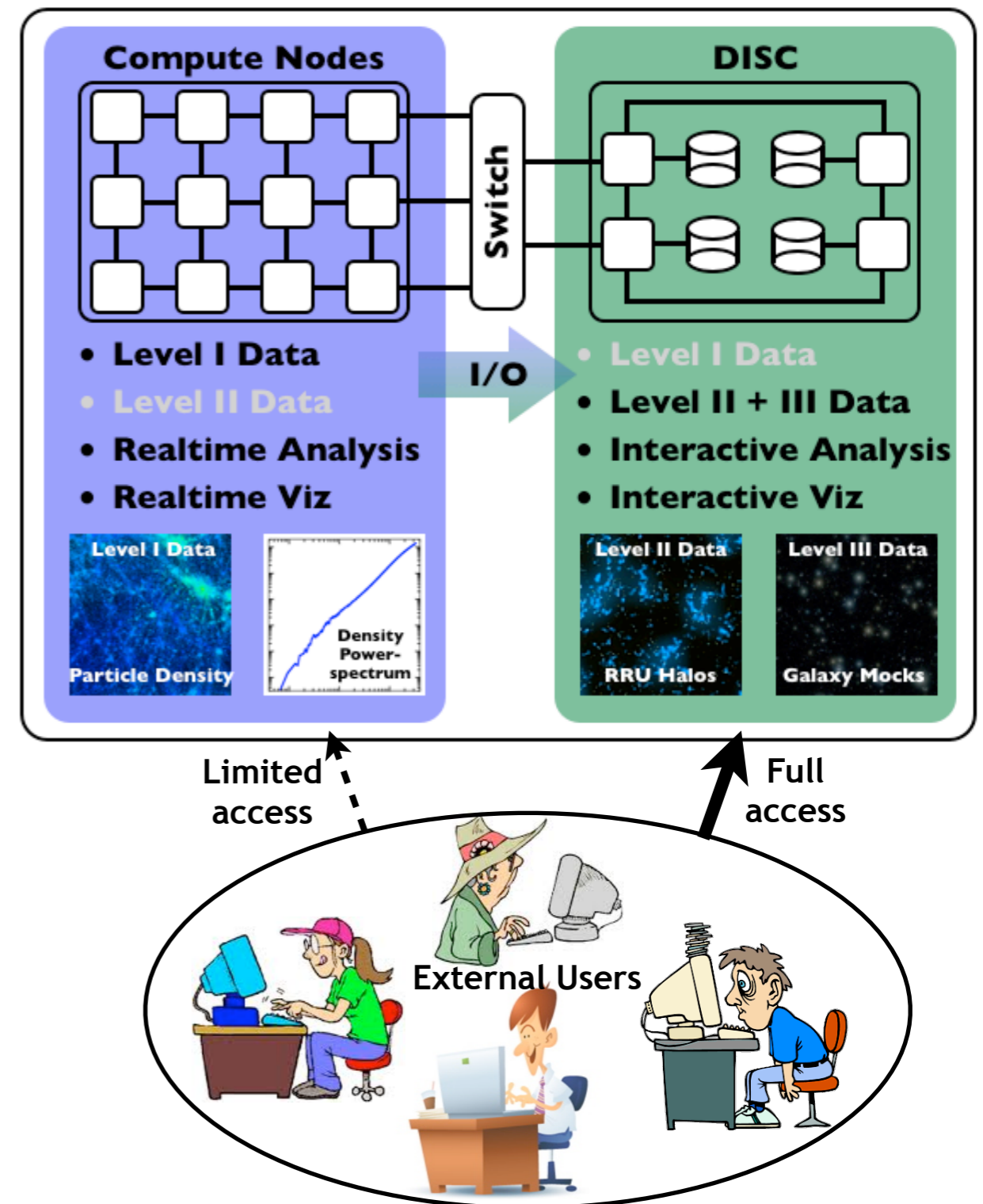
Current State-of-the-Art Hydro Simulations

Focus on simulations important for HEP Cosmic Frontier Projects

Current State-of-the-Art Hydro Simulations

Focus on simulations important for HEP Cosmic Frontier Projects

Analysis Challenges

- **Currently simulation data generation is constrained only by storage and I/O bandwidth, ~PB datasets already available**

  - *In situ* analysis: Large-scale analysis tasks on the compute platform that require access to raw data (Level I), only feasible if analysis task is well load-balanced

  - *In situ*/co-scheduling: Reduce raw data *in situ* to Level II data; co-schedule analysis of Level II data products

  - Post-processing: Analysis of Level II data to generate Level III data

- **How can we efficiently share data?**

  - Simulation campaigns are carried out at very few places (supercomputer centers)

  - Outputs are very science rich, many people can contribute to the analysis

  - Moving raw data is impractical, analysis often takes a lot of computing power

  - Need for making data *and* analysis opportunity available to the community

Compute Nodes — Switch — **DISC**

- Level I Data
- Level II Data
- Realtime Analysis
- Realtime Viz

I/O

- Level I Data
- Level II + III Data
- Interactive Analysis
- Interactive Viz

Level I Data / Particle Density — Density Power-spectrum

Level II Data / RRU Halos — Level III Data / Galaxy Mocks

Limited access — Full access

External Users

Summary and Outlook

- **Gravity-only simulations:**
  - ▶ Very important for survey science (synthetic skies), exploration of fundamental physics (dark energy, dark matter, neutrinos)
  - ▶ Handful of codes scale well on current LCF machines, will continue to scale in the future; other codes still rely on "standard hardware" (e.g., Gadget-2, publicly available)
  - ▶ Example of future code capability: Adaptation to new architectures, scaling to full machine, *in situ*/co-scheduling analysis capabilities (HACC has this now)

- **Hydrodynamics simulations:**
  - ▶ Crucial for understanding systematics in up-coming surveys, in particular on small scales
  - ▶ Work on porting different codes to future architectures ongoing
  - ▶ Uncertainties in subgrid model physics needs to be reduced (better data?)

- **Analysis Challenges:**
  - ▶ Data outputs are becoming very large, can only partially be stored
  - ▶ *In situ* data analysis capabilities can help, but simulations are very science rich, difficult to anticipate all possible analysis tasks in advance, not all analysis tools scale well
  - ▶ Data and tool sharing