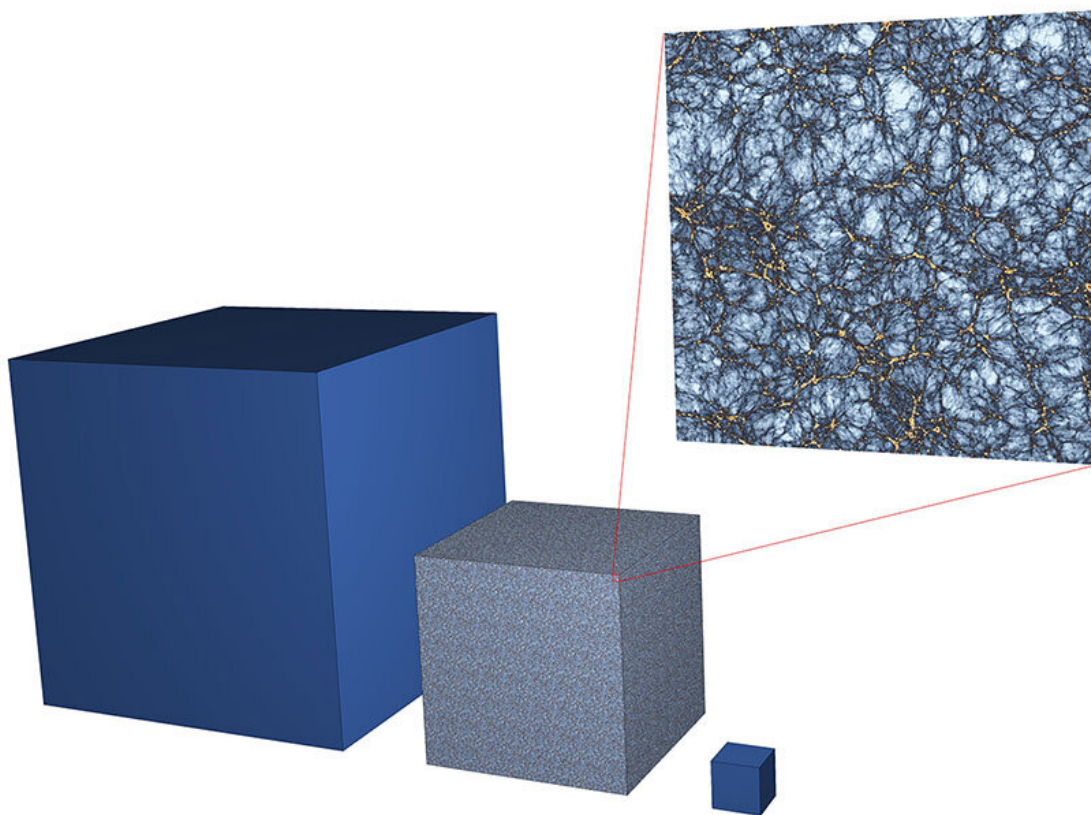# Researchers confirm accuracy of cosmological data analysis technique using mock data

March 19 2021

The volume of the challenge data prepared by the Japanese team (center), compared to the total volume of the real universe observable by light (left), and the volume of one of the largest galaxy survey catalogs, the Sloan Digital Sky Survey (right). From left to right, the volumes would be equivalent to cubes with edges of 75 billion light-years, 40 billion light-years, and 9 billion light-years.

The marbling in the simulation shows areas of high (red) and low (blue) density. Credit: Takahiro Nishimichi

Astronomers have played a game of guess-the-numbers with cosmological implications. Working from a mock catalog of galaxies prepared by a Japanese team, two American teams correctly guessed the cosmological parameters used to generate the catalog to within 1% accuracy. This gives us confidence that their methods will be able to determine the correct parameters of the real universe when applied to observational data.

The basic equations governing the evolution of the [universe](#) can be derived from theoretical calculations, but some of the numbers in those equations, the cosmological parameters, can only be derived through observations. Cosmological parameters tied to the unobservable parts of the universe, like the amount of dark matter or the expansion of the universe driven by [dark energy](#), must be inferred by looking at their effects on the distribution of visible galaxies. There is always uncertainty when working with the dark part of the universe, and it is hard to be sure that the models and data analysis are accurate.

To test the [data analysis](#), a Japanese team led by Takahiro Nishimichi at Kyoto University and the Kavli IPMU(note) at the University of Tokyo used the ATERUI II supercomputer at the National Astronomical Observatory of Japan to create 10 mock universes with a total volume 100 times greater than even the most extensive galaxy surveys so far. The large volume, large dynamical range, and high resolution achievable only with the world's most powerful supercomputer dedicated to astronomy were needed to separate systematic errors in the analysis models from random errors due to meaningless coincidences in the data. The [cosmological parameters](#) used to evolve these mock universes were

chosen randomly from the range of reasonably expected values. The Japanese team prepared a [catalog](#) listing the positions of the [galaxies](#) in the simulation similar to the catalogs produced by real telescopes observing the heavens. The Japanese team then challenged other astronomers to guess the numbers used to generate the catalog.

Two American teams accepted the challenge. Working independently and using different methods, both teams analyzed the Japanese data with tools used for real astronomy surveys. Each team had only one chance to guess the numbers, and both teams produced answers within 1% of the real values. This shows that these methods should give correct results when applied to real [observational data](#).

So what were the correct numbers? They're still secret so that more teams can play guess-the-numbers. In this way the challenge data will continue to support the development and testing of cosmic analysis techniques.

These results appeared as Nishimichi et al. "Blinded challenge for precision cosmology with large-scale structure: results from effective field theory for the redshift-space galaxy power spectrum" in *Physical Review D* on December 28, 2020.

Citation: Researchers confirm accuracy of cosmological data analysis technique using mock data