Multi-Fidelity Emulation for Matter Power Spectrum and Lyα Forest

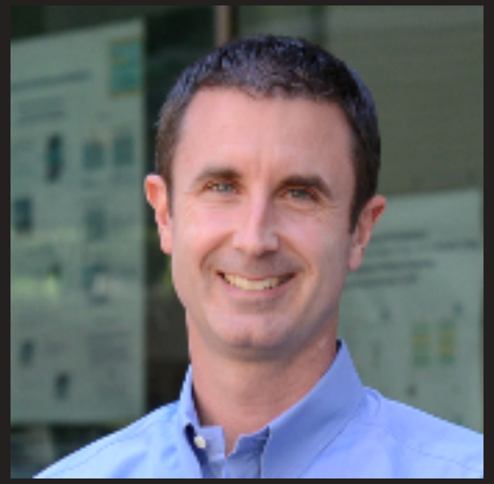
… using Astrid simulation code

Ming-Feng Ho
UC Riverside
NASA FINESST FI

1. Ho, Bird, Shelton (2022) arXiv:2105.01081
2. Fernandez, Ho, Bird (2022) arXiv:2207.06445
3. Ho, Bird, Fernandez, Shelton (in prep)

Simeon Bird (UCR Astro)

Martin A. Fernandez (UCR Astro)



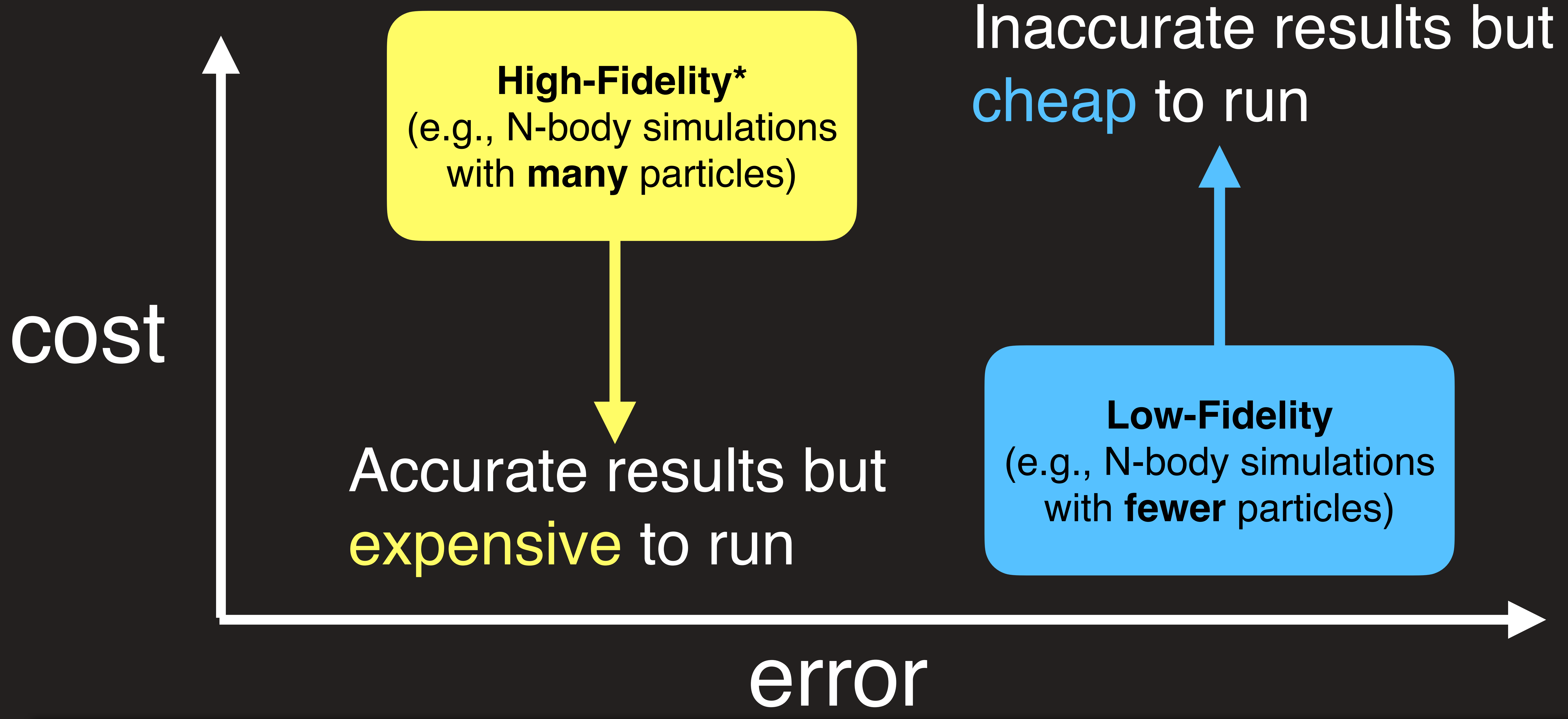
Christian Shelton (UCR CS)

December 2022 | New York City NY

Outline

- What's multi-fidelity emulation?
- Example 1: Matter power spectrum using DM only simulations with different number of particles + box sizes
- Example 2: Lyman alpha 1D flux power using Astrid simulations with different number of particles

What's multi-fidelity emulation?
The trade-off between speed and accuracy



*Could be anything you think is high-fidelity

Idea: Many LF + A few HF
= minimize the cost and maximize the accuracy.

Illustration credit: adapted from Perherstorfer et al (2018)
Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization

What's multi-fidelity emulation?
An analogy: University's hiring

High-Fidelity
Professors

Less-accurate results
but inexpensive to hire

cost

Accurate results but
expensive to hire

Low-Fidelity
Grad Student
Researchers

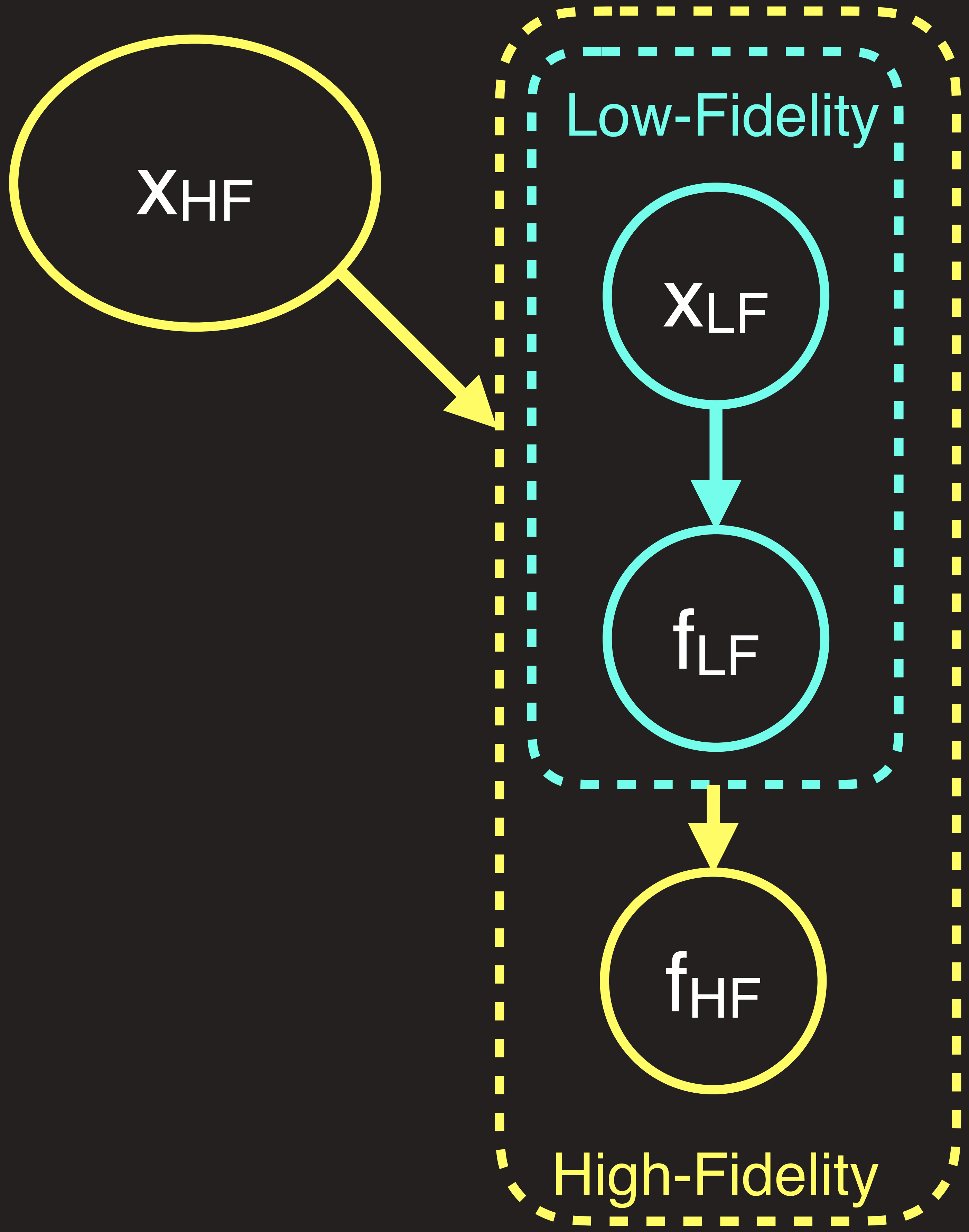
error

Idea: Many Grad Students + A few Professors
= minimize the cost and maximize the accuracy.

image credit: PHD comics

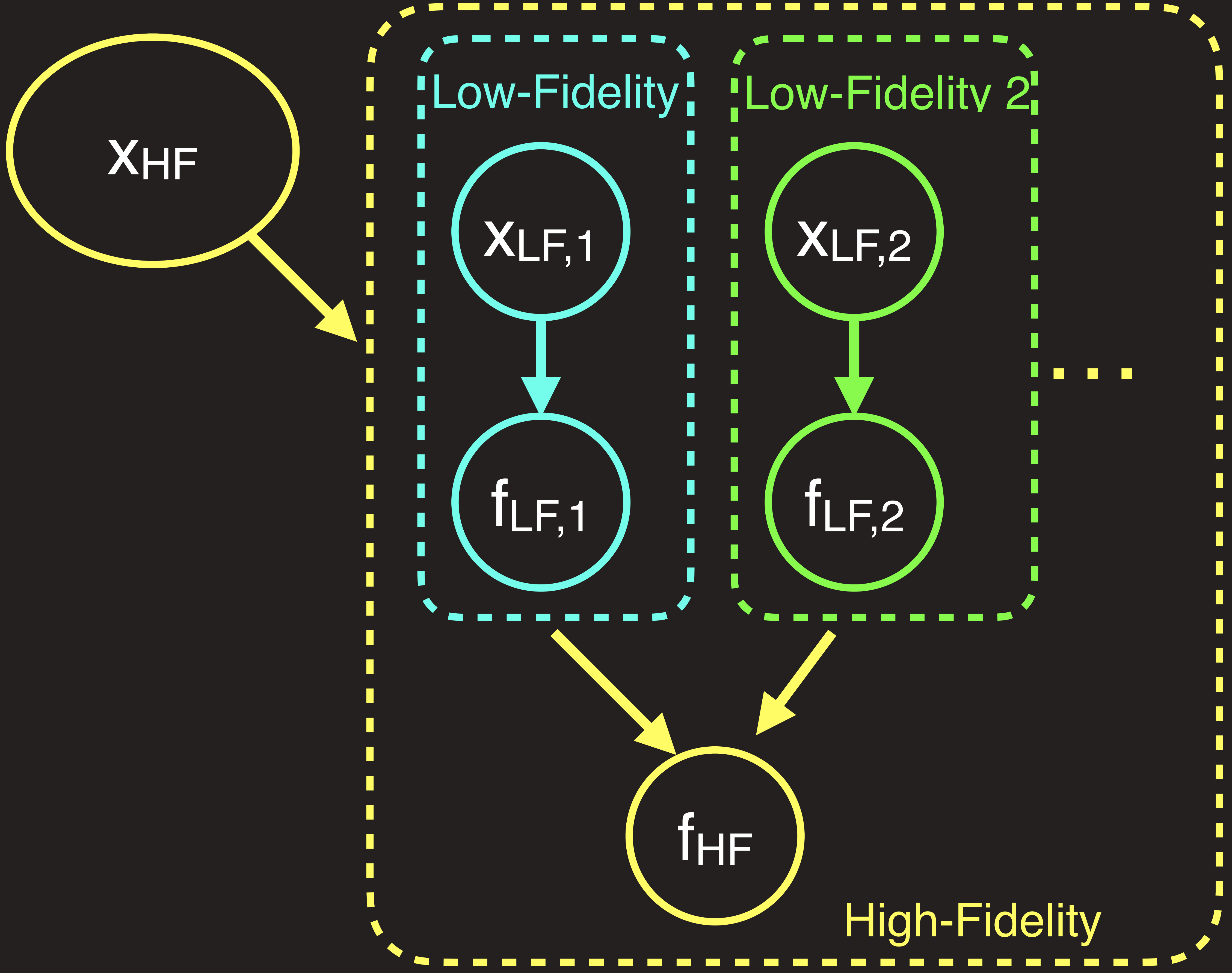
Multi-fidelity emulators: Transfer learn the information from low-fidelity

Kennedy & O’Hagan (2000)



Deep GP: $f_{\mathrm{HF}}(x) = \rho(x, f_{\mathrm{LF}}(x)) + \delta(x).$

Graphical model GP (2021)



$$f_{\mathrm{HF}}(x) = \rho(\{f_{\mathrm{LF},1}, f_{\mathrm{LF},2}\} \cup x) + \delta(x)$$

Multi-fidelity emulators:
Transfer learn the the simulations from different resolutions + box sizes

Low-fidelity:

128^3
256 Mpc/h

High-fidelity:

512^3
256 Mpc/h

Deep GP: $f_{\rm HF}(x) = \rho(x, f_{\rm LF}(x)) + \delta(x).$

Low-fidelity:

128^3
256 Mpc/h

Low-fidelity 2

128^3
100 Mpc/h

High-fidelity:

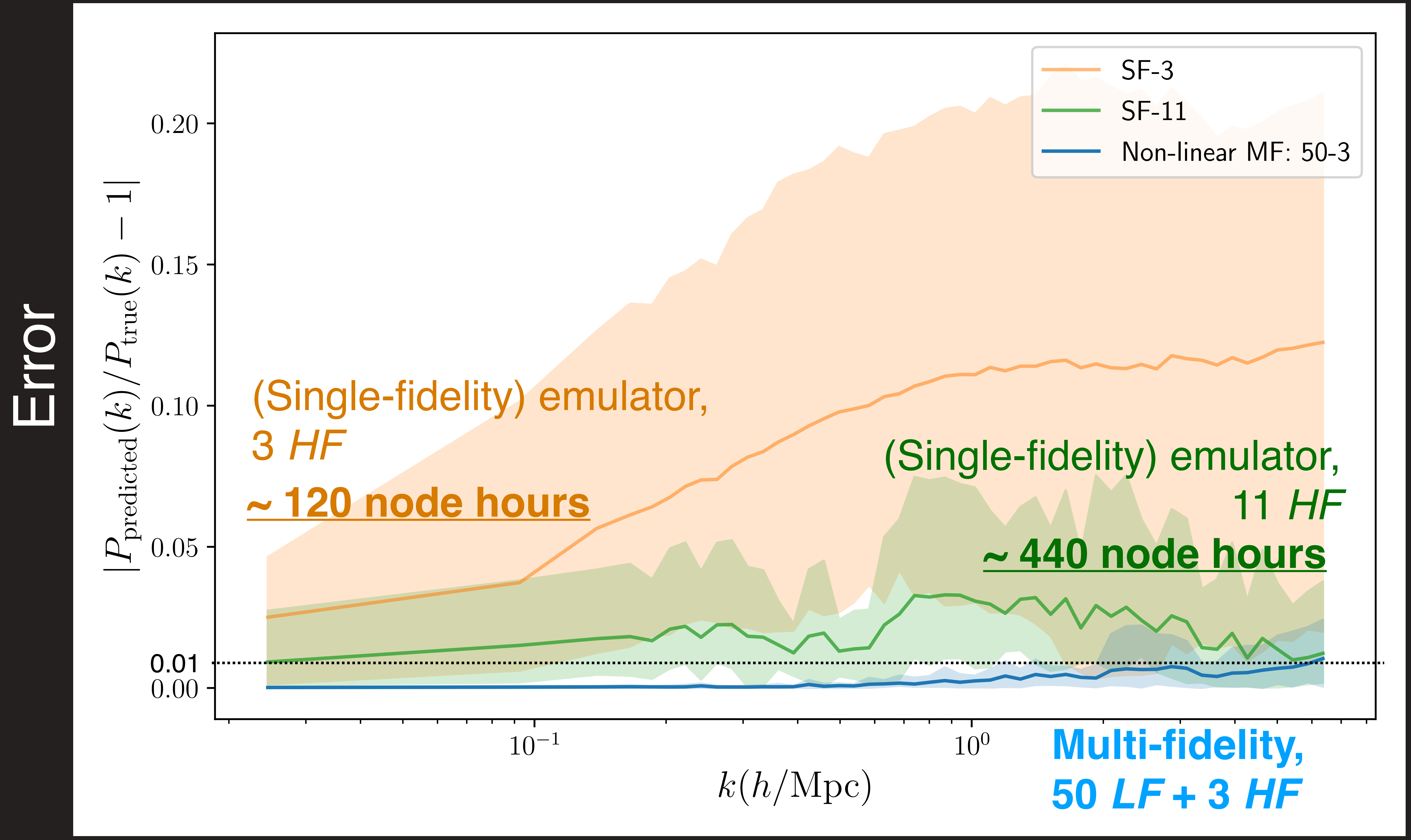
512^3
256 Mpc/h

$$f_{\rm HF}(x) = \rho(\{f_{\rm LF,1}, f_{\rm LF,2}\} \cup x) + \delta(x)$$

Example 1: matter power spectrum
Experimental design using DM-only simulations

- Parameters: $(h, \Omega_0, \Omega_b, A_s, n_s)$
- 50 *Low-fidelity*: space-filling strategy (Latin hypercube)
  - $128^3, 256\,\mathrm{Mpc\,h^{-1}}$
- 3 *High-fidelity*: a subset of low-fidelity runs
  - $512^3, 256\,\mathrm{Mpc\,h^{-1}}$
- *HF* choices were optimized using *LF* suite as a prior

z = 0

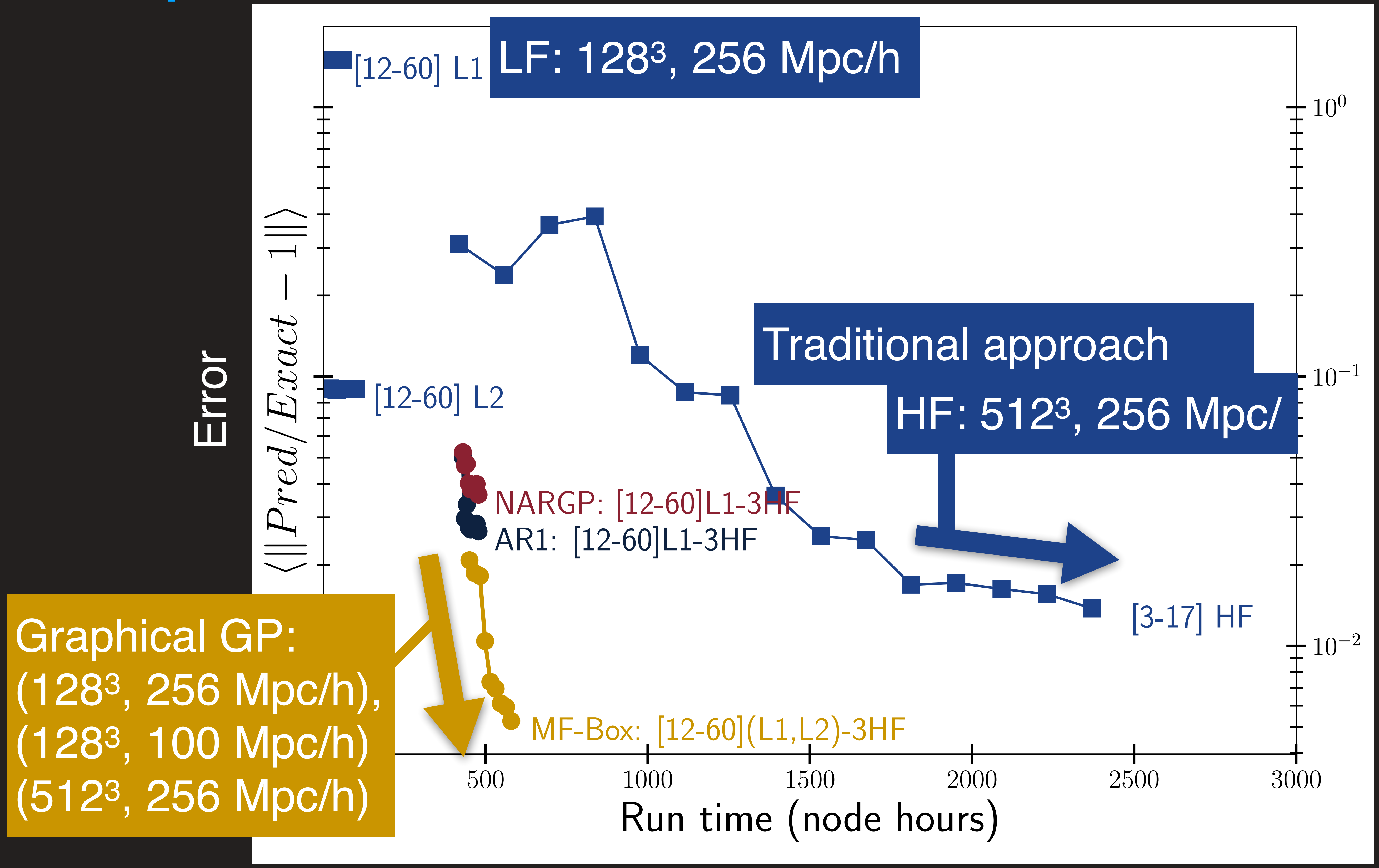


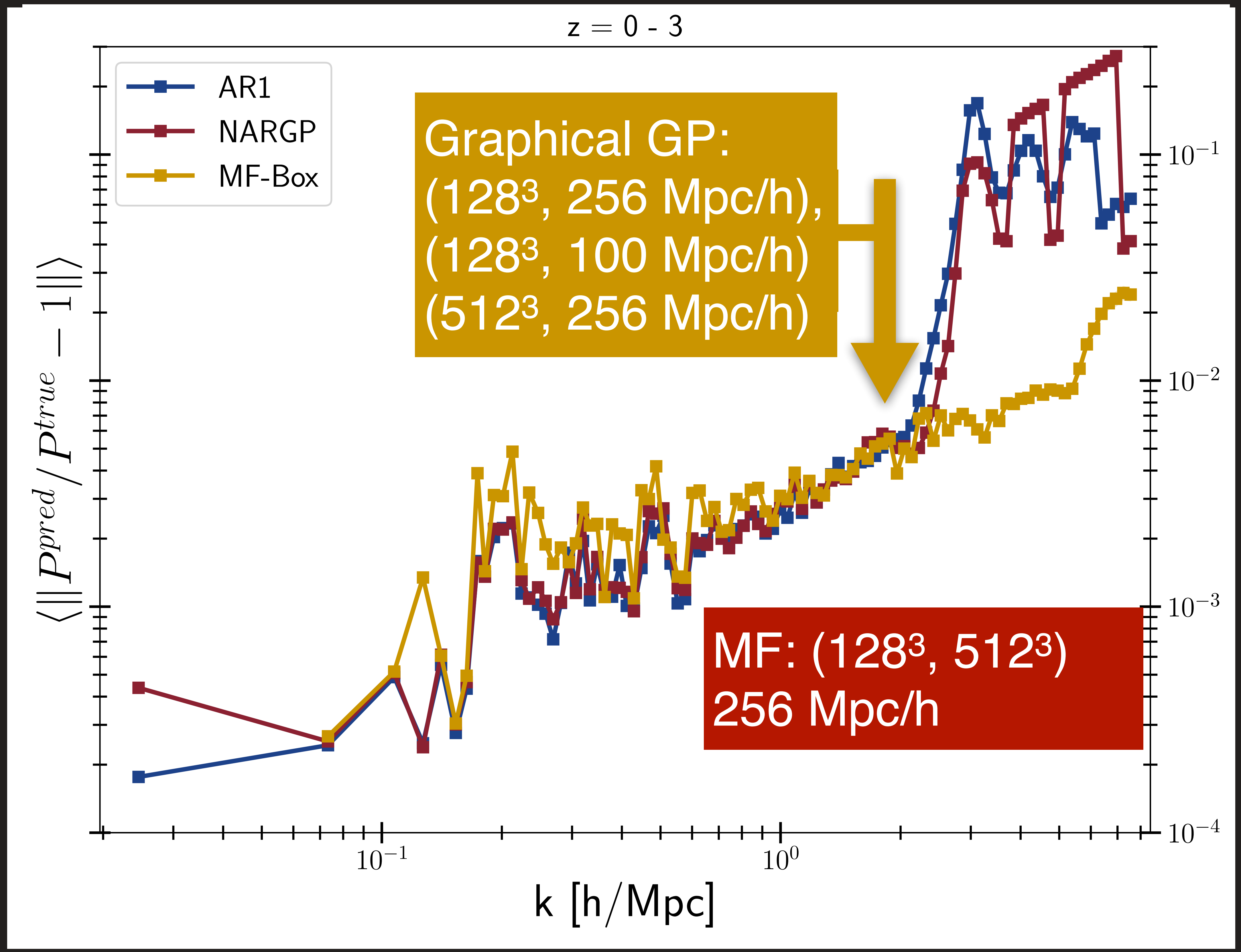
Multi-fidelity, 50 *LF* + 3 *HF*

~140 node hours

Ho, Bird, Shelton (2022)

Example 1: matter power spectrum
Graphical GP with simulations from different box sizes

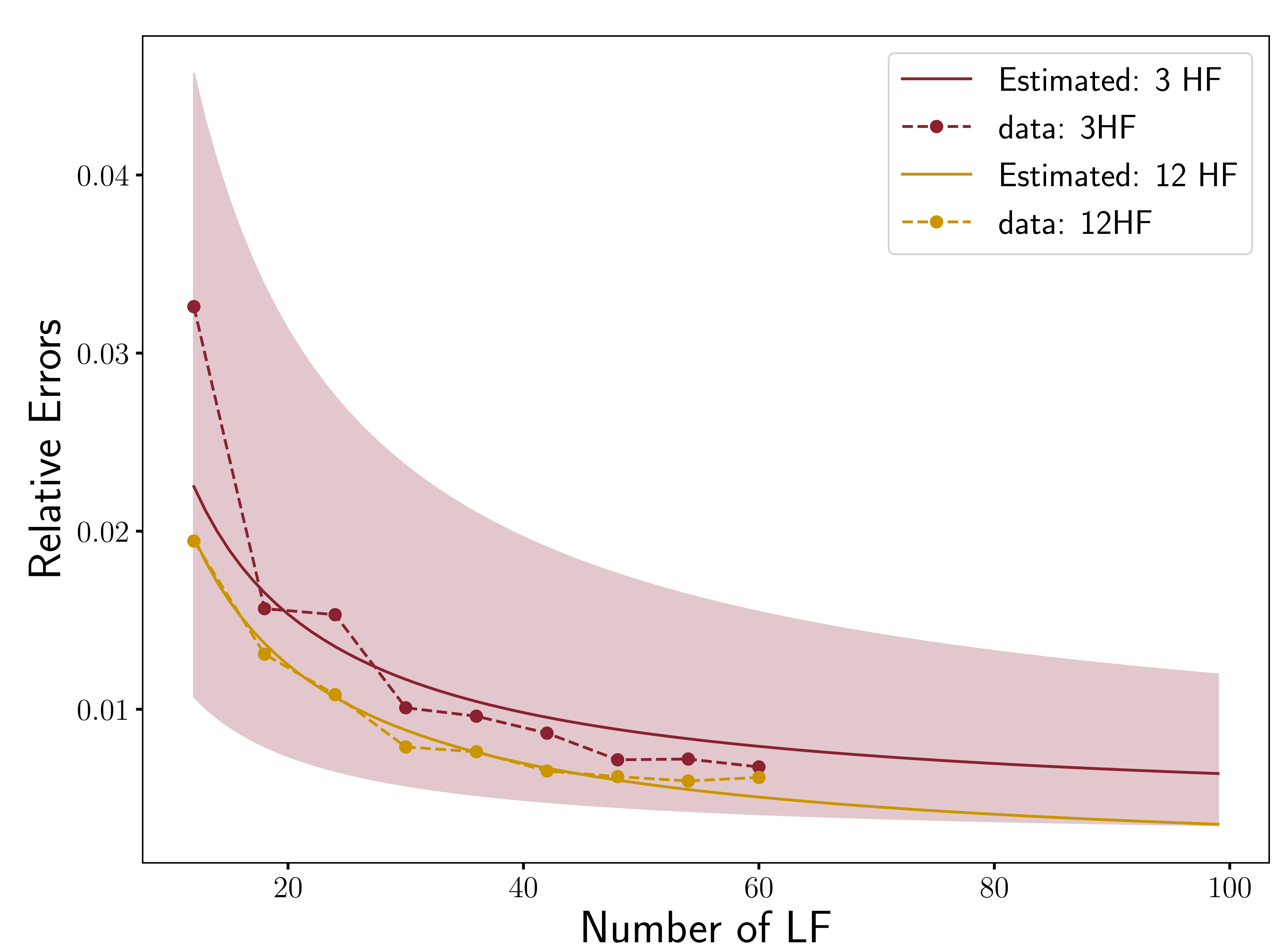




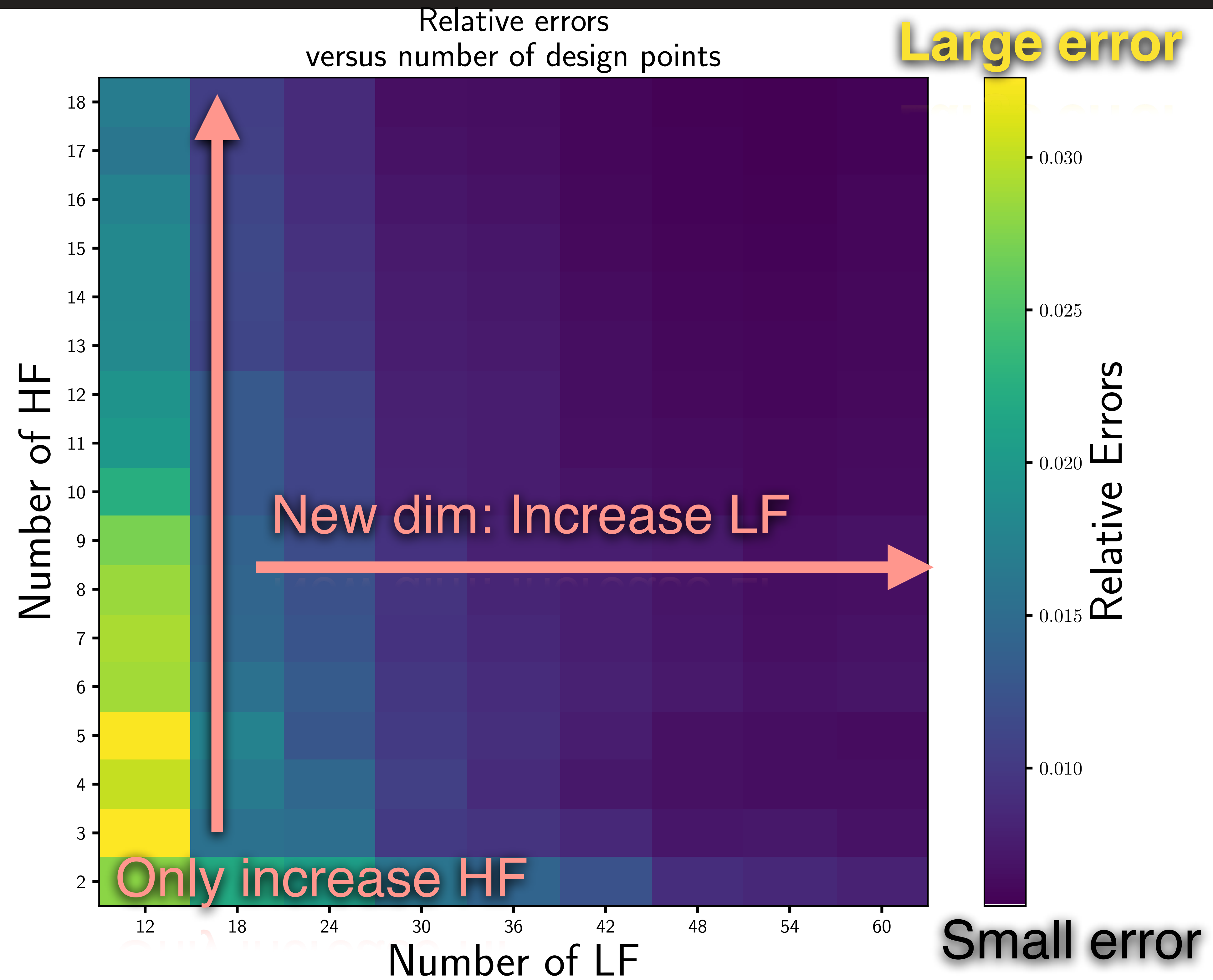
- MF approach economically achieves sub-percent error, outperforming traditional single-fidelity emulator.
- Graphical-GP bridges the information from different box sizes.

Error analysis and budget estimation

- GP error roughly scales as a power law of the number of training points
- Each multi-fidelity node opens a new dimension to improve the emulator accuracy



$$|f_{\mathrm{HF}}(\boldsymbol{\theta}) - m_{f_{\mathrm{HF}}}(\boldsymbol{\theta})| \propto (\rho \cdot n_{\mathrm{LF}}^{-\frac{\nu}{d}} + n_{\mathrm{HF}}^{-\frac{\nu}{d}}),$$



- Suggested budget:

$$n_{\mathrm{HF}} \propto \left(\frac{1}{C_{\mathrm{HF}}}\right)^{\frac{d}{\nu+d}}$$

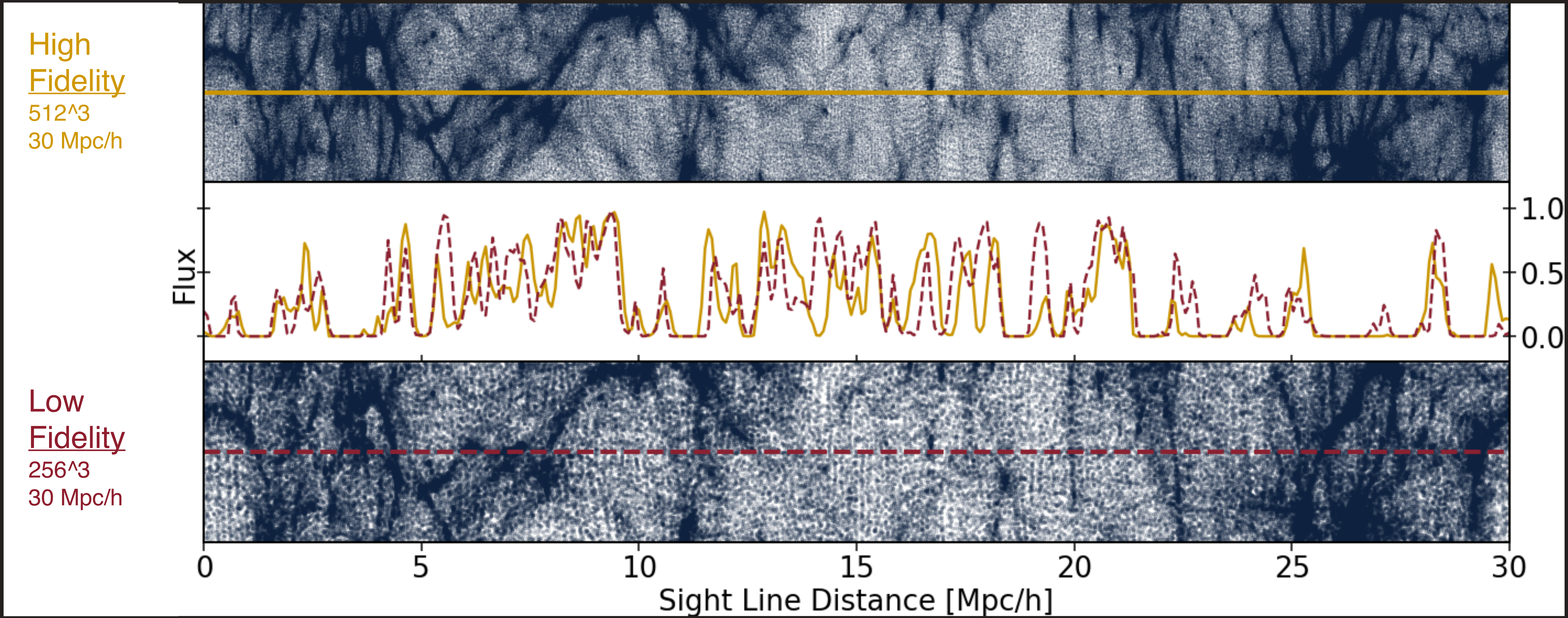
$$n_{\mathrm{LF}} \propto \left(\frac{\rho}{C_{\mathrm{LF}}}\right)^{\frac{d}{\nu+d}}$$

Based on: Ji (2021)

- n_{HF}: number of HF simulations
- ρ: correlation between LF and HF
- C_{HF}: cost of a HF simulation

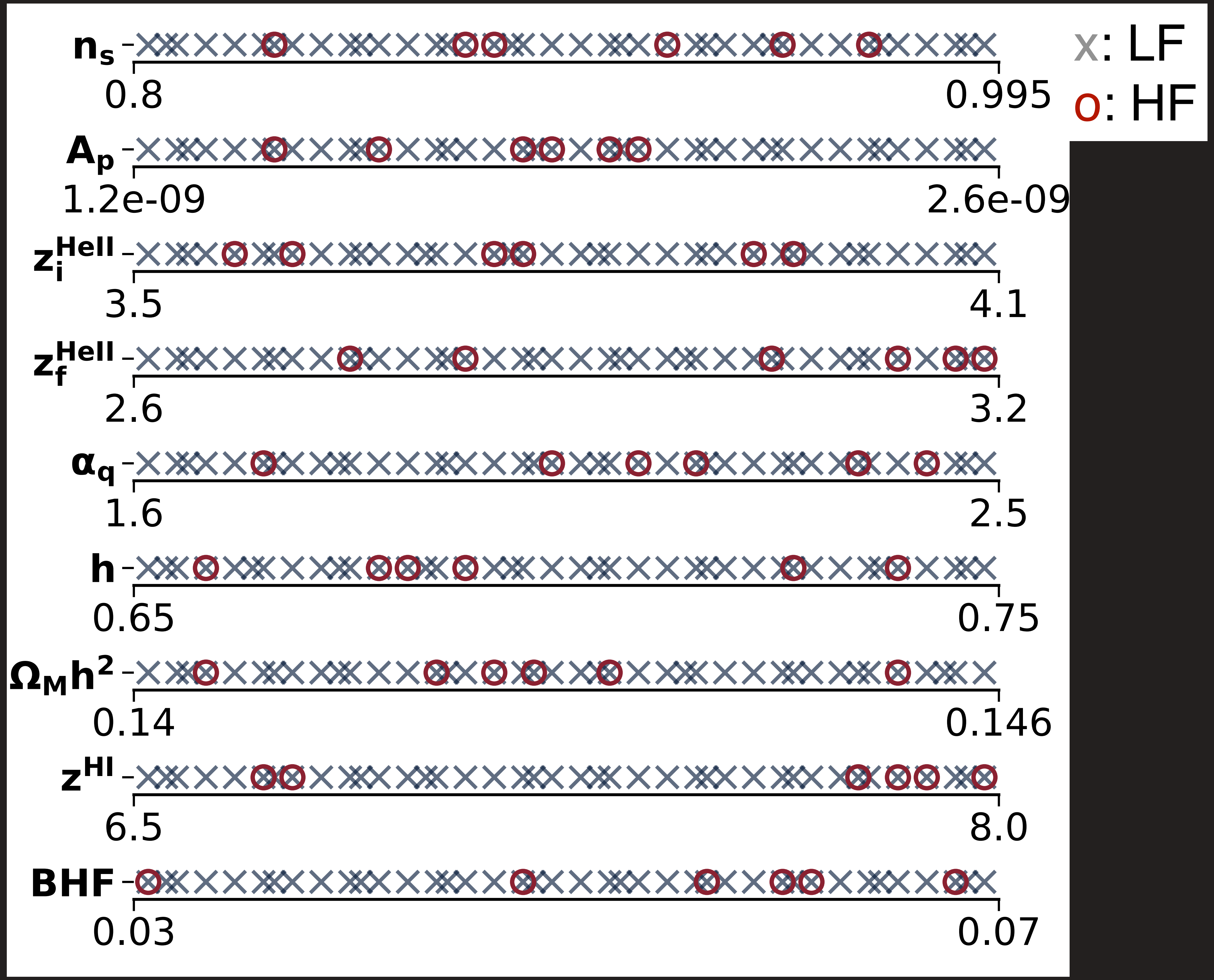
Example 2: Lya flux power spectrum

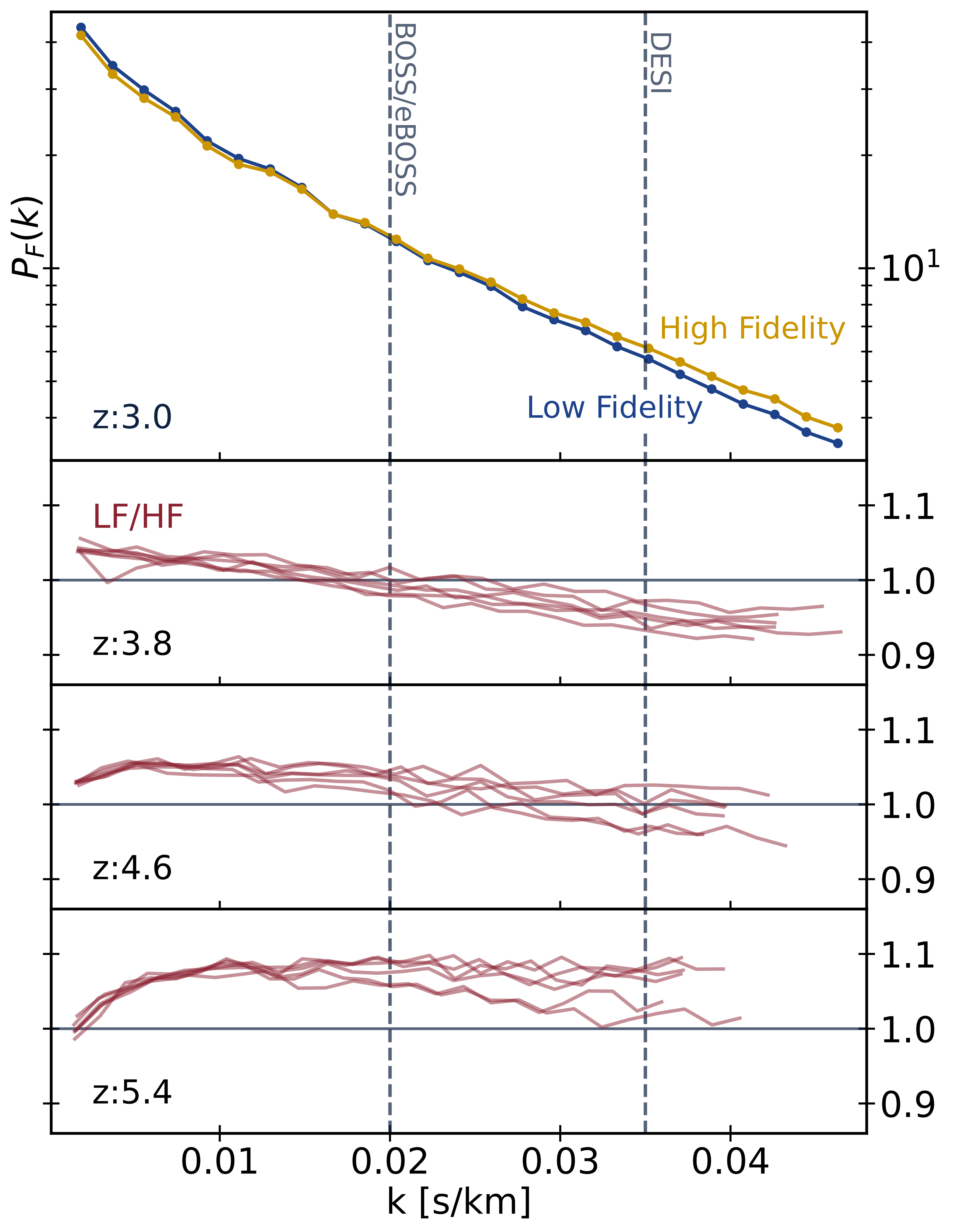
Simulated Lyα forest using Astrid simulations (30 Mpc/h)



- 32,000 simulated spectra per snapshot
- Lya flux power spectrum: Measure correlation between neutral hydrogen within a slightline

Example 2: Lya flux power spectrum
Experimental design in 9 dimensions

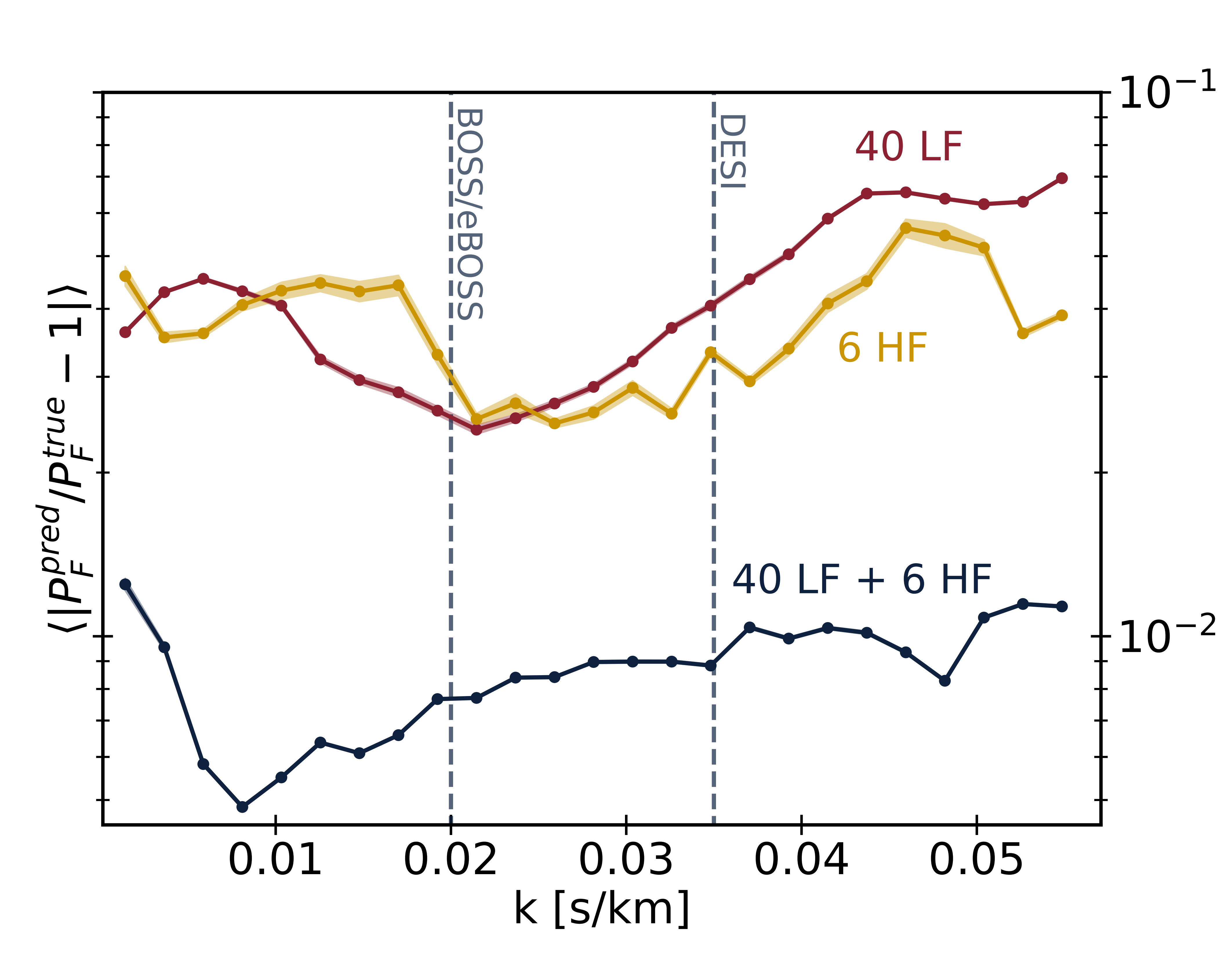


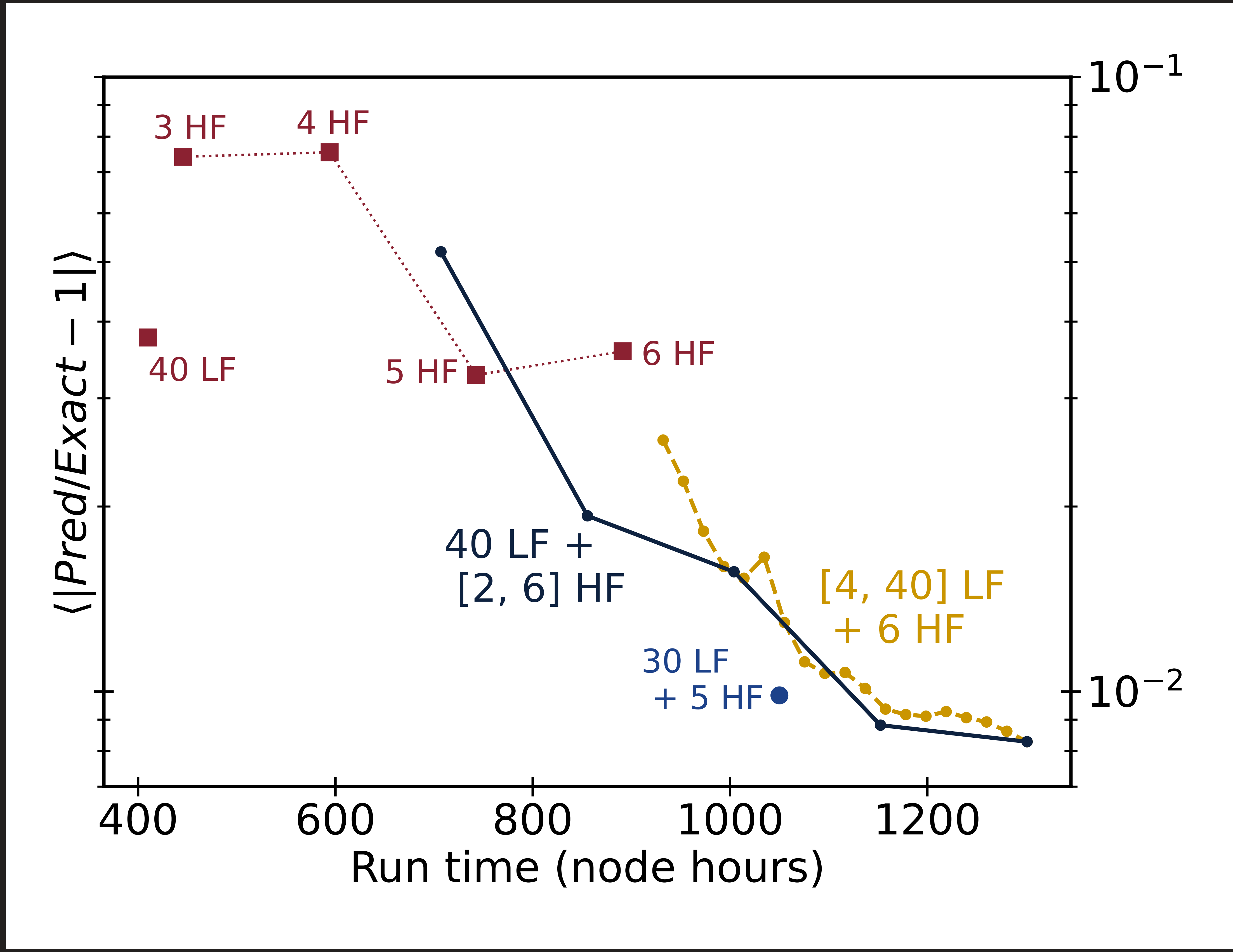


- 9 parameters (z = 2 - 5.4), including reionization parameters and black hole feedback
- Choice of HF optimized using LF suite as a prior

Fernandez, Ho, Bird (2022)

Example 2: Lya flux power spectrum
Application to Lya flux power





- MFEmulator with 40 LF + 6 HF has ≈ 1% accuracy.

Conclusion

- *Multi-fidelity emulation* economically uses simulations from different qualities
- Example 1: First application of MFEmulator to cosmology (DM only)
- Example 2: Application to Lya forest (Astrid) → currently running large-volume production runs
- Possibilities in applying to CAMELS
  - Help fill the parameter space of SB-28 using many more low-fidelity simulations
  - Bridge the information from different box sizes at the emulation level

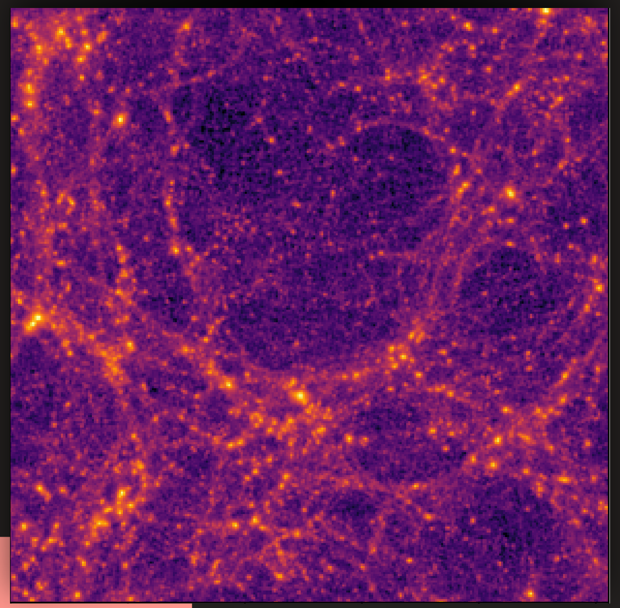
arXiv:2105.01081 | github.com/jibanCat/matter_multi_fidelity_emu
arXiv:2207.06445 | github.com/mafern/MFEmulatorLyaData
Paper for Graphical GP is expected to be submitted later this year.

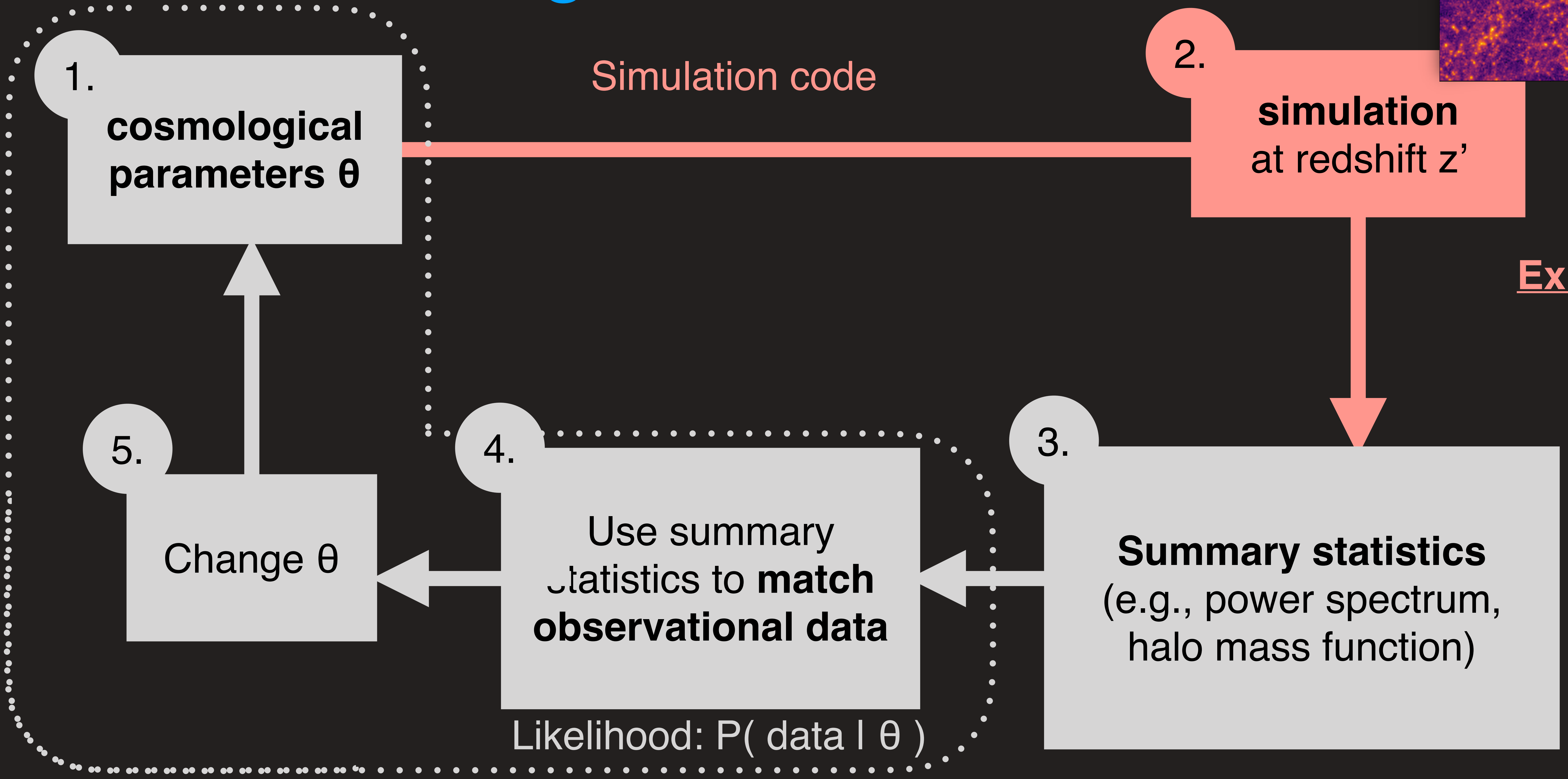
We thanks Yi Ji (Duke, Stat) and Simon Mak (Duke, Stat) for kindly providing the GMGP code in Python.

Backup slides

What's emulation?

Bayesian inference using simulations



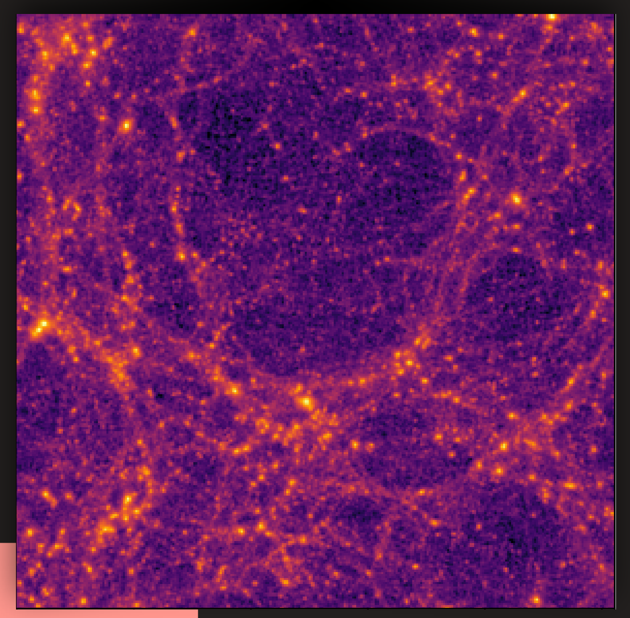


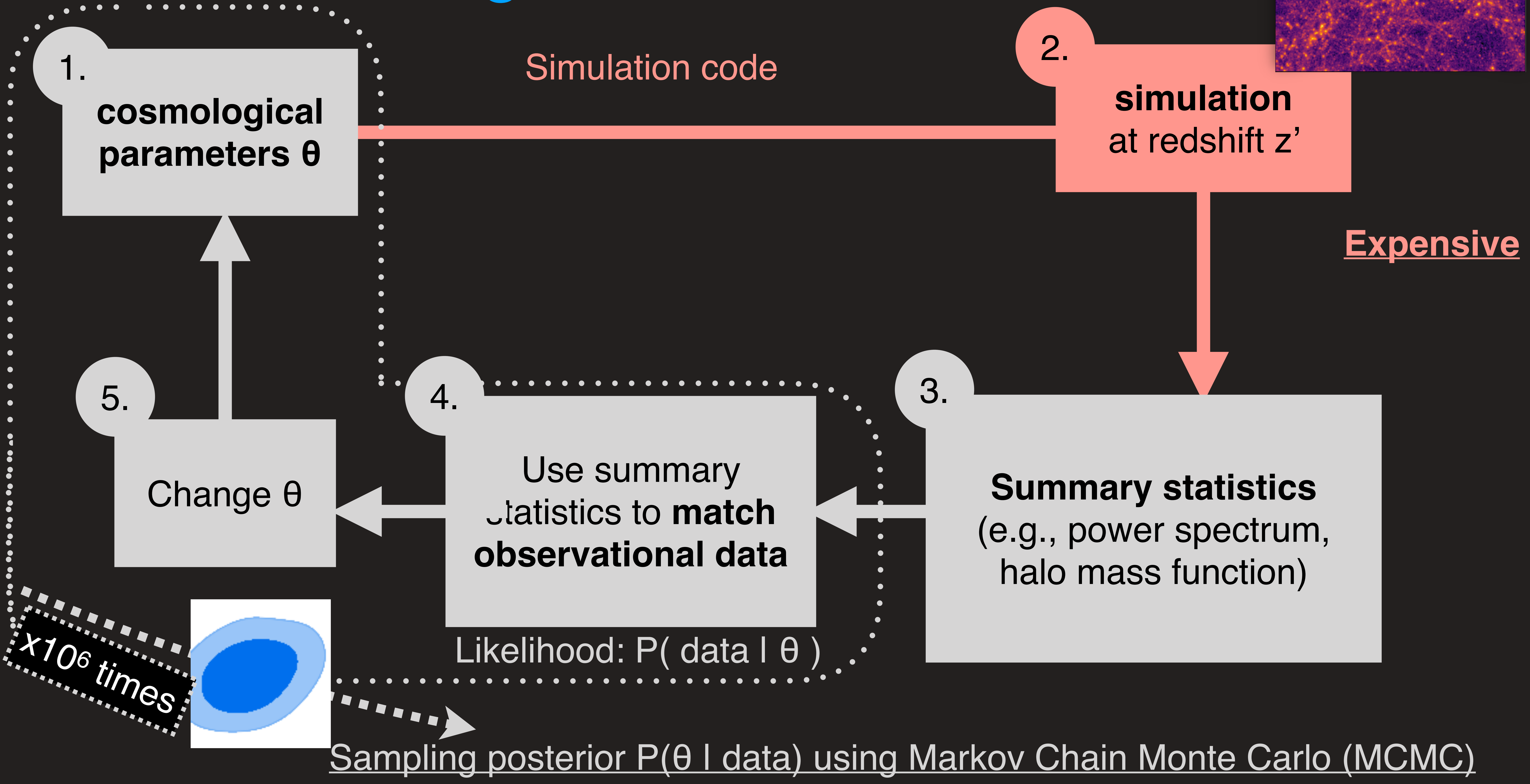
- Slide 4: figures b
big as th
also em
simulatio
do this i
simulatio
- An analo
-

image credit: wiki, cobaya

What’s emulation?

Bayesian inference using simulations





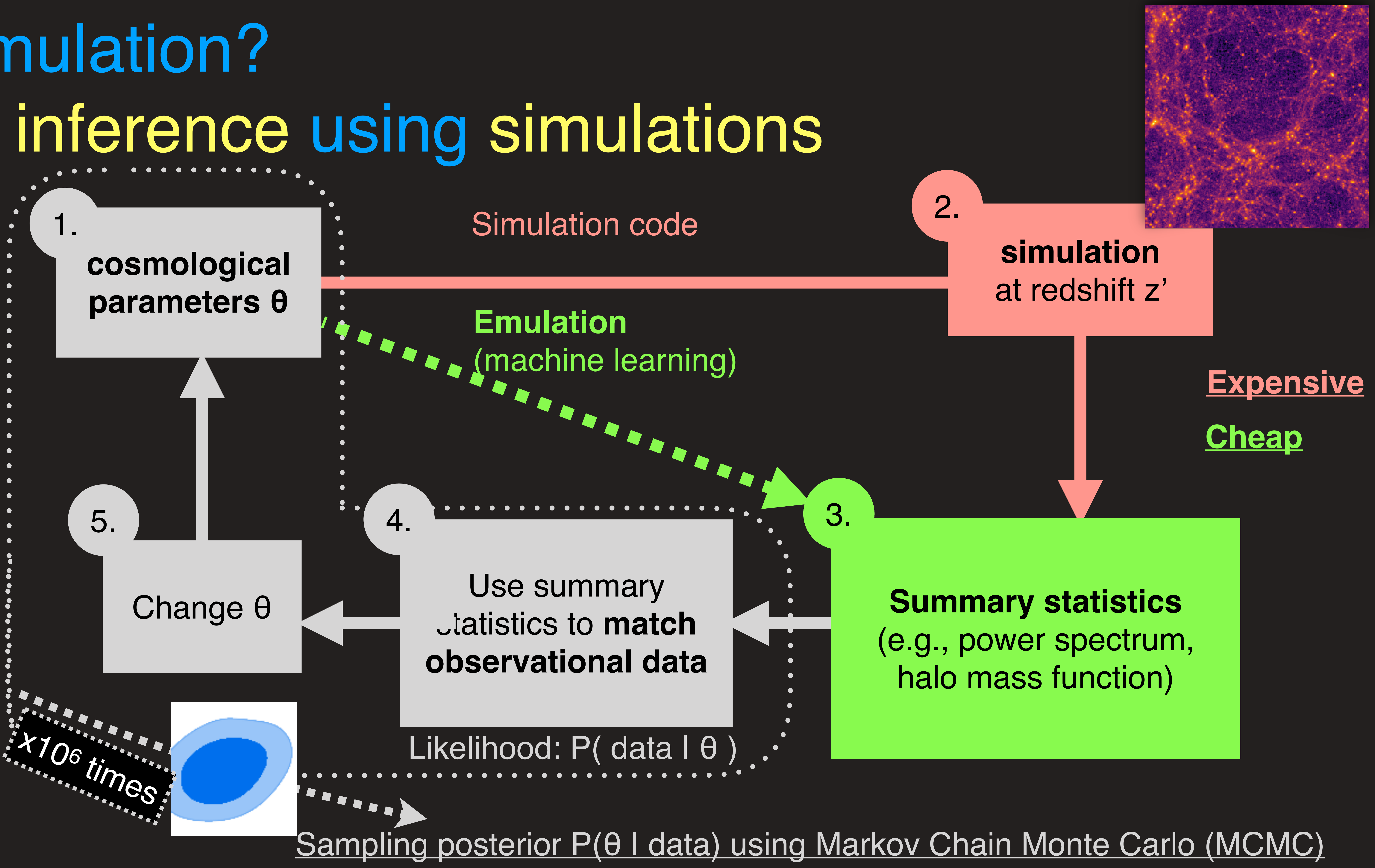
MCMC using forward simulations: require ~10^6 simulations

image credit: wiki, cobaya

- Slide 4: figures b big as th also em simulatio do this it simulatio
- An analo
-

What's emulation?

Bayesian inference using simulations



MCMC using forward simulations: require ~10^6 simulations

MCMC using emulation: require ~10^2 simulations

image credit: wiki, cobaya

- Slide 4:
figures b
big as th
also em
simulatio
do this i
simulatio
- An analo
-

15

What's emulation?

Reducing time on exploring parameter space

- The *everyday* scenario

- Your PI asks you to run 20 simulations before next meeting

- … but you only have time to run 5 simulations.

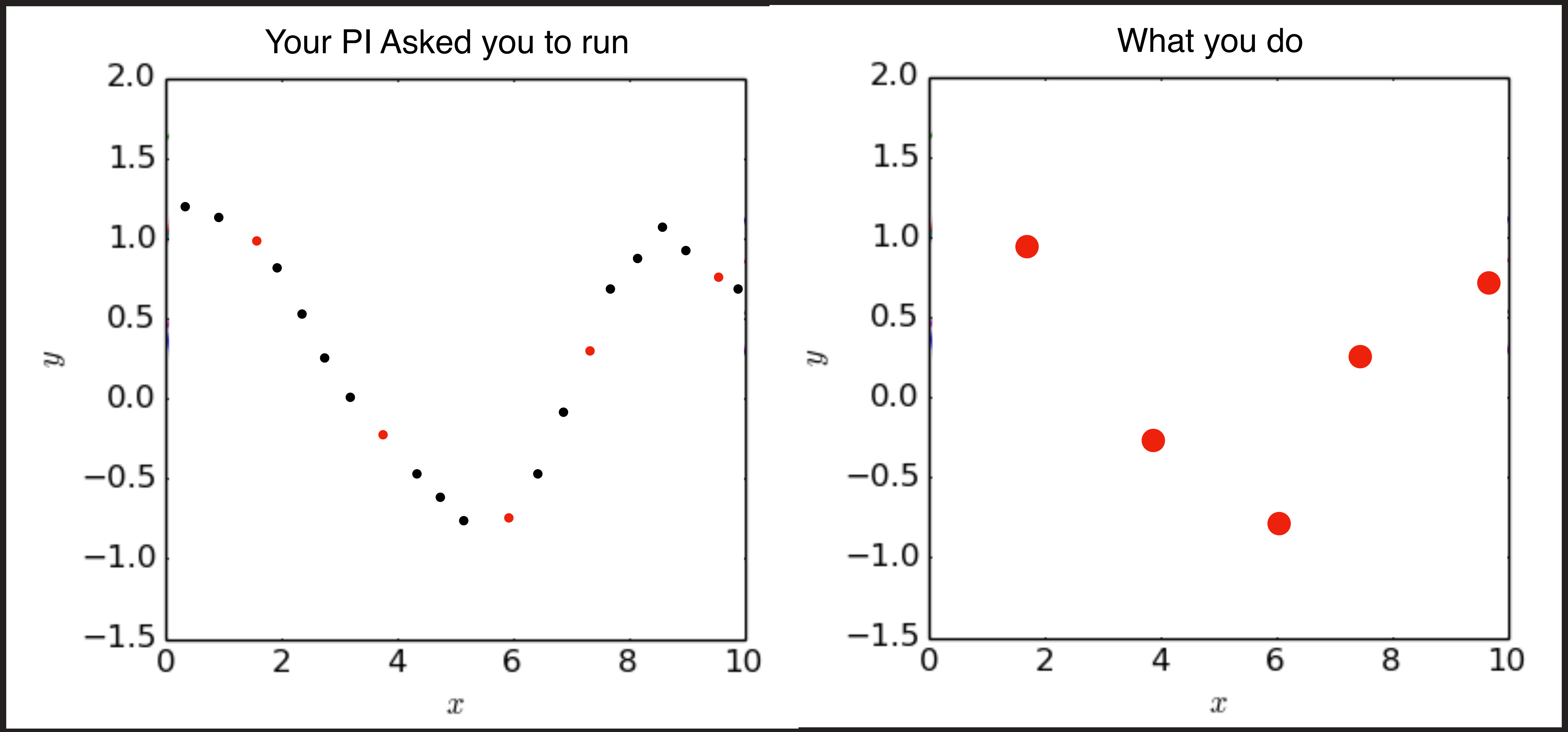


image credit: PHD comics

What’s emulation?

Reducing time on exploring parameter space

- The *everyday* scenario
  - Your PI asks you to run 20 simulations before next meeting
  - … but you only have time to run 5 simulations.

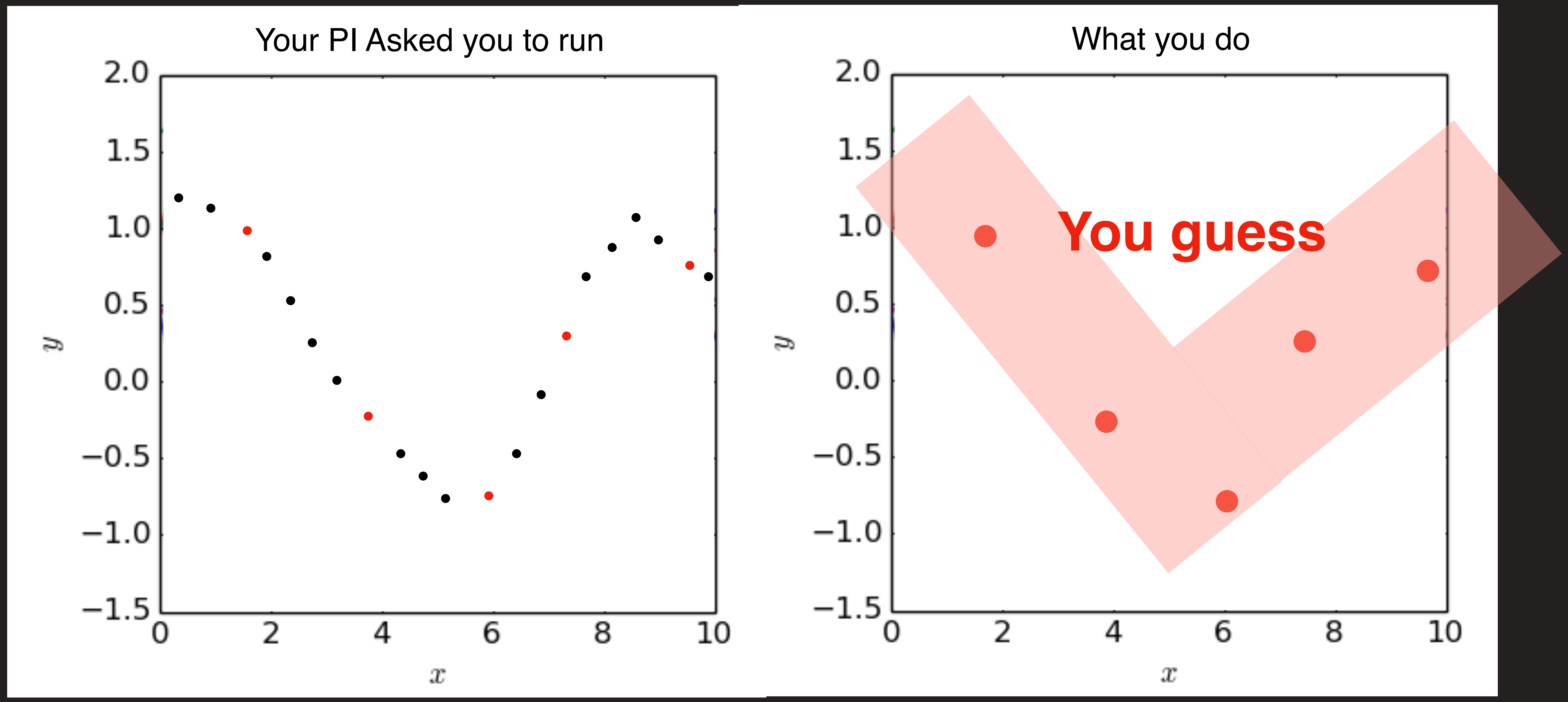


image credit: PHD comics

What's emulation?

It's your *Bayesian* prior/posterior

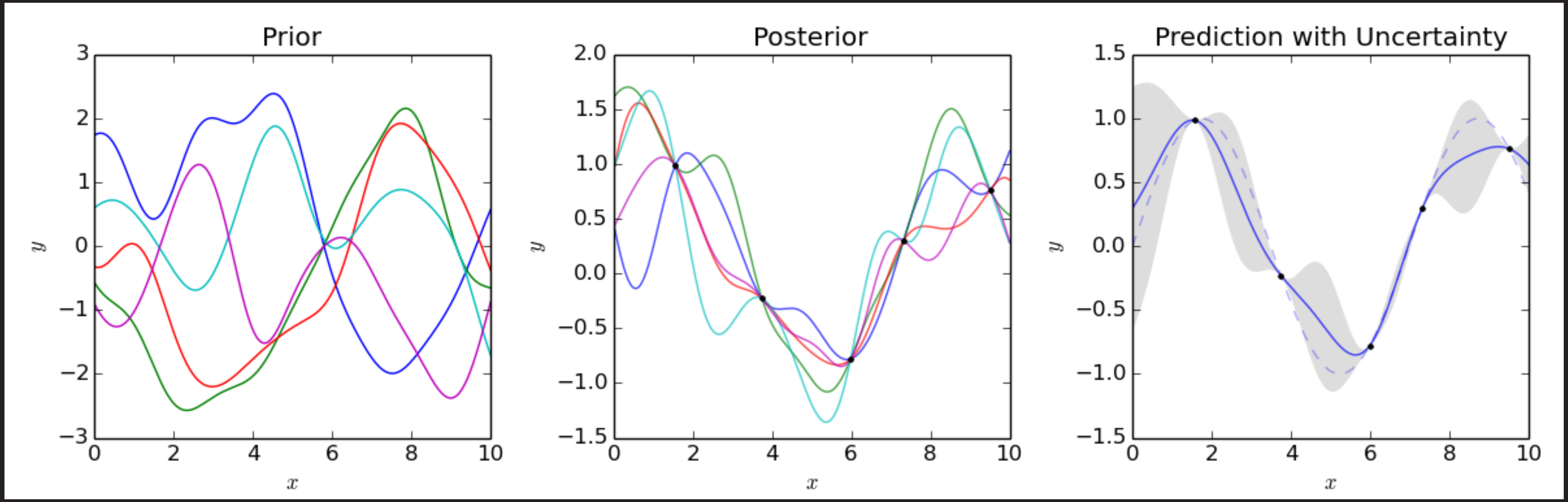
Emulation =

Posterior predictions given *Prior* and *Data*

Simulations

Simulations you haven't run

A distribution over smooth functions



- Gaussian process prior: Smoothness features of y(x) before data are collected.

17

What's emulation?
Cosmic calibration

- Key ingredients for emulation:
  - Surrogate modelling: interpolation
  - Experimental design: space-filling

- Better d
- How do

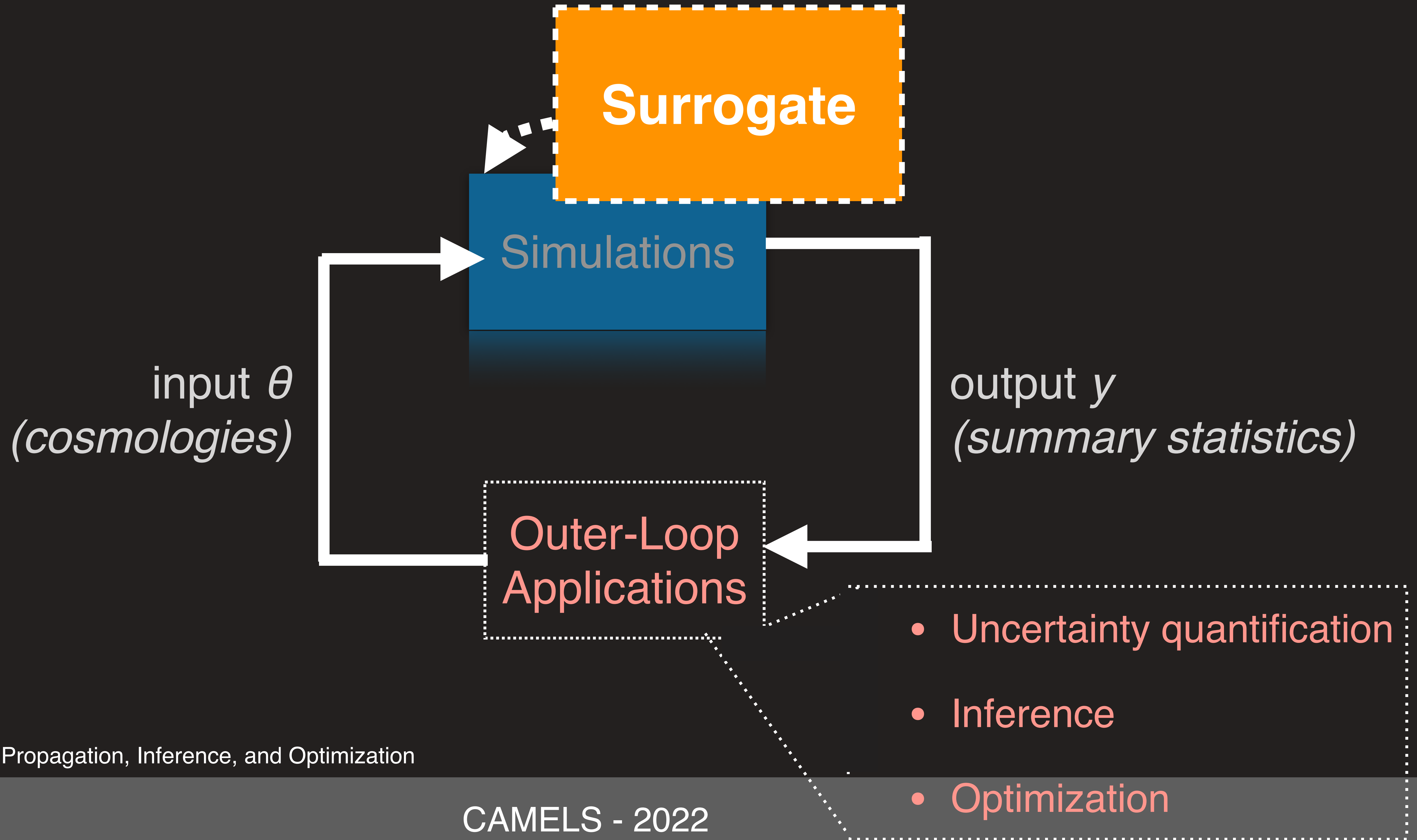


Illustration credit: Perherstorfer et al (2018)
Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization

What's emulation? ~~Cosmic calibration~~

Bayesian calibration for computer experiments

- Key ingredients for ~~emulation~~ *Bayesian* modeling:
  - Surrogate modelling: ~~interpolation~~ ***Prior***
  - Experimental design: ~~space filling~~ ***Data***

Simulations

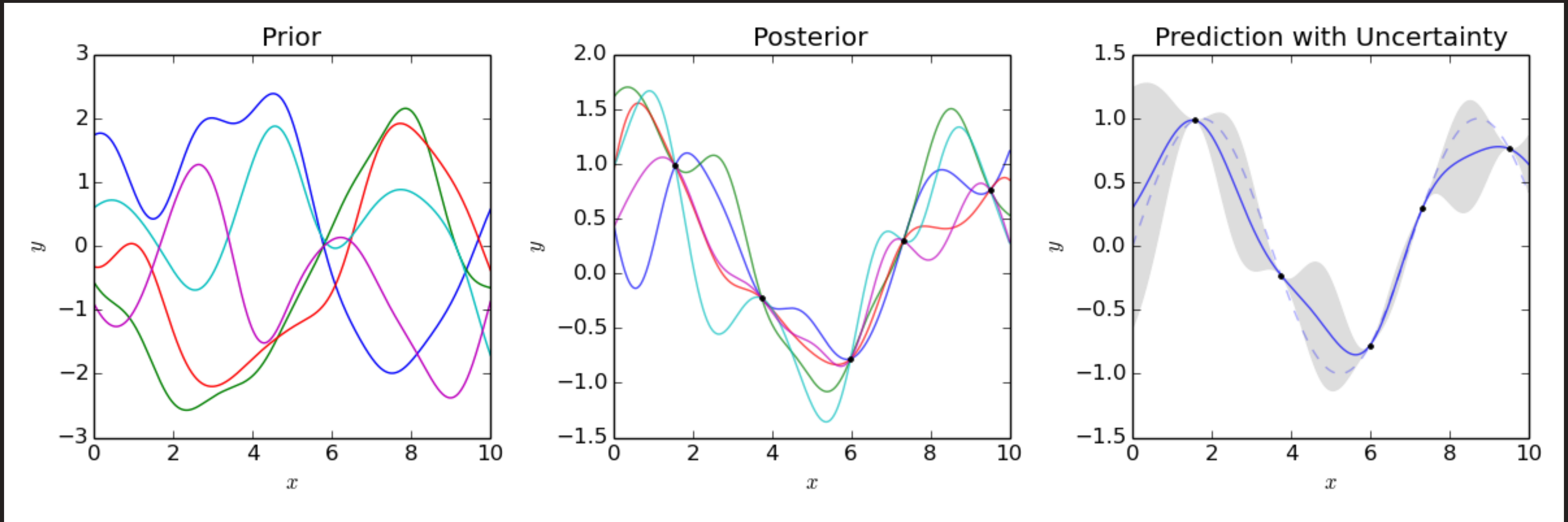
Emulation =

Posterior predictions given *Prior* and *Data*

Simulations you haven't run

A distribution over smooth functions

Gaussian process: Bayesian function prediction



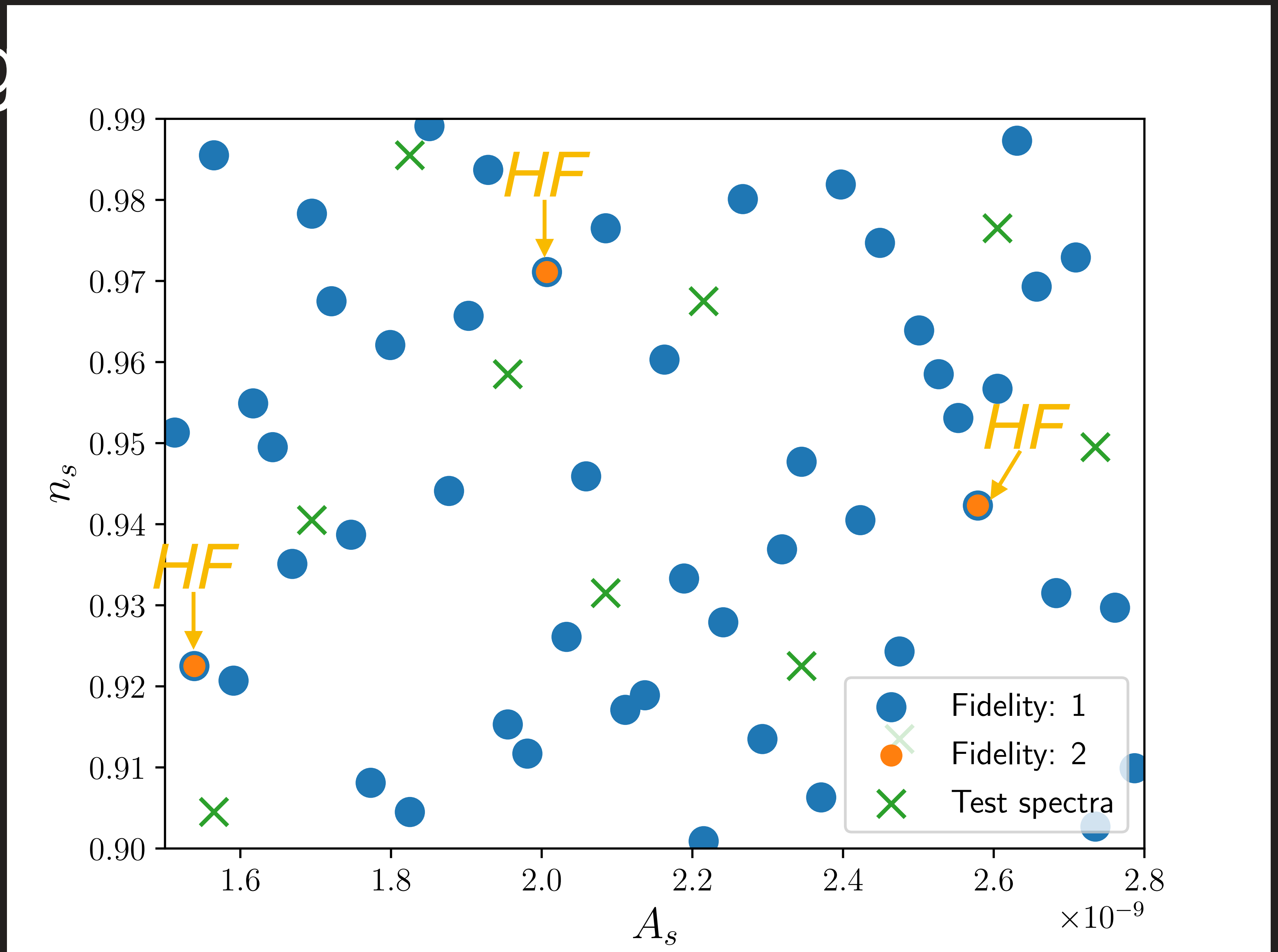
- Gaussian process prior: Smoothness features of y(x) before data are collected.

- Bayesian approach: Choose a flexible prior allowing many shapes of y(x), and let the Bayesian machinery to direct the details of the predictions.*

image credit: wikipedia
* This Bayesian attitude is mentioned in Santner (2003), The Design and Analysis of Computer Experiments

Example 1: matter power spectrum
Experimental design

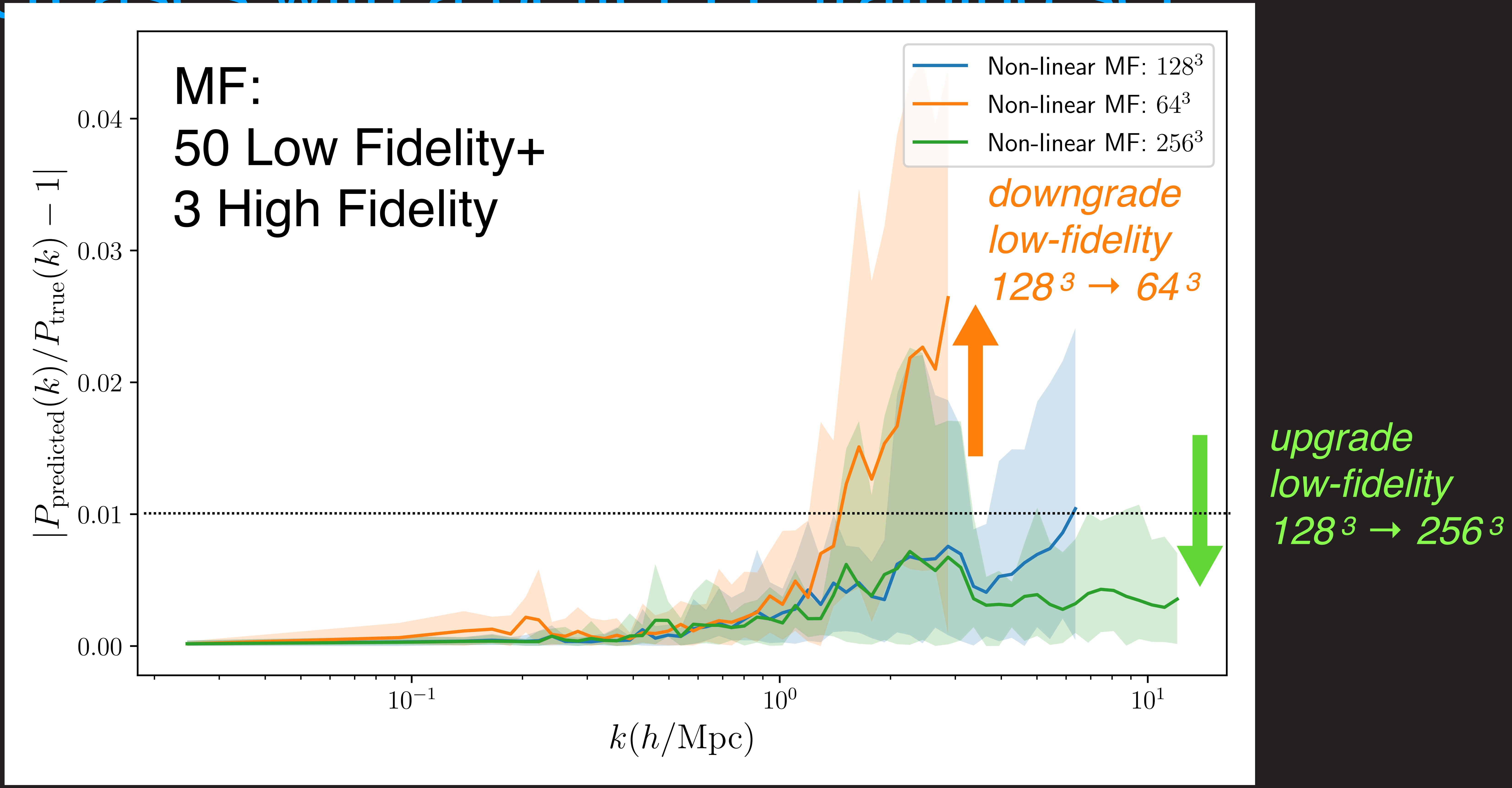
- Parameters: $(h, \Omega_0, \Omega_b, A_s, n_s)$
- *Low-fidelity*: space-filling strateg (Latin hypercube)
  - $.128^3, 256\,\mathrm{Mpc\,h^{-1}}$
- *High-fidelity*: a subset of low-fidelity runs
  - $.512^3, 256\,\mathrm{Mpc\,h^{-1}}$
- *HF* choices were optimized using *LF* simulations



Ho, Bird, Shelton (2022)

- **This loo**
- **Draw arr people w**
- **Prior vo here but can put i**
- **Stress th improve**
- **Optimize interpola**

Example 1: matter power spectrum
Accuracy increases with a better LF training set

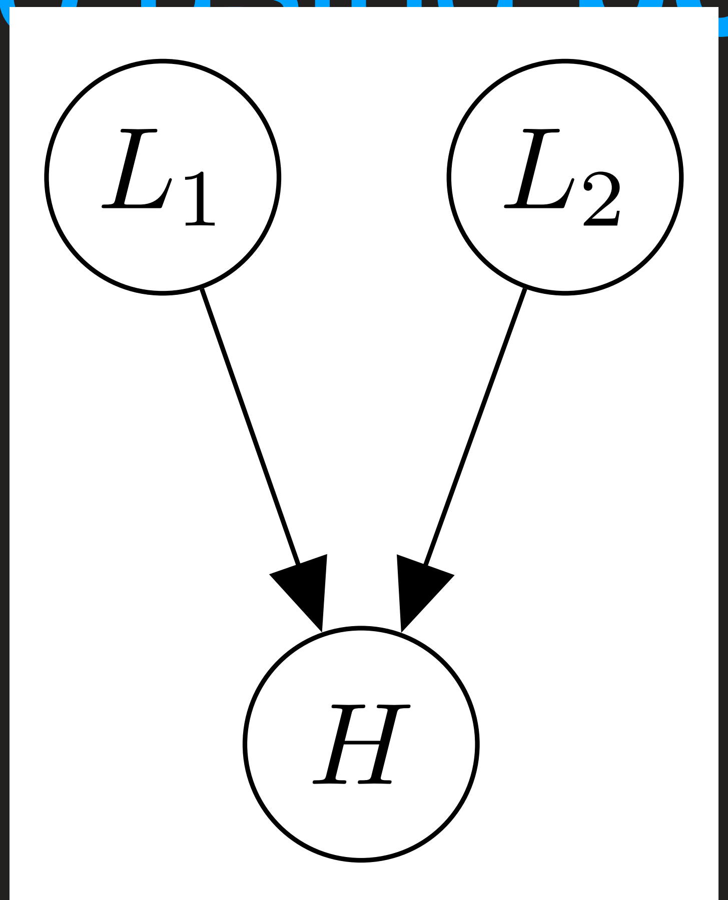


- The quality of LF simulations affects the accuracy of MF emulation
- Small scales emulation can be improved with a better quality of LF simulation suite. → *Question: Can we use small box LF to improve emulation?*

- **This ac**
- **Someo the 64^ accurat**
- **The poi design budget accura versatil**

22

Example 2: matter power spectrum
Extending to using **boxsize** as a fidelity



- L_1: 128^3, 256 Mpc/h
- L_2: 128^3, 100 Mpc/h
- H: 512^3, 256 Mpc/h

- Number of particles is not the only fidelity variable, ***boxsize*** is also a fidelity variable
- A smaller boxsize, ***better*** resolution at ***small scales***
- We can combine both ***large box (L1)*** and ***small box (L2)*** information through a graphical model construction.
- A graphical GP (Ji et al., 2021) allows us to do so.

- Did not l

Deep Graphical Multi-fidelity GP (dGMGP)

$$f_{\mathrm{HF}}(x) = \rho(\{f_{\mathrm{LF},1}, f_{\mathrm{LF},2}\} \cup x) + \delta(x)$$

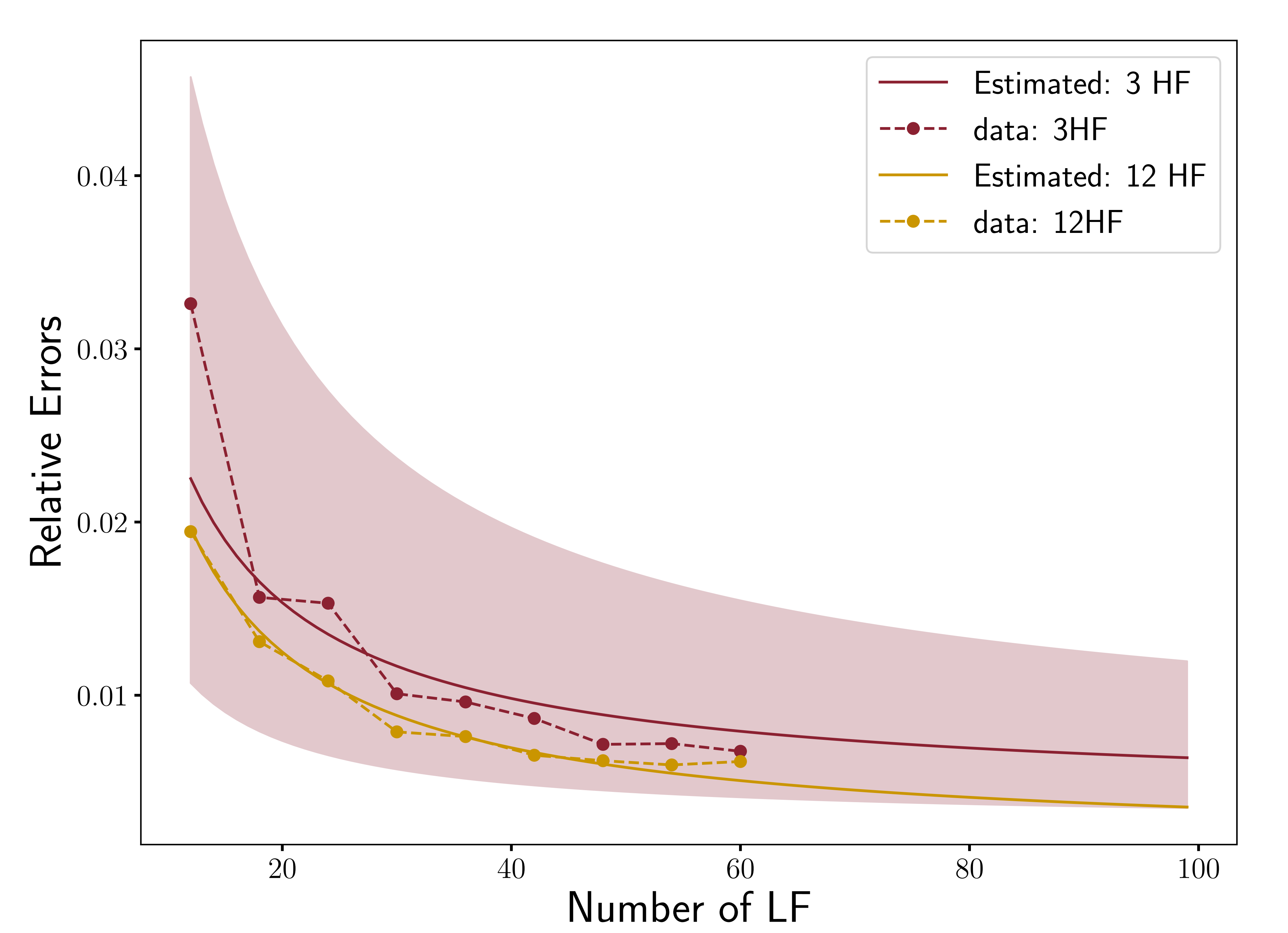
$$K([x, f_{\mathrm{LF}}], [x', f'_{\mathrm{LF}}]) = K_{\mathrm{SE}}(x, x')[K_{\mathrm{LIN}}(f_{\mathrm{LF}}, f'_{\mathrm{LF}}) + K_{\mathrm{SE}}(f_{\mathrm{LF}}, f'_{\mathrm{LF}})] + K_{\mathrm{SE}}(x, x')$$

K_{SE} : Squared-exponential kernel, guarantees smooth functions
K_{LIN} : Linear kernel, doing Bayesian linear regression

GMGP: Ji (2021)
A graphical multi-fidelity Gaussian process model, with application to emulation of expensive computer simulations

Example: matter power spectrum
Error analysis & optimal design

- Emulation error roughly scales as a power law



$$|f_{\mathrm{HF}}(\boldsymbol{\theta}) - m_{f_{\mathrm{HF}}}(\boldsymbol{\theta})| \propto (\rho \cdot n_{\mathrm{LF}}^{-\frac{\nu}{d}} + n_{\mathrm{HF}}^{-\frac{\nu}{d}}),$$

Theoretical basis: Wendland (2004), Ji (2021)

- Solve the Lagrangian multiplier with a fixed budget gives you the optimal design:

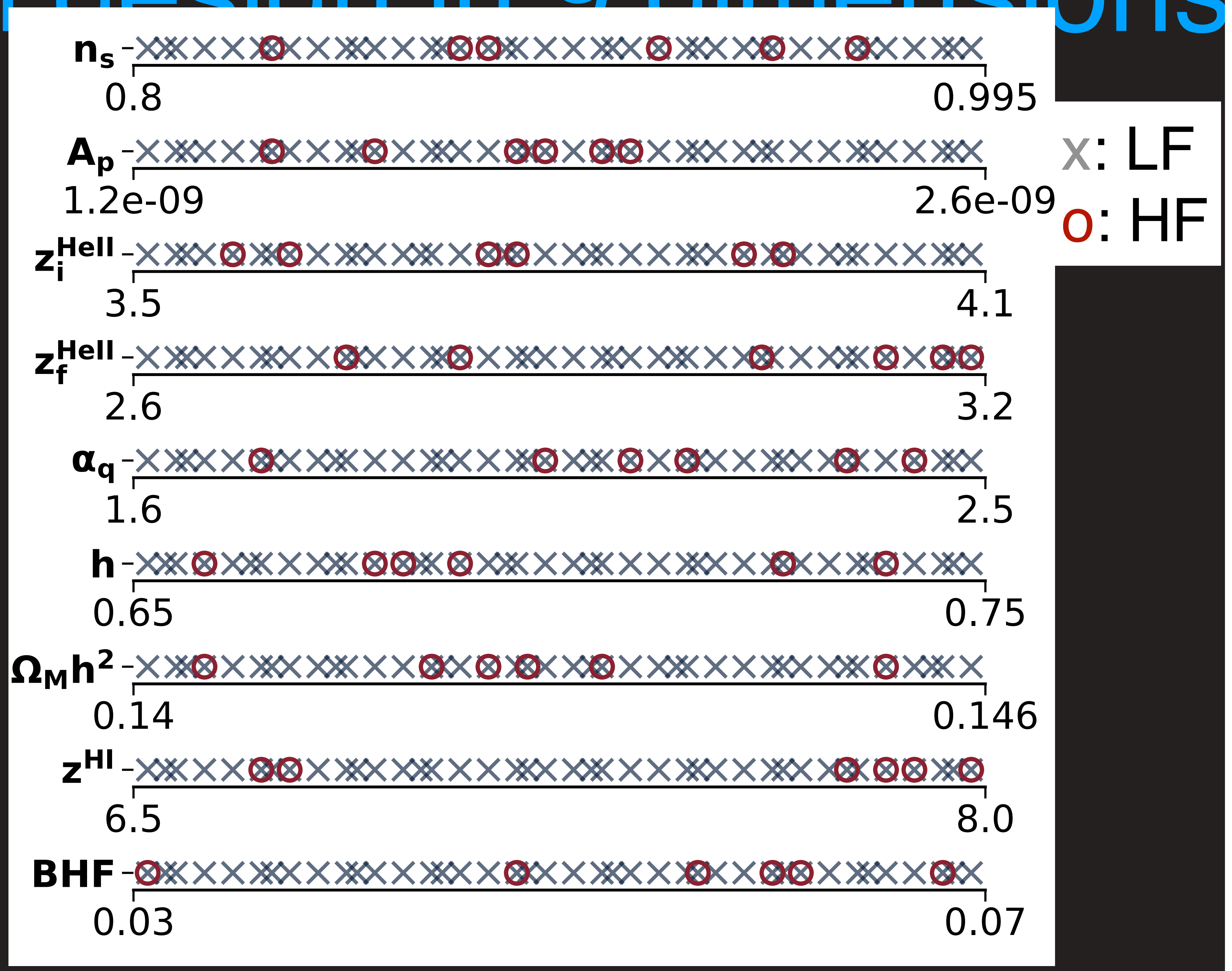
$$n_{\mathrm{HF}} \propto \left(\frac{1}{C_{\mathrm{HF}}}\right)^{\frac{d}{\nu+d}}$$

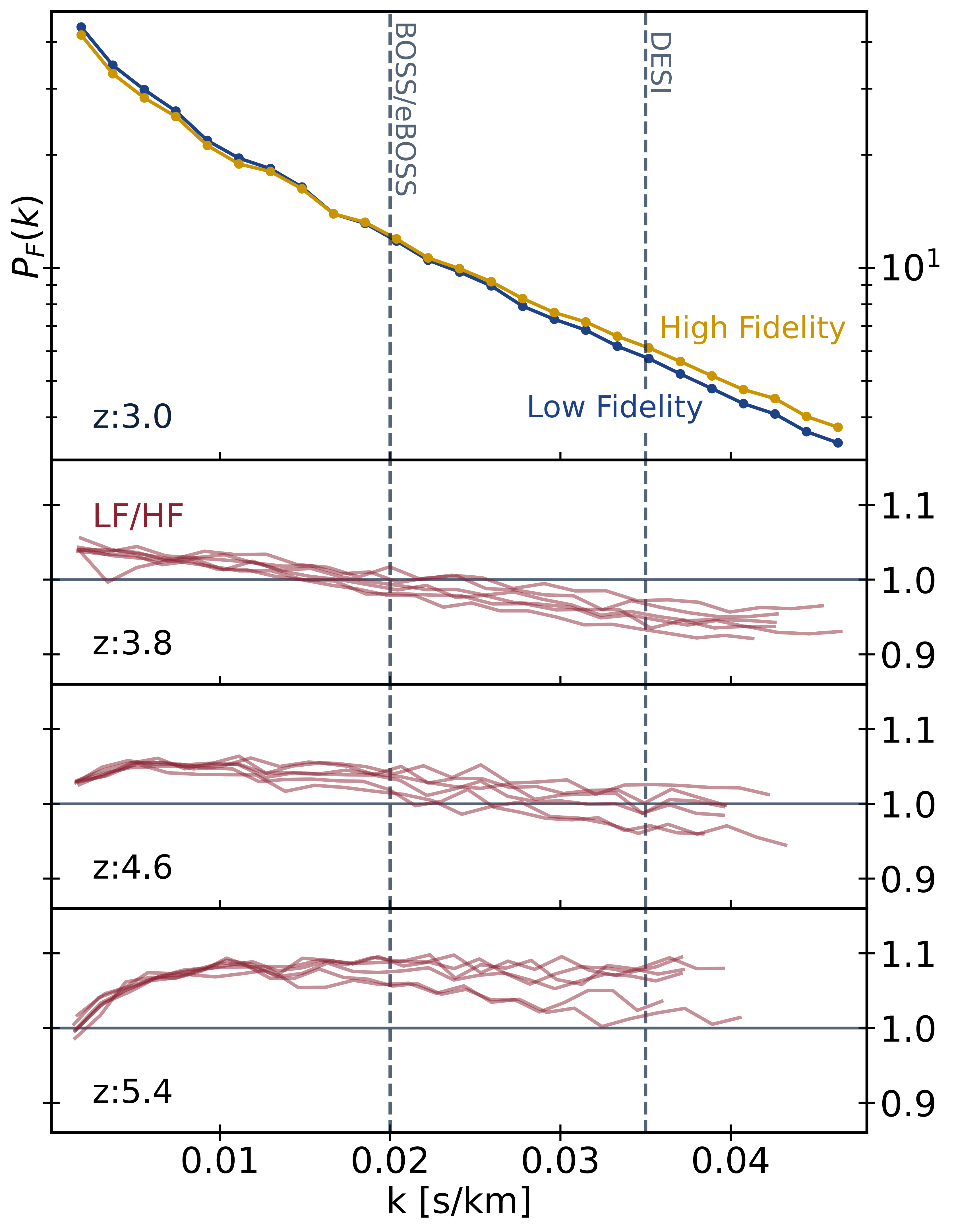
$$n_{\mathrm{LF}} \propto \left(\frac{\rho}{C_{\mathrm{LF}}}\right)^{\frac{d}{\nu+d}}$$

- n_{HF}: number of HF simulations
- C_{HF}: cost of a HF simulation
- ρ: correlation between LF and HF

Ho, Bird, Fernandez, Shelton (in prep)

Example 3: Lya flux power spectrum
Experimental design in 9 dimensions



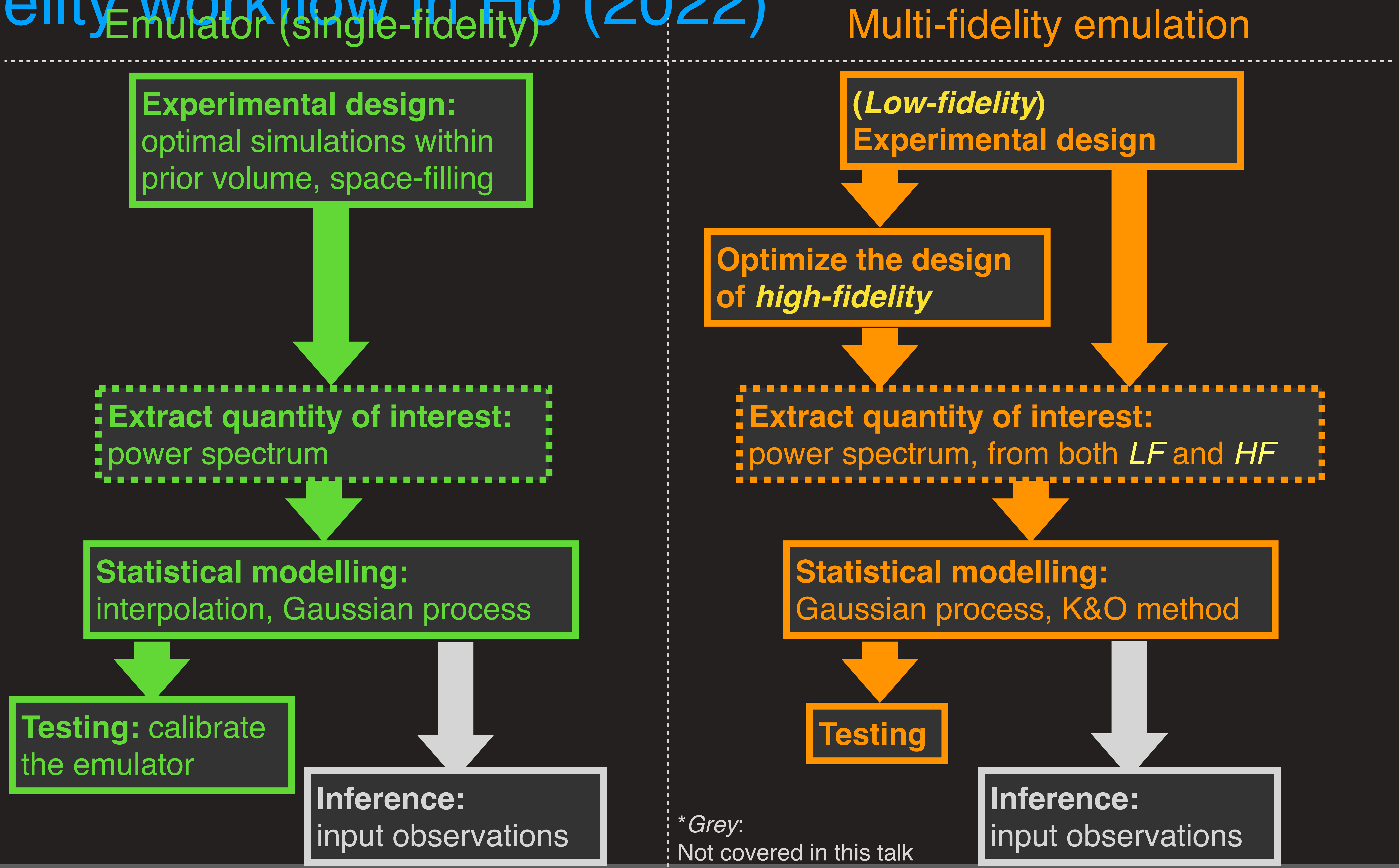


- 9 parameters (z = 2 - 5.4), including reionization parameters and black hole feedback
- The discrepancy between LF/HF appears across scales (k), varies with redshifts

figure credit: Martin Fernandez

Fernandez, Ho, Bird (2022)

What's multi-fidelity?
Multi-fidelity workflow in Ho (2022)



- Looks a
- Kind of
  Heitma
- not sur

Non-linear (NARGP) or linear (KO)?

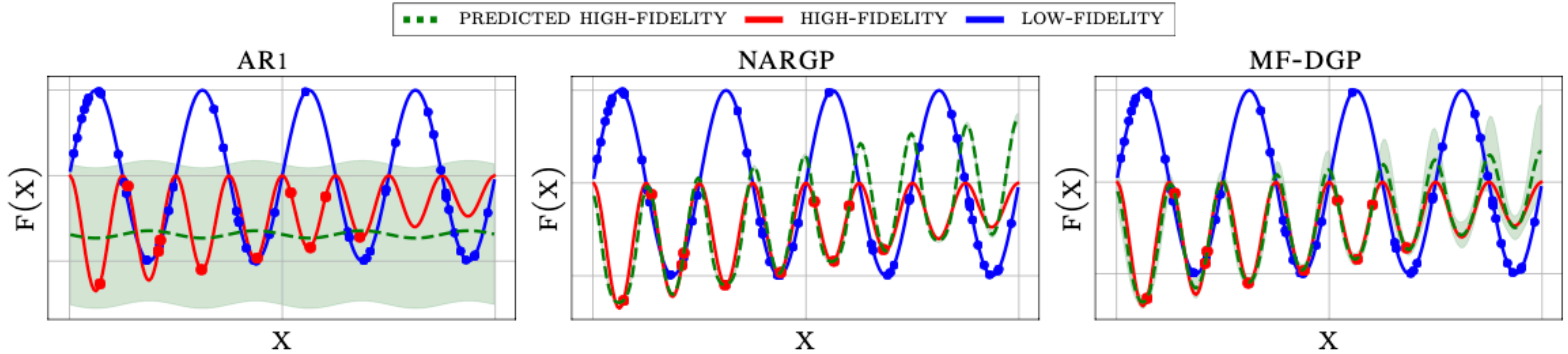


Figure 1: Limitations addressed and resolved jointly by our proposed MF-DGP architecture. Blue and red markers denote low and high-fidelity observations respectively. Shaded regions indicate the 95% confidence interval.

Uncertainty quantification depends on kernel choice

fig: Cutajar et al. - 2019 - Deep Gaussian Processes for Multi-fidelity Modeling