Multi-Fidelity Emulation for Matter Power Spectrum and Lyα Forest

... using Astrid simulation code

Ming-Feng Ho  
UC Riverside  
NASA FINESST FI

3. Ho, Bird, Fernandez, Shelton (in prep)

Simeon Bird (UCR Astro)  
Martin A. Fernandez (UCR Astro)  
Christian Shelton (UCR CS)
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

High-Fidelity* (e.g., N-body simulations with many particles)

Inaccurate results but cheap to run

Low-Fidelity (e.g., N-body simulations with fewer particles)

Accurate results but expensive to run

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

*Could be anything you think is high-fidelity
What’s multi-fidelity emulation?
An analogy: University’s hiring

High-Fidelity Professors

Less-accurate results but inexpensive to hire

Low-Fidelity Grad Student Researchers

Accurate results but expensive to hire

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


Graphical model GP (2021)

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

\[
f_{\text{HF}}(x) = \rho(\{f_{\text{LF},1}, f_{\text{LF},2}\} \cup x) + \delta(x)
\]
Multi-fidelity emulators:
Transfer learn the simulations from different resolutions + box sizes

Low-fidelity:
128$^3$
256 Mpc/h

High-fidelity:
512$^3$
256 Mpc/h

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

Low-fidelity:
128$^3$
256 Mpc/h

Low-fidelity 2
128$^3$
100 Mpc/h

High-fidelity:
512$^3$
256 Mpc/h

$f_{HF}(x) = \rho(\{f_{LF,1}, f_{LF,2}\} \cup x) + \delta(x)$.
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)
  - \(128^3, 256 \text{ Mpc} h^{-1}\)
- **3 High-fidelity**: a subset of low-fidelity runs
  - \(512^3, 256 \text{ Mpc} h^{-1}\)
- **HF** choices were optimized using **LF** suite as a prior

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.

Graphical GP: (128$^3$, 256 Mpc/h), (128$^3$, 100 Mpc/h), (512$^3$, 256 Mpc/h)

MF: (128$^3$, 512$^3$) 256 Mpc/h
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

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

- \( n_{HF} \): number of HF simulations
- \( \rho \): correlation between LF and HF
- \( C_{HF} \): cost of a HF simulation

Suggested budget:

\[ n_{HF} \propto \left( \frac{1}{C_{HF}} \right)^{\frac{d}{\nu+d}} \]
\[ n_{LF} \propto \left( \frac{\rho}{C_{LF}} \right)^{\frac{d}{\nu+d}} \]

Based on: Ji (2021)
Example 2: Lya flux power spectrum
Simulated Lya 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

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_e$</td>
<td>0.8</td>
</tr>
<tr>
<td>$A_p$</td>
<td>1.2e-09</td>
</tr>
<tr>
<td>$z_{\text{HeII}}$</td>
<td>3.5</td>
</tr>
<tr>
<td>$z_{\text{HeII}}$</td>
<td>2.6</td>
</tr>
<tr>
<td>$\alpha_q$</td>
<td>1.6</td>
</tr>
<tr>
<td>$h$</td>
<td>0.65</td>
</tr>
<tr>
<td>$\Omega_M h^2$</td>
<td>0.14</td>
</tr>
<tr>
<td>$z_{\text{HI}}$</td>
<td>6.5</td>
</tr>
<tr>
<td>BHF</td>
<td>0.03</td>
</tr>
</tbody>
</table>

$X$: LF
$\circ$: HF

Example 2: Lya flux power spectrum

$P_{\ell}(k)$ vs $k$ [s/km]
- Low Fidelity
- High Fidelity

$z$: 3.0, 3.8, 4.6, 5.4
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

1. **cosmological parameters** $\theta$

2. **simulation at redshift** $z'$

3. **Summary statistics**
   (e.g., power spectrum, halo mass function)

4. **Use summary statistics to match observational data**

5. **Change $\theta$**

Simulation code

Expensive

- Slide 4: I might make the simulation and power spectrum figures bigger. Don’t rearrange anything, just make them as big as they can be and still fit where they are. You could also emphasize how impossible running MCMC with actual simulation outputs is by pointing out that you would have to do this iteratively, i.e. take an MCMC step, run the relevant simulation, take another step, run the new simulation, etc.

- An analogy would be better.
What’s emulation? Bayesian inference using simulations

1. cosmological parameters $\theta$

2. simulation at redshift $z'$

3. Summary statistics (e.g., power spectrum, halo mass function)

4. Use summary statistics to match observational data

5. Change $\theta$

Simulation code

Sampling posterior $P(\theta \mid \text{data})$ using Markov Chain Monte Carlo (MCMC)

Expensive

MCMC using forward simulations: require $\sim 10^6$ simulations

- Slide 4: I might make the simulation and power spectrum figures bigger. Don’t rearrange anything, just make them as big as they can be and still fit where they are. You could also emphasize how impossible running MCMC with actual simulation outputs is by pointing out that you would have to do this iteratively, i.e., take an MCMC step, run the relevant simulation, take another step, run the new simulation, etc.

- An analogy would be better...
What's emulation?

*Bayesian* inference using simulations

1. **cosmological parameters** $\theta$
2. **simulation** at redshift $z'$
3. **Summary statistics** (e.g., power spectrum, halo mass function)
4. **Use summary statistics to match observational data**
5. **Change $\theta$**

Expensive

Cheap

- Slide 4: I might make the simulation and power spectrum figures bigger. Don't rearrange anything, just make them as big as they can be and still fit where they are. You could also emphasize how impossible running MCMC with actual simulation outputs is by pointing out that you would have to do this iteratively, i.e. take an MCMC step, run the relevant simulation, take another step, run the new simulation, etc.
- An analogy would be better
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.
What’s emulation?
It’s your *Bayesian* prior/posterior

**Emulation** =

*Posterior* predictions given *Prior* and *Data*

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

![Prior](image1.png) ![Posterior](image2.png) ![Prediction with Uncertainty](image3.png)

**Simulations**

Simulations you haven’t run

A distribution over smooth functions
What’s emulation?
Cosmic calibration

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

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

Input $\theta$ (cosmologies)
Outer-Loop Applications

Output $y$ (summary statistics)

Surrogate

- Better definition outer loop applications
- How do you know when to stop training?
- Uncertainty quantification
- Inference
- 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

**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.*


*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 strategy (Latin hypercube)
  - \(.128^3, 256 \text{ Mpc}\ h^{-1}\)

- **High-fidelity**: a subset of low-fidelity runs
  - \(.512^3, 256 \text{ Mpc}\ h^{-1}\)

- HF choices were optimized using LF simulations

Ho, Bird, Shelton (2022)
Example 1: matter power spectrum
Accuracy increases with a better LF training set

- This actually looks nice
- Someone might ask about the $64^3$ being not accurate
- The point is to say you can design based on your own budget and target accuracy, it's very versatile.

![Graph showing non-linear MF: 128^3, 64^3, 256^3](image)

- 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?

Ho, Bird, Shelton (2022)
Example 2: matter power spectrum
Extending to using boxsize as a fidelity

- 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.

- L1: $128^3$, 256 Mpc/h
- L2: $128^3$, 100 Mpc/h
- H: $512^3$, 256 Mpc/h

Deep Graphical Multi-fidelity GP (dGMGP)

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

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

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

A graphical multi-fidelity Gaussian process model, with application to emulation of expensive computer simulations

GMGP: Ji (2021)
Example: matter power spectrum
Error analysis & optimal design

- Emulation error roughly scales as a power law

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

\[
\begin{align*}
    n_{HF} &\propto \left( \frac{1}{C_{HF}} \right)^{\frac{d}{\nu+d}} \\
    n_{LF} &\propto \left( \frac{\rho}{C_{LF}} \right)^{\frac{d}{\nu+d}}
\end{align*}
\]

- \( n_{HF} \): number of HF simulations
- \( C_{HF} \): cost of a HF simulation
- \( \rho \): correlation between LF and HF

Theoretical basis: Wendland (2004), Ji (2021)
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
What’s multi-fidelity?
Multi-fidelity workflow in Ho (2022)

**Experimental design:**
- Optimal simulations within prior volume, space-filling

**Extract quantity of interest:**
- Power spectrum

**Statistical modelling:**
- Interpolation, Gaussian process

**Testing:**
- Calibrate the emulator

**Inference:**
- Input observations

---

**Multi-fidelity emulation**

**Experimental design:**
- (Low-fidelity)

**Optimize the design of high-fidelity**

**Extract quantity of interest:**
- Power spectrum, from both LF and HF

**Statistical modelling:**
- Gaussian process, K&O method

**Testing**

**Inference:**
- Input observations

---

- Looks a bit complicated
- Kind of following Heitmann’s slides
- Not sure how to improve
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