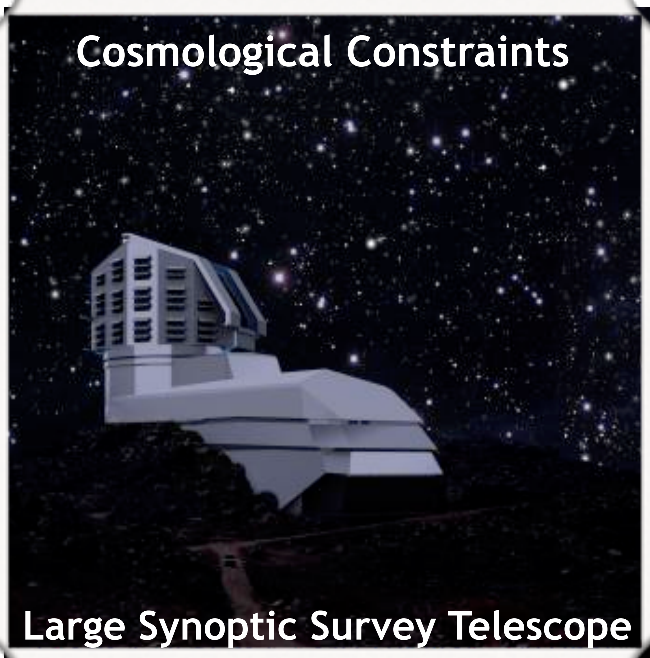
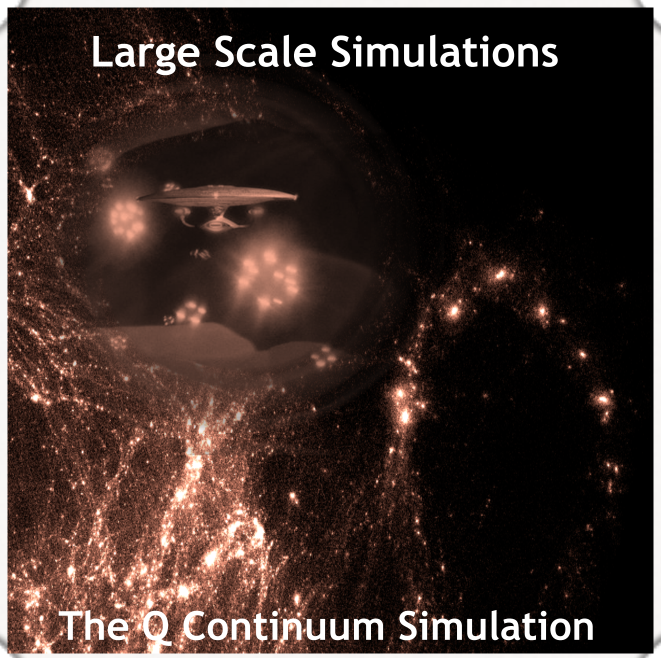
Cosmic Calibration

Katrin Heitmann

Statistical Challenges for Large-Scale Structure in the Era of LSST

Oxford , April 18, 2018





http://www.hep.anl.gov/cosmology/CosmicEmu/emu.html

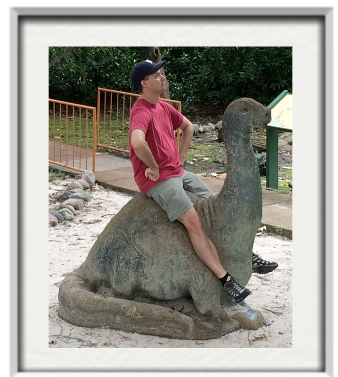
Thanks to many collaborators!




- **The Beginnings -- Proof of Concept**

  (Heitmann et al. 2006, Habib et al. 2007)



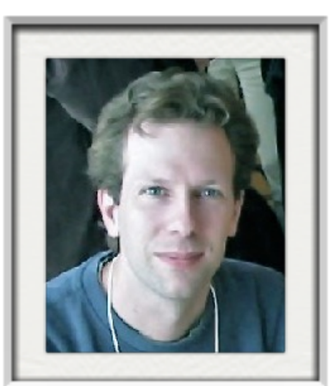



- **The Coyote Universe + Extension**

  (Heitmann et al. 2009, 2010, 2013, Lawrence et al. 2010)


- **Emulators beyond P(k)**

  (Kwan et al. 2013a,b)



- **The Mira-Titan Universe**

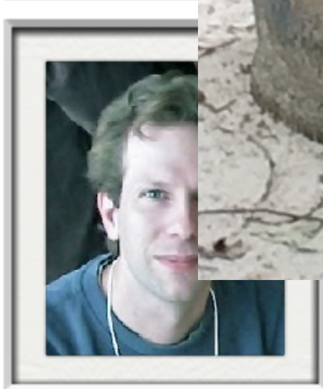
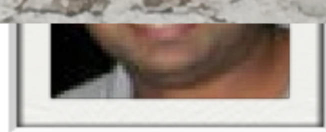
  (Heitmann et al. 2015, Lawrence et al. 2017, Kwan et al., Bocquet et al in prep.)

+ HACC team

ators!





+ HACC team

e Beginnings -- Proof of Concept

eitmann et al. 2006, Habib et al. 2007)

e Coyote Universe + Extension

eitmann et al. 2009, 2010, 2013,
wrence et al. 2010)

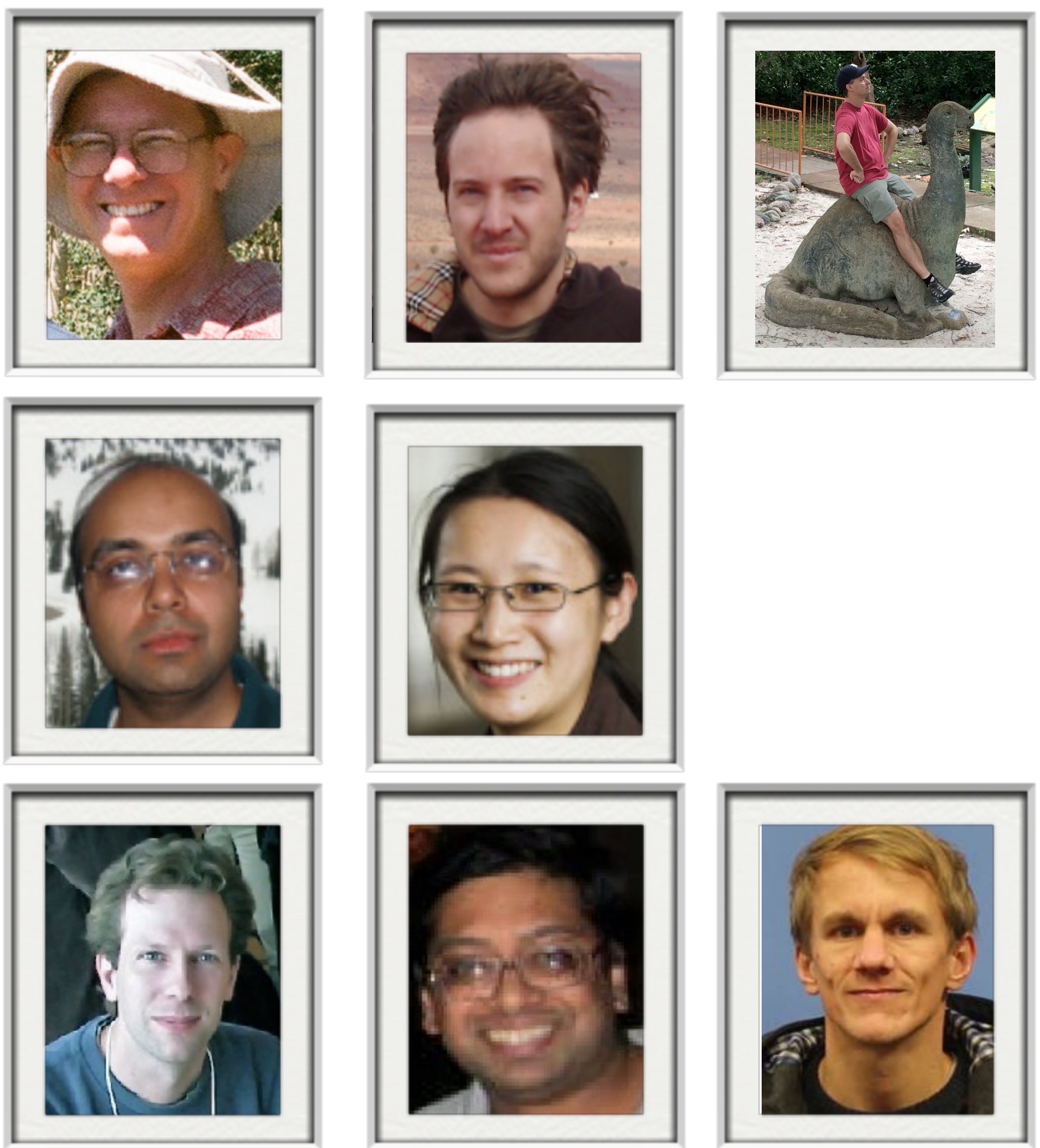
ulators beyond P(k)

an et al. 2013a,b)

e Mira-Titan Universe

itmann et al. 2015, Lawrence et al.
7, Kwan et al., Bocquet et al in prep.)

Thanks to many collaborators!



+ HACC team

- **The Beginnings -- Proof of Concept**
  (Heitmann et al. 2006, Habib et al. 2007)

- **The Coyote Universe + Extension**
  (Heitmann et al. 2009, 2010, 2013, Lawrence et al. 2010)

- **Emulators beyond P(k)**
  (Kwan et al. 2013a,b)

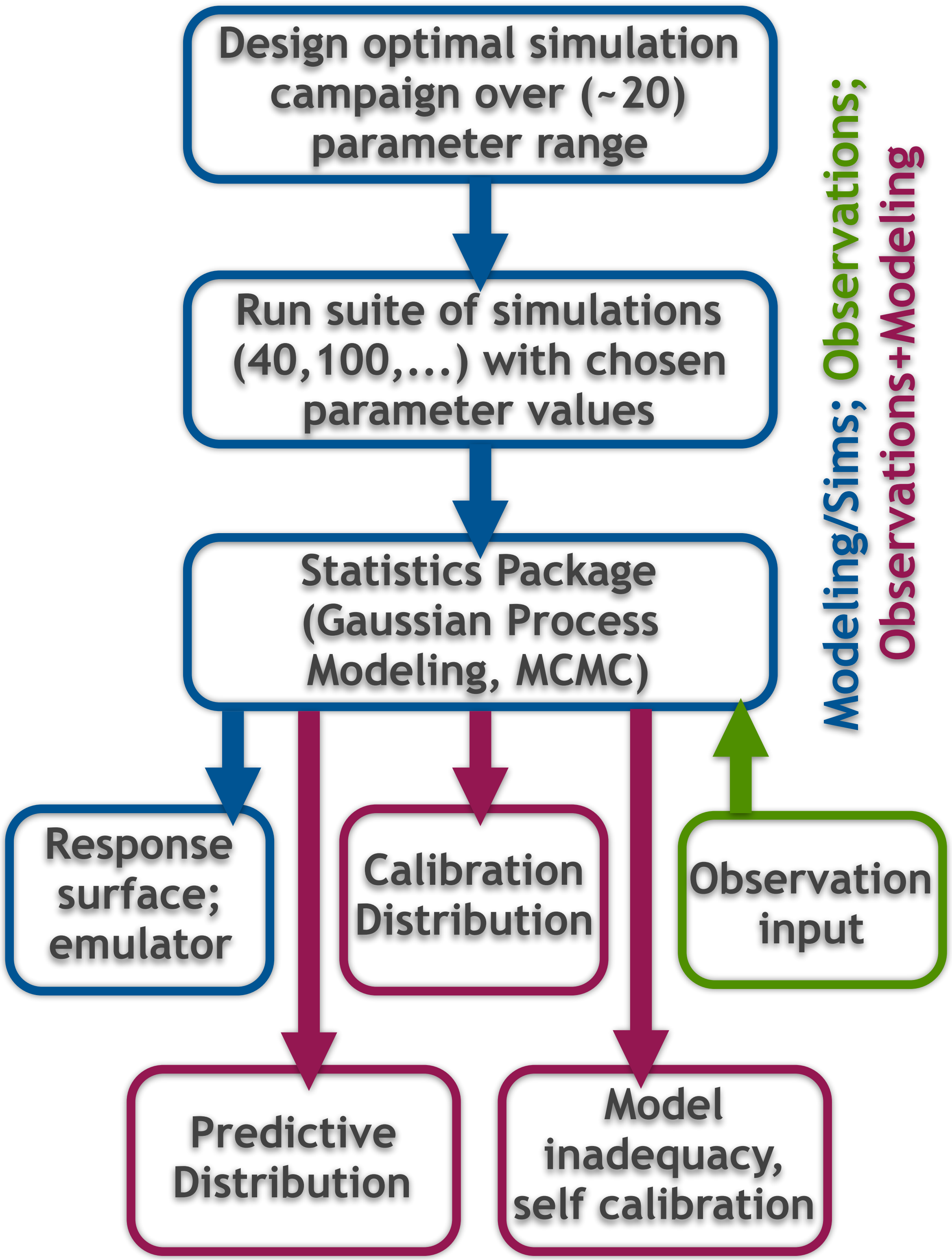
- **The Mira-Titan Universe**
  (Heitmann et al. 2015, Lawrence et al. 2017, Kwan et al., Bocquet et al in prep.)

Cosmic Calibration: Solving the Inverse Problem

- **Challenge:** To extract cosmological constraints from observations in nonlinear regime, need to run Marko Chain Monte Carlo code; input: 10,000 - 100,000 different models
- **Direct simulations:** Cost estimate
  - HACC on Titan (fastest system in the US, 5th in the world)
  - 10 simulations fit on full machine, 24 hours per simulation
  - For 100,000 simulations this translates to ~30 years
- **Current strategy:** Fitting functions for e.g. P(k), accurate at 10% level, this is not good enough!
- **Our alternative:** Emulators, fast prediction schemes built from a manageable set of simulations
- **"Ingredients":** Optimal sampling methods to decide which models to simulate, efficient representation of simulation outcome, powerful interpolation scheme
- **Example here: Power spectrum emulator**

Cosmic Calibration Framework

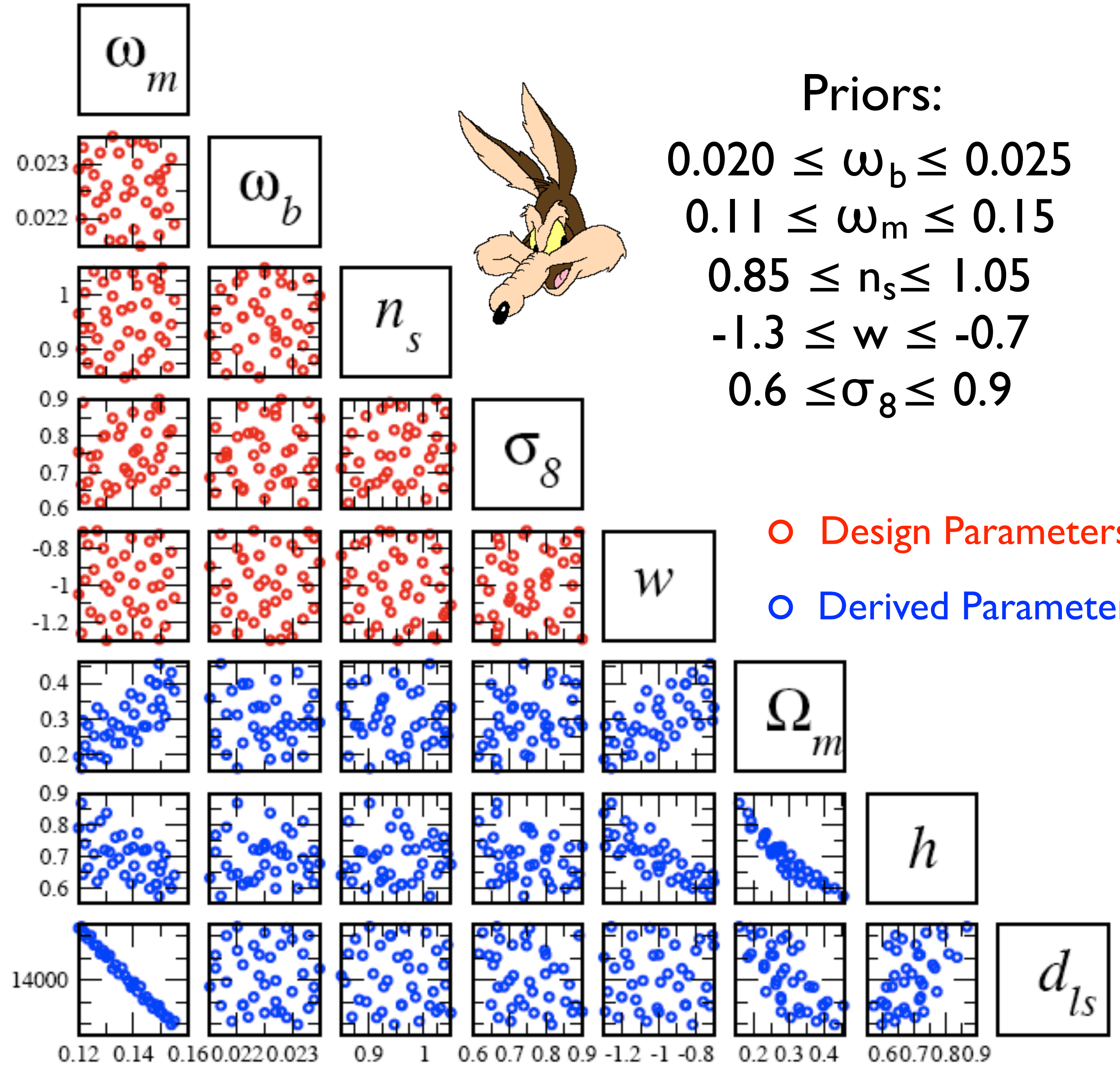
- Step 1: Design simulation campaign, rule of thumb: O(10) models for each parameter
- Step 2: Carry out simulation campaign and extract quantity of interest, in our case, power spectrum
- Step 3: Choose suitable interpolation scheme to interpolate between models, here Gaussian Processes
- Step 4: Build emulator
- Step 5: Use emulator to analyze data, determine model inadequacy, refine simulation and modeling strategy...



The Coyote Simulation Design for wCDM Cosmologies

- **Observational considerations**
  - CMB provides very accurate measurements of "vanilla parameters"
  - In particular, ω_b, ω_m, n_s known at the 2-3% level
  - w, σ_8 less well known
- **For good emulator performance from very small number of runs**
  - Not too broad priors
  - Not too many parameters

The (original) Coyote Universe



Priors:

$0.020 \le \omega_b \le 0.025$

$0.11 \le \omega_m \le 0.15$

$0.85 \le n_s \le 1.05$

$-1.3 \le w \le -0.7$

$0.6 \le \sigma_8 \le 0.9$

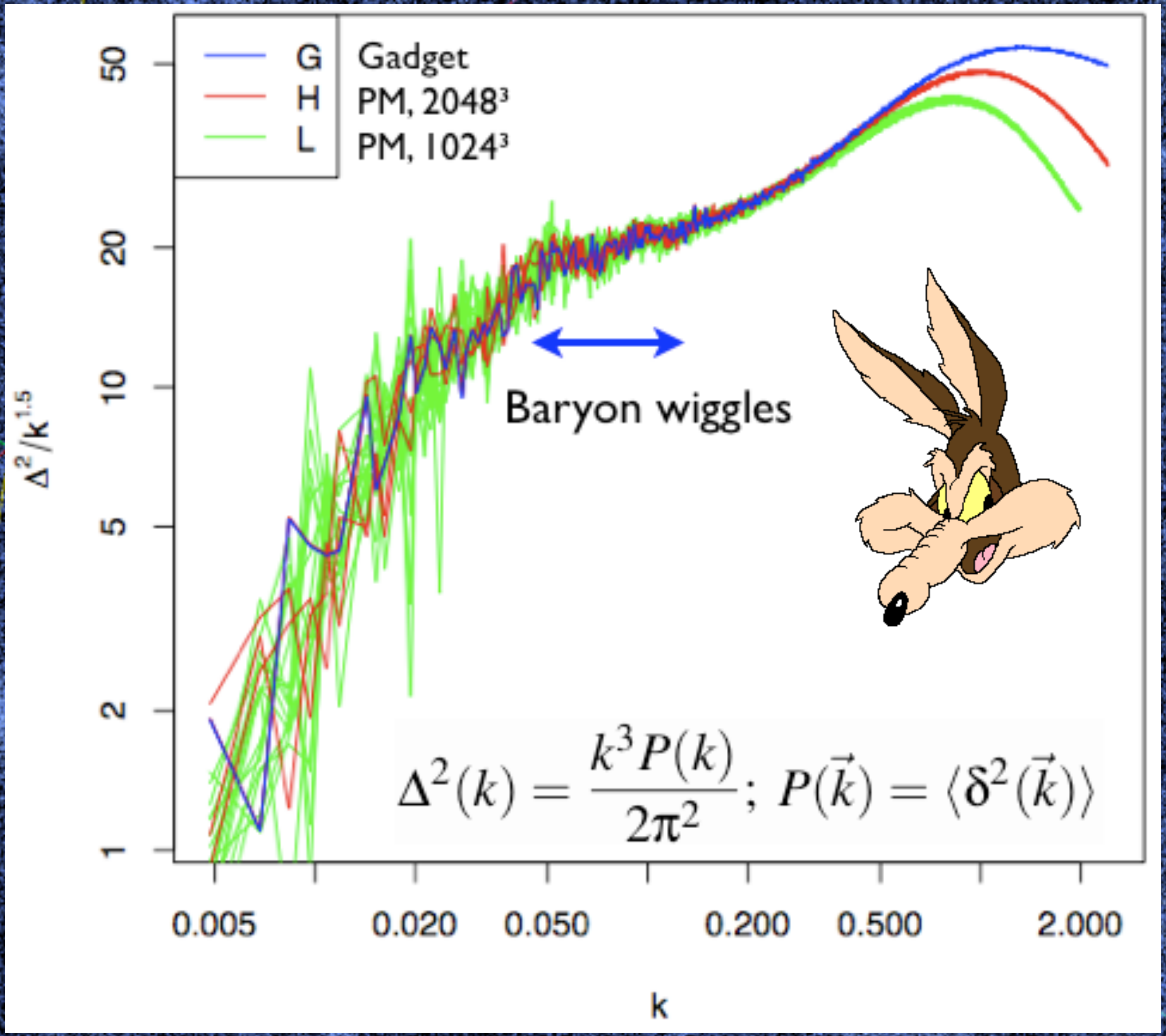
o Design Parameters

o Derived Parameters

The Coyote Universe

Priors:
$0.020 \le \omega_b \le 0.025$
$0.11 \le \omega_m \le 0.15$
$0.85 \le n_s \le 1.05$
$-1.3 \le w \le -0.7$
$0.6 \le \sigma_8 \le 0.9$

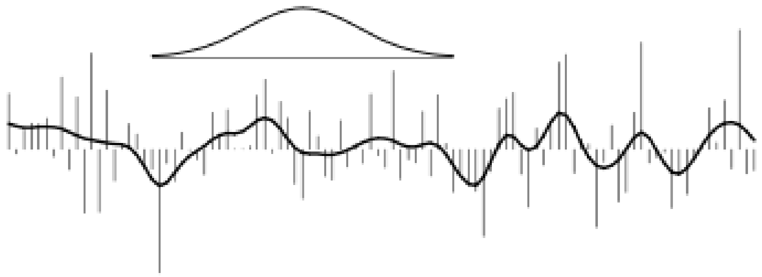
- 37 model runs + ΛCDM
  - 16 low resolution realizations (green)
  - 4 medium resolution realizations (red)
  - 1 high resolution realization (blue)
  - 11 outputs per run between z = 0 - 3
- Restricted priors to minimize necessary number of runs
- 1.3 Gpc boxes, $m_p \sim 10^{11} M_\odot$
- ~1000 simulations, 60TB

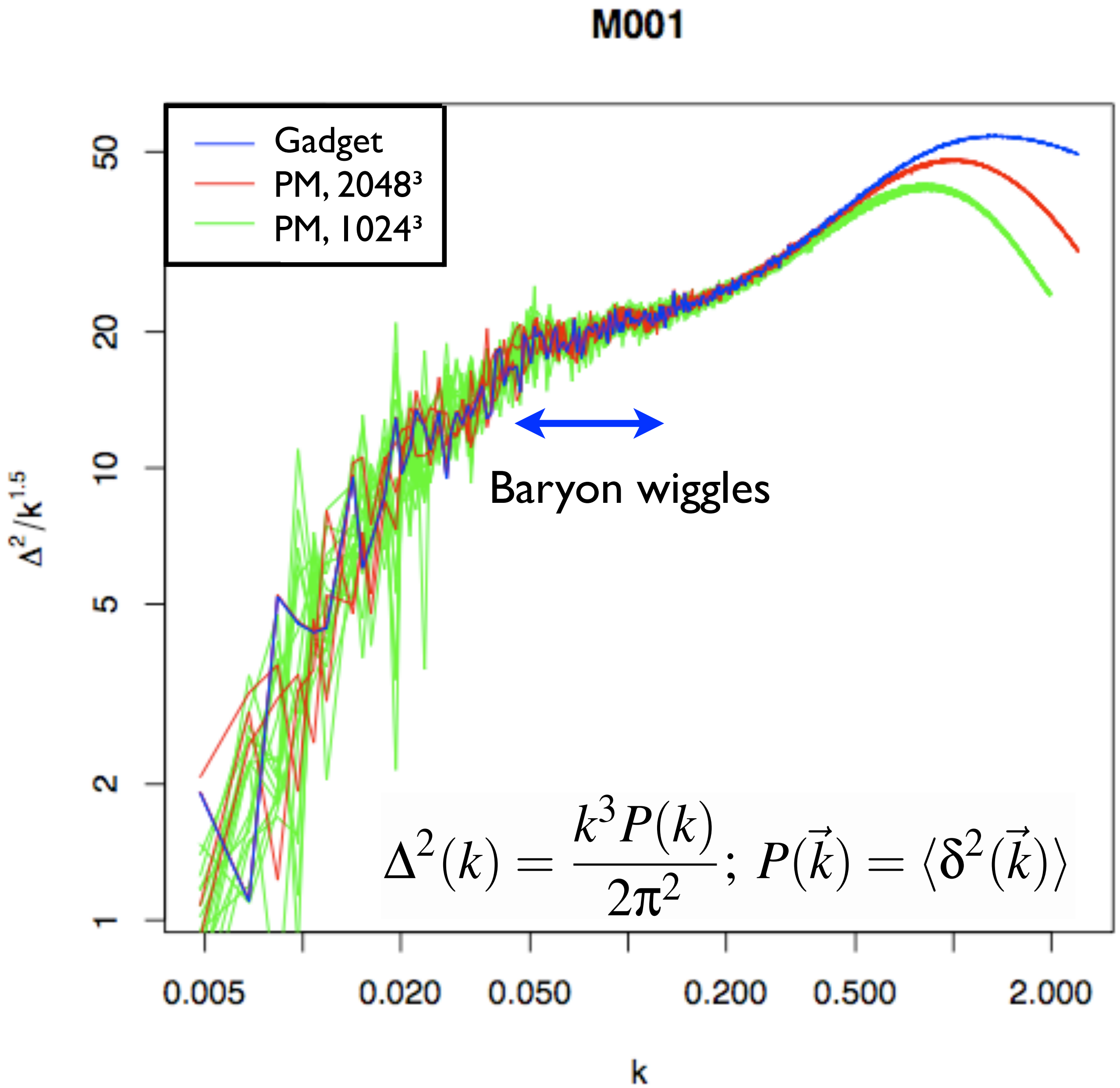


Background Visualization with ParaView by J. Woodring

Next step: Smooth Power Spectrum

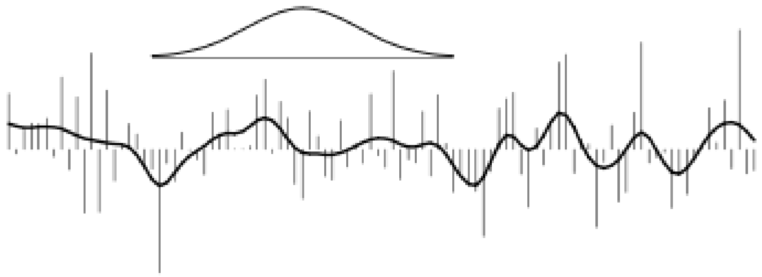
- Each simulation represents one possible realization of the Universe in a finite volume
- Need smooth prediction for building the emulator for each model
- Major challenge: Make sure that baryon features are not washed out or enhanced due to realization scatter
  - Construct smooth power spectra using a process convolution model (Higdon 2002)
  - Basic idea: calculate moving average using a kernel whose width is allowed to change to account for non-stationarity

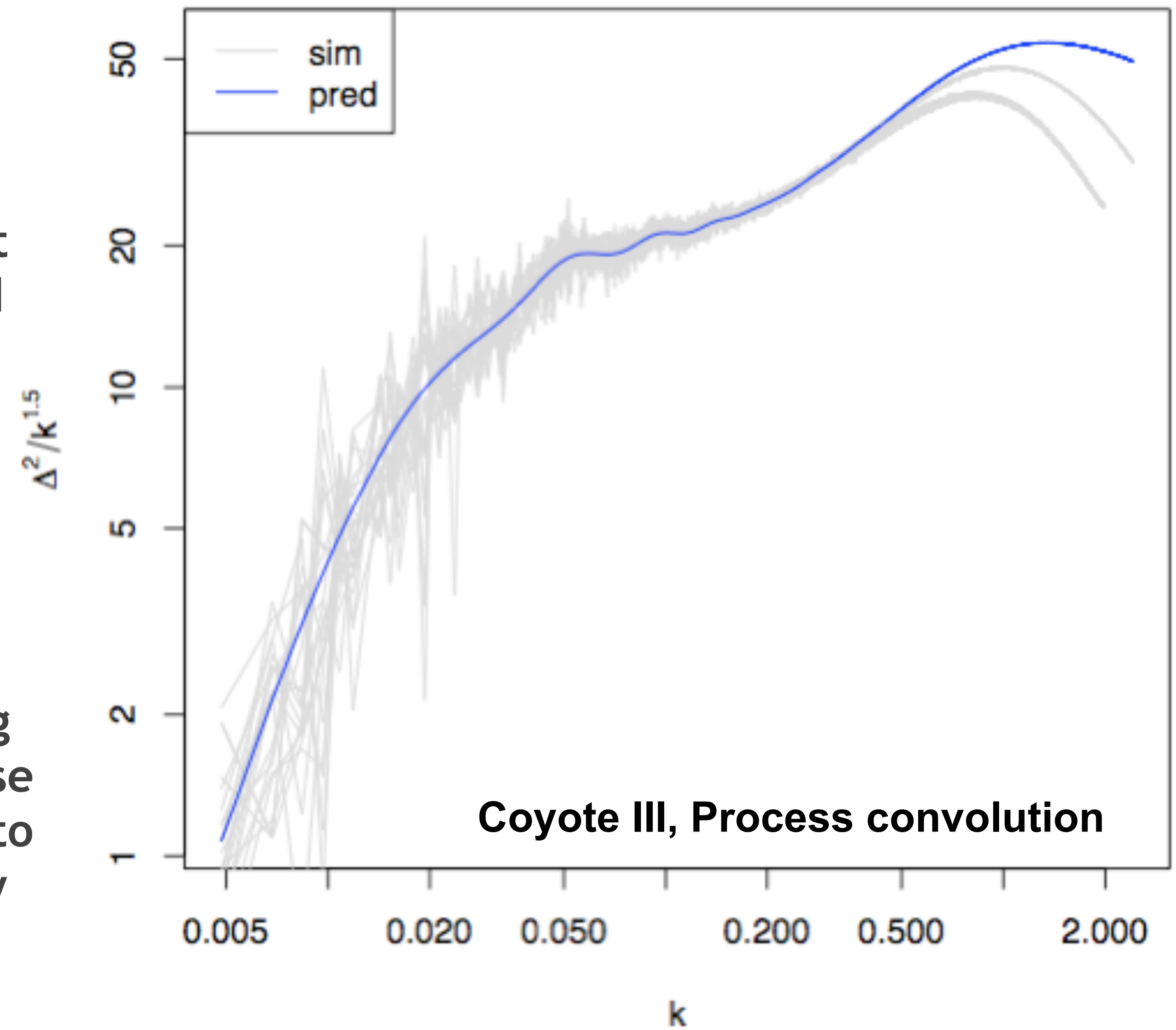




Next step: Smooth Power Spectrum

- Each simulation represents one possible realization of the Universe in a finite volume
- Need smooth prediction for building the emulator for each model
- Major challenge: Make sure that baryon features are not washed out or enhanced due to realization scatter
  - Construct smooth power spectra using a process convolution model (Higdon 2002)
  - Basic idea: calculate moving average using a kernel whose width is allowed to change to account for non-stationarity





The Interpolation Scheme: Gaussian Processing

- **After simulation design specification: Build interpolation scheme that yields predictions for any cosmology within the priors**
- **Model simulation outputs using a p_η - dimensional basis representation**
  - Find suitable set of orthogonal basis vectors $\phi_i(k,z)$, here: Principal Component Analysis
  - 5 PC bases needed, fifth PC basis pretty flat
  - Next step: modeling the weights
  - Here: Gaussian Process modeling (non-parametric regression approach, local interpolator; specified by mean function and covariance function)

$$\ln\left\{\frac{\Delta^2(k,z)}{2\pi k^{3/2}}\right\} = \sum_{i=1}^{p_\eta} \phi_i(k,z) w_i(\theta) + \epsilon \qquad \theta \in [0,1]^{p_\theta}$$

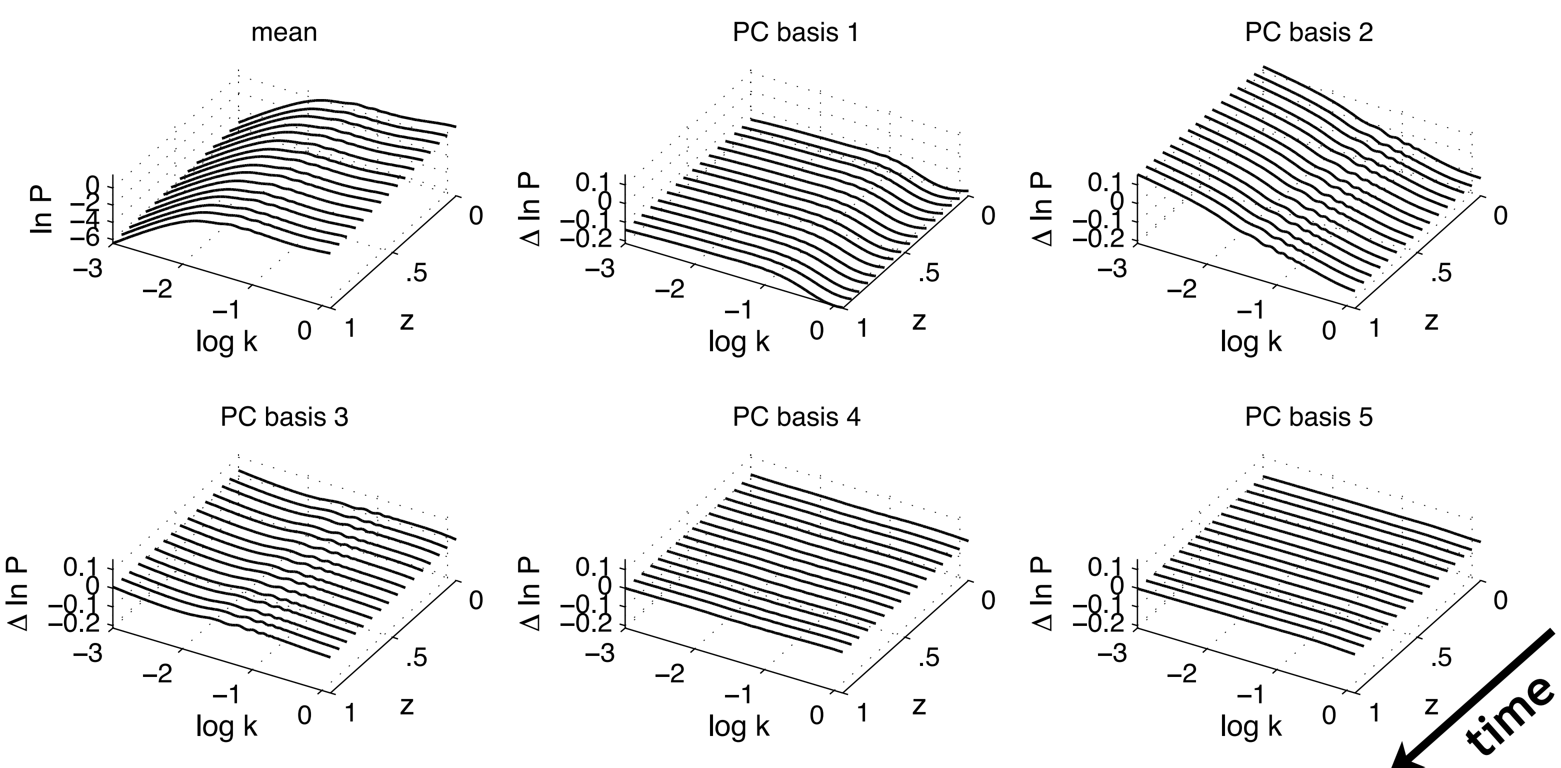
Number of basis functions, here: 5

Cosmological parameters

Number of parameters, 5

Basis functions, here: PC basis

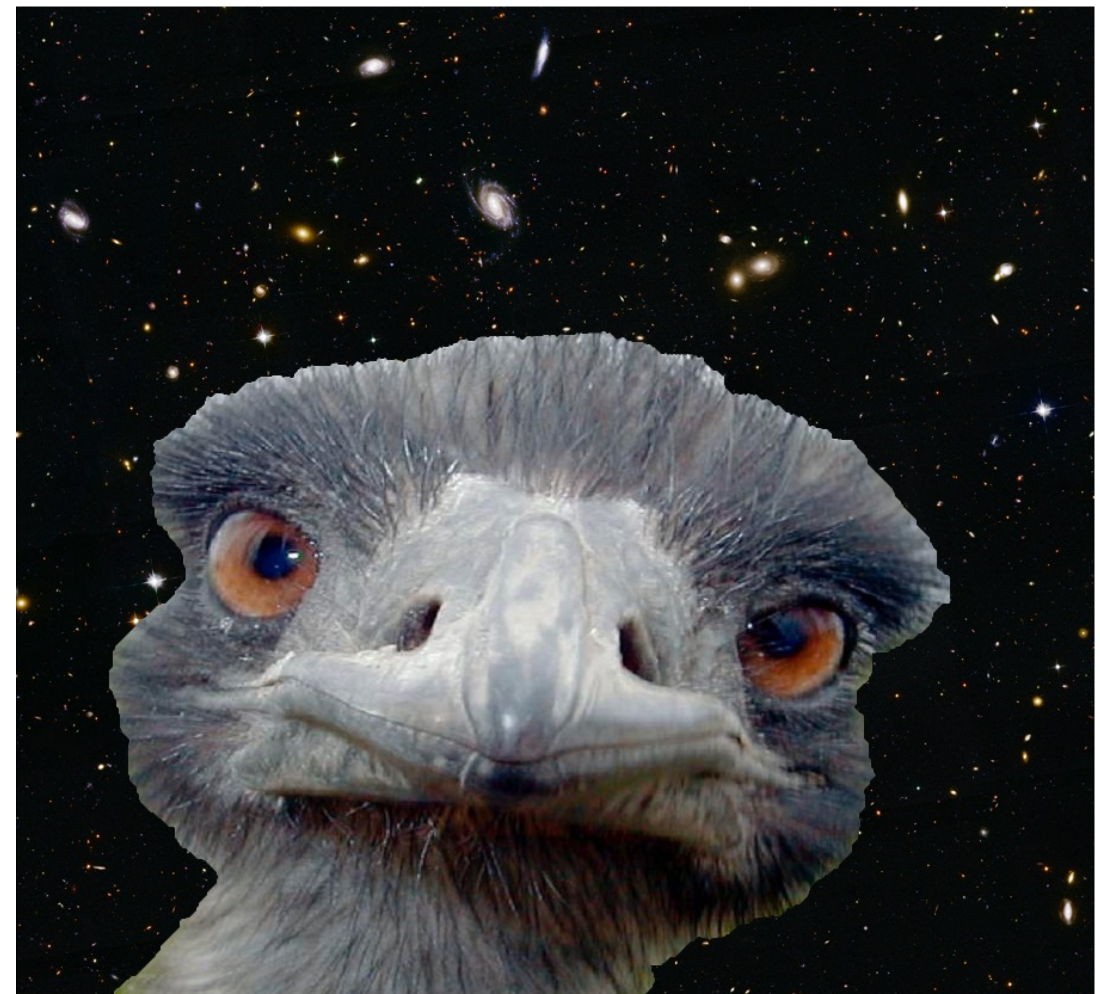
Weights, here: GP model



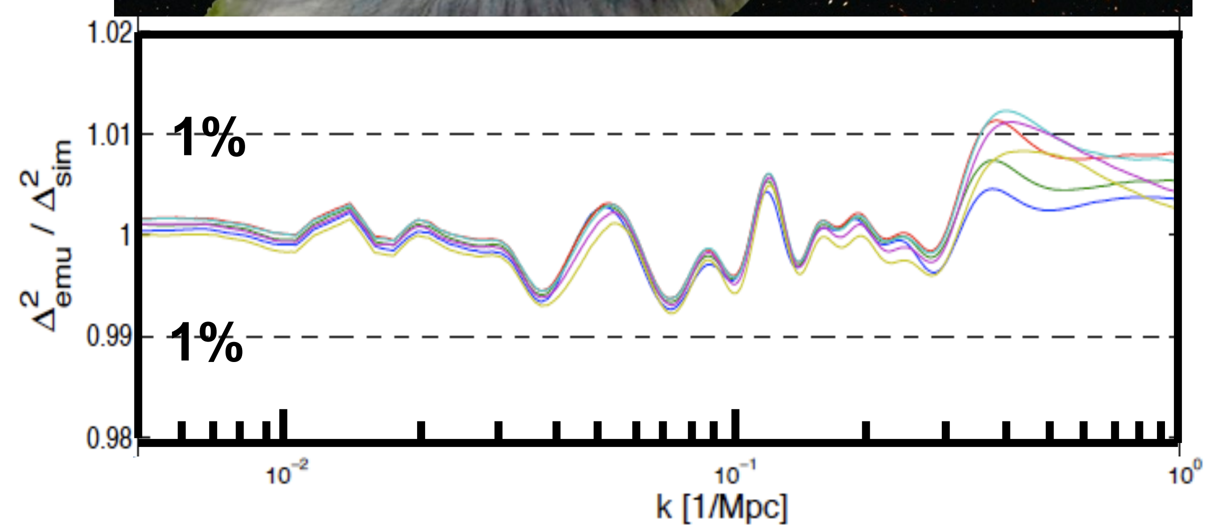
Basis functions

The Cosmic Emu(lator)

- Prediction tool for matter power spectrum has been constructed
- Accuracy within specified priors between z=0 and z=1 out to k=1 *h*/Mpc at the 1% level achieved
- Emulator has been publicly released, C code (Lawrence et al., 2010)
- Extension: Includes a additional parameter, covers smaller scales and earlier times (Heitmann et al., 2014)
  - Nested simulations to cover large k-range
  - Approach degrades accuracy to ~3%



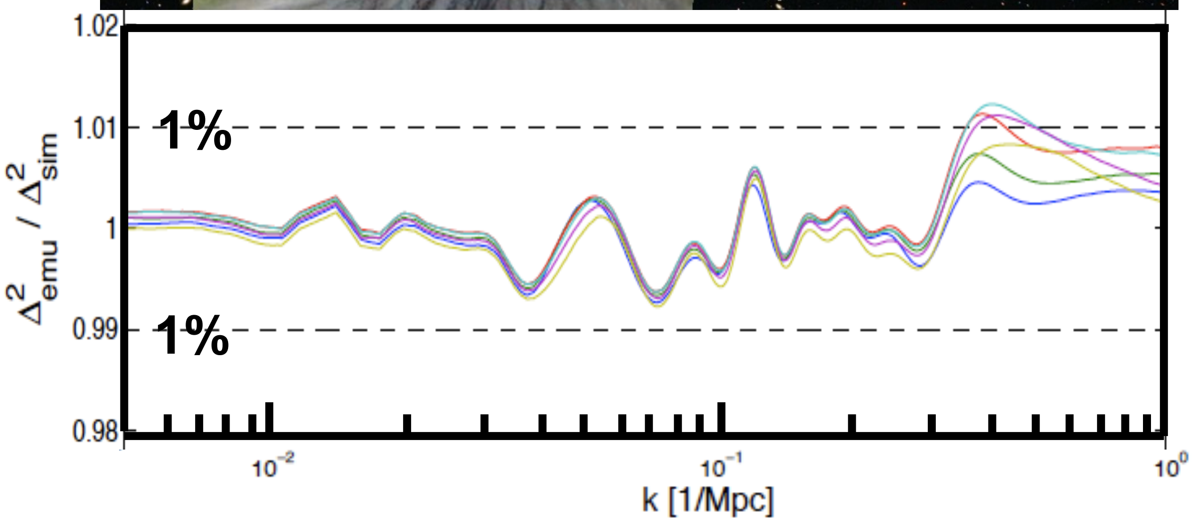
Emulator performance: Comparison of prediction and simulation output for a model not used to build emulator at 6 redshifts.



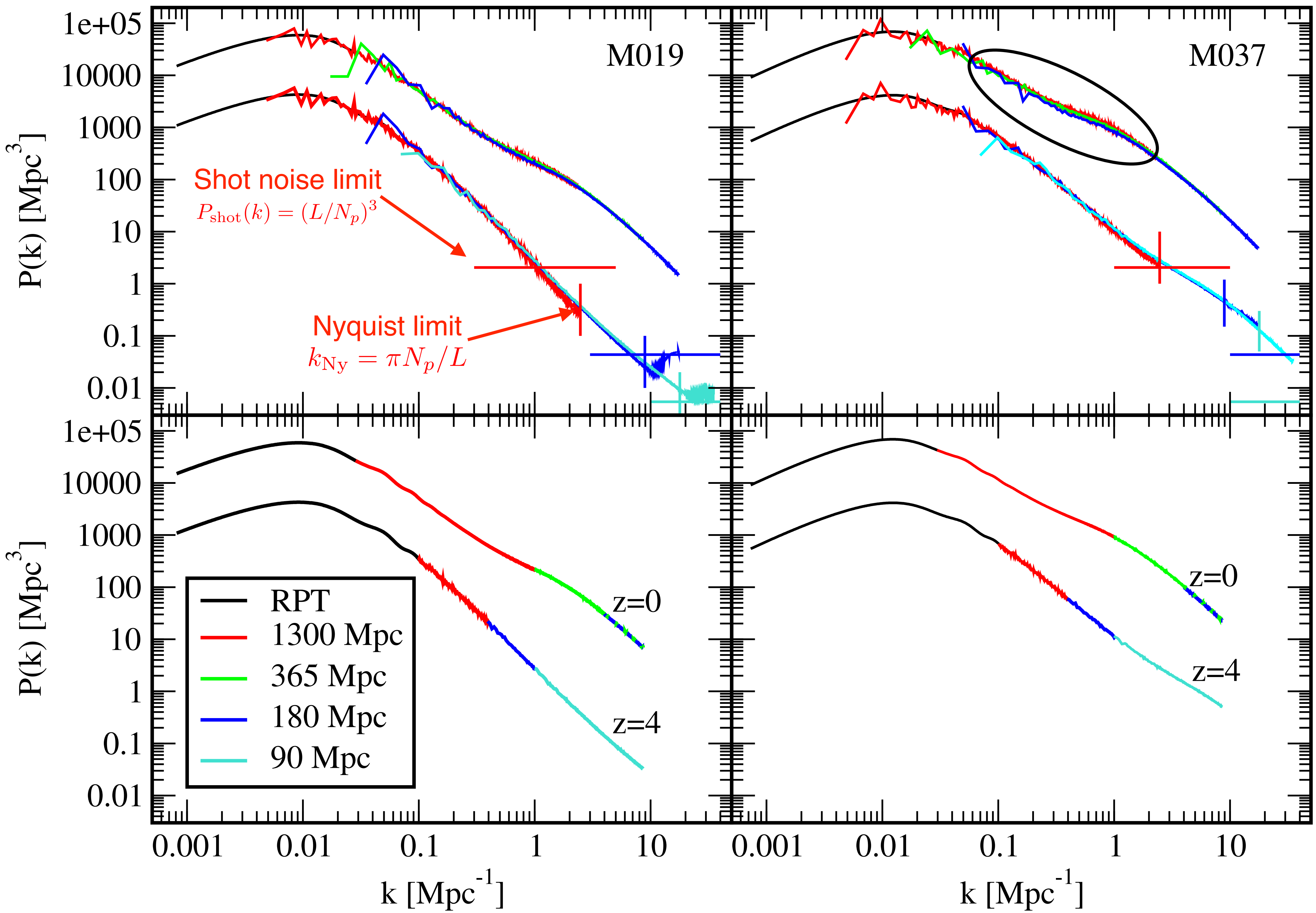
The Cosmic Emu(lator)

- Prediction tool for matter power spectrum has been constructed
- Accuracy within specified priors between z=0 and z=1 out to k=1 *h*/Mpc at the 1% level achieved
- Emulator has been publicly released, C code (Lawrence et al., 2010)
- Extension: Includes a additional parameter, covers smaller scales and earlier times (Heitmann et al., 2014)
  - Nested simulations to cover large k-range
  - Approach degrades accuracy to ~3%

Emulator performance: Comparison of prediction and simulation output for a model not used to build emulator at 6 redshifts.

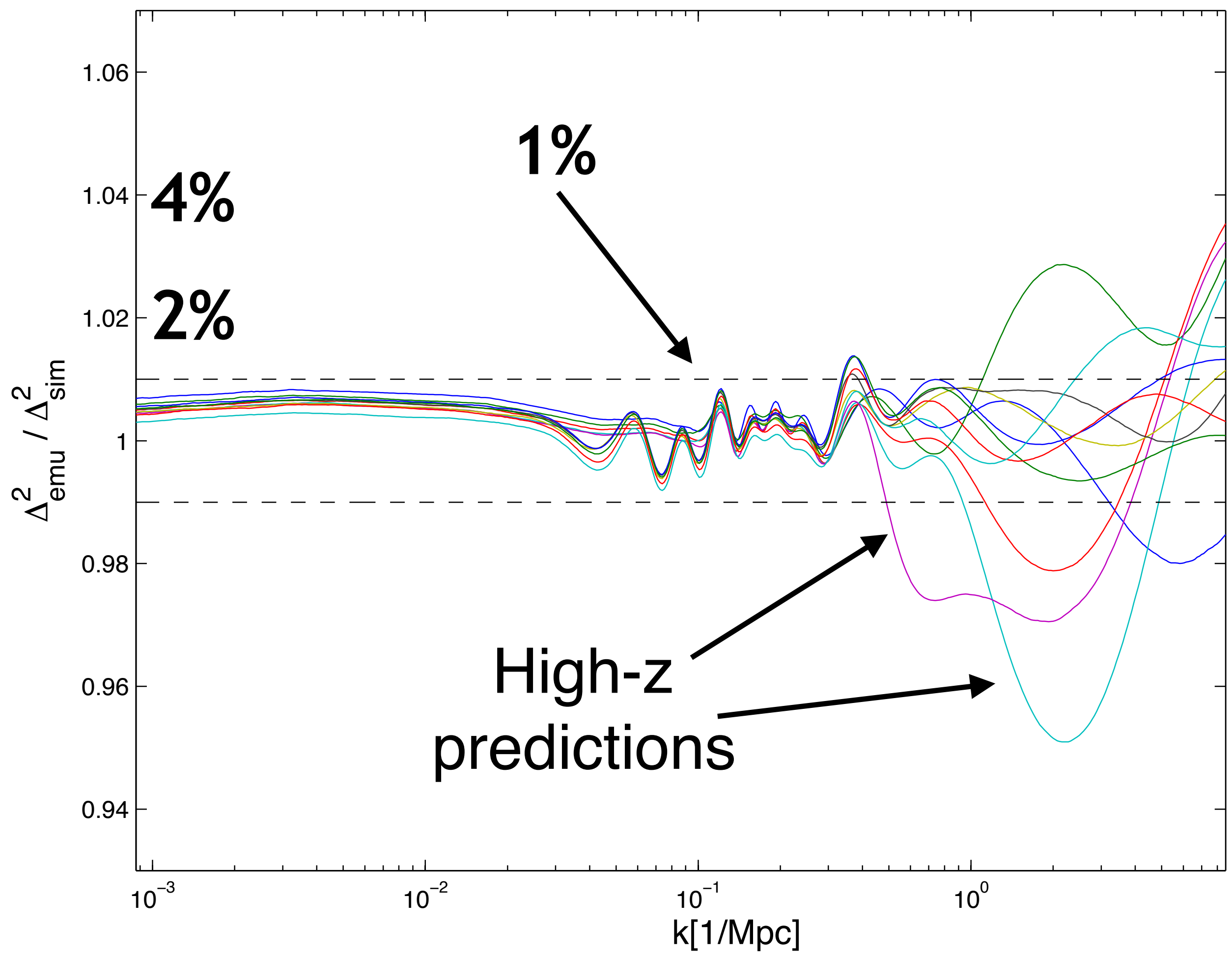


The FrankenEmu Challenge

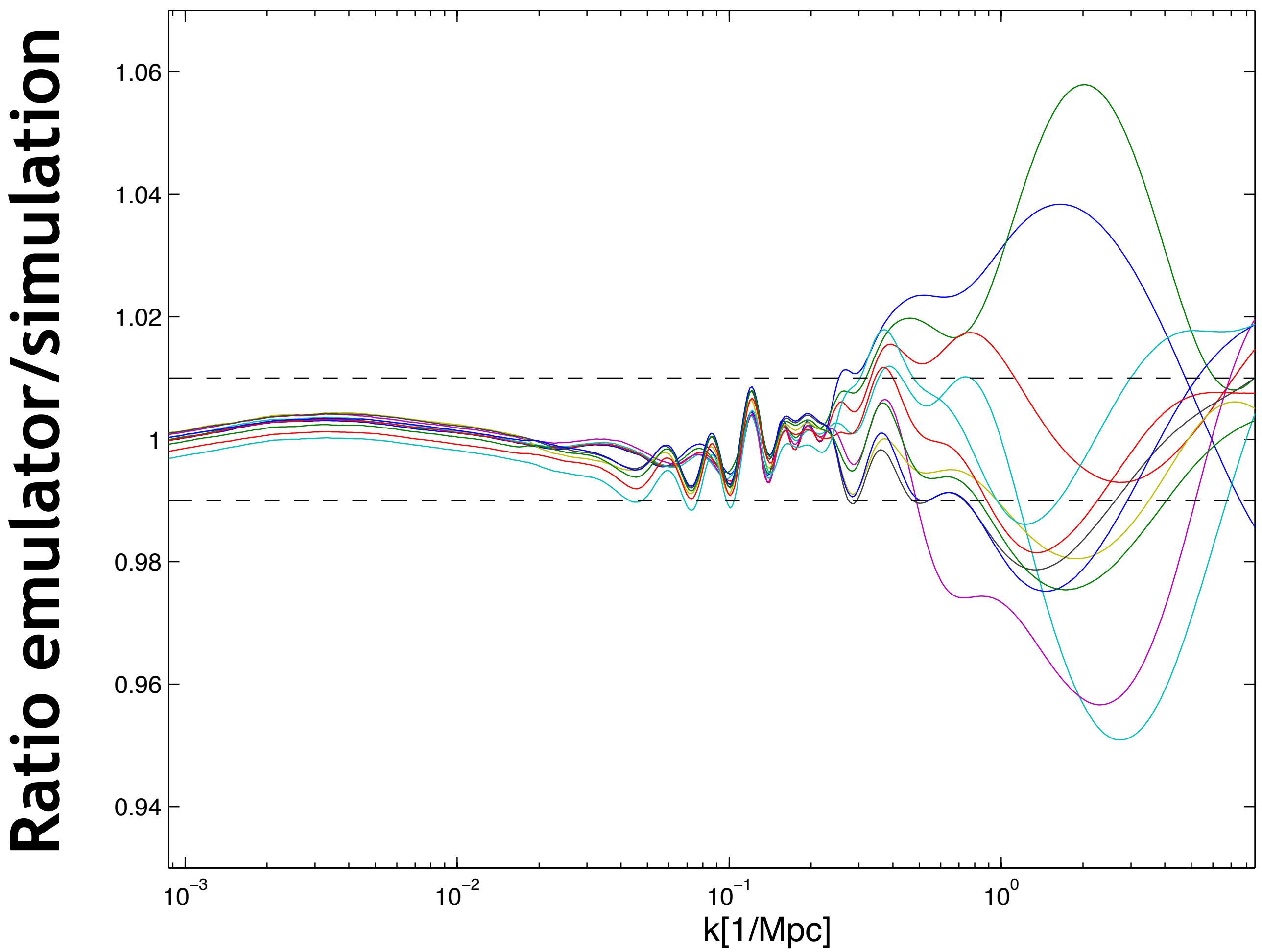


FrankenEmu Results

Extension in k- and z-range

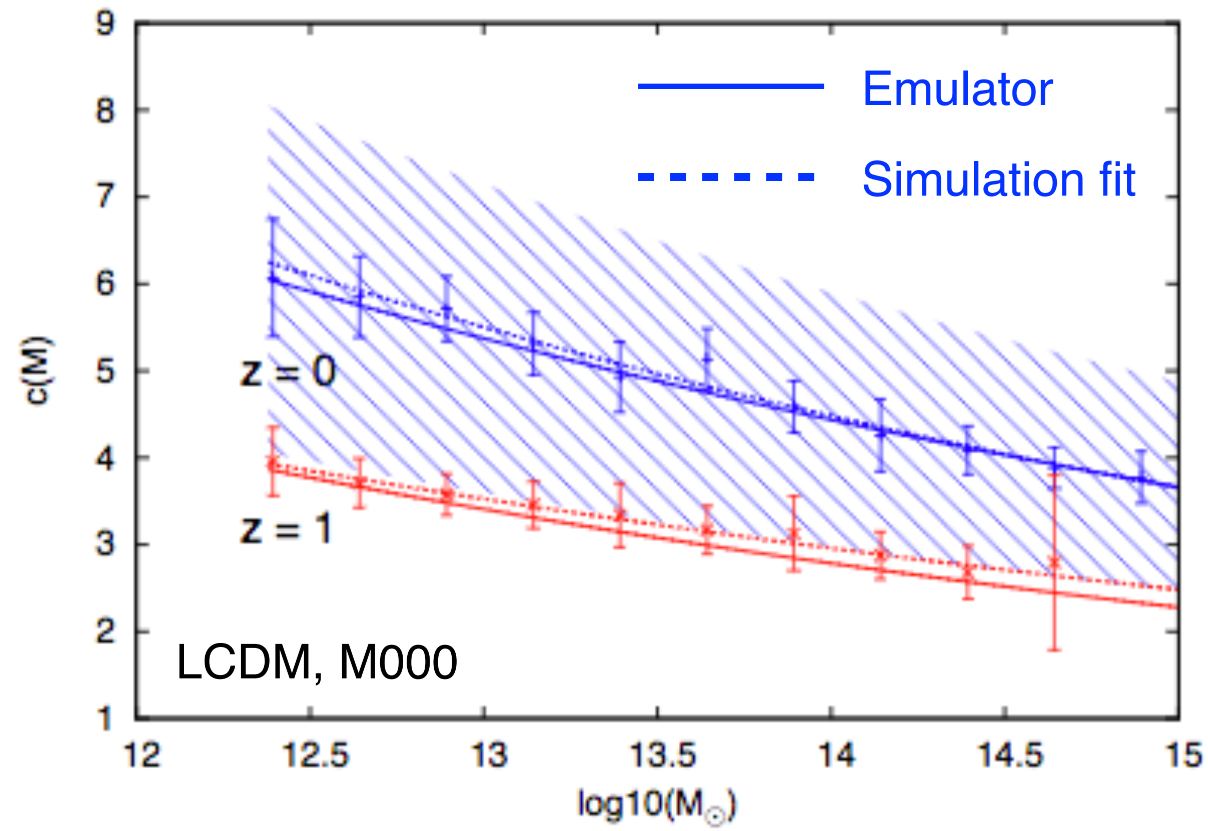


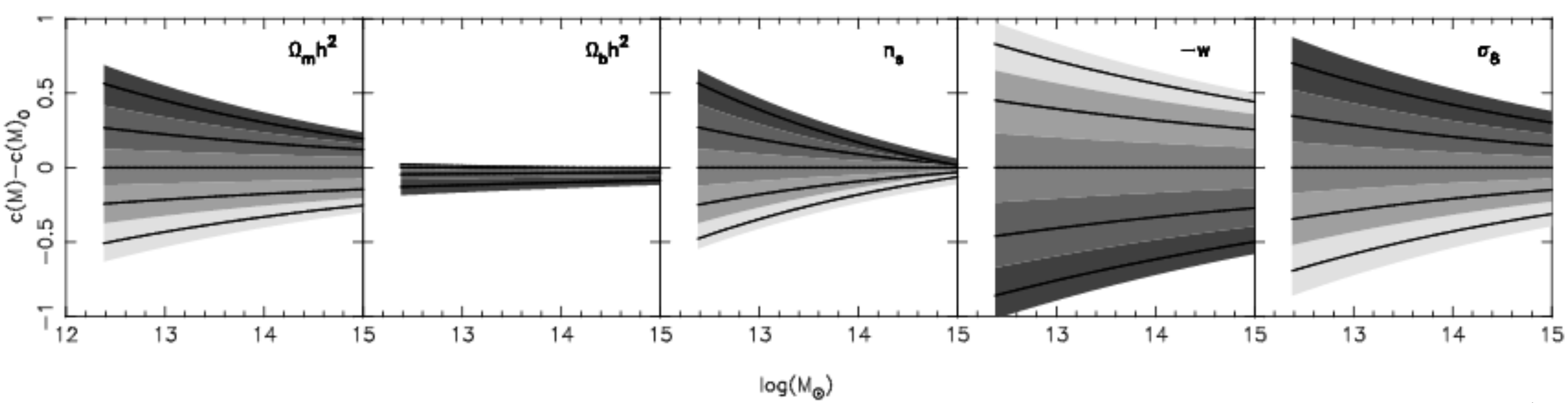
Hubble treated as free parameter



FrankenEmu Concentration Emulator

- Nested simulation provide high mass resolution and allow us to measure halo concentration for small halo masses
- Due to large variance in concentration measures, accuracy requirements are not as daunting
- Emulator for z-range 0-1, concetration variation between c~2 to c~8





Kwan et al. 2013

The Next Step: The Mira-Titan Universe

The Next Step: The Mira-Titan Universe

The Next Step: The Mira-Titan Universe

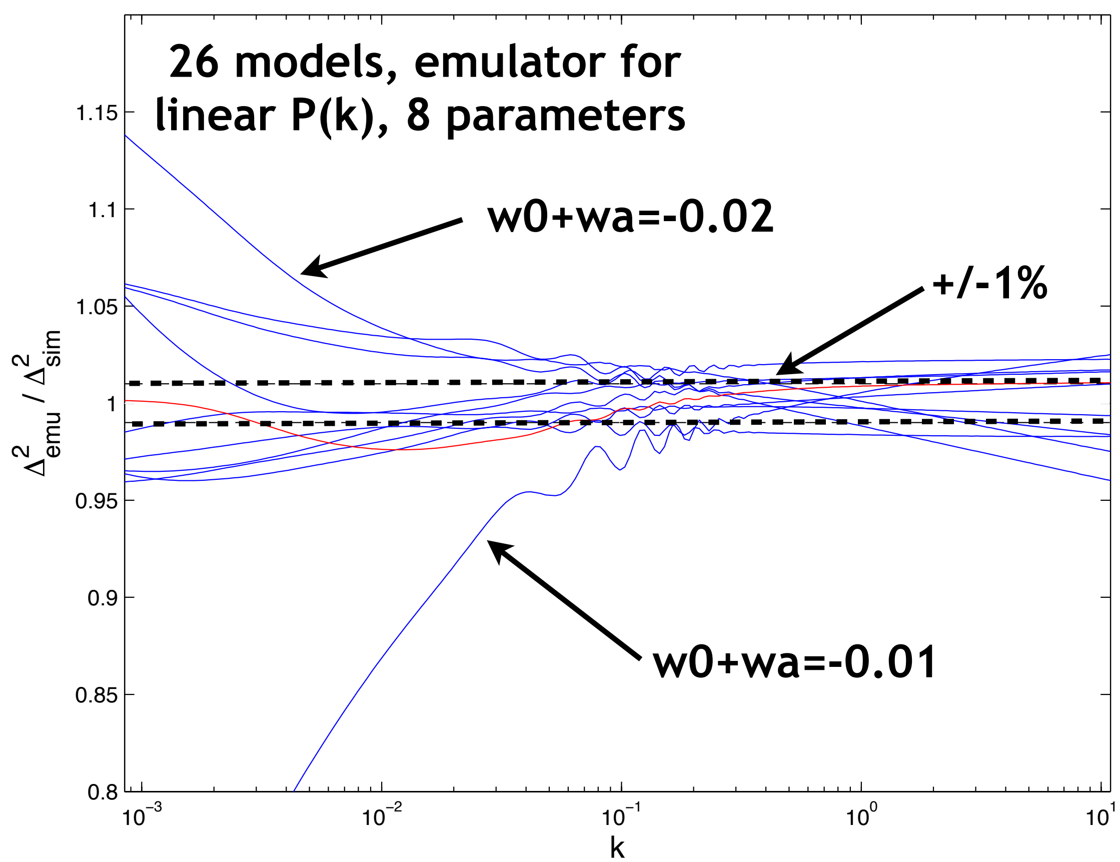
- Extend parameter space to include varying w(z) and massive neutrinos
- Build "nested designs": enable to build emulator from first set of 26 models, improve with additional 29 models, final precision with 101 models overall
- Various emulators for P(k), mass function, c-M relation, RSD predictions, derived quantities...

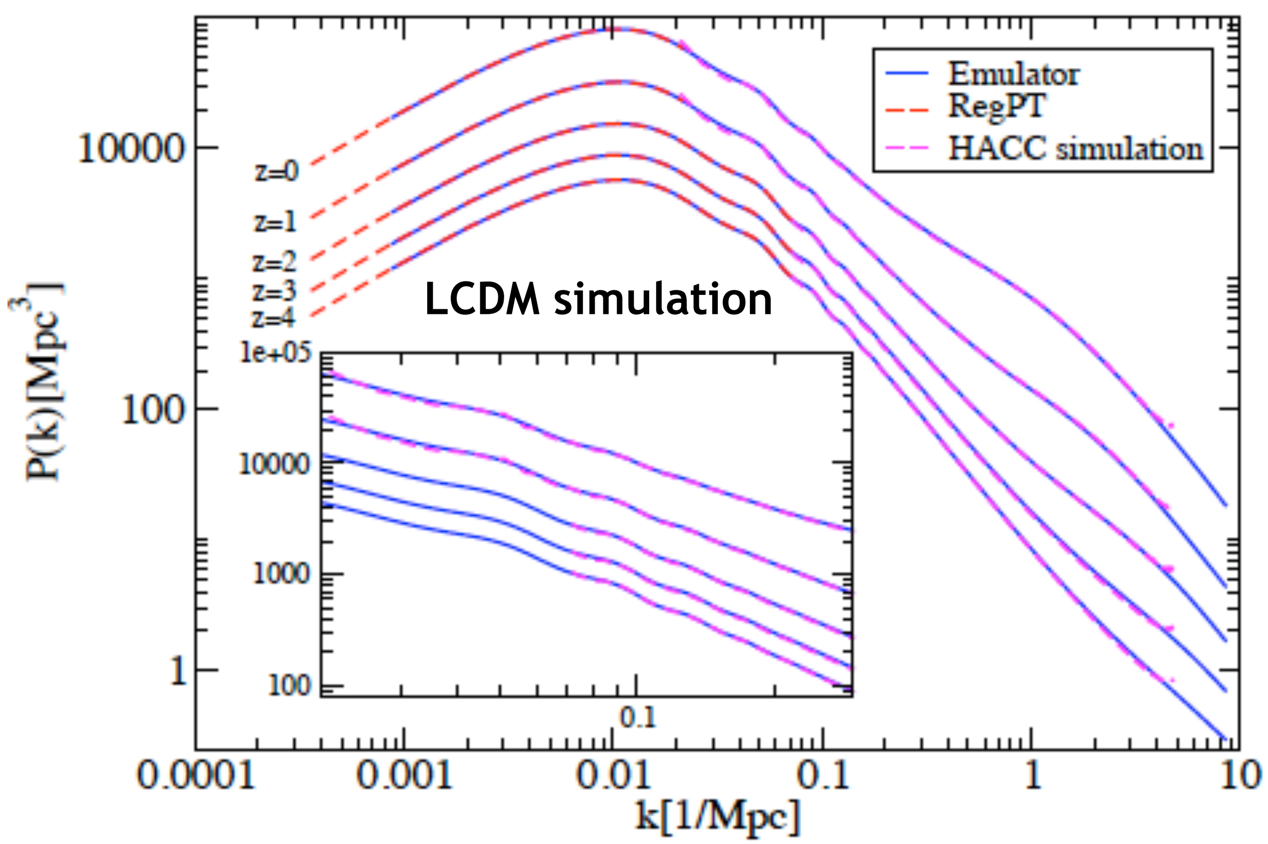
Key advance, Bergner PhD thesis, SFU, 2011

Parameters

Heitmann et al. 2015, proxy models

$$
\begin{aligned}
0.12 &\le \omega_m \le 0.155 \\
0.0215 &\le \omega_b \le 0.0235 \\
0.7 &\le \sigma_8 \le 0.9 \\
0.55 &\le h \le 0.85 \\
0.85 &\le n_s \le 1.05 \\
-1.3 &\le w_0 \le -0.7 \\
-1.5 &\le w_a \le 1.15 \\
0.0 &\le \omega_\nu \le 0.01.
\end{aligned}
$$





The Next Step: The Mira-Titan Universe

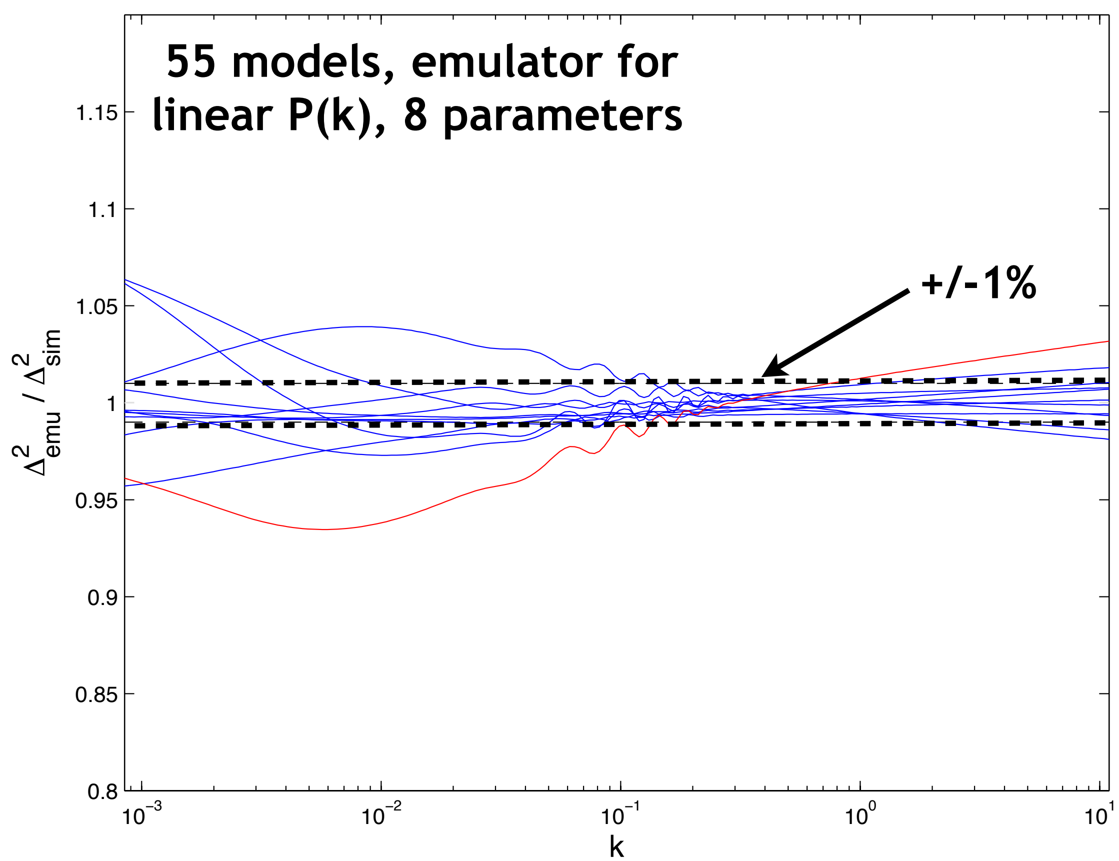
- Extend parameter space to include varying w(z) and massive neutrinos
- Build "nested designs": enable to build emulator from first set of 26 models, improve with additional 29 models, final precision with 101 models overall
- Various emulators for P(k), mass function, c-M relation, RSD predictions, derived quantities...

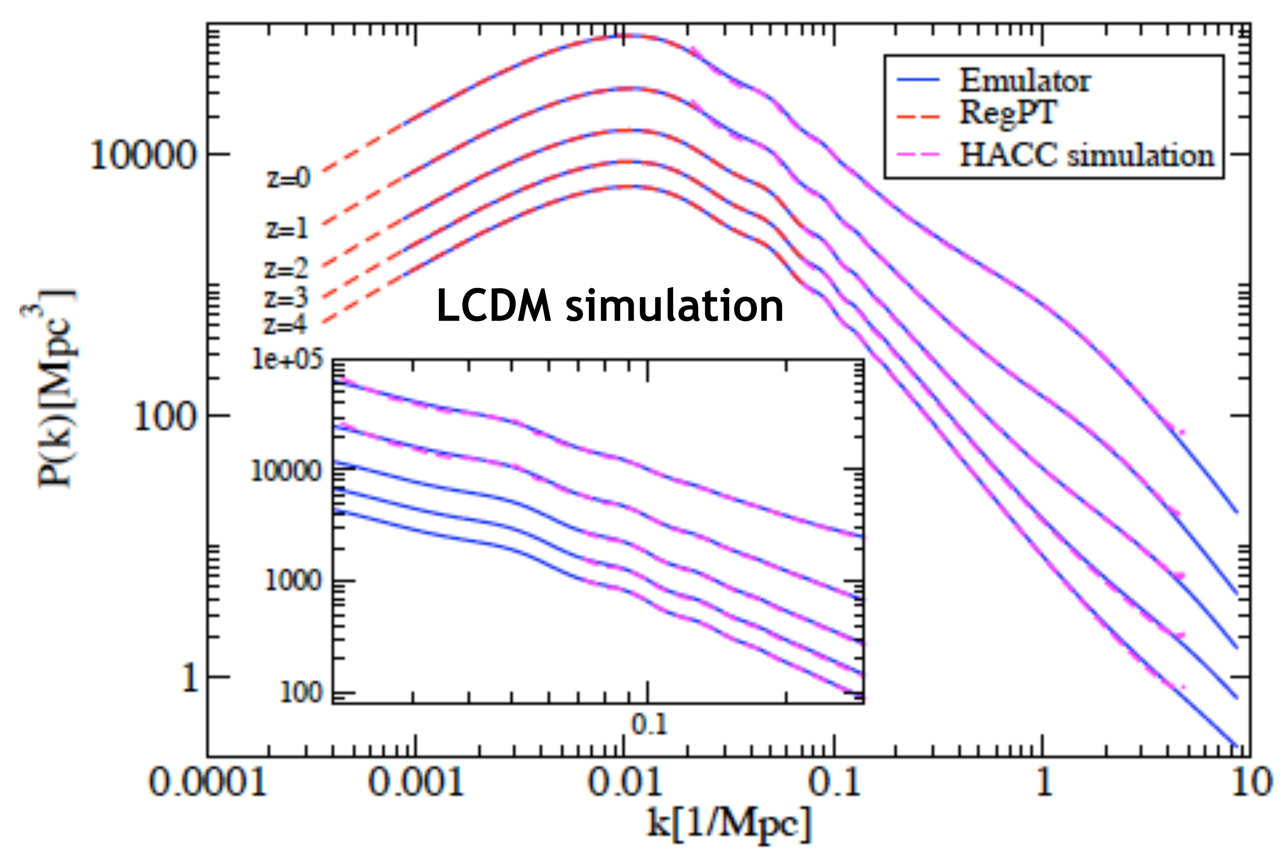
Key advance, Bergner PhD thesis, SFU, 2011

Parameters

$$0.12 \leq \omega_m \leq 0.155$$
$$0.0215 \leq \omega_b \leq 0.0235$$
$$0.7 \leq \sigma_8 \leq 0.9$$
$$0.55 \leq h \leq 0.85$$
$$0.85 \leq n_s \leq 1.05$$
$$-1.3 \leq w_0 \leq -0.7$$
$$-1.5 \leq w_a \leq 1.15$$
$$0.0 \leq \omega_\nu \leq 0.01.$$

Heitmann et al. 2015, proxy models





The Next Step: The Mira-Titan Universe

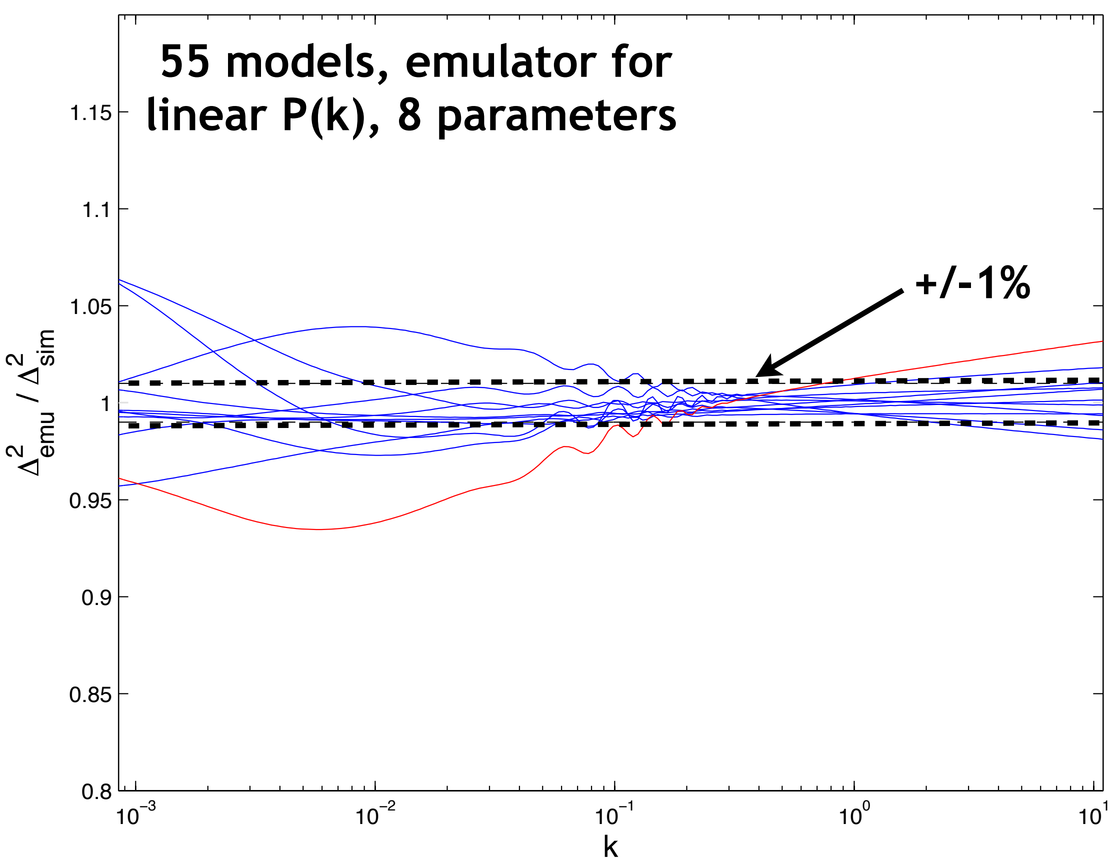
- Extend parameter space to include varying w(z) and massive neutrinos
- Build “nested designs”: enable to build emulator from first set of 26 models, improve with additional 29 models, final precision with 101 models overall
- Various emulators for P(k), mass function, c-M relation, RSD predictions, derived quantities...

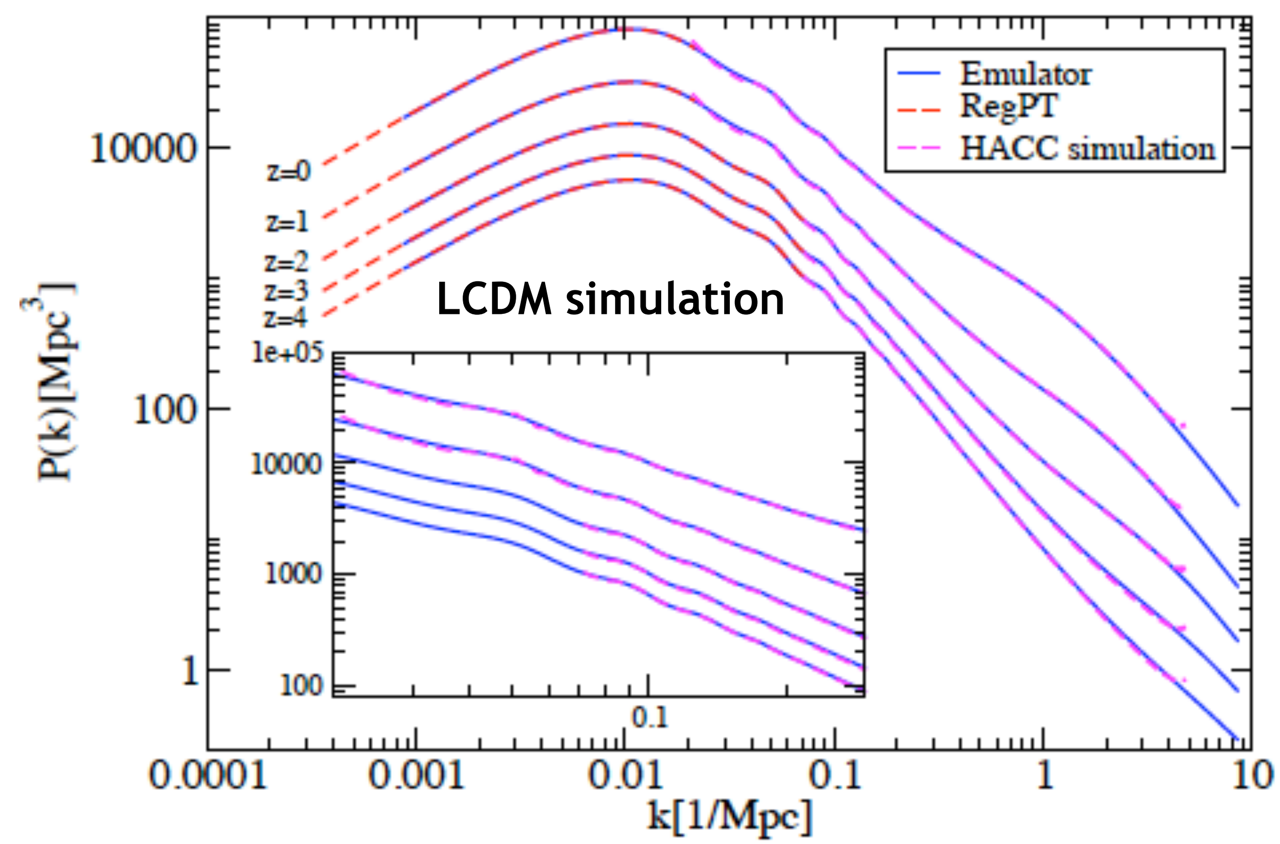
Key advance, Bergner PhD thesis, SFU, 2011

Parameters

Heitmann et al. 2015, proxy models

$$
\begin{aligned}
0.12 \leq\ & \omega_m \leq 0.155 \\
0.0215 \leq\ & \omega_b \leq 0.0235 \\
0.7 \leq\ & \sigma_8 \leq 0.9 \\
0.55 \leq\ & h \leq 0.85 \\
0.85 \leq\ & n_s \leq 1.05 \\
-1.3 \leq\ & w_0 \leq -0.7 \\
-1.5 \leq\ & w_a \leq 1.15 \\
0.0 \leq\ & \omega_\nu \leq 0.01.
\end{aligned}
$$





The Next Step: The Mira-Titan Universe

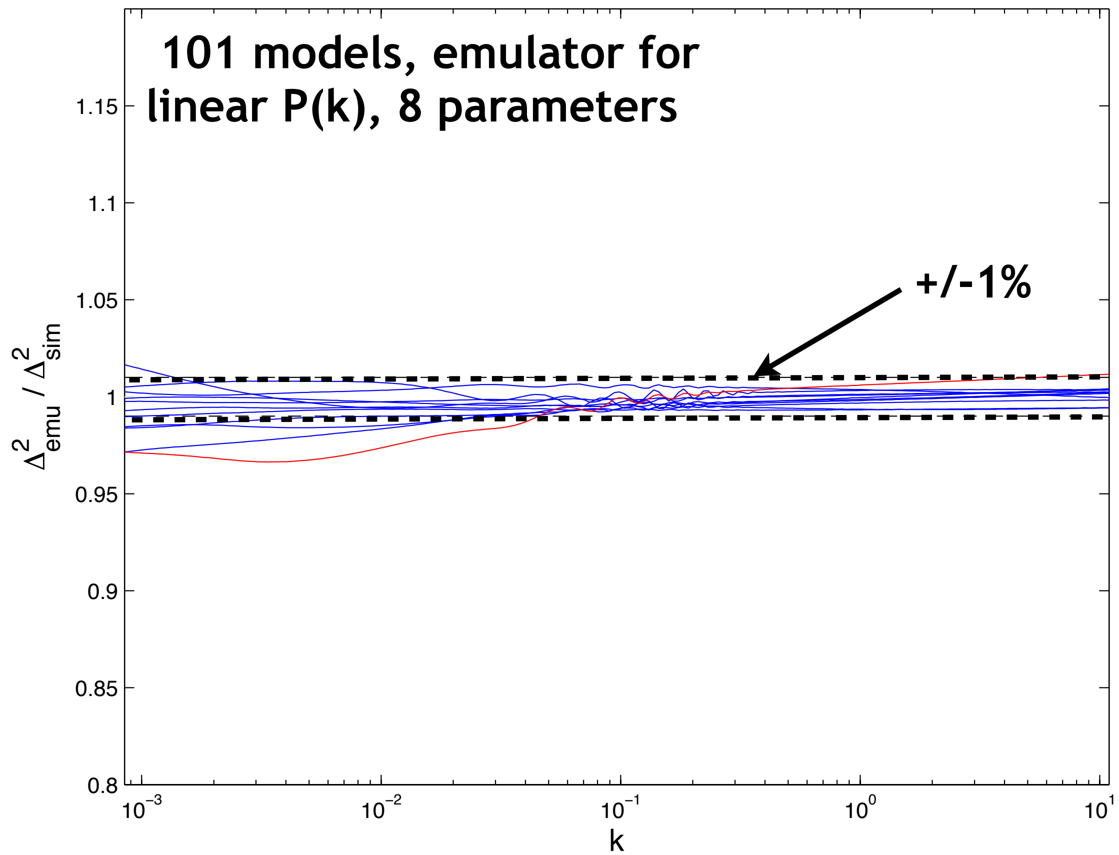
- Extend parameter space to include varying w(z) and massive neutrinos
- Build "nested designs": enable to build emulator from first set of 26 models, improve with additional 29 models, final precision with 101 models overall
- Various emulators for P(k), mass function, c-M relation, RSD predictions, derived quantities...

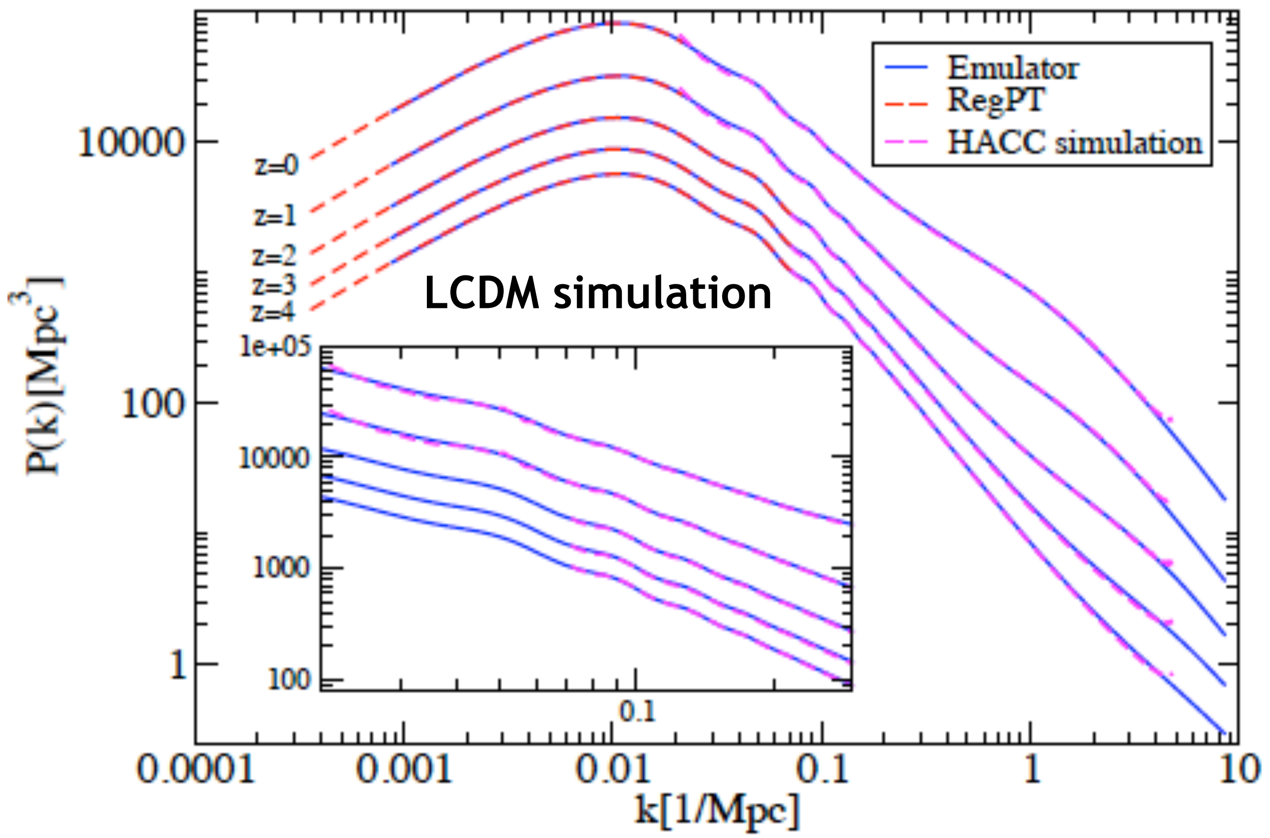
Key advance, Bergner PhD thesis, SFU, 2011

Parameters

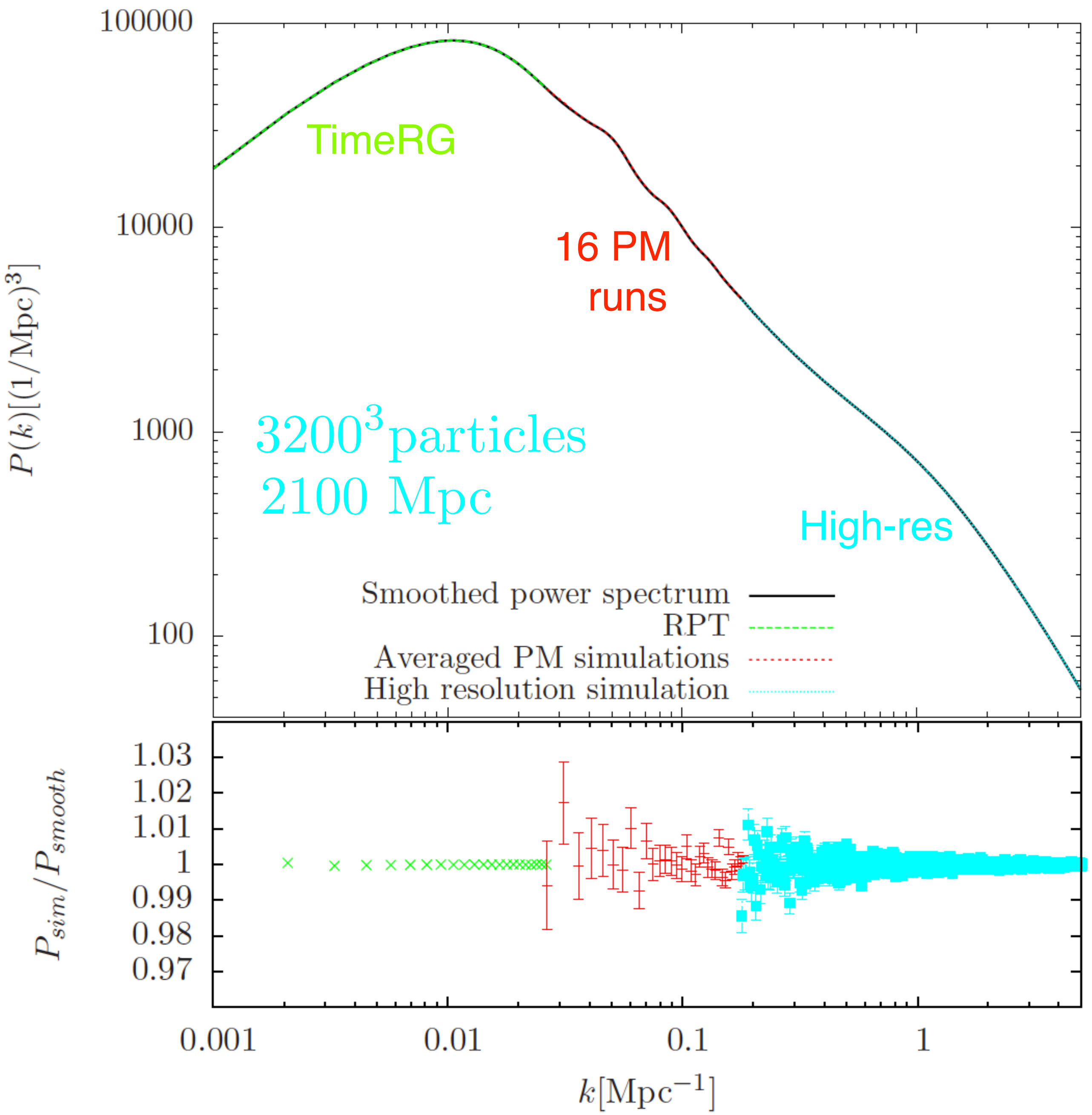
$$0.12 \leq \omega_m \leq 0.155$$
$$0.0215 \leq \omega_b \leq 0.0235$$
$$0.7 \leq \sigma_8 \leq 0.9$$
$$0.55 \leq h \leq 0.85$$
$$0.85 \leq n_s \leq 1.05$$
$$-1.3 \leq w_0 \leq -0.7$$
$$-1.5 \leq w_a \leq 1.15$$
$$0.0 \leq \omega_\nu \leq 0.01.$$

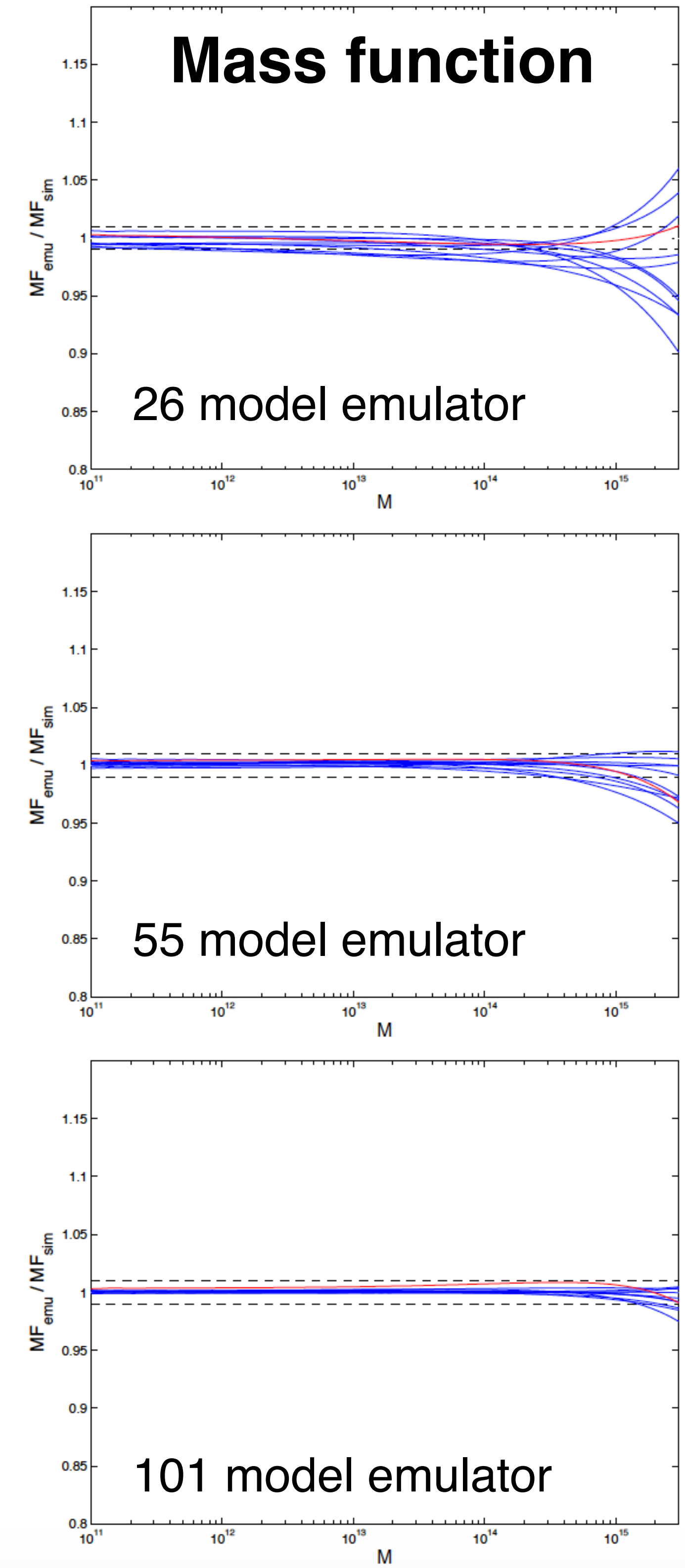
Heitmann et al. 2015, proxy models





The Next Step: The Mira-Titan Universe

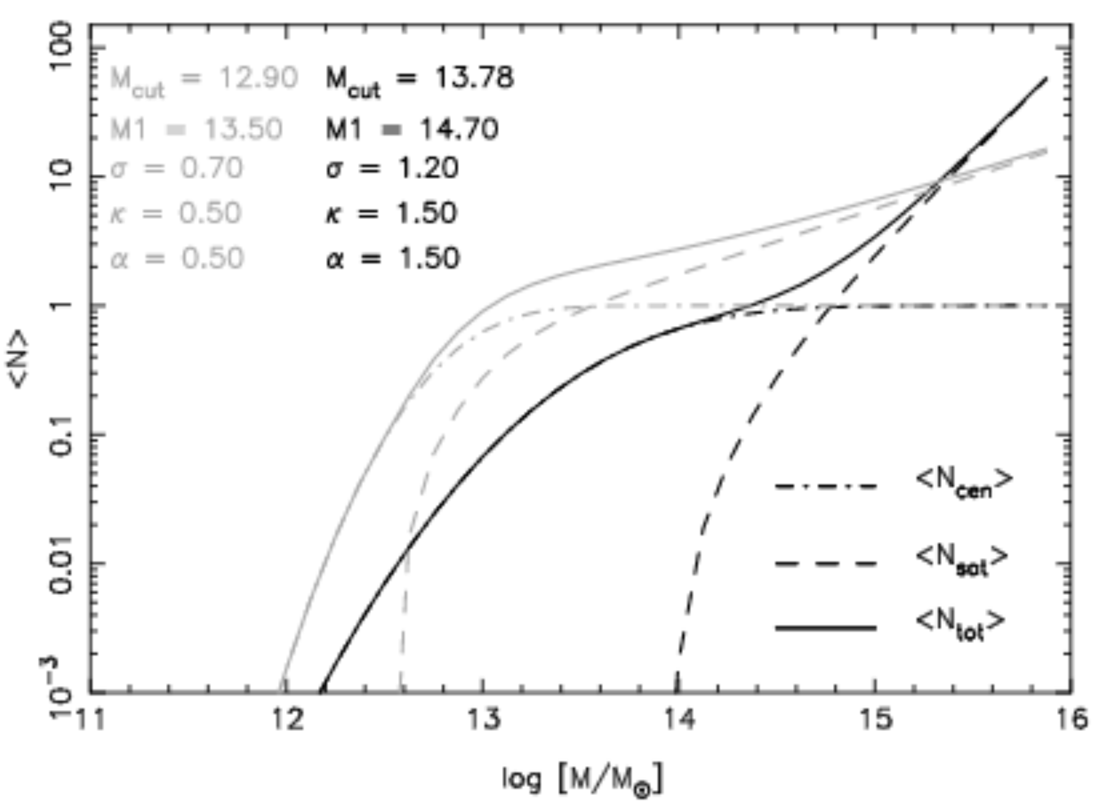




Execution terminated
Exit_status=0
resources_used.cput=00:00:30
resources_used.energy_used=0
resources_used.mem=80560kb
resources_used.vmem=291496kb
resources_used.walltime=08:18:43
Error_Path: titan-ext4:/lustre/atlas2/hep100/current/TitanU/Grid/M038/L2100/HACC000/run/run003/hacc_gpu_m038.err
Output_Path: titan-ext4:/lustre/atlas2/hep100/current/TitanU/Grid/M038/L2100/HACC000/run/run003/hacc_gpu_m038.log

Emulating the Galaxy Power Spectrum

- **First paper:** Keep cosmology fixed and only vary 5 HOD parameters
- **Emulators for:** Galaxy-galaxy auto, galaxy-dark matter cross power spectra and correlation function based on 100 HOD models
- **Accuracy:** 1-2% between z=0 and z=1 out to k=1/Mpc
- **Currently in preparation:** Currently extended emulators to take into account cosmology dependence

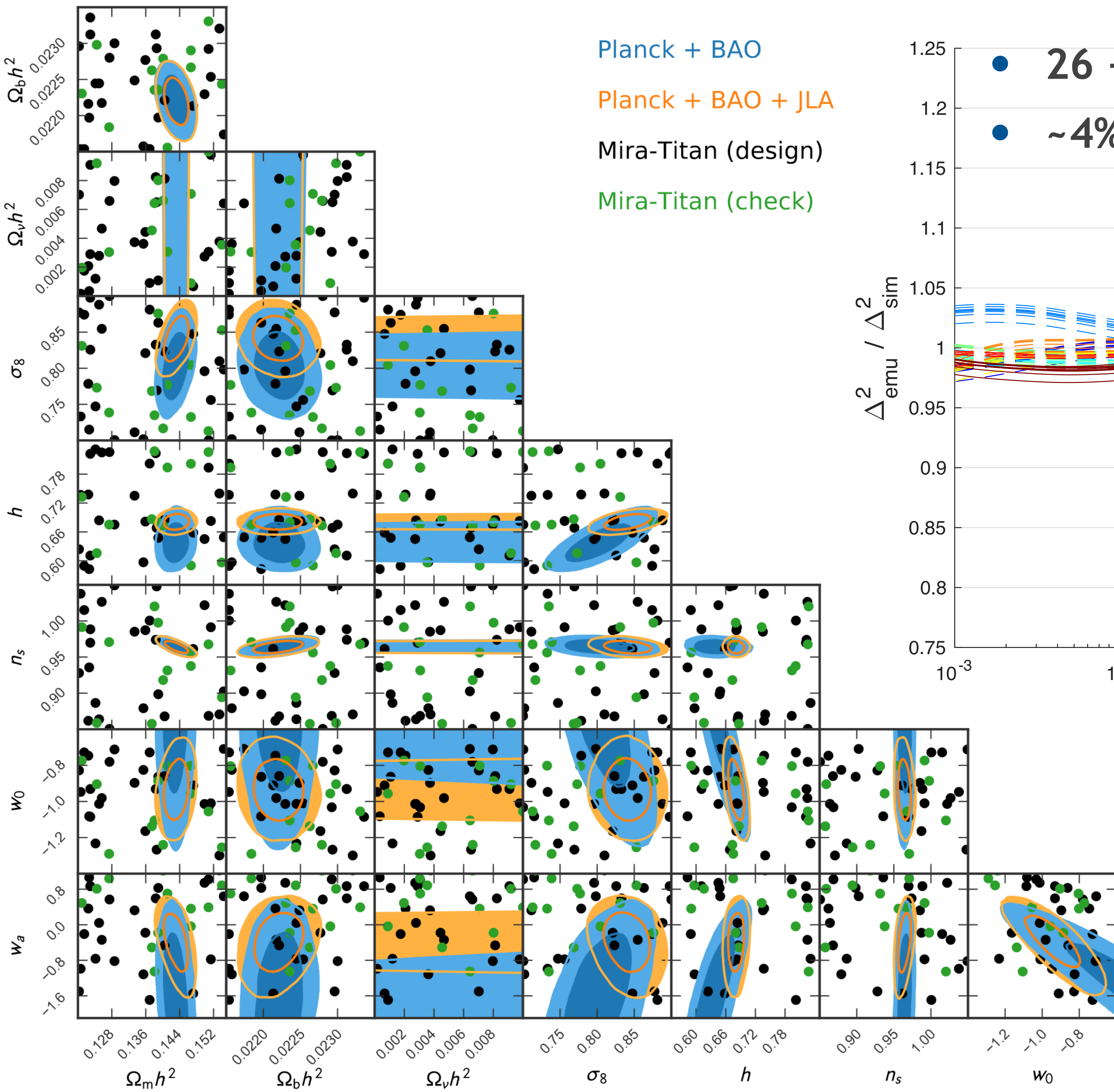


$$N_{\rm cen}(M) = \frac{1}{2}{\rm erfc}\left[\frac{\ln(M_{\rm cut}/M)}{\sqrt{2}\sigma}\right]$$

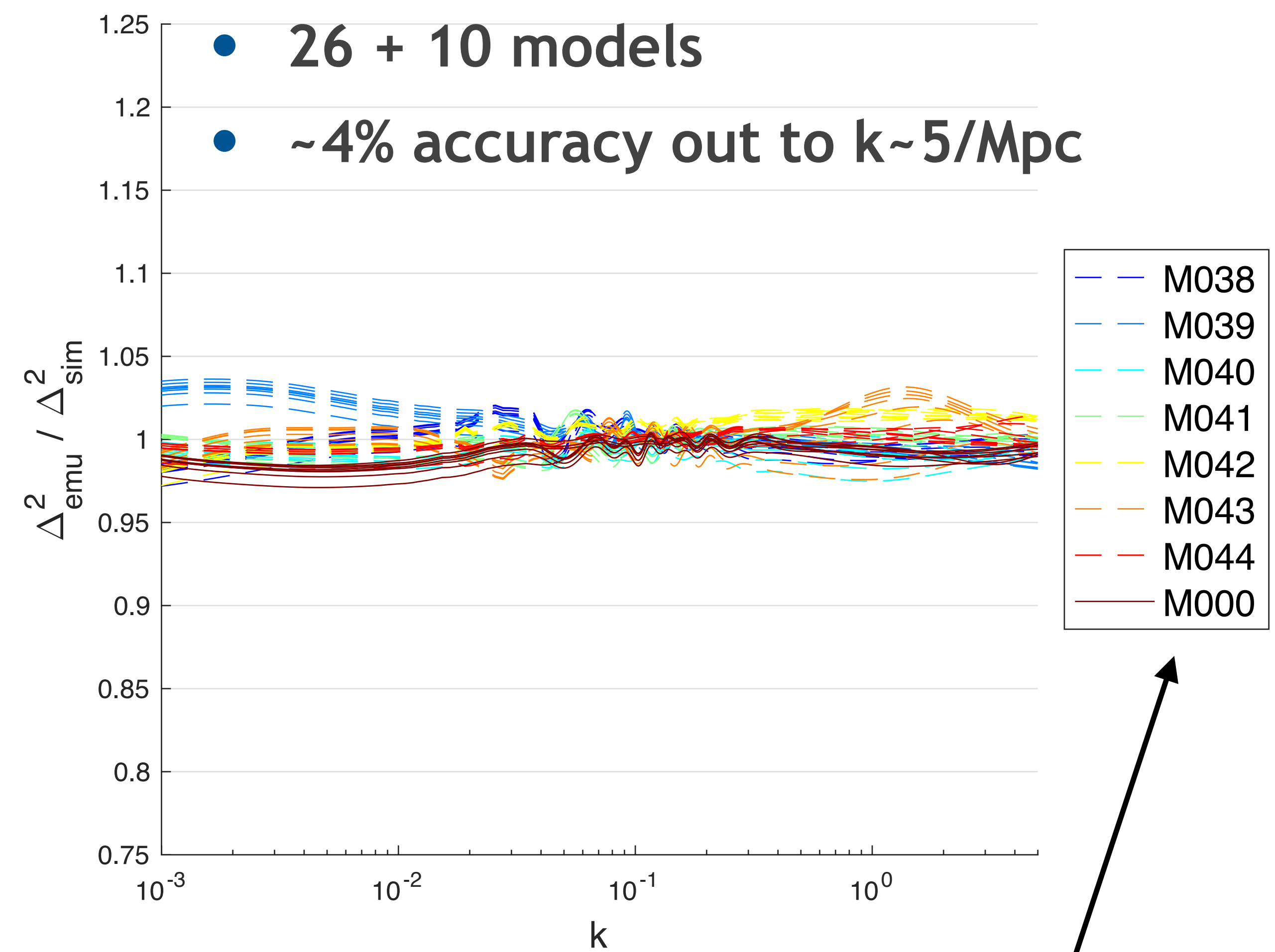
$$N_{\rm sat}(M) = \left(\frac{M - \kappa M_{\rm cut}}{M_1}\right)^{\alpha}$$

Kwan et al. 2015

The Mira-Titan Universe: Power Spectrum



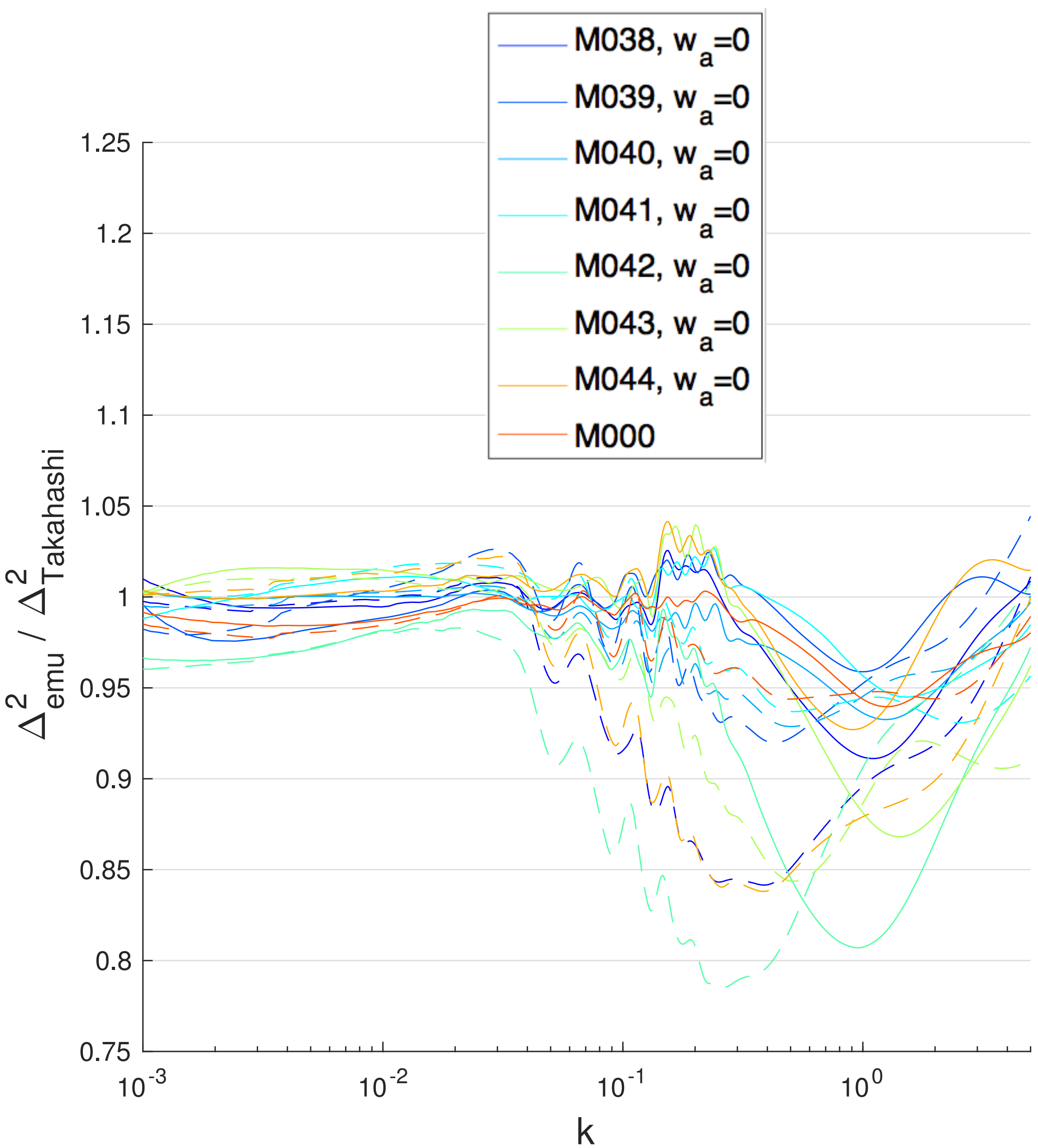
- 26 + 10 models
- ~4% accuracy out to k~5/Mpc



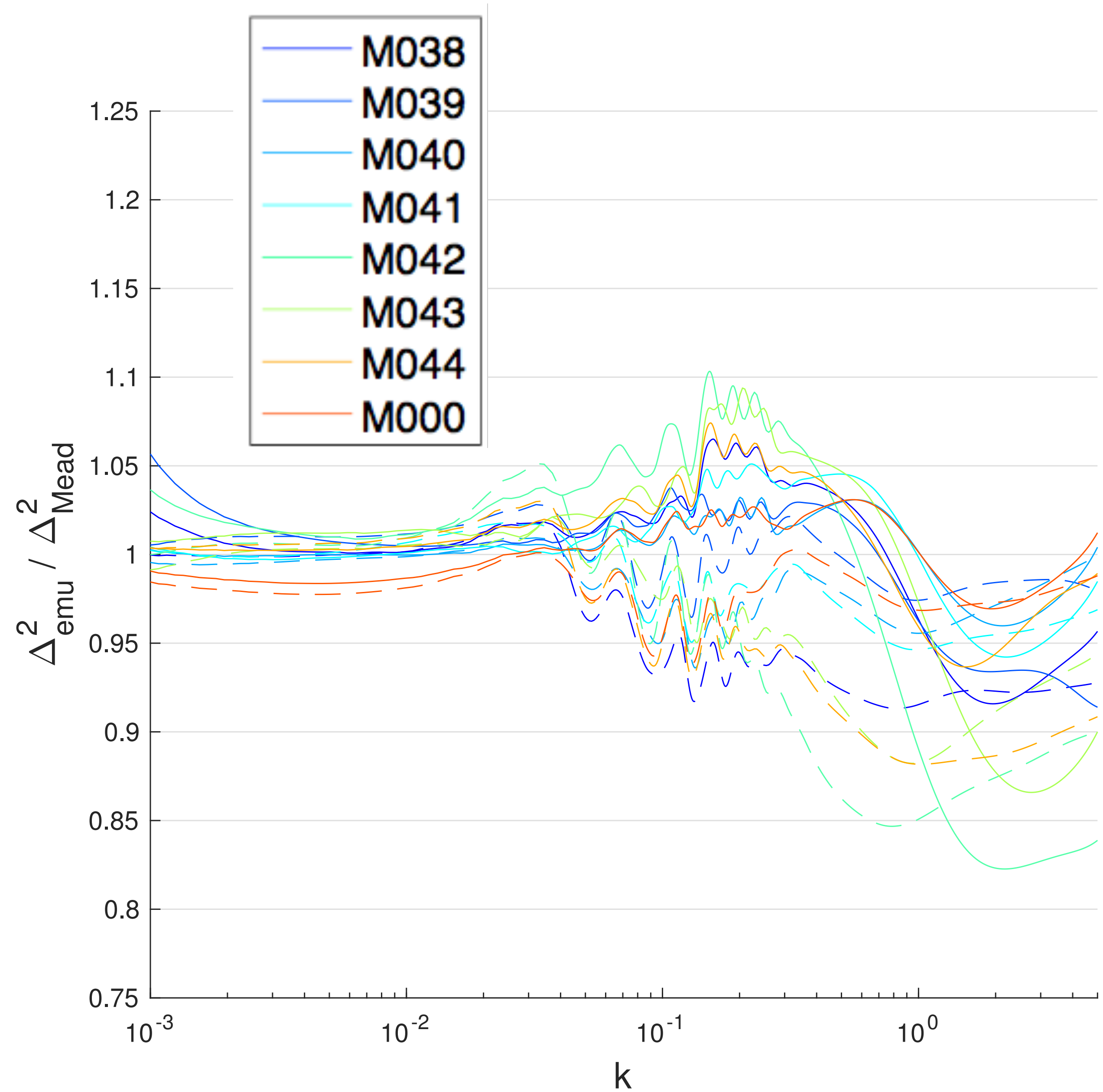
Models not used to build the emulator

Lawrence et al. 2017

Comparison with Other Methods



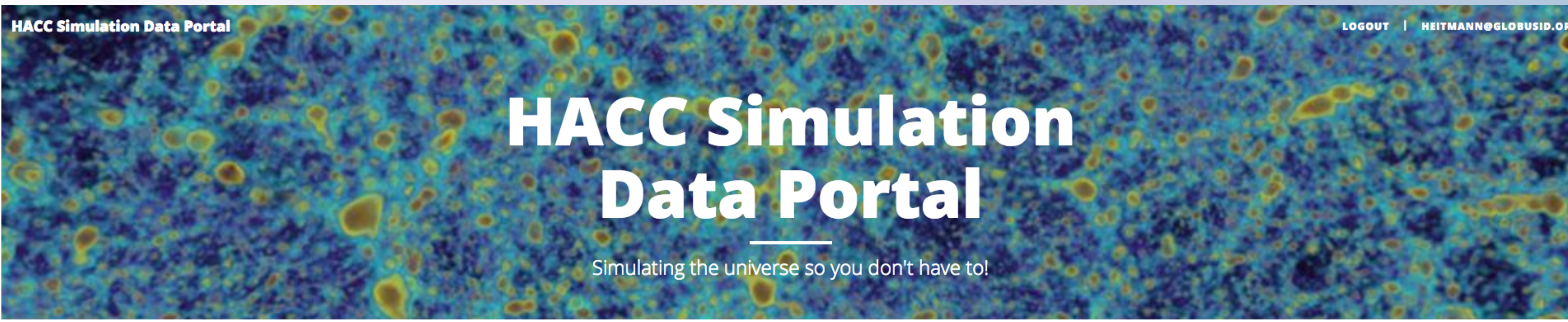
Takahashi et al. 2012



Mead et al. 2015

Lawrence et al. 2017

Mira-Titan Universe Release in Preparation



Mira/Titan Universe Simulation

text describing this simulation

OuterRim Simulation

text describing this simulation

Frequently Asked Questions

more text

LEARN MORE →

in collaboration with Tom Uram

Mira-Titan Universe Release in Preparation

HACC Simulation Data Portal

Repository

Select dataset(s) to transfer

Search:

	Name	omega_cdm	deut	omega_nu	hubble	ss8	ns	w_de	wa_de
▢	M000	0.22	0.02258	0.0	0.71	0.8	0.963	-1.0	0.0
▢	M001	0.3276	0.02261	0.0	0.6167	0.8778	0.9611	-0.7	0.6722
▢	M002	0.1997	0.02328	0.0	0.75	0.8556	1.05	-1.033	0.9111
▢	M003	0.259	0.02194	0.0	0.7167	0.9	0.8944	-1.1	-0.2833
▢	M004	0.2971	0.02283	0.0	0.5833	0.7889	0.8722	-1.167	1.15
▢	M005	0.1658	0.0235	0.0	0.85	0.7667	0.9833	-1.233	-0.04445
▢	M006	0.3643	0.0215	0.0	0.55	0.8333	0.9167	-0.7667	0.1944
▢	M007	0.19329867	0.02217	0.0	0.8167	0.8111	1.028	-0.8333	-1.0
▢	M008	0.207625252	0.02306	0.0	0.6833	0.7	1.006	-0.9	0.4333
▢	M009	0.278532533	0.02172	0.0	0.65	0.7444	0.85	-0.9667	-0.7611
▢	M010	0.17180095	0.02239	0.0	0.7833	0.7222	0.9389	-1.3	-0.5222

Showing 1 to 11 of 11 entries

- ▢ Include Halo particles
- ▢ Include BIG Halo particles
- ▢ Include Simulation particles
- ▢ Include Halo properties

Transfer

Mira-Titan Universe Release in Preparation

HACC Simulation Data Portal

LOGOUT | HEITMANN@GLOBUSID.ORG

Repository

Select dataset(s) to transfer

Search:

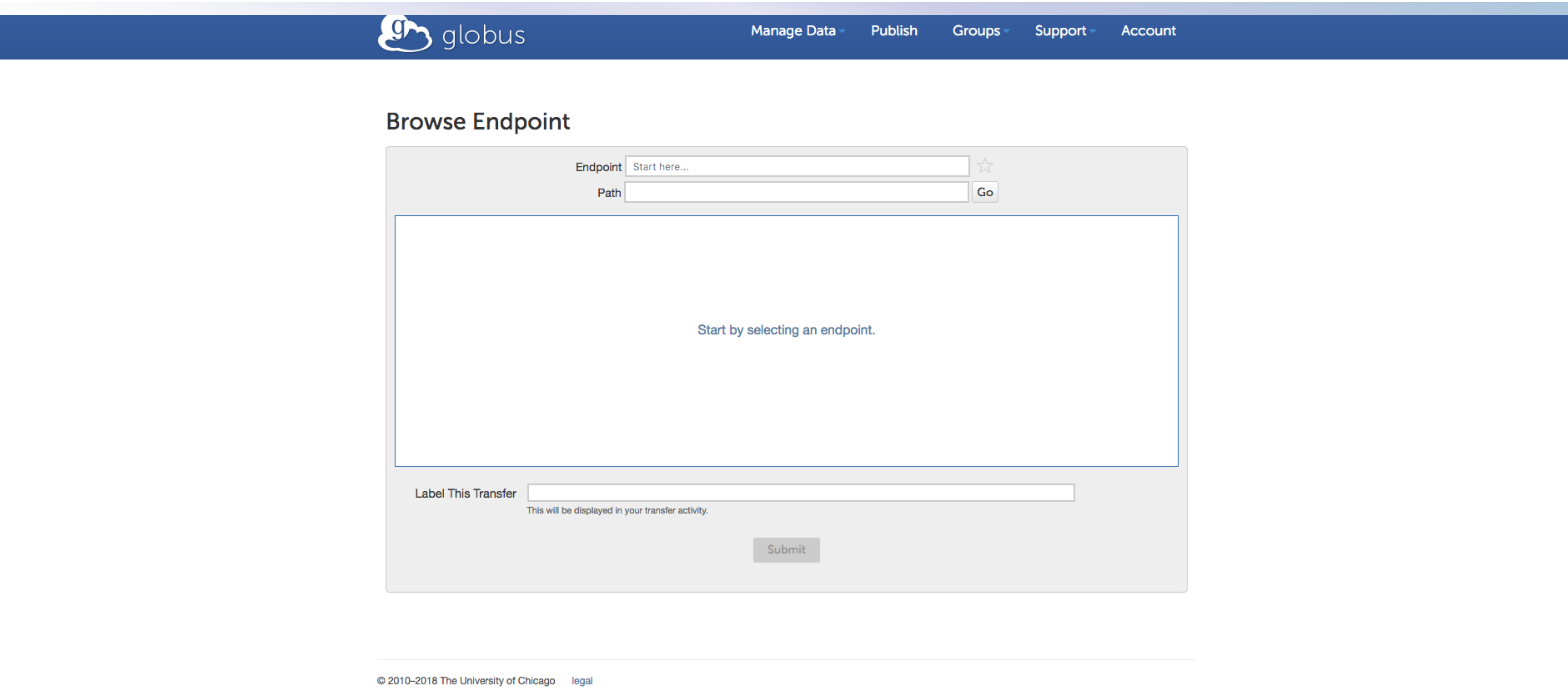
	Name	omega_cdm	deut	omega_nu	hubble	ss8	ns	w_de	wa_de
☑	M000	0.22	0.02258	0.0	0.71	0.8	0.963	-1.0	0.0
☐	M001	0.3276	0.02261	0.0	0.6167	0.8778	0.9611	-0.7	0.6722
☐	M002	0.1997	0.02328	0.0	0.75	0.8556	1.05	-1.033	0.9111
☐	M003	0.259	0.02194	0.0	0.7167	0.9	0.8944	-1.1	-0.2833
☐	M004	0.2971	0.02283	0.0	0.5833	0.7889	0.8722	-1.167	1.15
☑	M005	0.1658	0.0235	0.0	0.85	0.7667	0.9833	-1.233	-0.04445
☐	M006	0.3643	0.0215	0.0	0.55	0.8333	0.9167	-0.7667	0.1944
☐	M007	0.19329867	0.02217	0.0	0.8167	0.8111	1.028	-0.8333	-1.0
☐	M008	0.207625252	0.02306	0.0	0.6833	0.7	1.006	-0.9	0.4333
☐	M009	0.278532533	0.02172	0.0	0.65	0.7444	0.85	-0.9667	-0.7611
☐	M010	0.17180095	0.02239	0.0	0.7833	0.7222	0.9389	-1.3	-0.5222

Showing 1 to 11 of 11 entries 1 row selected

- ☐ Include Halo particles
- ☐ Include BIG Halo particles
- ☐ Include Simulation particles
- ☐ Include Halo properties

Transfer

Mira-Titan Universe Release in Preparation



Future Work and Open Questions

- **More Emulators:** Mass function, galaxy power spectrum and correlation function across cosmologies (real and redshift space), dark matter halo bias, …
- **Discrepancy Modeling:** What happens if our forward model isn't correct?
- **Nested/Adaptive Sampling:** Convergent/Learning approach to emulation
- **Covariance Emulation:** Emulate covariances rather than just the mean (observations are only for one realization!)
- **Accuracy Limits:** Theory for convergence (a posteriori so far)
- **Limits of Dimensionality:** How high can we go?
- **Cross-Correlations:** Optical X CMB, lensing X galaxy distribution, etc.
- **Galaxy Catalogs:** Emulation of statistics from galaxy formation models

Key References

- **General Method:**
  - K. Heitmann, D. Higdon, C. Nakhleh, and S. Habib, ApJ Lett **646**, 1 (2006) [short]
  - S. Habib, K. Heitmann, D. Higdon, C. Nakhleh, and B. Williams, Phys. Rev. D **76**, 083503 (2007) [technical]
  - D. Higdon, K. Heitmann, C. Nakhleh, and S. Habib, in the *Oxford Handbook of Applied Bayesian Analysis* edited by O' Hagan and West (Oxford, 2010) [review with a worked out inverse problem]
  - K. Heitmann, D. Bingham, E. Lawrence, S. Bergner, S. Habib, D. Higdon, A. Pope, R. Biswas, H. Finkel, N. Frontiere, and S. Bhattacharya, ApJ **820**, 108 (2016) [nested sampling, strong convergence]
- **Power Spectra:**
  - K. Heitmann, M. White, C. Wagner, S. Habib, and D. Higdon, ApJ **715**, 104 (2010) [Coyote Universe I]; K. Heitmann, D. Higdon, M. White, S. Habib, B.J. Williams, and C. Wagner, ApJ **705**, 156 (2009) [Coyote Universe II]; E. Lawrence, K. Heitmann, M. White, D. Higdon, C. Wagner, S. Habib, and B. Williams, ApJ **713**, 1322 (2010) [Coyote Universe III]; K. Heitmann, E. Lawrence, J. Kwan, S. Habib, and D. Higdon, ApJ **780**, 111 (2014) [Coyote Universe IV]
  - J. Kwan, K. Heitmann, S. Habib, N. Padmanabhan, E. Lawrence, H. Finkel, N. Frontiere, and A. Pope, Phys. Rev. D **810**, 35 (2015) [galaxy power spectrum]
  - E. Lawrence, K. Heitmann, J. Kwan, A. Upadhye, D. Bingham, S. Habib, D. Higdon, A. Pope, H. Finkel, and Nicholas Frontiere, ApJ **847**, 50 (2017) [Mira-Titan Universe II];
- **Other Examples:**
  - J. Kwan, S. Bhattacharya, K. Heitmann, and S. Habib, ApJ **768**, 123 (2013) [Halo shape emulation]
  - T. Holsclaw, U. Alam, B. Sanso, H. Lee, K. Heitmann, S. Habib, and D. Higdon, Phys. Rev. Lett. **105**, 241302 (2010) [Sn constraints on w(z) using GPs]
  - T. Holsclaw, U. Alam, B. Sanso, H. Lee, K. Heitmann, S. Habib, and D. Higdon, Phys. Rev. D **84**, 083501 (2011) [combining data sets]