Modeling HI at the field level

Andrej Obuljen (UZH)

with Marko Simonović (CERN) et al.

arXiv: 2207.12398

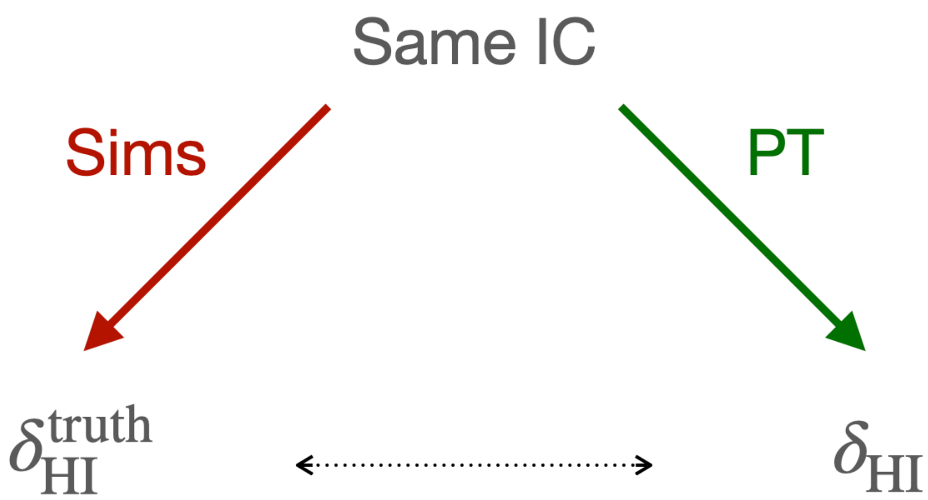
Motivation

- Test PT+bias models for HI at the field level (Schmittfull+18)
- 21cm IM surveys will be mainly sensitive to perturbative scales
- Explore HI noise properties
- Generate fast & accurate HI field level mocks

Field level approach

Advantages

- Pixel-by-pixel agreement → agreement of all summary statistics
- No overfitting
- Easy to isolate and study noise
- No cosmic variance for same IC, no need for large hydro sims.



PT model

Hybrid Lagrangian and Eulerian scheme, bulk flows included, only linear fields

$$\delta_{\rm HI}(\mathbf{k}) = \int d^3\mathbf{q}(1 + b_1^L\delta_1 + b_2^L(\delta_1^2 - \sigma_1^2) + b_{\mathcal{G}_2}^L\mathcal{G}_2 + \ldots - i\mathbf{k}\cdot\boldsymbol{\psi}_2 + \ldots)e^{-i\mathbf{k}(\mathbf{q}+\boldsymbol{\psi}_1)}$$

Define *shifted bias operators* in Eulerian space:

$$\tilde{\mathcal{O}}(\mathbf{k}) = \int d^3\mathbf{q}\,\mathcal{O}(\mathbf{q})e^{-i\mathbf{k}(\mathbf{q}+\boldsymbol{\psi}_1(\mathbf{q}))},\ \text{where } \mathcal{O} \in \{1, \delta_1, \delta_2, \mathcal{G}_2, \delta_3, \ldots\}$$

Stochastic noise

$$\delta_{\rm HI}(\mathbf{k}) = \beta_1(k)\tilde{\delta}_1(\mathbf{k}) + \beta_2(k)\tilde{\delta}_2^{\perp}(\mathbf{k}) + \beta_{\mathcal{G}_2}(k)\tilde{\mathcal{G}}_2^{\perp}(\mathbf{k}) + \ldots + \epsilon$$

Transfer functions

Matches 1-loop EFT & CLPT power spectrum

Measure of success

How to compare model & truth at the field level?

- Minimise mean-squared difference/residuals:

$$P_{\rm err}(k) \equiv \langle\, |\, \delta_{\rm HI}^{\rm truth}(\mathbf{k}) - \delta_{\rm HI}^{\rm model}(\mathbf{k})\,|^2\,\rangle$$

by doing least squares in each k-bin to obtain best-fit transfer functions:

$$\beta_i = \langle \mathcal{O}_i^{\perp} \delta_{\rm HI}^{\rm truth *} \rangle / \langle\, |\, \mathcal{O}_i^{\perp}\,|^2\,\rangle$$

- Example:

$$\delta_{\rm HI} = b_1 \delta + \epsilon$$

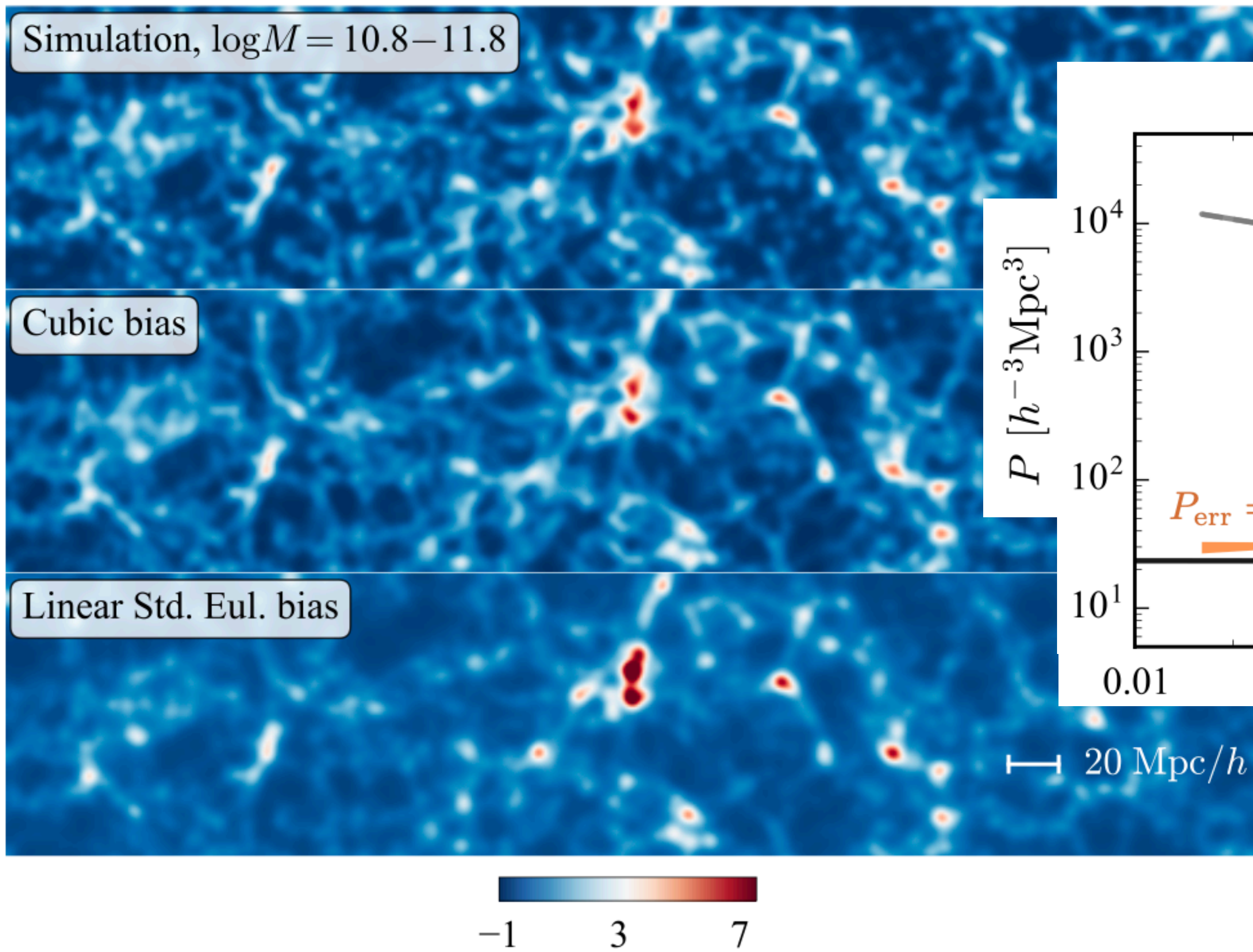
$$b_1(k) = \langle \delta_{\rm HI}^{\rm truth} \delta^* \rangle / \langle\, |\, \delta\,|^2\,\rangle$$

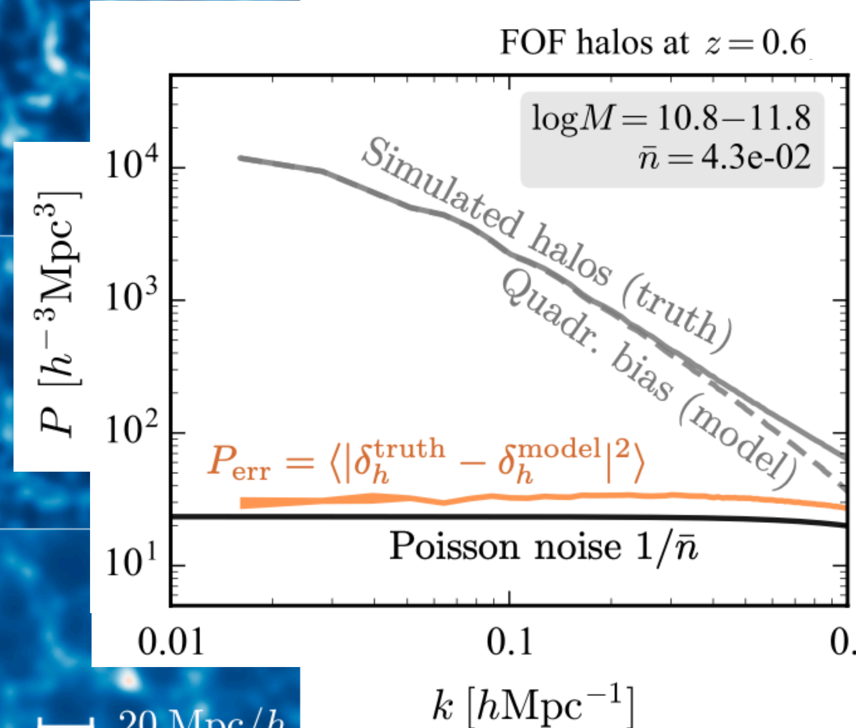
$$P_{\rm err}(k) = \langle\, |\, \epsilon\,|^2\,\rangle$$

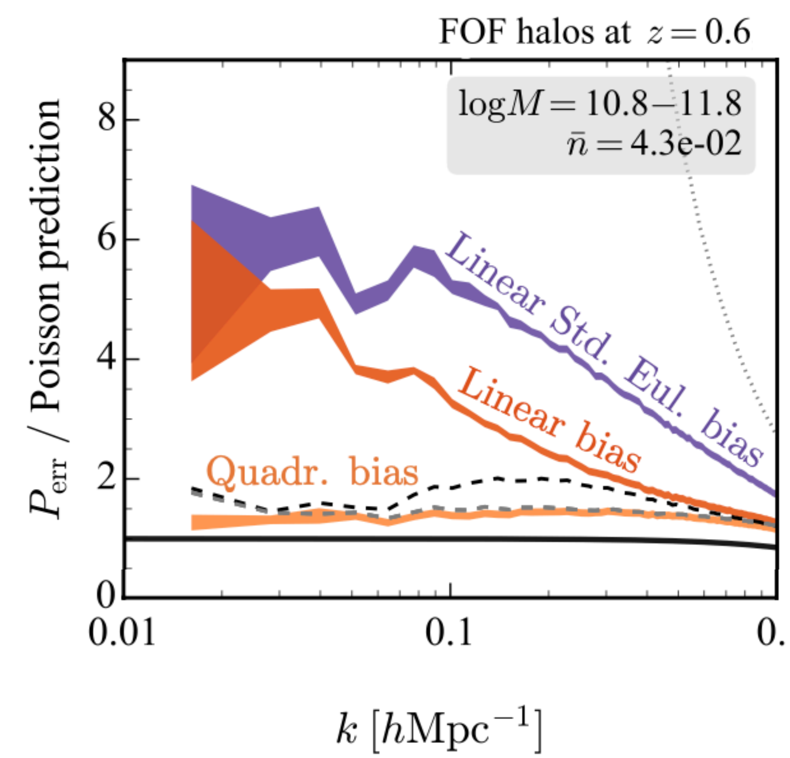
Halos

Schmittfull+18 (arXiv:1811.10640)

$L = 500\,h^{-1}\mathrm{Mpc}$



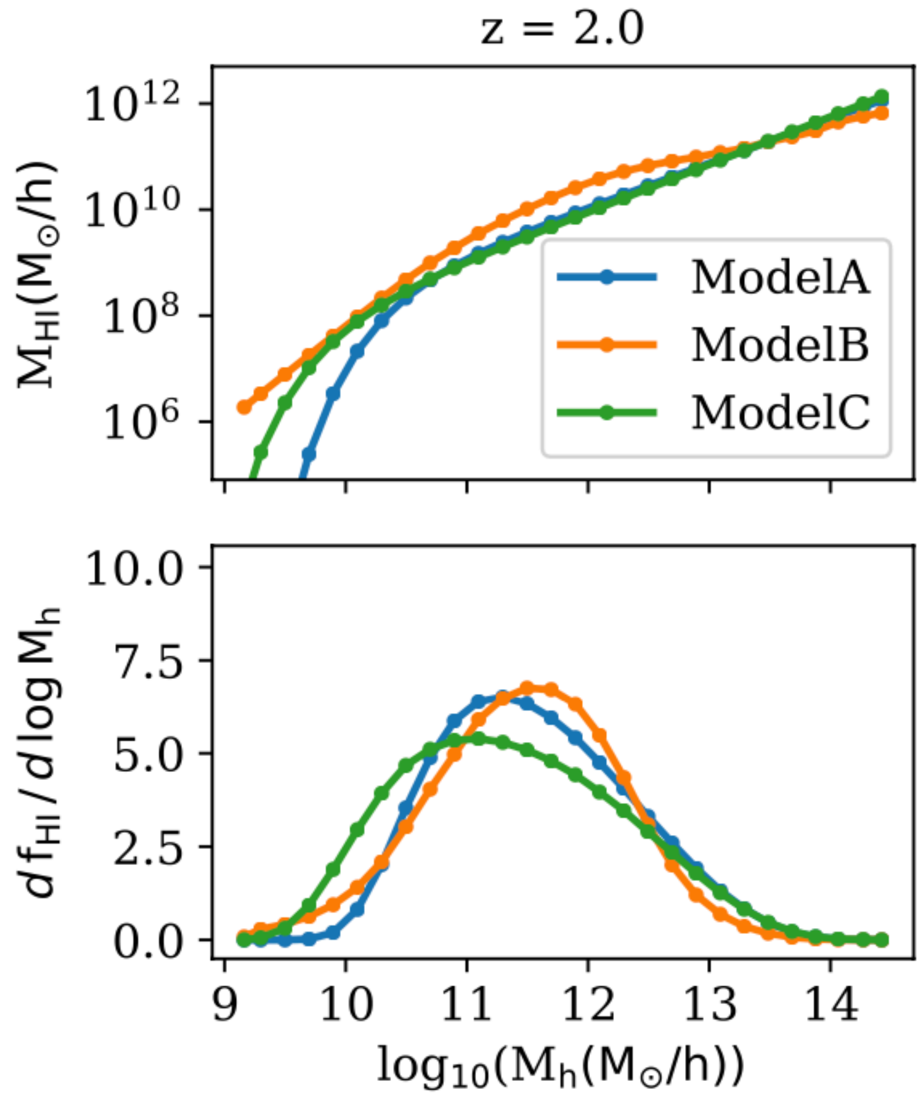




$$\delta_h(\boldsymbol{k}) = \beta_1(k)\ \tilde{\delta}_1(\boldsymbol{k}) + \beta_2(k)\ \tilde{\delta}_2^{\perp}(\boldsymbol{k}) + \beta_{\mathcal{G}_2}(k)\ \tilde{\mathcal{G}}_2^{\perp}(\boldsymbol{k}) + \cdots$$

Why is HI different/interesting?

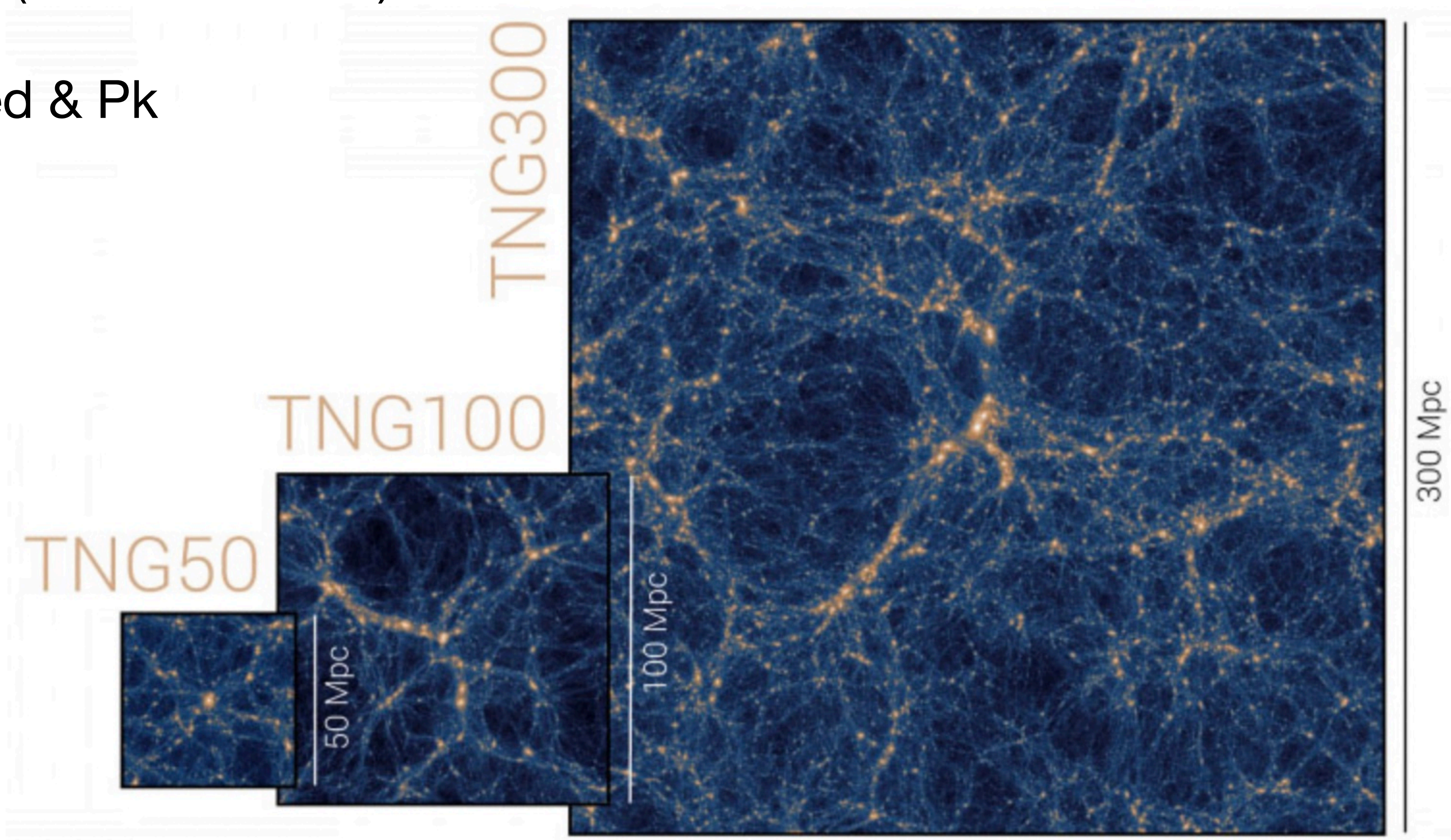
- HI signal from moderate halo masses
- Low shot-noise
- Moderate redshift space effects
- Field level ideal for 21cm maps (no need to discretise)



Modi+19

HI from IllustrisTNG

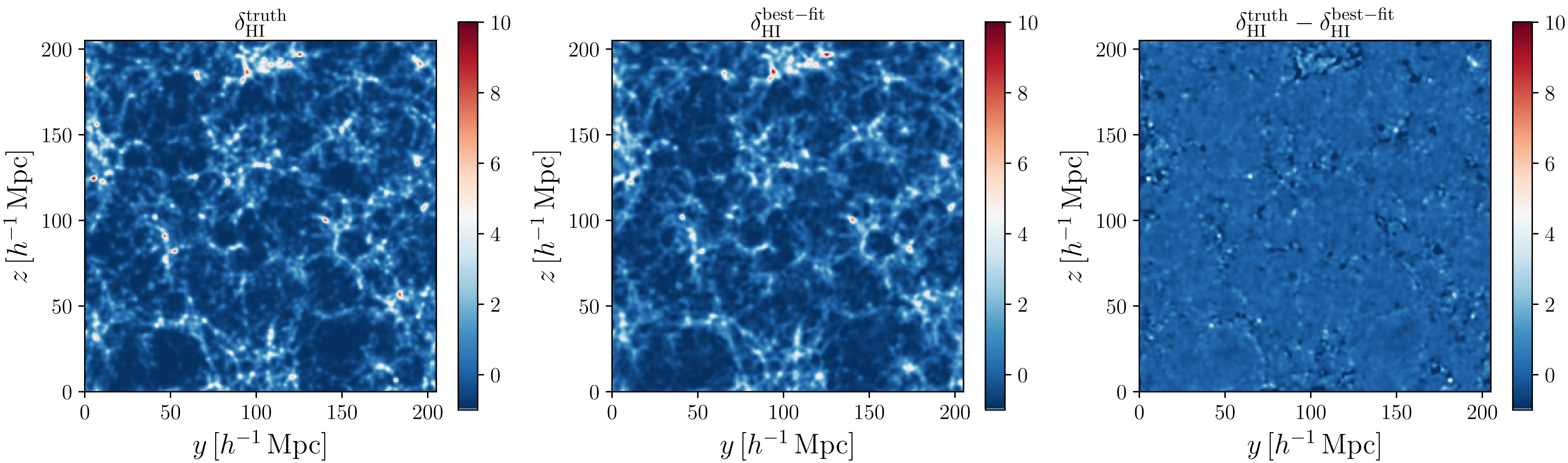
- Application to full hydro simulation
- TNG300-1 ($L = 205\,h^{-1}\mathrm{Mpc}$)
- HI in post-processing (Villaescusa+18)
- Same IC: random seed & Pk



Results

$z = 1$, real space

$$\delta_{\mathrm{HI}}(\mathbf{k}) = \beta_1(k)\tilde{\delta}_1(\mathbf{k}) + \beta_2(k)\tilde{\delta}_2^{\perp}(\mathbf{k}) + \beta_{\mathscr{G}_2}(k)\tilde{\mathscr{G}}_2^{\perp}(\mathbf{k}) + \beta_3(k)\tilde{\delta}_3^{\perp}(\mathbf{k}) + \ldots + \epsilon$$

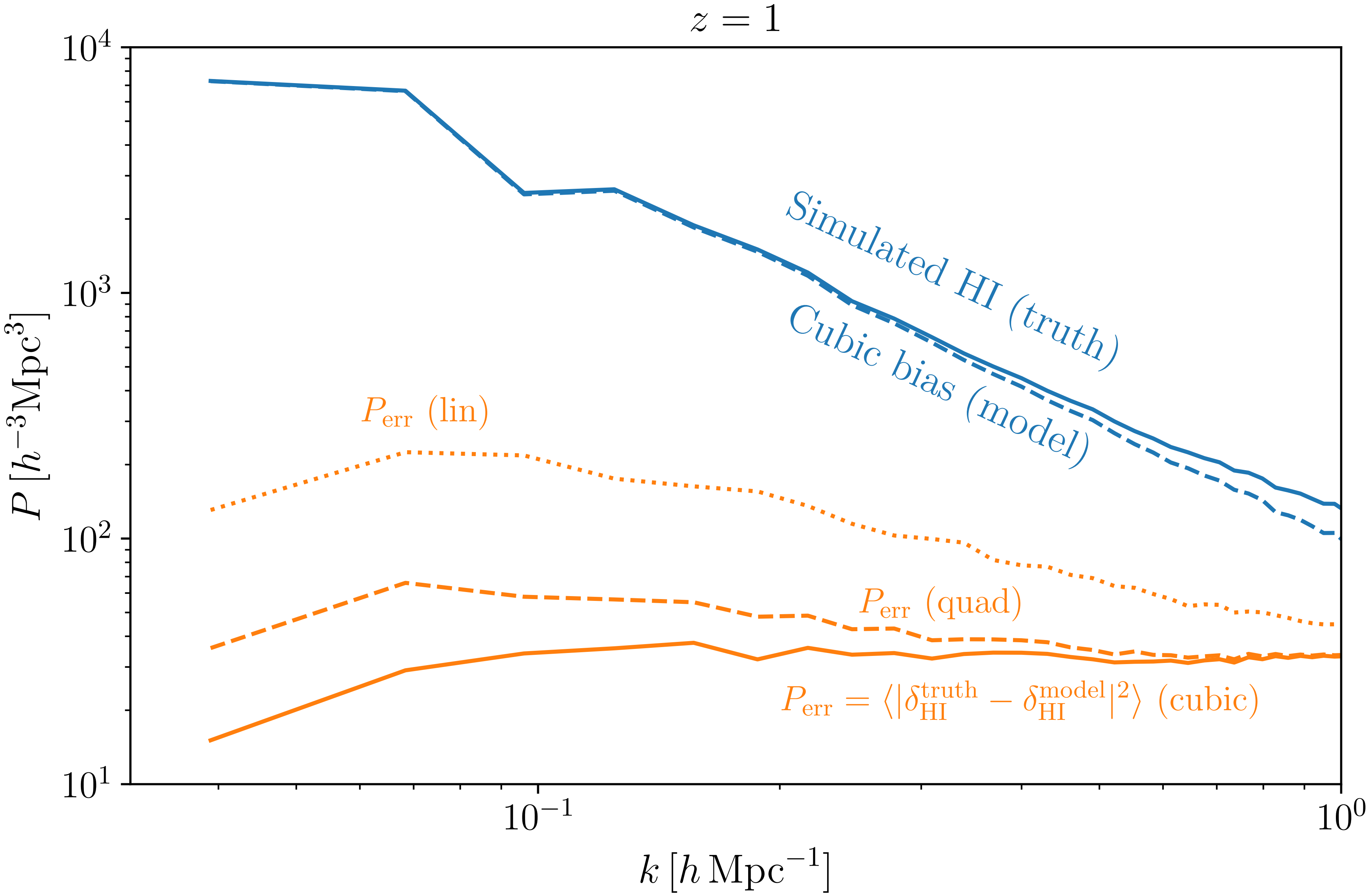


Slice depth 20 Mpc/h, smoothed 1 Mpc/h Gaussian

Results – power spectrum

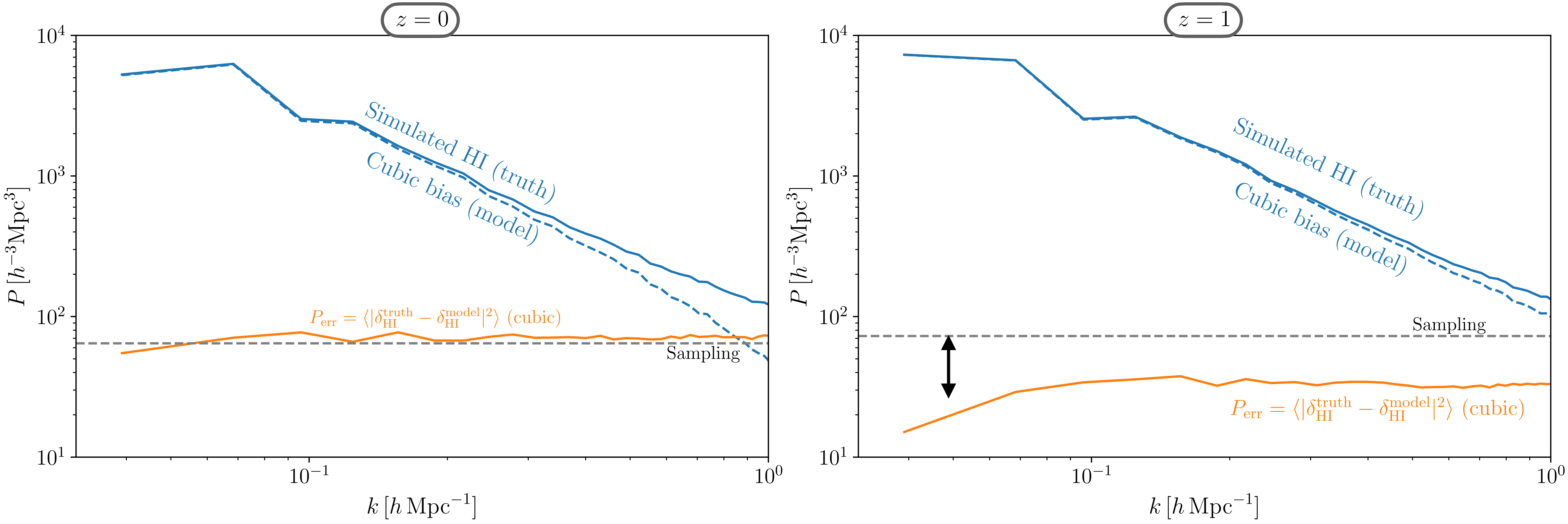
$z = 1$, real space

$$\delta_{\rm HI}(\mathbf{k}) = \beta_1(k)\tilde{\delta}_1(\mathbf{k}) + \beta_2(k)\tilde{\delta}_2^{\perp}(\mathbf{k}) + \beta_{\mathcal{G}_2}(k)\tilde{\mathcal{G}}_2^{\perp}(\mathbf{k}) + \beta_3(k)\tilde{\delta}_3^{\perp}(\mathbf{k}) + \ldots + \epsilon$$



HI noise properties

- $P_{\rm err}$ flat
- $P_{\rm err}$ comparable (but not equal!) to sampling noise (~1/nbar)

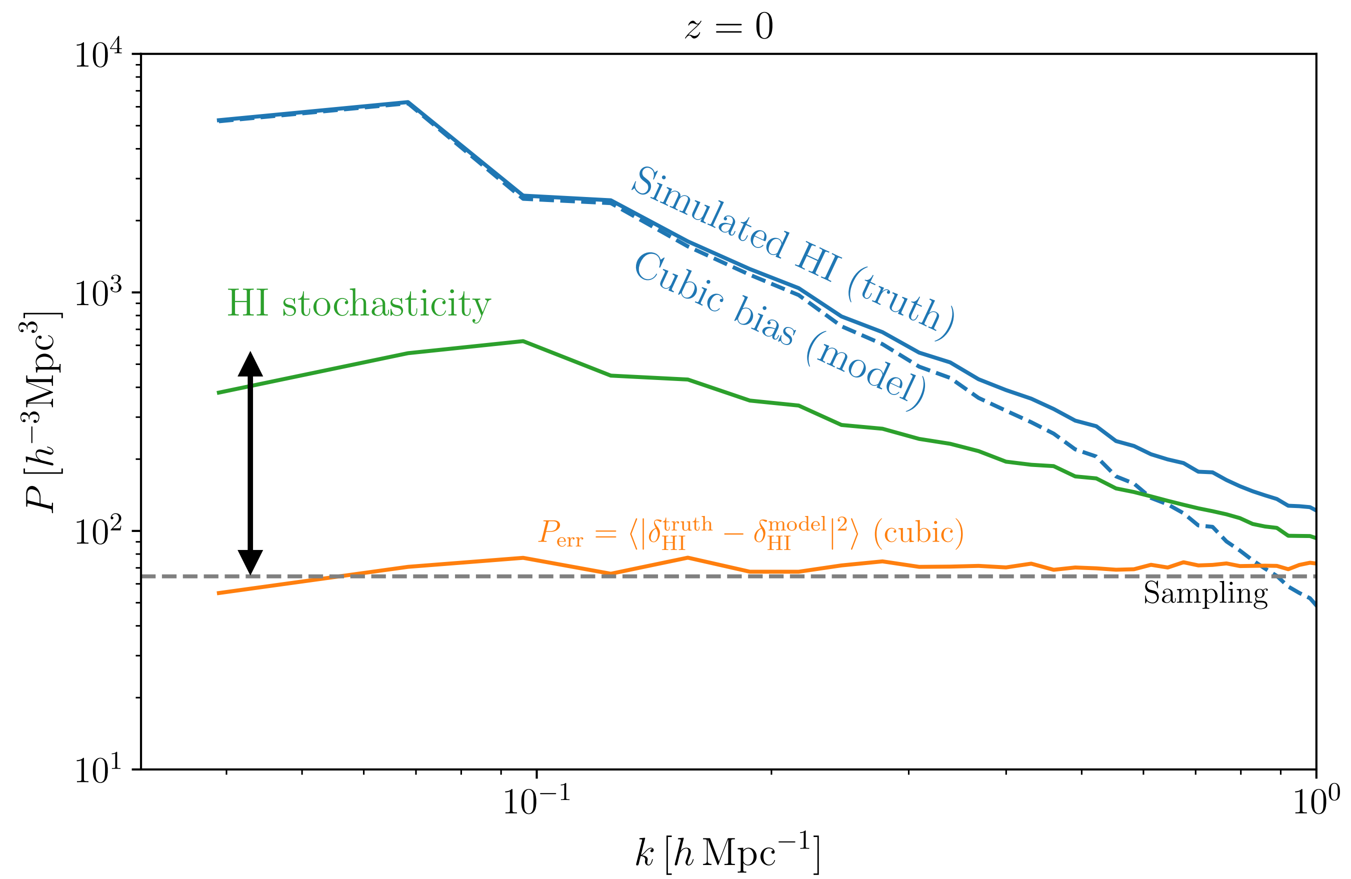


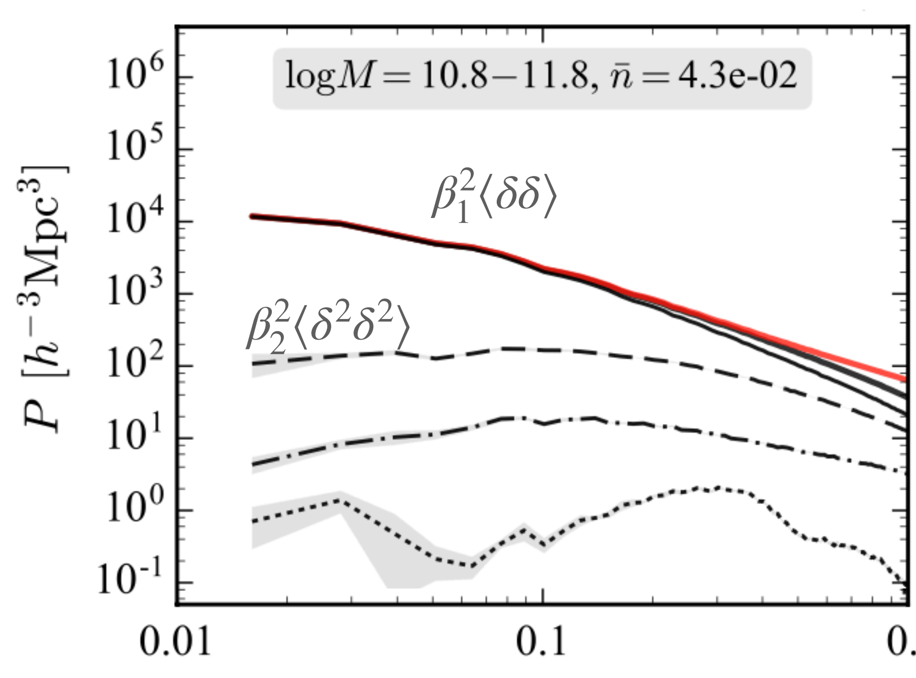
HI noise properties

- HI stochasticity: $\langle | \delta_{\rm HI}^{\rm truth} - b_1 \delta_{\rm m} |^2 \rangle$
- $P_{\rm err}$ lower than stochasticity

$$\delta_{\rm HI}(\mathbf{k}) - \beta_1 \tilde{\delta}_1(\mathbf{k}) = \beta_2 \tilde{\delta}_2^{\perp}(\mathbf{k}) + \beta_{\mathcal{G}_2} \tilde{\mathcal{G}}_2^{\perp}(\mathbf{k}) + \ldots + \epsilon$$

- In contrast to galaxy surveys, higher order terms dominate $P_{\rm err}$

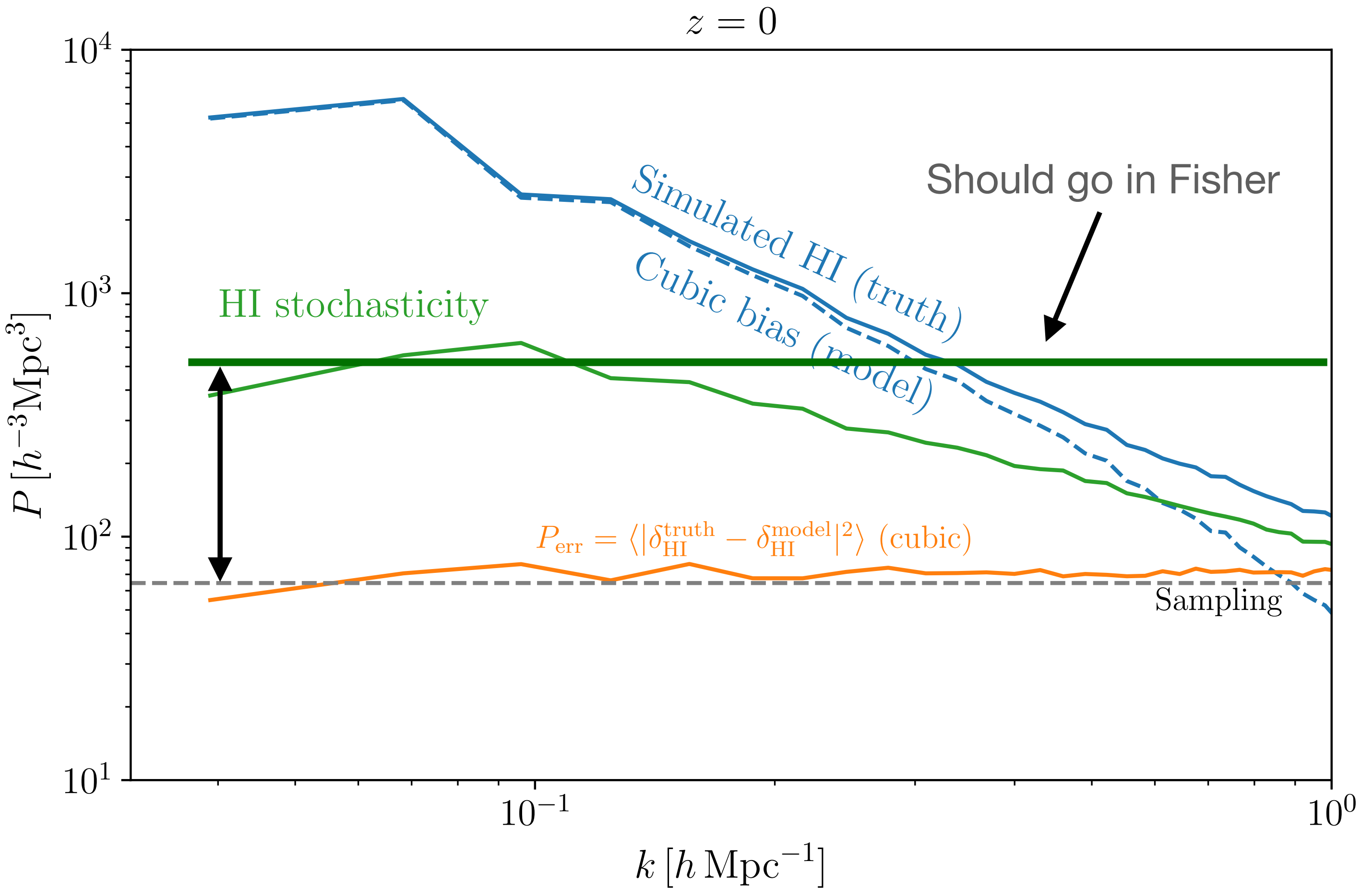




Schmittfull+18

HI noise properties

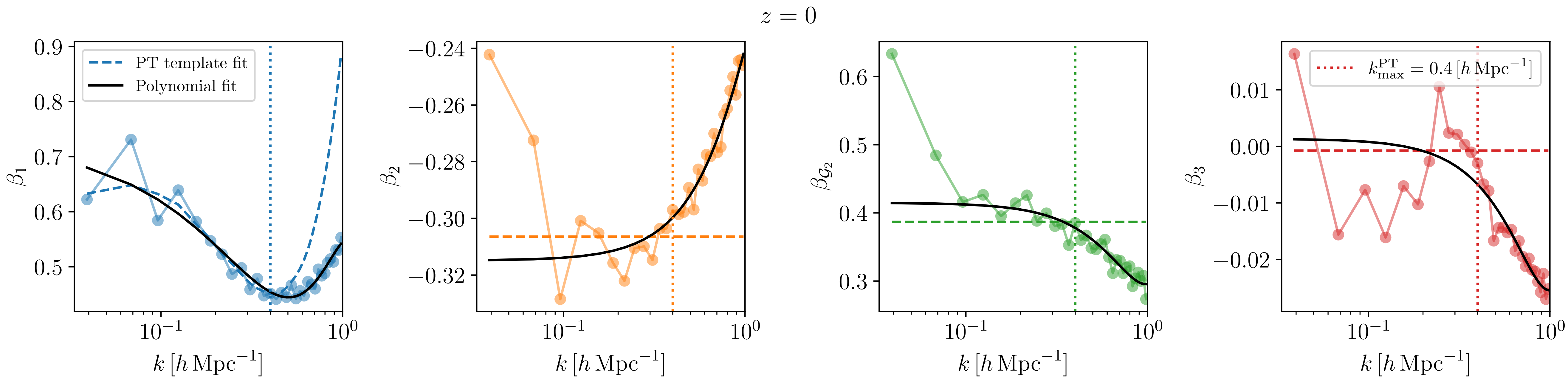
- For Pk, higher order terms flat & degenerate with noise
- Most Fisher forecasts assume sampling noise (optimistic)
- Field level may do better!



Transfer functions

$z = 0$, real space

$$\beta_1^{\rm real}(k) = a_0 + a_1 k + a_2 k^2 + a_4 k^4,$$
$$\beta_{i\neq 1}^{\rm real}(k) = a_0 + a_2 k^2 + a_4 k^4$$

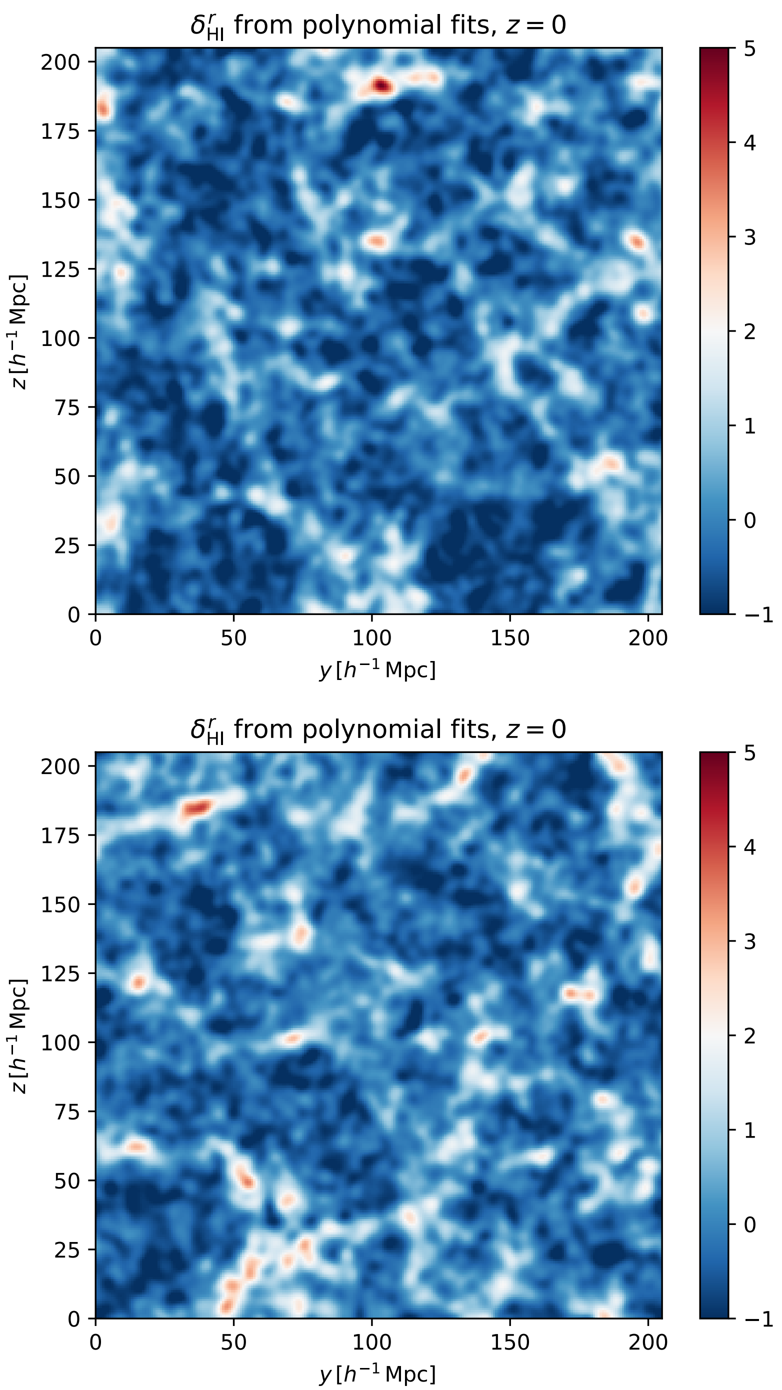


$$\delta_{\rm HI}(\mathbf{k}) = \beta_1(k)\tilde{\delta}_1(\mathbf{k}) + \beta_2(k)\tilde{\delta}_2^{\perp}(\mathbf{k}) + \beta_{\mathcal{G}_2}(k)\tilde{\mathcal{G}}_2^{\perp}(\mathbf{k}) + \beta_3(k)\tilde{\delta}_3^{\perp}(\mathbf{k}) + \ldots + \epsilon$$

Hi-Fi mocks

https://github.com/andrejobuljen/Hi-Fi_mocks

- Publicly available code
- Generate fast 3D HI field (Hi-Fi) level mocks
- Tuned to TNG HI clustering (k=0.03-1)
- Real & redshift space at z=0,1 (*more soon*)
- Few min. on a laptop (*improvements soon!*)
- Variables: BoxSize, IC seed, z…
- Please give it a try and get in touch!
- Hopefully useful for SKA WG



TNG and random IC seed at z=0

Conclusions

- HI is a biased tracer of matter field
- Cubic bias model 1% up to $k = 0.4\ (0.3)\ h/\mathrm{Mpc}$ in real (redshift) space

- HI noise flat & lower than stochasticity
- Higher order terms dominant at low-k, opposite to galaxies
- Case where field level analysis could be worthwhile (future work)

- Velocity dispersion estimates lower than for galaxies

- Public code to generate HI field level mocks: Hi-Fi mocks
- Improve future data analysis
- Useful for forecasts, mocks, covariances…

Thank you!