Sep 29, 2025 - Oct 03, 2025
New Physics from Galaxy Clustering
GGI

Ivana Nikolac
Fabian Schmidt
Beatriz Tucci

MAX PLANCK INSTITUTE
FOR ASTROPHYSICS

Connecting Field-level and Summary Statistics from Galaxy Clustering

Seeing the Universe Voxel by Voxel:
The Field-Level Approach

observations ⟶ 3D map of galaxy distribution ⟶ cosmological inference ⟶ cosmological and bias parameters

“”STANDARD” COSMOLOGICAL INFERENCE

Seeing the Universe Voxel by Voxel:
The Field-Level Approach

observations ⟶ 3D map of galaxy distribution ⟶ cosmological inference ⟶ cosmological and bias parameters

“”STANDARD” COSMOLOGICAL INFERENCE ⟶ SUMMARY STATISTICS

2- and 3- point correlation functions

Seeing the Universe Voxel by Voxel: The Field-Level Approach

"”STANDARD”
COSMOLOGICAL
INFERENCE

⟶

SUMMARY
STATISTICS

2- and 3- point
correlation functions

⟶

FIELD
LEVEL

$$\delta_g(x) = \frac{n_g(x)}{\bar{n}_g} - 1$$

Seeing the Universe Voxel by Voxel: The Field-Level Approach

How much reliable cosmological information can be extracted?

"”STANDARD” COSMOLOGICAL INFERENCE → SUMMARY STATISTICS → FIELD LEVEL

SUMMARY STATISTICS: 2- and 3- point correlation functions

FIELD LEVEL:

$$\delta_g(x) = \frac{n_g(x)}{\bar{n}_g} - 1$$

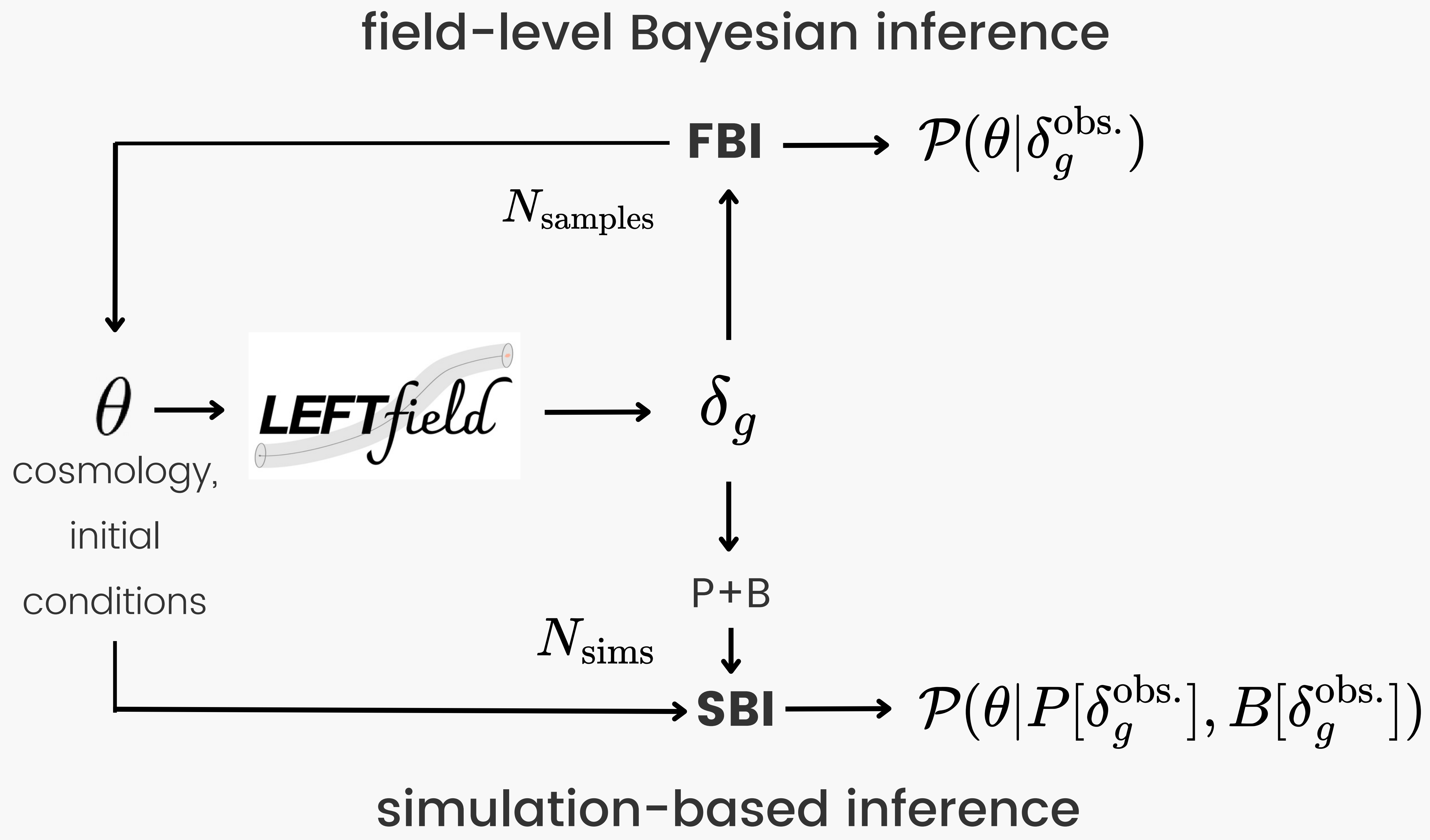
HOW MUCH INFORMATION CAN BE EXTRACTED FROM GALAXY CLUSTERING AT THE FIELD LEVEL?

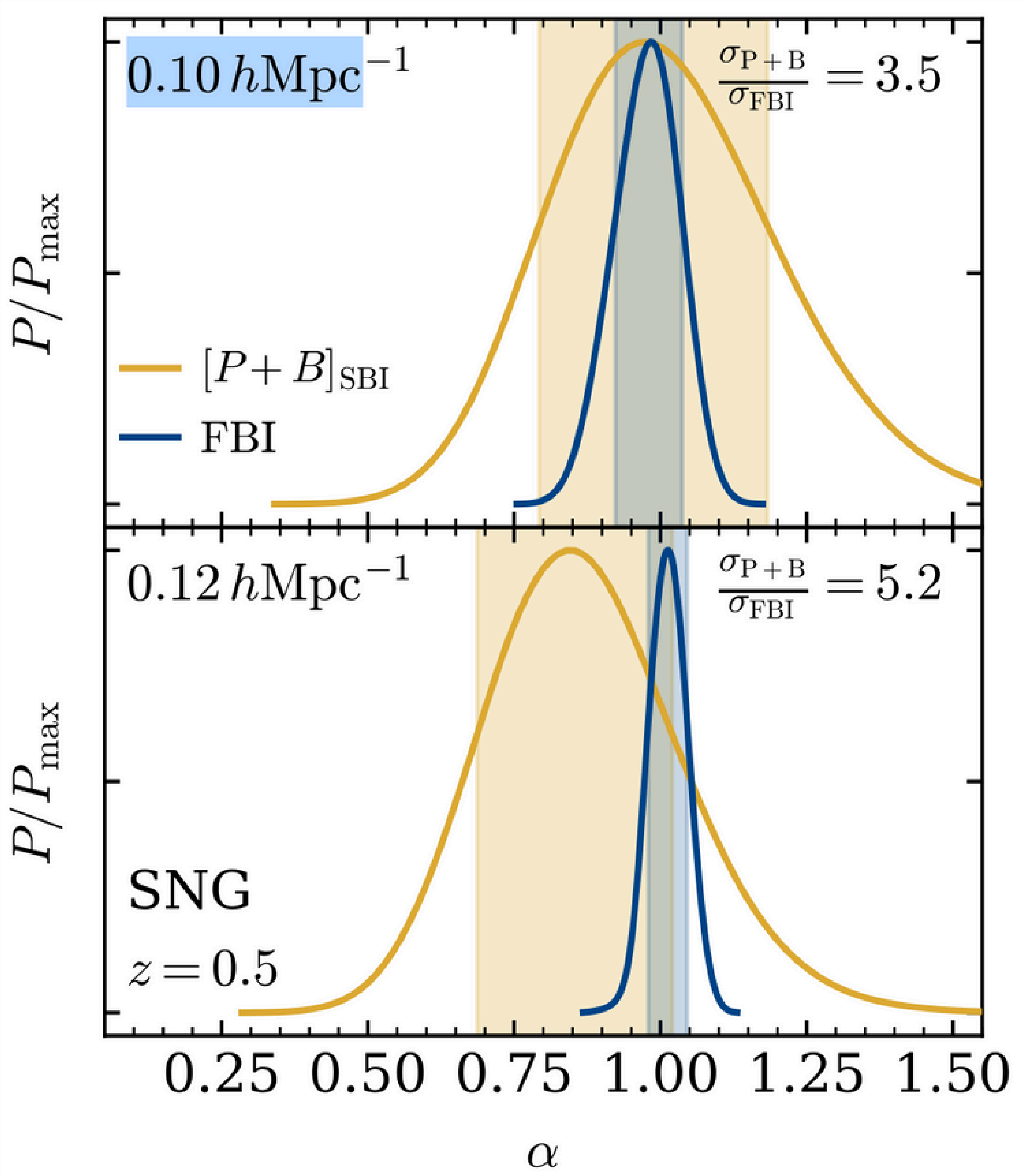
Nguyen, Schmidt, Tucci et al. (2024)


Nhat-Minh Nguyen
IPMU, University of Tokyo


Fabian Schmidt
MPA


Beatriz Tucci
MPA-Stanford





A TOY SCENARIO

as discussed in Cabass et al. (2023) and Schmidt (2025)

$$\delta_{g,\mathrm{det}}(\boldsymbol{x}) = b_\delta \delta^{(1)}(\boldsymbol{x}) + b_{\delta^2}[\delta^{(1)}(\boldsymbol{x}) - \langle \delta^{(1)} \rangle]^2 + \varepsilon(\boldsymbol{x})$$

A TOY SCENARIO

$$\delta_{g,\mathrm{det}}(\boldsymbol{x}) = b_\delta \delta^{(1)}(\boldsymbol{x}) + b_{\delta^2}[\delta^{(1)}(\boldsymbol{x}) - \langle \delta^{(1)} \rangle]^2 + \varepsilon(\boldsymbol{x})$$

A TOY SCENARIO

$$\delta_{g,\mathrm{det}}(\boldsymbol{x}) = b_\delta \delta^{(1)}(\boldsymbol{x}) + b_{\delta^2}[\delta^{(1)}(\boldsymbol{x}) - \langle \delta^{(1)} \rangle]^2 + \varepsilon(\boldsymbol{x})$$

assumptions: $k_{\mathrm{max}} = \Lambda$

$P_\varepsilon \to 0$

A TOY SCENARIO

$$\delta_{g,\mathrm{det}}(\boldsymbol{x}) = b_\delta \delta^{(1)}(\boldsymbol{x}) + b_{\delta^2}[\delta^{(1)}(\boldsymbol{x}) - \langle \delta^{(1)} \rangle]^2 + \varepsilon(\boldsymbol{x})$$

assumptions: $k_{\mathrm{max}} = \Lambda$

$P_\varepsilon \to 0$

MAP expression for b_{δ^2} : (at field level)

$$b_{\delta^2} = \frac{N[\delta_g]}{D[\delta_g]} \quad \text{with}$$

$$N[\delta_g] = \frac{1}{b_\delta^2} \int_{\boldsymbol{k}}^{\Lambda} \frac{1}{P_{\mathrm{L}}(k)} (\delta_g^2)(-\boldsymbol{k}) \delta_g(\boldsymbol{k}) - 2 \int_{\boldsymbol{x}} \delta_g(\boldsymbol{x})$$

$$D[\delta_g] = \frac{1}{b_\delta^4} \int_{\boldsymbol{k}}^{\Lambda} \frac{1}{P_{\mathrm{L}}(k)} (\delta_g^2)(-\boldsymbol{k}) (\delta_g^2)(\boldsymbol{k}) - \frac{8}{b_\delta^2} \int_{\boldsymbol{x}} \delta_g^2(\boldsymbol{x})$$

A TOY SCENARIO

$$\delta_{g,\mathrm{det}}(\boldsymbol{x}) = b_\delta \delta^{(1)}(\boldsymbol{x}) + b_{\delta^2}[\delta^{(1)}(\boldsymbol{x}) - \langle \delta^{(1)} \rangle]^2 + \varepsilon(\boldsymbol{x})$$

assumptions: $k_{\mathrm{max}} = \Lambda$

$P_\varepsilon \to 0$

MAP expression for b_{δ^2} :

(at field level)

$$b_{\delta^2} = \frac{N[\delta_g]}{D[\delta_g]} \quad \text{with}$$

$$N[\delta_g] = \frac{1}{b_\delta^2} \int_{\boldsymbol{k}}^{\Lambda} \frac{1}{P_{\mathrm{L}}(k)} (\delta_g^2)(-\boldsymbol{k}) \delta_g(\boldsymbol{k}) - 2 \int_{\boldsymbol{x}} \delta_g(\boldsymbol{x})$$

$$D[\delta_g] = \frac{1}{b_\delta^4} \int_{\boldsymbol{k}}^{\Lambda} \frac{1}{P_{\mathrm{L}}(k)} (\delta_g^2)(-\boldsymbol{k}) (\delta_g^2)(\boldsymbol{k}) - \frac{8}{b_\delta^2} \int_{\boldsymbol{x}} \delta_g^2(\boldsymbol{x})$$

OPERATOR CORRELATORS (OC)

MAP expression for b_{δ^2} : $b_{\delta^2} = \frac{N[\delta_g]}{D[\delta_g]}$ with

$$N[\delta_g] = \frac{1}{b_\delta^2} \int_{\boldsymbol{k}}^{\Lambda} \frac{1}{P_{\rm L}(k)} (\delta_g^2)(-\boldsymbol{k})\delta_g(\boldsymbol{k}) - 2 \int_{\boldsymbol{x}} \delta_g(\boldsymbol{x})$$

$$D[\delta_g] = \frac{1}{b_\delta^4} \int_{\boldsymbol{k}}^{\Lambda} \frac{1}{P_{\rm L}(k)} (\delta_g^2)(-\boldsymbol{k})(\delta_g^2)(\boldsymbol{k}) - \frac{8}{b_\delta^2} \int_{\boldsymbol{x}} \delta_g^2(\boldsymbol{x})$$

new summaries: $\{\langle O^{(n)}[\delta_g] \,|\, O^{(m)}[\delta_g] \rangle\}$

OPERATOR CORRELATORS (OC)

MAP expression for b_{δ^2} : $b_{\delta^2} = \dfrac{N[\delta_g]}{D[\delta_g]}$ with

$$N[\delta_g] = \frac{1}{b_\delta^2}\int_{\boldsymbol{k}}^{\Lambda} \frac{1}{P_{\rm L}(k)} (\delta_g^2)(-\boldsymbol{k})\delta_g(\boldsymbol{k}) - 2\int_{\boldsymbol{x}} \delta_g(\boldsymbol{x})$$

$$D[\delta_g] = \frac{1}{b_\delta^4}\int_{\boldsymbol{k}}^{\Lambda} \frac{1}{P_{\rm L}(k)} (\delta_g^2)(-\boldsymbol{k})(\delta_g^2)(\boldsymbol{k}) - \frac{8}{b_\delta^2}\int_{\boldsymbol{x}} \delta_g^2(\boldsymbol{x})$$

new summaries: $\{\langle O^{(n)}[\delta_g] \,|\, O^{(m)}[\delta_g]\rangle\}$

$\longrightarrow$ correlators of powers of δ_g :

OC 2nd, LO: $\{\langle\delta_g\delta_g\rangle, \langle\delta_g\delta_g^2\rangle\}$

OC 2nd, full: $\{\langle\delta_g\delta_g\rangle, \langle\delta_g\delta_g^2\rangle, \langle\delta_g^2\delta_g^2\rangle\}$

OPERATOR CORRELATORS (OC)

MAP expression for b_{δ^2} : $b_{\delta^2} = \frac{N[\delta_g]}{D[\delta_g]}$ with

$$N[\delta_g] = \frac{1}{b_\delta^2} \int_{\boldsymbol{k}}^{\Lambda} \frac{1}{P_{\mathrm{L}}(k)} (\delta_g^2)(-\boldsymbol{k})\delta_g(\boldsymbol{k}) - 2 \int_{\boldsymbol{x}} \delta_g(\boldsymbol{x})$$

$$D[\delta_g] = \frac{1}{b_\delta^4} \int_{\boldsymbol{k}}^{\Lambda} \frac{1}{P_{\mathrm{L}}(k)} (\delta_g^2)(-\boldsymbol{k})(\delta_g^2)(\boldsymbol{k}) - \frac{8}{b_\delta^2} \int_{\boldsymbol{x}} \delta_g^2(\boldsymbol{x})$$

new summaries: $\{\langle O^{(n)}[\delta_g] \,|\, O^{(m)}[\delta_g]\rangle\}$

$\longrightarrow$ correlators of powers of δ_g :

OC 2nd, LO: $\{\langle\delta_g\delta_g\rangle, \langle\delta_g\delta_g^2\rangle\}$ — P + tree-level bispectrum

OC 2nd, full: $\{\langle\delta_g\delta_g\rangle, \langle\delta_g\delta_g^2\rangle, \langle\delta_g^2\delta_g^2\rangle\}$ — add part of LO trispectrum

OPERATOR CORRELATORS (OC)

MAP expression for b_{δ^2} : $b_{\delta^2} = \dfrac{N[\delta_g]}{D[\delta_g]}$ with

$$N[\delta_g] = \frac{1}{b_\delta^2} \int_{\boldsymbol{k}}^{\Lambda} \frac{1}{P_{\rm L}(k)} (\delta_g^2)(-\boldsymbol{k})\delta_g(\boldsymbol{k}) - 2\int_{\boldsymbol{x}} \delta_g(\boldsymbol{x})$$

$$D[\delta_g] = \frac{1}{b_\delta^4} \int_{\boldsymbol{k}}^{\Lambda} \frac{1}{P_{\rm L}(k)} (\delta_g^2)(-\boldsymbol{k})(\delta_g^2)(\boldsymbol{k}) - \frac{8}{b_\delta^2}\int_{\boldsymbol{x}} \delta_g^2(\boldsymbol{x})$$

new summaries: $\{\langle O^{(n)}[\delta_g] \,|\, O^{(m)}[\delta_g]\rangle\}$

$\longrightarrow$ correlators of powers of δ_g :

OC 2nd, LO: $\{\langle\delta_g\delta_g\rangle, \langle\delta_g\delta_g^2\rangle\}$ P + tree-level bispectrum

OC 2nd, full: $\{\langle\delta_g\delta_g\rangle, \langle\delta_g\delta_g^2\rangle, \langle\delta_g^2\delta_g^2\rangle\}$ add part of LO trispectrum

OC 3rd, LO: $\{\langle\delta_g\delta_g\rangle, \langle\delta_g\delta_g^2\rangle, \langle\delta_g^2\delta_g^2\rangle, \langle\delta_g\delta_g^3\rangle\}$ P + tree-level B + tree-level T

OC 3rd, full: $\{\langle\delta_g\delta_g\rangle, \langle\delta_g\delta_g^2\rangle, \langle\delta_g^2\delta_g^2\rangle, \langle\delta_g\delta_g^3\rangle, \langle\delta_g^2\delta_g^3\rangle, \langle\delta_g^3\delta_g^3\rangle\}$ add part of 5- and 6-point functions

FIELD-LEVEL POSTERIOR

$$\mathcal{P}(\boldsymbol{\theta}, \delta_{\rm in}, \{b_O\}, \{\sigma_\varepsilon\} | \delta_g^{\rm obs.}) \quad \longrightarrow \quad \mathcal{P}(\{b_O\}, \sigma_\varepsilon | \delta_g^{\rm obs.})$$

FIELD-LEVEL POSTERIOR

$$\mathcal{P}(\boldsymbol{\theta}, \delta_{\rm in}, \{b_O\}, \{\sigma_\varepsilon\} | \delta_g^{\rm obs.}) \quad \longrightarrow \quad \mathcal{P}(\{b_O\}, \sigma_\varepsilon | \delta_g^{\rm obs.})$$

EFT likelihood

$$\ln\mathcal{L}_{\rm EFT}(\delta_g^{\rm obs.} | \boldsymbol{\theta}, \delta_{\rm in}, \{b_O\}, \{\sigma_\varepsilon\}) =$$

$$-\frac{1}{2} \sum_{\boldsymbol{k}>0}^{|\boldsymbol{k}|<k_{\rm max}} \left(\ln 2\pi\sigma_\varepsilon^2(k) + \frac{1}{\sigma_\varepsilon^2(k)} |\delta_g^{\rm obs.}(\boldsymbol{k}) - \sum_O b_O O[\boldsymbol{\theta}, \delta_{\rm in}](\boldsymbol{k})|^2 \right)$$

FIELD-LEVEL POSTERIOR

$$\mathcal{P}(\boldsymbol{\theta}, \delta_{\rm in}, \{b_O\}, \{\sigma_\varepsilon\} | \delta_g^{\rm obs.}) \quad \longrightarrow \quad \mathcal{P}(\{b_O\}, \sigma_\varepsilon | \delta_g^{\rm obs.})$$

EFT likelihood

$$\ln \mathcal{L}_{\rm EFT}(\delta_g^{\rm obs.} | \boldsymbol{\theta}, \delta_{\rm in}, \{b_O\}, \{\sigma_\varepsilon\}) =$$

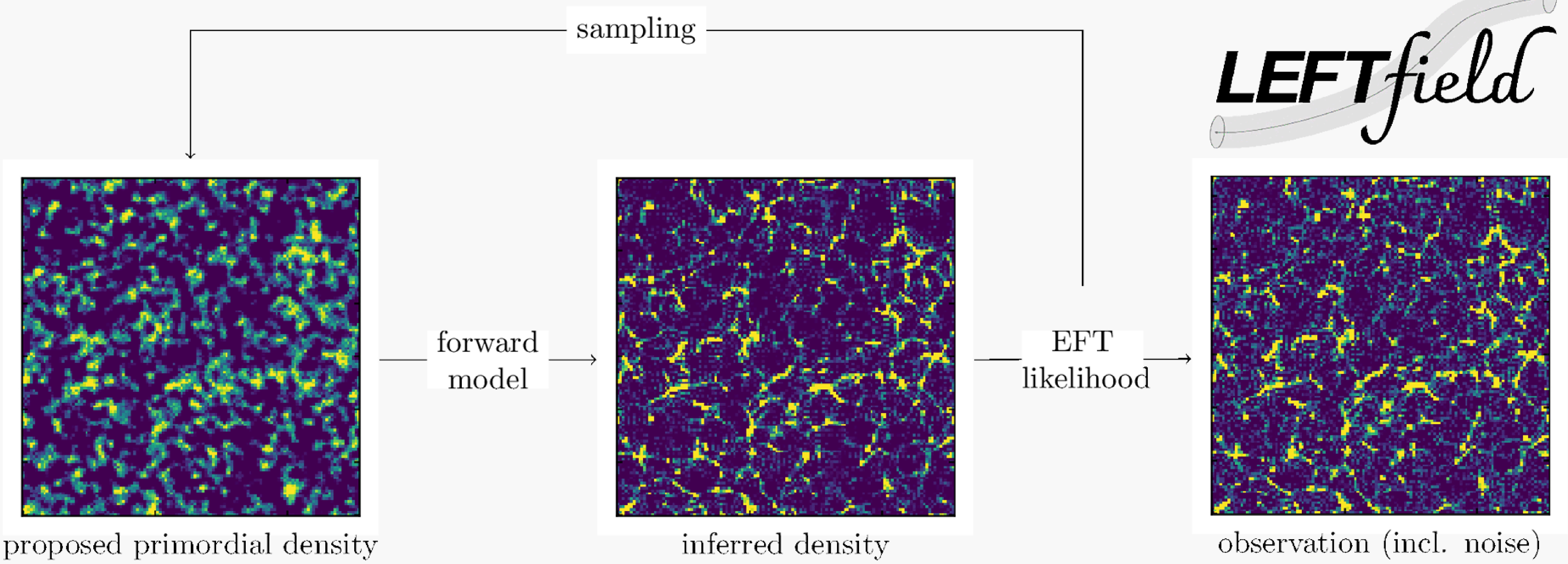
$$-\frac{1}{2} \sum_{\boldsymbol{k}>0}^{|\boldsymbol{k}|<k_{\rm max}} \left(\ln 2\pi\sigma_\varepsilon^2(k) + \frac{1}{\sigma_\varepsilon^2(k)} \left| \delta_g^{\rm obs.}(\boldsymbol{k}) - \sum_O b_O O[\boldsymbol{\theta}, \delta_{\rm in}](\boldsymbol{k}) \right|^2 \right)$$

data

model

FORWARD MODEL

FIELD-LEVEL BAYESIAN INFERENCE (FBI)



Schmidt (2021)

Credits: Julia Stadler

SIMULATION-BASED INFERENCE (SBI)

parameters drawn from prior

summary statistics

samples

$\theta \longrightarrow$ LEFTfield $\longrightarrow \delta_g \longrightarrow \mathbf{x} \longrightarrow$ density estimator

posterior

$\mathcal{P}(\boldsymbol{\theta}, \{b_O\}, \{\sigma_\varepsilon\} | \mathbf{x}^{\mathrm{obs}})$

$\mathbf{x}^{\mathrm{obs}}$

OC convergence test

$$k_{\max} = 0.14h\,\mathrm{Mpc}^{-1}$$
$$L_{\mathrm{box}} = 2000h^{-1}\,\mathrm{Mpc}$$
$$\Delta k \approx 2k_f$$

size of data vector:

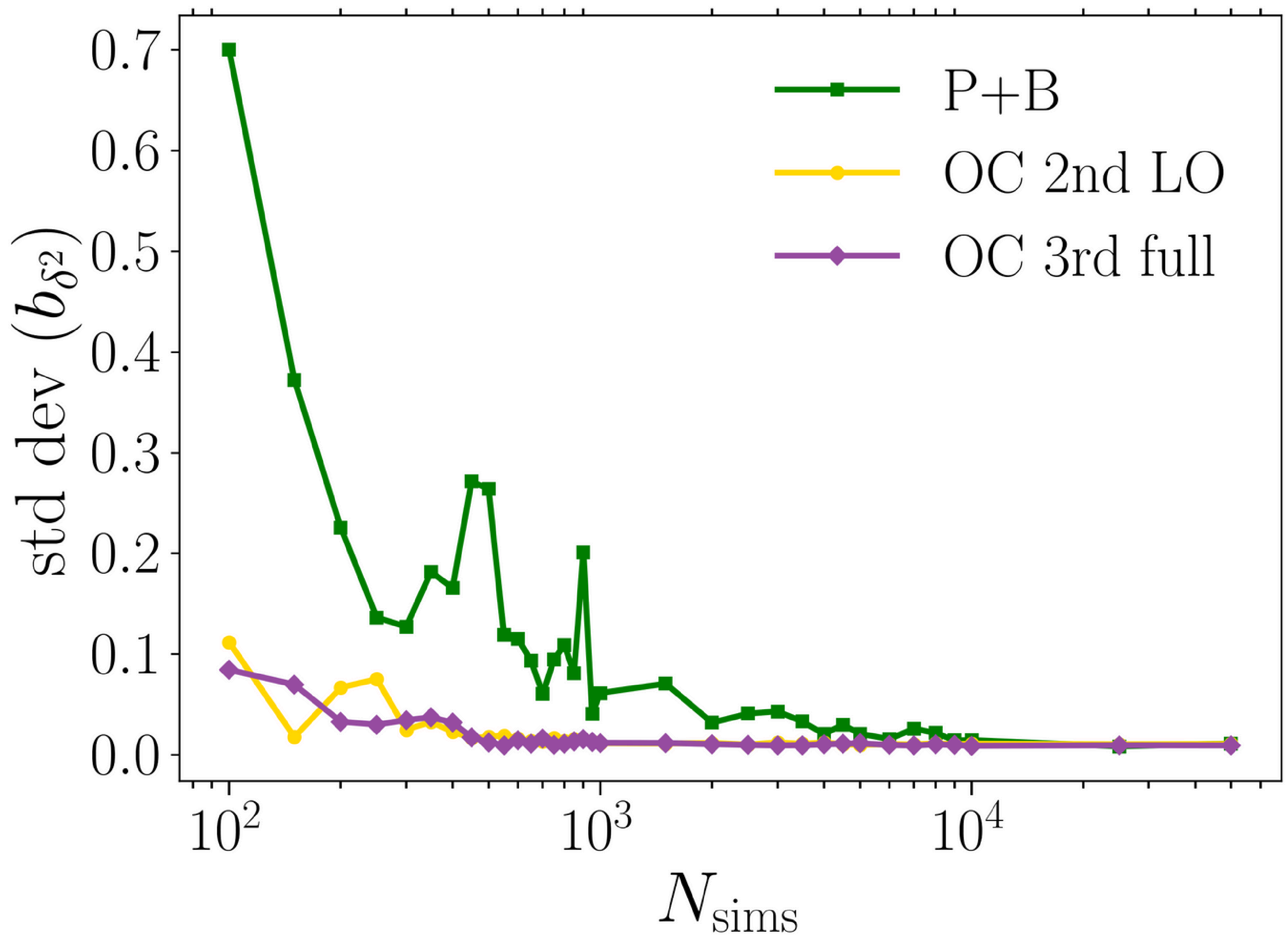
PB: 945 (PBT: ~ 15 000)

OC 2nd, LO: 40

OC 2nd, full: 60

OC 3rd, LO: 80

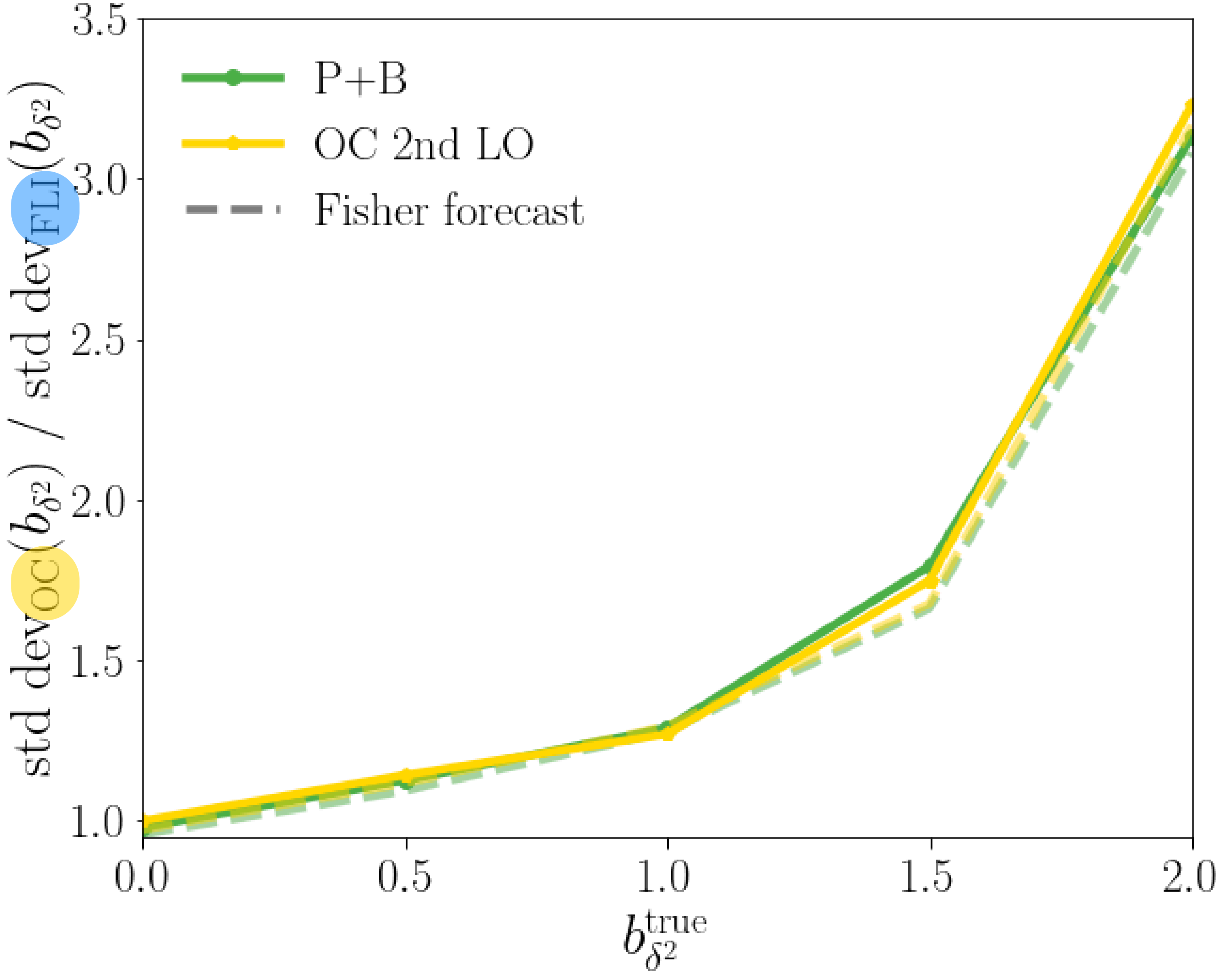
OC 3rd, full: 120

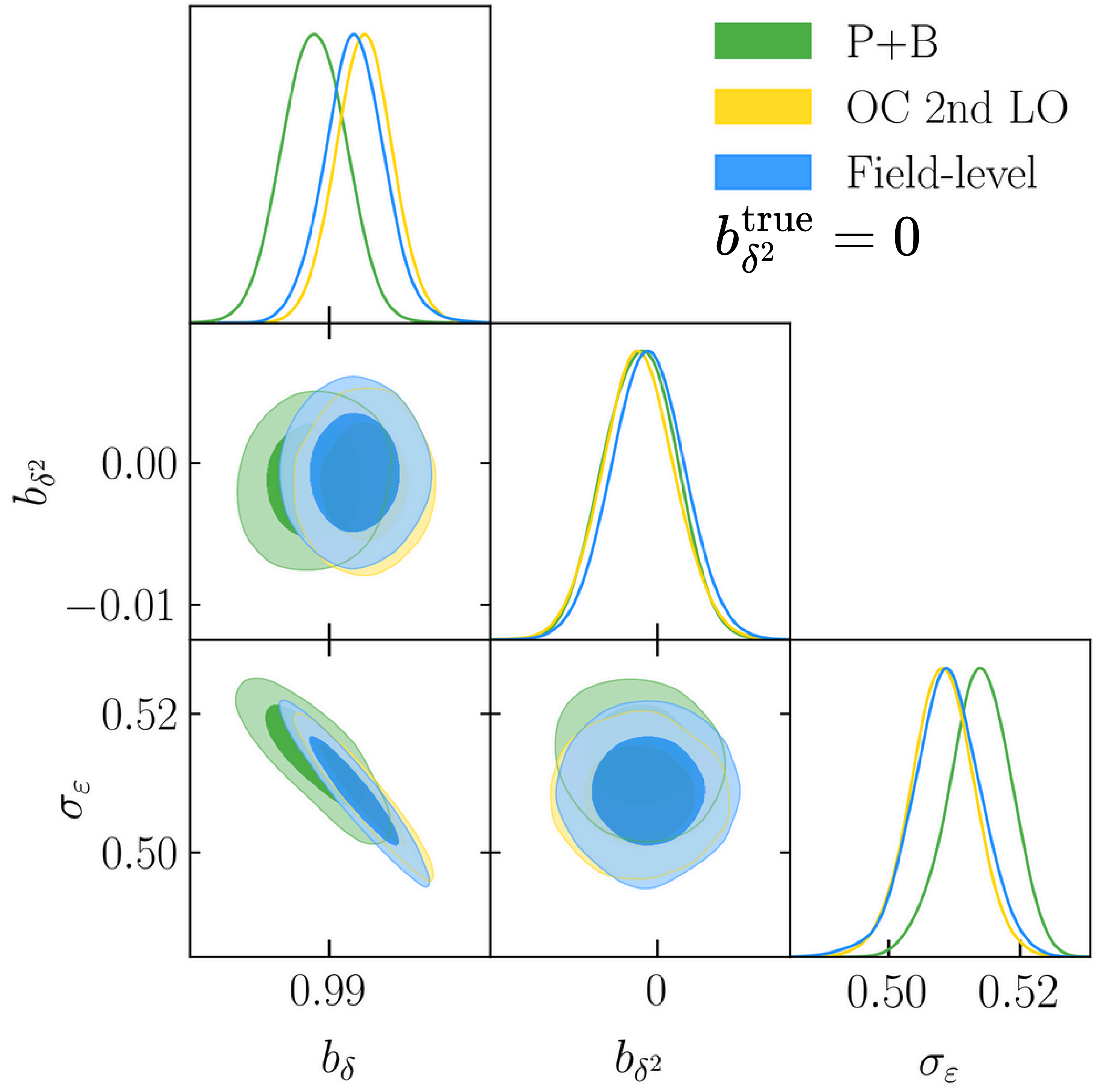


IN, Schmidt, Tucci, in prep.

convergence bellow 10^3 sims

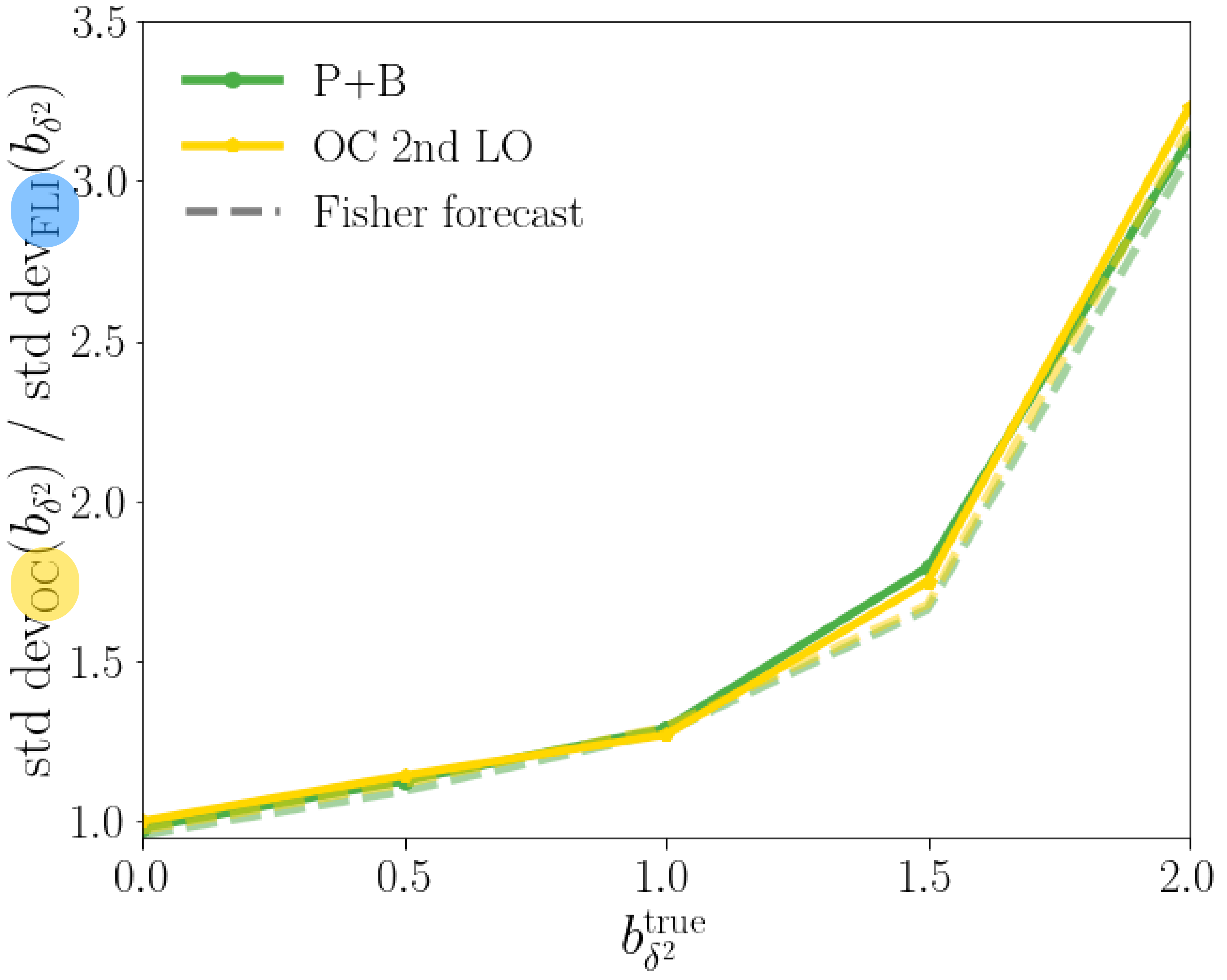
FBI to SBI comparison - OC2ndLO (vs P+B)

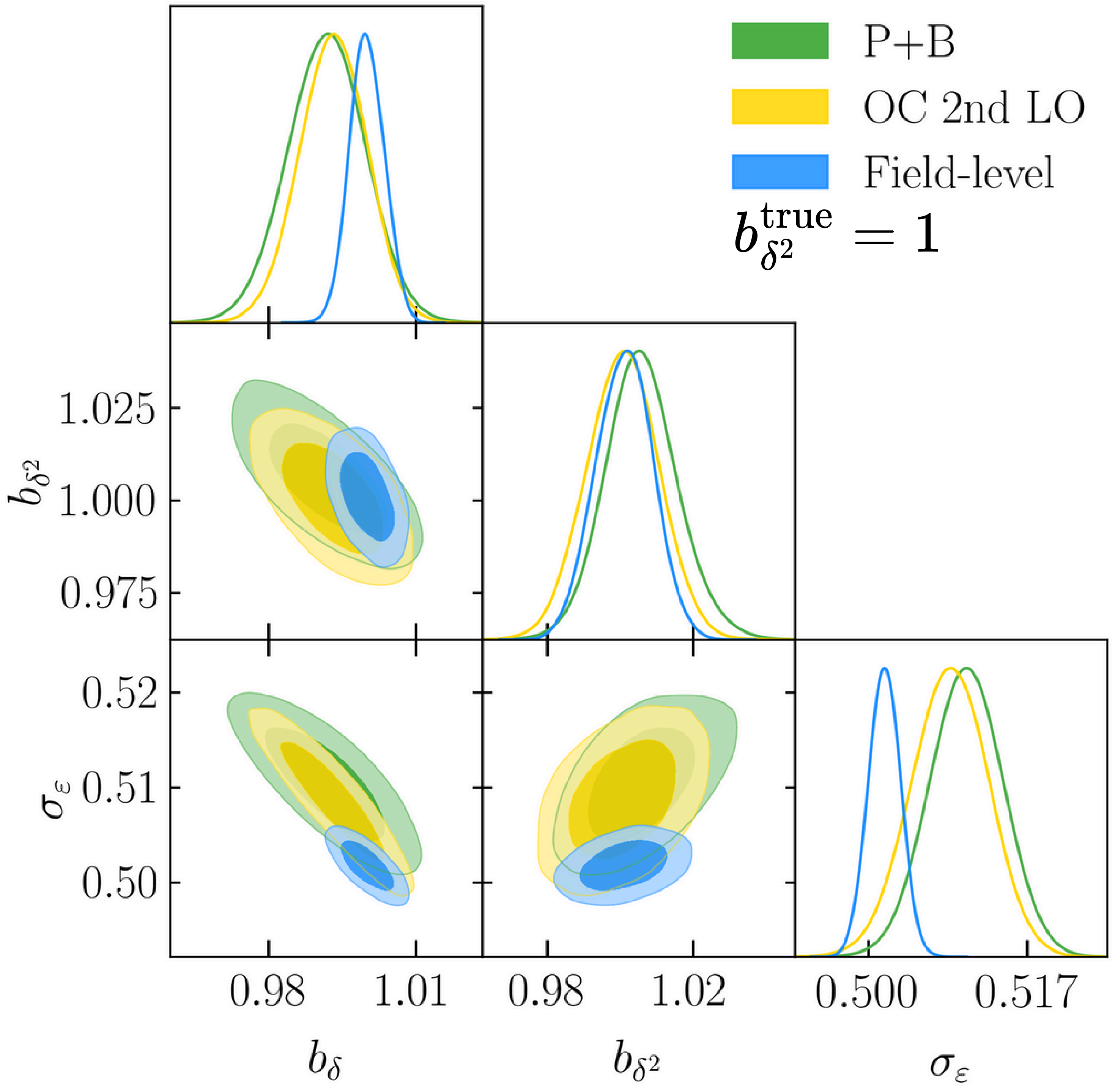




IN, Schmidt, Tucci, in prep.

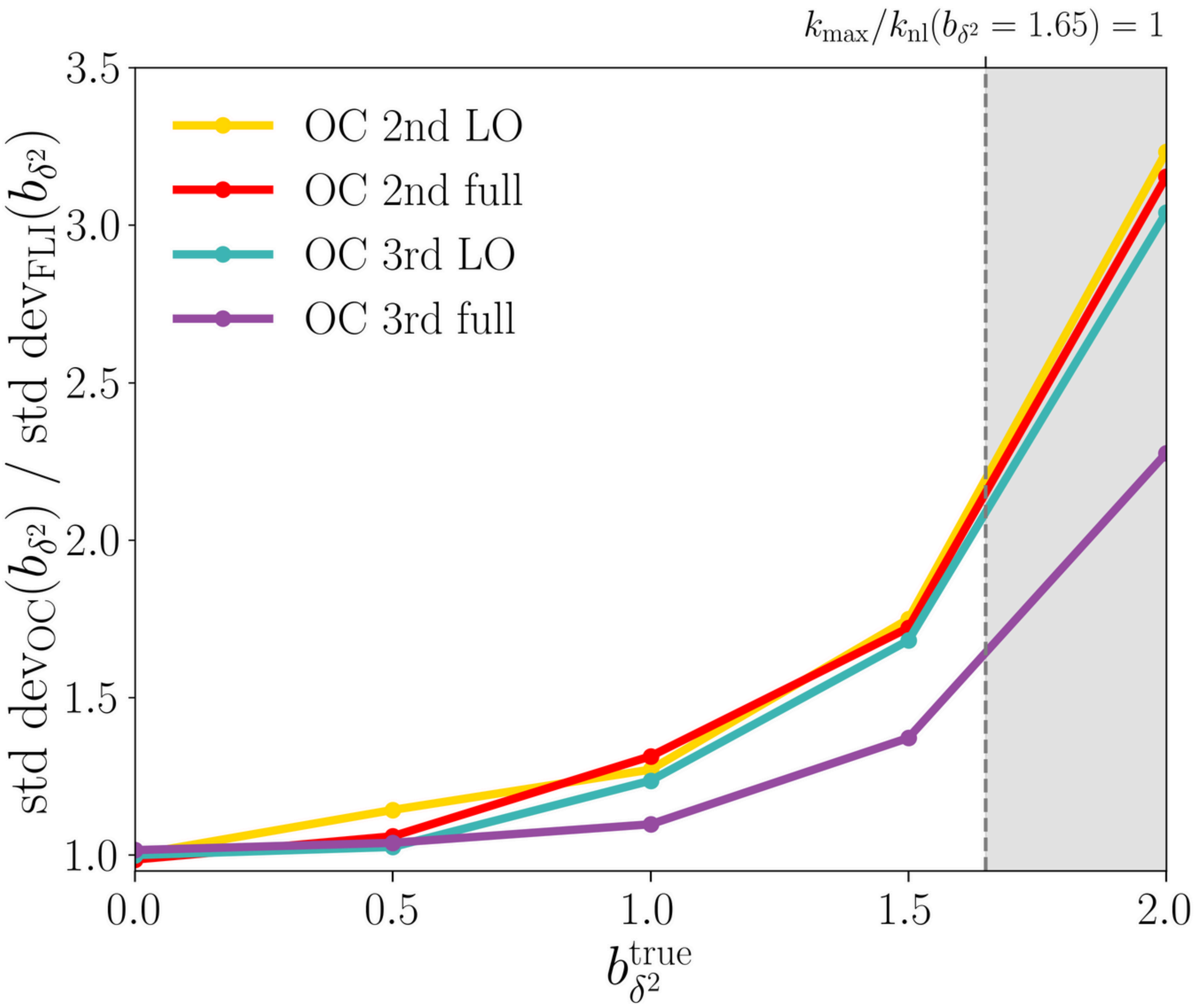
FBI to SBI comparison - OC2ndLO (vs P+B)





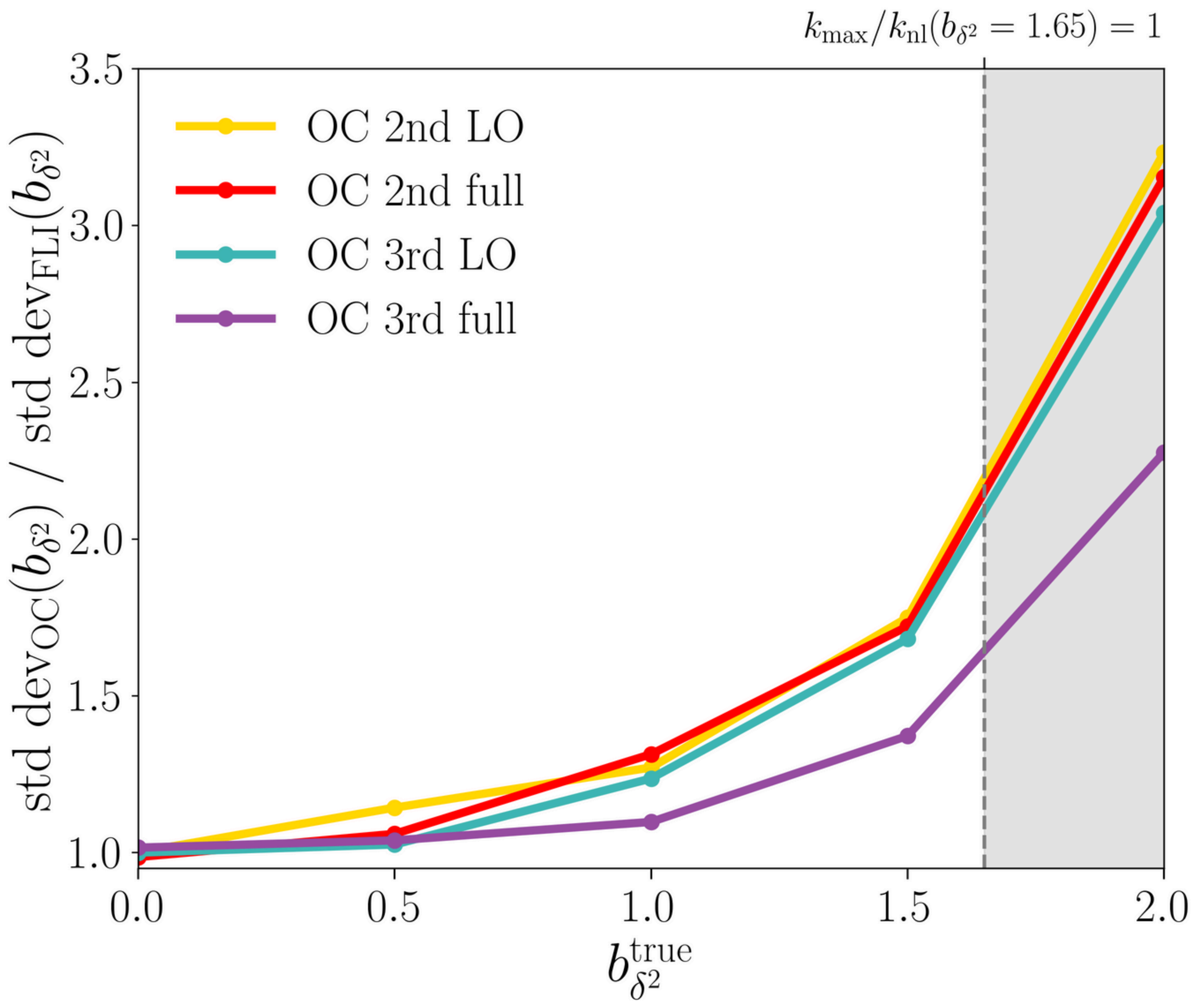
IN, Schmidt, Tucci, in prep.

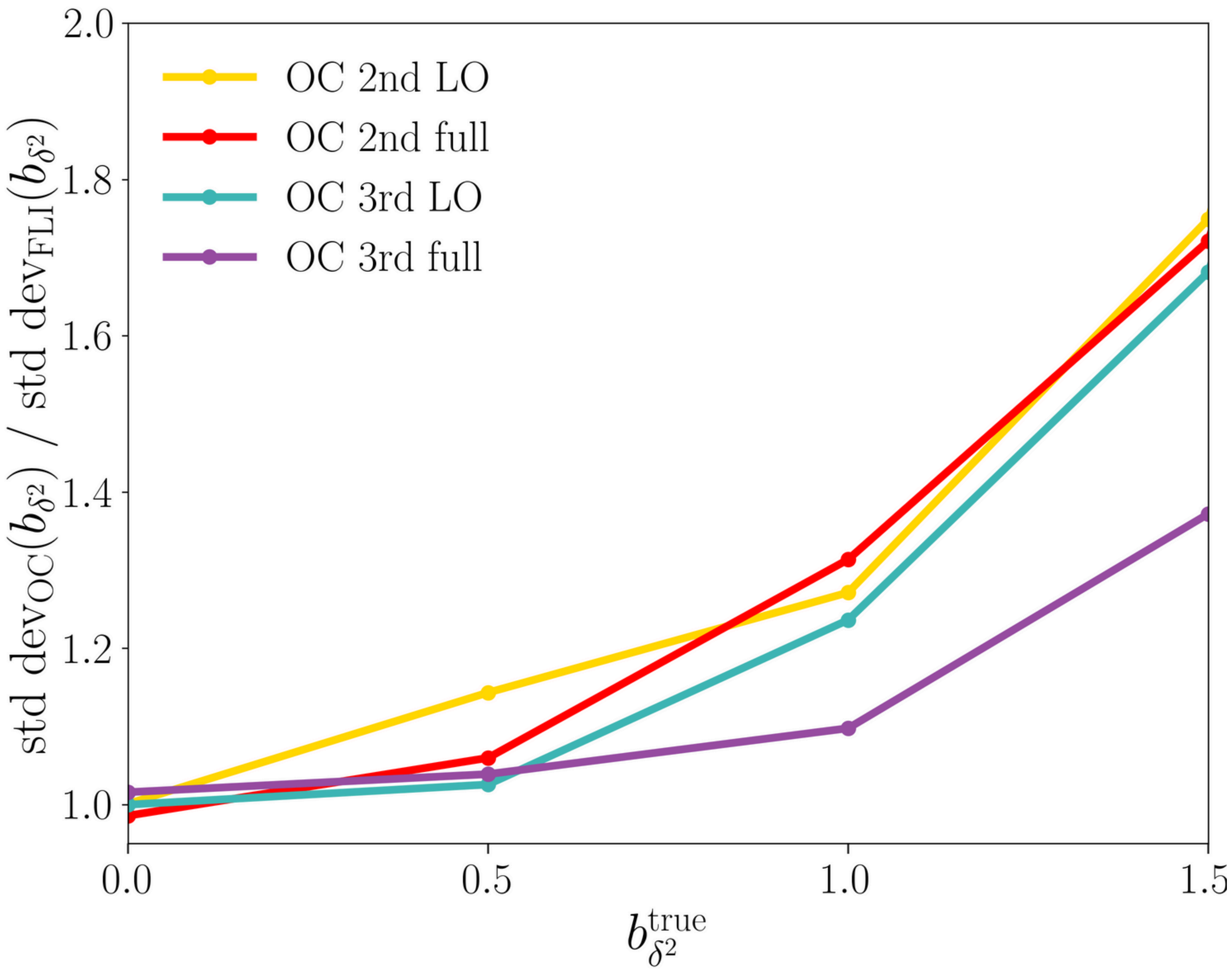
FBI to SBI comparison - OC cases



IN, Schmidt, Tucci, in prep.

FBI to SBI comparison - OC cases





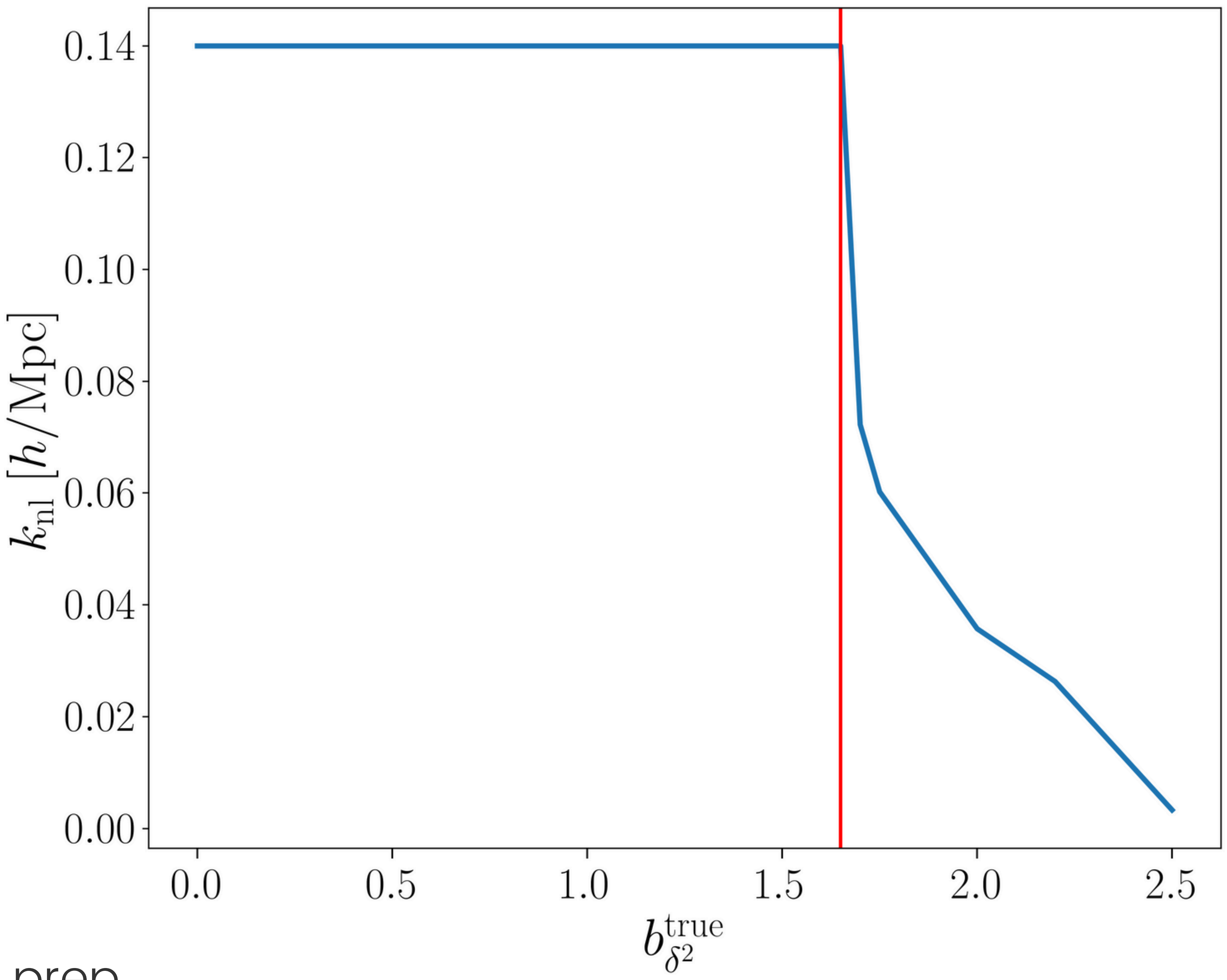
IN, Schmidt, Tucci, in prep.

KEY TAKEAWAYS

- Field-level inference with **LEFTfield** is a powerful tool for galaxy clustering analysis
- Apples-to-apples comparison of FBI and SBI summaries shows that there is a lot of reliable information beyond 2+3-point functions
- **Operator correlators (OC)** offer an efficient way of including higher order n-point functions in our analysis
- Going higher order in OCs extracts more information in the perturbative regime

DETERMINING k_nl

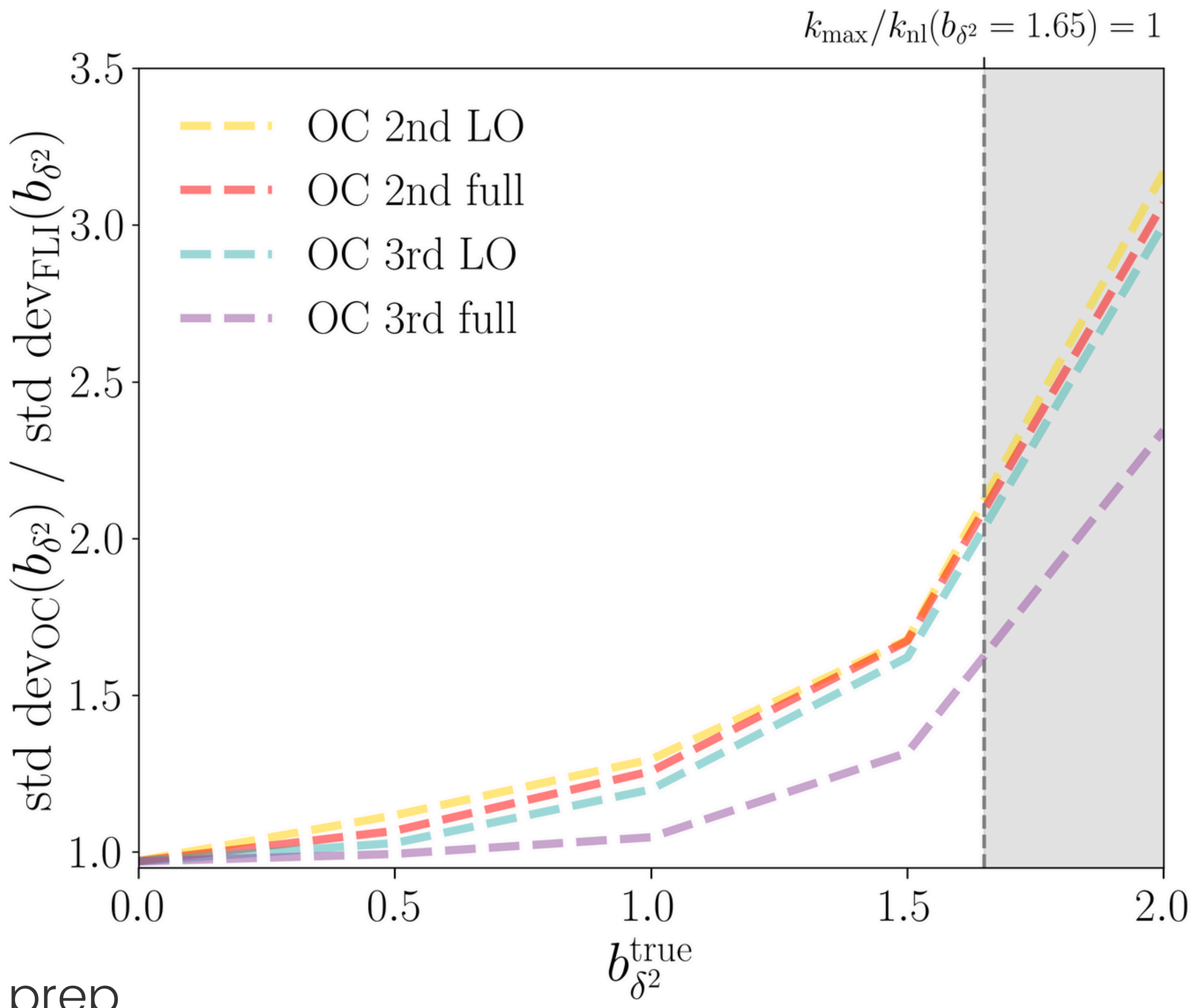
(backup slides)



$$\max_{k \le k_{\rm nl}} \frac{P_{\text{1-loop}}(k)}{P_{\rm tree}(k)} = 1$$

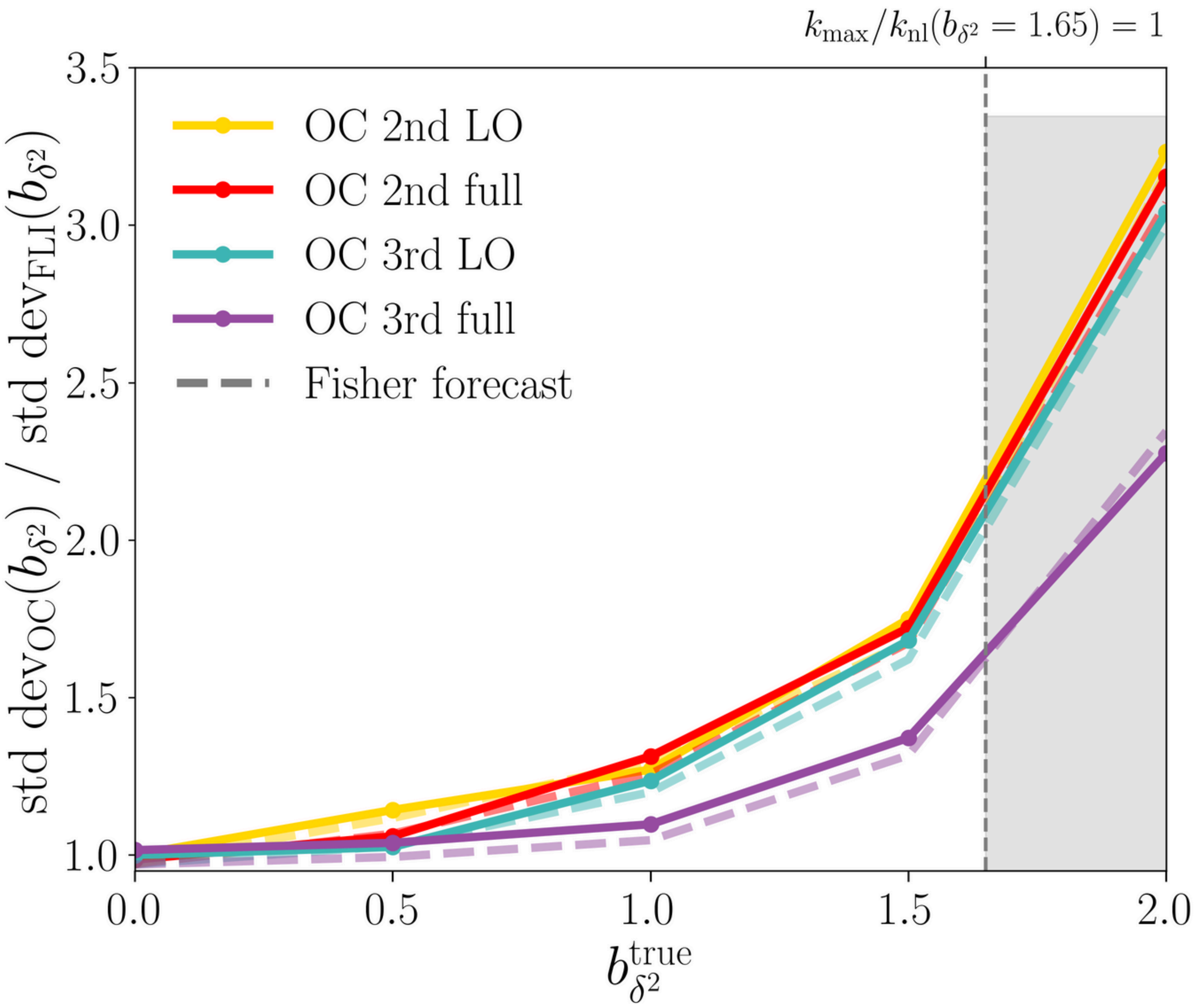
IN, Schmidt, Tucci, in prep.

FBI to SBI comparison - OC Fisher forecast (backup slides)



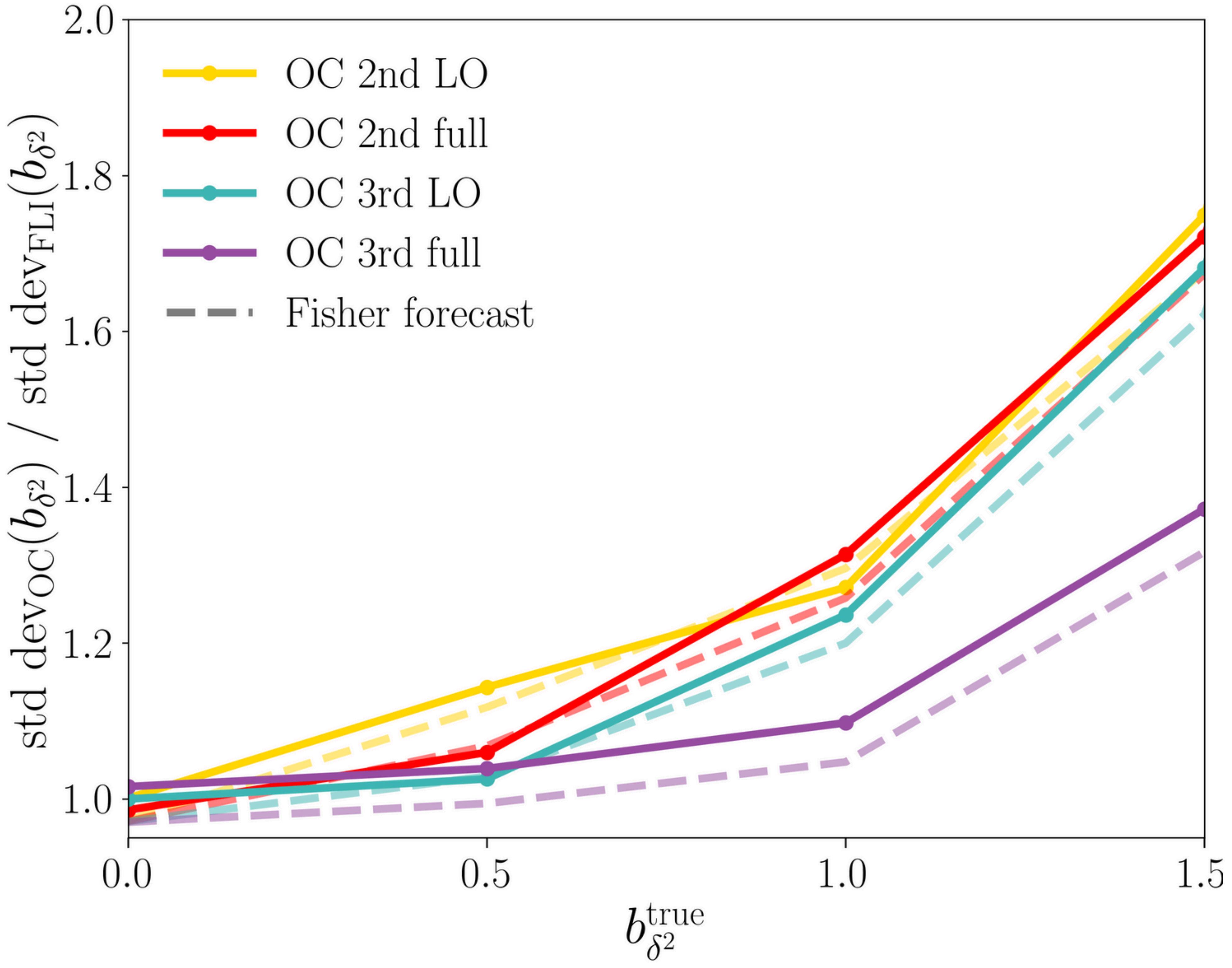
IN, Schmidt, Tucci, in prep.

FBI to SBI comparison - OC cases with Fisher (backup slides)



IN, Schmidt, Tucci, in prep.

FBI to SBI comparison - OC cases with Fisher (backup slides)



IN, Schmidt, Tucci, in prep.