BAYESIAN ESTIMATES FOR TH UNCERTAINTIES

Alexander Huss

Discussion of theoretical systematics in LHC precision measurements — February 26th 2024 — CERN

What is the Uncertainty Δ_{TH} of my Result?

- increasingly urgent to address with $\Delta_{\mathrm{EXP}} \searrow$ ($\leftrightsquigarrow$ HL-LHC)
  - what does 5σ mean if Δ_{TH} non-negligible?
  - interpretation of data in need for robust Δ_{TH}: PDF fits, χ^2 in ATLAS jets, …
- various sources that contribute to Δ_{TH}:
  - Δ_{α_s}, Δ_{param}: parametric uncertainties $\leftrightsquigarrow$ exp. extraction
  - Δ_{PDF}: parton distribution functions (PDFs) $\leftrightsquigarrow$ fits
  - $\Delta_{\text{non pert.}}$: hadronisation, UE, … $\leftrightsquigarrow$ parton showers [e.g. Herwig vs. Pythia]
  - Δ_{MHO}: *missing higher-order (MHO)* corrections

Focus here

Conventional Approach for Δ_{MHO} — Scale Variation

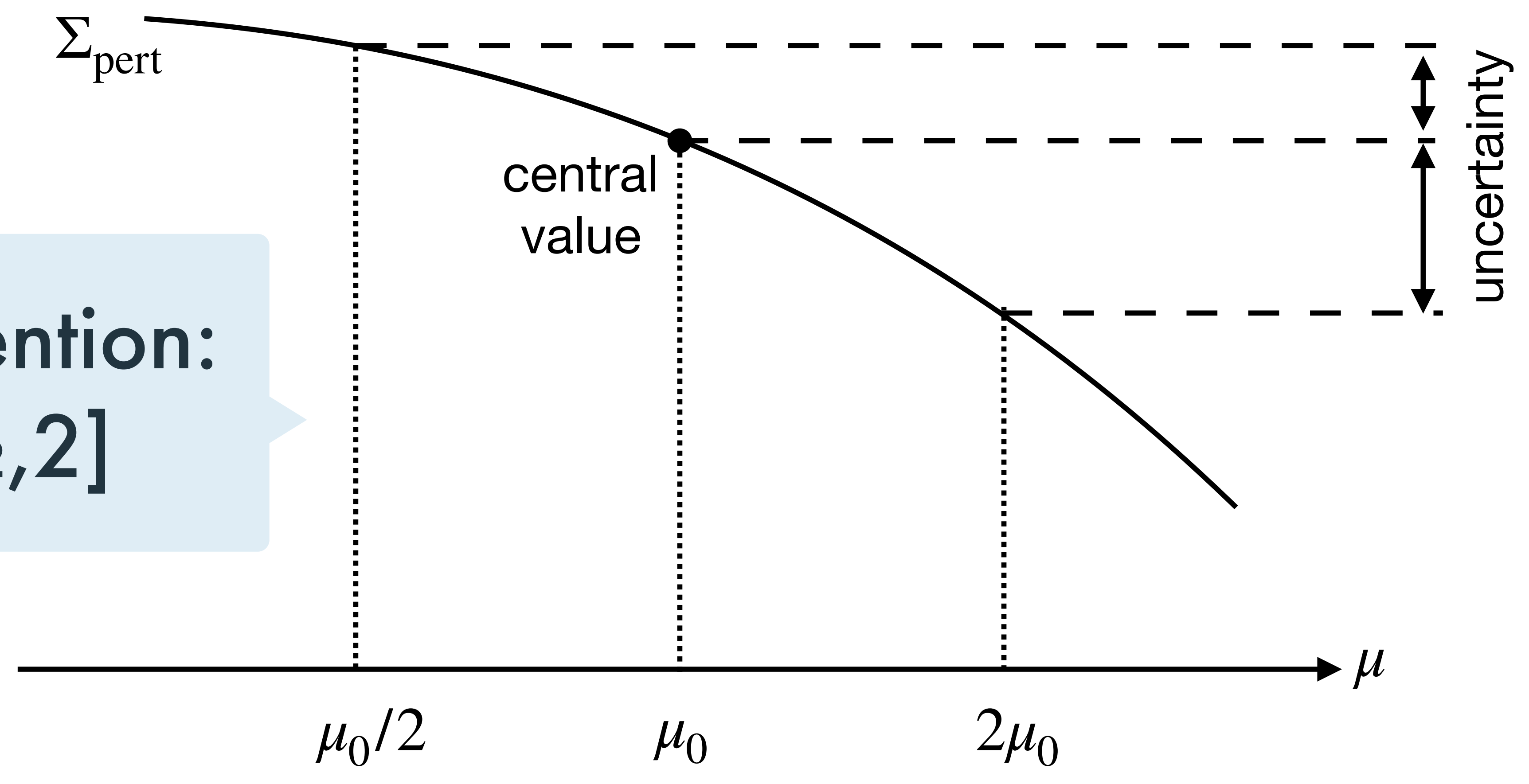
- approximation for an observable @ (next-to-)^n leading order:
  - N^{n}LO: $\Sigma \simeq \Sigma_n(\mu) = \sum_{k=0}^{n} \Sigma^{(k)}(\mu)$ $\quad \propto \alpha_s^{n_0+k}$

electroweak (EW):
↪ scheme dependence
↪ $\alpha \ll \alpha_s$

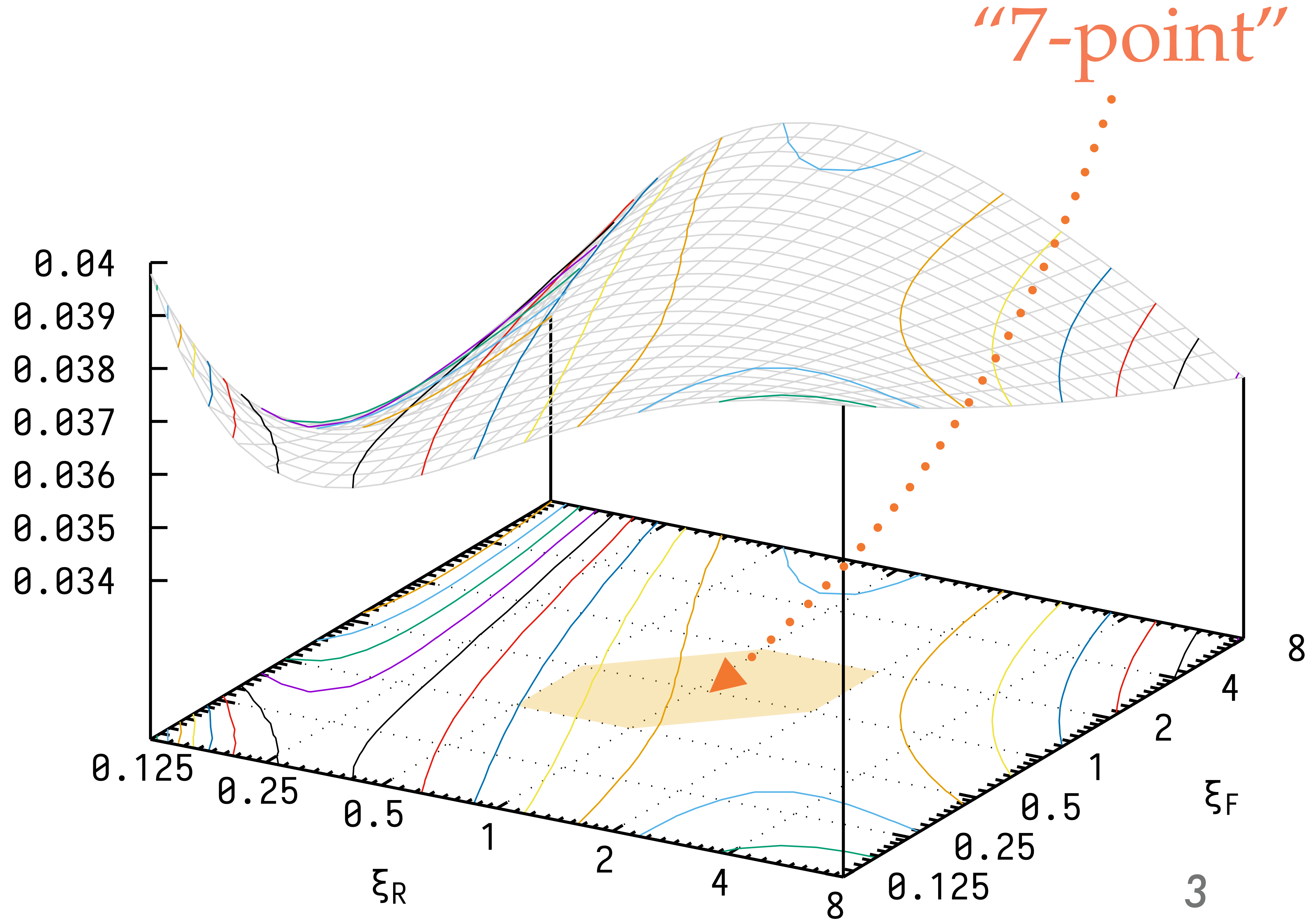
- truncation of series induces a sensitivity to terms of the next order

$$\mu \frac{\mathrm{d}}{\mathrm{d}\mu} \Sigma_n(\mu) = \mathcal{O}(\alpha_s^{n_0+n+1}) = \mathcal{O}(\Delta_{\mathrm{MHO}})$$

convention: [½,2]

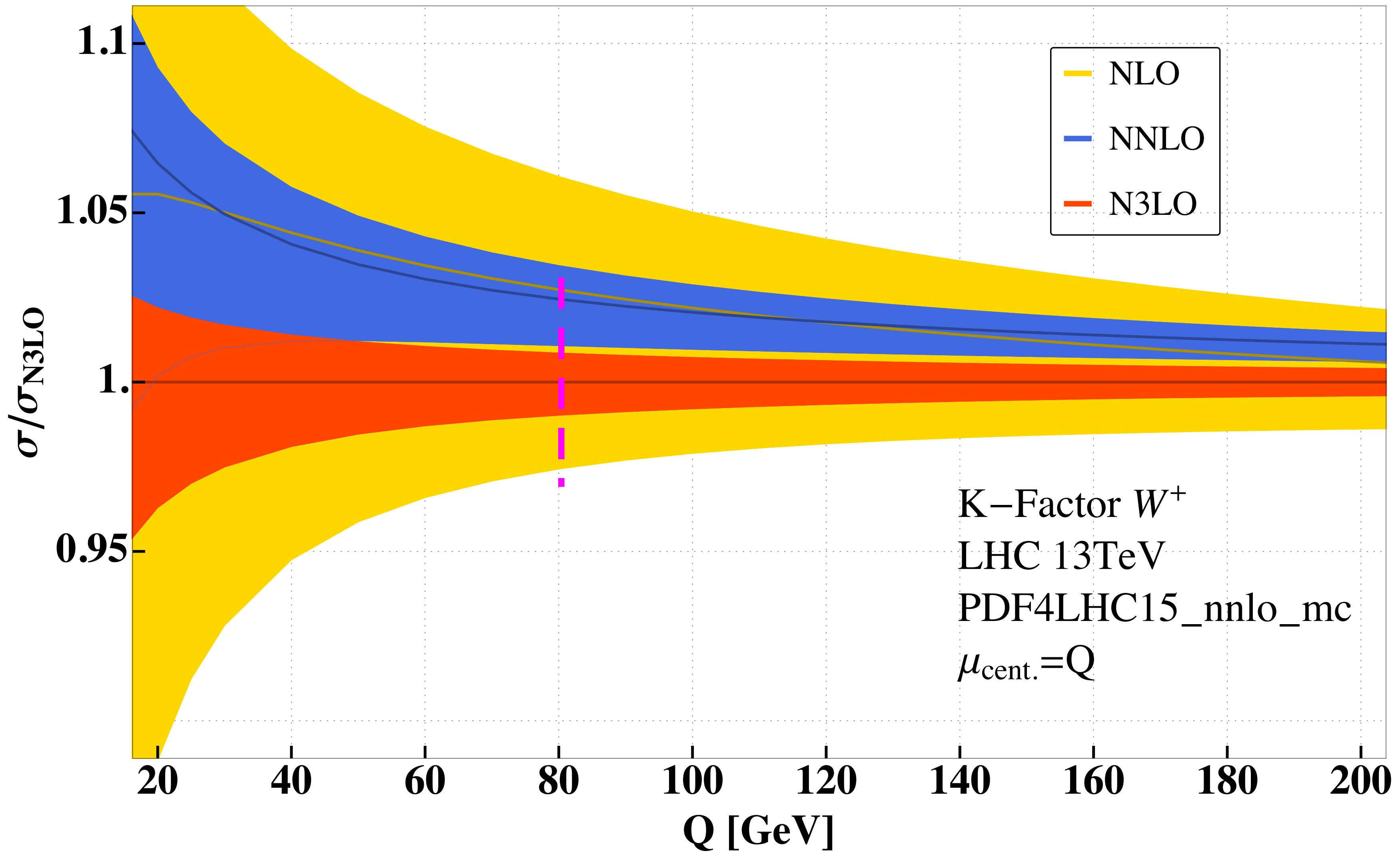


(μ_R, μ_F)



Issues with Standard Scale Variations

- known to be insufficient:
  - exclusive jet(s) (veto)
  - ratios (correlation?)
  - cancellations (e.g. $q\bar{q}$ vs. qg in DY)



[Duhr, Dulat, Mistlberger '20]

- choice of the central scale
  - fastest apparent convergence (FAC)
    $\hookrightarrow$ $\Sigma^{(n)}(\mu_{\text{FAC}}) = 0$
  - principle of minimal sensitivity (PMS)
    $\hookrightarrow$ $\frac{\partial}{\partial\mu}\Sigma^{(n)}(\mu)\Big|_{\mu_{\text{PMS}}} = 0$
  - BLM / PMC
    [Brodsky, Lepage, Mackenzie '83]; [Brodsky, Di Giustino '12]
  - …
- crucially: *no probabilistic interpretation!*
  $\rightsquigarrow$ can we do better?

Probability Distributions for Δ_{MHO}

[Cacciari, Houdeau '11]

- Sequence of perturbative *corrections* δ_k normalised w.r.t. LO (dimensionless)

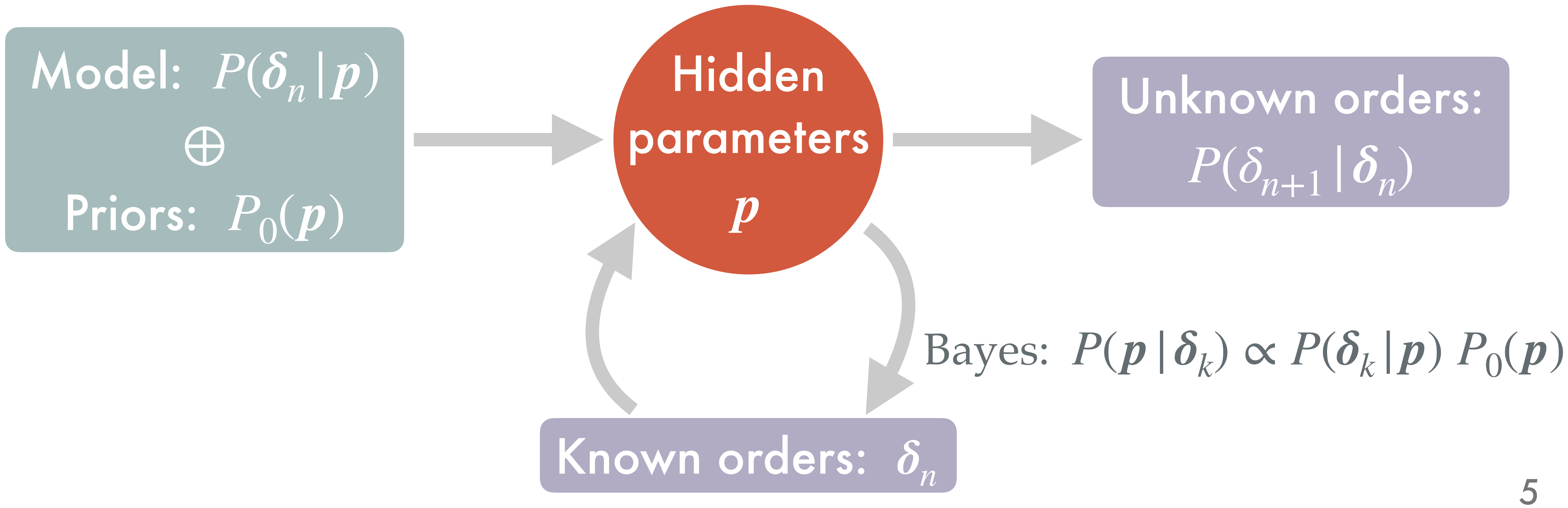
$$\Sigma_n = \Sigma^{(0)} \left(1 + \delta_1 + \ldots + \delta_n\right) \quad \rightsquigarrow \quad \delta_k = \mathcal{O}(\alpha_s^k)$$

- Probability distribution for δ_{n+1}, given $\boldsymbol{\delta}_n = (\delta_0, \delta_1, \ldots, \delta_n)$

$$P(\delta_{n+1} \,|\, \boldsymbol{\delta}_n) = \frac{P(\boldsymbol{\delta}_{n+1})}{P(\boldsymbol{\delta}_n)} = \frac{\int \mathrm{d}^m \boldsymbol{p}\; P(\boldsymbol{\delta}_{n+1} \,|\, \boldsymbol{p})\; P_0(\boldsymbol{p})}{\int \mathrm{d}^m \boldsymbol{p}\; P(\boldsymbol{\delta}_n \,|\, \boldsymbol{p})\; P_0(\boldsymbol{p})}$$

$$P(A,B) = P(A \,|\, B)\; P(B)$$

$$P(A) = \int \mathrm{d}B\; P(A,B)$$



The CH Model

[Cacciari, Houdeau '11]

- perturbative expansion $\delta_k = c_k\, \alpha_s^k$ bounded by a geometric series: $|c_k| \leq \bar{c} \quad \forall k$

$$\Big| \sum_k \delta_k \Big| \leq \sum_k |c_k|\, \alpha_s^k \leq \sum_k \bar{c}\, \alpha_s^k$$

  - one hidden parameter: $\bar{c}$
  - constrain upper bound $\bar{c}$ from known orders
    ⇝ constraint on unknown coefficients c_{n+1}

$c_k \sim \eta^k$
↪ survey of observables
[Bagnaschi, Cacciari, Guffanti, Jenniches '14]
↪ fitting
[Forte, Isgro, Vita '13]

- limitations:
  - α_s at what scale? why not: $\dfrac{\alpha_s}{\pi}$, $\dfrac{\alpha_s}{2\pi}$, $\alpha_s \ln^2(v)$, $\alpha_s \ln(v)$, … ?
- why not let the model figure out the expansion parameter itself?

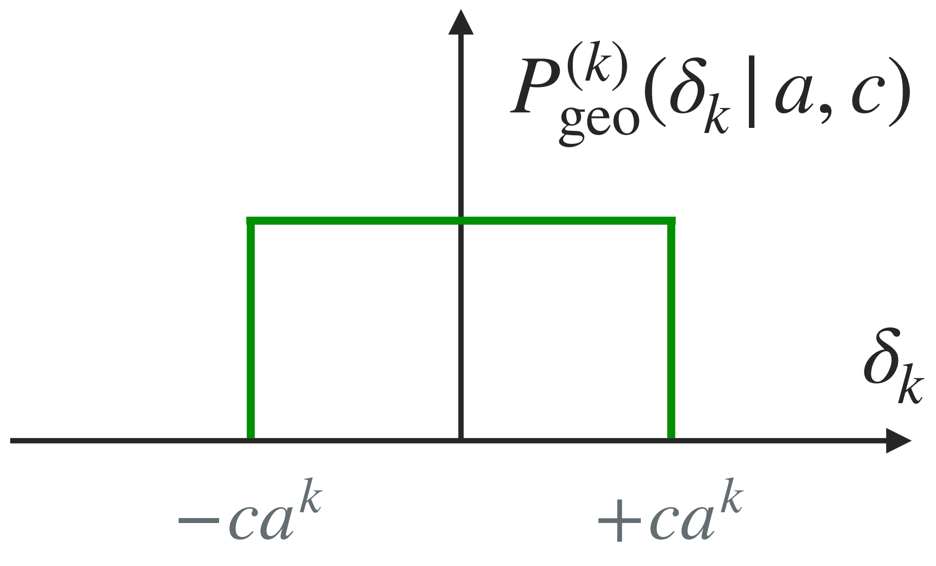
The Geometric Model

[Bonvini '20]

- bounded by a geometric series with expansion parameter a:

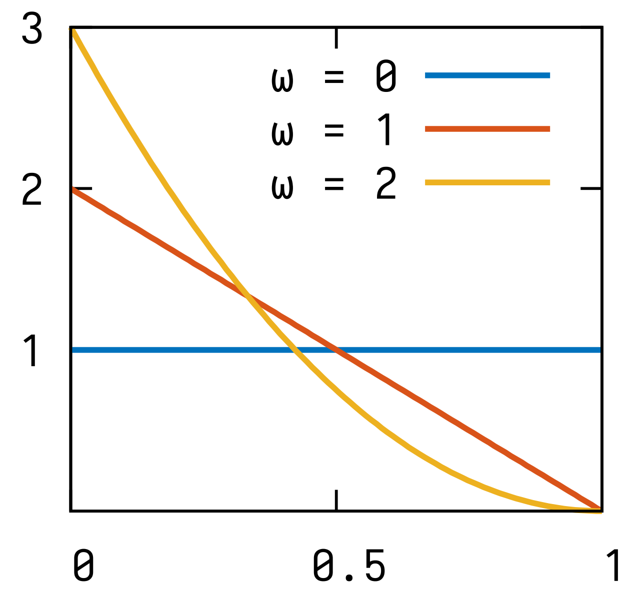
$$|\delta_k| \le c\, a^k \quad \forall k$$ ⇝ two model parameters: $a,\ c$

- **model:** $$P_{\rm geo}^{(k)}(\delta_k\,|\,a,c) = \frac{1}{2c\,a^k}\,\Theta\!\left(c - \frac{|\delta_k|}{a^k}\right)$$



- **priors:** $$P_0(a,c) = P_0(a)\,P_0(c)$$

$$P_0(a) = (1+\omega)\,(1-a)^\omega\,\Theta(a)\,\Theta(1-a)$$



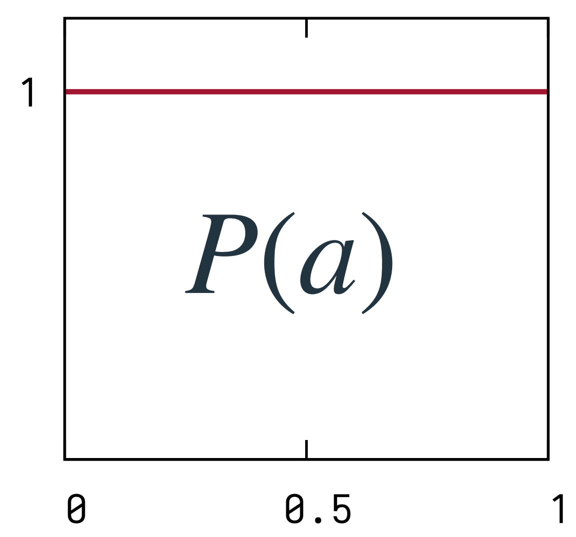
$$P_0(c) = \frac{\varepsilon}{c^{1+\varepsilon}}\,\Theta(c-1)$$ ⇝ $\mathrm{d}c/c \sim \mathrm{d}\ln(c)$ (ε: regulator)

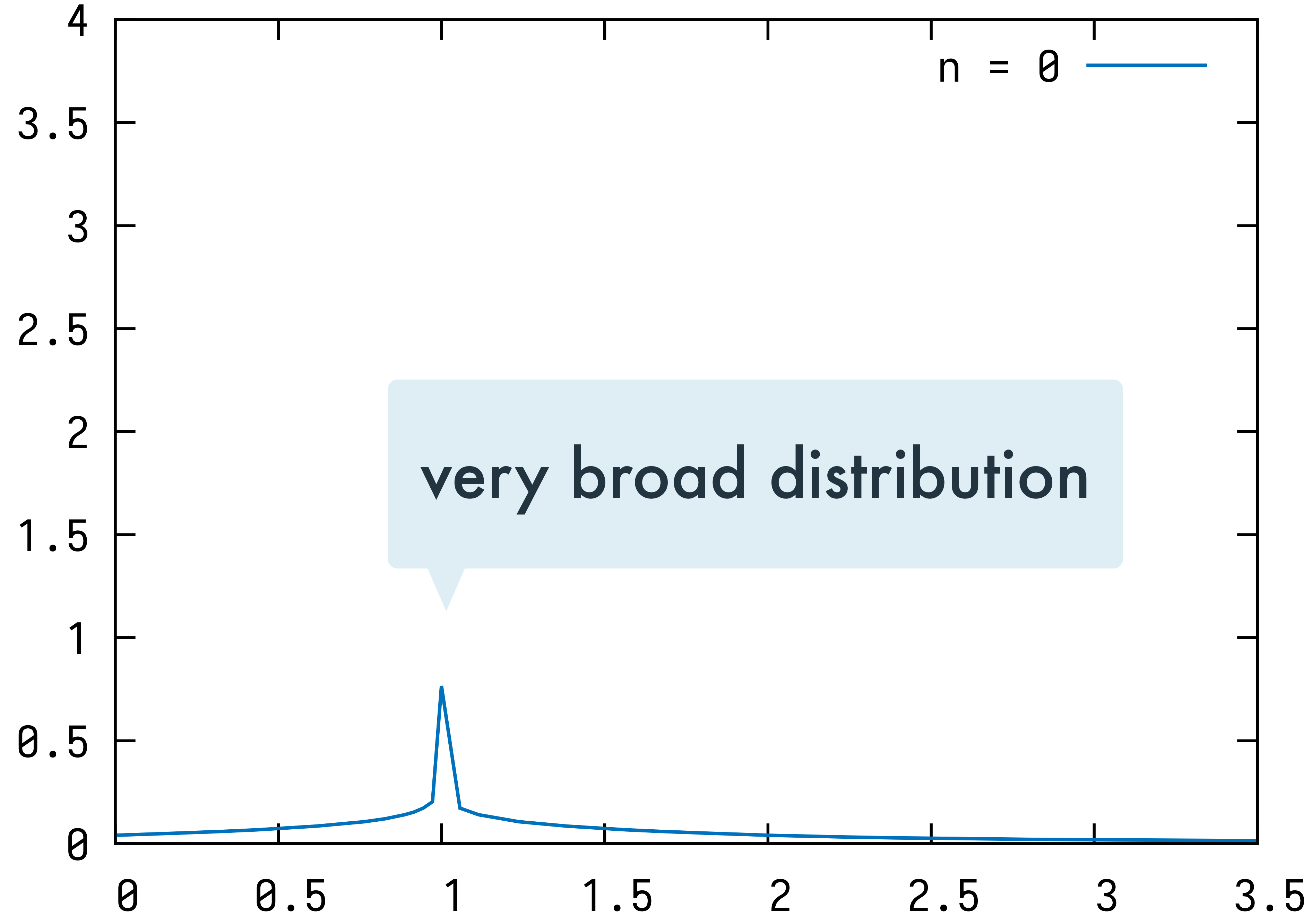
THE INFERENCE STEP — GEOMETRIC SERIES: $\delta_k = (0.7)^k$

- LO > $\delta_0 \equiv 1$

$$P_0(a,c) = \Theta(a)\ \Theta(1-a)\ P_0(c)$$

chose $\omega = 0$ for flat prior in a





no inference yet!
$P(\delta_1)$ entirely determined by the *model & priors*

$$P(\delta_1) = \int \mathrm{d}a \int \mathrm{d}c\ P^{(1)}_{\text{geo}}(\delta_1 \,|\, a,c)\ P_0(a,c)$$

The Inference Step — Geometric series: $\delta_k = (0.7)^k$

- **LO** > $\delta_0 \equiv 1$

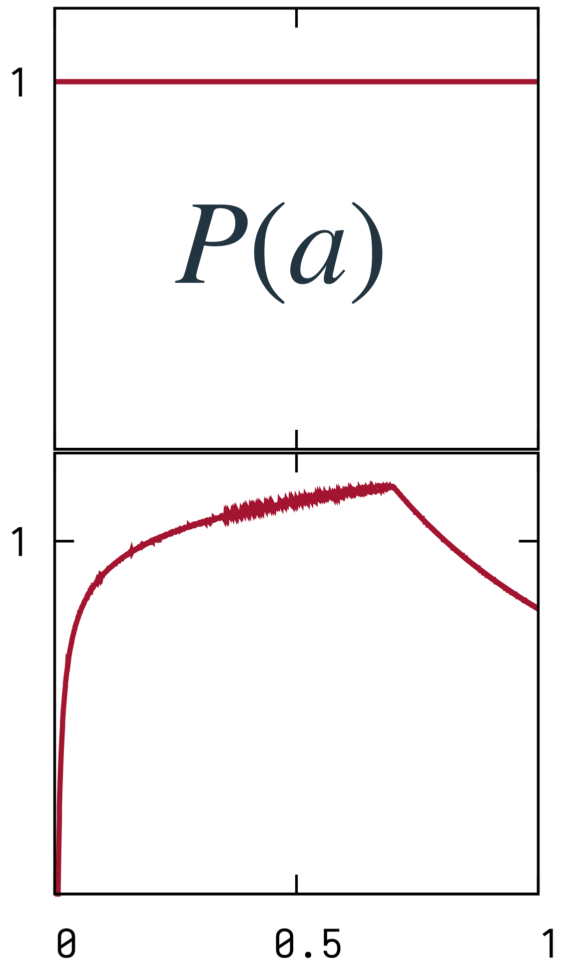
$$P_0(a,c) = \Theta(a)\,\Theta(1-a)\,P_0(c)$$

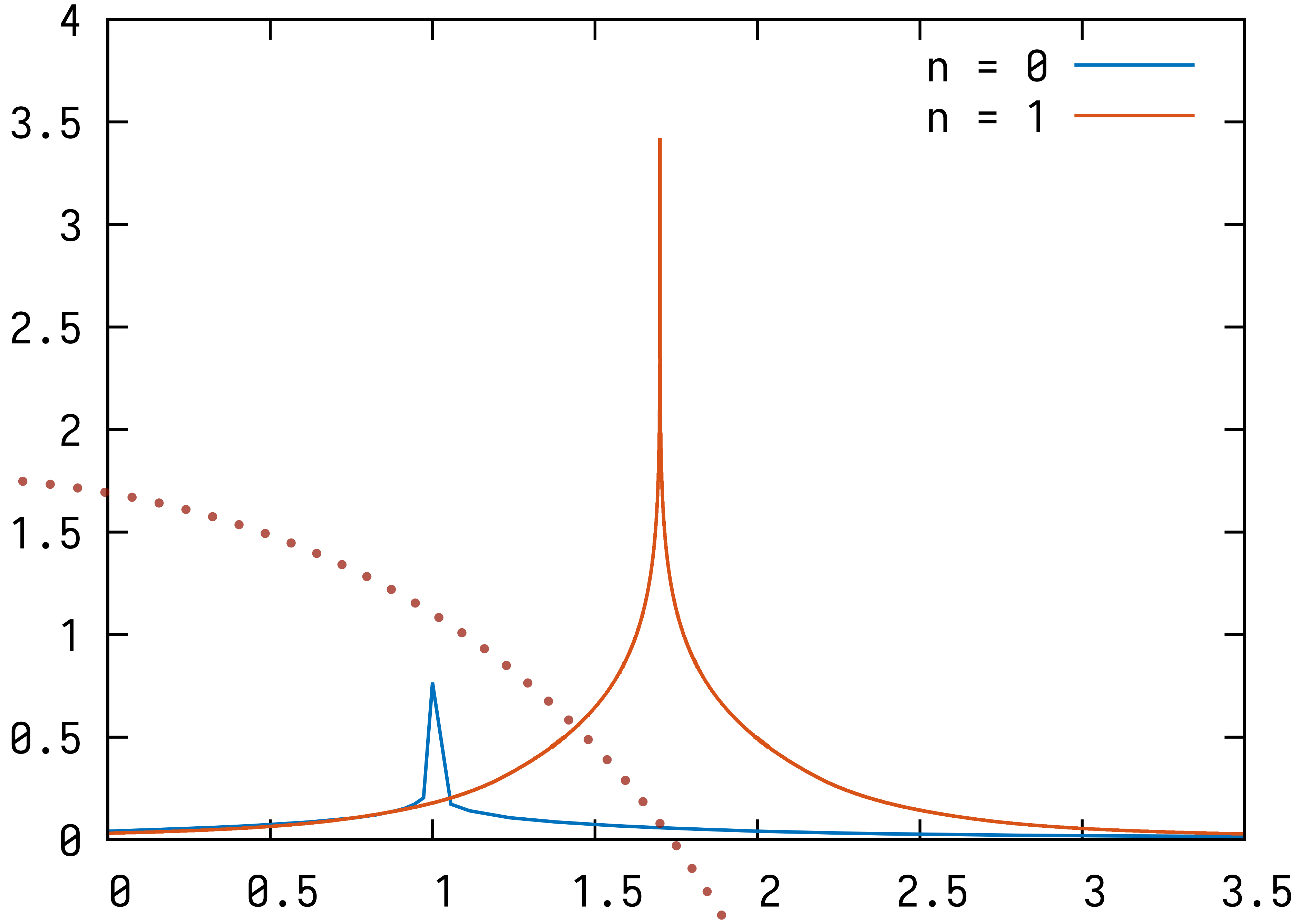
- **NLO** > $\delta_1 = 0.7$

$$\underbrace{P(a,c\,|\,\delta_1)}_{\text{posterior}} \propto \underbrace{P^{(1)}_{\text{geo}}(\delta_1\,|\,a,c)}_{\text{likelihood}}\;\underbrace{P_0(a,c)}_{\text{prior}}$$

Bayes' theorem:

$$P(A\,|\,B) = \frac{P(B\,|\,A)\,P(A)}{P(B)}$$





$$P(\delta_2\,|\,\delta_1) = \int da \int dc\; P(\delta_2\,|\,\delta_1,a,c)\;P(a,c\,|\,\delta_1)$$

$$\propto \int da \int dc\; P^{(2)}_{\text{geo}}(\delta_2\,|\,a,c)\;P^{(1)}_{\text{geo}}(\delta_1\,|\,a,c)\;P_0(a,c)$$

δ_k independent:

$$P(\delta_2\,|\,\delta_1) = P(\delta_2)$$

The Inference Step — Geometric series: $\delta_k = (0.7)^k$

- **LO** > $\delta_0 \equiv 1$

$$P_0(a,c) = \Theta(a)\,\Theta(1-a)\,P_0(c)$$

- **NLO** > $\delta_1 = 0.7$

$$P(a,c\,|\,\delta_1) \propto P^{(1)}_{\rm geo}(\delta_1\,|\,a,c)\;P_0(a,c)$$

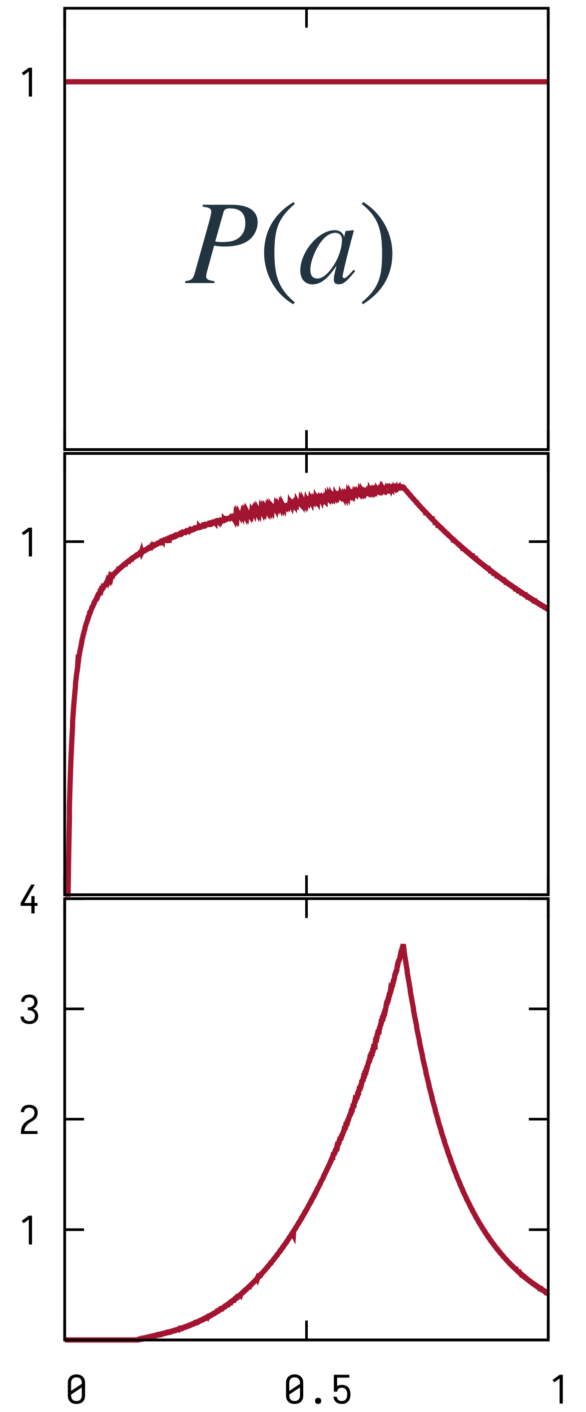
- **N²LO** > $\delta_2 = 0.7^2$

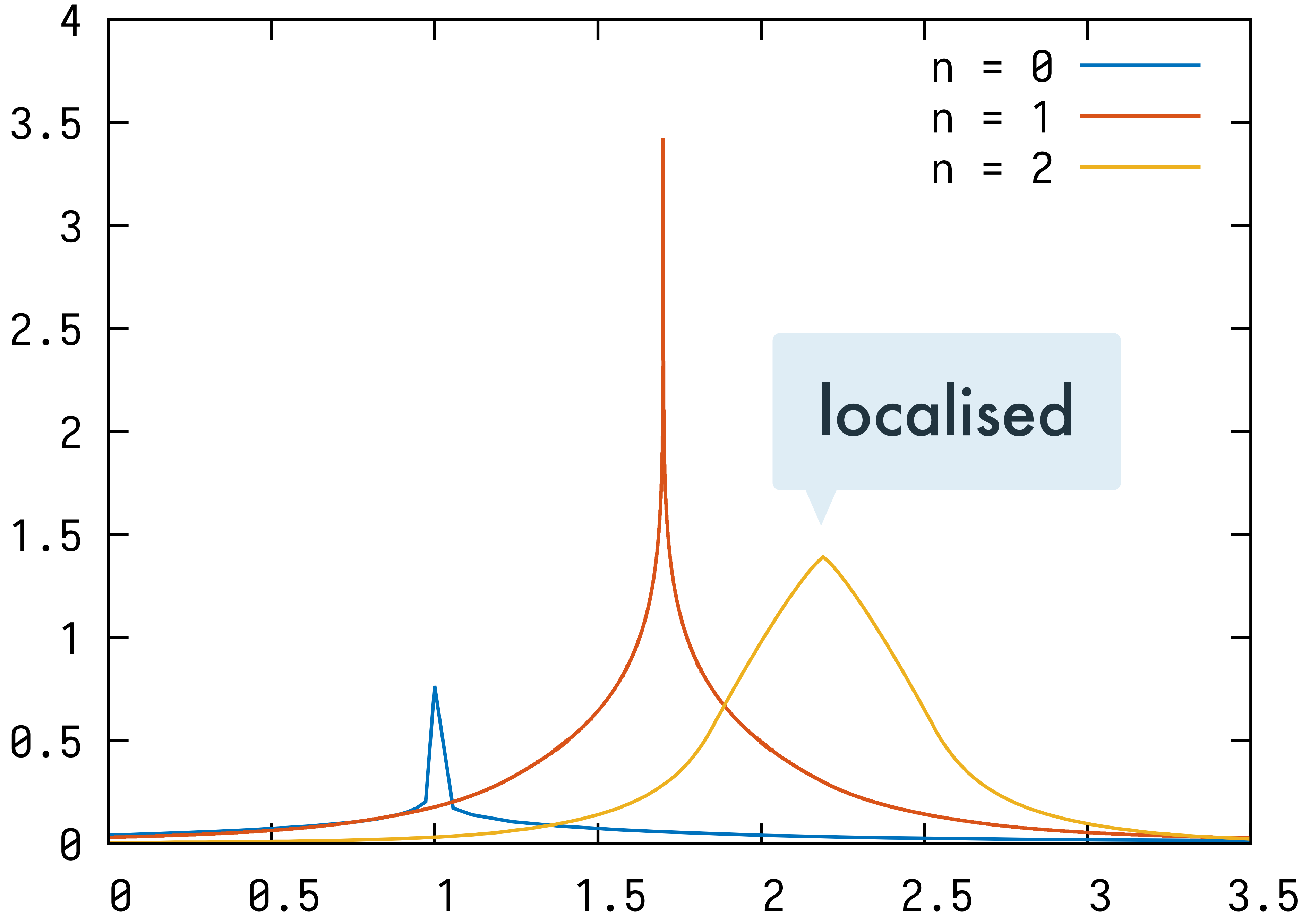
$$P(a,c\,|\,\delta_1,\delta_2) \propto P(\delta_2\,|\,\delta_1,a,c)\;P(a,c\,|\,\delta_1)$$

$$\propto P^{(2)}_{\rm geo}(\delta_2\,|\,a,c)\;P^{(1)}_{\rm geo}(\delta_1\,|\,a,c)\;P_0(a,c)$$

Bayes' theorem & independence

$a \sim 0.7$
also: $c \sim 1$





$$P(\delta_3\,|\,\delta_1,\delta_2) \propto \int {\rm d}a \int {\rm d}c\;\prod_{k=1}^{3}\left[P^{(k)}_{\rm geo}(\delta_k\,|\,a,c)\right] P_0(a,c)$$

The Inference Step — Geometric series: $\delta_k = (0.7)^k$

- **LO** > $\delta_0 \equiv 1$

  $$P_0(a,c) = \Theta(a)\,\Theta(1-a)\,P_0(c)$$

- **NLO** > $\delta_1 = 0.7$

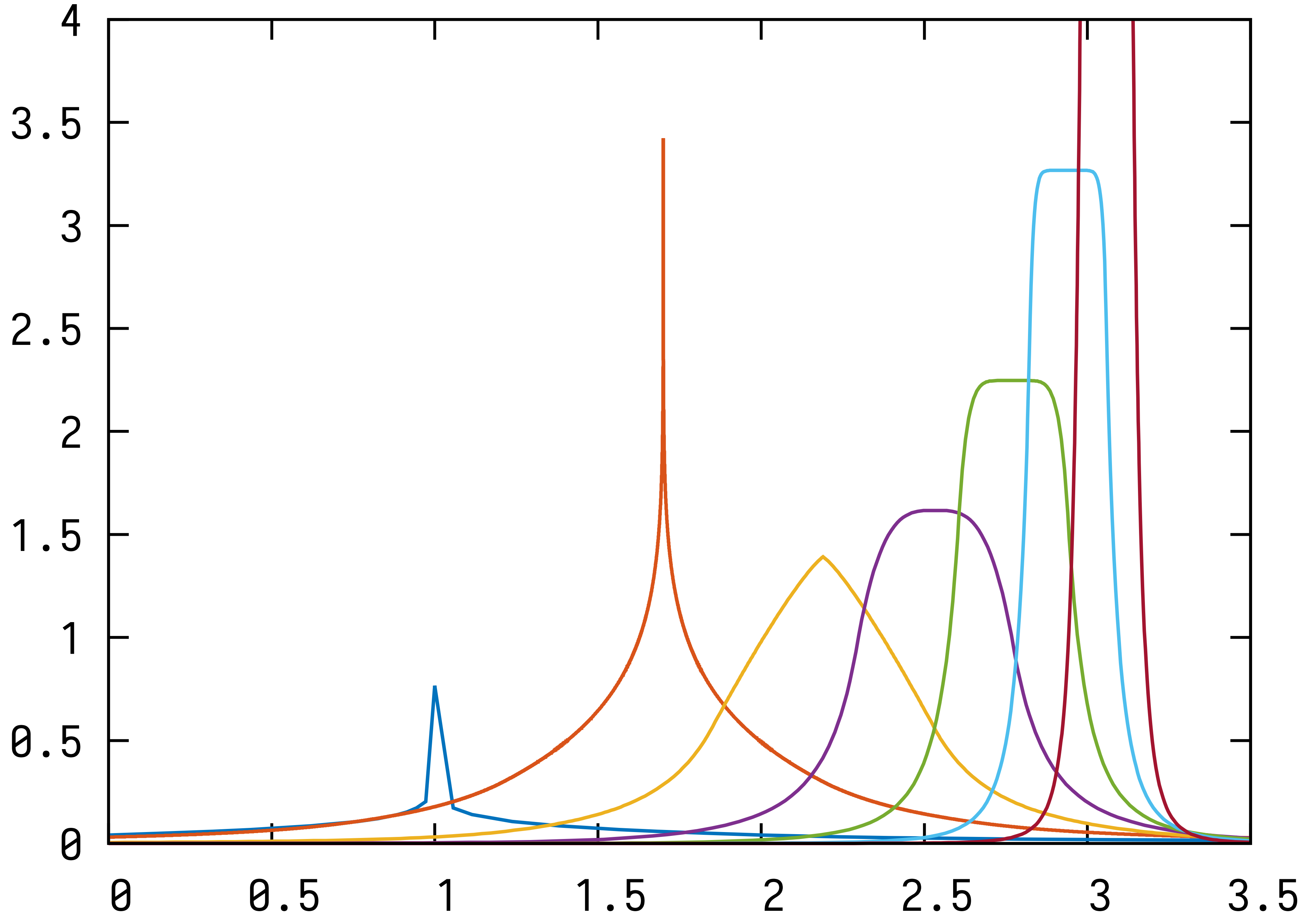
  $$P(a,c\,|\,\delta_1) \propto P^{(1)}_{\text{geo}}(\delta_1\,|\,a,c)\,P_0(a,c)$$

- **N²LO** > $\delta_2 = 0.7^2$

  $$P(a,c\,|\,\delta_1,\delta_2) \propto P(\delta_2\,|\,\delta_1,a,c)\,P(a,c\,|\,\delta_1)$$

  $$\propto P^{(2)}_{\text{geo}}(\delta_2\,|\,a,c)\,P^{(1)}_{\text{geo}}(\delta_1\,|\,a,c)\,P_0(a,c)$$

- …



$$P(\delta_{n+1}\,|\,\boldsymbol{\delta}_n) \propto \int \mathrm{d}a \int \mathrm{d}c\, \prod_{k=1}^{n} \left[P^{(k)}_{\text{geo}}(\delta_k\,|\,a,c)\right] P_0(a,c)$$

can be solved analytically

The *abc* Model — Asymmetric Geometric Model

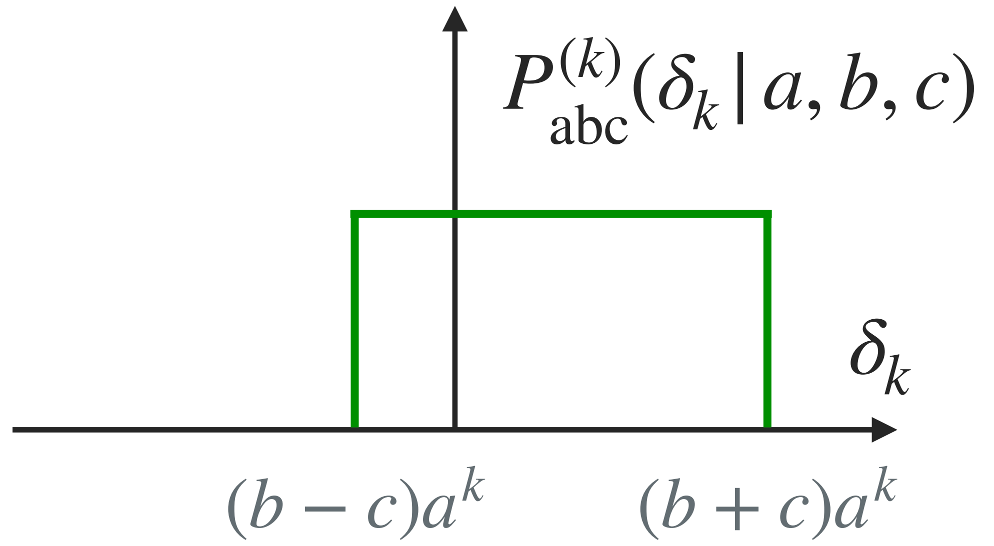
[Duhr, AH, Mazeliauskas, Szafron '21]

- geometric model is symmetric: $P(\delta_0, \ldots, \delta_n) = P(|\delta_0|, \ldots, |\delta_n|) \rightsquigarrow \langle \delta_{n+1} \rangle_{\text{geo}} = 0$
- allow for different lower & upper bound:

$$b - c \le \frac{\delta_k}{a^k} \le b + c \quad \forall k \quad \leftrightsquigarrow \quad \text{three model parameters: } a,\ b,\ c$$

bias/offset (→ b)

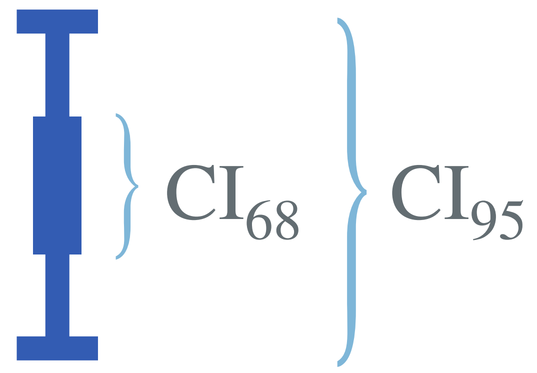
- **model:** $$P^{(k)}_{abc}(\delta_k | a, b, c) = \frac{1}{2c|a|^k}\, \Theta\left(c - \left|\frac{\delta_k}{a^k} - b\right|\right)$$



- **priors:** $$P_0(a,b,c) = P_0(a)\, P_0(b,c)$$

$$P_0(a) = \frac{1}{2}\,(1+\omega)\,(1-|a|)^\omega\, \Theta(1-|a|) \quad \leftrightsquigarrow \quad \text{support: } [-1,+1] \ (\text{alternating ✔})$$

$$P_0(b,c) = \frac{\varepsilon\eta^\varepsilon}{c^{1+\varepsilon}}\, \Theta(c-\eta)\, \frac{1}{2\xi c}\, \Theta(\xi c - b)$$

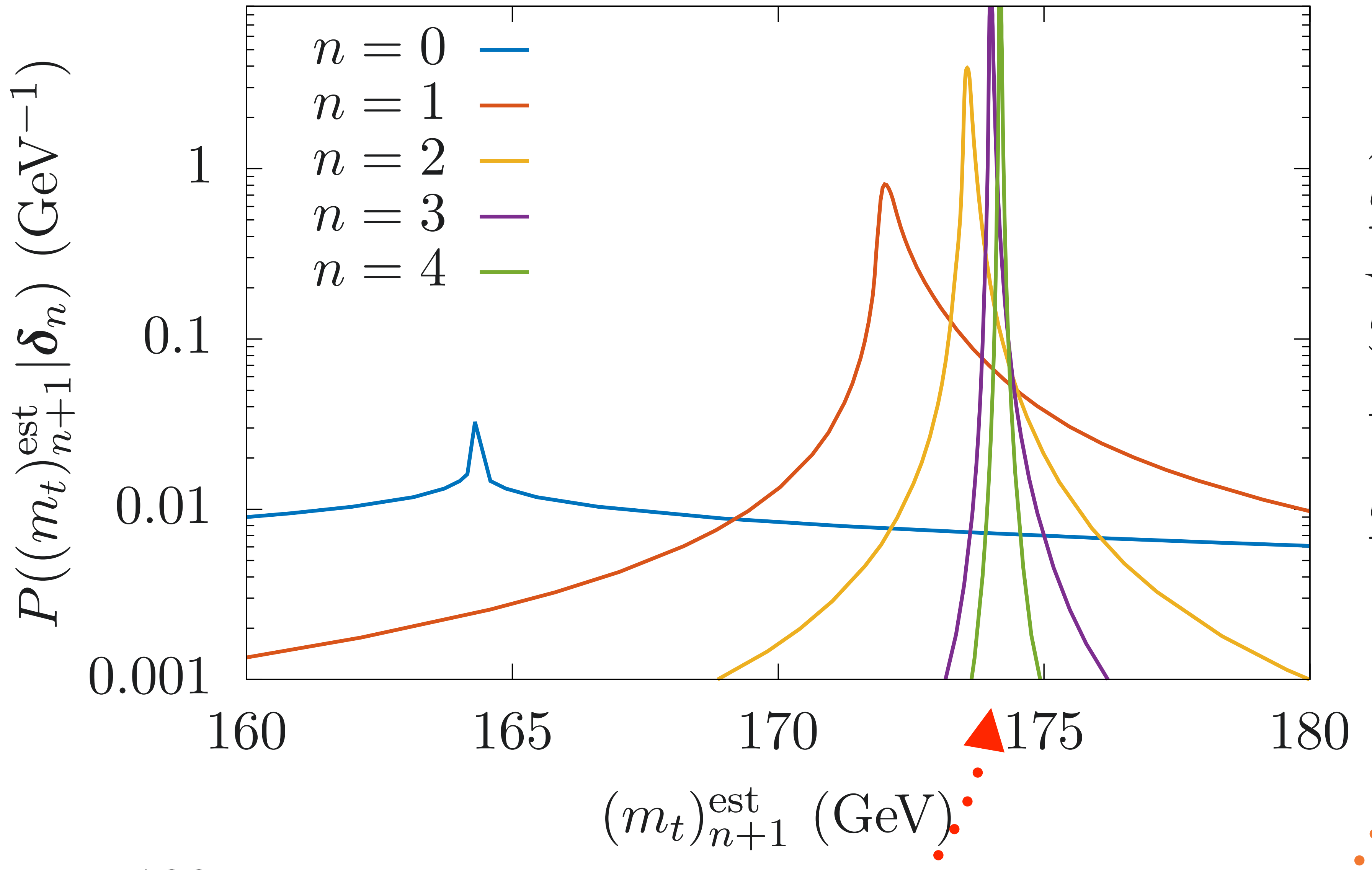


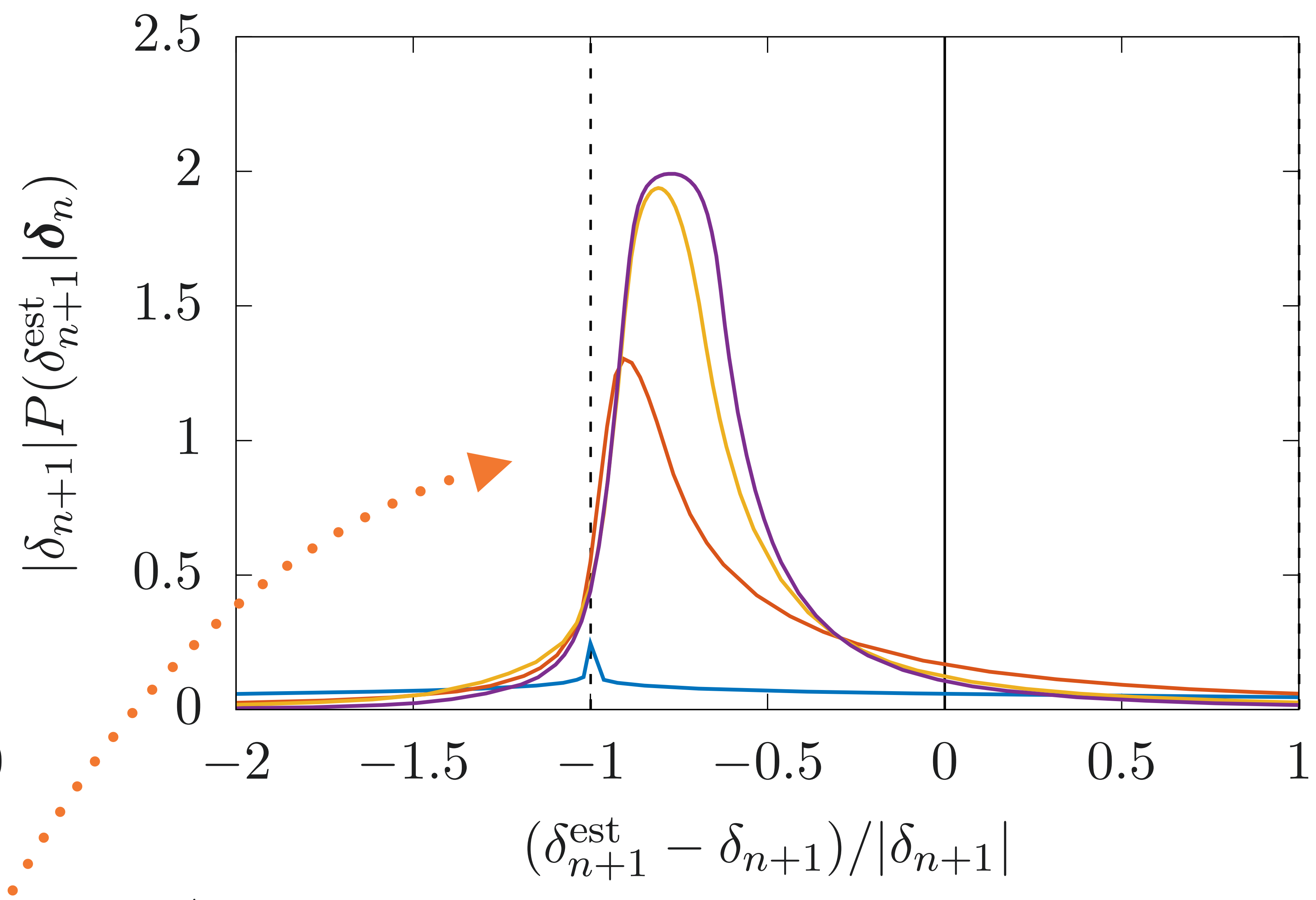
A Real-World Example — m_t (OS $\leftrightsquigarrow$ $\overline{\text{MS}}$)

$$m_t = \frac{Z_m^{\text{OS}}}{Z_m^{\overline{\text{MS}}}(\mu_R)}\,\overline{m}_t(\mu_R)$$

$$= \sum_k m_t^{(k)}$$

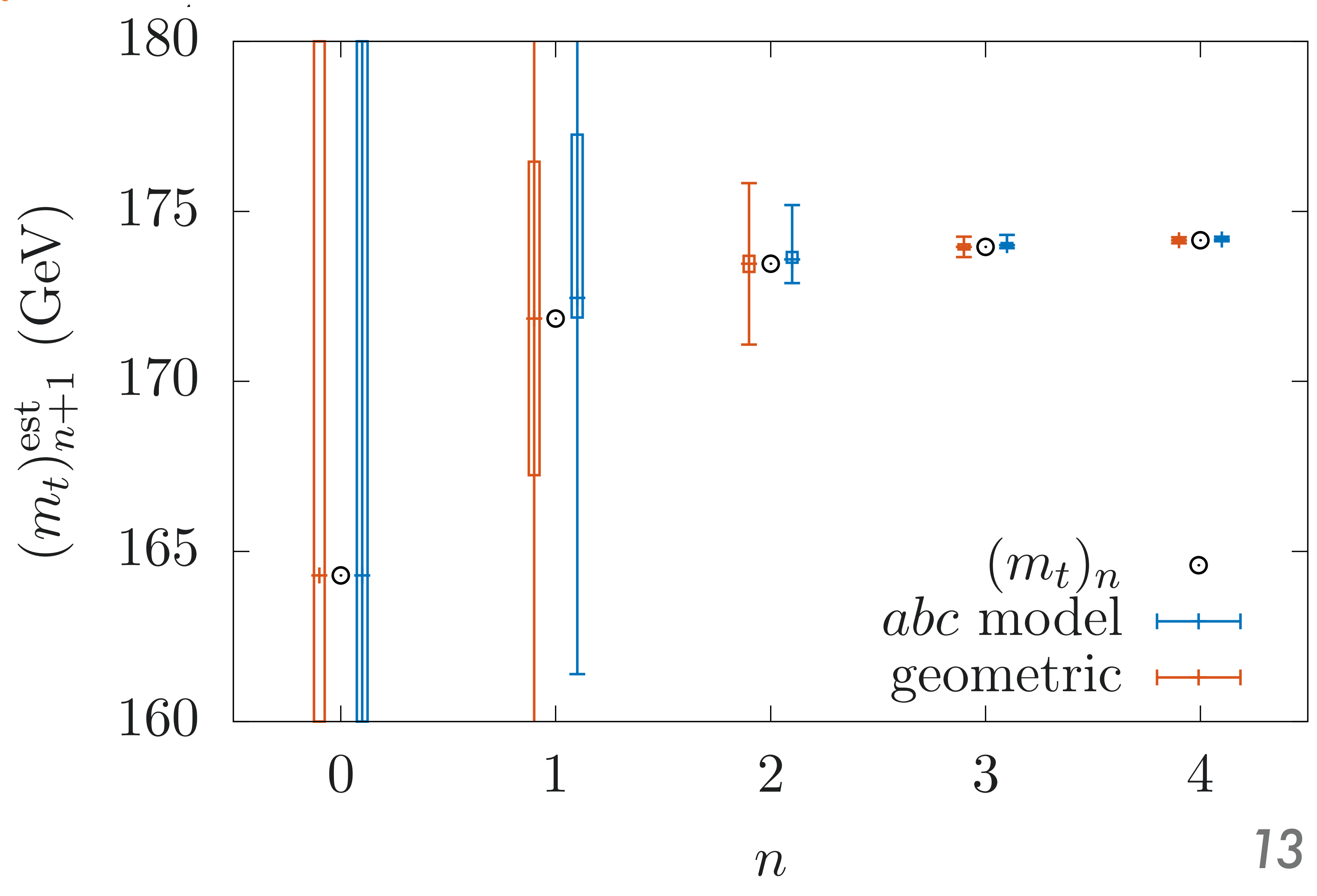
- estimate for $m_t - (m_t)_4$
  - $CI_{68} = [0.008, 0.046]$ GeV
  - $CI_{95} = [-0.027, 0.112]$ GeV





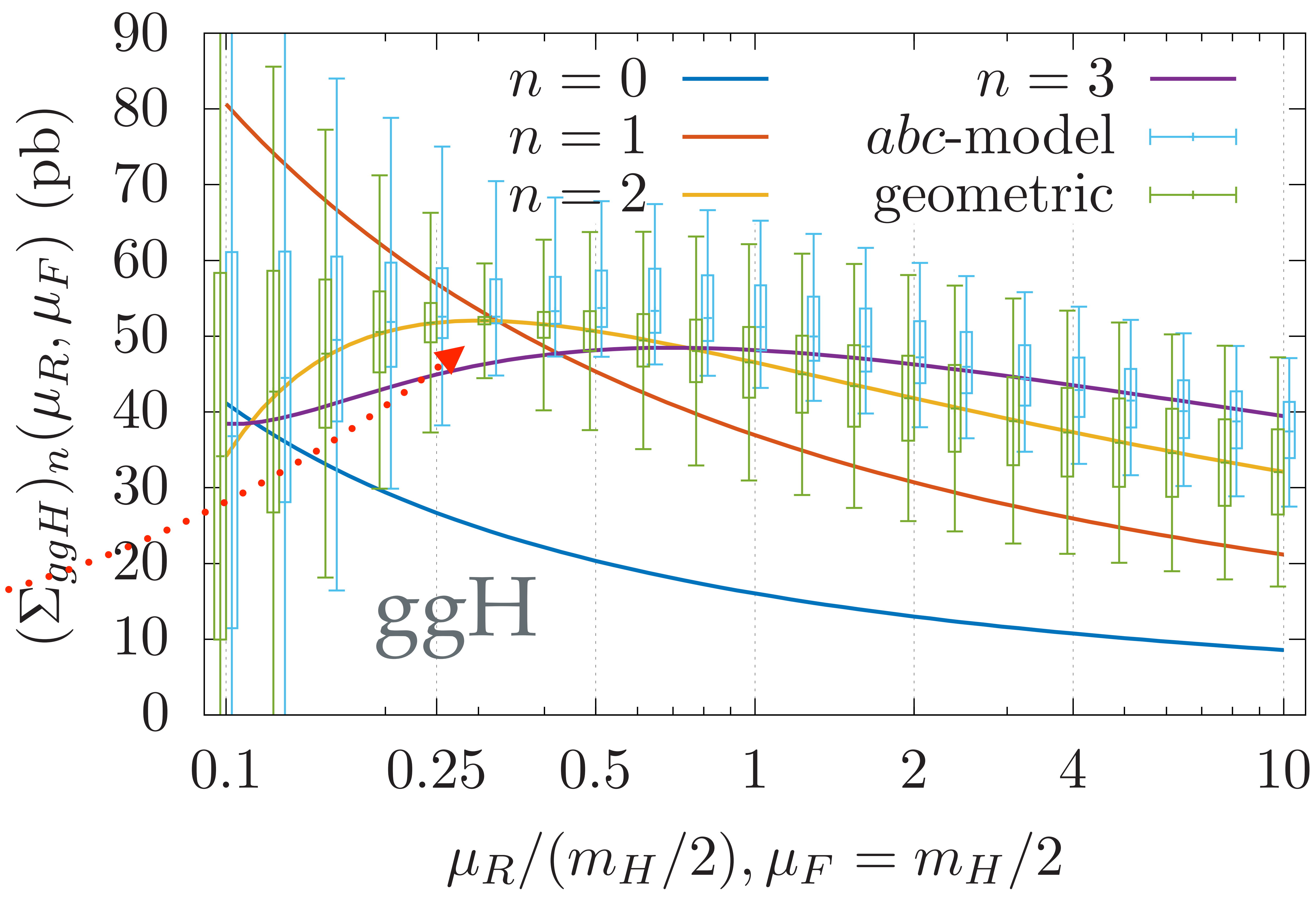
strongly peaked $n \nearrow$

positive corrections anticipated



What to do with the the scale μ?

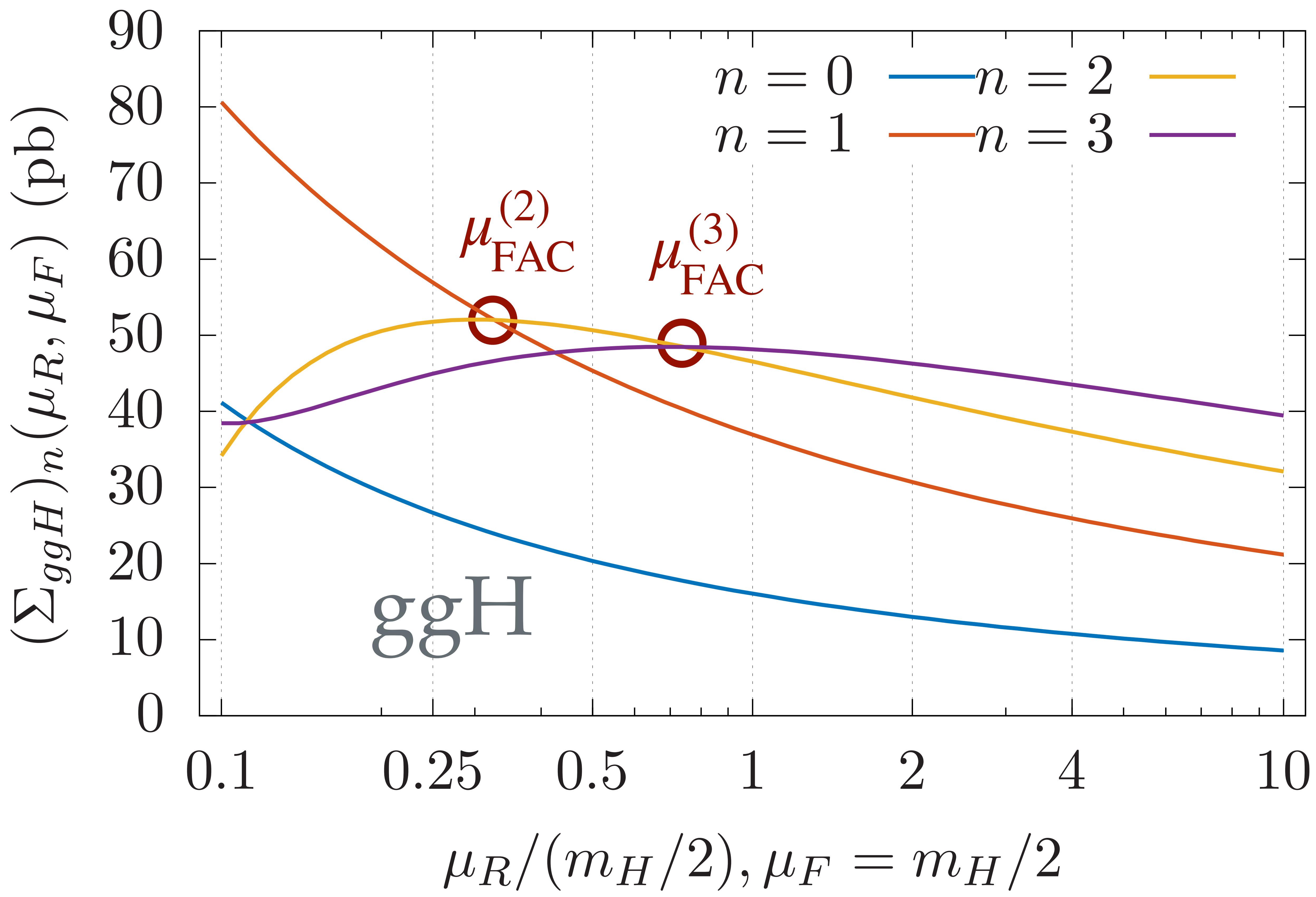
- $\forall \mu \rightsquigarrow P(\delta_3 \,|\, \delta_0, \delta_1, \delta_2; \mu)$
  - $CI_{68/95}$ (geo) (abc)
- geo
  - always entered around NNLO
  - very narrow peak
- *abc*
  - $\mu/\mu_0 \gtrsim 1 \rightsquigarrow$ anticipate pos. N3LO
  - $\mu/\mu_0 \lesssim 1 \rightsquigarrow$ bias slowly disappears



WHAT TO DO WITH THE THE SCALE μ?

- $\forall\mu \rightsquigarrow P(\delta_3 \,|\, \delta_0, \delta_1, \delta_2; \mu)$
  - $CI_{68/95}$ (geo) (abc)

- two options:
  1. invoke some *principle* to pick the *"optimal"* scale
     - FAC, PMS, PMC, …

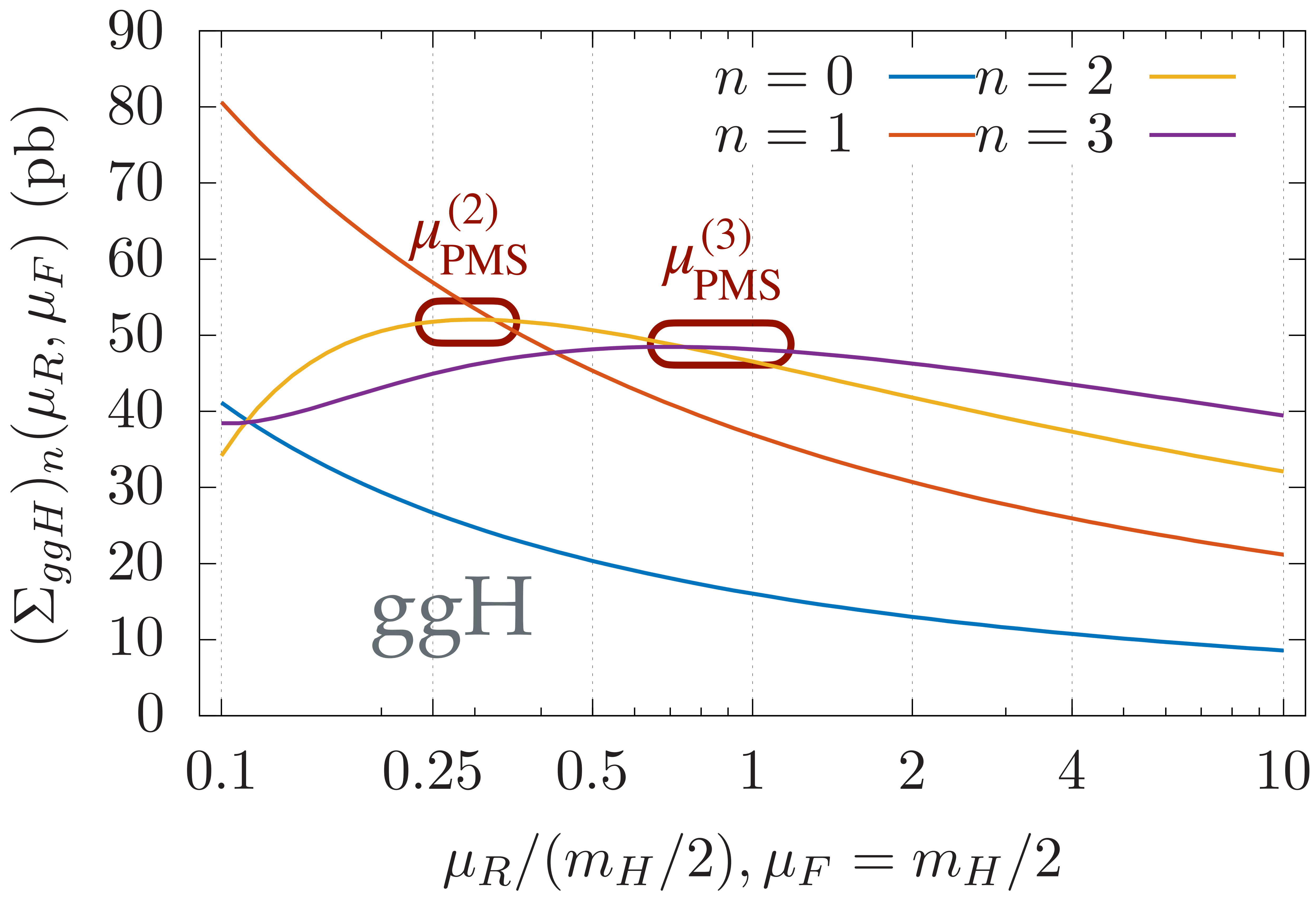


Fastest Apparent Convergence
$\Sigma_n(\mu_{\rm FAC}) = \Sigma_{n-1}(\mu_{\rm FAC})$

depends on order
might not be unique

What to do with the the scale μ?

- $\forall \mu \rightsquigarrow P(\delta_3 \,|\, \delta_0, \delta_1, \delta_2; \mu)$
  - $\mathrm{CI}_{68/95}$ (geo) (abc)
- two options:
  1. invoke some *principle* to pick the *"optimal"* scale
     - FAC, PMS, PMC, …



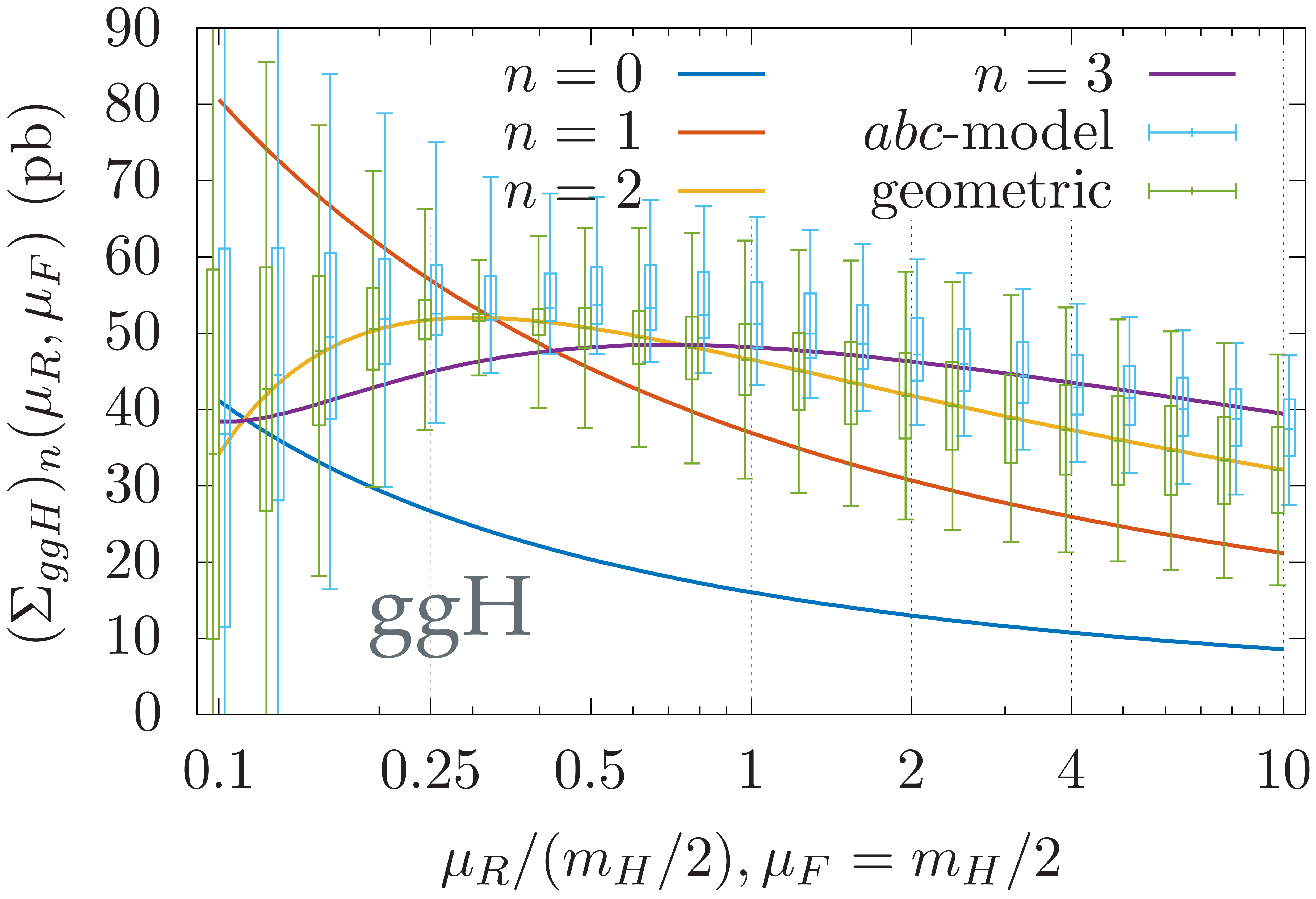
Principle of Minimal Sensitivity

$$\frac{\partial}{\partial \mu} \Sigma_n(\mu) \big|_{\mu_{\mathrm{PMS}}} = 0$$

depends on order
might not be unique

What to do with the the scale μ?

- $\forall\mu \rightsquigarrow P(\delta_3 \,|\, \delta_0, \delta_1, \delta_2; \mu)$
  - $\mathrm{CI}_{68/95}$ (geo) (abc)



- two options:
  1. invoke some *principle* to pick the *"optimal"* scale
     - FAC, PMS, PMC, …
  2. combine different $P(\delta_{n+1} \,|\, \boldsymbol{\delta}_n; \mu)$

pursued in the following

Prescriptions for Scales

Scale Marginalisation (sm):

[Bonvini '20]

- treat μ as a hidden model parameter
  & *marginalise* over it:

$$P_{\rm sm}(\delta_{n+1}\,|\,\boldsymbol{\delta}_n) = \int \mathrm{d}\mu \; P(\delta_{n+1}, \mu\,|\,\boldsymbol{\delta}_n)$$
$$= \int \mathrm{d}\mu \; P(\delta_{n+1}\,|\,\boldsymbol{\delta}_n; \mu)\; P(\mu\,|\,\boldsymbol{\delta}_n)$$

- $P(\mu\,|\,\boldsymbol{\delta}_n) \propto P(\boldsymbol{\delta}_n;\mu)\; P_0(\mu)$ with prior:

$$P_0(\mu) = \frac{1}{2\mu \ln F}\,\Theta\Big(\ln F - \Big|\ln\big(\tfrac{\mu}{\mu_0}\big)\Big|\Big)$$

Scale Average (sa):

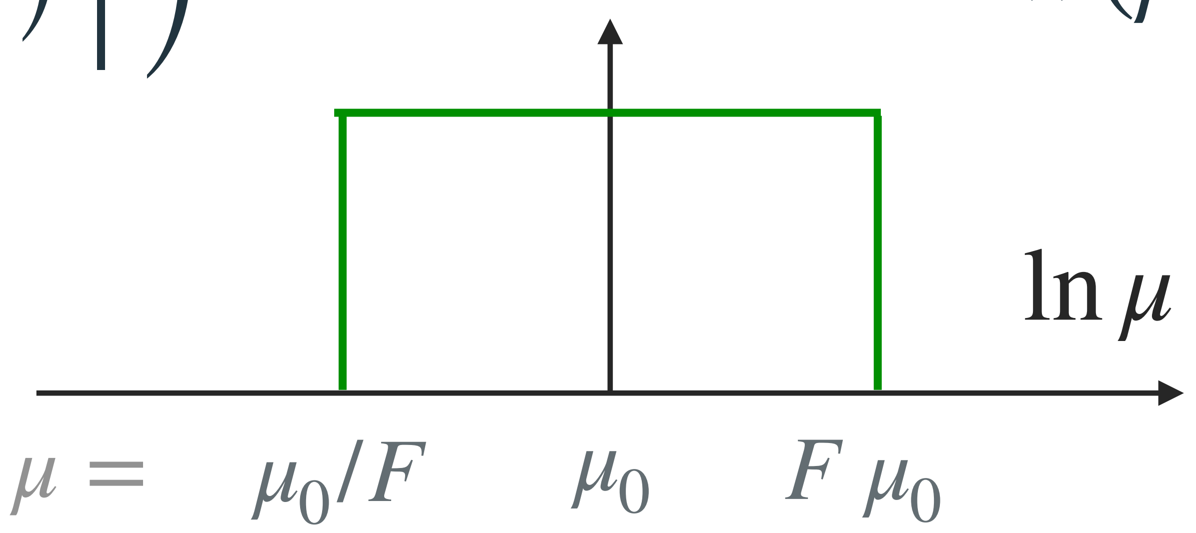
[Duhr, AH, Mazeliauskas, Szafron '21]

- μ has no probabilistic interpretation
  ⇝ *average* over it:

$$P_{\rm sa}(\delta_{n+1}\,|\,\boldsymbol{\delta}_n) = \int \mathrm{d}\mu \; w(\mu)\; P(\delta_{n+1}\,|\,\boldsymbol{\delta}_n; \mu)$$

- weight function:

$$w(\mu) = \frac{1}{2\mu \ln F}\,\Theta\Big(\ln F - \Big|\ln\big(\tfrac{\mu}{\mu_0}\big)\Big|\Big)$$



Peak of the Distributions*

[Duhr, AH, Mazeliauskas, Szafron '21]

Scale Marginalisation (sm):

- if $\mu_{\rm FAC} \in [\mu_0/F, F\,\mu_0]$ then $P_{\rm sm}(\delta_{n+1}\,|\,\boldsymbol{\delta}_n)$ peaks at $\Sigma_n(\mu_{\rm FAC})$
  - $P(\boldsymbol{\delta}_n\,|\,\mu)$ dominated by $(k=n)$ term
  - symmetric model
    ⇝ $\delta_n(\mu) = 0$ enhanced

Scale Average (sa):

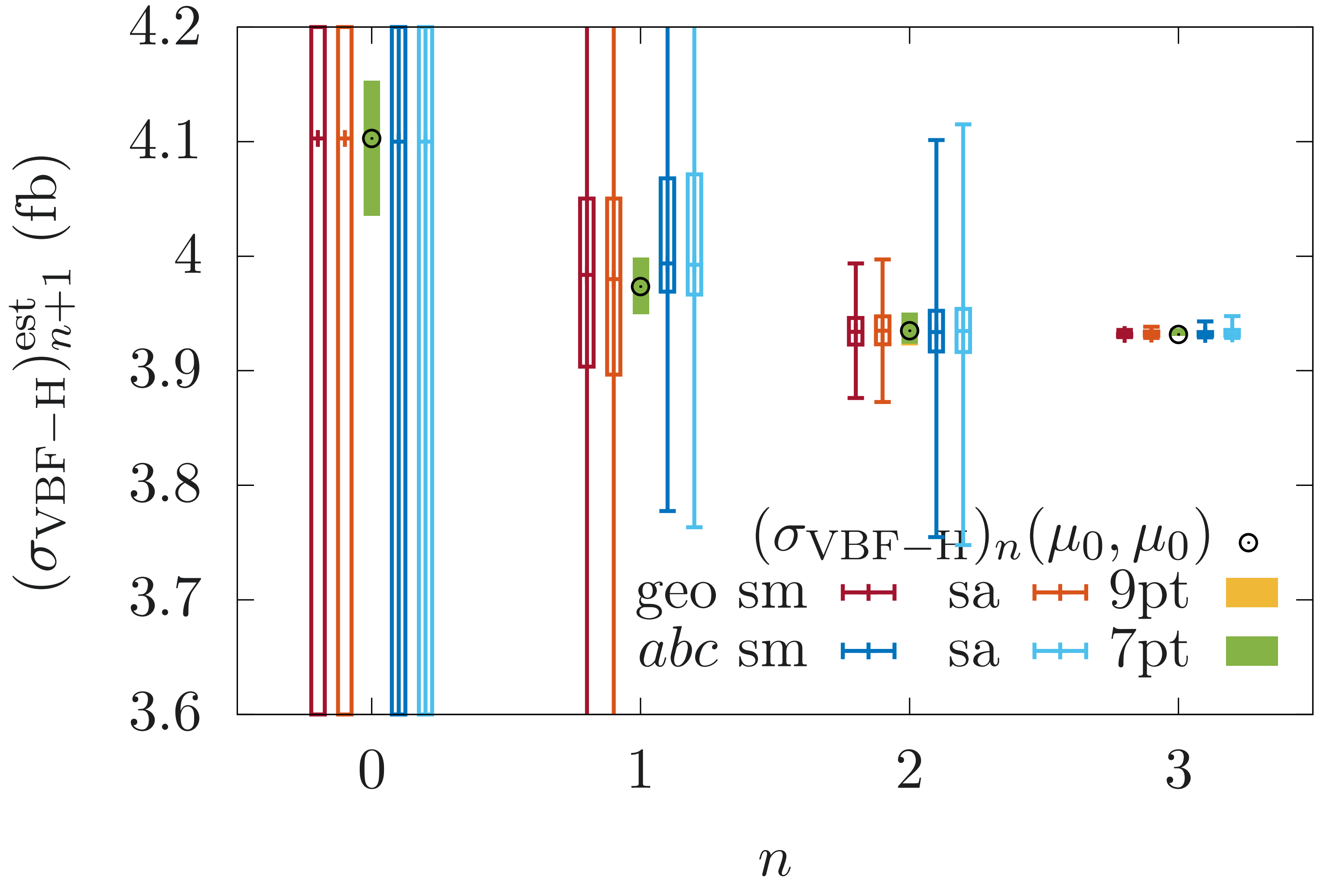
- if $\mu_{\rm PMS} \in [\mu_0/F, F\,\mu_0]$ then $P_{\rm sa}(\delta_{n+1}\,|\,\boldsymbol{\delta}_n)$ peaks at $\Sigma_n(\mu_{\rm PMS})$
  - overlap between $P(\delta_{n+1}\,|\,\boldsymbol{\delta}_n;\mu)$ enhanced at stationary point
    ⇝ $\Sigma_n'(\mu_{\rm PMS}) \approx 0$

Choice of how to interpret the scale has consequences for predictions!

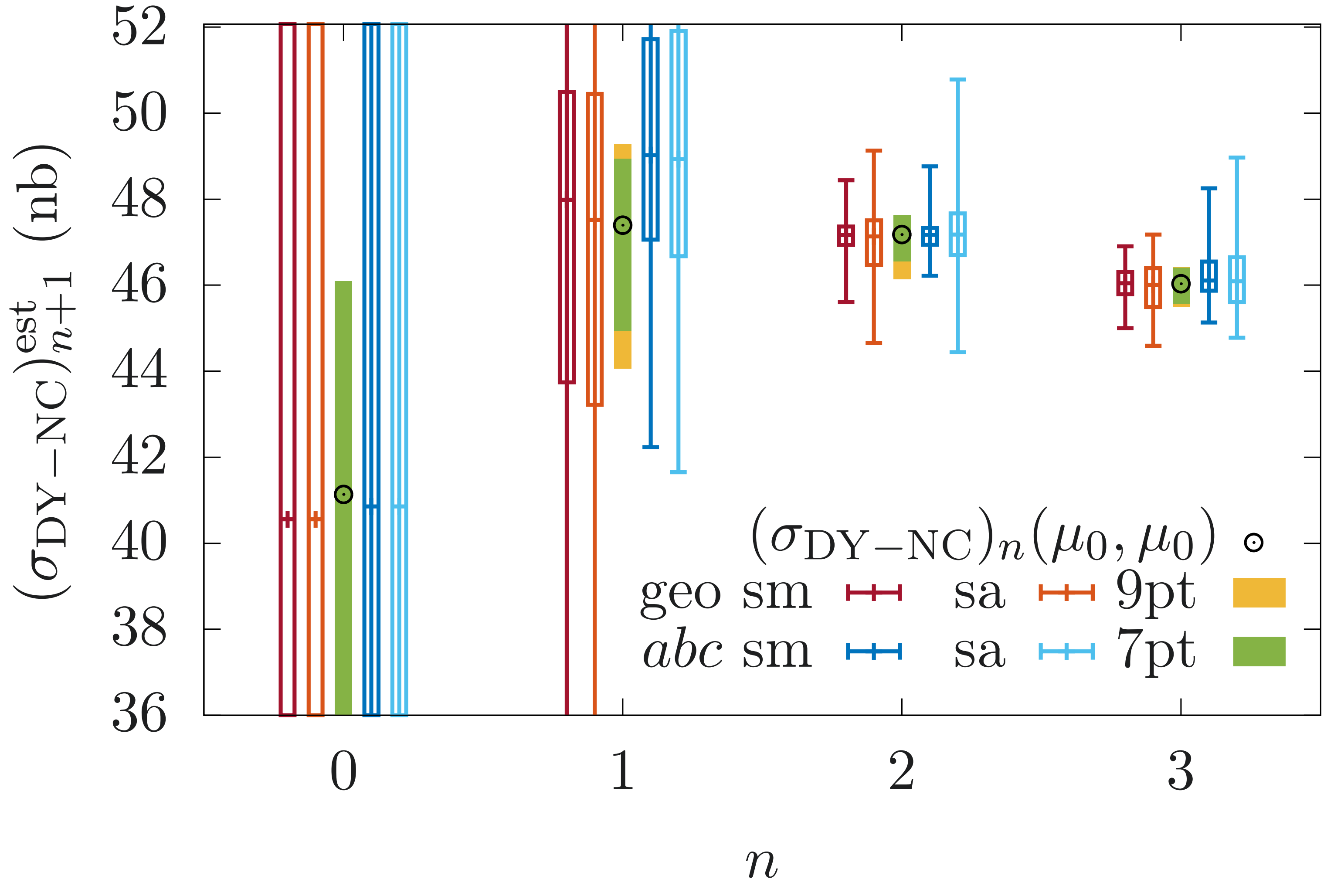
* for symmetric models, a convergent series, and reasonable assumptions

Inclusive Cross Sections up to N^{3}LO

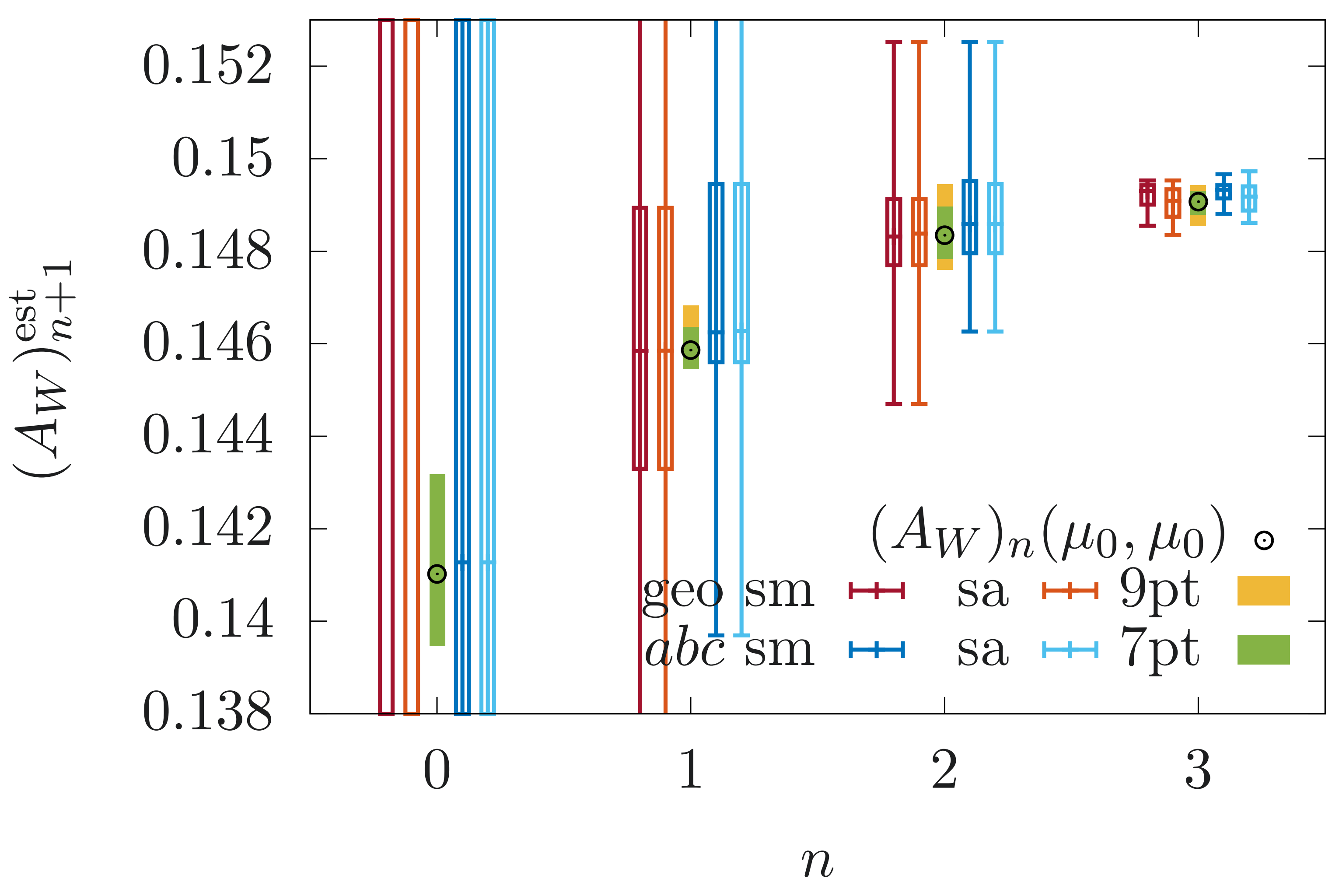
VBF–H



DY–NC



$A_W = \dfrac{W^+ - W^-}{W^+ + W^-}$



- $n < 2$: CI_{68} bigger than 9pt
- $\delta_1 < 0 \rightsquigarrow$ *abc* alternating
- $n > 2$: all prescriptions similar

- δ_3 is large and outside of 9pt!
- similar unc.: sa $\simeq$ 9pt
- $n = 2$: sm $\ll$ others (μ_{FAC})
- $n = 3$: all prescriptions similar

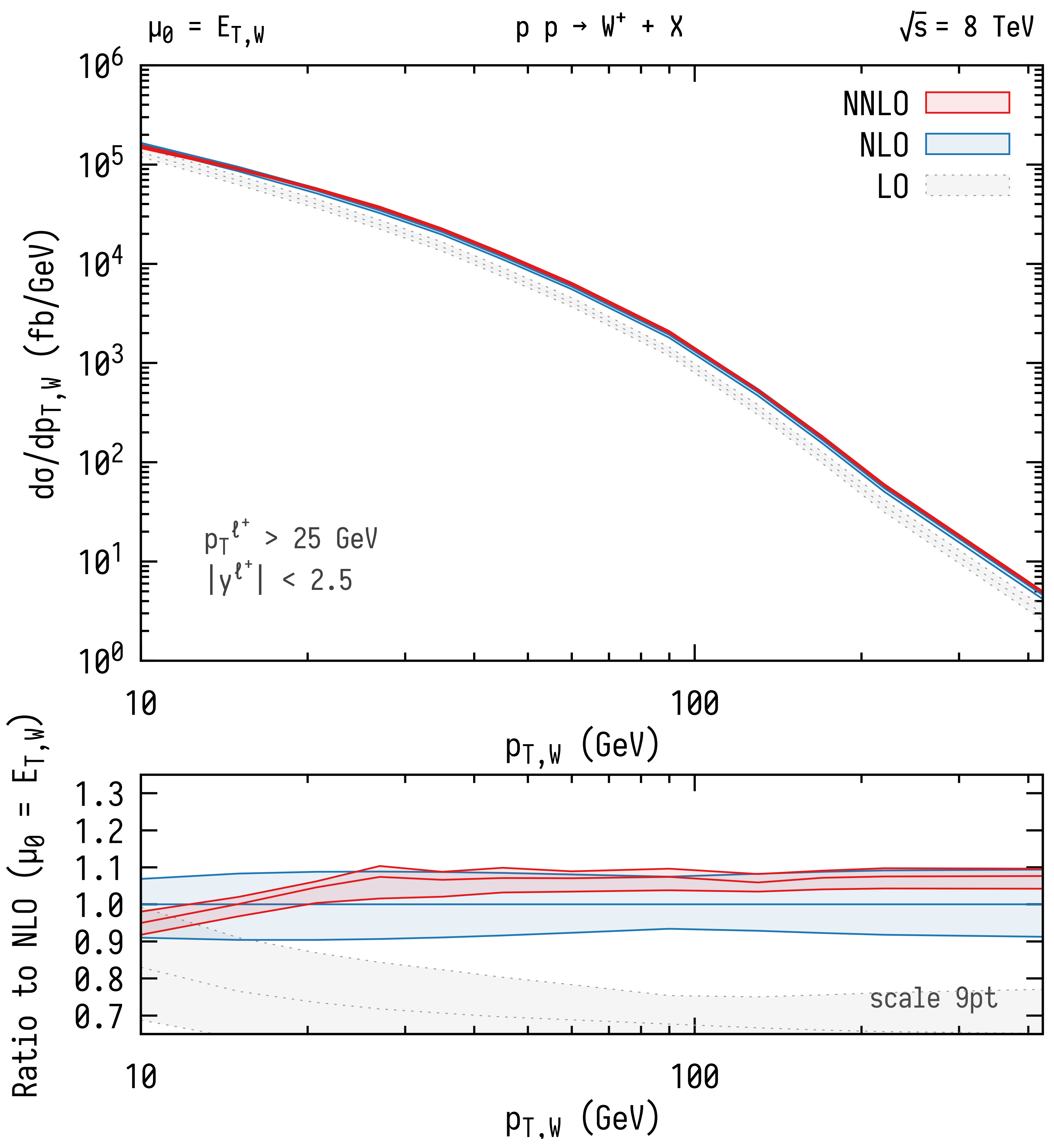
- large cancellations in the ratio
- $n < 2$: 9pt performs poorly
- $(A_W)_n \nearrow$ (anticipated by *abc*)
- size: *abc* $\lesssim$ others

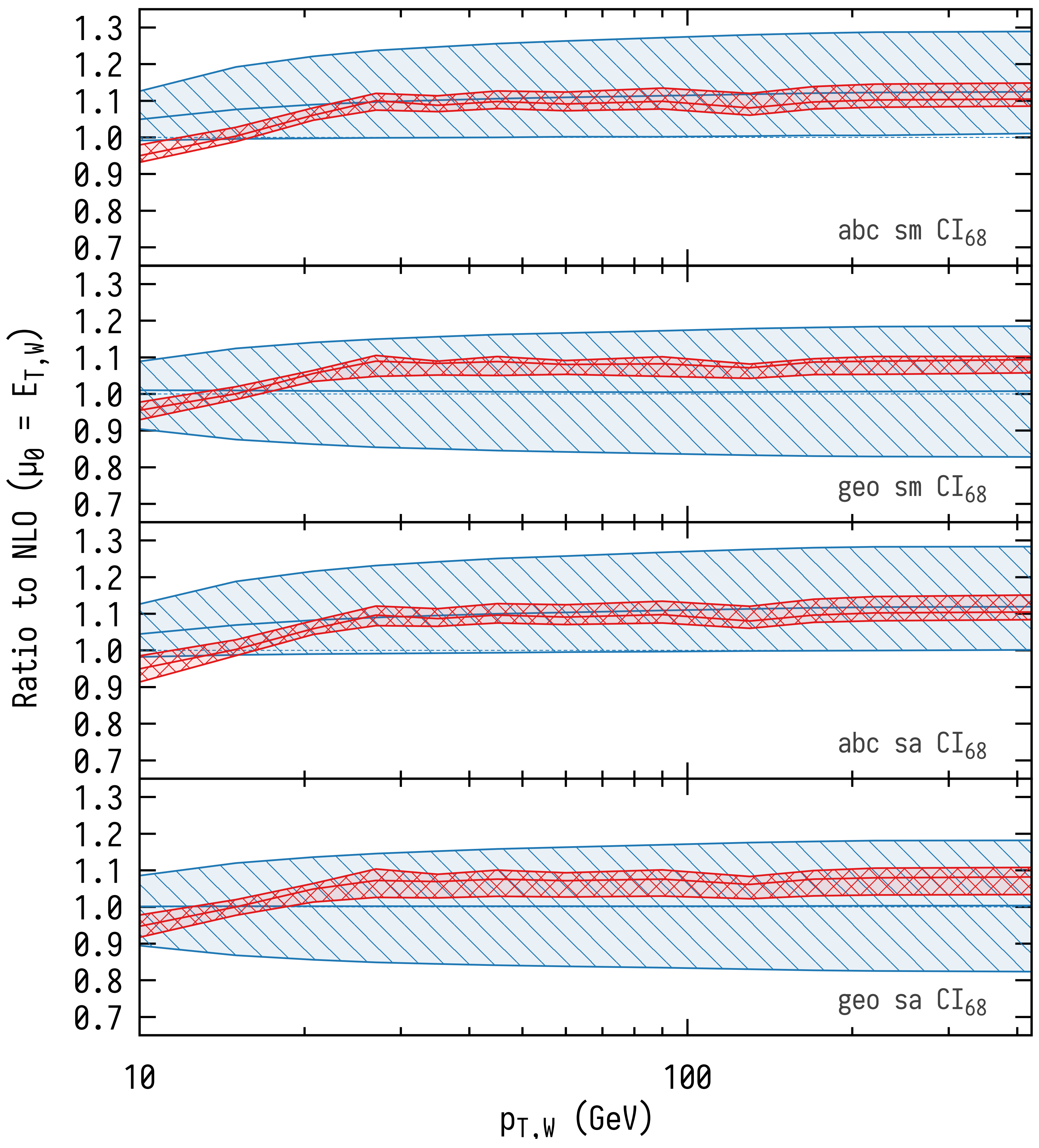
overall: not radically different estimates for Δ_{MHO} ($n \geq 2$)

DIFFERENTIAL DISTRIBUTIONS

- Bayesian approach also applicable to distributions
  - ⇝ treat each bin individually ⇜⇝ will not include correlations!
- new challenges
  - no longer "easy" to identify an appropriate hard scale μ_0 (up to rescaling)
    - ⇝ inclusive ggH: M_H vs. $\frac{1}{2}\,M_\mathrm{H}$? Just let the model figure it out.
  - differential distributions can probe different kinematic regimes
    - ⇝ dynamical scale choice ⇜⇝ *many choices!*
    - ⇝ e.g. in jet production: p_T^{j} , $p_\mathrm{T}^{j_1}$, $\langle p_\mathrm{T}^{j}\rangle_\mathrm{avg}$, $H_\mathrm{T} \equiv \sum_{i\in\mathrm{jets}} p_\mathrm{T}^{i}$, $\hat{H}_\mathrm{T} \equiv \sum_{i\in\mathrm{partons}} p_\mathrm{T}^{i}$, …
  - re-cycling via quadrature limited ⇝ ideally interpolation grids

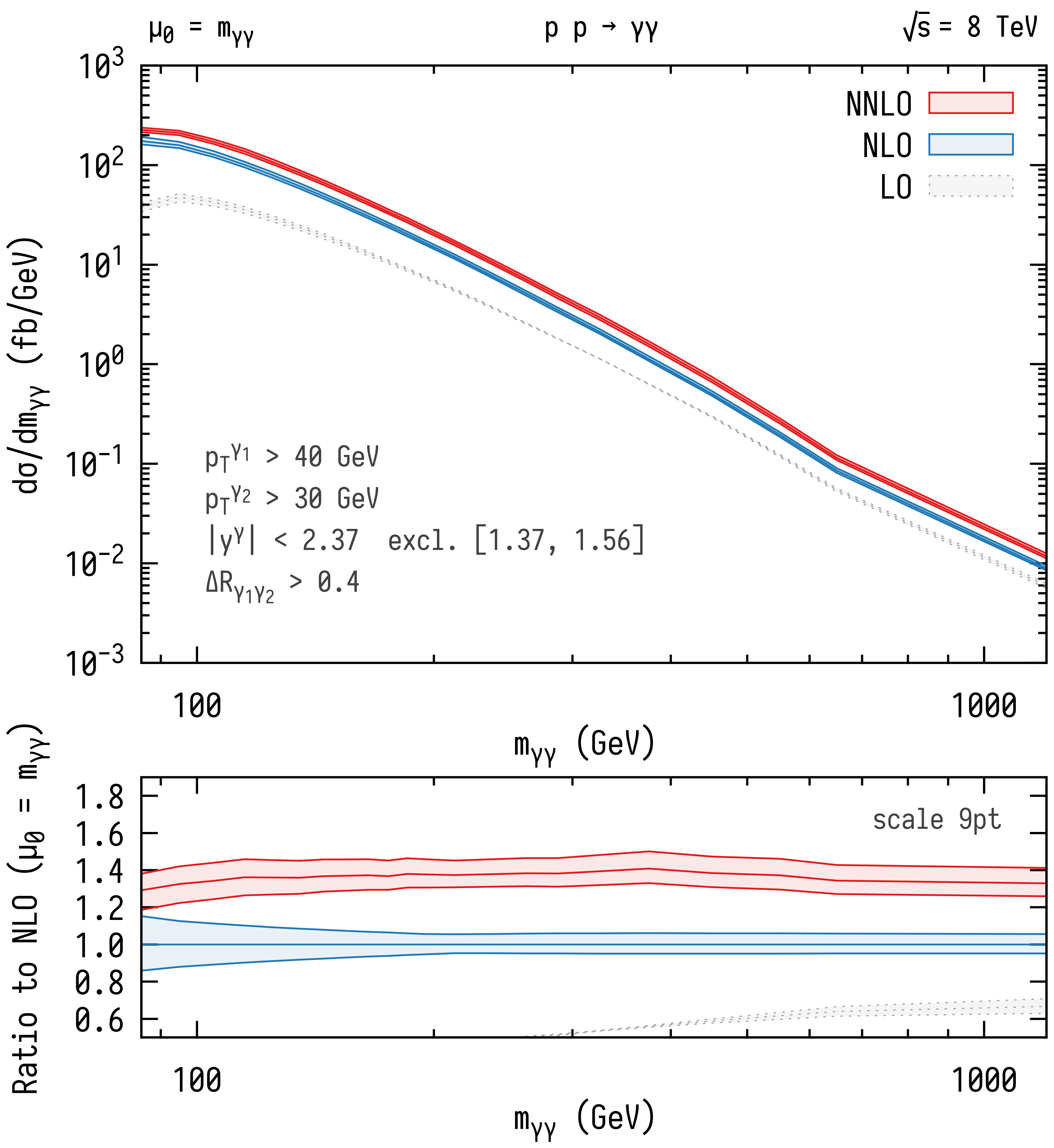
W-Boson + Jet Production

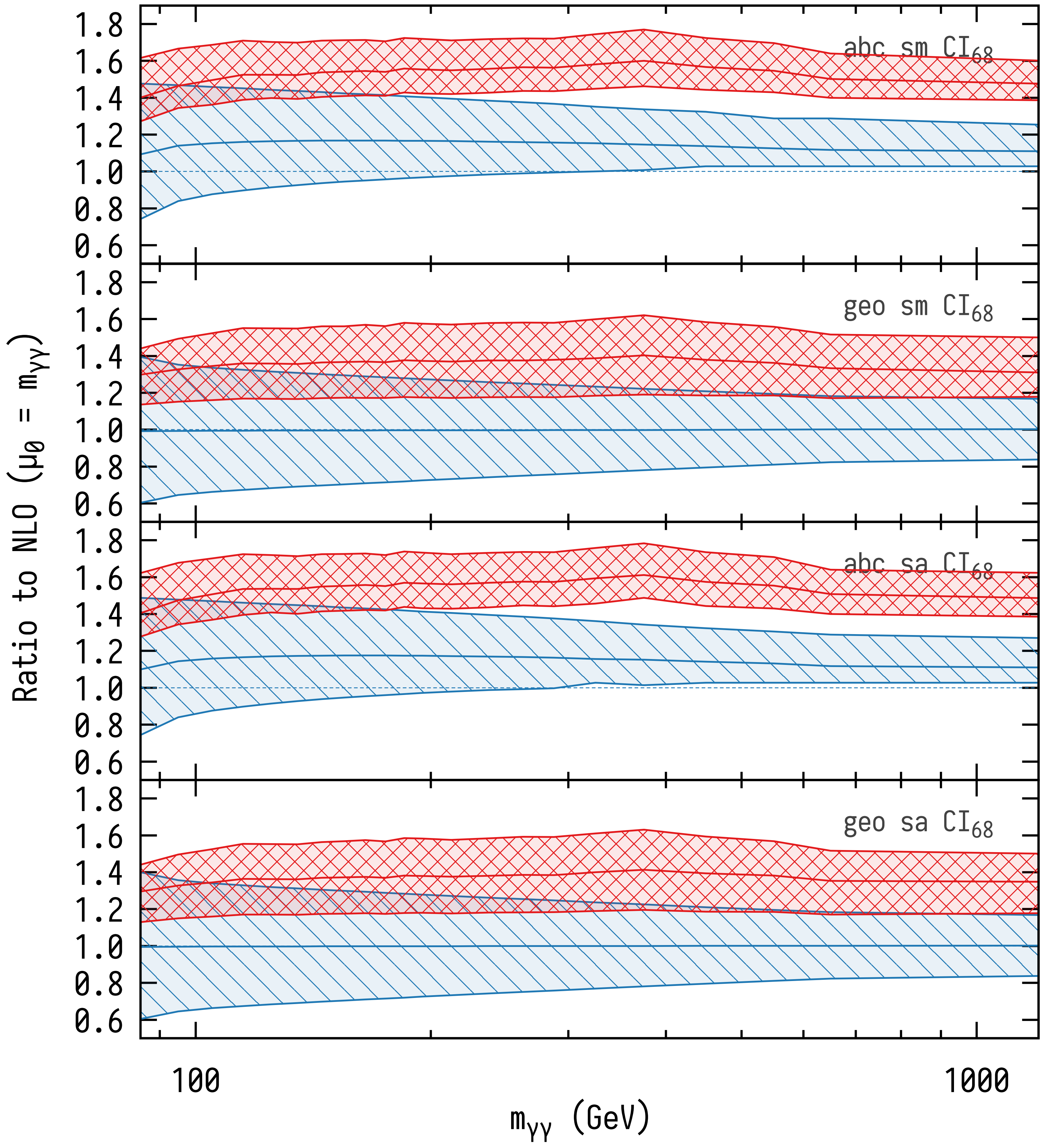




- $n < 2$:
  - CI_{68} bigger than 9pt
  - *abc* captures pos. shift
- $n = 2$:
  - almost identical bands
  - Δ_{MHO} very robust
- sm vs. sa
  - almost identical CI

Di-Photon Production

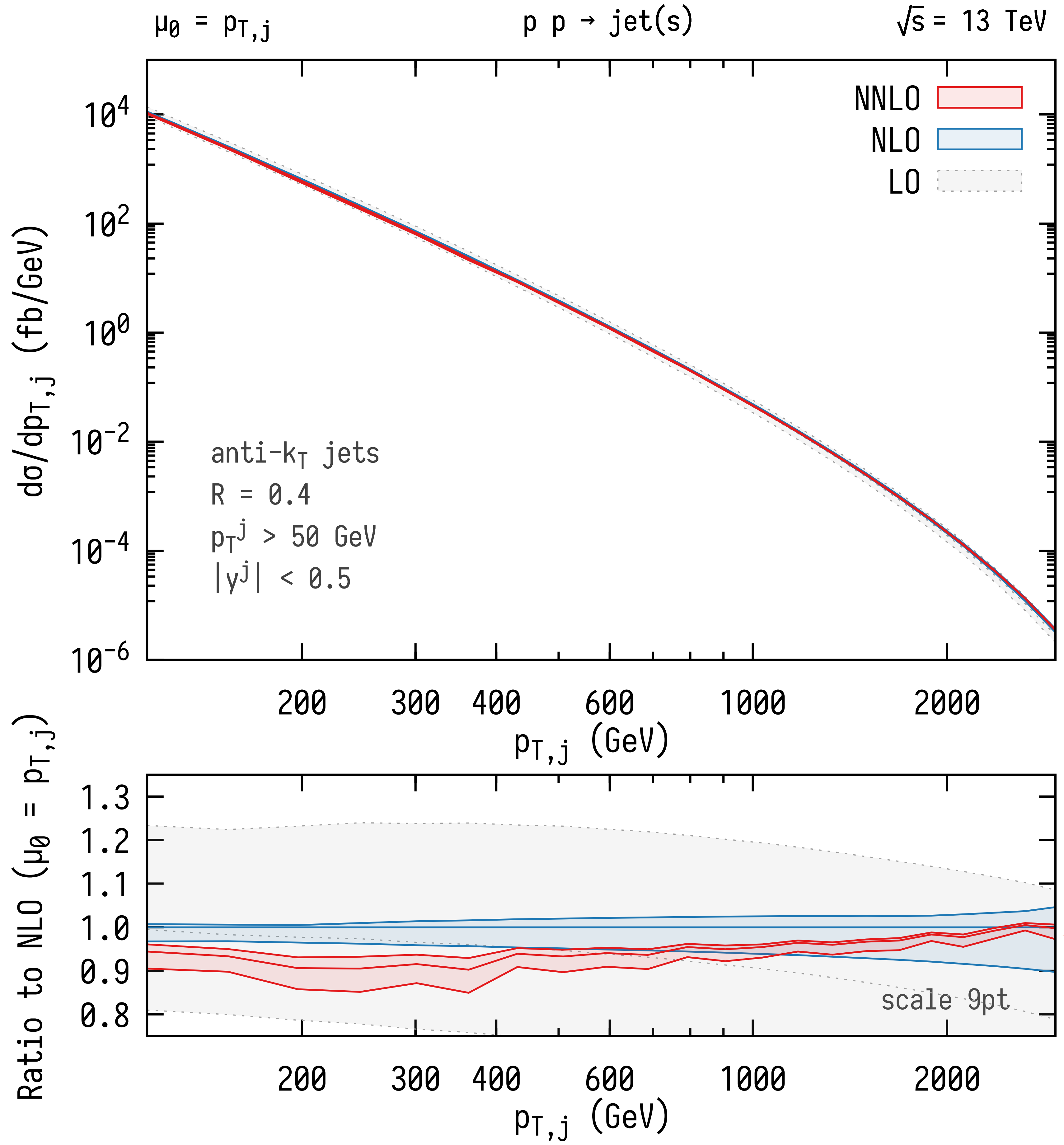




- example where 9pt fails
  - large corrections
  - $\Delta_{\text{MHO}}^{\text{NNLO}} \gtrsim \Delta_{\text{MHO}}^{\text{NLO}}$
  - no sign of convergence
- $n < 2$:
  - $\text{CI}_{68} \sim$ 2-3 $\times$ 9pt
- $n = 2$:
  - marginal overlap for geo
  - differences in *size* & *position*
  - ideally N3LO for robust Δ_{MHO}
- sm $\simeq$ sa
  - large corrections prohibit FAC points

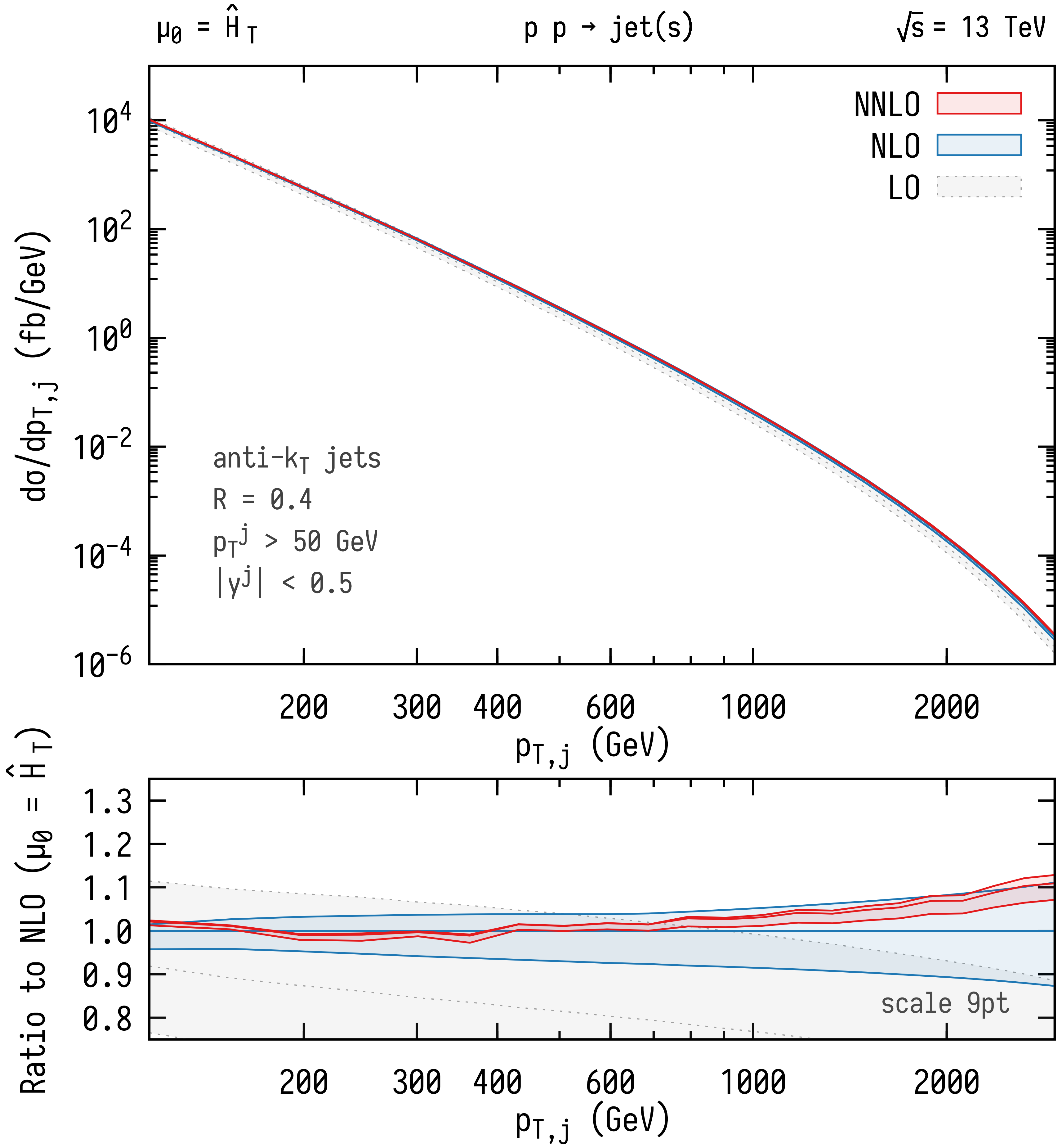
The Problem with Jets...

$\mu_0 = p_T^j$: infrared sensitivity



instability at low p_T

$\mu_0 = \hat{H}_T$: recommendation [Currie et al. '18]

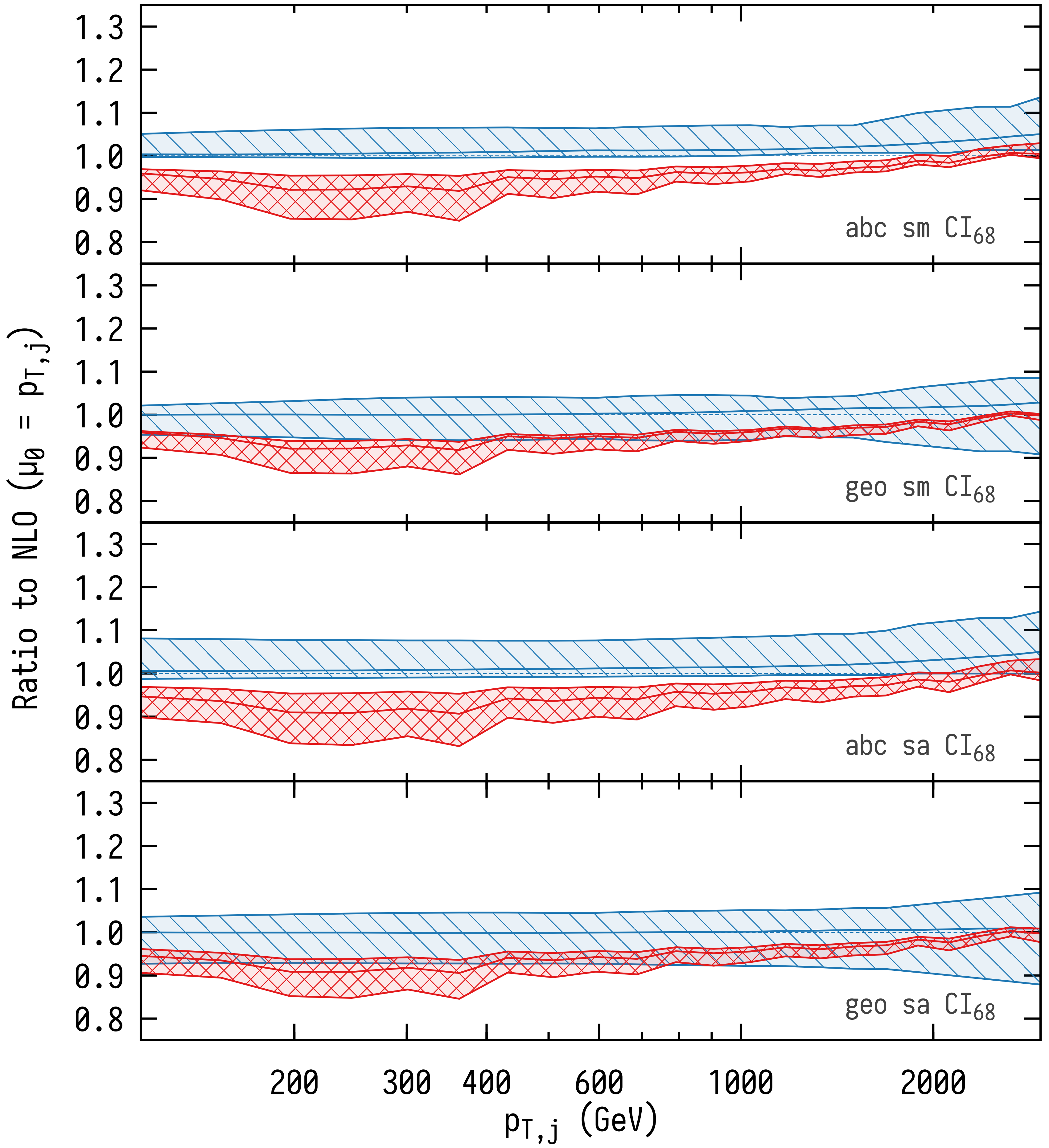


The Problem with Jets… Persists

non-trivial change of dynamical scales cannot be captured by a simple re-scaling

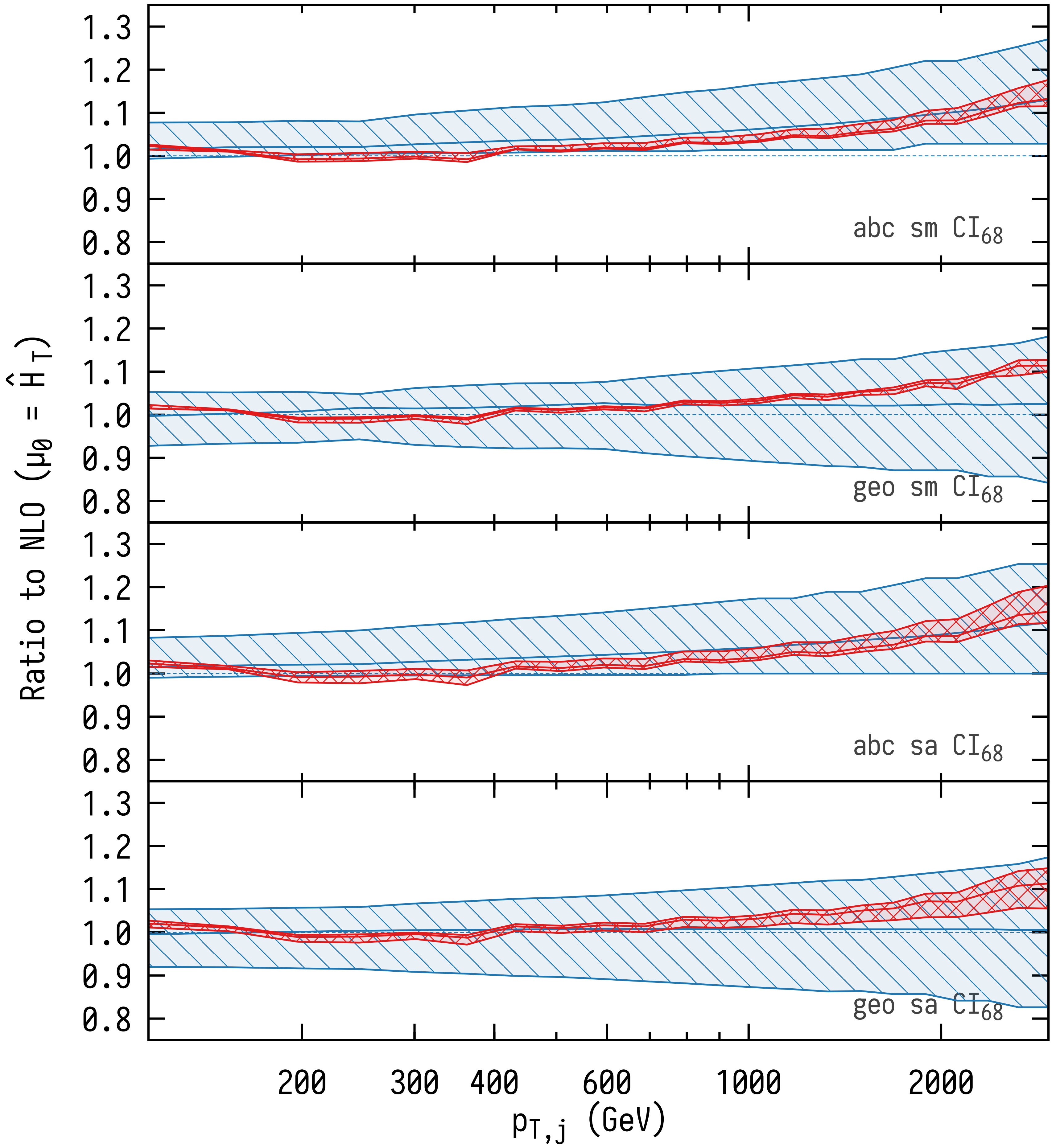
$\mu_0 = p_\mathrm{T}^j$: infrared sensitivity

$\mu_0 = \hat{H}_\mathrm{T}$: recommendation [Currie et al. '18]



larger NLO Δ_MHO

barely any difference at NNLO



abc captures positive corrections

Work in Progress — Correlations

[AH, Mazeliauskas w.i.p]

- idea: if two bins show similar (opposite) perturbative behaviour
  ↪ two bins should be partially (anti-)correlated.

- we want: joint probability distribution $P(x, y)$ for two bins x & y
  ↪ preserve projections for compatibility:

$$P(x) = \int \mathrm{d}y\ P(x, y) = \int \mathrm{d}z\ P(x, z)$$

  ↪ hidden parameter $-1 < c < +1$ to smoothly implement the correlation

- possibilities: algorithmic "earth movers distance"; map $P(x)$ onto $P(y)$, …
  ↪ can be done much simpler

Work in Progress – Correlation Model in miho

[AH, Mazeliauskas w.i.p]

- projections of multi-dim. Gaussians (+ correlation matrix) are again Gaussian
  ↪ map P_i onto Gaussians, implement correlations, map back

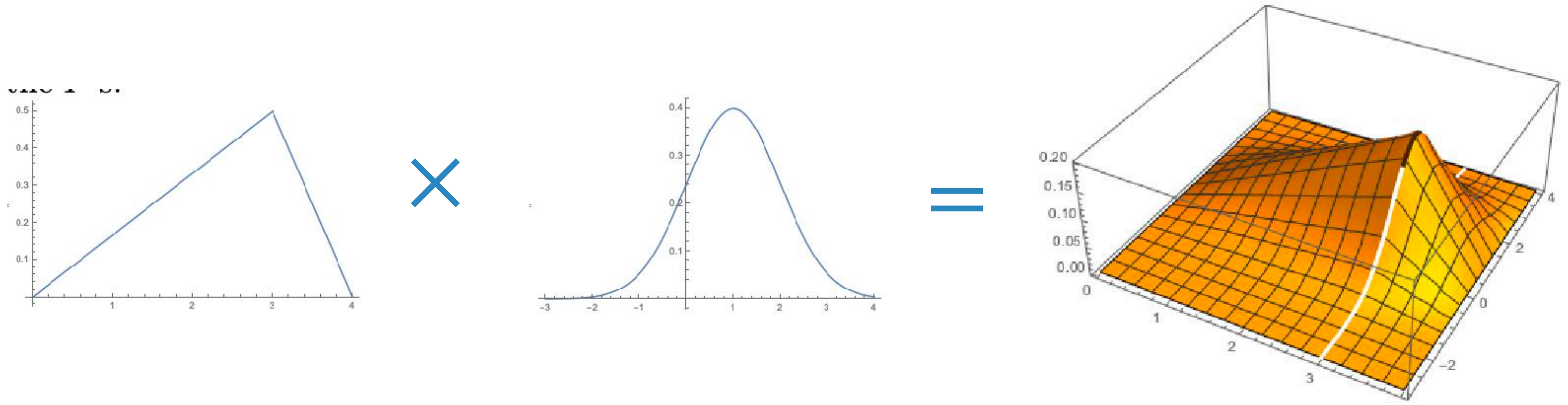
$$P(x,y) = P_1(x)P_2(y) \times \left.\frac{d\Phi^{-1}(\alpha)}{d\alpha}\right|_{\alpha=\Sigma_1(x)} \left.\frac{d\Phi^{-1}(\beta)}{d\beta}\right|_{\beta=\Sigma_2(y)} \times \frac{1}{2\pi\sqrt{1-c^2}} \exp\left(-\frac{1}{2(1-c^2)}\left[\xi(x)^2+\eta(y)^2-c2\xi(x)\eta(y)\right]\right)$$

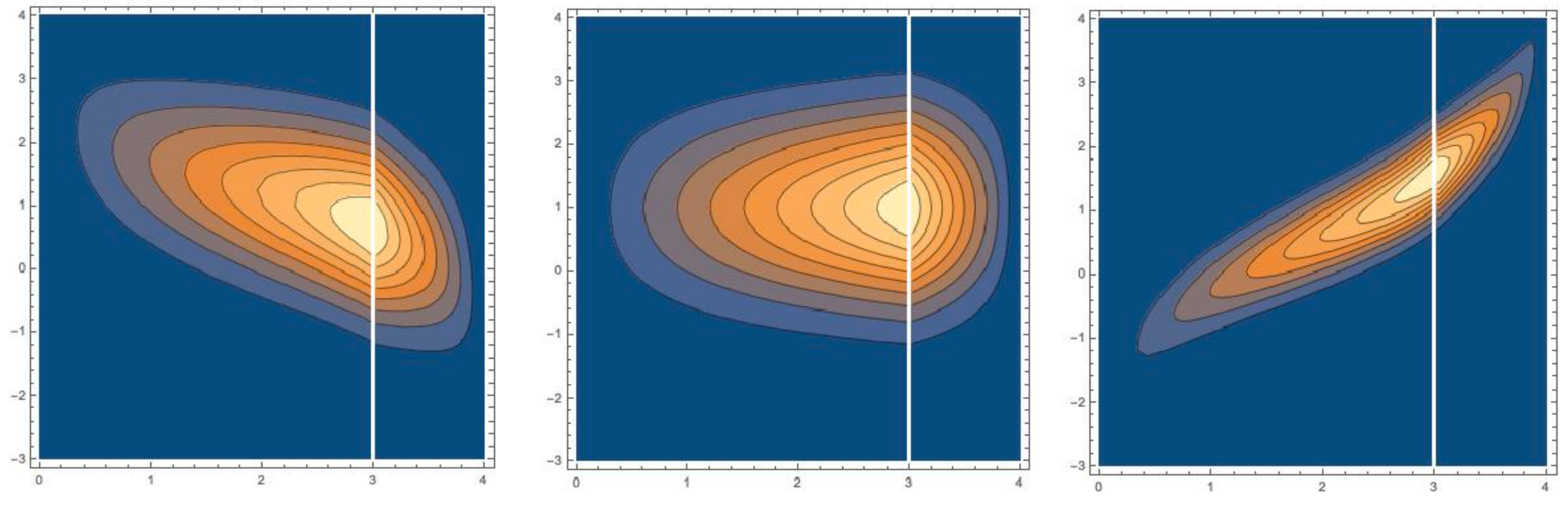
$$\Sigma_i(x) = \int_{-\infty}^{x} dx' P_i(x')$$

$$\Phi^{-1}(p) = \sqrt{2}\mathrm{Erf}^{-1}(-1+2p)$$

$$\xi(x) = \Phi^{-1}(\Sigma_1(x))$$

$$\eta(y) = \Phi^{-1}(\Sigma_2(y))$$





$c = -0.5$ $c = 0$ $c = 0.9$

use inference to constrain c

Conclusions & Outlook

- Bayesian inference is a powerful framework to estimate Δ_{MHO}
  - probabilistic interpretation ⇝ $P(\delta_{n+1}\,|\,\delta_n)$
  - exposes our *assumptions & biases* clearly ⇝ model & priors
  - *but:* it is not more reliable than scale variation ⇝ careful analysis required

relying on a single prescription for TH unc. in *precision measurements* that does not admit a probabilistic interpretation is potentially dangerous!

- typically for $n < 2$: $\mathrm{CI}_{68} > 9\mathrm{pt}$; $n \geq 2$: $\mathrm{CI}_{68} \simeq 9\mathrm{pt}$
- *public code:* ミホ (miho) ⇝ https://github.com/aykhuss/miho
- future directions
  - correlations (p_T^W/p_T^Z, p_T^Z vs. p_T^ℓ, PDF fits & data interpretation, …) [AH, Mazeliauskas w.i.p]
  - marginalisation over models, …

CONCLUSIONS & OUTLOOK

- Bayesian inference is a powerful framework to estimate Δ_{MHO}
  - probabilistic interpretation ⇝ $P(\delta_{n+1}\,|\,\delta_n)$
  - exposes our *assumptions & biases* clearly ⇝ model & priors
  - *but:* it is not more reliable than scale variation ⇝ careful analysis required

relying on a single prescription for TH unc. in *precision measurements* that does not admit a probabilistic interpretation is potentially dangerous!

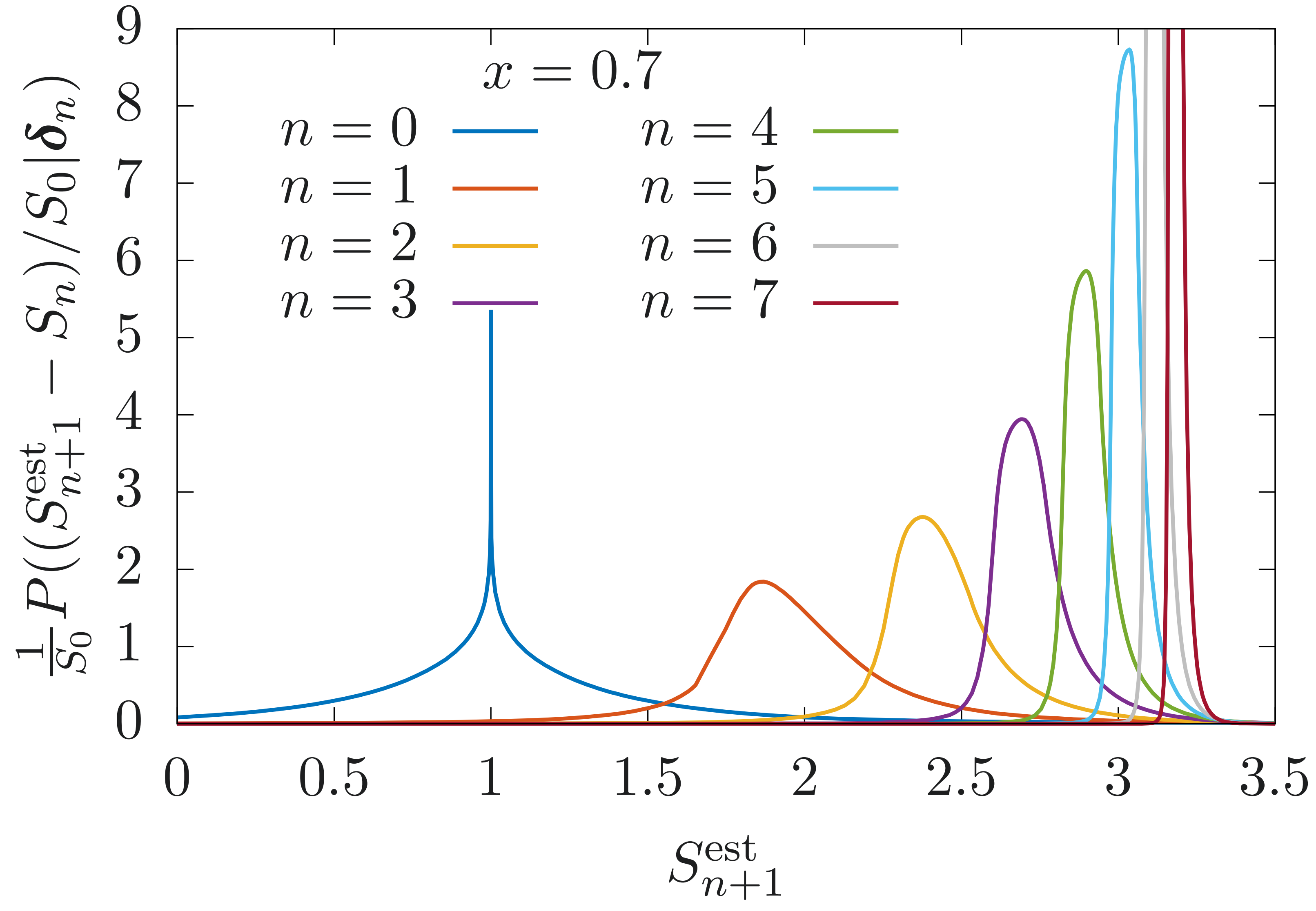
- typically for $n < 2$: $\mathrm{CI}_{68} > 9\mathrm{pt}$; $n \geq 2$: $\mathrm{CI}_{68} \simeq 9\mathrm{pt}$
- *public code:* ミホ (miho) ⇝ https://github.com/aykhuss/miho
- future directions
  - correlations $(p_{\mathrm{T}}^{W}/p_{\mathrm{T}}^{Z},\ p_{\mathrm{T}}^{Z}$ vs. p_{T}^{ℓ}, PDF fits & data interpretation, …) [AH, Mazeliauskas w.i.p]
  - marginalisation over models, …

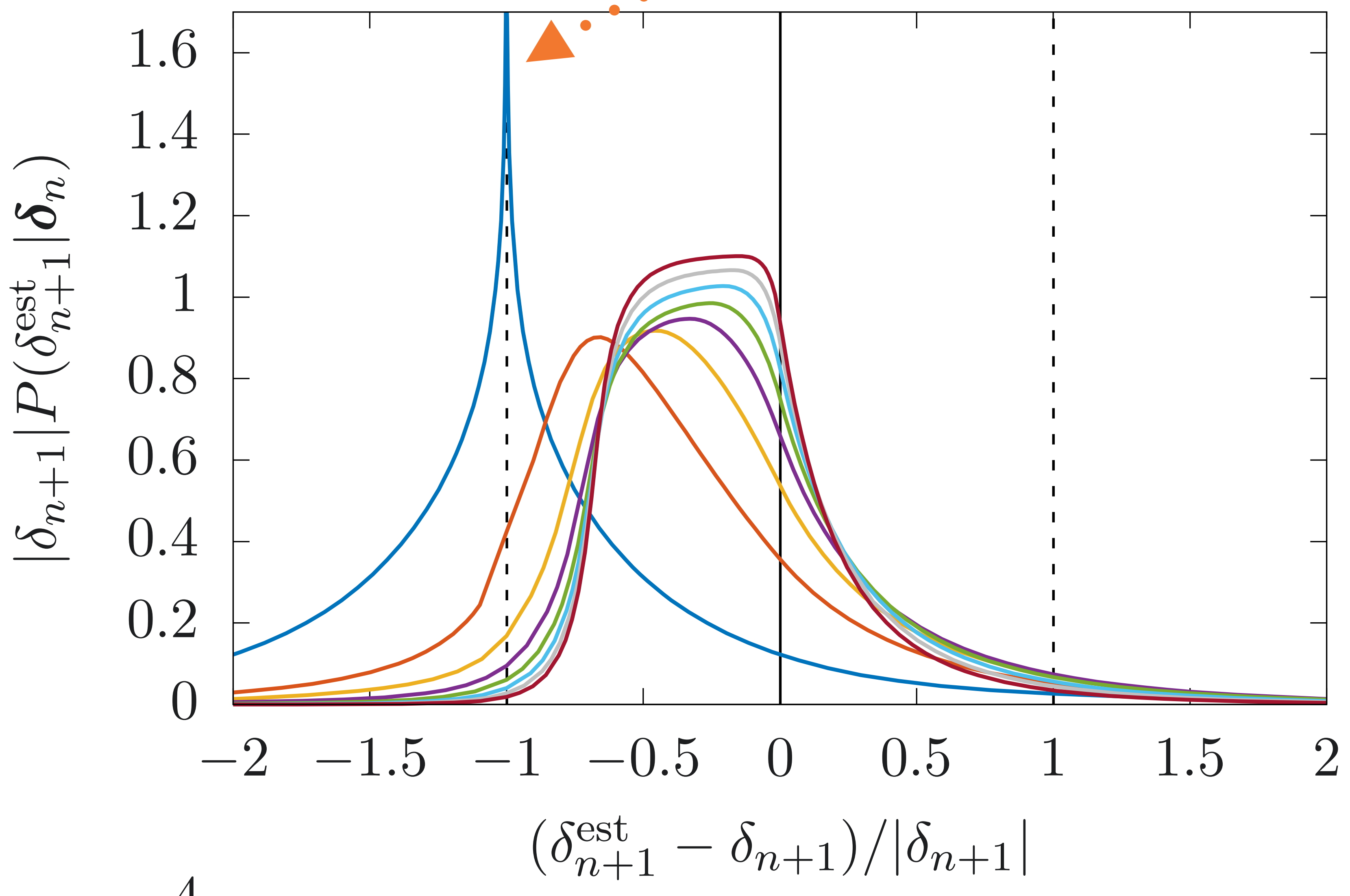
Thank you!

BACKUP

Toy Example – $\delta_k = (0.7)^k$

geometric would be localised here



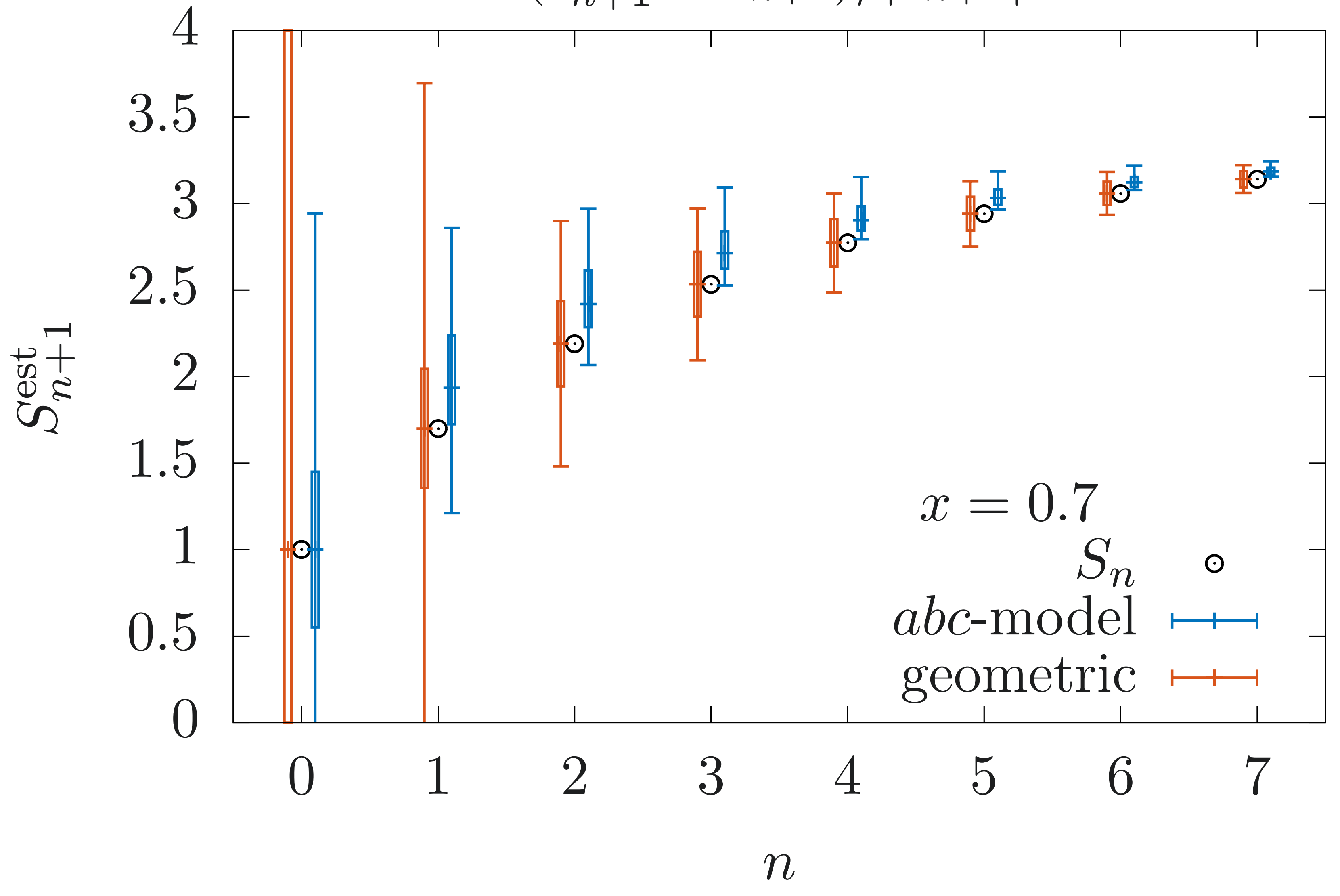


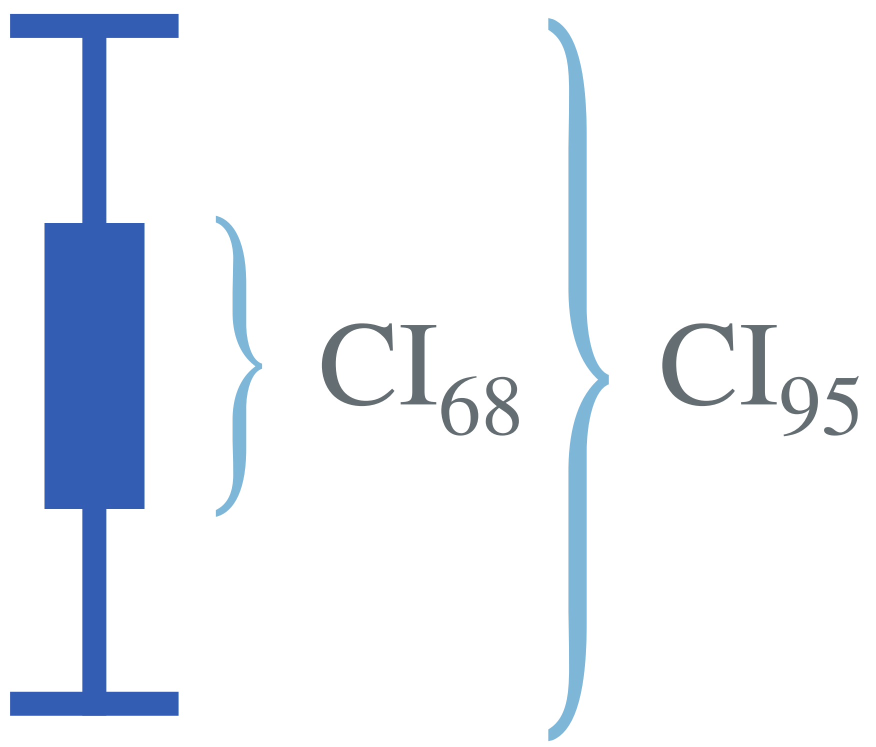
correctly anticipates positive δ_{n+1}

- Confidence Intervals

$$\text{CI}_x = [\Sigma_x^{\text{low}}, \Sigma_x^{\text{upp}}]$$

containing $x\,\%$ of the probability



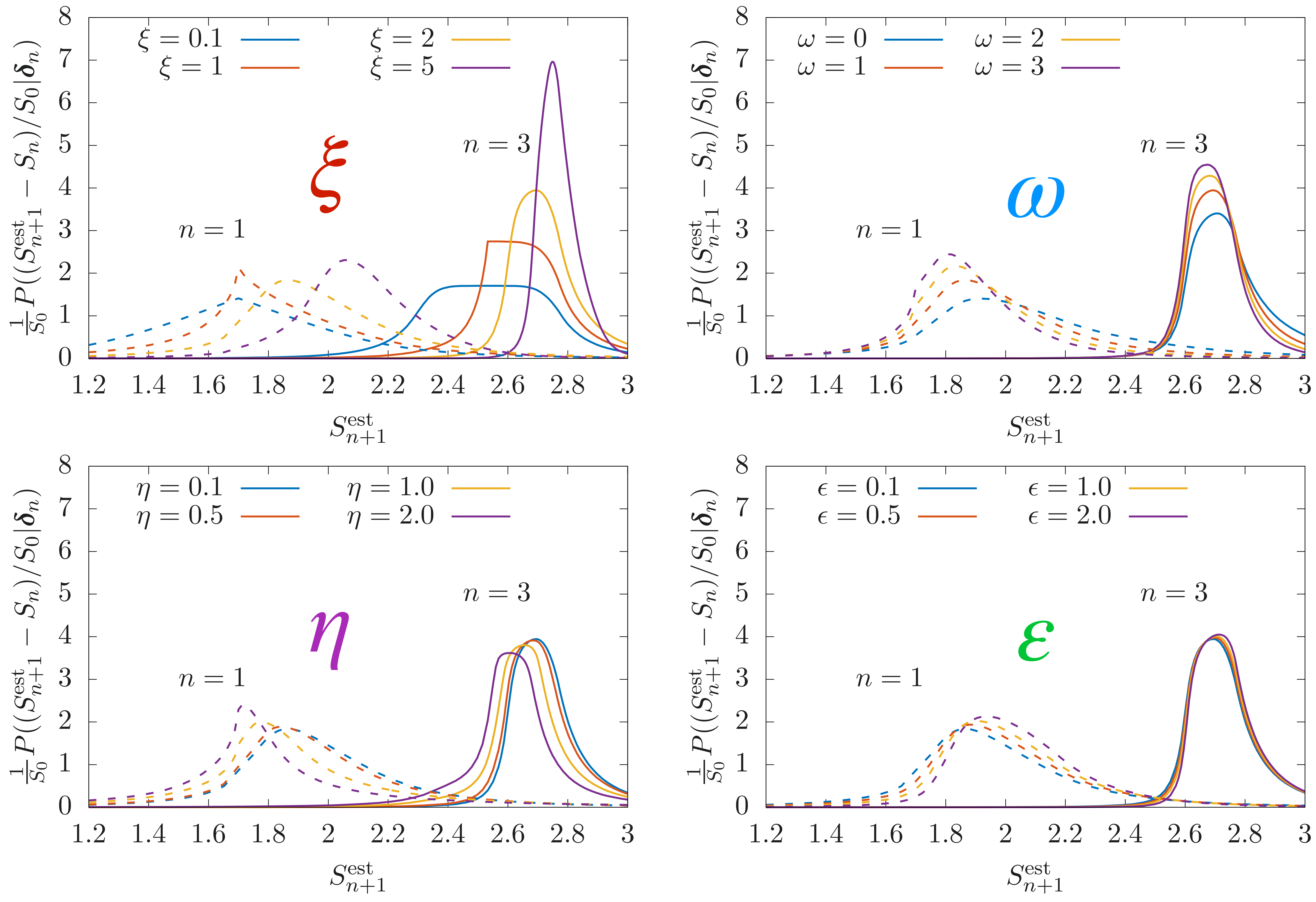


Sensitivity on the Priors — $(\varepsilon, \omega, \eta, \xi) \simeq (0.1, 1, 0.1, 2)$

$$P_0(a) = \tfrac{1}{2}\,(1+\omega)\,(1-|a|)^{\omega}\,\Theta(1-|a|) \qquad P_0(b,c) = \frac{\varepsilon\,\eta^{\varepsilon}}{2\xi c^{2+\varepsilon}}\,\Theta(c-\eta)\,\Theta(\xi c - b)$$

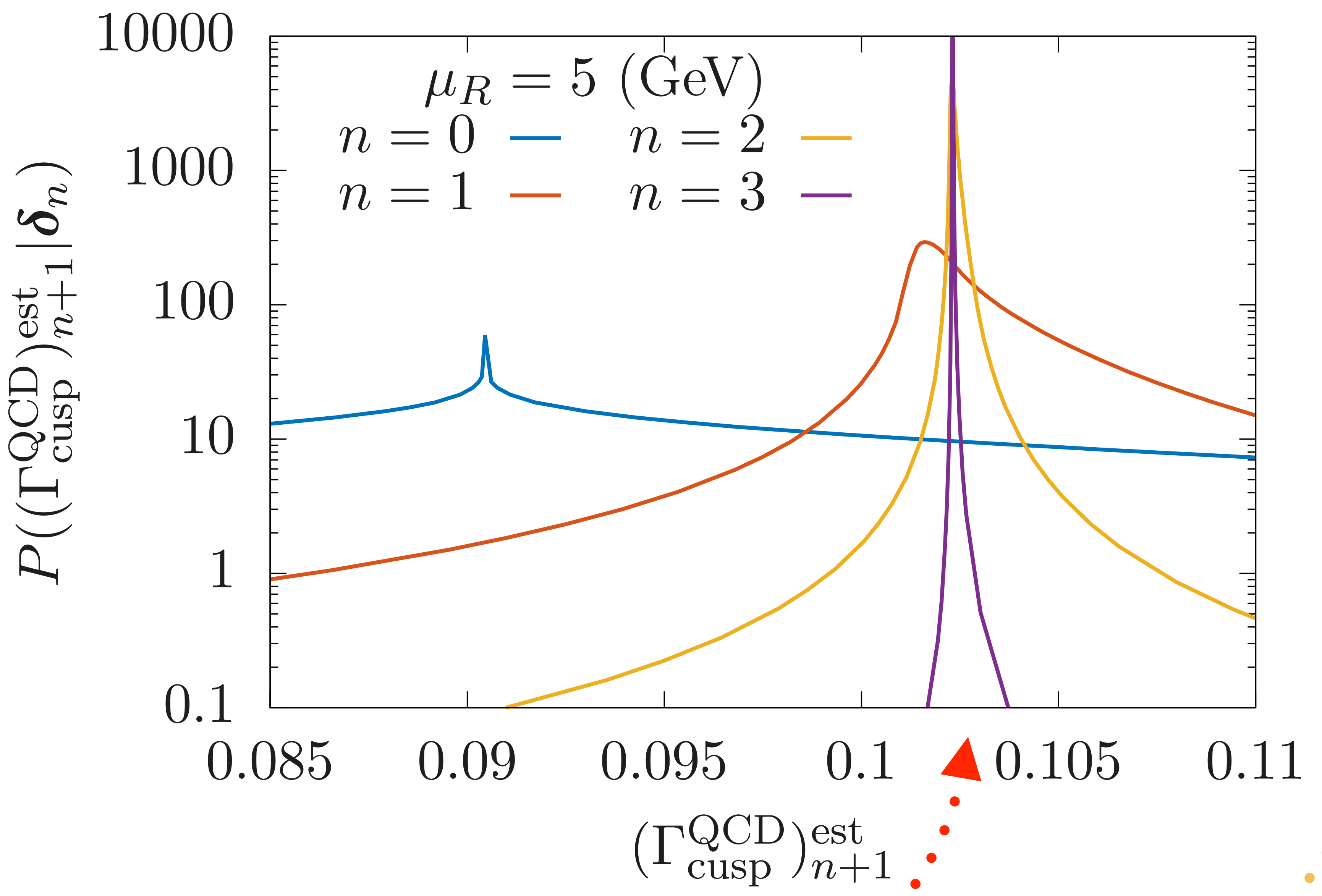
- dependence on priors decreases as n increases
- less for ξ as it controls the asymmetry of the distribution
- for geometric series:

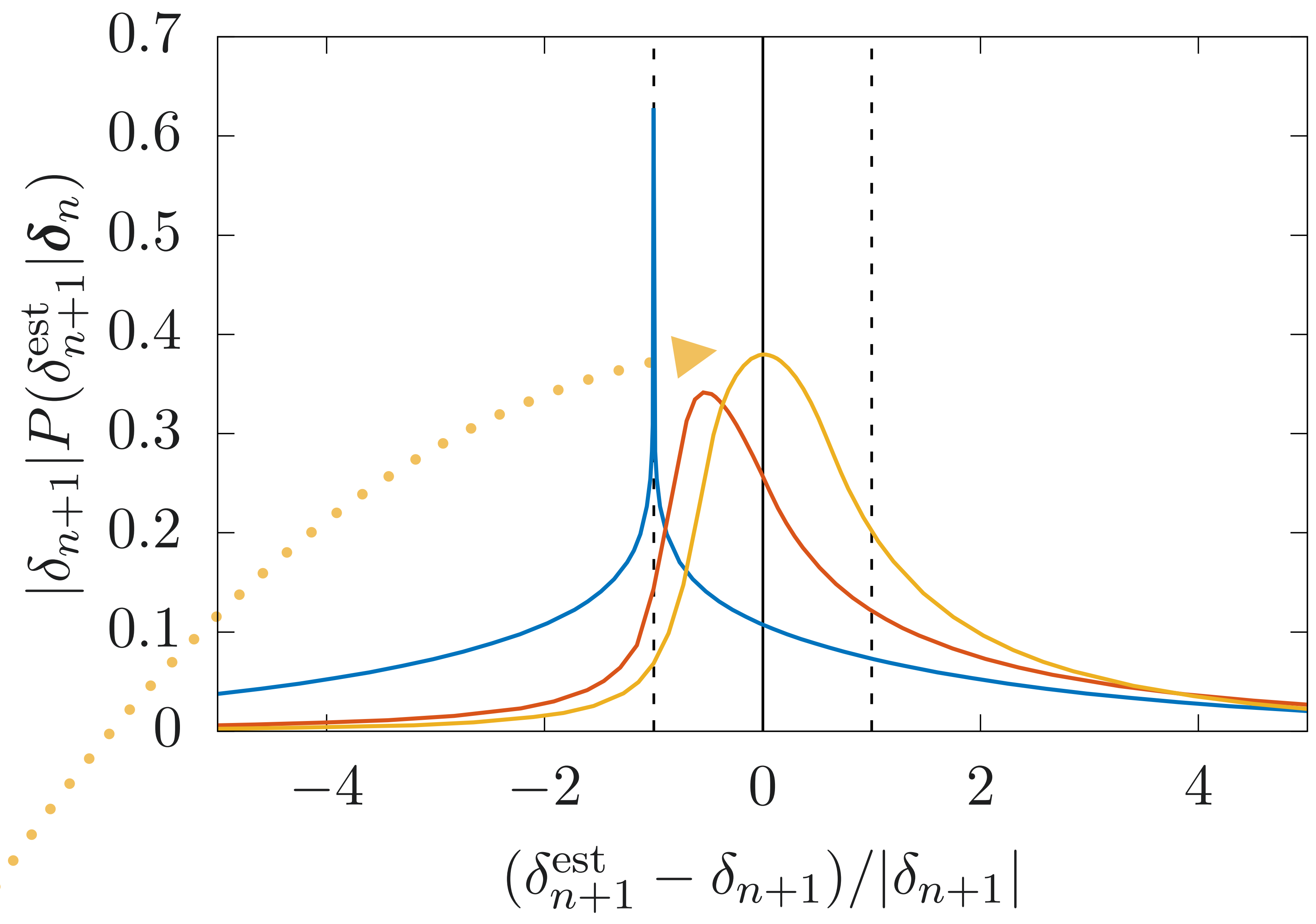
$$\left.\begin{matrix}\eta \to 0 \\ \xi \to \infty\end{matrix}\right\} \rightsquigarrow S_{n+1}^{\text{est}} \to S_{n+1}$$



A Real-World Example — Cusp Anomalous Dimension in QCD

$$\Gamma^{\rm QCD}_{\rm cusp} = C_F \sum_{k=1}^{\infty} \left(\frac{\alpha_s}{\pi}\right)^k \gamma^{(k)}_{\rm QCD}$$

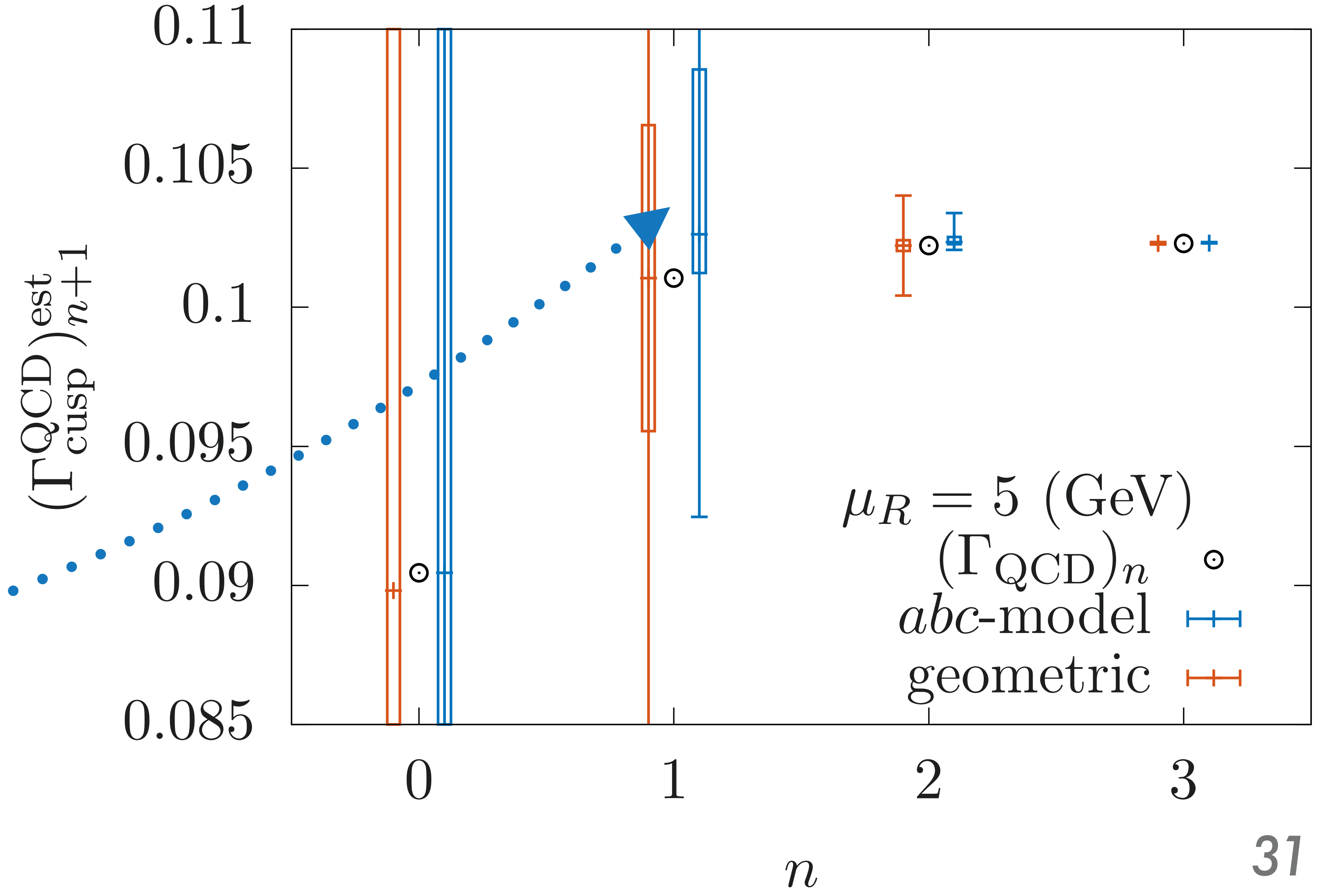




strongly peaked $n \nearrow$

$n = 2$: anticipates the exact known result

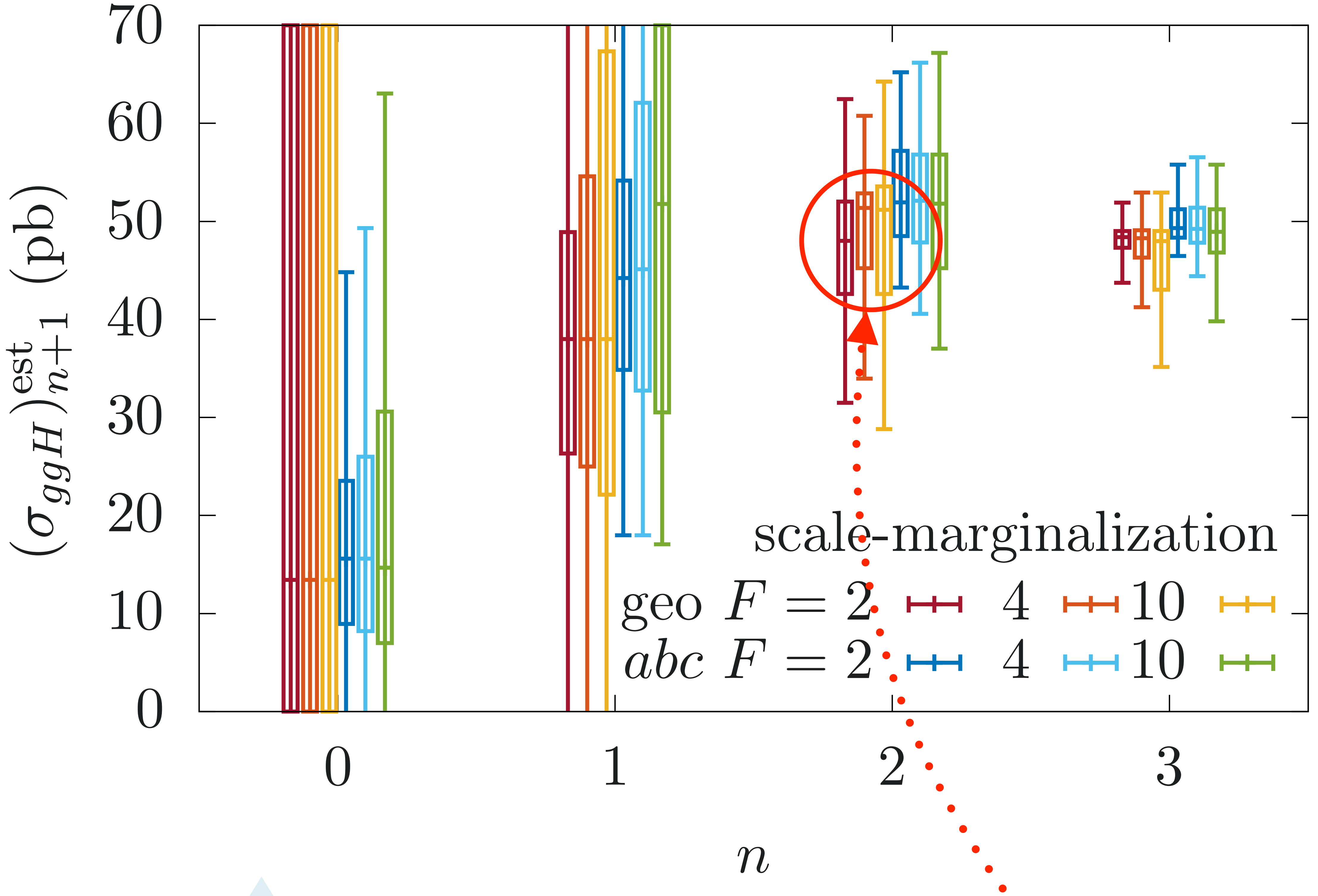
CI_{68} always captures next term



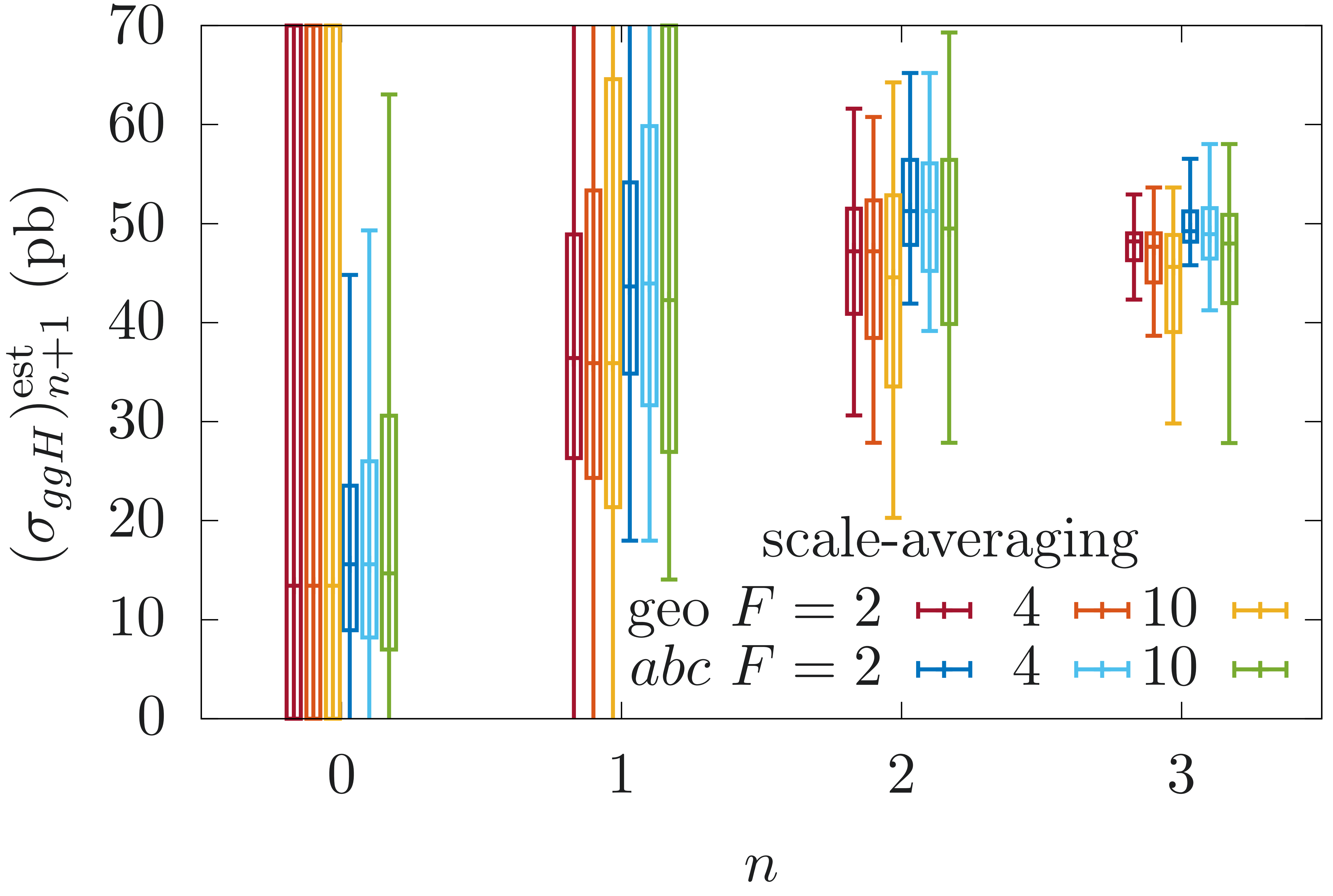
- estimate for $\Gamma^{\rm QCD}_{\rm cusp} - \left(\Gamma^{\rm QCD}_{\rm cusp}\right)_4$
  - $\mathrm{CI}_{68} = C_F \left(\frac{\alpha_s}{\pi}\right)^5 [2.1, 9.5]$
  - $\mathrm{CI}_{95} = C_F \left(\frac{\alpha_s}{\pi}\right)^5 [-0.38, 21]$

SENSITIVITY ON THE RANGE F

Scale Marginalisation (sm):



Scale Average (sa):



milder dependence on F in sm procedure

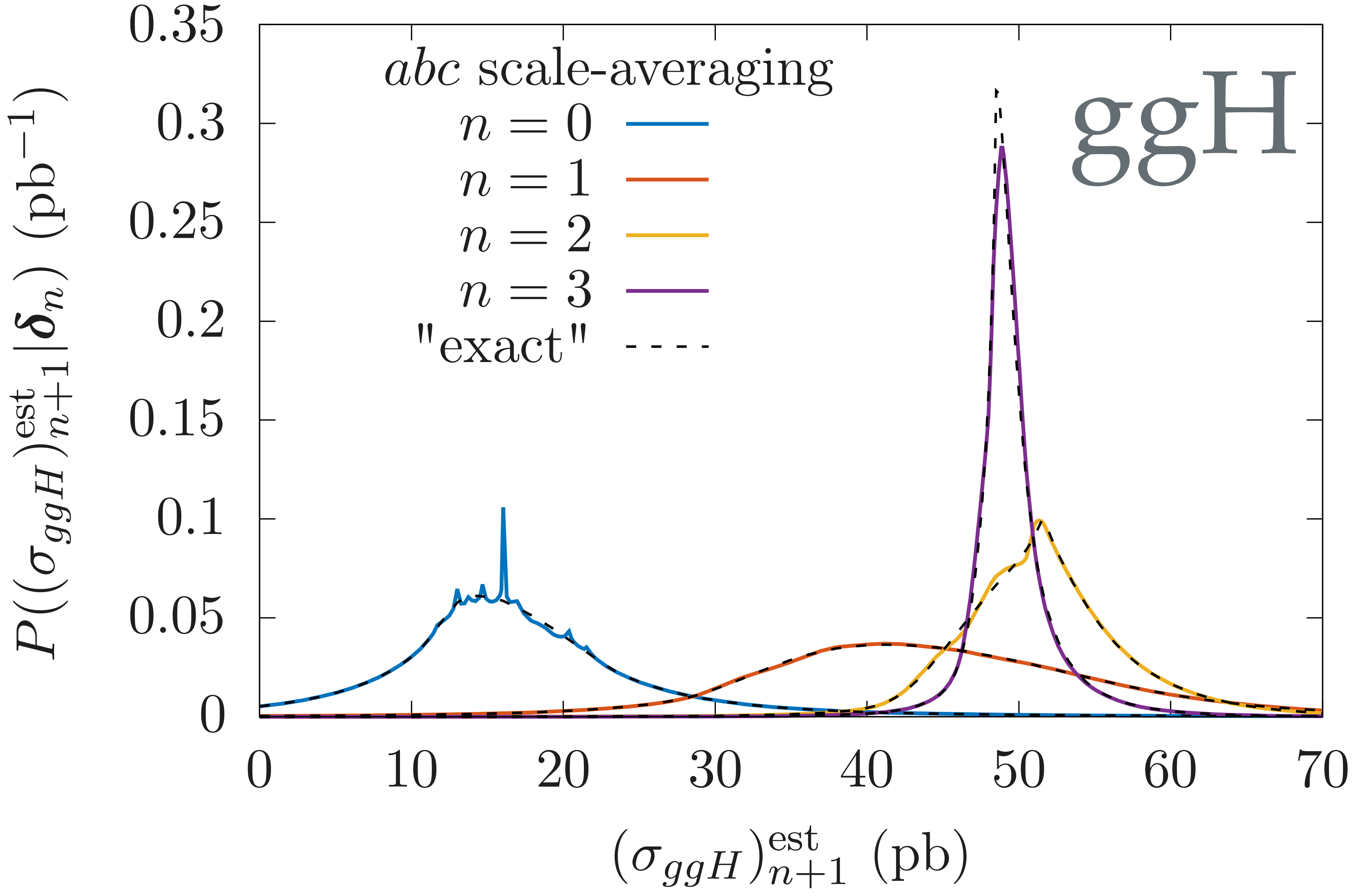
transition to $\mu_{\mathrm{FAC}} \in [\mu_0/F, F\,\mu_0]$

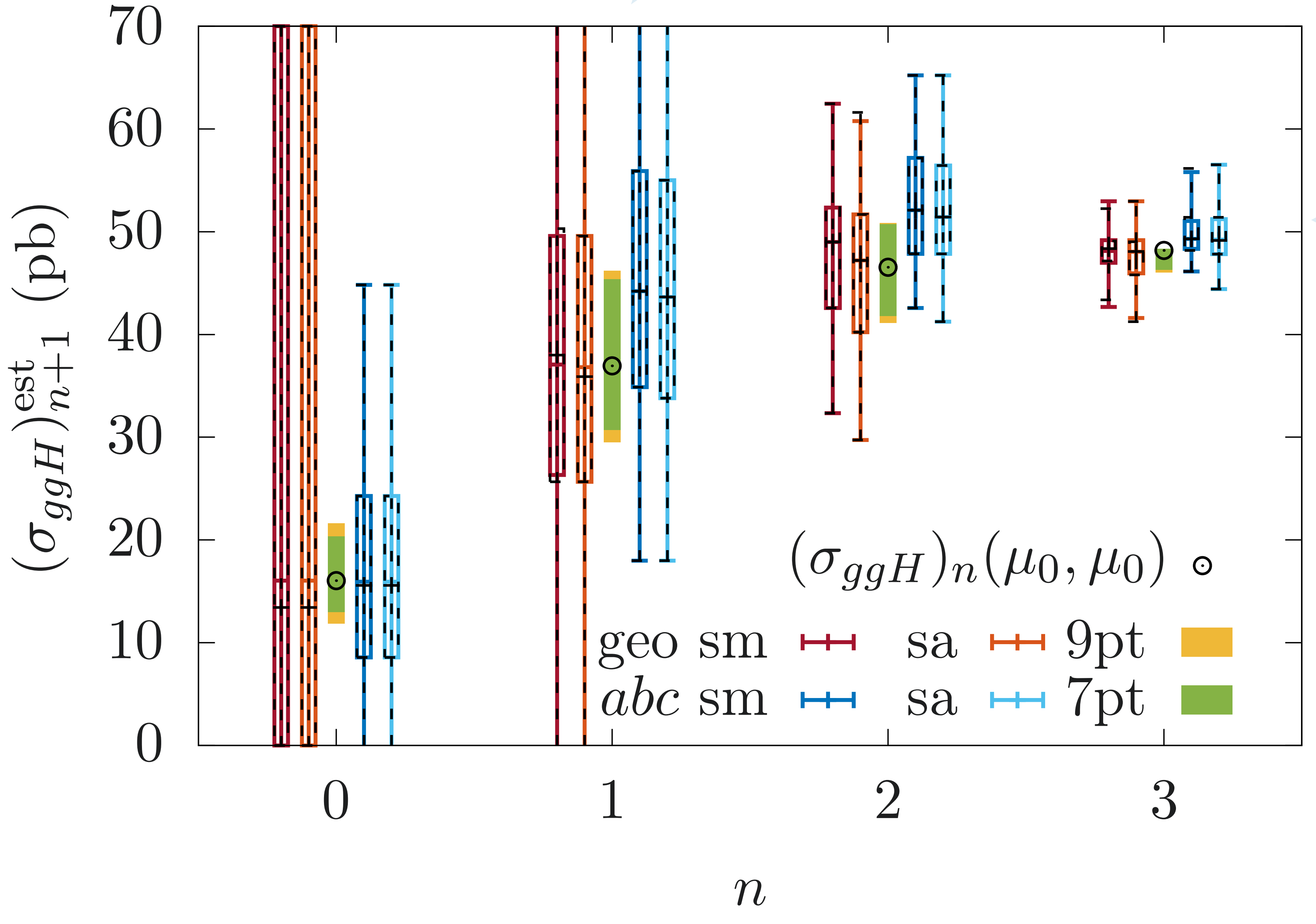
Intermezzo — Integration over μ

- the integration over μ is in general very costly (numerical)
  - ⇝ approximate it using *quadrature rule* (works well for $\mathrm{CI}_{68/95}$)
  - ⇝ recycle existing calculations done for $\{\mu_0/2,\, 2\,\mu_0\}$
- Gauss-Legendre $\left(w_0 = \frac{8}{18}\,;\ w_{\pm1} = \frac{5}{18}\right)$ ⇝ corresponds to $F \simeq 2.45$

$$\int \mathrm{d}\mu\ w(\mu)\, f(\mu) \simeq w_{-1}\ f(\mu_0/2) + w_0\ f(\mu_0) + w_{+1}\ f(2\mu_0)$$

more conservative up to NLO





from NNLO on similar to 7/9-pt

N3LO 7/9-pt asymmetric