A statistical description of theory uncertainty from missing higher orders

Marco Bonvini

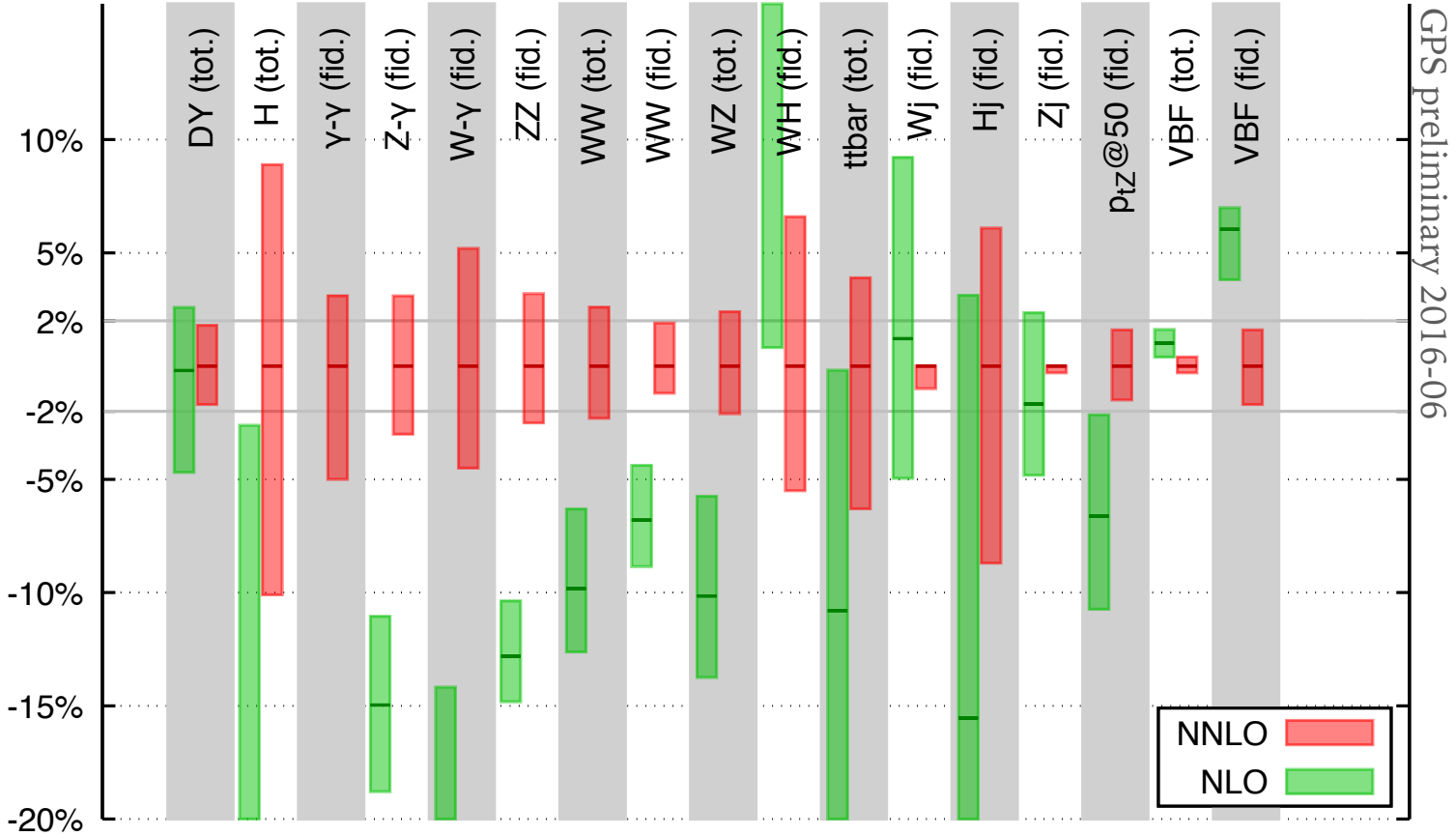
INFN, Rome 1 unit

DIS2019, 11 April 2019

How accurate are current theoretical predictions?

WHAT PRECISION AT NNLO?

Slide from Gavin Salam, PSR 2016



For many processes NNLO scale band is ~±2%

Though only in 3/17 cases is NNLO (central) within NLO scale band…

How do we determine theory uncertainties?

Scale variation: dependence on unphysical scales of a physical observable at N^{n}LO is of higher order

$$\Sigma_{\mathsf{N}^n\mathsf{LO}}(\mu) = \sum_{k=0}^{n} c_k(\mu)\alpha_s^k(\mu) \qquad \mu\frac{d}{d\mu}\Sigma_{\mathsf{N}^n\mathsf{LO}}(\mu) = \mathcal{O}(\alpha_s^{n+1})$$

Canonical uncertainty: variation by a factor of 2 about a "central" scale μ_0

$$\Sigma_{\mathsf{true}} \approx \Sigma_{\mathsf{N}^n\mathsf{LO}}(\mu_0) \pm \max_{\mu_0/2\leq\mu\leq 2\mu_0} |\Sigma_{\mathsf{N}^n\mathsf{LO}}(\mu) - \Sigma_{\mathsf{N}^n\mathsf{LO}}(\mu_0)|$$

Other (less used) methods:

- use the last known order: $\Sigma_{\mathsf{true}} \approx \Sigma_{\mathsf{N}^n\mathsf{LO}} \pm |\Sigma_{\mathsf{N}^n\mathsf{LO}} - \Sigma_{\mathsf{N}^{n-1}\mathsf{LO}}|$
- include other scales or parameters to get more variations (e.g. in resummed calculations)
- novel approach recently proposed by F.Tackmann [slides at SCET2019]
- ....

Caveats of canonical methods

$$\Sigma_{\text{true}} \approx \Sigma_{\text{N}^n\text{LO}}(\mu_0) \pm \max_{\mu_0/2 \leq \mu \leq 2\mu_0} |\Sigma_{\text{N}^n\text{LO}}(\mu) - \Sigma_{\text{N}^n\text{LO}}(\mu_0)|$$

Which central scale μ_0?

There are guiding principles, but several values are acceptable

How much should I vary the scale?

The factor of 2 is an arbitrary choice

How do I interpret the uncertainty?

I can say that it's a range where I'm sure the next order will lie, or the 1-sigma of a gaussian, but it's very arbitrary...

Need for a statistically-sound definition of theoretical uncertainties, which does not depend so much on arbitrary assumptions

Ideally, theory uncertainty from missing higher orders should be a **probability distribution**

This would also allow for a statistically meaningful comparison of theory predictions with data (e.g. in a PDF fit, or in precision tests of the SM)

A different approach based on Bayesian statistics

Cacciari and Houdeau [1105.5152] proposed a statistical model for the interpretation of theory uncertainties, from which one can compute the probability distribution for the observable Σ given the first few terms

$$\Sigma = \sum_k c_k \alpha_s^k$$

"We make the assumption that all the coefficients c_k in a perturbative series share some sort of upper bound $\bar{c} > 0$ to their absolute values, specific to the physical process studied. The calculated coefficients will give an estimate of this $\bar{c}$, restricting the possible values for the unknown c_k."

In other words, the model assumes that

$$|c_k| \leq \bar{c} \quad \forall k$$

implying that the perturbative expansion is bounded by a geometric series

$$|\Sigma| = \left| \sum_k c_k \alpha_s^k \right| \leq \sum_k \bar{c} \alpha_s^k = \frac{\bar{c}}{1 - \alpha_s}$$

Caveats of the Cacciari-Houdeau model

- it assumes a convergent perturbative expansion
  - acceptable from an asymptotic expansion point of view
  - attempt to treat the series as factorially divergent [Bagnaschi,Cacciari,Guffanti,Jenniches 1409.5036]
- if the coefficients grow as a power, $c_k \sim \eta^k$, which is very likely, the method cannot perform well
  - Cacciari-Houdeau proposed a modified version with η accounted for
  - in [Bagnaschi,Cacciari,Guffanti,Jenniches 1409.5036] η is determined from a survey on various observables
  - in an alternative approach [Forte,Isgrò,Vita 1312.6688] the value of η is fitted
- it still depends on the choice of the "central" scale

In this talk

$$\Sigma = \Sigma_{\mathsf{LO}}(\mu) \sum_{k \geq 0} \delta_k(\mu) \qquad \Sigma_{\mathsf{LO}}(\mu)\delta_k(\mu) = c_k(\mu)\alpha_s^k(\mu)$$

I propose two new methods based on the following assumptions:

In this talk

$$\Sigma = \Sigma_{\mathsf{LO}}(\mu) \sum_{k \geq 0} \delta_k(\mu) \qquad \Sigma_{\mathsf{LO}}(\mu)\delta_k(\mu) = c_k(\mu)\alpha_s^k(\mu)$$

I propose two new methods based on the following assumptions:

- **geometric behaviour model (i.e., improved Cacciari-Houdeau)**

$$|\delta_k(\mu)| \leq c\, a^k \qquad \forall k$$

depends on two hidden parameters c, a, and also on μ

In this talk

$$\Sigma = \Sigma_{\mathsf{LO}}(\mu) \sum_{k \geq 0} \delta_k(\mu) \qquad \Sigma_{\mathsf{LO}}(\mu)\delta_k(\mu) = c_k(\mu)\alpha_s^k(\mu)$$

I propose two new methods based on the following assumptions:

- **geometric behaviour model (i.e., improved Cacciari-Houdeau)**

$$|\delta_k(\mu)| \leq c\, a^k \qquad \forall k$$

  depends on two hidden parameters c, a, and also on μ

- **scale variation inspired model**

$$|\delta_k(\mu)| \leq \lambda\, r_{k-1}(\mu) \qquad \forall k, \qquad r_k(\mu) \simeq \left| \mu \frac{d}{d\mu} \log \Sigma_{\mathsf{N}^k\mathsf{LO}}(\mu) \right|$$

  depends on one hidden parameter λ, and also on μ

- Assumptions not contradictory: we can assume **both** at the same time!

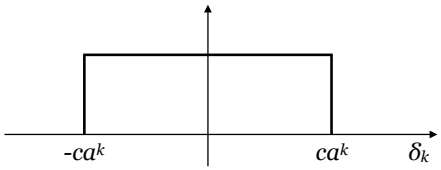
Model 1: Geometric behaviour

The condition

$$|\delta_k(\mu)| \leq ca^k$$

is implemented through the likelihood

$$P(\delta_k|c,a,\mu) \propto \theta(ca^k - |\delta_k(\mu)|) =$$



namely all values of δ_k within the allowed range are equally likely

Inference scheme:

$$\underbrace{\delta_0, ..., \delta_n}_{\text{known}} \xrightarrow{\text{inference}} c, a \xrightarrow{\text{inference}} \underbrace{\delta_{n+1}, \delta_{n+2}, ...}_{\text{unknown}} \xrightarrow{\text{sum}} \Sigma$$

Final output:

$$P(\Sigma|\delta_0, ..., \delta_n, \mu, \text{model}_1)$$

Note: The assumption is good for convergent series, but a power divergence is also possible if $a > 0$ is allowed

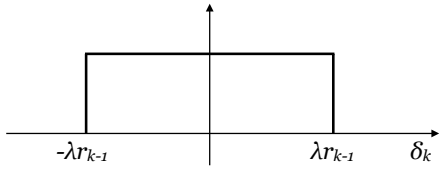
Model 2: Scale variation inspired

The condition

$$|\delta_k(\mu)| \leq \lambda r_{k-1}(\mu)$$

is implemented through the likelihood

$$P(\delta_k | r_{k-1}, \lambda, \mu) \propto \theta(\lambda r_{k-1} - |\delta_k(\mu)|) =$$



namely all values of δ_k within the allowed range are equally likely

Inference scheme:

$$\underbrace{\delta_0, ..., \delta_n, r_0, ..., r_{n-1}}_{\text{known}} \xrightarrow{\text{inference}} \lambda \xrightarrow{\text{inference}+r_n} \underbrace{\delta_{n+1}}_{\text{unknown}} \xrightarrow{\text{sum}} \Sigma_{\mathrm{N}^{n+1}\mathrm{LO}}$$

in this case only the first missing higher order can be predicted:

$$P(\Sigma_{\mathrm{N}^{n+1}\mathrm{LO}} | \delta_0, ..., \delta_n, \mu, \text{model}_2)$$

Note: Canonical scale variation is ∼recovered for $\lambda = \log 2$; having λ a free parameter generalizes canonical scale variation to any size of the variation

Construct a scale-independent result

Basic idea: treat the unphysical scale μ as a parameter of the model, and simply marginalize it

$$P(\Sigma|\delta_0, ..., \delta_n) = \int d\mu \; P(\Sigma|\delta_0, ..., \delta_n, \mu) P(\mu|\delta_0, ..., \delta_n)$$

where $P(\mu|\delta_0, ..., \delta_n)$ is the posterior distribution for μ given the known orders

This distribution depends on the model

In this approach, inference on μ selects the values that give the best convergence properties according to the model

The prior $P_0(\mu)$ contains our prejudices on what are the most appropriate scales, but the results are largely independent on the precise form and size of the prior $\Rightarrow$ **a lot of arbitrariness is removed!**

Examples/validation

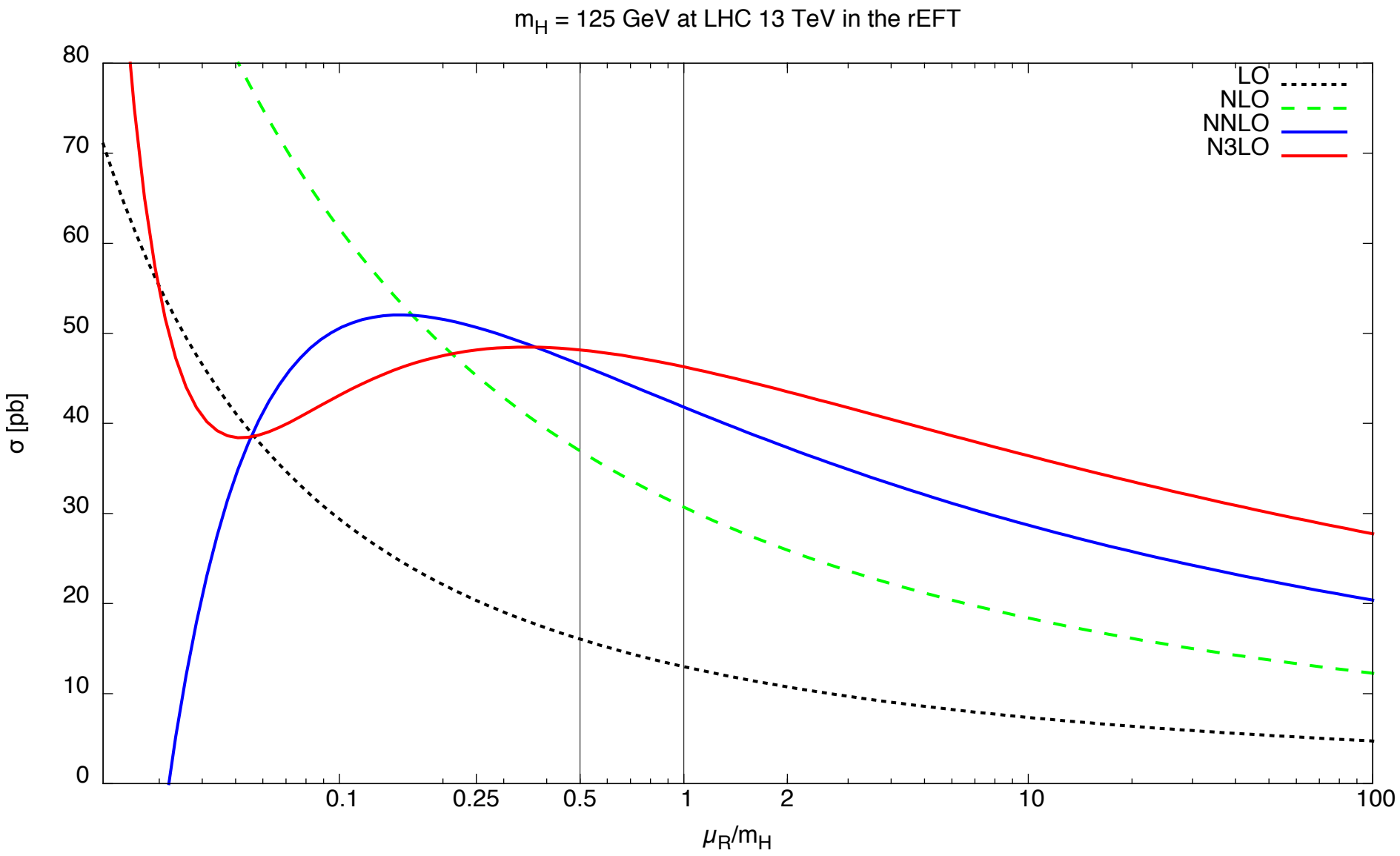
Some examples of observables known at high perturbative order in QCD:

- Higgs production in gluon fusion (ggH) inclusive cross section (N^{3}LO)
  (converges slowly)
- $e^+e^- \to$ hadrons (N^{4}LO)
  (converges fast)
- ....

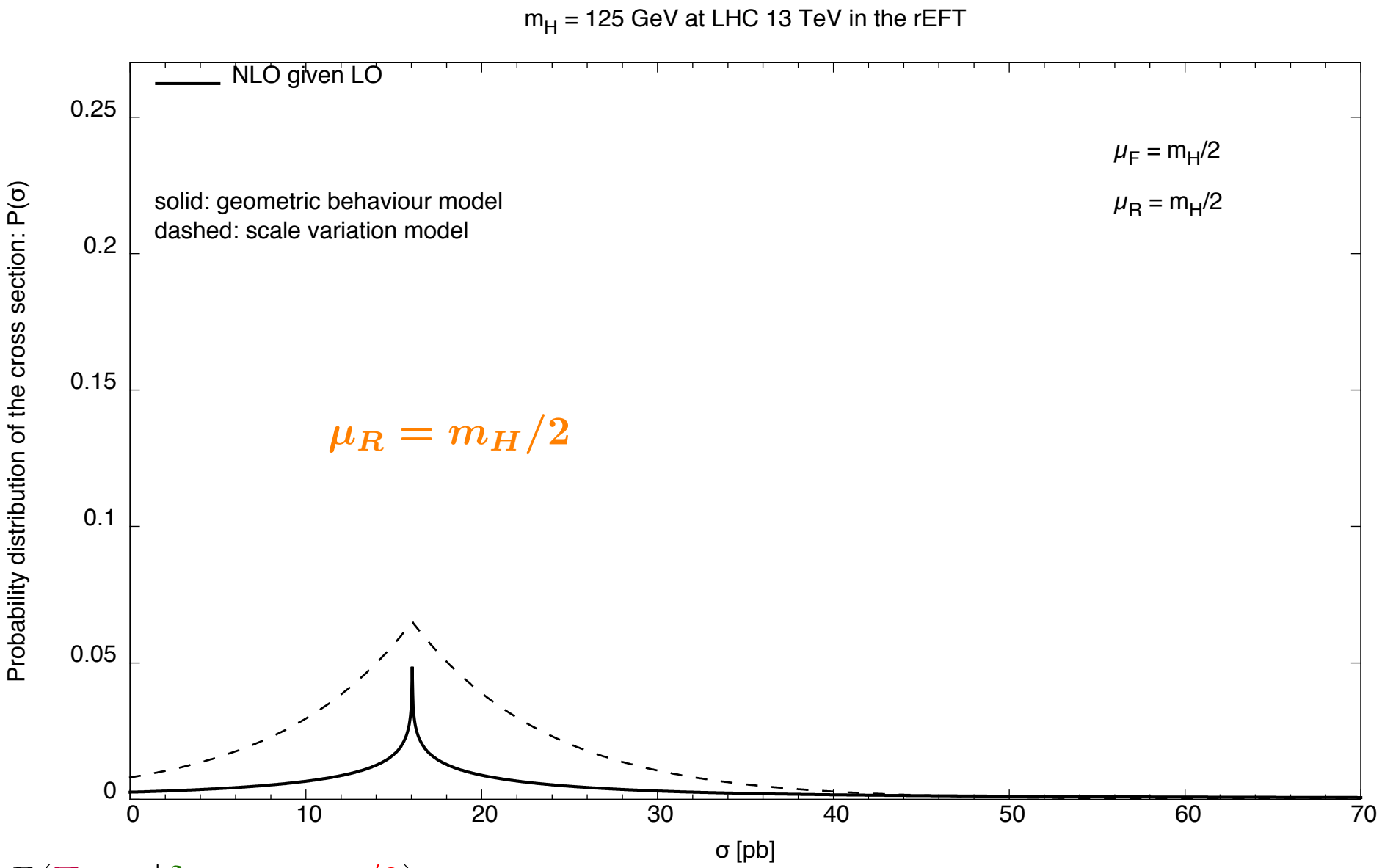
Some applications to series with known sums to verify the goodness of the model:

- purely resummed ggH at N^{3}LL, expanded in powers of α_s
- factorially divergent series with alternating signs $\sum_k (-1)^k k! \alpha_s^k$
- factorially divergent series with equal signs $\sum_k k! \alpha_s^k$
- ....

Higgs in gluon fusion at LHC: the "raw" result

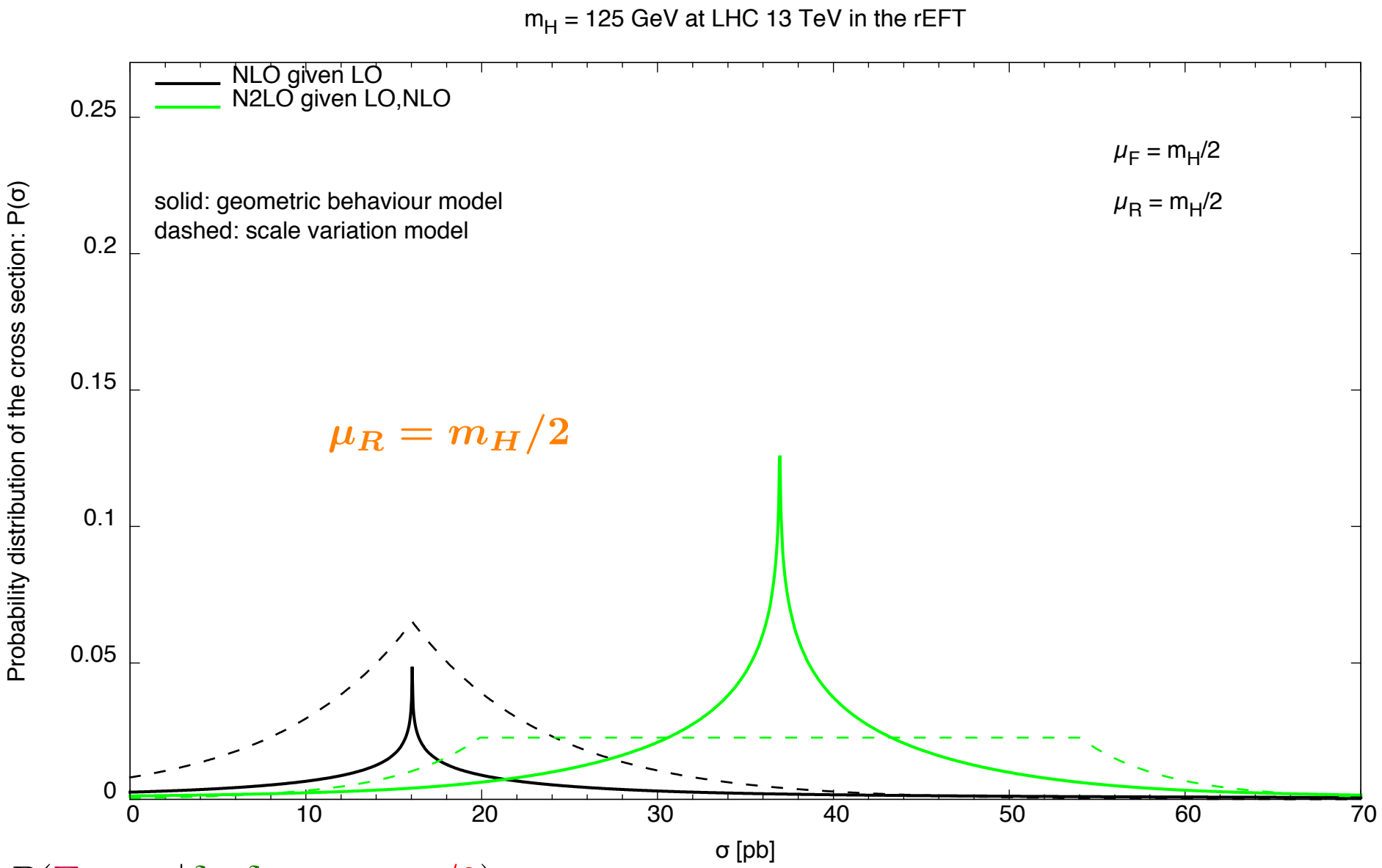


Higgs in gluon fusion at LHC: models at work



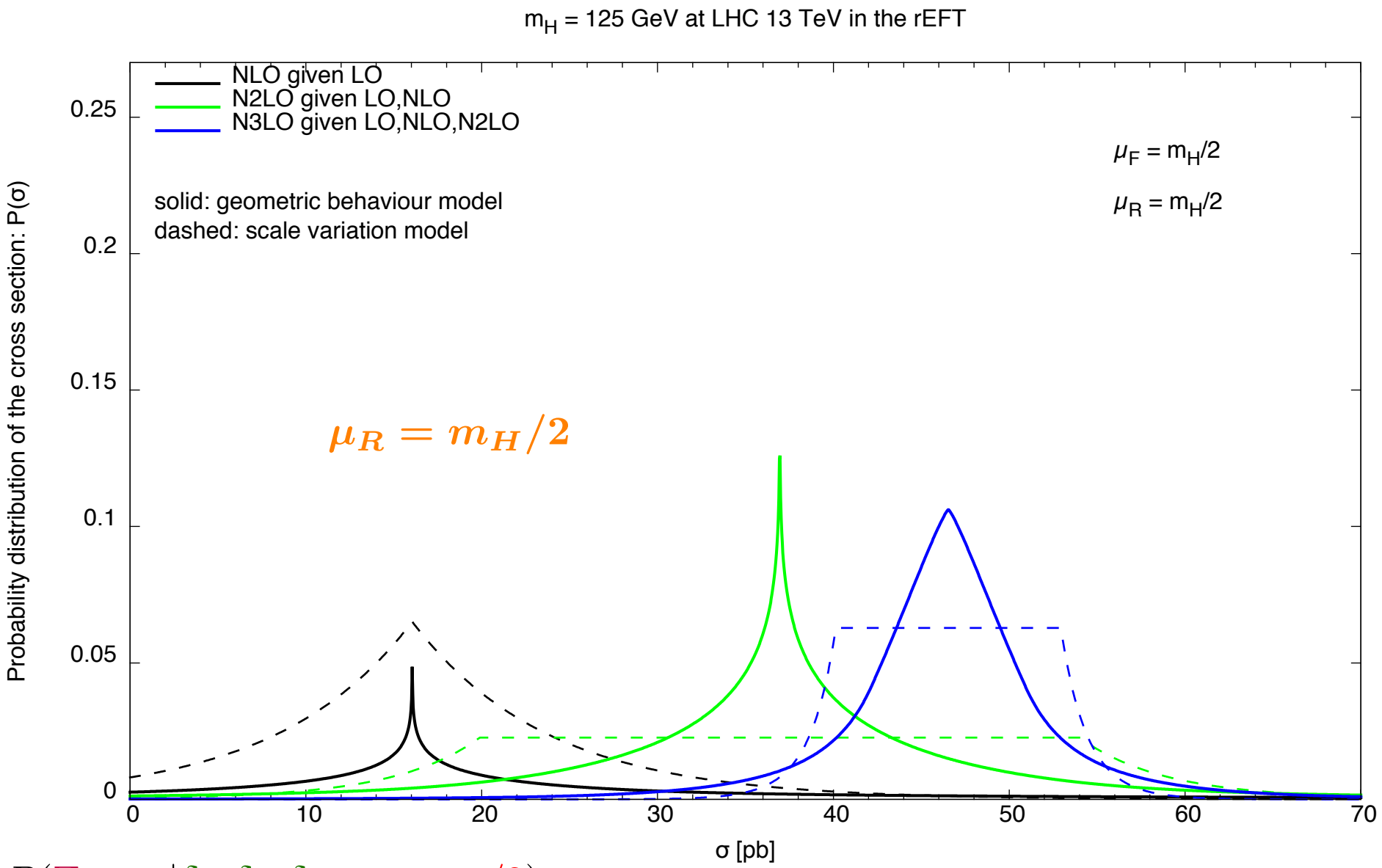
$P(\Sigma_{\rm NLO}|\delta_0, \mu = m_H/2)$

Higgs in gluon fusion at LHC: models at work



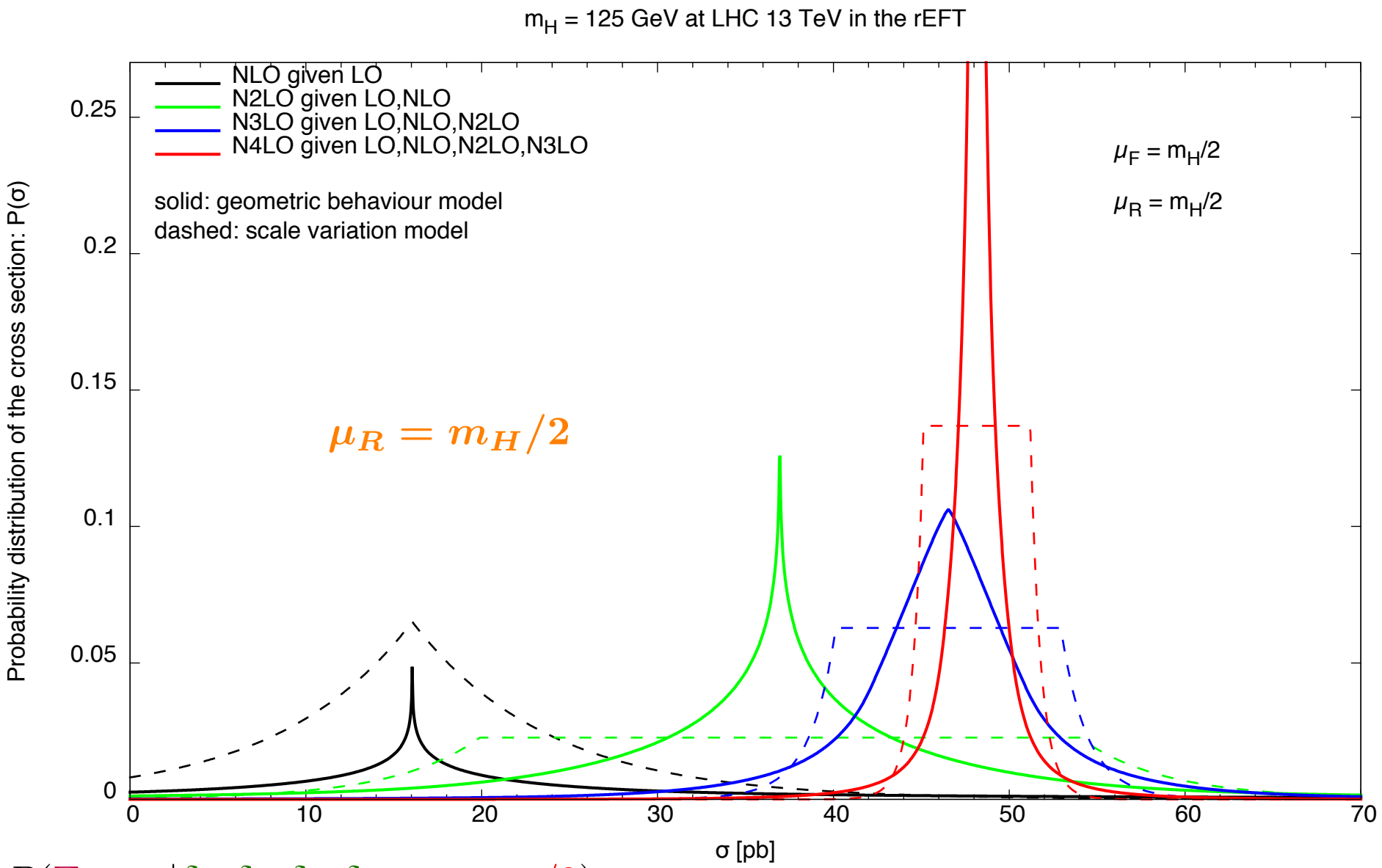
$$P(\Sigma_{\rm NNLO}|\delta_0,\delta_1,\mu=m_H/2)$$

Higgs in gluon fusion at LHC: models at work



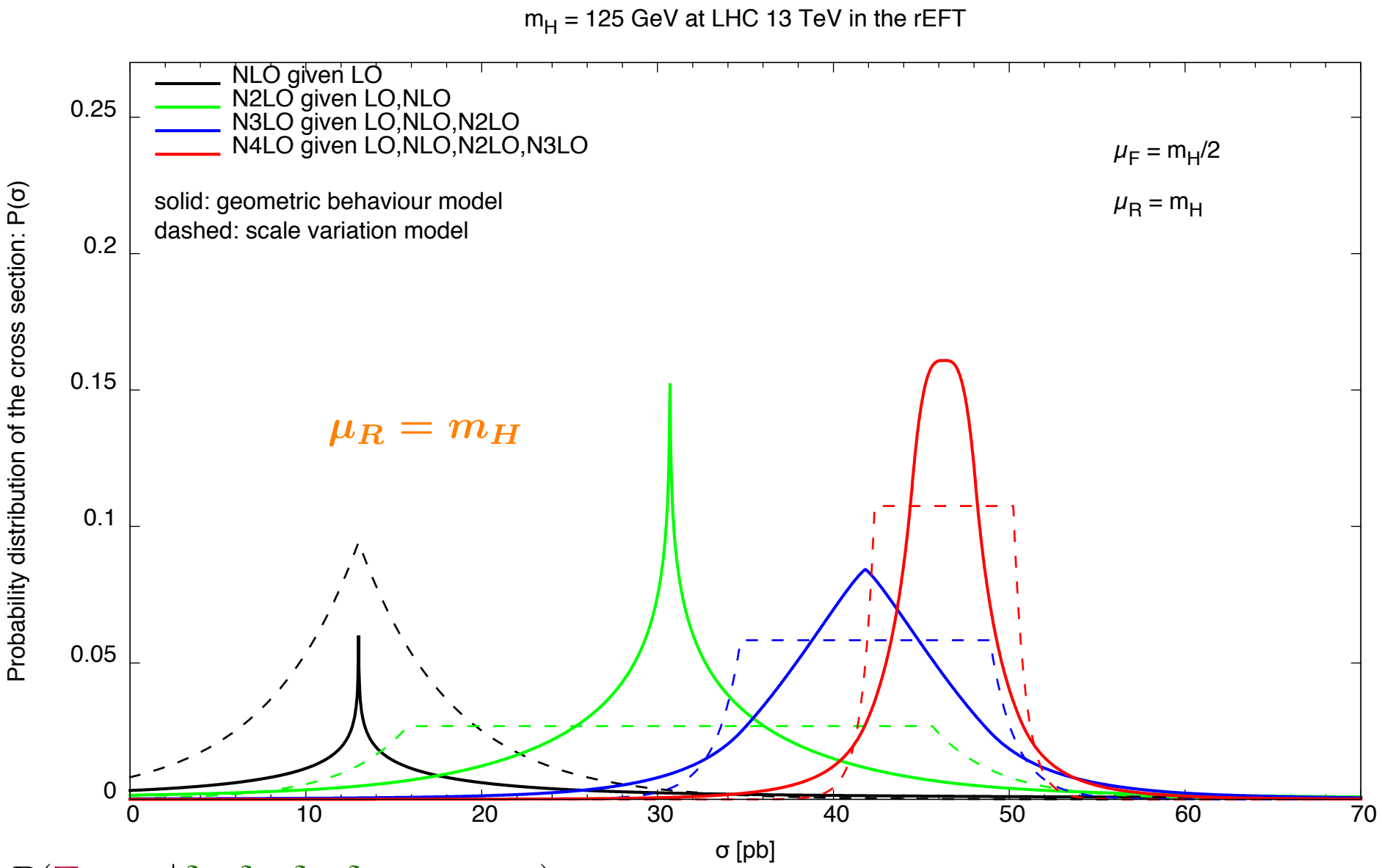
$$P(\Sigma_{\mathrm{N^3LO}}|\delta_0, \delta_1, \delta_2, \mu = m_H/2)$$

Higgs in gluon fusion at LHC: models at work



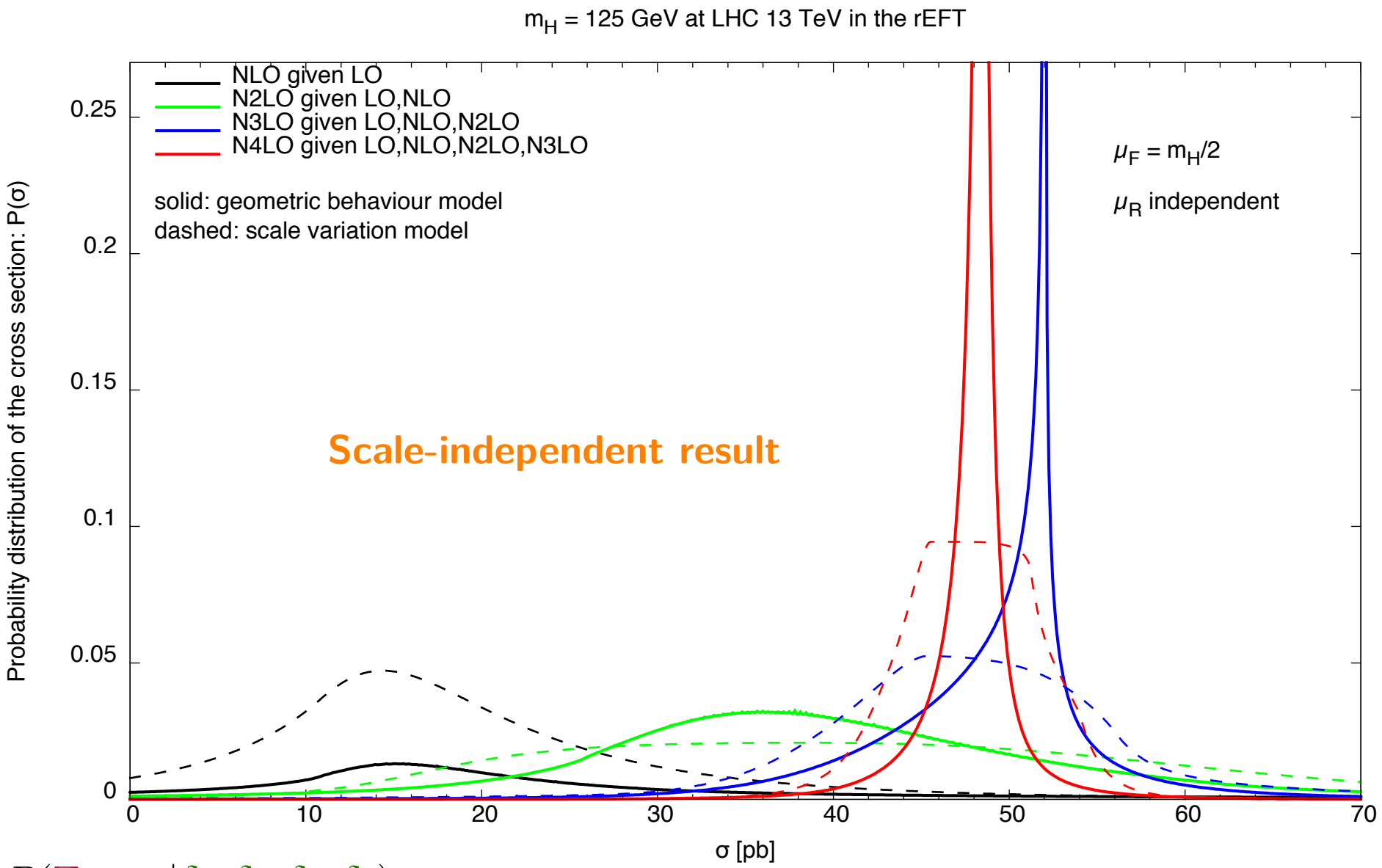
$$P(\Sigma_{\mathrm{N^4LO}}|\delta_0,\delta_1,\delta_2,\delta_3,\mu=m_H/2)$$

Higgs in gluon fusion at LHC: models at work



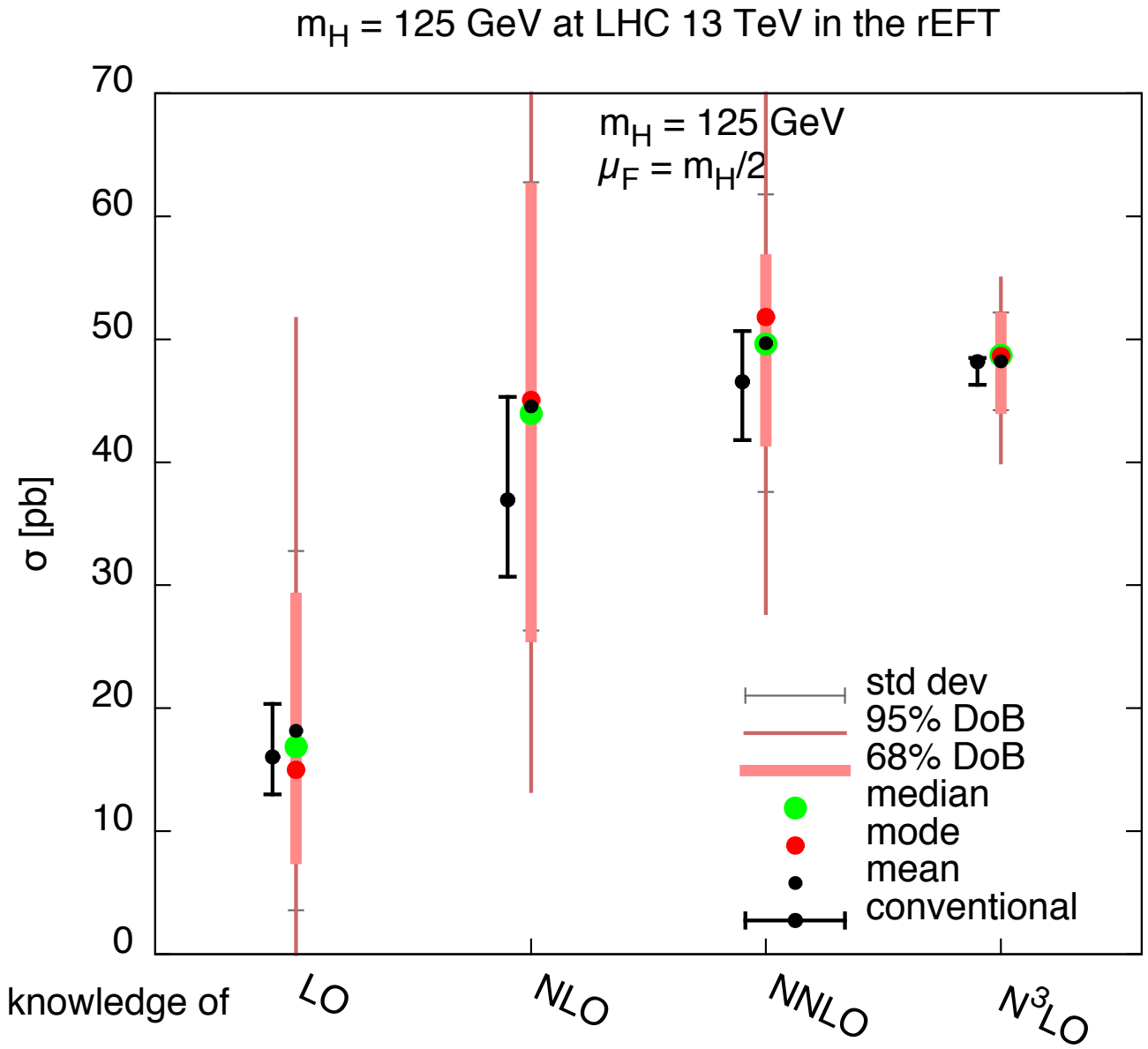
$$P(\Sigma_{\mathrm{N^4LO}}|\delta_0,\delta_1,\delta_2,\delta_3,\mu=m_H)$$

Higgs in gluon fusion at LHC: models at work



$$P(\Sigma_{\mathrm{N^4LO}}|\delta_0,\delta_1,\delta_2,\delta_3)$$

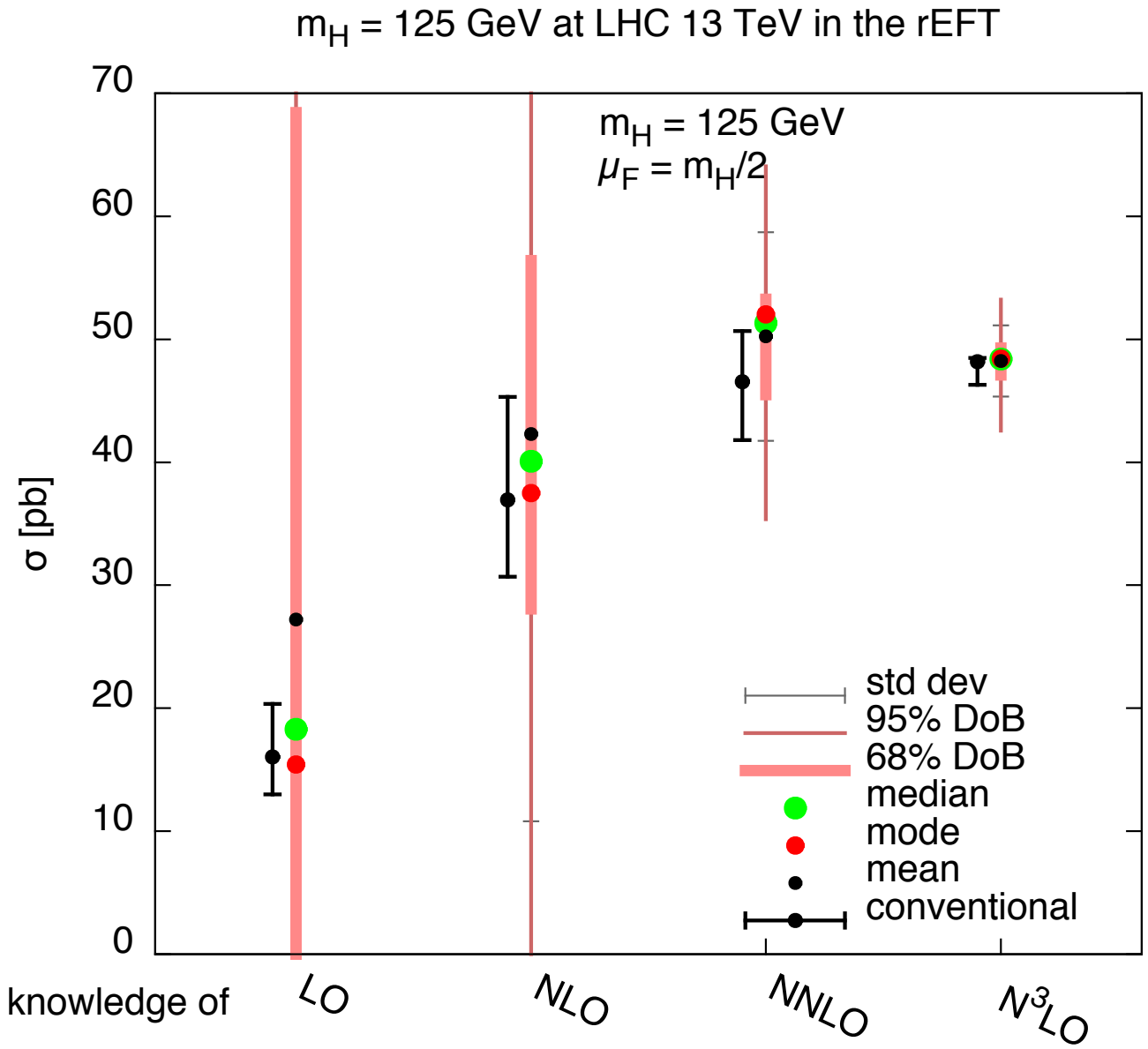
Higgs in gluon fusion at LHC: final results



conventional result: canonical scale variation by a factor of 2 about $\mu_R = m_H/2$ (best convergence properties)

new result: scale variation inspired model

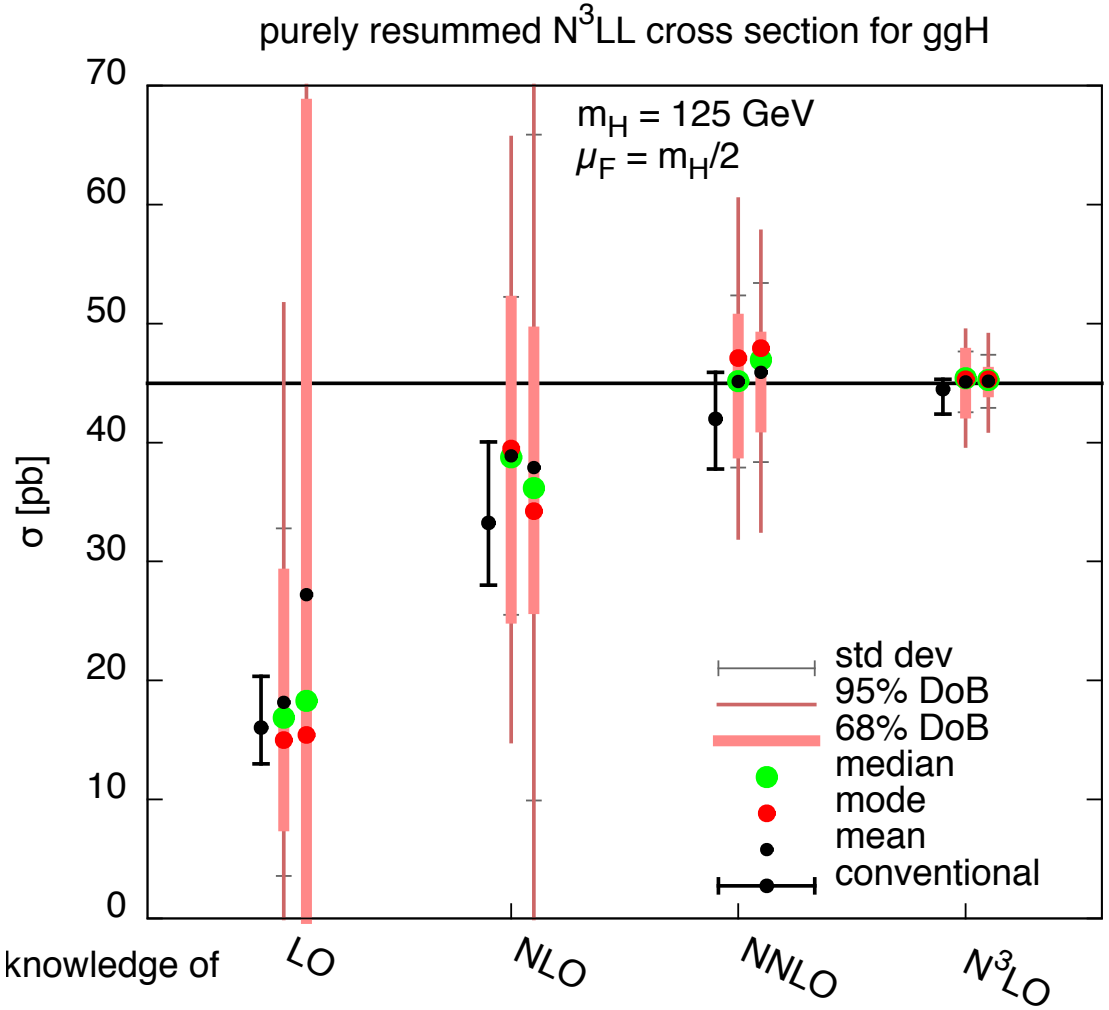
Higgs in gluon fusion at LHC: final results

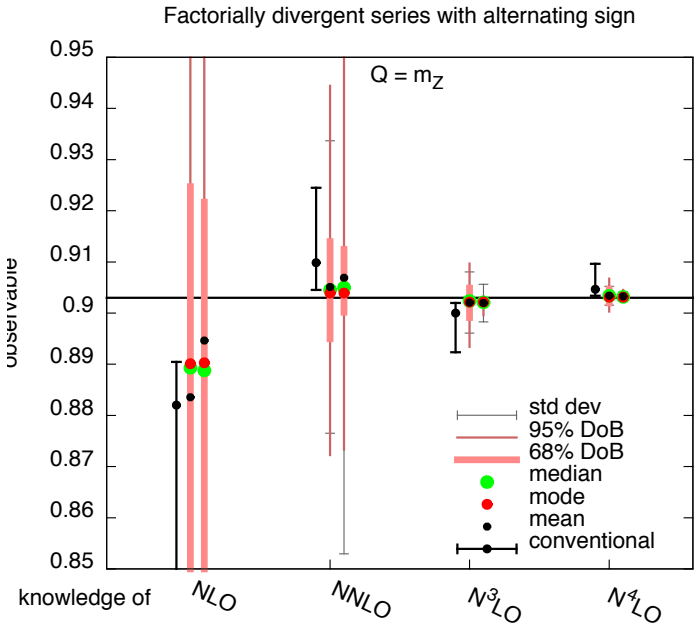


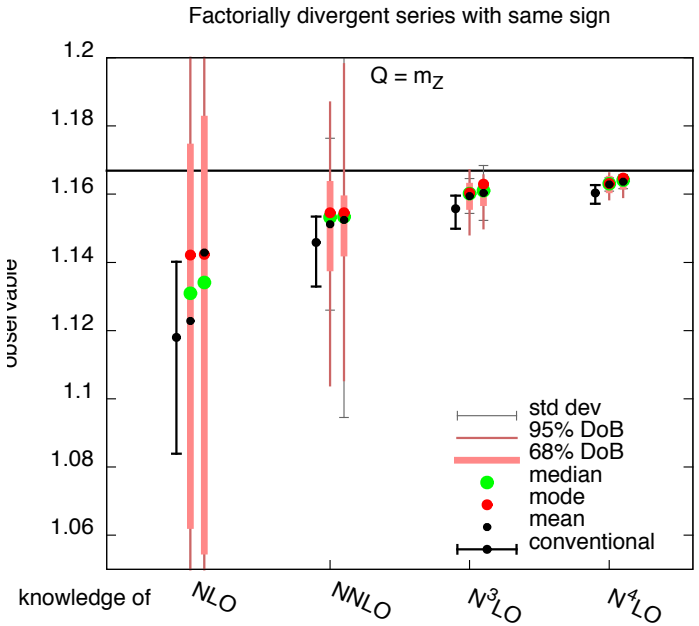
conventional result: canonical scale variation by a factor of 2 about $\mu_R = m_H/2$ (best convergence properties)

new result: geometric behaviour model

Validation using known sums







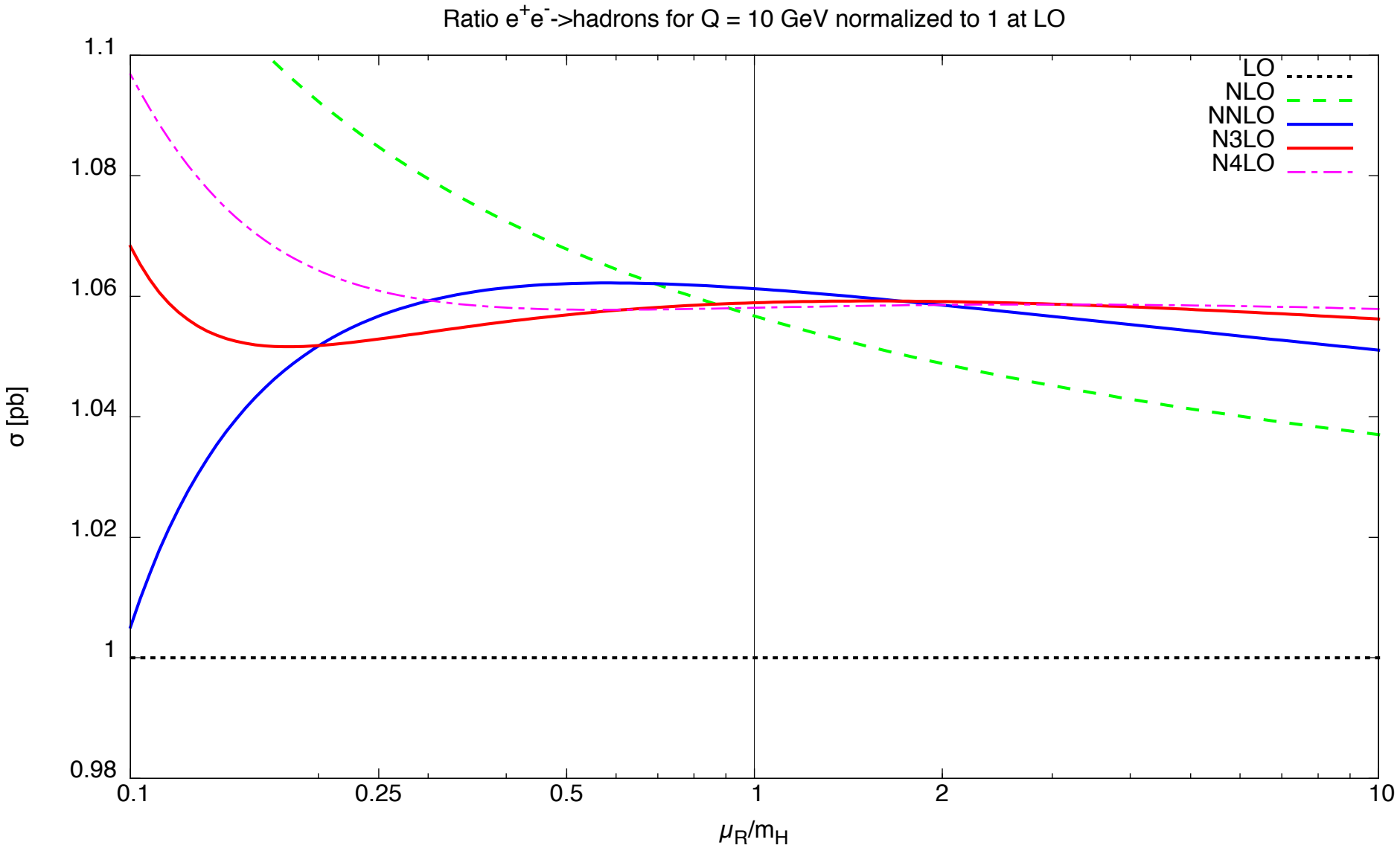
new results:
left: scale variation inspired model
right: geometric behaviour model

Summary

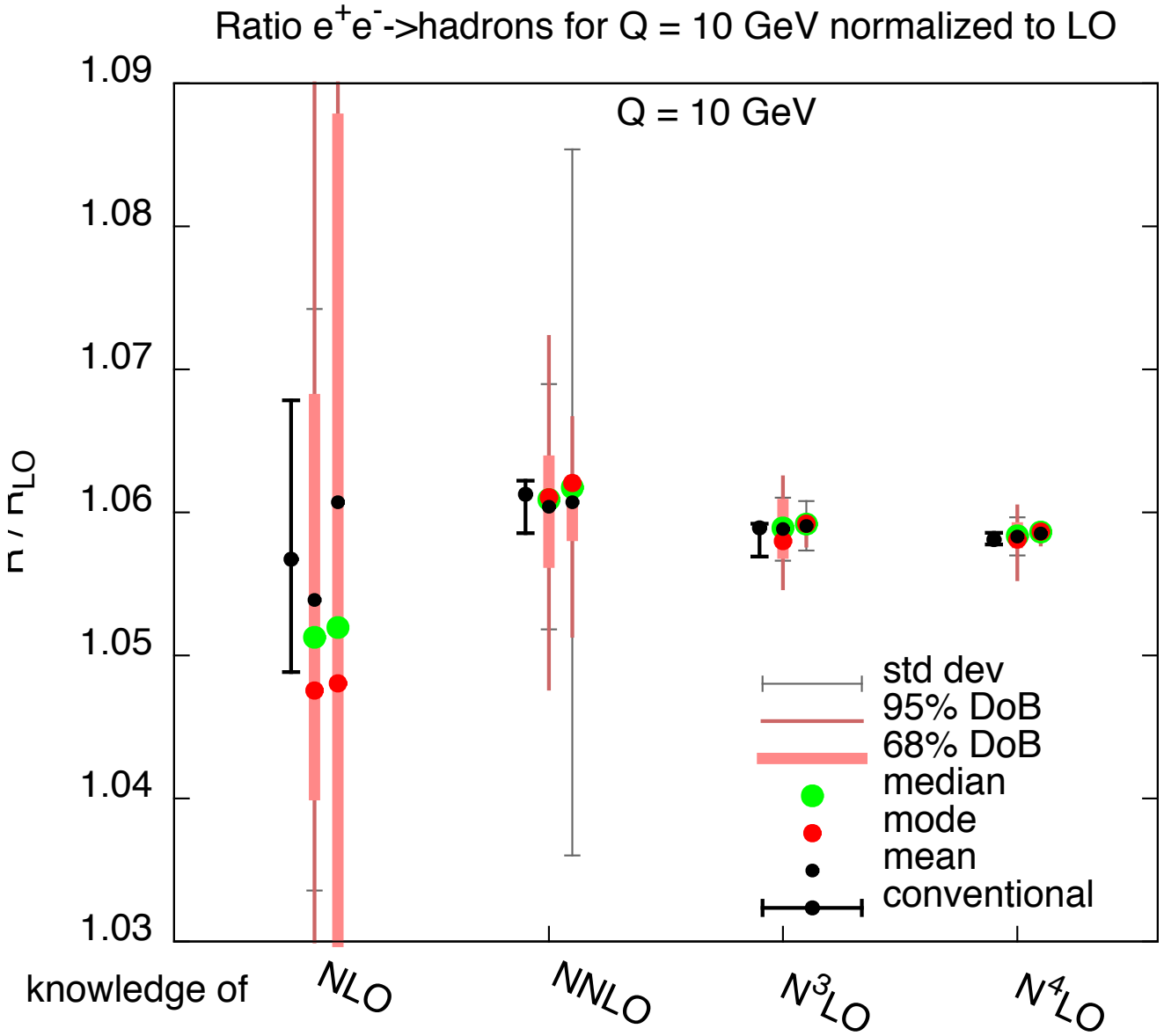
- Two new statistical models for theory uncertainties:
  - an improved version of Cacciari-Houdeau (geometric behaviour model)
  - a model inspired by scale variation
- A novel way to obtain **scale-independent results**
- Preliminary results very promising (geom. behaviour model somewhat better)
- What's next?
  - A systematic study of priors
  - Investigate potential improvements/combinations of the models
  - Release a **public code**
  - Consider **correlations** among different kinematic configurations
  - Applications

Backup slides

$e^+e^- \to$ hadrons: the "raw" result



$e^+e^- \to$ hadrons: final results



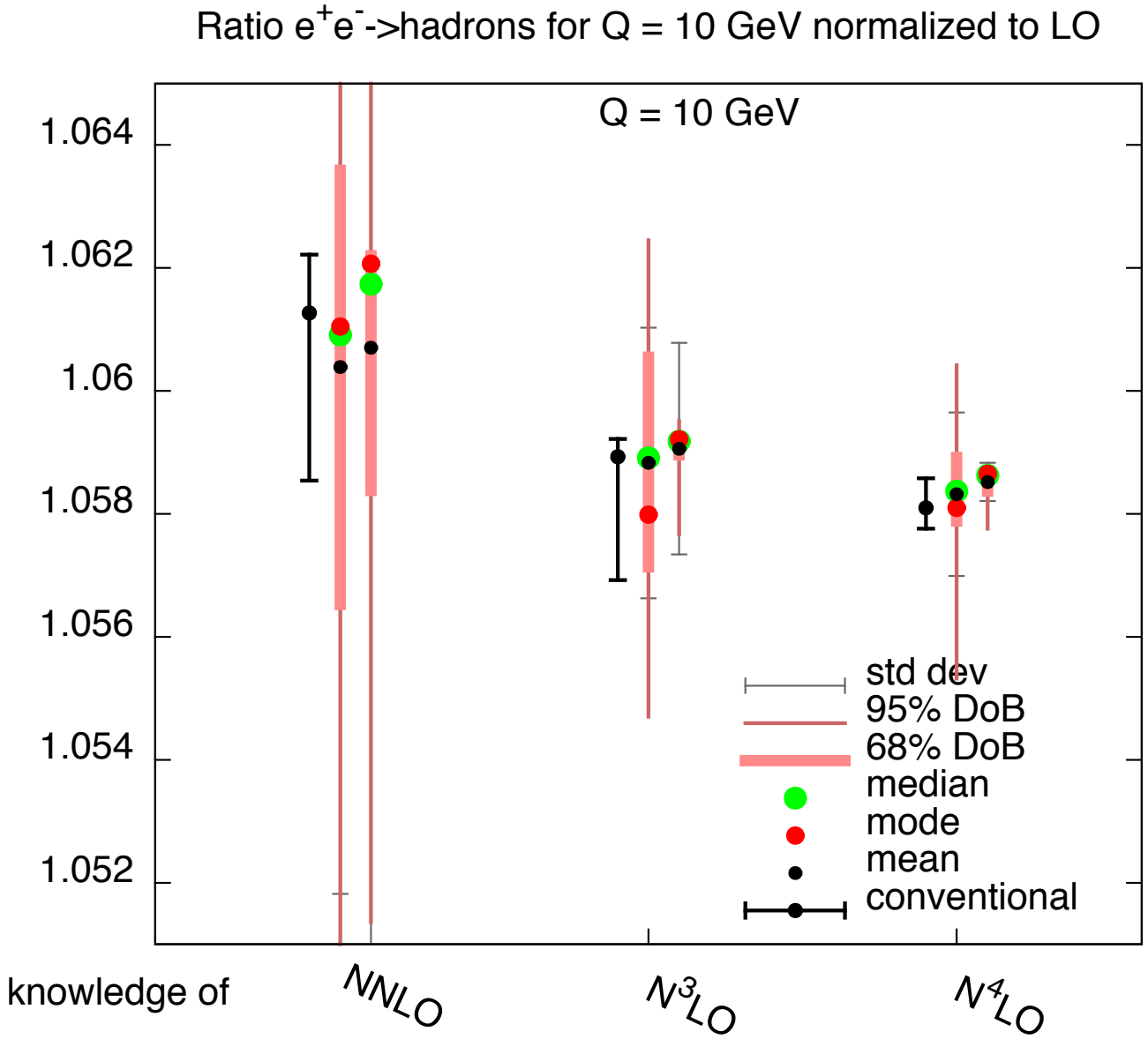
conventional result: canonical scale variation by a factor of 2 about $\mu_R = Q$

new results:

left: scale variation model

right: geometric behaviour model

$e^+e^- \to$ hadrons: final results



conventional result: canonical scale variation by a factor of 2 about $\mu_R = Q$

new results:

left: scale variation model

right: geometric behaviour model

The priors

- **geometric behaviour model**

$$P_0(c,a) = P_0(c)P_0(a)$$
$$P_0(c) = \frac{\epsilon}{c^{1+\epsilon}}\theta(c-1) \qquad \epsilon = 0.1$$
$$P_0(a) = 2(1-a)\theta(a)\theta(1-a)$$

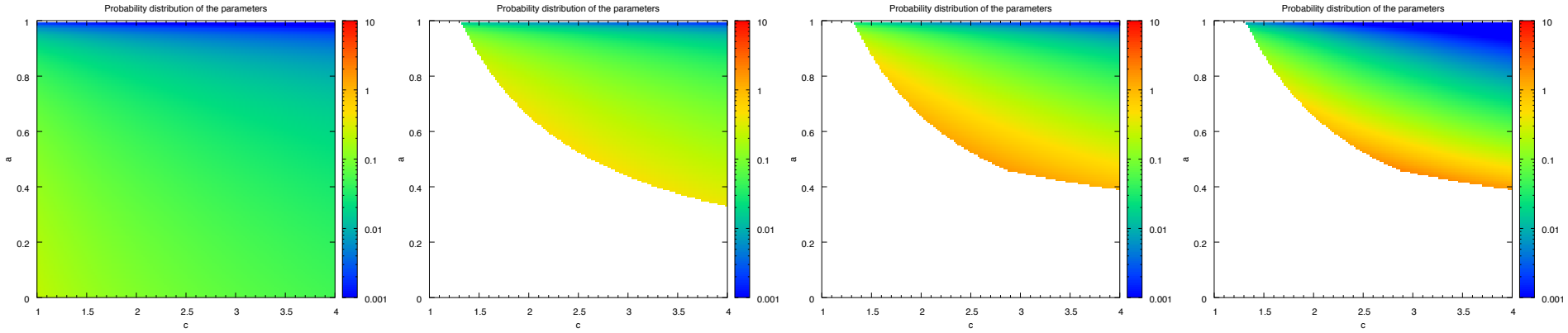
- **scale variation inspired model**

$$P_0(\lambda) = \lambda e^{-\lambda}$$

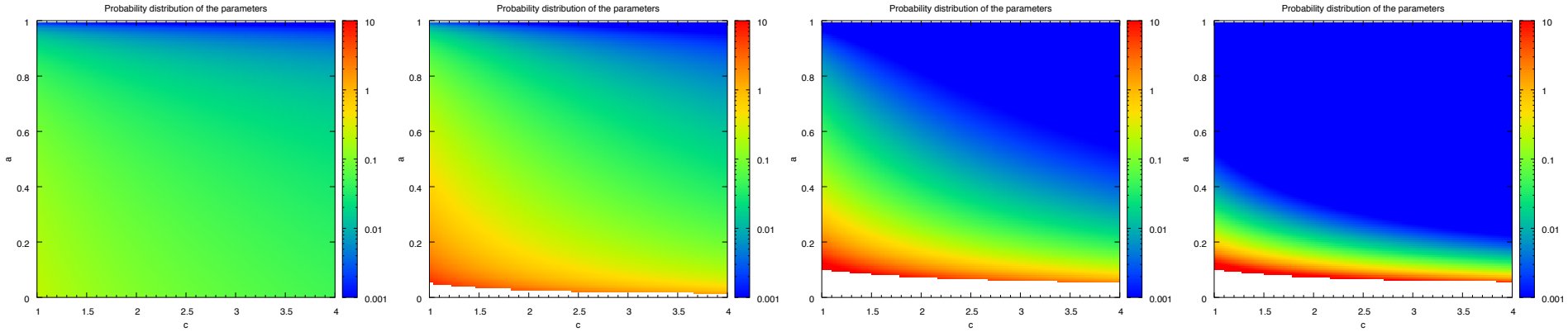
NB: These distributions assume the trivial knowledge of $\delta_0 = 1$. The first non-trivial information updating these distributions is δ_1

Posterior of the geometric-behaviour model parameters

ggH:

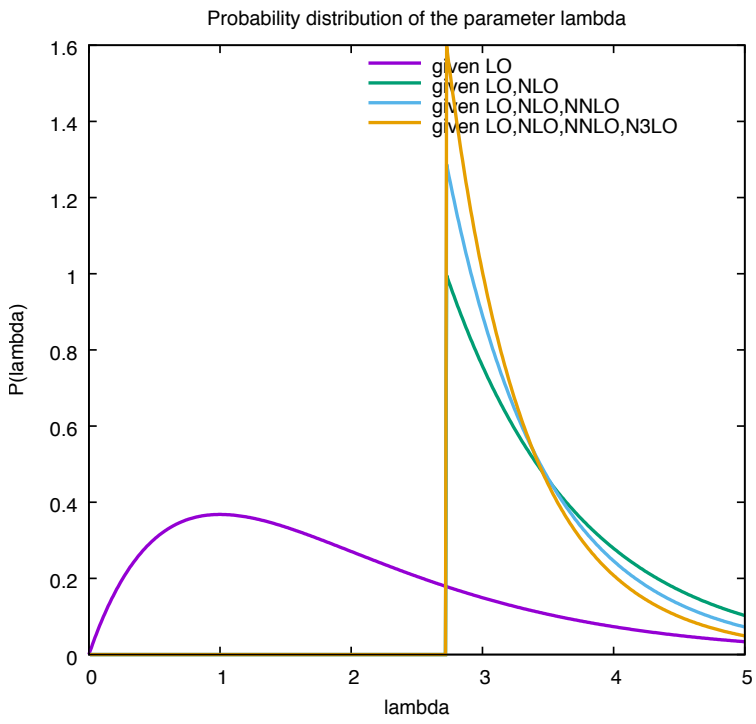


$e^+e^- \to$ hadrons:

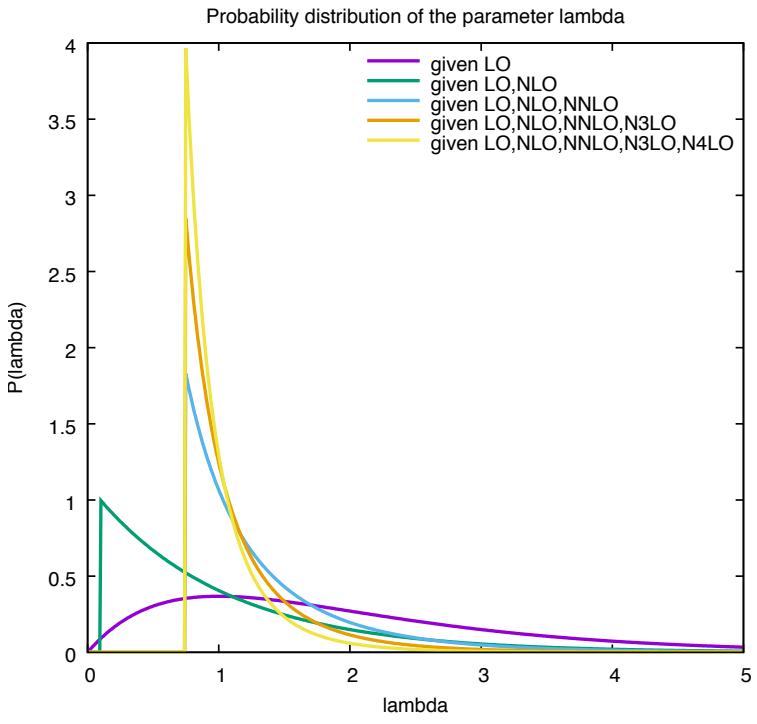


Posterior of the scale variation inspired model parameter

ggH:

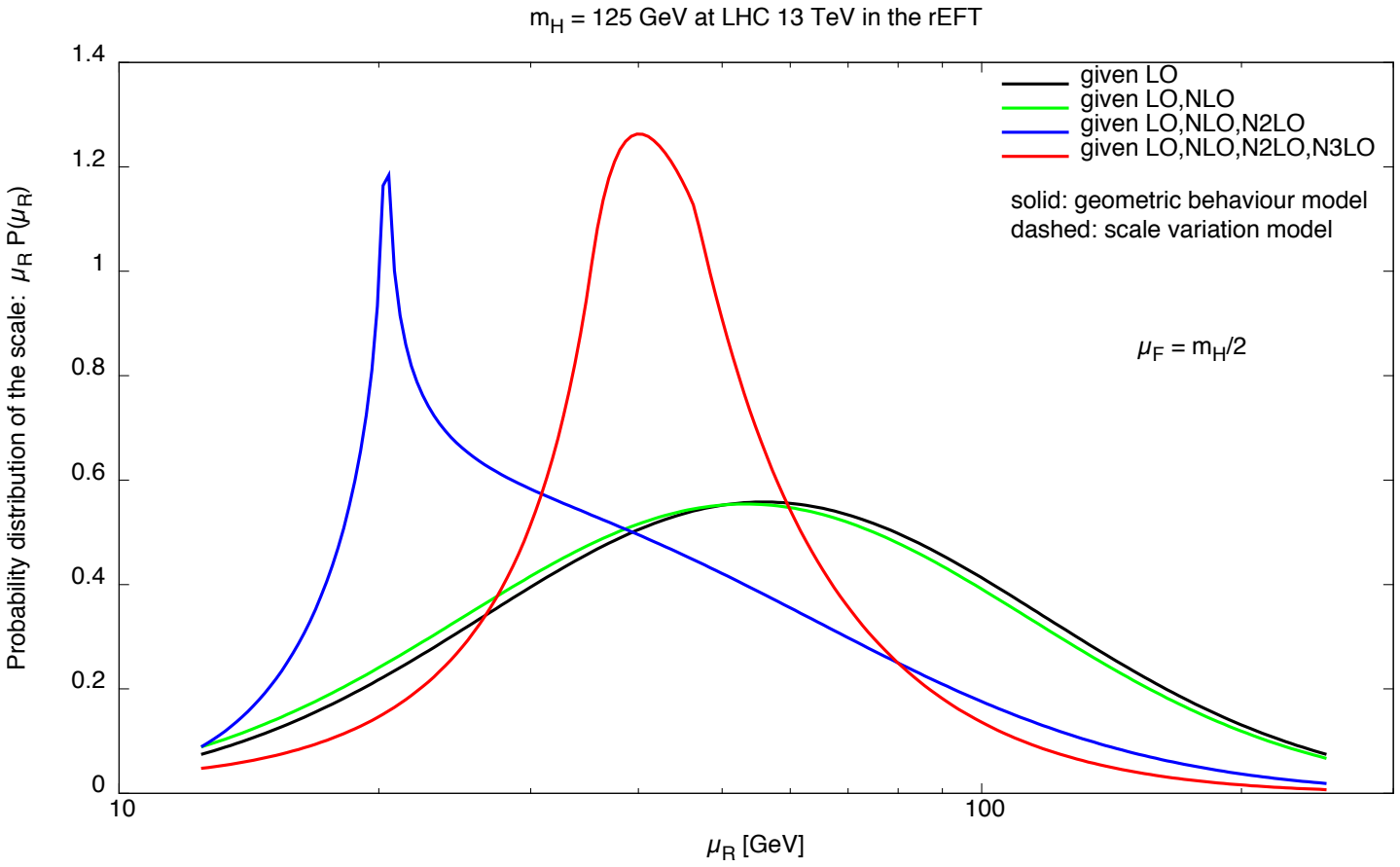


$e^+e^- \to$ hadrons:



Posterior of μ in ggH

Geometric-behaviour model



Support: $m_H/10 < \mu < 2m_H$