Improved Asymptotic Formulae

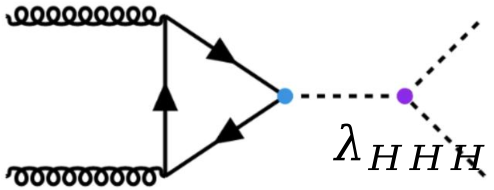
with an application to the Di-Higgs resonance search

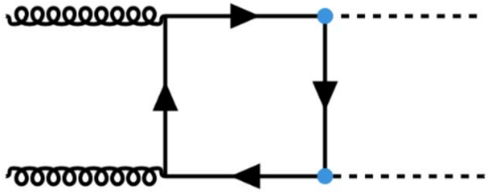
Ligang Xia
Nanjing University

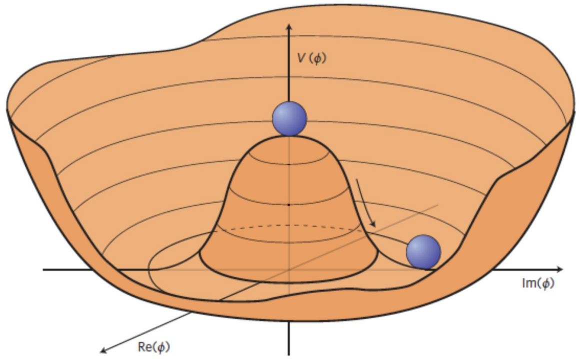
SYSU-PKU Collider Physics Forum for Young Scientists 5th,
June 8th, 2022

Di-Higgs production

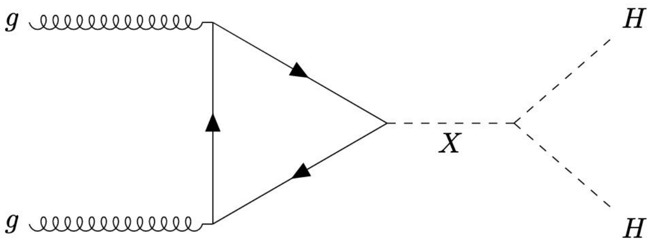
Non-resonant production of double higgs events   triple-higgs coupling strength, …

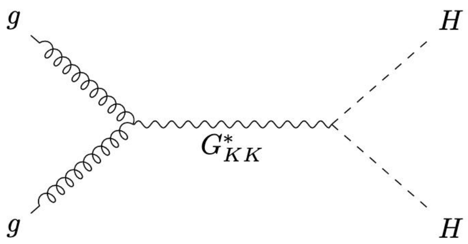






Resonant production of double higgs events   search for high-mass resonance, graviton, …

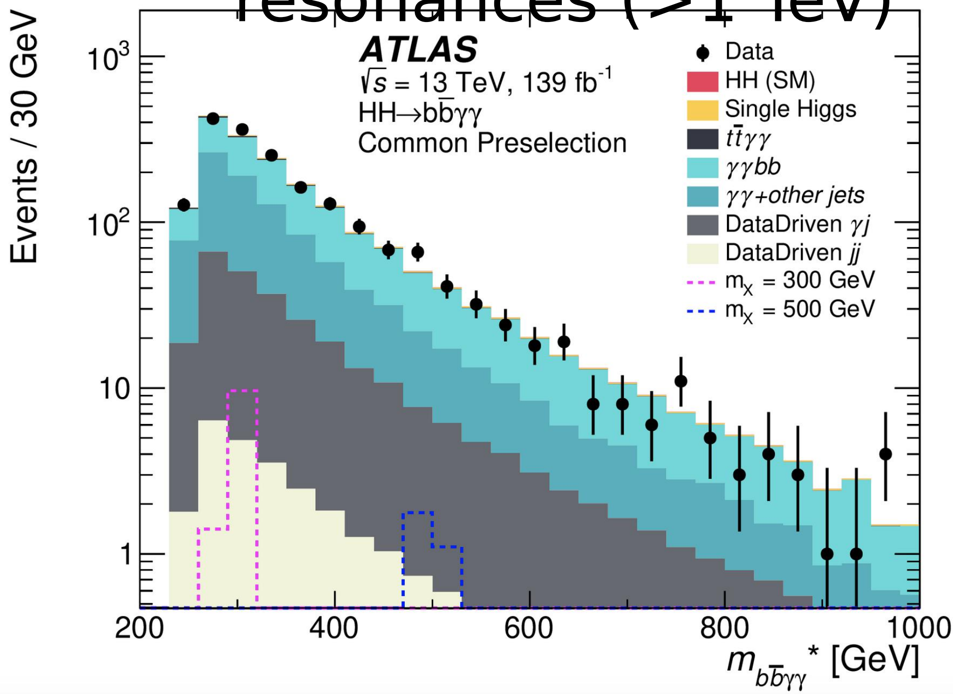


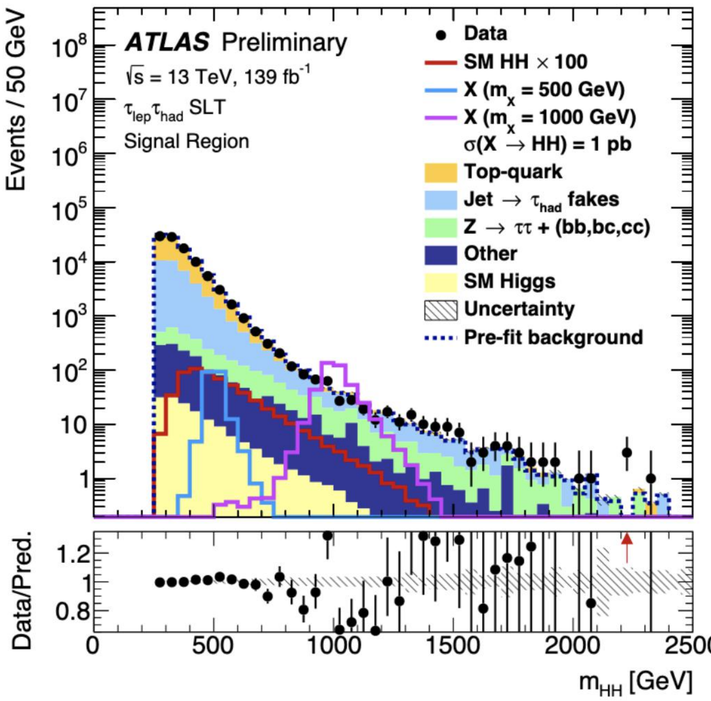


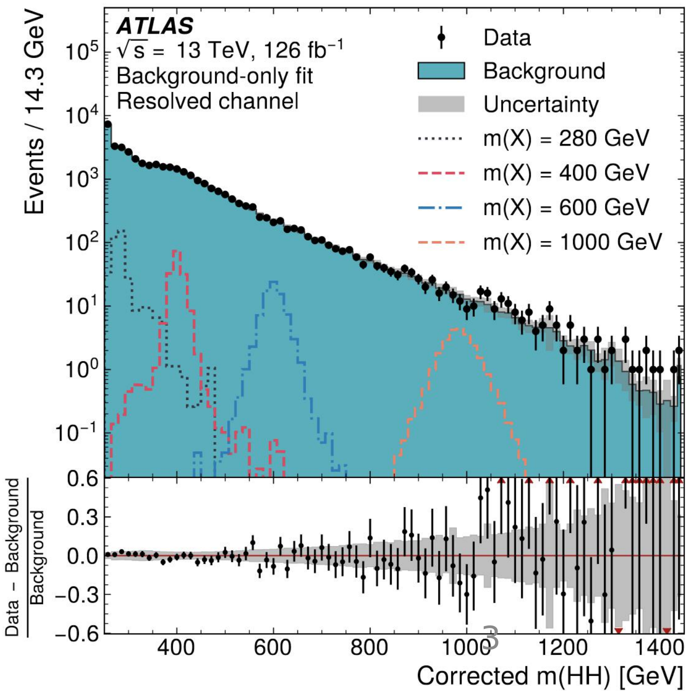
Di-Higgs resonance search in ATLAS

The search has been performed in several final states.

- (H bb)(H $\gamma\gamma$): low BR, clean, sensitive to low-mass resonances (< 1 TeV)
- (H bb)(H $\tau\tau$): sensitive to median-mass resonances
- (H bb)(H bb): high BR, dirty, sensitive to high-mass resonances (>1 TeV)



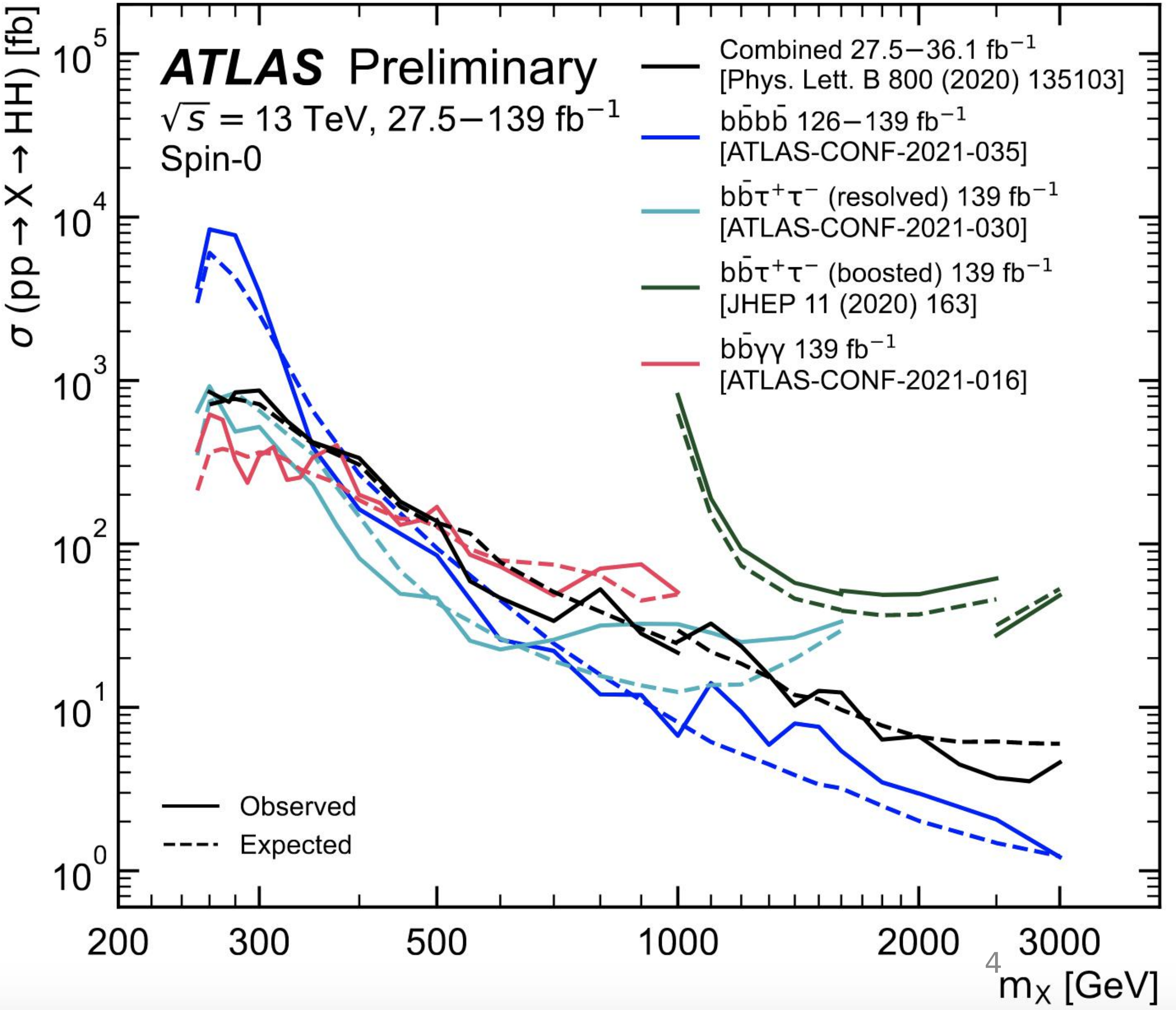




Di-Higgs resonance search in ATLAS

The search has been performed ir

- (H   bb)(H $\gamma\gamma$): sensitive to lo
- (H   bb)(H $\tau\tau$): sensitive to m
- (H   bb)(H   bb): sensitive to hi



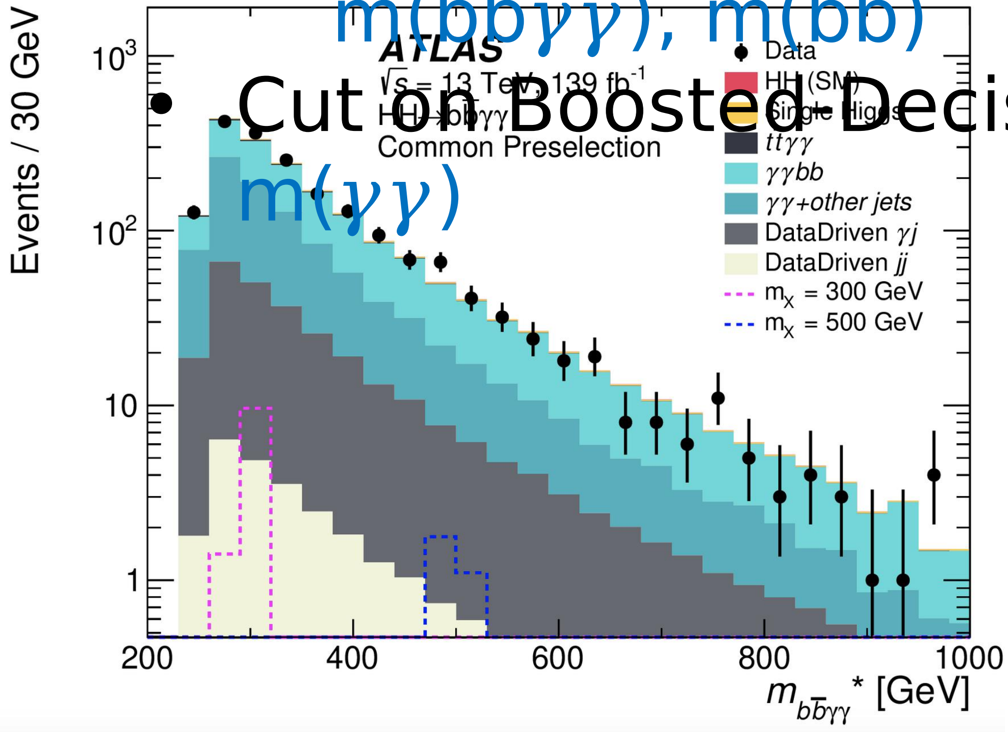
Current status:

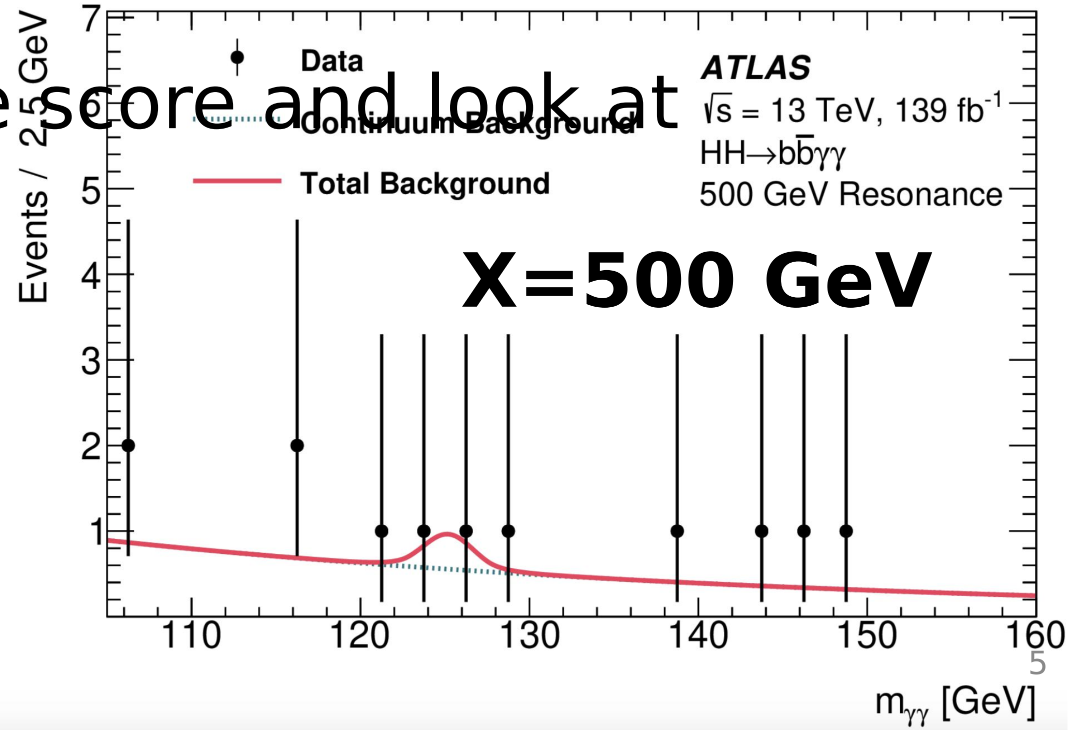
- Using full run2 dataset, no significant evidence for an HH resonance up to 3 TeV
- Upper limits on production cross section are provided.

Di-Higgs resonance search using bb $\gamma\gamma$

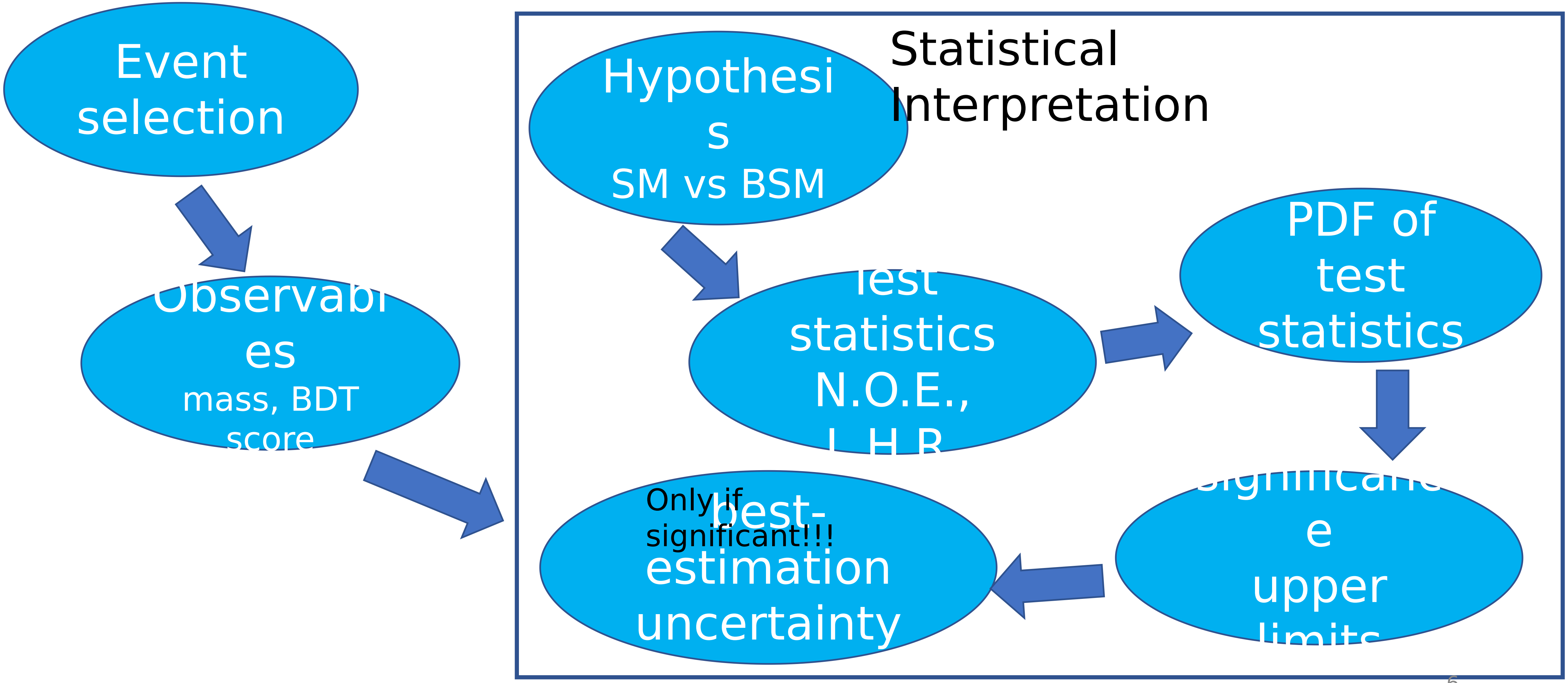
Search for HH resonance using (H bb, H $\gamma\gamma$) in ATLAS

- Multi-Variant-Analysis-based event selection:
  - training on many observables including m(bb$\gamma\gamma$), m(bb)
- Cut on Boosted-Decision-Tree score and look at m($\gamma\gamma$)

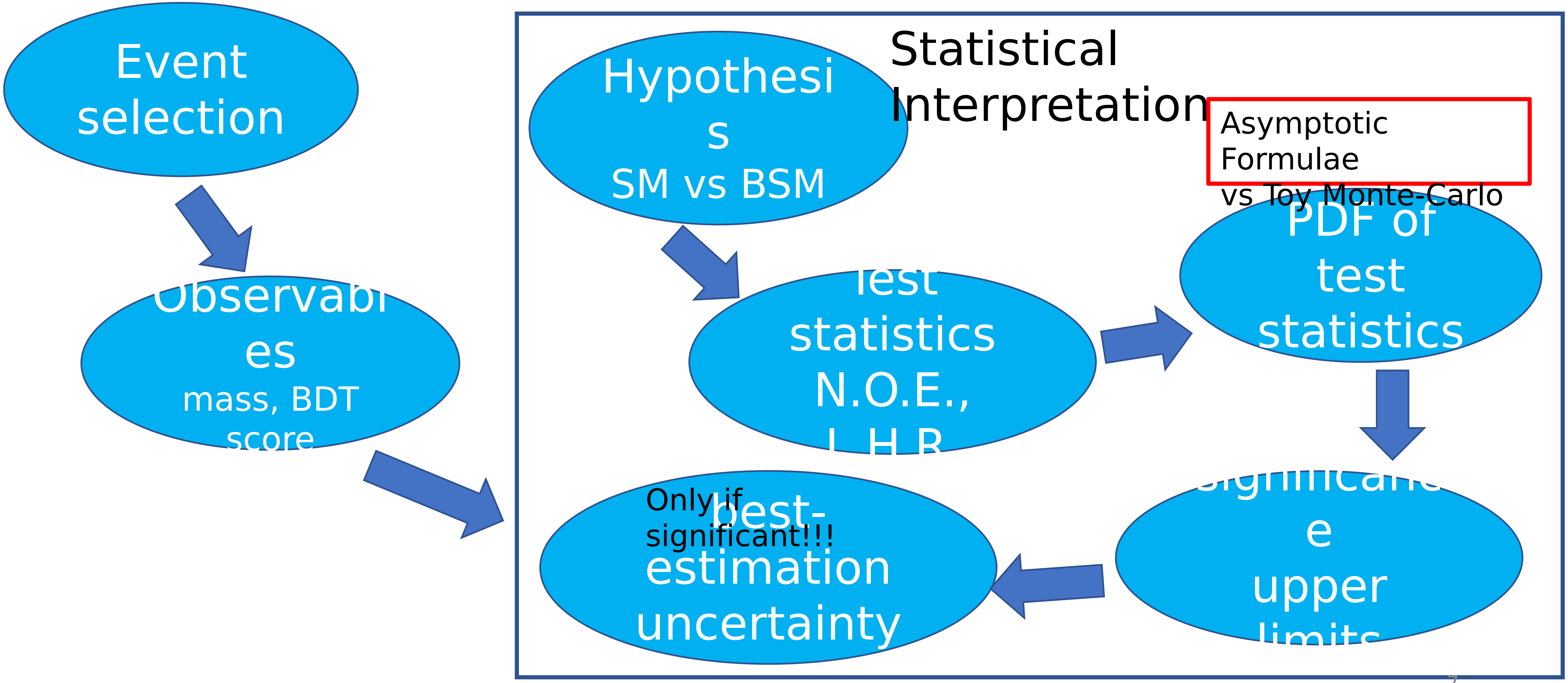




Statistical Interpretation



Statistical Interpretation



Summary of 6 Test statistics

Test statistic	Note
t_0	To establish the discovery of a signal
q_0	To establish the discovery of a positive signal
t_{μ}	To set an interval at a given confidence level
$\tilde{t}_{\mu}$	To set an interval for a positive signal at a given confidence level
q_{μ}	To set an upper limit of a signal at a given confidence level
$\tilde{q}_{\mu}$	To set an upper limit of a positive signal at a given confidence level

Classical Asymptotic Formulae

Eur. Phys. J. C (2011) 71: 1554
DOI 10.1140/epjc/s10052-011-1554-0

THE EUROPEAN PHYSICAL JOURNAL C

Special Article - Tools for Experiment and Theory

arXiv:1007.1727

Asymptotic formulae for likelihood-based tests of new physics

Glen Cowan[1], **Kyle Cranmer**[2], **Eilam Gross**[3], **Ofer Vitells**[3,a]

[1]Physics Department, Royal Holloway, University of London, Egham TW20 0EX, UK
[2]Physics Department, New York University, New York, NY 1
[3]Weizmann Institute of Science, Rehovot 76100, Israel

TESTS OF STATISTICAL HYPOTHESES CONCERNING SEVERAL PARAMETERS WHEN THE NUMBER OF OBSERVATIONS IS LARGE(1)

BY
ABRAHAM WALD

TABLE OF CONTENTS

1. Introduction 426
2. Assumptions on the density function $f(x, \theta)$ 428

Likelihood ratio in a measurement

Signal region:
sensitive to Parameter Of Interest ($\boldsymbol{\mu}$)

Control region:
sensitive to Nuisance Parameters (θ)

The likelihood function is the product of Poisson probabilities for all bins:

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} e^{-u_k} .$$

To test a hypothesized value of μ we consider the profile likelihood ratio

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} .$$

Most powerful test (Neyman-Pearson

The $\boldsymbol{\mu}$ value we want to test

Optimal estimate of $\boldsymbol{\mu}$

Wald's theorem

The likelihood function is the product of Poisson probabilities for all bins:

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} e^{-u_k} .$$

To test a hypothesized value of μ we consider the profile likelihood ratio

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} .$$

482 ABRAHAM WALD

same limit distribution. The limit distribution of $-2 \log \lambda_n$ is the χ^2-distribution with r degrees of freedom if the hypothesis to be tested is true. If the true parameter point θ_n is not an element of ω, the distribution of $-2 \log \lambda_n$ approaches the distribution of a sum of non-central squares

$$U^2 = u_1^2 + \cdots + u_r^2,$$

where the variates $u_1, \cdots, u_r$ are independently and normal
with unit variances and

$$\sum_{p=1}^{r} (Eu_p)^2 = n \sum$$

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

Can we go one small step further?

$$t_\mu \equiv -2 \ln \lambda(\mu) = \frac{(\hat{\mu} - \mu)^2}{\sigma^2} + \mathcal{O}(\frac{1}{\sqrt{N}}) \qquad (17)$$

Here $\hat{\mu}$ follows a Gaussian distribution with a mean μ' and standard deviation σ, and N represents the data sample size. The standard deviation σ of $\hat{\mu}$ is obtained from the covariance matrix of the estimators for all the parameters, $V_{ij} = \mathrm{cov}[\hat{\theta}_i, \hat{\theta}_j]$, where here the θ_i represent

Example: $\widetilde{q_\mu}$

$$\tilde{q}_\mu = \begin{cases} 0 & \hat{\mu} > \mu \,, \\ -2\ln \frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(\hat{\mu},\hat{\boldsymbol{\theta}})} & \mu \geq \hat{\mu} \geq 0 \\ -2\ln \frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \,. \end{cases}$$

$\widehat{\mu} > \mu$: not inconsistent with upper limit mu
$\widehat{\mu} < 0$: assume non-negative signal strength and we compare with 0

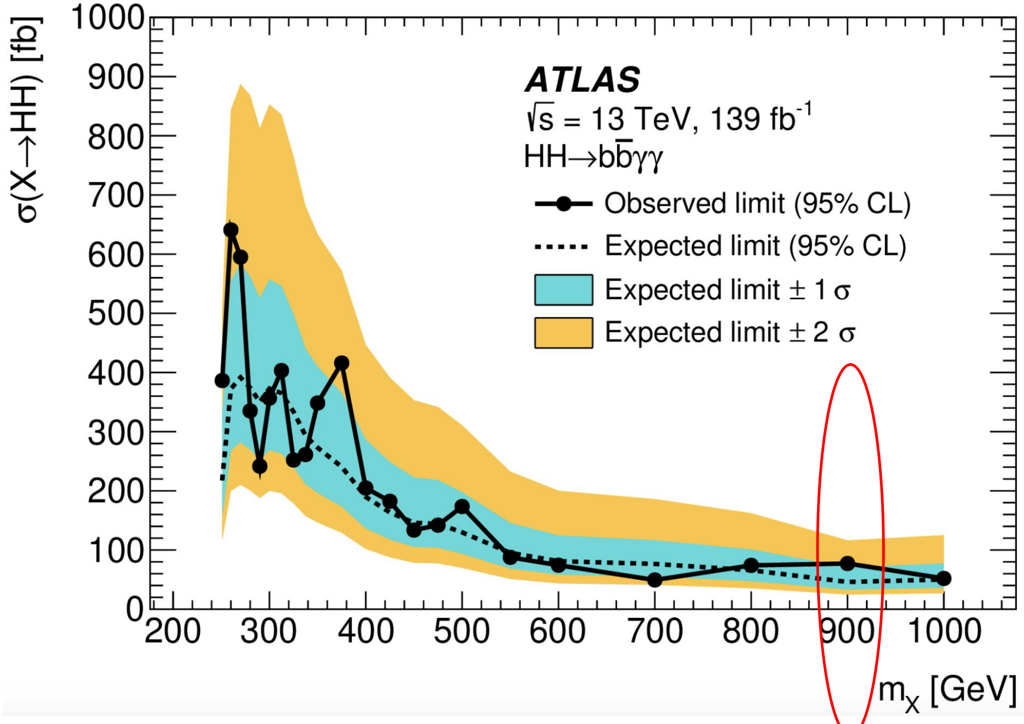
$$\tilde{q}_\mu = \begin{cases} 0 & \hat{\mu} > \mu \,, \\ \frac{(\hat{\mu}-\mu)^2}{\sigma^2} & \mu \geq \hat{\mu} \geq 0 \,, \\ \frac{\mu^2-2\mu\hat{\mu}}{\sigma^2} & \hat{\mu} < 0 \,. \end{cases}$$

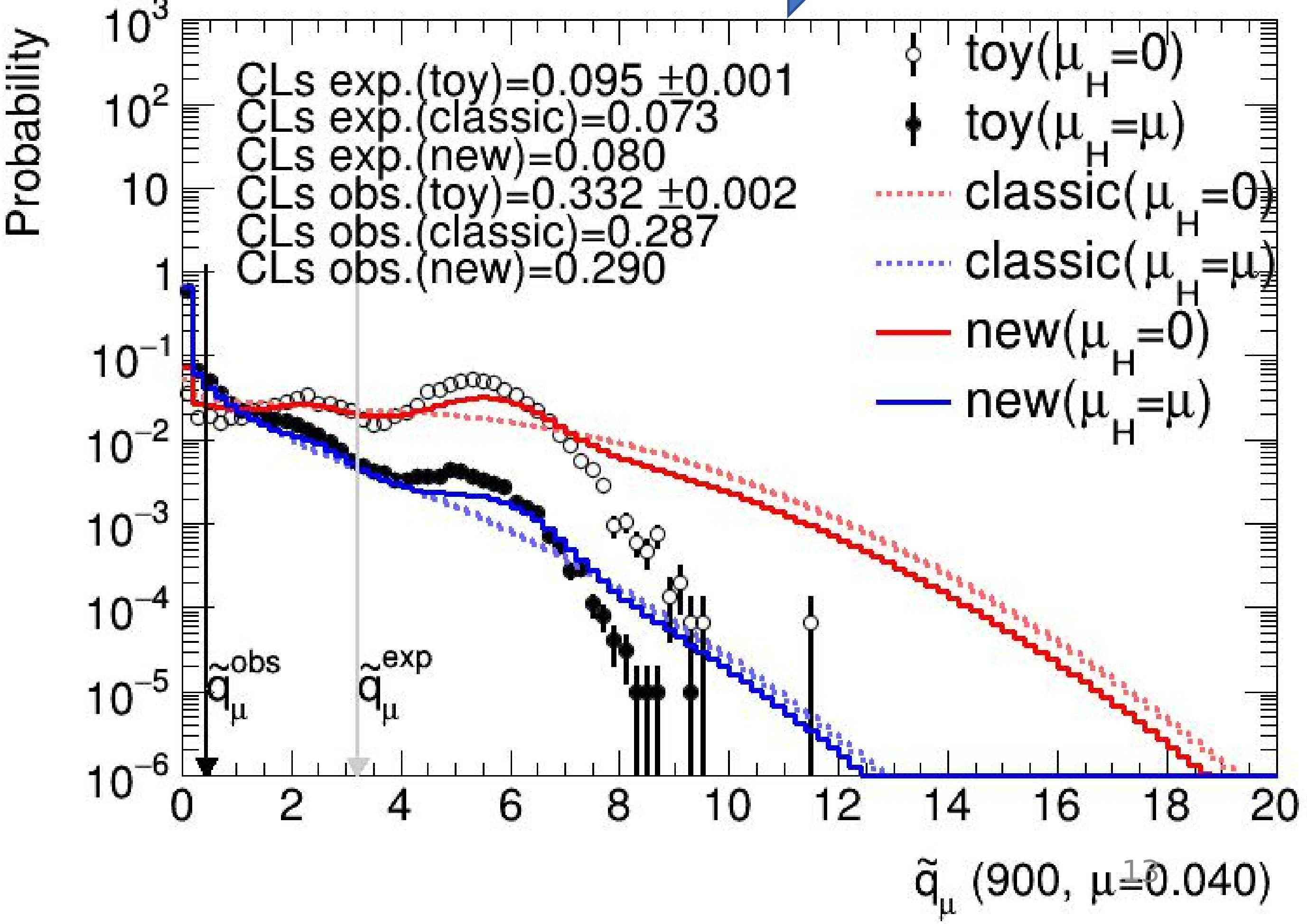
Classic asymptotic formulae

$$f(\tilde{q}_\mu|\mu') = \Phi\left(\frac{\mu'-\mu}{\sigma}\right)\delta(\tilde{q}_\mu) + \begin{cases} \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{q}_\mu}}\exp\left[-\frac{1}{2}\left(\sqrt{\tilde{q}_\mu}-\frac{\mu-\mu'}{\sigma}\right)^2\right] & 0 < \tilde{q}_\mu \leq \mu^2/\sigma^2 \,, \\ \frac{1}{\sqrt{2\pi}(2\mu/\sigma)}\exp\left[-\frac{1}{2}\frac{(\tilde{q}_\mu-(\mu^2-2\mu\mu')/\sigma^2)^2}{(2\mu/\sigma)^2}\right] & \tilde{q}_\mu > \mu^2/\sigma^2 \,. \end{cases}$$

Example: X(900GeV) HH$\gamma\gamma$ bb

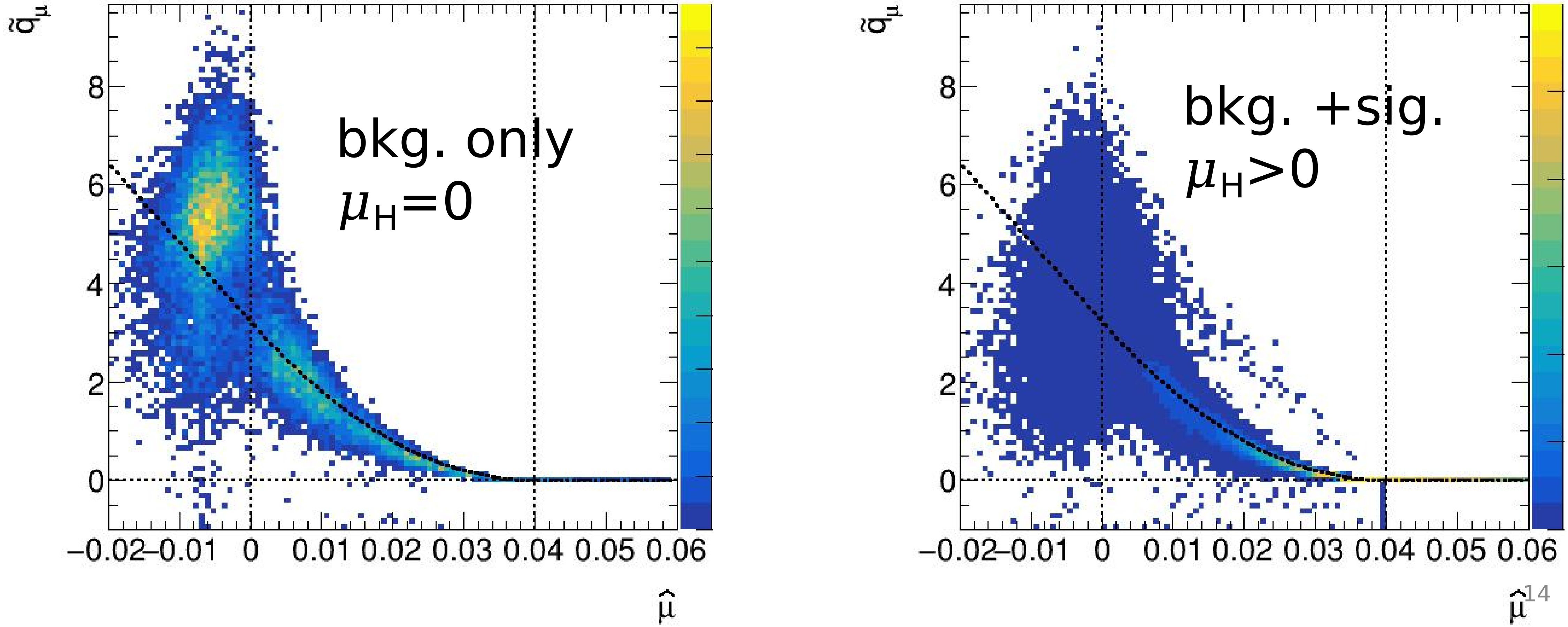
Signal like ← → Background like





Why “bumps” in PDF of

$$\tilde{q}_\mu = \begin{cases} 0 & \hat{\mu} > \mu\,, \\ \frac{(\hat{\mu}-\mu)^2}{\sigma^2} & \mu \geq \hat{\mu} \geq 0\,, \\ \frac{\mu^2 - 2\mu\hat{\mu}}{\sigma^2} & \hat{\mu} < 0\,. \end{cases}$$

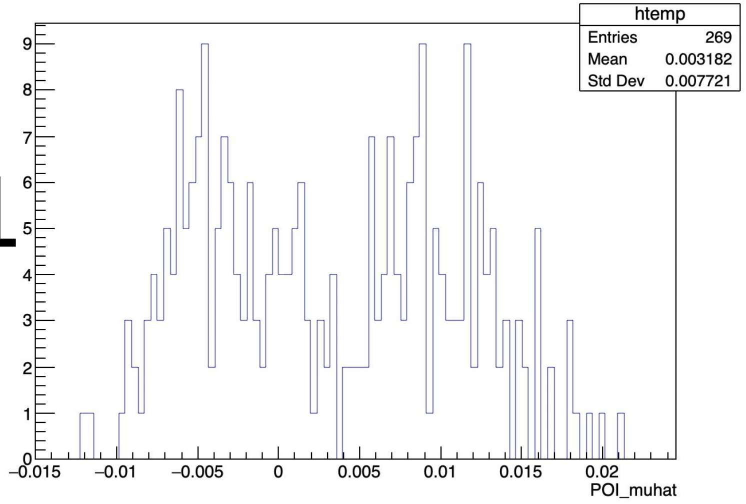


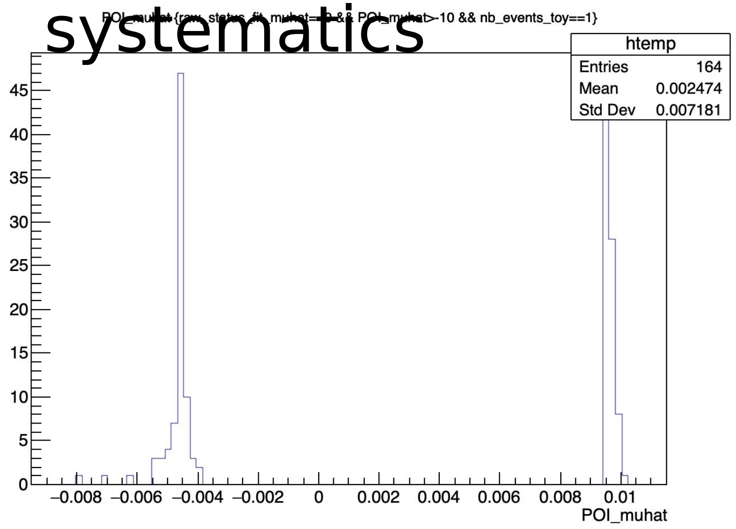
Why “bumps” in PDF of $\widetilde{q_{\mu}}$?

switch off systematics

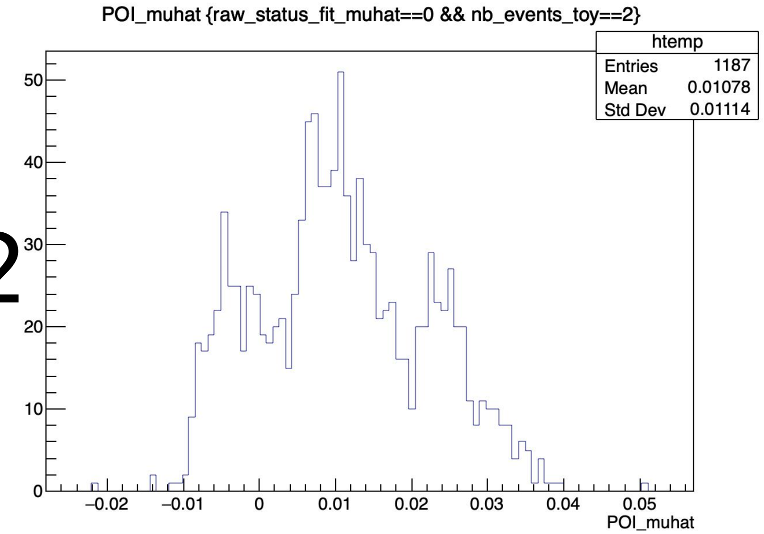
Things are clear if looking at the $\hat{\mu}$ distribution in **toys** for the number of events = 1, 2, …

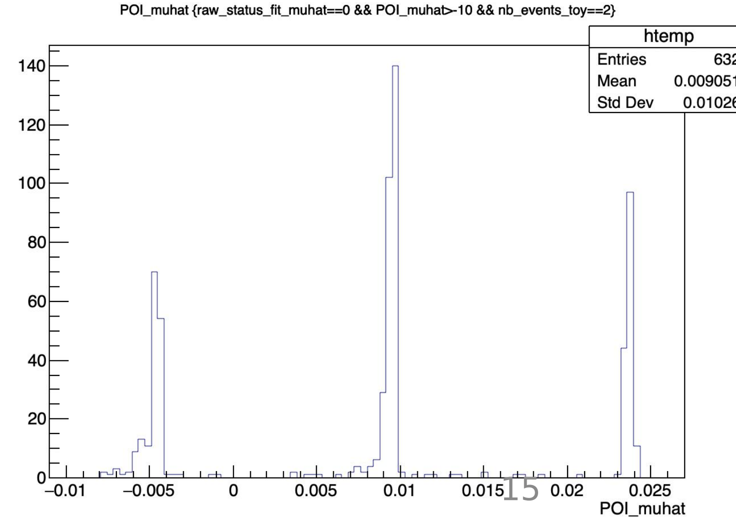
n=1



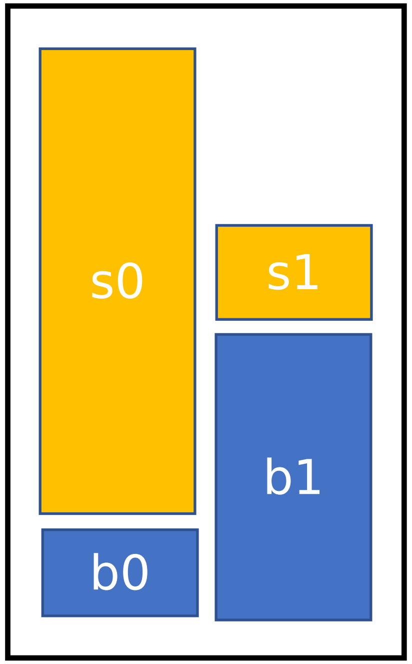


n=2





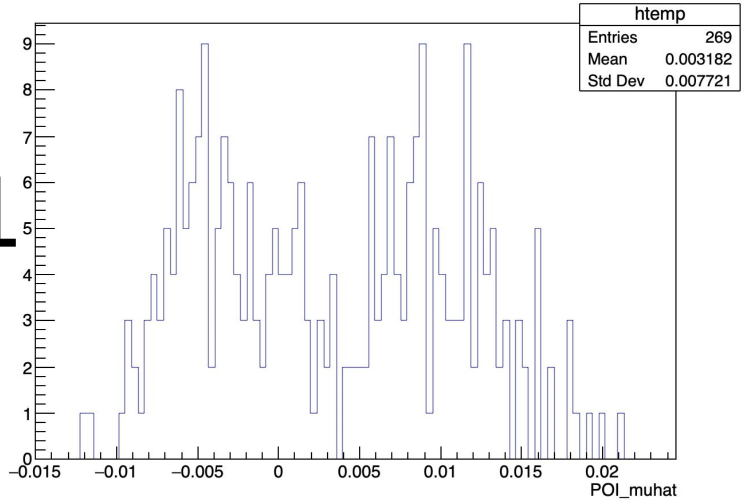
Why “bumps” in PDF of $\widetilde{q_\mu}$?

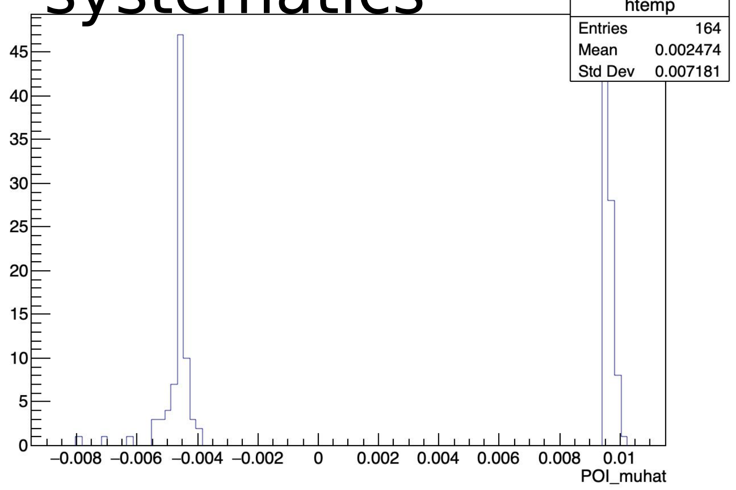


supposing a **2-bin** observable distribution

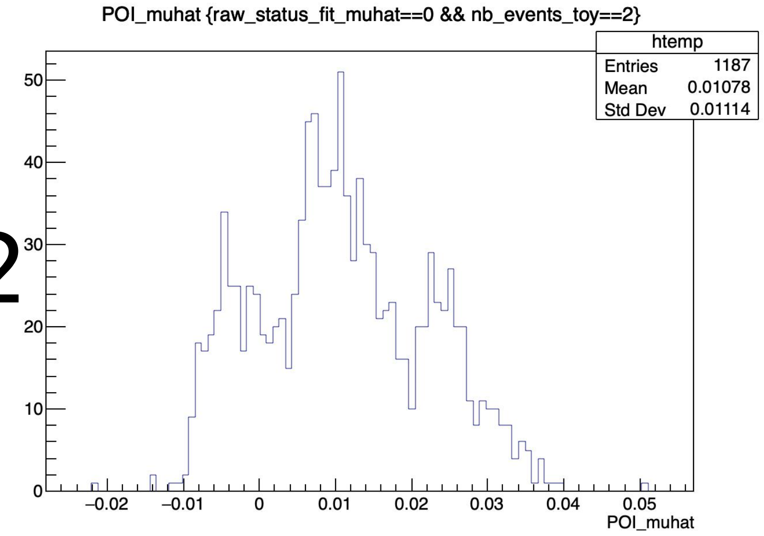
switch off systematics

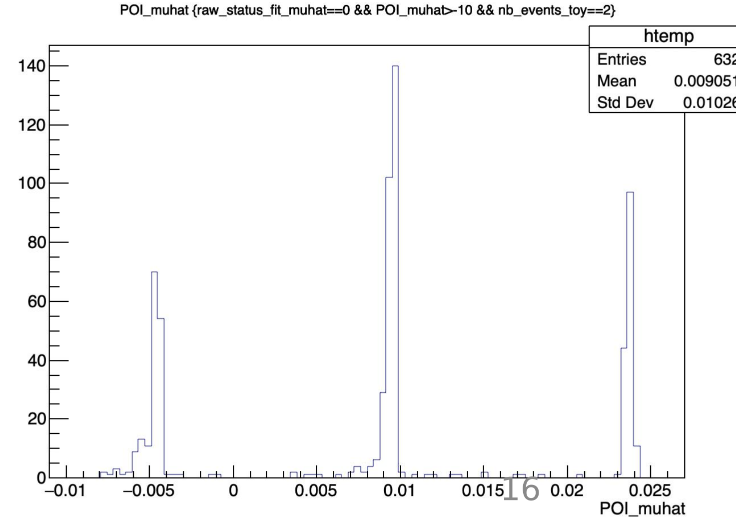
n=1





n=2

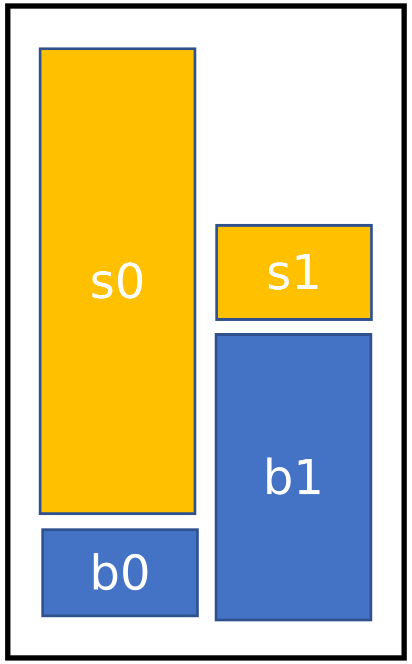




Why “bumps” in PDF of $\tilde{\mu}$?

$$\mathcal{L} = \Pi_{i=0}^{N} P(n_i | b_i + \mu s_i)$$

$$\log \mathcal{L} = \sum_{i=0}^{N} n_i \log(b_i + \mu s_i) - (b_i + \mu s_i)$$



supposing a **2-bin** observable distribution

$$\partial \log \mathcal{L} / \partial \mu = 0$$

$$A = 2 s_0 s_1$$

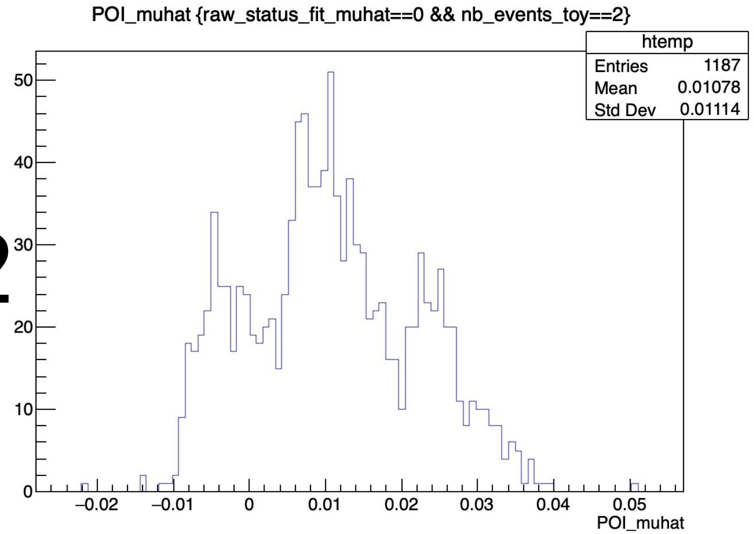
$$B = s_0 b_1 + s_1 b_0 - \frac{n_0 + n_1}{s_0 + s_1} s_0 s_1$$

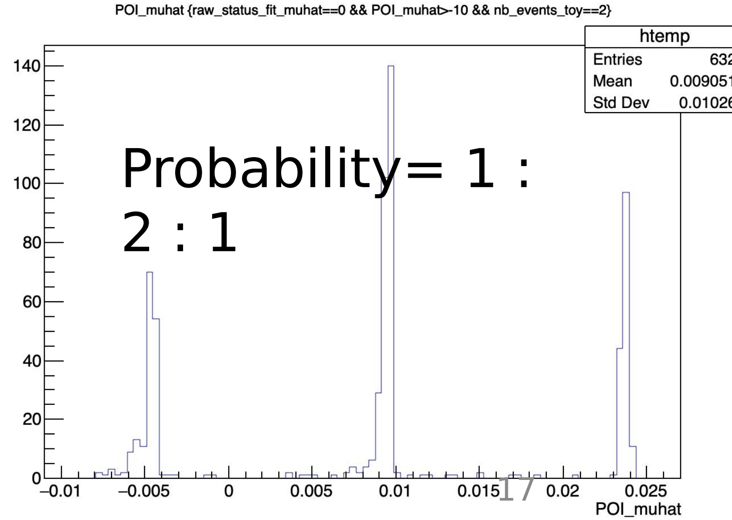
$$C = b_0 b_1 - \frac{n_0 s_0 b_1 + n_1 s_1 b_0}{s_0 + s_1}$$

$$\hat{\mu}(n_0, n_1) = \frac{-B + \sqrt{B^2 - 4AC}}{2A}$$

$$\hat{\mu}(k, n-k) \approx -\frac{b_0}{s_0} + \frac{k}{s_0 + s_1}$$

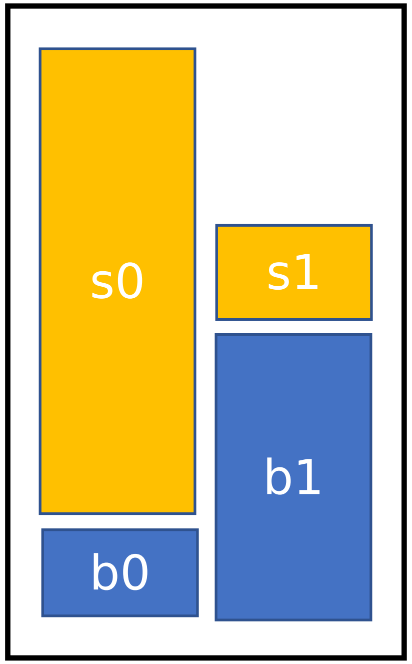
n=2





Probability= 1 : 2 : 1

New Asymptotic Formulae



- supposing a **6-bin** observable distribution
- Supposing $n_{small}=5$

$$f(T_\mu|\mu_H) = \sum_{n=0}^{+\infty} f(T_\mu|n,\mu_H)P(n|b+\mu_H s)$$
$$= \sum_{n=0}^{n_{\rm small}} f(T_\mu|n,\mu_H)P(n|b+\mu_H s) + \sum_{n>n_{\rm small}} f(T_\mu|n,\mu_H)P(n|b+\mu_H s)$$
$$\approx \sum_{n=0}^{n_{\rm small}} f_{\rm SS}(T_\mu|n,\mu_H)P(n|b+\mu_H s) + (1-\sum_{n=0}^{n_{\rm small}} P(n|b+\mu_H s))f_{\rm LS}(T_\mu|n,\mu_H)$$

n_{small}: threshold between large statistics and small statistics

Small-Statistics part:
- discrete
- analytical/numerical calculation

Large-Statistics part:
- continual
- classical asymptotic formulae

$$f_{\rm SS}(T_\mu|n,\mu_H) = \sum_{k_0+k_1+k_2+k_3+k_4+k_5=n} \frac{n!}{k_0!k_1!\cdots k_5!}\left(\frac{b_5}{b+\mu_H s}\right)^{k_5}\Pi_{i=0}^{4}\left(\frac{b_i+\mu_H s_i}{b+\mu_H s}\right)^{k_i}$$
$$\times f_{\rm binned}(T_\mu|n_i=k_i, i=0,1,2,3,4;\mu_H)\ ,$$

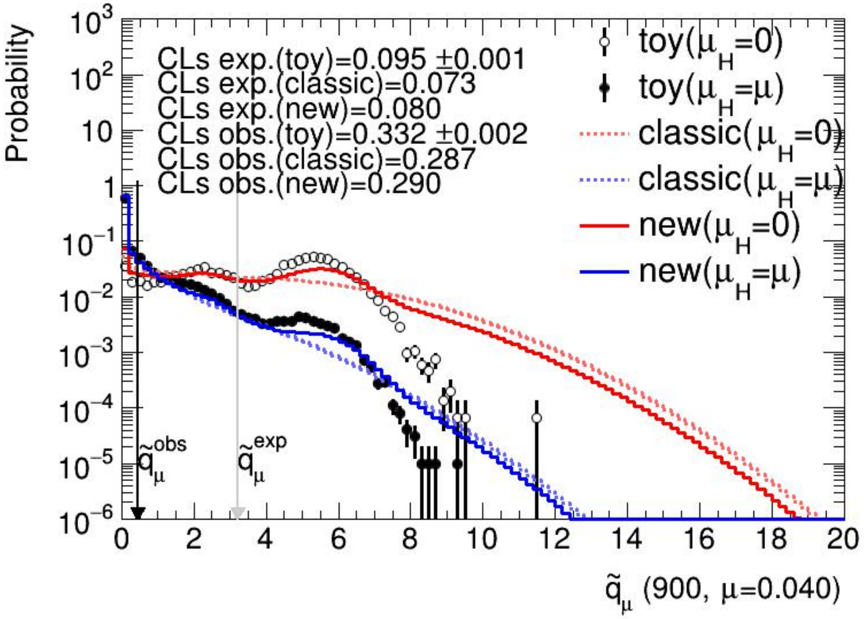
New Asymptotic Formulae

n_{small}: threshold between large statistics and small statistics

$$f(T_\mu|\mu_H) = \sum_{n=0}^{+\infty} f(T_\mu|n,\mu_H)P(n|b+\mu_H s)$$

$$\sum_{n=0}^{n_{\mathrm{small}}} f(T_\mu|n,\mu_H)P(n|b+\mu_H s) + \sum_{n>n_{\mathrm{small}}} f(T_\mu|n,\mu_H)P(n|b+\mu_H s)$$

$$\sum_{n=0}^{n_{\mathrm{small}}} f_{\mathrm{SS}}(T_\mu|n,\mu_H)P(n|b+\mu_H s) + (1-\sum_{n=0}^{n_{\mathrm{small}}} P(n|b+\mu_H s))f_{\mathrm{LS}}(T_\mu|n,\mu_H)$$



Small-Statistics part:
- discrete
- analytical/numerical calculation

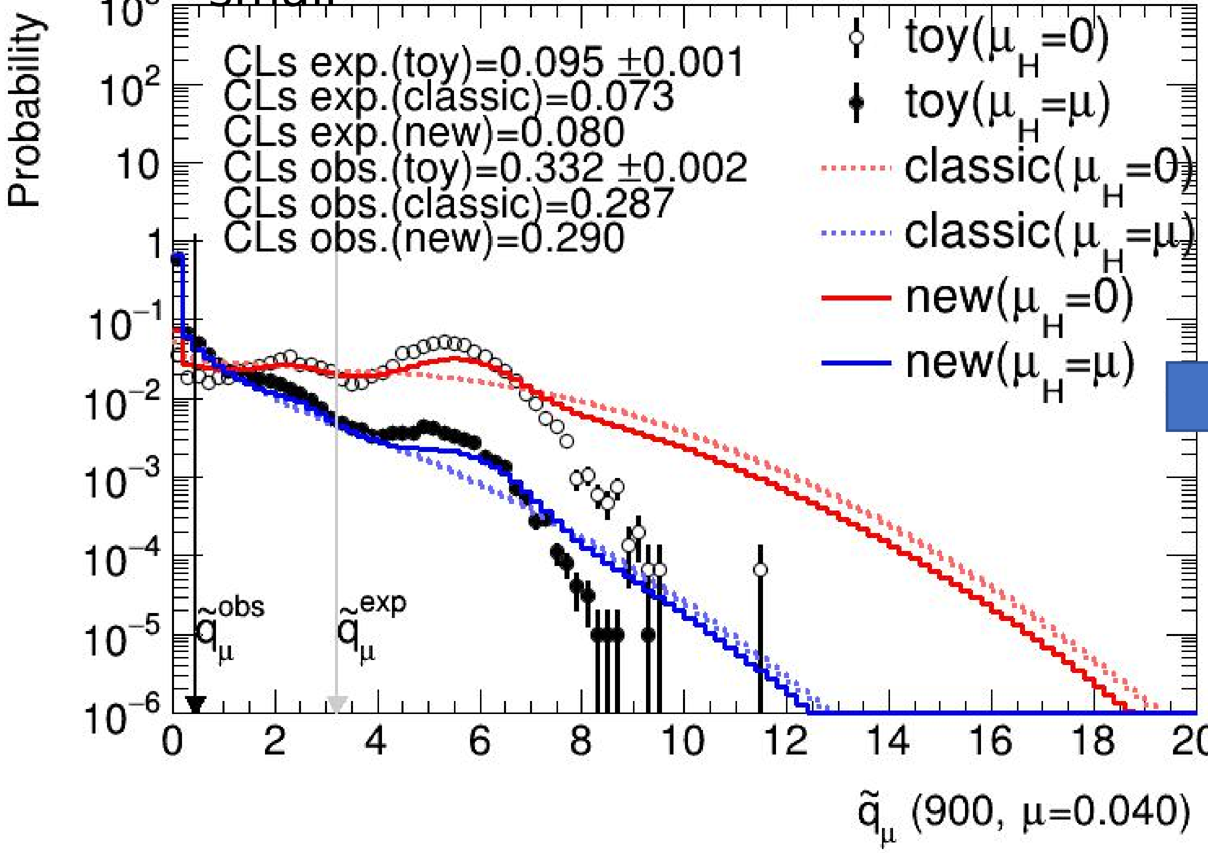
Large-Statistics part:
- continual
- classical asymptotic formulae

$$f_{\mathrm{SS}}(T_\mu|n,\mu_H) = \sum_{k_0+k_1+k_2+k_3+k_4+k_5=n} \frac{n!}{k_0!k_1!\cdots k_5!}\left(\frac{b_5}{b+\mu_H s}\right)^{k_5}\Pi_{i=0}^{4}\left(\frac{b_i+\mu_H s_i}{b+\mu_H s}\right)^{k_i}$$
$$\times f_{\mathrm{binned}}(T_\mu|n_i=k_i, i=0,1,2,3,4;\mu_H)\ ,$$

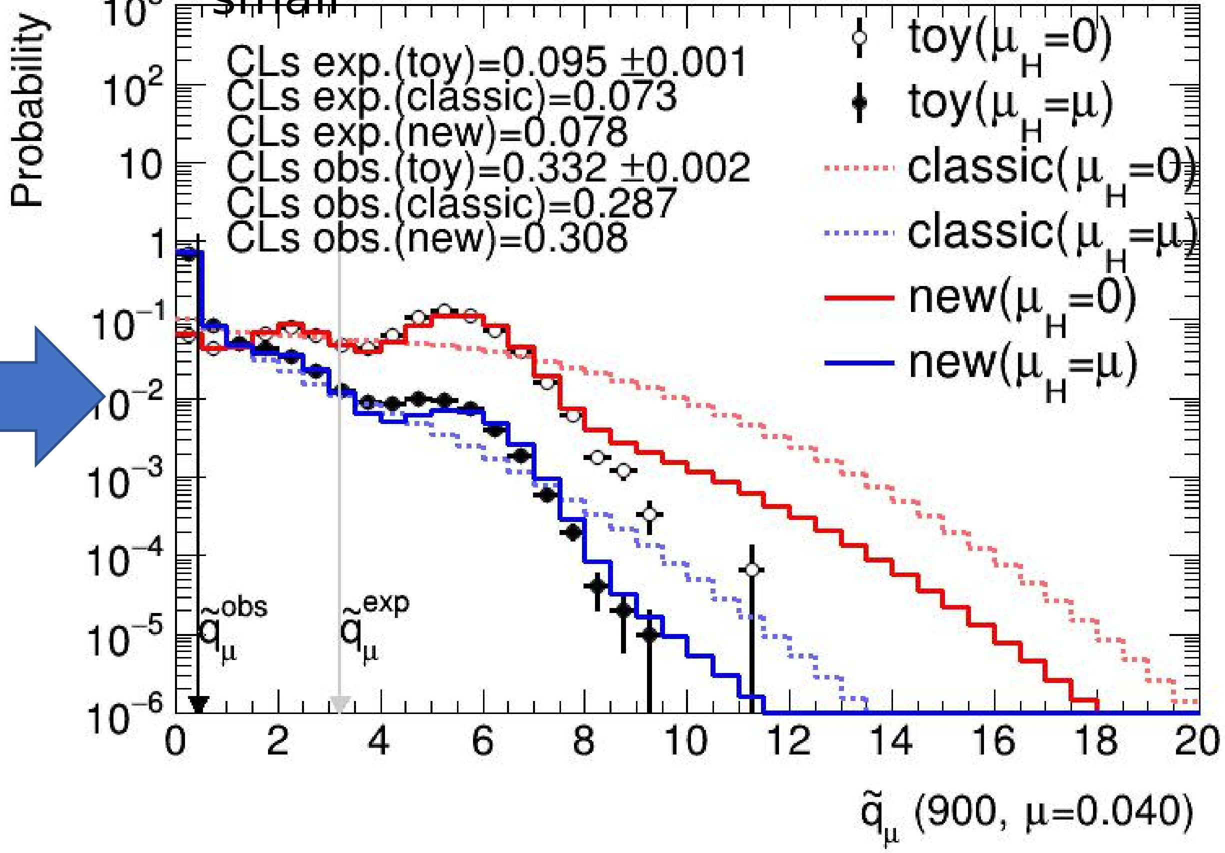
New Asymptotic Formulae

n_{small}: threshold between large statistics and small statistics

A conservative choice of n_{small}



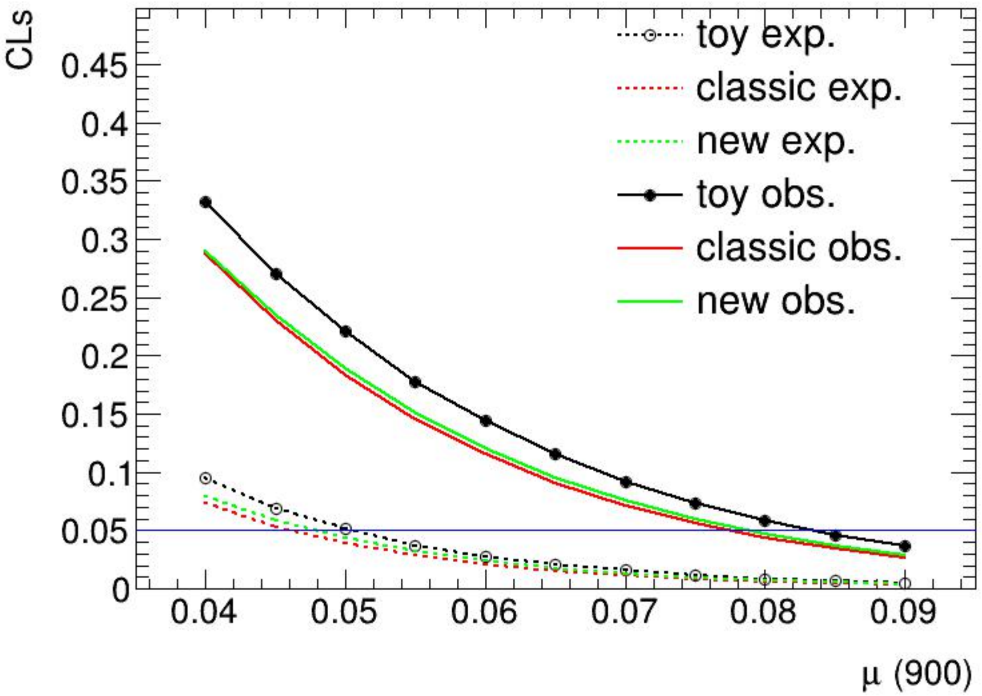
An aggressive choice of n_{small}



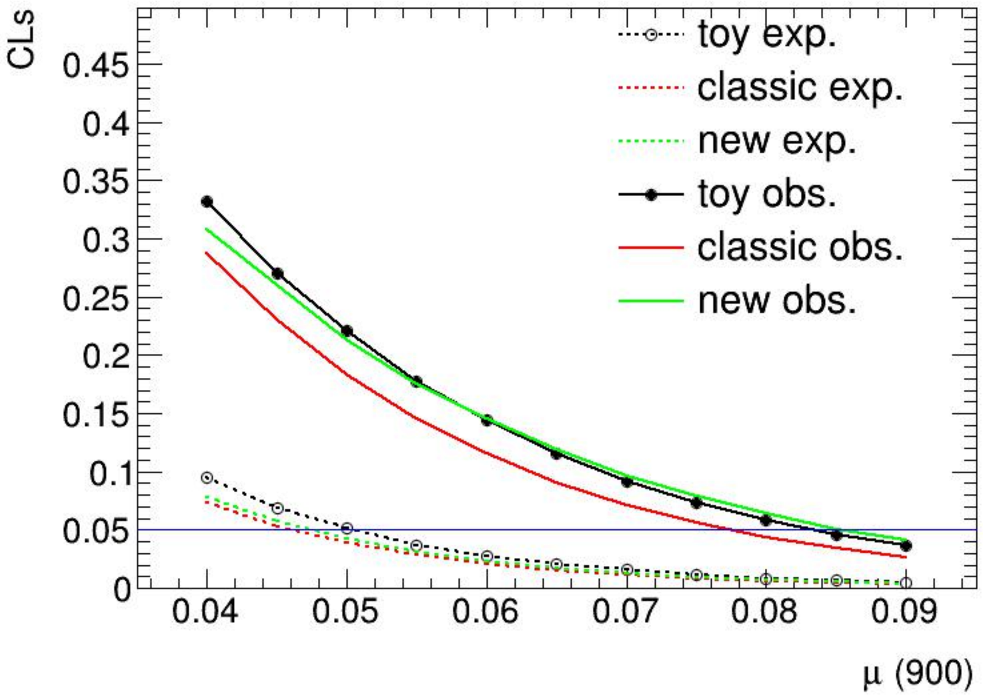
New Asymptotic Formulae

n_{small}: threshold between large statistics and small statistics

A conservative choice of n

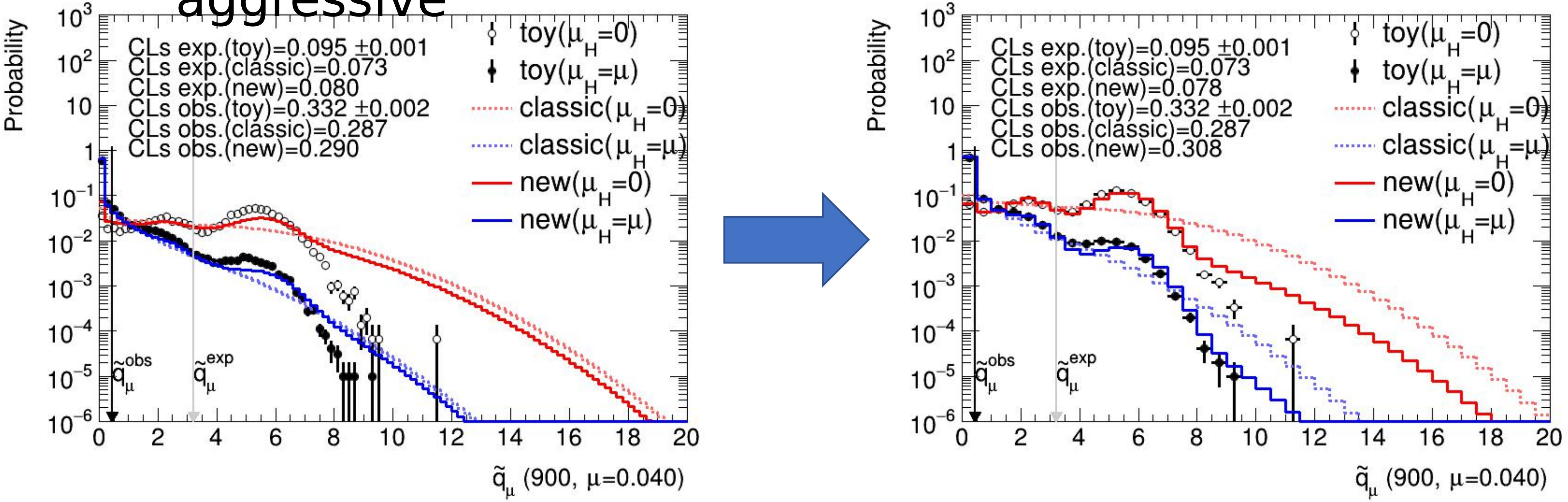


An aggressive choice of n



Comments on New Asymptotic Formulae

- Still more of education value due to
  - How to choose n_{small}?
  - How many bins in the Small-Statistics part?
  - How to correct the Large-Statistics part if we want to be aggressive



Summary

- New asymptotic formulae proposed (work in progress)
  - Large-statistics part: described by classical asymptotic formulae
  - Small-statistics part: described by analytic/numerical calculation

- Application to a real case: di-higgs resonance search X HH bbyy
  - Look promising
  - Still some issues to be fixed
  - More of education value for the moment

Thank you for your attention!