Taming Systematic Uncertainties in Experimental Physics Analyses

Georgios Alexandris

University of Bonn

ICPS 2024, 4–11 August 2024

Statistical VS Systematic Uncertainties

In summary:

- ➢ Statistical Uncertainties
  - **Decrease** with more data (~1/√N, usually)
  - Easy to estimate
- ➢ Systematic Uncertainties
  - **Might decrease** with more data (**no guarantees**)
  - Difficult to estimate
  - Introduce **correlations** in our measurements

Spectroscopy Accuracy Measurement Errors (hitachi-hightech.com)

Maximum Likelihood Estimation

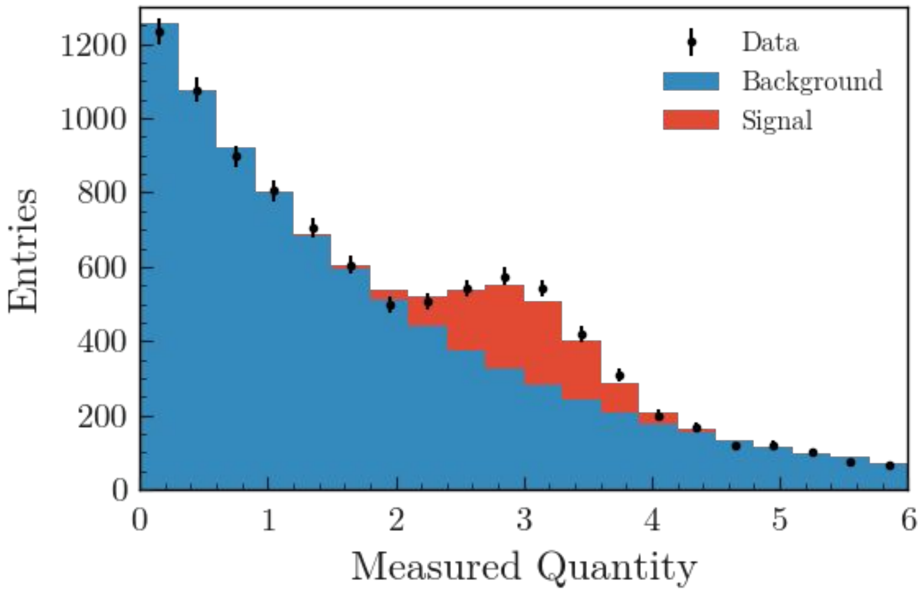
- Maximizes the probability of model to describe the data

$$\mathcal{L}(\vec{\theta}|\vec{x}) = \prod_i \mathrm{PDF}(x_i|\vec{\theta}) \qquad \text{e.g. } \mathcal{L}(\mu,\sigma|\vec{x}) = \prod_i e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

- For numerical stability, minimize $-2\ln\mathcal{L}(\vec{\theta}|\vec{x})$
- Gaussian case: $-2\ln\mathcal{L}\left(\mu,\sigma|\vec{x}\right) = \sum_i \frac{(x_i-\mu)^2}{\sigma^2} = \chi^2\left(\vec{\theta}\right)$
- <u>Advantage</u>: PDF can be anything

Templates

- ➢ Analytical form of distributions describing data not exactly known
- ➢ Two solutions
  - ○ Empirical PDFs (e.g. Gaussian signal + Exponential Background)
  - ○ Simulation, outputs are histograms called **Templates**



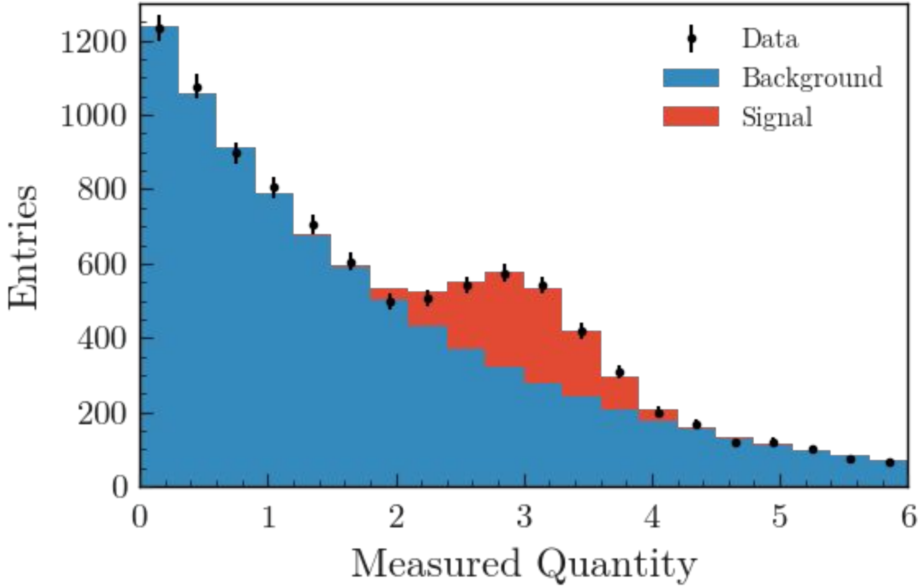
"Stat. only fit"

- Likelihood:

$$\mathcal{L}(\eta_{k=1,2}|\mathbf{n}) = \prod_{\mathrm{b}\in\mathrm{bins}} \mathcal{P}(n_{\mathrm{b}}|\nu_{\mathrm{b}})$$

$$\nu_{\mathrm{b}} = \sum_{k=1,2} f_{\mathrm{b}k}\eta_k = \sum_{k=1,2} \eta_{\mathrm{b}k}$$

- Minimize $-2\ln\mathcal{L}(\eta_1,\eta_2)$
- Profile Likelihood Scan

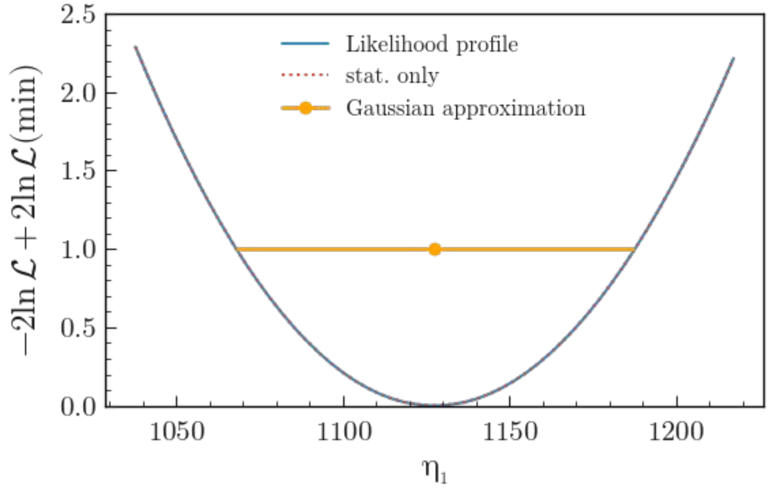


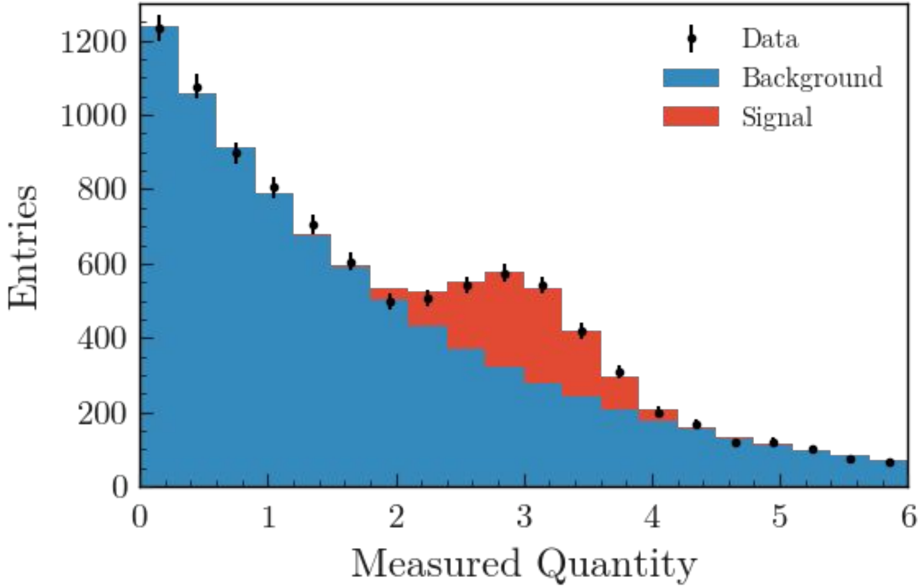
“Stat. only fit”

- Likelihood:

$$\mathcal{L}(\eta_{k=1,2}|\mathbf{n}) = \prod_{b\in \text{bins}} \mathcal{P}(n_b|\nu_b)$$

$$\nu_b = \sum_{k=1,2} f_{bk}\eta_k = \sum_{k=1,2} \eta_{bk}$$





- Param. of Interest (POI): $\eta_{sig} \equiv \eta_1$ (signal yield)

Systematic Uncertainties

Types of Systematic Uncertainties

Non-exhaustive list:

- ➢ Calibration
- ➢ Detector acceptance and efficiency
- ➢ Detector resolution
- ➢ Background estimation
- ➢ Modeling and Theory
- ➢ Measurements from others
- ➢ etc..

Covariance Matrices

- Assume known systematic sources & uncertainties
- How do we add this info to our measurements?
- One solution: **Covariance Matrix** for each source
- Final uncertainties, from sum of covariance matrices

$A = A_{true} + A_{stat} + \mathbf{A_{sys}}$

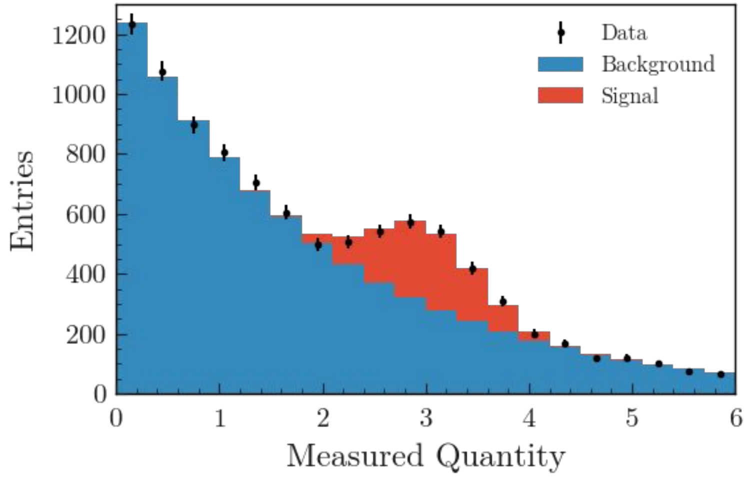
$B = B_{true} + B_{stat} + \mathbf{B_{sys}}$

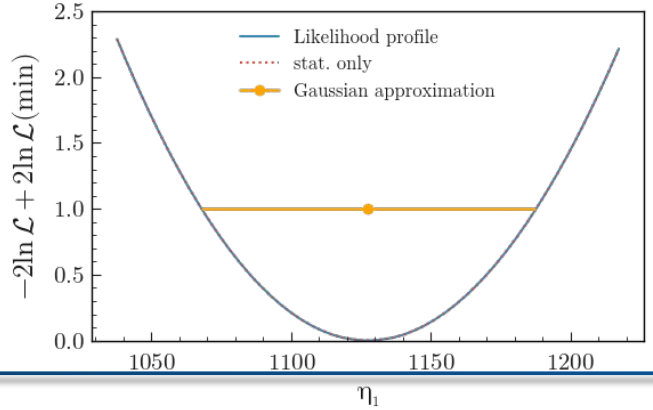
$$\longrightarrow \begin{pmatrix} \sigma^2_{A,stat} & 0 \\ 0 & \sigma^2_{B,stat} \end{pmatrix} + \begin{pmatrix} \sigma^2_{A,sys} & \sigma_{A,sys}\sigma_{B,sys} \\ \sigma_{A,sys}\sigma_{B,sys} & \sigma^2_{B,sys} \end{pmatrix}$$

Same source $\longrightarrow$ Correlated measurements

Nuisance Parameters

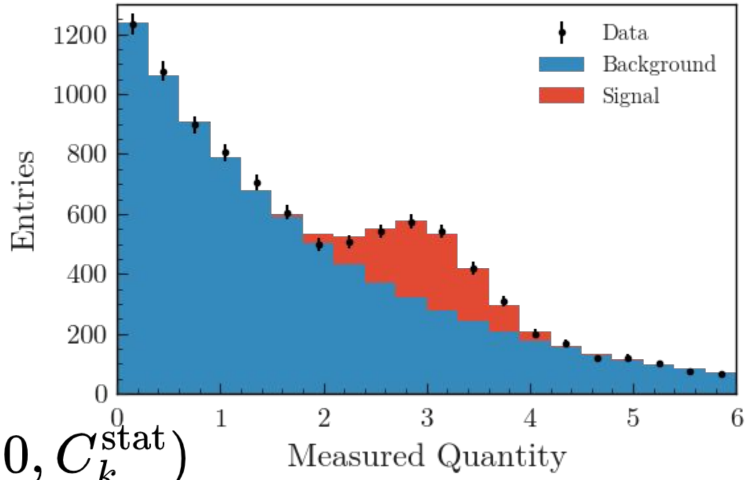
- Parameters our model depends on, but their exact value is not of interest (like $\eta_{bkg} \equiv \eta_2$)
- Systematics affect the model
- Must be included as Nuisance Parameters!





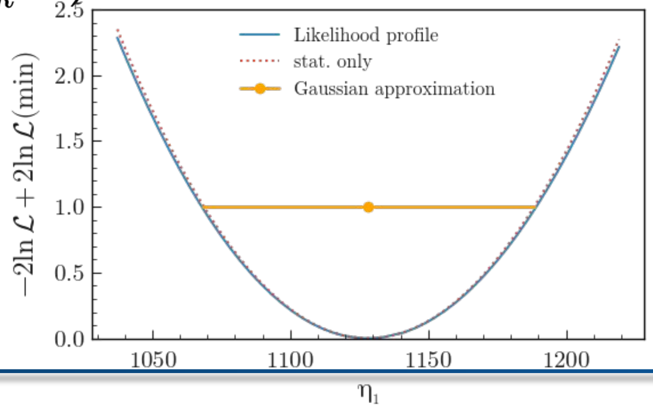
MC Uncertainties

- Templates have statistical uncertainties $\rightarrow \Sigma_k^{\text{stat}} = \text{Diag}(\boldsymbol{\sigma}_k^2)$
- Include Nuisance Parameters for each bin, for each Template: $\theta_{\text{b}k}$



$$\mathcal{L}(\eta_{k=1,2}, \boldsymbol{\theta}_{k=1,2}) = \prod_{\text{b}\in\text{bins}} \mathcal{P}(n_\text{b}|\nu_\text{b}(\boldsymbol{\theta})) \prod_{k=1,2} \mathcal{N}(\boldsymbol{\theta}_k|0, C_k^{\text{stat}})$$

$$\eta_{\text{b}k}(\theta) = \frac{\eta_{\text{b}k}\,(1+\theta_{\text{b}k}\varepsilon_{\text{b}k})}{\sum_j \eta_{jk}\,(1+\theta_{jk}\varepsilon_{jk})}$$



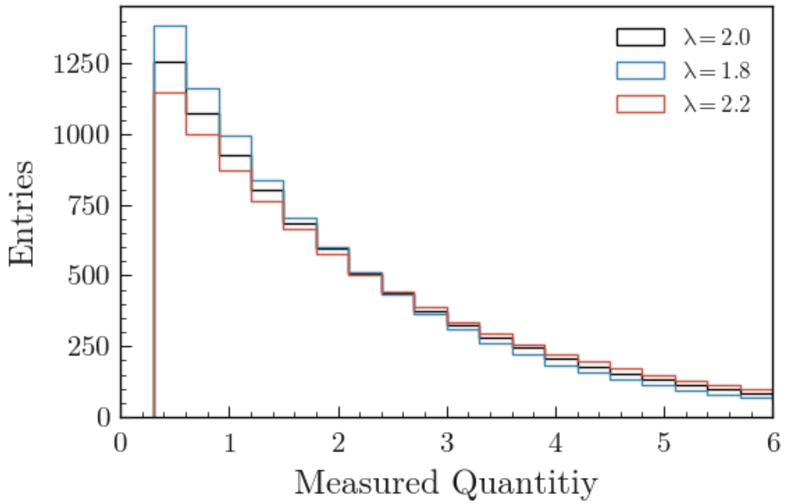
Model (Shape) Uncertainties

- Modeling errors, e.g. λ=2.0±0.2, bkg only
- Covariance matrix fully correlated between bins

$$\Sigma_k^{\mathrm{syst}} = \boldsymbol{\sigma}_k \otimes \boldsymbol{\sigma}_k$$

- Likelihood has same form, but

$$\Sigma_k^{\mathrm{stat}} \to \Sigma_k = \Sigma_k^{\mathrm{stat}} + \Sigma_k^{\mathrm{syst}}$$



Model (Shape) Uncertainties

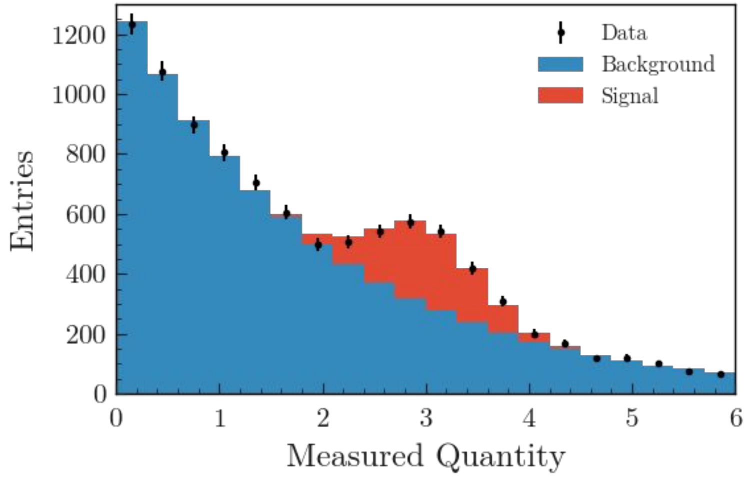
- Modeling errors, e.g. λ=2.0±0.2, bkg only

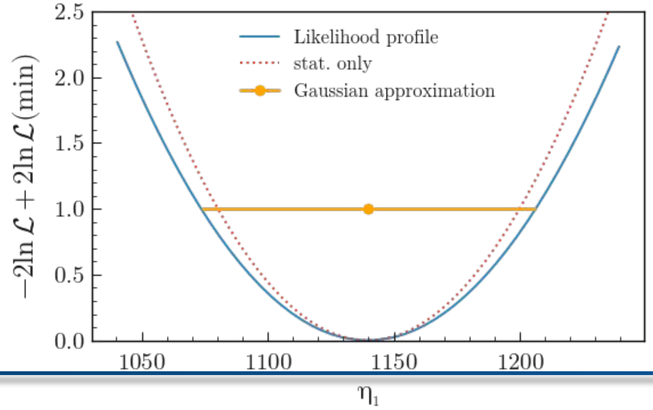
- Covariance matrix fully correlated between bins

$$\Sigma_k^{\text{syst}} = \boldsymbol{\sigma}_k \otimes \boldsymbol{\sigma}_k$$

- Likelihood has same form, but

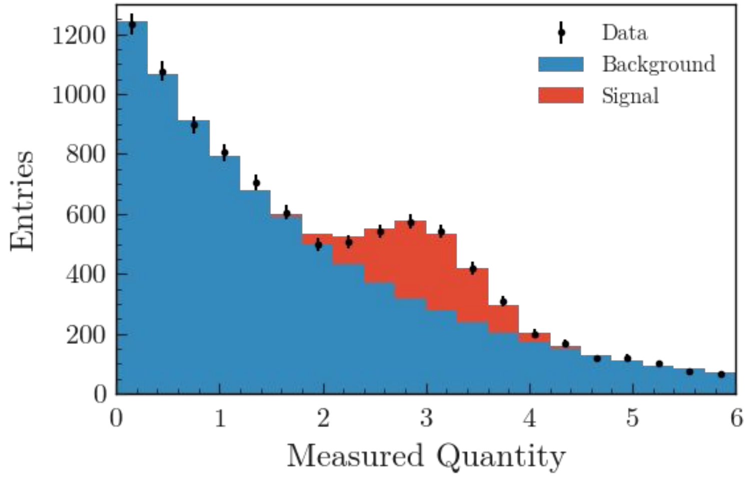
$$\Sigma_k^{\text{stat}} \rightarrow \Sigma_k = \Sigma_k^{\text{stat}} + \Sigma_k^{\text{syst}}$$



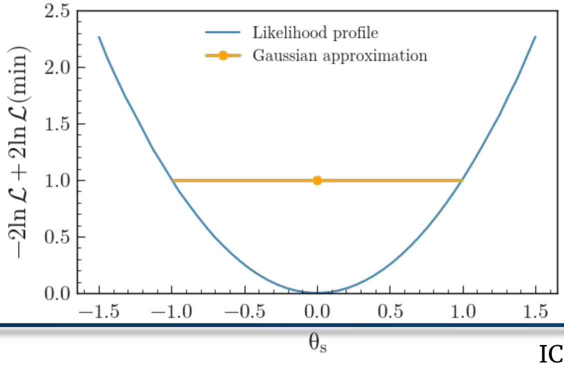


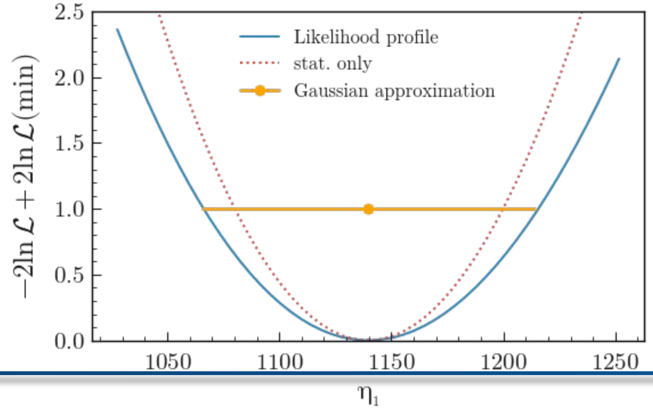
Normalization (Multiplicative) Uncertainties

- Errors on total scale factors (luminosity, efficiency, etc)
- Let's introduce 2% uncertainty on signal tracking efficiency
- Modify signal yield



$$\eta_1(\boldsymbol{\theta}_1, \theta_s) = \eta_1(\boldsymbol{\theta}_1)(1 + \theta_s \varepsilon_s)$$





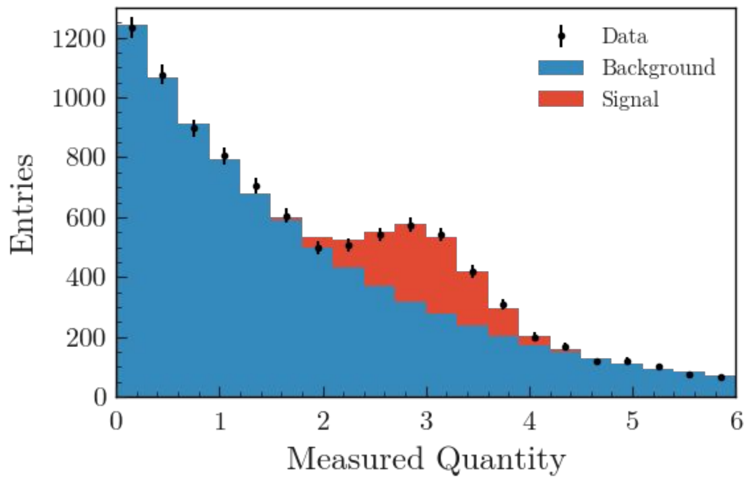
Final Likelihood

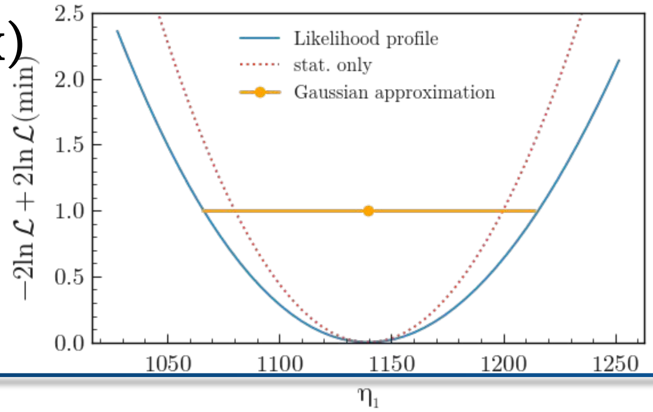
$$\mathcal{L}(\eta_1, \eta_2, \theta_s, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \prod_{\mathrm{b}\in\mathrm{bins}} \mathcal{P}\left(n_{\mathrm{b}} | \nu_{\mathrm{b}}(\eta_1, \eta_2, \theta_s, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)\right) \times \prod_{k=1,2} \mathcal{N}(\boldsymbol{\theta}_k | 0, C_k) \times \mathcal{N}(\theta_s | 0, 1)$$

$$\nu_{\mathrm{b}}(\eta_{k=1,2}, \theta_s, \boldsymbol{\theta}_{k=1,2}) = \sum_{k=1,2} \frac{\eta_{\mathrm{b}k}(1 + \theta_{\mathrm{b}k}\varepsilon_{\mathrm{b}k})}{\sum_j \eta_{\mathrm{j}k}(1 + \theta_{\mathrm{j}k}\varepsilon_{\mathrm{j}k})}(1 + \theta_s \varepsilon_s \delta_{1k})$$

$$\Sigma_k = \Sigma_k^{MC} + \Sigma_k^{syst} \rightarrow C_k$$ (Correlation Matrix)

$\varepsilon_{\mathrm{b}k/s}$ (Relative Uncertainties)





Some References

- Very good modern book: Data Analysis in High Energy Physics
  - Olaf Behnke, Kevin Kröninger, Grégory Schott, and Thomas Schörner-Sadenius
- Classic #1: Statistical Data Analysis
  - Glen Cowan
- Classic #2: Statistics
  - R. J. Barlow
- Hidden Gem: Systematic uncertainties and profiling
  - Wouter Verkerke

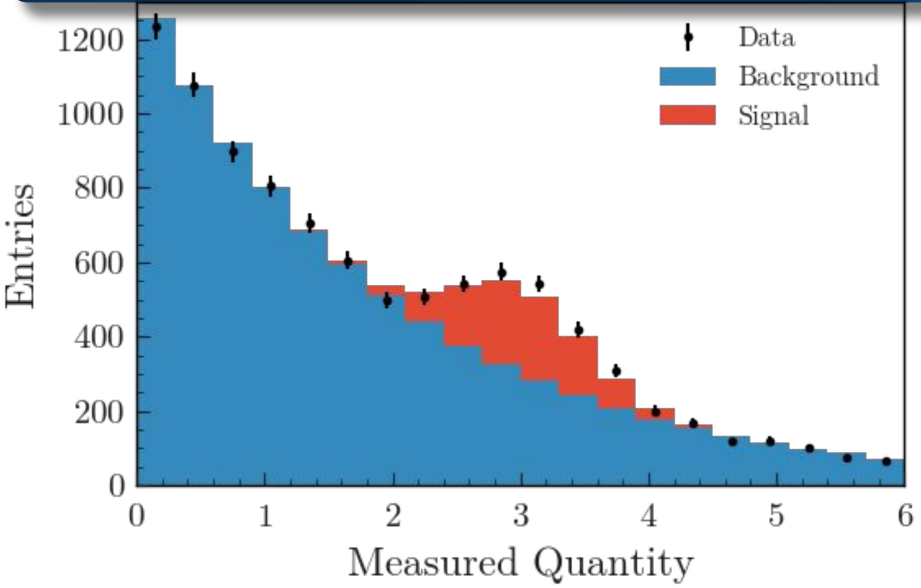
Thank you for the attention!

Questions?

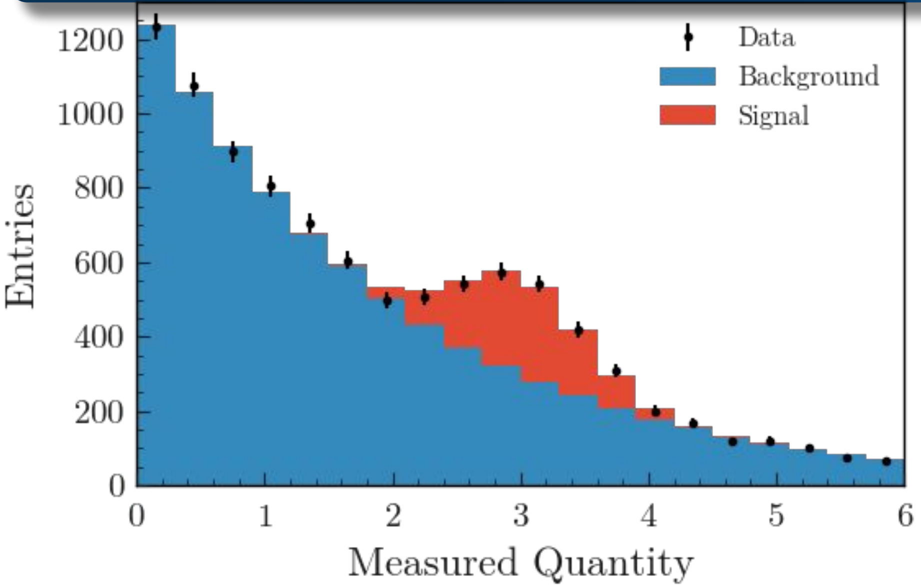
Backup

Plots-only procedure

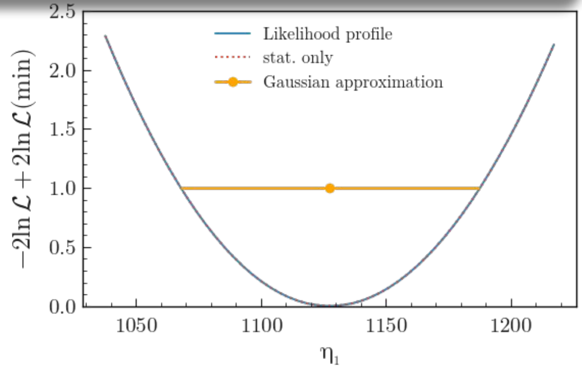
Initial Template



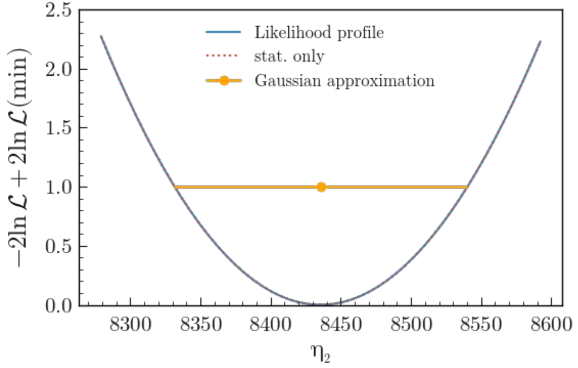
Stat. only



Data
Background
Signal
Entries
Measured Quantity

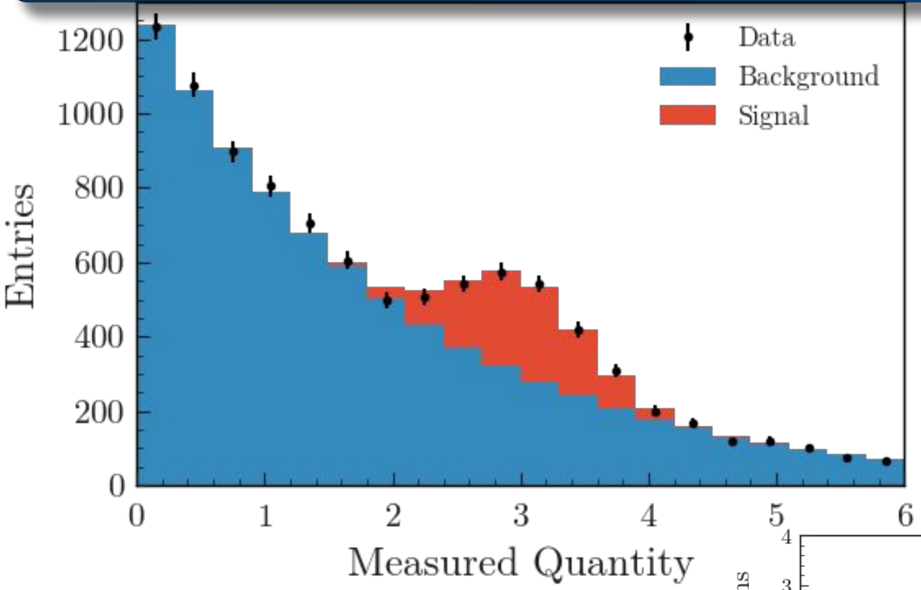


Likelihood profile
stat. only
Gaussian approximation
$-2\ln\mathcal{L} + 2\ln\mathcal{L}(\min)$
η_1

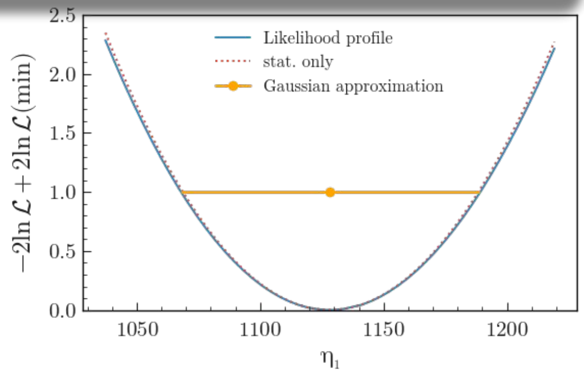


Likelihood profile
stat. only
Gaussian approximation
$-2\ln\mathcal{L} + 2\ln\mathcal{L}(\min)$
η_2

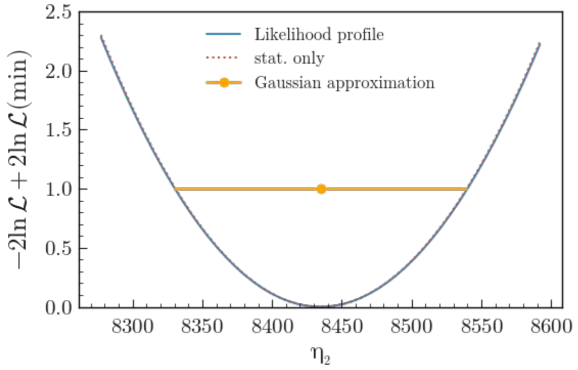
Stat. + MC



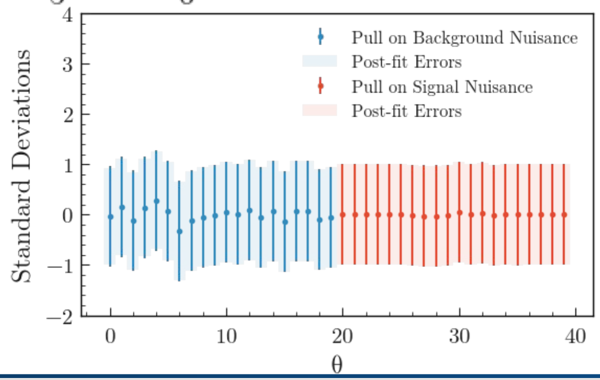
Data
Background
Signal
Entries
Measured Quantity



Likelihood profile
stat. only
Gaussian approximation
$-2\ln\mathcal{L}+2\ln\mathcal{L}(\min)$
η_1

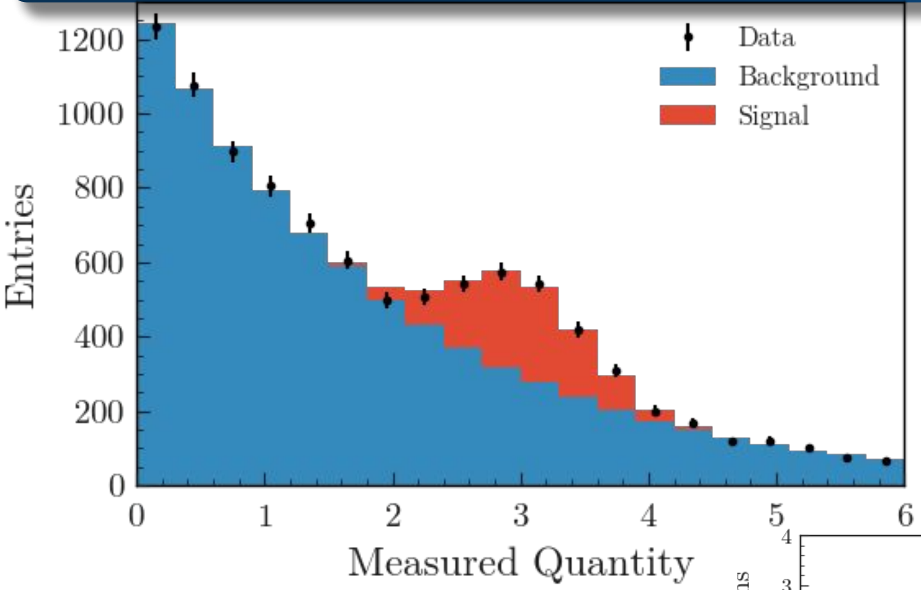


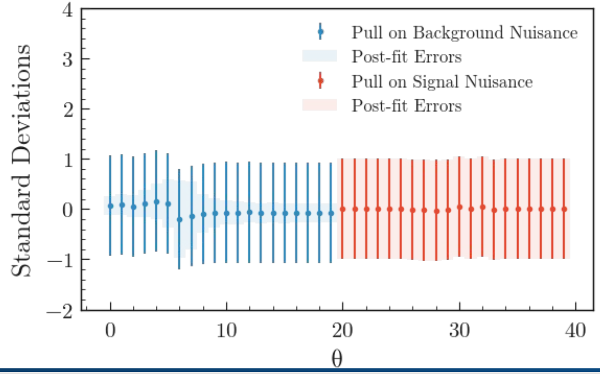
Likelihood profile
stat. only
Gaussian approximation
$-2\ln\mathcal{L}+2\ln\mathcal{L}(\min)$
η_2

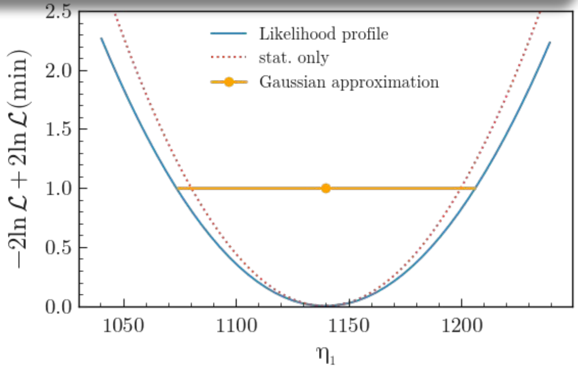


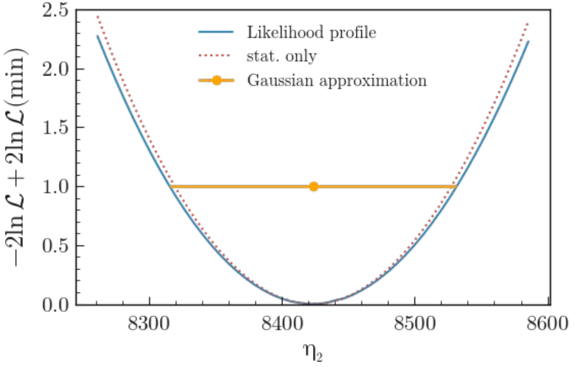
Pull on Background Nuisance
Post-fit Errors
Pull on Signal Nuisance
Post-fit Errors
Standard Deviations
θ

Stat. + MC + Model

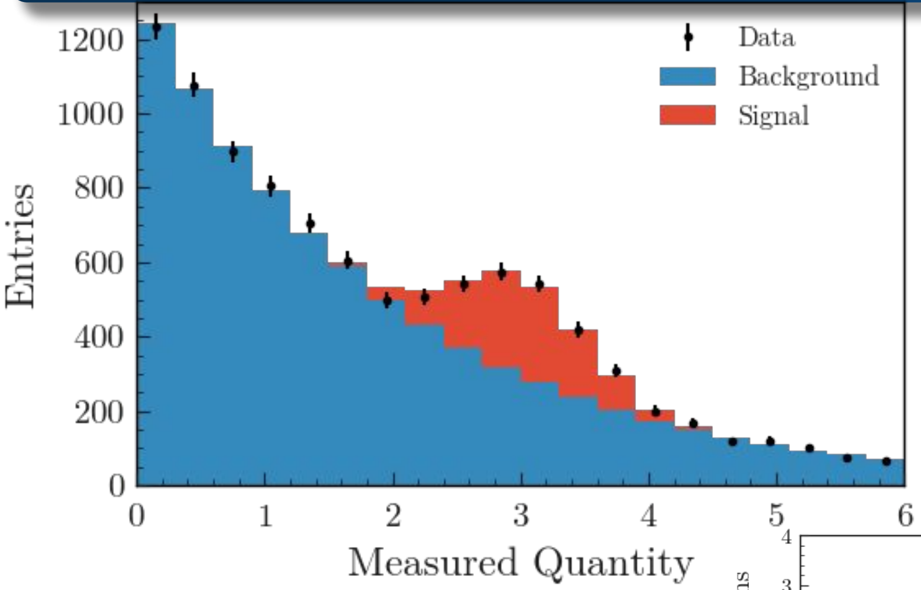


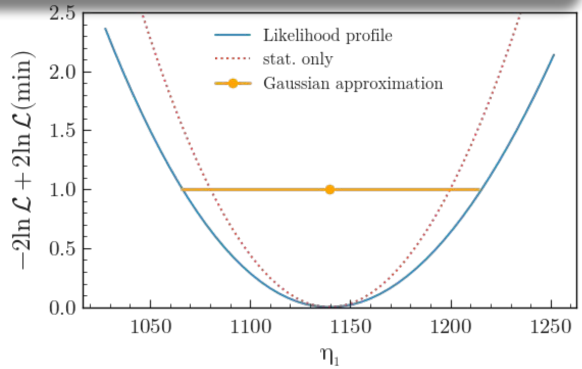


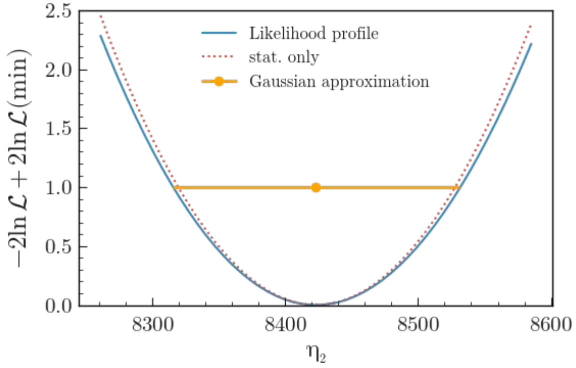


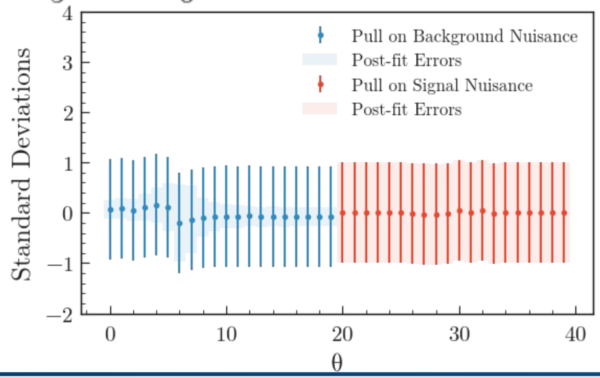


Stat. + MC + Model + Normalization







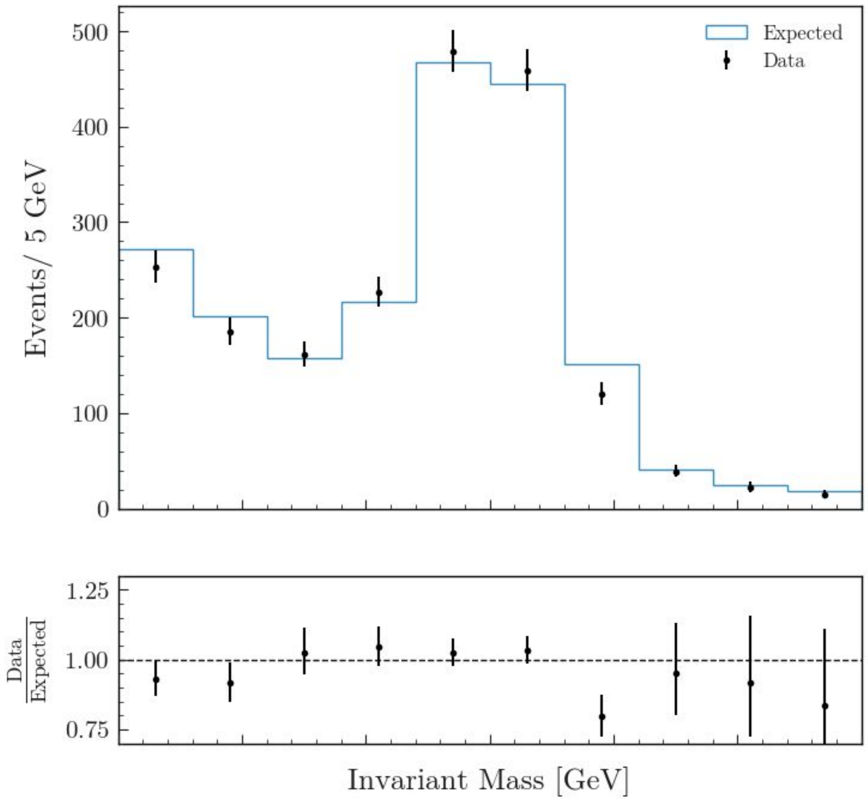


Fit Uncertainties

Uncertainties (from fit)

➢ Let's fit the following model

$$f_{\text{tot}}(x|s, b, \mu, \sigma, \tau) = s \cdot \text{Gauss}(x|\mu, \sigma) + b \cdot \text{Exp}(x|\tau)$$

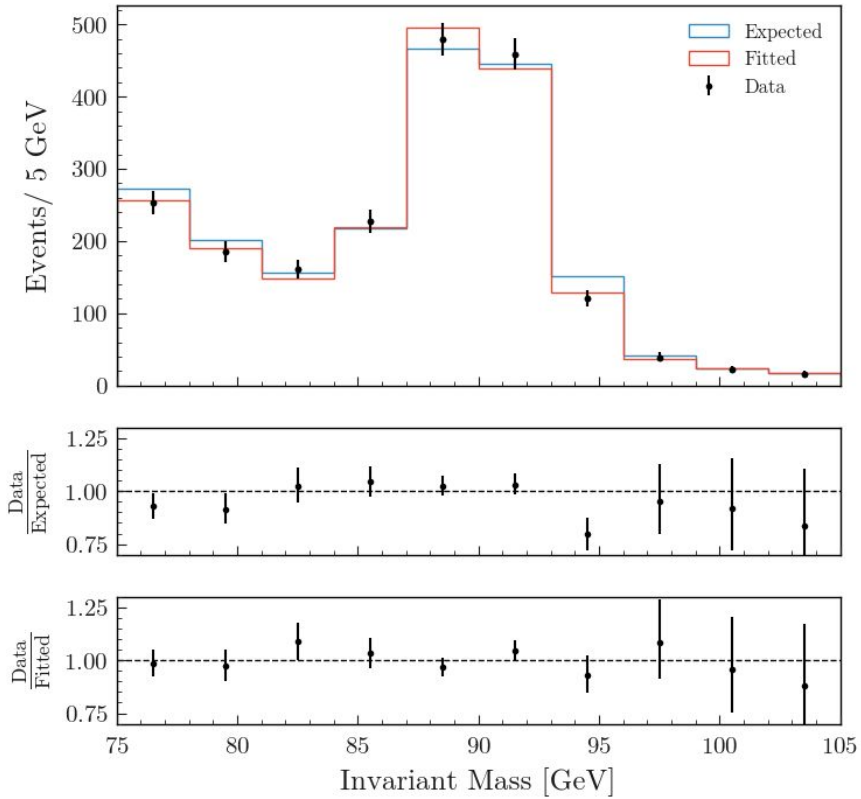


Uncertainties (from fit)

- ➢ Let's fit the following model:

$$f_{\text{tot}}(x|s,b,\mu,\sigma,\tau) = s \cdot \text{Gauss}(x|\mu,\sigma) + b \cdot \text{Exp}(x|\tau)$$

- ➢ Minimize $-2\ln\mathcal{L}(s,b,\mu,\sigma,\tau|\text{data})$
  - ○ 5D likelihood!



Uncertainties (from fit)

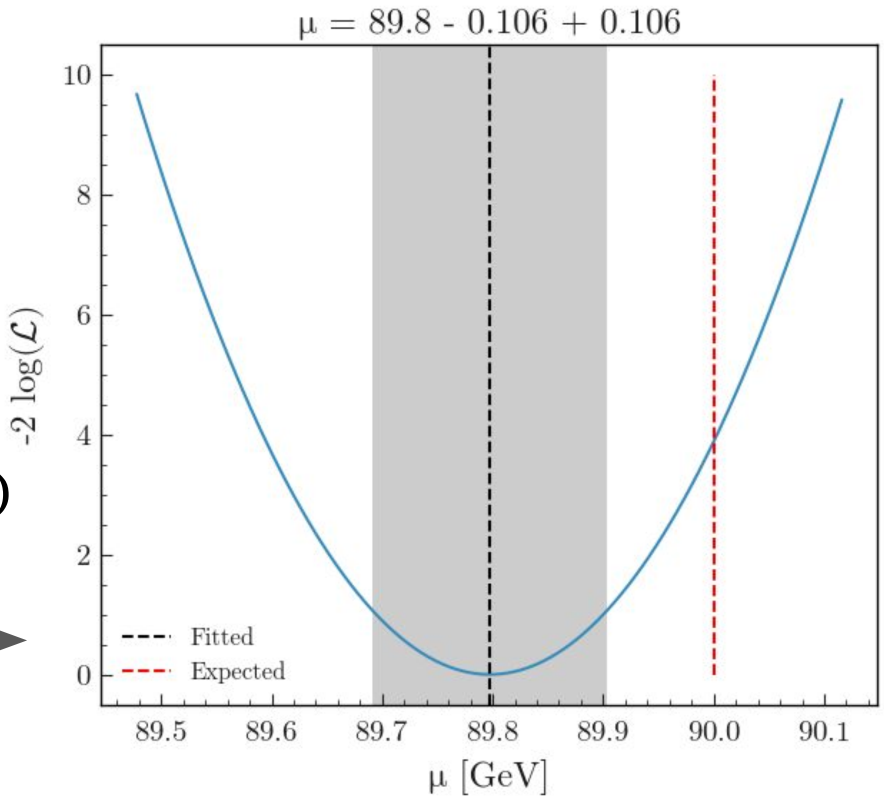
➢ Let's fit the following model:

$$f_{\text{tot}}(x|s,b,\mu,\sigma,\tau) = s \cdot \text{Gauss}(x|\mu,\sigma) + b \cdot \text{Exp}(x|\tau)$$

➢ Minimize $-2\ln\mathcal{L}(s,b,\mu,\sigma,\tau|\text{data})$

- 5D likelihood!
- Show projections

1D scan of the likelihood over 1 parameter (μ)



➢ Get (stat) uncertainties from points where $-2\ln\mathcal{L} = -2\ln\mathcal{L}_{\min} + 1$

Uncertainties (from fit)

- Let's fit the following model:

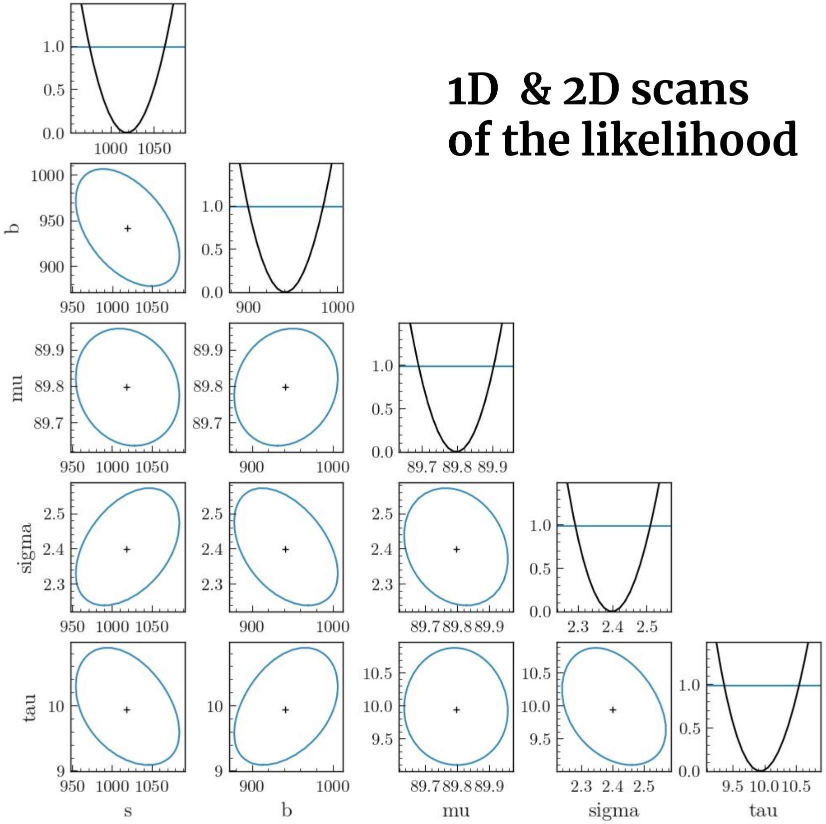
$$f_{\text{tot}}(x|s,b,\mu,\sigma,\tau) = s \cdot \text{Gauss}(x|\mu,\sigma) + b \cdot \text{Exp}(x|\tau)$$

- Minimize $-2\ln\mathcal{L}(s,b,\mu,\sigma,\tau|\text{data})$
  - 5D likelihood!
  - Show projections

- Another estimate of uncertainties: inverse Hessian

$$V(\hat{\theta}) \xrightarrow{n\to\infty} = \left\{ -n \left\langle \frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2} \right\rangle_{\theta=\hat{\theta}} \right\}^{-1}$$

(Cramer-Rao bound)



1D & 2D scans of the likelihood

Evaluation of Sys. Uncertainties

Evaluation of Systematic Uncertainties (a common case)

- Uncertainties on external input parameters
- Simulate at ±1σ for each parameter
- Resulting variations are systematic uncertainties
- e.g. Resonance width for MC simulation, $\Gamma = 2.5 \pm 0.5$ GeV