30/10/2024

ROYAL HOLLOWAY UNIVERSITY OF LONDON

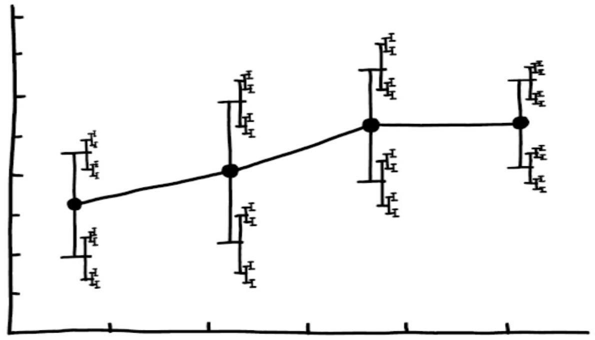
Errors-on-errors (Gamma Variance Model)

Enzo Canonero, Glen Cowan

“Uncertain systematics”

“Errors-on-errors”

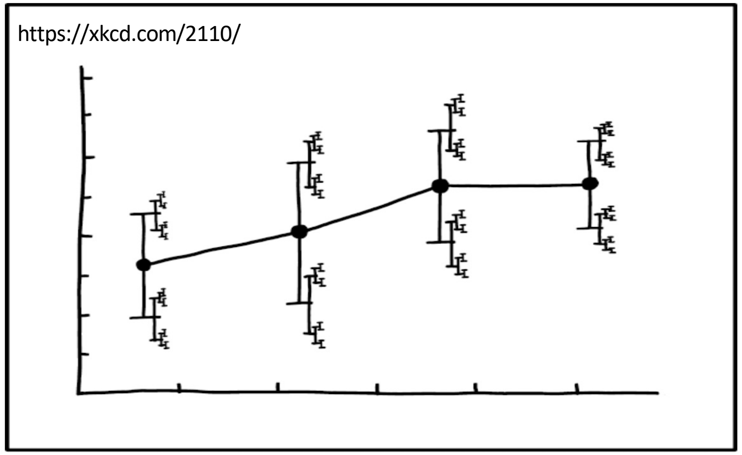
https://xkcd.com/2110/



References: Eur. Phys. J. C (2019) 79:133, Eur. Phys. J. C (2023) 83:1100, arXiv: 2407.05322

"Uncertain systematics"

"Errors-on-errors"

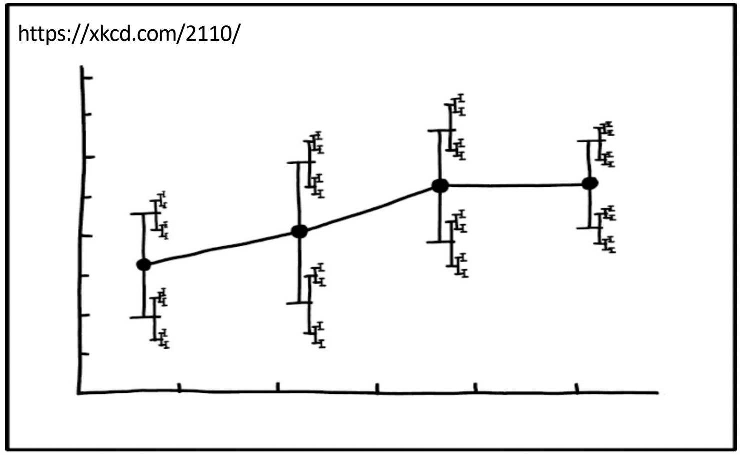


1) *Some **systematic uncertainties** can be well estimated:*
   - **Related to size of control measurements dataset**
   - **Related to size of MC sample**

References: Eur. Phys. J. C (2019) 79:133, Eur. Phys. J. C (2023) 83:1100, arXiv: 2407.05322

"Uncertain systematics"

"Errors-on-errors"



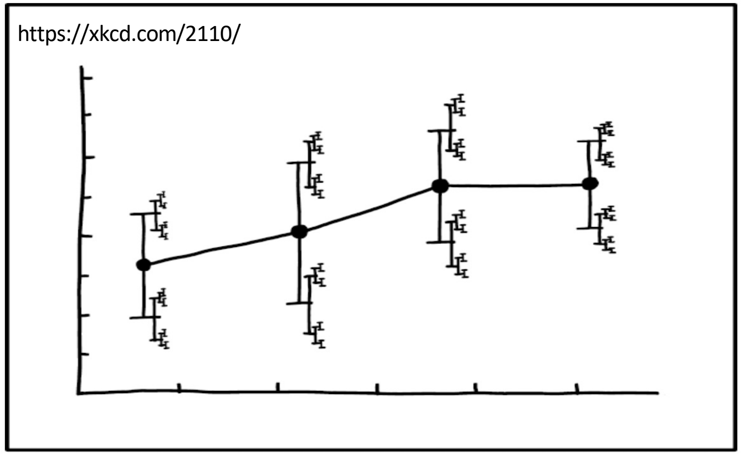
1) *Some **systematic uncertainties*** can be well estimated:
   - **Related to size of control measurements dataset**
   - **Related to size of MC sample**

2) But they can also be ***quite uncertain***:
   - **Theory systematics**
   - **Two points systematics**

References: Eur. Phys. J. C (2019) 79:133, Eur. Phys. J. C (2023) 83:1100, arXiv: 2407.05322

"Uncertain systematics"

"Errors-on-errors"



1) *Some **systematic uncertainties*** can be well estimated:
   - **Related to size of control measurements dataset**
   - **Related to size of MC sample**

2) But they can also be ***quite uncertain***:
   - **Theory systematics**
   - **Two points systematics**

→ **Non-trivial effects**

References: Eur. Phys. J. C (2019) 79:133, Eur. Phys. J. C (2023) 83:1100, arXiv: 2407.05322

Formulation of the problem

ROYAL HOLLOWAY UNIVERSITY OF LONDON

$$L(\boldsymbol{\mu}, \boldsymbol{\theta}) = P(\boldsymbol{y}, \boldsymbol{u} | \boldsymbol{\mu}, \boldsymbol{\theta}) = P(\boldsymbol{y} | \boldsymbol{\mu}, \boldsymbol{\theta}) \times \prod_i \frac{1}{\sqrt{2\pi}\sigma_{u_i}} e^{-(u_i - \theta_i)^2 / 2\sigma_{u_i}^2}$$

Formulation of the problem

ROYAL HOLLOWAY UNIVERSITY OF LONDON

$$L(\boldsymbol{\mu}, \boldsymbol{\theta}) = P(\boldsymbol{y}, \boldsymbol{u} | \boldsymbol{\mu}, \boldsymbol{\theta}) = \boxed{P(\boldsymbol{y} | \boldsymbol{\mu}, \boldsymbol{\theta})} \times \prod_i \frac{1}{\sqrt{2\pi}\sigma_{u_i}} e^{-(u_i - \theta_i)^2 / 2\sigma_{u_i}^2}$$

- $\boldsymbol{y}$ = Main measurements
- $\boldsymbol{\mu}$ = Parameters of interest
- $\boldsymbol{\theta}$ = Nuisance parameters (Systematic effects)

Formulation of the problem

$$L(\boldsymbol{\mu},\boldsymbol{\theta}) = P(\boldsymbol{y},\boldsymbol{u}|\boldsymbol{\mu},\boldsymbol{\theta}) = P(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\theta})\times\prod_i \frac{1}{\sqrt{2\pi}\sigma_{u_i}} e^{-(u_i-\theta_i)^2/2\sigma_{u_i}^2}$$

- $\boldsymbol{y}$ = Main measurements
- $\boldsymbol{\mu}$ = Parameters of interest
- $\boldsymbol{\theta}$ = Nuisance parameters (Systematic effects)

- Gaussian constrains for systematic effects
- $\boldsymbol{u_i}$ = Auxiliary measurements constraining nuisance parameters (often set = 0)

Formulation of the problem

$$L(\boldsymbol{\mu}, \boldsymbol{\theta}) = P(\boldsymbol{y}, \boldsymbol{u} | \boldsymbol{\mu}, \boldsymbol{\theta}) = P(\boldsymbol{y} | \boldsymbol{\mu}, \boldsymbol{\theta}) \times \prod_i \frac{1}{\sqrt{2\pi}\sigma_{u_i}} e^{-(u_i - \theta_i)^2 / 2\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2}$$

- $\boldsymbol{y} =$ Main measurements
- $\boldsymbol{\mu} =$ Parameters of interest
- $\boldsymbol{\theta} =$ Nuisance parameters (Systematic effects)

- Gaussian constrains for systematic effects
- $\boldsymbol{u}_i =$ Auxiliary measurements constraining nuisance parameters (often set = 0)
- $\boldsymbol{\sigma}_{\boldsymbol{u}_i} =$ Systematic errors

Formulation of the problem

$$\log L(\boldsymbol{\mu}, \boldsymbol{\theta}) = \log P(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\theta}) - \sum \frac{(\boldsymbol{u_i} - \boldsymbol{\theta_i})^2}{2\boldsymbol{\sigma_{u_i}^2}}$$

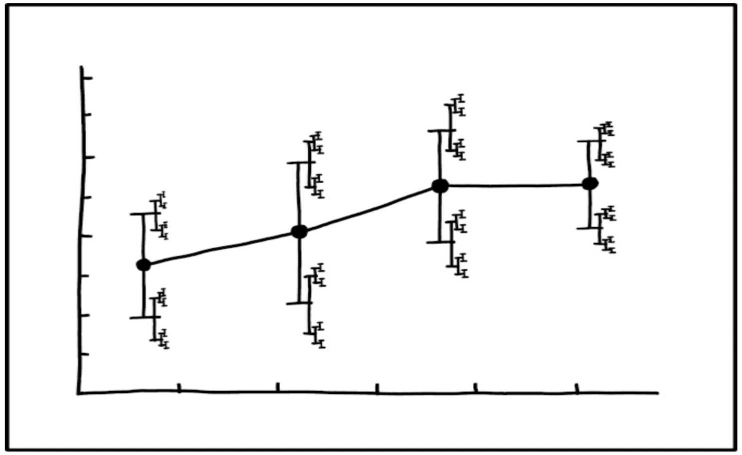
Let systematic errors be potentially uncertain!

Formulation of the problem

$$\log L(\boldsymbol{\mu}, \boldsymbol{\theta}) = \log P(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\theta}) - \sum \frac{(\boldsymbol{u_i} - \boldsymbol{\theta_i})^2}{2\boldsymbol{\sigma_{u_i}^2}}$$

Let systematic errors be potentially uncertain!

- We do not literally introduce error bars on errors bars

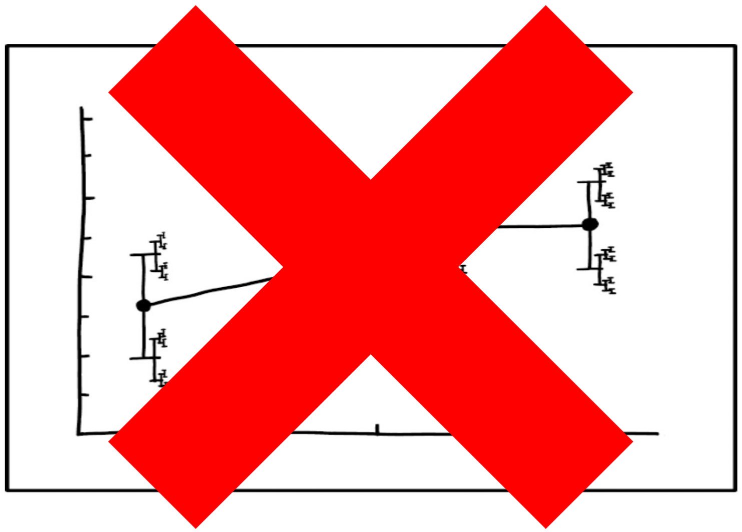


Formulation of the problem

$$\log L(\boldsymbol{\mu}, \boldsymbol{\theta}) = \log P(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\theta}) - \sum \frac{(\boldsymbol{u_i} - \boldsymbol{\theta_i})^2}{2\boldsymbol{\sigma_{u_i}^2}}$$

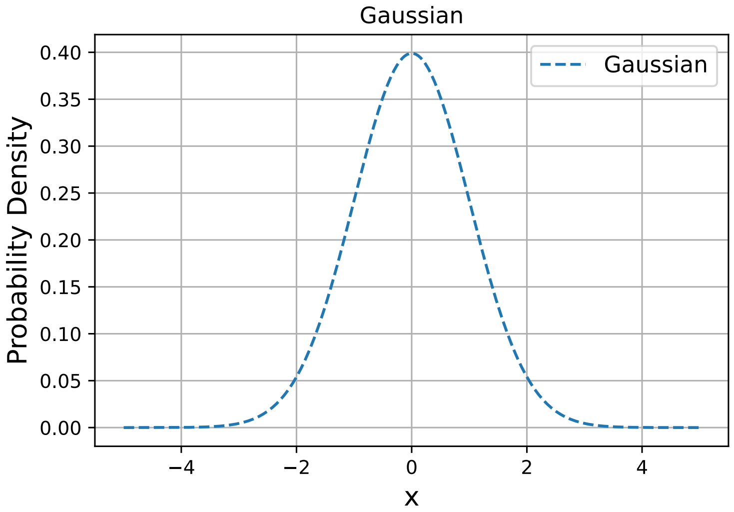
Let systematic errors be potentially uncertain!

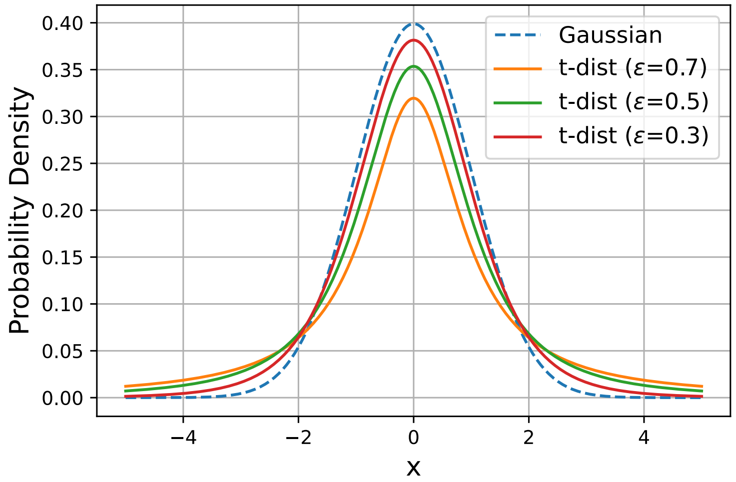
- We do not literally introduce error bars on errors bars



Formulation of the problem

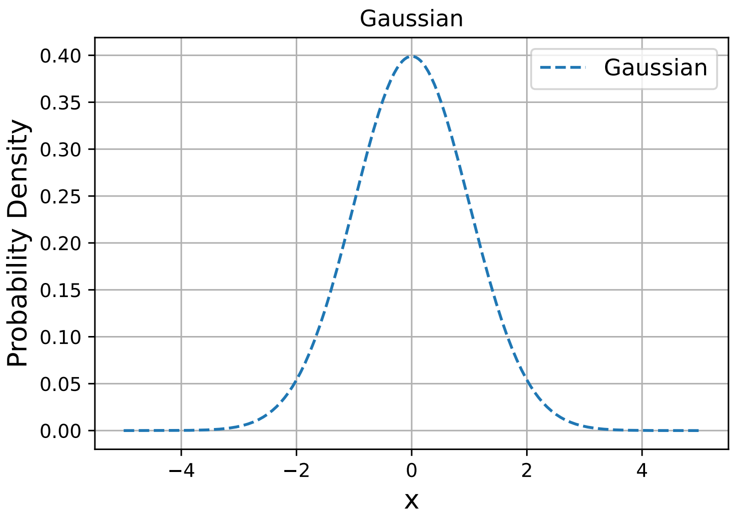
- Effectively, we modify the distribution for modelling systematic effects

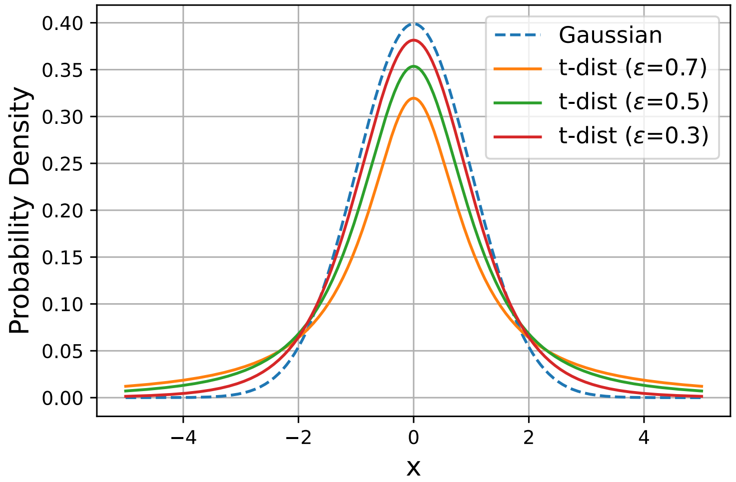




Formulation of the problem

- Effectively, we modify the distribution for modelling systematic effects





- We use a Student’s t-distribution instead of a Gaussian distribution.

Gamma Variance Model (GVM)

- The original **quadratic terms** in the log likelihood replaced by a **logarithmic terms:**

$$\sum_i \frac{(u_i - \theta_i)^2}{2\sigma_{u_i}^2}$$

Gamma Variance Model (GVM)

- The original **quadratic terms** in the log likelihood replaced by a **logarithmic terms:**

$$\sum_i \frac{(u_i - \theta_i)^2}{2\sigma_{u_i}^2} \longrightarrow \sum_i \frac{1}{2}\left(1 + \frac{1}{2\varepsilon_i^2}\right)\log\left(1 + 2\varepsilon_i^2 \frac{(u_i - \theta_i)^2}{\sigma_{u_i}^2}\right)$$

ε = error-on-error parameter

ε = 0.3 means 30% uncertainty on systematic error σ_{u_i}

Gamma Variance Model (GVM)

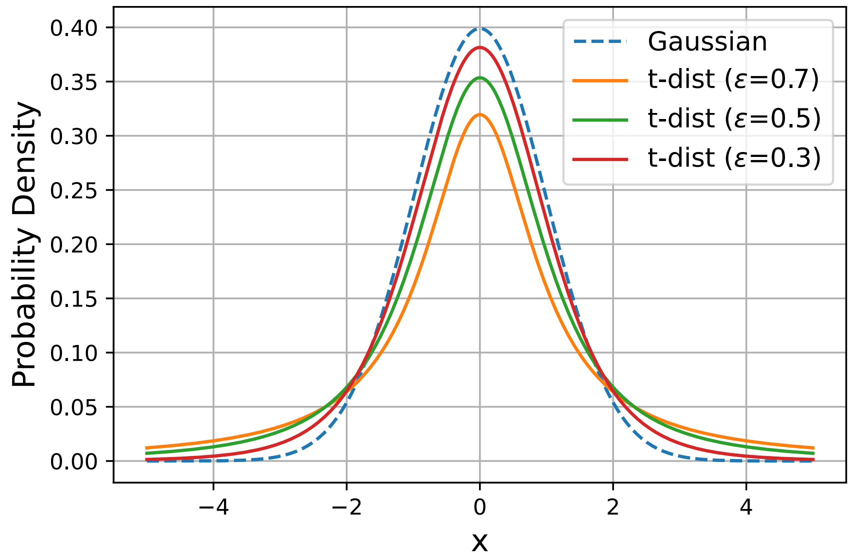
- The original **quadratic terms** in the log likelihood replaced by a **logarithmic terms:**

$$\sum_i \frac{(u_i - \theta_i)^2}{2\sigma_{u_i}^2} \longrightarrow \sum_i \frac{1}{2}\left(1 + \frac{1}{2\varepsilon_i^2}\right)\log\left(1 + 2\varepsilon_i^2 \frac{(u_i - \theta_i)^2}{\sigma_{u_i}^2}\right)$$

ε = error-on-error parameter

ε = 0.3 means 30% uncertainty on systematic error σ_{u_i}

- The student t distribution has longer tails for larger values of ε
- Outliers are **de-weighted** in the fit/combination



For more details: [Eur. Phys. J. C (2019) 79:133]

Formal interpretation

- The Updated log-likelihood

$$\log L_P(\boldsymbol{\mu}, \boldsymbol{\theta}) = \log P(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\theta}) - \frac{1}{2}\sum_i \left(\mathbf{1} + \frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\varepsilon}_i^2}\right)\mathbf{log}\left(\mathbf{1} + \mathbf{2}\boldsymbol{\varepsilon}_i^2 \frac{(\boldsymbol{u}_i - \boldsymbol{\theta}_i)^2}{\boldsymbol{v}_i}\right)$$

Formal interpretation

ROYAL HOLLOWAY UNIVERSITY OF LONDON

- The Updated log-likelihood

$$\log L_P(\boldsymbol{\mu}, \boldsymbol{\theta}) \;=\; \log P(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\theta}) - \frac{1}{2}\sum_i \left(\mathbf{1} + \frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\varepsilon}_i^2}\right) \mathbf{log}\left(\mathbf{1} + \mathbf{2}\boldsymbol{\varepsilon}_i^2 \frac{(\boldsymbol{u}_i - \boldsymbol{\theta}_i)^2}{\boldsymbol{v}_i}\right)$$

is a profile model, of

$$L(\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\sigma}_{\boldsymbol{u}_i}^2) = P(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\theta}) \times \prod_i \frac{1}{\sqrt{2\pi}\sigma_{u_i}} e^{-(u_i-\theta_i)^2/2\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2} \times \frac{\boldsymbol{\beta}_i^{\boldsymbol{\alpha}_i}}{\boldsymbol{\Gamma}(\boldsymbol{\alpha}_i)} \boldsymbol{v}_i^{\boldsymbol{\alpha}_i - \mathbf{1}} \boldsymbol{e}^{-\boldsymbol{\beta}_i \boldsymbol{v}_i}$$

Formal interpretation

- The Updated log-likelihood

$$\log L_P(\boldsymbol{\mu},\boldsymbol{\theta}) = \log P(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\theta}) - \frac{1}{2}\sum_i \left(\mathbf{1} + \frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\varepsilon}_i^2}\right)\mathbf{log}\left(\mathbf{1} + \mathbf{2}\boldsymbol{\varepsilon}_i^2 \frac{(\boldsymbol{u}_i - \boldsymbol{\theta}_i)^2}{\boldsymbol{v}_i}\right)$$

is a profile model, of

$$L(\boldsymbol{\mu},\boldsymbol{\theta},\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2) = P(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\theta}) \times \prod_i \frac{1}{\sqrt{2\pi}\sigma_{u_i}} e^{-(u_i-\theta_i)^2/2\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2} \times \frac{\boldsymbol{\beta}_i^{\boldsymbol{\alpha}_i}}{\boldsymbol{\Gamma}(\boldsymbol{\alpha}_i)} \boldsymbol{v}_i^{\boldsymbol{\alpha}_i - \mathbf{1}} \boldsymbol{e}^{-\boldsymbol{\beta}_i \boldsymbol{v}_i}$$

- $\boldsymbol{v}_i = \boldsymbol{E}[\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2]$ is the best estimate of $\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2$

- $\boldsymbol{v}_i \sim \frac{\boldsymbol{\beta}_i^{\boldsymbol{\alpha}_i}}{\boldsymbol{\Gamma}(\boldsymbol{\alpha}_i)} \boldsymbol{v}_i^{\boldsymbol{\alpha}_i - \mathbf{1}} \boldsymbol{e}^{-\boldsymbol{\beta}_i \boldsymbol{v}_i}$ is Gamma distributed

*In this interpretation $\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2$ is replaced by $\boldsymbol{v}_i$ in the profile likelihood

Formal interpretation

- The Updated log-likelihood

$$\log L_P(\boldsymbol{\mu},\boldsymbol{\theta}) = \log P(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\theta}) - \frac{1}{2}\sum_i \left(\mathbf{1}+\frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\varepsilon}_i^2}\right)\mathbf{log}\left(\mathbf{1}+\mathbf{2}\boldsymbol{\varepsilon}_i^2\frac{(\boldsymbol{u}_i-\boldsymbol{\theta}_i)^2}{\boldsymbol{v}_i}\right)$$

is a profile model, of

$$L(\boldsymbol{\mu},\boldsymbol{\theta},\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2) = P(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\theta}) \times \prod_i \frac{1}{\sqrt{2\pi}\sigma_{u_i}} e^{-(u_i-\theta_i)^2/2\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2} \times \frac{\boldsymbol{\beta}_i^{\boldsymbol{\alpha}_i}}{\boldsymbol{\Gamma}(\boldsymbol{\alpha}_i)} \boldsymbol{v}_i^{\boldsymbol{\alpha}_i-\mathbf{1}} \boldsymbol{e}^{-\boldsymbol{\beta}_i\boldsymbol{v}_i}$$

Gamma Variance Model

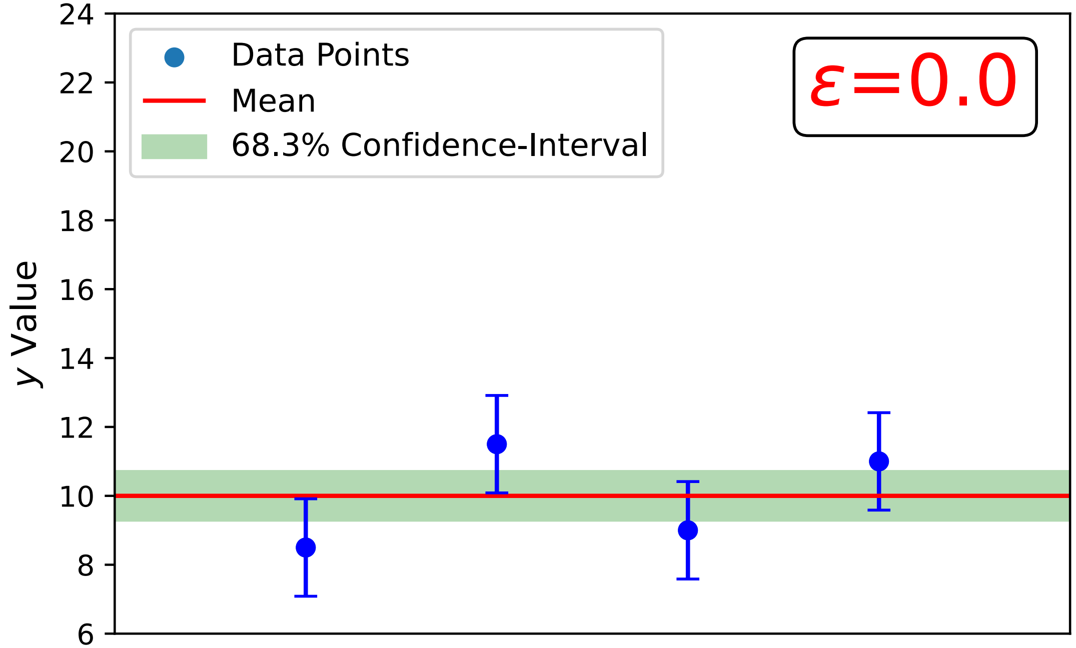
- $\boldsymbol{v}_i = \boldsymbol{E}[\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2]$ is the best estimate of $\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2$
- $\boldsymbol{v}_i \sim \frac{\boldsymbol{\beta}_i^{\boldsymbol{\alpha}_i}}{\boldsymbol{\Gamma}(\boldsymbol{\alpha}_i)} \boldsymbol{v}_i^{\boldsymbol{\alpha}_i-\mathbf{1}} \boldsymbol{e}^{-\boldsymbol{\beta}_i\boldsymbol{v}_i}$ is Gamma distributed

*In this interpretation $\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2$ is replaced by $\boldsymbol{v}_i$ in the profile likelihood

Sensitivity to outliers

- Suppose we want to average 4 measurements
- **Syst errors** all have equal ***errors-on-errors*** $\boldsymbol{\varepsilon}$:

$$\log \mathrm{L}(\boldsymbol{\mu}, \boldsymbol{\theta}) = -\frac{1}{2}\sum_i \frac{(y_i - \mu - \theta_i)^2}{\sigma_{y_i}^2} - \frac{1}{2}\sum_i \left(1 + \frac{1}{2\varepsilon_i^2}\right) \log\left(1 + 2\varepsilon_i^2 \frac{(u_i - \theta_i)^2}{\sigma_{u_i}^2}\right)$$

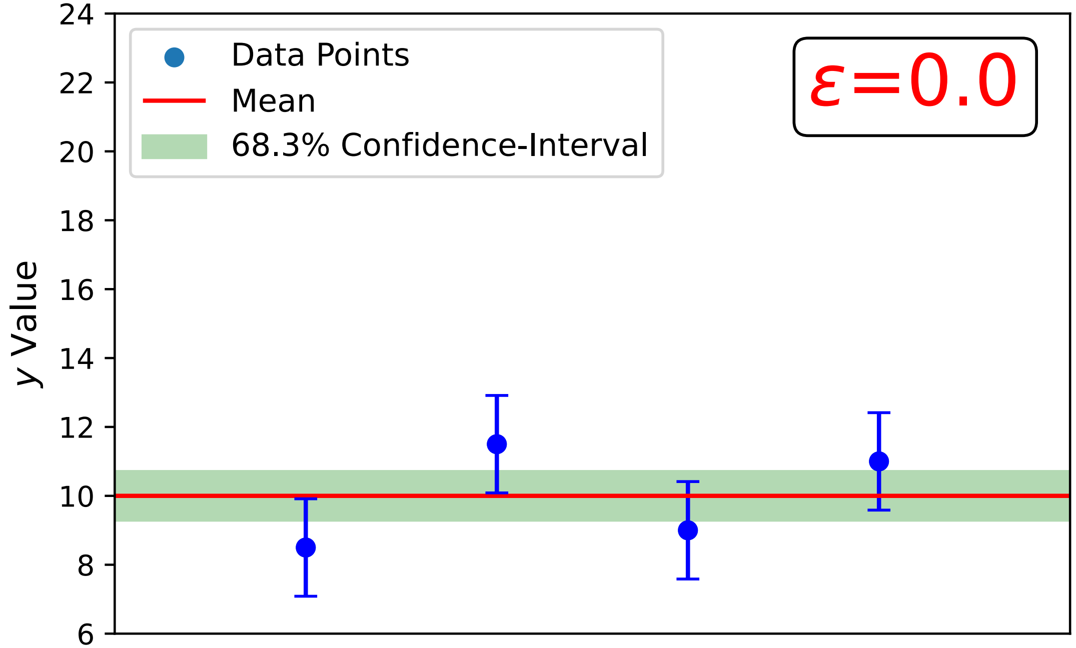


Sensitivity to outliers

- Suppose we want to average 4 measurements
- **Syst errors** all have equal ***errors-on-errors*** $\boldsymbol{\varepsilon}$:

$$\log \mathrm{L}(\boldsymbol{\mu}, \boldsymbol{\theta}) = -\frac{1}{2}\sum_i \frac{(\boldsymbol{y_i} - \boldsymbol{\mu} - \boldsymbol{\theta_i})^2}{\boldsymbol{\sigma_{y_i}^2}} - \frac{1}{2}\sum_i \left(1 + \frac{1}{2\boldsymbol{\varepsilon}_i^2}\right)\log\left(1 + 2\boldsymbol{\varepsilon}_i^2 \frac{(u_i - \theta_i)^2}{\sigma_{u_i}^2}\right)$$

Measurements internally compatible

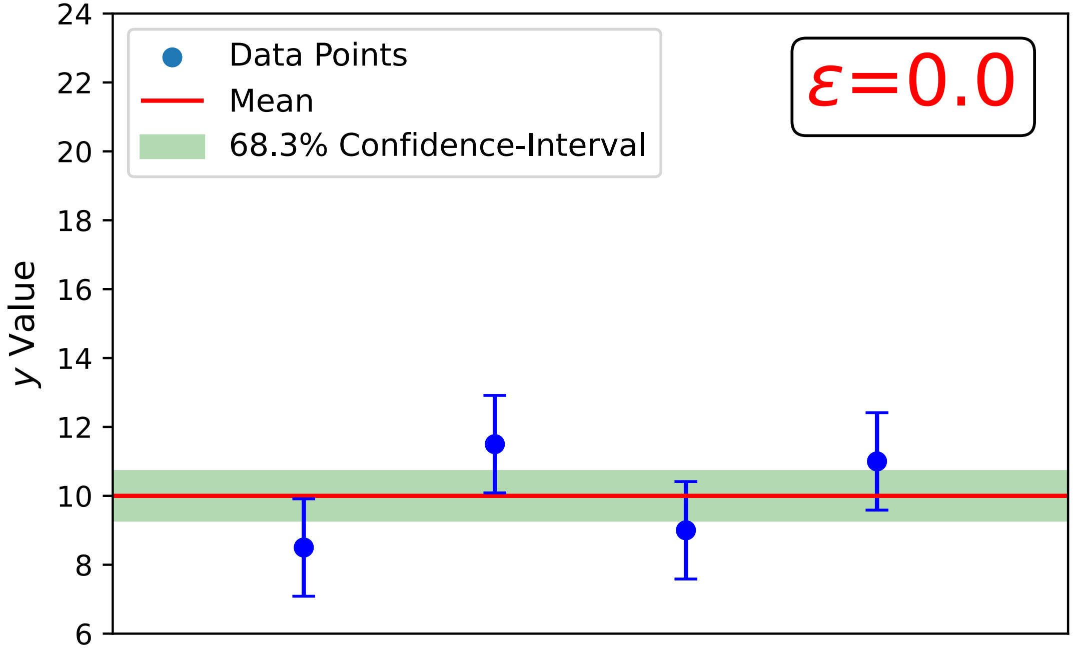


Sensitivity to outliers

- Suppose we want to average 4 measurements
- **Syst errors** all have equal ***errors-on-errors*** $\boldsymbol{\varepsilon}$:

$$\log \mathrm{L}(\boldsymbol{\mu}, \boldsymbol{\theta}) = -\frac{1}{2}\sum_i \frac{(\boldsymbol{y_i} - \boldsymbol{\mu} - \boldsymbol{\theta_i})^2}{\boldsymbol{\sigma^2_{y_i}}} - \frac{1}{2}\sum_i \left(1 + \frac{1}{2\boldsymbol{\varepsilon}_i^2}\right)\log\left(1 + 2\boldsymbol{\varepsilon}_i^2 \frac{(u_i - \theta_i)^2}{\sigma^2_{u_i}}\right)$$

Measurements internally compatible



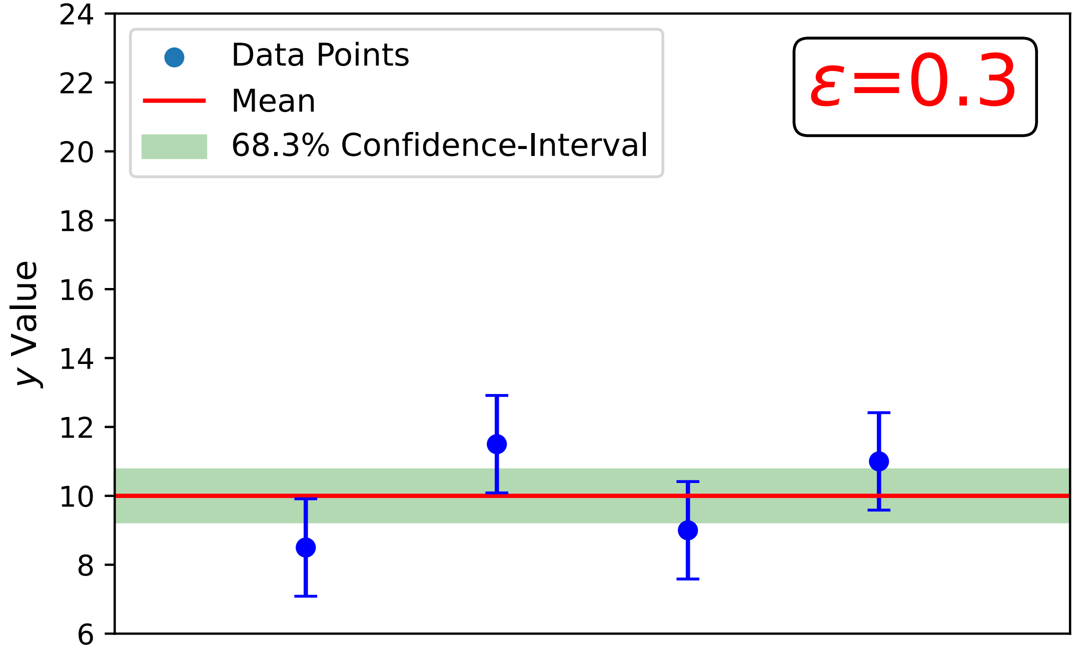
Let's increase $\boldsymbol{\varepsilon}$!

Sensitivity to outliers

- Suppose we want to average 4 measurements
- **Syst errors** all have equal ***errors-on-errors*** $\boldsymbol{\varepsilon}$:

$$\log \mathrm{L}(\boldsymbol{\mu}, \boldsymbol{\theta}) = -\frac{1}{2}\sum_i \frac{(y_i - \mu - \theta_i)^2}{\sigma_{y_i}^2} - \frac{1}{2}\sum_i \left(1 + \frac{1}{2\varepsilon_i^2}\right) \log\left(1 + 2\varepsilon_i^2 \frac{(u_i - \theta_i)^2}{\sigma_{u_i}^2}\right)$$

Measurements internally compatible



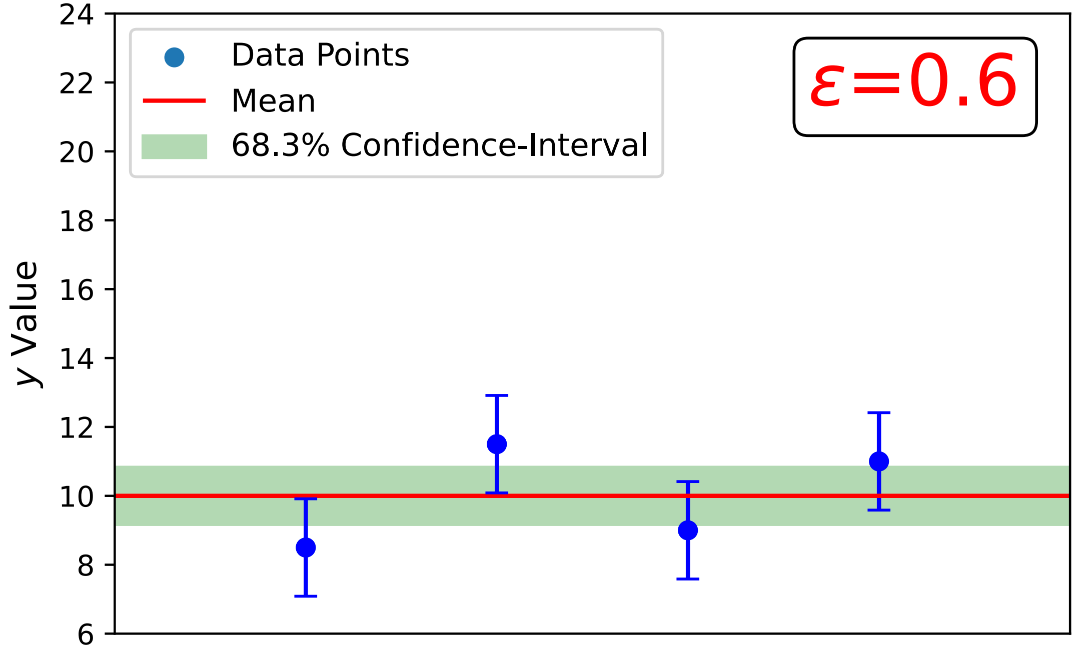
Let's increase ε!

Sensitivity to outliers

- Suppose we want to average 4 measurements
- **Syst errors** all have equal ***errors-on-errors*** $\boldsymbol{\varepsilon}$:

$$\log \mathrm{L}(\boldsymbol{\mu}, \boldsymbol{\theta}) = -\frac{1}{2}\sum_i \frac{(y_i - \mu - \theta_i)^2}{\sigma_{y_i}^2} - \frac{1}{2}\sum_i \left(1 + \frac{1}{2\varepsilon_i^2}\right) \log\left(1 + 2\varepsilon_i^2 \frac{(u_i - \theta_i)^2}{\sigma_{u_i}^2}\right)$$

Measurements internally compatible



Let's increase ε!

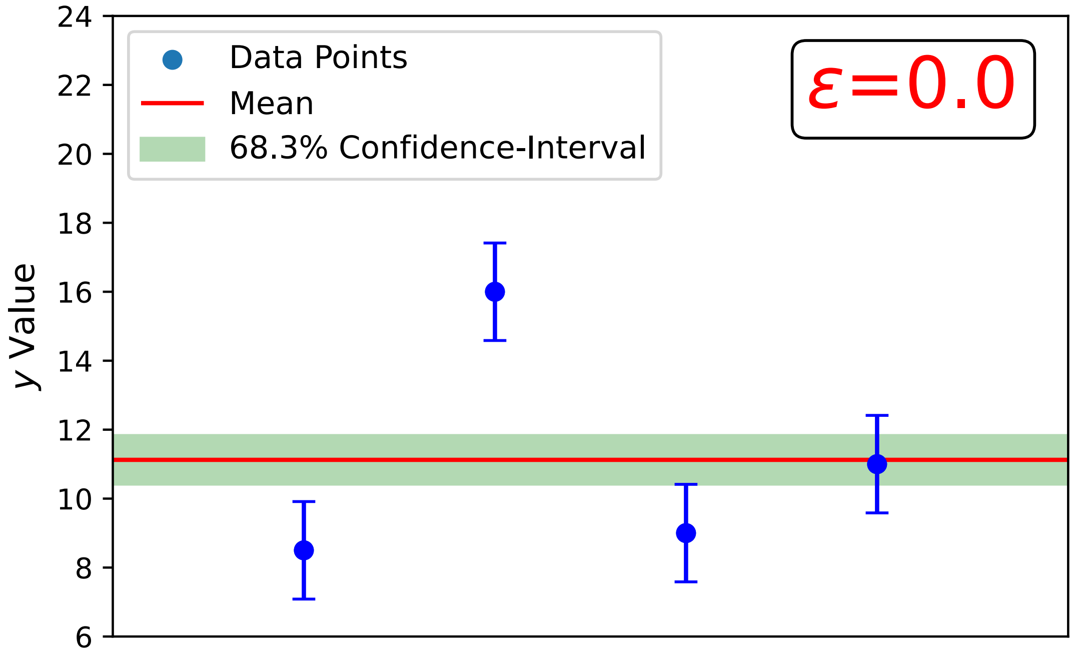
Sensitivity to outliers

ROYAL HOLLOWAY UNIVERSITY OF LONDON

If data are internally compatible results are only slightly modified

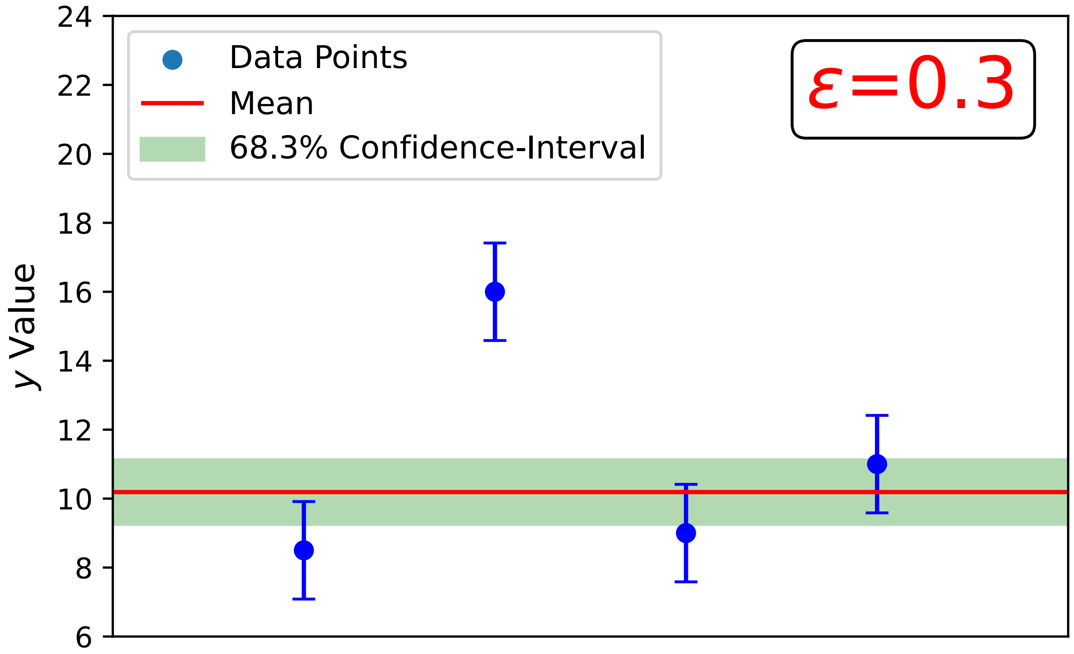
Sensitivity to outliers

- Suppose one of the measurements is an outlier
- If data are internally incompatible important changes can be observed



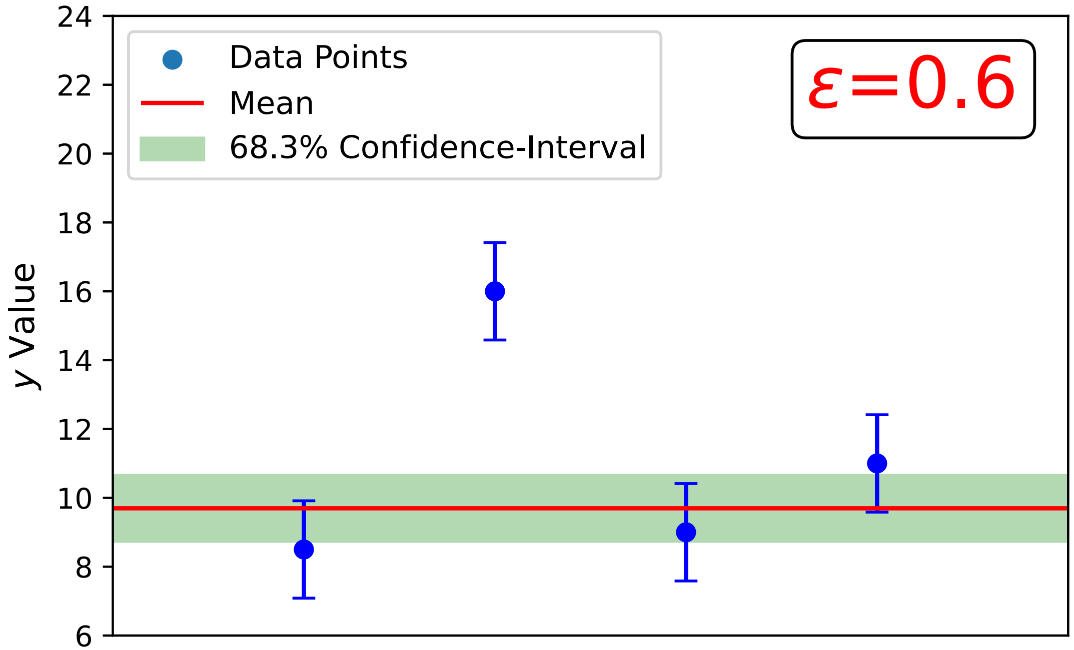
Sensitivity to outliers

- Suppose one of the measurements is an outlier
- If data are internally incompatible important changes can be observed

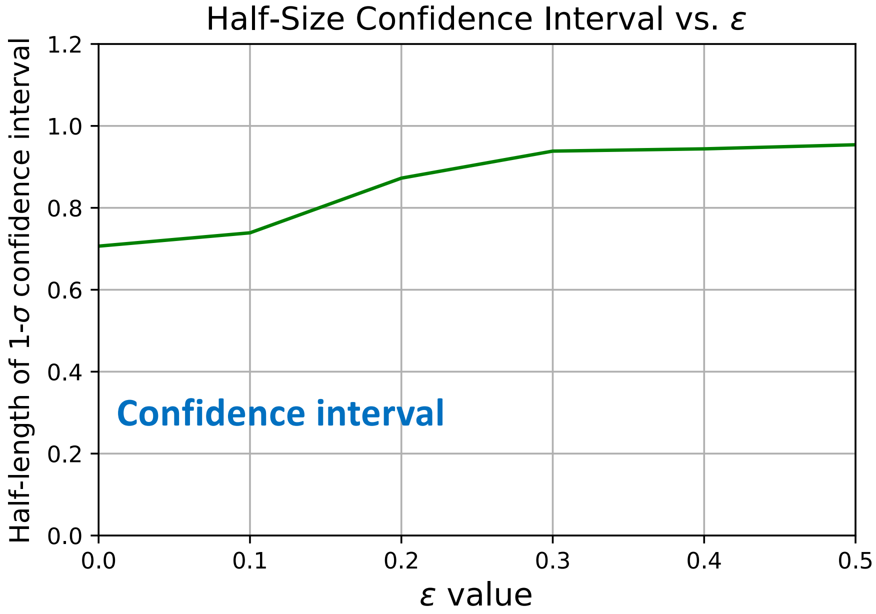


Sensitivity to outliers

- Suppose one of the measurements is an outlier
- If data are internally incompatible important changes can be observed



Sensitivity to outliers



1. The outlier is de-weighted

2. **The confidence interval grows**: the model treats internal incompatibility as an additional source of uncertainty

Conclusion: The model is sensitive to internal compatibility of the data

Beyond Wilks' theorem

- We use the likelihood to construct test statistics to compute CI (Confidence Intervals) or evaluate GOFs (Goodness-Of-Fits):

- Profile likelihood ratio (CI):

$$w_{\mu} = 2[\log L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\theta}}) - \log L(\boldsymbol{\mu}, \hat{\hat{\boldsymbol{\theta}}})]$$

- goodness-of-fit (chi2):

$$q = -2 \log L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\theta}})$$

Beyond Wilks' theorem

- To compute CIs and GOFs often asymptotics properties are used:

*P = number POIs
N = number of measurements

$$w_\mu \sim \chi^2_P$$

$$q \sim \chi^2_{N-P}$$

- **Problem**: Our likelihood is not quadratic; test statistics deviate from asymptotic formulas by $\mathcal{O}(\epsilon^2)$**,** as the likelihood is not anymore quadratic because of the presence of terms like $\left(1+\frac{1}{2\varepsilon_i^2}\right)\log\left(1+2\varepsilon_i^2\frac{(u_i-\theta_i)^2}{\sigma_{u_i}^2}\right)$

- **Equivalence:** Equivalent problem of having a small sample size: $n_{eff} = 1+\frac{1}{2\varepsilon^2}$

Effective sample size

Bartlett Corrections

Modify the test statistic t so that its distribution is closer to a χ^2 :

[Bartlett](#)
[Gauss M. Cordeiro, Francisco Cribari-Neto](#)

$$t \longrightarrow t^* = t\frac{N_{dof}}{E[t]}$$

$$t \sim \chi^2 + \mathcal{O}(\epsilon^2)$$

$$t^* \sim \chi^2 + \mathcal{O}(\epsilon^4)$$

[Canonero, Cowan, Brazzale](#)

Bartlett Corrections

ROYAL HOLLOWAY UNIVERSITY OF LONDON

Modify the test statistic t so that its distribution is closer to a χ^2 :

[Bartlett](#)
[Gauss M. Cordeiro, Francisco Cribari-Neto](#)

$$t \longrightarrow t^* = t\frac{N_{dof}}{E[t]}$$

Expectation value in the asymptotic limit (degrees of freedom of χ^2)

Exact expectation value

$$t \sim \chi^2 + \mathcal{O}(\epsilon^2)$$

$$t^* \sim \chi^2 + \mathcal{O}(\epsilon^4)$$

[Canonero, Cowan, Brazzale](#)

Bartlett Corrections

Modify the test statistic t so that its distribution is closer to a χ^2 :

Bartlett
Gauss M. Cordeiro, Francisco Cribari-Neto

$$t \longrightarrow t^* = t \frac{N_{dof}}{E[t]}$$

Expectation value in the asymptotic limit (degrees of freedom of χ^2)

Exact expectation value

$$t \sim \chi^2 + \mathcal{O}(\epsilon^2)$$

$$t^* \sim \chi^2 + \mathcal{O}(\epsilon^4)$$

Canonero, Cowan, Brazzale

***Bartlett correction originally developed for small sample sizes problems**

Bartlett Corrections

ROYAL HOLLOWAY UNIVERSITY OF LONDON

- All these expectation values can be computed analytically up to $\mathcal{O}(\epsilon^4)$:

$$E[q] = N_{dof} + b_q$$

$$b_q = \sum_s \left(3 - 4\frac{\sigma_{\hat{\theta}_s}^2}{\sigma_s^2} - \frac{\sigma_{\hat{\theta}_s}^4}{\sigma_s^4} \right) \epsilon_s^2 + \mathcal{O}(\epsilon^4)$$

- We have recently computed $E[q]$ and $E[\omega_\mu]$ up to $\mathcal{O}(\epsilon^4)$

Correlations in Combinations - Recap

- BLUE (Best Linear Unbiased Estimators) approach to fits:

$$-2logL = \chi^2$$

$$\chi^2 = \sum_i (y_i - \mu) W_{ij}^{-1} (y_j - \mu)$$

$$W_{ij} = \boldsymbol{V}_{ij} + \boldsymbol{U}_{ij}^{(syst)}$$

Correlations in Combinations - Recap

- BLUE (Best Linear Unbiased Estimators) approach to fits:

$$-2logL = \chi^2$$

$$\chi^2 = \sum_i (y_i - \mu) W_{ij}^{-1} (y_j - \mu)$$

$$W_{ij} = \boldsymbol{V}_{ij} + \boldsymbol{U}_{ij}^{(syst)}$$

- $\boldsymbol{V}_{ij}$: Statistical covariance matrix.

Correlations in Combinations - Recap

- BLUE (Best Linear Unbiased Estimators) approach to fits:

$$-2logL = \chi^2$$

$$\chi^2 = \sum_i (y_i - \mu) W_{ij}^{-1} (y_j - \mu)$$

$$W_{ij} = V_{ij} + U_{ij}^{(syst)}$$

- V_{ij}: Statistical covariance matrix.
- $U_{ij}^{(syst)}$: Covariance matrix induced by systematic source.
- $U_{ij}^{(syst)} = \sum_s U_{ij}^{(s)}$

Correlations in Combinations - Recap

- Nuisance parameters approach:

$$\chi^2 = \sum_{ij} \left(y_i - \mu - \sum_s \boldsymbol{\Gamma}_i^s \theta_s \right) V_{ij}^{-1} \left(y_j - \mu - \sum_s \boldsymbol{\Gamma}_j^s \theta_s \right) + \sum_s \frac{(\boldsymbol{u_s} - \theta_s)^2}{\boldsymbol{\sigma_s^2}}$$

- $\boldsymbol{u_s}$ set to 0: know biases already removed
- $\boldsymbol{\sigma_s^2}$ set to 1: absorbed in the definition of $\boldsymbol{\Gamma}$

Correlations in Combinations - Recap

- Nuisance parameters approach:

$$\chi^2 = \sum_{ij} \left(y_i - \mu - \sum\nolimits_s \boldsymbol{\Gamma}_i^s \theta_s \right) V_{ij}^{-1} \left(y_j - \mu - \sum\nolimits_s \boldsymbol{\Gamma}_j^s \theta_s \right) + \sum_s \frac{(\boldsymbol{u_s} - \theta_s)^2}{\boldsymbol{\sigma_s^2}}$$

- $\boldsymbol{u_s}$ set to 0: know biases already removed
- $\boldsymbol{\sigma_s^2}$ set to 1: absorbed in the definition of $\boldsymbol{\Gamma}$

- Connection: $\boldsymbol{U}_{ij}^{(\boldsymbol{s})} = \boldsymbol{\sigma_s^2}\, \boldsymbol{\Gamma}_i^s \boldsymbol{\Gamma}_j^s$

Correlations in Combinations - Recap

- Nuisance parameters approach:

$$\chi^2 = \sum_{ij} \left(y_i - \mu - \sum_s \boldsymbol{\Gamma}_i^s \theta_s \right) V_{ij}^{-1} \left(y_j - \mu - \sum_s \boldsymbol{\Gamma}_j^s \theta_s \right) + \sum_s \frac{(\boldsymbol{u_s} - \theta_s)^2}{\boldsymbol{\sigma_s^2}}$$

- $\boldsymbol{u_s}$ set to 0: know biases already removed
- $\boldsymbol{\sigma_s^2}$ set to 1: absorbed in the definition of $\boldsymbol{\Gamma}$

- Connection: $\boldsymbol{U_{ij}^{(s)}} = \boldsymbol{\sigma_s^2}\, \boldsymbol{\Gamma_i^s \Gamma_j^s}$

- Systematic uncertainties can induce ± 1 correlations: $\rho_{ij}^{(s)} = \frac{V_{ij}^{(s)}}{\sqrt{V_{ii}^{(s)} V_{jj}^{(s)}}} = \frac{\Gamma_i^s \Gamma_j^s}{|\Gamma_i^s||\Gamma_j^s|} = \pm 1$

Non-Trivial Correlations in Combinations

- What to do if you want to use non-trivial correlation assumptions? $\rho_{ij}^{(s)} \neq \pm 1$

Non-Trivial Correlations in Combinations

- What to do if you want to use non-trivial correlation assumptions? $\rho_{ij}^{(s)} \neq \pm 1$
- Let's do an easy example with only one systematic $\boldsymbol{\theta_s}$

$$\chi^2 = \sum_{ij} (y_i - \mu - \Gamma_i^s \boldsymbol{\theta_s}) V_{ij}^{-1} \left(y_j - \mu - \Gamma_i^s \boldsymbol{\theta_s}\right) + \frac{(u_s - \boldsymbol{\theta_s})^2}{\sigma_s^2}$$

Non-Trivial Correlations in Combinations

- What to do if you want to use non-trivial correlation assumptions? $\rho_{ij}^{(s)} \neq \pm 1$

- Let's do an easy example with only one systematic $\boldsymbol{\theta_s}$

$$\chi^2 = \sum_{ij} (y_i - \mu - \Gamma_i^s \boldsymbol{\theta_s}) V_{ij}^{-1} \left(y_j - \mu - \Gamma_i^s \boldsymbol{\theta_s} \right) + \frac{(u_s - \boldsymbol{\theta_s})^2}{\sigma_s^2}$$

- Define new parameters $\boldsymbol{\theta_s^{(i)}}$

Non-Trivial Correlations in Combinations

- What to do if you want to use non-trivial correlation assumptions? $\rho_{ij}^{(s)} \neq \pm 1$

- Let's do an easy example with only one systematic $\boldsymbol{\theta_s}$

$$\chi^2 = \sum_{ij} (y_i - \mu - \Gamma_i^s \boldsymbol{\theta_s}) V_{ij}^{-1} (y_j - \mu - \Gamma_i^s \boldsymbol{\theta_s}) + \frac{(u_s - \boldsymbol{\theta_s})^2}{\sigma_s^2}$$

- Define new parameters $\boldsymbol{\theta_s^{(i)}}$

$$\chi^2 = \sum_{ij} \left(y_i - \mu - \Gamma_i^s \boldsymbol{\theta_s^{(i)}}\right) V_{ij}^{-1} \left(y_j - \mu - \Gamma_j^s \boldsymbol{\theta_s^{(j)}}\right) + \sum_{ij} (u_s^{(i)} - \boldsymbol{\theta_s^{(i)}}) \boldsymbol{C_{ij}^{(s)}}^{-1} (u_s^{(j)} - \boldsymbol{\theta_s^{(j)}})$$

Non-Trivial Correlations in Combinations

$$\chi^2 = \sum_{ij} \left(y_i - \mu - \Gamma_i^s \boldsymbol{\theta}_{\boldsymbol{s}}^{(i)} \right) V_{ij}^{-1} \left(y_j - \mu - \Gamma_j^s \boldsymbol{\theta}_{\boldsymbol{s}}^{(j)} \right) + \sum_{ij} (u_s^{(i)} - \boldsymbol{\theta}_{\boldsymbol{s}}^{(i)}) {\boldsymbol{C}_{\boldsymbol{ij}}^{(\boldsymbol{s})}}^{-1} (u_s^{(j)} - \boldsymbol{\theta}_{\boldsymbol{s}}^{(j)})$$

- The new matrix $\boldsymbol{C}_{\mathrm{ij}}^{(\boldsymbol{s})}$ is defined as

$$\boldsymbol{C}_{\mathrm{ij}}^{(\boldsymbol{s})} = \boldsymbol{\sigma}_{\boldsymbol{s}}^{\mathbf{2}} \, \rho_{ij}^{(s)}$$

Non-Trivial Correlations in Combinations

ROYAL HOLLOWAY UNIVERSITY OF LONDON

$$\chi^2 = \sum_{ij} \left(y_i - \mu - \Gamma_i^s \boldsymbol{\theta_s^{(i)}}\right) V_{ij}^{-1} \left(y_j - \mu - \Gamma_j^s \boldsymbol{\theta_s^{(j)}}\right) + \sum_{ij} (u_s^{(i)} - \boldsymbol{\theta_s^{(i)}}) \boldsymbol{C_{ij}^{(s)}}^{-1} (u_s^{(j)} - \boldsymbol{\theta_s^{(j)}})$$

- The new matrix $\boldsymbol{C}_{\mathrm{ij}}^{(s)}$ is defined as

$$\boldsymbol{C}_{\mathrm{ij}}^{(s)} = \boldsymbol{\sigma_s^2}\, \rho_{ij}^{(s)}$$

- Generalization to errors-on-errors framework:

$$\chi^2 = \sum_{i} \left(y_i - \mu - \Gamma_i^s \boldsymbol{\theta_s^{(i)}}\right) V_{ij}^{-1} \left(y_j - \mu - \Gamma_j^s \boldsymbol{\theta_s^{(j)}}\right) + \left(\boldsymbol{N} + \frac{\boldsymbol{1}}{\boldsymbol{2\epsilon_s^2}}\right) \boldsymbol{log}\left(\boldsymbol{1} + \boldsymbol{2\epsilon_s^2} \sum_{ij} (u_s^{(i)} - \boldsymbol{\theta_s^{(i)}}) \boldsymbol{C_{ij}^{(s)}}^{-1} (u_s^{(j)} - \boldsymbol{\theta_s^{(j)}})\right)$$

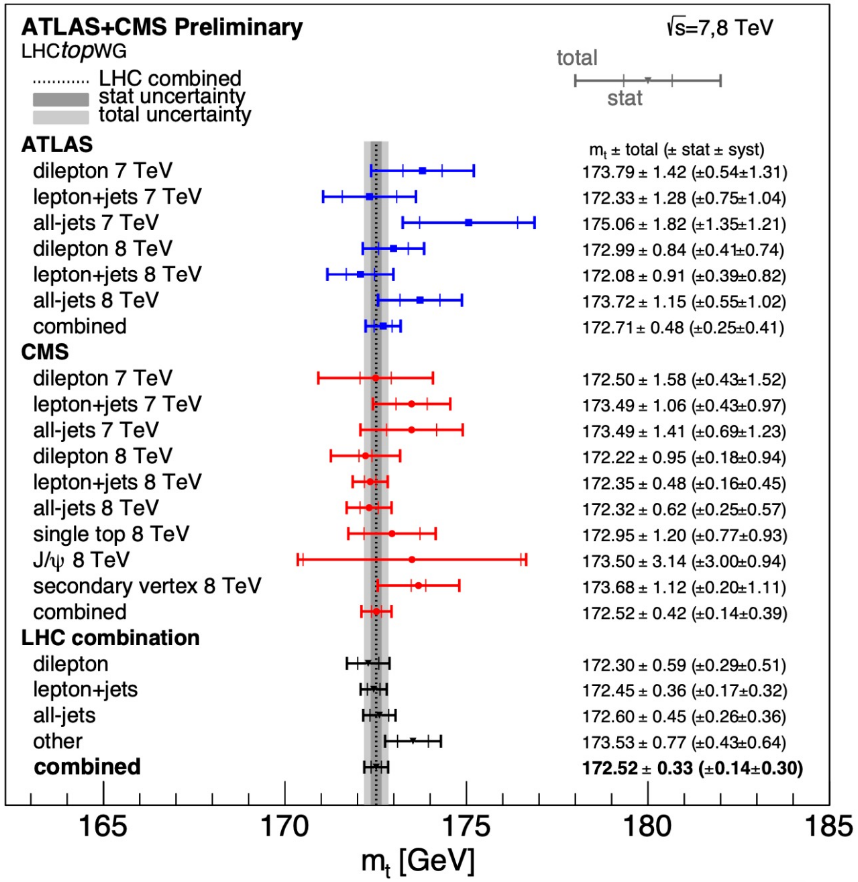
New paper for more details! [arXiv: 2407.05322]

7-8 TeV ATLAS-CMS top-quark mass combination

Goals

- The combination was performed with the BLUE approach:

→ Reproduce it with the nuisance parameters approach described earlier



Combination paper: [Phys. Rev. Lett. 132 (2024) 261902]

Our in parallel analysis (using public results): [arXiv: 2407.05322]

7-8 TeV ATLAS-CMS top-quark mass combination

ROYAL HOLLOWAY UNIVERSITY OF LONDON

Goals

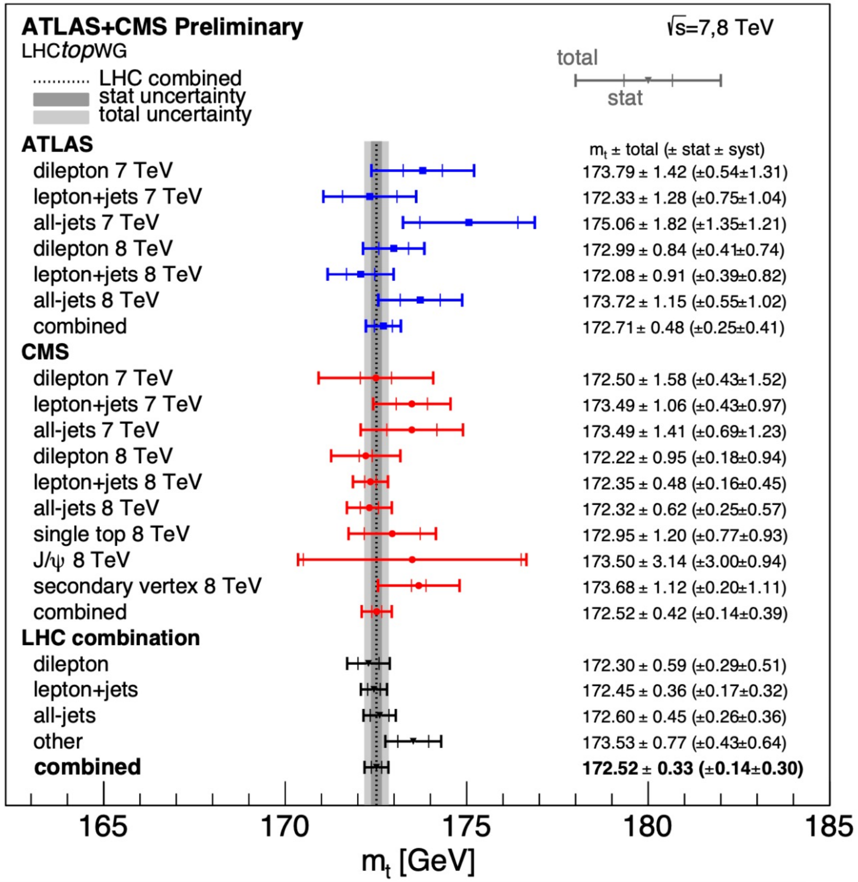
- The combination was performed with the BLUE approach:

→ Reproduce it with the nuisance parameters approach described earlier

---

- Top mass measurements are becoming systematics dominated
- Potentially affected by QCD modelling systematics

→ Study errors-on-errors impact on the combination



Combination paper: [Phys. Rev. Lett. 132 (2024) 261902]

Our in parallel analysis (using public results): [arXiv: 2407.05322]

7-8 TeV ATLAS-CMS top-quark mass combination

ROYAL HOLLOWAY UNIVERSITY OF LONDON

Robustness-test: Check if the combination is sensitive when any major systematic uncertainty is itself uncertain

Uncertainty category	Uncertainty impact [GeV]		
	LHC	ATLAS	CMS
LHC b-JES	0.18	0.17	0.25
b tagging	0.09	0.16	0.03
ME generator	0.08	0.13	0.14
LHC JES 1	0.08	0.18	0.06
LHC JES 2	0.08	0.11	0.10
Method	0.07	0.06	0.09
CMS B hadron BR	0.07	—	0.12
LHC radiation	0.06	0.07	0.10
Leptons	0.05	0.08	0.07
JER	0.05	0.09	0.02
Top quark p_T	0.05	—	0.07
Background (data)	0.05	0.04	0.06
Color reconnection	0.04	0.08	0.03
Underlying event	0.04	0.03	0.05
LHC g-JES	0.03	0.02	0.04
Background (MC)	0.03	0.07	0.01
Other	0.03	0.06	0.01
LHC l-JES	0.03	0.01	0.05
CMS JES 1	0.03	—	0.04
Pileup	0.03	0.07	0.03
LHC JES 3	0.02	0.07	0.01
LHC hadronization	0.02	0.01	0.01
p_T^{miss}	0.02	0.04	0.01
PDF	0.02	0.06	<0.01
Trigger	0.01	0.01	0.01
Total systematics	0.30	0.41	0.39
Statistical	0.14	0.25	0.14
Total	0.33	0.48	0.42

From: arXiv:2402.08713

7-8 TeV ATLAS-CMS top-quark mass combination

ROYAL HOLLOWAY UNIVERSITY OF LONDON

Robustness-test: Check if the combination is sensitive when any major systematic uncertainty is itself uncertain

1. Treat the eight largest systematic uncertainties as potentially uncertain one at a time.

2. Assign an error-on-error parameter ε_s to each one.

3. Study how varying each ε_s individually affects the central value and confidence interval.

Uncertainty category	Uncertainty impact [GeV]		
	LHC	ATLAS	CMS
LHC b-JES	0.18	0.17	0.25
b tagging	0.09	0.16	0.03
ME generator	0.08	0.13	0.14
LHC JES 1	0.08	0.18	0.06
LHC JES 2	0.08	0.11	0.10
Method	0.07	0.06	0.09
CMS B hadron BR	0.07	—	0.12
LHC radiation	0.06	0.07	0.10
Leptons	0.05	0.08	0.07
JER	0.05	0.09	0.02
Top quark p_T	0.05	—	0.07
Background (data)	0.05	0.04	0.06
Color reconnection	0.04	0.08	0.03
Underlying event	0.04	0.03	0.05
LHC g-JES	0.03	0.02	0.04
Background (MC)	0.03	0.07	0.01
Other	0.03	0.06	0.01
LHC l-JES	0.03	0.01	0.05
CMS JES 1	0.03	—	0.04
Pileup	0.03	0.07	0.03
LHC JES 3	0.02	0.07	0.01
LHC hadronization	0.02	0.01	0.01
p_T^{miss}	0.02	0.04	0.01
PDF	0.02	0.06	<0.01
Trigger	0.01	0.01	0.01
Total systematics	0.30	0.41	0.39
Statistical	0.14	0.25	0.14
Total	0.33	0.48	0.42

From: arXiv:2402.08713

7-8 TeV ATLAS-CMS top-quark mass combination

ROYAL HOLLOWAY UNIVERSITY OF LONDON

Robustness-test: Check if the combination is sensitive when any major systematic uncertainty is itself uncertain

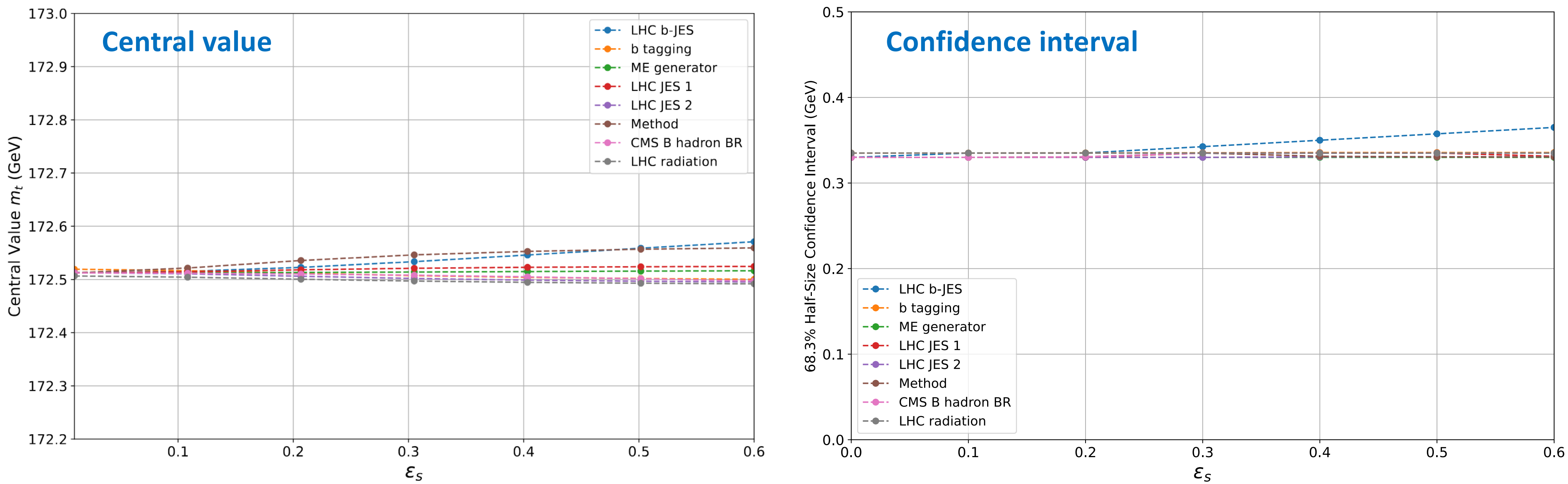
1. Treat the eight largest systematic uncertainties as potentially uncertain one at a time.
2. Assign an error-on-error parameter ε_s to each one.
3. Study how varying each ε_s individually affects the central value and confidence interval.

Goal: Show a possible way to use the errors-on-errors model in an analysis

Uncertainty category	Uncertainty impact [GeV]		
	LHC	ATLAS	CMS
LHC b-JES	0.18	0.17	0.25
b tagging	0.09	0.16	0.03
ME generator	0.08	0.13	0.14
LHC JES 1	0.08	0.18	0.06
LHC JES 2	0.08	0.11	0.10
Method	0.07	0.06	0.09
CMS B hadron BR	0.07	—	0.12
LHC radiation	0.06	0.07	0.10
Leptons	0.05	0.08	0.07
JER	0.05	0.09	0.02
Top quark p_T	0.05	—	0.07
Background (data)	0.05	0.04	0.06
Color reconnection	0.04	0.08	0.03
Underlying event	0.04	0.03	0.05
LHC g-JES	0.03	0.02	0.04
Background (MC)	0.03	0.07	0.01
Other	0.03	0.06	0.01
LHC l-JES	0.03	0.01	0.05
CMS JES 1	0.03	—	0.04
Pileup	0.03	0.07	0.03
LHC JES 3	0.02	0.07	0.01
LHC hadronization	0.02	0.01	0.01
p_T^{miss}	0.02	0.04	0.01
PDF	0.02	0.06	<0.01
Trigger	0.01	0.01	0.01
Total systematics	0.30	0.41	0.39
Statistical	0.14	0.25	0.14
Total	0.33	0.48	0.42

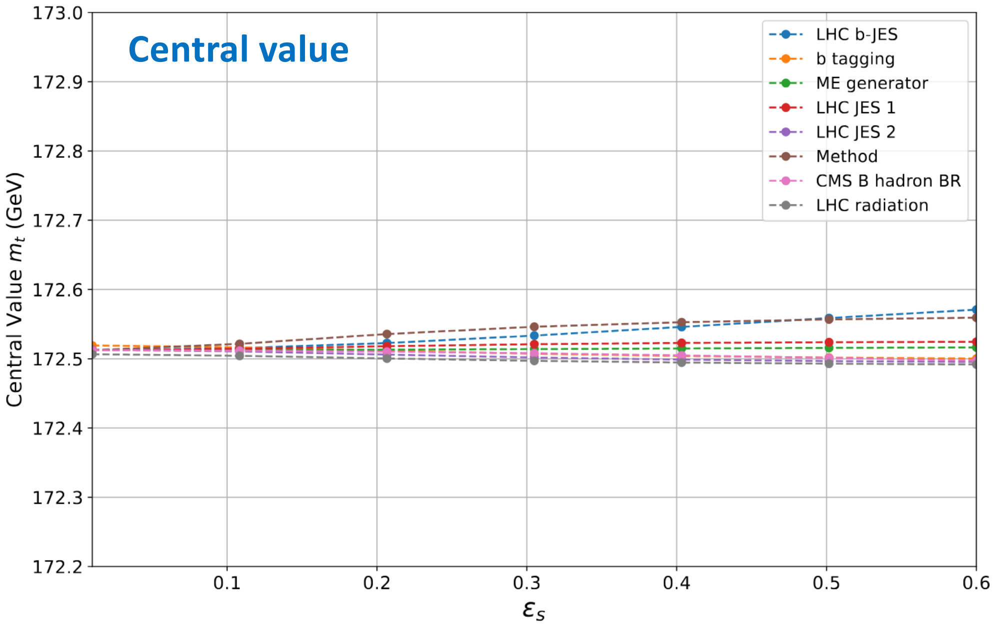
From: arXiv:2402.08713

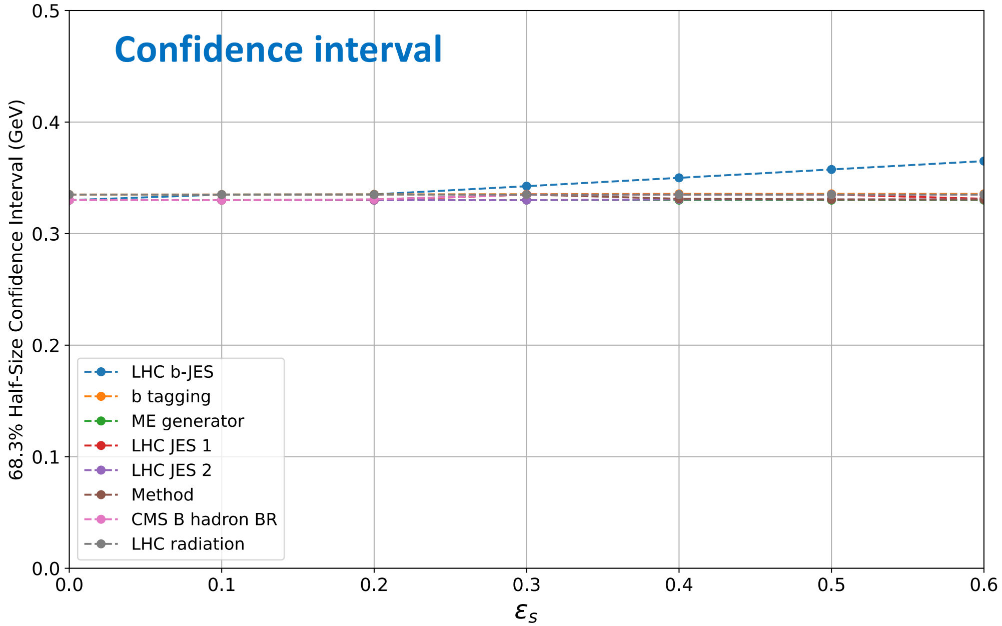
7-8 TeV ATLAS-CMS top-quark mass combination



- As $\boldsymbol{\varepsilon_s \rightarrow 0}$ we recover the results of the combination. Treatment of correlations consistent.

7-8 TeV ATLAS-CMS top-quark mass combination

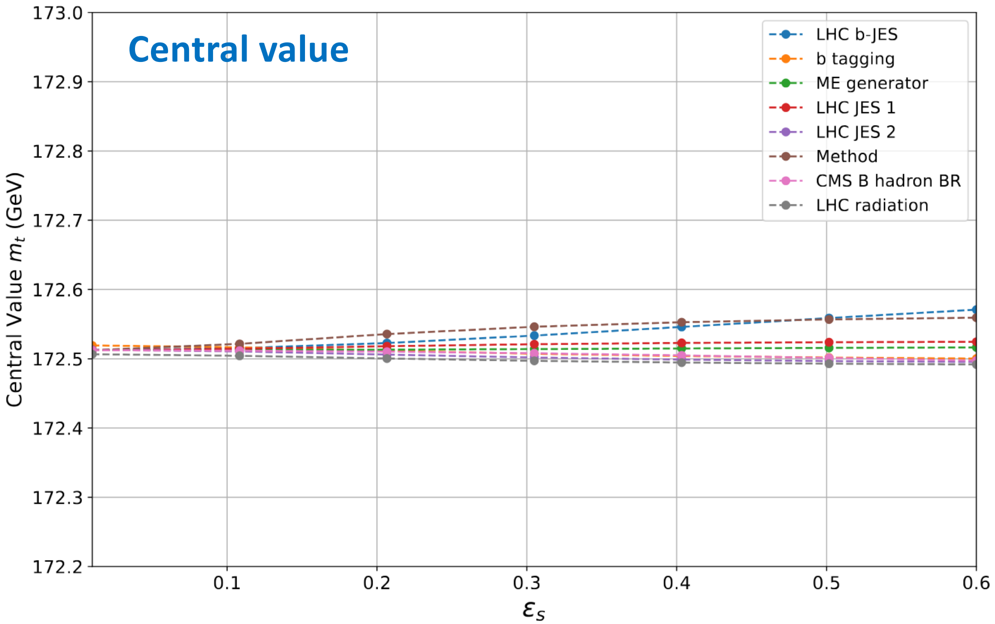


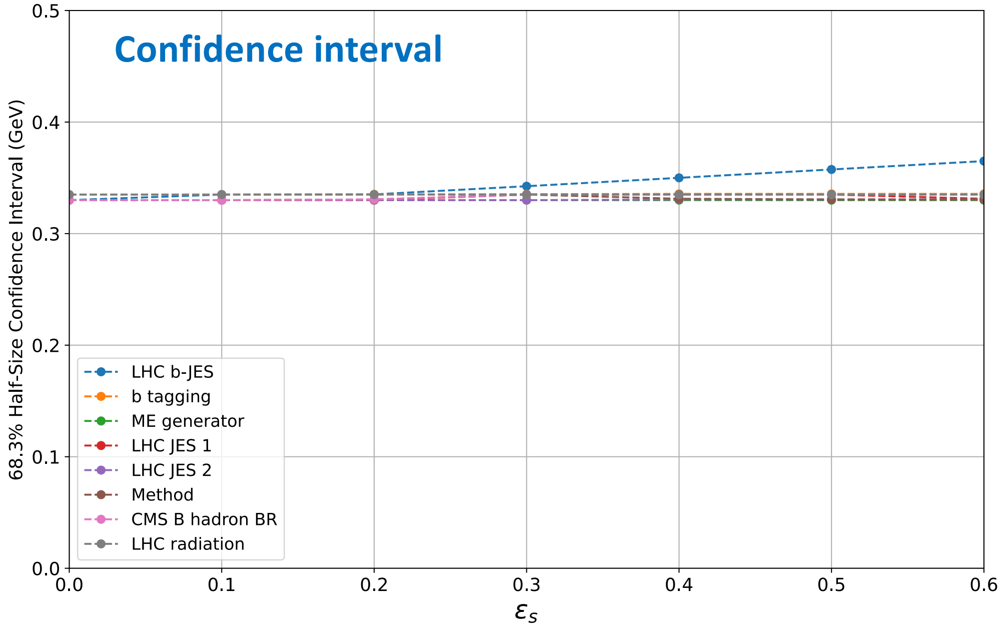


- As $\boldsymbol{\varepsilon_s \to 0}$ we recover the results of the combination. Treatment of correlations consistent.
- The central value is robust to the presence of uncertain systematic errors:

  ⟶ The change in the central value remains always within *0.1* GeV, well within the confidence interval of approximately *0.3* GeV.

7-8 TeV ATLAS-CMS top-quark mass combination





- As $\boldsymbol{\varepsilon_s \to 0}$ we recover the results of the combination. Treatment of correlations consistent.
- The central value is robust to the presence of uncertain systematic errors:
  → The change in the central value remains always within *0.1* GeV, well within the confidence interval of approximately *0.3* GeV.
- The confidence interval is also stable, though it exhibits a *10%* inflation when the *LHC b-JES* uncertainty has an error the error

Sensitivity to outliers

- **Goal**: Show how the combination is affected if any input measurement conflicts with others

Sensitivity to outliers

ROYAL HOLLOWAY UNIVERSITY OF LONDON

- **Goal**: Show how the combination is affected if any input measurement conflicts with others
- **Future relevance**: Relevant for future LHC–Tevatron combinations or LHC Run 2 combinations including the top mass measurement using a leptonic invariant mass (*J. High Energ. Phys.* **2023**, 19).

Sensitivity to outliers

ROYAL HOLLOWAY UNIVERSITY OF LONDON

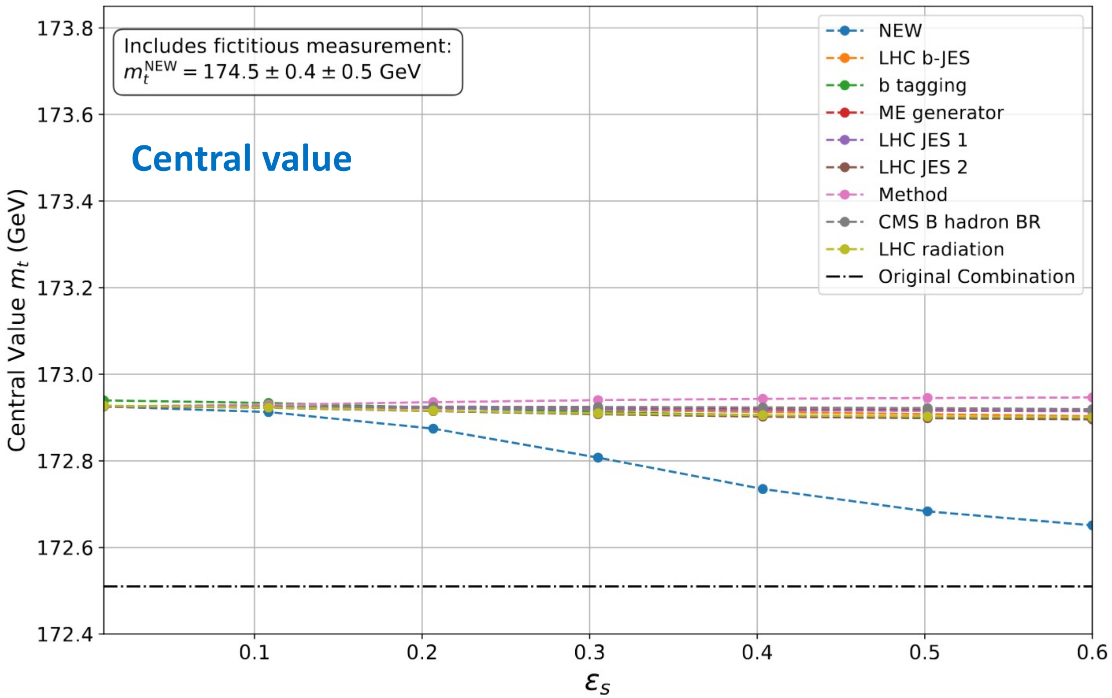
- **Goal**: Show how the combination is affected if any input measurement conflicts with others

- **Future relevance**: Relevant for future LHC–Tevatron combinations or LHC Run 2 combinations including the top mass measurement using a leptonic invariant mass (*J. High Energ. Phys.* **2023**, 19).

- **Introduce fictitious measurement**: Add a hypothetical measurement to explore the model properties in these scenarios.

Sensitivity to outliers

ROYAL HOLLOWAY UNIVERSITY OF LONDON



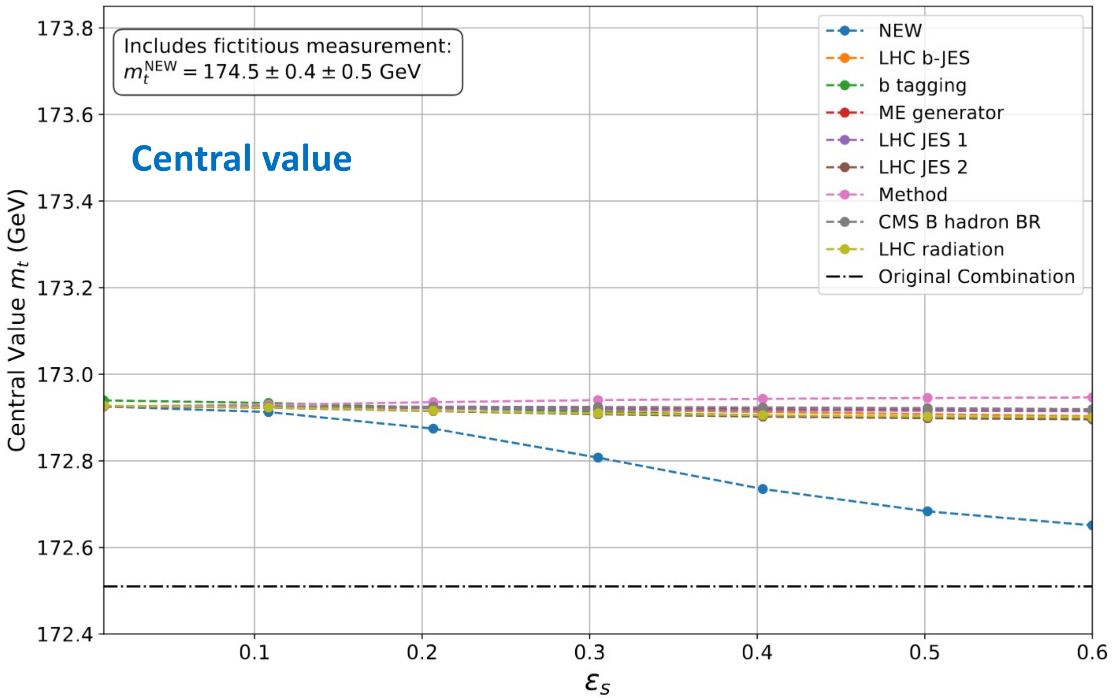
$$m_t^{NEW} = 174.5 \pm 0.4 \pm 0.5\ GeV$$

Independent systematic

- When ε_s is zero, the central value of the combination is pulled by the new measurement:

From $172.52\ GeV$ → **To** $172.91\ GeV$

Sensitivity to outliers



$$m_t^{NEW} = 174.5 \pm 0.4 \pm 0.5\ GeV$$

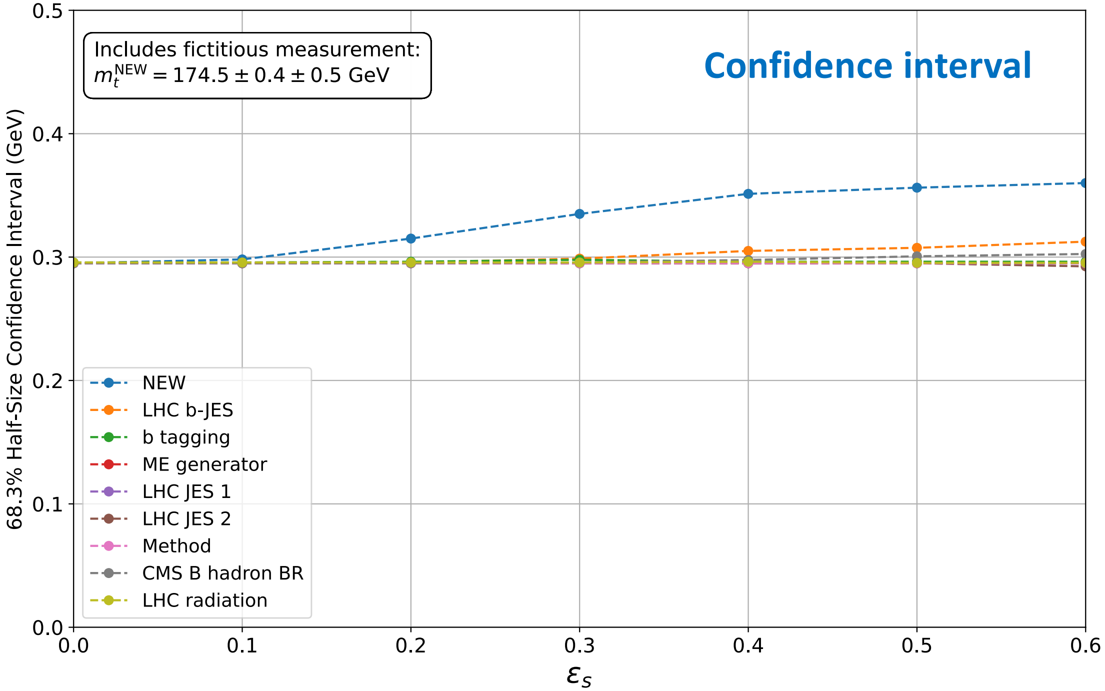
Independent systematic

- When $\boldsymbol{\varepsilon_s}$ is zero, the central value of the combination is pulled by the new measurement:

From $172.52\ GeV$ → **To** $172.91\ GeV$

- If the new measurement is affected by a large uncertain systematic, it shifts back to the original value

Sensitivity to outliers

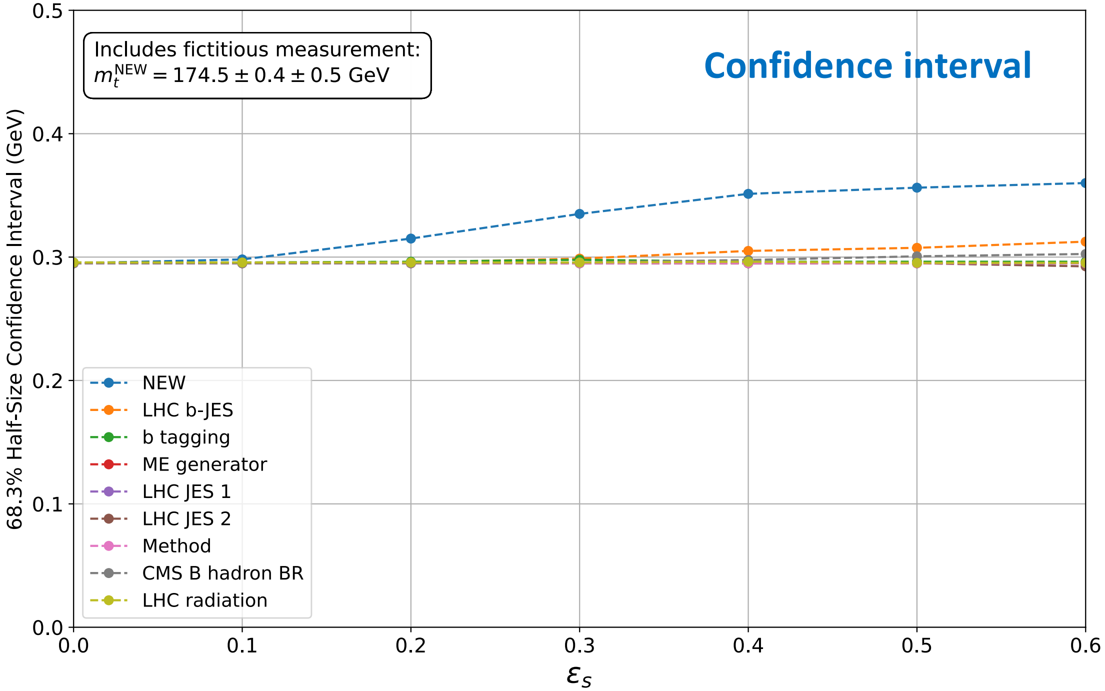


$$m_t^{NEW} = 174.5 \pm 0.4 \pm 0.5\ GeV$$

Independent systematic

- When ε_s is zero, adding the new measurement shrinks the CI: $0.33\ GeV \rightarrow 0.29\ GeV$

Sensitivity to outliers

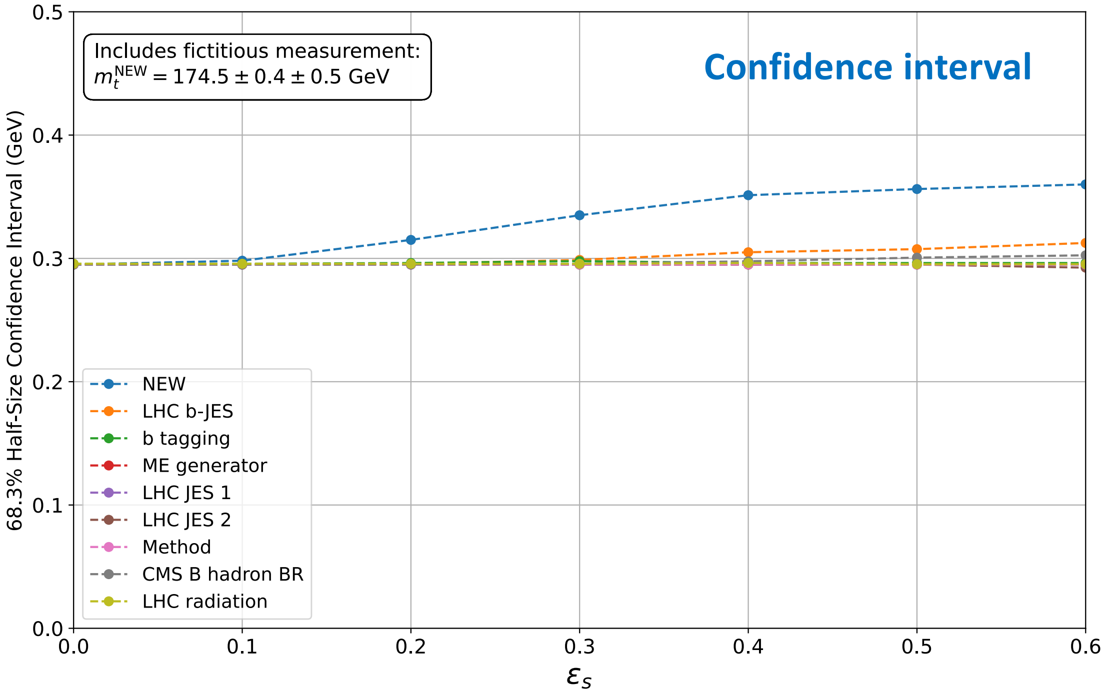


$$m_t^{NEW} = 174.5 \pm 0.4 \pm 0.5\ GeV$$

Independent systematic

- When ε_s is zero, adding the new measurement shrinks the CI: $0.33\ GeV \rightarrow 0.29\ GeV$
- If the new measurement is affected by a large uncertain systematic, the CI inflates

Sensitivity to outliers



$$m_t^{NEW} = 174.5 \pm 0.4 \pm 0.5\ GeV$$

Independent systematic

- When ε_s is zero, adding the new measurement shrinks the CI: $0.33\ GeV$ → $0.29\ GeV$
- If the new measurement is affected by a large uncertain systematic, the CI inflates
- The tension in the dataset in treated as an additional source of uncertainty

Summary conclusions

ROYAL HOLLOWAY UNIVERSITY OF LONDON

Gamma variance model

- A model to account from theory uncertainties, two points systematics, etc..
- The primary advantage of this approach is that it reduces the sensitivity of the fits to outliers.
- The presence of incompatible data is reflected by inflated error bars on the final results.

Bartlett correction

- Used to correct for deviations from Wilks' theorem
- Useful tool to correct for small sample sizes errors

"Non-trivial" correlations

- A method to use nuisance parameters for non-trivial systematics assumptions

Thank you for your attention

ROYAL HOLLOWAY UNIVERSITY OF LONDON

Back-up slides

Motivation for the GVM

- Gamma distributions allow to parametrize distributions of positive defined variables (like estimates of variances)

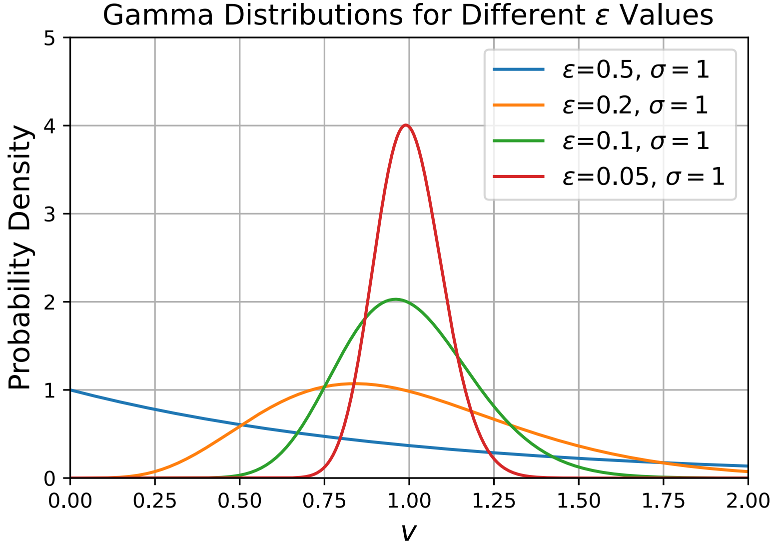
- Using Gamma distributions it is possible to profile in close form over σ_i^2

Gamma distribution

To implement "errors-on-errors" suppose the systematic variances $\sigma_{u_i}^2$ are *adjustable parameters*, and their best estimates v_i are gamma distributed:

$$v \sim \frac{\beta^{\alpha}}{\Gamma(\alpha)} v^{\alpha-1} e^{-\beta v}$$

$$\alpha = \frac{1}{4\varepsilon_i^2} \qquad \beta = \frac{1}{4\varepsilon_i^2 \sigma_{u_i}^2}$$



- $\sigma_{u_i}^2$ Expectation value of v_i
- ε_i: relative error on σ_{u_i}: "***Error on error***"*

*ε used to be r in previous references

Gamma Variance Model (GVM)

- The likelihood is modified as follows:

$$L(\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\sigma}_{\boldsymbol{u}_i}^2) = P(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\theta}) \times \prod_i \frac{1}{\sqrt{2\pi}\sigma_{u_i}} e^{-(u_i-\theta_i)^2/2\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2} \times \frac{\boldsymbol{\beta}_i^{\boldsymbol{\alpha}_i}}{\boldsymbol{\Gamma}(\boldsymbol{\alpha}_i)} \boldsymbol{v}_i^{\boldsymbol{\alpha}_i - 1} \boldsymbol{e}^{-\boldsymbol{\beta}_i \boldsymbol{v}_i}$$

- One can profile over $\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2$ in closed form:

$$\log L_P(\boldsymbol{\mu}, \boldsymbol{\theta}) \;=\; \log P(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\theta}) - \frac{1}{2} \sum_i \left(\mathbf{1} + \frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\varepsilon}_i^2}\right) \mathbf{log}\left(\mathbf{1} + \mathbf{2}\boldsymbol{\varepsilon}_i^2 \frac{(\boldsymbol{u}_i - \boldsymbol{\theta}_i)^2}{\boldsymbol{v}_i}\right)$$

Gamma Variance Model (GVM)

- The likelihood is modified as follows:

$$L(\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\sigma}_{\boldsymbol{u}_i}^2) = P(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\theta}) \times \prod_i \frac{1}{\sqrt{2\pi}\sigma_{u_i}} e^{-(u_i-\theta_i)^2/2\sigma_{u_i}^2} \times \frac{\boldsymbol{\beta}_i^{\boldsymbol{\alpha}_i}}{\boldsymbol{\Gamma}(\boldsymbol{\alpha}_i)} \boldsymbol{v}_i^{\boldsymbol{\alpha}_i - 1} \boldsymbol{e}^{-\boldsymbol{\beta}_i \boldsymbol{v}_i}$$

- One can profile over $\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2$ in closed form:

$$\log L_P(\boldsymbol{\mu}, \boldsymbol{\theta}) = \log P(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\theta}) - \frac{1}{2} \sum_i \left(\mathbf{1} + \frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\varepsilon}_i^2}\right) \mathbf{log} \left(\mathbf{1} + \mathbf{2}\boldsymbol{\varepsilon}_i^2 \frac{(\boldsymbol{u}_i - \boldsymbol{\theta}_i)^2}{\boldsymbol{v}_i}\right)$$

- Profiling means computing

$$L_P(\boldsymbol{\mu}, \boldsymbol{\theta}) = L\left(\boldsymbol{\mu}, \boldsymbol{\theta}, \widehat{\widehat{\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2}}\right), \qquad \widehat{\widehat{\boldsymbol{\sigma}_{\boldsymbol{u}_i}^2}} = argmax_{\sigma_{u_i}^2}\left(L(\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\sigma}_{\boldsymbol{u}_i}^2)\right)$$

Motivation for the GVM

ROYAL HOLLOWAY UNIVERSITY OF LONDON

- Gamma distributions include the case where the variance is estimate from a real dataset of control measurements:

$$v_i = \frac{1}{n_i - 1} \sum \left(u_{i,j} - \overline{u_i}\right)^2$$

- $(n-1)v_i/\sigma_{u_i}^2$ follows a χ_{n-1}^2 distribution and v_i a Gamma distribution with:

$$\alpha_i = \frac{n_i - 1}{2}$$

$$\beta_i = \frac{n_i - 1}{2\sigma_{u_i}^2}$$

Calculation of confidence intervals

- The likelihood function can be used to construct the profile likelihood ratio test statistic:

$$w_{\boldsymbol{\mu}} = -2ln\frac{L\left(\boldsymbol{\mu}, \hat{\hat{\boldsymbol{\theta}}}\right)}{L\left(\hat{\boldsymbol{\mu}},\ \hat{\boldsymbol{\theta}}\right)}$$

- Use the p-value:

$$p_{\boldsymbol{\mu}} = \int_{w_{\boldsymbol{\mu},obs}}^{\infty} f\left(w_{\boldsymbol{\mu}}|\boldsymbol{\mu}\right) dw_{\boldsymbol{\mu}}$$

- Include $\boldsymbol{\mu}$ such that:

$$p_{\boldsymbol{\mu}} < \alpha$$

Calculation of confidence intervals

- Modify the likelihood ratio w directly so that its distribution is closer to the asymptotic form:

$$w_{\mu} \longrightarrow w_{\mu}^{*} = w_{\mu}\frac{M}{E[w]}$$

To compute confidence intervals, rescale the results obtained with Standard methods, such as the Hessian method, by $\frac{M}{E[w]}$

$$w \sim \chi_M^2 + \mathcal{O}(\boldsymbol{n}^{-1})$$

$$w^{*} \sim \chi_M^2 + \mathcal{O}(\boldsymbol{n}^{-2})$$