Challenges of fitting cross-section data

Peelle's pertinent puzzle

Jaafar Chakrani, Stephen Dolan, Margherita Buizza Avanzini

Lawrence Berkeley National Laboratory (LBNL)

<jchakrani@lbl.gov>

NuXTract Workshop 2023

October 5th, 2023

Based on work performed in the context of arXiv:2308.01838

BERKELEY LAB

Nu Tract

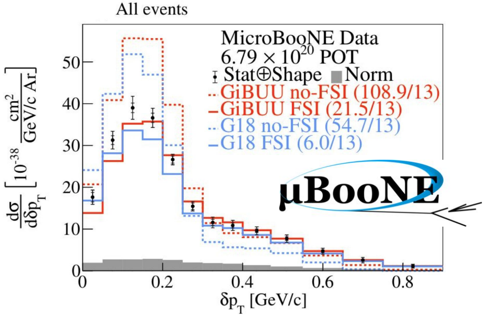
Introduction

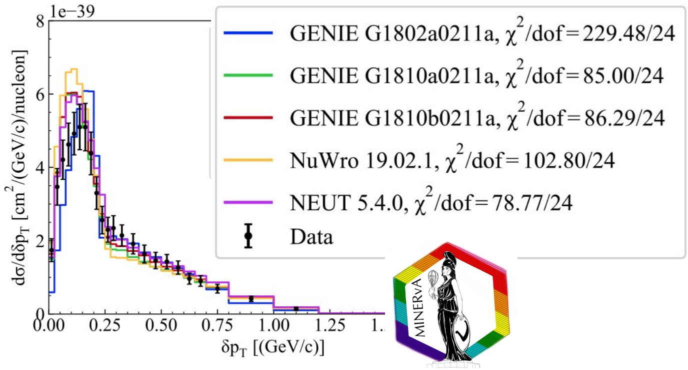
Introduction

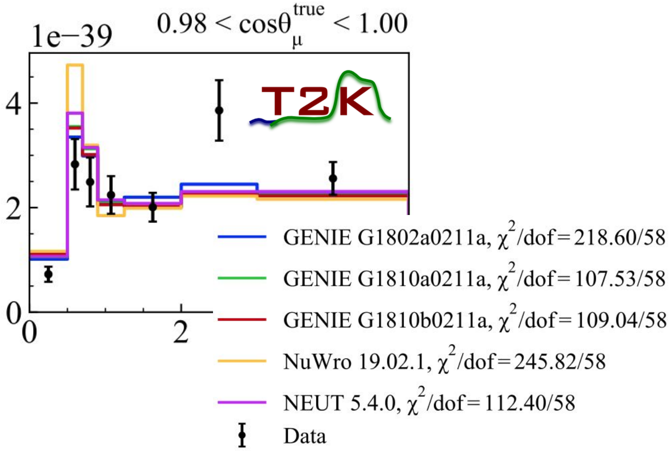
- Throughout the workshop, various techniques of cross-section extraction were presented
- Once the cross-section measurement is published, it can be used to **benchmark our models** from MC generators (NEUT, GENIE, NuWro, …)
- We find that, almost all the times, the models do not give a good description of the available cross-section data
- We often attempt to **tweak the models** using theoretically-motivated (or not!) parameters to better describe the data through **chi-square fits**
- A few recent examples:
  - GENIE: Phys. Rev. D 106, 112001, Phys. Rev. D 105, 072001, … (see talks by Julia and Michael)
  - NEUT: arXiv:2308.01838, …
  - NuWro: Phys. Rev. C 102, 015502 (2020), …

Challenge: models ≠ data

- No neutrino MC event generator is able to give a satisfactory description of neutrino-nucleus cross-section data (see TENSIONS 2019 report)*



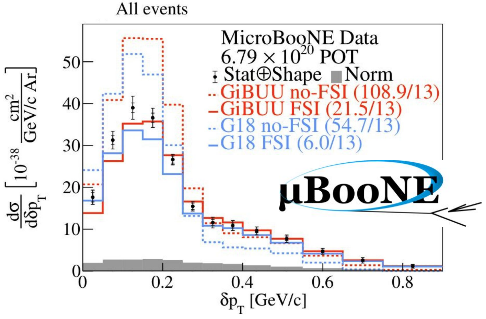


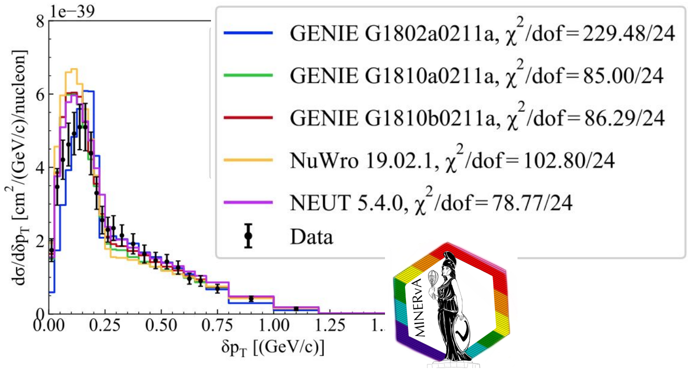


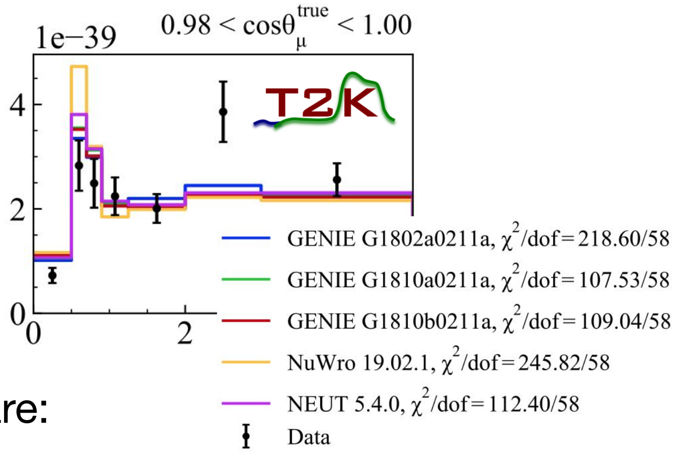
* This remains true even when we consider uncertainties within the models

Challenge: models ≠ data

- No neutrino MC event generator is able to give a satisfactory description of neutrino-nucleus cross-section data (see TENSIONS 2019 report)







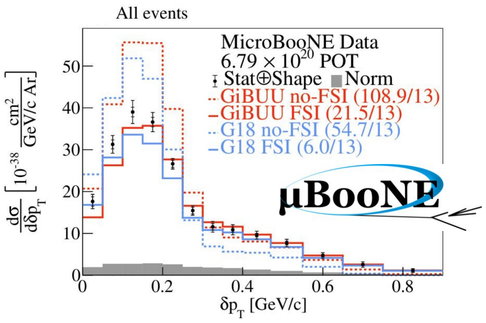
- We quantify the MC/data agreement using the chi-square:

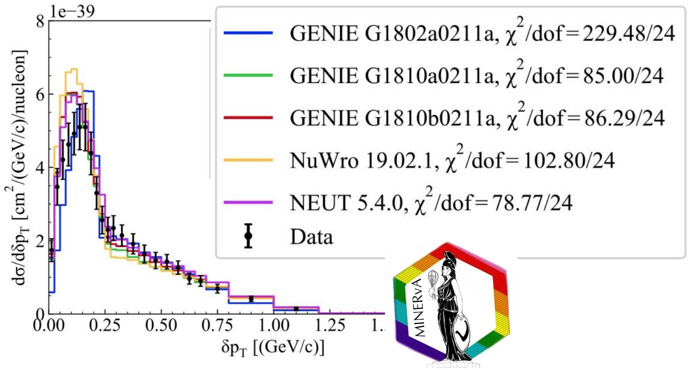
$$\chi^2 = \sum_{ij} \left(B_i - B_i^{\mathrm{MC}}\right) M_{ij}^{-1} \left(B_j - B_j^{\mathrm{MC}}\right)$$

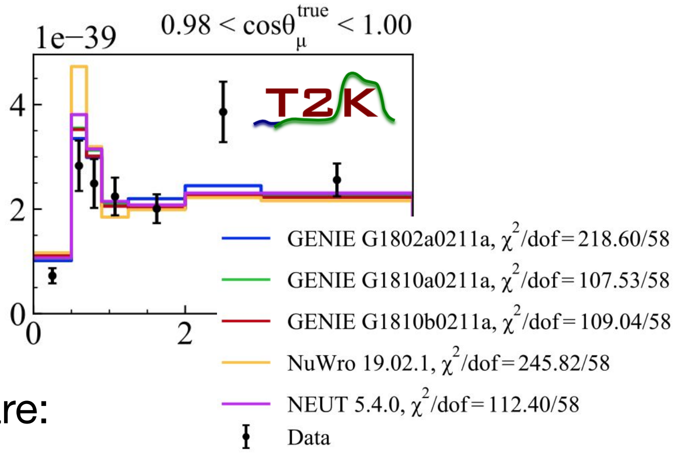
* This remains true even when we consider uncertainties within the models

Challenge: models ≠ data

- No neutrino MC event generator is able to give a satisfactory description of neutrino-nucleus cross-section data (see TENSIONS 2019 report)







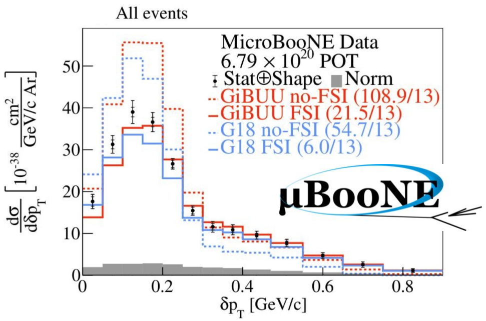
- We quantify the MC/data agreement using the chi-square:

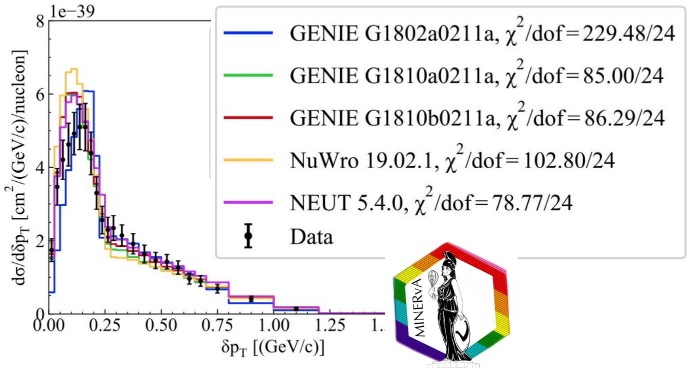
$$\chi^2 = \sum_{ij} \left(\boxed{B_i} - B_i^{\mathrm{MC}}\right) \boxed{M_{ij}^{-1}} \left(\boxed{B_j} - B_j^{\mathrm{MC}}\right)$$

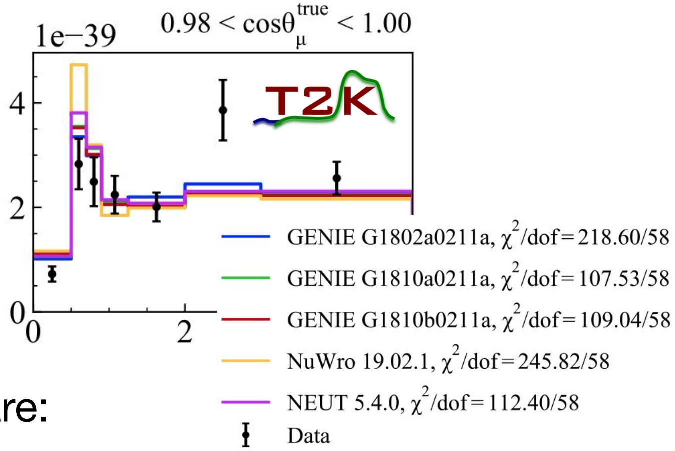
Published data: bin content + bin-to-bin covariance

Challenge: models ≠ data

- No neutrino MC event generator is able to give a satisfactory description of neutrino-nucleus cross-section data (see TENSIONS 2019 report)







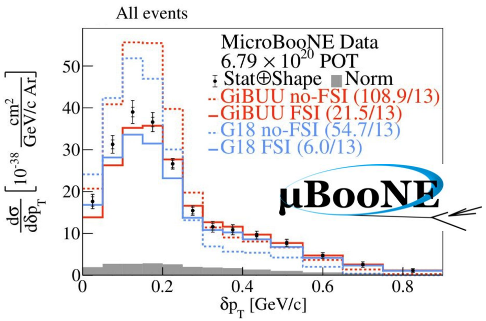
- We quantify the MC/data agreement using the chi-square:

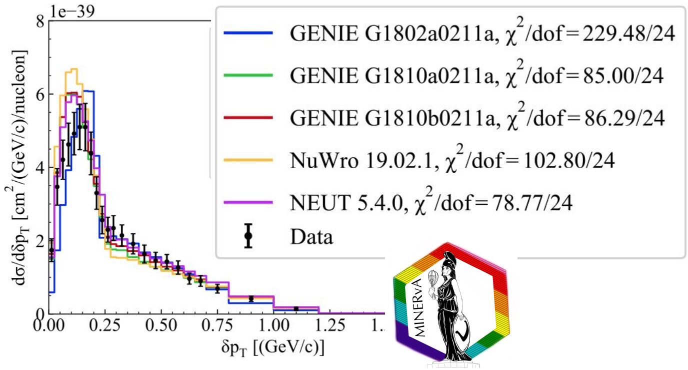
$$\chi^2 = \sum_{ij} \left(B_i - \boxed{B_i^{\mathrm{MC}}}\right) M_{ij}^{-1} \left(B_j - \boxed{B_j^{\mathrm{MC}}}\right)$$

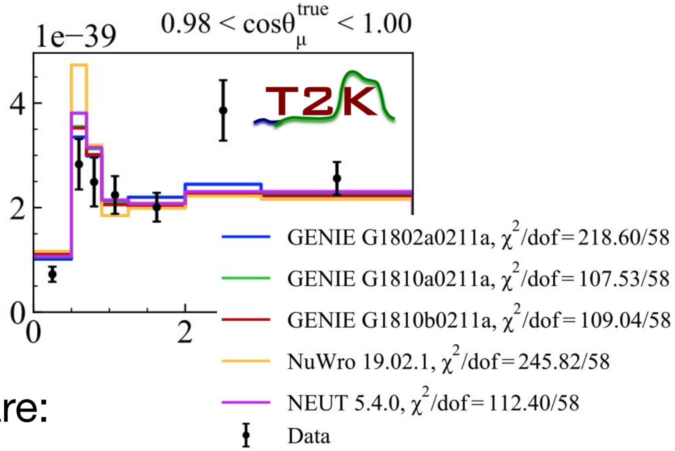
Bin content predicted by the model

Challenge: models ≠ data

- No neutrino MC event generator is able to give a satisfactory description of neutrino-nucleus cross-section data (see TENSIONS 2019 report)





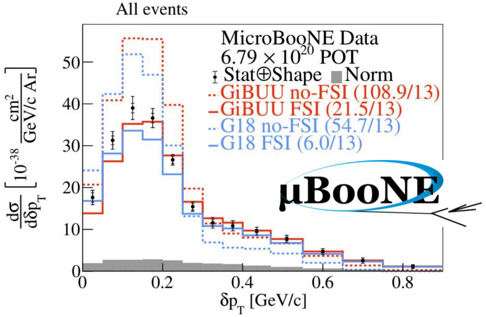


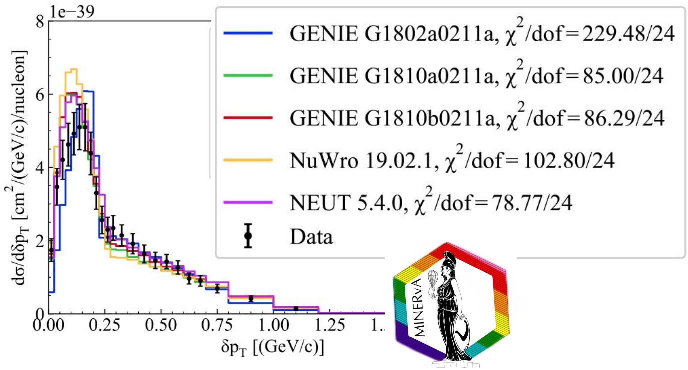
- We quantify the MC/data agreement using the chi-square:

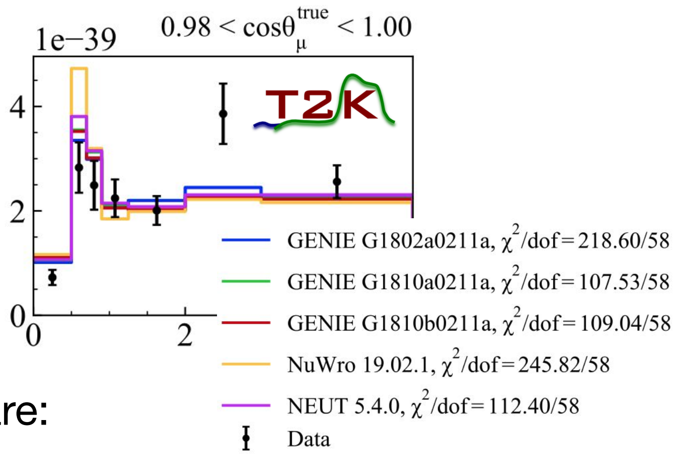
$$\chi^2 = \sum_{ij} \left(B_i - B_i^{\mathrm{MC}}\right) M_{ij}^{-1} \left(B_j - B_j^{\mathrm{MC}}\right)$$

Challenge: models ≠ data

- No neutrino MC event generator is able to give a satisfactory description of neutrino-nucleus cross-section data (see TENSIONS 2019 report)







- We quantify the MC/data agreement using the chi-square:

$$\chi^2(\vec{x}) = \sum_{ij} \left(B_i - B_i^{\mathrm{MC}}(\vec{x})\right) M_{ij}^{-1} \left(B_j - B_j^{\mathrm{MC}}(\vec{x})\right)$$

- To improve this agreement, we use parameters (knobs/dials) that tweak the model predictions and perform fits to the data

Typical fitting analysis in NUISANCE

JINST 12 P01016 (2017)

Allows us to:
- Add neutrino cross-section data
- Implement custom parameters
- Interact with MC reweight engines
- Compare/fit models to data
- …

(See talks by Laura and Luke)

Typical fitting analysis in NUISANCE

[JINST 12 P01016 (2017)](#)

- Ingredients to fit cross-section data:
  - Model

GENIE NEUT NuWro

$$\chi^2(\vec{x}) = \sum_{ij} \left(B_i - \boxed{B_i^{\mathrm{MC}}}(\vec{x})\right) M_{ij}^{-1} \left(B_j - \boxed{B_j^{\mathrm{MC}}}(\vec{x})\right)$$

Allows us to:
- Add neutrino cross-section data
- Implement custom parameters
- Interact with MC reweight engines
- Compare/fit models to data
- …

(See talks by Laura and Luke)

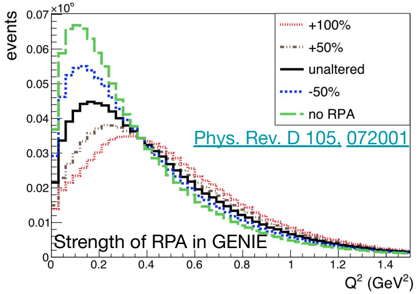
Typical fitting analysis in NUISANCE

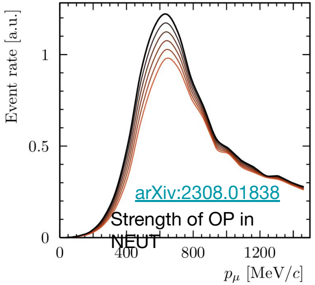
JINST 12 P01016 (2017)

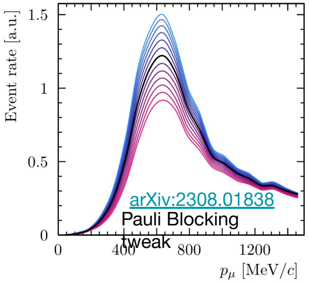
- Ingredients to fit cross-section data:
  - Model

GENIE NEUT NuWro

  - Parameters







$$\chi^2(\boxed{\vec{x}}) = \sum_{ij} \left(B_i - B_i^{\mathrm{MC}}(\boxed{\vec{x}})\right) M_{ij}^{-1} \left(B_j - B_j^{\mathrm{MC}}(\boxed{\vec{x}})\right)$$

Allows us to:
- Add neutrino cross-section data
- Implement custom parameters
- Interact with MC reweight engines
- Compare/fit models to data
- …

(See talks by Laura and Luke)

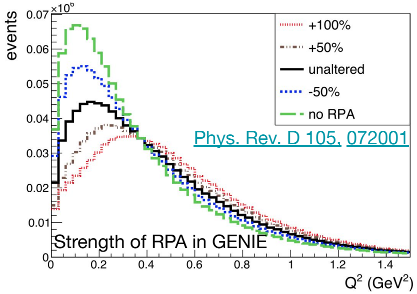
Typical fitting analysis in NUISANCE

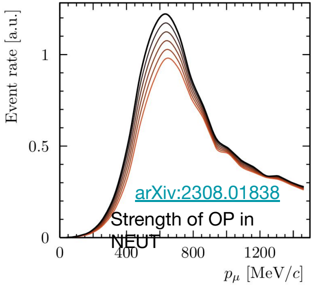
JINST 12 P01016 (2017)

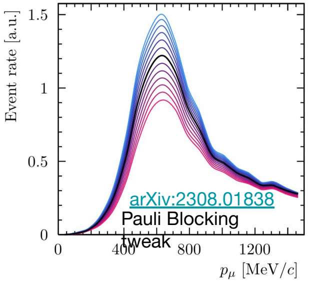
- Ingredients to fit cross-section data:
  - Model

    GENIE NEUT NuWro

  - Parameters







  - Data (bin content + bin-to-bin correlation)

$$\chi^2(\vec{x}) = \sum_{ij} \left(\boxed{B_i} - B_i^{\mathrm{MC}}(\vec{x})\right) \boxed{M_{ij}^{-1}} \left(\boxed{B_j} - B_j^{\mathrm{MC}}(\vec{x})\right)$$

Allows us to:
- Add neutrino cross-section data
- Implement custom parameters
- Interact with MC reweight engines
- Compare/fit models to data
- …

(See talks by Laura and Luke)

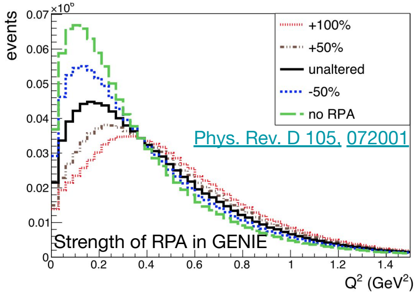
Typical fitting analysis in NUISANCE

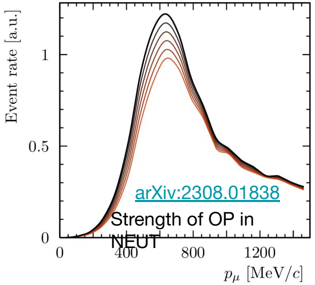
JINST 12 P01016 (2017)

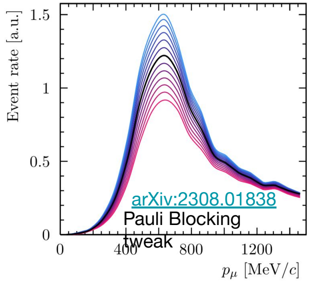
- Ingredients to fit cross-section data:
  - Model

GENIE NEUT NuWro

  - Parameters







  - Data (bin content + bin-to-bin correlation)

$$\chi^2(\vec{x}) = \sum_{ij} \left(B_i - B_i^{\mathrm{MC}}(\vec{x})\right) M_{ij}^{-1} \left(B_j - B_j^{\mathrm{MC}}(\vec{x})\right)$$

- This chi-square is then minimized as a function of $\vec{x}$

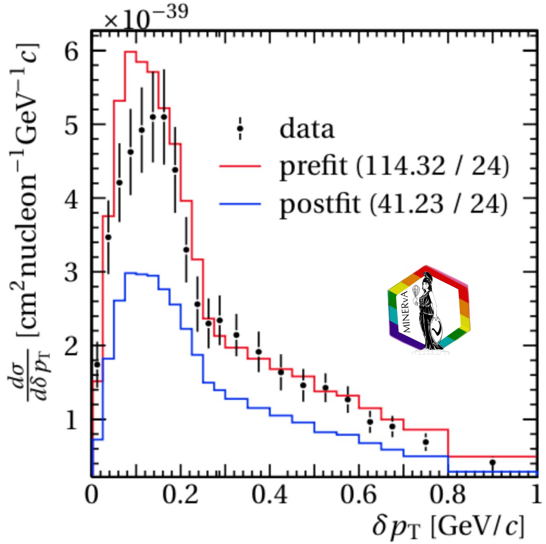
Allows us to:
- Add neutrino cross-section data
- Implement custom parameters
- Interact with MC reweight engines
- Compare/fit models to data
- ...

(See talks by Laura and Luke)

Peelle’s pertinent puzzle (PPP)

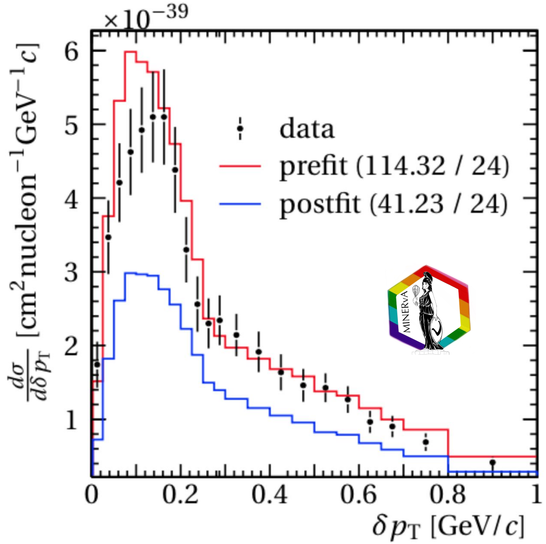
Peelle's Pertinent Puzzle

- Peelle's pertinent puzzle in one picture



Peelle's Pertinent Puzzle

- Peelle's pertinent puzzle in one picture

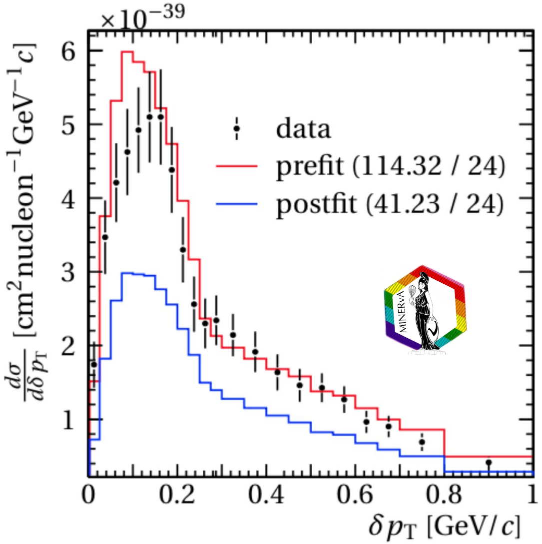


I chose my model (NEUT), implemented a few parameters, and ran a fit in NUISANCE to MINERvA data…

The postfit model I obtain is with a very small normalization!?

Peelle's Pertinent Puzzle

- Peelle's pertinent puzzle in one picture



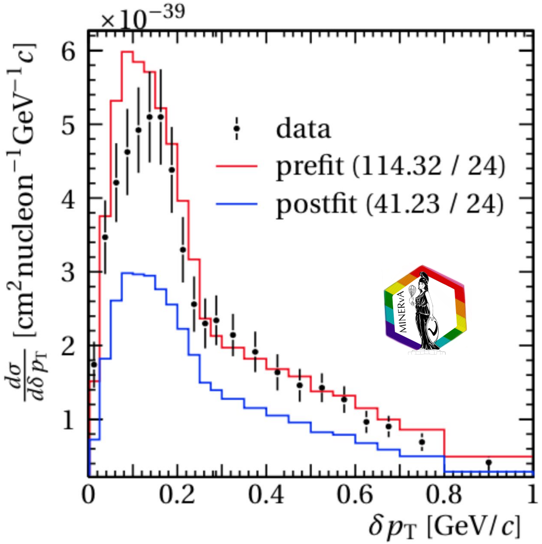
I chose my model (NEUT), implemented a few parameters, and ran a fit in NUISANCE to MINERvA data…

The postfit model I obtain is with a very small normalization!?

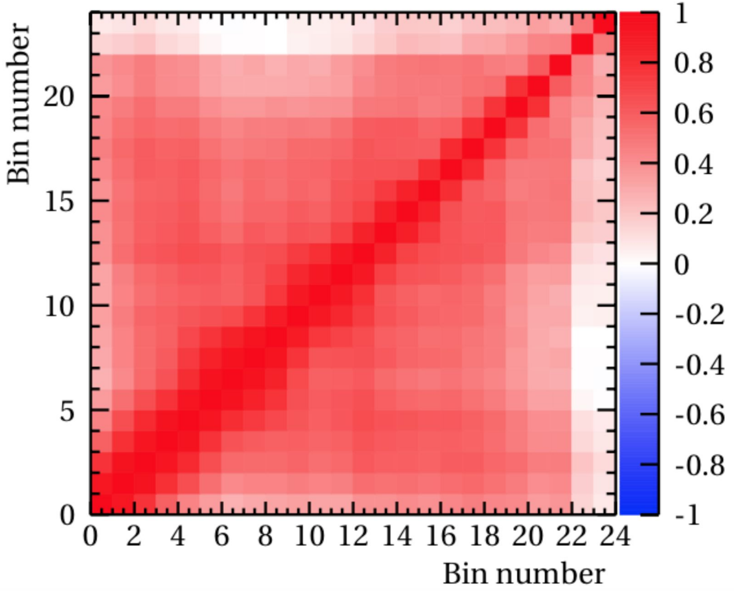
- This is the result of two things:
  - "Flawed" model that is **unable to perfectly describe the data** (even with the introduced free parameters)

Peelle's Pertinent Puzzle

- Peelle's pertinent puzzle in one picture



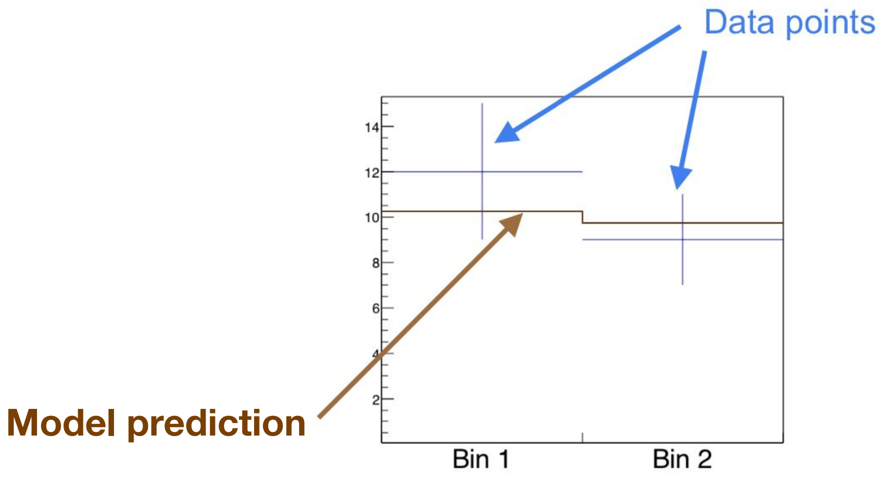
Bin-to-bin correlation matrix



- This is the result of two things:
  - "Flawed" model that is **unable to perfectly describe the data** (even with the introduced free parameters)
  - Highly correlated uncertainties between the bins summarized under **Gaussian** assumptions

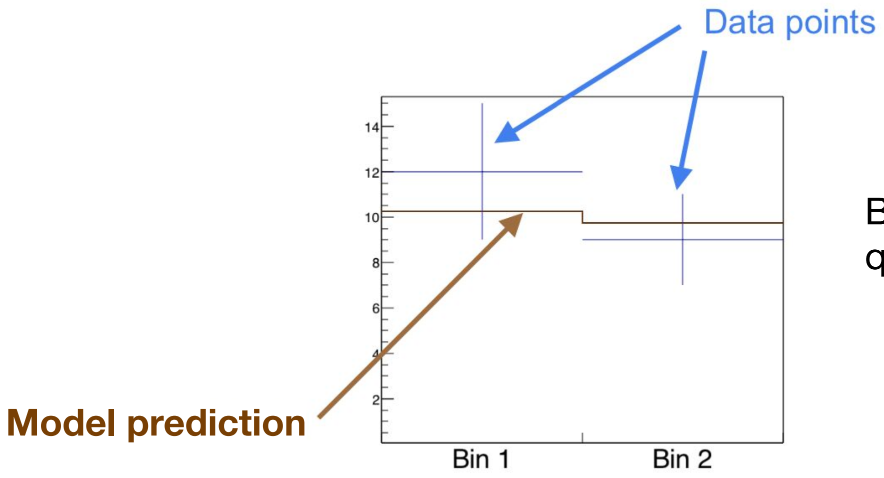
Why does this happen?

- Let's have a look at an example of a 2-bin measurement



Why does this happen?

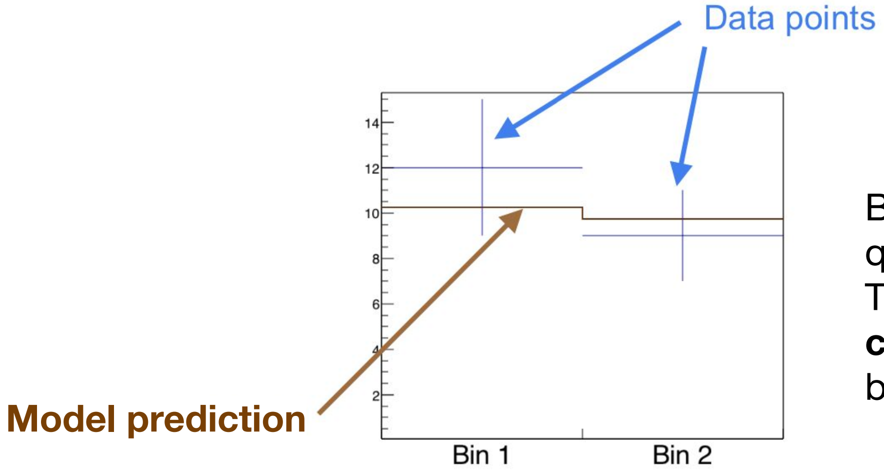
- Let's have a look at an example of a 2-bin measurement



By eye, the agreement looks quite good, **but…**

Why does this happen?

- Let’s have a look at an example of a 2-bin measurement

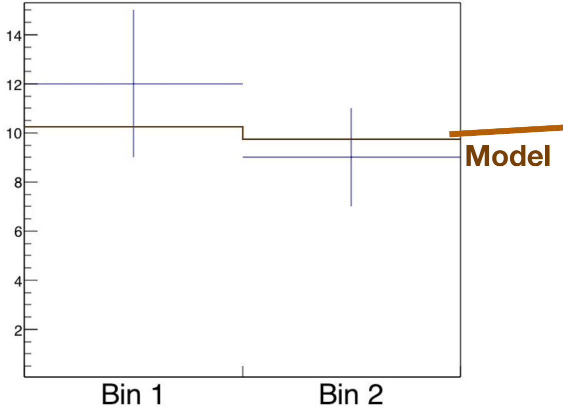


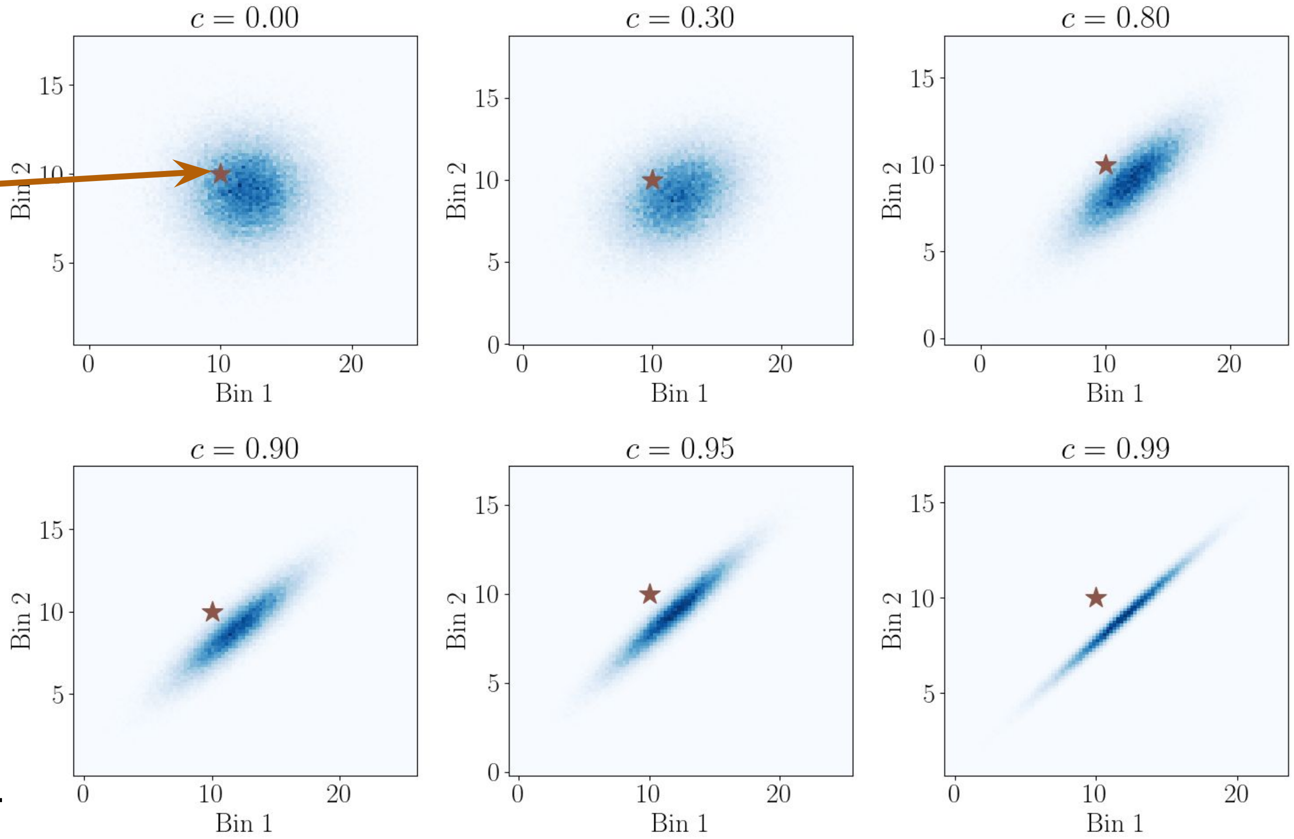
By eye, the agreement looks quite good, **but…**
This actually depends on the **correlation** between the two bins!

correlation = 0.00	correlation = 0.30	correlation = 0.80	correlation = 0.90	correlation = 0.95
chi2 = 0.69	chi2 = 0.98	chi2 = 3.41	chi2 = 6.81	chi2 = 13.62

Why does this happen?

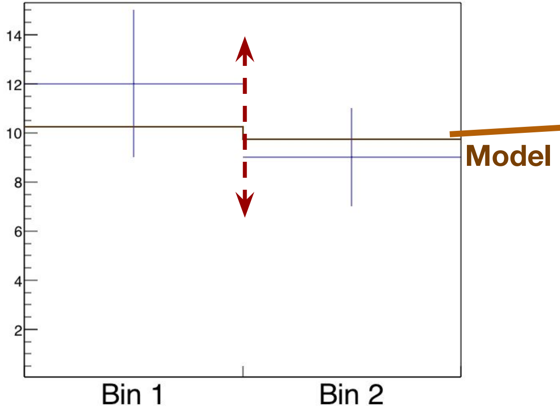
- In the two-dimensional space of the two bins

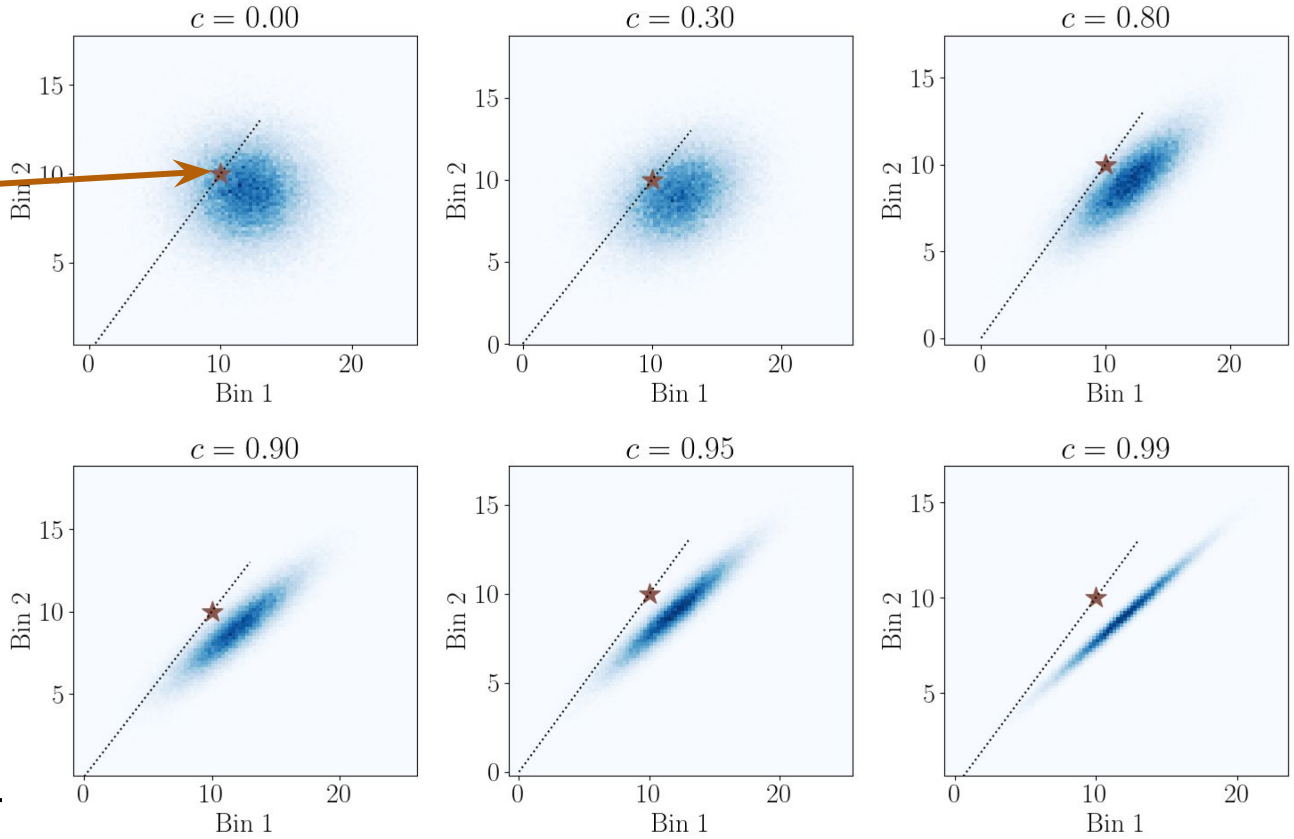




Why does this happen?

- In the two-dimensional space of the two bins

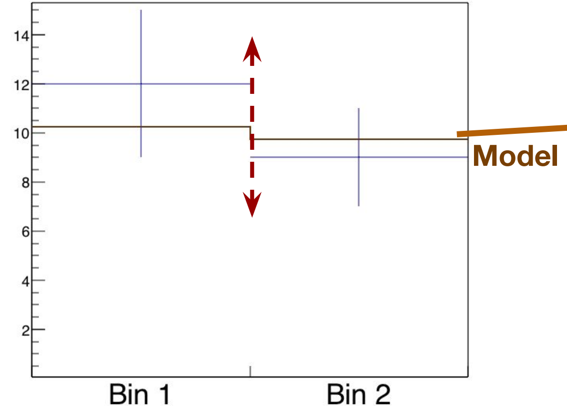


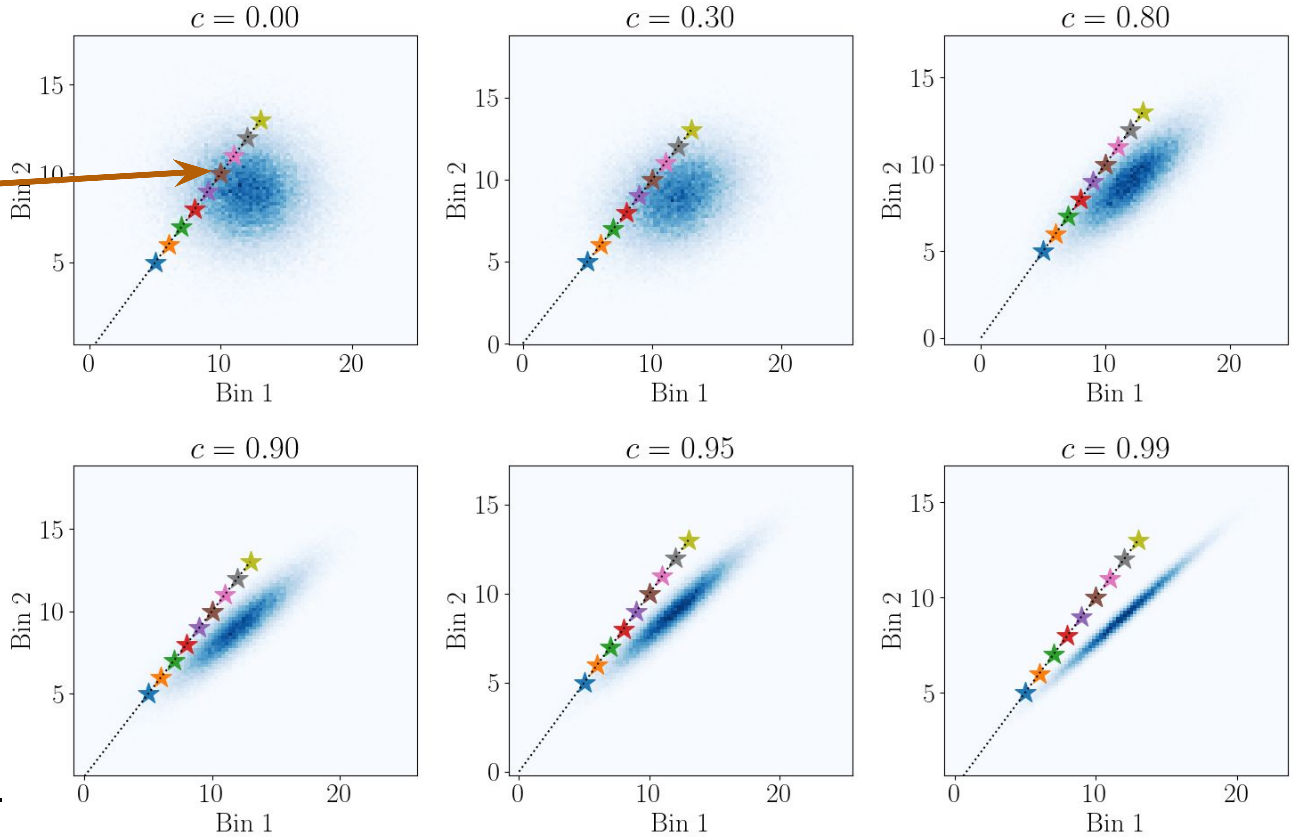


Let's introduce now a global normalization parameter *N*

Why does this happen?

- In the two-dimensional space of the two bins

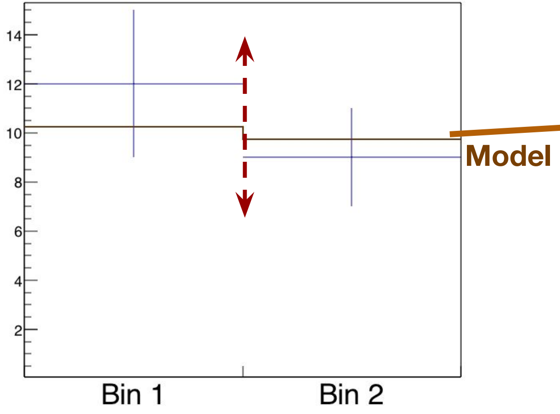


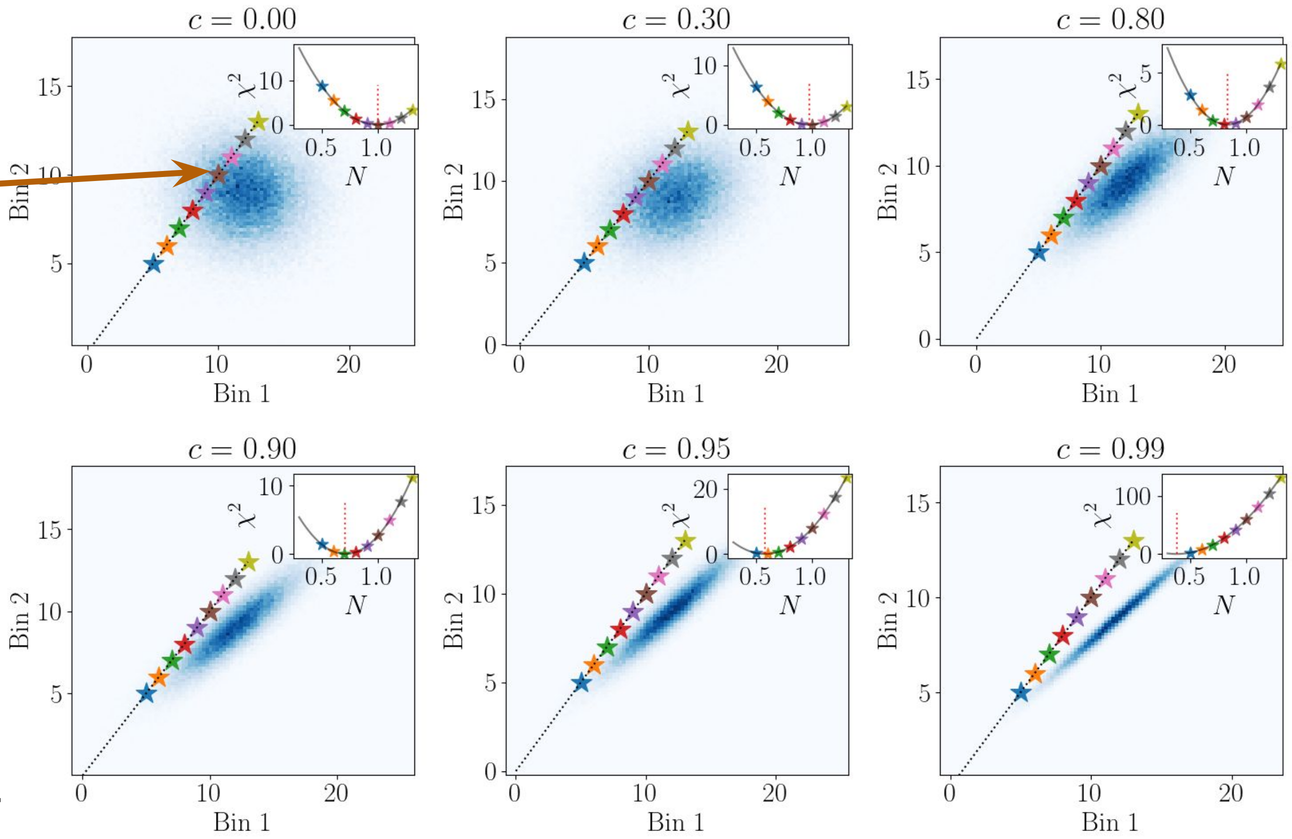


Let’s introduce now a global normalization parameter *N*

Why does this happen?

- In the two-dimensional space of the two bins





Let's introduce now a global normalization parameter *N*

As the correlation increases, the **low normalizations** are preferred

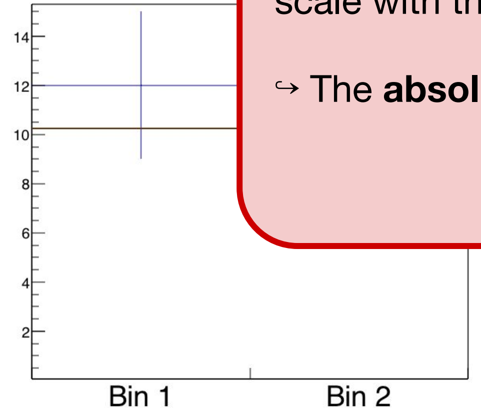
Why does this happen?

- In the t

With the standard covariance approach, the uncertainty on the shape does not scale with the changes to the normalization

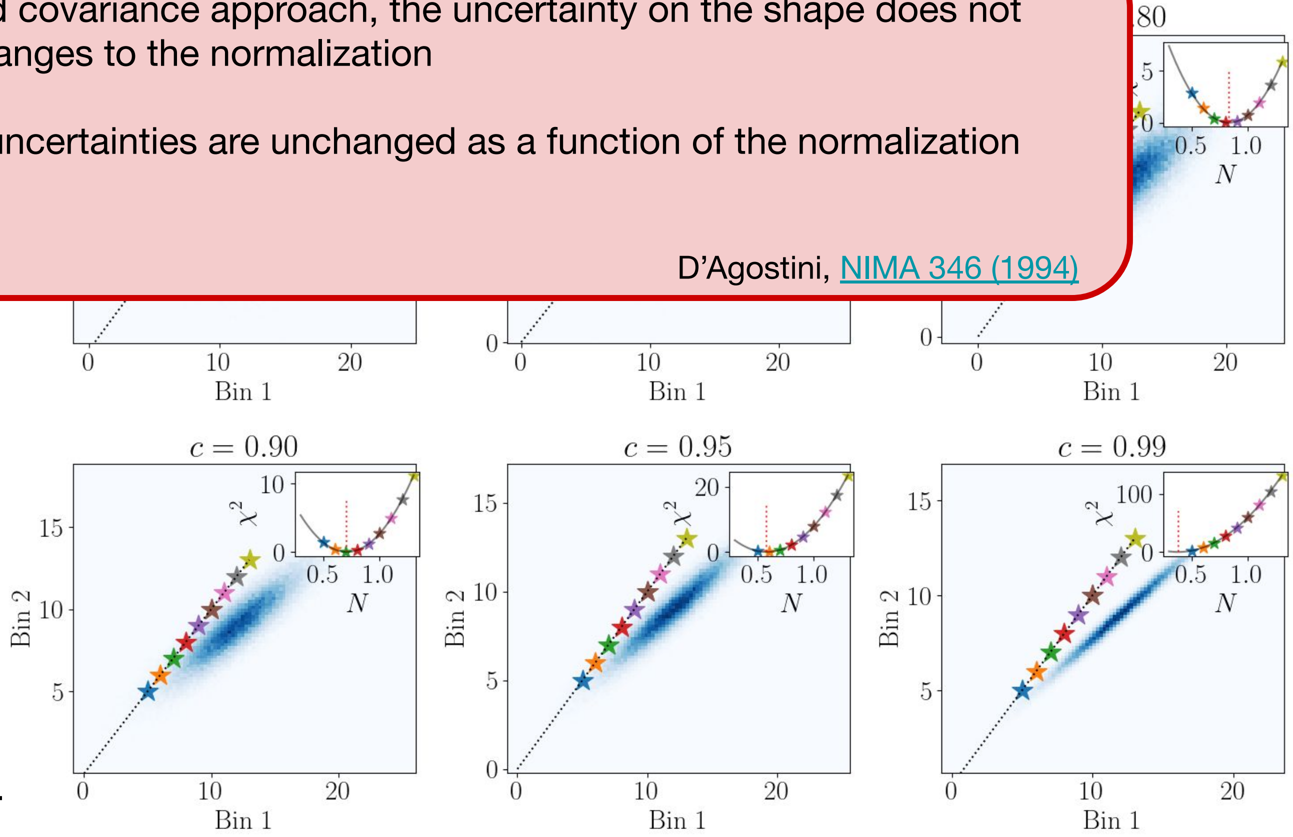
↪ The **absolute** uncertainties are unchanged as a function of the normalization

D'Agostini, [NIMA 346 (1994)]



Let's introduce now a global normalization parameter N

As the correlation increases, the **low normalizations** are preferred



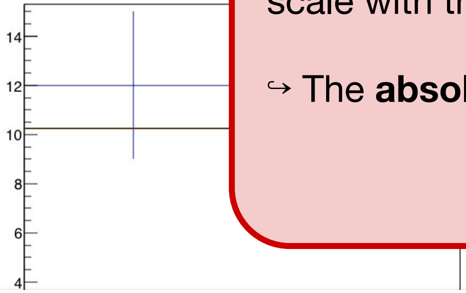
Why does this happen?

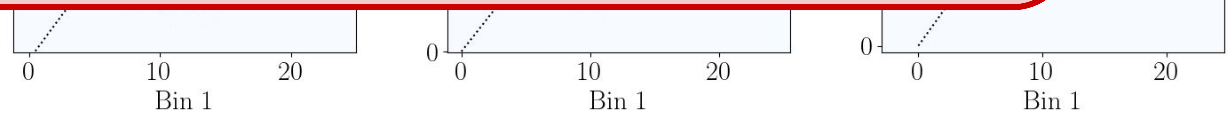
- In the t[…]

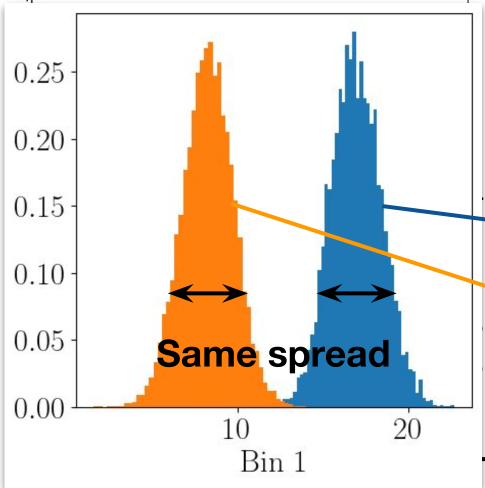
With the standard covariance approach, the uncertainty on the shape does not scale with the changes to the normalization

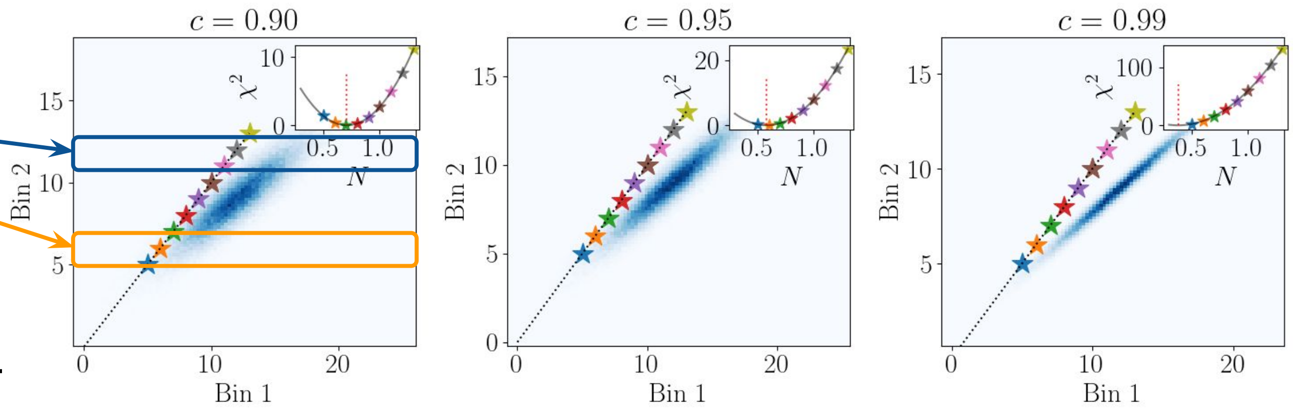
↪ The **absolute** uncertainties are unchanged as a function of the normalization

D'Agostini, [NIMA 346 (1994)]()



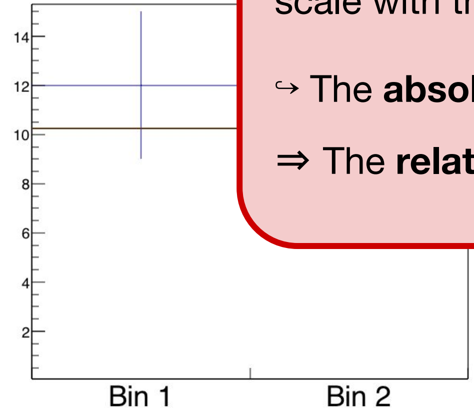






Why does this happen?

- In the t

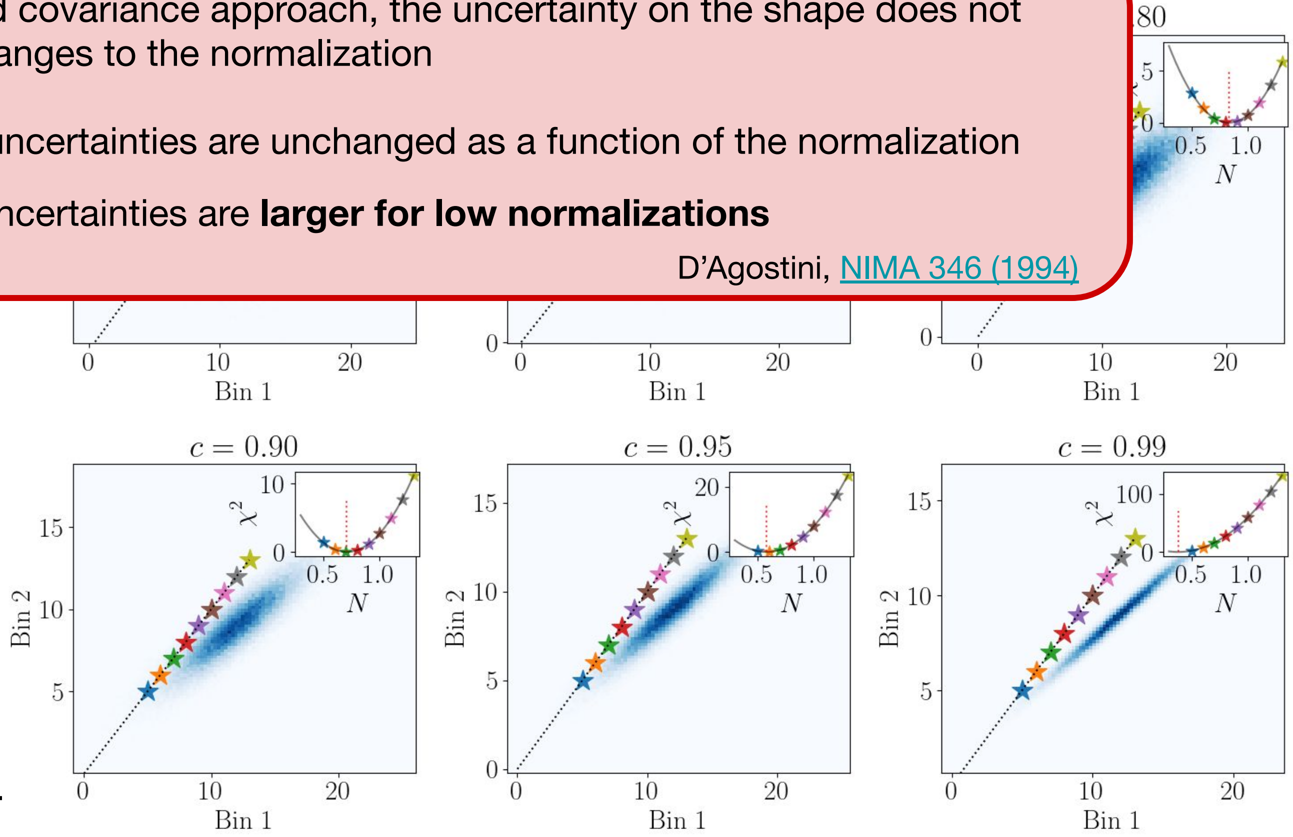


With the standard covariance approach, the uncertainty on the shape does not scale with the changes to the normalization

↪ The **absolute** uncertainties are unchanged as a function of the normalization

⇒ The **relative** uncertainties are **larger for low normalizations**

D'Agostini, NIMA 346 (1994)



Let's introduce now a global normalization parameter N

As the correlation increases, the **low normalizations** are preferred

Norm-shape covariance

- One way to overcome this is building a covariance that keeps the **relative** uncertainties constant (instead of the absolute)
- This can be satisfied if the covariance is quoted in the “norm” and the “shape” as in:

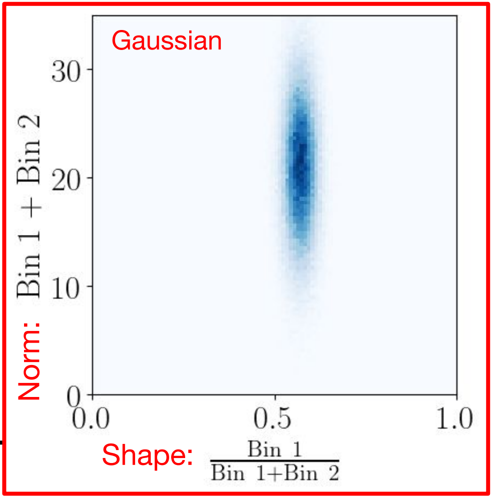
$$(\text{Bin } 1, \text{Bin } 2) \rightarrow \left(\text{Bin } 1 + \text{Bin } 2, \frac{\text{Bin } 1}{\text{Bin } 1 + \text{Bin } 2} \right)$$

Norm-shape covariance

- One way to overcome this is building a covariance that keeps the **relative** uncertainties constant (instead of the absolute)
- This can be satisfied if the covariance is quoted in the "norm" and the "shape" as in:

$$(\mathrm{Bin}\ 1, \mathrm{Bin}\ 2) \rightarrow \boxed{\left(\mathrm{Bin}\ 1 + \mathrm{Bin}\ 2, \frac{\mathrm{Bin}\ 1}{\mathrm{Bin}\ 1 + \mathrm{Bin}\ 2}\right)}$$

Gaussian uncertainties in the norm/shape



Norm-shape covariance

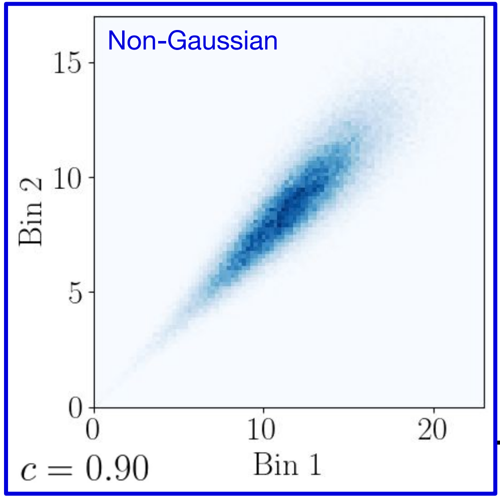
- One way to overcome this is building a covariance that keeps the **relative** uncertainties constant (instead of the absolute)
- This can be satisfied if the covariance is quoted in the "norm" and the "shape" as in:

$$\boxed{(\text{Bin }1, \text{Bin }2)} \rightarrow \boxed{\left(\text{Bin }1 + \text{Bin }2, \frac{\text{Bin }1}{\text{Bin }1 + \text{Bin }2}\right)}$$

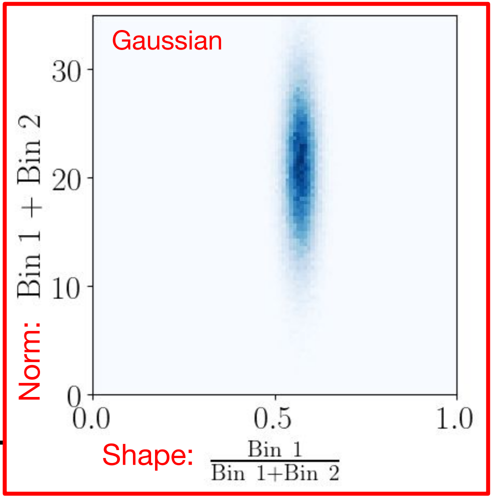
Non-Gaussian uncertainties

Gaussian uncertainties in the norm/shape

Non-linear transformation



Fixed relative uncertainty



Norm-shape covariance

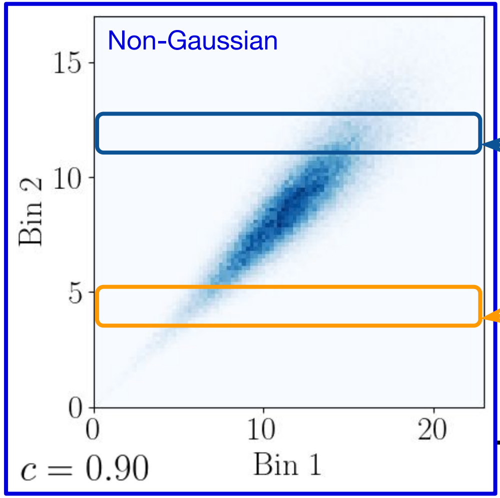
- One way to overcome this is building a covariance that keeps the **relative** uncertainties constant (instead of the absolute)
- This can be satisfied if the covariance is quoted in the “norm” and the “shape” as in:

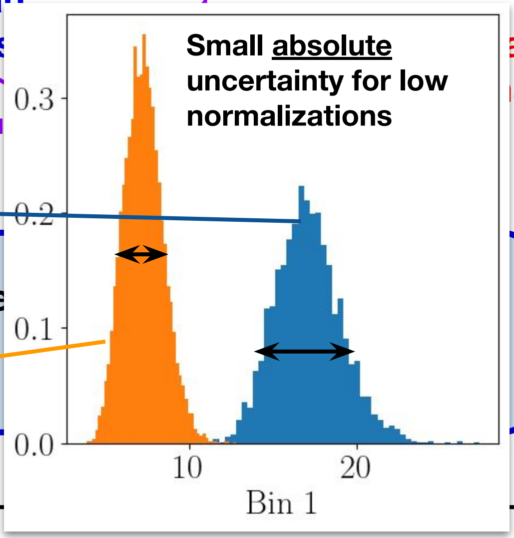
$$\boxed{(\mathrm{Bin}\ 1, \mathrm{Bin}\ 2)} \rightarrow \boxed{\left(\mathrm{Bin}\ 1 + \mathrm{Bin}\ 2, \frac{\mathrm{Bin}\ 1}{\mathrm{Bin}\ 1 + \mathrm{Bin}\ 2}\right)}$$

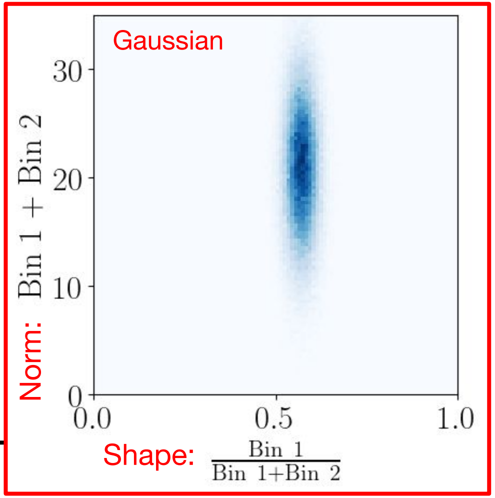
Non-Gaussian uncertainties

...ainties in the ...ape

Fixed re...







Norm-shape covariance

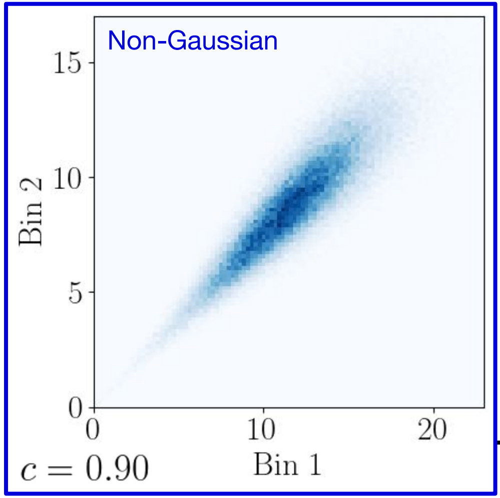
- One way to overcome this is building a covariance that keeps the **relative** uncertainties constant (instead of the absolute)
- This can be satisfied if the covariance is quoted in the “norm” and the “shape” as in:

$$\boxed{(\mathrm{Bin}\ 1, \mathrm{Bin}\ 2)} \rightarrow \boxed{\left(\mathrm{Bin}\ 1 + \mathrm{Bin}\ 2, \frac{\mathrm{Bin}\ 1}{\mathrm{Bin}\ 1 + \mathrm{Bin}\ 2}\right)}$$

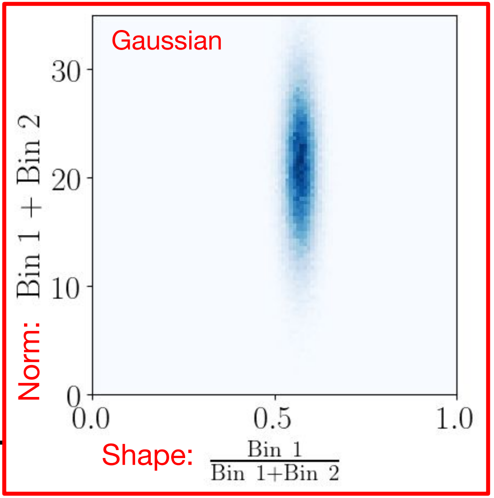
Non-Gaussian uncertainties

Gaussian uncertainties in the norm/shape

Non-linear transformation



Fixed relative uncertainty



Norm-shape covariance

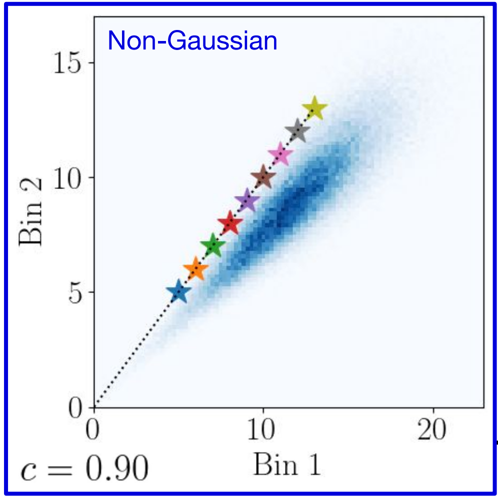
- One way to overcome this is building a covariance that keeps the **relative** uncertainties constant (instead of the absolute)
- This can be satisfied if the covariance is quoted in the "norm" and the "shape" as in:

$$\boxed{(\mathrm{Bin}\ 1, \mathrm{Bin}\ 2)} \to \boxed{\left(\mathrm{Bin}\ 1 + \mathrm{Bin}\ 2, \frac{\mathrm{Bin}\ 1}{\mathrm{Bin}\ 1 + \mathrm{Bin}\ 2}\right)}$$

Non-Gaussian uncertainties

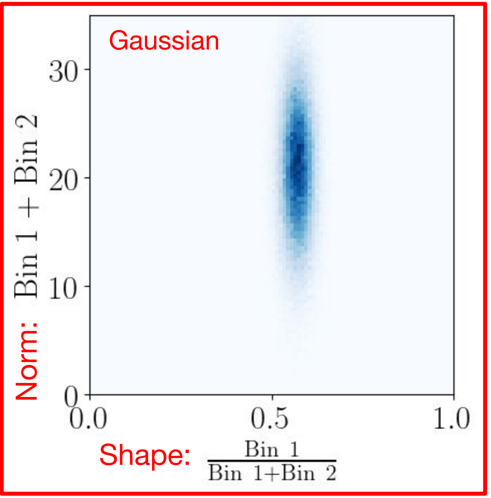
Gaussian uncertainties in the norm/shape

Non-linear transformation



Fixed relative uncertainty

⇒ No more preference for low norms



Norm-shape covariance

- One way to overcome this is building a covariance that keeps the **relative** uncertainties constant (instead of the absolute)
- This can be satisfied if the covariance is quoted in the “norm” and the “shape” as in:

$$(\text{Bin }1, \text{Bin }2) \rightarrow \left(\text{Bin }1 + \text{Bin }2, \frac{\text{Bin }1}{\text{Bin }1 + \text{Bin }2}\right)$$

arXiv:2308.01838

- This can be generalized for *n* bins, with the following transformation:

$$C_i = f(B_i) = \begin{cases} \alpha \frac{B_i}{\sum_k B_k}, & 1 \leq i \leq n-1 \\ B_T = \sum_k B_k, & i = n \end{cases}$$

$$N \approx J(f) \times M \times J(f)^T$$

Norm-shape covariance

- One way to overcome this is building a covariance that keeps the **relative** uncertainties constant (instead of the absolute)
- This can be satisfied if the covariance is quoted in the "norm" and the "shape" as in:

$$(\mathrm{Bin}\ 1, \mathrm{Bin}\ 2) \rightarrow \left(\mathrm{Bin}\ 1 + \mathrm{Bin}\ 2, \frac{\mathrm{Bin}\ 1}{\mathrm{Bin}\ 1 + \mathrm{Bin}\ 2}\right)$$

arXiv:2308.01838

- This can be generalized for *n* bins, with the following transformation:

$$C_i = f(B_i) = \begin{cases} \alpha \frac{B_i}{\sum_k B_k}, & 1 \leq i \leq n-1 \\ B_T = \sum_k B_k, & i = n \end{cases}$$

New basis

Data release histogram

New covariance

Jacobian of the transformation

$$N \approx J(f) \times M \times J(f)^T$$

Data release covariance

Linear approximation as f is non linear

Norm-shape covariance

- One way to overcome this is building a covariance that keeps the **relative** uncertainties constant (instead of the absolute)
- This can be satisfied if the covariance is quoted in the "norm" and the "shape" as in:

$$(\text{Bin } 1, \text{Bin } 2) \to \left(\text{Bin } 1 + \text{Bin } 2, \frac{\text{Bin } 1}{\text{Bin } 1 + \text{Bin } 2}\right)$$

arXiv:2308.01838

- This can be generalized for *n* bins, with the following transformation:

$$C_i = f(B_i) = \begin{cases} \alpha \frac{B_i}{\sum_k B_k}, & 1 \leq i \leq n-1 \\ B_T = \sum_k B_k, & i = n \end{cases}$$

New basis

Data release histogram

$$N \approx J(f) \times M \times J(f)^T$$

New covariance

Jacobian of the transformation

Data release covariance

Linear approximation as f is non linear

$$\chi^2_{\text{NS}} = \sum_{1 \leq i,j \leq n} \left(C_i - C_i^{\text{MC}}\right) \left(N^{-1}\right)_{i,j} \left(C_j - C_j^{\text{MC}}\right)$$

Norm-shape covariance

- One way to overcome this is building a covariance that keeps the **relative** uncertainties constant (instead of the absolute)
- This can be satisfied if the covariance is quoted in the “norm” and the “shape” as in:

$$(\text{Bin }1, \text{Bin }2) \rightarrow \left(\text{Bin }1 + \text{Bin }2, \frac{\text{Bin }1}{\text{Bin }1 + \text{Bin }2}\right)$$

arXiv:2308.01838

- This can be generalized for *n* bins, with the following transformation:

$$C_i = f(B_i) = \begin{cases} \alpha \frac{B_i}{\sum_k B_k}, & 1 \leq i \leq n-1 \\ B_T = \sum_k B_k, & i = n \end{cases}$$

New basis

Data release histogram

New covariance

Jacobian of the transformation

$$N \approx J(f) \times M \times J(f)^T$$

Data release covariance

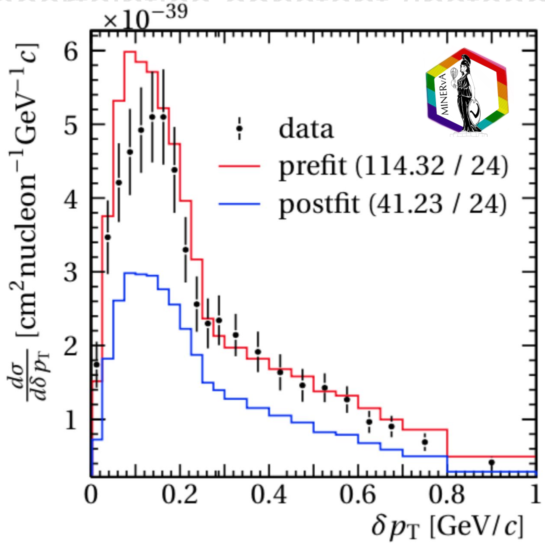
Linear approximation as f is non linear

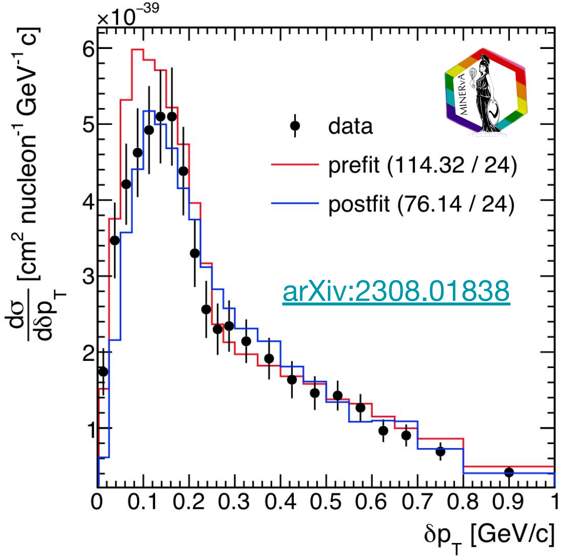
$$\chi^2_{\text{NS}} = \sum_{1 \leq i,j \leq n} \left(C_i - C_i^{\text{MC}}\right) \left(N^{-1}\right)_{i,j} \left(C_j - C_j^{\text{MC}}\right)$$

Can be easily used in NUISANCE!

Norm-shape covariance

- One way to overcome this is building a covariance that keeps the **relative** uncertainties constant (instead of the absolute)





New basis

Data release histogram

Data release covariance

Linear approximation as f is non linear

$$\chi^2_{\mathrm{NS}} = \sum_{1 \le i,j \le n} \left(C_i - C_i^{\mathrm{MC}}\right)\left(N^{-1}\right)_{i,j}\left(C_j - C_j^{\mathrm{MC}}\right)$$

Can be easily used in NUISANCE!

Is this satisfactory?

- Short answer… no

Is this satisfactory?

- Short answer… no
- The Norm+Shape decomposition is one workaround, but:
  - It is an approximate calculation $N \approx J(f) \times M \times J(f)^T$
  - The choice of assuming Gaussian uncertainties on the norm+shape decomposition is rather arbitrary, the actual distributions should be dictated by the measurement

Is this satisfactory?

- Short answer… no
- The Norm+Shape decomposition is one workaround, but:
  - It is an approximate calculation $N \approx J(f) \times M \times J(f)^T$
  - The choice of assuming Gaussian uncertainties on the norm+shape decomposition is rather arbitrary, the actual distributions should be dictated by the measurement
- Other workarounds have been considered in other studies like calculating a shape-only chi-square or even completely removing the bin-to-bin correlations…

Is this satisfactory?

- Short answer… no
- The Norm+Shape decomposition is one workaround, but:
  - It is an approximate calculation $N \approx J(f) \times M \times J(f)^T$
  - The choice of assuming Gaussian uncertainties on the norm+shape decomposition is rather arbitrary, the actual distributions should be dictated by the measurement
- Other workarounds have been considered in other studies like calculating a shape-only

- The bin-to-bin covariance provided by experiments is typically produced by varying all the uncertainties a large number of times, and summarizing the average and the spread
- **If these full variations (toys/universes) were to be provided in the data release, it would be possible to test these assumptions and tailor a dedicated test statistic**

Conclusion

Conclusion

- Fitting neutrino interaction models to cross-section measurements allows us to test and benchmark our neutrino event generators
- One of the challenges encountered in this procedure is Peelle's pertinent puzzle due to the imperfect models and the Gaussian assumptions of the errors in the data releases
- There are some workarounds that can be done to attempt to mitigate this issue using the published covariances, but…
- **An ideal data release would also contain the toys/universes used to propagate the uncertainties so that the full distributions can be accessed and used beyond the simple Gaussian assumptions**