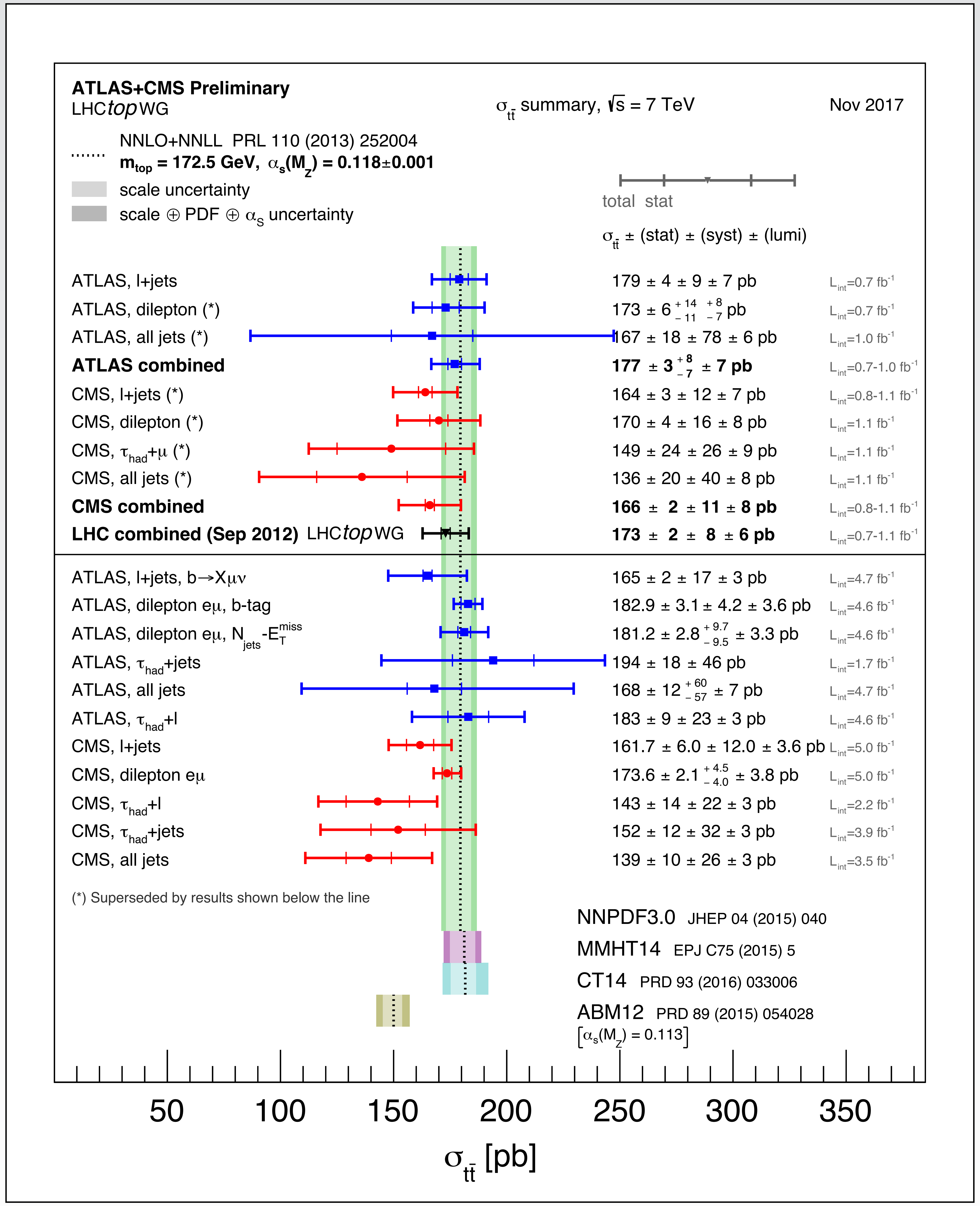
Data Combination
- In Practice -

Jan Kieseler

29.3.2022

Outline

- Full likelihood combinations (only some pointers)
- General common pitfalls and challenges when performing combinations
- Example of likelihood approximation
- Method validation based on toys
- Combined hands-on/hands-off examples using [Convino](https://github.com/jkiesele/Convino) (https://github.com/jkiesele/Convino)



Convino Setup

- Please check **now** if you can ‘ssh **-X** to’ lxplus.cern.ch
  - Make sure to use X forwarding if you can

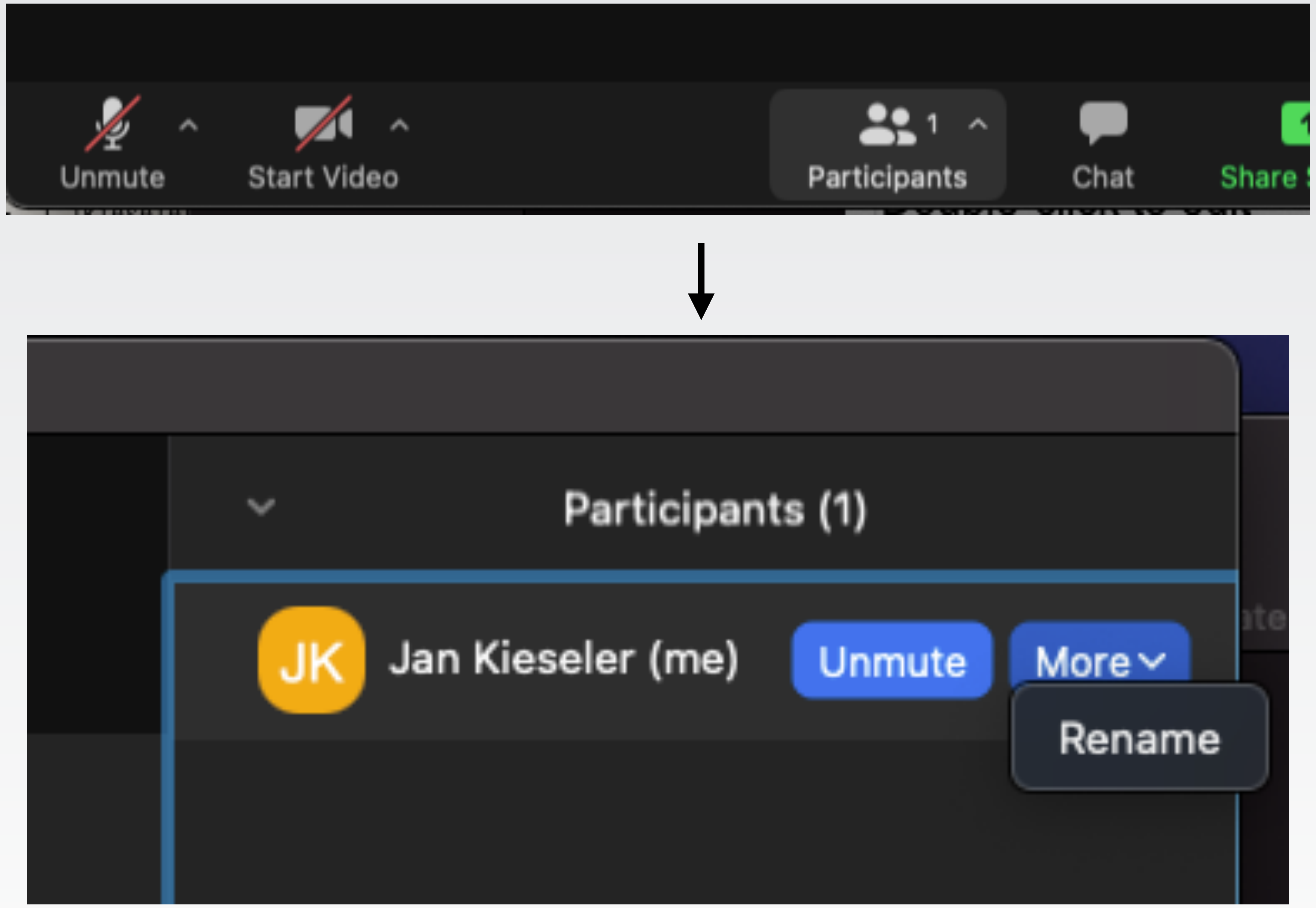
```
> bash
> cd /afs/cern.ch/user/j/jkiesele/public/Convino/latest
> source lxplus_env.sh
> cd
> mkdir convino_tutorial
> cd convino_tutorial
> convino /afs/cern.ch/user/j/jkiesele/public/Convino/latest/examples/exampleconfig.txt
```

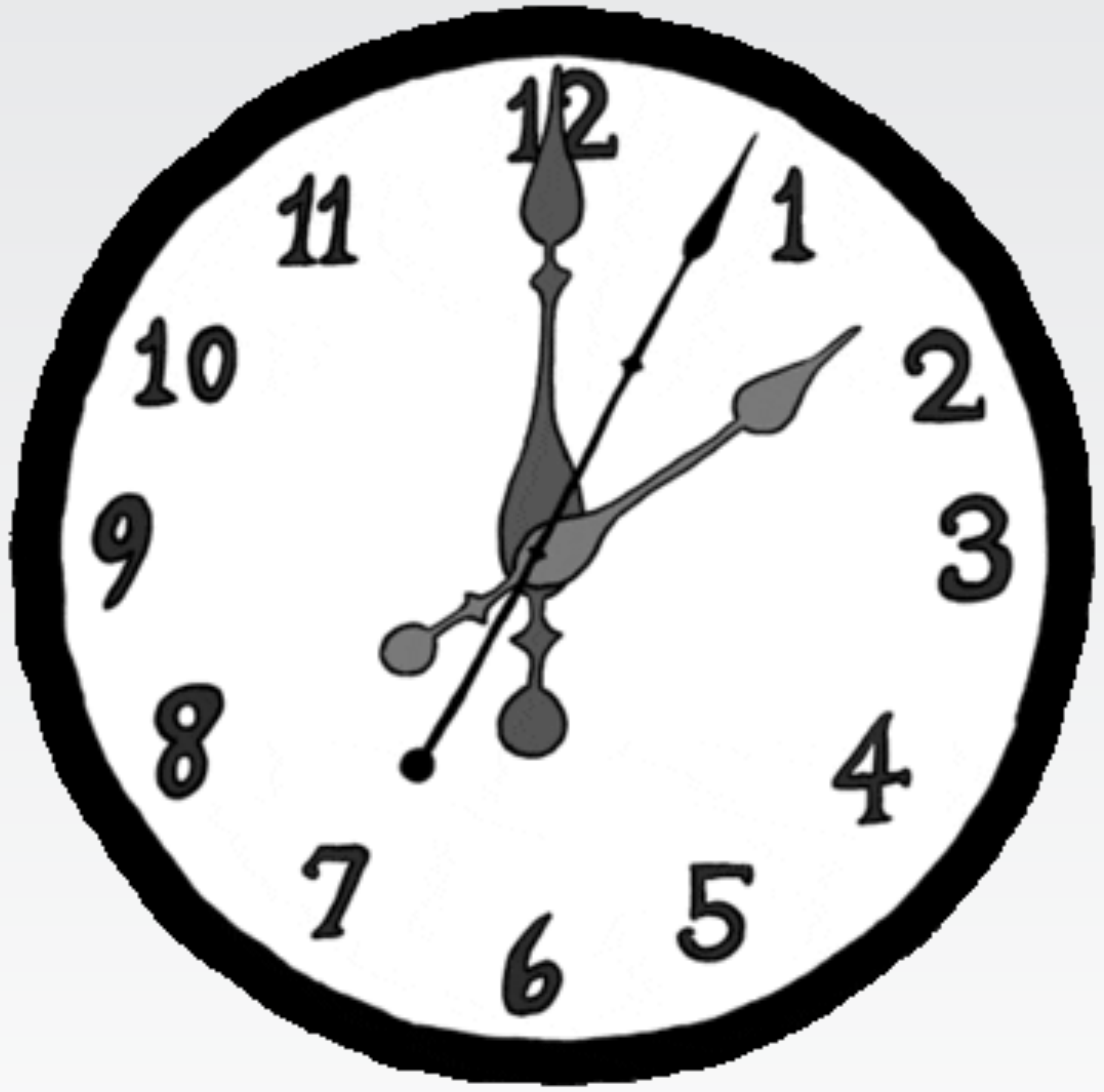
- If you **don’t have access** or anything does not work, **you can still follow this lecture fully**

- There is no need to spend the whole lecture trying to get it to work if it does not work out of the box

Rename

Please indicate if you can run the example by adding a “1_C ” in front of your name





Combining full likelihoods: Combine and HistFactory

- “Cleanest way” of performing the combination: works for limits and measurements
- Standard tools used for likelihood fits and full likelihood combinations in CMS and ATLAS
- Build on top of ROOT and RooFit
  - Likelihood is persisted in RooFit workspace
  - Inputs of different human readable sorts (text files+ROOT histograms, JSON)

Combine

- 3-day tutorial (CMS only): https://indico.cern.ch/event/976099/timetable/?view=standard
- Public documentation: https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/

HistFactory-based, e.g. TRExFitter:

- Original Histfactory document: https://cds.cern.ch/record/1456844?ln=en
- Tutorial/docs: https://trexfitter-docs.web.cern.ch/trexfitter-docs/

I am not an expert here!

Other pythonic approaches using automatic differentiation

Combine-tensorflow
Github

https://github.com/zfit/zfit

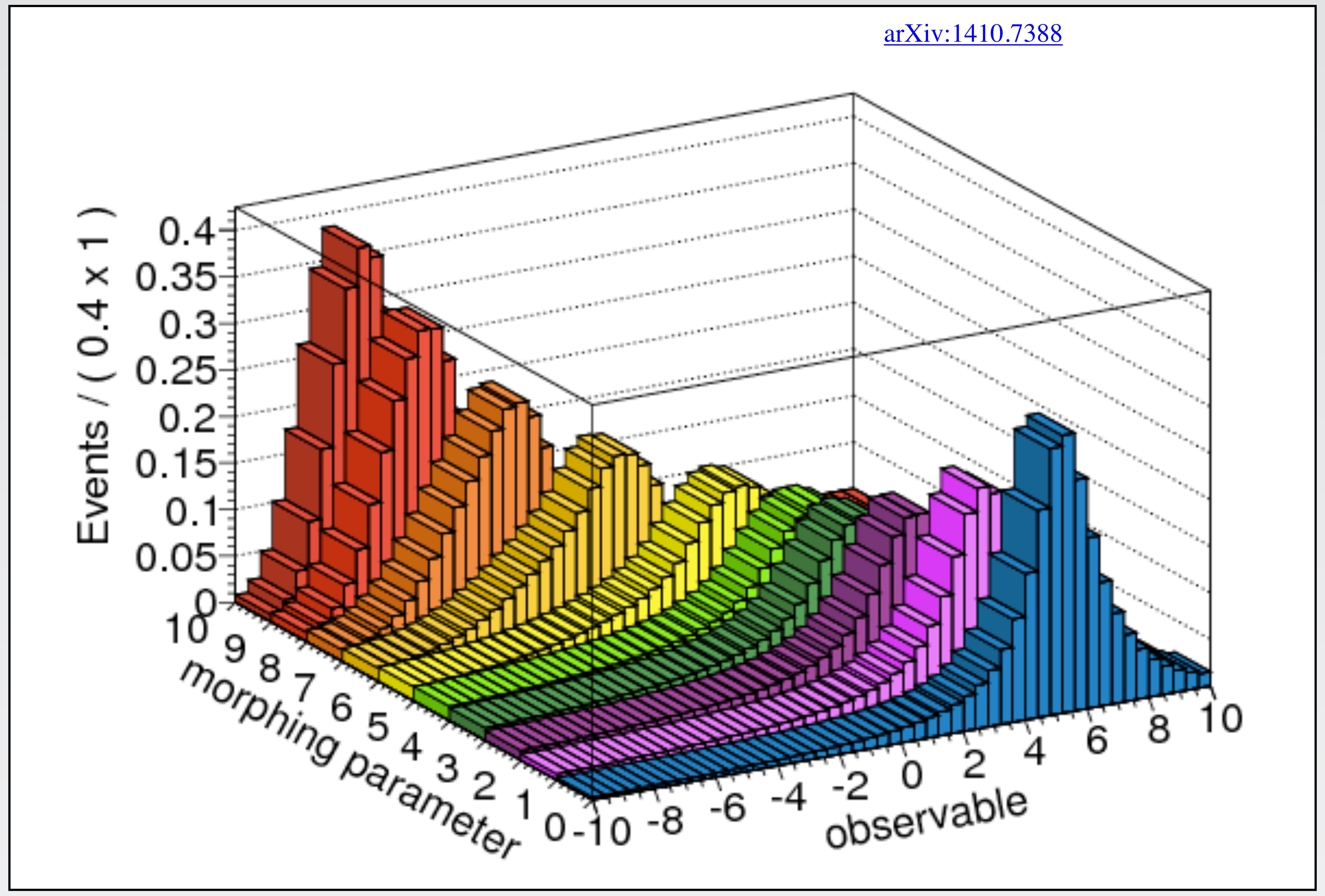
https://cds.cern.ch/record/2752552?ln=en

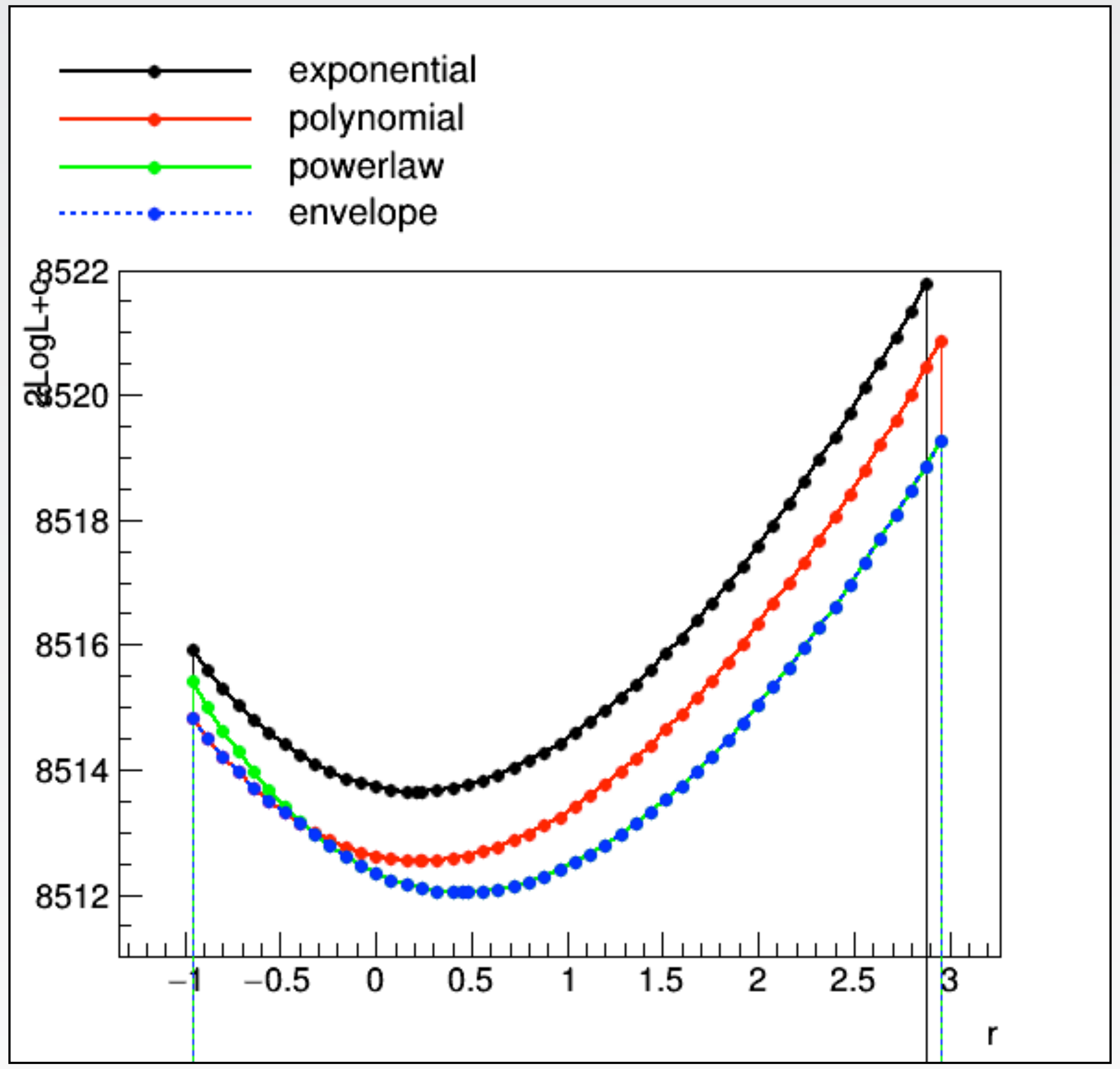
The suite of tools

- Modelling of the likelihood
  - Template morphing
  - Analytic functional forms
  - Uncertainties due to limited statistics in the nominal MC samples

- Different result representations
  - Limit setting (asymptotic+toy based)
  - Significance / p-value determination
  - Confidence intervals
  - Discrete profiling
  - Unfolding (combine)

- Diagnostics
  - Uncertainty impacts
  - nuisance parameter constraints and pulls
  - Goodness of fit tests
  - Checks on toy / Asimov data





- Very powerful tools are available: impossible to cover them here

Introduce Correlation Assumptions

- Consider two measurements from two different experiments
  - one or more parameters of interest
  - nuisance parameters, representing sources of systematic uncertainties or analysis repeated for each variation
  - $-2\ln L^{\alpha}(\mu,\theta^{\alpha}) = -2\ln P(y^{\alpha}|\mu,\theta^{\alpha}) + \sum_i \theta_i^{\alpha 2}/\sigma_i^{\alpha 2}$ or $-2\ln L^{\alpha}(\mu,\lambda^{\alpha}) = -2\ln P(y^{\alpha}|\mu,\lambda^{\alpha}) + \sum_i \lambda_i^{\alpha 2}$

- Correlations between uncertainty sources for different experiments

> "Different parameter settings and strategies are used to evaluate the systematic uncertainties due to initial and final state radiation. Preliminary investigations indicate that the methodologies used are approximately equivalent, and describe to a large extent the same physics aspects. Moreover, different baseline Monte Carlo programs and hadronisation models are used for the evaluation of the MC modelling systematics. In the presence of these underlying differences, we opt to reduce the ***assumed correlations across experiments for the MC and Rad categories*** from 100% to 50%."
>
> CMS-PAS-TOP-12-001

- Express as $\sum_i \lambda_i^{\alpha 2} \rightarrow \lambda^T (C^{\alpha})^{-1} \lambda$

- Then when combining the likelihoods this becomes

$$C = \begin{pmatrix} \mathbb{1} & A & \cdots \\ A & \mathbb{1} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix},$$

Pitfalls in practice: Ill-defined Assumptions

- Assumed strong correlations between
  - λ_0 and λ_1
  - λ_1 and λ_2

- This is equivalent to stating:
  - λ_0 and λ_1 describe (almost) the same uncertainty
  - λ_1 and λ_2 describe (almost) the same uncertainty
  - λ_0 and λ_2 have no correlation whatsoever

$$C = \begin{pmatrix} 1 & & \cdots \\ 0.99 & 1 & \\ 0 & 0.99 & 1 \end{pmatrix}$$

- In practice, similar situations can occur quite frequently when assigning correlations

- These can easily lead to C being not positive definite and/or not invertible
  - If this occurs, and the fit fails, this is not an issue of the program but a result of ill defined assumptions

"Different parameter settings and strategies are used to evaluate the systematic uncertainties due to initial and final state radiation. Preliminary investigations indicate that the methodologies used are approximately equivalent, and describe to a large extent the same physics aspects. Moreover, different baseline Monte Carlo programs and hadronisation models are used for the evaluation of the MC modelling systematics. In the presence of these underlying differences, we opt to reduce the ***assumed correlations across experiments for the MC and Rad categories*** from 100% to 50%."

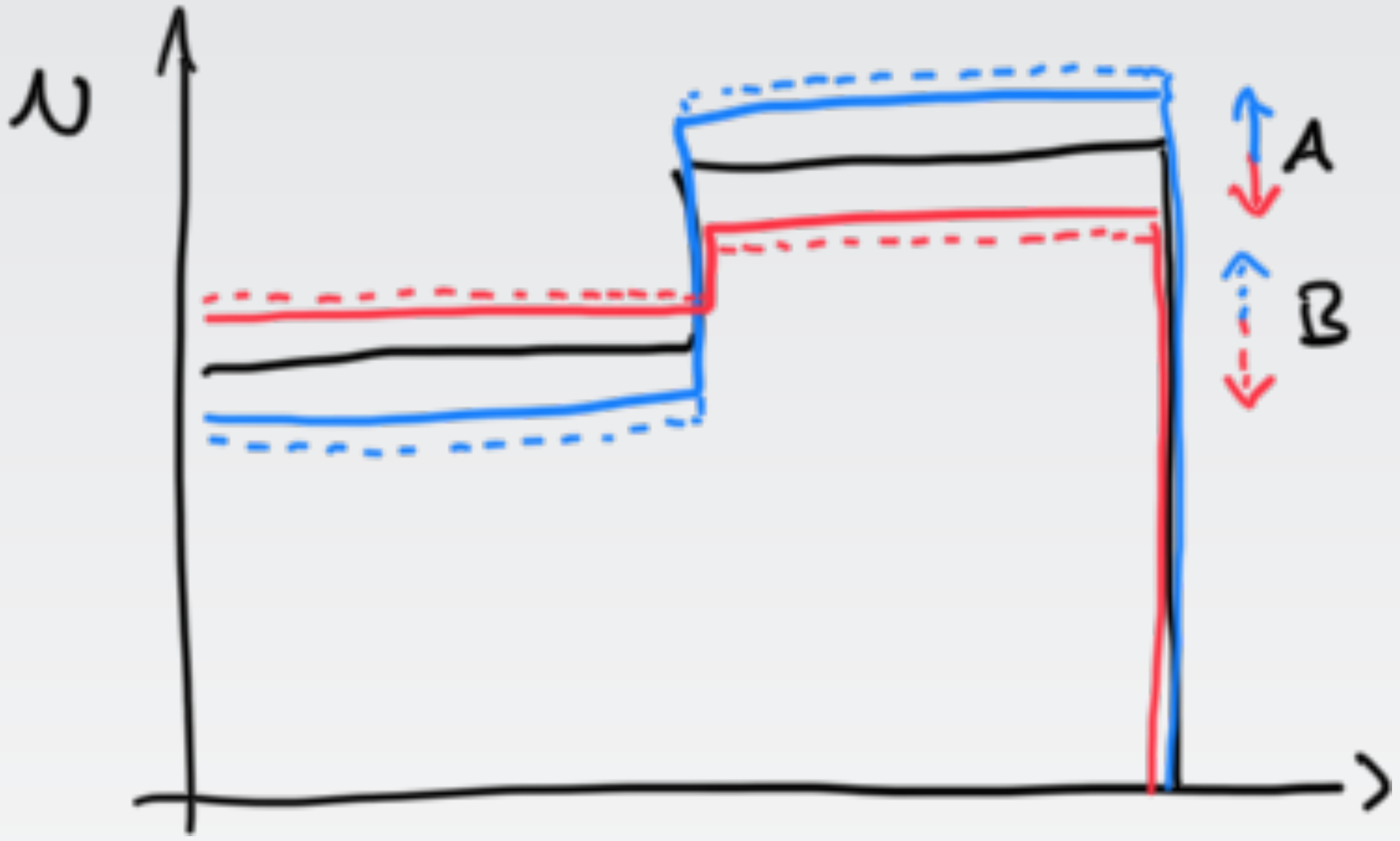
CMS-PAS-TOP-12-001

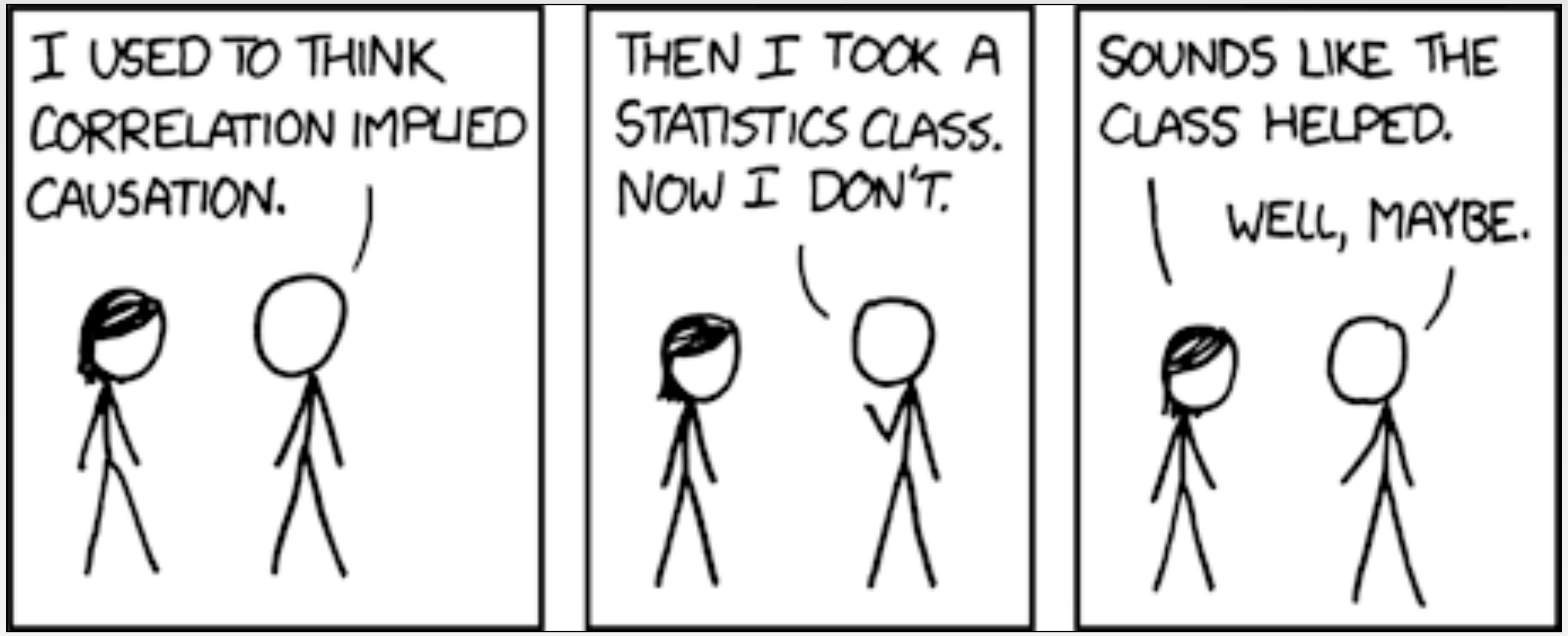
➡Even if C is invertible and p.d. (e.g. smaller correlations), it might still be worth thinking about the assumption again

Pitfalls in practice: Cause and Effect

"Uncertainty A is correlated with uncertainty B"

- What does it really (not) mean?

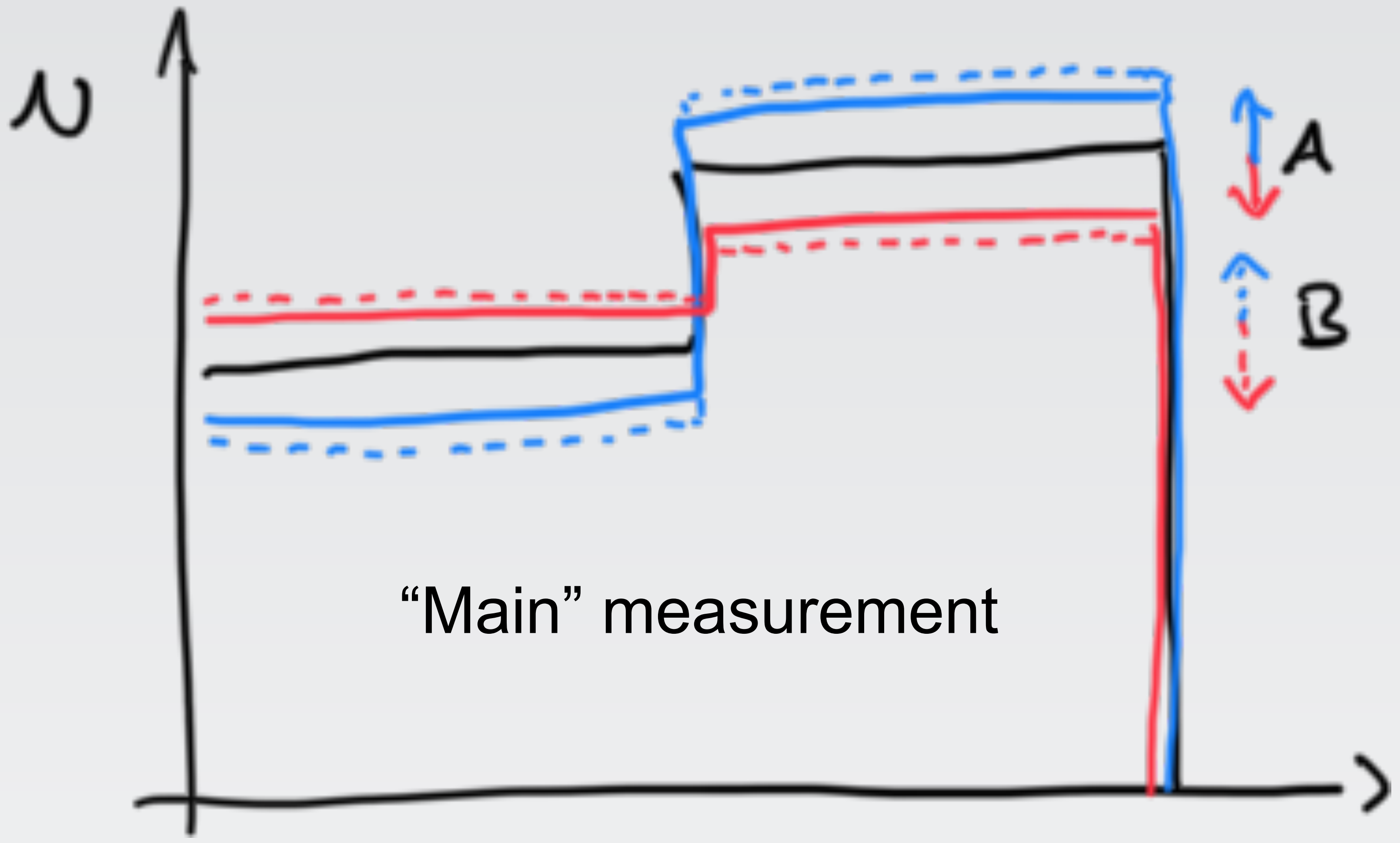




- The **variation of sources** of uncertainties having a **similar effect** on a measurement (or limit) **does not imply** anything w.r.t. the correlations of the sources

- The measured distribution cannot tell us anything about the correlations between uncertainty sources

Ancillary Measurements

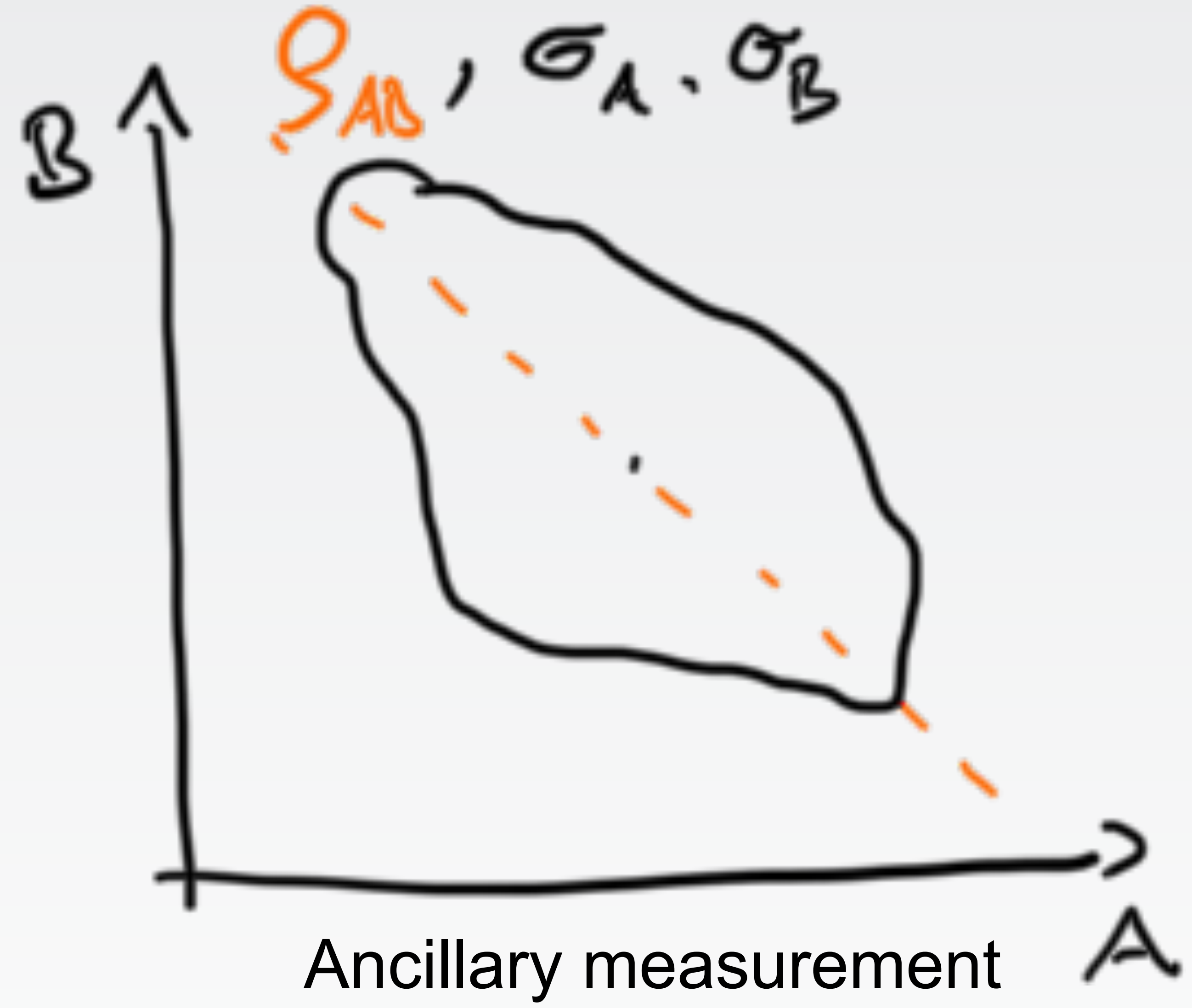


Approximation

$\rho_{AB}, \sigma_A . \sigma_B$

Glen's talk
slide 8

$$P(\boldsymbol{u}\,|\theta, \lambda, \xi)$$



Ancillary measurement
of A and B

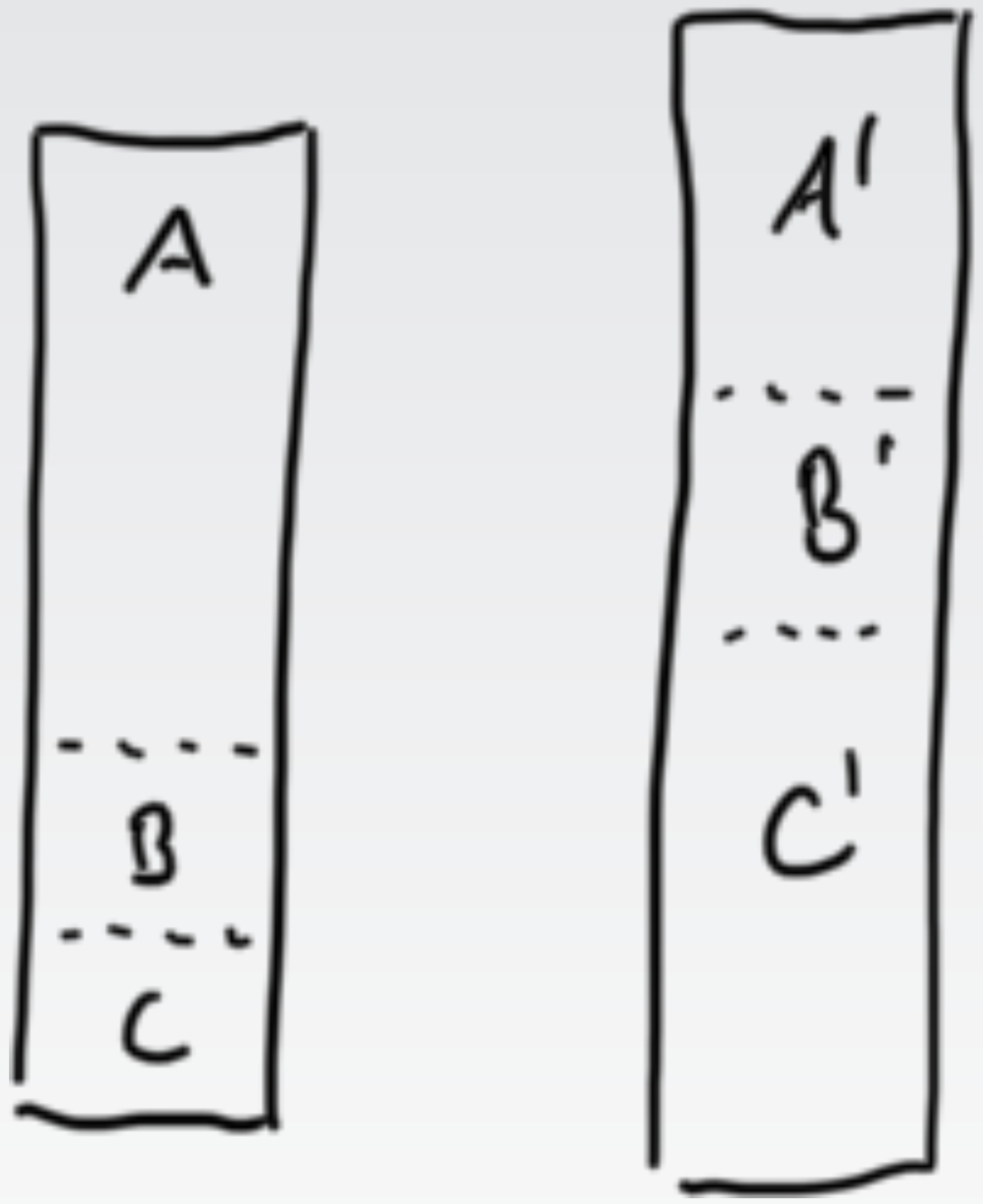
- When applying systematics variations, we just approximate the original ancillary measurements leading to the uncertainties
- **Only** those ancillary measurements of A and B can be used to judge correlations between the sources A and B

Pitfalls in Practice: Grouped Correlations

- Estimating the correlations between individual sources often challenging

- Often used fall-back: correlations between groups of uncertainties
  “The modelling uncertainty sources of the measurements performed by ATLAS and CMS should be somewhat correlated”

- Often right from a physics perspective, but also ill defined
  - Impact on different measurements
  - Signs are lost
  - Impossible to treat fully consistently if correlations between sources in the same group exist



‘Quasi-solutions’

- Try to avoid assigning correlations to groups in the first place and try to do it source-by-source
- Possibly try to infer source-by-source correlations from group correlations
- Make sure to check robustness of combined result against these assumptions

Approximate Likelihood Combinations of Measurements

- Consider a measurement of an observable that results in a value X and and total uncertainty σ
  - Performed by repeating the measurement for each uncertainty source
  - (From here on fully Gaussian)

- Can be easily expressed as an approximate likelihood (/chi2)

$$-2\ln\tilde{L}(\mu,\theta) = (x-(\mu+\theta))^2/\sigma^2_{\text{stat}} + \sum_i \theta_i^2/\sigma^2_{\theta_i}$$

$$-2\ln\tilde{L}(\mu,\theta) = \chi^2_s(\mu,\theta) + \chi^2_P(\theta)$$

	rel. unc. [%]
syst. A	5
syst. B	3
stat.	2

- Can be used for the combination directly

- For more complex measurements: Taylor expansion around maximum
  - The first term is constant → does not affect the result
  - The second term (first derivatives) are zero, since we maximised the likelihood to perform the measurement
  - The third term (second derivatives) does not vanish ~ Hessian

- **Example** of approximate likelihood: try to express as 3 terms (convino method)

$$-2\ln\tilde{L} = \chi^2 = \chi^2_s(\mu,\theta) + \chi^2_u(?) + \chi^2_P(\theta)$$

"Everything more complex": nuisance parameter constraints, correlations

Convino Approach

$-2\ln\tilde{L} = \chi^2 = \chi^2_s(\mu,\theta) + \chi^2_u(?) + \chi^2_P(\theta)$ $\lambda = \theta/\sigma_\theta \rightarrow -2\ln\tilde{L} = \chi^2 = \chi^2_s(\mu,\lambda) + \chi^2_u(?) + \chi^2_P(\lambda)$

- Assume
  - $\chi^2_u(?) = \chi^2_u(\lambda) = \lambda^T D\lambda$ Correlations between nuisance parameters and additional constraints on them
  - $\chi^2_P(\lambda) = \lambda^T 1\lambda = \sum_i P_i(\lambda)$ "Original" penalty terms

(Will help introduce correlation assumptions later)

- Assume the covariance of the inputs measurements to be known *(minimum requirement, more in the next lecture)*

- Re-organise Hessian of measurement α in the following form

$$H^\alpha_{\text{in}} = \begin{pmatrix} \tilde{C} & \kappa^T \\ \kappa & M \end{pmatrix}^\alpha$$

- Derive parameters of the approximate likelihood by comparing analytical Hessian and input Hessian

$$-2\ln(L^\alpha) = \chi^2_\alpha = \sum_{\mu\nu} \boxed{\tilde{M}^\alpha_{\mu\nu}} \frac{\xi^\alpha_\mu \xi^\alpha_\nu}{\tau^\alpha_\mu \tau^\alpha_\nu} + \sum_{ij} \lambda_i \boxed{D^\alpha_{ij}} \lambda_j + \sum_i \boxed{P_i^2(\lambda_i)}, \text{ with}$$

$$\xi^\alpha_\mu = x^\alpha_\mu - \left(\bar{x}_\mu \prod_i (\lambda_i \boxed{K^\alpha_{\mu i}} / x^\alpha_\mu + 1) + \sum_i \lambda_i \boxed{k^\alpha_{\mu i}} \right). \quad \tau^\alpha_\mu = \frac{\bar{x}_\mu}{x^\alpha_\mu}$$

$$H^\alpha_{\text{in}} = \begin{pmatrix} \tilde{C} & \kappa^T \\ \kappa & M \end{pmatrix}^\alpha$$

$$\tilde{H}^\alpha_{\mu\nu}(0) = \frac{1}{2} \left(\frac{\partial^2}{\partial \Delta x^\alpha_\mu \partial \Delta x^\alpha_\nu} \chi^2_\alpha \right) \Bigg|_{\lambda_i = 0,\ \Delta x^\alpha_\mu = 0\ \forall\ i,\ \mu} = \boxed{\tilde{M}_{\mu\nu}}, \quad \Leftrightarrow \quad M$$

$$\tilde{H}^\alpha_{\mu i}(0) = \frac{1}{2} \left(\frac{\partial^2}{\partial \Delta x^\alpha_\mu \partial \lambda_i} \chi^2_\alpha \right) \Bigg|_{\lambda_i = 0,\ \Delta x^\alpha_\mu = 0\ \forall\ i,\ \mu} = \sum_\nu \tilde{M}_{\mu\nu} (-\tilde{k}^\alpha_{\nu i}),$$

$$\Leftrightarrow \quad \boxed{\tilde{k}^\alpha_{\nu i}} = -\sum_\mu ((M^\alpha)^{-1})_{\mu\nu}\ \kappa^\alpha_{\mu i}.$$

Identify χ^2_u

$$-2\ln(L^\alpha) = \chi^2_\alpha = \sum_{\mu\nu} \boxed{\tilde{M}^\alpha_{\mu\nu}} \frac{\xi^\alpha_\mu \xi^\alpha_\nu}{\tau^\alpha_\mu \tau^\alpha_\nu} + \sum_{ij} \lambda_i \boxed{D^\alpha_{ij}} \lambda_j + \sum_i \boxed{P^2_i(\lambda_i)}, \text{ with}$$

$$\xi^\alpha_\mu = x^\alpha_\mu - \left(\bar{x}_\mu \prod_i (\lambda_i \boxed{K^\alpha_{\mu i}}/x^\alpha_\mu + 1) + \sum_i \lambda_i \boxed{k^\alpha_{\mu i}}\right). \quad \tau^\alpha_\mu = \frac{\bar{x}_\mu}{x^\alpha_\mu}$$

$$H^\alpha_{\text{in}} = \begin{pmatrix} \tilde{C} & \kappa^T \\ \kappa & M \end{pmatrix}^\alpha$$

$$\tilde{H}^\alpha_{ij}(0) = \frac{1}{2}\left(\frac{\partial^2}{\partial\lambda_i\partial\lambda_j}\chi^2_\alpha\right)\Bigg|_{\lambda_i=0,\ \Delta x^\alpha_\mu=0\ \forall\ i,\ \mu} = D^\alpha_{ij} + \delta_{ij}\frac{1}{2}\frac{\partial^2}{\partial\lambda_i^2}P_i^2\Bigg|_{\lambda_i=0} + \sum_{\mu\nu} M^\alpha_{\mu\nu}\tilde{k}^\alpha_{\nu i}\tilde{k}^\alpha_{\mu j}.$$

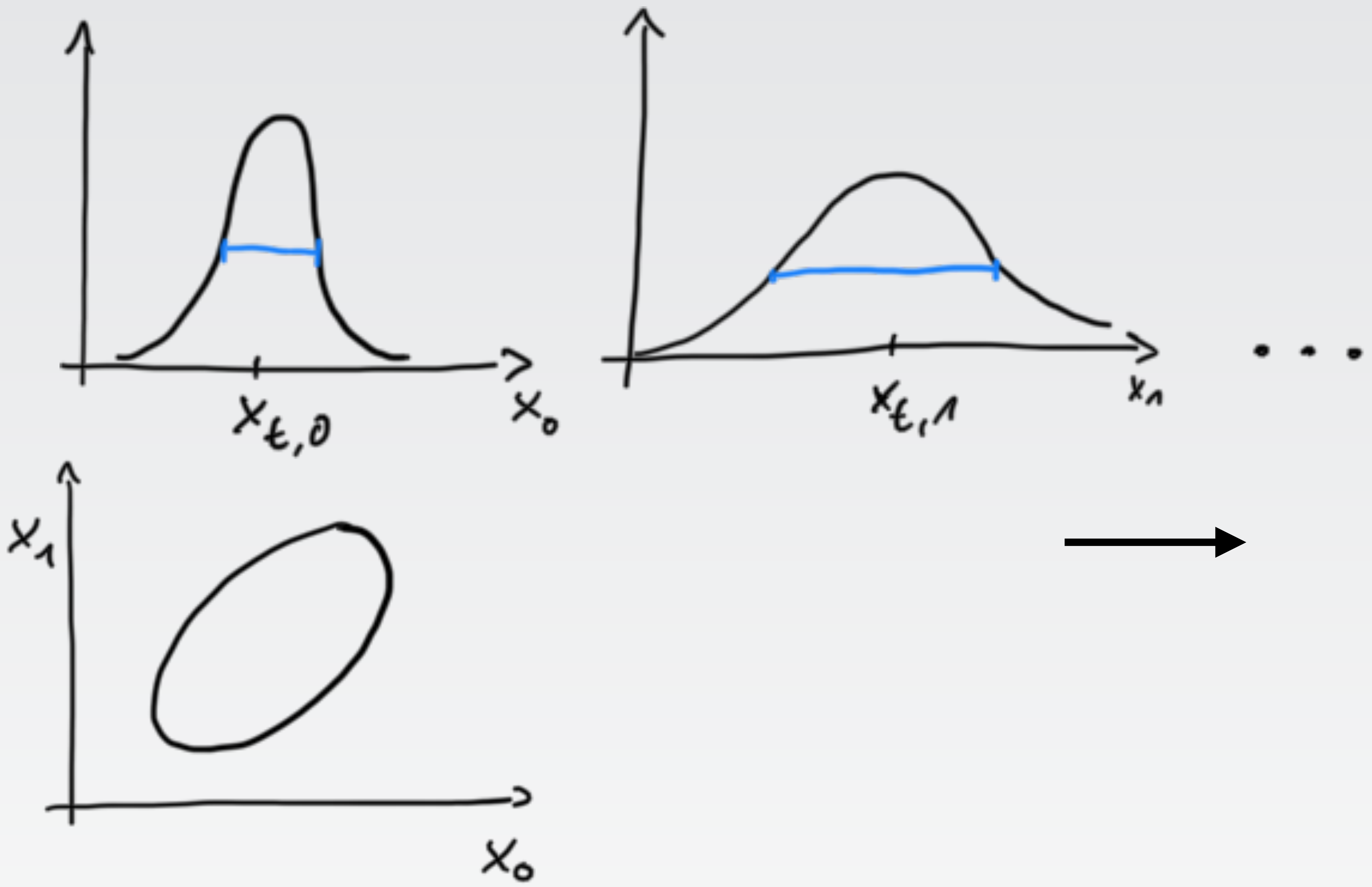
$$\Leftrightarrow \boxed{D^\alpha_{ij}} = \tilde{C}^\alpha_{ij} - \delta_{ij} - \sum_{\mu\nu} M^\alpha_{\mu\nu}\tilde{k}^\alpha_{\nu i}\tilde{k}^\alpha_{\mu j}.$$

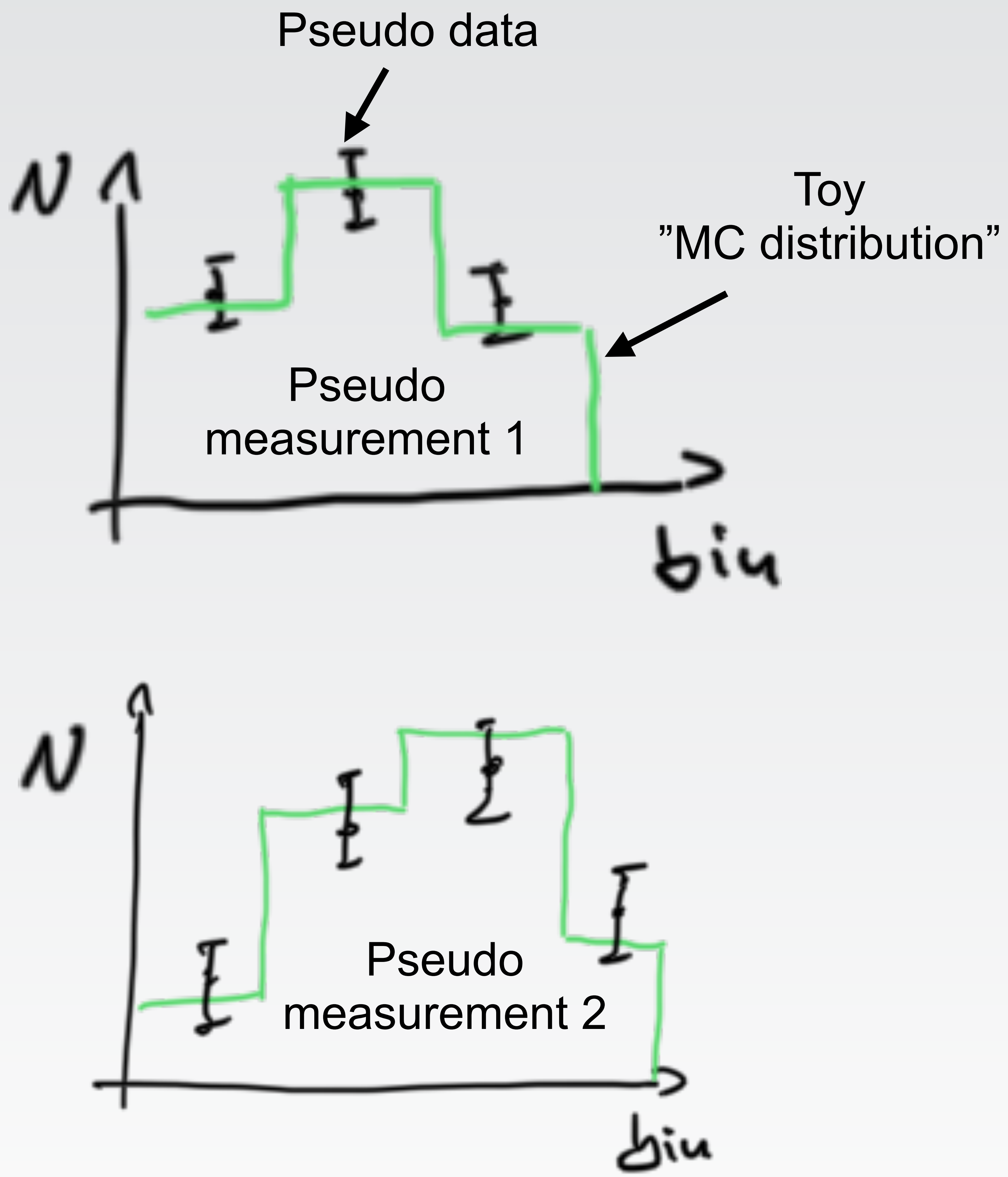
- $\chi^2_P(\lambda) = \lambda^T 1 \lambda = \sum_i P_i(\lambda)$ by definition, can use same method to introduce correlations discussed before ($\sum_i \lambda^{\alpha 2}_i \rightarrow \lambda^T (C^\alpha)^{-1}\lambda$)

➡All parameters identified, combination can be performed

Validation

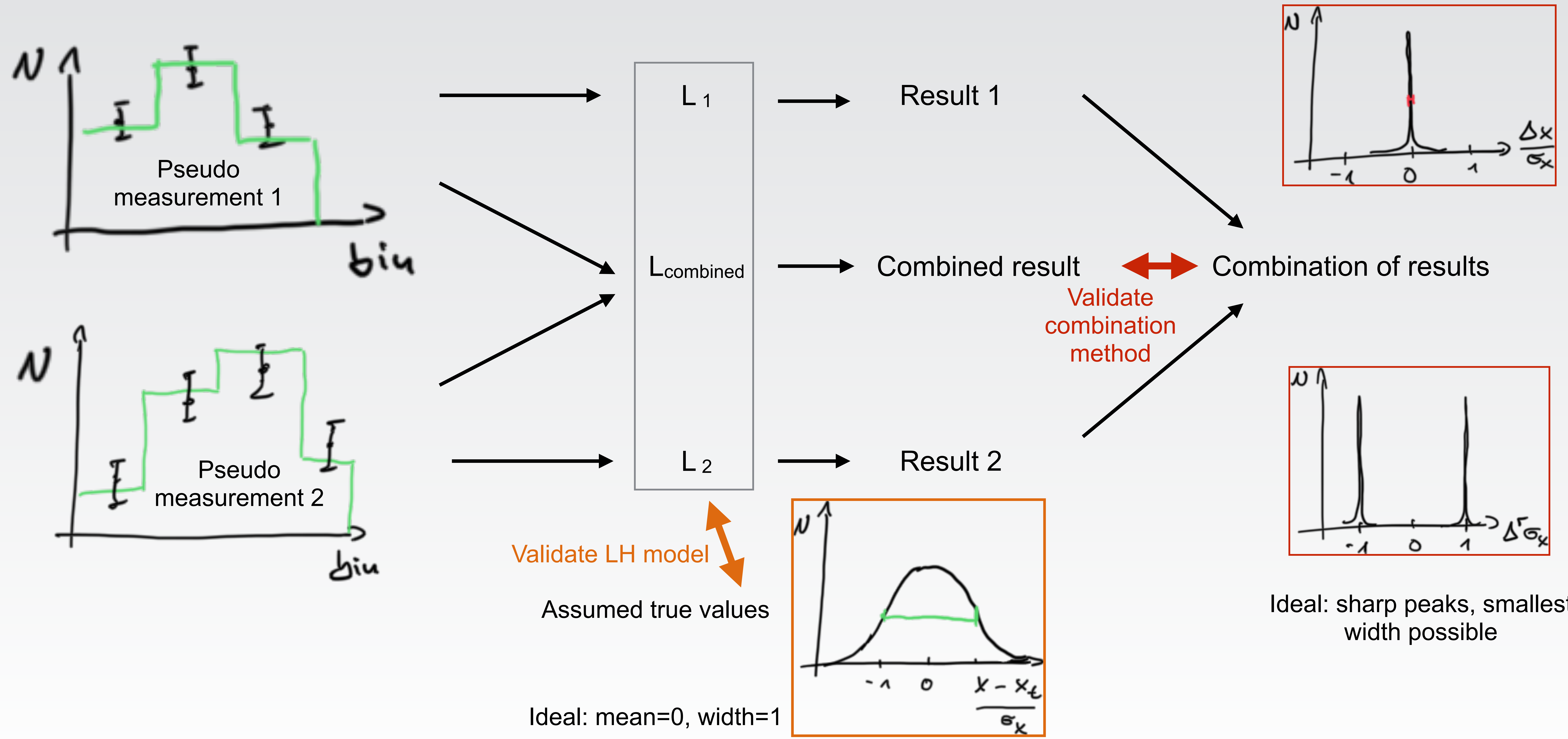
- Ultimate test of a model: toy experiments with known ‘true’ reference value
- E.g. generate independent bins, given an expectation value for each bin
  - $X_{t,i.}$ can also be chosen randomly



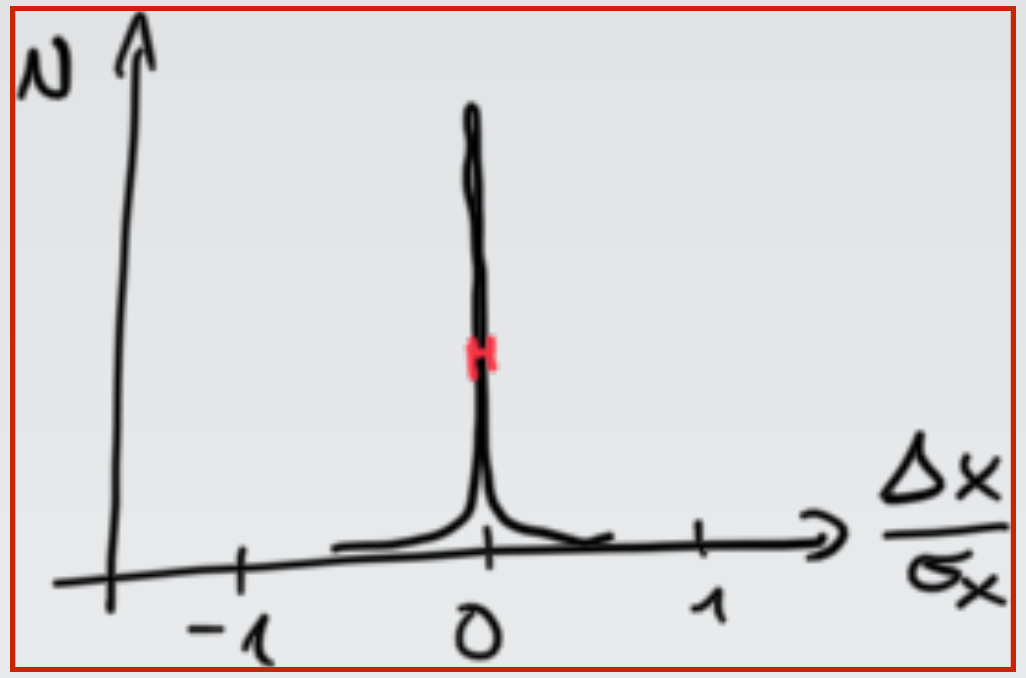


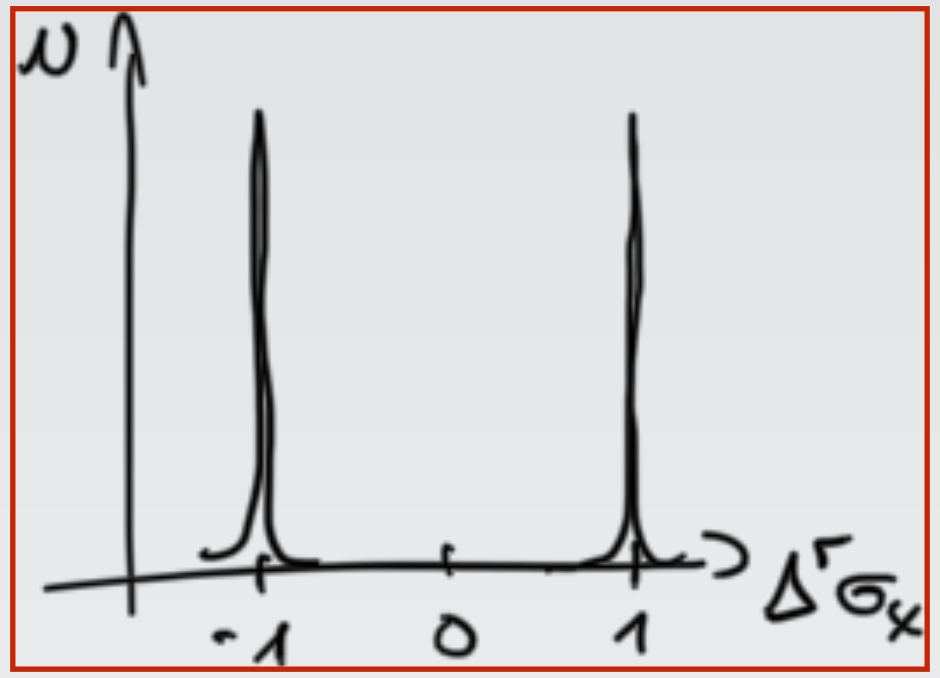
- The toy needs to include all possible variations that are expected to occur in practice
  - Statistical uncertainties over a wider range (low statistics, Gaussian limits), correlations
  - Systematic dependencies in each bin, for each x

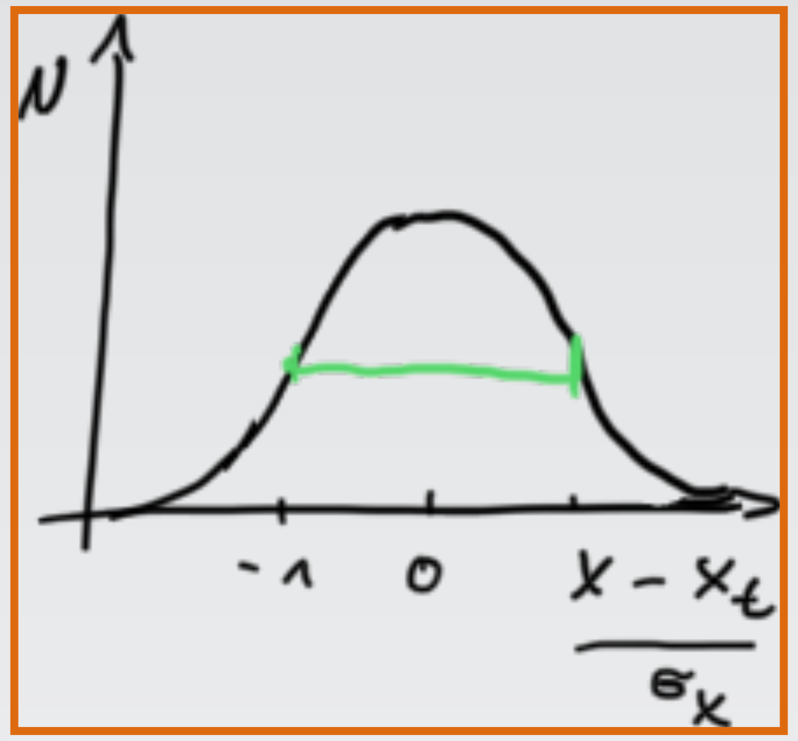
Validation workflow



Interpretation of the validation results







Validate the combination of measurements w.r.t. a combined measurement

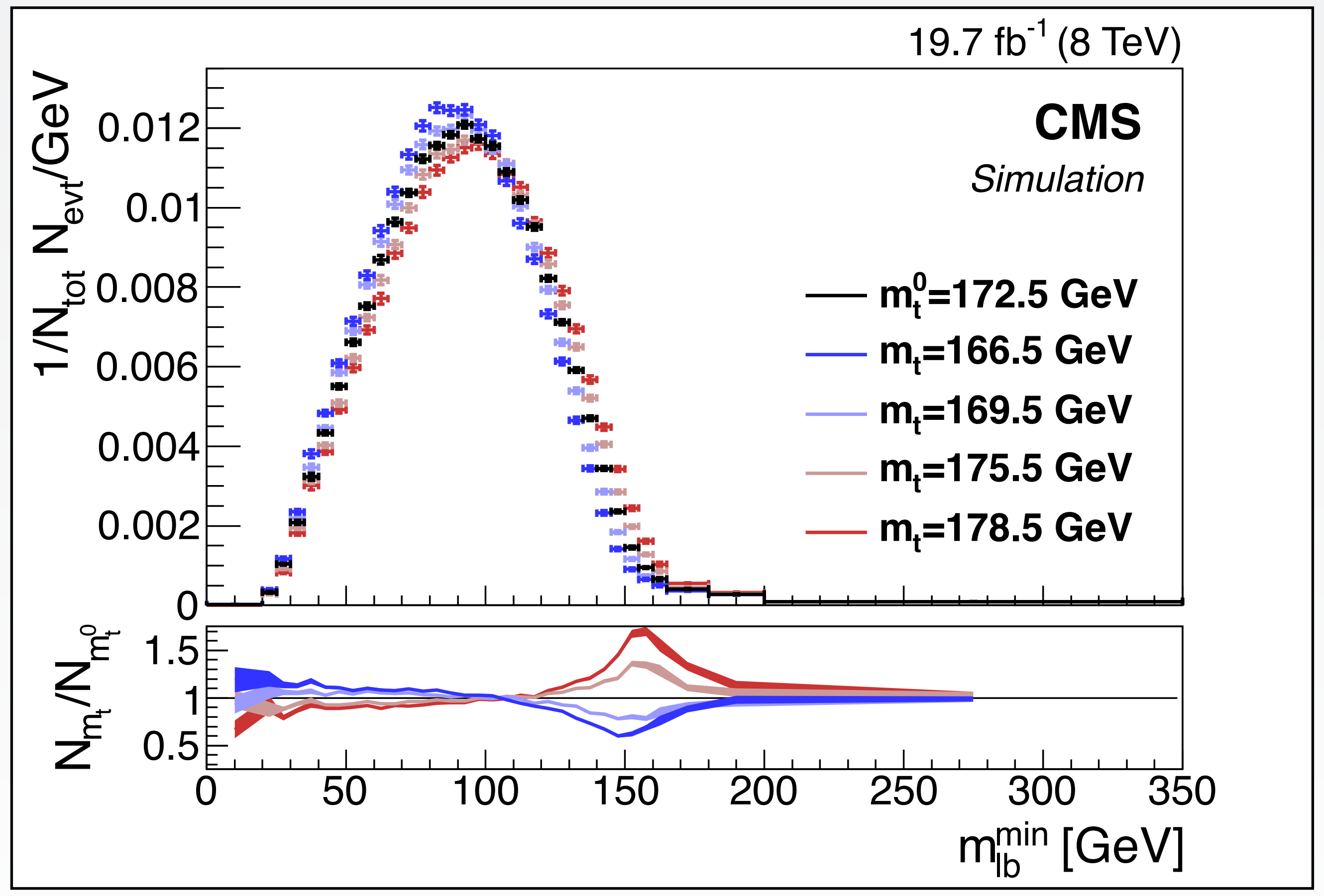
- How much does the (approximate) combination method bias the central result
- How well is the uncertainty (possibly asymmetric) modelled by the combination method

Validate the model fit to the data

- Does the model have a bias
- Does the model describe the fluctuations correctly and assigns the uncertainty correctly

➡**In either case, performing toy-based validation is usually very valuable**

➡**Can be used to validate even models that are not generally correct, but might still work for the specific case**



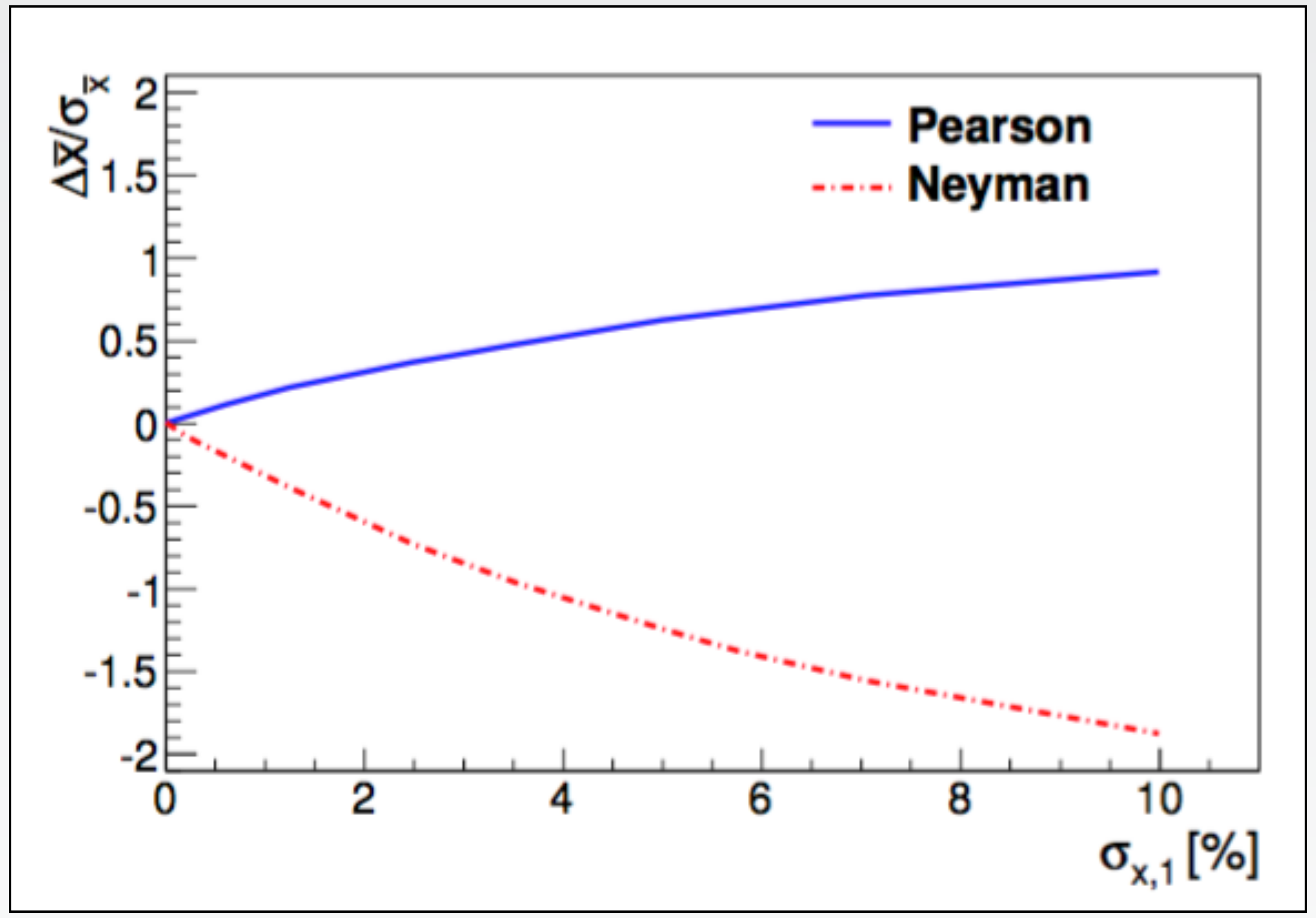
CMS PAS TOP-14-014:
Standard chi2 (no correlations) fully sufficient to determine mt from a normalised distribution

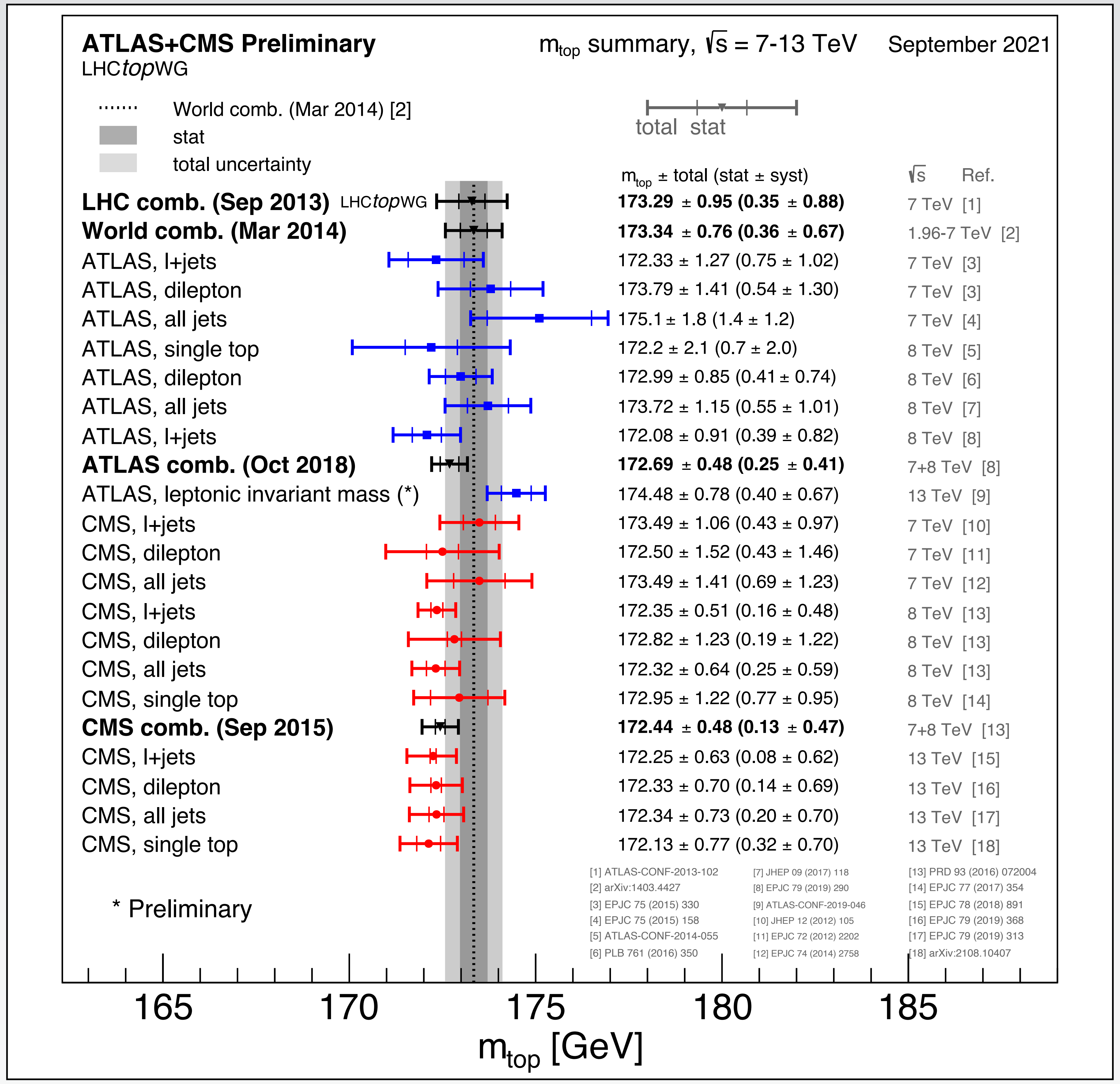
In practice: Convino Validation

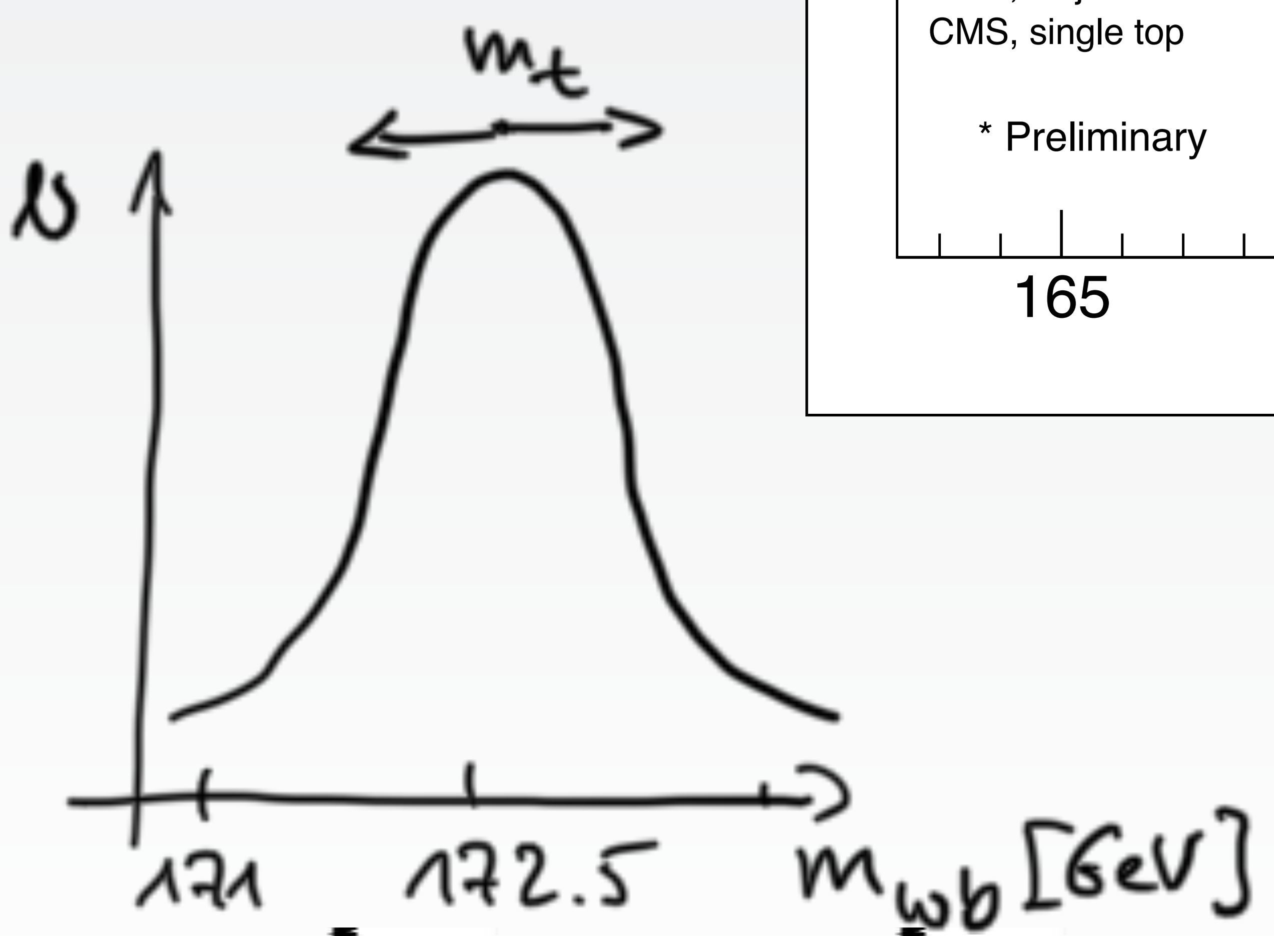
- Combine pseudo ‘cross section’ measurements
  - Binned distributions
  - The measured value scales with N
- No systematic uncertainties

$$\begin{aligned} x^a &= s \cdot 100, \\ x^b &= x^a + 10\sqrt{x^a}, \end{aligned}$$

- Bias as expected from a chi2 approach





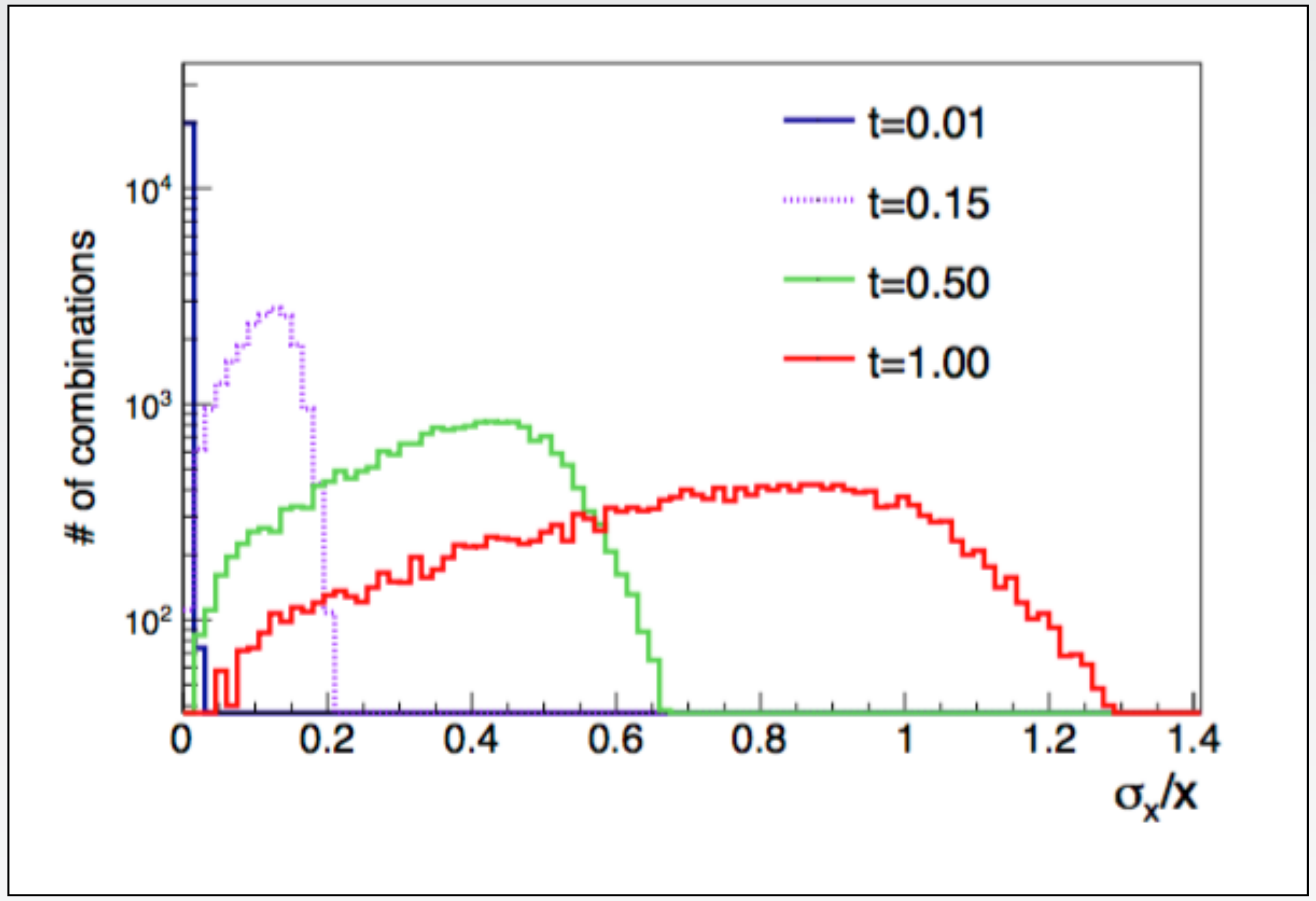


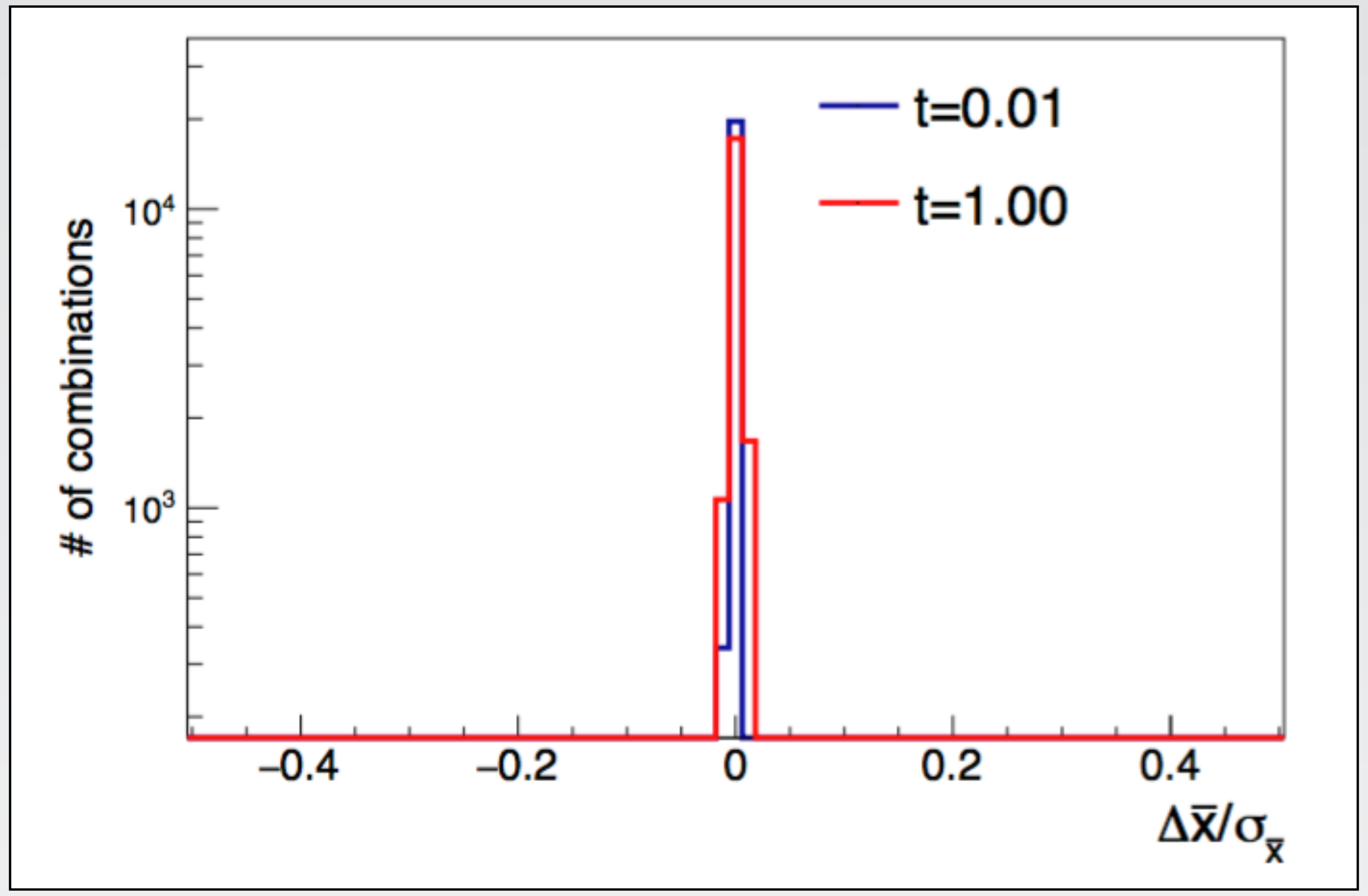
- For other combinations, the statistical uncertainty of the measurements might not scale with the number of events

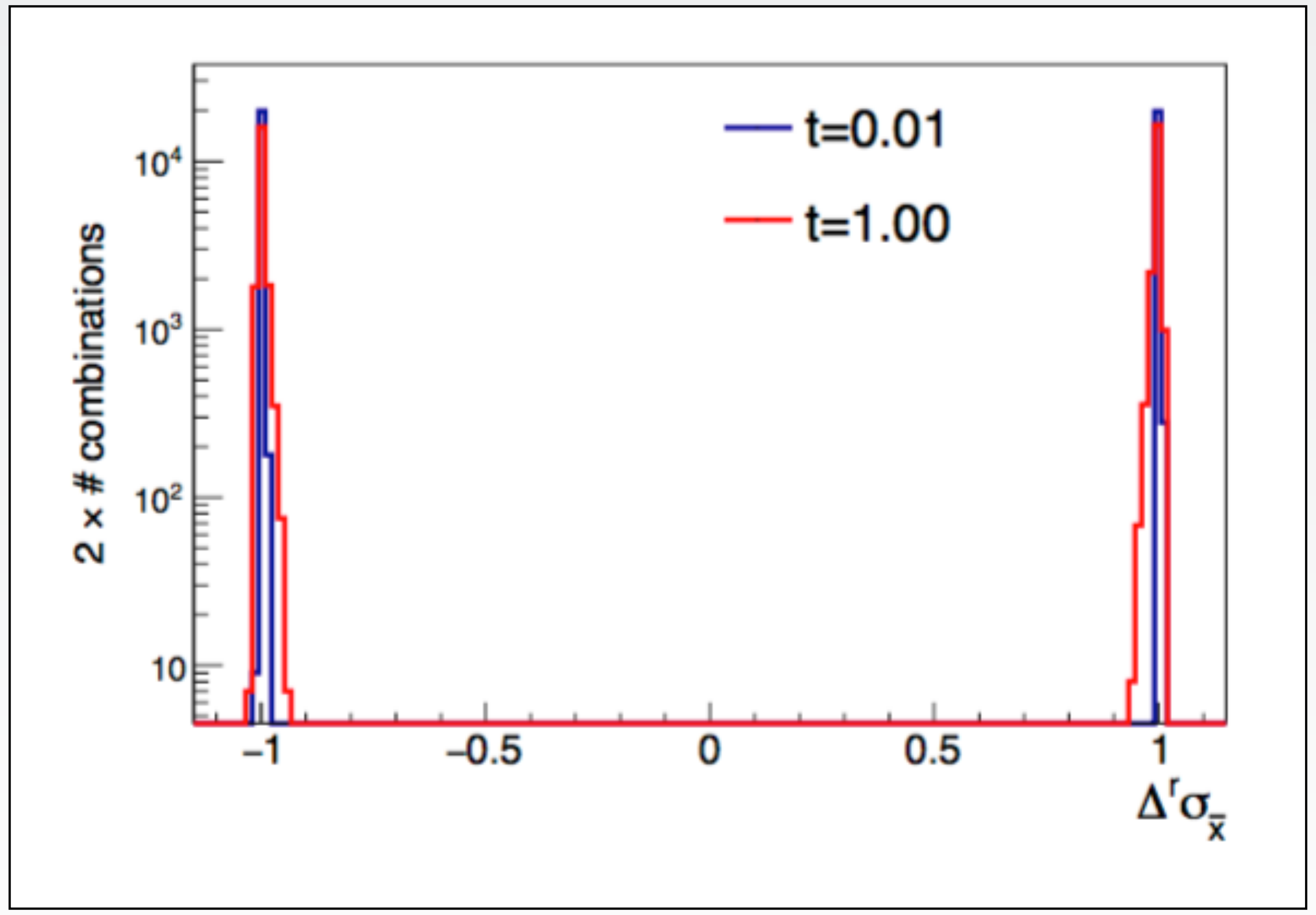
➡Choice needs to be adapted to the observable

Systematic uncertainties

- Reduce the relative impact from statistical uncertainties to 0.5%
- Introduce 2 systematic uncertainties per pseudo measurement
- Limit contribution per uncertainty to t
- Correlate one uncertainty of one with one of the other







➡Also validated using toys for many other scenarios

COMBINED HANDS-ON/ HANDS-OFF COMBINATION EXAMPLES

```
> bash
> cd /afs/cern.ch/user/j/jkiesele/public/Convino/latest
> source lxplus_env.sh
> cd
> mkdir convino_tutorial
> cd convino_tutorial
> convino /afs/cern.ch/user/j/jkiesele/public/Convino/latest/examples/exampleconfig.txt
> cp -r /afs/cern.ch/user/j/jkiesele/public/Convino/tutorial/* .
```

The Convino Inputs

```
> cd ~/convino_tutorial/1
```

```
# This is the 'base file' for example 1
# The header can and should contain some description


[global]
   #information mostly for extra options in differential measurements. Can be left empty in most cases

[end global]



[inputs]

    nFiles = 2

    file0 = measurement_1.txt
    file1 = measurement_2.txt

[end inputs]


[observables]

        # This block defines which estimates should be combined.
        # The combined name of the observable can be chosen freely.

        combined_a = estimate_a1 + estimate_a2

[end observables]

[correlations]

    # define correlations with the following syntax
    # lumi_a1 = (0.3) lumi_a2
    # this line assigns a correlation of 0.3 between lumi_a1 and lumi_a2

    # if there were more uncertainties, one can define multiple correlations in one line, e.g.
    # lumi_a1 = (0.3) lumi_a2  + (0.4) JES_a2

[end correlations]


[uncertainty impacts]

  # this is optional to evaluate uncertainty impacts (can take longer)
  # example (commented out)
  # lumi_total = lumi_a1 + lumi_a2


[end uncertainty impacts]
```

measurement_1.txt

```
[not fitted]

  # the uncertainties listed here describe orthogonal uncertainties,
  # e.g. when repeating the measurement after varying different sources of uncertainties.
  # the "stat" keyword is reserved and must be given last

                   lumi_a1   stat
  estimate_a1      4.        0.1

[end not fitted]


[systematics]

  # can be used to define whether an uncertainty is absolute or relative.
  # the default is absolute; example (commented out):
  # lumi_a1 = relative

[end systematics]


[estimates]

  # The central value of each estimates is defined here

  n_estimates = 1

  name_0      = estimate_a1
  value_0     = 100;

[end estimates]
```

measurement_2.txt

```
[not fitted]

                   lumi_a2    stat
  estimate_a2      4           0.1

[end not fitted]


[systematics]

[end systematics]


[estimates]

  # The central value of each estimates is defined here

  n_estimates = 1

  name_0      = estimate_a2
  value_0     = 100;


[end estimates]
```

```
> convino base_file.txt
> less result.txt
```

Example 1

result.txt

```
[pre-combine systematics correlations]
lumi_a1  1.0000000 0.0000000
lumi_a2 0.0000000  1.0000000
[end pre-combine systematics correlations]

[post-combine systematics correlations]
multiplied by: 1000
lumi_a1  1000.0000  998.74870
lumi_a2  998.74870  1000.0000
[end post-combine systematics correlations]

[pre-combine estimate correlations]

[end pre-combine estimate correlations]

[post-combine result correlations]
combined_a  1.0000000
[end post-combine result correlations]

[post-combine result covariance]
combined_a  8.0050000
[end post-combine result covariance]

combined (minimum chi^2=0):
combined_a: 100 +2.82931 -2.82931
[full correlation matrix]
multiplied by: 1000
lumi_a1      1000.0000  998.74870  999.37415
lumi_a2      998.74870  1000.0000  999.37415
combined_a   999.37415  999.37415  1000.0000
[end full correlation matrix]

[full covariance matrix]
lumi_a1      0.4991720  0.4985479  1.9954389
lumi_a2      0.4985479  0.4991720  1.9954389
combined_a   1.9954389  1.9954389  7.9867499
[end full covariance matrix]

Name       pull    constraint
lumi_a1   0.000     0.706
lumi_a2   0.000     0.706
Simple impact table: name, impact [%]
                combined_a        |
lumi_a1          2.8275401868783  |
lumi_a2          2.8275401868783  |


merged impacts
```

No correlations assumed

Post combination: large correlations

One result with large correlation with itself ;)

Central result and total uncertainty: how large is the reduction?

Full covariance (nuisance parameters and combined value)

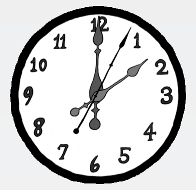
Pulls and constraints

Simple impacts (calculated from correlation coefficients)

- Introduce a correlation between the luminosity uncertainties of experiments A and B:

base_file.txt

```
[correlations]

    # define correlations with t
    # lumi_a1 = (0.3) lumi_a2
    # this line assigns a correl
```



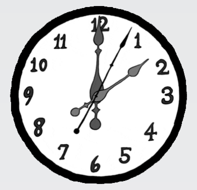
- Run convino with convino --prefix withcorr base_file.txt to create withcorr_result.txt
- How does the combined result change?
  - Printout
  - Result file

Correlation Scan

- Use the scan option of convino, adapt base_file.txt

```
lumi_a1 = (0 & -0.99 : 0.99) lumi_a2
```

scans the correlation coefficient from -0.99 to 0.99, and assumes a correlation of zero as the central value



```
> convino -sp base_file.txt
> mupdf scan_results/combined_a_lumi_a1_lumi_a2_0.pdf
```

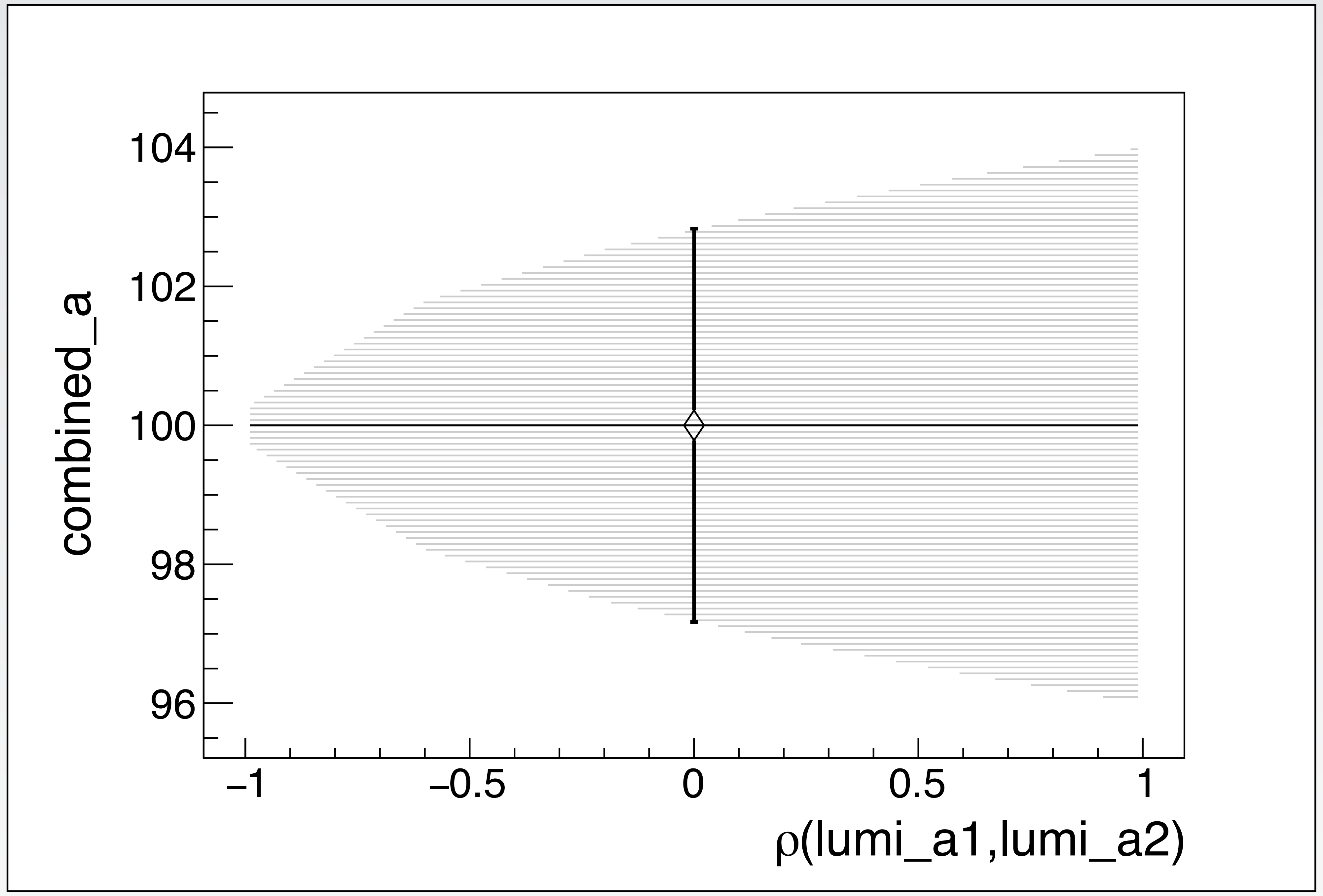
- What do we see? [scan_results/combined_a_lumi_a1_lumi_a2_0.pdf]

Correlation Scan

- Use the scan option of convino, adapt base_file.txt

```
lumi_a1 = (0 & -0.99 : 0.99) lumi_a2
```

scans the correlation coefficient from -0.99 to 0.99, and assumes a correlation of zero as the central value



- With larger correlation coefficient, the uncertainty increases.
- **This is not true in general. Large correlation != conservative**

Example 2: Correlations Reduce Uncertainty

- NB: Of course, if the impacts have opposite sign, the combined result gets more precise with larger correlation
- But, also in other cases… assume measurements with very different impact of uncertainties (example 2)

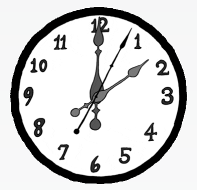
measurement_1.txt

```
[not fitted]

                   sys_b1    stat
   estimate_a1     0.5       0.1

[end not fitted]
```

measurement_2.txt

```
[not fitted]

                  lumi_a2  syst_a2  stat
  estimate_a2     1        3.       0.1

[end not fitted]
```

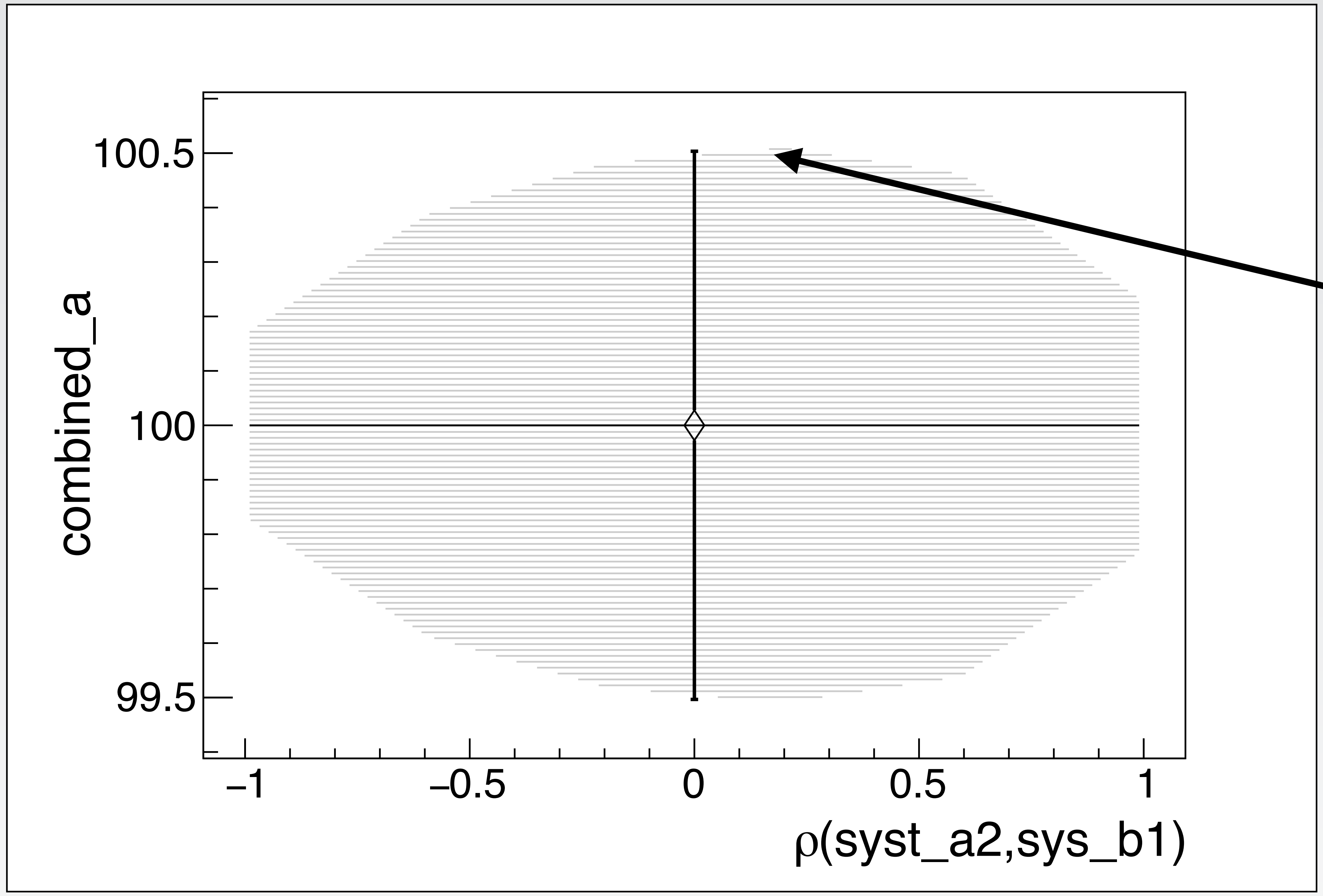
```
[correlations]

    syst_a2 =  (0 & -0.99 : 0.99) sys_b1

[end correlations]
```



```
> cd ../2
> convino -sp base_file.txt
> mupdf scan_results/combined_a_syst_a2_sys_b1_0.pdf
```

- What do we see?

Results Example 2



Largest uncertainty around rho = 0.2

- Large correlations does not equal conservative

➡Assume correlations carefully, and check the dependence (e.g. through scans)

Example 3: Ill-defined assumptions

measurement_1.txt

```
[not fitted]

                syst_a1    syst_b1    stat
  estimate_a1   3.         2.         0.1

[end not fitted]
```

measurement_2.txt — Edited

```
[not fitted]

                lumi_a2    syst_a2    stat
  estimate_a2   1          3.         0.1

[end not fitted]
```

```
[correlations]

    syst_a1 =  (0.99) lumi_a2 + (0.99) syst_a2
    syst_b1 = (0.99) lumi_a2

[end correlations]
```

$$C = \begin{array}{l} \text{syst\_a1} \\ \text{syst\_b1} \\ \text{lumi\_a2} \\ \text{syst\_a2} \end{array} \begin{pmatrix} 1 & & & \\ 0 & 1 & & \cdots \\ 0.99 & 0.99 & 1 & \\ 0.99 & 0 & 0 & 1 \end{pmatrix}$$

- Run the example
- What is the problem?
- Does removing the line syst_b1 = … help?

Example 4: Correlated Nuisances

measurement_1.txt

```
[correlation matrix]

  sys_a2       (1)    1
  sys_b2       (1)    0.3          1
  sys_c2       (0.1)  0.4          0.3          1
  sys_d2       (0.9)  0.3          0.2          0.6          1
  estimate_a1  (10.6) 0.2026484018 0.2652968037 0.8479452055 0.105936073  1

[end correlation matrix]
```

measurement_2.txt

```
[not fitted]

                 lumi_a2  syst_a2  stat
  estimate_a2    4        10.        0.1

[end not fitted]
```

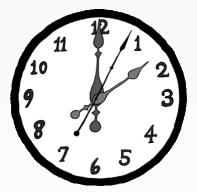
Constraint/total uncertainty

Largest contribution
from sys_c2,
Others much smaller

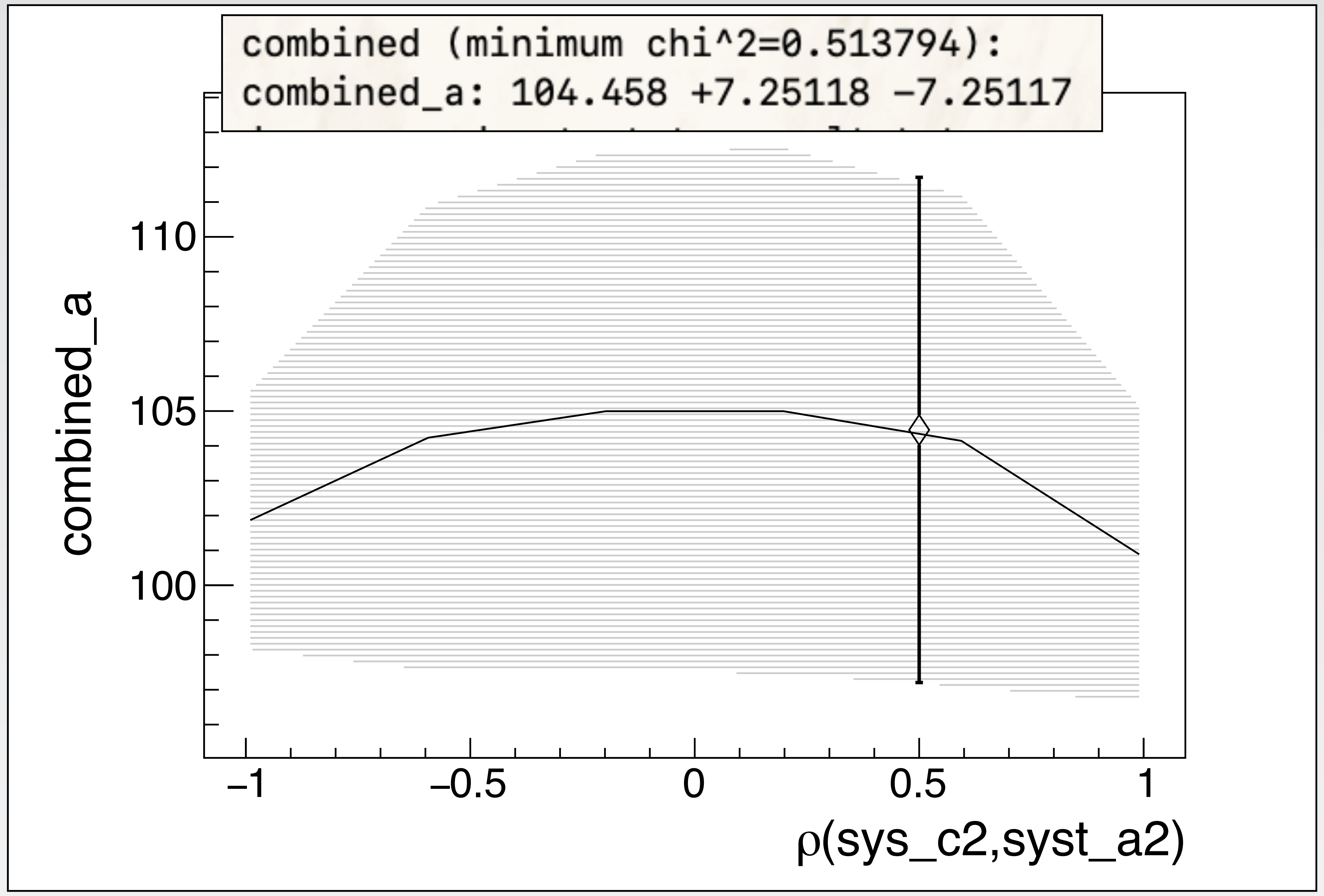
- More complex example
  - Main contribution from sys_c2
  - Moderate correlations between the nuisance parameters of measurement 1

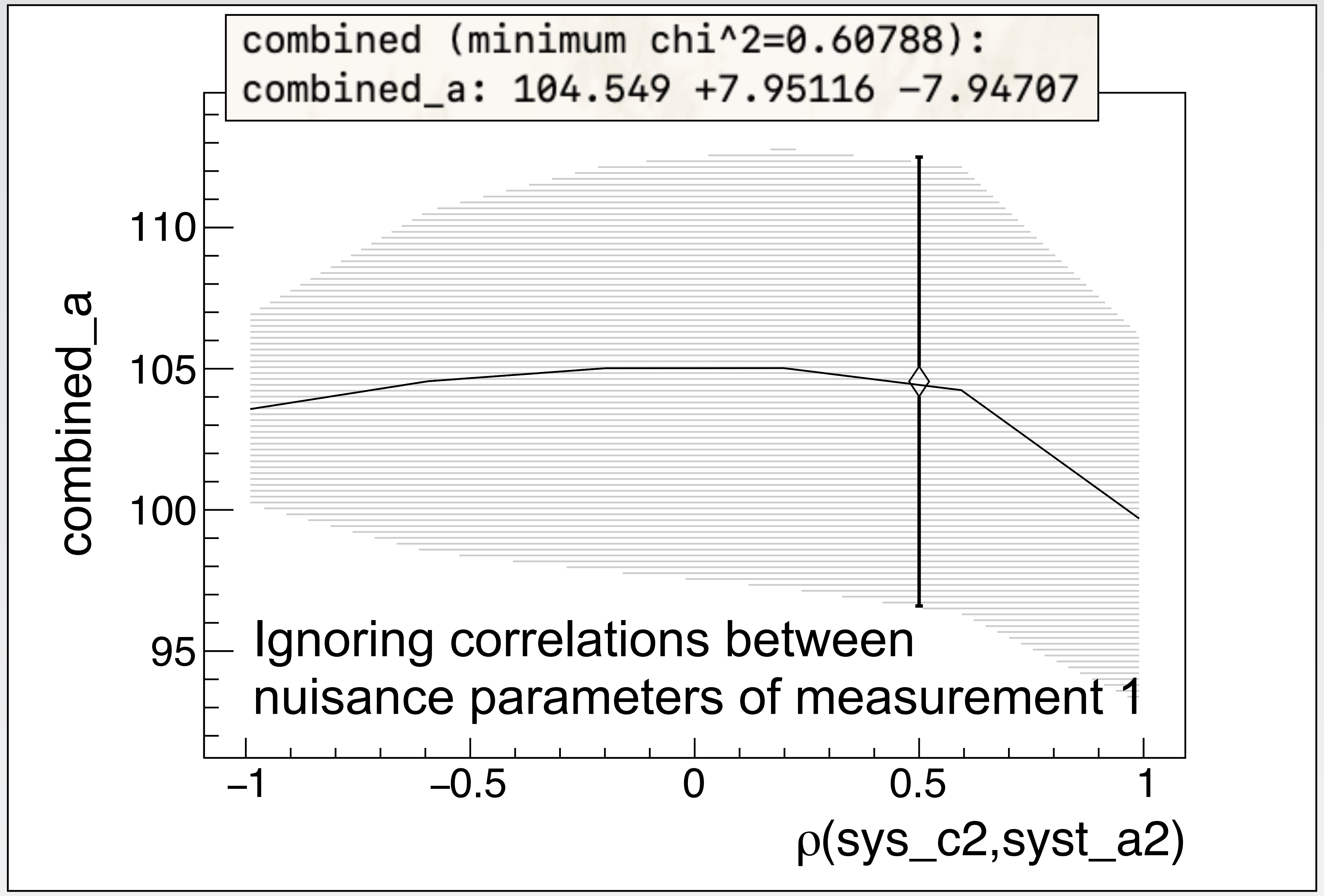
- Often these moderate correlations are considered as removable to apply methods and tools that cannot treat these correlations (e.g. BLUE)

- Compare the correlation scan and central results when including or ignoring these correlations (do a scan)
  - For convenience, the correlation matrix without correlations between nuisance parameters is in the same file
  - To create distinct outputs you can use the ‘prefix’ option: convino -sp --prefix ignorecorr base_file.txt



Results Example 4



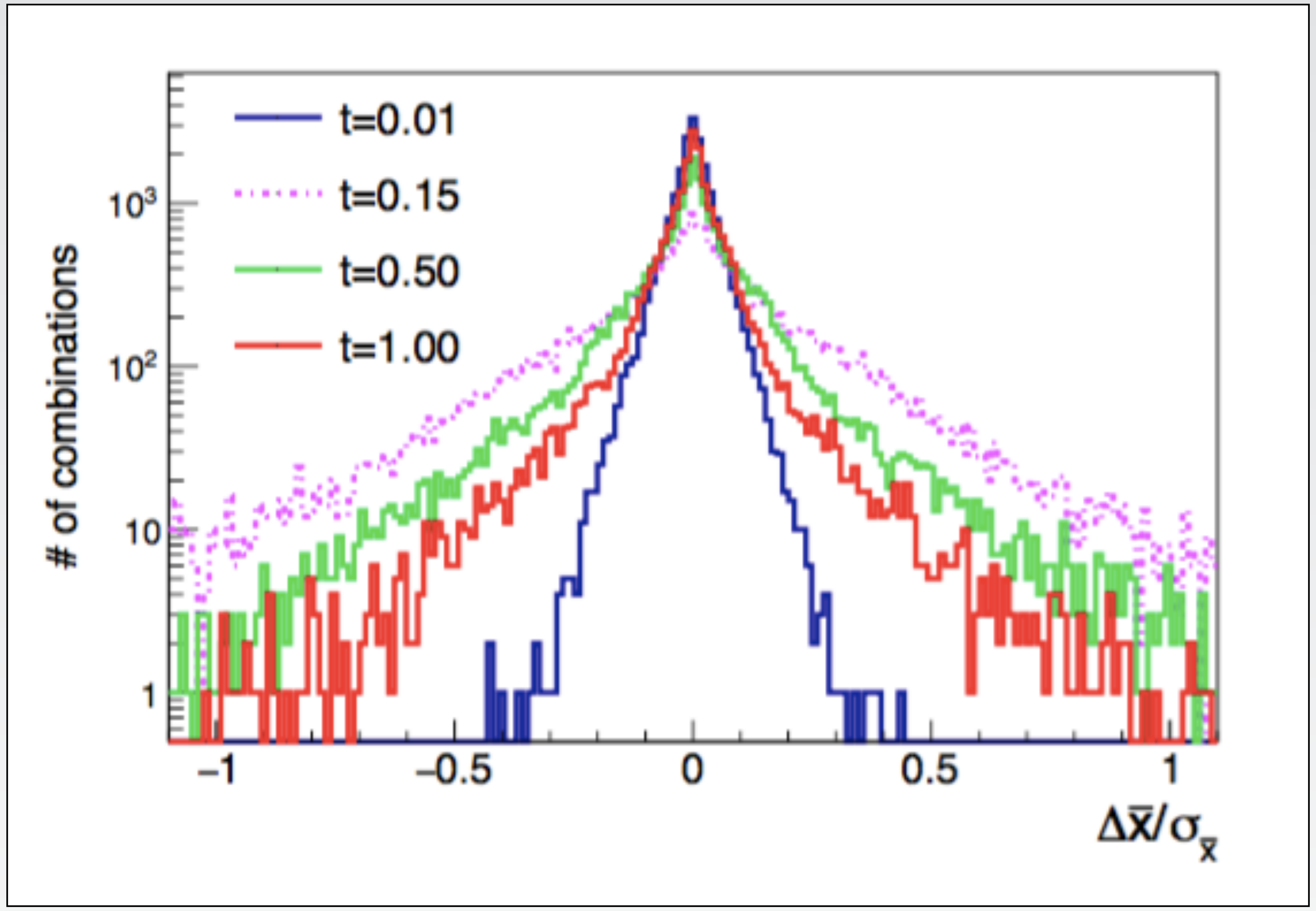


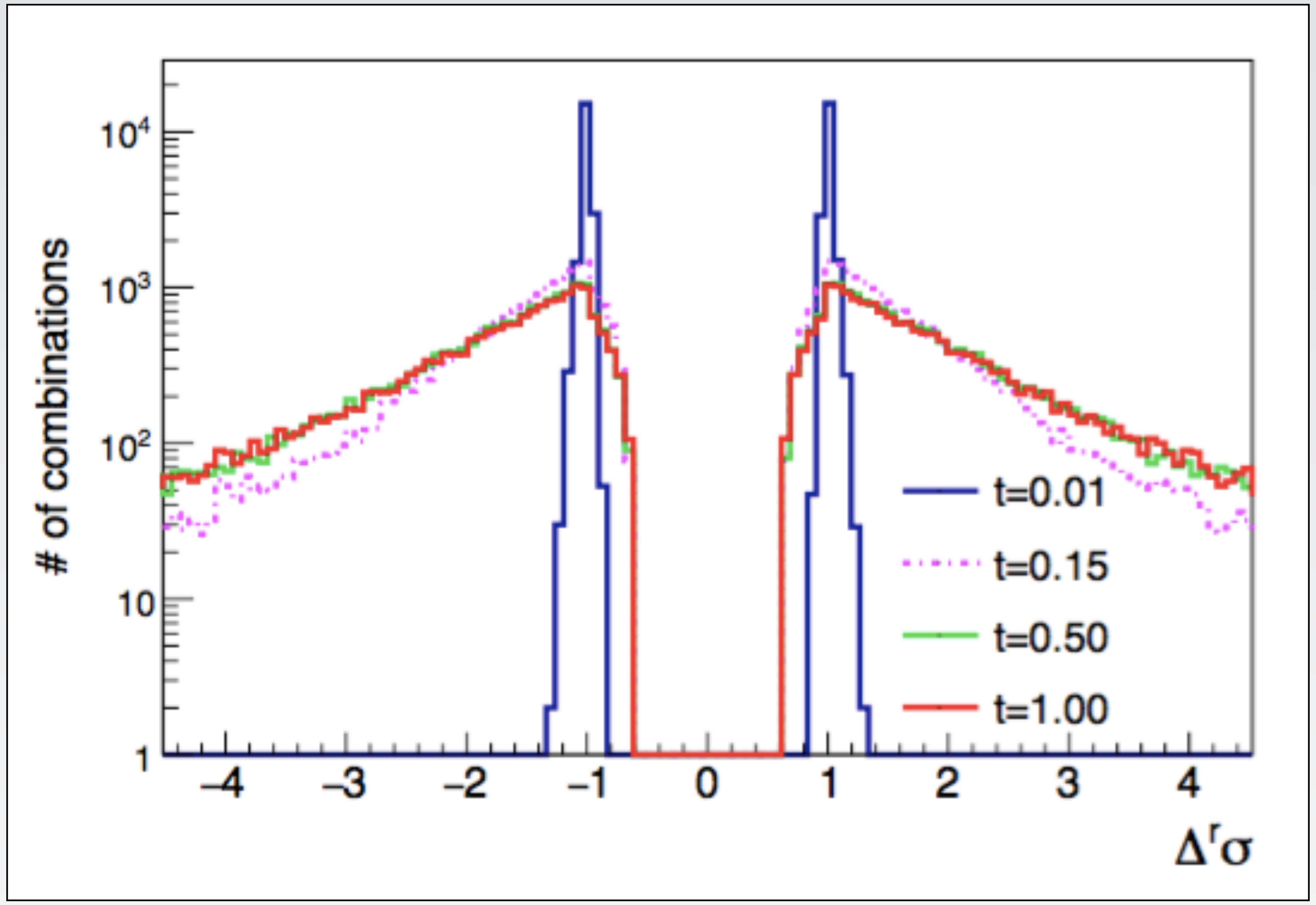
- Dependence and results differ quite significantly
  - Here, ignoring the correlations gives a larger uncertainty
  - This is not true in general

- These correlations can be more important than they seem at first glance

```
[correlation matrix]

 sys_a2       (1)    1
 sys_b2       (1)    0.3          1
 sys_c2       (0.1)  0.4          0.3          1
 sys_d2       (0.9)  0.3          0.2          0.6          1
 estimate_a1  (10.6) 0.2026484018 0.2652968037 0.8479452055 0.105936073  1

[end correlation matrix]
```

Ignoring Correlations: Toy Example





- Generate pseudo measurements with different contributions of systematic uncertainties
- Compare to combined likelihood

- In particular uncertainties can be largely over- or underestimated.

Back to grouped uncertainties

measurement_1.txt

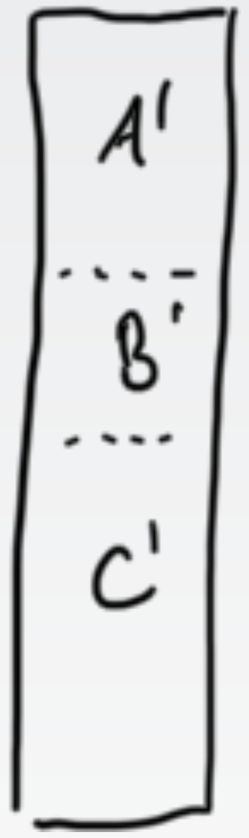
```
[not fitted]

                 syst_a1   syst_b1   stat
   estimate_a1   3.        2.        0.1

[end not fitted]
```

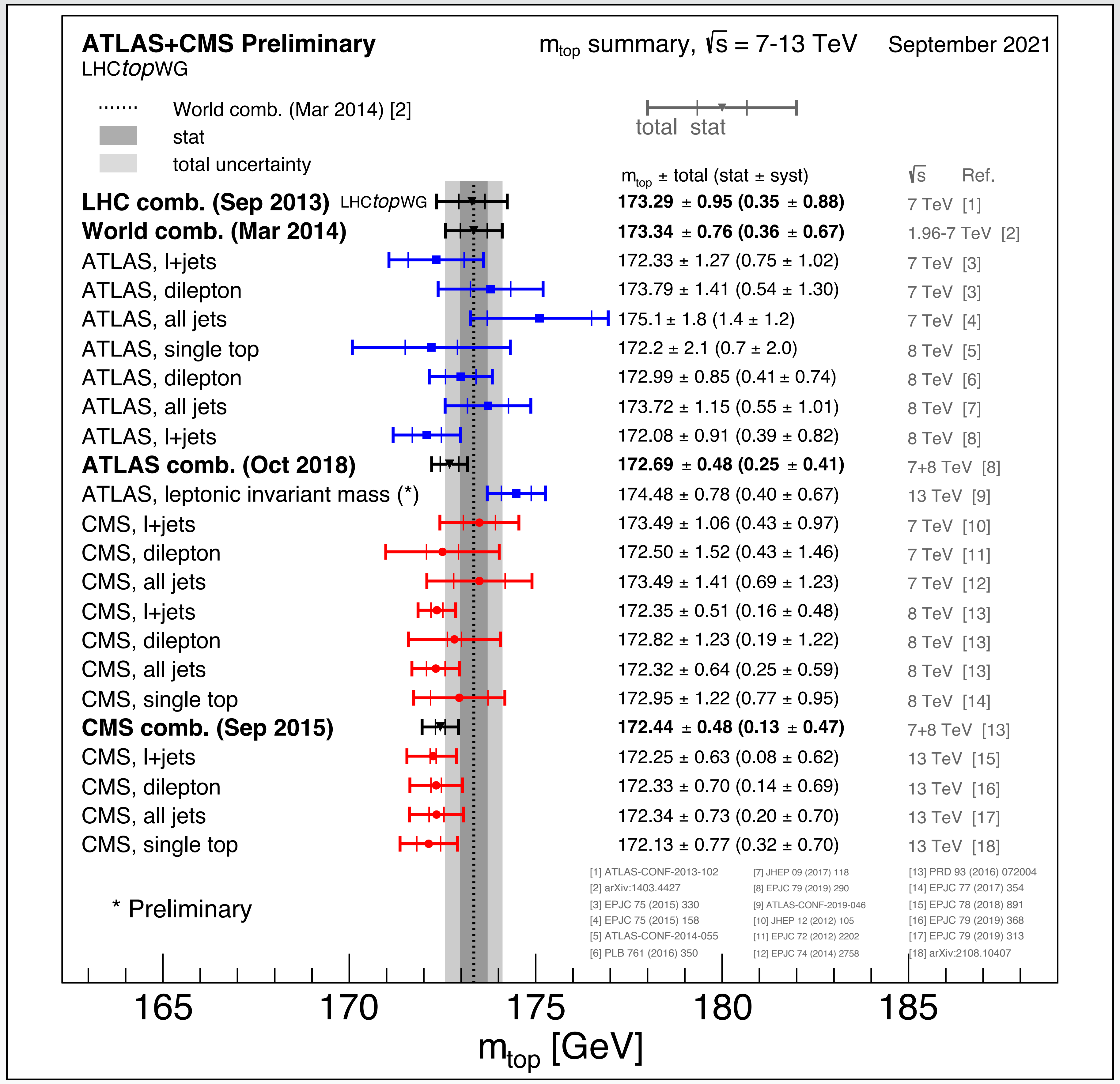
measurement_2.txt — Edited

```
[not fitted]

                 lumi_a2   syst_a2   stat
   estimate_a2   1         3.        0.1

[end not fitted]
```

- Despite it not being fully well defined, orthogonal uncertainties can be grouped by summing them up quadratically
- In practice, if the measurements to be combined have a similar split into sub contributions, this might be ok

```
[correlation matrix]

  sys_a2       (1)    1
  sys_b2       (1)    0.3          1
  sys_c2       (0.1)  0.4          0.3          1
  sys_d2       (0.9)  0.3          0.2          0.6          1
  estimate_a1  (10.6) 0.2026484018 0.2652968037 0.8479452055 0.105936073  1

[end correlation matrix]
```





- The concept breaks down entirely when there are correlations between the nuisance parameters of one individual measurement → any attempt results in information loss
- We should help avoid this grouping in the future

Summary

Common pitfalls are often related to (consistency of) correlation assumptions

It is instrumental to precisely distinguish between cause and effect

Well designed toy experiments are incredibly helpful for combinations as well as measurements and limits

For the combination of measurements approximate likelihood methods can be instructive and fully sufficient

All methods rely on having sufficient information accessible and being able to make use of it