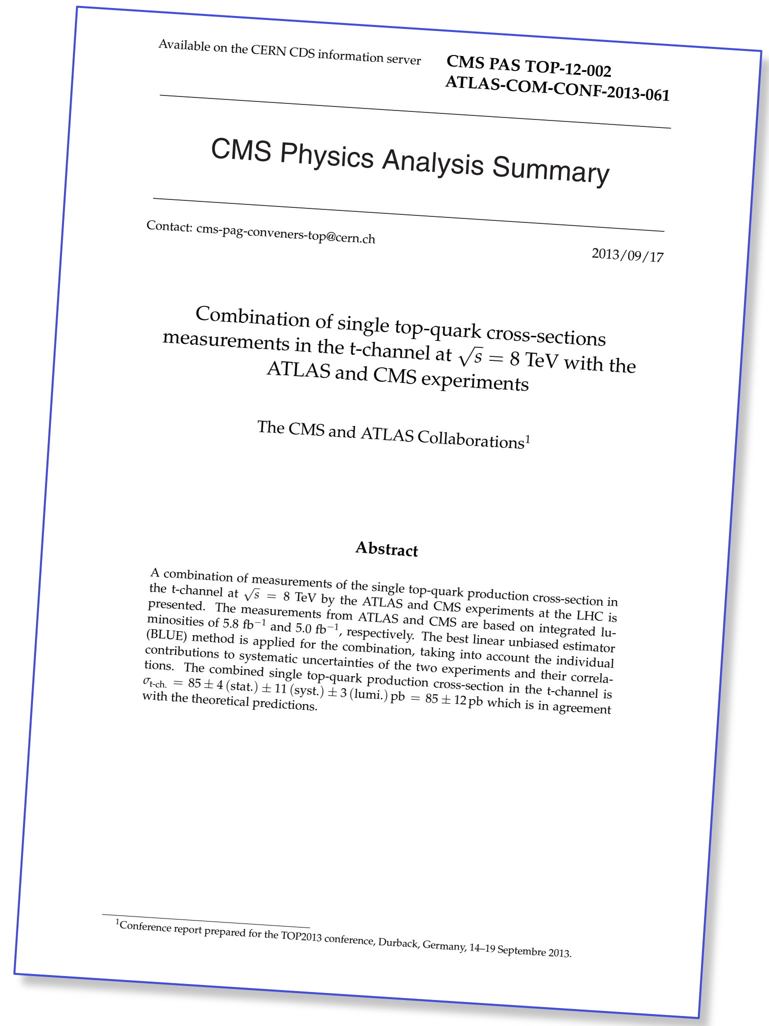
The Bias of the Unbiased Estimator

Julien Donini,
LPC/Univ. Blaise Pascal - Clermont-Ferrand
Luca Lista
INFN, Napoli

Introduction

- The ATLAS+CMS combination of single-top production cross-section measurements in the t channel was performed using the BLUE (Best Linear Unbiased Estimator)[*] method, as many other similar combinations
  - [*] L. Lyons et al. NIMA270 (1988) 110
- Compared to its original formulation, relevant contributions to the total uncertainty are known as relative uncertainties
  - Typically: systematic uncertainties
- This spoils the Gaussian assumption of the original formulation and may introduce a bias in BLUE estimate in case of sizable uncertainties (~10% or more)
- The iterative application of the BLUE method was proposed[**] in the case of a lifetime measurement to reduce such bias
  - [*] L. Lyons et al. PRD41(1990)982985
  - For each iteration, uncertainties known as relative contributions are rescaled to the combined value
- We performed tests to motivate this choice for the single-top combination quantifying the bias in the two cases of the 'plain' vs iterative BLUE implementations


Available on the CERN CDS information server

CMS PAS TOP-12-002
ATLAS-COM-CONF-2013-061

CMS Physics Analysis Summary

Contact: cms-pag-conveners-top@cern.ch

2013/09/17

Combination of single top-quark cross-sections measurements in the t-channel at $\sqrt{s} = 8$ TeV with the ATLAS and CMS experiments

The CMS and ATLAS Collaborations[1]

Abstract

A combination of measurements of the single top-quark production cross-section in the t-channel at $\sqrt{s} = 8$ TeV by the ATLAS and CMS experiments at the LHC is presented. The measurements from ATLAS and CMS are based on integrated luminosities of 5.8 fb^{-1} and 5.0 fb^{-1}, respectively. The best linear unbiased estimator (BLUE) method is applied for the combination, taking into account the individual contributions to systematic uncertainties of the two experiments and their correlations. The combined single top-quark production cross-section in the t-channel is $\sigma_{\text{t-ch.}} = 85 \pm 4\,(\text{stat.}) \pm 11\,(\text{syst.}) \pm 3\,(\text{lumi.})$ pb $= 85 \pm 12$ pb which is in agreement with the theoretical predictions.

[1]Conference report prepared for the TOP2013 conference, Durback, Germany, 14–19 Septembre 2013.

BLUE reminder

– Find linear combination of available measurements: $x = \Sigma\, w_i x_i$
  - No bias implies $\Sigma\, w_i = 1$
– Choose weights to minimize the variance of estimator
  - Take properly into account correlations between measurements!
– Equivalent to χ^2 minimization or maximum likelihood for Gaussian uncertainties

Simple example:

L.Lyons et al. NIM A270 (1988) 110

- Two measurements: $x_1 \pm \sigma_1$, $x_2 \pm \sigma_2$ with correlation ρ
- The weights that minimize the χ^2:

Cov. matrix

$$\chi^2 = \begin{pmatrix} x_1 - x & x_2 - x \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}^{-1} \begin{pmatrix} x_1 - x \\ x_2 - x \end{pmatrix}$$

are:

$$w_1 = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2} \qquad w_2 = \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2} \qquad (w_1 + w_2 = 1)$$

- The uncertainty of the combined value is:

$$\sigma_x = \sqrt{\frac{\sigma_1^2\sigma_2^2(1-\rho^2)}{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}}$$

Iterative BLUE

- Basic idea
  - Consider simple weighted average $\hat{\tau} = \frac{\sum \tau_i / \sigma_i^2}{\sum 1/\sigma_i^2}$
  - Errors (σ_i) are supposed to be true errors
  - But what are available are estimated errors ... that may vary with estimated central value (τ_i) $\hat{\tau} = \frac{\sum N_i / \tau_i}{\sum N_i / \tau_i^2}$
  - Violates 'combination principle': combination of partial combinations differs from combination of all results
- Solution proposed

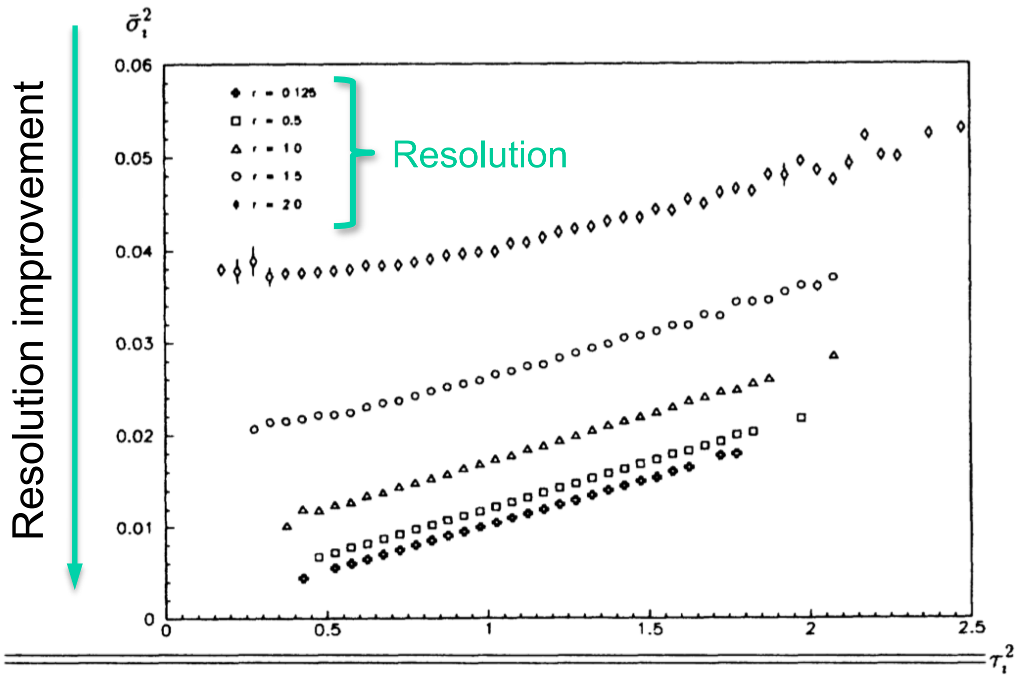
L. Lyons et al., Phys. Rev. D41 (1990) 982985

  - LMS: *Locally Matched Solution* (we prefer iterative BLUE)
  - Bias reduced if covariance matrix determined as if the central value is the one obtained from combination:
    - Rescale uncertainties to combined value e.g.: if we know a relative uncertainty:

$$\sigma^{\text{rescaled}} = \sigma \cdot x_{\text{blue}} / x_1$$

    - Iterate until central value converges to stable value

Application to B lifetime (Lyons)



r		$\langle \hat{\tau} \rangle$	$\langle \Delta \rangle$	$\langle \Delta^2 \rangle^{1/2}$
0.50	(I)	0.991±0.002	0.010	0.014
	(II)	1.000±0.002	0.001	0.001
1.00	(I)	0.989±0.003	0.010	0.014
	(II)	0.997±0.003	0.002	0.003
0.5−1.5	(I)	0.991±0.003	0.011	0.015
	(II)	1.000±0.003	0.002	0.003
2.00	(I)	0.990±0.002	0.011	0.015
	(II)	0.999±0.002	0.003	0.004

- Dependence of average estimated statistical uncertainties σ_i on lifetime estimate τ_i (squared) determined using toy experiments (100 decays each) and parameterized as:

Resolution

$$\sigma_i^2(\hat{\tau}) = \frac{k_i \hat{\tau}^2}{N_i} + \frac{r_i^2}{N_i}$$

Comparison of LMS method (II) with least square procedure (I).

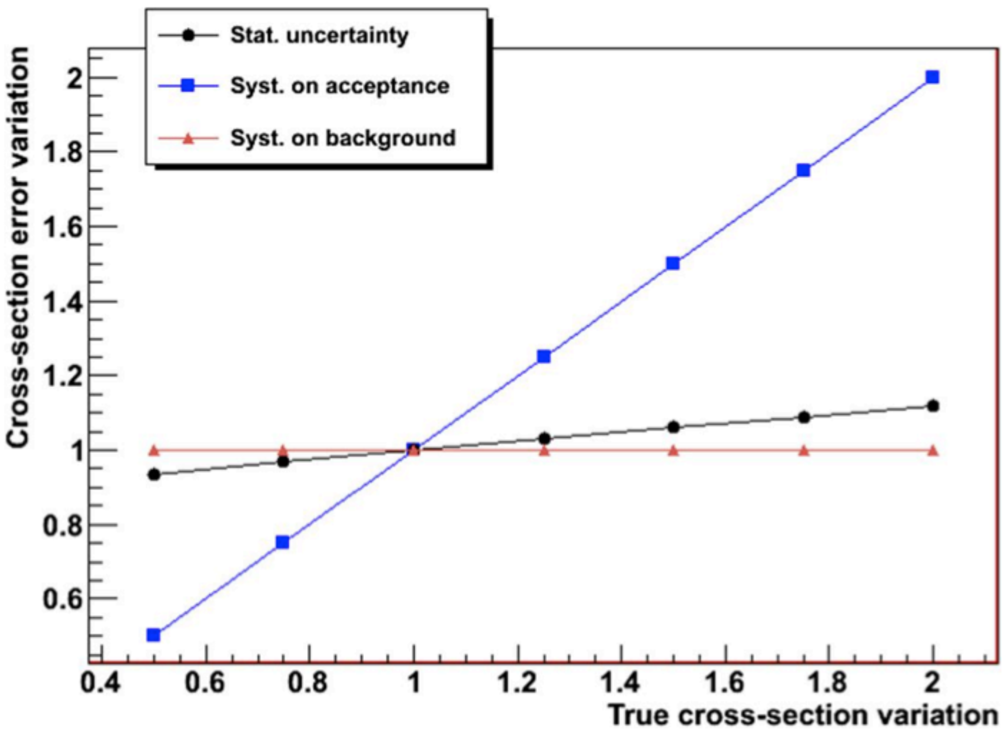
$\Delta = \tau_{\mathrm{fit}} - \tau$ with τ_{fit} the result from one single combined experiment ('combination principle')

Application to single-top comb.

- Cross section measurements given by:

$$\sigma = \frac{N - B}{\mathcal{L} \cdot A}$$

sel. events → N; bkg. estimate → B; Int. lumi. → $\mathcal{L}$; acceptance → A



σ varied from: 40 → 160 pb

B = 3000,

A = 0.25%

L= 5 fb^{-1}

δA =5%, δB =5%

ATLAS+CMS combination

- ATLAS (ATLAS-CONF-2013-098):

  $$\sigma_{\text{t-ch.}} = 95.1 \pm 2.4\,(\text{stat.}) \pm 18.0\,(\text{syst.})\,\text{pb} = 95.1 \pm 18.1\,\text{pb}$$

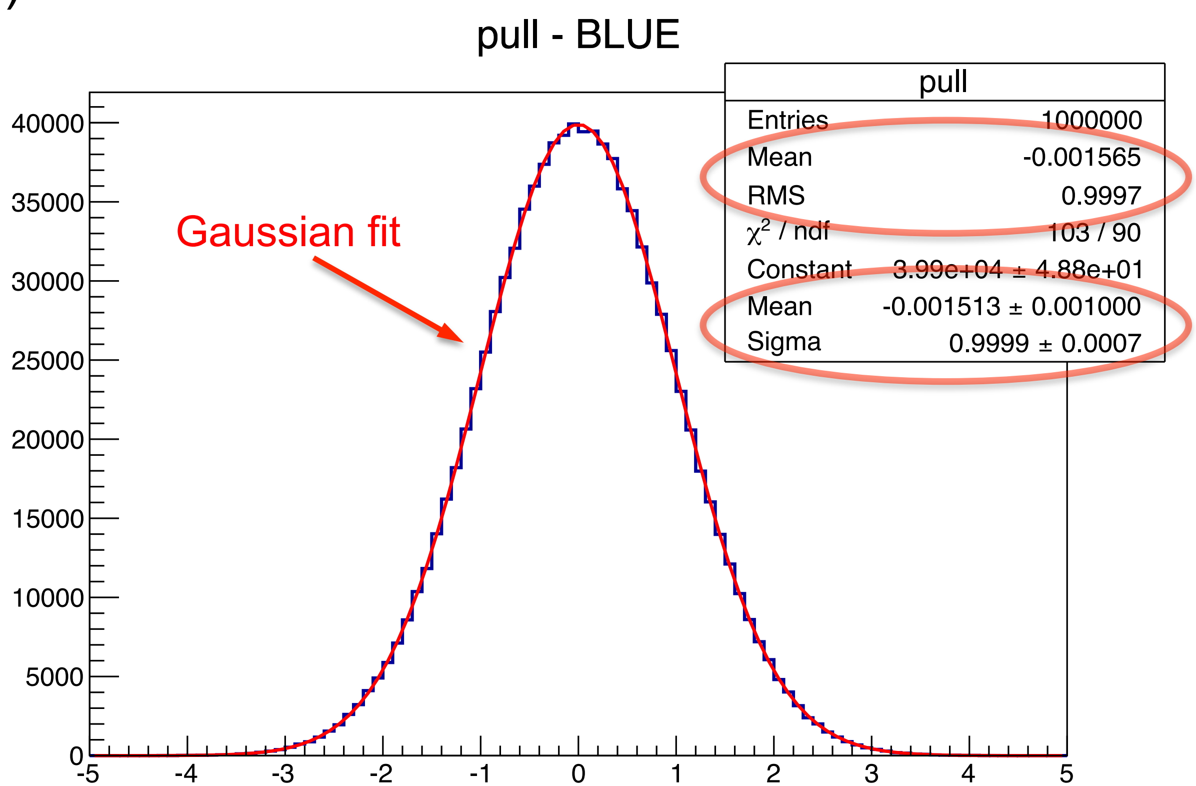
  – stat: 2.5% (fixed), syst.: 20% (relative uncert.)
- CMS (CMS-PAS-TOP-12-002):

  $$\sigma_{\text{t-ch.}} = 80.1 \pm 5.7\,(\text{stat.}) \pm 11.0\,(\text{syst.}) \pm 4.0\,(\text{lumi.})\,\text{pb} = 80.1 \pm 13.0\,\text{pb}$$

  – stat: 7.1% (fixed), syst.: 15% (relative uncert.)
- Correlation: $\rho = 0.38$ (total uncert.), $\rho = 0.42$ (syst. only)
- Plain BLUE combination:
  – $\sigma_{\text{t-ch.}} = 83.6 \pm 12.1$ pb
- Iterative BLUE combination:
  – $\sigma_{\text{t-ch.}} = 85.3 \pm 12.2$ pb

Plain BLUE method

- Assuming as true uncertainties ATLAS and CMS estimates, with their correlation
- Generate toys according including correlation and apply BLUE (plain)



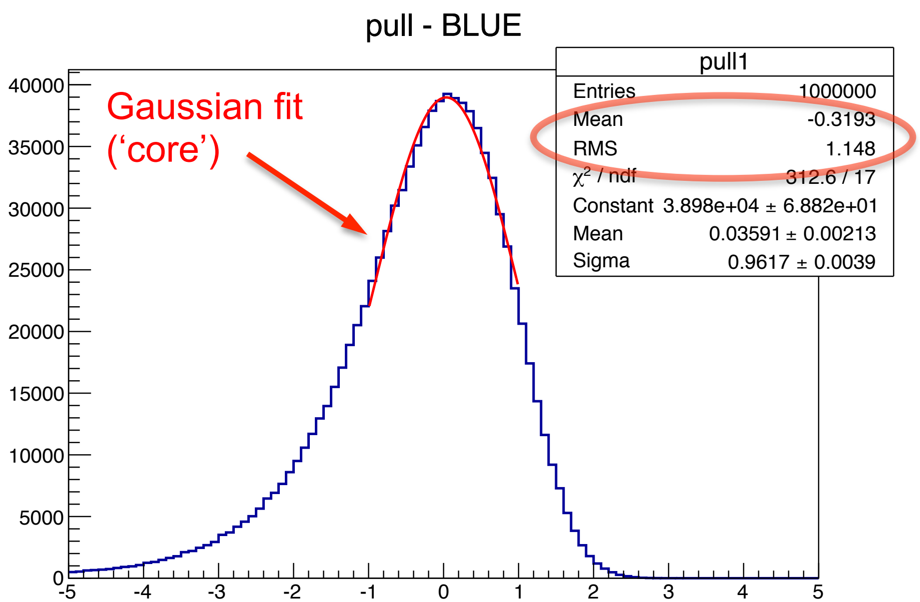
Test with Toy Monte Carlo

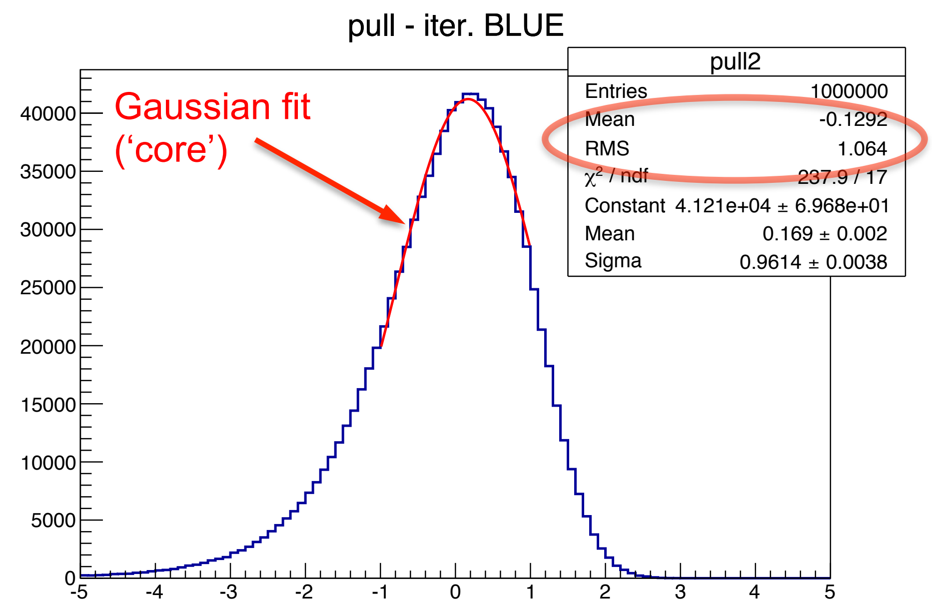
- To assess the possible bias using the two estimators, a simple toy generator was used:

- Assume as true value the combined one: 85 pb
- Assume as true fluctuations the nominal uncertainties of the two experiments with their correlation
  - Nominal uncertainty: stat + syst added in quadrature of each experiment
- Two central values are extracted for ATLAS and CMS according to the know uncertainties, their correlation and the true central value
  - Fixed stat errors, rel. syst errors
- Relative (systematic) uncertainties are rescaled to the extracted central values for ATLAS and CMS

Iterative vs plain BLUE

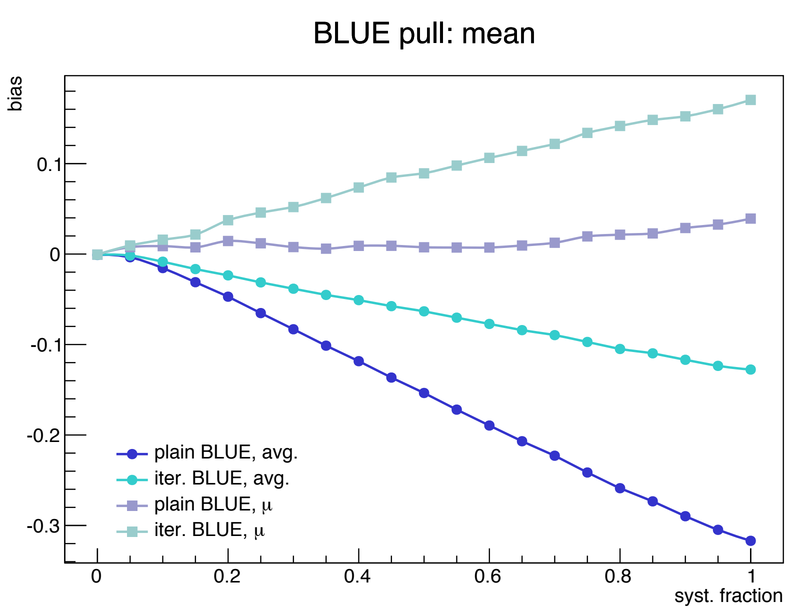
- Bias: −0.32σ ~ 3.9pb (plain) → −0.13σ ~ 1.6pb (iterative)
- RMS: 1.15 (plain) → 1.06 (iterative): uncertainty underestimate
- Gaussian μ: 0.04 (plain) → 0.169 (iterative)
- The iterative BLUE method reduces the bias by shifting the 'core' of the distribution and by shrinking the non-Gaussian tails

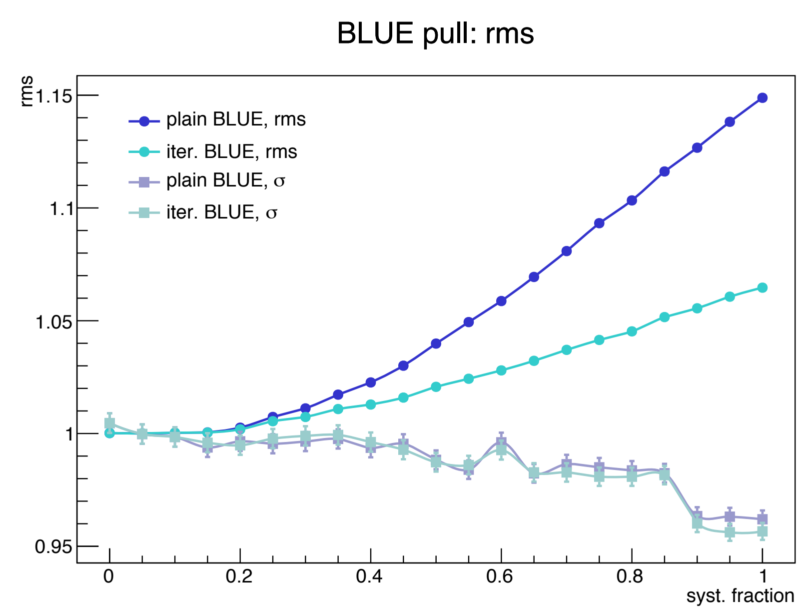




Dependence on syst. fraction

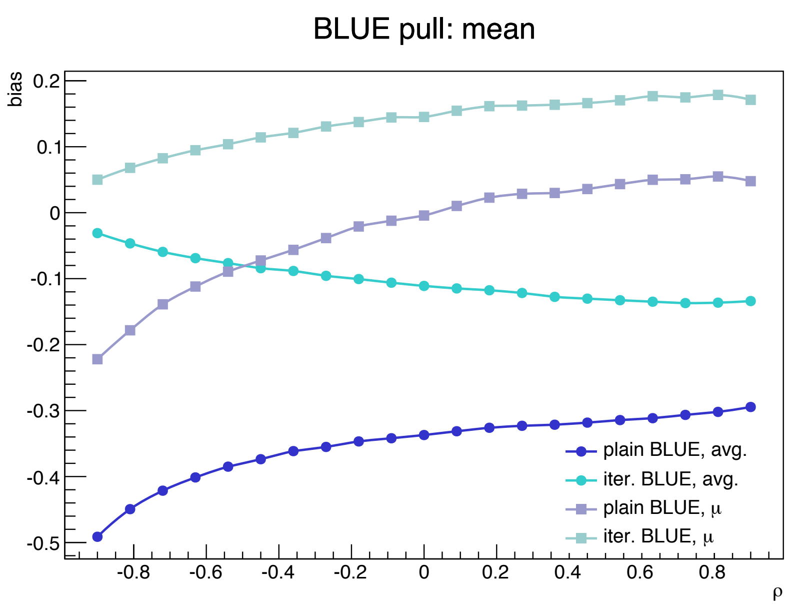
- The systematic contribution was varied from the nominal values to zero rescaling it by a fraction *f* going from 0 to 1
- Bias was measured varying the fraction *f* of the actual uncertainty
- As expected, the bias goes to zero as *f* goes to zero
  - Iterative BLUE is not an issue for precision measurements (e.g.: m_t)
- The iterative BLUE has in general smaller RMS and less bias

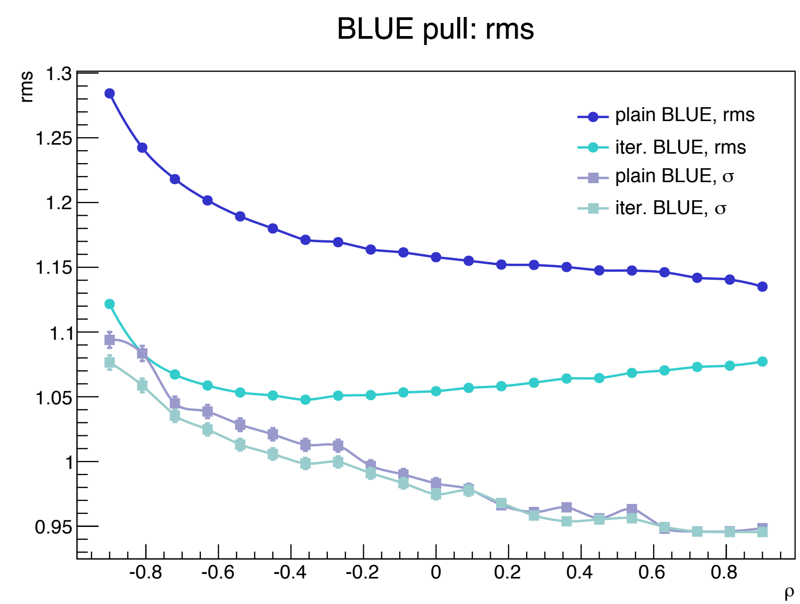




Dependence on correlation

- Correlation between systematic uncertainties was varied from ρ=0 to ρ=1
- Bias was measured varying the correlation ρ
- The iterative BLUE has in general smaller RMS and less bias





What we have learned

- Iterative BLUE method in general reduce biases
  - Better behaved than standard BLUE method, but not perfect either
  - In case of single top combination: difference is minor
- Not all uncertainties should be rescaled
  - Data-driven bkg. uncertainties
  - Stat uncertainty (at 1st order)
  - Others ?
- More investigations ongoing…