Usage of weakly correlated observables for nuisance parameter fits

Lars Stietz[1], Patrick Connor[2,3], Johannes Lange[2],
Peter Schleper[2], Hartmut Stadie[2]
[1] Hamburg University of Technology, Institute of Mathematics, Chair of Computational Mathematics
[2] University of Hamburg, Institute of Experimental Physics
[3] University of Hamburg, Center for Data and Computing in Natural Science

Contact Information:
lars.stietz@tuhh.de

TUHH Hamburg University of Technology

Universität Hamburg — DER FORSCHUNG | DER LEHRE | DER BILDUNG

DASHH — Data Science in Hamburg HELMHOLTZ Graduate School for the Structure of Matter

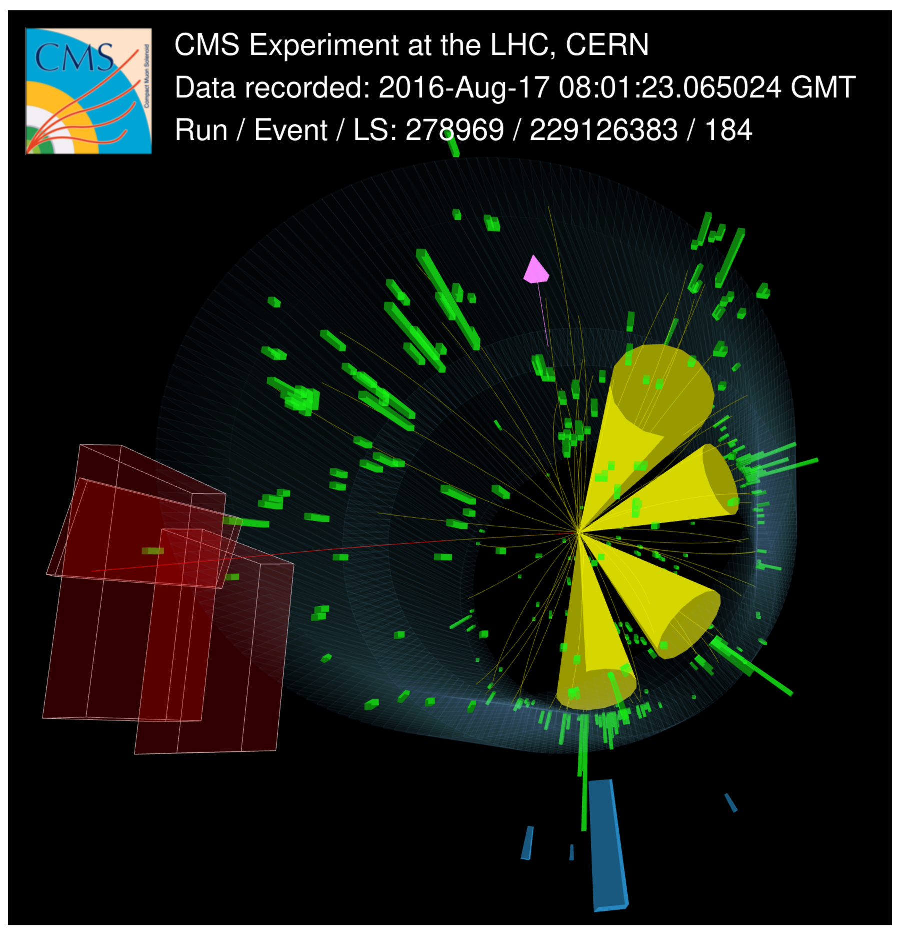


Figure 1: Proton-Proton Collision Event Display

Overview

- Current top mass is measured at 172.52 ± 0.33 GeV
- Multiple observables are used for the measurement
- Profile likelihood fits including nuisance parameters could constrain systematic uncertainties on the measurement
- Independency of used observables is assumed
- Including more observables is expected to improve the accuracy and precision of the measurement
- Assumption of independent observables could be violated

Single Event Likelihood

- Individual events will be stochastically independent
- Measured observables in an event are potentially correlated
- Increase accuracy and precision through joint probability density of multiple observables per event

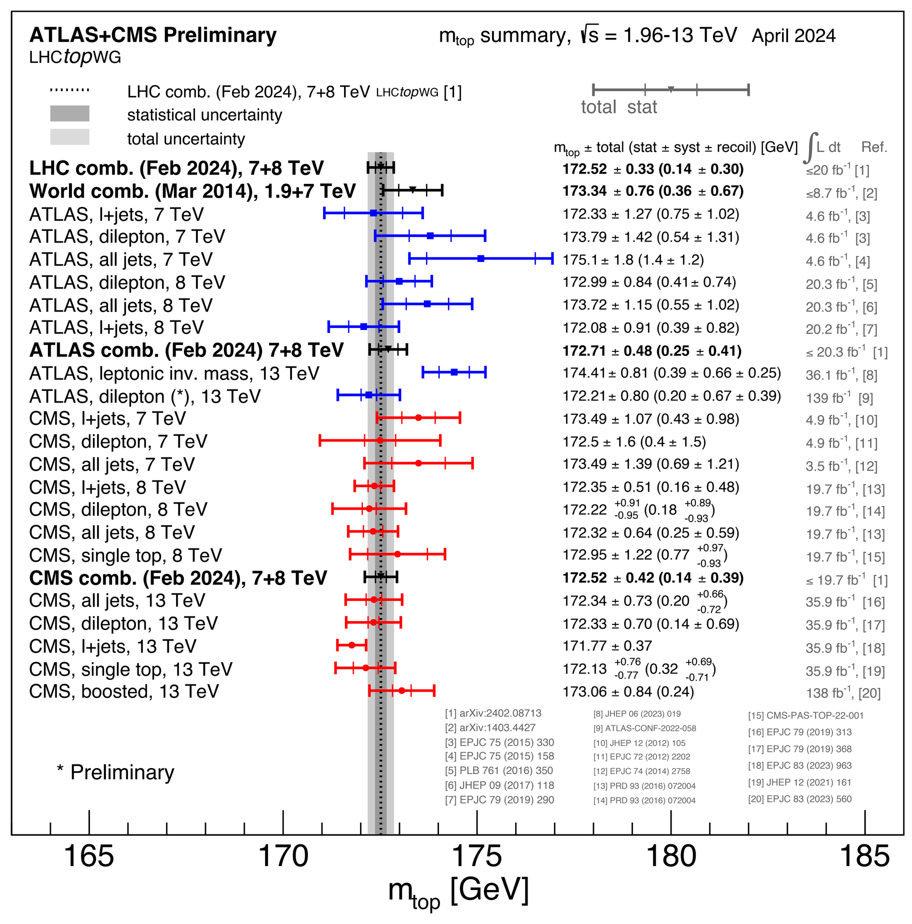


Figure 2: Top mass measurement results

Toy Model Setup (Inspired by CMS-TOP-20-008)

Goal: Reconstruct the top mass from "measuring" two observables

Observables: $m_t \sim P(\mu_t(\theta), \sigma_t(\theta))$ and $m_W \sim Q(\mu_W(\theta), \sigma_W(\theta))$

Parameters of interest: $\theta = (\Delta m_t, \Delta\text{JSF}, \Delta\text{JRF})$

Reconstruction: Minimize the negative log-likelihood with respect to Δm_t given data samples from P and Q

Use nuisance parameters $\Delta\text{JSF}, \ \Delta\text{JRF}$ to include systematic uncertainty

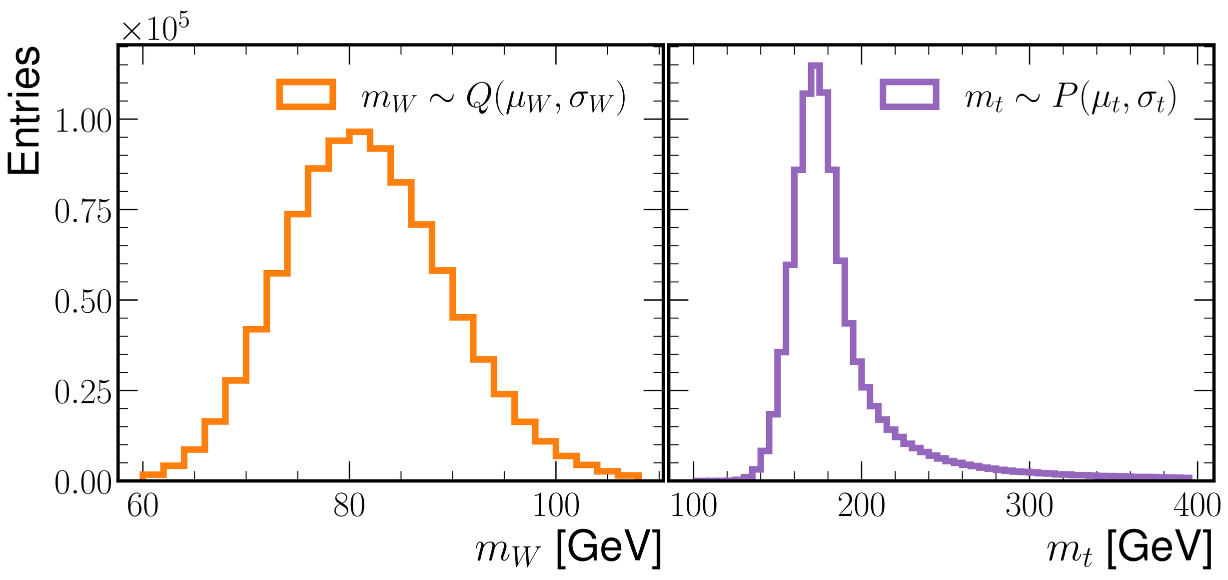


Figure 3: Example of sampled distributions for m_W and m_t

Likelihood formulation

$$-\log \mathcal{L}(\theta \mid m_t, m_W) = -\sum_i \log \mathcal{L}_P(\mu_t(\theta), \sigma_t(\theta) \mid m_t^i) - \sum_i \log \mathcal{L}_Q(\mu_W(\theta), \sigma_W(\theta) \mid m_W^i) - \mathcal{L}_{\text{JSF}}(\Delta\text{JSF}) - \mathcal{L}_{\text{JRF}}(\Delta\text{JRF})$$

Independency of m_t and m_W needed

Correlated Sampling

1. Sample θ_i independently and compute $\mu_t, \sigma_t, \mu_W, \sigma_W$
2. Sample sets of $m_t \sim P(\mu_t, \sigma_t)$ and $m_W' \sim Q(\mu_W, \sigma_W)$ independently
3. Normalize m_t, m_W', i.e., $\hat{m}_\star = (m_\star - \bar{\mu}_\star)/(\bar{\sigma}_\star)$
4. Given correlation coefficient $\rho \in [0,1]$ calculate
$$\tilde{m}_W = \rho \hat{m}_t + \sqrt{1-\rho^2}\hat{m}_W'$$
5. Renormalize to get $m_W = \tilde{m}_W \sigma_W + \mu_W$

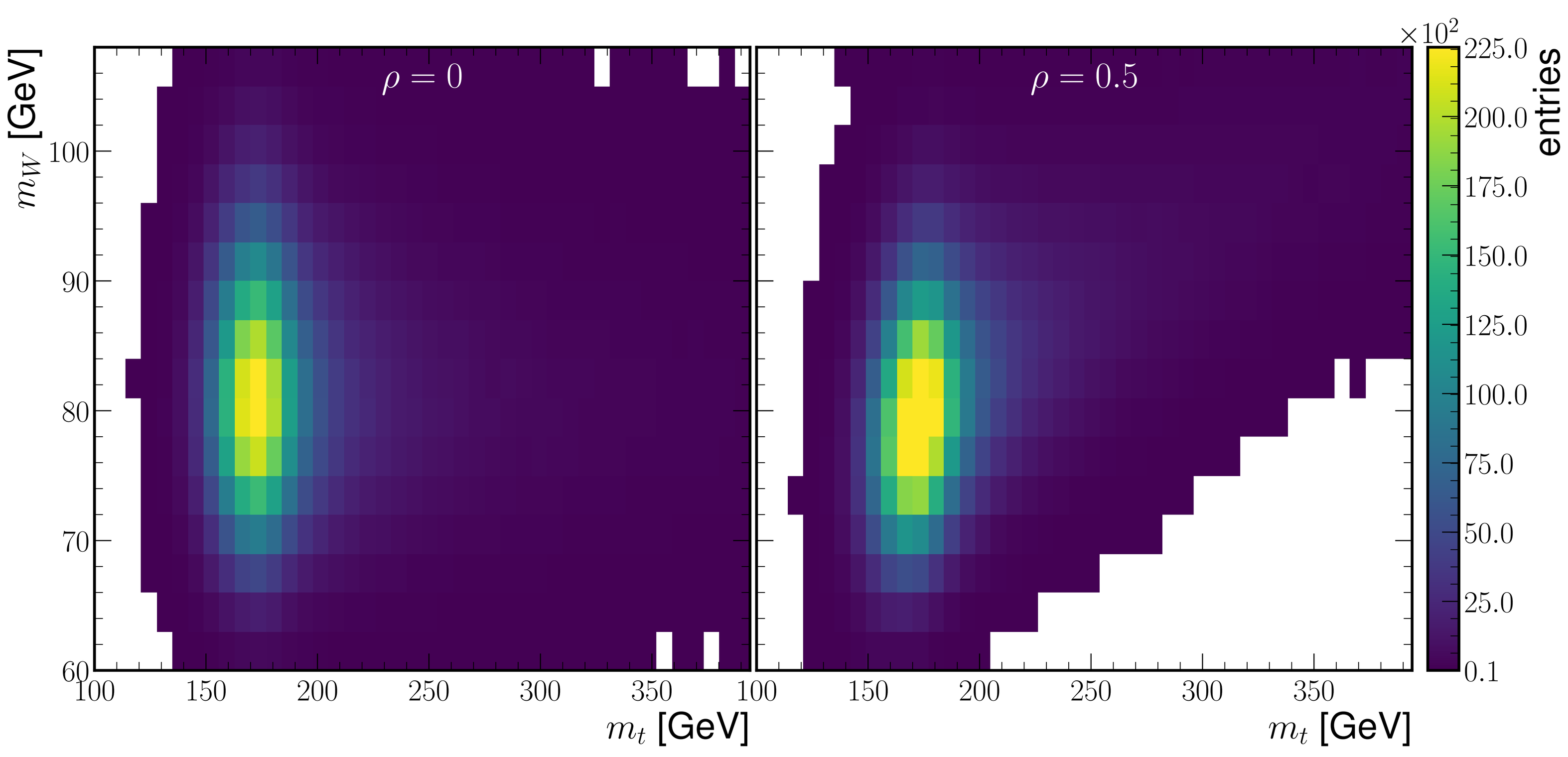


Figure 4: 2D-Histogram from the joint distribution of m_W, m_t for correlations 0 and 0.5

Toy Experiment Result

1) Calibration

Before fitting the parameters of interest one has to calibrate the distributions used, i.e., find $\mu_t^0, s_{\star,t}^\mu$ from simulations with known parameters of interest $(\Delta m_t, \Delta\text{JSF}, \Delta JRF \in \{-1, 0, +1\}$:

$$\mu_t'(\theta) = \mu_t^0 \cdot (1 + s_{t,t}^\mu \Delta m_t) \cdot (1 + s_{\text{JSF},t}^\mu \Delta\text{JSF}) \cdot (1 + s_{\text{JRF},t}^\mu \Delta\text{JRF})$$

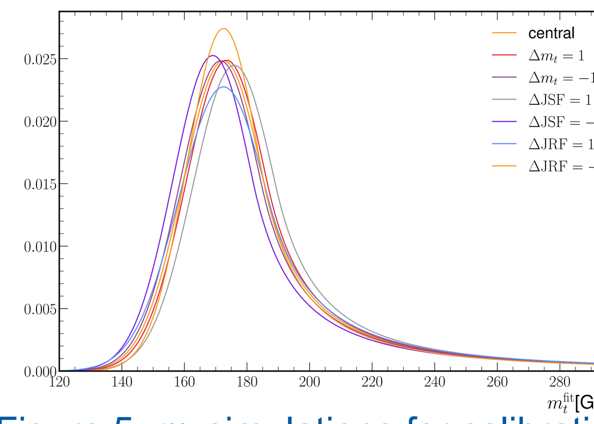


Figure 5: m_t simulations for calibration

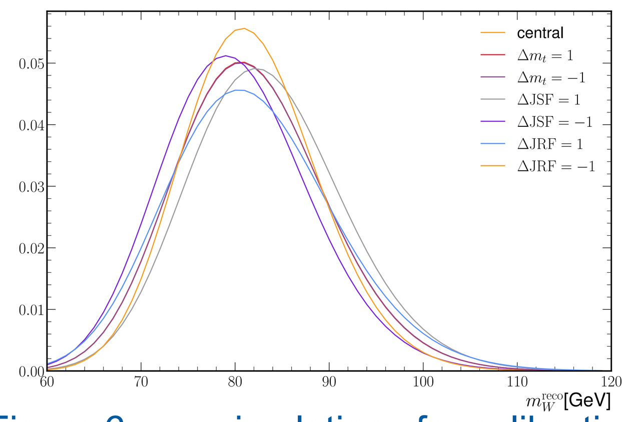


Figure 6: m_W simulations for calibration

2) Update calibrated parameters in likelihood function

3) Minimize negative log-likelihood wrt θ

Alternative Approach: Binned likelihood fit with pyhf

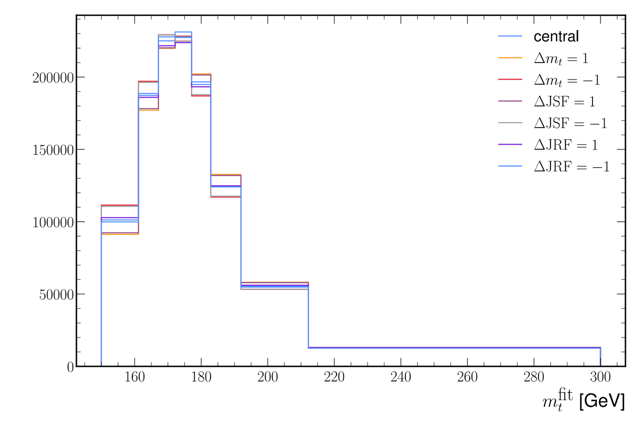


Figure 7: m_t simulations for calibration (binned)

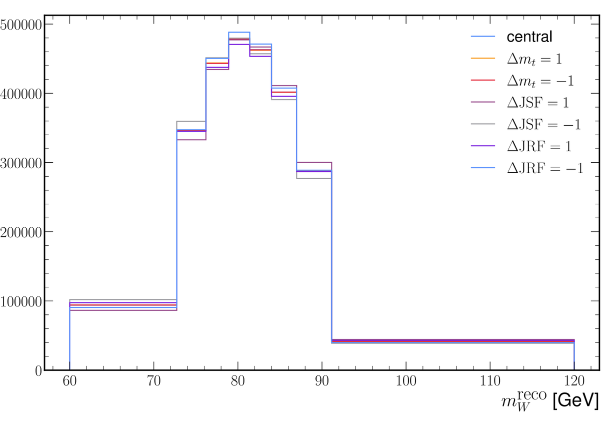
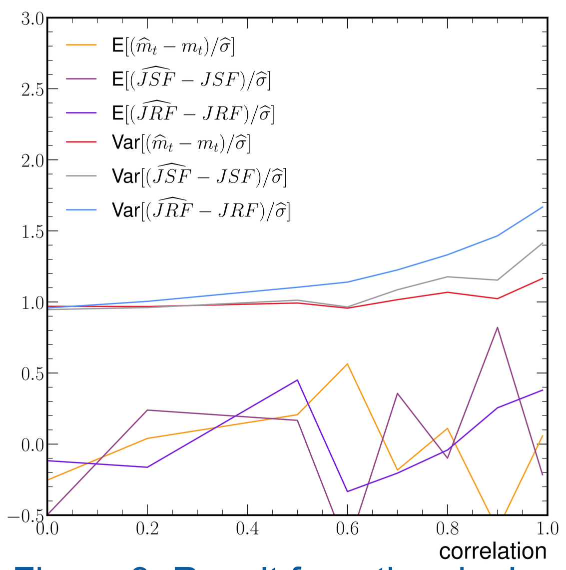


Figure 8: m_W simulations for calibration (binned)

Evaluation of the experiment



- 10^4 pseudoexperiments
- Evaluate pull:
$$\text{Pull}(\theta_i) = \frac{\hat{\theta}_i - \theta_i^{\text{true}}}{\hat{\sigma}_i}$$
- Variance of pull describes accuracy
- Mean of pull describes bias

Figure 9: Result from the single event likelihood fit

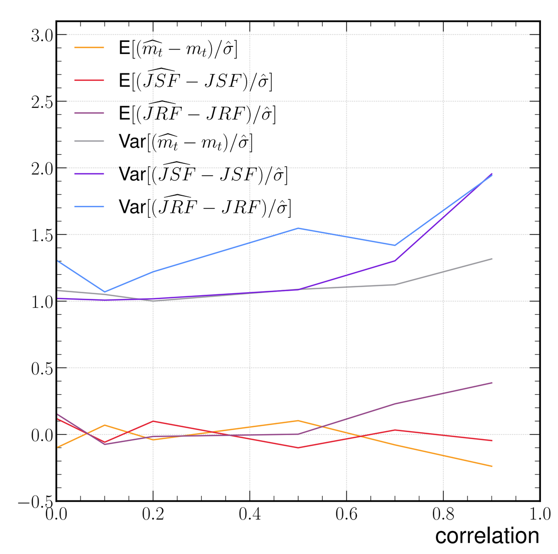


Figure 10: Result from the binned likelihood fit

Conclusion

- With increasing correlation pull width increases from 1
- Most significant impact are in parameters affected by both distributions
- Measurement needs careful selection of observables or multi-dimensional joint probability density functions

Outlook to Normalizing Flows (NF)

- Normalizing flows can be used to model complicated distributions
- Giving a transformation of variables from a complicated distribution to a simple distribution (i.e. Gaussian) NF can be used to calculate exact likelihoods
- Use NF to learn the combined distribution of $(m_t, m_W) \sim D$

Travel sponsored by PHYSTAT in the context of "PHYSTAT for the Future"

HELMHOLTZ
Grant No. HIDSS-0002 (Helmholtz Association)