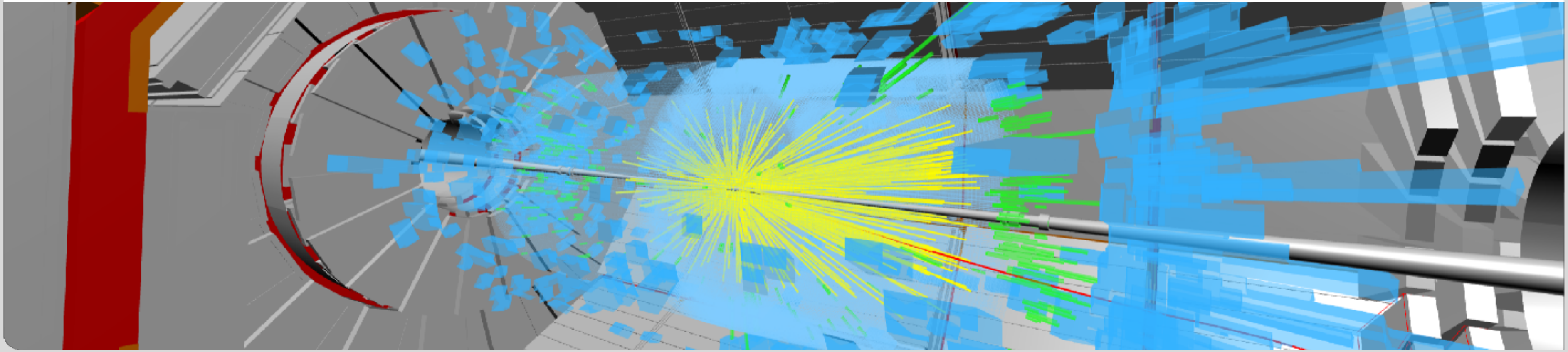
Tools and Techniques for High-p_T Physics

7th ENHEP School on High Energy Physics
Ain Shams University, Cairo, January 26–31, 2019

Ulrich Husemann, Institute of Experimental Particle Physics, Karlsruhe Institute of Technology



Outline

Karlsruhe Institute of Technology

- From Raw Data to Physics Results
- Monte-Carlo Event Generation
- Physics Objects
- Background Estimation
- Advanced Signal Analysis

Ulrich Husemann
Institute of Experimental Particle Physics

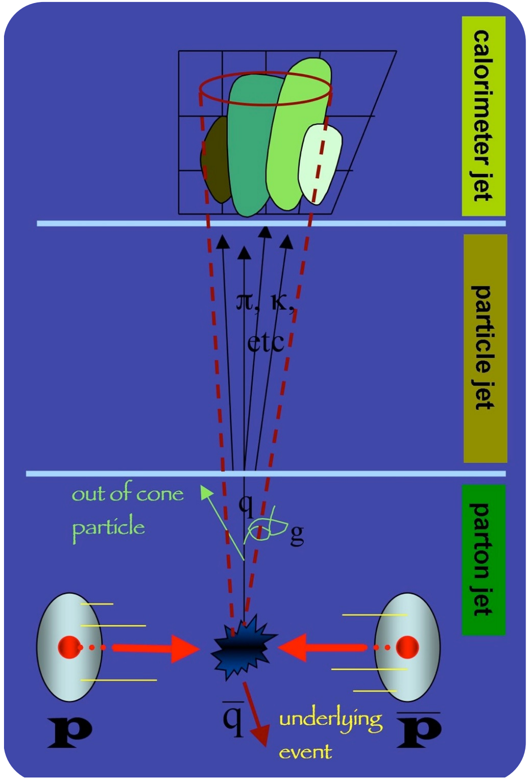
PHYSICS OBJECTS

Ulrich Husemann
Institute of Experimental Particle Physics

Short Recap

- Physics objects:
  - Many classes: charged leptons, jets, b-jets, boosted objects …
  - Lepton ID: **multivariate** discriminant, isolation
- All objects must be properly calibrated:
  - Tracks: **alignment** of tracking detectors (not shown)
  - Leptons: **scale factors**, e.g. via tag&probe method
  - Jets: today…

What is a Jet?



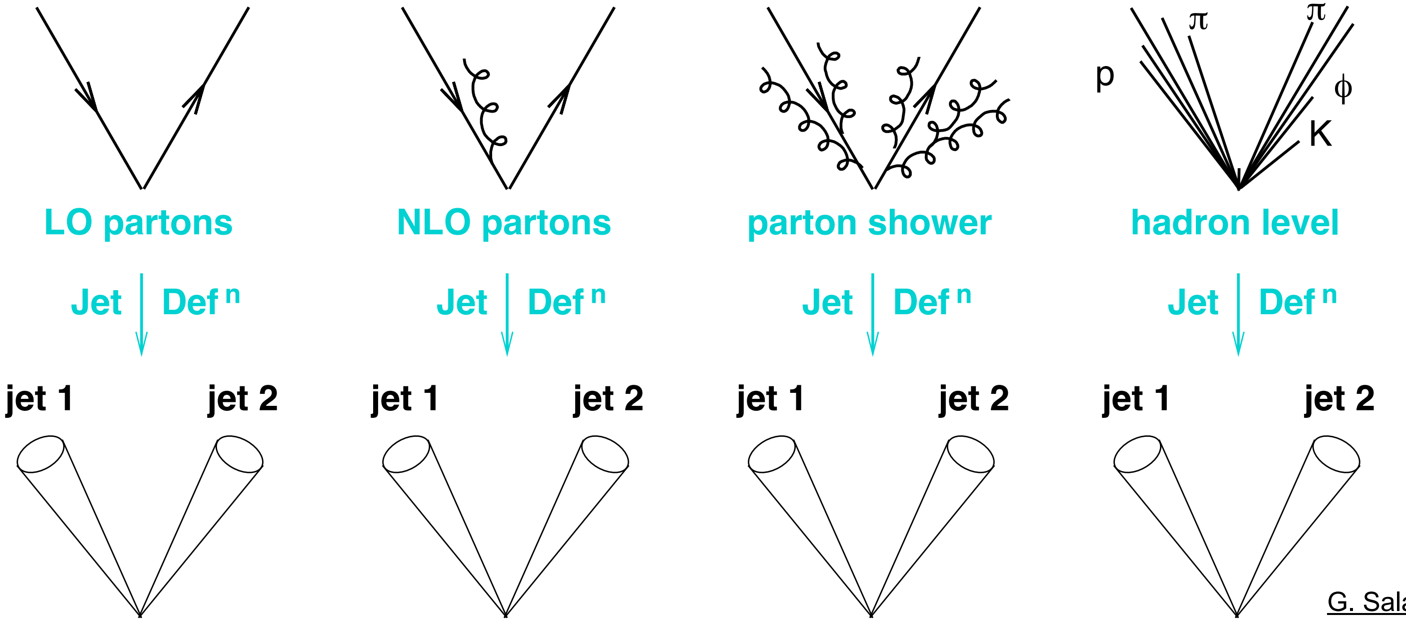
www-cdf.fnal.gov

- Jets can be defined on different **technical** levels:
  - **Parton level**: for calculations in perturbative QCD ("theory jets")
  - **Particle level**: jets reconstructed from stable hadrons
  - **Detector level**: jets reconstructed from energy deposits in calorimeter and/or tracks in tracking detectors
- Design of successful **jet algorithms**:
  - **Independent** of technical level
  - **Invariant** under Lorentz boosts
  - Comparison with theory: **infrared and collinear safe** → find same jet even after emitting soft/collinear radiation

Ulrich Husemann
Institute of Experimental Particle Physics

Infrared and Collinear Safety

- Goal: jet definition on **robust on all technical levels** against additional radiation (low momentum or small angle)



G. Salam

Ulrich Husemann
Institute of Experimental Particle Physics

Sequential Recombination

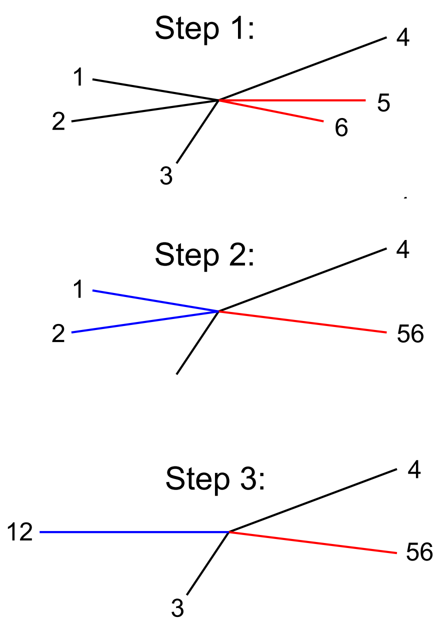
- Jets at the LHC: **sequential recombination**
  - **Infrared/collinear safe** by construction
  - Define **distance measure d_{ij}** between particles i, j and distance of particle i to beam axis d_{iB}
  - LHC standard: **"anti-k_t algorithm"**

$$d_{ij} = \min(k_{t,i}^{-2}, k_{t,j}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \qquad d_{iB} = k_{t,i}^{-2}$$

LHC Run 2:
$R = 0.4$ ("ak4")

- Sequential recombination algorithm:
  - Compute **d_{ij} for all pairs** of (pseudo-)particles, **combine pairs** with $d_{ij} < d_{iB}$ to new pseudo-particles
  - Termination condition: pseudo-particle → **jet** if d_{iB} is the smallest distance d_{ij}

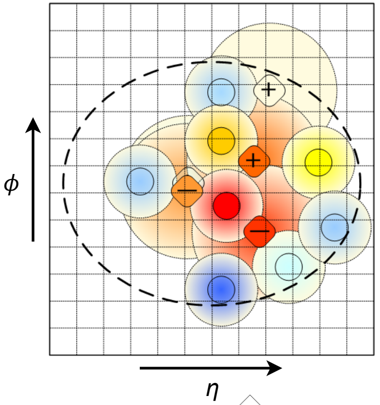
Sequential recombination



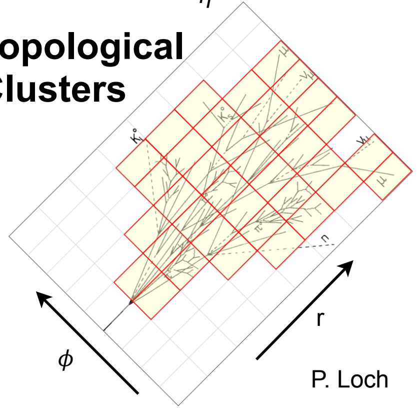
Ulrich Husemann
Institute of Experimental Particle Physics

Jet Reconstruction

Calorimeter Towers



Topological Clusters

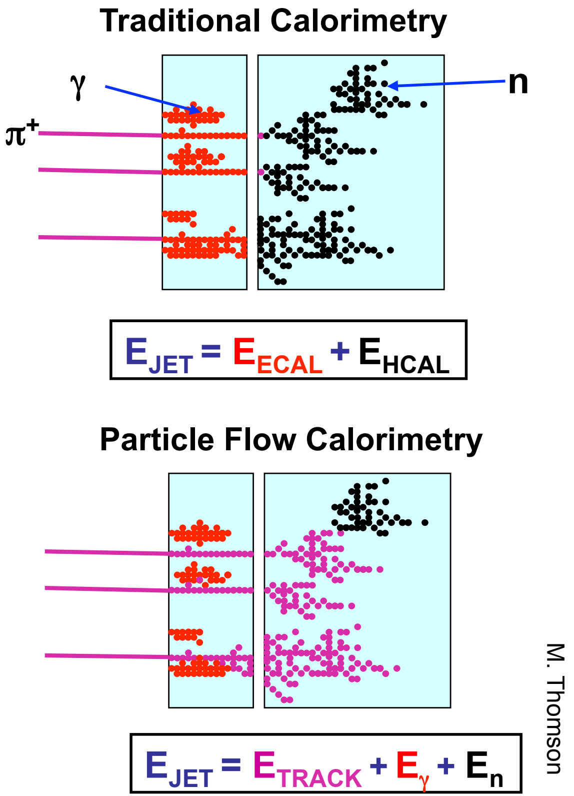


P. Loch

- Starting points for energy reconstruction in calorimeter:
  - **Calorimeter towers**: fixed grid grouping calorimeter cells, e.g. $\Delta\eta\times\Delta\phi = 0.1\times0.1$
  - **Topological clusters** ("topo clusters"): groups of cells with energy deposits
- Jet reconstruction strategies:
  - Pure **calorimeter jets** or **track jets**
  - **Combination** of calorimeter and tracker information
  - **Particle flow**: optimal combination of subdetectors for reconstruction of each particle type

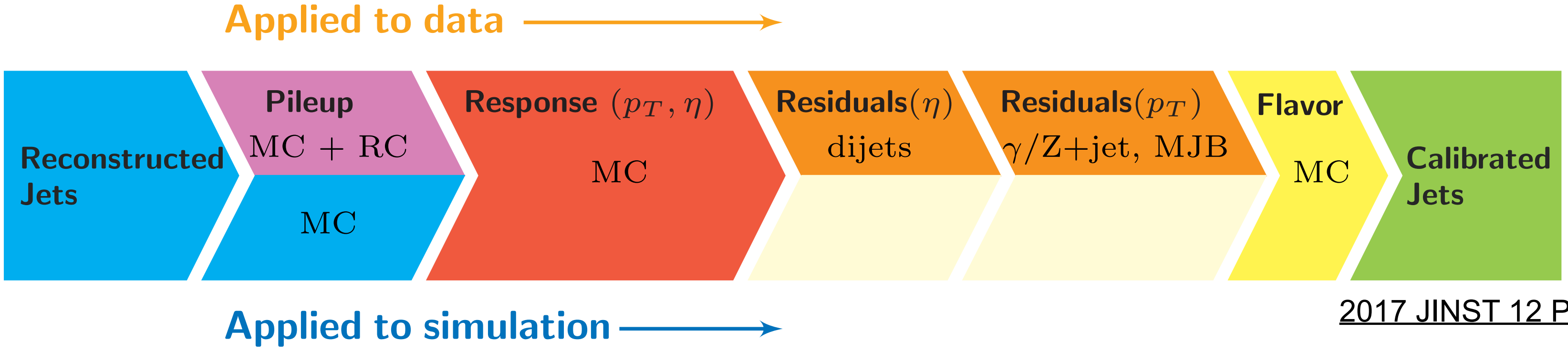
Particle Flow Calorimetry

- Jet reconstruction with **traditional** calorimetry:
  - Reconstruct jet energies from ECAL and HCAL energy deposits
  - Approx. 70% from HCAL with rather poor energy resolution ($\gtrsim 50\%/\sqrt{E}$)
- **Particle flow** calorimetry: jet reconstruction exploiting **each sub-detector optimally**
  - Tracker: charged particle momentum
  - ECAL: photon energy and bremsstrahlung
  - HCAL: only neutral hadrons (n, K_L) → only 10% of jet energy reconstructed with rather bad resolution



Jet Calibration: CMS Method

- Multi-stage **jet energy corrections** in CMS:

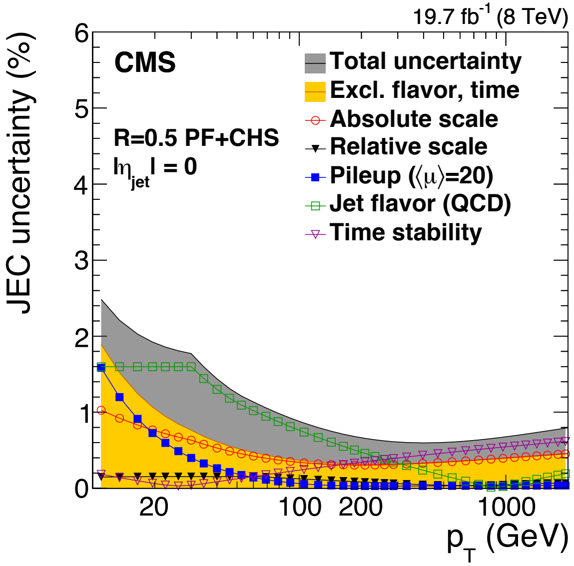


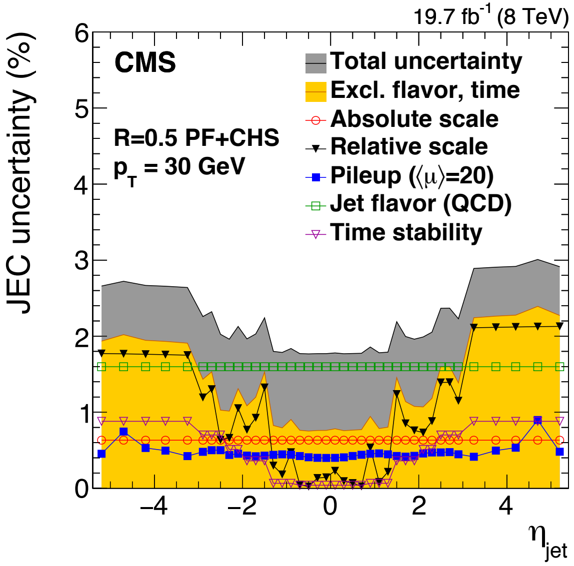
2017 JINST 12 P02014

- **Pileup** and electronic **noise** correction (MC-based + additional “residual correction” for data)
- **Response** correction: uniform response as a function of p_T and η (MC)
- **Residual** corrections for data in η (dijet events with one well-calibrated jet in the barrel region)
- **Residual** corrections for data in p_T (balance of Z or γ recoiling against jet)
- (optional) **Jet flavor** correction: different response from light quarks, gluons, heavy quarks (MC)

Ulrich Husemann
Institute of Experimental Particle Physics

Jet Calibration: Results





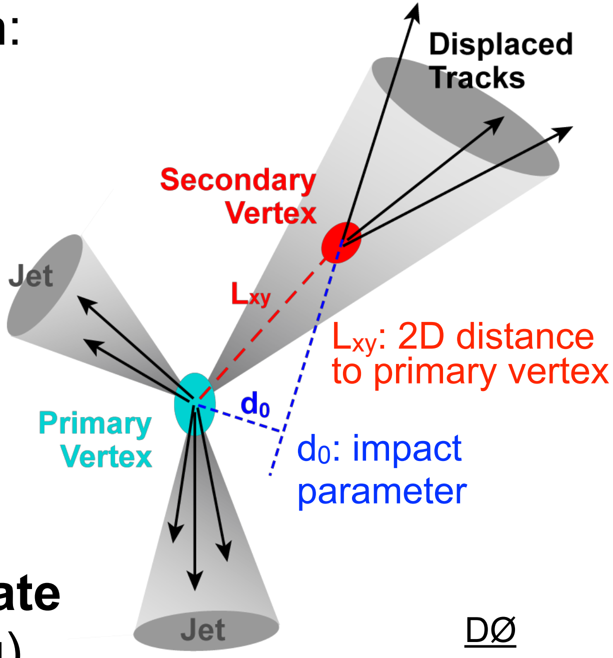
2017 JINST 12 P02014

- Typical **uncertainties** of jet energy corrections: **1–2%**
- Uncertainties propagated into more complex observables, e.g. E_T^{miss}
- Jet energy **resolution** in data worse than in MC → “smear” MC

Ulrich Husemann
Institute of Experimental Particle Physics

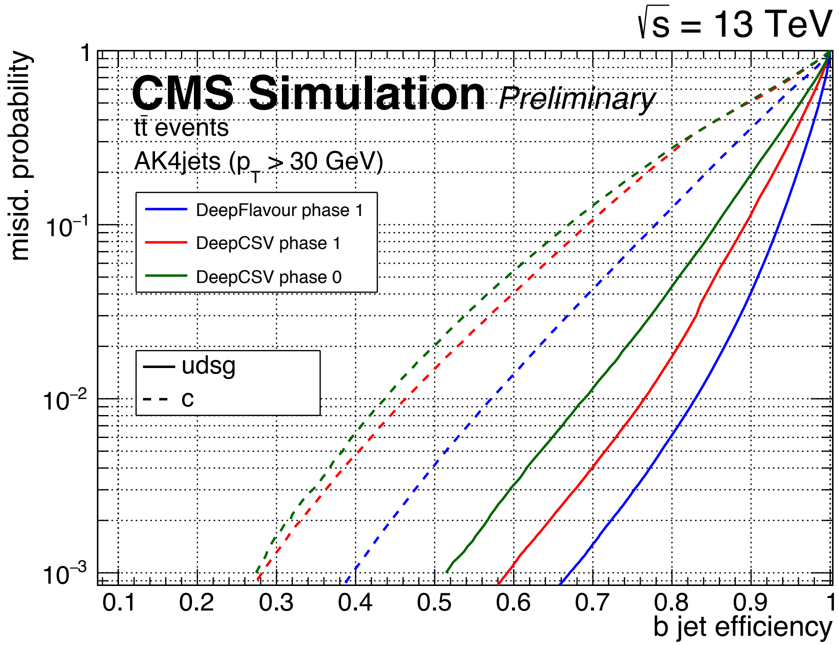
B-Tagging

- B-tagging algorithms at hadron colliders based on:
  - **Secondary vertices** and tracks with **large impact parameter**: long B hadron lifetime (picoseconds)
  - **Soft leptons**: semileptonic decays $B \to \ell \nu X$
  - **Large b-quark mass:** wider jets, large relative p_T of lepton in $B \to \ell \nu X$
  - **Hard b-quark fragmentation**: B hadron carries most of b-quark energy
- LHC Run 2: above criteria combined in **multivariate** discriminant (increasingly based on deep learning)
- Relevant e.g. for $t \to Wb$, $H \to b\bar{b}$, $Z \to b\bar{b}$



B-Tagging Performance

CMS Expected B-Tagging Performance 2017

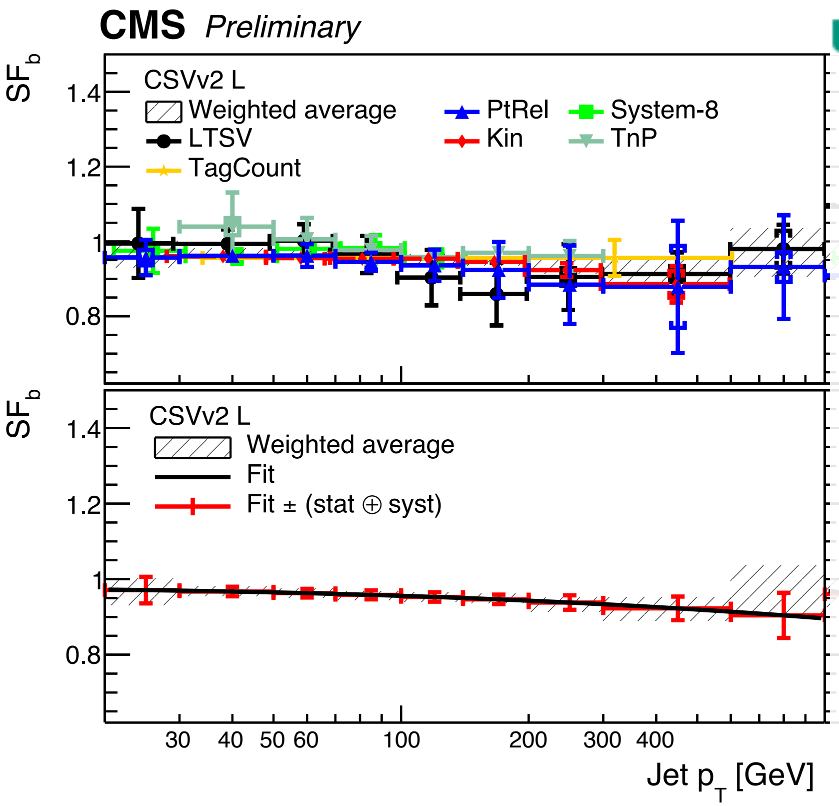


CMS DP-2018/33

- Performance of b-tagging algorithms:
  - **b-jet tagging efficiency**: fraction of true b-jets tagged as b-jets
  - **Misidentification probability** ("mistag rate"): fraction of true light-flavor (uds), charm (c), or gluon (g) jets wrongly tagged as b-jets
  - Depends on **physics process** considered, popular benchmark: $t\bar{t}$
  - Representation: **receiver-operating characteristic** (ROC)

Ulrich Husemann
Institute of Experimental Particle Physics

B-Tagging Calibration

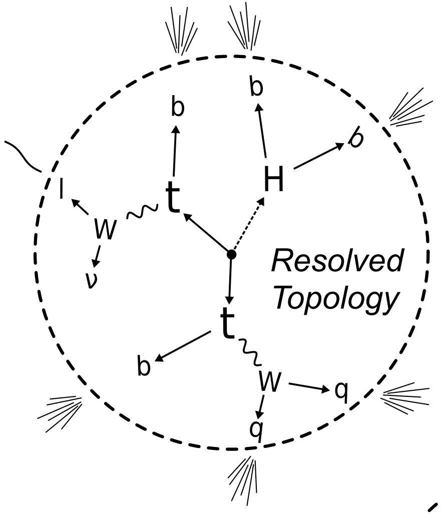


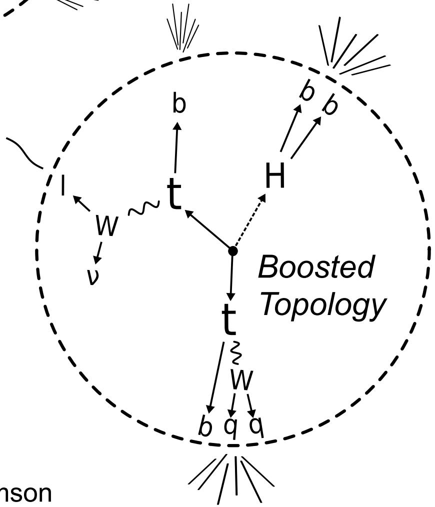
- Calibration of b-tagging algorithms:
  - Method: **measure** tagging efficiency and mis-ID probability for benchmark processes (e.g. QCD dijets, $t\bar{t}$) with **different methods** in **data and MC** → **correct MC** with scale factor SF_b depending on jet kinematics
  - Several **working points** (= cuts on b-tagging discriminant) with fixed mis-identification probability, e.g. 10% (“loose”), 1% (“medium”), 0.1% (“tight”)
  - Very useful: calibration of full b-tagging discriminant **shape**

Ulrich Husemann
Institute of Experimental Particle Physics

Boosted Objects

- Significant fraction of LHC events contains decays of heavy particles with **large transverse momenta** ($p_T \gtrsim 200$ GeV)
  - Decay products **strongly collimated** through large Lorentz boost
  - Need **novel algorithms** to reconstruct **jet substructure** and **tag** W/Z/H/top
  - Standard model example: associated $t\bar{t}H$ production with $H \rightarrow b\bar{b}$ decay
  - BSM example: heavy Z' decay





Courtesy of S. Williamson

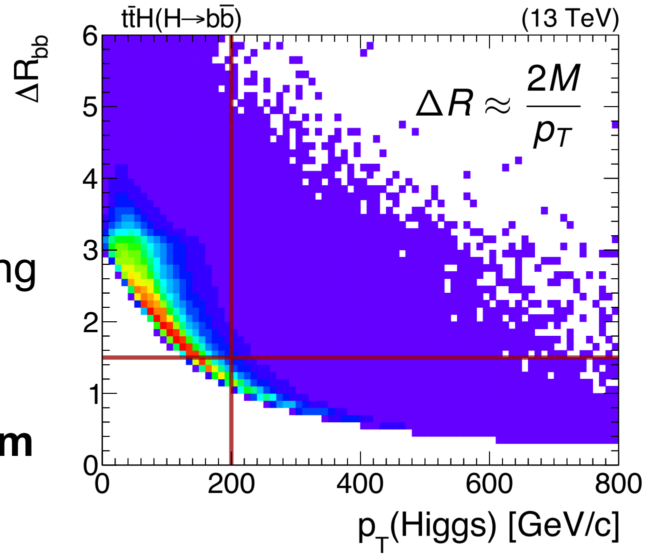
Ulrich Husemann
Institute of Experimental Particle Physics

Jet Substructure Algorithms

- Reconstruction of boosted objects: **very active field** at the LHC
  - General idea: reconstruct **“fat jets”** and study their **substructure**
  - More general sequential recombination jet algorithm:

$$d_{ij} = \min(k_{t,i}^{2n}, k_{t,j}^{2n}) \frac{\Delta R_{ij}^2}{R^2} \qquad d_{iB} = k_{t,i}^{2n}$$

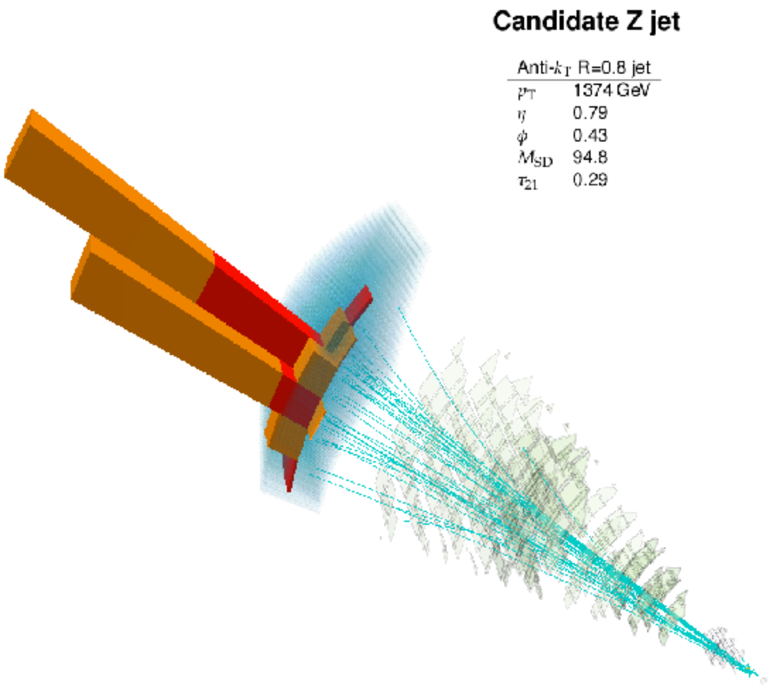
    (n = 0: Cambridge/Aachen, n = +1: k_t, n = –1: anti-k_t)
  - Anti-k_t jets: hard particle first clustered with surrounding soft particles → reversion of clustering steps (“declustering”) not meaningful
  - Jet clustering with **Cambridge/Aachen** or **k_t algorithm** → declustering reveals **substructure**
  - Typical fat-jet radius parameters: R = 0.8–1.5



S. Williamson, Dissertation, KIT (2016)

Jet Substructure Algorithms

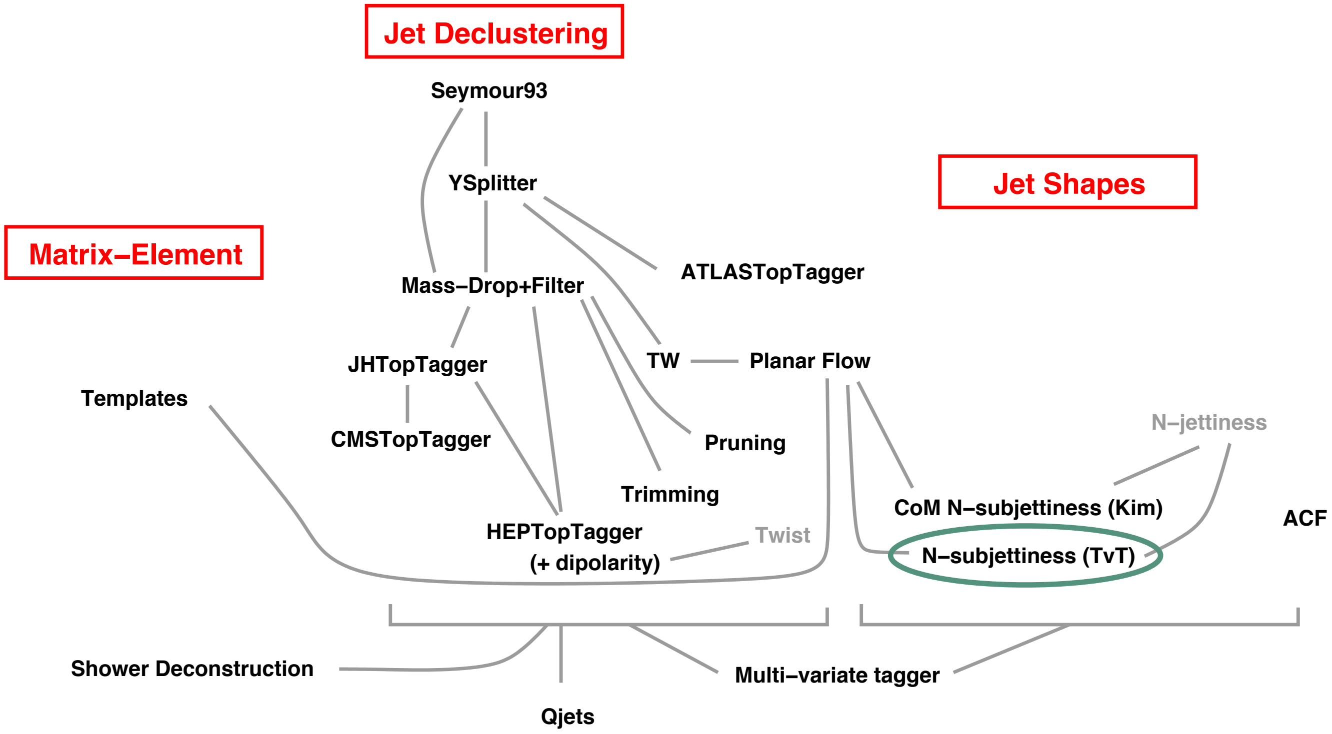
- Analysis of jet substructure:
  - Iterative **declustering** of fat jet
  - **Grooming**: removal of uncorrelated wide-angle soft emission from fat jet → better mass resolution, reduced pileup dependence
  - Jet **shape** algorithms, e.g. **N-subjettiness**
  - More involved algorithms used for tagging top and Higgs: **combination** of jet shape algorithms with grooming



CMS-PAS-B2G-17-001

Ulrich Husemann
Institute of Experimental Particle Physics

Jet Substructure Landscape



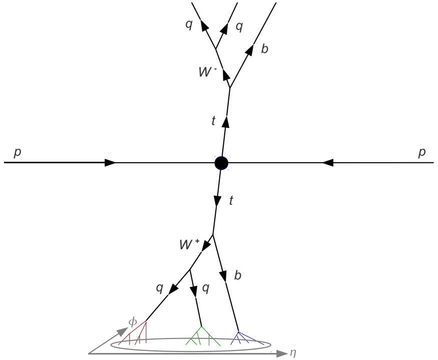
G. Salam, BOOST 2012

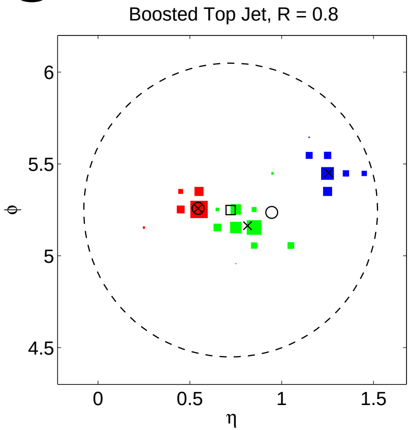
apologies for omitted taggers, arguable links, etc.

Ulrich Husemann
Institute of Experimental Particle Physics

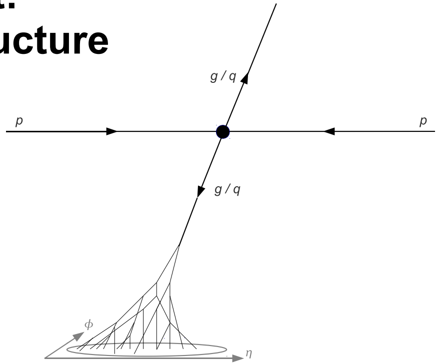
N-Subjettiness: Signatures

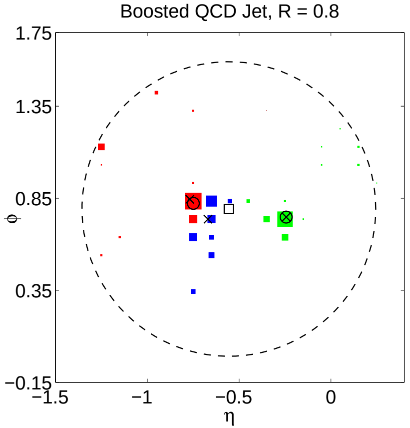
Top-quark decay: 3-prong structure





QCD quark/gluon jet: diffuse, no clear structure





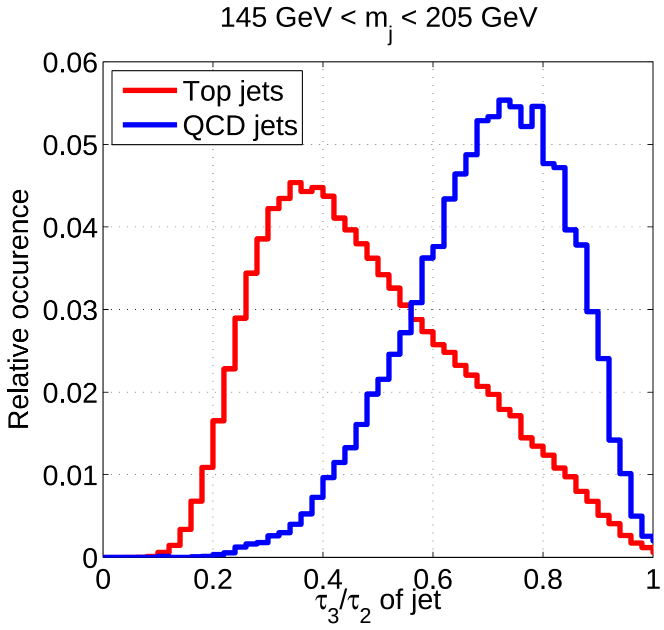
JHEP 03 (2011) 015

N-Subjettiness: Definition

- Jet-shape variable N-subjettiness τ_N: **energy flow** inside fat jets
- Jet with M particles → **deviation of energy flow** from ***N*** **subjet axes**

$$\tau_N = \frac{\sum_{i=1}^{M} p_{T,i} \min\{\Delta R_{i1}, \dots, \Delta R_{iN}\}}{\sum_{i=1}^{M} p_{T,i} R_0}$$

  - ΔR_{iJ}: η-ϕ distance from subjet axis J
  - Smaller τ_N → better description with N (or fewer) subjet axes
  - Good observables: **ratios** τ_N/τ_{N-1}, e.g. τ_3/τ_2 for top, τ_2/τ_1 for W, Z, H



JHEP 03 (2011) 015

Ulrich Husemann
Institute of Experimental Particle Physics

Short Summary

Karlsruhe Institute of Technology

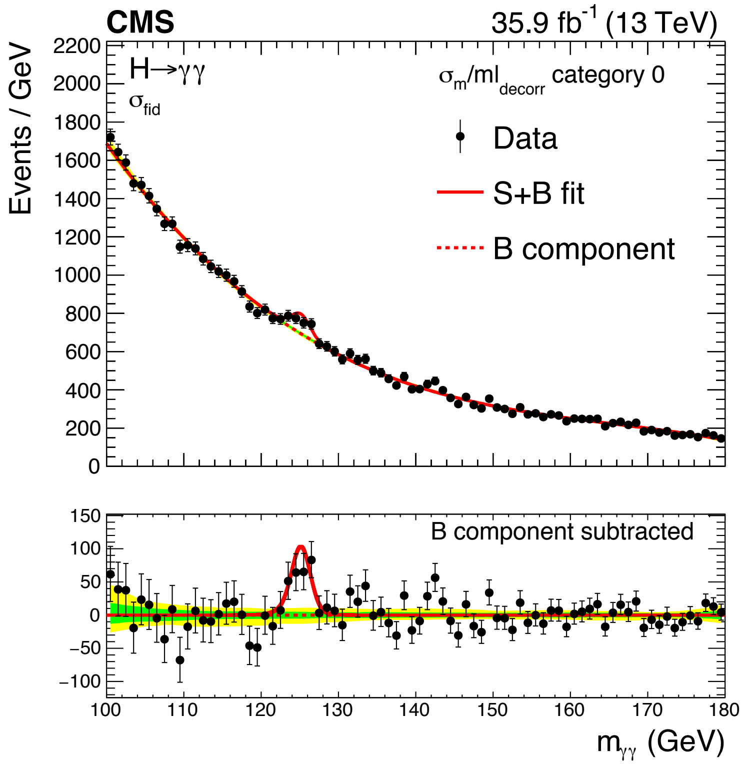
- Jet reconstruction:
  - Sequential recombination: infrared and collinear safe
  - Multi-stage calibration of **energy scale**
- Identification of **b-jets**: key to many process
  - Multivariate b-tagging algorithms
  - **Scale factors** for differences between data and MC
- **Boosted** jets: special treatment

Ulrich Husemann
Institute of Experimental Particle Physics

Background Estimation

Sidebands

- General idea: extract information on background in phase space region that is signal-depleted but **representative** of signal region
- **Sideband** techniques:
  - Use case: signal = **narrow** invariant mass peak on large (often combinatorial) background (e.g. $H \to \gamma\gamma$)
  - Estimate background normalization (sometimes: also shape) from the same invariant mass spectrum **outside** peak



arXiv:1807.03825, accepted by JHEP

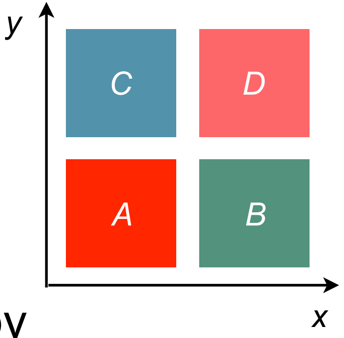
Ulrich Husemann
Institute of Experimental Particle Physics

Control Regions

Karlsruhe Institute of Technology

- Use case: MC simulation of background in signal region **unreliable** (e.g. corner of phase space that is not well modeled or lacking statistics)
- Event selection for control region such that signal and control regions are **mutually exclusive** (jargon: “orthogonal”), e.g. by **inverting** certain selection criteria
- **Measure** background normalization (and shape) **in control region**, **transfer to signal region** (usually using MC simulation)

Control Regions: ABCD Method



- Idea: measure background in three control regions $\rightarrow$ predict background in signal region
  - Define four regions A, B, C, D in space of **uncorrelated** observables x and y $\rightarrow$ signal in A (small x AND small y)
  - Measure background b in B, C, D $\rightarrow$ background in A given by

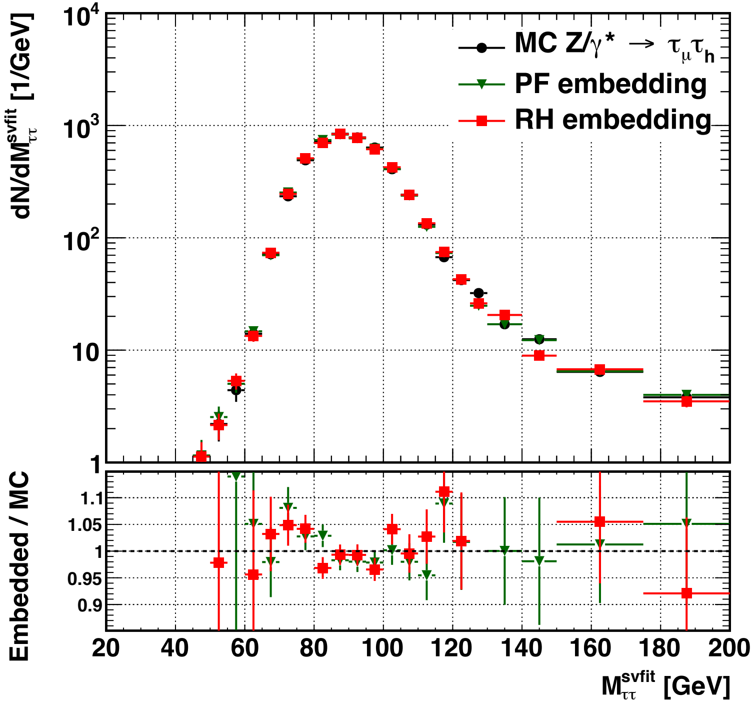
$$b(A) = \frac{b(B) \times b(C)}{b(D)}$$

- Example: CMS $H \rightarrow \tau\tau$ analysis – QCD multijets in $H \rightarrow \tau_{\mathrm{h}}\tau_{\ell}$
  - Observables: charge sign (same vs. opposite charge) and isolation of hadron and lepton (tight vs. relaxed)

Ulrich Husemann
Institute of Experimental Particle Physics

Closure Test

- Background estimation from sidebands or control regions: **consistency check** required
  - **Closure test**: does the method “close” on simulated events (i.e. is **known** background process predicted **accurately**, are there biases)?
  - Often: amount of “non-closure” used as **systematic uncertainty** due to background estimation method
- Example: closure test of tau modeling method in $H \rightarrow \tau\tau$ (“embedding”)

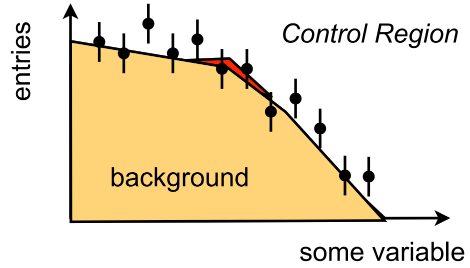


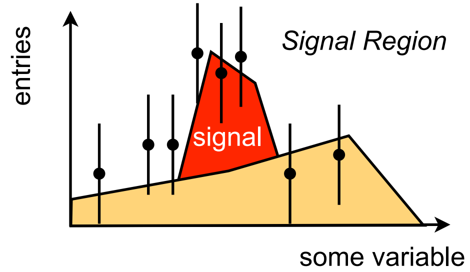
A. Burgmeier, Dissertation, KIT (2014)

Ulrich Husemann
Institute of Experimental Particle Physics

In-Situ Background Determination

- So far: **a-priori** expectation of background normalization of shape in signal region f**rom MC simulation** or **data-driven method**
- Signal extraction in many analyses: **profile-likelihood fit**
  - **Simultaneous fit** of signal and control region(s) → background in control region(s) **constrains** background in signal region
  - Systematic uncertainties included as **nuisance parameters** (e.g. normalization and shape from ABCD = nominal value, uncertainty from closure test)
  - Assumptions: fit model **adequate**, **correlations** between signal and control regions **well modeled**





Ulrich Husemann
Institute of Experimental Particle Physics

Summary

- Various methods to determine backgrounds: **MC simulation** or **data-driven** methods
- Data-driven methods: Measure background in signal-depleted region in data → estimate background in signal-enriched region (closure test required)
- **Various** techniques, **specific** to analysis and background:
  - Combinatorial background underneath mass peak: fit to parametric signal and background model
  - Continuum background: control regions, ABCD method, …
  - If background is hard to model: embedding of MC objects in data events
  - Simultaneous profile-likelihood fits to signal and control regions
  - …

Advanced Signal Analysis

Ulrich Husemann
Institute of Experimental Particle Physics

Differential Cross Section

- So far: **reconstructed** distributions of kinematic observables compared to expected distributions (from MC and/or data)
  - All physics effects **forward-folded** with detector effects (e.g. resolution)
  - Problem: distributions **cannot be compared** between experiments
- Way out: measurements presented as **differential cross sections** = cross sections as a function of one or more kinematic observable
  - Detector effects corrected by **unfolding** procedure
  - Typical result: **fiducial differential cross section** on level of **stable particles**
  - Differential distributions contain **more information** on physics processes than inclusive cross sections → **more detailed comparison** with theory

Ulrich Husemann
Institute of Experimental Particle Physics

Unfolding Techniques

- Determine **true** distribution $f(\mathbf{y})$ from **reconstructed** distribution $g(\mathbf{x})$:
  - Relation: Fredholm integral equation

$$g(\mathbf{x}) = \int R(\mathbf{x}|\mathbf{y})\, f(\mathbf{y})\, \mathrm{d}\mathbf{y} + b(\mathbf{x}) = \int A(\mathbf{x}|\mathbf{y})\, \epsilon(\mathbf{y})\, f(\mathbf{y})\, \mathrm{d}\mathbf{y} + b(\mathbf{x})$$

  - **x**: observed (reconstruction-level) kinematics, **y**: “true” kinematics
  - $R(\mathbf{x}|\mathbf{y})$ **transfer function** (“translation” from true to reconstructed kinematics), can be factorized in **acceptance function** $A(\mathbf{x}|\mathbf{y})$ and **efficiency function** $\epsilon(\mathbf{y})$
  - $b(\mathbf{x})$ **background** distribution

Unfolding Techniques

- Unfolding = solving integral equation for $f(y)$
  $\rightarrow$ **ill-posed** mathematical problem, typical solutions:
  - First step: discretization (= histograms), **response/migration matrix** R

$$g_i = \sum_{j=1}^{m} R_{ij} f_j + b_i$$

  - If R ~ diagonal: **bin-by-bin** correction factors c_i may be sufficient:

$$g_i = c_i f_j + b_i$$

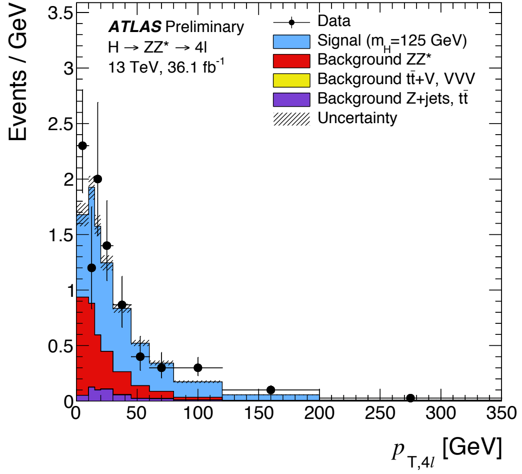
  - **Matrix inversion** of R: numerically unstable due to **statistical fluctuations**
    $\rightarrow$ additional assumption: smooth distributions ("regularization")

Ulrich Husemann
Institute of Experimental Particle Physics

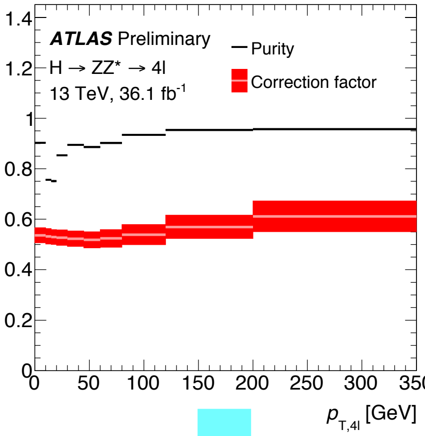
Bin-by-bin Unfolding: $H \to ZZ \to 4\ell$

$H \to ZZ \to 4\ell$: **clean** signature, high **purity** (i.e. response matrix almost diagonal)
→ **bin-by-bin** unfolding

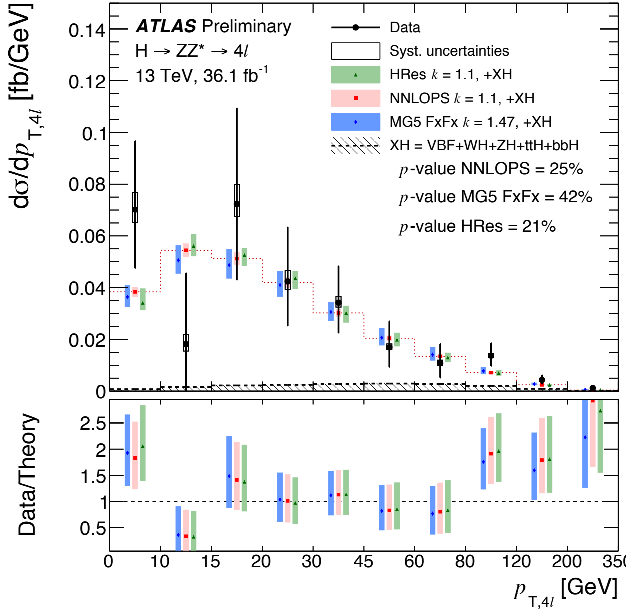
Reconstructed $p_T(4\ell)$



Correction Factors



Unfolded $p_T(4\ell)$

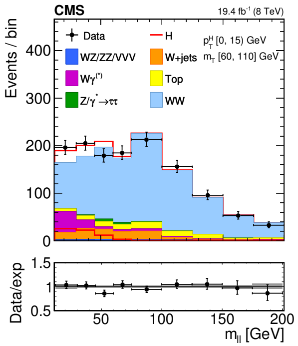


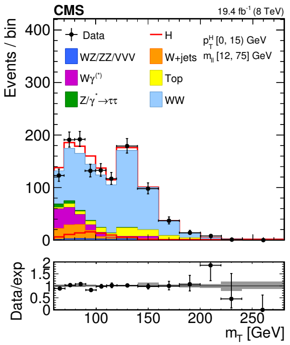
ATLAS-CONF-2017-032

Ulrich Husemann
Institute of Experimental Particle Physics

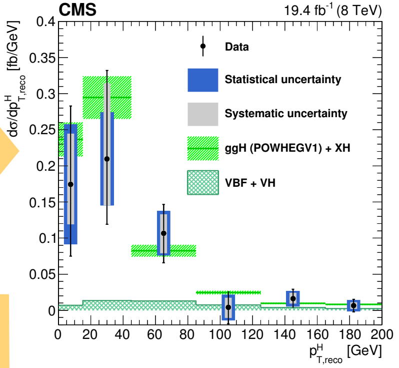
Matrix Unfolding: $H \to WW$

Fit to $m_{\ell\ell}$ and m_T in bins of $p_T(H)$

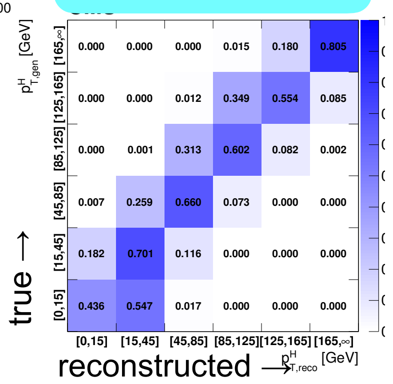




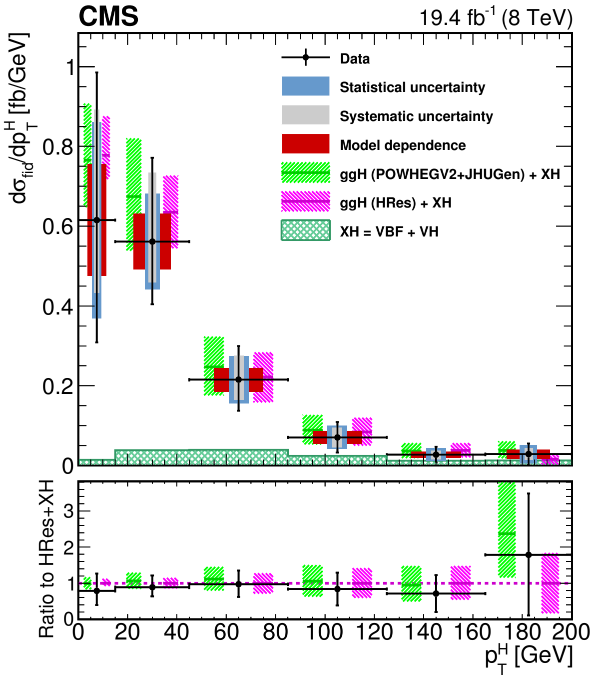
Reconstructed $p_T(H)$ distribution



Response matrix



Unfolded $p_T(H)$ distribution



JHEP 03 (2017) 032

Matrix-Element Method

- **Entire** parton-level kinematics of a process contained in **squared scattering amplitude** ("matrix element", ME)
- **Matrix-element method** (MEM): construct **event-based likelihood** discriminant that **fully exploits** all information from matrix element
  - Likelihood function for a given process contains hard-scattering matrix element for that process (→ next slide)
  - For each event: **ratio of likelihood functions** for observed set of kinematic variables $\mathbf{x}$ under signal hypothesis S and background hypotheses B_i

$$R(\mathbf{x}) = \frac{L(\mathbf{x}|S)}{L(\mathbf{x}|S) + \sum_i c_i \, L(\mathbf{x}|B_i)}$$

  - Full discriminant: **product** of event-based discriminants **for all events**

Ulrich Husemann
Institute of Experimental Particle Physics

Matrix Element and Phase Space

- Main ingredient of event-based likelihood: **differential parton-level cross sections** for signal and (main) backgrounds
  - Consider cross section for all processes $pp \to y$ with parton-level kinematics y that could have led to the reconstruction-level final state x with kinematics x

$$\sigma(pp \to y) = \sum_{jk}^{\text{partons}} \int_0^1 \mathrm{d}z_j \, \mathrm{d}z_k \underbrace{f_j(z_j)\, f_k(z_k)}_{\text{PDFs}} \frac{(2\pi)^4}{z_j z_k s} \underbrace{\left|M(jk \to y)\right|^2}_{\text{matrix element}} \underbrace{\mathrm{d}\Phi}_{\text{phase space}}.$$

  - Approach uses **QCD factorization**: PDFs f_j, f_k, (squared) hard-scattering matrix element M, Lorentz-invariant phase space measure $\mathrm{d}\Phi$
  - Current implementations: **LO matrix elements** (NLO in the works)
  - **Integration** over all unobserved variables in the event: momentum fractions of colliding partons, phase space integral → often **numerically expensive**

Ulrich Husemann
Institute of Experimental Particle Physics

Transfer Functions

- Translation of parton-level final state to **reconstruction level**:
  - Folding with **transfer functions** $W(\mathbf{x}|\mathbf{y})$ determined from MC simulation
  - $W(\mathbf{x}|\mathbf{y})$ accounts for **limited detector resolution** and **combinatorics** in matching parton level and reconstruction level objects (especially quarks/ gluons → jets)

$$\sigma(pp \to x) = \int \sigma(pp \to y)\, W(\mathbf{x}|\mathbf{y})\, \mathrm{d}\mathbf{y}$$

- Normalization to (fiducial) cross section of process $P = S, B$

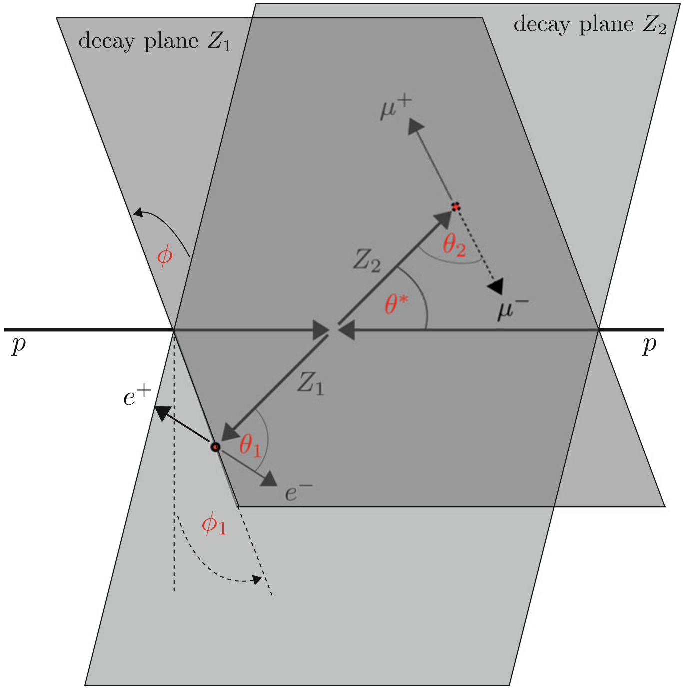
$$L(\mathbf{x}|P) = \frac{\sigma(pp \to x)}{\sigma^P_{\mathrm{obs}}} \quad \text{with} \quad \sigma^P_{\mathrm{obs}} = \int \sigma(pp \to y)\, W(\mathbf{x}|\mathbf{y}) \cdot f_{\mathrm{acc}}(\mathbf{x})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y}$$

and $f_{\mathrm{acc}}(\mathbf{x}) = 0;1$ acceptance for single event with kinematics $\mathbf{x}$

MEM Application: $H \to ZZ \to 4\ell$

Karlsruhe Institute of Technology

- **$H \to ZZ \to 4\ell$ angular analysis**
  $\to$ Higgs-boson spin and parity
- Kinematics fully determined by
  - 2 masses m_{Z1}, m_{Z2}
  - Decay planes of $Z_{1,2}$
    $\to$ 5 angles $\boldsymbol{\Omega} = (\theta^*, \phi_1, \phi, \theta_1, \theta_2)$
    - Polar angle of Z bosons (θ^*)
    - Azimuthal angle of Z_1 plane (ϕ_1)
    - Azimuthal angle of Z_2 plane relative to Z_1 plane (ϕ)
    - Polar angles of leptons relative to $Z_{1,2}$ ($\theta_{1,2}$)



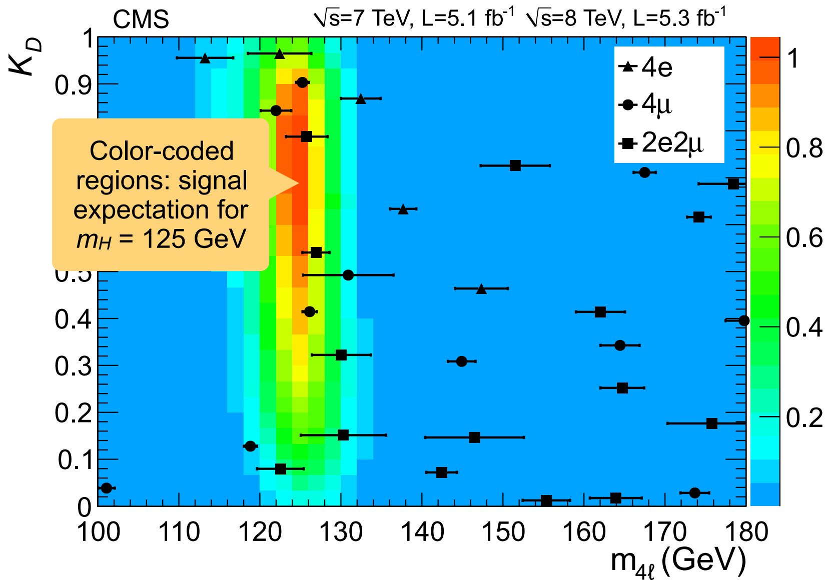
R. Wolf

Ulrich Husemann
Institute of Experimental Particle Physics

MEM Application: MELA

- Application of MEM to angular analysis of $H \rightarrow ZZ \rightarrow 4\ell$
  - MELA: Matrix Element Likelihood Analysis (based on PRD 81 (2010) 075022) → already applied for CMS Higgs discovery analysis
  - Purely leptonic final state: no phase space integration and transfer functions required
  - MELA discriminant:

$$K_D = \frac{L(m_{Z1}, m_{Z2}, \Omega; m_{4\ell}|S)}{L(m_{Z1}, m_{Z2}, \Omega; m_{4\ell}|S) + L(m_{Z1}, m_{Z2}, \Omega; m_{4\ell}|B)}$$



Phys. Lett. B716 (2012) 30

Ulrich Husemann
Institute of Experimental Particle Physics

Conclusions

Karlsruhe Institute of Technology

- Experimental particle physics: **many tools and techniques**, e.g.
  - Simulation of collision processes: **Monte-Carlo** event generators
  - Reconstruction, ID, calibration of all **physics objects**
  - Treatment of **background** processes
- **Many opportunities** for you to dig deep into particle physics

Tutorial

TAG&PROBE EFFICIENCY

Ulrich Husemann
Institute of Experimental Particle Physics

Measuring the Trigger Efficiency

- Tag&probe trigger efficiency definition:

$$\epsilon = \frac{N(\text{reference \&\& probe})}{N(\text{reference})} = \frac{N(\texttt{HLT\_IsoMu20 \&\& HLT\_PFJet500})}{N(\texttt{HLT\_IsoMu20})}$$

Motivate this definition by discussing the following questions:

- What is the purpose of the reference trigger?

  Cannot measure absolute efficiency in data.

- Why do we require the muon of the reference trigger to be isolated? What could happen if one considers any muon (also non-isolated)?

  Tag should be independent of probe → muon.
  Non-isolated muon may be part of a jet → not independent.

Ulrich Husemann
Institute of Experimental Particle Physics

Measuring the Trigger Efficiency

Karlsruhe Institute of Technology

- Tag&probe trigger efficiency definition:

$$\epsilon = \frac{N(\text{reference \&\& probe})}{N(\text{reference})} = \frac{N(\texttt{HLT\_IsoMu20 \&\& HLT\_PFJet500})}{N(\texttt{HLT\_IsoMu20})}$$

  - Could we use a trigger that fires at random as the reference trigger?

    Yes, but the rate will be much too low → see later.

  - Could we, instead of using the single-muon trigger, use a trigger that requires the presence of a jet with a p_T threshold lower than the threshold of HLT PFJet500 as the reference trigger, e.g. a trigger requiring a jet with $p_T > 300$ GeV?

    Yes, but but reference trigger must be constant → see later.

Ulrich Husemann
Institute of Experimental Particle Physics

Measuring the Trigger Efficiency

- Inspect the two jet p_T histograms and discuss the following questions:
  - Why does the number of entries per bin decrease towards large p_T?

    Differential cross sections of all processes decrease with increasing p_T

  - What is the reason for the turn-on at low p_T?

    Events recorded with lepton trigger are mainly W+jets events,
    minimum momentum transfer $\rightarrow$ minimal p_T of order of $m_W/2$

Measuring the Trigger Efficiency

Karlsruhe Institute of Technology

- Inspect the trigger efficiency plot and answer the following questions:
  - What is the efficiency of the HLT PFJet500 trigger path?

    Heavily p_T-dependent: zero efficiency → turn-on → plateau
    (if efficiency is quoted: plateau efficiency)

  - Why is there a smooth turn-on region around 500 GeV where the efficiency gradually increases? Why does the trigger not reach its maximum efficiency instantly at p_T = 500 GeV?

    Online reconstruction of p_T (in trigger) and offline reconstructed
    (for plotted p_T) different (→ resolution effect)

Ulrich Husemann
Institute of Experimental Particle Physics

Measuring the Trigger Efficiency

- In the light of the turn-on feature of a trigger efficiency, consider again the question:
  - Could we, instead of using the single-muon trigger, use a trigger that requires the presence of a jet with a lower p_T threshold than HLT PFJet500, e.g. a trigger that requires a jet with $p_T > 300$ GeV?

    Reference trigger must be in plateau (not necessarily fully efficient!), i.e. beyond on turn-on, before turn-on of probe trigger starts

  - Which condition must be satisfied when a trigger with lower threshold is used as the reference trigger?

    Need sufficient difference in thresholds. Exact conditions depend on the offline-vs-online resolution of objects, typically large for jets, small for leptons

Uncertainty of the Efficiency

Karlsruhe Institute of Technology

- Have a look at the error bars in the efficiency plot produced in the previous exercise:
  - Are they reasonable? Not really, they extend below 0 and above 1.
  - How are they calculated? Error propagation of Poisson uncertainties.
- You can switch to using binomial uncertainties by adding the "B" option to the TH1::Divide method.
  - How do the error bars change? Error bars vanish for for 0 and 1.
  - Is this reasonable? Yes, except for 0 and 1.
- Adjust calculate eff.py to use Clopper–Pearson intervals as error bars: How do the error bars change? Is this reasonable?

  Yes, for entire efficiency interval [0,1].

When is Trigger "Fully Efficient"?

- Given the turn-on feature, we can assume the trigger efficiency to remain constant for large p_T far above the turn-on region. Why?
  Jet trigger and reconstruction stays fully efficient to very high p_T.
- A suitable function to fit the turn-on is

$$f(p_T; a_0, a_1, a_2) = \frac{1}{2} \cdot a_2 \cdot \left[\mathrm{erf}\left(\frac{1}{\sqrt{2}a_0}(p_T - a_1)\right) + 1 \right]$$

  - How is the error function erf(x) defined and why is it suitable in this case? (Remember again what the reason for the turn-on feature was!)
    Cumulative distribution of the normal (Gaussian) distribution.
    Suitable because turn-on is a (Gaussian) resolution effect.
  - What is the interpretation of the parameters a_i? (Which trigger threshold do you find? What is the efficiency of the trigger above the threshold?)
    a_0: width of the Gaussian
    a_1: turn-on point (50% of plateau efficiency)
    a_2: normalization → plateau efficiency

Efficiency of a Different Trigger

- We want to use the tools developed above to measure the efficiency of a single-jet trigger with a different threshold. Adjust your calculate eff.py to measure the efficiency of the HLT PFJet60 trigger What do you observe?

This should be possible in principle
(but not with the way the histograms in histos.root were prepared).
However, certain triggers may have been pre-scaled
(i.e. only every *n*-th event recorded)
→ observable consequence: plateau much below 1