Jet&MET overview at CMS

Robert Schöfbeck for JetMET

JME Workshop Helsinki 17/05/09

intro

- the last two years have been a fun ride!!
  - I want to thank everybody who has worked with me!

- Think global, act local
  - quote from our host
    - not really, rather from P. Geddes, 1915

- Goal of the 4th JME workshop
  - get ready for 2017 data taking
  - collect and fix all analysis details where we can improve
  - establish a plan until 2020 exploiting fully the Run-II dataset

team

JME
Robert Schoefbeck
Seema Sharma

JERC
Anastasia Karavdina
Henning Kirschenmann

MET
Zeynep Demiragli
Loukas Gouskos

JMAR
Kevin Nash
Gregor Kasieczka

JetMET DQM
Giuseppe Latino
Raman Khurana

JetMET Trigger
Aruna Nayak
Milos Djordjevic

Contacts

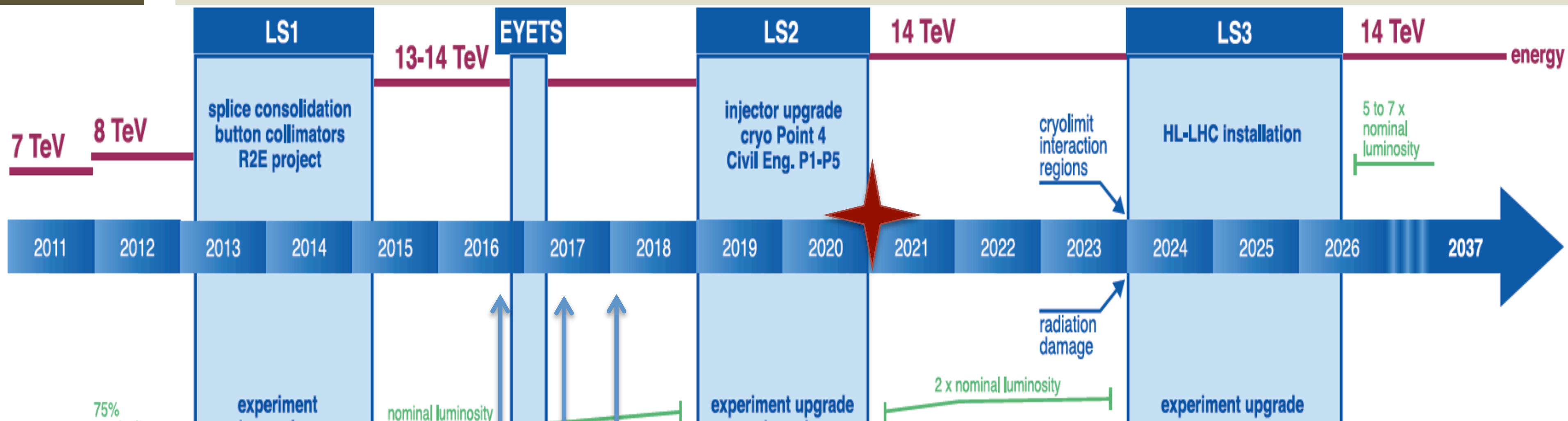
Generators
Yorgos Tsipolitis
Alca
Sangeun Lee
Jet RECO
James Dolen
Andreas Hinzmann
Salvatore Rappoccio
MET RECO
Matthieu Marionneau
TOP
Hannu Siikonen
Andrey Popov
EXO
Yutaro Liyama
Keith Rose
SMP
Stephanie Brandt
Patrick Connor
SUS
Andrew Whitback
Batool Safarzadeh

special thanks to the outgoing conveners
Ia Iashvili, Mariarosaria D'Alfonso, Jordan Damgov,
Andreas Hinzmann, James Dolen, Jane Nachtmann, Matthieu Marionneau

documentation / papers

- 8 TeV paper out !
  https://arxiv.org/abs/1607.03663
- particle flow @ 8 TeV (with JME contributions) is imminent
- in the pipeline with 36/fb
  - jet energy corrections and performance
  - MET performance
  - 'pile-up' paper on Puppi vs. CHS
- PASes/DPS
  https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsJME

LHC schedule



20/fb

36/fb

50/fb?

50/fb?

0.75 10^{34} $cm^{-2}s^{-1}$
50ns spacing
PU 10-35

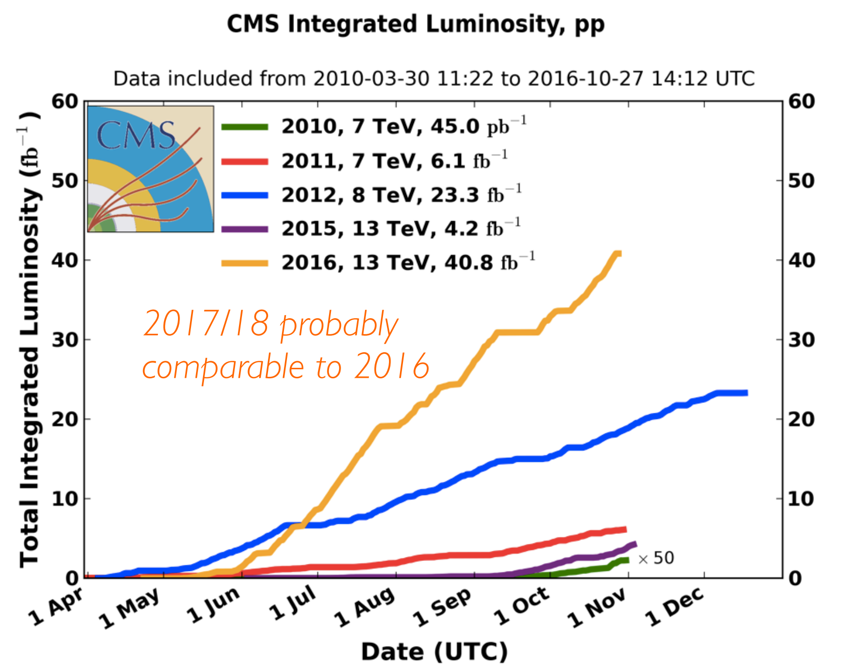
1.5 10^{34} $cm^{-2}s^{-1}$
25ns spacing
PU up to 40

HE upgr.
HPDs to
SiPM

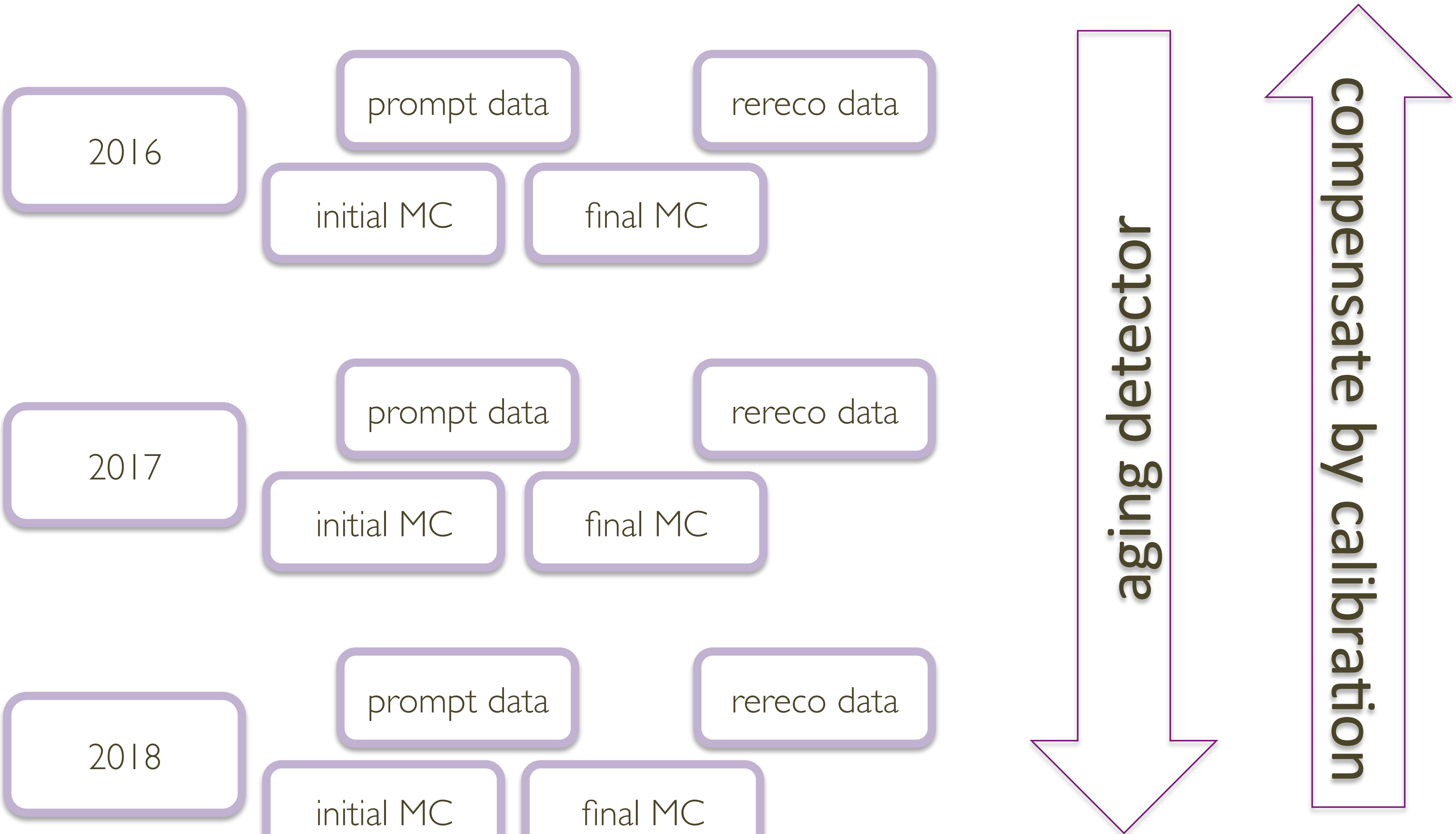
1.9 10^{34} $cm^{-2}s^{-1}$?
new pixel,
HF electronics
<PU> 40?

2016/17/18

- PU will increase
- lumi profile similar
- data will be different!

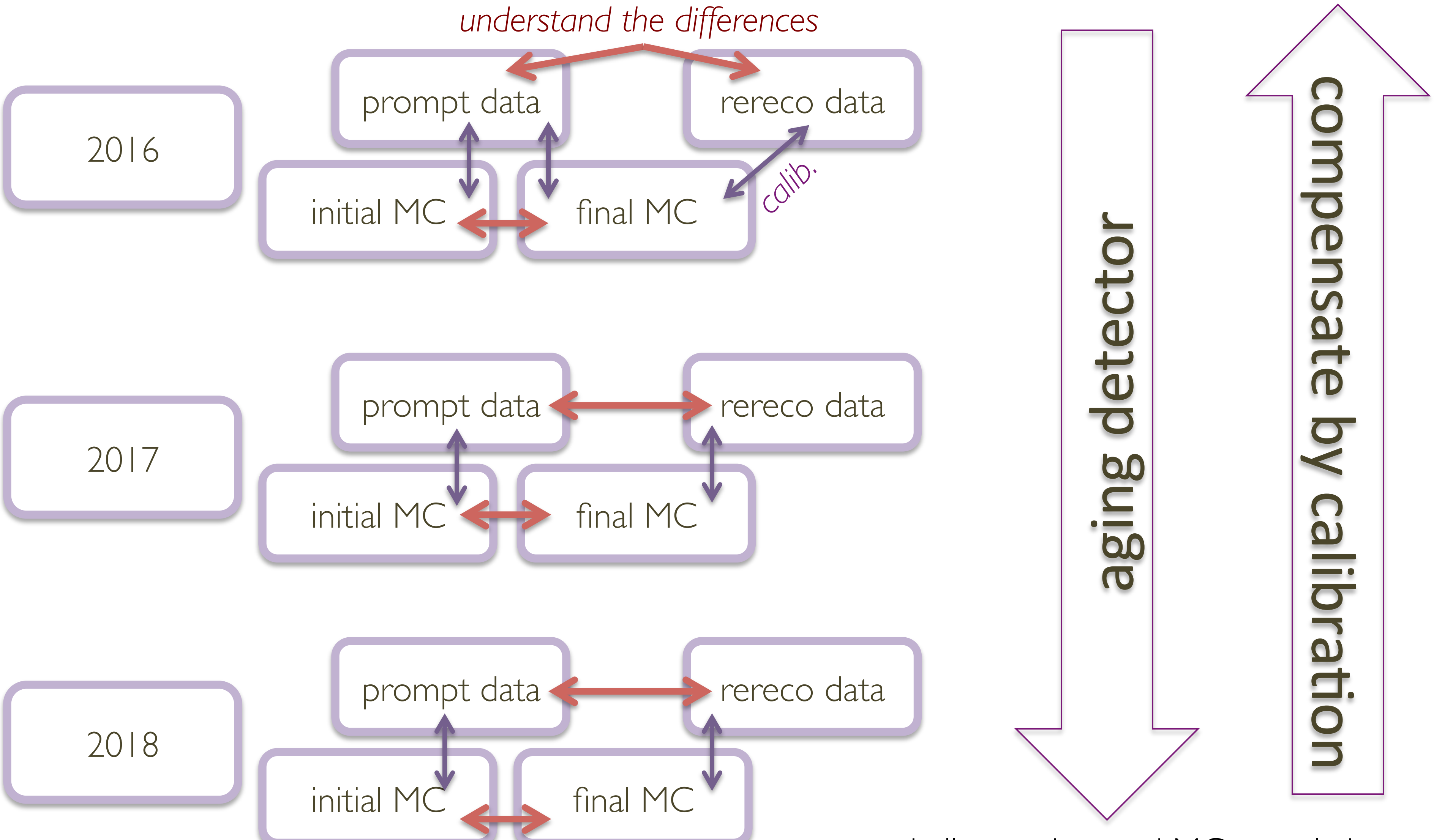


correlations and homogeneity



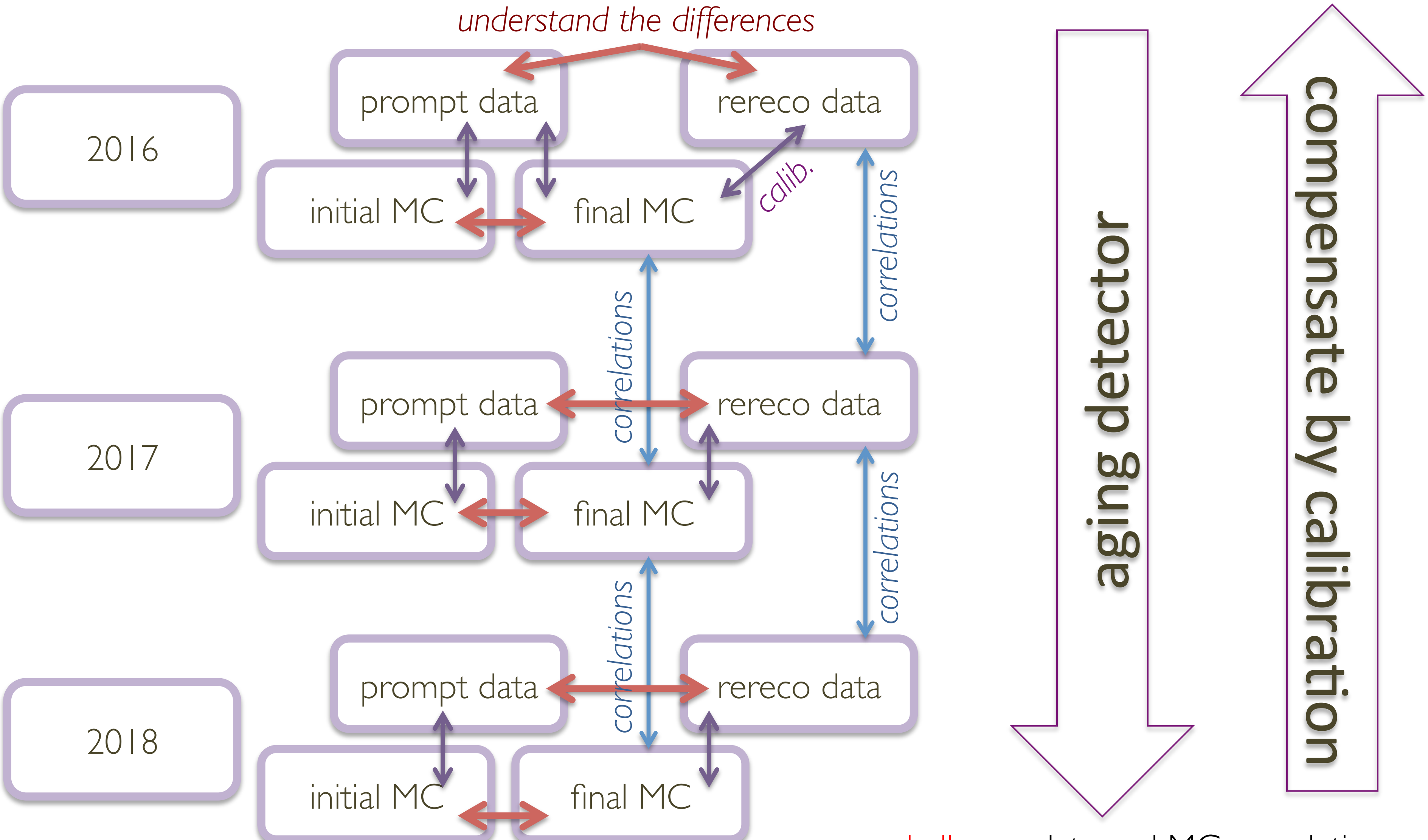
new challenge: data and MC correlations

correlations and homogeneity



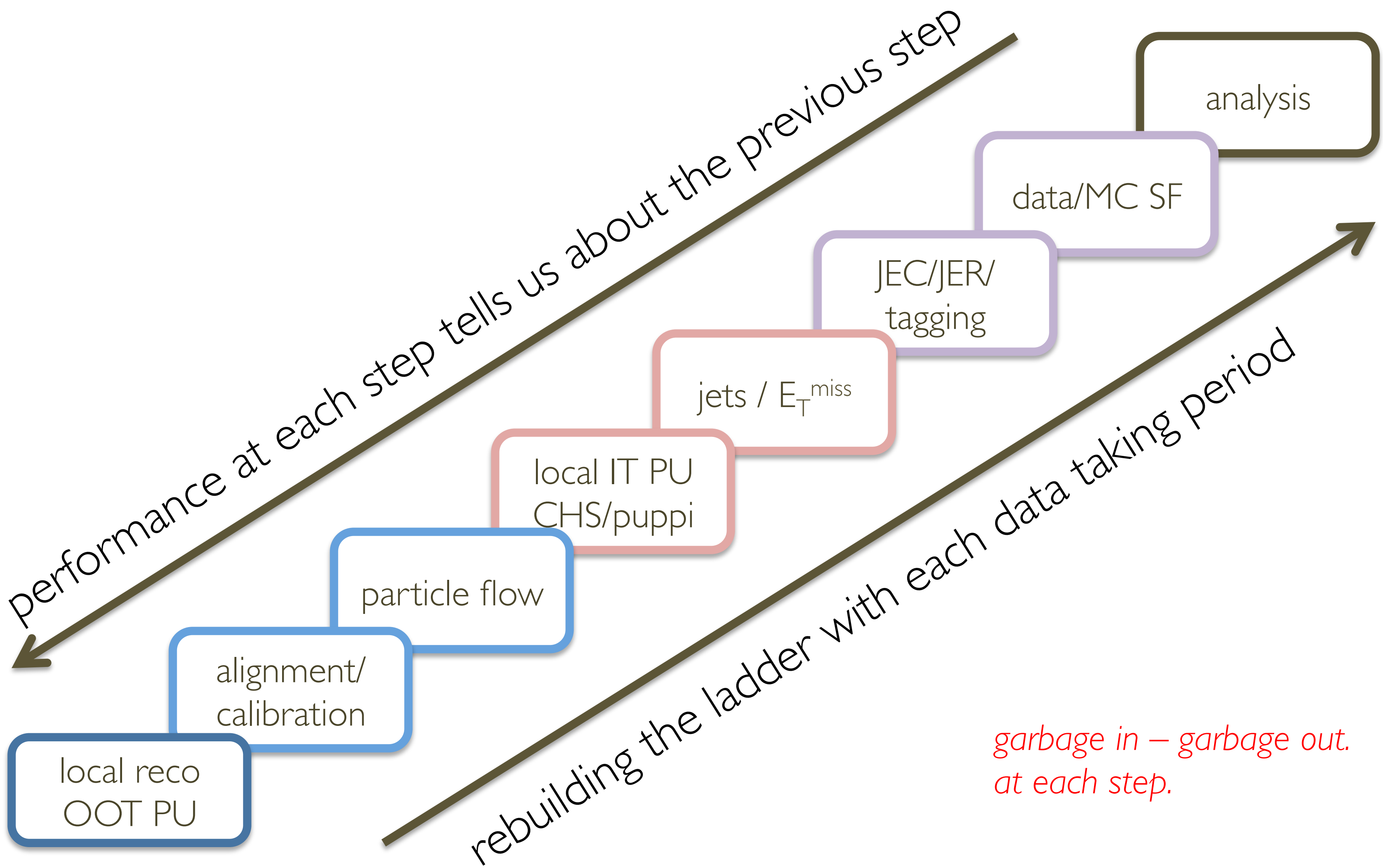
new challenge: data and MC correlations

correlations and homogeneity

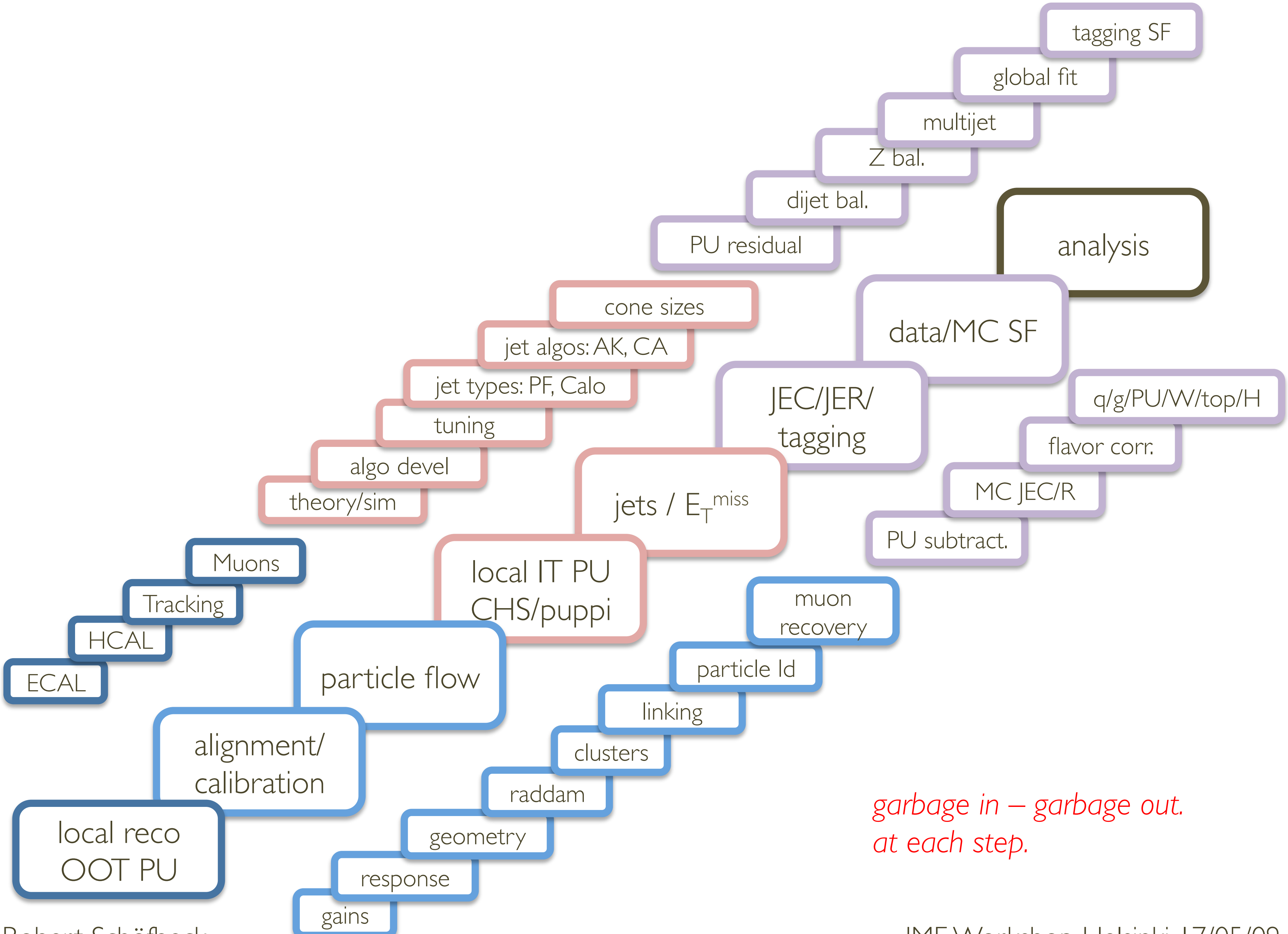


new challenge: data and MC correlations

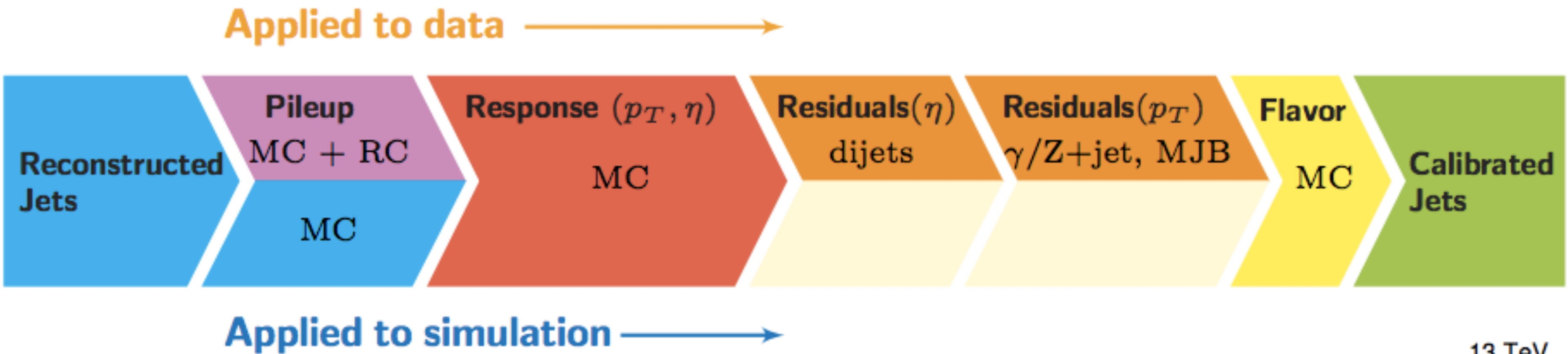
reconstruction ladder



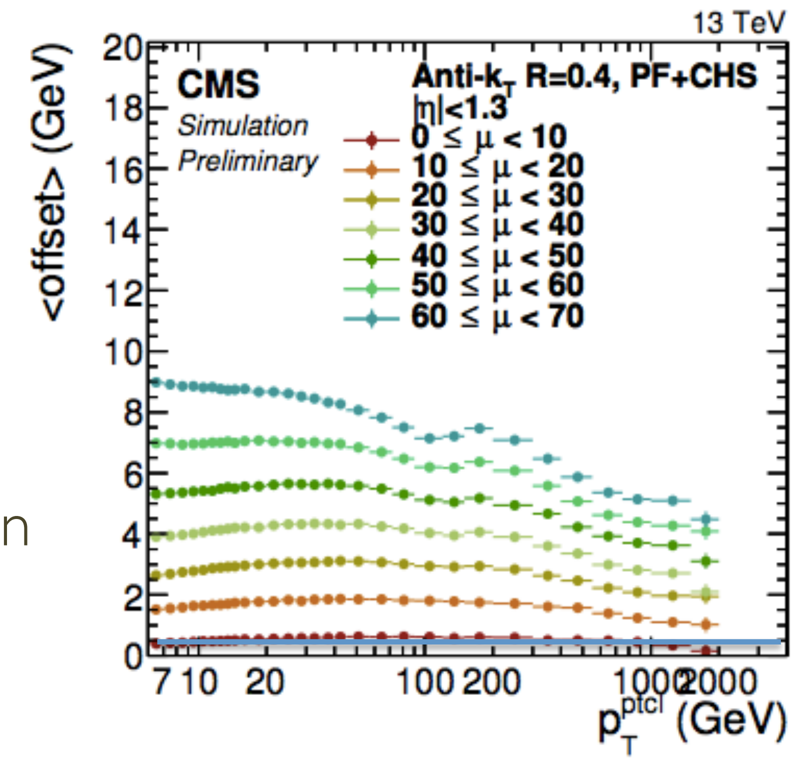
garbage in – garbage out.
at each step.


tagging SF
global fit
multijet
Z bal.
dijet bal.
PU residual
analysis
data/MC SF
q/g/PU/W/top/H
JEC/JER/
tagging
flavor corr.
MC JEC/R
PU subtract.
cone sizes
jet algos: AK, CA
jet types: PF, Calo
tuning
algo devel
theory/sim
jets / E_T^{miss}
local IT PU
CHS/puppi
Muons
Tracking
HCAL
ECAL
particle flow
alignment/
calibration
local reco
OOT PU
muon
recovery
particle Id
linking
clusters
raddam
geometry
response
gains
garbage in – garbage out.
at each step.

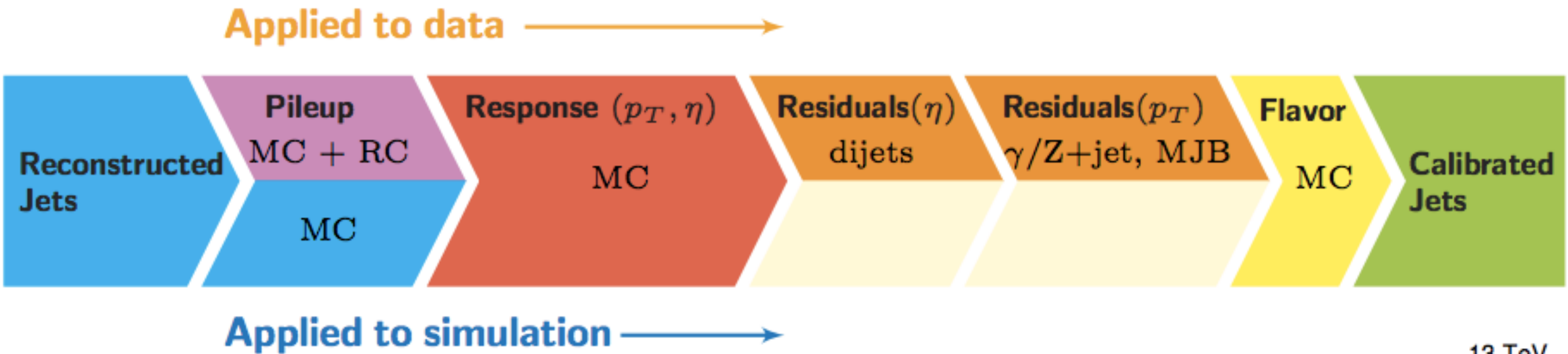
jets at 13 TeV



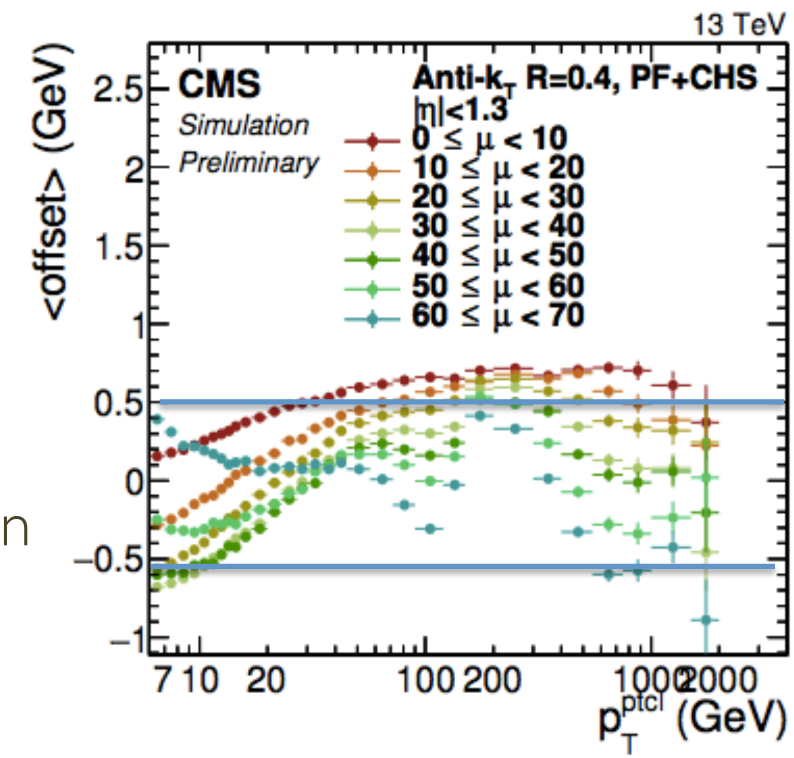
- subtract simulated PU
- subtle nonlinear effects
  - PU un-zerosuppression, shadowing
  - small dependence on jet p_T
  - first subtraction based on simulation
  - scaled by data/MC ratio of RC offset energy density



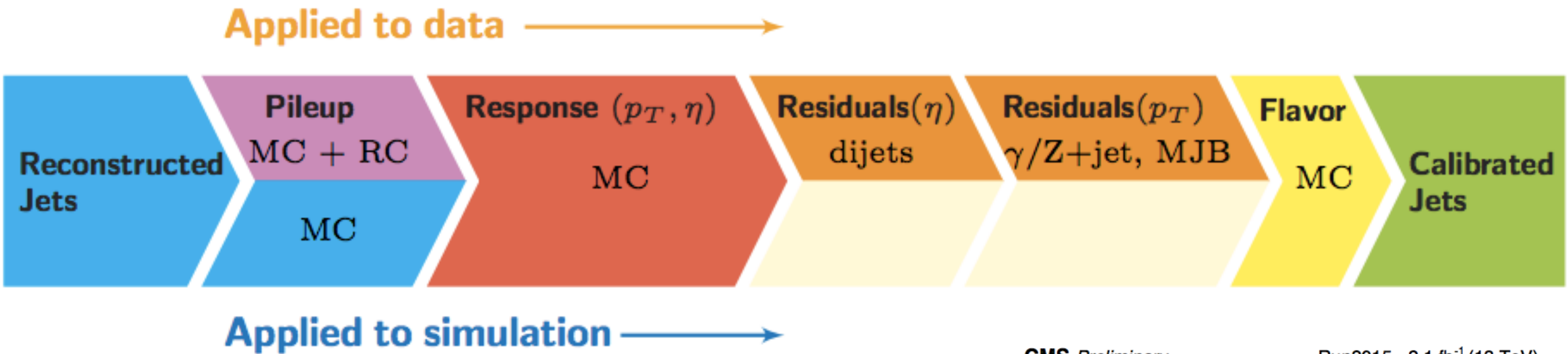
jets at 13 TeV



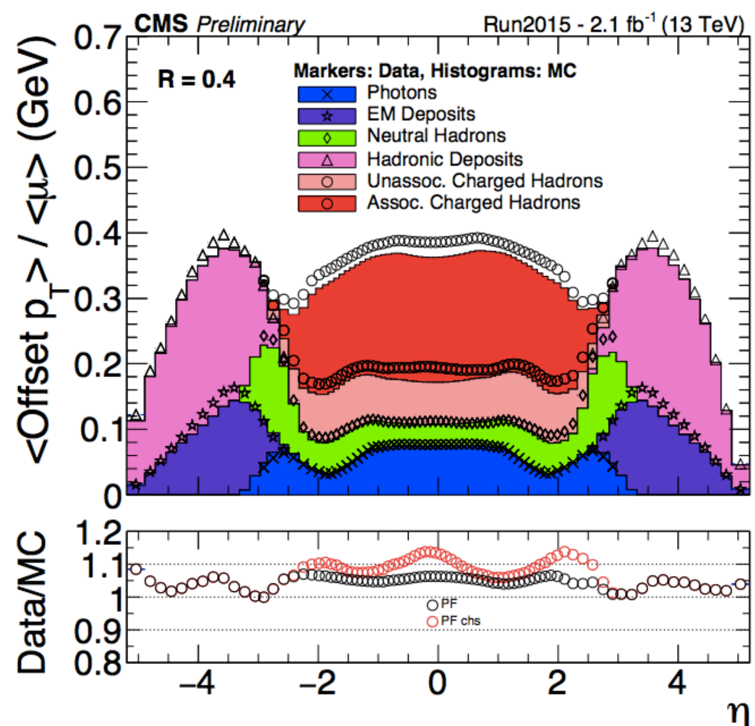
- subtract simulated PU
- subtle nonlinear effects
  - PU un-zerosuppression, shadowing
  - small dependence on jet p_T
  - first subtraction based on simulation
  - scaled by data/MC ratio of RC offset energy density



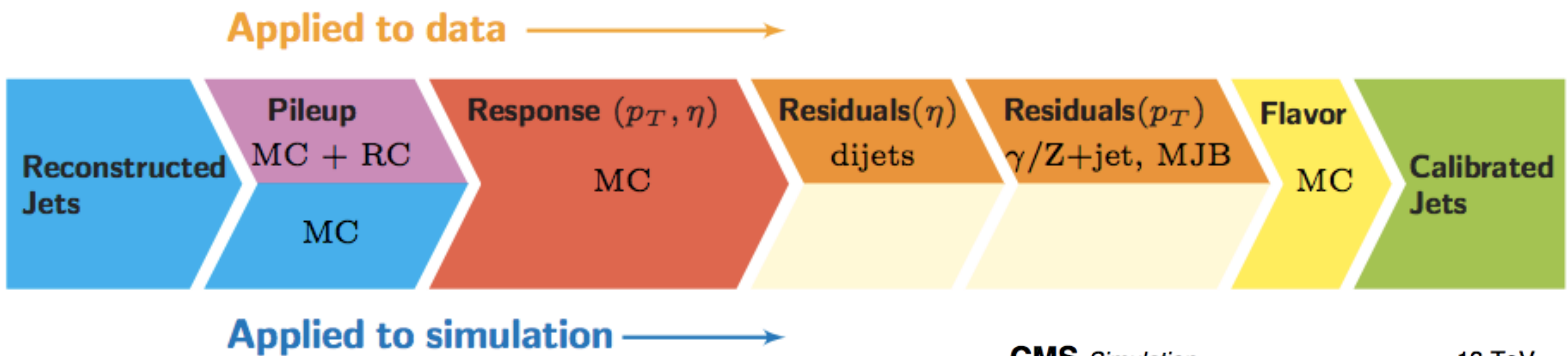
jets at 13 TeV



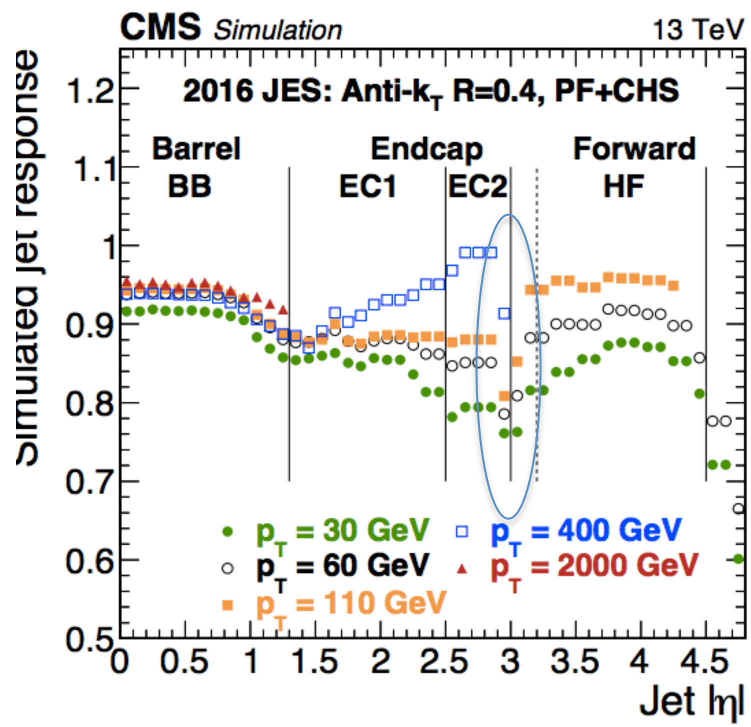
- subtract simulated PU
- subtle nonlinear effects
  - PU un-zerosuppression, shadowing
  - small dependence on jet p_T
  - first subtraction based on simulation
  - scaled by data/MC ratio of RC offset energy density



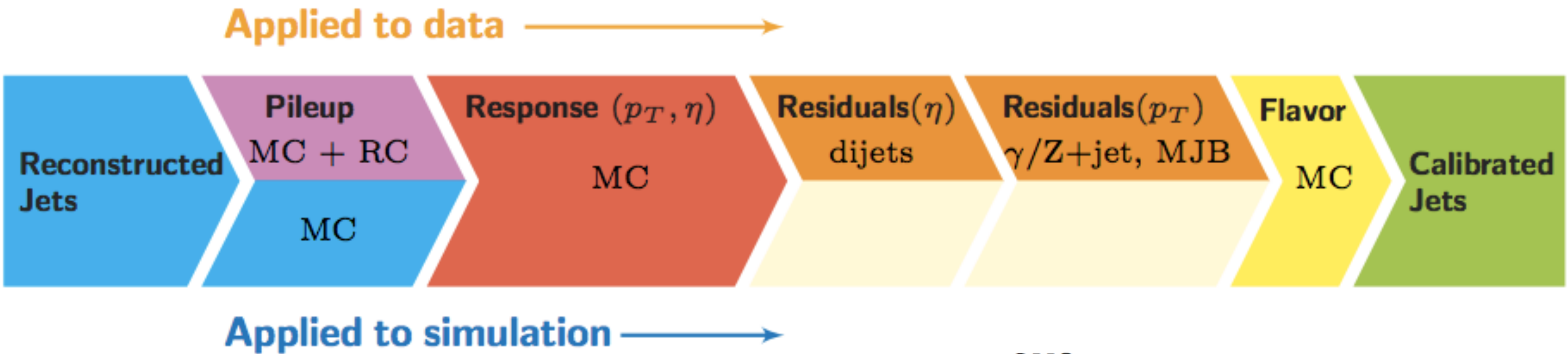
jets at 13 TeV



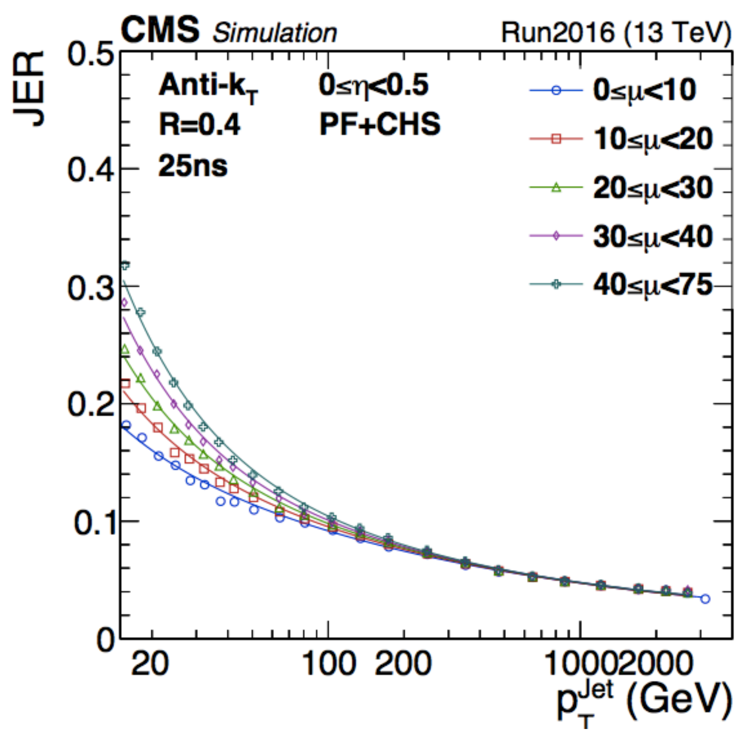
- response in simulation
  - response jumps at the end of tracker coverage and HE/HF boundary can cause eta bias and bin migration
- simulated resolution
  - rather stable over run periods



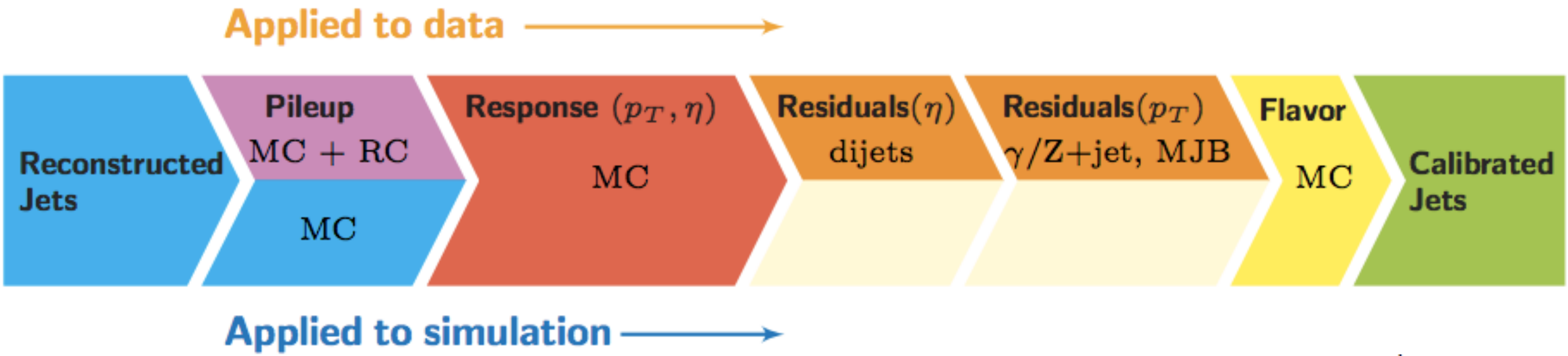
jets at 13 TeV



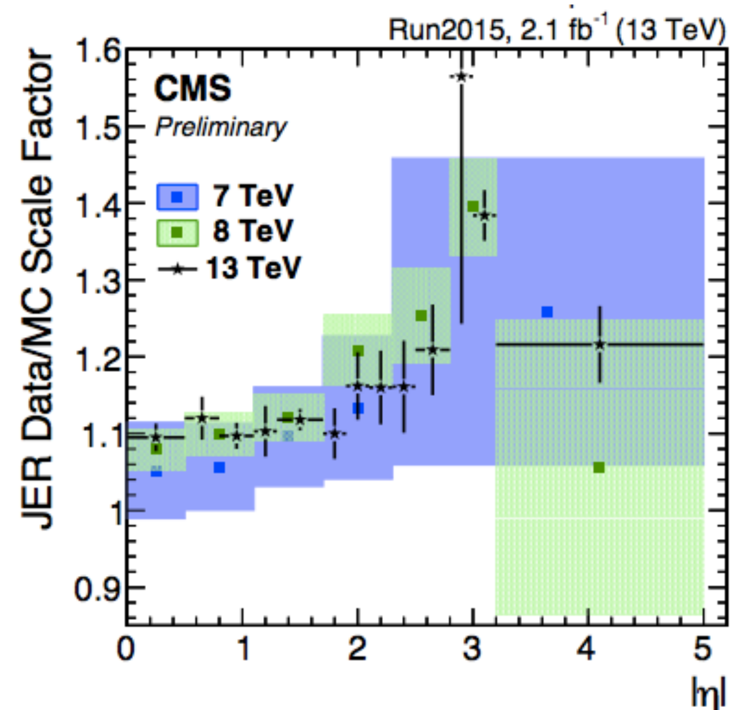
- response in simulation
  - response jumps at the end of tracker coverage and HE/HF boundary can cause eta bias and bin migration
- simulated resolution
  - rather stable over run periods



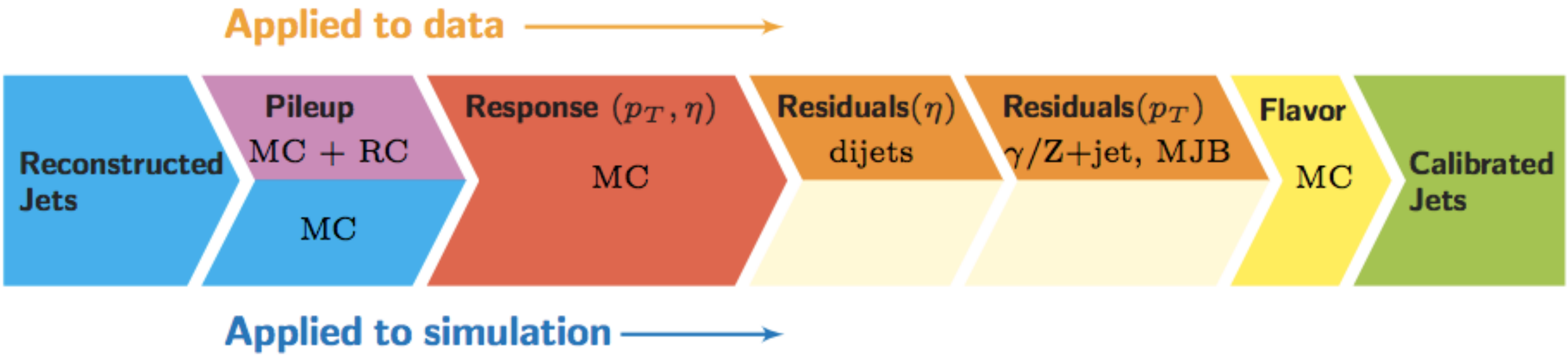
jets at 13 TeV

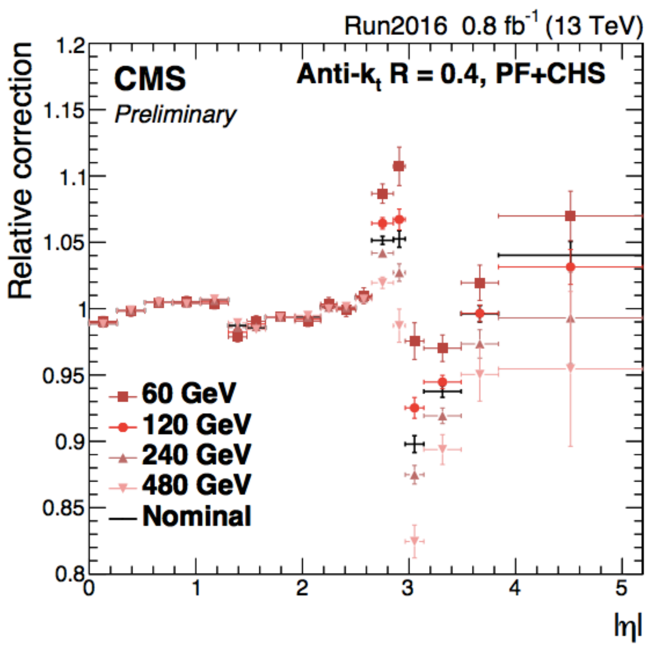


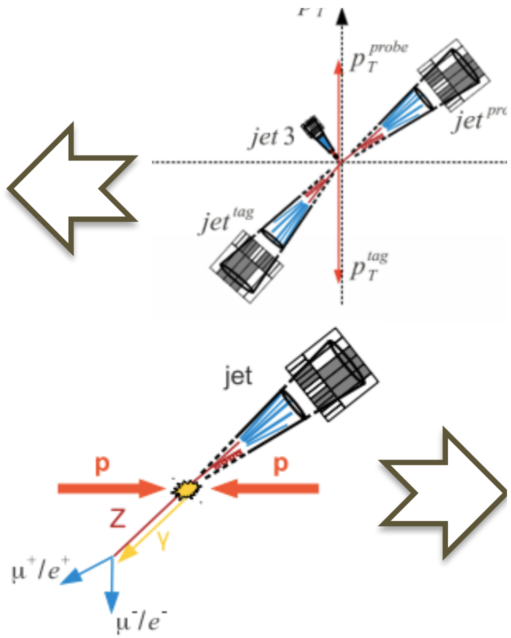
- response in simulation
  - response jumps at the end of tracker coverage and HE/HF boundary can cause eta bias and bin migration
- simulated resolution
  - rather stable over run periods

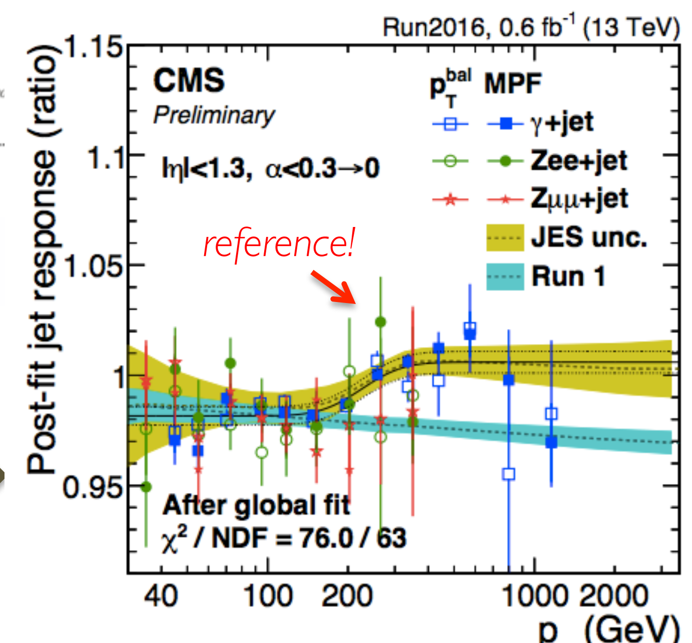


jets at 13 TeV

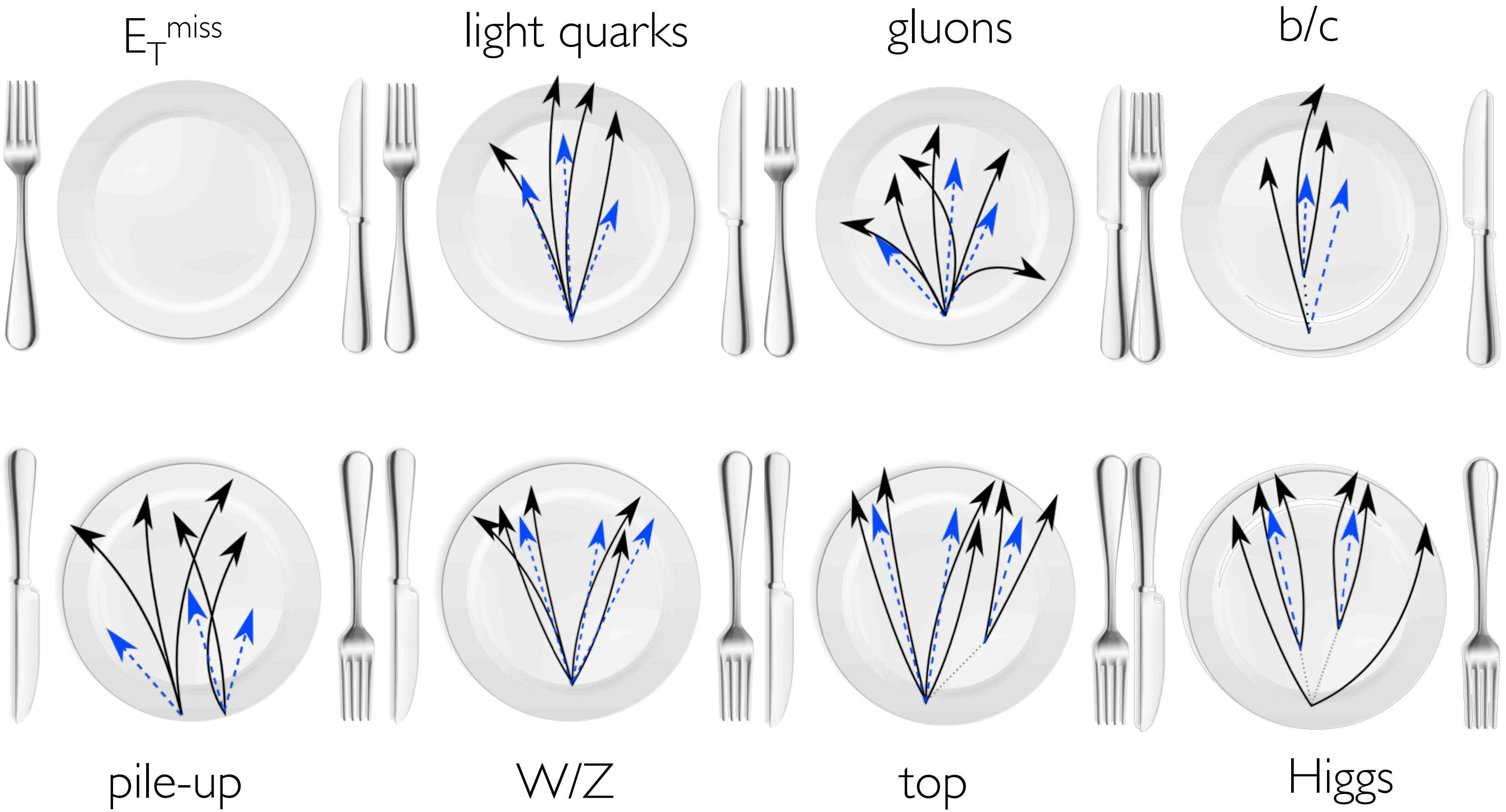






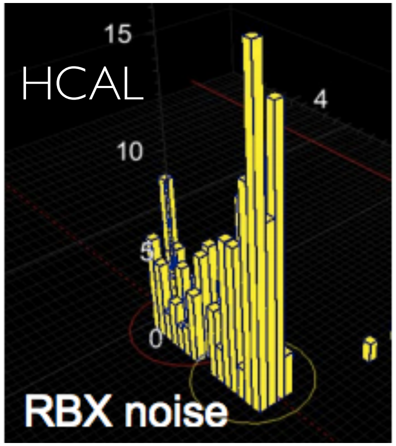


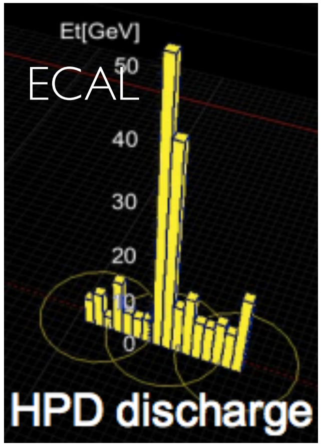
From the menu

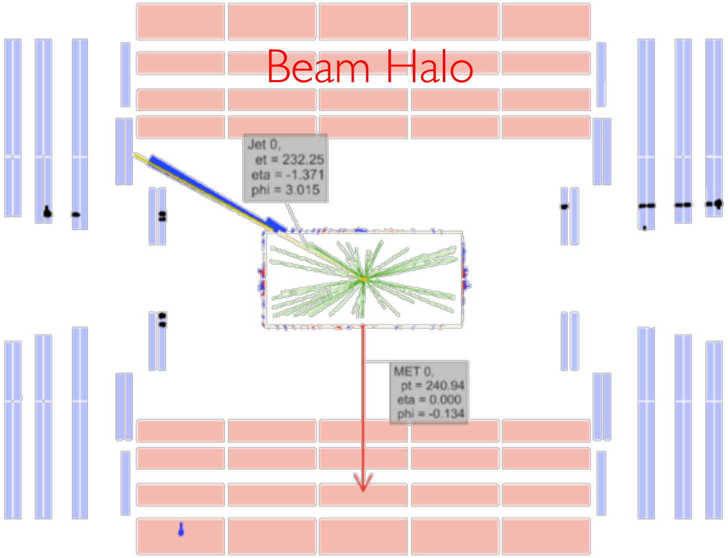


E_T^{miss} and jet tails

- very fast turn around in E_T^{miss} tails crucial for performance
- all the know problems under control
  - HBHE noise (RBX, HPD discharge, ion fb.)
  - Beam Halo, non instrumented regions
- numerous new issues found
  - ECAL gain switch at high energy
    - mismeasure a small fraction of high energy e/γ
  - tracking dynamic inefficiency
    - biases charged fraction
    - loosening tracking induced spurious muons



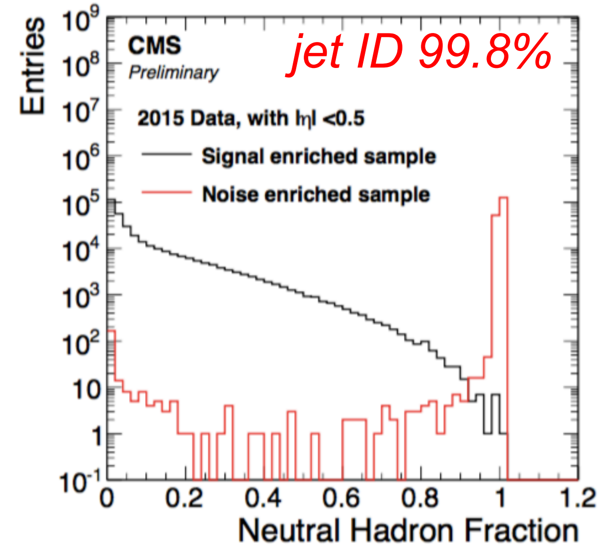


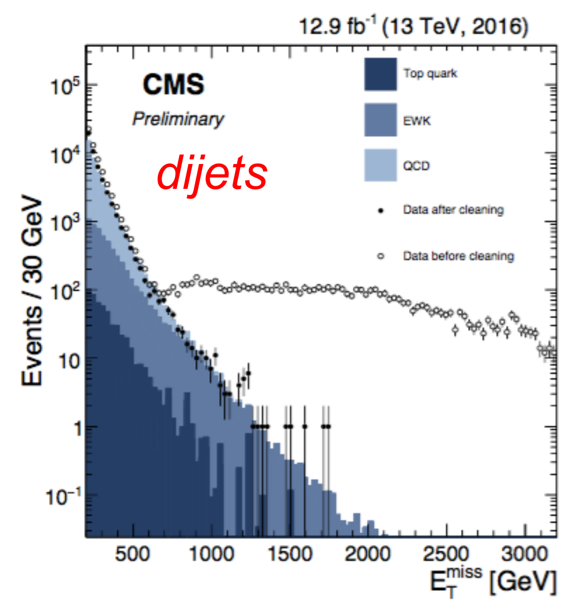


E_T^{miss} and jet tails

CMS-PAS-JME-16-003, CMS-PAS-JME-16-004

- very fast turn around in E_T^{miss} tails crucial for performance
- all the know problems under control
  - HBHE noise (RBX, HPD discharge, ion fb.)
  - Beam Halo, non instrumented regions
- numerous new issues found
  - ECAL gain switch at high energy
    - mismeasure a small fraction of high energy e/γ
  - tracking dynamic inefficiency
    - biases charged fraction
    - loosening tracking induced spurious muons
- highly efficient jet ID, clean E_T^{miss} tail
- we succeeded to deliver high quality objects on an aggressive time scale

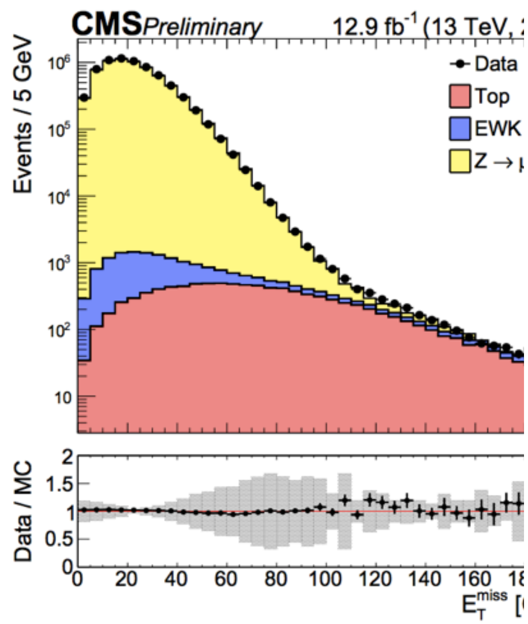


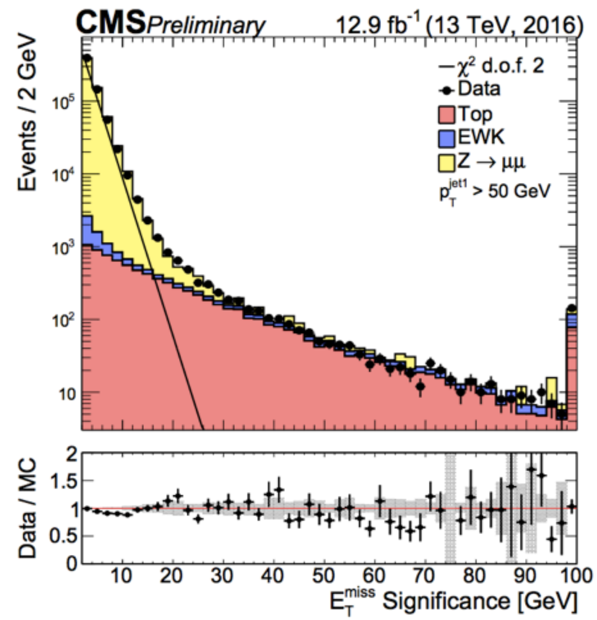


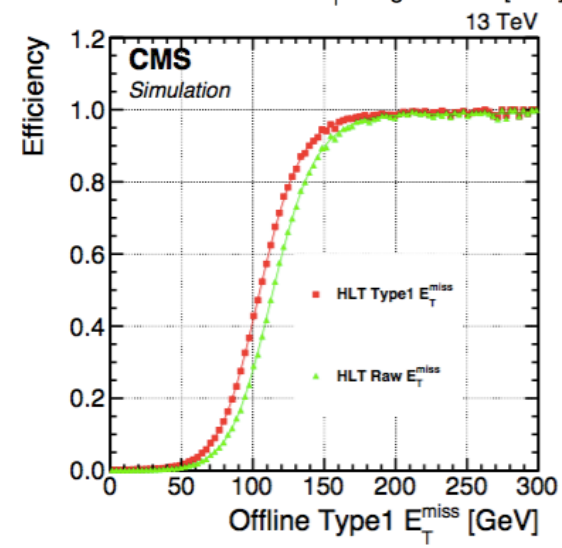
E_T^{miss} performance

CMS-PAS-JME-16-004

- bulk E_T^{miss} performance great out of the box
- directly benefited a wealth of results at ICHEP16 and Moriond17
- this is straight from reconstruction
  - sum up PF particles, add JEC
  - no further corrections
  - JEC corrected E_T^{miss} now running at HLT

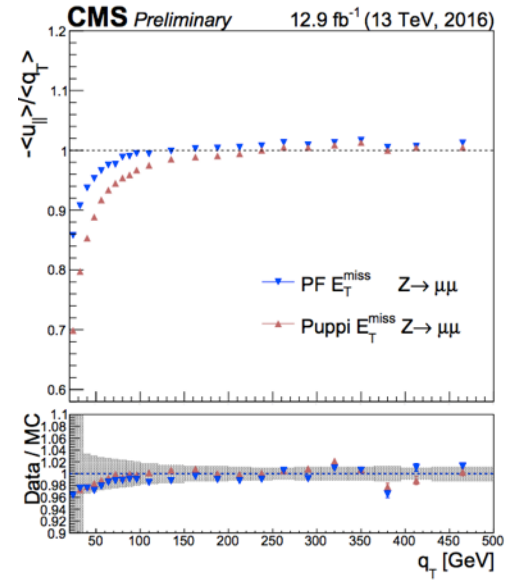


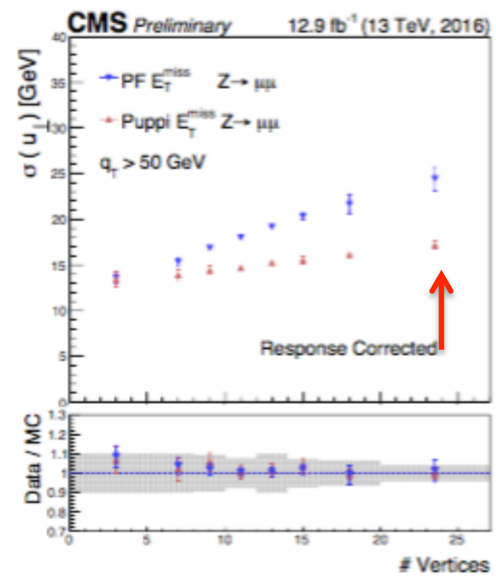




E_T^{miss} performance

- scale and resolution under control
- battling PU: commission puppi
  - Define $$\alpha_i = \log \sum_{\substack{j \in Ch,PV \\ j \neq i}} \left(\frac{p_{Tj}}{\Delta R_{ij}}\right)^2 \Theta(R_0 - \Delta R_{ij})$$ which encodes the PU-probability of a particle
  - distribution of **α** is *measured* using charged component in each event and *applied* to the neutrals. Extrapolate to outside of TRK.
  - reweight neutrals according to PU probability $$w_i \cdot (p_{Ti} > (A + B \cdot n_{PV}))$$
  - Take care with isolated high p_T photons & leptons
  - Makes optimal use of PF particle level
- It pays off: global event interpretation with PU subtraction within reach!





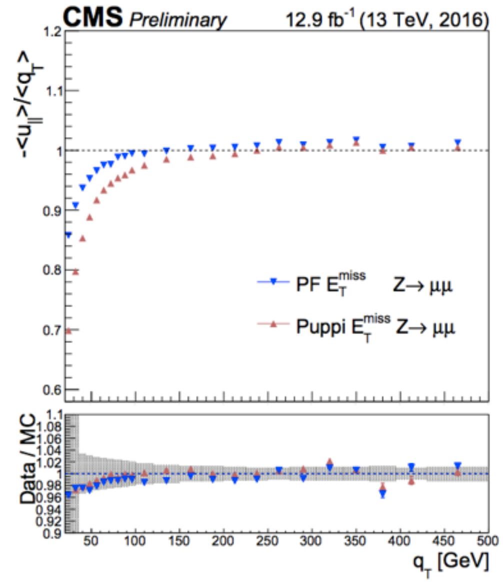
E_T^{miss} performance

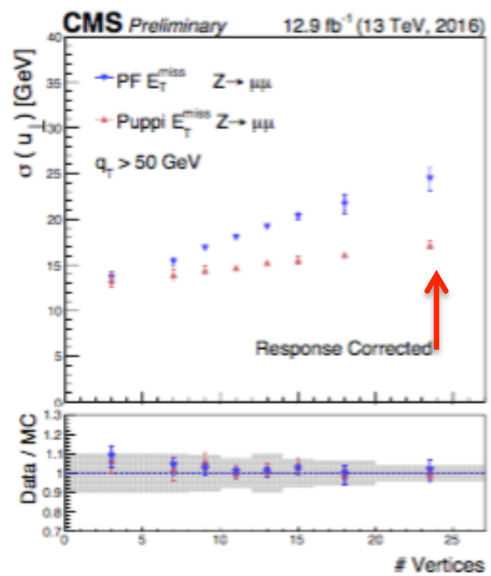
CMS-PAS-JME-16-004

- Advantages from redundancy in algorithms

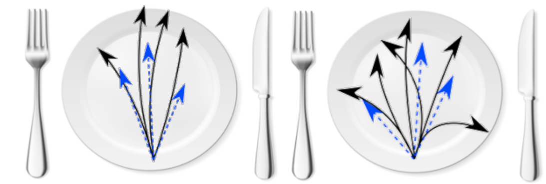
$$w_i \cdot (p_{Ti} > (A + B \cdot n_{PV}))$$

  - dynamic thresholds for neutrals are important tuning parameters
    - effectively, a coarse model of neutral had. resolution vs. PU
  - can this be refined?
    - dominant systematic is unclustered energy
    - balance insensitivity wrt. PU vs. tails
    - should improve MET significance
  - revisit/refine CHS at high PU



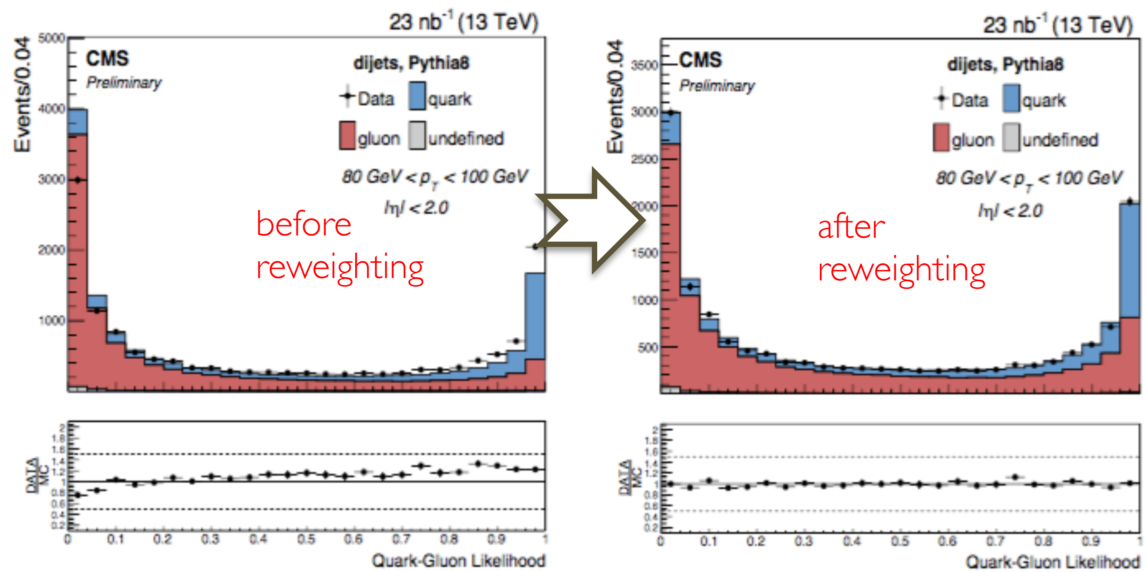


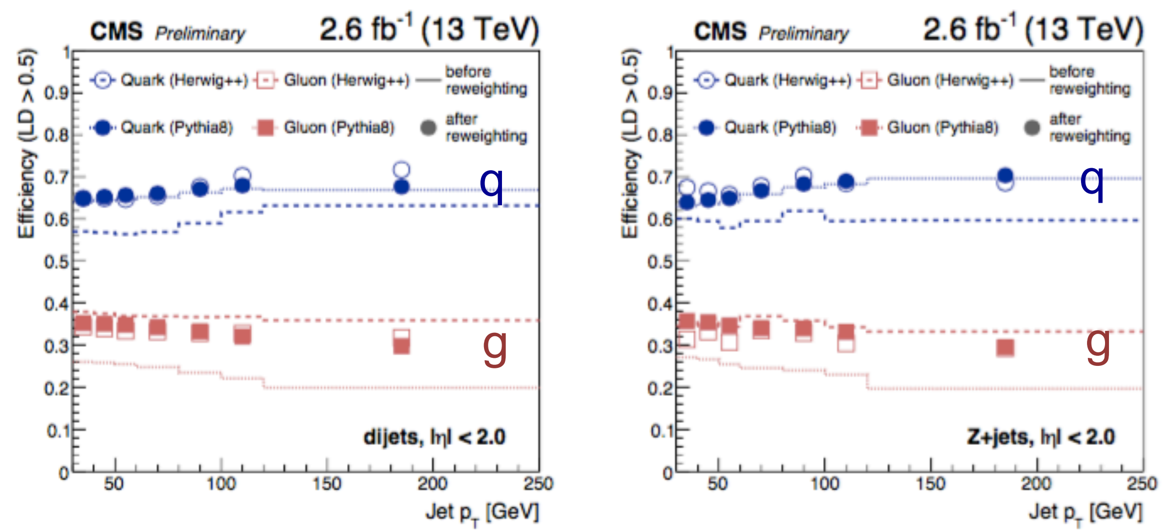
q/g tagging likelihood



CMS-PAS-JME-16-003

- simple log likelihood of ch. mult, σ_2, $p_{T,D}$
- since a long time used in analysis
- Herwig++/Pythia8 shower model differ for gluon jet properties
- reweighting in dijet/ Z+jets CR reduces sys by a factor 10

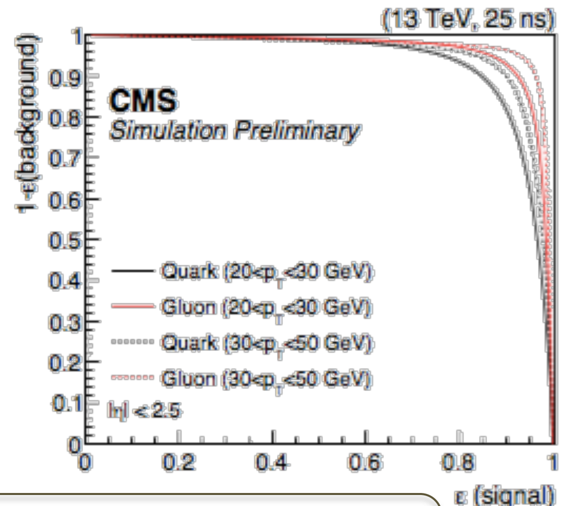


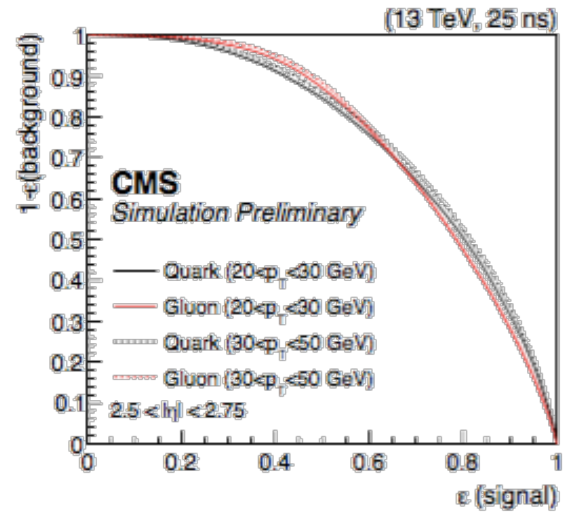


PU Jet ID: BDT

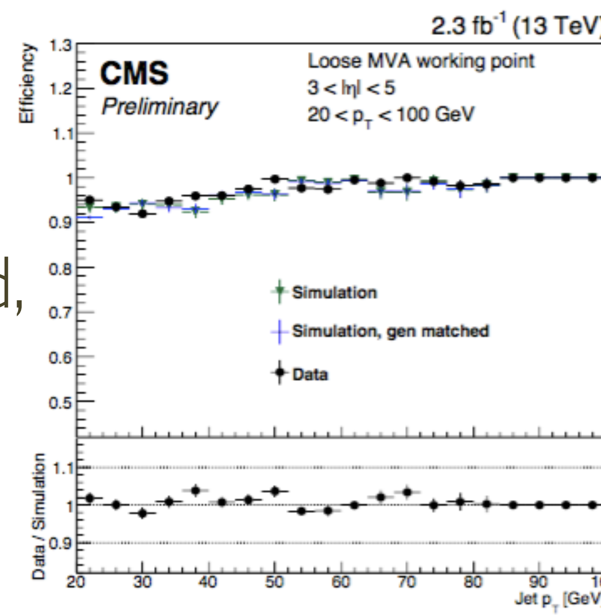
CMS-PAS-JME-16-003

- run-II main changes
  - AK5 → AK4
  - charge hadron subtr.
- most PU jets within TRK already removed





$$\beta = \frac{\sum_{i\in PV} p_{Ti}}{\sum_i p_{Ti}}, \quad p_T^{lead}/p_T^{jet}, \; |\vec{m}|, \; N_{total}, \; N_{charged},$$
$$N_{vertices}, \; \langle \Delta R^2 \rangle, \; f_{ring0}, \; f_{ring1}, \; f_{ring2}, \; f_{ring3},$$

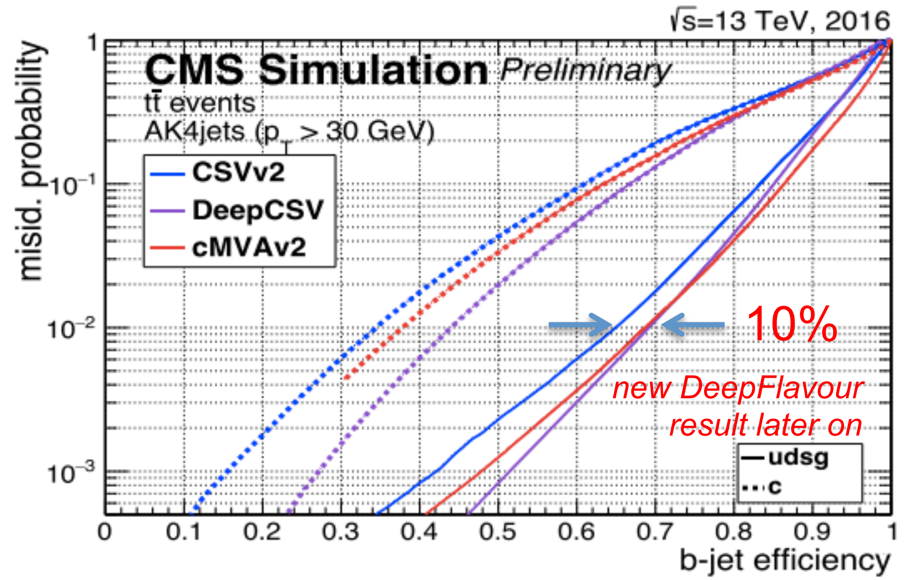


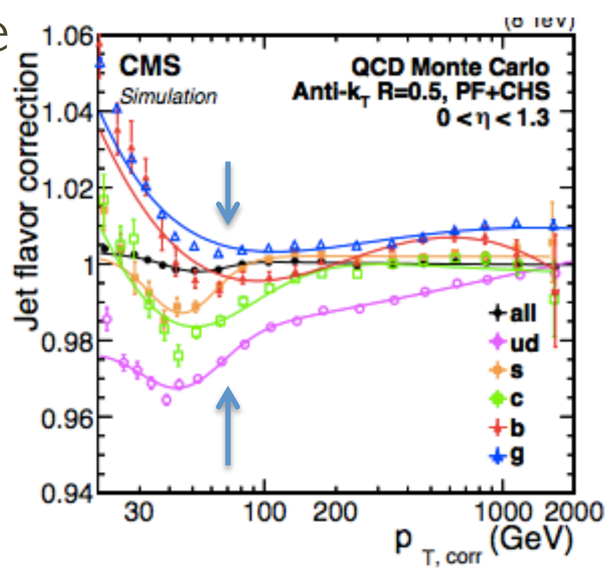
- BDT with 14 variables. Well understood, well calibrated. Is it optimal?
  - use observables from larger cone size
  - should really profit from ML gold rush

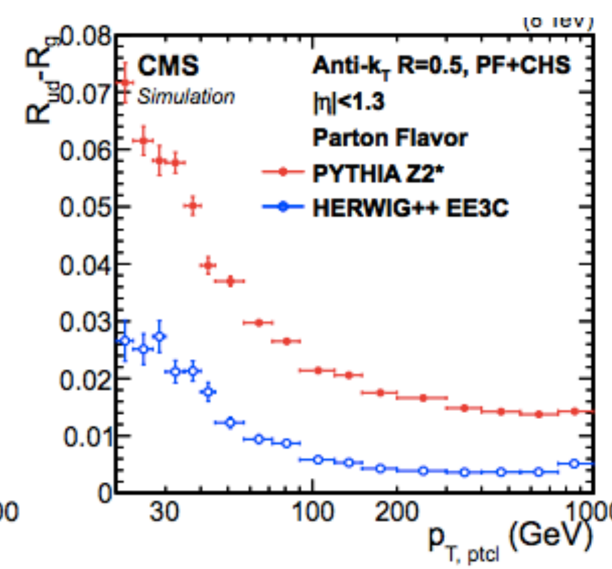
machine learning

CMS DP 2017-005

- 10% gain with DNN (DeepCSV) shown in 2016
  - uses PF candidates
  - do same for q/g and PU Jet Id?
- flavor response regression
  - reduce flavor response and generator difference
  - same physics!
- NNs offer multi-classification
  - train g/uds/c/b/PU
  - or W/Z/H/t ?

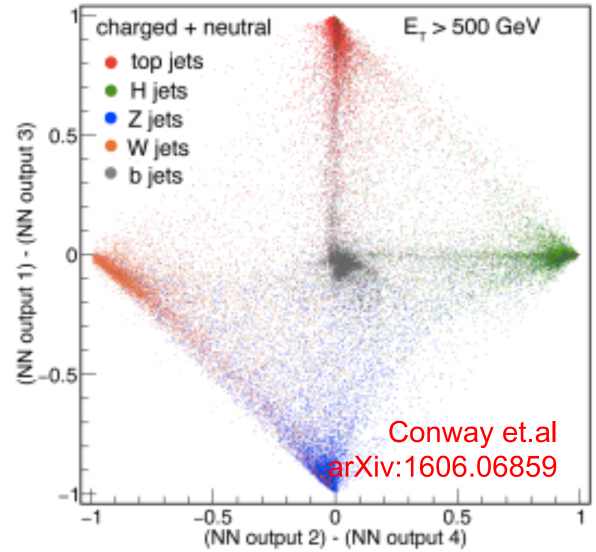


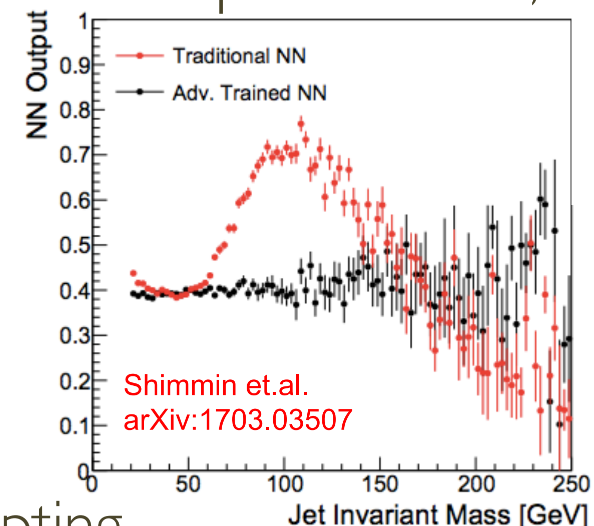


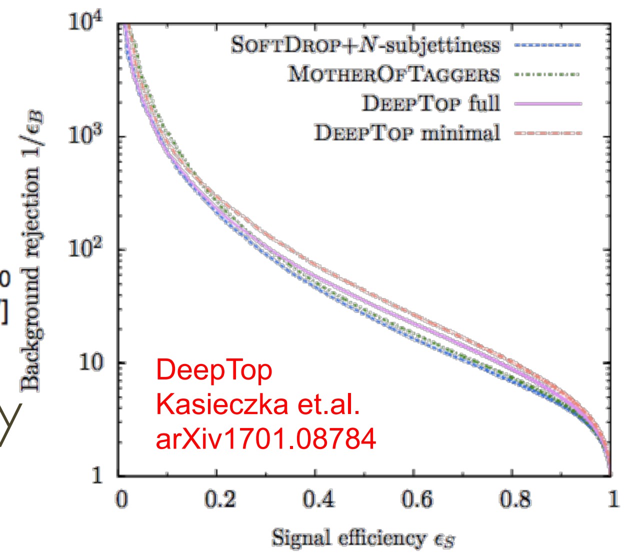


machine learning, biased view

- multiclassification considered seriously
- DeepTop: use jet images in deep NN
  - another example of raw input
- real deal: stability wrt to experimental, modelling unc., PU etc.
- adversarial NN: decorrelate taggers
  - force flat efficiency to reduce mass sculpting
  - trade some performance for gain in stability





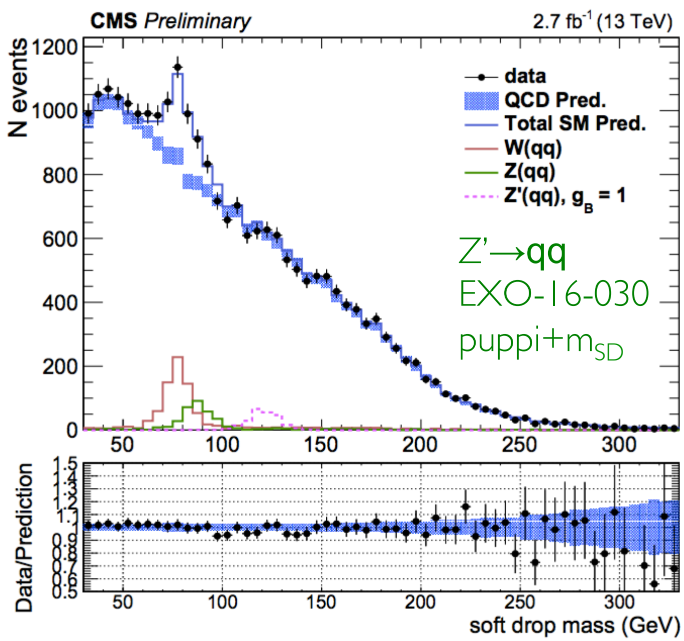


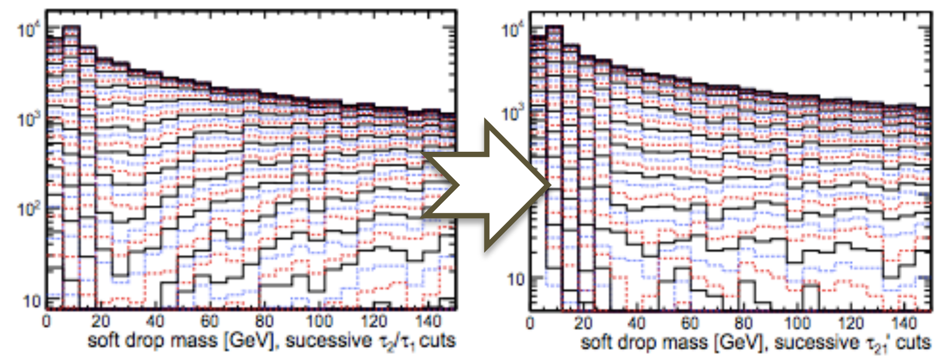
decorrelated substructure

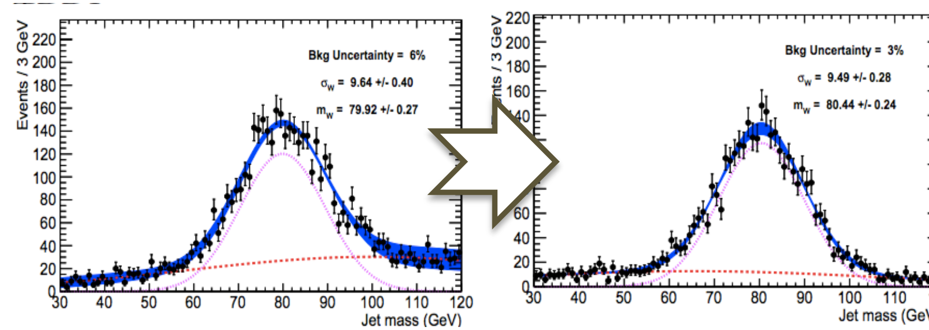
J.Dolen et.al., arXiv:1603.00027, CMS-PAS-EXO-16-030

- 2/1-subjettiness ratio excellent for finding 2-prong structure
- some correlation with m_{SD}, remove with simple reparametrization

$$\tau_{21}^{\rm DDT} = \tau_2/\tau_1 - M \times \log\left(\frac{m^2}{p_T \mu}\right)$$





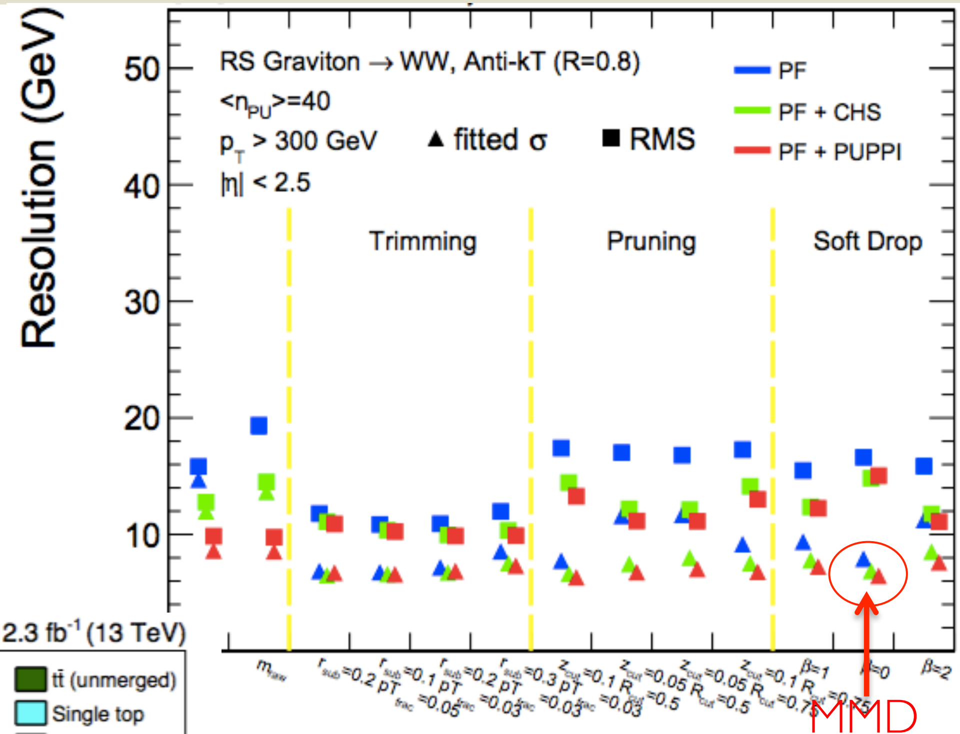


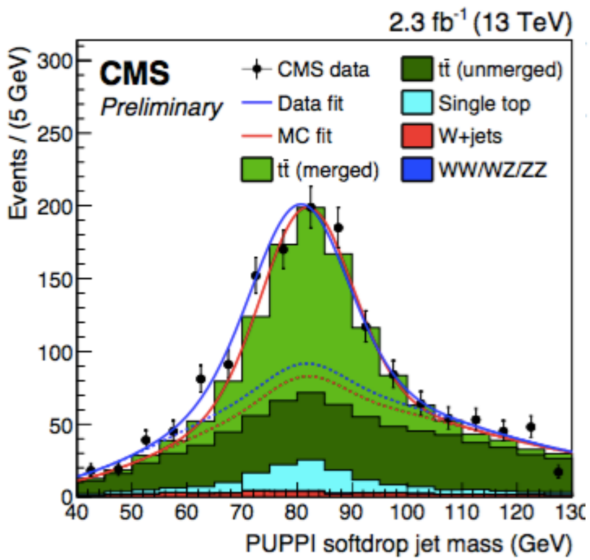
- shall be happy with simple, physics motivated trafo, if no MVA gain

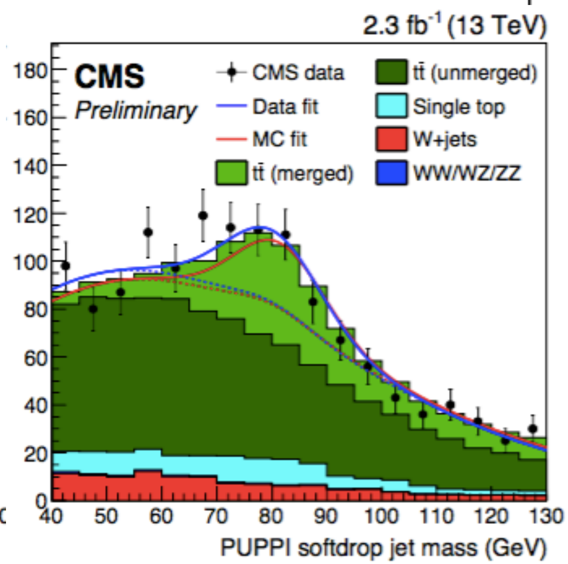
tagging boosted W/Z

CMS-EXO-PAS-16-030

- extensive MC studies of groomers and PU mitigation in Run-1, LS1 and until now
- AK8 PF+CHS/Puppi jets and SD(β =0) (i.e. MMD) with τ_{21} is CMS default







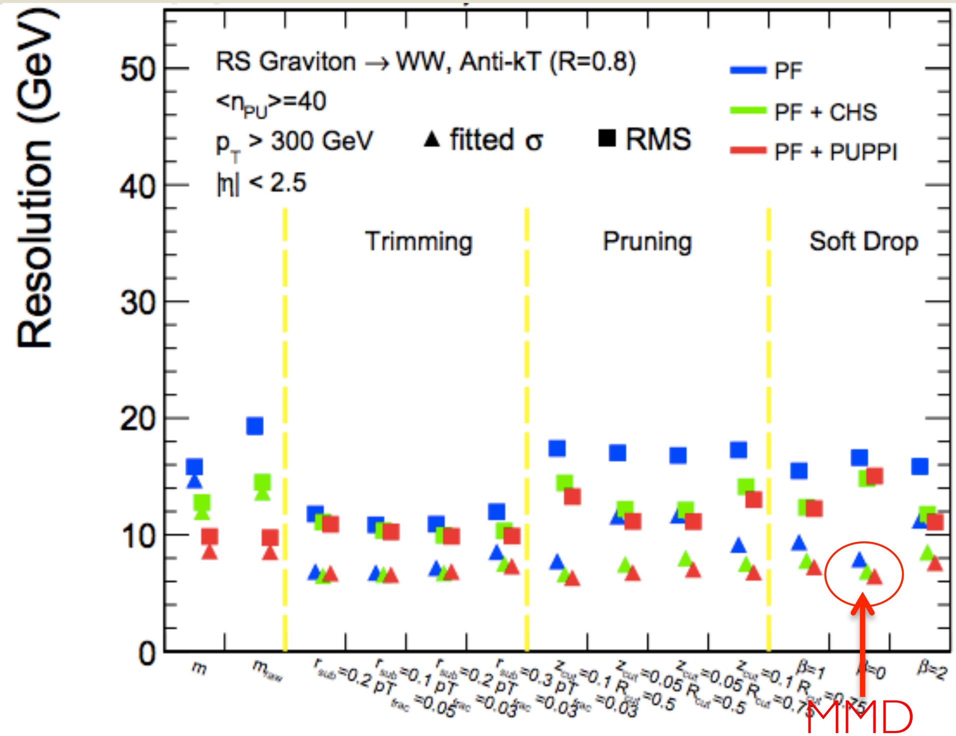
… and studied in data!

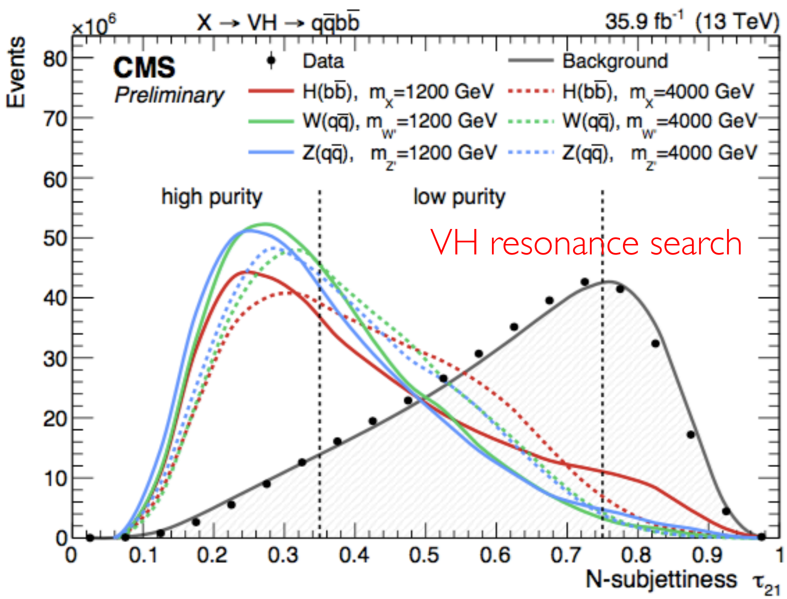
- O(10%) SF measured routinely
- factorized, universal tools, e.g. H(bb) tagger

tagging boosted W/Z

CMS-PAS-B2G-17-002

- extensive MC studies of groomers and PU mitigation in Run-1, LS1 and until now
- AK8 PF+CHS/Puppi jets and SD(β =0) (i.e. MMD) with τ_{21} is CMS default





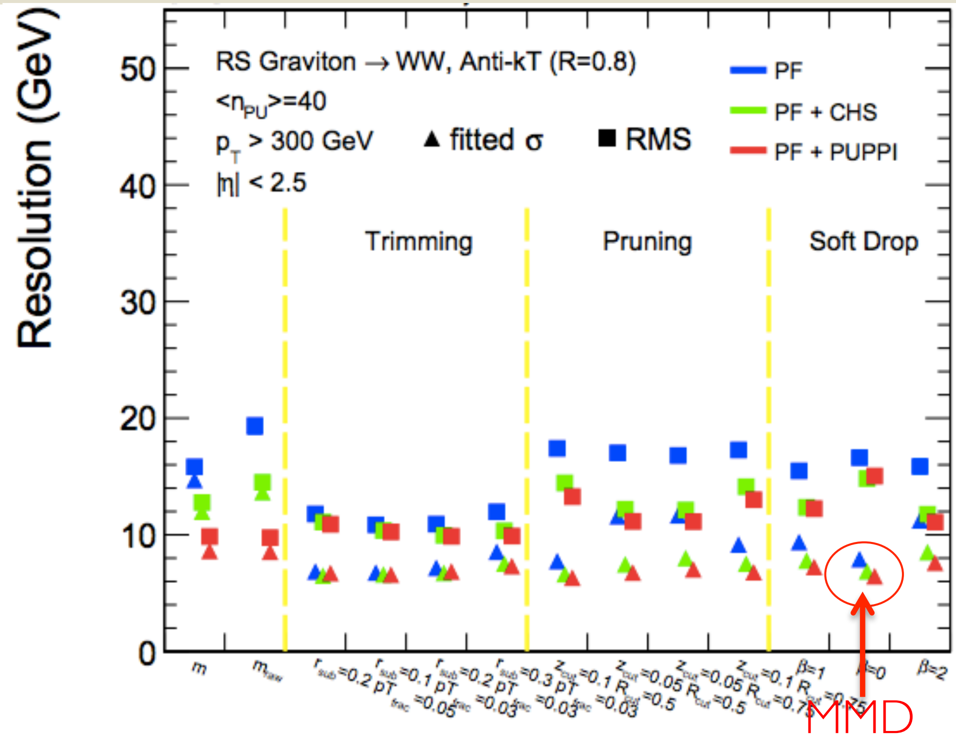
… and studied in data!

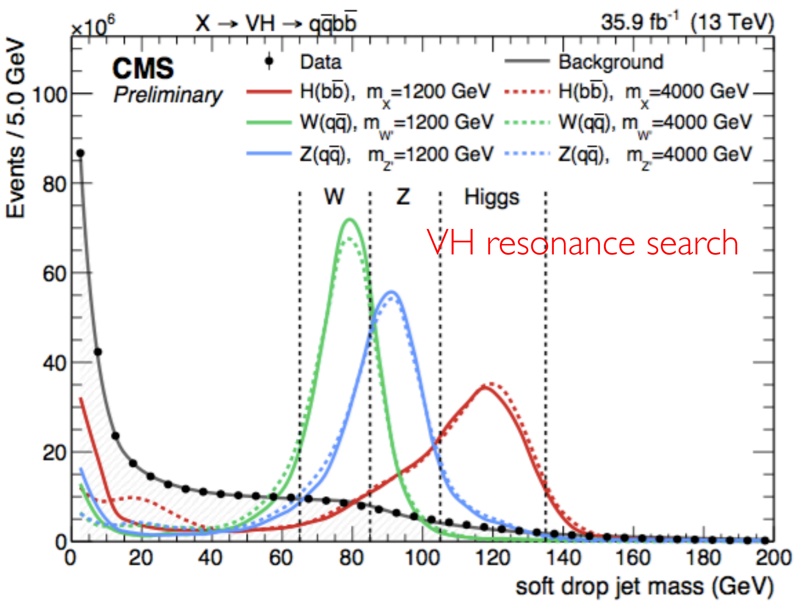
- O(10%) SF measured routinely
- factorized, universal tools, e.g. H(bb) tagger

tagging boosted W/Z

CMS-PAS-B2G-17-002

- extensive MC studies of groomers and PU mitigation in Run-1, LS1 and until now
- AK8 PF+CHS/Puppi jets and SD(β =0) (i.e. MMD) with τ_{21} is CMS default



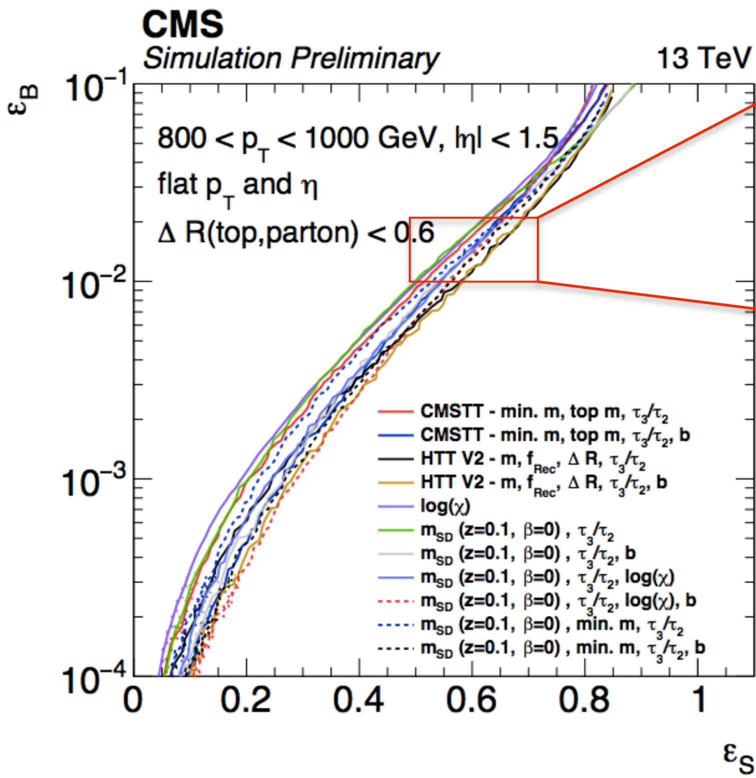


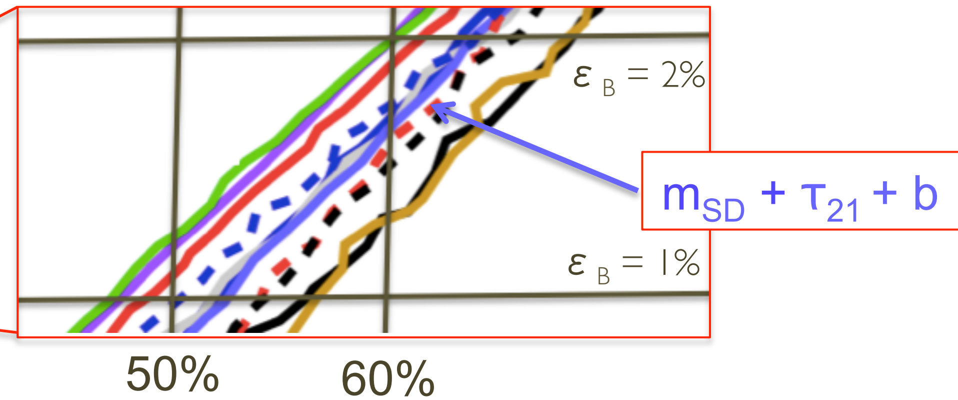
… and studied in data!

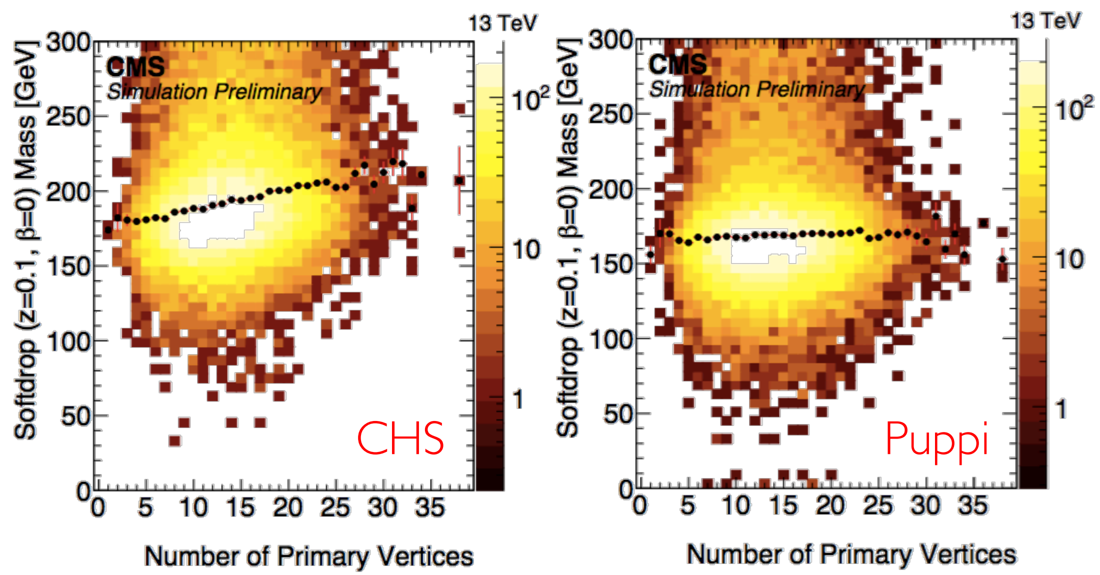
- O(10%) SF measured routinely
- factorized, universal tools, e.g. H(bb) tagger

boosted top tagging

CMS-PAS-JME-15-002, CMS-DP-2015/034







- systematic comparison no dramatic differences of combined taggers
- ‘safe’ default: $m_{SD} + \tau_{21} + b$
- cut-based or in BDT. Stability vs PU, jet p_T, and subjet b-tag efficiency

how did we get here?

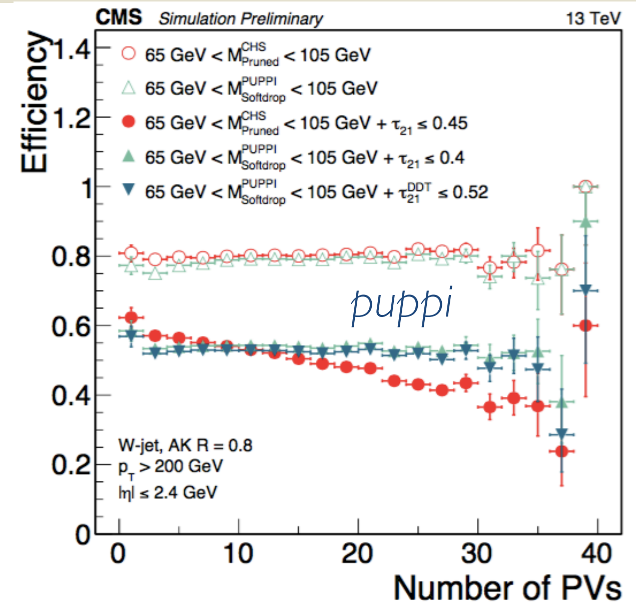
CMS-PAS-JME-16-003, CMS-PAS-JME-14-002

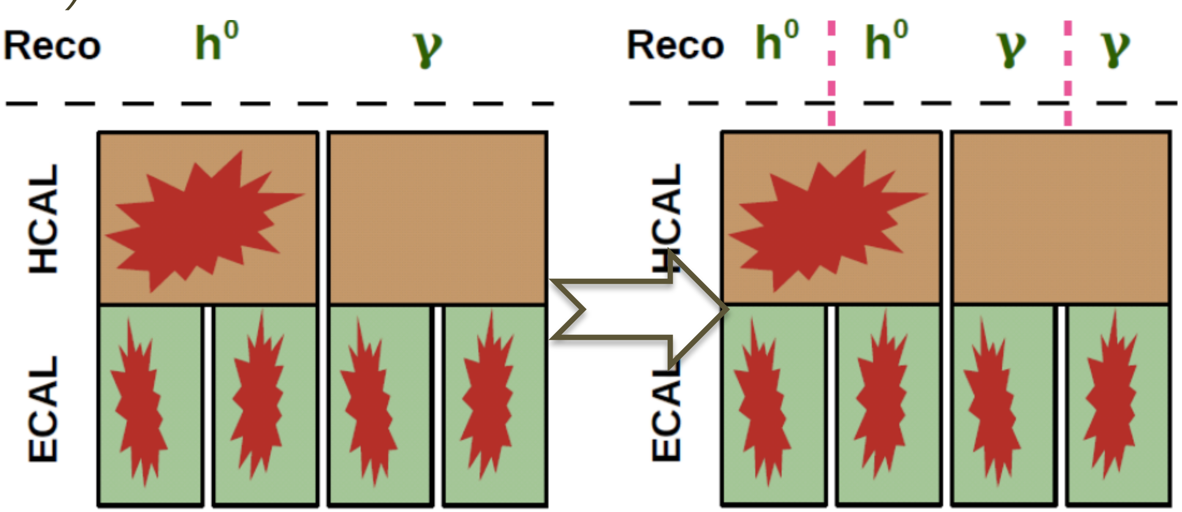
a) we can subtract PU locally

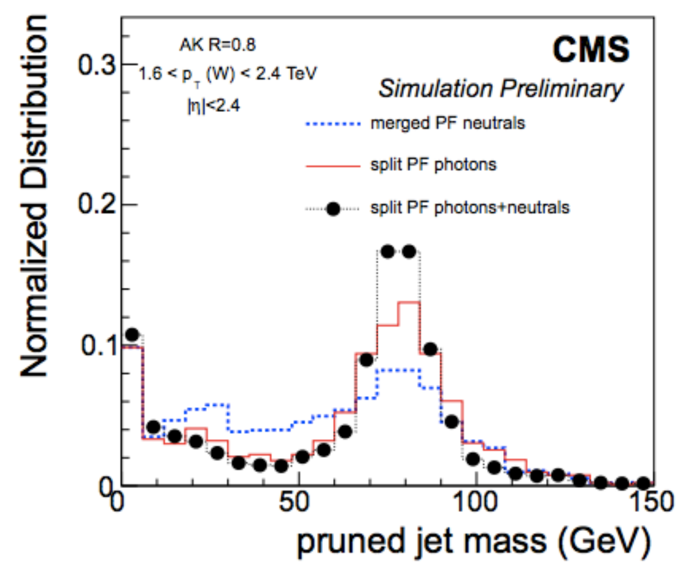
b) finely tuned local reconstruction

- dedicated jet core tracking: recover soft tracks in dense cores
- pixel cluster splitting: reconstruct overlapping tracks

c) dedicated PF for substructure







how did we get here?

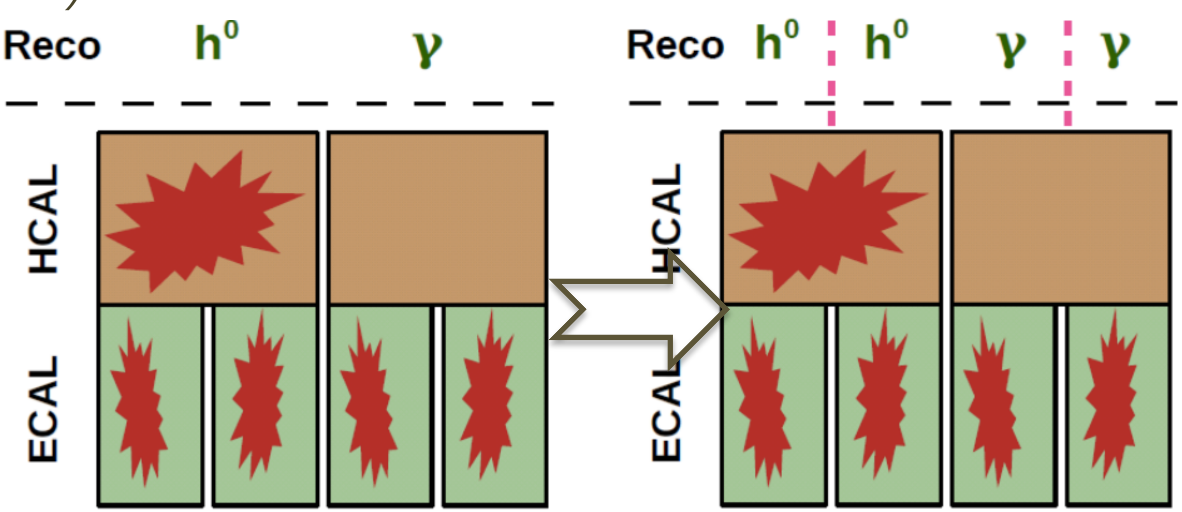
CMS-PAS-JME-16-003, CMS-PAS-JME-14-002

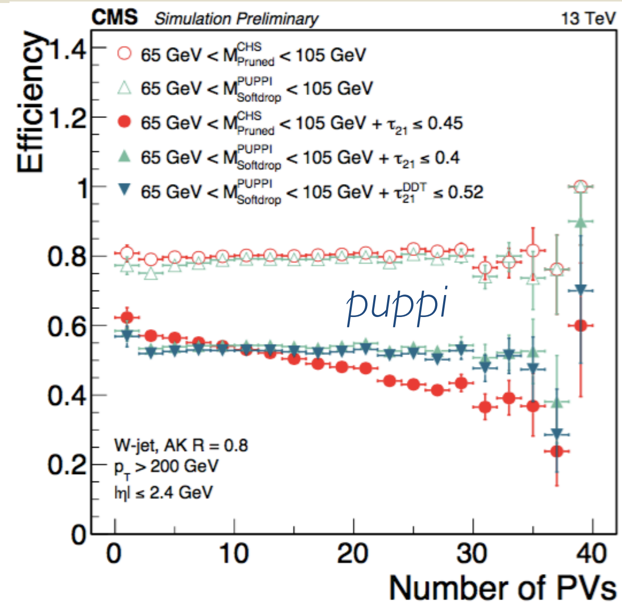
a) we can subtract PU locally

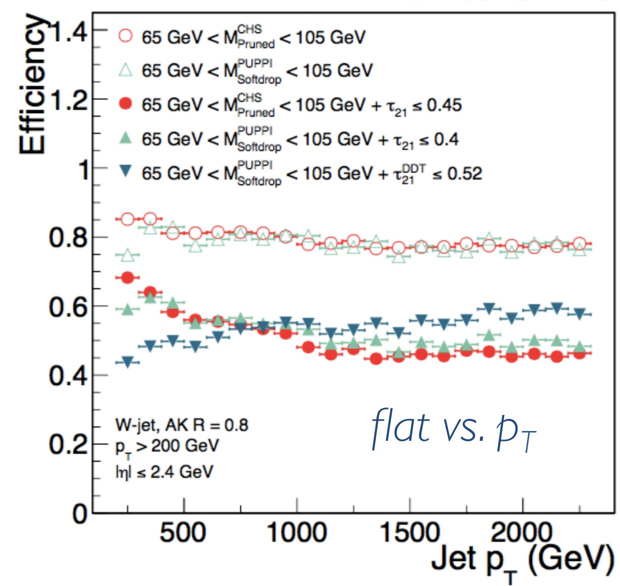
b) finely tuned local reconstruction

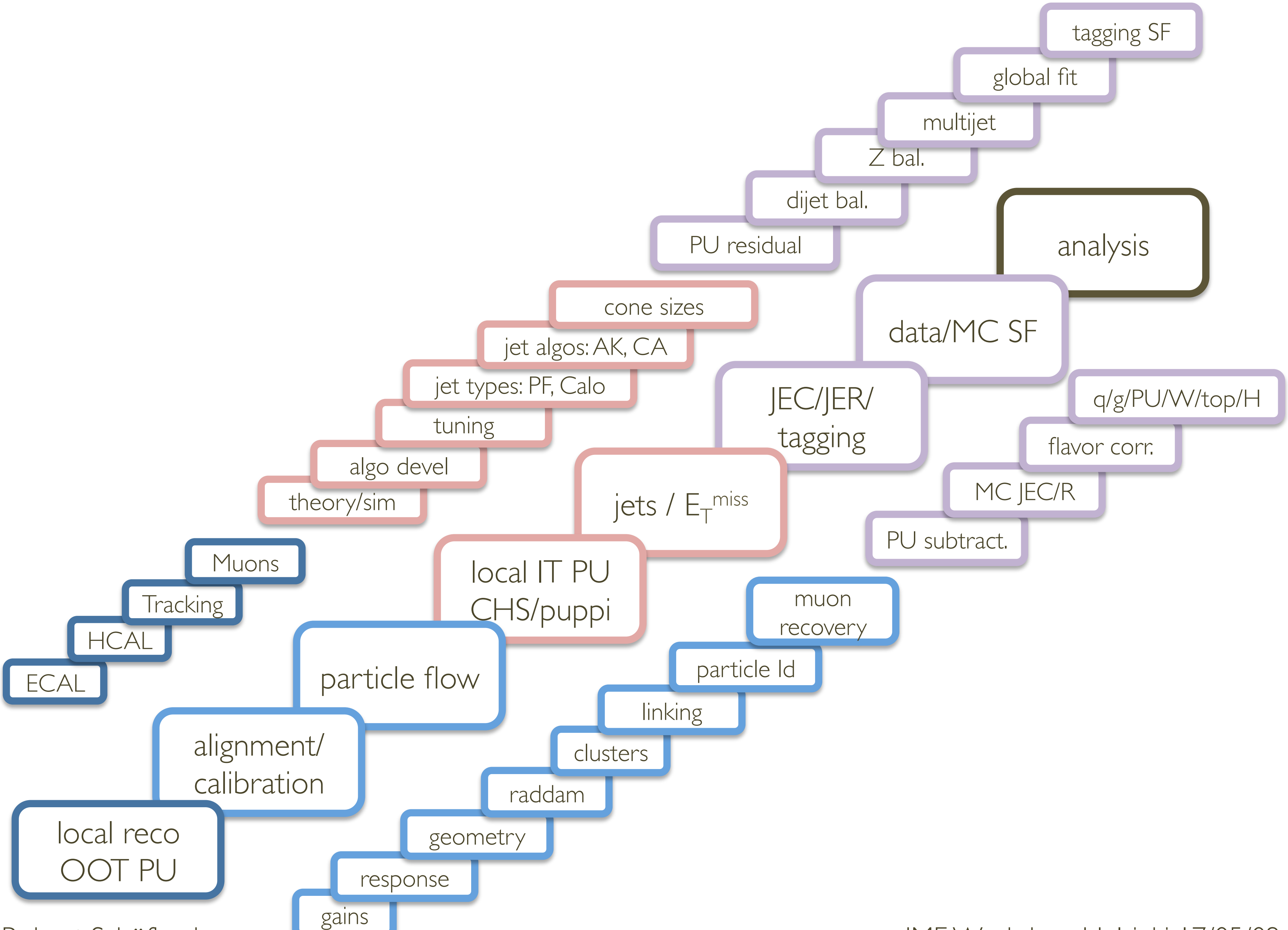
- dedicated jet core tracking: recover soft tracks in dense cores
- pixel cluster splitting: reconstruct overlapping tracks

c) dedicated PF for substructure

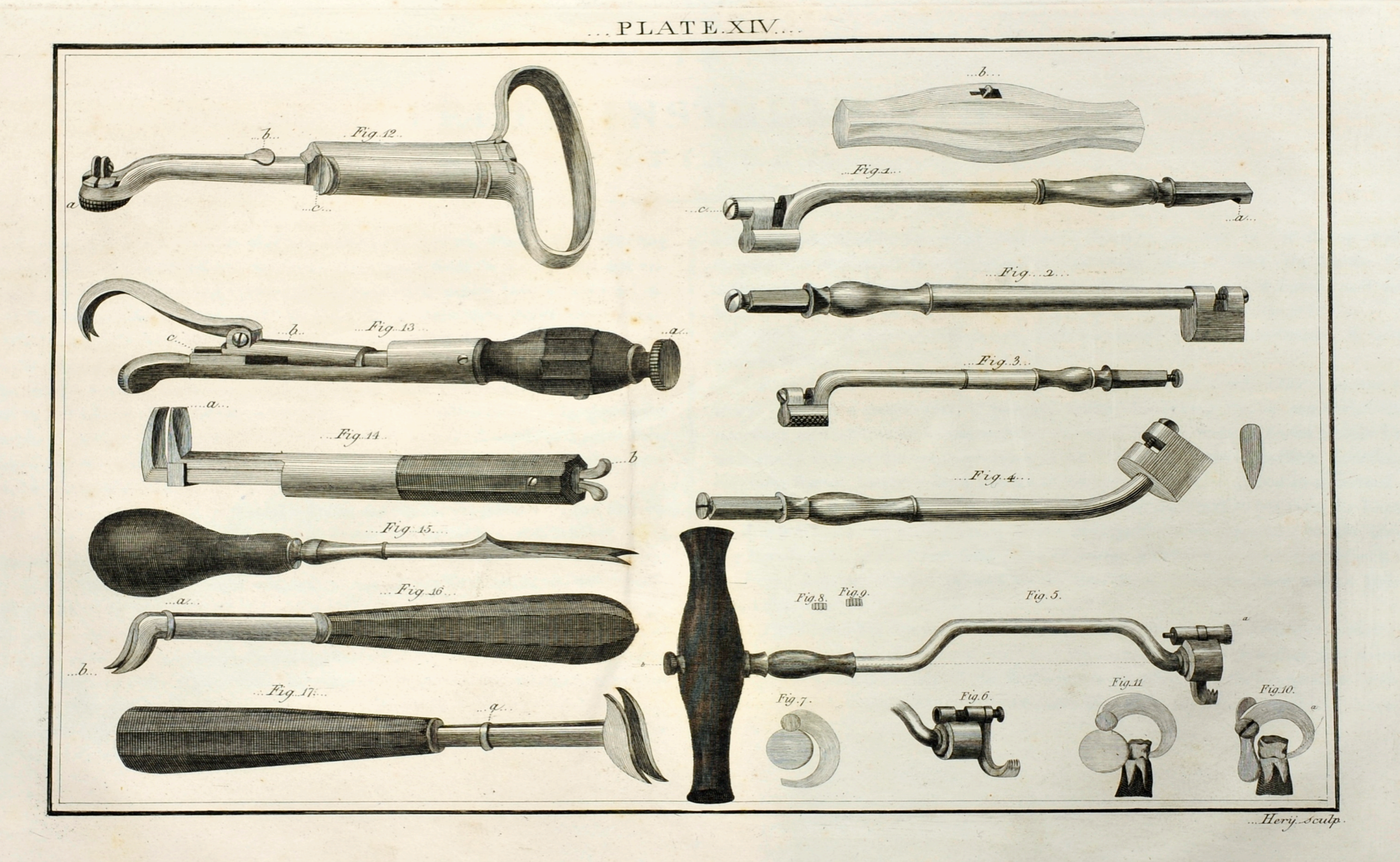






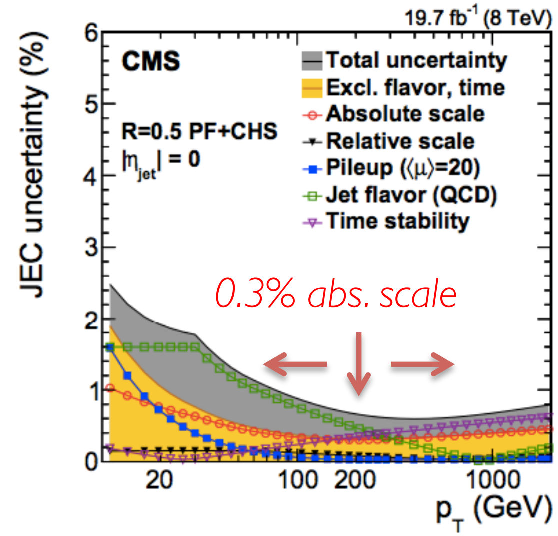

tagging SF
global fit
multijet
Z bal.
dijet bal.
PU residual
analysis
data/MC SF
q/g/PU/W/top/H
flavor corr.
MC JEC/R
PU subtract.
JEC/JER/
tagging
cone sizes
jet algos: AK, CA
jet types: PF, Calo
tuning
algo devel
theory/sim
jets / E_T^{miss}
local IT PU
CHS/puppi
Muons
Tracking
HCAL
ECAL
particle flow
muon
recovery
particle Id
linking
clusters
raddam
geometry
response
gains
alignment/
calibration
local reco
OOT PU

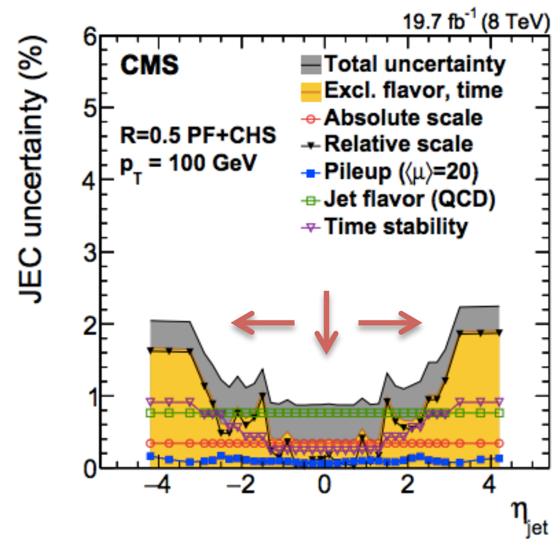
precision devices



jet energy accuracy in Run-II

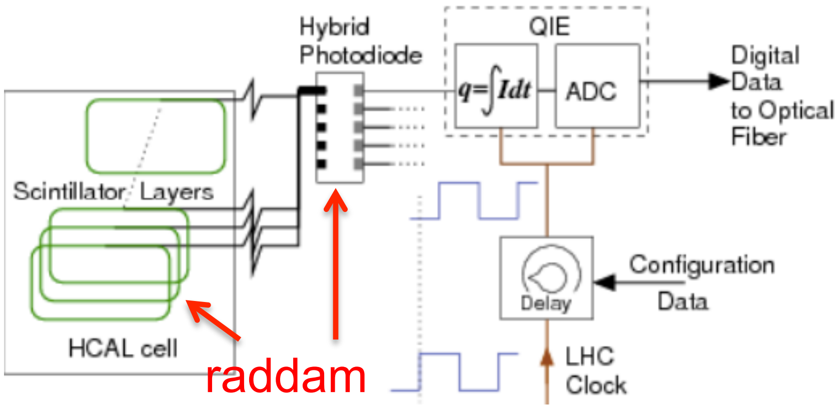
- Run-1: min. 0.32% abs. scale uncertainty
  - Factor 10 more Z!
  - 0.1% at Run-II?!
- Optimal calibration at working point
  - 200 GeV at η=0 and 20% gluons
  - then move away and control systematics
- need to 'fix everything' to get even close



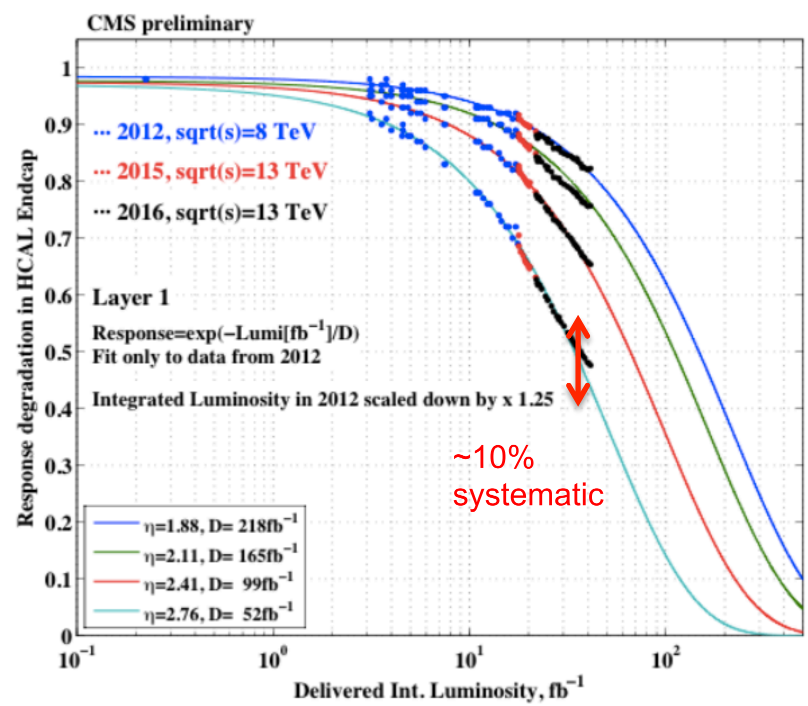


low level systematics

Jay Lawhorn, US LHC user mtg. 2016, CERN-CMS-DP-2016-052



- both, HCal scintillators and HPDs degrade when irradiated
- total raddam well understood



- systematic in the raddam. peculiar pattern in jet response PU/eta/E/time
- parameterize response variation and constrain nuisance in JEC global fit
- need similar approach for other subdets

reference objects

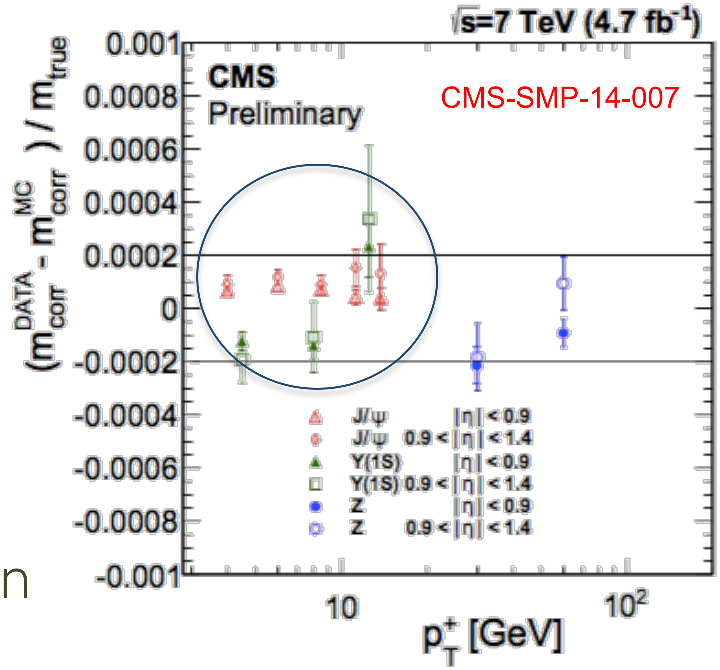
CMS-SMP-14-007, arXiv:1502.02702

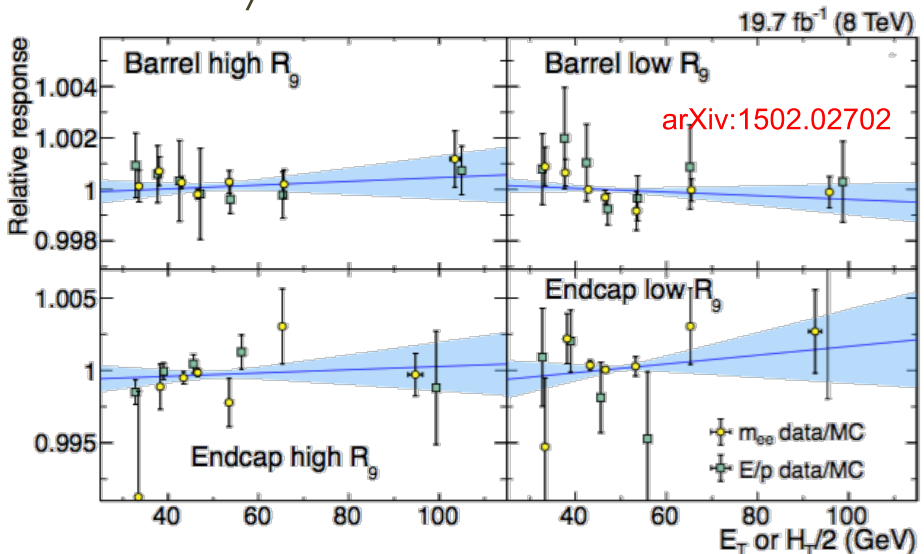
- Make use of analysis work
  - Muon calibration from 'W like mass' curvature corrections

$$k^c = (A-1)k + qM + \frac{k}{1+k\epsilon\sin\theta}$$

go down to 2 x 10^{-4}.

  - impressive closure test on Z
  - likely to hold in TRK dominated region

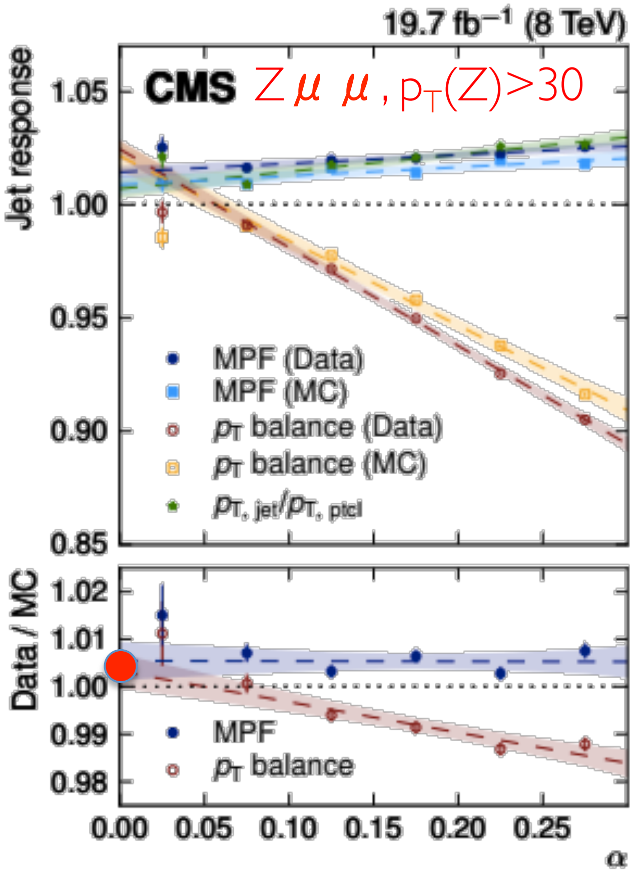




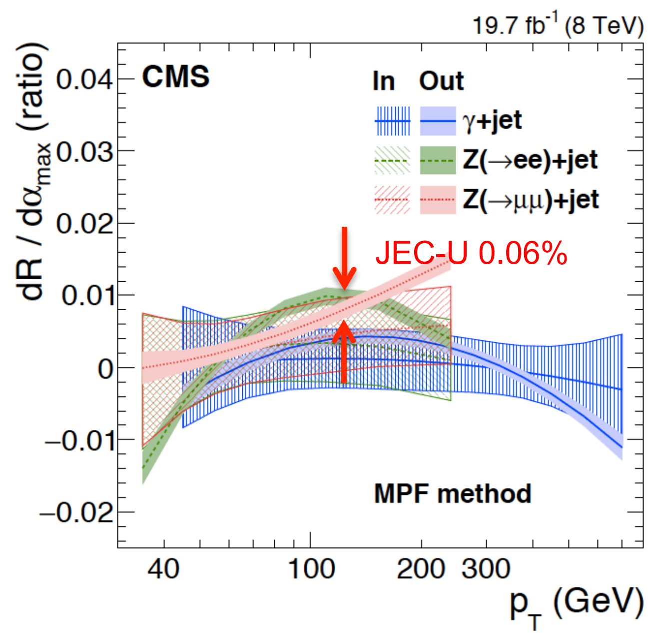
photons: m_H and Z'(ee) searches

- Run-1 photon scale Run-1 at 10^{-3} at m_Z/2.
- 2016 may suffer from gain switch issue, now quote ~2% (~1%) at 1 TeV in EB/EE

accuracy in balancing



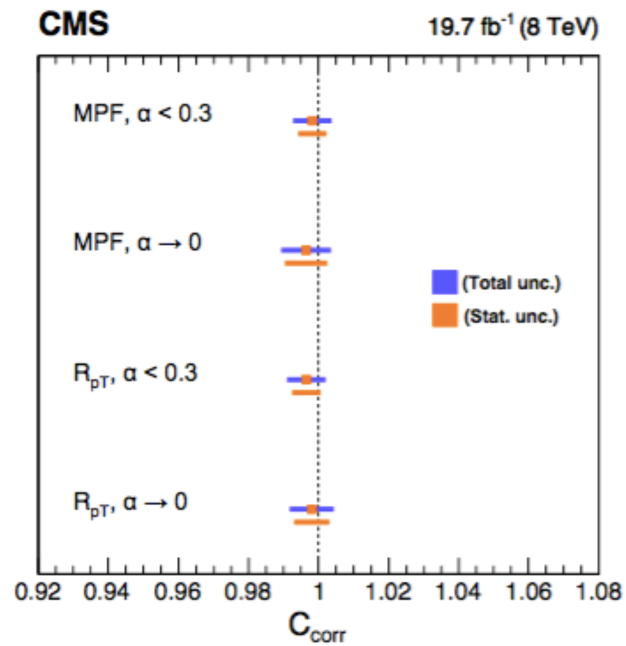
- extrapolation to zero extra activity
  - single point from global fit now
  - extend to p_T bins for each sample
  - can narrow down to 10^{-3} level?
  - jets beyond HF: gen comparisons, long. profiles

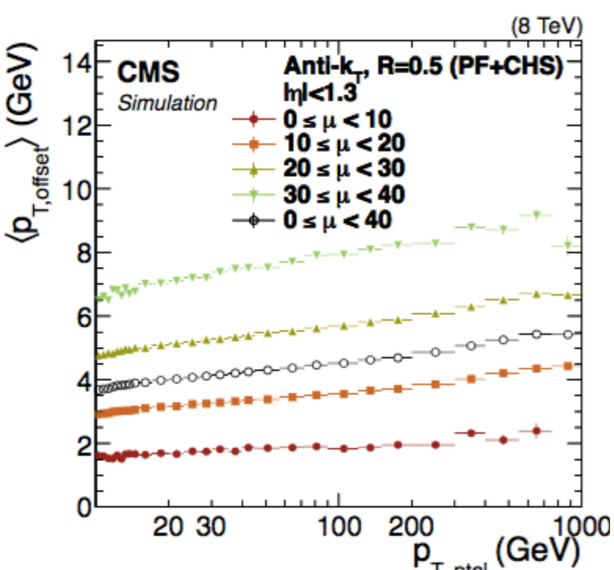


- dijet balancing
  - global fit of MPF and $p_T$bal
  - improve lever arm by adding Z's between 30-70, include $|\eta|>1.3$
  - properly parameterize uncertainties, constraints from PF composition?

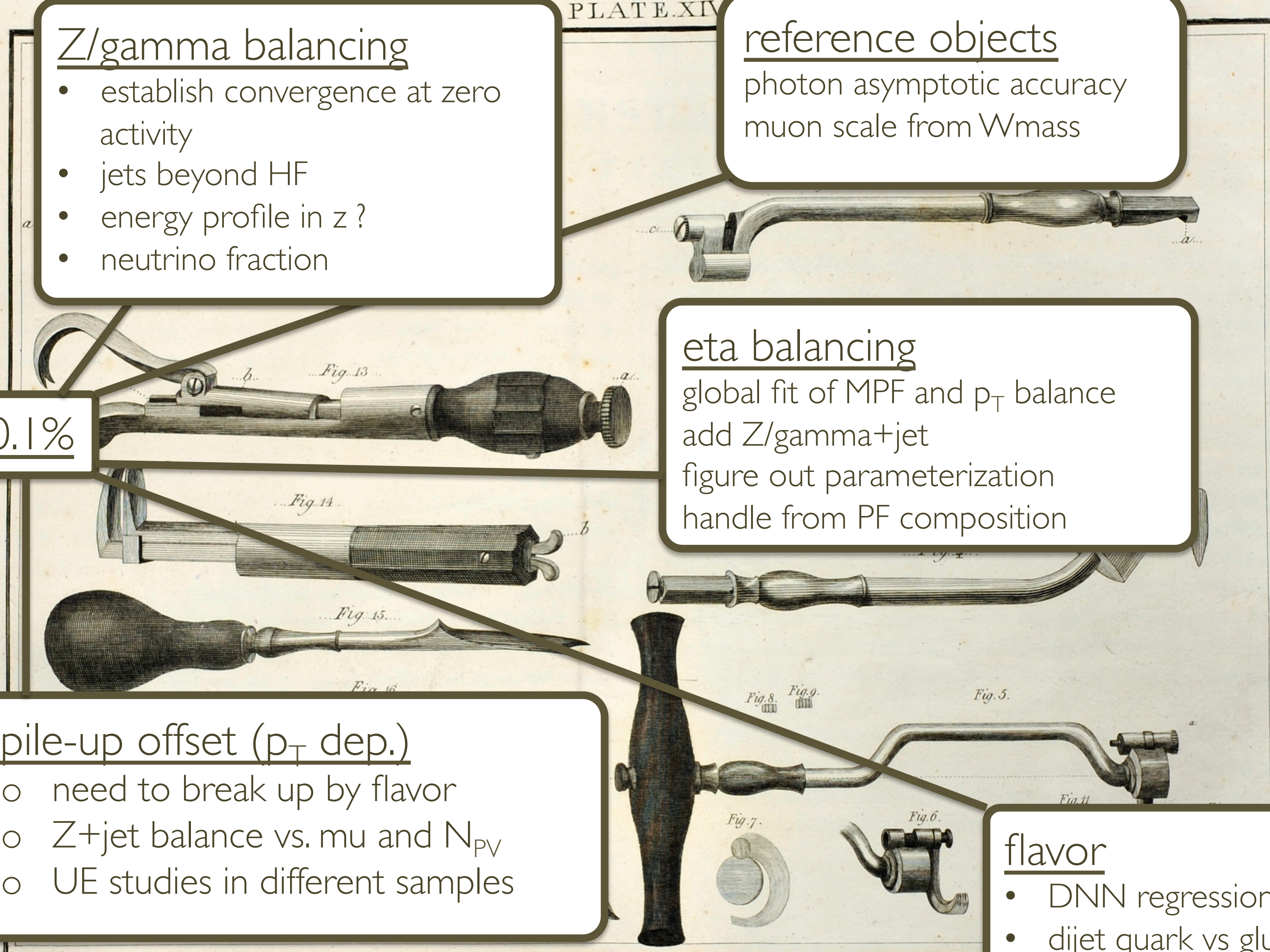
reducing flavor and PU offset

- gluon fraction ~20% in both g+jet and Z+jet
  - yet, need gluon scale for incl. jets and b-jet scale for top mass
- gluon/quark response difference in dijets with q-tag?
- Redo Z+b balancing: 0.4%/sqrt(10) ~ 0.12%
- improve MC flavor response with DNN





- Pile up offset
  - origin jet p_T dependence unclear
  - handle in data not easy
  - need to break up by flavor
  - Z+jet balance vs. mu and N_{PV}
  - UE studies in different samples



Z/gamma balancing

- establish convergence at zero activity
- jets beyond HF
- energy profile in z ?
- neutrino fraction

reference objects

photon asymptotic accuracy
muon scale from Wmass

0.1%

eta balancing

global fit of MPF and p_T balance
add Z/gamma+jet
figure out parameterization
handle from PF composition

pile-up offset (p_T dep.)

- need to break up by flavor
- Z+jet balance vs. mu and N_{PV}
- UE studies in different samples

flavor

- DNN regression
- dijet quark vs gluon
- Z+b response

Summary

- It's not about m_t
  - accurate jet energy calibration will benefit everything by making sure we walk the calibration ladder all the way *down*
- Let us
  - organize & benefit from ML developments
  - continue to move offline procedures to online
  - continue to monitor performance with a fast turnaround and fix the tails quickly
- I wish good luck and all the fun for Seema, Zeynep and the rest of the team!

particle flow

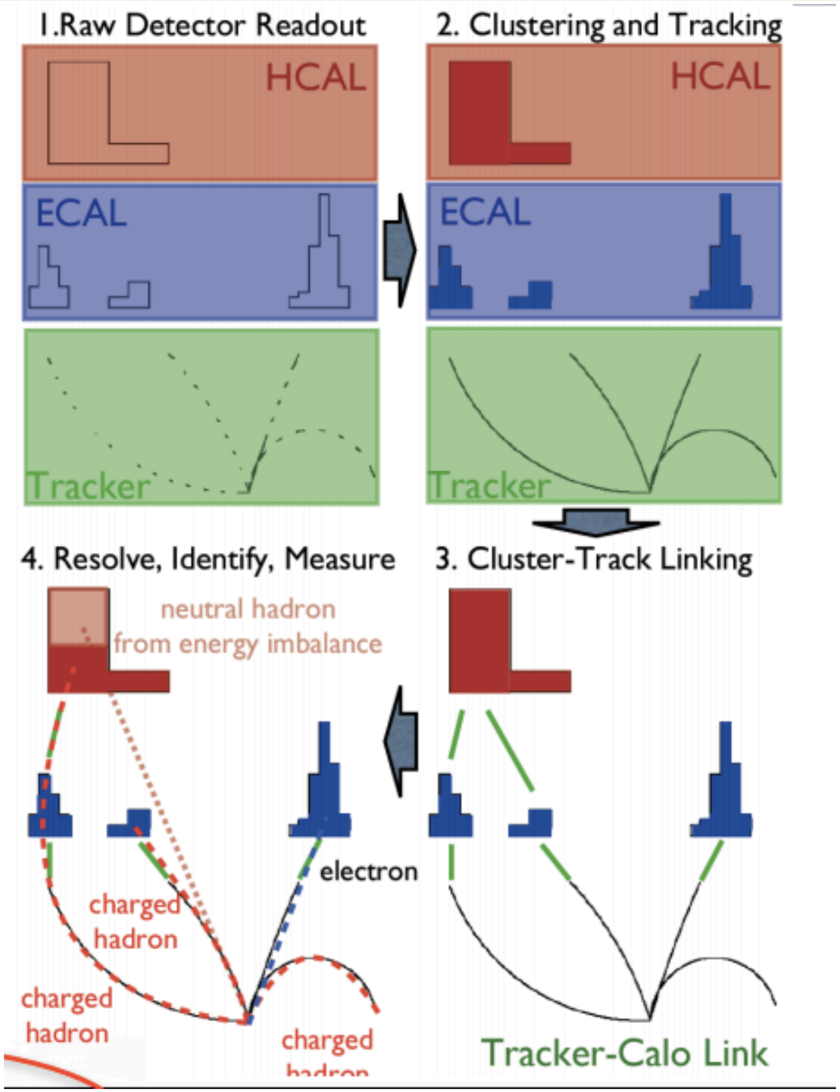
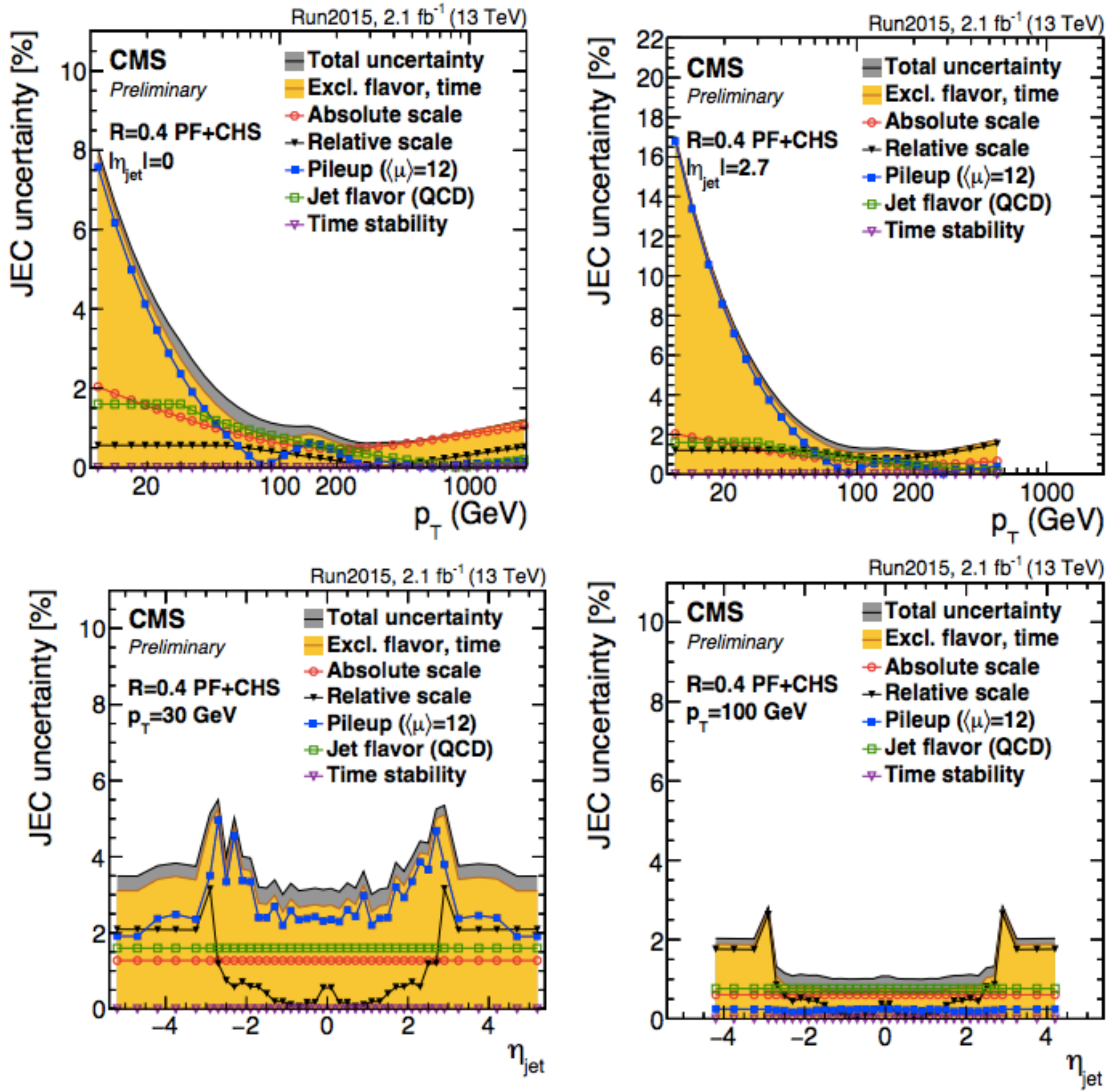


Figure is stolen from somebody who stole it and didn't put a reference

JEC uncertainties in 2015

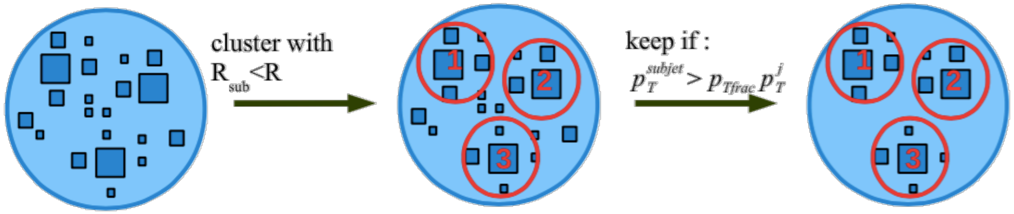
- Pileup uncertainty dominant below 50 GeV
- JetFlavorQCD sizable uncertainty for inclusive jets, but smaller for other analyses
- Other important uncertainties: absolute scale within |η|<3 and relative scale at |η|>3
- Minimum uncertainty of ~0.7% at p_T=300 GeV and |η|<3



PU mitigation for jet mass

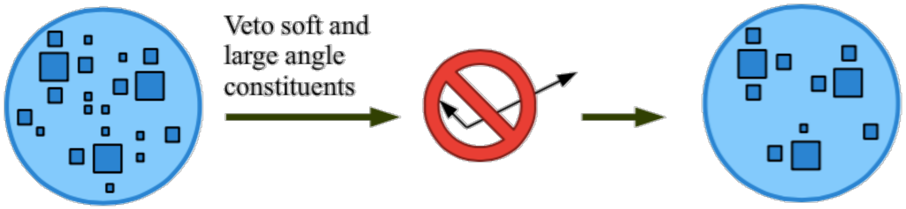
JME-14-001

- Grooming techniques
  - Trimming
    - recluster with AK R_{sub}<R
    - keep constituent jets above a dynamic threshold



$$p_T^{subjet} > p_{Tfrac}\, p_T^{j}$$

  - Pruning
    - Recluster with CA
    - remove soft and large angle constituents at each step



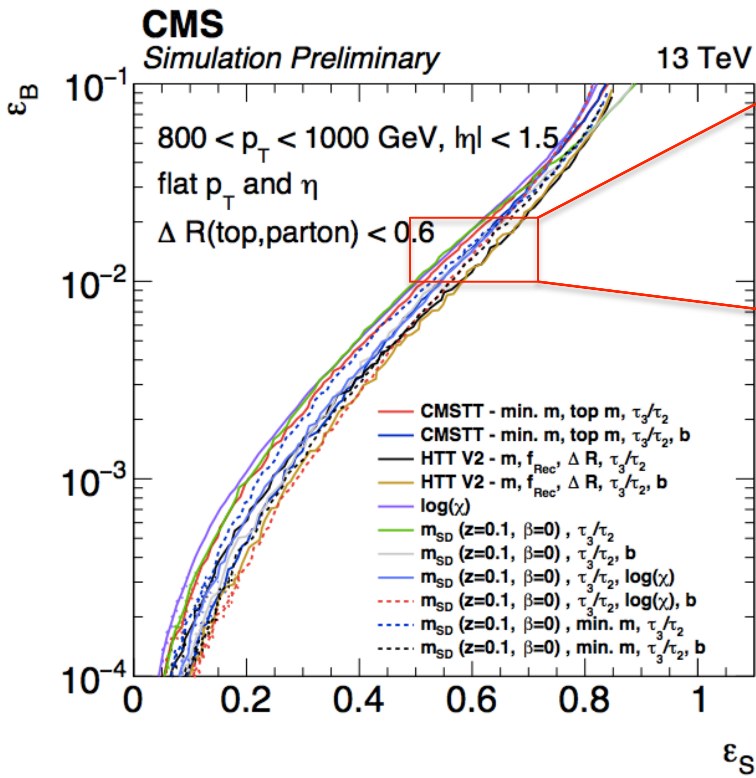
$$\begin{cases} z_{ij} = \dfrac{min\{p_T^i, p_T^j\}}{p_T^i + p_T^j} < z_{cut} \\ \Delta R_{ij} > D_{cut} = \dfrac{2 r_{cut} m^{jet}}{p_T^{jet}} \end{cases}$$

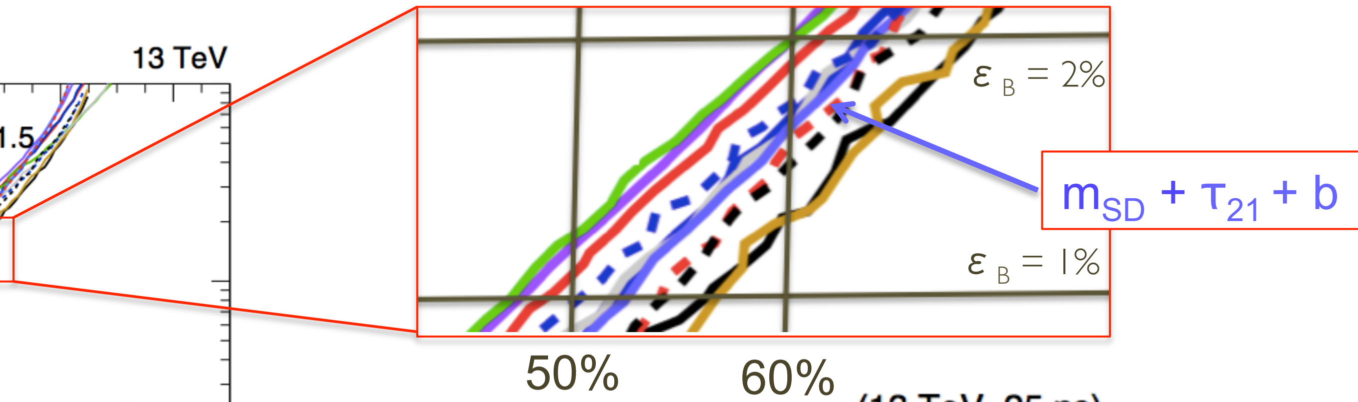
  - Soft Drop/Modified Mass Drop
    - jet is clustered with CA, then declustered and the softer of two jets is dropped if

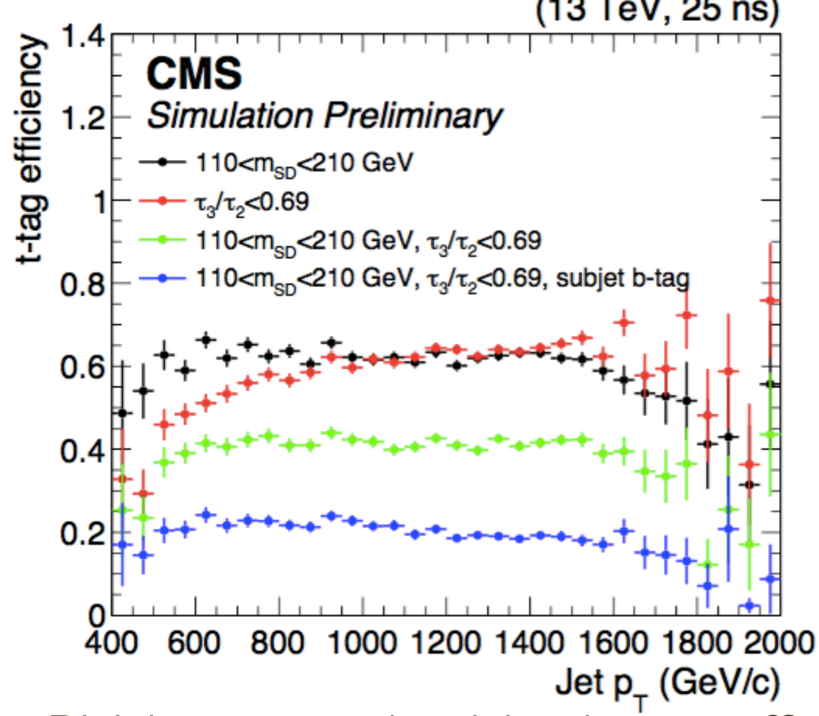
$$\frac{min\{p_T^{j1}, p_T^{j2}\}}{p_T^{j1} + p_T^{j2}} > z_{cut} \left(\frac{\Delta R_{12}}{R}\right)^{\beta}$$

boosted top tagging

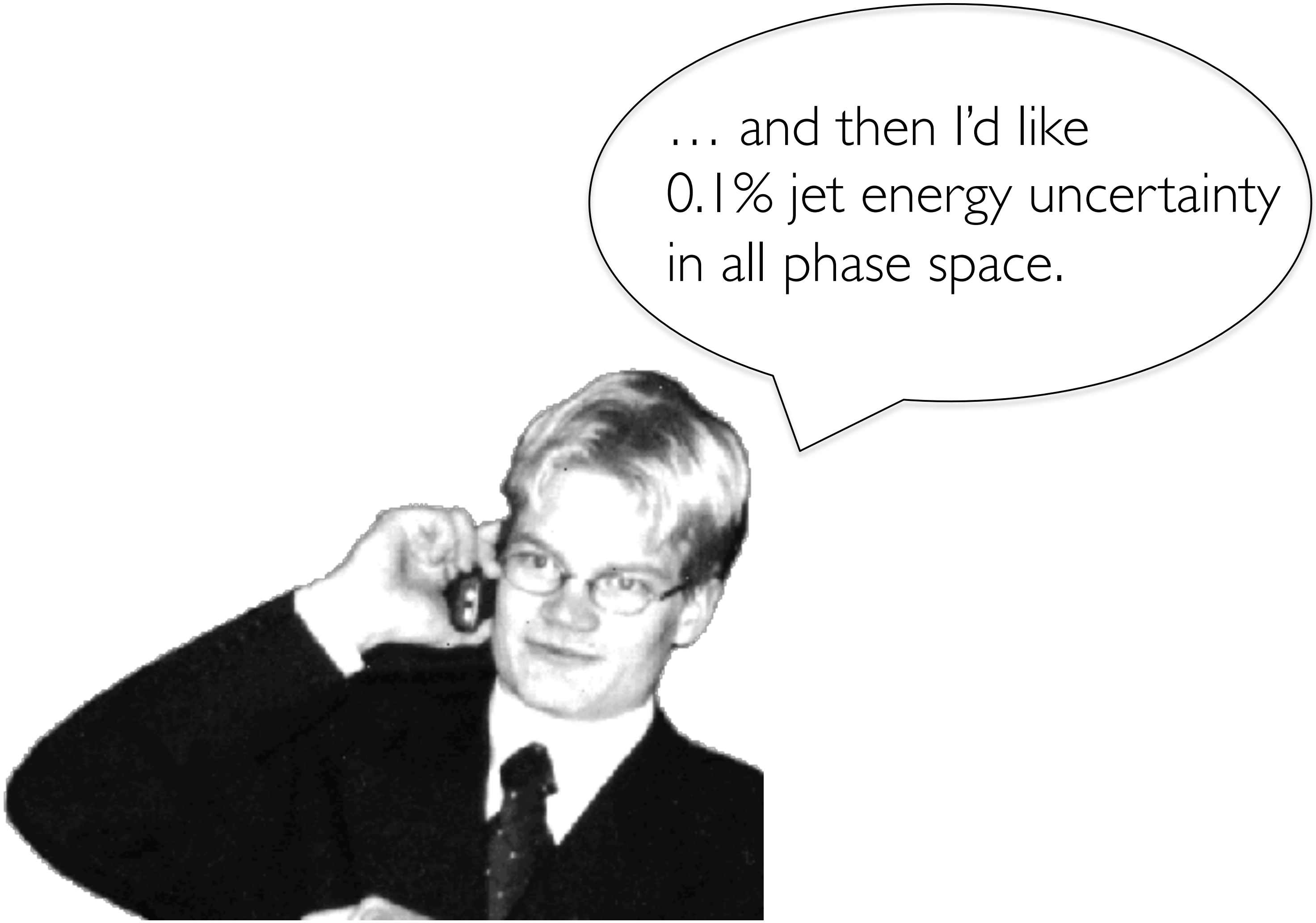
CMS-PAS-JME-15-002, CMS-DP-2015/034







- systematic comparison
  no dramatic differences
  of combined taggers
- ‘safe’ default: $m_{SD} + \tau_{21} + b$
- cut-based or in BDT. Stability vs PU, jet p_T, and subjet b-tag efficiency


… and then I'd like
0.1% jet energy uncertainty
in all phase space.