JET SUBSTRUCTURE FOR HIGGS PHYSICS

Simone Marzani
Università di Genova & INFN Sezione di Genova

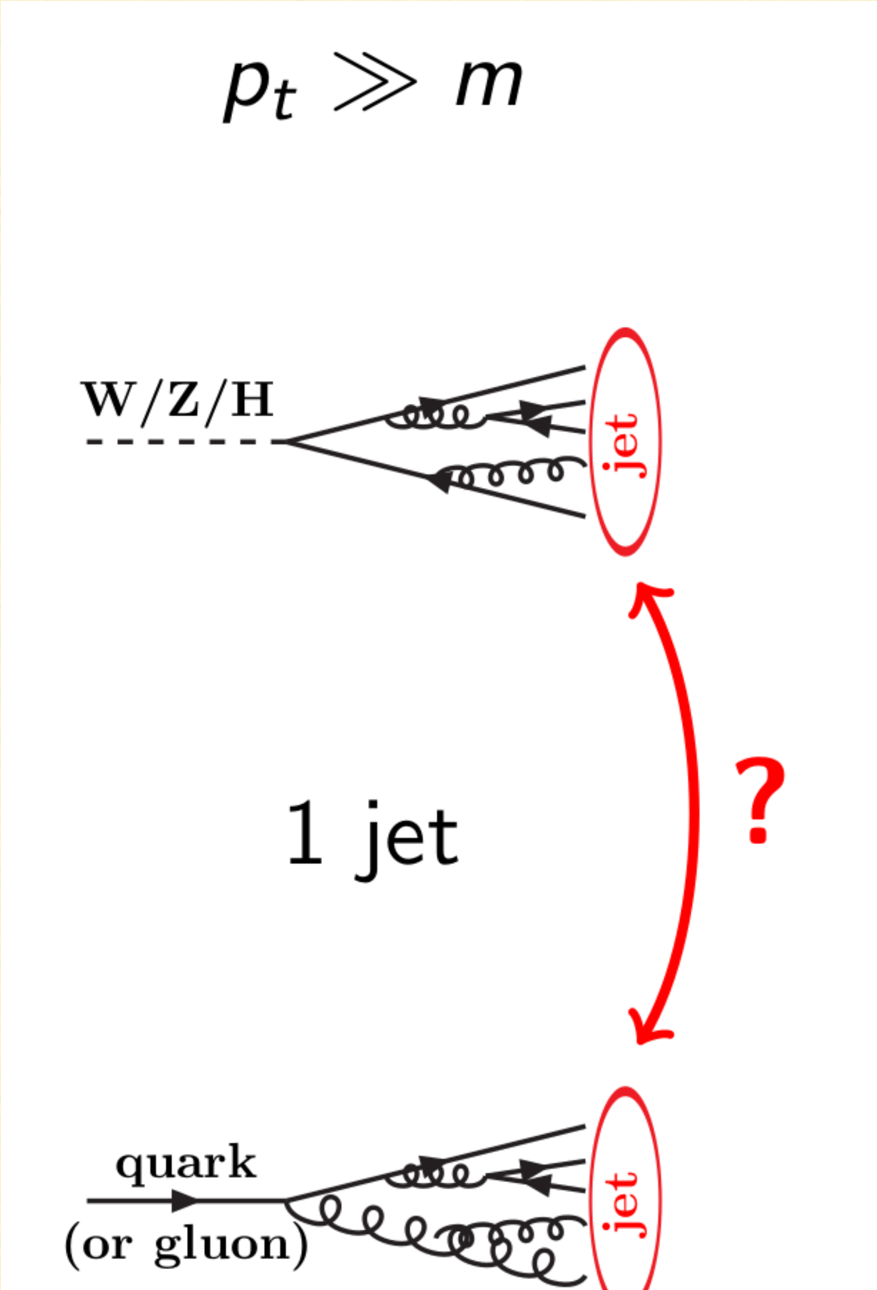
Higgs Couplings 2019
30 September - 4 October 2019
University of Oxford

OUTLINE

- (Higgs) boson tagging with jet substructure: where we are
- Augmenting performance: machine-learning for jet physics
- Conclusions and Open Questions

LOOKING INSIDE JETS

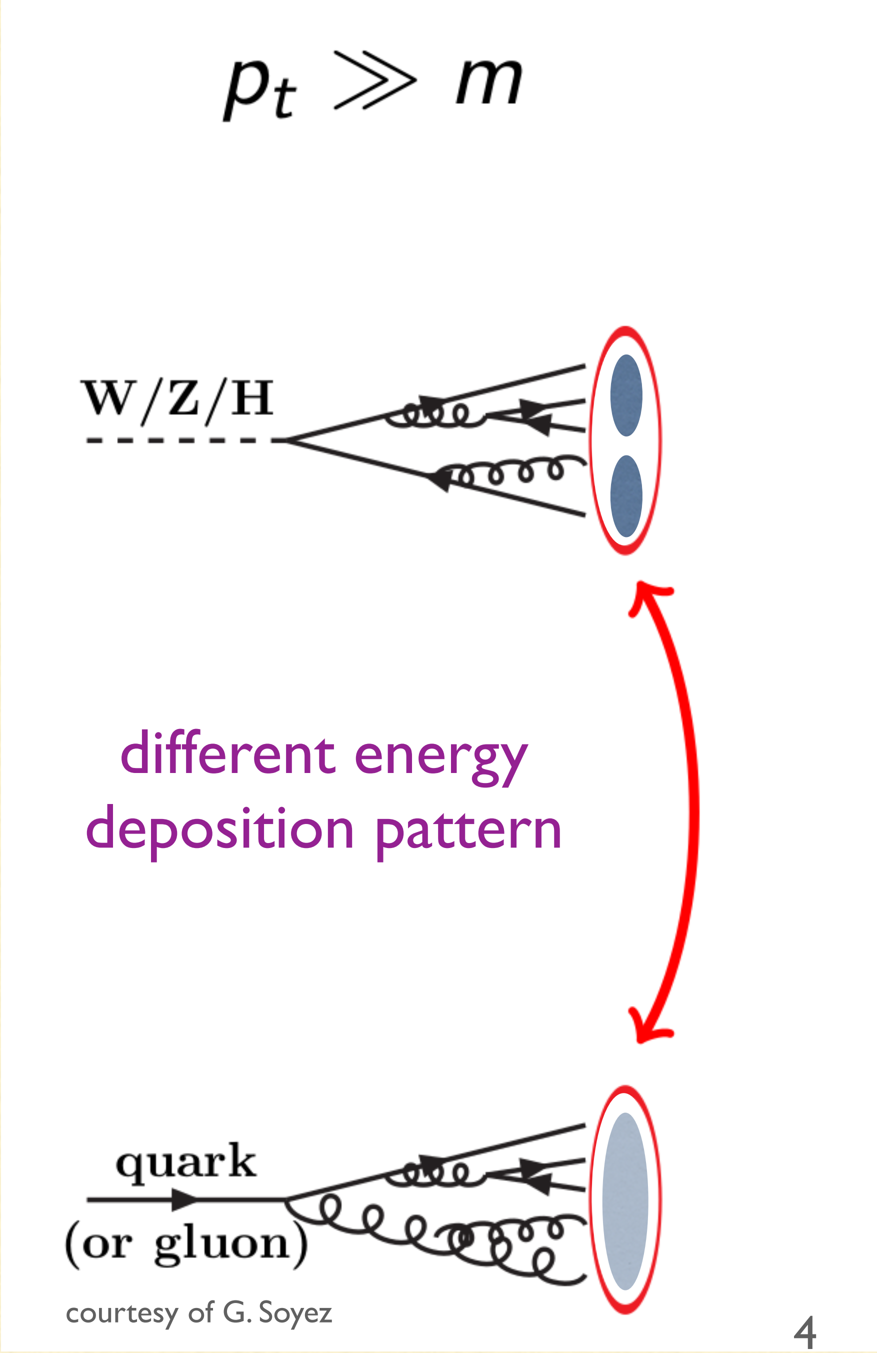
- the two major goals of the LHC
  - search for new particles
  - characterise the particles we know
- jets can be formed by QCD particles but also by the decay of massive particles (if they are sufficiently boosted)
- how can we distinguish signal jets from background ones?



courtesy of G. Soyez

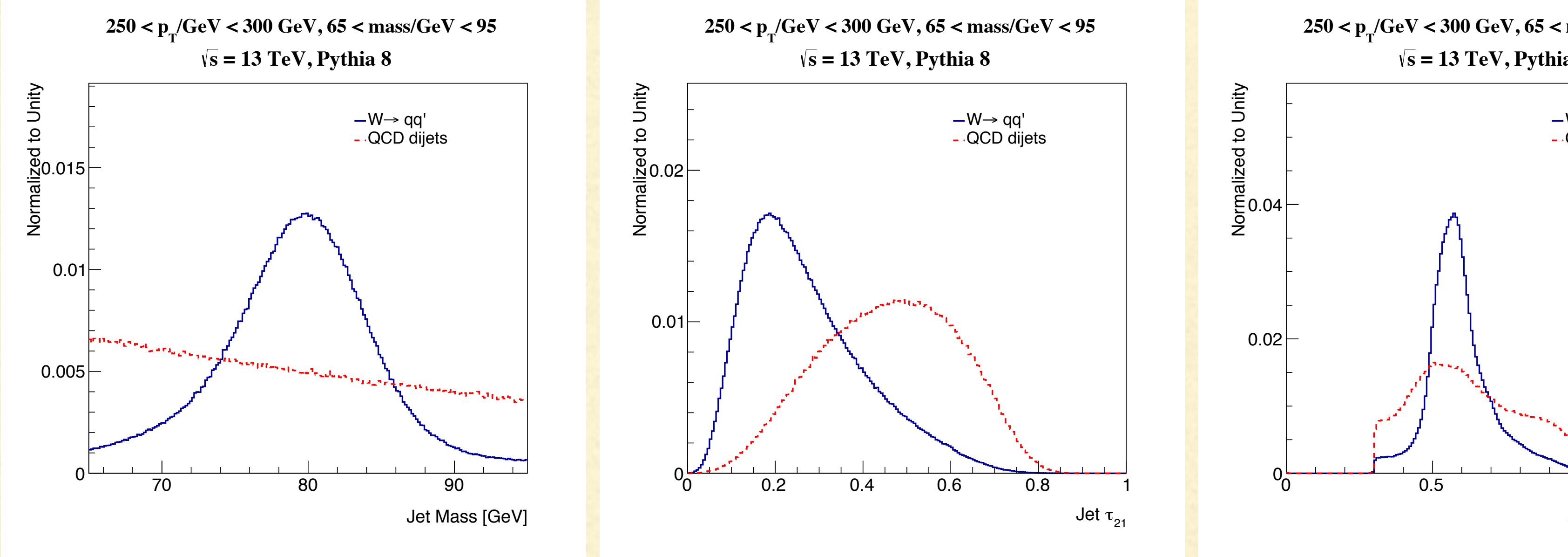
SUBSTRUCTURE IN A NUTSHELL

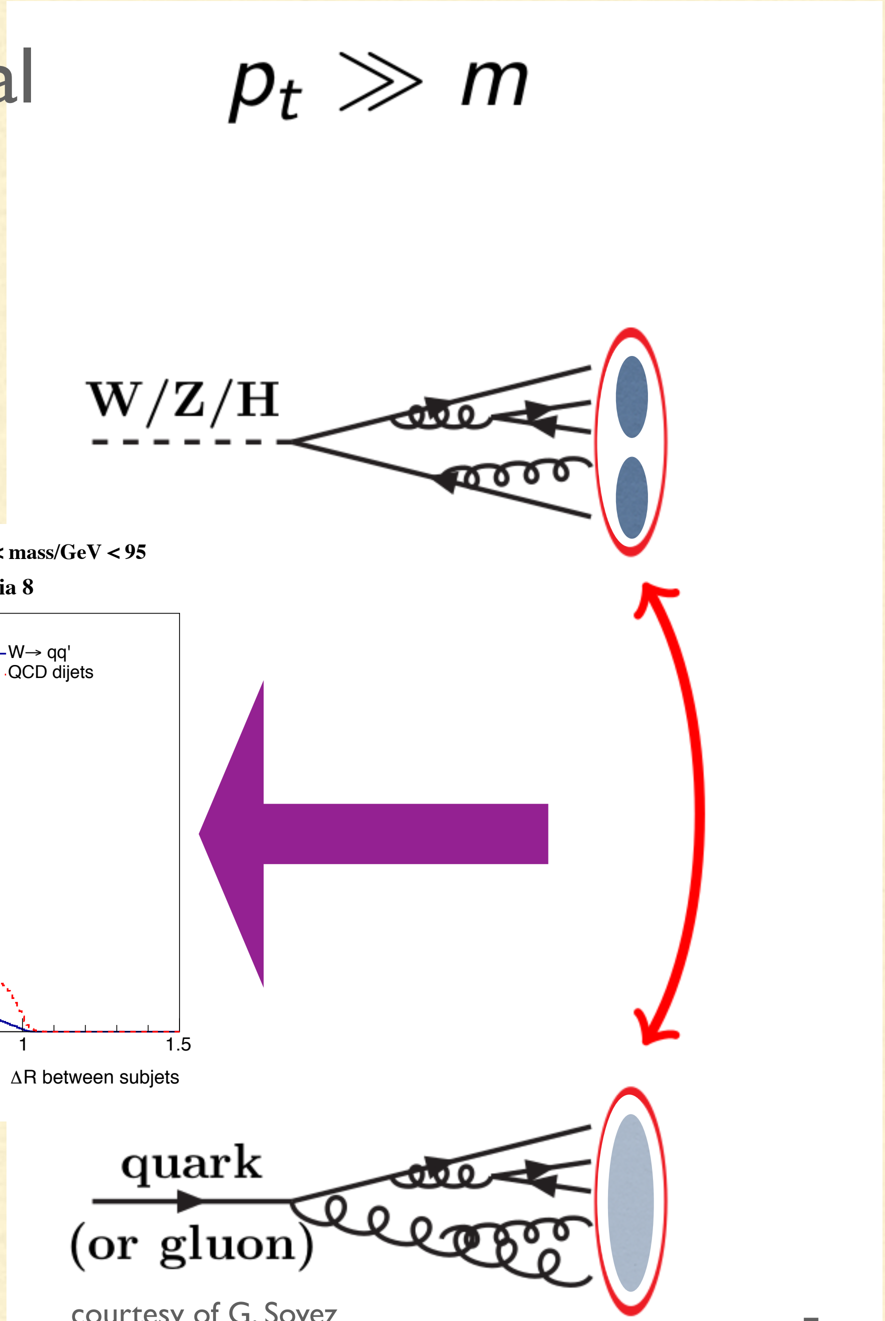
- the final energy deposition pattern is influenced by the originating splitting
- hard vs soft translate into 2-prong vs 1-prong structure
- picture is mudded by many effects (hadronisation, Underlying Event, pileup)
- two-step procedure:
  - *grooming:* clean the jets up by removing soft radiation
  - *tagging:* identify the features of hard decays and cut on them



A THEORIST'S JOB

- devise clever ways to project the multi-dimensional parameter space of final-state momenta into suitable lower dimensional (typically 1-D) distributions

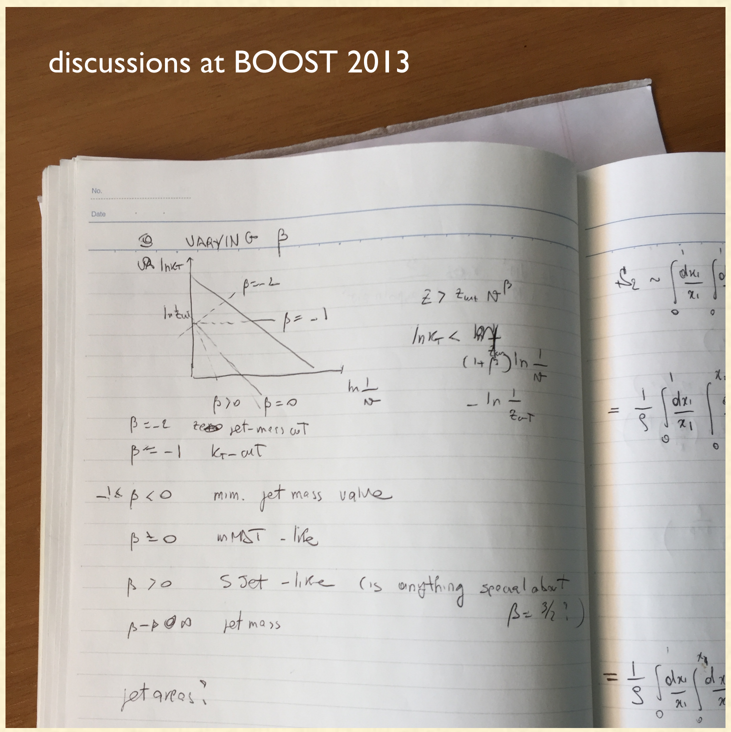


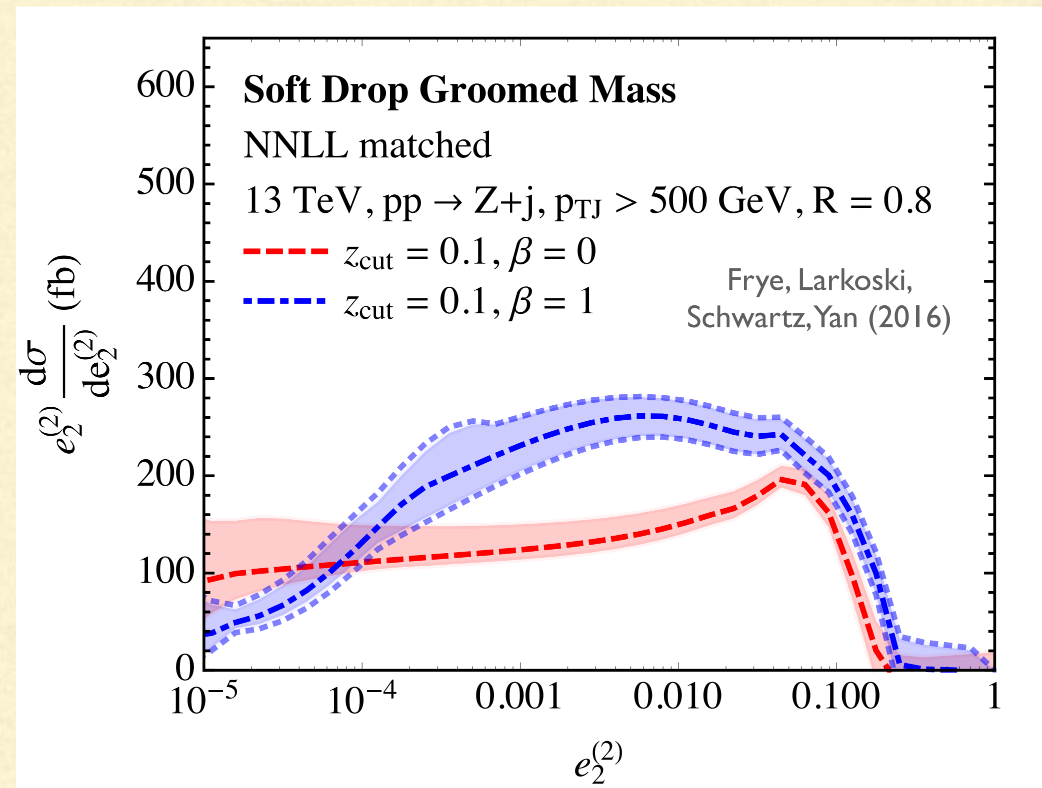


courtesy of G. Soyez

for an introduction see SM, Soyez, Spannowsky

FROM IDEAS TO PRECISION



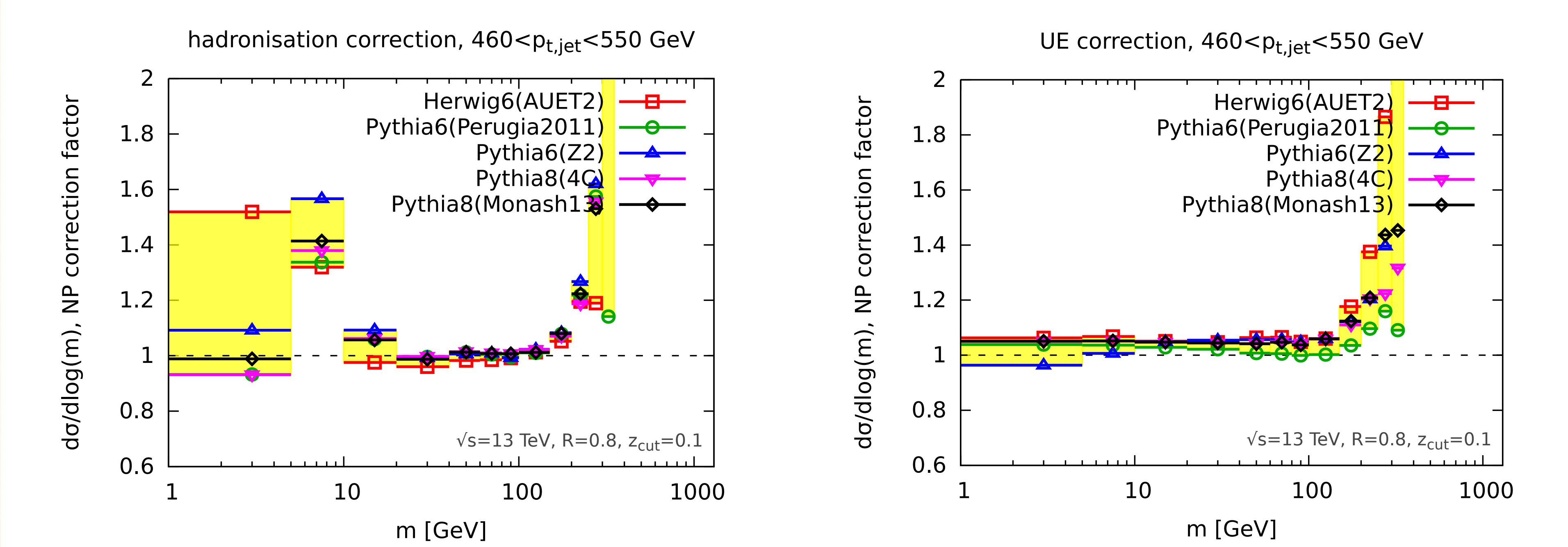


- understanding of groomers and taggers led to the definition of theory-friendly efficient tools, e.g. soft drop:
  - good perturbative properties (convergence, absence of soft effects such as non-global logs)
  - small (but non-trivial) non-perturbative corrections

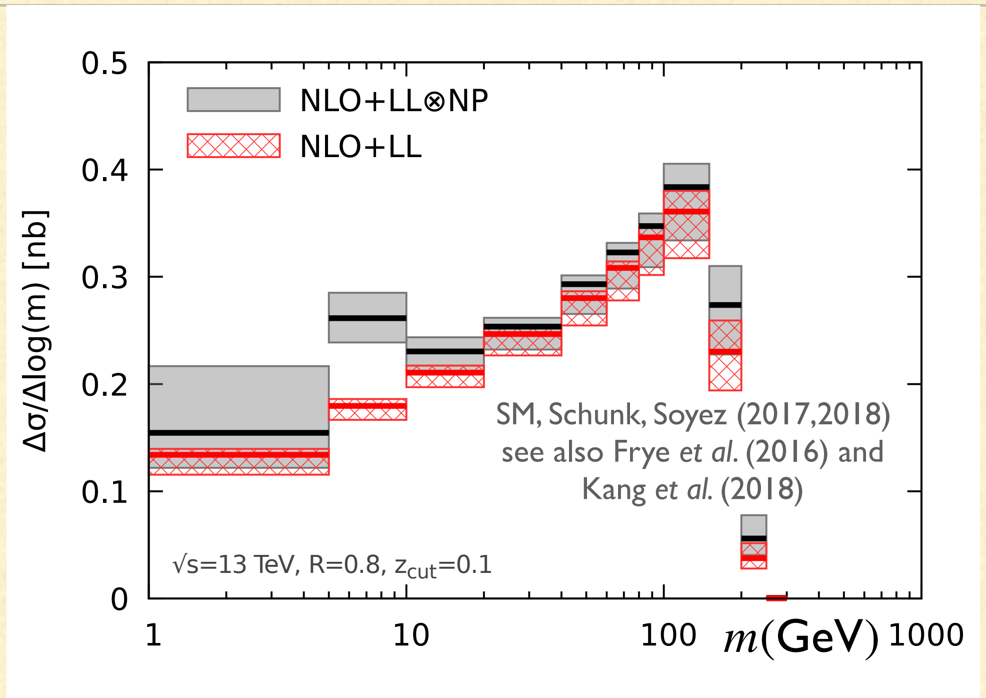
Hoang, Mantry, Pathak, Stewart (2019)

FROM THEORY TO DATA

- time is mature for theory / data comparison
- reduced sensitivity to non-pert physics (hadronisation and UE) should make the comparison more meaningful
- what is the value of unfolded measurements / theory comparisons for "discovery" tools?
  - understanding systematics (e.g. kinks and bumps)
  - where non-pert. corrections are small, test perturbative showers in MCs
  - at low mass, hadronisation is large but UE is small: TUNE!

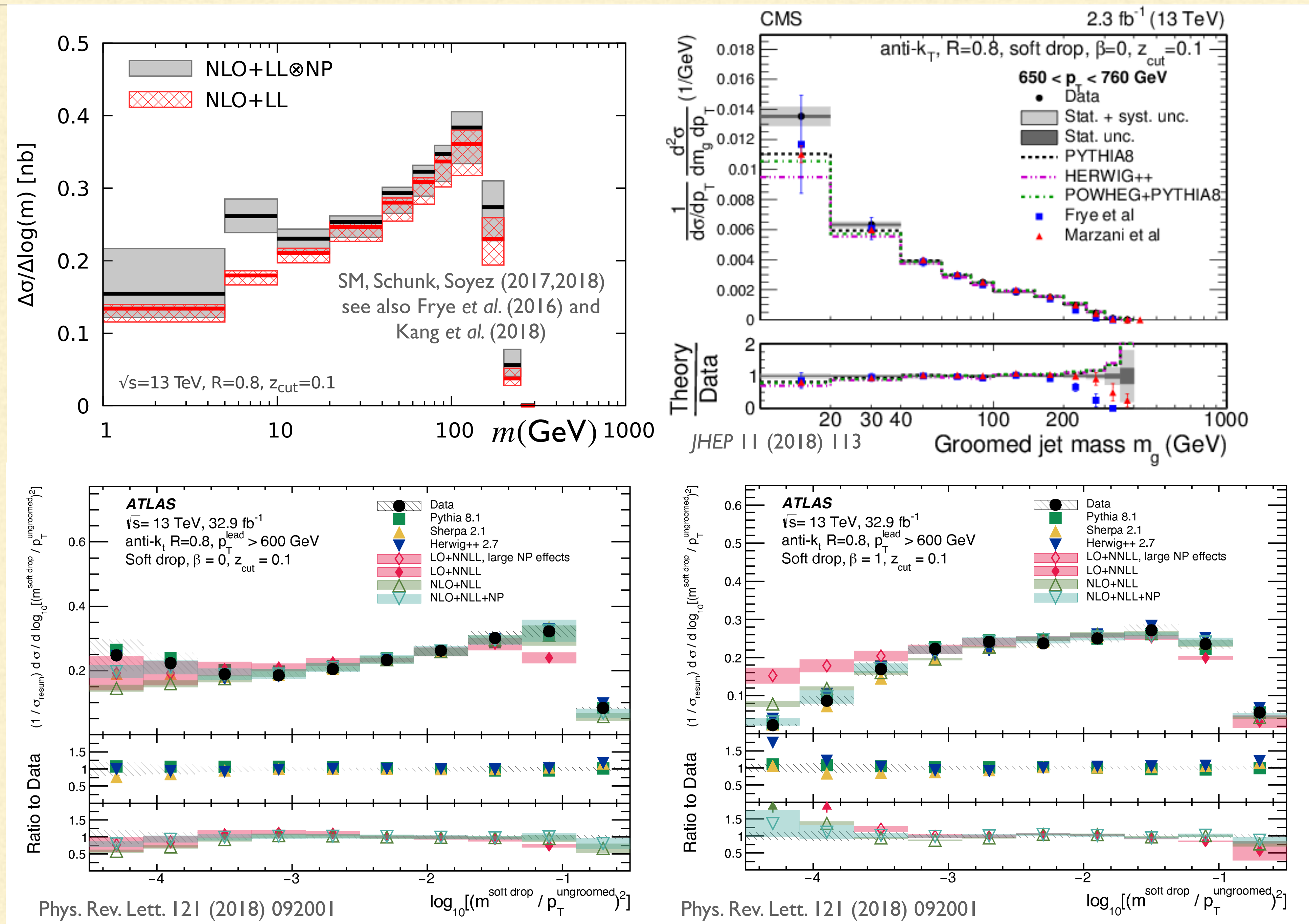


THEORY PREDICTIONS…



- large range of masses where non-pert. corrections are small and we can trust resummation
- they can be included through MC or analytical modelling

…AND THE DATA


NLO+LL⊗NP
NLO+LL
Δσ/Δlog(m) [nb]
SM, Schunk, Soyez (2017,2018)
see also Frye et al. (2016) and
Kang et al. (2018)
√s=13 TeV, R=0.8, zcut=0.1
m(GeV)
CMS
2.3 fb-1 (13 TeV)
anti-kT, R=0.8, soft drop, β=0, zcut=0.1
650 < pT < 760 GeV
Data
Stat. + syst. unc.
Stat. unc.
PYTHIA8
HERWIG++
POWHEG+PYTHIA8
Frye et al
Marzani et al
Theory / Data
JHEP 11 (2018) 113
Groomed jet mass mg (GeV)
ATLAS
√s= 13 TeV, 32.9 fb-1
anti-kt R=0.8, pT lead > 600 GeV
Soft drop, β = 0, zcut = 0.1
Data
Pythia 8.1
Sherpa 2.1
Herwig++ 2.7
LO+NNLL, large NP effects
LO+NNLL
NLO+NLL
NLO+NLL+NP
Ratio to Data
Phys. Rev. Lett. 121 (2018) 092001
log10[(m soft drop / pT ungroomed)2]
ATLAS
√s= 13 TeV, 32.9 fb-1
anti-kt R=0.8, pT lead > 600 GeV
Soft drop, β = 1, zcut = 0.1
Ratio to Data
Phys. Rev. Lett. 121 (2018) 092001
log10[(m soft drop / pT ungroomed)2]

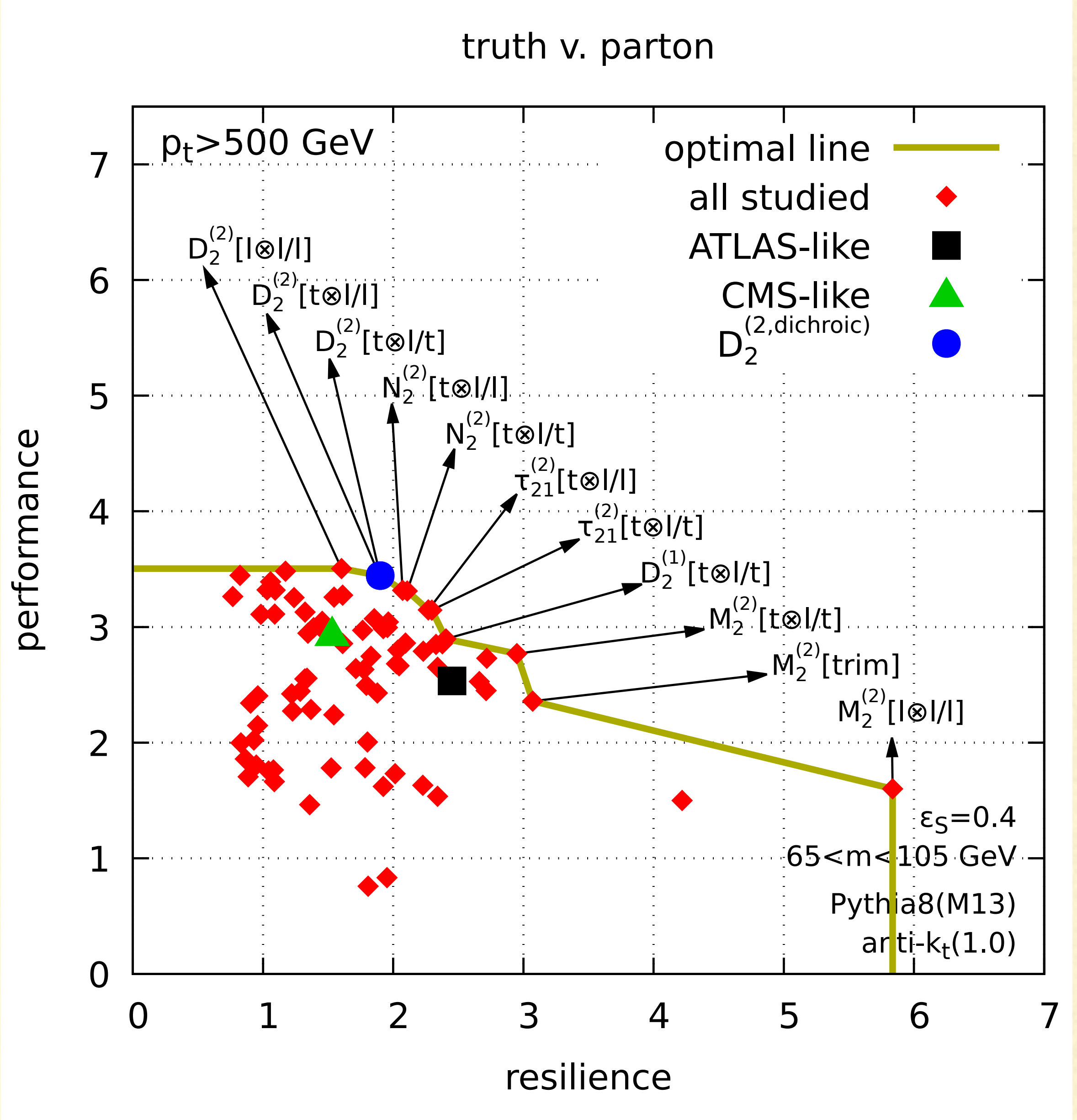
PERFORMANCE & RESILIENCE

- first-principle understanding of groomers' and taggers' perturbative properties has reached remarkable levels
- resilience measures a tagger's robustness against non-perturbative effects (hadronisation and UE)
- it is defined in terms of signal/background efficiencies with/without non-pert. contributions Looking inside jets

$$\zeta = \left(\frac{\Delta\epsilon_S^2}{\langle\epsilon\rangle_S^2} + \frac{\Delta\epsilon_B^2}{\langle\epsilon\rangle_B^2} \right)^{-1/2}$$

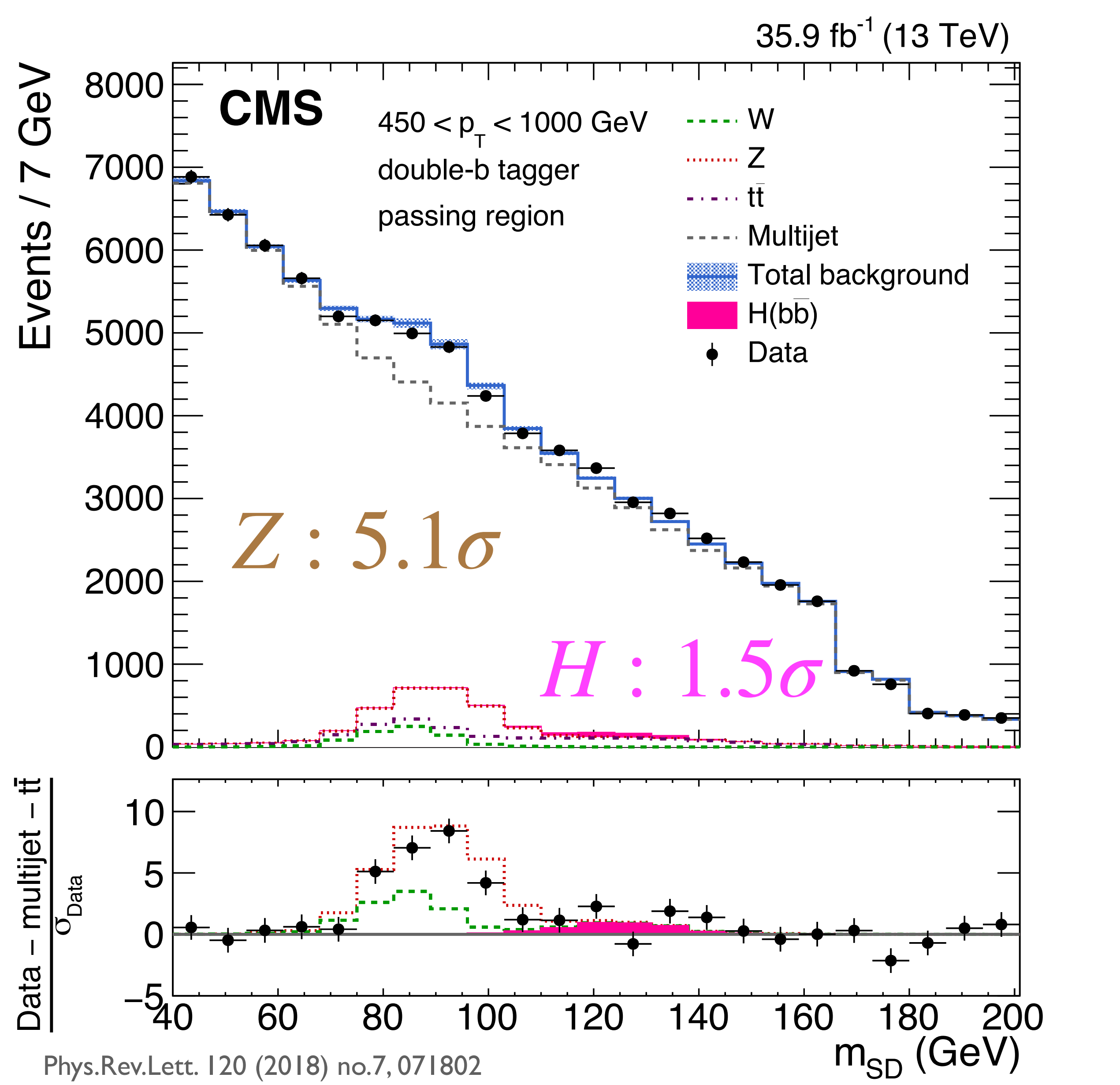
$$\Delta\epsilon_{S,B} = \epsilon_{S,B} - \epsilon'_{S,B},$$

$$\langle\epsilon\rangle_{S,B} = \frac{1}{2}\left(\epsilon_{S,B} + \epsilon'_{S,B}\right)$$



HARD WORK DOES PAY OFF

- QCD and EW corrections to obtain Z+jets and W+jets
- Higgs p_T spectrum corrected for finite top mass effects
- inclusion of N^3LO normalisation
- matching NLO-PS
- state-of-the arts PDFs

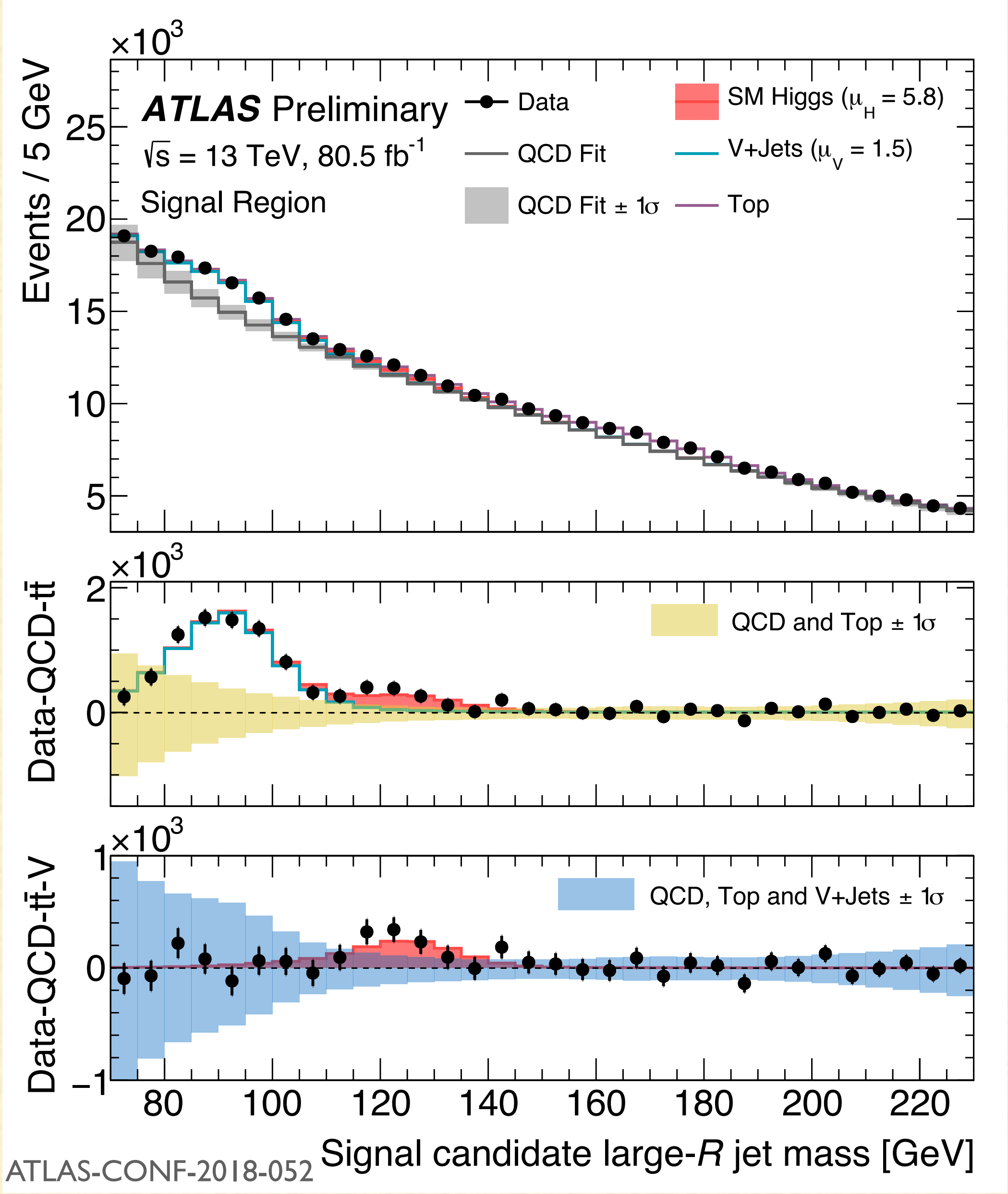


Phys.Rev.Lett. 120 (2018) no.7, 071802

- state-of-the art jet reconstruction (anti-k_t & particle-flow)
- b-tagging
- soft-drop grooming
- 2-prong jets identified with energy correlation function $N^1{}_2$
- decorrelation: $N^1{}_2 \rightarrow N^{1,DDT}{}_2$

HARD WORK DOES PAY OFF

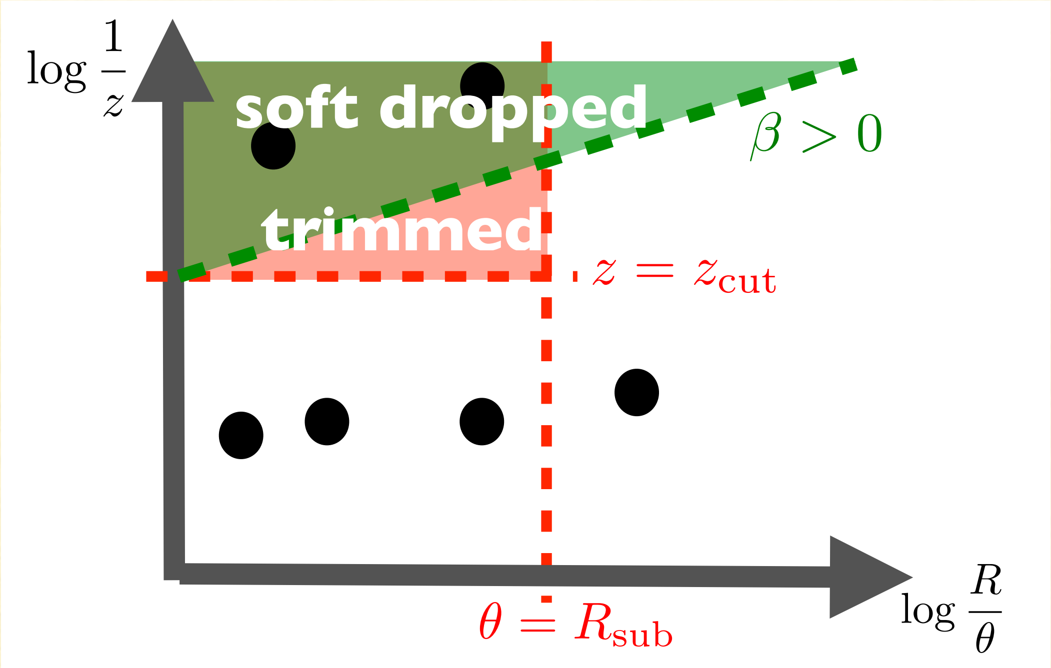
- QCD and EW corrections to obtain Z+jets and W+jets
- Higgs p_T spectrum corrected for finite top mass effects
- ~~inclusion of N^3LO normalisation~~
- matching NLO-PS
- state-of-the arts PDFs

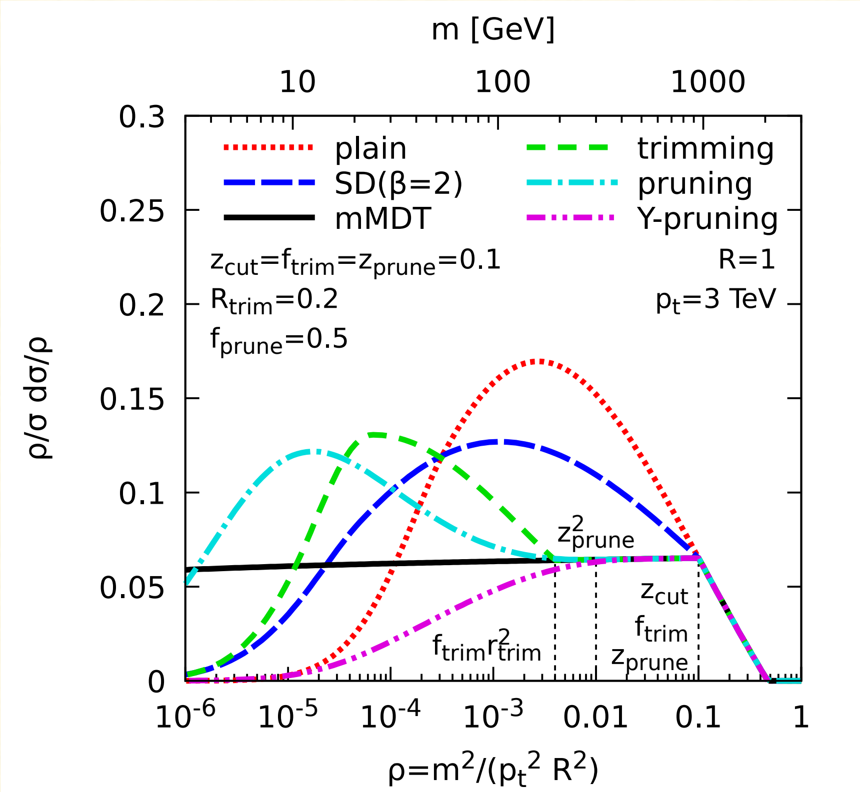


- state-of-the art jet reconstruction (anti-k_t & topoclusters)
- b-tagging
- trimming
- 2-prong jets identified by requiring two track subjets with variable R

DIFFERENCES IN GROOMING: SOFT-DROP VS TRIMMING

- CMS favours soft drop, ATLAS trimming
- Performance depends on the detail of the jet reconstruction procedure / detector
- However, performance is not the only criterion





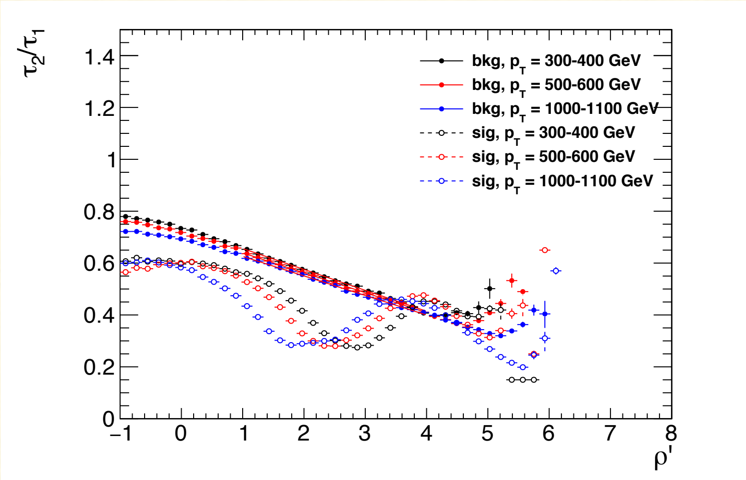
- trimming has an abrupt change of behaviour due to fixed R_{sub}
- loss of efficiency at high p_T
- in SD angular resolution controlled by the exponent β: phase-space appears smoother
- SD under better theory control

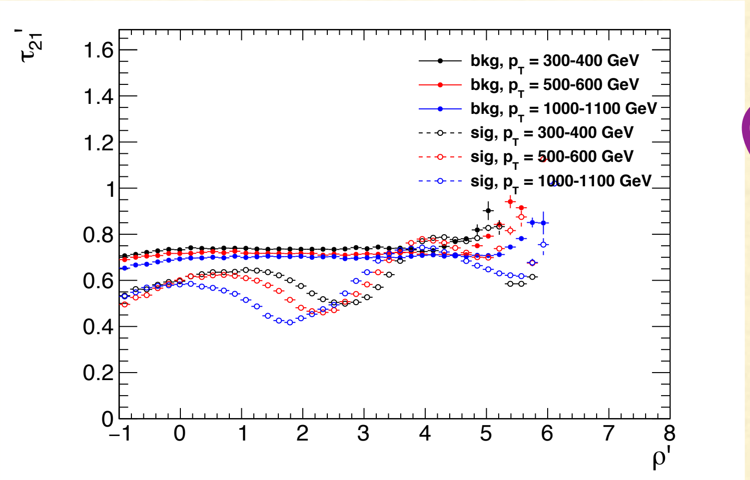
DIFFERENCES IN TAGGING: SHAPE VS VARIABLE-R

- CMS analysis cuts on a shape to isolate 2-pronged jets
- N^1_2 is a ratio of generalised energy correlation functions optimised to work after grooming

  Moult, Necib, Thaler (2016)
- DDT is a procedure to de-correlate the mass from the jet shape cut, reducing sculpting

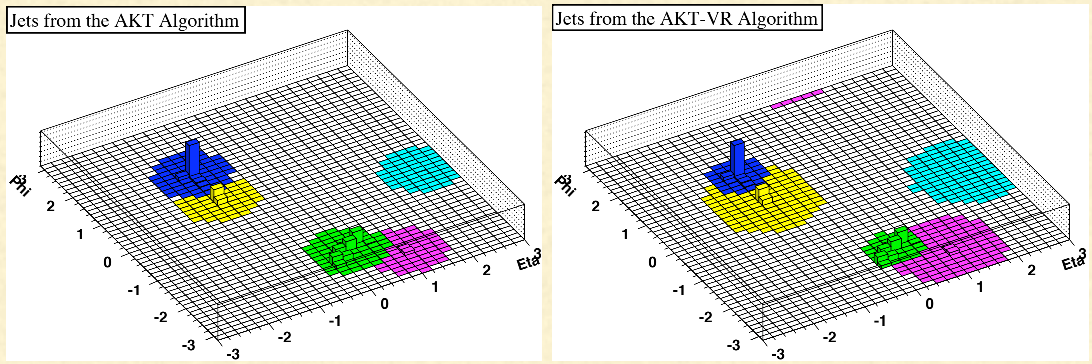
  Dolen, Harris, SM, Nhan, Rappoccio (2016)





- ATLAS analysis looks for 2 track jets using variable-R jets

  Krohn, Thaler, Wang (2009)



$$d_{ij} = \min\left[p_{Ti}^{2n}, p_{Tj}^{2n}\right] R_{ij}^2, \qquad d_{iB} = p_{Ti}^{2n} R_{\text{eff}}(p_{Ti})^2$$

$$R_{\text{eff}}(p_T) = \min\left[\frac{\rho}{p_T}, R_{\max}\right]$$

30 GeV

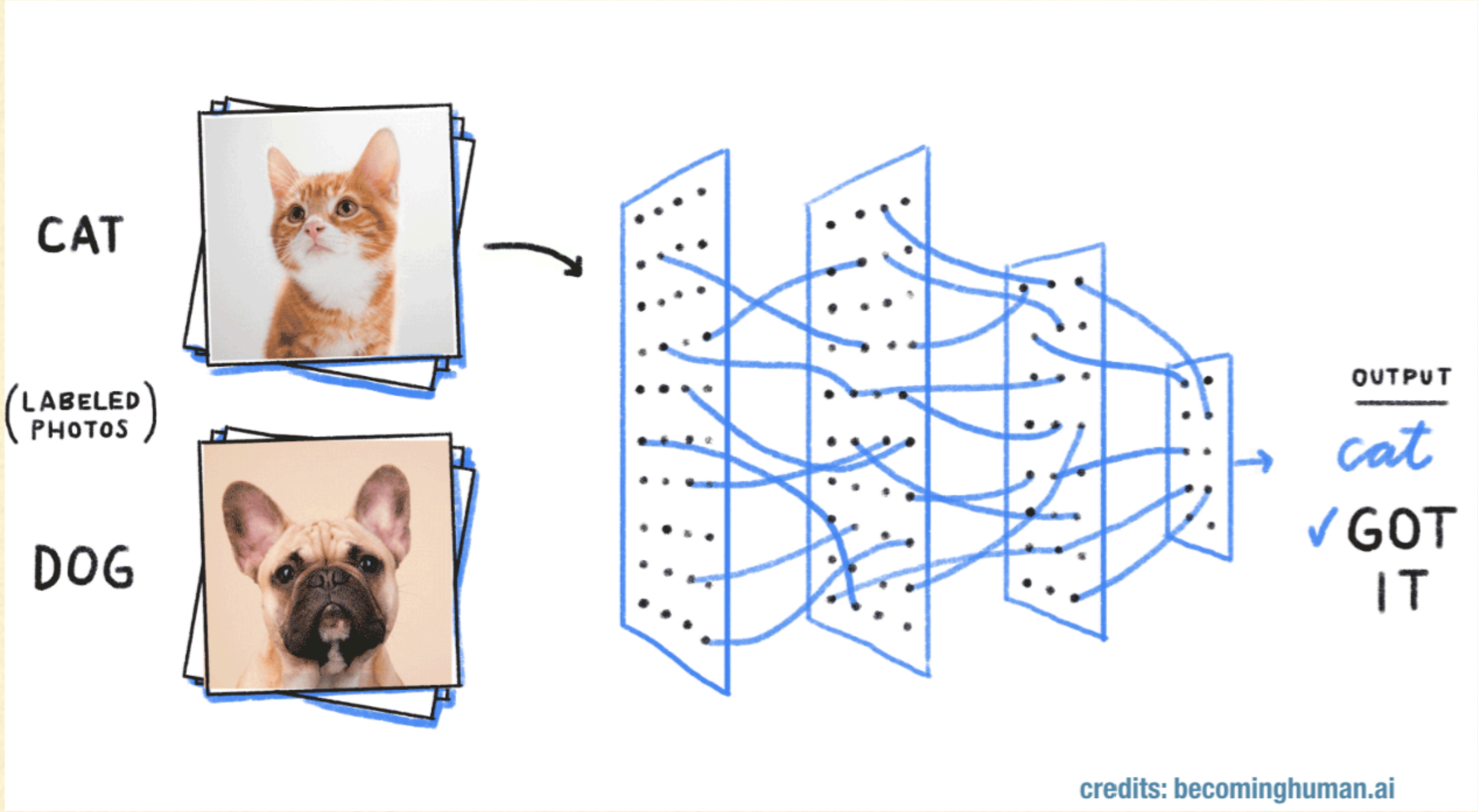
0.4

WHAT'S LEFT TO DO?

- $H \rightarrow bb$ is the holy grail of jet substructure, where it all started … embarrassingly it's not been observed yet!
- Need more efficient tools?
    - enter machine learning!

DEEP LEARNING

- a wave of machine learning algorithms has hit jet physics in the past 3/4 years
- ML algorithms are powerful tools for classification, can we then apply them to our task?

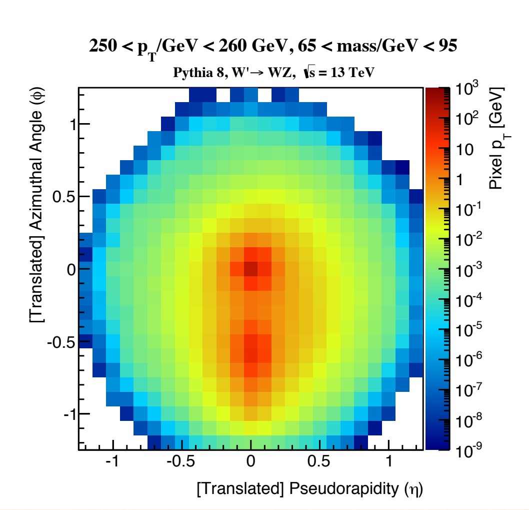


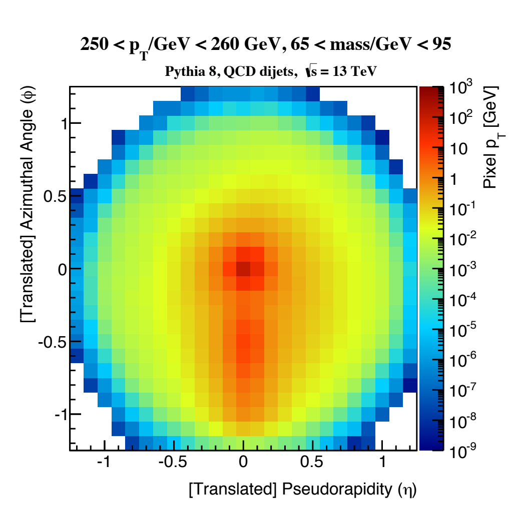
- if an algorithm can distinguish pictures of cats and dogs, can it also distinguish QCD jets from boosted-objects?
- number of papers trying to answer this question has recently exploded!
- very active and fast-developing field

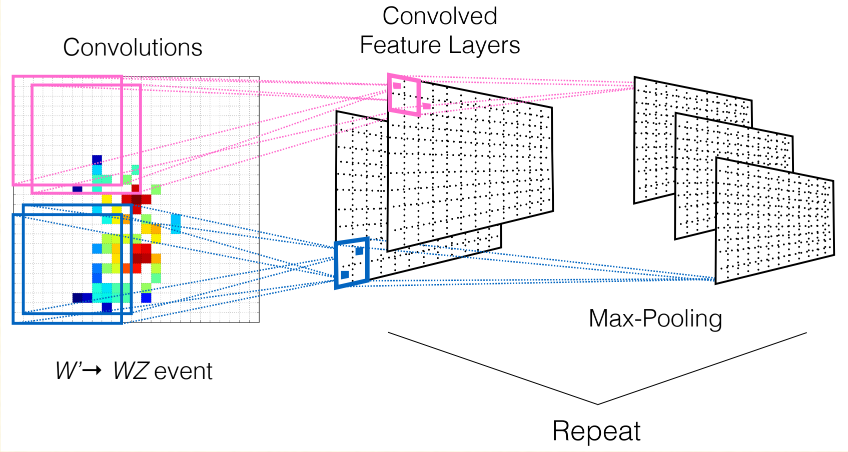
JETS AS IMAGES

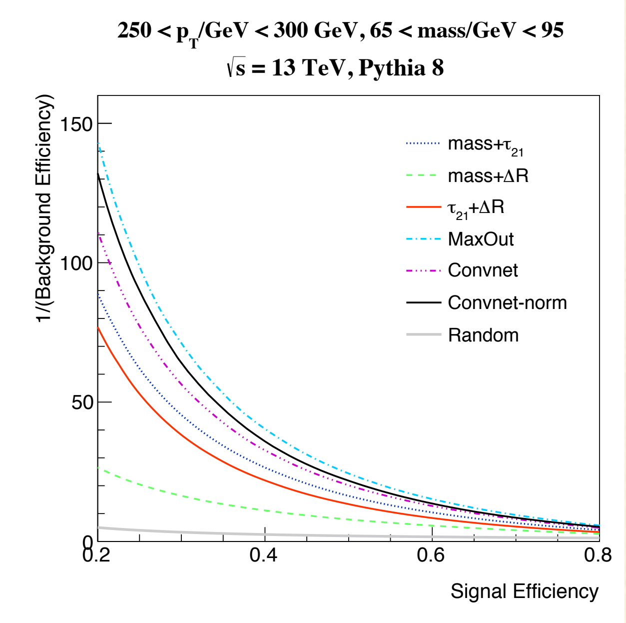
- jet images do what they say: project the jet into a *nxn* pixel image, where intensity is given by energy deposition
- use convolutional neural network (CNN) to classify
- right pre-processing is crucial for many reasons: we average over many events and Lorentz symmetry would wash away any pattern

Cogan, Kagan, Strauss, Schwartzman (2015)
de Olivera, Kagan, Mackey, Nachman, Schwartzman (2016)



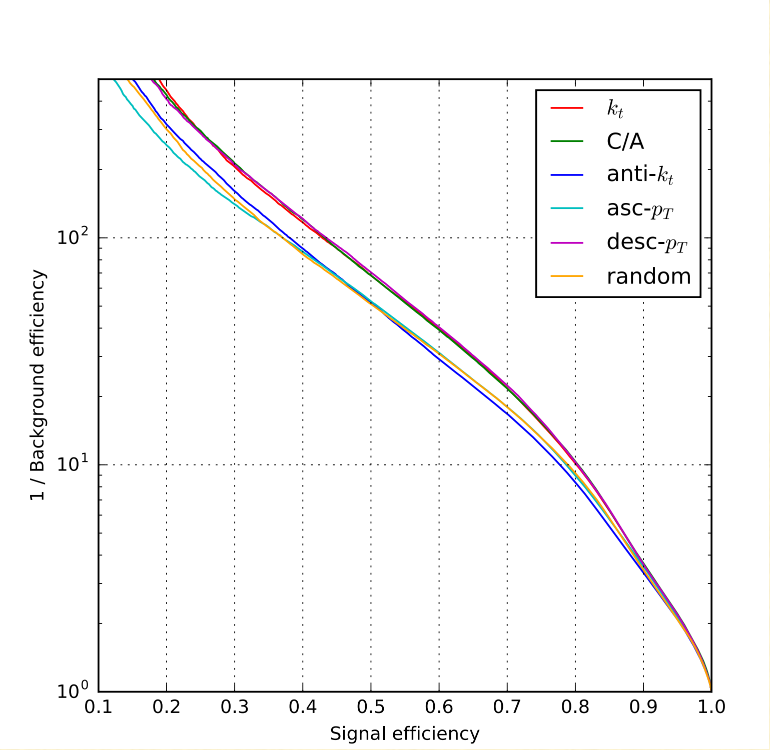


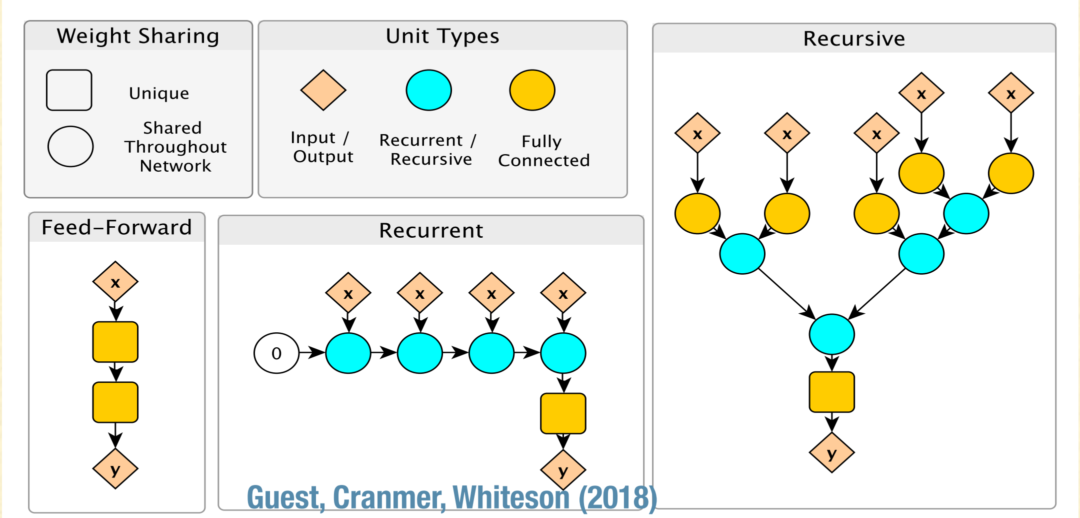




BEYOND IMAGES: 4-MOMENTA

- analyses typically have access to more information than energy deposit in the calorimeter: e.g. particle id, tracks, clustering history in a jet, etc.
- build network that take 4-momenta as inputs:
  - clever N-body phase-space parametrisation to maximise information Datta, Larkoski (2017)
  - recurrent / recursive neural networks to model jet clustering history (using techniques borrowed from language recognition) Louppe, Cho, Cranmer (2017)

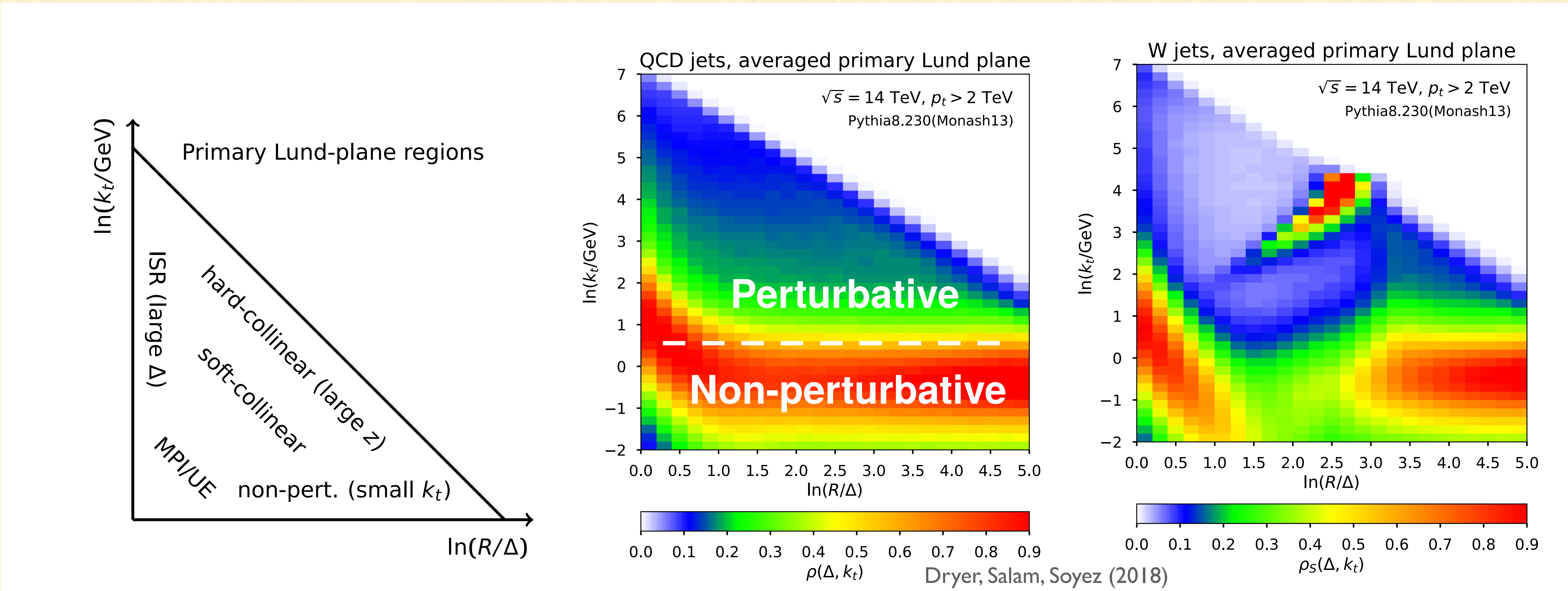




Guest, Cranmer, Whiteson (2018)

DEEP LEARNING MEETS DEEP THINKING: LUND JET PLANE

- inputs of ML algorithms can be low-level (calorimeter cells/particle 4-momenta) but also higher-level variables
- physics intuition can lead us to construct better representations of a jet: the Lund jet plane
  - de-cluster the jet following the hard branch and record (kt, Δ) at each step
  - feed this representation to a log-likelihood or a ML algorithm

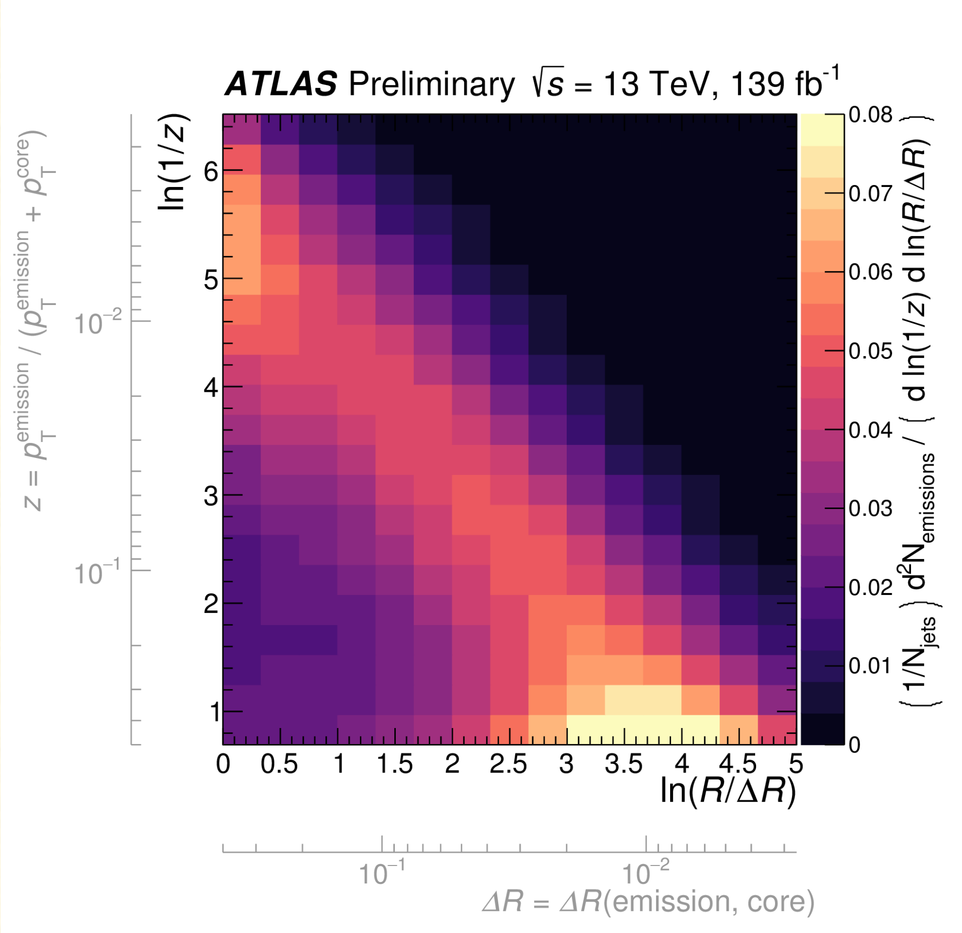


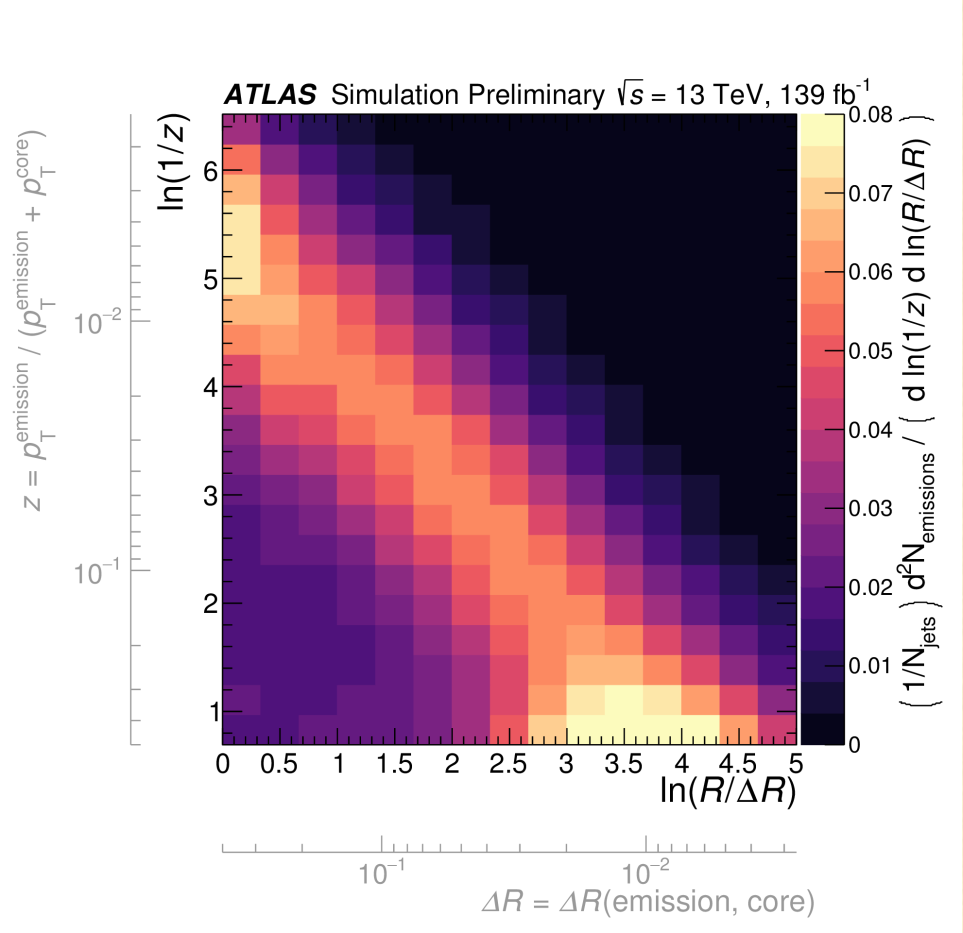
Dryer, Salam, Soyez (2018)

MAPPING OUT THE LUND PLANE

- ATLAS presented at BOOST 2019 the first experimental measurement of the Lund plane (note the different coordinates)

ATLAS-CONF-2019-035



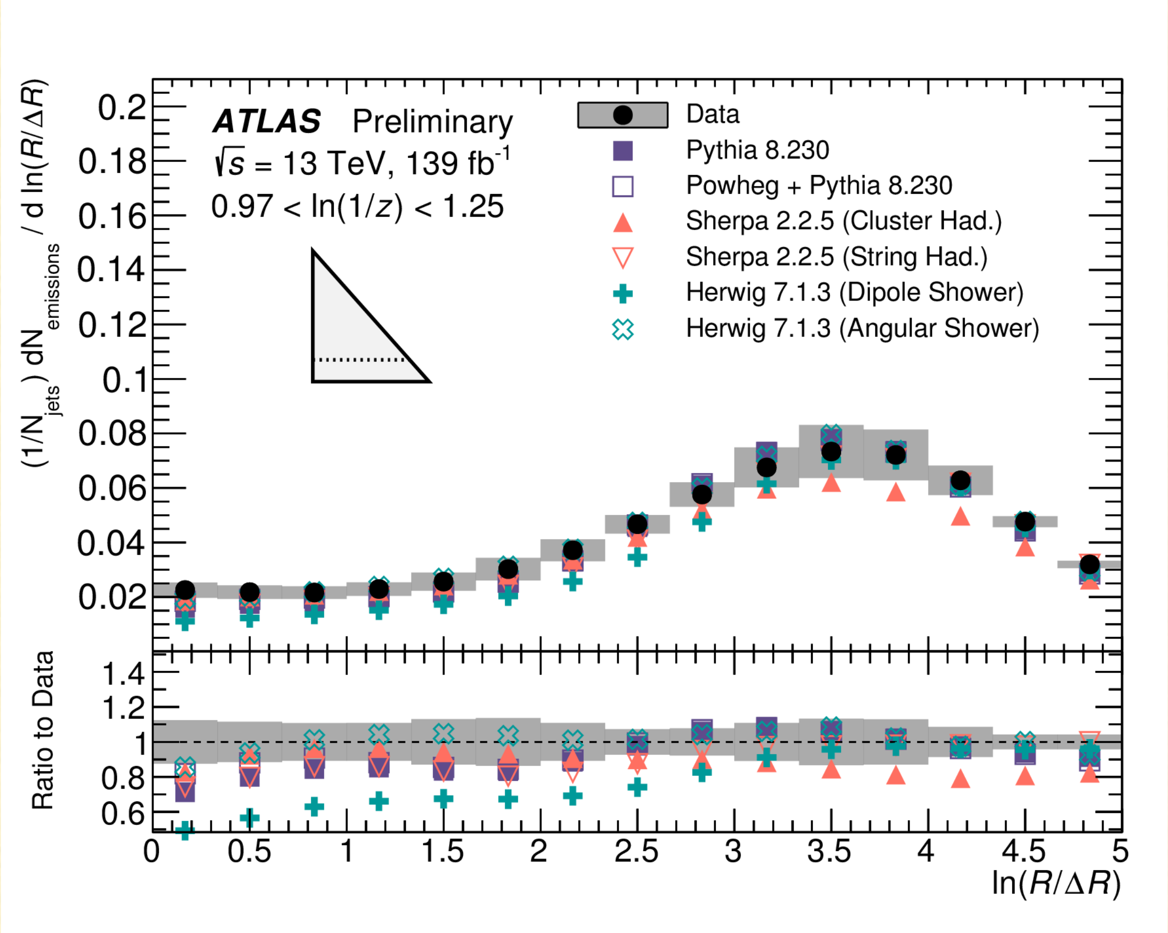


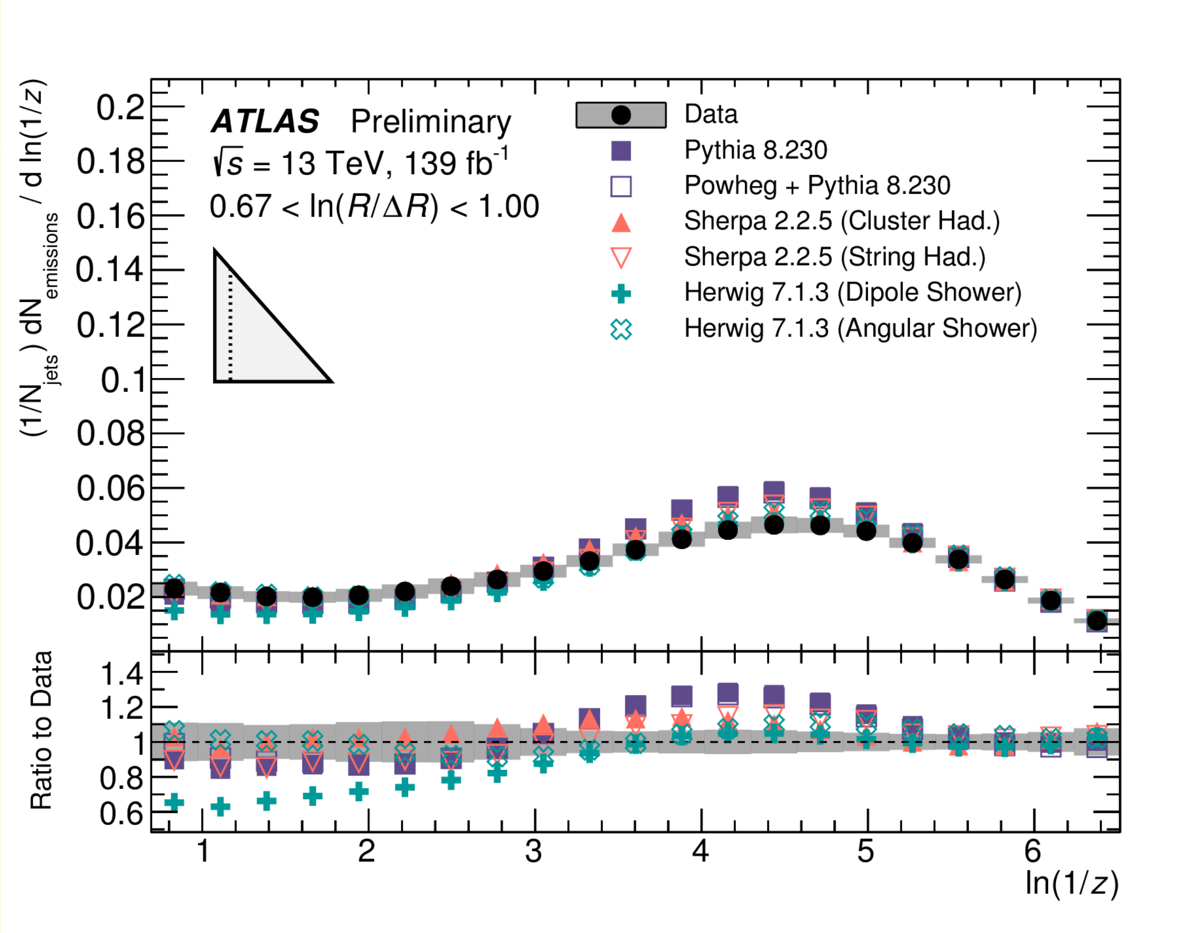
MAPPING OUT THE LUND PLANE

- ATLAS presented at BOOST 2019 the first experimental measurement of the Lund plane (note the different coordinates)

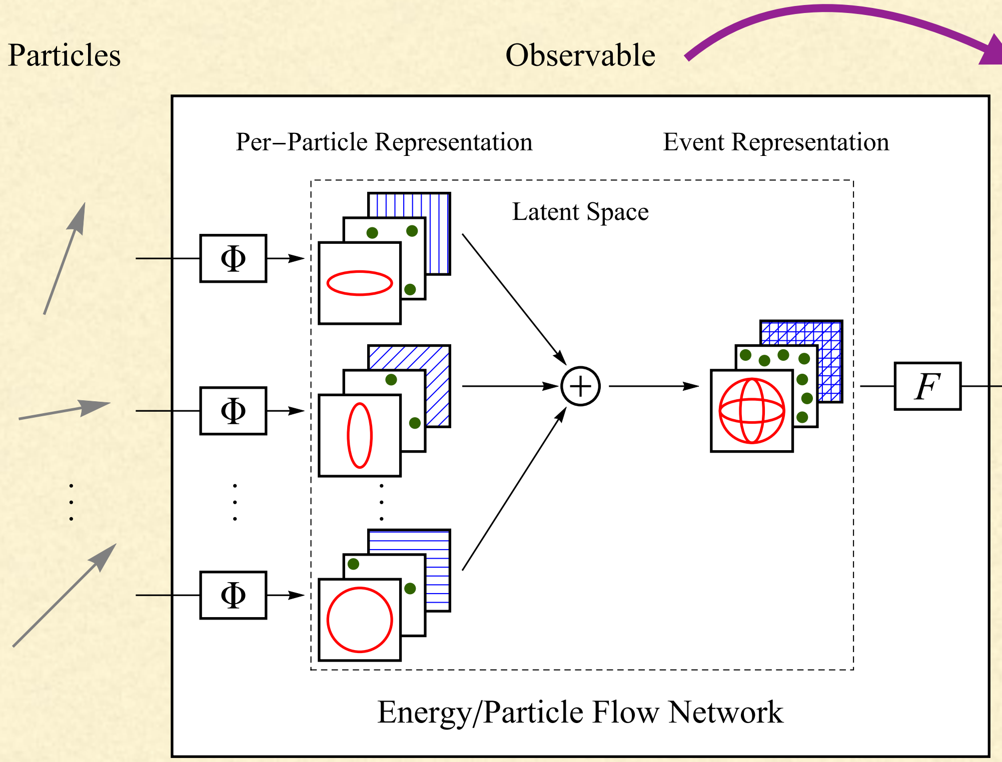
ATLAS-CONF-2019-035

- and for the benefit of us theorists they even provided 1-D projections



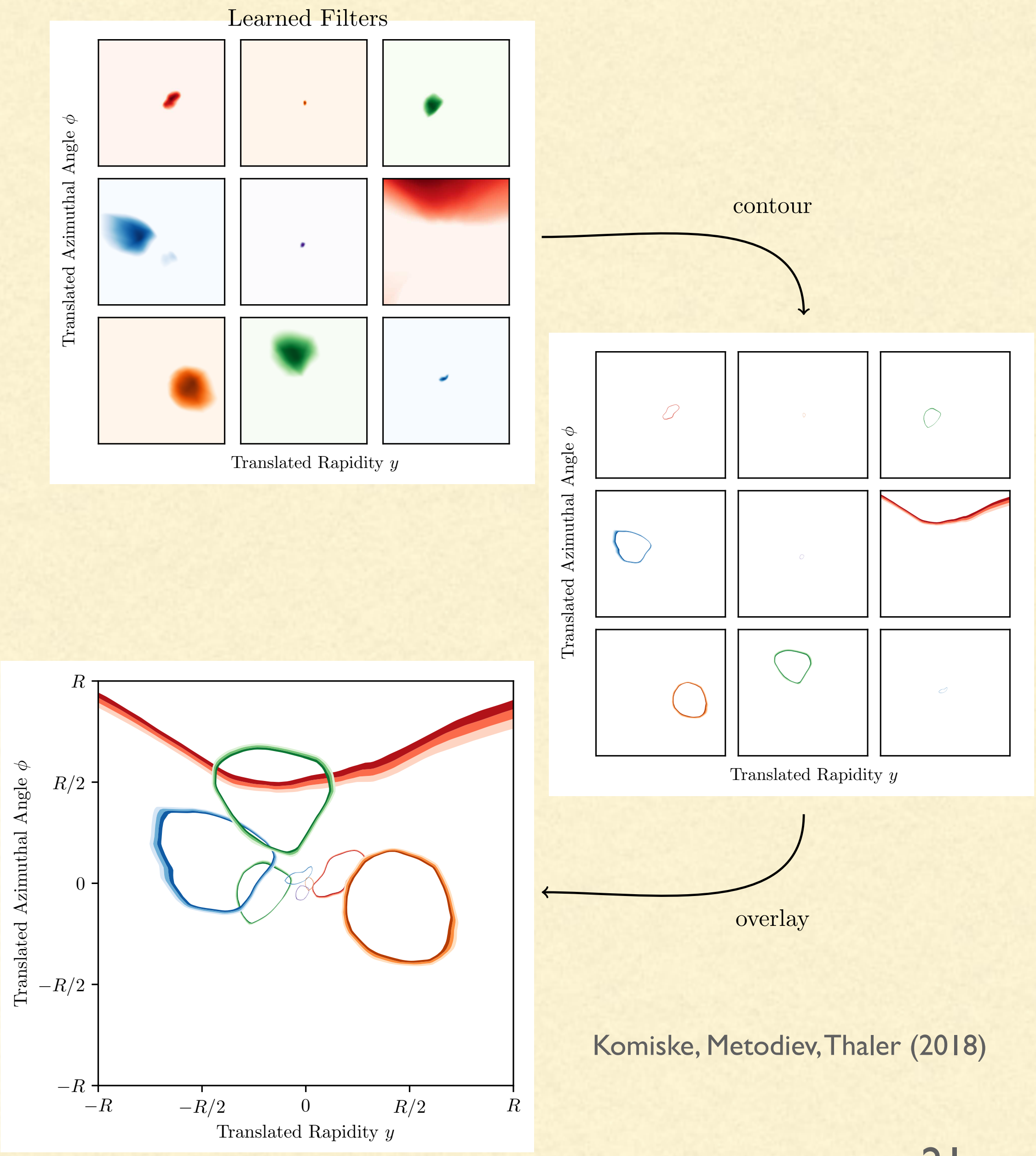


DEEP LEARNING MEETS DEEP THINKING: ENERGY FLOW NET



$$EFN : F\left(\sum_{i=1}^{M} z_i \Phi(\theta_i, \phi_i)\right)$$

$$PFN : F\left(\sum_{i=1}^{M} \Phi(p_i)\right)$$

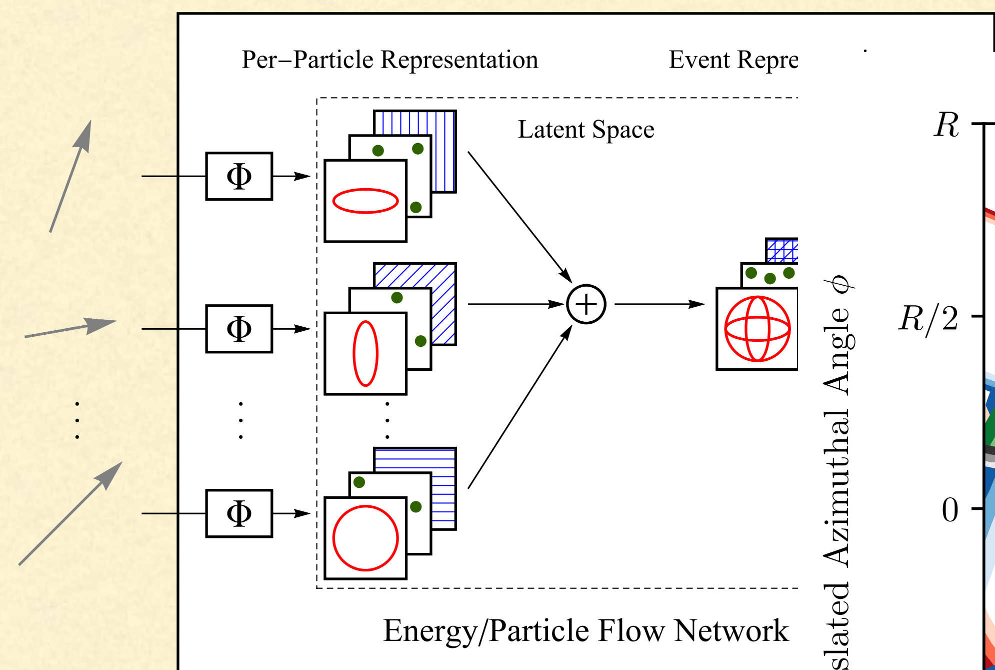


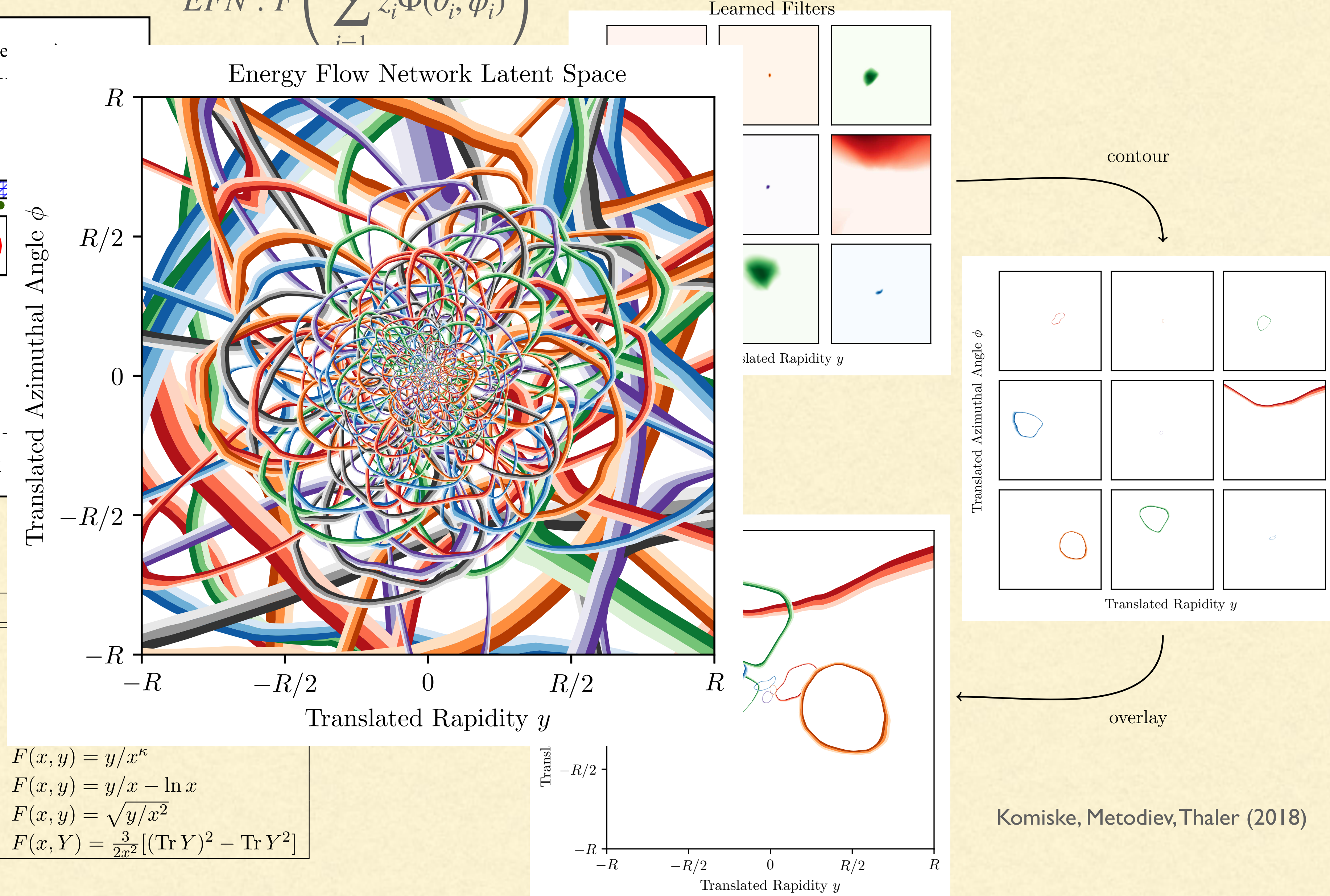
Observable $\mathcal{O}$		Map Φ	Function F
Mass	m	p^μ	$F(x^\mu) = \sqrt{x^\mu x_\mu}$
Multiplicity	M	1	$F(x) = x$
Track Mass	m_{track}	$p^\mu \mathbb{I}_{\text{track}}$	$F(x^\mu) = \sqrt{x^\mu x_\mu}$
Track Multiplicity	M_{track}	$\mathbb{I}_{\text{track}}$	$F(x) = x$
Jet Charge [72]	$\mathcal{Q}_\kappa$	$(p_T, Q\, p_T^\kappa)$	$F(x, y) = y/x^\kappa$
Eventropy [74]	$z \ln z$	$(p_T, p_T \ln p_T)$	$F(x, y) = y/x - \ln x$
Momentum Dispersion [93]	p_T^D	(p_T, p_T^2)	$F(x, y) = \sqrt{y/x^2}$
C parameter [94]	C	$(\lvert\vec{p}\rvert, \vec{p} \otimes \vec{p}/\lvert\vec{p}\rvert)$	$F(x, Y) = \frac{3}{2x^2}[(\operatorname{Tr} Y)^2 - \operatorname{Tr} Y^2]$

Komiske, Metodiev, Thaler (2018)

DEEP LEARNING MEETS DEEP THINKING: ENERGY FLOW NET

$$EFN : F\left(\sum_{i=1}^{M} z_i \Phi(\theta_i, \phi_i)\right)$$



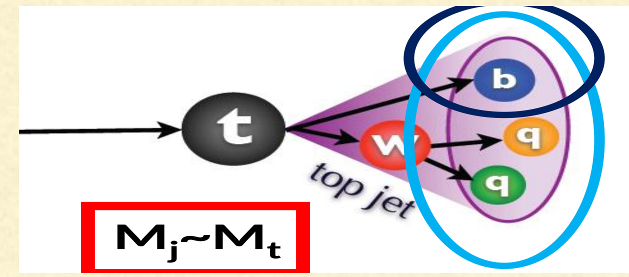


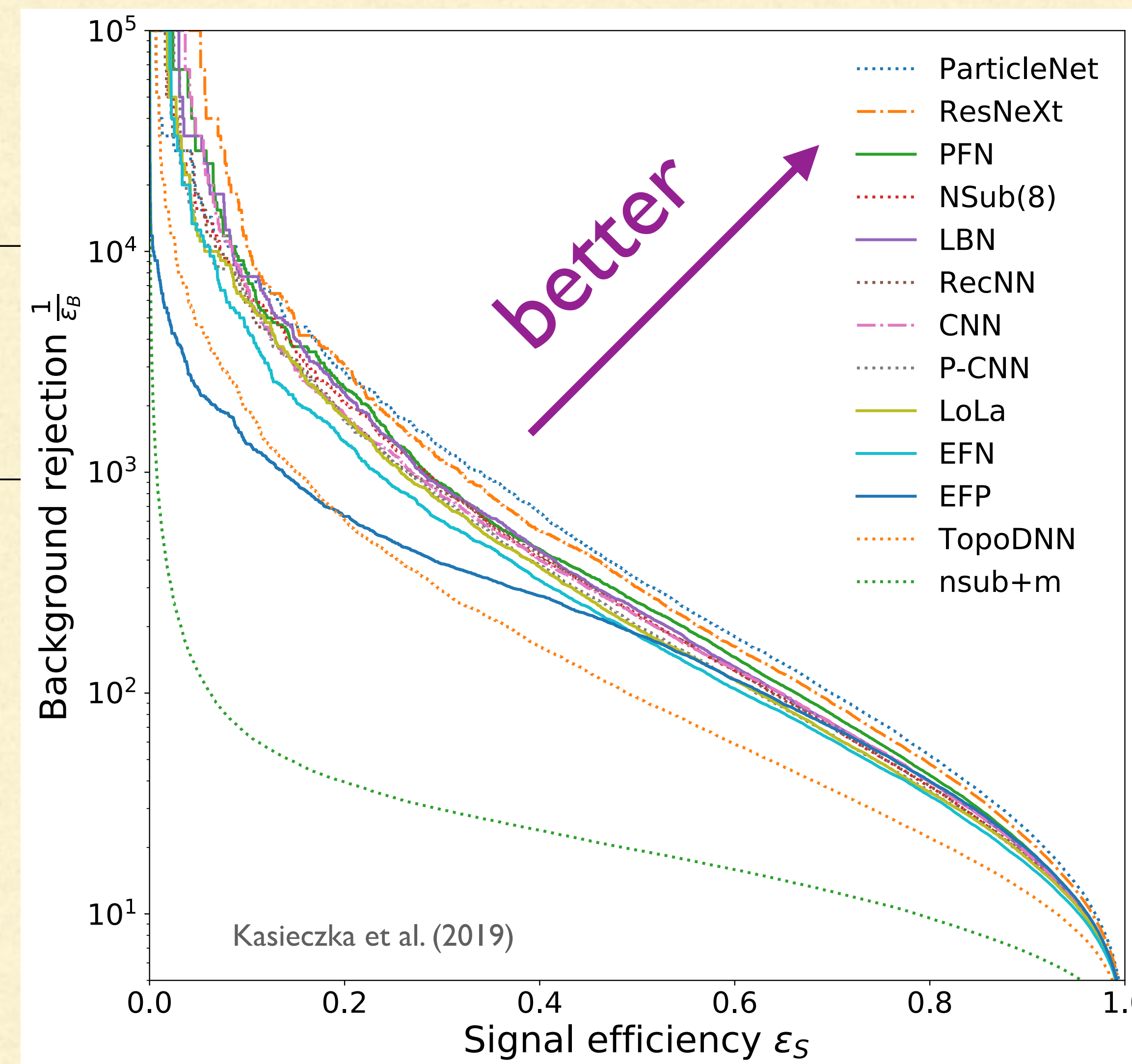
Observable $\mathcal{O}$		Map Φ	
Mass	m	p^μ	
Multiplicity	M	1	
Track Mass	m_{track}	$p^\mu \mathbb{I}_{\text{track}}$	
Track Multiplicity	M_{track}	$\mathbb{I}_{\text{track}}$	
Jet Charge [72]	$\mathcal{Q}_\kappa$	$(p_T, Q\, p_T^\kappa)$	$F(x,y) = y/x^\kappa$
Eventropy [74]	$z \ln z$	$(p_T, p_T \ln p_T)$	$F(x,y) = y/x - \ln x$
Momentum Dispersion [93]	p_T^D	(p_T, p_T^2)	$F(x,y) = \sqrt{y/x^2}$
C parameter [94]	C	$(\lvert\vec{p}\rvert, \vec{p}\otimes\vec{p}/\lvert\vec{p}\rvert)$	$F(x,Y) = \frac{3}{2x^2}[(\operatorname{Tr} Y)^2 - \operatorname{Tr} Y^2]$

Komiske, Metodiev, Thaler (2018)

ML SURVEY FOR TOP TAGGING

		AUC	Accuracy	$1/\epsilon_B$ ($\epsilon_S = 0.3$)	#Parameters
images	CNN [16]	0.981	0.930	780	610k
	ResNeXt [32]	0.984	0.936	1140	1.46M
four-momenta	TopoDNN [18]	0.972	0.916	290	59k
	Multi-body N-subjettiness 6 [24]	0.979	0.922	856	57k
	Multi-body N-subjettiness 8 [24]	0.981	0.929	860	58k
	RecNN	0.981	0.929	810	13k
	P-CNN	0.980	0.930	760	348k
	ParticleNet [45]	0.985	0.938	1280	498k
theory-inspired	LBN [19]	0.981	0.931	860	705k
	LoLa [22]	0.980	0.929	730	127k
	Energy Flow Polynomials [21]	0.980	0.932	380	1k
	Energy Flow Network [23]	0.979	0.927	600	82k
	Particle Flow Network [23]	0.982	0.932	880	82k





- all solutions offer big improvement over standard analysis (nsub+m)
- similar performances
- physics intuition useful to match performance of highly-sophisticated architectures

TOWARDS UNDERSTANDING

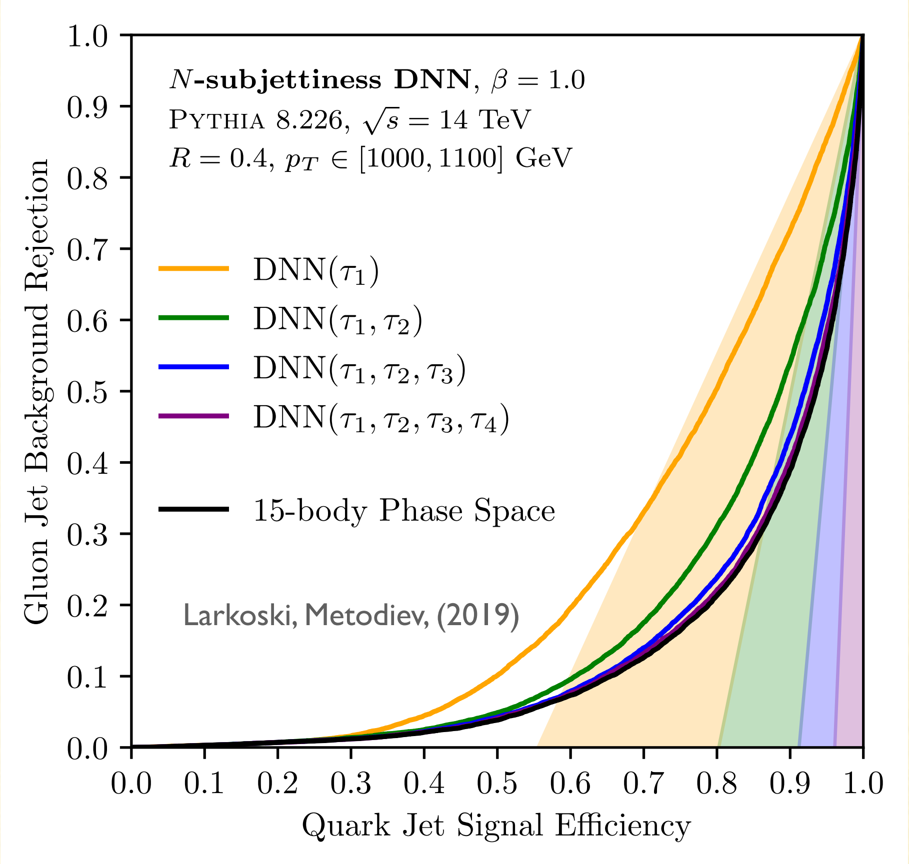
- ML techniques do bring significant improvement but also many questions
- Theory community (within and outside jet physics) reacted in different ways
- Recently first attempts to “open the black box” have appeared
- Calculable (IRC safe) input allows for (some) first-principle understanding

TOWARDS UNDERSTANDING

- ML techniques do bring significant improvement but also many questions
- Theory community (within and outside jet physics) reacted in different ways
- Recently first attempts to "open the black box" have appeared
- Calculable (IRC safe) input allows for (some) first-principle understanding

- Theory of q/g discrimination studied using N-subjettines variables
- Likelihood ratios, ROC, reducibility factors can be computed
- A bound on the Area Under the Curve can be obtained

$$AUC \geq \frac{\kappa_S + \kappa_B - 2\kappa_S\kappa_B}{2 - \kappa_S\kappa_B} = \left(\frac{C_F}{C_A}\right)^n$$



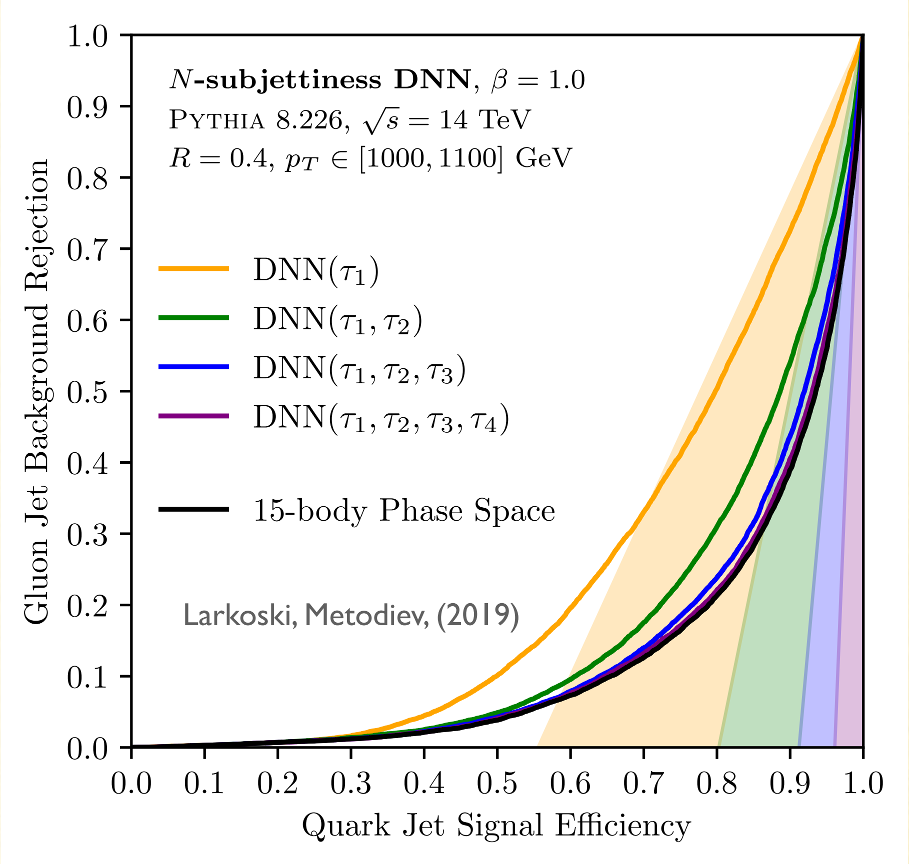
TOWARDS UNDERSTANDING

- ML techniques do bring significant improvement but also many questions
- Theory community (within and outside jet physics) reacted in different ways
- Recently first attempts to "open the black box" have appeared
- Calculable (IRC safe) input allows for (some) first-principle understanding

- Theory of q/g discrimination studied using N-subjettines variables
- Likelihood ratios, ROC, reducibility factors can be computed
- A bound on the Area Under the Curve can be obtained

$$AUC \geq \frac{\kappa_S + \kappa_B - 2\kappa_S\kappa_B}{2 - \kappa_S\kappa_B} = \left(\frac{C_F}{C_A}\right)^n$$

- "a first step in a theoretical effort to deconstruct machine learning for particle physics"



CONCLUSIONS & OUTLOOK

- What is needed to boost sensitivity to Hbb?
- Are traditional tools/approach sufficient or do we need to resort to ML?
- In the context of ML, are we suspicious of black-boxes? Should we?
  - can we move from machine-learning to learning-from-machines? Interpretable neural networks? Prescriptive analytics?
  - can we devise ML learning algorithms that preserve calculability? (jet topics, grooming through reinforcement learning …)
- What's the best use of first-principle knowledge in jet physics?
  - extraction of SM parameters? PDFs with q/g tagging?
  - jet substructure probes of quark-gluon plasma in heavy ion collisions

(there are links to things I hadn't time to discuss)

CONCLUSIONS & OUTLOOK

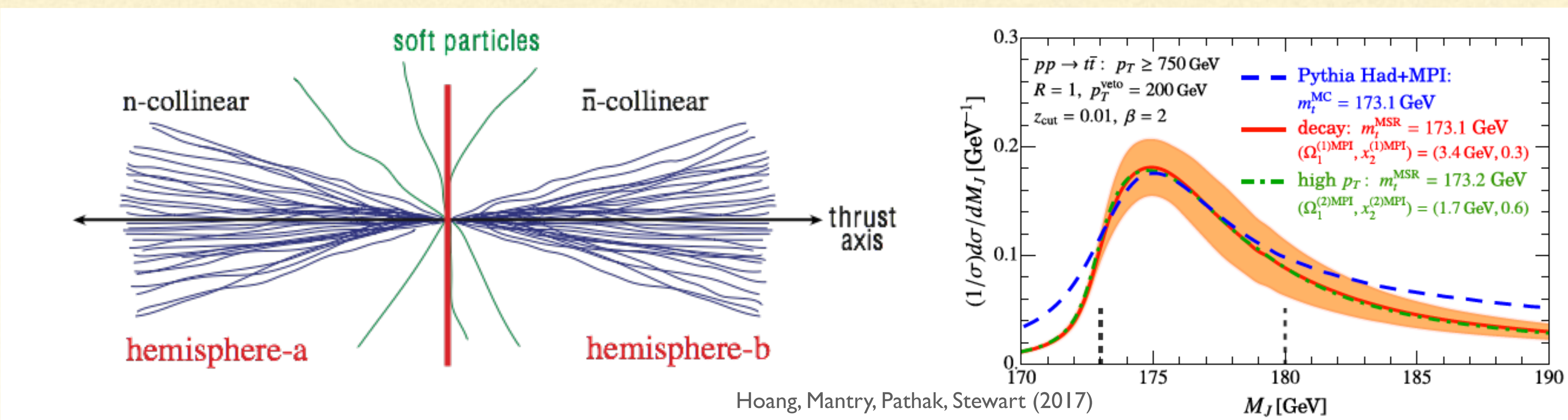
- What is needed to boost sensitivity to Hbb?
- Are traditional tools/approach sufficient or do we need to resort to ML?
- In the context of ML, are we suspicious of black-boxes? Should we?
  - can we move from machine-learning to learning-from-machines? Interpretable neural networks? Prescriptive analytics?
  - can we devise ML learning algorithms that preserve calculability? (jet topics, grooming through reinforcement learning …)
- What's the best use of first-principle knowledge in jet physics?
  - extraction of SM parameters? PDFs with q/g tagging?
  - jet substructure probes of quark-gluon plasma in heavy ion collisions

(there are links to things I hadn't time to discuss) **THANK YOU !**

BACKUP SLIDES

TOP MASS WITH SOFT-DROP JETS

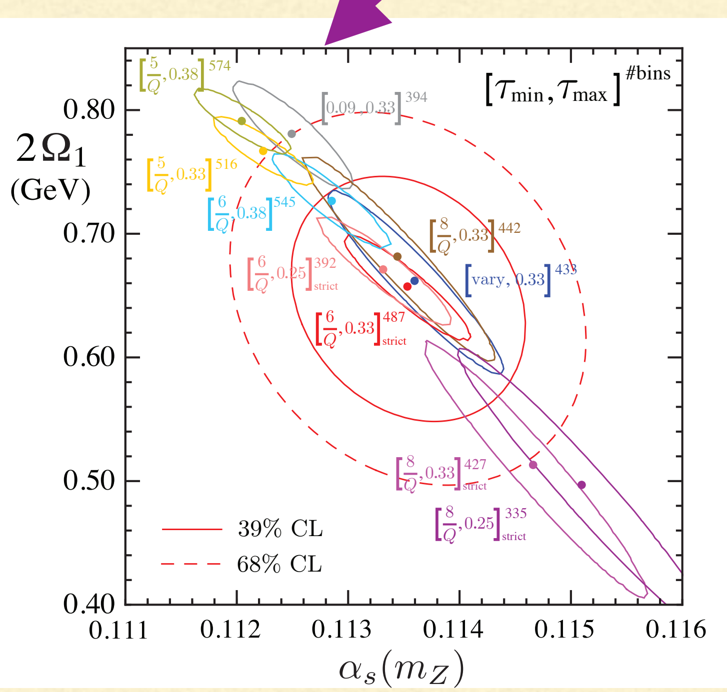
- determination of other fundamental parameters may benefit from grooming, e.g. the top quark mass
- in the context of e^+e^- collisions SCET factorisation theorems allow for a precision-determination of the top-jet mass
- the picture at *pp* collisions is polluted by wide-angle soft radiation
- grooming "turns" *pp* observables into e^+e^- ones

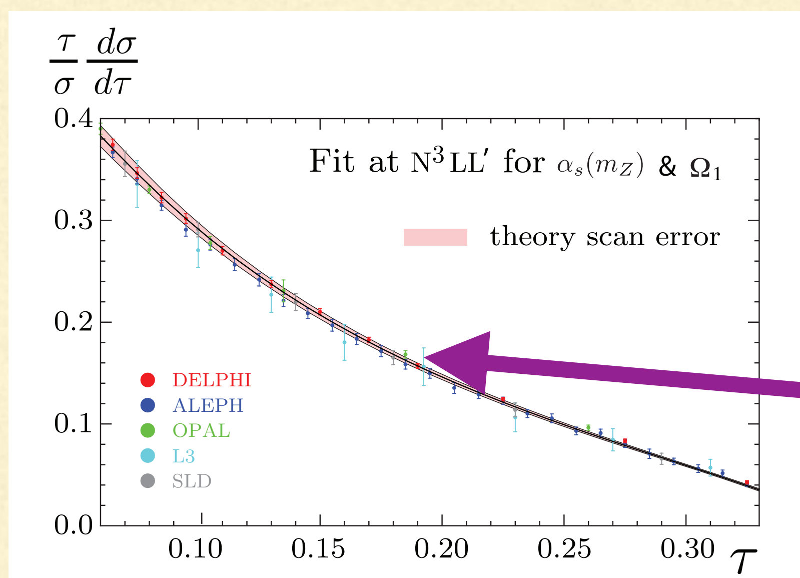


Hoang, Mantry, Pathak, Stewart (2017)

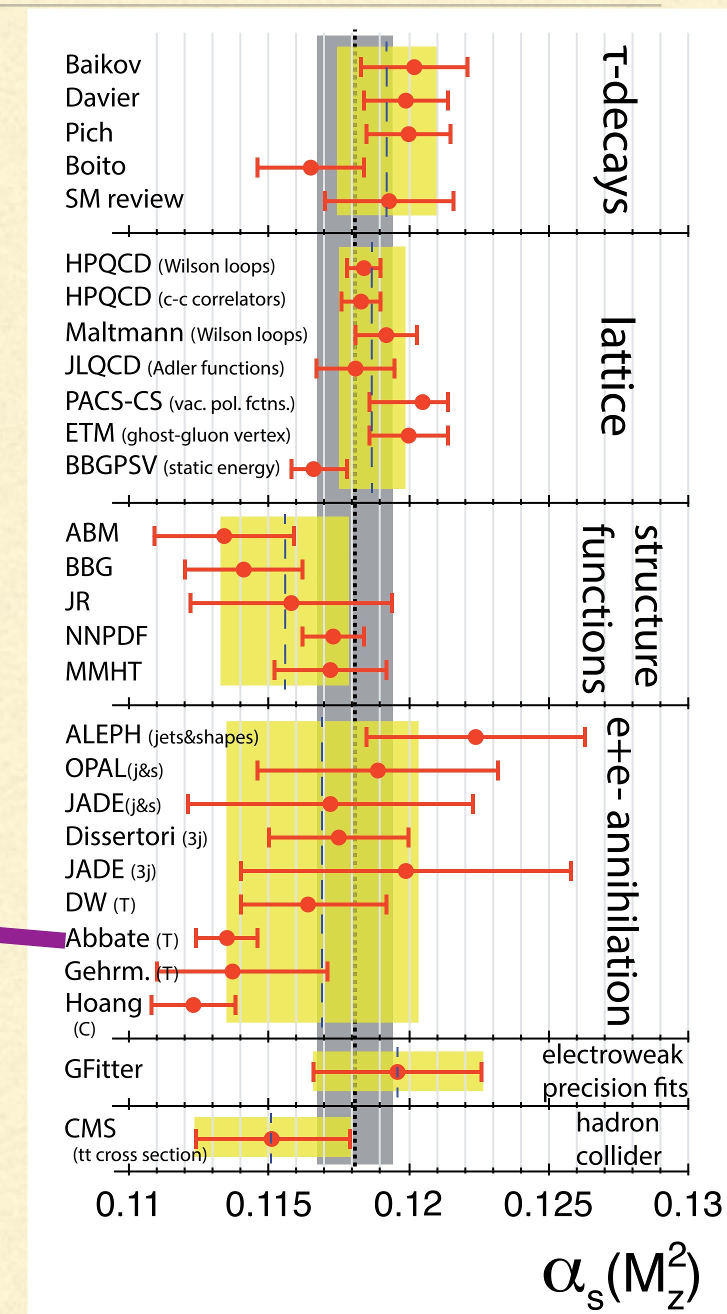
MEASURING THE STRONG COUPLING

- current precision below 1%, dominated by lattice extractions
- LEP event shapes also very precise (5%)
- however they are in tension with the world average
- thrust (and C parameter) known with outstanding accuracy

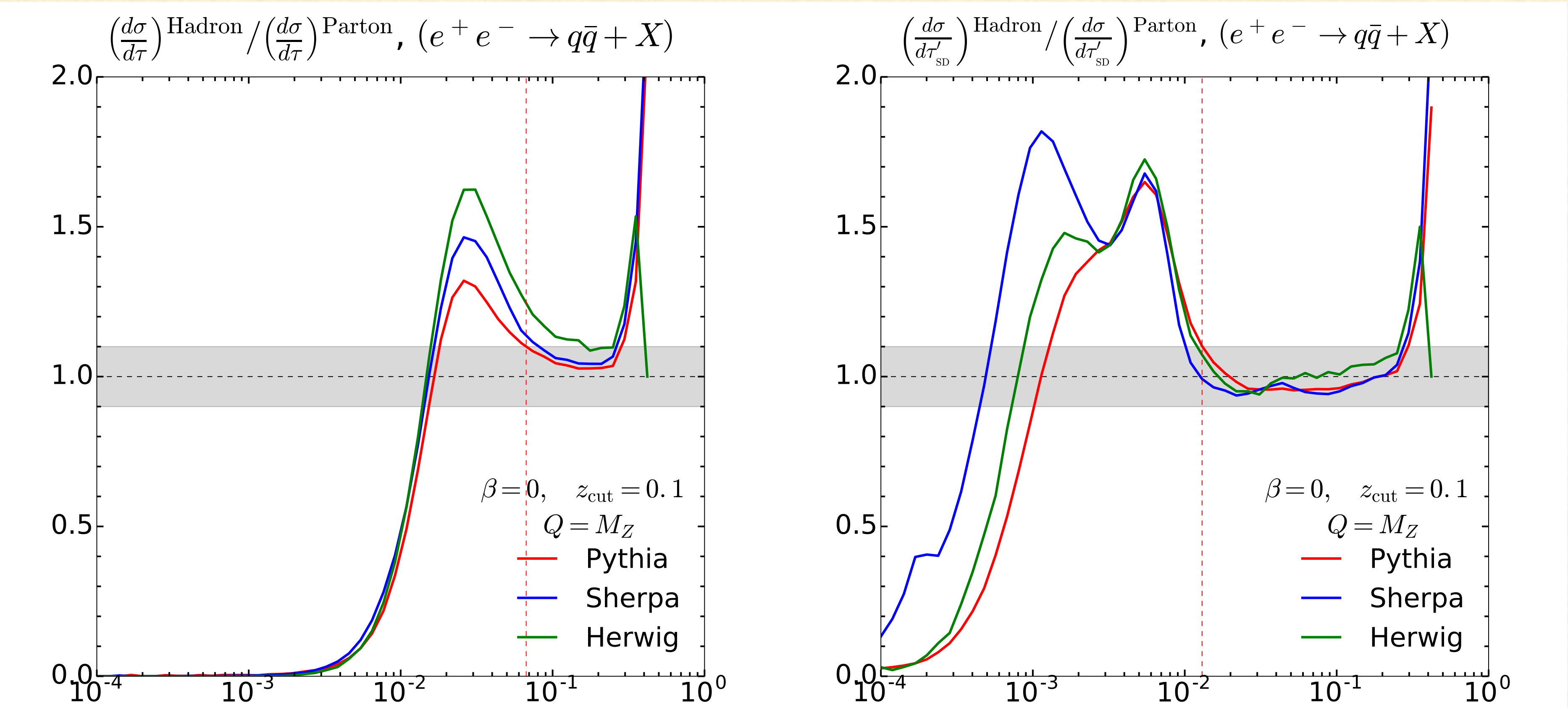




strong correlation with non-perturbative parameter



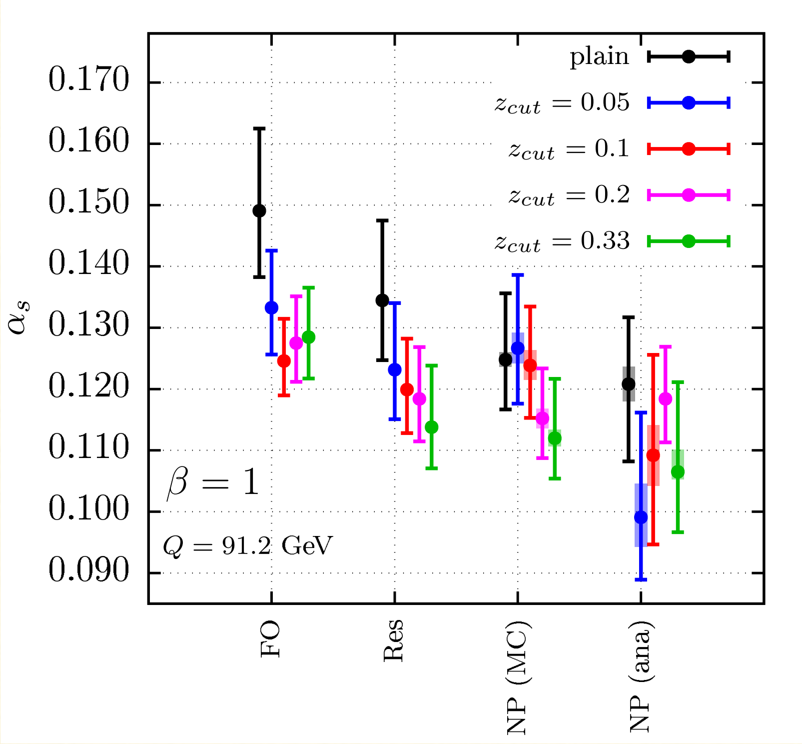
SOFT-DROP EVENT SHAPES



Baron, SM, Theeuwes (2018)

- noticeable reduction of non-pert. corrections may allow to disentangle the degeneracy
- can we compute it at the same accuracy as standard event shapes?
- NNLO calculations recently performed Kardos, Somogyi, Trocsanyi (2018)

α_s WITH SOFT-DROP THRUST



- fits to pseudo-data generated by SHERPA
- results shows reduced dependence on non-pert. corrections
- subleading effects are under investigation

SM, Reichelt, Schumann, Soyez, and Theeuwes (2019)

- soft-drop allows us to extend the fit range
- Generale question: is there a natural way to define soft-drop event shapes? e.g. bottom-up soft-drop

Dreyer, Necib, Soyez, Thaler (2018)
Baron (in preparation)