Optimizing PYTHIA 8 event generation with RHIC data

Raghav Kunnawalkam Elayavalli (Yale/BNL)

BNL Nuclear Physics seminar

Nov 9th, 2021

Aguilar, Chang, RKE, Fatemi, He, Ji, Kalinkin,
Kelsey, Mooney and Verkest

2110.09447

Wright Laboratory

raghavke.me

Brookhaven National Laboratory

Introduction / Motivation

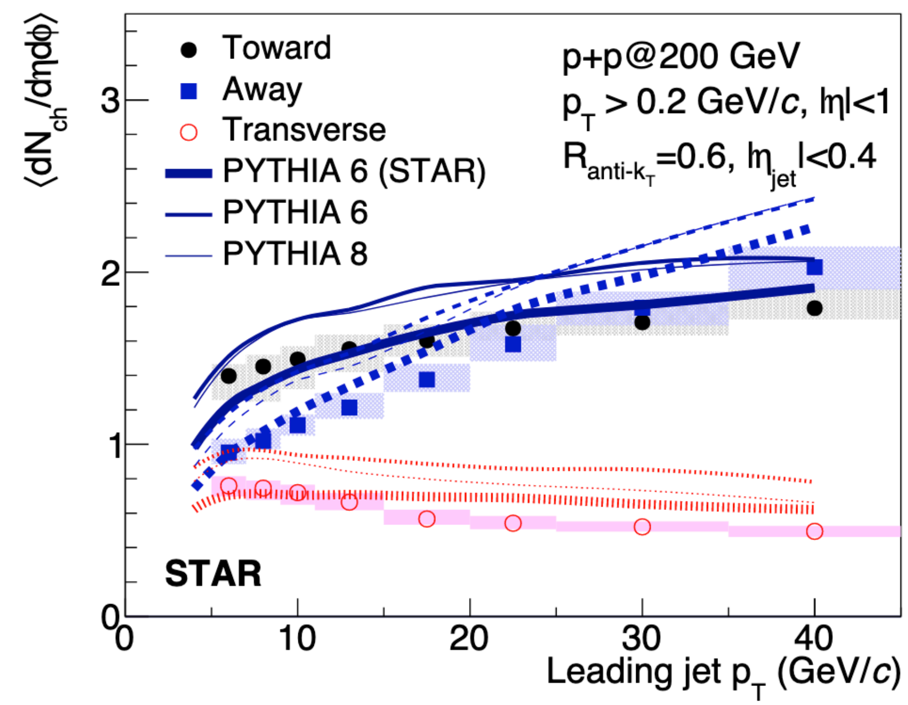
Tuning Procedure

Detroit Tune

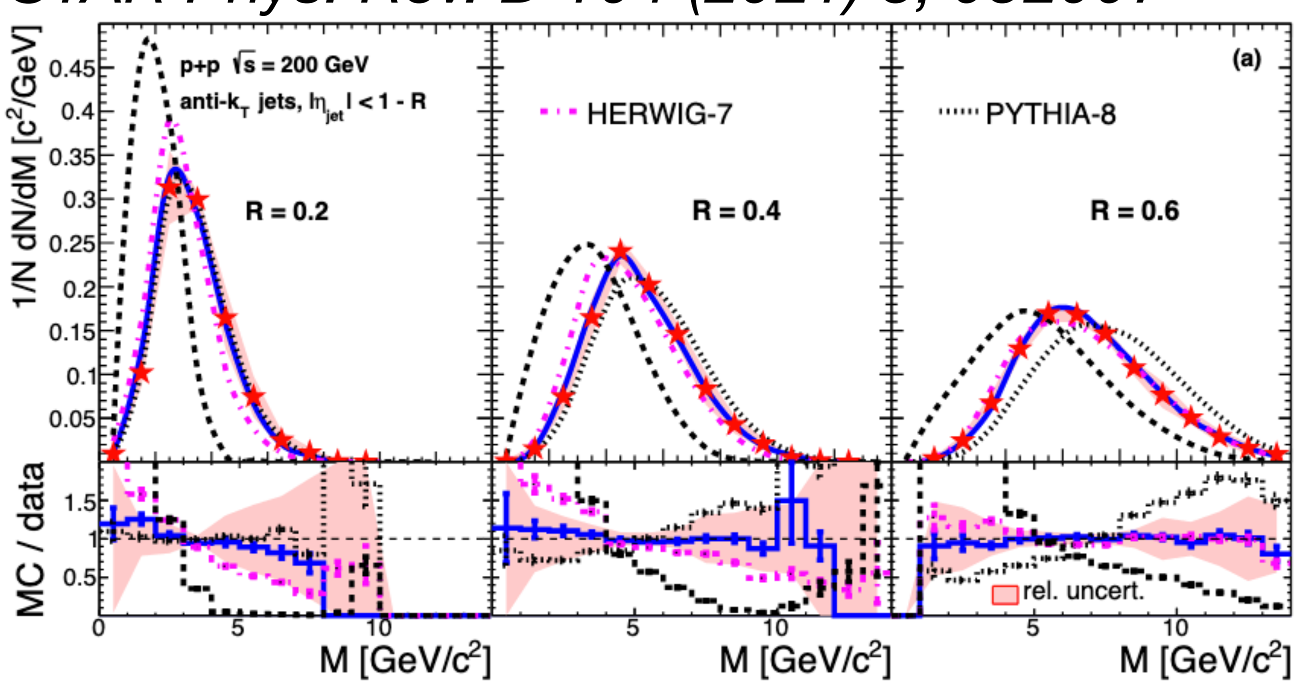
Outlook

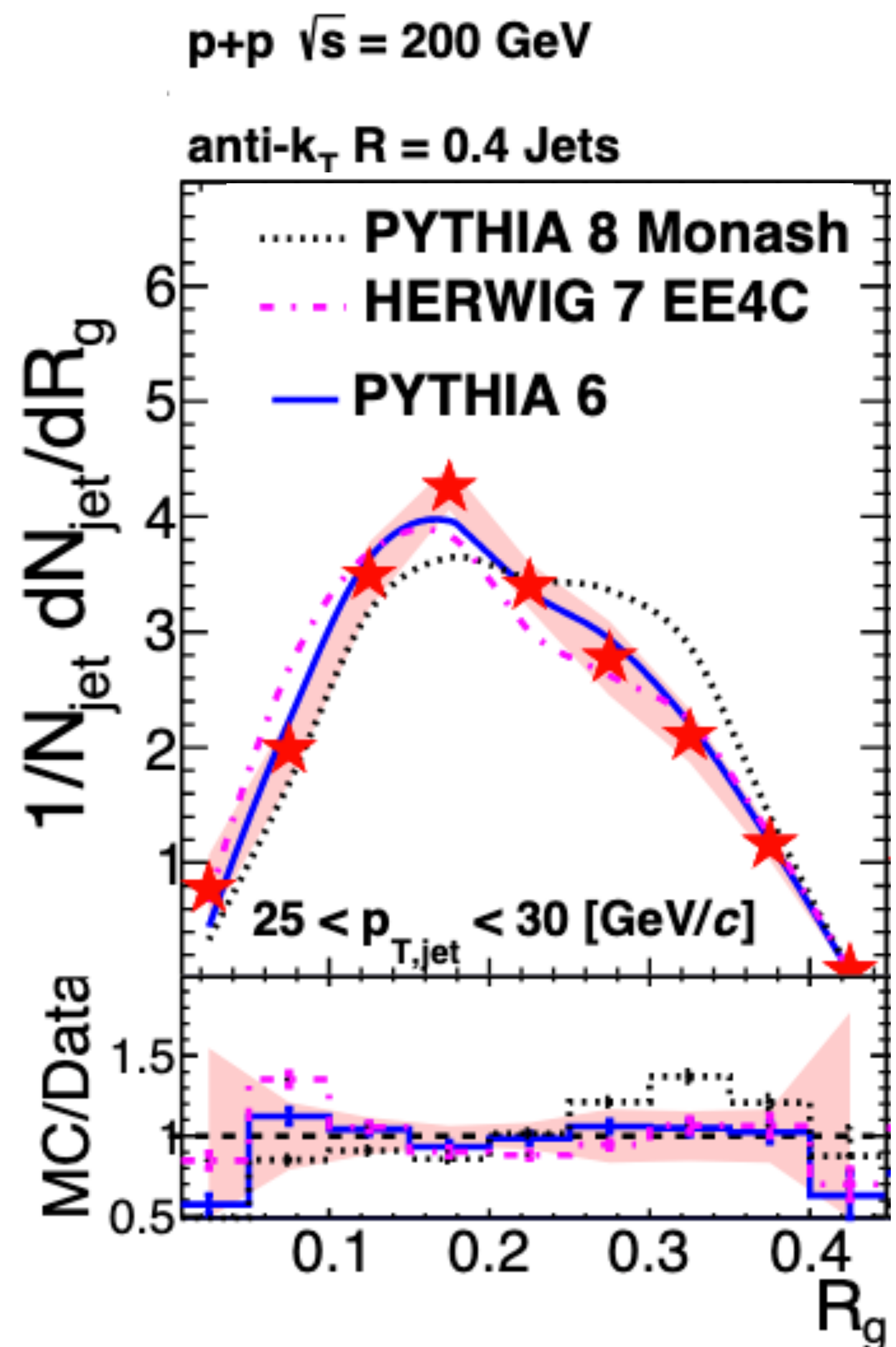
Comparison of recent RHIC data to PYTHIA tunes

STAR Phys. Rev. D 101 (2020) 5, 052004



STAR Phys. Rev. D 104 (2021) 5, 052007

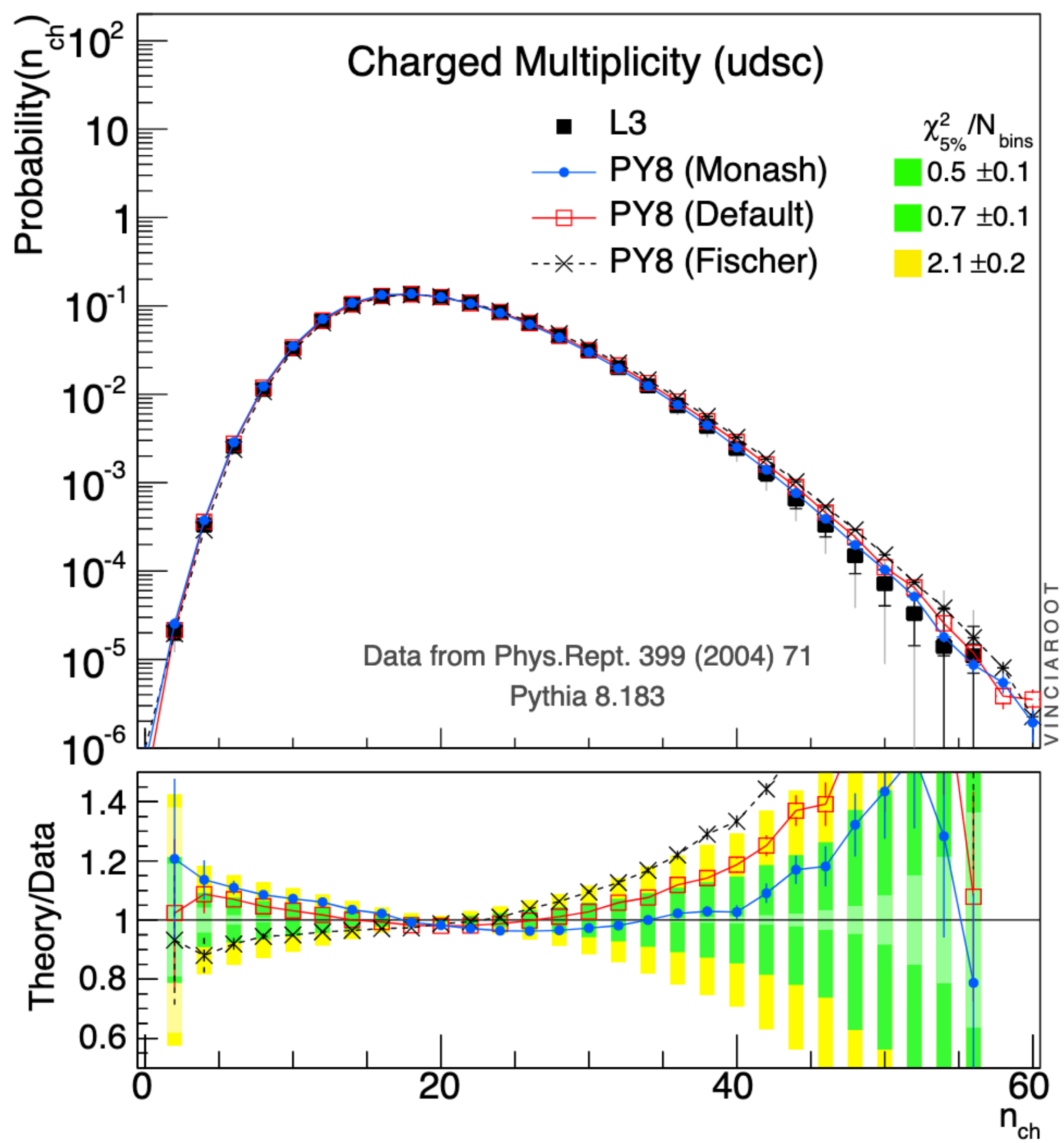


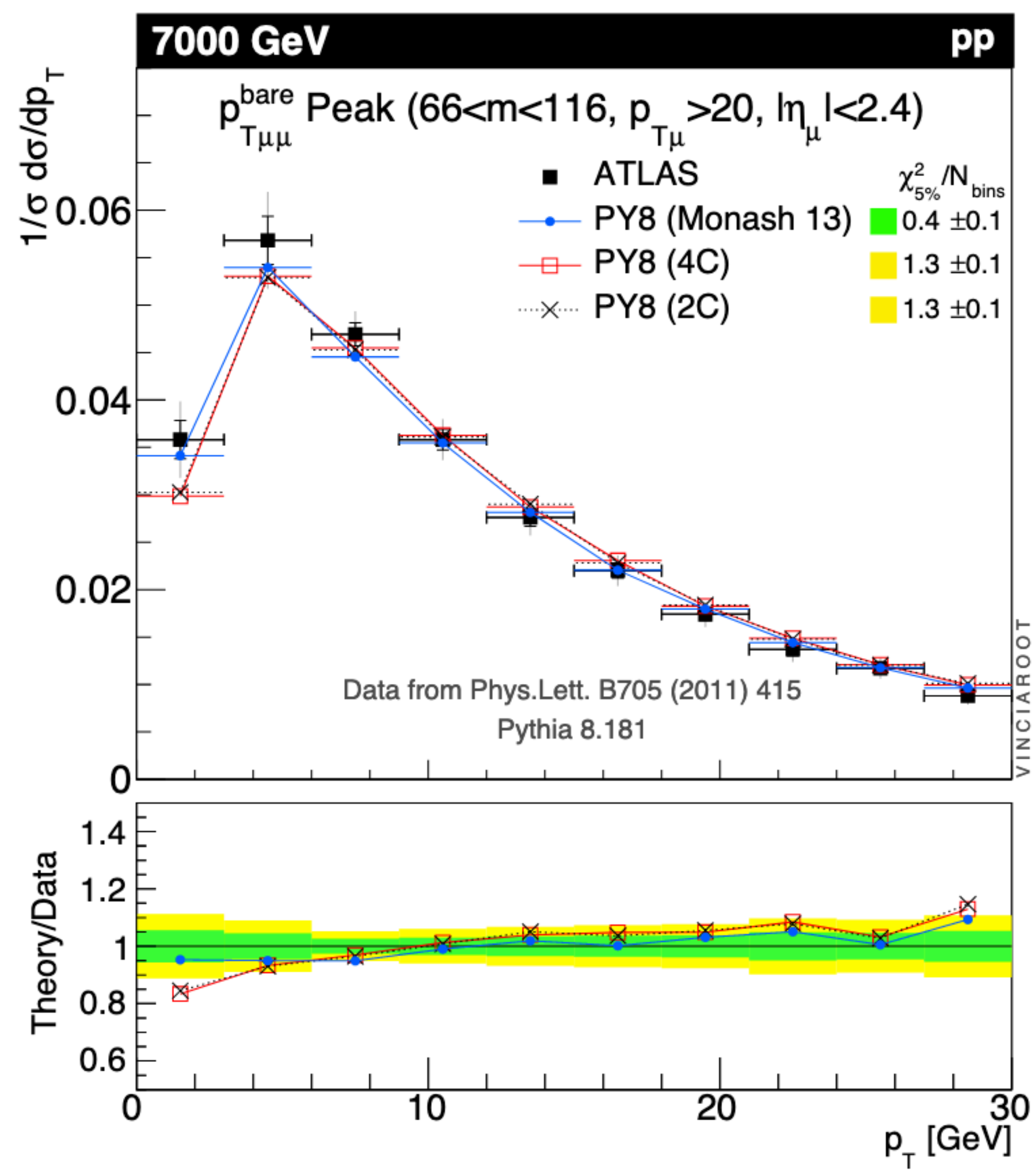


STAR Phys. Lett. B 811 (2020) 135846

PYTHIA 8 Monash Tune

Skands, Carrazza and Rojo Eur.Phys.J.C 74 (2014) 8, 3024






```
StringZ:aLund          =  0.68
StringZ:bLund          =  0.98
StringZ:aExtraDiquark  =  0.97
StringZ:aExtraSquark   =  0.00
```

```
StringPT:sigma              = 0.335
StringPT:enhancedFraction   = 0.01
StringPT:enhancedWidth      = 2.0
```

STAR PYTHIA 6 Tune

$$p_{T,0} = p_{T,0}^{Ref} \left(\frac{\sqrt{s}}{\sqrt{s_{Ref}}} \right)^{ecmPow}$$

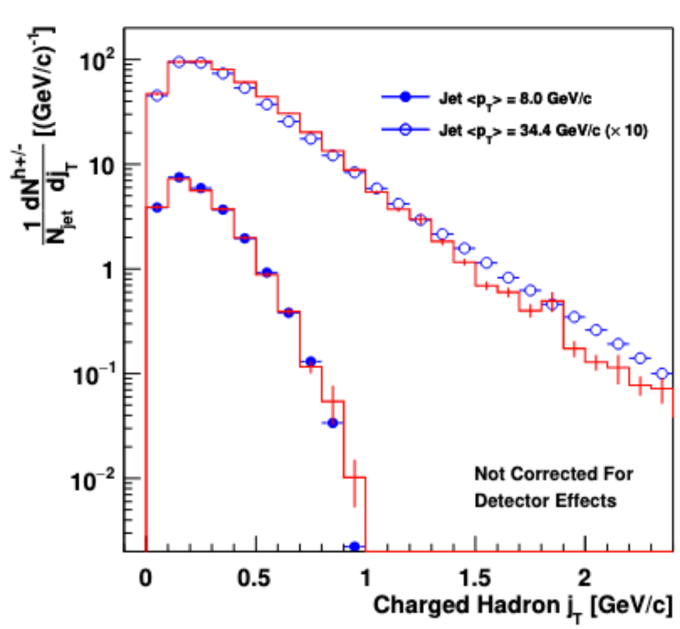
Adjusted power law extrapolation parameter (PARP(90)=0.24→ 0.213) to match low-p_T yields

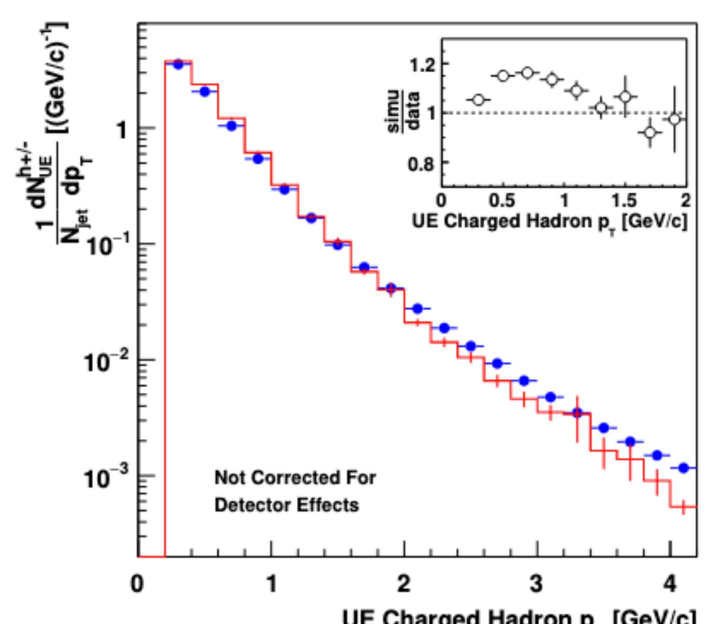
STAR PYTHIA 6 Tune

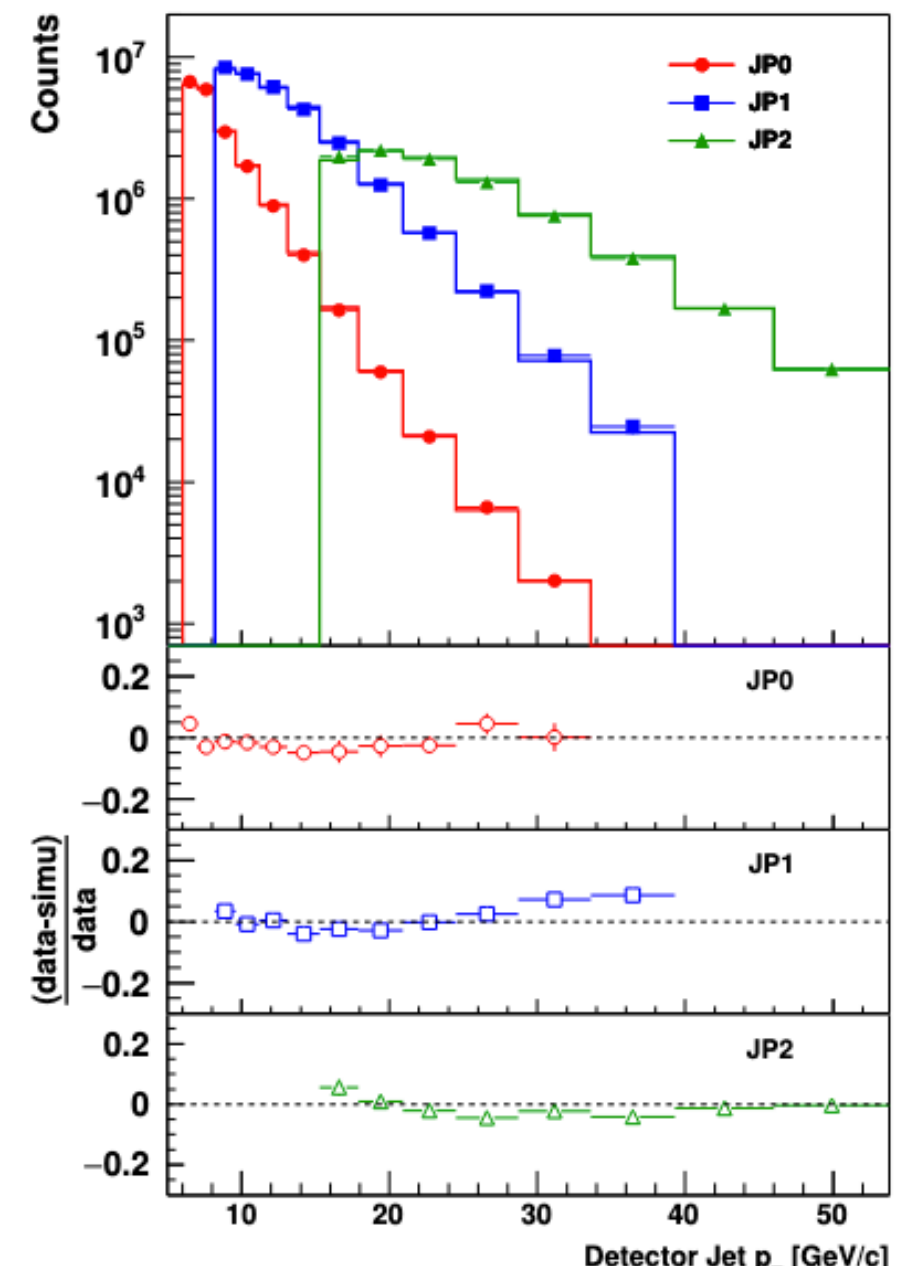
$$p_{T,0} = p_{T,0}^{Ref}\left(\frac{\sqrt{s}}{\sqrt{s_{Ref}}}\right)^{ecmPow}$$

Adjusted power law extrapolation parameter (PARP(90)=0.24→ 0.213) to match low-p_T yields

STAR Phys.Rev.D 100 (2019) 5, 052005







Recent tunes of PYTHIA-8 w/ CMS pp Data

PYTHIA8 parameter	CP1	CP2
PDF Set	NNPDF3.1 LO	NNPDF3.1 LO
$\alpha_S(m_Z)$	0.130	0.130
`SpaceShower:rapidityOrder`	off	off
`MultipartonInteractions:EcmRef` [GeV]	7000	7000
$\alpha_S^{ISR}(m_Z)$ value/order	0.1365/LO	0.130/LO
$\alpha_S^{FSR}(m_Z)$ value/order	0.1365/LO	0.130/LO
$\alpha_S^{MPI}(m_Z)$ value/order	0.130/LO	0.130/LO
$\alpha_S^{ME}(m_Z)$ value/order	0.130/LO	0.130/LO
`MultipartonInteractions:pT0Ref` [GeV]	2.4	2.3
`MultipartonInteractions:ecmPow`	0.15	0.14
`MultipartonInteractions:coreRadius`	0.54	0.38
`MultipartonInteractions:coreFraction`	0.68	0.33
`ColorReconnection:range`	2.63	2.32
χ^2/dof	0.89	0.54

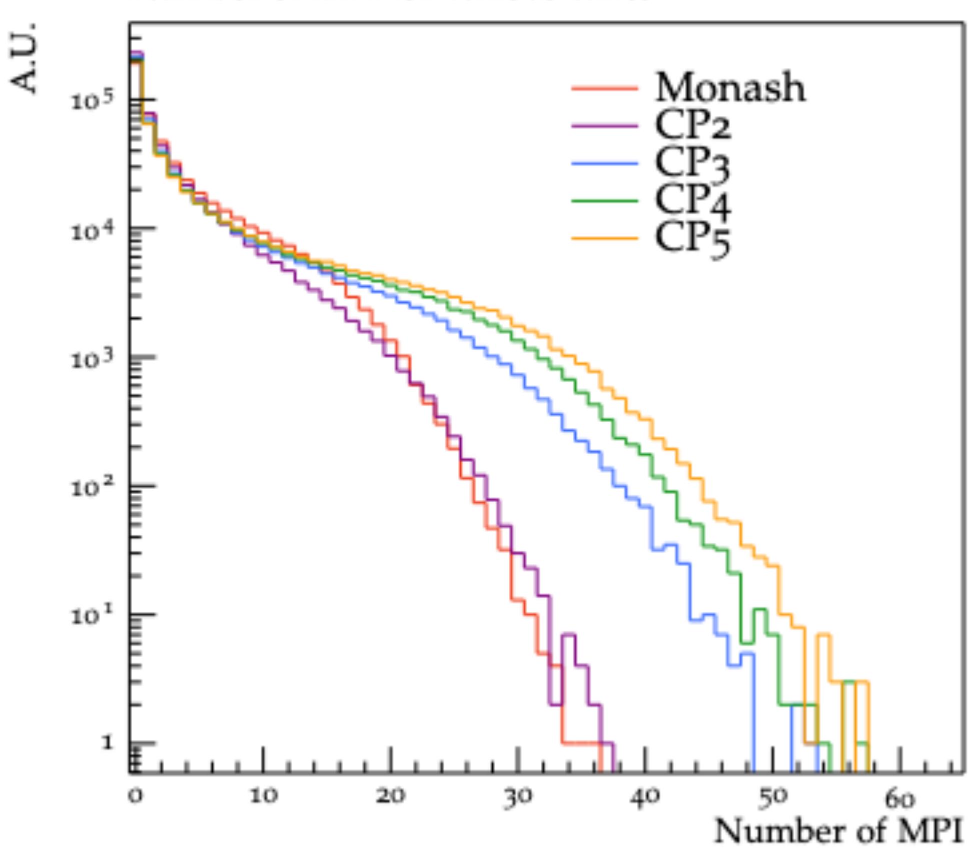
CMS Eur. Phys. J. C 80 (2020) 1, 4

Recent tunes of PYTHIA-8 w/ CMS pp Data

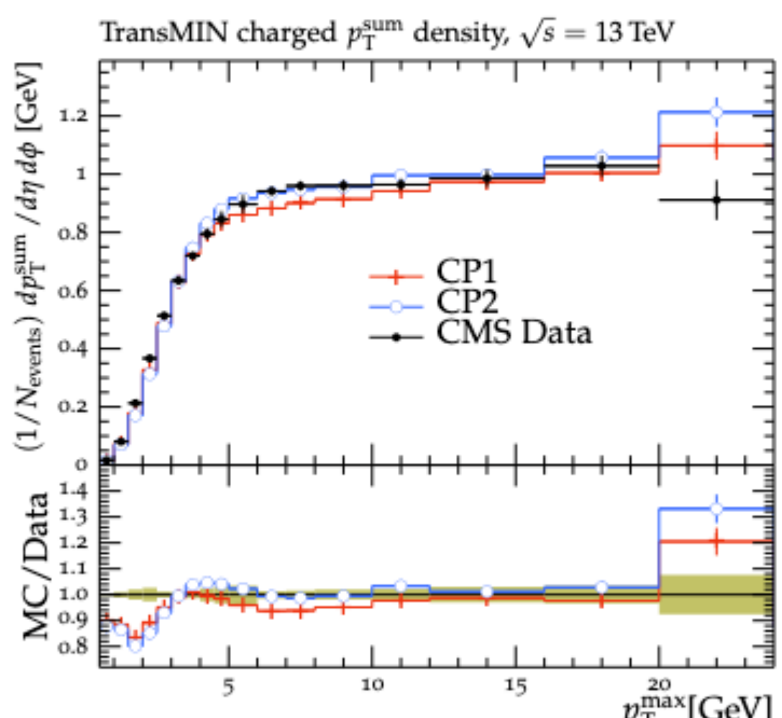
PYTHIA8 parameter	CP1	CP2	CP3	CP4	CP5
PDF Set	NNPDF3.1 LO	NNPDF3.1 LO	NNPDF3.1 NLO	NNPDF3.1 NNLO	NNPDF3.1 NNLO
$\alpha_S(m_Z)$	0.130	0.130	0.118	0.118	0.118
`SpaceShower:rapidityOrder`	off	off	off	off	on
`MultipartonInteractions:EcmRef` [GeV]	7000	7000	7000	7000	7000
$\alpha_S^{ISR}(m_Z)$ value/order	0.1365/LO	0.130/LO	0.118/NLO	0.118/NLO	0.118/NLO
$\alpha_S^{FSR}(m_Z)$ value/order	0.1365/LO	0.130/LO	0.118/NLO	0.118/NLO	0.118/NLO
$\alpha_S^{MPI}(m_Z)$ value/order	0.130/LO	0.130/LO	0.118/NLO	0.118/NLO	0.118/NLO
$\alpha_S^{ME}(m_Z)$ value/order	0.130/LO	0.130/LO	0.118/NLO	0.118/NLO	0.118/NLO
`MultipartonInteractions:pT0Ref` [GeV]	2.4	2.3	1.52	1.48	1.41
`MultipartonInteractions:ecmPow`	0.15	0.14	0.02	0.02	0.03
`MultipartonInteractions:coreRadius`	0.54	0.38	0.54	0.60	0.76
`MultipartonInteractions:coreFraction`	0.68	0.33	0.39	0.30	0.63
`ColorReconnection:range`	2.63	2.32	4.73	5.61	5.18
χ^2/dof	0.89	0.54	0.76	0.80	1.04

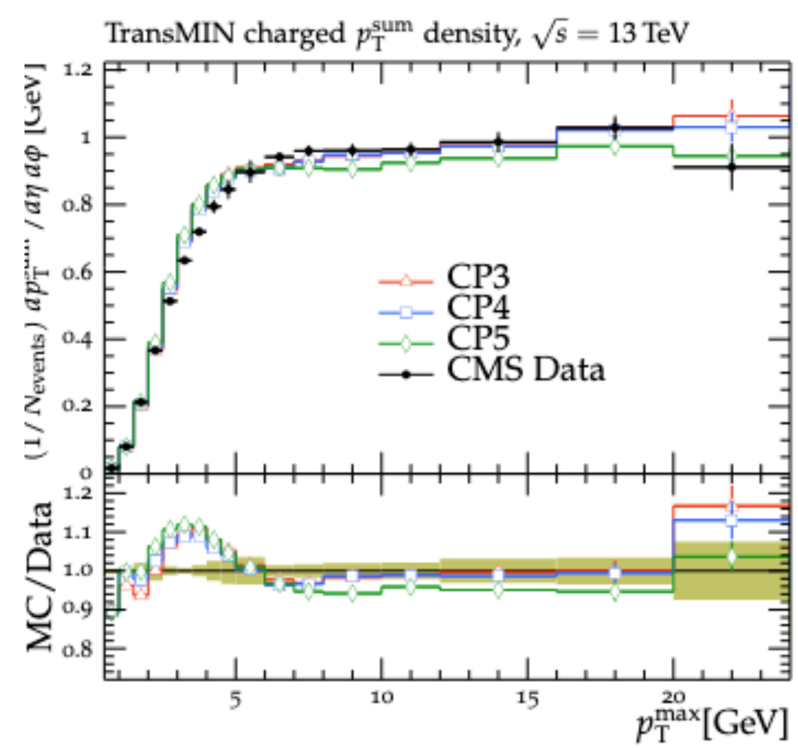
CMS Eur. Phys. J. C 80 (2020) 1, 4

Number of MPI for various tunes

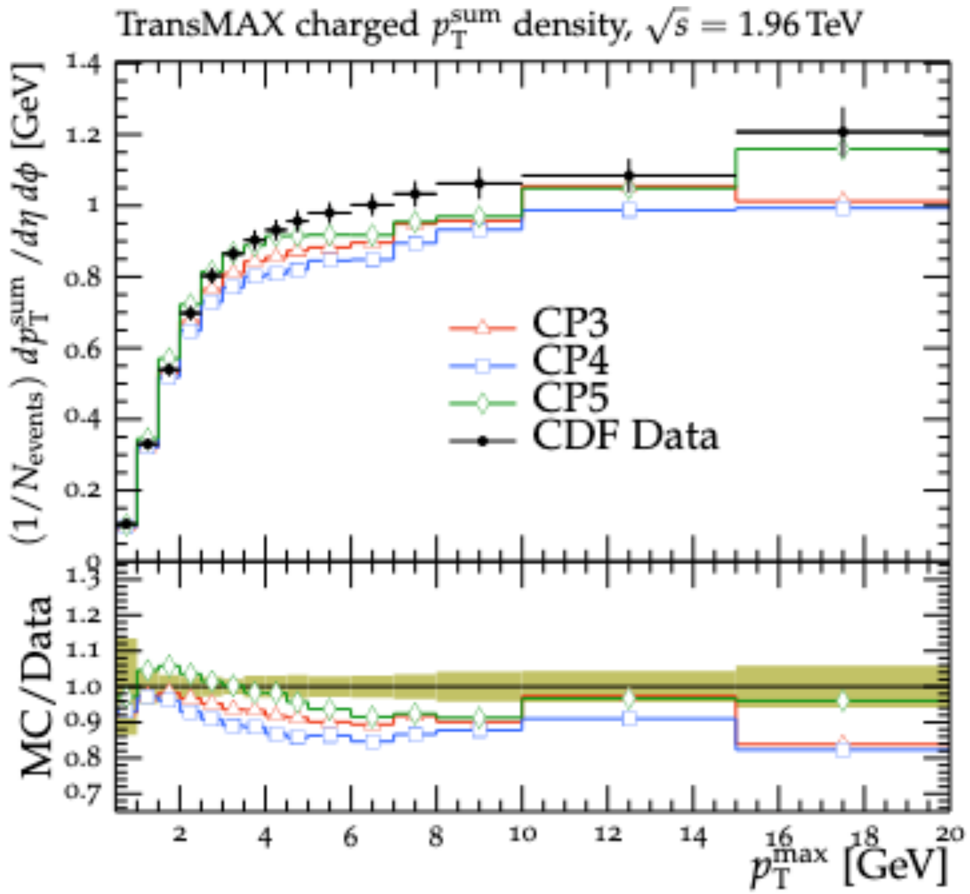


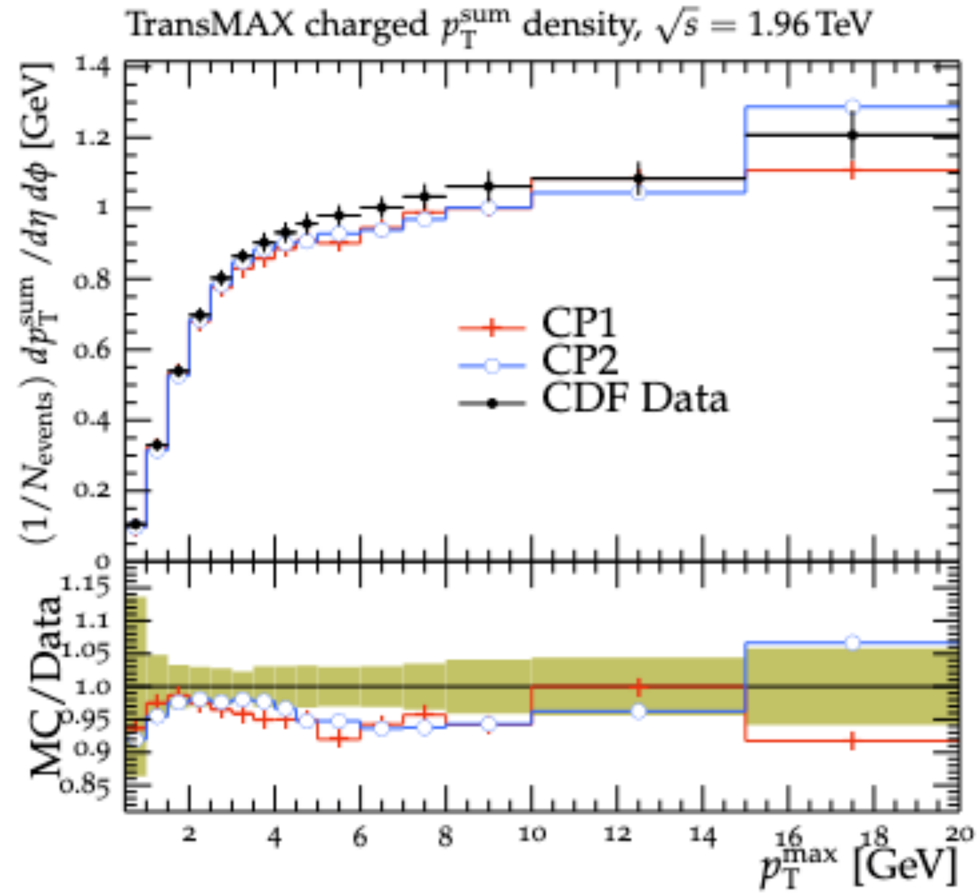
$$\sqrt{s} = 13 \text{ TeV}$$





$\sqrt{s} = 1.96$ TeV





CMS Eur. Phys. J. C 80 (2020) 1, 4

- Parameters involved in the generation of multi-parton interactions (MPI) govern particle production in the underlying event (UE)
- Regularization of the $p_{T,0}^{\mathrm{Ref}}$ parameter controls the handoff from parton shower to MPI
- Reference energy usually set to $\sqrt{s} = 7$ TeV - extrapolation down to 200 GeV is often not considered
- Collision energy dependence of energy scale power law function - tune that with lower energy data

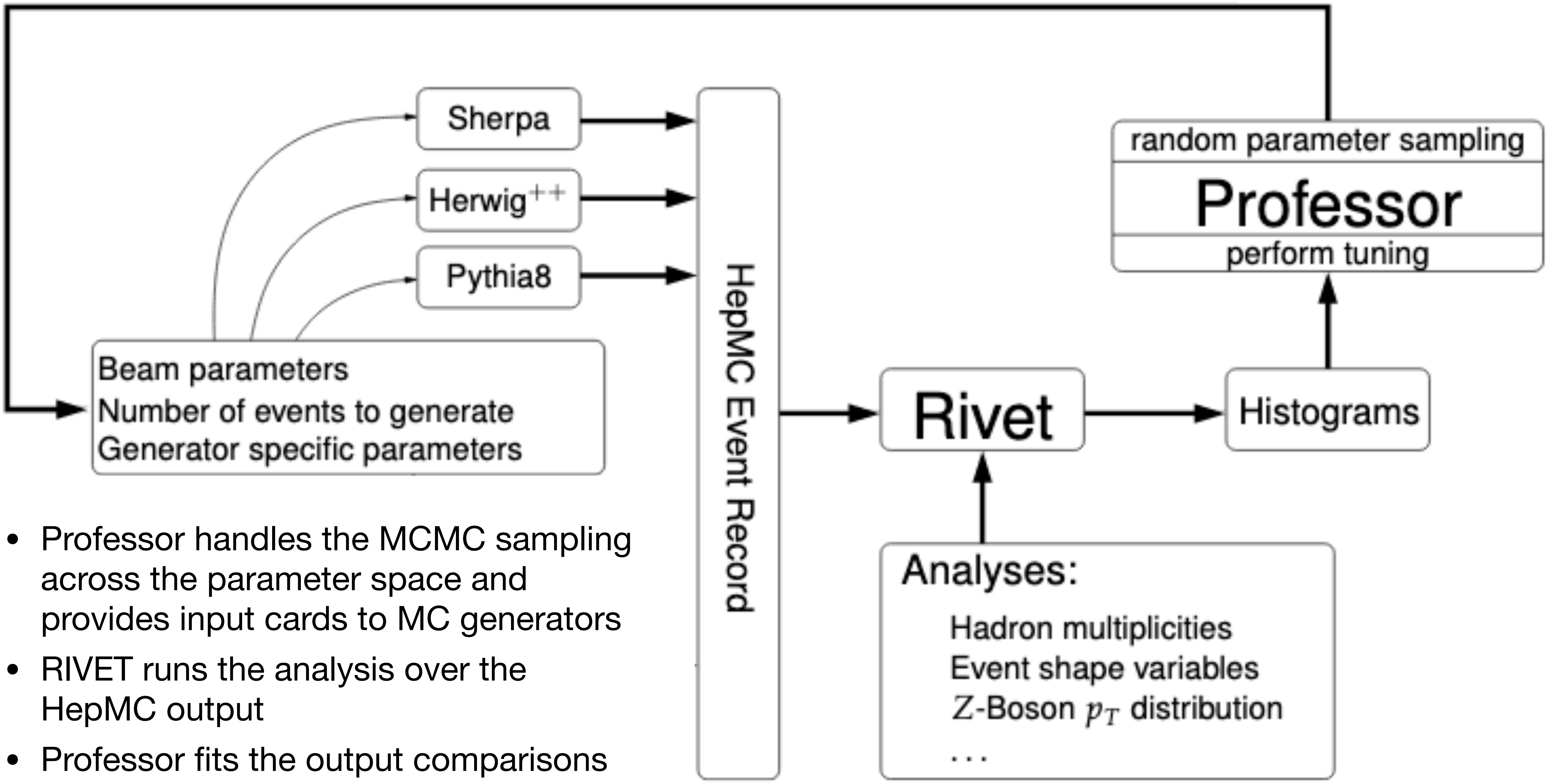
Introduction / Motivation

Tuning Procedure

Detroit Tune

Outlook

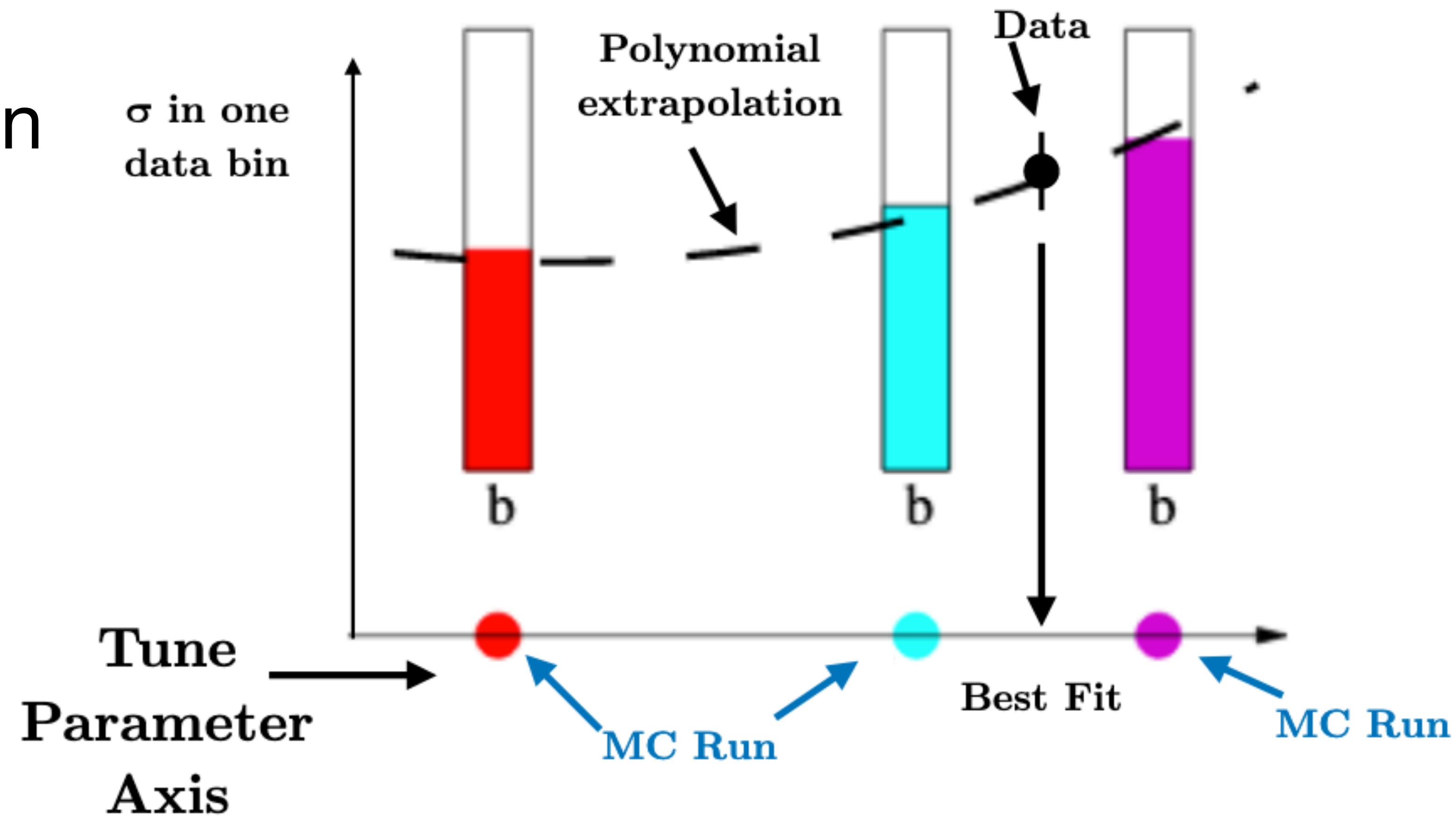
Parametrization-based tuning methodology professor.hepforge.org



- Professor handles the MCMC sampling across the parameter space and provides input cards to MC generators
- RIVET runs the analysis over the HepMC output
- Professor fits the output comparisons of data to MC

Buckley, 0809.4638

- Polynomial parameterization of generated Monte Carlo (MC) data w/ parameter variation
- χ^2 min. w.r.t. data

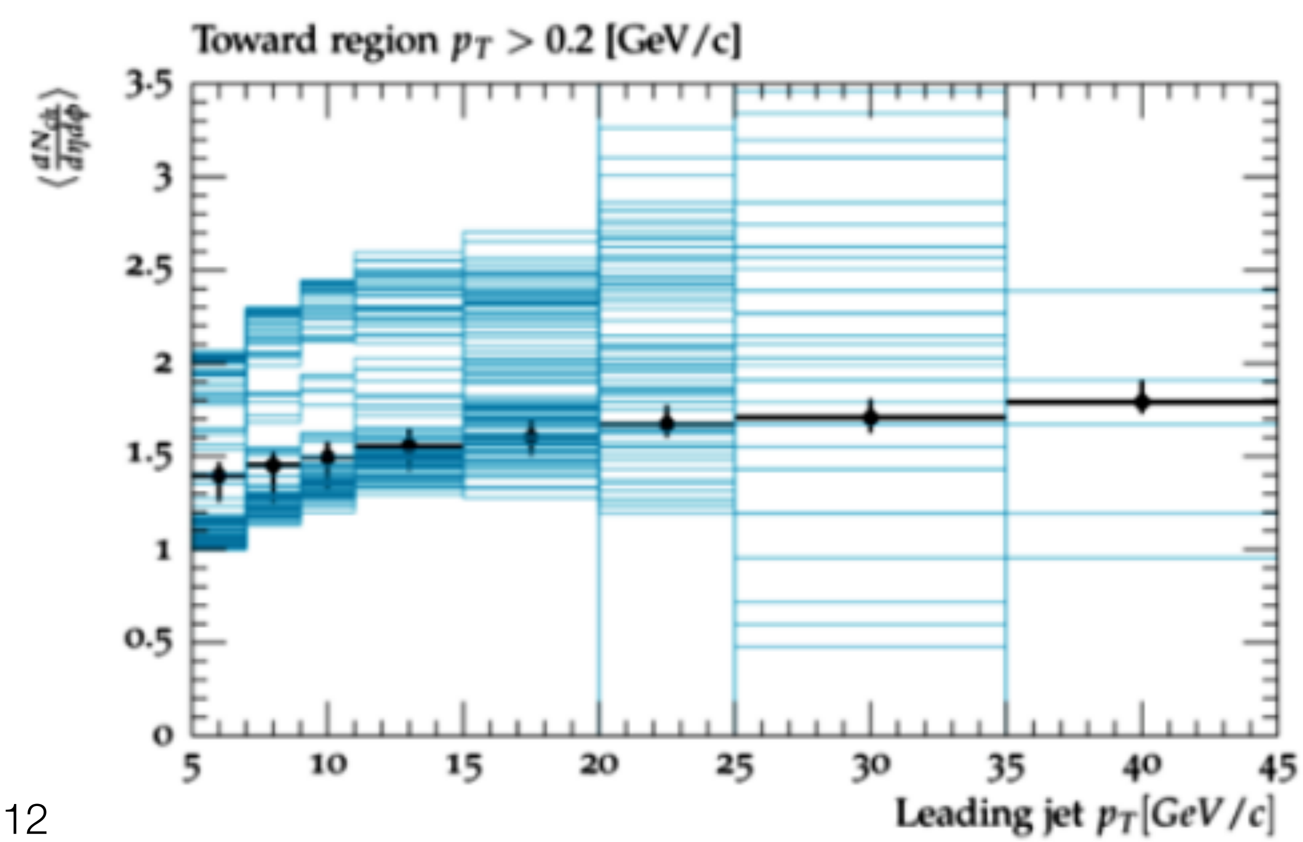


p ≡ **tunable parameter space**

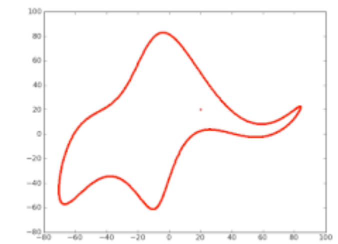
$$\mathrm{MC}_b(\boldsymbol{p}) \approx f^{(b)}(\boldsymbol{p}) = \alpha_0^{(b)} + \sum_i \beta_i^{(b)} p'_i + \sum_{i \leq j} \gamma_{ij}^{(b)} p'_i p'_j$$

MC response in one data bin *b*

$$\chi^2(\boldsymbol{p}) = \sum_{\mathcal{O}} w_{\mathcal{O}} \sum_{b \in \mathcal{O}} \frac{(f^{(b)}(\boldsymbol{p}) - \mathcal{R}_b)^2}{\Delta_b^2}$$



Parameters within PYTHIA that we tune



Setting	Default	New
PDF:pSet	13	17
MultipartonInteractions:ecmRef	7 TeV	200 GeV
MultipartonInteractions:bprofile	3	2
Tuning Parameter	**Default**	**Range**
MultipartonInteractions:pT0Ref	2.28 GeV	0.5-2.5 GeV
MultipartonInteractions:ecmPow	0.215	0.0-0.25
MultipartonInteractions:coreRadius	0.4	0.1-1.0
MultipartonInteractions:coreFraction	0.5	0.0-1.0
ColourReconnection:range	1.8	1.0-9.0

`mode` **MultipartonInteractions:bProfile** (`default = `**`3`**`; minimum = 0; maximum = 4`)
Choice of impact parameter profile for the incoming hadron beams.
`option` **0** : no impact parameter dependence at all.
`option` **1** : a simple Gaussian matter distribution; no free parameters.
`option` **2** : a double Gaussian matter distribution, with the two free parameters *coreRadius* and *coreFraction*.
`option` **3** : an overlap function (i.e. the convolution of the matter distributions of the two incoming hadrons) of the form *exp(- b^expPow)*, where *expPow* is a free parameter.

`parm` **ColourReconnection:range** (`default = `**`1.8`**`; minimum = 0.; maximum = 10.`)
The `range` parameter defined above. The higher this number is the more reconnections can occur. For values above unity the reconnection rate tends to saturate, since then most systems are already connected with each other. This is why 10 is set as an effective upper limit, beyond which it is not meaningful to let the parameter go.

- Start from PYTHIA 8.303 w/ Monash tune with updated PDF to NNPDF 3.1 w/ LO $\alpha_s(M_z) = 0.13$
- Reference Energy set to 200 GeV

- String parameters are left to the default Monash values
- Note: whatever Monash can't do w.r.t strangeness or baryons etc… Detroit also cannot do.

Data used for the tuning procedure

Particle spectra	Underlying event	Jet substructure

Data used for the tuning procedure

Particle spectra

Underlying event

Jet substructure

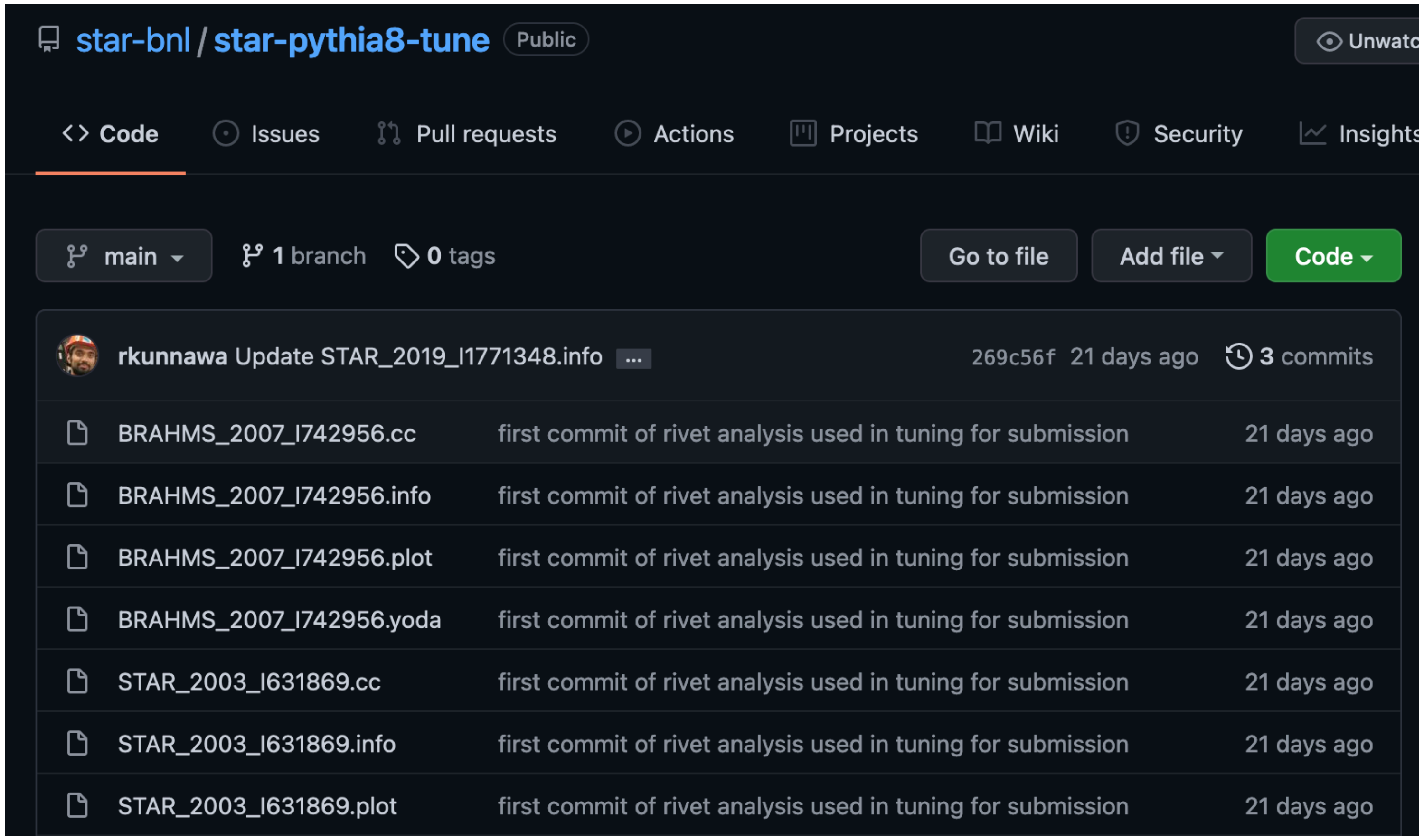
Experiment	$\sqrt{s}$ (GeV)	Observable
STAR	200	$\pi^{\pm}$ cross sections vs. p_T
PHENIX	200	Di-muon pairs from Drell-Yan vs. di-muon p_T
STAR	200	Average charged particle multiplicities and p_T vs. leading jet p_T in the forward, transverse, and away regions
CDF	300, 900, 1960	Charge particle density and $\sum p_T$ vs. leading hadron p_T in transverse region
STAR	200	SoftDrop groomed jet sub-structure (z_g and R_g)
STAR	200	Inclusive and groomed jet mass

RIVET Analysis

- https://rivet.hepforge.org/

- The Rivet toolkit (Robust Independent Validation of Experiment and Theory) is a system for validation of Monte Carlo event generators. It provides a large (and ever growing) set of experimental analyses useful for MC generator development, validation, and tuning, as well as a convenient infrastructure for adding your own analyses.

- https://github.com/star-bnl/star-pythia8-tune



Introduction / Motivation

Tuning Procedure

Detroit Tune

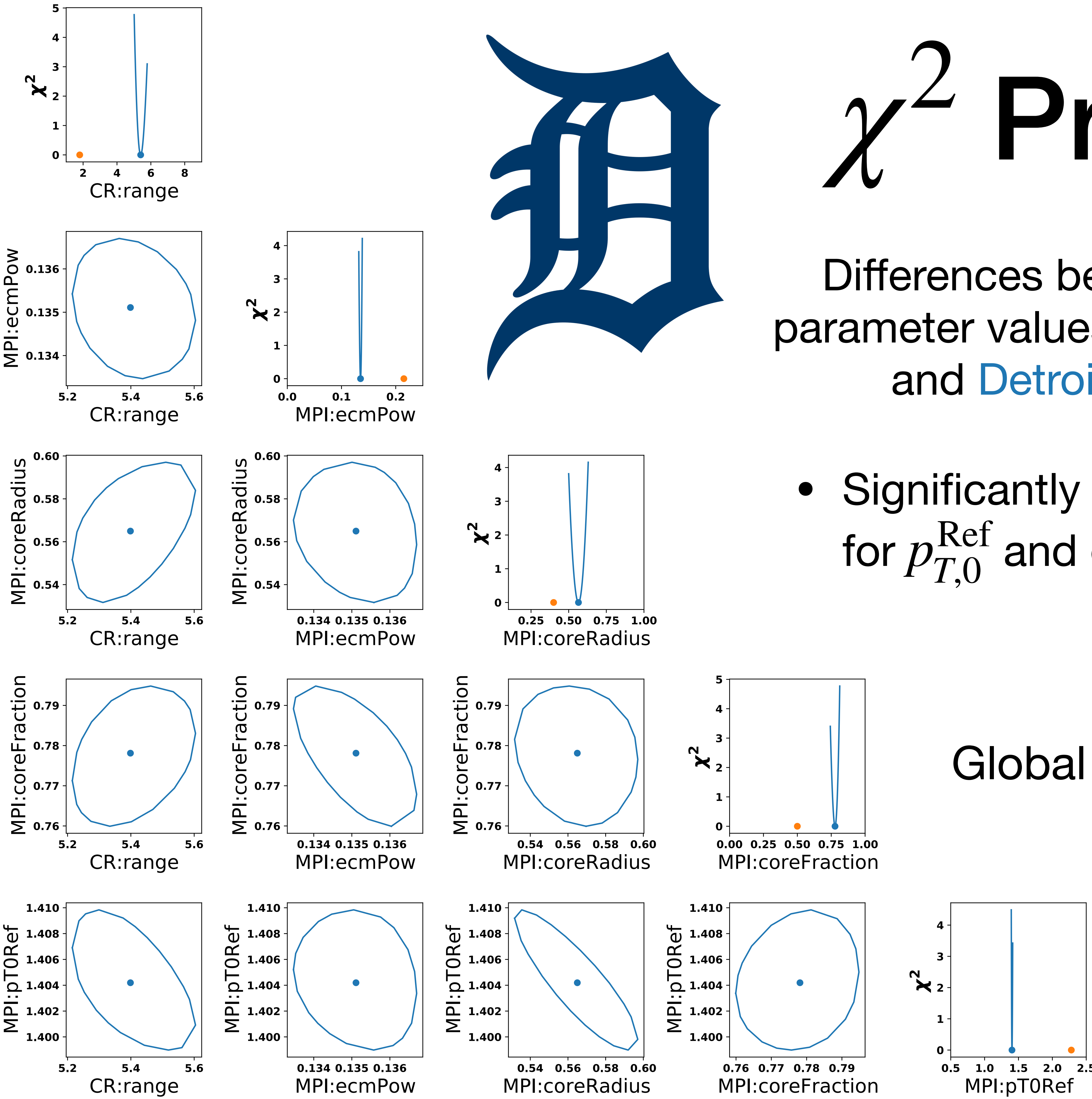
Outlook

χ^2 Profiles

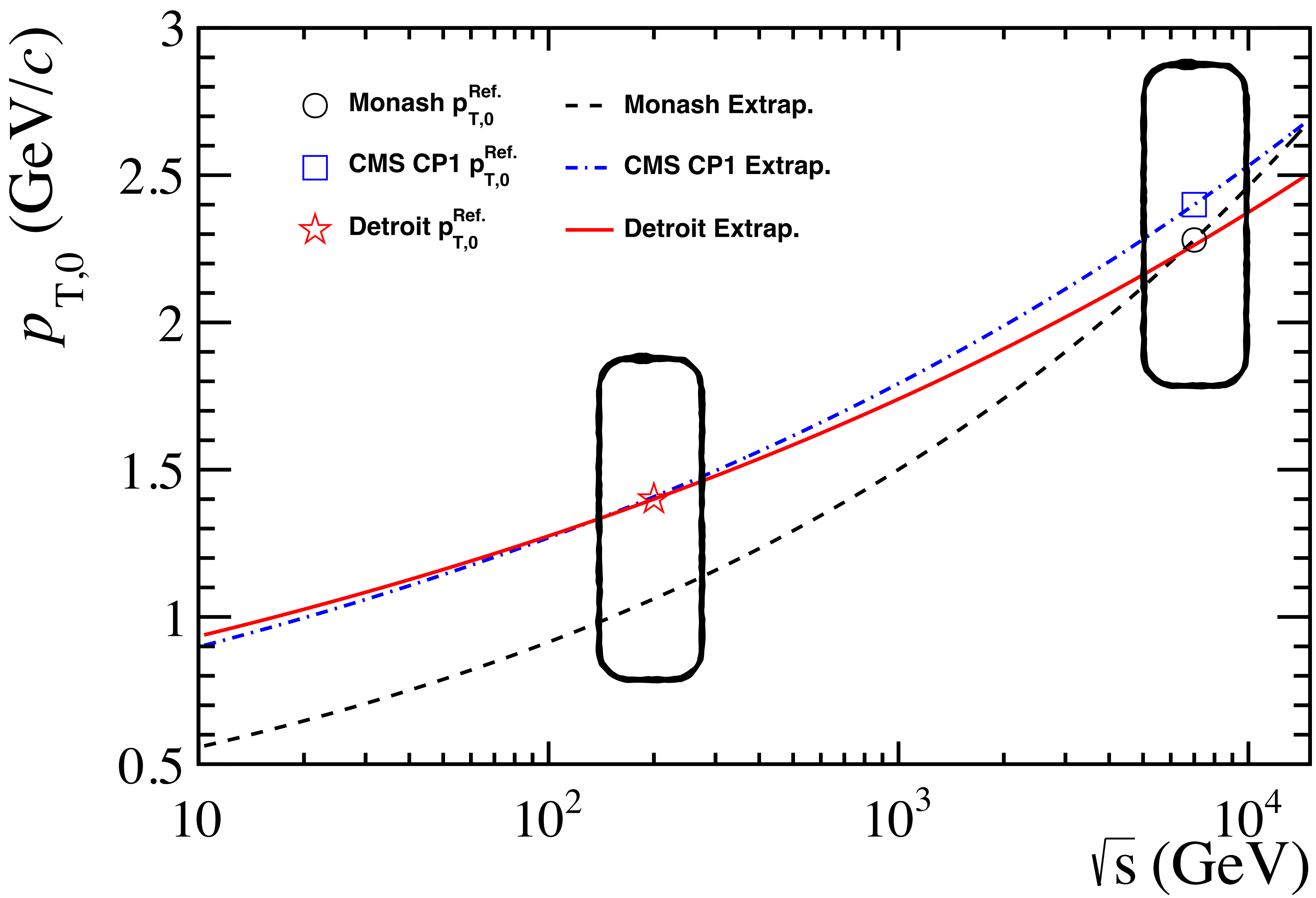
Differences between the parameter values for Monash and Detroit tunes

- Significantly smaller values for $p_{T,0}^{\mathrm{Ref}}$ and ecm exponent

Global χ^2/n.d.f. = 1.2

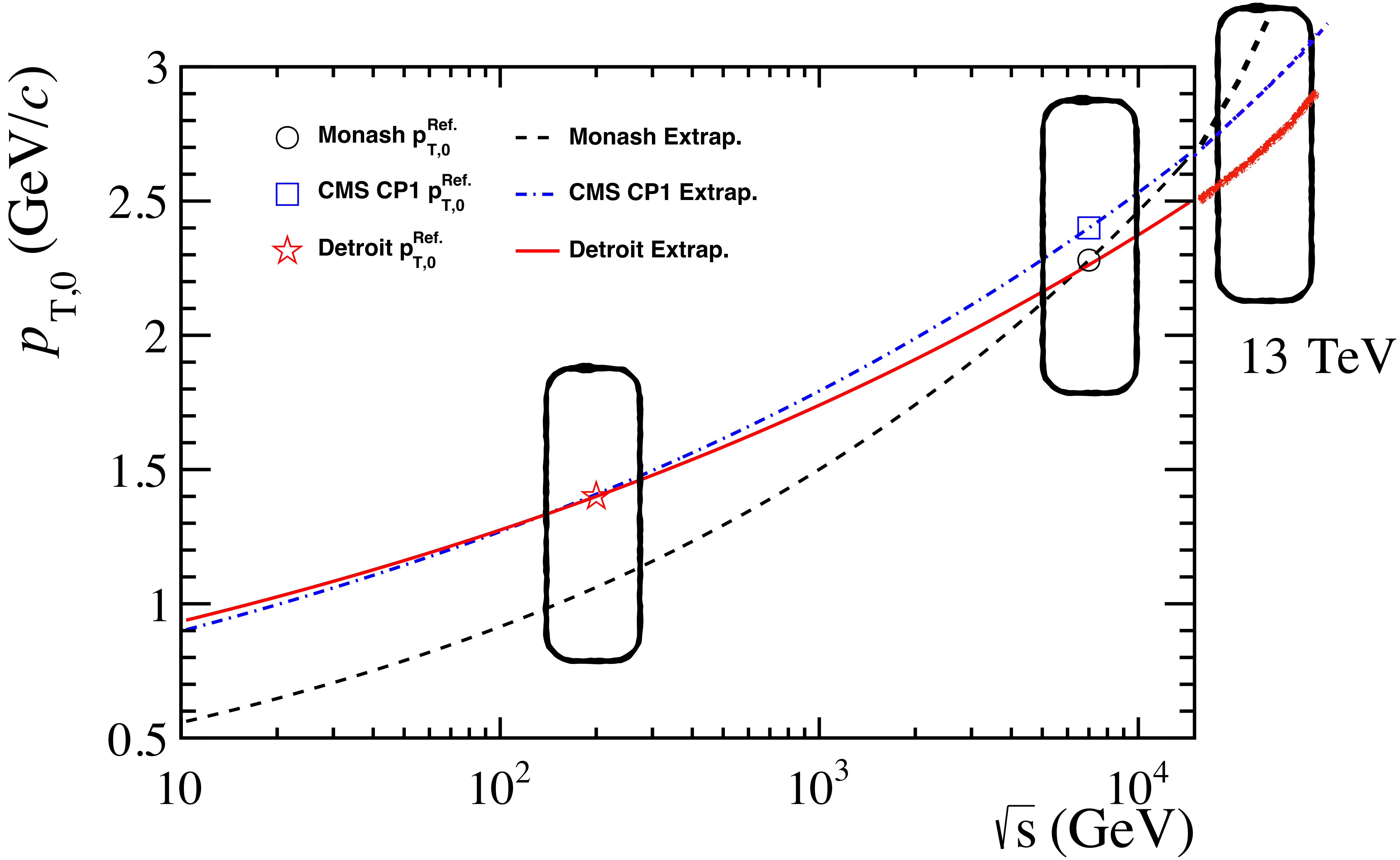


Energy extrapolation



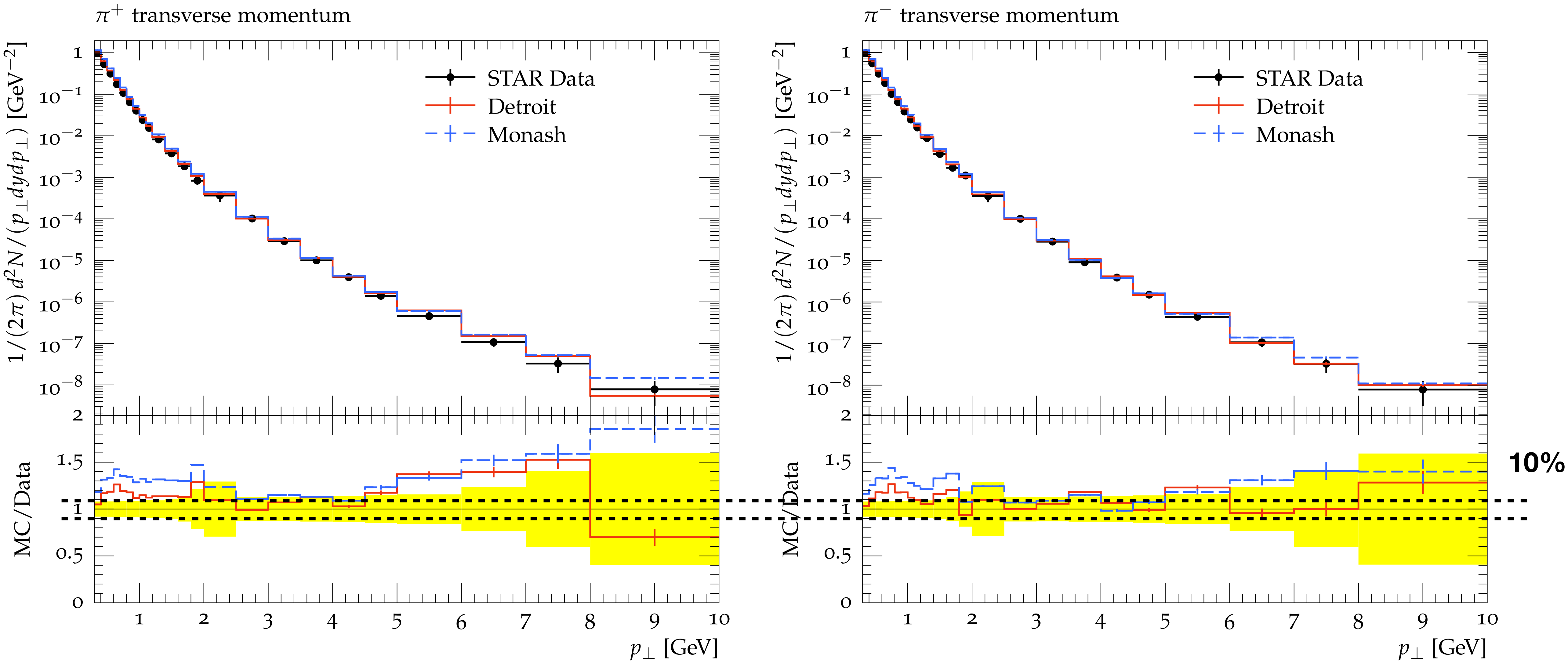
- Monash severely undershoots the $p_{T,0}$ scale at RHIC energies
- CMS CP1 and Detroit have similar values at $\sqrt{s}$ = 200 GeV but significantly differ in the extrapolation at LHC energies and beyond

Energy extrapolation



- Monash severely undershoots the $p_{T,0}$ scale at RHIC energies
- CMS CP1 and Detroit have similar values at $\sqrt{s} = 200$ GeV but significantly differ in the extrapolation at LHC energies and beyond

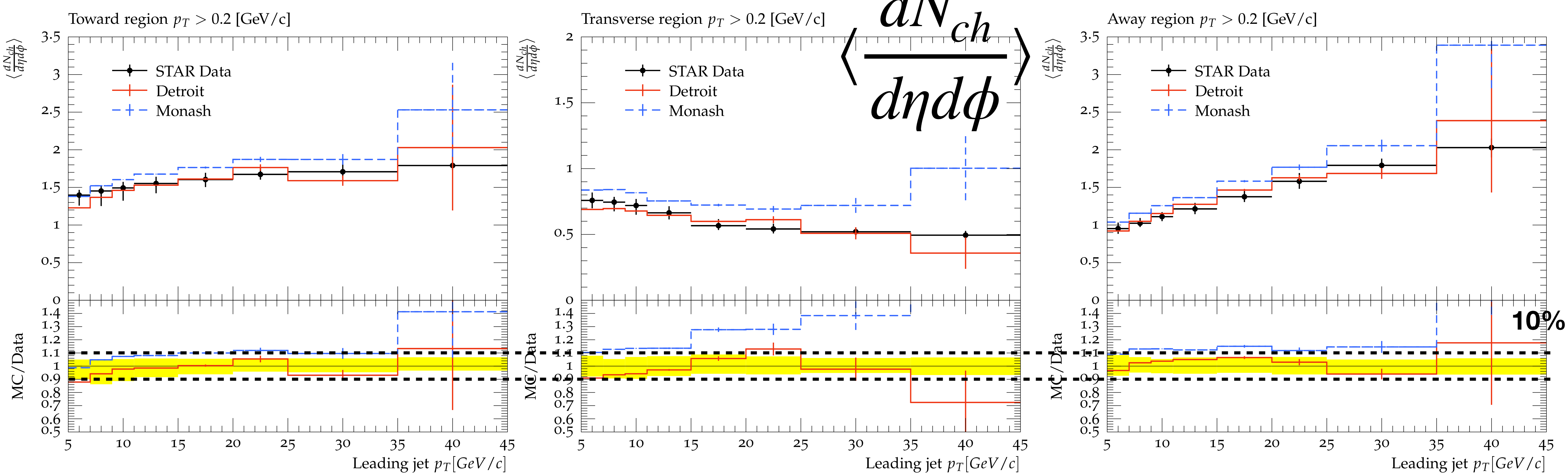
RHIC Comparisons
$\pi^{\pm}$ production @ STAR

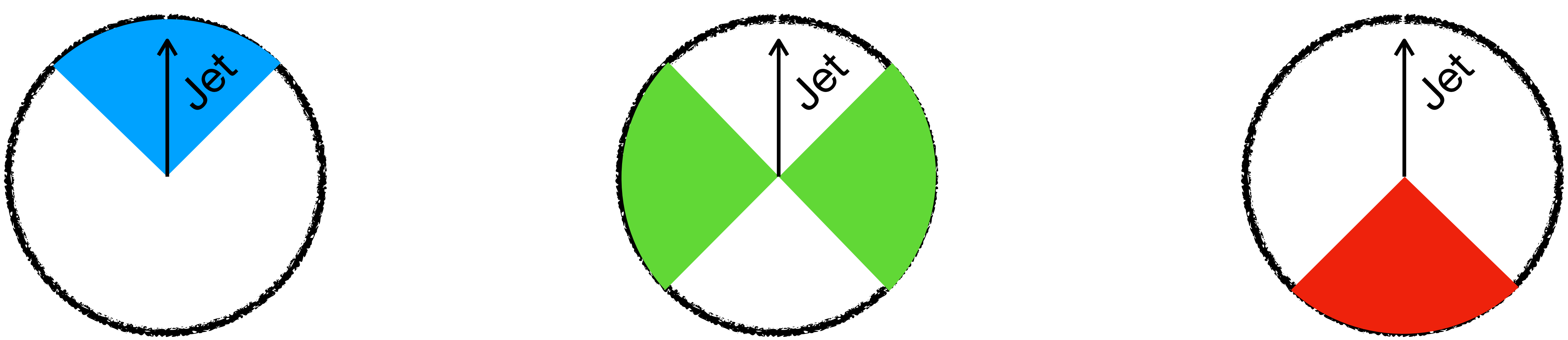


- Inclusive charged π production at low p_T shows significant reduction in the yield with the Detroit tune compared to Monash - similar to the observation in the PYTHIA 6 tuning exercise - first checkpoint!

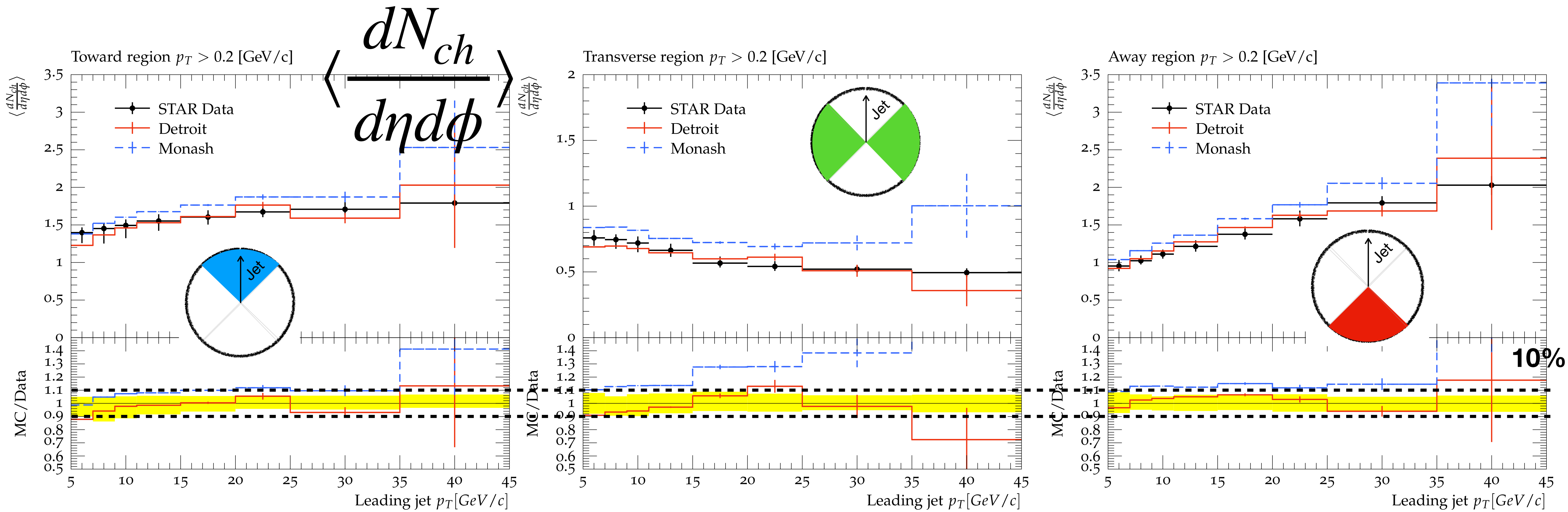
RHIC Comparisons
Underlying event @ STAR - I

$$\left\langle \frac{dN_{ch}}{d\eta d\phi} \right\rangle$$





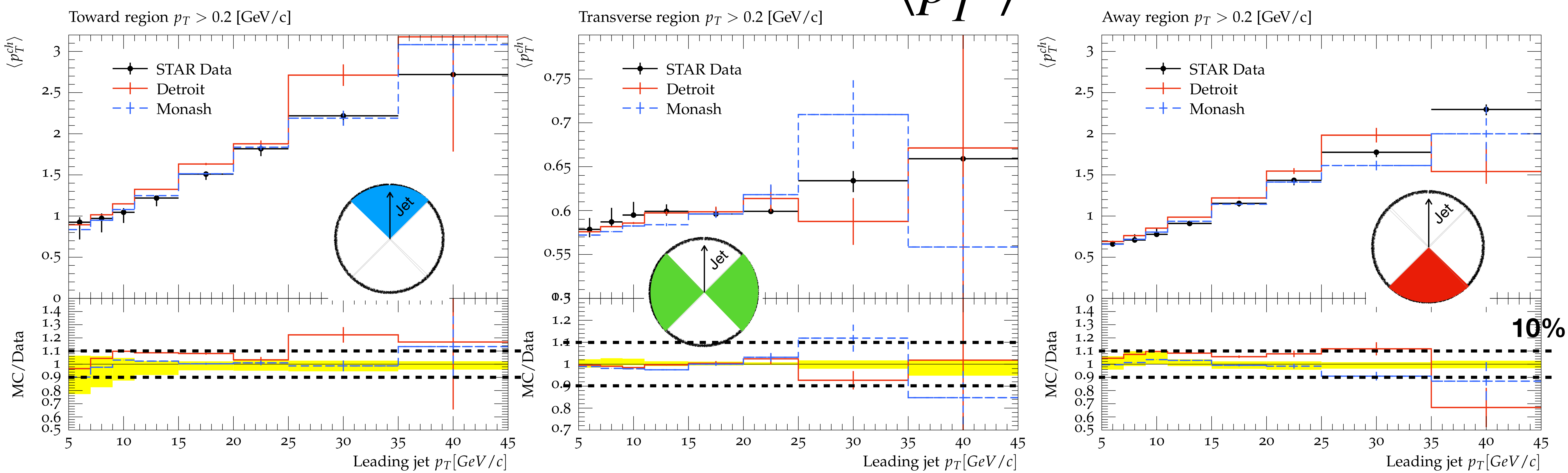
RHIC Comparisons
Underlying event @ STAR - I



- Consistent reduction for the mean multiplicity in events tagged by the leading jet momenta for all regions w.r.t jet
- Disagreements with Monash grows stronger in the transverse region

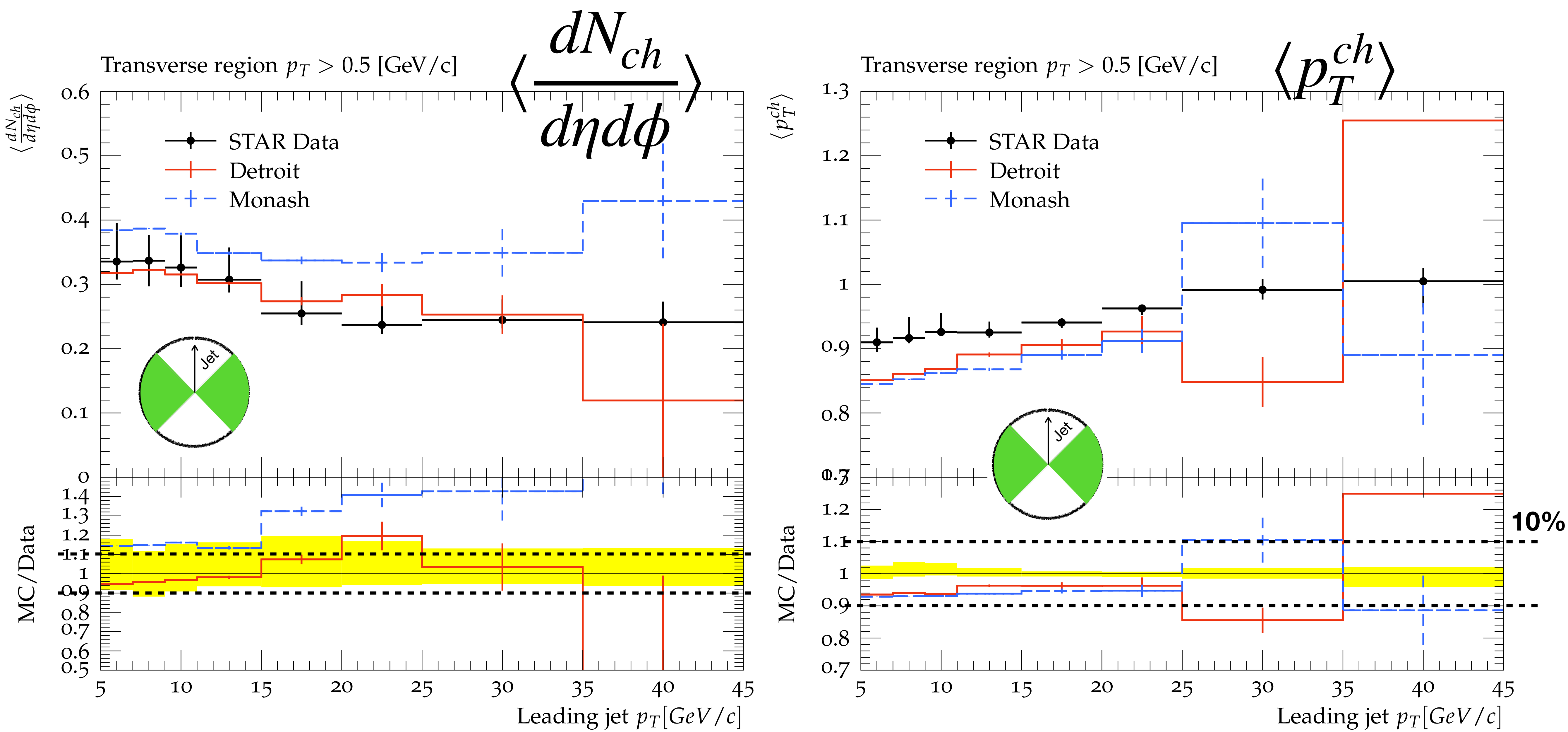
RHIC Comparisons Underlying event @ STAR - II

$\langle p_T^{ch} \rangle$



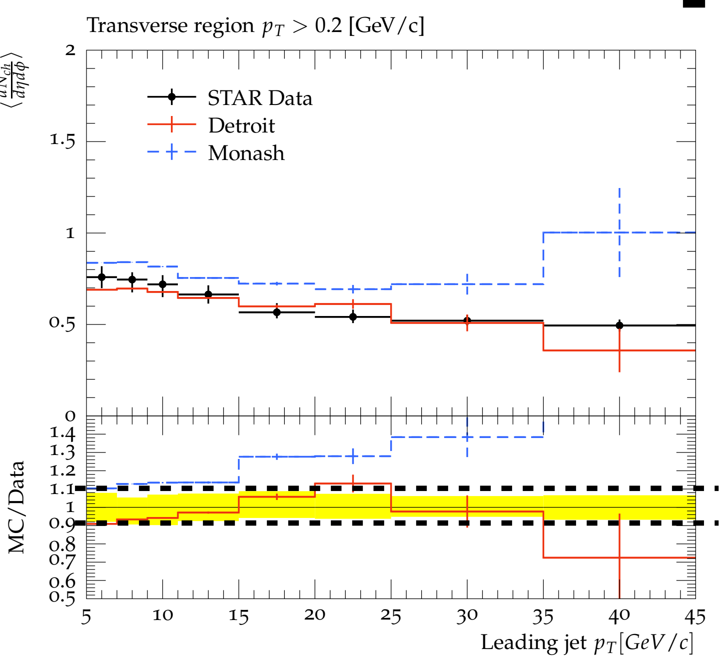
- Intriguing effect seen in the average p_T of the particles in that Monash and Detroit are similar (w/ Monash even slightly better but differences are on the order of ~ 5-10%)

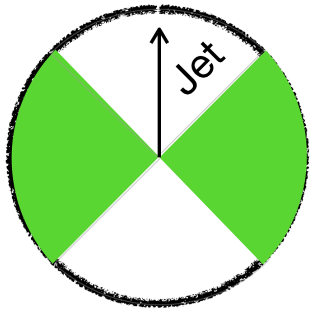
RHIC Comparisons
Underlying event @ STAR - III



- Extending to higher track momentum - similar take away for the mean multiplicity but both tunes unable to describe mean momenta for events with lower jet momenta

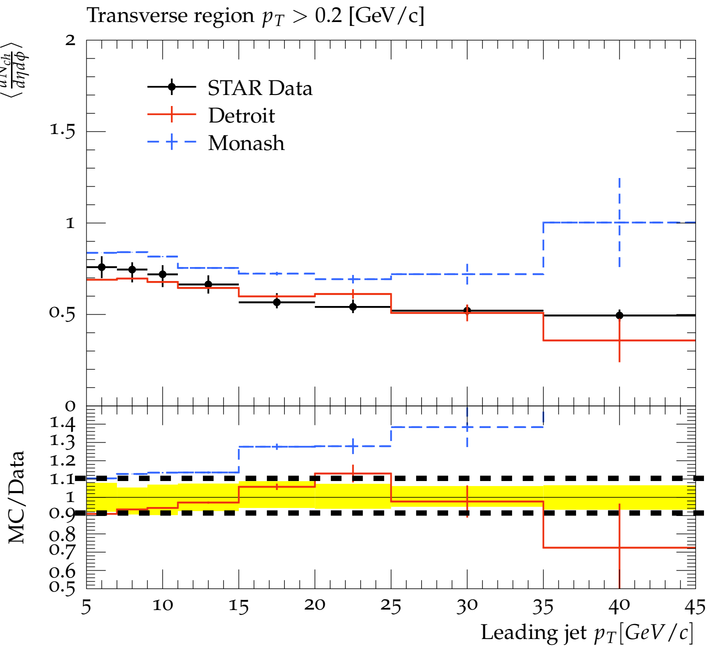
Underlying events from RHIC to Tevatron

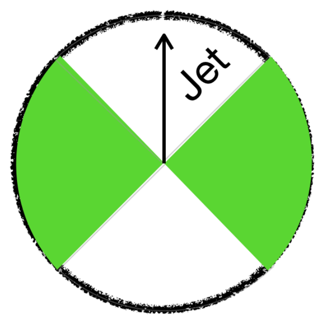




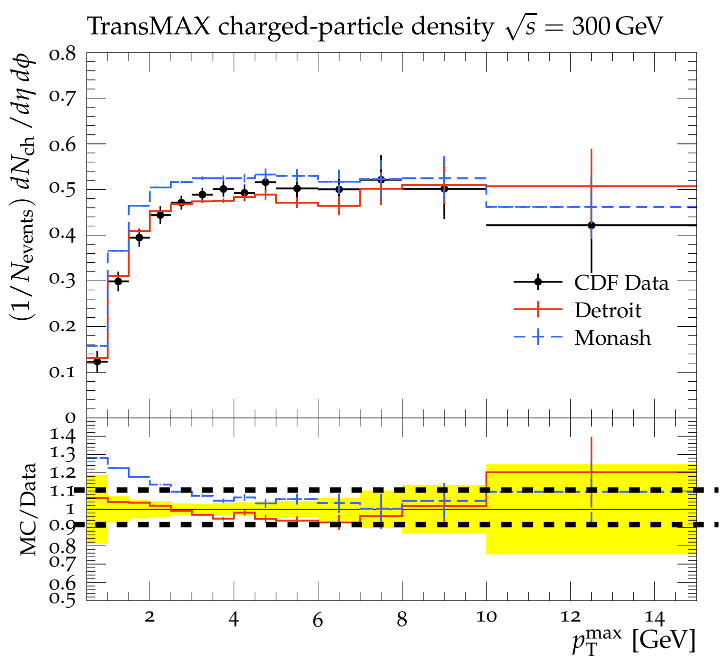
$$\langle \frac{dN_{ch}}{d\eta d\phi} \rangle$$

Underlying events from RHIC to Tevatron



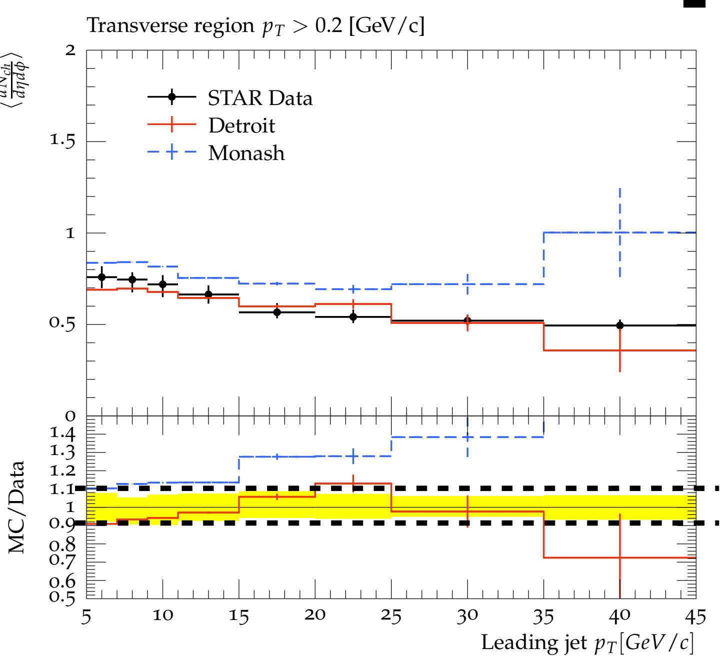


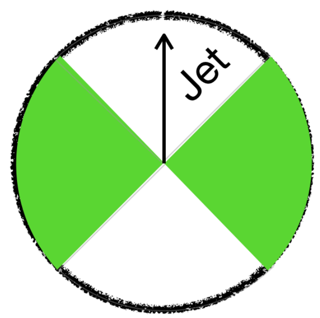
$$\langle \frac{dN_{ch}}{d\eta d\phi} \rangle$$



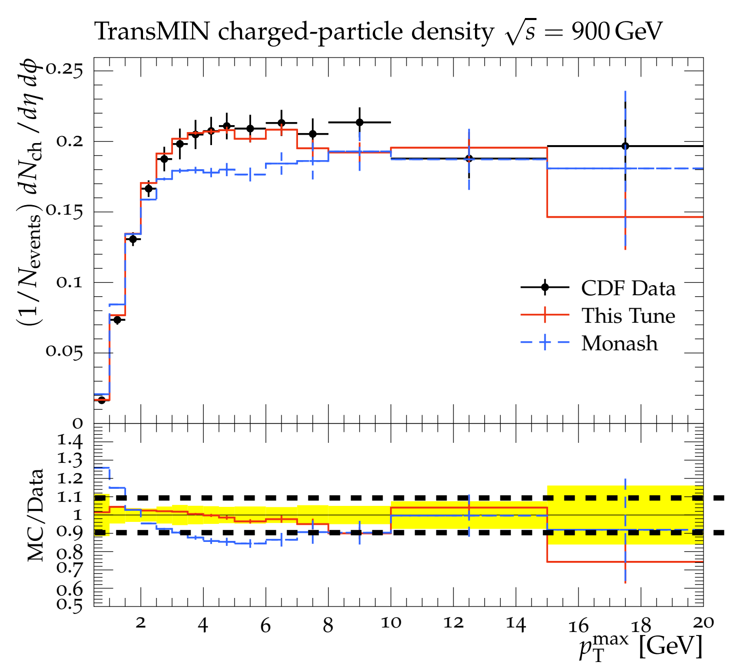
$\sqrt{s} =$ 200 300 900 1960 [GeV]

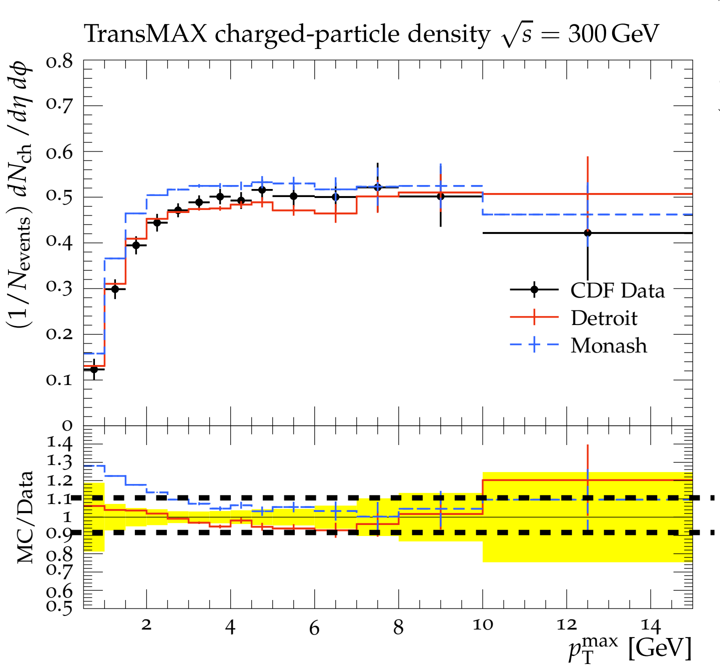
Underlying events from RHIC to Tevatron





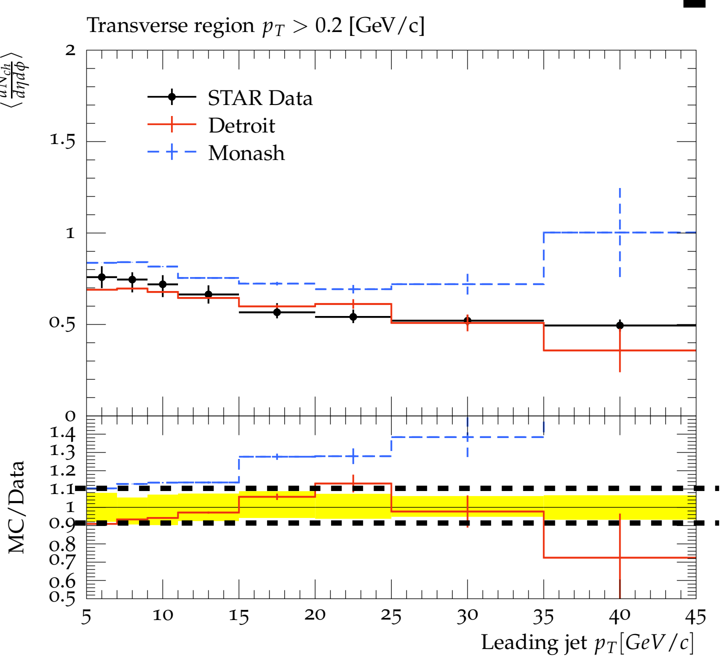
$$\langle \frac{dN_{ch}}{d\eta d\phi} \rangle$$

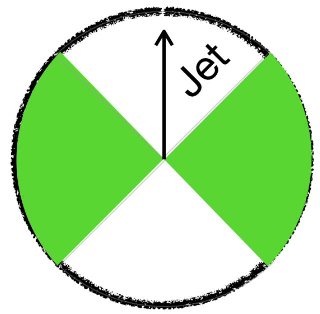




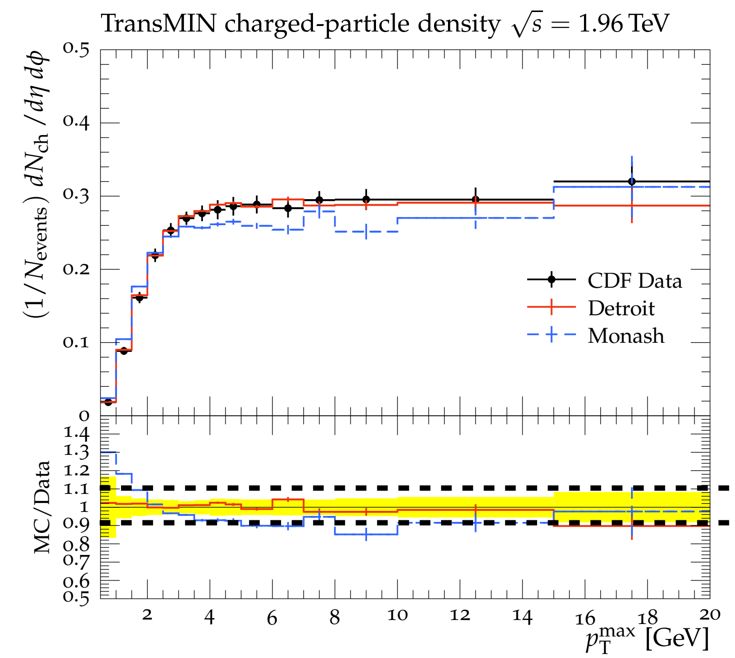
$\sqrt{s} =$ 200 300 900 1960 [GeV]

Underlying events from RHIC to Tevatron

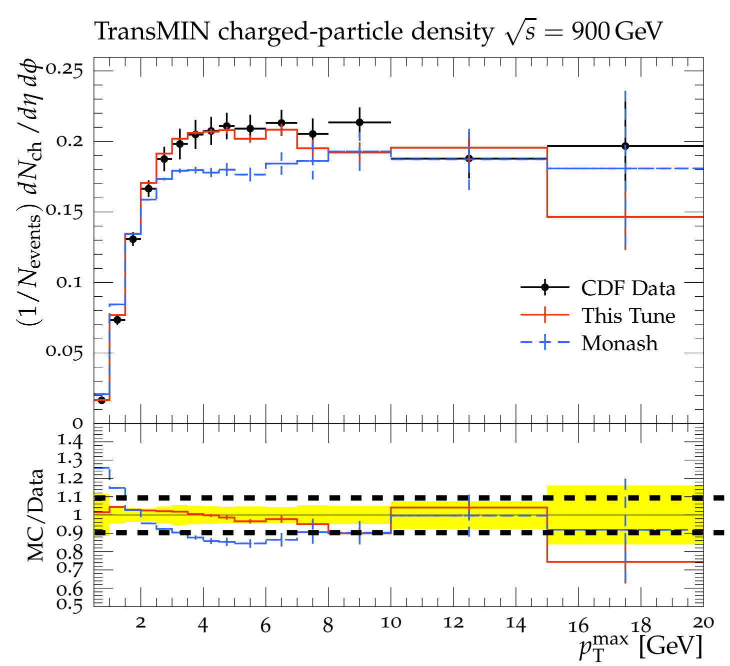


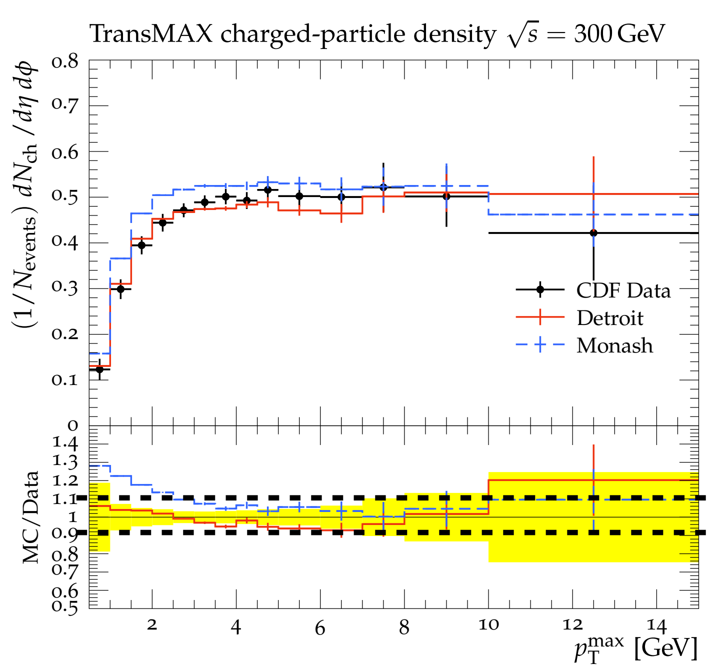


$$\left\langle \frac{dN_{ch}}{d\eta d\phi} \right\rangle$$



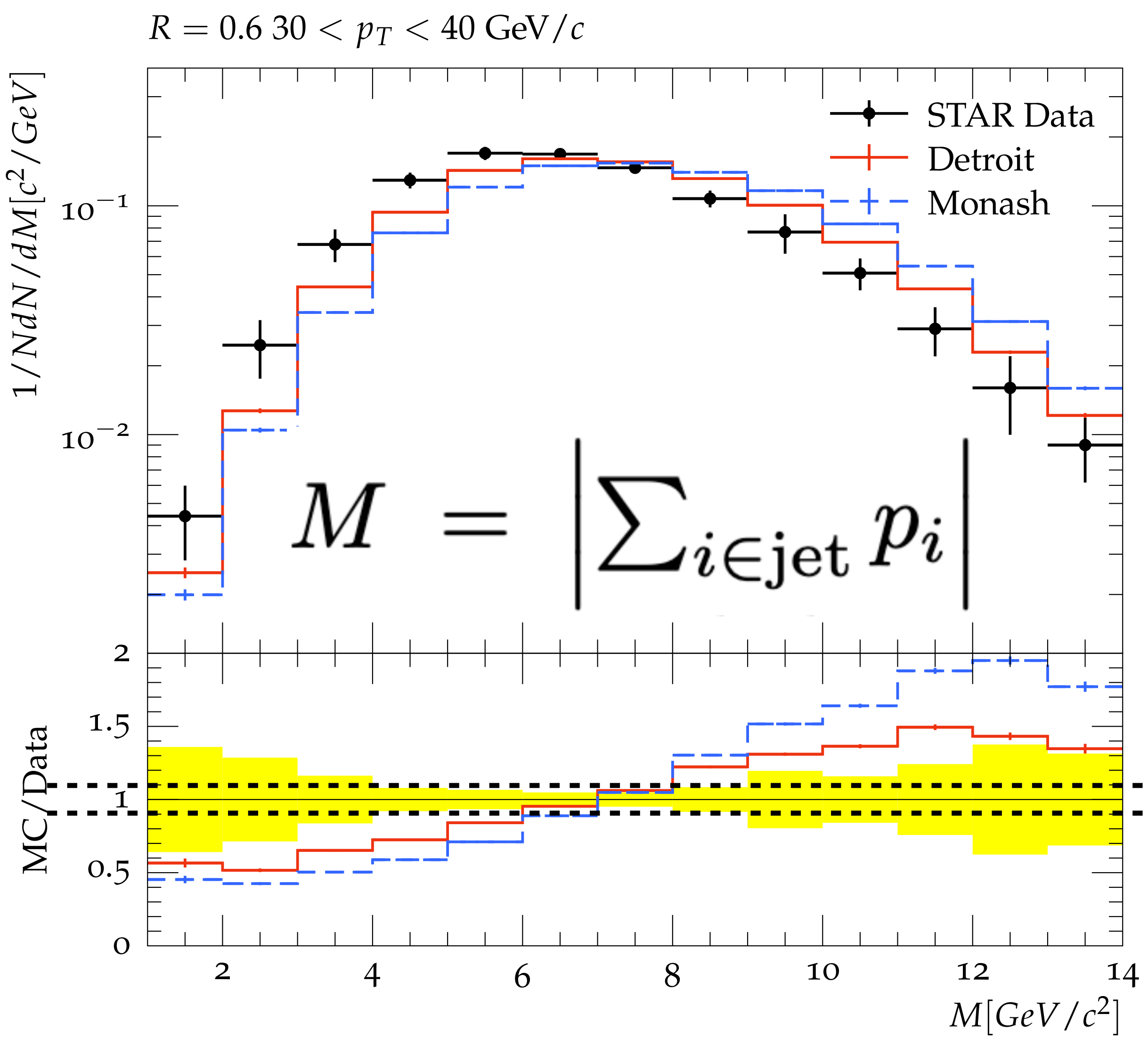
10%

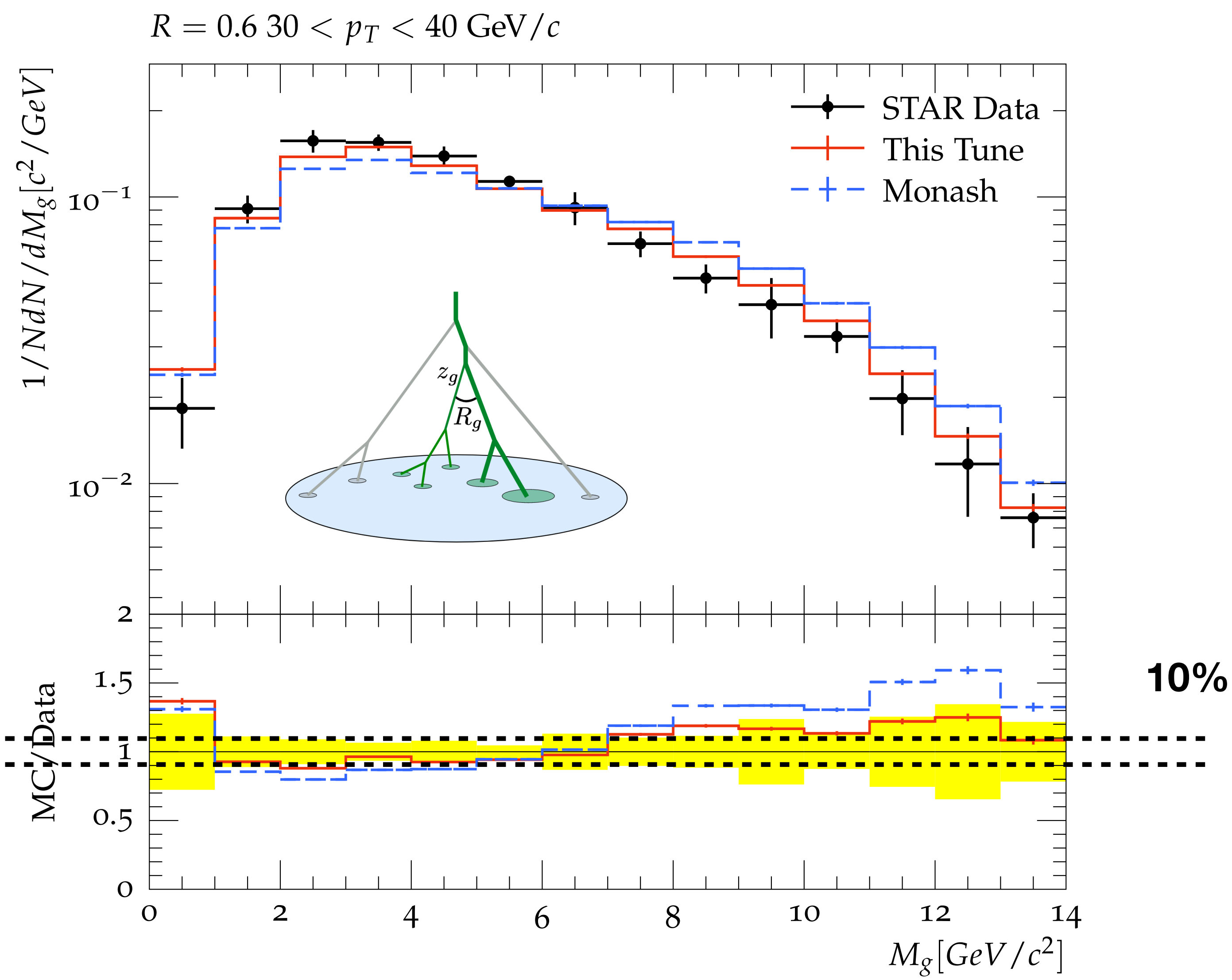




$\sqrt{s} =$ 200 300 900 1960 [GeV]

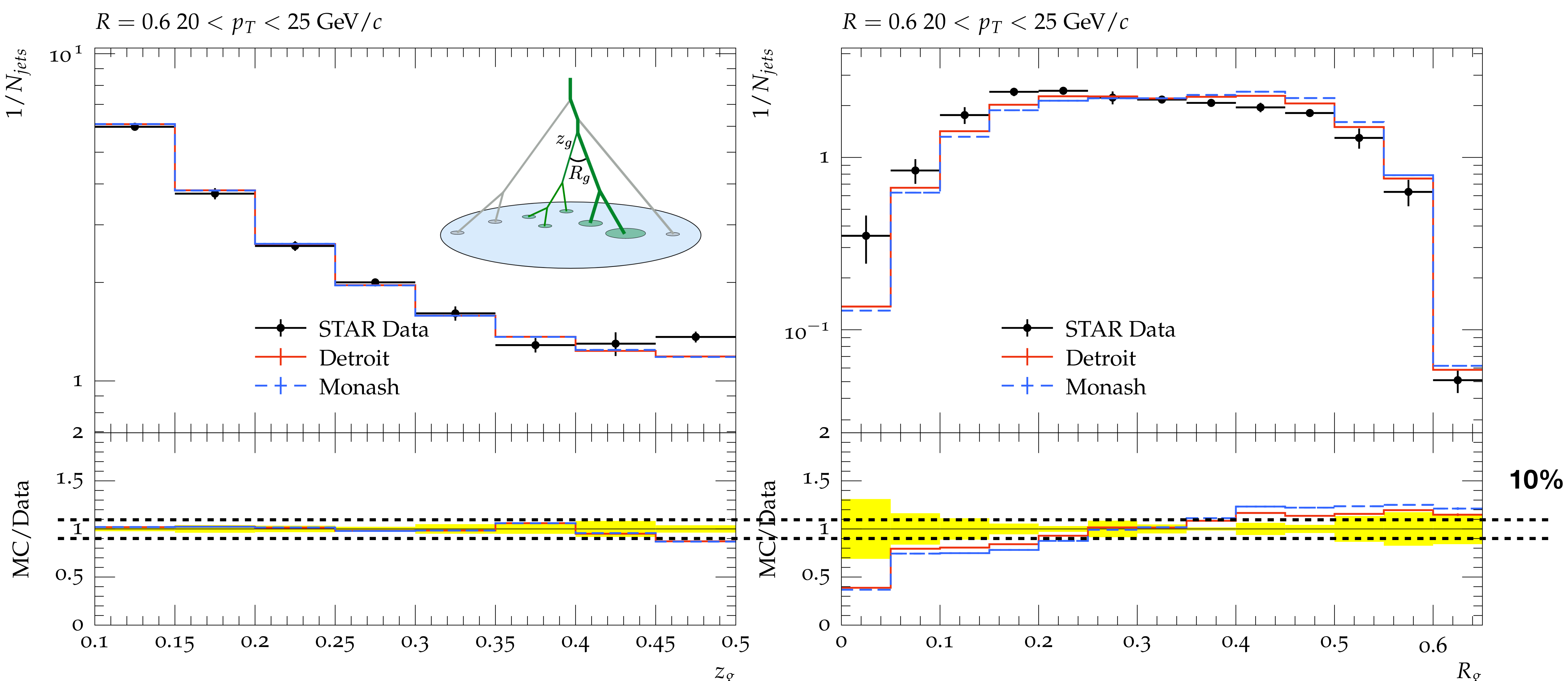
RHIC Comparisons
Inclusive and groomed jet mass @ STAR





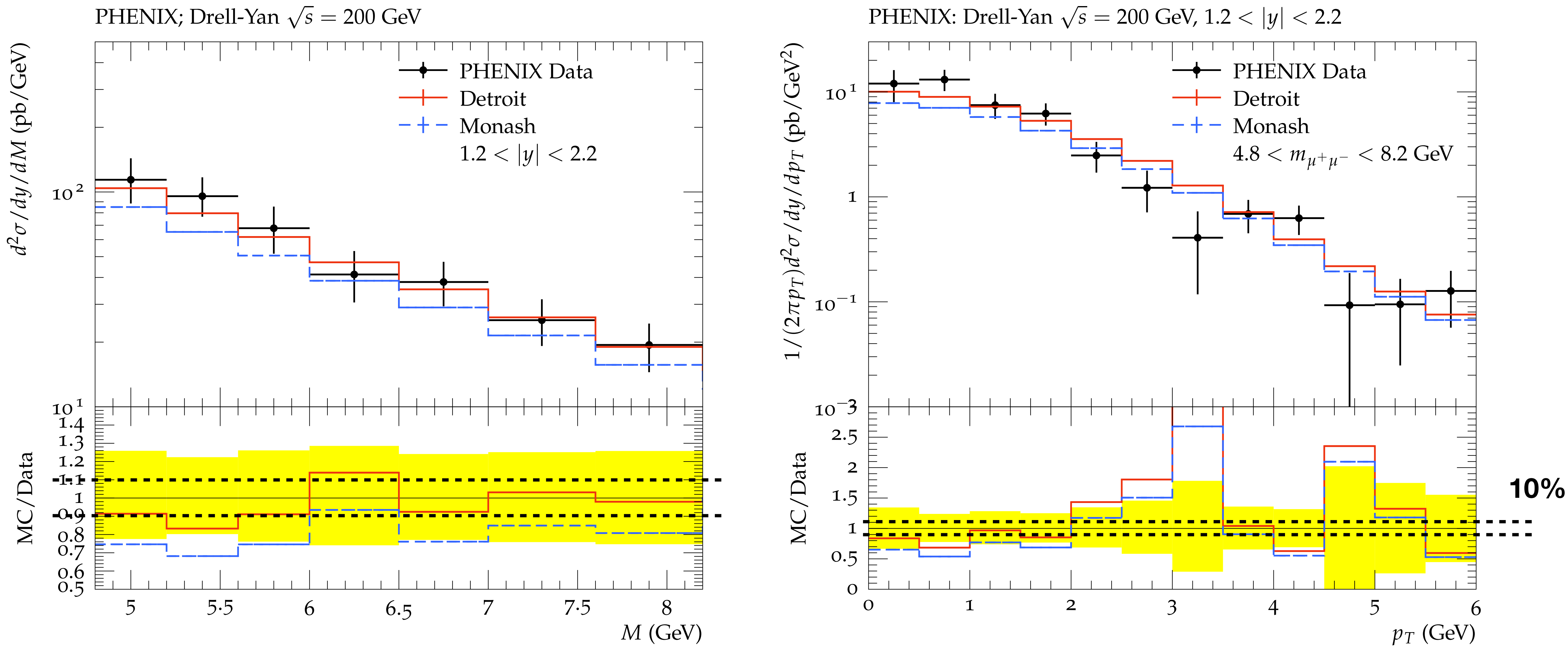
- Slight improvements in the description of the jet mass - shifts the distribution to the left (smaller masses)
- Grooming removes softer particles and hence we only see a small change with the Detroit tune for large groomed masses

RHIC Comparisons
Jet substructure @ STAR



- Mass as a jet angularity $\approx z_g\ R_g^2$ and so we see almost negligible improvement with the opening angle whereas the momentum fraction is effectively insensitive to the underlying event contribution after grooming

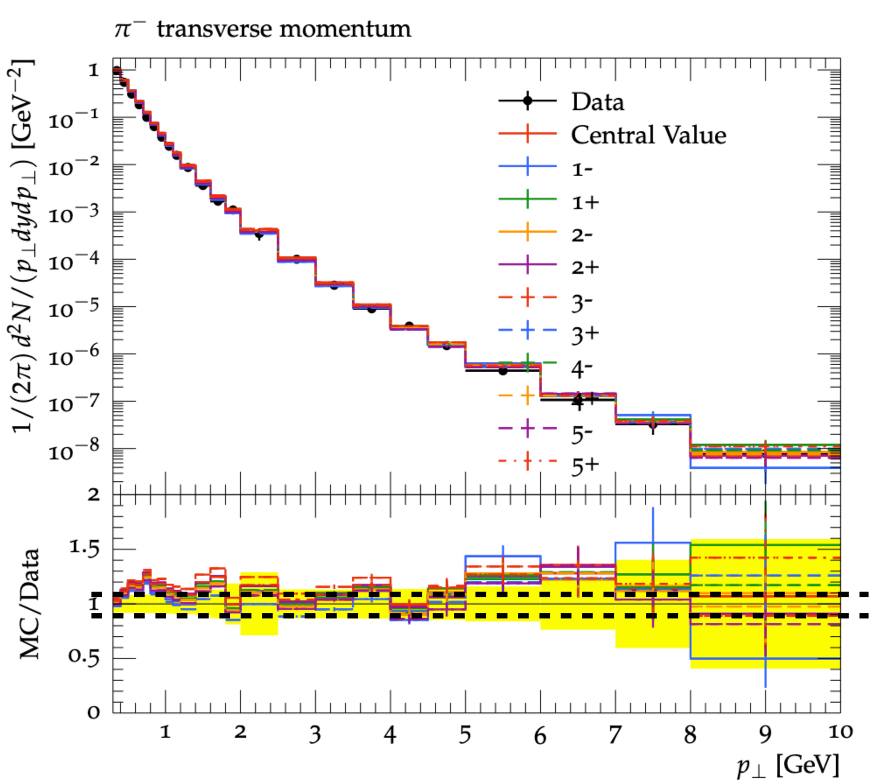
RHIC Comparisons
Drell-Yan production @ PHENIX

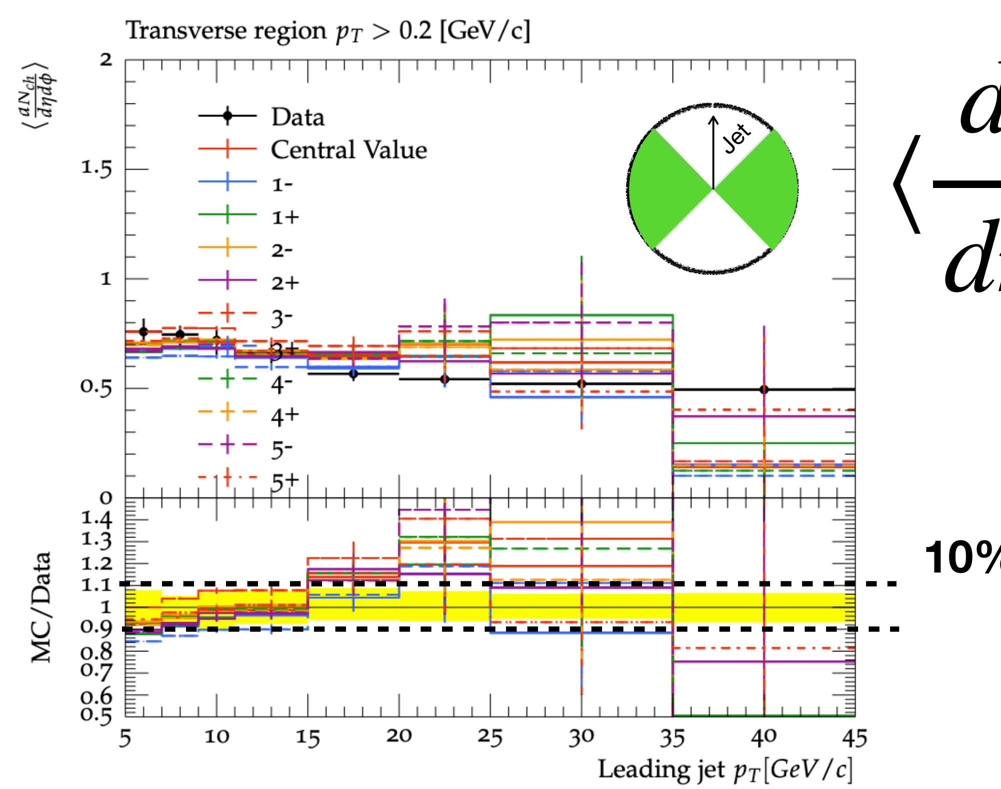


- Improvement in the di-μ pair invariant mass M and their p_T spectra
- Possible extensions to specifically tuning heavy flavor production within the same framework

Eigentunes

Tuning Parameter	1+	1-	2+	2-	3+	3-	4+	4-	5+	5-
MultipartonInteractions:pT0Ref (GeV)	1.37	1.43	1.38	1.42	1.44	1.37	1.41	1.40	1.40	1.41
MultipartonInteractions:ecmPow	0.132	0.138	0.135	0.135	0.119	0.150	0.145	0.126	0.148	0.125
MultipartonInteractions:coreRadius	0.74	0.41	0.77	0.41	0.57	0.56	0.57	0.56	0.51	0.60
MultipartonInteractions:coreFraction	0.84	0.72	0.72	0.82	0.78	0.78	0.78	0.78	0.60	0.90
ColourReconnection:range	7.50	3.61	5.38	5.41	5.40	5.40	5.40	5.40	5.41	5.40





$$\langle \frac{dN_{ch}}{d\eta d\phi} \rangle$$

- Variations in the respective 1σ boundaries of the individual parameters and their corresponding impact on the best fit values
- Can be considered as an inherent uncertainty in the way the individual parameters talk to each other in the event generation

Introduction / Motivation

Tuning Procedure

Detroit Tune

Outlook

Recap

- PYTHIA 8 Monash tune systematically mis-represented data at RHIC energies
- MPI related parameters ($p_{T,0}^{\mathrm{Ref}}$) and $\sqrt{s}$ extrapolation are important for universality of particle production
- Detroit tune significantly outperforms Monash at RHIC and Tevatron energies

Outlook

- What about the Detroit tune at top LHC energy?
- Particle production across the whole event - mid and forward rapidities
- Extensions of the tuning procedure

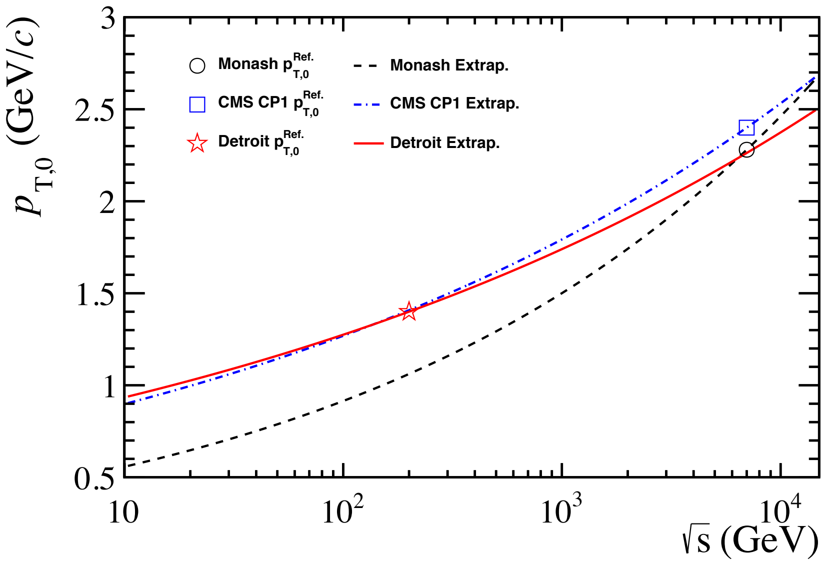
Recap

- PYTHIA 8 Monash tune systematically mis-represented data at RHIC energies
- MPI related parameters ($p_{T,0}^{\mathrm{Ref}}$) and $\sqrt{s}$ extrapolation are important for universality of particle production
- Detroit tune significantly outperforms Monash at RHIC and Tevatron energies

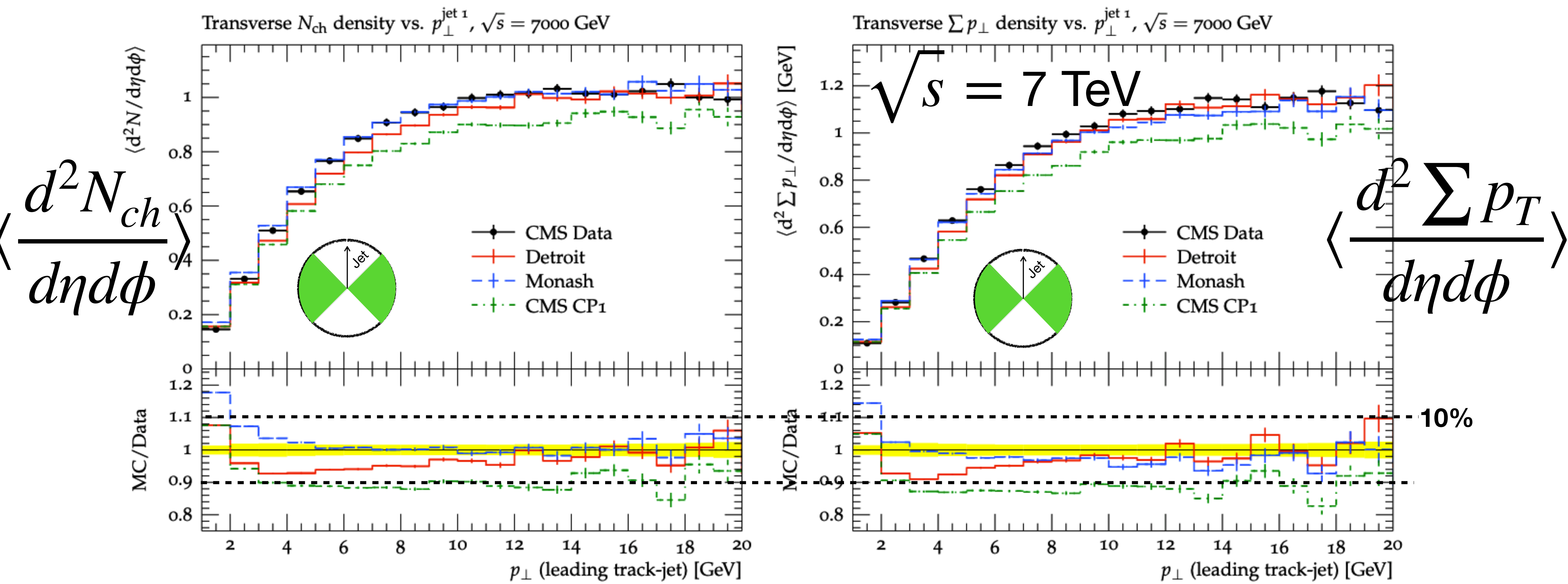
Outlook

- What about the Detroit tune at top LHC energy?
- Particle production across the whole event - mid and forward rapidities
- Extensions of the tuning procedure

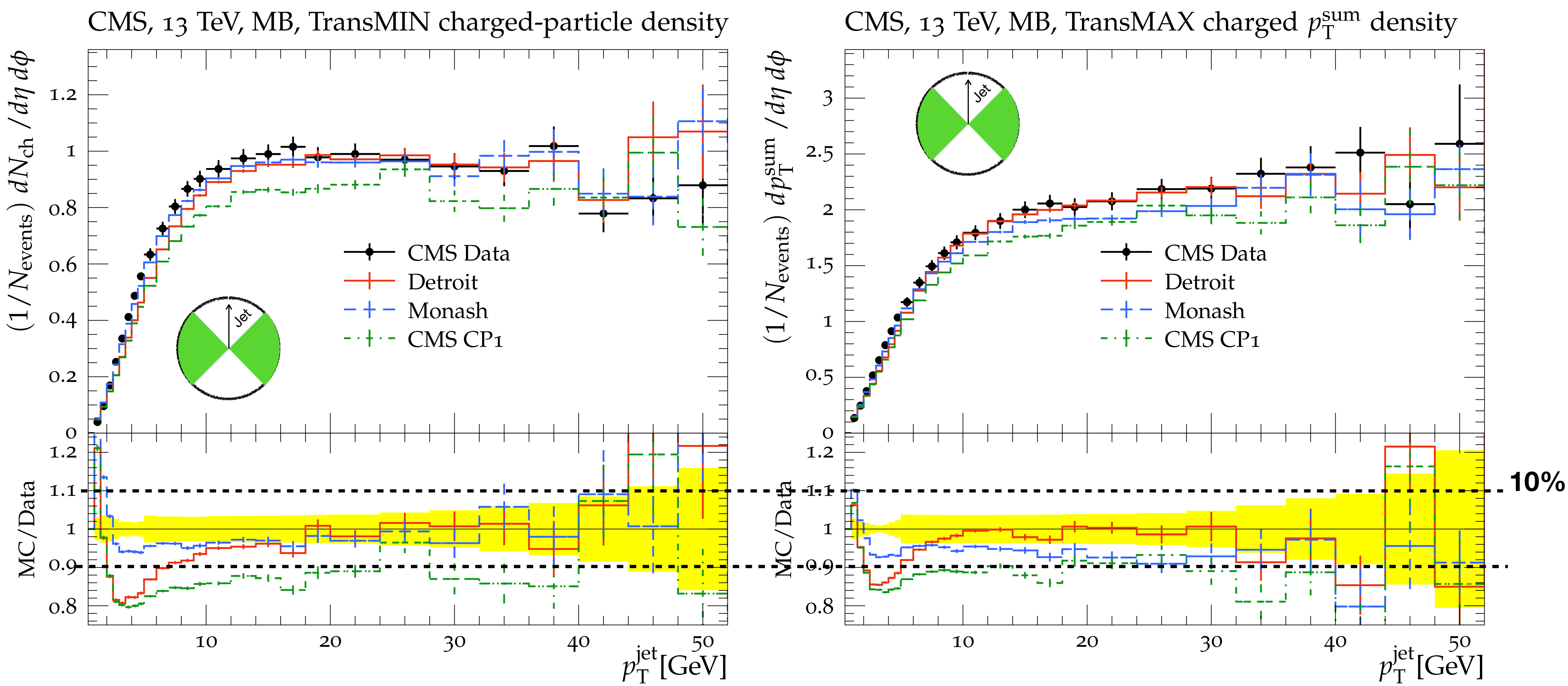
Comparison at the LHC energies



- Under-predicts data for low energies but better comparison (similar to Monash) at high p_T
- Significant differences between comparisons to data including weak decayed particles

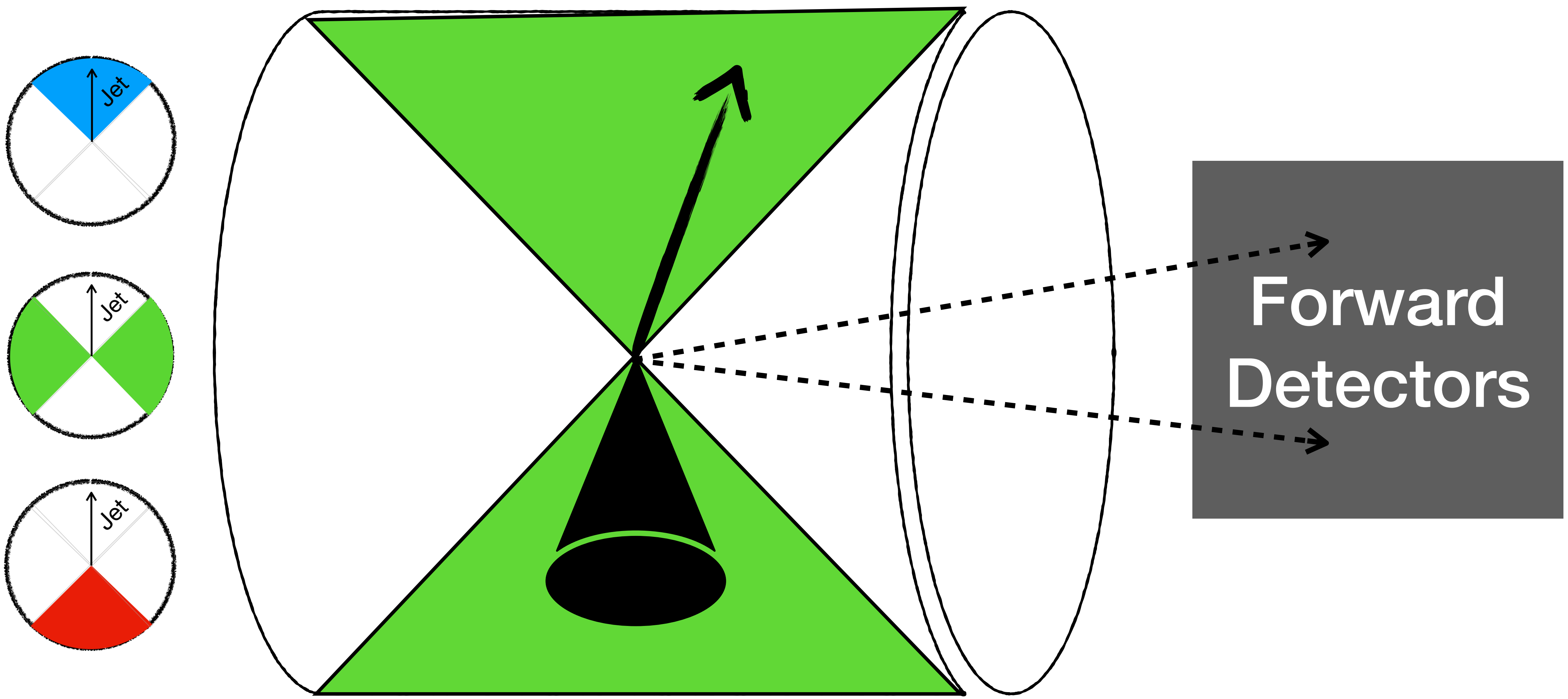


How about $\sqrt{s} = 13$ TeV?

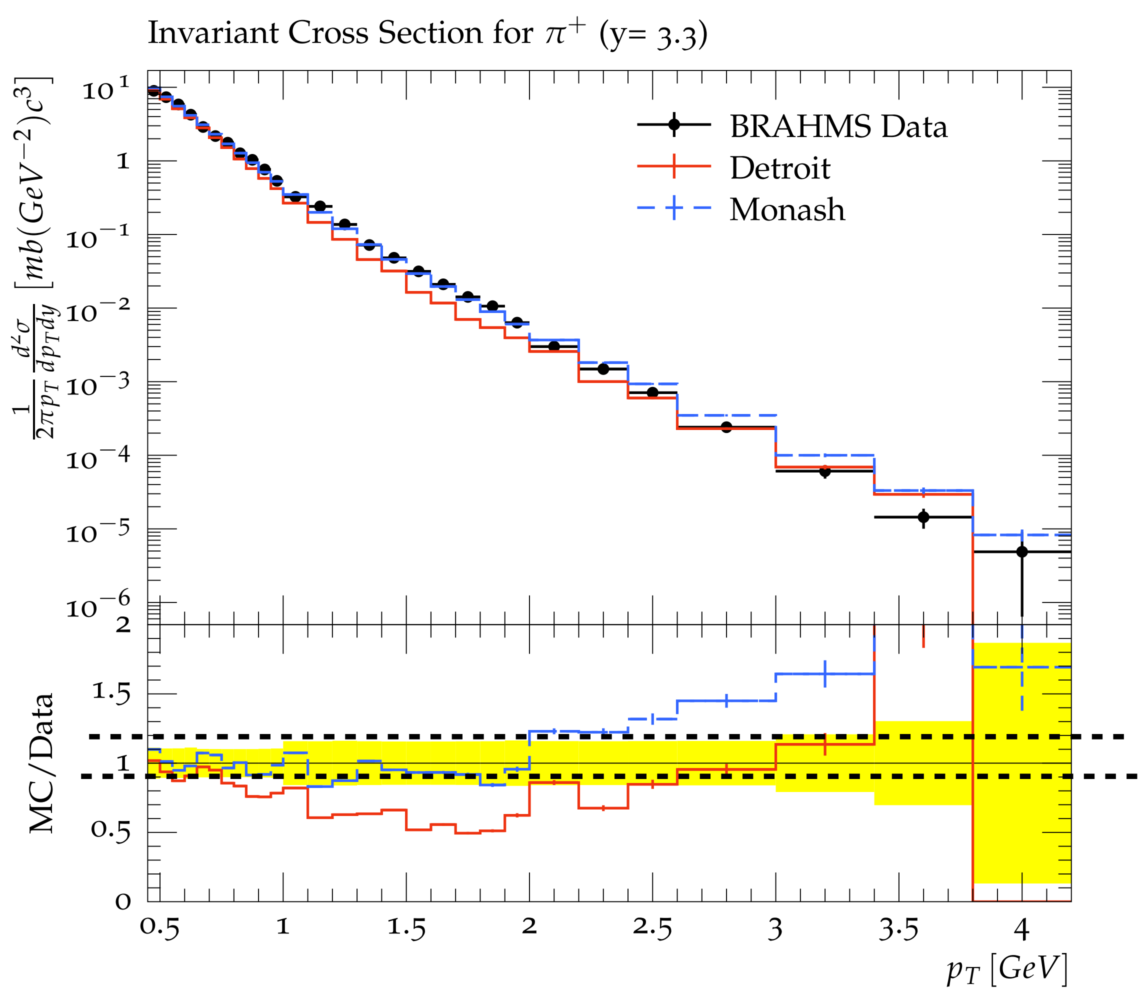


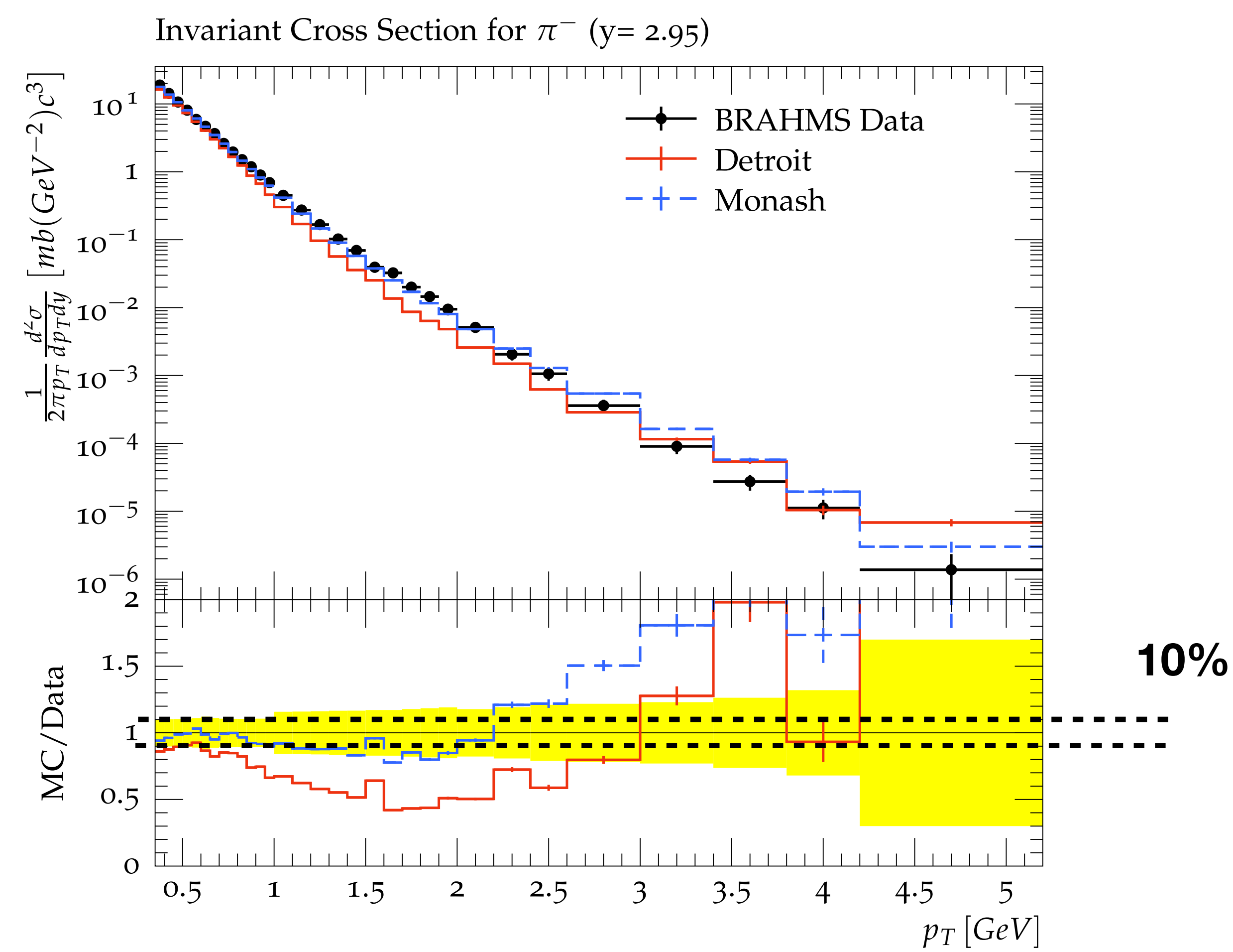
- Events with low p_T^{jet} - Monash tune better in describing data
- Events with high p_T^{jet} - Detroit outperforms all other tunes

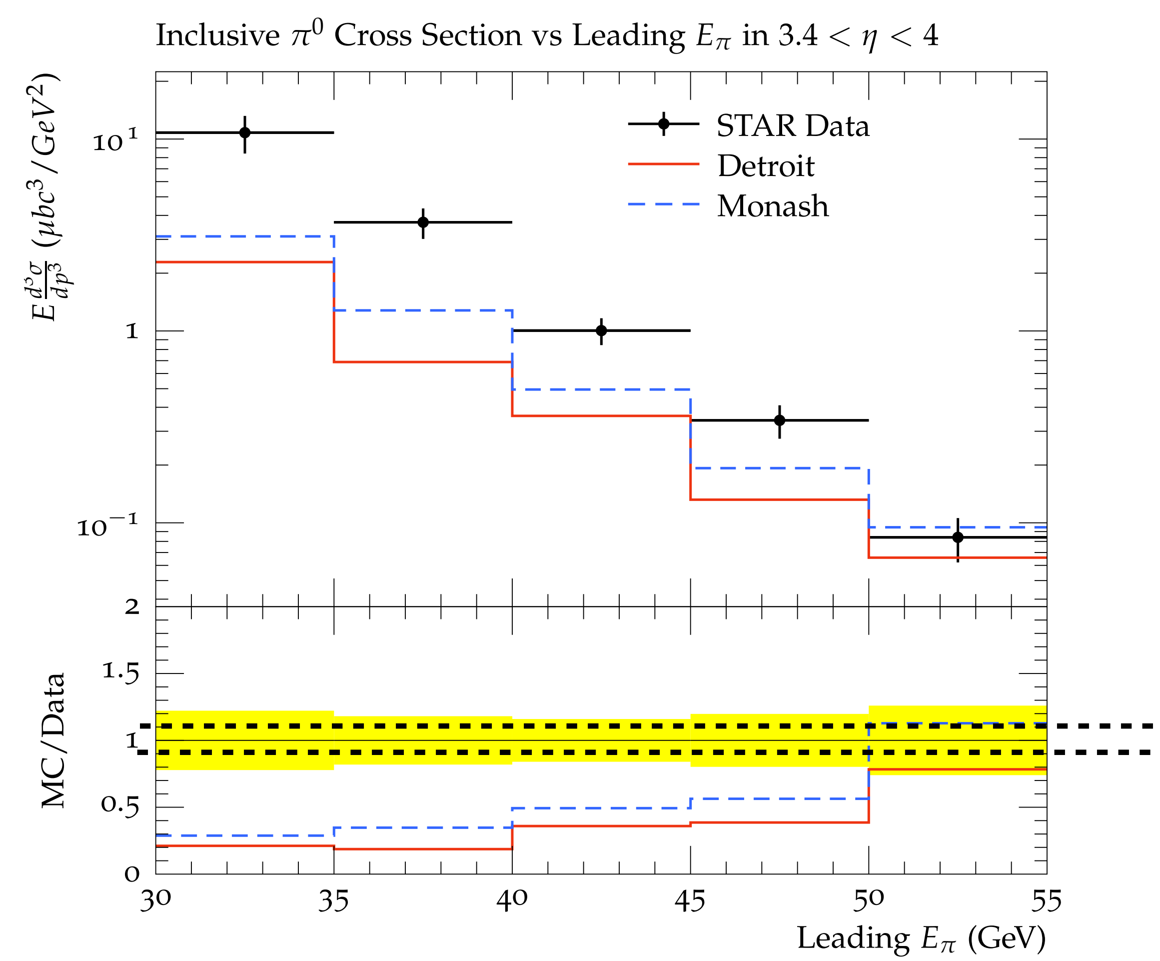
Lets move 'forward'



- Everything we have shown so far is only using the mid-rapidity particle production and jets
- Describing both kinematic ranges simultaneously is a crucial requirement for upcoming science goals both in the near and long term future







- Both Monash and Detroit tunes can't describe forward pion spectra from BRAHMS or STAR
- New tune does worse than Monash
- Simultaneous tune with mid-rapidity and larger tune-able phase space (ISR) unable to recover MC/data agreement

Summary

- First exercise in tuning PYTHIA 8 event generator with RHIC data

- ‘Detroit’ tune showcases the universality in MC generators that potentially work across 2 orders of magnitude in collision energies, without additional tuning

    - Recommend current and upcoming experiments at both RHIC and LHC to utilize this tune - reduction in the mis-match between data and simulation leads to more precision in measurements

- Simultaneous tuning of mid-rapidity and forward-rapidity particle production does not converge!

    - Impact of PDFs non negligible - will be improved with data from STAR-forward

    - Maybe we are hitting the inherent limitations of the parameterization framework of particle production in PYTHIA 8…

Next Steps

- The Monash tuning exercise was significantly broad in its parameter space with tuning string breaking and Lund models in addition to MPI
- We have RHIC data on strangeness production and baryon/meson ratios which PYTHIA
- Regarding Heavy Flavor production - we have a separate STAR HF tune discussed in recent publications which could be incorporated
- Relatively straightforward to expand the procedure to include these new analysis (in RIVET) and run the tuning
- Also expand to other MC generators such as HERWIG or SHERPA - explore the impact of varying hadronization models in particle production across rapidities

```
#----------------------------------------------------------------------------
# Relative production ratio vector/pseudoscalar for charm and bottom mesons
# This was originally PARJ(13) where PARJ(13) = V/(PS+V) that is the
# vector meson  fraction of primary charm+bottom mesons.
# Andre David (CERN/NA60) made an exhaustive study and found that the
# world data supports 0.6 while PYTHIA default was PARJ(13) = 3/4 = 0.75
# from simple spin counting.
# In PYTHIA8 we now use V/PS not V/(PS+V)
# Documentation: <pyhiadir>/htmldoc/FlavourSelection.html
#----------------------------------------------------------------------------
StringFlav:mesonCvector = 1.5    ! same as PARJ(13)=0.6
StringFlav:mesonBvector = 3      ! leave at 0.75

#----------------------------------------------------------------------------
# Heavy quark masses.
# Note that this should match with the ones used in the PDF.
# The masses are listed in the header of the refering PDF file.
# Documentation: <pyhiadir>/htmldoc/ParticleDataScheme.html
# Documentation: <pyhiadir>/htmldoc/ParticleData.html
#----------------------------------------------------------------------------
4:m0 = 1.43
5:m0 = 4.30
```

```
# EE4C Tune
set /Herwig/UnderlyingEvent/MPIHandler:EnergyExtrapolation Power
set /Herwig/UnderlyingEvent/MPIHandler:ReferenceScale 7000.*GeV
set /Herwig/UnderlyingEvent/MPIHandler:Power 0.24
set /Herwig/UnderlyingEvent/MPIHandler:pTmin0 3.91*GeV
#Colour reconnection settings
set /Herwig/Hadronization/ColourReconnector:ColourReconnection Yes
set /Herwig/Hadronization/ColourReconnector:ReconnectionProbability 0.61
#Colour Disrupt settings
set /Herwig/Partons/RemnantDecayer:colourDisrupt 0.75
#inverse hadron radius
set /Herwig/UnderlyingEvent/MPIHandler:InvRadius 1.35
#MPI model settings
set /Herwig/UnderlyingEvent/MPIHandler:softInt Yes
set /Herwig/UnderlyingEvent/MPIHandler:twoComp Yes
set /Herwig/UnderlyingEvent/MPIHandler:DLmode 2
```