Scale dependence: some issues regarding topologies/kinematics for inclusive jets/W+jet/H+jets

J. Huston
Michigan State University

Taming Unphysical Scales for Physical Predictions Cambridge 2017

Scale choices

- Motivated choice of scale very important at NLO
- Still relevant at NNLO
  - see Nigel's talk at PSR; James Currie's talk here
- In some cases, relevant scale choice is obvious
  - p_T^{jet} for inclusive jet cross sections
  - or should it be $p_T^{jet}/2$
  - …or event $p_T^{leadjet}$
- Does the preferred scale choice depend on the kinematics and/or the topology of the event?
  - if so, to what extent should this be taken into account?
- More and more NNLO calculations are becoming available allowing more data to be used in NNLO PDF fits
  - best scales to use?
- What about more complex situations, such as W+5 jets, or Higgs+3 jets, where there are multiple scales in the event?
- Does a global scale, like $H_T/2$ work (all the time)?
  - does it also depend on kinematics/topology?
- Why does it work?
- Can we make a connection to MC-like scales, such as with using a MINLO-approach?
  - see Stefan H's talk here
- I'll also touch on some issues regarding data incompatibilities discussed in the PSR workshop

Some pedagogy: write out the explicit scale dependent terms (at NLO; more terms at NNLO)

- Write cross section indicating explicit scale-dependent terms
- First term (lowest order) in (3) leads to monotonically decreasing behavior as scale increases (the LO piece)
- Second term is negative for $\mu<p_T$, positive for $\mu>p_T$
- Third term is negative for factorization scale $M < p_T$
- Fourth term has same dependence as lowest order term
- Thus, lines one and four give contributions which decrease monotonically with increasing scale while lines two and three start out negative, reach zero when the scales are equal to p_T, and are positive for larger scales
- At NLO, with $\mu_R=\mu_F$, result is a roughly parabolic behavior (if the kinematics are favorable)
  - 2D is a saddle shape

Consider a large transverse momentum process such as the single jet inclusive cross section involving only massless partons. Furthermore, in order to simplify the notation, suppose that the transverse momentum is sufficiently large that only the quark distributions need be considered. In the following, a sum over quark flavors is implied. Schematically, one can write the lowest order cross section as

$$E\frac{d^3\sigma}{dp^3} \equiv \sigma = a^2(\mu)\,\hat{\sigma}_B \otimes q(M) \otimes q(M) \tag{1}$$

where $a(\mu) = \alpha_s(\mu)/2\pi$ and the lowest order parton-parton scattering cross section is denoted by $\hat{\sigma}_B$. The renormalization and factorization scales are denoted by μ *and* M, respectively. In addition, various overall factors have been absorbed into the definition of $\hat{\sigma}_B$. The symbol $\otimes$ denotes a convolution defined as

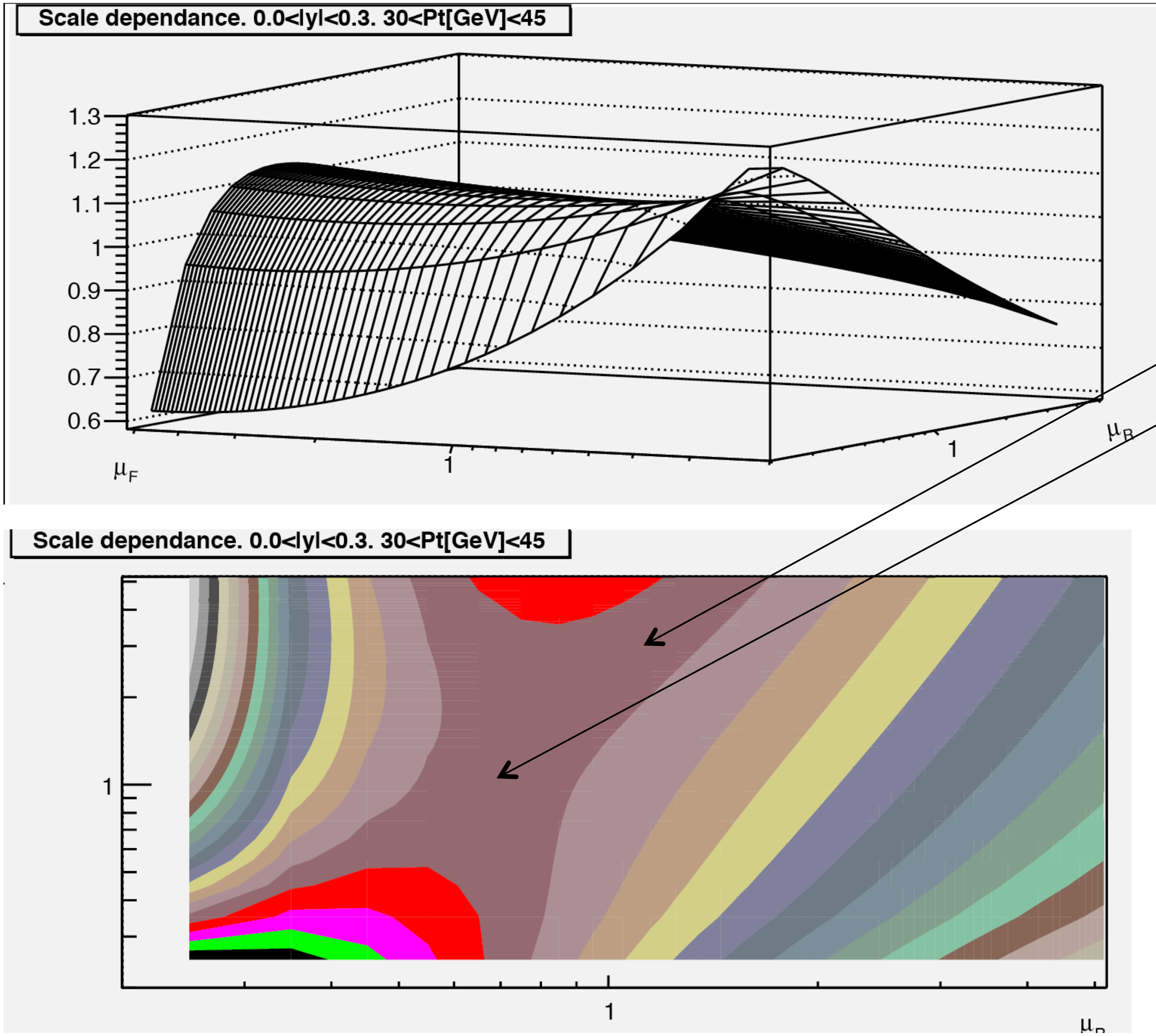
$$f \otimes g = \int_x^1 \frac{dy}{y}\, f(\frac{x}{y})\, g(y). \tag{2}$$

When one calculates the $\mathcal{O}(\alpha_s^3)$ contributions to the inclusive cross section, the result can be written as

$$\begin{aligned}
(1)\quad \sigma &= a^2(\mu)\,\hat{\sigma}_B \otimes q(M) \otimes q(M) \\
(2)\quad &+ 2a^3(\mu)\, b \ln(\mu/p_T)\hat{\sigma}_B \otimes q(M) \otimes q(M) \\
(3)\quad &+ 2a^3(\mu)\, \ln(p_T/M) P_{qq} \otimes \hat{\sigma}_B \otimes q(M) \otimes q(M) \\
(4)\quad &+ a^3(\mu)\, K \otimes q(M) \otimes q(M).
\end{aligned} \tag{3}$$

In writing Eq. (3), specific logarithms associated with the running coupling and the scale dependence of the parton distributions have been explicitly displayed; the remaining higher order corrections have been collected in the function K in the last line of Eq. (3). The μ

2D scale dependence: use log-log scale



- …since perturbative QCD is logarithmic
- Note that there's a saddle region, and a saddle point, where locally there is no slope for the cross section with respect to the two scales
- This is kind of the 'golden point' and typically around the expected scale (p_T^{jet} in this case)

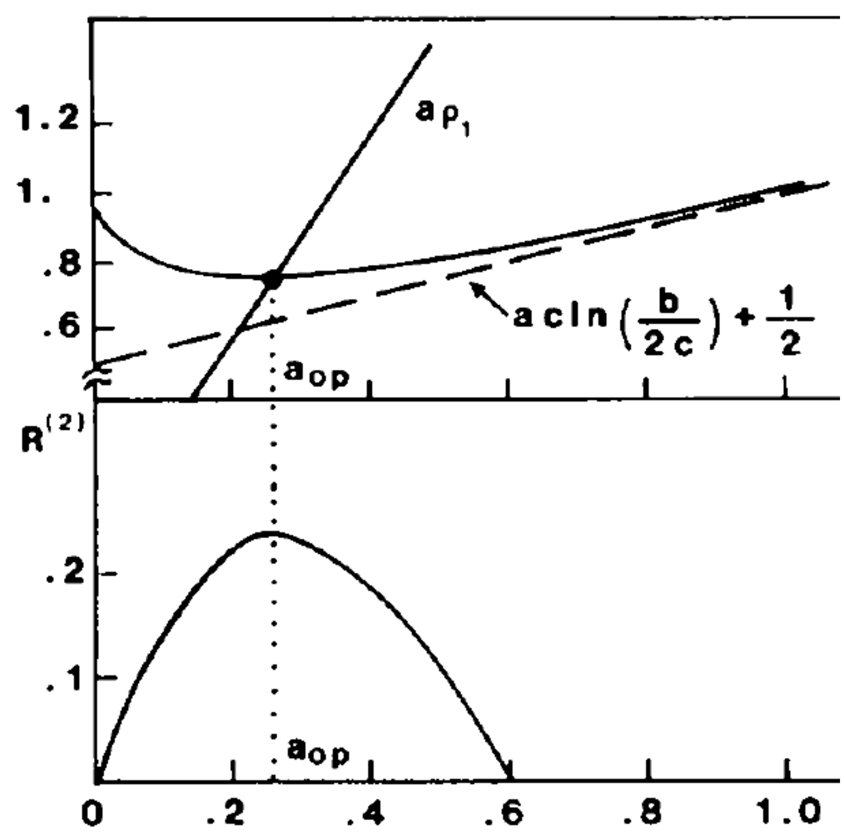
Looking for saddle points

- Can find (renormalization scale) saddle point analytically by solving a transcendental equation

$$\tau + \frac{1}{2}\frac{c}{1+ca} = \rho_1$$

- …where ρ_1 is a dimensionless form of the jet cross section, and τ depends on the scale μ and on Λ
- Can find saddle point for factorization scale by following ridge line.

P. Aurenche et al. / Higher order QCD prediction

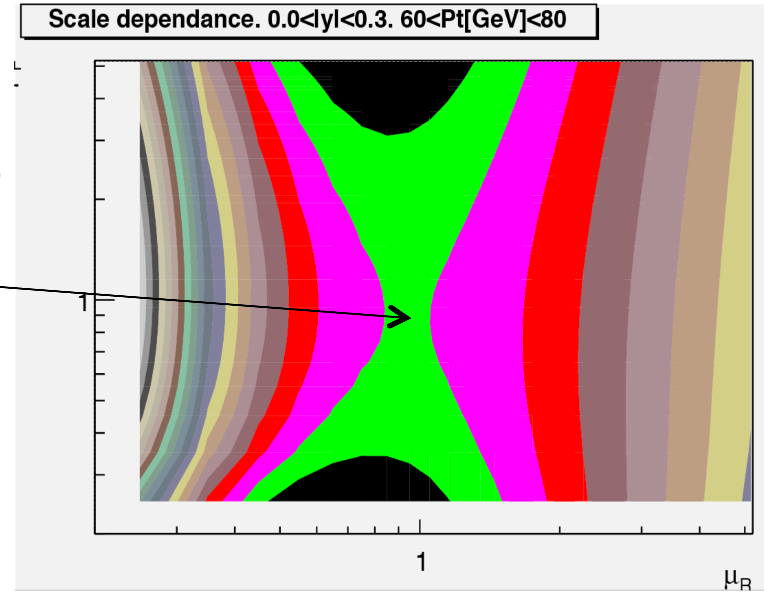


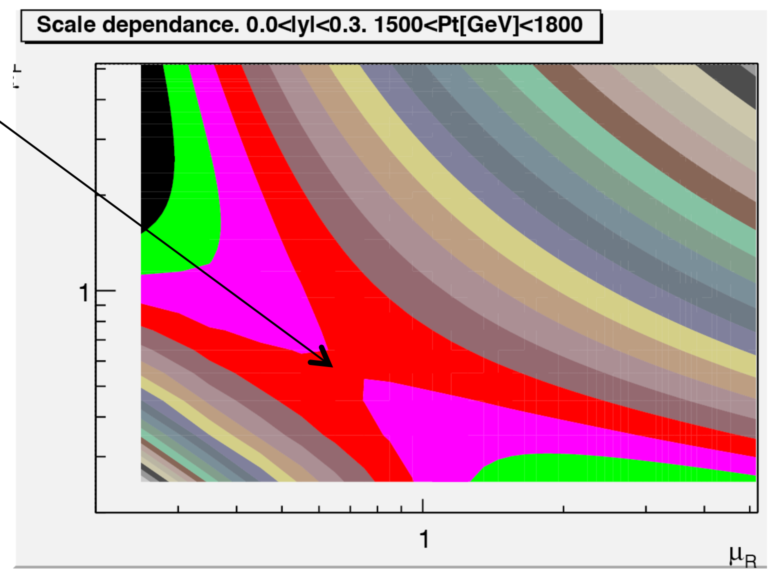
tion of Stevenson's equation for a_{op}; (b) plot of th
a function of a.

Scale choices

- Take inclusive jet production at the LHC
- Canonical scale choice at the LHC is $\mu_r=\mu_f=1.0*p_T$
  - CTEQ6.6 used $0.5p_T$ for determination of PDFs
  - CT10/14 uses p_T
- Close to saddle point for low p_T
- But saddle point moves down for higher p_T (and the saddle region rotates)
- Saddle point also moves down for larger jet size
- For some kinematics, there is no saddle point
- Cross sections can even become negative for some scales of p_T

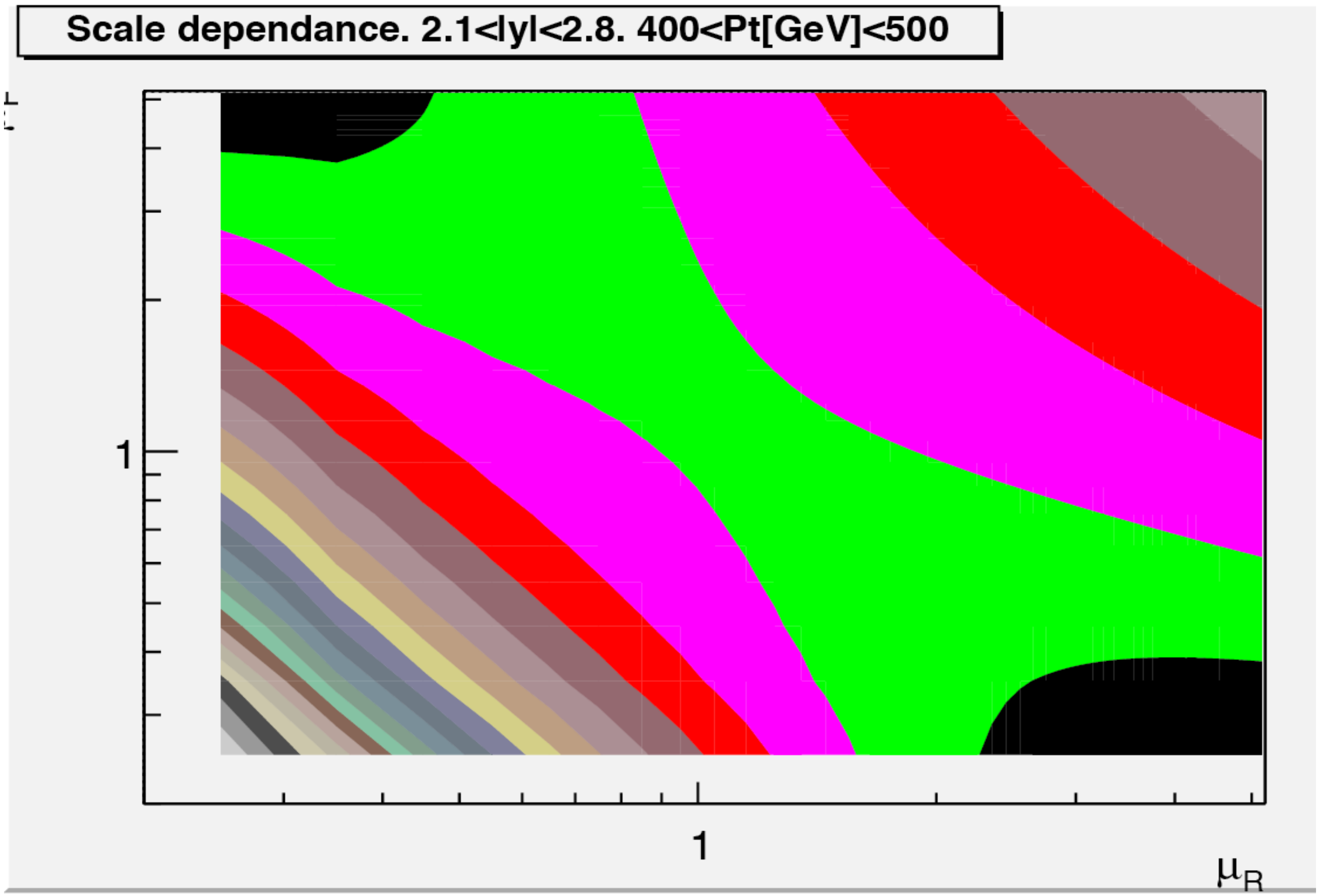
R=0.4
antikT





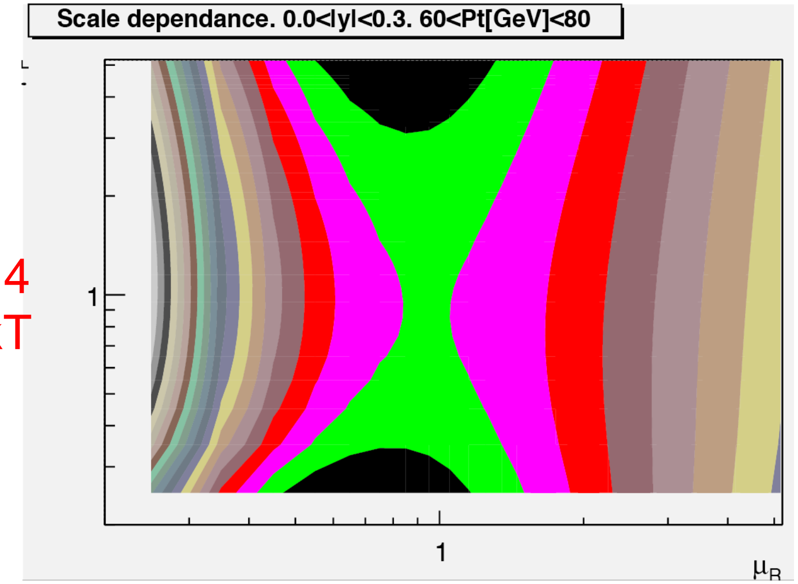
Scale dependence depends on rapidity

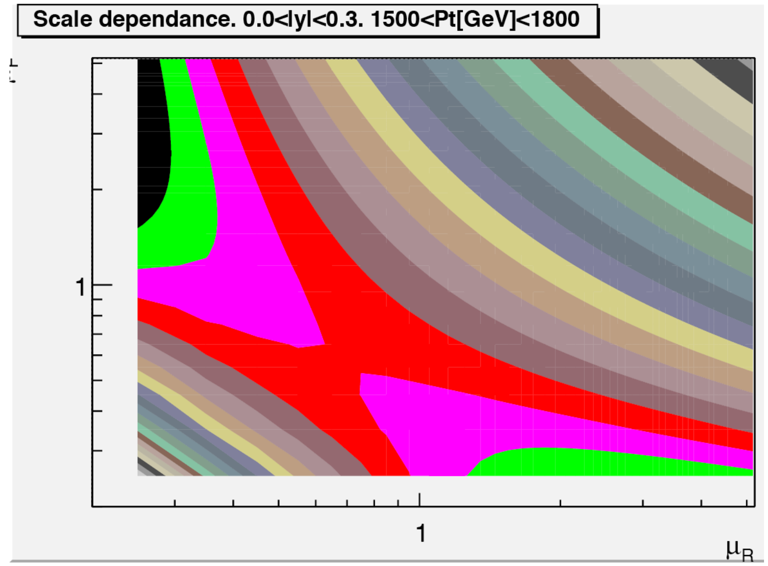
- The saddle point tends to move upwards in scale as the rapidity increases
- Is the physics changing; no, just the kinematics



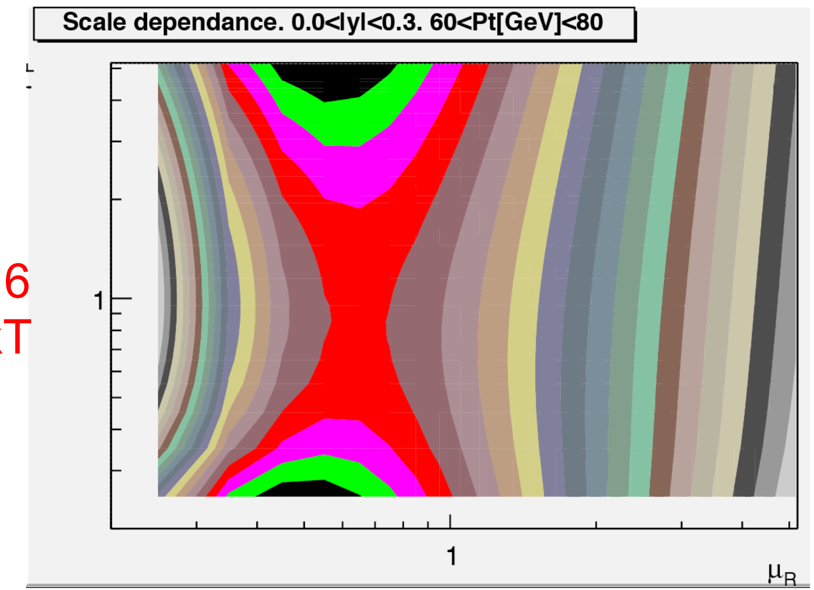
Scale dependence also depends on jet size; again see equation on previous page

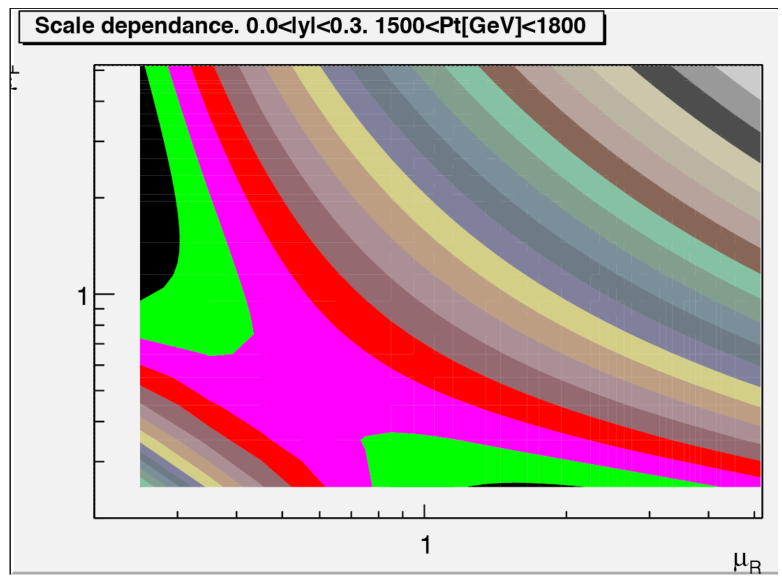
R=0.4
antikT





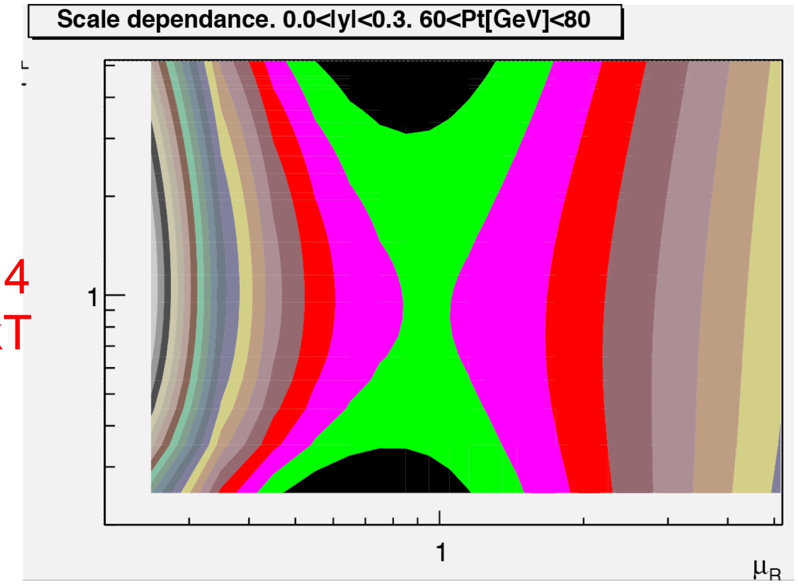
R=0.6
antikT

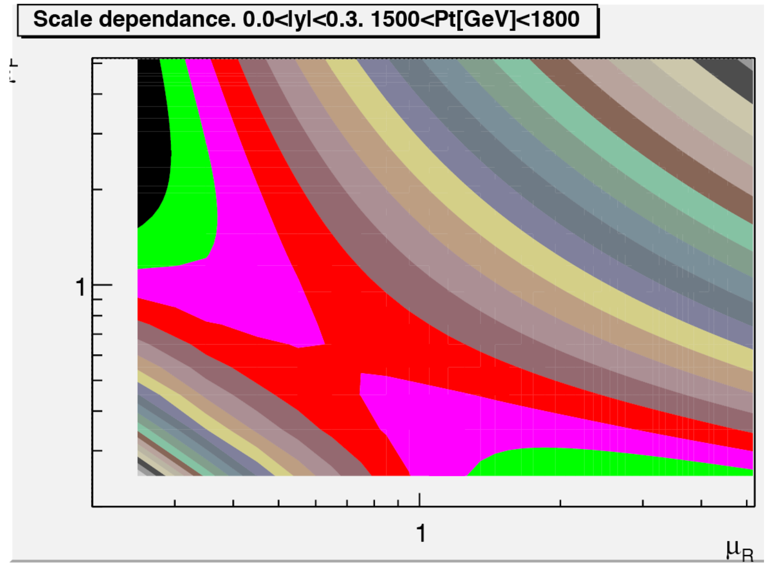




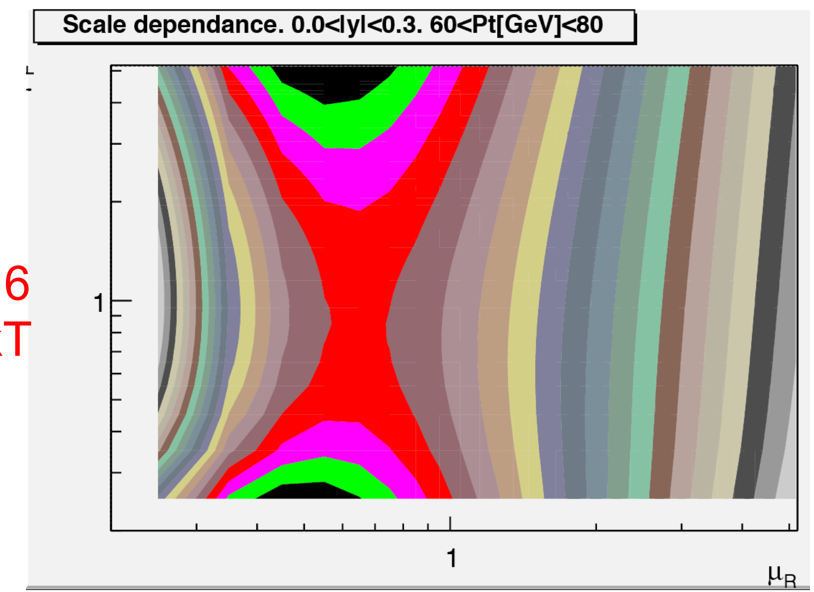
Note rotation of -45°; gluon distribution relatively stable at small x (p_T/y) to changes in scale

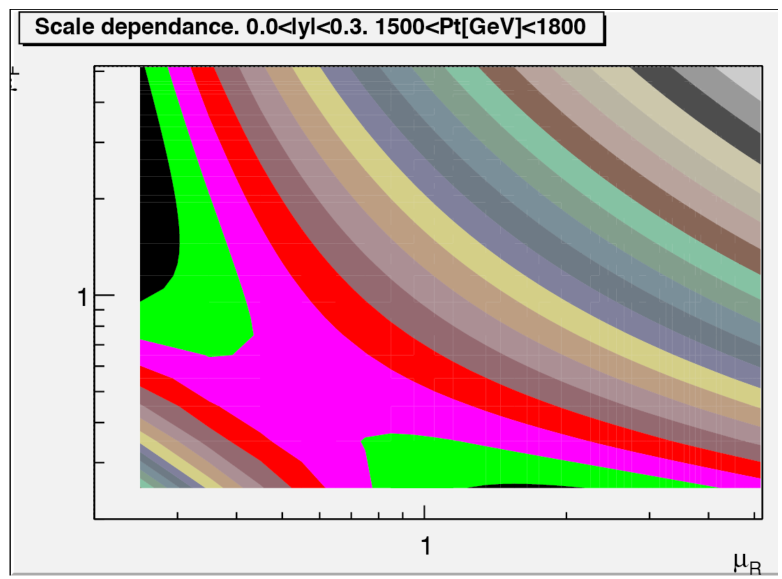
R=0.4
antikT





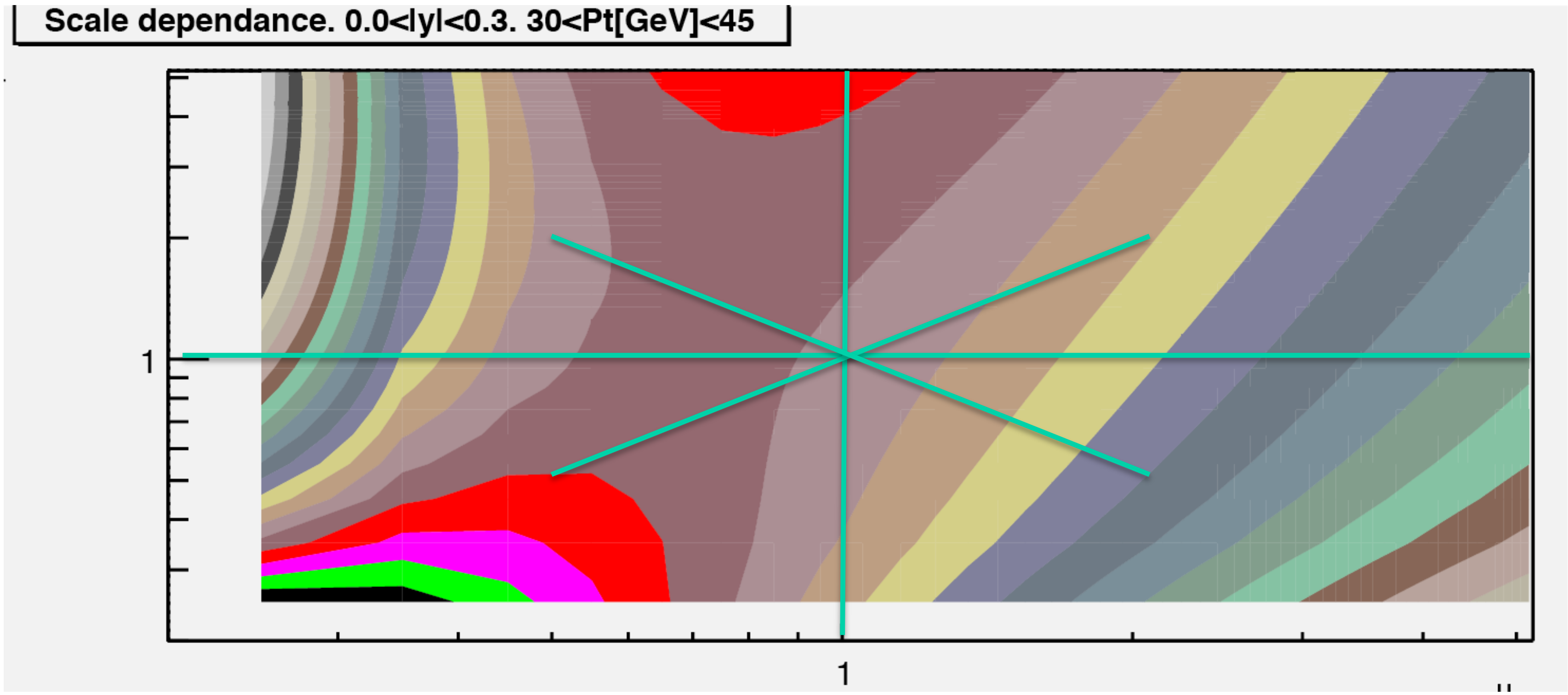
R=0.6
antikT





Scale variations

- Note that you may come up with a different estimate of the scale uncertainty depending on whether you use coherent or incoherent scale variations
- For many cross section, incoherent leads to smaller scale uncertainties



Another scheme

- F. Olness and D. Soper, arXiv: 0907.5052
- Define x_1 and x_2

$$x_1 = \log_2\left(\frac{\mu_{\rm uv}}{P_T/2}\right)$$

$$x_2 = \log_2\left(\frac{\mu_{\rm col}}{P_T/2}\right)$$

- Make a circle of radius |x|=2 around a central scale (could be saddle point, or could be some canonical scale) and evaluate the scale uncertainty

$$\left[\frac{d\sigma(x_1,x_2)}{dP_T}\right]_{\rm NLO} \approx \left[\frac{d\sigma(0,0)}{dP_T}\right]_{\rm NLO} \left[1 + P(\vec{x})\right]$$

where

$$P(\vec{x}) = \sum_J x_J A_J + \sum_{J,K} x_J M_{JK} x_K$$

A_J and M_{JK} carry information on the scale dependence beyond NLO

$$\mathcal{E}^2_{\rm scale} = \frac{1}{2\pi}\int_0^{2\pi} d\theta\ P(|\vec{x}|\cos\theta, |\vec{x}|\sin\theta)^2$$

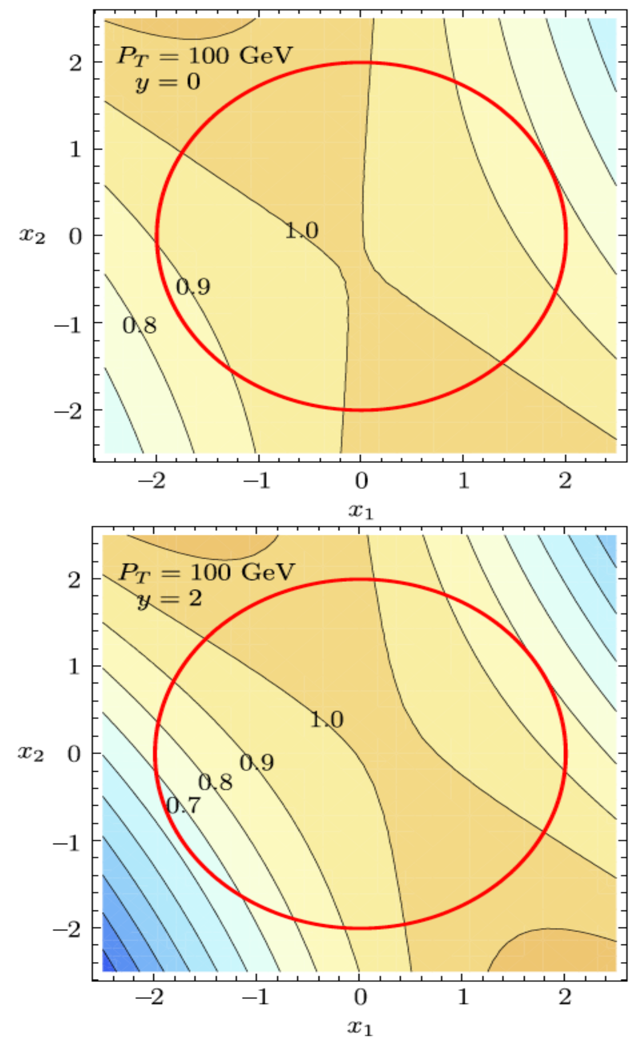
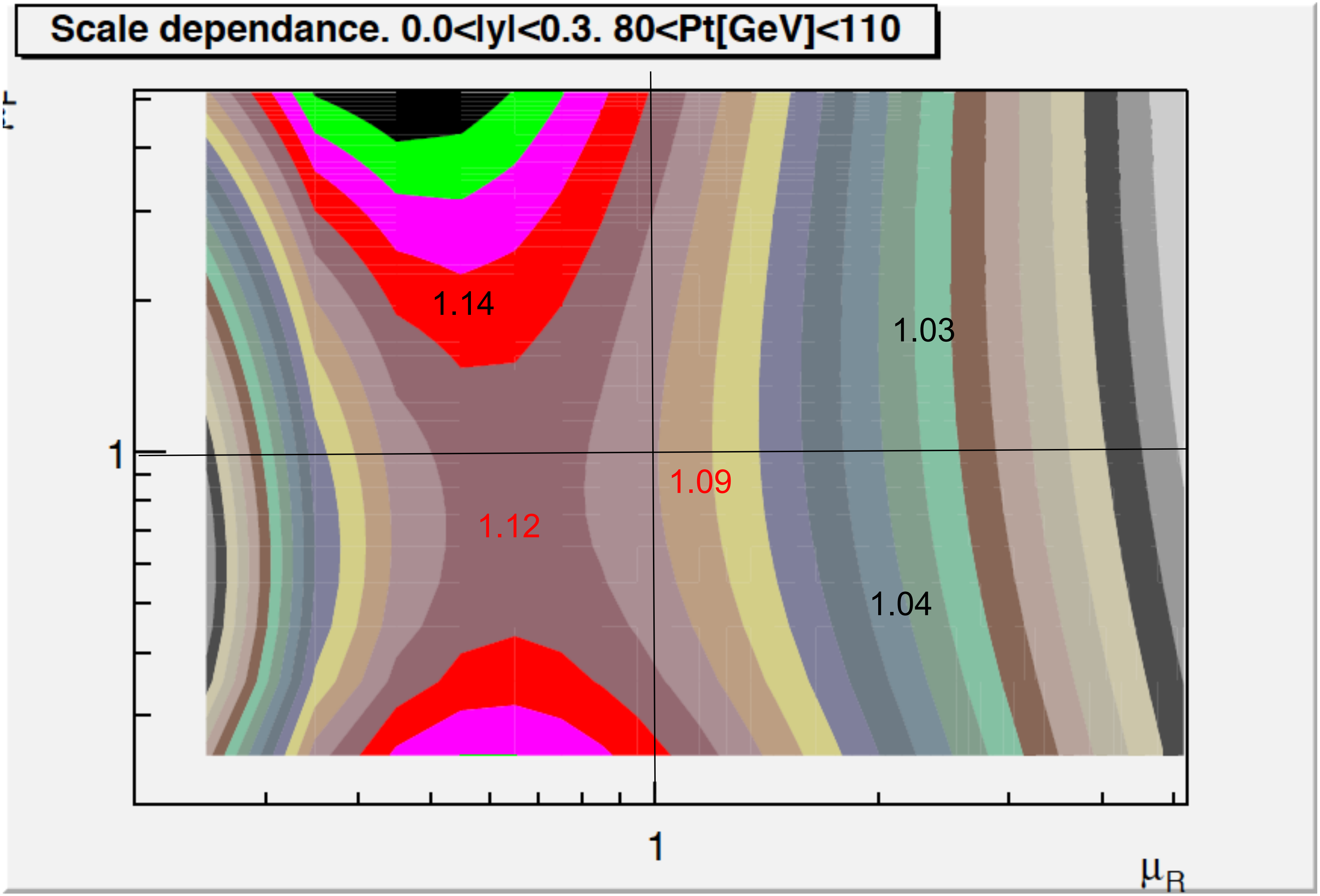
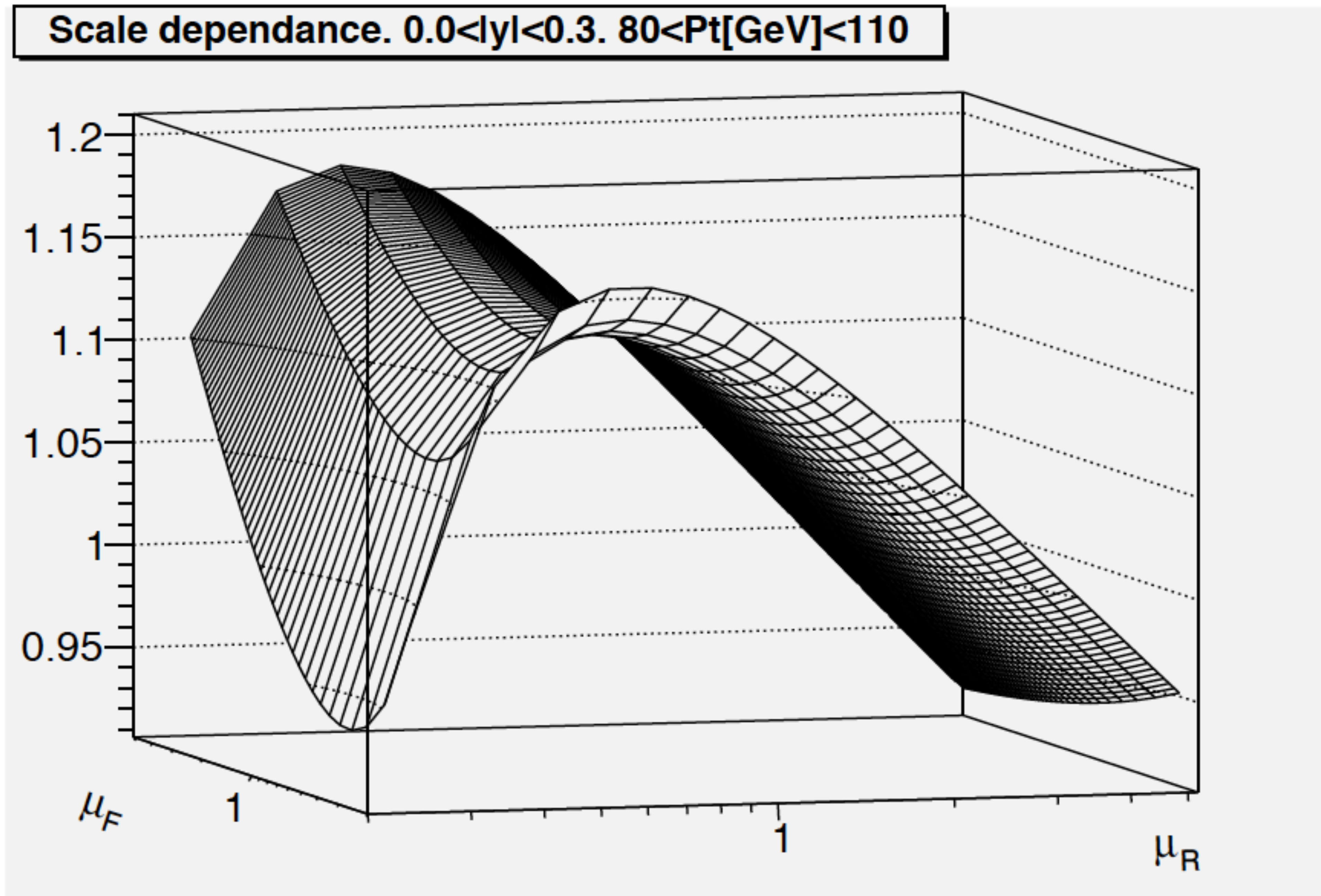


Figure 2: Contour plot of the jet cross section in the $\{x_1, x_2\}$ plane for the Tevatron ($\sqrt{s} = 1960$ GeV) with $P_T = 100$ GeV and a) central rapidity $y = 0$ and b) forward rapidity $y = 2$. We plot the ratio of the cross section compared to the central value at $\{x_1, x_2\} = \{0, 0\}$. Contour lines are drawn at intervals of 0.10. The (red) circle is at radius $|x| = 2$.

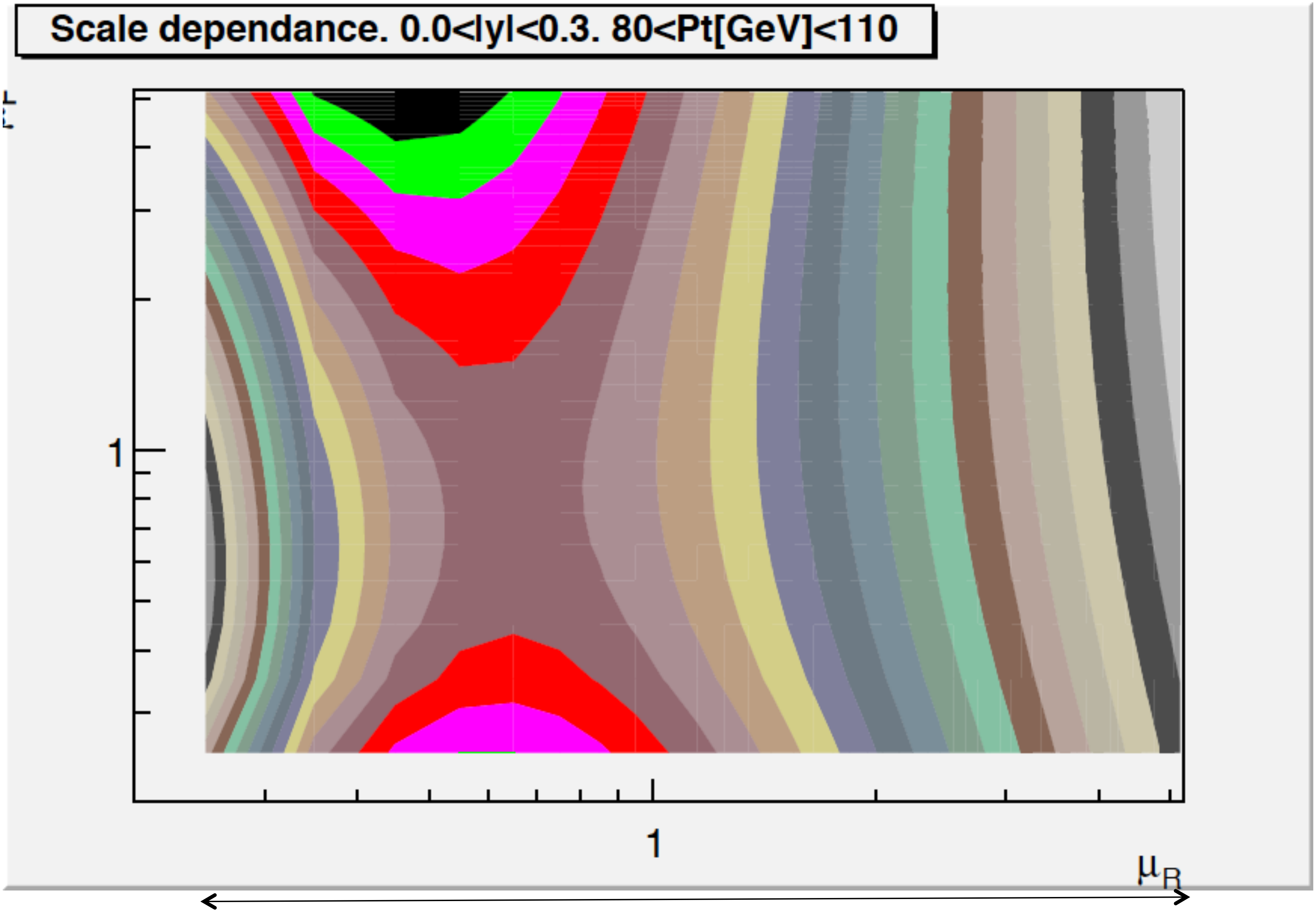
Scale variations for 7 TeV inclusive jets (R=0.6)

- R=0.6 because that's what PDF fitters will tend to use for the ATLAS data (CT will)
  - smaller dependence on higher order corrections; smaller differences between p_{Tjet} and p_{Tmax}
  - ROOT files using a 50X50 grid for μ_R and μ_F, from 0.1 to 50, with the cross sections normalized to the point (2.5,2.5) generated using NLOJET++ and applgrid (by Pavel Starovoitov)
- In the following slides, indicate the relative cross section at scales
  - (1.0,1.0) canonical choice
  - Saddle point
  - ($e^{0.3y}$,$e^{0.3y}$): rapidity dependent scale as advocated for example by Thorne
  - scale variations of (0.5,0.5), (2.0,2.0), (0.5,2.0) and (2.0,0.5) shown in black

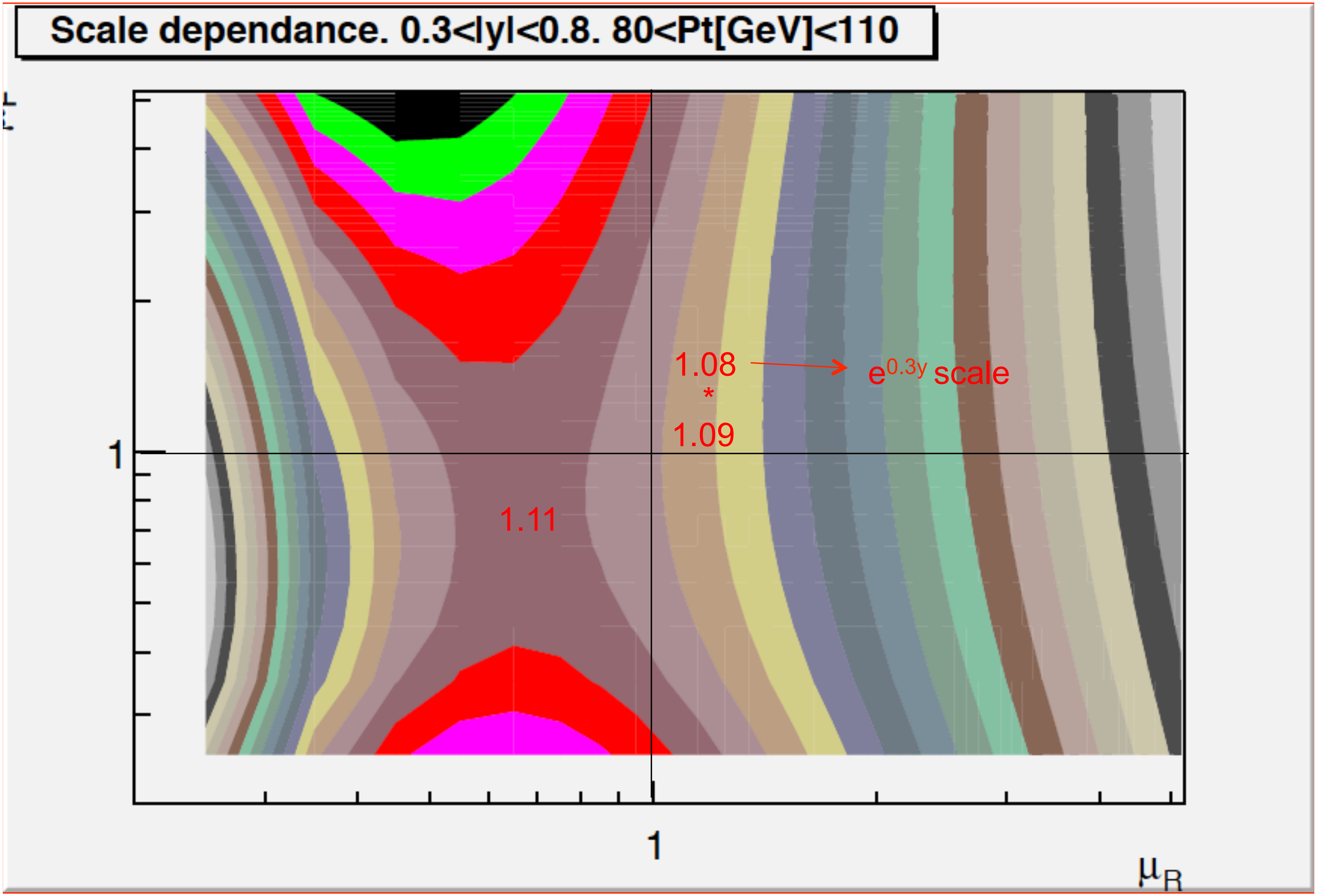

Scale dependance. 0.0<|y|<0.3. 80<Pt[GeV]<110
1.14
1.03
1
1.09
1.12
1.04
1
μ_R



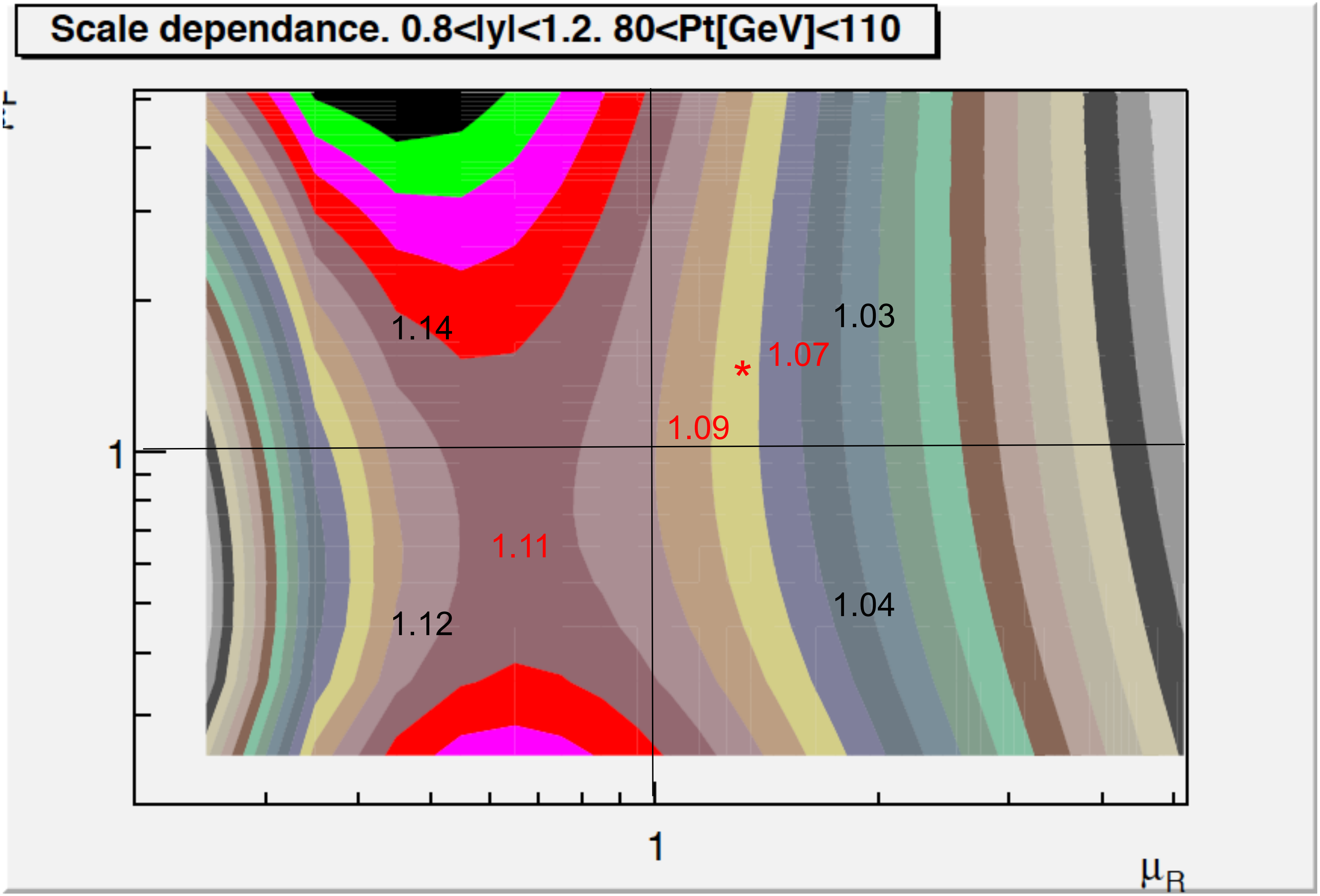
Scale dependance. 0.0<|y|<0.3. 80<Pt[GeV]<110
1.2
1.15
1.1
1.05
1
0.95
μ_F
1
1
μ_R



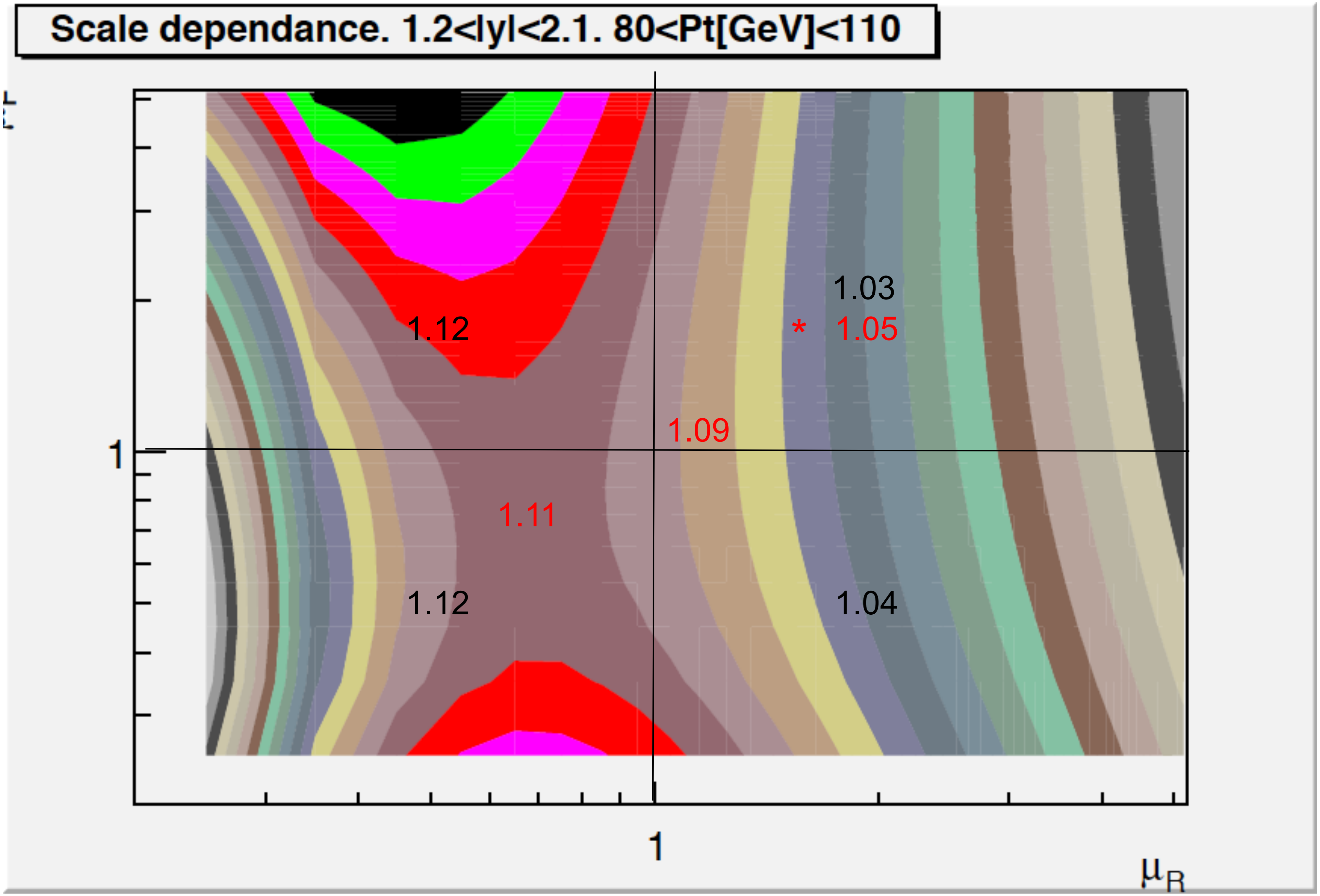
maximum variation in cross section is ~30%

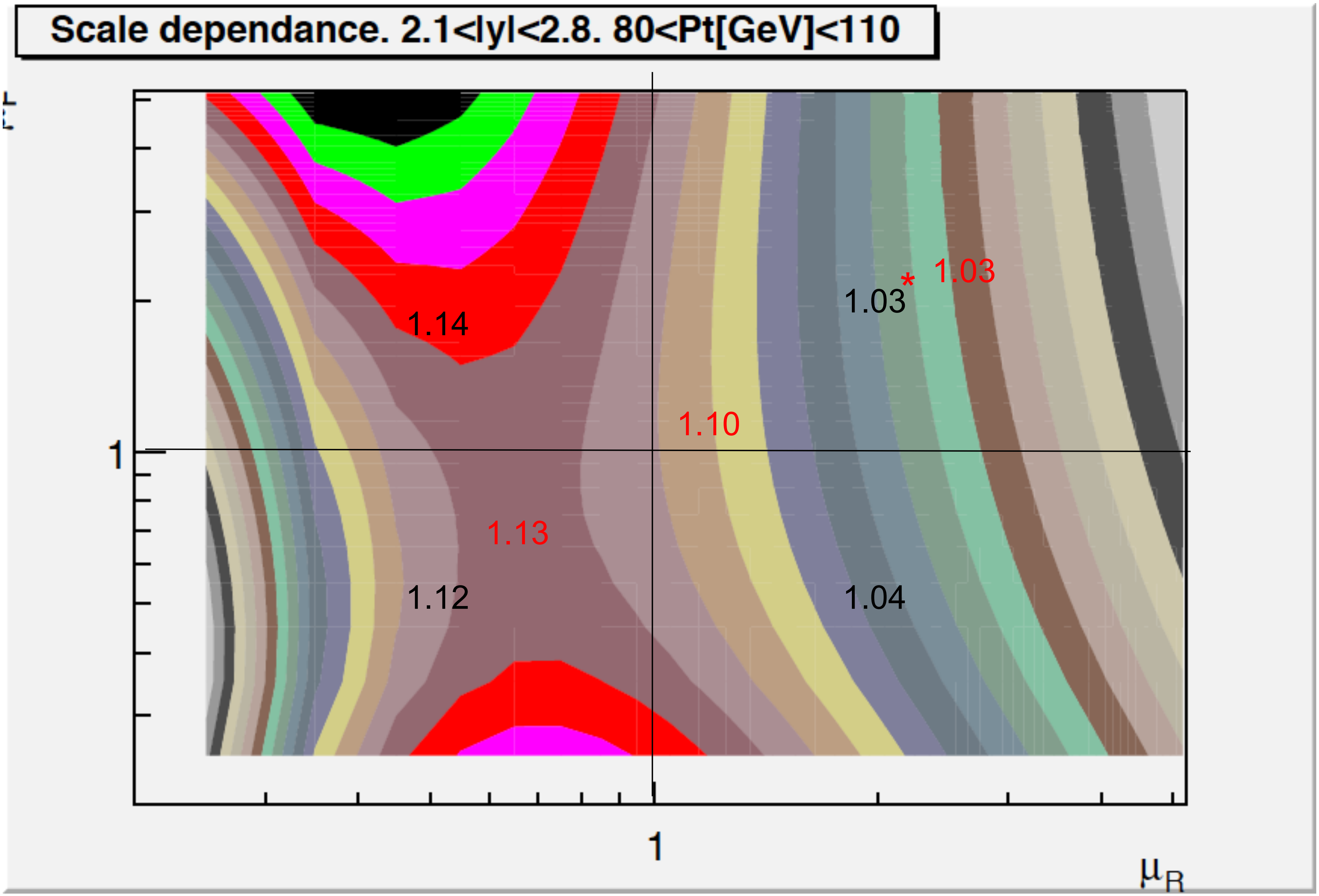


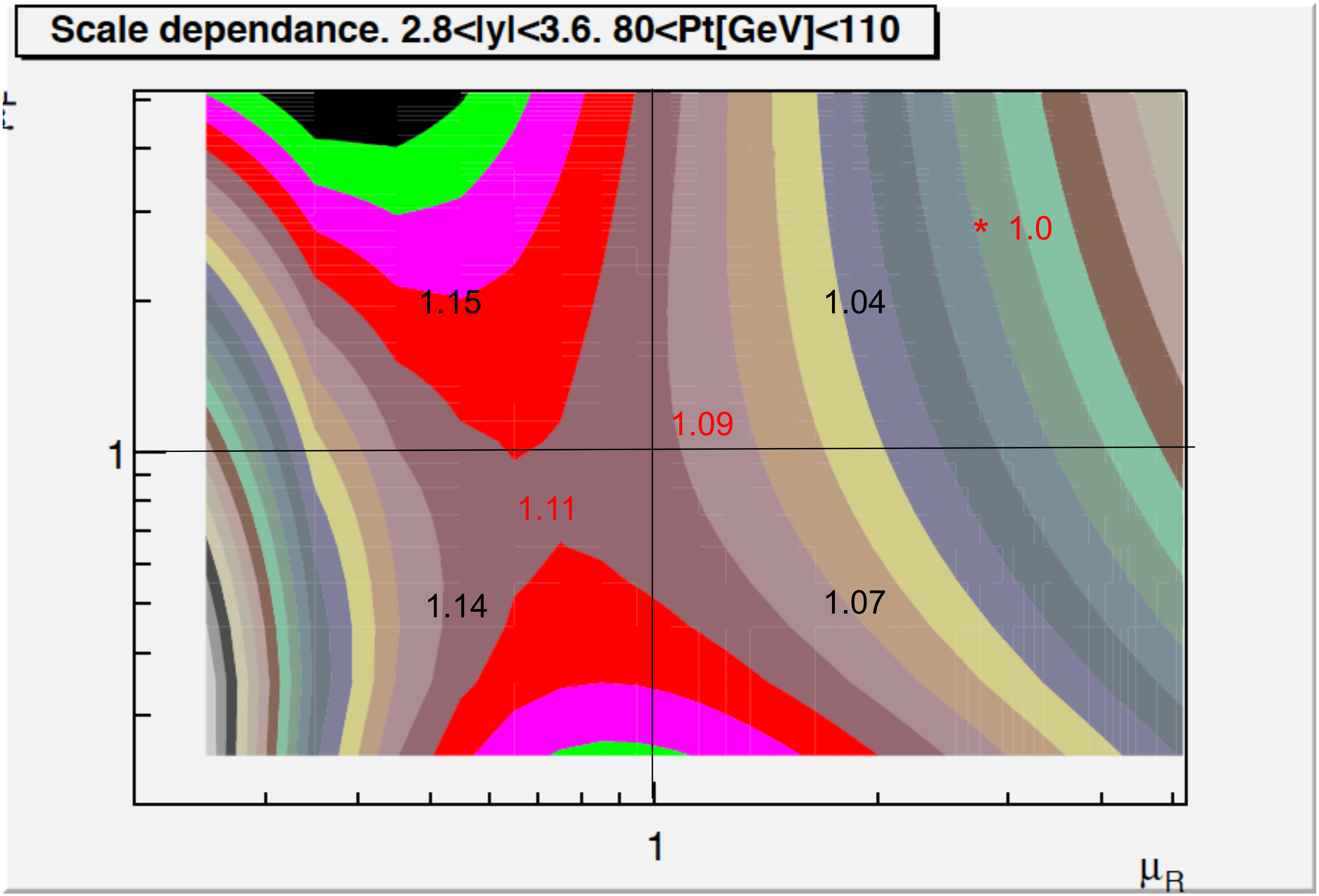
Scale dependance. 0.3<lyl<0.8. 80<Pt[GeV]<110
1.08
*
1.09
$e^{0.3y}$ scale
1.11
1
1
μ_R



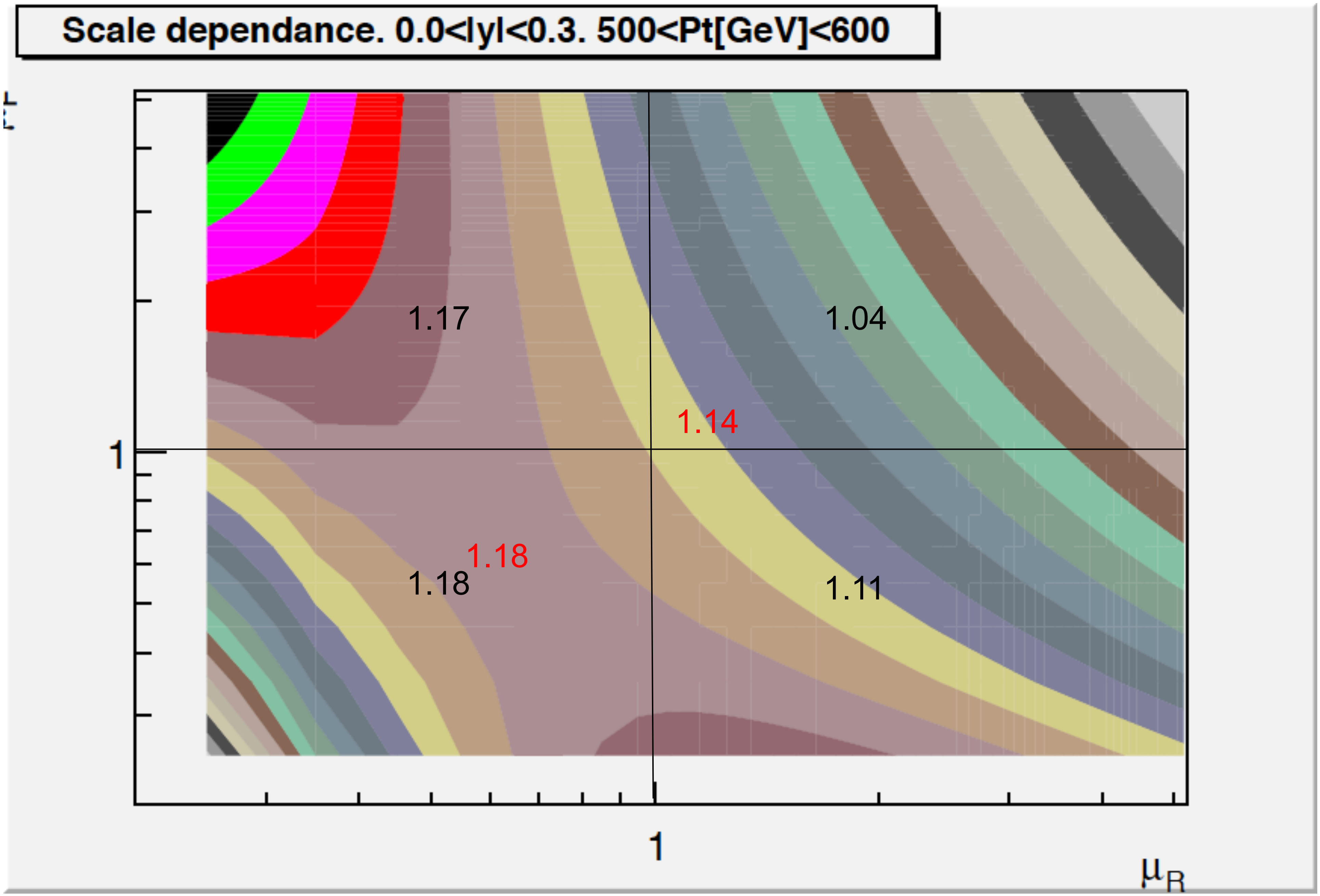
Scale dependance. 0.8<|y|<1.2. 80<Pt[GeV]<110
1.14
1.03
1.07
*
1.09
1
1.11
1.12
1.04
1
μ_R


Scale dependance. 1.2<|y|<2.1. 80<Pt[GeV]<110
1.03
1.12
* 1.05
1.09
1
1.11
1.12
1.04
1
μ_R

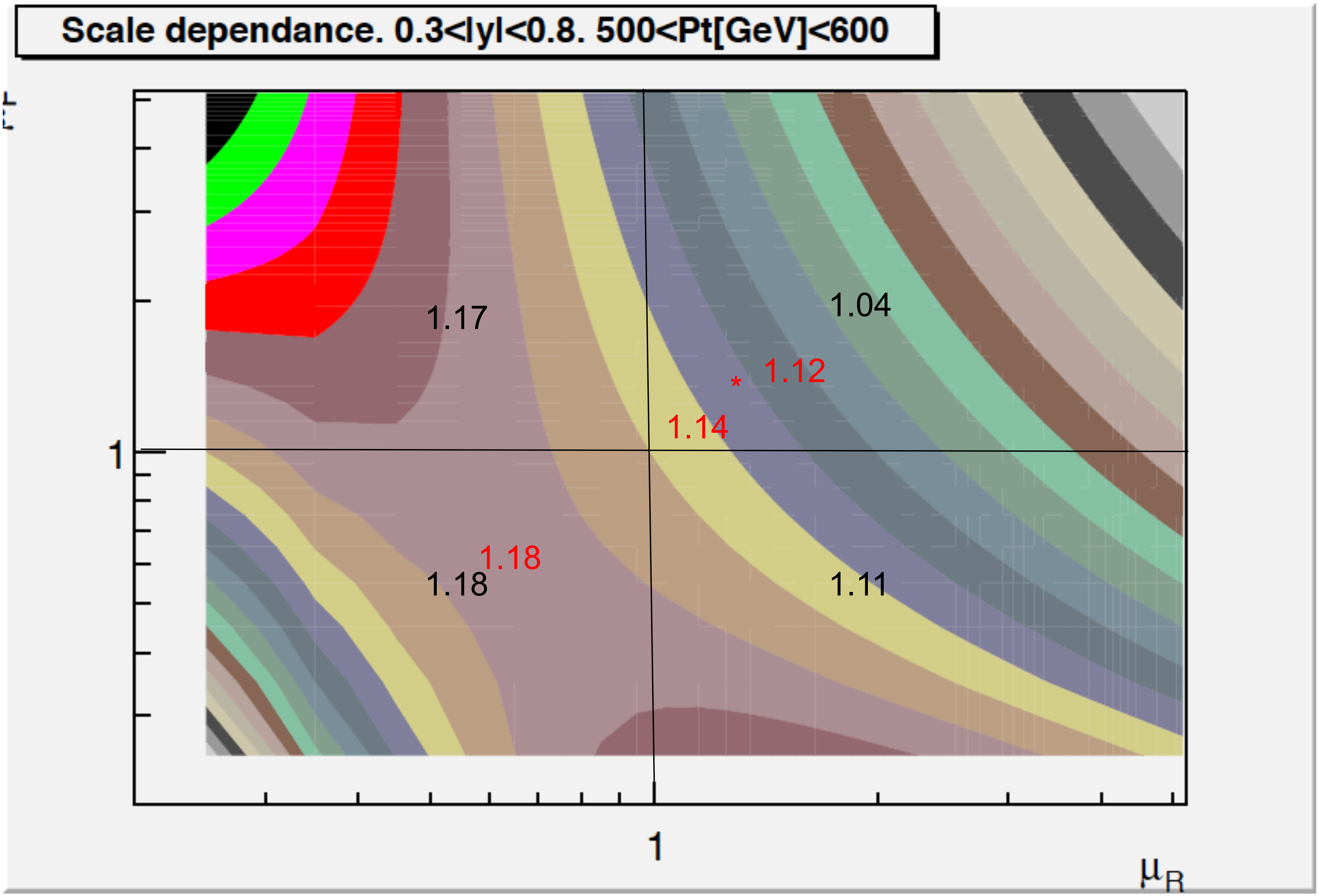

Scale dependance. 2.1<|y|<2.8. 80<Pt[GeV]<110
1.03
1.03
1.14
1.10
1
1.13
1.12
1.04
1
μR



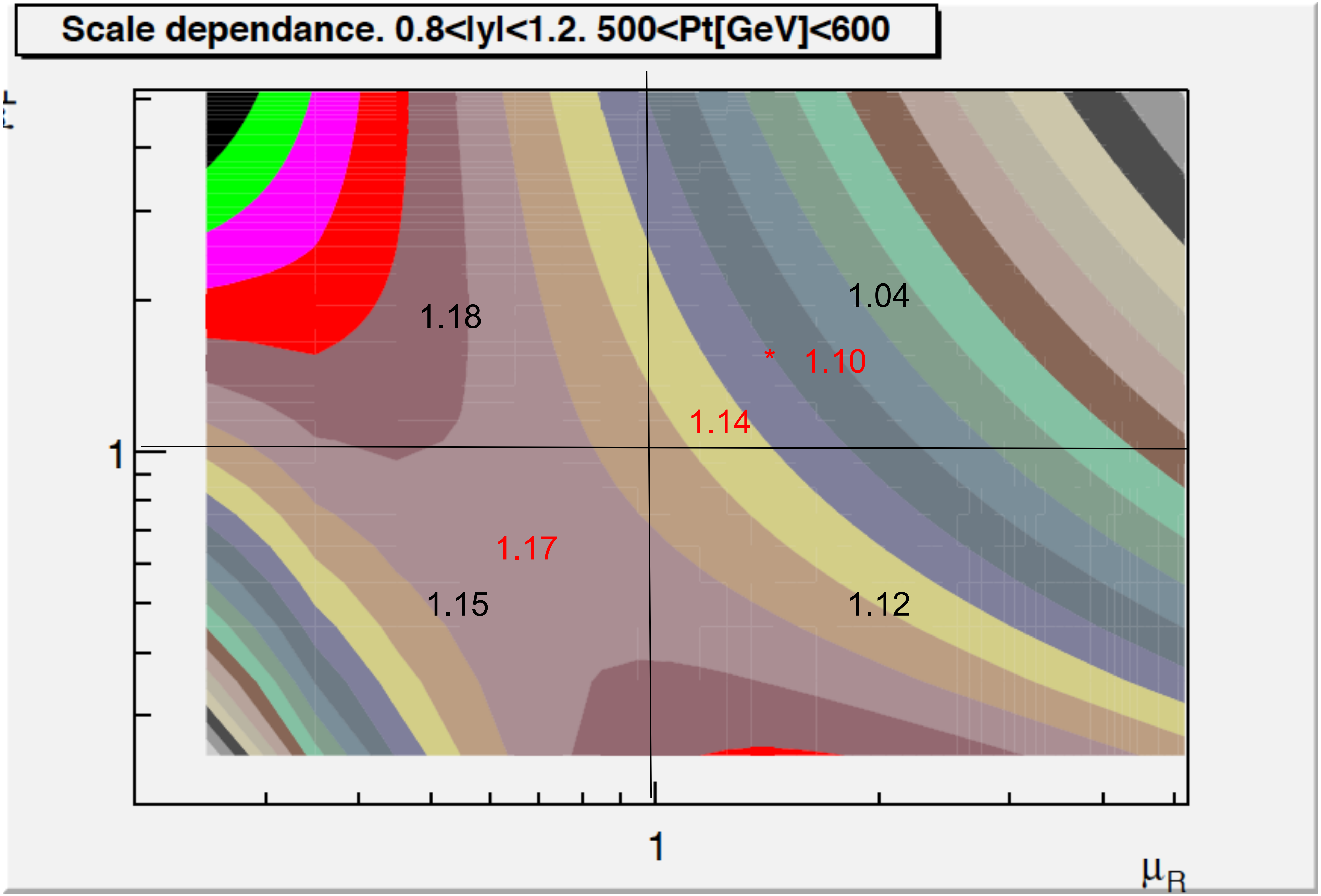
Scale dependance. 2.8<|y|<3.6. 80<Pt[GeV]<110
1.15
1.04
* 1.0
1.09
1.11
1.14
1.07
1
1
μ_R

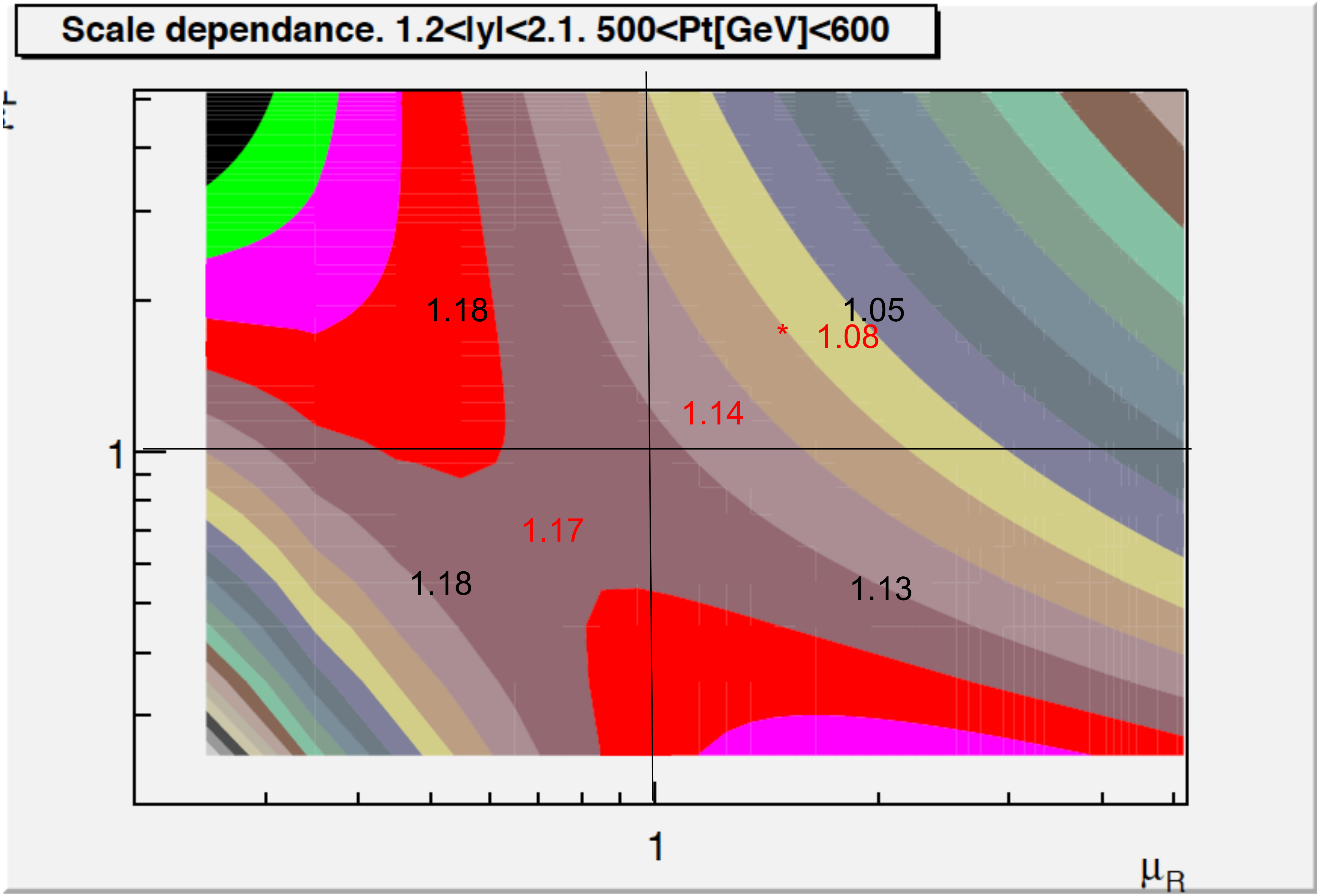


Scale dependance. 0.0<lyl<0.3. 500<Pt[GeV]<600
1.17
1.04
1.14
1.18
1.18
1.11
1
1
μ_R

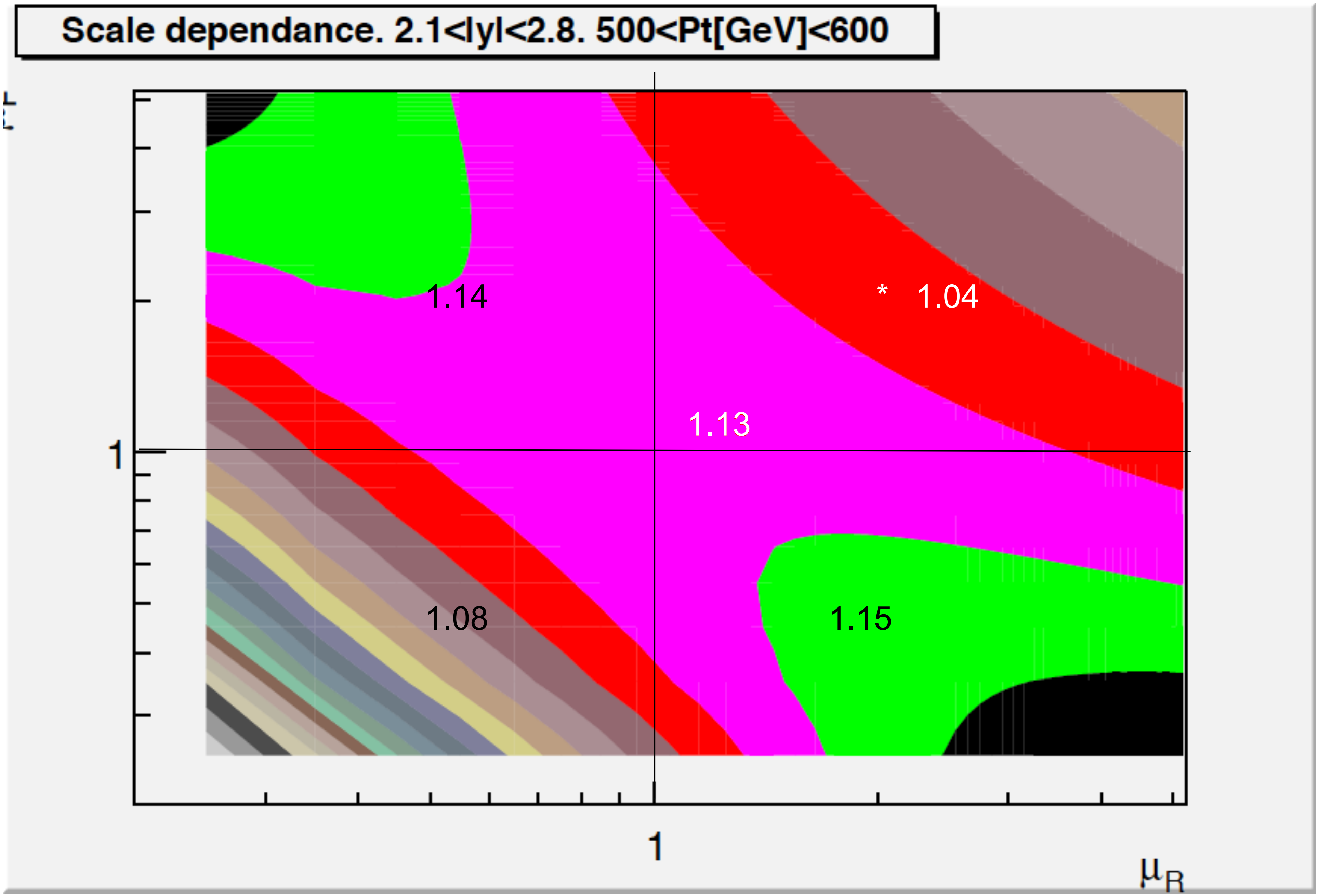


Scale dependance. 0.3<|y|<0.8. 500<Pt[GeV]<600
1.17
1.04
1.12
1.14
1
1.18
1.18
1.11
1
μ_R

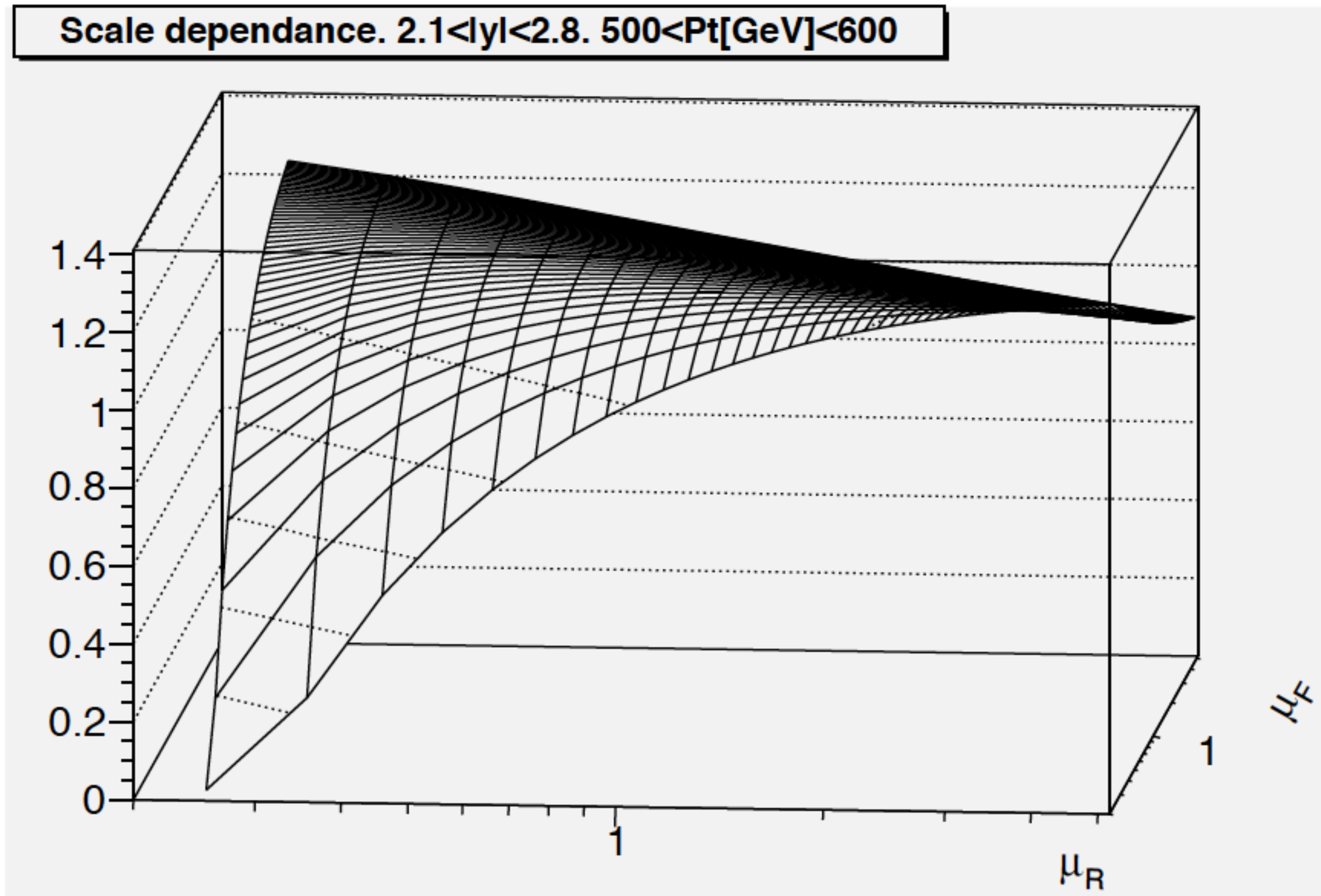

Scale dependance. 0.8<lyl<1.2. 500<Pt[GeV]<600
1.18
1.04
*
1.10
1.14
1
1.17
1.15
1.12
1
μ_R

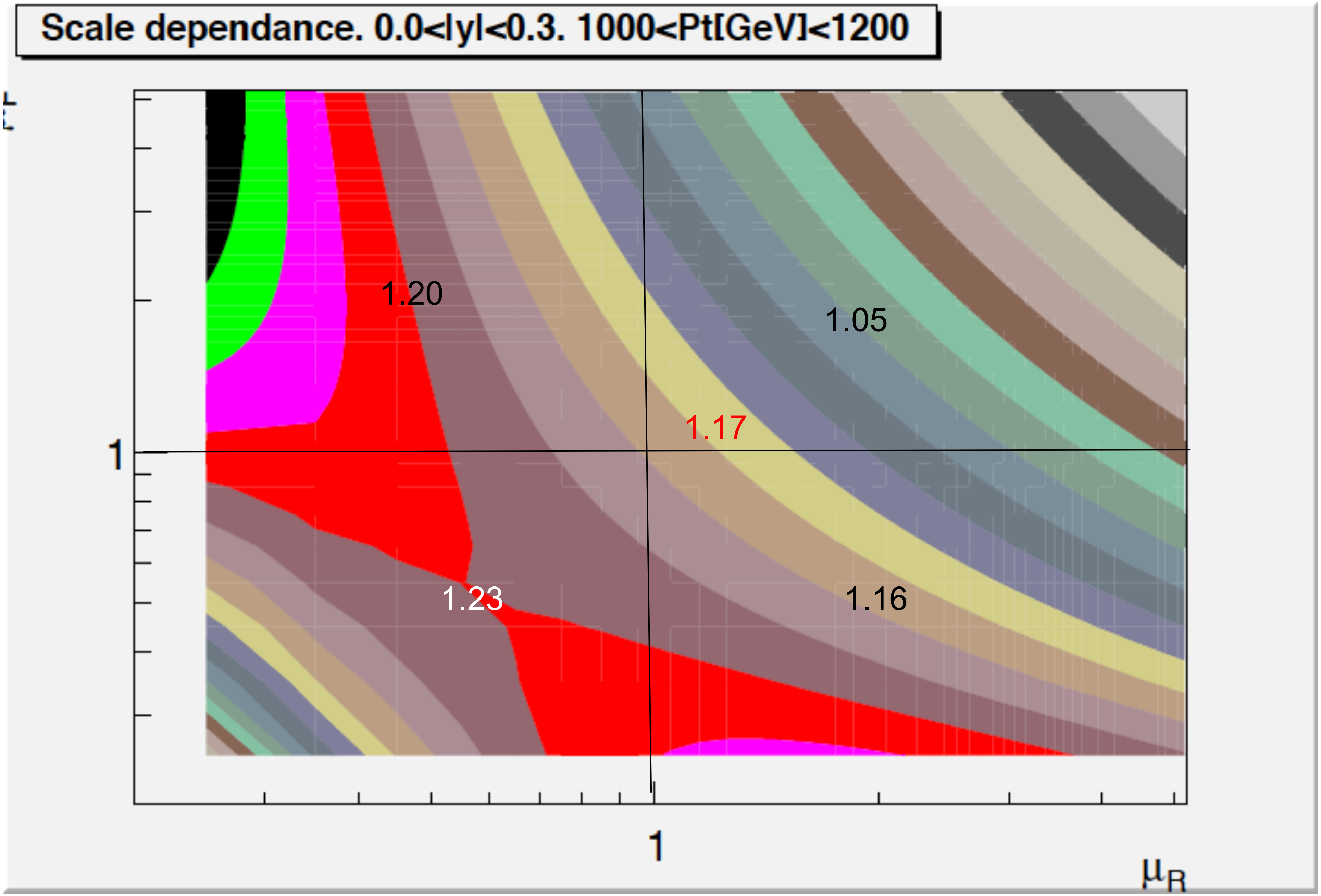


Scale dependance. 1.2<|y|<2.1. 500<Pt[GeV]<600
1.18
1.05
* 1.08
1.14
1.17
1.18
1.13
1
1
μ_R

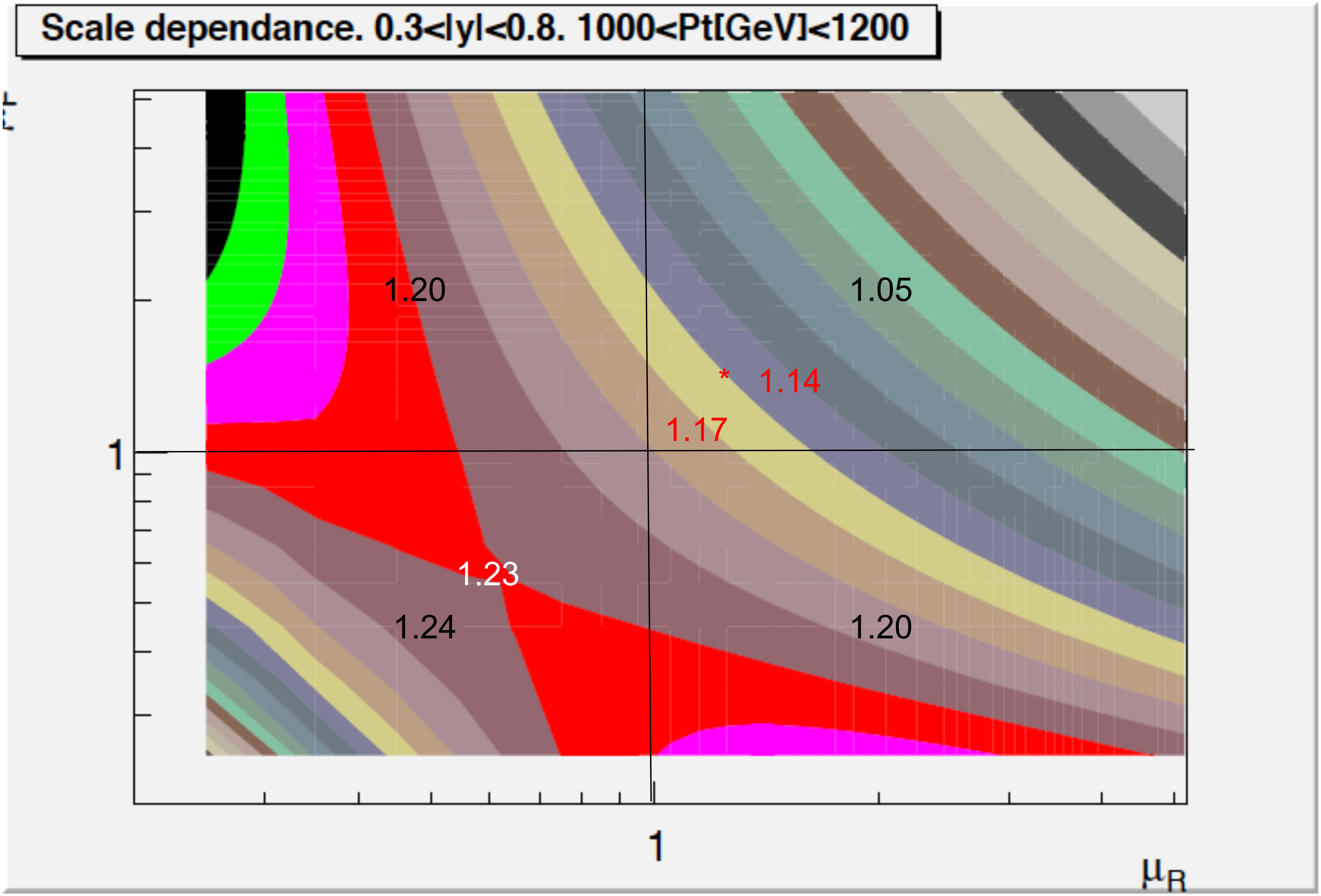


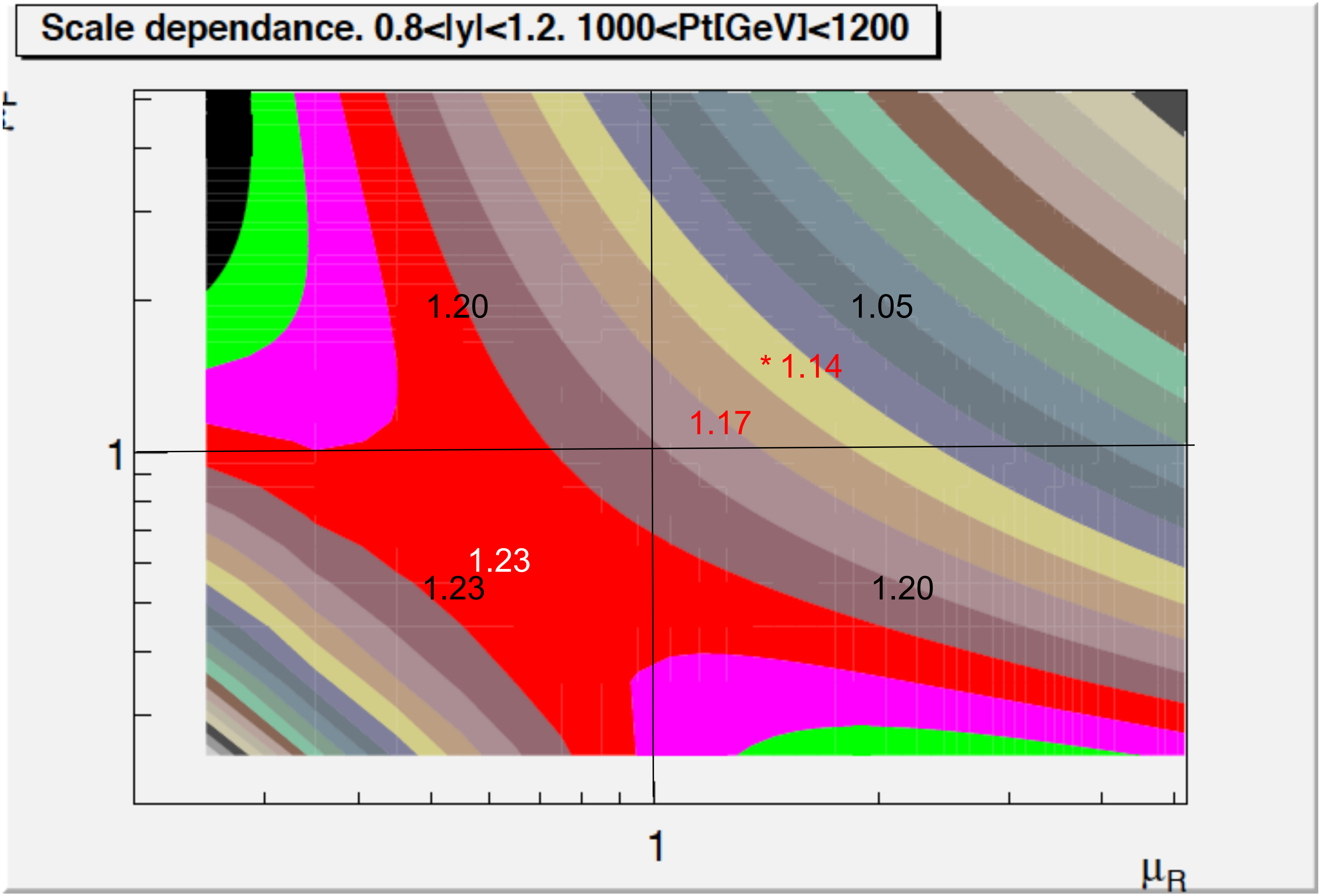
Scale dependance. 2.1<|y|<2.8. 500<Pt[GeV]<600
1.14
* 1.04
1.13
1.08
1.15
1
1
μR


Scale dependance. 2.1<|y|<2.8. 500<Pt[GeV]<600
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1
μ_R
1
μ_F

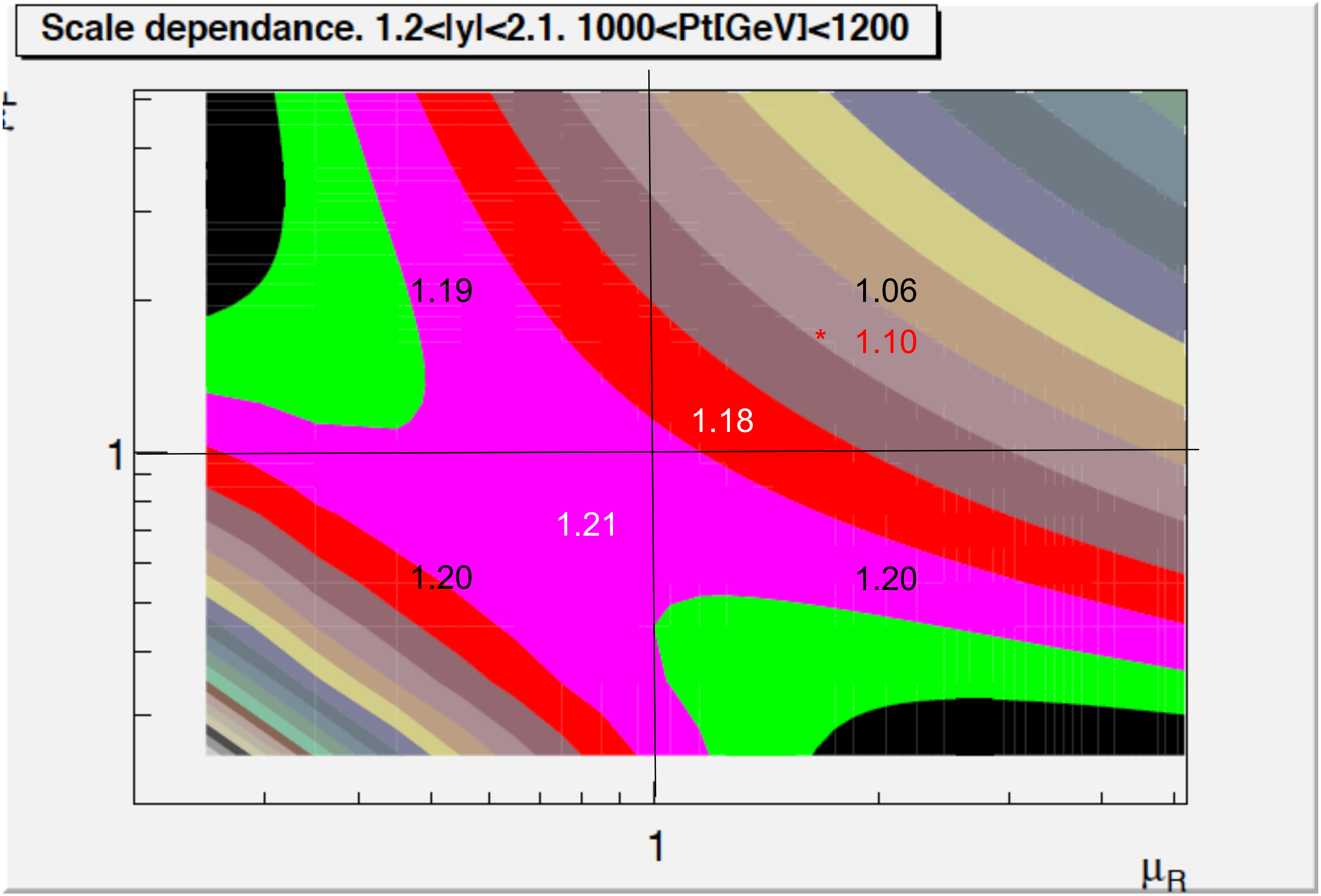


Scale dependance. 0.0<|y|<0.3. 1000<Pt[GeV]<1200
1.20
1.05
1.17
1
1.23
1.16
1
μ_R

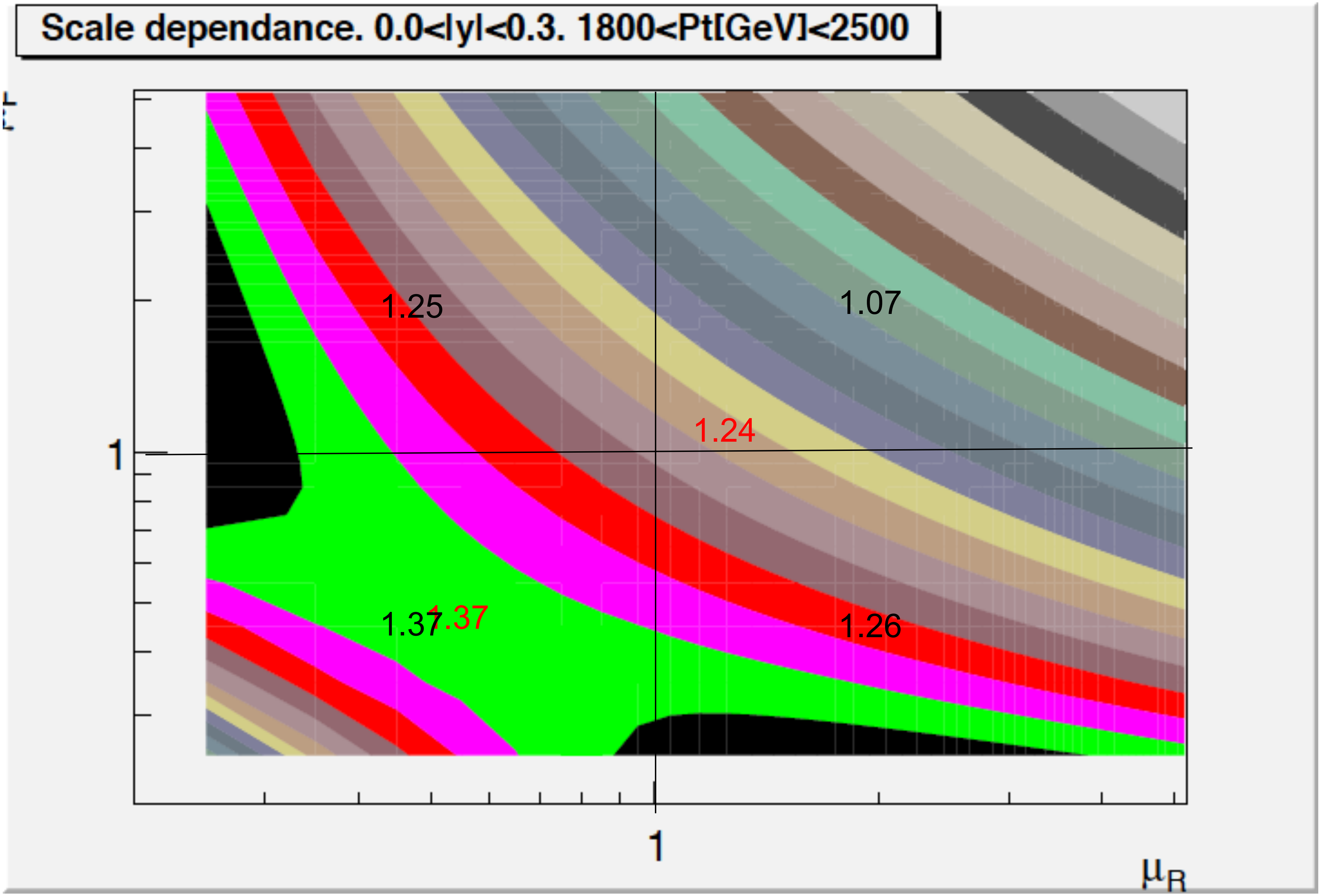

Scale dependance. 0.3<|y|<0.8. 1000<Pt[GeV]<1200
1.20
1.05
* 1.14
1.17
1
1.23
1.24
1.20
1
μ_R



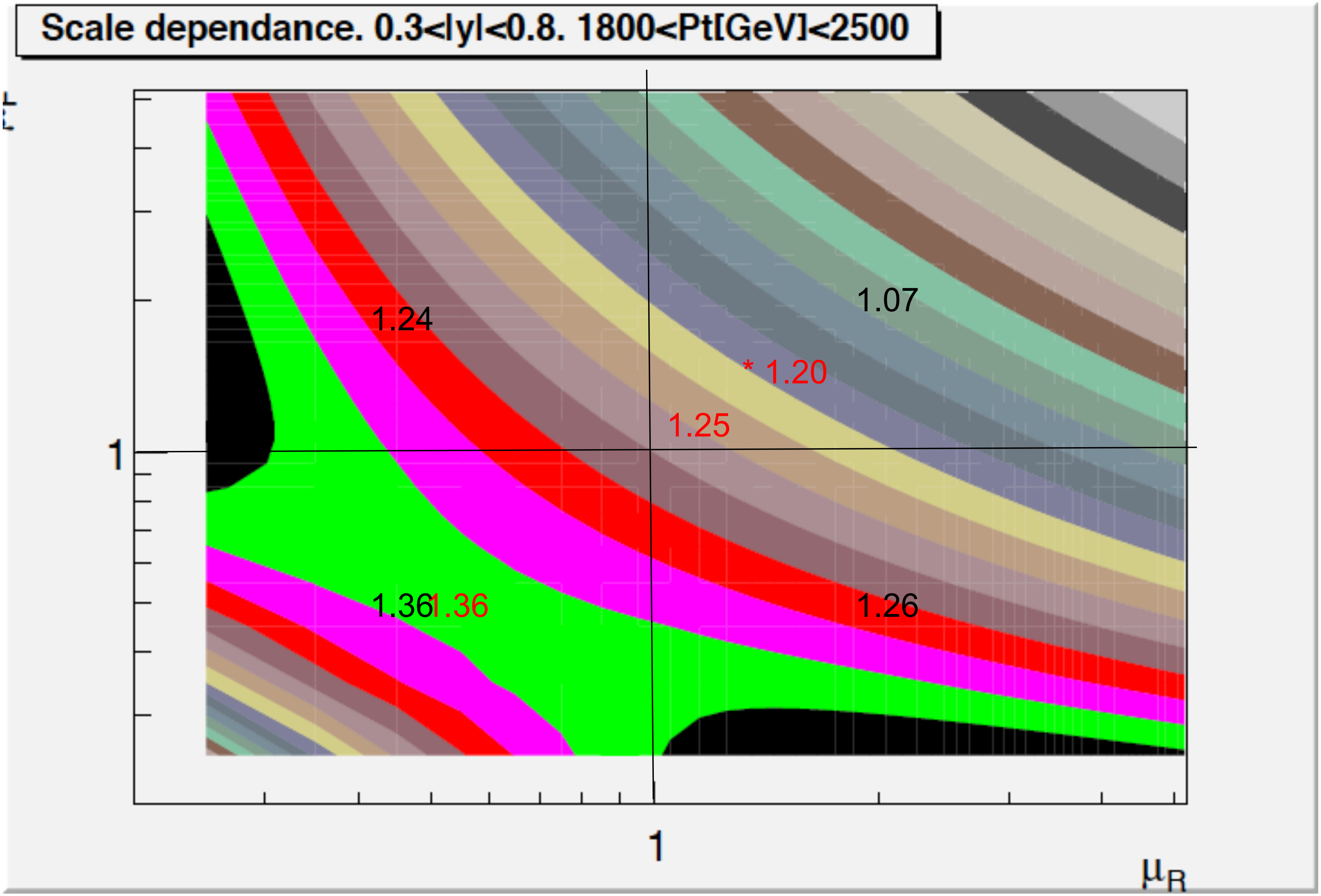
Scale dependance. 0.8<|y|<1.2. 1000<Pt[GeV]<1200
1.20
1.05
* 1.14
1.17
1.23
1.23
1.20
1
1
μR



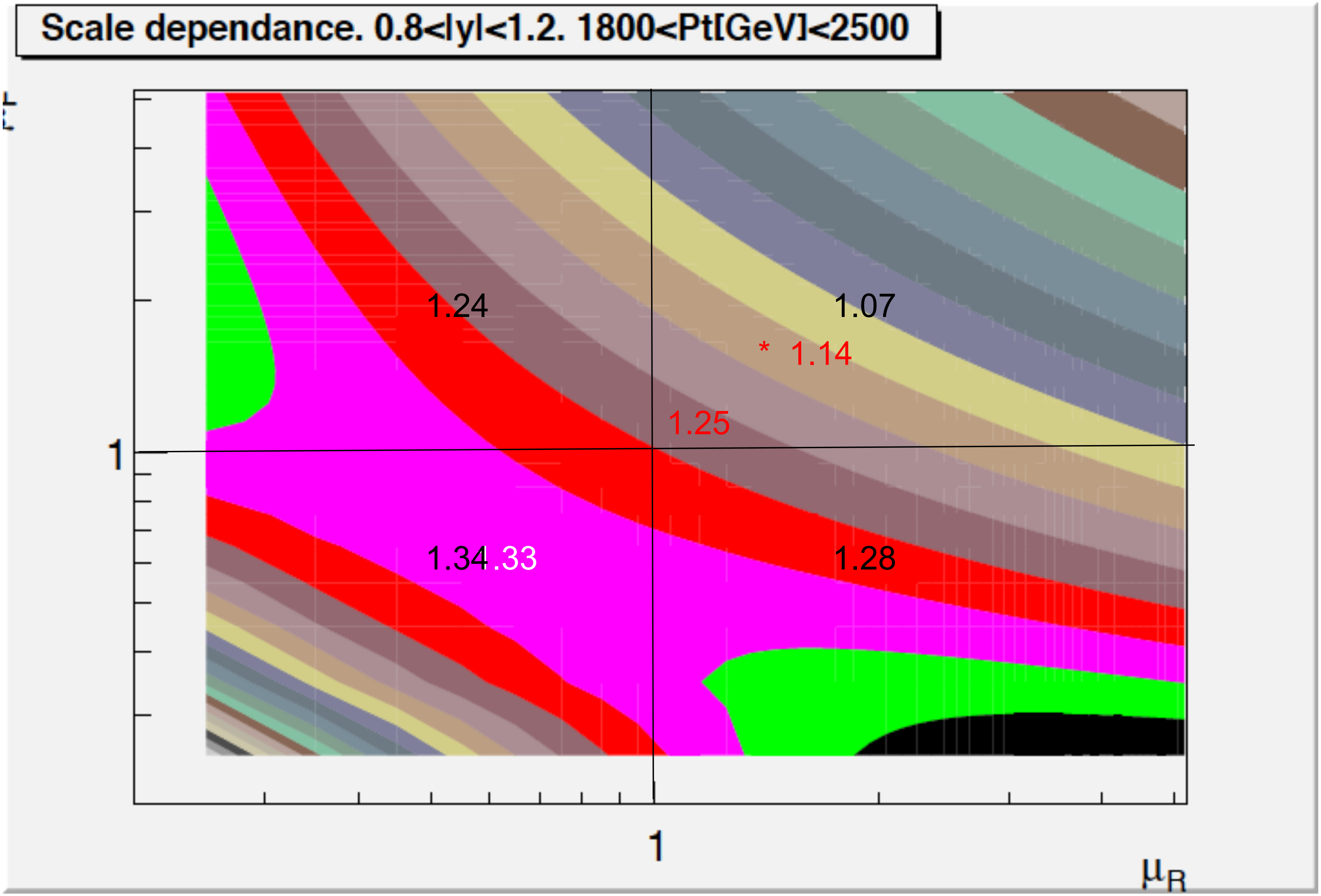
Scale dependance. 1.2<|y|<2.1. 1000<Pt[GeV]<1200
1.19
1.06
* 1.10
1.18
1.21
1.20
1.20
1
1
μ_R

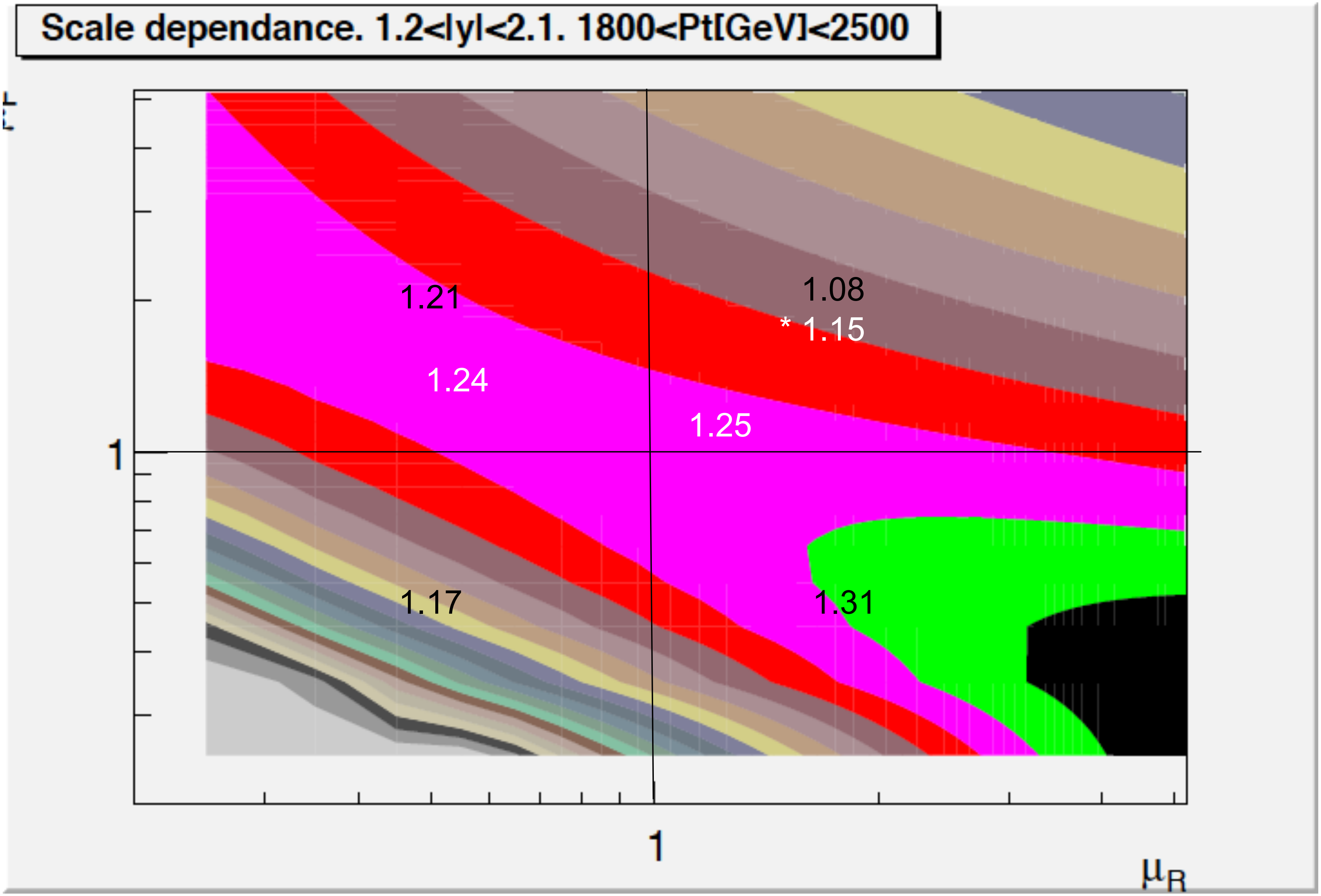


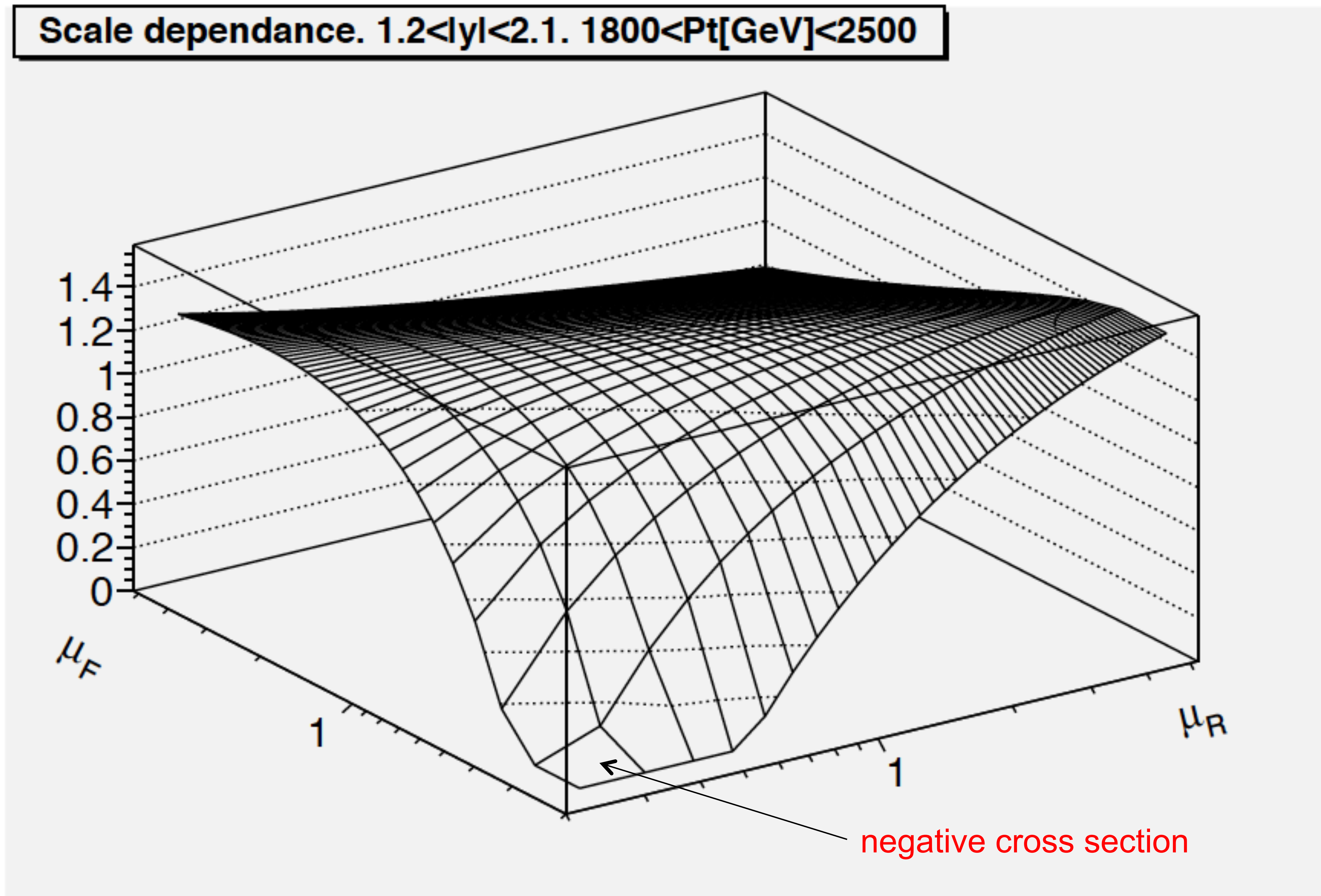
Scale dependance. 0.0<|y|<0.3. 1800<Pt[GeV]<2500
1.25
1.07
1.24
1.37
1.37
1.26
1
1
μR



Scale dependance. 0.3<lyl<0.8. 1800<Pt[GeV]<2500
1.24
1.07
* 1.20
1.25
1.36
1.36
1.26
1
1
μR


Scale dependance. 0.8<lyl<1.2. 1800<Pt[GeV]<2500
1.24
1.07
* 1.14
1.25
1.34
1.33
1.28
1
1
μ_R


Scale dependance. 1.2<|y|<2.1. 1800<Pt[GeV]<2500
1.08
* 1.15
1.21
1.24
1.25
1.17
1.31
1
1
μR

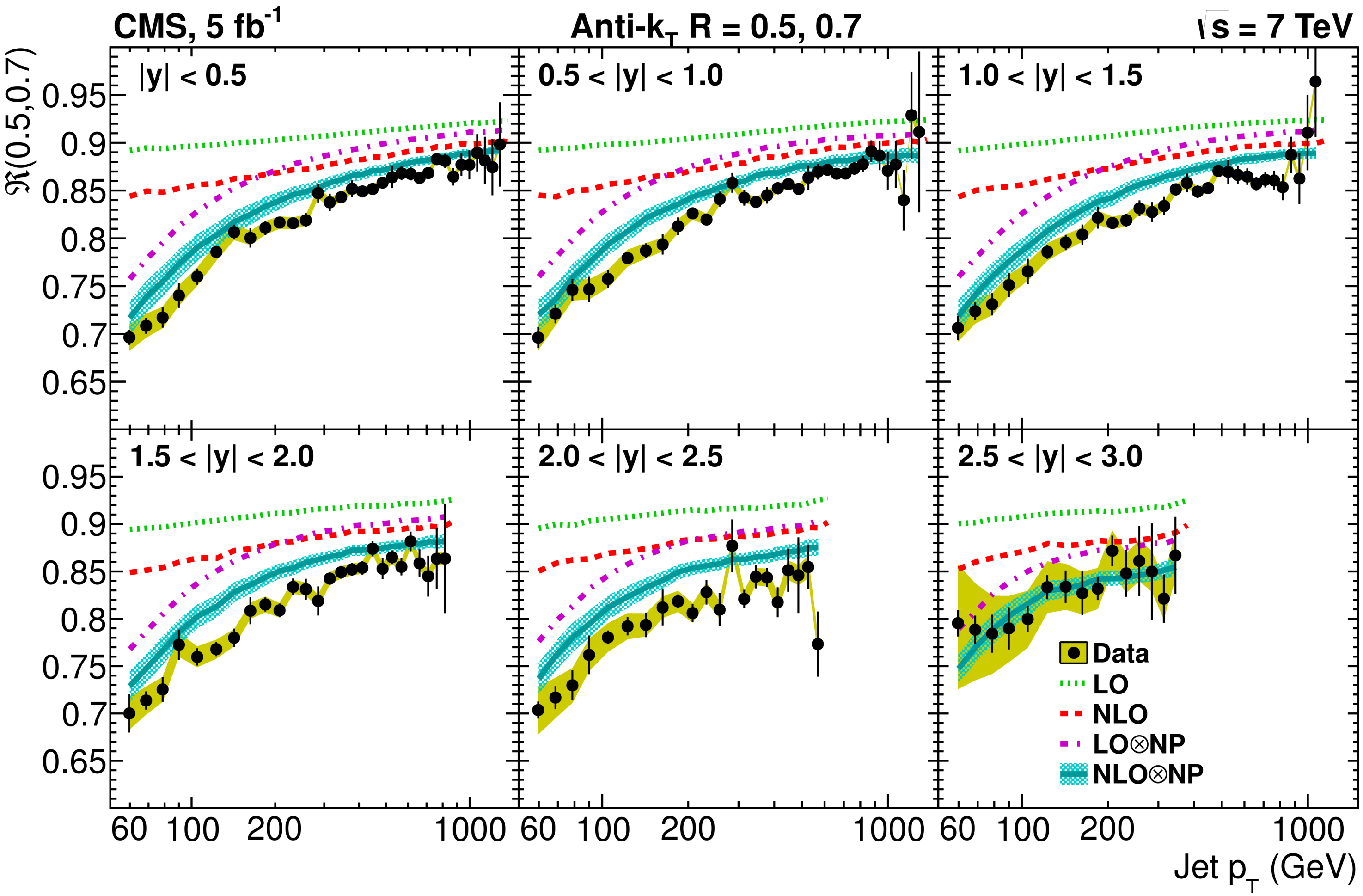


Scale dependance. 1.2<lyl<2.1. 1800<Pt[GeV]<2500
1.4
1.2
1
0.8
0.6
0.4
0.2
0
μ_F
1
μ_R
1
negative cross section

What else (for inclusive jets)?

- Would be nice to reproduce this study at NNLO
  - possible with applgrid storage for NNLO cross sections
  - maybe not jet size variations
- Data, especially at 13 TeV, is getting more precise, as is the theory
- Easy to benchmark theory predictions
- What about the data?
- There are tensions between ATLAS and CMS jet data
- There are tensions within the ATLAS jet data
- Even with very precise data, it may be difficult to reduce PDF uncertainties in global PDF fits, due to these conflicts; this requires a full exploration by PDF groups
- We know that the jet data as a function of rapidity have to be consistent, i.e. PDF's are universal
  - we are working on this; see Bogdan's talk from this morning
  - note that jitter in the data can mask the general constraining power of the data, i.e. the global PDF fit tries to follow the jitter
- There are also known changes in the theory cross sections as the jet size is varied; this is another check that our data analysis and theoretical frameworks are self-consistent

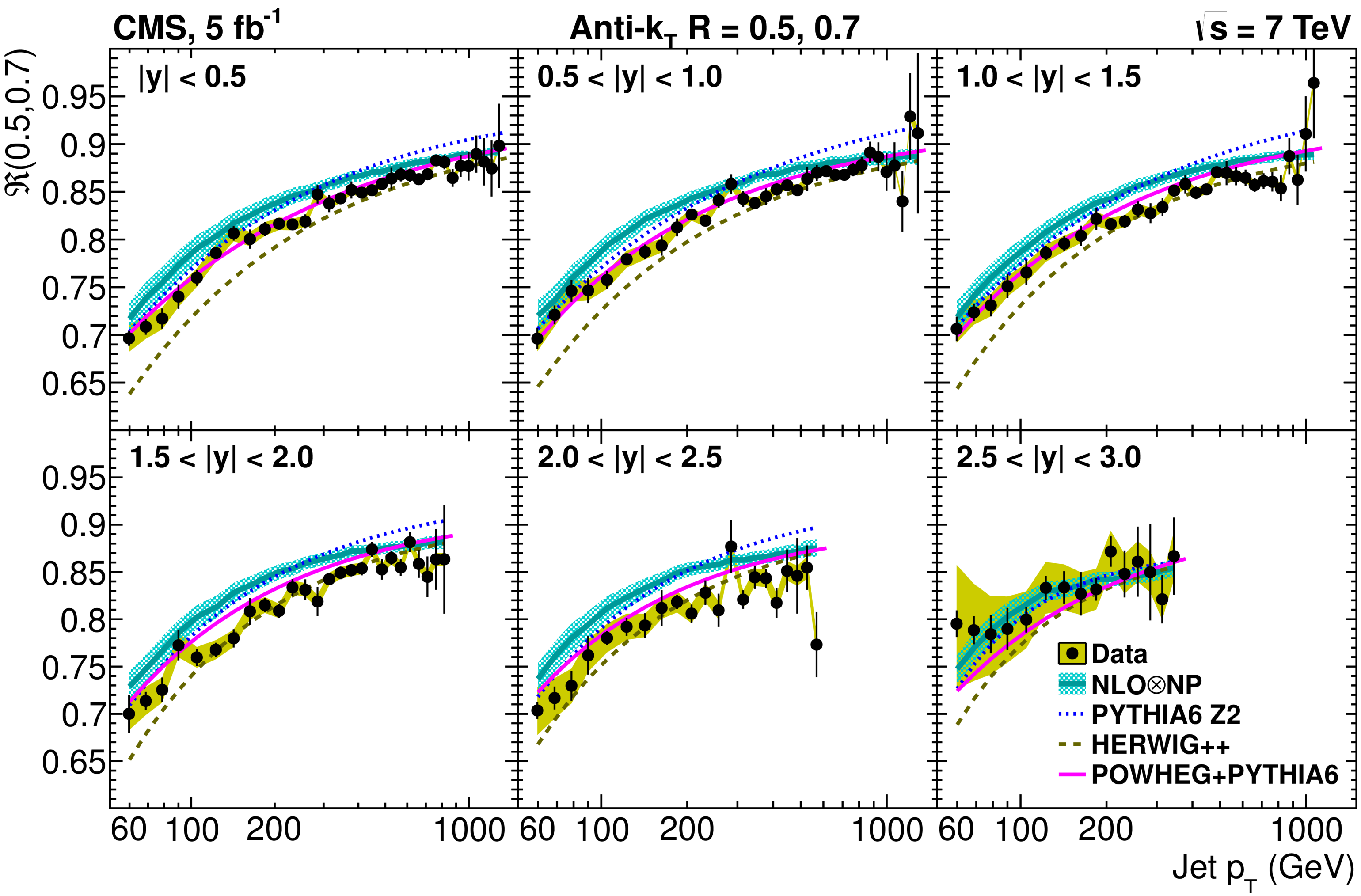
CMS ratio of antikT5 to antikT7:arXiv:1406.0324



NLO+NP works reasonably well, but not perfectly

extend this to NNLO

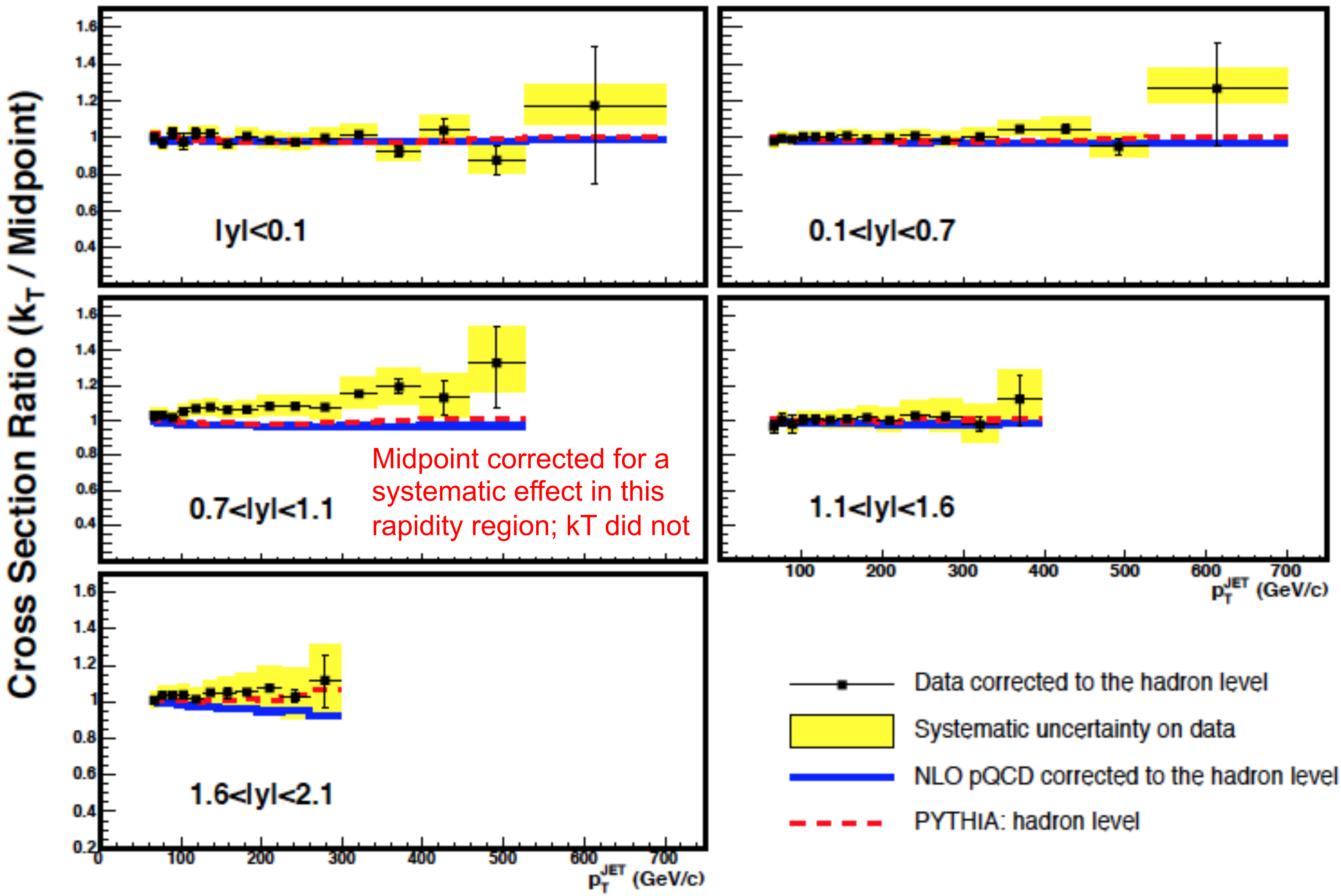
A bit closer with Powheg+Pythia



Note that the implicit assumption in hadronization corrections to fixed order predictions is that NLO=PS as far jet shape is concerned; worked well at Tevatron, especially as R=0.7 in common use; again, expect better agreement with NNLO

Algorithm dependence

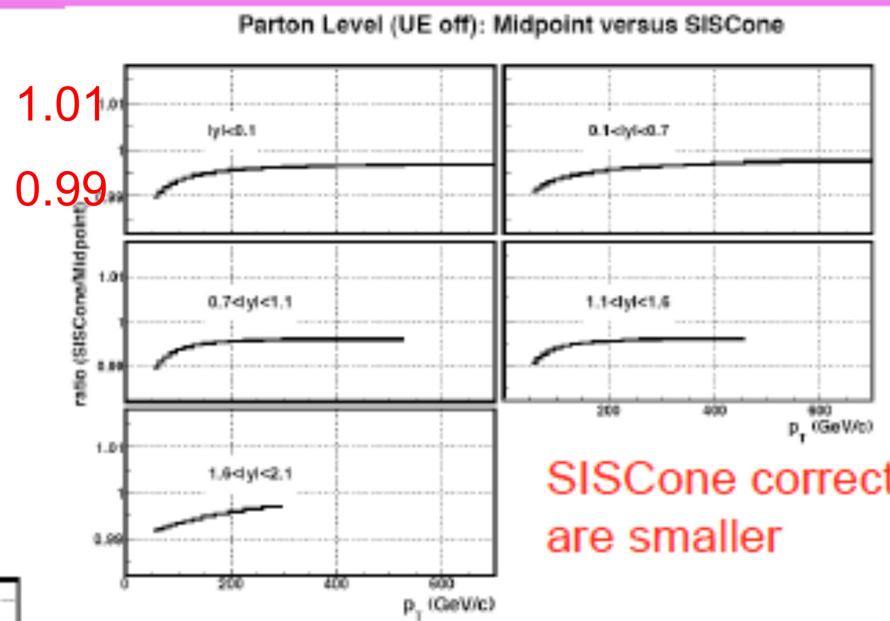
- There are also clear differences in the cross sections predicted for different jet algorithms, that again serves as a check on the measurement and on the theory
- This is almost completely unknown at the LHC as the antikT algorithm is universal
- It happened once at the Tevatron where CDF measured the same cross section using the MidPoint cone algorithm and the kT algorithm, both with R=0.7
- The agreement is exactly as expected as shown on the next slide



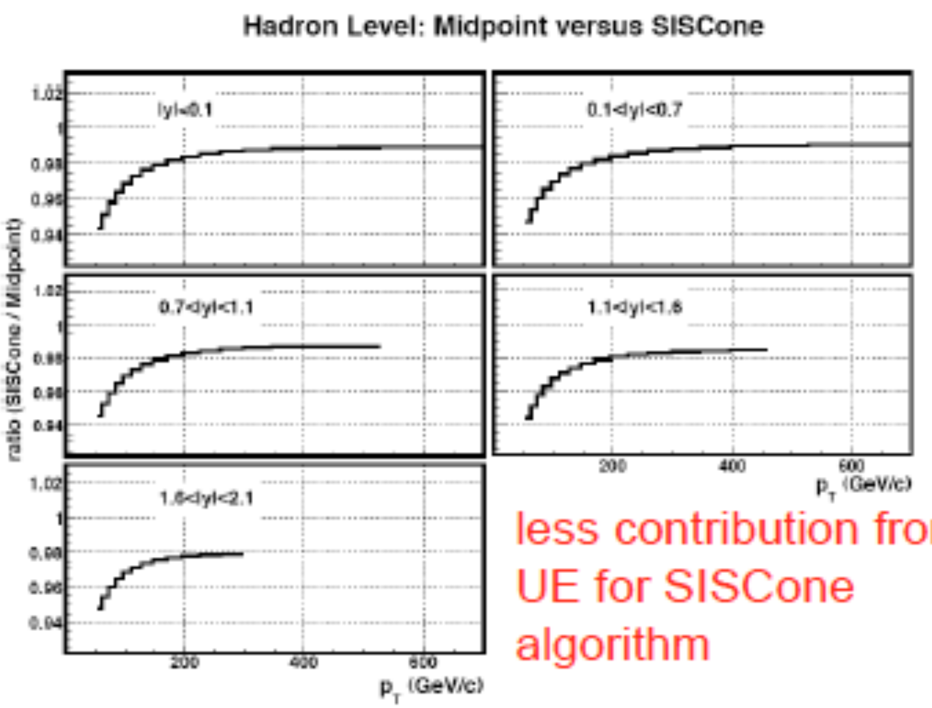
Cross Section Ratio (k_T / Midpoint)
|y|<0.1
0.1<|y|<0.7
0.7<|y|<1.1
1.1<|y|<1.6
1.6<|y|<2.1
Midpoint corrected for a
systematic effect in this
rapidity region; kT did not
p_T^{JET} (GeV/c)
Data corrected to the hadron level
Systematic uncertainty on data
NLO pQCD corrected to the hadron level
PYTHIA: hadron level

SISCone vs Midpoint

- The SISCone jet algorithm developed by Salam et al is preferred from a theoretical basis, as there is less IR sensitivity from not requiring any seeds as the starting point of a jet



SISCone corrections are smaller



less contribution from UE for SISCone algorithm

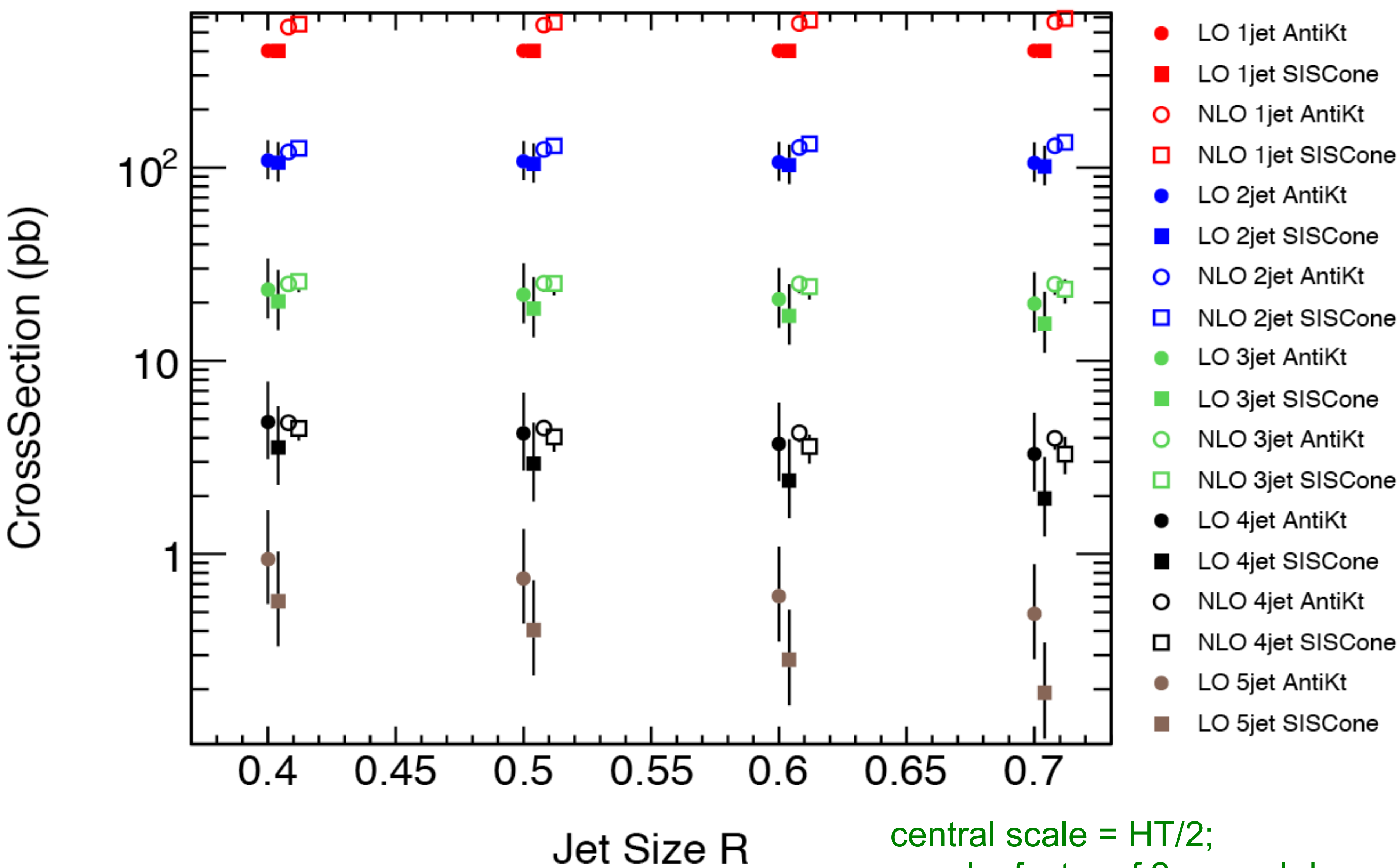
- So far, at the Tevatron, we have not explicitly measured a jet cross section using the SISCone algorithm, although studies are underway, but we have done some Monte Carlo comparisons for the inclusive cros sections
- Differences of the order of a few percent at the hadron level reduce to <1% at the parton level

W+jets

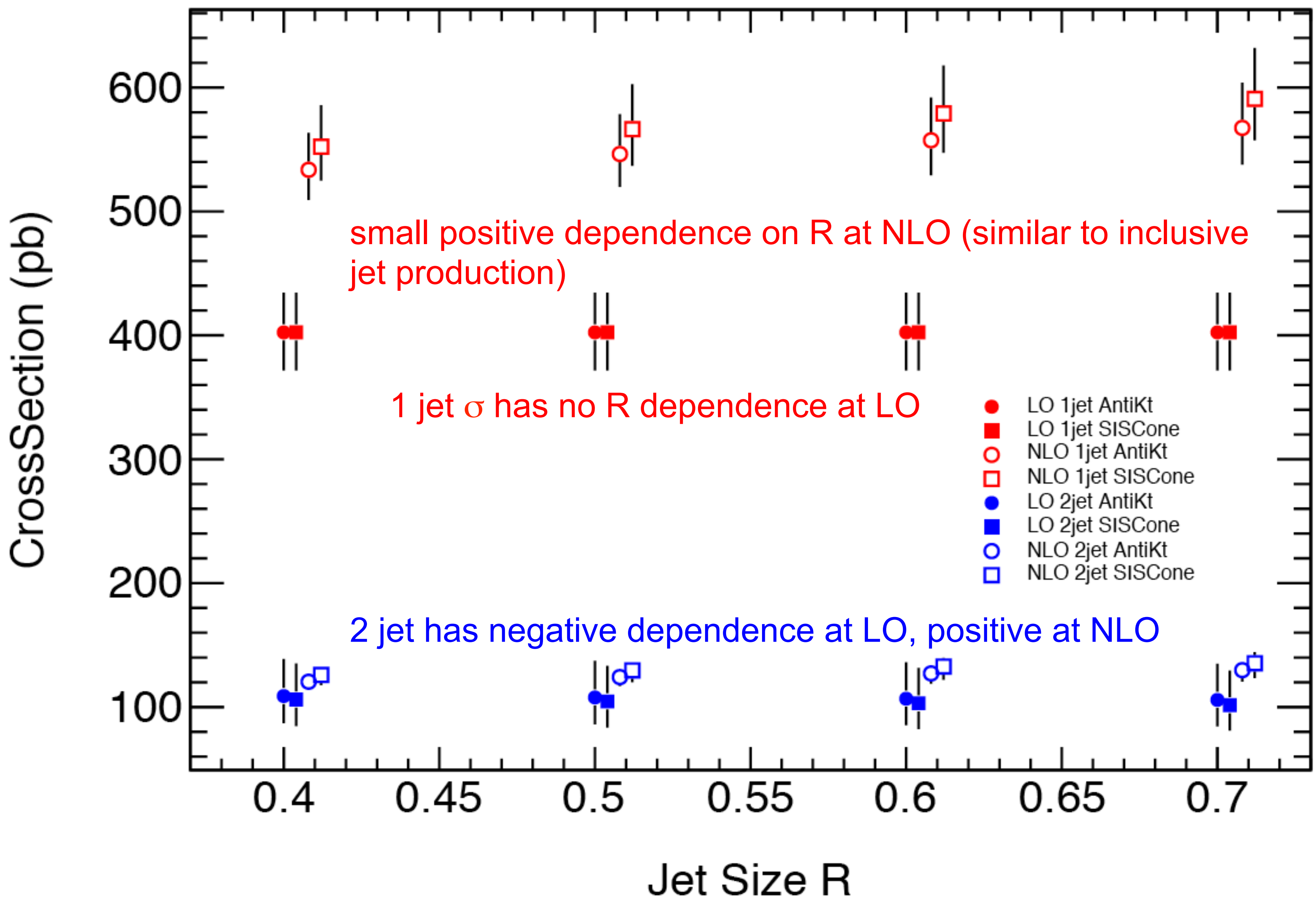
- Now turn to W/Z+jets (at 7 TeV)
- Use Blackhat+Sherpa ROOT ntuples to vary scales/jet sizes/jet algorithms
  - arxiv:1310.7439
- So far, only 1D plots, i.e. $\mu_R = \mu_F$

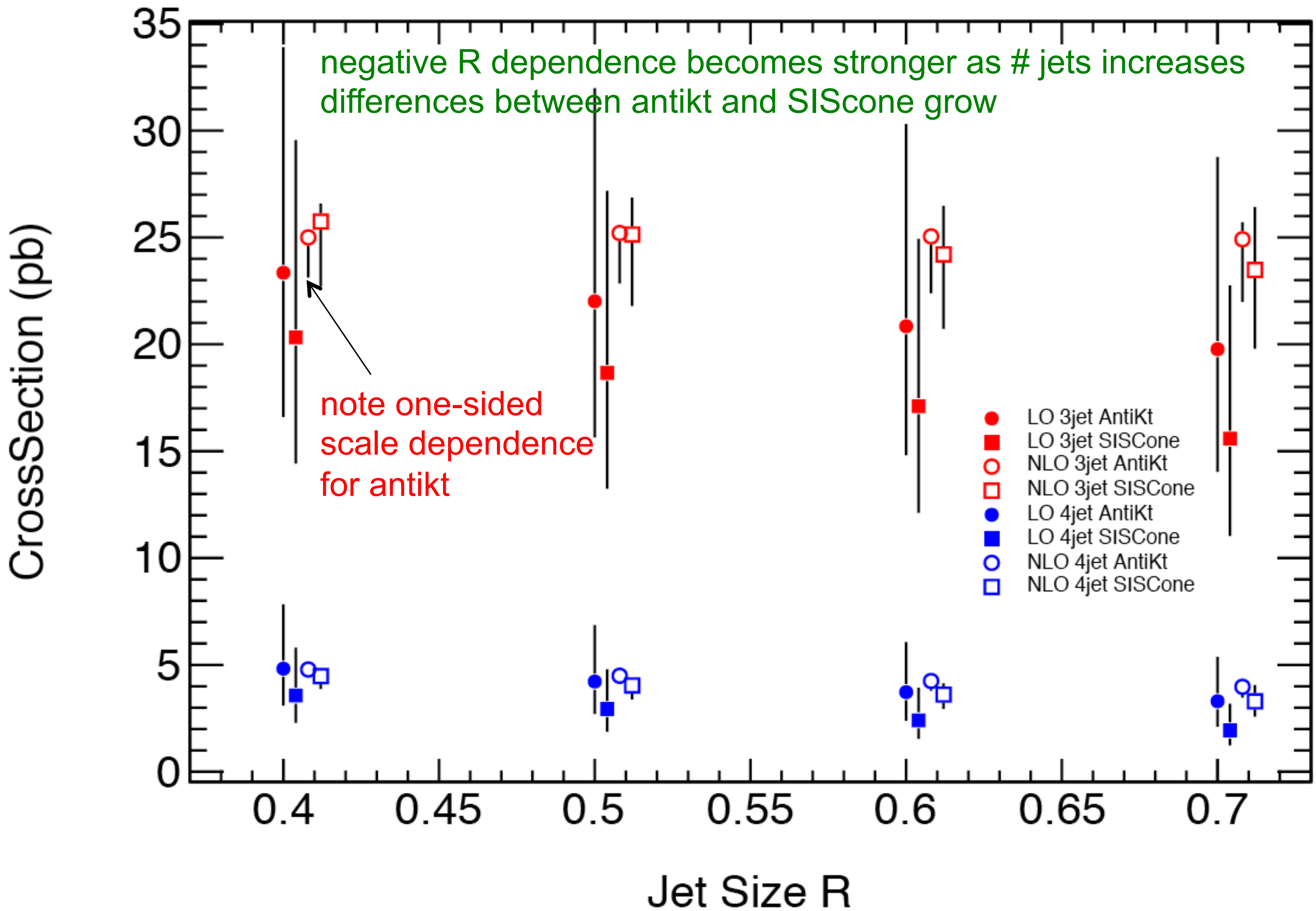
branch name	type	Notes
id	I	id of the event. Real events and their associated counterterms share the same id. This allows for the correct treatment of statistical errors.
nparticle	I	number of particles in the final state
px	F[nparticle]	array of the x components of the final state particles
py	F[nparticle]	array of the y components of the final state particles
pz	F[nparticle]	array of the z components of the final state particles
E	F[nparticle]	array of the energy components of the final state particles
alphas	D	$alpha_s$ value used for this event
kf	I	PDG codes of the final state particles
weight	D	weight of the event
weight2	D	weight of the event to be used to treat the statistical errors correctly in the real part
me_wgt	D	matrix element weight, the same as weight but without pdf factors
me_wgt2	D	matrix element weight, the same as weight2 but without pdf factors
x1	D	fraction of the hadron momentum carried by the first incoming parton
x2	D	fraction of the hadron momentum carried by the second incoming parton
x1p	D	second momentum fraction used in the integrated real part
x2p	D	second momentum fraction used in the integrated real part
id1	I	PDG code of the firt incoming parton
id2	I	PDG code of the second incoming parton
fac_scale	D	factorization scale used
ren_scale	D	renormalization scale used
nuwgt	I	number of additional weights
usr_wgts	D[nuwgt]	additional weights needed to change the scale

Look at jet size, algorithm dependences; scale uncertainty



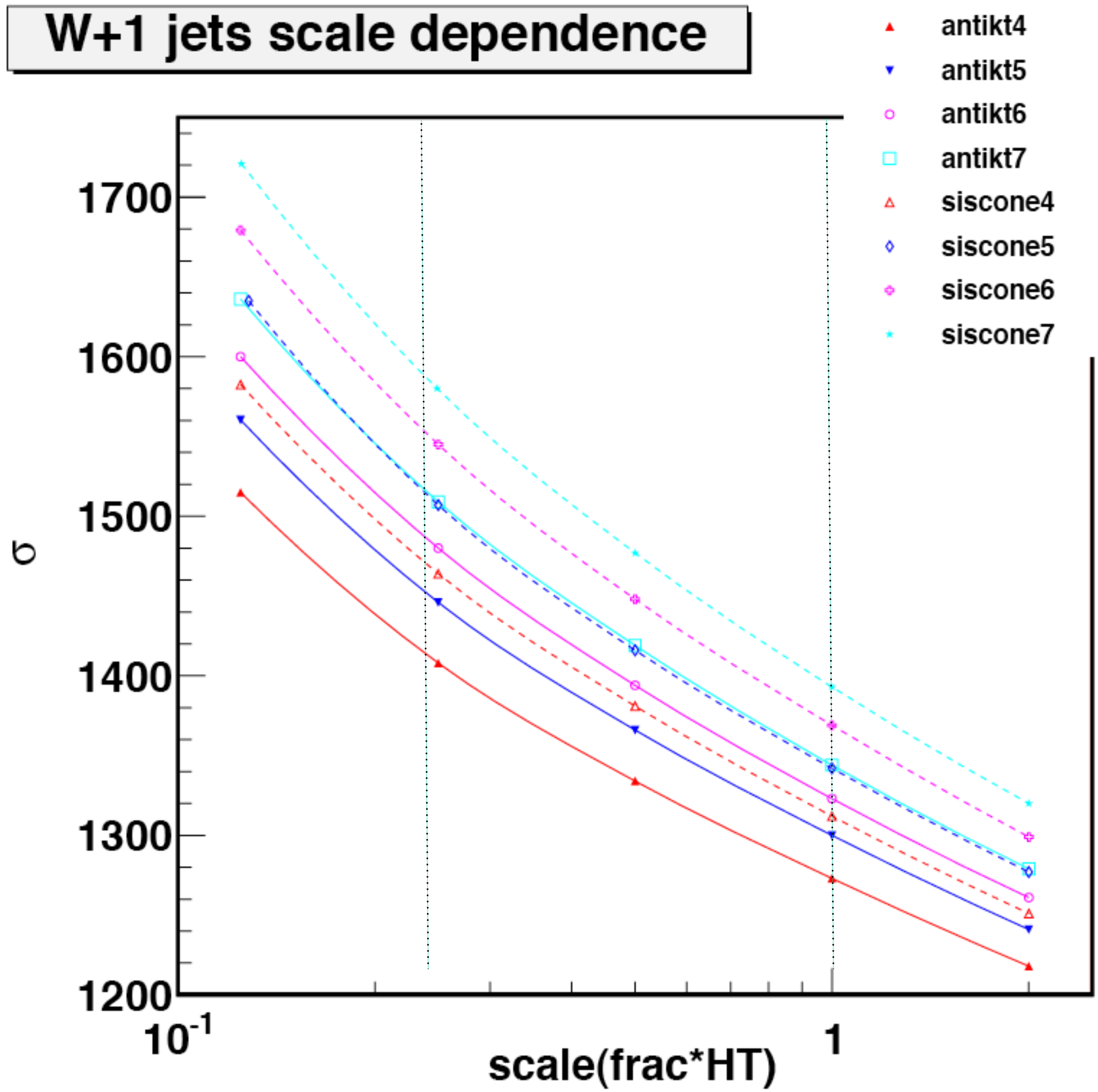
central scale = HT/2;
vary by factor of 2 up and down


small positive dependence on R at NLO (similar to inclusive jet production)
1 jet σ has no R dependence at LO
2 jet has negative dependence at LO, positive at NLO
LO 1jet AntiKt
LO 1jet SISCone
NLO 1jet AntiKt
NLO 1jet SISCone
LO 2jet AntiKt
LO 2jet SISCone
NLO 2jet AntiKt
NLO 2jet SISCone
CrossSection (pb)
Jet Size R
600
500
400
300
200
100
0.4
0.45
0.5
0.55
0.6
0.65
0.7



negative R dependence becomes stronger as # jets increases
differences between antikt and SIScone grow
note one-sided
scale dependence
for antikt
LO 3jet AntiKt
LO 3jet SISCone
NLO 3jet AntiKt
NLO 3jet SISCone
LO 4jet AntiKt
LO 4jet SISCone
NLO 4jet AntiKt
NLO 4jet SISCone
CrossSection (pb)
Jet Size R
0
5
10
15
20
25
30
35
0.4
0.45
0.5
0.55
0.6
0.65
0.7

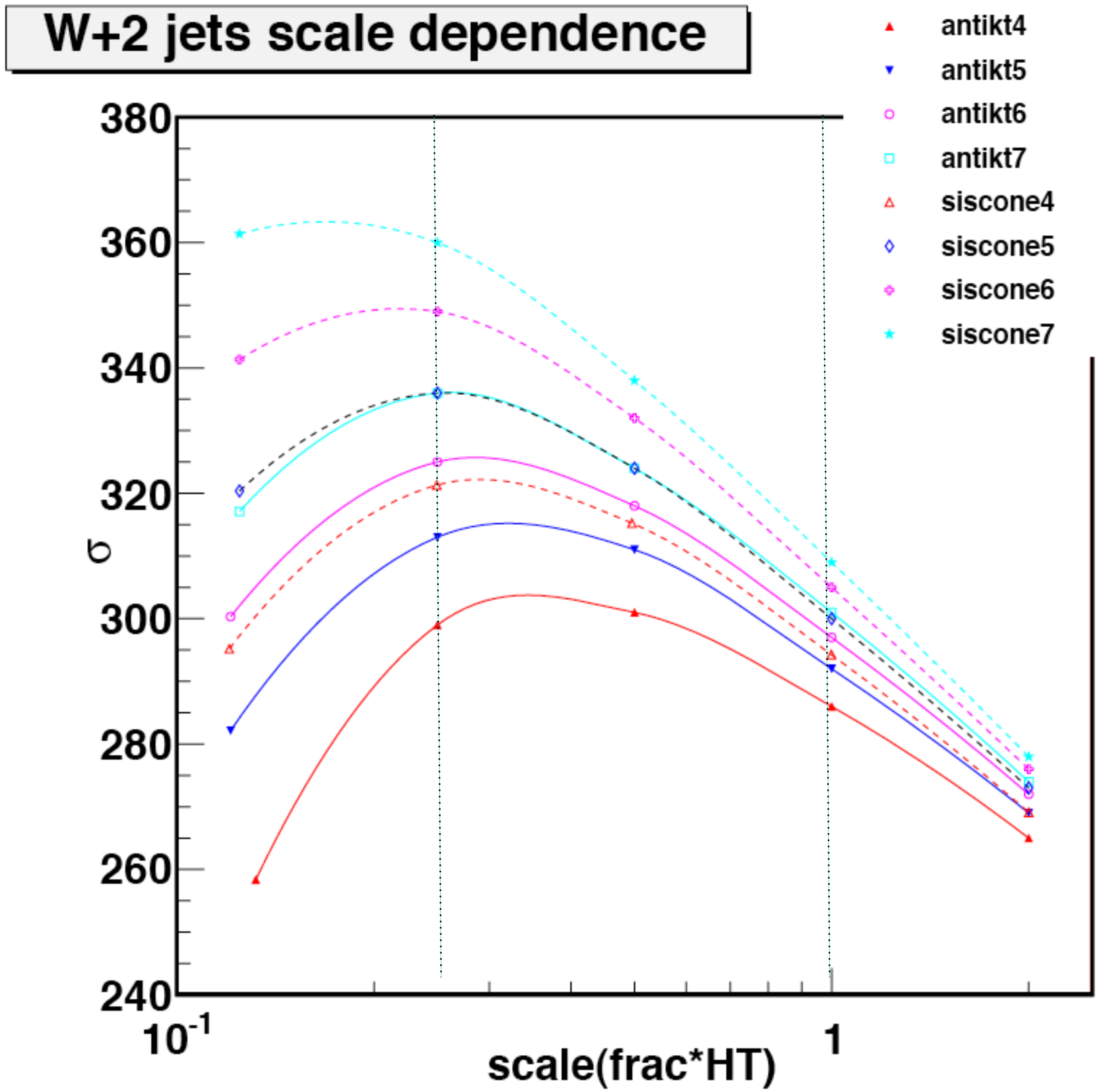
note monotonic scale dependence at NLO, similar to what is seen in a typical LO calculation



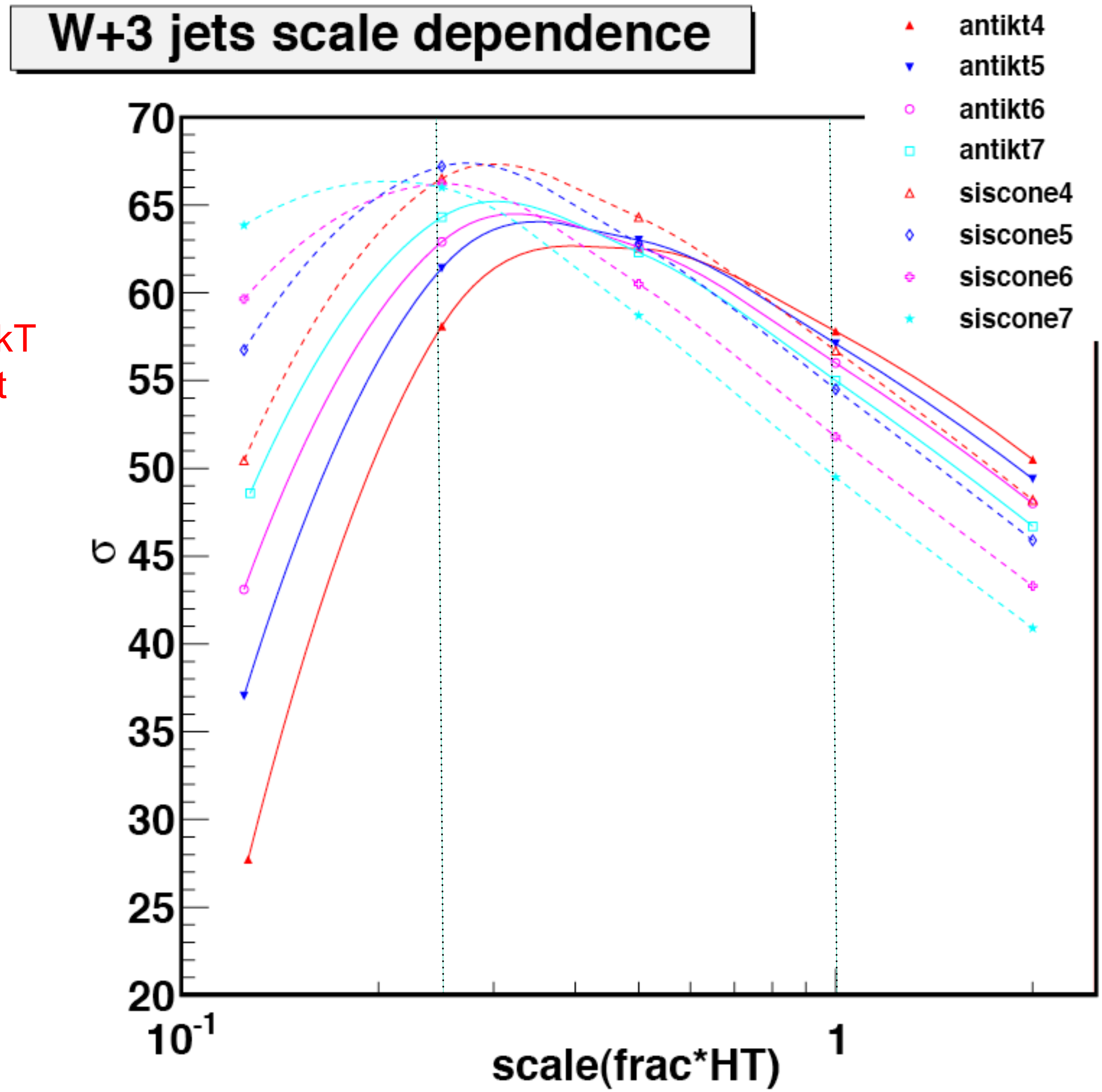
See parabolic shape for W+2 jets

Note that for antikT4, the scale $H_T/2$ is at the same point as $H_T/4$; scale dependence will appear smaller

may be better to look for max/min over scale range



Note again that SISCone and antikT algorithms peak at different scales



arXiv:1004.1659: Blackhat+Sherpa (Z+3 jets)

A. Total cross sections

Tevatron

# of jets	CDF midpoint	LO parton SISCone	NLO parton SISCone	LO parton anti-k_T	NLO parton anti-k_T
1	$7003 \pm 146^{+483}_{-470} \pm 406$	$4635(2)^{+928}_{-715}$	$6080(12)^{+354}_{-402}$	$4635(2)^{+928}_{-715}$	$5783(12)^{+257}_{-334}$
2	$695 \pm 37^{+59}_{-60} \pm 40$	$429.8(0.3)^{+171.7}_{-111.4}$	$564(2)^{+59}_{-70}$	$481.2(0.4)^{+191}_{-124}$	$567(2)^{+31}_{-57}$
3	$60 \pm 11^{+8}_{-8} \pm 3.5$	$24.6(0.03)^{+14.5}_{-8.2}$	$36.8(0.2)^{+8.8}_{-7.8}$	[illegible]7.88(0.04)[illegible]	$44.7(0.24)^{+5.1}_{-6.8}$

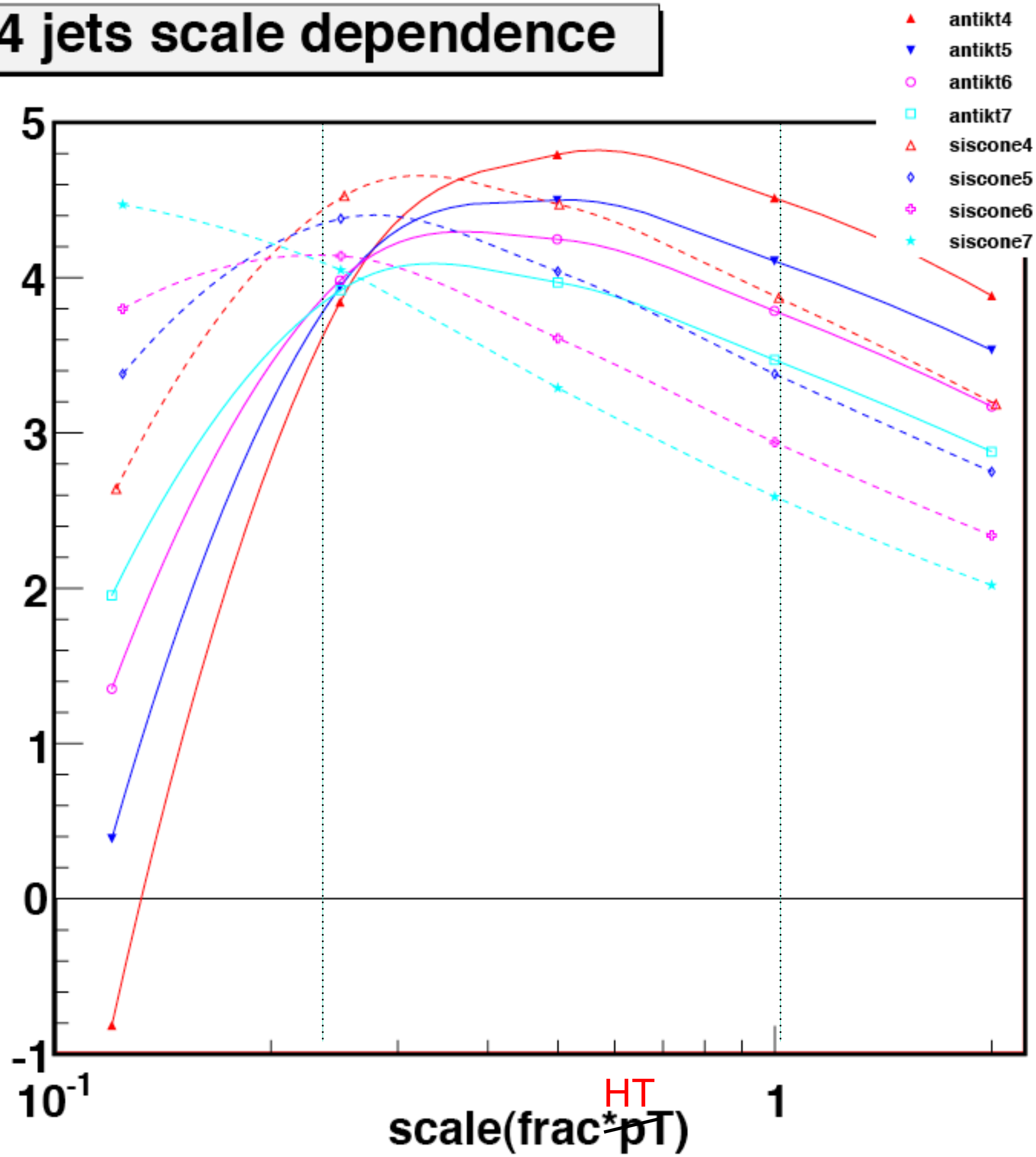
TABLE II: Z, γ^* + 1, 2, 3-jet production (inclusive) cross section (in fb) at CDF. The column labeled CDF gives the hadron-level results from ref. [2], using a midpoint jet algorithm. The experimental uncertainties are statistical, systematics (upper and lower) and luminosity. The columns labeled by LO parton and NLO parton contain the parton-level results for the SISCone and anti-k_T jet algorithms. The central scale choice for the theoretical prediction is $\mu = \hat{H}_T/2$, the numerical integration uncertainty is in parentheses, and the scale dependence is quoted in super- and subscripts. Non-perturbative corrections should be accounted for prior to comparing the CDF measurement to parton-level NLO theory.

Note that it appears that antikT cross section is much larger than similar SISCone cross section…and uncertainty is much smaller

SISCone cross section peaks at scale smaller than $H_T/2$; peak cross sections are similar; uncertainties about peak are similar

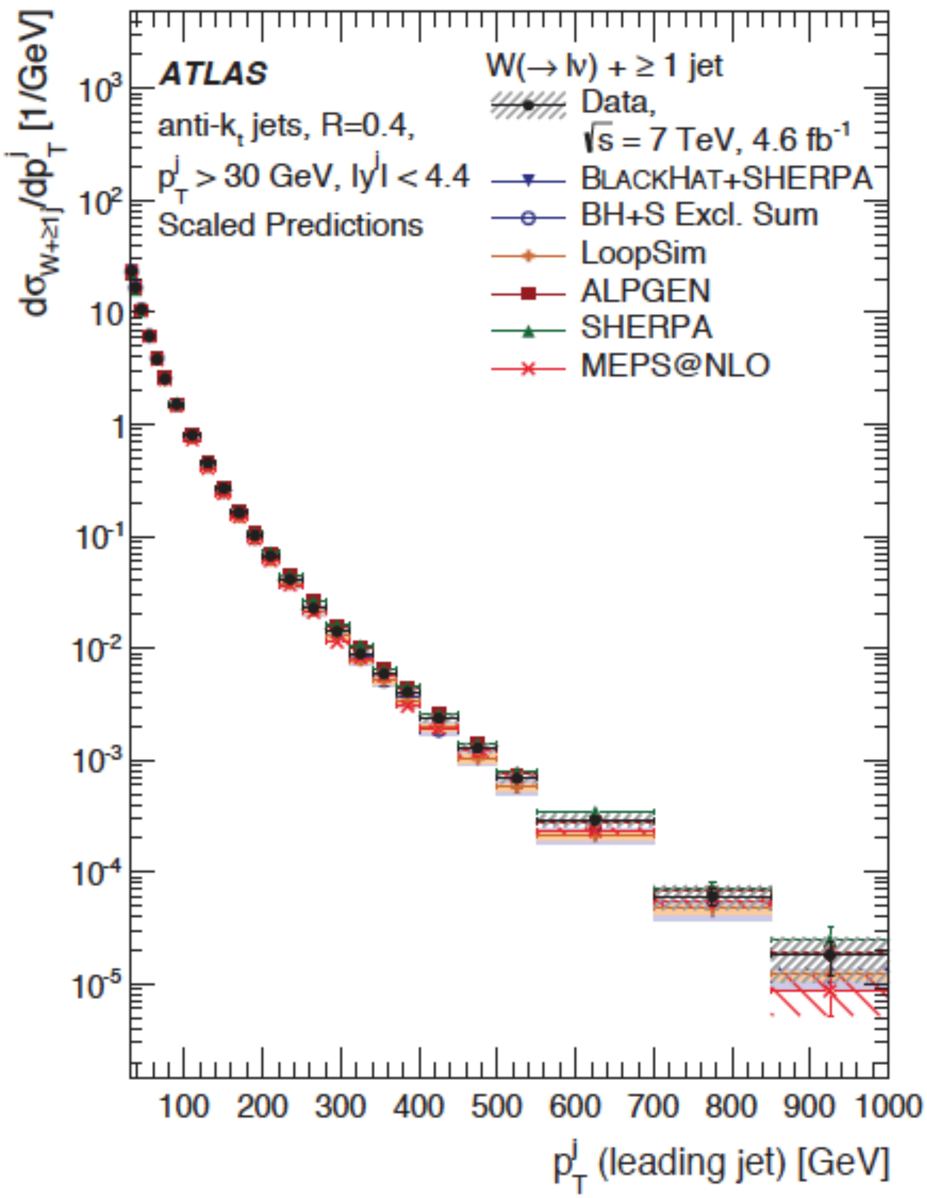
W+4 jets scale dependence

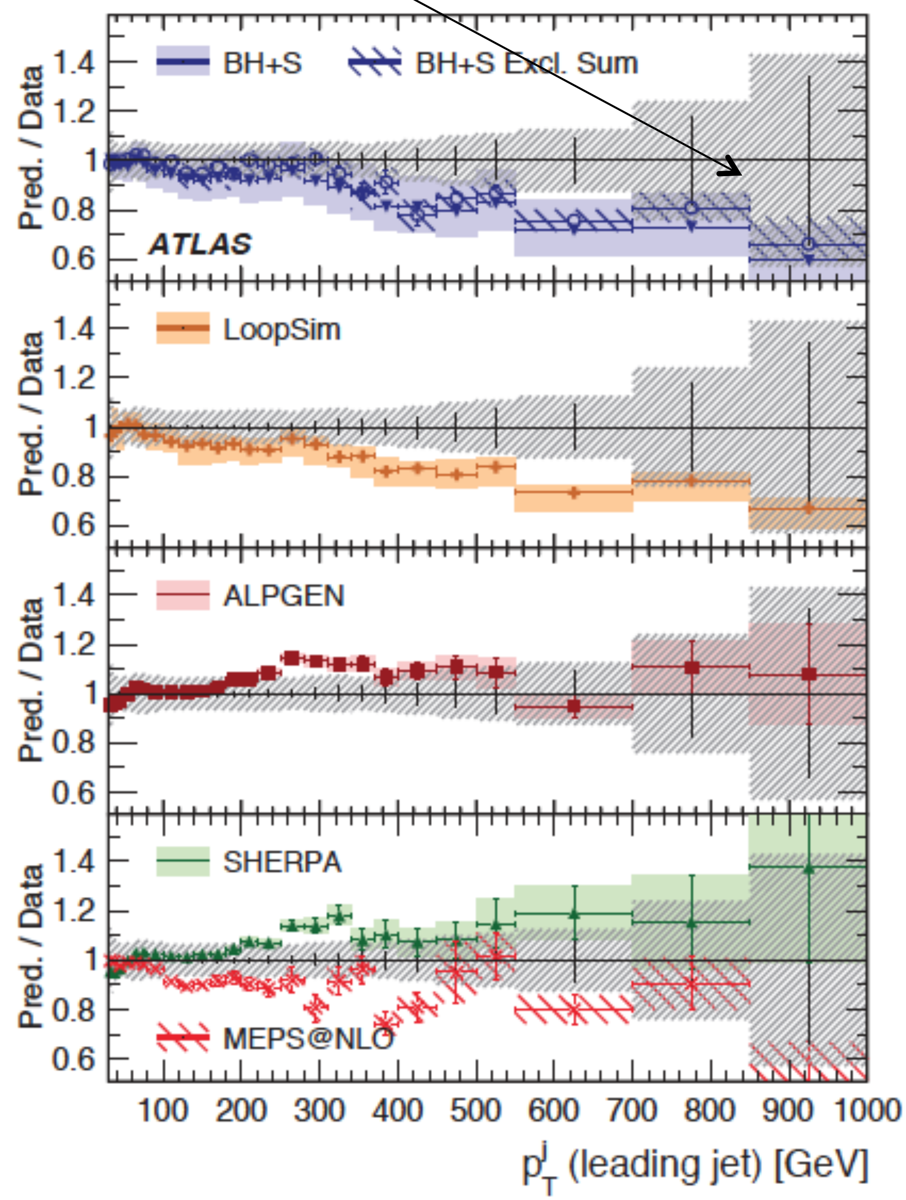
- A scale of HT/2 is ~ the peak for antikt4; so all deviations are negative
- Siscone peaks around HT/3
- Moves to smaller scales for larger R
- @HT/4, all antikt R give same result; that scale seems to be around HT/5 for siscone
- it is difficult to make conclusions about the uncertainty of any particular W + n jet cross section without understanding the scale dependence as the jet size/algorithm is varied

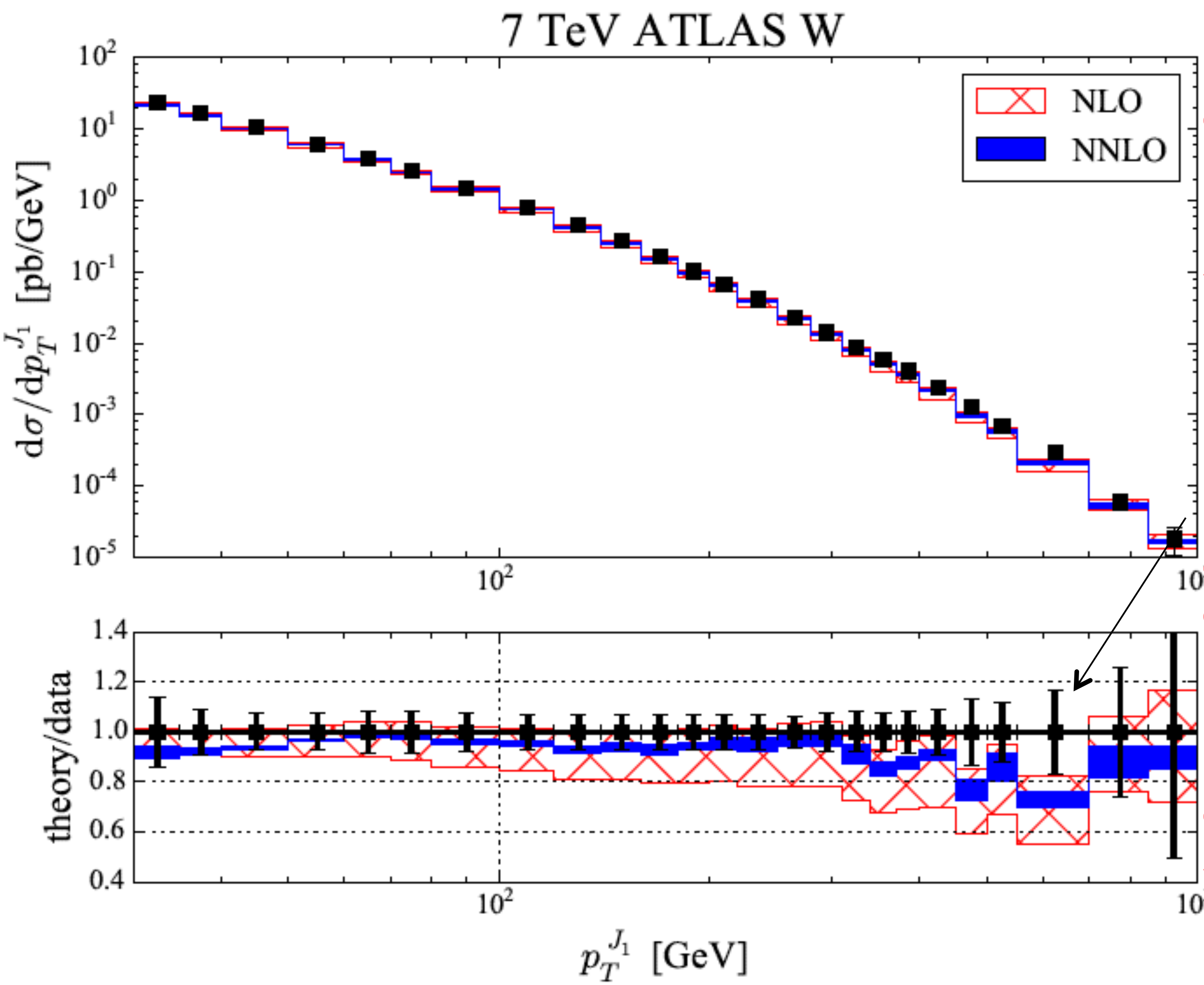


Now look at the data;
unfortunately only at R=0.4
In ATLAS (7 TeV)

Scale HT/2: NLO below data at
1 TeV







but NNLO (and NLO) describe the high p_T data fairly well, using a different scale

$$\mu_0 = \sqrt{M_{l\nu}^2 + \sum_i (p_T^{J_i})^2}$$

but the two scales don't seem that far off (numerically)

how can NLO predictions look so different for similar scales

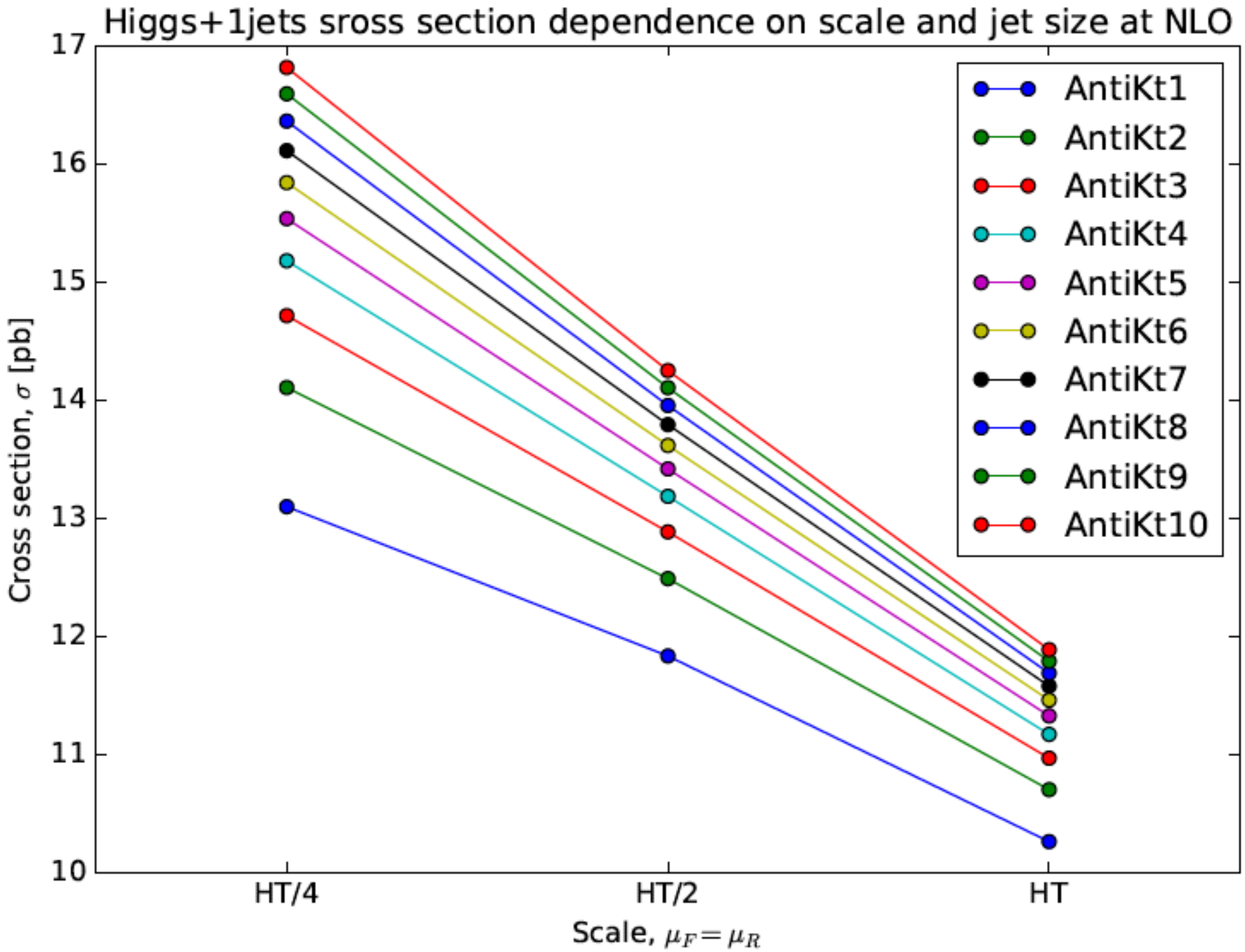
Higgs+jets at NLO (13 TeV)

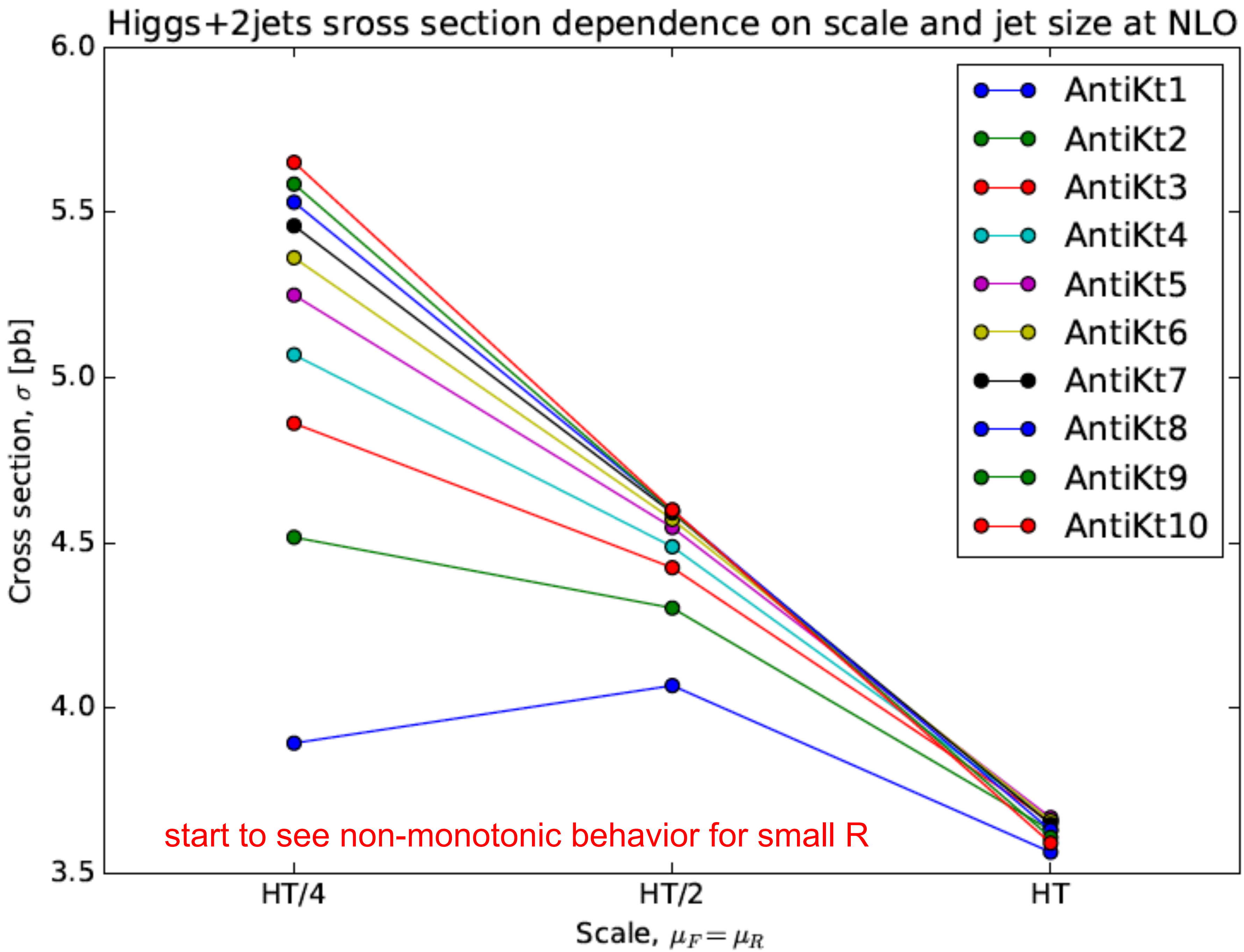
- Using gosam ntuples (same structure as B+S); again only 1-D for the moment

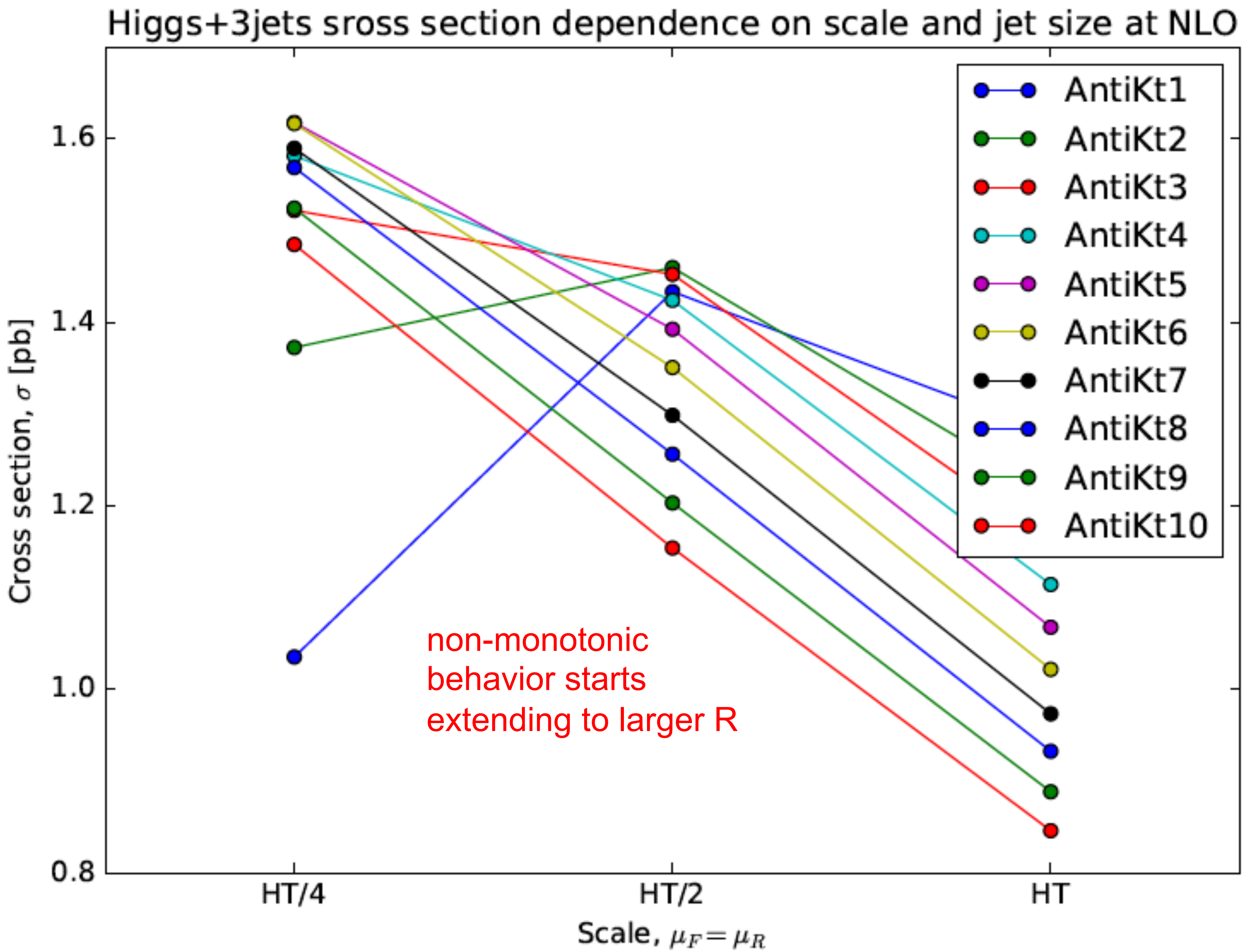
montonic behavior, with scale and with jet size

…ignoring any lnR terms that may have to be resummed

small R dependence from 0.4 to 0.7, similar to W+jet




Higgs+2jets sross section dependence on scale and jet size at NLO
6.0
5.5
5.0
4.5
4.0
3.5
Cross section, σ [pb]
HT/4
HT/2
HT
Scale, $\mu_F = \mu_R$
AntiKt1
AntiKt2
AntiKt3
AntiKt4
AntiKt5
AntiKt6
AntiKt7
AntiKt8
AntiKt9
AntiKt10
start to see non-monotonic behavior for small R


Higgs+3jets sross section dependence on scale and jet size at NLO
Cross section, σ [pb]
1.6
1.4
1.2
1.0
0.8
HT/4
HT/2
HT
Scale, $\mu_F = \mu_R$
AntiKt1
AntiKt2
AntiKt3
AntiKt4
AntiKt5
AntiKt6
AntiKt7
AntiKt8
AntiKt9
AntiKt10
non-monotonic
behavior starts
extending to larger R

Summary

- Precision LHC data requires precision LHC predictions
- NNLO provides that, for 2->2 processes, but still at NLO for more complex processes
- Scale dependence depends on kinematics and topology
  - 1 size may not fill all (uncertainties)
  - even if only 1 jet size is measured, variation of jet size in the theory may give a better understanding of the *fundamental* scale uncertainty
- Simple '1-D' scale evaluations may not give the full scope of the scale uncertainty
  - 2-D provides more information; a Olness-Soper type variation may give a more realistic estimate
- Applgrid/fastNLO storage trivial for NLO
- Applgrid/fastNLO storage may provide similar benefits for NNLO, especially for PDF fits
- Studies here will be extended to 2-D, and NNLO (if possible, for H/W+>=1 jets)
- Connection with MINLO-type scales may also provide insight