Fake leptons in SUSY

Giovanni Zevi Della Porta

Muon Fakes Workshop
https://indico.cern.ch/event/695762/

12 February 2018

Outline

1) Review the results of the SUS Fake Lepton Working Group

- Based on report given at CMS week in December 2016

2) Comment on recent implementations in the SUS group

2018 updates

“Fakes” include:

1) Leptons from decays of heavy-flavor and light-flavor hadrons
2) Hadrons misidentified as leptons
3) Conversions of γ in jets

Disclaimer: focus on Fake estimation, not Fake reduction, but the SUSY group has also advanced significantly in this direction, studying new variables:

MiniIso, p_T^{ratio}, p_T^{rel}, Rel+AbsIso, IP3D, σ_{IP3D}, LeptonMVA (nearest-jet: b-discr., #tracks), …

Goals and Results

Mandate of fake lepton working group (as of 17.1.2014)

Review existing procedures for estimating the fake+non-prompt lepton contribution and its uncertainty in the SUSY group

- Identify commonalities
- Understand differences (which are relevant, which not)

Establish benchmarks and common procedures

Contribute to Run 2 trigger menu

Design and implement supporting triggers for the method's Control Regions

Summarized at SUSY Workshop

- https://indico.cern.ch/event/339666/session/4/contribution/1/material/slides/0.pdf (Slides 38, 39)

Produced an Analysis Note summarizing results:

AN2014/261

Participating groups:

Brussels, Catania, Cantabria, CERN, ETH, Florida, Oviedo, UCSB/UCSD/FNAL

Coordinated by Frank Golf and Boris Mangano

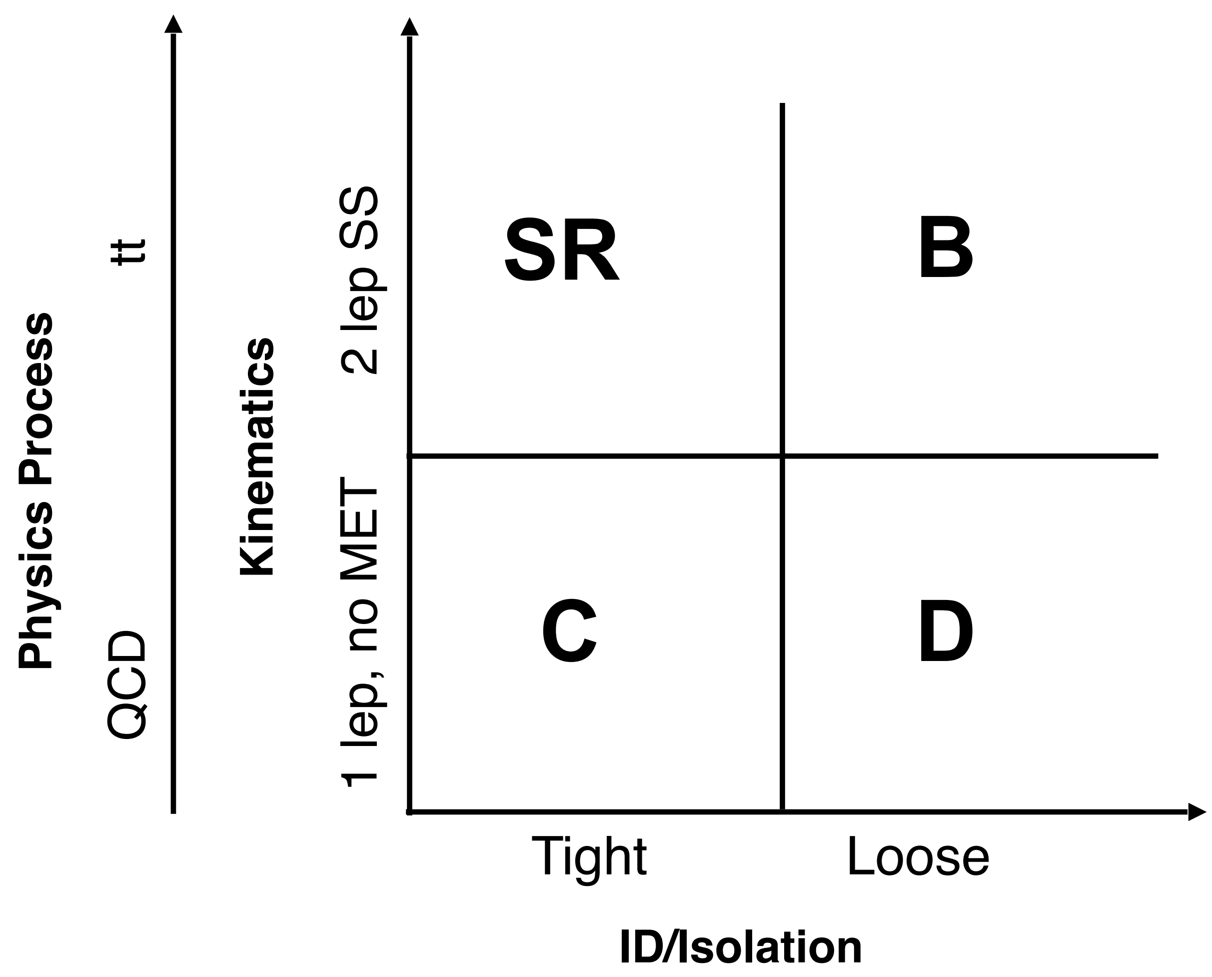
The tight-to-loose method

Focus on "tight-to-loose" or "fake rate" method

Most groups used a variation of this method in 2012.

The basic method: ABCD

- Assume B,C,D are fake dominated
- Assume two axes are uncorrelated
  —> SR/B = C/D = "Fake Rate"
  or "Transfer Factor" <—



Assumptions are not perfect:

1) Prompt lepton contamination in control regions (B, C, D)

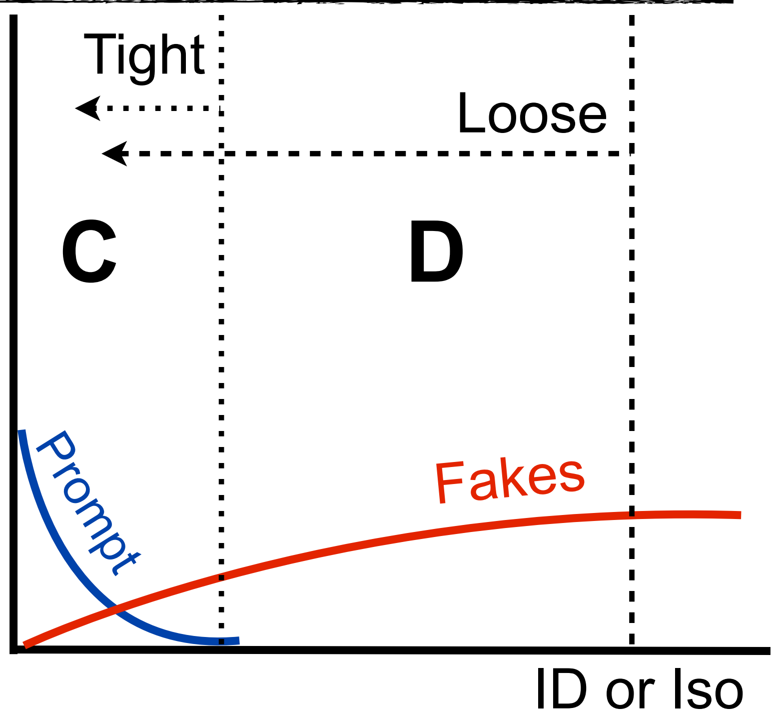
2) SR/B != C/D (even with zero prompt contamination)

- This is what we call "non-universality", which causes non-closure of the method

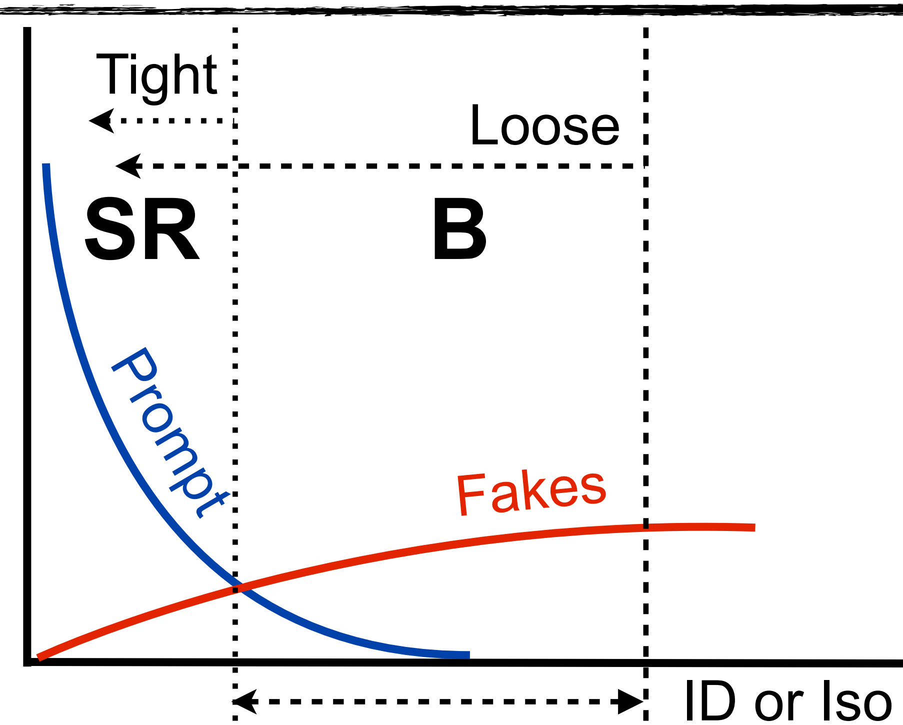
Another way to see it

Step-by-step procedure, including prompt contamination:

Measurement Region (Fake dominated)



Signal Region / Application Region



1 - account for prompt contribution
2 - extract FR = N_{tight} / N_{loose} for fakes

3 - account for prompts in Loose-not-Tight (B)
4 - count fakes in Loose-not-Tight
5 - use FR to predict Fakes in Tight region

"Fakes" = fake + non-prompt

Knobs and Choices

Many choices can be made, opportunities to diverge

a-- define Tight and Loose leptons

- tight/loose should be uncorrelated to sensitive variables in analysis, or parametrized
- loose should have ~ 10x more statistics than tight

b-- define a measurement region (fake-rich) in which to measure the tight/loose ratio

- consider available triggers and datasets
- design additional triggers, for example for low-pT single leptons, if necessary

a and b are analysis dependent, but the rest can be common to all analyses

Fixed a common "a" and "b", used this common language to compare methods to:

c-- deal with prompt lepton contamination in measurement region

d-- deal with prompt contamination in application region

e-- define a closure test

f-- determine systematic uncertainties

g-- understand non-closure and improve the method

“Fakes” = fake + non-prompt

c: prompt leptons in measurement region

Having defined a measurement region, need to

(1) remove prompt contamination

Simplest approach (MT<20, MET<20) included in baseline definition

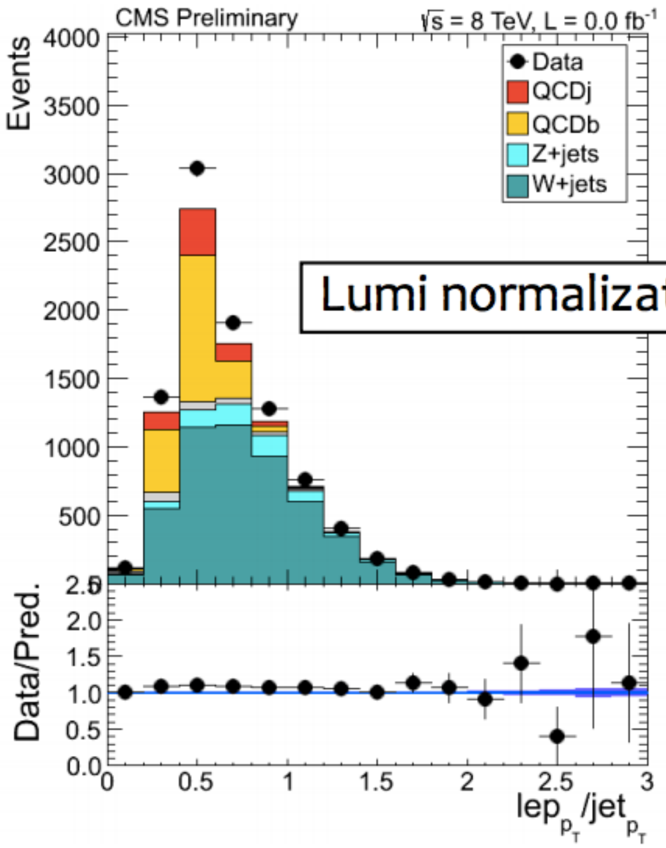
Additional recoil jet selections discussed (b-tag, $\Delta\varphi_{(jet,lep)}$, p_T^{lep}/p_T^{jet})

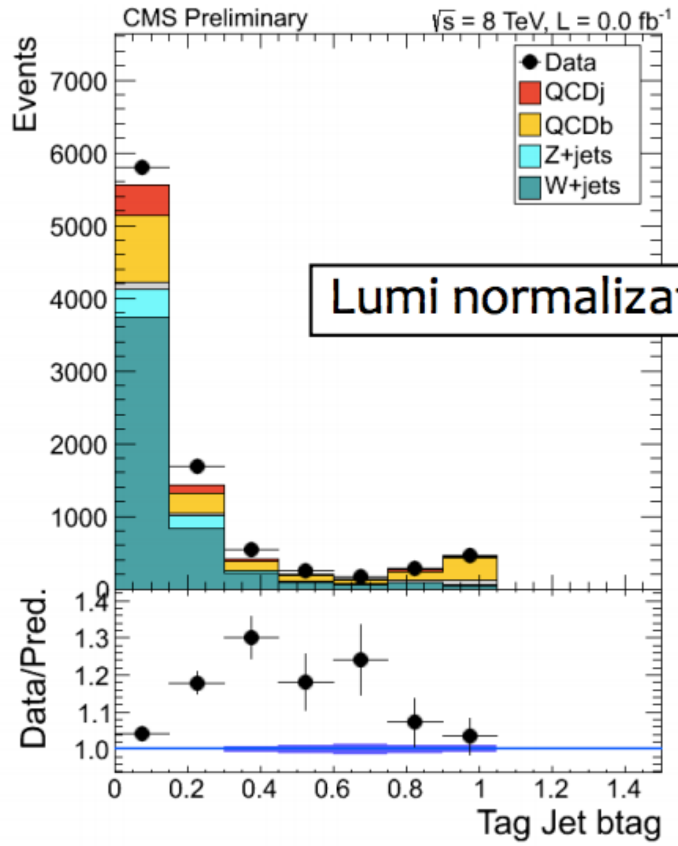
- https://indico.cern.ch/event/314612/contribution/2/material/slides/0.pdf
- https://indico.cern.ch/event/313140/contribution/1/material/slides/0.pdf
- Pros: lower prompt contamination
- Cons: higher statistical uncertainty, possible sample bias

Use of additional variables will depend on analysis

Examples of additional variable for prompt suppression:

- pTlep/pTrecoilJet,
- b-tag discriminator





c: prompt leptons in measurement region

Having defined a measurement region, need to

(1) remove prompt contamination

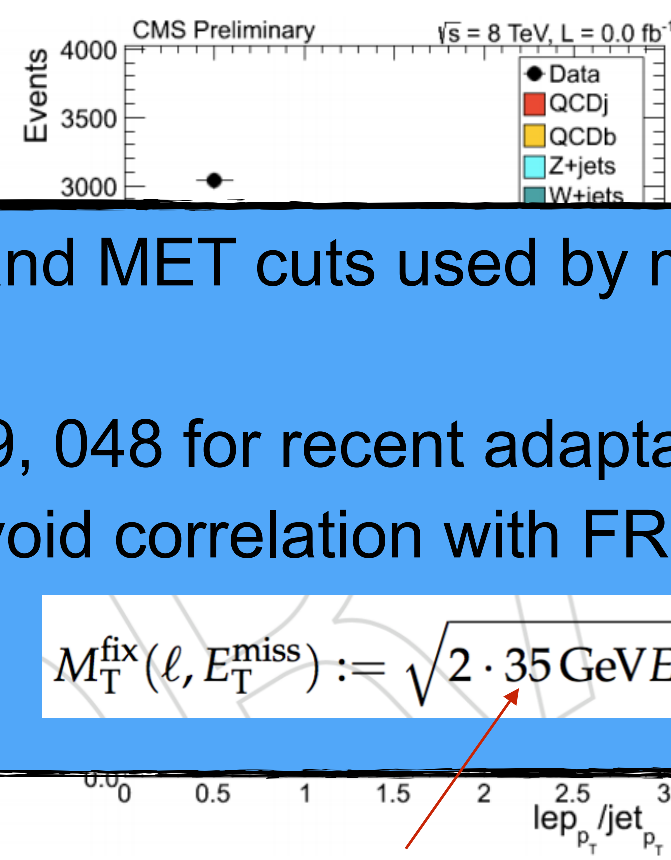
Simplest approach (MT<20, MET<20) included in baseline definition

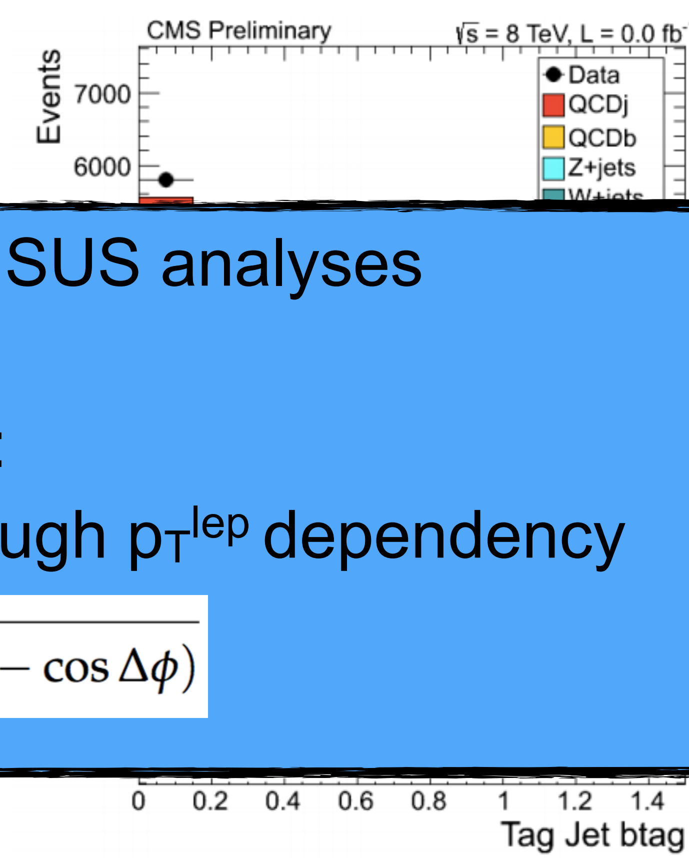
Additional recoil jet selections discussed (b-tag, $\Delta\varphi_{(jet,lep)}$, p_T^{lep}/p_T^{jet})

- https://indico.cern.ch/event/314612/contribution/2/material/slides/0.pdf
- https://indico.cern.ch/event/313140/contribution/1/material/slides/0.pdf
- Pros: lower prompt contamination
- Cons: higher statistical uncertainty, possible sample bias

Use of additional variables will depend on analysis

Examples of additional
va
su
- p
- b





2018 update: MT and MET cuts used by most SUS analyses

See SUS-16-039, 048 for recent adaptation:

Use m_T^{fix} to avoid correlation with FR through p_T^{lep} dependency

$$M_T^{fix}(\ell, E_T^{miss}) := \sqrt{2 \cdot 35\,\mathrm{GeV} E_T^{miss}(1 - \cos\Delta\phi)}$$

c: prompt leptons in measurement region

Having defined a measurement region, need to

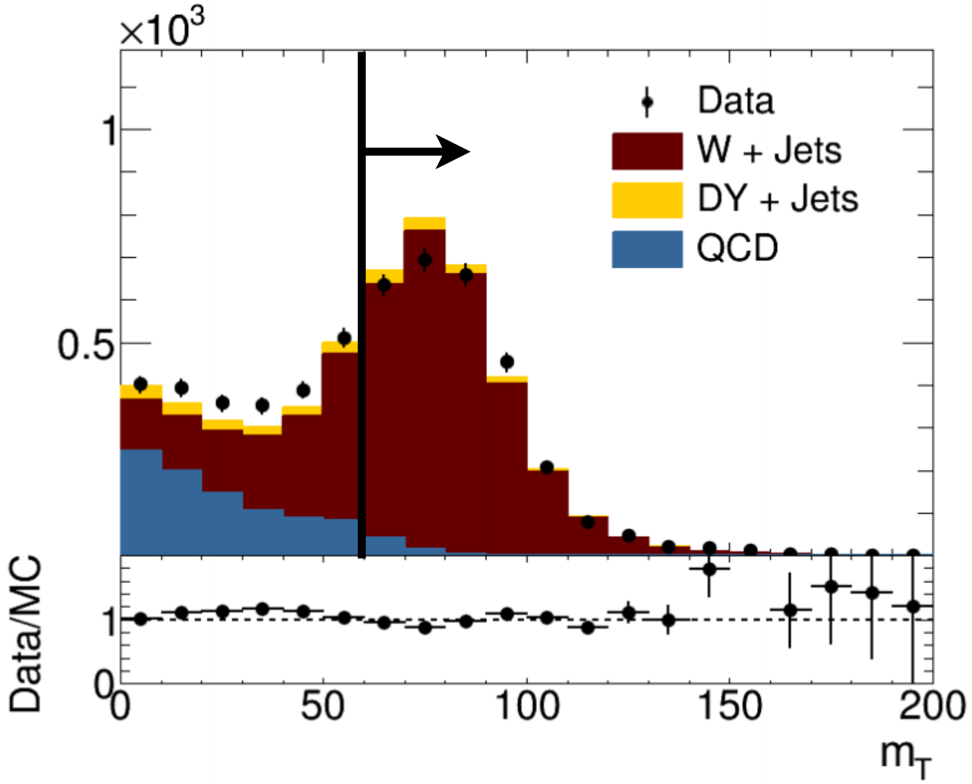
(2) estimate residual prompt contamination, subtract it

Initial estimate based on pure MC (W,Z,tt)

- Based on this, decide whether to subtract (if large) or just apply uncertainty

Subtraction is based on MC normalized to data

- Different methods used to normalize prompt MC for subtraction (https://indico.cern.ch/event/309096/)
- We recommend the simplest one as a starting point
  - Normalize prompt MC in high-MET, high-MT region (ignore QCD there)
  - Take 50% of difference from default cross-section as uncertainty



MET>30 GeV, MT>60 GeV, Tight
Used to set normalization of prompt MC

c: prompt leptons in measurement region

Having defined a measurement region, need to

(2) estimate residual prompt contamination, subtract it

Initial estimate based on pure MC (W,Z,tt)

- Based on this, decide whether to subtract (if large) or just apply uncertainty

Subtraction is based on MC normalized to data

- Different methods used to normalize prompt MC for subtraction (https://indico.cern.ch/event/309096/)
- We recommend the simplest one as a starting point
  - Normalize prompt MC in high-MET, high-MT region (ignore QCD there)
  - Take 50% of difference from default cross-section as uncertainty

2018 update:

Why not take MC normalization directly?

More work: measure SFs for “Loose” definition, make sure trigger pre-scales are handled correctly

Does it make sense to take 50% of Δ(Data, MC) as uncertainty?

No. More defensible approaches used by some analyses:

- Difference between Data normalized in Z and W peaks (accounts for lep. eff)
- Use MC directly (with SFs and pre-scales), and apply standard RECO unc.
- Fit m_T distribution for both Tight and Loose leptons

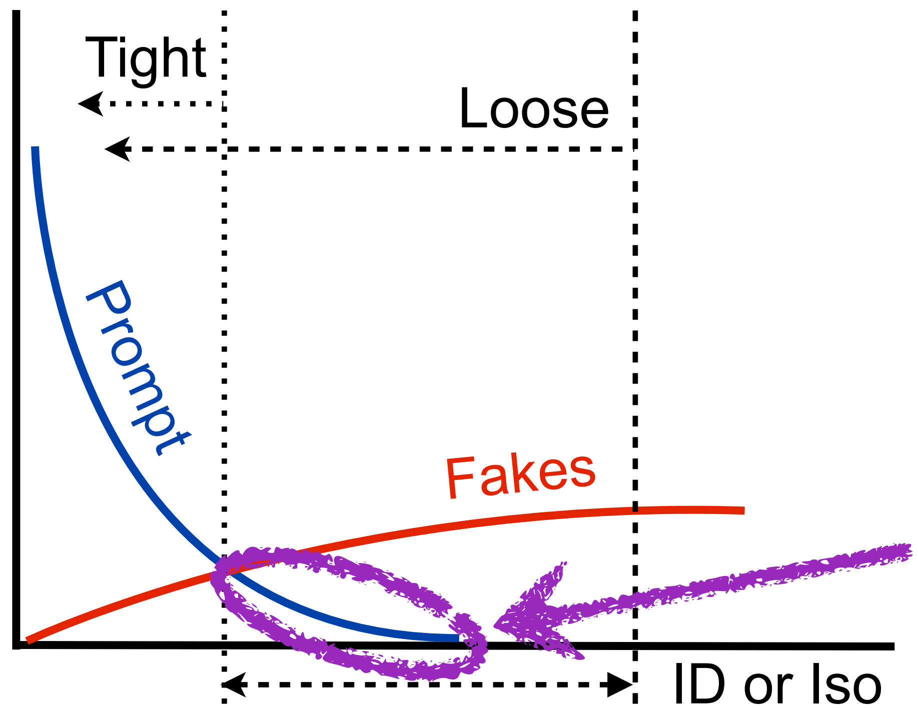
d: prompt leptons in application sideband

Two approaches for prompt contamination treatment (small effect)

Discussion: https://indico.cern.ch/event/319663/

1. Include a "prompt rate" factor in the fake rate equation

- CMS AN-2010/261
- Pro: Data-driven, based on tag/probe efficiency and leptons in Tight region
- Cons: Efficiency in Z sample is not the same as ttbar sample



2. Use MC prediction

- Pro: not relying on absolute efficiency in Z events, just SF between Data and MC
- Cons: prompt MC prediction can be wrong

Recommendation is to use either, since both have drawbacks and have similar performance and effect is small.

2018 update: All analyses use (2).

e: what is a closure test?

Basic recipe

1. Determine FR in measurement region using QCD MC
   - No prompt contamination in this sample (it's QCD!)

2. Apply it to estimate fake leptons in Signal Region
   - Use gen-info to make sure we are applying FR to the fake lepton (prompt contamination = 0)
   - Use full MC "soup" (tt, W+jets, any other process which generates fakes in this SR)

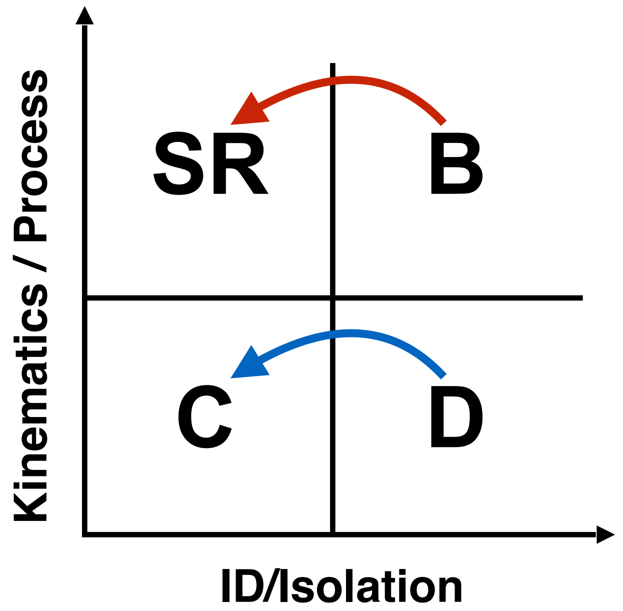
3. Compare estimate with actual number of fakes predicted by MC

4. Study kinematic dependence within Signal Region, to test limits of closure

Use closure test to validate the main assumption:

—> "the two axes are uncorrelated" <—

$$\frac{C}{D} \stackrel{?}{=} \frac{SR}{B}$$



Closure test example (SUS-15-035)

Based on ABCD, treating MC as data

Prediction [788]

$t\bar{t}$ [711]

W+jets [85]

Actual yield of Fakes based on MC truth info.

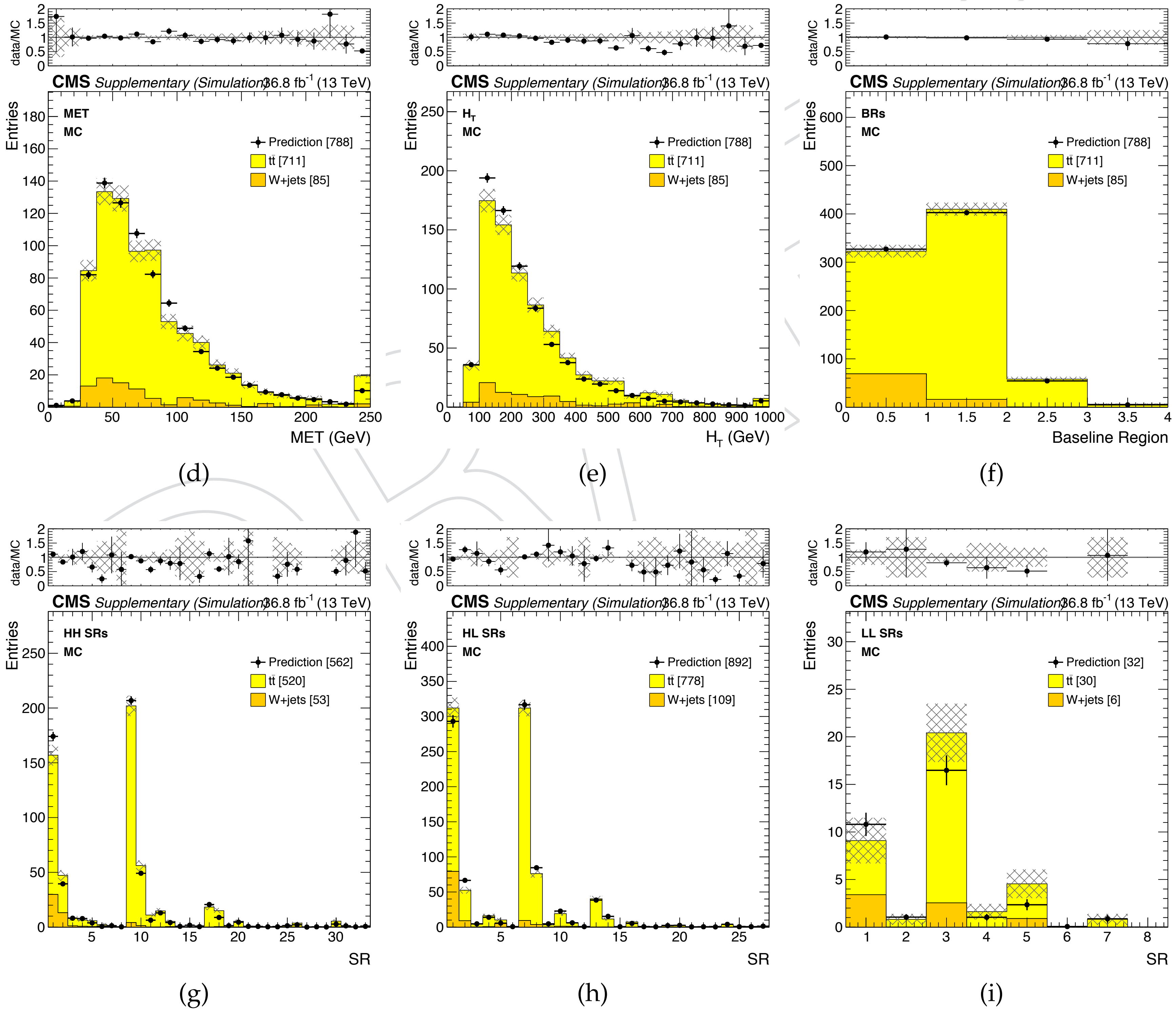


Figure 27: Electron+muon fake rate closure for QCD measurement in MC soup: trailing lepton p_T, M_T^{min}, E_T^{miss} and H_T distributions combining isolated and non-isolated triggers. Samples used: $t\bar{t}$ POWHEG, W+jets

Contributors to non-closure

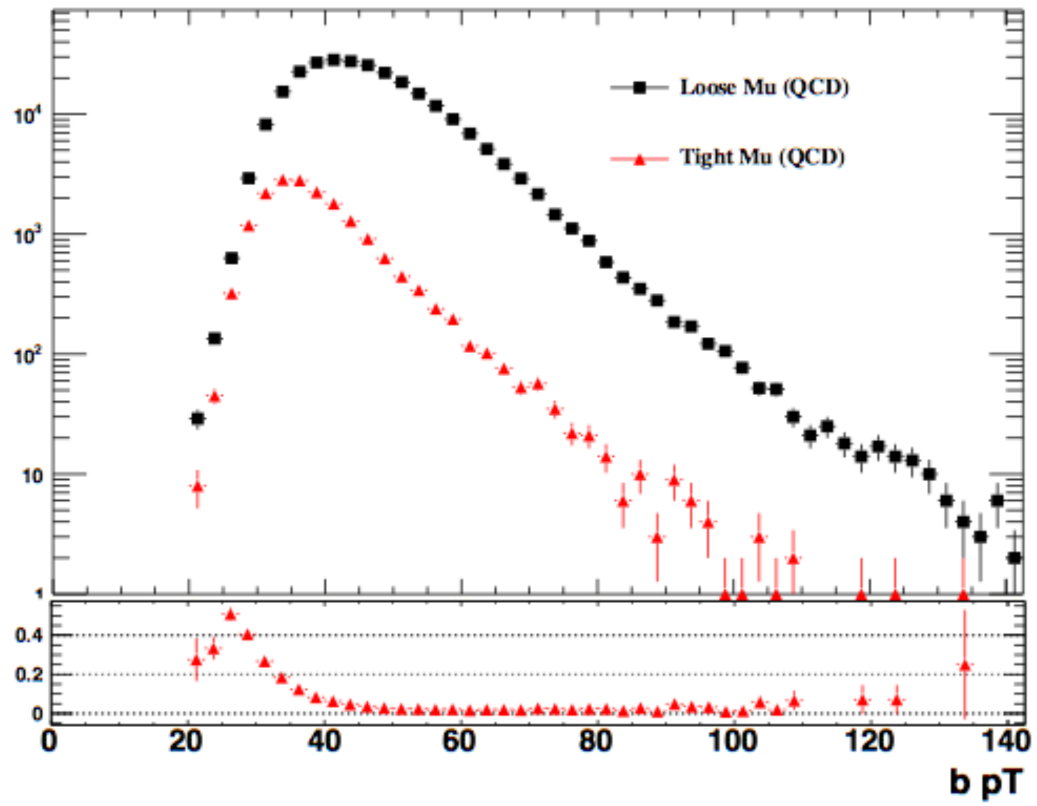
Lack of closure (discussed in next section) usually related to the <u>combination of</u>:

1. Correlation between a hidden variable and the definition of Tight vs. Loose leptons
   - The Fake Rate (Tight/Loose) depends on this variable
2. The same variable is different in Measurement and Application region

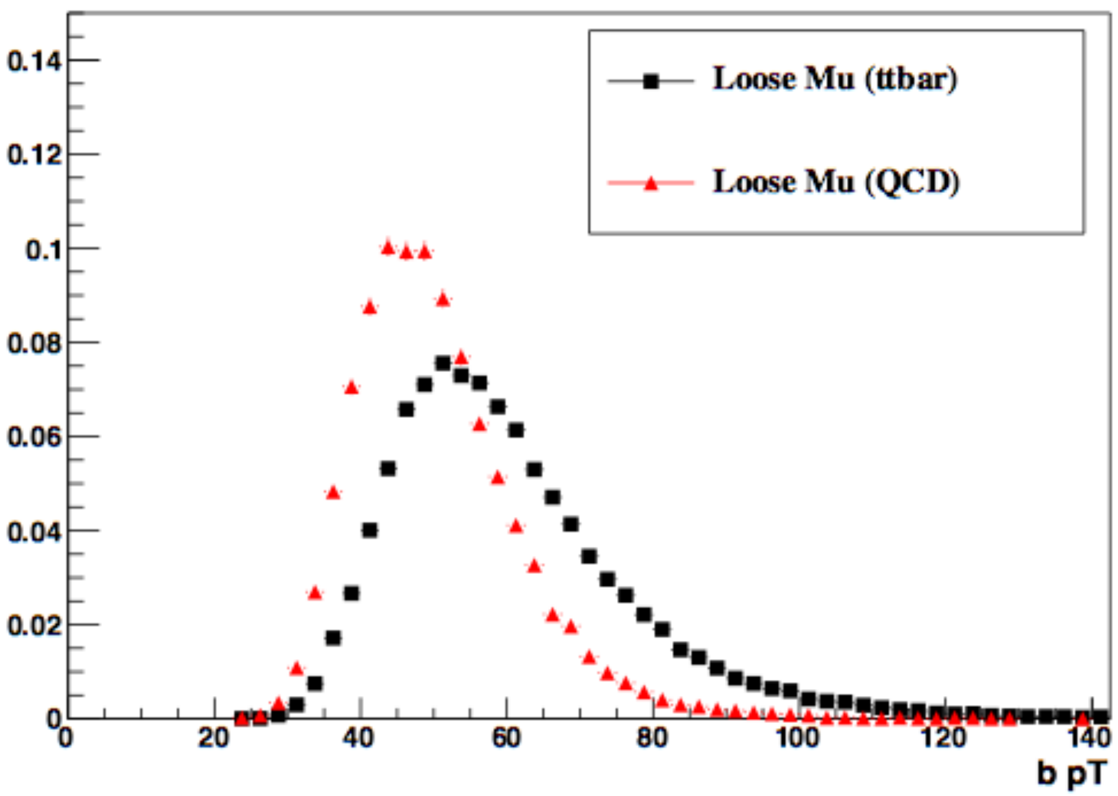
Main examples are the <u>lepton origin</u> (flavor composition) and its <u>mother parton pT</u>

For example, mother b-quark pT for muons with pT 20-25 GeV

Fake Rate is not flat vs b-quark pT



b-quark pT is different in QCD and ttbar



f: Systematics

Non-universality of tight-to-loose ratio is main systematic effect

Account for leading differences between application and measurement region

- mother parton pT, flavor composition are main effects
- effect will depend on definition of tight and loose

For each of these, we suggest variations

- Mother parton pT:
  - Vary recoil jet pT in measurement region
- Flavor composition (in progress):
  - Require or veto b-tag on recoil jet in measurement region

Prompt contamination in measurement and application regions

Uncertainty on prompt lepton subtraction also enters systematics

Systematics, closure, improvements

Use the results of closure test to confirm validity of systematics

- If closure is good (± systematics), then continue using systematics
- If closure is outside of systematics range, then one effect is missing
  - Option A: study and understand missing effect, add a systematic for it
  - Option B: take non-closure as single systematic
- *If trend is observed when closing in different regions, consider adding systematic for it*

Improvements need to show their power compared to these systematics

- Improvements that aim to account for flavor dependency and parton pT should result in smaller systematics
- Other effects (currently subdominant) might become more relevant if improvements reduce these uncertainties

Systematics, closure, improvements

Use the results of closure test to confirm validity of systematics

If closure is good (± systematics), then continue using systematics

If closure is outside of systematics range, then one effect is missing

- Option A: study and understand missing effect, add a systematic for it
- Option B: take non-closure as single systematic

If trend is observed when closing in different regions, consider adding systematic for it

2018: Since these studies, all SUSY analyses have based systematics on:

- Closure in MC (including kinematic dependence, and covering each signal region)
- Data/Prediction agreement in a fake-enriched Validation Region
- Additional systematic for Prompt Contamination

Studying non-closure

Understand the main effects

Large (~200%) non-closure observed after basic method

	ee	mm	em	em (e fake)	em (m fake)	ll
Loose!Tight	133.94	1346.35	1314.73	176.85	1137.87	2795.02
pred	65.48 +/- 2.83	42.18 +/- 0.71	121.72 +/- 3.71	86.30 +/- 3.66	35.42 +/- 0.60	229.38 +/- 4.72
obs	23.02	16.55	41.14	26.99	14.15	80.70
pred/obs	2.84 +/- 0.16	2.55 +/- 0.12	2.96 +/- 0.12	3.20 +/- 0.17	2.50 +/- 0.12	2.84 +/- 0.08
(p-o)/p	0.65 +/- 0.02	0.61 +/- 0.02	0.66 +/- 0.01	0.69 +/- 0.02	0.60 +/- 0.02	0.65 +/- 0.01

Remove effects of flavor composition

- Require B-hadron mother, ΔR match status3 b-quark (avoiding g-->bb in parton shower)

b-only	1.38 +/- 0.17	2.28 +/- 0.11	1.79 +/- 0.13	1.56 +/- 0.19	2.24 +/- 0.12	1.78 +/- 0.08

- Improvement in electrons, almost no change for muons (already dominated by b)

Remove effects of mother parton pT

- Reweight pTb-quark in measurement region to match Signal Region B

b-only, pTb	0.93 +/- 0.12	1.34 +/- 0.10	1.15 +/- 0.10	1.06 +/- 0.14	1.32 +/- 0.10	1.13 +/- 0.06

- Improvement in both electrons and muons

Addressing the main effects

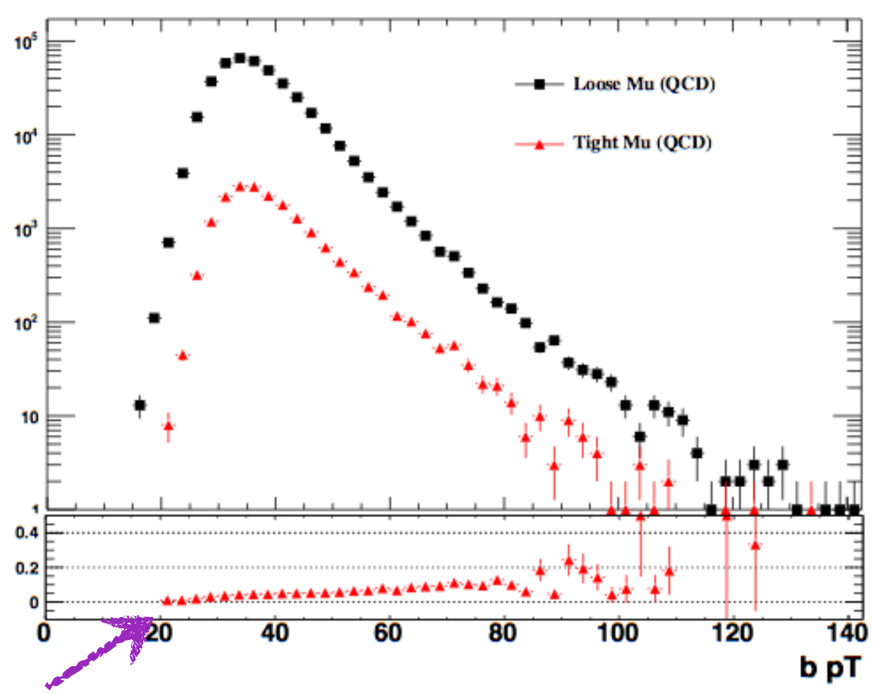
In principle, non-closure causes could be parametrized

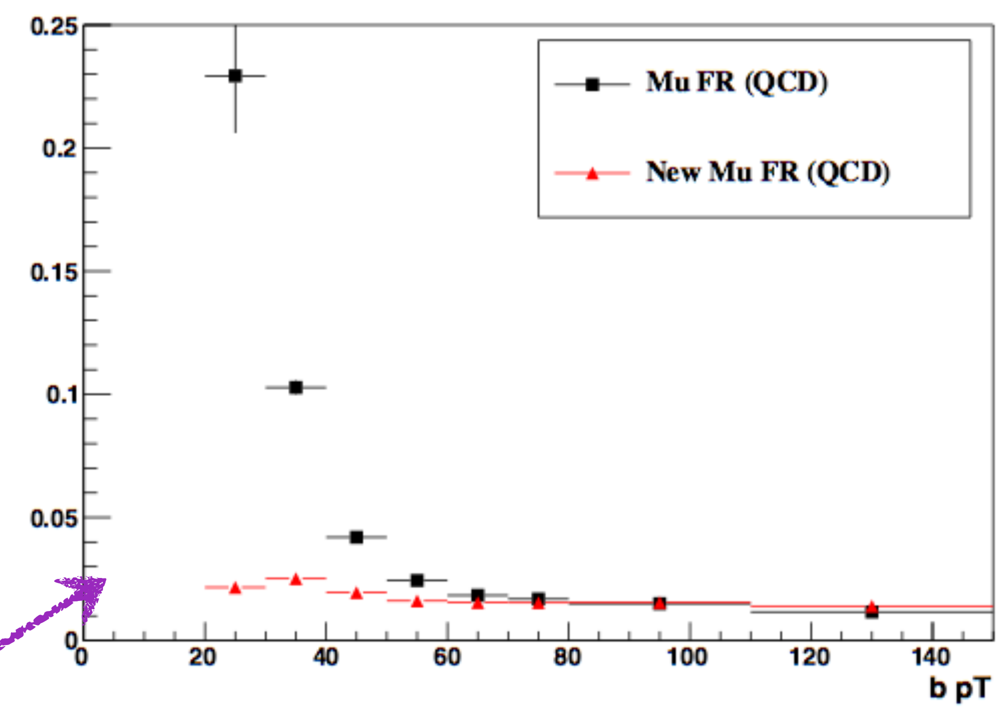
- For example, we already do this with lepton kinematics: (p_T^l, η^l)
- Can we do this also for p_T and flavor (b/c/light) of lepton's parton parent (p)?

$$TF(p_T^l, \eta^l, p_T^p, f^p)$$

Use two tricks to avoid a 4D Fake Rate

- Tune the "Loose" selection (mainly important for electrons): $FR|_b = FR|_c = FR|_l$
- Combine p_T^l and p_T^p: $p_T^{cone} = \text{Tight} \; ? \; p_T^l : p_T^l(1 + \text{RelIso})$
- Can use a 2D FR: (p_T^{cone}, η^l)





Much smaller dependency of Tight/Loose ratio on pTb

Addressing the main effects

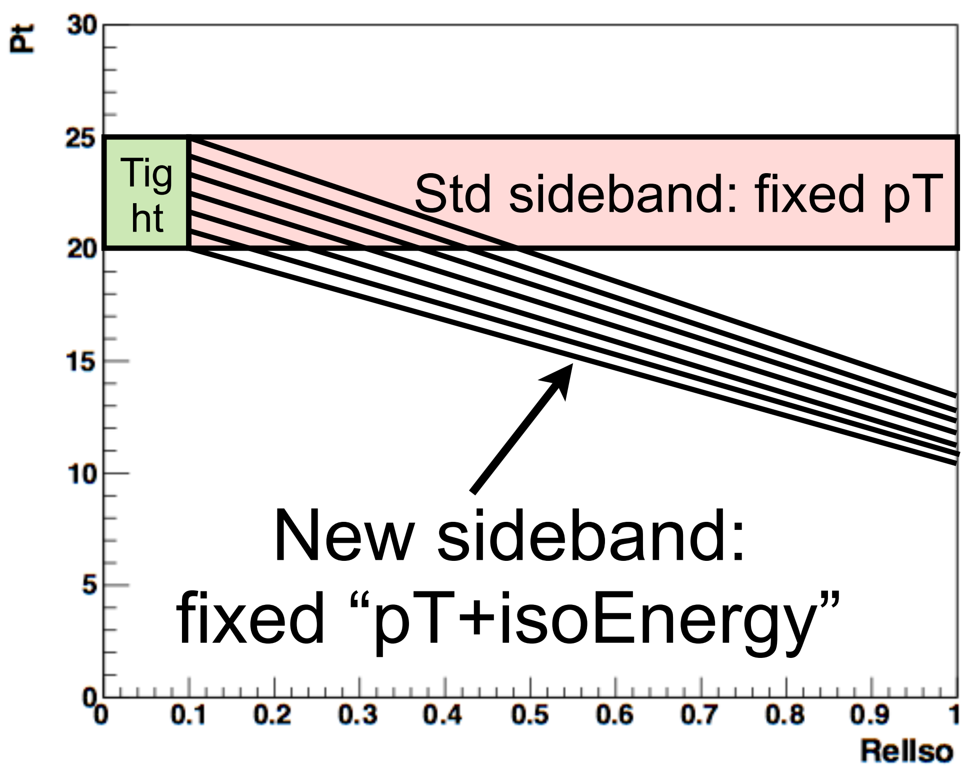
In principle, non-closure causes could be parametrized

- For example, we already do this with lepton kinematics: (p_T^l, η^l)
- Can we do this also for p_T and flavor (b/c/light) of lepton's parton parent (p)?

$$TF(p_T^l, \eta^l, p_T^p, f^p)$$

Use two tricks to avoid a 4D Fake Rate

- Tune the "Loose" selection (mainly important for electrons): $FR|_b = FR|_c = FR|_l$
- Combine p_T^l and p_T^p: $p_T^{cone} = \text{Tight} \; ? \; p_T^l : p_T^l(1 + \text{RelIso})$
- Can use a 2D FR: (p_T^{cone}, η^l)



Improvements bring "closure" of ABCD well within ±30%, across sample kinematics

Note: p_T^{cone} requires lowering lepton p_T, so can be affected by trigger turn-on.

Main effects understood, but closure is still not perfect. Why?

Study residual non-closure

Many investigations to study into residual non-closure

- Reconstruction level effects (PU subtraction)
- Move the study to Gen level, where we have full control
  - kinematics of fragmentation/decay: pT(b-quark) vs pT(muon) and pT(GenJet)
  - focus on B decay products to calculate “Isolation”
  - underlying event
  - neighboring jets (ΔR with closest GenJet)
  - total activity (SumPt of status1 particle in event)
  - total activity at Reco level (using charged PFcandidates)

Studies in backup, several small effects observed, most important is the “total activity” in the event

Further improvements in closure would have to reduce or parametrize away this dependency

Conclusions and analysis dependence

These studies have helped the SUSY group adopt a common and detailed approach for estimating Fakes, and greatly reduced systematic uncertainties

Fake estimation is still not straightforward, because (a) and (b) are analysis dependent

a-- define Tight and Loose leptons

b-- define a measurement region (fake-rich) in which to measure the tight/loose ratio

As long as lepton IDs continue to improve and analyses reach new phase spaces, new analyses will have to re-optimize the Loose definition and Measurement region choice

A few examples on the next slide

Examples from recent SUSY analyses

Analysis	Tight	Loose	Measurement Region	Validation Region
SUS-16-035	"MultiIso" (MiniIso, pTratio, pTrel)	MiniIso (with p_T^{cone} parametrization)	Pre-scaled single-lepton triggers	NA
SUS-15-008	"MultiIso" (MiniIso, pTratio, pTrel)	MiniIso (with p_T^{cone} parametrization)	As above + cross-check region: IP (d0/dz) sideband	NA
SUS-16-041	"MultiIso" (MiniIso, pTratio, pTrel)	MiniIso (with p_T^{cone} parametrization)	As above + cross-check region: IP (d0/dz) sideband	N_{jets} : 1-2 $N_{b\text{-}jets}$: 1
SUS-17-005 (pT < 25)	Iso03 < 5 GeV d0, dz < 0.01, 0.2 cm	RelIso03 < 20 GeV, d0, dz < 0.1, 0.5 cm	HT > 900 GeV (HT-triggered)	NA
SUS-16-048 (pT < 30)	Iso03 < 5 GeV, RelIso03 < 0.5, tight IP cuts	RelIso03 < 20 GeV + $300/p_T^{lep}$, relax IP	Pre-scaled single-lepton triggers	Same-sign
SUS-16-039	"LeptonMVA" (iso and nearby jet info)	Relaxed LeptonMVA, with p_T^{cone}-like parametrization	Pre-scaled single-lepton triggers	NA

Backup

a and b: Definitions and Samples

Common tight and loose lepton definition for e and μ

POG-based selections with improvements from 2012 SS analyses

Twiki: https://twiki.cern.ch/twiki/bin/viewauth/CMS/SUSYFakeLeptonWorkingGroup

Common samples and selection

2012 Data (periodD) and MC

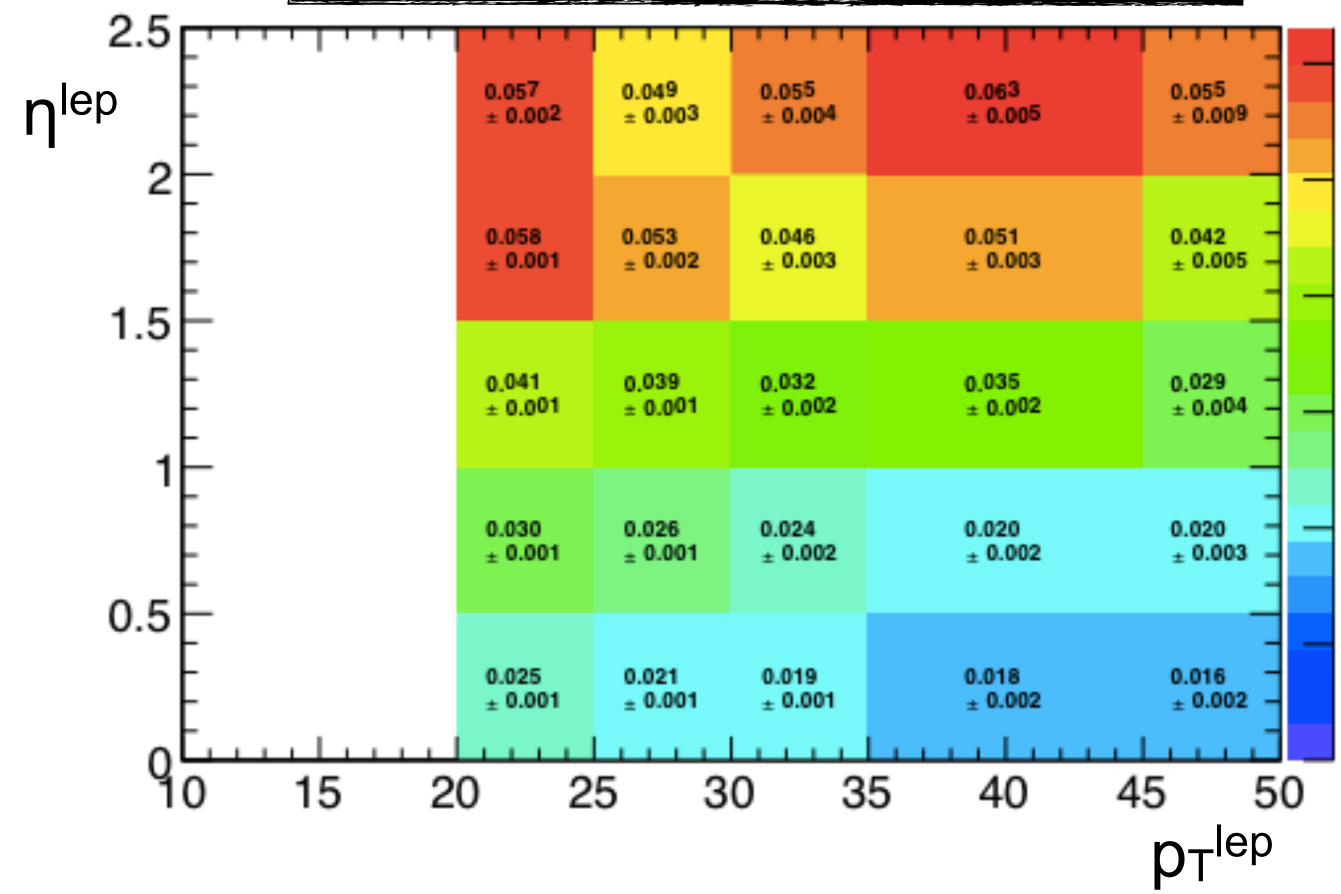
Measurement region (QCD sample)

- NLep=1
- pTjet>40, DR(jet,lep)>1
- MET<20, MT(lep,MET)<20

Application regions for closure test:

- MET>30GeV, NLep=2
- SRA (WJets): Any sign, NJ≥1, NB=0
- SRB (ttbar): Same sign, NJ≥2, NB≥1

Fake Rate map for measurement region (MC)



Non-closure investigations

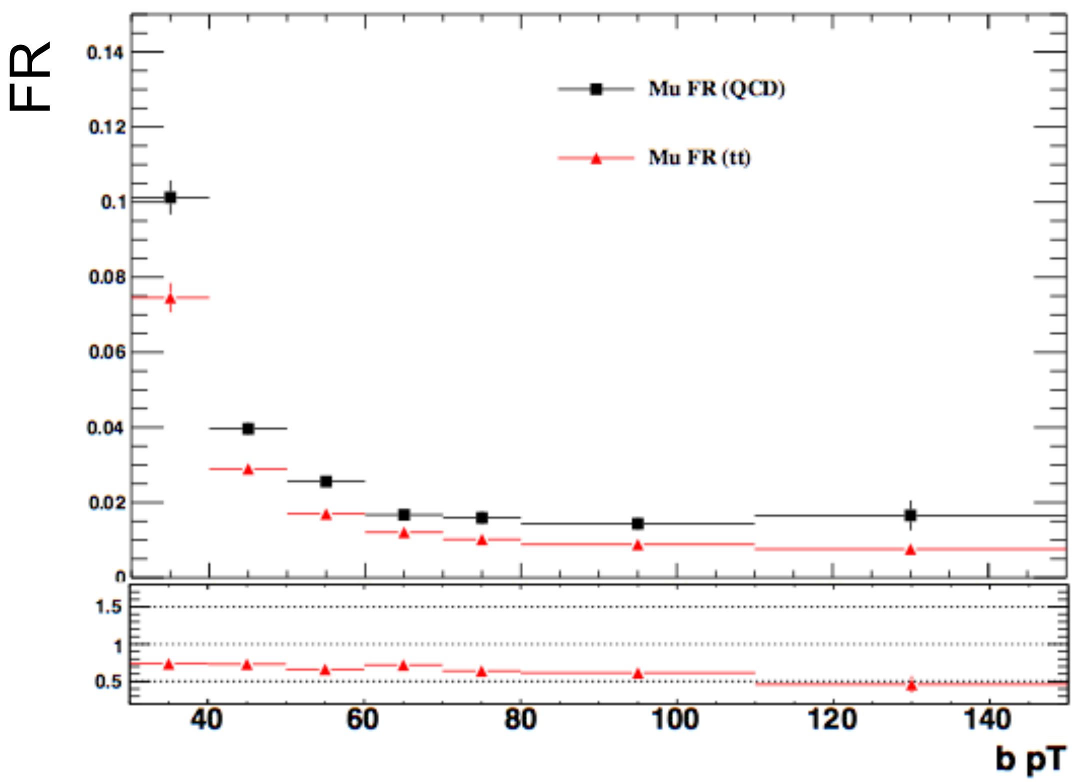
What about residual non-closure

Even using Gen info, closure is not perfect

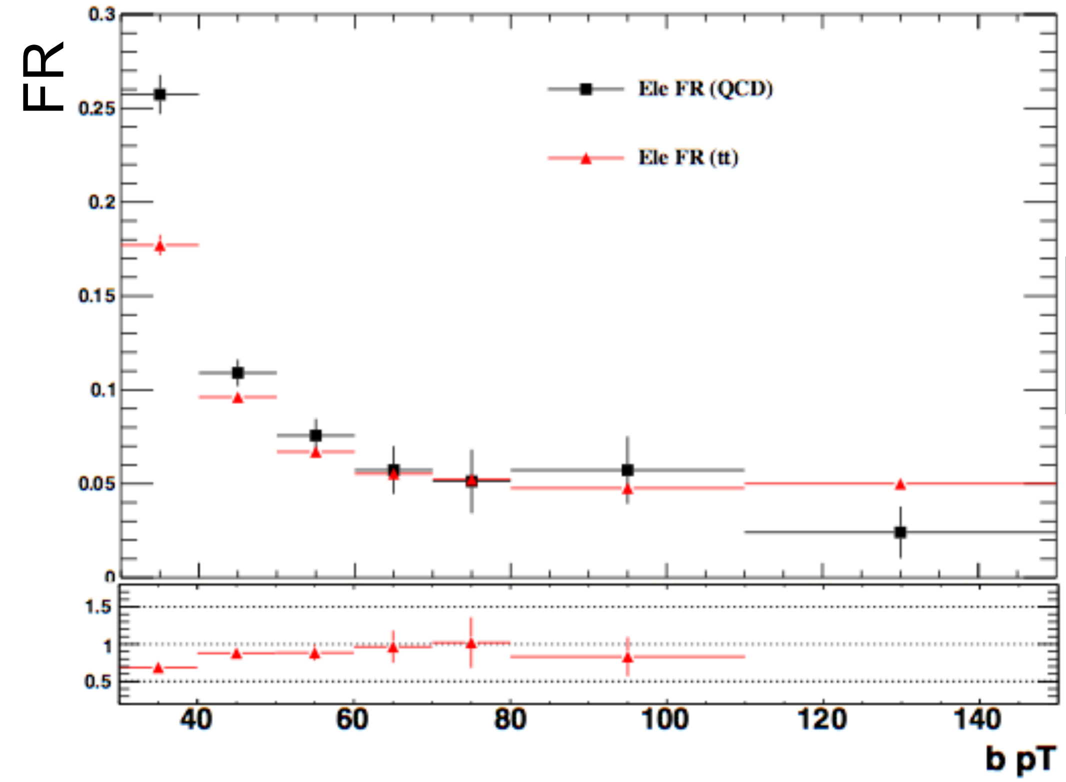
pred/obs	ee	mm	em	em (e fake)	em (m fake)	ll
b-only, pTb	0.93 +/- 0.12	1.34 +/- 0.10	1.15 +/- 0.10	1.06 +/- 0.14	1.32 +/- 0.10	1.13 +/- 0.06

Another way to see it: Fake Rate vs pTb, in b-only samples

Muons



Electrons



$|\eta^{lep}| < 0.9$ and iso-sideband only for all closure slides

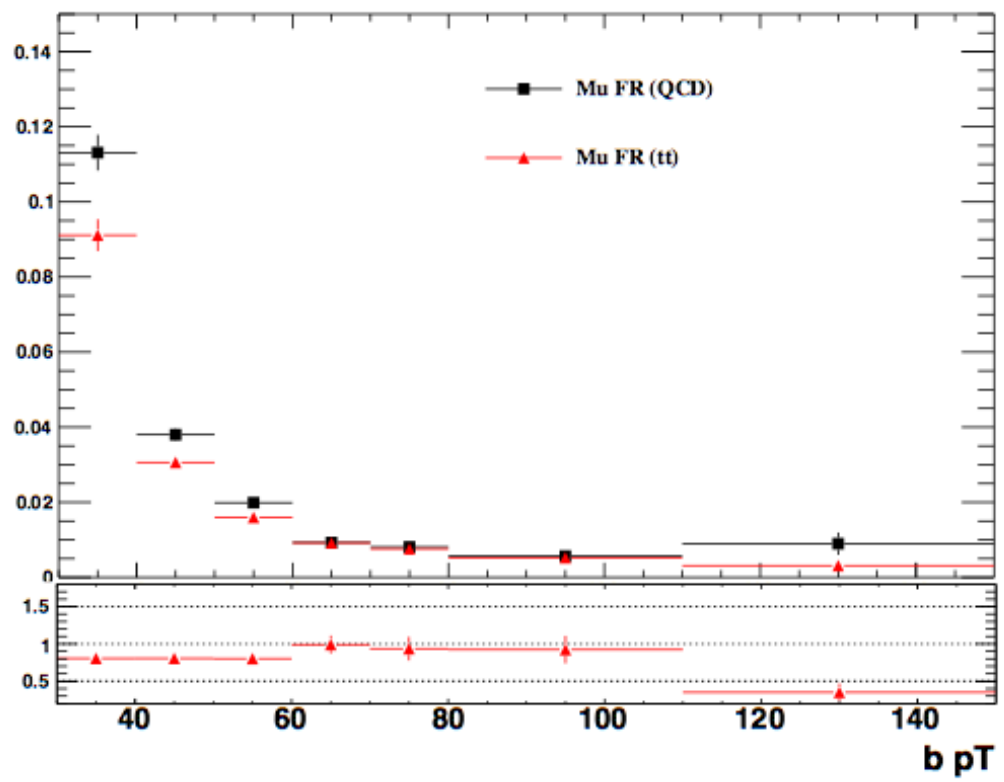
Isolation of μ in b-quarks is sample dependent. Why?

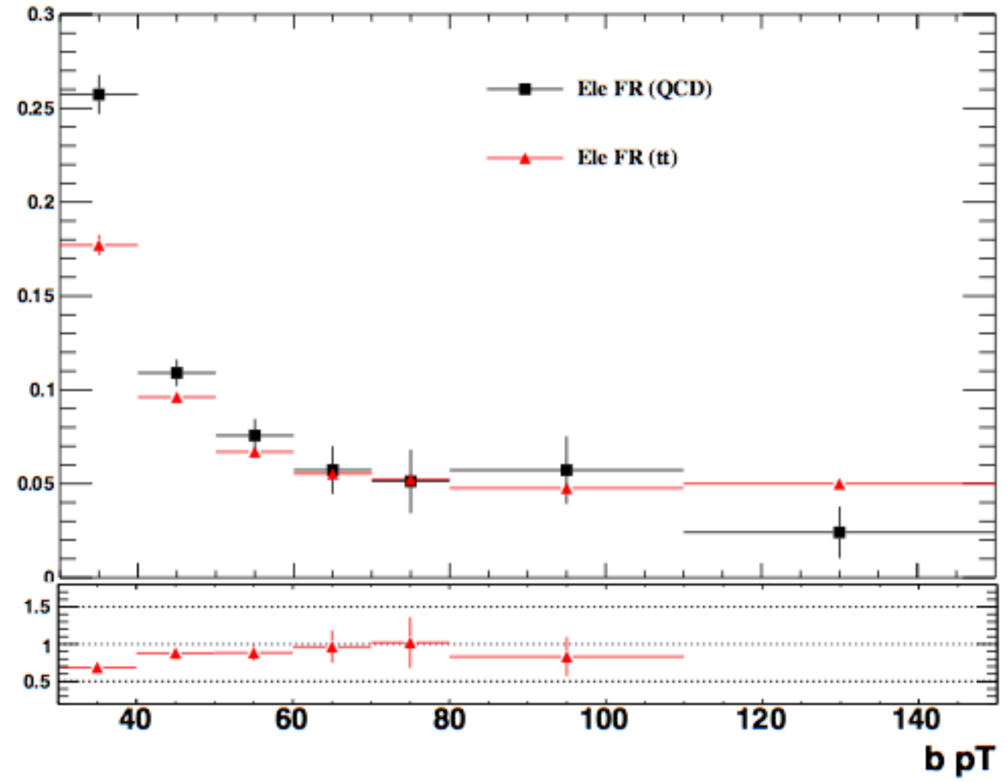
Reco effects: PU subtraction

At high pTb, PU subtraction method has a non-negligible effect

Using EffectiveArea method for muons yields better agreement than Δβ

- Δβ relies on primary vertex resolution. Different in ttbar and QCD?
- ρ is sensitive to overall event energy



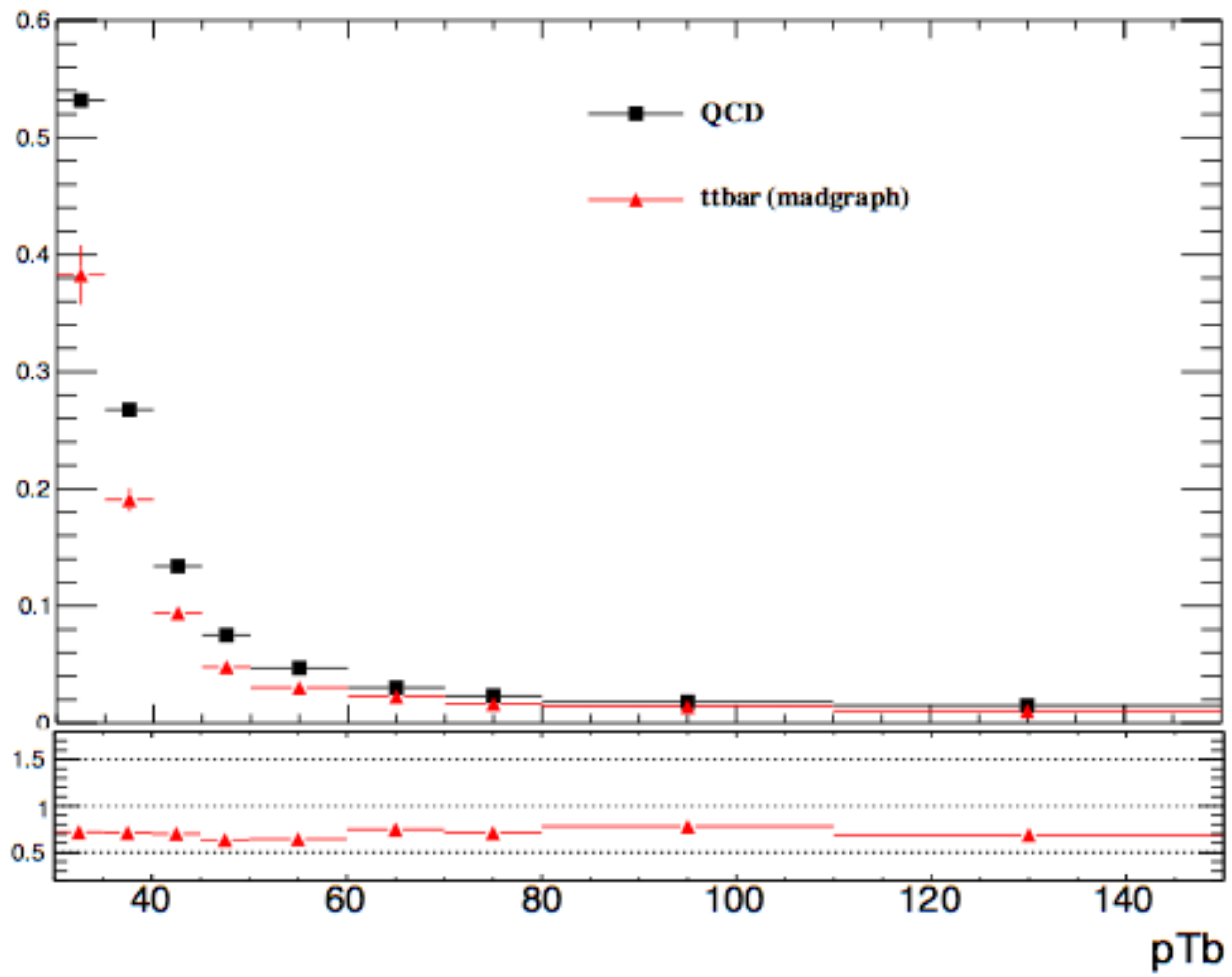


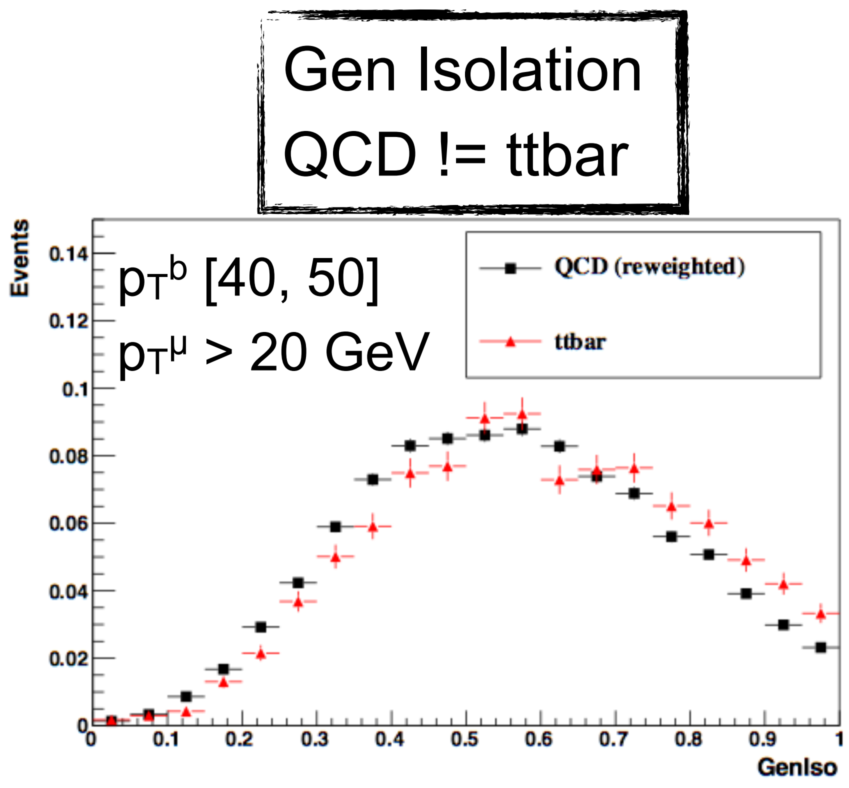
To rule out Reco effects and pile-up, need to check at gen-level

Discrepancy at Gen-Level

To rule out Reco effects and pile-up, check at gen-level

- Use Gen muons, with no ID requirement except gen isolation
- Gen Isolation = sum of status1 (except neutrinos). Tight < 0.3, Loose < 1.0.





Discrepancy present across pTb, stable at ~30%

Checked a pythia ttbar sample, same effect there

Gen checks: underlying event

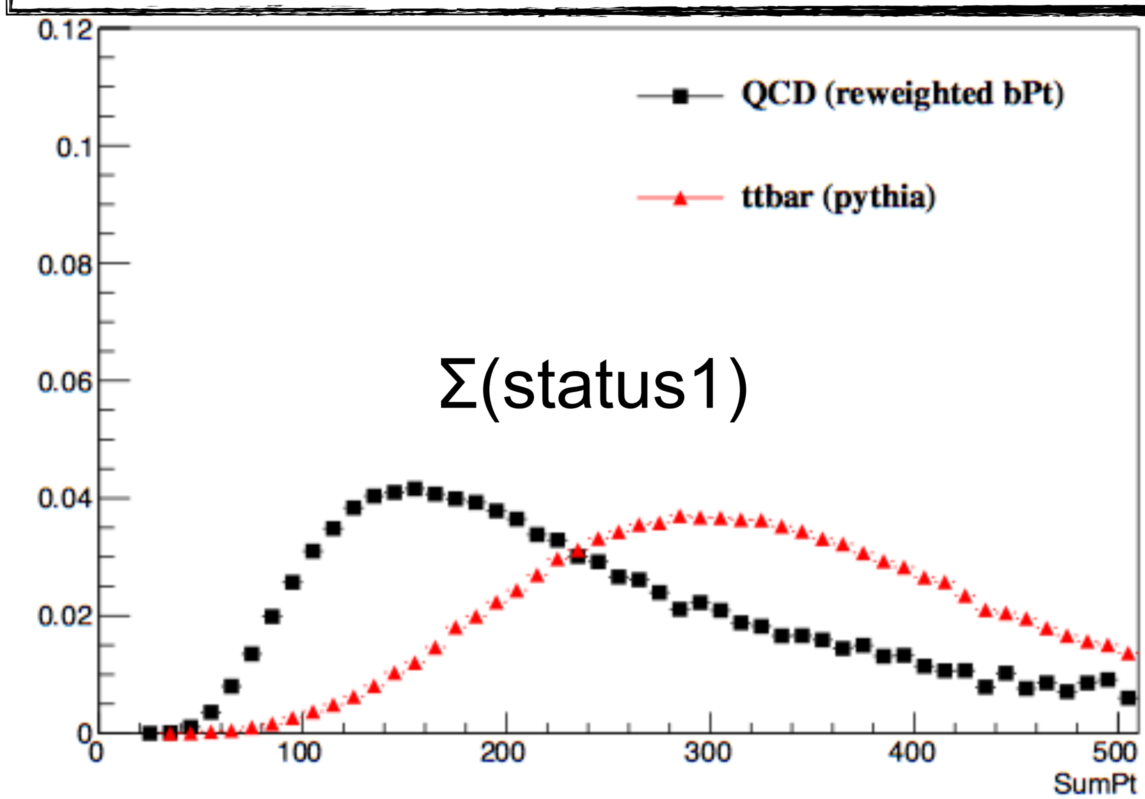
UE definition inclusive of ISR/FSR (as in TOP-13-007)

Sum all status1 particles within eta < 1.2 (don't count neutrinos)

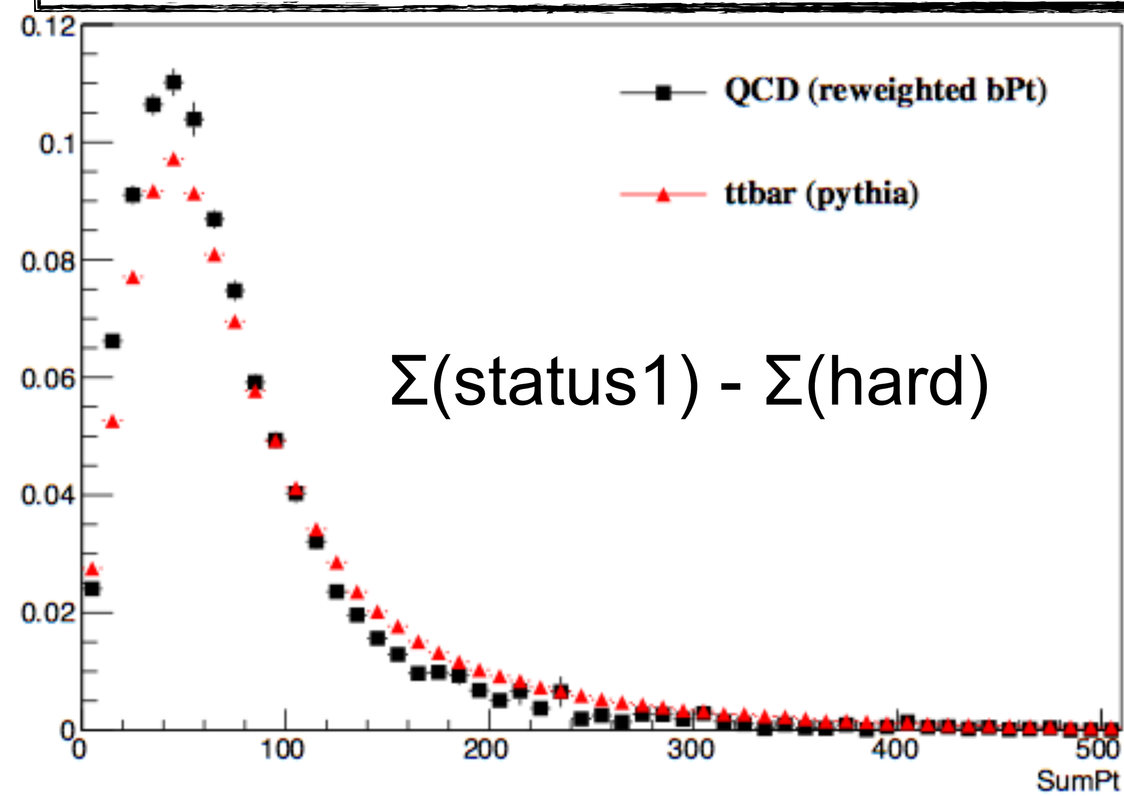
- Subtract status3 leptons (these are for sure from ttbar)
- Subtract status1 particles within DR<0.4 of status3 quarks (udcsb) or gluons

pTb distribution is accounted for by reweighting QCD events

SumPt quite different: expected, since it contains hard scattering



UE similar in QCD and ttbar, further reweighting has no effect on closure



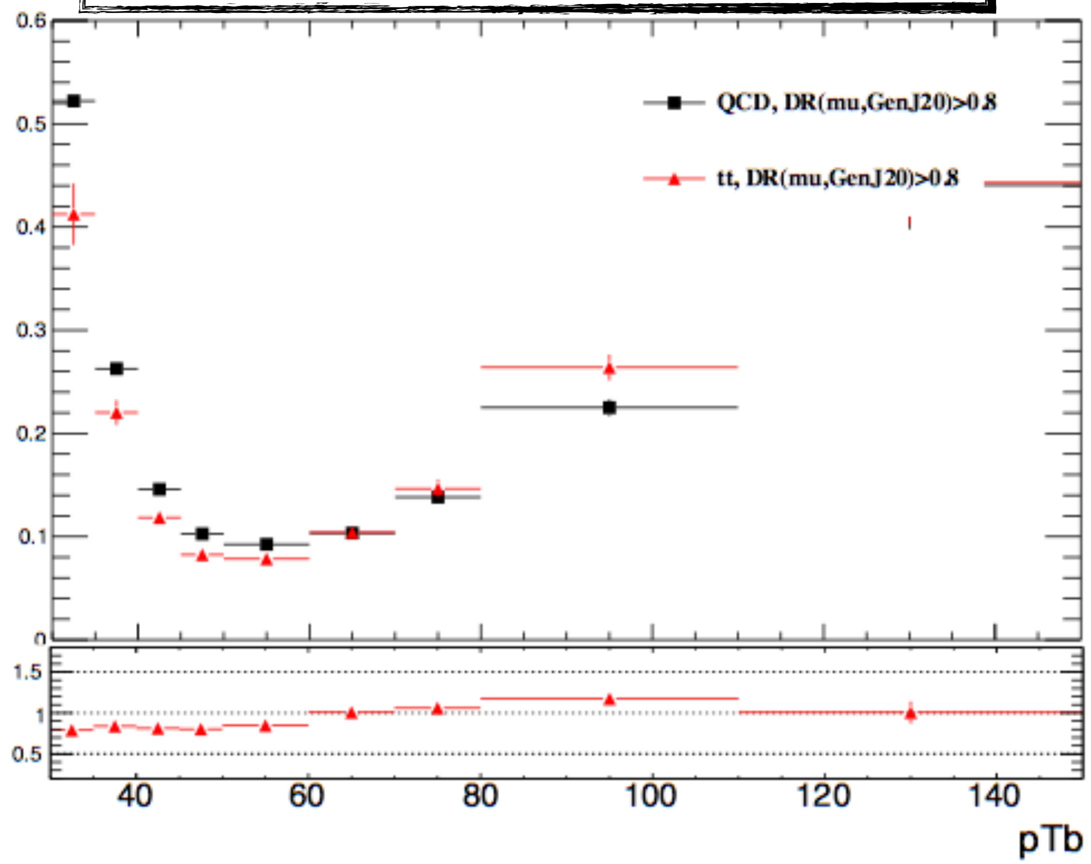
Gen checks: neighboring activity

What about **accidental overlaps with neighboring activity?**

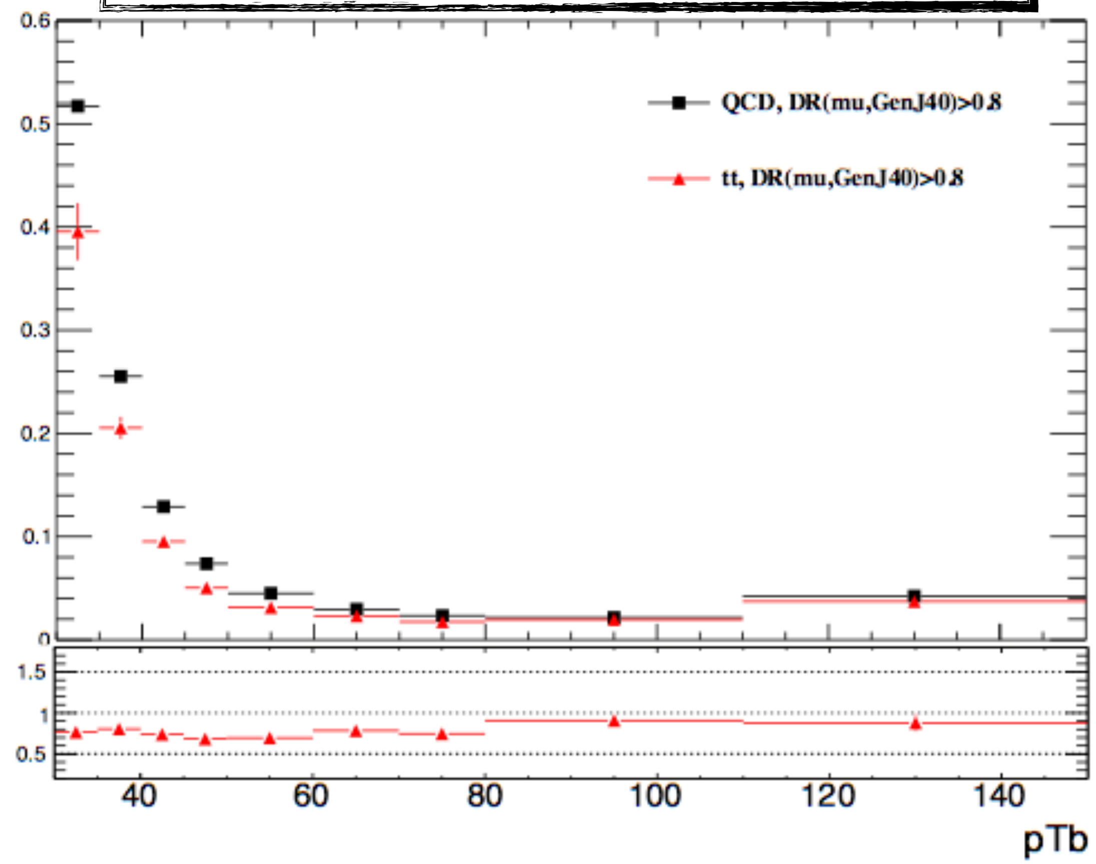
Enforce ΔR(lep, GenJet)>0.8 using different GenJet pT

- ~0.5 (jet cone) + 0.3 (isolation cone)

ΔR (μ, GenJet pT > 20 GeV) > 0.8



ΔR (μ, GenJet pT > 40 GeV) > 0.8



FR at high pTb is strongly affected: agreement improves

Disagreement at low pTb remains

Gen checks: Fragmentation/Decay

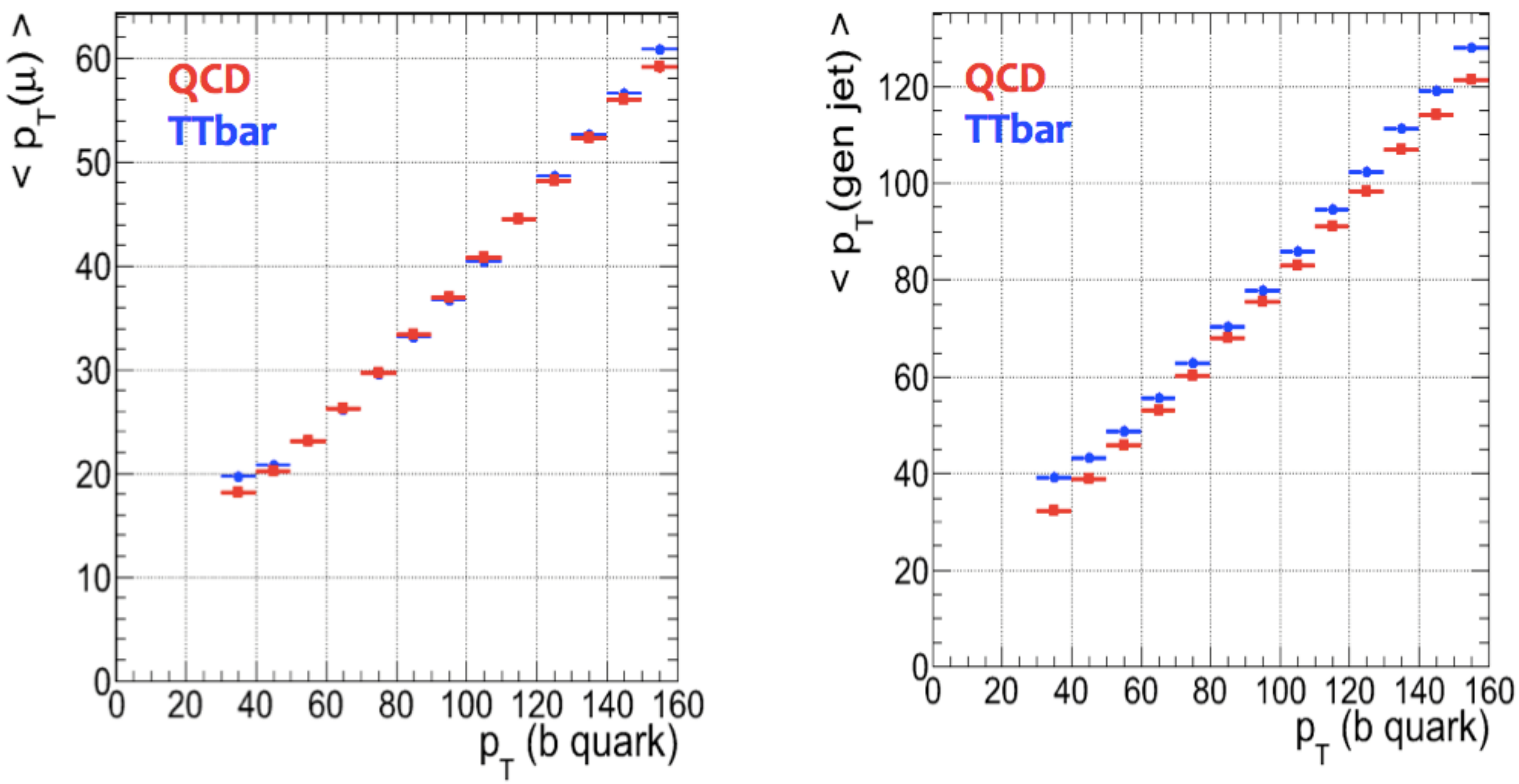
Fragmentation/Decay

mean (reco) lepton pT as a function of b-quark pT

- same in ttbar and QCD

mean GenJet pT as a function of b-quark pT

- harder GenJet in ttbar than QCD for same b-quark pT
- more neighboring status1 particles?
  - consistent with ttbar leptons being less isolated

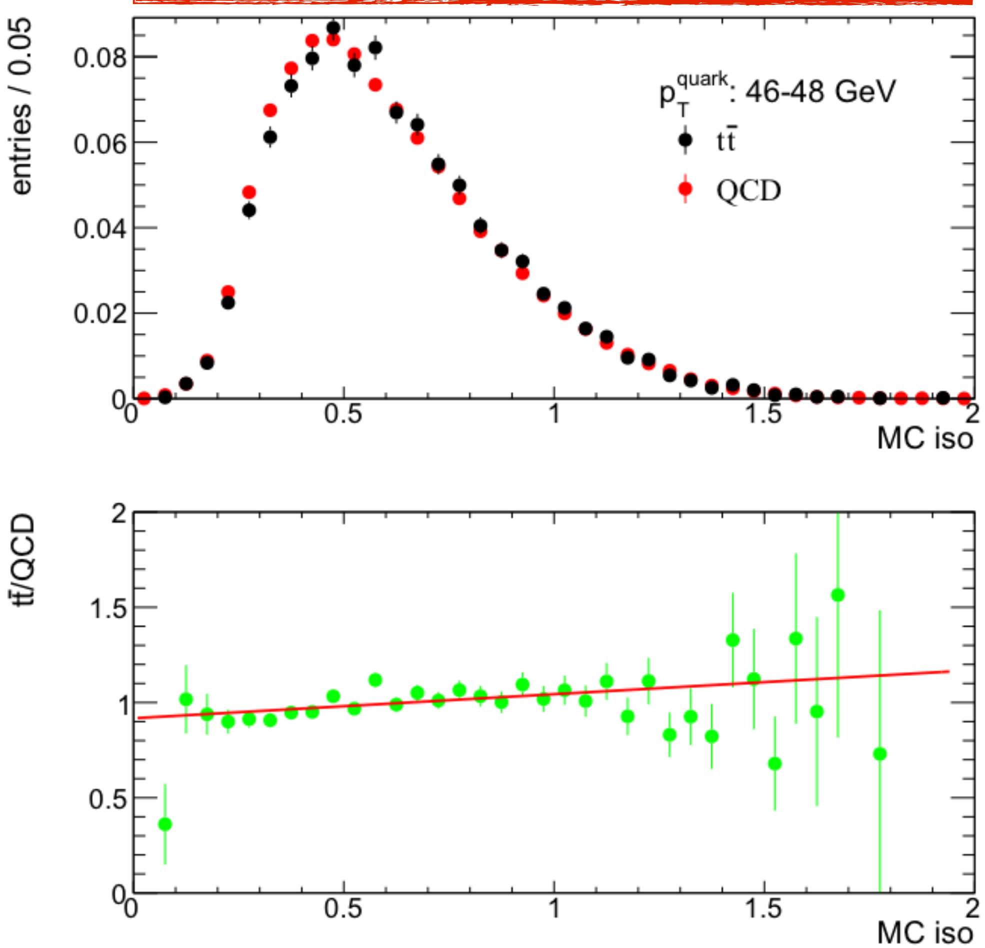


https://indico.cern.ch/event/332064/contribution/0/material/slides/0.pdf

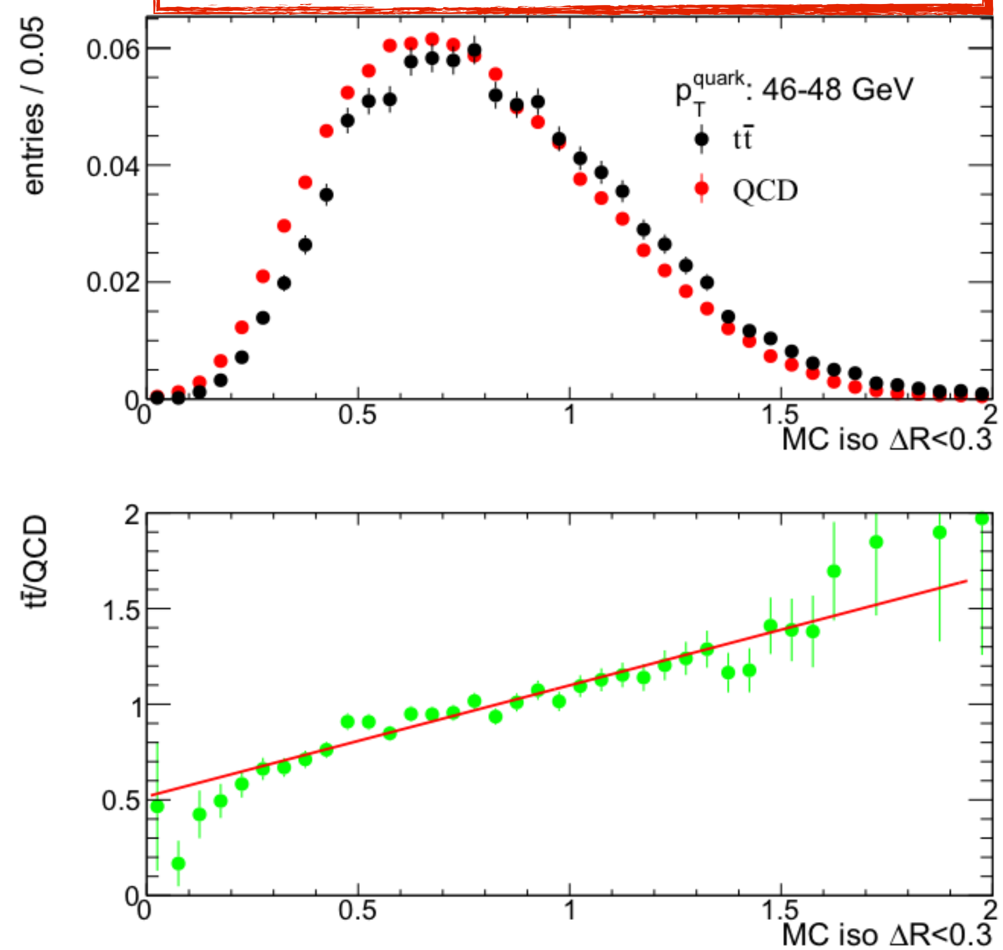
Gen checks: only use the B decay producs

Redefine isolation to only include B decay products

Σ(B decay products in cone): very similar in tt and QCD



Σ(visible status1 in cone): QCD much lower than tt



Effect is dominated by particles other than B-decay products

- Higher multiplicity in ttbar b-jets, leading to differences in isolation and jet pT
- Even using only B product, small discrepancy in tt and QCD

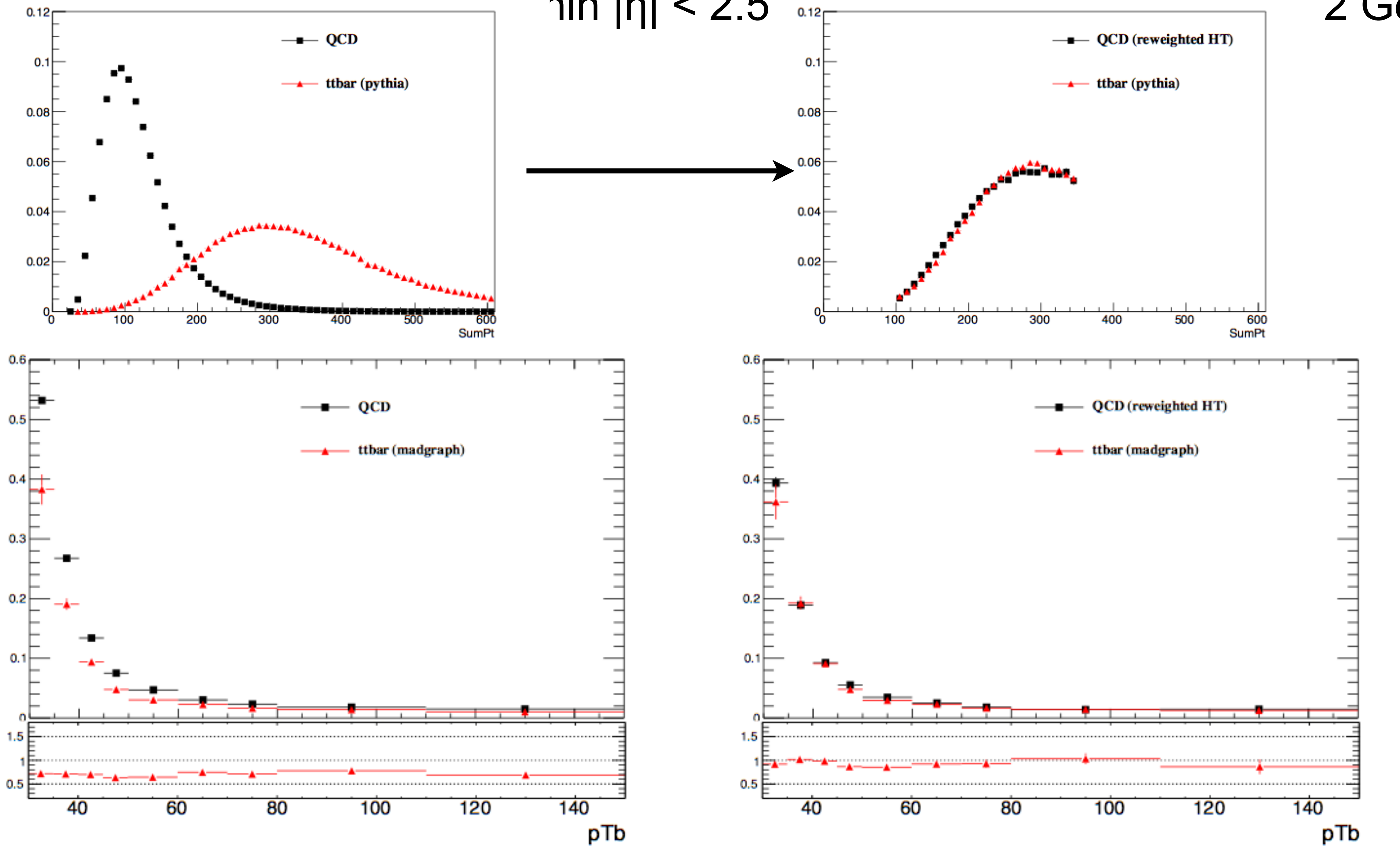
http://lesya.web.cern.ch/lesya/12-16-FR.pdf

Gen checks: SumPt in event

SumPt(status1) plays a large effect even at low pTb

Reweighting SumPt(status1) brings very good agreement at Gen level

- ıin |η| < 2.5 … 2 GeV

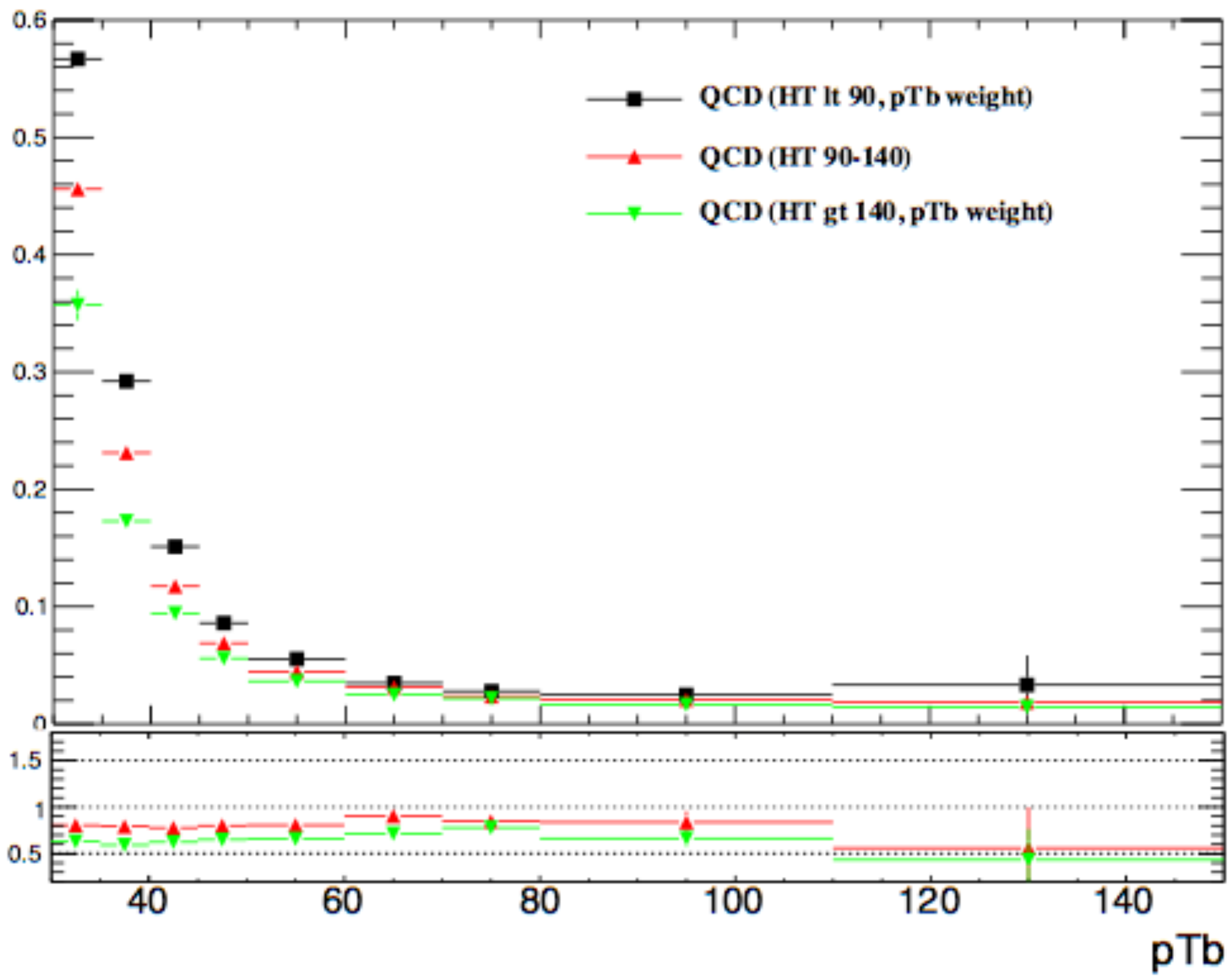


SumPt in QCD sample

Define 3 SumPt regions in QCD

To make sure pTb has small effect, pTb is reweighted to be the same

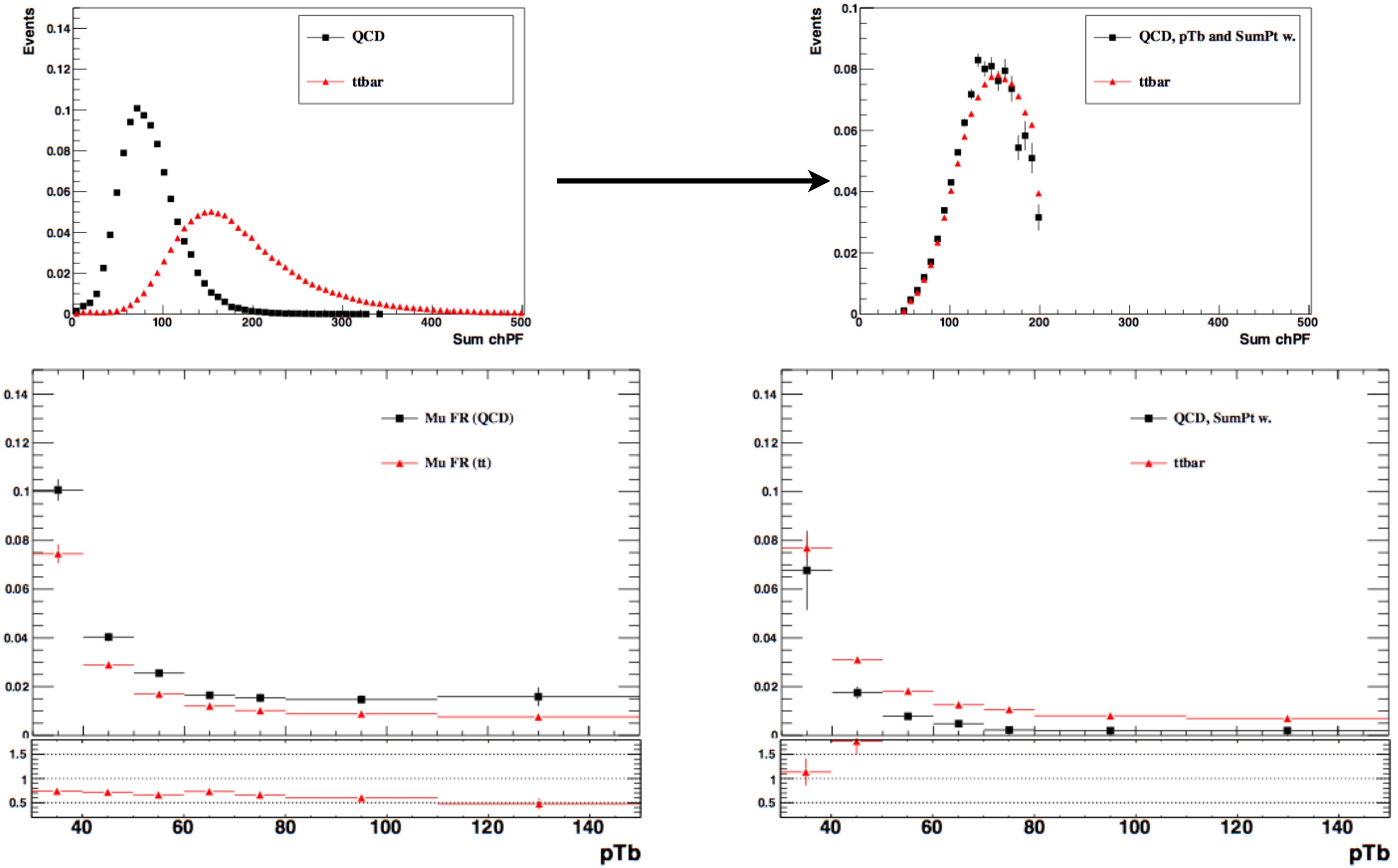
Effect of SumPt is very visible at low pTb



SumPt at Reco level

Reco level: sum charged PF candidates from PV, |η| < 2.5

Simple reweighting leads to large over-correction at Reco level



Non-closure studies: conclusions?

Main reasons for non-closure are understood

Differences in flavor composition and mother parton pT between measurement and application region

- Next section will introduce ideas to reduce these differences

Residual non-closure is more difficult to track down

Reco-level effects play a role

- PU subtraction technique relevant at high pTb

Some discrepancy already visible at Gen level with this Tight/Loose definitions

- Underlying Event has very small effect
- ΔR with closest jets affects high pTb, while HT(GenJets>30GeV) has small effect
- SumPt(status1) spectrum has a large effect, even at low pTb
  - Reweighting SumPt(status1) spectrum improves agreement

Reco level SumPt(chPF) plays a similar role: affects Tight/Loose ratio at low pTb

- Reweighting SumPt(chPF) spectrum is too drastic --> over-correction!

SumPt might become important once first order effects are taken into account.