Fully-Hadronic Search for Standard Model Four-Top Production at CMS

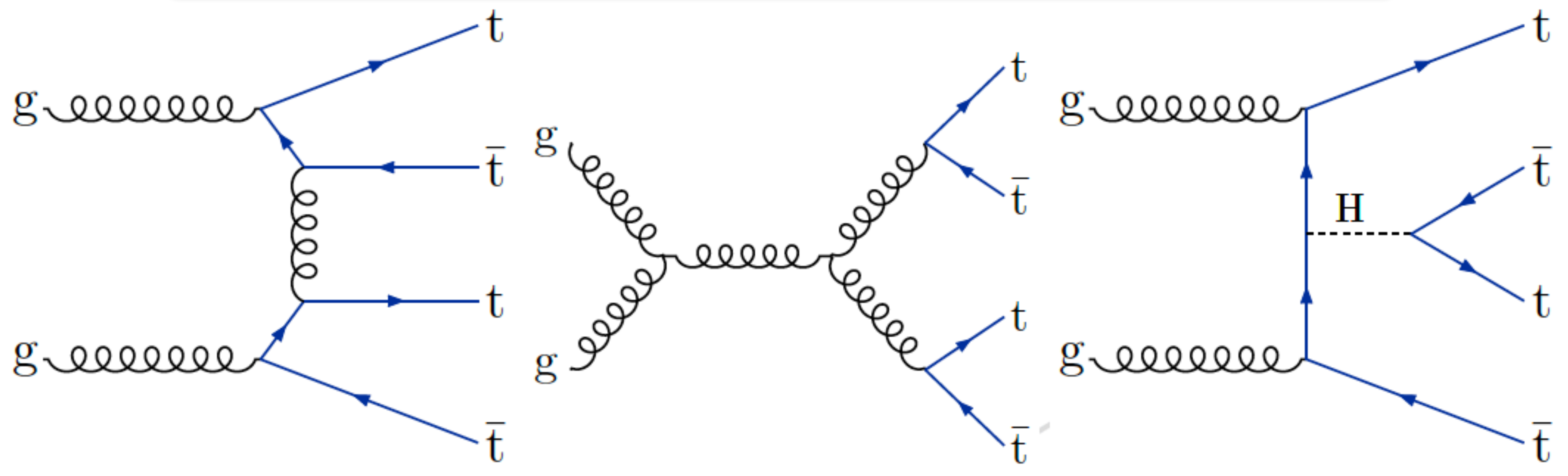
Suyong Choi+, Valentina Dutta*, Joseph Incandela*,
Chang Whan Jung+, Hayoung Oh+,
Melissa Quinnan*, Jae Hyeok Yoo+

New Perspectives 2.0 @ Fermilab
August 24-25 2020

*UCSB, +KU

Four-Top Physics

4-top production in the Standard Model
(gluon fusion)



- Cross-section is ~12 fb in the Standard Model
- Enhancement of cross section predicted by some BSM physics
- Promising probe of top-quark Yukawa coupling

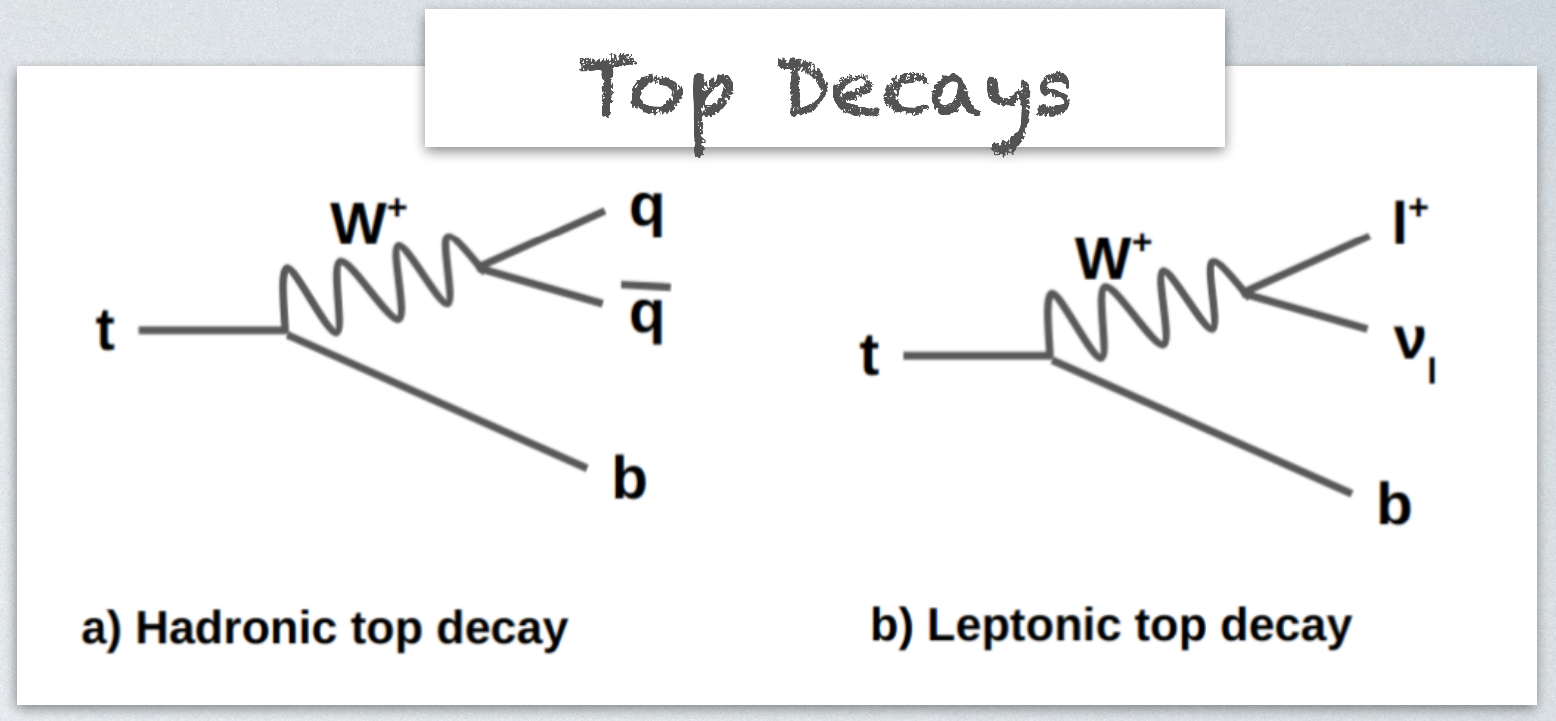
4-top Channels

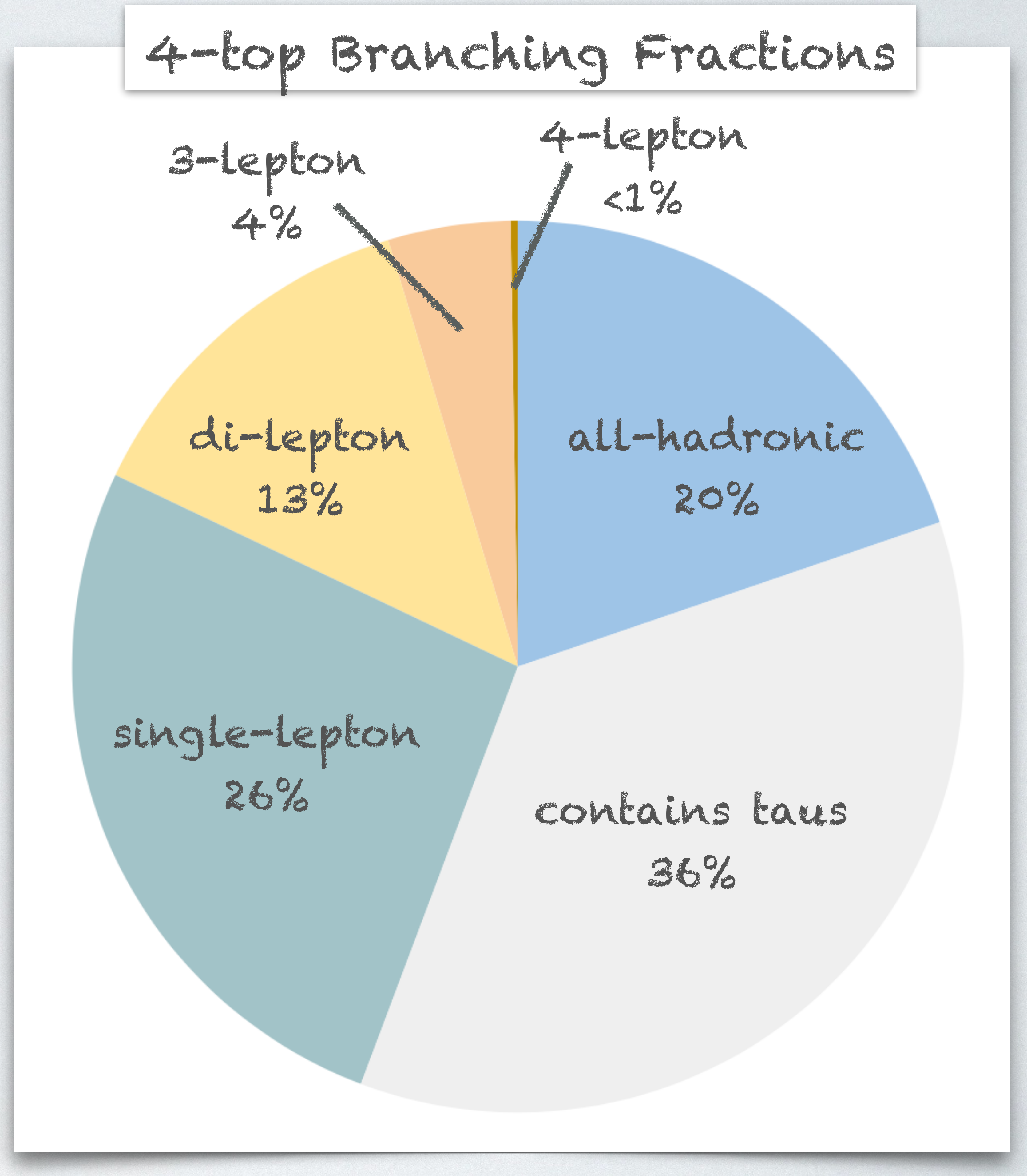
- All-Hadronic Channel
  - Major background(s): QCD + hadronic TT
- Same-Sign DiLepton + MultiLepton Channels
  - **Run II result published** (arXiv:1908.06463)
    - Observed (expected) significance: **2.6(2.7)σ**
- Single-Lepton + Opposite-Sign DiLepton Channels
  - Run II analysis ongoing

Combined CMS Run II Result

- ATLAS Run II Result submitted (arXiv:2007.14858)
- Same-Sign DiLepton + MultiLepton Channels ONLY
  - Observed (expected) significance: **4.3(2.4)σ**

ATLAS Run II Result



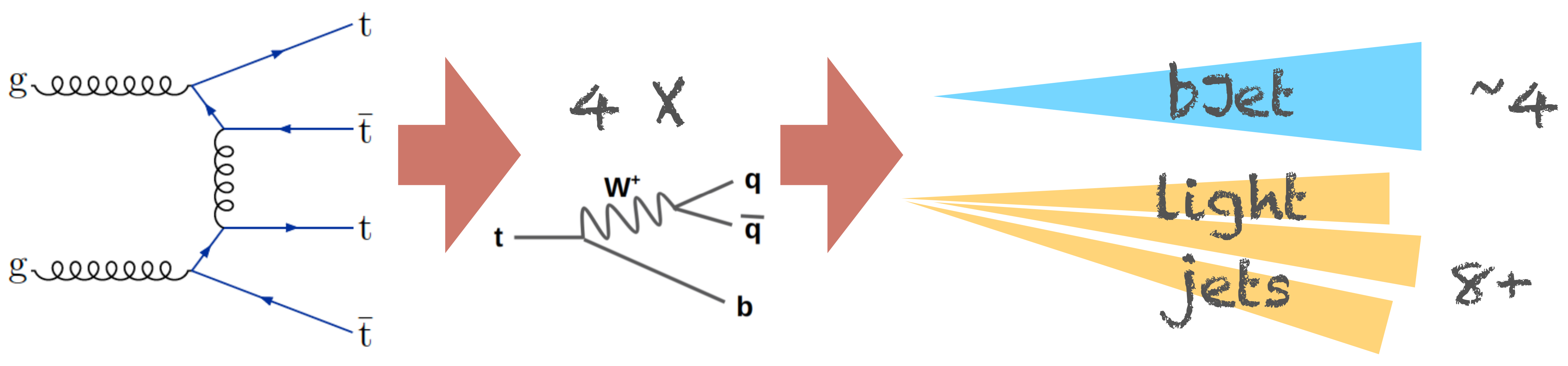


Analysis Strategy

- In order to address QCD and TT backgrounds we need some special techniques…

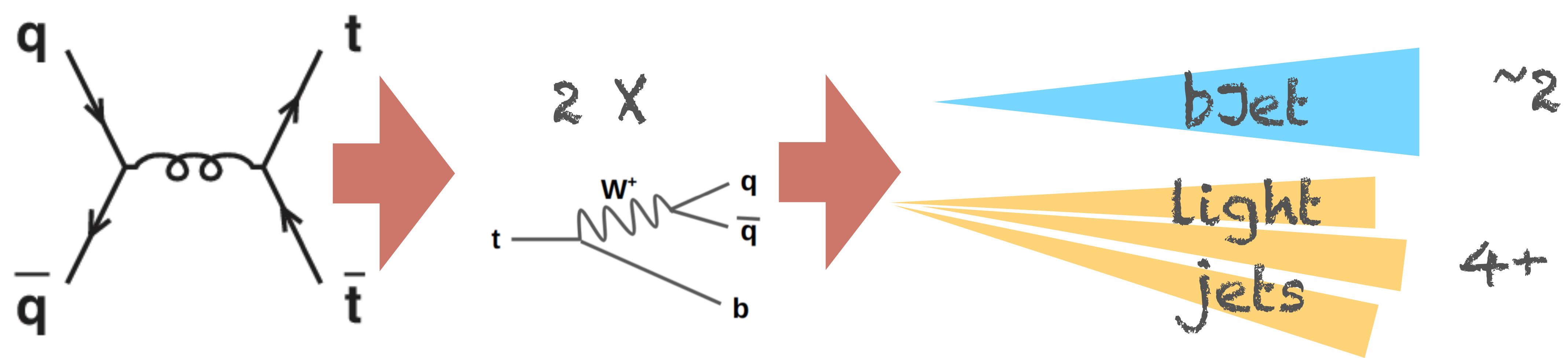
Hadronic TTTT Signal

$\sigma \cong 0.012$ pb



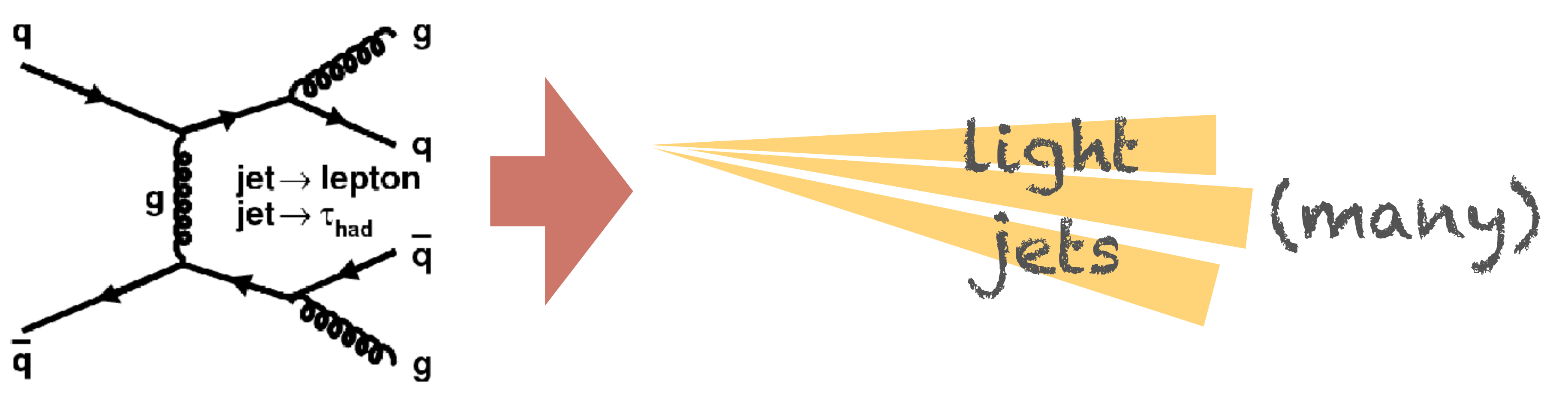
Hadronic TT Background

$\sigma \cong 800$ pb



QCD Background

$\sigma \cong 1$ million pb

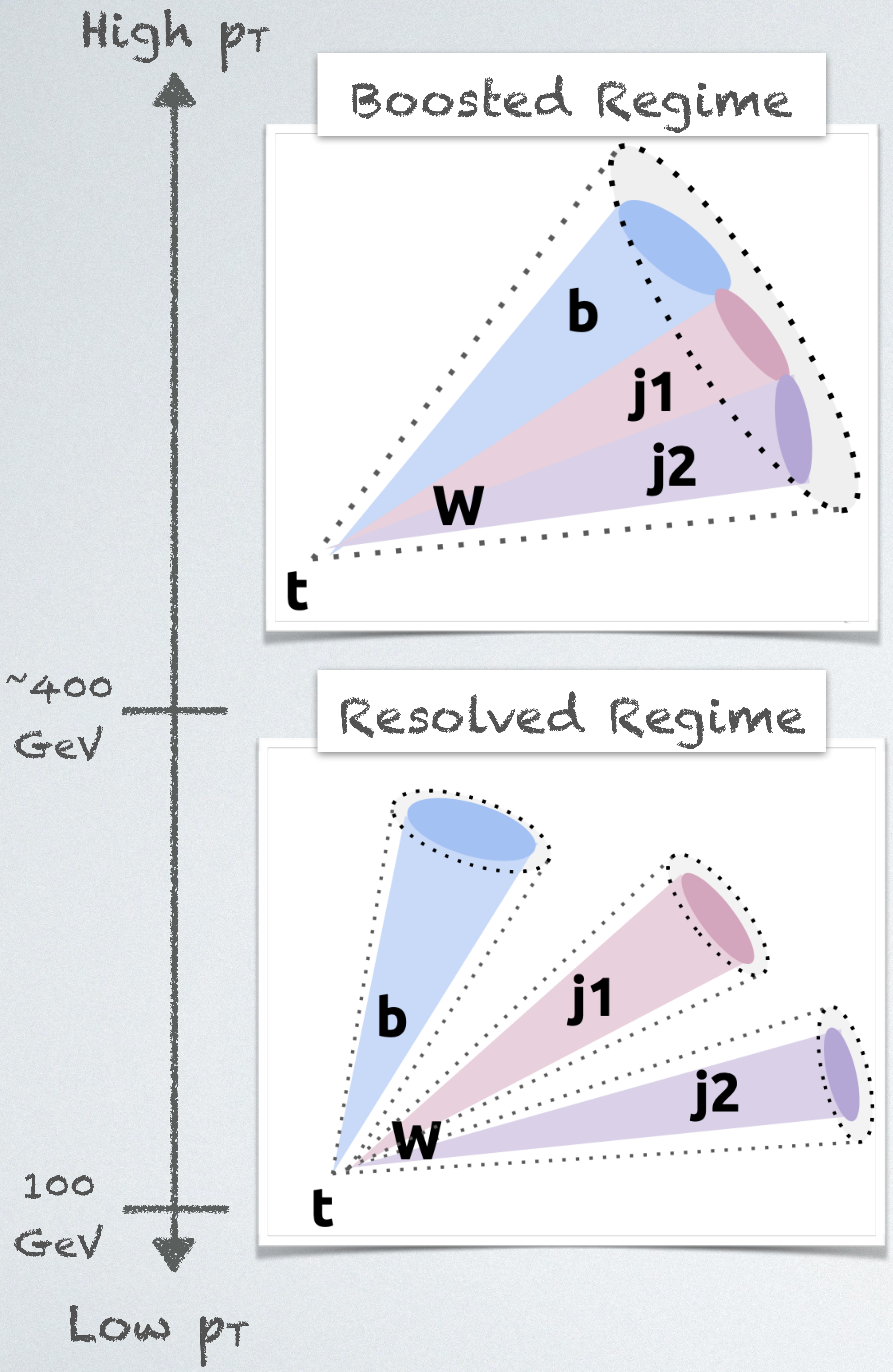


BASELINE:
N(Jets+bJets)≥9, N(bJets)≥3, HT≥700GeV
No leptons

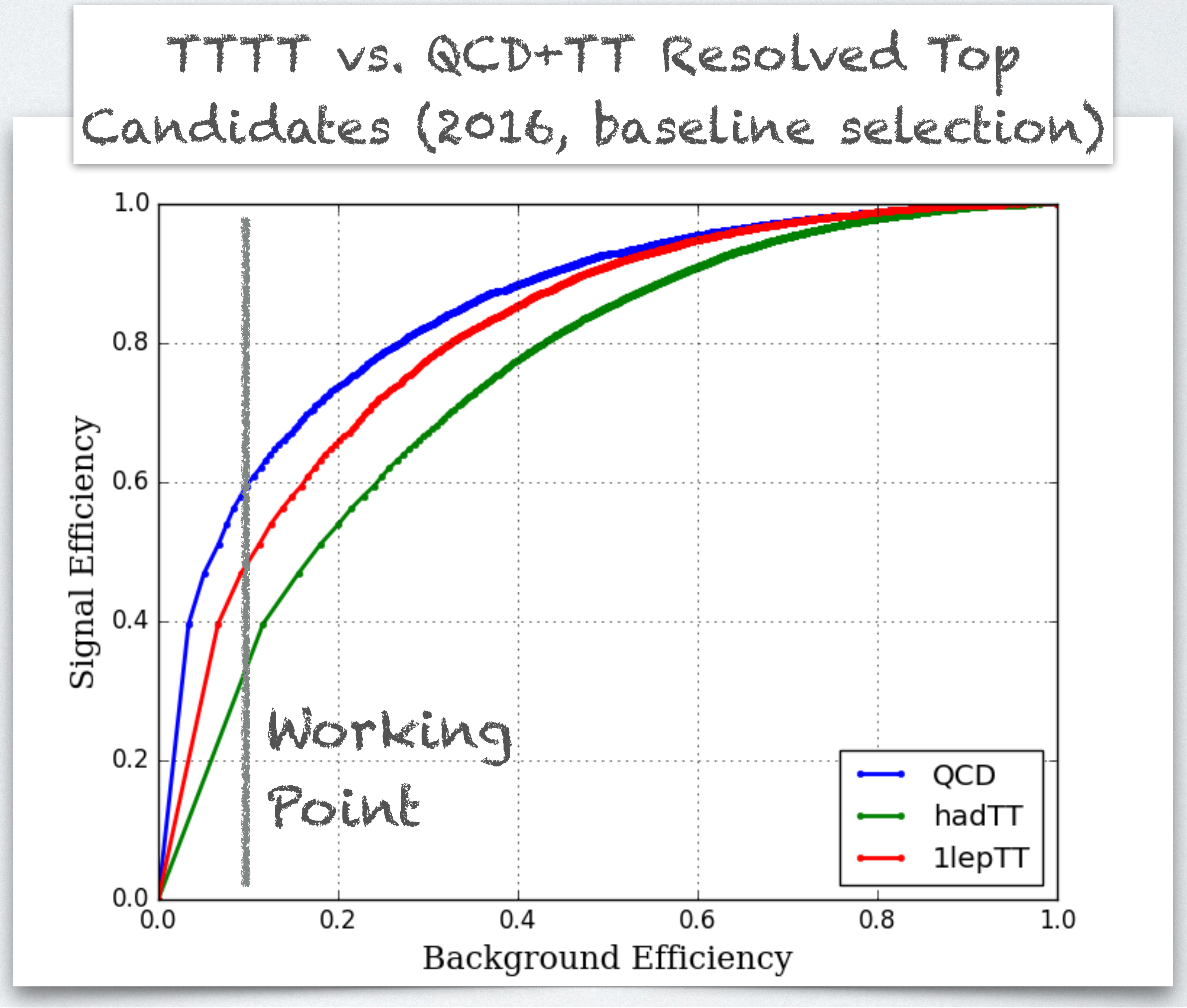
Overall Strategy

1. Start with data and apply baseline.
2. Tag resolved and boosted tops.
3. Evaluate BDT to discriminate 4-top signal vs. TT+QCD backgrounds.
4. Bin signal region in numbers of resolved and boosted tops. Weight bins by the BDT shape.
5. Estimate TT & QCD backgrounds in signal region from data in control regions.

Top Tagging

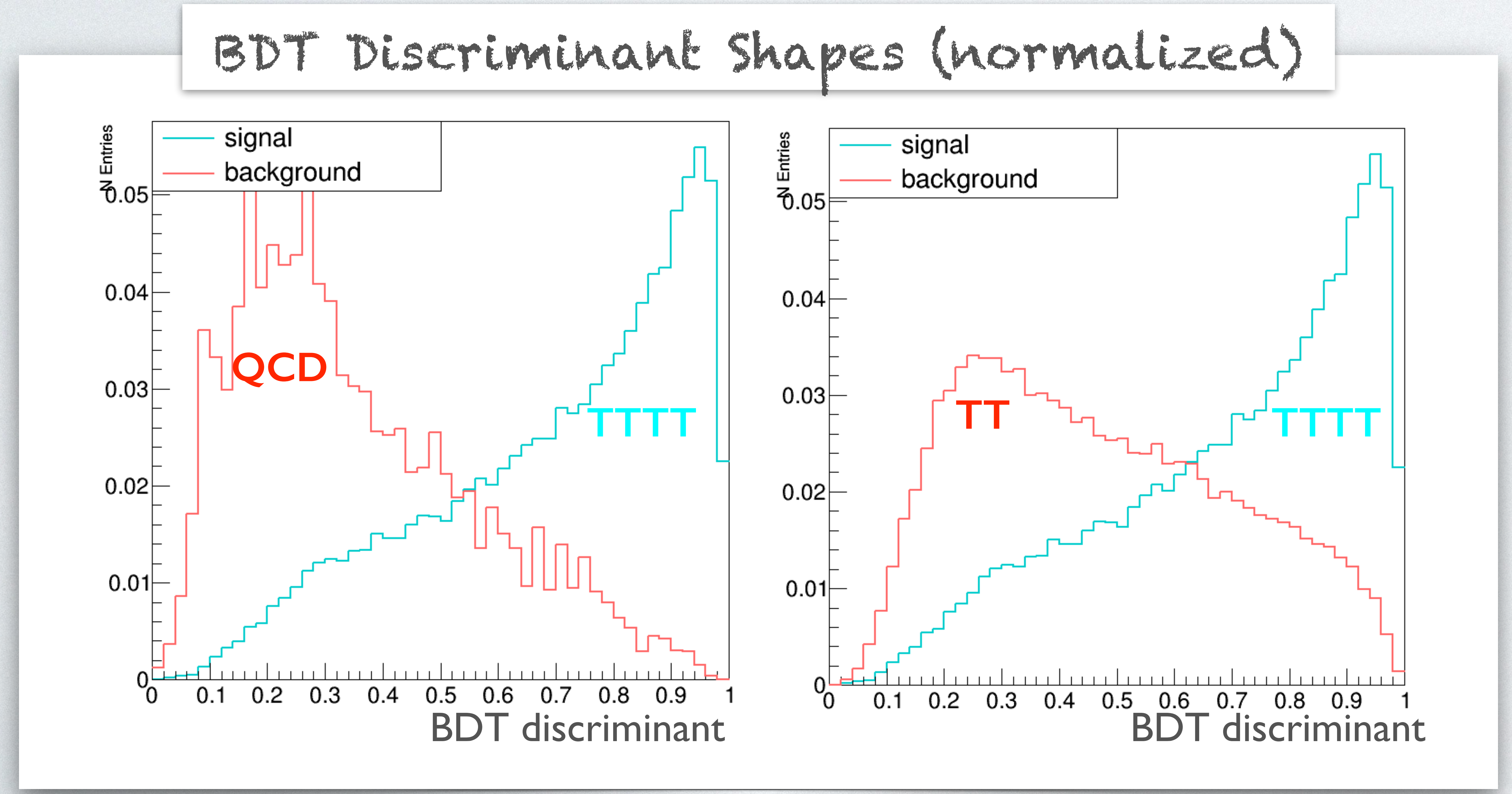


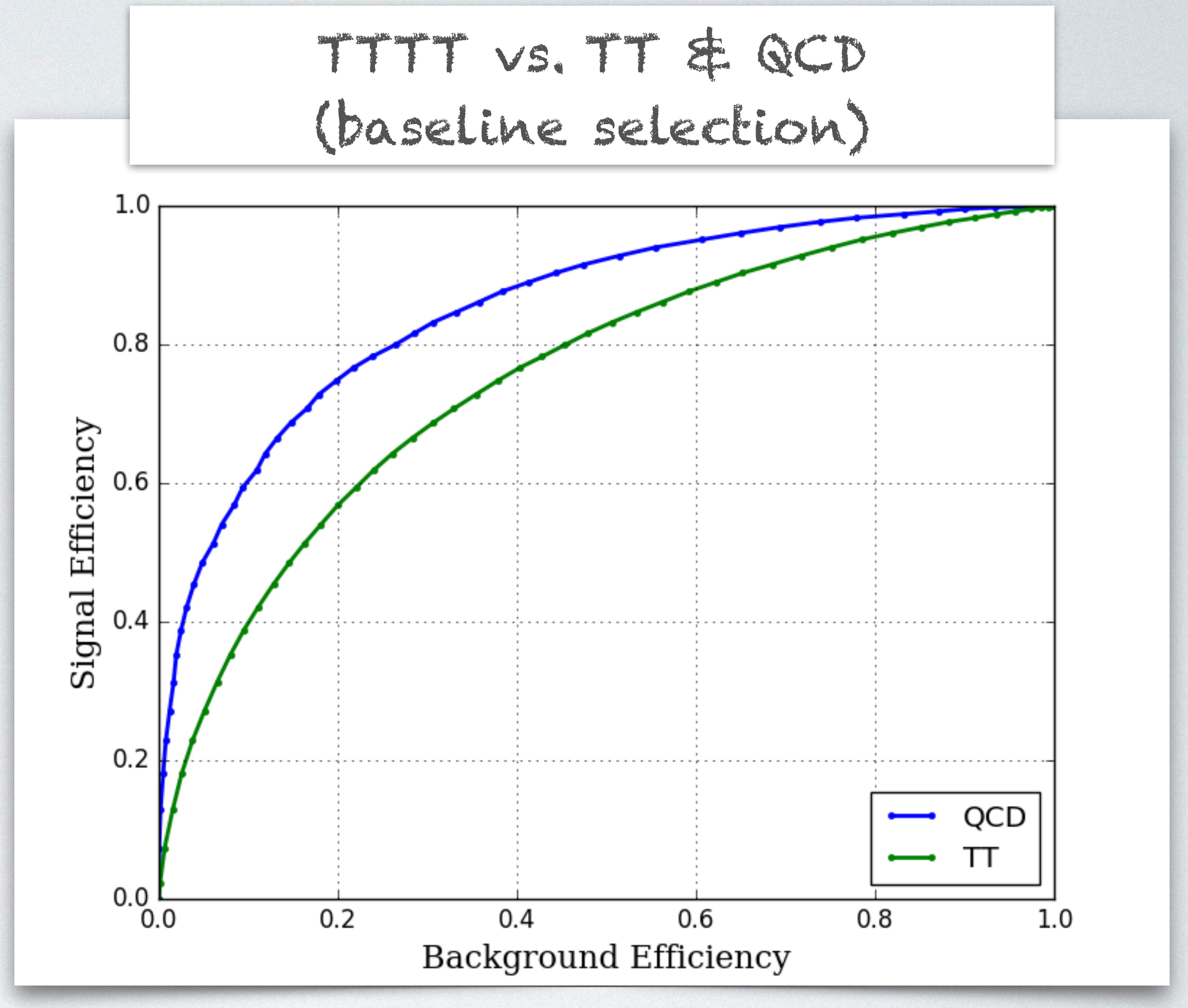
- Boosted top tagger: **DeepAK8** (CMS-JME-18-002)
  - Deep Neural Net (DNN) that inputs low level jet substructure and flavor information.
- Resolved top tagger: **Custom BDT built for this search**
  - Inputs Neural Net (NN) based b and c tagging information and jet kinematics



BDT Targeting QCD & TT Backgrounds

- Event-level BDT trained to discriminate between TTTT signal and TT+QCD backgrounds
- Inputs: jet and resolved top kinematics & angular variables
- Discriminant shape used to weight signal region bins

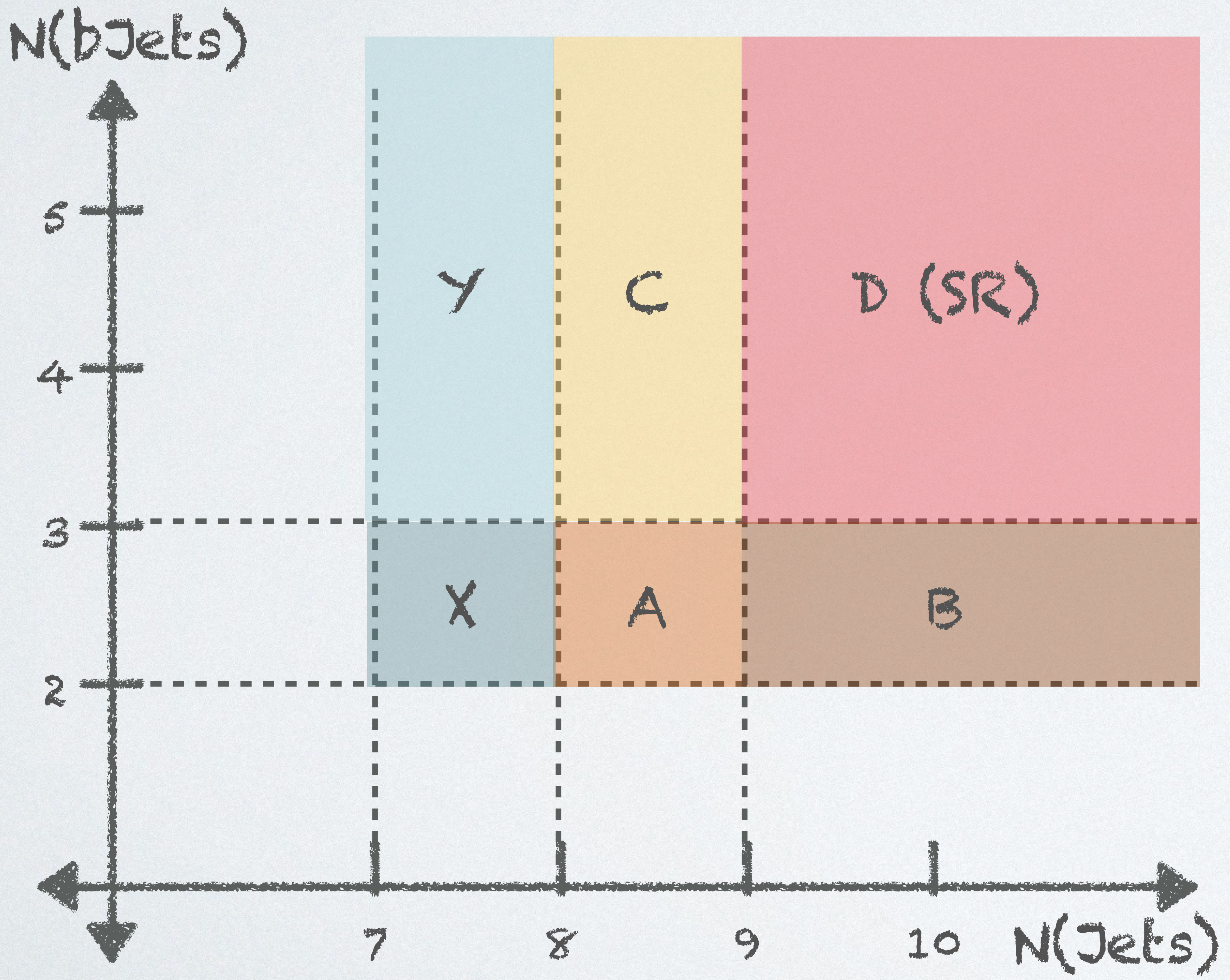


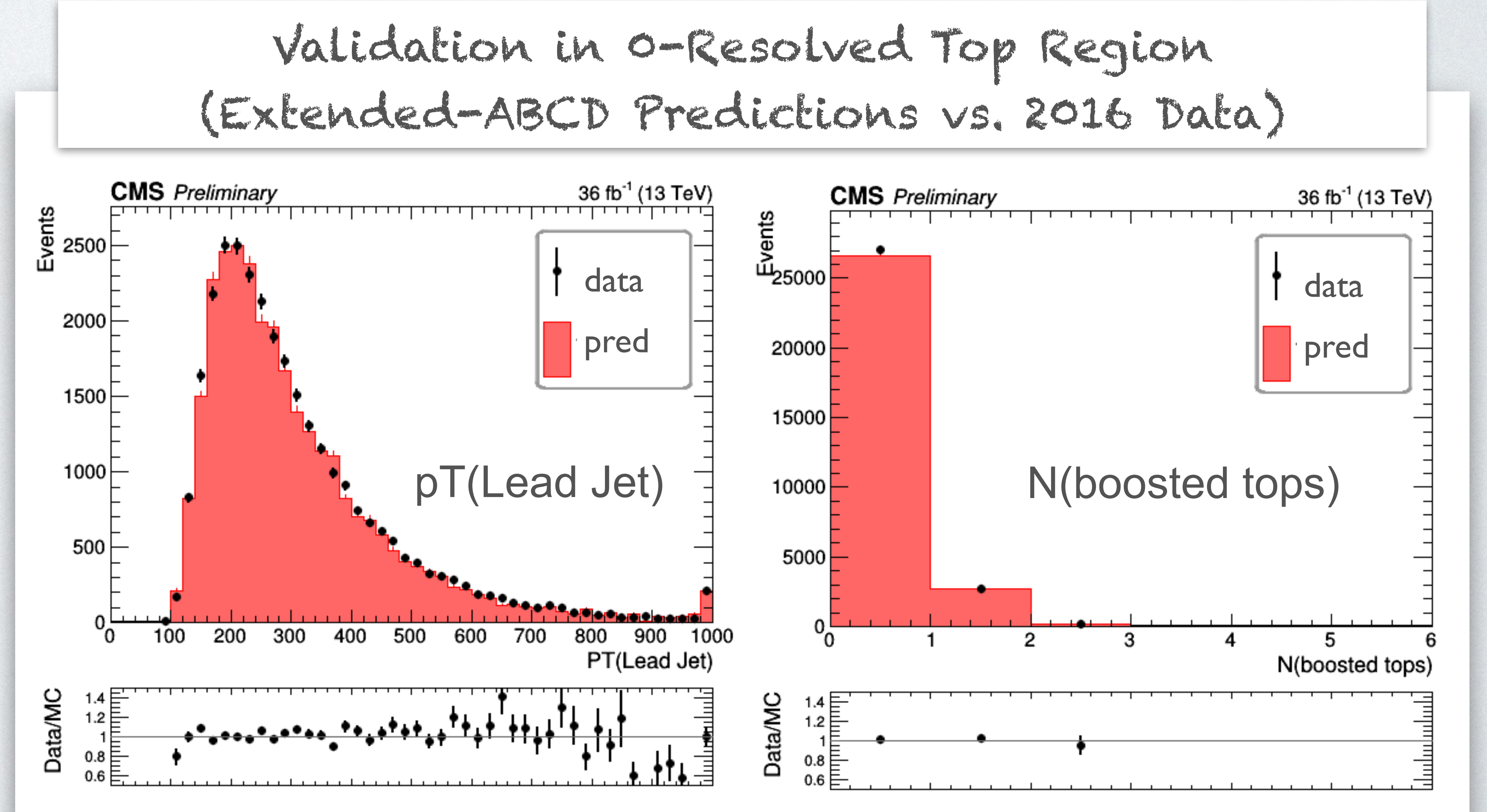


Data-Driven Background Estimates

- **Extended ABCD method**: predict QCD+TT background in signal region (SR) from data in control regions (CRs) (arXiv:1906.10831)

$$D\,(SR) \cong \left(\frac{B\;C}{A}\right) * \left(\frac{B\;C}{A}\right)\left(\frac{B\;Y}{X}\right)^{-1}$$

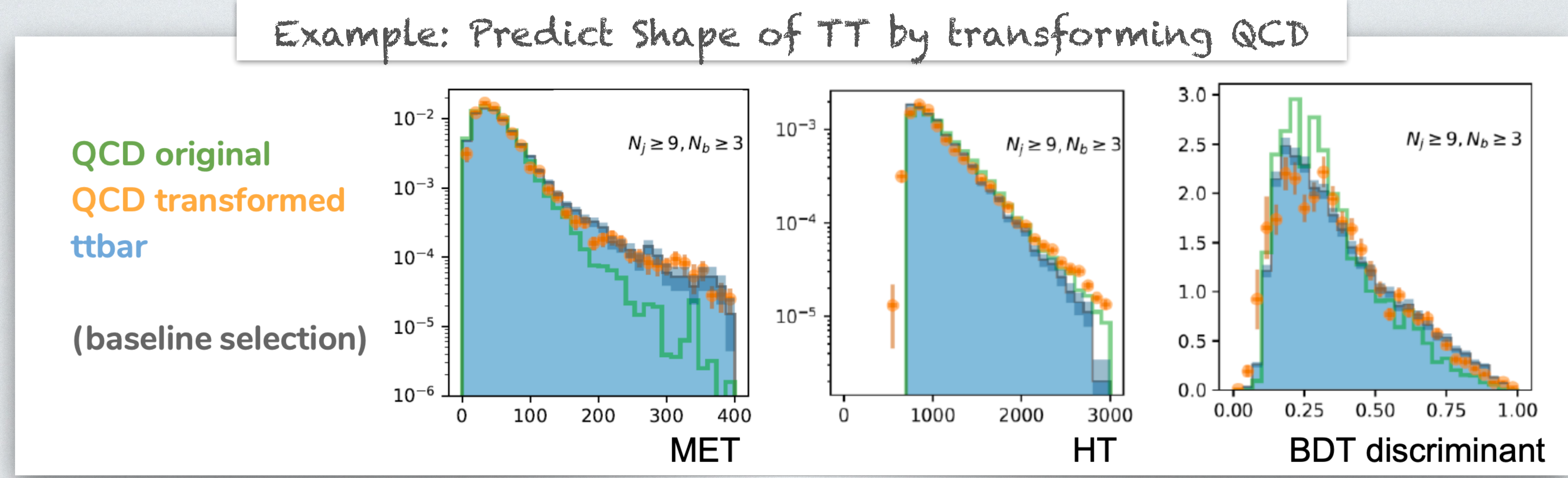




selection: baseline + N(resolved tops=0)

Predicting BDT Shape from Data with Normalizing Flows

- Problem: Extended-ABCD method can fail if variables are correlated, distribution shape changes too fast, or low statistics in one phase space
  - BDT discriminant shape is a case of this!



- Solution: **Normalizing Flows** (arXiv:1804.00779)
- Adversarial neural net that learns transformation from one distribution to another
  - Predict BDT shape in SR by learning MC → data shape differences in CR

$$\mathcal{T}(\vec{x}\,|\vec{x}_0;\vec{c})f_{MC}(\vec{x}_0;\vec{c}) = f_{Data}(\vec{x};\vec{c})$$

doesn't learn physics ("f"), only "T"

Expected Sensitivity

SR Top Bins

		N(resolved tops) (N_{RT})	
		=1	≥2
N(boosted tops) (N_{BT})	=0	$N_{RT} = 1$ $N_{BT} = 0$	$N_{RT} \geq 2$ $N_{BT} = 0$
	≥1	$N_{RT} = 1$ $N_{BT} \geq 1$	$N_{RT} \geq 2$ $N_{BT} \geq 1$

Can further split bins by HT, N(Jets), or other kinematic variables

Channel	Expected Run II Limit x SM σ (139 fb^{-1})
All-Hadronic (preliminary)	~3.0
Single-Lepton *scaled from 2016 result	~3.7
Opposite-Sign DiLepton *scaled from 2016 result	~2.9
Same-Sign Di+Multi Lepton (run II result)	1.6

*re-optimized full Run 2 result should improve on this scaling by luminosity

- Split SR bins by number of boosted and resolved tops and weight by shape of BDT
- Sensitivity estimate based on MC study, considering statistical and rough estimate of systematic uncertainties.
- Adds sensitivity when combined with other channels

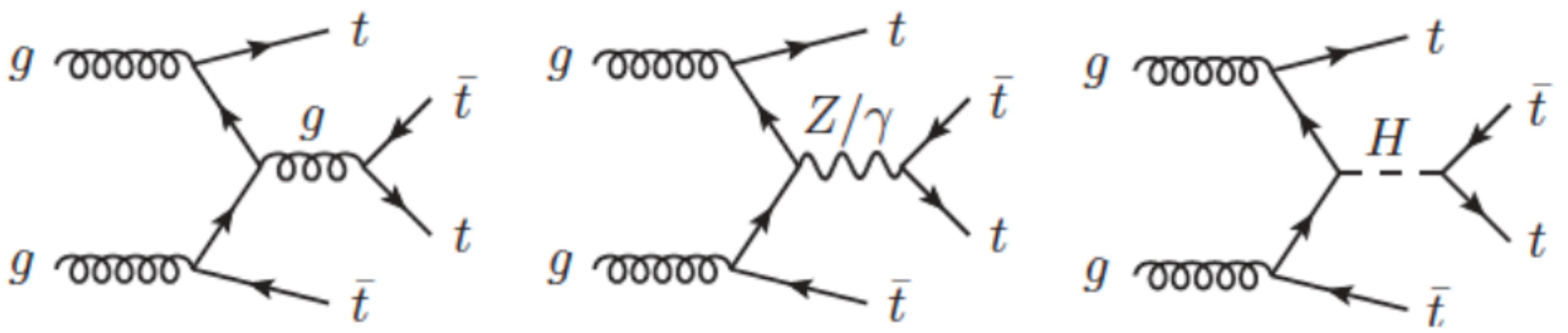
Outlook

- Multifaceted and novel approach to a challenging channel
- Contributing to combined CMS result by end of the year
  - **Very close to "evidence" of 4-top production in CMS!**
- BSM interpretations also of great interest


BACKUP

4-top Cross Section Phenomenology

$$\sigma(t\bar{t}t\bar{t}) = \sigma^{\mathrm{SM}}(t\bar{t}t\bar{t})_{g+Z/\gamma} + \kappa_t^2 \sigma_{\mathrm{int}}^{\mathrm{SM}} + \kappa_t^4 \sigma^{\mathrm{SM}}(t\bar{t}t\bar{t})_H$$



$$\sigma^{\mathrm{SM}}(t\bar{t}t\bar{t})_{g+Z/\gamma} \propto \left|\mathcal{M}_g + \mathcal{M}_{Z/\gamma}\right|^2,$$
$$\sigma^{\mathrm{SM}}(t\bar{t}t\bar{t})_H \propto \left|\mathcal{M}_H\right|^2,$$
$$\sigma^{\mathrm{SM}}(t\bar{t}t\bar{t})_{\mathrm{int}} \propto \mathcal{M}_{g+Z/\gamma}\mathcal{M}_H^\dagger + \mathcal{M}_{g+Z/\gamma}^\dagger \mathcal{M}_H$$

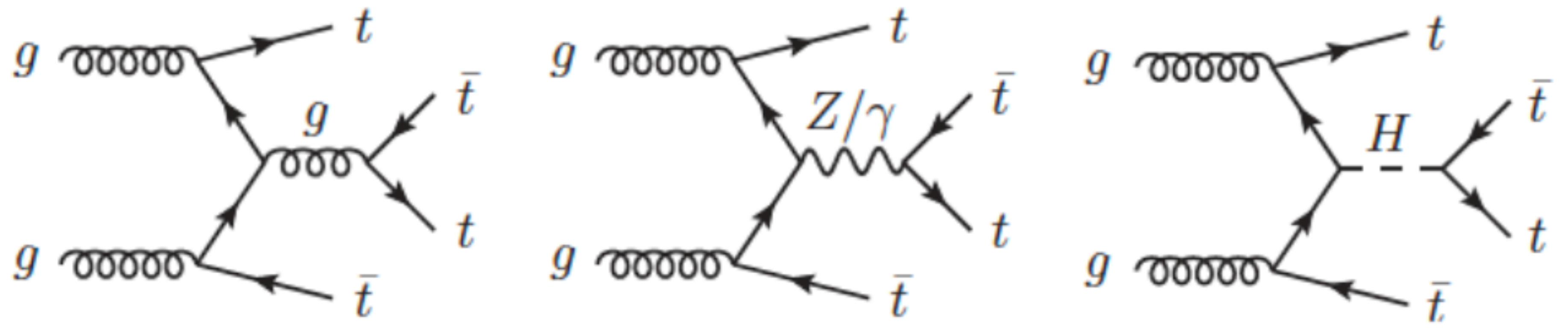
- $k_t = y_T/y^{SM}_T$ where y_T is the top Yukawa coupling, and y^{SM}_T is its value in the SM.
- Strong, electroweak, and higgs terms contribute:
  - σ_{int} is interference between strong+EW and H terms
  - Higgs term proportional to yukawa coupling to 4th power:

$$\sigma(t\bar{t}t\bar{t})_H \propto \kappa_t^4 \sigma^{\mathrm{SM}}(t\bar{t}t\bar{t})_H$$

  - Sensitive to top yukawa coupling, but independent of Higgs decay
    - Deviations from standard model predictions (k_t not =1) would be red flag for BSM physics
- More information: arXiv:1602.01934 and arXiv:1908.06463, Giovanni Zevi Della Porta's presentation

4-top Cross Section Phenomenology

$$\sigma(t\bar{t}t\bar{t}) = \sigma^{\rm SM}(t\bar{t}t\bar{t})_{g+Z/\gamma} + \kappa_t^2 \sigma_{\rm int}^{\rm SM} + \kappa_t^4 \sigma^{\rm SM}(t\bar{t}t\bar{t})_H$$



$$\begin{aligned}
\sigma^{\rm SM}(t\bar{t}t\bar{t})_{g+Z/\gamma} &\propto \left|\mathcal{M}_g + \mathcal{M}_{Z/\gamma}\right|^2, \\
\sigma^{\rm SM}(t\bar{t}t\bar{t})_H &\propto |\mathcal{M}_H|^2, \\
\sigma^{\rm SM}(t\bar{t}t\bar{t})_{\rm int} &\propto \mathcal{M}_{g+Z/\gamma}\mathcal{M}_H^\dagger + \mathcal{M}_{g+Z/\gamma}^\dagger \mathcal{M}_H
\end{aligned}$$

Leading Order Cross sections @ 13 TeV

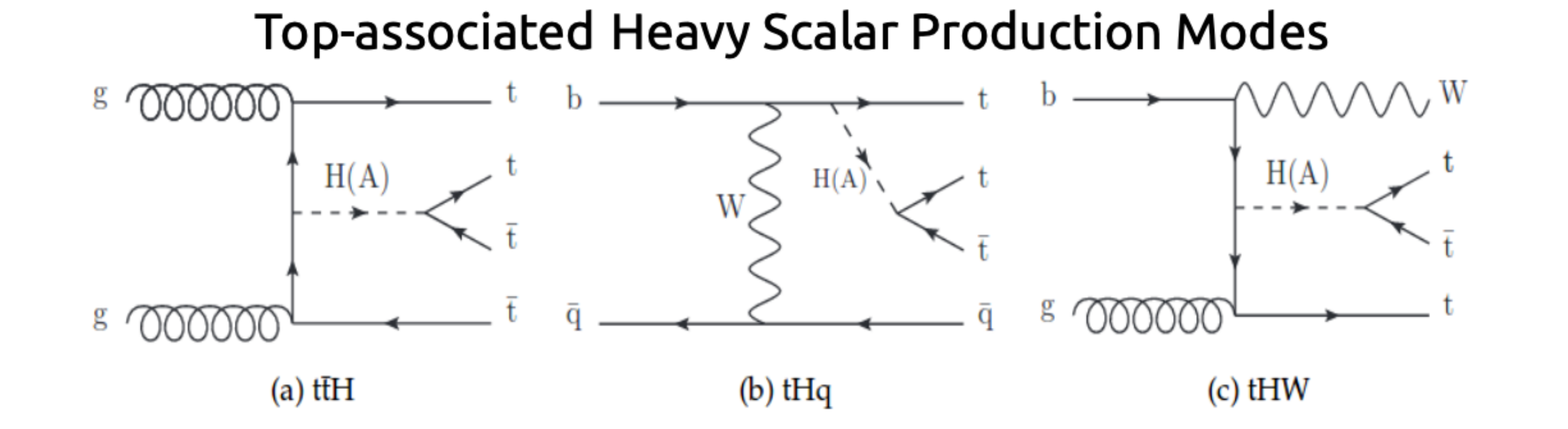
$$\begin{aligned}
\sigma^{\rm SM}(t\bar{t}t\bar{t})_{g+Z/\gamma} &: \; 9.997 \text{ fb}, \\
\sigma^{\rm SM}(t\bar{t}t\bar{t})_H &: \; 1.168 \text{ fb}, \\
\sigma^{\rm SM}(t\bar{t}t\bar{t})_{\rm int} &: \; -1.547 \text{ fb}, \\
\text{Total} &: \; 9.6 \text{ fb*}
\end{aligned}$$

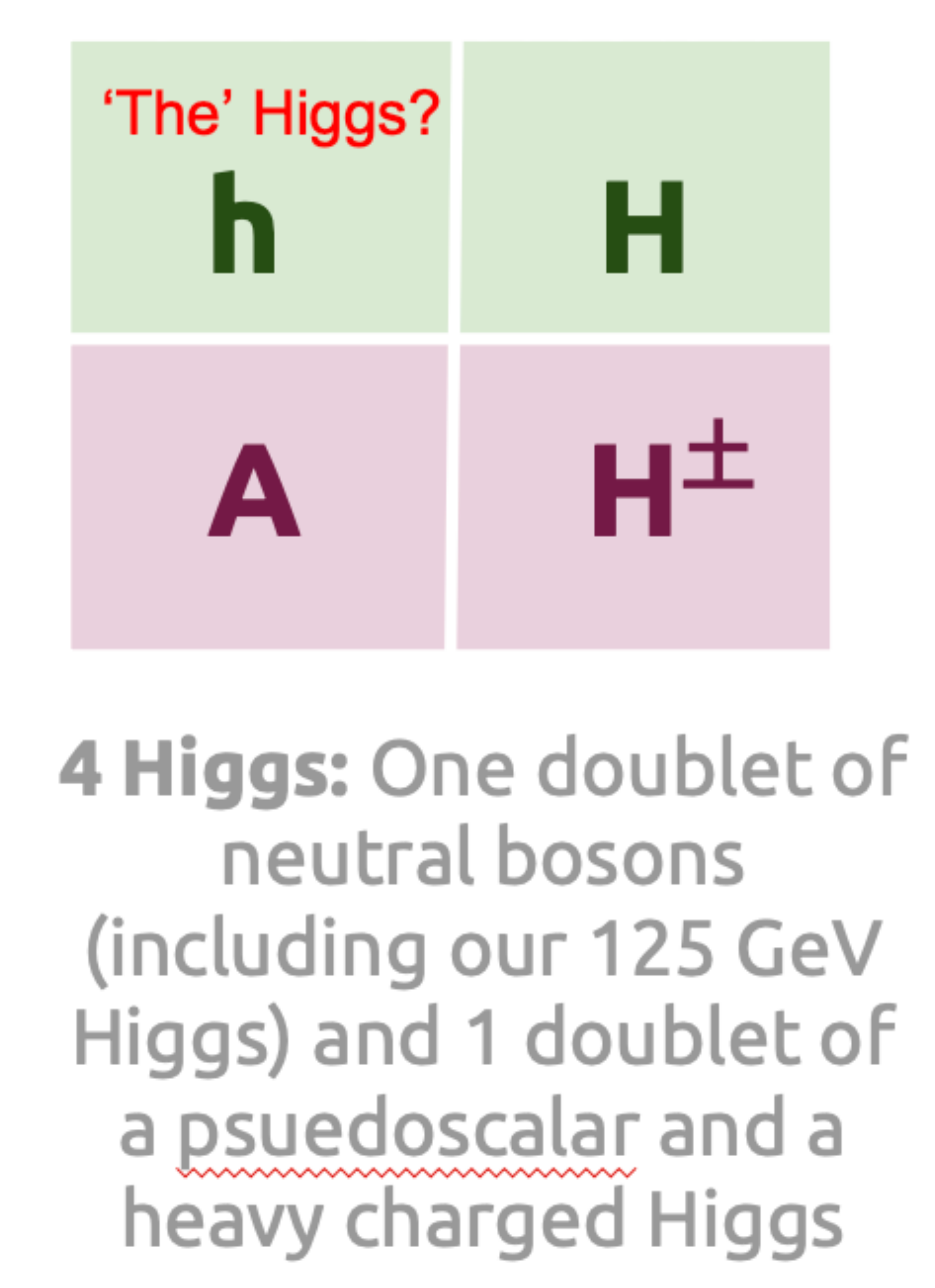
*Note: cross section on slide 2 includes NLO calculations, which is why it doesn't match the "Total" value. Detailed info here: arXiv:1711.02116

LO strong + LO weak + LO Higgs = 9.6 fb
LO+NLO strong + LO weak + LO Higgs = 9.2 fb
LO+NLO strong + LO+NLO weak + LO Higgs = 12.0 fb

2-Higgs-Doublet-Models (2HDMs)

- Simple extension of the Standard Model (SM) with 2 doublets of Higgs-like bosons rather than one
- A heavy scalar or psuedo-scalar boson decaying to tops is produced in association with 1 or 2 tops
- Signal is 4 top quarks (ttH/A) or 3 top quarks (tHq, tHW)
- Leads to enhanced tttt cross section signature
- Key ingredient in minimal supersymmetric standard model (MSSM)
- More information: arXiv:1106.0034

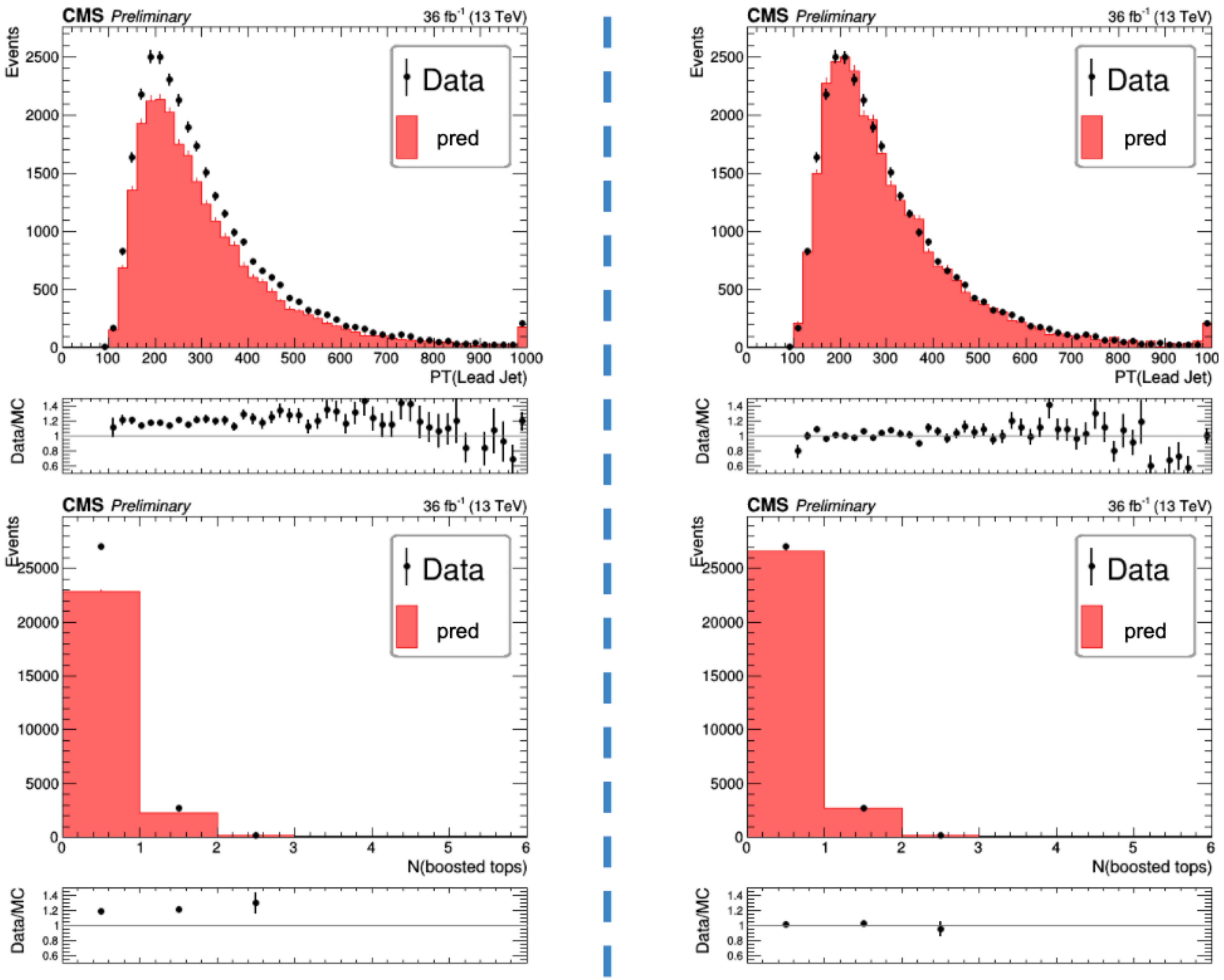




ABCD vs. Extended ABCD

ABCD method: $D \approx BC/A$

ext-ABCD method: $D \approx (BC/A)^2(A'/BC')$



Normalizing Flows

- More Details (arXiv:2008.03636v1)

- Through normalizing flows (NF), feature variables are transformed through multidimensional invertible bijections.
  - This transformation is implemented as a NN and learned during training.
  - We use the neural autoregressive flow (NAF), to construct a multidimensional invertible function

- More appealing and acceptable than GAN, since only transformation learned and it doesn't try to reproduce the underlying physics.

- Can predict different variables simultaneously

Example: Predict Shape of TT by transforming QCD