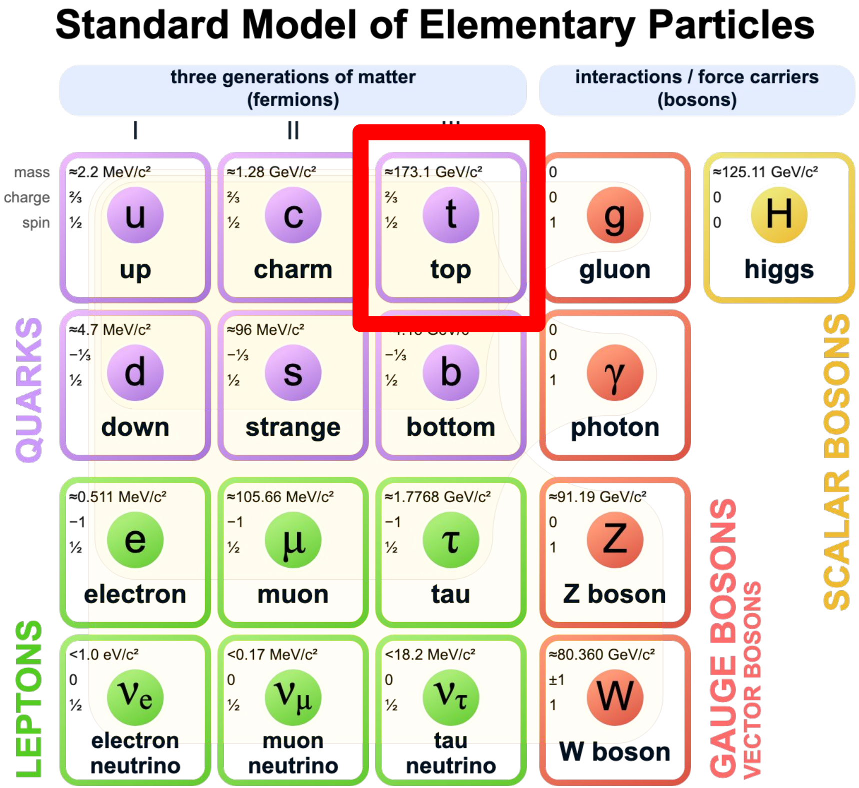
CP violation and tqg FCNC in single top quark t-channel

Byeonghak Ko

(University of Seoul)

Top quark

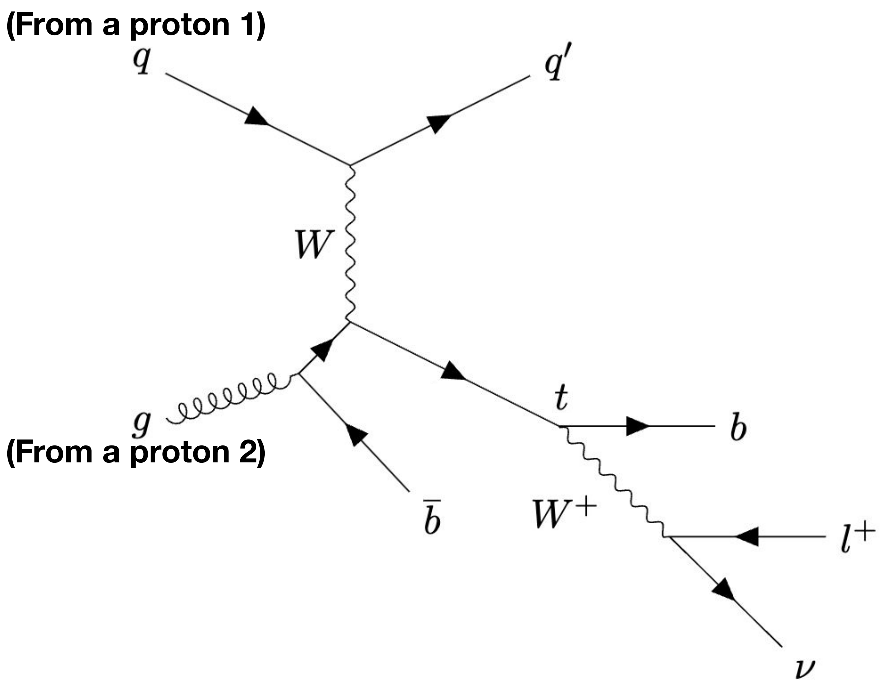


(From https://en.wikipedia.org/wiki/Standard_Model)

- Top quark: One of the fundamental particles in Standard Model
  - Discovered in 1995 (CDF, D0)
- The heaviest elementary particle in Standard Model
- Top quark physics is one of most active field in particle physics
  - Top quark properties (e.g., mass) are sensitive to high energy phenomena, such as Higgs physics and Beyond Standard Model (BSM)
  - Many studies of search for BSM relate to top quark sector

Top quark doll, in sale!

Single top quark t-channel process

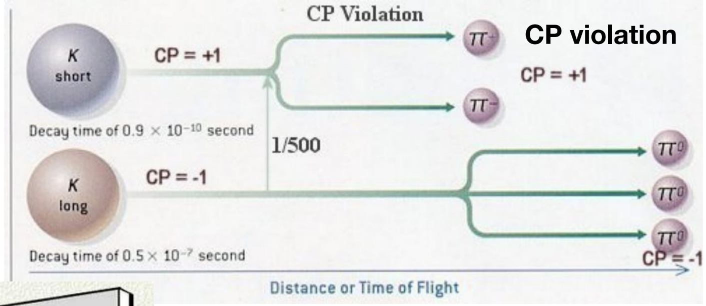


- Single top quark t-channel process: One of process producing only one top quark
  - The others: tW, s-channel
- From the collision, one top quark and a quark is produced
  - The top quark decays immediately to bottom quark and W boson
  - The W boson decays to a pair of quark (**hadronic decay**) or lepton+neutrino (**leptonic decay**)
  - In my studies, only the case of leptonic decay is considered
    - Hadronic decay mode is hard to be extracted from backgrounds

CP violation

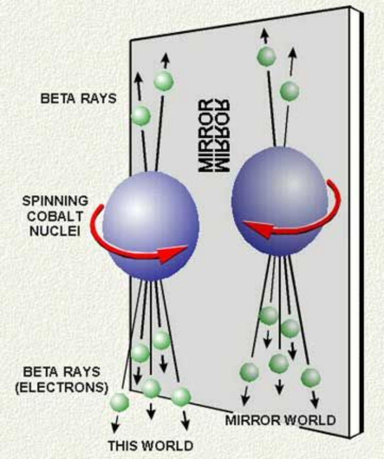
About CP violation

CP violation



https://www.physi.uni-heidelberg.de/~reygers/seminars/2015/nobel_prizes_in_particle_physics/talks/paul_cp_violation.pdf

Parity violation



https://www.aps.org/publications/apsnews/200112/history.cfm

- In 1956, Wu has found a **parity violation** in an experiment (Co-60 experiment [1])
    - Weak interaction occurs only with left-handed fermions
- In 1964, Cronin and Fitch have found a **CP violation** in an experiment of neutral Kaon decay [2]

[1] Wu, C. S.; Ambler, E.; Hayward, R. W.; Hoppes, D. D.; Hudson, R. P. (1957). "Experimental Test of Parity Conservation in Beta Decay". *Physical Review*. 105 (4): 1413–1415

[2] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, "Evidence for the 2π Decay of the K20 Meson," Phys. Rev. Lett. 13, 138 (1964).

Why CP violation?

https://www.quantumdiaries.org/2011/08/28/why-b-physics/

- Since the discovery of antimatter (1932, C. D. Anderson), this question follows physicists:

  Why does the observable universe have more matter than antimatter?

- Sakharov suggested Sakharov conditions to solve this puzzle, which requires significant violation of CP symmetry
- Since its discovery, but still with a too few fraction to explain the puzzle, physicists are looking for more source of CP violation

Introduction to CP violation in top decay

Effective Lagrangian of tWb vertex:

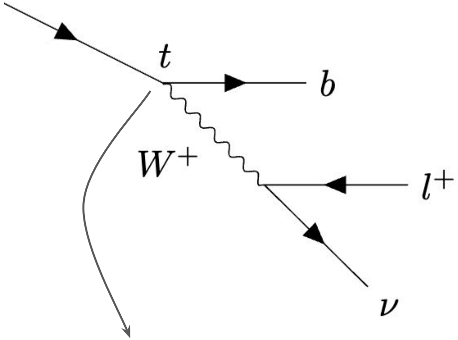
$$\mathscr{L}_{tWb} = -\frac{g}{\sqrt{2}}\bar{b}\gamma^{\mu}(V_L P_L + V_R P_R)tW_{\mu}^{-}$$
$$-\frac{g}{\sqrt{2}}\bar{b}\frac{i\sigma^{\mu\nu}q_{\nu}}{M_W}(g_L P_L + \underline{g_R P_R)tW_{\mu}^{-}} + \text{h.c.},$$

There might be an anomalous coupling g_R, of which the complex phase makes the term **CP violating**

(See arXiv:1005.5382)

In SM: $g_R = (-7.17-1.23i)\times10^{-3}$

The tWb vertex of top decay



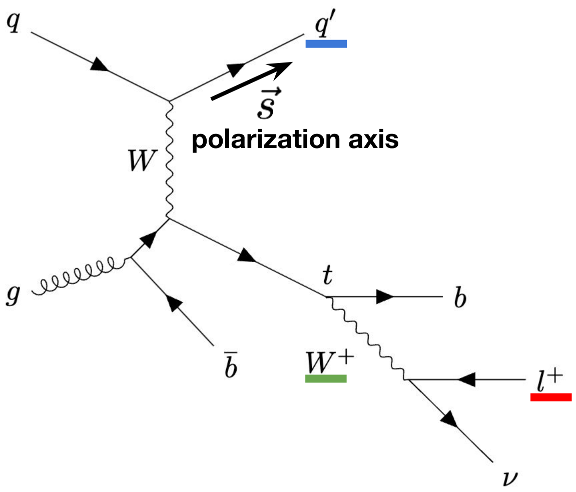
The strength of CP violating term in the tWb vertex can be estimated by measuring a forward-backward asymmetry (A^N_{FB})

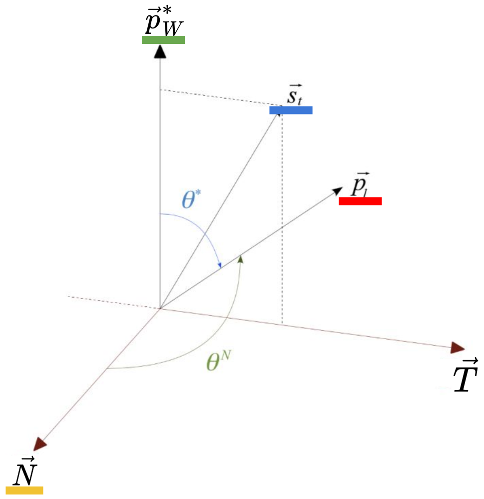
$$\boxed{\begin{array}{c} A^N_{FB} = -0.64P\,\text{Im}(V_L g_R^*) \\ (P:\text{top quark polarization}) \end{array}}$$

(Assuming $V_L = 1$, $P = 0.9$)

SM expectation: $A^N_{FB} = -0.708\times10^{-3}$

Measurement of A^{N}_{FB} in t-channel process





(In the W rest frame)

$\vec{p}_l$: the momentum of the lepton

(In the top quark rest frame)

$\vec{p}^{\,*}_W$: the momentum of W boson

$\vec{s}_t$: top quark polarization axis

(The direction of q')

$\vec{N}$: $\vec{s}_t \times \vec{p}^{\,*}_W$

$\vec{T}$: $\vec{p}^{\,*}_W \times \vec{N}$

(Following notation of arXiv:1005.5382)

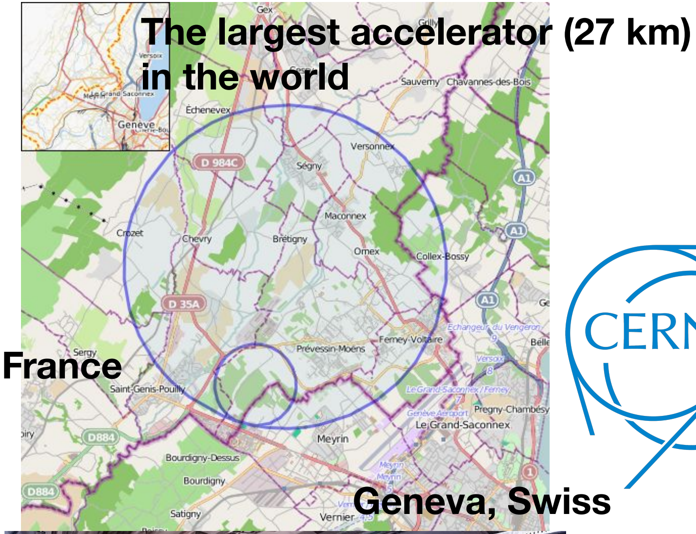
Target variable

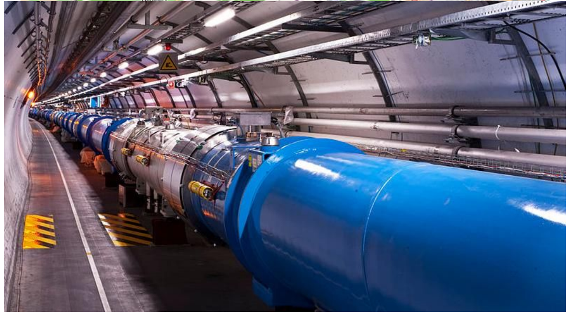
$$A^{N}_{FB} = \frac{\#(\cos\theta^N > 0) - \#(\cos\theta^N < 0)}{\#(\cos\theta^N > 0) + \#(\cos\theta^N < 0)},$$

where θ^N is the opening angle between $\vec{N}$ and $\vec{p}_l$

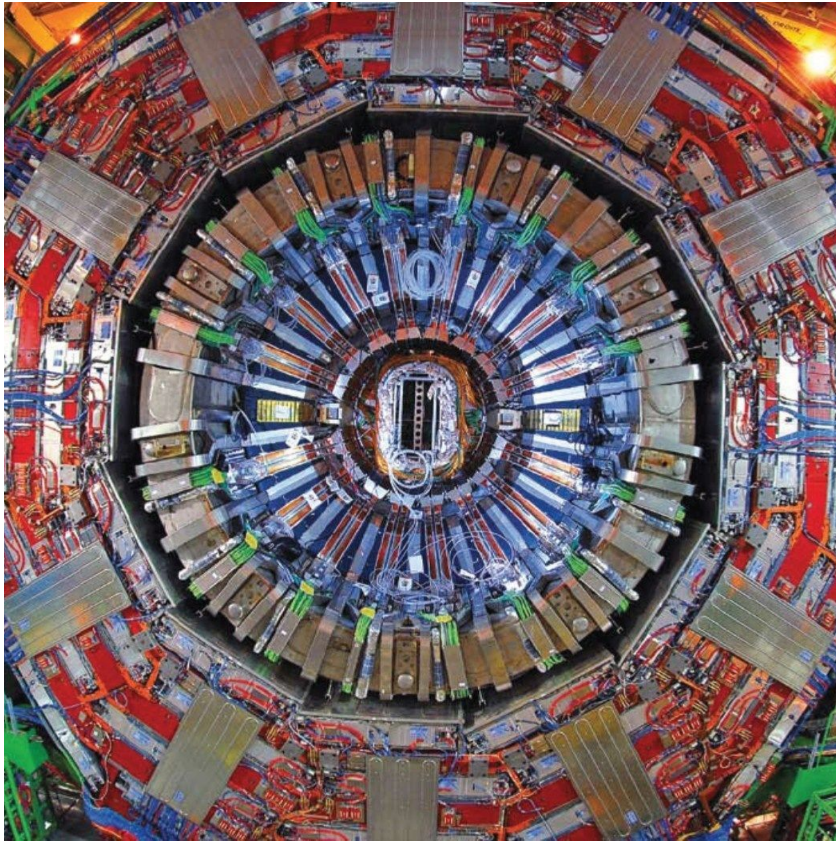
A result on Run I:
Search for CP violation in single top quark events in pp collisions at √s = 7 TeV with the ATLAS detector (ATLAS-CONF-2013-032)

About LHC and CMS



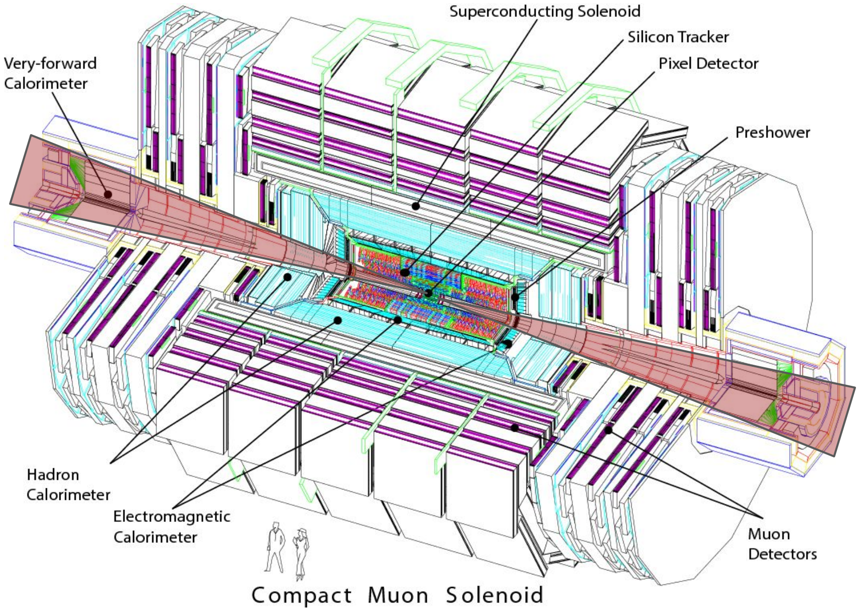


LHC
(Large Hadron Collider)



CMS
(Compact Muon Solenoid)

About LHC and CMS

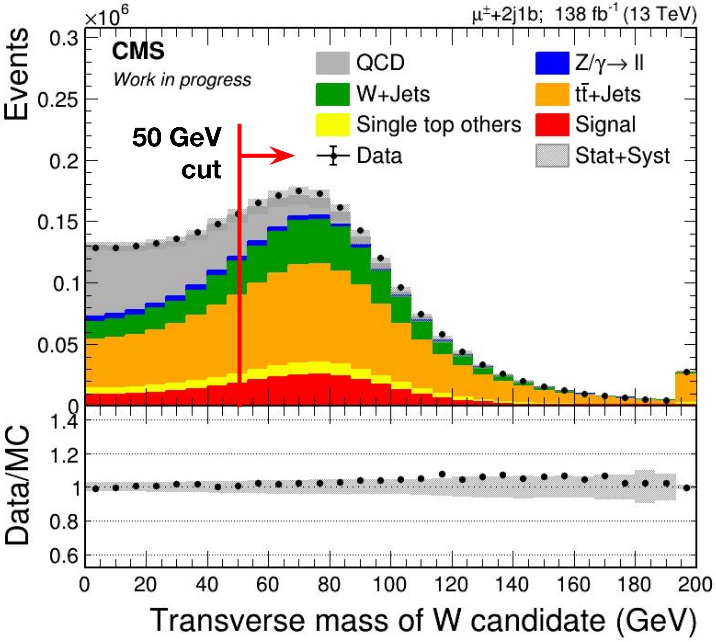


- CMS is one of general-purpose detectors of LHC
- One collision per 25 ns
- Although only triggered data are taken, CMS collects 4 PB of data per year
- For this analysis 380 TB of data (1.2 x 10^9 events) taken in 2017 are used
  - Only single muon or single electron triggered events

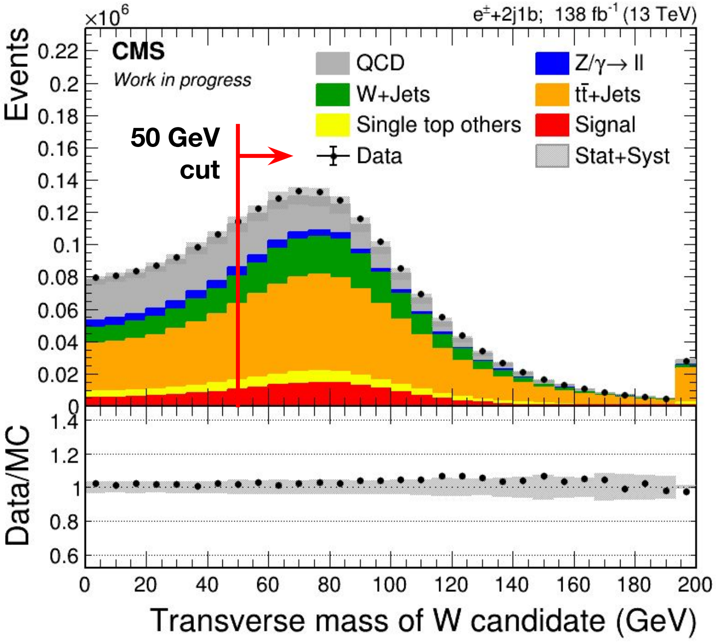
Data and event selection

- Target: Full Run II (pp-collision, 2016-2018, 138 fb^{-1}) in CMS
- Monte-Carlo (MC) simulations
  - Signal
    - Single top quark t-channel process
  - Background
    - Top quark pair production, single top quark associated with W boson (tW)
    - W+jets, DY+jets
- Object selection
  - **Lepton**: Muon ($p_T > 26$ GeV, $|\eta| < 2.4$) or electron ($p_T > 29$ GeV, $|\eta| < 2.4$)
  - **Jet**: Reconstructed by anti-k_T (cone size 0.4), $p_T > 40$ GeV,
  - **b-tagged jet**: DeepJet for b-tagging (eff. > 0.5)
- Event selection for signal region
  - Exactly one muon or electron, no additional lepton with looser condition
  - 2j1b (exactly 2 jets, exactly 1 b-tagged jet and 1 non-b-tagged jet)
- All backgrounds except QCD are taken from MC simulation, while QCD background is taken from a data-driven method

Transverse W mass distributions



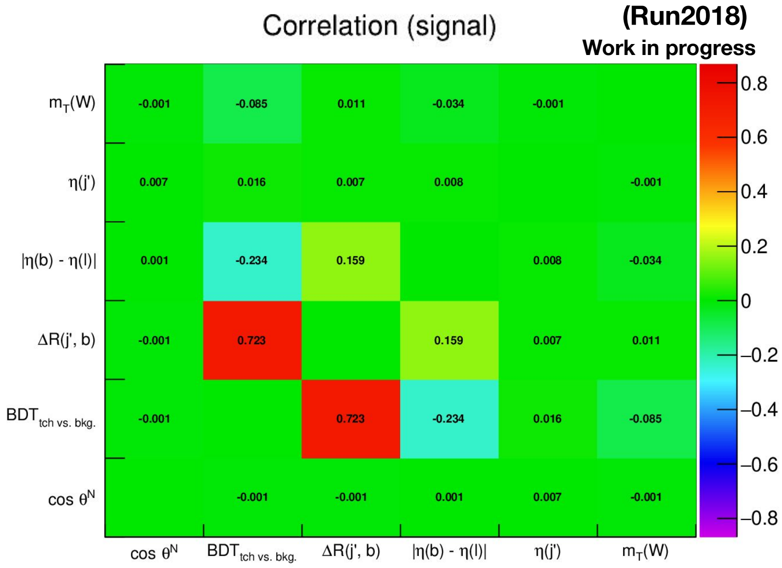
Muon channel



Electron channel

Signal extraction (MVA configuration)

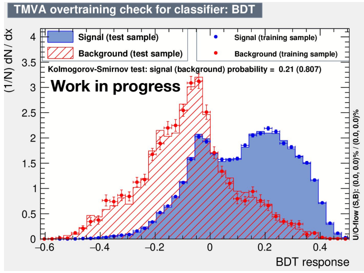
- Used variables
  - $\Delta R(j', b)$ (j’: the associated quark)
  - $\Delta\eta(l, b)$ (l: lepton, b: the b-quark)
  - $\eta(j')$
  - $m_T(W)$
- Used method
  - BDT (with adaBoost)
  - Hyperparameter: Default
  - Cut: 2 jets (one of them b-tagged), $m_T(W) > 50$ GeV
- Signal: t-channel process
  Background: All others
- No significant correlation between the BDT score and $\cos\theta^N$
- A cut at the BDT score giving the maximal significance is applied



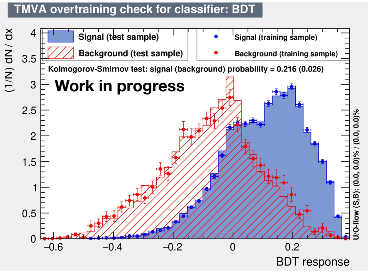
Correlation (signal)

(Run2018)
Work in progress

	$\cos\theta^N$	$BDT_{tch\ vs.\ bkg.}$	$\Delta R(j', b)$	$\lvert\eta(b) - \eta(l)\rvert$	$\eta(j')$	$m_T(W)$
$m_T(W)$	-0.001	-0.085	0.011	-0.034	-0.001	
$\eta(j')$	0.007	0.016	0.007	0.008		-0.001
$\lvert\eta(b) - \eta(l)\rvert$	0.001	-0.234	0.159		0.008	-0.034
$\Delta R(j', b)$	-0.001	0.723		0.159	0.007	0.011
$BDT_{tch\ vs.\ bkg.}$	-0.001		0.723	-0.234	0.016	-0.085
$\cos\theta^N$		-0.001	-0.001	0.001	0.007	-0.001



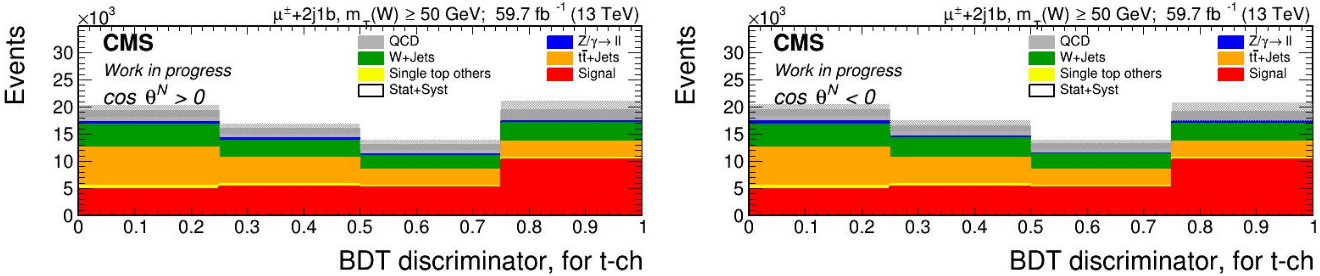
Muon channel



Electron channel

Estimation of A^N_{FB} at parton level

- Binned likelihood fit with profiling is employed to unfold $cos\ \theta^N$ at parton level
- BDT score is used for the template variable for binned likelihood fit
- To avoid a difficulty to deal with scale uncertainties, we express positive and negative $cos\ \theta^N{}_{\text{parton}}$ yields with $A^N{}_{FB}$ and the scale of signal and perform the fit procedure to estimate $A^N{}_{FB}$ directly
- The fit is performed in each lepton channels and eras simultaneously
- Uncertainties are taken account by Conways method [1] with Barlow-Beeston method [2]
- Higgs combine tool is used to unfold $cos\ \theta^N$ distribution at parton level

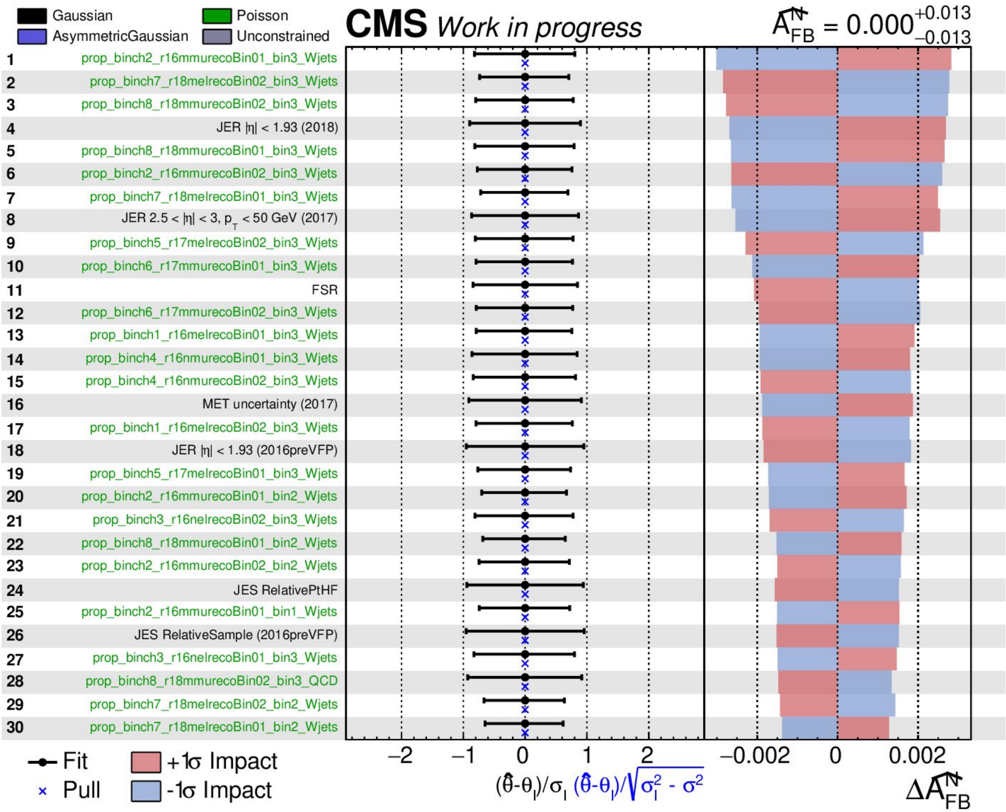


Template variable distributions in cos θ^N > 0 and cos θ^N < 0 (in Run2018 muon channel)

---

[1] J. S. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra", in PHYSTAT 2011, pp. 115–120. 2011. arXiv:1103.0354.
[2] R. J. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples", Comput. Phys. Commun. 77 (1993) 219–228, doi:10.1016/0010-4655(93)90005-W.

Preliminary result: Pull and impact



- Pull and impacts (top 30)
- Asimov dataset is used
- Green: From MC stat. uncertainty
    - Mostly from lack of W+jets MC sample
- Parts of JES, JER, MET, and FSR are dominant

Expected precision (using Asimov dataset):

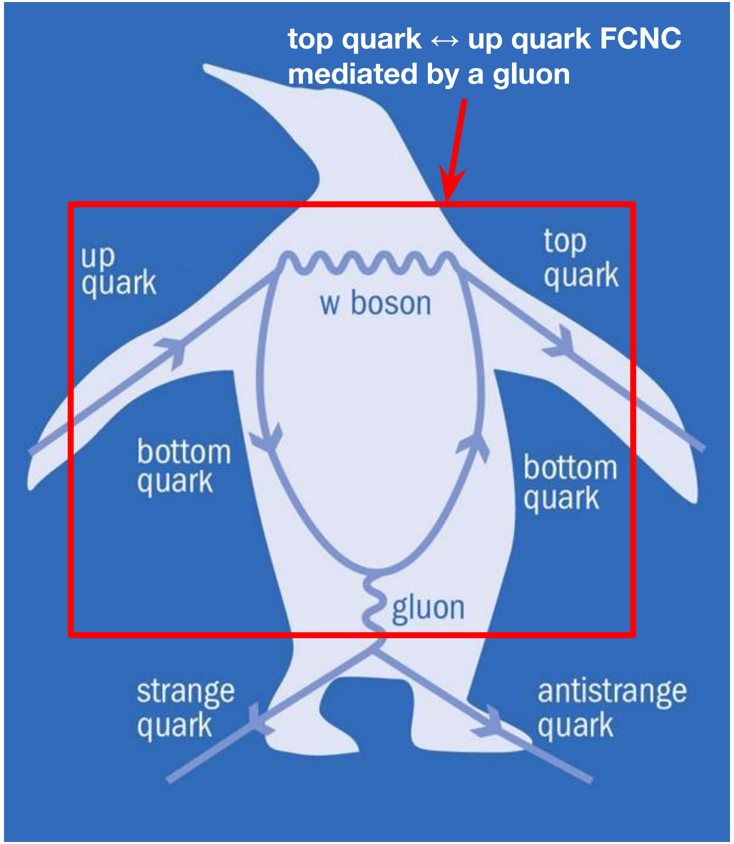
$$\Delta A^{N}_{FB} = \pm 0.013$$

Previous result from ATLAS Run I (ATLAS-CONF-2013-032):

$$A^{\mathrm{N}}_{\mathrm{FB}} = 0.031 \pm 0.065\,(\mathrm{stat.})\,^{+0.029}_{-0.031}\,(\mathrm{syst.})$$

tqg FCNC

FCNC

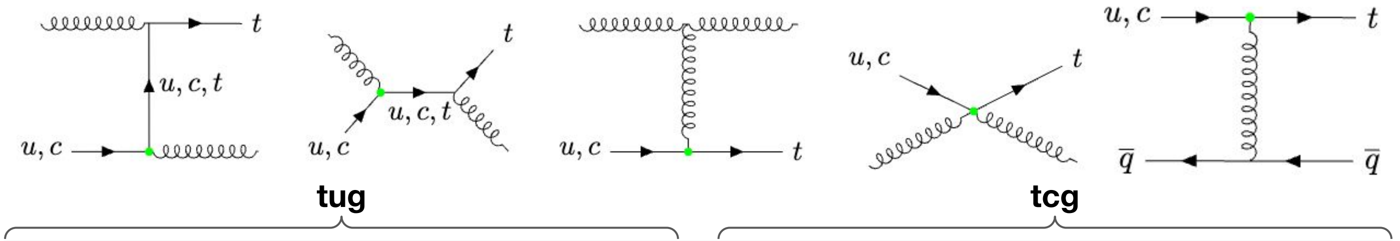


(But this process is suppressed in Standard Model...)

- Flavor-changing neutral current (FCNC)
  - An interaction changing the type (flavor) of an incoming particle without changing its charge
    (E.g., top quark → up quark,
    top quark → charm quark)
- Suppressed in Standard Model
  - Not allowed in tree level (i.e., the lowest perturbation order)
  - Suppressed in the next-leading order by GIM mechanism (or Glashow–Iliopoulos–Maiani mechanism)

Introduction

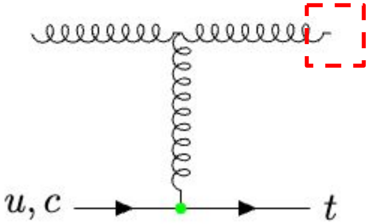
- Several beyond-Standard Models (BSM), e.g., 2HDM, MSSM, SUSY with R-parity violation, predict a significant increase of flavor-changing neutral current (FCNC) strength
- **Target**: Discrimination of tg/tq production from tqg FCNC from Standard Model (SM) backgrounds
- The predicted upper limit of $|\kappa_{tqg}|/\Lambda$ (q = u, c) : ~10^{-7} TeV^{-1} (SM) → ~10^{-3} TeV^{-1} (BSM)



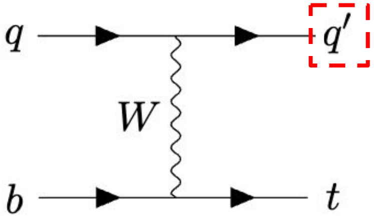
$$\boxed{\left(\frac{\kappa_{tug}}{\Lambda}\bar{q}_1\sigma^{\mu\nu}t^a t + \frac{(\kappa_{tug})^*}{\Lambda}\bar{t}\sigma^{\mu\nu}t^a q_1\right) g_s G^a_{\mu\nu} + \left(\frac{\kappa_{tcg}}{\Lambda}\bar{q}_2\sigma^{\mu\nu}t^a t + \frac{(\kappa_{tcg})^*}{\Lambda}\bar{t}\sigma^{\mu\nu}t^a q_2\right) g_s G^a_{\mu\nu}}$$

The 5-dim effective Lagrangian of theory-independent tqg FCNC interaction
Experimental limit: $|\kappa_{tug}|/\Lambda < 4.1\times10^{-3}$ TeV^{-1}, $|\kappa_{tcg}|/\Lambda < 1.8\times10^{-2}$ TeV^{-1}
(arXiv:1610.03545)

Signal and Backgrounds



tg production
with tqg FCNC vertex



Single top quark
t-channel process

- The presence of tqg FCNC yields tg productions, which is significantly different from SM events with top quark
  - tg production is ~80% (~50%) from tug (tcg) vertex
- Three signal samples of tqg FCNC processes
  - Only from tug vertex
  - Only from tcg vertex
- Main SM backgrounds
  - Single top quark t-channel process, top quark pair production, W+jets

Machine Learning Techniques

Boosted decision tree (BDT)

- A widely used machine learning technique in particle physics field
- Used for benchmark
- Used variables
    - Lepton kinematics & PID & charge
    - leading light jet, leading b-tagged jet kinematics
    - MET, MET ϕ
    - Reconstructed top quark kinematics
    - H_T, Several angular variables between objects
    - **Jet shape variables** of leading light jet (See the table on the right)

Self-attention jet-parton assignment (SAJA)[1]

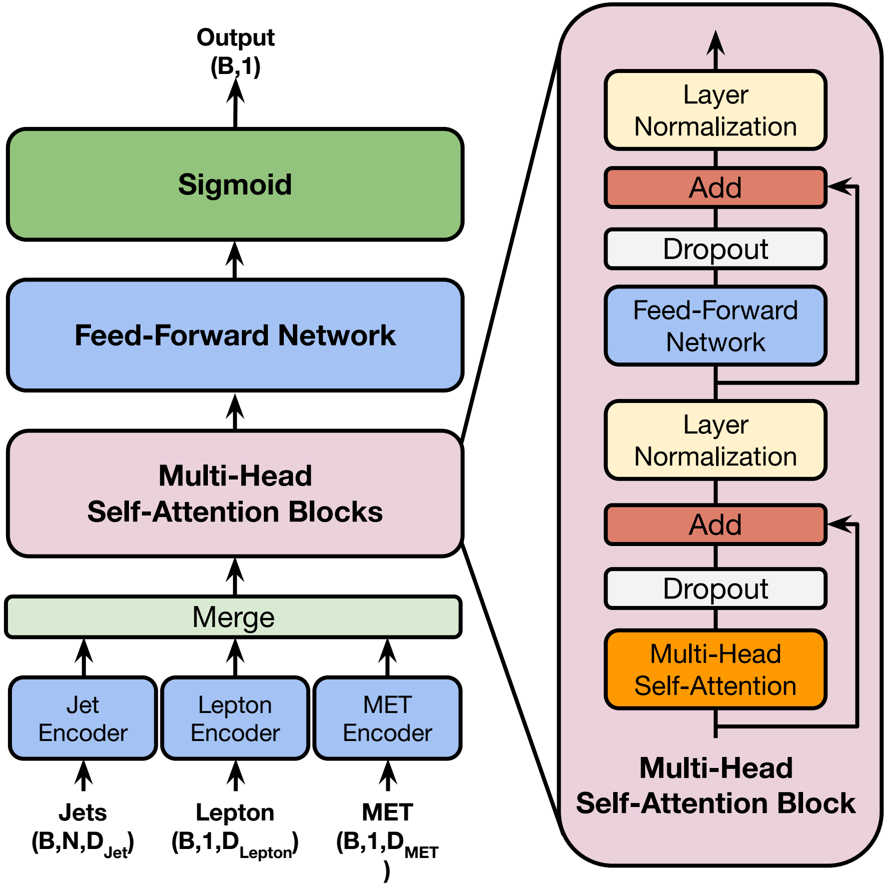
- Self-attention technique for assignment of physical objects in particle physics
- It takes into account the event topology and the all jet information together
    - This property gives an advantage on the use of the jet information for jet flavor determination
- Used variables
    -

Lepton kinematics & PID & charge, MET, MET ϕ		
Jets (all)	p_T, η, ϕ, mass b-tagging flag	**Jet shape variables** # of charged hadron # of neutral hadron # of electron # of muon # of photon major axis, minor axis p_TD

For each BDT and SAJA, two models with/without **jet shape variables** are trained

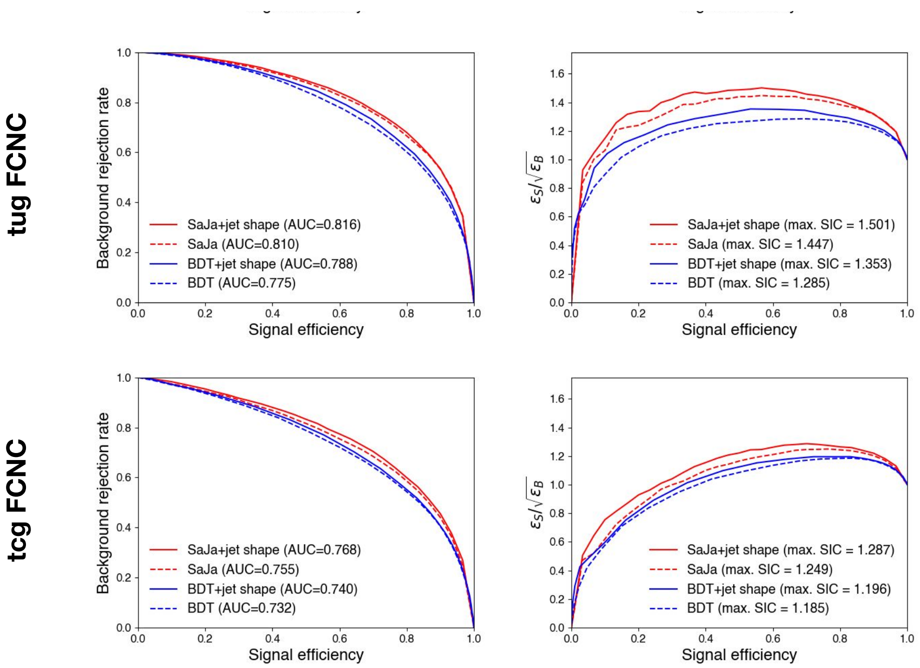
[1] J. S. H. Lee, I. Park, I. J. Watson, S. Yang, Zero-Permutation Jet-Parton Assignment using a Self-Attention Network, arXiv:2012.03542

SAJA: Architecture



- Modified model from SAJA
  - Modified for single lepton channel
- Variables of physical objects are entered to encoder and concatenated to pass through a sequence of self-attention blocks
- The output is a binary classification between signal (FCNC events) and background (SM)

Results



- Trains with tg production samples, tug FCNC samples, and tcg FCNC samples
- ROC curves and SIC from the trainings
  - SIC: Significance Improvement Curve (= [efficiency of signal (ε_S)] / √[efficiency of background (ε_B)])
- The maximal significance of SAJA+jet shapes is significantly higher than others
  - The difference of quark jet and gluon jet is crucial

Summary

- Studies in single top quark t-channel process
- CP violation
    - Estimation of the strength of CP-violation by A^{N}_{FB} measurement is performed in the t-channel production of single top quarks in proton-proton collision at 13 TeV using data in full Run II (2016-2018) collected by the CMS detector
    - Expected precision has been estimated preliminarily using MC simulation
- tqg FCNC
    - Discrimination of tqg FCNC events from Standard Model (SM) backgrounds
    - The enrichment of tg production in tqg FCNC processes will be useful for this purpose
    - We have observed an improvement using SAJA with jet shape information

- Talks
    - Measurement of CP violation in single top t-channel production at 13 TeV (2023 KPS Spring Meeting)
    - Identification of tqg FCNC process using machine learning techniques
    - (2023 KPS Fall Meeting)

Backup

Object and event selection

- Object selection

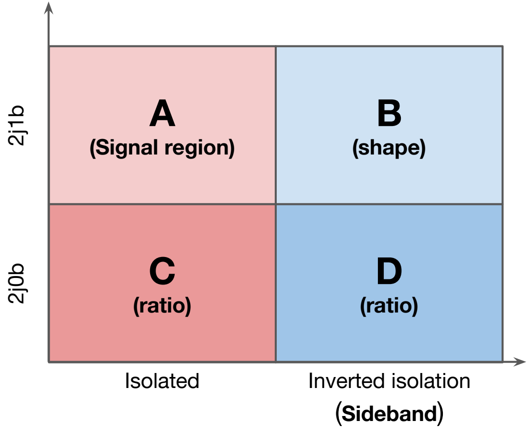
Muon:	$p_T > 29$ GeV (2017), $\lvert\eta\rvert < 2.4$, Tight ID, rel. Iso < 0.06 26 GeV (2016, 2018)
Electron:	$p_T > 29$ GeV (2016), $\lvert\eta\rvert < 2.4$, cutBased TightWP 35 GeV (2017, 2018)
Veto muon:	$p_T > 10$ GeV, $\lvert\eta\rvert < 2.4$, Loose ID, rel. Iso < 0.2
Veto electron:	$p_T > 15$ GeV, $\lvert\eta\rvert < 2.4$, cutBased VetoWP
Jet:	$p_T > 40$ GeV, $\lvert\eta\rvert < 4.7$ (2.4 for b-tagged), Loose ID, ΔR(lepton, jet) > 0.4, DeepJet TightWP

- Lepton selection: Exactly one “tight” lepton (muon, electron), no additional veto lepton
- Corrections: Using correctionlib with JSON POG integration

  Muon efficiency corrections for iso < 0.06 provided by Matteo (#)

Background Estimations

- Control regions by the number of jets and b-tagged jets
  - 2j0b: QCD control region
- QCD: ABCD method is applied to take shapes and scales
  - QCD shape is taken from sideband region
    - Sideband region is set by inverse of lepton isolation
  - The scale is estimated in 2j0b region
    - Fitting with $m_T(W)$ distribution
  - For muon channel, the estimation is performed in $|\eta_{muon}| < 1.4$ and $|\eta_{muon}| > 1.4$, separately
- Other backgrounds are from MC



	Isolated	Inverted isolation (Sideband)
2j1b	**A** **(Signal region)**	**B** **(shape)**
2j0b	**C** **(ratio)**	**D** **(ratio)**

ABCD regions for QCD estimation

Estimation of A^N_{FB} at parton level

$$L = \prod_{i}^{\text{years, bins}} \prod_{j\in\{+,-\}} \text{Poi}\left(\left(\frac{s}{2}\right)(1+A^N_{FB})S_{ij+}(\vec{\theta}) + \left(\frac{s}{2}\right)(1-A^N_{FB})S_{ij-}(\vec{\theta}) + B_{ij}(\vec{\theta}) \middle| N_{ij}\right) \times (\text{constraints})$$

$N_{i\pm}$: Event yield of data in bin i with $\pm\cos\theta^N_{\text{reco}} > 0$

$B_{i\pm}$: Background yield in bin i with $\pm\cos\theta^N_{\text{reco}} > 0$

$S_{i\pm+}$: Signal yield in bin i with $\pm\cos\theta^N_{\text{reco}} > 0$ and $\cos\theta^N_{\text{parton}} > 0$

$S_{i\pm-}$: Signal yield in bin i with $\pm\cos\theta^N_{\text{reco}} > 0$ and $\cos\theta^N_{\text{parton}} < 0$

$\vec{\theta}$: Nuisance parameters

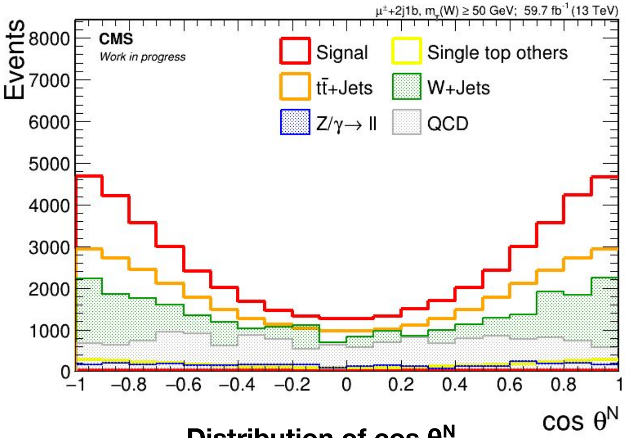
(Powered by
Higgs combine tool)

- To estimated A^N_{FB}, binned likelihood with profiling is employed
- The parameters of yields with positive and negative $\cos\theta^N_{\text{parton}}$ in the fit model are set to be $(s/2)(1+A^N_{FB})$ and $(s/2)(1-A^N_{FB})$, respectively
- Letting A^N_{FB} be the only POI (Parameter Of Interest)
- Uncertainties are taken account by Conways method [1] with Barlow-Beeston method [2]
- The fit is performed in each lepton channels and eras simultaneously

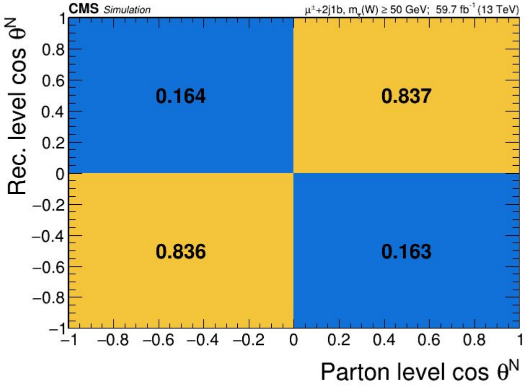
[1] J. S. Conway, “Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra”, in PHYSTAT 2011, pp. 115–120. 2011. arXiv:1103.0354.
[2] R. J. Barlow and C. Beeston, “Fitting using finite Monte Carlo samples”, Comput. Phys. Commun. 77 (1993) 219–228, doi:10.1016/0010-4655(93)90005-W.

Unfolding setup for $\cos\theta^N$

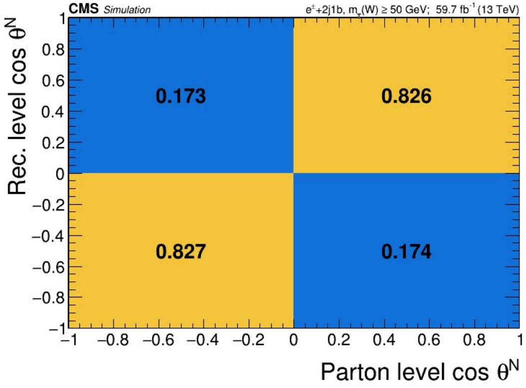
(Run2018)



Distribution of $\cos\theta^N$
($m_T(W)$ > 50 GeV)



Response matrix
(2018, Muon channel)



Response matrix
(2018, Electron channel)

- Response matrices of $\cos\theta^N$ in each lepton channels
  - The off-diagonal components are smaller than 0.2
  - No regularization is used

Systematic uncertainties

Experimental uncertainties

- Luminosity
- Pileup
- Trigger efficiency
- Muon ID efficiency
- Muon isolation efficiency
- Electron ID efficiency
- Electron reconstruction efficiency
- b-tagging efficiency
- Jet energy resolution
- Jet energy scale
- Missing transverse energy reconstruction
- Other minor sources

Theoretical uncertainties

- Parton Shower (`ISR`); t-ch, ttbar
- Parton Shower (`FSR`); t-ch, ttbar
- Matrix element scale μ_R; t-ch, ttbar, W+jet
- Matrix element scale μ_F; t-ch, ttbar, W+jet
- Resummation (h_{damp}); ttbar
- UE tune (t-ch, ttbar)
- top quark mass; t-ch, tt/tW (±1 GeV)
- PDF
- top quark pair p_T reweight

MC sample size uncertainties

Event Selection

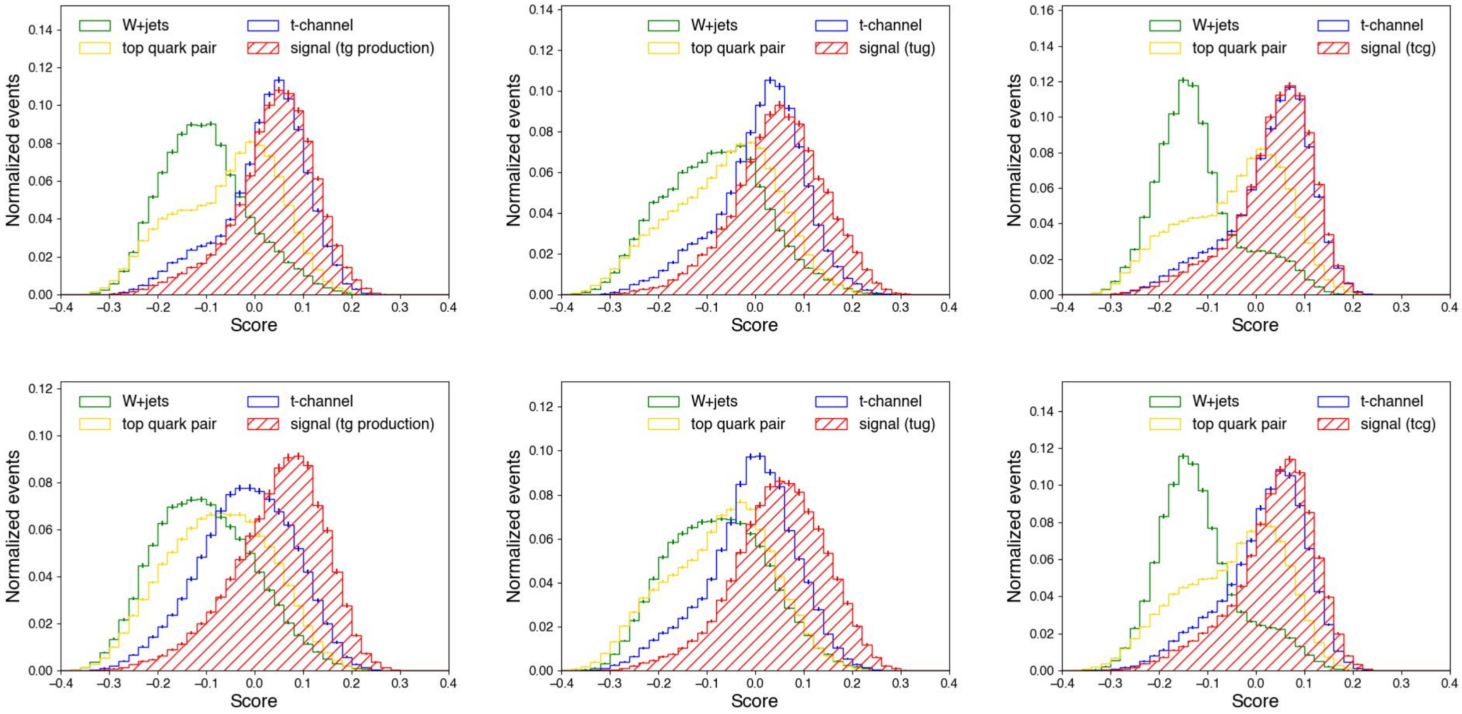
- Following a CMS study of single top quark t-channel process[1]
- Lepton
  - Muon: $p_T > 30$ GeV, $|\eta| < 2.4$, rel. $\text{iso}_{\Delta R<0.4} < 0.06$
  - Electron: $p_T > 30$ GeV, $|\eta| < 2.4$, rel. $\text{iso}_{\Delta R<0.3} < 0.06$
  - Veto muon: $p_T > 15$ GeV, $|\eta| < 2.4$, rel. iso < 0.20
  - Veto electron: $p_T > 15$ GeV, $|\eta| < 2.4$, rel. iso < 0.20
- Jet
  - anti-k_T 0.4, $p_T > 40$ GeV, $|\eta| < 2.4$, $\Delta R(\text{lepton, jet}) > 0.4$
- b-tagged jet
  - $p_T > 40$ GeV, $|\eta| < 2.4$, $\Delta R(\text{lepton, jet}) > 0.4$
  - b-tagging efficiency: Like CMS CSVv2 medium working point
- Only one lepton without any additional veto lepton is required
- (the number of jets) ≥ 2, (the number of b-tagged jets) ≥ 1
- $m_T(W) > 50$ GeV
  - In this region, we can ignore QCD background

- Two machine learning techniques, boosted decision tree (BDT) and self-attention jet-parton assignment (SAJA), are used to discriminate the FCNC events and SM events

---

[1] CMS Collaboration, Measurement of differential cross sections and charge ratios for t-channel single top quark production in proton-proton collisions at $\sqrt{s} = 13$ TeV, arXiv:1907.08330

Results



BDT

BDT+jet shape

tg production
from tug + tcg FCNC

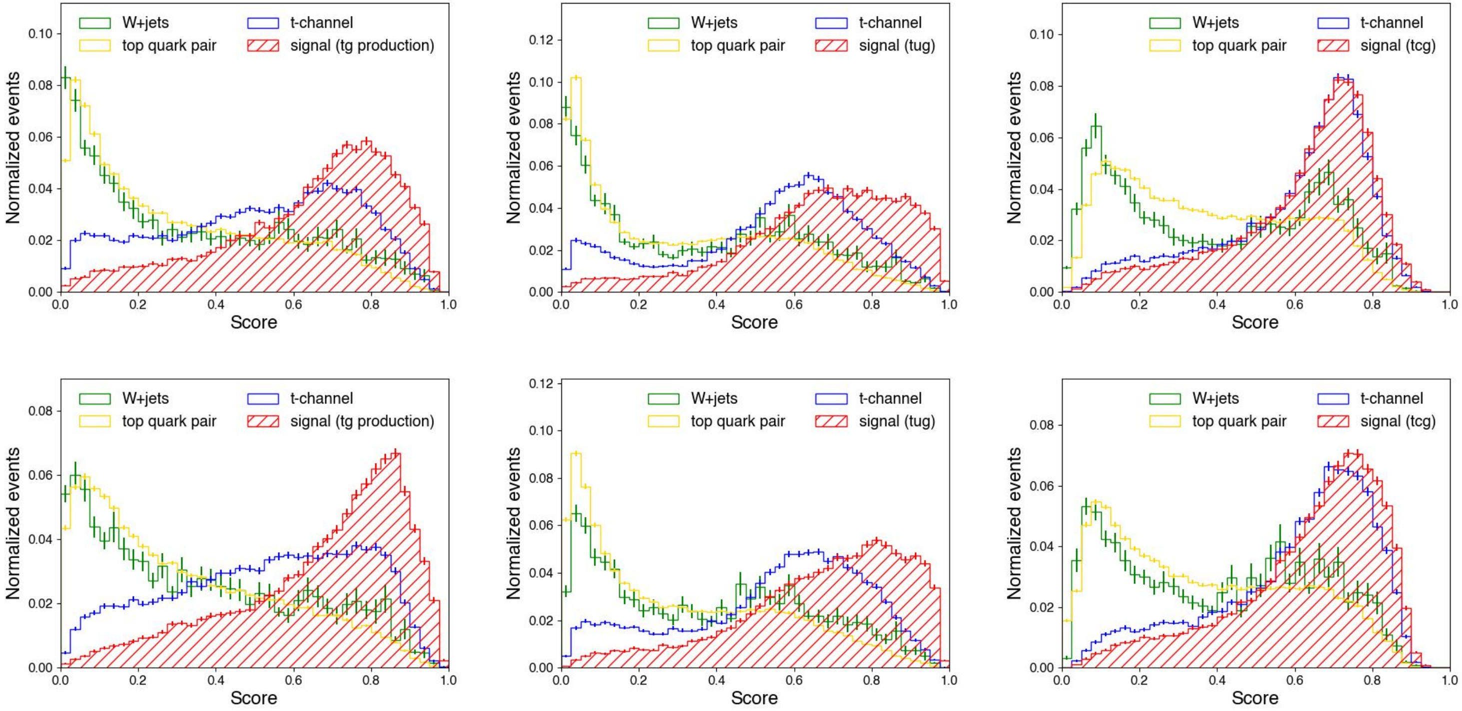
tg/tq production
from tug FCNC

tg/tq production
from tcg FCNC

Results

SAJA

SAJA+jet shape



tg production
from tug + tcg FCNC

tg/tq production
from tug FCNC

tg/tq production
from tcg FCNC