# A measurement of top quark production at $\sqrt{s} = 13$ TeV with LHCb data


UNIVERSITY OF LIVERPOOL

**James Vincent Mead**

Department of Physics

University of Liverpool

This dissertation is submitted for the degree of

*Doctor of Philosophy*

March 2021

I dedicate this thesis to my nephews with the hope they one day revel in products of their own fascination, wherever it may find them.

# Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text.

James Vincent Mead
March 2021

# Acknowledgements

I would like to acknowledge the considerable lengths to which **Dr Stephen Farry**, **Dr David Hutchcroft** and **Prof Tara Shears** took towards this project and affording me the opportunities I had throughout my studentship. The work of **Dr William Barter**, **Dr Philip Ilten**, **Dr Murilo Rangel**, **Dr Oscar Francisco**, **Dr Lorenzo Sestini** and **Dr Daniel Craik** was heavily relied upon for the development of the work presented in this thesis and each have been generous with their time and energy. Additionally, the patience and endless discussion offered by **Dr Matthew Sullivan** and **Dr Vinícius Franco** since making their acquaintance was very much invaluable. Finally, in the closing days of this almost surreal annus horribilis, I must thank **Simon Swarbrick** and **Dr Jack Williams** for providing me with considerably more than just good company in trying times.

# Abstract

■ The LHCb experiment provides unique detector coverage of the highest energy proton-proton interactions ever produced. Designed to study $b$-& $c$-hadron physics at the LHC, the detector is fully instrumented in the forward region, $2.0 < \eta < 4.5$, with excellent tracking, vertex resolution and particle identification. The increased centre of mass energy in Run II gives rise to 3-fold increased inclusive top cross-section over Run I at the LHC, corresponding to a 10-fold increase within the LHCb acceptance. The top quark is the heaviest fundamental particle and is expected to play a special role in new physics scenarios.

Higher order interference mechanisms, sensitive to physics beyond the reach of current colliders, result in a charge asymmetry in the relative angular distributions of $t\bar{t}$ pairs. The LHCb acceptance offers greater sensitivity to $A_C^{t\bar{t}}$ due to reduced dilution from gluon-gluon fusion. Top quarks are identified through the presence of a high $p_T$ muon and $b$-jet in the final state. Forward production was first observed with Run I data at LHCb in this channel. Top pairs may be identified with an additional opposite-sign lepton or $b$-jet.

The increase in available statistics with Run II, as well as improved signal to background ratio, enables differential measurements of heavy flavour tagged $W$+jet yields in muon pseudorapidity. New running conditions necessitated re-optimisation of jet input selection for reconstruction as well as renewal of heavy flavour tagging algorithms, achieved using deep learning techniques. Together these provide the first full Run II top cross-section in the $\mu + b$ channel at LHCb and the first top asymmetry measurement in the forward region. Each use data corresponding to an integrated luminosity of $5.4\,\text{fb}^{-1}$ (5% systematic).

$$\begin{aligned} \sigma(t)\ [13\,\text{TeV}] &= 0.89 \pm 0.06\ (\text{stat}) \pm 0.18\ (\text{syst})\ \text{pb}\ , \\ \sigma(\bar{t})\ [13\,\text{TeV}] &= 0.66 \pm 0.05\ (\text{stat}) \pm 0.17\ (\text{syst})\ \text{pb}\ . \end{aligned}$$

$$A_C^{\text{top}}\ [13\,\text{TeV}] = 0.14 \pm 0.05\ (\text{stat}) \pm 0.05\ (\text{syst})\ .$$

While the latter measurement was inconclusive with respect to the $t\bar{t}$ asymmetry, the combined asymmetry was observed to $2.1\sigma$ above zero. Differential cross-sections were found to be within $1\sigma$ per bin of NLO standard model predictions. The precision of both sets of measurements are systematically limited, with largest contributions from heavy flavour yields ($\sim 20\%$) and SV-tagging efficiencies (10%) despite at least partially cancelling in $\delta A_C$.

# Table of contents

# List of figures

# List of tables

# Introduction

Particle physics is the discipline of interrogating the fundamental constituents of the universe, from the properties of sub-atomic particles to the forces dictating their laws of motion. The Standard Model encompasses a quantum theory of electrodynamics, the weak and strong nuclear forces, as well as the origin of the massive behaviour of known particles. However, there remain crucial facets of the universe which cannot emerge through such a model. Extensions to the Standard Model or some overarching framework may subsume general relativity, a persisting theory of gravitation, but have thus far produced no verified predictions of new physics. Processes by which, on cosmological scales, the one part per billion matter-antimatter asymmetry may arise are yet to be discovered. Establishing the non-zero mass of neutrinos through their flavour oscillations, in spite of the absence of chiral partners, raises further foundational questions. In addition, experiments have not detected any candidates explaining astrophysical observations of either dark matter or dark energy.

Collider experiments rely on the principle of mass-energy equivalence through which high energy particle interactions are able to produce particles otherwise seen only in rare instances or in the early universe. While these particles rapidly decay into lighter more stable particles, their properties and their interactions may be probed in order to test theoretical predictions of the Standard Model and of new physics. While testing such models through measurements in data, a direct approach allows for constraints on the parameters of a given model and comparisons of quantities compared to its precise predictions. Depending on the assumptions behind such predictions, indirect tests access higher energy scales with contributions beyond tree-level where new physics may yet lie.

The jump in centre-of-mass energy of the LHC to $\sqrt{s} = 13\,\text{TeV}$ corresponds to a substantial gain in statistical power, increasing production cross-sections and thus the breadth of physics reach delivered by modern collider experiments. This thesis presents a measurement of 13 TeV top production using the LHCb detector, with access to the forward region complimentary to boosted regimes of other collider experiments, using the full Run II data set collected from 2015-2018, corresponding to an integrated luminosity of $5.4\,\text{fb}^{-1}$. The combined differential single top and top pair production cross-sections provide a Standard

Model test at extremes of the proton-proton collision phase space. Measuring the top quark charge asymmetry, the $t\bar{t}$ component of which arises through interference terms via a purely quantum mechanical or beyond tree-level process, tests both perturbative quantum chromodynamics and for signs of new physics.

A summary of the relevant background in particle physics and collider phenomenology is to be found in Chapter 1. The experimental hardware and software relied upon for collecting this data set and the techniques for reconstructing and analysing it are is discussed in Chapters 2 & 3 respectively. In efforts towards these measurements in the $\mu + b$-jet final state, a new jet reconstruction configuration was produced, optimised and validated for 13 TeV running conditions, as shown in Chapter 4. Furthermore, new secondary decay vertex based machine learning models employing deep neural networks were developed based upon the new configuration applied in Run II data, as per Chapter 5. The resultant Run II top quark measurements are laid out in Chapter 6. The latter three chapters contain studies unique to this thesis.

# Chapter 1

# Theoretical overview

In this chapter, the theoretical description of particle physics is discussed. An overview of the current theory, encapsulated in the Standard Model (SM), is given in Section 1.1 which includes the contents of the SM and how the observed interactions are understood followed by the motivation for new physics tests and precision measurements. Practical implications of hadron collider physics at the Large Hadron Collider (LHC), its limitations and prospects for discovery are covered in Section 1.2 together with the application of the SM and computational techniques to predicting physical processes in high energy experiments, applied to top physics in Section 1.3.

## 1.1 The Standard Model of particle physics

The equations of motion of a field theory are encoded within its Lagrangian density. The invariance of a Lagrangian under groups of transformations implies the symmetries of the system. These prescribe conserved currents which, to the best of our understanding, underpin the predictability of nature. The most stringently tested scientific predictions, which concern the fundamental particle interactions, rely on gauge-invariant quantum field theories (QFTs) constructed to describe physical observables consistent with the measured universe.

The $\mathrm{ISO}_+(3,1)$ group describes the symmetries of a relativistic space-time that combine with quantum mechanics within a QFT. The Standard Model (SM) is one such QFT based on the product of gauge groups $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$, representing elementary particles and their interactions through continuous fields. $\mathrm{SU}(3)_C$ refers to the theory of colour charge, quantum chromodynamics (QCD). $\mathrm{U}(1)_Y$ is the group of the weak-hypercharge while $\mathrm{SU}(2)_L$ pertains to a symmetry group of weak isospin; these unify within electroweak (EW) theory, from which a $\mathrm{U}(1)$ description of quantum electrodynamics (QED) emerges.

Table 1.1 Generations of quarks, their masses and electric charge arranged by family [1].

| Family | Gen. | Quark | Mass [MeV$^{-1}$] | | | Charge [e$^{-1}$] | Spin [ħ$^{-1}$] |
|---|---|---|---|---|---|---|---|
| $q$ | 1 | $u$ | 2.16 | $^{+0.49}_{-0.26}$ | | $^{2}/_{3}$ | $^{1}/_{2}$ |
| | 2 | $c$ | 1.27 | $^{+0.02}_{-0.02}$ | $\times 10^{3}$ | $^{2}/_{3}$ | $^{1}/_{2}$ |
| | 3 | $t$ | 172.9 | $^{+0.4}_{-0.4}$ | $\times 10^{3}$ | $^{2}/_{3}$ | $^{1}/_{2}$ |
| $q'$ | 1 | $d$ | 4.67 | $^{+0.48}_{-0.17}$ | | $-^{1}/_{3}$ | $^{1}/_{2}$ |
| | 2 | $s$ | 93 | $^{+11}_{-5}$ | | $-^{1}/_{3}$ | $^{1}/_{2}$ |
| | 3 | $b$ | 4.18 | $^{+0.03}_{-0.02}$ | $\times 10^{3}$ | $-^{1}/_{3}$ | $^{1}/_{2}$ |

Table 1.2 Generations of leptons, their masses and electric charge arranged by family [1].

| Family | Gen. | Lepton | | Mass [MeV$^{-1}$] | | Charge [e$^{-1}$] | Spin [ħ$^{-1}$] |
|---|---|---|---|---|---|---|---|
| $l^{\pm}$ | 1 | $e^{-}$ | | 0.510998946 1(31) | | $-1$ | $^{1}/_{2}$ |
| | 2 | $\mu^{-}$ | | 105.658374 5(24) | | $-1$ | $^{1}/_{2}$ |
| | 3 | $\tau^{-}$ | | 1776.86(12) | | $-1$ | $^{1}/_{2}$ |
| $l^{0}$ | 1 | $\nu_e$ | $<$ | 2 | $\times 10^{-6}$ | 0 | $^{1}/_{2}$ |
| | 2 | $\nu_\mu$ | $<$ | 0.19 | | 0 | $^{1}/_{2}$ |
| | 3 | $\nu_\tau$ | $<$ | 18.2 | | 0 | $^{1}/_{2}$ |

### 1.1.1 Particle content

The SM is the framework physicists use to describe the observed matter and three of the known forces of nature. All matter is built from point-like spin-$\frac{1}{2}$ particles known as fermions, where spin refers to their intrinsic angular momentum. These are split into two groups of six quarks and six leptons listed in Tables 1.1 & 1.2. Both groups are arranged into two families, 'up-type' $q$ and 'down-type' $q'$ quarks, charged leptons and neutrinos, each with three generations where the masses of the consecutive generations increase. The force carriers from Table 1.3 produce relative strengths and effective ranges summarised in Table 1.4.

The lepton families are distinct between the neutral neutrinos ($\nu_e$, $\nu_\mu$, $\nu_\tau$) and those carrying the unit electric charge ($e$, $\mu$, $\tau$) with corresponding generations. The quark families carry fractional unit charge; the up-type ($u$, $c$, $t$) carrying +$\frac{2}{3}$ paired with the down-type ($d$, $s$, $b$) carrying -$\frac{1}{3}$. Every fermion in the SM has an anti-matter partner with the same mass and spin but opposite charge. The charged fermions can be further attributed left- or right-handedness, depending on how their spin projections undergo space-time transformations.

Table 1.3 Boson masses, electric charge and spin [1].

| Boson | Mass $[\mathrm{GeV}^{-1}]$ | Charge $[\mathrm{e}^{-1}]$ | Spin $[\hbar^{-1}]$ |
|---|---|---|---|
| $g$ | 0 | 0 | 1 |
| $\gamma$ | 0 | 0 | 1 |
| Z | 91.1876(21) | 0 | 1 |
| W | 80.379(12) | $\pm 1$ | 1 |
| H | 125.10(14) | 0 | 0 |

Table 1.4 Relative strengths and effective ranges of fundamental forces ($Q^2 = 0$) [2, 3].

| Force | Strength $[\alpha_s^{-1}]$ | Range $[\mathrm{m}^{-1}]$ | Propagator |
|---|---|---|---|
| Strong | 1 | $\mathscr{O}(10^{-15})$ | $g$ |
| EM | $^1/_{137}$ | $\infty$ | $\gamma$ |
| Weak | $\mathscr{O}(10^{-6})$ | $\mathscr{O}(10^{-18})$ | Z |
| | | | W |
| Gravity | $\mathscr{O}(10^{-38})$ | $\infty$ | |

The forces corresponding to the strong, electromagnetic (EM) and weak interactions manifest through the exchange of integer spin force-carrying particles known as bosons. These interactions conserve the charges to which each force is coupled. Where quarks carry the charges required to couple to all these forces, leptons couple only to the EM and weak bosons with only left (right) handed (anti-)neutrinos observed via the weak interaction. The SM offers no prescription for a fundamental force of gravity.

The massive $W^{\pm}$ and $\mathrm{Z}^0$ electroweak bosons mediate the weak interaction, each possessing spin-1. Only left-handed matter and right-handed anti-matter states experience any coupling to $W$ and an enhanced coupling to the $Z$. The massless gluon, $g$, and photon, $\gamma$ mediate the strong and EM interaction respectively, again with spin-1. While the photon is electrically neutral, the gluon carries colour charge resulting in a very different phenomenology, as discussed in Section 1.1.3. The spin-0 Higgs field is responsible for the masses of the fermions and weak bosons, with its corresponding boson, $H^0$.

Gauge invariance is a feature of specific field theories relating scalar and vector potentials in which physical dynamics are unaffected by redundant degrees of freedom. Such transformations, defining the symmetry group upon which the theory is based, are known as gauge transformations. The SM is a gauge-invariant QFT, defining the particles of the theory as excitations in respective fundamental fields. The fermion fields can exhibit spatially-dependent

(local) gauge invariance given the introduction of gauge boson fields. Redundant degrees of freedom are, in other words, underlying symmetries of the theory Lagrangian and, according to Noether's theorem, correspond to conservation laws built into a given model such as, in the SM case, the various charge conserving interactions.

Despite its widely regarded success in the precise description of experimental particle physics, the SM provides an incomplete picture with the absence of gravity. The SM fails to account for a mechanism by which the matter-antimatter asymmetry of the universe may arise since the Big Bang. Cosmological observations imply a dominant matter component whose only visible interaction is through gravity, coined dark matter. The accelerating expansion of the universe has established an additional component known as dark energy, which comprises the majority of this universe [4].

### 1.1.2 Quantum electrodynamics

The Lagrangian density, $\mathscr{L}$, for a free Dirac field, $\psi$, which describes the behaviour of a fermion of mass $m$, where $\not{\partial} = \gamma^\mu \partial_\mu$ is the partial derivative in Einstein notation with Dirac $\gamma$-matrices, is as follows:

$$\mathscr{L}_f = \bar{\psi}(x)(i \not{\partial} - m)\psi(x). \tag{1.1}$$

If invariance under spatially-independent (global) $\mathrm{U}(1)$ gauge transformations (Equation 1.2) with coupling constant $g_e$ is assumed, where $\theta$ is the continuous parameter, this facilitates the conservation of charge $Q$ and its associated current.

$$\psi(x) \to e^{-ig_eQ\theta}\psi(x). \tag{1.2}$$

If allowed spatial dependence, the phase $\theta(x)$ breaks invariance through the derivative term. Replacing $\partial_\mu$ with $D_\mu$ such that the derivative becomes co-variant (Equation 1.3), transforming with the field, preserves the invariance of the Lagrangian. This requires the introduction of the vector field, $A$ (Equations 1.4 & 1.5), in order to instantiate a local symmetry.

$$D_\mu\psi(x) \to e^{-ig_eQ\theta(x)}D_\mu\psi(x) \tag{1.3}$$

$$D_\mu = (\partial_\mu - ig_eQA_\mu) \tag{1.4}$$

$$A_\mu \to A_\mu - \partial_\mu\theta(x) \tag{1.5}$$

The simplest unitary group U(1) is abelian, maintaining the commutivity of the group operation. As a result, $[A_\mu, A_\nu] = 0$ and the charged terms from the field strength tensor of the vector field, $F_{\mu\nu}$ (Equation 1.7), are dropped, thus preventing the self-interaction of the field and resembling that of classical EM. The non-invariant mass term forces $m_A$ to be set to zero to maintain gauge invariance and its resulting conservation law.

$$\mathscr{L}_b = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{m_A^2}{2}A_\mu A^\mu \tag{1.6}$$

$$\begin{aligned} F_{\mu\nu} &\equiv \frac{1}{ig_eQ}[D_\mu, D_\nu] \\ F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu + ig_e[A_\mu, A_\nu] \end{aligned} \tag{1.7}$$

Combining the free fermion Lagrangian (Equation 1.1) with that of the vector boson (Equation 1.6) produces a description of quantum electrodynamics (QED) with a framework for electric charge conserving fermion fields interacting via a massless particle where charge Q couples to the neutral vector field through $g_e \propto e$, the fundamental electric charge.

$$\mathscr{L}_{QED} = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.8}$$

### 1.1.3 Quantum chromodynamics

The interactions between quarks via gluons, governed by the strong force, is described by quantum chromodynamics (QCD), a SU(3) gauge theory. A unitary group U($N$) has generators ($N \times N$ matrices) with those of special unitary groups, SU($N$), having a determinant equal to 1. In general, unitary groups contain U(1) phases $\theta^a$, where $a = 1 \rightarrow (N^2 - 1)$ [5].

$$\psi \rightarrow e^{-ig_s\theta^a(x)\lambda^a}\psi \tag{1.9}$$

Requiring the free fields to be invariant under local SU(3) gauge transformations (Equation 1.9), where $\lambda^a$ represents the 8 generators of the group and $g_s$ is the strong coupling constant, necessitates the introduction of as many vector fields $G_\mu^a$.

$$D_\mu = \partial_\mu - g_s\lambda^a G_\mu^a \tag{1.10}$$

The QCD Lagrangian requires co-variant derivative, $D_\mu$ (Equation 1.10), and boson field tensor, $F_{\mu\nu}^a$ (Equation 1.11) where the self-coupling term does not vanish as the generators are non-commutative. This is prescribed by the non-zero structure functions, $f^{abc}$ (Equation

1.12), indicative of a non-abelian gauge group, which is true for U($N$) with $N \geqslant 2$ [5].

$$F^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f^{abc} G^b_\mu G^c_\nu \tag{1.11}$$

$$\left[\lambda^a, \lambda^b\right] = i2 f^{abc} \lambda^c \tag{1.12}$$

The fields are written as triplets in a complex space dubbed colour charge. Such a theory features bosons carrying the charge of the symmetry group, unlike the electrically neutral photon. The unbroken symmetries in $\theta^a$ provide 8 massless gluons carrying the colour - anti-colour combinations to mediate the strong interactions between quarks and one another. Quarks are the only SM fermions to exhibit colour charge.

$$\mathscr{L}_{QCD} = \bar{\psi}(i \not{D} - m)\psi - \frac{1}{4} F^a_{\mu\nu} F^{\mu\nu}_a. \tag{1.13}$$

### 1.1.4 Electroweak theory

The interactions of the weak force along with those of QED may each be understood as an emergent property of the more fundamental eletroweak (EW) theory. The Lagrangian may be broken down into terms (Equation 1.14) discussed further below where they are defined (Equations 1.18, 1.21, 1.31 & 1.32).

$$\mathscr{L}_{\text{EW}} = \mathscr{L}_{\text{Gauge}} + \mathscr{L}_{\text{Fermions}} + \mathscr{L}_{\text{Higgs}} + \mathscr{L}_{\text{Yukawa}} \tag{1.14}$$

Under a unified SU(2) $\times$ U(1) scheme reintroducing the abelian subgroup and, again, requiring local gauge invariance (Equation 1.15) necessitates a co-variant derivative $D_\mu$ (Equation 1.16).

$$\psi \to \mathrm{e}^{-\frac{i}{2} g_w W^a(x) - \frac{i}{2} g_b B(x)} \psi \tag{1.15}$$

$$D_\mu = \partial_\mu - \frac{i}{2} g_w T \sigma_a W^a_\mu - \frac{i}{2} g_b Y B_\mu \tag{1.16}$$

In SU(2), a vector field $W$ has generators $\sigma_{a=1\to3}$, the $2 \times 2$ Pauli spin matrices, with $\mathrm{Tr}\left(\sigma_i \sigma_j\right) = 2\delta_{ij}$. Three massless self-interacting gauge bosons $W^a$ are produced and $T$ is the conserved current of the SU(2) group, isospin. The generator of U(1) produces a fourth gauge boson, $B$, with coupling $g_b$ and $Y$, weak hypercharge, being conserved. The resultant field strength tensors (Equation1.17) with the corresponding SU(2) structure functions $\varepsilon^{abc}$

allow self coupling of the $W$ field through $F^a$ while $B$ remains abelian with $F'$.

$$\begin{aligned} F^a_{\mu\nu} &= \partial_\mu W^a_\nu - \partial_\mu W^a_\mu + g_w \varepsilon^{abc} W^b_\mu W^c_\nu \\ F'_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \end{aligned} \tag{1.17}$$

$$\mathscr{L}_{\text{Gauge}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W^a_{\mu\nu} W^{a,\mu\nu} \tag{1.18}$$

Fermion fields may be treated in $\mathrm{SU}(2)$ doublet representation arranged by generation, with the lepton and quark families coupling to the EW groups differently.

### Chirality

The SM EW theory decomposes fermions into left- and right-handed projections via the chirality operators. The SM, not observing parity or P-symmetry, is a chiral theory; Equation 1.16 becomes Equations 1.20 imposing $\mathrm{SU}(2)_L$ such that left-handed states remain isospin doublets. In contrast, right-handed states form weak hypercharge singlets. The fermion field charges are summarised in Table 1.5. Interactions via the $W$ bosons are observed proceeding only through left-handed fermions and right-handed anti-fermions, violating parity symmetry. The chiral symmetry, allowing left and right components of massless fermion fields to be transformed independently, is said to be broken but preserving the quantum number for electric charge, $Q = T + Y/2$.

$$\psi_L \equiv P_L \psi \equiv \frac{1-\gamma^5}{2}\psi, \quad \psi_R \equiv P_R \psi \equiv \frac{1+\gamma^5}{2}\psi \tag{1.19}$$

$$D^L_\mu = \partial_\mu + \frac{i}{2} g_w \sigma_a W^a_\mu + i g_b Y B_\mu, \quad D^R_\mu \;= \partial_\mu + i g_b Y B_\mu \tag{1.20}$$

$$\mathscr{L}_{\text{Fermions}} = \sum_{\text{f}} \bar{\psi}_L \gamma^\mu D^L_\mu \psi_L + \bar{\psi}_R \gamma^\mu D^R_\mu \psi_R \tag{1.21}$$

### Mixing and CP-violation

One might assume EW theory, while violating parity or P-symmetry, preserves charge-parity or CP-symmetry between matter and anti-matter. However, in a model containing at least three generations, if flavour field $D'$ is not an observable mass eigenstate but rather a superposition of the mass eigenstates $D$, then the mixing between quark flavours is possible [6]. In the SM this is parameterised by the Cabibbo-Kobayashi-Maskawa (CKM) matrix, $V$ (Equation 1.22 & 1.23 [1]). The elements, $V_{ij}$, act as coefficients to the interaction strength

Table 1.5 Charges attributed to SM fermion families (Tables 1.1 & 1.2) under unified EW theory.

| Fermion fields | $\mathrm{SU}(2)_L$ $T$ | $\mathrm{U}(1)_Y$ $Y$ | $\mathrm{U}(1)_Q$ $Q$ |
|---|---|---|---|
| $\begin{pmatrix} l^0 \\ l^\pm \end{pmatrix}_L$ | $\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$ | $-1$ | $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$ |
| $\left(l^\pm\right)_R$ | 0 | $-2$ | $-1$ |
| $\begin{pmatrix} q \\ q' \end{pmatrix}_L$ | $\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$ | $+1/3$ | $\begin{pmatrix} +2/3 \\ -1/3 \end{pmatrix}$ |
| $\left(q\right)_R$ | 0 | $+4/3$ | $+2/3$ |
| $\left(q'\right)_R$ | 0 | $-2/3$ | $-1/3$ |

and off-diagonal terms contain not only mixing angles but a CP-symmetry violating complex phase:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.22}$$

$$V_{\mathrm{CKM}} = \begin{pmatrix} 0.97446 \pm (10) & 0.22452 \pm (44) & 0.00365 \pm (12) \\ 0.22438 \pm (44) & 0.97359 \ \left(^{+10}_{-11}\right) & 0.04214 \pm (76) \\ 0.00896 \ \left(^{+24}_{-23}\right) & 0.04133 \pm (74) & 0.999105 \pm (032) \end{pmatrix} \tag{1.23}$$

As exclusively left-handed $W$ interactions occur, we expect to observe only left-handed neutrinos and right-handed anti-neutrinos. Observations of neutrino flavour mixing over large distances imply that neutrinos are massive, requiring different flavour versus mass eigenstates. Parameterisation of mixing in the lepton sector, namely the neutrino masses and mixing matrix, is an ongoing pursuit of physics beyond the Standard Model.

### The Higgs mechanism

The Higgs mechanism allows contributions to the new gauge boson masses with a scalar kinetic term spontaneously breaking EW symmetry to $\mathrm{U}(1)_Q$. The EW Lagrangian, in addition to coupling to the gauge fields through $D_\mu$, features a complex scalar $\mathrm{SU}(2)$ doublet $\phi$ (Equation 1.24) which may be introduced with a potential term $V(\phi)$ (Equation 1.25) appearing in the Lagrangian in the place of $\frac{m^2}{2}\phi^\dagger\phi$. The simplest allowed form in terms of $\eta$, the mass of the field, and $\rho$, a positive dimensionless constant is assumed.

If $\eta^2 > 0$, the theory mimics the massless vector boson case preserving phase invariance $\theta(x)$, its vacuum state remaining at $\phi = 0$ (Figure 1.1). However, if $\eta^2 < 0$ then a continuum of minima form with a non-zero vacuum expectation value, $v$; this meta-stable form for $\phi = 0$ taken by the Higgs field results in spontaneous symmetry breaking [7].

$$\begin{aligned} \phi &= \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \\ &= \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \end{aligned} \tag{1.24}$$

$$V(\phi) = \eta^2|\phi|^2 + \rho\left(|\phi|^2\right)^2 \tag{1.25}$$

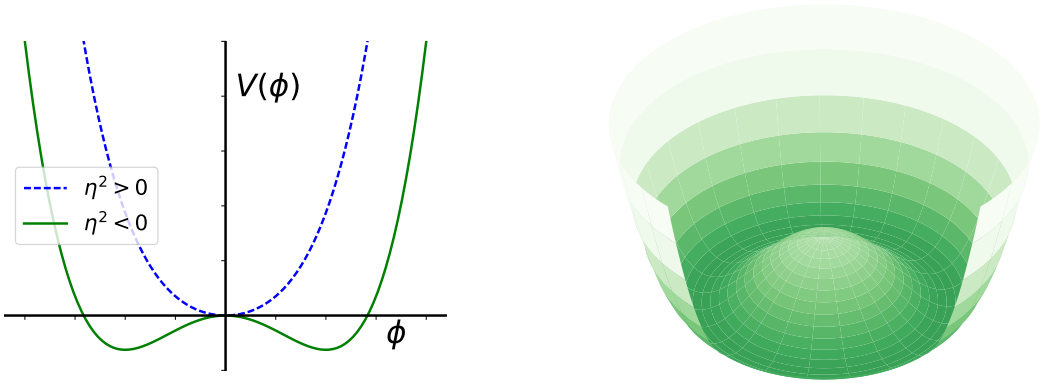



Fig. 1.1 The non-zero expectation value of $\phi$ exists at a minimum in $V(\phi)$ as shown on the left in green and on the right projected, to help visualise the transverse and longitudinal degrees of freedom, in the complex plane.

As the vacuum relaxes to a value of $|\phi| > 0$, the global U(1) gauge symmetry will be broken; choosing a gauge where $\phi_{1,2,4} = 0$ forms infinite degenerate states at $|\phi| = \frac{v}{\sqrt{2}}$. While the Lagrangian remains invariant, perturbations around the new vacuum state are not symmetric. If $h$ denotes the transverse excitations about the Higgs vacuum expectation, then the resulting theory acquires a massive spin-0 boson, $H^0$, with $m_H^2 = 2\rho v^2$, isospin $T = \frac{1}{2}$ and hypercharge $Y = 1$.

Three Goldstone bosons, $\chi$, emerge in the Lagrangian through the invariance from longitudinal degrees of freedom in $V(v)$, one for each broken generator. With a 1-to-1 correspondence of non-zero mass eigenstates to the vector bosons, and the choice of gauge (Equation 1.26), it is said the additional fields are absorbed into $W^a$ and $B$ through transformations (Equation 1.27). They then represent additional polarisation states of the vector bosons only physical for massive particles. The four-vector bosons undergo mixing

to produce the $W^{\pm}$, $Z$ and $\gamma$ fields of the SM and their respective masses, proportional to $v$ (Equation 1.28).

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+h \end{pmatrix} e^{-\frac{i}{v}\chi} \tag{1.26}$$

$$\begin{aligned} B_\mu &\to B_\mu - \frac{1}{g_b}\partial_\mu \chi \\ W^a_\mu &\to W^a_\mu - \frac{1}{g_w}\partial_\mu \chi^a + \varepsilon^{abc} W^b_\mu \chi^c \end{aligned} \tag{1.27}$$

$$\begin{aligned} W^{\pm}_\mu &\equiv \frac{1}{\sqrt{2}}\left(W^1_\mu \mp i W^2_\mu\right), \quad & Z_\mu &\equiv \frac{-g_b B_\mu + g_w W^3_\mu}{\sqrt{g_w^2 + g_b^2}}, \quad & A_\mu &\equiv \frac{g_w B_\mu + g_b W^3_\mu}{\sqrt{g_w^2 + g_b^2} W^3_\mu} \\ m^2_W &= \frac{1}{2} v g_b, & m_Z &= \frac{1}{2} v \sqrt{g_w^2 + g_b^2}, & m_A &= 0 \end{aligned} \tag{1.28}$$

The weak mixing angle $\theta_W$ between the $Z$ and $\gamma$ eigenstates of the mass matrix (Equation 1.29) can be used to define the electroweak coupling constants with respect to the QED coupling, $g_e = g_w \sin\theta_W = g_b \cos\theta_W$. The resultant co-variant derivative (Equation 1.30) ensures the massive $W^{\pm}$ and $Z$ and massless $\gamma$ consistent with observation. The couplings of the $Z$, $g_R = -Q\sin\theta_W$ and $g_R = T - Q\sin\theta_W$ which produce vector and axial-vector couplings $g_v = g_L + g_R$ and $g_a = g_L - g_R$ respectively, predict the correct fermion family dependant neutral current interactions.

$$\begin{pmatrix} W^+_\mu \\ W^-_\mu \\ Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \cos\theta_W & -\sin\theta_W \\ 0 & 0 & \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W^1_\mu \\ W^2_\mu \\ W^3_\mu \\ B_\mu \end{pmatrix} \tag{1.29}$$

$$D_\mu \phi = \begin{pmatrix} \frac{i g_w}{\sqrt{2}} W^+_\mu (h+v) \\ \left(\partial_\mu - \frac{i}{2}(g_b + g_w) Z_\mu\right)(h+v) \end{pmatrix} \tag{1.30}$$

$$\mathscr{L}_{Higgs} = \left(D_\mu \phi\right)^\dagger \left(D^\mu \phi\right) + V(\phi) \tag{1.31}$$

### Yukawa interactions

The fermion masses are generated under the exchange of hyper-charge with the Higgs field through Yukawa interactions. The spontaneously broken symmetry allows the mixing of massless $f_{L,R}$ projections of weak hyper-charged states to produce the physical particle

consistent with observation. A fermion, $f$, obtains mass proportional to the Higgs vacuum expectation value and its Yukawa coupling, $m_f \propto v g_f$, which is interpreted as an expression of the rate of mixing or virtual Higgs exchange. $g_f$ are considered free parameters of the SM.

$$\mathscr{L}_{\text{Yukawa}} = -g_f v \bar{\psi}_f \psi_f - g_f \bar{\psi}_f \psi_f h \tag{1.32}$$

### 1.1.5 Observables

Feynman diagrams are used to represent the mathematical expressions for processes between incoming and outgoing particles. Feynman rules, derived from the Lagrangian density of the SM, establish the allowed paths to contribute to a transition (Figure 1.2). Represented in the vertical and horizontal axis respectively are space and time. The arms to the exterior of the diagram represent the currents of the incoming and outgoing states. Vertices are formed where currents meet and represent particle interactions. Internal currents between vertices are considered virtual particles.

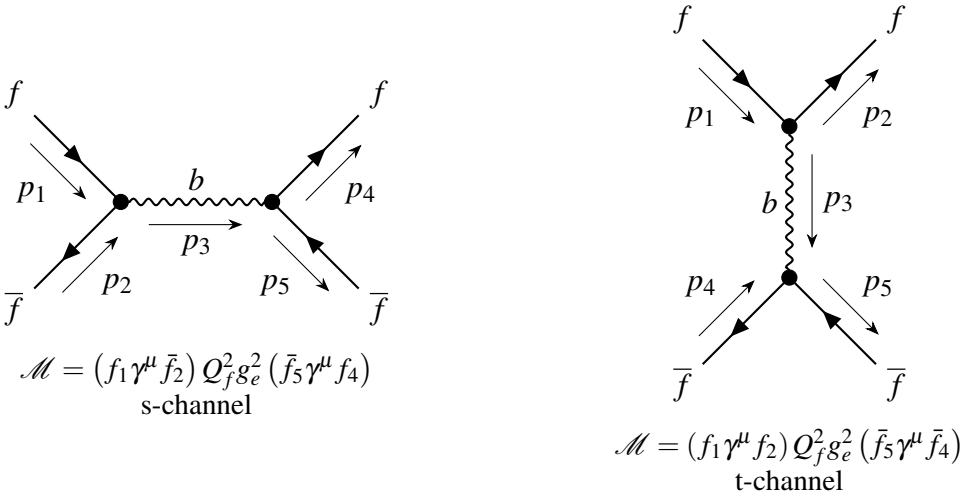



Fig. 1.2 Feynman diagrams with charged fermions, $f$, annihilating or scattering via a photon, $b$ in the s-channel (left) and t-channel (right).

Each contribution to a transition amplitude, $\mathscr{M}$, is the product of the components of a given diagram. Exchanges between fields at vertices conserve the currents of the symmetry groups within the theory, introducing factors proportional to the dimensionless coupling constants. Virtual particles, considered 'off-shell', act between the physical states and each other, integrated across all time and space. Diagrammatic classification of contributions allows the simplified probabilistic combination of their amplitudes.

### Perturbation theory

Physical scattering amplitudes can be predicted by functional integration over all distinct paths between initial and final states or approximated through time-dependent perturbation theory. The full Hamiltonian to a corresponding Lagrangian density may be broken down into free and interacting components, $H(t) = H_0 + H_{int}(t)$ where $H_{int} = -L_{int}$. An interaction may be described in terms of a Taylor expansion if $H_{int} << H_0$.

The scattering-matrix $S$ can be calculated with a Taylor expansion in powers of $H_{int}$ represented by the series of Feynman diagrams where the order corresponds to the number of vertices. Leading order (LO), based on the simplest diagrams for $\alpha \to \beta$, provides amplitudes expected from classical field theories. Higher orders contribute to $\alpha \to \beta$ with diminishing amplitudes analogous to products of combined probabilities related to the coupling strengths at each additional vertex.

$$S_{\beta\alpha} = \delta_{\beta\alpha} - 2\pi i \delta^4 \left(p_\beta - p_\alpha\right) \mathscr{M}_{\beta\alpha} \tag{1.33}$$

$$\frac{1}{\varepsilon} d\sigma' = |\mathscr{M}|^2 dx \tag{1.34}$$

Perturbation theory provides a series expansion of the potential diagrams in powers of N-couplings to predict S-matrix elements, $\mathscr{M}$ (Equation 1.33). The $\delta$-functions factorised out ensure momentum exchange and conservation. Integrating across phase space $x$ provides a cross-section, $\sigma$, which corrected for detector effects and collision environment by the factor $\varepsilon$, may be compared to the measured cross-section $\sigma'$ (Equation 1.34).

### Renormalisation

The summation of all allowed diagrams, from $N = 1 \to \infty$, would converge on the physical observable of the QFT. Truncating the series for practical computation introduces new problems. Both the convergence rate of the series and the order to which it is calculated affect the precision of any prediction. Beyond leading order, QED and QCD produce ultra-violet (UV) and infra-red (IR) divergences respectively, through self-energy terms and loop diagrams requiring integration over infinite momenta.

The theory may be regularised, parameterising out divergent terms to operate in the regime of a chosen scale through momentum cut-off or limiting dimensionality. Before regularisation, a calculation at all orders would exhibit no scale dependence. Renormalisation systematically incorporates higher-order contributions into the effective definitions of fundamental quantities themselves to provide physical predictions.

Redefinitions may be formulated such that they depend only on the physical quantities. Finite predictions of observables are then produced relative to known experimental values from a specific regime used to renormalise the theory. However, this introduces an unphysical renormalisation scale dependence in both the coupling and the matrix element.

### Running couplings

In QED, QCD and EW theory, the coupling strengths of the forces are energy scale dependant. Pair production of electrically charged virtual particles leads to vacuum polarisation around charged fermions and consequently a distance-dependent shielding of their fundamental electric charge. This shielding results in the running of the EM coupling constant, $g_e$, increasing with the energy scale of the interaction, $Q^2$, and decreasing length-scale. The running of QCD and EW coupling constants follow the opposite trend, at least up to the symmetry breaking scale, due to the inclusion of dominant boson self-coupling terms.

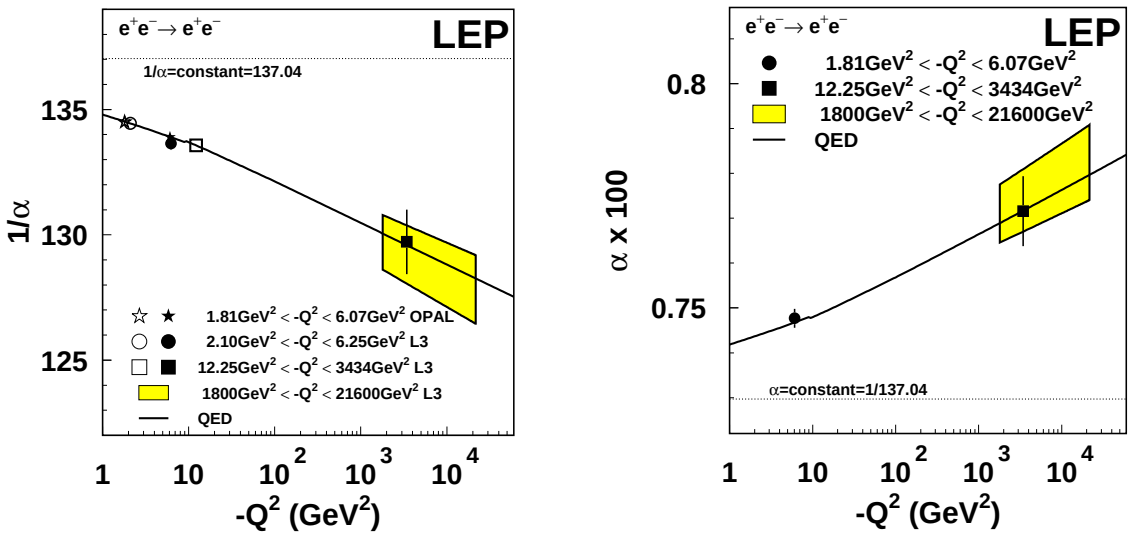



Fig. 1.3 Summary of LEP results of the running of the electromagnetic coupling [8].

The equivalent effect for colour charge results in the playoff between quark pair production and gluon self-coupling contributions (cubic and quartic), dictating the effective range of QCD. If, as in nature, $2N_f - 11N_c \leqslant 0$ ($N_f = 6, N_c = 3$) then quarks will experience asymptotic freedom such that $g_s$ increases with length-scale and with decreasing $Q^2$ [9]. As a result of the running of the strong coupling strength, the perturbative approximation for low energy interactions is invalid. Bare quarks will produce $q\bar{q}$ pairs from the vacuum to exist as colour charge-neutral states in a process known as confinement. These colourless states are known as hadrons which include $q\bar{q}$ (mesons) and $qqq\,/\,\bar{q}\bar{q}\bar{q}$ (baryons).

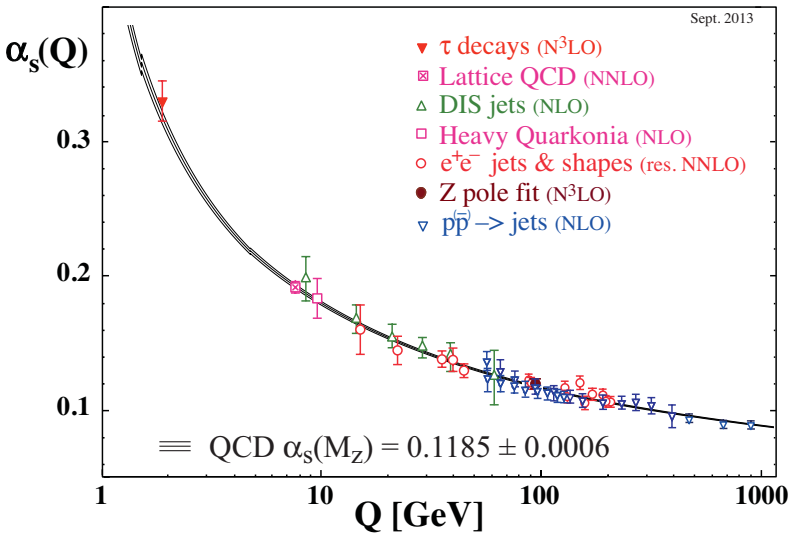



Fig. 1.4 Summary of measurements of $\alpha_S$ as a function of the energy scale $Q$ [10].

## 1.2 Hadron collider physics

Hadrons are the bound colourless states of quarks confined by gluon interactions, of which the proton is a stable example. Collisions in high energy experiments result in the 'hard', or high-$Q^2$, scattering of accelerated particles dominating interactions between composite particles, which produce additional soft-QCD processes. This results in a perturbative high energy process and a non-perturbative low energy background present in proton-proton collisions.

### 1.2.1 Factorisation theorem

At sufficiently high energies, interactions between the constituents of a proton are neglected. This allows factorisation of the perturbatively calculable partonic cross-section from that of the overall interaction, assuming asymptotic freedom for each possible set of initial states. Each contribution is weighted by the relevant parton distribution functions (PDFs), which act as the parameterisation of the contents of hadrons in the collisions taking place.

$$\sigma_{AB\to X} = \int dx_A dx_B \, f_a\left(x_a, Q^2\right) f_b\left(x_b, Q^2\right) \, \sigma_{ab\to X} \tag{1.35}$$

The factorisation theorem provides the hadronic cross-section in these terms (Equation 1.35) where $A$ and $B$ are the colliding hadrons, $a$ and $b$ are the hard scattered partons and $f_a(x_a, Q^2)$

and $f_b(x_b, Q^2)$ are the PDFs for partons $a$ and $b$. The partonic cross-section, $\sigma_{ab\to X}$, depends upon $\alpha_s(\mu_R^2)$, where $\mu_R$ is the renormalisation scale, and $\mu_F$, the factorisation scale separating long range effects absorbed into the PDFs of the proton and short range effects of the primary interaction. $X$ is independent of the number of final state particles or kinematic configuration, the cross-section is said to be inclusive.

### 1.2.2 Parton distribution functions

The hard process is described by the partonic cross-section, composed of diagrams with simplified initial states. The availability of these initial state particles in the collision may be parameterised through probability densities of the real and virtual constituents of the proton, its PDFs. These non-perturbative functions are defined in terms of the longitudinal momentum fraction carried by the parton, $x$, and depend on the scale of the interaction, $Q^2$. The functions are fitted to experimental data ($ep$, $p\bar{p}$, $pp$) and, for producing theory predictions, evolved through DGLAP equations [11] to the relevant scale.

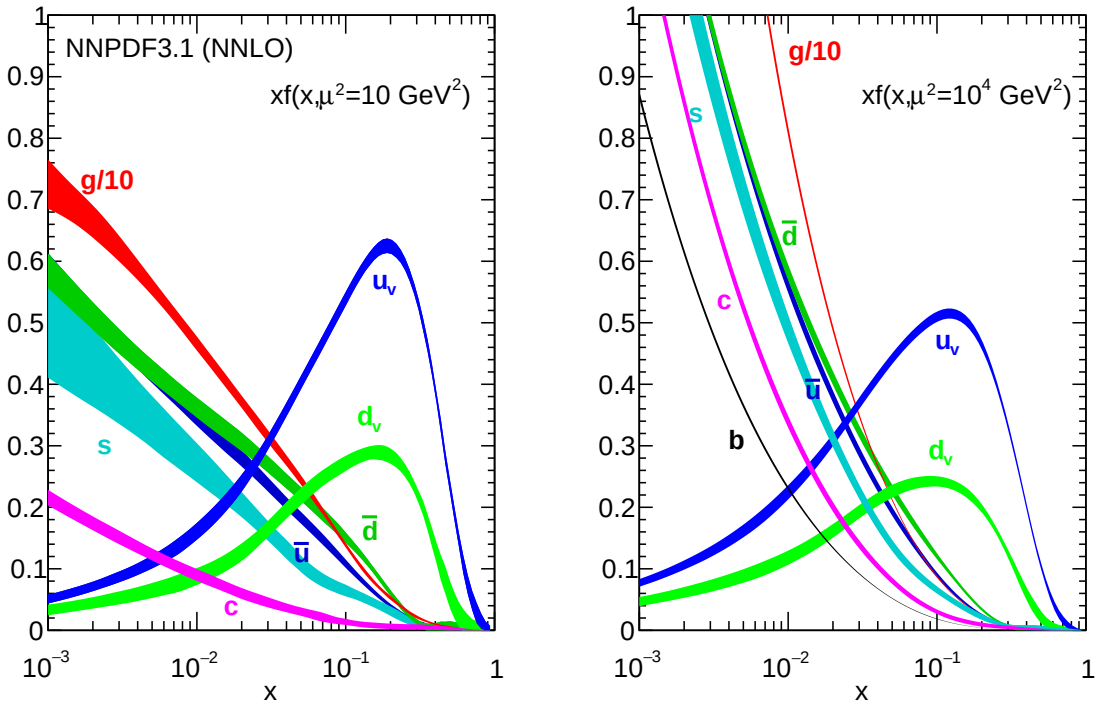



Fig. 1.5 NNPDF3.1 NNLO PDFs for the gluon and various (anti-)quark content evaluated at $Q^2 = 10, 10^4\,\text{GeV}^2$ [12].

### 1.2.3 Showering & hadronisation

The remaining partons following the $pp$ scatter will evolve, radiating photons and gluons or pair producing quarks. These additional processes require description using higher-order real-emission diagrams, thus contributing to the inclusive cross-section of the final state in question. Those occurring before and after the hard process are initial and final state radiation (ISR & FSR) respectively. The relevant emissions are shown in Figure 1.6. With the interaction scale decreasing at each emission vertex, or branching, eventually confinement results in the hadronisation of the resultant colour-charged states. Of the variety of colour neutral resonances produced, many have relatively short lifetimes and will decay in flight resulting in further radiation.

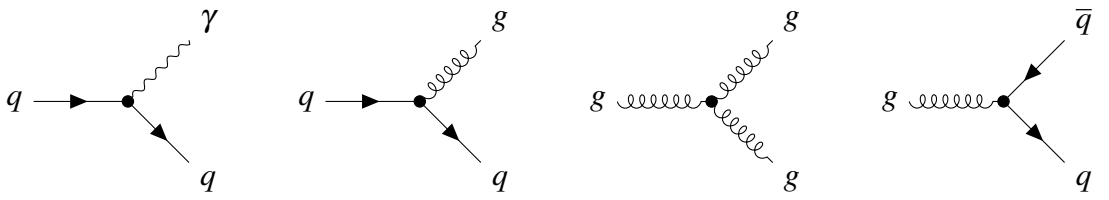



Fig. 1.6 Showering contributions of bare parton radiation or splitting.

The matrix element for gluon emission experiences infrared divergence ($E \to 0$) and collinear divergence ($\theta \to (0, \pi)$). The integral is regularised in the soft/collinear limit ($E\theta \approx p_T$), removing these divergences with a cutoff, $p_T > \Lambda$, where $\Lambda$ is the perturbative limit. At each order $N$, terms of the form $\alpha_s^N \ln^N \left(p_T^2/\Lambda^2\right)$ are produced which, for $N \to \infty$, may be combined analytically [13]. Thereby, fixed order calculations can be modified to include subsequent emission of partons generating corrected modelling of transverse momenta, at least for high energies.

In simulation, these splittings provide phase-space contributions beyond fixed-order calculations and maybe sequentially modelled through parton showering algorithms. The probabilities of the particles splitting are associated with their energy dependant couplings. The probability of evolution without branching as a function of scale is described by Sudakov Form Factors [14]. For each subsequent system at a given scale, this process will be repeated until the limit $\Lambda$ is reached. For initial-state showers, additional consideration must be made for the PDF evolution.

### 1.2.4 Computational techniques

Monte Carlo generators use stochastic modelling to predict event counts by integrating differential cross-sections over specific phase-space regions, bridging the gap between theory

and experimental measurement. Modern event generators allow the double-counting of fixed higher-order processes evolved with parton showering to be subtracted so that hard processes may be calculated independently of the showering and hadronisation procedure. The software relied upon for the sample production detailed in Chapter 6 is expanded upon below.

### NNPDF 3.1

The PDF sets used in the studies presented are accessible through the LHAPDF interface [15]. The global fits from this latest iteration of NNPDF incorporates experimental $t\bar{t}$ differential cross-sections and the $Z$ $p_T$ spectrum data while theoretical inputs now include NNLO QCD corrections. These updates, along with the independent parameterisation of charm content of the proton, demonstrate improvement to the light-quark separation and gluon precision.

Like previous versions, NNPDF 3.1 [12] implements an artificial neural network as an unbiased modeling tool, propagating experimental uncertainties and correlations from data. Uncertainties from data are used to define the variances of a multi-dimensional Gaussian distribution to provide replica sets of the input data and its statistical distribution. A projection into the PDF space is produced through a minimisation procedure for each replica set.

### POWHEG & aMC@NLO

Two NLO generators, POWHEG and aMC@NLO, are used in the work presented in this thesis. These methods differ in their treatment of the subtraction of divergent sub-leading colour terms and their choice of scale in the Sudakov form factor, resulting in differences between their predictions [16].

A fixed order matrix calculation at NLO may be provided through the POsitive Weight Hardest Emission Generator (POWHEG). Calculations using POWHEG are independent of parton showering, allowing them to be interfaced with subsequent showering MC. For $p_T$ ordered showering, the evolution scale for further emissions uses a starting point fixed to that of the $p_T$ of the original POWHEG event [17].

The aMC@NLO framework, the successor to MadGraph5, provides the automated computation of differential cross-sections up to NLO in QCD in association with parton shower matching, solving the issue of exponentiation of non-leading terms experienced by POWHEG [18]. The decay of the top quarks is performed using MADSPIN such that spin correlations for leptonic $t\bar{t}$ final states are modelled [19].

**Pythia8**

Pythia, a LO process MC generator, allows flexible use of LO PDF sets and phase space selection to simulate high statistics samples efficiently. Pythia is commonly incorporated into NLO generators to produce accurate showering and hadronisation consistent with ISR and FSR in data [20]. The string fragmentation approach implemented for hadronisation in Pythia is known as the Lund model [21]. Additionally, Pythia provides simulation of beam-remnants, a non-negligible colour-connected contribution to the underlying event. The multi-parton interactions have a shared or 'interleaved' evolution with ISR (& FSR) scaled in decreasing transverse momentum.

## 1.3 Top physics

The top quark is the highest mass fundamental particle in the SM and the only fermion at the electroweak symmetry breaking scale. Expected to provide indirect sensitivity to processes beyond the reach of current colliders through its coupling to the Higgs, the top quark holds special significance in many potential beyond SM scenarios, typically through modified Higgs phenomenology [22]. Combined with exceptionally precisely predicted behaviour, to NNLO in QCD at EW level, its unique phenomenology makes the top quark a powerful probe of high energy physics.

### 1.3.1 Production

Top quarks can be pair produced at colliders through quark anti-quark annihilation ($q\bar{q}$) and gluon-gluon fusion ($gg$). Quark-gluon production ($qg$) can occur via either mechanism, preceded by gluon radiation or gluon to quark splitting respectively, resulting in an additional quark jet in the event (Figure A.2). At 13 TeV top quarks are produced in LHC collisions with a cross-section of $\sim 830\,\text{pb}$ for $t\bar{t}$ [23, 24] and $\sim 280\,\text{pb}$ for single-$t$ [25, 26]. The channels for single top are depicted in Figure 1.7 where $q' = b$ and take place almost exclusively via the $Wtb$-vertex ($V_{tb}$ in Equation 1.22). The components of the single-$t$ cross-sections are dominated by the t-channel. Pair production occurs predominantly through gluon fusion (Figure 1.8).

### 1.3.2 Top pair asymmetry

Higher-order corrections, calculable in perturbative QCD and measurable with the top pair production cross-section, may be tested for enhancements from new physics. Previous

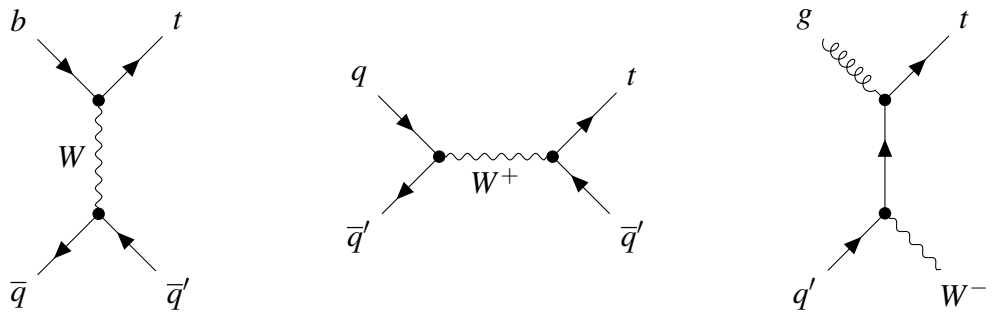



Fig. 1.7 Single-$t$ production Feynman diagrams via the t-channel (left), s-channel (centre) and $tW$ production mechanisms (right).

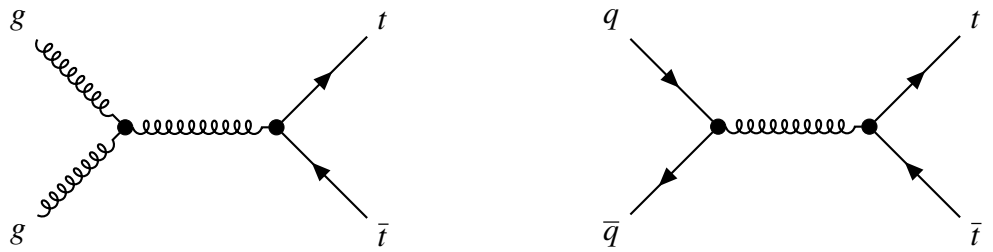



Fig. 1.8 LO top pair production Feynman diagrams via gluon fusion (left) and quark annihilation (right).

measurements by CDF and D0 at the Tevatron have shown some contention with NLO predictions which are partially resolved at NNLO [27]. Whereas colliding $p\bar{p}$ produced a forward-backwards charge asymmetry in the final state, the initial state at the LHC is symmetric in quark content with momentum fraction disparity between the valance quarks and the sea of quark anti-quarks pairs in the collision environment.

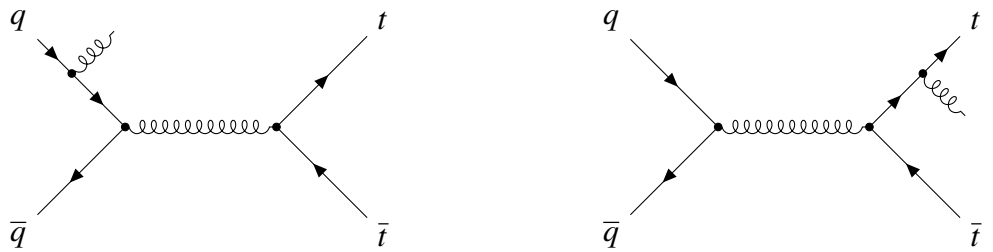



Fig. 1.9 Top pair production ISR (left) and FSR (right) Feynman diagrams which interfere to produce a negative asymmetry contribution.

The interference between Born and box diagrams (Figures 1.8 & 1.10) leads to a boost of top quarks over anti-tops relative to the beamline and a charge asymmetry arises. The initial and final state radiative processes (ISR & FSR) (Figure 1.9) interfere, generating an opposite asymmetry. The soft, virtual process dominates the real hard radiation. A positive

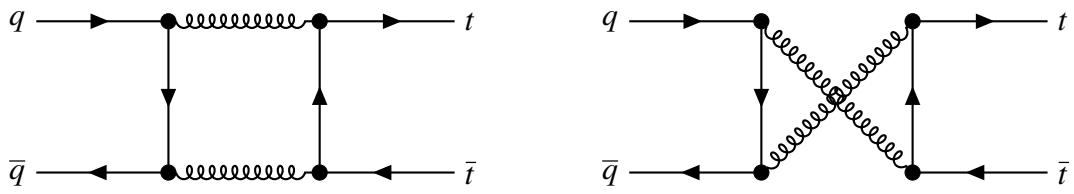



Fig. 1.10 Top pair producition box diagrams at NLO which interfere with the $q\bar{q}$ Born diagram to produce a dominant positive asymmetry.

charge asymmetry is established, which may be interpreted as quark initiated pair production processes producing a corollary momentum dependence of the products upon that of the annihilating partons of the same sign. The symmetry of the dominant gluon fusion process dilutes the quark initiated pair production processes and thus the charge asymmetry.

At the Tevatron, $p\bar{p} \to t\bar{t}$ was dominated by valence-$q$ valence-$\bar{q}$ interactions. This results in a boost of $t$ ($\bar{t}$) in the $z$-direction of the incoming $p$ ($\bar{p}$) producing a forward-backward asymmetry, $A_{FB}$. The asymmetry measured by DØ and CDF [28] is defined by Equation 1.36 where $N$ is the number of events passing the criteria for the $t\bar{t}$ rapidity difference, $\Delta y$, and each event contributes a top pair.

$$A_{FB} = \frac{N(\Delta y > 0) - N(\Delta y < 0)}{N(\Delta y > 0) + N(\Delta y < 0)}, \qquad \Delta y = y_t - y_{\bar{t}} \tag{1.36}$$

At the LHC, $pp \to t\bar{t}$ processes are dominated by symmetric $gg$ interactions. Unlike in the Tevatron $p\bar{p}$ collisions, the $q\bar{q}$ annihilations occur with asymmetric contributions in $x$ from respective PDFs. This produces a boost of $t$ in both directions along the $z$-axis from incoming protons resulting in a central-forwards asymmetry instead of a forward-backwards asymmetry. Using hermetic detectors such as ATLAS and CMS, the charge asymmetry may be defined as shown in Equation 1.37 in order to measure the relative components of $t$ and $\bar{t}$ in the longitudinal axis versus transverse plane, where products of asymmetric $q\bar{q}$ interactions are boosted to rapidities of the same sign [28].

$$A_C = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)}, \qquad \Delta|y| = |y_t| - |y_{\bar{t}}| \tag{1.37}$$

Using a detector with asymmetric coverage in $y$ such as LHCb, statistics can in part be recovered through using partial event reconstruction. Reconstructing one top at a time, without access to $\Delta y$, each contributes to another definition for the charge asymmetry, laid out in Equation 1.38 when calculated for a specific range in $y$ [29]. In Section 1.3.3, the consequences of relying on partial reconstruction are discussed along with dominant

background processes relevant to top analysis at LHCb in the $\mu b$ final state.

$$A_C = \left(\frac{N_t - N_{\bar{t}}}{N_t + N_{\bar{t}}}\right)_y \tag{1.38}$$

### 1.3.3 Decay signatures

Due to its large mass, the top quark decays before hadronisation occurs. Due to the fact that the coupling controlling the decay rate of the top to each down-type quark is dominated by the mixing element $V_{tb}$, which is approximately 1, tops decay with $\sim 100\%$ branching fraction to $Wb$. The $b$ will produce a secondary decay vertex and a jet. The $W$ will decay to a $q\bar{q}'$ or $l\nu_l$ pair (Figure 1.11). In leptonic $W$ decays, the final state $l$ provides charge reconstruction of the seeding top. The decay of the bare quarks results in the preservation of spin correlations passed to the leptons via each $W$. The dominant background for a reconstructed top decay is the $Wb$ contribution. Several processes contribute to the $W$ and associated jet final states through NLO pQCD and RunI measurements at LHCb have found the $W$+$(b,c)$ cross-sections and asymmetries consistent with the SM [30].

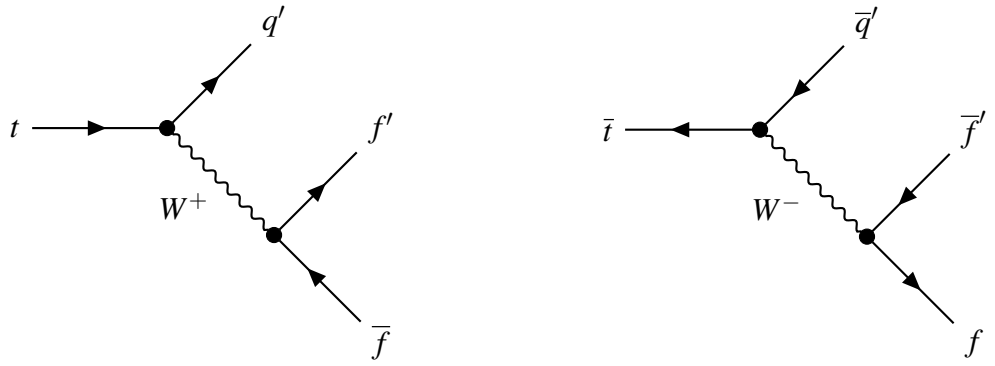



Fig. 1.11 Feynman diagrams of the decay of top quarks, whereby the extreme mass of the top quark leads to EW decay before it undergoes hadronisation.

At the cost of disentangling the top charge asymmetry from the single-$t$ contribution dominated by the positive charge of the collision environment, the relative abundance of $lb$ events provides the opportunity for differential measurements of the asymmetry into kinematic regions with enhanced sensitivity, such as high $\eta$. Top quarks were first observed in the forward region in this channel using LHCb RunI data [31], discussed in Section 6.3.1.

# Chapter 2

# Experimental environment

This chapter pertains to the design and function of the Large Hadron Collider Beauty (LHCb) experiment at the LHC. A description of the LHC, which accelerates and collides hadrons in the experiment, found in Section 2.1, is followed by details of the LHCb detector, its trigger system, the computing and software environments in Section 2.2. The LHCb experiment is a fully instrumented single-arm spectrometer specialised in the study of heavy-flavour hadrons. It is also a general-purpose detector for the forward region. LHCb is one of the four largest experiments at the European Organisation for Nuclear Research (CERN) which exploit high energy particle collisions at the LHC.

## 2.1 The Large Hadron Collider

The LHC is a synchrotron accelerator housed in an approximately circular tunnel, 27 km in circumference, at around 100 m below the French-Swiss border. The LHC is the highest-energy particle accelerator in the world. It provides proton-proton collisions for the detector experiments on the LHC ring at centre of mass energies of initially 7 & 8 TeV in Run I and most recently 13 TeV for Run II. Plans to reach 14 TeV are well underway, including work on the individual experiments to enable them to take advantage of the prospective running capabilities of the LHC. The work presented throughout this thesis is based on the study of the expected performance in and analysis of Run II data taken by LHCb.

### 2.1.1 Accelerator complex

Protons are provided from stored hydrogen gas which is first ionised and then sent through a series of accelerators. First, the Linear Accelerator (LINAC) 2 brings the proton energy to 50 MeV. Next, the Proton Synchrotron Booster (PSB) accelerates the protons to 1.4 GeV and provides a transverse emittance required by the Proton Synchrotron (PS). The PS then further accelerates the protons to 25 GeV and produces a longitudinal structure or bunch train beam. These beams are fed into the Super Proton Synchrotron (SPS) which accelerates the protons to 450 GeV and creates beams with structured bunches for injection to the LHC [32].

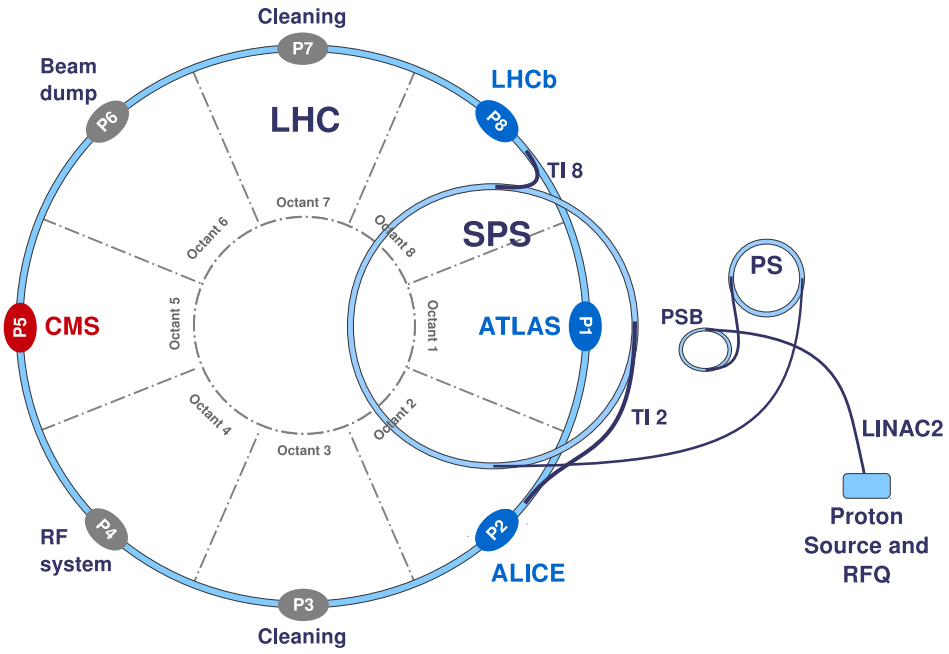



Fig. 2.1 Schematic view of the CERN accelerator complex (not to scale) [33].

The LHC beams consist of bunches of protons, designed to be separated by a 25 ns time interval providing a crossing frequency of 40 MHz. Each bunch contains $\sim 1.15 \times 10^{11}$ protons, and there are 2808 bunches per beam. The beams travel through ultra-high vacuum conditions of pressures $\mathcal{O}(10^{-9}\,\text{mbar})$. They are deflected around the ring using 1232 superconducting dipole magnets, each producing a field of 8.4 T and cooled to 1.9 K using super-fluid helium. The protons are accelerated through Radio Frequency (RF) cavities around the ring up to 6.5 TeV to provide a centre of mass energy of 13 TeV. The beams are typically a few millimetres wide in the transverse plane and are focused to a width of approximately 14 µm using quadrupole magnets at the collision points.

Luminosity is a measure of collision frequency per unit area ($\text{cm}^2$). The LHC luminosity, $\mathscr{L}$ (Equation 2.1), can be expressed in terms of the beam parameters, where $f$ is the beam

crossing frequency, $N_1$ and $N_2$ are the numbers of particles in each colliding bunch, and $\sigma_x$ and $\sigma_y$ describe the transverse beam profiles in the horizontal and vertical planes.

$$\mathscr{L} = f\frac{N_1 N_2}{4\pi\sigma_x\sigma_y} \tag{2.1}$$

With a known integrated luminosity and the yield found for a particular process, the respective cross-section can be calculated.

### 2.1.2 Run II performance

Accelerating two counter-rotating beams of tightly packed proton bunches through the LHC and forcing them to collide at points surrounded by particle detectors provides data describing particle interactions through their products and properties. Maximised luminosity is the aim of ATLAS [34] and CMS [35], the general-purpose detectors (GPDs) which have accumulated $\sim 140\,\text{fb}^{-1}$ each in Run II. They achieve these results with a higher $\mu$, the average number of collisions per bunch-crossing. Additional interactions, where $\mu > 1$, and their effects are collectively known as pile-up.

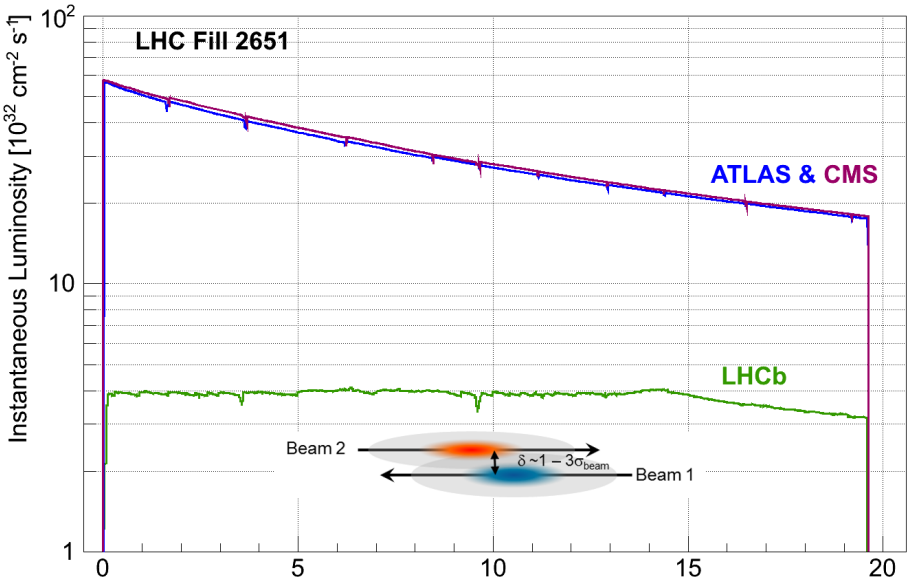



Fig. 2.2 Development of the instantaneous luminosity for ATLAS, CMS and LHCb during LHC fill 2651 demonstrating the adjustment to the transverse overlap of the beams at LHCb used to minismise the effects of luminosity decay [36].

Pile-up would degrade the primary and secondary vertex reconstruction precision of LHCb, which is required to isolate the signal from background processes. To this end, LHCb is designed to run at a lower luminosity by offsetting the colliding beams, reducing the active

collision area and thus the pile-up. LHCb employs luminosity levelling; as the proton beams diminish, the beam-beam offset may be reduced to maintain constant luminosity throughout data taking (Figure 2.2). This allows LHCb to maintain its trigger configuration throughout a fill and reduce systematic uncertainties related to changes in detector occupancy [36].

The beam profile may be monitored using the dependence of the collision rate upon beam displacement. Vertex reconstruction from protons interacting with residual gas released into the vacuum around the interaction region provides a secondary measurement. The intersection of the beam profiles allows the luminosity to be calculated using Equation 2.1. For 2015 and 2016 data, the integrated luminosity had an uncertainty of 3.9% [37].

In Run II, LHCb has lowered its $\mu$ from 1.7 to 1.1 in order to reduce systematic uncertainties associated with the mis-association of particles. The reduction of $\mu$ also increases tracking and trigger efficiencies through reduced average occupancy and limits radiation damage to detectors adjacent to the beamline. The impact on the data rate is counteracted by the LHC wide halved bunch spacing to 25ns moving to Run II; the number of interactions per bunch crossing decreases by a factor of 1.5 while the number of crossings for any given data-taking period doubles. LHCb has a data set corresponding to an integrated luminosity ($L_{int}$) of $5.4\,\text{fb}^{-1}$ from 13 TeV collisions in Run II.

## 2.2 The LHCb detector

A spectrometer reconstructs the momentum of charged particles with a tracking system flanking a bending magnet. Using Equation 2.2, a Lorentz invariant angular coordinate, rapidity ($y$), may be approximated using the angle from the beamline ($\theta$), to a value called pseudorapidity ($\eta$). In the furthest positive pseudorapidity, $\eta$ (Equation 2.2), a detector will be oriented in planes perpendicular to the $z$-axis in layers downstream of the collisions. This forward geometry allows the products of interactions with asymmetric initial momenta to be reconstructed. In the case of LHCb, this design was originally to exploit the forward-backwards dominated production of $b\bar{b}$. The LHCb detector only extends downstream with asymmetric pseudorapidity coverage to fit within the dimensions of the cavern vacated by the LEP experiment DELPHI. It is hence known as a single-arm spectrometer.

$$\eta \equiv \frac{1}{2}\ln\left(\frac{|\vec{p}|+p_z}{|\vec{p}|-p_z}\right) = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right), \quad y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) \tag{2.2}$$

Surrounding the collision point at the far left of Figure 2.3, a retractable silicon strip tracker, the vertex locator (VELO), provides precise charged particle hit reconstruction and discrimination between primary and secondary vertices. In addition to the tracking stations

on either side of the magnet, Ring Imaging Cherenkov (RICH) detectors are included for reconstructing the masses of pion, kaon and proton candidates for particle identification (PID). In addition to the specialised PID sub-detectors, electromagnetic and hadronic calorimetry systems provide energy reconstruction and a veto based on information from sub-detectors lying at increasing interaction lengths along the beamline. At the furthest end of the detector from the VELO are the muon stations, the final tracking sub-detectors sandwiched by iron filters which provide shielding from hadronic showers.

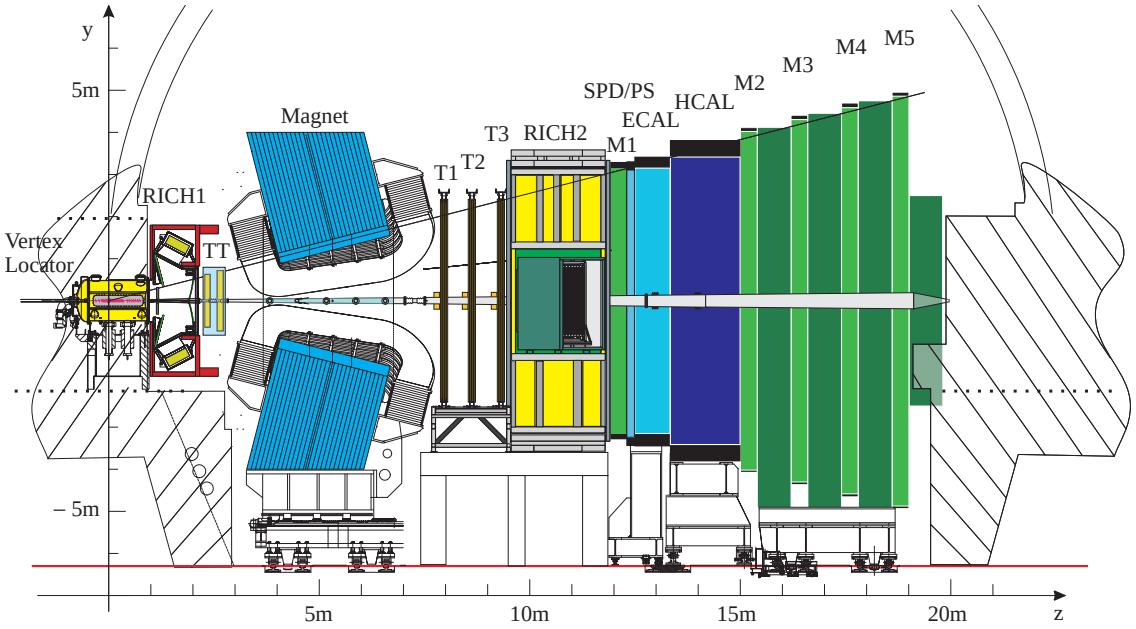



Fig. 2.3 LHCb detector layout, showing the Vertex Locator (VELO), the dipole magnet, the two RICH detectors, the four tracking stations TT and T1-T3, the Scintillating Pad Detector (SPD), Preshower (PS), Electromagnetic (ECAL) and Hadronic (HCAL) calorimeters, and the five muon stations M1-M5 all with respect to the y and z coordinate axes as indicated in the right-handed framework [38].

The LHCb coordinate system is defined as follows. Collisions take place at $z \sim 0$ at the interaction region to the left of Figure 2.3 surrounded by the VELO. The angular acceptance of the detector, covering $\pm 300$ and $\pm 250\,\text{mrad}$ in $x$- & $y$-planes respectively, lies in the positive $z$-direction. The extent of the fully instrumented acceptance at LHCb corresponds to a range of $2.0 < \eta < 4.5$ ($< 4\%$ solid angle) . The direction of particles propagating along the $z$-axis from the interaction point denotes the forward direction, positive $\eta$, where moving from left to right in Figure 2.3 is described as upstream to downstream.

The detector elements of particular relevance to this work are: the VELO, allowing $c$- and $b$-hadron tagging from the emergence of secondary decay vertices, their $\mathcal{O}$(ps) lifetime resulting in an average flight distance $\sim 1\,\text{cm}$; a tracking system providing momentum, $p$, of charged particles with hits in trackers either side of the magnet; the calorimeter systems

which provide improved energy resolution in reconstructed particle showers as well as vetoing against mis-ID hadrons punching through the calorimeters to the muon stations, where 250 GeV pions and kaons have $\sim 1\%$ punch through rate [39]; the muon stations in turn match hits with existing tracks to identify muons.

### 2.2.1 Vertex locator

The VELO resides within the vacuum surrounding the interaction region where the beams cross. The trajectories of particles resulting from the $pp$-collisions are precisely reconstructed into primary and secondary vertices using information from this specialised tracker. Silicon microstrip technology, as used in the VELO, relies on high energy charged particles creating electron holes in the depleted detector material. The electric current produced in the presence of an electron-hole by the high voltage (HV) across the strip indicates the point of incidence of a particle upon the detector plane.

Each sensor is 300 μm thick with a semi-circular layout extending from 42 mm radius down to 8 mm with an increased strip density towards the beamline (Figure 2.4). A module is made up of back-to-back sensors with radial and concentric silicon strips providing ($\phi$,$r$)-measurements which, when combined with the ($r$,$z$)-coordinates of a module, reconstructs hits in 3D space. The resulting tracks are used to reconstruct primary vertices to within $\mathcal{O}(10\,\mu\text{m})$ and provide an impact parameter resolution on the same scale [40]. Strip detectors are susceptible to producing ghosts, which result from alternate combinations of coordinates from concurrent hits in a given detector plane being included alongside the real or noise-based hits. This effect correlates track reconstruction ghost-rate with the average detector occupancy [40].

The VELO comprises 21 opposing pairs of modules staggered along the $z$-axis with the modules, each providing spatial reconstruction in its cylindrical geometry. The left and right arrays of modules are designed to overlap (Figures 2.4 & 2.5) to reduce edge-effects and regions of incomplete detector geometry. Additionally, it provides information for detector alignment, known to $\mathcal{O}(1\,\mu\text{m})$ with individual hit resolution on a comparable scale. The modules are arranged to provide a $1.6 < \eta < 4.9$ coverage of tracks produced from vertices within the central $\pm 2\sigma$ (Figure 2.5) of the interaction region such that charged particles inside the nominal LHCb acceptance cross at least three modules [40].

For protection from the EM fields of the LHC and to prevent potential gas leaks polluting its vacuum, the VELO is housed in a corrugated aluminium box of $300\,\mu$m thickness known as the RF-foil, containing a secondary vacuum with a pressure of $2 \times 10^{-7}$ mbar. The VELO and RF-foil combined make up 17.5% of the average radiation length of particles traversing them [41]. Before the beam has been properly focused for data-taking, at such proximity,

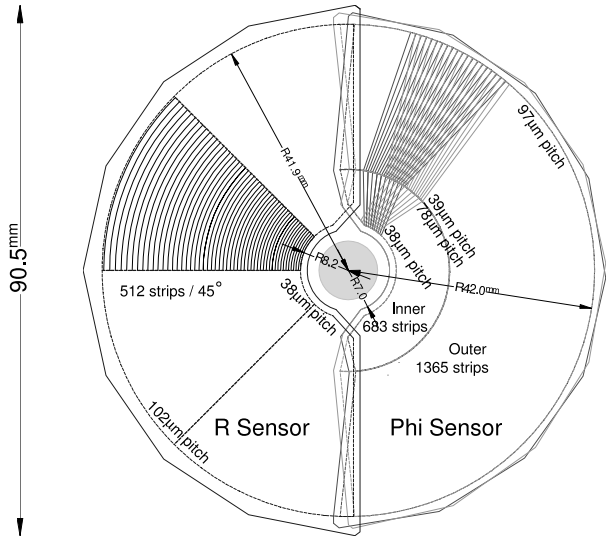



Fig. 2.4 Silicon strip arrangement in the $R$ & $\phi$ sensor sides to each VELO module, where the relative position of the left and right halves during data taking is displayed [41].

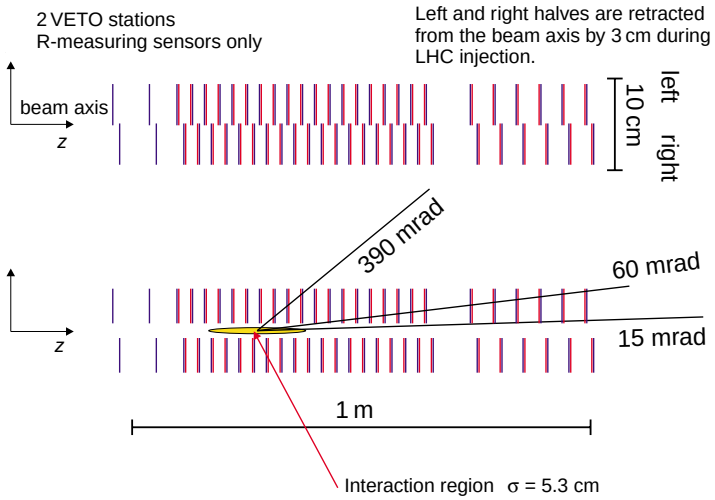



Fig. 2.5 Detector arrangement along the beam axis, where the VELO setup seen from above (top) indicates the overlap between the left and right halves and a cross section of the setup at $x = 0$ (bottom) shows the nominal position of the interaction area [42].

the VELO must remain retracted to a distance of 29 mm for its protection before returning to its operational position (Figures 2.4 & 2.5). An evaporative $CO_2$ cooling system is used to operate at the voltage required for $> 99\%$ hit efficiency while remaining under -5$C^\circ$, not only preventing thermal runaway but maintaining a 20:1 signal to noise ratio [42].

The VELO was originally intended for use only during Run I. During its total operation, the VELO will have been exposed to $5 \times 10^{13} \cdot 1\,\text{MeV}$ neutron equivalents $/\text{cm}^2$ per $\text{fb}^{-1}$ with a heavy radial dependence of the dose across the detector $\propto r^{-1.9}$. Despite having operated at twice the design luminosity of Run I, monitoring for radiation damage showed no degradation to the tracking efficiency within $\pm 0.3\%$ and the VELO remained fully operational throughout Run II [40].

### 2.2.2 Tracking

By providing vertexing capabilities of LHCb, the VELO forms an important component within the LHCb tracking systems. The initial trajectories and points of association are established between the tracks in the rest of the detector. Downstream of the interaction region is the Silicon Tracker (ST), composed of sub-detectors on either side of the magnet: a large-area silicon-strip detector at each of the two upstream stations, the Tracker Turicensis (TT) a & b; downstream of the magnet, a system of smaller area trackers, known as the Inner Tracker (IT), concentrated around the beamline at high $\eta$ reside at each of the T1-3 stations. Where much lower occupancy is expected in the remainder of the acceptance of T1-3, the stations are filled with the straw drift tubes of the Outer Tracker (OT).

Combining tracks reconstructed on either side of the magnet provides a measurement of particle charge to momentum ratio based on their deflection through the known $B$-field. The dipole magnet between the TT and T1 has a bending power of about 4 Tm integrated from 0-10 m in $z$ (Figure 2.6). The magnet-down and magnet-up configurations correspond to the direction of the $y$-component of the magnetic field. Each is used for approximately half of data taking to constrain detection asymmetries to $\mathcal{O}(10^{-3})$ [43].

The tracking system provides a measurement of charged particle momentum precision, $\delta(p)/p \sim 0.5\text{-}1.0\%$ for $p$ up to 200 GeV. The impact parameter (IP), the minimum perpendicular distance of a track to a primary vertex, is determined with a resolution of $(15 + 29/p_T)\,\mu\text{m}$, where $p_T$ is the component of the momentum transverse to the beam, in GeV [40]. Tracks are classified based on the presence of hits in each sub-detector (Figure 2.6): VELO tracks have hits only in the VELO; upstream tracks have hits in both the VELO and TT; long tracks have hits present in the VELO, TT and T-stations; downstream tracks have hits in the TT and T-stations.

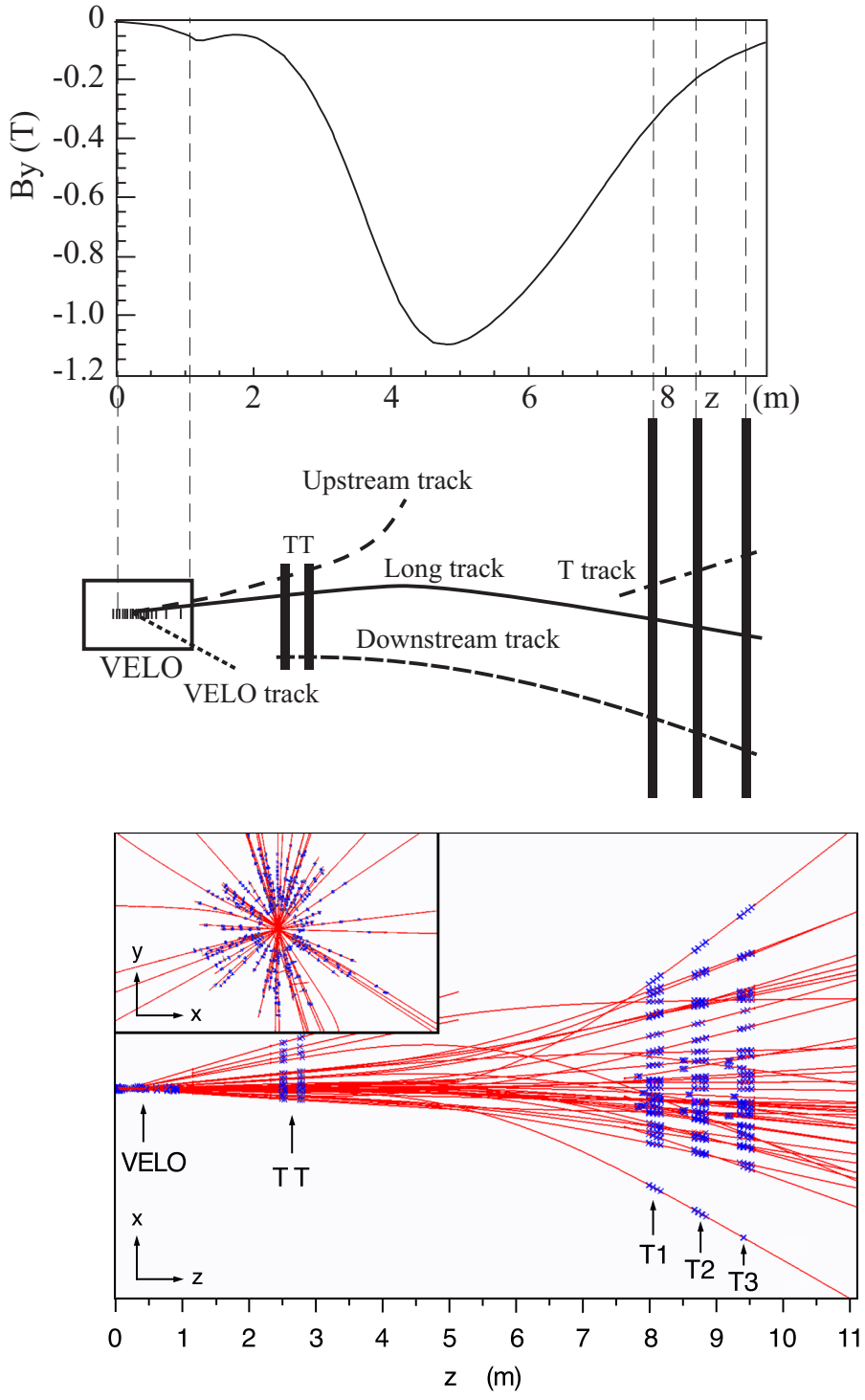



Fig. 2.6 A schematic illustration of the various track types: long, upstream, downstream, VELO and T-tracks, with the vertical component of the $B$-field ($B_y$) plotted as a function of the $z$-coordinate aligned to a display of hits and tracks reconstructed viewed from above with an insert zoomed into the VELO in the transverse plane [36].

### Tracker Turicensis

The TT ensures detection of low momentum particles deflected outside of the remaining detector acceptance; these upstream tracks are reconstructed in the trigger (Section 2.2.6). The two TT stations, each containing a pair of layers, are separated by a distance of 0.3 m in the $z$-direction. The layers comprise half modules, each made up of nine columns of eleven 500 µm thick silicon sensors. The columns extend from the outer vertical edges to either side of the ($x$,$z$)-plane. Adjacent modules are staggered $\sim 1$ cm along the $z$-axis and four half columns, two above and below the beamline centred at $x = 0$, of five sensors each cover the remaining area. As in the VELO, the overlap provides information to the alignment procedure [38].

With the dimensions shown in Figure 2.7, the TT operates an active area of 8.4 m$^2$. Strips in the second and third layers are rotated by $\pm 5^\circ$ with respect to those in the first and fourth layers to provide both coordinates in the transverse plane (Figure 2.7). The third and fourth layers, TTb, have four additional sensors, one in each of the half-columns above and below the beamline. The overall material of the TT contributes just $\sim 4\%$ of the average track radiation length, and its single hit resolution may be measured at a value of 59 µm [44]. The residual magnetic field means the TT alone provides tracks with a 30% $p_T$ resolution which is sufficient for simplified reconstruction and track segment matching (Section 2.2.6). When combined with a VELO track to produce upstream tracks, the momentum resolution is improved to 15%, enabling tracks to be extrapolated to the T-stations for full reconstruction.

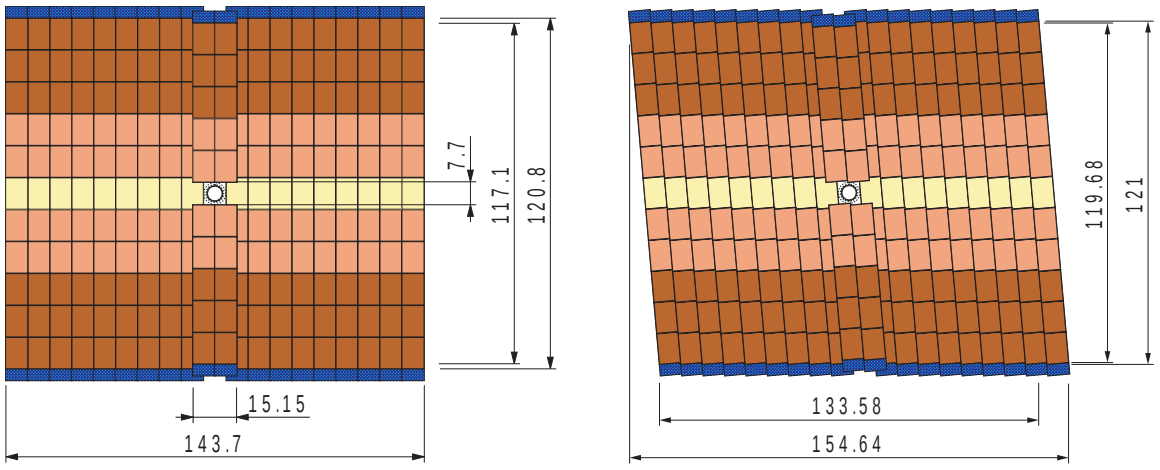



Fig. 2.7 Layout of the TTa $x$- and $u$-layers (+5°) where dimensions labelled are in cm, a corresponding TTb station has (-5°) $v$- and $x$- layers moving downstream [38].

### Dipole magnet

A room temperature dipole magnet is located between the TT and T1. Two saddle-shaped coils are positioned opposite one another above and below the beamline (Figure 2.8); these provide LHCb with 4 Tm bending power. Each $B_y$ configuration, dictated by the polarity of the dipole, is mapped at a precision of $< 4 \times 10^{-4}$ T from the VELO to RICH2 to within $\mathcal{O}(1\,\text{mm})$. The agreement between the measured field and the calculated field is better than 1%, and the variation due to hysteresis effects from the regular polarity flips is limited to at least the same precision [41]. There is a residual field $< 50$ mT around the RICH systems with a peak field strength of 1.1 T between the TT and the T1 stations (Figure 2.6).

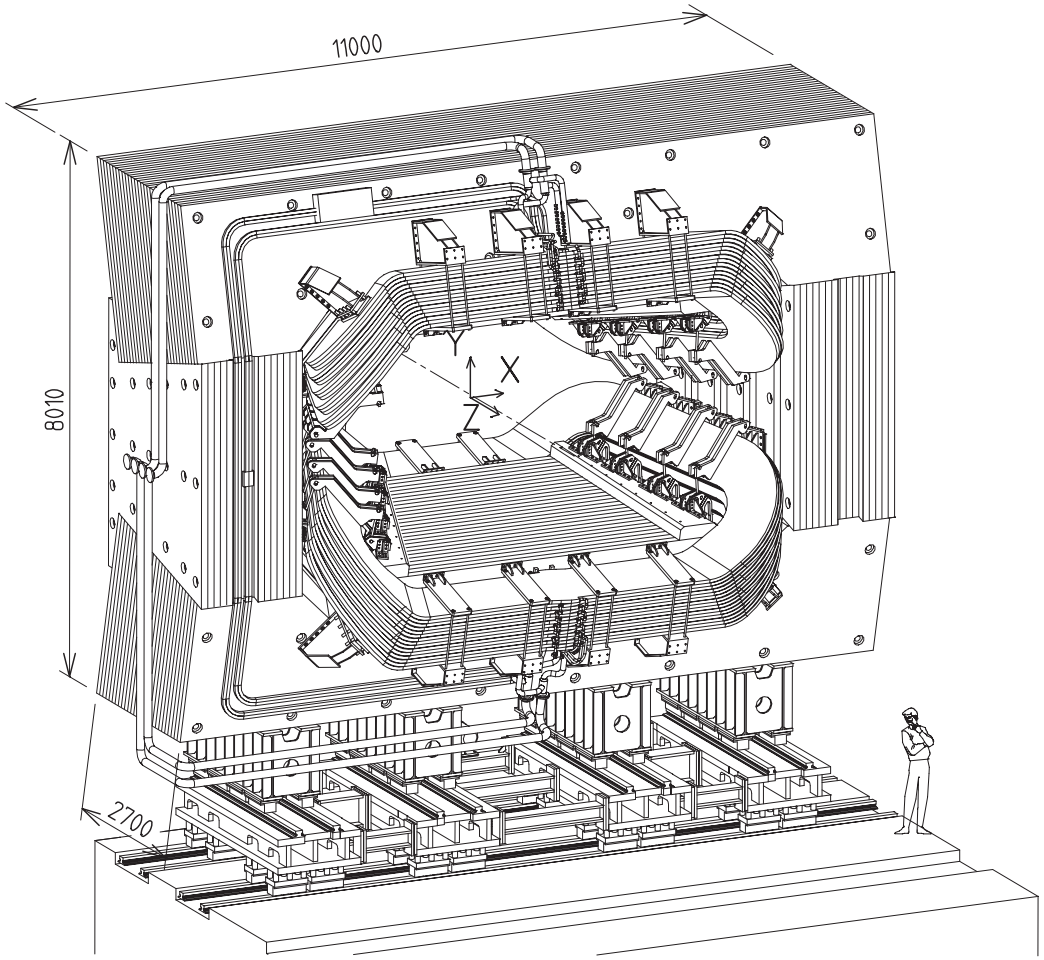



Fig. 2.8 Perspective view of the LHCb dipole magnet from downstream, in its yolk with its current and water connections, where the interaction point lies behind the magnet [38].

### Inner tracker

Found in the regions of T1-3 directly around the beam-pipe, the IT covers only 2% of the overall tracking acceptance. However, at such high charged particle flux, with around 20% of tracks passing through the IT, drift tube technology would suffer from inefficiency due to increased occupancy. The increased granularity and reduced latency offered through the use of Si-strip technology alleviates this problem. The active area of $44.0\,\mathrm{m}^2$ is positioned where both the track density and the impact of spatial resolution on momentum determination is highest to provide comparable performance to the TT for the highest momentum tracks [41].

The IT, located in all three T-stations, covers a reduced cross-shaped region made of four individual segments: one row of modules above and below; two rows of modules left and right of the beampipe (Figure 2.9). One-sensor modules are 320 µm thick; those for two-sensor modules are 410 µm thick. They correspond to $\sim 3.5\%$ of the mean radiation length of a track passing through the IT, providing minimal impact on the material budget. The silicon tracker (ST) detectors, the TT and IT, were optimised to achieve single-hit resolutions of around 50 µm; sufficient for the momentum resolution to be dominated by multiple scattering over the full track momentum range [41].

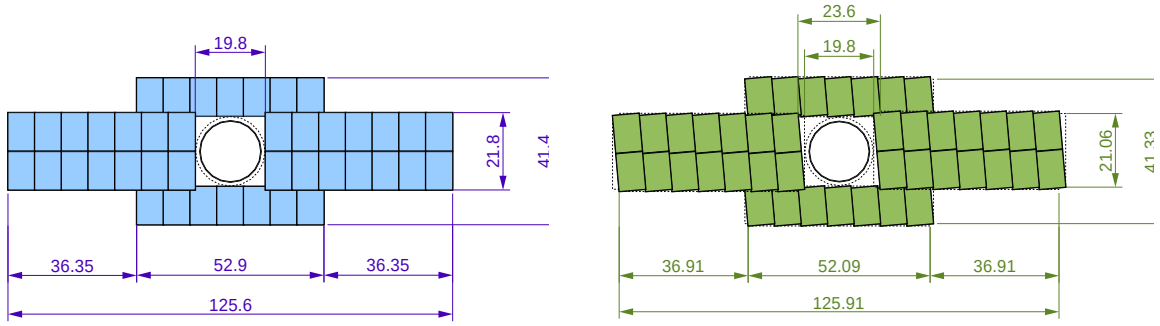



Fig. 2.9 Layout of the IT $x$- and $u$-layers (+5°) where dimensions labelled are in cm, the corresponding (-5°) $v$- and $x'$- layers moving downstream, all of which are combined to provide four layer coverage (mimicking TTa+TTb) per T-station: 1-3 [38].

### Outer tracker

The OT is a drift-time tracker; these detectors use the relative time of charge collection in layers of electrically biased inert gas-filled straws to improve spatial resolution to below the radius of the straws themselves. Each station consists of 55,000 straw channels covering an area of around $597 \times 485\,\mathrm{cm}$. Three OT stations, contained in the combined IT and OT stations T1-3, define the tracking acceptances $2.0 < \eta < 4.5$ vertically and $1.8 < \eta < 3.4$

horizontally. They use the same $\pm 5^\circ$ rotated geometry of staggered layers (Figure 2.10) as in the TT and the IT at the centres of T1-3.

Relativistic charged particles traversing the tubes ionise the gas inside. A high potential difference results in an avalanche effect as liberated electrons drift to the anode wire at the centre providing an electrical signal in the detector. Modules are composed of two layers of straws (Figure 2.10) and the delay in signal between them can be used to determine the point at which the track traverses the detector. For the closely overlapped planes in a module, the sum of drift radii is a constant, thus providing the drift times and trajectory from a tangent to circles centred on each straw with the reconstructed drift radii.

The straws contain a 25 µm diameter gold-plated tungsten wire at their centre and have an inner diameter of 4.9 mm. The straws are pressurised with a mixture of argon and carbon dioxide with a small fraction of oxygen included offsetting degradation effects. The time between ionisation and detection is dominated by the drift-time and known to be $\sim 50$ ns. As the drift time of the OT is around double the minimum LHC bunch spacing, the effects of spillover (collisions from adjacent beam crossings) have become more significant in Run II. The single hit resolution for the OT of 220 µm provides a $\delta(p)/p \approx 0.4\%$ with an average event occupancy of 13% in Run II [41, 45]. This performance is maintained despite having been designed with occupancy at nominal luminosity limited to 10% in Run I.

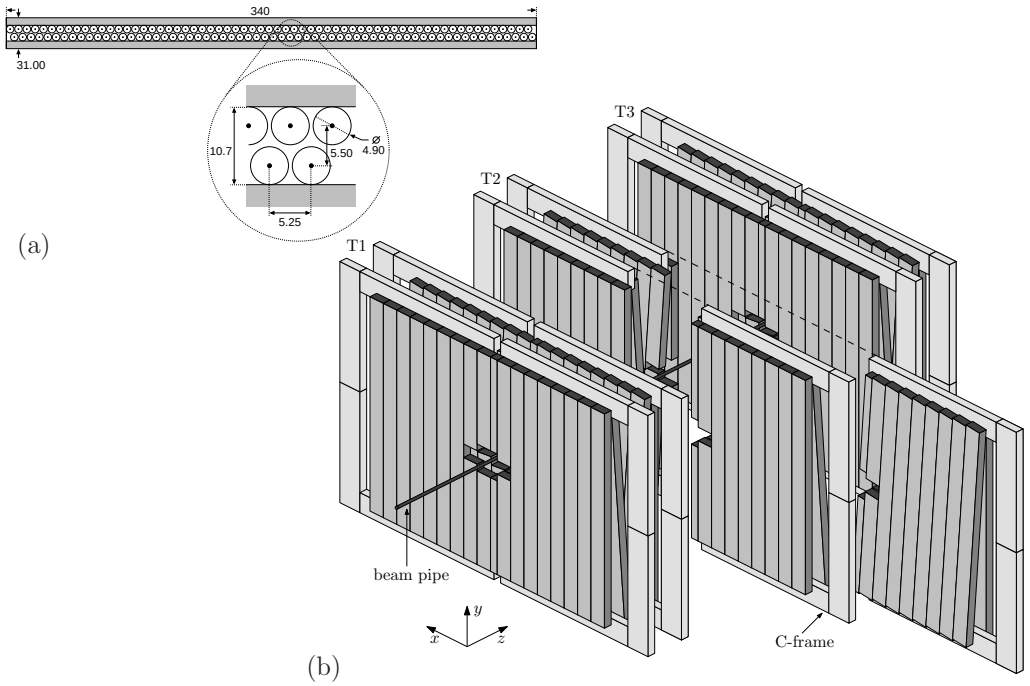



Fig. 2.10 Structure of the OT stations with a (a) cross-section of the paired mono-layers and (b) perspective view of the straw-tube module arrangement along the $z$-direction in the T1-3 stations [46].

### 2.2.3 RICH detectors

Charged hadrons are identified using two ring-imaging Cherenkov detectors. The Cherenkov radiation is the name for the cones of light produced from a medium due to charged particles moving at superluminal velocities through it. Equation 2.3 defines the polar angle of the cone in terms of $n$, the refractive index of the material, and $\beta = v/c$ where $v$ is the velocity of the particle.

$$\cos\theta_C = 1/n\beta \tag{2.3}$$

A measured velocity can be combined with the momentum information from the tracking system to estimate a mass and hence the particle species. RICH1 is positioned directly downstream of the VELO, just ahead of the TT at $\sim 1\,\text{m}$ along the $z$-axis. RICH2 lies between T3 and the calorimeters at 9.5 m in $z$. Each detector is housed in magnetic shielding to reduce the field strength to $< 2.4\,\text{mT}$ and $< 0.6\,\text{mT}$ respectively [47].

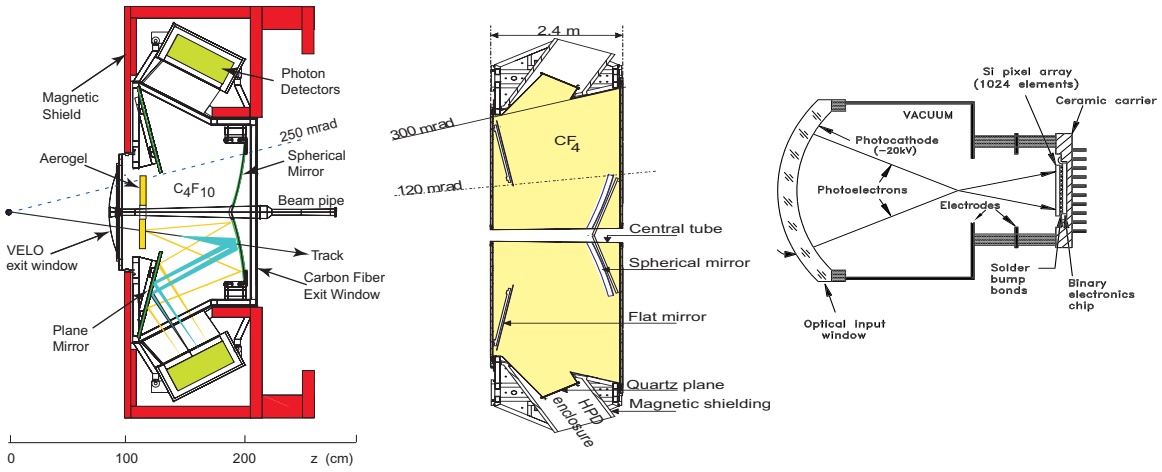



Fig. 2.11 Schematic layout of the RICH 1 (left) & 2 (centre) detectors as seen from above, the relative scale as indicated at approximately 2:1, and a schematic of a hybrid photo-detector (right) [41].

Each RICH detector at LHCb is optimised for particle ID across a specific momentum range by using radiating materials of specific refractive index (gaseous $C_4F_{10}$ and $CF_4$, where the relative indices are $n_{(C_4F)} \approx 1.7\,n_{(C_4F_{10})}$). RICH1 provides kaon versus pion discrimination for momenta between 2-40 GeV while RICH2 extends sensitivity to 15-100 GeV [47]. Spherical mirrors are used to focus the Cherenkov light emitted, with flat mirrors for redirection, onto photo-detectors (Figure 2.11). Both RICHs use hybrid pixel photon detectors of $500 \times 500$ μm pixels, able to distinguish individual photons with high-efficiency and their $\mathcal{O}$(mrad) resolution distinguishes pions and kaons to the $3\sigma$ level overall [41]. However, performance is strongly dependant on track multiplicity in an event, particularly for $>$50 GeV kaons [47].

### 2.2.4 Calorimeters

A system of four sub-detectors measures the energy of both charged and neutral particles. These include the scintillator pad detector (SPD), the and pre-shower (PS) and the electromagnetic and hadronic calorimeters (ECAL & HCAL). Scintillators exhibit luminescence when excited by ionising radiation. The detectors each use scintillating materials, some with a relatively short interaction length, to produce showers through the detectors. When absorbed, these cascades of secondary particles produce additional photons to be collected by photo-multiplier tubes (PMTs).

The scintillation light is transmitted to photo-multiplier tubes by wavelength shifting (WLS) clear plastic fibres running through each module. The scale of the scintillation response provides information regarding the particle shower energy deposition in the calorimeter. In Equation 2.4, where $E_i$ is the energy deposited and $\theta_i$ is the angle from the $z$-axis to calorimeter cell $i$ for $2 \times 2$ cell clusters, the definition for transverse energy, combining corresponding ECAL and HCAL cells, is used for hadron, electron and photon candidate trigger criteria in hardware-based Level-0 [48]. The ECAL and HCAL employ an LED calibration system to compensate for PMT gain drift through active HV corrections [49].

$$E_T = \sum_{i=1}^{4} E_i \sin\theta_i \tag{2.4}$$

The detectors utilise alternate layers of scintillation material and absorbers. This design further reduces the radiation length and are known as sampling-calorimeters. As well as using layers of absorbers within the detectors themselves, additional material designed to instigate showering or shield subsequent layers is positioned between the SPD and PS as well as each of the muon stations downstream of the HCAL [41]. EM showers are produced through Bremsstrahlung and pair-production, whereas hadronic showers proceed by the strong interaction (Chapter 1).

The extent to which a shower penetrates the detector depends on the radiation length of the material or, in the case of hadrons, the nuclear absorption length of the material, which is typically longer. As components of the PID system, the calorimeters form layers consecutively exceeding interaction lengths of different particle types (Figure 2.12). In this way, LHCb can differentiate between hadrons and leptons and, in combination with the tracking system, between charged and neutral particles [41].

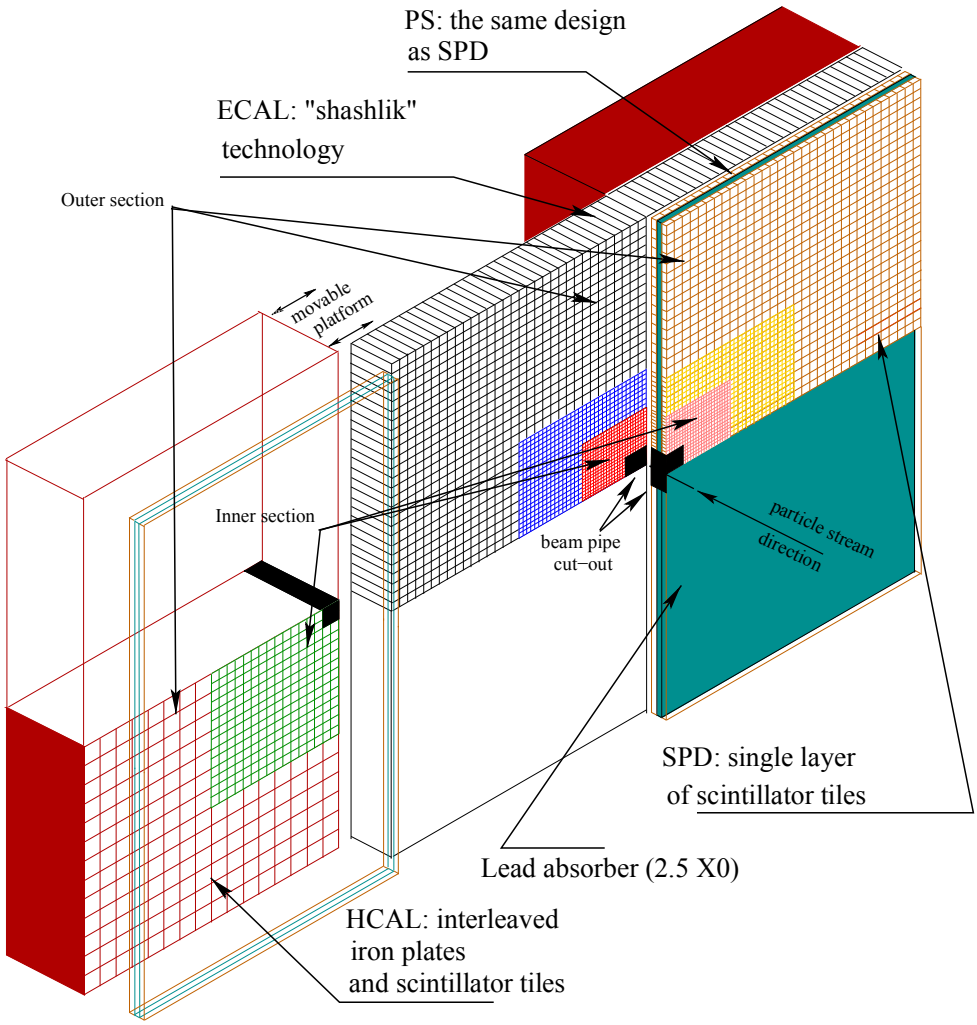



Fig. 2.12 Layout of the LHCb calorimeter system [49].

### Scintillator Pad Detector and Pre-shower

The SPD is situated after RICH2. The PS is separated from the SPD by a 15 mm lead converter. With no significant material to traverse to convert energy into showers, the SPD provides charged track multiplicity and discrimination between electrons and photons while suppressing the contribution from neutral pions. The converter depth corresponds to 2.5 (0.01) electron (pion) interaction lengths; a shower initiated in the converter and detected in the PS is likely an EM shower distinguishing pions from electrons and photons. This combined system provides $> 99\%$ rejection of pions for an electron efficiency of at least 90%, with increased performance at higher energies [41].

The detectors are composed of planes of scintillator cells, wound internally with WLS fibres, arranged with increased granularity towards the beamline (Figure 2.12 & 2.13). The SPD covers $2.1 < \eta < 4.4$ vertically and $1.9 < \eta < 4.4$ horizontally, with an active area of $\sim 33.4\,\text{m}$. By design, the PS is approximately 0.45% larger than the SPD with one-to-one projective correspondence with respect to trajectories from the interaction region. This relationship extends to the ECAL modules, positioned directly after the PS [41].

### Electromagnetic calorimeter

The ECAL is designed to provide transverse energy measurements, $E_T$, and discrimination between photons and electrons. The calorimeter is composed of lateral tiles: 2 mm lead and 4 mm scintillator separated by 120 μm thick reflective paper, in repeating layers. As shown in Figure 2.13, the granularity increases towards the highest particle density in regions adjacent to the beamline. The total depth of 42 cm corresponds to 25 electron radiation lengths and is therefore expected to contain EM showers and provide good energy and transverse energy resolution of electron and photon candidates. The energy resolution of the ECAL is $\sigma_E/E = 10\%/\sqrt{E/\text{GeV}} \oplus 1\%$ [41].

Events with a neutral cluster pointing back to a track indicate Bremsstrahlung emission; the ECAL provides information necessary to preserve the combined energy. The gain is optimised for $B$-physics, predominantly at low energies. As a result, cells experience saturation above 10 GeV and energy resolution is degraded at high energies due to losses of information through large Bremsstahlung depositions.

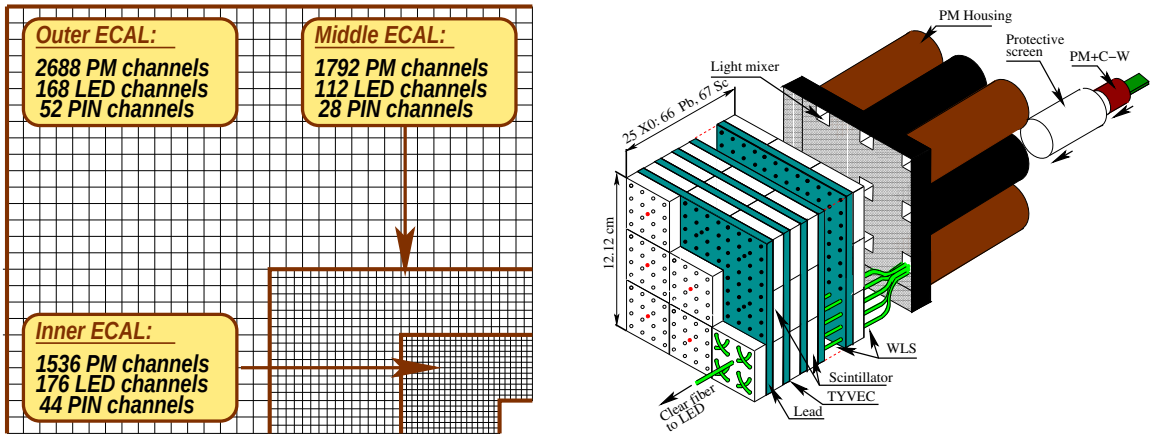



Fig. 2.13 ECAL layout per quadrant showing lateral dependant of channel density (left) and the structure of an inner module (right) from shashlik style sampling layers, permiated with the WLS fibres running parallel to the beamline through every layer, to the PMTs [50].

Downstream of the ECAL is the HCAL. Individual cells in the detector consist of five rows of square tiles, oriented perpendicular to the $x$-direction in 26 (13 half-tile) layers for the outer (inner) modules (Figure 2.14). The layers consist of three 0.3 cm polystyrene scintillator tiles alternating with 0.4 cm iron absorbers and are traversed by WLS fibres running along the edge of the layers. Each module extends downstream in what is referred to as a stack. The order of layers is inverted between rows such that, for paths parallel to the $z$-axis, three sets of scintillator and absorber tiles alternate along a stack [41].

The longitudinal orientation of the tiles provides 20.2 cm of iron along the trajectory of a particle, each corresponding to 1.0 interaction length, per row. The five stacks of a cell extends to a depth of 128.3 cm and the total HCAL material imposes 5.6 hadron interaction lengths [51]. The vertical and horizontal coverage of this detector reaches $1.8 < \eta < 4.2$ and $2.1 < \eta < 4.2$ respectively. As shown in Figure 2.14, the granularity is reduced compared to the previous calorimeters, with just 152 outer- and 215 inner-cells, accounting for the larger size of hadronic showers [36].

HCAL is used to differentiate between electrons and charged hadrons, providing energy deposition measurements for the latter. The inclusion of HCAL information significantly improves jet $p_T$ resolution in Run II as shown in Chapter 4. The detector provides not only an energy resolution for clusters in the HCAL of $\sigma_E/E = (69 \pm 5)/\sqrt{(E/\text{GeV})} \oplus (9 \pm 2)\%$ but also stopping power, limited by spatial constraints on cell depth, containing most hadronic showers prior to the muon chambers [41].

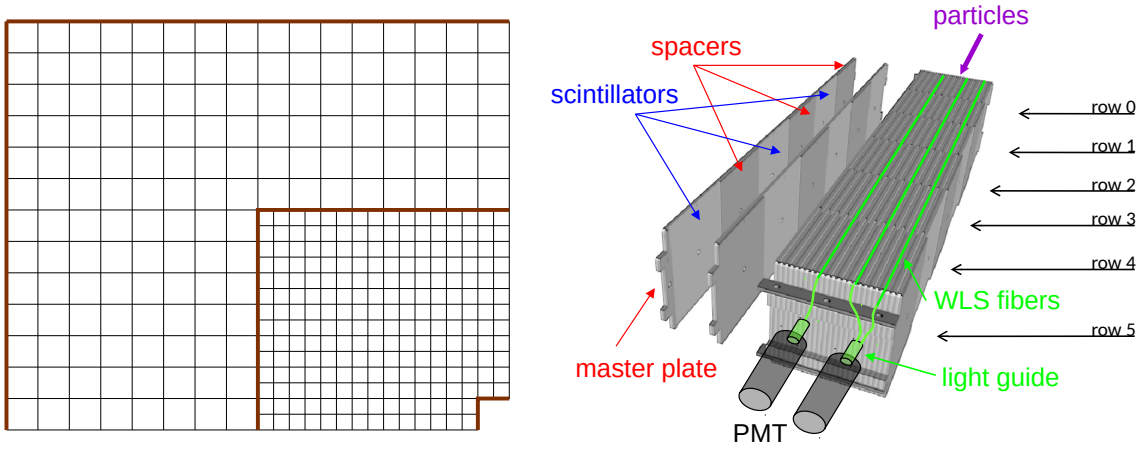



Fig. 2.14 HCAL layout per quadrant showing lateral dependant of channel density with square inner-cells of 13.13 cm and outer-cells of double the size (left) [52] and the structure of an inner module (right) with the WLS fibres running parallel to the beamline through every row between the tiles to the PMTs [49].

### 2.2.5 Muon stations

The first of five muon stations, M1, lies ahead of the calorimeters. The remaining stations are positioned downstream of the HCAL, alternating with 80 cm thick iron filters placed ahead of M3-5 (Figure 2.15). The detectors themselves use multi-wire proportional chambers (MWPCs), analogous to straw drift chambers, which collect an electron cascade from gas ionisation at an anode element. The innermost region of the M1 station instead employs a Triple-GEM (Gas-Electron Multiplier) to provide comparable tracking performance with a higher radiation tolerance [41]. M1-5 are spit into chambers of differing size and granularity to provide constant occupancy from beamline to outer acceptance (Figure 2.15).

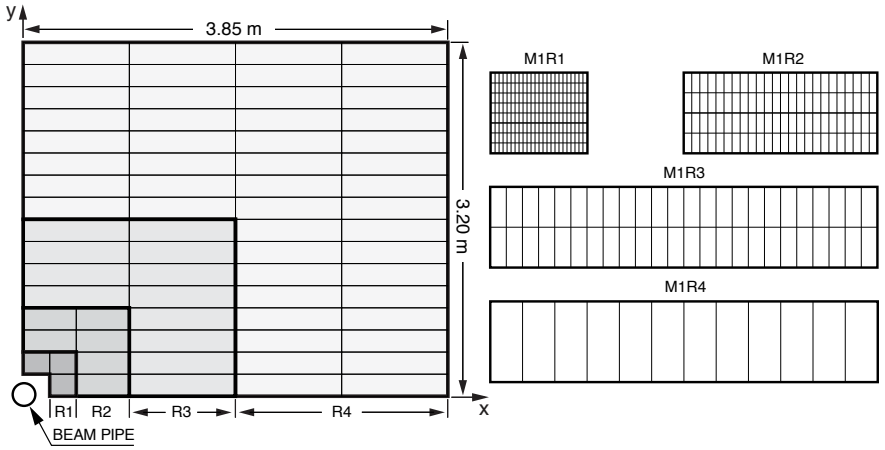



Fig. 2.15 A quadrant of the M1 station with rectangles representing a single chamber (left) division of the chambers of the four regions in the M1 detector planes and their respective pad distribution per chamber (right) increasing with $\eta$ [53].

The chambers through M1-5 are scaled to offer one-to-one $\eta$ mapping between stations, $2.0 < \eta < 4.8$ vertically and $1.9 < \eta < 4.6$ horizontally. They are optimised for the momentum resolution of muons with $p > 6\,\text{GeV}$, the minimum required to pass through all five stations [54]. The first three chambers provide high precision measurements of $p_T$ through their spatial resolution in the $(x,z)$-plane. Muon $p_T$ resolution using M1-5 is 20%, a 10% improvement afforded over just using information from M2-5 [53]. M4&5 are primarily used for particle ID and, given the inclusion of the filters, the M1, calorimeter systems and M2-5 combined material corresponds to 20 hadron interaction lengths [36].

### 2.2.6 Trigger system

The unfiltered data readout of LHCb greatly exceeds its bandwidth and storage capacities. The reduced $\mu$ results in an interaction rate comparable to that of the LHC bunch crossing,

40 MHz. To achieve a data acquisition rate of $\sim$ 12.5 kHz that can be analysed and stored, the data flow is reduced through real-time event selection provided by rapid event-by-event analysis. The three-tier trigger system (Figure 2.16) imposes requirements to isolate signal processes of predetermined interest with increasing complexity at each stage [48].

First applied is the hardware-based Level-0 (L0), using information from the calorimeter and muon systems. Events passing the L0 trigger are subject to the High Level Trigger (HLT), a two-part software-based stage implementing selections based on full event reconstruction performed online in HLT2, following the real-time alignment and calibration enabled by partial reconstruction HLT1 [48]. Trigger efficiencies must be determined for each final state.

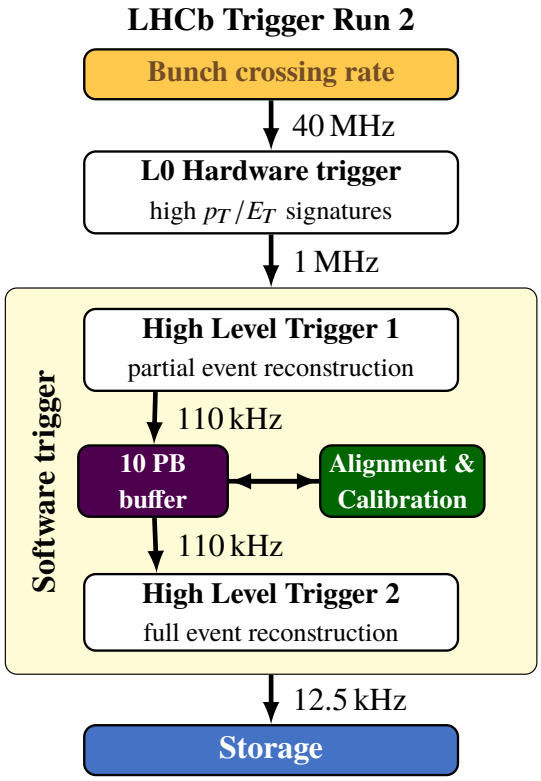



Fig. 2.16 Overview of the Run II trigger system with data output rates at each level displayed [55].

### Level-0

At L0, the trigger is passed if the event contains a high $p_T$ electron, photon, hadron or muon observed in the calorimeters or muon stations respectively. In addition, a pile-up system composed of the additional two VELO stations provides a veto on the measured number of PVs (primary vertices) in an event. With a decision latency of $< 4$ µs, the L0 stage can

reduce the data rate to 1 MHz, where events triggered by electrons or photons, hadrons and muons make up 150, 450 and 400 kHz respectively [55].

The highest $E_T$ $2 \times 2$ cell cluster in the ECAL with matched hits in the PS form the electron, those without are the photon candidates. The highest $E_T$ $2 \times 2$ cell cluster in the HCAL with a combined ECAL and HCAL energy above threshold determines the hadron candidates. The two highest $p_T$ tracks exceeding L0 threshold in each muon station quadrant, and traversing all five stations, provide the muon candidates. The candidate thresholds may change slightly by year through Run II [56].

### High level trigger

For Run II, the time required to process an event passing L0 is 50 ms. The HLT must first reduce the L0 output rate to 110 kHz (HLT1) to implement its run-by-run automated alignment and calibration ahead of a full event reconstruction stage (HLT2), outputting the 12.5 kHz capable of being sent to permanent storage [56]. The event filter farm (EFF), which benefited from a significant upgrade ahead of Run II, allows the output of HLT1 to be written to a local buffer, providing the software trigger with offline reconstruction quality information. HLT1 has a decision time of 35 ms per event with an average size of events passing HLT1 of 55 kB. A predetermined subset of events are used for the alignment and calibration via dedicated exclusive trigger lines [56].

HLT1 performs track reconstruction to produce long tracks which are fitted and evaluated to reject fake tracks. The fitted VELO tracks are then used to reconstruct PVs. Using long tracks with $p_T > 0.5$ GeV, HLT1 can carry out an inclusive selection of 1-2 track events, events containing muon candidates displaced from the PVs or events containing di-muon candidates [56]. Muon ID may be performed in HLT1 using fitted tracks with at least two hits in the muon stations and momentum greater than 3 GeV.

With full offline quality reconstruction, HLT2 can reconstruct 2-4 track vertices and select those with sufficient $p_T$ and a significant displacement from any primary interaction to signify $b$-hadron decay candidates. Additionally, both prompt and displaced muons or di-muon events may be selected using an HLT2 muon-ID procedure identical to offline reconstruction. The output rate is divided approximately 40% to inclusive topological trigger lines, another 40% to exclusive $c$-hadron triggers. The remainder is made up of di-muon, electroweak physics, exotic searches, and specific lines [56].

### Turbo-stream

Full raw event storage accounts for just half of the output from HLT2; such events require reprocessing to access physics objects for analysis purposes. As full offline quality reconstruction is provided in HLT2, a dedicated processing stream known as 'Turbo' that performs physics analysis online while discarding the raw event, was introduced [57]. This secondary stream accounts for another third of the output; the reduced event format of Turbo is accumulated in a dedicated data bank and allows an increased output rate, higher efficiencies and smaller selection biases for such events [58].

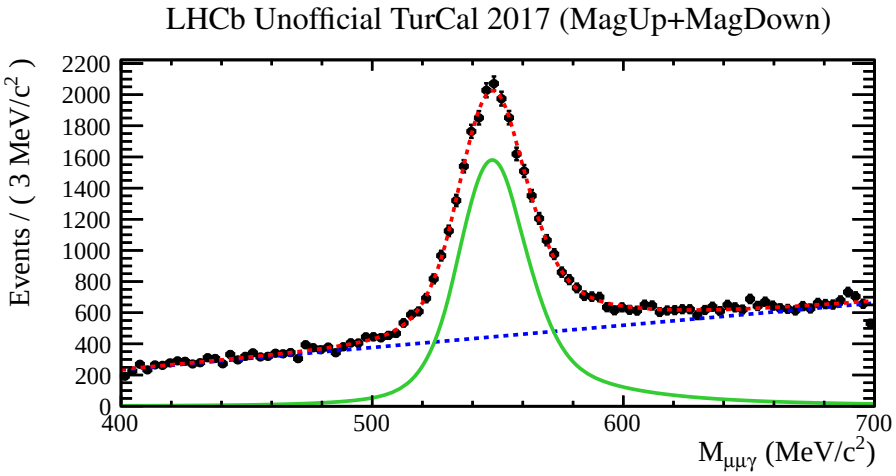
### Turbo-calibration

The remaining trigger rate is reserved for the Turbo-calibration (TurCal) where both the reduced and full formats are kept. TurCal provides large dedicated samples for detector calibration required for precision analyses in the Turbo-stream. Calibration channels have been included in TurCal since the beginning of 2017 data-taking. One example of a neutral calibration mode is $\eta \to \mu\mu\gamma$ (Figure 2.17a) which provides a calibration for soft photon reconstruction [57].

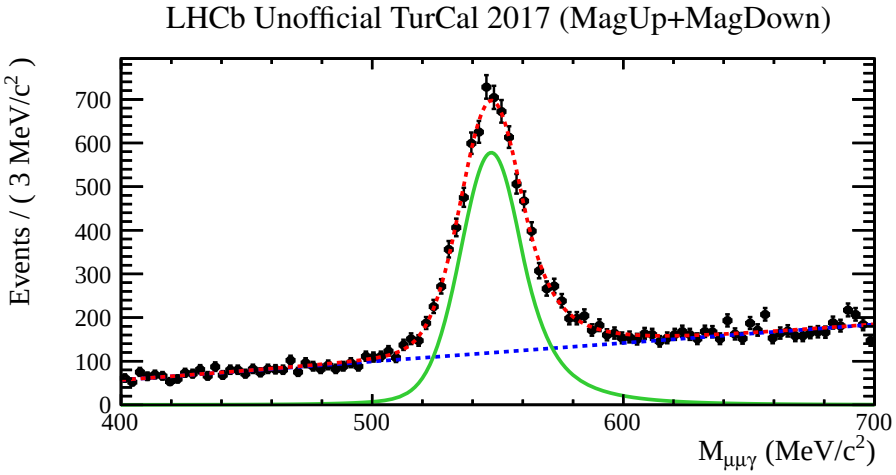
Photons undergoing pair production, or converted photons, before the magnet are reconstructed as a pair of electron tracks. Figure 2.17a shows crystal-ball function fitted to the reconstructed $M_\eta$ distribution in Turbo data implementing calibration coefficients (dictating the relative energy response of the calorimeter sub-detectors) for non-converted photons and re-optimised for di-photon mass resolution. Though converted photons typically offer a better resolution than calorimetric photons due to the tracking system performance, low energy photons such as those reconstructed to produce Figure 2.17b can offer at least comparable resolution.

### Software framework

HLT is written in the same framework as the software used in the offline reconstruction of events for physics analyses. Simulated $pp$ collisions generated with PYTHIA 8 use a specific LHCb configuration [59]. Decays of hadronic particles are described by EVTGEN [60], in which final-state radiation is generated using PHOTOS [61]. The interaction of the generated particles with, and the response of, the detector are implemented using the GEANT4 toolkit [62, 63] for the full LHCb setup [64].




(a) Double-sided crystal-ball fit to the Turbo calibration $\eta \to \mu\mu\gamma$ sample




(b) Double-sided crystal-ball fit to the Turbo calibration $\eta \to \mu\mu\gamma_{\not\ell}$ sample

Fig. 2.17 Signal plus combinatorial background fits to 2017 TurCal samples for the $\eta \to \mu\mu\gamma$ process including selecting for unconverted photons, $\gamma_{\not\ell}$.

Details of reconstruction algorithms and criteria are discussed in Chapter 3. The performance of jet configurations available in Turbostream are discussed in Chapter 4. The constraints on jet reconstruction included in HLT are mentioned in detail in Chapter 4.

# Chapter 3

# Event reconstruction

Raw outputs from the detector sub-systems, signals left by particles, are grouped into detector objects. Section 3.1 outlines the process of combining hits into tracks and tracks into interaction vertices. Section 3.2 details the reconstruction of jets, a detector level description of collimated radiation, and their association with secondary decay vertices. High-level objects like these are used to identify the products and event topology resulting from $pp$-collisions. Online reconstruction is performed in real-time, parallel with data taking [57], while offline reconstruction occurs once written to disk. Stringent budgeting of the available bandwidth and processing time limits online reconstruction to detector alignment and monitoring, Turbo-stream data taking, and producing the information necessary for trigger lines to discriminate relevant events to store with full detector information for offline processing.

## 3.1 Pattern recognition

The means to filter raw events ahead of each increasingly demanding processing stage is dependant on fast and efficient track reconstruction. Track-building employs pattern recognition techniques, associating low-level tracking detector information into trajectories that can be fit based on the magnetic field and scattering properties throughout the detector. The resultant tracks undergo selection to reduce those arising from double counting and detector coincidences. Identification of the particle associated with a track may be based on the presence of corresponding signals in other sub-detectors, such as the calorimeters or muon stations.

### 3.1.1 Tracks

The path of a charged particle through the tracking system, which includes the VELO, TT, IT and OT, will result in hits corresponding to a series of positions relative to the known detector geometry. Seeding algorithms reconstruct track segments: from VELO hits in $(r,\phi,z)$-coordinates, where straight lines are extrapolated towards the interaction point and down the beamline, producing VELO seeds; in the T-stations where the magnetic field is in effect, lines of hits forming lines in the $(x,z)$-plane are similarly selected, producing T-station seeds. Further algorithms responsible for the grouping of tracking information operate in the following sequence:

- *Forward tracking* - VELO seeds are combined with remaining single hits in the forward tracking stations to provide trajectories. Additional hits are then included if they are consistent with the reconstructed path.
- *Track matching* - Remaining VELO and T-station seeds are extrapolated through the magnet and combined if compatible, after which further hits in the TT are added if consistent with the track trajectory.
- *Up- & Downstream tracking* - Unpaired seeds, from the VELO & TT respectively, are extrapolated to the $(y,z)$-plane of the TT-stations where individual hits are each used to calculate resultant momenta. Only trajectories with at least three TT-hits of compatible momenta form a track.

The groups of hits associated with the candidates are input to the track fitting procedure, which provides the momentum of the particle, the sign of its charge from the curvature of its trajectory and the quality of the fitted track. Momentum resolution varies from 0.5-1.0% between 5-200 GeV with an average track reconstruction efficiency of 96%.

Trajectories are treated as discrete dynamical systems in steps of $z$, allowing the use of a Kalman filter, a discrete-data linear estimator, to fit the tracks. The filter provides recursive optimisation equivalent to a least-squares minimisation while accounting for multiple scattering in the detector and energy loss due to ionisation [56] . As shown in Figure 3.1, tracks form different classes depending on the sub-detector information associated with them. These include:

- *Long tracks* - information from the VELO and T-stations;
- *Downstream tracks* - information in the TT and T-stations only;
- *Upstream tracks* - information in the VELO and the TT only;

- *VELO tracks* - VELO information only;
- *T tracks* - information from the T-stations only.

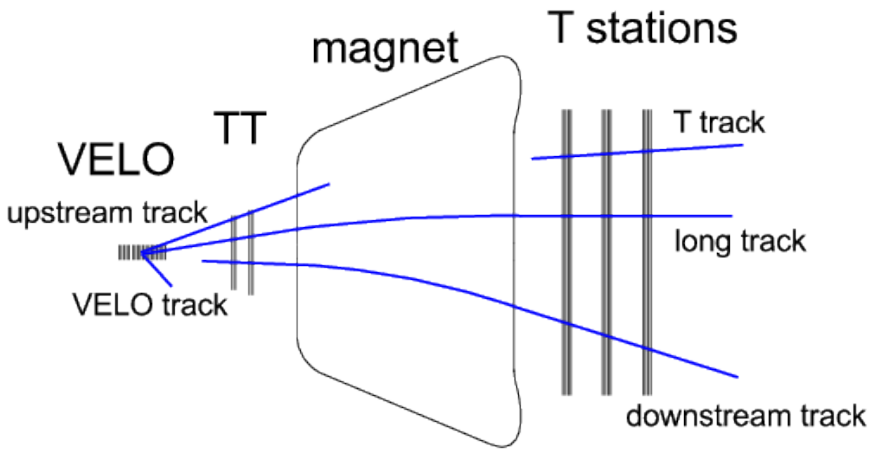



Fig. 3.1 Representation of reconstructed track types in the LHCb detector [65].

The algorithms aim to maximise the number of well-identified tracks that fall into these classes while minimising false tracks. The component arising from the mismatch of track coordinates in the $(r,\phi)$-layers, the combination of random hits from the underlying event and pile-up effects or detector noise, known as ghosts, are removed. Additionally, if more than one track shares the same hits, a clone-killing algorithm is subsequently run, prioritising tracks with the most hits and avoiding double-counting. The best-fitting long and downstream tracks undergo Kalman filtering across the full magnetic field range. VELO tracks, free of the magnetic field, enable using a simplified straight line Kalman filter for vertex reconstruction.

### 3.1.2 Vertices

Reconstructed track intersection points provide the vertices of particle interactions taking place in the VELO. A hard scatter will produce a primary vertex (PV) which, once reconstructed, may be associated to its constituent tracks by their distance of closest approach $< 1\,\text{mm}$. The groups of tracks satisfying this requirement may be extrapolated to a shared candidate PV called a seed. Only those tracks with a distance of closest approach of $< 30\,\text{mm}$ to the seed remain associated; this distance with respect to the vertex is known as the impact parameter (IP). Determination of the PV position utilises an adaptive weighted least-square minimisation of summed impact parameter significance ($\chi^2_{IP}$ values) from the tracks of the seed, allowing impact from displaced products of secondary interactions and ghosts to be

minimised. When $|\Delta z| < 0.5$ μm, where $\Delta z$ is the shift of the $z$-coordinate of the PV after each iteration, and at least five tracks have been assigned to the PV with non-zero weights, the minimisation terminates [66].

The selection in the trigger places requirements on lifetime parameters which, when computed online, use primary vertices constructed only using VELO tracks. While improvements can be made in the primary vertex resolution offline by including long tracks, the reduced systematic effects due to a consistent treatment of the selection variables provide a sufficient motivation to also restrict the offline primary vertex reconstruction to VELO tracks [55]. Like track reconstruction, the parameterisation of seed clustering has been tuned to maximise efficiency and minimise fake track reconstruction, providing an impact parameter resolution of $(15 + 29/p_T)$ μm.

### 3.1.3 Particle identification

The particle identification (PID) at LHCb relies on information from the RICH detectors, the electromagnetic and hadronic calorimeters and the muon system combined:

- *RICH I&II* - Tracks produced by hadrons will be associated with patterns in the photo-detectors of the RICHs which are used to test against various mass hypotheses allowing the distinction between kaons, pions and protons. When combined, the two RICH systems provide sensitivity to charged particle masses within the momentum range 2–100 GeV/c.

- *ECAL* - Energy deposits are matched against well-reconstructed tracks, and a position matching estimator is used to distinguish neutral from charged clusters. Where the ECAL fully absorbs particles in EM showers, neutral clusters not associated with tracks or SPS information are identified as photons. Conversely, charged clusters may be identified as electrons provided their energy matches between the shower and track. Multiple reconstructed photons may be traced back to a Bremsstrahlung process in the active region of the magnet and used in electron candidate cluster energy matching.

- *HCAL* - Energy reconstruction from hadronisation processes that pass through the ECAL reach the HCAL, which in turn acts as a shield to the muon chambers ensuring minimal hadronic contamination in the drift chambers.

- *M1:5* - If a track is associated with enough muon chamber hits and a consistent fitted momentum, then it is considered a muon, e.g. $p_\mu > 10\,GeV$ requires hits in M2, M3, M4 and M5. The result is a kaon rejection efficiency $> 95\%$, pion misidentification fraction of 10% and a muon efficiency $> 97\%$ increasing with momentum.

A set of likelihoods may be used in further selection criteria for specific PID based on collated information from all the sub-detector systems. Additionally, neural network approaches, tuned using Turbo-stream calibration samples, provide probability-like PID variables, `ProbNNX`.

## 3.2 Jet reconstruction

Jets are the label given to collimated radiation resulting from boosted particle interactions. Such events are prevalent at proton-proton collision experiments and observable with the LHCb detector. They provide objects with measurable properties tied to the kinematics of the individual seeding partons, obfuscated by the hadronisation process, providing a comparison to calculable predictions. LHC experiments generally use reconstruction algorithms implemented in the FastJet package [67]. The anti-$k_T$ algorithm, used to cluster hard (high $p_T$) particles preferentially, is proficient at resolving jets with a regular cone-like structure (Figure 3.1) of chosen radius while avoiding singularities [68]. When jet boundaries overlap, the hardest jets take precedent. LHCb implements anti-$k_T$ clustering with a default cone radius $R = 0.5$ in $(\eta,\phi)$-space, for preferable reconstruction performance in Run I and ease of comparison with CMS results, for its forward region studies.

### 3.2.1 Clustering

Essential requirements of the clustering algorithms to be shared between theorists and experimentalists include: consistent definitions; provision of finite results at all orders of pQCD, or infrared (IR) safety; predictions invariant to the soft emission or splitting of partons, or collinear (UV) safety. Where standard cone algorithms cannot treat overlap between jet boundaries, defining a fixed solid angle within which radiation is grouped, the iterative and split merge cone algorithms sacrifice UV and IR safety respectively to do so [68].

Sequential algorithms use preferentially ordered clustering, which preserves IR and UV safety. Three such orderings are described by Equation 3.1 by using different values for $c$ [69]. Iterating $i$ and $j$ over the jet inputs, if $d_{iB} > d_{ij}$ then $\vec{i}$ is replaced by $(\vec{i}+\vec{j})$ and $j$ is removed from the list. Once $d_{iB} < d_{ij}$ then the final $\vec{i}$ is defined as a jet and the process begins again until all inputs are clustered. The parameter $c$ defines the relative power of momentum against geometric separation in the preferential clustering [68].

$$\begin{aligned} d_{ij} &= \min\left(p_{T\,i}^{2c}, p_{T\,j}^{2c}\right) \Delta(\eta,\phi)_{ij}^{2}/R^2 \\ d_{iB} &= p_{T\,i}^{2c} \end{aligned} \tag{3.1}$$

$k_T$-clustering resolves sub-jets due to ascending $p_T$ ordering ($c = 1$); Cambridge / Aachen orders by distance only ($c = 0$) thus providing the best reconstruction for substructure; however each of these result in irregular jet size and shape which, in *hh* rather than *ee* collisions, degrades energy information due to pile-up and underlying event contributions proportional to $(\eta, \phi)$-area (Figure 3.1). For anti-$k_T$, inputs are ordered in descending $p_T$ ($c = -1$) providing regular jet size and shape as found in cones but with the IR and UV safety of other sequential clustering algorithms [69].

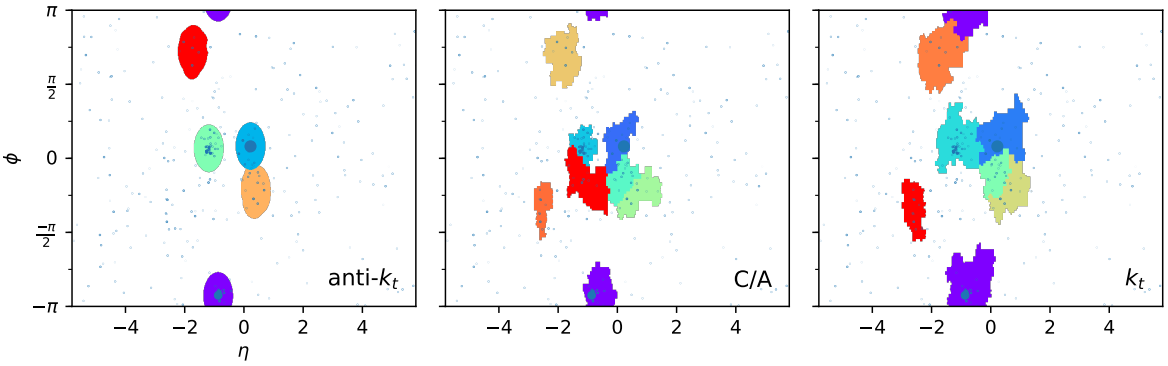



Fig. 3.2 Jet clustering with $R = 0.5$ using the same pyjet test data event [70], $c = -1$ (left), $c = 0$ (centre), $c = 1$ (right) demonstrating regular clone structure of inverse $p_T$ ordered iterative clustering.

### 3.2.2 Particle flow

The stage of reconstruction prior to jet clustering implements a particle flow algorithm (PF) [71, 72]. Its job is to sort and define jet inputs to prevent double-counting by enforcing a sequence to event reconstruction (Figure 3.3) as well as providing neutral energy recovery (NER). Both charged and neutral detector objects undergo matching between the tracks and calorimeter clusters. The charged energy depositions (track associated clusters) may be subtracted and the remaining neutral energy depositions may be recovered. The expected energy depositions of tracks are obtained with an $E/p$ calibration based on isolated tracks in Run II minimum bias data. This allows a parameterisation of the detector response, known as energy response functions (ERFs), to include cumulative degradation effects due to radiation.

Before matching, the detector objects must undergo selections specific to their presence in various sub-detectors and optimised to the energy content and resolution of jets produced. This selection reduces fake contributions to jet daughters and prevents fake jet reconstruction; fake jets are more costly to reject once fully reconstructed. This provides particle candidates (tracks & clusters), which make up the inputs to the PF, feeding the jet clustering algorithm.

This procedure is summarised in Figure 3.3:

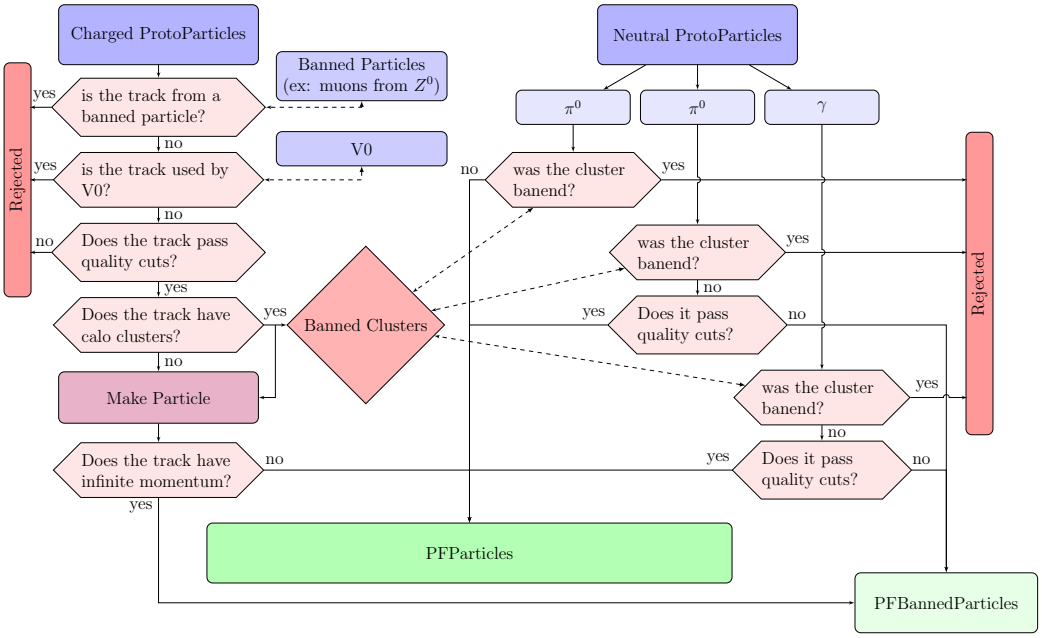



Fig. 3.3 Workflow of the `ParticleFlow` algorithm [73].

- Reconstructed tracks and ECAL clusters are paired with PID information into objects;
- Tracks containing a VELO segment are associated with the primary vertex to which they have the smallest impact parameter significance;
- Tracks and their closest associated calorimeter cluster make up the charged objects;
- ECAL calorimeter clusters that match with tracks are marked as charged clusters and banned from further usage as neutral components;
- Inputs are sorted, track associated clusters subtracted and the remaining energy is defined as neutral energy recovered;
- Calorimeter clusters with no associated tracks become the neutral objects;
- Classification of the neutral components treats the PID hypothesis of the clusters in the order of the highest constraint;
- HCAL clusters not banned during track matching are converted to particle candidates.

The standard jet configuration (Std) provides full reconstruction offline using the original PF algorithm implemented in Run I. The PF was redesigned for trigger rate reconstruction for online jets as well as a more flexible input selection. Running full reconstruction in

the software trigger necessitates careful CPU time budgeting. Fake track rejection before PID and decay reconstruction in the software trigger is essential in this regard [74]. Using the Run II PF, the Turbo-stream jet configuration (Turbo) provides online jet reconstruction and the high level trigger jet configuration (HLT) provides offline reconstruction. Each of the new configurations lacks in-built selection criteria; optimisation studies based on Run II conditions were therefore carried out (Chapter 4).

The jet configuration in TURBO runs without the information required for NER, sacrificing 10% in energy resolution (Chapter 4). With this information included, its HLT based jets may be shown to perform similarly to the offline configuration (Std) given the right quality of input selection. The problem of sufficiently fast fake rejection for trigger level reconstruction instigated the development of a fast neural-net (NN) for fake track ID [74]. A study of jet performance in Run II is detailed in Chapter 4.

### 3.2.3 Jet identification

Cuts on global jet variables reject badly reconstructed, lepton-seeded (clustered around a dominant high-$p_T$ track from non-QCD processes) and pile-up or calorimeter noise-based jets. Reconstruction level $(Z \to \mu\mu)$+jet and di-jet MC samples, produced using PYTHIA, EvtGen and GEANT4, are each used to define background as $\Delta R > 0.5$ between the reconstructed jet and any true jets. Events considered for the $(Z \to \mu\mu)$+jet sample require $\Delta R(\mu, j) > 1$ in order to isolate the jet contents from $Z$ decay muon pairs while for the di-jet sample, $\Delta R(j_a, j_b) > 1$ to remove soft radiation processes; both require a true PV reconstructed within $\Delta x < 0.3\,\text{mm}$. A linear regression [75] is used to optimise the selection of signal over background based on simulation samples described providing the jet identification (JetID) criteria.

### 3.2.4 Jet energy corrections

A correction factor for the systematic offset of the reconstructed energy of a jet may be derived from two functions: the MC correction factor, $C_{MC}$, and the residual correction factor, $C_{RES}$. The aim is to calibrate the jet energy to the truth level jet energy with $C_{MC}$, taking into account noise, pile-up and the non-uniformity of the detector. Then $C_{RES}$ accounts for the differences between jet $p_T$ in data and MC. The corrected momentum of jets in data may be defined $p'_T = C_{MC} C_{RES} p_T$.

### 3.2.5 Secondary vertex tagging

Jets produced via heavy flavour production are tagged using an algorithm through the presence of a secondary vertex. The SV-tagger first selects displaced tracks, those with large $\chi^2_{\text{IP}}$, to combine them into two- and three-body vertices. Those vertices which share tracks are linked to form $n$-body SVs. A quality selection is then applied to suppress backgrounds from strange decays and material interactions. A jet is successfully SV-tagged if it contains a vertex passing the selection and falling within $\Delta R < 0.5$ of the jet axis [76]. Events including an SV-tag may employ fits to dedicated heavy flavour BDT responses to extract flavour content as explained further in Chapter 5. The jet tagging undergoes a data-driven calibration by comparing fits of the corrected mass and SV-track multiplicity between data and MC.

# Chapter 4

# Run II jet reconstruction

Having revised its particle flow algorithm to enable speeds required for front end analysis, LHCb was also provided with more flexible quality control over jet inputs. Section 4.1 outlines the constraints for jet reconstruction in the high level trigger, motivating the subsequent studies of its performance under Run II conditions. Section 4.2 presents the optimisation of an input filter for the particle flow algorithm, including fast neural network based ghost rejection. Performance studies of new and existing configurations, from clustering, JetID cuts and energy corrections, finalise the coherent approach between online and offline jet reconstruction; Section 4.4 establishes these as the default configurations for Run II legacy.

## 4.1 Jets in the Run II trigger

As explored in Chapter 2, the LHCb trigger system consists of three stages: the hardware-based L0 followed by software HLT 1 and HLT 2. Improvements made to the event filter farm and HLT data flow for Run II provided low $p_T$ tracking without IP cuts in HLT 1, as well as full event reconstruction in HLT 2 [77]. To facilitate real-time reconstruction, LHCb became the first high energy physics experiment to implement a fully automatic tracking system alignment, PID calibration and offline-equivalent reconstruction in the trigger. Hence, it was possible to perform physics analyses directly with the information calculated by the HLT event reconstruction in what has been dubbed the 'Turbo-stream', or TURBO [77]. Once full event reconstruction was moved to the HLT, both the particle flow algorithm and processes behind its input selection were revised to run at the necessary speeds to support online analysis. The revision also provided the opportunity for new performance studies of jet reconstruction with new configurations while the standard configuration (Std), the default since Run I, serves as a benchmark. Both the jet definitions and suitability of established input selections under Run II conditions were explored in MC.

The challenge of sufficiently fast fake track rejection for trigger level reconstruction instigated the development of a neural network (NN) using computationally light activation functions to increase speed and reduce processing load for fake track ID [74]. The resulting 'ghost probability' algorithm, `GhostProb`, was able to supersede existing multivariate techniques in the track quality selection for the new HLT particle flow algorithm which were not applicable in the HLT. The TURBO jet configuration (Turbo jets) became the online default in Run II by saving processing time, running without the HCAL information required for track and cluster matching or neutral energy recovery (NER, Section 3.2.2). Without this information, TURBO can be shown to sacrifice 5-10% in energy resolution (Section 4.3.2). Retaining this information while using the new particle flow in an offline configuration (HLT jets) is shown to perform similarly to the Run I configuration (Std jets) given sufficient input selection.

## 4.2 Ghost tracks and jet input selection

In silicon strip detectors, such as the VELO, a coincidence of real hits with each other or noise in a single layer can result in 'ghost tracks'. These can be made up of hits in the module registering an event with the mismatched coordinate (e.g. $R$ and $\phi$) combinations of the true recorded hits from the same layer and time. A proportion of charged particle candidates contain noise and coincidences included in their tracking information, producing poor fits or fake tracks. Cutting on the response of the `GhostProb` NN, developed to identify ghost tracks for HLT in Run II, provides ghost rejection for charged particle candidates. Not only can a `GhostProb` requirement be placed on tracks of all types, but a unique track ID (UTID) requirement can be applied; of tracks found to share a track segment in the VELO, only one is selected. In Std, UTID was based on the smallest track $\chi^2$ value whereas, in the new HLT and Turbo configurations, the track with the lowest `GhostProb` passes.

In addition to a uniform cut to all track types based upon the predecessor to `GhostProb`, the Std configuration imposed a maximum $\chi^2/NDF$ and minimum $p_T$ upon tracks. These requirements were tighter for upstream and downstream track types (Figure 3.1). Each of the chosen requirements from Std replicated in the particle candidate filter of the new particle flow are summarised in Table 4.1, applied to tracks by type and on neutral HCAL clusters. Track-type specific selections had not previously been implemented in either HLT and Turbo configurations and are shown to reduce the rate of fake jet reconstruction in Section 4.3.1. Studies using a simulated 13 TeV samples of upstream, long and downstream tracks over the range $2.0 < \eta < 4.5$ showed that, while the $\chi^2$/NDF and `GhostProb` cuts each had direct impact on ghost content, the track $p_T$ and $\Delta p/p$ requirements were not over a range suitable

for ghost rejection. These cuts were assumed to define the limits for reliable reconstruction, with impacts manifesting at the jet reconstruction and JetID stages.

Studies using simulated 13 TeV samples of jets demonstrated that the majority of the remaining $\eta$ dependence of the fake jet rate was addressed with the HCAL $E_T$ threshold. Such a requirement had not previously been implemented in either HLT and Turbo configurations. Tightening the `GhostProb` requirement further was shown to offer no improvements. While ghost rates could be reduced by cutting further on `GhostProb` alone, a uniform application to all track types canf be shown to provide no noticeable improvement when combined with the various track-type specific requirements from Table 4.1 As shown in Sections 4.3.1 & 4.3.2, an input filter based on these criteria significantly improves the performance of the HLT based configurations.

Table 4.1 Requirements on track types and calorimeter clusters in Std jet reconstruction, replicated in Run II configurations.

| Type | Upstream | Downstream | Long | HCAL |
|---|---|---|---|---|
| Track $\chi^2$/NDF | $< 1.5$ | $< 1.5$ | - | - |
| Track $p_T$ (GeV) | $> 0.1$ | - | - | - |
| $\Delta p/p$ | $< 0.5$ | $< 0.1$ | $< 0.1$ | - |
| $E_T$ (GeV) | - | - | - | $> 2.5$ |

## 4.3 Jet reconstruction with the particle flow filter

The following study was performed using a 13 TeV Pythia $t\bar{t}$ sample of W$\rightarrow \mu$ triggered events with an associated $R = 0.5$ anti-$k_T$ jet. The selection requires the vector sum $p(\mu_j +$ jet$)_T > 20$ GeV (a proxy for $\not{E}_T$ where $\mu_j$ is a anti-$k_T$ clustered object with $R = 0.5$ seeded by the muon), a minimum muon $p_T$ of 20 GeV for high reconstruction efficiency and the jet and muon have a $\Delta R > 0.5$ such that the muon is isolated at least from the primary jet in the event. This allows for clean, efficient discrimination of EW jets from QCD backgrounds as exploited in Chapter 6. This allows tests of pile-up jet rejection in the post-reconstruction selection (Section 3.2.3) and provides a broad $p_T$ spectrum over which to assess performance.

### 4.3.1 Fake jet rate and reconstruction efficiency

The inputs passing the particle flow are clustered into jets whereupon a reconstructed $p_T$ threshold of 12.5 GeV is applied for Std, HLT and Turbo configurations. Fake jets are defined as those reconstructed without a truth level jet associated (within a $\Delta R = 0.5$ with $p_T > 12.5\,\text{GeV}$). The definitions in Equation 4.1 form the primary metrics for assessing the impact of input selection criteria.

$$\text{Fakerate} = 1 - (\text{jets}^{Rec}_{MC} / \text{jets}^{Rec}), \qquad \text{Efficiency} \quad = \text{jets}^{Rec}_{MC} / \text{jets}_{MC} \tag{4.1}$$

The above terms correspond to count of: $\text{jets}_{MC}$, true jets; $\text{jets}^{Rec}_{MC}$, reconstructed true jets; $\text{jets}^{Rec}$, reconstructed jets. It can be shown that the reconstruction efficiency outside the range $2.2 < \eta < 4.2$ drops sharply; for Std jets the losses are 5% for $2.0 < \eta < 2.2$ and $> 10\%$ at $4.2 < \eta < 4.5$. For the remainder of the study, the jet acceptance will be limited within $2.2 < \eta < 4.2$ to avoid regions of irregular detector geometry.

Besides requirements placed on tracks and calorimeter clusters at their trigger level reconstruction, such as `GhostProb` $< 0.4$, the particle flow algorithm in HLT2 initially provided the HLT and Turbo configurations with unfiltered jet inputs. Jet configurations are compared in Figure 4.1 as functions of true jet $\eta$ and $p_T$, where only the Std jet particle flow has an inherent input selection, represented in part by the criteria in Table 4.1. As a result, the HLT based jets have $\sim$2% improved efficiency while suffering $\sim$2% higher fake rate, exacerbated for $3.0 < \eta < 4.5$, rising to nearly 6% over that of Std.

Applying quality requirements to the tracks and calorimeter clusters used in the particle flow algorithm is shown to substantially reduce fake jet reconstruction for both HLT configuration and TURBO (Figures 4.1 & 4.2). The fake jet rate at $p_T < 30\,\text{GeV}$ is approximately halved while above 30 GeV the fake rate becomes negligible. HLT and Turbo experience $\sim 2\%$ and $\sim 1\%$ loss in efficiency respectively in jets with $p_T < 30\,\text{GeV}$. The heavy $\eta$ dependence of the fake rate in the forward most bins for HLT and Turbo jets is alleviated. Increasing from 1% to 2% for Turbo, a reconstruction inefficiency of up to 1% is introduced in HLT for 2.5-4.5 jet $\eta$. For HLT and Turbo jets, the fake rate and inefficiency are negligible for jets of $p_T > 50\,\text{GeV}$ while remaining sub per cent across the $\eta$ acceptance.

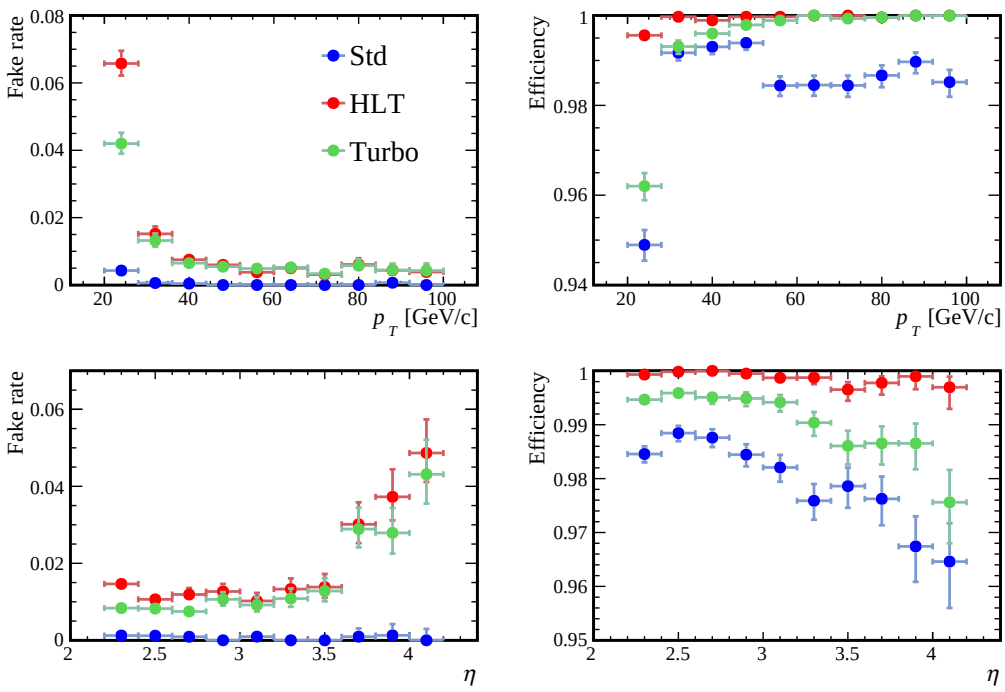



Fig. 4.1 $p_T$ (top) and $\eta$ (bottom) dependence of jet fake rate (left) and efficiency (right) for Std (red), HLT (blue) and TURBO (green), without new input filter requirements for the HLT particle flow.

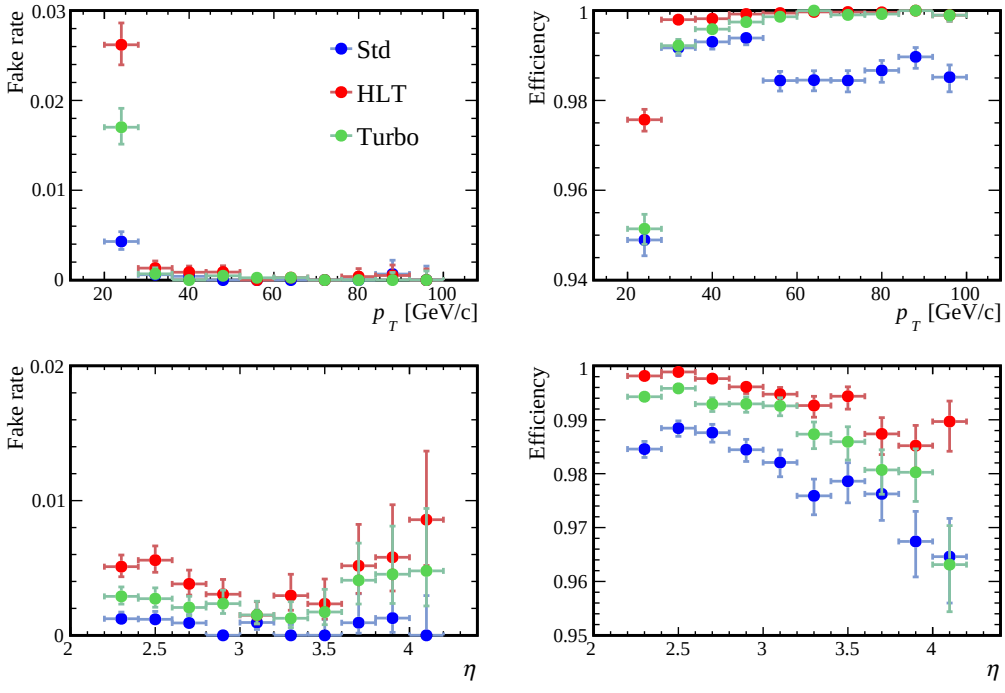
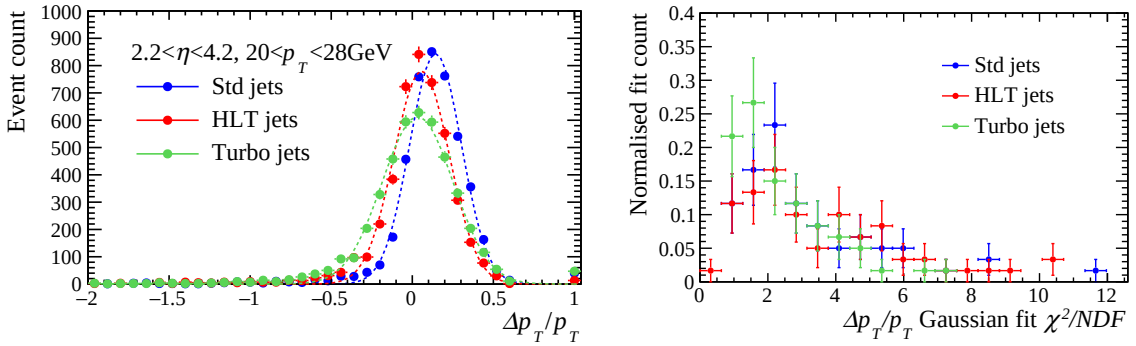



Fig. 4.2 $p_T$ (top) and $\eta$ (bottom) dependence of jet fake rate (left) and efficiency (right) for Std (blue), HLT (red) and TURBO (green), with the HLT particle flow input filter applied.

### 4.3.2 Transverse momentum and directional resolutions

The $p_T$ resolution is defined by the width of the distribution of fractional residuals between true and reconstructed jet $p_T$. The spatial resolutions, in ($\eta$,$\phi$)-coordinates and $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}$ between the true and reconstructed jets, are also accessible in MC. The fits used to compare the estimated resolutions of each configuration are illustrated in Figures 4.3 & 4.4. They show examples of the fits used to extract central or peaked values, $\mu$ (the Gaussian mean, Lorentzian mode and Landau location), and widths, $\sigma$ (the Gaussian standard deviation, Lorentzian full width at half maximum and the Landau scale), for true minus reconstructed: transverse momentum as a fraction of the true value, jet pseudorapidity, azimuthal angle; and their separation in $R$.

Gaussian fits to $\Delta p_T / p_T$ proved favourable to Lorentzian and Voigtian[1] alternatives based on the $\chi^2$ values shown in Figure 4.3. Both single and double crystal-ball functions, using a Gaussian core, resulted in fit instability. When defining the position resolution in $(\eta, \phi)$-space, the preference for a Voigtian (Figure 4.4 a & c) is likely due to the $\Delta\phi$ and $\Delta\eta$ profiles having heavier tails for which the $\chi^2$ values in Figure 4.4 b & d imply the Voigitan provides a good description. These features could be due to the transformation between the spatial resolution of the detector in an Euclidean lab frame to the jet daughters ($\eta$,$\phi$)-coordinates or an artefact of anti-$k_T$ clustering in ($\eta$,$\phi$)-space. Gaussian and Lorentzian distributions were found to produce worse $\chi^2$ fits to $\Delta\phi$ and $\Delta\eta$. The combination of $\Delta\phi$ and $\Delta\eta$ into the distribution for $\Delta R$ is used to provide an absolute offset and resolution estimated using a Landau fit. By comparison, a crystal-ball function, gamma function and log-normal distribution provided disfavourable $\chi^2$ values.




(a) (b)

Fig. 4.3 Fits to the difference in true jet and reconstructed $p_T$, where (a) shows an example of Gaussian fits and (b) shows the corresponding $\chi^2$ of fits (across $p_T$ and $\eta$ bins) for each configuration.

[1]The convolution of Gaussian and Lorentzian distributions approximated by a linear combination of the two functions [78] where they share $\mu$ and $\sigma$ parameters in the fit.

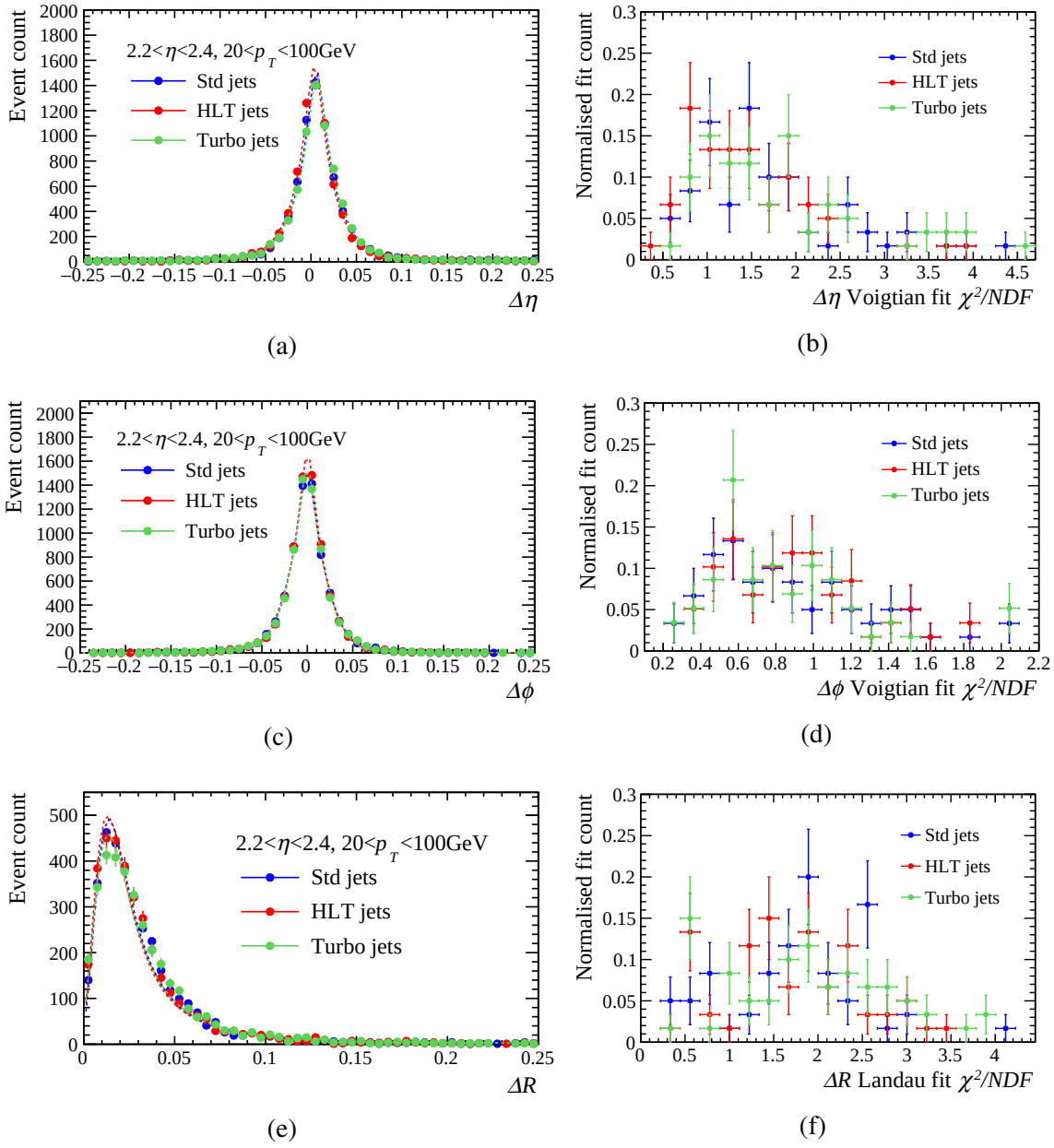



Fig. 4.4 Fits to the difference in reconstructed spatial variables to those of the associated true jet, where (a) & (c) display examples of Voigtian fits to $\eta$ & $\phi$ respectively, (e) displays an example of a Landau fit to $\Delta R$, and the corresponding $\chi^2$ of fits (across $p_T$ and $\eta$ bins) for each configuration are in (b), (d) & (f).

Figures 4.5-4.8 correspond to the jets compared in Figure 4.2, showing the $p_T$ and spatial resolutions of Std, HLT and Turbo jets as functions of true jet $p_T$ and $\eta$. Figure 4.5 shows that the HLT $p_T$ resolution increases by up to 1% for $p_T < 50\,\text{GeV}$ and 2% for $\eta > 3.5$ compared to Std. The differences between HLT and Std input selection are also demonstrated

by the reduced positive $\Delta p_T/p_T$ offset. Though jet energy corrections later address most of this reconstruction bias for each configuration, HLT and Turbo produce a smaller such systematic offset.

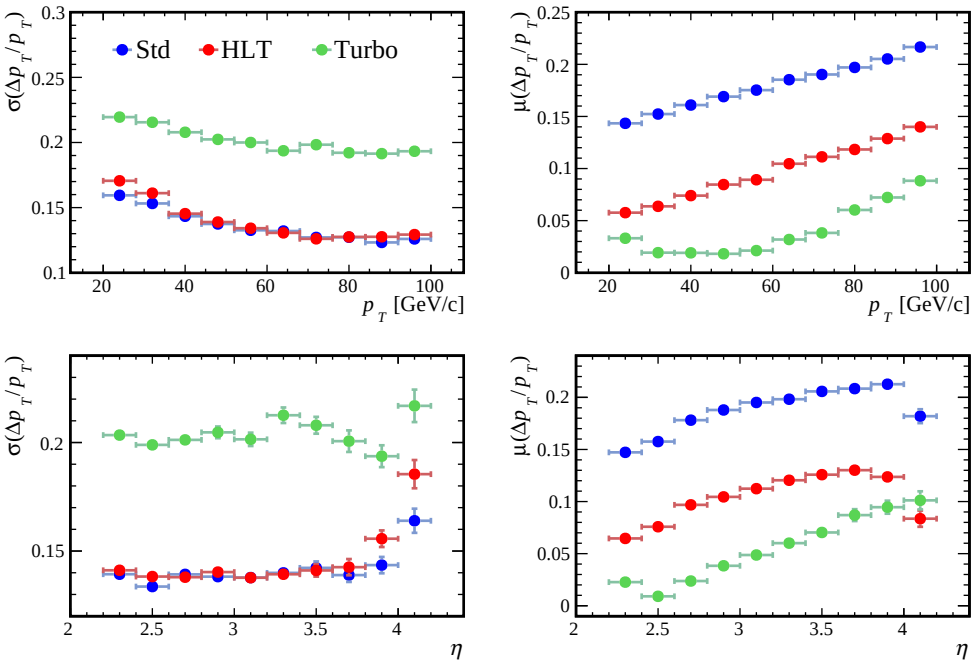



Fig. 4.5 $p_T$ (top) and $\eta$ (bottom) dependence of MC jet to reconstructed jet $\Delta p_T/p_T$ resolution (left) and offset (right) estimated from fitted Gaussian width & mean corresponding to Figure 4.2.

The $\eta$, $\phi$ and $R$ resolutions of HLT jets (Figures 4.6, 4.7 and 4.8) each improve on Std jet resolutions by roughly 5, 10, 20% respectively across the jet $p_T$ range, while Turbo offers 20% reduction for jets in bins $3.8 < \eta < 4.0$. While no significant bias is shown in $\phi$ reconstruction, the mean $\Delta\eta$ is consistently positive and up to 15% larger for Turbo than Std at low $p_T$ while 20% smaller for HLT than Std. For jets with $p_T > 50\,\text{GeV}$, the $\Delta R$ between the true and reconstructed jets is reduced by up to 15% moving to HLT or Turbo jets from Std, with smaller improvements offered by HLT for lower $p_T$ jets too. HLT and Turbo demonstrate reduced truth to reconstructed $\Delta R$, improving substantially over Std as shown in Figure 4.8. The offsets are reduced by approximately 10, 30% for bins $3.8 < \eta < 4.0$. All configurations are shown to have a $\Delta R$ resolution of $< 4\%$ of maximum jet width ($R = 0.5$) and offset of $< 8\%$ maximum jet width across all $\eta$ and $p_T$ bins. These reconstruction level differences are dominated by the $p_T$ resolution and, as discussed in Chapter 6 Section 6.7.2, they are absorbed into the systematic uncertainty on the acceptance factor between truth and reconstructed jets, on the $\mathcal{O}(1\%)$.

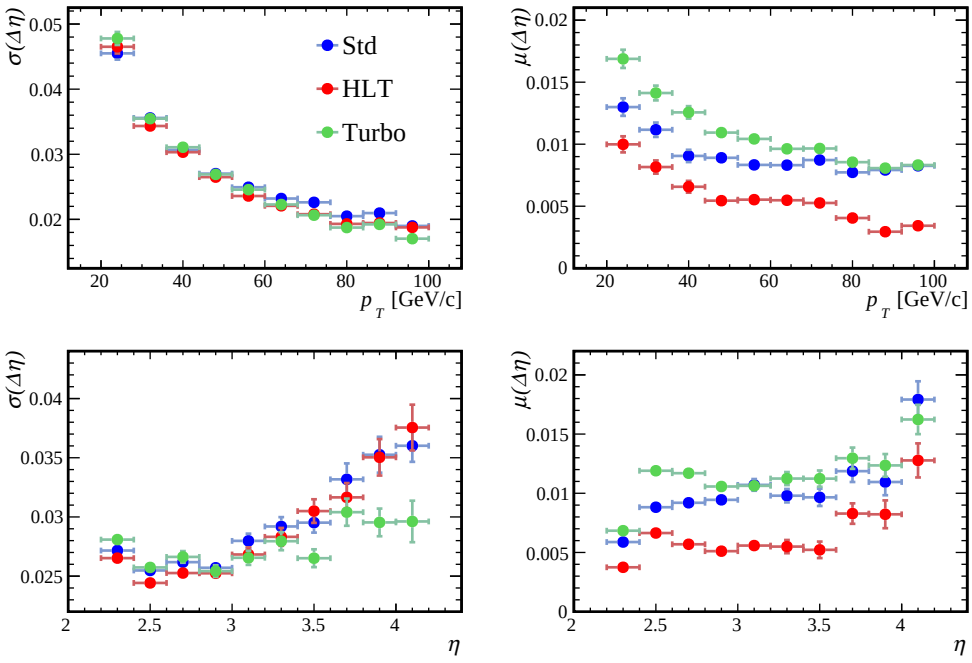



Fig. 4.6 $p_T$ (top) and $\eta$ (bottom) dependence of MC jet to reconstructed jet $\eta$ resolution and offset (estimate from fitted Lorentzian width & mode) corresponding to Figure 4.2.

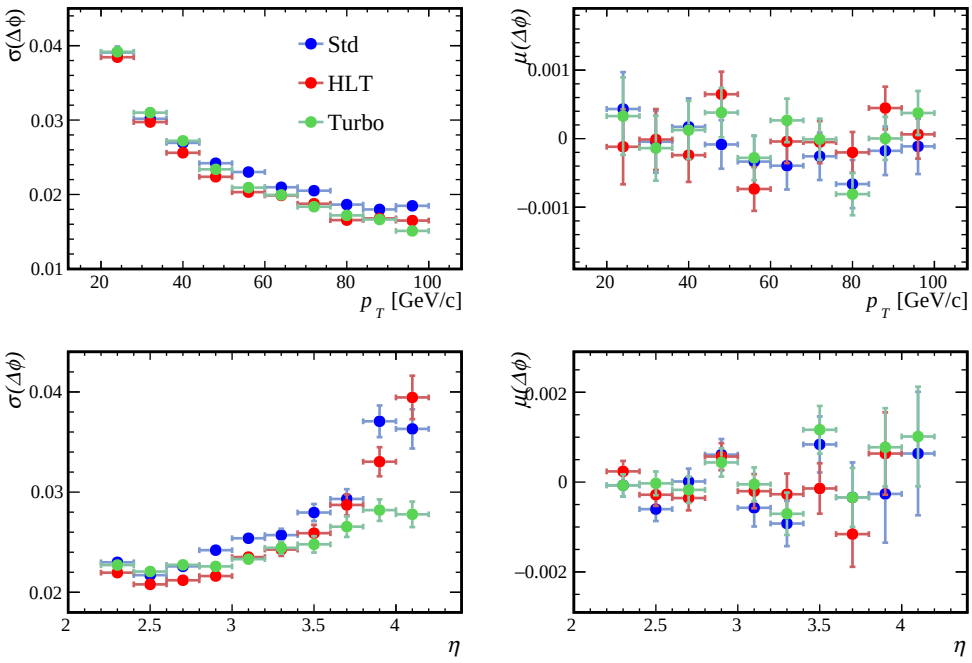



Fig. 4.7 $p_T$ (top) and $\eta$ (bottom) dependence of MC jet to reconstructed jet $\phi$ resolution (left) and offset (right) estimated from fitted Lorentzian width & mode corresponding to Figure 4.2.

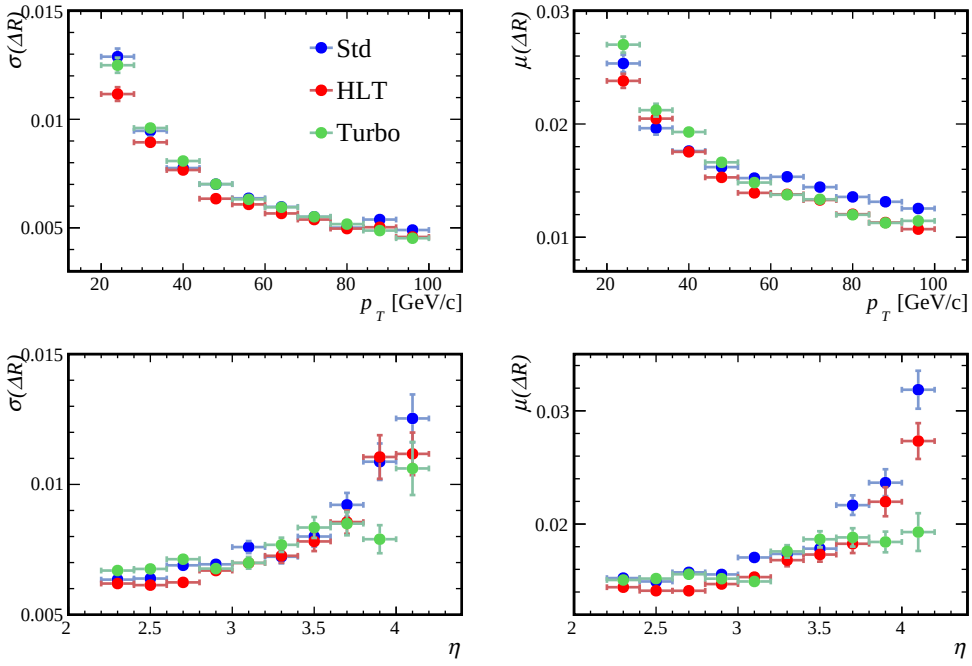



Fig. 4.8 $p_T$ (top) and $\eta$ (bottom) dependence of MC jet to reconstructed jet $\Delta R$ resolution (left) and offset (right) estimated from fitted Landau scale & location parameters corresponding to Figure 4.2.

### 4.3.3 Anti-$k_T$ radius

Given the maximum clustering radius, $R = 0.5$, had been used throughout Run I and into Run II, Std jets are tested under Run II conditions along with the newly filtered HLT and Turbo configurations. Figures 4.9-4.14 correspond to the jets Section 4.3.1 (Figure 4.2) split by configuration and compared to jets of varied radius. Figures 4.9, 4.11 & 4.13 demonstrate a loss in efficiency in each configuration when moving to $R = 0.4$ and a significant increase in fake jet rate with $R = 0.7$. Jets with both $R = 0.4, 0.7$ are shown by Figures 4.10, 4.12 & 4.14 to degrade $p_T$ resolution compared to $R = 0.5$. Given the range of performance observed across jet radii, the optimisation of jet selection criteria post-reconstruction (Section 3.2.3) is carried out independently for each anti-$k_T$ radius. The need for radius specific JECs (Section 4.4.2) are also demonstrated by the $\mu(\Delta p_T / p_T)$ variation within Figures 4.10, 4.12 & 4.14.

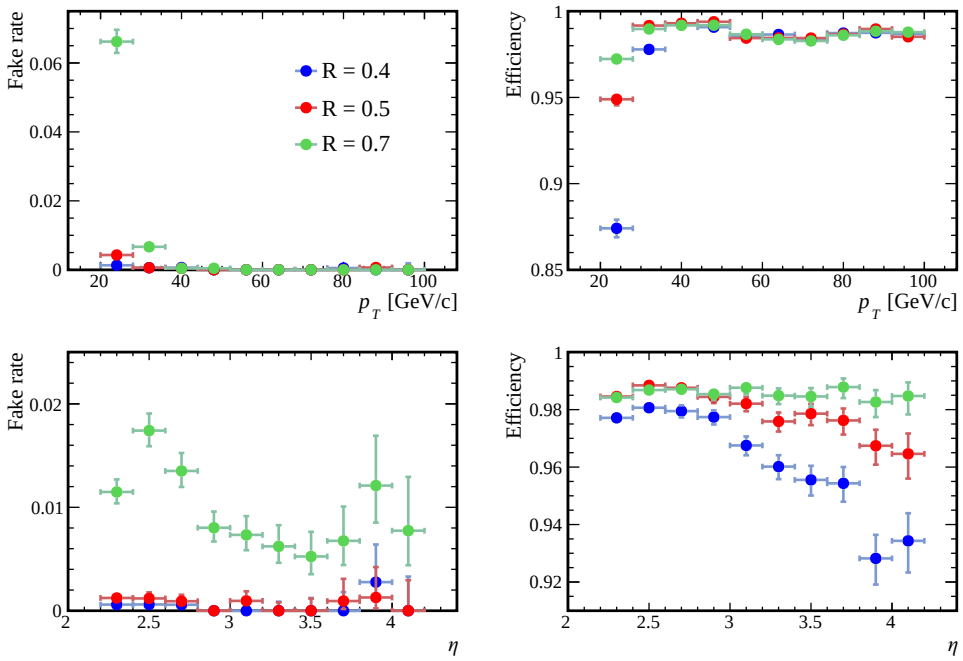



Fig. 4.9 $p_T$ (top) and $\eta$ (bottom) dependence of Run II jet reconstruction fake rate (left) and efficiency (right) for Std jets of anti-$k_T$ radii: $R = 0.4, 0.5, 0.7$ (blue, red, green).

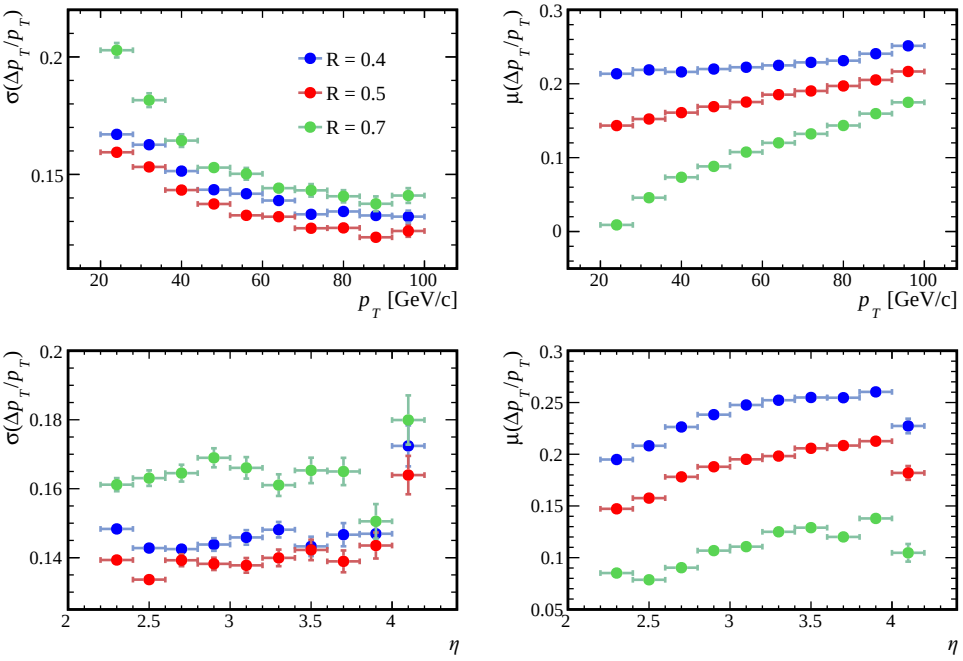



Fig. 4.10 $p_T$ (top) and $\eta$ (bottom) dependence of MC jet to reconstructed jet $\Delta p_T/p_T$ resolution (left) and offset (right) estimated from fitted Gaussian width & mean for Std jets of anti-$k_T$ radii: $R = 0.4, 0.5, 0.7$ (blue, red, green).

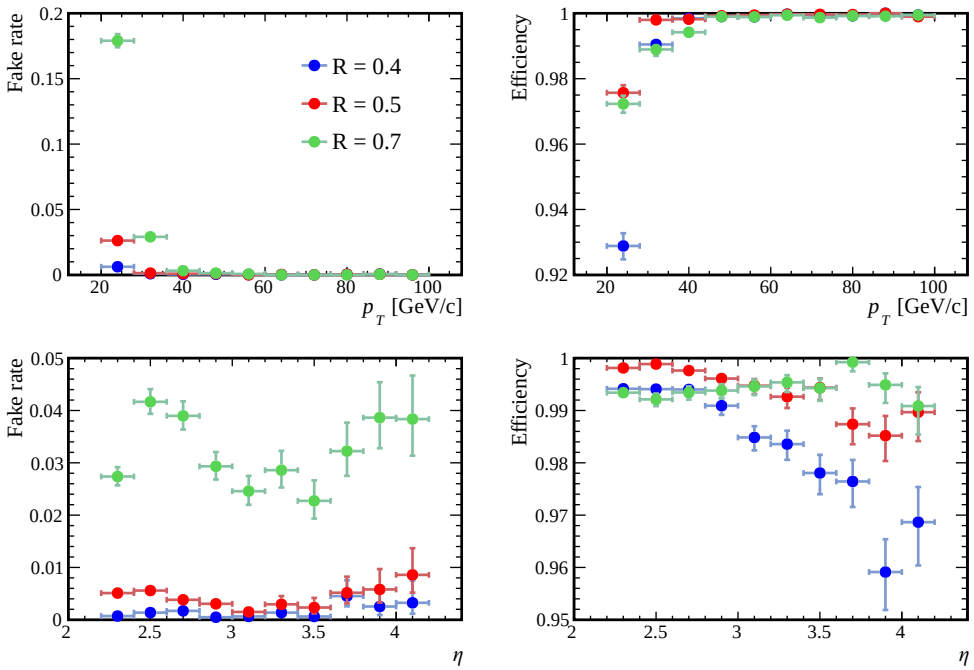



Fig. 4.11 $p_T$ (top) and $\eta$ (bottom) dependence of Run II jet reconstruction fake rate (left) and efficiency (right) for HLT jets of anti-$k_T$ radii: $R = 0.4, 0.5, 0.7$ (blue, red, green).

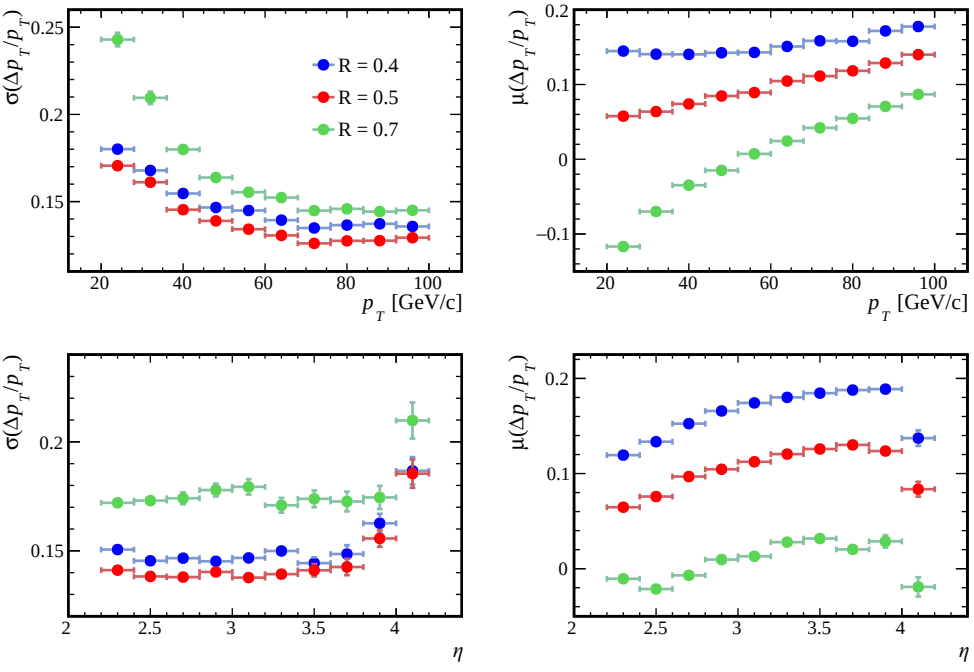



Fig. 4.12 $p_T$ (top) and $\eta$ (bottom) dependence of MC jet to reconstructed jet $\Delta p_T/p_T$ resolution (left) and offset (right) estimated from fitted Gaussian width & mean for HLT jets of anti-$k_T$ radii: $R = 0.4, 0.5, 0.7$ (blue, red, green).

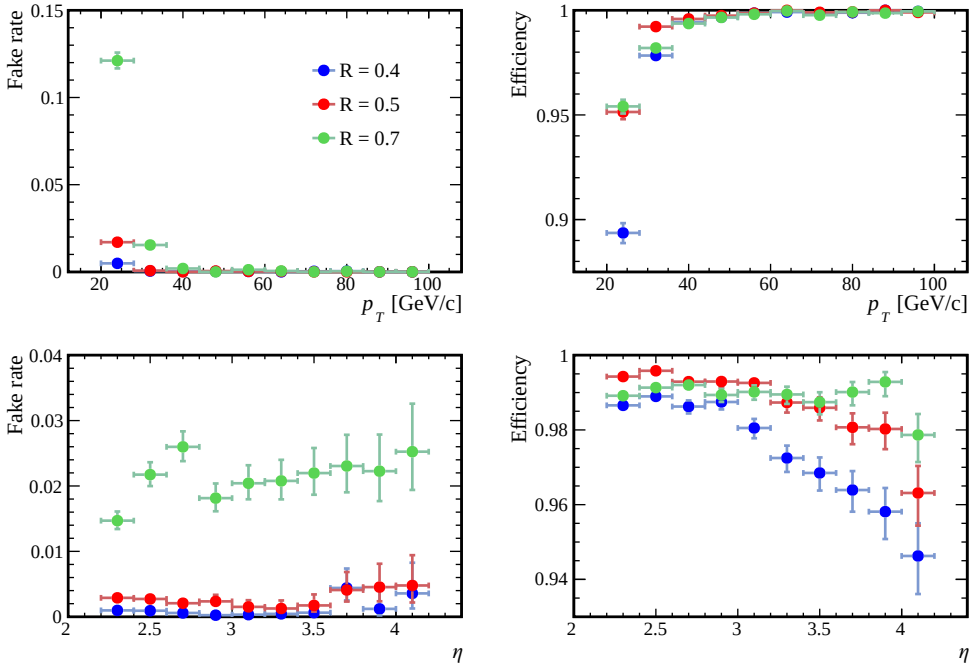



Fig. 4.13 $p_T$ (top) and $\eta$ (bottom) dependence of Run II jet reconstruction fake rate (left) and efficiency (right) for Turbo jets of different anti-$k_T$ radii: $R = 0.4, 0.5, 0.7$ (blue, red, green).

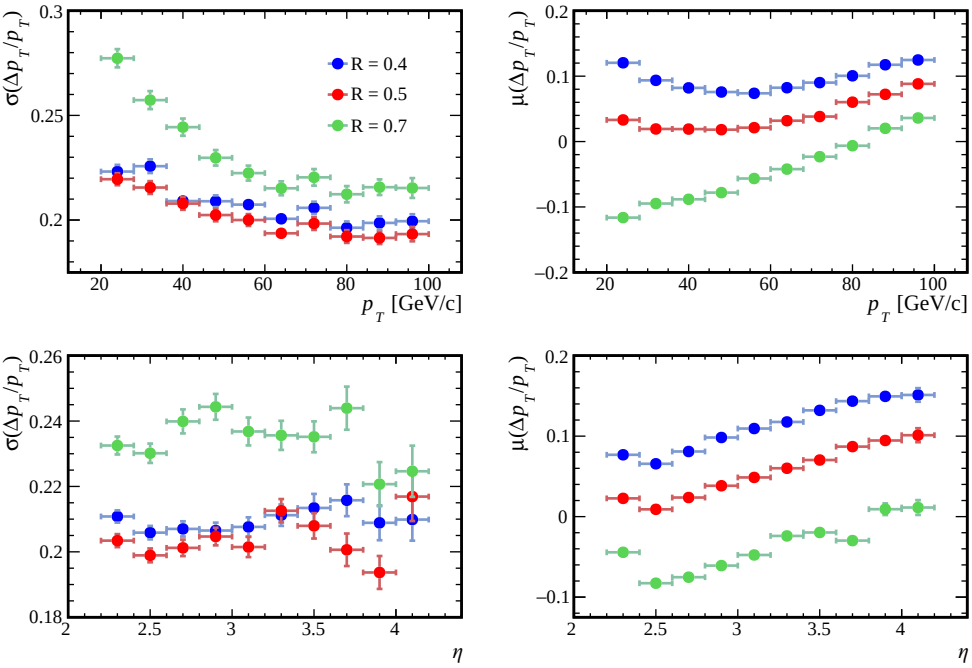



Fig. 4.14 $p_T$ (top) and $\eta$ (bottom) dependence of MC jet to reconstructed jet $\Delta p_T/p_T$ resolution (left) and offset (right) estimated from fitted Gaussian width & mean for Turbo jets of anti-$k_T$ radii: $R = 0.4, 0.5, 0.7$ (blue, red, green).

## 4.4 Reconstructed jet selection and energy corrections

The new definitions for HLT and Turbo jet configurations are optimised for reconstruction level performance using a set jet radius to match and even exceed the performance of Std. The following studies introduced post-reconstruction selections and jet energy corrections assessing the fake rate, efficiency and $p_T$ and directional resolutions at each stage.

### 4.4.1 Jet identification

Using an anti-$k_T$ $R = 0.5$, the performance of HLT and Turbo jets without filtered particle flow inputs is compared to equivalent jets reconstructed using filtered inputs, both with and without JetID selections (Chapter 3, Section 3.2.3). Figures 4.15-4.18 demonstrate the expected performance using particle flow input selection and post-reconstruction JetID cuts. The impact of the input filter upon fake rates and efficiencies has been discussed in Section 4.3.1 comparing Figures 4.1 and 4.2.

Following the JetID selection, as shown in Figure 4.15 & 4.17, the fake rate decreases $\mathcal{O}(0.1\%)$ and the reconstruction efficiency reduces by 4-5%. Despite such losses, the relative improvement in jet reconstruction efficiency of the input filtered HLT and Turbo jets over Std is maintained following JetID cuts as Std experiences similar losses. The inclusion of the input filter degrades the $p_T$ resolution of HLT by 0.5-1.0% (Figure 4.16) while improving that of Turbo 1-2% (Figure 4.18). The JetID requirements then degrade the $p_T$ resolution of HLT by $\mathcal{O}(0.1\%)$ and Turbo by $\mathcal{O}(1\%)$. It can be shown that any change to relative directional resolutions for each configuration with the JetID cuts applied is negligible and that Std experiences similar effects to its $p_T$ resolution as the HLT configuration.

The filtered input HLT and Turbo jets with JetID cuts applied are compared directly to Std jets with their own JetID cuts in Figures 4.19-4.20. Figure 4.19 shows the sub-percent fake rate of both HLT and Turbo configurations relatively unaffected by the application of JetID requirements; for each configuration, the fake jets are still concentrated $p_T < 30\,\text{GeV}$. The cost in terms of efficiency for applying JetID cuts is also fairly consistent across configurations, with the efficiency advantage of HLT and Turbo over Std jets preserved at $\sim$2%. Figure 4.20 shows the performance of HLT and Turbo in terms of $p_T$ resolution, relative to that of Std, remains unaffected by JetID requirements when compared to just including the input filter. While the offset in reconstructed $p_T$ due to JetID is less significant than that of including just the input filter, differences across the configurations are addressed with energy calibrations specific to each (Section 4.4.2).

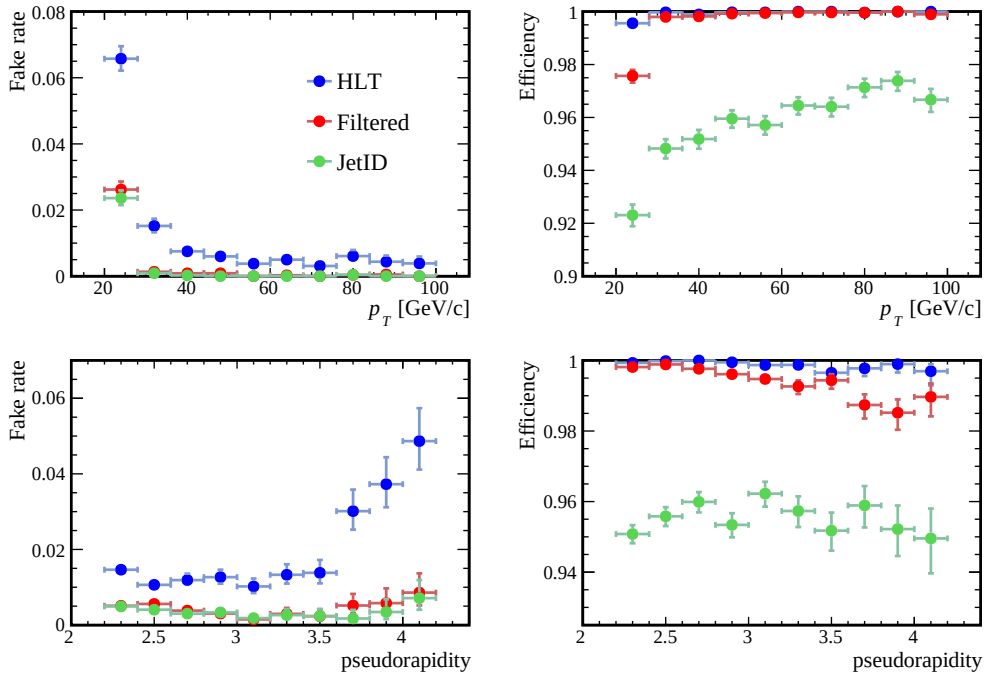



Fig. 4.15 $p_T$ (top) and $\eta$ (bottom) dependence of jet fake rate (left) and efficiency (right) for HLT (blue), with the new input configuration (red) and including JetID (green).

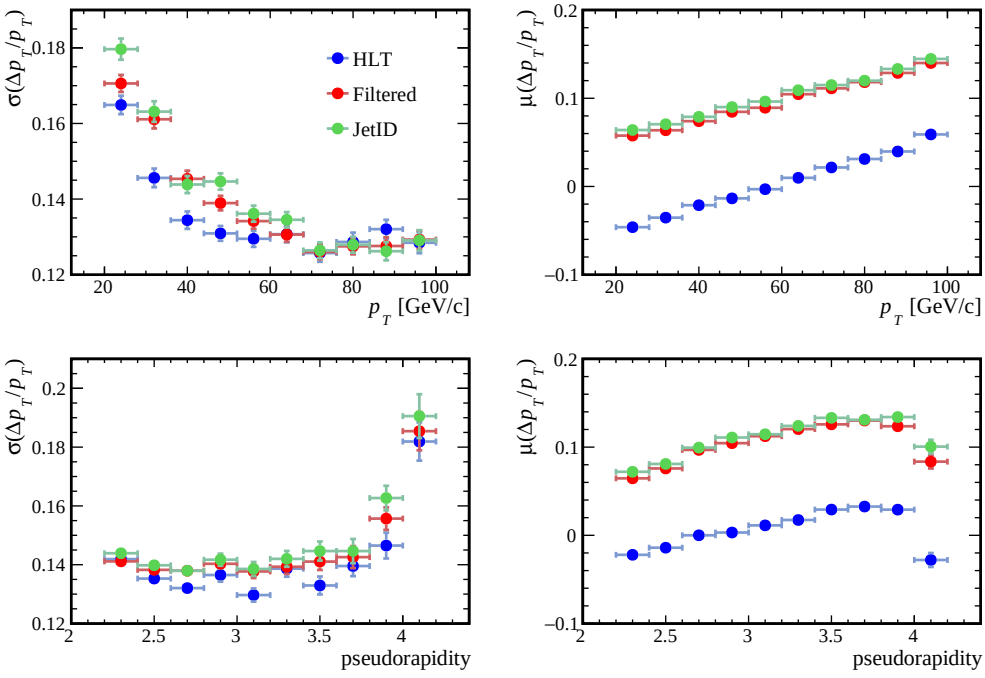



Fig. 4.16 $p_T$ (top) and $\eta$ (bottom) dependence of MC jet to reconstructed jet $\Delta p_T/p_T$ resolution and offset (estimate from fitted Gaussian width & mean) for HLT (blue), with the new input configuration (red) and including JetID (green).

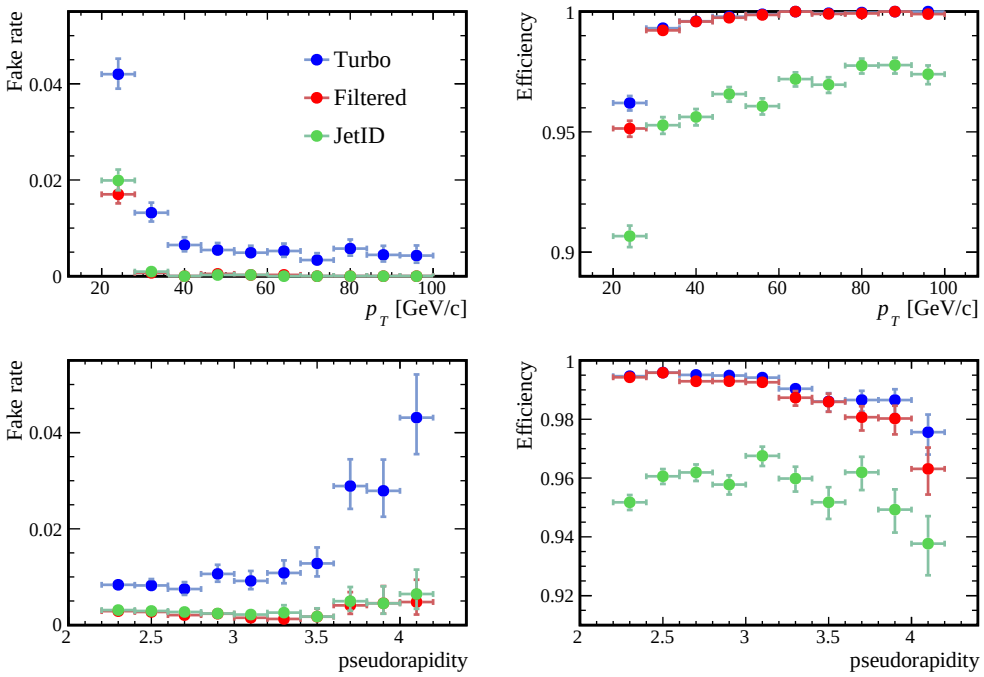



Fig. 4.17 $p_T$ (top) and $\eta$ (bottom) dependence of jet fake rate (left) and efficiency (right) for TURBO (blue), with the new input configuration (red) and including JetID (green).

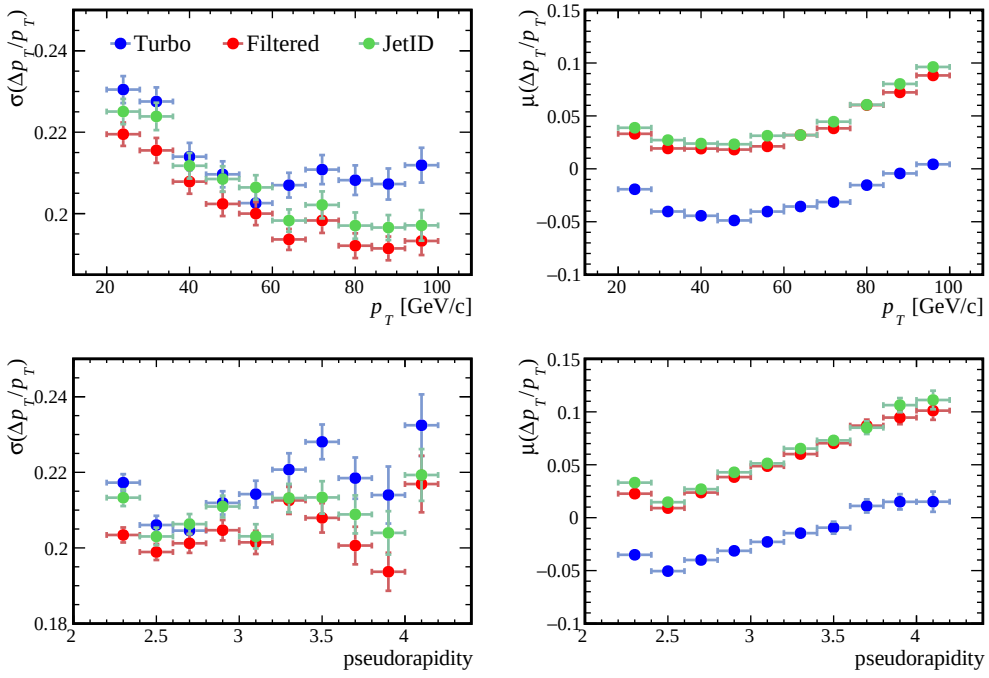



Fig. 4.18 $p_T$ (top) and $\eta$ (bottom) dependence of jet MC jet to reconstructed jet $\Delta p_T/p_T$ resolution and offset (estimate from fitted Gaussian width & mean) for HLT (blue), with the new input configuration (red) and including JetID (green).

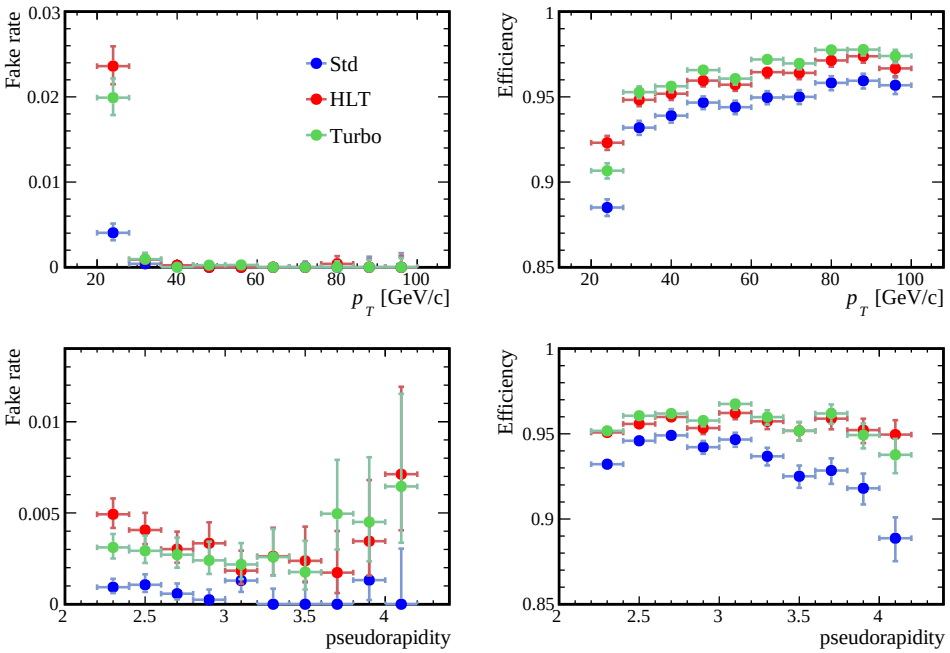



Fig. 4.19 $p_T$ (top) and $\eta$ (bottom) dependence of jet fake rate (left) and efficiency (right) for Std (blue), HLT (red) and TURBO (green) with their respective JetID cuts applied.

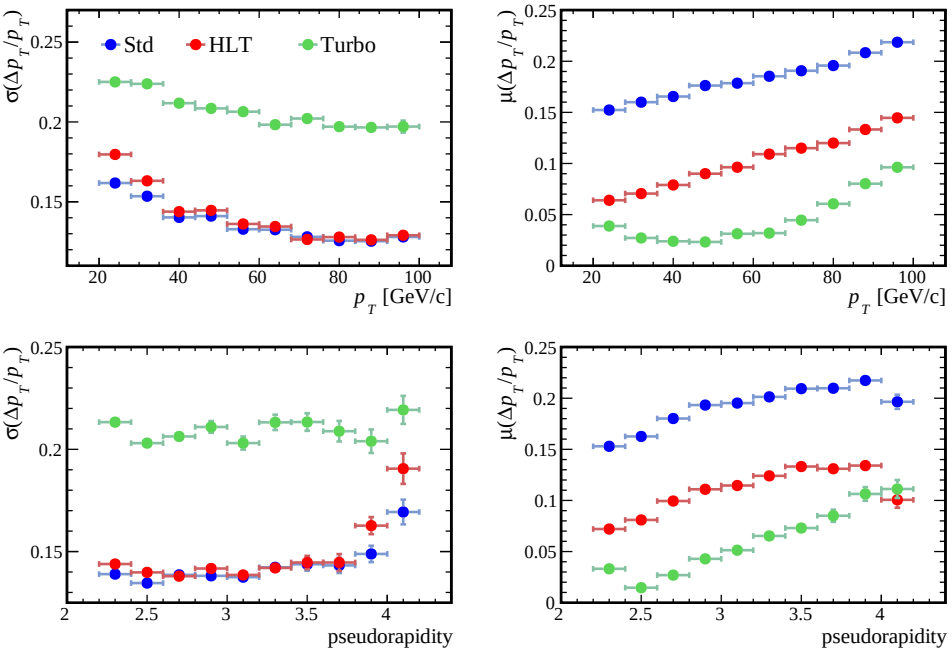
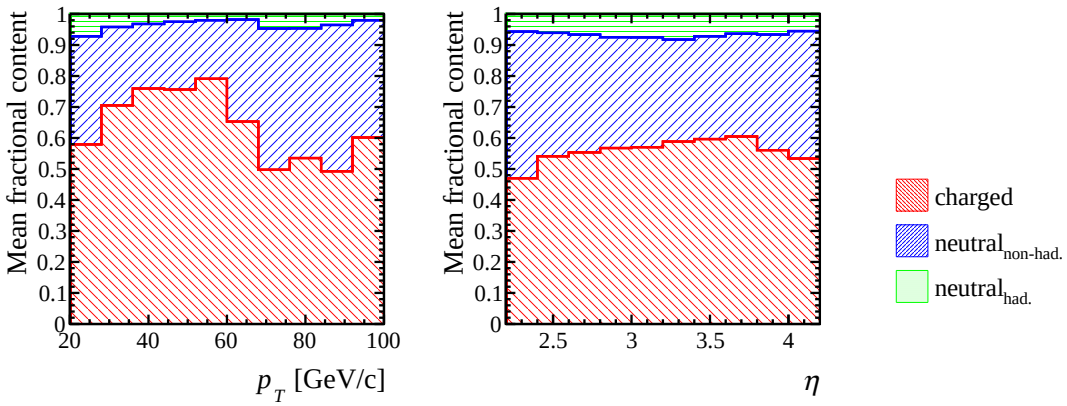
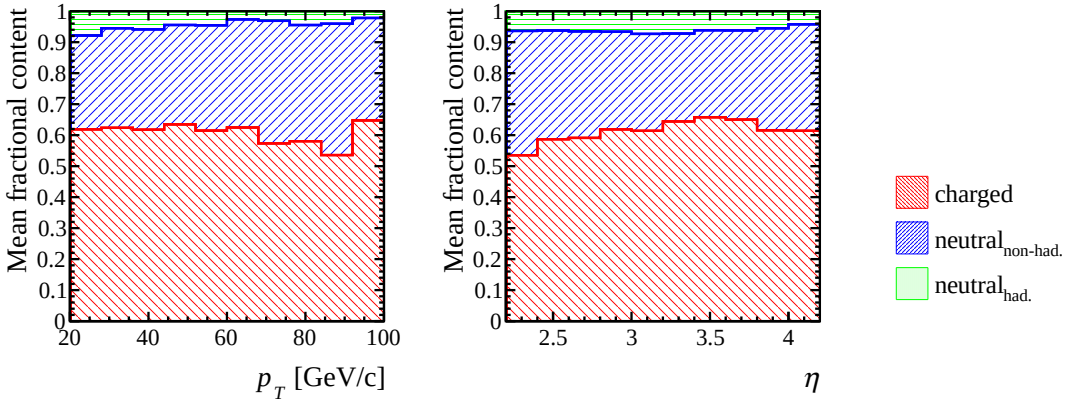



Fig. 4.20 $p_T$ (top) and $\eta$ (bottom) dependence of MC jet to reconstructed jet $p_T$ resolution and offset (estimate from fitted Gaussian width & mean) for for Std (blue), HLT (red) and TURBO (green) with their respective JetID cuts applied.

The composition of jets, based on both the constituents identified within the detectors and the respective fractions of the energy of the jet they carry, can be used to better understand the impact of the particle flow filter and the JetID requirements. Of the daughters produced within the jet, charged species include tracks and track-associated clusters, neutral species are split into hadronic and non-hadronic where remaining unmatched energy depositions are defined as neutral energy recovery (NER, Chapter 4). The average fractions of jet energy carried by daughters of different species registered in the detectors (therefore neutral non-hadronic includes $\pi^0 \to \gamma\gamma$ and excludes $\nu$) are shown in Figures 4.21-4.23 as functions of jet $p_T$ and $\eta$. The kinematic dependence of the fractional content is greatly reduced by JetID requirements, producing an approximate 60:40 split between charged (tracks and track-associated clusters) and neutral (track-isolated clusters and NER) components, the latter of which includes a 5-10% neutral hadronic contribution. While NER is offered in Std, the definition is absorbed into the non-hadronic neutral component for this configuration.

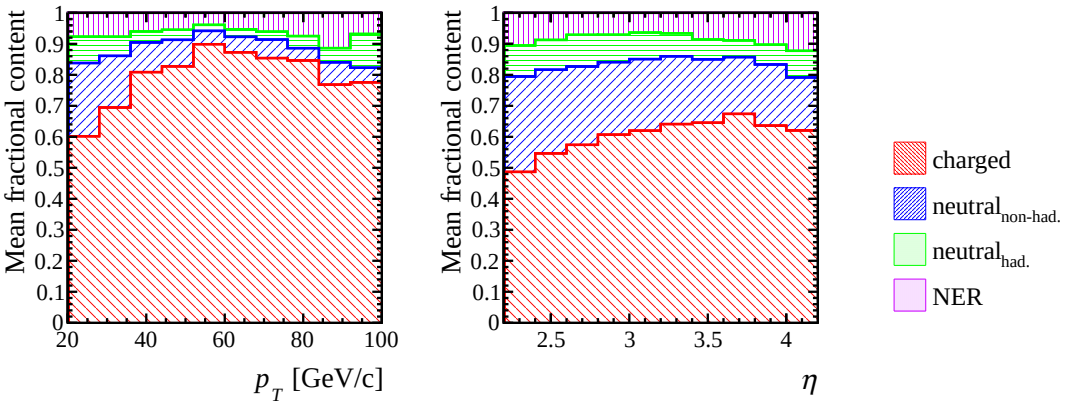



(a) Average energy composition of Std jet daughters

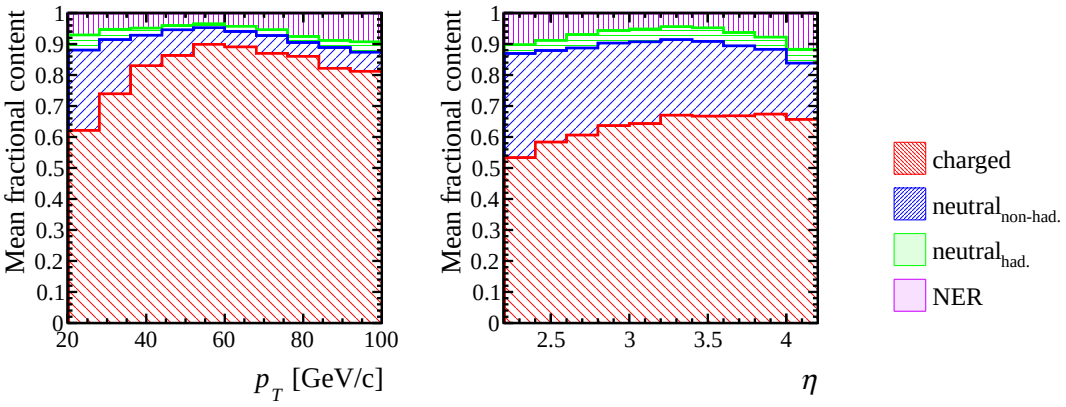



(b) Average energy composition of Std jets passing JetID

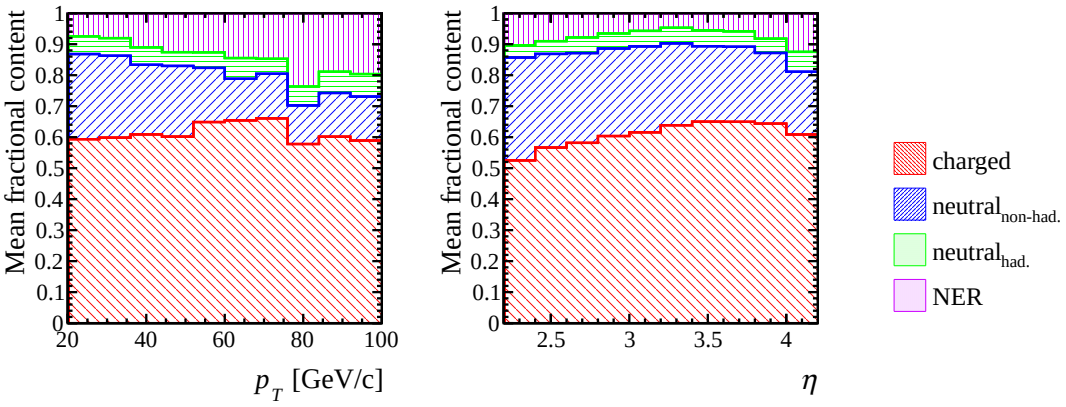
Fig. 4.21 Fractional energy content as a function of jet $p_T$ (left) and $\eta$ (right), in terms of charged (red), neutral hadronic (green) and neutral non-hadronic (blue) daughters for Std jets.




(a) Average energy composition of HLT jet daughters

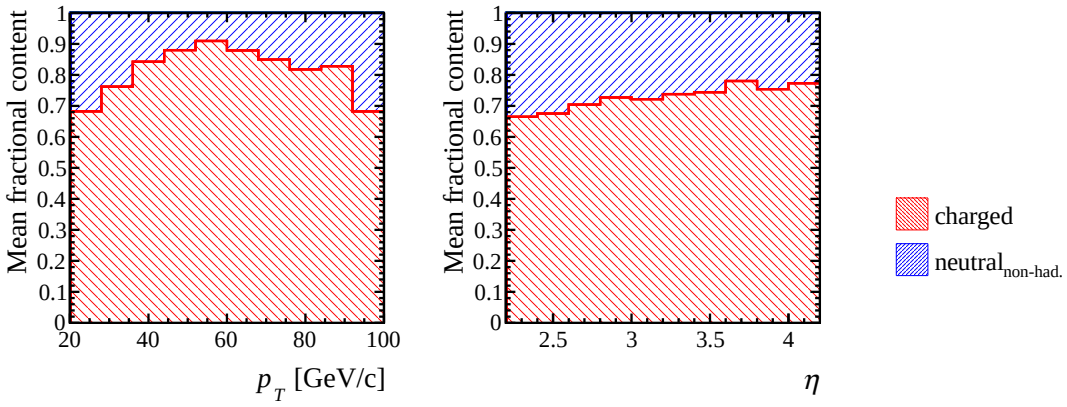



(b) Average energy composition of filtered HLT jets

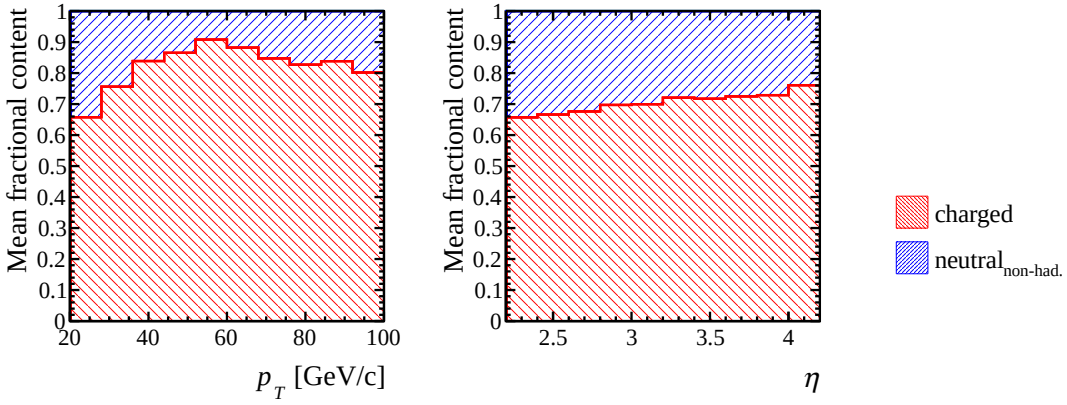



(c) Average energy composition of filtered HLT jets passing JetID

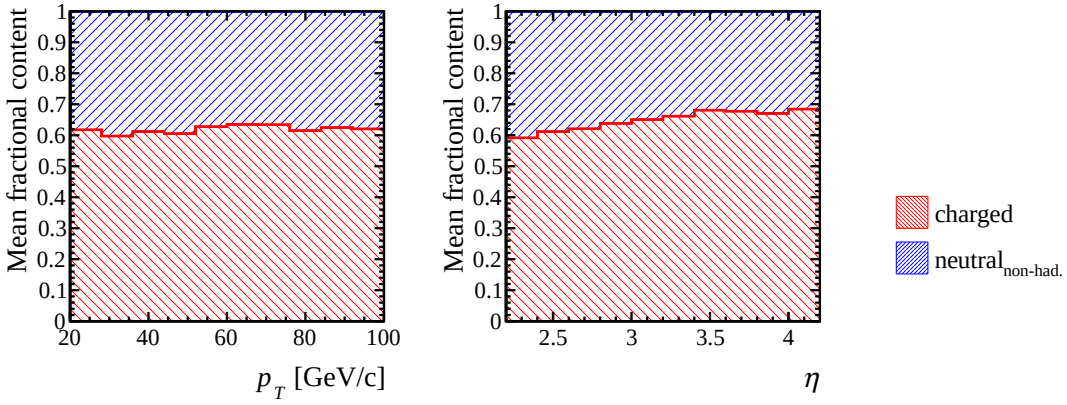
Fig. 4.22 Fractional energy content as a function of jet $p_T$ (left) and $\eta$ (right), in terms of charged (red), neutral hadronic (green) and neutral non-hadronic (blue) daughters and neutral energy recovery (purple) for HLT jets.




(a) Average energy composition of Turbo jet daughters




(b) Average energy composition of filtered Turbo jets




(c) Average energy composition of filtered Turbo jets passing JetID

Fig. 4.23 Fractional energy content as a function of jet $p_T$ (left) and $\eta$ (right), in terms of charged (red) and neutral (blue) daughters for Turbo jets.

Comparing Figure 4.22 (a) & (b), the input filter is shown to have little impact on the makeup of HLT jets and are shown to be dominated by their charged component in both. However, with JetID applied in (c), HLT replicates the flat 60:40 split with the constant neutral hadronic component observed in Std. The absence of HCAL information and, therefore, NER and hadronic components in Turbo jets are evident by Figure 4.23. Again, the largest change is between filtered input jets with and without JetID applied, and the post-reconstruction selection produces a 60:40 split with minimal dependence on jet $p_T$. The changes observed with the inclusion of JetID selections could be due to the removal of poorly reconstructed jets, which form outliers in the energy composition and skew the mean values. As discussed in Chapter 3, one example would be jets dominated by a single high energy lepton track. Despite the sacrifices in terms of efficiency, the provision of predictable $p_T$ independent jet composition justifies the inclusion of JetID cuts.

### 4.4.2 Jet energy corrections

As demonstrated in the 1D plots of $\mu(\Delta p_T/p_T)$ (Sections 4.3.2-4.4.1), there is a systematic offset in MC between the reconstructed jet $p_T$ and its true value. As discussed in Chapter 3, jet energy corrections (JECs) are calculated in MC and fitted as a function of uncorrected $p_T$. These functions are then interpolated to provide an energy correction to the jet four vector, preserving its direction.

The plots in Figures 4.24-4.27 include the normalised distributions, fake rates, Gaussian fitted $p_T$ resolutions and offsets of the jets using HLT particle flow. This provides comparisons to those with the JetID-based selection and with JECs, calculated and applied in terms of variables defined in Chapter 3 (including $p_T$ and $\eta$). With no significant impact on the fake rates in Figures 4.24-4.27, the reconstructed $p_T$ offset is flattened with respect to $p_T$ and $\eta$ at the cost of increasing $p_T$ resolution by 2-4% in HLT (Figures 4.24-4.25) and a 0.5-1.0% in Turbo (Figures 4.24). Figure 4.27 also shows that for $p_T < 50\,\text{GeV}$, this corresponds to a 0.5-1.0% resolution improvement but up to 4% degradation for jets around 100 GeV.

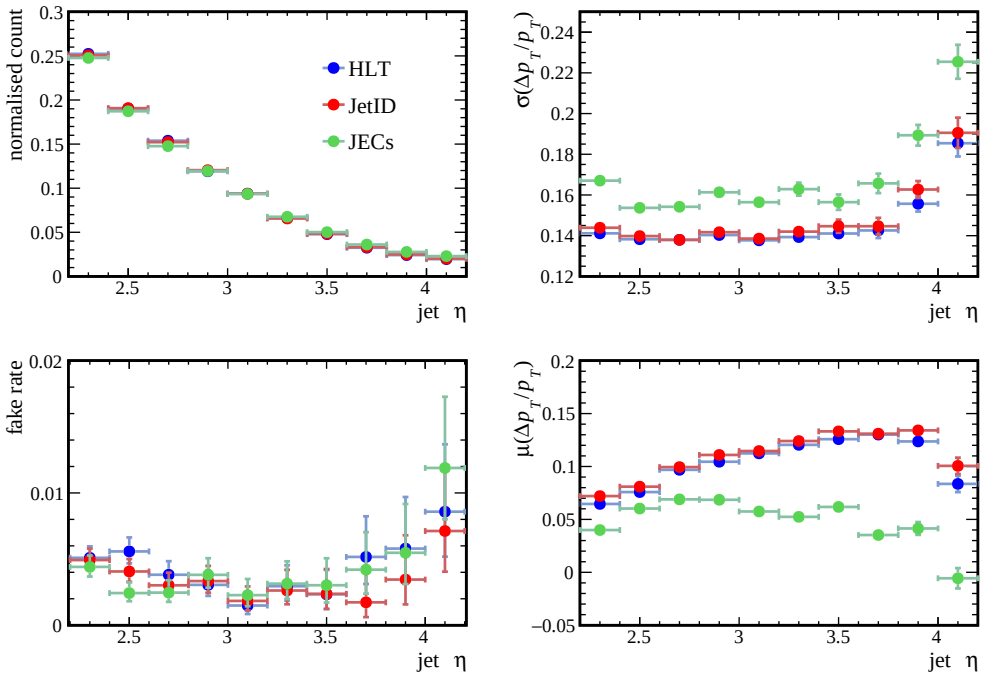



Fig. 4.24 HLT jets with filtered particle flow (blue), including JetID cuts (red) and jet energy corrections (green), their $\eta$ distribution and differential fake rate, $p_T$ resolution and offset.

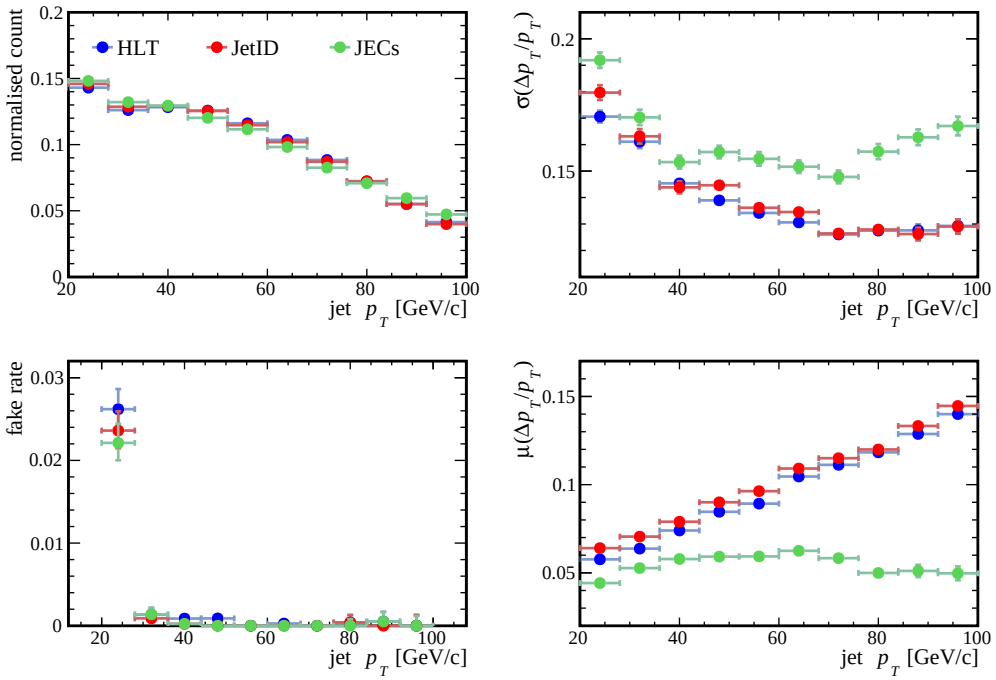



Fig. 4.25 HLT jets with filtered particle flow (blue), including JetID cuts (red) and jet energy corrections (green), their $p_T$ distribution and differential fake rate, $p_T$ resolution and offset.

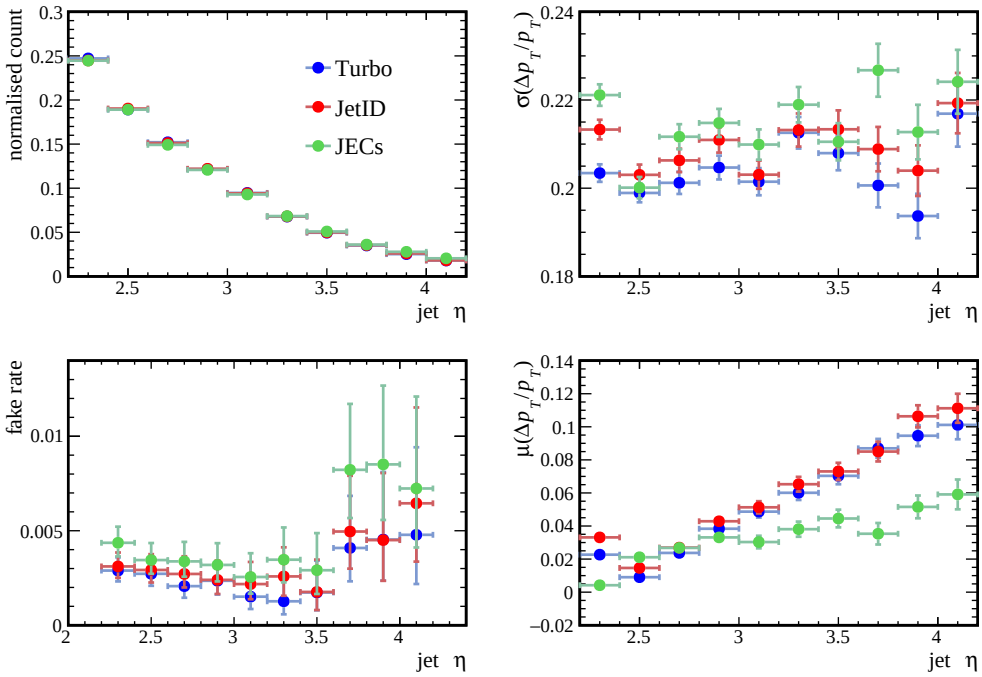



Fig. 4.26 Turbo jets with filtered particle flow (blue), including JetID cuts (red) and jet energy corrections (green), their $\eta$ distribution and differential fake rate, $p_T$ resolution and offset.

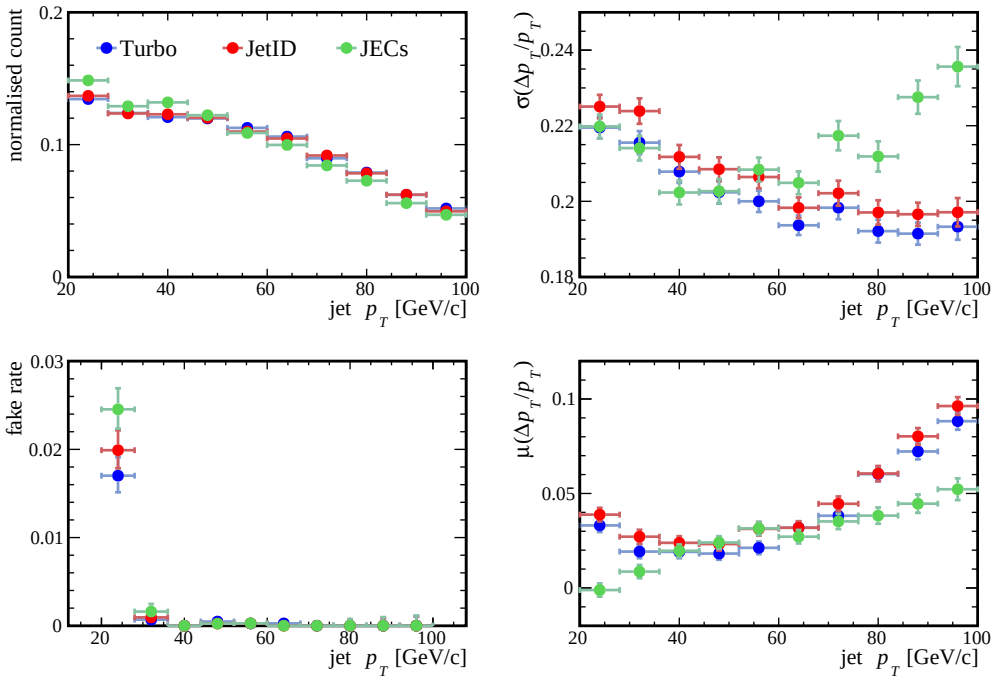



Fig. 4.27 Turbo jets with filtered particle flow (blue), including JetID cuts (red) and jet energy corrections (green), their $p_T$ distribution and differential fake rate, $p_T$ resolution and offset.

## Future work

Transitioning to HLT from Std provides a 1-2% boost in efficiency with negligible impact to fake jet rate for jets passing the top decay $b$-jet $p_T$ threshold of 50 GeV (Figure 4.2). However, for more general jet physics at lower energies, some areas may be investigated further to improve overall reconstruction performance. For the studies in Section 4.3.3, the jets of each radius (0.4, 0.5, 0.7) have the same input quality control applied optimised for $R = 0.5$. Full input selection optimisations for HLT and Turbo jets using different anti-$k_T$ radii may provide a range of jet definitions suited to various studies with different physics goals.

If the track-type composition of the charged daughters of a jet is ascertained, whether downstream tracks are a relatively significant contribution could be discerned. It could then be inferred whether or not the remaining fake jet reconstruction excess, $\mathcal{O}(1\%)$ for jet $p_T < 30$ GeV (Figure 4.2), observed in new configurations over Std jets is not ghost-track seeded in nature. This could be based on the higher ghost content of downstream tracks and their negligible impact on the fake jet rate. Assuming the relevance of persistent ghost content remains ambiguous, a more comprehensive study of the track-based requirements (Figure C.1) may yet address their effects on jet reconstruction performance. A compromise between training a classifier specific to each track-type and relying on a generalised algorithm may also be reached. Though the information from each track based variable in Table 4.1 is, in principle, encoded within the general track discrimination power of `GhostProb`, one course of action might be to perform linear regressions to optimise a selection placed on `GhostProb` and its input variables for each track-type individually.

An $E_T$ threshold applied to HCAL clusters corresponding to values of $\eta > 4$ in Std is not replicated in HLT jets to maintain uniform input requirements across the acceptance. Given the almost uniform distribution of fake jets as a function of $\eta$ (Figure 4.2), this is unlikely to account for the increased fake reconstruction over Std. However, the $\sim 1.5\%$ increase in $p_T$ resolution observed in HLT with $\eta > 3.5$ (Figure 4.6) could correspond to this difference in selection. Therefore, it may be advisable to investigate the impact of calorimeter resolution as a function of $\eta$ & $\phi$. This information would help quantify effects, potentially due to incomplete detector geometry, which would not be accounted for in the ERFs and map regions where stricter selection criteria should be applied. The nature of the edge effects may be confirmed by comparing the reconstructed jet composition in the same format as the reconstructed jet configurations compared in Figures 4.21-4.23.

The potential detector geometry effects and those of the inverse $p_T$ ordered clustering on the spatial resolution profiles of $\eta$ and $\phi$ is not yet clear. Observing the relative weight of the Gaussian and Lorentzian components to the $\Delta\eta$ and $\Delta\phi$ pseudo-Voigt fits as functions of $p_T$

and $\eta$ may provide a better understanding. An analytical argument may yet be constructed based on the coordinate system definitions to better inform the choice of distributions. A comparison may be drawn between the cumulative distribution function of hyperbolic secant function (HSF) and the equation relating angle from the beamline, $\theta$, and pseudorapidity, $\eta$, shown in Equation 4.2. Given the resemblance of the Voigtian distribution to the HSF [78], this may go some way in explaining why $\chi^2$ values favour the pseudo-Voigt distributions used to fit the $(\eta,\phi)$-residuals. The Gaussian component in the Voigtian fits has provided a convenient measure of the width of the distribution of residuals.

$$\begin{aligned} HSF(x) &= \frac{2}{\pi}\arctan\left(\exp\left[\frac{\pi}{2}x\right]\right) \\ \theta &= 2\arctan\left(\exp[\eta]\right) \end{aligned} \tag{4.2}$$

The impact of spatial resolutions is absorbed into acceptance factor systematic uncertainties, which are dominated by the $p_T$ resolution. As shown in Chapter 6, the systematic uncertainty associated with the jet reconstruction resolutions is $\mathcal{O}(1\%)$ and is sub-leading for the analysis discussed. However, more rigorous jet diagnostics may become relevant for future analyses.

LHCb GitLab

# Chapter 5

# Run II heavy flavour tagging

Heavy flavour quarks travel a short but discernible distance before decaying. As described in Section 5.1, the VELO provides LHCb with secondary vertex (SV) reconstruction by identifying the shared point of origin of the decay products displaced from the primary interaction. SVs are indicative of $b$- & $c$-quarks (heavy flavour) content and can be reconstructed within jets. Flavour classification models, including boosted decision trees and deep neural networks, have been developed using Run II MC, as shown in Section 5.2, to identify heavy flavour jets based on secondary vertex tagged jet information. The flavour classification performance using SV-tagged HLT jets is assessed in Section 5.3 for use in the final analysis.

## 5.1 Secondary vertex reconstruction

Following the reconstruction of PVs in the VELO, tracks are subject to an additional $p_T$ threshold of 0.5 GeV before being used in SV reconstruction. An $\chi^2_{IP} > 16$ is required between the remaining tracks and the PV to ensure they are dissociated. Any remaining pairs of tracks with distance of closest approach $< 0.2$ mm are combined into two-track SVs. These 2-body vertices must fit with their own $\chi^2_{IP} < 10$ and satisfy a two-body mass $0.4\,\text{GeV} < M < m(B_0)$, where $m(B_0)$ is the nominal $B_0$ mass. Those SVs passing this selection are associated with any jet with which they have $\Delta R < 0.5$ (though individual tracks may lie outside the radius relative to the jet axis) [76].

A linking procedure iterates over any 2-body SVs in a jet and those which share tracks are merged into $n$-body vertices until no remaining SVs with $\Delta R < 0.5$ share tracks. The merged $n$-track SV positions are taken as the average of their constituent two-track SV positions weighted by the inverse of their $\chi^2$ from the vertex fit ($\chi^2_V$). The $n$-body SVs are also required to provide significant spatial separation from the PV and lie within a region consistent with ($b$,$c$)-hadron decays. The candidate SVs are then rejected if they either contain more than

one track outside the jet radius or are composed of only two tracks and reconstructed with a mass consistent with a $K_s^0$ meson ($497.611 \pm 0.013$ MeV [1]). Scattering events and $s$-decays are rejected by requiring $d_f/p_T < 1.5\,\text{mm/GeV}$ [76] using flight distance ($d_f$) as a proxy for hadron lifetime.

SV-tagged events from MC samples of $Z$+jet, containing $b$- and $c$-jets and light-jets ($udsg$) are used to assess the performance of new and existing machine learning (ML) models (Section 5.2). Firstly, the SV-tag rate of heavy-flavour jets and mis-tag rate of light-jets is calculated (Figure 5.1). This demonstrates the inherent efficiency limitations imposed by dependence on existing SV-jet definitions, with $b$-jets SV-tag rate $\sim 70\%$ and $c$-jets $\sim 25\%$. However, the efficacy with which an SV-tag requirement suppresses light-jets is also shown, with an SV-tag rate $\mathcal{O}(1\%)$.

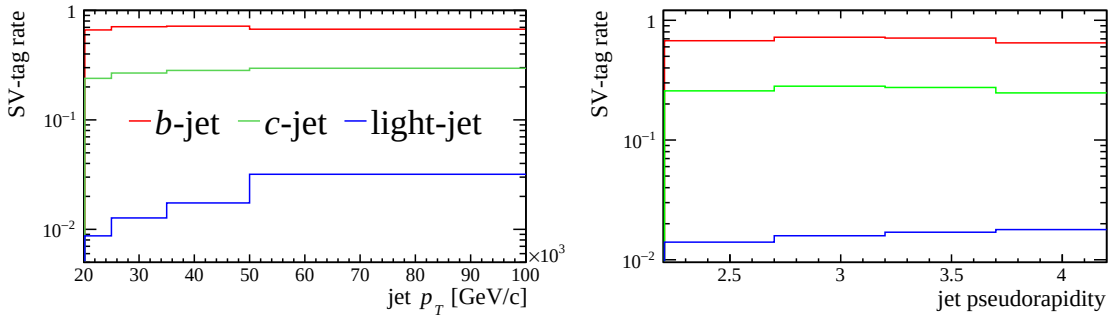



Fig. 5.1 SV-tag efficiencies for $b$- & $c$-jets and mis-tag rate for light-jets as a function of jet $p_T$ (left) and of jet pseudorapidity (right).

## 5.2 Multivariate classification

Techniques for optimising a decision process, reliant upon differential distributions of observables, include regression and classification; these offer continuous and discreet outputs respectively [79]. A variety of ML applications provide solutions to classification problems for data with non-linear relationships in a multi-dimensional space, whereby a model of input data is improved upon through an iterative training process. Using such methods, the information from a range of observables can be combined, aiding separation between defined classes of events. Variables which offer the biggest differences between classes, often signal versus background processes, provide the greatest sensitivity for discrimination.

The SV corrected mass ($M_{cor}$), the minimum mass that accounts for missing particles and satisfies the flight direction, is defined as:

$$M_{\text{cor}} = \sqrt{M^2 + p^2 \sin^2\theta} + p\sin\theta, \tag{5.1}$$

where $M$ and $p$ are the invariant mass and momentum of the particles that form the SV and $\theta$ is the angle between the momentum and the direction of flight of the SV [76]. $M_{cor}$ is one of the variables used for the Run I HF-tagger inputs. Others include: $\chi^2_{FD}$, the measured flight distance divided by its uncertainty and the sum of track $\chi^2_{IP}$ of the SV [76]. Those variables used as inputs for producing new MVA classifiers of $(udsg|bc)$ for light-rejection and $(c|b)$ for HF-discrimination are as follows:

- `tau` - the secondary vertex lifetime
- `m` - the SV mass
- `mCor` - the SV corrected mass
- `mCorErr` - the uncertainty on the SV corrected mass
- `ptSvrJet` - the fraction of the jet pT carried by the tracks of the SV
- `pt` - the SV $p_T$
- `nTrk` - the number of tracks in the SV
- `nTrkJet` - the number of SV tracks with $\Delta R < 0.5$ relative to the jet axis
- `fdrMin` - $\Delta R$ between the SV flight direction and the jet
- `drSvrJet` - the transverse flight distance of the two-track SV closest to the PV
- `fdChi2` - the flight distance $\chi^2$
- `ipChi2Sum` - the sum of all SV track $\chi^2_{IP}$

The MVA input variables, the information each class sample provides in training, are required of each event for the model prediction. The parameterisation of the learning procedure, the structure and imposed constraints, are coined the hyper-parameters. The extent to which each set of parameters may impact the efficiency of the learning process and final performance of the model is assumed to be problem-specific. The quality of information available is dictated by the input parameters and the choice of samples. The ML

algorithm relies upon a representative training set to form a model to apply to unseen data. ML algorithms rely upon monitoring the cost of their correct and incorrect decisions. In the case of those discussed in this chapter, training data provides the true answers (supervised learning) allowing the model to map a function (loss) upon which it optimises itself. On the other hand, hyper-parameters are typically tuned to optimise the training process in terms of both minimising iterations required and maximising absolute performance reached.

### 5.2.1 Gradient boosted decision trees

Decision trees are used to approximate a function of a given parameter space to distinguish between classes. At each layer of a decision tree, it bifurcates (forming two branches) into more specific regions called nodes that form the input of the next layer. A decision tree will provide a binary output for each event based on which classification the final node is defined as [80]. Single trees suffer from systematic uncertainties due to migrations across class boundaries. These produce behaviour which is sensitive to even small changes in the training sample including statistical variation, known as model instability. Splitting the data set and training many trees over independent sub-samples provides the means to take a weighted combination of an ensemble of tree outputs. This avoids problems associated with over-specifying a model to noisy data sets (over-fitting) or a strict binary output (involving trade-offs between variance and bias) [81]. Applying this technique provides a continuous output that acts as a new differential distribution, combining the separation powers of the learned inputs.

Rather than fully accumulating and weighting a randomly generated ensemble (random forest), trees can be added to an ensemble as they are generated in a process known as boosting. If each tree is based on the residuals of the previous tree, in terms of components of the gradient of the loss function, then producing and combining the individual models with poor performance (weak-learners) is known as gradient boosting [82]. A ROOT integrated Toolkit for MultiVariate Analysis (TMVA) provides a gradient boosting implementation [75] for the boosted decision trees (BDTs) [80] produced in this work.

### 5.2.2 Deep neural networks

Designed to emulate the neural pathways of the brain by propagating signals for pattern analysis, artificial neural networks (NNs) are composed of inter-connected computational units, or nodes (Figure 5.2). The series of connections between nodes define consecutive layers in a network. The weighted inputs of each node provide the signal strength at that point in the network; the weights are updated iteratively during the learning process. The sum

of the input weights are converted at each node using an activation function, which provides a normalised signal to propagate to subsequent nodes [83]. Typically, the first layer of nodes takes in the training variable values as the initial signals. The hidden layer is expected to discern relationships between inputs through its connections with the previous layer and conveying those to the output. Extracted at the final layer, the results are derived from the values of the nodes. The model of the data encoded within the neural network is refined through the information it assimilates into the tensor of weights throughout training.

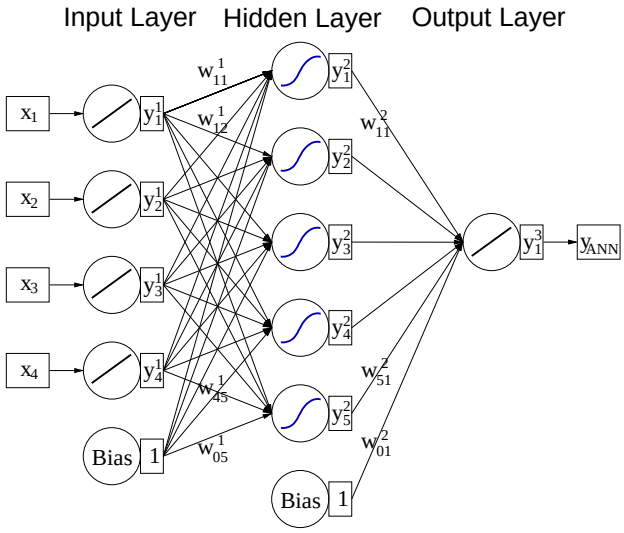



Fig. 5.2 Structure of a neural network, where $x_i$ are the inputs providing signals propagated through the network where the nodes, represented by circles (indicating typical activation functions), are connected in consecutive layers and attributed weights, $w^j_{mn}$, applied to each node output, $y^j_i$ [75].

By adding intermediate hidden layers, networks can access attributes that would have acted as output from the previous layer [84]. These attributes can become more and more abstract the more layers the network has. As a result, 'deep' networks can exploit features of data regardless of human intuition or awareness. The characterisation of any learning a model might achieve, contained within its potentially vast number of parameters, becomes obfuscated. As a model becomes increasingly complex, its susceptibility to over-fitting becomes more of a concern. How each layer is connected with each of the others and the size of (or number of nodes in) each is a problem-specific issue for optimisation; even activation functions may take a variety of forms. The breadth of potential configurations (the number of layers, nodes per layer, choice of activation functions) for neural networks is limited in use by both the practicalities of producing such a model with available computing resources and maintaining an understanding of the its behaviour and performance.

Sequential models are formed from layers connected consecutively. Dense layers are defined as operating with nodes connected to every node in the previous layer and subsequent layer. Each layer may have different numbers of nodes connected to both the input, output and one another. Each node can be assigned a probability to propagate a signal value of zero at random in a process called dropout regularisation. Regularisation allows models to develop in a more generalised way. Dropout is a computationally efficient way of avoiding over-fitting. A penalisation summed into a loss function can be applied to large weights in a layer in order to combat over-fitting; this is known as kernel regularisation [85]. For the DNNs in this work, such functionality is provided by Keras [86], an application programming interface for Tensorflow [87]. Keras offers several minimisation algorithms which, applied to the loss function, reach a local minimum efficiently. ADAM, implemented in the model training later in this chapter, provides adaptive learning rates suitable for gradient descent, a first-order iterative optimisation algorithm, in deep learning tasks [86].

### 5.2.3 Feature selection

Training times and tendency towards over-fitting can be reduced through the implementation of simplified models. Correlation analysis of used sample classes can be used to determine which variables to retain. The correlation matrices, $M$, of the training classes were investigated in an attempt to identify potential candidates for removal, simplifying the model by reducing the number of parameters. Only the SV lifetime, SV $p_T$ and SV corrected mass uncertainty had not previously been included for the Run I classifiers. It is assumed that if the absolute value of an element of the correlation matrix ($M_{ij}^{|\cdot|}$) is large then the parameter with the weaker correlations with the remaining inputs should be retained. However, it is also assumed that maximising the difference between signal and background class matrices ($\Delta(M)_{sig,bkg}$) is beneficial to discrimination power, therefore $M_{sig}^{|\cdot|}$, $M_{bkg}^{|\cdot|}$, $\Delta(M)_{sig,bkg}$ and $\Delta(M^{|\cdot|})_{sig,bkg}$ were each considered.

All features were shown to provide at least one correlation difference between classes of greater than $\pm 10\%$. Although consistently correlated with one another, the variables `fdChi2Sum` (`pt`) and `ipChi2Sum` (`ptSvrJet`) each have dominant correlations across all variables in the light-rejection and HF-discriminating samples respectively; for the sake of using consistent inputs between classifiers however, each of them are retained. All the variables listed at the start of Section 5.2 are included in the training of models discussed in subsequent sections.

### 5.2.4 Sample pre-processing

The input distributions of each of the data-sets undergo transformation such that the signal class shapes take on a normal distribution centred at zero with a variance of one [75]. While BDTs are scale and shift invariant, the rate of convergence for NNs can be shown to increase using inputs over a regular range [88]. Scanning over the cumulative MVA responses of each class provides signal efficiency versus background rejection functions known as receiving operator characteristic (ROC) curves. The ROC curve also provides a metric of performance by taking its integral, the area beneath it (AUC), with values 0.5-1.0 ranging from random assignment to perfect discrimination. ROC AUC forms the primary metric for model assessment in these studies.

A method called cross-validation enables each model to be optimised with unseen data, used to calculate the loss during training. The available data is split into statistically independent subsets for training versus validation performance. In terms of loss per iteration, the divergence of the training sample performance from the validation sample indicates over-training. When the validation loss function gradient with iteration (or epoch) reaches zero, further training is unnecessary. When over-training occurs, the validation loss as a function of the epoch may start to rise. The comparison of training versus validation loss is monitored to curtail training for each model appropriately.

Training samples for the classifiers were taken from di-jet events in MC, where the two jets share a common PV from $pp$ collisions at 13 TeV. Each jet was considered independently, with $20 < p_T < 100\,\text{GeV}$ and $2.2 < \eta < 4.2$ required and the truth level flavour dictating the classes. Compared to heavy-flavour discrimination ($b \sim 1.9 \times 10^6 : c \sim 8.7 \times 10^5$), the training samples for light-rejection are very statistically imbalanced ($HF \sim 2.7 \times 10^6 : udsg \sim 6.5 \times 10^4$) as coincidence SVs in light-jets are relatively rare. Under-sampling involves the larger class sample only making use of events equal in number to the size of the limited class sample. The shape of the DNN response is closer to expectation when using under-sampled training sets. This is also found to provide improved performance in terms of ROC AUC (Figure D.6) for both BDTs and DNNs and, as a result, training for light-jet rejection is performed using under-sampled classes ($HF \sim 6.5 \times 10^4 : udsg \sim 6.5 \times 10^4$).

Taking the logarithm of variables with sharp, skewed peaks as replacement inputs improves the performance of the TMVA BDT models. No such improvement was demonstrated for the DNNs. For model development and comparison, several variables will remain transformed such that the benchmark set by the Run II retrained TMVA BDTs includes this improvement. Future models, including DNNs, are based on these transformed variables.

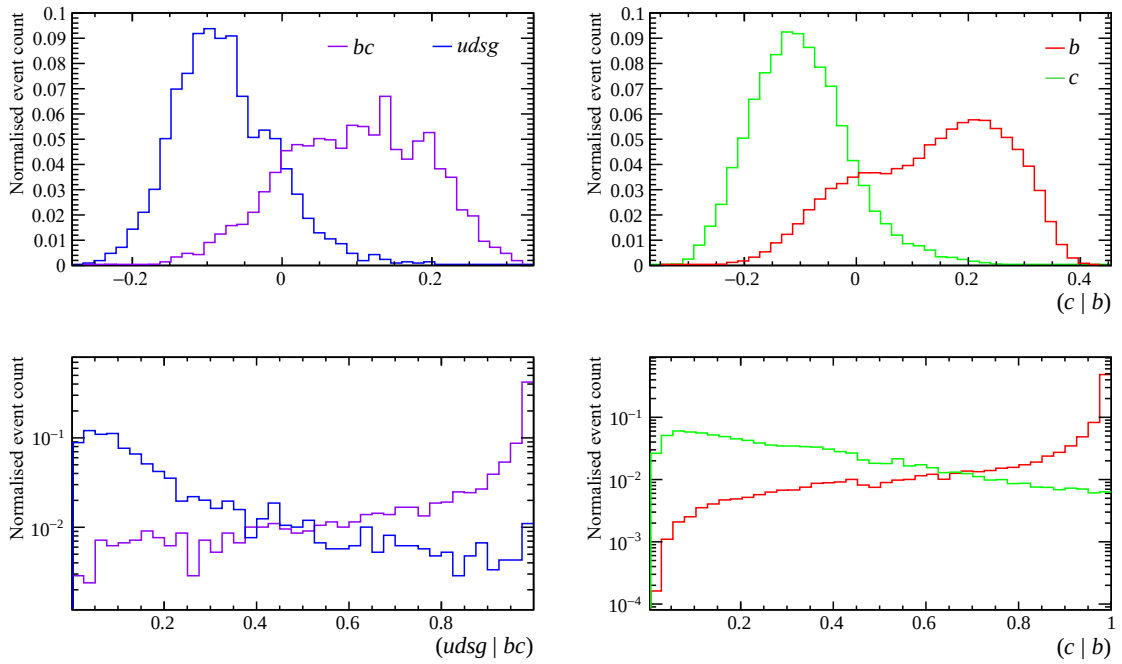



Fig. 5.3 Di-jet training class response for the highest performing BDTs (top) and DNNs (bottom), where $(udsg)$ is blue, $(bc)$ is purple (left), $(c)$ is green and $(b)$ is red (right)

### 5.2.5 Model comparison

Fits to MVA response in $W$+jet and top candidate data using templates from $t\bar{t}$ MC, $W$+jet MC and $Z$+jet MC, each split by truth level $b$, $c$ and light ID, were projected in $\eta$ and $p_T$ of the muon and the jet. The $Z$+jet sample demonstrated the best agreement with data and provided another sample, unseen by the model in training, with which to assess model performance. $Z$+jet templates are used to produce flavour-specific ROC curves. As presented in Run I studies [76], these can then include factors of SV-efficiency estimating performance for $b$- & $c$-tagging individually as functions of each background-mistag. Templates produced from MC binned in jet $p_T$ and $\eta$ may also be used to compare the differential performance of each model with the area under the curve (ROC AUC) using the same method.

The Run II tagging algorithms outperform the default Std jets Run I trained BDTs. The new HLT jets DNNs consistently outperform the new BDTs. The Run II BDTs therefore set a benchmark for optimisation studies of training the DNNs. This optimisation will involve tuning the training samples as well as the hyper-parameters of each network. The performance of the models is consistently weakest in the range $p_T > 50\,\text{GeV}$, the same threshold used for rejecting $Wb$-backgrounds for studies of top decays. In Section 5.2.6, three input MC selections (all reconstructed jets, $p_T > 20\,\text{GeV}$, $p_T > 50\,\text{GeV}$) are optimised

individually to find if the balance between generalisation through a broader sample and specificity in high-$p_T$ training for HF-decays may be exploited.

### 5.2.6 Hyper-parameter tuning

The DNN hyper-parameters were optimised based on the performance outcomes of models trained using variety of constraints upon the network. For a range of ratios in which the class-labelled data is split between training and validation samples (test:train), a grid-scan of the following hyper-parameters was performed: number of hidden layers (of equal size); the number of nodes per hidden layer; learning rate. The grid used was sparse to save processing time (nodes in units of 10, learning rate in orders of magnitude). These scans were performed for the three training sample $p_T$ thresholds: 10, 20 and 50 GeV.

Adjusting the training sample minimum $p_T$ to 10 and 50 GeV each reduced the performance of the DNNs across all $p_T$ bins. While improvements to the BDTs were offered by the $p_T > 50$ GeV training set, the DNNs trained with each threshold outperformed the high-$p_T$ BDTs on jets with $p_T > 50$ GeV. The SV-efficiency scaled ROC curves[1] of each model trained with $p_T > 20$ GeV jets are shown in Figure 5.4 for jets $(20, 50) < p_T < 100$ GeV. The finalised DNN responses to events with a jet $p_T > 20\,(50)$ GeV are displayed in Figure 5.5. These samples are used as 2D templates with which to fit to the combined DNN responses in data to extract HF-yields as demonstrated in Section 5.3.

## 5.3 Heavy flavour yield extraction

By fitting to data using the same two-dimensional MVA response method from Run I [76], with the newly developed DNN models and 2D EW+jet MC templates, the *b*- and *c*-yields are provided through the fitted normalisation of each component. This method is tested using $(W \to \mu)$+jet events in 13 TeV data. This is necessary for identifying *b*-jets from top decays in Chapter 6 but, in the future, may also provide a cross-check of negligible charm contributions from top decays. Producing additional flavour templates in bins of jet pseudorapidity and $p_T$ provides a check of fit stability across the binned kinematic acceptance, both for jets in the top analysis acceptance ($2.2 < \eta < 4.2$ & $50 < p_T < 100$ GeV) and future differential jet measurements. Projections of the fits in the MVA training variables provides $\chi^2$ values to isolate potential sources of MC to data discrepancy. The $\chi^2$ values of fits to data, either binned in jet kinematics or integrated across the acceptance, and the consistency in the total yield between them attempts to localise the effect and assess its impact.

[1] SV-efficiency as a functions of MVA response is used to calculate an efficiency versus mis-tag curve.

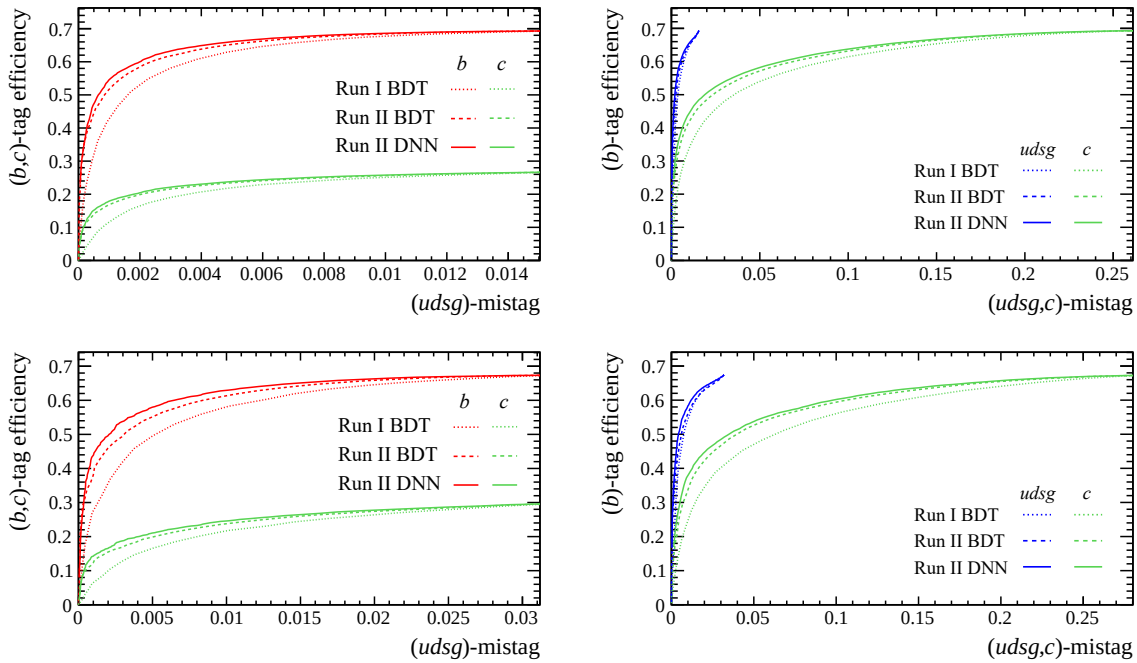



Fig. 5.4 SV-tag efficiency corrected ROC curves of light rejection (left) and $b$-tagging (right) for jets with $(20,50) < p_T < 100\,\text{GeV}$ (top, bottom).

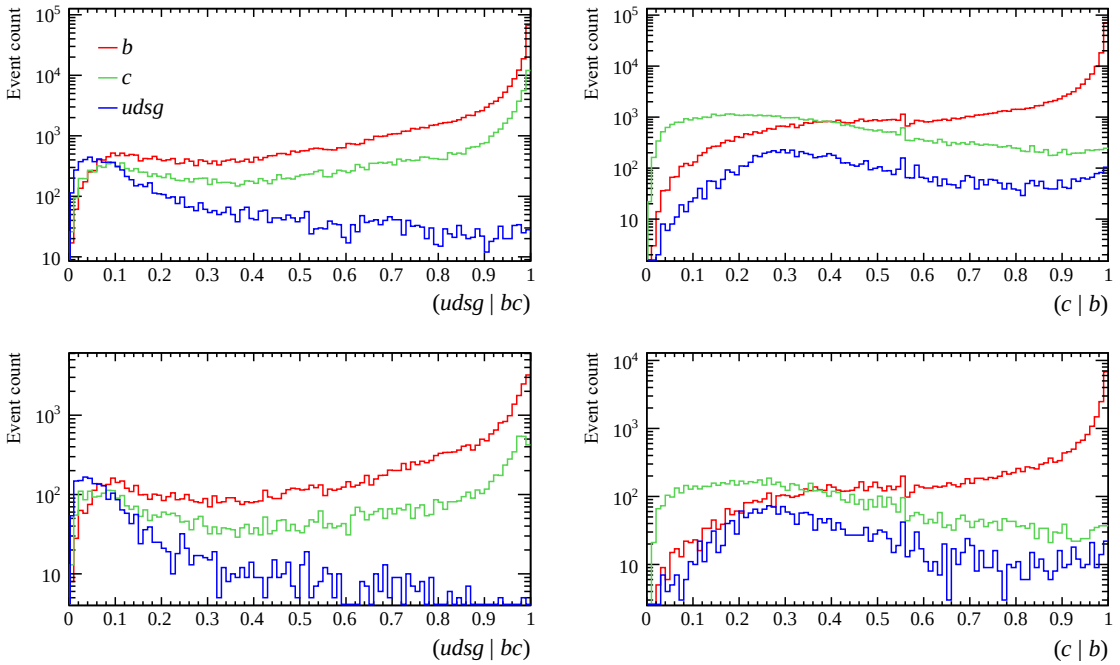



Fig. 5.5 Run II Z+($b$,$c$,light)-jet (red, green, blue) MC sample light rejection and $b$-tagging (left, right) DNN response templates to events with jet $(20,50) < p_T < 100\,\text{GeV}$ (top, bottom).

### 5.3.1 Template fits

2D distributions (Figures 5.6) are used in likelihood fits to data to extract the relative normalisations and therefore the HF-yields. By producing 1D projections of the data and the templates using the fitted normalisations, both in the DNN axes (Figure 5.6) and in each of the training variable axes, the disagreement between data and MC is assessed through the $\chi^2$ of the best fit projections for each MVA input. The projections of the fit in each MVA axis and the training variables are compared between the finalised MVA and an alternative to better gauge this discrepancy. The extent to which the DNN is susceptible to this difference and its impact on the template fit method is assessed using alternative fits.

Replacing the template axes with the corrected mass and number of SV-tracks (Figure 5.7) provides an alternative fit (ALT) allowing an estimate of a systematic uncertainty on the yields [76]. It was found that the $\chi^2$/NDF values of the training variable projections from these alternative template fits implied closer agreement with data in `drSvrJet`, `tau`, `fdrMin` and `ptSvrJet` while the DNN fit had closer agreement in `fdChi2`, `m`, `ipChi2Sum`, `nTrkJet`, `pt` and `mCorErr`. Yields from the two template fits were found to be inconsistent within error, thus the variation under the exchange of the choice of 2D templates is used as a systematic on the fitted yields, as discussed in Chapter 6. The jets in both data and MC templates satisfy $2.2 < \eta < 4.2$ and $20 < p_T < 100\,\text{GeV}$.

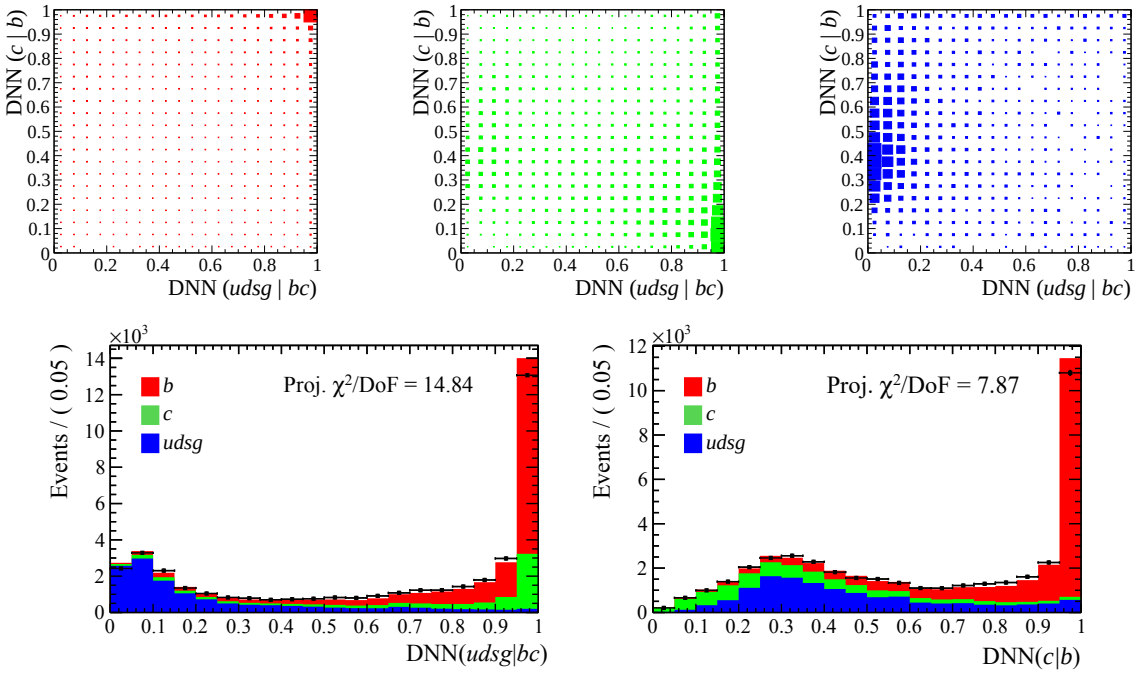



Fig. 5.6 $(b, c, udsg)$-templates (red, green, blue) for DNN 2D fits to SV-tagged $W$+jet data projected in each fitted axis (2D $\chi^2$/NDF $\sim$ 12.2).

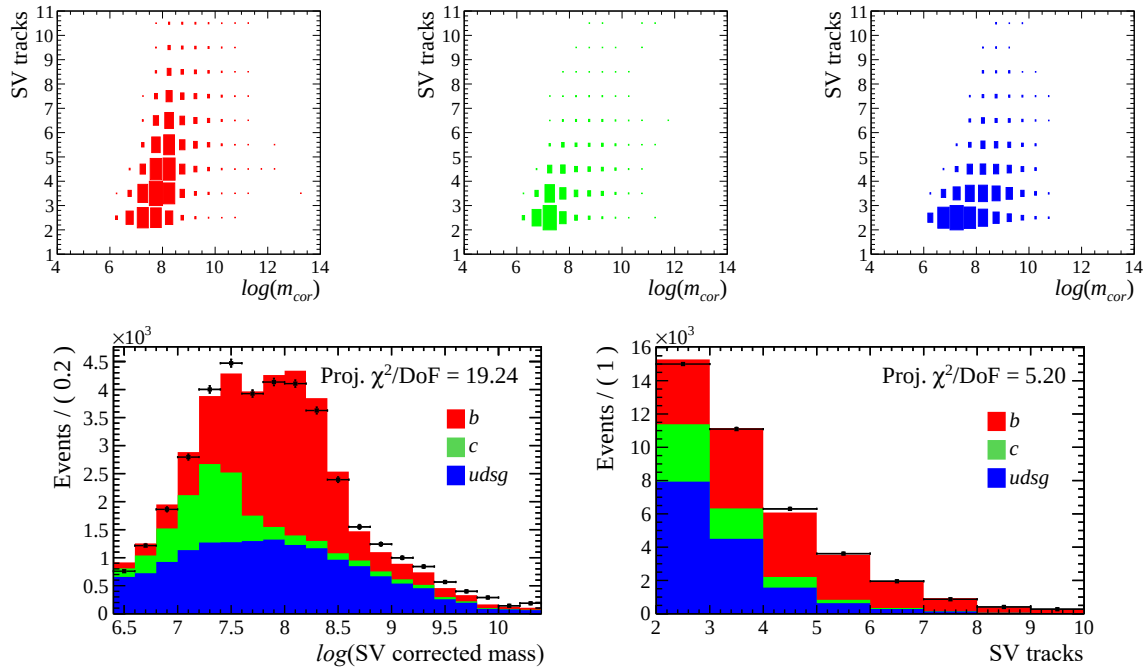



Fig. 5.7 $(b,c,udsg)$-templates (red, green, blue) for alternative 2D fits to SV-tagged $W$+jet data projected in each fitted axis (2D $\chi^2$/NDF $\sim$ 28.7).

### 5.3.2 $(\eta, p_T)$-binned templates

Templates are produced from MC in bins of jet $\eta$ and $p_T$ and are used to fit $W$+SV-jet data under the same binning scheme. Binned fits allow the stability of the procedure and the robustness of the model to be assessed based on the fit performance across the jet acceptance of LHCb. The only projection $\chi^2$/NDF > 10 is for the $(bc,udsg)$-DNN axis across the total acceptance (14.8); the 2D fit which the projections are taken from is also the only one with $\chi^2$/NDF > 10 (12.2). The $\chi^2$/NDF values for the 2D fits are provided in Figure 5.8 alongside those for the alternative fit with an equivalent binning scheme. The alternative 2D fit $\chi^2$/NDF values demonstrate only comparable or higher values than the DNN 2D fits.

Consistency checks between the summed yields from binned fits versus integrated fits were also performed. The sum of binned fits contain more information and are expected to outperform the integrated fits in terms of resulting $\chi^2$. This is in part due to the reduction in relative statistical uncertainty resulting in the inflated the significance of the MC to data discrepancy. This method may be used in tandem with the $\chi^2$ tests to localise MC to data differences in terms of jet kinematics and enable future attempts at binned re-weighting or smearing to be assessed. The $\chi^2$ values of 2D fits to binned events, and those to events

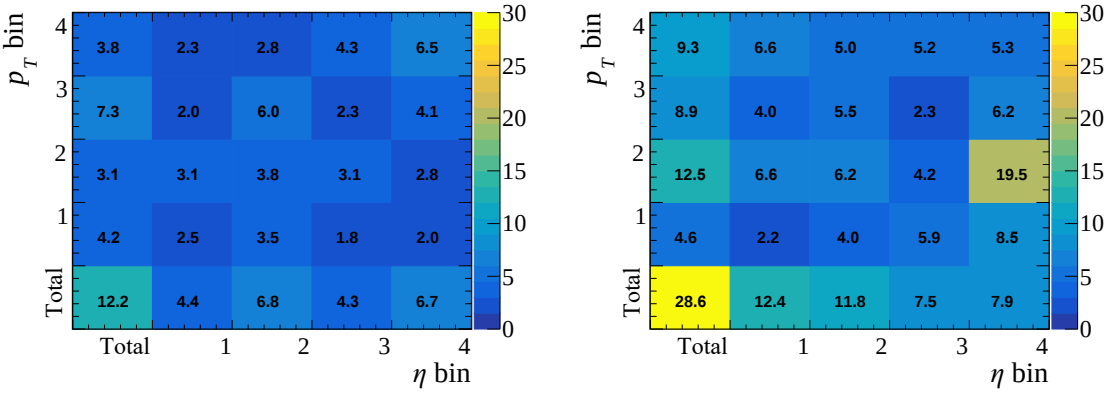



Fig. 5.8 2D fit $\chi^2/NDF$ values using DNN (left) and alternative (right) templates where bin number corresponds to the $\eta$ and $p_T$ boundaries $[2.2, 2.7, 3.2, 3.7, 4.2]$ and $[20, 25, 35, 50, 100]$ GeV respectively and the zeroth bin represents the fit(s) to the total in that axis.

integrated across the full acceptance, are combined with a consistency check of the sum of binned yields against the integrated fit yields to better understand MC to data discrepancy.

Only the $p_T$-binned $c$-yield for jets with $3.7 < \eta < 4.2$ resulted in any inconsistency with integrated fits for the DNN templates. For the alternative fits, binned and integrated yields produce larger deviations and larger uncertainties than the DNN fits. For $b$-yields, they remain consistent but $c$-yields are inconsistent for $\eta$ binning of jets $50 < p_T < 100$ GeV and for $p_T$ binning of jets $3.2 < \eta < 4.2$. The performance of the $b$-yield extraction was deemed sufficient for studies of top production subject to the application of MC shape or template based systematic uncertainties. The last $p_T$ bin ($50 < p_T < 100$ GeV) across the total $\eta$ acceptance from Figure 5.8, corresponding to the jet flavour templates used in the top analysis in Chapter 6, has 2D $\chi^2/NDF$ of 3.8 and projections in each DNN axis as shown by Figure 5.9.

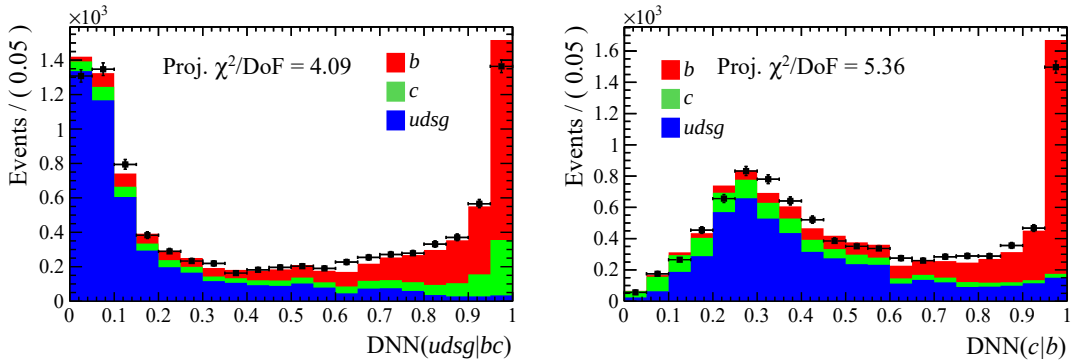



Fig. 5.9 $(b, c, udsg)$-templates (red, green, blue) for DNN 2D fits to $50 < p_T(j) < 100$ GeV SV-tagged $W$+jet data projected in each fitted axis (2D $\chi^2$/NDF $\sim 3.78$).

## Future work

The SV selection and the calibration of the SV-efficiency were optimised from studies of $Z$+$c$ production in Run II. In future, a $b$-decay specific (or more general) calibration may be of benefit. Any future ML techniques applied to SV and jet information will be subject to the selections developed for HLT jets and Run II SVs. Reduced training variable options for binary classifiers, including using SV-variable information only, had also been produced. While SV-only training may be shown to sacrifice overall performance (marginally improved $c$-rejection for $b$-jets in exchange for reduced light rejection), it may be of interest to pursue if a reduced dependence of performance upon jet variables may be demonstrated. Should such reduction be desirable for HF classification studies in the future, the implementation of uniformity boosting techniques [89] may also provide flattened differential performance in terms of jet features.

Future comparisons of HF-tagging models might include multi-classifiers [90], which provide multiple outputs corresponding to the number of classes instead of one output discriminating two classes. During the investigation of the class imbalance strategies (Section 5.2.4), with only the smaller data-set available, concurrent studies implemented a multi-classifier model (instead of dual-binary classifiers). This single model was found to require larger data-sets to learn effectively despite tuning efforts. This used a framework for training, tuning and comparing XGBoost [91] & Keras DNN [86] binary and multi-class models. For further development of alternative HF-tagging models, it is recommended that the larger data-set is put to use testing these approaches as multi-classifiers and the learning algorithms training them may offer further improvements. In addition, a comparison could be made to a topological trigger, initially investigated in Run I but, with potential to be updated using newer boosting techniques (XGBoost [91] & LightGBM [92]) or deep learning. For more expedient hyper-parameter tuning, potentially offering improved model performance, the implementation of a randomised adaptive grid-scan, genetic algorithm or Bayesian optimiser may prove desirable.

As discussed in Section 5.2.3, if producing each binary classifier with different inputs then `ipChi2Sum` & `ptSvrJet` may be removed from the light rejection inputs and `fdChi2Sum` & `pt` from the HF discriminating inputs. While it may be recommended to simplify future analogous binary classifiers, these variables would still be expected to offer discrimination to a multi-classifier model assigned to the same tasks. The potential benefit of automated methods towards feature selection, including principal component analysis or feature importance score [75] based elimination algorithms, have not been ruled out and may be found to be beneficial in future efforts. However, throughout BDT training, the model-dependent feature-importance scores were found to be non-zero for all variables used.

It may be informative to investigate how the model improvement offered by using undersampling for training is impacted by the size of the class imbalance relative to the training samples. Additionally, in Section 5.2.4 it was demonstrated that performing a logarithmic transformation of inputs which are asymmetric random variables may be shown to increase discrimination. For future iterations of the classifiers, it may be beneficial to also perform such a transform of the SV mass (`m`) $p_T$ (`pt`) variables. Caution may be necessary, however, as this may also lead, or already have led, to exaggerated differences between data and MC for the models being trained. These differences may result in poorly fitting MC templates to the model response in data. Initial tests of 2D alternative template fits using un-transformed corrected mass versus SV tracks templates imply this does not improve fit stability or agreement between the two yields. Discussions of smearing input variables to more closely match data, or adapting the handling of uncertainties on the templates used in fits are found in Chapter 6 when considering future work on reducing yield fit systematics.

LHCb GitLab

# Chapter 6

# Top quark cross-section measurements

This chapter presents a measurement of the forward top cross-section and charge asymmetry in the $\mu$+$b$-jet final state using the full Run II data set, corresponding to $5.40\,\text{fb}^{-1}$ collected at centre-of-mass energy of 13 TeV. It includes: the motivation for studying top production; an overview of accessible channels in Run II; a summary of past results from LHCb; details of the selection and background determination; the procedure for the cross-section and asymmetry measurements and their respective systematics; the final results compared to theoretical predictions; and a conclusion with suggestions for further work.

## 6.1 Motivation

As of the conclusion of Run II, top physics in the forward region has been enabled exclusively by the LHCb detector. Chapter 2 provides details of the particle identification, forward tracking and low pile-up environment at LHCb, which, despite a relatively small acceptance, allows for forward top production measurements. The VELO provides vertex reconstruction capabilities of particular benefit to heavy flavour tagging in jets. Developments in machine learning methods for $b$-tagging were explored in Chapter 5 which benefited from a low $b$-jet SV-mistag rate provided using the Run II selection outlined in Chapter 3.

The unique coverage of the detector extends the physics scope of the LHC to new kinematic regions of high energy particle interactions. As discussed in Chapter 1, LHCb makes measurements of production at high-$x$ and low-$x$, which are used to constrain PDFs [93]. The top pair cross-section depends upon the gluon PDF, particularly at high-$x$ (Figure 6.1). Due to their high mass, top production is sensitive to high-$x$ required for their production, which is where the $g$-PDF itself is poorly constrained [94]. Consequently, forward top production measurements could constrain the $g$-PDF by $\mathcal{O}(20\%)$ [95].

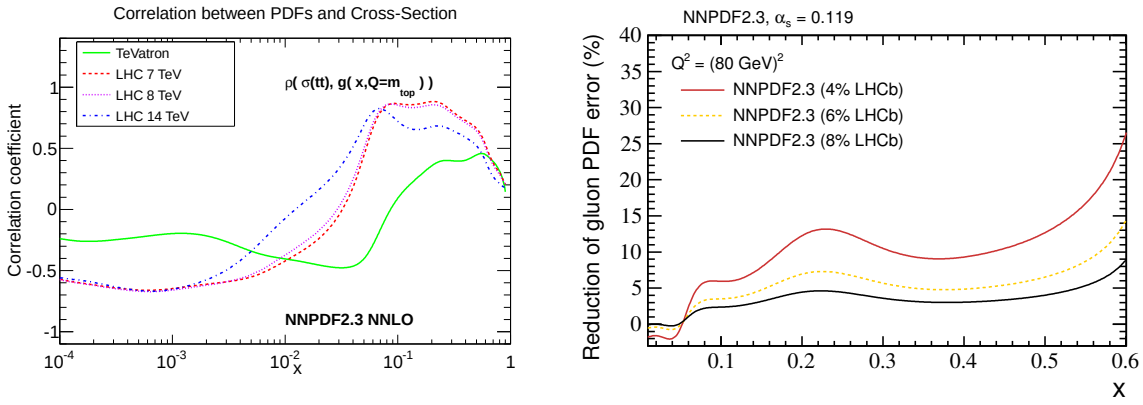



Fig. 6.1 Total NNLO+NNLL top quark production cross-section dependence on the gluon-PDF, driven by $gg$ density, at the LHC for different centre of mass energies [94] (left) and the potential constraint on gluon PDF for NNPDF2.3 with the inclusion of an LHCb semi-inclusive measurement with an associated uncertainty of 4-8% [95] (right).

Besides measuring top production, the opportunity for an LHC $t\bar{t}$ charge asymmetry first observation arises from Run II onward with $\mathcal{O}(1\%)$ sensitivity expected by Run III [96]. The SM predicts that interference effects at NLO produce a positive asymmetry in top pairs from quark-initiated production; new physics could enhance this asymmetry [97]. As shown in Figure 6.2, pair production occurs predominantly through gluon fusion, which is a charge symmetric process. Figure 6.2 also shows the quark initiated pair production contribution, which is charge symmetric at LO. The single-$t$ has a relatively large LO asymmetry, producing a significant positive offset to the combined top asymmetry.

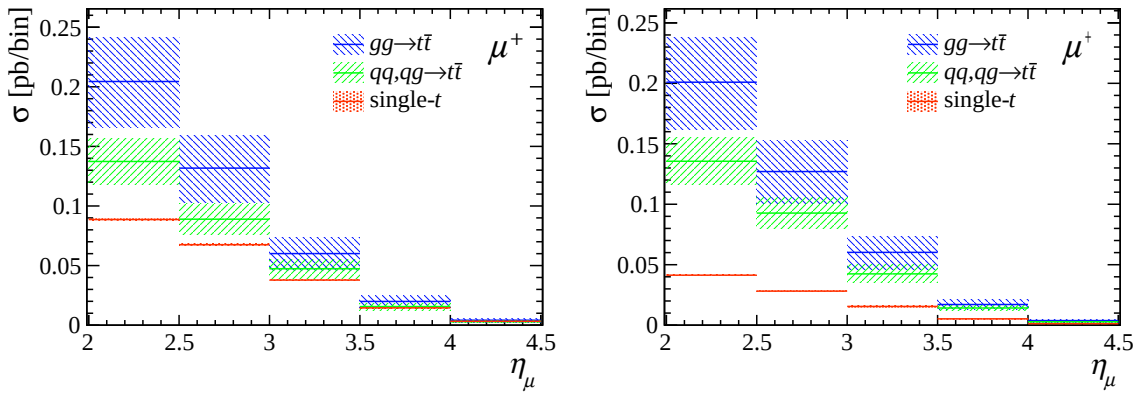



Fig. 6.2 NLO top production cross-sections calculated using POWHEG for 13 TeV in the $\eta_\mu$ acceptance of LHCb, for final state $\mu^+$ (left) and $\mu^-$ (right) with the uncertainty accounting for variation of PDF, scale and $\alpha_S$.

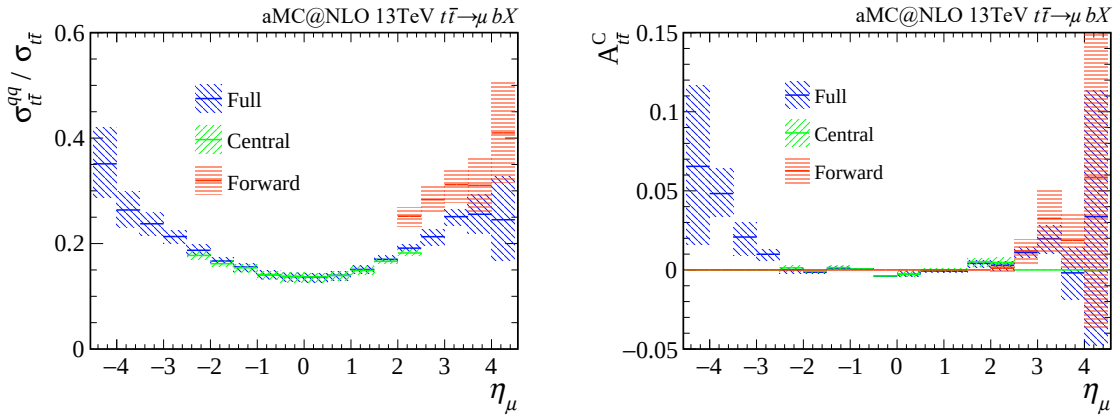



Fig. 6.3 Generator level $t\bar{t}$ production ratio between quark initiated and gluon fusion (left) and top charge asymmetry (right) at 13 TeV as a function of $\eta_\mu$ [41, 98]. $b$-jet acceptance is limited to $\eta_b < |5|$ for Full, $\eta_b < |2.5|$ for Central [99] and $2.2 < \eta_b < 4.2$ for Forward [76]. Values were calculated using Equation 1.38 with the uncertainty accounting for variation of PDF, scale and $\alpha_S$.

The dilution of the $t\bar{t}$ asymmetry by the dominant gluon fusion process reduces in the forward region, resulting in a larger asymmetry with respect to the central region [29]. The fraction of pair production for which gluon fusion is not responsible and its impact on the magnitude of $A_C^{t\bar{t}}$ is shown in Figure 6.3. Contributions from LHCb will become increasingly valuable in setting indirect constraints on new physics using top data from Runs III & IV [100]; reinterpretation of future results may provide model-independent constraints through complementary contributions to global fits [101]. Thus far, $A_C^{t\bar{t}}$ measurements at the LHC have been inconclusive, with preliminary $4\sigma$ evidence from ATLAS only (Figure 6.4).

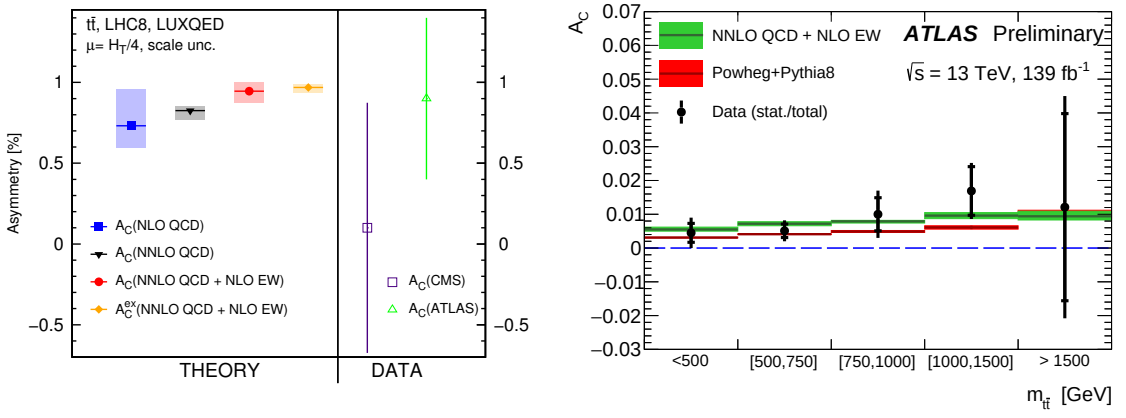



Fig. 6.4 Inclusive charge asymmetry, $A_C$, for the LHC at 8 TeV in NLO QCD, NNLO QCD and NNLO QCD + NLO EW versus CMS and ATLAS measurements [102] (left) and the preliminary differential charge asymmetry measurement as a function of the longitudinal boost of the top pair system in Run II data [103] (right).

## 6.2 Decay channels

Top quarks decay into a $W$ boson and a $b$ quark before hadronising; the $W$ will decay to a $q\bar{q}'$ (branching fraction, $BF = 68\%$) or $l\nu_l$ ($BF = 32\%$) pair [1]. Leptonic decays provide favourable modes for LHCb to trigger upon, where the sign of the $W$ boson corresponds to the charge of the top quark. Though LHCb cannot access missing transverse energy for full kinematic reconstruction of leptonic decays, the top yield is increased through partial reconstruction of each $l\nu b$ and by requiring only a subset of the $t\bar{t}$ final state in the detector. Selecting events with a high $p_T$ $b$-jet accesses relatively high statistics through $lbX$ and suppresses backgrounds for $llbX$. In the $lb$ channel, events produced through single-$t$ and $t\bar{t}$ cannot be separated, so a combined measurement of single and pair produced top quarks is performed instead.

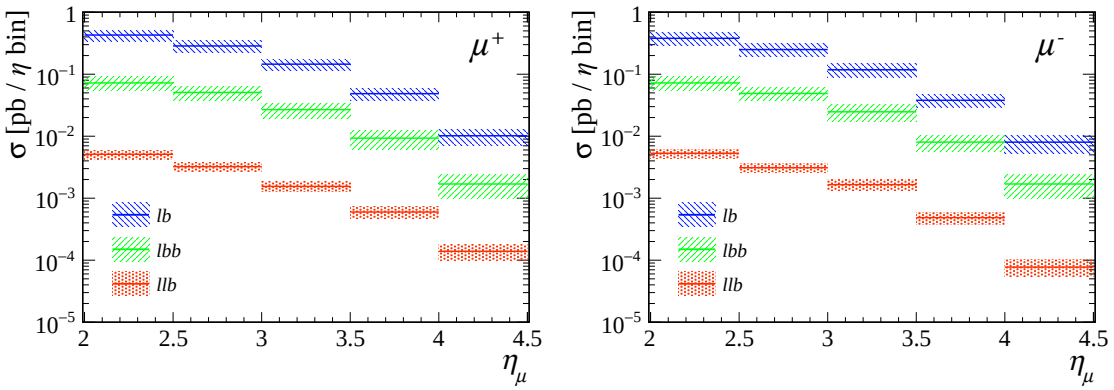



Fig. 6.5 Inclusive POWHEG predictions for partially reconstructed top channels ($\mu b$, $\mu bb$ & $\mu eb$) at 13 TeV in the LHCb $\eta_\mu$ acceptance, where $llbb$ is negligible (uncertainty from PDF, scale and $\alpha_S$ variation).

Muons rather than electrons are used in single lepton channels due to their more reliable ID, greater reconstruction efficiency and more precise energy resolution (Chapter 2). For dilepton channels, an opposite sign $\mu e$ pair are required, recovering twice the branching fraction of $\mu\mu$ while suppressing Drell-Yan backgrounds [37]. The $lb$ channel in Figure 6.5 is inclusive of $lb$, $lbj$, $lbb$ and $llb$, the latter two of which are plotted for comparison and where $llbb$ is expected to be negligible. These predictions for the $\mu bX$ channel combine $t\bar{t}$ channels with an irreducible single-$t$ ($\rightarrow \mu b$) background.

Compared to the more abundant $lb$ channel, the $lbb$ channel compounds the $b$-jet SV-tag efficiency and associated systematic uncertainties. While the $llb$ channel has proven to provide a pure supply of $t\bar{t}$ events for Run II analysis [37], the larger $lb$ channel yield brings differential measurements of the top cross-section within reach. The dominant background

for the reconstruction of one top quark is the $Wb$ contribution, shown in Figure 6.6 alongside the single-$t$ and $t\bar{t}$ expectations for each muon sign. The SM predicts that $\sim 80\%$ of $t \to Wb$ decays in the forward region are due to top pair production. The charge asymmetric t-channel dominates single-$t$ production (Figure 6.7) with the s-channel and $W$ associated single-$t$ contributing at the percent and sub-percent level respectively, consistent with the fractional composition of Run I centre-of-mass energy collisions [31].

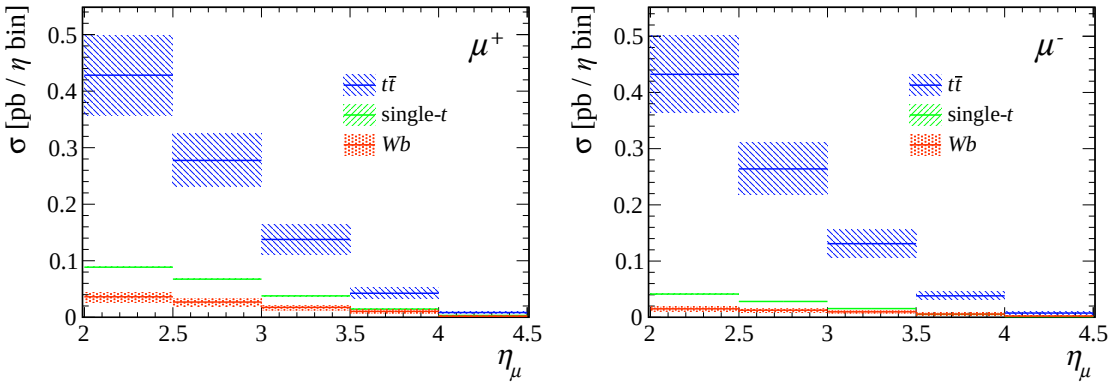



Fig. 6.6 Leading contributions to top candidate events at 13 TeV in the LHCb $\eta_\mu$ acceptance, for final state $\mu^\pm$ (uncertainty from PDF, scale and $\alpha_S$ variation).

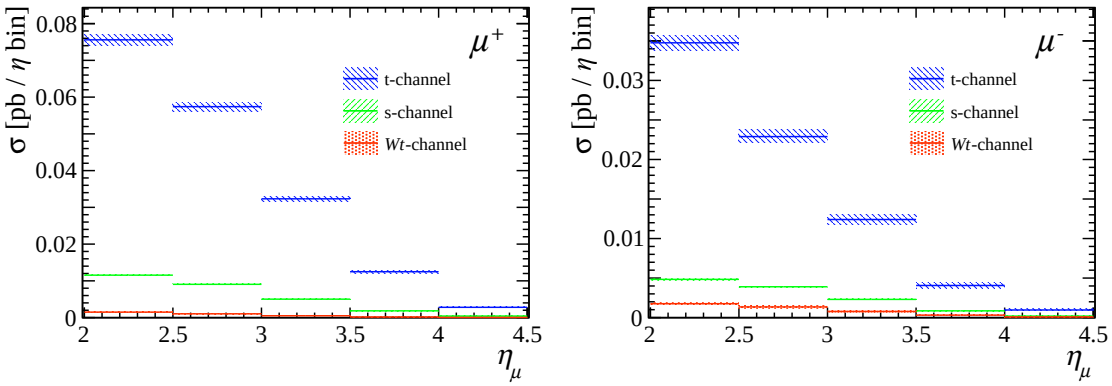



Fig. 6.7 Single-top production cross-sections calculated using POWHEG at 13 TeV in the LHCb $\eta_\mu$ acceptance, for final state $\mu^\pm$ (uncertainty from PDF, scale and $\alpha_S$ variation).

At the cost of disentangling the top charge asymmetry from the single-$t$ contribution [1], the relative abundance of $lb$ events provides differential measurements of the asymmetry over kinematic regions with enhanced sensitivity, such as high $\eta$. Figure 6.8 demonstrates the relative sizes and impact on the overall value of the $t\bar{t}$ and single-$t$ asymmetries and how

[1] Dominated by the net positive charge of the $pp$ collision environment.

predictions for final states excluding single-$t$ background compare. Though sub-dominant, the single-$t$ has a significant asymmetry, complicating the extraction of the $t\bar{t}$ component or distinguishing it from zero. Statistical limitations in the available NLO samples result in noise, observed particularly in the $llb$ asymmetry prediction. A purely $t\bar{t}$ asymmetry may yet be resolved in data in the $lbb$ or $llb$ channels; Sections 6.5-6.7 concern new measurements in the $lb$ channel.

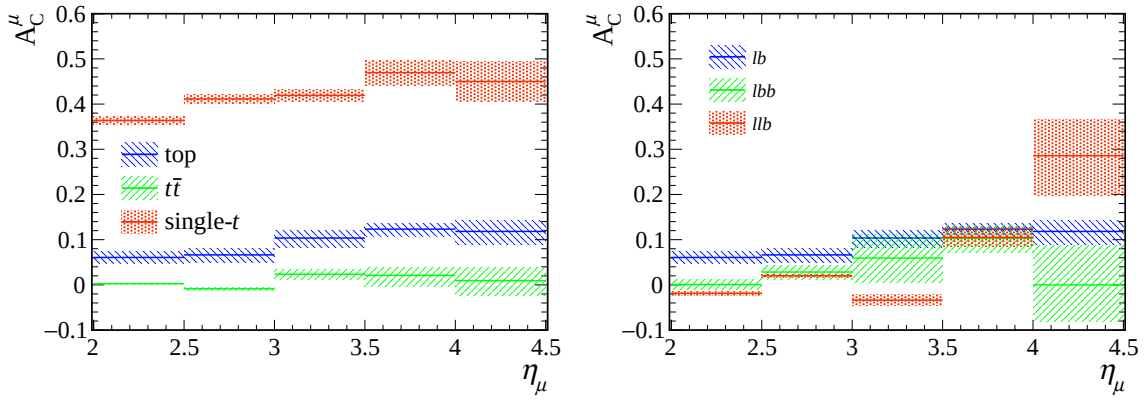



Fig. 6.8 POWHEG predictions for components of $A_C$ in the $\mu b$ final state (left) at 13 TeV in the LHCb $\eta_\mu$ acceptance with partially reconstructed top channel ($\mu b$, $\mu bb$ & $\mu eb$) $A_C$ (right) where $lb$ on the right includes single-$t$ (uncertainty from PDF, scale and $\alpha_S$ variation).

## 6.3 Past results from LHCb

The following results on forward region top production, measured with the LHCb detector, are from the: $\mu b$ channel using 7 & 8 TeV data; $\mu bb$ channel using 8 TeV data; and $\mu eb$ channel using 13 TeV data from 2015 & 2016. In principle, each of these measurements could be improved upon using the full Run II data set [104], as in the $\mu eb$ update [37].

### 6.3.1 First observation, $\mu + b$ final state

Measurements of $W$+jet events are performed using data corresponding to an integrated luminosity of 1.0 and 2.0 fb$^{-1}$ collected in Run I at 7 and 8 TeV. Chapter 6 discusses an analogous kinematic and fiducial selection in more detail. A data-driven template for the QCD contribution is taken from a control region of events with a final state balanced in the transverse plane. A profile likelihood fit to a variable describing the relative isolation of the final state muon was performed to extract EW yields. Having subtracted simulated $\sigma(Wb)/\sigma(Wj)$ normalised to data, the remaining events provide the top cross-section.

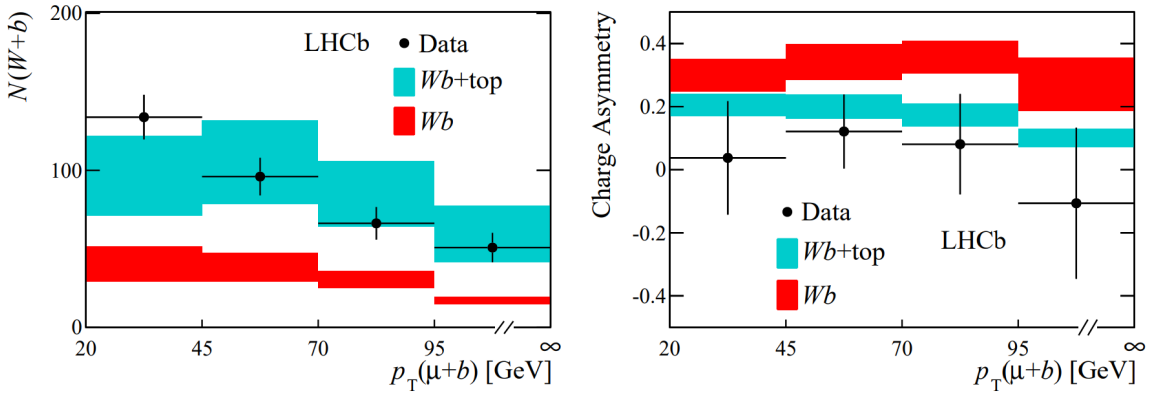



Fig. 6.9 Results for the $W$+$b$ yield (left) and charge asymmetry (right) versus $p_T(\mu+b)$ compared to SM predictions obtained at NLO using MCFM for $L_{int} = 1.0, 2.0\,\text{fb}^{-1}$ at $7, 8\,\text{TeV}$ [31].

The resulting inclusive top production cross-sections, observed to $5.4\sigma$, in the fiducial region defined by: $p_T(\mu) > 25\,\text{GeV}$; $2.0 < \eta(\mu) < 4.5$; $50 < p_T(b) < 100\,\text{GeV}$; $2.2 < \eta(b) < 4.2$; $\Delta\text{R}(\mu, b) > 0.5$; and $p_T(\mu{+}b) > 20\,\text{GeV}$ are:

$$\begin{aligned} \sigma(\text{top})[7\,\text{TeV}] &= 239 \pm 53\,(\text{stat}) \pm 33\,(\text{syst}) \pm 24\,(\text{theory})\,\text{fb}\,, \\ \sigma(\text{top})[8\,\text{TeV}] &= 289 \pm 43\,(\text{stat}) \pm 40\,(\text{syst}) \pm 29\,(\text{theory})\,\text{fb}\,. \qquad [31] \end{aligned}$$

The uncertainty is dominated by the $b$-tagging efficiency. These results, including differential yields and charge asymmetries, are in agreement with SM predictions at NLO.

### 6.3.2 Pair production, $l + bb$ final state

A simultaneous four-dimensional fit to 8 TeV data, corresponding to an integrated luminosity of $2\,\text{fb}^{-1}$, provides access to the $t\bar{t}$, $W$+$bb$ and $W$+$cc$ cross-sections. These results are in agreement with SM predictions and imply $t\bar{t}$ observation to $4.9\sigma$ significance. The resulting inclusive $t\bar{t}$ production cross-section in the fiducial region defined by $p_T(\mu, e) > (20, 15)\,\text{GeV}$; $2.0 < \eta(\mu, e) < (4.5, 4.25)$; $12.5 < p_T(j) < 100\,\text{GeV}$; $2.2 < \eta(j) < 4.2$; $\Delta\text{R}(l, j) > 0.5$; $\Delta\text{R}(j_1, j_2) > 0.5$; $p_T(l + j_1 + j_2) > 15\,\text{GeV}$ is:

$$\sigma(t\bar{t})[8\,\text{TeV}] = 0.05\,^{+0.02}_{-0.01}\,(\text{stat})\,^{+0.02}_{-0.01}\,(\text{syst})\,\text{pb}\,[105].$$

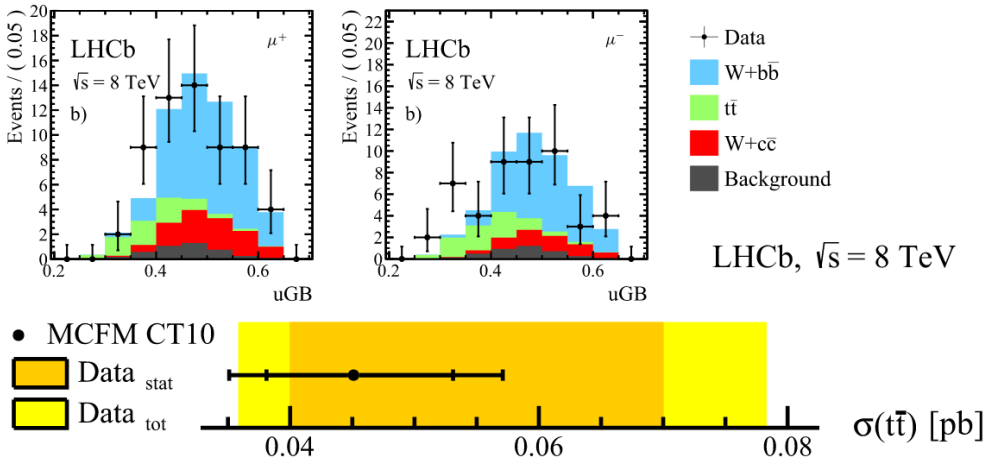



Fig. 6.10 Result of simultaneous 4D fit in terms of uniformity boosted MVA response (uGB) for $\mu^+$ (left) and $\mu^-$ (right); observed and expected cross-sections compared, where inner uncertainties on theory represent the scale errors (bottom) for $L_{int} = 2.0\,\text{fb}^{-1}$ at 8 TeV [105].

### 6.3.3 Run II, $\mu e + b$ final state

The greater centre of mass energy of Run II provides an increased yield of $t\bar{t}$ with up to a factor of 10 gained in the forward region. With just $\sim 2\,\text{fb}^{-1}$ of data, the purest channel, previously inaccessible, becomes viable for analysis. The inclusion of a second lepton suppresses both $Wb$ and multi-jet QCD while the different flavours suppress the $Z$+jet background. A subtraction of expected backgrounds is performed and simulated $t\bar{t}$ is normalised to the remainder of the 44 events in data (Figure 6.11). The same procedure yields 118 events with the full Run II data (Figure 6.12).

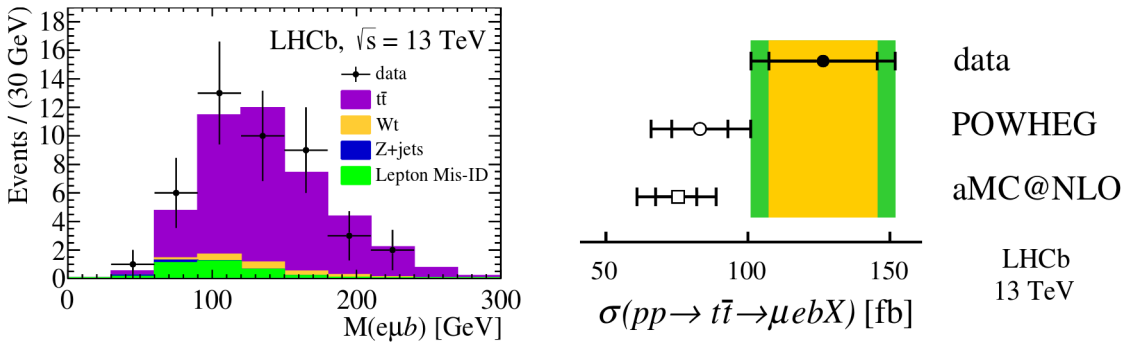



Fig. 6.11 Fit to invariant mass distribution of $\mu eb$ final state (left) with comparison (right) of theory (measured) cross-section, where the inner band represents the scale (statistical) uncertainty for $L_{int} = 2.0\,\text{fb}^{-1}$ at 13 TeV [106].

The resulting top pair production cross-sections in the fiducial region defined by: $p_T(l) > 20\,\text{GeV}$; $2.0 < \eta(\mu,e) < (4.5,4.25)$; $20 < p_T(j) < 100\,\text{GeV}$; $2.2 < \eta(j) < 4.2$; $\Delta\text{R}(l,j) > 0.5$; $\Delta\text{R}(\mu,e) > 0.1$, $\text{IP}_l < 0.04\,\text{mm}$ and $p_T(j_l) > 5\,\text{GeV}$ with 13 TeV data-sets are:

$$\begin{aligned}\sigma(t\bar{t})[13\,\text{TeV}]\ (L_{int} = 1.9\,\text{fb}^{-1}) &= 126 \pm 19\ (\text{stat}) \pm 16\ (\text{syst}) \pm 5\ (\text{lumi})\ \text{fb}\ [106],\\ \sigma(t\bar{t})[13\,\text{TeV}]\ (L_{int} = 5.4\,\text{fb}^{-1}) &= 117 \pm 10\ (\text{stat}) \pm 15\ (\text{syst}) \pm 5\ (\text{lumi})\ \text{fb}\ [37].\end{aligned}$$

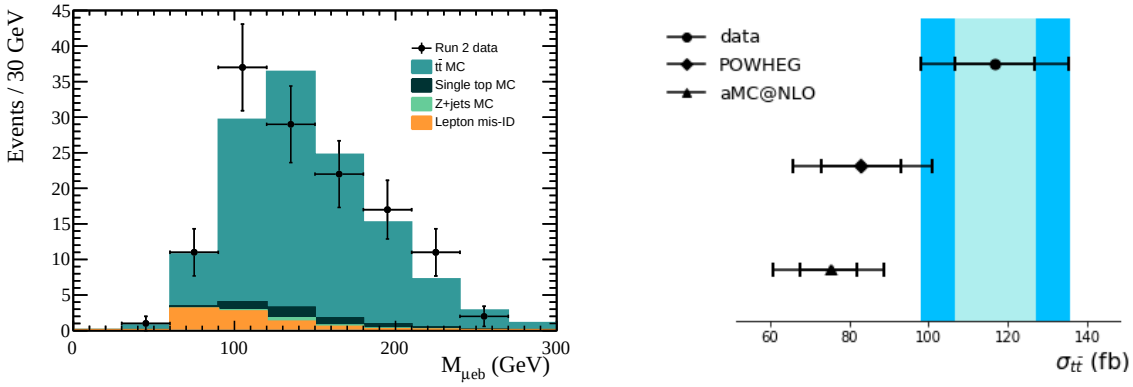



Fig. 6.12 The measured fiducial cross-section, where inner band represents the statistical uncertainty and the outer band represents the total, compared with NLO theoretical predictions from POWHEG and aMC@NLO, where the inner bands represent the scale uncertainty and the outer bands represent the total (right) for $L_{int} = 5.4\,\text{fb}^{-1}$ at 13 TeV [37].

## 6.4 Data samples

Due to no change to the muon trigger or detector level systems between 2016 and 2018, simulated sample production was configured to 2016 running conditions throughout. Fully simulated samples of $t\bar{t}$, $(W \to \mu)$+jet, $(Z \to \mu\mu)$+jet and $Z \to \mu\mu$ were produced at LO for use in the analysis. The $\mu$+jet in the final state are required to be separated from one another by $\Delta R > 0.5$ (maximum jet clustering radius). NLO predictions for $W$+$b$-jet fraction of $W$+jet production are used to normalise the measured $W$+jet yield in data and provide an expected yield of the dominant $Wb$ background [31]. This method is preferable because the SM prediction for $\sigma(Wb)/\sigma(Wj)$ has a smaller relative uncertainty than $\sigma(Wb)$ since theory uncertainties partially cancel. The cross-section for $W \to \tau$ contributions to the $W \to \mu$ channel is corrected for in the final analysis.

As detailed in Chapter 1, the PDF sets used are the NNPDF 3.1 and the LO generator used is Pythia8 which also provides simulated showering. The $t\bar{t}$ samples include corrections in pQCD at NLO using aMC@NLO [107] and POWHEG [108, 109] in order to provide: NLO

re-weighting for efficiency calculations; $k$-factor systematics on jet acceptance corrections; and predictions for the theory comparison to measurement. For the aMC@NLO samples, the decay is performed using MADSPIN [110][19] to account for the influence of $t \to (W \to l)$ spin correlations on final state kinematics.

At the time of this analysis, the full Run II 13 TeV data set collected by LHCb in 2015 through 2018 only had a luminosity calibration available for years 2015 and 2016. As a result, the integrated luminosity ($L_{int}$) for the rest of Run II was calculated using the number of fully reconstructed $Z \to \mu\mu$ events (Table 6.1) to extrapolate subsequent years, where $R$ is the ratio normalised to 2016. The Run II data, therefore, corresponds to an $L_{int} = 5.40\,\text{fb}^{-1}$. This value is considered accurate to within the 4% of 2015 & 2016, assuming no significant changes in muon reconstruction efficiency moving into 2017 and 2018. A 5% systematic was assumed until the new central estimate could be applied.

Table 6.1 Integrated luminosity calibration by year using $Z \to \mu\mu$ events in Run II.

| Year | N($Z \to \mu\mu$) | DV Lumi $\left[\text{pb}^{-1}\right]$ | $R$ | Lumi $\left[\text{pb}^{-1}\right]$ |
|---|---|---|---|---|
| 2015 | 37,930 | 245.35 | 1.001 | 245.57 |
| 2016 | 245,176 | 15587.35 | 1.000 | 1587.35 |
| 2017 | 248,976 | 1002.53 | 1.607 | 1611.954 |
| 2018 | 302,568 | 1215.16 | 1.612 | 1958.926 |
| Total | 834,650 | | | 5403.802 |

## 6.5 Event selection and backgrounds

Events passing the trigger requirements for EW production (Table 6.2) undergo selection for one high-$p_T$ muon produced in association with a high-$p_T$ jet. The muon track is required to lie in the pseudorapidity range $2.0 < \eta < 4.5$. As discussed in Chapter 3, the muons are identified through hits in the four outermost muon stations, depositing more than 50 MeV in the pre-shower (PR), more than 10% of its energy in the ECAL and less than 5% of its energy in the HCAL. To reduce contamination from pile-up events, only tracks associated with the same PV as the final state muon are clustered into a jet. The jets are reconstructed using the HLT particle flow algorithm and clustered using the anti-$k_T$ algorithm using R=0.5 and implemented with `Fastjet` [67], as detailed in Chapter 4. The muon and jet must be separated by $\Delta R > 0.5$. The full event selection is detailed in Table 6.3.

The jet acceptance is trimmed to $2.2 < \eta < 4.2$ to reduce the $\eta$ dependence of jet reconstruction efficiency loss due to incomplete detector geometry, as discussed in Chapter 4.

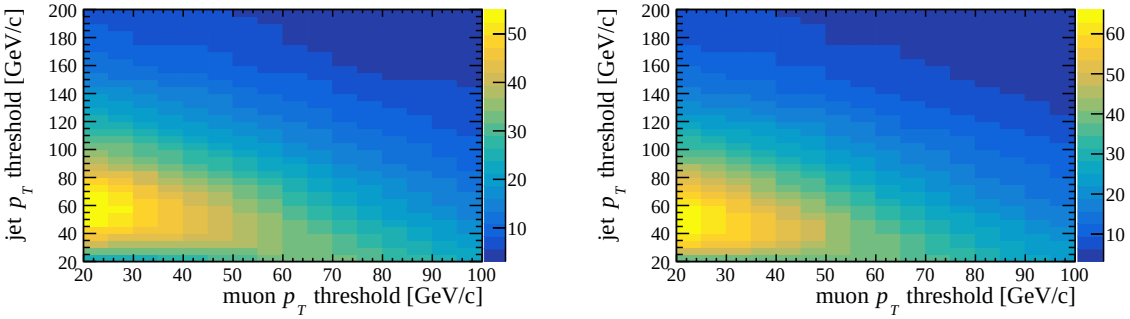
Reconstructing SVs with a direction of flight within the $(\eta, \phi)$-cone of jets within the defined acceptance provides the SV-tagged events. Considering the SV-tagged jet with the highest $p_T$, rather than only the leading jet with a subsequent SV-tag requirement, provides a 5-10% boost in statistics. The same requirements are applied to theoretical predictions[2].

Table 6.2 Trigger on signal (TOS) decision requirements for the $W \to \mu$ events.

| Trigger line | Selection |
|---|---|
| L0MuonEW | nSPD < 10000<br>$p_T > 6.0$ GeV |
| HLT1SingleMuonHighPT | $p > 8.0$ GeV<br>$p_T > 6.0$ GeV<br>Track $\chi^2 < 4.0$ |
| Hlt2SingleMuonHighPT | IP $< 0.25$ mm<br>IP $\chi^2 < 100$<br>$p_T > 15.0$ GeV |

Table 6.3 Selection requirements on top candidate ($W \to \mu$)+jet events.

| Object | Selection |
|---|---|
| Muon | $2.0 < \eta < 4.5$<br>$p_T > 25$ GeV<br>$IP < 0.04$ mm<br>$E/p < 0.04$<br>$\sigma(p)^2/p^2 < 0.01$<br>ProbNNmu $> 0.98$<br>$Z \to \mu\mu$ veto |
| Jet | $2.2 < \eta < 4.2$<br>$p_T > 50$ GeV<br>$\Delta R(\mu, j) > 0.5$<br>SV-tagged |




(a) $\mu^+ + b$-jet $(N_{t\bar{t}}/\sqrt{N_{t\bar{t}} + N_{Wb}})$ (5% $Wb$ syst.) (b) $\mu^- + b$-jet $N_{t\bar{t}}/\sqrt{N_{t\bar{t}} + N_{Wb}}$ (5% $Wb$ syst.)

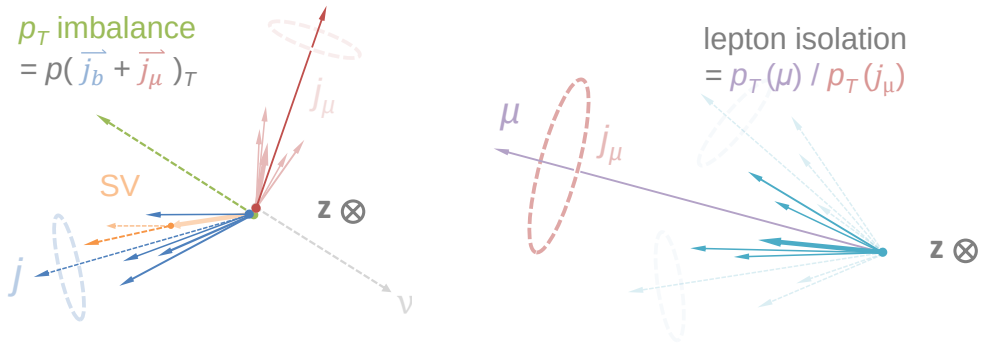
Fig. 6.13 Top significance over $Wb$ for final state kinematic thresholds using an assumed $Wb$-background yield systematic of 5% applied to 13 TeV generator level samples normalised to 5.4 fb$^{-1}$.

The $p_T$ requirements, replicated from the Run I analysis [31], use muon and jet $p_T$ thresholds to limit multi-jet QCD and $Wb$ backgrounds respectively. Figure 6.13 demonstrates the impact of minimum $p_T$ upon generator level significance of top versus $Wb$ events

[2]Some statistics are recovered in $t\bar{t}$ events where either $b$-jet may contribute to the final state.

normalised to $5.4\,\text{fb}^{-1}$ in the $\mu$+$b$ final state. With minimum $p_T$ cuts applied, the fraction of $(W \to \mu)$+$b$ events from top quark decays in the LHCb acceptance is $\sim 90\%$.

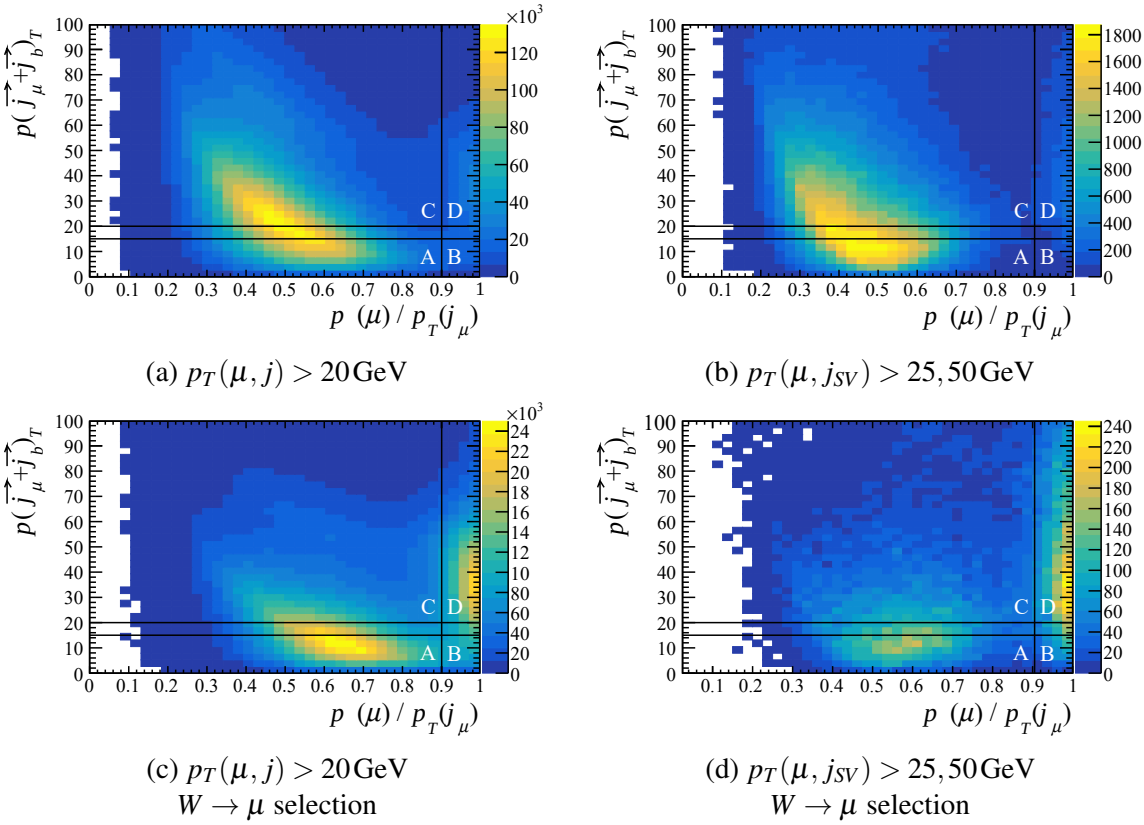
The leptonic decay of a $W$ in association with a jet is expected to produce missing transverse energy compared to multi-jet QCD processes. Due to the lack of $4\pi$ detector coverage, the transverse component of the vector sum of a muon and jet momenta provides a proxy for missing transverse energy carried by the neutrino escaping the detector. A jet-object (clustered with anti-$k_T$ but not subject to JetID) reconstructed around the muon, $j_\mu$, allows modification of the fiducial requirement, $p(\vec{j_b} + \vec{j_\mu})_T$, where $\mu$ has been replaced with $j_\mu$. This further suppresses the background from di-jets which tend to be balanced in the transverse plane (Figure 6.14). On this basis, the final state fiducial acceptance also requires $p(\vec{j_b} + \vec{j_\mu})_T > 20\,\text{GeV}$ [31]. The $j_\mu$ object also provides an estimate of lepton isolation, defined as the fraction of the $p_T(j_\mu)$ carried by the muon.




(a) final state imbalance $= p(\tilde{j_b} + \tilde{j_\mu})_T$ (b) muon isolation $= p_T(\mu)/p_T(j_\mu)$

Fig. 6.14 Diagrams illustrating the definitions of $\mu$-isolation and $p_T$-imbalance in the transverse plane.

Backgrounds from semi-leptonic decays of heavy-flavour hadrons are suppressed by requiring an IP of the muon track with respect to the PV $< 0.04\,\text{mm}$ [93]. Further requirements rejecting hadron mis-ID, replicated from the Run I analysis, include a maximum calorimeter deposition energy fraction $(E/p)$ of 4%, `ProbNNmu` $> 0.98$ and maximum curvature error $(\sigma(p)^2/p^2)$ of 1% applied to the muon [31]. The $Z$+jet background is suppressed with a veto on opposite sign high-$p_T$ muons with $M(\mu\mu) > 40\,\text{GeV}$. These criteria are summarised in Table 6.3 and each is associated with an efficiency estimate outlined in Section 6.7.1. Figure 6.15 compares the impact of these requirements to the fiducial $p_T$ thresholds where the low isolation, low $p_T$ imbalance peak is expected to be background. In contrast, the highly isolated region of events is expected to be made up of EW processes. The full event selection is shown to result in the relative suppression of the background dominated regions.




(a) $p_T(\mu, j) > 20\,\text{GeV}$

(b) $p_T(\mu, j_{SV}) > 25, 50\,\text{GeV}$

(c) $p_T(\mu, j) > 20\,\text{GeV}$
$W \to \mu$ selection

(d) $p_T(\mu, j_{SV}) > 25, 50\,\text{GeV}$
$W \to \mu$ selection

Fig. 6.15 Muon triggered jet events from Run II data in muon isolation versus final state $p_T$ imbalance, where (a) & (b) include the $\mu$+jet and $\mu + b$jet fiducial selections respectively, (c) & (d) represent the same data with EW production selection placed upon the muon, and region boundaries for the signal region and ABCD data driven background subtraction definitions overlaid.

## 6.6 Analysis strategy

The method for extracting of the top yields from data is as follows, where MC samples have a correction applied in bins of muon isolation from the ratio of $Z \to \mu\mu$ in data and MC:

- Count $\mu$+jet events in signal and control regions split by muon charge and in bins of muon pseudorapidity, with and without a SV-tag;
- Extract $b$-yields with 2D flavour MVA template fits to SV-tagged events;
- Subtract a data-driven multi-jet QCD background, estimated from the control regions;
- Subtract $Z$+jet expectations, taken from $(Z \to \mu)$+jet / $(Z \to \mu\mu)$+jet in MC normalised to $(Z \to \mu\mu)$+jet in data;

- Estimate expected $Wb$ background using NLO ratio of $W$+$b$-jet to $W$+jet production from theory (and corrected with SV-tag efficiency) to normalise the $(W \to \mu)$+jet yield (having subtracted QCD & $Z$ backgrounds);
- Following the subtraction of $Wb$ background estimate from the $W$+$b$-jet yield, only top quark decays remain, with their sign indicated by the muon .

### 6.6.1 Side-band counting

Due to the uncertainty surrounding the modelling of multi-jet (non-EW) backgrounds, this analysis utilises a side-band counting technique, coined ABCD, to correct for background contributions extrapolated from a control region [111, 112]. The background, expected to be non-negligible even after tight selection requirements, is estimated by data-driven means. The region boundaries illustrated in Figure 6.15 are defined as follows:

- A - anti-isolated control region $p(\vec{j_b} + \vec{j_\mu})_T < 15\,\text{GeV}$ & $p_T(\mu)/p_T(j_\mu) < 0.9$
- B - control region $p(\vec{j_b} + \vec{j_\mu})_T < 15\,\text{GeV}$ & $p_T(\mu)/p_T(j_\mu) > 0.9$
- C - anti-isolated region $p(\vec{j_b} + \vec{j_\mu})_T > 20\,\text{GeV}$ & $p_T(\mu)/p_T(j_\mu) < 0.9$
- D - signal region $p(\vec{j_b} + \vec{j_\mu})_T > 20\,\text{GeV}$ & $p_T(\mu)/p_T(j_\mu) > 0.9$

Equation 6.1 shows the equation used for splitting the EW and QCD contributions using the side-band technique. The coefficients $c_{A,B,C}$, the expected EW content relative to the D-region, are taken from ratios in MC and normalised to the signal content in $N_D^{\text{sig}}$, accounting for the EW process contamination in the A,B,C side-bands [111]. For a small number of cases where the statistics in certain regions of MC and data samples are limited, the convergence fails. A culprit may be that the analytical solutions, of which there are usually one positive and one negative, break the $N_D^{\text{bgd}} \geqslant 0$ restriction. In these cases, the value of $N_D^{\text{bgd}}$ is set to half the statistical uncertainty of $N_D$ with an uncertainty of 100%.

$$N_D^{\text{sig}\prime} = N_D - \left(N_C - c_C N_D^{\text{sig}\prime}\right) \frac{\left(N_B - c_B N_D^{\text{sig}\prime}\right)}{\left(N_A - c_A N_D^{\text{sig}\prime}\right)}, \qquad c_{A,B,C} = N_{A,B,C}^{\text{MC}} / N_D^{\text{MC}} \tag{6.1}$$

This method, applied to obtain the central yields, assumes no correlation between the axes used (Figure 6.15). A re-weighting of the anti-isolated control to the respective signal region is expected to reduce this correlation. The weights, $N_C/N_A$, are determined as a function of the $p_T$ of the muon-jet, $p_T(j_\mu)$. In equation 6.1, describing ABCD, the weighting cancels in the term containing $N_B/N_A$, assuming no correlation.

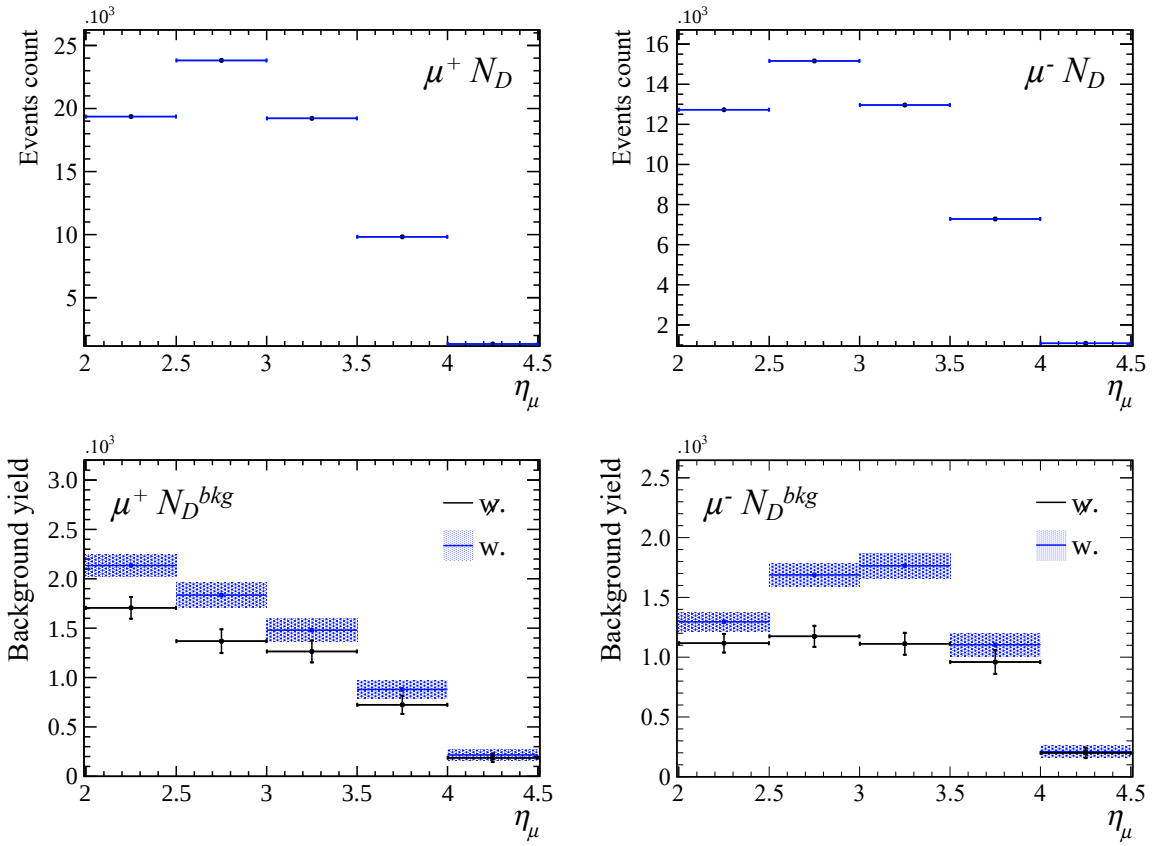
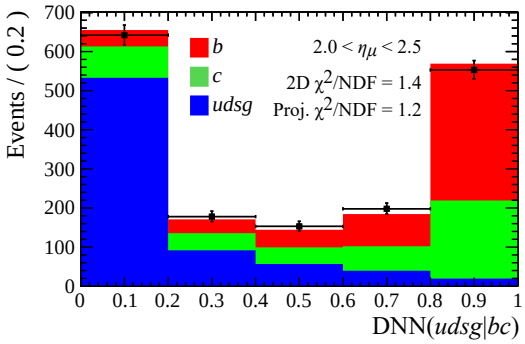


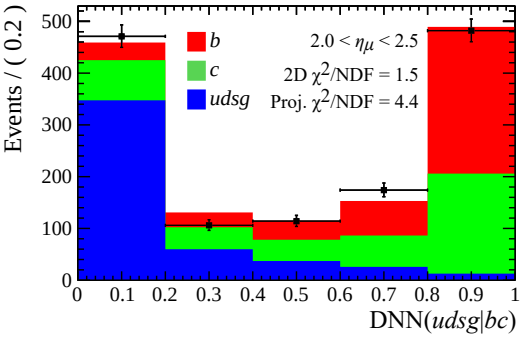

Fig. 6.16 Multi-jet QCD background expectation ($\sim 10\%$) for $\mu$+jet events in isolated $p_T$ imbalanced (D) region calculated with ($w$) and without ($\not{w}$) control region re-weighting.

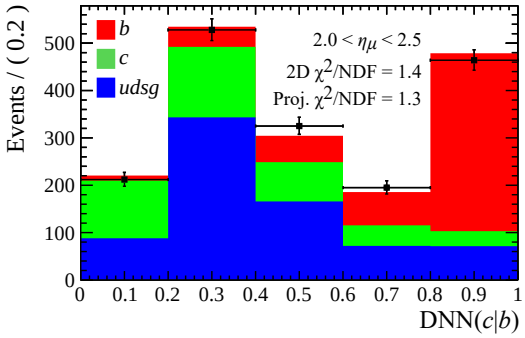
Figure 6.16 shows the data driven estimates of the multi-jet QCD background as a function of $\eta_\mu$ for $\mu^\pm$. These are defined as $N_D^{bkg} = N_D - N_D^{sig}$, as based on measured $\mu$+jet yields ($N_D$) and signal yield ($N_D^{sig}$) calculated with and without control region re-weighting.

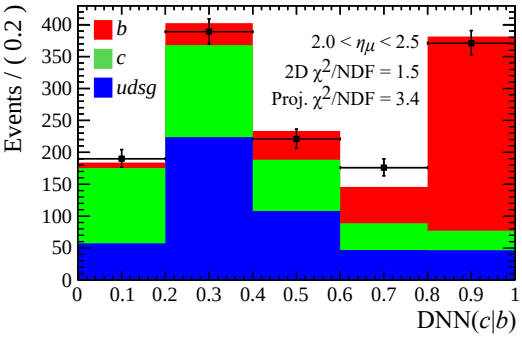
### 6.6.2 Heavy flavour yields

The $b$-, $c$- and light-jet yields for each muon charge and $\eta_\mu$ bin are extracted from the 2D fitting procedure detailed in Chapter 5. SV-tagged jets provide the inputs to the two DNN classifiers. The statistical uncertainty is taken as the error on the yield from the minimiser [113, 114]. This fitting procedure does not consider statistical errors on the templates. The fits are performed upon unweighted data from each region and repeated on re-weighted data from the $p_T$-balanced control regions (A & B). An additional set of yields are determined using the alternative fit (ALT) from Chapter 5, using corrected SV mass versus SV track multiplicity 2D templates to fit to unweighted data from each region.

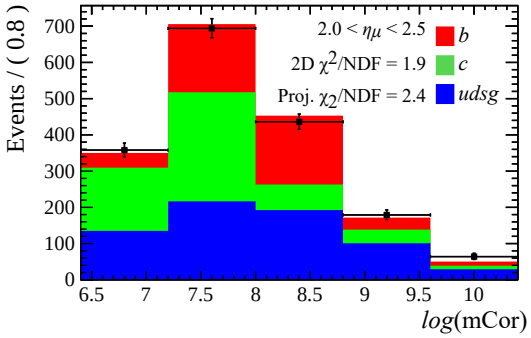


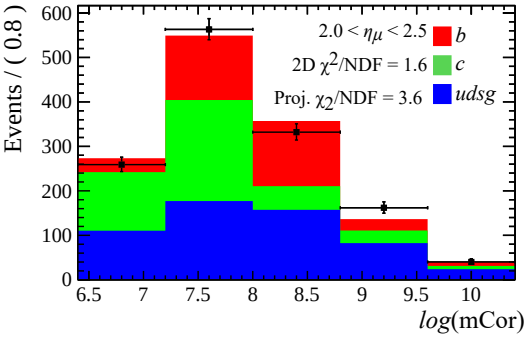



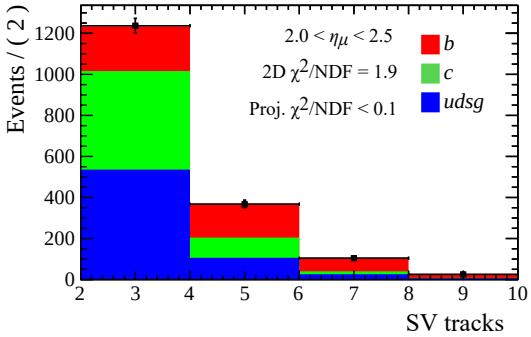



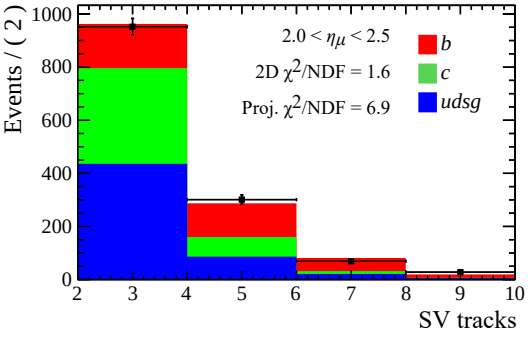




(a) $\mu^+$+SV-jet DNN fit (b) $\mu^-$+SV-jet DNN fit




(c) $\mu^+$+SV-jet ALT fit (d) $\mu^-$+SV-jet ALT fit

Fig. 6.17 Projections of example signal region (D) SV-jet flavour template fits to 2D response space of DNNs and of corrected SV mass and SV tracks for $\mu^{\pm}$ events from the $2.0 < \eta_{\mu} < 2.5$ bin.

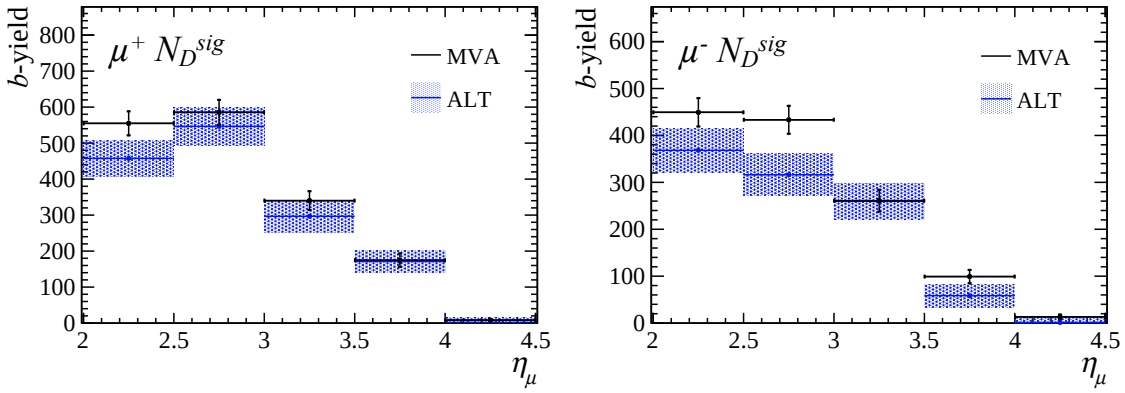



Fig. 6.18 EW+jet $\mu$+$b$-yields in the $p_T$-imbalanced (D) region data for $\mu^\pm$ compared between DNN fit and alternative fit results.

Projections in each DNN axis of 2D fit results are shown in Figure 6.17 for both muon signs. This example is taken from the $\eta$-binned signal region fits shown in Figure 6.18, including systematic uncertainties discussed in Section 6.6.4.

### 6.6.3 Background subtraction

The following steps are performed on the $\mu$+jet and $\mu+b$-jet yields independently:

- Equation 6.1 is solved for the signal yield, $N_D^{sig}$, using the measured yield, $N_D$. These events are assumed to be comprised of $W,Z$+jet contributions.
- An expectation for the residual $Z$+jet background events, $N_D^Z$ is calculated using ($Z \to \mu\mu$)+jet events in data normalised with the ratio of ($Z \to \mu$)+jet and ($Z \to \mu\mu$)+jet from MC in bins of muon isolation for each muon sign[3] This contribution (Figure 6.19) is subtracted leaving only the $W$+jet events, $N_D^W$.

For both the $\mu$+jet and $\mu+b$-jet event yields, the steps above are repeated with a re-weighting applied to the $p_T$-balanced control regions (A & B) used in the $N_D^{sig}$ calculation.

- Each $W$+jet measurement is then used to normalise a NLO $(Wb/Wj)$ ratio from theory. Corrected using the SV-jet reconstruction efficiency, this factor (Figure 6.20) provides the $W$+$b$-jet background expectation. This contribution is deducted from $N_D^W$ leaving only top decays, $N_D^t$.

This final subtraction provides the top and anti-top quark yield central values.

[3]To provide greater statistics, the ($Z \to \mu\mu$)+jet samples used for this correction are not required to have a SV-tag or $b$-jet, where the MC and data ratio is assumed consistent for events passing these criteria.

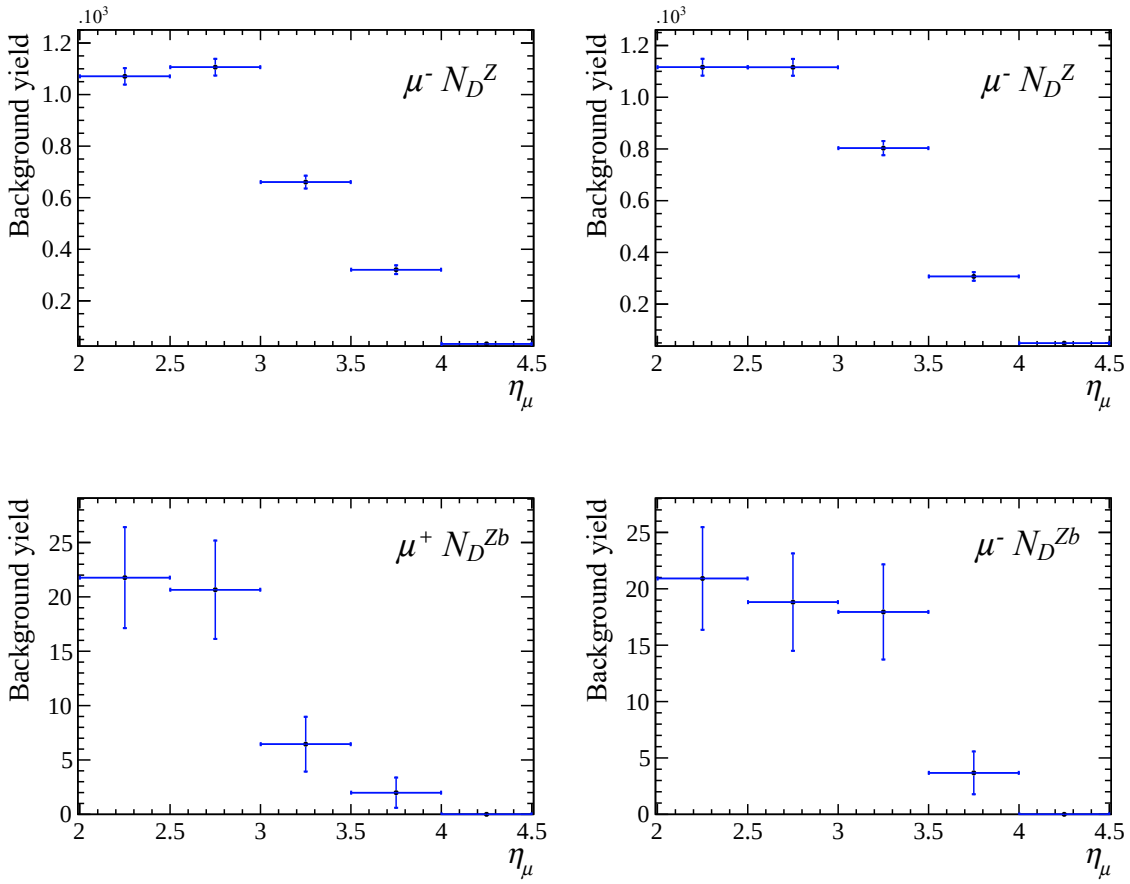



Fig. 6.19 $Z$+jet background expectations taken from MC and normalised to $Z \to \mu\mu$ in data.

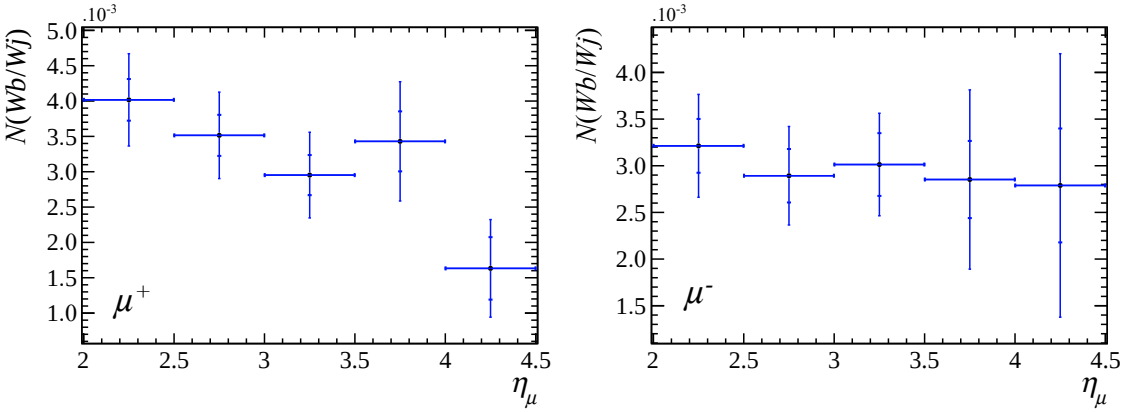



Fig. 6.20 SV-efficiency corrected NLO theory ($Wb/Wj$) normalisation to apply to the measured $W$+jet yield for a $W$+$b$-jet background expectation.

### 6.6.4 Top yield systematic uncertainties

The following procedures were used to determine systematic uncertainties on the top and anti-top yields, through which they are propagated to the cross-section measurement:

- Templates - To account for the limited sample size of the 2D templates, a systematic uncertainty is estimated using a bootstrapping method [115]. The results of 100 replica fits, using templates sampled from the original at random including repeats, provide residuals from the central value from which a Gaussian width and offset may be combined in quadrature into a template systematic uncertainty [116]. This uncertainty is only propagated for the unweighted signal region (central value) fits (Figure 6.18);
- Fits - Performing ALT fits in each ABCD region provides an independent calculation of the signal and background yields simultaneously. The envelope of the 'background-subtracted' yields ($N_D - N_D^{bkg}$) between DNN and ALT fits (Figure 6.18) accounts for correlations between $N_D$ and $N^{A,B,C}$ $b$-yield uncertainties. As a result, this is used as a systematic uncertainty on the HF-yields;
- $N(Wb)$ - The systematic uncertainty on the top yield following the $Wb$ background subtraction, originating from the theory uncertainty on the $Wb/Wj$ ratio and the SV-efficiency uncertainty used to correct the measured $Wj$ used for the normalisation;
- ABCD - The envelope on $N_D^t$, defined as the difference in re-weighted $N_D^{Wj} \cdot N(Wb/Wj)$ subtracted from re-weighted $N_D^{Wb}$ to that of the central unweighted value, accounts for correlations in uncertainty between the $W$+jet and $W$+$b$-jet measurements. This provides the systematic uncertainty on the top yields from the multi-jet QCD background and axes of the ABCD regions.

The DNN and ALT fit signal yields for $\mu$+$b$-jet include the templates uncertainty and provide the fits uncertainty. The SV-efficiency corrected $N(Wb/Wj)$ (Figure 6.20) contains the $N(Wb)$ uncertainty and is multiplied by the $W$+jet $N_D^{sig}$ . The re-weighted variations on the $W$+$b$-jet yields are determined using the ratio of re-weighted to unweighted $W$+jet to alleviate statistical fluctuations in the weights (Figure 6.21). The re-weighted variation in normalised $W$+jet, $N(Wb)$, and the resulting background subtracted top yields from measured $W$+$b$jet minus $N_D^{Wb}$ (Figure 6.22) are shown to be secondary to the fits systematic included on the unweighted central yield.

The ALT yields are also carried through to their own $\mu^{\pm}$ cross-sections and an envelope between DNN and ALT value for $A_C^{\text{top}}$ is taken to account for correlations from the heavy flavour yield fits. This uncertainty was reduced $O(0.01)$ by adjusting the ALT template under- & overflow boundaries and allowing the template normalisations to float below zero.

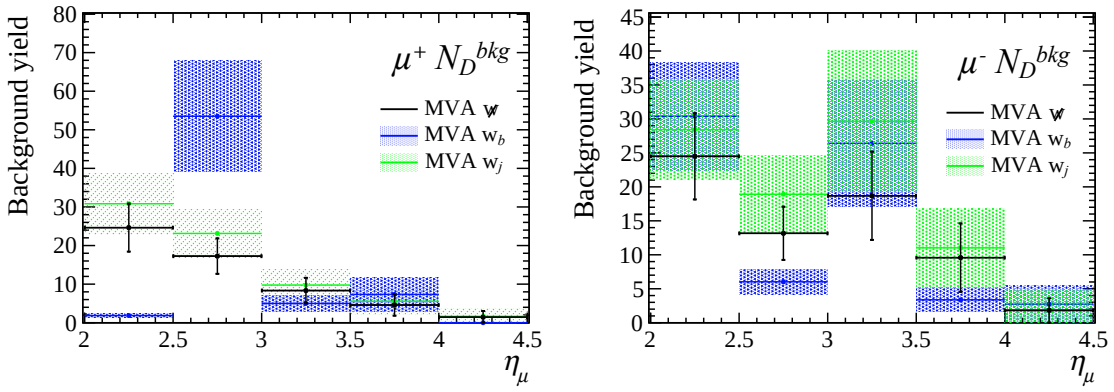



Fig. 6.21 $\mu$+$b$-jet multi-jet QCD background, including a systematic uncertainty from alternative fit variation, in isolated $p_T$ imbalanced (D) region, with re-weighted variations (blue) and unweighted variations normalised with the ratio of re-weighted to unweighted $W$+jet $N_D^{bkg}$ (green).

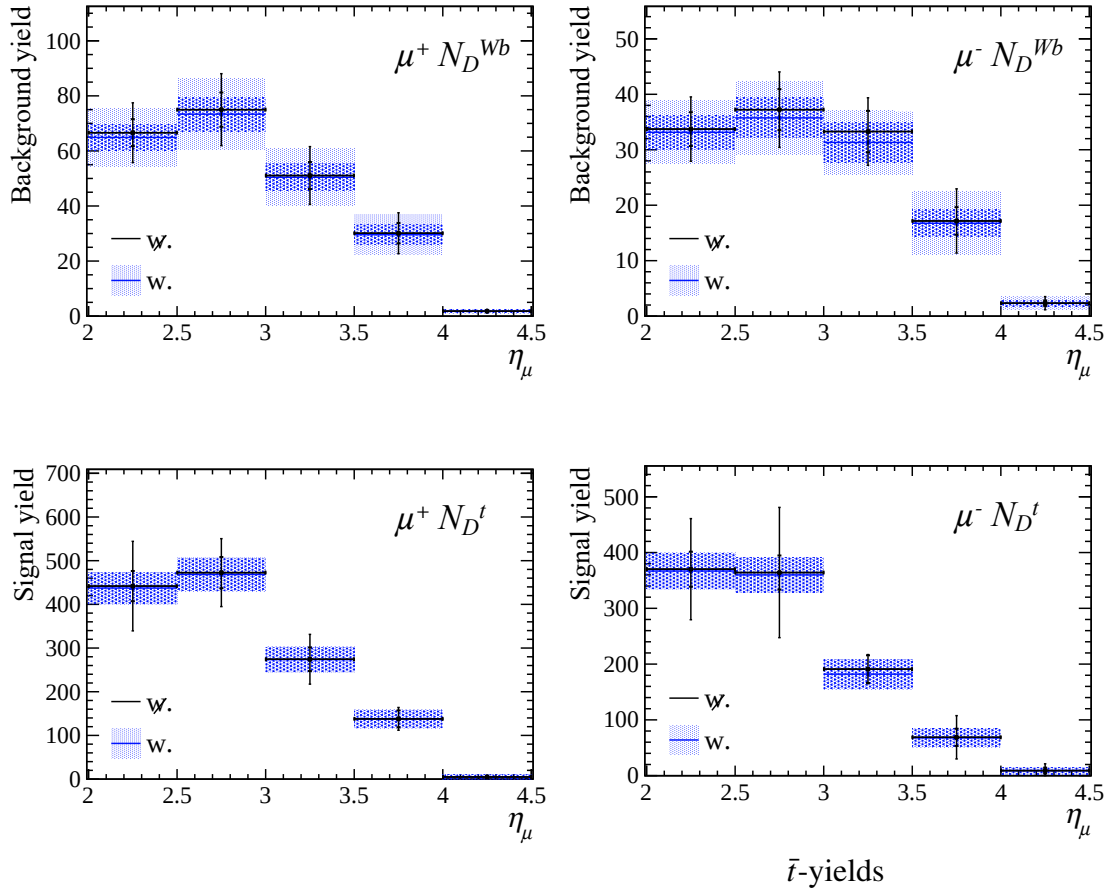



Fig. 6.22 The $Wb$ background expectation and resultant top yields, where the central unweighted values include the variation from the alternative fits as the dominant systematic.

## 6.7 Cross-sections and asymmetry measurements

Equation 6.2 provides the definition for the fiducial top cross-section to be calculated from the measured yields. The term for the signal yield, $(N - N^{\text{bkg}})$, corresponds to the central top yields (Figure 6.23[4]). The signal yield is then corrected for reconstruction and selection efficiencies, given by $\varepsilon_{rec}$ and $\varepsilon_{sel}$ respectively (Section 6.7.1). The total signal yield is then divided by the integrated luminosity, given by $\mathscr{L}$ and detailed in Section 6.4, to provide a preliminary cross-section. The acceptance factor, $\mathscr{A}$, accounts for migrations in and out of the fiducial region (Section 6.7.2). A final theory-based correction is used to remove top decays passing the final state selection with $(W \to \tau)$+jet producing a muon passing selection is included in $\mathscr{A}$. These components are outlined in the following sections.

$$\sigma_{\text{top}}^{fid} = \frac{\left(N - N^{\text{bkg}}\right) \cdot \mathscr{A}}{\mathscr{L} \cdot \varepsilon_{\text{sel}} \cdot \varepsilon_{\text{rec}}} \tag{6.2}$$

The top cross-sections, split by the charge of the final state muon, may be used to calculate the charge asymmetry as per Equation 6.3. The statistical and systematic uncertainties are calculated independently in the second line before being added in quadrature.

$$\begin{aligned} A_C^{\text{top}} &= \left(\frac{\sigma_t - \sigma_{\bar{t}}}{\sigma_t + \sigma_{\bar{t}}}\right)_{\eta(\mu)} = \left(\frac{R_C - 1}{R_C + 1}\right)_{\eta(\mu)}, \quad R_C = \sigma_t/\sigma_{\bar{t}} \\ \delta A_C^{\text{top}} &= 2\sqrt{(\delta\sigma_{\bar{t}})^2(\sigma_t)^2 + (\delta\sigma_t)^2(\sigma_{\bar{t}})^2}/(\sigma_t + \sigma_{\bar{t}})^2 \end{aligned} \tag{6.3}$$

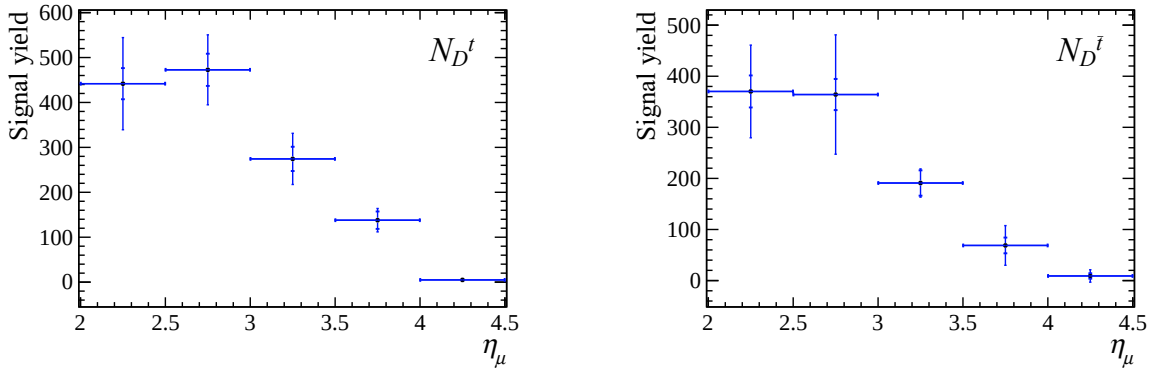



Fig. 6.23 Background subtracted top yields used in the cross-section calculation including systematics from templates, fits, $N(Wb)$ and ABCD.

[4]Including the ABCD systematic variation between the central and weighted control results (Figure 6.22).

### 6.7.1 Efficiencies

Muon reconstruction, trigger and identification efficiencies are determined as a function of $\eta_\mu$ using $Z \to \mu\mu$ data. By exploiting a well known di-muon resonance, in this case the Z mass [93], the combination of final state particles can provide an in-situ tag and probe measurement. One reconstructed muon track, triggering the event, defines a tag while the additional reconstructed object in the $Z \to \mu\mu$ final state defines a probe [117]. The fraction of probes that satisfy the given requirements on the tagged events within the fiducial acceptance is the efficiency. The details of the event selections and chosen tags & probes of each procedure may be found in Table 6.4.

Table 6.4 Tag-and-probe criteria applied to inclusive $Z \to \mu\mu$ data for muon tracking, trigger and identification efficiency measurements for Run II [37].

| Efficiency | Tag | Probe | Event |
|---|---|---|---|
| Tracking | Long track | Muon TT track | $70 < M_{\mu\mu} < 110\,\text{GeV}$ |
| | Triggered | - | $\chi^2_{PV}/\text{NDF} < 5$ |
| | `isMuon` | - | $\lvert\Delta\phi\rvert > 0.1\text{rad}$ |
| | $p_T > 20\,\text{GeV}$ | $p_T > 20\,\text{GeV}$ | - |
| | $p_T^{\text{cone}} < 2\,\text{GeV}$ | $p_T^{\text{cone}} < 2\,\text{GeV}$ | - |
| | $2.0 < \eta < 4.5$ | $2.0 < \eta < 4.5$ | - |
| Trigger | Long track | Long track | $60 < M_{\mu\mu} < 120\,\text{GeV}$ |
| | Triggered | - | $\chi^2_{PV}/\text{NDF} < 5$ |
| | `isMuon` | `isMuon` | - |
| | $p_T > 20\,\text{GeV}$ | $p_T > 20\,\text{GeV}$ | - |
| | $p_T^{\text{cone}} < 2\,\text{GeV}$ | $p_T^{\text{cone}} < 2\,\text{GeV}$ | - |
| | $2.0 < \eta < 4.5$ | $2.0 < \eta < 4.5$ | - |
| ID | Long track | Long track | $60 < M_{\mu\mu} < 120\,\text{GeV}$ |
| | Triggered | Triggered | $\chi^2_{PV}/\text{NDF} < 5$ |
| | `isMuon` | - | $\lvert\Delta\phi\rvert > 2.7\text{rad}$ |
| | $p_T > 20\,\text{GeV}$ | $p_T > 20\,\text{GeV}$ | - |
| | $p_T^{\text{cone}} < 2\,\text{GeV}$ | $p_T^{\text{cone}} < 2\,\text{GeV}$ | - |
| | $2.0 < \eta < 4.5$ | $2.0 < \eta < 4.5$ | - |

Given the muon stations and TT are not included in long track reconstruction, the tracking efficiency tag is a triggered and identified muon, with a matched Muon TT track providing the probe. Correction factors between MC and data to account for material interactions and detector misalignment are calculated using an analogous method [37].

Using events with both muons passing the `isMuon` requirement, the tag for the trigger efficiency must pass all three trigger stages (`L0EWMuon`, `HLT1SingleMuonHighPT` and `HLT2EWSingleMuonVHighPt`) as described in Section 6.5. With the additional conditions summarised in Table 6.4 applied to ensure sample purity, the probe must also fire all three triggers. Differences between the data and simulation are mostly due to the L0 hardware trigger efficiency and tracking misalignment, which is exaggerated at high $\eta$ [37].

Muons from events passing all three triggers and positive `isMuon` identification are used as the tag for the ID efficiency. A long track passing the `isMuon` requirement is used as the probe. A discrepancy at low $p_T$ is caused by the larger background in data, reducing the efficiency near the threshold. The reduced acceptance of muon stations adjacent to the beampipe degrades the efficiency in the highest $\eta$ bin for data and MC [37].

The reconstruction efficiency for jets is taken from $t\bar{t}$ MC by reconstructing and identifying truth level jets in the fiducial acceptance. The uncertainty on the efficiency is determined by varying individual JetID selection requirements in $Z$+jet data and MC using the fiducial acceptance defined in Table 6.5. As previously defined in Chapter 3, applying tighter cuts to $CPF$, $MPT$, $MTF$ and $N_{point}$ provides a variation in the ratio between data and MC efficiencies. These are summed in quadrature to provide the systematic on the efficiency [37]. The efficiencies over the full $\eta_\mu$ acceptance are provided in Table 6.6 while the individual and combined efficiencies in bins of $\eta_\mu$ are shown in Figure 6.24.

Table 6.5 Selection criteria applied to the $(Z \to \mu\mu)$+jet data and MC for determining jet reconstruction efficiencies [37].

| | |
|---|---|
| Trigger | `L0MuonEW`<br>`HLT1SingleMuonHighPT`<br>`HLT2EWSingleMuonVHighPT` |
| Muon | `isMuon`<br>$p_T > 20\,\text{GeV}$<br>$2.0 < \eta < 4.5$ |
| Boson | $60 < M_{\mu\mu} < 120\,\text{GeV}$ |
| Jet | $p_T > 20\,\text{GeV}$<br>$2.2 < \eta < 4.2$ |

As outlined in Chapter 5, the HF-tagging MVAs are only applicable to jets with a SV within their anti-$k_T$ radius. As discussed in Chapter 3, investigations of the SV-tagger may be conducted using events enriched in $b$- and $c$-decays requiring jets contain a muon or

have an associated $B$ or $D$ meson fully reconstructed. A combined fit to all samples then estimates the flavour content of the jets before and after applying the SV-requirement, hence providing SV efficiencies [76]. A validation of the combined fitting method was performed using $(W \to \mu\nu)$+jet data and MC with the same selection as outlined in Chapter 4 but including the isolation requirement outlined in Section 6.5. This constrains the variables `fdrMin`, `fdChi2`, `tau` (Chapter 5) and the maximum z-coordinate of the SV in order to demonstrate variations between data and MC are of $\mathcal{O}(0.1\%)$. Existing studies show the agreement between data and MC is consistent between Runs I & II. The tagging efficiency is taken from 2016 MC and is expected to be of comparable accuracy to Run I [37]. As such, a conservative estimate of the uncertainty on the jet tagging efficiency is set at 10%.

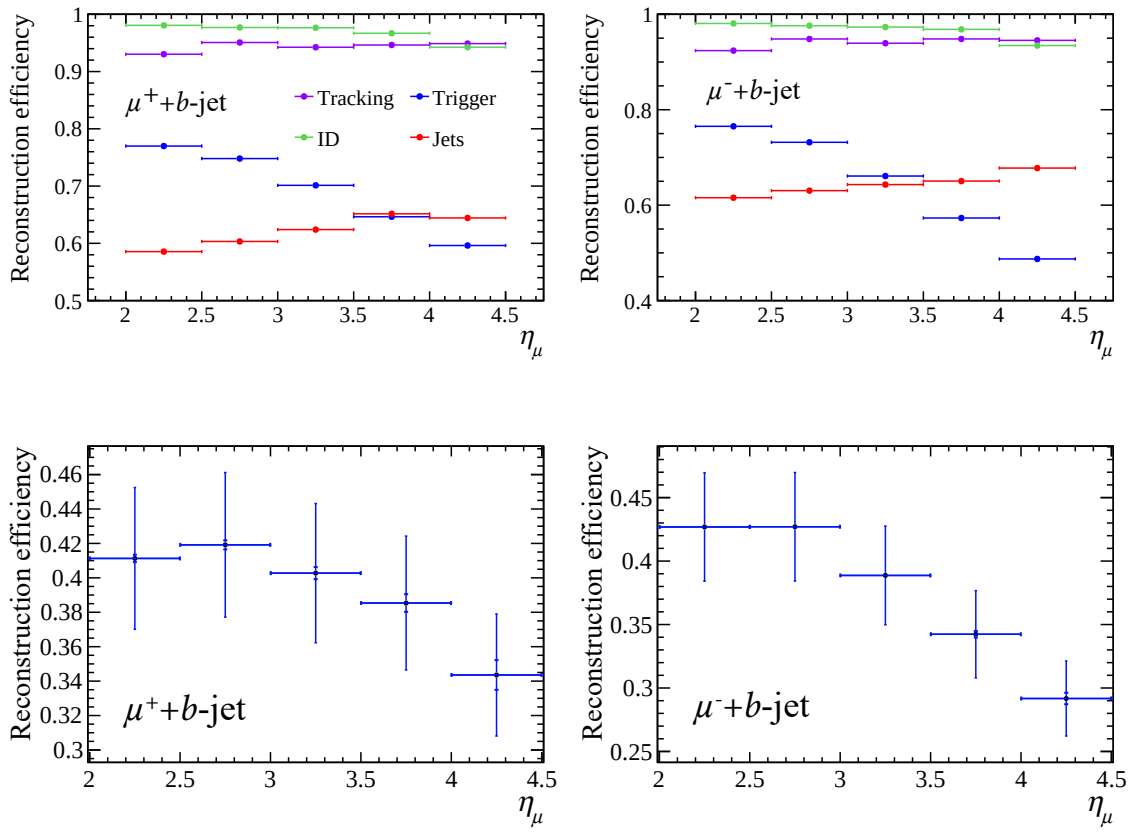



Fig. 6.24 The individual (top) and total (bottom) reconstruction efficiencies listed in Table 6.6 as a function of $\eta_\mu$ for $\mu^\pm$ for top decays to $(W \to \mu)$+$b$-jet final state in Run II.

Table 6.6 Tag and probe efficiencies of top decay to $(W \to \mu)$+$b$-jet event reconstruction in Run II in descending order displaying statistical uncertainties only.

| Requirement | $\mu^+$ | $\mu^-$ |
|---|---|---|
| muon ID | $0.9780 \pm 0.0004$ | $0.9769 \pm 0.0002$ |
| muon tracking | $0.9395 \pm 0.0007$ | $0.9302 \pm 0.0003$ |
| trigger | $0.7453 \pm 0.0011$ | $0.7350 \pm 0.0005$ |
| jets reconstruction | $0.6034 \pm 0.0020$ | $0.6267 \pm 0.0009$ |
| Reconstruction | $0.4132 \pm 0.0015$ | $0.4186 \pm 0.0007$ |

Selection efficiencies correspond to the requirements imposed upon the signal region events outlined in Section 6.5. The impact parameter is tuned to account for mis-modelling using $Z \to \mu\mu$ events passing the selection in Table 6.7 to improve the match of MC to data. The mean and width of a Gaussian used to shift and smear the $x$ and $y$ IP components in bins of $\eta$ and $\phi$, are determined using a $\chi^2$ minimisation between data and MC [37]. A $t\bar{t}$ MC sample, with the tuning applied to final state muons event by event, is used to find the efficiency of the impact parameter cut. The systematic uncertainty is taken as half of the difference between the efficiency with and without the tuning, $\sim 2\%$ [37]. A 2% systematic on the IP selection efficiency is applied uniformly across each $\eta_\mu$ bin. The individual cumulative selection efficiencies are shown in Figure 6.25 as summarised in Table 6.8.

Table 6.7 Selection for events in $Z \to \mu\mu$ data and MC for muon IP tune for the corresponding selection efficiency [37].

| | |
|---|---|
| Trigger | `LOMuonEW`<br>`HLT1SingleMuonHighPT`<br>`HLT2EWSingleMuonVHighPT` |
| Muon | `isMuon`<br>$p_T > 20\,\text{GeV}$<br>$2.0 < \eta < 4.5$ |
| Boson | $60 < M_{\mu\mu} < 120\,\text{GeV}$ |

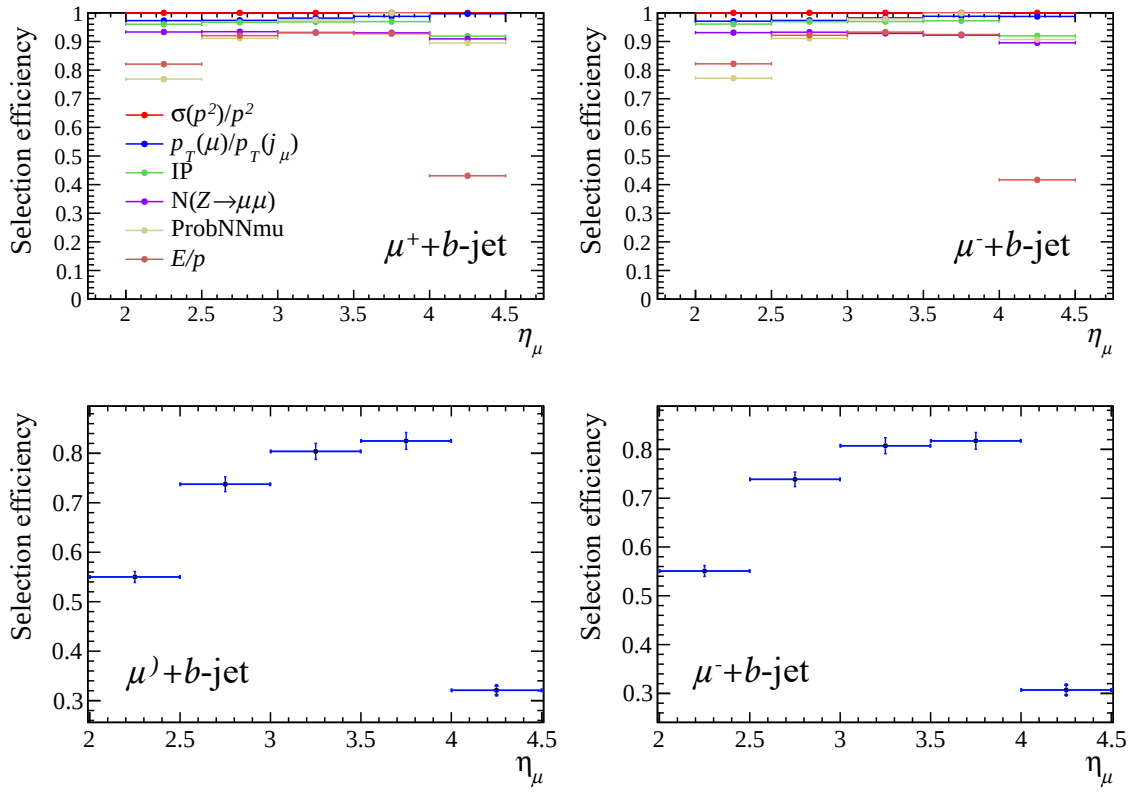



Fig. 6.25 Individual cumulative and combined reconstruction efficiencies for the $\mu$+$b$-jet final state in $t\bar{t}$ MC events.

Table 6.8 Tag-and-probe efficiencies selection requirements for the $(W \to \mu)$+$b$-jet events in Run II in descending order displaying statistical uncertainties only.

| Requirement | $\mu^+$ | $\mu^-$ |
|---|---|---|
| $\sigma(p)^2/p^2$ \| fid. | $1.0000 \pm 0.0000$ | $1.0000 \pm 0.0000$ |
| `ProbNNmu` \| $\sigma(p)^2/p^2$ | $0.8467 \pm 0.0008$ | $0.8454 \pm 0.0009$ |
| $E/p$ \| `ProbNNmu` | $0.7042 \pm 0.0011$ | $0.7125 \pm 0.0013$ |
| IP \| $E/p$ | $0.9630 \pm 0.0005$ | $0.9642 \pm 0.0006$ |
| $Z_{\mu\mu}$ veto \| IP | $0.9329 \pm 0.0007$ | $0.9308 \pm 0.0008$ |
| muon iso. \| $Z_{\mu\mu}$ veto | $0.9731 \pm 0.0004$ | $0.9718 \pm 0.0005$ |
| Selection | $0.5213 \pm 0.0011$ | $0.5254 \pm 0.0013$ |

### 6.7.2 Acceptance factors

An acceptance factor applied in the cross-section calculation (Equation 6.2) is used to account for migrations between bins and in and out of the fiducial acceptance due to imperfect detector resolution as well as differences in fiducial definitions. These factors are calculated using $t\bar{t}$ MC. The effects from muon reconstruction are considered negligible compared to the dominant contribution from the jet $p_T$ resolution. The general equation for each component is shown in Equation 6.4 in terms of true and reconstructed event counts:

$$\mathscr{A} = N_{true}/N_{rec.} \tag{6.4}$$

A sub-leading factor in $\mathscr{A}$ is due to the use of jet momentum in the definition of the $p_T$-imbalance requirement on the $j_\mu$ (Section 6.5). A factor taken from the ratio of $\mathscr{A}_{j_\mu}$ in $(Z \to \mu\mu)$+jet data to MC provides a systematic variation on the value from simulated $t\bar{t}$ (Figure 6.26).

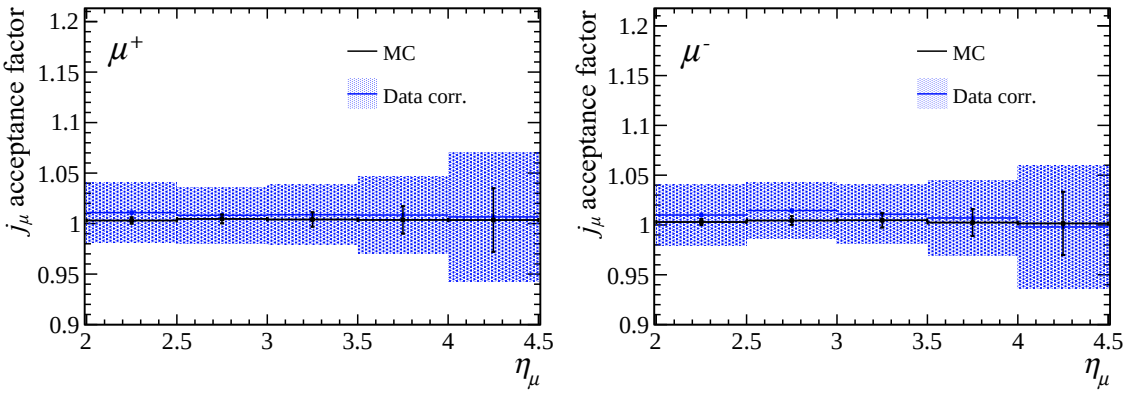



Fig. 6.26 Fiducial acceptance factor on the $p(\vec{j}_b + \vec{j}_\mu)_T$ requirement for comparison to $p(\vec{j}_b + \vec{\mu})_T$ in theory, providing a systematic uncertainty from the difference between $(Z \to \mu)$+jet data and MC.

Another $\mathscr{A}$ contribution, also derived from the jet $p_T$ resolution, accounts for the impact of the 50 GeV $p_T$ threshold. The jet acceptance factor, $\mathscr{A}_{jet}$, is calculated from simulation as the ratio of truth to reconstruction level jets in events satisfying the fiducial requirements. The jet $p_T$ is smeared and shifted using Crystal Ball (CB) fitted values to the $(Z \to \mu\mu)$+jet data and MC (Figure 6.27), where the tail of the distribution accounts for multi-jet events. This choice provides a relatively clean channel in which the $Z$ balances the jet in the transverse plane to test the resolution. The differences between the Gaussian centres from MC and data provide a smear and offset (Figure 6.28) where the parameters of the fit to smeared MC are cross-checked against data. $\mathscr{A}_{jet}$ is the dominant factor in $\mathscr{A}$.

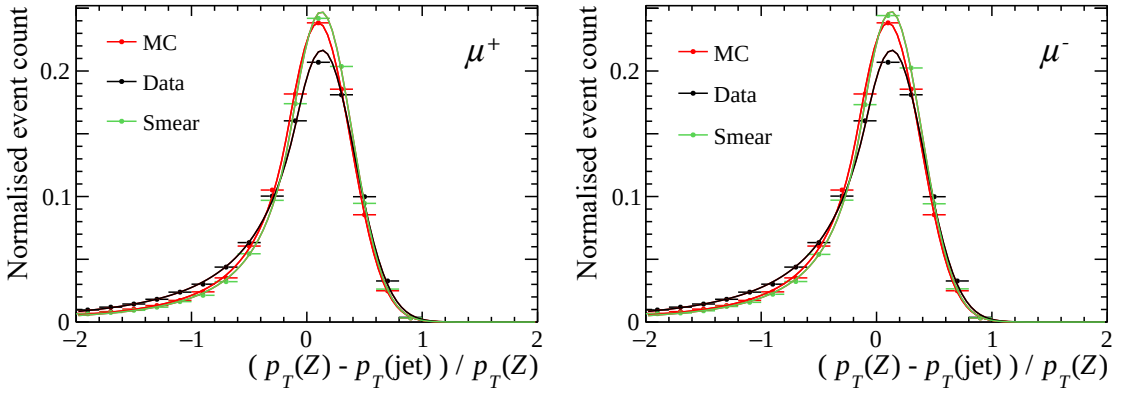



Fig. 6.27 Example Crystal Ball fits to the Data, MC and Smeared MC to extract MC to data smearing and quantify its associated systematic uncertainty.

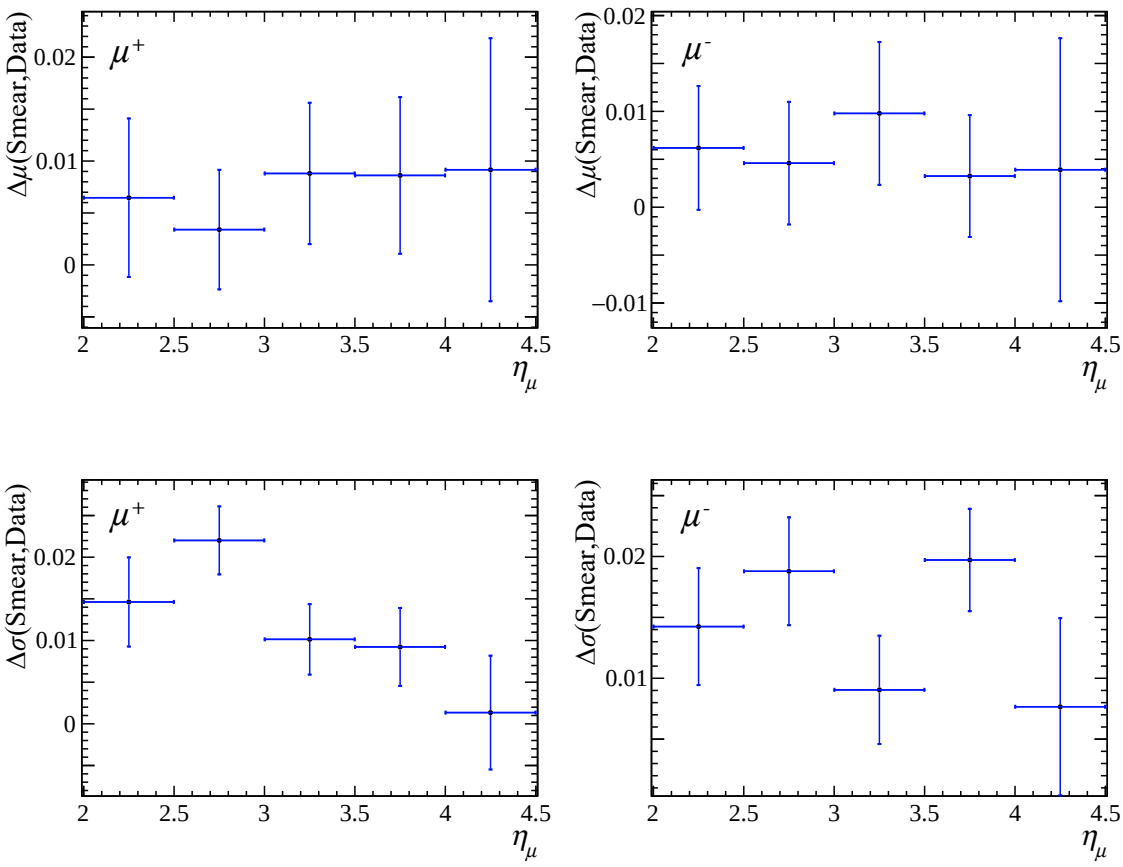



Fig. 6.28 The difference in mean and width of Gaussian cores of Crystal Ball fits used to provide the associated systematic to the MC smearing.

An NLO re-weighting is performed in $(p_T, \eta)$ bins of the jet and a systematic is derived from the difference in $\mathscr{A}_{jet}$ with and without the $k$-factor correction included (Figure 6.29).

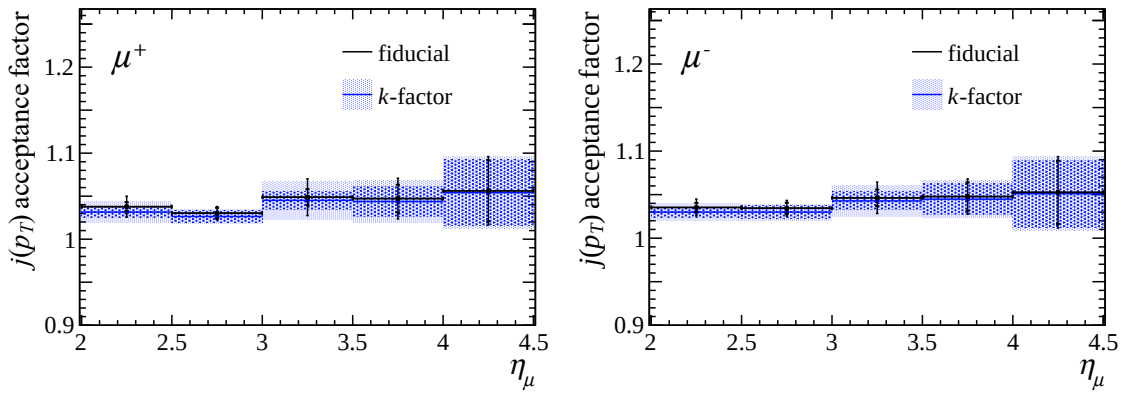



Fig. 6.29 The $p_T$ smeared jet acceptance factor from $(Z \to \mu\mu)$+jet balanced events and its systematic derived from the variation provided by $k$-factor weights in bins of jet $p_T$ and $\eta$.

As a non-negligible number of $W \to \tau\nu$ events decay to a muon that passes the fiducial requirements and selection criteria. A final contribution to $\mathscr{A}$ corrects the measured cross-section to a purely $W \to \mu\nu$ branching fraction, allowing for comparison to purely $W \to \mu\nu$ theory predictions. The values for this $\tau$ exclusion factor are shown as a function of $\eta_\mu$ in Figure 6.30. Uncertainties have been extrapolated into the forward most bin using a conservative factor of two increase from the previous bin.

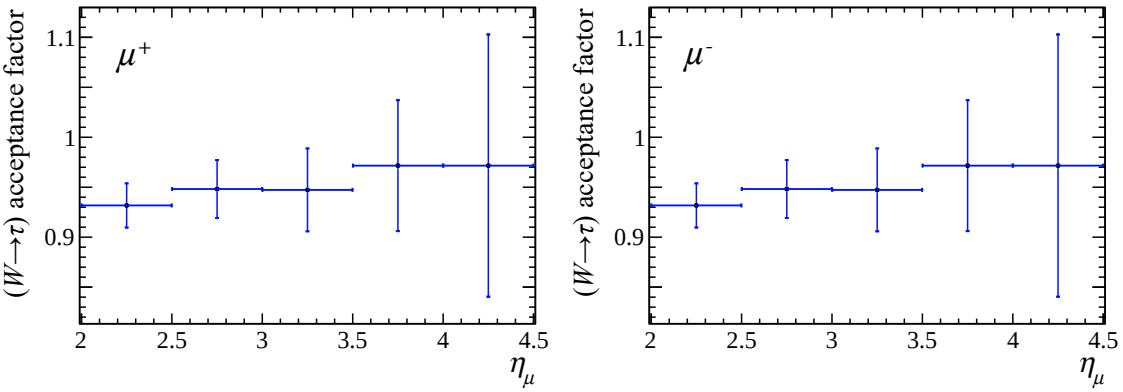



Fig. 6.30 The correction factor from $t\bar{t}$ MC to exclude a measured cross-section component from $(W \to \tau)$+jet for comparison to $(W \to \mu)$+jet theory predictions.

### 6.7.3 Results

The $\sigma$(top) measurements are limited by systematic uncertainties, with overall uncertainties comparable to POWHEG theory uncertainties. The $t$ and $\bar{t}$ cross-sections, observed to $4.6\sigma$ and $3.7\sigma$ respectively, are consistent with the NLO POWHEG expectation. The total cross-section for $t \to \mu$+$b$-jet split by muon sign for $2.0 < \eta < 4.5$ are as follows:

$$\sigma(t)\ [13\,\text{TeV}] \quad = \quad 0.89 \pm 0.06\ (\text{stat}) \pm 0.18\ (\text{syst})\ \text{pb}\ ,$$
$$\sigma(\bar{t})\ [13\,\text{TeV}] \quad = \quad 0.66 \pm 0.05\ (\text{stat}) \pm 0.17\ (\text{syst})\ \text{pb}\ .$$

The cross-sections, as a function of muon $\eta$, are provided in Figure 6.31 with the combined total statistical (inner) and systematic (outer) uncertainties. While each is consistent with the SM in each bin, the distributions in data appear skewed towards higher $\eta$ compared to theory. A summary of the systematic uncertainties provided in Table 6.9 shows that the 2D template fit systematic has the most impact on the measurement ($> 20\%$) followed by the uncertainties on the SV-tag efficiency (10%) and the Run II luminosity (5%).

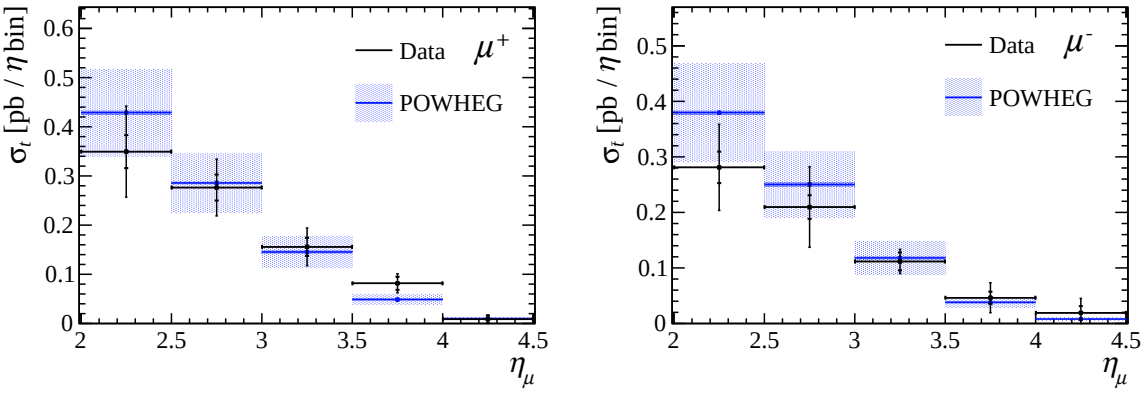



Fig. 6.31 $\sigma_{\text{top}}$ in the $\mu$+$b$-jet final state and POWHEG predictions for $\mu^{\pm}$.

Of the systematic uncertainties present in the cross-sections result, the SV-tag efficiency ($e_{SV}$), muon impact parameter tuning ($e_{IP}$) and integrated luminosity ($L_{int}$) systematics (Table 6.9) are assumed to be correlated between muon signs. Using Equation 6.3 for the asymmetry, they are expected to cancel. The dominant systematic from the HF-yield (Fits) is found to partially cancel when calculated in terms of the asymmetry difference resulting from ALT versus DNN fits (Fits$_{A_C}$). The top cross-section charge asymmetry is as follows:

$$A_C^{\text{top}}\ [13\,\text{TeV}] \quad = \quad 0.14 \pm 0.05\ (\text{stat}) \pm 0.05\ (\text{syst})\ .$$

The $A_C$ value measured across $2.0 < \eta < 4.5$ in data lies $1.0\sigma$ above the NLO SM hypothesis while $1.1\sigma$ above the zero $A_C^{t\bar{t}}$ hypothesis. Although this implies a preference for the SM value for the $t\bar{t}$ asymmetry ($7.3 \pm 2.5\%$), these results remain inconclusive until greater

precision is achieved. The combined top asymmetry in data lies $2.1\sigma$ above a value of zero. In bins of $\eta_\mu$, the asymmetry demonstrates positive excess across most bins each within $1\sigma$ (Figure 6.32) and where the envelope between DNN and ALT fits for the asymmetry reduces the systematic uncertainty in all but the last bin.

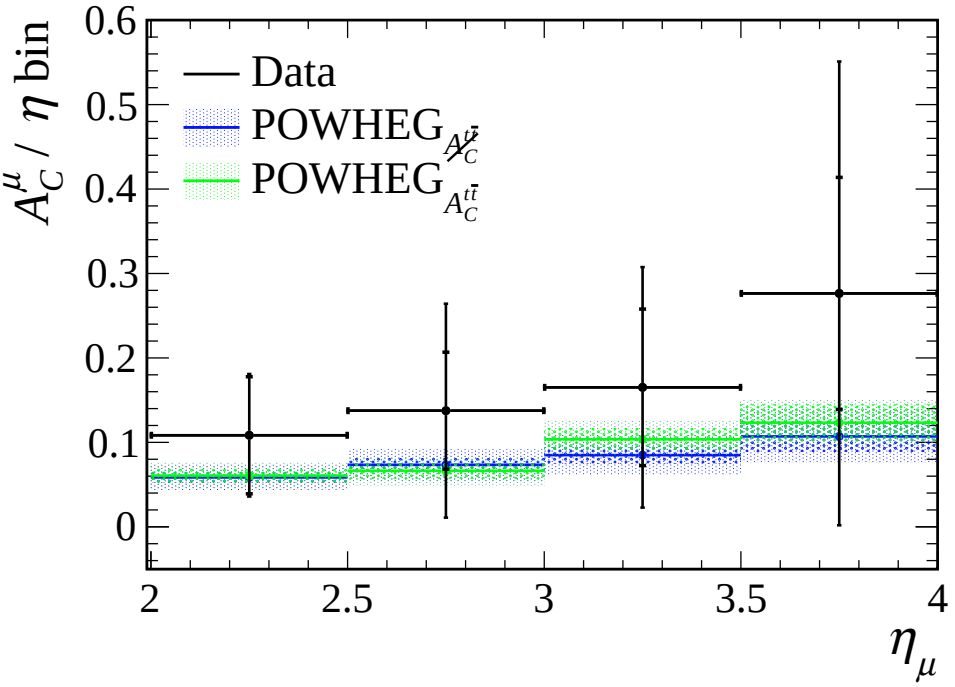



Fig. 6.32 Top cross-section $A_C$ in the $\mu$+$b$-jet final state compared to POWHEG predictions using SM $A_C^{t\bar{t}}$ (green) and $A_C^{t\bar{t}} = 0$ (blue) hypotheses, where the $4.0 < \eta < 4.5$ bin (with $\delta A_C^{syst} > 1$) is cropped from view.

A summary of the overall systematic uncertainties on the asymmetry by $\eta$ bin are also provided in Table 6.9 where: Templates, Fits, $N_{(Wb)}$ and ABCD arise from the yield and background determination; factors used calculate the cross-section, including the SV efficiency $e_{SV}$, the IP efficiency tune $e_{IP}$ and the $L_{int}$, each incur a conservative flat systematic; the acceptance factors from $\mathscr{A}$, including the muon jet, $j_\mu$, the jet $p_T$ threshold and tau exclusion, $\mathscr{T}$; and Fits$_{R_C}$ which represents the $A_C^\mu$ envelope resulting from the ALT and DNN flavour template fits.

Table 6.9 Systematic uncertainties associated with the top cross-section (fractional, %) and asymmetry (absolute) measurements in regular bins of muon pseudorapidity between 2.0 and 4.5, where ~~100~~ corresponds to uncertainties exceeding 100%.

| $\%(\sigma_t)$ | | | $\eta(\mu^+)$ | | | | | $\eta(\mu^-)$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Temp. | 1.42 | 1.63 | 1.75 | 1.73 | 4.76 | 1.46 | 1.49 | 2.17 | 2.63 | 3.32 |
| Fits | 21.63 | 14.25 | 17.77 | 11.13 | 37.64 | 22.84 | 30.84 | 1.26 | 50.87 | ~~100.00~~ |
| $N_{(Wb)}$ | 2.18 | 2.42 | 3.36 | 4.63 | 11.82 | 1.33 | 1.56 | 2.50 | 7.57 | 11.75 |
| ABCD | 1.01 | 0.90 | 0.29 | 0.33 | 3.52 | 0.90 | 1.17 | 4.71 | 1.50 | 0.71 |
| $e_{SV}$ | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 | 10.00 |
| $e_{IP}$ | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 |
| $L_{int}$ | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 |
| $j_\mu$ | 0.77 | 0.36 | 0.46 | 0.47 | 0.31 | 0.69 | 0.98 | 0.60 | 0.45 | 0.35 |
| $p_T$ | 0.61 | 0.37 | 0.35 | 0.33 | 0.21 | 0.52 | 0.43 | 0.34 | 0.29 | 0.16 |
| ~~$\not{x}$~~ | 0.29 | 0.20 | 0.34 | 0.59 | 1.18 | 0.29 | 0.20 | 0.34 | 0.59 | 1.18 |
| $\frac{\delta(\sigma_t)}{\sigma_t}$ | 24.60 | 18.48 | 21.43 | 16.66 | 41.48 | 25.61 | 32.96 | 12.80 | 52.76 | ~~100.00~~ |
| $\text{Fits}_{R_C}$ | 0.01 | 0.10 | 0.10 | 0.23 | ~~1.00~~ | | | | | |
| $\delta A_C$ | 0.02 | 0.11 | 0.11 | 0.24 | ~~1.00~~ | | | | | |

## 6.8 Summary

The cross-sections of combined single-$t$ and $t\bar{t}$ decays in the $\mu$+$b$-jet final state are found to be within $1\sigma$ of NLO standard model prediction. As they stand, the DNNs produced for $b$-tagging in Run II data have sufficient disparities between the central and alternative fit that the top measurements are dominated by 15-25% systematics on the $b$-jet yields. Despite employing a procedure to at least partially cancel the largest of the systematic uncertainties from the cross-section (Fits, $e_{SV}^{REC}$, $L_{int}$), each expected to be correlated between muons of each charge, the measurement of the combined top charge asymmetry presented in this chapter is inconclusive. The statistical and systematic uncertainties on $A_C$ are shown to contribute in roughly equal terms.

## Future work

While the DNN template fits provide better $\chi^2$ values than those of the alternative fits, both provide good agreement with data despite their disagreement with one another. As is demonstrated by the fit projections in the MVA inputs (Chapter 5), a disagreement between data and MC may be to blame. It is unclear whether this may be alleviated with improved SV selection or more rigorous pre-processing for the MVAs. No improvement to the yield systematic was provided by the following:

- Producing flavour templates with, and fitting to, the Run II BDT outputs (Chapter 5);
- Deferring to template fits using the Run I BDT outputs [76];
- Using ALT fit templates f($M_{cor}$), rather than f(log($M_{cor}$)) as used in Run II MVAs;
- Allowing the floating HF-yield parameters to vary below zero.

Adapting the limits of the ALT fit templates, at which the under- & overflow are summed into the first and last bins [76], had a negligible impact on the central values for the cross-sections and asymmetry. The precision of the cross-sections is both improved and degraded bin to bin, $\mathcal{O}$(1-10%), by this change. The degree to which the leading systematic (from the HF-yields) is cancelled between charges for the asymmetry measurement proved sensitive to this change. For an integrated $A_C$ measurement, the total systematic is reduced to $\sim 3\%$ leaving the 5% statistical uncertainty dominant; for a differential measurement, improvements bin to bin vary considerably. This could imply that the instability of the results between fits arises predominantly from the definition of the ALT fit template axis ranges.

While a reduced $\eta_\mu$ binning does not affect the cross-section systematics from the HF-yields (remaining $> 20\%$), preliminary studies demonstrate that the cancellation of this systematic in the asymmetry measurement (reducing to $\sim 3\%$) can also produce a statistics limited measurement. Such results imply that, even with the current room for improvement in understanding and constraining cross-section systematics, the sensitivity to resolve the $A_C$ should be in reach with sufficiently large data-sets.

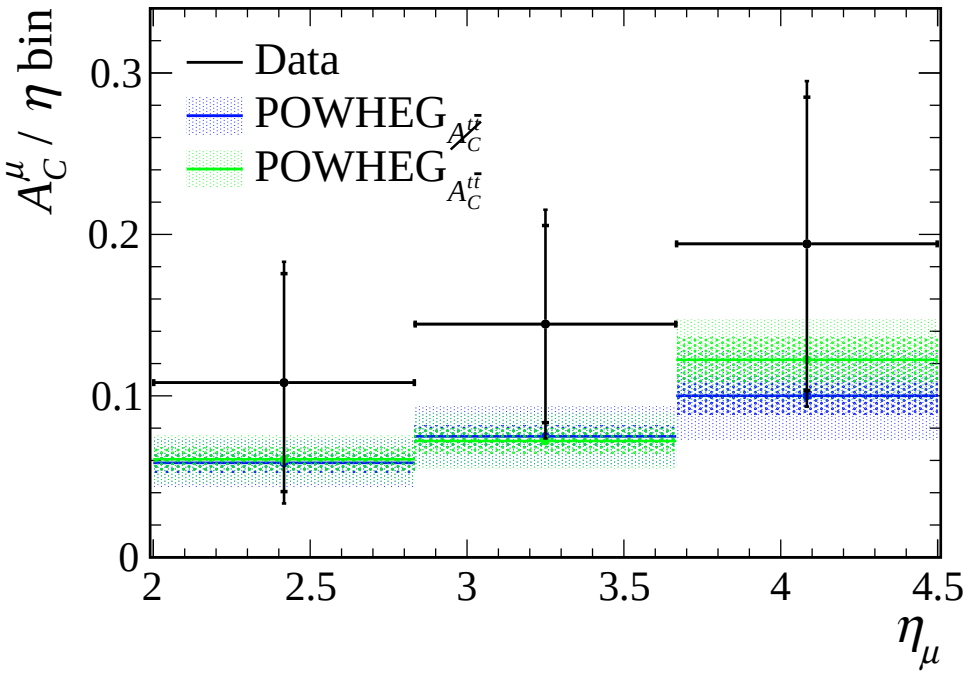



Fig. 6.33 Preliminary cross-section $A_C$ for top quark decay in the $\mu$+$b$-jet final state compared to POWHEG predictions using SM $A_C^{t\bar{t}}$ (green) and $A_C^{t\bar{t}} = 0$ hypotheses (blue) with a reduced $\eta_\mu$ binning scheme demonstrating a statistically limited measurement.

Further steps to improve upon the results from Section 6.7.3 include:

- A smearing procedure, applied to the MC to resemble data, may address a mismatch between them. If MC proves difficult to transform appropriately in one or more variables, then both data and MC could be smeared in those specific inputs to reduce model dependence upon them. By sacrificing discrimination power based on these variables, the smearing may provide fits with reduced systematics.

- Barlow-Beeston fitting [118] may be used to reduce flavour fit systematics; constructing subsidiary Poisson measurements to constrain the histogram parameters provides variations of the bin contents within the statistical uncertainty of the MC. This implementation would absorb the template statistical uncertainty, previously estimated using bootstrapped variations, into the fit itself.

- Templates produced from models only trained on SV-information may provide reduced systematics. As discussed in Chapter 5, SV-only $b$-tagging showed marginally increased discrimination against $c$-jets while SV-only light rejection experienced a more

significant decrease in performance. Combining a SV-jet ($udsg|bc$) classifier with a SV-only ($b|c$) classifier for 2D fits may improve HF-yield precision.

- Large samples of SV-tagged light-jets may be produced through projecting backwards reconstructed SVs from data forwards ($z' = -z$), superimposed onto abundant light-jet MC. While this was discounted for MVA training purposes to avoid introducing further room for MC to data discrepancy, this may yet provide light-jet templates for the response of the existing MVA to events which more accurately represent data.

- A cross-check for potential biases in the HF-tagging may be performed using the $W$+$c$ yield. By performing a subtraction of NLO $Wc/Wj$ normalised $W$+jet events, analogous to that in the analysis procedure (Chapter 6), the remaining $c$-jet yield may be tested for consistency with the negligible expectation coming from top decays [31].

- Systematic uncertainties on the muon calorimeter $E/p$ requirement efficiency, due to its heavy $\eta$ dependence, and the trigger efficiency, due to its considerable asymmetry between $\mu^+$ and $\mu^-$, should be investigated along with correlations between systematic uncertainties for future asymmetry measurements.

- The SV efficiency for Run II is to be calibrated to data using independently dedicated $c$- & $b$-tagging studies, which are already underway and the full Run II integrated luminosity measurement, having recently become available, should also be included.

- Closure tests should be developed for continued efforts in the $lb$ channel and a binned likelihood ratio test may provide clearer comparison between the SM and zero-$A_C^{t\bar{t}}$ hypotheses for an improved differential $\mu$+$b$ asymmetry measurement.

Analysis of top quark channels, including existing Run II measurements presented in this work and $t\bar{t}$ extensions (Section 6.3), provide the opportunity for further measurements:

- With the addition of efficiencies calculated from $W$+HF-jet and $W$+jet samples, the existing analysis for $t \to \mu b$ would provide differential cross-sections and asymmetries for these 'subsidiary' processes.

- The inclusion of a second SV-tagged jet for the study of the $t\bar{t} \to \mu bb$ provides lower statistics and would have the SV-tag efficiency, contributing a significant systematic uncertainty, included twice in the cross-section calculation. However, the use of flavour tagging on a combination of high $p_T$ jets [116] may provide reduced $b(b)$-yield systematics. In principle, the methods discussed in Chapter 6 may be extended to the $lbb$ channel and could provide an asymmetry excluding single-$t$ contamination.

- The $lbb$ channel provides colour flow in $t\bar{t}$ by measuring relative jet pull angle in the di-jet final state [119].
- The $llb$ channel provides spin correlations preserved in the di-lepton channel through the $\Delta\phi$ between lepton tracks [120].

As the proportion of $t\bar{t}$ from gluon fusion increases with centre of mass energy, the magnitude of the charge asymmetry is expected to decrease at higher running energies. As statistical limitations become less relevant in Run III and beyond, a better understanding of underlying processes and known background contributions, as well as the ability to constrain systematic uncertainties, will increase in importance. Projections for the sensitivity to the asymmetry based on the statistics of future Runs at 14 TeV with LHCb are shown in Figure 6.34. This demonstrates the utility of such forward region measurements with larger data sets and their potential contributions to the LHC top program.

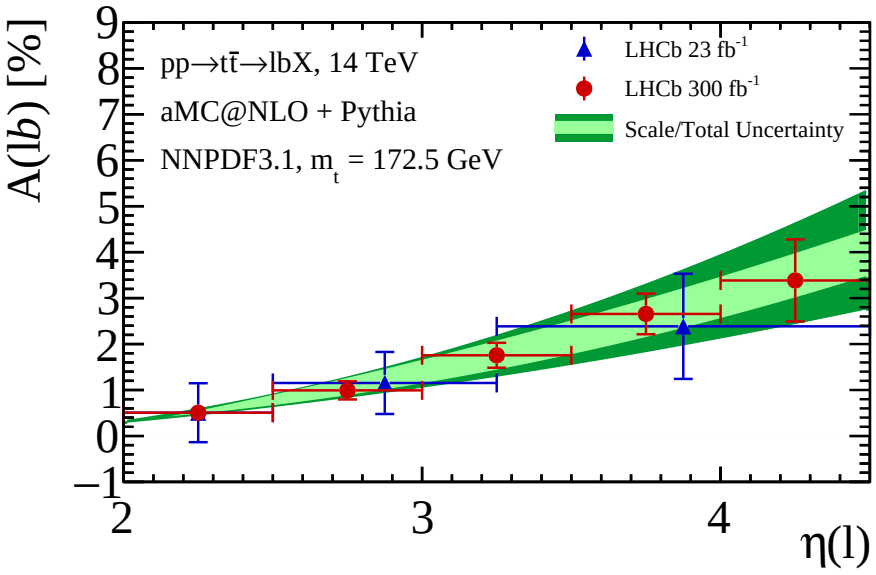



Fig. 6.34 Top asymmetry (Equation 1.38) in $lbX$ final state at 14 TeV in the LHCb $\eta_\mu$ acceptance where bands show the uncertainty on the theoretical predictions due to scale variations (green) and due to combined scale, PDF and $\alpha_S$ variations (yellow) and with statistical uncertainties projected for Runs 3 & 5 (23 & 300 $\text{fb}^{-1}$ of data) [96].

LHCb GitLab

# Conclusion

This thesis presented the differential measurement of the top production cross-section, performed as a test of perturbative quantum chromodynamics in the Standard Model, to provide forward region top data to constraint $g$-PDF uncertainties and make an LHCb first measurement of the charge asymmetry. The leptonic top quark decays in the $\mu + b$-jet final state, reconstructed using forward region data from the LHCb detector, used in the analysis correspond to an integrated luminosity of $5.4\,\text{fb}^{-1}$ at $\sqrt{s} = 13\,\text{TeV}$.

Final state jets tagged with secondary decay vertices are crucial to top reconstruction at LHCb and, to this end, the jet reconstruction configuration and heavy flavour tagging procedure were each updated and optimised to Run II conditions. The studies developing the input filter for the high level trigger integrated particle flow algorithm and subsequent jet reconstruction performance were presented in Chapter 4 (Future work, Page 82). This configuration became the new default for Run II legacy analyses and was implemented in the Monte Carlo based studies, training and validating heavy flavour tagging deep neural networks for Run II. The development and assessment of deep learning models, to be implemented for $b$-tagging in the top analysis, was communicated in Chapter 5 (Future work, Page 98).

With uncertainties on the cross-section measurement dominated by heavy flavour fitting systematics, the precision was restricted to $\mathcal{O}(10\%)$. While this was sufficient to confirm the production of combined $t\bar{t}$ and single-$t$, the charge asymmetry was unable to be resolved beyond the single-$t$ positive offset. The results of the top quark analysis were presented in Chapter 6 (Future work, Page 133). Given sufficient progress on systematics, this analysis should aim to be the last statistically limited top quark measurement at LHCb. Run III data from LHCb will contribute more substantially to the LHC top program and, if Run II can provide a <10% precision on $A_C$ that is statistically limited, then by the end of Long Shutdown 2 the world will have gained another precision top physics instrument.

# References


[1] Particle Data Group, M. Tanabashi *et al.*, *Review of Particle Physics*, Physics Review Letters D **98** (2018) 030001.

[2] J. Rohlf, *Modern physics from $a_\alpha$ to $Z^0$*, John Wiley and Sons, New York (1994) .

[3] A. Salam, *Unification of Fundamental Forces: The First 1988 Dirac Memorial Lecture*, Cambridge University Press (2005) .

[4] S. Perlmutter, M. S. Turner, and M. White, *Constraining dark energy with type-Ia supernovae and large-scale structure*, Physical Review Letters **83** (1999) 670.

[5] J. F. Cornwell, *Group theory in physics: An introduction*, Academic press, (1997).

[6] A. Pich, *CP violation*, ICTP Series Theoretical Physics **10** (1994) 14, arXiv:hep-ph/9312297.

[7] S. Dawson, *Introduction to electroweak symmetry breaking*, BNL-HET-99/1 (1999) arXiv:hep-ph/9901280.

[8] P. Achard *et al.*, *Measurement of the running of the electromagnetic coupling at large momentum-transfer at LEP*, Physics Letters B **623** (2005) 26.

[9] A. Pich, *Quantum chromodynamics*, FTUV/95-19; IFIC/95-19 (1995) arXiv:hep-ph/9505231.

[10] S. Alekhin *et al.*, *Proceedings, High-Precision $\alpha_s$ Measurements from LHC to FCC-ee*, CERN-PH-TH-2015-299 (2015) arXiv:1512.05194.

[11] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nuclear Physics B **126** (1977) 298.

[12] R. Ball *et al.*, *Parton distributions from high-precision collider data*, The European Physical Journal C **77** (2017) 663.

[13] G. P. Salam, *An introduction to leading and next-to-leading BFKL*, Acta Phys. Polon. (1999) arXiv:hep-ph/9910492.

[14] J. C. Collins, *Sudakov form factors*, World Scientific: Perturbative QCD (1989) arXiv:hep-ph/0312336.

[15] M. R. Whalley, D. Bourilkov, and R. C. Group, *HERA and the LHC: A Workshop on the Implications of HERA and LHC Physics*, The Les Houches accord PDFs (LHAPDF) and LHAGLUE (2005) , arXiv:hep-ph/0508110.

[16] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *A critical appraisal of NLO+PS matching methods*, Journal of High Energy Physics **2012** (2012) 49.

[17] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, Journal of High Energy Physics **2007** (2007) 070.

[18] J. Alwall *et al.*, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, Journal of High Energy Physics **2014** (2014) 79.

[19] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, Journal of High Energy Physics **2013** (2013) 15.

[20] R. Corke and T. Sjöstrand, *Interleaved parton showers and tuning prospects*, Journal of High Energy Physics **2011** (2011) 32.

[21] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, *Parton fragmentation and string dynamics*, Physics Reports **97** (1983) 31.

[22] A. Buckley *et al.*, *Global fit of top quark effective theory to data*, Physical Review D **92** (2015) 091501.

[23] M. Czakon and A. Mitov, *Top++: a program for the calculation of the top-pair cross-section at hadron colliders*, Computer Physics Communications **185** (2014) 2930.

[24] ATLAS Collaboration, G. Aad *et al.*, *Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s} = 13\,TeV$ with the ATLAS experiment*, Physics Letters B **810** (2020) 135797, arXiv:2006.13076.

[25] A. M. Sirunyan *et al.*, *Measurement of the single top quark and anti-quark production cross sections in the t channel and their ratio in proton-proton collisions at $\sqrt{s} = 13\,TeV$*, Physics Letters B **800** (2020) 135042.

[26] CMS Collaboration, A. M. Sirunyan *et al.*, *Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at $\sqrt{s} = 13\,TeV$*, JHEP **10** (2018) 117, arXiv:1805.07399.

[27] M. Czakon, P. Fiedler, and A. Mitov, *Resolving the Tevatron top quark forward-backward asymmetry puzzle: fully differential next-to-next-to-leading-order calculation*, Physical Review Letters **115** (2015) 052001.

[28] Y. Peters, *Top anti-top Asymmetries at the Tevatron and the LHC*, FERMILAB-CONF-12-616-E (2012) arXiv:1211.6028.

[29] R. Gauld, *Measuring top quark production asymmetries at LHCb*, LHCb-PUB-2013-009 (2013) cds.cern.ch/record/1557385.

[30] R. Aaij *et al.*, *Study of W boson production in association with beauty and charm*, Physical Review D **92** (2015) 052001.

[31] R. Aaij *et al.*, *First observation of top quark production in the forward region*, Physical review letters **115** (2015) 112001.

[32] L. Evans and P. Bryant, *LHC machine*, Journal of instrumentation **3** (2008) S08001.

[33] J. Gruschke, *Detection of top quarks and first measurement of the $t\bar{t}$ cross-section at a center of mass energy of 7 TeV with the CMS experiment at the LHC*, KIT (2011) cds.cern.ch/record/1359140.

[34] G. Aad *et al.*, *The ATLAS experiment at the CERN large hadron collider*, Jinst **3** (2008) S08003.

[35] CMS Collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, (2008). doi: 10.1088/1748-0221/3/08/S08004.

[36] LHCb Collaboration, R. Aaij *et al.*, *LHCb detector performance*, International Journal of Modern Physics A **30** (2015) 1530022.

[37] H. Wark, *Measurement of the top quark pair production cross-section in the $\mu eb$ final state at $\sqrt{s} = 13$ TeV with the LHCb detector*, University of Liverpool (2020) .

[38] S. Cadeddu *et al.*, *LHCb reoptimized detector design and performance: Technical Design Report*, LHCb-TDR-009 (2003) .

[39] J. S. Anderson, *Testing the electroweak sector and determining the absolute luminosity at LHCb using dimuon final states*, University College Dublin (2008) cds.cern.ch/record/1170478.

[40] R. Aaij *et al.*, *Performance of the LHCb vertex locator*, Journal of Instrumentation **9** (2014) P09007, arXiv:1405.7808.

[41] A. A. Alves Jr *et al.*, *The LHCb detector at the LHC*, Journal of instrumentation **3** (2008) S08005.

[42] Y. V. Pavlenko *et al.*, *LHCb VELO (VErtex LOcator): Technical Design Report*, LHCb-TDR-005; CERN-LHCC-2001-011 (2001) cds.cern.ch/record/504321.

[43] M. Vesterinen, *Considerations on the LHCb dipole magnet polarity reversal*, LHCb Public Notes, CERN-LHCb-PUB-2014-006 (2014) cds.cern.ch/record/1642153, url = https://cds.cern.ch/record/1642153.

[44] M. Tobin, *Performance of the LHCb tracking detectors*, The 21st International Workshop on Vertex Detectors **167** (2013) 047.

[45] P. d'Argent *et al.*, *Improved performance of the LHCb outer tracker in LHC Run 2*, Journal of Instrumentation **12** (2017) P11016.

[46] R. Arink *et al.*, *Performance of the LHCb outer tracker*, JINST **9** (2013) P01002, arXiv:1311.3893.

[47] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, The European Physical Journal C **73** (2013) 1.

[48] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, Journal of Instrumentation **8** (2013) P04022.

[49] C. A. Beteta *et al.*, *Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC*, CERN Report, CERN-LHCb-DP-2020-001 (2020) arXiv:2008.11556.

[50] LHCb, I. Machikhiliyan, *The LHCb electromagnetic calorimeter*, Journal Physics Conference Series **160** (2009) 012047.

[51] LHCb Collaboration, O. Omelaenko, *LHCb calorimeters: Technical design report*, CERN Reports, CERN-LHCC-2000-036 (2000) cds.cern.ch/record/494264.

[52] I. Machikhiliyan *et al.*, *Current status and performance of the LHCb electromagnetic and hadron calorimeters*, Journal of Physics: Conference Series, IOP Publishing **293** (2011) 012052.

[53] A. A. Alves Jr *et al.*, *Performance of the LHCb muon system*, Journal of Instrumentation **8** (2013) P02022.

[54] LHCb Collaboration, G. Wilkinson, R. Lindner *et al.*, *LHCb PID upgrade technical design report*, CERN Report, LHCB-TDR-014 (2013) cds.cern.ch/record/1624074.

[55] R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, Journal of Instrumentation **14** (2019) P04013.

[56] R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the lhc*, Journal of Instrumentation **14** (2019) P04013.

[57] R. Aaij *et al.*, *A comprehensive real-time analysis model at the LHCb experiment*, Journal of Instrumentation **14** (2019) P04006.

[58] S. Benson, V. Gligorov, M. A. Vesterinen, and J. M. Williams, *The LHCb Turbo Stream*, J. Phys. Conf. Ser **664** (2015) 082004.

[59] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, Journal of Physics: Conference Series, IOP Publishing **331** (2011) 032047.

[60] D. J. Lange, *The EvtGen particle decay simulation package*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **462** (2001) 152.

[61] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, The European Physical Journal C-Particles and Fields **45** (2006) 97.

[62] GEANT Collaboration, S. Agostinelli *et al.*, *GEANT4 - a simulation toolkit*, Nuclear Instruments and Methods in Physics Research A **506** (2003) 0.

[63] J. Allison *et al.*, *GEANT4 developments and applications*, IEEE Transactions on nuclear science **53** (2006) 270.

[64] LHCb Collaboration, M. Clemencic *et al.*, *The LHCb simulation application, Gauss: design, evolution and experience*, Journal of Physics: Conference Series,IOP Publishing **331** (2011) 032023.

[65] LHCb Collaboration, M. Stahl, *Machine learning and parallelism in the reconstruction of LHCb and its upgrade*, Journal of Physics: Conference Series **898** (2017) 042042, arXiv:1710.08947.

[66] M. Kucharczyk, P. Morawski, and M. Witek, *Primary vertex reconstruction at LHCb*, LHCb Public Notes, CERN-LHCb-PUB-2014-044 (2014) cds.cern.ch/record/1756296.

[67] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, The European Physical Journal C **72** (2012) 1896.

[68] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-$k_T$ jet clustering algorithm*, Journal of High Energy Physics **2008** (2008) 063.

[69] G. P. Salam, *Towards jetography*, The European Physical Journal C **67** (2010) 637.

[70] E. Rodrigues, *The Scikit-HEP Project*, EPJ Web of Conferences, EDP Sciences **214** (2019) 06005.

[71] R. Aaij *et al.*, *Study of forward Z+jet production in pp collisions at $\sqrt{s} = 7\,TeV$*, Journal of High Energy Physics **2014** (2014) 33.

[72] CMS collaboration, F. Beaudette, *The CMS particle flow algorithm*, International Conference on Calorimetry for the High Energy Frontier (2013) arXiv:1401.8155.

[73] W. Barter, *Z boson and associated jet production at the LHCb experiment*, CERN-THESIS-2014-178, Cambridge University (2014) cds.cern.ch/record/1970903.

[74] M. De Cian, S. Farry, P. Seyfert, and S. Stahl, *Fast neural-net based fake track rejection in the LHCb reconstruction*, LHCb-PUB-2017-011, CERN-LHCb-PUB-2017-011 (2017) cds.cern.ch/record/2255039.

[75] A. Hoecker *et al.*, *TMVA-toolkit for multivariate data analysis*, CERN-OPEN-2007-007 (2007) arXiv:physics/0703039.

[76] LHCb Collaboration, R. Aaij *et al.*, *Identification of beauty and charm quark jets at LHCb*, Journal of Instrumentation **10** (2015) P06013.

[77] B. Sciascia, *LHCb Run 2 Trigger Performance*, PoS: BEAUTY2016, LHCb-PROC-2016-020. CERN-LHCb-PROC-2016-020 (2016) cds.cern.ch/record/2208038.

[78] T. Ida, M. Ando, and H. Toraya, *Extended pseudo-Voigt function for approximating the Voigt profile*, Journal of Applied Crystallography **33** (2000) 1311.

[79] C. Chatfield and A. Collins, *Introduction to multivariate analysis*, vol. 1, CRC Press, (1981).

[80] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees*, CRC press, (1984).

[81] L. Rokach, *Decision forest: Twenty years of research*, Information Fusion **27** (2016) 111.

[82] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **543** (2005) 577.

[83] A. K. Jain, J. Mao, and K. M. Mohiuddin, *Artificial neural networks: A tutorial*, Computer **29** (1996) 31.

[84] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*, vol. 1, MIT press Cambridge, (2016).

[85] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*, O'Reilly Media, (2019).

[86] A. Gulli and S. Pal, *Deep learning with Keras*, Packt Publishing Ltd, (2017).

[87] M. Abadi *et al.*, *Tensorflow: A system for large-scale machine learning*, 12th USENIX symposium on operating systems design and implementation (OSDI 16) (2016) 265, arXiv:1605.08695.

[88] Y. A. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller, *Efficient backprop*, Neural networks: Tricks of the trade (2012) 9.

[89] J. Stevens and M. Williams, *uBoost: A boosting method for producing uniform selection efficiencies from multivariate classifiers*, Journal of Instrumentation **8** (2013) P12013.

[90] T.-F. Wu, C.-J. Lin, and R. C. Weng, *Probability estimates for multi-class classification by pairwise coupling*, Journal of Machine Learning Research **5** (2004) 975. doi:10.5555/1005332.1016791.

[91] T. Chen and C. Guestrin, *XGBoost: A scalable tree boosting system*, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2016) 785.

[92] G. Ke *et al.*, *LightGBM: A highly efficient gradient boosting decision tree*, Advances in neural information processing systems (2017) 3146. doi:10.5555/3294996.3295074.

[93] LHCb Collaboration, R. Aaij *et al.*, *Measurement of forward $W$ and $Z$ boson production in association with jets in proton-proton collisions at $\sqrt{s} = 8\,TeV$*, JHEP **05** (2016) 131, arXiv:1605.00951.

[94] M. Czakon, M. L. Mangano, A. Mitov, and J. Rojo, *Constraints on the gluon PDF from top quark pair production at hadron colliders*, Journal of High Energy Physics **2013** (2013) 167.

[95] R. Gauld, *Feasibility of top quark measurements at LHCb and constraints on the large-x gluon PDF*, Journal of High Energy Physics **2014** (2014) 126.

[96] LHCb, R. Aaij *et al.*, *Physics case for an LHCb Upgrade II-Opportunities in flavour physics, and beyond, in the HL-LHC era*, LHCb Public Notes (2018) , arXiv:1808.08865.

[97] J. A. Aguilar-Saavedra and M. Pérez-Victoria, *Simple models for the top asymmetry: constraints and predictions*, Journal of High Energy Physics **2011** (2011) 97.

[98] G. Aad *et al.*, *Muon recontruction performance of the ATLAS detector in proton-proton collision data at* $\sqrt{s} = 13\,TeV$, The European Physical Journal C **76** (2016) 292.

[99] A. Miucci *et al.*, *ATLAS b-jet identification performance and efficiency measurement with* $t\bar{t}$ *events in pp collisions at* $\sqrt{s} = 13\,TeV$, The European physical journal C **79** (2019) .

[100] C. Zhang and S. Willenbrock, *Effective-field-theory approach to top-quark production and decay*, Physical Review D **83** (2011) 034006.

[101] N. P. Hartland *et al.*, *A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector*, Journal of High Energy Physics **2019** (2019) 100.

[102] M. Czakon *et al.*, *Top-quark charge asymmetry at the LHC and Tevatron through NNLO QCD and NLO EW*, Physical Review D **98** (2018) 014003.

[103] ATLAS Collaboration, H. J, *Inclusive and differential measurement of the charge asymmetry in* $t\bar{t}$ *events at* $13\,TeV$ *with the ATLAS detector*, EPS-HEP (2019) .

[104] J. V. Mead, *Tops in the forward region*, TOP2018 proceedings (2019) arXiv:1901.03648.

[105] R. Aaij *et al.*, *Measurement of forward* $t\bar{t}$, $w + b\bar{b}$ *and* $w + c\bar{c}$ *production in pp collisions at* $\sqrt{s} = 8\,TeV$, Physics Letters B **767** (2017) 110.

[106] R. Aaij *et al.*, *Measurement of forward top pair production in the dilepton channel in pp collisions at* $\sqrt{s} = 13\,TeV$, Journal of High Energy Physics **2018** (2018) 174.

[107] J. Alwall *et al.*, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, Journal of High Energy Physics **2014** (2014) 79.

[108] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, Journal of High Energy Physics **2010** (2010) 43.

[109] S. Alioli, P. Nason, C. Oleari, and E. Re, *Vector boson plus one jet production in POWHEG*, Journal of High Energy Physics **2011** (2011) 95.

[110] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, Journal of High Energy Physics **2007** (2007) 081.

[111] G. Aad *et al.*, *Measurement of the inclusive isolated prompt photon cross-section in pp collisions at* $\sqrt{s} = 7\,TeV$ *using* $35\,pb^{-1}$ *of ATLAS data*, Physics Letters B **706** (2011) 150.

[112] R. Aaij *et al.*, *Search for massive long-lived particles decaying semi-leptonically in the LHCb detector*, The European Physical Journal C **77** (2017) 224.

[113] F. James and M. Roos, *MINUIT: a system for function minimization and analysis of the parameter errors and corrections*, Computer Physics Communications **10** (1975) 343. `cds.cern.ch/record/310399`.

[114] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, Computing in High Energy and Nuclear Physics (CHEP03) (2003) `arXiv:physics/0306116`.

[115] B. Efron, *Bootstrap methods: another look at the jackknife*, Breakthroughs in statistics (1992) 569.

[116] LHCb, R. Aaij *et al.*, *Measurement of differential $b\bar{b}$- and $c\bar{c}$-dijet cross-sections in the forward region of pp collisions at $\sqrt{s}$ = 13 TeV*, JHEP **02** (2021) 023, `arXiv:2010.09437`.

[117] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 7\,TeV$*, Journal of High Energy Physics **2015** (2015) 39.

[118] R. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Computer Physics Communications **77** (1993) 219.

[119] G. Aad *et al.*, *Measurement of colour flow with the jet pull angle in $t\bar{t}$ events using the ATLAS detector at $\sqrt{s} = 8$ TeV*, Physics Letters B **750** (2015) 475.

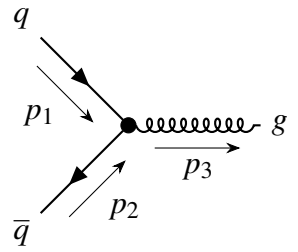
[120] ATLAS, M. Aaboud *et al.*, *Measurements of top-quark pair spin correlations in the $e\mu$ channel at $\sqrt{s} = 13$ TeV using pp collisions in the ATLAS detector*, The European Physical Journal C **80** (2020) 754, `arXiv:1903.07570`.

# Appendix A

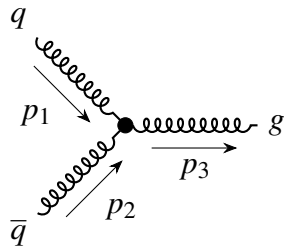
# Theoretical overview

This section provides explicit forms for vertices (and their couplings) and $qg \to t\bar{t}(j)$ Feynman diagrams referred to in the main text (Chapter 1), found in Figures A.1 & A.2 respectively.

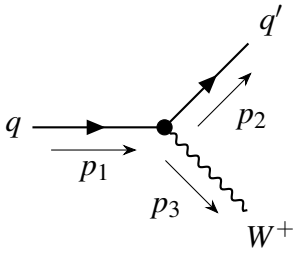
## Observables



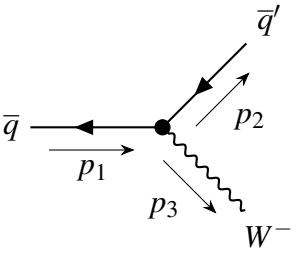

$$-\frac{ig_s}{2}\bar{q}_i^f \gamma^\mu \lambda_{ij}^a q_f^j G_\mu^a$$



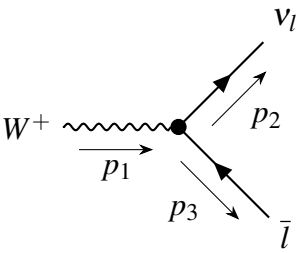

$$-ig_s f^{abc}\left(\partial^\mu G_a^\nu - \partial^\nu G_a^\mu\right) G_\mu^b G_\nu^c$$



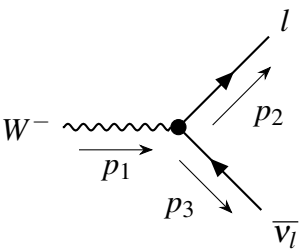

$$\frac{-ig_w V_{ij}}{2\sqrt{2}}\gamma^\mu\left(1+\gamma^5\right)$$




$$\frac{-ig_w V_{ij}}{2\sqrt{2}}\gamma^\mu\left(1-\gamma^5\right)$$




$$\frac{-ig_w}{2\sqrt{2}}\gamma^\mu\left(1+\gamma^5\right)$$




$$\frac{-ig_w}{2\sqrt{2}}\gamma^\mu\left(1-\gamma^5\right)$$

Fig. A.1 Couplings at vertices associated with top production and decay diagrams.

## Production

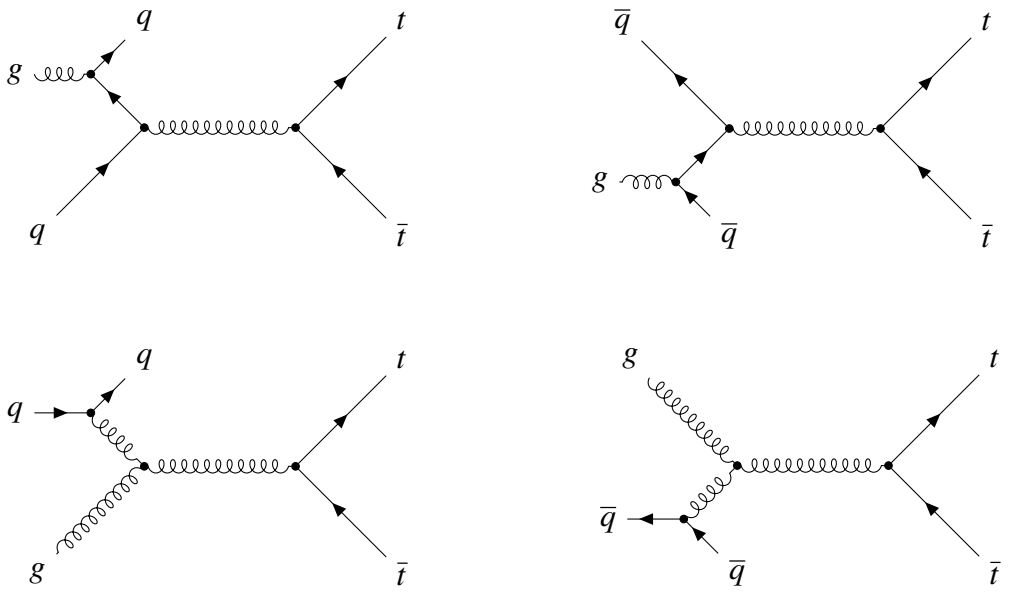



Fig. A.2 $qg$-initiated top pair production diagrams which produce interference with the NLO box diagrams leading to a positive asymmetry contribution.

# Appendix B

# Event reconstruction

This section provides plots and selections, produced by collaborators of the author, pertaining to jet reconstruction procedures referred to in the main text (Chapter 3).

## Energy response functions

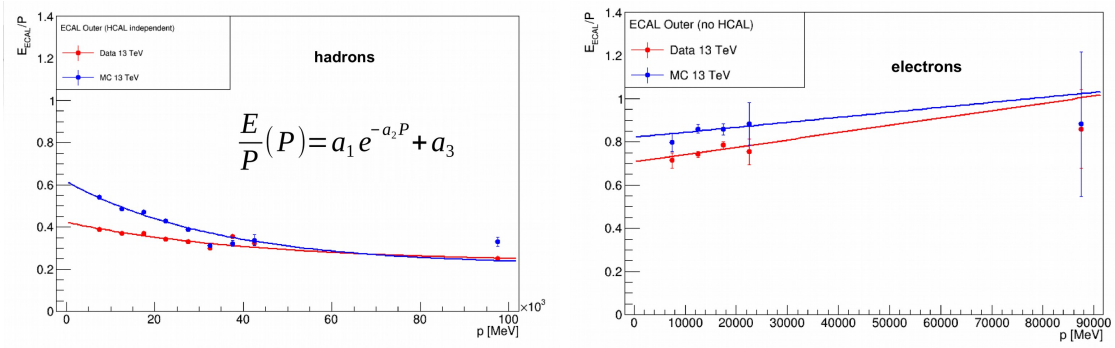



Fig. B.1 Fitted functions for data and MC ERFs for hadrons (left) and electrons (right) from studies performed by Lorenzo Sestini in Run II.

## Jet identification

A variety of variables were defined and investigated in this study; the optimal selection was found to apply to a subset of the variables, three of which are shown in Figure B.2. The requirements applied to reconstructed jets and dubbed JetID are listed below:

- Number of tracks in the jet ($nTrk > 1$);
- Maximum $p_T$ fraction carried by a single particle ($mpf < 0.75$);
- Maximum $p_T$ carried by a track ($mpt > 1.4\,\text{GeV}$);
- Charged particle $p_T$ fraction ($cpf > 0.06$).

The requirement on $nTrk$ is expected to reject pile-up jets; $mpf$ to suppress reconstructed jets originating from a high-$p_T$ isolated lepton; $mpt$ to suppress jets with high pile-up content and $cpf$ to suppress jets with high calorimeter noise.

## Jet energy corrections

The $C_{MC}$ is evaluated using simulation in bins of $log(p_T)$, $nPoint$, $\eta$, $\phi$, and $cpf$. A cubic function is fitted with respect to the logarithm of the uncorrected $p_T$ in each ($N_{PV}$, $\eta$, $\phi$, $cpf$)-bin (Figure B.3) in order to interpolate the correction factor to be applied to the four vectors of the jets, preserving their direction. The resultant jet energy resolution may be estimated in these same bins as a function of $p'_T$ (Figure B.3).

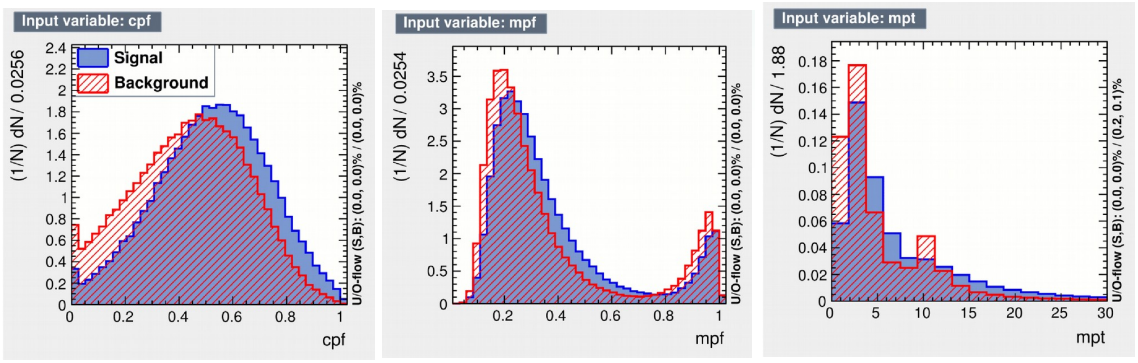



Fig. B.2 Signal and background distributions of reconstructed jet charged particle $p_T$ fractions ($cpf$, left), maximum fraction of $p_T$ from a single particle ($mpf$, centre) and the maximum track $p_T$ ($mpt$, right) from the Run II optimisation performed by Oscar Francisco.

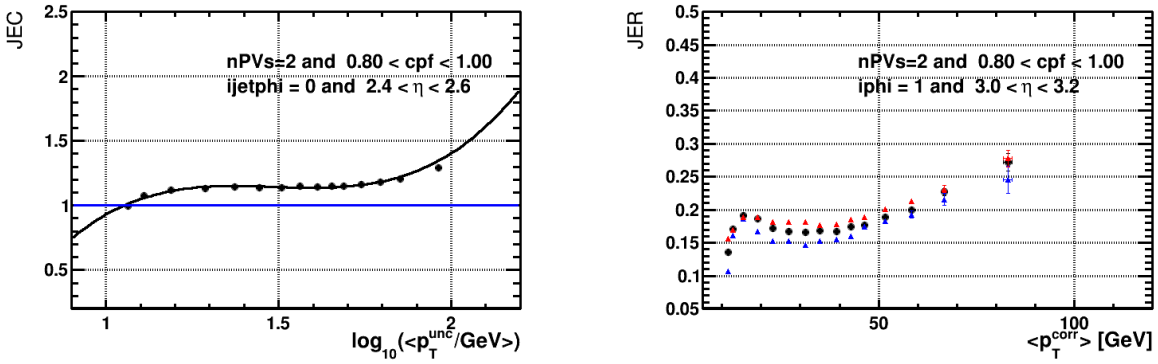



Fig. B.3 The correction factor for a specific bin in 4D as a function of uncorrected $p_T$ with fitted third order polynomial (left) and the resultant jet energy resolution from 3 estimators, Gaussian width (blue), RMS (red) and central 68% integral width (black), as a function of the corrected jet $p_T$ (right) from studies performed by Murilo Rangel.

## Secondary vertex tagging

A summary of the selection requirements imposed at each of the three stages is given in Table D.1 where: pseudo-lifetime, $\tau_z = \Delta z \cdot m/p_z$ where $\Delta z$ is the separation between the PV and SV in the $z$-axis; minimum radial two-body flight distance, min.$FD_T$; material distance ($MD$); corrected mass, $m_{\text{cor}} = \sqrt{m^2 + p_\perp^2} + p_\perp$, where $p_\perp$ defined as the momentum perpendicular to the direction of flight; the number of tracks in the vertex, $N_{trk}$; distance of closest approach for the constituent tracks ($DOCA$); the response of a NN (`GhostProb`) trained to reject reconstructed ghost tracks defined as containing less than 70% of the VELO clusters from the same simulated particle [40].

Table B.1 A summary of the requirements placed in order to SV-tag a jet from studies by Daniel Craik.

| tracks | (2,3)-body vertices | $n$-body vertices |
|---|---|---|
| `GhostProb` $< 0.2$ | $DOCA < 0.2\,\text{mm}$ | min.$FD_T < 15\,\text{mm}$ |
| $\chi^2_{IP} > 9$ | $\chi^2_V < 10$ | $\chi^2_{FD} > 32$ |
| $p_T > 500\,\text{MeV}$ | $400 < M < 5279.4\,\text{MeV}$ | $p_T(\text{SV}) > 2\,\text{GeV}$ |
| | $M_{cor} > 600\,\text{MeV}$ | $N_{trk} \leq 4$ |
| | $\sigma(M_{cor}) < 500\,\text{MeV}$ | $\tau_z < 10\,\text{ps}$ |
| | | $z < 200\,\text{mm}$ |
| | | $MD > 0.5\,\text{mm}$ |

Using samples enriched in $b$- and $c$-jets, selecting for events with a jet containing either a muon or a fully reconstructed $B$- or $D$-hadron, bottom and charm yields and light jet ($u$,$d$,$s$,$g$) background can be extracted to assess the performance of the tagger [76]. In events without an SV-tag, flavour yields were provided by fitting to $\chi^2_{\text{IP}}$ of the highest $p_T$ track in the jet. The performance of SV-tagging was compared between 2016 data and MC. Validation of the $b$-tagging efficiency is performed using $B \to (J/\psi)K$ events, where the $J/\psi$ is reconstructed through its decay to muon pairs and compared using 2016 MC and data. A conservative uncertainty of 10% applied to account for differences between data and simulation, most prominent at low jet $p_T$ [37].

# Appendix C

# Run II jet reconstruction

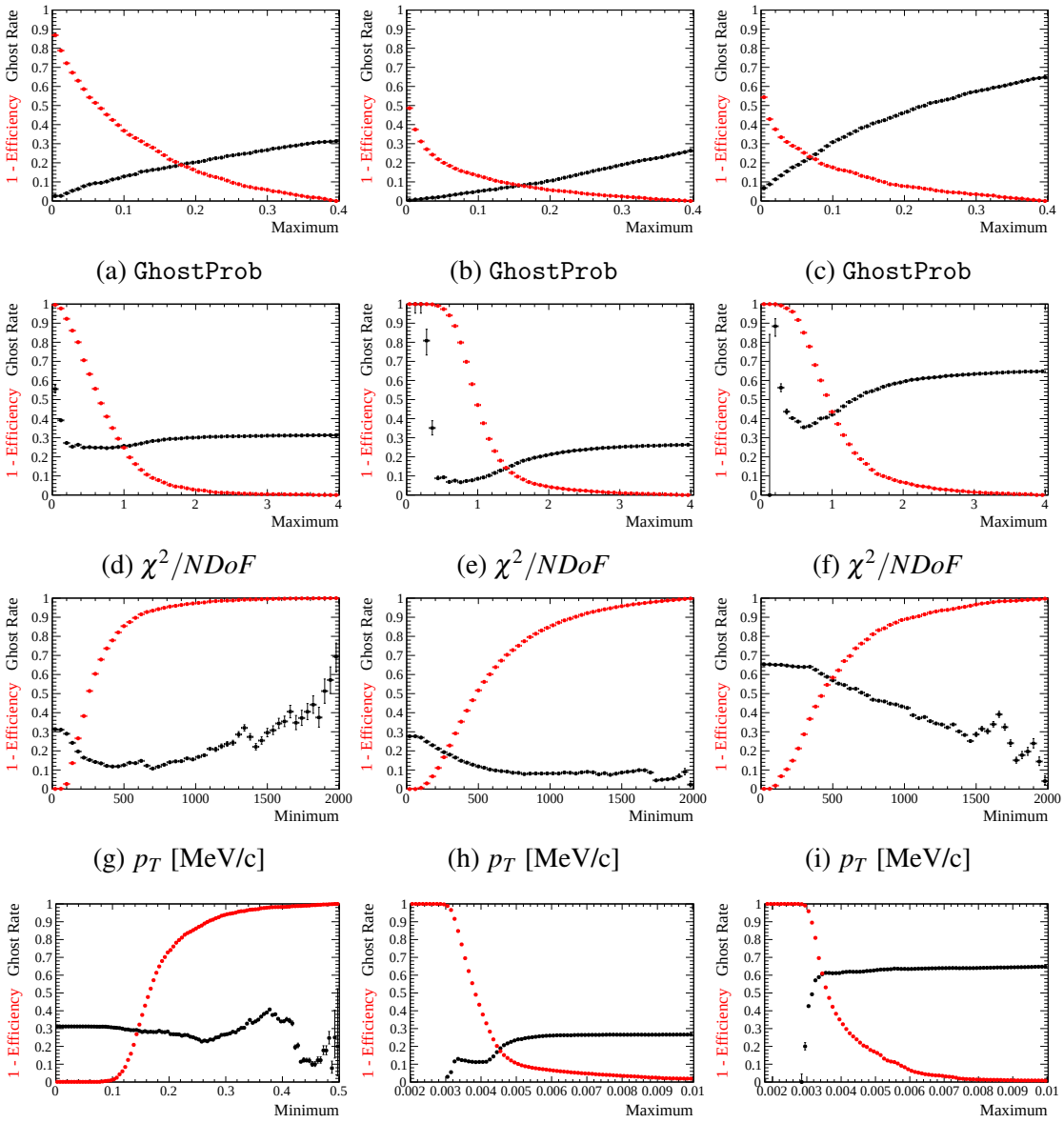
This section provides plots of ghost track rate and track reconstruction efficiency based on global cuts and track-type specific cuts from studies referred to in the main text (Chapter 4) as well as a demonstration of the reduced jet reconstruction performance outside of the chosen acceptance for the same chapter.

## Ghost tracks and track selection

Optimising input selection criteria based on jet reconstruction metrics is slowed by the processing of sufficiently large samples repeatedly with different input filters before jet clustering and MC truth matching. If the remaining ghost track contribution to fake jet reconstruction is significant, then a more rigorous optimisation procedure, purely based on the rejection of expected ghost content per track-type, may offer further improvement without the need for repeated sample processing. Figure C.1 demonstrates the impact of individual track-type based track selections on ghost rate and true track efficiency for each track-type.

It should be noted that the tracks used to calculate the ghost rates and inefficiencies displayed in Figure C.1 already include the global cut on `GhostProb`, hence the inefficiency reaching zero at the imposed maximum of 0.4 in plots (a), (b) & (c). The range of the cuts applied in plots (g)-(i) is increased, and for (j)-(l) is reduced, relative to their respective cuts from Table 4.1. The $\Delta p/p$ cuts for upstream tracks in (j) are also inverted. These scans of track performance imply that, while the existing $\chi^2/NDoF$ and NN output requirements are on a suitable scale to address ghost tracks, the minimum $p_T$ and curvature error requirements are intended to address other track-based effects, including detector noise and trajectories associated with large energy or position uncertainties. The impact of their inclusion would emerge at the jet reconstruction level either in terms of energy resolution or by seeding fake jets. As a result these criteria are carried forward into HLT and Turbo.

The ghost rate for downstream tracks, being approximately a factor of two higher than that of upstream or long tracks, was considered a culprit for seeding fake jets. Tightening the input requirement on the downstream tracks to a `GhostProb`$< 0.1$ would reduce the ghost rate to $\sim 30\%$, bringing it in line with upstream and long tracks, albeit with a 20% inefficiency $\subset (\texttt{GhostProb} < 0.4)$. However, having already imposed the input selection outlined in Table 4.1, the jet performance was investigated with downstream tracks excluded from particle flow inputs entirely, providing negligible improvements on fake jet reconstruction. Introducing charged input selection alone showed a more significant reduction in the fake rate for Turbo than HLT. This implied that a fake component, dependent on calorimeter information absent in Turbo, was not rejected in HLT jets. The $E_T$ threshold also shown in Table 4.1 for the hadronic calorimeter clusters replicates a requirement applied in the Std configuration and addresses the majority of the $\eta$ dependant fake jet component.




(a) `GhostProb` (b) `GhostProb` (c) `GhostProb`

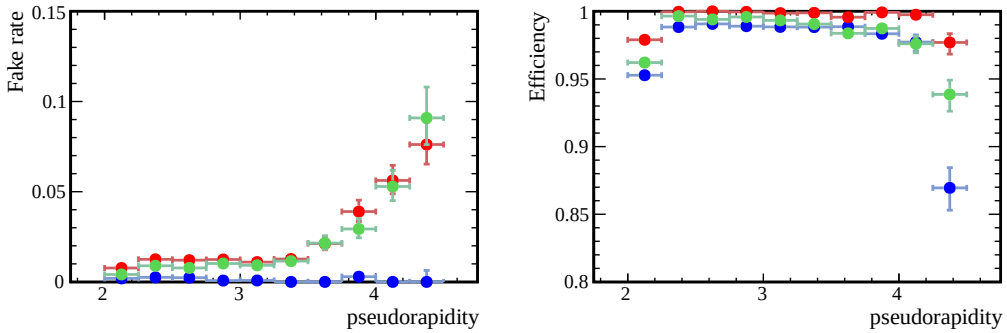
(d) $\chi^2/NDoF$ (e) $\chi^2/NDoF$ (f) $\chi^2/NDoF$

(g) $p_T$ [MeV/c] (h) $p_T$ [MeV/c] (i) $p_T$ [MeV/c]

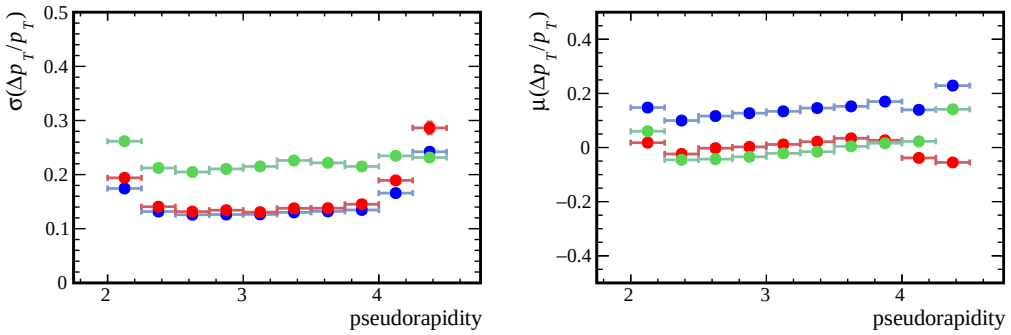
(j) $\Delta p/p$ (k) $\Delta p/p$ (l) $\Delta p/p$

Fig. C.1 Upstream (left), Long (centre) and Downstream (right) track 'ghost rate' (black) and 'inefficiency' (red) with cuts to be applied during pre-PF quality control using samples taken from 13 TeV upstream, long and downstream tracks respectively, where for each track $\chi^2/NDoF$ is the goodness of fit, $p_T$ is the transverse momentum and $\Delta p/p$ is the fractional error on the momentum.

## Jet fiducial acceptance




(a) $p_T$ (top) and $\eta$ (bottom) dependence of jet fake rate (left) and efficiency (right) for Std (red), HLT (blue) and TURBO (green) with Run II ERFs.




(b) $p_T$ (top) and $\eta$ (bottom) dependence of jet fake rate (left) and efficiency (right) for Std (blue), HLT (red) and TURBO (green) demonstrating edge effects of $\eta$ acceptance.

# Appendix D

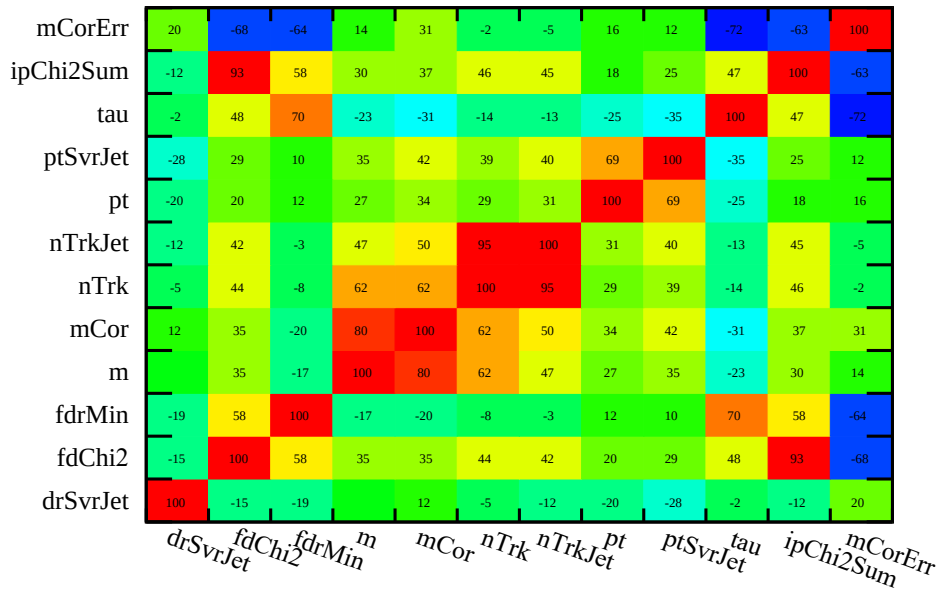
# Run II heavy flavour tagging

This section provides details of: the MVA training set correlation matrices and their relationships between classes; model responses and ROC curves for intermediate models produced during pre-processing test phases; the integrated and differential model performances prior to tuning; the hyper-parameter tuning stages and resulting differential performance at each training jet $p_T$ threshold; and the projection $\chi^2$ values in each of the MVA training variables compared between DNN and ALT fits of (20,50)$< p_T <$ 100 GeV jets for assessing MC to data agreement, each referred to in the main text (Chapter 5).

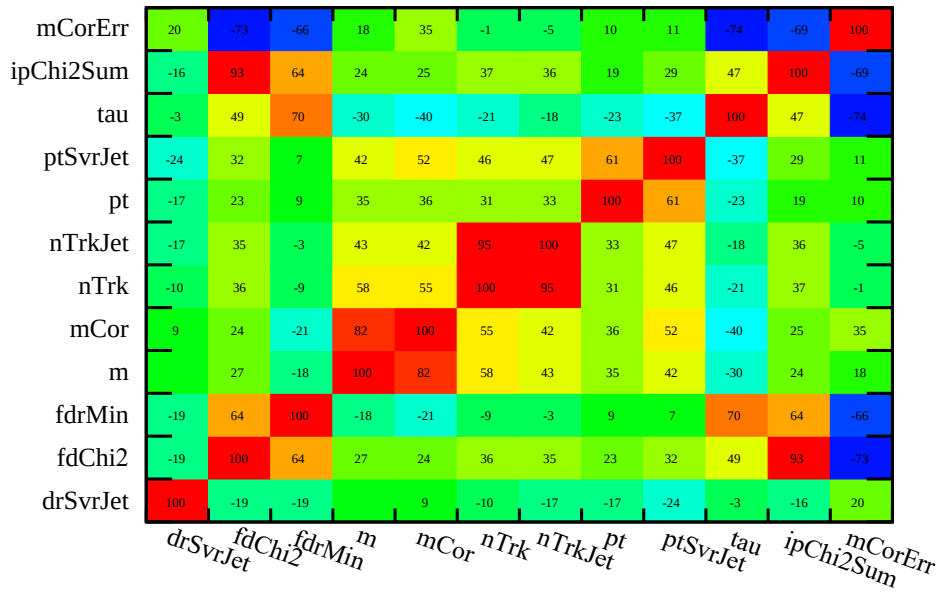
## Feature selection

The kinematic variables provided as inputs to the MVA-classifier training are shown with their correlation matrices in Figures D.1 & D.2. Using the binary classes used in Run I [76], Figure D.1 shows the correlations of $(bc|udsg)$ and Figure D.2 shows the correlations of $(b|c)$. The difference between the signal and background class correlation matrices of each classifier are in Figure D.3. Matrices of the difference in absolute correlation (or relationship strength, $|M^{a,b}_{cov}|$) are in Figure D.4. The matrices demonstrate that each variable shown offers sufficient information for discrimination across one or both classification problems. As a result, all those included in Figures D.1-D.3 are used in training for SV+jet models.

Figure D.3 shows that if the value of the difference between classes is positive (negative) then a positively (negatively) correlated signal (background) is dominant or there are opposite correlations between classes. Figure D.4 shows that if the value of the difference between classes is positive (negative) then the signal (background) correlation magnitude is dominant. When comparing variables $a$ and $b$, if $\Delta M^{a}_{cov} < 0 < \Delta M^{b}_{cov}$ then the dominance in same sign correlation strength between those variables is inverted for signal versus background. The same can be said for $\Delta|M^{a,b}_{cov}|$ irrespective of sign.

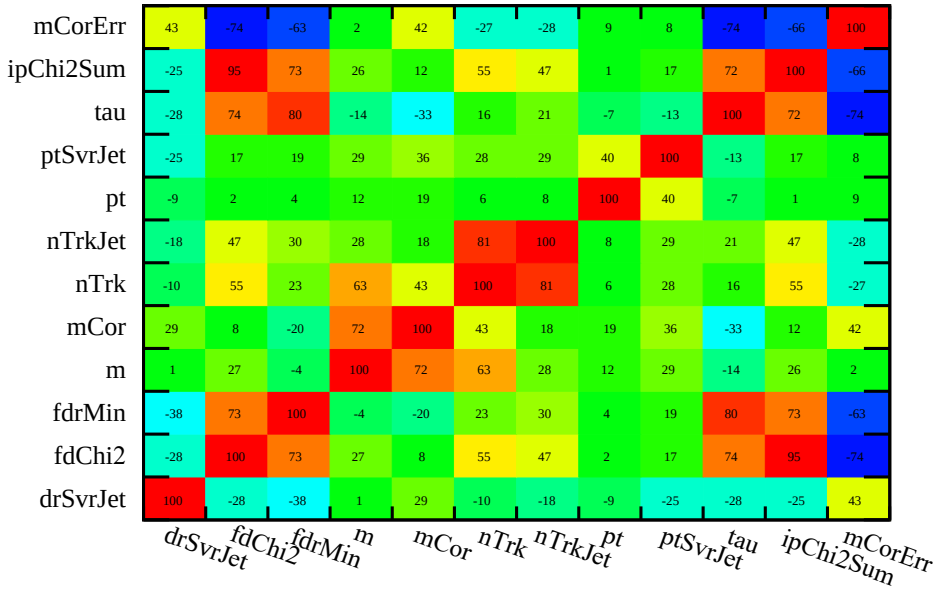



(a) Heavy versus light-jet classifier signal sample correlation matrix.

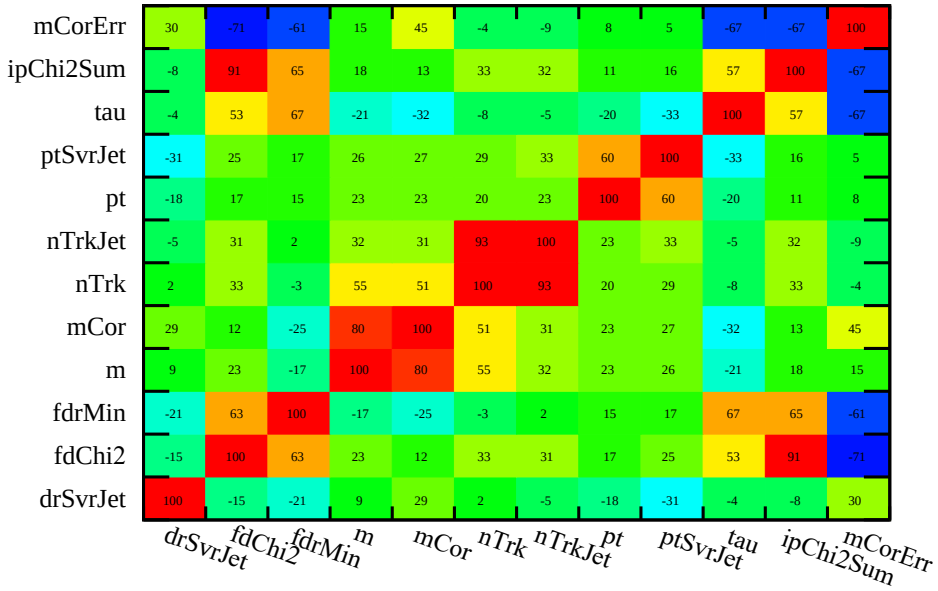



(b) Bottom versus charm classifier signal sample correlation matrix.

Fig. D.1 Training class signal sample correlation matrices for each binary classifier.

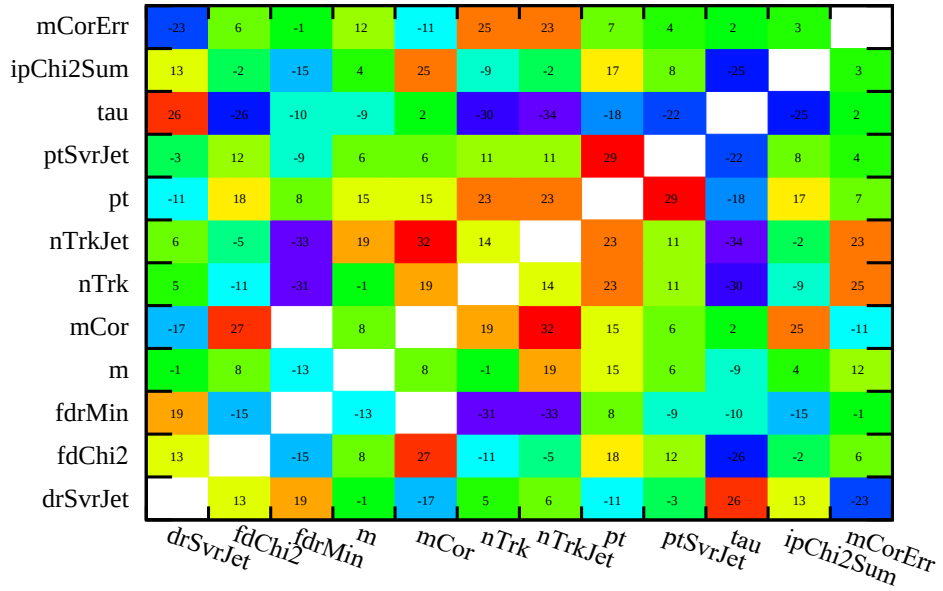



(a) Heavy versus light-jet classifier background sample correlation matrix.

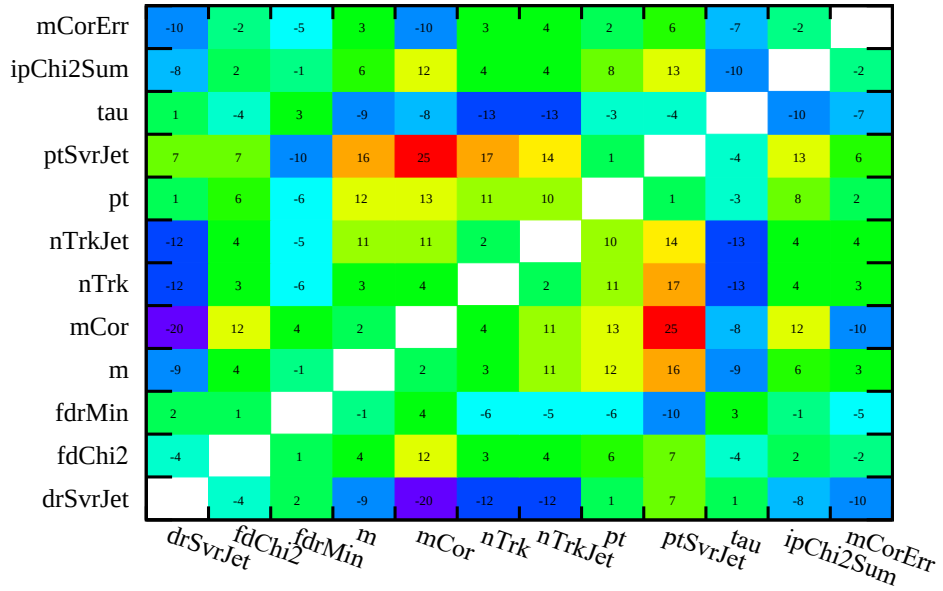



(b) Bottom versus charm classifier background sample correlation matrix.

Fig. D.2 Training variable covariance between MVA class di-jet MC samples.

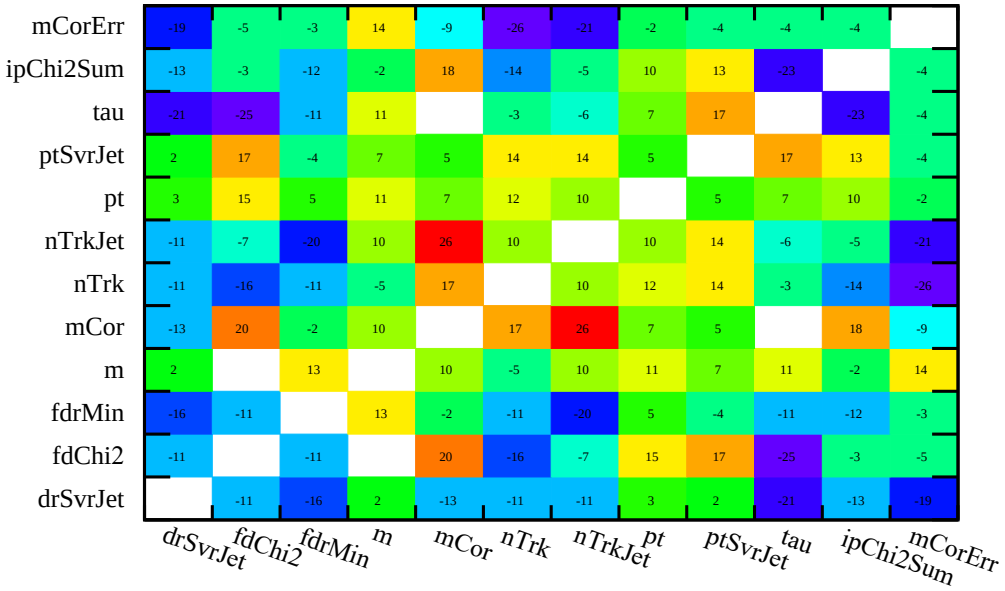



(a) Heavy versus light sample correlation matrix difference.

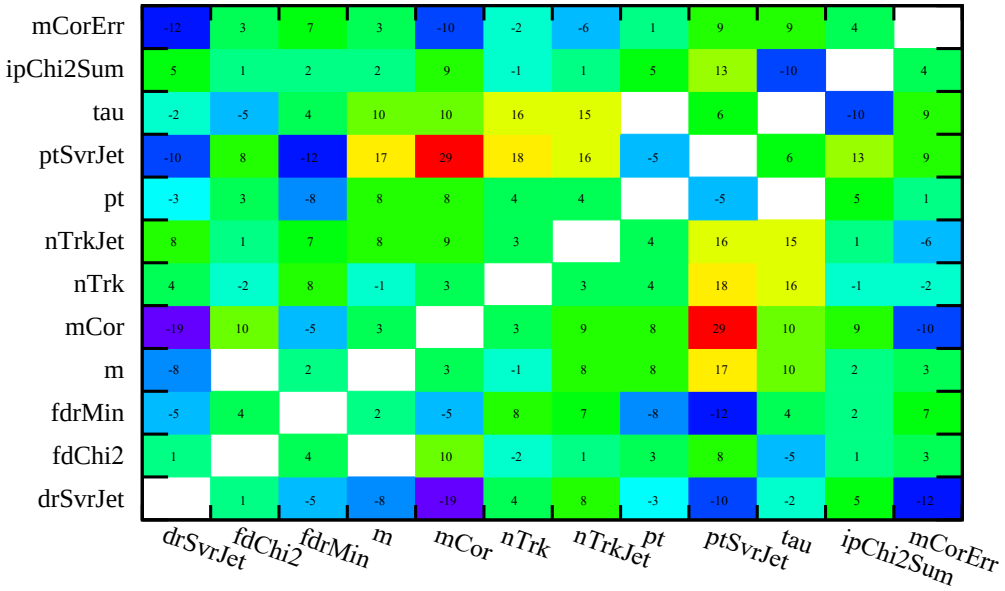



(b) Bottom versus charm sample correlation matrix difference.

Fig. D.3 Training sample class correlation matrix difference for each binary classifier.




(a) Heavy versus light sample absolute correlation matrix difference.




(b) Bottom versus charm sample absolute correlation matrix difference.

Fig. D.4 Training sample class absolute correlation matrix difference for each binary classifier.

## Sample pre-processing

Event weights scaled according to the ratio between the class normalisation can be applied to address this. ($HF \sim 780k : udsg \sim 6k$, $b \sim 520k : c \sim 260k$) When models were trained using these samples, artefacts were observed in the DNN responses (Figure D.5) and the shapes of each class while the BDT remained relatively unaffected. It was anticipated that, with an increased sample size for the light jet class in particular, discrimination could improve upon the (*udsg*)-rejection models even before testing alternative algorithms. However, the artefacts were replicated using a set of larger samples as demonstrated in Figure D.6, comparing the responses and ROC curves of light jet rejecting models, implying class imbalance is the culprit. Models trained with the inclusion of weighted samples (WS) or under-sampled HF-signal (US) are compared to models trained using neither technique in Figure D.5.

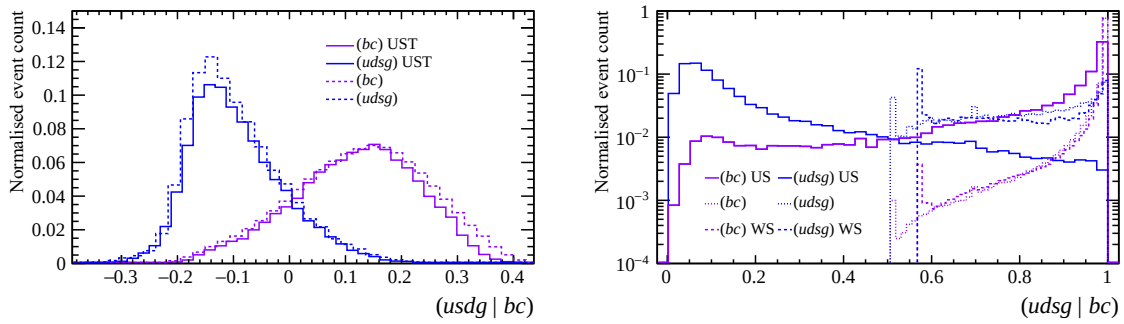



Fig. D.5 BDT (left) & DNN (right) responses of competing models each applying strategies for imbalanced training samples: default training (dotted line), under-sampled (US) training (solid line) and weighted-sample (WS) training (dashed line); for class templates for heavy flavour (purple) and light jets (blue) where default training (solid line) is shown overlaid with under-sampled (US) training (dashed line).

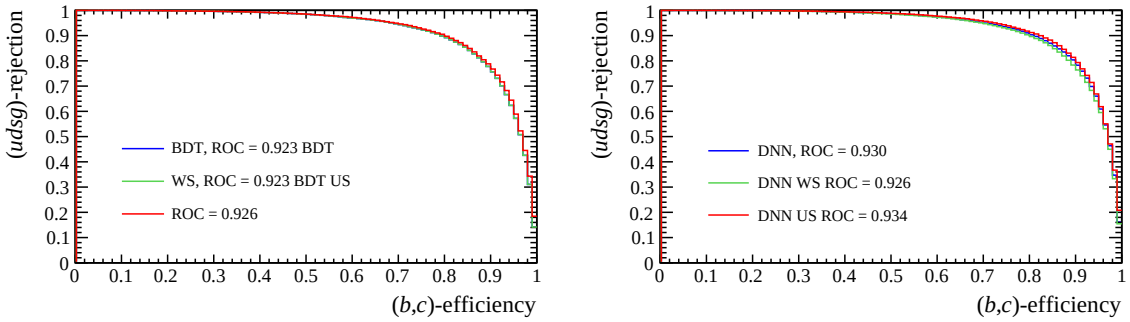



Fig. D.6 ROC curves for light jet rejection BDT (left) & DNN (right) models applying strategies for imbalanced training samples: default (AUC=92.3% & 93.0%) training (blue), weighted-sample (AUC=92.3% & 92.6%) training (green) and under-sampled (AUC=92.6% & 93.4%) training (red).

It was shown that using the logarithm of variables with sharp, skewed peaks improved the performance of the TMVA BDT models. Some of the transformed variables from the training sample are shown in Figure D.7 split by flavour. Figure D.8 compares the ROC curves of these models with and without including log-transformations. As might be expected based on the continuous weights and activation of the NNs, there was no impact on the performance of each equivalent DNN. Despite this, even with the improvements offered to the BDT through this additional pre-processing step, +7% for $(udsg|bc)$ and +3% for $(c|b)$, the DNN maintains a higher ROC curve integral (Figure D.8).

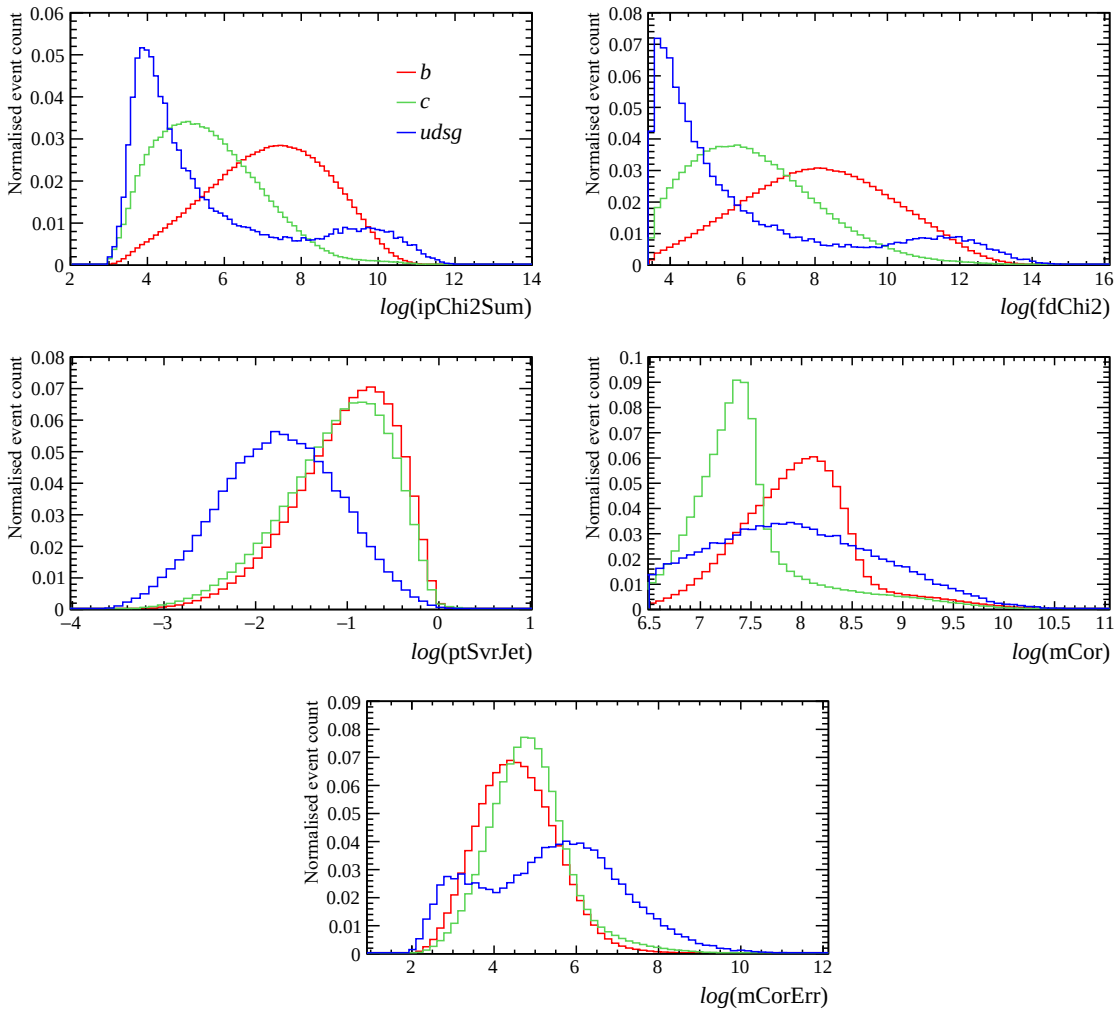



Fig. D.7 Examples of logarithm transformed variables, normalised by flavour in di-jet MC sample.

The improvement of the BDT may be due to grid search method by which TMVA tests cuts at each tree split. Transforming these variables may increase sampling density over regions of interest. For model development and comparison, these variables will remain transformed such that the benchmark set by the Run II retrained TMVA BDTs includes this improvement.

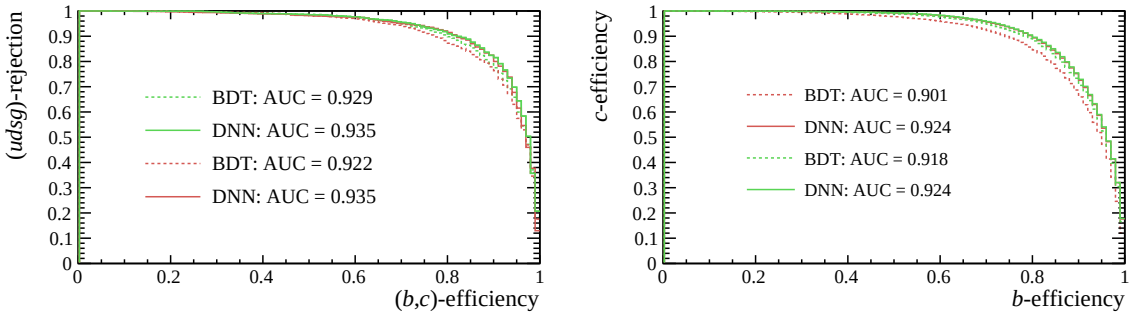



Fig. D.8 Receiver operating characteristic (ROC) curves for competing models in light jet rejection (left) and b vs. c distinction (right) training sets. The dashed line indicates BDT models and solid DNNs, where red lines indicate the use of default variables whereas green lines indicate use of $log(x)$ selected asymmetric random variables.

## Model comparison

Figure D.9 shows the flavour templates for the Run I BDTs and Run II BDTs & DNNs templates, demonstrating improved discrimination from the Run II models as expected having trained on HLT jets in Run II MC. The $(c|b)$ BDT shows similar shapes for each class as the Run I equivalent, even replicating the shouldered peak to the $b$ distribution [76] as demonstrated in Section 5.2.5 where the Run I BDT is compared to the Run II BDT and DNN normalised responses to training data split by flavour. The plots in From Figures D.10 are analogous to ROC curves with the axis swapped and a downward scaling in the signal efficiency axis as discussed in the main text. Figure D.11 presents the ROC AUC of the models as functions of jet kinematics.

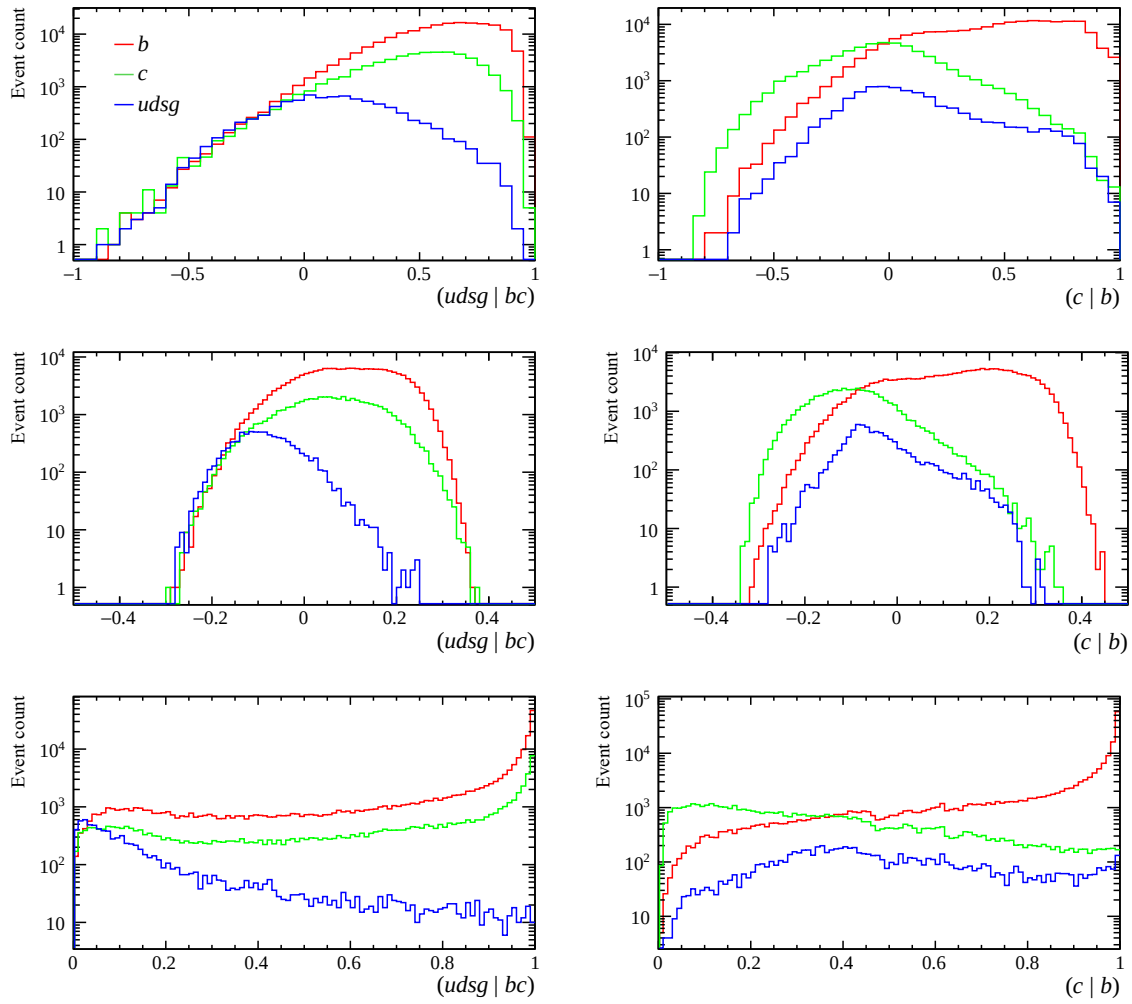



Fig. D.9 Light rejection (left) and $b$-tagging (right) Run I BDT (top), Run II BDT (centre) and DNN (bottom) responses for Run II Z+jet MC of SV-tagged light (blue), $c$- (green) and $b$- (red) jets.

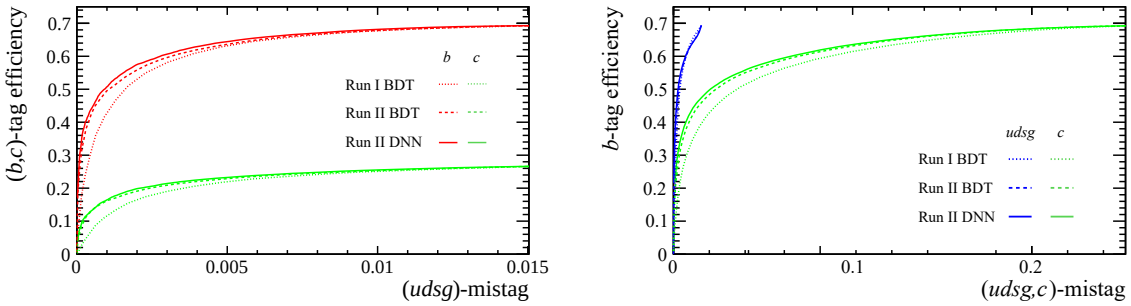



Fig. D.10 ROC curves where light rejection (left) tagging $(b,c)$-jets (red, green) and $(b)$-tagging (right) for (light,$c$)-mistag (blue, green) for Run I BDT (dotted), Run II BDT (dashed) and DNN (solid).

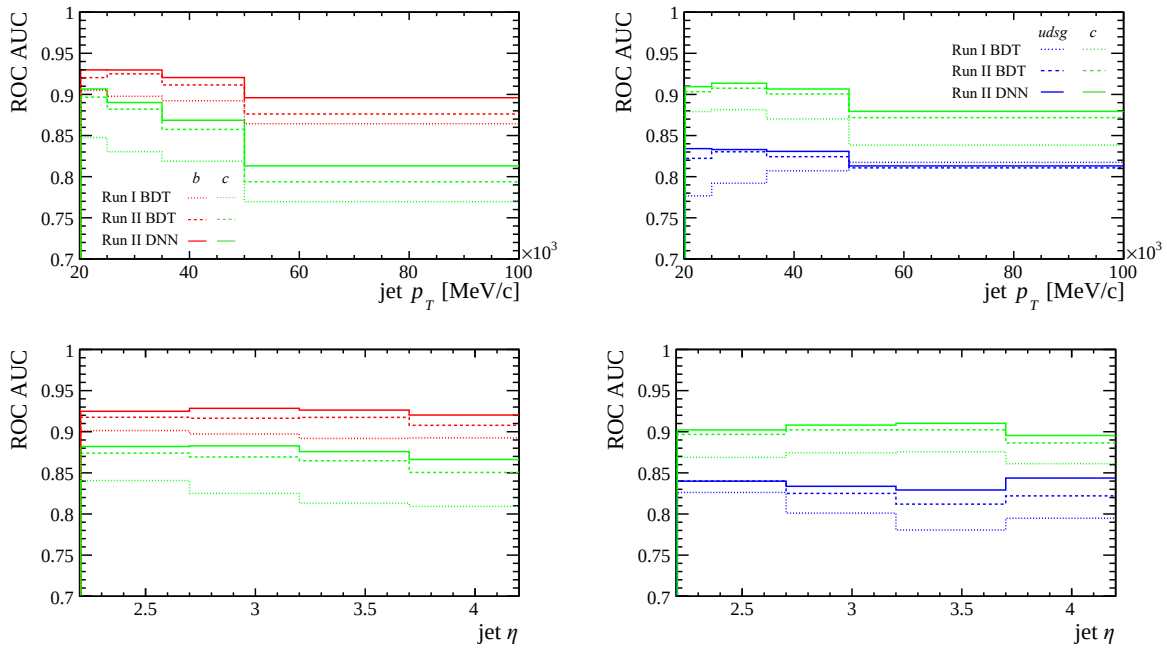



Fig. D.11 ROC AUC as functions of jet $p_T$ (top) and $\eta$ (bottom), for light rejection or ($b$)-tagging performance of Run I BDTs, Run II BDTs and Run II DNNs split by (light,$c$)-mistag.

## Hyper-parameter tuning

For DNN models which did not fail (AUC $< 80\%$), the average ROC AUC (%) is projected as a function of training sample compared to a BDT trained on the same split (Figure D.12). Figure D.12 demonstrates a dependence of the light-jet rejecting model AUCs upon the choice of training sample fraction, with variation $\sim 0.5\%$ consistent between BDT and DNNs. This relationship may be due to the smaller sample sizes used with an under-sampled approach. The HF identification models seem to show a weaker dependence on Test:Train split, with a change in performance $< 0.1\%$ independently varying between BDT and DNNs with no apparent pattern. The hyper-parameter grids in Figure D.13 imply that light rejecting models are improved using a simpler network structure than the HF discriminating models. The models used to proceed for each threshold are the best performing models out of the grid-scans of each classification problem. The ROC AUCs in Figure D.14 imply that BDT performance above 50 GeV are improved by O(0.1%), while incurring more significant losses across lower $p_T$ bins, by increasing the training set $p_T$ threshold.

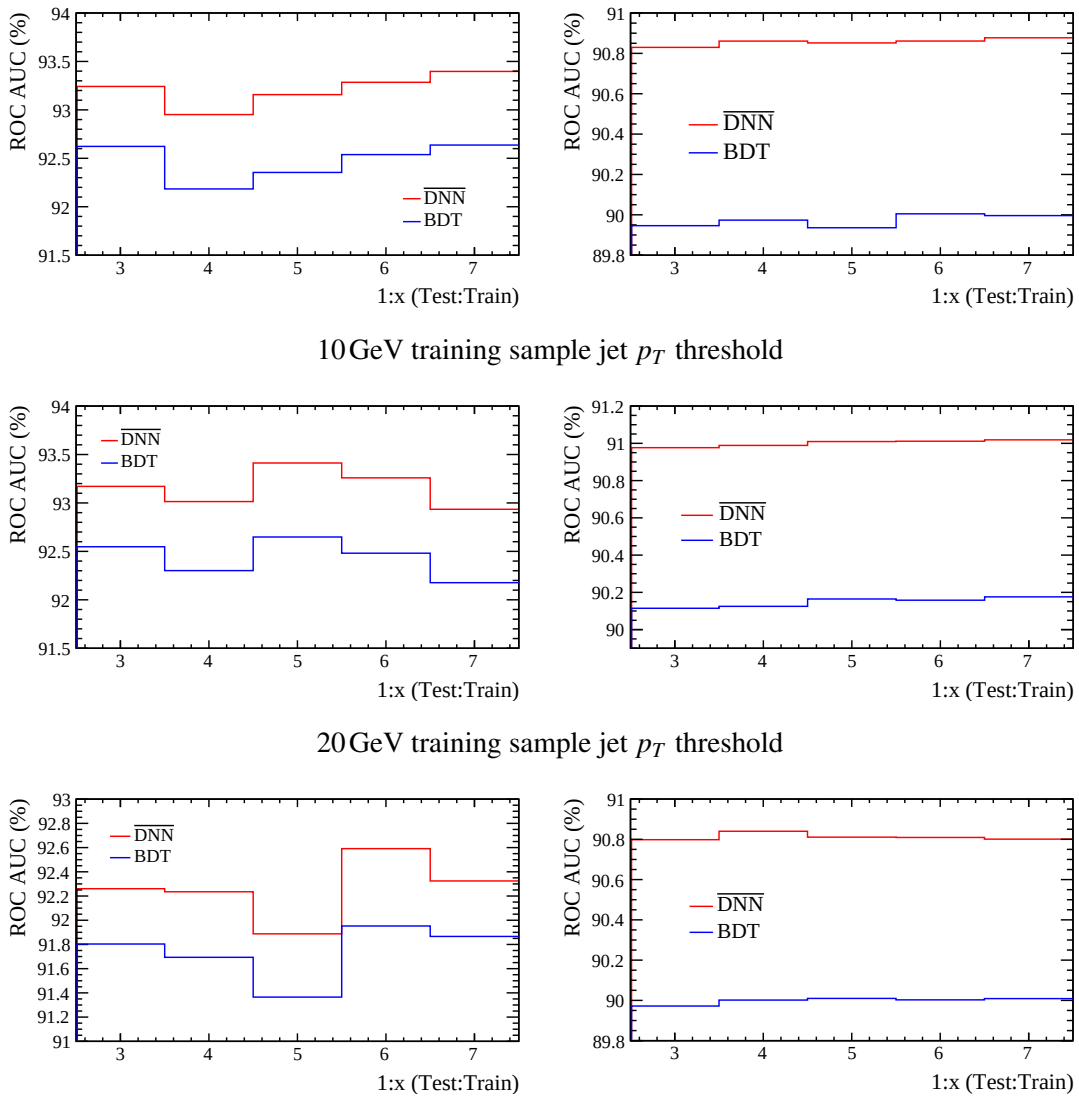
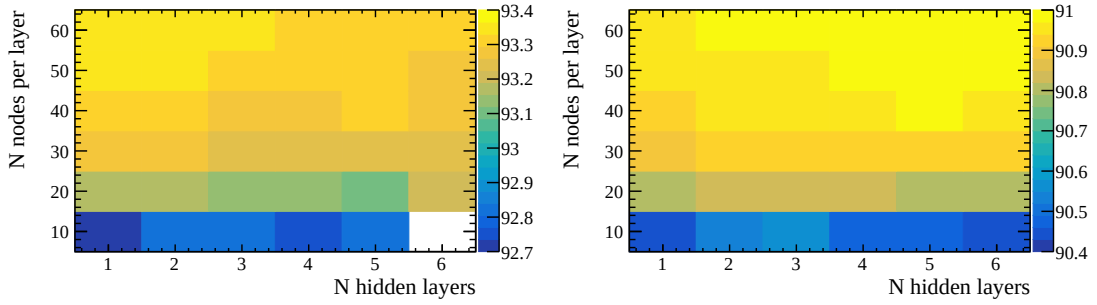
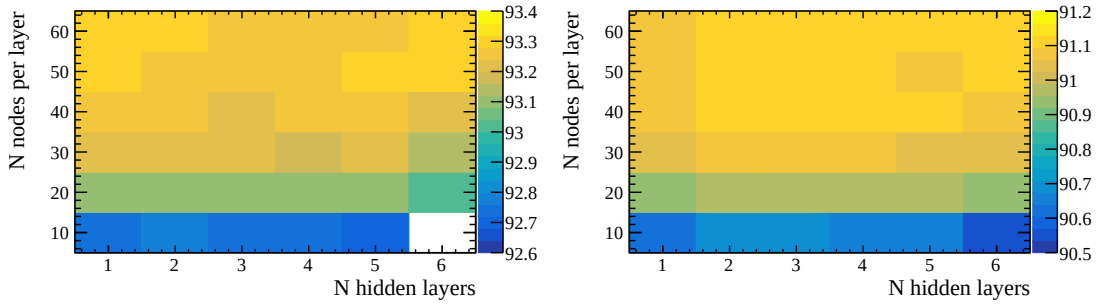


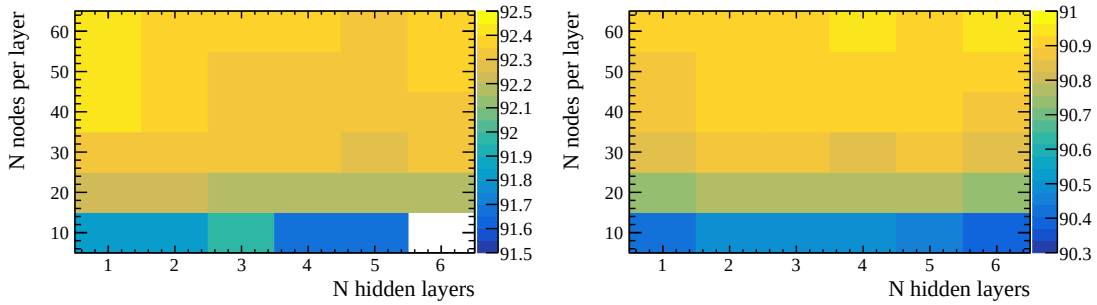

Fig. D.12 Light rejection (left) and $b$-tagging (right) BDT and average DNN ROC AUCs as a function of training sample fraction, trained on samples with $10, 20, 50$ GeV jet $p_T$ thresholds.




10 GeV training sample jet $p_T$ threshold




20 GeV training sample jet $p_T$ threshold




50 GeV training sample jet $p_T$ threshold

Fig. D.13 Light rejection (left) and $b$-tagging (right) learning rate averaged DNN ROC AUC in grids of hidden layers versus nodes per layer, trained on samples with $(10, 20, 50\,\text{GeV})$ jet $p_T$ thresholds.

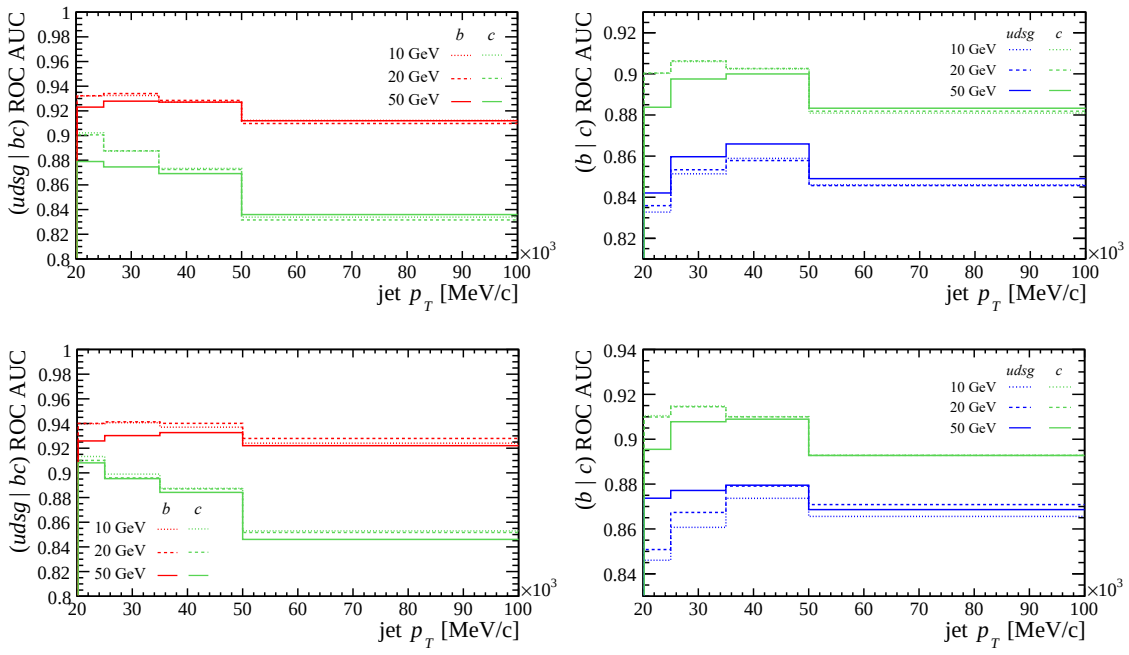



Fig. D.14 ROC AUCs for BDT (top) and DNN (bottom) as function of jet $p_T$ for a models with training sample jet $p_T$ thresholds of 10 GeV (dotted), 20 GeV (dashed) and 50 GeV (solid) for light rejection (left) for $(b,c)$-jet (red, green) and $b$-tagging (right) for (light, c)-mistag (blue, green).

Table D.1 The $\chi^2$-values for the projections of 2D fits in the DNN training axes.

| Variable projections | 20 GeV DNN | 20 GeV ALT* | 50 GeV DNN | 50 GeV ALT* |
|---|---|---|---|---|
| mCor* | 24.0 | 19.2 | 6.37 | 5.4 |
| ntrk* | 3.1 | 5.2 | 7.2 | 5.9 |
| mCorErr | 19.8 | 21.8 | 10.0 | 10.6 |
| nTrkJet | 1.5 | 3.17 | 6.1 | 5.9 |
| m | 22.6 | 24.6 | 6.9 | 7.9 |
| tau | 21.8 | 21.7 | 7.2 | 7.4 |
| pt | 9.1 | 18.7 | 3.9 | 5.0 |
| ptSvrJet | 17.1 | 13.3 | 7.1 | 5.1 |
| fdrMin | 23.6 | 18.1 | 11.5 | 7.4 |
| drSvrJet | 14.0 | 11.9 | 4.4 | 3.0 |
| fdChi2 | 22.0 | 32.5 | 11.6 | 14.0 |
| ipChi2Sum | 22.6 | 23.9 | 11.6 | 14.7 |

## $(\eta, p_T)$-binned templates

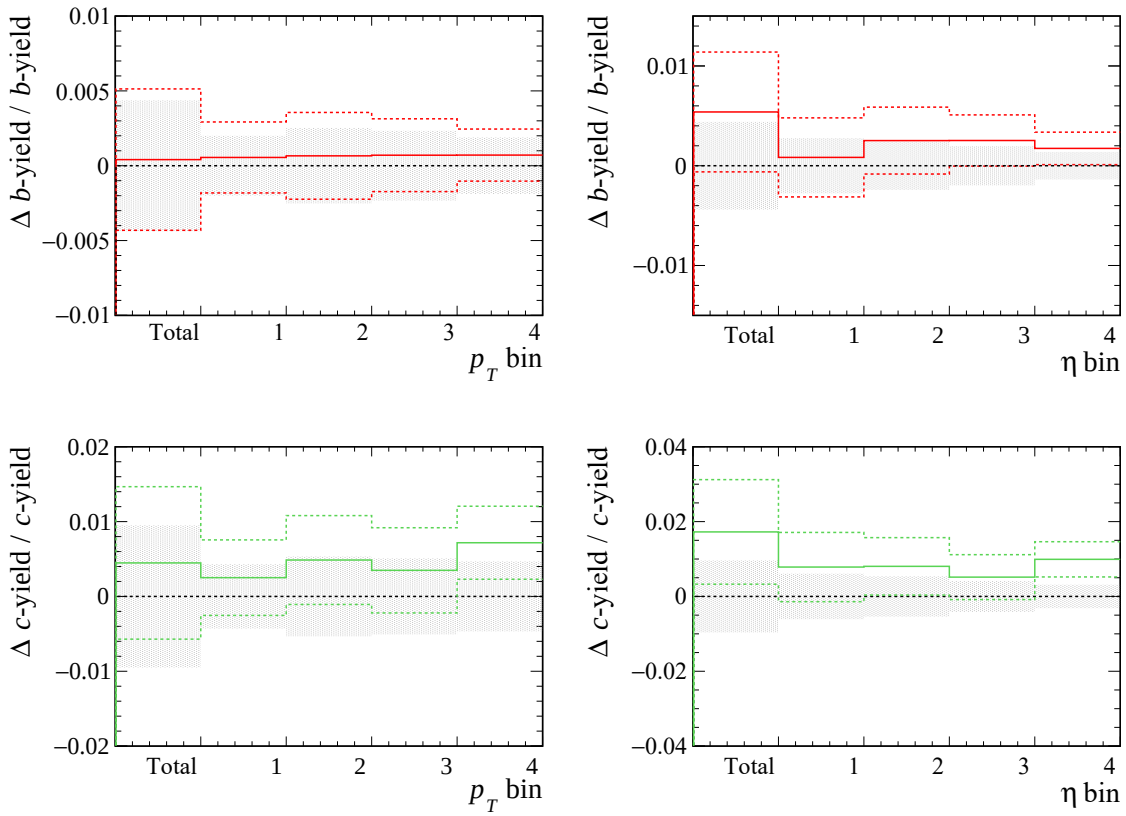



Fig. D.15 Consistency checks between ($b$,$c$)-yields in $p_T$ ($\eta$) regions from integrated DNN fits (grey band) and the sum of yields fitted in $\eta$ ($p_T$) bins respectively across that region (coloured band), minus the integrated fit yield and divided by it, each displaying $1\sigma$ fit uncertainty, where bin number corresponds to the $\eta$ and $p_T$ boundaries $[2.2, 2.7, 3.2, 3.7, 4.2]$ and $[20, 25, 35, 50, 100]$ GeV respectively and the zeroth bin represents the fit(s) to the total in that x-axis.

The DNN fit $b$-yield from $\eta$ bins was within 0.1% of integrated fits in both the full $p_T$ range and individual bins, with the $c$-yield difference $\mathcal{O}(0.1\%)$. For the $b$- and $c$-yield, the integrated fits and the sum of $p_T$ binned fits differ by $\mathcal{O}(0.1\%)$ and 1-2% respectively. Fitting to $\eta$ ($p_T$) binned data shows an increase in both $(b,c)$-yield across all $p_T$ ($\eta$) regions. For the ALT fit the sum of the fitted binned yields is consistently higher than fit over the total in a given range.

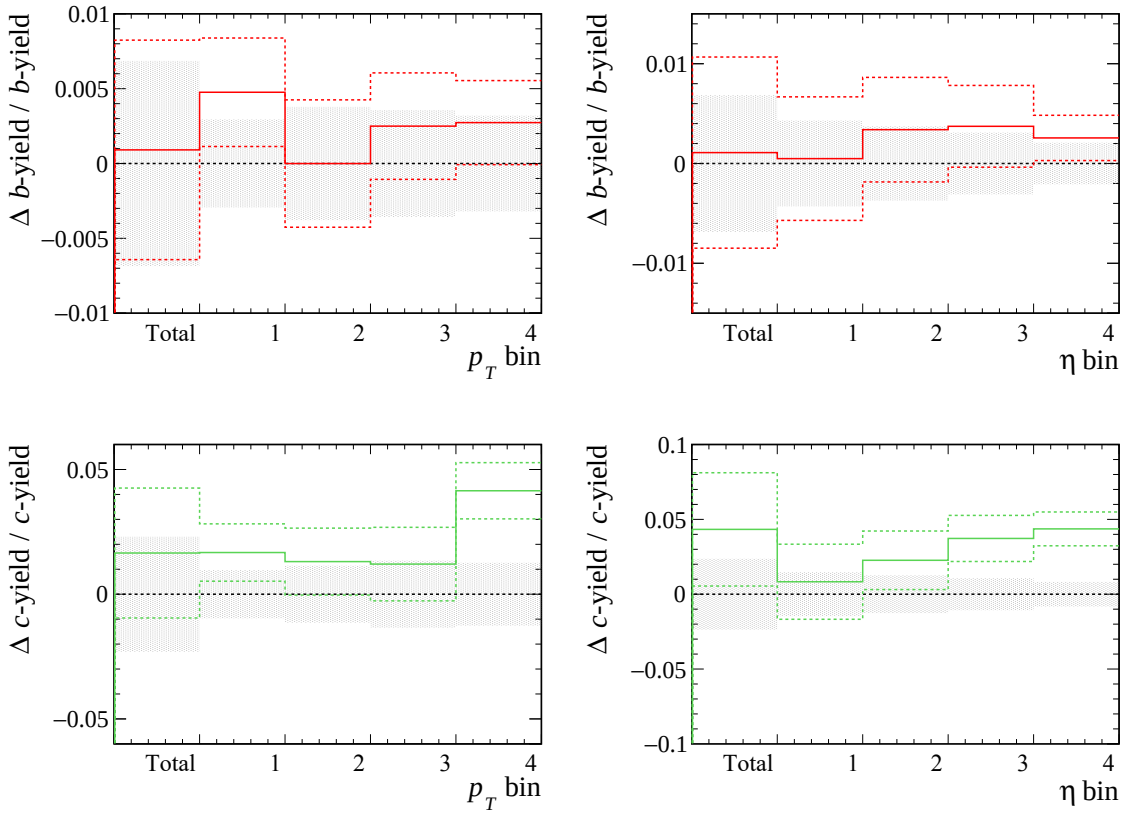



Fig. D.16 Consistency checks between ($b$,$c$)-yields in $p_T$ ($\eta$) regions from integrated ALT fits (grey band) and the sum of yields fitted in $\eta$ ($p_T$) bins respectively across that region (coloured band), minus the integrated fit yield and divided by it, each displaying $1\sigma$ fit uncertainty, where bin number corresponds to the $\eta$ and $p_T$ boundaries $[2.2, 2.7, 3.2, 3.7, 4.2]$ and $[20, 25, 35, 50, 100]$ GeV respectively and the zeroth bin represents the fit(s) to the total in that x-axis.

# Appendix E

# Top quark cross-section measurements

## Decay channels

Table E.1 Inclusive $t\bar{t}$ cross-section channels within the LHCb acceptance with quoted uncertainty accounting for variation of scale, PDF and shower modelling uncertainty [29] and the mean Bjorken-$x$ probed by partially reconstructed $t\bar{t}$ events [96].

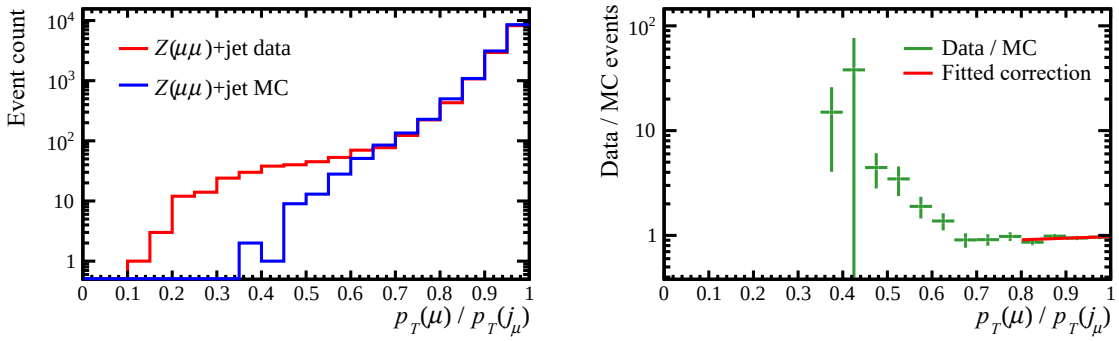
| $d\sigma$(fb) | 7 TeV | 8 TeV | 14 TeV | $< x >$ |
|---|---|---|---|---|
| $lb$ | $285 \pm 52$ | $504 \pm 94$ | $4366 \pm 663$ | 0.295 |
| $lbj$ | $97 \pm 21$ | $198 \pm 35$ | $2335 \pm 323$ | |
| $lbb$ | $32 \pm 6$ | $65 \pm 12$ | $870 \pm 116$ | 0.368 |
| $l^+l^-b$ | $19 \pm 4$ | $39 \pm 8$ | $417 \pm 79$ | 0.348 |

## Background yields

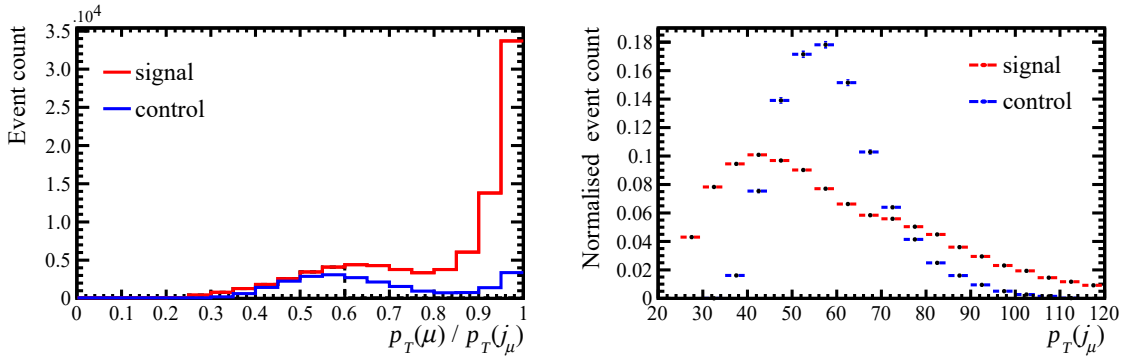
For solving Equation 6.1 for ABCD, two methods were used to extract the value of $N_D^{\text{sig}}$: one analytical, the other substituting the solution for $N_D^{\text{sig}}$ (initially assuming $c_{A,B,C} = 0$) into $N_D^{\text{sig}'}$ from Equation 6.1 recursively, restricting the convergence to values $N_D^{\text{sig}} > 0$ and $N_D^{\text{bgd}} \geqslant 0$.

A method that involves fitting to muon isolation using corrected signal MC and data-driven background templates was also explored. This had provided a cross-check for ABCD [111] and the central yield estimate [30, 31] in analogous studies. Limited understanding of the observed structure to the background contributions prevented sufficient improvements to the background muon isolation template to take advantage of this method.

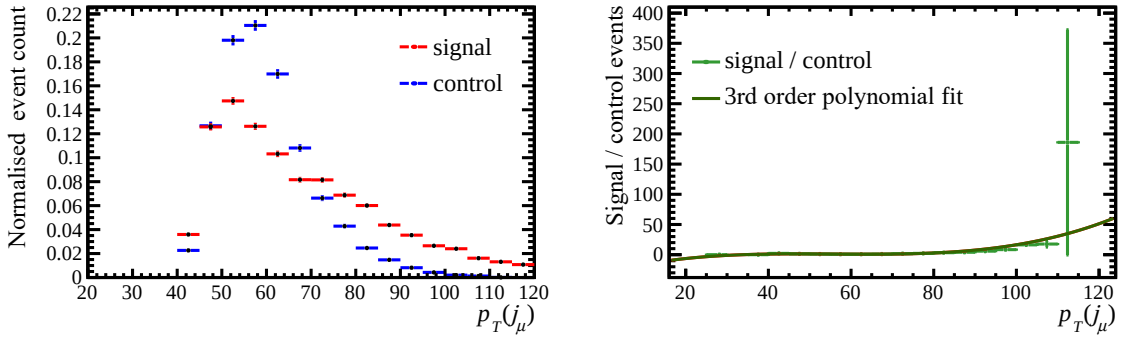
Extracting EW and associated jet production from muon isolation distributions was employed in RunI for studies of both W+jet and top production [30] [31] [93]. The $W \rightarrow \mu(\nu)$+jet and $Z \rightarrow \mu(\mu)$+jet shapes may be provided through MC corrected using ($Z \rightarrow \mu\mu$)+jet MC and data (Figure E.1a). Multi-jet QCD events are expected to lie further down the muon isolation range and may have been predicted in shape through data driven methods (Figures E.1b & E.1c). The multi-peaked structure in $p_T(j_\mu)$ in the signal region exaggerated in the anti-isolated data is not well understood and may be responsible for difficulties with producing a reliable prediction for the background shape in data.




(a) Correction to EW isolation template shapes taken from a fitted ratio between $Z \to \mu\mu$ MC and data across high isolation range assumed free of contamination

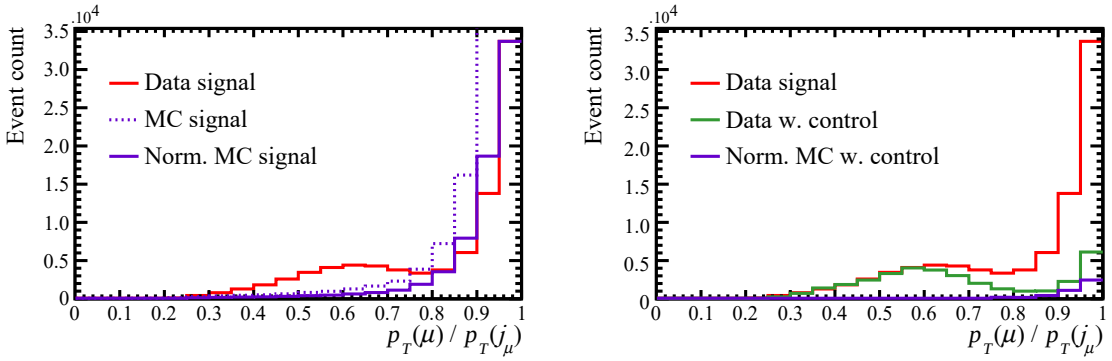



(b) $\mu$-isolation and $p_T(j_\mu)$ in data where events passing $p(j_\mu + j_b)_T > 20\,\text{GeV}$ constitute signal and $p(j_\mu + j_b)_T < 15\,\text{GeV}$ is regarded as the control region

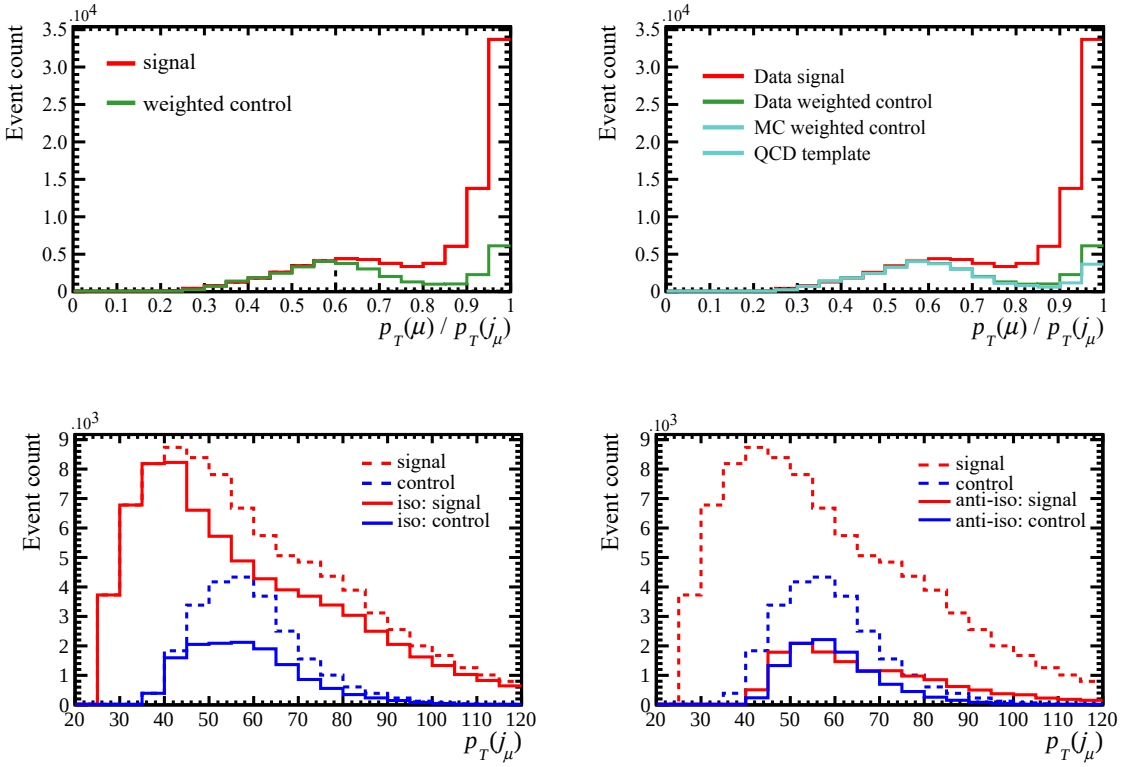



(c) $p_T(j_\mu)$ in anti-isolated data ($p_T(\mu)/p_T(j_\mu)$) used as the weights to estimate the multi-jet QCD content in the signal region by applying them to the control in $p_T(j_\mu)$ across the full isolation range

Fig. E.1 Initial stages of retrieving muon isolation templates for fitting to $p_T$-imbalanced data resemble past implementations in terms of the expected shapes of $Z$ MC to data correction and signal and control regions in data, while further investigation into the data driven weights reveal poorly understood structure in $p_T(j_\mu)$.




(a) Reweighted control contribution expected from W MC normalised assuming 100% signal in data with isolation $> 0.95$ and, below, the background template resulting from the subtraction.



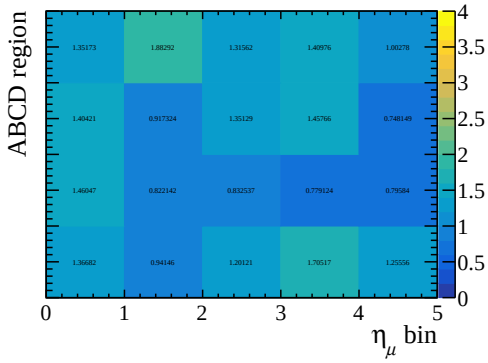

(b) Breakdown of the isolated (left) and anti-isolated (right) contributions to signal and control region $p_T(j_\mu)$ profiles where the secondary peaked structure in signal extends the tails

Fig. E.2 Latter stages of producing isolation templates including refining the background estimate by accounting for contamination by EW signal in the control region going into the data driven background, where normalised to the assumed pure signal content in the last isolation bin in data, it may be subtracted reducing the background template peak at high isolation, as well as a breakdown of the unexpected shapes in $p_T(j_\mu)$ distributions.

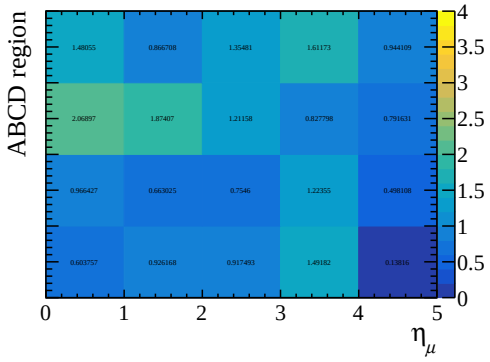
The background shape is estimated across the muon isolation range using control events re-weighted by the ratio of anti-isolated signal to anti-isolated control in $p_T(j_\mu)$. The signal contamination in the control region going into this calculation are accounted for by producing a background estimate from the corrected $p_T$-balanced W MC events, initially normalised to the events in isolation > 0.95 assuming pure signal but later included as a negative contribution to the isolation fit itself with the value normalised to the fitted yield of the total W MC contribution. In each isolation bin in data the flavour tagging procedure described in Chapter 5, and applied analogously to RunI analyses [31] [30], allows the SV-tagged data to be normalised bin by bin to the HF-yield of interest.

The procedure for accurately modelling the background shape was insufficient and the fits were found to be failing regardless of choice of signal template. Fitting $t\bar{t}$ and $W$ templates independently was also carried out but was considered too sensitive to the poorly understood underlying event to provide understanding of uncertainties relating the relative top and $W$-yields (while simultaneously suffering from the same issues with background fitting).

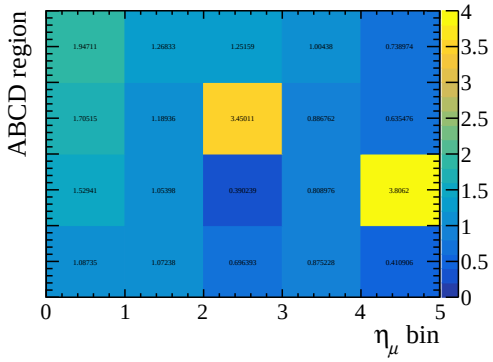
## Heavy flavour yields



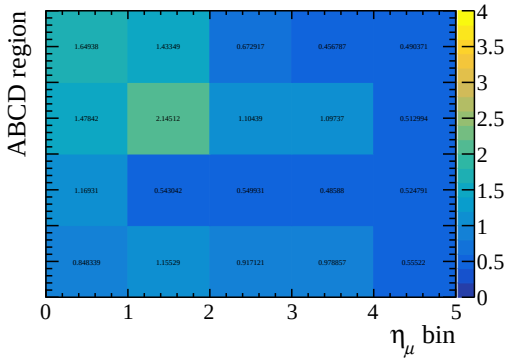

(a) MVA fit to $\mu^+$+SV-events




(b) MVA fit for $\mu^-$+SV-events




(c) Alternate fit to $\mu^+$+SV-events




(d) Alternate fit to $\mu^-$+SV-events

Fig. E.3 $\chi^2$ per degree of freedom for the 2D flavour template fits with three floating parameters corresponding to the relative normalisation of the $b$, $c$ and light components binned in muon $\eta$ for each ABCD region, the empty bins are those in which the $\chi^2$ calculation produces infinities.