



# CERN Open Data Portal

Tibor Šimko  
CERN IT

IPPOG - 4 October 2018

# CERN Open Data portal

- `opendata.cern.ch` launched in November 2014
- LHC collaboration data policies
  - restricted → embargo period (~5 years) → open
- over 1.5 Petabytes of open particle physics data
  - datasets, software, VMs, configuration, documentation, ...
- users
  - education: general public, high-school students, masterclasses
  - research: data scientists, physicists

*Developed by CERN-IT and CERN-SIS  
in close collaboration with Experiments*



# CERN Open Data portal

opendata  
CERN

About

Explore more than **1 petabyte**  
of open data from particle physics!

Start typing...

Search

search examples: [collision datasets](#), [keywords:education](#), [energy:7TeV](#)

**Explore**

- [datasets](#)
- [software](#)
- [environments](#)
- [documentation](#)

**Focus on**

- [ATLAS](#)
- [ALICE](#)
- [CMS](#)
- [LHCb](#)

Get started

<http://opendata.cern.ch/>

# Information organisation

open data CERN Search About

## Mu primary dataset in AOD format from RunB of 2010 (/Mu/Run2010B-Apr21ReReco-v1/AOD)

/Mu/Run2010B-Apr21ReReco-v1/AOD, CMS collaboration

Cite as: CMS collaboration (2014). Mu primary dataset in AOD format from RunB of 2010 (/Mu/Run2010B-Apr21ReReco-v1/AOD). CERN Open Data Portal. DOI:10.7483/OPENDATA.CMS.B8MR.C4A2

Home Contact CMS Software energy 1TeV Accelerator CERN LHC Fermilab Dataset /Mu/Run2010B-Apr21ReReco-v1/AOD

### Description

Mu primary dataset in AOD format from RunB of 2010

### Notes

This dataset contains all runs from 2010 RunB. The list of validated runs, which must be applied to all analyses, can be found in [CMS list of validated runs Cert\\_136033-149442\\_7TeV\\_Apr21ReReco\\_Collisions10\\_JSON\\_v2.txt](#)

### Related Datasets

/Mu/Run2010B-v1/RAW

### Characteristics

Dataset: **32376291** events **2979** files **3.2 TB** in total

### System Details

Global tag: FT\_R\_42\_V10A:AB  
Recommended release for analysis: CMSSW\_4\_2\_1\_patch1

## How were these data selected?

There are four categories of triggers in the Mu dataset (with significant overlaps):

- 70% inclusive single muon triggers with varying trigger pt threshold 3.5,7.9,11,13,15,17,19,21 GeV plus a few with loosened quality cuts.
- 20% isolated single muon triggers with varying trigger pt threshold 9,11,13,15,17 GeV.
- 10% inclusive dimuon triggers with varying trigger pt threshold 3.5 GeV plus one Z-muon trigger with loosened quality cuts.
- 20% combinations of muon triggers with various pt thresholds 3.5,7.8,9,11 GeV with some EM/e/gamma or hadronic/jet energy deposit with thresholds 6-100 GeV.

## How were these data validated?

During data taking all the runs recorded by CMS are certified as good for physics analysis if all subdetectors, trigger, lumi and physics objects (tracking, electron, muon, photon, jet and MET) show the expected performance. Certification is based first on the offline shifters evaluation and later on the feedback provided by detector and Physics Object Group experts. Based on the above information, which is stored in a specific database called Run Registry, the Data Quality Monitoring group verifies the consistency of the certification and prepares a json file of certified runs to be used for physics analysis. For each reprocessing of the raw data, the above mentioned steps are repeated. For more information see:

[CMS data quality monitoring: Systems and experiences](#)

[The CMS Data Quality Monitoring software experience and future improvements](#)

[The CMS data quality monitoring software: experience and future prospects](#)

## How can you use these data?

You can access these data through the CMS Virtual Machine. See the instructions for setting up the Virtual Machine and getting started in

[How to install the CMS Virtual Machine](#)

[Getting started with CMS open data](#)

Context information about data selection, validation, use

# Information discovery

The screenshot shows the OpenData CERN website interface. At the top left is the 'opendata CERN' logo. A search bar is located at the top center. On the right, there is an 'About' link. Below the search bar, there are filters for 'Sort by' (Most recent, asc.) and 'Display' (detailed, 20 results). A pagination bar shows 'Found 3778 results.' with page numbers 1 through 9. The main content area displays two search results. The first result is for a simulated dataset: '/TTjets\_MSDecays\_scaleup\_mt172\_5\_7TeV-madgraph-tauola/Summer11LegDR-PU\_S13\_START53\_LV6-v1/AODSIM'. Below the title, it says 'Simulated dataset TTjets\_MSDecays\_scaleup\_mt172\_5\_7TeV-madgraph-tauola in AODSIM format for 2011 collision data (SM Systematic Variations)'. There is a link to 'See the description of the sim...' and three tags: 'Dataset', 'Simulated', and 'CMS'. The second result is for another simulated dataset: '/Vector1MTtoZZto4L\_M-125p6\_7TeV-JHUGenV3-pythia6/Summer11LegDR-PU\_S13\_START53\_LV6-v1/AODSIM'. It says 'Simulated dataset Vector1MTtoZZto4L\_M-125p6\_7TeV-JHUGenV3-pythia6 in AODSIM format for 2011 collision data (SM Inclusive)'. There is a link to 'See the description of the simulated dataset nam...' and three tags: 'Dataset', 'Simulated', and 'CMS'. On the left side, there is a 'Filter by type' section with a tree view of categories and their counts: Dataset (997), Documentation (56), Environment (19), Software (33), and Supplementary (2642). Each category has sub-items with their respective counts.

opendata CERN

Search

About

Filter by type

- Dataset 997
  - Collision 100
  - Derived 173
  - Simulated 723
- Documentation 56
  - About 8
  - Activities 19
  - Authors 3
  - Guide 16
  - Help 2
  - Policy 4
  - Report 1
- Environment 19
  - Condition 5
  - VM 11
  - Validation 3
  - Glossary 22
  - News 9
- Software 33
  - Analysis 16
  - Framework 4
  - Tool 8
  - Validation 5
- Supplementary 2642
  - Configuration 917
  - Luminosity 3
  - Trigger 1722

Sort by: Most recent asc. Display: detailed 20 results

Found 3778 results.

< 1 2 3 4 5 6 7 8 9 >

/TTjets\_MSDecays\_scaleup\_mt172\_5\_7TeV-madgraph-tauola/Summer11LegDR-PU\_S13\_START53\_LV6-v1/AODSIM

Simulated dataset TTjets\_MSDecays\_scaleup\_mt172\_5\_7TeV-madgraph-tauola in AODSIM format for 2011 collision data (SM Systematic Variations)

See the description of the sim...

Dataset Simulated CMS

/Vector1MTtoZZto4L\_M-125p6\_7TeV-JHUGenV3-pythia6/Summer11LegDR-PU\_S13\_START53\_LV6-v1/AODSIM

Simulated dataset Vector1MTtoZZto4L\_M-125p6\_7TeV-JHUGenV3-pythia6 in AODSIM format for 2011 collision data (SM Inclusive)

See the description of the simulated dataset nam...

Dataset Simulated CMS

/VBFHiggs0PToGG\_M-125p6\_7TeV-JHUGenV4-pythia6-tauola/Summer11LegDR-PU\_S13\_START53\_LV6-v1/AODSIM

Explore a variety of data, software, VMs, supplementary material. . .

# Masterclasses

The screenshot shows the CERN Open Data portal interface. At the top, there is a search bar with the text 'Search' and a magnifying glass icon. The 'opendata CERN' logo is on the left, and an 'About' link is on the right. Below the search bar, there are several filter sections on the left and a main results area on the right.

**Filter by type**

- Dataset: 58
- Derived: 58
- Documentation: 4
- Activities: 4
- Environment: 1
- VM: 1
- Software: 3
- Analysis: 2
- Tool: 1

**Filter by experiment**

- ALICE: 4
- ATLAS: 57
- CMS: 1
- LHCb: 4

**Filter by year**

- 2011: 18
- 2012: 37

**Filter by file type**

- jpg: 1
- root: 3

**Filter by keywords**

- education: 13
- external resource: 9
- masterclass: 66
- teaching: 7

**Sort by:** Best match | asc. |

**Display:** detailed | 20 results |

Found 66 results.

**ATLAS ZPath 2015 Masterclass dataset 1**  
A dataset of 1000 event display files accessible via HYPATIA as part of the Masterclasses Z-Path. The events were recorded in 2012 by the ATLAS detector....  
[Dataset](#) [Server](#) [ATLAS](#)

**ATLAS WPath 2015 Masterclass dataset 5**  
A dataset of 1000 event display files accessible via MINERVA as part of the Masterclasses W-Path. The events were recorded in 2011 by the ATLAS detector....  
[Dataset](#) [Server](#) [ATLAS](#)

**ATLAS WPath 2015 Masterclass dataset 7**  
A dataset of 1000 event display files accessible via MINERVA as part of the Masterclasses W-Path. The events were recorded in 2011 by the ATLAS detector....  
[Dataset](#) [Server](#) [ATLAS](#)

**ATLAS WPath 2014 Masterclass dataset 5**  
A dataset of 1000 events taken in 2011 by the ATLAS Detector, used in the Physics Masterclasses W Path....  
[Dataset](#) [Server](#) [ATLAS](#)

Masterclass resources held on the CERN Open Data portal

# Visualise detector events

opendata CERN Search About

Need HELP?

iSpy WebGL DoubleMuon:Events/Run\_167674/Event\_255544818 [3 of 25]

Detector

- Pixel Barrel
- Pixel Endcap (+)
- Pixel Endcap (-)
- Tracker Inner Barrel
- Tracker Outer Barrel
- Tracker Inner Detector (+)
- Tracker Inner Detector (-)
- Tracker Endcap (+)

CMS Experiment at the LHC, CERN  
Data recorded: 2011-Jun-25 00:15:00.683123 GMT  
Run / Event / LS: 167674 / 255544818 / 209

Click on a name under "Provenance", "Tracking", "ECAL", "HCAL", "Muon", and "Physics" to view contents in table

Interactive event display for high-level derived datasets

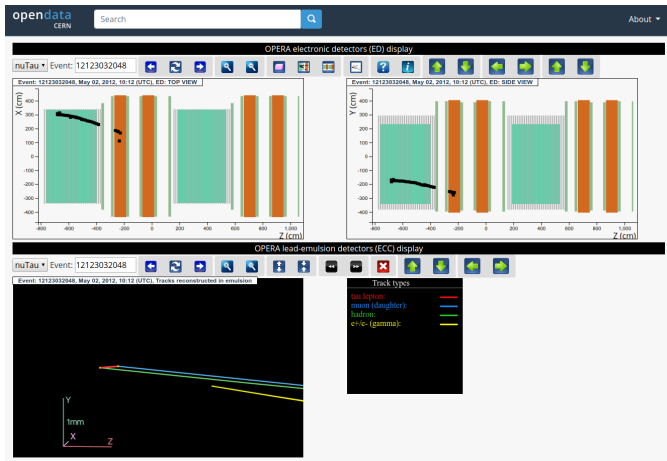
# Visualise histograms



Interactive histogramming for high-level derived datasets



# First non-LHC data releases



OPERA tau neutrino candidate and multiplicity studies data

# Jupyter notebooks

The screenshot shows a GitHub repository page for 'cms-opensource / cms-jupyter-materials-english'. The repository is on the 'master' branch. The file being viewed is 'Normfit-transversemomentum+pseudorapidity.ipynb'. The notebook content is as follows:

### Creating a normfit and transverse momentum+pseudorapidity

The point of this exercise is to learn to create a normal distribution fit for the data, and to learn what are transverse momentum and pseudorapidity (and how are they linked together). The data used is open data released by the [CMS](#) experiment.

#### First the fit

Let's begin by loading the needed modules, data and creating a histogram of the data to see the more interesting points (the area for which we want to create the fit).

```
In [ ]: # This is needed to create the fit
from scipy.stats import norm

import pandas as pd
import numpy as np
import matplotlib.mlab as mlab
import matplotlib.pyplot as plt

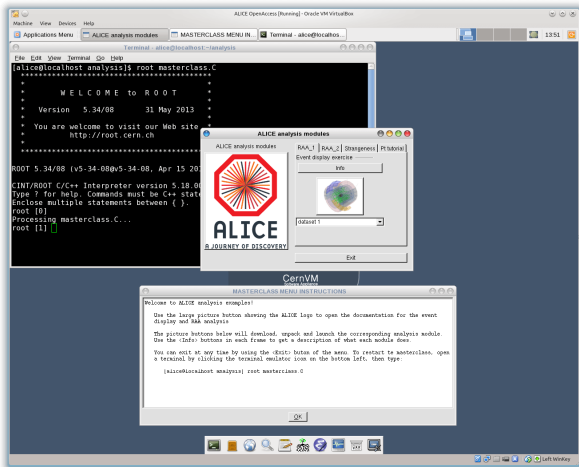
In [ ]: # Let's choose Dimaon_DoubleMu.csv
data = pd.read_csv('http://opendata.cern.ch/record/545/files/Dimaon_DoubleMu.csv')

# And save the invariant masses to iMass
iMass = data['M']

# Plus draw the histogram
n, bins, patches = plt.hist(iMass, 300, facecolor='g')
plt.xlabel('Invariant Mass (GeV)')
plt.ylabel('Amount')
```

CMS education activities using notebooks and CMS open data

# Virtual machines



Install CernVM virtual machines to explore primary datasets

# Analysis examples

## Higgs-to-four-lepton analysis example using 2011-2012 data

Jomhari, Nur Zulaiha; Geiser, Achim; Bin Anuar, Afiq Aizuddin;

Cite as: Jomhari, Nur Zulaiha; Geiser, Achim; Bin Anuar, Afiq Aizuddin; (2017). Higgs-to-four-lepton analysis example using 2011-2012 data. CERN Open Data Portal. DOI:10.7483/OPENDATA.CMS\_JKB8\_RR42

Software Analysis CMS Accelerator CERN/LHC

### Description

This research level example is a strongly simplified reimplemention of parts of the original CMS Higgs to four lepton analysis published in *Phys.Lett. B716 (2012) 30-61*, arXiv:1207.7235.

The published reference plot which is being approximated in this example is [https://inspirehep.net/record/1124338/files/H4l\\_mass\\_3.png](https://inspirehep.net/record/1124338/files/H4l_mass_3.png). Other Higgs final states (e.g. Higgs to two photons), which were also part of the same CMS paper and strongly contributed to the Higgs boson discovery, are not covered by this example.

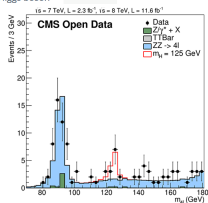
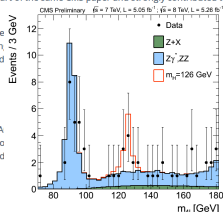
The example consists of different levels of complexity. The highest level minimal understanding of the content of this paper and of the meanin educational exercises. The lower levels might also be interesting for ed with the linux operating system and the ROOT analysis tool.

### Use with

The example uses legacy versions of the original CMS datasets in the A publication due to improved calibrations. It also uses legacy versions o but not identical to, the ones in the original publication. These legacy d in many later CMS publications.

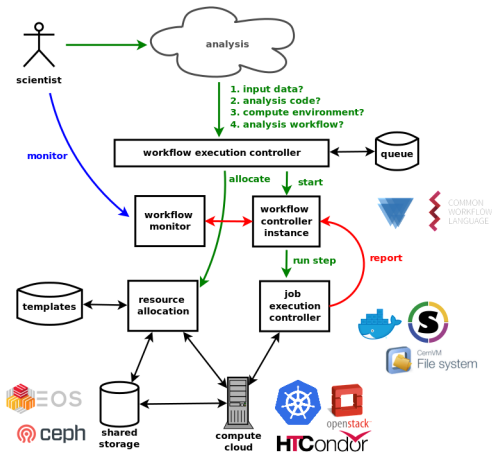
/DoubleElectron/Run2011A-12Oct2013-v1/AOD

/DoubleMu/Run2011A-12Oct2013-v1/AOD



Run realistic physics analysis examples

# Run analyses on the cloud



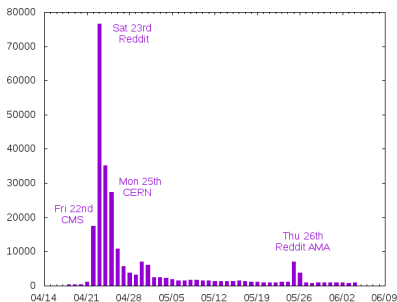
<http://www.reana.io/>

REANA reproducible research data analysis platform

# Release-driven usage patterns

The screenshot shows a TechCrunch news article. At the top, there's a navigation bar with 'News Video Events Crunchbase' and a search box. Below that, a banner for '10TH ANNUAL CRUNCHIES AWARDS' is visible. The article title is 'CERN releases 300TB of Large Hadron Collider data into open access', posted on April 20, 2016, by Devin Coldewey. Below the title are social media sharing icons for Facebook, Twitter, LinkedIn, Google+, YouTube, Email, and Print. The main image shows a 3D visualization of the LHC tunnel with a bright yellow starburst of data points. Below the image, the text reads: 'Cancel your plans for this weekend! CERN just dropped 300 terabytes of hot collider data on the world and you know you want to take a look.'

Open data releases are widely covered by general media



Six weeks in 2016: 200K users, 40K viewed records, 70K used event display, 3K used histogramming

# Research made open data

PHYSICAL REVIEW LETTERS

PHYSICAL REVIEW D

## Exposing the QCD Splitting Function with CMS Open Data

Andrew Larkoski,<sup>1,2</sup> Simeon Mrazek,<sup>1,2</sup> Jesse Thaler,<sup>1,2</sup> Anshu Tripathi,<sup>1,2</sup> and Wei Xue<sup>1,2</sup>

<sup>1</sup>Princeton University, 807 GFDL, Princeton, New Jersey 08542, USA  
<sup>2</sup>University of Virginia, The Stone Center of New York, Buffalo, New York 14260-2900, USA  
<sup>3</sup>Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA  
(Received 1 May 2017; revised manuscript received 27 July 2017; published 26 September 2017)

The splitting function is a universal property of quantum chromodynamics (QCD) which describes how energy is shared between particles. Despite its ubiquitous appearance in many QCD calculations, the splitting function cannot be measured directly, since it always appears multiplied by a collinear singularity factor. Recently, however, a new jet substructure observable was introduced which circumvents the jet splitting function for sufficiently light jet energies. This provides a way to expose the splitting function through an observable measurement at the Large Hadron Collider. In this Letter, we use public data released by the CMS experiment to study the two-prong substructure of jets and use the  $1 \rightarrow 2$  splitting function of QCD. To our knowledge, this is the first ever physics analysis based on the CMS Open Data.

DOI: 10.1103/PhysRevLett.119.132003

Quantum chromodynamics (QCD), like any weakly coupled gauge theory, exhibits universal behavior in the small angle limit. When two particles become collinear in QCD, the cross section for a  $2 \rightarrow n$  scattering process factorizes into a  $2 \rightarrow n-1$  scattering cross section multiplied by a universal  $1 \rightarrow 2$  splitting probability, with corrections suppressed by the degree of collinearity. Collinear universality is a fundamental property of QCD and appears in many applications, most famously in deriving the Dokshitzer-Gribov-Lipatov-Altarelli-Feynman evolution equations [1–3] (see also [4–5]), and in the heart of the factorization theorem in hadron-hadron collisions [6,15,16]. In addition, parton shower generators are based on iteratively applying  $1 \rightarrow 2$  splitting [16–18], fixed-order resummation schemes utilize the  $1 \rightarrow 2$  splitting functions [19–21], and the resummation of  $1 \rightarrow 2$  splitting [22–24] is a nonperturbative [22–24]. Collinear universality can be extended to multiparticle splittings at tree level and beyond [25–41], however, in all-order validity [42–47] is spoiled in the presence of Glauber modes [48–51]. More recently, jet substructure techniques [52–57] have been introduced to distinguish  $1 \rightarrow n$  decays of heavy particles from  $1 \rightarrow n$  splittings in QCD in order to enhance the search for new physics at the Large Hadron Collider (LHC) [55–56].

Despite its ubiquity, however, the  $1 \rightarrow 2$  splitting function cannot be directly measured at a collider, since collinear universality is inseparable from the existence of collinear singularities and only related nonambiguously to fragmentation functions. Specifically, when two particles are separated by an angle  $\theta$ , the  $1 \rightarrow 2$  splitting probability takes the form

$$dP_{1 \rightarrow 2} = \frac{d\sigma}{d\sigma_0} dP_{1 \rightarrow 2}(\zeta), \quad (1)$$

where the  $P_{1 \rightarrow 2}$  are the Altarelli-Parisi QCD splitting functions [1] which depend on the momentum fraction  $z$  and the parent flavor,  $f$ , and  $d\sigma_0$ . Crucially, this expression has an entanglement singularity in the  $\theta \rightarrow 0$  limit, as required to cancel corresponding virtual singularities from loop diagrams. In this sense, then, it is not so directly measurable the splitting function  $P_{1 \rightarrow 2}(\zeta)$  in data, though there is a coarse overabundance indirect evidence that  $P_{1 \rightarrow 2}(\zeta)$  is a universal function from the many successes of QCD in describing high-energy scattering (see, e.g., [57–67]).

In this Letter, we present a concrete method to test the  $1 \rightarrow 2$  splitting function in QCD by studying the two-prong substructure of jets. Our method is based on soft drop recombination [68] (also [69]), which recursively removes soft radiation from a jet until hard two-prong substructure is found. When applied to collinear quark and gluon related jets with intrinsic substructure, soft drop exposes the collinear core of the jet. As shown in Ref. [71], the momentum sharing between the two prongs (denoted  $z_{12}$ ) is closely related to the momentum fraction  $z$  appearing in Fig. 1(i), and the cross section for  $z_{12}$  corresponds to the QCD splitting function in the high-energy limit. While variants of  $z_{12}$  have appeared in many jet substructure studies, notably the  $\beta$  parameter in Refs. [57,72], to the best of our knowledge, no particular distribution has ever been proposed using actual collider data, though there are preliminary  $z_{12}$  studies from CMS [73], STAR [74], and ALICE [75] Collaborations. Here, we present the first analysis of  $z_{12}$  using LHC data, taking advantage of the first time of public data released by the CMS experiment [76].

The CMS Open Data are derived from a  $7 \text{ TeV}$  center-of-mass proton-proton collisions recorded in 2010 and released to the public on the CERN Open Data Portal in November 2014 [77]. The data are provided in analysis object (AOD) format, which is a CMS-specific data scheme based

PHYSICAL REVIEW D

119

## Jet substructure studies with CMS open data

Anshu Tripathi,<sup>1,2</sup> Wei Xue,<sup>1,2</sup> Andrew Larkoski,<sup>1,2</sup> Simeon Mrazek,<sup>1,2</sup> and Jesse Thaler<sup>1,2</sup>

<sup>1</sup>Center for Theoretical Physics, Massachusetts Institute of Technology,

Cambridge, Massachusetts 02139, USA

<sup>2</sup>Physics Department, Reed College, Portland, Oregon 97202, USA

<sup>3</sup>University of Virginia, The Stone Center of New York, Buffalo, New York 14260-2900, USA

(Received 1 May 2017; published 1 October 2017)

We use public data from the CMS experiment to study the two-prong substructure of jets. The CMS open data are based on  $31.8 \text{ fb}^{-1}$  of  $7 \text{ TeV}$  proton-proton collisions recorded at the Large Hadron Collider in 2010, holding a sample of 706,667 events containing a high quality control jet with intrinsic momentum larger than  $85 \text{ GeV}$ . Using CMS's particle flow reconstruction algorithm to obtain jet constituents, we extract the two-prong substructure of the leading jet using soft drop recombination. We find good agreement between results obtained from the CMS open data and those obtained from proton shower generators, and we also attempt to analyze jet substructure calculations performed on modified leading-jet-parton momenta. Although the 2010 CMS open data do not include simulated data for soft collinear singularity resummations, we use track-level observables to validate these substructure studies.

DOI: 10.1103/PhysRevD.96.104003

## I. INTRODUCTION

In November 2014, the CMS experiment at the Large Hadron Collider (LHC) announced the CMS Open Data project [1]. To our knowledge, this is the first time in the history of particle physics that research-grade collider data has been made publicly available for use outside of an official experimental collaboration. The CMS open data were reconstructed from  $7 \text{ TeV}$  proton-proton collisions in 2010, corresponding to a single low-luminosity running environment where pile-up contributions were minimal and trigger thresholds were relatively low. The CMS open data present an occasion especially to the particle physics community, both for performing physics studies that would be more difficult at higher luminosities and for demonstrating the scientific value of open data releases.

In this paper, we use the CMS open data to analyze the substructure of jets. Jets are collimated sprays of particles that are optimally produced by LHC collisions, and by studying the substructure of jets, one can gain valuable information about their parentage [2–10]. A key application of jet substructure is measuring heavy objects like top quarks [11–13] and electroweak bosons [14,15,17,20–26,56]. To successfully tag such objects, though, one first has to understand the radiation patterns of ordinary quark and gluon jets [26,60–70], which are the main backgrounds to boosted objects. The CMS open data are a fantastic

resource for performing these baseline quark/gluon studies. Using the Jet Primary Dataset [78], we perform initial investigations of the two-prong substructure of jets as well as present a general analysis framework to facilitate future studies. The effort is complementary to the growing catalog of jet substructure measurements performed within the ATLAS and CMS collaborations [77–106].

The core of our analysis is based on soft-drop declustering [79], which is a jet grooming technique [70–202] that mitigates jet contamination from initial state radiation (ISR), underlying event (UE), and pileup. For the studies in this paper, we use the soft-drop parameter  $\beta$  equal to zero, such that the soft drop behaves like the modified mass drop trigger (MDT) [80,203]. After soft dropping, a jet is composed of two well-defined subjets, which can then be used to derive various two-prong substructure observables. In addition to computing the CMS open data to probe shower generators, we perform first-principle calculations of soft-dropped observables using recently developed analytic techniques [40,203,206]. In a companion paper, we use soft drop to expose the QCD splitting function in the CMS open data [207], a similar strategy was used to perturbative CMS [107], STAR [108], and ALICE [109] heavy ion studies to test for possible modifications to the splitting function from the dense QCD medium [110,111].

For studying jet substructure, the key features of the CMS open data is that they contain full information about particle

<sup>1</sup>publicly-released  
<sup>2</sup>collaboration-wide  
<sup>3</sup>internationally-released  
<sup>4</sup>scientifically-released  
<sup>5</sup>physicists-released

<sup>1</sup>To highlight the ubiquity of the field, we have attempted to list all published jet substructure measurements from ATLAS and CMS. Please contact us if we missed a reference.  
<sup>2</sup>The original mass-drop trigger [80] was implemented using techniques for jet substructure not also applicable with Refs. [20–26].

0031-9007/17/119.13(2003)

132003-1

© 2017 American Physical Society

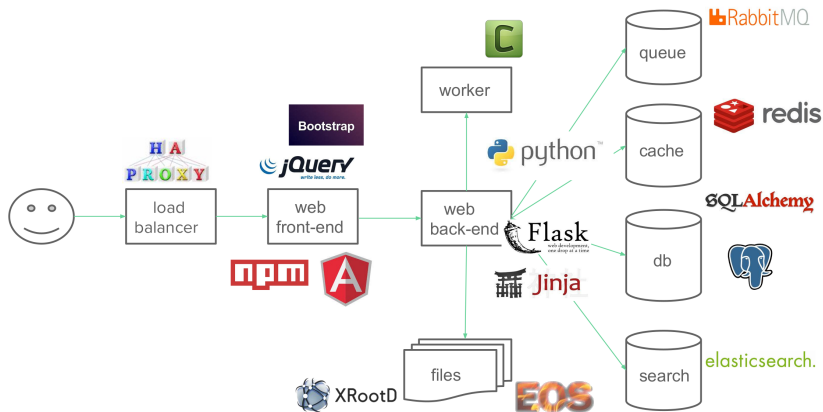
2478-0033/2017/96(7):104003(3)

04000-1

© 2017 American Physical Society

# First interesting analyses by theorists (esse et hae et al, MIT)

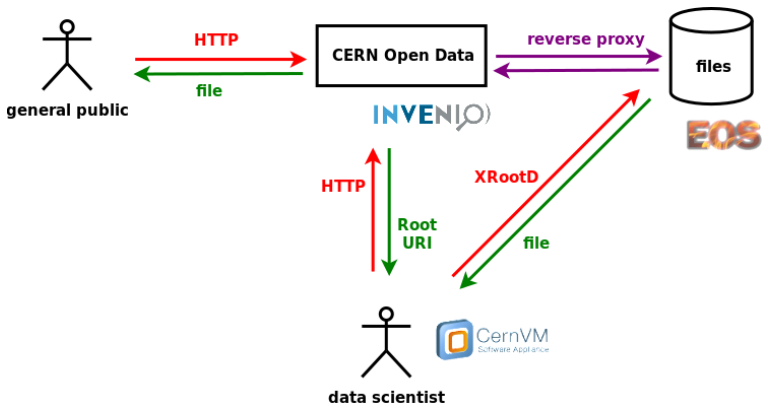
# Technology



Technology stack using **INVENIQ** digital repository



# Data exposure



HTTP and XRootD access scenarios

# Synergies for outreach

## ■ digital repository use cases?

- upload/download data, VM, documentation
- version control, mint and cite persistent identifiers (DOI)
- organise and search through data

## ■ reaching wider scope?

- traditionally hosting LHC experiments
- first non-LHC data by OPERA
- forthcoming Machine Learning reference datasets

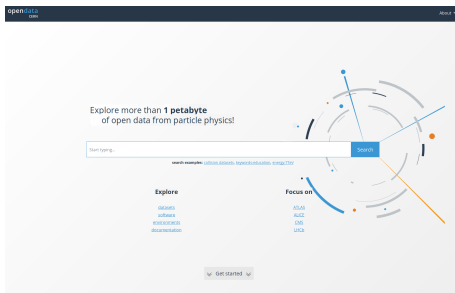
## ■ reaching wider audience?

- O(200K) “visitors” in six weeks; O(40K) deeper, O(3K) heavy use

## ■ run realistic examples?

- in-browser visualisations and histogramming
- analysis examples via Virtual Machines (ROOT, non-ROOT)
- analysis examples via containers (ROOT, non-ROOT)
- run locally or on cloud?

# CERN Open Data



<http://opendata.cern.ch>



<http://github.com/cernopendata>

**CERN IT** D. Rodriguez, T. Šimko · **CERN SIS** S. Dallmeier-Tiessen, S. Feger, P. Fokianos, A. Lavasa, I. Tsanaktsidis, A. Trzcinska · **ALICE** Y. Foka, M. Gheata, C. Grigoras, M. Zimmermann · **ATLAS** K. Cranmer, L. Heinrich, A. Sanchez Pineda, D. Rousseau, F. Socher · **CMS** H. Bittencourt, A. Calderon, E. Carrera, A. Geiser, A. Huffman, K. Lassila-Perini, T. McCauley, A. Rao, A. Rodriguez Marrero · **LHCb** S. Amerio, C. Burr, B. Couturier, S. Neubert, C. Parkes, S. Roiser, A. Trisovic · **OPERA** G. De Lellis, S. Dmitrievsky · **CERN CernVM** J. Blomer · **CERN EOS** L. Mascetti, H. Rousseau · **CERN OpenShift** A. Lossent, A. Peon