CERN-THESIS-2017-477

SEARCH FOR LOW-MASS AND HIGH-MASS NARROW DIJET RESONANCES WITH THE CMS DETECTOR AT $\sqrt{s}$ = 13 TeV

By

DAVID GRUFT SHEFFIELD

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Physics and Astronomy

written under the direction of

Amitabh Lath

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2017

ABSTRACT OF THE DISSERTATION

Search for Low-Mass and High-Mass Narrow Dijet Resonances with the CMS Detector at $\sqrt{s}$ = 13 TeV

By DAVID GRUFT SHEFFIELD

Dissertation Director:

Amitabh Lath

A search for narrow resonances decaying to the dijet final state at the CMS experiment. The data consists of $12.9\,\text{fb}^{-1}$ of proton–proton collisions. Data scouting is used to extend the search to lower masses. The dijet mass spectra at low and high masses are well parametrized by smoothly falling functions. No significant evidence of a resonance is observed. Upper limits on the product of cross section, branching fraction, and experimental acceptance at the 95% confidence level are set on resonances with masses between 0.6 and 7.5 TeV. The limits exclude benchmark models for string resonances below 7.4 TeV, scalar diquarks below 6.9 TeV, axigluons and colorons below 5.5 TeV, excited quarks below 5.4 TeV, color-octet scalars below 3.0 TeV, W' bosons below 2.7 TeV, Z' bosons below 2.1 TeV and from 2.3 to 2.6 TeV, vector and axial-vector DM mediators below 2.0 TeV, and RS gravitons below 1.9 TeV. These limits expand the excluded region beyond past searches and are the first limits set on DM mediators by the dijet channel.

Acknowledgments

I would like to thank my adviser, Amit Lath, for his guidance over the last five years. And to the professors, postdoctoral researchers, graduate students, and undergraduates that I have worked with and learned from during my time at Rutgers University.

This analysis was only possible through the work of my fellow members of CMS's dijet analysis group. In particular, I would like to thank Robert Harris and Francesco Santanastasio, who lead the group. I had the pleasure to work with Robert at Fermi National Accelerator Laboratory during most of this work and am grateful for the direction he provided. My thanks to the other members of the dijet analysis group: Dustin Anderson, Artur Apresyan, Lucija Bajan, Federico De Guio, Javier Duarte, Dinko Ferencek, Bora Isildak, Dimitris Karaasavvas, Sung-Won Lee, Juska Pekkanen, Maurizio Pierini, Federico Preiato, Niki Saoulidou, Eirini Tziaferi, Mikko Voutilainen, Tyler Zhixing Wang, and Si Xie. This analysis and its cross checks were only possible through the joint efforts of our team.

I would also like to thank all of the scientists, engineers, and workers who designed, built, operated, and calibrated CMS, the LHC, and the software we use. These are large and complicated collaborations. Without these thousands of people and billions of dollars in funding from government agencies around the world, this analysis would not be possible.

Dedication

For my parents

Table of Contents

List of Tables

List of Figures

Glossary

AK4 jets clustered by the anti-k_t algorithm with distance parameter $R = 0.4$.

ALICE A Large Ion Collider Experiment.

AOD analysis object data.

ATLAS A Toroidal LHC ApparatuS.

BCM1F fast beam conditions monitor.

CDF Collider Detector at Fermilab.

CERN European Organization for Nuclear Research.

CHS charged-hadron subtraction.

CL confidence level.

CMS Compact Muon Solenoid.

CSC cathode strip chamber.

DAQ data acquisition.

DM dark matter.

dof number of degrees of freedom.

DT drift tube.

EB ECAL barrel.

ECAL electromagnetic calorimeter.

EE ECAL endcap.

EM electromagnetic.

EW electroweak.

HB HCAL barrel.

HCAL hadron calorimeter.

HE HCAL endcap.

HF forward calorimeter.

HLT high-level trigger.

HO outer hadron calorimeter.

IP interaction point.

JEC jet energy correction.

JER jet energy resolution.

JES jet energy scale.

KK Kaluza–Klein.

L1 level-1 trigger.

LEP Large Electron–Positron Collider.

LHC Large Hadron Collider.

MC Monte Carlo.

MET missing transverse energy.

Offline occurring after data taking.

Online occurring during data taking.

PDF parton distribution function.

pdf probability density function.

PF particle-flow.

PLT pixel luminosity telescope.

PS Proton Synchrotron.

PU pileup.

PV primary vertex.

QCD quantum chromodynamics.

QED quantum electrodynamics.

QIE charge integrator and encoder.

RECO data fully reconstructed offline.

RF radio-frequency.

RPC resistive plate capacitor.

RS Randall–Sundrum.

RSS residual sum of squares.

SD spin-dependent.

SI spin-independent.

SL super layer.

SM standard model.

SPS Super Proton Synchrotron.

TEC tracker end cap.

TIB tracker inner barrel.

TID tracker inner disk.

TOB tracker outer barrel.

UA1 Underground Area 1.

UA2 Underground Area 2.

Chapter 1

Introduction

All particles predicted by the standard model (SM) of particle physics have been observed. The theory was developed in the 1960s to explain the particles and interactions discovered by the new field, which emerged in the wake of World War II. In the ensuing decades, the predicted particles were slowly found, until the final one was detected in 2012. Particle physics now has many theories and some indication of what might come next, but we do not know which of the paths nature will lead us down.

Ancient Greek and later Roman atomists speculated that the world consisted of particles and the void. In contrast, other philosophers believed that materials were made out of continua. The first truly elementary particle, the electron, was discovered by J. J. Thomson in 1897. Later, light was shown to be made up of photons, muons were discovered from cosmic rays, and the electron neutrinos were detected in beta decays. During the same time numerous composite particles were discovered including protons and neutrons, which, together with electrons, form atoms. At the time, the composite nature of those particles was not clear.

The SM emerged out of that theory of quarks, called quantum chromodynamics (QCD); quantum electrodynamics (QED); and electroweak (EW) theory. The theory incorporated the already known particles: photon, electron, muon, electron neutrino, muon neutrino, up quark, down quark, and strange quark. In 1974, the charm quark was confirmed with the discovery of the J/ψ meson. The tau was discovered in 1975, although the tau neutrino was not discovered until 2000. The discovery of the upsilon in 1977 confirmed the existence of bottom quarks. The gluon was discovered in 1979,

and the W and Z bosons were discovered in 1983 using accelerators still used by the Large Hadron Collider (LHC). In 1995, the predecessors to the Compact Muon Solenoid (CMS) experiment discovered the top quark. The last missing particle from the SM, the Higgs boson, was discovered in 2012.

The SM has been a major success for particle physics over nearly half a century. It includes the most precise agreement between theory and experiment in science—the electron anomalous magnetic dipole, where theory and experiment agree to more than one part in a trillion. However, we know that there is more to the fundamental laws of physics than the SM. Neutrino oscillations have shown that neutrinos are massive unlike the massless neutrinos of the model. Astronomical observations have shown that most of the mass in the universe comes from dark matter (DM), which cannot be attributed to the known particles. The reason the forces have large disparities in their strengths is unknown. New physics beyond the SM must exist to cause these effects.

There are some clues to where we will find new physics and many theories of what it could be. We are now trying to find direct evidence. This search for dijet resonances is one of many attempts to find evidence of new particles. The properties of the particles would indicate what the new physics is.

The first particles to be discovered were observed from naturally occurring sources. Electrons are stable and only need to be stripped from atoms. Muons are unstable, so they had to be created and observed by an experiment before they decayed. However, the first muons that were detected were not created by humans; they came from cosmic rays hitting the Earth's atmosphere, which produces a shower of particles including muons. Experiments such as direct-detection searches for DM continue to look for naturally occurring particles.

CMS primarily observes particles created by humans through the collision of high-energy protons. The collisions create many particles that have already been observed, but they have the potential to produce particles that have escaped detection so far. The LHC produced the then unobserved Higgs boson in its collisions, which lead to the

particle's discovery. Colliders have been successfully used in the second half of the 20th century to discover several particles.

The LHC accelerates two counter-rotating beams of protons to high energies and crosses them to allow the gluons and quarks (collectively called partons) inside the protons of opposing beams to interact. The interacting partons can produce particles. The maximum mass of the particles produced depends on the energy in the partons—the higher the energy the collider can reach, the more massive the particles it can make. The particles can then decay and the detectors surrounding the collision can measure the decay products to determine the properties of the particles that were temporarily created.

The particles the LHC produces come from the interaction of two partons. Therefore, we know that the particle can also decay back to two partons. Each parton creates a shower of particles in the detector called a *jet*. A search for dijets resonances provides a broad way to look for new particles regardless of what that particle may be. The downside is that as a hadron collider, the LHC creates copious numbers of jets. This creates a large background of dijet events that can swamp the signal of a dijet resonance. Figure 1.1 shows the Feynman diagram for a dijet resonance.
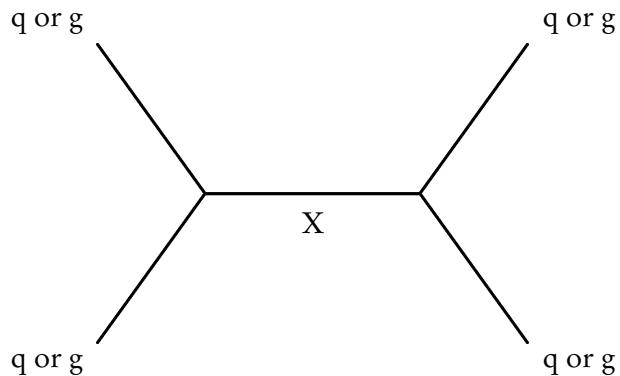
Figure 1.1: Feynman diagram for dijet resonance. Time increases to the right. The interaction of two gluons, two quarks, or a quark and a gluon creates unstable particle X. Particle X can then decay back to quarks and gluons, which create two jets. If the lifetime of X is small, it is seen as a resonance.

1.1 Motivation

Our goal is to find the signatures of new particles. The properties of a new particle would indicate what physics exists beyond the SM. We know some effects of new physics including neutrino masses, matter–antimatter asymmetry, gravity, DM, dark energy, and unnaturally large disparities between the strengths of the forces. There are many theories to explain these observations including supersymmetry, technicolor, string theory, and others meant to account for parts of these.

The SM leaves neutrinos massless, their observed nonzero masses indicates that there must be some mechanism not in the SM that gives them mass. The Universe, including ourselves, is overwhelming composed of matter. In the Big Bang, matter and antimatter should have be created in nearly equal amounts. The created matter and antimatter then annihilated, but an excess of matter still remains. Therefore, more matter must have been created than antimatter and there must be some new physics causing that asymmetry.

Astronomical observations have indicated the existence of DM, something that has mass but does not interact with electromagnetic (EM) radiation—i.e., it is dark. The effect of DM was first seen in galaxy rotation curves—the velocities of stars as a function of their distance from the center of the galaxy. The curves showed stars moving faster at larger distances than the mass determined from luminous matter could explain. The explanation is a halo of non-luminous DM that provides mass but not light. Since then, this massive but non-luminous DM has been observed by the gravitational lensing of light from behind galaxy clusters.

SM particles like neutrinos have been eliminated as possible explanations for DM. The observation of lensing in places without luminous matter has mostly eliminated modified theories of gravity that might have explained parts of the anomaly. We can therefore be confident that there exists a stable particle or particles that interact weakly with SM particles (through gravity and possibly some other weak interaction).

Direct-detection experiments attempt to observe DM through nuclear recoils as the

DM particles interact with SM particles in nuclei. Indirect-detection experiments attempt to observe the SM products of DM annihilation. Collider production attempts to produce DM particles through the interaction of SM particles and observe the result. CMS and ATLAS (A Toroidal LHC ApparatuS) are analyzing data to detect the signature of DM production at the LHC.

The DM particles could interact with SM particles through the existence of a mediator. The mediator could interact with partons and therefore appear as a resonance in the dijet mass spectrum.

In December 2015 CMS and ATLAS announced an excess in the diphoton mass spectrum at $750\,\text{GeV}$ [1].[1] This resulted in over 500 theory papers to explain this unexpected result [2] and analyses in multiple decay modes at both experiments looked for evidence of a new particle at $750\,\text{GeV}$. The dijet analysis is one of the searches that has looked in the region of this excess for evidence of a new particle decaying to partons.

The diphoton excess was seen by both experiments in $13\,\text{TeV}$ data. The analyses were then performed over a combination of 8 and $13\,\text{TeV}$ data. The excess that CMS saw based on both sets of data had a local significance at $750\,\text{GeV}$ of 3.4 standard deviations for a scalar (spin-0) narrow-width ($\Gamma/m = 1.4 \times 10^{-4}$) resonance. For a wide resonance ($\Gamma/m = 5.6 \times 10^{-2}$), the local significance was 2.3 standard deviations. The results were similar for spin-2 resonances. The global significance, which accounts for the "look elsewhere effect" of searching for the range of signal hypotheses in the the analysis, was 1.6 standard deviations [3].

ATLAS's results favored a wide-width hypothesis. At $750\,\text{GeV}$, a scalar resonance with $\Gamma/m = 7.6 \times 10^{-2}$ had a local significance of 3.8 standard deviations. For a spin-2 resonance, the local significance was 3.9 standard deviations for $\Gamma/m = 6.0 \times 10^{-2}$. The global significance for both was 2.1 standard deviations [4].

The cross sections measured by CMS and ATLAS for 8 and $13\,\text{TeV}$, separately, are

---

[1] An electronvolt eV is the standard unit of energy in particle physics $1\,\text{eV} \approx 1.6 \times 10^{-19}\,\text{J}$. It is the energy that an electron gains or loses as it moves across a $1\,\text{V}$ electric potential difference.

given in Table 1.1. The excesses are greater in the 13 TeV data. One explanation is that the resonance is produced by gluon–gluon fusion. Increasing the center-of-mass energy from 8 TeV to 13 TeV will increase the cross section—especially for resonances produced by gluons. The other explanation is that the 13 TeV data is a statistical fluctuation and the stand-alone 8 TeV result that there is no statistically significant excess at the order of 1 fb is correct.

Table 1.1: Cross section in femtobarns of diphoton excesses at 750 GeV seen by CMS and ATLAS at 8 and 13 TeV for fits to narrow and wide resonances. From Ref. [2].

| | $\sqrt{s} = 8$ TeV | | $\sqrt{s} = 13$ TeV | |
	narrow	wide	narrow	wide
CMS	0.63 ± 0.31	0.99 ± 1.05	4.8 ± 2.1	7.7 ± 4.8
ATLAS	0.21 ± 0.22	0.88 ± 0.46	5.5 ± 1.5	7.6 ± 1.9

1.2 Related searches

Dijet searches for new physics have been performed at several hadron colliders over the past 30 years [5]. The earliest dijet searches at a hadron collider were performed with with proton–antiproton collisions from the Super Proton Synchrotron (SPS) collider, which is now part of the LHC accelerator complex. From the 1980s, the Underground Area 1 (UA1) experiment published results looking in the mass range 70–300 GeV with center-of-mass energy $\sqrt{s} = 630$ GeV [6,7]. Underground Area 2 (UA2), the other detector at SPS, searched in the mass range 50–300 GeV in the early 1990s [8, 9].

Analyses at the Tevatron, which collided protons and antiprotons at $\sqrt{s} = 1.8$ and 1.96 TeV, were performed at the Collider Detector at Fermilab (CDF) [10–13] and D0 [14] in the 1990s and 2000s. These analyses developed the techniques used in this search.

CMS and ATLAS have performed similar studies since 2010 at the LHC. CMS performed studies at $\sqrt{s} = 7$ TeV [15–17], 8 TeV [18, 19], and 13 TeV [20]. ATLAS performed studies at 7 TeV [21–24], 8 TeV [25], and 13 TeV [26].

CMS used data scouting for the first time at $\sqrt{s} = 8\,\text{TeV}$ to search at lower masses than the standard CMS analysis could probe [27]. ATLAS has adopted a similar trigger-level analysis at $\sqrt{s} = 13\,\text{TeV}$ [28].

A list of past dijet resonance searches is presented in Table 1.2. Benchmark models have been excluded by dijet searches at the 95% CL in the ranges 0.26–7.0 TeV for string resonances, 0.29–6.0 TeV for scalar diquarks, 0.14–5.2 TeV for excited quarks, 0.12–5.1 TeV for axigluons and colorons, 0.10–0.26 and 0.28–2.6 TeV for W' bosons, 0.13–0.25 and 0.32–1.7 TeV for Z' bosons, and 0.5–1.6 TeV for Randall–Sundrum (RS) gravitons. Past searches have not set limits on DM mediators. This analysis extends CMS's past analyses to lower cross section and higher mass. It also presents the first limits on DM set with the dijet channel and uses data scouting to look for a resonance at 750 GeV.

Table 1.2: Past dijet resonance searches showing the experiment, year, reference, center-of-mass energy $\sqrt{s}$, integrated luminosity L, and mass ranges excluded at the 95% CL for axigluon, coloron, excited quark, W', Z', E_6 diquark, string, and RS graviton resonances.

Experiment	Year	Ref.	$\sqrt{s}$ [TeV]	L [fb^{-1}]	Axigluon or coloron [TeV]	Excited quark [TeV]	W' [TeV]	Z' [TeV]	E_6 diquark [TeV]	String [TeV]	RS graviton [TeV]
UA1	'86	[29]	0.63	0.00026	0.13–0.28	—	—	—	—	—	—
UA1	'88	[7]	0.63	0.00049	0.15–0.31	—	—	—	—	—	—
CDF	'90	[30]	1.8	0.000026	0.12–0.21	—	—	—	—	—	—
UA2	'90	[8]	0.63	0.0047	0.22–0.64	—	0.10–0.16	—	—	—	—
CDF	'93	[10]	1.8	0.0042	—	—	—	—	—	—	—
UA2	'93	[9]	0.63	0.011	—	0.14–0.29	0.13–0.26	0.13–0.25	—	—	—
CDF	'95	[11]	1.8	0.019	0.20–0.87	0.20–0.56	—	—	—	—	—
CDF	'97	[12]	1.8	0.106	0.20–0.98	0.20–0.52, 0.58–0.76	0.30–0.42	—	0.29–0.42	—	—
D0	'04	[14]	1.8	0.109	—	0.20–0.78	0.30–0.80	0.40–0.64	—	—	—
CDF	'09	[13]	1.96	1.13	0.26–1.25	0.26–0.87	0.28–0.84	0.32–0.74	0.29–0.63	0.26–1.4	—
ATLAS	'10	[24]	7	0.00032	—	0.30–1.26	—	—	—	—	—
CMS	'10	[17]	7	0.0029	0.50–1.17, 1.47–1.52	0.50–1.58	—	—	0.50–0.58, 0.97–1.08, 1.45–1.60	0.50–2.50	—
ATLAS	'11	[23]	7	0.036	0.60–2.10	0.60–2.15	—	—	—	—	—
CMS	'11	[16]	7	1	1.00–2.47	1.00–2.49	1.00–1.51	—	1.00–3.52	1.00–4.00	—
ATLAS	'11	[21]	7	1	0.80–3.32	0.80–2.99	—	—	—	—	—
ATLAS	'12	[22]	7	4.8	—	1–2.83	1–1.68	—	—	—	—
CMS	'12	[15]	7	5	1–3.36	1–3.32	1–1.92	1–1.47	1–3.75	1–4.31	—
CMS	'13	[19]	8	4	1–3.27	1–3.19	1–1.73	1–1.62	1–4.28	1–4.78	1–1.45
ATLAS	'14	[25]	8	20.3	—	0.8–4.06	0.8–2.45	—	—	—	—
CMS	'15	[18]	8	19.7	1.3–3.6	1.2–3.5	1.2–1.9, 2.0–2.2	1.2–1.7	1.2–4.7	1.2–5.0	1.2–1.6
CMS	'15	[20]	13	2.4	1.5–5.1	1.5–5.0	1.5–2.6	—	—	1.5–7.0	—
ATLAS	'16	[26]	13	3.6	—	2.0–5.2	1.25–2.6	—	—	—	—
CMS	'16	[27]	8	18.8	0.5–1.6	0.5–1.6	0.5–1.6	0.5–1.6	—	—	0.5–1.6

Chapter 2

Theory

2.1 Standard model

The standard model (SM) was developed starting in the 1960s in response to the discoveries of early particle physics. It was an attempt to describe the electromagnetic, weak, and strong forces and the particles observed at the time. It predicted the existence of several particles—the last of which was confirmed in 2012. The elementary particles of the theory are shown in Fig. 2.1.

Figure 2.1: Table of SM particles. Quarks (purple) and leptons (green) are in columns by generation. Shaded loops show which fermions couple to the gauge bosons (red). Quark and gluon colors and antiparticles are not shown. From Ref. [31].

The SM is a gauge theory with the symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The

three forces in the theory are mediated by spin-1 gauge bosons. The electromagnetic force is mediated through the massless photon γ, which does not carry electric charge. The weak force is mediated through three massive bosons, the charged $W^{\pm}$ bosons and the neutral Z^0 boson. These four particles arise out of the $SU(2) \times U(1)$ portion of the theory. From the $SU(3)$ portion come eight gluons g. While massless and electrically neutral, the gluons carry color charge.

The final elementary boson is the Higgs H. It is the result of the spontaneous breaking of the EW $SU(2) \times U(1)$ symmetry, which gives the massive gauge bosons and fermions their masses. The particle was discovered by CMS and ATLAS in 2012. After EW symmetry breaking, the SM symmetry group becomes $SU(3) \times U(1)$.

The first elementary particle to be discovered was the electron e. Two additional particles similar to the electron except for their greater masses were discovered. They are the muon μ and the tau τ. They are referred to as *flavors*. These couple to the electromagnetic and weak forces. There are three related particles called neutrinos. In the SM these particles are massless. But neutrino oscillations demonstrated that they must be massive—massless particles would not oscillate flavors. These oscillations arise because the mass eigenstates for the neutrinos (ν_1, ν_2, and ν_3) are a mixture of the flavor eigenstates (ν_e, ν_μ, and ν_τ). Neutrinos are neutral and only interact through the weak force. Unlike other fermions, which come in both right- and left-handed versions of their spin, neutrinos only come left-handed (with right-handed antineutrinos). Together, these six particles are called *leptons*.

Starting with the proton, a collection of strongly interacting particles called hadrons were observed. The theory of quarks and their interactions, QCD, was developed to explain the properties of the many hadrons. It initially, included three quarks (up u, down d, and strange s) and brought order to the strongly interacting particles. The theory made predictions for new particles which have since been observed. Hadrons are composite particles—the bound states of quarks held together by the strong interaction mediated by gluons.

The discovery of new hadrons indicated the existence of additional quarks—the charm c and the bottom b. The existence of the bottom indicated a sixth flavor of quark, the top t. Its discovery provided further evidence for the SM.

The SM is an effective field theory; it is not a complete theory of physics. While three forces are included, the fourth, gravity, is not. At current energies, gravity does not have a relevant effect on particles—it is significantly weaker than the other forces.

2.2   Quantum chromodynamics

The strong force is described by quantum chromodynamics (QCD). It governs the interactions of quarks and gluons which carry the color charge of this force.

The QCD Lagrangian density

$$\mathscr{L}_{\mathrm{QCD}} = \sum_q \overline{\psi}_{q,a} \left(i \gamma^\mu \partial_\mu \delta_{ab} - g_s \gamma^\mu t_{ab}^C \mathscr{A}_\mu^C - m_q \delta_{ab} \right) \psi_{q,b} - \frac{1}{4} F_{\mu\nu}^A F^{A\,\mu\nu}, \qquad (2.1)$$

where repeated indices are summed over [32]. The quark-field spinors $\psi_{q,a}$ have quark flavor q with mass m_q and a is one of three color charges. The sum is over the six quark flavors. The factors γ^μ are the Dirac gamma matrices, t_{ab}^C are the generators of SU(3),[1] and g_s is the coupling parameter such that $\alpha_s = g_s^2/4\pi$. The gluon field is $\mathscr{A}_\mu^C$ with C running over the eight types of gluons. The field tensor

$$F_{\mu\nu}^A = \partial_\mu \mathscr{A}_\nu^A - \partial_\nu \mathscr{A}_\mu^A - g_s f_{ABC} \mathscr{A}_\mu^B \mathscr{A}_\nu^C, \qquad (2.2)$$

where f_{ABC} are the structure constants of SU(3) such that $[t^A, t^B] = i f_{ABC} t^C$.

QCD is a gauge theory with gauge group SU(3), which forms the SU(3)$_C$ component of the SU(3)$_C$ × SU(2)$_L$ × U(1)$_Y$ symmetry of the SM. Because the structure constants for SU(3) are non-zero,[2] the group is non-Abelian.

---

[1] The Gell-Mann matrices are related to the generators of SU(3) by $t_{ab}^C = \lambda_{ab}^C/2$.
[2] $f_{123} = 1$, $f_{147} = -f_{156} = f_{246} = f_{257} = f_{345} = -f_{367} = 1/2$, and $f_{458} = f_{678} = \sqrt{3}/2$.

The charge of QCD is called color charge in analogy to primary colors. The canonical names of the colors are red, green, and blue. Combining the three colors together gives something that is *colorless*. Similarly, combining a color (e.g., red) with its anti color (anti-red or, in terms of real colors, cyan) is also colorless. The quarks each carry one color and the antiquarks carry one anticolor. The gluons carry both a color and anticolor. They take one of eight combinations of colors and anticolors dictated by the generators of SU(3). The representation based on the Gell-Mann matrices is

$$
\begin{array}{ll}
\left(r\overline{b} + b\overline{r}\right)/\sqrt{2}, & -\mathrm{i}\left(r\overline{b} - b\overline{r}\right)/\sqrt{2}, \\[2mm]
(r\overline{g} + g\overline{r})/\sqrt{2}, & -\mathrm{i}(r\overline{g} - g\overline{r})/\sqrt{2}, \\[2mm]
\left(b\overline{g} + g\overline{b}\right)/\sqrt{2}, & -\mathrm{i}\left(b\overline{g} - g\overline{b}\right)/\sqrt{2}, \\[2mm]
\left(r\overline{r} - b\overline{b}\right)/\sqrt{2}, & \left(r\overline{r} + b\overline{b} - 2g\overline{g}\right)/\sqrt{6}.
\end{array}
\tag{2.3}
$$

A gauge transformation will rotate the three colors in SU(3) and change the representations of the colors of the quarks and gluons.

A gauge theory is one in which the Lagrangian is invariant as so-called gauge fields transform under a local continuous group. The Lagrangian remains the same so gauge transformations of the gauge fields do not produce observable effects in the physics. The force carrying particles are excitations of the gauge fields. The gauge field of QCD $\mathscr{A}_\mu^C$ can be transformed under the symmetry group SU(3) and leads to the eight gluons. The most familiar gauge field is the one that gives rise to the photon. QED has a U(1) gauge symmetry that leads to the gauge invariance in Maxwell's equations—the oldest instance of gauge symmetry.

One of the defining features of QCD is asymptotic freedom. The coupling is weak at large momentum transfers and strong at small momentum transfers. This is the opposite of the coupling of QED and limits the perturbative study of QCD.

No free quarks or gluons have ever been directly observed. Any colored particle will *hadronize*—form hadrons—on too short of a time scale to be observed free. This is called

confinement. Only the top quark does not hadronize; it decays before hadronization can occur. Free particles must be colorless singlet states. Two types of gluons are colorless: $(r\overline{r} - b\overline{b})/\sqrt{2}$ and $(r\overline{r} + b\overline{b} - 2g\overline{g})/\sqrt{6}$. However, these are part of a color octet (the eight gluon color states) not a color singlet, so they, like the other gluons, will be confined.

Hadrons are composite particles of quarks bound by the strong force. They have some number of valence quarks, which give the composite particle its quantum numbers. The hadrons also contain a sea of quarks and gluons (partons). The probably of a certain momentum for a parton inside the hadron is given by a parton distribution function (PDF). The PDF $q_f(x)$ for a quark of flavor f and momentum fraction x satisfy

$$\int_0^1 q_f(x)\,\mathrm{d}x = N_f, \tag{2.4}$$

where N_f is the number of valence quarks of that flavor. And similarly for antiquarks with PDF $\overline{q}_f(x)$ and number $\overline{N}_f$.

The simplest hadrons are mesons, which have one quark and one antiquark. The quarks have three color states with the fundamental representation **3**, while the antiquarks have three anticolor states $\overline{\mathbf{3}}$. Combining a quark and antiquark together gives

$$\mathbf{3} \otimes \overline{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1}. \tag{2.5}$$

This is identical to the combination of color and anticolor for gluons, where **8** is the color octet that gives rise to Eq. (2.3).[3] The meson must be a color singlet **1**, which is

$$\left(r\overline{r} + b\overline{b} + g\overline{g}\right)/\sqrt{3}. \tag{2.6}$$

Hadrons with three quarks or three anti-quarks are called baryons. For baryons with

---

[3]If there were a color singlet gluon in addition to the color-octet types, the singlet gluon would not be confined and it would be massless. The strong force would extend to infinite range like the EM force and QCD would have U(3) gauge symmetry. These are not observed.

three quarks and three colors

$$3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1 \tag{2.7}$$

and similarly for antibaryons. The color singlet state **1**

$$(rgb - rbg + gbr - grb + brg - bgr)/\sqrt{6} \tag{2.8}$$

is the only combination that would not be confined. In baryons, the analogue of color charge to actual color is most apparent. The presence of red, green, and blue color charges makes the state colorless just as adding red, green, and blue light creates the color white to humans. The most familiar baryons, in that we are made from them, are the proton and neutron. The proton contains two up and one down valence quarks, while the neutron has one up and two down valence quarks.

Recently, hadrons with four and five valence quarks have been observed. Like mesons and baryons, these tetraquarks and pentaquarks must be color singlets to not be confined.

As partons from the hard interaction at the LHC hadronize, they create mesons and baryons. These are what we observe (or, if they are unstable relative to the time of flight like B mesons, we observe their decay products). Hadronization occurs with low momentum transfer. The running of the coupling means QCD is non-perturbative in that regime, so the process of hadronization is not completely understood. Models are used when simulating events to match the observed phenomenology of hadronization.

The color factor for gluon emission by a quark is 4/3. For gluon emission by a gluon, the color factor is 3 [32]. Because of the larger color factor, gluons are more likely to radiate, which will affect the signal shapes in Chapter 9.

2.3 Resonances

Resonances are peaks in the differential cross section that result from short-lived particles. Stable particles and unstable ones with long enough lifetimes will interact with detectors. Particles that decay before they can be directly detected can only be observed through their decay products.

Energy and momentum are conserved, so the four-momentum of the mother particle must equal the sum of the four-momenta of all daughter particles from the decay. Because the four-momenta are the same, the mass, which is the norm of the four-momentum $m = \sqrt{E^2 - p^2}$, is invariant. Calculating the invariant mass of all daughter particles will give (roughly) the mass of the mother particle.

A resonances has finite lifetime τ, where the probability of the particle to have not decayed by time t is $P(t) = e^{-t/\tau}$. The uncertainty principle applies to energy just as it does to momentum, so

$$\Delta E \Delta t \geq \frac{1}{2}. \tag{2.9}$$

An uncertainty in the time of order τ leads to an uncertainty in the energy and the mass on the order of $1/2\tau$. A histogram of particle masses have a peak at the resonance mass but the peak will have some finite width due to the uncertainty. The width associated with the distribution of the mass of a resonance will be twice this uncertainty

$$\Gamma = \frac{1}{\tau}. \tag{2.10}$$

The Breit–Wigner distribution (also known as the Cauchy distribution)

$$\frac{M\Gamma}{(m^2 - M^2)^2 + M^2\Gamma^2} \tag{2.11}$$

is used to describe resonances in the differential cross section. This distribution is maximum at mass $m = M$ and the full width at half maximum is Γ.

2.4 Beyond the Standard Model

The SM is not a complete theory of physics. While it describes three of the forces, gravity is absent from the theory. Because the gravitational interaction is so weak compared to the other three forces, it can be safely ignored in most conditions.

We have discovered that neutrinos are massive, but the SM does not provide a mechanism to give them their mass. Astronomical observations have indicated the presence of DM, which is massive and only weakly interacting. DM must be composed of a particle or particles not in the SM.

There is a hierarchy in the strength of the forces. The weakness of gravity relative to the EW interaction is unnaturally large. This discrepancy is called the hierarchy problem. Multiple theories have been developed to provide a mechanism that would take the fundamental energy scales of the gravitational and EW interactions, which are of the same order of magnitude in the theories, and cause their observed energy scales to be at much different values.

Physics could contain additional gauge symmetries and new gauge bosons associated with these groups. A major area of research is a symmetry between fermions and bosons called supersymmetry.

2.4.1 Excited quarks

Some models predict that quarks are composite particles instead of elementary ones. Evidence for compositeness could come from the observation of excited states q^*. The energy scale of the substructure is at Λ. The spin and weak isospin do not necessarily have to be the same as for quarks, although in the simplest models they are [33, 34].

The Lagrangian density with the interaction of an excited quark with an ordinary quark and gluon

$$\mathscr{L} \supset \frac{1}{2\Lambda} \overline{q}_R^* \sigma^{\mu\nu} \left(g_s f_s t_a G^a_{\mu\nu} + g f \frac{\tau}{2} \cdot W_{\mu\nu} + g' f' \frac{Y}{2} B_{\mu\nu} \right) q_L + \text{h.c.,} \qquad (2.12)$$

where g_s, g, and g' are the gauge coupling constants; f_s, f, and f' are dimensionless parameters for the deviations of the composite dynamics from SM behavior; t_a and τ are the generators of SU(3) and SU(2), respectively; $G^a_{\mu\nu}$, $W_{\mu\nu}$, and $B_{\mu\nu}$ are the gauge fields; and Y is the hypercharge. For a mass m^*, the partial decay width is

$$\Gamma(q^* \to qg) = \frac{1}{3}\alpha_s f_s^2 \frac{m^{*3}}{\Lambda^2}. \tag{2.13}$$

2.4.2 Randall–Sundrum gravitons

Randall and Sundrum proposed a modified theory of gravity with an extra space dimension to solve the hierarchy problem between the EW and Planck (gravity) energy scales [35, 36]. In the theory, both forces are at fundamentally the same scale. The hierarchy is created by including an exponential factor to the metric

$$ds^2 = e^{-2kr_c\phi}\eta_{\mu\nu}\,dx^\mu\,dx^\nu + r_c^2\,d\phi^2, \tag{2.14}$$

where k is the curvature of the extra dimension and on the order of the Planck scale, r_c is the compactification radius of the extra dimension, $\phi \in [0, \pi]$ is the coordinate of the extra dimension, and x^μ remains the coordinate of the four apparent spacetime dimensions.

The exponential warp factor takes the five-dimensional mass M at the EW scale and gives us the effective reduced Planck mass in four dimensions

$$\overline{M}_{\mathrm{Pl}}^2 = \frac{M^3}{k}\left(1 - e^{-2kr_c\pi}\right). \tag{2.15}$$

The extra dimension is an orbifold S^1/Z_2. At either end is a 3-brane. All SM particles reside on a "TeV" brane at $\phi = \pi$. Another brane lies at $\phi = 0$ and is called the "Planck" brane. Only gravitons can propagate through the bulk, everything else is confined to the $3+1$ dimensions on the branes.

We can expand spacetime fluctuations $h_{\mu\nu}(x,\phi)$ around the Minkowski metric as [37]

$$\hat{G}_{\mu\nu} = e^{-2kr_c|\phi|}\left(\eta_{\mu\nu} + \frac{2}{M^{3/2}}h_{\mu\nu}\right). \tag{2.16}$$

The fluctuations can be decomposed

$$h_{\mu\nu}(x,\phi) = \sum_{n=0}^{\infty} h_{\mu\nu}^{(n)}(x)\frac{\chi^{(n)}(\phi)}{\sqrt{r_c}}, \tag{2.17}$$

where $h_{\mu\nu}^{(n)}(x)$ are Kaluza–Klein (KK) modes in four-dimensional spacetime and $\chi^{(n)}(\phi)$ are fluctuations along the extra dimension [37].[4] The boundary conditions of the extra dimension determine $\chi^{(n)}(\phi)$. The interaction Lagrangian density on the TeV brane is

$$\mathscr{L}_{\text{int}} = -\frac{1}{M^{3/2}}T^{\mu\nu}(x)h_{\mu\nu}(x,\phi=\pi), \tag{2.18}$$

where $T^{\mu\nu}(x)$ is the Minkowski space energy-momentum tensor. Expanding in Eq. (2.17) with the solutions to $\chi^{(n)}(\phi)$ gives

$$\mathscr{L}_{\text{int}} = -\frac{1}{\overline{M}_{\text{Pl}}}T^{\mu\nu}(x)h_{\mu\nu}^{(0)}(x) - \frac{1}{\Lambda_\pi}T^{\mu\nu}(x)\sum_{n=1}^{\infty}h_{\mu\nu}^{(n)}(x). \tag{2.19}$$

The massless graviton mode $h_{\mu\nu}^{(0)}(x)$ couples with the usual $\overline{M}_{\text{Pl}}^{-1}$ strength. All other modes produce massive KK gravitons that couple with strength Λ_π^{-1}, where the energy scale

$$\Lambda_\pi = e^{-kr_c\pi}\overline{M}_{\text{Pl}} \tag{2.20}$$

is much lower and at the TeV level.

---

[4]Kaluza and Klein developed theories that attempted to explain electromagnetism with an extra fifth dimension. With a finite extra dimension fluctuations can be decomposed into modes with quantized momentum in that dimension p_5. If the particle is fundamentally massless $m = (E^2 - p^2 - p_5^2)^{1/2} = 0$, the KK mode will appear to have mass $m = (E^2 - p^2)^{1/2} = p_5$ under the false assumption that there are four dimension. This leads to a tower of KK modes at increasing masses.

The KK gravitons have masses

$$m_n = k x_n e^{-k r_c \pi},\qquad(2.21)$$

where x_n is the nth root of the Bessel function $J_1(x)$. The coupling can then be expressed in terms of the parameter m_n and the ratio $k/\overline{M}_{\mathrm{Pl}}$

$$g = \frac{1}{\Lambda_\pi} = x_1 \frac{k/\overline{M}_{\mathrm{Pl}}}{m_n}.\qquad(2.22)$$

We are interested in the case where $n = 1$—the lowest mass KK graviton.

Randall–Sundrum (RS) gravitons can interact with both gluon and quark–antiquark pairs. The partial widths for the least massive KK graviton are

$$\Gamma(\mathrm{G}\to\mathrm{gg}) = \frac{x_1^2}{10\pi}\left(\frac{k}{\overline{M}_{\mathrm{Pl}}}\right)^2 m_1$$

$$\Gamma(\mathrm{G}\to\mathrm{q\bar{q}}) = \frac{3x_1^2}{160\pi}\left(\frac{k}{\overline{M}_{\mathrm{Pl}}}\right)^2 m_1\left[\left(1 - 4\frac{m_{\mathrm{q}}^2}{m_1^2}\right)^{3/2} + \left(1 + \frac{8m_{\mathrm{q}}^2}{3m_1^2}\right)\right],\qquad(2.23)$$

where in the graviton mass range we probe, $m_1 \gg m_{\mathrm{q}}$ makes the bracketed factor approximately unity.

2.4.3 Axigluons

Axigluons arise from a model where the symmetry group of the strong sector has been expanded from the SU(3) of the SM to a chiral color group $\mathrm{SU}(3)_L \times \mathrm{SU}(3)_R$ [38]. At some energy scale the chiral group breaks and produces the familiar diagonal SU(3). Left- and right-handed fermions $\psi_{L,R} = (1 \mp \gamma_5)\psi/2$ transform differently. The left-handed fermions transform under $\mathrm{SU}(3)_L$ with the generators T_L^a, and the right-handed fermions transform under $\mathrm{SU}(3)_R$ with the generators T_R^a. The generators can be transformed into a vectorial generator $T_V^a = T_L^a + T_R^a$ and an axial generator $T_A^a = T_L^a - T_R^a$ [5].

The vectorial gauge field is equivalent to the field in QCD and gives us the gluons. The axial gauge field and the axial generators T_A^a result in a second color octet called axigluons. The theory also requires new massive particles to cancel out triangular anomalies.

Axigluons can decay to quark–antiquark pairs, but decays to gluon–gluon pairs is prevented by parity conservation. The interaction term of the Lagrangian density

$$\mathscr{L}_{\mathrm{int}} = -\mathrm{i}g_s \left(\sum_{ij} \overline{q}_i \gamma_5 \gamma_\mu t_a q_j \right) \mathscr{A}^{a\mu}, \tag{2.24}$$

where t^a are the generators of SU(3) and $\mathscr{A}^a$ are the axigluon fields [5]. Strong parity requires the left and right couplings to be equal, which force the coupling to be the same as the strong coupling g_s. The axigluon width [29]

$$\Gamma(A \rightarrow q\overline{q}) = \frac{N\alpha_s M_A}{6}, \tag{2.25}$$

where N is the number of open decay channels and M_A is the axigluon mass.

2.4.4 Colorons

Like the axigluon model, the coloron model expands the symmetry group of the strong sector to $\mathrm{SU}(3)_1 \times \mathrm{SU}(3)_2$ with coupling ξ_1 and ξ_2 such that $\xi_2 \gg \xi_1$ [39,40]. The quarks are triplet representations of the $\mathrm{SU}(3)_2$ group. A scalar boson Φ breaks the symmetry of the two groups at an energy scale f resulting in the unbroken SU(3) of QCD. The scalar field develops a vacuum expectation value of f.

Once the symmetry is broken, an octet of massless gluons is recovered with coupling g_3 equal to the standard QCD coupling. An octet of massive *colorons* is also produced. The colorons interact with quarks through the QCD-like Lagrangian density term

$$\mathscr{L}_{\mathrm{int}} = -g_s \cot\theta \sum_f \overline{q}_f \gamma_\mu t_a q_f C^{\mu a}, \tag{2.26}$$

where $\cot\theta = \xi_2/\xi_1 > 1$ is the gauge boson mixing angle. The mass of the coloron is

$$M_C = \frac{g_3}{\sin\theta\cos\theta}f.$$

(2.27)

The decay width of the coloron is

$$\Gamma(C \to q\bar{q}) \approx \frac{n}{6}\alpha_s M_C \cot^2\theta$$

(2.28)

with $\alpha_s = g_3^2/4\pi$ and n flavors of quarks lighter than $M_C/2$ [41].

2.4.5 Color-octet scalars

A color-octet scalar is a type of exotic color resonance. The interaction term in the Lagrangian density for color-octet scalars and gluons

$$\mathcal{L}_{\text{int}} = g_s \frac{k_s}{\Lambda} d^{abc} S_8^a G_{\mu\nu}^b G^{c,\mu\nu},$$

(2.29)

where g_s is the strong coupling, k_s is the anomalous scalar coupling, Λ is the characteristic scale of the interaction, d^{abc} are structure constants of SU(3) such that the generators $\{t^a, t^b\} = \delta^{ab}/3 + d^{abc}t^c$, S_8 is the color-octet scalar field, and $G_{\mu\nu}$ is the gluon field tensor [42]. The width of the resonance is

$$\Gamma(S_8 \to gg) = \frac{5}{6}\alpha_s k_s^2 \frac{M^3}{\Lambda^2}.$$

(2.30)

2.4.6 New heavy gauge bosons

New gauge bosons W' and Z' are predicted by a number of models [43, 44]. They are heavy versions of the SM W and Z bosons.

After EW symmetry breaking, the weak interaction terms for the SM Lagrangian

density

$$\mathcal{L}_{\text{int}} = -\frac{g}{2\sqrt{2}}\sum_i \overline{\Psi}_i \gamma^\mu (1-\gamma^5)(T^+ W^+_\mu + T - W^-_\mu)\Psi_i$$
$$-\frac{g}{2\cos\theta_W}\sum_i \overline{\psi}_i \gamma^\mu (g^i_V - g^i_A \gamma^5)\psi_i Z_\mu, \qquad (2.31)$$

where g is the coupling constant for the SU(2) sector of the EW group, θ_W is the weak mixing angle, and the sum is over fermions with fields ψ_i [32]. The left handed fermion fields are combined into doublets

$$\Psi_i = \begin{pmatrix} \nu_i \\ \ell^-_i \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} u_i \\ d'_i \end{pmatrix}, \qquad (2.32)$$

where $d'_i = \sum_j V_{ij} d_j$ with Cabibbo–Kobayashi–Maskawa matrix V_{ij}. The fields $W^\pm_\mu$ and Z_μ lead to the bosons $W^\pm$ and Z making the first term the charged-current weak interaction between a W and two fermions and the second term the neutral-current weak interaction between a Z and two fermions. In the first term, $T^\pm$ are the weak isospin raising and lowering operators. In the second term, the vector and axial-vector couplings

$$g^i_V = t_{3L}(i) - 2Q_i \sin^2\theta_W$$
$$g^i_A = t_{3L}(i), \qquad (2.33)$$

where $t_{3L}(i)$ is the weak isospin and Q_i is the charge of fermion ψ_i. The weak isospin for u_i and ν_i is $+1/2$, while it is $-1/2$ for d_i and ℓ_i.

Models for new heavy gauge bosons can include extensions of the EW $SU(2)_L \times U(1)_Y$ gauge symmetry group. They can also come about from KK modes of the W and Z bosons in models with extra dimensions. The new gauge bosons carry the same charge and spin as their SM analogues. They may also couple to quarks and leptons in the same way that the W and Z boson do, which makes their cross section proportional to the SM

ones with

$$G_F' = G_F \left(\frac{M}{M'}\right)^2,\tag{2.34}$$

where G_F is the Fermi constant, M is the mass of the W or Z boson, and M' is the mass of the W$'$ or Z$'$ boson.

Models can also include *leptophobic* coupling to only quarks and not to leptons. This is useful because this analysis only looks at the interaction of a resonance to quarks, not to leptons. We can measure the possible universal coupling to quarks g_q' with the assumption of a leptophobic Z$'$ boson that can be applied to other models.

2.4.7   String resonances

In string theory, particles are vibration states of strings. The strings have mass M_s, which are generally at the Planck scale. In some theories with large extra dimensions, the string mass can drop down to the TeV scale [45, 46].

The Veneziano form factor

$$V(\hat{s}, \hat{t}, \hat{u}) = \frac{\Gamma(1-\hat{s}/M_s^2)\Gamma(1-\hat{u}/M_s^2)}{\Gamma(1-\hat{t}/M_s^2)},\tag{2.35}$$

where $\hat{s}$, $\hat{t}$, and $\hat{u}$ are Mandelstam variables and $\Gamma(x)$ is the Γ-function, is for two-to-two parton scattering. The poles from expanding in the s-channel give virtual Regge resonances with mass $\sqrt{n}M_s$. The first resonance therefore has a mass of M_s. If M_s is at the TeV scale, it could be produced by the LHC. The cross section and width for gluon–gluon production is model independent if the string coupling is small. The model independence is only approximate for gluon–gluon resonances.

2.4.8 E_6 scalar diquarks

Some superstring theory models in 10 dimensions predict the grand unification symmetry group to be E_6. These models predict color-triplet scalar diquarks D and D^c [47]. The diquarks have charge $-1/3$ and $+1/3$, respectively, and couple to the u and d

quarks [48].

The interaction between the light quarks and E_6 diquarks is described by

$$\mathscr{L} \supset \lambda \epsilon_{ijk} \overline{u}^{ci} \frac{1}{2}(1-\gamma_5)d^j D^k + \frac{\lambda_c}{2}\epsilon_{ijk}\overline{u}^i \frac{1}{2}(1+\gamma_5)d^{cj}D^{ck} + \text{h.c.}, \qquad (2.36)$$

where i, j, k are color indices and λ, λ_c are coupling parameters. The decay widths [5]

$$\Gamma(D \to q\overline{q}) = \alpha_\lambda M_D = \frac{\lambda^2}{4\pi}M_D \qquad (2.37)$$

$$\Gamma(D^c \to q\overline{q}) = \alpha_{\lambda_c} M_{D^c} = \frac{\lambda_c^2}{4\pi}M_{D^c}. \qquad (2.38)$$

2.4.9 Dark matter

The effects of dark matter (DM) have been observed in the rotation of galaxies, gravitational lensing, and the large-scale structure of the Universe among other astronomical observations. Information can be discerned about the nature of DM particles by observing its effects. However, no DM particles have been detected.

There are three basic methods of observing DM particles presented in Fig. 2.2. There is an interaction between two SM and two DM particles. Which are the initial particles and which are the final particles determines the method of detection is being used. Direct-detection experiments rely on DM particles present in the galaxy scattering off a target of SM particles. Direct-detection experiments tend to be underground to limit background from cosmic rays and contain large targets to probe lower cross sections. Indirect detection involves the observation of SM particles that result from the annihilation of DM particles. Indirect-detection experiments are astronomical searches that tend to observe areas of dense DM.

The final method of detection is production in a collider. The traditional method is to search for missing transverse momentum present from the created DM particles, which do not interact with the detector, that recoil off an SM particle used for the trigger. We use a simplified model for dark matter Fig. 2.3 [49, 50]. The DM particle is a
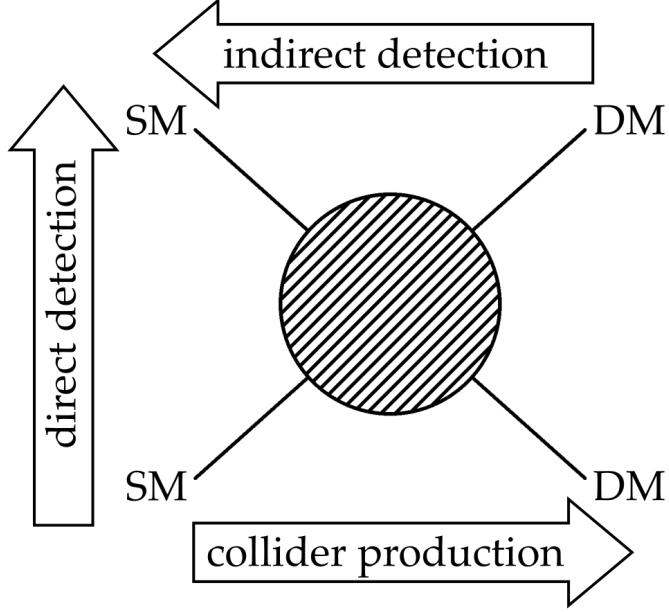
Figure 2.2: Feynman diagram of the interaction of two DM and two SM particles. Three methods of detection result from running time along each of the arrows.

Dirac fermion χ that interacts with SM particles through a spin-1 mediator Φ. The mediator only interacts with DM and quarks. Instead of searching for the signal of the DM particles disappearing in the detector, we look for a dijet resonance from the mediator.

The Lagrangian density for this simplified model [49,51]

$$\mathscr{L} \supset \Phi_\mu \overline{\chi}\gamma^\mu(g_{\mathrm{DM}}^V - g_{\mathrm{DM}}^A \gamma^5)\chi + \sum_i \Phi_\mu \overline{q}_i \gamma^\mu(g_{\mathrm{q}}^V - g_{\mathrm{q}}^A \gamma^5)q_i, \tag{2.39}$$

where the index i is a summed over the six quark flavors; g_{DM}^V and g_{DM}^A are the vector and axial-vector, respectively, couplings of the mediator to the DM particles; and g_{q}^V and g_{q}^A are the vector and axial-vector couplings of the mediator to the quarks. We assume the mediator has equal coupling to all quarks g_{q}. The partial widths for vector mediators

$$\Gamma(\Phi_{\mathrm{vector}} \to \chi\overline{\chi}) = \frac{g_{\mathrm{DM}}^2 M_{\mathrm{Med}}}{12\pi}\sqrt{1-4z_{\mathrm{DM}}}(1+2z_{\mathrm{DM}})$$

$$\Gamma(\Phi_{\mathrm{vector}} \to q\overline{q}) = \frac{g_{\mathrm{q}}^2 M_{\mathrm{Med}}}{4\pi}\sqrt{1-4z_{\mathrm{q}}}(1+2z_{\mathrm{q}}) \tag{2.40}$$
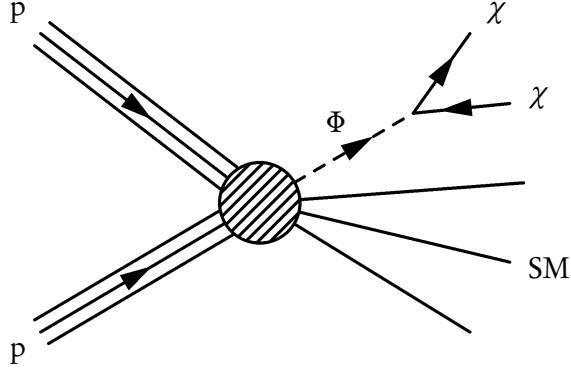
Figure 2.3: Feynman diagram of simplified model used for DM. SM particles interact with DM mediator Φ. The mediator interacts with DM particles χ.

and for axial-vector mediators

$$
\Gamma(\Phi_{\text{axial-vector}} \to \chi\overline{\chi}) = \frac{g_{\text{DM}}^2 M_{\text{Med}}}{12\pi}(1 - 4z_{\text{DM}})^{3/2}
$$
$$
\Gamma(\Phi_{\text{axial-vector}} \to q\overline{q}) = \frac{g_q^2 M_{\text{Med}}}{4\pi}(1 - 4z_q)^{3/2},
$$

(2.41)

where $z_q = m_q^2/M_{\text{Med}}^2$ and $z_{\text{DM}} = m_{\text{DM}}^2/M_{\text{Med}}^2$.

When $M_{\text{Med}} < 2m_{\text{DM}}$, the mediator is kinematically prohibited from decaying to DM particles. The result is a process in which two quarks produce the DM mediator that then decays back to two quarks. This behavior is exactly like the leptophobic Z' boson in Section 2.4.6 where $M_{Z'} = M_{\text{Med}}$ and the Z' boson coupling is equal to the DM mediator's coupling to quarks $g_q' = g_q$.

In the regime $M_{\text{Med}} \geq 2m_{\text{DM}}$, the mediator can decay to DM particles. The DM mediators are still being produced through quarks, so the production remains the same—some of the decays are just invisible to the detector. The observed events will fall as the branching fraction of the DM mediator to quarks $\mathcal{B}(\Phi \to q\overline{q}) = \sum_i \Gamma(\Phi \to q_i\overline{q}_i)/\Gamma$, where $\Gamma = \sum_i \Gamma(\Phi \to q_i\overline{q}_i) + \Gamma(\Phi \to \chi\overline{\chi})$ and we sum over the effective number of quark flavors that the DM mediator can decay to N_f. These models will be equivalent to

the leptophobic model with the same mass and

$$g'_q = g_q \sqrt{\mathcal{B}(\Phi \to q\bar{q})}. \tag{2.42}$$

The branching fraction differs between vector and axial-vector DM mediators except for two cases when both models give the same result: $M_{\mathrm{Med}} < 2m_{\mathrm{DM}}$ when $\mathcal{B}(\Phi \to q\bar{q})|_{M_{\mathrm{Med}} < 2m_{\mathrm{DM}}} = 1$ and $m_{\mathrm{DM}} = 0$ when

$$\mathcal{B}(\Phi \to q\bar{q})|_{m_{\mathrm{DM}}=0} \approx \frac{g_q^2}{g_q^2 + g_{\mathrm{DM}}^2/(3N_f)}, \tag{2.43}$$

with the assumption that $M_{\mathrm{Med}} \gg m_{\mathrm{q}}$.

2.4.10 Benchmark models

We use benchmark models in setting limits. The limits are set on model-independent decay modes and then interpreted as limits on particular versions of models with defined parameters. This allows us to set a finite number of limits and compare results between different searches.

The benchmark models are summarized in Table 2.1. The parameters used for the benchmark models are:

- Excited quarks: SM couplings $f_s = f = f' = 1$ and compositeness scale $\Lambda = M^*$, the excited quark mass.

- RS gravitons: the ratio $k/\overline{M}_{\mathrm{Pl}} = 1$.

- Axigluons: $N = 6$ open decay channels for axigluon decays.

- Colorons: $N = 6$ open decay channels for coloron decays with gauge boson mixing angle $\cot \theta = 1$.

- Color-octet scalars: the anomalous scalar coupling $k_s^2 = 1/2$ and the characteristic scale of the interaction $\Lambda = M$, the resonance mass.

- New gauge bosons: SM couplings.

- E_6 diquarks: the diquark couplings are set to the EM coupling $\alpha_\lambda = \alpha_{\lambda_c} = \alpha_e$.

- DM: couplings $g_q = 0.25$ and $g_{DM} = 1.0$.

Past searches at CMS used $k_s^2 = 1$ for the color-octet scalar benchmark model [20], which has twice the width and cross section as the model used here.

Table 2.1: Summary of benchmark models with symbol, spin and parity J^P, color multiplet, and channels that the resonance can decay to.

Resonance	Symbol	J^P	Color multiplet	Decay channels
Excited quark	q^*	$1/2^+$	triplet	qg
Axigluon	A	1^+	octet	$q\bar{q}$
Coloron	C	1^-	octet	$q\bar{q}$
RS graviton	G	2^-	singlet	$q\bar{q}$, gg
E_6 diquark	D_6	0^+	triplet	ud
Color-octet scalar	S_8	0^+	octet	gg
Heavy W	W'	1^-	singlet	$q\bar{q}$
Heavy Z	Z'	1^-	singlet	$q\bar{q}$
String resonance	S	various	various	$q\bar{q}$, qg, gg
DM mediator	Φ	$1^+, 1^-$	singlet	$q\bar{q}$

2.4.11 X750 models

A 750 GeV diphoton resonance was not a widely expected signal of new physics. Subsequent to the announcement of the 2015 diphoton excess by CMS and ATLAS, over 500 theory papers were produced—some released quickly after the announcement [2]. Because photons are spin-1 particles, a diphoton resonance must either have spin 0 or spin 2. Diphoton searches were originally optimized to search for the Higgs boson, which has spin 0 and a mass of 125 GeV. After the observation of the excess, the searches were optimized to look for a 750 GeV signal in new data.

Both experiments had previously searched in this region at 8 TeV and saw no excess at the time. Those results were combined with the 13 TeV ones. The absence of an excess

at lower energies pointed towards a resonance that was produced by gluon–gluon fusion rather than production from quarks. As the center-of-mass energy rises, the PDFs favor gluons. At 13 TeV, more of the interactions come from gluons than they did at 8 TeV. If the resonance were produced by gluons, this would explain seeing a signal only at the higher energy.

Figure 2.4 shows the allowed region for the diphoton resonance. It is constrained on one side by the assumption that the resonance couples only to gluons and photons, so $\Gamma = \Gamma_{gg} + \Gamma_{\gamma\gamma}$. It is constrained on the other side by the case where $\Gamma_{\gamma\gamma} \gg \Gamma_{gg}$, so $\Gamma_{\gamma\gamma} \approx \Gamma$. The ratio $\Gamma/M \approx 0.06$ is set by data. Another boundary is set by the ratio of the cross sections at 13 TeV to 8 TeV σ_{13}/σ_8. Lower ratios are disfavored by the absence of an excess in the 8 TeV searches. This provides a region where a diphoton resonance could live [2, 52]. Limits have been set on it by the dijet searches at 8 TeV.
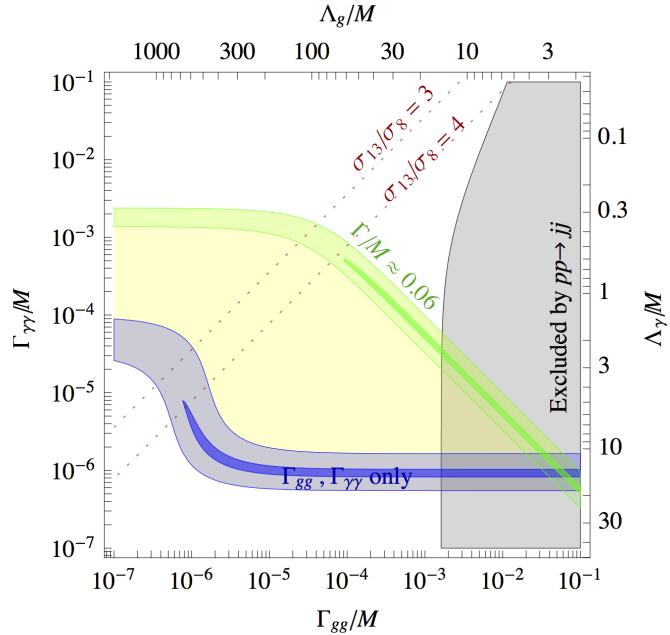
Figure 2.4: Plot of partial width to photons $\Gamma_{\gamma\gamma}$ versus partial width to gluons Γ_{gg} relative to diphoton resonance mass M. Blue band is $\Gamma = \Gamma_{gg} + \Gamma_{\gamma\gamma}$, the green band is $\Gamma \approx \Gamma_{\gamma\gamma} = 0.06M$, and the grey region is excluded by 8 TeV dijet searches assuming coupling only to gluons and photons. The allowed region for the resonance is in yellow. Dashed lines show constant ratios of the cross section at 13 TeV to that at 8 TeV. From Ref. [52].

Another explanation for an excess at 13 TeV but none at 8 TeV is coupling to bottom quarks [52]. Like gluons, their PDF will also increase at higher center-of-mass energy.

If there were a particle at 750 GeV that decays to photons, its properties would be determined by studying the diphoton signal and looking for decays in other channels such as the dijet one.

Chapter 3

Experiment

3.1 Large Hadron Collider

The Large Hadron Collider (LHC) is a synchrotron at the European Organization for Nuclear Research (CERN). It is located approximately 100 m under the ground and extends from Geneva, Switzerland across the border into France over a circumference of 26.7 km. It produces proton–proton (pp) collisions at center-of-mass energy $\sqrt{s} = 13$ TeV. It is also capable of lead–lead and proton–lead collisions. A diagram of the LHC and the CERN accelerator complex is shown in Fig. 3.1.
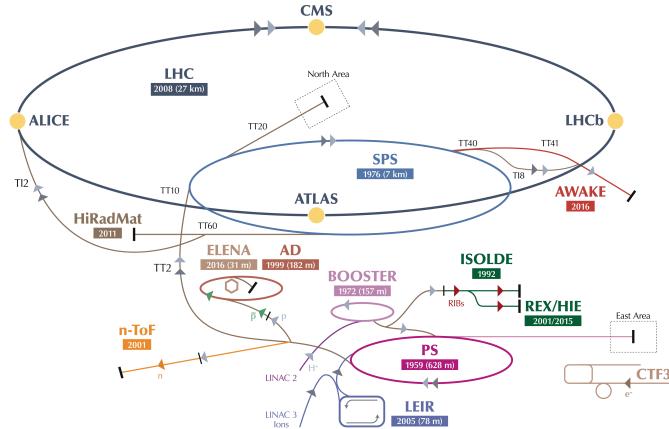
Figure 3.1: CERN accelerator complex with the LHC main ring, injector chain, and experiments labeled. Accelerators in the complex used by other experiments are also shown.

The protons are obtained from a tank of hydrogen gas. The electrons are stripped from the hydrogen atoms leaving protons. The protons are accelerated over a series of stages to bring them to the final energy [53]. The first stage is the linear accelerator

LINAC 2, which gives the protons 50 MeV of energy. LINAC 2 was first built in 1978. Because the ability of linear accelerators to add energy is limited by their lengths, all of the other stages use synchrotrons to accelerate bunches of protons as they pass multiple times through the same components, each time increasing the energy further. Multiple synchrotrons must be used because they cannot reliably operate at the low magnetic fields required to keep low-energy protons in a ring of that size. Larger rings are needed to reduce the magnetic field required by the dipole magnets to keep the protons circulating. This sets upper and lower limits to the energy at which each synchrotron can circulate proton beams.

After leaving the linear accelerator, the protons enter the Proton Synchrotron Booster, which began operation in 1972. It allows the next stage, the Proton Synchrotron (PS), to accept more bunches than is possible at a starting energy of 50 MeV. Four synchrotron rings accelerate the protons to an energy of 1.4 GeV after which they are injected into the PS. The PS accelerates the protons up to an energy of 26 GeV with a spacing between proton bunches of 25 ns. It was the first synchrotron built at CERN in 1959 and the highest-energy accelerator in the world at the time. The PS has a circumference of 628 m and uses 277 room-temperature electromagnets including 100 dipole magnets.

The final accelerator before the protons reach the main ring of LHC is the Super Proton Synchrotron (SPS). It was commissioned in 1976. The SPS has 1317 room-temperature electromagnets including 744 dipole magnets along the nearly 7 km circumference ring. While operating as a proton–antiproton collider, its UA1 and UA2 experiments discovered the W and Z bosons 1983. The SPS accelerates the bunches of protons up to an energy of 450 GeV, which the the LHC main ring can receive [54].

The LHC was built in the tunnel of the Large Electron–Positron Collider (LEP). LEP operated from 1989 through 2000, when it was decommissioned to begin construction on the LHC. Originally envisioned as sharing the tunnel with LEP, the first official meeting about the LHC took place in 1984 [55]. The LHC was approved in 1994 and

began operation in 2008. Shortly afterward, an electrical fault between two dipole magnets lead to the explosive vaporization of two metric tons of liquid helium. After repairs, the collider became the highest energy accelerator in the world in 2010. It later collected data at 7 TeV and 8 TeV. Starting in 2015 and continuing in 2016, the LHC circulated beams at 6.5 TeV for a center-of-mass energy of 13 TeV.

Protons are kept circulating in the ring by 1232 superconducting dipole magnets that are 14.3 m in length. When the protons are injected at 450 GeV, the dipole magnets have a field of 0.535 T. When they will guide 7 TeV beams ($\sqrt{s} = 14$ TeV) the magnets will operate at 8.33 T. At 6.5 GeV, 7.7 T is required. The dipole magnets have a bending radius of 2803.95 m. The current in the dipole coils is 13 kA.

Two magnetic fields are required to bend the two beams of protons moving in opposite directions because particles in both beams have the same charge. Because of limited space in the tunnel, the bending magnets are two-in-one dipole magnets, producing an upward magnetic field through one beam pipe and a downward magnetic field through the other. The current flowing in opposite directions through parallel wires creates a force of up to 1.7 MN/m on each quadrant trying to pull the magnet apart.
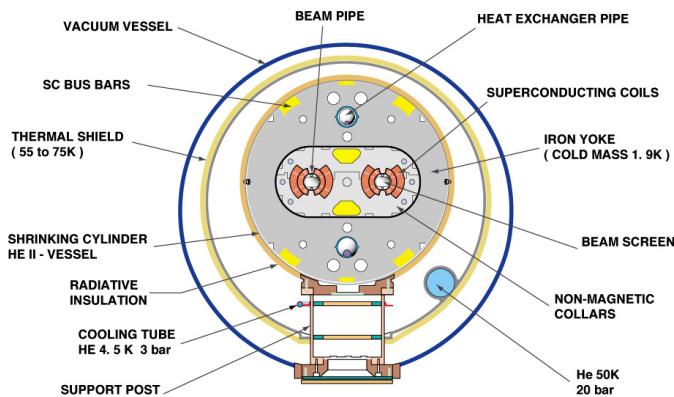
Figure 3.2: Cross section of dipole magnet with labeled components. The beams circulate through the two beam pipes in the center of the magnet.

To achieve the currents needed to produce a strong enough magnetic field, niobium-titanium (NbTi) wire was used. The magnets are cooled to 1.8 K, at which temperature

the wire is superconducting. While NbTi has a critical temperature of 9.2 K, it is cooled further to allow the magnets to achieve a higher magnetic field. The liquid helium used for cooling is a superfluid at 1.8 K, which aids in heat conduction. The total cold mass of the LHC is 36 800 metric tons.

The beam is focused by quadrupole magnets. The quadrupole focuses along one axis while defocusing along the other. By alternating quadrupole magnets that focus in the horizontal direction while defocusing in the vertical direction with magnets that focus in the vertical direction while defocusing in the horizontal direction, the beam can be confined. The focusing and defocusing produces betatron oscillations as the beam's envelope expands and contracts. The oscillations are tuned to avoid a resonance. The beams make 64.31 horizontal and 59.32 vertical oscillations for every orbit of the ring. The beams are focused tightly at the interaction points (IPs) to maximize the luminosity.

The protons are accelerated by radio-frequency (RF) cavities. An RF cavity is a metallic chamber that acts as a resonator for microwaves. Microwaves at 400 MHz are injected into the cavity and build up until each one has an electric field with up to a 2 MV electric potential difference. The strength of this field and direction vary with time as the standing wave inside the RF cavity oscillates. If the bunch passes through when there is no electric field, the energy of the protons will remain the same. If the electric field is present as the protons pass through the cavity, they will be accelerated or decelerated depending on the direction of the electric field. This feedback keeps the protons in the bunches together. It is also used in accelerating the protons up to their final energy by timing their arrival so the electric field is pointing in the forward direction as they pass. The LHC has 16 RF cavities—eight for each beam.

The beams contain 2800 bunches of 10^{11} protons with a spacing between bunches of 25 ns. The total energy in each beam is 362 MJ. The beam is less than $1\,\text{mm}^2$ in area. After the beams are injected and accelerated to the final energy, the collider enters a regime of stable beams when the beams are crossed to start collisions at the IPs inside the detectors. Collisions are maintained until the luminosity is reduced roughly in half.

This takes from several hours up to over a day. The beam pipe is maintained at a vacuum equivalent to 10^{15} hydrogen molecules per cubic meter to limit the collisions between the protons in the beam and gas in the pipe. The vacuum is kept at 10^{13} H_2 per cubic meter near the IPs to minimize beam–gas interactions. This is equivalent at room temperature to 10^{-8} to 10^{-9} Pa [54].

The rate of collisions in a collider

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \sigma_{int}\mathscr{L},$$ (3.1)

where σ_{int} is the cross section of the interaction and $\mathscr{L}$ is the luminosity. Cross section is measured in units of area. The instantaneous luminosity

$$\mathscr{L} = f_{rev}\frac{n_b N_b^2}{4\pi\sigma_x\sigma_y}R$$ (3.2)

is dependent on the revolution frequency f_{rev}, number of bunches in each beam n_b, number of particles in each bunch N_b, the transverse root-mean-square of the beam size σ_x and σ_y, and a reduction factor R that depends on the property of the beams and their crossing angle [32]. The amplitude function β is a parameter in the beam's betatron oscillations with β^* being the value at the IP. It is dependent on the accelerator's beam optics. The emittance is a measure of the spread of the beam in distance and momentum. The transverse emittance for beams with a gaussian distribution of particles

$$\epsilon_x = \frac{\sigma_x^2}{\beta_x} \qquad \text{and} \qquad \epsilon_y = \frac{\sigma_y^2}{\beta_y}.$$ (3.3)

The luminosity can be rewritten in terms of these beam parameters

$$\mathscr{L} = f_{rev}\frac{n_b N_b^2}{4\pi\sqrt{\epsilon_x\beta_x^*\epsilon_y\beta_y^*}}R(\theta_c, \epsilon_i, \beta_i^*, \sigma_z),$$ (3.4)

where θ_c is the crossing angle between the beams and σ_z is the root-mean-square of the

proton bunch in the longitudinal direction. The luminosity can be raised by increasing the beam current $f_{\text{rev}} n_b N_b$, increasing the brightness $N_b/\sqrt{\epsilon_x \epsilon_y}$, decreasing β^*, or reducing the loss from the reduction factor with the crossing angle and other beam effects. For the LHC, $\beta^* = 0.8\,\text{m}$ and $\epsilon_x = \epsilon_y = 0.5\,\text{nm}$.

We measure the amount of data collected in terms of integrated luminosity L as it easily gives us the number of events by integrating Eq. (3.1)

$$N = \sigma \int \mathscr{L}\,\mathrm{d}t = \sigma L. \tag{3.5}$$

This is true not only for the total number of collisions and σ_{int} but also for the number of expected events for a process with any cross section. Cross section and integrated luminosity are typically measured in the unit barn and inverse barn, respectively, where $1\,\text{b} = 10^{-28}\,\text{m}^2$ is roughly the area of an atomic nucleus.

The LHC collides particles for four main experiments. The two general-purpose detectors are CMS and ATLAS. ATLAS is located in Switzerland, roughly the other side of the ring from CMS. It searches for the same physics that CMS does. Consequently, its general design is similar to that of CMS and detectors at the Tevatron and SPS. However, the ATLAS collaboration decided on different priorities and made different choices, so the particular detectors and magnets are different from what was used in CMS. The independent designs of the two detectors and independent analyses from the two collaborations provide a way to check the results of both experiments.

Two smaller collaborations also rely on collisions from the LHC. The LHCb experiment is a bottom quark physics experiment that probes CP violation.[1] Unlike CMS and ATLAS, LHCb is not hermetic and is designed to study the forward region in only one direction. ALICE (A Large Ion Collider Experiment) studies heavy-ion lead–lead collisions. There are also three other small experiments that use the IPs of the four main experiments.

---

[1]CP is a symmetry under flipping charge C (matter to antimatter) and parity P (opposite of all spacial dimensions).

3.2 Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is a general-purpose detector designed to measure particles from proton–proton and heavy-ion collisions. Its main feature is a superconducting solenoid that provides a magnetic field for the experiment. Inside the solenoid are silicon pixel and strip trackers, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL). A forward calorimeter (HF) covers high pseudorapidity and gas-ionization detectors are interspersed with the steel magnetic return yoke.

The CMS is located in Cessy, France. The cavern it resides in is approximately 100 m beneath the ground, in line with the LHC's beam pipe. The detector's letter of intent from 1992 states [56]:

> We propose to build a general purpose detector designed to run at the highest luminosity at the LHC. The Compact Muon Solenoid (CMS) detector has been optimized for the search of the SM Higgs boson over a mass range from 90 GeV to 1 TeV, but it also allows detection of a wide range of possible signatures from alternative electro-weak symmetry breaking mechanisms. CMS is also well adapted for the study of top, beauty and tau physics at lower luminosities and will cover several important aspects of the heavy ion physics programme. We have chosen to identify and measure muons, photons and electrons with high precision.

The CMS detector has accomplished these goals. In 2012, it and ATLAS discovered the Higgs boson at 125 GeV [57]. The detector was designed to provide accurate measurements of photons to discover the Higgs boson in the diphoton final state and accurately measure leptons. These considerations lead to the quality of the tracker and the electromagnetic calorimeter at the expense of the hadronic calorimeter (N.B., jets were not included in the list of objects to be measured with high precision). However, CMS was designed as a general-purpose detector capable of searching for new physics at the EW energy scale, including with jets.

Bunches of protons collide inside CMS every 25 ns. The detector must measure its channels at this rate. However, the response of parts of the detector to particles takes

longer. This leads to out-of-time pileup (PU), where the effect of previous events are still being measured more than 25 ns later. Additionally, there are tens of interactions during each crossing. This is in-time PU, where the results of many interactions occur effectively at once inside the detector. The interactions helpfully occur at different vertices along the beam pipe. It is highly unlikely that more than one of these interactions is of interest to us, they just provide additional noise. The detector and its software are designed to mitigate the effects of both kinds of PU. PU could be almost completely eliminated by using different run conditions, but we have chosen to operate in this regime to reach larger luminosities.

Figure 3.3 shows a cutaway of CMS with the main subdetectors labeled. A cross section of the detector and how particles interact with the subdetectors is shown in Fig. 3.4. A full description of the detector can be found in Ref. [58].
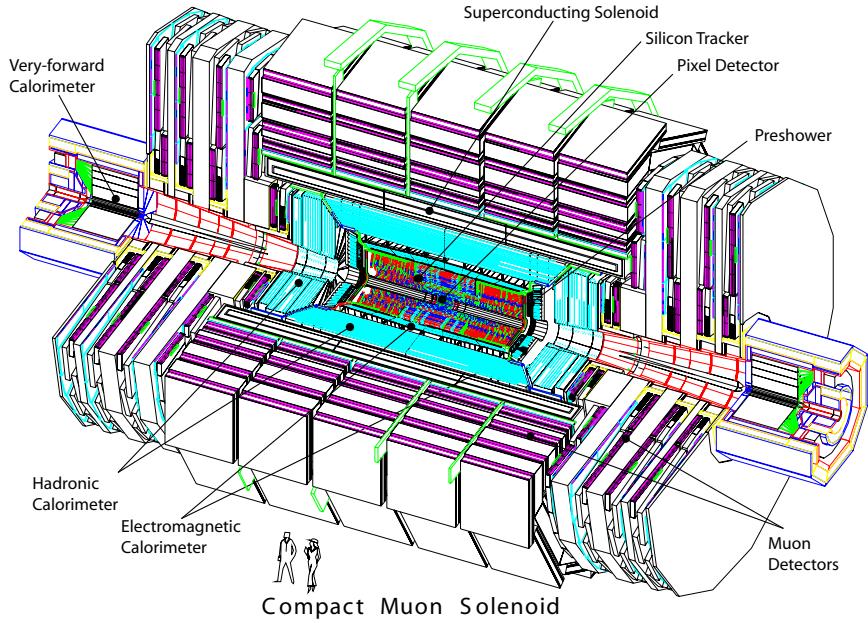
Figure 3.3: Cutaway of the the CMS detector with labeled subdetectors. Humans are shown for scale.
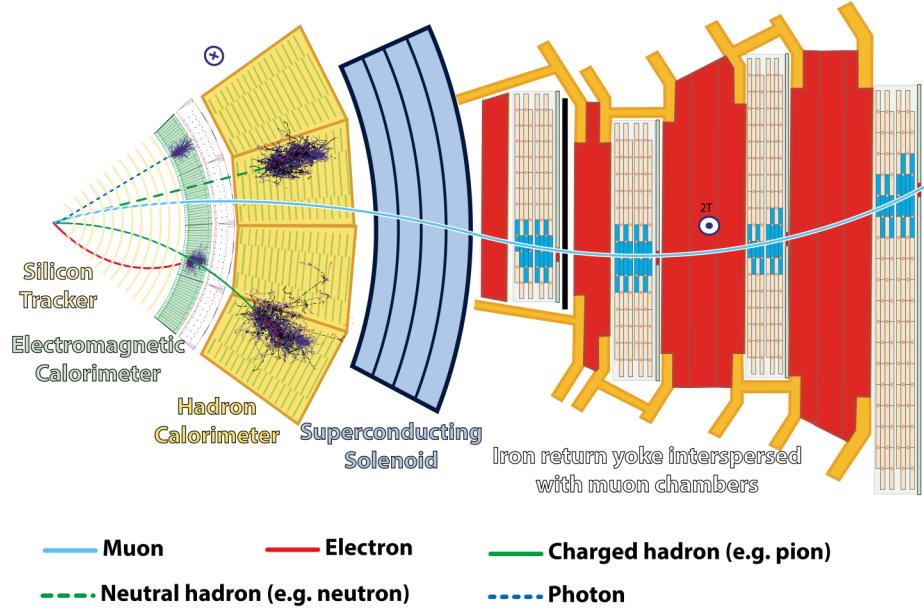
Figure 3.4: Transverse section of the CMS detector showing the interaction of particles with the sub detectors. Particles with solid paths produce tracks in the silicon tracker—dashed paths do not. The magnetic field points into the page on the left and out of the page on the right.

3.2.1 Coordinate system

CMS uses a right-handed coordinate system. The x axis points radially inward to the center of the main ring, the y axis points upward, and the z axis points in the counterclockwise direction when the collider is viewed from above. The azimuthal angle ϕ is measured in the x-y plane with $\phi = 0$ along the positive x axis. The polar angle is measured from the positive z axis.

Pseudorapidity η is a more useful parametrization of the polar angle:

$$\eta = -\log\left(\tan\frac{\theta}{2}\right) = \frac{1}{2}\log\left(\frac{|p| + p_{\mathrm{L}}}{|p| - p_{\mathrm{L}}}\right), \tag{3.6}$$

where $p_{\mathrm{L}} = p_z$ is the longitudinal momentum. Differences in pseudorapidity $\Delta\eta$ are invariant under longitudinal Lorentz boosts. Particle production in a hadron collider is roughly constant in η. The longitudinal momentum of the interacting partons can vary, boosting the products in the longitudinal direction. Figure 3.5 shows values of η relative

to the CMS detector.
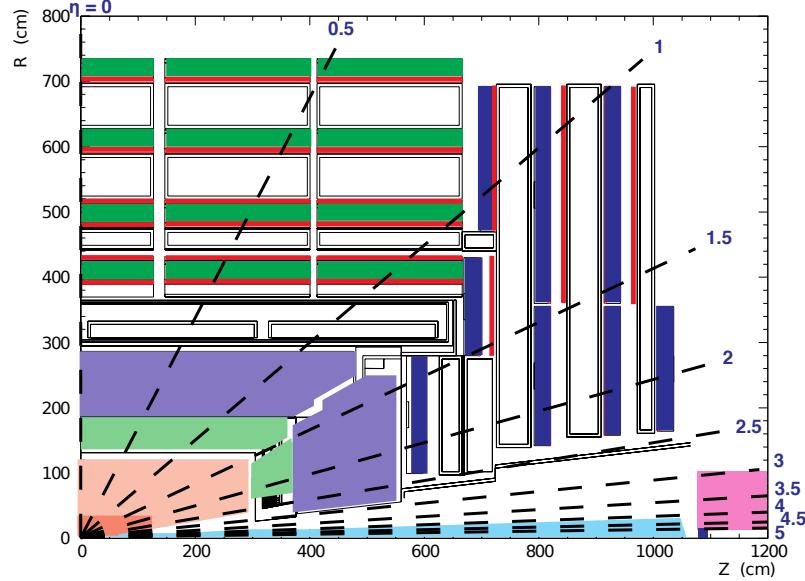
Figure 3.5: Schematic of a quarter of CMS in the r-z plane with constant pseudorapidity along dashed lines.

Rapidity is defined as

$$y = \frac{1}{2} \log\left(\frac{E + p_{\mathrm{L}}}{E - p_{\mathrm{L}}}\right). \tag{3.7}$$

The pseudorapidity η is approximately equal to the rapidity y in the limit where $p \gg m$ and $\theta \gg 1/\gamma$ (for a massless particle, they are identical). In special relativity, rapidity is defined as

$$w = \frac{1}{2} \log\left(\frac{E + |p|}{E - |p|}\right). \tag{3.8}$$

It is the hyperbolic angle between two references frames. The Lorentz factor $\gamma = \cosh w$. Boosting by w transforms between the lab frame and the particle's rest frame. Boosting by y transforms between the lab frame and the frame in which the particle is only moving transverse to the beam. Pseudorapidity is a close approximation for the relativistic particles studies here and relies only on knowing where the particle was in the detector.

The protons from opposing beams have momenta equal in magnitude and nearly opposite in direction. However, the momentum shared between the partons is randomly distributed according to PDFs, which can produce large boosts along the longitudinal

direction if they are disparate. There is little momentum transverse to the beams. This makes it useful to define transverse quantities such as the transverse momentum

$$p_{\mathrm{T}} = p_x + p_y = |\boldsymbol{p}| \sin\theta = |\boldsymbol{p}| \cosh\eta \tag{3.9}$$

and transverse energy

$$E_{\mathrm{T}} = E \sin\theta = E \cosh\eta. \tag{3.10}$$

Direction is measured in η-ϕ space with ϕ in radians. The distance between two points in this space is

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{3.11}$$

With approximately no p_{T} in the incoming partons, conservation of momentum requires that the vectorial sum of p_{T} from all outgoing particles $\sum_i p_{\mathrm{T}i} = 0$. The CMS detector is designed to be hermetic to limit how many particles can escape detection and throw this calculation off by carrying away transverse momentum. Nonetheless, it is impossible to create an entirely hermetic detector. The beam pipes must be left clear for the beams to enter, allowing particles to leave undetected at high $|\eta|$, which helpfully means smaller p_{T} for a particle of given energy. The detector does also have gaps to let necessary components through the layers. The difference is measured by the missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$, which is the magnitude of the vectorial sum

$$\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_i \boldsymbol{p}_{\mathrm{T}i}. \tag{3.12}$$

Particles such as neutrinos cannot be measured by the CMS detector. They carry away some p_{T} resulting in nonzero $p_{\mathrm{T}}^{\mathrm{miss}}$ (unless there is another undetectable particle carrying aware equal p_{T} in the opposite direction). This quantity is also called missing transverse energy (MET).

3.2.2 Magnet

The eponymous feature of CMS is its superconducting solenoid magnet. The requirement that the tracker and calorimeters fit inside the magnet forces the compactness of the detector. The muon system is one of the few parts of the detector to lie outside of the solenoid.

The solenoid is 12.9 m in length with an inner diameter of 5.9 m. It is made from 2168 turns of NbTi. The magnet is designed for a current of 19.5 kA producing a field of 4 T inside the solenoid. To increase longevity, the current has been run at 18.16 kA, producing a magnetic field of 3.8 T. The magnet stores 2.3 GJ of energy.

The magnetic field outside of the solenoid is controlled by a steel return yoke that is interspersed with the muon system. There is more iron in the yoke than in the Eiffel Tower. Although ATLAS is significantly larger in volume (it is not compact), CMS is nearly twice as massive.

Charged particles moving through this magnetic field will experience a Lorentz force (unless it is moving parallel to the longitudinal magnetic field). The transverse momentum of a charged particle with charge $\pm e$ is

$$p_{\mathrm{T}} = \left(0.300\,\mathrm{GeV\,m^{-1}\,T^{-1}}\right) B\rho, \tag{3.13}$$

where B is the magnetic field and ρ is the radius of curvature of the particle's path.

3.2.3 Inner tracking system

The innermost subdetector in CMS is the inner tracking system, which detects the passage of ionizing particles. Its primary purpose is to reconstruct the momentum of the particles. Measuring the bending radius of the reconstructed track of a charged particle gives the particle's p_{T} from Eq. (3.13). The tracks also enable the reconstruction of vertices and indicate which vertex an object comes from. Vertices are points that tracks emerge from. This is used both for the primary vertex (PV), where the hard interaction

occurred, and secondary vertices from the decay of B hadrons and other unstable particles. Because only charged particles will be detected by the tracking system, it provides one means of particle identification. The tracker covers $|\eta| < 2.4$.

The tracker is made of silicon detectors. The innermost layers are pixel detectors to handle the high particle flux. Surrounding these are strip detectors. The entire tracker has roughly $210\,\text{m}^2$ of silicon sensors and ten million channels, making it the largest silicon tracker ever built. Figure 3.6 shows a cross section of the tracker.
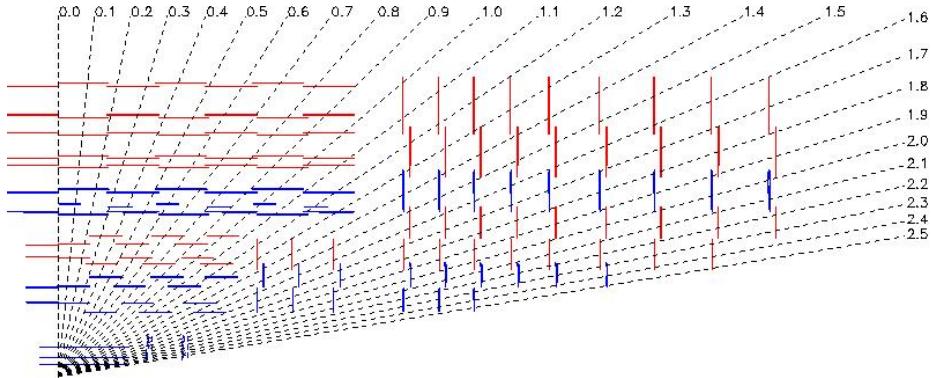
Figure 3.6: Schematic of silicon pixel and strip detectors in a quarter of the tracker in the r-z plane. Single layer strip detectors are in red and double layer "stereo" strip detectors are in blue.

Closest to the point of interaction are three barrel layers and two endcap disks of silicon pixel detectors. The closest layer is at a radius of $4.4\,\text{cm}$. The outermost barrel of pixels is at $10.2\,\text{cm}$ and the outermost disks are at $|z| = 46.5\,\text{cm}$. An ionizing particle passing through the silicon will strip electrons out of the atoms. The electrons and holes are separated by the electric field of the p–n junction. The holes move to the pixels of p-type silicon where the current is read out. Despite a charged particle flux of $10^8\,\text{cm}^{-2}\,\text{s}^{-1}$ at the innermost layer, the pixels have an occupancy of 0.01% for pp collisions. The pixels cover an area of approximately $1\,\text{m}^2$.

Outside of the pixel detectors is the tracker inner barrel (TIB), which has four layers. The strips are $10\,\text{cm}$ long, $320\,\mu\text{m}$ thick, and with a pitch that varies between 80 and $120\,\mu\text{m}$. The first two layers are "stereo" modules made from overlapping silicon strip

detectors oriented at an angle of 100 mrad (5.7°), which allows the particle's track to be localized to a smaller region in z than the 10 cm length of the strip. The resolution is between 23 and 34 μm in ϕ and 230 μm in z.

The outermost layer of the tracker is the tracker outer barrel (TOB), which extends to a radius of 1.1 m. The lower radiation allows a thicker 500 μm strip that is 25 cm long and has a pitch between 120 and 180 μm. It is comprised of six layers with the first two made from "stereo" modules. These modules have a resolution of 35–52 μm in ϕ and 530 μm in z.

The tracker inner disk (TID) is the endcap for the TIB. It consists of three disks. The innermost two rings of detectors in the TID are "stereo" modules. The endcap of the TOB is the tracker end cap (TEC), which is made of nine disks out to $|z| = 280$ cm. The first two rings and the fifth ring are "stereo" modules.

To reduce radiation damage, the silicon is operated at -20 °C.

3.2.4 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) surrounds the inner tracking system. It is comprised of two parts, the barrel (EB) and endcap (EE). The EB covers pseudorapidity $|\eta| < 1.479$ and the EE covers $1.479 < |\eta| < 3.0$. A third component, the preshower (ES), sits in front of the EE covering the range of $1.653 < |\eta| < 2.6$. It helps in identifying neutral pions and the direction of electrons and photons through the endcaps. Figure 3.7 shows a schematic of the ECAL's design.

The ECAL is a homogenous calorimeter that uses lead tungstate ($PbWO_4$) crystals. The crystals have a radiation length $X_0 = 0.89$ cm, which is the average distance an electron radiates all but 1/e of its initial energy through bremsstrahlung and 7/9 of the mean free path for pair production of a high-energy photon. The crystals in the EB have a length of 230 mm or $25.8X_0$. The short radiation length of $PbWO_4$ allows a large number to fit within the confines of the solenoid magnet. In the EE, the crystals are $24.7X_0$ long.
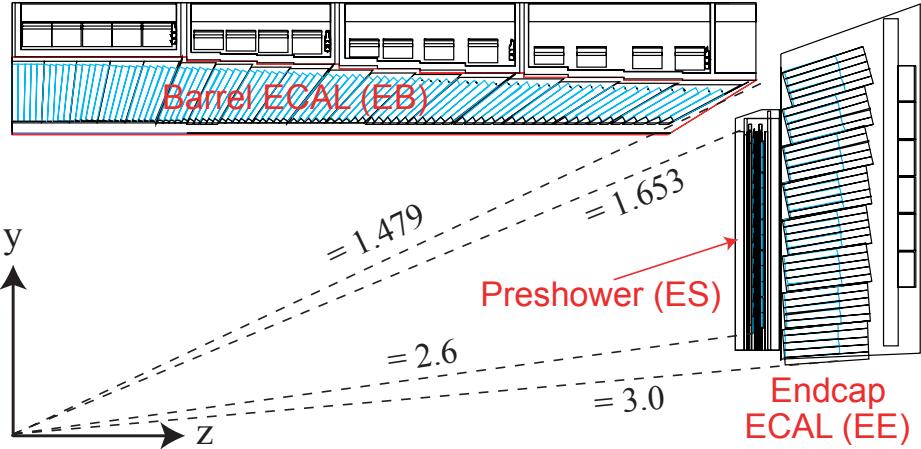
Figure 3.7: Transverse quarter of the ECAL in the r-z plane.

The crystals are truncated pyramids. In the EB, each crystal covers roughly 0.0174×0.0174 in η-ϕ (360 segments in ϕ and 85 on each side of the detector in η). The crystals are $22\,\mathrm{mm} \times 22\,\mathrm{mm}$ on the front face and $26\,\mathrm{mm} \times 26\,\mathrm{mm}$ at the rear. In the EE, they are $28.62\,\mathrm{mm} \times 28.62\,\mathrm{mm}$ on the front face and $30\,\mathrm{mm} \times 30\,\mathrm{mm}$ on the rear face. The width of the crystals is comparable to the $22\,\mathrm{mm}$ Molière radius of the material. The Molière radius is a measure of the characteristic transverse size of an EM shower. Approximately 90% of the shower's energy is contained in one Molière radius.

The crystals are arranged in a quasi-projective geometry as can be seen in Fig. 3.7. They are oriented 3° off axis in both directions from a truly projective geometry. The crystals are organized into 5×5 groups called supercrystals.

All but one side of the crystal are polished to improve the total internal reflection of light. The light collection is not uniform across the length of the truncated pyramid, so one polished side is depolished the appropriate amount along the crystal's length to make light collection uniform. Scintillation light is detected by avalanche photodiodes in the EB, which are affixed to the rear, unpolished faces of the crystals. The EE uses vacuum phototriodes to detect scintillation light. The amount of scintillation light varies by $-1.9\%/1\,\mathrm{K}$ at $18\,^{\circ}\mathrm{C}$, so the ECAL contains a cooling system to maintain its temperature.

The preshower detector is a two-layer sampling calorimeter. It uses silicon strip

detectors behind lead absorber to measure the shower. At $\eta = 1.653$, the first layer of lead is $2X_0$ followed by a second layer of $1X_0$.

Energy resolution σ_E/E can be divided into three terms [59]. Shower development is stochastic with the uncertainty in the measurement increasing as the square root of the scintillated photons observed. This makes the energy resolution dependent on a stochastic term proportional to $1/\sqrt{E}$ after the uncertainty has been divided by E. There is electronic noise present in the readout of the calorimeter. This does not scale with energy, so the noise term is proportional to $1/E$. Finally, there is a constant term that captures effects that scale proportionally with the energy in the shower. The energy resolution in the EB

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{12\%}{E} \oplus 0.3\%, \tag{3.14}$$

where the terms are added in quadrature, was measured with test beams [60, 61].

3.2.5 Hadron calorimeter

The hadron calorimeter (HCAL) is the outermost part of the detector that still lies within the magnetic solenoid. It is comprised of two parts: the barrel (HB) covers $|\eta| < 1.4$ and the endcap (HE) covers $1.3 < |\eta| < 3.0$. The calorimeter is designed to be hermetic, hence the overlap of the two sections. The HCAL spans a solid angle of 3.98π sr. Outside the magnetic solenoid is the outer hadron calorimeter (HO), which covers $|\eta| < 1.26$.

The HCAL is a sampling calorimeter. The HB consists of 17 layers of plastic scintillator interspersed with absorber. The absorber in all but two layers is made from C26000 cartridge brass with 70% copper and 30% zinc. The interaction length, the average distance for the shower to lose all but $1/e$ of its energy, for the brass is $\lambda_I = 16.42$ cm. The first and last layer of absorber are made from stainless steel for structural strength. At $\eta = 0$, the HB is $5.82\lambda_I$ thick. At $|\eta| = 1.3$, the thickness increases to $10.6\lambda_I$. The ECAL is approximately $1.1\lambda_I$ [62].

The front of the HB is an active layer of scintillator to measure the shower from the material in between the ECAL and HCAL. It is 9 mm of Bicron BC408 plastic scintillator. It and the final active layer are thicker than the intermediate layers to oversample the early and late shower. The first layer of absorber is 61 mm of stainless steel. The intermediate active layers are 3.7 mm thick Kuraray SCSN81 plastic scintillator. The first nine are separated by 50.5 mm thick brass plates. The remaining ones are separated by 56.5 mm thick brass plates except for the last layer of absorber, which is 75 mm of stainless steel. Behind the stainless steel for all but the most forward tower is 9 mm of Kuraray SCSN81 plastic scintillator.

The scintillator plates are segmented in η and ϕ into tiles with $\Delta\eta = 0.087$ and $\Delta\phi = 0.087$. Outside of $\eta = 1.740$, the size of towers in the HE is doubled to $\Delta\phi = 0.175$ with $\Delta\eta$ also increasing. The segmentation of the HCAL into layers and towers is shown in Fig. 3.8.
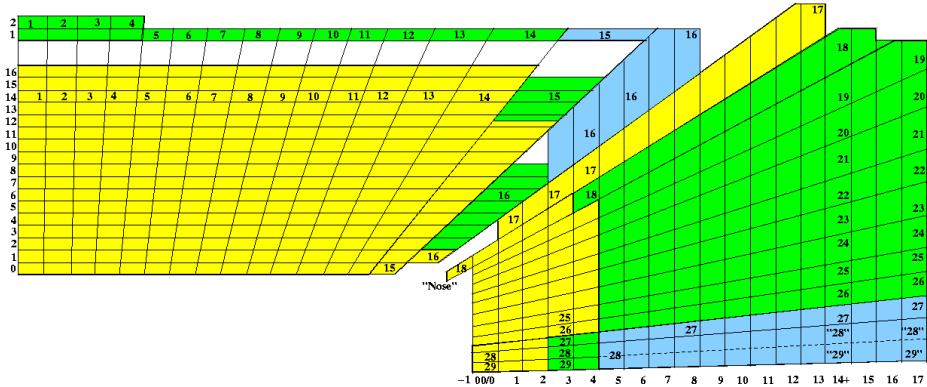
Figure 3.8: Schematic of the towers and layers in a quarter of the HCAL in the r-z plane. The HB extends to the front of tower 16. The HE starts in the rear half of tower 16. The HO is the outermost layer or two through tower 15.

Scintillated light is collected by a wavelength shifting fiber that runs through grooves in the scintillator tiles. The 0.94 mm diameter fiber is double-clad and made from Kuraray Y11. Without the wavelength shifting fiber absorbing light and isotropically emitting it at a different wavelength, light from the scintillator would pass through the embedded fiber without being funneled out to be detected. Some of the time, the light will

be emitted at an angle where there is total internal reflection. Only that light will travel along the fiber—the rest is not measured. One end of the fiber is mirrored. The other is spliced to clear fiber to transport the light to the photodetectors. The light is optically added into towers and measured by hybrid photodiodes.

The photodetectors used in the HCAL and other subdetectors in the experiment produce an analog electrical signal proportional to the number of photons they receive. The electrical signal is integrated and digitized by a charge integrator and encoder (QIE) application-specific integrated circuit, which is an analog-to-digital converter that integrates the charge in capacitors and then converts it into a seven-bit pseudo-logarithmic scale. The encoded data from several QIE channels are combined and sent out of the experiment with an optical link.

The HO is located outside the magnet in the return yoke and is divided into five rings like the yoke and muon system. A 10 mm layer of scintillator is placed behind 18 cm of steel. The central ring ($|\eta| < 0.348$) also contains an additional layer of scintillator in front of the iron absorber. The HO has larger gaps than the HB and HE to allow for the support structure for the experiment. The HO increases the depth of the calorimeter to $11\lambda_I$ in the central region of the detector.

Like the resolution of the ECAL, the energy resolution of the HCAL can be broken into terms. The HCAL measures less energy from the hadronic component of showers than it does from the EM component, which reduces the energy resolution of the detector. This difference is quantified as the response to the EM component divided by the response to the hadronic component e/h. For the HCAL, $e/h = 1.4$. The ECAL has $e/h = 1.6$. Neither is a compensating calorimeter (one designed such that $e/h \approx 1$), so the measured energy depends on how the shower evolves. The energy resolution of test-beam pions in the HCAL was measured to be [63]

$$\frac{\sigma_E}{E} = \frac{115\%}{\sqrt{E}} \oplus 5.5\%. \tag{3.15}$$

3.2.6 Forward calorimeter

The forward calorimeter (HF) is located 11.2 m away from the IP. It covers a pseudo-rapidity of $2.9 < |\eta| < 5$, extending the calorimeters to cover 3.9996π sr of solid angle. It is a sampling calorimeter made from steel with quartz fibers that produce and collect Cherenkov light. The fibers lie parallel to the z axis in a non-projective geometry.

The quartz fibers are divided into two types: long and short lengths—1.65 m and 1.43 m, respectively. The fibers are aligned in the back, so the long fibers reach closer towards the front of the absorber than the short fibers, which do not begin until 22 cm into the absorber. Showers from EM particles will be absorbed more quickly than those from hadrons, so comparing the energy sampled from the full length of the HF (long fibers) to that sampled from back portion (short fibers) provides a handle for particle identification. The total depth is approximately $10\lambda_I$.

The fibers are spaced 5 mm on center in a square grid with alternating long and short fibers. The Cherenkov light is funneled out through the quartz fibers, passes through light guides, and is measured by photomultiplier tubes.

3.2.7 Muon system

The muon system lies outside the magnetic solenoid where it is interspersed with an iron magnetic return yoke. It consists of three types of gas detectors to measure muons, which generally travel through the inner layers of the detector and pass through this outermost system. The muon system is divided into a barrel for $|\eta| < 1.2$ and an endcap that extends to $|\eta| < 2.4$.

The barrel muon detector consists of four concentric layers called "stations" and is divided along the z axis into 5 "wheels". The two innermost stations are made of a drift tube (DT) chamber surrounded in the front and rear by two resistive plate capacitors (RPCs). The outer two stations are made from a DT chamber and 1, 2, or 4 RPCs.

A DT is a variation of a wire chamber where each anode wire is in a separate cell

with cathode strips on either side. The spacing between the cathode partitions is 4.2 cm. There are also anode strips on the other two sides of the cell. The cells are filled with a mixture of argon (85%), which becomes ionized, and carbon dioxide (15%), which quenches the avalanche of ionization.

Four layers of DTs form a super layer (SL). The DTs in successive layers are shifted by half a tube. The first three stations are made from three SLs. The innermost SL is aligned so the wires are parallel to the beam axis to allow it to measure the azimuthal position of the muon. A spacer separates this from the two other SLs to give a longer lever arm. The second SL is oriented with the wires perpendicular to the beam to measure the position of the track in η. The outermost SL is oriented like the innermost one to measure the track in ϕ. The outermost DT chamber is composed of 2 SLs that are both oriented to measure the track in ϕ.

The RPCs are made from two similar chambers. Both feature two plates of bakelite plastic separated by a gap filled with gas. Readout strips are placed on the back side of the chambers.

The endcaps are also divided into four stations. The stations are composed of cathode strip chambers (CSCs) along with RPCs in the outer rings of the stations. Each CSC is made from seven trapezoidal planes with cathode strips running in the radial direction that create six gas gaps. In between the planes are wires that run in the azimuthal direction. The wires provide timing data to locate the track in r, the strips provide position in ϕ, and the planes give the position in z.

3.2.8   Trigger system

The LHC collides protons inside CMS at a rate of 40 MHz (once every 25 ns). The average size of the raw data from a recorded event is 630 kB. CMS does not have the ability to save approximately 25 TB/s nor the storage space for exabytes of data. Neither does the experiment have the resources to reconstruct the raw data at this rate.

However, most events contain common processes that have been studied in other

experiments. That allows us to select only those events likely to contain new physics or SM physics of interest to the members of the CMS collaboration.

CMS uses a two-level trigger system to decide which events should be saved [64]. The level-1 (L1) trigger reduces the rate of events to below 100 kHz using custom electronics. The high-level trigger (HLT) then partially reconstructs the events and makes a final decision on whether the event is saved.

The triggers cannot make decisions within 25 ns. Therefore, the data from each event must be collected from the detector and buffered while each system makes a determination. The L1 trigger has 3.2 µs to decide to keep an event. The whole L1 system is a pipeline with no downtime. A new event enters every 25 ns so a decision must be made in step every 25 ns to keep the pipeline from backing up.

The L1 trigger does not have time to construct tracks from the inner tracking system or perform complex algorithms. It looks at the energy deposited in sections of the calorimeters and hits in the muon system as well as global quantities like the global sum of E_T and $p_\mathrm{T}^\mathrm{miss}$. To pass, these quantities must exceed p_T or E_T thresholds.

Data are sent from the detector in the experimental cavern to a service cavern that contains the electronics used to process the data for the L1 trigger. The trigger uses custom application-specific integrated circuits, field-programmable gate arrays, and programmable logic devices to process the data [64].

The calorimeter trigger takes energy from the ECAL, HCAL, and HF that has been summed from the individual towers by trigger primitive generators. The regional calorimeter trigger finds candidate jets, electron, photons, and taus. The candidates are sent to the global calorimeter trigger, which orders the objects and calculates global E_T and $p_\mathrm{T}^\mathrm{miss}$. It sends the leading four of each object and the energy sums to the global trigger.

The muon trigger is divided into the three types of detectors in the muon system (DT, RPC, and CSC). The trigger finds segments in the individual parts of the muon system and then correlates them with other segments to find candidate tracks. These

muon tracks are sent to the global muon trigger, which combines them, sorts the candidates, and sends the four top muons to the global trigger.

The global trigger combines the results of the global calorimeter and muon triggers to filter events. Events that pass the L1 trigger proceed to data acquisition and are sent to the HLT. Information from the L1 trigger is also passed along to the HLT, which gives it regions of interest for it to begin processing first.

The HLT reduces the rate of events further to 100 Hz, a rate that we are able to store and reconstruct. To do this the HLT runs on a computer farm with standard processors rather than custom electronics like the L1 trigger or level-2 triggers in other experiments. The HLT was designed to run on processors and combine the level-2 and level-3 triggers of most experiments to allow the trigger to have full access to all of the data from an event and provide flexibility in what can be run—only constrained by computing resources. The trigger operates in stages to process events quickly by throwing out ones that fail at early stages before they consume processor time with later stages. The data from events being processed at the HLT are buffered in random-access memory while the trigger makes its decision. It takes approximately 1 s for the HLT to process an event [65].

The data acquisition system merges the data from the detector into a single, complete event in the event builder. The HLT then reconstructs parts of the event and makes selections on the objects it creates. This process is performed by algorithms running on the computer farm's processors rather than with circuits and firmware, so these algorithms can be updated more easily to change the behavior in the HLT.

The HLT has more capabilities and time to process events than the L1 trigger does, but it is constrained by having to reduce the data from 100 kHz to 100 Hz with finite resources while keeping up with the flow of incoming data. The algorithms run at the HLT must be efficient in their use of resources. They must not use too much computing time or the system would fall behind. The algorithms also need to be efficient in their usage of the available memory and cannot access external resources as that would overload the network.

Events are first filtered based on the calorimeters and muon system as would happen in the level-2 trigger in most experiments [66]. Events that pass move on to a level-3-like stage where tracks are reconstructed. Tracking, in particular, is time intensive so the HLT saves time by only reconstructing some regions in the detector rather than the whole event.

3.2.9   Luminosity measurement

Five detectors are used to measure the luminosity [67]. The pixel detector in the inner tracking system and the barrel DTs from the muon system are two of the methods of determine luminosity. Their output is read out using the normal data acquisition (DAQ) system. Both have low occupancy. The HF, fast beam conditions monitor (BCM1F), and pixel luminosity telescope (PLT) are read out by a fast readout system separate from CMS's normal DAQ system to provide real-time luminosity measurements.

Pixel cluster counting is used to measure the luminosity with the pixel tracker [68]. The high granularity of the pixel tracker means that there is low occupancy during collisions—i.e., hits occur in a low fraction of the pixels at any one time. The average number of pixel clusters

$$\langle N_{\text{cluster}} \rangle = \langle N_{\text{pixel/interaction}} \rangle \mu. \tag{3.16}$$

The average number of interactions in each bunch crossing

$$\mu = \frac{\sigma_{\text{interaction}}}{f_{\text{rev}}} \mathscr{L}, \tag{3.17}$$

where $\sigma_{\text{interaction}}$ is the interaction cross section and $f_{\text{rev}} = 11\,246\,\text{Hz}$ is the revolution frequency of the LHC. Using $\sigma_{\text{cluster}} = \sigma_{\text{interaction}} \langle N_{\text{pixel/interaction}} \rangle$ Eq. (3.16) becomes

$$\langle N_{\text{cluster}} \rangle = \frac{\sigma_{\text{cluster}}}{f_{\text{rev}}} \mathscr{L}. \tag{3.18}$$

The absolute luminosity was measured from a Van der Meer scan. The two beams

are scanned across one another and the rate of collisions is recorded as a function of the displacement between the beams [69]. The value of $\langle N_{\text{cluster}} \rangle$ is found from the number of clusters with zero displacement and $\mathscr{L}$ comes from the scan [68]. Equation (3.18) gives σ_{cluster}, which can then be used with $\langle N_{\text{cluster}} \rangle$ during data taking to measure the instantaneous luminosity.

The BCM1F is made from single-crystal chemical vapor deposition diamond sensors [70]. The sensors are located inside the inner tracking system along the beam pipe on opposite sides of the IP. PLT consists of three layers of silicon pixel detectors also located along the beam pipe [71]. Coincidence between the layers allows the PLT to distinguish particles from the IP versus those from the beam halo, which should not contribute to the luminosity measurement.

Chapter 4

Data scouting

The data acquisition of the experiment is limited and cannot record and reconstruct every collision. Triggers are needed to decide which events to keep. An effect of this is a lower limit on what dijet mass we can explore in this analysis. *Data scouting* is a technique to record information from low-mass events that would otherwise be lost.

As described in Section 3.2.8, CMS uses two levels of triggers to determine which events should be saved. The first, the L1 trigger, selects which events will pass on to the second, HLT, where the data for the events are quickly reconstructed *online* and trigger decisions are made. If the event passes the HLT, the event is reconstructed more fully *offline*. These reconstructed data sets (called RECO) are saved to disk. Additionally, subsets of the data with the more commonly used reconstructed objects are saved to facilitate their use in analyses that do not need to run over the less used information present in larger data sets. The analysis-level data sets are analysis object data (AOD) and the smaller MiniAOD. The raw data from the detector is saved for these events allowing them to undergo a new version of reconstruction in the future.

The triggers use basic kinematic- and object-based selection to decide which events to keep. In this analysis, the trigger uses the kinematic variable H_T, which is the sum of the hadronic energy in the event. The L1 trigger passes events with H_T above certain thresholds and marks the event with which thresholds were passed (e.g., $H_T > 125$, 150, or 175 GeV). The particle-flow (PF) RECO data comes from the PFHT800 trigger. The HLT reconstructs PF jets out of the data and then selects events with $H_T > 800$ GeV. The events that pass are saved and continue on for offline reconstruction. Events that

fail this criterion and those of all other triggers at the HLT are lost.

In data scouting, the objects reconstructed online by the HLT are saved with a limited event size. This bypasses the bottlenecks in offline reconstruction and write speed. Scouting adds little to the time it takes the HLT to reconstruct events since the quantities of interest to us are normally reconstructed for use by the trigger. We largely are opportunistically saving information for events that would have been lost when they failed the standard triggers.

The downside to this is that the information we have for each even is limited. The events store calorimeter jets with information regarding the four-vector $(p_{\mathrm{T}}, \eta, \phi, m)$; area of the jet; maximum energy in ECAL and HCAL towers; hadronic energy in the HB, HE, and HF; EM energy in the EB, EE, HF; area of the towers; whether the jet is from the PV; and a variable to b-tag the jet if it is available.[1] The events also contain other information such as the trigger results, vertices when available, and missing transverse momentum. Most of the reconstructed information available in RECO or MiniAOD formats is not available (including everything for objects other than jets). The raw data are also not generally saved, so unlike the standard data sets, there is no way to ever reconstruct the event again.

For the data sets used in this analysis, the raw data takes up approximately 630 kB per event. After it is reconstructed, the event size rises to 1400 kB. The MiniAOD samples used for the high-mass part of this analysis are a subset of that data that averages 27 kB per event. The average event for calorimeter scouting uses 1 kB—approximately 2.5 times the size of this paragraph encoded in ASCII.

Calorimeter scouting uses an $H_{\mathrm{T}} > 250\,\mathrm{GeV}$ trigger—well under the 800 GeV trigger used for the high-mass analysis. This still limits the lowest mass we are able to probe, but it gives us access to approximately a 600 GeV range of the dijet mass spectrum that we would be lost without data scouting. The H_{T} trigger is limited by the L1 triggers.

---

[1]The data scouting calorimeter jet class is available at `https://github.com/cms-sw/cmssw/blob/25ea639141c2a4ddf38c2dd6eda5f274ff809551/DataFormats/Scouting/interface/ScoutingCaloJet.h`

Reducing the H_T threshold at the HLT would not help. We have also investigated a trigger that selects events based on the dijet mass m_{jj} and found that it provides only a little improvement with the current L1 triggers.

In addition to calorimeter scouting, we have also implemented PF scouting, which saves PF objects from the HLT. It saves information for electrons, photons, and muons as well as jets. Although the reconstruction is online, the HLT PF algorithm still provides better resolution and more accurate information than the calorimeter jets do. However, the PF algorithm is more computationally intensive, so we must use a higher $H_T > 410\,\mathrm{GeV}$ trigger to prevent data scouting from using too much of the computing resources. This results in the trigger not becoming fully efficient until a higher mass is reached—one that is too high for a search for $750\,\mathrm{GeV}$ resonances.

We are able to save the raw data for some scouting events in what is called *parking*. Events with $H_T > 410\,\mathrm{GeV}$ can be saved directly to disk. No reconstruction is possible at the time of collection, but should CMS decide to allocate the resources based on a scouting analysis, these events could be reconstructed offline. However, this still leaves events with $250 < H_T < 410\,\mathrm{GeV}$ with calorimeter scouting as our only record.

There are supporting data sets for data scouting. One is a parking and scouting monitoring stream. It producing scouting files for a subset of events and also sends those same events to offline reconstruction. This allows us to measure the difference between the HLT calorimeter jets and RECO PF jets for the same data. There is also a commissioning stream that provides additional trigger information on a subset of events, which we can use to measure the trigger efficiency.

Chapter 5
Data and simulated samples

5.1 Data

All data were collected in 2016 during Run 2 of the LHC from the start of data taking in May through July with center-of-mass energy $\sqrt{s} = 13\,\text{TeV}$. Two data sets are used to search in different mass ranges. For high-mass resonances, we use a standard PF, fully reconstructed data set (PF RECO). To search for low-mass resonances, we use a scouting data set that employs calorimeter jets rather than PF jets (calorimeter scouting).

All events must be in certified luminosity sections where the detector was operating nominally. The certified data were determined by the "Golden" JSON `Cert_271036-276811_13TeV_PromptReco_Collisions16_JSON.txt`. The integrated luminosity of both the PF RECO and calorimeter scouting data is $12.9\,\text{fb}^{-1}$.

The PF RECO samples used for the high-mass analysis are listed in Table 5.1. Calorimeter scouting uses three data sets: ScoutingCaloHT for the event data, ScoutingCaloCommissioning for the trigger turn-on, and ParkingScoutingMonitor for the comparison of HLT calorimeter jets to PF jets reconstructed offline. The calorimeter scouting data sets used for the low-mass analysis are presented in Table 5.2.

Table 5.1: List of data sets used in the high-mass analysis.

Data set
/JetHT/Run2016B-PromptReco-v2/MINIAOD
/JetHT/Run2016C-PromptReco-v2/MINIAOD
/JetHT/Run2016D-PromptReco-v2/MINIAOD

Table 5.2: List of data sets used in the low-mass analysis for the dijet mass data (top), trigger efficiency (middle), and HLT to RECO comparison (bottom).

Data set
/ScoutingCaloHT/Run2016B-v2/RAW
/ScoutingCaloHT/Run2016C-v2/RAW
/ScoutingCaloHT/Run2016D-v2/RAW
/ScoutingCaloCommissioning/Run2016B-v2/RAW
/ScoutingCaloCommissioning/Run2016C-v2/RAW
/ScoutingCaloCommissioning/Run2016D-v2/RAW
/ParkingScoutingMonitor/Run2016B-v2/RAW
/ParkingScoutingMonitor/Run2016B-PromptReco-v2/MINIAOD
/ParkingScoutingMonitor/Run2016C-v2/RAW
/ParkingScoutingMonitor/Run2016C-PromptReco-v2/MINIAOD
/ParkingScoutingMonitor/Run2016D-v2/RAW
/ParkingScoutingMonitor/Run2016D-PromptReco-v2/MINIAOD

The samples were reconstructed using the CMS software version CMSSW_8_0_X.[1]

Checks were performed on the data to ensure its quality. We found that kinematic and other distributions from data were consistent with expectations from QCD simulation.

5.2 Monte Carlo simulations

The Monte Carlo (MC) simulation samples were created in the Spring16 campaign of production, which used CMSSW_7_1_X[2] for generation and simulation and CMSSW_8_0_X for digitization and reconstruction. The samples were generated with PYTHIA 8.205 [72] using the CUETP8M1 tune [73]. Appendix A presents the code used to generate the samples.

The events were simulated with a CMS software package based on GEANT 4 [74] using a model of the detector. Bunch crossings were simulated at 25 ns spacing with

---

[1] https://github.com/cms-sw/cmssw/tree/CMSSW_8_0_X
[2] https://github.com/cms-sw/cmssw/tree/CMSSW_7_1_X

asymptotic PU conditions. The jets were fully reconstructed with PU mitigation in the form of charged-hadron subtraction (CHS) [75].

The samples for the background from jets produced through QCD interactions were generated in bins of $\hat{p}_T$, which is the transverse momentum in the center-of-mass rest frame, starting at 50 GeV. A list of the data sets with cross sections, numbers of generated events, and the equivalent integrated luminosities of those numbers are in Table 5.3.

The signal MC samples cover three decay modes: $gg \rightarrow G \rightarrow gg$, $qq \rightarrow G \rightarrow qq$, and $qg \rightarrow q^* \rightarrow qg$. The gg and qq processes use an RS graviton model (Section 2.4.2) and the qg process uses an excited quark model (Section 2.4.1).

The samples were produced with narrow resonances, so that the resonance width is negligible compared to the the experimental mass resolution. The RS graviton samples were produced with $k/\overline{M}_{Pl} = 0.01$, so Eq. (2.23) gives $\Gamma(G \rightarrow gg)/m_G = 4.67 \times 10^{-5}$ and $\Gamma(G \rightarrow qq)/m_G = 8.76 \times 10^{-6}$. The excited quarks were produced with $f_s = 1$ and $\Lambda = m^*$, so Eq. (2.13) gives $\Gamma(G \rightarrow qg)/m^* \approx 0.03$.

The samples were produced at 11 masses: 500 GeV, 750 GeV, and 1 through 9 TeV in 1 TeV increments. The signal MC samples are listed in Table 5.4.

Table 5.3: QCD MC data sets with cross sections, numbers of generated events, and equivalent integrated luminosities.

Data set	Cross section [pb]	Events	Equivalent luminosity [fb^{-1}]
/QCD_Pt_50to80_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	19 204 300	9 968 410	5.191×10^{-4}
/QCD_Pt_80to120_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	2 762 530	9 968 410	0.003608
/QCD_Pt_120to170_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	471 100	6 863 827	0.01457
/QCD_Pt_170to300_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	117 276	6 914 086	0.05896
/QCD_Pt_300to470_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	7823	5 970 600	0.7632
/QCD_Pt_470to600_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	648.2	3 928 870	6.061
/QCD_Pt_600to800_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	186.9	3 959 768	21.19
/QCD_Pt_800to1000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	32.293	3 924 080	121.5
/QCD_Pt_1000to1400_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	9.4183	2 999 069	318.4
/QCD_Pt_1400to1800_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	0.84265	396 409	470.4
/QCD_Pt_1800to2400_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	0.114943	396 100	3.446×10^{3}
/QCD_Pt_2400to3200_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	0.00682981	399 226	5.845×10^{4}
/QCD_Pt_3200toInf_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	0.000 165 445	383 926	2.321×10^{6}

Table 5.4: Signal MC data sets with cross sections and numbers of generated events.

Data set	Cross section [pb]	Events
/RSGravitonToGluonGluon_kMpl01_M_500_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	209.6	100 000
/RSGravitonToGluonGluon_kMpl01_M_750_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	35.25	100 000
/RSGravitonToGluonGluon_kMpl01_M_1000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	6.186	99 674
/RSGravitonToGluonGluon_kMpl01_M_2000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	0.0775	100 000
/RSGravitonToGluonGluon_kMpl01_M_3000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	3.58×10^{-3}	99 702
/RSGravitonToGluonGluon_kMpl01_M_4000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	2.25×10^{-3}	99 854
/RSGravitonToGluonGluon_kMpl01_M_5000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	1.69×10^{-5}	99 929
/RSGravitonToGluonGluon_kMpl01_M_6000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	1.36×10^{-6}	99 952
/RSGravitonToGluonGluon_kMpl01_M_7000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	1.68×10^{-7}	100 000
/RSGravitonToGluonGluon_kMpl01_M_8000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	2.64×10^{-8}	100 000
/RSGravitonToGluonGluon_kMpl01_M_9000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	8.55×10^{-9}	99 853
/RSGravitonToQuarkQuark_kMpl01_M_500_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	22.32	99 004
/RSGravitonToQuarkQuark_kMpl01_M_750_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	5.39	99 598
/RSGravitonToQuarkQuark_kMpl01_M_1000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	1.61	99 106
/RSGravitonToQuarkQuark_kMpl01_M_2000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	0.0663	100 000
/RSGravitonToQuarkQuark_kMpl01_M_3000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	6.49×10^{-3}	100 000
/RSGravitonToQuarkQuark_kMpl01_M_4000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	7.47×10^{-4}	99 817
/RSGravitonToQuarkQuark_kMpl01_M_5000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	6.78×10^{-5}	100 000
/RSGravitonToQuarkQuark_kMpl01_M_6000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	7.82×10^{-6}	99 901
/RSGravitonToQuarkQuark_kMpl01_M_7000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	8.38×10^{-7}	99 632
/RSGravitonToQuarkQuark_kMpl01_M_8000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	1.05×10^{-7}	99 956
/RSGravitonToQuarkQuark_kMpl01_M_9000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	2.07×10^{-8}	100 000
/QstarToJJ_M_500_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	16260	98 980
/QstarToJJ_M_750_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	2931	91 550
/QstarToJJ_M_1000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	807.7	100 000
/QstarToJJ_M_2000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	21.49	99 856
/QstarToJJ_M_3000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	1.625	99 903
/QstarToJJ_M_4000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	0.176	100 000
/QstarToJJ_M_5000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	0.0182	99 915
/QstarToJJ_M_6000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	1.86×10^{-3}	99 520
/QstarToJJ_M_7000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	2.94×10^{-4}	97 840
/QstarToJJ_M_8000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	4.15×10^{-5}	97 520
/QstarToJJ_M_9000_TuneCUETP8M1_13TeV_pythia8/RunIISpring16MiniAODv2-PUSpring16_80X_mcRun2_asymptotic_2016_miniAODv2_v0-v1/MINIAODSIM	1.07×10^{-5}	99 808

Chapter 6

Event reconstruction

Events were reconstructed with software version CMSSW_8_0_X. The raw data from the subdetectors are reconstructed into hits and those hits are combined together based on the behavior of the particles to reconstruct the type of stable particle that passed through the detector and measure its four-momentum. "Stable" particle for CMS includes the muon, which while unstable has $c\tau = 659\,\text{m}$ and will exit the detector most of the time before it decays [32].

6.1 Particle-flow algorithm

The particle-flow (PF) algorithm reconstructs particles using an optimized combination of subdetector information based on how different particles traverse the detector [76,77]. First other algorithms create *elements* that the PF algorithm can use such as tracks from hits in the tracking system or clusters of energy in the calorimeters. A linking algorithm then compares pairs of elements based on distance and produces *blocks* of elements that are linked together (usually with one to three elements). Blocks can contain a mix of elements from the inner tracker, calorimeters, and muon system. The PF algorithm then interprets these blocks to form reconstructed particles.

A track that when extrapolated back leads to a cluster of energy in the ECAL with more clusters tangential to the curving track from bremsstrahlung is reconstructed as an electron. A photon is identified by a cluster of energy in the ECAL that is not linked with a track. A track in the inner tracking system linked with another in the muon system results in a muon. Charged hadrons are reconstructed when there is a track from

a charged particle that has not been identified as an electron or muon. Neutral hadrons are derived from calorimeter clusters not linked to a charged hadron track or energy in clusters in excess of what is expected from the linked charged hadron's p_T. No tracker is in front of the HF, so there is no information to determine whether the particles are charged or neutral. The energy deposited in the long versus short fibers in the HF determine whether they are the result of an EM particle or hadron.

The four-momentum is calculated for each of the particles. The direction comes from the location of the hits in the detector. The p_T for charged particles is determined from the curvature of their tracks and energy deposits in the calorimeters. The energy for neutral particles is determined from the calorimeters. The particles are then clustered to form jets.

6.2   Calorimeter jet algorithm

Calorimeter jets are formed by clustering the energy deposited in the calorimeters rather than by clustering reconstructed particles from the PF algorithm. Calorimeter towers are made from one cell in the HCAL and the 5×5 supercrystal array of ECAL crystals in front of it. The energy from both calorimeters are summed to produce the calorimeter-tower energy. The magnitude of the momentum is equal to the calorimeter tower's energy and its direction is determined by the tower's position in η and ϕ. The jet-clustering algorithm is applied to the calorimeter towers to create jets. The jet energy is the sum of the energy in the calorimeter towers, and the jet momentum is the vectorial sum of the momenta of the calorimeter towers. Only calorimeter information is used.

Calorimeter jets are more susceptible to PU than PF jets. The tracks from charged constituent particles coming from in-time PU (i.e., originating from a different vertex than the jet) can be removed from PF jets. But without the tracking information to remove these particles, energy from PU is included in the calorimeter jet. The jet energy resolution for PF jets is typically 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV. For

calorimeter jets, it is approximately 40% at $10\,\text{GeV}$, 12% at $100\,\text{GeV}$, and 5% at $1\,\text{TeV}$. The difference in resolution is greatest as lower energies.

6.3   Clustering

Several algorithms exist to cluster energy deposits or particles into jets. These include the k_T, anti-k_T [78], Cambridge–Aachen, and cone algorithms. An iterative cone algorithm starts with the hardest object and creates a cone around it of radius R in η-ϕ space. The four-momenta of the particles inside the cone are summed and if it is close enough to the four-momentum of the seed particle in ΔR, the particles inside the cone are a jet. The particles in the jet are removed and this process is iterated starting with the hardest remaining object as the seed for the next cone.

A concern with cone algorithms is that they tend not to be collinear- or infrared-safe. Collinear splitting and soft radiation, respectively, should not affect what jets result from clustering. Sequential clustering algorithms are designed to avoid this problem. These algorithms use two distance parameters [78]. One is the distance between two objects

$$d_{ij} = \min\left(k_{\text{T}i}^{2p}, k_{\text{T}j}^{2p}\right)\frac{\Delta_{ij}^2}{R^2}, \tag{6.1}$$

where k_T is the transverse momentum, $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, and R is a radius parameter. The other is the distance between an object and the beam

$$d_{iB} = k_{\text{T}i}^{2p}. \tag{6.2}$$

The objets include particles that have not been clustered yet and pseudo-jets made from clustered particles as the algorithm proceeds. The distances are calculated for every object and pair of objects. If d_{iB} is the smallest distance, then object i is considered a jet and removed. If d_{ij} is the smallest distance, objects i and j are combined. This process is repeated until all particles are in jets. Changing parameter p results in different clustering

behaviors.

The Cambridge–Aachen algorithm uses $p = 0$. This results in clustering based only on separation in y-ϕ as the distance between two objects becomes $d_{ij} = \Delta_{ij}^2 / R^2$. The k_{T} algorithm uses $p = 1$. This makes the clustering depend not just on spatial separation but also on the transverse momenta. When particles are equally close to two jets, they will be clustered into the jet with the smaller transverse momentum.

CMS primarily uses the anti-k_{T} algorithm, where $p = -1$. Particles equally separated between two jets will be clustered into the jet with larger transverse momentum. This results in hard jets that tend to be circular in y-ϕ space with a radius of R. Softer jets also tend to be circular but any overlap with harder jets is removed from the circle. The jets produced by the anti-k_{T} algorithm are less sensitive to the presence of PU [78].

The jets used in this analysis were clustered using the anti-k_{T} algorithm with a distance parameter of $R = 0.4$ (called AK4 jets) using the FASTJET package [79]. After identifying the two leading AK4 jets, we create wide jets to pick up radiation. The clustering for wide jets does not need to be as sophisticated as the anti-k_{T} algorithm. We cluster all jets within a radius of $\Delta R = 1.1$ of the leading AK4 jets into the two wide jets. Because dijet events will tend to produce back-to-back leading jets and we only create two wide jets, we do not need sophisticated handling of overlap. We will get two circular wide jets of radius 1.1 that do not touch.

The PF jets are made with CHS to mitigate the effect of PU on the reconstructed jets. Charged-particle PF candidates that do not originate from the primary vertex (PV) are removed before the jets are clustered. The PV is defined as the one with the greatest scalar sum of p_{T}^2 from its tracks.

6.4   Jet energy corrections

The measured energy of jets is dependent on PU, detector response, transverse momentum, and pseudorapidity. The energy of reconstructed jets must be corrected to give

consistent results at the right energy. Jet energy corrections (JECs) provide correction factors based on p_T, η, energy density ρ, and jet area A that are applied to the jets' four-momenta. To find the JECs, data are compared to the results of MC simulations [80].

The η dependance comes from dijet events. The absolute energy scale is determined for $|\eta| < 1.3$ and $30 < p_\mathrm{T} < 800\,\mathrm{GeV}$ by events with a photon or Z boson that recoils off a jet. Multijet events are used to obtain the absolute energy scale for $p_\mathrm{T} > 800\,\mathrm{GeV}$. The p_T of the jet is balanced against the p_T of the object it recoiled against (another jet, photon, or lepton pair from a Z boson decay). This relies on the transverse momentum of the event being conserved.

The JECs used for the high-mass PF RECO analysis (Spring16_25nsV6) are standard. The L1 FastJet, L2 relative, and L3 absolute corrections come from MC simulations. The L2L3 residual correction comes from PF data. The corrections for the low-mass calorimeter scouting analysis use corrections for HLT jets (80X_dataRun2_HLT_frozen_v12) except for the L2L3 residual correction which is the same as used for the high-mass analysis.

Calorimeter jets from the HLT are not standard analysis objects in CMS. Analyses normally use PF jets reconstructed offline like the high-mass search here does. The importance of HLT jets to the trigger means they are as accurate as possible but the trigger has constrained resources that limit what can be done for their quality. To compare data from HLT calorimeter jets to MC simulations using CHS PF jets that were fully reconstructed, we must correct for the difference.

The ParkingScoutingMonitor data set from Section 5.1 provides us with a subset of the data that has events with both scouting data containing its HLT calorimeter jets and an offline reconstruction of the event with CHS PF jets. These two type of jets can be compared to produce a correction to remove their average difference.

The comparison gives us the correction factor

$$c(p_T) = \begin{cases} \left(a_0 + a_1 \log p_T + a_2 \log^2 p_T\right)^{-1} & p_T \leq 993.264\,\text{GeV} \\ c(993.264\,\text{GeV}) & p_T > 993.264\,\text{GeV} \end{cases}, \qquad (6.3)$$

where $a_0 = 0.682\,802$, $a_1 = 0.085\,861\,1$, and $a_2 = -0.006\,220\,92$, that we can multiply to the four-momenta of wide jets made from HLT calorimeter jets to approximate those made from PF jets reconstructed offline. Appendix B shows the derivation of this factor.

6.5 Jet energy resolution

Jet energy resolution (JER) will be different for PF jets reconstructed offline and calorimeter jets reconstructed at the HLT even if JECs mean their jet energy scale (JES) agrees on average. The dijet mass resolution $\sigma(m_{jj})/m_{jj} \approx \sigma(A)$, where $\sigma(A)$ is the resolution of the asymmetry of p_T between the leading two jets

$$A = \frac{p_T^{j_1} - p_T^{j_2}}{p_T^{j_1} + p_T^{j_2}}. \qquad (6.4)$$

If the remaining jets in the event do not contribute much, the two leading jets should have balanced p_T ($A = 0$) because of conservation of momentum.

To measure the resolution, we require the leading two AK4 jets have $p_T > 10\,\text{GeV}$ and $|\eta| < 2.5$. The wide jets must have $p_T > 10\,\text{GeV}$, $|\eta| < 2.5$, $\Delta\phi > 2.8$, and $p_T^{j_3} < 0.3\,p_T^{\text{ave}}$, where $p_T^{\text{ave}} = (p_T^{j_1} + p_T^{j_2})/2$. Using the ParkingScoutingMonitor data set from Section 5.1 we compare the resolution of calorimeter jets from the HLT to RECO PF jets by matching them in η-ϕ space with the requirement that $\Delta R < 0.3$ between the two types of jets.

The distributions of p_T asymmetry A for bins in m_{jj} were fit to a Gaussian with exponential tails to determine $\sigma(A)$. Figure 6.1 shows the distribution of $\sigma(A) \approx \sigma(m_{jj})/m_{jj}$ versus the dijet mass for HLT and RECO jets. The ratio of $\sigma(m_{jj})/m_{jj}$ for HLT to that

for RECO provides a scale factor to transform signal shapes using PF jets to those needed for calorimeter jets from the HLT in Chapter 9.
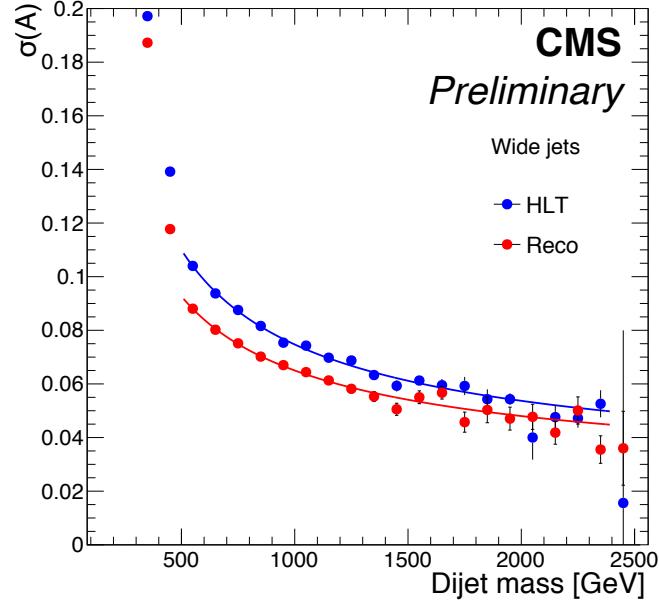
Figure 6.1: Resolution of the p_T asymmetry $\sigma(A) \approx \sigma(m_{jj})/m_{jj}$ for wide jets from HLT calorimeter jets and RECO PF jets as a function of the dijet invariant mass.

Chapter 7

Event selection

7.1 Preselection

The calorimeter scouting trigger at the HLT uses an $H_T > 250\,\mathrm{GeV}$ threshold. The H_T is calculated from the scalar sum of the p_T of jets with $p_T > 40\,\mathrm{GeV}$ and $|\eta| < 3$. The high-mass analysis uses a threshold of $H_T > 800\,\mathrm{GeV}$, where only jets with $p_T > 30\,\mathrm{GeV}$ and $|\eta| < 3$ are used in the H_T sum.

7.2 Selection

All events must be in luminosity sections that have been certified as being collected while the accelerator and detector were functioning nominally. All AK4 jets are required to have $p_T > 30\,\mathrm{GeV}$ and $|\eta| < 2.5$. In the high-mass analysis, at least one reconstructed vertex is required.

The two jets with the highest p_T (the leading jets) must pass a jet identification. For the PF jets, we use the "tight lepton veto" jet identification criteria that

- Neutral hadron fraction < 0.9

- Neutral EM fraction < 0.9

- Number of constituents > 1

and for jets with $|\eta| < 2.4$, a region that the tracker covers so determining whether a particle is charged is possible,

- Charged hadron fraction > 0

- Charge multiplicity > 0

- Charged EM fraction < 0.9

- Muon fraction < 0.8

where these are fractions of the jet's energy [81]. For calorimeter jets, we require the leading two jets to have

- EM fraction < 0.95

- Hadron fraction < 0.95

The wide jets are required to be separated in pseudorapidity by $|\Delta\eta| = |\eta_1 - \eta_2| < 1.3$. This criterion is to reduce t-channel QCD background, in which leading jets tend to be separated by larger $|\Delta\eta|$ than they are for signal. This is equivalent to the requirement on

$$\cos\theta^* = \tanh\frac{y_1 - y_2}{2} \rightarrow \tanh\frac{\Delta\eta}{2} \tag{7.1}$$

used in some past dijet analyses in the relativistic limit, where θ^* is the scattering angle in the center-of-mass frame.

The lower mass boundaries are determined by the triggers for the two data sets such that the triggers in the search regions are fully efficient. We require $m_{jj} > 453\,\text{GeV}$ for calorimeter scouting and $m_{jj} > 1058\,\text{GeV}$ for PF RECO. While the PF RECO analysis extends to the highest mass events we observed, the calorimeter scouting analysis limits the dijet mass to the range $453 < m_{jj} < 2037\,\text{GeV}$. The upper boundary at $2037\,\text{GeV}$ allows us to set limits up to the point where the high-mass analysis begins setting limits.

Chapter 8

Trigger efficiency

The fit to the data in Chapter 10 assumes that we are fully efficient in collecting all of the events in the fitted region. Otherwise, a trigger efficiency factor would need to be included in the fit function, which would introduce additional systematic uncertainties.

We can measure the efficiency of the trigger by comparing it to a sample with either an orthogonal trigger or one with a low enough H_T threshold that the turn-on of the second trigger does not affect the measurement of the efficiency of the first one.[1]

For the high-mass PF RECO analysis, we use all of the JetHT data regardless of the trigger. In 2016, all of the data were collected with the PFHT800 trigger. This was not the case in 2015, when some high-mass events passed the PFJet500 trigger. We are able to measure the efficiency with respect to the Mu45 and PFHT475 triggers. The efficiency is defined as the ratio of the number of events in the PFHT800 and Mu45 triggers divided by the number of events that pass the Mu45 trigger. Similarly for the PFHT475 trigger.

The measurements of the trigger efficiency for the high-mass analysis as a function of dijet mass are shown in Fig. 8.1 with a fit to a sigmoid function. The PFHT800 trigger reaches an efficiency greater than 99% above 1058 GeV as measured by both methods. Therefore, we use a threshold of $m_{jj} > 1058$ GeV for the high-mass analysis.

In the low-mass calorimeter scouting analysis, we use all of the ScoutingCaloHT data regardless of the trigger. In 2016 all of the events passed the HT250 trigger. Additional triggers in the ScoutingCaloCommissioning data set are used to measure the efficiency of HT250. The trigger efficiency is measured against the L1HTT trigger, which uses the

---

[1]The lower trigger is prescaled (i.e., only run on a fraction of events). Otherwise, we would be using the lower H_T trigger for the analysis.
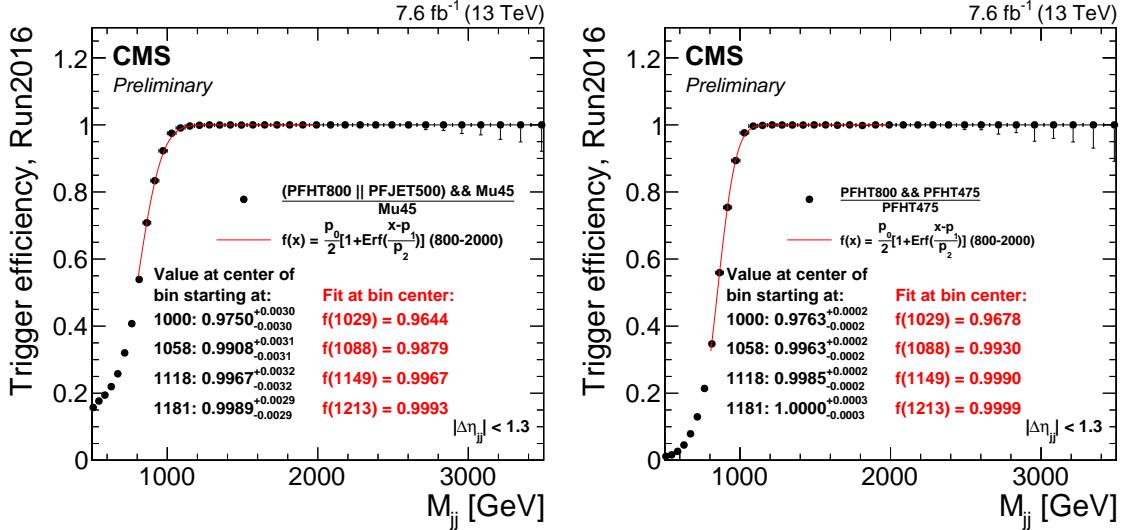
Figure 8.1: Efficiency verses dijet mass of the PFHT800 trigger used in the high-mass analysis measured against the Mu45 trigger (left) and the PFHT475 trigger (right).

same L1 seeds as HT250 but applies no selection at the HLT. We measure the efficiency of the L1HTT trigger against a zero-bias trigger at L1 that requires a jet with $p_T > 40\,\mathrm{GeV}$ at the HLT. Table 8.1 lists the trigger and their details.

Table 8.1: Scouting triggers with the L1 seeds, prescale, selection criteria at the HLT, and purpose.

Name	L1 Seeds	HLT Prescale	Selection	Purpose
HT250	HTT125 or HTT150 or HTT175	1	$H_T > 250\,\mathrm{GeV}$	Main analysis trigger
L1HTT	HTT125 or HTT150 or HTT175	1000	None	Measure HLT efficiency
CaloJet40	ZeroBias	10	calorimeter jet with $p_T > 40\,\mathrm{GeV}$	Measure L1 efficiency

The trigger efficiencies for the HT250 and the L1HTT triggers are shown in Fig. 8.2. Because of the large number of events at low dijet mass, even an efficiency of 99.0% would mean that there would be about half a million missing events in the first dijet

mass bin. We therefore, chose the threshold $m_{jj} > 453\,\text{GeV}$ for the calorimeter scouting analysis as the point where the trigger is sufficiently efficient for our purposes.
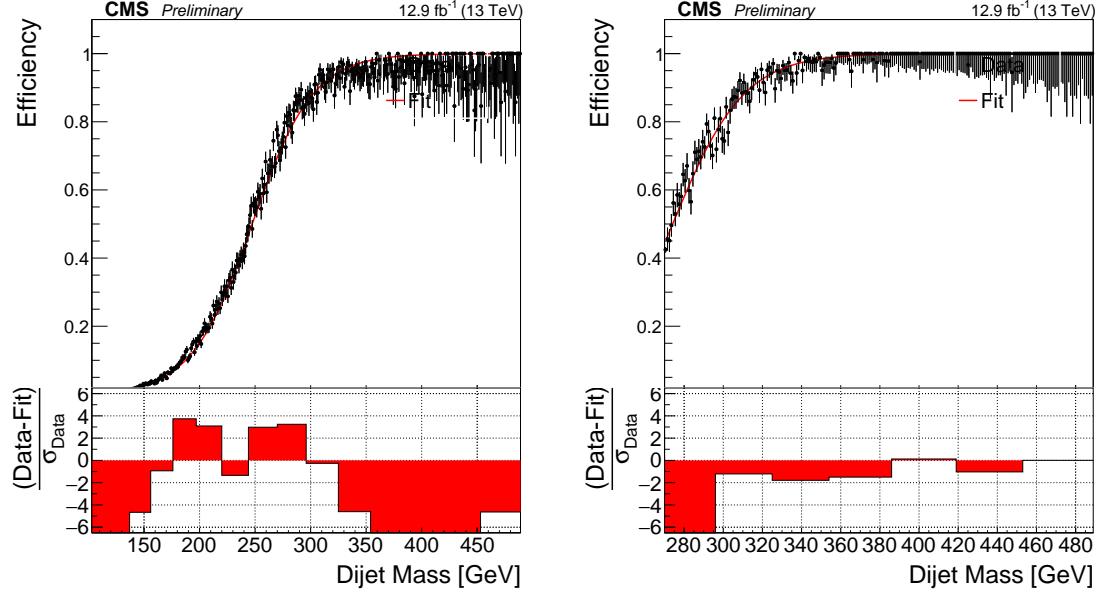
Figure 8.2: Efficiency versus dijet mass for the L1 (left) and L1+HLT (right) triggers used in the low-mass analysis. Both are measured relative to the CaloJet40 trigger. The lower panels show the difference between the data and the sigmoid fit divided by the statistical uncertainty of the data.

Chapter 9

Signal shapes

The shape of a signal resonance depends on the decay mode as quarks and gluons radiate differently. The natural width of the resonance does not affect the shape because it is chosen to be negligible compared to the dijet mass resolution (i.e., a narrow resonance).

Some past analyses have used Breit–Wigner and Gaussian signal shapes [5]. But dijet analyses now use MC simulations to produce more physically realistic signal shapes. That is the technique used here.

The signal MC samples were described in Section 5.2. They cover the processes $gg \to G \to gg$, $qg \to q^* \to qg$, and $qq \to G \to qq$. These three models are used for the shapes for general gg, qg, and qq resonances. We also use Gaussian signal shapes, which, while non-physical, provide a reference point useful for theorists.

Eleven mass points were generated for each process: 500–1000 GeV in 250 GeV steps and 1–9 TeV in 1 TeV steps. Shapes in between these mass points are determined by interpolation. The probability density function (pdf) of the m_{jj} distributions of the MC samples are found in terms of the parameter $x = m_{jj}/M_{res}$, where M_{res} is the resonance mass. Then for an interpolated mass M, the pdf is

$$p_M(x) = p_{M_1}(x) + \left[p_{M_2}(x) - p_{M_1}(x) \right] \frac{M - M_1}{M_2 - M_1}, \qquad (9.1)$$

where $M_1 < M_2$ are the two neighboring masses with MC samples.

The shapes for wide jets after applying the analysis selection criteria are shown for some masses in Fig. 9.1. The resonance shapes are dependent on the mode of decay. Gluons radiate more than quarks do, so the width of the shape increases with the number

of gluons in the decay channel. The lower tail is caused by final-state radiation that has not been caught by the wide jets and the PDF having higher parton luminosity at low mass than at high mass. The upper tail is smaller and caused by initial-state radiation that is clustered into a wide jet. Initial-state radiation is when an incoming parton radiates off a gluon before the resonance; final-state radiation is when an outgoing parton from the decay of the resonant particle radiates a gluon. The peak of the resonance shape is lower than the resonance mass because of losses due to radiation.
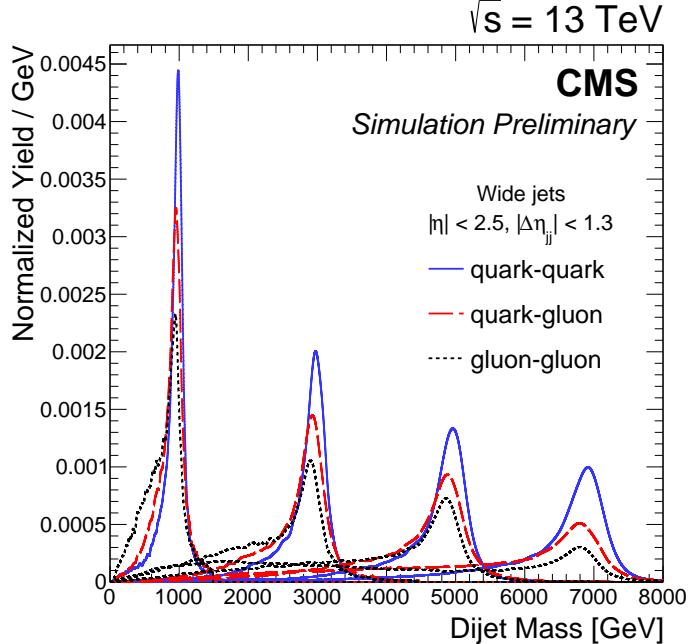
Figure 9.1: Signal shapes for the high-mass analysis from $gg \to G \to gg$, $qg \to q^* \to qg$, and $qq \to G \to qq$ resonances at 1, 3, 5, and 7 TeV. The integrals of all shapes have been normalized to unity.

While the shapes will vary depending on the model, they will be similar to the shapes from the RS graviton or excited quark models as long as the widths of the resonances are narrow compared to the detector resolution.

The calorimeter scouting data had its energy scale corrected but its energy resolution is larger than the JER PF jets. We therefore smeared the shapes from MC simulation with a Gaussian. The dijet mass for each event in the MC sample was multiplied by a

factor chosen randomly from a Gaussian distribution of mean 1 and standard deviation

$$\left(\frac{\sigma(m_{jj})}{m_{jj}}\right)_{\text{RECO}} \sqrt{\frac{[\sigma(m_{jj})/m_{jj}]^2_{\text{HLT}}}{[\sigma(m_{jj})/m_{jj}]^2_{\text{RECO}}} - 1}, \tag{9.2}$$

where RECO is the resolution of PF jets from offline reconstruction and HLT is the resolution of calorimeter jets from the HLT. The relative mass resolution $\sigma(m_{jj})/m_{jj}$ is found from the dijet p_{T} asymmetry in Fig. 6.1. The smeared shapes for calorimeter scouting are shown in Fig. 9.2.
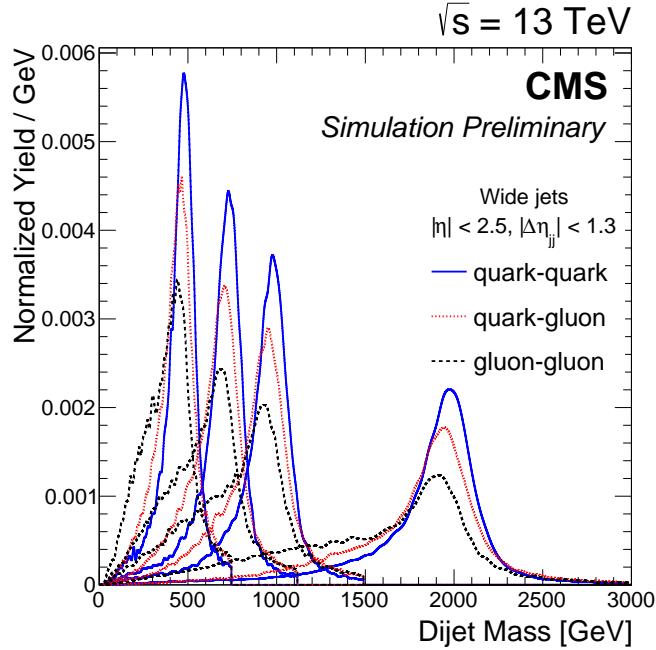
Figure 9.2: Signal shapes for the low-mass analysis smeared to approximate calorimeter jets from the HLT. Shapes for $gg \to G \to gg$, $qg \to q^* \to qg$, and $qq \to G \to qq$ resonances at 500, 750, 1000, and 2000 GeV. The integrals of all shapes have been normalized to unity.

We also set limits with a Gaussian for ease of interpretation by theorists. The physical shapes are not Gaussian, but their cores can be approximated by one. We can therefore fit the cores to obtain an approximation of the mean and width that would be observed if the tails of the signal were obscured by the background leaving only a Gaussian-shaped core. The Gaussian mean and standard deviation for the the cores of the signal shapes are

shown in Fig. 9.3 as a ratio of resonance mass. The peaks of all shapes occur below the resonance mass, although at large masses the effect is smaller. The quark–quark shapes are the closest because quarks radiate less than gluons.
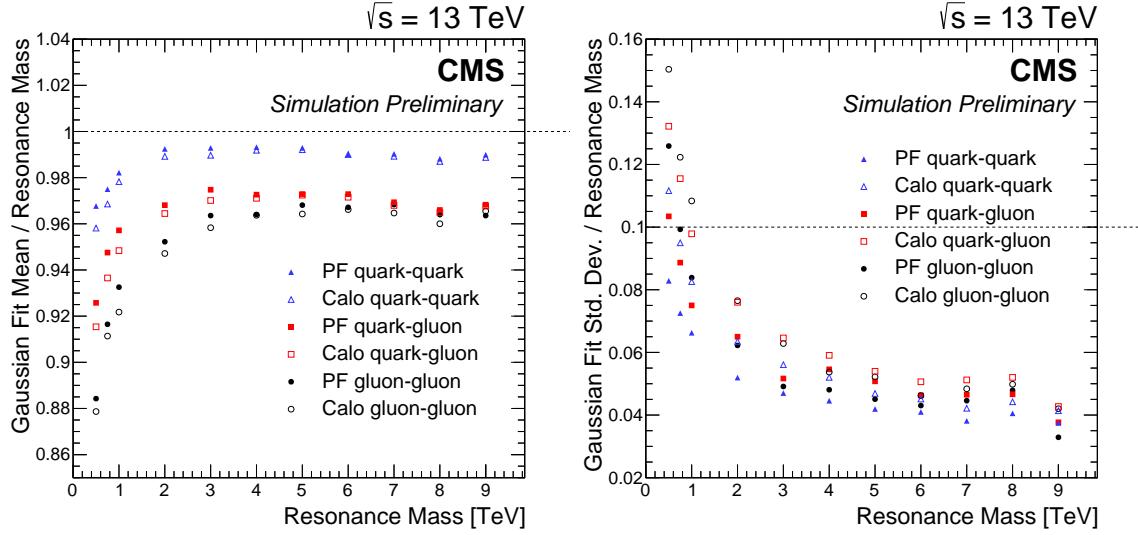
Figure 9.3: The ratio of the mean (left) and standard deviation (right) to the resonance mass for Gaussian fits to the core of PF RECO and calorimeter scouting signal shapes for $gg \to G \to gg$, $qg \to q^* \to qg$, and $qq \to G \to qq$ resonances. Values for 10% Gaussian shapes are on the dashed line.

Figure 9.4 shows a comparison of all of the shapes at $750\,\text{GeV}$.
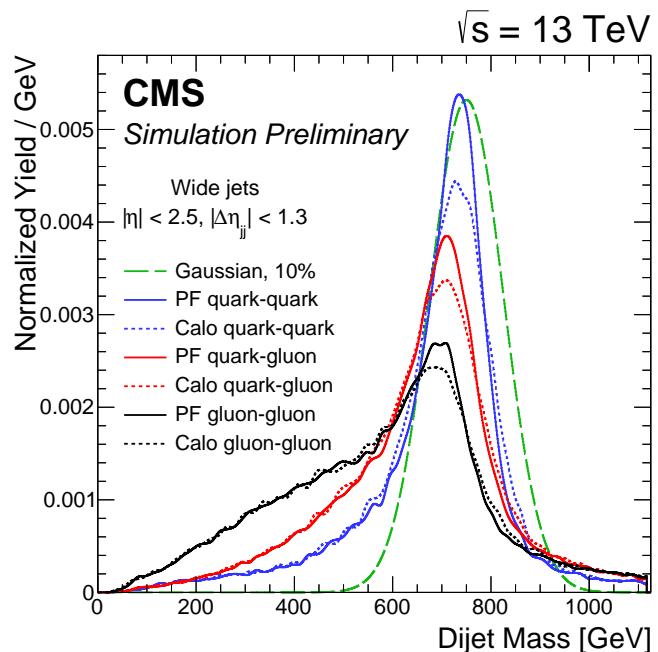
Figure 9.4: Comparison of signal shapes from gg → G → gg, qg → q* → qg, and qq → G → qq resonances at 750 GeV. PF RECO (solid), calorimeter scouting (dotted), and 10% Gaussian (dashed). The integrals of all shapes have been normalized to unity.

Chapter 10

Dijet mass spectrum

The background dijet mass spectrum smoothly falls with increasing invariant mass. We search for bumps in this spectrum indicative of a resonance modeled by the signal shapes in Chapter 9.

Dijet searches have taken different approaches to modeling the background. Early searches used QCD MC simulations for the background. However, QCD is insufficiently well understood to produce fully reliable simulations. Instead, we fit the data with a smoothly falling function to determine the shape of the background.

The dijet mass resolution and the number of events decrease with mass, so we use a variable binning for the dijet mass spectrum with bins increasing in width as mass goes up.[1] Additionally, the number of events in a given bin will change depending on how much integrated luminosity has been collected for a given analysis. Because of this, we look at the differential cross section $d\sigma / dm_{jj}$. The number of events N in a bin is related to this by

$$N = L \int_{m_{low}}^{m_{up}} \frac{d\sigma}{dm_{jj}} \, dm_{jj} \qquad (10.1)$$

for integrated luminosity L.

---

[1]The bin boundaries are at: 1, 3, 6, 10, 16, 23, 31, 40, 50, 61, 74, 88, 103, 119, 137, 156, 176, 197, 220, 244, 270, 296, 325, 354, 386, 419, 453, 489, 526, 565, 606, 649, 693, 740, 788, 838, 890, 944, 1000, 1058, 1118, 1181, 1246, 1313, 1383, 1455, 1530, 1607, 1687, 1770, 1856, 1945, 2037, 2132, 2231, 2332, 2438, 2546, 2659, 2775, 2895, 3019, 3147, 3279, 3416, 3558, 3704, 3854, 4010, 4171, 4337, 4509, 4686, 4869, 5058, 5253, 5455, 5663, 5877, 6099, 6328, 6564, 6808, 7060, 7320, 7589, 7866, 8152, 8447, 8752, 9067, 9391, 9726, 10 072, 10 430, 10 798, 11 179, 11 571, 11 977, 12 395, 12 827, 13 272, 13 732, 14 000 GeV.

We use the four-parameter function

$$\frac{d\sigma}{dm_{jj}} = p_0 \frac{(1 - m_{jj}/\sqrt{s})^{p_1}}{(m_{jj}/\sqrt{s})^{p_2 + p_3 \log(m_{jj}/\sqrt{s})}} \tag{10.2}$$

to model the falling background. The parameters p_0, p_1, p_2, and p_3 are found in a fit to data. Equation (10.2) is a function of the dimensionless parameter $m_{jj}/\sqrt{s}$.

Other functions have been used in past experiments. UA2 [8] used

$$\frac{d\sigma}{dm_{jj}} = \frac{p_0}{m_{jj}^{p_1}} e^{-p_2 m_{jj} + p_3 m_{jj}^2}. \tag{10.3}$$

CDF [11] used

$$\frac{d\sigma}{dm_{jj}} = p_0 \frac{(1 - m_{jj}/\sqrt{s})^{p_1}}{m_{jj}^{p_2}} \tag{10.4}$$

based on the $m_{jj}^{-p_2}$ behavior of the QCD matrix elements and $(1 - m_{jj}/\sqrt{s})^{p_2}$ dependence of parton distributions. It later used the function

$$\frac{d\sigma}{dm_{jj}} = p_0 \frac{(1 - m_{jj}/\sqrt{s} + p_3 m_{jj}^2/s)^{p_1}}{m_{jj}^{p_2}} \tag{10.5}$$

with the additional p_3 term added to better fit the data [12]. Finally, Eq. (10.2) was developed by CDF to provide a better fit to the dijet mass spectrum [13]. CMS and ATLAS have used Eq. (10.2) for all searches (ATLAS once used it with $p_3 = 0$ [26]).

Four parameters were chosen for Eq. (10.2) because they provide a significant improvement over the fit of a three-parameter function ($p_3 = 0$) while a five-parameter function (adding $p_4 \log^2[m_{jj}/\sqrt{s}]$ in the denominator's exponent) does not provide a significant improvement. See Appendix C for the test used to determine the necessary number of parameters.

We performed signal injection tests using pseudo-data generated with the parametriza-tion

$$\frac{\mathrm{d}\sigma}{\mathrm{d}m_{\mathrm{jj}}} = p_0 e^{p_1(m_{\mathrm{jj}}/\sqrt{s})^{p_2} + p_1(1-m_{\mathrm{jj}}/\sqrt{s})^{p_3}} \tag{10.6}$$

that was then fit to Eq. (10.2). We found the bias in the extracted signal from our choice of four-parameter function to be negligible.

The fit of Eq. (10.2) to data is performed with RooFit [82] as an extended, binned, maximum-likelihood fit. The fit is first performed for the hypothesis that there is no signal, only background. We use the likelihood

$$\mathcal{L}(x_i|b_i(\theta)) = \prod_{i=1}^{n_b} \mathrm{Poisson}(x_i|b_i(\theta)) = \prod_{i=1}^{n_b} \frac{b_i(\theta)^{x_i} e^{-b_i(\theta)}}{x_i!}, \tag{10.7}$$

where x_i are the data for the differential cross section of each bin, θ is a vector of the four nuisance parameters (the four parameters of the background function), n_b is the number of bins, and the expected number of background events in the bins

$$b_i(\theta) = N_b \int_{m_{i,\mathrm{low}}}^{m_{i,\mathrm{up}}} p(m_{\mathrm{jj}}) \mathrm{d}m_{\mathrm{jj}}, \tag{10.8}$$

where N_b is the total number of events and $p(m_{\mathrm{jj}})$ is Eq. (10.2) normalized to unity. The parameters are varied to minimize the negative logarithm of the likelihood function $-\log\mathcal{L}(x_i|b_i(\theta))$.

To evaluate the goodness of fit, we use two measures: χ^2 and a likelihood ratio. The first is defined as

$$\chi^2 = \sum_{i=1}^{n_b} \left(\frac{x_i - b_i}{\sigma_i}\right)^2, \tag{10.9}$$

where σ_i is the statistical uncertainty on the data in bin i. It can be found from the 68%

CL of a Poisson distribution

$$\sigma_i = \begin{cases} D_c^{-1}(\alpha/2, x_i + 1) & b_i > x_i \\ D^{-1}(\alpha/2, x_i) & b_i < x_i \end{cases},$$ (10.10)

where $\alpha = 1 - 0.687$, $D^{-1}(\alpha/2, x_i)$ is the quantile function for the lower tail of the gamma distribution, and $D_c^{-1}(\alpha/2, x_i + 1)$ is the quantile function of the upper tail of the gamma distribution. The quantile functions are the inverses of the cumulative distribution functions

$$D(\alpha/2, x_i) = \int_{-\infty}^{\alpha/2} \frac{1}{\Gamma(x_i)} z^{x_i - 1} e^{-z} \, dz$$ (10.11)

and

$$D_c(\alpha/2, x_i + 1) = \int_{\alpha/2}^{\infty} \frac{1}{\Gamma(x_i + 1)} z^{x_i} e^{-z} \, dz.$$ (10.12)

The second goodness-of-fit measure is the likelihood ratio of the fit compared to a *saturated model*, where the background for each bin is exactly what was observed,

$$-2 \log \lambda = -2 \log \frac{\mathscr{L}(x_i | b_i(\theta))}{\mathscr{L}(x_i | x_i)} = 2 \sum_{i=1}^{n_b} \left[b_i(\theta) - x_i + x_i \log \frac{x_i}{b_i(\theta)} \right].$$ (10.13)

We generated 10 000 toy data sets from the best-fit model parameters and fit the background function to each toy data set to generate a distribution against which we can evaluate the χ^2 and $-2 \log \lambda$ tests. If the values of the test statistics lie off on the tails of the distributions, the fits is not accurately describing the data.

To measure the presence of signal, we fit the four-parameter function plus a signal shape at a set mass with signal strength as a function of a new parameter μ.

The background-only fit to the PF RECO data set is shown in Fig. 10.1. The fit was performed on the range $1058 < m_{jj} < 7866 \, \text{GeV}$. It is shown with signal shapes for gg, qg, and qq resonances whose cross sections are equal to those at the 95% CL upper limit set in Chapter 12. The goodness-of-fit statistics for this fit are $\chi^2 = 33.3$ with 42 degrees of freedom and $-2 \log \lambda = 35.1$. Figure 10.2 shows the distributions of these measures in

toy data. The χ^2 toys can be fit to a χ^2 distribution with 37.9 ± 0.1 effective degrees of freedom. The p-values for χ^2 and $-2\log\lambda$ are 0.87 and 0.84, respectively.
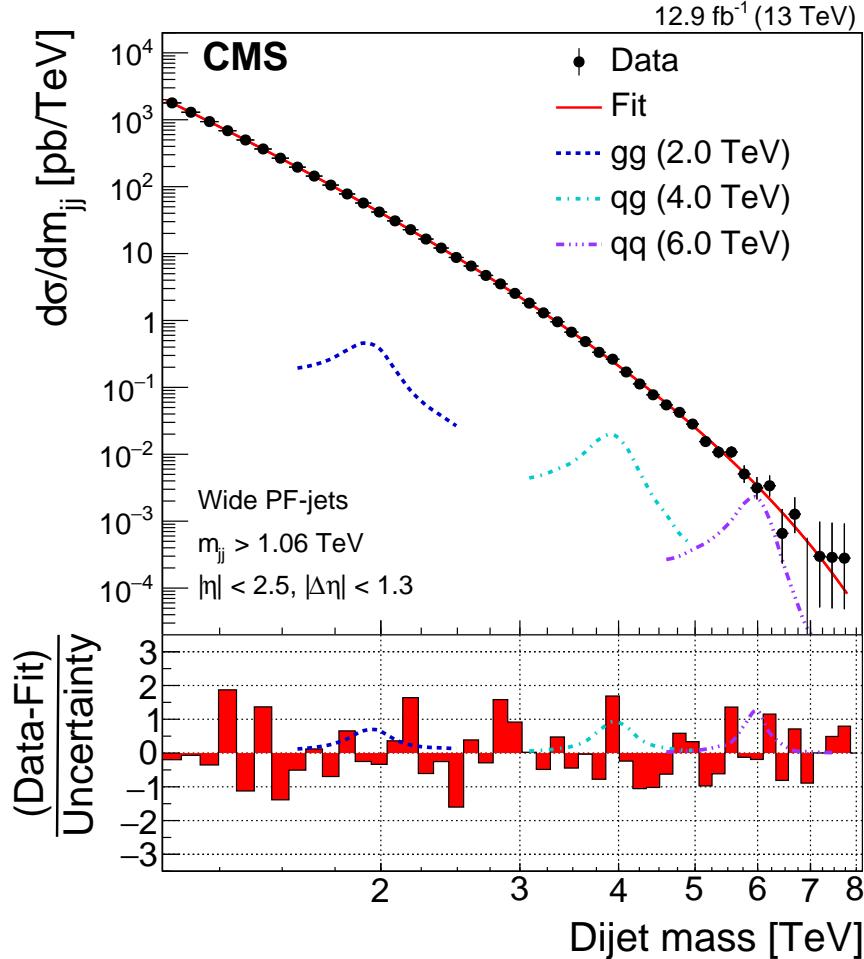
Figure 10.1: Background-only fit to the PF RECO dijet mass spectrum. Signal shapes for gg, qg, and qq resonances are displayed at the 95% CL upper limit cross sections for their masses. The lower panel shows the difference between the data and the fitted background function divided by the statistical uncertainty of the data.

The background-only fit to the calorimeter scouting data set is presented in Fig. 10.3. The fit was performed on the range $453 < m_{jj} < 2037\,\text{GeV}$. It is also shown along side signal shapes with cross sections at the 95% CL upper limits for the given processes and masses. The fit results in $\chi^2 = 17.3$ with 22 degrees of freedom and $-2\log\lambda = 17.3$. Toy distributions for the goodness-of-fit statistics are shown in Fig. 10.4, where a fit to the χ^2
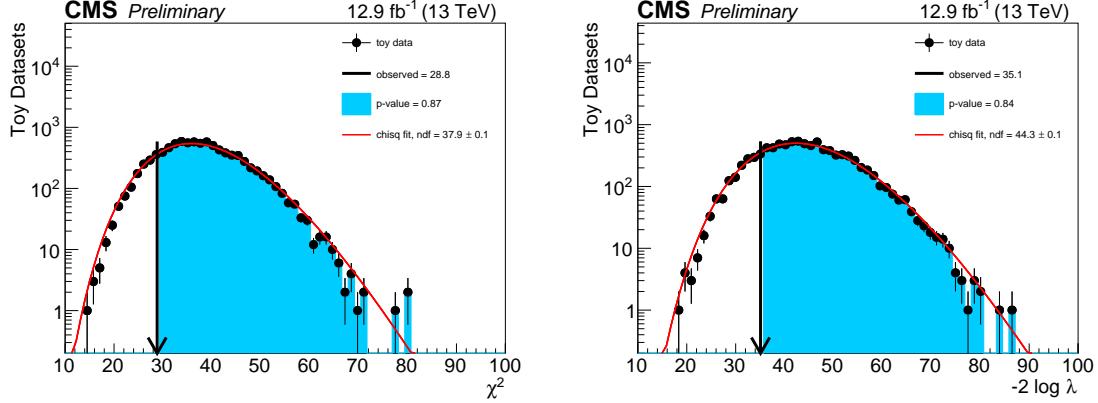
Figure 10.2: Toy distributions for the goodness-of-fit tests χ^2 (left) and $-2\log\lambda$ (right) fit to χ^2 distributions for the PF RECO fit. The observed value is shown at the arrow with its p-value calculated from the toy distributions.

toys gives 24.0 ± 0.1 effective degrees of freedom. The p-values for both χ^2 and $-2\log\lambda$ are 0.84.
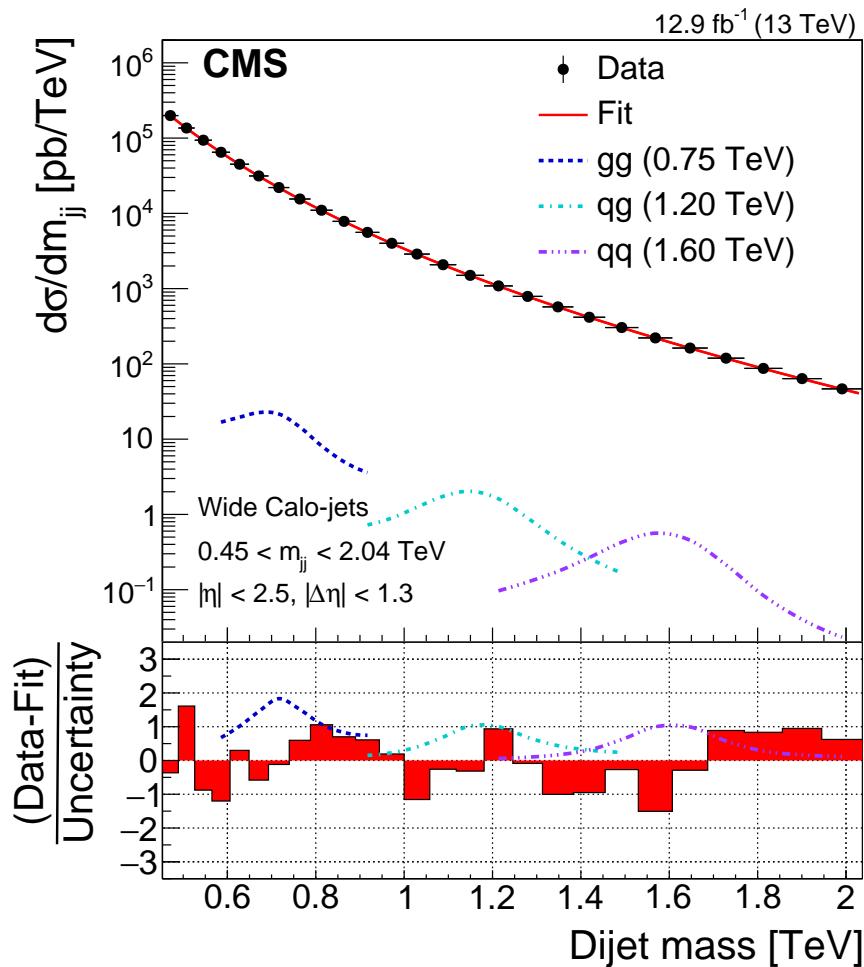
Figure 10.3: Background-only fit to the calorimeter scouting dijet mass spectrum. Signal shapes for gg, qg, and qq are displayed at the 95% CL upper limit cross sections for their masses. The lower panel shows the difference between the data and the fitted background function divided by the statistical uncertainty of the data.
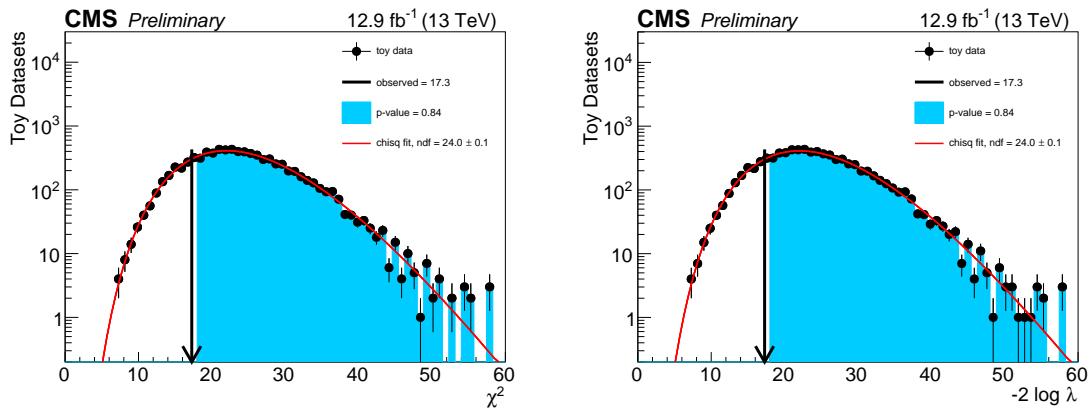
Figure 10.4: Toy distributions for the goodness-of-fit tests χ^2 (left) and $-2\log\lambda$ (right) fit to χ^2 distributions for the calorimeter scouting fit. The observed value is shown at the arrow with its p-value calculated from the toy distributions.

Chapter 11

Systematic uncertainties and limit setting

11.1 Systematic uncertainties

The dominant sources of systematic uncertainties come from the JER, JES, integrated luminosity, and background estimation.

The uncertainty on JER results in a 10% uncertainty on the resolution of the dijet mass (either for RECO PF jets or HLT calorimeter jets) [80]. The systematic uncertainty is propagated by increasing and decreasing the width of the signal shapes by 10%. The uncertainty of the JES is 2% as determined from Run 2 data with the methods described in Ref. [80]. It is propagated by scaling the mass of the signal shapes up and down by 2%. Figure 11.1 shows the effects of the systematic uncertainties for JER and JES on the signal shapes.

The systematic uncertainty in the integrated luminosity is 6.2%. It is propagated with the normalization of the signal. The systematic uncertainty in the background parametrization is handled through profiling. The background is fit at every value of signal strength. Of all the systematic uncertainties, the background uncertainty has the largest effect on the limits. The effect is greatest for gg resonances because of their greater width. It is most prominent at smaller masses for all resonances.

11.2 Limit setting procedure

Upper limits are found using the modified frequentist CL_s method [83, 84]. The limits are calculated by the Higgs Combine tool [85] with the asymptotic approximation of
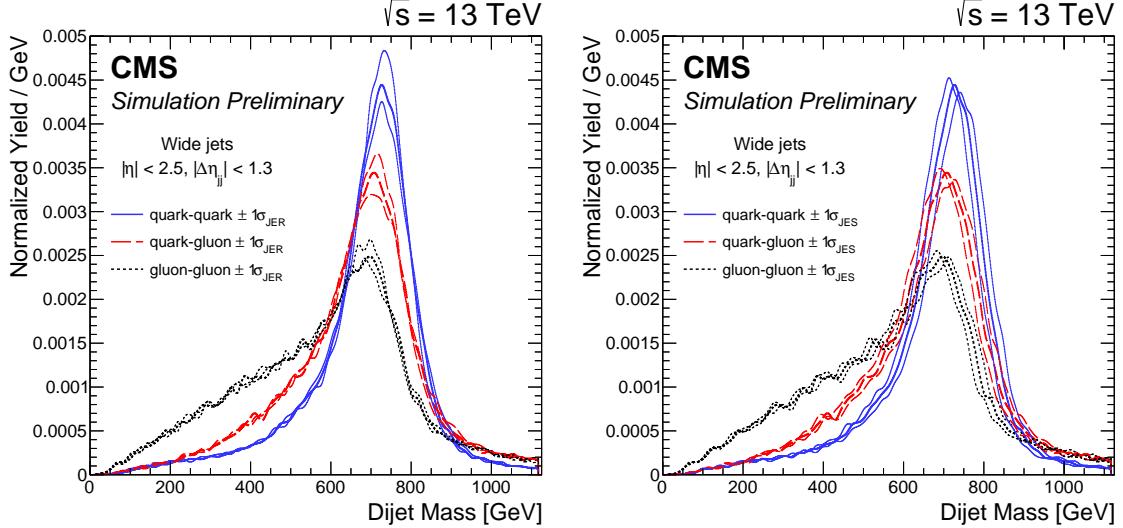
Figure 11.1: Signal shapes for gg, qg, and qq resonances at 750 GeV (thick lines) with the effect of modifying the JER (left) and JES (right) by ±1 standard deviation σ (thin lines). The integrals of all shapes have been normalized to unity.

CL_s [86]. Past dijet searches by CMS have used a Bayesian methodology to set limits, but we have found that CL_s gives similar results to the Bayesian approach.

In the CL_s method, confidence levels (CLs) are determined by the pdfs of a test statistic $-2 \log Q$, where

$$Q = \frac{\mathscr{L}(\text{data}|s+b)}{\mathscr{L}(\text{data}|b)} \tag{11.1}$$

is the ratio of the likelihood of the data for signal plus background to the likelihood of the data when there is only background. The likelihood of the data for signal strength μ and nuisance parameters θ is

$$\mathscr{L}(\text{data}|\mu, \theta) = \rho(\tilde{\theta}|\theta) \prod_{i=1}^{n_b} \text{Poisson}\left(x_i | \mu s_i(\theta) + b_i(\theta)\right), \tag{11.2}$$

where $\rho(\tilde{\theta}|\theta)$ is the systematic uncertainty pdf with default value $\tilde{\theta}$. The test statistic used is

$$\tilde{q}_\mu = -2 \log \frac{\mathscr{L}(\text{data}|\mu, \hat{\hat{\theta}}_\mu)}{\mathscr{L}(\text{data}|\hat{\mu}, \hat{\theta})} \tag{11.3}$$

with the constraint $0 \leq \hat{\mu} \leq \mu$, where $\hat{\theta}_\mu$ are the conditional maximum likelihood estimators of θ for signal strength μ and either data or pseudo-data depending on context. The parameters $\hat{\mu}$ and $\hat{\theta}$ are the best fit parameters for the signal strength and nuisance parameters. The constraint $0 \leq \hat{\mu}$ ensures that the signal rate is nonnegative. The constraint $\hat{\mu} \leq \mu$ is to produce a one-sided confidence interval.

The observed test statistic $\tilde{q}_\mu^{\text{obs}}$ for a given μ can then be calculated from the fit to data. From the pdfs for the background-only and signal plus background hypotheses we can calculate

$$CL_{s+b} = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}} | \text{signal} + \text{background}) = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} f(\tilde{q}_\mu | \mu, \hat{\theta}_\mu^{\text{obs}}) \, d\tilde{q}_\mu \qquad (11.4)$$

and

$$CL_b = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}} | \text{background-only}) = \int_{\tilde{q}_0^{\text{obs}}}^{\infty} f(\tilde{q}_\mu | 0, \hat{\theta}_0^{\text{obs}}) \, d\tilde{q}_\mu. \qquad (11.5)$$

From this we can get the ratio

$$CL_s(\mu) = \frac{CL_{s+b}}{CL_b} \qquad (11.6)$$

as a function of signal strength. The observed 95% CL upper limit on the signal strength is the value of μ such that $CL_s(\mu^{95\% \text{ CL}}) = 0.05$.

The pdfs for $\tilde{q}_\mu$ can be found from the distributions of toy pseudo-data generated about the best-fist parameters. However, if we remove the requirement that $\mu \geq 0$, then Wilks's theorem [87] says that in the asymptotic limit, where the number of events is large, $\tilde{q}_\mu$ is distributed a χ^2 distribution with one degree of freedom. Therefore, we know the value of $\tilde{q}_\mu$ needed to give $CL_s = 0.05$ without resorting to generating a set of pseudo-data.

The physical case with the constraint $\mu \geq 0$ is more complicated but has the same result of allowing the asymptotic CL_s method used in Higgs Combine to more quickly calculate the CL [85]. The pdfs are calculated using the known parameters and an "Asimov" data set. In an Asimov data set, the most representative value is substituted for the

result found from pseudo-data. This is used to evaluate $q_{\mu,A}$ for the case with expected background and nominal nuisance parameters. In the asymptotic limit $\tilde{q}_\mu$ is equivalent to

$$q_\mu = -2\log \frac{\mathscr{L}(\text{data}|\mu,\hat{\theta}_\mu)}{\mathscr{L}(\text{data}|0,\hat{\theta}_0)}. \tag{11.7}$$

We can then use

$$CL_s(\mu) = \frac{1-\Phi\left(\sqrt{q_\mu}\right)}{\Phi\left(\sqrt{q_{\mu,A}}-\sqrt{q_\mu}\right)}, \tag{11.8}$$

where $\Phi(x)$ is the cumulative distribution function of the standard Gaussian with mean 0 and standard deviation 1, to obtain the 95% CL upper limit on the signal strength.

Expected limits and 1 and 2 standard deviations bands can be found by generating toy pseudo-data sets from the best-fit parameters. The value of $\mu^{95\% \text{ CL}}$ is calculated for each toy data set and a cumulative probability distribution is created from these. The expected limit is at the median value, where the cumulative distribution crosses 50%. The ± 1 standard deviation (68%) band is found from the crossing of the 16% and 84% quantiles. The ± 2 standard deviations (95%) band is found from the crossings of the 2.5% and 97.5% quantiles. In the asymptotic limit, this can be calculated without generating pseudo-data by using

$$\sqrt{q_{\mu,A}} = \Phi^{-1}(1-\alpha\Phi(N)+N), \tag{11.9}$$

where $\Phi^{-1}(x)$ is the quantile function for the standard Gaussian, $\alpha = 0.05$ for the 95% CL, and N is the number of standard deviations from the median. The expected limit is the case where $N = 0$. The 1 and 2 standard deviation bands are calculated with $N = \pm 1$ and ± 2.

The meaning of the 95% CL is that if an ensemble of experiments were performed, the true value of the signal strength would lie within the 95% CLs in 95% of the experiments in the ensemble. It does not imply that there is a 95% probability that the true value of the signal strength lies within the CL of an individual experiment. That is a Bayesian claim that is dependent on prior probabilities.

The significance of a signal is measured by the p-value, the probability that the null hypothesis produces at least as large a signal as was observed. We typically express this probability in terms of a Gaussian with the number of standard deviations in one tail that would result in that probability

$$Z = \Phi^{-1}(1-p). \tag{11.10}$$

The background-only hypothesis is typically rejected at $Z = 5$ standard deviations, which corresponds to $p = 2.87 \times 10^{-7}$. An important benchmark is at $Z = 3$ standard deviations with $p = 1.34 \times 10^{-3}$. We reject signal hypotheses at the 95% CL, which is at $Z = 1.64$ standard deviations.

We determine the local significance of signal with the modified test statistic

$$q_0 = \begin{cases} -2\log \frac{\mathscr{L}(\text{data}|0,\hat{\hat{\theta}}_0)}{\mathscr{L}(\text{data}|\hat{\mu},\hat{\theta})} & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases}. \tag{11.11}$$

The observed local significance

$$Z = \sqrt{q_0}. \tag{11.12}$$

The significance is 0 when the best-fit signal strength is negative, as the least contribution a physical signal can make is no contribution. It increases with signal strength.

Chapter 12
Results

The data are well-described by the smoothly falling four-parameter functions fit in Figs. 10.1 and 10.3. There is no evidence for the signal of a resonance above the background. We can set limits on the models from this absence of evidence.

The significance of resonances are shown in Fig. 12.1 for the low-mass analysis and Fig. 12.2 for the high-mass analysis. The most significant excess has a local significance of 2.58 standard deviations for a 850 GeV gg resonance in the low-mass analysis. The largest local significance in the high-mass analysis is 1.68 standard deviations at 2.9 TeV for a qq resonance.

We set limits on the product of cross section, branching fraction, and experimental acceptance $\sigma \mathscr{B} A$. This allows other models to be compared where their branching fractions or acceptance of signal events by our selection criteria are different of our benchmark models. The acceptance for isotropic decays to pass $|\eta| < 2.5$ and $|\Delta\eta| < 1.3$ is approximately 0.6. The acceptance for the signal samples from benchmark models were closer to 0.5. The m_{jj} threshold is accounted for and is not a factor in the acceptance.

For the low-mass analysis, we set limits from 600 to 1600 GeV in 50 GeV steps, and for the high-mass analysis, we set limits from 1.6 to 7.5 TeV in 100 GeV steps. The limits are set for model-independent parton modes as well as for Gaussian shapes with standard deviations that are 10% the resonance mass.

The 95% CL upper limits for gg, qg, and qq low-mass resonances are presented in Fig. 12.3. Figure 12.4 shows the same for the high-mass resonances. These limits are combined to cover a range of 600 GeV to 7.5 TeV in Fig. 12.5, which summarizes the
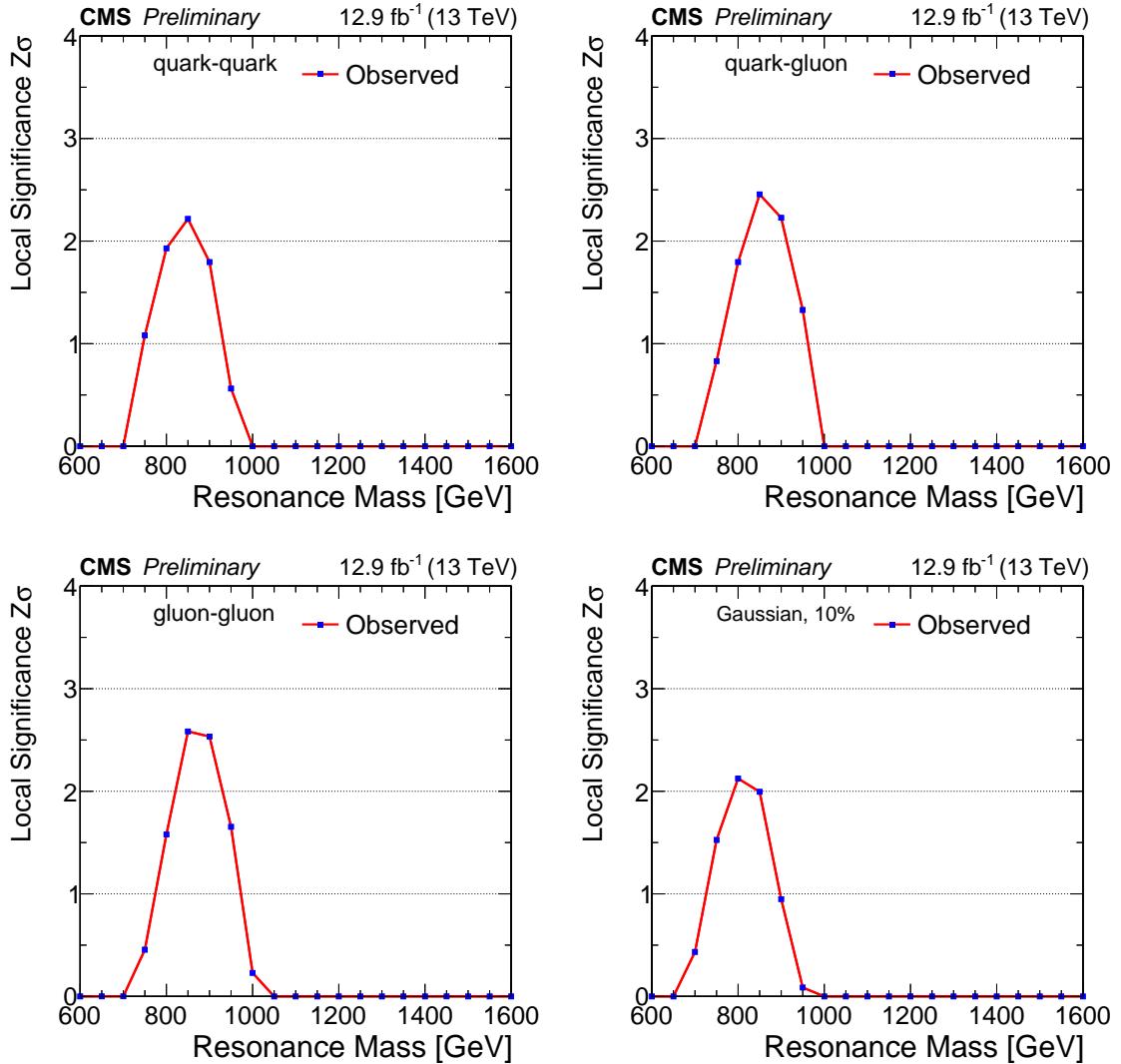
Figure 12.1: Local significance for qq resonances (top left), qg resonances (top right), gg resonances (bottom left), and Gaussian shapes (bottom right) from the low-mass analysis.
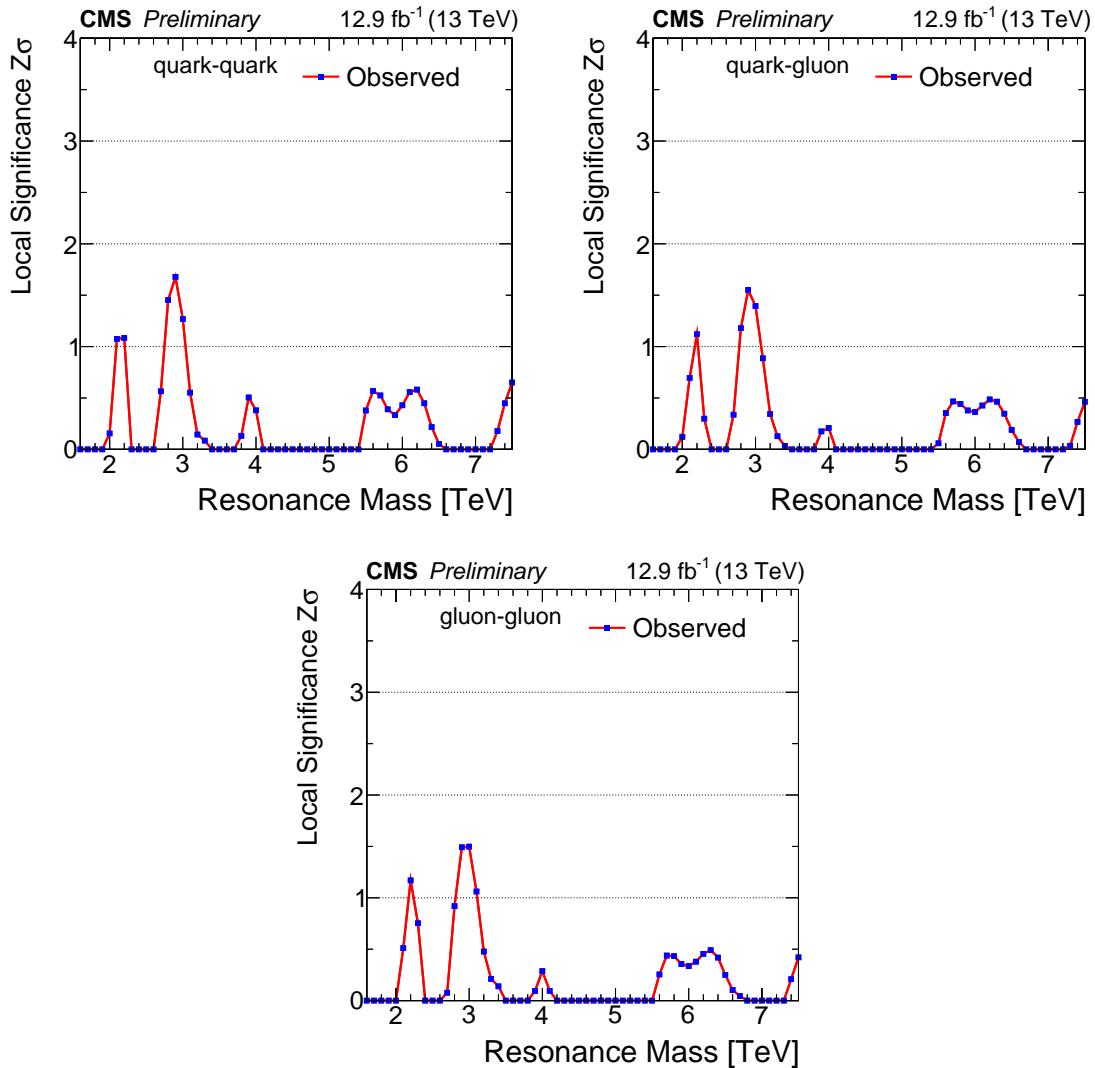
Figure 12.2: Local significance for qq resonances (top left), qg resonances (top right), and gg resonances (bottom) from the high-mass analysis.

results. The predicted values of $\sigma \mathscr{B} A$ for string resonances, excited quarks, axigluons, colorons, scalar diquarks, color-octet scalars, W$'$ bosons, Z$'$ bosons, DM mediators, and RS gravitons are presented on the plots for the parton resonances that they decay to.

The limits are dependent on the decay mode. The gg resonances have larger widths and consequently we obtain weaker limits than qq resonances. The more strongly peaked the distribution (i.e., the narrower the width), the lower the limit we can set. Additionally, for a given resonance mass, the signal shape will peak at different masses with gg peaking at the lowest mass and qq peaking at the highest mass for physical shapes. Although qq resonances still peak below the resonance mass. This also affects the limits and can be seen in shifted features for each decay mode in Fig. 12.5.

The predicted $\sigma \mathscr{B} A$ for the benchmark models in Figs. 12.3 and 12.4 are calculated in the narrow-width approximation [5] using the leading-order CTEQ6L1 PDF [88]. Next-to-leading-order K factor corrections of approximately 1.3 are applied for W$'$ and Z$'$ while a factor of approximately 1.2 is applied for axigluons and colorons [40]. The branching fractions include decays to gluons and the five lightest quarks. Top quark decays are not included except in calculating the resonance widths.

We exclude values of $\sigma \mathscr{B} A$ greater than the observed upper limits. Therefore, we are able to exclude mass ranges for models where the predicted $\sigma \mathscr{B} A$ is greater than our upper limits. Table 12.1 lists the observed and expected limits we set on the benchmark models at the 95% CL. The RS graviton decays to both gg and q$\bar{\text{q}}$. We set its limit from a combination of the gg and qq mode limits based on their contribution to the RS graviton cross section (60% gg and 40% q$\bar{\text{q}}$).

Figure 12.6 shows the limits on universal quark coupling g'_q of a leptophobic Z$'$ boson (it decays only to quarks) as a function of mass. There is an equivalence between the simplified DM and leptophobic Z$'$ boson models with Z$'$ boson mass equal to the DM mediator's mass. Their couplings are related by Eq. (2.42), which depends on the branching fraction of the mediator to quarks. When $M_{\text{Med}} < 2m_{\text{DM}}$, the mediator is
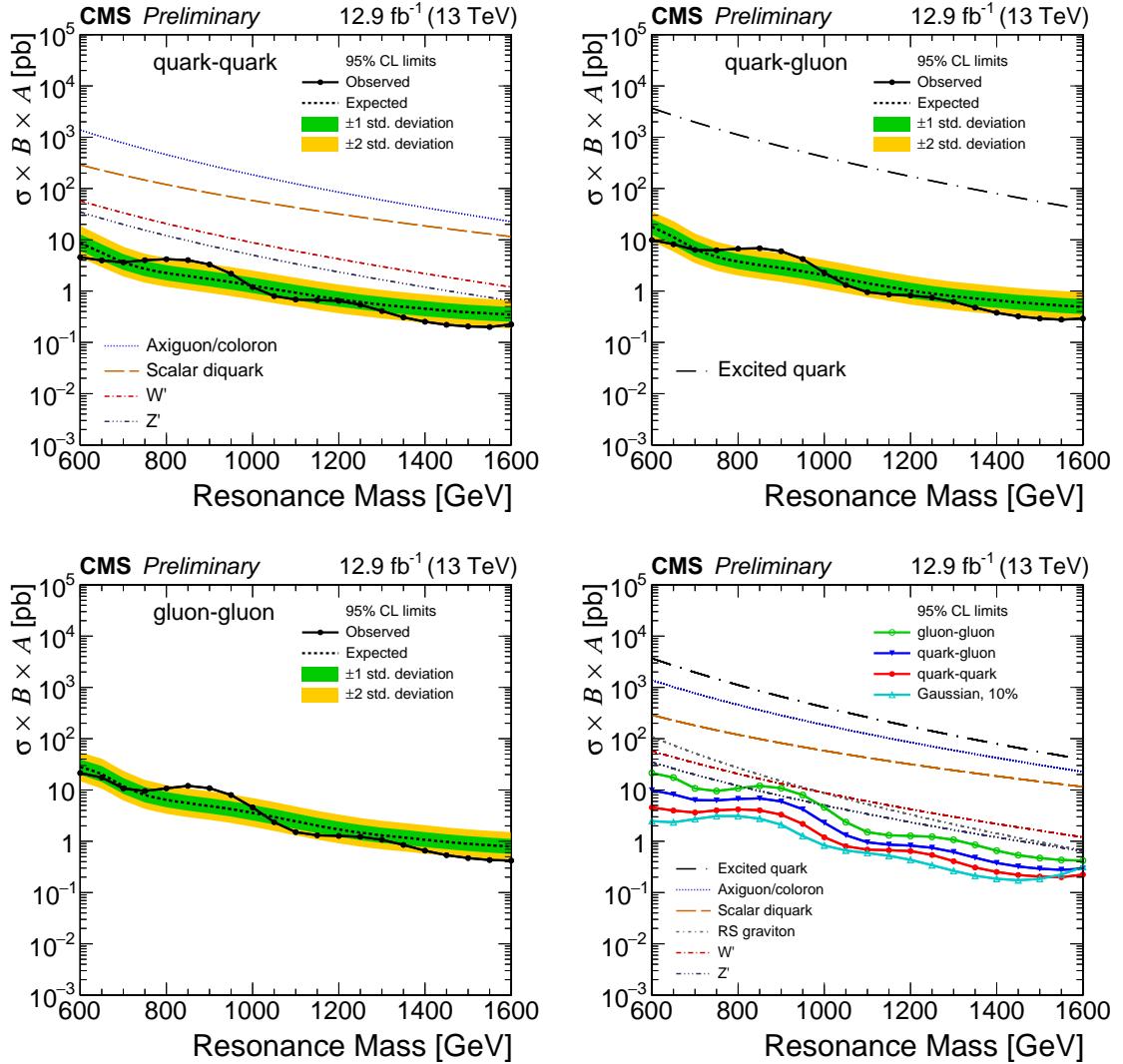
Figure 12.3: Limits on the product of cross section, branching fraction, and acceptance from the low-mass calorimeter scouting analysis for qq (top left), qg (top right), and gg (bottom left) resonances. A summary of the three parton modes and the limits for Gaussian shapes (bottom right). The observed 95% CL upper limits are the solid lines and the dashed lines are the expected values of those limits with 1 and 2 standard deviations in shaded bands. The predicted $\sigma \mathcal{B} A$ for excited quarks, axigluons/colorons, scalar diquarks, RS gravitons, W' bosons, and Z' bosons are presented for the relevant modes.
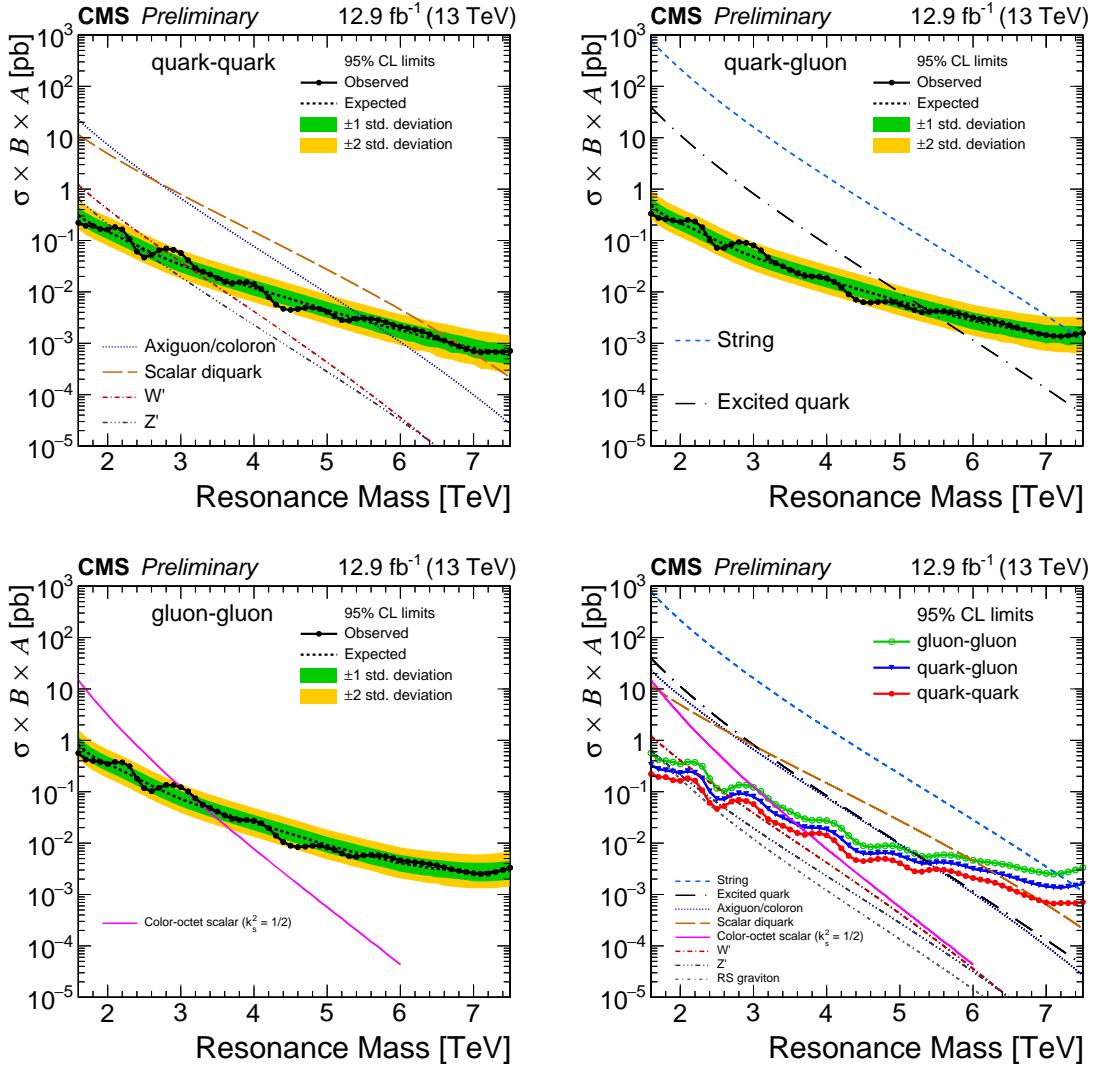
Figure 12.4: Limits on the product of cross section, branching fraction, and acceptance from the high-mass PF RECO analysis for qq (top left), qg (top right), and gg (bottom left). A summary of the three parton modes (bottom right). The observed 95% CL upper limits are the solid lines and the dashed lines are the expected values of those limits with 1 and 2 standard deviations in shaded bands. The predicted $\sigma \, \mathscr{B} A$ for string resonances, excited quarks, axigluons/colorons, scalar diquarks, color-octet scalars, W′ bosons, Z′ bosons, and RS gravitons are presented for the relevant modes.
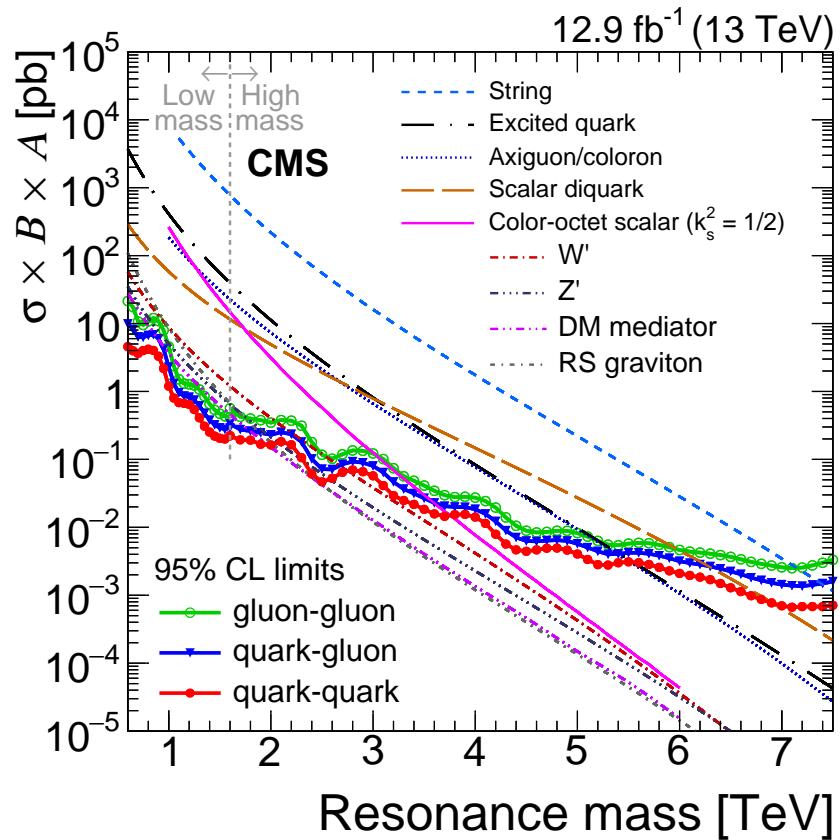
Figure 12.5: Combination of limits on the product of cross section, branching fraction, and acceptance from the low- and high-mass analyses for qq, qg, and gg resonances. The observed 95% CL upper limits are the solid lines. The predicted $\sigma \mathscr{B} A$ for string resonances, excited quarks, axigluons/colorons, scalar diquarks, color-octet scalars, W$'$ bosons, Z$'$ bosons, DM mediators, and RS gravitons are presented.

Table 12.1: Observed and expected ranges of benchmark model masses excluded at the 95% CL.

Model	Final State	Observed mass limit [TeV]	Expected mass limit [TeV]
String	qg	0.6–7.4	0.6–7.4
Scalar diquark	qq	0.6–6.9	0.6–6.8
Axigluon/coloron	$q\bar{q}$	0.6–5.5	0.6–5.6
Excited quark	qg	0.6–5.4	0.6–5.4
Color-octet scalar	gg	0.6–3.0	0.6–3.3
W$'$	$q\bar{q}$	0.6–2.7	0.6–3.1
Z$'$	$q\bar{q}$	0.6–2.1, 2.3–2.6	0.6–2.3
DM mediator ($m_{\mathrm{DM}} = 1\,\mathrm{GeV}$)	$q\bar{q}$	0.6–2.0	0.6–2.0
RS graviton	$q\bar{q}$, gg	0.6–1.9	0.6–1.8

kinematically excluded from decaying to DM particles, so it behaves the same as a leptophobic Z$'$ boson with $g_q' = g_q = 0.25$. The limits are the upper dotted line in Fig. 12.6. The lower dotted line corresponds to the simplified model with $m_{\mathrm{DM}} = 0$. With our benchmark DM model's couplings, Eq. (2.43) results in $g_q' \approx 0.182$. These two lines are the same for vector and axial-vector models.

The 95% CL limits on the simplified DM model with $g_q = 0.25$ and $q_{\mathrm{DM}} = 1$ are presented in Fig. 12.7 in the DM mass versus DM mediator mass plane. The excluded region is derived from the leptophobic Z$'$ boson limits using the branching fractions for vector and axial-vector DM mediators to quarks. The diagonal line $M_{\mathrm{Med}} = 2m_{\mathrm{DM}}$ corresponds to the upper dotted horizontal line in Fig. 12.6. For a DM mass $m_{\mathrm{DM}} = 1\,\mathrm{GeV}$, DM mediators with mass between 0.6 and 2.0 TeV are excluded at the 95% CL. The excluded mediator region increases with greater m_{DM} because of larger branching fraction to $q\bar{q}$. The maximum excluded mediator mass range is from 0.6 to 2.65 TeV as the exclusion region crosses the diagonal line. Above where it crosses the diagonal line ($m_{\mathrm{DM}} > 1.325\,\mathrm{TeV}$) the excluded mediator mass range is constant with respect to m_{DM}. The results are presented alongside constraints from cosmological relic density of DM from Ref. [89]. The constraints are based on astronomical measurements [90, 91] and
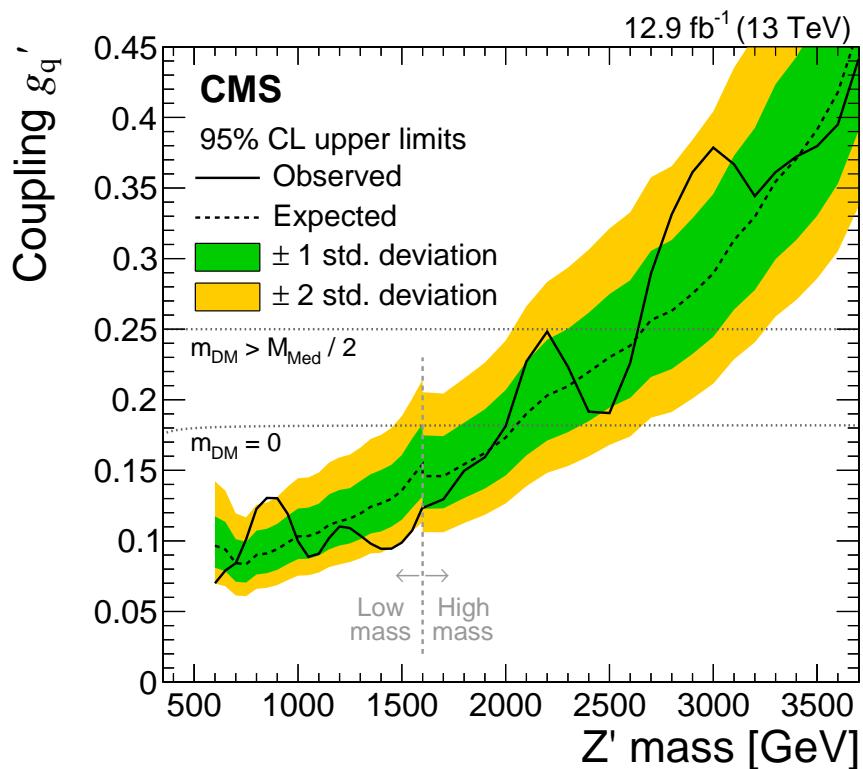
Figure 12.6: Limits on the universal quark coupling g_q' as a function of mass for a leptophobic Z' boson that only couples to quarks. The observed 95% CL upper limits are the solid lines while the dashed lines are the expected limits with 1 and 2 standard deviations in shaded bands. Dotted horizontal lines show special cases where the simplified DM model has the same cross section for dijet DM mediator resonances as the leptophobic Z' boson model.

MADDM version 2.0.6 simulation [92, 93].
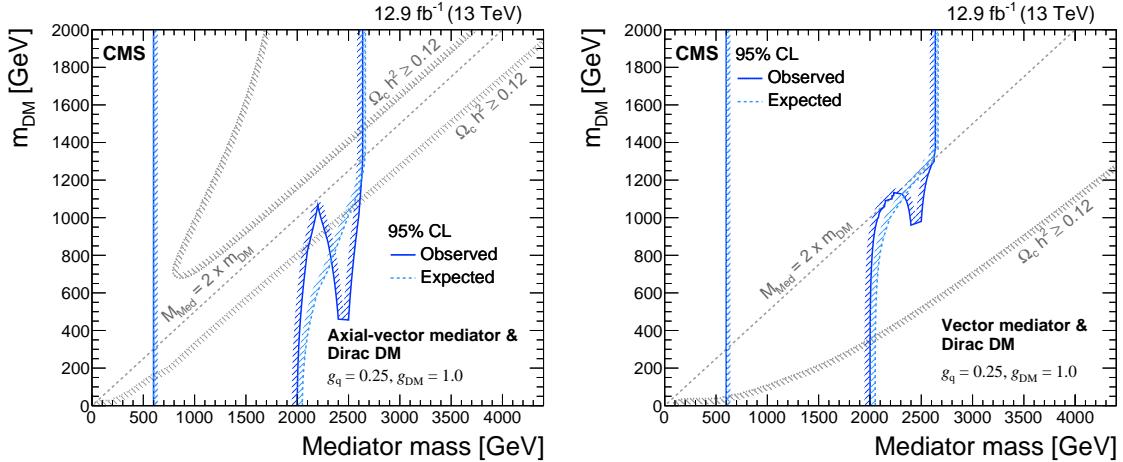
Figure 12.7: Limits at the 95% CL on DM mass and DM mediator mass for axial-vector mediator (left) and vector mediator (right) with $g_q = 0.25$ and $g_{DM} = 1$. The observed limits are solid and expected limits are dashed. The diagonal diagonal grey dashed line is at $M_{Med} = 2m_{DM}$, above which the mediator cannot decay to two DM particles. The grey limits are constraints from the cosmological relic density of DM from Ref. [89].

Direct-detection experiments rely on the interaction of DM with nuclei. They report their results in the DM–nucleon cross section versus DM mass plane using the 90% CL. Figure 12.8 shows the excluded regions along with direct-detection results. Axial-vector models give rise to spin-dependent (SD) interactions with limits set by the PICO experiments [94, 95], IceCube [96], and Super-Kamiokande [97]. Vector models give rise to spin-independent (SI) interactions with limits set by LUX 2016 [98], PandaX-II [99], CDMSLite 2015 [100], and CRESST-II 2015 [101].

The highest mass dijet event was at 7.7 TeV. A display of the event is shown in Fig. 12.9. The leading jet had $p_T = 3.61$ TeV and $\eta = 0.32$. The subleading jet has $p_T = 3.38$ TeV and $\eta = -0.56$. The jets were separated by $|\Delta\eta| = 0.88$ and $|\Delta\phi| = 3.13$.
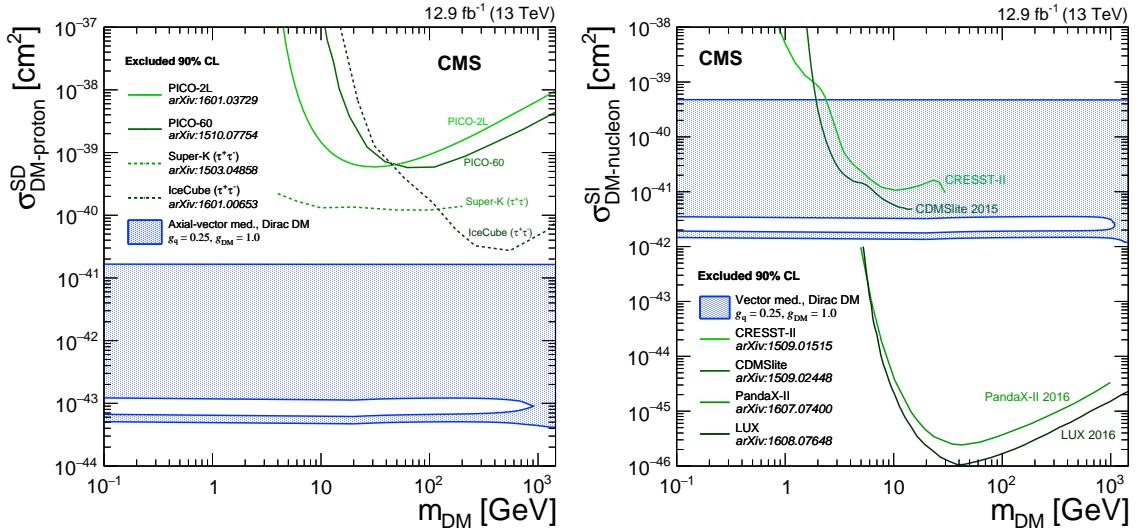
Figure 12.8: Limits at the 90% CL on DM–nucleon interaction cross section and DM mass. The shaded regions are excluded for axial-vector (left) and vector (right) mediators with Dirac fermion DM and couplings $g_q = 0.25$ and $g_{DM} = 1$. The excluded region is compared with direct-detection limits.
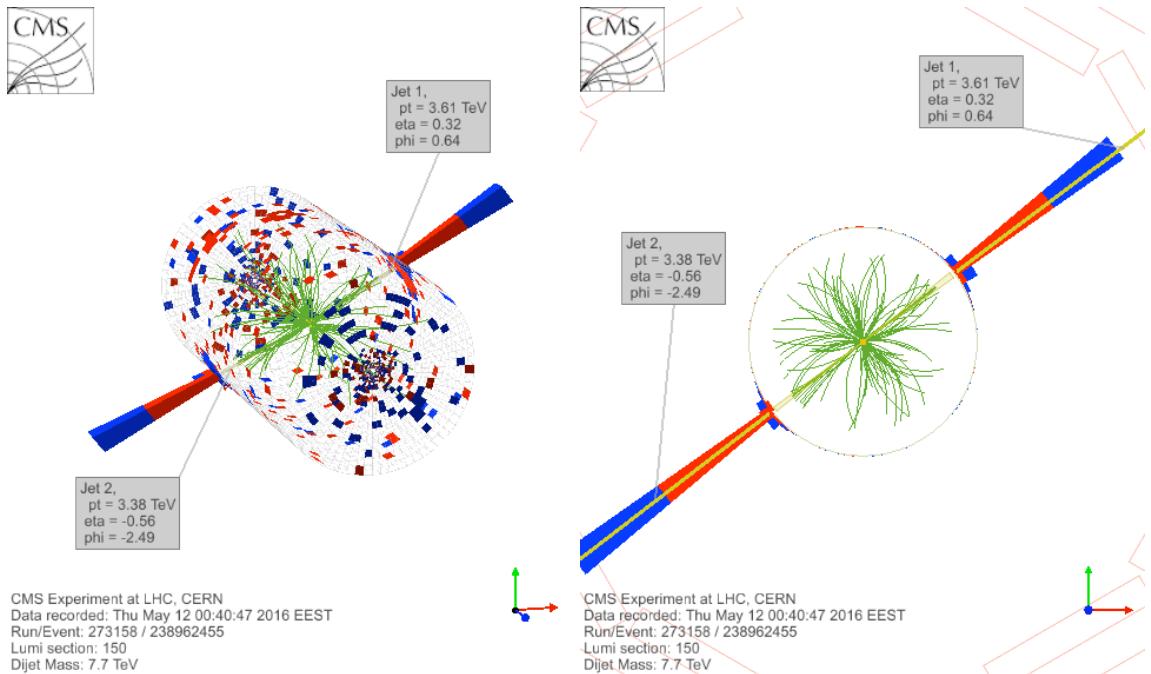
Figure 12.9: Event with the largest dijet mass, 7.7 TeV. Tracks are shown in green, energy deposited in the ECAL in red, and energy deposited in the HCAL in blue.

Chapter 13

Conclusion

We performed two related analyses for low-mass and high-mass narrow dijet resonances. Using $12.9\,\mathrm{fb}^{-1}$ of $\sqrt{s} = 13\,\mathrm{TeV}$ data collected by CMS, we found the dijet mass spectrum to be smoothly falling with no significant excess. The analysis using calorimeter scouting set limits on low-mass resonances including at $750\,\mathrm{GeV}$. The 95% CL on $\sigma\,\mathscr{B}\,A$ for gluon–gluon resonances at $750\,\mathrm{GeV}$ is $9.54\,\mathrm{pb}$. This limit is not as strong as the one set with $\sqrt{s} = 8\,\mathrm{TeV}$ data scouting, which had a greater integrated luminosity of $18.8\,\mathrm{fb}^{-1}$ [27]. The high-mass analysis excluded benchmark models to higher masses than past analyses have.

No evidence was observed of narrow dijet resonances. We exclude (starting at $600\,\mathrm{GeV}$) string resonances below $7.4\,\mathrm{TeV}$, scalar diquarks below $6.9\,\mathrm{TeV}$, axigluons and colorons below $5.5\,\mathrm{TeV}$, excited quarks below $5.4\,\mathrm{TeV}$, color-octet scalars below $3.0\,\mathrm{TeV}$, W' bosons below $2.7\,\mathrm{TeV}$, Z' bosons below $1.9\,\mathrm{TeV}$ and from 2.3 to $2.6\,\mathrm{TeV}$, and RS gravitons below $1.9\,\mathrm{TeV}$. We set the first dijet limits on DM mediators, excluding them below $2.0\,\mathrm{TeV}$ for vector and axial-vector mediators with $g_q = 0.25$ and $g_{\mathrm{DM}} = 1$. Limits on DM cross section derived from limits on the mediator are more sensitive than limits from direct-detection experiments from spin-dependent cross sections. These limits depend on the model and parameters. Changing the parameters will affect the limits that can be derived from this analysis.

At the same time as the release of this analysis in August 2016, CMS and ATLAS released the results of their most recent diphoton analyses. They found no excess in the new data [102, 103]. The excess seen in 2015 was likely upward statistical fluctuations in

both experiments rather than a resonance from new physics.

This analysis has been submitted to *Phys. Lett. B* for publication [104].

Bibliography

[1] ATLAS and CMS Collaborations, "ATLAS and CMS physics results from Run 2", last modified 15 December 2015, `https://indico.cern.ch/event/442432/`.

[2] A. Strumia, "Interpreting the 750 GeV digamma excess: a review", 2016, `arXiv:1605.09401`.

[3] CMS Collaboration, "Search for resonant production of high-mass photon pairs in proton–proton collisions at $\sqrt{s} = 8$ and 13 TeV", *Phys. Rev. Lett.* **117** (2016) 051802, `doi:10.1103/PhysRevLett.117.051802`, `arXiv:1606.04093`.

[4] ATLAS Collaboration, "Search for resonances in diphoton events at $\sqrt{s} = 13$ TeV with the ATLAS detector", *JHEP* **09** (2016) 001, `doi:10.1007/JHEP09(2016)001`, `arXiv:1606.03833`.

[5] R. M. Harris and K. Kousouris, "Searches for dijet resonances at hadron colliders", *Int. J. Mod. Phys. A* **26** (2011) 5005, `doi:10.1142/S0217751X11054905`, `arXiv:1110.5302`.

[6] UA1 Collaboration, "Measurement of the inclusive jet cross-section at the CERN p$\bar{\text{p}}$ collider", *Phys. Lett. B* **172** (1986) 461, `doi:10.1016/0370-2693(86)90290-X`.

[7] UA1 Collaboration, "Two-jet mass distributions at the CERN proton–antiproton collider", *Phys. Lett. B* **209** (1988) 127, `doi:10.1016/0370-2693(88)91843-6`.

[8] UA2 Collaboration, "A measurement of two-jet decays of the W and Z bosons at the CERN $\bar{\text{p}}$p collider", *Z. Phys. C* **49** (1991) 17, `doi:10.1007/BF01570793`.

[9] UA2 Collaboration, "A search for new intermediate vector mesons and excited quarks decaying to two-jets at the CERN $\bar{\text{p}}$p collider", *Nucl. Phys. B* **400** (1993) 3, `doi:10.1016/0550-3213(93)90395-6`.

[10] CDF Collaboration, "Search for quark compositeness, axigluons and heavy particles using the dijet invariant mass spectrum observed in p$\bar{\text{p}}$ collisions", *Phys. Rev. Lett.* **71** (1993) 2542, `doi:10.1103/PhysRevLett.71.2542`.

[11] CDF Collaboration, "Search for new particles decaying to dijets in p$\bar{\text{p}}$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **74** (1995) 3538, `doi:10.1103/PhysRevLett.74.3538, arXiv:hep-ex/9501001`.

[12] CDF Collaboration, "Search for new particles decaying to dijets at CDF", *Phys. Rev. D* **55** (1997) R5263, `doi:10.1103/PhysRevD.55.R5263`, `arXiv:hep-ex/9702004`.

[13] CDF Collaboration, "Search for new particles decaying into dijets in proton–antiproton collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev. D* **79** (2009) 112002, `doi:10.1103/PhysRevD.79.112002, arXiv:0812.4036`.

[14] D0 Collaboration, "Search for new particles in the two-jet decay channel with the D0 detector", *Phys. Rev. D* **69** (2004) 111101, `doi:10.1103/PhysRevD.69.111101, arXiv:hep-ex/0308033`.

[15] CMS Collaboration, "Search for narrow resonances and quantum black holes in inclusive and b-tagged dijet mass spectra from pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **01** (2013) 013, `doi:10.1007/JHEP01(2013)013`, `arXiv:1210.2387`.

[16] CMS Collaboration, "Search for resonances in the dijet mass spectrum from 7 TeV pp collisions at CMS", *Phys. Lett. B* **704** (2011) 123, `doi:10.1016/j.physletb.2011.09.015, arXiv:1107.4771`.

[17] CMS Collaboration, "Search for dijet resonances in 7 TeV pp collisions at CMS", *Phys. Rev. Lett.* **105** (2010) 211801, `doi:10.1103/PhysRevLett.105.211801,10.1103/PhysRevLett.106.029902, arXiv:1010.0203`.

[18] CMS Collaboration, "Search for resonances and quantum black holes using dijet mass spectra in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev. D* **91** (2015) 052009, `doi:10.1103/PhysRevD.91.052009, arXiv:1501.04198`.

[19] CMS Collaboration, "Search for narrow resonances using the dijet mass spectrum in pp collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev. D* **87** (2013) 114015, `doi:10.1103/PhysRevD.87.114015, arXiv:1302.4794`.

[20] CMS Collaboration, "Search for narrow resonances decaying to dijets in proton–proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. Lett.* **116** (2016) 071801, `doi:10.1103/PhysRevLett.116.071801, arXiv:1512.01224`.

[21] ATLAS Collaboration, "Search for new physics in the dijet mass distribution using 1 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV collected by the ATLAS detector", *Phys. Lett. B* **708** (2012) 37, `doi:10.1016/j.physletb.2012.01.035`, `arXiv:1108.6311`.

[22] ATLAS Collaboration, "ATLAS search for new phenomena in dijet mass and angular distributions using pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **01** (2013) 029, `doi:10.1007/JHEP01(2013)029, arXiv:1210.1718`.

[23] ATLAS Collaboration, "Search for new physics in dijet mass and angular distributions in pp collisions at $\sqrt{s} = 7$ TeV measured with the ATLAS detector", *New J. Phys.* **13** (2011) 053044, `doi:10.1088/1367-2630/13/5/053044`, `arXiv:1103.3864`.

[24] ATLAS Collaboration, "Search for new particles in two-jet final states in 7 TeV proton–proton collisions with the ATLAS detector at the LHC", *Phys. Rev. Lett.* **105** (2010) 161801, `doi:10.1103/PhysRevLett.105.161801`, `arXiv:1008.2461`.

[25] ATLAS Collaboration, "Search for new phenomena in the dijet mass distribution using pp collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Rev. D* **91** (2015) 052007, `doi:10.1103/PhysRevD.91.052007`, `arXiv:1407.1376`.

[26] ATLAS Collaboration, "Search for new phenomena in dijet mass and angular distributions from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Lett. B* **754** (2016) 302, `doi:10.1016/j.physletb.2016.01.032`, `arXiv:1512.01530`.

[27] CMS Collaboration, "Search for narrow resonances in dijet final states at $\sqrt{s} = 8$ TeV with the novel CMS technique of data scouting", *Phys. Rev. Lett.* **117** (2016) 031802, `doi:10.1103/PhysRevLett.117.031802`, `arXiv:1604.08907`.

[28] ATLAS Collaboration, "Search for light dijet resonances with the ATLAS detector using a Trigger-object Level Analysis in LHC pp collisions at $\sqrt{s} = 13$ TeV", ATLAS Conference Note ATLAS-CONF-2016-030, 2016.

[29] J. Bagger, C. Schmidt, and S. King, "Axigluon production in hadronic collisions", *Phys. Rev. D* **37** (1988) 1188, `doi:10.1103/PhysRevD.37.1188`.

[30] CDF Collaboration, "Two-jet invariant-mass distribution at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. D* **41** (1990) 1722, `doi:10.1103/PhysRevD.41.1722`.

[31] Wikimedia Commons, "Standard Model of Elementary Particles", last modified 1 October 2014, `https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg`.

[32] Particle Data Group Collaboration, "Review of particle physics", *Chin. Phys. C* **40** (2016) 100001, `doi:10.1088/1674-1137/40/10/100001`.

[33] U. Baur, I. Hinchliffe, and D. Zeppenfeld, "Excited quark production at hadron colliders", *Int. J. Mod. Phys. A* **2** (1987) 1285, `doi:10.1142/S0217751X87000661`.

[34] U. Baur, M. Spira, and P. M. Zerwas, "Excited quark and lepton production at hadron colliders", *Phys. Rev. D* **42** (1990) 815, `doi:10.1103/PhysRevD.42.815`.

[35] L. Randall and R. Sundrum, "Large mass hierarchy from a small extra dimension", *Phys. Rev. Lett.* **83** (1999) 3370, `doi:10.1103/PhysRevLett.83.3370`, `arXiv:hep-ph/9905221`.

[36] L. Randall and R. Sundrum, "An alternative to compactification", *Phys. Rev. Lett.* **83** (1999) 4690, `doi:10.1103/PhysRevLett.83.4690`, `arXiv:hep-th/9906064`.

[37] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo, "Phenomenology of the Randall–Sundrum gauge hierarchy model", *Phys. Rev. Lett.* **84** (2000) 2080, `doi:10.1103/PhysRevLett.84.2080`, `arXiv:hep-ph/9909255`.

[38] P. H. Frampton and S. L. Glashow, "Chiral color: An alternative to the standard model", *Phys. Lett. B* **190** (1987) 157, `doi:10.1016/0370-2693(87)90859-8`.

[39] R. S. Chivukula, A. G. Cohen, and E. H. Simmons, "New strong interactions at the Tevatron?", *Phys. Lett. B* **380** (1996) 92, `doi:10.1016/0370-2693(96)00464-9`, `arXiv:hep-ph/9603311`.

[40] R. S. Chivukula, A. Farzinnia, J. Ren, and E. H. Simmons, "Hadron collider production of massive color-octet vector bosons at next-to-leading order", *Phys. Rev. D* **87** (2013) 094011, `doi:10.1103/PhysRevD.87.094011`, `arXiv:1303.1120`.

[41] E. H. Simmons, "Coloron phenomenology", *Phys. Rev. D* **55** (1997) 1678, `doi:10.1103/PhysRevD.55.1678`, `arXiv:hep-ph/9608269`.

[42] R. S. Chivukula, E. H. Simmons, and N. Vignaroli, "Distinguishing dijet resonances at the LHC", *Phys. Rev. D* **91** (2015) 055019, `doi:10.1103/PhysRevD.91.055019`, `arXiv:1412.3094`.

[43] E. Eichten, I. Hinchliffe, K. D. Lane, and C. Quigg, "Supercollider physics", *Rev. Mod. Phys.* **56** (1984) 579, `doi:10.1103/RevModPhys.56.579`. [Erratum: *Rev. Mod. Phys.* **58** (1986) 1065–1073 `doi:10.1103/RevModPhys.58.1065`].

[44] B. A. Dobrescu and F. Yu, "Coupling-mass mapping of dijet peak searches", *Phys. Rev. D* **88** (2013) 035021, `doi:10.1103/PhysRevD.88.035021`, `arXiv:1306.2629`. [Erratum: *Phys. Rev. D* **90** (2014) 079901 `doi:10.1103/PhysRevD.90.079901`].

[45] S. Cullen, M. Perelstein, and M. E. Peskin, "TeV strings and collider probes of large extra dimensions", *Phys. Rev. D* **62** (2000) 055012, `doi:10.1103/PhysRevD.62.055012`, `arXiv:hep-ph/0001166`.

[46] L. A. Anchordoqui et al., "Dijet signals for low mass strings at the LHC", *Phys. Rev. Lett.* **101** (2008) 241803, `doi:10.1103/PhysRevLett.101.241803`, `arXiv:0808.0497`.

[47] J. L. Hewett and T. G. Rizzo, "Low-energy phenomenology of superstring-inspired E_6 models", *Phys. Rept.* **183** (1989) 193, `doi:10.1016/0370-1573(89)90071-9`.

[48] G. Katsilieris, O. Korakiantis, and S. D. P. Vlassopulos, "Two-jet rates affected by exotic particles at hadron colliders", *Phys. Lett. B* **288** (1992) 221, `doi:10.1016/0370-2693(92)91980-N`.

[49] G. Busoni et al., "Recommendations on presenting LHC searches for missing transverse energy signals using simplified s-channel models of dark matter", 2016, `arXiv:1603.04156`.

[50] J. Abdallah et al., "Simplified models for dark matter searches at the LHC", *Phys. Dark Univ.* **9–10** (2015) 8, `doi:10.1016/j.dark.2015.08.001`, `arXiv:1506.03116`.

[51] M. Chala et al., "Constraining dark sectors with monojets and dijets", *JHEP* **07** (2015) 089, `doi:10.1007/JHEP07(2015)089`, `arXiv:1503.05916`.

[52] R. Franceschini et al., "What is the $\gamma\gamma$ resonance at 750 GeV?", *JHEP* **03** (2016) 144, `doi:10.1007/JHEP03(2016)144`, `arXiv:1512.04933`.

[53] M. Benedikt et al., "LHC design report Vol. 3: The LHC injector chain", Technical Report CERN-2004-003-V-3, 2004.

[54] O. S. Bruning et al., "LHC design report Vol. 1: The LHC main ring", Technical Report CERN-2004-003-V1, 2004.

[55] G. Brianti, "Large Hadron Collider in the LEP Tunnel: A feasibility study of possible options", *Conf. Proc. C* **8405141** (1984) 207.

[56] CMS Collaboration, "Letter of intent for a general purpose detector at the LHC", Technical Report CERN-LHCC-92-03, 1992.

[57] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, `arXiv:1207.7235`.

[58] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, `doi:10.1088/1748-0221/3/08/S08004`.

[59] C. W. Fabjan and F. Gianotti, "Calorimetry for particle physics", *Rev. Mod. Phys.* **75** (2003) 1243, `doi:10.1103/RevModPhys.75.1243`.

[60] P. Adzic et al., "Energy resolution of the barrel of the CMS electromagnetic calorimeter", *JINST* **2** (2007) P04004, `doi:10.1088/1748-0221/2/04/P04004`.

[61] CMS Collaboration, "Energy calibration and resolution of the CMS electromagnetic calorimeter in pp Collisions at $\sqrt{s} = 7\,\text{TeV}$", *JINST* **8** (2013) P09009, `doi:10.1088/1748-0221/8/09/P09009, arXiv:1306.2016`.

[62] CMS Collaboration, "Design, performance, and calibration of CMS hadron-barrel calorimeter wedges", CMS Note CMS-NOTE-2006-138, 2007.

[63] V. D. Elvira, "Measurement of the pion energy response and resolution in the CMS HCAL test beam 2002 experiment", CMS Note CMS-NOTE-2004-020, 2004.

[64] CMS Collaboration, "CMS, The TriDAS project, Technical design report, Vol. 1: The trigger systems", Technical Report CMS-TDR-6-1, 2000.

[65] CMS Collaboration, "CMS, The TriDAS project, Technical design report, Vol. 2: Data acquisition and high-level trigger", Technical Report CMS-TDR-6, 2002.

[66] CMS Collaboration, "The CMS high level trigger", *Eur. Phys. J.* **C46** (2006) 605, `doi:10.1140/epjc/s2006-02495-8, arXiv:hep-ex/0512077`.

[67] CMS Collaboration, "CMS luminosity measurement for the 2015 data taking period", CMS Physics Analysis Summary CMS-PAS-LUM-15-001, 2016.

[68] CMS Collaboration, "CMS luminosity based on pixel cluster counting—Summer 2012 update", CMS Physics Analysis Summary CMS-PAS-LUM-12-001, 2012.

[69] S. van der Meer, "Calibration of the effective beam height in the ISR", Technical Report CERN-ISR-PO-68-31, 1968.

[70] A. Bell et al., "Fast beam conditions monitor BCM1F for the CMS experiment", *Nucl. Instrum. Meth. A* **614** (2010) 433, `doi:10.1016/j.nima.2009.12.056, arXiv:0911.2480`.

[71] CMS Collaboration, "The CMS pixel luminosity telescope", *Nucl. Instrum. Meth. A* **824** (2016) 304, `doi:10.1016/j.nima.2015.09.104`.

[72] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A brief introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852, `doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820`.

[73] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016) 155, `doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815`.

[74] GEANT4 Collaboration, "GEANT4: A simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, `doi:10.1016/S0168-9002(03)01368-8`.

[75] CMS Collaboration, "Jet performance in CMS", *PoS* **EPS-HEP2013** (2013) 433.

[76] CMS Collaboration, "Particle-flow event reconstruction in CMS and performance for jets, taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.

[77] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", CMS Physics Analysis Summary CMS-PAS-PFT-10-001, 2010.

[78] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, `doi:10.1088/1126-6708/2008/04/063`, `arXiv:0802.1189`.

[79] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, `doi:10.1140/epjc/s10052-012-1896-2`, `arXiv:1111.6097`.

[80] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, `doi:10.1088/1748-0221/12/02/P02014`, `arXiv:1607.03663`.

[81] CMS Collaboration, "Jet algorithms performance in 13 TeV data", CMS Physics Analysis Summary CMS-PAS-JME-16-003, 2017.

[82] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling", eConf C **0303241** (2003) MOLT007, `arXiv:physics/0306116`.

[83] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, `doi:10.1016/S0168-9002(99)00498-2`, `arXiv:hep-ex/9902006`.

[84] A. L. Read, "Presentation of search results: The CL_s technique", *J. Phys. G* **28** (2002) 2693, `doi:10.1088/0954-3899/28/10/313`.

[85] ATLAS Collaboration, CMS Collaboration, LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in summer 2011", CMS Note/ATLAS Pub CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-011, 2011.

[86] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, `doi:10.1140/epjc/s10052-011-1554-0`, `arXiv:1007.1727`. [Erratum: *Eur. Phys. J. C* **73** (2013) 2501 `doi:10.1140/epjc/s10052-013-2501-z`].

[87] S. S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses", *Annals Math. Statist.* **9** (1938) 60, `doi:10.1214/aoms/1177732360`.

[88] J. Pumplin et al., "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP* **07** (2002) 012, `doi:10.1088/1126-6708/2002/07/012`, `arXiv:hep-ph/0201195`.

[89] T. du Pree, K. Hahn, P. Harris, and C. Roskas, "Cosmological constraints on dark matter models for collider searches", 2016, `arXiv:1603.08525`.

[90] WMAP Collaboration, "Three-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Implications for cosmology", *Astrophys. J. Suppl.* **170** (2007) 377, `doi:10.1086/513700`, `arXiv:astro-ph/0603449`.

[91] Planck Collaboration, "Planck 2013 results. XVI. Cosmological parameters", *Astron. Astrophys.* **571** (2014) A16, `doi:10.1051/0004-6361/201321591`, `arXiv:1303.5076`.

[92] M. Backović, K. Kong, and M. McCaskey, "MadDM v.1.0: Computation of dark matter relic abundance using MadGraph 5", *Phys. Dark Univ.* **5–6** (2014) 18, `doi:10.1016/j.dark.2014.04.001`, `arXiv:1308.4955`.

[93] M. Backović et al., "Direct detection of dark matter with MadDM v.2.0", *Phys. Dark Univ.* **9–10** (2015) 37, `doi:10.1016/j.dark.2015.09.001`, `arXiv:1505.04190`.

[94] PICO Collaboration, "Improved dark matter search results from PICO-2L Run 2", *Phys. Rev. D* **93** (2016) 061101, `doi:10.1103/PhysRevD.93.061101`, `arXiv:1601.03729`.

[95] PICO Collaboration, "Dark matter search results from the PICO-60 CF_3I bubble chamber", *Phys. Rev. D* **93** (2016) 052014, `doi:10.1103/PhysRevD.93.052014`, `arXiv:1510.07754`.

[96] IceCube Collaboration, "Improved limits on dark matter annihilation in the Sun with the 79-string IceCube detector and implications for supersymmetry", *JCAP* **1604** (2016) 022, `doi:10.1088/1475-7516/2016/04/022`, `arXiv:1601.00653`.

[97] Super-Kamiokande Collaboration, "Search for neutrinos from annihilation of captured low-mass dark matter particles in the Sun by Super-Kamiokande", *Phys. Rev. Lett.* **114** (2015) 141301, `doi:10.1103/PhysRevLett.114.141301`, `arXiv:1503.04858`.

[98] LUX Collaboration, "Results from a search for dark matter in the complete LUX exposure", *Phys. Rev. Lett.* **118** (2017) 021303, `doi:10.1103/PhysRevLett.118.021303`, `arXiv:1608.07648`.

[99] PandaX-II Collaboration, "Dark matter results from first 98.7 days of data from the PandaX-II experiment", *Phys. Rev. Lett.* **117** (2016) 121303, `doi:10.1103/PhysRevLett.117.121303`, `arXiv:1607.07400`.

[100] SuperCDMS Collaboration, "New results from the search for low-mass weakly interacting massive particles with the CDMS low ionization threshold experiment", *Phys. Rev. Lett.* **116** (2016) 071301, `doi:10.1103/PhysRevLett.116.071301`, `arXiv:1509.02448`.

[101] CRESST Collaboration, "Results on light dark matter particles with a low-threshold CRESST-II detector", *Eur. Phys. J. C* **76** (2016) 25, `doi:10.1140/epjc/s10052-016-3877-3`, `arXiv:1509.01515`.

[102] CMS Collaboration, "Search for high-mass diphoton resonances in proton–proton collisions at 13 TeV and combination with 8 TeV search", *Phys. Lett. B* **767** (2017) 147, `doi:10.1016/j.physletb.2017.01.027`, `arXiv:1609.02507`.

[103] ATLAS Collaboration, "Search for scalar diphoton resonances with 15.4 fb^{-1} of data collected at $\sqrt{s} = 13$ TeV in 2015 and 2016 with the ATLAS detector", ATLAS Conference Note ATLAS-CONF-2016-059, 2016.

[104] CMS Collaboration, "Search for dijet resonances in proton–proton collisions at $\sqrt{s} = 13$ TeV and constraints on dark matter and other models", *Phys. Lett. B* (2016) `doi:10.1016/j.physletb.2017.02.012`, `arXiv:1611.03568`.

Appendix A
Generator code

The generator fragments used to produce the MC samples in Section 5.2 are presented here. They are written for the PYTHIA 8 interface of CMSSW_7_1_X.

QCD

Code for QCD background with $800 < \hat{p}_\mathrm{T} < 1000\,\mathrm{GeV}$. Other samples were generated in different bins of $\hat{p}_\mathrm{T}$.

```python
import FWCore.ParameterSet.Config as cms


from Configuration.Generator.Pythia8CommonSettings_cfi \
    import *
from Configuration.Generator.Pythia8CUEP8M1Settings_cfi \
    import *


generator = cms.EDFilter("Pythia8GeneratorFilter",
    maxEventsToPrint = cms.untracked.int32(1),
    pythiaPylistVerbosity = cms.untracked.int32(1),
    filterEfficiency = cms.untracked.double(1.0),
    pythiaHepMCVerbosity = cms.untracked.bool(False),
    comEnergy = cms.double(13000.0), # [GeV]

    crossSection = cms.untracked.double(32.2928),
```

```
PythiaParameters = cms.PSet(

    pythia8CommonSettingsBlock,

    pythia8CUEP8M1SettingsBlock, # Use CUETP8M1 tune

    processParameters = cms.vstring(

        # Turn on all QCD hard processes

        'HardQCD:all = on',

        # pT hat range

        'PhaseSpace:pTHatMin = 800  ',

        'PhaseSpace:pTHatMax = 1000  ',

    ),

    parameterSets = cms.vstring('pythia8CommonSettings',

        'pythia8CUEP8M1Settings',

        'processParameters',

    )

)

)


configurationMetadata = cms.untracked.PSet(

    version = cms.untracked.string('\$Revision$'),

    name = cms.untracked.string('\$Source$'),

    annotation = cms.untracked.string(

        'QCD pthat 800to1000 GeV, 13 TeV, TuneCUETP8M1')

)
```

$gg \rightarrow G \rightarrow gg$

Code for the process $gg \rightarrow G \rightarrow gg$ with graviton mass $m_G = m_1 = 1\,\mathrm{TeV}$ and $k/\overline{M}_{Pl} = 0.01$, which implies

$$x m_G = \frac{\sqrt{2} x_1 k}{\overline{M}_{Pl}} = 0.54, \tag{A.1}$$

where $x_1 \approx 3.83$ is the first root of the Bessel function $J_1(x)$. Other samples were produced with varying m_G with the ratio $k/\overline{M}_{Pl}$ fixed to 0.01.

```python
import FWCore.ParameterSet.Config as cms


from Configuration.Generator.Pythia8CommonSettings_cfi \
    import *
from Configuration.Generator.Pythia8CUEP8M1Settings_cfi \
    import *


generator = cms.EDFilter("Pythia8GeneratorFilter",
    comEnergy = cms.double(13000.0), # [GeV]
    crossSection = cms.untracked.double(6.186),
    filterEfficiency = cms.untracked.double(1),
    maxEventsToPrint = cms.untracked.int32(0),
    pythiaHepMCVerbosity = cms.untracked.bool(False),
    pythiaPylistVerbosity = cms.untracked.int32(1),
    PythiaParameters = cms.PSet(
        pythia8CommonSettingsBlock,
        pythia8CUEP8M1SettingsBlock, # Use CUETP8M1 tune
        processParameters = cms.vstring(
            # Produce with gluon fusion
            'ExtraDimensionsG*:gg2G* = on',
            # kappa/M_Pl = 0.01
```

```
            'ExtraDimensionsG*:kappaMG = 0.54',

            '5100039:m0 = 1000', # Graviton mass [GeV]

            '5100039:onMode = off', # Turn off all decays

            '5100039:onIfAny = 21' # Turn on decays to gluons

        ),

        parameterSets = cms.vstring('pythia8CommonSettings',

            'pythia8CUEP8M1Settings',

            'processParameters',

        )

    )

)


ProductionFilterSequence = cms.Sequence(generator)
```

qq → G → qq

Code for the process $qq \to G \to qq$ with graviton mass $m_G = 1\,\text{TeV}$ and $k/\overline{M}_{Pl} = 0.01$. Other samples were produced with varying m_G with the ratio $k/\overline{M}_{Pl}$ fixed to 0.01.

```
import FWCore.ParameterSet.Config as cms


from Configuration.Generator.Pythia8CommonSettings_cfi \
    import *
from Configuration.Generator.Pythia8CUEP8M1Settings_cfi \
    import *


generator = cms.EDFilter("Pythia8GeneratorFilter",
    comEnergy = cms.double(13000.0), # [GeV]
    crossSection = cms.untracked.double(1.61),
```

```
        filterEfficiency = cms.untracked.double(1),

        maxEventsToPrint = cms.untracked.int32(0),

        pythiaHepMCVerbosity = cms.untracked.bool(False),

        pythiaPylistVerbosity = cms.untracked.int32(1),

        PythiaParameters = cms.PSet(

            pythia8CommonSettingsBlock,

            pythia8CUEP8M1SettingsBlock, # Use CUETP8M1 tune

            processParameters = cms.vstring(

                # Produce with quark fusion

                'ExtraDimensionsG*:ffbar2G* = on',

                # kappa/M_Pl = 0.01

                'ExtraDimensionsG*:kappaMG = 0.54',

                '5100039:m0 = 1000', # Graviton mass [GeV]

                '5100039:onMode = off', # Turn off all decays

                # Turn on decays to quarks

                '5100039:onIfAny = 1 2 3 4 5'

            ),

            parameterSets = cms.vstring('pythia8CommonSettings',

                'pythia8CUEP8M1Settings',

                'processParameters',

            )

        )

    )


ProductionFilterSequence = cms.Sequence(generator)
```

qg → q* → qg

Code for the processes qg → q* → qg with excited quark mass mass $m^* = 1\,\mathrm{TeV}$, compositeness scale $\Lambda = 1\,\mathrm{TeV}$, and strength of strong, EM, and weak couplings $f_s = f = f' = 1$. Other samples were produced with varying $m^* = \Lambda$ and strengths $f_s = f = f'$ fixed to 1.

```python
import FWCore.ParameterSet.Config as cms


from Configuration.Generator.Pythia8CommonSettings_cfi \
    import *
from Configuration.Generator.Pythia8CUEP8M1Settings_cfi \
    import *


generator = cms.EDFilter("Pythia8GeneratorFilter",
    comEnergy = cms.double(13000.0), # [GeV]
    crossSection = cms.untracked.double(807.7),
    filterEfficiency = cms.untracked.double(1),
    maxEventsToPrint = cms.untracked.int32(0),
    pythiaHepMCVerbosity = cms.untracked.bool(False),
    pythiaPylistVerbosity = cms.untracked.int32(1),
    PythiaParameters = cms.PSet(
        pythia8CommonSettingsBlock,
        pythia8CUEP8M1SettingsBlock, # Use CUETP8M1 tune
        processParameters = cms.vstring(
            # Quark-gluon fusion
            'ExcitedFermion:ug2uStar = on',
            'ExcitedFermion:dg2dStar = on',
            '4000001:m0 = 1000.', # dStar mass
            '4000001:onMode = off', # Turn off decays
```

```
            '4000001:onIfMatch = 21 1', # Decay to dg

            '4000002:m0 = 1000.', # uStar mass

            '4000002:onMode = off', # Turn off decays

            '4000002:onIfMatch = 21 2', # Decay to ug

            # Compositeness scale [GeV]

            'ExcitedFermion:Lambda = 1000.',

            'ExcitedFermion:coupFprime = 1.', # U(1) strength

            'ExcitedFermion:coupF = 1.', # SU(2) strength

            'ExcitedFermion:coupFcol = 1.' # SU(3) strength

        ),

        parameterSets = cms.vstring('pythia8CommonSettings',

            'pythia8CUEP8M1Settings',

            'processParameters',

        )

    )

)


ProductionFilterSequence = cms.Sequence(generator)
```

Appendix B
HLT to RECO corrections

We derived corrections between the wide jets constructed from HLT AK4 calorimeter jets used in the low-mass scouting analysis and those constructed from AK4 PF jets that were fully reconstructed in the MC samples and the high-mass analysis. We use the ParkingScoutingMonitor data set from Section 5.1, which gives us events with both data scouting objects and ones that have been reconstructed offline. The wide jets derived from data scouting calorimeter jets are labeled as HLT, while those reconstructed offline are labeled as RECO.

The corrections are derived using a tag-and-probe method similar to the derivation of JECs for AK4 jets (Section 6.4). One of the HLT jets is labeled the *probe* jet. The RECO data set has two jets matched to the HLT ones in η-ϕ space. The one that coincides with the HLT probe jet is the RECO probe jet. The other one is the RECO *tag* jet. To double the number of events used to derive the corrections we repeat this with the other HLT jet labeled the probe causing the two RECO jets switch between tag and probe. The goal is to derive corrections for the four-momentum of the HLT probe jets to what a RECO jet would be by comparing their p_T to that of the RECO tag and probe jets.

To begin with we calculate three ratios

$$R_1 = \frac{p_T(\text{HLT, probe})}{p_T(\text{RECO, tag})}, \tag{B.1}$$

$$R_2 = \frac{p_T(\text{RECO, probe})}{p_T(\text{RECO, tag})}, \tag{B.2}$$

$$R_3 = \frac{p_T(\text{HLT, probe})}{p_T(\text{RECO, probe})}. \tag{B.3}$$

The average values of these ratios are sown in Fig. B.1 as a function of $p_T(\text{RECO, tag})$. The values are calculated for results based on whether they passed the ZeroBias, DoubleMu, CaloJet40, L1HT, or HT250 triggers.
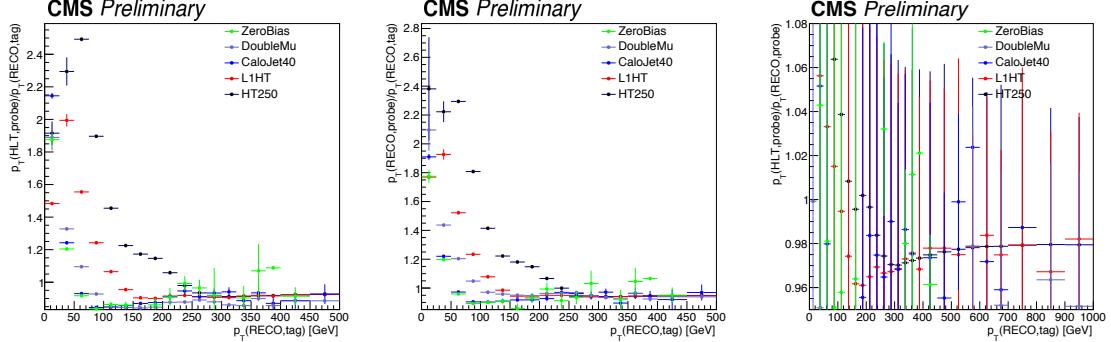

Figure B.1: Ratios R_1 (left), R_2 (center), and R_3 (right) versus $p_T(\text{RECO, tag})$ for events that pass the ZeroBias, DoubleMu, CaloJet40, L1HT, or HT250 triggers.

Dividing the averages of R_1 by those of R_2 gives us

$$\frac{\langle R_1 \rangle}{\langle R_2 \rangle} = \left\langle \frac{p_T(\text{HLT, probe})}{p_T(\text{RECO, tag})} \right\rangle \Big/ \left\langle \frac{p_T(\text{RECO, tag})}{p_T(\text{RECO, probe})} \right\rangle = \frac{\langle p_T(\text{HLT, probe}) \rangle}{\langle p_T(\text{RECO, probe}) \rangle}. \tag{B.4}$$

The error on this can be taken from the error on R_3 in Fig. B.1 (right). We can use the value of $\langle p_T(\text{HLT, probe}) \rangle$ as a function of $p_T(\text{RECO, tag})$ in Fig. B.2 to get $\langle R_1 \rangle / \langle R_2 \rangle$ as a function of the HLT probe jet's p_T. Figure B.3 shows this ratio as a function of $p_T(\text{RECO, tag})$ and $\langle p_T(\text{HLT, probe}) \rangle$.

To derive corrections, we want to use triggers that are fully efficient. The behavior at low p_T is biased as the triggers turn on. We therefore choose two regions: the data for $140 < p_T < 300\,\text{GeV}$ comes from the L1HT trigger and the data for $p_T > 300\,\text{GeV}$ comes from the HT250 trigger.

The data points from Fig. B.3 (right) for the two trigger regions are presented again in Fig. B.4 along with a fit to the data of the function

$$f(x) = p_0 + p_1 \log x + p_2 \log^2 x. \tag{B.5}$$
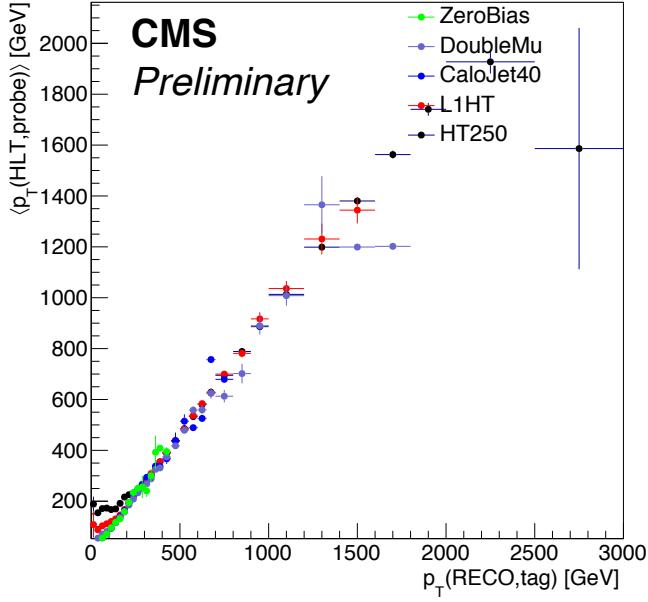
Figure B.2: The distribution of $\langle p_{\mathrm{T}}(\mathrm{HLT}, \mathrm{probe})\rangle$ as a function of $p_{\mathrm{T}}(\mathrm{RECO}, \mathrm{tag})$ for events that pass the ZeroBias, DoubleMu, CaloJet40, L1HT, or HT250 triggers.
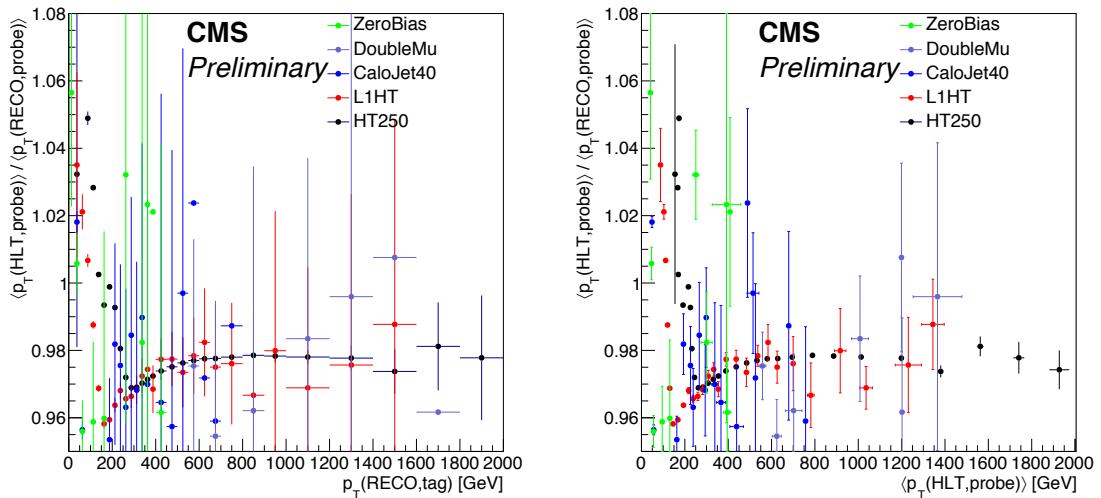
Figure B.3: The distribution of $\langle R_1\rangle/\langle R_2\rangle$ versus $p_{\mathrm{T}}(\mathrm{RECO}, \mathrm{tag})$ (left) and $\langle p_{\mathrm{T}}(\mathrm{HLT}, \mathrm{probe})\rangle$ (right) for events that pass the ZeroBias, DoubleMu, CaloJet40, L1HT, or HT250 triggers.

The best fit was with $p_0 = -31.7198$, $p_1 = 8.586\,11$, and $p_2 = -0.622\,092$ with $\chi^2 = 76$ and 22 degrees of freedom. The function has a stationary point at $e^{-p_1/2p_2} = 993.264\,\text{GeV}$. We correct the four-momenta of HLT jets by the factor $1/[1 + f(p_T)/100]$ for $p_T < 993.264\,\text{GeV}$ and by the constant $1/[1 + f(993.264\,\text{GeV})/100] = 1.02138$ for $p_T \geq 993.264\,\text{GeV}$.
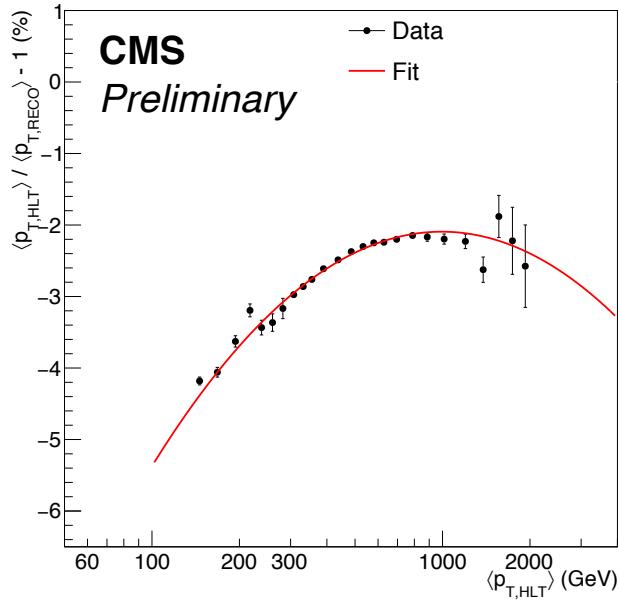
Figure B.4: The value $\langle p_T(\text{HLT, probe})\rangle / \langle p_T(\text{RECO, probe})\rangle - 1$ as a percentage versus $\langle p_T(\text{HLT, probe})\rangle$ for trigger regions with full efficiency along with fit to derive HLT to RECO corrections.

Appendix C
Number of parameters for background function

Many functions could be used to fit the falling dijet mass spectrum with various numbers of parameters. We used Eq. (10.2), which has been used in past dijet searches at CMS, ATLAS, and CDF. However, using fewer or more parameters with similar functions could produce nearly as good a fit to the data. Because adding more parameters will necessarily allow the function to better fit the data, we want to use the function with the fewest parameters that fits the data well. We use an F test to determine when adding additional parameters does not provide a significant increase to the quality of the fit.

The test is run on four nested functions

$$
\begin{aligned}
f_2 &= \frac{p_0}{(m_{jj}/\sqrt{s})^{p_2}} \\
f_3 &= p_0 \frac{(1-m_{jj}/\sqrt{s})^{p_1}}{(m_{jj}/\sqrt{s})^{p_2}} \\
f_4 &= p_0 \frac{(1-m_{jj}/\sqrt{s})^{p_1}}{(m_{jj}/\sqrt{s})^{p_2+p_3\log(m_{jj}/\sqrt{s})}} \\
f_5 &= p_0 \frac{(1-m_{jj}/\sqrt{s})^{p_1}}{(m_{jj}/\sqrt{s})^{p_2+p_3\log(m_{jj}/\sqrt{s})+p_4\log^2(m_{jj}/\sqrt{s})}},
\end{aligned}
\tag{C.1}
$$

where f_i is f_{i+1} with one parameter set to zero. The fits of these four functions to the low-mass calorimeter scouting and high-mass PF RECO dijet mass spectra are shown in Fig. C.1. The best-fit parameters are given in Table C.1.

Table C.1: Parameters, χ^2, number of degrees of freedom (dof), residual sum of squares RSS, number of parameters n, and number of non-zero bins N for two-, three-, four-, and five-parameter fits to PF RECO and calorimeter scouting dijet mass distributions.

| Data set	Function	p_0	p_1	p_2	p_3	p_4	χ^2	dof	RSS	n	N
PF RECO	f_2	3.1677×10^{-7}	0	6.2790	0	0	25360	45	3.8550×10^9	2	
	f_3	1.5134×10^{-5}	9.0590	5.0257	0	0	49.697	44	1.0452×10^7	3	46
	f_4	3.4640×10^{-6}	7.3496	5.9652	0.16361	0	34.173	43	4.1060×10^6	4	
	f_5	2.3800×10^{-5}	8.8764	4.1833	-0.45789	-0.078487	34.010	43	4.0775×10^6	5	
Calorimeter scouting	f_2	3.3111×10^{-6}	0	5.4047	0	0	545980	24	1.0143×10^{13}	2	
	f_3	8.5895×10^{-5}	14.314	4.5751	0	0	3710.3	23	8.2331×10^{10}	3	26
	f_4	9.4999×10^{-7}	5.8409	6.8323	0.29972	0	17.377	22	3.0590×10^8	4	
	f_5	9.5274×10^{-7}	5.8478	6.8313	0.29977	3.5109×10^{-5}	17.334	21	3.0439×10^8	5	

Figure C.1: Fits of two-, three-, four-, and five-parameter functions to the dijet mass spectrum from calorimeter scouting (left) and PF RECO (right). The lower panels show the difference between data and the fits divided by the statistical uncertainty of the data.

The F statistic for nested functions f_i and f_{i+1} is

$$F_{i+1,i} = \frac{\left(\frac{RSS_i - RSS_{i+1}}{n_{i+1} - n_i}\right)}{\left(\frac{RSS_{i+1}}{N - n_{i+1}}\right)}, \tag{C.2}$$

where N is the number of non-zero bins in the region being fit, n_i is the number of parameters in function f_i, and RSS_i is the residual sum of squares for function f_i. The residual sum of squares

$$RSS = \sum_{\text{bins}} \left(\text{data}_{\text{bin}} - \text{fit}_{\text{bin}}\right)^2, \tag{C.3}$$

where data_{bin} is the number of events in the bin, fit_{bin} is the integral of f_i over the bin times the integrated luminosity, and the sum skips bins that are empty.

The F test is performed by starting with the null hypothesis that f_{i+1} with its additional parameter does not provide a significantly better fit than f_i. The F statistic

has an F-distribution with $(n_{i+1} - n_i, N - n_{i+1}) = (1, N - n_{i+1})$ degrees of freedom $f(F; 1, N - n_{i+1})$. If the confidence level

$$CL_{i+1,i} = 1 - \int_{-\infty}^{F_{i+1,i}} f(F; 1, N - n_{i+1}) \, \mathrm{d}F < \alpha = 0.05 \tag{C.4}$$

we reject the null hypothesis—the additional parameter in f_{i+1} provides a significantly better fit than f_i. This process is iterated until the null hypothesis is not rejected. We are left with a function that gives a significantly better fit than the nested functions with fewer parameters but adding an additional parameter does not provide a significant increase in the quality of the fit.

The F statistic and CLs for the fits to PF RECO and calorimeter scouting data are shown in Table C.2. For both data sets, the three-parameter function provides a significant improvement over the two-parameter function. Likewise for the four-parameter function with respect to the three-parameter one. However, $CL_{54} > 0.05$ for both data sets, so we can say that the five-parameter function does not provide a significant improvement over the four-parameter one. We therefore use the four-parameter function in Chapter 10 to fit the dijet mass spectrum.

Table C.2: The F statistics and confidence levels for PF RECO and calorimeter scouting data. We are limited in computing CL below 1.1102×10^{-16}.

Data set	Functions	F	CL
PF RECO	f_3 vs. f_2	15816	$<1.1102 \times 10^{-16}$
	f_4 vs. f_3	64.915	4.6594×10^{-10}
	f_5 vs. f_4	0.28641	0.59542
Calorimeter scouting	f_3 vs. f_2	2810.6	$<1.1102 \times 10^{-16}$
	f_4 vs. f_3	5899.2	$<1.1102 \times 10^{-16}$
	f_5 vs. f_4	0.10363	0.750702

Appendix D

Comparison of limits in 2015 and 2016 data

The first dijet analysis of 13 TeV data by CMS was performed in 2015 [20]. That work was based on $2.4\,\mathrm{fb}^{-1}$ and only included a high-mass analysis. It also used a Bayesian methodology to set limits. This analysis uses $12.9\,\mathrm{fb}^{-1}$ of data collected in 2016 with limits determined by CL_s. There is no overlap in the data sets.

The 95% CL upper limits on $\sigma\,\mathscr{B}\,A$ set here with 2016 data should be better than those from the 2015 analysis by the square root of the ratios of the integrated luminosities of their data sets. Therefore, the upper limits on $\sigma\,\mathscr{B}\,A$ from 2016 should be roughly a factor of 0.43 lower than the ones set in 2015. Figure D.1 shows a comparison of the limits from 2015 and 2016. The ratios of the expected limits are close to 0.43. The ratios for the observed limits also tend to be close, but there are large fluctuations due to the stochastic nature of the data.
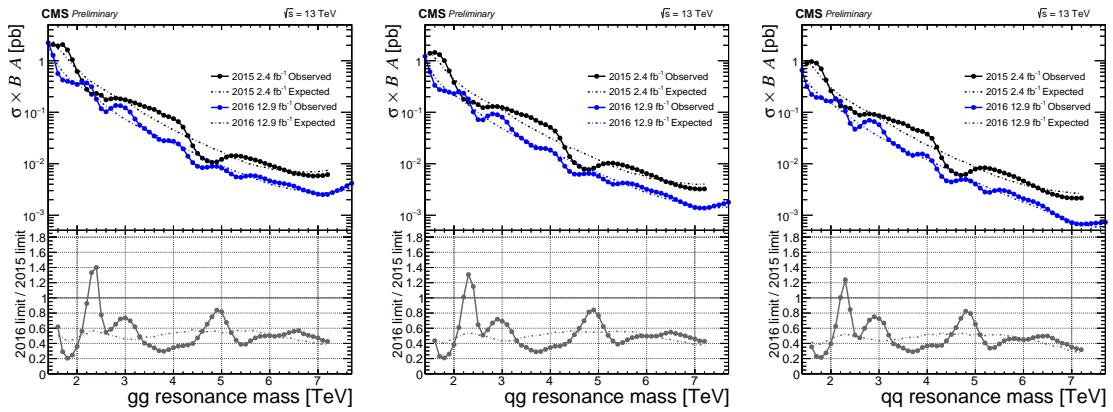

Figure D.1: Observed and expected 95% CL upper limits on $\sigma\,\mathscr{B}\,A$ from 2015 and 2016 data for gg (left), qg (center), and qq (right) resonances. The lower panels show the ratios of limits from 2016 to those from 2015 in the range of masses in which both analyses set limits.

Including systematic uncertainties when setting limits is necessary to obtain results that do not overestimate our confidence (see Chapter 11). To evaluate the effects of the systematic uncertainties, we can set limits without them (using only statistical uncertainties) and take the ratio of the limit with both statistical and systematic uncertainties to the limit with only statistical uncertainties. Figure D.2 shows a comparison of these limit ratios for 2015 and 2016 data. The limit ratios for 2015 and 2016 are similar for both observed and expected limits.
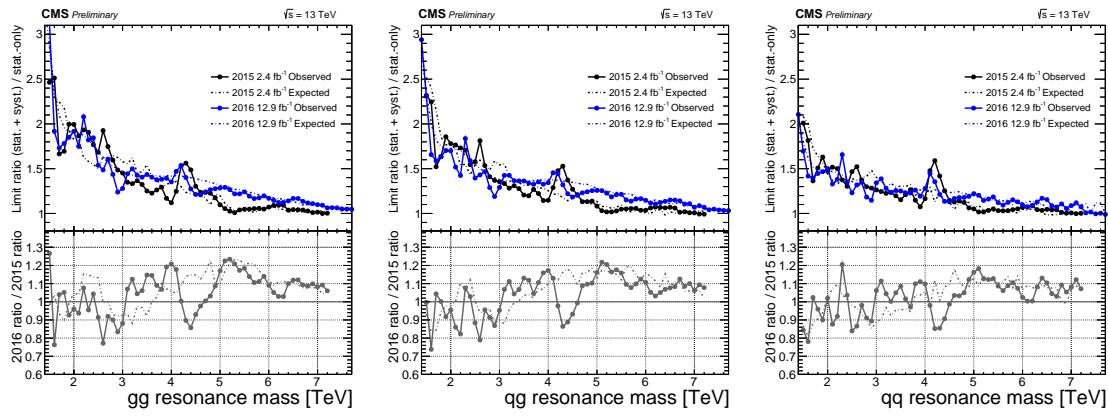

Figure D.2: Ratios of observed and expected 95% CL upper limits with and without systematic uncertainties from 2015 and 2016 data for gg (left), qg (center), and qq (right) resonances. The lower panels show the ratio of the limits for 2016 data divided by the ratio of the limits for 2015 data in the range of masses in which both analyses set limits.

This analysis and the one from 2015 provide consistent 95% CL limits, mainly differing by the available integrated luminosity in their data sets. The effects of systematic uncertainties on those limits are similar for both analyses.