CMS Physics Analysis Summary

Contact: cms-pag-conveners-susy@cern.ch 2016/06/13

Search for *R*-parity violating supersymmetry with displaced vertices

The CMS Collaboration

Abstract

A search for *R*-parity violating supersymmetry has been performed using proton-proton collision data collected by the CMS experiment at a center-of-mass energy of $\sqrt{s} = 8$ TeV. The data analyzed correspond to an integrated luminosity of 17.6 fb^{-1}. This search assumes a minimal flavor violating model where the lightest supersymmetric particle is a long-lived neutralino or gluino, leading to a signal with jets emanating from displaced vertices. Based on a sample of events with two displaced vertices, the pair production cross section is bounded as a function of mass and lifetime of the neutralino or gluino. For a mass of 400 GeV and mean proper decay length of 10 mm, the analysis excludes cross sections above 0.6 fb at 95% confidence level.

1 Introduction

In spite of extensive efforts by the ATLAS and CMS Collaborations at the CERN LHC, no evidence has been found for the superpartners of standard model (SM) particles predicted by supersymmetry (SUSY) [1, 2]. The majority of searches to date have assumed the conservation of R-parity [3]; if superpartners are produced and R-parity is conserved, the lightest supersymmetric particle (LSP) leaves the detector unnoticed aside from a large amount of missing transverse energy. R-parity conservation is not strictly necessary to explain experimental bounds such as the limits on proton lifetime, as long as at least one of lepton or baryon number is conserved, or if the associated R-parity violating (RPV) [4] terms in the Lagrangian are extremely small. However, no obvious mechanism suppresses the RPV terms.

In minimal flavor violating (MFV) models of RPV SUSY [5, 6], the Yukawa couplings between superpartners and SM particles are the sole source of flavor symmetry violation. The amplitudes for lepton- and baryon-number changing interactions are then small. At the LHC, the LSP typically decays within the detector volume, so there is no large missing transverse energy. The production processes of the superpartners are still nearly the same as in the minimal supersymmetric standard model, e.g., superpartners are produced in pairs, but the phenomenology depends on the identity of the LSP.

This analysis uses as a benchmark a model in Ref. [6] that assumes that the LSP is a neutralino ($\tilde{\chi}^0$) or gluino ($\tilde{g}$) heavy enough to decay into a top antiquark and virtual top squark ($\tilde{t}$) as depicted in Fig. 1, which is the dominant decay mode. The top squark then decays to strange and bottom antiquarks, with other partial widths smaller by a factor of 100 or more. This results in the LSP being long-lived due to the Yukawa-suppressed coupling of the top squark to the bottom and strange antiquarks, with the lifetime depending on the model parameters. For large parts of the parameter space, pair-produced LSPs lead to interesting signals. These include increased top quark production rates; events with many jets, especially b quark jets; and displaced vertices.

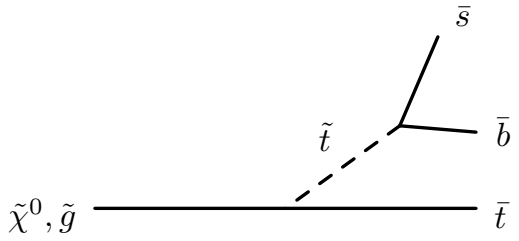

Figure 1: This search uses as a benchmark a model in which neutralino or gluino LSPs are pair-produced, and then decay as shown.

The decay of the LSP presents a distinct signature of multiple jets emerging from a displaced vertex, often with wide opening angles. In this analysis, a custom vertex reconstruction algorithm is used to identify these vertices. Standard algorithms to identify b quark jets [7] assume a single jet whose momentum is aligned with the vertex displacement from the primary vertex. By using a custom algorithm, the vertex reconstruction can be optimized for the distinctive features of this signal. The signature is two of these vertices, well separated in space. Monte Carlo (MC) simulation shows that SM background events rarely have even one reconstructed displaced vertex; in the even rarer events with two displaced vertices, the vertices are usually not well separated.

Another search by the CMS Collaboration has looked for pairs of displaced jets from a single vertex [8]. This analysis searches for a pair of displaced vertices in conjunction with jets. It extends to shorter lifetimes than a similar analysis reported by the ATLAS Collaboration [9]. A search for the same final state with the LSP decaying promptly, using the distributions of the number of jets and the number of b-tagged jets, was carried out by the CMS Collaboration [10].

This analysis applies not only to the MFV model described here, but more generally to models for physics beyond the SM with long-lived particles decaying to multiple jets. In addition to the results of the search with a neutralino or gluino LSP, we present a method for reinterpretation of the analysis.

2 The CMS detector

The central feature of the CMS detector is a superconducting solenoid providing a magnetic field of 3.8 T aligned with the proton beam direction. Contained within the field volume of the solenoid are a silicon pixel and strip tracker, a lead tungstate electromagnetic calorimeter (ECAL), and a brass and scintillator hadronic calorimeter (HCAL). Outside the solenoid is the steel magnetic return yoke, interspersed with muon tracking chambers. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [11].

The silicon tracker, which is particularly relevant to this analysis, measures the tracks of charged particles in the range of pseudorapidity, η, up to $|\eta| < 2.5$. For nonisolated particles with transverse momentum, p_{T}, of 1 to 10 GeV and $|\eta| < 1.4$, the track resolutions are typically 1.5% in p_{T}, 25–90 μm in the impact parameter in the transverse plane, and 45–150 μm in the impact parameter in the longitudinal direction [12]. When combining information from the entire detector, the jet energy resolution amounts typically to 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV, to be compared to about 40%, 12%, and 5% obtained when the ECAL and HCAL calorimeters alone are used [13].

The first level (L1) of the CMS trigger system, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select the most interesting events in a fixed time interval of less than 4 μs. The high-level trigger (HLT) processor farm further decreases the event rate from around 100 kHz to less than 1 kHz, before data storage.

3 Event samples

The data used in this analysis correspond to an integrated luminosity of 17.6 fb^{-1}, collected in proton-proton (pp) collisions at a center-of-mass energy of $\sqrt{s} = 8$ TeV in 2012. Events are selected using a trigger requiring the presence of at least four jets reconstructed from energy deposits in the calorimeters. At the first trigger level L1, the jets are required to have transverse momentum $p_{\mathrm{T}} > 40$ GeV, while in the HLT the threshold is $p_{\mathrm{T}} > 50$ GeV. The latter threshold is afforded by a special data-taking strategy called "data parking" [14], in which the triggered events were saved but not promptly reconstructed, allowing a higher event rate. The data included in this analysis represent the fraction of the 2012 LHC operation for which this strategy was implemented.

Simulated events are used to model both the signal and background processes. Using PYTHIA 8.165 [15], signal samples with varying neutralino masses (200–1500 GeV) and lifetimes τ ($c\tau =$ 0.1–30 mm) were produced. In these samples, neutralinos are produced in pairs, and each neutralino is forced to undergo a three-body decay into top, bottom, and strange (anti-)quarks.

Backgrounds arising from SM processes are dominated by multijet and top quark pair ($t\bar{t}$) events. The multijet processes include b quark pair events. Smaller contributions come from single top quark production (single t), vector boson production in association with additional jets (V+jets), diboson production (VV), and top quark pairs with a radiated vector boson ($t\bar{t}$ +V). Processes with a single vector boson include virtual photons, W bosons, or Z bosons, while the diboson processes include WW, WZ, and ZZ. Single top events were simulated with POWHEG 1.0 [16–20]; diboson events were simulated with PYTHIA 6.426 [21]; all other backgrounds were simulated using MADGRAPH 5.1 [22]. For all samples, hadronization and showering were done using PYTHIA 6.426 with tune Z2*. The Z2* tune is derived from the Z1 tune [23], which uses the CTEQ5L parton distribution set, whereas Z2* adopts CTEQ6L [24]. The detector response, for all simulated samples, was modeled using a GEANT4-based simulation [25] of the CMS detector. The effects of additional pp interactions per bunch crossing ("pileup") were included by overlaying additional simulated minimum-bias events, reweighting events such that the distribution of the number of interactions matches that in the data.

4 Event preselection

To ensure that the four-jet trigger efficiency is high and well understood, more stringent jet requirements are applied offline, requiring at least four jets in the calorimeter with $p_T > 60\,\text{GeV}$. These jets are reconstructed from the energy deposits in the calorimeter, clustered by the anti-k_t algorithm [26, 27] with a distance parameter of 0.5. The trigger efficiency, determined using events satisfying a single-muon trigger, is $(96.2 \pm 0.2)\%$ for events with four offline jets with $p_T > 60\,\text{GeV}$. The simulation overestimates this efficiency by a factor of 1.022 ± 0.002, so, where used, its normalization is corrected by this amount.

Jets considered in the rest of the analysis are those obtained in the full event reconstruction, performed using a particle-flow (PF) algorithm [28, 29]. The PF algorithm reconstructs and identifies photons, electrons, muons, and charged and neutral hadrons with an optimized combination of information from the various elements of the CMS detector. Before clustering the PF candidates into jets, charged PF candidates are excluded if they originate from a pp interaction vertex other than the primary vertex, which is the one with the largest scalar $\Sigma|p_T|^2$. The resulting particles are clustered into jets by the anti-k_t algorithm with a distance parameter of 0.5. Jets used in the analysis must satisfy $p_T > 20\,\text{GeV}$ and $|\eta| < 2.5$.

For an event to be selected for further analysis, the scalar sum of the p_T of jets in the event, H_T, is required to be at least 500 GeV. This requirement has little impact on signal events, but is useful for suppressing SM background.

5 Vertex reconstruction, variables, and selection

5.1 Vertex reconstruction

Displaced vertices are reconstructed from tracks in the CMS silicon tracker. These tracks are required to have $p_T > 1\,\text{GeV}$; at least eight hits in the tracker, at least one of which is in the pixel detector; and a transverse impact parameter with respect to the beam axis of at least 100 μm. The impact parameter requirement favors vertices that are displaced from the primary vertex. The vertex reconstruction algorithm starts by forming seed vertices from all pairs of tracks that satisfy these requirements. Each vertex is fitted with the Kalman filter approach [30], and a fit is considered successful if it has χ^2 per degree of freedom (χ^2/dof) less than 5. The vertices are then merged iteratively until no pair of vertices shares tracks: for each pair of vertices that

shares one or more tracks, if the three-dimensional (3D) distance between the vertices is less than four times the uncertainty in that distance, a vertex is fit to the tracks from both and they are replaced by the merged vertex if the fit has $\chi^2/\text{dof} < 5$. Otherwise, each track is assigned to one vertex or the other depending on its 3D impact parameter significance with respect to each of the vertices, as follows:

- if the track is consistent with both vertices (both impact parameters less than 1.5 standard deviations), assign it to the vertex that has more tracks already;
- if the track's impact parameter is greater than 5 standard deviations from either vertex, drop it from that vertex;
- otherwise, assign the track to the vertex to which it has a smaller impact parameter significance.

Each remaining vertex is then refit, and if the fit satisfies the requirement of $\chi^2/\text{dof} < 5$, the old vertex is replaced with the new one; otherwise it is dropped entirely.

This algorithm is similar in many regards to those used to identify ("tag") b quark jets [7]. Typical b tagging algorithms, however, are optimized for identifying the decay in flight of a particle into a single jet, and consequently, make requirements that degrade sensitivity to the multijet final states sought here. For example, b tagging algorithms generally require that the tracks assigned to a vertex are approximately aligned with the flight direction from the primary vertex to the decay point, which is inefficient when there are multiple jets in the final state, including some that may be directed at large angles with respect to the flight path. The b tagging algorithms also discard tracks with impact parameters beyond those typical for b quark daughters (>2 mm), thereby significantly reducing efficiency for finding vertices with large displacements.

5.2 Vertex variables and selection

The vertexing procedure produces multiple vertices per event, only some of which are consistent with the signal. In order to select quality vertices, we impose additional requirements on the vertex and its associated tracks and jets. The requirements for each vertex are:

- at least five tracks;
- at least three tracks with $p_{\mathrm{T}} > 3\,\text{GeV}$;
- at least one pair of tracks with separation $\Delta R < 0.4$ to favor vertices that include multiple tracks from a single jet, where $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$;
- at least one pair of tracks with $\Delta R > 1.2$ to favor vertices involving multiple jets;
- $\Delta R < 4$ for all pairs of tracks, to suppress wide-angle track coincidences;
- at least one jet that shares one or more tracks with the vertex;
- displacement in x-y of the vertex from the detector origin of less than 25 mm, to suppress vertices from interactions in the beam pipe or detector material;
- uncertainty in the x-y distance of the vertex from the beam axis of less than 25 μm.

In the observed data, 181,076 events have one vertex satisfying the above requirements, 251 have two of them, and no events have more than two. The two-vertex events comprise the candidate sample.

5.3 Signal region

The signal is extracted from the two-vertex events using the spatial separation between the vertices. The x-y distance of each vertex from the beam axis is defined as d_{BV}. Figure 2 shows the distribution of d_{BV} for the control region defined by events with exactly one vertex. The data and simulated background are in good agreement.

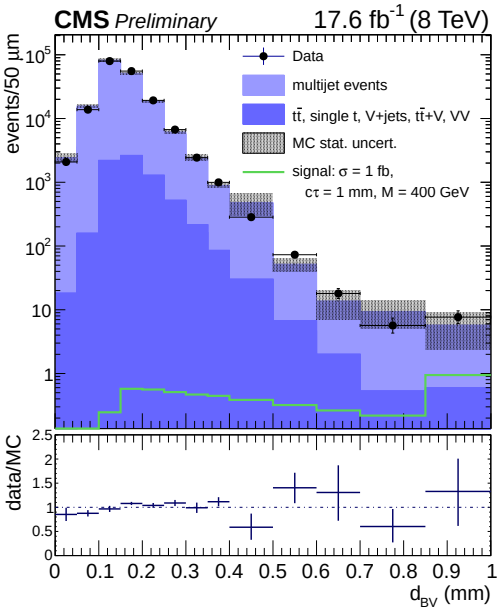

Figure 2: In events with exactly one vertex, the x-y distance from the beam axis to the vertex, d_{BV}, for data, simulated background normalized to data, and a simulated signal with $c\tau = 1\,\mathrm{mm}$, $M = 400\,\mathrm{GeV}$, and production cross section 1 fb. Event preselection and vertex selection criteria have been applied. The last bin includes the overflow events. These events are used to construct the two-vertex background template as explained in the text.

In signal events, the two LSPs are emitted approximately back-to-back, leading to large separation between the vertices. We define the distance between the two vertices in the x-y plane as d_{VV}, and fit this distribution to extract the signal. The fit to the d_{VV} distribution in data is described in Section 8. In signal simulation, fewer than 10% of events in the candidate sample have more than two selected vertices. For these events, the two vertices with the highest number of tracks are selected for the d_{VV} calculation, and in the case where two vertices have the same number of tracks, the vertex with the higher mass is chosen. The vertex mass is reconstructed using the momenta of the associated tracks, assuming that the particles associated to the tracks have the charged pion mass. Figure 3 shows the d_{VV} distribution of the simulated background and an example simulated signal with $c\tau = 1\,\mathrm{mm}$, $M = 400\,\mathrm{GeV}$, and production cross section 1 fb. The bins in d_{VV} are chosen to be sensitive to the peaking nature of the background at low d_{VV}; five 200 μm bins are used from 0 to 1 mm, then one bin from 1 to 50 mm where the long-lived signal dominates.

The signal d_{VV} templates are taken directly from simulation, with a distinct template for each LSP mass and lifetime. The background template is constructed from data as described in the following section. The signal region is defined by $d_{\mathrm{VV}} > 600\,\mu\mathrm{m}$. Figure 4 shows the signal efficiency in the signal region as a function of LSP mass and lifetime. The signal efficiency generally increases as lifetime increases, until more vertices appear beyond our fiducial limit

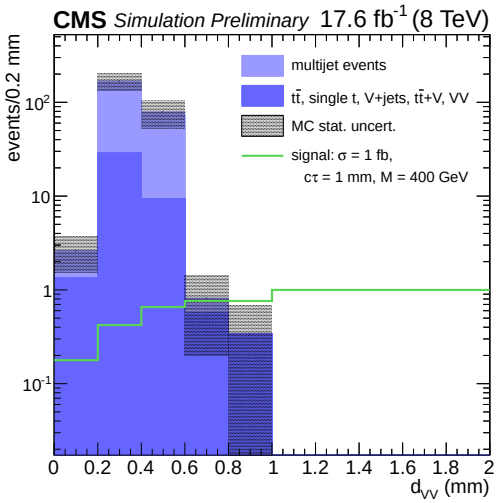

Figure 3: The x-y distance between vertices, d_{VV}, for simulated background normalized to data, and a simulated signal with $c\tau = 1\,\mathrm{mm}$, $M = 400\,\mathrm{GeV}$, and production cross section 1 fb. All vertex and event selection criteria have been applied. The last bin includes the overflow events.

at the beam pipe. The efficiency also generally increases as mass increases, until about 800 GeV where it begins to slightly decrease because of the event quality selection imposed, particularly the requirement on the maximum opening angle between tracks in a vertex.

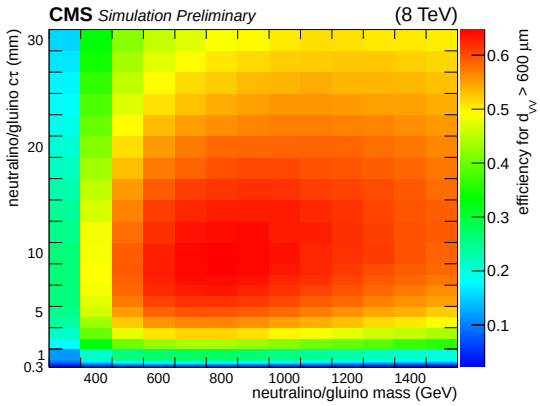

Figure 4: Signal efficiency as a function of neutralino/gluino mass and lifetime, requiring two quality vertices and $d_{\mathrm{VV}} > 600\,\mu\mathrm{m}$.

6 Background template

The background in the two-vertex sample comes primarily from multijet events and $\mathrm{t\bar{t}}$ production. In simulation, these account for approximately 85% and 15% of events, respectively. Other sources of background, such as V+jets and single t events, are negligible. Approximately half of the events include one or more b quark jets, as their displaced decay daughters can seed a vertex when combined with other, often misreconstructed, tracks.

The background d_{VV} template, denoted by $d_{\mathrm{VV}}^{\mathrm{C}}$, is constructed using the 181,076 events in the data that have only a single vertex. Each value of $d_{\mathrm{VV}}^{\mathrm{C}}$ is calculated as the distance between two

random vertices, which each have a random value of d_{BV} and a random value of ϕ_V. The two values of d_{BV} are drawn from the one-vertex d_{BV} distribution. The two values of ϕ_V are chosen using information about the jet directions in a one-vertex event. To choose each angle, we select a jet at random with p_T-weighted probability, and draw the angle from a Gaussian distribution with width 0.4 radians, centered on a direction perpendicular to the jet in the transverse plane. Background vertices tend to be displaced perpendicular to jet momenta in the x-y plane because they arise from poorly measured tracks in jets.

The vertex reconstruction algorithm merges neighboring vertices. To emulate this behavior in our background template construction procedure, we discard pairs of vertices that are not sufficiently separated. We keep pairs of vertices with a probability parameterized by a Gaussian error function with mean μ_{clear} and width σ_{clear}. The values of μ_{clear} and σ_{clear}, which are related to the position uncertainties of the tracks, are varied in the fit to the d_{VV} distribution in data. The values found in the fit are $\mu_{clear} = 320\,\mu$m and $\sigma_{clear} = 110\,\mu$m.

In simulated events, the agreement between the constructed d_{VV}^{C} template and nominal d_{VV} distribution is well within the statistical uncertainty. When normalized to the data, the difference in their yields in the signal region ($d_{VV} > 600\,\mu$m) is 0.6 ± 2.6 events. Figure 5 shows a comparison of the d_{VV}^{C} and d_{VV} distributions in simulated events, showing the variation in d_{VV}^{C} for the range of values of μ_{clear} and σ_{clear} that are within one standard deviation of the fit values.

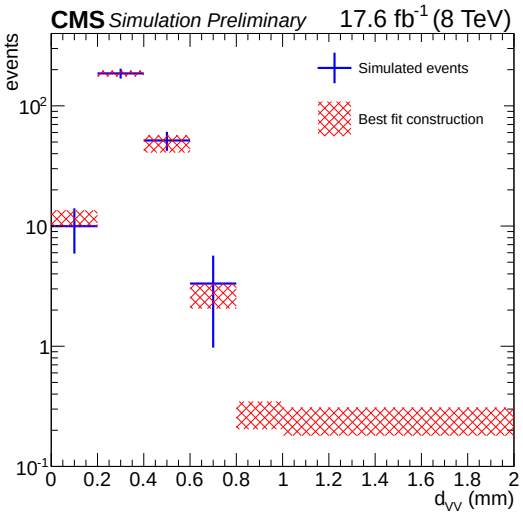

Figure 5: The x-y distance between vertices, d_{VV}, for simulated background events (blue crosses), compared with the best fit construction d_{VV}^{C} (red crosshatches). The distributions are normalized to the number of two-vertex events in the data. The error bars for the simulated events represent only the statistical uncertainty, while the shaded region for the constructed distribution is the result of varying μ_{clear} and σ_{clear} within one standard deviation of the values from the fit.

7 Systematic uncertainties

The signal is extracted from a fit of a weighted sum of the signal and background templates to the d_{VV} distribution in data. For the signal, the simulation provides both the d_{VV} distribution and its normalization, and systematic uncertainties arise from sources such as vertex reconstruction efficiency, track reconstruction, track multiplicity, pileup conditions, the detector alignment, and the jet energies. For the background, for which the template is derived

from a control sample in data, the systematic uncertainties come from effects that could cause a discrepancy between the constructed d_{VV}^{C} distribution and the nominal d_{VV} distribution.

7.1 Systematic uncertainties related to signal distribution and efficiency

The dominant systematic uncertainty in the signal normalization arises from the difference between the vertexing efficiencies in the simulation and data. This is evaluated in an independent study in which artificial signal-like vertices are produced by displacing tracks in background events by a known displacement vector, and then vertex reconstruction is performed. The magnitude of the displacement vector is sampled from an exponential with scale parameter 1 mm, restricted to values between 0.3 and 25 mm, similar to the expected distribution of signal vertices. The direction is calculated from the momentum of the jets in the event, but is smeared to emulate the difference between the flight and momentum directions in simulated signal events due to track-level inefficiency and unaccounted neutral particles. Events are required to satisfy the preselection requirements described in Section 4, and the displaced jets satisfy $p_T > 50\,\text{GeV}$ and $\Delta R < 4$ for all pairs. We find that in events in which the tracks from two light parton jets are displaced, a quality vertex, with the requirements described in Section 5.2, is reconstructed within 50 μm of the correct location in 69.8% of simulated events and in 64.0% of data events, with negligible statistical uncertainties. The former efficiency is comparable to that of LSP decay vertices in simulated signal events. The ratio of efficiencies between data and simulation gives an 8.4% uncertainty per vertex. Agreement improves for larger numbers of displaced light parton or b quark jets, and is independent of the size of the displacement. For two-vertex events, the uncertainty is 17%.

Additional studies explore the sensitivity of other effects that could alter the signal template. The vertex clustering depends on the number of charged particles in the event. This number of charged particles can vary based on the model of the underlying event, as represented by the chosen tune of PYTHIA, e.g. those described in Ref. [31]. The resulting signal templates differ by no more than 1% in any bin and the overall efficiency changes by no more than 3%. This 3% is taken as a systematic uncertainty.

To test sensitivity to a possible misalignment, the signal samples have been reconstructed using several tracker misalignment scenarios corresponding to various "weak modes": coherent distortions of the tracker geometry left over by the alignment procedure that lead to a systematic bias in the track parameters for no penalty in χ^2 of the overall alignment fit [32]. These misalignments change the overall efficiency by no more than 2%. This 2% is taken as a systematic uncertainty.

To study sensitivity to the pileup distribution, we vary the inelastic pp cross section used in the pileup weighting by $\pm 5\%$ [33]. This variation is found to have an effect of $<1\%$ on the signal efficiency.

The uncertainty in the jet energy scale affects the total energy measured, and could, for example, enable an event with low jet activity to pass the jet p_T or H_T selections. This effect is studied by varying the jet energy scale and resolution [13], and is found to change the signal efficiency by $<1\%$. A 2.6% uncertainty [34] is associated with the integrated luminosity for the 2012 dataset and thus in the derived signal cross section. The uncertainty in the trigger efficiency is $<1\%$.

Table 1 summarizes the systematic uncertainties in the signal efficiency. We assume there are no correlations among them, so we add them in quadrature to obtain the overall uncertainty.

Table 1: Summary of systematic uncertainties in the signal efficiency.

Systematic effect	Uncertainty
Vertex reconstruction	17%
Underlying event	3%
Tracker misalignment	2%
Pileup	<1%
Jet energy scale/resolution	<1%
Integrated luminosity	2.6%
Trigger efficiency	<1%
Overall	18%

7.2 Systematic uncertainties related to background estimate

The d_{VV}^{C} background template is constructed from a large sample of events in data with a single vertex. Systematic uncertainties in the d_{VV}^{C} template are estimated by varying the d_{VV}^{C} construction method and taking the difference between the d_{VV}^{C} distributions using the default and alternate methods. The method for constructing d_{VV}^{C} involves drawing two values of d_{BV} and two values of ϕ_V, with an angle between vertices $\Delta\phi_{VV}$, so the main uncertainties come from effects related to the d_{BV} and $\Delta\phi_{VV}$ distributions.

The production of b quarks in pairs introduces a correlation between the vertex distances in two-vertex events that is not accounted for when single vertices are paired at random. In simulation, events without b quarks have a mean d_{BV} of $\sim$160 μm, while events with b quarks, which account for 15% of one-vertex events, have a mean d_{BV} of $\sim$190 μm. We quantify this effect by sorting the simulated background events into those with and without b quarks, constructing the d_{VV}^{C} distributions for each, and then combining them in the proportions 45:55, which is the ratio of b-quark to non-b-quark events in two-vertex background events determined from simulation. The systematic uncertainty is taken to be the difference between the yields obtained with this procedure and the standard one. Both procedures use simulated events, scaled to the two-vertex yield in data.

The d_{VV}^{C} construction method discards pairs of vertices that would overlap, consistently leading to a two-vertex angular distribution that peaks at $\pm\pi$ radians. To assess the systematic uncertainty related to assumptions about the angular distribution between vertices, we draw $\Delta\phi_{VV}$ from the angular distribution between vertices in simulated two-vertex background events. This leads to a d_{VV}^{C} distribution with a more strongly peaked $\Delta\phi_{VV}$ distribution, and provides a conservative estimate of the uncertainty.

The statistical uncertainty from the limited number of one-vertex events in the data that are used to construct the two-vertex distribution is studied using a resampling method. Using the d_{BV} distribution as the parent, we randomly sample ten new d_{BV} pseudo-datasets, and use each to construct a d_{VV}^{C} distribution. The root-mean-square variation in bin-by-bin yields in the set of distributions gives the statistical uncertainty.

There is a small contribution to the uncertainty in the prediction of d_{VV}^{C} due to the binning of the d_{BV} parent distribution; moving the d_{BV} tail bin edges around by an amount compatible with the vertex position resolution, 20 μm, varies the prediction in d_{VV}^{C} appreciably only in the last two bins: by 0.06 events in the 0.8–1.0 mm bin, and by 0.09 events in the 1.0–50 mm bin.

The results of these four studies are summarized in Table 2. In assessing the overall systematic uncertainty in the background template, we add in quadrature the values and their uncertain-

ties, assuming no correlations.

Table 2: Uncertainties in the predicted event yield in each d_{VV}^{C} bin, arising from the correlations between vertex distances, the modeling of $\Delta\phi_{VV}$, the d_{BV} sample size, and the choice of d_{BV} binning. In the first two cases, shifts are given with their statistical uncertainty. The last row gives the overall systematic uncertainties, assuming no correlations. All yields are normalized to the total number of two-vertex events in the data.

Systematic effect	0.0–0.2 mm	0.2–0.4 mm	0.4–0.6 mm	0.6–0.8 mm	0.8–1.0 mm	1.0–50 mm
d_{BV} correlations	-0.65 ± 0.05	-3.60 ± 1.01	3.59 ± 0.76	0.63 ± 0.18	0.01 ± 0.07	0.01 ± 0.04
$\Delta\phi_{VV}$ variation	0.74 ± 0.01	1.00 ± 0.00	1.04 ± 0.01	1.18 ± 0.07	-0.01 ± 0.06	-0.01 ± 0.04
d_{BV} sample size	0.05	0.54	0.51	0.17	0.04	0.07
d_{BV} binning	-	-	-	-	0.06	0.09
Total events	1.0	3.9	3.8	1.4	0.1	0.1

In principle, the background template can be affected by the underlying event, tracker alignment, and pileup conditions as described in Section 7.1. To assess the impact of the underlying event, we generated 5 million all-hadronic $t\bar{t}$ events varying the PYTHIA tune, but observed no change in d_{VV}^{C} larger than 1%. Likewise, we reconstructed samples with tracker misalignments, also with negligible effect. Varying the inelastic pp cross section used in pileup weighting by $\pm 5\%$ and varying the number of pileup interactions also reveals no significant trends. None of these effects contributes significantly to the systematic uncertainty. Since the normalization of the template is a free parameter of the fit, uncertainties such as those in the integrated luminosity, trigger efficiency, and vertex reconstruction efficiency do not enter.

8 Fitting, signal extraction, and statistical interpretation

The distribution of d_{VV}, the separation between vertices in the x-y plane for two-vertex events, is used to discriminate signal from background, with the signal templates taken directly from the MC simulation and the background template constructed from the one-vertex event sample in data. In the following sections, we describe the fitting and statistical procedures used for the search.

8.1 Fitting procedure

To estimate the number of signal and background events present in the observed data, a binned shape fit is performed using an extended maximum likelihood method. Initially neglecting terms arising from uncertainty in the templates, the likelihood function is given by

$$-2\log\mathcal{L}(\mathbf{n}|s,b,\boldsymbol{\nu}) = \sum_i \left[n_i \log a_i(s,b,\boldsymbol{\nu}) - a_i(s,b,\boldsymbol{\nu})\right], \tag{1}$$

where n_i is the number of observed events in bin i, s and b are the normalizations of the signal and background templates corresponding to the yields, $\boldsymbol{\nu}$ denotes the shape parameters μ_{clear} and σ_{clear} used in the background template construction procedure as described in Section 6, and

$$a_i(s,b,\boldsymbol{\nu}) = s a_i^{(s)} + b a_i^{(b)}(\boldsymbol{\nu}) \tag{2}$$

is the sum of the signal and background shapes $a^{(s)}$ and $a^{(b)}$ in bin i.

The only assumed shape uncertainty in the signal templates is that due to the finite MC statistics; the uncertainty is as high as 20% for the lowest lifetime and mass samples, but is generally no more than 1% in any bin for the majority of the templates. For the background templates, a Gaussian uncertainty is assumed in the value of the template in each bin, truncated at zero. To incorporate these uncertainties in the signal and background templates, a procedure similar to that of Barlow and Beeston [35] is followed, modified to allow a bin-by-bin Gaussian uncertainty in the background shape [36]. The final likelihood function is then given by

$$\begin{aligned} -2\log\mathcal{L}(\mathbf{n}|s,b,\boldsymbol{\nu},A_i^{(s)},A_i^{(b)}) = &\sum_i n_i \log A_i - A_i \\ &+ \sum_i Ma_i^{(s)} \log MA_i^{(s)} - MA_i^{(s)} \\ &+ \sum_i -\frac{1}{2}\left(\frac{a_i^{(b)} - A_i^{(b)}}{\sigma_i^{(b)}}\right)^2, \end{aligned} \tag{3}$$

with $A_i = sA_i^{(s)} + bA_i^{(b)}$. The $A_i^{(s)}$ and $A_i^{(b)}$ replace the $a_i^{(s)}$ and $a_i^{(b)}$ from above in the shape fit to the data, and are allowed to vary as either Poisson ($A_i^{(s)}$) or Gaussian ($A_i^{(b)}$) distributed parameters. The quantity M is the number of events from the MC signal sample that produced the $a_i^{(s)}$ estimates, and $\sigma_i^{(b)}$ are the widths of the Gaussians, taken to be the relative sizes of the uncertainties listed in Table 2. The modified Barlow-Beeston procedure finds the $A_i^{(s)}$ and $A_i^{(b)}$ that maximize $\log\mathcal{L}$ given $(s,b,\boldsymbol{\nu})$; the difference here is that the $A_i^{(b)}$ are Gaussian distributed parameters.

The likelihood function is only weakly dependent on the background shape parameters $\boldsymbol{\nu}$, and when signal is injected, the best fit values $\hat{\boldsymbol{\nu}}$ agree well with the background-only values. The fit is well behaved: for most signal templates, in pseudo-experiments where the true signal and background strengths are known, the distribution of the fitted yields for s and b have means consistent with those input, and the widths of the distributions as measured by their root-mean-square are consistent with the uncertainties in the fits. For the signal templates with low lifetimes, however, the signal yield is biased downward when an injected signal is present. This is due to the background shape being allowed to vary up in the tail within the uncertainties assigned. When no injected signal is present, there is a larger bias toward obtaining $s > 0$ when fitting using templates with $c\tau < 300\,\mu$m. Therefore, we only consider signals with $c\tau \geq 300\,\mu$m in the fit and search.

8.2 Statistical analysis

The test statistic q used to quantify any excess of signal events over the expected background is given by a profile likelihood ratio [37]:

$$q = \log\frac{\max_{s\geq 0, b\geq 0}\mathcal{L}(\mathbf{n}|s,b,\hat{\boldsymbol{\nu}},\hat{A}_i^{(s)},\hat{A}_i^{(b)})}{\max_{b\geq 0}\mathcal{L}(\mathbf{n}|s=0,b,\hat{\boldsymbol{\nu}},\hat{A}_i^{(s)},\hat{A}_i^{(b)})}, \tag{4}$$

where for each value of s and b the nuisance parameters $\hat{A}_i^{(s)}$, $\hat{A}_i^{(b)}$, and $\hat{\boldsymbol{\nu}}$ are found that maximize the relevant likelihood. The probability under the background-only hypothesis, p_0, to obtain a value of the test statistic at least as large as that in the data, q_{obs}, is estimated as the fraction of 10,000 pseudo-experiments with $q \geq q_{\text{obs}}$. This is referred to as the p-value for

a particular signal hypothesis. The pseudo-experiments are generated using the background $d_{\mathrm{VV}}^{\mathrm{C}}$ distribution corresponding to $\hat{\boldsymbol{\nu}}_b$, and background count b drawn from a Poisson distribution with mean equal to the number of events in the data n. The nuisance parameters, $\boldsymbol{\nu}$, $A_i^{(s)}$ and $A_i^{(b)}$, are drawn from their corresponding Poisson or Gaussian distributions in each pseudo-experiment.

We obtain limits on the signal yield, which can be converted into limits on the product of the cross section for neutralino or gluino pair production and the square of the branching fraction for decay via the channel under study, denoted by $\sigma\mathcal{B}^2$. To obtain limits on $\sigma\mathcal{B}^2$, for a given number of signal events s_0 we calculate the probability for the null hypothesis of $s = s_0$ versus the alternative that $s < s_0$, denoted by p_{s_0}. We do this in practice by generating 10,000 pseudo-experiments with s drawn from a Poisson distribution with mean s_0 and b drawn from a Poisson distribution with mean $n - s_0$. The background shape $d_{\mathrm{VV}}^{\mathrm{C}}$ is taken from the $\boldsymbol{\nu}$ from the original fit and signal shape corresponding to the signal hypothesis in question, with $A_i^{(b)}$ from their Gaussian distributions. p_{s_0} is then the fraction of pseudo-experiments where $q \geq q(s_0)$. We protect against downward fluctuations in the data by using the $\mathrm{CL_S}$ criterion [38, 39], defining the statistic as

$$\mathrm{CL_S} = \frac{p_{s_0}}{1 - p_0}, \tag{5}$$

The 95% confidence level (CL) upper limit on s is then the biggest s_0 for which $\mathrm{CL_S}$ is still greater than 0.05.

The limit on the signal yield is converted to a limit on $\sigma\mathcal{B}^2$ using the efficiencies calculated from simulation and the integrated luminosity of the data sample, 17.6 fb^{-1}. We include the effect of the estimated 18% signal efficiency uncertainty by varying the cross section in each pseudo-experiment by the value sampled from a log-normal density with location parameter 1 and scale parameter 0.18.

8.3 Results of the fit

The result of the fit to data is shown in Fig. 6, for the $c\tau = 1\,\mathrm{mm}$, $M = 400\,\mathrm{GeV}$ signal template. The observed data counts in each bin, along with the predictions from the background-only fit and the related uncertainties, are listed in Table 3. There is a small excess of events with $0.6 < d_{\mathrm{VV}} < 50\,\mathrm{mm}$: 7 in the data, while the background-only fit predicts 4.1 ± 1.4, where the uncertainty is the overall systematic uncertainty discussed in Section 7. In the signal+background fits, a typical value for the signal yield is 1.7 ± 1.9, obtained with the $c\tau = 1\,\mathrm{mm}$, $M = 400\,\mathrm{GeV}$ signal hypothesis. The associated p-value obtained from pseudo-experiments is in the range 0.05–0.14 for signals with $0.3 \leq c\tau \leq 30\,\mathrm{mm}$, with the larger p-values coming from those with longer lifetimes.

8.4 Upper limits on signal cross section

Figure 7 shows the observed 95% CL upper limits on $\sigma\mathcal{B}^2$. As an example, for a neutralino with mass of 400 GeV and $c\tau$ of 10 mm, the observed 95% CL upper limit on $\sigma\mathcal{B}^2$ is 0.6 fb.

Exclusion curves are overlaid, assuming the gluino pair production cross section [40–44]. In the context of the MFV model that we are studying, the LSP being a neutralino or gluino can result in the same final state.

The scan in $c\tau$ is in steps of 100 μm from 300 μm to 1 mm, then in 1 mm steps up to 10 mm, and in 2 mm steps to 30 mm; the mass points are spaced by 100 GeV. The exclusion curves

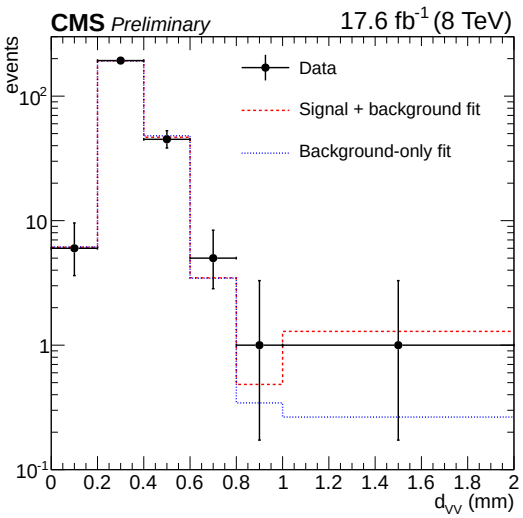

Figure 6: Fit to the data using the $c\tau = 1\,\text{mm}$, $M = 400\,\text{GeV}$ signal hypothesis, with the background-only and signal+background fits shown with blue dotted and red dashed lines, respectively. The last bin contains events with $1 < d_{\text{VV}} < 50\,\text{mm}$, but is shortened for visualization.

Table 3: Observed and mean expected background-only counts in each bin. The uncertainty is the sum in quadrature of the statistical and systematic uncertainties.

Bin i	d_{VV} range	Observed n_i	Mean expected count
1	0.0–0.2 mm	6	6.2 ± 1.0
2	0.2–0.4 mm	193	192.2 ± 3.9
3	0.4–0.6 mm	45	48.0 ± 3.8
4	0.6–0.8 mm	5	3.5 ± 1.4
5	0.8–1.0 mm	1	0.3 ± 0.1
6	1.0–50 mm	1	0.3 ± 0.1

are produced by linear interpolation of the limit scan, finding the set of points for which the interpolated upper limit is less than the gluino pair production cross section (the neutralino pair production cross section is expected to be much smaller).

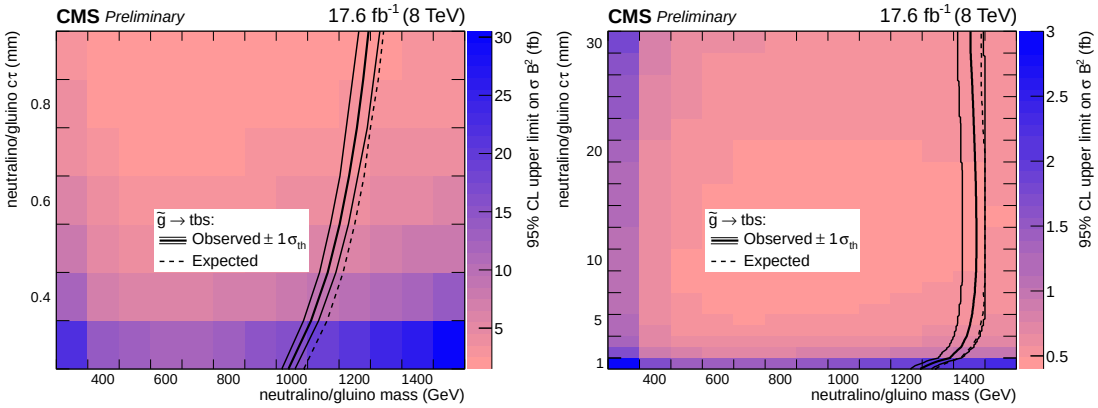

Figure 7: Observed 95% CL upper limits on cross section times branching fraction, with overlaid curves assuming gluino pair production cross section, for both observed (solid), with ± 1 standard deviation theoretical uncertainties, and expected (dashed) limits. The search excludes masses to the left of the curve. The left plot spans $c\tau$ from 300 through 900 μm, while the right plot ranges from 1 to 30 mm.

9 Extending the search to other signal models

The search for displaced vertices applies to other types of long-lived particles decaying to multiple jets. Here we present a generator-level selection that can be used to reinterpret the results of our analysis. For signal models in which there are two well-separated displaced vertices, this generator-level selection approximately replicates the reconstruction-level efficiency. The selection is based on the displacements of the long-lived particles, and the momenta and angular distributions of their daughter particles, which are taken to be u, d, s, c, and b quarks; electrons; and muons. The daughter particles are said to be "accepted" if they satisfy $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.5$, and "displaced" if their transverse impact parameter with respect to the origin is at least 100 μm. The criteria of the generator-level selection are:

- at least four accepted quarks with $p_{\mathrm{T}} > 60\,\mathrm{GeV}$;
- H_{T} of accepted quarks $> 500\,\mathrm{GeV}$;
- for each vertex:
  - x-y distance from beam axis $< 25\,\mathrm{mm}$;
  - at least one pair of accepted displaced daughter particles with $\Delta R > 1.2$;
  - $\Delta R < 4$ for all pairs of accepted displaced daughter particles;
  - at least one accepted displaced daughter quark;
  - $\sum p_{\mathrm{T}}$ of accepted displaced daughter particles $> 200\,\mathrm{GeV}$;
- x-y distance between vertices $> 600\,\mu$m.

In the region with $d_{\mathrm{VV}} > 600\,\mu$m, the background level is well determined and insensitive to fit parameters. Use of this generator-level selection replicates the reconstruction-level efficiency with an accuracy of 20% or better for a selection of models for which the signal efficiency is high (>10%). The selection may underestimate the trigger efficiency because it does not take into

account effects such as initial- and final-state radiation, and may overestimate the efficiency for reconstructing vertices with b quark secondaries, since the b quark lifetime can impede the assocation of their decay products with the reconstructed vertices.

10 Conclusions

A search has been performed using data corresponding to an integrated luminosity of 17.6 fb^{-1} collected with the CMS detector at $\sqrt{s} = 8$ TeV in 2012. The data were collected with a trigger requiring the presence of at least four jets. No evidence has been found for events in which pairs of massive particles with intermediate lifetimes decay into multijet final states. At 95% confidence level, the data exclude $\sigma\mathcal{B}^2$ above approximately 1 fb for pairs of particles with masses between 400 and 1500 GeV and $c\tau$ between 1 and 30 mm. While the search specifically addresses the production and decay of neutralino or gluino LSPs in R-parity violating SUSY, the results are relevant to other massive particles that decay to two or more jets. These are the best bounds to date on the production and decay of pairs of such massive particles with intermediate lifetimes.

References

[1] H. P. Nilles, "Supersymmetry, supergravity and particle physics", *Phys. Rept.* **110** (1984) 1, doi:10.1016/0370-1573(84)90008-5.

[2] H. E. Haber and G. L. Kane, "The search for supersymmetry: probing physics beyond the standard model", *Phys. Rept.* **117** (1985) 75, doi:10.1016/0370-1573(85)90051-1.

[3] G. R. Farrar and P. Fayet, "Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry", *Phys. Lett. B* **76** (1978) 575, doi:10.1016/0370-2693(78)90858-4.

[4] R. Barbier et al., "R-parity violating supersymmetry", *Phys. Rept.* **420** (2005) 1, doi:10.1016/j.physrep.2005.08.006, arXiv:hep-ph/0406039.

[5] E. Nikolidakis and C. Smith, "Minimal flavor violation, seesaw mechanism, and R-parity", *Phys. Rev. D* **77** (2008) 015021, doi:10.1103/PhysRevD.77.015021, arXiv:0710.3129.

[6] C. Csáki, Y. Grossman, and B. Heidenreich, "Minimal flavor violation supersymmetry: a natural theory for R-parity violation", *Phys. Rev. D* **85** (2012) 095009, doi:10.1103/PhysRevD.85.095009, arXiv:1111.1239.

[7] CMS Collaboration, "Identification of b quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[8] CMS Collaboration, "Search for long-lived neutral particles decaying to quark-antiquark pairs in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev. D* **91** (2015) 012007, doi:10.1103/PhysRevD.91.012007, arXiv:1411.6530.

[9] ATLAS Collaboration, "Search for massive, long-lived particles using multitrack displaced vertices or displaced lepton pairs in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Rev. D* **92** (2015) 072004, doi:10.1103/PhysRevD.92.072004, arXiv:1504.05162.

[10] CMS Collaboration, "Search for R-parity violating supersymmetry in proton-proton collisions at $\sqrt{s} = 8$ TeV in events with large jet and b-jet multiplicity", CMS Physics Analysis Summary CMS-PAS-SUS-14-003, 2015.

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

[12] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[13] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[14] CMS Collaboration, "Data parking and data scouting at the CMS experiment", CMS Performance Note CMS-DP-2012-022, 2012.

[15] T. Sjöstrand, S. Mrenna, and P. 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.

[16] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[17] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[18] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[19] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **09** (2009) 111, doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076.

[20] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[21] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[22] J. Alwall et al., "MadGraph 5: going beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.

[23] R. Field, "Early LHC underlying event data—findings and surprises", in *Hadron Collider Physics. Proceedings, 22nd Conference, HCP 2010, Toronto, Canada, August 23-27, 2010*. 2010. arXiv:1010.3558.

[24] 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.

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

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] R. Frühwirth, "Application of Kalman filtering to track and vertex fitting", *Nucl. Instrum. Meth. A* **262** (1987) 444, doi:10.1016/0168-9002(87)90887-4.

[31] 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.

[32] CMS Collaboration, "Alignment of the CMS tracker with LHC and cosmic ray data", *JINST* **9** (2014) P06009, doi:10.1088/1748-0221/9/06/P06009, arXiv:1403.2286.

[33] CMS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **722** (2013) 5, doi:10.1016/j.physletb.2013.03.024, arXiv:1210.6718.

[34] CMS Collaboration, "CMS luminosity based on pixel cluster counting—summer 2013 update", CMS Physics Analysis Summary CMS-PAS-LUM-13-001, 2013.

[35] R. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples", *Comput. Phys. Commun.* **77** (1993) 219, doi:10.1016/0010-4655(93)90005-W.

[36] J. S. Conway, "Incorporating nuisance parameters in likelihoods for multisource spectra", in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva, Switzerland 17-20 January 2011*. 2011. arXiv:1103.0354. doi:10.5170/CERN-2011-006.115.

[37] W. A. Rolke, A. M. López, and J. Conrad, "Limits and confidence intervals in the presence of nuisance parameters", *Nucl. Instrum. Meth. A* **551** (2005) 493, doi:10.1016/j.nima.2005.05.068, arXiv:physics/0403059.

[38] A. L. Read, "Presentation of search results: the $\mathrm{CL_S}$ technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[39] 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.

[40] W. Beenakker, R. Hopker, M. Spira, and P. M. Zerwas, "Squark and gluino production at hadron colliders", *Nucl. Phys. B* **492** (1997) 51, doi:10.1016/S0550-3213(97)80027-2, arXiv:hep-ph/9610490.

[41] A. Kulesza and L. Motyka, "Threshold resummation for squark-antisquark and gluino-pair production at the LHC", *Phys. Rev. Lett.* **102** (2009) 111802, doi:10.1103/PhysRevLett.102.111802, arXiv:0807.2405.

[42] A. Kulesza and L. Motyka, "Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC", *Phys. Rev. D* **D80** (2009) 095004, doi:10.1103/PhysRevD.80.095004, arXiv:0905.4749.

[43] W. Beenakker et al., "Soft-gluon resummation for squark and gluino hadroproduction", *JHEP* **12** (2009) 041, doi:10.1088/1126-6708/2009/12/041, arXiv:0909.4418.

[44] W. Beenakker et al., "Squark and gluino hadroproduction", *Int. J. Mod. Phys. A* **26** (2011) 2637, doi:10.1142/S0217751X11053560, arXiv:1105.1110.