

Identifying Dark Matter Event Topologies at the LHC

Yang Bai^a and Hsin-Chia Cheng^b

^a*Theoretical Physics Department, SLAC, Menlo Park, CA 94025, USA*

^b*Department of Physics, University of California, Davis, CA 95616, USA*

Abstract

Assuming dark matter particles can be pair-produced at the LHC from cascade decays of heavy particles, we investigate strategies to identify the event topologies based on the kinematic information of final state visible particles. This should be the first step towards measuring the masses and spins of the new particles in the decay chains including the dark matter particle. As a concrete example, we study in detail the final states with 4 jets plus missing energy. This is a particularly challenging scenario because of large experimental smearing effects and no fundamental distinction among the 4 jets. Based on the fact that the invariant mass of particles on the same decay chain has an end point in its distribution, we define several functions which can distinguish different topologies depending on whether they exhibit the end-point structure. We show that all possible topologies (*e.g.*, two jets on each decay chain or three jets on one chain and the other jet on the other chain, and so on) in principle can be identified from the distributions of these functions of the visible particle momenta. We also consider cases with one jet from the initial state radiation as well as off-shell decays. Our studies show that the event topology may be identified with as few as several hundred signal events after basic cuts. The method can be readily generalized to other event topologies. In particular, event topologies including leptons will be easier because the end points are expected to be sharper and there are more distinct invariant mass distributions from different charges.

1 Introduction

The Large Hadron Collider (LHC) have started taking data and many beyond the standard model (SM) theories will be tested. Many of these new models contain dark matter candidates whose stability is protected by some unbroken \mathbb{Z}_2 symmetry. Examples are the R -parity in the supersymmetric model [1] and the KK-parity in models with Universal Extra Dimensions (UEDs) [2, 3, 4]. In these models, the dark matter particle is the lightest \mathbb{Z}_2 -odd particle and may be produced in pairs at the LHC via cascade decays of heavier \mathbb{Z}_2 -odd particles. At colliders, the dark matter particles escape the detectors and result in missing energies in the events. Knowing the missing particle mass and its interactions with the SM particles is crucial for testing whether it can indeed be the major component of the dark matter in the universe.

Because only the transverse part of the sum of the missing particle momenta can be measured at hadron colliders, the kinematics can not be fully reconstructed on an event-by-event basis. It is a challenging task to measure the dark matter particle mass at hadron colliders. Nonetheless, there have been a lot of progress in developing new techniques to determine the dark matter particle mass at hadron colliders despite the partially unknown dark matter particle momenta (see [5] for a review). On a single decay chain, one can use the end points of visible particle invariant mass distributions to obtain relations among the masses of the particles on the same decay chain [6, 7, 8, 9, 10, 11, 12, 13]. If one uses both decay chains, the end point of the variable M_{T2} can also be used to determine missing particle masses [14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26]. If the decay chains are long enough and the event topology is known, one can even solve the kinematics of the full system by combining events [27, 28, 29, 30, 31].

In almost all of those analyses, the event topology has been assumed to be known from the beginning. However, it frequently happens that the same final states can come from different event topologies arising from the same model or different models. If a wrong topology is assumed, these mass determination methods will not give sensible results. Therefore, the *first* step towards measuring the dark matter mass should be identifying the dark matter event topology. In this paper, we are trying to clear the topology ambiguities and close the gap between the real experimental data and the different dark matter mass measurement techniques.

For a given final state, there are discrete choices of corresponding event topologies. The kinematic distributions are in general functions of the masses, spins and couplings of all particles in the process, in addition to the event topology. Therefore, to distinguish topologies we would like to use binary information in these distributions which is independent of the details of the model. An important

observation is that the kinematics of the visible particles from a single decay chain is constrained by the mother particle mass, and hence the invariant mass distributions of these visible particles will have end points. On the other hand, the invariant mass combinations of particles from different chains are not constrained (except by the center-of-mass energy) and are not expected to show the end-point structure. Using these facts we can design functions of the invariant mass combinations of the visible particles to exhibit end points for certain topologies but not for the others. The correct topology can emerge if there are enough such functions to differentiate all possible topologies.

If there are different types of visible particles in the event sample (for example, there are both jets and leptons or leptons with different flavors and different charges), many different kinds of invariant mass combinations can be formed. By examining whether each of them has an end point or not, one can easily tell if these objects in the corresponding invariant mass combination belong to the same decay chain. The task is more difficult if all the visible objects are jets. Not only there are fewer different types of invariant mass combinations, but the jets also suffer more from the experimental smearing effects. In this paper, rather than exhausting all possible final states we choose the quite challenging 4 jets plus missing transverse energy (MET) scenario as a case study. This final state can have a large production cross section at the LHC and a good discovery chance even with early LHC data [32] [33]. Without considering the case in which two jets are from W or Z boson decays, there are three different topologies for $4j + \cancel{E}_T$ even without the contamination of the initial state radiation (ISR): two jets on each decay chain ($2\oplus 2$); three jets on one chain and the other jet on the second chain ($3\oplus 1$); all four jets on a single chain and only the dark matter particle on the other chain ($4\oplus 0$). The $4\oplus 0$ topology in principle can be identified by looking at the invariant mass distribution of all four jets, since only $4\oplus 0$ can have an end point under this function. To identify the other cases one needs to cleverly define certain functions of the 2-jet or 3-jet invariant masses which can preserve the end-point structure given that we do not know which 2 jets or 3 jets to use *a priori*.

In our detailed study, we consider the case that the four jets have similar energies so that they can not be divided into different groups based on their E_T 's. We include the detector effects on the simulations as well as the initial and final state radiations to make the study more realistic. Furthermore, we will consider two other cases which give rise to the same final state: two jets on one chain and one jet on the other chain with the fourth jet coming from ISR; three jets on one chain and zero jet on the other chain with the fourth jet from ISR. We will show that with the help of two additional functions those two topologies may also be distinguished.

Our paper is organized as follows. In Section 2, we first define four invariant-mass functions for distinguishing the $4j + \cancel{E}_T$ event topologies based on some theoretical motivations, and we perform

the detailed parton-level studies to show that they indeed exhibit the behaviors that are expected. In Section 3 we perform the realistic particle-level analysis on these functions. Even though the showering, hadronization, and the detector smearing make the end points harder to identify, we show that by fitting the slopes of the distributions and comparing the orders we can still distinguish different topologies. We also study the two ISR cases: $2\oplus 1\oplus \text{ISR}$ and $3\oplus 0\oplus \text{ISR}$ in Section 3.1 and off-shell decays in Section 3.2. The cases with mixtures of topologies are considered in Section 3.3. We then discuss strategies to identify event topologies for other final states and conclude our paper in Section 4.

2 Kinematic Functions for $4j + \cancel{E}_T$ Event Topologies Identification

In this section, we construct functions of visible particle momenta which can be useful to identify event topologies. We use the $4j + \cancel{E}_T$ final state for a detailed case study. Similar functions can easily be constructed for other final states. As motivated by the SUSY-like models, we consider that the signal events come from the pair-production of heavy parity-odd particles, and then they go through cascade decays which end at the dark matter particle. We will first consider all those 4 jets coming from heavy parity-odd particle cascade decays and defer the cases with one jet from ISR in a later section.

Under these assumptions, there are three topologies for this final state: four jets from one chain and zero jet from the other chain; three jets from one chain and one jet from the other chain; and two jets from both chains. The Feynman diagrams for those three cases are depicted in Fig. 1. In these diagrams, we denote the dark matter particle as χ and the decaying particles in the cascading chains as Y_i . The intermediate particles in each decay chain can be either on-shell or off-shell, both of which will be covered in the following sections.

Although we do not restrict ourselves to any particular model, we note that all three topologies can arise in the popular SUSY scenario. In SUSY models, the lightest neutralino is usually the lightest supersymmetric particle (LSP) and also the dark matter particle. The $4\oplus 0$ case can come from the associated production of the lightest neutralino and gluino with the gluino going through the cascade decays: $pp \rightarrow \tilde{g} + \chi$ with $\tilde{g} \rightarrow \tilde{u}_L + 1j \rightarrow \chi_2 + 2j \rightarrow \tilde{u}_R + 3j \rightarrow \chi + 4j$. The $3\oplus 1$ case can arise from the squark pair production as $pp \rightarrow \tilde{u}_R + \tilde{u}_L$ with two squarks decaying as $\tilde{u}_R \rightarrow \chi + j$ and $\tilde{u}_L \rightarrow \chi_2 + 1j \rightarrow \tilde{u}_R + 2j \rightarrow \chi + 3j$. Finally, the $2\oplus 2$ case can come from the gluino pair production, $pp \rightarrow \tilde{g} + \tilde{g}$, with the same decaying processes on both chains as $\tilde{g} \rightarrow \tilde{u}_R + 1j \rightarrow \chi + 2j$. Whether any of these processes occur and dominate the $4j + \cancel{E}_T$ signal events depends on the spectrum and details of the specific model. It is also possible that more than one processes give comparable contributions to the particular final state. In this section we are looking for functions which can identify the event

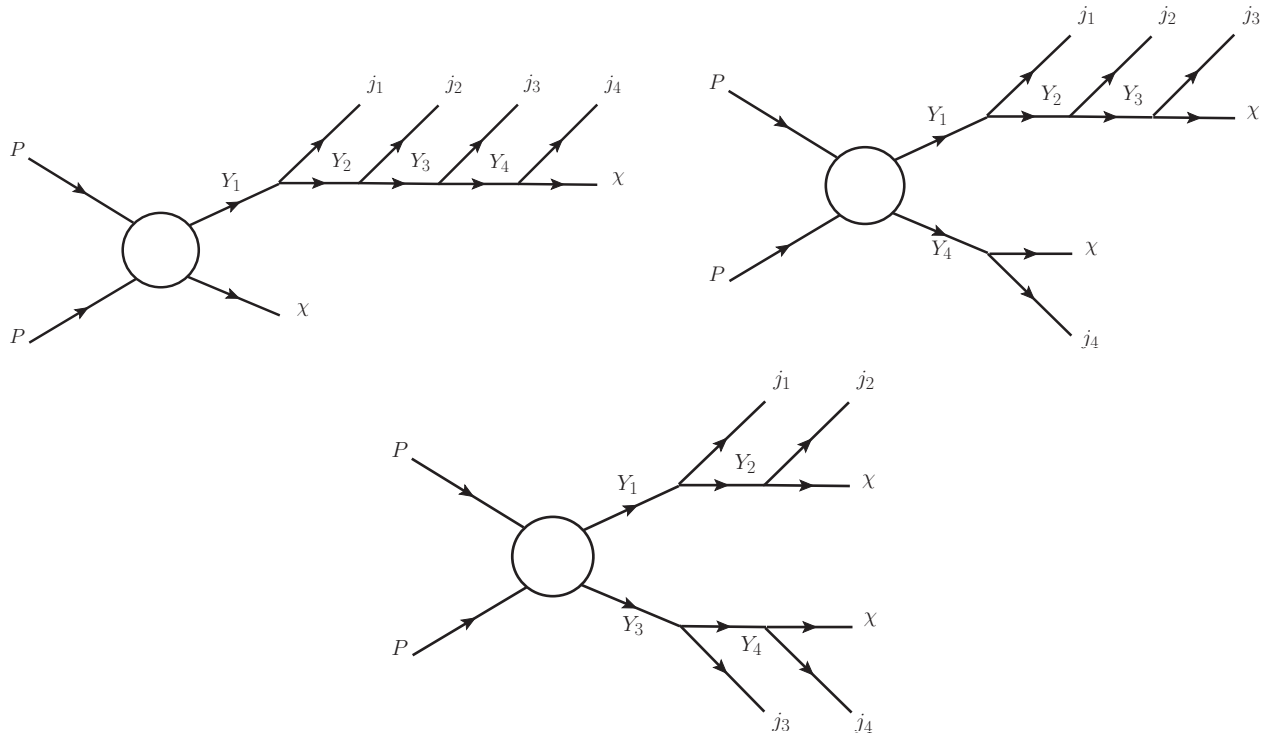


Figure 1: The three event topologies, $4\oplus 0$, $3\oplus 1$ and $2\oplus 2$, for $4j + \cancel{E}_T$ without specifying the spins of particles. The production mechanisms are not specified here and represented by a circle. The dark matter particle is denoted by χ . The intermediated particles in each chain can be either on-shell or off-shell.

topologies, so we will consider the idealized case where all events come from a single process, and leave the more complicated situations to Section 3.3 after we discuss the realistic implementation of the strategy.

For the illustration purpose, in the following when we define our kinematic functions, we demonstrate them with a SUSY spectrum such that the LSP mass is $m_\chi = 200$ GeV and in each decay the mother parity-odd particle is heavier than the neighboring daughter parity-odd particle by 200 GeV in the decay chains for each topology. The jets coming from these on-shell decays will have similar energies and hence are indistinguishable. This is the most challenging scenario. If instead there are large hierarchies among these four jets, we can consider separately the invariant mass combinations of the jets based on their energy hierarchies and obtain more handles on whether the harder jets and/or softer jets come from the same decay chain. The partonic event simulations are generated with the Madgraph/MadEvents [34] package for the 14 TeV center of mass energy of LHC with the CTEQ 6L1

[35] parton distribution functions.

The kinematic functions that we are looking for need to be able to distinguish different event topologies. As mentioned in the Introduction, the invariant mass combinations of visible particles coming from the same decay chain are constrained by the mass of the decaying mother particle and hence have end points in their distributions, while the invariant mass combinations of particles from different decay chains are not expected to exhibit the end-point structure. Therefore we focus on functions of various invariant mass combinations of the 4 jets. The end-point formulae for various invariant mass combinations of particles from the same decay chain are given in Appendix A. We found that the invariant mass distributions are sufficient and fast enough to achieve our goal. It might be worth exploring more complicated strategies and other kinematic variables to improve the results.

To identify the $4\oplus 0$ topology, the obvious function to use is the total invariant mass distribution of all four jets. We should anticipate a sharp end point for the $4\oplus 0$ case but not for the other two cases. Therefore, we define the first function, which is specifically sensitive to the $4\oplus 0$ topology, as

$$F_1(p_1, p_2, p_3, p_4) = \text{inv}[p_1, p_2, p_3, p_4]. \quad (1)$$

Here, p_i is defined according to the ordering of jet E_T in each event and $\text{inv}[\dots]$ means the total invariant mass of all momenta in the bracket, $\sqrt{(\sum_i p_i)^2}$. The F_1 distributions for three different topologies are shown in Fig. 2. One can see a sharp end-point at 800 GeV for the $4\oplus 0$ topology in

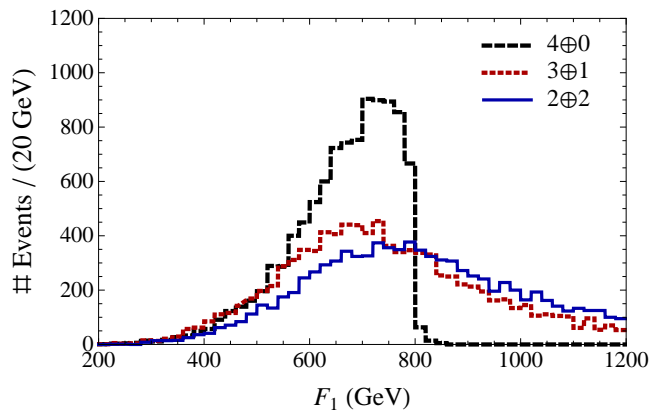


Figure 2: The number of events in terms of the function F_1 defined in Eq. (1) at the parton level. There are totally 10000 events for each topology generated in this plot.

Fig. 2. The sharp end point is expected to get smeared after showering, hadronization, and detector resolution effects. The strategy to deal with the realistic situations will be discussed in the next section.

For the $3\oplus 1$ topology, we have found two potentially useful functions. The first one is the smallest 3-particle invariant mass combination:

$$F_2(p_1, p_2, p_3, p_4) = \min\{\text{inv}[p_1, p_2, p_3], \text{inv}[p_1, p_2, p_4], \text{inv}[p_1, p_3, p_4], \text{inv}[p_2, p_3, p_4]\}. \quad (2)$$

The minimum of the invariant masses of the four combinations has a high probability to find the set of three jets on the same chain. Even if it sometimes picks the wrong set, it will not exceed the expected end point since it is smaller than the correct combination. Plotting the event numbers as functions of F_2 , one may expect a sharp end point for the $3\oplus 1$ case but not for the $2\oplus 2$ case. An end point for the $4\oplus 0$ topology is also expected since all visible particles come from the same chain. Another function uses 2-particle invariant masses. Since the invariant mass of 2 particles from different chains can be very large, we consider the invariant mass of the pair of the particles which is opposite to the pair that forms the largest invariant mass:

$$F_3(p_1, p_2, p_3, p_4) = \text{inv}[p_k, p_l],$$

such that $\epsilon^{kl ij} \neq 0$ and $\text{inv}[p_i, p_j] = \max\left\{\bigcup_{m,n} \text{inv}[p_m, p_n]\right\}$.

(3)

Here, $i, j, k, l, m, n = 1, \dots, 4$ and $\epsilon^{kl ij}$ is the totally antisymmetric tensor. For the $3\oplus 1$ case, this pair of particles has a large chance to come from the same chain. Even if occasionally they come from different chains, their invariant mass is bounded by an invariant mass from the same chain and hence will not exceed the corresponding end point. On the other hand, this combination for the $2\oplus 2$ case is likely to come from opposite chains and is not expected to have an end point. The event distributions of these two functions are shown in Fig. 3, where one can see that both $4\oplus 0$ and $3\oplus 1$ topologies have obvious end points at around 600 GeV for F_2 and around 400 GeV for F_3 . (The exact expected values can be found in Appendix A.) On the contrary, the $2\oplus 2$ distributions have long tails without end points. In the next section we find that F_2 seems to work better than F_3 after the experimental smearing. However, F_3 may still be useful for some other final states (*e.g.*, leptons which do not suffer too much from smearing effects).

To identify the $2\oplus 2$ topology, we define the following function which is sensitive to this topology:

$$F_4(p_1, p_2, p_3, p_4) = \min\left\{\bigcup_{i,j} \max(\text{inv}[i, j], \text{inv}[k, l])\right\} \quad \text{for } \epsilon^{kl ij} \neq 0. \quad (4)$$

For each event there are 3 ways to pair those four jets. One first chooses the pair with the larger invariant mass for each way of pairing, then calculate the minimum of the larger invariant masses

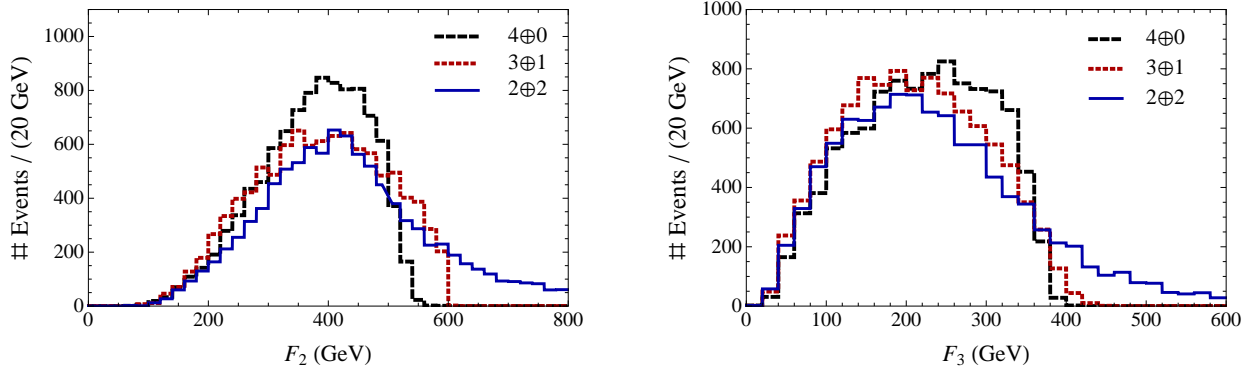


Figure 3: Left panel: the number of events in terms of the function F_2 at the parton level. There are totally 10000 events generated for each topology in this plot. Right panel: the same as the left one but in terms of F_3 .

among those three partitions. For the $2\oplus 2$ case, the partition that each pair comes from the same decay chain will have both invariant masses of the two pairs to be bounded from above. Therefore, the correct partition is likely to have a smaller value for the maximum of the two invariant masses among the three partitions. On the other hand for the $3\oplus 1$ topology, it is quite possible that the larger invariant mass pairs for the three partitions all include the only jet from the shorter decay chain and therefore no end point is expected. The simulated results are shown in Fig. 4. Indeed one can see

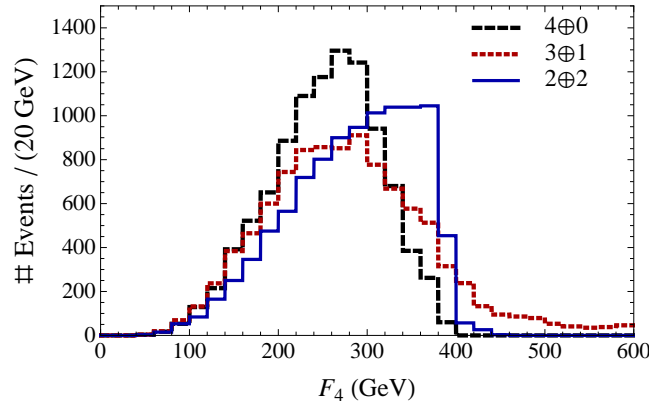


Figure 4: The number of events in terms of the function F_4 defined in Eq. (4) at the parton level. There are totally 10000 events generated for this plot.

a sharp end point at around 400 GeV for the $2\oplus 2$ topology but not for the $3\oplus 1$ topology. The $4\oplus 0$ topology does have an end point but it is not very sharp.

At the parton level, one can see that the three different topologies can be easily identified by checking whether there are end points in those four functions F_1-F_4 . However, including the showering, hadronization, and other experimental smearing effects these end points become less distinct and more concrete strategies need to be developed to identify the topologies, which is the subject of the next section.

3 Dealing with Realistic Particle-Level Event Distributions

In reality, the kinematic distributions proposed in the previous section will receive significant experimental smearing effects from showering, hadronization, detector resolutions, backgrounds, and so on. It is important to check whether the distinctive features of these functions can survive the smearing effects. We also need an unambiguous procedure to identify the event topology from these kinematic functions rather than just looking for end points by naked eyes.

To include the experimental smearing effects for our particle-level analysis, we first generate parton-level events using Madgraph/MadEvents as before. We then process the parton level events with PYTHIA [36] for showering and hadronization including initial and final state radiations, and PGS [37] (with the default CMS detector card) for the detector simulation. Some basic cuts are imposed on all signal events after the detector simulation.

We start with the same spectra used in the previous section with all on-shell decays and generate 10000 events for each topology. We require at least four jets with $E_T > 100$ GeV and the missing transverse energy $\cancel{E}_T > 200$ GeV on the events. This set of cuts are utilized just for the illustration purpose, and our following analysis is insensitive to those cuts. The signal selection efficiencies of $4\oplus 0$, $3\oplus 1$ and $2\oplus 2$ are 19.3%, 10.1% and 13.1%, respectively.

The event distributions in terms of those four functions F_i with $i = 1, \dots, 4$ are shown in Fig. 5. Compared to the parton level distributions in Fig. 2, 3, 4, we can still roughly tell some of the end points for the first three functions. The end point structure for the last function F_4 is not so clear on the other hand. Nevertheless, for the distributions which were supposed to have end points, the after-peak slopes look steeper than the ones without end points. We will explore this observation to come up with a concrete procedure to identify different topologies.

To find the falling slopes after the peaks of these distributions, we will try to fit the histograms in Fig. 5 with some simple functions. The first fitting function that we consider is the following log-normal function:

$$h_1(F_i) = a_0 \times \exp \left\{ - \left(\frac{\ln[F_i/a_1]}{a_2} \right)^2 \right\}. \quad (5)$$

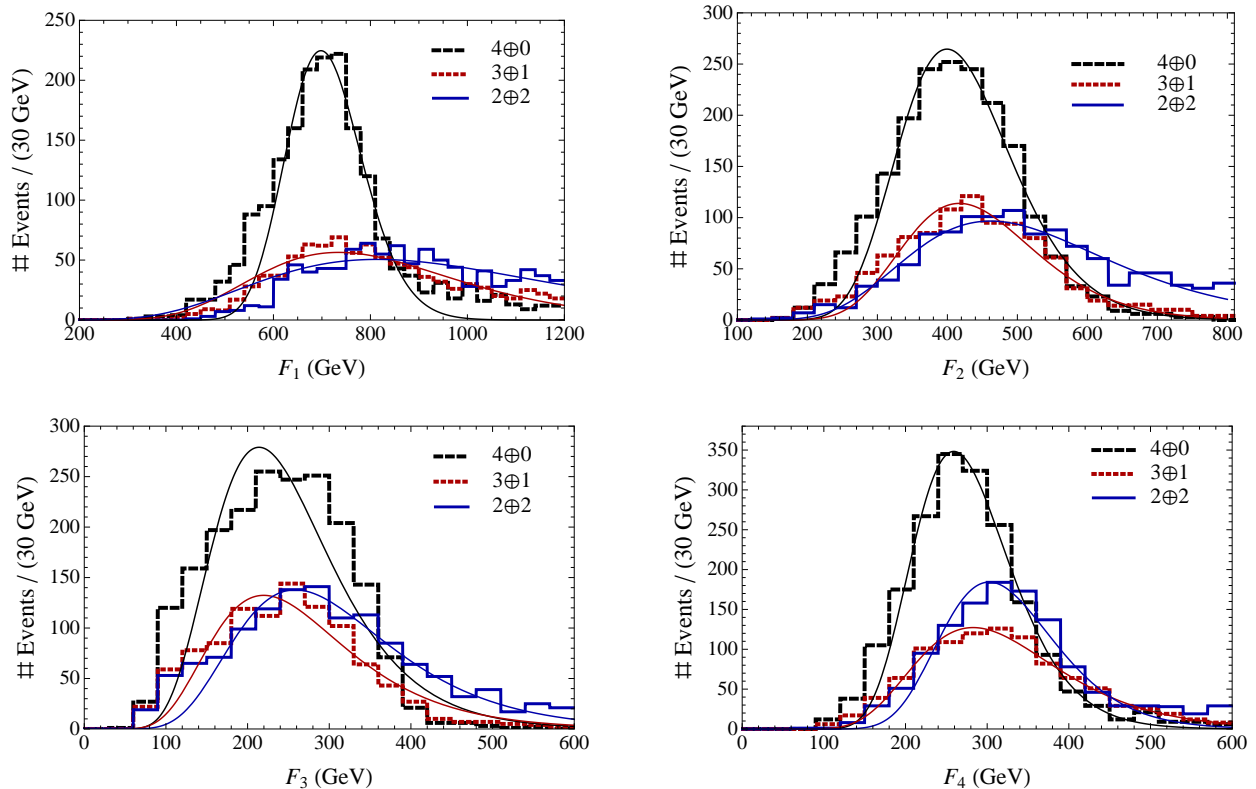


Figure 5: The number of events as functions of those four variables after passing PYTHIA and PGS. We select the signal events by requiring: at least 4 jets with $E_T > 100$ GeV and missing transverse energy $\cancel{E}_T > 200$ GeV. The continuous lines are fitted lines using the log-normal function. Only bins with a height above 1/2 of the peak height on the left side and above 1/4 of the peak height on the right side are included in the fit.

This log-normal function has a peak structure with asymmetric half-height widths. The three parameters a_0 , a_1 and a_2 determine the peak height, the peak location and the falling slope after the peak. The last parameter a_2 can be used to distinguish different topologies. A smaller value of a_2 means a steeper slope after the peak. Given the large statistical uncertainties on the tails of those distributions and also potentially large contaminations from backgrounds, we only include bins with a height above one half of the peak height on the left side and above one quarter of the peak height on the right side into the fit. The χ^2 in our fit is defined as

$$\chi^2 = \sum_{i_{\text{bin}}} \frac{(h_1^{i_{\text{bin}}} - s^{i_{\text{bin}}})^2}{s^{i_{\text{bin}}}}, \quad (6)$$

to only take into account the statistical uncertainties from the signal events. The asymmetric choice

of bins around the peak is to minimize the effects of cuts, which affect the left side of the peak more severely. The fitted curves are shown in the continuous lines in Fig. 5 for the 30 GeV bin size. As one can see, except the F_3 distributions all other histograms are well fitted by the log-normal functions.

The fitted values of a_2 are listed in Table 1 for 30 GeV and 40 GeV bin sizes. Smaller values of the

Topologies	bin size = 30 GeV				bin size = 40 GeV			
	$a_2(F_1)$	$a_2(F_2)$	$a_2(F_3)$	$a_2(F_4)$	$a_2(F_1)$	$a_2(F_2)$	$a_2(F_3)$	$a_2(F_4)$
$4\oplus 0$	0.16	0.29	0.48	0.32	0.16	0.29	0.52	0.30
$3\oplus 1$	0.40	0.32	0.53	0.43	0.39	0.30	0.51	0.43
$2\oplus 2$	0.52	0.44	0.52	0.32	0.52	0.48	0.51	0.29

Table 1: The fitted values for a_2 of the log-normal function in Eq. (5), which determine the steepness of the slopes of the histograms after the peak. The bin size has been chosen to be 30 GeV (left), and 40 GeV (right).

bin size introduce larger fluctuations of the number of events in each bin, and make the fit unreliable. Comparing those numbers for the 30 GeV and the 40 GeV bin sizes, we can see that the fitted slopes are fairly consistent.

To examine the sensitivity on the functions used to do the fits, we also try to fit the histograms with two straight lines around the peak. This fit also gives an estimate of the end point value at the point when the right straight line crosses zero. The “broken-line” function used in the fits is given by

$$h_2(F_i) = \begin{cases} -c(F_i - a) + b & \text{for } F_i \geq a, \\ d(F_i - a) + b & \text{for } F_i < a, \end{cases} \quad (7)$$

with all parameters being positive. The parameters “ (a, b) ” determine the location of the peak and “ c ” and “ d ” determine the slopes of curves on the right side and the left side of the peak. For the 30 GeV bin size, we compare the fitted results with the simulated distributions in Fig. 6. We see that the “broken-line” function can fit all histograms including F_3 pretty well. The slope “ c ” depends on the number of events and the overall scale of the invariant mass functions. To quantify the differences of the slopes on the right side of the peaks for different topologies and functions, we define the dimensionless combination $b/(ac)$ to be inverse of the normalized slope, which is independent of the number of events and the overall scale. A smaller value of $b/(ac)$ means a steeper slope on the right side of the peak. For two different bin sizes, 30 GeV and 40 GeV, we present the fitted values of $b/(ac)$ in Table 2. Similarly, the fitted results are insensitive to the choices of bin sizes.

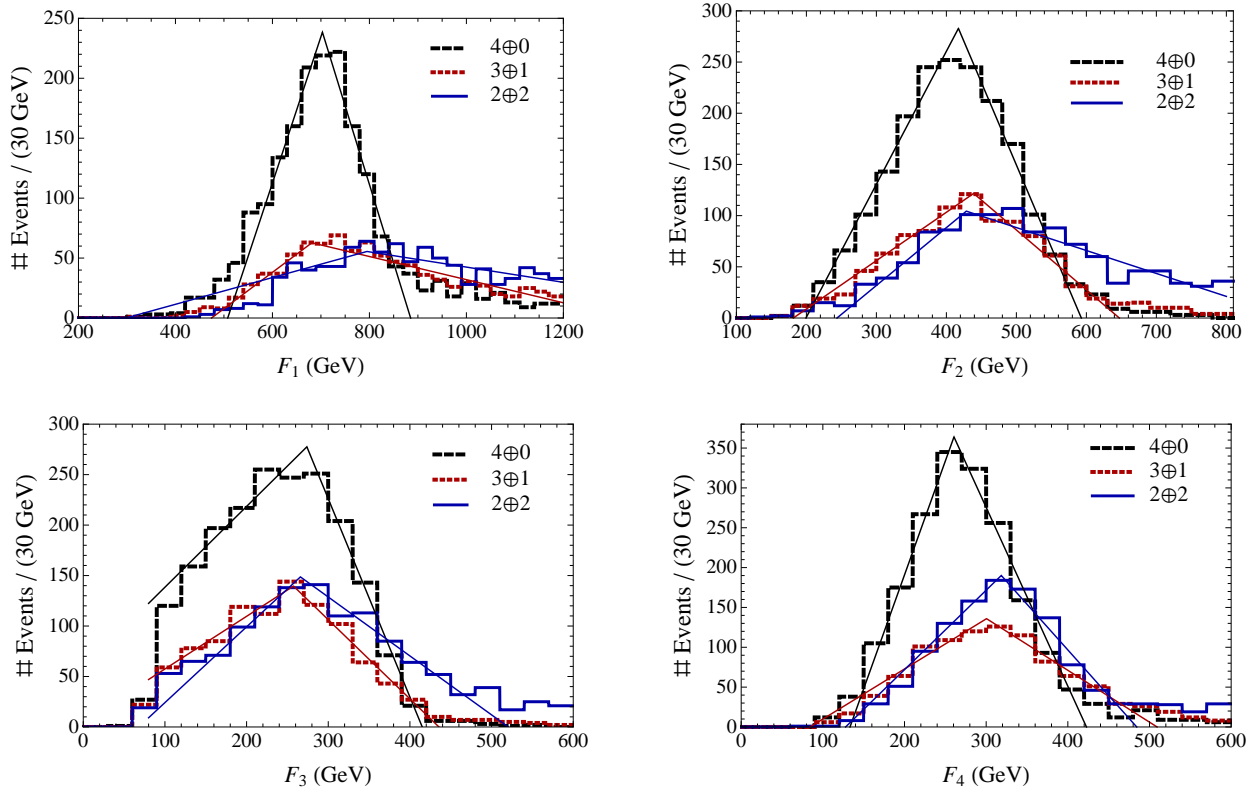


Figure 6: The same as Fig. 5 but using the broken-line function defined in Eq. (7) to fit the distributions.

Topologies	bin size = 30 GeV				bin size = 40 GeV			
	$\frac{b}{ac}(F_1)$	$\frac{b}{ac}(F_2)$	$\frac{b}{ac}(F_3)$	$\frac{b}{ac}(F_4)$	$\frac{b}{ac}(F_1)$	$\frac{b}{ac}(F_2)$	$\frac{b}{ac}(F_3)$	$\frac{b}{ac}(F_4)$
4⊕0	0.26 ± 0.02	0.42 ± 0.04	0.52 ± 0.03	0.62 ± 0.04	0.24	0.48	0.47	0.49
3⊕1	0.95 ± 0.19	0.47 ± 0.06	0.69 ± 0.09	0.70 ± 0.10	0.95	0.49	0.70	0.70
2⊕2	1.10 ± 0.30	1.09 ± 0.15	0.96 ± 0.14	0.52 ± 0.05	1.09	1.09	1.21	0.53

Table 2: The fitted values for $b/(ac)$ of the “broken-line” function in Eq. (7), which determine the normalized steepness of the slope of the histogram after the peak. The bin size has been chosen to be 30 GeV (left), and 40 GeV (right). We have also shown the 1σ statistical uncertainties for the 30 GeV bin size, based on 1932, 1013 and 1313 events after cuts for 4⊕0, 3⊕1 and 2⊕2, respectively.

One can naïvely define the end point to be the intersecting point of the right branch of the broken-line with the x -axis, which is given by $F_i^{\text{end}} \equiv a + b/c$, as a simple estimate of the true end point. The fitted values of the end points in terms of different functions are listed in Table 3. As one might have

	bin size = 30 GeV				bin size = 40 GeV			
	F_1^{end}	F_2^{end}	F_3^{end}	F_4^{end}	F_1^{end}	F_2^{end}	F_3^{end}	F_4^{end}
4 \oplus 0	885 \pm 19 (800)	592 \pm 19 (600)	415 \pm 10 (394)	422 \pm 11 (397)	886	603	408	409
3 \oplus 1	1332	647 \pm 30 (600)	436 \pm 25 (< 458)	509	1329	652	441	512
2 \oplus 2	1668	894	520	484 \pm 17 (387)	1670	898	554	483

Table 3: The fitted values (in GeV) of the end points, which are the right intersection point of broken lines with the x -axis. The bin size has been chosen to be 30 GeV (left), and 40 GeV (right). The numbers in the parenthesis are the end-point values at the parton level. We have also shown the 1σ statistical uncertainties for the same event samples as in Table 2 for the 30 GeV bin size.

expected, this definition tends to give a larger value of the end point than the actual value for most of the cases having a real end point.

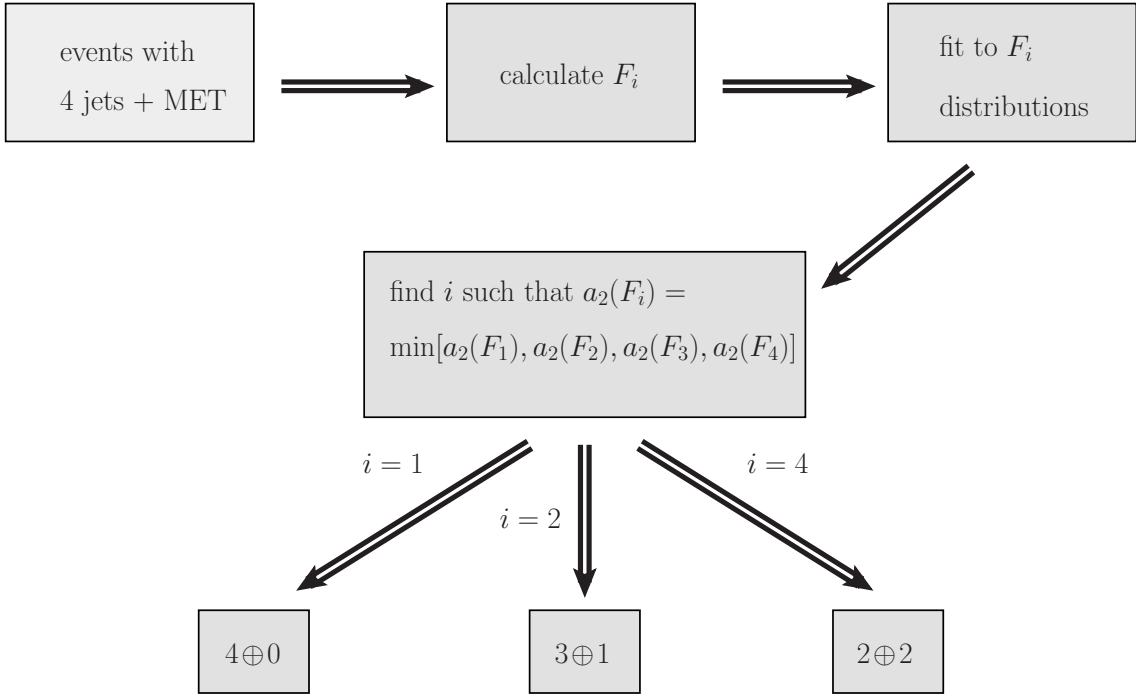


Figure 7: The flow chart of our procedures to identify the event topologies with $4j + \cancel{E}_T$ when using the log-normal function to fit the F_i distributions. For the broken-line fit, one needs to replace a_2 by $b/(ac)$.

After fitting the shapes of distributions, we can now use the information in Table 1 and 2 to identify different topologies. From these tables, one can see that the topology can be identified by finding “ i ”

of the minimum of all four $a_2(F_i)$ for the log-normal fit or $b/(ac)(F_i)$ for the broken-line fit. If $i = 1$, then the topology is $4\oplus 0$; if $i = 2$ (or 3), it is $3\oplus 1$; if $i = 4$, it is $2\oplus 2$. The procedure described above to identify the dark matter event topologies with $4j + \cancel{E}_T$ is summarized in the flow chart in Fig. 7. The two choices of the fitting functions give similar results, so we will only use the broken-line fits in the rest of the paper.

Based on the central values and 1σ statistical uncertainties in Table 2, we estimate the required numbers of events to achieve 90% of time correct for the topology $4\oplus 0$ to be 446, $3\oplus 1$ to be 777 and $2\oplus 2$ to be 740. All of those numbers are defined after the basic kinematic cuts and assuming negligible SM backgrounds. A more complete analysis requires simulations of SM backgrounds and is beyond the scope of this paper.

3.1 Initial State Radiation

There exist other possibilities to have the $4j + \cancel{E}_T$ final state after passing the basic kinematic cuts. For example, two of the jets come from a resonance like the W boson. For this case, one can first check the two-body invariant mass distribution which should have an obvious mass peak, and then translate the $4j + \cancel{E}_T$ final state to the $W + 2j + \cancel{E}_T$ final state. It could be easier to identify the three topologies for the $W + 2j + \cancel{E}_T$ final state following a similar method described before. Another possibility is to have the partonic events to be $5j + \cancel{E}_T$ but with one jet being soft and lost by basic cuts. If this indeed happens a lot, two dark matter parity-odd particles in the same decay chain may have masses close to each other. For this case, one can only treat those events as real $4j + \cancel{E}_T$ events and apply the procedure proposed earlier to at least understand the relative positions of the four hard jets, though the end points become less sharp and the formulae in Appendix A may not be accurate.

The third possibility is to have only 3 jets coming from the decay chains but the fourth jet from ISR. There are two types of topologies, $3\oplus 0\oplus \text{ISR}$ (three jets from a single chain) and $2\oplus 1\oplus \text{ISR}$ (two jets from one chain and another jet from the other chain). In this subsection, we describe the additional procedures for identifying those two new topologies with one ISR jet.

The $3\oplus 0\oplus \text{ISR}$ topology is very similar to the $3\oplus 1$ topology. The only difference is that the isolated jet comes from ISR instead of from the decay of a heavy particle. Indeed, the fitted slopes for the $3\oplus 0\oplus \text{ISR}$ topology have the same pattern as the $3\oplus 1$ topology, with the F_2 function giving the steepest slope. To distinguish it from the $3\oplus 1$ topology, we should examine the E_T distribution of the isolated jet. If it comes from the ISR, it should have a falling E_T distribution starting from the kinematic cut. On the other hand, a jet coming from a heavy particle decay should have a peak value determined by the mass difference. Because the 3 jets from the same decay chain are likely to have a

smaller invariant mass, we define the following $3\oplus 0\oplus \text{ISR}$ -specified function:

$$F_5(p_1, p_2, p_3, p_4) = E_T(p_k),$$

such that $\epsilon^{klj} \neq 0$ and $\text{inv}[p_l, p_i, p_j] = \min\left\{\bigcup_{r,s,t} \text{inv}[p_r, p_s, p_t]\right\}.$ (8)

The distributions of F_5 for different topologies are shown in Fig. 8. There, we use the broken-line

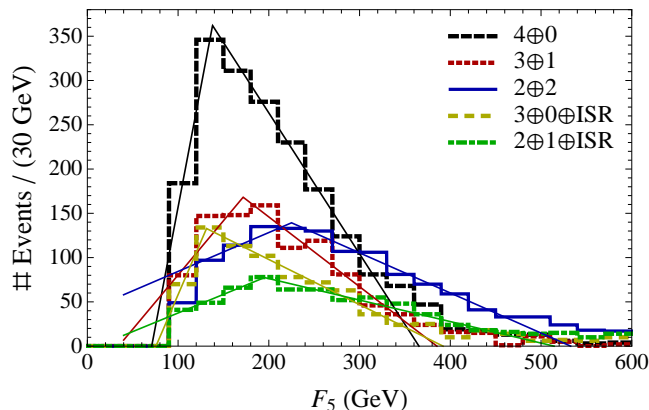


Figure 8: The distributions of F_5 for different topologies. 10000 events are generated for each topology before the cuts. There are 826 $3\oplus 0\oplus \text{ISR}$ events and 699 $2\oplus 1\oplus \text{ISR}$ events after the cuts.

function to fit various distributions with 30 GeV bin size. As can be seen from this plot, $3\oplus 1$ and $3\oplus 0\oplus \text{ISR}$ have different locations of the peak with a lower value for the $3\oplus 0\oplus \text{ISR}$ case. To quantify this difference, we compare the dimensionless parameter d/c , which measuring the ratio of the rising slope before the peak and the falling slope after the peak. The numerical values of the $b/(ac)$ and d/c for those five topologies and five functions are listed in Table 4. We see that the $3\oplus 0\oplus \text{ISR}$ and $3\oplus 1$ can be distinguished by examining $d/c(F_i)$. The $3\oplus 0\oplus \text{ISR}$ has the largest value in F_5 while the largest value for $3\oplus 1$ occurs at F_1 . For the $2\oplus 1\oplus \text{ISR}$ topology, it is not expected to have an end point for any of these functions. Indeed, we find that all four $b/(ac)(F_i)$ are comparable and none of them takes a value as small as those ones with end points for the corresponding topology. That it is not any of the other topologies which should have at least one end point among the 4 functions can be taken as a sign for the $2\oplus 1\oplus \text{ISR}$ topology. It is also not so easy to identify the ISR jet for this topology. To unambiguously distinguish it from other topologies is more challenging and probably requires additional functions. We will have more discussion on this topology in Appendix B.

	$\frac{b}{ac}(F_1)$	$\frac{b}{ac}(F_2)$	$\frac{b}{ac}(F_3)$	$\frac{b}{ac}(F_4)$	$\frac{b}{ac}(F_5)$	$\frac{d}{c}(F_1)$	$\frac{d}{c}(F_2)$	$\frac{d}{c}(F_3)$	$\frac{d}{c}(F_4)$	$\frac{d}{c}(F_5)$
4 \oplus 0	0.26	0.42	0.52	0.62	1.65	0.93	0.81	0.41	1.27	3.39
3 \oplus 1	0.95	0.47	0.69	0.70	1.24	3.14	0.81	0.67	1.36	1.56
2 \oplus 2	1.10	1.10	0.96	0.52	1.37	1.76	2.52	1.29	0.95	0.97
3 \oplus 0 \oplus ISR	0.84	0.48	0.60	0.93	1.98	1.51	0.98	0.46	2.39	4.68
2 \oplus 1 \oplus ISR	1.00	0.96	1.14	1.31	1.63	1.18	1.70	0.90	2.29	1.73

Table 4: The fitted values of $b/(ac)$, which determine the normalized steepness of the slopes of the histograms after the peak, and the fitted values of d/c , which determine the ratios of the rising slopes and the falling slopes around the peak. The broken-line function in Eq. (7) is used to fit the distributions. The bin size has been chosen to be 30 GeV.

3.2 Off-Shell Decays

It frequently happens in models with dark matter that some particles have the dominant decays to be three-body processes. For example in SUSY, if squarks are heavier than the gluino, the dominant decay channel of the gluino is to two quarks plus one neutralino via off-shell squarks. In this subsection, we use the SUSY model to generate events with off-shell decay processes in the decay chains with the assistance of the Monte-Carlo tool BRIDGE [38]. For the 2 \oplus 2 event topology, the events are generated by pair producing two gluinos, and then decaying each one to $\bar{u} + u + \chi$ via off-shell squarks. For 3 \oplus 1, we first generate events with $\tilde{u}_L + \tilde{u}_R$ in the final state, and then we require $\tilde{u}_L \rightarrow \bar{u} + u + \tilde{u}_R$ via the off-shell gluino or neutralino and $\tilde{u}_R \rightarrow u + \chi$. For 4 \oplus 0, we first generate events with $\tilde{g} + \chi$ in the final state, and then we require $\tilde{g} \rightarrow \bar{u} + u + \chi_2$ and $\chi_2 \rightarrow \bar{u} + u + \chi$ via off-shell squarks. To produce events with similar visible kinematics for all three topologies, the LSP χ is fixed to have a mass of 200 GeV, and the mass difference between the mother superparticle and the daughter superparticle is chosen to be 200 GeV for two-body decays and 400 GeV for three-body decays. As a consequence, all four jets should have similar E_T distributions on average.

After the basic cuts on all signal events: at least four jets with $E_T > 100$ GeV and the missing transverse energy $\cancel{E}_T > 200$ GeV, the acceptance efficiencies are 11.5%, 6.7% and 9.2% for 4 \oplus 0, 3 \oplus 1 and 2 \oplus 2, respectively. Those efficiencies are lower than the on-shell decay cases simply because jets in the off-shell case have a larger probability to become soft.

Repeating the same procedure as in the on-shell decay case, we have fitted the slopes of the distributions after the peak in Table 5. Comparing Table 5 with Table 2 for the on-shell case, we can

Topologies	bin size = 25 GeV				bin size = 30 GeV			
	$\frac{b}{ac}(F_1)$	$\frac{b}{ac}(F_2)$	$\frac{b}{ac}(F_3)$	$\frac{b}{ac}(F_4)$	$\frac{b}{ac}(F_1)$	$\frac{b}{ac}(F_2)$	$\frac{b}{ac}(F_3)$	$\frac{b}{ac}(F_4)$
$4\oplus 0$	0.40	0.93	0.92	0.82	0.44	0.85	0.82	0.78
$3\oplus 1$	0.96	0.62	0.80	0.80	0.80	0.67	0.58	0.84
$2\oplus 2$	1.47	1.09	0.84	0.74	1.47	1.15	0.83	0.81

Table 5: The fitted values of $b/(ac)$ for the off-shell decay case. The broken-line fit is used to obtain those numbers. Again each topology has 10000 events before the cuts, and there are 1152, 671 and 915 events passing the cuts for $4\oplus 0$, $3\oplus 1$ and $2\oplus 2$ respectively.

see that the differences of $b/(ac)(F_i)$ are reduced for the off-shell case. Although a similar selection criterion like Fig. 7 can still be used to identify those three topologies, more signal events or a higher integrated luminosity at the LHC are required to identify the topologies for the off-shell case. After the event topology is identified, the on-shell and off-shell decay cases may be distinguished by examining the correlations of the invariant masses [12] and/or the function $M_{T2,\max}$ [14, 15, 17, 18, 19].

For completeness, we also report the fitted end-point values in Table 6 for the 30 GeV bin size.

	bin size = 30 GeV			
	F_1^{end}	F_2^{end}	F_3^{end}	F_4^{end}
$4\oplus 0$	977 (800)	660 (< 754*)	411 (< 754*)	438 (< 754*)
$3\oplus 1$	1343	669 (600)	413 (< 600*)	538
$2\oplus 2$	1680	806	485	483 (400)

Table 6: The fitted values (in GeV) of the end points, which are the right intersection points of broken lines with the x -axis. The bin size has been chosen to be 30 GeV. The numbers in the parenthesis are the end-point values at the parton level. The numbers with * occur in the soft limit of some jet(s) which in practice will not pass the cuts. Therefore the actual end points are much smaller.

3.3 The Cases of Mixtures of Different Event Topologies

It may also happen that the signal events of the same final state come from a combination of two or more different event topologies. The method discussed in this paper should work when one topology dominates over the others. To quantify the limit at which the topology can be identified by our method, we use a mixture of $2\oplus 2$ and $3\oplus 1$ topologies as a case study. Fixing the total number of

events after the basic cuts to be 1000, we study the patterns of the fitted $b/(ac)$ as a function of the mixture fraction, x . For $x = 0$, all events are from the $3\oplus 1$ topology, while for $x = 1$ all events are from the $2\oplus 2$ topology. For the 4 functions F_1-F_4 , the fitted values of $b/(ac)$ as functions of the mixture fraction are shown in Fig. 9. As x increases, one can see from Fig. 9 that there is a transition

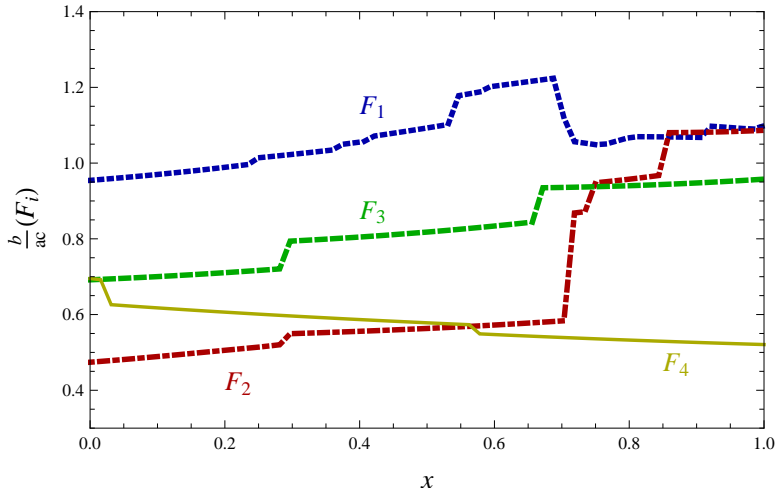


Figure 9: The fitted values of $b/(ac)$ for different functions F_i as functions of x , where x is the fraction of $2\oplus 2$ type events in a mixture of $3\oplus 1$ and $2\oplus 2$ topologies with 1000 combined events (after the basic cuts). The bin size is 30 GeV. Only bins with a height above $1/2$ of the peak height on the left side and above $1/4$ of the peak height on the right side are included in the fit.

of the function with the smallest $b/(ac)$ from F_2 to F_4 . The transition happens around $x \approx 50\%$ with a mixture of equal amount of $2\oplus 2$ and $3\oplus 1$ events. It is not surprising that if the signal is indeed a mixture of $2\oplus 2$ and $3\oplus 1$ with comparable weights, one can not identify the topology unless additional kinematic information can split the signal events into the two categories. The abrupt changes of lines in Fig. 9 are due to changes of bins during the fit procedure by only including bins with a height above $1/2$ of the peak height on the left side and above $1/4$ of the peak height on the right side.

In our analysis, the SM backgrounds are neglected. Since there is no reason to anticipate SM backgrounds to have end points in terms of those functions F_i , the pattern of the fitted slopes for different F_i should not be modified much if the signal events dominate over the backgrounds. As an illustration, we treat the events from $2\oplus 1\oplus \text{ISR}$ as the backgrounds of $2\oplus 2$ signals, and show the patterns as a function of percentage of $2\oplus 2$ in Fig. 10. If the signal events are over around 50% of the total number of events, the $b/(ac)(F_4)$ is the smallest among all F_i functions. In this situation, we still can determine the signal events are from the $2\oplus 2$ topology.

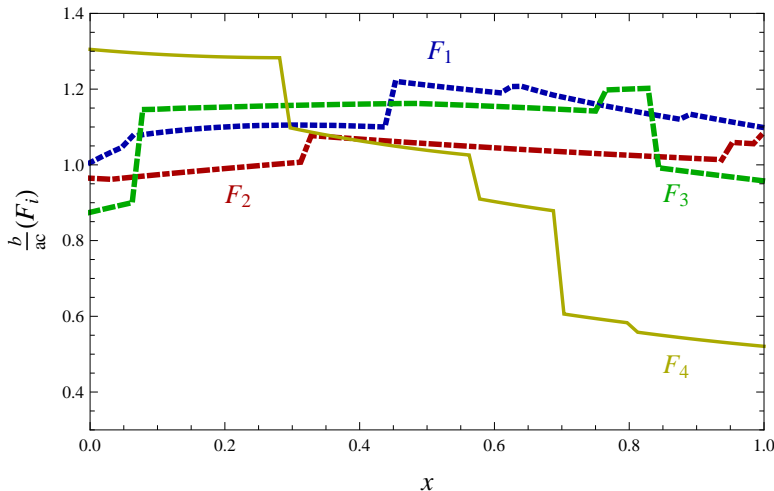


Figure 10: The same as Fig. 9 but for a mixture of $2\oplus 1\oplus \text{ISR}$ and $2\oplus 2$ topologies with 1000 combined events (after the cuts). Here, x is the fraction of $2\oplus 2$ type events in a mixture of $2\oplus 1\oplus \text{ISR}$ and $2\oplus 2$ topologies.

4 Discussion and Conclusions

Although we focus on the $4j + \cancel{E}_T$ final state in this paper, a similar set of functions can be used to identify event topologies of $nj + \cancel{E}_T$ with $n \geq 3$. Without ISR, there are $[n/2] + 1$ different topologies with two decay chains. For $n = 2k + 1$, one can define $k + 1$ functions to find the minimum of all possible invariant masses of m jets for $m = k + 1, \dots, 2k + 1$. For $n = 2k$, other than the k functions defined to be the minimum of all possible invariant masses of m jets for $m = k + 1, \dots, 2k + 1$, one additional function can be defined as the minimum of the maximum of two invariant masses of k jets among all possible combinations (similar to the function F_4). One can also define functions analogous to F_3 by removing some jets which form the largest invariant mass and looking at the invariant mass of the other jets. For the cases with ISR, the general strategy would be to first identify the existence of an ISR jet by the jet E_T distribution and then define functions based on the remaining $n - 1$ jets. From our study of the $4j + \cancel{E}_T$ example, we found that some cases with ISR are difficult to identify without lowering the E_T cuts on jets. For $2j + \cancel{E}_T$, the general strategy does not work. The case with 2 jets on the same chain may be checked by finding the end point of the 2-jet invariant mass. On the other hand, the topology with one jet on each decay chain can not be identified with invariant mass functions. Other kinematic functions like M_{T2} or M_{CT} [39] might be useful for this topology.

If the final states also contain leptons in addition to jets plus \cancel{E}_T , their charges, flavors and good energy resolutions can give better handles for the event topology identification. The lepton number and flavor conservations (if they are good symmetries) may already give us some hints on the type of topologies. For instance, consider the $2j + 2\ell^+ + 2\ell^- + \cancel{E}_T$ final state. The SUSY-like theories suggest that the leptons from the same decay chain should have the same flavor and opposite charges. However, one can imagine more general models which may contain doubly charged particles and/or large lepton flavor violations and hence invalidate the argument. Nevertheless, the strategies taken in this paper should be able to identify event topologies without these assumptions. We can first take the four leptons and use the methods in this paper to find out their distribution on the two chains. The case with 4 leptons on the same chain can be identified with the function F_1 . The function F_2 (which is the minimum of the 3-lepton invariant masses) can be used to identify the case with 3 leptons from one chain and the other lepton from the other chain. The functions F_3 and F_4 can be divided into more functions based on whether we take the invariant mass of the 2 leptons of the same charge or the opposite charges. For example, the function F_4 restricted to opposite-charge invariant masses can identify the case with 2 opposite-charged leptons on each chain, while the pair of the same-charge 2-lepton invariant masses can identify the case with the same-sign leptons on each chain. Once the relative distribution of the leptons on the 2 chains has been fixed, one can check the existence of the end point for the invariant mass distribution of two jets. If there is no end point, one simply assigns one jet on each chain. If there is an end point, the 2 jets are on the same chain. If the lepton distribution is not symmetric (for example, 3 leptons on one chain and one lepton on the other chain), one can then check the invariant mass combination of the 2 jets together with the 3 leptons to see if they come from the same chain. In this way, we know how the visible particles distributed between the two chains. To remove the order ambiguity of the jets and leptons on a single chain probably requires much more sophisticated analysis which may depend on the details of the event kinematic distributions.

In this paper we have generated events from SUSY models with certain specific spectra. The strategies employed in this paper should work for general dark matter models with an unbroken \mathbb{Z}_2 discrete symmetry [3, 4, 40, 41, 42, 43], or even models with more complicated symmetries (*e.g.*, Ref. [44]) as long as the events contain 2 decay chains which end with missing particles. Additional missing particles like neutrinos coming from the decay chains may be difficult to be identified themselves, but should not prevent us from identifying the topology of the visible particle part. There could also be accompanied events with neutrinos replaced by its charged lepton partners which may be used to identify the topology. The invariant mass distributions may have different behaviors for different models.

Depending on the spins of the intermediate particles, the sharpness of the end points of invariant mass distributions varies [45]. For example, comparing the processes of $\tilde{g} \rightarrow \bar{q} + \tilde{q} \rightarrow \bar{q} + q + \chi$ in SUSY and $g^{(1)} \rightarrow \bar{q} + q^{(1)} \rightarrow \bar{q} + q + B^{(1)}$ in the UED model, the invariant mass distribution of two jets in the UED model has a sharper end-point than in SUSY because it has a spin-1/2 intermediate particle, in contrast to the spin-0 particle in the SUSY case. In addition, the production cross section can vary a lot among different models with different spectra. Therefore, the actual required luminosity at the LHC to identify a particular topology depends on the specific model and requires a detailed study for the individual model. (For a general study of model discriminations at the LHC, see [46] [47].)

Although we have concentrated on identifying event topologies in this paper, we would like to point out that the functions defined in this paper can also be useful for reducing the combinatorial problems. Once a particular topology is identified. The functions in which this topology has end points can be used to determine the end points of some invariant masses of particles from the same decay chain. Then the end points can be used to cut wrong combinations [30, 48, 49] by performing an event-by-event analysis and removing combinations with the invariant masses above the corresponding end-point value. The order of the visible particles from a single chain for each individual event remains a difficult problem without additional handles. We also note that the signal events often have peak structure in these functions. The backgrounds, on the other hand, are not expected to have peak structure in general and should be falling rapidly above the kinematic cut. A cut around the peak of the signal region may increase the signal-to-background ratio which could help the discovery and/or the follow-up signal analysis.

In conclusion, we have studied how to identify different event topologies with dark matter particles produced in pairs from cascade decays of heavy particles. Setting $4j + \cancel{E}_T$ as a case study, we have shown that one can identify all event topologies based on the existence of end points of several functions of invariant mass distributions defined in this paper. We have also extended our studies to include the cases with ISR and off-shell particles in the decay chains. It is found that most of those topologies can be identified with $\mathcal{O}(10^3)$ signal events after basic kinematic cuts. We believe that this study should be the *first* step towards measuring the masses and spins of the dark matter particles, and similar studies should be performed for events with other final states.

Acknowledgments

We would like to thank Jared Kaplan, Joe Lykken, Michael Peskin and Jay Wacker for useful discussions and comments. H.-C. C. thanks the hospitality of Fermilab where this work was initiated. H.-C. C. is supported by the Department of Energy Grant DE-FG02-91ER40674. SLAC is operated

by Stanford University for the US Department of Energy under contract DE-AC02-76SF00515.

A Invariant Mass End Point Formulae

In this Appendix we summarize the formulae for various (up to four particles) invariant mass end points in a decay chain, with either on-shell or off-shell intermediate particles. We start with a general consideration. Consider that a mother particle M goes through cascade decays by emitting several visible particles and end up with a single missing particle A as shown in Fig. 11. Denote the total



Figure 11: The general Feynman diagram with a mother particle M cascade decays into visible particles and a single missing particle A .

momentum of all visible particles by p_V , the invariant mass-squared of all visible particles is given by

$$\begin{aligned}
 p_V^2 &= (p_M - p_A)^2 = m_M^2 + m_A^2 - 2p_M p_A \\
 &\rightarrow m_M^2 + m_A^2 - 2m_M E_A, \quad \text{in the rest frame of the mother particle } M. \quad (9)
 \end{aligned}$$

We see that the maximum of p_V^2 occurs when E_A is minimized in the rest frame of M . In particular, if it is possible to make A at rest in the M rest frame, then $E_{A,\min} = m_A$ and $p_{V,\max}^2 = (m_M - m_A)^2$. However, it is not always possible to make A at rest if one of the on-shell decay produces a large boost in some direction which can not be compensated by the boosts in the opposite direction from other stages of the decays. In this case, to reach the maximum of p_V^2 we still would like to have the other decays to boost A in the opposite direction from the largest-boost decay to minimize E_A in the M rest frame. We will see some examples in the invariant mass end point formulae.

Many invariant mass end point formulae (up to three particles) can be found in the literature [7, 8, 9, 10, 11, 12, 13]. For our study, we extend the formulae to any invariant mass combinations in a decay chain with four visible particles, shown in Fig. 12. For simplicity we assume that all visible particles are massless, $p_1^2 = p_2^2 = p_3^2 = p_4^2 = 0$, which is a good approximation for most cases. The intermediate particles D, C, B may be off shell. The modifications of the end point formulae for off-shell decays

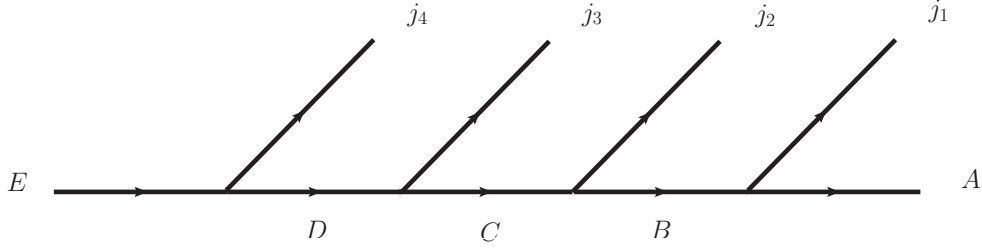


Figure 12: The Feynman diagram with a mother particle E cascade decays into four jets j_i and a single missing particle A .

will be remarked when it is relevant. As a rule of thumb, when an end point formula contains a mass parameter explicitly, it applies to the case when the corresponding intermediate particle is on shell. If a mass parameter does not appear in an end point formula, then the formula applies for either that particle being on shell or off shell. We follow the notations in Ref. [13] by defining $R_{ij} \equiv m_i^2/m_j^2$, where $i, j = A, B, C, D, E$.

A.1 Two-particle invariant masses

1. $m_{12,\max}^2$:

The extremal configuration is shown in Fig. 13. The invariant mass end point for on-shell B is

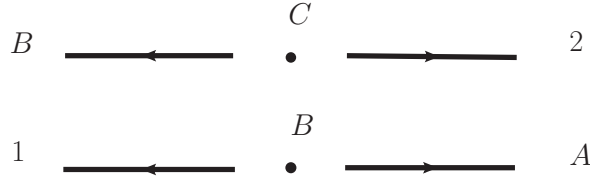


Figure 13: The extremal configuration for $m_{12,\max}^2$.

given by the well-known formula

$$m_{12,\max}^2 = \frac{(m_C^2 - m_B^2)(m_B^2 - m_A^2)}{m_B^2} = m_C^2(1 - R_{BC})(1 - R_{AB}). \quad (10)$$

In the case that B is off shell, it is possible to make A at rest in the C rest frame, so

$$m_{12,\max}^2 = (m_C - m_A)^2 = m_C^2(1 - \sqrt{R_{AC}})^2, \quad \text{if } B \text{ is off shell.} \quad (11)$$

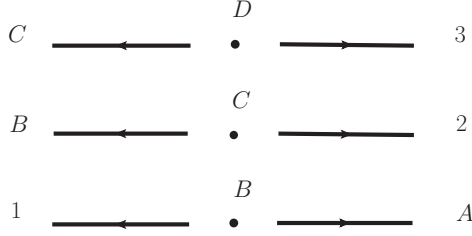


Figure 14: The extremal configuration for $m_{13,\max}^2$.

2. $m_{13,\max}^2$:

The extremal configuration for on-shell decays is shown in Fig. 14. The corresponding invariant mass end point is

$$m_{13,\max}^2 = \frac{(m_D^2 - m_C^2)(m_B^2 - m_A^2)}{m_B^2} = m_D^2(1 - R_{CD})(1 - R_{AB}). \quad (12)$$

If C or B is off shell, then the invariant mass end point occurs in the soft limit of visible particle 2, $p_2 \rightarrow 0$. The invariant mass end point formula can be obtained in the same way as m_{12}^2 ,

$$m_{13,\max}^2 = \begin{cases} m_D^2(1 - R_{BD})(1 - R_{AB}), & \text{if } C \text{ is off shell,} \\ m_D^2(1 - R_{CD})(1 - R_{AC}), & \text{if } B \text{ is off shell.} \end{cases} \quad (13)$$

However, unless the initial particle D is highly boosted, the particle 2 is likely to be too soft to pass the cut in this limit. The end point distribution will not be very sharp on an event sample which includes a hard-enough particle 2. The formula only works as an upper bound in practice and the actual end point may be much smaller depending on the jet E_T cut. Similar consideration also applies to other cases where only one of the two visible particles coming from an off-shell decay is included in the invariant mass calculation. The invariant mass maximum occurs in the soft limit of the other visible particle from that decay. The invariant mass end point is the same as the case with that soft visible particle removed, though in practice it will be smaller and not be as sharp.

3. $m_{14,\max}^2$:

The extremal configuration for on-shell decays is shown in Fig. 15. The corresponding invariant mass end point is

$$m_{14,\max}^2 = m_E^2(1 - R_{DE})(1 - R_{AB}). \quad (14)$$

When some of the D, C, B intermediate particles are off shell, $m_{14,\max}^2$ occurs in the soft limit of

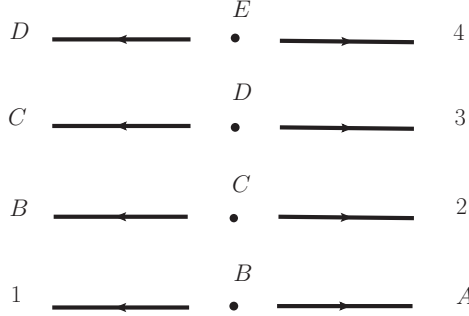


Figure 15: The extremal configuration for $m_{14,\max}^2$.

particle 2 or 3. The configurations become equivalent to the previous cases and one can use the results in these cases with replacements of the appropriate masses.

A.2 Three-particle invariant masses

1. $m_{123,\max}^2$:

First we consider that all three decays are on shell. If A can be put at rest in the D rest frame by combining three boosts, then the end point formula is simply given by $(m_D - m_A)^2$. On the other hand, if one of the decays gives a large boost which can not be compensated by the other two boosts, the end point occurs when the two smaller boosts are in the opposite direction of the largest boost. If one of the intermediate particle is off shell, the two visible particles coming from that three-body decay can have a boost ranging from 0 to the maximal value. Therefore, A can not be put at rest only if the other on-shell decay gives a boost larger than the maximum boost from the three-body decay.

- (a) If $R_{CD} < R_{AC}$, the boost from $D \rightarrow C$ decay can not be compensated by the boost from $C \rightarrow A$ (irrespective of whether B is on shell or off shell). The end point of m_{123}^2 is given by

$$m_{123,\max}^2 = m_D^2(1 - R_{CD})(1 - R_{AC}). \quad (15)$$

- (b) If $R_{BC} < R_{AB}R_{CD}$, the largest boost comes from $C \rightarrow B$ decay and the end point of m_{123}^2 is given by

$$m_{123,\max}^2 = m_D^2(1 - R_{BC})(1 - R_{AB}R_{CD}). \quad (16)$$

- (c) If $R_{AB} < R_{BD}$, irrespective of whether C is on shell or not, the end point of m_{123}^2 is given

by

$$m_{123,\max}^2 = m_D^2(1 - R_{AB})(1 - R_{BD}). \quad (17)$$

- (d) In other cases when there is no single large boost and A can be put at rest in the D rest frame, the end point of m_{123}^2 is given by the standard formula,

$$m_{123,\max}^2 = m_D^2(1 - \sqrt{R_{AD}})^2. \quad (18)$$

2. $m_{124,\max}^2$:

If C or D is off-shell, the end point occurs in the soft limit of visible particle 3. Then it

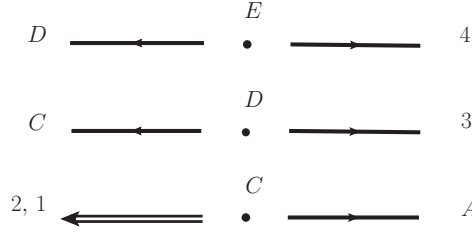


Figure 16: The general extremal configuration for $m_{124,\max}^2$.

reduces to the previous case (case 1) and the end point formulae can be easily obtained by replacing the appropriate masses. In the following discussion we assume that both C and D are on shell. The general extremal configuration is shown in Fig. 16. The double arrow represents the total momentum of particles 1 and 2, while individual particles may go in different directions depending on the relative mass parameters.

- (a) If $R_{AC}(1 - R_{DE} + R_{CE}) > R_{CE}$, the extremal configuration is that particles 1 and 2 move in the same direction, which is opposite to particles 3 and 4. The end point of m_{124}^2 in this case is given by

$$m_{124,\max}^2 = m_E^2(1 - R_{DE})(1 - R_{AC}). \quad (19)$$

The intermediate particle B can be either on shell or off shell, while other intermediate particles C and D need to be on shell.

- (b) If all intermediate particles (B, C, D) are on shell and $R_{BC}^2(1 - R_{DE} + R_{CE}) > R_{AE}$, and if $R_{AB} > R_{BC}$, the extremal configuration is that particles 1, 3 and 4 move in the same direction which is opposite to particle 2. The end point of m_{124}^2 in this case is given by

$$m_{124,\max}^2 = m_E^2(1 - R_{BC})[1 - R_{DE} + R_{CE}(1 - R_{AB})]. \quad (20)$$

- (c) If all intermediate particles are on shell and $R_{AB}(1 - R_{DE} + R_{CE}) > R_{BE}$, and if $R_{AB} < R_{BC}$, the extremal configuration is that particles 2, 3 and 4 move in the same direction which is opposite to particle 1. The end point of m_{124}^2 in this case is given by

$$m_{124,\max}^2 = m_E^2(1 - R_{AB})[1 - R_{DE} + R_{CE}(1 - R_{BC})]. \quad (21)$$

- (d) In other cases, particles 1, 2 and 4 are not collinear in the extremal configuration. The end point is given by

$$m_{124,\max}^2 = m_E^2(\sqrt{1 - R_{DE} + R_{CE}} - \sqrt{R_{AE}})^2. \quad (22)$$

3. $m_{134,\max}^2$:

If B or C is off-shell, the end point occurs in the soft limit of visible particle 2. It also reduces

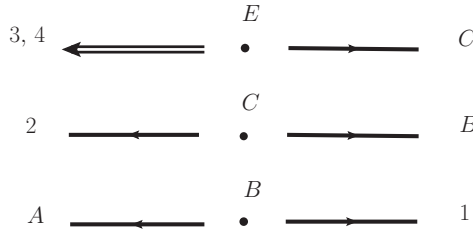


Figure 17: The general extremal configuration for $m_{134,\max}^2$.

to the case 1 and the end point formulae can be easily obtained by replacing the appropriate masses. In the following discussion we assume that both B and C are on shell. The general extremal configuration is shown in Fig. 17. The double arrow represents the total momentum of particles 3 and 4, while individual particles may go in different directions depending on the relative mass parameters.

- (a) If $\sqrt{R_{BE}}(1 - R_{AB}) > (1 - R_{CE})$, the extremal configuration is that particles 2, 3 and 4 move in the same direction, which is opposite to particle 1. This is independent whether D is on shell or not. The end point of m_{134}^2 in this case is given by

$$m_{134,\max}^2 = m_E^2(1 - R_{CE})(1 - R_{AB}). \quad (23)$$

- (b) If all intermediate particles are on shell and $\sqrt{R_{BE}}(1 - R_{AB}) < |R_{CD} - R_{DE}|$, and if $R_{CD} > R_{DE}$, the extremal configuration is that particles 1 and 3 move in the same direction, which is opposite to particles 2 and 4. The end point of m_{134}^2 in this case is given by

$$m_{134,\max}^2 = m_E^2(1 - R_{DE})(1 - R_{CD}R_{AB}). \quad (24)$$

- (c) If all intermediate particles are on shell and $\sqrt{R_{BE}}(1 - R_{AB}) < |R_{CD} - R_{DE}|$, and if $R_{CD} < R_{DE}$, the extremal configuration is that particles 1 and 4 move in the same direction, which is opposite to particles 2 and 3. The end point of m_{134}^2 in this case is given by

$$m_{134,\max}^2 = m_E^2(1 - R_{CD})(1 - R_{DE}R_{AB}). \quad (25)$$

- (d) In other cases, particles 1, 3 and 4 are not collinear in the extremal configuration. The end point is given by

$$m_{134,\max}^2 = m_E^2(1 - \sqrt{R_{CE}R_{AB}})^2. \quad (26)$$

A.3 Four-particle invariant masses

1. $m_{1234,\max}^2$:

Again, the end point depends on whether there is a large boost from one of the on-shell decays which can not be compensated by the combined boost from other decays.

- (a) If $R_{DE} < R_{AD}$ (B, C can be on shell or off shell), the end point is given by

$$m_{1234,\max}^2 = m_E^2(1 - R_{DE})(1 - R_{AD}). \quad (27)$$

- (b) If $R_{CD} < R_{AC}R_{DE}$ (B can be on shell or off shell), the end point is given by

$$m_{1234,\max}^2 = m_E^2(1 - R_{CD})(1 - R_{AC}R_{DE}). \quad (28)$$

- (c) If $R_{BC} < R_{AB}R_{CE}$ (D can be on shell or off shell), the end point is given by

$$m_{1234,\max}^2 = m_E^2(1 - R_{BC})(1 - R_{AB}R_{CE}). \quad (29)$$

- (d) If $R_{AB} < R_{BE}$ (C, D can be on shell or off shell), the end point is given by

$$m_{1234,\max}^2 = m_E^2(1 - R_{AB})(1 - R_{BE}). \quad (30)$$

- (e) In other cases when there is no single large boost and A can be made at rest in the E rest frame, the end point of m_{1234}^2 is given by the standard formula,

$$m_{1234,\max}^2 = m_E^2(1 - \sqrt{R_{AE}})^2. \quad (31)$$

A.4 End point formulae for functions F_1, F_2, F_3, F_4

Now we can write down the formulae for the end points (if they exist) for the functions F_1, F_2, F_3, F_4 defined in Section 2 for various topologies.

For the $4\oplus 0$ topology, the invariant mass end points come from the decay chain with 4 jets. It has exactly the topology in Fig. 12. The end points for F_1-F_4 are

$$F_{1,\max} = \sqrt{m_{1234,\max}^2}, \quad (32)$$

$$F_{2,\max} = \min \left\{ \sqrt{m_{123,\max}^2}, \sqrt{m_{124,\max}^2}, \sqrt{m_{134,\max}^2}, \sqrt{m_{234,\max}^2} \right\}, \quad (33)$$

$$F_{3,\max} = \sqrt{m_{ij,\max}^2}, \text{ with } \max \left\{ \bigcup_{m,n} m_{mn,\max}^2 \right\} = m_{kl,\max}^2 \text{ and } \epsilon^{ijkl} \neq 0, \quad (34)$$

$$F_{4,\max} = \min \left\{ \bigcup_{i,j} \max \left(\sqrt{m_{ij,\max}^2}, \sqrt{m_{kl,\max}^2} \right) \right\} \text{ for } \epsilon^{kl ij} \neq 0. \quad (35)$$

For the $3\oplus 1$ and $3\oplus 0\oplus \text{ISR}$ topologies, the invariant mass end points come from the decay chain with 3 jets, which are labeled as 1, 2, 3. The end points for F_1-F_4 are

$$F_{1,\max} : \text{ no end point}, \quad (36)$$

$$F_{2,\max} = \sqrt{m_{123,\max}^2}, \quad (37)$$

$$F_{3,\max} \leq \max \left\{ \sqrt{m_{12,\max}^2}, \sqrt{m_{13,\max}^2}, \sqrt{m_{23,\max}^2} \right\}, \quad (38)$$

$$F_{4,\max} : \text{ no end point}. \quad (39)$$

For the three visible particles from the same decay chain, the maximum of $\sqrt{m_{12,\max}^2}, \sqrt{m_{13,\max}^2}, \sqrt{m_{23,\max}^2}$ occurs when two of the three particles are parallel and the other one is anti-parallel. However, the definition of F_3 will take the more energetic one away from the 2 parallel particles to pair with the particle from the other decay chain, so the actual $F_{3,\max}$ will in general be smaller than the above formula.

For the $2\oplus 2$ topology, the invariant mass end points can come from either decay chain. We assume that the two decay chains are symmetric and label the two jets on the same chain as 1 and 2. The end points for F_1-F_4 are

$$F_{1,\max} : \text{ no end point}, \quad (40)$$

$$F_{2,\max} : \text{ no end point}, \quad (41)$$

$$F_{3,\max} : \text{ no end point}, \quad (42)$$

$$F_{4,\max} = \sqrt{m_{12,\max}^2}. \quad (43)$$

The $2\oplus 1\oplus \text{ISR}$ topology is not expected to have a sharp end point in any of F_1 – F_4 functions.

For the parameters used in Sec. 2, $m_A = 200$ GeV, $m_B = 400$ GeV, $m_C = 600$ GeV, $m_D = 800$ GeV, and $m_E = 1000$ GeV, the various end points (in GeV) are listed in Table 7.

Topology	F_1	F_2	F_3	F_4
$4\oplus 0$	800	600	394	397
$3\oplus 1$	–	600	< 458	–
$2\oplus 2$	–	–	–	387

Table 7: Theoretical predictions of the end points in terms of the functions F_i for the spectrum chosen in Sec. 2. All numbers are in GeV.

B Identifying the $2\oplus 1\oplus \text{ISR}$ topology

To identify the $2\oplus 1\oplus \text{ISR}$ topology, we need to find a way to pick out the ISR jet and distinguish it from the other cases without ISR. The ISR jets should have a falling E_T distribution, while other jets from heavy particle on-shell decays should have an E_T distribution with peak structure if the peak is above the cut. In order to keep the peak structure of those energetic jets, the cut on the jet E_T can not be too strong. Therefore, in this section, we impose a softer cut on the basic kinematics: at least 4 jets with $E_T > 50$ GeV and missing transverse energy $\cancel{E}_T > 100$ GeV.

To increase the probability of finding the ISR jet for the $2\oplus 1\oplus \text{ISR}$ topology, we only choose those events in which the two pairs with the largest invariant masses contain the same jet. We then plot the E_T distribution of the remaining one which does not appear in the two largest invariant masses. We use the following function to describe this procedure

$$F_6(p_1, p_2, p_3, p_4) = E_T(j_l),$$

such that $\epsilon^{ijkl} \neq 0$ and $m_{ij}, m_{ik} =$ two largest invariant masses. (44)

The event distributions for different topologies in terms of F_6 are shown in Fig. 18. By choosing a proper bin size, we can see from Fig. 18 that the two topologies with ISR have the largest number of events at the first bin, which can be used to distinguish them from topologies without ISR.

To further distinguish $2\oplus 1\oplus \text{ISR}$ from $3\oplus 0\oplus \text{ISR}$, we can simply check their patterns of the parameters of the first four functions. If $b/(ac)(F_2)$ is smaller than the other values, we can then identify the

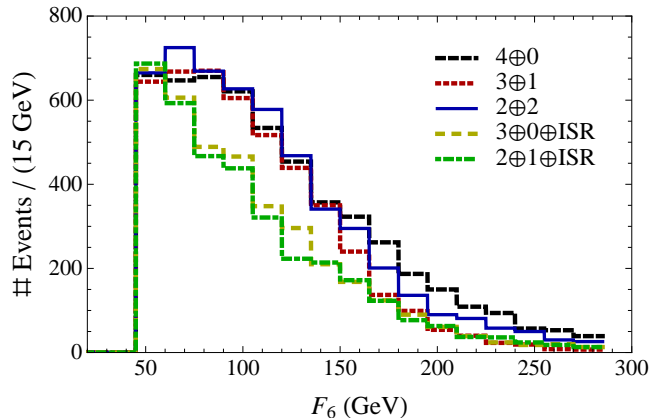


Figure 18: The distributions of F_6 for different topologies. We require at least 4 jets with $E_T > 50$ GeV and missing transverse energy $\cancel{E}_T > 100$ GeV.

Topologies	bin size = 40 GeV				bin size = 50 GeV			
	$\frac{b}{ac}(F_1)$	$\frac{b}{ac}(F_2)$	$\frac{b}{ac}(F_3)$	$\frac{b}{ac}(F_4)$	$\frac{b}{ac}(F_1)$	$\frac{b}{ac}(F_2)$	$\frac{b}{ac}(F_3)$	$\frac{b}{ac}(F_4)$
$3\oplus 0\oplus \text{ISR}$	0.99	0.81	1.58	0.98	1.00	0.81	1.65	0.97
$2\oplus 1\oplus \text{ISR}$	1.11	1.37	1.54	1.27	1.20	1.37	2.20	1.35

Table 8: The fitted values of $b/(ac)$, which determine the normalized steepness of the slopes of the histograms after the peak. The broken-line function in Eq. (7) is used to fit distributions. The bin size has been chosen to be 40 GeV (left), and 50 GeV (right). Again each topology has 10000 events before the cuts, and there are 3841 and 3632 events passing the cuts for $3\oplus 0\oplus \text{ISR}$ and $2\oplus 1\oplus \text{ISR}$, respectively.

topology as $3\oplus 0\oplus \text{ISR}$. On the other hand for the $2\oplus 1\oplus \text{ISR}$ topology, none of the normalized inverse slopes exhibit a particularly small value and $b/(ac)(F_2)$ is not significantly smaller than others. In this case the topology can be $2\oplus 1\oplus \text{ISR}$ or with even more ISRs,

References

- [1] S. Dimopoulos and H. Georgi, *Softly Broken Supersymmetry and SU(5)*, *Nucl. Phys.* **B193** (1981) 150.
- [2] T. Appelquist, H.-C. Cheng, and B. A. Dobrescu, *Bounds on universal extra dimensions*, *Phys.Rev.* **D64** (2001) 035002, [[hep-ph/0012100](#)].

- [3] H.-C. Cheng, K. T. Matchev, and M. Schmaltz, *Radiative corrections to Kaluza-Klein masses*, *Phys. Rev.* **D66** (2002) 036005, [[hep-ph/0204342](#)].
- [4] H.-C. Cheng, K. T. Matchev, and M. Schmaltz, *Bosonic supersymmetry? Getting fooled at the CERN LHC*, *Phys. Rev.* **D66** (2002) 056006, [[hep-ph/0205314](#)].
- [5] A. J. Barr and C. G. Lester, *A Review of the Mass Measurement Techniques proposed for the Large Hadron Collider*, *J.Phys.G* **G37** (2010) 123001, [[arXiv:1004.2732](#)].
- [6] F. E. Paige, *Determining SUSY particle masses at LHC*, [hep-ph/9609373](#).
- [7] I. Hinchliffe, F. Paige, M. Shapiro, J. Soderqvist, and W. Yao, *Precision SUSY measurements at CERN LHC*, *Phys.Rev.* **D55** (1997) 5520–5540, [[hep-ph/9610544](#)].
- [8] B. Allanach, C. Lester, M. A. Parker, and B. Webber, *Measuring sparticle masses in nonuniversal string inspired models at the LHC*, *JHEP* **0009** (2000) 004, [[hep-ph/0007009](#)].
- [9] B. K. Gjelsten, D. J. Miller, 2, and P. Osland, *Measurement of SUSY masses via cascade decays for SPS 1a*, *JHEP* **12** (2004) 003, [[hep-ph/0410303](#)].
- [10] B. K. Gjelsten, D. J. Miller, 2, and P. Osland, *Measurement of the gluino mass via cascade decays for SPS 1a*, *JHEP* **06** (2005) 015, [[hep-ph/0501033](#)].
- [11] D. J. Miller, 2, P. Osland, and A. R. Raklev, *Invariant mass distributions in cascade decays*, *JHEP* **03** (2006) 034, [[hep-ph/0510356](#)].
- [12] M. Burns, K. T. Matchev, and M. Park, *Using kinematic boundary lines for particle mass measurements and disambiguation in SUSY-like events with missing energy*, *JHEP* **0905** (2009) 094, [[arXiv:0903.4371](#)].
- [13] K. T. Matchev, F. Moortgat, L. Pape, and M. Park, *Precise reconstruction of sparticle masses without ambiguities*, *JHEP* **0908** (2009) 104, [[arXiv:0906.2417](#)].
- [14] C. Lester and D. Summers, *Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders*, *Phys.Lett.* **B463** (1999) 99–103, [[hep-ph/9906349](#)].
- [15] A. Barr, C. Lester, and P. Stephens, *$m(T_2)$: The Truth behind the glamour*, *J.Phys.G* **G29** (2003) 2343–2363, [[hep-ph/0304226](#)].

- [16] C. Lester and A. Barr, *MTGEN: Mass scale measurements in pair-production at colliders*, *JHEP* **0712** (2007) 102, [[arXiv:0708.1028](#)].
- [17] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, *Gluino Stransverse Mass*, *Phys.Rev.Lett.* **100** (2008) 171801, [[arXiv:0709.0288](#)].
- [18] A. J. Barr, B. Gripaios, and C. G. Lester, *Weighing Wimps with Kinks at Colliders: Invisible Particle Mass Measurements from Endpoints*, *JHEP* **02** (2008) 014, [[arXiv:0711.4008](#)].
- [19] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, *Measuring superparticle masses at hadron collider using the transverse mass kink*, *JHEP* **02** (2008) 035, [[arXiv:0711.4526](#)].
- [20] H.-C. Cheng and Z. Han, *Minimal Kinematic Constraints and $m(T_2)$* , *JHEP* **0812** (2008) 063, [[arXiv:0810.5178](#)].
- [21] M. M. Nojiri, Y. Shimizu, S. Okada, and K. Kawagoe, *Inclusive transverse mass analysis for squark and gluino mass determination*, *JHEP* **06** (2008) 035, [[arXiv:0802.2412](#)].
- [22] M. Burns, K. Kong, K. T. Matchev, and M. Park, *Using Subsystem MT_2 for Complete Mass Determinations in Decay Chains with Missing Energy at Hadron Colliders*, *JHEP* **0903** (2009) 143, [[arXiv:0810.5576](#)].
- [23] Y. Bai and Z. Han, *Measuring the Dark Force at the LHC*, *Phys.Rev.Lett.* **103** (2009) 051801, [[arXiv:0902.0006](#)].
- [24] J. Alwall, K. Hiramatsu, M. M. Nojiri, and Y. Shimizu, *Novel reconstruction technique for New Physics processes with initial state radiation*, *Phys.Rev.Lett.* **103** (2009) 151802, [[arXiv:0905.1201](#)].
- [25] P. Konar, K. Kong, K. T. Matchev, and M. Park, *Superpartner Mass Measurement Technique using 1D Orthogonal Decompositions of the Cambridge Transverse Mass Variable MT_2* , *Phys.Rev.Lett.* **105** (2010) 051802, [[arXiv:0910.3679](#)].
- [26] P. Konar, K. Kong, K. T. Matchev, and M. Park, *Dark Matter Particle Spectroscopy at the LHC: Generalizing $M(T_2)$ to Asymmetric Event Topologies*, *JHEP* **1004** (2010) 086, [[arXiv:0911.4126](#)].
- [27] K. Kawagoe, M. M. Nojiri, and G. Polesello, *A new SUSY mass reconstruction method at the CERN LHC*, *Phys. Rev.* **D71** (2005) 035008, [[hep-ph/0410160](#)].

- [28] H.-C. Cheng, J. F. Gunion, Z. Han, G. Marandella, and B. McElrath, *Mass determination in SUSY-like events with missing energy*, *JHEP* **0712** (2007) 076, [arXiv:0707.0030].
- [29] H.-C. Cheng, D. Engelhardt, J. F. Gunion, Z. Han, and B. McElrath, *Accurate Mass Determinations in Decay Chains with Missing Energy*, *Phys.Rev.Lett.* **100** (2008) 252001, [arXiv:0802.4290].
- [30] H.-C. Cheng, J. F. Gunion, Z. Han, and B. McElrath, *Accurate Mass Determinations in Decay Chains with Missing Energy. II*, *Phys.Rev.* **D80** (2009) 035020, [arXiv:0905.1344].
- [31] M. M. Nojiri, K. Sakurai, and B. R. Webber, *Reconstructing particle masses from pairs of decay chains*, *JHEP* **06** (2010) 069, [arXiv:1005.2532].
- [32] E. Izaguirre, M. Manhart, and J. G. Wacker, *Bigger, Better, Faster, More at the LHC*, arXiv:1003.3886.
- [33] D. S. Alves, E. Izaguirre, and J. G. Wacker, *It's On: Early Interpretations of ATLAS Results in Jets and Missing Energy Searches*, arXiv:1008.0407.
- [34] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, *et. al.*, *MadGraph/MadEvent v4: The New Web Generation*, *JHEP* **0709** (2007) 028, [0706.2334].
- [35] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, *et. al.*, *New generation of parton distributions with uncertainties from global QCD analysis*, *JHEP* **0207** (2002) 012, [hep-ph/0201195].
- [36] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **0605** (2006) 026, [hep-ph/0603175].
- [37] J. S. Conway, *Pretty Good Simulation of high-energy collisions, 090401 release* (2009). <http://physics.ucdavis.edu/~conway/research/software/pgs/pgs4-general.htm>.
- [38] P. Meade and M. Reece, *BRIDGE: Branching ratio inquiry / decay generated events*, hep-ph/0703031.
- [39] D. R. Tovey, *On measuring the masses of pair-produced semi-invisibly decaying particles at hadron colliders*, *JHEP* **0804** (2008) 034, [arXiv:0802.2879].
- [40] H.-C. Cheng and I. Low, *TeV symmetry and the little hierarchy problem*, *JHEP* **09** (2003) 051, [hep-ph/0308199].

- [41] H.-C. Cheng and I. Low, *Little hierarchy, little Higgses, and a little symmetry*, *JHEP* **08** (2004) 061, [[hep-ph/0405243](#)].
- [42] I. Low, *T parity and the littlest Higgs*, *JHEP* **10** (2004) 067, [[hep-ph/0409025](#)].
- [43] Y. Bai and R. J. Hill, *Weakly Interacting Stable Pions*, [arXiv:1005.0008](#).
- [44] K. Agashe, D. Kim, M. Toharia, and D. G. Walker, *Distinguishing Dark Matter Stabilization Symmetries Using Multiple Kinematic Edges and Cusps*, *Phys.Rev.* **D82** (2010) 015007, [[arXiv:1003.0899](#)].
- [45] L.-T. Wang and I. Yavin, *Spin Measurements in Cascade Decays at the LHC*, *JHEP* **04** (2007) 032, [[hep-ph/0605296](#)].
- [46] J. Hubisz, J. Lykken, M. Pierini, and M. Spiropulu, *Missing energy look-alikes with 100 pb^{-1} at the LHC*, *Phys. Rev.* **D78** (2008) 075008, [[arXiv:0805.2398](#)].
- [47] G. Hallenbeck, M. Perelstein, C. Spethmann, J. Thom, and J. Vaughan, *Model Discrimination with the CMS Detector: A Case Study*, *Phys.Rev.* **D79** (2009) 075024, [[arXiv:0812.3135](#)].
- [48] M. Blanke, D. Curtin, and M. Perelstein, *SUSY-Yukawa Sum Rule at the LHC*, *Phys. Rev.* **D82** (2010) 035020, [[arXiv:1004.5350](#)].
- [49] A. Rajaraman and F. Yu, *A New Method for Resolving Combinatorial Ambiguities at Hadron Colliders*, [arXiv:1009.2751](#).