Novel Networks in Collider Searches for New Physics

Anna Jane Mullin

Supervisors: A/Prof. M.J. White
Prof. P.D. Jackson

School of Physical Sciences
University of Adelaide

This dissertation is submitted for the degree of
Master of Philosophy

April 2021

Table of contents

Abstract

Beyond-Standard Model (BSM) physics searches at the LHC are limited by the amount of information available to distinguish a new physics process from its backgrounds. Analyses apply a range of classification algorithms to obtain sensitivity to rare signals, but are challenged to obtain enough information in a broad parameter space without relying on heavy optimisation in narrow search regions. LHC event classification techniques become more powerful when they can be applied broadly to diverse models, requiring a large number of independent variables sensitive to anomalous signals. In our prototype ATLAS search, we create new variables that target information not used in current methods. Whereas typical variables treat events in isolation, we obtain further discrimination from the “similarity” between event pairs by evaluating “distances” in a kinematic space. A map of event similarities forms a graph network, which provides a convenient range of network variables able to quantify local topologies. In networks constructed from nodes of LHC events, we aim to use network variables to increase sensitivity to anomalous topologies local to BSM events. Our proof-of-principle analysis reveals that BSM physics events may populate distinct distributions compared with Standard Model events in several types of network variables, including measures of local centrality and clustering, using supersymmetry searches as examples. Graph network analysis may contribute power to existing methods of event classification and increase sensitivity to anomalous signatures.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Acknowledgements

I acknowledge the original conceptualisation of graph networks in LHC searches for new physics by my supervisor A/Prof. Martin White, who has consistently supported my progress as a research student. I am also grateful for the support and feedback of my secondary supervisor Prof. Paul Jackson, and for the members of the Department of Physics at the University of Adelaide, who continue to inspire me to be curious. Thirdly, I would like to acknowledge the many contributions to this project by our colleagues at the University of Cambridge, Prof. M.A. Parker, Dr. Sarah Williams and Dr. Holly Pacey, and by Stuart Nicholls.

Introduction

Many viable improvements to the Standard Model theory of particle physics have been proposed over recent years, offering solutions to the limitations of the current model. Although the Standard Model provides sound descriptions of all known particles and accurate predictions of their interactions, it cannot solve several remaining problems, including the nature of dark matter, gravity and the matter-antimatter asymmetry. These and other conundrums are driving beyond-Standard Model research to seek the true theory using an experimentally verifiable framework.

One proposed extension to the Standard Model is the principle of supersymmetry, which may in one of its possible forms provide a dark matter candidate, stabilise the Higgs boson mass and answer several other pressing questions. If the correct "beyond the Standard Model" scenario involves supersymmetry, then the Large Hadron Collider has a high chance of detecting direct evidence for it within the coming years, especially after undergoing planned upgrades reaching higher energy and luminosity. Searches for strong and electroweak supersymmetry both have merit: whereas strongly coupled supersymmetry is characterised by larger production rates, electroweak processes are likely to benefit from more distinctive signals and identifiable final states.

While the true extension to the current framework remains elusive, we must focus experimental efforts on searching for evidence of new physics, which has a large free parameter space even when restricted to theories of supersymmetry. Progress in these analyses requires powerful methods of isolating small signals in one of many possible forms. The logical solution is to increase sensitivity to as many scenarios as possible, driving us to develop creative new methods of isolating anomalies against the Standard Model background. Not only do we need model-independent search strategies, but we also benefit from coupling new methods with existing analysis techinques, combining their sensitivity to develop powerful, broad searches.

The work in this thesis tests a new method for beyond-Standard Model physics analyses in collider data using graph theory techniques. Graph networks have proven useful in a range of social, biological and economic applications, but within particle physics have mostly been applied in deep learning models, and have never been used to study collider topology in particle collision event networks. Wider research shows that networks are an effective tool for characterising and quantifying the extent of topological structures; the connections between elements of a dataset reveal more information about the data than we can obtain from studying the elements in isolation. Supersymmetry presents a good example on which to test this novel technique for new physics searches with the Large Hadron Collider.

This thesis evaluates the potential for using a network analysis approach in beyond-Standard Model searches, seeking to characterise the network behaviour of small signals in particle collider data using examples from the strong and electroweak sectors of supersymmetry. The analysis exploits convenient variables defined by graph theory to identify distinct network topologies, isolating unique signatures of signal events against the predicted Standard Model background. The network variables target aspects of clustering, closeness and other centrality measures, which may reveal new information depending on the scale of connections. Further information is obtained from selecting sensitive basis variables underlying the graph space, and by calculating a range of spatial distances. Exploiting these various possibilities may allow particle collider event networks to become effective methods in searches, and further apply to model-independent anomaly detection.

Chapter 1 describes the successes and limitations of the Standard Model, identifying the gaps which may be explained by theories of supersymmetry presented in Chapter 2. In Chapter 3, the Large Hadron Collider and ATLAS experiment designs are described, followed by an outline of current analysis strategies for new physics searches with ATLAS data. Chapter 4 discusses graph networks in a particle context, presenting our methods for constructing networks using collider events and their kinematic attributes. The methods then feature in several contexts throughout Chapters 5-7, first in two preliminary tests involving electroweak and strongly coupled supersymmetry in Chapters 5 and 6, and later in Chapter 7 as a test to determine whether network methods yield reliable results when constructing graphs using real ATLAS data.

Chapter 8 discusses the conclusions and future applications of network analysis to ATLAS searches. Since the completion of the work presented in Chapters 5 and 6, these results are now published in the Journal of High Energy Physics [1]. I contributed personally to the development of this project at all core stages including design and concepts, modelling and simulation, refinement and generation of results, and the writing and editing of the publication.

Chapter 1

The Standard Model

1.1 Introduction to the Standard Model

The Standard Model of particle physics (SM) is a gauge field theory describing the constituents of matter and their strong, weak and electromagnetic interactions, based on symmetry laws. Particles are allowed to interact only if the process conserves the required symmetries, so in the quantum field theory of the SM, invariance under a local symmetry transformation restricts the form of the interactions.

Since the laws of physics must never depend on the frame of an observer, all physics in the model must obey global Poincaré symmetry, consisting of translational invariance, rotational invariance and invariance under the change of inertial reference frame. A different type of symmetry, this time local not global, restricts the form of the three fundamental interactions, and is a type of internal $SU(3) \times SU(2) \times U(1)$ gauge symmetry. The SM is *gauge* invariant because it is locally invariant under transformations belonging to this product of groups.

All possible interactions can be written within the Lagrangian mathematical formalism, which provides a concise statement of the physical properties of particles in the SM. It represents matter as fields that couple to one another with different relative strengths. The Lagrangian comprises free terms for every field, including kinetic terms that describe dynamic motion, in addition to coupling terms that describe the possible interactions permitted by the symmetries and their relative strengths. Any symmetry operation belonging to the theory must leave the time integral of the Lagrangian (i.e. the action) invariant, so the terms allowed in the Lagrangian are constrained by the required symmetries.

Another requirement of the model is renormalisability, a property that relates to the highest energy processes the theory can possibly describe. We can define D as a distance cutoff such that effects involving fields varying on shorter distance scales than D are ignored, and consider the continuum limit as $D \rightarrow 0$ at some momentum scale p. If the SM is renormalisable, then it can describe the physical values of constant parameters in the high energy limit (at small distance scales) using only a finite number of bare coupling parameters. To satisfy this requirement, all terms in the Lagrangian must be renormalisable.

The quantum field theory of the SM is the culmination of decades of international effort and careful reasoning. Subatomic physics theory in the 20th century progressively unified explanations of matter and interactions until the current SM was achieved, unifying the electromagnetic and weak forces but leaving the strong force in a separate sector. Theoretical calculations can now predict many experimental measurements to high precision, verifying the model at energies up to those probed by the LHC.

1.2 Symmetry groups and particle content

The symmetries driving the SM are based on transformations that conserve some properties that we call the quantum numbers of the particles. If the SM has a symmetry under a particular transformation, then a type of quantum number is conserved. This is a consequence of Noether's theorem, which states that a continuous symmetry implies that a conserved current exists with an associated conserved charge [2].

The SM is split into two sectors: the electroweak sector and the strong sector. The theory is invariant under transformations in the subgroup $SU(2)_L \times U(1)_Y$, which describes the electroweak symmetry group, and the $SU(3)_C$ gauge theory, which describes strong interactions belonging to the theory of quantum chromodynamics (QCD). The Lagrangian can be written as the simple sum of terms belonging separately to the electroweak and QCD sectors: $L_{\text{SM}} = L_{\text{EW}} + L_{\text{QCD}}$. The subscripts C, L and Y refer respectively to colour, left chirality and weak hypercharge, and are properties of the subsets of matter that interact in each case.

In any interaction, matter particles called fermions interact with force-carrying particles called bosons, which are shown in blue in Figure 1.1.

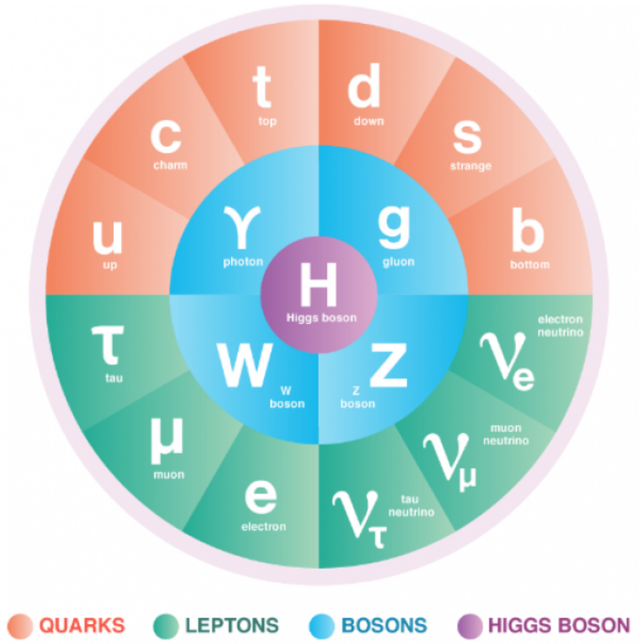


Fig. 1.1 Table of particles in the Standard Model [3].

Matter comprises quarks and leptons, and the SM categorises these as fermion fields according to their spin statistics, which are Fermi-Dirac statistics due to their half-integer spin. Bosons instead obey Bose-Einstein statistics, have integer spin values, and mediate the fermion interactions. If the integer spin of a particle is zero, it is called a scalar boson, whereas a spin-1 particle is called a vector boson. The Lagrangian contains all the types of renormalisable interactions of four mass dimensions between scalar, fermion and vector fields which are invariant under transformations belonging to the local symmetry group of the SM.

The photon (γ), Z, $W^{\pm}$ and gluons are all force-carrying bosons. The $W^{\pm}$ states are responsible for carrying the weak charge-current interactions, and the neutral Z mediates the neutral-current weak interactions, while gluons mediate the strong interactions. The final boson is the Higgs particle, the most recent additional SM particle measured experimentally. The Higgs boson was discovered at a mass of 125.36 ± 0.41 GeV by the ATLAS [4] and CMS [5] experiments some decades after it was introduced in the theory by Brout, Englert and Higgs to give mass to the fermions and weak bosons via the Higgs mechanism, described in Section 1.6. The discovery confirmed the picture of electroweak symmetry breaking in the SM, and added the final element to complete the current theory. Since the Higgs boson was confirmed, all experimental observables in the SM can now be calculated from

SM parameters, although some (such as the anomalous magnetic moment of the muon) are non-perturbative and remain difficult to calculate.

1.3 Constructing the Standard Model

The fermions and bosons are combined in a gauge invariant quantum field theory that describes their properties and interactions using symmetries. The symmetry groups are classified either as Abelian or non-Abelian, according to whether their elements commute or do not commute. Developing a free fermion field theory into an interacting model via these two types of continuous gauge groups, we find that particle content comes out differently from Abelian and non-Abelian theories. In the SM, the theory of electrodynamics governed by symmetry under transformations belonging to the $U(1)_Q$ subgroup is an Abelian theory, whereas the weak and strong interactions come from non-Abelian subgroups. The different nature and description of the bosons involved in these interactions is accurately developed by constructing the theories starting from the same free fermion fields with these differences in mind.

1.3.1 Quantum electrodynamics

The need to combine Maxwell's equations of electromagnetism with quantum field theory led to the development of quantum electrodynamics (QED) in the 1940s, a gauge theory that now forms part of the Standard Model. The principle of symmetry by gauge invariance was already present in Maxwell's theory, and is developed in QED into a symmetry by local phase invariance. Gauge invariance under local transformations of fermion fields with the $U(1)$ group is equivalent to invariance under phase rotations at every point x in spacetime such that every point in the space is phase invariant. The transformations proceed for a field ϕ as shown in Equations 1.1 and 1.2. The phase invariance of the electromagnetic field under these transformations implies charge conservation, following from Noether's theorem.

$$\phi \to \phi^{'} = e^{iq\alpha(x)}\phi \tag{1.1}$$

$$\phi^{\dagger} \to \phi^{'\dagger} = e^{-iq\alpha(x)}\phi^{\dagger} \tag{1.2}$$

When the local phase transformations of the $U(1)$ group are applied to terms in the Lagrangian that describes the properties and dynamics of a free, complex scalar field, some terms are naively not gauge invariant. For a dynamic field, the Lagrangian contains a kinetic term that is proportional to the derivative product $(\partial_\mu \phi^\dagger)(\partial^\mu \phi)$, where the partial derivative introduces a term proportional to $\partial_\mu \alpha(x)$. This term removes the gauge invariance of the theory, and a significant step towards formalising QED involves adapting it into covariant form so the theory is invariant under these local transformations.

The covariant derivative

The Lagrangian becomes gauge invariant in this case by introducing a derivative which transforms covariantly, denoted D_μ as shown in Equation 1.3. By making the derivative covariant, we remove the term proportional to $\partial_\mu \alpha(x)$ and 'gauge' local transformations to maintain the required gauge invariance in QED. The Lagrangian consequently includes an additional vector gauge field A_μ with the property in Equation 1.4, introducing a new interaction that promotes the free theory to an interacting theory. The form of the interaction is fixed in terms of the new vector field.

$$D_\mu = \partial_\mu + iqA_\mu \tag{1.3}$$

$$A_\mu(x) \rightarrow A'_\mu = A_\mu(x) - \partial_\mu \alpha(x) \tag{1.4}$$

A gauge boson is a quantisation of the gauge field, and one boson is added to the model of QED corresponding to the generator inserted above. Generally, local gauge invariance requires the introduction of massless vector gauge bosons, whose interactions are fixed. The boson mediating the electromagnetic interaction pertaining to the A_μ field is the photon, γ, so local phase invariance in QED leads to the existence of the photon field. By this process, QED comes from a theory of free Fermi fields via the continuous, Abelian $U(1)$ gauge group. The symmetry does not permit a quadratic term in the vector field, which would not be invariant under the transformation, so the QED Lagrangian is missing a mass term of the form $A^\mu A_\mu$. Since the photon is massless, the absence of a mass term is an accurate depiction of reality.

1.3.2 The non-Abelian Standard Model theory

The electroweak and QCD theories are both non-Abelian, where the elements of the group do not commute. The combination of $SU(3) \times SU(2) \times U(1)$ groups then produces a fundamentally non-Abelian gauge theory. Developing the non-Abelian theory using the same steps that were taken to develop the Abelian theory of QED now leads to a more complicated result. In the electroweak theory, starting with the Lagrangian density of free fermions leads us to again make the derivative covariant, but now that the group structure involves multiple generators we find that after several additional steps there is now more than one vector boson added to the model. These bosons are the three $SU(2)_L$ gauge bosons $W^{1,2,3}$. Similarly, in QCD the same process applied to the non-Abelian theory of strong interactions involves more generators, where each of the eight gauge fields adds a new boson, a gluon g^{α}. An important consequence of non-Abelian symmetries is their implication that the gauge bosons may interact.

As in the Abelian theory of QED, mass terms in the non-Abelian theories of electroweak and strong interactions are not invariant under the group transformations and are therefore not allowed. The eight gluons always remain massless, but the $W^{1,2,3}$ and B are massless only in the limit of exact electroweak symmetry, and obtain mass via the Higgs mechanism described later in Section 1.6. The combination of non-Abelian gauge groups which form the SM produces a model that agrees with the experimental constraints on the masses and other properties of the measured particles.

1.4 Symmetries in the Standard Model

By unifying electromagnetic interactions and weak interactions in the electroweak theory belonging to the Glashow-Weinberg-Salam (GWS) model [6], the SM describes its three constituent forces in simpler terms using only two main symmetry subgroups. An important consideration in understanding the SM symmetry groups and forces is the chirality of the model. Particle states are classified according to their spin properties as either left-handed or right-handed fermion fields, and these states transform differently within the two gauge groups and the Lorentz group. This section considers the transformation properties resulting from symmetries in the SM subgroups.

1.4.1 Symmetries in the electroweak theory

Whereas electric charge is conserved in transformations belonging to $U(1)_Q$ electromagnetism, two different charges are conserved in the electroweak group: the quantum numbers hypercharge and weak isospin. The $U(1)_Y$ transformations act on hypercharge and $SU(2)_L$ transformations act on weak isospin.

Weak isospin is denoted T, representing the conserved charge of states interacting via the weak nuclear force, and is derived from a relation to the Pauli matrices σ_i. Particles which form part of a quantum system with an $SU(2)$ symmetry have intrinsic angular momentum which can also be described using the Pauli matrices. In the SM, all weak interactions must conserve the third component of the left chiral weak isospin T_{3L} [7]. Hypercharge instead represents the conserved charge of the $U(1)_Y$ component of the electroweak gauge group, and is denoted Y. For any fermion field, hypercharge is related to the electromagnetic charge Q and T_{3L} by the relationship: $Y = 2Q - 2T_{3L}$. When vector currents are written as projections of the left- and right-handed components into separate terms, interactions via the $U(1)_Y$ group conserve the two hypercharge values Y_L and Y_R.

The weakly-interacting fermions are typically projected as left- and right-handed components to write the SM Lagrangian more easily. Under $SU(2)_L$ in the SM, left chiral fermion fields transform as doublets so that we can construct theories that are invariant under $SU(2)$ transformations that mix the components. Right chiral fields instead form singlets, as they do not interact weakly and have $T_3 = 0$ for their third component of weak isospin. In contrast with the weak current, the electromagnetic current has the same charge for both left and right chiral components. Therefore, combining the weak and electromagnetic interactions into the same terms in the SM Lagrangian requires us to use states of definite chirality, where the weak current resembles the vector form of currents from electromagnetic interactions.

1.4.2 Symmetries in the strong theory

The non-Abelian gauge theory of QCD was constructed by Yang and Mills in 1954 [8] using local isospin symmetry transformations, comprising independent isospin rotations at every point in spacetime. The physics of the strong interactions must be invariant under $SU(2)$ transformations that mix the proton and neutron wavefunctions in isospin space. By treating nucleon wavefunctions as components of a two-component nucleon ‘isospinor’, we can construct a theory which satisfies this invariance. Furthermore, if we consider the proton and neutron as composite baryons of constituent quarks, we can extend the invariance of the strong interaction under rotations of the nucleon isospinor to include transformations on the scale of the quarks. This smaller-scale theory leads to the current model of transformations between quark flavours, which is an approximate global symmetry of strong interactions. The transformation acts on a 3-component vector in the fundamental representation of the $SU(3)_f$ group, pertaining to transformations mixing the up, down and strange quark flavours.

The QCD Lagrangian of the SM has another symmetry under transformations mixing the quark colours, involving quarks as colour triplets in the fundamental $SU(3)$ representation. Colour is simply a label for states in an $SU(3)$ representation, where states may possess one of three possible colours called “red", “green" and “blue". Whereas the $SU(3)_f$ symmetry of the QCD Lagrangian is approximate, the colour transformations yield an exact symmetry, such that the strong interactions always proceed identically regardless of which colour charges are present. The $SU(3)_c$ group must conserve the three colour charges, where colour is the Noether conserved charge of the QCD symmetry group.

1.5 Transformation properties

The properties of the particles under the symmetry transformations dictate which terms are permitted in the SM Lagrangian, where a term that is not invariant under transformations belonging to the group $SU(3)_c \times SU(2)_L \times U(1)_Y$ is not a physical process and cannot be added to the model. All permitted transformations preserve the required symmetries of the SM. This section describes these properties of the electroweak and strong transformations in further detail.

1.5.1 Electroweak transformations

The left and right chiral fermion fields transform in the electroweak gauge sector of the SM with the behaviour described in Equations 1.5 and 1.6. The relative strengths of the weak and electromagnetic forces appear in the gauge couplings of the $SU(2)_L$ and $U(1)_Y$ groups, denoted g_2 and g_Y respectively. Left-handed weak transformations are described using g_2, the gauge coupling function $\vec{\alpha}_2(x)$ and the Pauli matrices $\vec{\tau}$ acting in the weak isospin doublet representation space. The transformations under the $U(1)_Y$ group involve a different gauge coupling function $\alpha_Y(x)$ and the hypercharge Y.

$$f_L(x) \rightarrow e^{-ig_Y\alpha_Y(x)Y/2}e^{-ig_2\vec{\alpha}_2(x)\cdot\vec{\tau}/2}f_L(x) \tag{1.5}$$

$$f_R(x) \rightarrow e^{-ig_Y\alpha_Y(x)Y/2}f_R(x) \tag{1.6}$$

The summary of the fermion electroweak quantum numbers in Table 1.1 shows the left-handed components as doublets and the right-handed components as singlets. Quarks and charged leptons interact via both the weak nuclear force and electromagnetism. Neutrinos, which are massless in the SM and have no electric charge on which electromagnetic interactions can act, only interact via the weak force.

Table 1.1 SM fermion electroweak quantum numbers.

Fermion		Y	T	T_3	Q
Leptons	e^-_R	-2	0	0	-1
	$\begin{pmatrix}\nu_e\\ e^-\end{pmatrix}_L$	-1	$\frac{1}{2}$	$\frac{1}{2}$	0
		-1	$\frac{1}{2}$	$-\frac{1}{2}$	-1
Quarks	u_R	$\frac{4}{3}$	0	0	$\frac{2}{3}$
	d_R	$-\frac{2}{3}$	0	0	$-\frac{1}{3}$
	$\begin{pmatrix}u\\ d\end{pmatrix}_L$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{3}$
		$\frac{1}{3}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{3}$

1.5.2 Strong transformations

All quark (q) and lepton (l) fields have left- and right-handed components transforming by Equations 1.7 and 1.8 under the colour rotations in the QCD sector. The lepton fields are colour singlets under $SU(3)$, so leptons are unaffected by the strong interactions. Here, g_s and α_s^a are the gauge coupling and corresponding function of $SU(3)_C$, and λ^a are the Gell-Mann matrices acting in the triplet representation space.

$$q_{L,R}(x) \rightarrow e^{-g_s\alpha_s^a(x)\lambda^a/2}q_{L,R}(x) \tag{1.7}$$

$$l_{L,R}(x) \rightarrow l_{L,R}(x) \tag{1.8}$$

The strong force also transforms bosons in different ways. While the photon and $W^\pm$ and Z bosons have colour charge zero like the leptons and do not participate in strong interactions, the gluons g^α are colour charged and therefore interact via the strong force as well as mediating it.

The QCD gauge coupling sets the strength of the strong interactions, and is much larger than the electroweak coupling, leading to larger loop corrections than we find in the electroweak sector. A distinguishing characteristic of QCD interactions comes from the behaviour of the coupling function at different scales of momentum transfer, which informs our understanding of hadronic interactions at long and short distances. At large momentum (or short distances), perturbative calculations are feasible because the coupling function is small and quarks are asymptotically free. In the long distance domain, perturbative QCD is no longer possible because the coupling function grows large when momentum transfer decreases, creating the effect of quark confinement. The running of the coupling is a unique property of the strong interactions that generates interesting behaviour in hadronic interactions, where quarks are confined to exist naturally in colour-neutral hadrons and never in isolation.

1.6 Spontaneous symmetry breaking

A symmetry in a Lagrangian is *spontaneously broken* if the Lagrangian is symmetric but the symmetry is not conserved by the physical vacuum. If the Lagrangian contains an exact continuous symmetry that is not shared by the physical vacuum, then one or more massless spin-0 particles emerge, called *Goldstone bosons* [9]. If the Lagrangian instead contains a local gauge symmetry, then it must have obtained interaction gauge fields when we constructed a covariant derviative using the process in Section 1.3.1. When a local symmetry is broken by the physical vacuum, the interplay between the “would-be" Goldstone bosons and the massless gauge bosons removes the Goldstone bosons from the physical spectrum and implies masses for the gauge bosons.

Spontaneous breakdown of the electroweak symmetry group has important effects on the physical boson and fermion states in the SM, many of which are observed with masses generated by the mechanism of symmetry breaking. For example, several bosons in the electroweak sector have measured masses, but the fields derived from the steps described up until this point are massless. Three out of four of the standard electroweak generators come from the $SU(2)$ group and one from $U(1)$, which together generate the four massless fields labelled W^1, W^2, W^3 and B. Below the weak scale at 246 GeV, the physical electroweak bosons acquire mass from spontaneous symmetry breaking, where the Higgs mechanism breaks $SU(2)_L \times U(1)_Y$ into $U(1)_Q$ electromagnetism [10]. The B quantum field mixes with W^3 to produce two observed integer-spin particles: the photon and the Z gauge boson. Similarly, the massless W fields mix to produce the physical $W^{\pm}$ bosons. Although fermions are already allowed to have mass terms in the SM Lagrangian, they obtain additional mass terms from this process. This section describes spontaneous symmetry breaking via the Higgs mechanism.

1.6.1 The Higgs mechanism

The SM fermions and the $W^{\pm}$, Z and Higgs bosons are all measured at observed masses which agree with the terms they acquire via the spontaneous breakdown of the electroweak gauge symmetry. The Higgs mechanism allows us to build an experimentally accurate gauge invariant theory, using interactions with a complex scalar field. The mechanism proceeds differently for Abelian and non-Abelian symmetries, reminiscent of the differences in the gauge theory. The Abelian case involves one vector field coupled to a single complex scalar field, whereas the true Higgs mechanism of the SM is non-Abelian involving non-commuting group elements, and comprises several vector fields $A_{\mu}^{\alpha(x)}$ and two scalar fields, $\phi_i(x)$. The SM is non-Abelian in general, as for electroweak and QCD interactions.

The Higgs sector in the SM introduces an $SU(2)_L$ Higgs doublet Φ with weak isospin $T = 1/2$ and hypercharge $Y = 1$. The Higgs potential V is typically described using a Mexican hat function:

$$V(\Phi) = \lambda[\Phi^{\dagger}\Phi - \frac{v^2}{2}]^2 \tag{1.9}$$

where λ is an overall scaling factor. We can describe the Higgs field in a physical gauge using the form:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+H \end{pmatrix} \tag{1.10}$$

where H relates the fluctuation field from the ground state and v parameterises the Higgs field vacuum expectation value (vev), which can take many possible complex values but is often chosen for convenience as a point on the real axis. Denoting the Higgs mass as M_H, we can write the Higgs potential as:

$$V(H) = \frac{1}{2}M_H^2 H^2 + \frac{M_H^2}{2v}H^3 + \frac{M_H^2}{8v^2}H^4 \tag{1.11}$$

Shifting to the minimum of the Higgs potential generates mass terms that fully determine the masses of the gauge bosons. If we expand Φ around the minimum of the potential, we obtain one massive scalar particle (the Higgs boson), plus three massless Goldstone bosons which are absorbed to produce masses for the $W^{\pm}$ and Z bosons.

Mass is generated in this case if transformation of the vacuum state by the generators causes it no longer to share the symmetry of the Lagrangian. When we shift to the minimum of the Higgs potential, the non-Abelian Higgs mechanism of the SM produces a mass matrix (in place of the single mass term that we would expect from the Abelian mechanism). By diagonalising this matrix, we obtain one of two distinct possible solutions. First, the generators may annihilate the vacuum so the symmetry remains unbroken, producing massless vector bosons. Alternatively, the generators may not annihilate the vacuum, instead transforming the vacuum so it is no longer symmetric under the action of the group of symmetries. In this case where the symmetry is broken, we obtain massive vector bosons and the SM Higgs boson couples proportional to mass.

1.6.2 Mixed bosonic states

In the non-Abelian broken symmetry of the electroweak sector, the physical bosonic states observed in experiment are combinations of field contributions from the $W^{1,2,3}$ and B. All electroweak interactions depend on the mixing of these fields, determining which interactions are permitted. For example, the Z interacts with all fermions, containing contributions from both the $U(1)_Y$ B boson and the $SU(2)_L$ W^3. The strength of these interactions depends on the relative contributions from the B and W^3 fields; in this case, stronger interactions are observed between the Z and the left-handed fermions. Equations 1.12 and 1.13 parameterise the mixing contributions in the Z boson and photon (A) states using the Weinberg angle θ_W, which is known as the angle by which spontaneous symmetry breaking rotates the plane of massless vector bosons.

$$Z_\mu = \cos(\theta_W)W_\mu^3 - \sin(\theta_W)B_\mu \tag{1.12}$$

$$A_\mu = \cos(\theta_W)B_\mu + \sin(\theta_W)W_\mu^3 \tag{1.13}$$

The relationship between the Weinberg angle and the electroweak gauge couplings reveals further information about the relative strengths of the two terms in the equations above. In Equations 1.14 and 1.15, the gauge coupling strength for $SU(2)_L$ is g_2 and the coupling for $U(1)_Y$ is g_Y.

$$\cos(\theta_W) = \frac{g_2}{\sqrt{g_2^2 + g_Y^2}} \tag{1.14}$$

$$\sin(\theta_W) = \frac{g_Y}{\sqrt{g_2^2 + g_Y^2}} \tag{1.15}$$

Similar mixing of the original electroweak gauge fields produces the physical $W^{\pm}$ bosons. Equation 1.16 shows the physical $W^{\pm}$ states mixed entirely from $SU(2)_L$ bosons, corresponding to our observation of their electroweak interactions in experiment.

$$W^{\pm} = \frac{1}{\sqrt{2}}(W_{\mu}^{1} \mp iW_{\mu}^{2}) \tag{1.16}$$

The masses of the $W^{\pm}$ and Z bosons obtained through the Higgs mechanism are in agreement with the short distance range of the weak nuclear force.

1.6.3 Fermion mass terms

The electroweak gauge symmetry of the SM forbids fermion mass terms in the fundamental Lagrangian because they are constructed from left and right components, where the left are doublets and the right are singlets, and therefore cannot be written in a way that is gauge invariant. The fermions obtain mass from interactions with the Higgs scalar fields, and all interactions between fermions and scalar fields are described in the SM Lagrangian by *Yukawa terms*. Through spontaneous symmetry breaking, the fermions in these interactions acquire a mass proportional to the vev of the Higgs field. Equation 1.17 contains an example of the Yukawa terms describing the interactions between an SM fermion field and the Higgs Φ, where left- and right-handed components of the fermion field are separated because their different spin quantum numbers require us to treat them as different particles. The fermion field denoted e_R represents the right-handed electron singlet, and Ψ_{eL} is the left-handed doublet. The Higgs coupling with the electron is written as h_e.

$$L_{YU} = -h_e(\overline{\Psi}_{eL}\Phi e_R + \overline{e}_R\Phi^{\dagger}\Psi_{eL}) \tag{1.17}$$

After spontaneous symmetry breaking, rewriting the Higgs field Φ in terms of its vev (v) reveals which terms in the Yukawa component of the Lagrangian contain the Higgs vev, and are therefore the mass terms of the electron field. They are identified in Equation 1.18, which is written in terms of the component e_L of the left-handed electron field doublet.

$$L_{YU} = -h_e(\frac{v}{\sqrt{2}})(\bar{e}_L e_R + \bar{e}_R e_L) + [\text{other terms}] \tag{1.18}$$

The electron mass is determined by the effects of spontaneous symmetry breaking on the Yukawa terms, and can be discerned directly from Equation 1.18 in terms of the Yukawa coupling constant and the Higgs field vev:

$$m_e = \frac{h_e v}{\sqrt{2}} \tag{1.19}$$

The mass equations proceed in a similar way for the muon and tau. The quarks have similar Yukawa couplings to the leptons, but the right-chiral quark components add terms to the Yukawa couplings that are not present in the leptonic couplings (for models where right-handed neutrinos do not exist). The consequence of the right-chiral quark components in the Yukawa couplings is to form a mass matrix for the down-type quarks, instead of a mass term like for the leptons. Diagonalising the mass matrix using a biunitary transformation reveals the masses of the physical down-type quarks through the eigenvectors of the matrix.

1.7 Parameters in the Standard Model

Diagonalising the quark mass matrix requires us to apply a rotation to the fermion fields, where the same rotation must be applied to both the charged current and neutral current components. Since the charged current couples different objects (such as up and down quarks), the rotations do not cancel out, resulting in quark sector flavour mixing. This is the origin of the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which describes the extent to which the quantum states of quarks are mixed by the weak interactions [11, 12]. The CKM matrix contains the information required to predict how much of each quark mass eigenstate exists in the weak eigenstates, using rotation angles called the quark mixing angles.

The CKM matrix is a unitary matrix containing fundamental parameters of the SM, and its unitarity hinges upon the values of these parameters. For example, the most sensitive test of unitarity in the SM with three quark generations is provided by the relation,

$$|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 1 - \delta \tag{1.20}$$

where $|V_{ud}|$ describes the probability for a d quark to transition to a u quark, and so on. Here, a non-zero value of δ indicates that new physics is present. These components of the matrix V_{CKM} are generally parameterised by the three real quark mixing angles and one imaginary phase, whose precise values are determined fundamentally by the couplings in the charged current part of the Yukawa interactions in the electroweak Lagrangian. Consequently, the unitarity of the CKM matrix is a strong experimental test of the merit of the SM. Tests of CKM unitarity using experimental decay rates and lattice field theory have found that the SM is an accurate prediction of the parameters of the CKM matrix. Strict constraints have been placed on the CKM matrix as a result of these searches [13].

Combined with the four parameters defining the CKM matrix, the SM has nineteen other free parameters that must be measured experimentally. These are: the Higgs vev and mass, the Weinberg angle θ_W, the masses of the nine charged fermions, three gauge couplings and four parameters defining the relationships between the weak interaction eigenstates of the neutrinos [14].

The latest measurements of the SM reveal no signs of inaccuracies up to the weak scale. Searches at the LHC and other colliders have tightly constrained its fundamental parameters, including the fermion and boson masses which have all now been measured experimentally. Since the discovery of the Higgs boson confirmed how boson and fermion mass contributions come from interactions with the Higgs field, the SM has been considered the appropriate field theory for encapsulating physics on the scale of the weak interactions, $m_{weak} \sim 246$ GeV. Increasingly precise data continues to tighten the constraints on the parameters, repeatedly confirming that the theory is complete at low enough energies.

1.8 Limitations of the Standard Model

Historically, experimental advances have allowed us to delve into progressively more microscopic theories on shorter distance scales. We can then use these more fundamental theories to backtrack and calculate the masses of particles from outdated theories. For example, developing the present-day SM enabled calculations to be made of the pion mass, a particle that had previously been discovered experimentally but was unknown in origin. Once the current account was formulated describing the pion as a composite particle, its mass was calculated from the quark masses. There is little reason to assume the same will not be true for the SM, or that physics stops at the scales already probed. The various shortcomings of the current model provide further evidence of a more fundamental microscopic theory that will allow predictions to be made of parameters it cannot explain.

The main shortcomings of the SM can be split into two main classes: observational and structural problems. Observation of the matter existing today, which is comprised of the one particle per billion that still remains after the Big Bang, indicates that there is very little antimatter. Matter and antimatter are found to behave differently in experiments, including the first observations in 2019 of CP violation for the charm quark in D0 meson decays at LHCb [15]. However, the matter-antimatter asymmetry is unexplained by the SM. The size of CP violation in SM interactions in the quark sector is not large enough to explain the measured imbalance, implying there are additional, unknown sources of CP violation. Another unexplained observation is neutrino oscillations, which indicates that they have mass. The SM predicts no mass for neutrinos, but Yukawa couplings can give them very small masses. Solving the problem by introducing lepton number violation with higher dimensional operators leads to unitarity violation at higher energies, requiring new degrees of freedom. Adding interactions to the SM may solve both of these observational problems.

Additional degrees of freedom may also explain the nature of dark matter, a critical observational problem the SM cannot address without new physics. Cosmological evidence from descriptions of the formation of galaxy clusters in the observable universe, as well as the evolution of galaxies and galactic collisions on a more local scale, indicate that a missing mass constitutes 85% of total mass. Dark energy and dark matter combined must form 95% of the total mass-energy content. The observational evidence from gravitational interactions places stringent limits on the type of matter that is missing, with popular candidates being a new type of weakly interacting massive particle (WIMP), or axions [16, 17]. For example, a natural WIMP-like dark matter candidate is proposed by supersymmetric extensions to the SM.

Supersymmetric extensions can also unify the electroweak and strong forces, which cannot be achieved without a mechanism changing their high-energy dependence. Grand unified theories (GUTs) are another method of performing this change, relating the free parameters in the SM to reduce them to a single gauge coupling. Gravitational interactions are not described in the SM, so the process of unifying the gravitational force with the strong and electroweak forces at any scale is already highly uncertain. Gravitational interference is vital to inferring the location of dark matter, so its absense from the SM alongside dark matter poses questions about the shortcomings of our current model.

1.8.1 Structural issues & the Hierarchy Problem

New interactions may solve the structural problems with the SM, which are shaped by the unexplained hierarchies in the generations and scales of the theory. For example, the three-generational hierarchy in the masses of the quarks and leptons could be explained by a more microscopic theory. Labelling the particles by their gauge couplings produces three copies in an exponential hierarchy, due to a pattern in the Yukawa couplings. The generational patterns are compelling evidence for an underlying theory containing the information required to derive the masses from more fundamental principles. A similar argument is relevant for the Higgs mass, which has a bare mass that could in future be calculated from fundamental parameters.

Other structural conundrums abound, typically appearing as unexplained patterns similar to the quark and lepton masses, or as disagreements between theory and experiment. One disagreement is the cause of the strong CP problem, where no CP violation is observed in the strong sector even though no known symmetry holds it at zero. There are natural terms in the QCD Lagrangian able to break CP symmetry that suggest CP violation should be observed at the magnitude of those terms, but no evidence has been found. Another structural problem, this time an unexplained pattern, is the hierarchy between the Planck scale with mass $M_P = \frac{1}{\sqrt{8\pi G_N}} \approx 1.2 \times 10^{19}$ GeV and Fermi's constant on the weak scale [18]. These two parameters operate at energy scales that differ by 16 orders of magnitude. The difference between the two scales forces SM parameters to be finely tuned to eliminate the quadratic divergence of their contributions to the Higgs mass. If the Planck scale is the fundamental scale of nature, then the bare Higgs mass must be on the order of the Planck scale, requiring an unnatural tuning of one part in 10^{17} between the contributions to the Higgs mass to cancel the quadratic divergence and obtain a mass on the electroweak scale [19].

The LHC has revealed that the Higgs indeed has a light mass at 125 GeV, which allows the quartic coupling of the Higgs to remain small so that the scattering of the longitudinal components of the W bosons, which grows with energy, does not break down at exponentially higher energies. Otherwise, this extra degree of freedom belonging to the massive spin-0 particles would result in a Landau pole at high energies, breaking down the theory.

In this respect, the light mass of the Higgs is a requirement of the high-energy theory, but it is also the basis of the hierarchy problem. Such a light Higgs is not predicted by the SM, which gives it couplings that break any symmetry that could have kept it light. The option to finely-tune these couplings to cancel their contributions is unconvincing. In the

SM, the Higgs couplings to the fermions is proportional to the fermion masses, reflecting the hierarchy of the generations. In Figure 1.2, the one-loop correction to the Higgs mass squared parameter m_H^2 due to the Dirac fermion f with mass m_f is written

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda_{UV}^2 + ... \tag{1.21}$$

where the quadratic divergence of the Higgs loop corrections causes the mass to become highly ultraviolet (UV)-sensitive. A convenient method of handling the UV divergences in the SM is provided by dimensional regularisation and then renormalisation. The loop integral is regulated by the UV momentum cutoff Λ_{UV}, which is typically the energy scale of new physics. The terms in this sum which are proportional to the fermion mass cause large corrections to the Higgs scalar boson squared mass because of the direct quadratic sensitivity to the cutoff Λ_{UV}.

Fig. 1.2 One-loop corrections to the Higgs boson mass, due to a fermion f and a scalar S [20].

The question remains whether the Higgs may break flavour universality as part of a beyond-SM (BSM) theory. If the mass of the Higgs boson is calculable from a more fundamental microscopic theory, the cutoff terms imply that new physics should emerge at around 400 GeV, or around 2 TeV in the gauge sector, to which the Higgs mass is sensitive. No such physics has yet been observed, but this discovery would naturally follow the progression of particle physics through history as experiments have moved to higher energy scales.

If the Higgs is not a composite particle, its positive parity lends it only one degree of freedom, making it the first elementary scalar found in nature. If it is a composite particle then its mass, like the mass of the pion, may be calculated in the future from its component particles. If the Higgs is composite, we expect three main phenomenological signatures to appear at high energies: $\rho^{'}$ (rho prime) vectors, top partners and Modified Higgs Couplings. Searches for all three are ongoing at the LHC, but none are yet supported by compelling evidence. Regardless, it is convincing to argue that solving the hierarchy problem requires new high-energy physics that couples to the Higgs.

The idea that degrees of freedom may reveal completely new objects at higher energies is not new. Already, the theory of gravity cannot hold at high energies, and on the Planck scale on the order of 10^{19} GeV, a quantum description of gravity breaks down. Combining quantum field theory with gravity generates problems with renormalisability, and quantum computations involving gravity at these energies are non-renormalisable. Therefore, a UV cutoff must lie either at or before the Planck scale, at which point the theory requires new physics. Effective field theories (EFTs) are a general method of describing effective interactions at SM energies that parametrise the effects of a higher energy theory, an example of which is the Fermi theory. In the case of gravity, an EFT yields a renormalisable quantised theory of gravitational interactions up to energies below the Planck scale, at which point it breaks down. The effective interactions in an EFT involve the low-energy SM fields only, so the precise nature of the new physics is irrelevant to the parametrisation.

The limitations of the SM prompt us to return to developments in theory and experiment to describe the physics model which best explains the gaps. Although many scenarios remain possible, only some are both theoretically well-motivated and phenomenologically viable, able to solve several remaining questions and offer a chance of discovery at the current experimental energy frontier.

Chapter 2

Supersymmetry

2.1 Introduction to supersymmetry

Supersymmetry is a class of extensions to the SM that aims to incorporate the existing gauge field theory into a larger model by combining observed physics with additional particles and interactions. Supersymmetric extensions are praised for their mathematical simplicity, which derives from their consistent structure and their reliance on only a few key concepts. The primary idea behind these theories is the existence of an additional symmetry of nature linking fermions and bosons, which results in several benefits, including a possible solution to the hierarchy problem of Section 1.8.1.

The principle of supersymmetry extends the existing Poincaré spacetime symmetry structure of the SM to include an as-yet-unobserved fermionic spacetime symmetry. The extended supersymmetric SM is consequently invariant under generalised spacetime transformations linking fermions and bosons. Phenomenologically, a quantum field theory that requires supersymmetry is one in which every fermion (or boson) degree of freedom is the “partner” of a boson (or fermion) degree of freedom that has the same set of non-mass quantum numbers. If the *superpartner* particles are also mass degenerate, where the new supersymmetric state (“sparticle") and its SM counterpart have the same mass, then the symmetry is *exact*. Although this extension implies the addition of at least 17 extra particles, it does so by tightening constraints on the structure of the theory which ultimately simplifies the model.

Supersymmetry is a common target for collider physics searches because theories that incorporate it have several plausible properties. Initial interest in supersymmetry during the

1970s and ‘80s focused on the nonrenormalisation of some supersymmetric interactions and the higher UV convergence of the field theories, as well as the link to gravity [21]. Soon to follow, further theoretical motivation came from the idea that we observe only a low energy manifestation of Planck scale physics, and that supersymmetric parameters at high energies may solve the hierarchy problem [22]. The experimental discovery of supersymmetry was therefore highly anticipated by phenomenological studies at the advent of the LHC, and continues to motivate new search strategies in the hope of a discovery in the near future.

2.2 Key motivations

If supersymmetry implies the emergence of degrees of freedom at energy scales above the current observed range, and if the Higgs mass is sensitive to them, then the fine-tuning between the loop corrections to the Higgs mass are no longer necessary. The observed masses of the Higgs and W and Z bosons are preserved without unnatural fine-tuning [23]. For each leading order correction shown in Figure 1.2, supersymmetry requires another that comes from the superpartner of the old one. The Feynman rules require that a -1 is added with every fermion loop, which causes all such pairs of corrections to cancel exactly, leading to theories that allow the Higgs to remain light in spite of its many couplings to massive particles. The cancellation is exact only when the superpartners are equal in mass to their SM counterparts, and the supersymmetry is *softly broken* when the superpartners are heavier within constraints that preserve the leading order cancellation. This cancellation is one of the primary advantages of supersymmetric extensions to the SM, because it provides a natural alternative to fine-tuning.

In addition, various models of supersymmetry are capable of unifying the fundamental interactions. Some unify all particles and forces except gravity on the order of $10^{15} - 10^{16}$ [24, 22], but other models include gravity. Supergravity theories generalise gravity using a local gauged supersymmetry, which automatically produces Einstein’s general theory of relativity as a necessary consequence [25]. This method of unifying gravity with the other fundamental interactions is a natural motivation for local supersymmetry theories. String theories are often cited as methods of achieving unification, and superstring theory incorporates supersymmetry as an integral component. Supersymmetry is a vital ingredient of various prominent theories unifying all the couplings of the strong, electroweak and gravitational interactions, discussed further in Section 2.5.

2.3 Symmetry properties

The generator for the additional symmetry transformation is the supercharge operator Q. The operator is a spinor, carrying spin angular momentum $\frac{1}{2}$. The properties of the superpartners come from the commutation and anti-commutation relations for Q; SM particles alone do not satisfy the properties of all the particles implied by the enhanced symmetry. When acting on a boson state Q generates a fermion state, and vice versa, where the operation produces superpartner states that differ in spin from their counterparts by $\pm\frac{1}{2}$. Consequently, the superpartners of the fermions are scalars, and are called sleptons (scalar leptons) and squarks (scalar quarks). The superpartners of the SM bosons are spin-$\frac{1}{2}$ fermions, and these fermionic sparticles have physical states called the charginos, neutralino, gluino and Higgsino. Equation 2.1 summarises the operation of the symmetry transformation.

$$Q\left|\text{boson}\right\rangle = \left|\text{fermion}\right\rangle, \qquad Q\left|\text{fermion}\right\rangle = \left|\text{boson}\right\rangle \tag{2.1}$$

The property of the operators Q_α and $Q^\dagger_\alpha$ where they transform quantum mechanical states between fermionic and bosonic statistics is encoded in the anticommutation relations. Equations 2.2 - 2.4 are the relations satisfied by the generators in supersymmetry. Here, four-component Dirac spinors are split into 2 two-component spinors denoted by indices α or β for the first two components and $\dot{\alpha}$ or $\dot{\beta}$ for the second. The four-momentum P^μ is the generator of spacetime translations under Lorentz boosts and rotations.

$$\{Q_\alpha, Q^\dagger_{\dot{\alpha}}\} = -2\sigma^\mu_{\alpha\dot{\alpha}} P_\mu \tag{2.2}$$

$$\{Q_\alpha, Q_\beta\} = \{Q^\dagger_{\dot{\alpha}}, Q^\dagger_{\dot{\beta}}\} = 0, \tag{2.3}$$

$$[Q_\alpha, P^\mu] = [Q^\dagger_{\dot{\alpha}}, P^\mu] = 0 \tag{2.4}$$

Following Equation 2.4, operators Q_α and $Q^\dagger_\alpha$ commute with the squared-mass operator $-P^2$ such that, for example, $[P^2, Q_\alpha] = 0$. The squared-mass operator also commutes with all spacetime translation and rotation operators. These relations require that the superpartners have the same mass as their SM counterparts. Although this rule must hold in exact supersymmetry, superpartner masses are allowed to differ when the symmetry is broken, described later in Section 2.4.3. Furthermore, the Q and $Q^\dagger$ operators commute with the momentum operator and with the generators of the gauge transformations which gives the superpartners the same values of weak isospin, electric and colour charge as the particles in the SM. In general, every supersymmetric particle inherits the same (non-mass) quantum numbers as its SM counterpart as a result of the continuous symmetry of the theory, which commutes with the SM gauge group.

2.4 The MSSM

The Minimal Supersymmetric Standard Model (MSSM) is the “simplest” supersymmetric model that requires several constraints. All the particle content of the SM must be contained in the MSSM, and it must preserve the SM conservation laws. The most general phenomenologically viable type of MSSM is an $N = 1$ supersymmetry, and this is the type most frequently studied. Here, N is the number of supersymmetries in the model, equivalent to the number of distinct copies of the generators Q, $Q^\dagger$. Extended models of supersymmetry exist where $N > 1$ copies of the generators cause unphysical phenomenologies in four-dimensions, but which may describe real physics in higher-dimensional field theories with careful treatment of extra dimensions [26]. The MSSM is the minimal extension required to incorporate a softly-broken $N = 1$ global supersymmetry, where extra dimensions are not required.

The $N = 1$ MSSM is motivated by its desirable properties including conservation of charge parity, minimal flavour violation at the TeV scale and suppressed flavour-changing neutral currents at tree level, which constrain some of the possible free parameters in supersymmetry-space. Although the number of free parameters is greater than 100 at first, extra simplifying assumptions sometimes decrease this number to comprise only 10 sfermion masses, 3 Higgs and Higgsino parameters, 3 gaugino masses and 3 trilinear couplings [27]. These simplified models narrow the experimental search region.

Combining supersymmetry with gravity as mentioned in Section 2.2 is possible as part of a phenomenologically viable MSSM. The theory must promote supersymmetry to a local symmetry and unify the resulting transformations with the spacetime symmetries of general relativity. When an MSSM scenario is extended to include gravity it typically involves non-renormalisable interactions, but these can be ignored in the phenomenology at LHC energies because their effects are suppressed by inverse powers of the high energy cutoff, λ_{UV} [28]. Supergravity therefore remains an interesting method of incorporating gravity into the quantum field theory, and minimal supergravity (MSUGRA) is a benchmark scenario for experimental collider search limits.

Superpartner states are stored in two-component objects called supermultiplets, which contain the pairs of bosonic and fermionic states that are related via a supersymmetric transformation. Whereas the irreducible representations of the symmetry group of the SM algebra were multiplets, the same concept is referred to in supersymmetry as supermultiplets. Supermultiplets contain both the SM particles and the superpartners of those states, containing both fermion and boson members. The components within a supermultiplet must possess the same electric charge, mass, weak isospin and colour degrees of freedom. It is always possible to write $N = 1$ supersymmetric theories in terms of only two different types of supermultiplet, called chiral and gauge supermultiplets. Both types contain two spin-$\frac{1}{2}$ fermions, but whereas chiral supermultiplets also contain two scalars, gauge supermultiplets contain one massless spin-1 boson instead. Superfields are the fields of the superpartners of the SM particles, and they are sorted into the chiral or gauge supermultiplets depending on their transformation properties.

2.4.1 Chiral & gauge superfields

Chiral supermultiplets treat left- and right-handed fermionic components differently, whereas gauge supermultiplets treat them the same. The quarks and leptons in the SM belong in chiral supermultiplets, because they have left and right states that transform differently under Lorentz transformations. Since their superpartners must reside in the same supermultiplets, the superpartners of the SM fermions are also in chiral supermultiplets and are scalars with spins half a unit *less* than their SM counterparts, as shown in Table 2.1. The final members of Table 2.1 are the SM scalar Higgs particles and their superpartners, the Higgsinos, which are chiral superfields with spin $\frac{1}{2}$, this time half a unit of spin *more* than their SM counterparts. In Table 2.2, the gauge supermultiplets are shown, which must contain the spin-1 SM vector bosons according to their transformation properties. The superpartners of the SM gauge bosons belong in the same gauge supermultiplets and have spins half a unit less.

If the SM were extended to include a particle description of gravity, then the spin-2 graviton would adopt a superpartner called the gravitino with spin $\frac{3}{2}$. The supermultiplet structure in the graviton-gravitino sector is therefore different to the chiral and gauge supermultiplet structure for the SM particle fields and their superpartners.

The properties of the fields belonging to the supermultiplets in the MSSM are clearly not accounted for by the SM fields alone, which motivates the addition of the new states. For example, a spin-$\frac{1}{2}$ neutrino cannot form the superpartner of the spin-1 photon, because the

Table 2.1 The chiral supermultiplets of the MSSM [20].

Names		spin 0	spin 1/2
squarks, quarks	Q	$(\tilde{u}_L\ \tilde{d}_L)$	$(u_L\ d_L)$
×3 families	$\overline{u}$	$\tilde{u}^*_R$	$u^\dagger_R$
	$\overline{d}$	$\tilde{d}^*_R$	$d^\dagger_R$
sleptons, leptons	L	$(\tilde{\nu}_L\ \tilde{e}_L)$	$(\nu_L\ e_L)$
×3 families	$\overline{e}$	$\tilde{e}^*_R$	$e^\dagger_R$
Higgs, Higgsinos	H_u	$(H_u^+\ H_u^0)$	$(\tilde{H}_u^+\ \tilde{H}_u^0)$
	H_d	$(H_d^0\ H_d^-)$	$(\tilde{H}_d^0\ \tilde{H}_d^-)$

Table 2.2 The gauge supermultiplets of the MSSM [20].

Names	spin 1/2	spin 1
gluino, gluon	$\tilde{g}$	g
winos, W bosons	$\tilde{W}^\pm\ \tilde{W}^0$	$W^\pm\ W^0$
bino, B boson	$\tilde{B}^0$	B^0

photon is neutral under weak isospin whereas the neutrino forms a doublet. Despite the name "minimal", the MSSM does not contain the minimum possible new particle content, because it introduces a new superpartner for every SM particle. Not including the gravitino, the MSSM contains 32 distinct mass parameters for undiscovered particles. Some superpartners seem to share the properties of SM particles; a Higgs doublet with hypercharge -1 transforms identically to the $SU(2)$ (ν_L, e_L) doublet and therefore could be contained in the same chiral supermultiplet. However, if this is true and the sneutrino is indeed the Higgs boson, then several phenomenological problems arise, including an incorrect neutrino mass and lepton number violation. All SM particles must consequently belong in separate supermultiplets, partnered by a new degree of freedom.

Electroweak symmetry breaking in the MSSM performs a similar function to its role in the SM and mixes the higgsinos with the electroweak gauginos. The zinos and winos, which are the superpartners of the electroweak Z and W bosons and are together called the *electroweakinos*, mix with the higgsino to form the physical states, the charginos and neutralinos. The gluinos are the strongly-interacting fermion partners to the SM gluons, and are spin-$\frac{1}{2}$ gauginos that do not mix. The supersymmetric states transform in interactions as either chiral or gauge supermultiplets via the transformations contained in the gauge groups, analogous to the transformations of multiplets in the SM.

2.4.2 Gauge groups & the Lagrangian

Abelian gauge transformations on the chiral superfields describe the interactions in supersymmetric quantum electrodynamics (*SQED*), and non-Abelian gauge transformations describe supersymmetric quantum chromodynamics (*SQCD*) and supersymmetric chiral gauge theory ($S\chi GT$). The latter two are extensions of the SM gauge field theories based on the strong $SU(3)_C$ and electroweak $SU(2)_L \times U(1)_Y$ gauge groups. Interactions are described in the *SQCD* and $S\chi GT$ Lagrangian densities, resulting in a list of allowed supersymmetric vertices and their Feynman rules, informing the ways in which sparticles are predicted to be experimentally observed through their interactions with each other and with SM particles. The interactions are contained in the MSSM Lagrangian.

The Lagrangian of the MSSM can be written generally as the sum $L_{MSSM} = L_{SUSY} + L_{soft}$. L_{SUSY} contains the interaction and mass terms coming from the exact supersymmetrisation of the SM; in other words, it contains all gauge and Yukawa interactions and preserves supersymmetry invariance [26]. L_{soft} describes the heavier masses of the sparticles than their partner particles, containing terms derived from supersymmetry breaking. The supersymmetric part of the MSSM Lagrangian decomposes into pure gauge, matter and Higgs-Yukawa parts, $L_{SUSY} = L_g + L_M + L_H$, which contain terms with different relevant field strengths for the spinorial superfields. The gauge part of the supersymmetric Lagrangian is a sum of interactions between the chiral and vector superfields.

2.4.3 Supersymmetry breaking

Electroweak symmetry breaking in the MSSM is analogous to the mechanism of the SM and is required to give mass to the standard gauge bosons, quarks and leptons. In supersymmetry, however, two Higgs doublets are typically required instead of one, cancelling the higgsino contributions to the gauge anomaly, and also allowing holomorphic Yukawa couplings. Like spontaneous symmetry breaking in the SM (described in Section 2.4.3), the spontaneous mechanism in the MSSM can only give masses to particles via their Yukawa interactions, even though the superpartners of the SM quarks and leptons must be more massive than their counterparts to explain why they have not yet been observed. The difference in mass between the superpartners must come from another form of supersymmetry breaking.

The group transformations in a quantum theory with an exact symmetry leave the vacuum state Ψ_0 invariant, so it must be annihilated by the symmetry group generators. In an exact supersymmetry with mass degeneracy between the superpartners, the vacuum state is annihilated by the generators Q as in Equation 2.5. The annihilation of the vacuum state implies that the vacuum state energy of the Hamiltonian also vanishes, because the Hamiltonian is expressed in terms of the supercharges, $H = \{Q, Q^{\dagger}\}$. However, this is no longer true in a spontaneously broken supersymmetry, when the vacuum state energy of the Hamiltonian becomes non-zero.

$$Q\Psi_0 = Q^{\dagger}\Psi_0 = 0 \tag{2.5}$$

In models with spontaneous supersymmetry breaking, by definition the vacuum state must not be invariant under supersymmetry transformations. In particular, the vacuum must have positive energy, because the inner product with the Hamiltonian operator H acting on $\Psi_0 = |0\rangle$ gives $\langle 0| H |0\rangle > 0$ if the Hilbert space is required to have positive norm.

As a global continuous symmetry, supersymmetry can be broken in one of three ways, of which the spontaneous breaking mechanism of electroweak symmetry breaking is only part. The other two possibilities are the explicit addition of terms to L_{soft}, and the combination of both explicit and spontaneous symmetry breaking. Explicit supersymmetry breaking leaves the vacuum invariant and keeps supermultiplets intact except for mass splittings, characteristic of what is known as the Heisenberg-Wigner mode. Spontaneous breaking instead occurs in the Nambu-Goldstone mode, which causes the vacuum symmetry to break and multiplet structure to be destroyed. These are the possible mechanisms for breaking mass degeneracy of the superpartners, although spontaneous supersymmetry breaking is a favoured option. The spontaneous mechanism adds new particles and interactions at high mass scales such that supersymmetry is hidden at low energies just as the electroweak symmetry is hidden below the appropriate mass scale in the SM.

Explicit versus spontaneous breaking

In explicit supersymmetry breaking, superfields do not break supersymmetry spontaneously at the weak scale by themselves; some soft supersymmetry breaking operators must be added to the Lagrangian as part of L_{soft} so that the superpartners obtain heavier masses. The coefficients of the operators in L_{soft} are treated as unknown parameters of the MSSM. The soft supersymmetry-breaking coupling parameters and mass terms must have positive mass dimensions so they can maintian a hierarchy between the electroweak and Planck scales and enable a solution to the hierarchy problem. This requirement on the mass dimensions in particular prohibits dimensionless supersymmetry-breaking couplings, of which there would otherwise be several notable possibilities. These terms only involve scalars and gauginos and not their respective superpartners, so they must break the supersymmetry. The added soft terms can give mass to all of the scalars and gauginos, even if the gauge bosons and fermions in chiral supermultiplets are massless.

Spontaneous supersymmetry breaking is different, as mentioned above, in that it instead involves a positive ground state energy that produces non-degenerate superpartner mass states, where no equivalent unknown parameters are added to the theory. In this sense, spontaneous supersymmetry breaking is preferred in the search for a theory that explains the mechanism rather than merely parameterising it. Explicit supersymmetry breaking adds a large number of free parameters, so that the broken MSSM obtains around 105 mass, phase and mixing angle parameters which cannot be removed by redefining the basis for the supermultiplets, even though the unbroken MSSM has only a similar number of free parameters to the SM. The number of free parameters added by models with spontaneous supersymmetry breaking is an order of magnitude smaller.

Only soft supersymmetry breaking terms with mass dimension less than 4 are permitted in the weak scale effective Lagrangian, to control the quadratic divergence in the radiative correction to the Higgs mass in the SM. These terms are a consequence of the spontaneous breakdown of supersymmetry at a high energy cutoff scale, Λ_S. The information from the Λ_S scale appears in the physics of our low energy world via messengers, which act on some mass scale M_M. If the supersymmetry breaking terms in the weak scale Lagrangian are on the mass scale M_S, then the mass scale of the messengers is approximately $M_M \sim \frac{\Lambda_S^2}{M_S}$. The high scale mechanism of supersymmetry breaking fixes the values of the soft supersymmetry breaking parameters on the weak scale. The weak scale values of these parameters control superpartner masses and, in the case where mixing occurs, they influence the superpartner couplings too. Pinpointing the values of these weak scale parameters could rely on measuring the scalar superpartner masses and determining the gaugino-higgsino mixing parameters, amongst other possible measurements. In turn, these values point to the high energy physics responsible for supersymmetry breaking at the Λ_S scale.

2.4.4 Natural supersymmetry

Since supersymmetry must be broken to allow room for discovering superpartners in a higher mass range, the perfect cancellation of leading order contributions to the Higgs mass is not possible. In “natural supersymmetry” theories, the couplings of the sparticles to the Higgs balance well with those of their SM partners, still allowing the model to bypass any need for unnatural fine-tuning. To preserve naturalness, the masses of the superpartners are therefore constrained by the Higgs mass and cannot be too heavy: the mass splitting between partners becomes severely limited, especially between the top and the stop (which is the lightest squark according to prediction) [29]. To observe a Higgs at 125 GeV, the stop mass cannot be greater than around 400 GeV without reintroducing a large degree of tuning between this correction and the others that remain after pair cancellation.

In general, these considerations for natural electroweak symmetry breaking indicate that stops, the gluino and higgsinos are close to the weak scale. Therefore, search efforts focusing on natural supersymmetry target these lighter states, but allow the remaining superpartners to have masses in excess of 1 TeV. In Figure 2.1, the superpartners to the right are decoupled at high mass, but the superpartners on the left are not: the gluino enters the Higgs mass matrix at two-loop level, the stop and sbottom at one-loop and the Higgsino mass parameter (μ) at tree-level. In addition, natural supersymmetry has a weak isospin symmetry demanding a left-chiral sbottom. LHC searches focus on these states, and produce experimental constraints on the superpartner masses. Various stop searches suggest that the stop mass is larger than 1 TeV, indicating that some degree of fine-tuning may be required on the order of the percent level or below, despite attempts to promote a natural theory. The possibility remains that a small amount of fine-tuning exists in the theory, and that these new particles will soon be detected at fairly low mass.

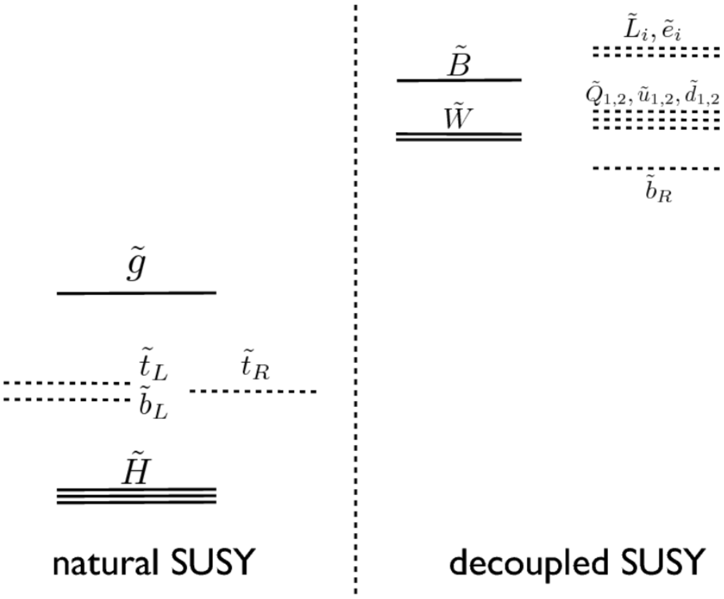


Fig. 2.1 A mass spectrum comparing parameters of natural supersymmetry (SUSY) with the possibility of decoupled superpartners at high mass [30].

2.4.5 R-parity symmetry

Any new physics interactions must have good symmetry under R-parity transformations. R-parity is a discrete symmetry relating the fermion number, baryon number and lepton number:

$$R_p = (-1)^R = (-1)^{F+3B+L}$$

Requiring R-parity symmetry in supersymmetric models avoids proton decays, which are not experimentally observed. SM particles have $R_p = +1$ whereas the superpartners have $R_p = -1$, which means that supersymmetric particles can only be pair-produced at SM particle colliders, until a pair of superpartners remains: the lightest supersymmetric particle (LSP). This pair is stable, because any further decays into SM particles with opposite R_p are kinematically forbidden. In many models, the LSP is the lightest neutralino, but the gravitino has also been proposed in supersymmetric models containing gravity, and in either case the stable LSP is an excellent dark matter candidate. Different incarnations of supersymmetry treat R-parity conservation and R-parity breaking in different ways, but in each case it is a useful symmetry, with the important phenomenological consequence that supersymmetric particles must appear at the interaction vertex in pairs.

2.5 Further advantages of supersymmetry

Many advantages of supersymmetry relate to its simplification of the SM, extending the symmetries to place tighter constraints on calculations and produce uncomplicated solutions. Calculating quantum corrections is simpler, since the corrections to the Higgs mass are no longer quadratically divergent in the cutoff scale; the Higgs mass renormalises multiplicatively, but not additively, due to cancellations between corrections from SM and superpartner particles. New states at the Planck scale would not impact the Higgs mass, so supersymmetry is a hierarchy enabler. There is no longer a need to bring the new mass scales down towards the weak scale to achieve a small hierarchy, because supersymmetry permits interactions between particles of such different mass scales.

Superpartners also enable the unification of the gauge couplings, which is a coveted feature in BSM theories that cannot be achieved in the SM, and is one of the original perceived benefits motivating supersymmetry. The renormalisation group in the MSSM can cause the three couplings of the gauge interactions to coincide at a scale $M_U \sim 1.5 \times 10^{16}$ GeV [20, 31]. The approximate unification at M_U is sometimes considered evidence to support a Grand Unified Theory (GUT) or supersymmetric models of string theory. Within the bounds of group theory typically used in GUTs, the MSSM can achieve this unification much more precisely than the SM. Although no convincing unification so far can achieve absolute precision, several studies have shown that the small differences that remain in the couplings at this scale can be accounted for with supersymmetric threshold corrections [32]. Where GUTs incorporating the SM theory are able to produce reasonable unification between the gauge couplings, replacing the SM with the MSSM can achieve a stronger result.

2.6 Searching for supersymmetry

The search for supersymmetry signals has so far been confined to setting limits, which are mostly based on simplified models and pre-defined search regions. Every iteration of the MSSM has a large number of free parameters, but we can narrow the search by constraining these parameters. Typically, theories contain tree-level relations between the quartic Higgs couplings and the other couplings, which are useful for predicting the masses of the MSSM Higgs particles. These Higgs masses are sensitive to all the particles in the model (when radiative corrections are included), which constrains the parameter space even before any supersymmetric particles are discovered. The Higgs mass in the SM has been measured to high precision at the LHC from different decay modes, originally $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{*} \rightarrow 4l$, and is now known to the level of a few hundred MeV [33]. The precision of the SM Higgs mass measurement has consequences for the parameters of the MSSM: if supersymmetry is detected, then the mass and couplings of the SM Higgs boson will help to constrain those MSSM parameters that may be too large to access directly in experiments. In the likely scenario where some superpartner masses are larger than 1 TeV, the parameters of the Higgs will provide these constraints by acting as precision observables. Notably, the theoretical predictions of the Higgs mass should be increased to match the precision of experiment, including two-loop effects.

Measurements at the LHC

Supersymmetry could provide the high-energy theory that will allow us to comprehensively solve the remaining problems with the SM and perhaps to calculate a composite Higgs mass from more fundamental parameters. We explore possible supersymmetry scenarios experimentally both by probing higher energies directly and by measuring rare or forbidden processes in the SM indirectly by targeting the contributions from supersymmetric particle loops. To discover the nature of the true extension, which remains likely to involve some iteration of supersymmetry, high energy experiments must operate at or beyond the energy scale of the SM. The LHC was designed to operate at a maximum centre-of-mass energy of 14 TeV, but could be sensitive to the corrections from more microscopic processes contributing higher energy degrees of freedom. In direct search efforts, the stop quark in particular is a likely candidate for discovery due to its low mass relative to other sparticles in natural supersymmetry models. Third generation squark searches are motivated by the possibility that supersymmetry can be considered natural with only these sparticles accessible at LHC energies. In models with large stop mixing, only one squark may be light enough for observation [34]. If supersymmetry adds the next layer to the microscopic fundamental theory, it may be probed with the greatest sensitivity currently possible at the LHC.

If superpartners are accessible, they could be discovered directly through one of three main searches: R-parity conserving searches, R-parity violating regimes or searches for long-lived particles. These strategies are based on predictions that sparticles will be discernible from standard physics by their unique and unusual kinematic behaviour and decays. The first type focuses on pair-production proceeding through the strong or electroweak interaction, including decays to the stable LSP, characterised by high transverse momentum Standard Model decay products and large missing transverse energy (defined in Section 3.2.3). R-parity violating models instead assume the lightest supersymmetric particle decays to Standard Model particles, which may be detected through multi-jet and multi-lepton searches. The third main possibility for detecting supersymmetry is long-lived particles, in which superpartners have small couplings, or there are mass degeneracies in the spectrum, such that superpartners are produced with long lifetimes. In these cases, we may conduct searches that include displaced vertices from late decays of the LSP. With the LHC operating at 13 TeV, the search strategy is focused on superpartners with the largest predicted production cross-sections and cleanest channels. The three main search strategies dedicated to detecting evidence of superpartners at the high energy LHC are ongoing, and continue to be refined by new results.

Chapter 3

Collider experiments

International efforts to discover evidence of BSM signals, with a focus on supersymmetry, rely on high-energy collider experiments. The largest is the LHC, a circular hadron collider that accelerates beams of high-energy protons to produce showers of secondary particles whose kinematics are measured at experimental sites around the ring. Searches in the recorded data seek evidence of small signals from SM and BSM processes which may lead to more precise measurements and new discoveries.

Once data is collected, an LHC analysis proceeds according to the parameters of the target model. Generally, measured objects are calibrated, the distributions of observables are built, these observables are compared with background expectations and in BSM searches, if no deviations from the SM are discovered, the group sets limits constraining the models of new physics. The analysis group may choose one or more methods that balance between greater sensitivity and a broader search. New strategies are informed by the design capacity of the detector, the physics reach of the collider, the types of information available, the phenomenology of target physics models and any relevant existing search strategies. This chapter describes these considerations regarding the LHC and its largest experiment, ATLAS. The data considered here is simulated later in Chapters 5 and 6, and extracted directly from ATLAS for the analysis in Chapter 7.

3.1 The Large Hadron Collider

The design of the LHC was intended to maximise sensitivity to small SM and BSM signals under the constraints of its size, power of instrumentation, shape and the capabilities of other existing detectors. It is a proton-proton collider operating up to a total beam energy of 14 TeV at design capacity, superseding the Fermilab Tevatron proton-antiproton collider. The LHC design was particularly optimised for measuring the Higgs boson mass for the first time, and now for exploring further Higgs physics. The sensitivity of the experiment to new physics requires a high centre-of-mass energy, high luminosity, high bandwidth and a strong ability to identify particle signatures.

3.1.1 LHC design specifications

The LHC uses powerful dipole magnets to deflect beams of protons around a ring 27 kilometres in diameter, before colliding the beams at Interaction Points (IPs). Massive and short-lived particle states often appear in BSM physics, so for a chance to detect these states the LHC must use precision timing with a high data rate and bandwidth. The most recent data-taking period spanned 2015 to 2018 (known as Run 2), when protons were accelerated in two counter-rotating beams of 6.5 TeV each, producing a nominal combined energy of 13 TeV. Strong magnets shape and direct the proton beams; a total of 1232 superconducting dipole magnets are responsible for maintaining the beams on their circular paths, and 392 quadupole magnets focus the beams at the IPs.

Four main experiments are located at the IPs around the ring, comprising two general-purpose detectors (ATLAS and CMS), a flavour physics experiment (LHCb) and a heavy ion detector (ALICE), the positions of which are shown in Figure 3.1. The four main experiments permit collaborations to analyse vastly different sets of data simultaneously, searching in realms as different as B physics, Higgs, cosmic rays, heavy ions, supersymmetry searches and more.

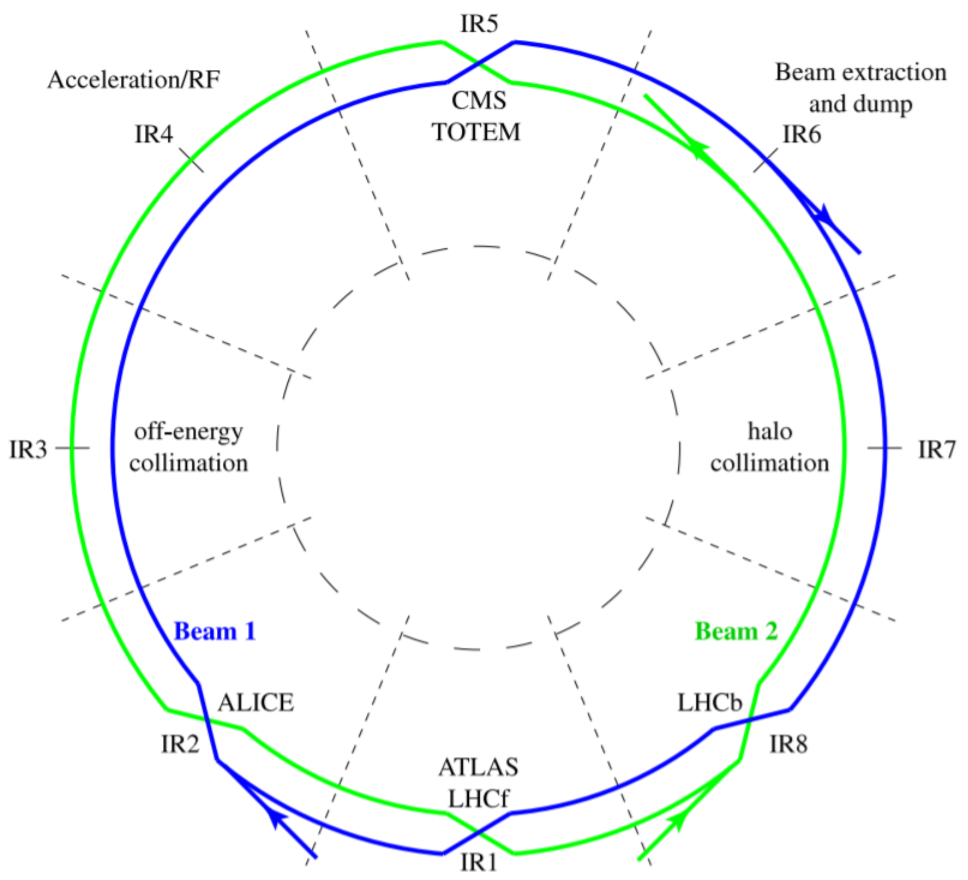


Fig. 3.1 Schematic diagram of the four main experiments located at interaction points around the LHC ring. Also identified are the eight arc sections of the ring and locations for beam collimation, acceleration and extraction [35].

The key objectives of these experiments include precise measurements of existing signatures and discoveries of new physics signatures at TeV energies. Achieving these objectives requires collisions that cover a wide energy range, for which a hadron collider design is the ideal instrument. Partly, this is because hadrons are able to achieve higher energy collisions than, for example, lepton colliders. The LHC is constructed within the tunnel that was first operated in 1989 to collide electrons and positrons as part of the Large Electron-Positron (LEP) experiment. The original lepton collider was designed to succeed in different objectives to the LHC; colliding elementary point-like particles achieved a well-defined collision energy that was suited for precision measurements. However, the LEP design also limited the maximum beam energy and collision rate compared with the LHC. Since the design capacity for beam energy at the LHC is 14 TeV, analyses can directly probe any energy up to this value.

The event rate is another consideration for achieving good significance in experimental measurements. Experiments colliding a beam of particles with their anti-particles can accommodate both oppositely-charged beams within the same vacuum chamber and magnets (as LEP did), but the advantage of this efficient design comes at the expense of a lower collision rate. Particle and anti-particle colliders are intrinsically limited by the rate at which anti-particles can be generated, so the LHC was designed to collide high-energy proton beams with a two ring design containing separate vacuum and magnet systems in each ring. The protons collide in bunches containing a certain number of particles each. In Run 2, the LHC was colliding 2808 bunches per proton beam, beginning the run with 1.2×10^{11} protons per bunch [36]. Per second, the number of collisions was approximately 1 billion. Such a high collision rate was made possible by the two-ring design of the LHC.

The amount of data produced in these collisions is vast and demands significant effort for identifying which is “interesting” to be stored for analysis and which is irrelevant and discarded. When beam particles scatter, information about the products of this interaction may be pieced together and stored as an *event*. Relevant information is selected by the trigger system, which stores an event if it is interesting enough to have been listed in the trigger menu which was decided in advance of data taking. Circulating selected events internationally for analysis is managed by data sharing on the Worldwide LHC Computing Grid (WLCG), which provides computing resources for storing, distributing and analysing many petabytes of data. The global collaborations connected to the WLCG comprise around 170 computing centres spread across more than 40 countries, all requiring high-statistics measurements to perform searches for interesting new signatures.

Luminosity and pileup

Luminosity is a main measure of the performance of a collider. The instantaneous luminosity (L) of the LHC is a measure of the collision rate at the IPs, and is determined by the geometry of the apparatus and the particle flux. It has dimensions of $(\text{area})^{-1}(\text{time})^{-1}$, typically $\text{cm}^{-2}\text{s}^{-1}$. An increase in luminosity equates to a higher rate of event production per unit area and time, producing higher statistics for rarer processes. A useful calculation for experiments is the integral of luminosity over time, called *integrated luminosity*, which indicates the number of collisions that occurred over the time interval. Integrated luminosity is measured in units of inverse femtobarns (fb^{-1}), where the barn is a measure of area such that 1 b $\equiv 10^{-24}\text{cm}^2 = 10^{-28}\text{m}^2$. If the time interval for integration is $t - t_0$, then the integrated luminosity is:

$$L_{int} = \int_{t_0}^{t} L(\tau)\, d\tau \tag{3.1}$$

Since the beam sizes and bunch intensities of the LHC vary with time, the integrated luminosity measures the total number of events possibly generated in $t - t_0$, which is a useful indicator of the performance of the LHC and its sub-detectors in a given run. At the end of Run 2, the LHC had delivered a total integrated luminosity of 156 fb^{-1}, of which 139 fb^{-1} can be used in physics analysis, where the increase in luminosity over time is shown in Figure 3.2. The ATLAS experiment records around 95.5% of the luminosity delivered by the LHC.

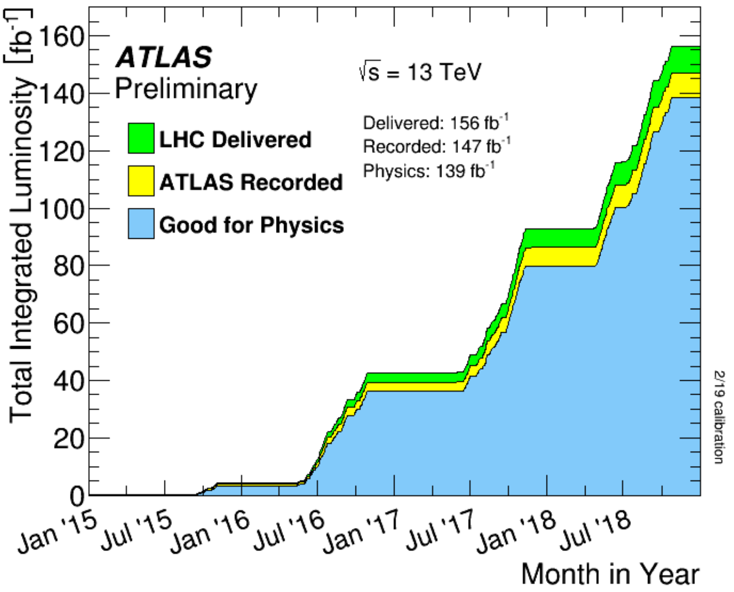


Fig. 3.2 Total integrated luminosity of the LHC over the Run 2 data-taking period comprising the years 2015-2018 [37].

To understand how the luminosity relates to bunch crossing statistics, consider the luminosity for head-on collisions in Equation 3.2. Luminosity depends on several parameters: N_1 and N_2 are the numbers of particles per bunch in beams 1 and 2 and n_b is the number of bunches, while f_{rev} is the frequency of revolutions. The effective overlap cross-section between the two beams at the IP is denoted A.

$$L = \frac{N_1 N_2 n_b f_{rev}}{A} \tag{3.2}$$

If Equation 3.2 is used to calculate the Run 2 luminosity of the ATLAS Experiment, it must first be manipulated to accommodate the non-zero crossing angle between the proton beams at the LHC. To understand how it obtains angular dependence when collisions are no longer head-on, consider Equation 3.3. At the LHC, the same number of protons per bunch is injected into both beams, so $N_1 = N_2 = N$. The number of bunches per beam is $n_b = 2808$, which is multiplied by the revolution frequency $f_{rev} \approx 11.245$ kHz and divided by the product of the areas of the two Gaussian beams at the IP, σ_x and σ_y, each of which is valued around $10-20$ μm [38].

$$L = \frac{N^2 n_b f_{rev}}{4\pi\sigma_x^* \sigma_y^*} F \tag{3.3}$$

The product of beam areas can be written in terms of useful parameters: the normalised emittance ε_n, the β function and the relativistic Lorentz factor γ proportional to beam energy.

$$\sigma_x^* \sigma_y^* = \frac{\varepsilon_n \beta^*}{\gamma} \tag{3.4}$$

where the β function is the envelope of proton oscillations, and β^* denotes its minimum value. These oscillations are called betatron oscillations, and they depend on the configuration of magnets in the apparatus which are designed to minimise the β function at the IP where the beams are focused. In the ATLAS experiment, focusing is achieved using a triplet of quadrupole magnets. The emittance ε is the cross-sectional area of the beam, approximated as an ellipse that contains a defined percentage of the particle phase space. Magnetic fields change the shape of the ellipse, but leave its area unchanged. The area ε is controlled instead by a process called adiabatic damping, which relies on the property that the beam cross-section decreases as the acceleration of the beam increases. We compare the emittances at different energies by calculating the normalised cross-sectional area (ε_n).

The geometric reduction factor F is valued between 0.3 and 1, so this fraction decreases the luminosity by a value that depends on the non-zero crossing angle between the interacting beams. For beams that are directed at different angles instead of aligned head-on, the luminosity is reduced by a larger fraction. The effect of geometric reduction is controlled in the experiment apparatus by using crab cavities to rotate the beam angle. Equation 3.5 describes the geometric reduction factor as a function of crossing angle (ϕ):

$$F = \frac{1}{\sqrt{1 + (\frac{\sigma_s}{\sigma_t}\frac{\phi}{2})^2}} \tag{3.5}$$

where the reduction factor depends inversely on the ratio of longitudinal bunch length σ_s and transverse bunch size σ_t. As the crossing angle increases, the reduction factor decreases, so collisions with a large transverse offset or crossing angle reduce the instantaneous luminosity.

The LHC is designed for a high luminosity to produce high statistics analyses, but the collision rate and detection is limited by several factors, including hardware performance, detector effects and collective effects such as beam instabilities. It is maximised by:

- Optimising the overlap area between the beams at the IPs,
- Minimising the size of the beams at the IPs,
- Increasing the number of particles per bunch, and
- Increasing the number of bunches.

A difficulty of achieving high luminosity is the increase in the number of simultaneous interactions per beam crossing, called *pile-up*. Simultaneous interactions are difficult to extract, as the overlap between detections confuses event reconstruction, a process described later in Section 3.2.3. An experiment may choose to separate the beams to decrease pile-up using a technique known as luminosity leveling, which reduces the luminosity on purpose when it is naturally high and then ceases to temper beam interaction at the time when luminosity would naturally decay. To achieve this effect, the experiment may begin with a high value of the beta function at the IP before decreasing β^* when there are fewer protons left in the beam to spread the rate of interactions more evenly over time.

The LHC therefore functions optimally when the distribution of collisions over time is constant, to avoid large pile-up, and the total number of collisions over time is high. Using the integrated luminosity (Equation 3.1), we can calculate the number of events resulting from a particular physics process N within a time period by multiplying by the cross-section as a function of the total beam energy of the LHC [39].

$$N = \sigma(\sqrt{s}) \times \int_{t_0}^{t} L(\tau)\, d\tau \tag{3.6}$$

The ATLAS and CMS experiments are the highest luminosity experiments, requiring $L > 10^{34}$ cm^{-2}s^{-1}, whereas LHCb requires a medium luminosity of $L \sim 10^{32}$ cm^{-2}s^{-1}, and the ALICE experiment for ion collisions requires low luminosities $L \sim 10^{29}$ cm^{-2}s^{-1}. The ATLAS experiment has a high design luminosity to permit the search for many benchmark processes with small predicted rates of production, including Higgs boson production and decay, and new physics scenarios such as supersymmetry.

LHC magnets

The main magnets on the LHC must be cryogenically cooled to temperatures near absolute zero, operating at 1.9K. The cryo-cooling capacity is affected by effects such as the heat load, power and impedance of other systems, including beam-induced heating. All magnets in the LHC are electromagnets, in which the magnetic field is produced by the flow of electric current. Superconducting materials, which conduct electricity without resistance, avoid overheating in the high-current (11,850 amp) magnet coils [39]. Superconductivity is achieved through the cryogenic systems where niobium-titanium wires in the coils are maintained at low enough temperatures to reach a superconducting state using a closed liquid-helium circuit. The magnets require 120 tonnes of liquid helium flowing through this circuit to maintain their temperature. The helium refrigeration capacity influences the ability of the magnets to accelerate and control the proton beams.

To correctly control the beams at the experimental collision sites, maintaining the magnets is a primary concern. Early in Run II, fast losses were observed in this system, in some cases requiring beam dumps where the energy of the protons was absorbed to control the beams [40]. Mitigating the losses, the collision points were shifted by “orbit bumps”, but the lossy object and the orbit bump remained for the full Run II. The maximum beam current was also limited by other losses, limiting the total intensity per proton bunch in the beams. The cause in this case was a vacuum leak following the exchange of a magnet. Although this limited the bunch intensity during 2017 and 2018, luminosity was increased by using beams inducing a low heat-load with small emittance. The luminosity did not become severely compromised by detector losses, and high quality functionality was maintained.

3.1.2 LHC physics capabilities

Proton beams in the LHC are collided with high enough energies to generate interesting results in the required centre-of-mass energy range for new physics processes. Supersymmetry and other BSM scenarios may be produced within the current energy range, or they may appear as contributions from higher energy processes. Understanding the sources of proton collisions, the phenomenology of the target physics scenarios and the range of signatures possible within our current energy capacity enables optimisation for future discoveries. Planned increases to the LHC integrated luminosity suggest that if the correct BSM scenario is supersymmetry described in Chapter 2 then the LHC has a high chance of detecting direct evidence for it within the next few years.

Proton collisions

Accelerating protons to such high energies causes them to behave as groups of unbound partons. The partons usually interact during collisions in hard scattering processes where the remaining partons are scattered away or interact at lower energies. In LHC collisions, the interacting partons are not restricted to the valence quarks of the proton (two up quarks and a down), but also include a sea of quarks and gluons within the proton that arise from QCD interactions. Each of these particles carries some fraction of the total beam energy, according to the parton distribution function (PDF) model.

The initial parton momenta in hadron collider experiments are fundamentally unknown, but could take a range of possible values as described by their PDFs. The collisions at the LHC do not exclusively have centre-of-mass energy at the limit of the LHC energy, but instead span the range of energies up to the LHC energy, where the hard scattering interaction is accompanied by multi-parton interactions and radiative corrections to the initial and final states coming from gluons and photons, amongst other processes. These corrections are called initial and final state radiation (ISR and FSR), and are detected at the LHC and included in analyses. The range of effects and energies means that a proton beam energy far in excess of 1 TeV is necessary to probe physics on the TeV scale, motivating the nominal LHC design energies of 7 TeV per beam [38].

Production of supersymmetric particles

At hadron colliders, superpartners can be produced in pairs from parton collisions of electroweak or QCD strength. If R-parity is conserved, then supersymmetric particles are always pair-produced and never appear in isolation, as per Section 2.4.5. The products include charginos ($\tilde{\chi}^{+}_{i=1,2}$), neutralinos ($\tilde{\chi}^{0}_{i=1,2,3,4}$), sleptons ($\tilde{l}$) and sneutrinos ($\tilde{\nu}$), in the electroweak sector. In the strong sector, reactions produce gluinos ($\tilde{g}$), squarks ($\tilde{q}$) and anti-squarks ($\tilde{q}^{*}$). Production mechanisms of electroweakinos include [20]:

$$\begin{aligned} q\bar{q} &\to \tilde{\chi}_i^+\tilde{\chi}_j^-,\ \tilde{\chi}_i^0\tilde{\chi}_j^0, \quad u\bar{d} \to \tilde{\chi}_i^+\tilde{\chi}_j^0, \quad d\bar{u} \to \tilde{\chi}_i^-\tilde{\chi}_j^0, \\ q\bar{q} &\to \tilde{l}_i^+\tilde{l}_j^-,\ \tilde{\nu}_l\tilde{\nu}_l^*, \quad u\bar{d} \to \tilde{l}_L^+\tilde{\nu}_l, \quad d\bar{u} \to \tilde{l}_L^-\tilde{\nu}_l^*, \end{aligned} \tag{3.7}$$

and QCD strength reactions involve the following supersymmetric products:

$$\begin{aligned} gg &\to \tilde{g}\tilde{g},\ \tilde{q}_i\tilde{q}_j^*, \\ gq &\to \tilde{g}\tilde{q}_i, \\ q\bar{q} &\to \tilde{g}\tilde{g},\ \tilde{q}_i\tilde{q}_j^*, \\ qq &\to \tilde{q}_i\tilde{q}_j. \end{aligned} \tag{3.8}$$

where i, j are the mass eigenstate indices of the charginos and neutralinos, and the generation indices of the sleptons and squarks.

LHC supersymmetry signals are always an inclusive combination of all types of parton collisions, but generally the LHC mostly operates as a gluon-gluon and gluon-quark collider instead of a quark-antiquark collider, because it collides two proton beams instead of colliding protons with antiprotons. QCD strength reactions already dominate over electroweak strength processes, so the most likely reactions are gluon-gluon and gluon-quark fusion in Equation 3.8. Of these, stop quark (top squark) production may be amongst the most common, given our current knowledge of the allowed masses of the other squarks and gluinos.

Predictions suggest that parton scattering produces supersymmetric particles very rarely, so signals are scarce. Scattering cross-sections, which quantify the rate of occurrence, are measured in units of barn common to the inverse femtobarn units for integrated luminosity introduced in Section 3.1.1. The total proton-proton scattering cross-section, ignoring its weak (logarithmic) energy dependence, is around $\sigma \sim \pi r_p^2 = 3\ \text{fm}^2 = 30$ mb, and the cross-section of any process of interest will be far smaller [38]. For example, the typical scale for weak interactions is Fermi's constant, $G_F \approx 1$ nb, which indicates that the rate of W boson production is around 1 per million proton collisions. If supersymmetric particles are produced in the collisions, then they have smaller cross-sections by several orders of magnitude [41]. At the LHC, the largest supersymmetric cross-sections belong to the QCD strength reactions described above, followed by lower cross-sections predicted for processes with the associated production of a chargino or neutralino with a squark or gluino. Their small predicted cross-sections motivate the high luminosity of the LHC, which should permit their discovery.

In addition to the limitations produced by small cross-sections on the maximum amount of available data for BSM processes, interference from similar physics processes further limits the number of available events. The event rate is reduced by the need to select final states that have good resolution and are distinct from as many SM background events as possible. If the optimal final states by these criteria are the states with small *branching ratio*, which is the fraction of time for which a particle decays to this state, then the event rate is severely limited. For example, during Run 2, the LHC produced almost 8 million Higgs bosons, but the number of selected events was far smaller for these reasons. Similar constraints on physics models and final states apply in supersymmetry searches.

Collider reach

The LHC must not only record enough events for analyses to isolate supersymmetry signals, but also must operate at energies higher than the superpartner masses. These masses determine the discovery potential, where the current collision energy of the LHC at $\sqrt{s} = 13$ TeV has a certain reach in a chosen discovery channel for a given model, defining upper bounds on the superpartner masses below which they could be observed with high enough statistical accuracy to be confirmed a discovery. There are no stringent limits on the degree of fine-tuning that is acceptable in a natural model of supersymmetry, so the overlap between the reach of the LHC and the possible parameter space of natural supersymmetry is somewhat arbitrary. Even so, when natural models are defined by a chosen amount of fine-tuning to have quantum corrections to the Higgs mass that fall below some limits, the range of possible superpartner masses is also bounded above, revealing the collider energy required to produce them experimentally.

In Figure 3.3 the region below the grey band is not considered finely-tuned, and the upper region where the gluino mass is greater than 6 TeV belongs to a finely-tuned model. Run 2 data from the 13 TeV LHC can exclude the green region with 95% confidence level, and is unable to reach the remaining part of the allowed mass range. Natural supersymmetry defined by these limits requires a hadron collider like the proposed High-Energy-LHC (HE-LHC) with $\sqrt{s} \sim 33$ TeV and 1000 fb^{-1} to either discover or falsify its existence over the full mass range, where the top squark signal will be accessible even if the gluino is out of reach [42]. However, the superpartner masses would ideally fall below these upper limits, because this reduces fine-tuning. For making discoveries in the lower mass range, the Run 2 dataset is ideal.

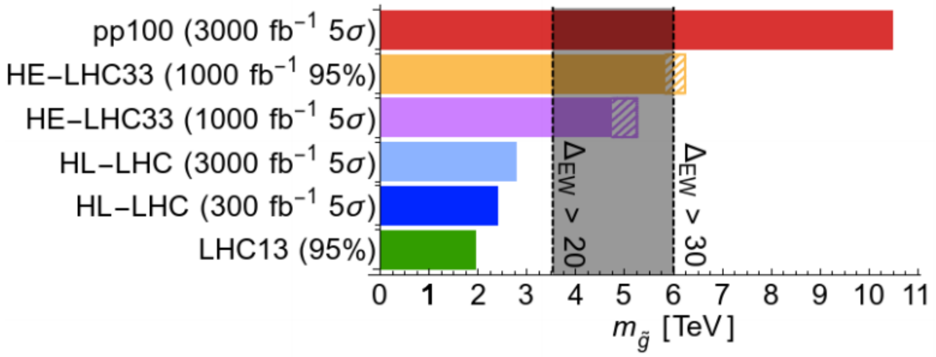


Fig. 3.3 The level of fine-tuning of the gluino mass required for its detection in current and possible future experiments [43].

The effective reach of experiments can be increased through a strategic choice of analysis techniques. Examples of strategies that search for signatures outside of the current energy range include precision measurements of contact interactions, and indirect searches for particles produced in quantum loops. Indirect searches are not limited to the LHC centre-of-mass energy, because energy conservation can be violated for a very short time in quantum loops. In considering possibilities for future colliders, direct and indirect searches for new physics complement each other: an excess in indirect searches informs the decision to build a new direct detector that will target relevant processes. Alongside targeted strategies, developing model-agnostic approaches and testing them on LHC Run-2 data is vital preparation for future discoveries. Such global fits allowing sparticle decays to be more complex than those encountered in simplified models have previously found ample parameter space for light squarks [44–48]. Testing old and new strategies on current data, in searches that reach as much of the new physics parameter space as possible, is important both to discover evidence of new physics in current data and to design future experiments for the coming years.

3.2 Detection with ATLAS

The ATLAS experiment operated with successful detector performance throughout Run 2 compared with design expectations. In spite of the challenging conditions produced by recording an unprecedented integrated luminosity of almost 96% of the LHC total, the data quality was high. Considering complications arising from pile-up, which amounted to 10-70 interactions per bunch crossing, performance was outstanding and all subdetectors ran smoothly with high data taking efficiencies.

High performance of the ATLAS detector in Run 2 was intended to achieve excellent lepton and photon resolution, good jet and missing energy resolution and standalone muon measurement, amongst other benefits [49]. These focus points enabled ATLAS to precisely measure the results in Figure 3.4, which shows a summary plot of the progress in measuring production cross-sections of the SM processes from the end of Run 2 in 2020. Cross-sections are measured in picobarns and are fiducial, meaning that they are calculated for a subset of phase space in which the process signatures are visible because the detector is sensitive to them in this region.

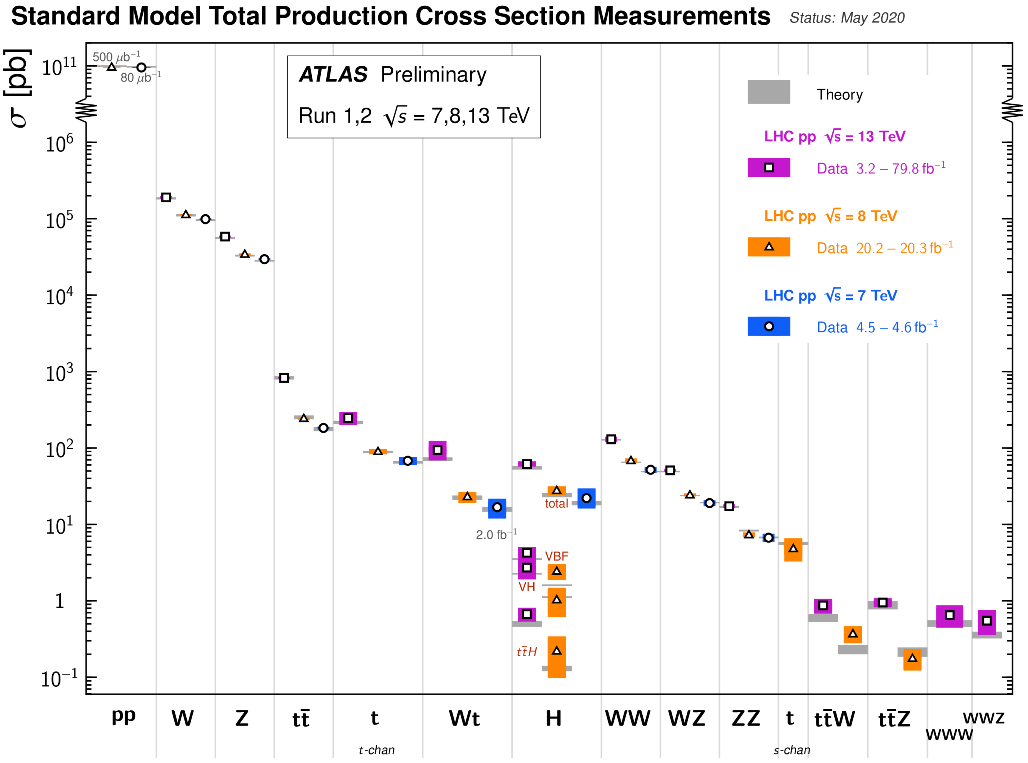


Fig. 3.4 Fiducial production cross-section measurements of SM processes compared with theoretical expectations calculated with next-to-leading order (NLO) or higher loop corrections [50].

The overlap between data and theory in Figure 3.4 represents strong agreement between experimental measurements with the ATLAS detector and the predicted values of each process from the SM. Cross-sections are measured in picobarns, and the dark-colour error bar represents statistical sources of uncertainty related to the available number of detected events in the analysis, while the light-colour bar represents the full uncertainty, including systematic effects. Processes with a high cross-section are clearly easier to measure precisely and in general do not require such a high luminosity to produce clear results as the processes lower in the table. These results have since been further improved by additional analysis, but the summary remains strong evidence of the detector (and analysis) performance during Run 2.

3.2.1 ATLAS detector design & setup

The ATLAS and CMS general-purpose detectors are the largest at the LHC and have the same basic design. In ATLAS, an inner detector close to the beamline measures tracks of charged particles, and is followed by a liquid argon (LAr) electromagnetic (EM) calorimeter which mostly measures electron and photon energy. Further from the beamline, a hadronic calorimeter made of plastic scintillator tiles and iron measures the energies of protons and neutrons, and outside of this is a muon detector. The inner detector tracking system is made up of silicon pixel detectors close to the beamline, surrounded by a silicon semiconductor tracker and a transition radiation tracker. The ATLAS detector uses magnets in several capacities, both toroidal and solenoidal, with a 2 Tesla strong magnetic field able to deflect and focus the proton beams [51].

The energy deposited in the hadronic and EM calorimeters by particles passing through the material is converted to electrical signals. Consequently, those particles that leave tracks are identifiable through the length and curvature of their trajectories. In Figure 3.5, the locations and relative sizes of detector components are visible, including the semiconductor and transition radiation trackers (SCT and TRT), which together with the pixel detector make up the three main components of the inner detector. Respectively, these three comprise a silicon microstrip tracker with 60 m^2 of silicon distributed over several barrel layers and endcap discs (SCT), straw tubes with gold-plated tungsten wire in the centre providing additional information on particle type with the ability to distinguish electrons from pions (TRT), and 80 million pixels (pixel detector) [52, 53]. The three together provide precise measurement of the direction, momentum and charge of electrically-charged particles.

The overall size of the ATLAS detector is dominated by the muon spectrometer, which accommodates large monitored drift tube (MDT) chambers composed of aluminium tubes designed to measure the curvature of the muons. Muons typically pass undetected through the inner detector and calorimeter, and then leave long paths in the MDT chambers because they are highly energetic at around 200 times the mass of the electron. They are tracked as they deposit energy in the spectrometer, but are too energetic for the detector to stop their motion, so their total momentum is measured instead from the curvature of their helical trajectories in the strong solenoidal magnetic field. The apparatus also measures their precise coordinates as they leave the detector using cathode strip chambers.

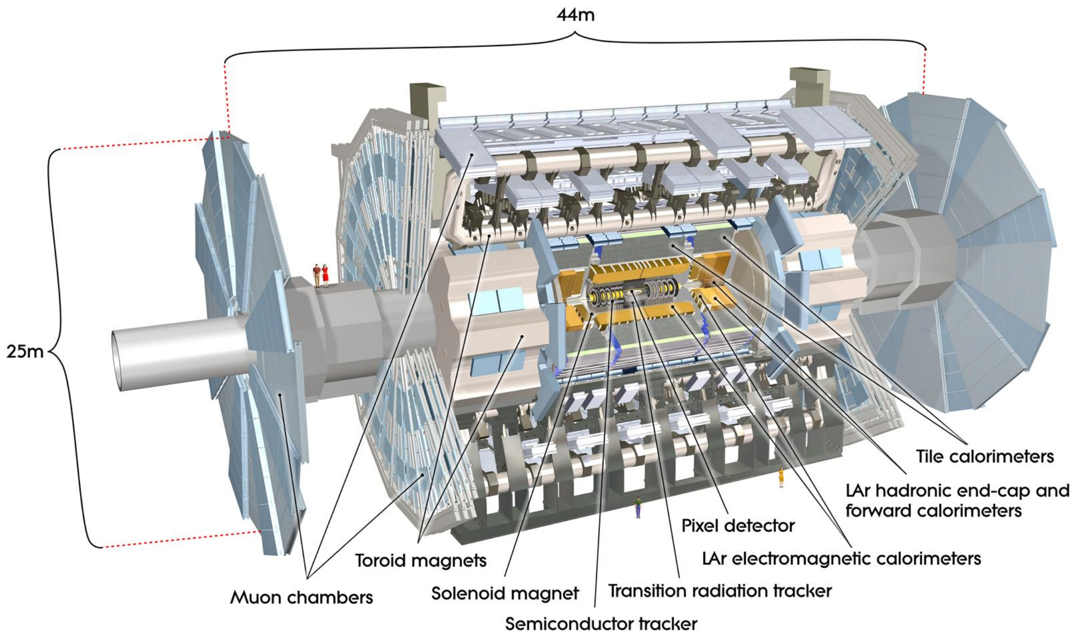


Fig. 3.5 A cutaway diagram of the ATLAS detector identifying the detector components, including hadronic and EM calorimeters and the large muon chamber [54].

The aim of the above combination of detector components is to precisely track and measure the energy and momenta of as many identifiable products as possible from hard scattering reactions and subsequent decays. When a hard scattering process takes place, the scattered particles move off at transverse directions to the beams. For clear results, we seek quantities that are measured in the frame of reference of the scattered products. The centre-of-mass frame of these particles is not in general the same as the lab frame in which measurements are made, so we choose measurements which have the same values in the lab and partonic centre-of-mass frames. These quantities are invariant under Lorentz boosts along the beam line, or the z- (longitudinal) direction, and may be derived from various levels of the detector system. Observables are functions of these kinematic measurements, revealing distinguishing characteristics of the collision products. These are combined to produce additional variables for selecting data and performing analysis.

First, discerning a comprehensive picture of particle kinematics is vital for creating useful variables. The momenta of scattering partons are fractions of the momenta of colliding protons, as described in Section 3.1.2. According to the parton model, the momentum fractions of partons colliding in a hard scattering process from within two different protons are distributed probabilistically and independently. Therefore, the probability of a parton in one proton found with momentum fraction x_1 of the total momentum P_1^μ of the first proton is independent of a parton found with momentum fraction x_2 of the second proton P_2^μ. In this case, the parton momenta p_1 and p_2 are written:

$$p_1^\mu = x_1 P_1^\mu, \quad p_2^\mu = x_2 P_2^\mu \tag{3.9}$$

Since the colliding partons carry different fractions of longitudinal momentum, the rest frames of the parton collisions have different longitudinal boosts.

Applying a Lorentz boost along the longitudinal direction reveals which components of energy and momentum are boost invariant and therefore particularly interesting for analysis. Under a Lorentz boost parameterised by β, the components of a parton four-momentum $p^\mu = (E, p_x, p_y, p_z)$ transform as:

$$E \rightarrow E\cosh\beta + p_z\sinh\beta, \quad p_x \rightarrow p_x, \quad p_y \rightarrow p_y, \quad p_z \rightarrow p_z\cosh\beta + E\sinh\beta \tag{3.10}$$

The transverse momenta p_x and p_y give the boost invariant vector and scalar quantities, $\vec{p}_\mathrm{T} \equiv (p_x, p_y)$ and $p_\mathrm{T} \equiv |p_\mathrm{T}|$, which are typical in analyses, and are useful for calculating more sophisticated variables. By remaining in the parton centre-of-mass frame, boost invariant variables are useful for disentangling the outcomes of the hard collisions from the protons that produced them. Using this method, we can produce final collision events characterised by kinematic attributes, amassing all the discernable measurements and piecing them together into physics objects by the reconstruction process. As part of this process, the ATLAS tracking system relies on a boost-invariant coordinate system to distinguish collision products, also calculated from the transverse momenta.

The coordinate system

The trajectories of scattered particles and decay products are tracked through the detector components in an angular coordinate system. The system is based on the azimuthal ϕ and polar θ angles standard in cylindrical coordinate systems, but since the polar angle is not boost invariant, further calculations are useful. The azimuthal angle is calculated from the components of transverse momentum:

$$\phi \equiv \tan^{-1}\frac{p_x}{p_y} \tag{3.11}$$

and is therefore already a boost invariant quantity. An additional component of the trajectory coordinates is *rapidity* y, which is a function of the energy and z-component of momentum.

$$y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z}$$

The difference of two rapidities for two momenta is boost invariant, so the difference Δy is part of the definition of *angular separation* ΔR, a useful variable for distinguishing particles produced in the hard collision.

$$\Delta R = \sqrt{(\Delta y)^2 + (\Delta\phi)^2} \tag{3.12}$$

Whereas rapidity is a kinematic quantity defined from components of the four-momentum vector, and therefore so is angular separation, there is an additional quantity which is entirely geometric known as the pseudorapidity η. Pseudorapidity is a function of polar angle θ:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right)$$

where massless particles have $\eta = y$. However, massive particles do not have the property that the difference between their pseudorapidities is boost invariant, unlike the rapidity difference Δy. A standard definition of angular separation uses the pseudorapidity difference $\Delta\eta$ in place of rapidity difference, as in Equation 3.13.

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{3.13}$$

The angular coordinate system composed of ϕ, θ and η for a cylindrical detector apparatus is represented in Figure 3.6.

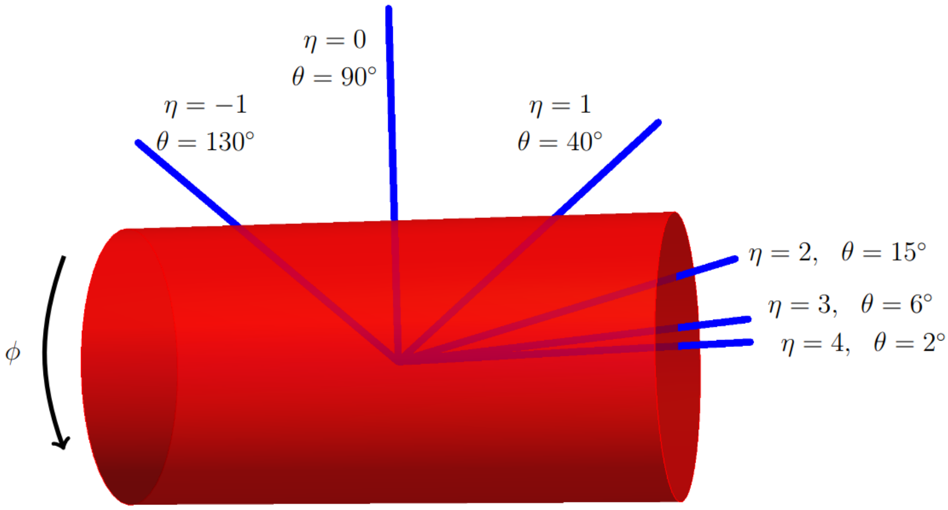


Fig. 3.6 The coordinate system of the ATLAS detector, comprising pseudorapidity η and azimuthal angle ϕ [38].

A pseudorapidity of 0 corresponds to a particle moving with no longitudinal component directly away from the beam line. In contrast, particles with $\eta \sim \pm 5$ are moving effectively along the beam line. The ATLAS detector calorimeters cover the region $|\eta| < 4.9$, and its charged particle trackers cover $|\eta| < 2.5$.

Detector capabilities

Certain particles have properties that cannot be measured with the ATLAS detector design. The calorimeters measure the energy and direction (η, ϕ) of particles that are stable on detector timescales, which does not include heavy particles such as the tau lepton because they decay too quickly to be observed. Neutrinos are stable, but they are also undetected because they do not interact with the calorimetry system. Because neutrinos are invisible, their momenta cannot be discerned if there is more than one produced in the same collision.

3.2.2 Data acquisition & the trigger system

The data available for analysis is restricted to fill the available storage space, retaining only the most interesting physics events. The data acquisition and trigger systems record signals and decide which regions of the geometric and kinematic space are retained. Data acquisition

(DAQ) refers to the programmable software used to measure electrical signals from the detector apparatus and convert them to digital. On the order of 100 million channels acquire data from ATLAS, contributing around 1 MB of raw data per measurement, recorded every 25 ns [55]. A small percentage of this total is chosen by a trigger system that selects events in real-time, where "interesting" physics such as electroweak and top physics starts at 6-8 orders of magnitude rarer than a typical collision. Trigger and DAQ systems always introduce some inefficiencies, which may be relevant for the physics program, although the ATLAS experiment in Run 2 employed a highly efficient trigger system.

The ATLAS experiment trigger is two-level, comprising a level-1 hardware system and a high-level software-based system. They operate around 1000 triggers during data-taking and together filter the 40 MHz crossing frequency of the LHC down to 1 kHz of physics collisions [56]. Pre-determined choices are compiled into a trigger menu, which specifies not only the selection algorithms but also their allocated rate and bandwidth. The trigger menu can be flexible to adjustments, following the priorities determined by the physics strategy; for example, the main triggers for supersymmetry searches involve generic jet, missing transverse energy and lepton selections. Sufficient events must be acquired belonging to all the relevant search classes. Sometimes, more events of interest are recorded than necessary or allowed by the bandwidth for their type. In this case, a pre-scale factor of N selects one event out of every N events of its type fulfilling the requirements, discarding the remainder. Regardless of how many events are detected, two factors are necessary for selections: the instantaneous luminosity of the LHC, and the limitations of the ATLAS detector readout. Both factors inform the trigger menu and are vital to overcoming challenges in the system. A major challenge for the ATLAS trigger group in 2016 was to maintain trigger performance with increasing luminosity and pile-up conditions, and measuring trigger efficiency is an important part of the experiment.

The ATLAS level-1 trigger hardware system is fed mostly by input from the calorimeter and muon detector components. Within regions of interest up to the limits of the detector hardware, the combined system makes a selection from the set of objects identified as muons, jets, taus and clusters. Its *topological* trigger component receives combined information from two other parts, the calorimeter muon triggers, which the topological trigger uses to calculate variables for further triggering. Events that are accepted by the level-1 trigger are passed to the High-Level Trigger (HLT), the software-based system that reconstructs them with greater precision.

3.2.3 Event reconstruction

Reconstruction algorithms translate electronic signatures from the trackers, calorimeters and muon chambers into the type, four-momentum and charge of the physics that deposited them. These processes occur inside both the high-level trigger and the computing Grid. They inform physics analysis of a particle's type, location of production and the type of its parent particles. The answers are determined by the combination of at least three classes of reconstruction, including tracking, vertexing (grouping particles into vertices) and particle identification (ID) of each track.

Invisible particles carry away missing energy from the system, but are identified from momentum conservation and theoretical knowledge of processes that involve the visible detected particles. Although precise outcomes from scattering are impossible to predict, expectations from probability classify which reactions produced the detected particles to within a certain confidence level. Additional discrepancies may arise from detector resolution and from imperfect reconstruction efficiencies.

The objects reconstructed from the measurements of particles shown in Figure 3.7 can be quite different from the original particles. For example, when the proton and neutron pass into the hadronic calorimeter, they leave a highly collimated cone of hadrons which have their tracks reconstructed as a jet object. The presence of a jet in the assembly of reconstructed objects indicates that charged or neutral hadrons were likely formed from a quark or gluon, although there is no way to uniquely identify their origin.

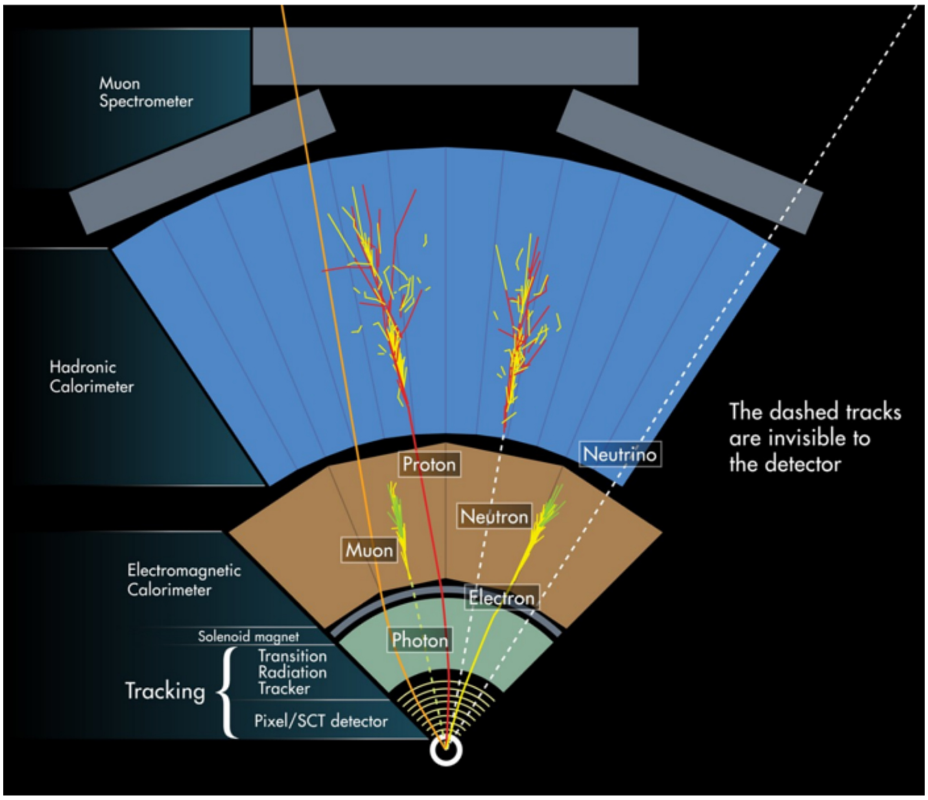


Fig. 3.7 Schematic diagram of identifiable particles detected in different layers of the detector. The hadrons produce jets in the hadronic calorimeter, which are later reconstructed alongside the muon, electron and photon [57].

Detector data is complicated for several further reasons, especially because it involves products from many different parton collisions, where multiple scattering reactions produce indistinct final state particles. Background processes are those which appear similar to the state of interest and must be removed through analysis to clean up the signal. The two causes of background processes in an LHC dataset are:

1. different reactions that produce the same physics objects, and
2. incorrect classification of objects because they mimic the properties of other particles.

These produce two different types of background process, the first imposing a hard limit which cannot be improved by perfecting the reconstruction and the second resulting from performance inefficiencies. The two types are:

1. irreducible backgrounds, with the same final state as the signal, and

2. reducible backgrounds, with a different final state.

The reducible backgrounds are caused by imperfectly identifying the type of reconstructed particles to produce “fakes”. Often, fakes include apparent jets which are actually electrons, or charged pions appearing as electrons. A lower fake rate for a particular type of particle can be achieved by tightening the reconstruction requirements, therefore increasing the proportion of “identified” signatures that were truly left by the desired particle. The reconstruction requirements are tuned by applying quality cuts, using combinations of tracker and calorimeter information to reject data from regions where the detector was malfunctioning.

Another feature that contributes to reducible backgrounds is misidentifying non-prompt particles. Prompt particles originated in the primary vertex of the hard collision, which coincides with the interaction point, whereas non-prompt particles are created from the decay of longer lived particles and originate in a secondary vertex. Some particles that commonly cause displaced vertices include τ-leptons, b and c-hadrons, due to their longer lifetimes. For example, non-prompt electrons come from processes including photon conversion, semi-leptonic heavy-flavour decays and decays of charged pions or K-mesons, and are considered a background for prompt electrons.

Irreducible backgrounds, unrelated to detector or reconstruction performance, are removed in analysis by modelling their contributions and subtracting them from the data. This method applies generally to performance inefficiencies that can be modelled probabilistically. If the recorded rate of a process is spuriously high, the additional contribution may be modelled and subtracted, extenuating the error.

Generally, the event reconstruction process is complicated by several performance inefficiencies, which are worsened by memory limitations. Higher luminosities strain the load on CPUs, causing Grid computing to turn towards parallel strategies. However, parallelising reconstruction algorithms is not always possible when serial segments must be performed sequentially, including tracking. Efficient strategies remain a primary consideration for managing memory across the WLCG without sacrificing reconstruction performance.

Observables

The following series of collider observables includes a range of the objects that can be identified in reconstructing LHC detector measurements.

Electrons have signatures comprising reconstructed tracks associated with localised clusters of energy in the EM calorimeter. When passing through the EM calorimeter, an electron produces a shower with a distinct shape according to a series of variables, which include the shower width, energy leakage, fraction of energy deposits in each calorimeter layer and quality of the track fit. These variables are used in *loose* and *medium* levels of electron identification.

Clean reconstruction of electrons requires good methods of distinguishing them from photons that produce an electron-positron pair, and from misidentified pions. Other sources of fakes include jets misidentified as electrons, and non-prompt electrons identified as electrons. Importantly, real electrons lose only a small amount of energy when accelerated by the electric field of another charged particle, called bremsstrahlung radiation. Consequently, the ratio between the energy they deposit in the EM calorimeter and their track momentum, called the E/p ratio, should be close to unity [51]. This *tight* criterion for electron identification distinguishes them from hadrons, which deposit a large fraction of energy in the hadronic calorimeter.

Muons have similar characteristics to electrons, but lose a comparatively small fraction of energy in collisions due to their larger mass. They are reconstructed from their interaction in the muon chamber. Fake muons arise from energetic pions which have "punched through" the hadronic calorimeter to the muon chambers.

Jets are formed when partons hadronise, where gluons split into quark-anti-quark pairs and quarks emit gluons, forming colour-neutral hadrons detected in the hadronic calorimeter. The jets track the energy flow of the hadrons, and are not fundamental objects but constructions containing the QCD interactions of pair-produced collimated hadrons. Consequently, there is no unique definition of a jet, and several mappings exist from a set of hadrons with given four-momenta to a set of recombined jets with a different series of four-momenta. The rich jet substructure can become hidden once a jet is defined, but improvements to our theoretical description of jet substructure from QCD are prompting more analyses to choose techniques that exploit its complexity by designing targeted observables.

Many physics processes are distinguished by B-tagging, which is the process of identifying jets containing B mesons (or originating from a b-quark). B-hadrons have long lifetimes, travelling further than others before decaying, and bottom quarks are heavier than all others except the top quark, so they produce more charged particles during decays. From these properties, reconstructing the secondary vertex for the B-hadron decay allows computers to identify B-jets. Similarly, c-tagging algorithms identify jets originating from charm quarks through a more difficult process, which is also useful in analysis.

Missing transverse energy is the magnitude of the vector quantity missing transverse momentum, $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$. Initially, LHC beams have zero transverse momentum, so by conservation of momentum the final state particles after collision must be balanced transverse to the beam axis. The transverse energy is $E_{\mathrm{T}} = \sqrt{m^2 + p_{\mathrm{T}}^2}$, and is important for identifying neutrinos. Neutrinos are the only SM particles which do not have either electric charge or colour-charged constituents, and therefore do not interact with any component of the detector, so they are attributed any imbalance in transverse energy. The phenomenology of R-parity conserving supersymmetry models typically presents two massive neutral candidates that contribute a large and measurable amount of missing transeverse momentum.

3.3 Techniques for data analysis

The observables identify a combination of kinematic properties unique to the pathway for event production. Using this data, there is a need for creative new analysis methods for extracting small signals from the dominant SM backgrounds. An essential step in developing new techniques is to evaluate the current statistical methods. Some standard approaches are considered here with the purpose of motivating future improvements to our current sensitivity.

Any improvement to current methods must treat statistics carefully to avoid false discoveries. Rigorous statistical techniques are important in LHC physics, which is clear through pertinent examples. In a model-independent resonance search (“bump hunt”), the analysis searches for an excess in data without an expected outcome for comparison, so obtaining a large enough event count is a particular consideration. Such a method is possible only if a (new) particle decays to visible SM decay products. The false discovery of di-photon production in 2015 through a bump hunt was caused by a small overall excess that was later attributed to random statistical fluctuations after more data were analysed [58]. Generic searches for supersymmetric particles may detect a wide range of new phenomena, and may be classified as anomaly searches, a growing area of LHC physics seeking any kind of significant deviations from expectation [59]. A statistically significant measurement must be recorded and accommodated within a plausible theoretical model for the new particles to be “discovered”. The most applicable search strategies include those that search broadly and approach kinematic methods resourcefully to improve sensitivity to many physics scenarios.

3.3.1 Kinematic analysis variables

Selection cuts are placed on kinematic variables to retain high sensitivity to the signal, searching in data regions where backgrounds are minimised. These region definitions are tightened by selecting on a combination of the basic four-vector components of the final state objects and several functions constructed from these four-vectors. Preselection criteria are designed to maximise signal purity prior to analysis, which is most easily achieved by targeting the dominant backgrounds.

Ideally, the selection only passes events that match the signal final state so that only irreducible backgrounds remain, but clearly some reducible background events pass the cuts due to object misidentification and other reconstruction inefficiencies. These are removed if possible using analysis methods described in later sections. Some variables well-motivated for supersymmetry searches were introduced in Section 2.6, and the current section will continue this discussion of convenient selections for achieving clean supersymmetry signals, considering examples of supersymmetry searches for electroweakinos and stop quarks motivated by the later analyses in Chapters 5-7.

Electroweak supersymmetry analyses

The number and type of dominant backgrounds strongly depend on whether the chosen signature contains zero, one, two or three light leptons (electrons or muons). In a two lepton electroweakino search, there are typically three sources of SM backgrounds that mimic the supersymmetric final state: diboson production, top-antitop production, and the production of a Z-boson plus jets. Alternatively, in a three lepton electroweakino search, the preselection criteria may define a signal region (SR) that largely excludes the latter two backgrounds so that only diboson production dominates.

Several other preselection criteria define the SR, including cuts on the number of jets. There may be no b-tagged jets allowed, and minimal (for example, zero or one) non-b-tagged jets. Any jet or lepton belonging to a passing event may obey transverse momentum criteria that typically demand greater than around $p_{\mathrm{T}} \sim 25$ GeV, as well as criteria requiring pseudorapidity less than some value, for example $|\eta| < 2.5$. Geometric bounds are often chosen to eliminate events that involve particle trajectories passing through detector regions that are notoriously difficult to reconstruct.

Searches usually require large missing transverse energy, if the supersymmetric extension of interest predicts a stable LSP escaping with non-zero momentum. Once the SR is defined, the variables that show the most disparity between signal and backgrounds are analysed to constrain the signal model. Methods of generating these results are discussed later in this and the following chapters. Some useful variables in electroweakino analyses include the following list. These are not unique to event topologies found in electroweakino searches, nor is this an exhaustive list; they are simply well-motivated for this search and therefore used later in the electroweakino analysis in Chapter 5.

- The **transverse mass** in a two-particle system, defined as:

  $$m_{\mathrm{T}} = \sqrt{(E_{\mathrm{T}}^1 + E_{\mathrm{T}}^2)^2 - (p_{\mathrm{T}}^1 + p_{\mathrm{T}}^2)^2}$$

  where p^1, p^2 and E_{T}^1, E_{T}^2 denote the transverse momenta and transverse energies of two daughter particles in a decay. Full mass reconstruction may not be possible in events with missing or mismeasured particles (or noise contamination from pileup), so when one particle cannot be detected directly but is indicated by missing transverse energy, this definition of transverse mass applies. The given definition was created by the Underground Area 1 (UA1) experiment [60] to eliminate dependence on unknown quantities such as the total energy.

  At the UA1 experiment, the W boson was detected through its decay into a charged lepton and a neutrino, where the neutrino escapes with missing momentum. The W mass therefore could not be directly obtained from the lepton and neutrino momenta in any given event, so the transverse mass variable was used to place a lower bound on its mass, given by the property $m_{\mathrm{T}}^2 \leq m_{\mathrm{W}}^2$. Equality is possible if the lepton and neutrino are produced with identical rapidity. Using m_{T}, the W mass was extracted at UA1 (and UA2) in a fairly model-independent way by accumulating enough events close to the upper limit of m_{W}.

  The transverse mass is derived from the definition of the invariant mass of a relativistic particle, $M^2 = E_{\mathrm{T}}^2 - (\vec{P}_{\mathrm{T}})^2$, and since it is defined only using quantities which are invariant under a Lorentz boost in the z-direction, it is also longitudinally boost invariant. If p_1 and p_2 are purely transverse, which is true if $\eta = 0$, then $m_{\mathrm{T}} = m$. Alternatively, if p_1 and p_2 are instead purely longitudinal then $m_{\mathrm{T}} = 0$. Otherwise, transverse mass lies in the range $0 < m_{\mathrm{T}} < m$.

- The **minimum transverse mass** of a lepton,

  $$m_{\mathrm{T}}^{l,\mathrm{min}} = min(m_{\mathrm{T}}(l_1, E_{\mathrm{T}}^{\mathrm{miss}}), m_{\mathrm{T}}(l_2, E_{\mathrm{T}}^{\mathrm{miss}}), m_{\mathrm{T}}(l_3, E_{\mathrm{T}}^{\mathrm{miss}}))$$

  where the transverse mass of a lepton is first calculated in events with missing transverse energy for each lepton in the event, before taking the minimum.

- The reconstructed **transverse momentum** of the Z boson $p_{\mathrm{T}}(Z)$.

- The **azimuthal angle** $\Delta\Phi$ between several different pairs of particles.
  - $\Delta\Phi(l_Z^+, l_Z^-)$ between the two leptons associated with the Z boson. The standard method requires the two chosen leptons to be a same-flavour, opposite-sign pair such that their dilepton invariant mass is closest to the Z-boson mass. The third lepton remaining in the event is assigned to the W-boson.
  - $\Delta\Phi(Z, l_W)$ between the reconstructed Z boson and the lepton associated with the W boson.

Third-generation squark search

The parameter space of stop quark searches is more constrained than the corresponding space for electroweakino searches, mostly because the signal cross-sections for electroweakino production are smaller. Stop quark masses and decay modes that are not excluded are those that closely resemble top quark pair processes, including scenarios where the mass difference between the stop quark and the neutralino is smaller than the top quark and consequently each stop quark decays via a 3-body process into a b quark, a W boson and a neutralino. ATLAS searches for stop quarks have included significant progress in 0-lepton, 1-lepton and 2-lepton final states.

Some variables described for electroweak supersymmetry are again useful here to target common features. The missing transverse momentum is a clear example of a variable that is broadly applicable across a range of event topologies, motivated here again by the stable LSP with R-parity conservation. Additional variables are defined as follows, and are again motivated by standard stop searches and are applied later in the analysis contained in Chapter 6.

- The **minimum value of the transverse mass** formed by the two b-tagged jets and missing transverse energy in the event, $m_{\mathrm{T}}^{b,min}$, using the following definition of m_{T}^{b}.

  $$m_{\mathrm{T}}^{b} = \sqrt{2p_{\mathrm{T}}^{b}E_{\mathrm{T}}^{\mathrm{miss}}[1-\cos(\Delta\phi)]}$$

  The transverse momentum of each b-jet is p_{T}^{b}, and $\Delta\phi$ is the difference in ϕ between each b-jet and the missing transverse momentum.

- The **asymmetric m_{T2} value**, where m_{T2} is a generalisation of the transverse mass for use with signatures where two or more particles are not directly undetected [61, 62]. It is useful for measuring the masses of pair-produced particles where each decays to one particle that is observable directly and another that is invisible and carries away missing transverse momentum. The m_{T2} may become a function of the mass of a missing heavy particle, if such a mass is known, or may be calculated from Equation 3.14 with two sets of one or more visible particles denoted a and b. The maximum of their transverse masses is minimised over all possible combinations of missing transverse momenta $\vec{p}_{\mathrm{Ta}}$ and $\vec{p}_{\mathrm{Tb}}$, with the requirement $\vec{p}_{\mathrm{Ta}} + \vec{p}_{\mathrm{Tb}} = \vec{E}_{\mathrm{T}}^{\mathrm{miss}}$.

  $$m_{T2} \equiv \min_{\vec{p}_{\mathrm{Ta}}+\vec{p}_{\mathrm{Tb}}=\vec{E}_{\mathrm{T}}^{\mathrm{miss}}} \{\max(m_{\mathrm{Ta}}, m_{\mathrm{Tb}})\} \tag{3.14}$$

  The kinematic endpoint of the m_{T2} distribution provides a convenient evaluation of the mass of the parent particles as a function of the daughter particles, which is a useful identifier of events in analysis [63]. In a supersymmetric process, it carries model-independent information about the mass difference between the primary and secondary supersymmetric particles.

  The term "asymmetric" refers to event topologies involving two supersymmetric decay chains terminating in different daughter particles, which is a more general approach than the commonly assumed decay pathway corresponding to two identical missing particles. For example, it is often useful for suppressing dileptonic $t\bar{t}$ events where one lepton goes undetected. For these top production events, the asymmetric m_{T2} (am_{T2}) has a kinematic endpoint at the top quark mass, but extends to higher values when there are more sources of missing energy in the process.

- The **minimum of the invariant mass** formed by the lepton and each of the two b-jets in an event, $m_{\mathrm{bl}}^{\mathrm{min}}$. The value $m_{\mathrm{bl}}^{\mathrm{min}}$ reaches a maximum in top pair production events, but is not bounded by the same value in the corresponding supersymmetric stop quark production events.

- The **scalar sum of transverse momenta** H_T, calculated by adding the moduli of the transverse momenta for the physics objects in the event.

  $$H_\mathrm{T} = \sum_j |p_{\mathrm{T},j}|$$

  The sum over the physics objects can in general include every object in the event, where j is the object index. More specific definitions may sum over, for example, just one lepton and two b-jets, if these are part of the pre-selection requirements. Sometimes, in analyses without leptons, the sum is restricted to jet objects only.

Another common variable is the effective mass (m_eff), which in its simplest form is the sum of the transverse momenta (p_T) of every jet and lepton in the event, plus the $E_\mathrm{T}^\mathrm{miss}$. Other variables are calculated for semi-invisible particle searches using intricate methods that account for unknown boosts along the beam direction and complicated cascade decay chains. We can apply additional assumptions about decay chains and intermediate masses to apply constraints on parent masses.

Recursive Jigsaw Reconstruction

The *Recursive Jigsaw Reconstruction* (RJR) method for generating uncorrelated analysis variables is capable of reconstructing particles in the presence of combinatoric or kinematic ambiguities [64]. Kinematic unknowns arise from unmeasured particles, and combinatoric unknowns come from indistinguishable particles. RJR infers both types of missing information by approximating the rest frames of every intermediate particle state in an event and evaluating the energy and momenta of interesting objects from within these rest frames. The reconstructed decays are entirely resolved using algorithms known as Jigsaw Rules, which can be applied to any event topology. These algorithms generate a set of recursive jigsaw (RJ) variables, a powerful approximation of the average kinematic behaviour which the reconstructed missing objects would possess in the rest frames. The RJR method is applied in the later analysis in Chapter 7.

The approach assumes a particular decay topology for the events from several possibilities by choosing the option that best captures the kinematic features of the signal topology under analysis. The chosen topology then sets the identities of the intermediate particle states, and the method chooses a rest frame for all particles in the decay tree by iteratively minimising their particle masses. The momenta that were observed in an event are then boosted to one of these frames, and the boosted four-vectors are used to construct the RJ variables, which either have units of mass or are dimensionless quantities such as angles between objects.

Figure 3.8 is a simplified decay tree showing a generic event where the decay state produces a single invisible particle. Assuming the visible state is identified and reconstructed by the detector with a four-momentum measured in the lab frame, RJR allows us to infer the mass and momentum of the invisible particle along the beam axis in the lab frame.

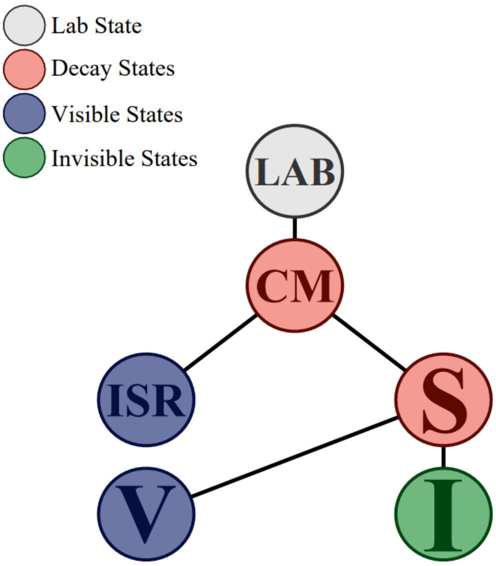


Fig. 3.8 A simplified decay tree, where invisible particles (green) are unknowns [65].

The same is true in the example of a decay tree for a single W production in Figure 3.9, where assumptions must be made about the kinematics of the neutrino. The RJR approach resolves the momentum and decay kinematics of the W boson without any prior assumptions about the kinematics or masses of the event, instead considering the underconstrained four-momentum of the neutrino as unknown components in the boost velocity between the lab frame and the rest frame of the W boson.

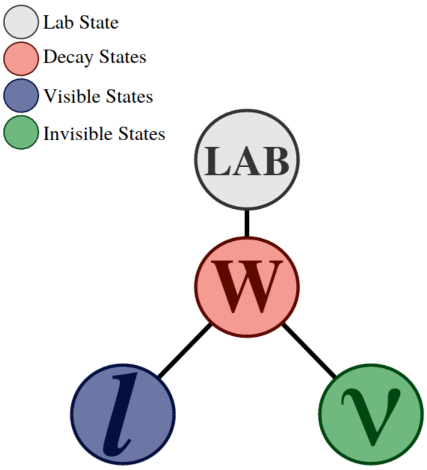


Fig. 3.9 A decay tree for W production in the lab frame and subsequent decay into a lepton and neutrino [64].

In general, the kinematics of a decay tree are fully specified by the velocities relating adjacent reference frames plus the measured four vectors of visible particles, without the need for information about the mass or kinematics of any invisible particles. The specific Jigsaw Rules required to resolve all the unknowns in an event and the order in which they should be applied is automated by a code package called RestFrames [66]. After analysis, the RestFrames object assigned to any particle in the decay tree can be queried about the particle's mass, decay angles and momentum reconstructed in any reference frame.

The observables calculated in a particular frame are largely independent of those calculated in all other frames of reference, so the variables calculated by RJR are a strong basis for analysis. The basis of observables is created in RestFrames when partitioning available information for each reconstructed event into approximately uncorrelated variables [64]. In tests of these observables on supersymmetric examples, they have shown evidence of being able to distinguish between different decays of similar supersymmetric particles in chosen models of supersymmetry, and are also sensitive to expected differences between these signals and SM backgrounds.

Application to analysis

An analysis group identifies measured and derived (calculated) kinematic variables to target interesting signatures. After applying preselection criteria, they plot the distributions of relevant variables to isolate features that are indicative of signal events. The kinematic variables described above provide a series of both measured and derived quantities designed to reveal independent features of the data. Ideally, different variables are sensitive to different kinematic properties, and analyses can set stronger exclusion limits when they use a larger set of minimally correlated variables whose combined sensitivity may isolate rarer signatures.

There is further scope to develop creative new variables that can add discriminatory power when applied alongside those existing. We require high sensitivity to small signals such as weakly-produced supersymmetry to avoid the need for heavy optimisation of BSM scenarios [67]. For the best use of Run-2 data and beyond, we must add power additional to the current variables.

3.3.2 Modelling physics processes for analysis

Signal and background processes are computer-simulated to predict their detector phenomenology, with the aim of bridging the gap between theoretical first-principles calculations and complex detector signatures in experimental data. This is achieved by simulating particle physics events as they occur in nature. Experiments depend on accurate models, and their accuracy is the subject of constant improvement. Simulations model events based on probability calculations of likely decay topologies and kinematics, which are limited by theoretical uncertainties. The generated events are compared with experimental data both to improve the simulations, in the case of SM processes, and to measure excesses, in the search for BSM signals. Simulations may be used in place of real data as part of a preliminary proof-of-principle analysis when developing a new technique, such as the methods applied later in Chapters 5-7.

Theoretical and technical improvements are ongoing to increase the order of perturbative calculations and explore new structures for parton showers, amongst other benefits. Simulations are increasing systematic inclusion of terms corresponding to sub-leading corrections and fixed-order calculations at increasing precision. Calculations at next-to-next-to leading order (NNLO) or better in QCD, mixed QCD-electroweak and electroweak theory are under heavy development, reducing theoretical uncertainties.

The detector response is modelled on top of the underlying decay process to account for the effect of finite resolution and reconstruction inefficiencies on particle signatures. Event generators must simulate the same performance capabilities as detectors, including their inability to identify intermediate states in a decay tree or attach precise mass and kinematics to all particles. The ATLAS detector has been fully simulated with high accuracy in programs such as Geant4 [68], which incorporate its predicted response to supersymmetric particles. Simulated event displays reveal predicted detector effects for particles not yet observed as well as consolidating our knowledge of the processes that produced real particle collision data. Accurate simulations are of particular importance for rare processes and their dominant backgrounds, including sparticle production, as they require more precise predictions to overcome the challenges arising from a low event rate.

Simulated events contain constituents which represent real collider data, including a list of stable particles (hadrons, photons, leptons) with their associated four-momenta. An example of an ATLAS event display for Higgs production is presented in Figure 3.10, where elements of the final state are identified in the legend: two b-jets are shown in blue, and an escaping muon appears in red. The display is a simulated version of a real event recorded by ATLAS in 2017 that contributed to the first measurement of the Higgs through a new channel, the decay to two b quarks, which was a challenging signature to identify despite comprising almost 60% of Higgs decays [69]. Although the neutrino would not be visible in the real detector, its proposed track is marked in the simulation. Yellow boxes represent clusters of energy deposited in the calorimeter, and Inner Detector tracks appear in orange. Reconstruction effects are precisely tuned using these detector response simulations.

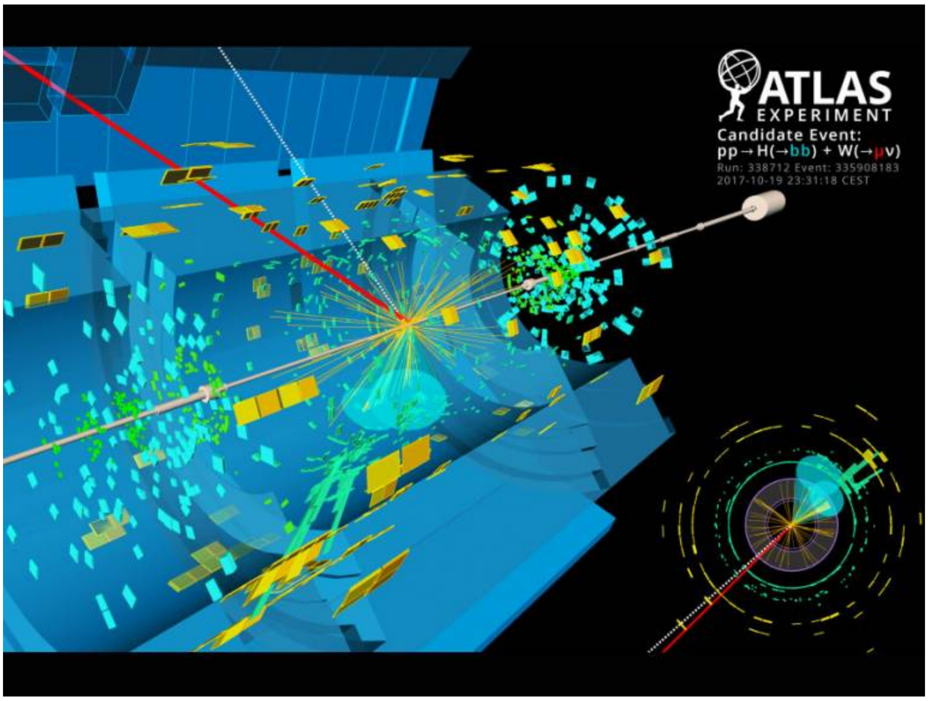


Fig. 3.10 Simulation of a candidate ATLAS event contributing to the Higgs boson measurement via a pair of b quarks published in 2018 [69]. The display identifies a range of reconstructed particles [70].

Monte Carlo simulations

The Monte Carlo (MC) method of event simulation is a numerical integration technique useful for representing the quantum mechanical variability that produces wildly different multiparticle final states under fixed initial conditions. It is capable of first integrating over the probability density of the physical process and then simulating the detector and subsequent trigger and reconstruction algorithms. Several separate stages of simulation contribute to this process: modelling hard processes, resonance decays, parton showers, the underlying event, hadronisation and others. Different MC event generator algorithms are maximally effective at each of these tasks.

After an algorithm integrates over the relevant probability density, it passes the hard process through to a parton shower algorithm before hadronising the results to include the full set of measurable objects, including jets, in reconstruction. Interactions with the detector are digitised, translating them into realistic signals. General purpose MC event generators are capable of performing many of these tasks, and specialised programs perform parts of the process with greater accuracy. One common generator is Pythia [71], widely used for its general purpose capabilities. To add a simulation of the ATLAS detector, Pythia can be interfaced with a detector simulation such as Delphes [72]. This interface is sufficient for the hard process and detector simulations required for the analysis in Chapters 5 and 6.

Event generators that use MC techniques such as Pythia are fast and reasonably accurate: they converge fast in many dimensions, and their accuracy improves with every additional point. They benefit from a straightforward error esimate, which scales with the number of points requested by the user. The error always converges proportional to $1/\sqrt{N}$, regardless of the number of dimensions, which is a strong benefit when performing the high-dimensional phase space integrals required to generate the momenta of many particles per LHC event [73]. A method called *importance sampling* further improves the rate of convergence by performing a Jacobian transform, producing a new integration variable in which the integral is flat, and thereby reducing the error.

Modelling with MC methods is undergoing ongoing improvements, and only a small number of non-standard backgrounds remain unavailable through MC simulations. Significant recent progress has been made in calculating processes at NLO with higher jet multiplicities, which are useful for supersymmetry and other BSM searches. Calculations at NLO involve a combination of explicit infrared divergences from the loop integral with divergences arising from integration over the phase space. To successfully integrate the differential cross-section at NLO with this combination, an MC generator must introduce a subtraction term with the same infrared pole structure as the real matrix element. It then integrates the result over the phase space that was unresolved, eliminating the nuisance divergences. MC event generators are able to accept arbitrarily complex integration regions by such methods, and require only a few points to obtain first estimates.

An MC generator is always subject to the under- or over-estimation of some physics processes in certain regions. Comparing simulated events at reconstruction-level with those at generator-level allows us to quantify the efficiencies and resolutions, and to consider theory systematics. One area of improvement is the *jet reconstruction performance* of event generators. This is described by measuring the jet energy scale (JES), a value given by the ratio and mean of the comparison between the reconstructed jets and the generator-level jets. The JES indicates how well on average the underlying events are corrected in the reconstruction. The detector has finite jet energy resolution, which smears out detailed jet structures (and is the width of the jet energy response), and MC is known to underestimate this resolution. Unfolding is the process of restoring the features that were smeared out by detector resolutions, so we perform unfolding on the MC to give an exact theoretical description of jet resolution and give it the same artifacts as the data.

If an analysis region is undersampled by the simulation, the user can benefit from increasing MC statistics in the chosen region by several useful methods. One is smearing the MC events, and another is re-weighting them after biasing the phase space. In the first option, events that pass the selection criteria for the region are copied perhaps ~ 100 times, then objects are smeared in a window around their expected response. For example, in a hadronic process, the leading jets may be smeared around their expected jet response. The window is small enough not to bias event topology, so the kinematic distributions of smeared and pre-smeared events match. Consequently, the existing events are made to represent a larger sample without requiring the user to generate more events.

The option for increased sampling where the user biases the phase space is another valuable technique, beneficial in cases where computing power or memory restrict the maximum number of events. Throughout an MC simulation, the user has the option of biasing the likelihood of events at parton-level (on the hard process) or at detector-level. This artificially increases the likelihood of the chosen kinematics, oversampling in this region compared with others that are less important to the analysis. By later assigning low weights to these contributions, the user easily decreases statistical error in their analysis region when computational limitations are preventing them from generating enough events to do so in all regions simultaneously.

Event weighting

Weights are assigned to simulated LHC samples to maintain the total number of events close to the true number expected in data. One purpose mentioned above is to re-weight oversampled phase space regions to increase analysis statistics. Alternatively, weights may be added to all events of a particular process in a sample to obtain an effective number of events that correctly reflects the LHC luminosity and the cross-section relative to other processes.

Assigning weights after oversampling follows a similar process to prescaling at trigger level in experiment to reduce the fraction of common events that are stored for analysis. However, in this case the resulting bias towards the region of interest is removed by applying a different normalisation factor to the divided subsets of phase space. The normalisation factors weight the oversampled regions down by their cross-section to represent the appropriate number of events that would be expected to fall into that region in a true experiment. Consequently, statistical reliability is distributed more evenly across the space without skewing results. The outcome is equivalent to replacing the basic cross-sections encoded in the MC event generator.

Weights are calculated in a simple case for an LHC integrated luminosity L as follows. The cross-section for a process x with N_x events contained in a larger total sample is:

$$\sigma_x = \frac{N_x}{L} \tag{3.15}$$

where N_x is the number of correctly weighted events. If the number of events of process x before weighting is N_{x0}, then the weight W on each event is equal to N_x/N_{x0}. Expressed in terms of the desired luminosity and cross-section for process x, the weight becomes:

$$W = \frac{\sigma_x L}{N_{x0}} \tag{3.16}$$

Using only the predicted cross-section and integrated luminosity, an MC event sample can be weighted according to this simple procedure. A more involved re-weighting process applies to more complicated samples, including those with oversampled regions of phase space. Occasionally, MC simulations are required to assign negative weights to events whose contributions must be removed, for example when subtraction terms have been introduced in NLO calculations. These more complex cases are treated later in Chapters 4 and 7, respectively.

3.3.3 Standard statistical analysis techniques

Statistical power is optimised so that analyses use as much of the available data as possible. Several separate SRs may contain a high signal-to-background ratio, and are combined to improve the overall signal sensitivity. ATLAS analyses often exploit multiple SRs, propagating errors appropriately between regions. In any SR, the predicted signal is extracted from its dominant and subdominant backgrounds by estimating and removing those backgrounds. Then, statistical methods are used to determine the significance of a discovery or exclusion.

Statistical versus systematic uncertainties

Statistical and systematic types of uncertainties are both applied to background models generated through simulation. The statistical uncertainty is derived from the finite number of MC events in a sample, and can be improved by increasing the sampling in regions of interest. Systematic uncertainties include detector and reconstruction effects, along with theoretical uncertainties derived from imperfect understanding of the SM.

To calculate systematic uncertainties, we vary systematic quantities and observe the model's response. Sources of experimental systematic uncertainty include the jet, lepton and photon energy scales and resolutions, as well as b-tagging efficiency and jet energy scale [74, 75]. Systematics also include trigger efficiencies and the imperfect measurement of the overall luminosity. Both experimental and theoretical systematic uncertainties are defined using this method of varying parameters within their expected range and quantifying the response. For theoretical uncertainties, the process involves varying cross-sections, comparing different MC generator yields, varying the renormalisation and factorisation scales for QCD processes, and varying the parton distribution functions [76]. The relevant systematic uncertainties are propagated through to the histograms of variables used in the analysis.

Data-driven background estimation

If backgrounds are estimated using data-driven methods, the dominant backgrounds in the SRs are modelled with MC simulations and normalised by signal-depleted regions called control regions (CRs). The backgrounds in the CRs are fitted to the data in order to calculate the normalisation factors. Separate regions called validation regions (VRs) are also chosen with the characteristics that they possess similar kinematics to the SRs but lower expected signal-to-background ratios, and they are used to validate the data-driven background estimation. The normalisation factors from the CRs are propagated to the SRs and VRs, and the fits are compared when the results are ‘unblinded’. After unblinding, we check for deviations in the data from the backgrounds, which may take the form of an excess that stands out against a smooth fitted background function.

A background yield estimation for an SR using MC in a dedicated high-statistics CR requires the analysis to calculate a transfer factor composed of the ratio between the number of Monte Carlo events in the SR and that in the CR. The result is a central estimate for the background yield in the SR, as shown in Equation 3.17.

$$N^{\mathrm{SR}}_{\mathrm{Background}} = \frac{N^{\mathrm{SR}}_{\mathrm{MC}}}{N^{\mathrm{CR}}_{\mathrm{MC}}} N^{\mathrm{CR}}_{\mathrm{Data}} \tag{3.17}$$

An alternative approach to background estimation

The MC background prediction can only be scaled by a CR normalisation factor if a valid CR exists which contains the same backgrounds as the SR but a depleted signal, which is not always true. If the backgrounds in the SR cannot be evaluated by these data-driven methods, pure Monte Carlo is a practical and dependable alternative. However, a pure MC approach requires some additional error propagation, because MC is affected by theoretical systematic uncertainties in addition to the experimental sources. The sensitivity of a result to each source of uncertainty depends on the analysis type and the SRs, so some are better designed to accommodate pure MC background estimations.

Likelihood fit

The next stage of the analysis aims to extract the significance of the chosen signal model given the background estimation by applying a simultaneous likelihood fit of the models to the observed distribution. Ideally, the likelihood function contains only one free parameter which corresponds to the value the analysis seeks to measure: the parameter of interest (POI). In this case, the value of the POI in the data may be determined by evaluating the likelihood function in a background-only region, and comparing with its value in the SR. Realistically, however, several parameters are added to the function by the systematic uncertainties from detector simulation, theory and statistics, which are called ‘nuisance parameters’ (NPs), and must be constrained. By definition, systematic uncertainties are unconstrained by the data and cannot be obtained from the measurement, so they are instead constrained in the likelihood by multiplying by estimated constraint factors. After this step, all uncertainties can be treated as statistical.

Constraining the background NPs and incorporating them into the likelihood function is the same process as including the backgrounds from the CRs in the fit. This process of using data-driven best-fit values to incorporate a description of systematic uncertainties into the statistical likelihood function is called ‘profiling’ the systematics, and can significantly reduce the total uncertainty in a measurement. The NPs are typically modelled using either a Gaussian or Poisson distribution, which allows the relevant Gaussian or Poisson statistical uncertainties to take the place of the unknown systematic uncertainties. The approach can be used in a resonance search, where the shape of a continuum background is determined in the likelihood from its shape in CRs.

Hypothesis testing

The likelihood function compares two hypotheses which suggest alternative physics scenarios that may be present in the data. For each signal hypothesis, we can calculate a p-value quantifying the compatibility between the data and the background-only hypothesis. If the likelihood function has one free POI μ, we use the function to compare the likelihood of the background-only null hypothesis H_0 with a signal hypothesis H_1 by evaluating it at the relevant values of the POI: $\mu = \mu_0$ under background-only assumptions, and $\mu = \mu_1$ from the data. The ratio between these two values gives the p-value, calculated by evaluating the likelihood under the two conditions, and is the most powerful test statistic for discriminating between two hypotheses according to the *Neyman-Pearson lemma* [77]. The test statistic is an extension from basic number counting analyses, in which the likelihood of observing n events can be found from the Poisson distribution.

The Asimov dataset is a simple method to obtain the median experimental sensitivity of a measurement, considering the fluctuations about the expectation. For a given set of model parameters (including POIs and NPs), the Asimov dataset is defined as the set generated when all the parameters are set to their best-fit values, so when the dataset is binned the event count in each bin is set to the expected yield. The approach then provides the median significance required to reject the background-only hypothesis under the assumption of a certain signal model.

3.3.4 Supersymmetry searches with the ATLAS experiment

The signal models predicted for supersymmetry searches depend on many precise factors, including the details of the supersymmetry breaking mechanism. Measuring sparticle masses and couplings on the weak scale is sensitive to this high scale mechanism. The experimental picture of supersymmetry depends on theoretical considerations which broadly include: the particle content and mass hierarchy, whether R-parity is conserved or violated, the types of particles targeted and their mixing angles and coupling constants to the Standard Model particles. Many analyses choose to target a highly specific set of particles and apply constraints to specify a particular realisation of the theory. Constraints include imposing a certain mass hierarchy and equalities that reduce the number of possible parameters in the MSSM.

In many supersymmetric scenarios, heavy sparticles decay to several massive Standard Model top quarks or bosons, which subsequently decay into final-state leptons and jets with significant branching ratios. The decay products of heavy particles are generally boosted, and subsequent decays are collimated, which can be captured in a large R jet. This information is used to design searches for the decays of very heavy particles, including top quarks. In general, due to features such as high missing momentum and large object multiplicity, the most powerful probes of natural supersymmetry often come from jets plus missing energy and same-sign dileptons. The typical irreducible backgrounds with final states common to supersymmetry decays of interest include QCD with mis-measured jets, W+jets and top quarks, which are modelled precisely to quantify their contribution to the data according to the above statistical methods.

The null results that have been produced by supersymmetry searches with ATLAS so far are narrowing the broad parameter space and range of complex particle spectra, but progress remains slow, and there are even some cases where LHC results provide no new constraints. This situation arose in a recent global fit of the electroweakino sector in the MSSM [78], which concluded no significant exclusion on any range of electroweakino masses in general. Studies like these are hampered by the very low rate of weakly-produced supersymmetric signal compared to the dominant SM backgrounds.

Current limits on supersymmetry searches with the ATLAS Experiment

No supersymmetry searches to date have found a significant deviation in the data from the background estimate. The analysis group sets limits on models using the fit and hypothesis testing methods described in Section 3.3.3. Until July 2020, stop quark searches had eliminated the regions defined in Figure 3.11. Analyses are employing steadily more sensitive techniques to explore previously-inaccessible signal models, aided by larger datasets. This combination inspires further hope for a discovery after Run 3 begins in 2022 with the intention of reaching a higher luminosity. If the complex decay patterns and obscure detector signatures of supersymmetry are hiding the evidence from view, we must develop progressively more advanced methods of analysis in preparation for a discovery in the near future, with a focus on decreasing their signal dependence.

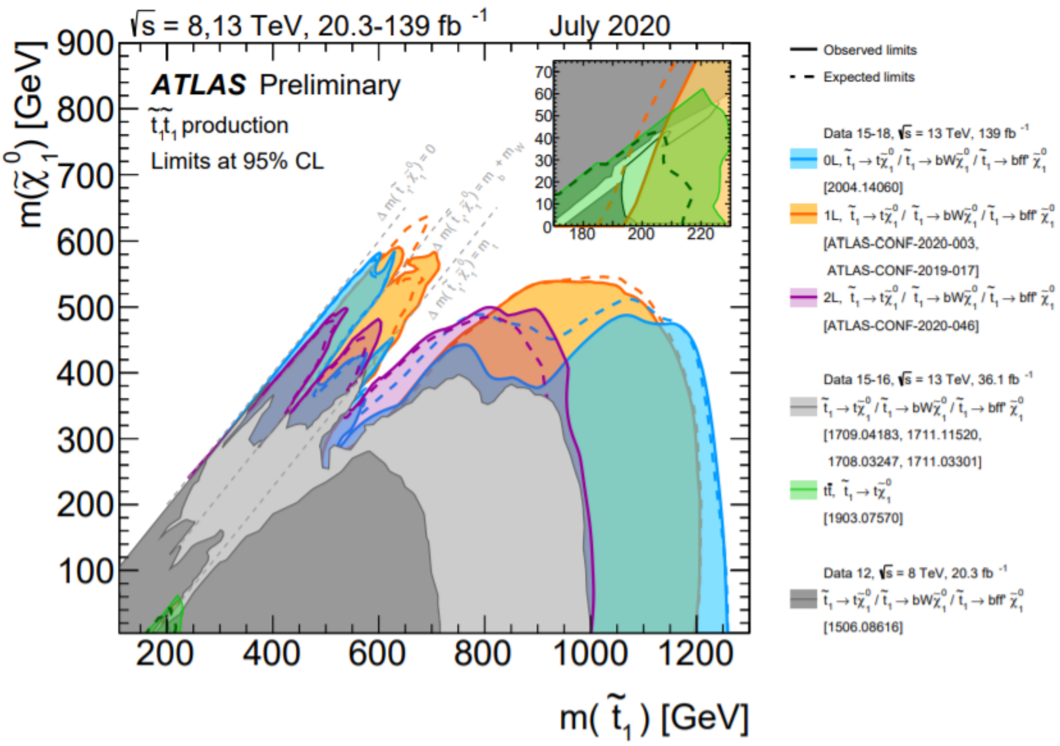


Fig. 3.11 Exclusion limits on the stop particle mass from the ATLAS Experiment, up-to-date in July 2020. The mass of the lightest neutralino is plotted as a function of the stop mass for measurements of stop production proceeding via several possible pathways [79].

3.3.5 Advances in analysis strategies

Advances in ATLAS analysis strategies have recently come to rely heavily on machine learning techniques, multi-channel statistical analyses and multi-variate analyses. An analysis develops a set of minimally correlated variables that can ideally be used as part of the selection criteria revealing an excess of signal, and these variables can equally be used in a machine-learning classification [80–82]. The more sophisticated cut analyses use multivariate algorithms, such as neural networks (NNs), to define more complicated SRs. They have the power to improve statistics and sometimes open up phase space that was previously unexplored.

The aim of advances in analysis, including machine learning, is to find new ways of constructing sensitive variables and to improve existing strategies. One method of increasing the sensitivity of standard variables is principal component analysis (PCA), which finds the axis of least variance between the SM-only and the SM+BSM model and then projects along this axis to produce the most sensitive variable one can construct using only those input features [83]. Consequently, differences between an SM-only and an SM-plus-supersymmetry model, for example, will appear most prominently in measurements of this new lower-dimensional variable. A machine learning technique that can further improve the basic PCA strategy is deep neural networks (DNNs). DNNs take the original variables and construct a new discriminant with even more input features, labelling the output as 0 for background and 1 for signal. The networks can be tested with different sets of inputs to find those best suited for a particular analysis [84].

Three deep learning techniques are the most widely used in particle physics analysis: the fully connected network (FCN), convolutional neutral network (CNN) and recurrent neutral network (RNN). Each has a unique definition for a layered network, with important differences between the weighting they place on relations between entities in close proximity with one another in the input's coordinate space, and those that are further away. FCNs relate all input components, whereas CNNs define only local relations, and RNNs define sequential relations. The data format required by FCNs, CNNs and RNNs is a space of vectors, sequences or grids. However, particle collider data commonly comprises a space of unordered elements (events). Unlike the three network types mentioned above, techniques that instead employ *graph* networks create pairwise connections between sets of elements, and are applicable to collider physics problems that involve unordered datasets with rich relations and interactions between the elements. Graph network methods have only minimally and recently been applied in a particle physics context, particularly in the form of graph neural network machine learning techniques [85].

An increasingly standard machine learning method of classifying signal and background events when those events do not have all the characteristics of either signal or background is the Boosted Decision Tree (BDT). A decision tree recursively checks if additional criteria can help to classify events correctly. Even if events initially fail a certain signal criterion, they are not immediately rejected and may later be classified as signal according to the additional criteria. A BDT requires a good understanding of the model describing the data, and the algorithm can only be trained if there are enough events available. They are specifically trained to discriminate against the dominant backgrounds, and are used alongside selection-based analyses to provide additional sensitivity.

A useful growing area of LHC analysis lies in model-agnostic searches using anomaly detection methods. The subject comprises several regimes for searching broadly for any signal-like events that are not explained by the SM background, one of the most robust of which is the variational autoencoder (VAE) [86]. An encoder represents the features of the original data in a new representation with fewer dimensions, similar to the reduced dimensionality approach of PCA. While an encoder compresses the data, the next stage of a VAE is a decoder which decompresses it, with some losses. These generative models are used for fast simulation in ATLAS [87], and also trained on SM processes so that they can detect anomalies with no assumptions on the signal [88].

Future outlook on analysis

There is clearly a strong motivation to revisit particle search techniques, and find new ways of extracting small signals from LHC data. Creative new search techniques will be useful first for searching in current Run 2 LHC data, which has already been partly scrutinised using standard analysis strategies, and then for analysing additional data produced after the planned upgrades to collider reach. Run 3 will modify the trigger system to record rare events in regions of parameter space that remain unexplored, and increase the luminosity. Following this, the high-luminosity LHC (HL-LHC) is the successor to the LHC scheduled to begin operation at the end of 2027 [89]. It is designed to increase the dataset by more than 10 times, key to analysing rare new phenomena that may already be present in the energy range currently under exploration. Progress will be made by synchronising advances in collider luminosity and energy reach with the development of creative search strategies capable of probing effectively for theoretically-motivated signals in the new data. We aim to use novel techniques to target the most likely areas for imminent discoveries.

Chapter 4

Graph network analysis

4.1 Introduction

Graph networks explore the connections between elements in a dataset, where the user has the power to choose the scale of these connections. Networks are a discrete representation of data sampled from a continuous underlying space, containing discrete *nodes* and the *edges* (connections) between them. Two nodes are connected if they are “close” in the space and disconnected otherwise, where closeness is defined by some measure of *similarity* specific to the data. Patterns in connections on a chosen scale reveal properties of the data that characterise either the individual nodes or the dataset as a whole, and these patterns are quantified by convenient calculations of *network metrics* defined in graph theory. The extensive set of metrics provides a powerful method of analysing the similarities between nodes, and comparing these measures has provided useful evidence of graph behaviours and characteristics for datasets from the social and physical sciences [90–92]. Network analysis has never before been applied to event-by-event particle collider analyses, and we show that by calculating the network properties of particle collider data, we can apply standard graph methods in novel ways to improve sensitivity to BSM physics. We formulate a series of discriminators between signal and background events which are sensitive to small signals across a range of physics scenarios.

Network analysis reveals more information about the data than standard particle physics searches. ATLAS searches typically calculate discriminating variables to compare disconnected events, prescribing a set of attributes to each event which treat them individually without quantifying properties of their situation in the dataset as a whole. The classification of signal and background may involve simple quality cuts or more elaborate algorithms. However, typical observables often reveal no difference between disparate event topologies, so there is a need to explore the structure of collider data in greater depth than we can achieve with these observables alone. By evaluating relative kinematics through graph networks instead, we characterise topologies of events on a chosen scale in the kinematic space as a whole and extend standard BSM LHC analyses to increase their sensitivity.

Our research proposes a new method of exploring LHC data by calculating effective distances between events, and uses this information to create networks which define novel event-by-event variables sensitive to event topologies. These variables are based on typical kinematics but contain additional information encoded in the similarities and differences between the events on some scale, which becomes available only through studying their network connections. Our networks comprise nodes of LHC events mapped as graphs in *collider space*, the underlying multi-dimensional space of original kinematic variables. Signals we seek to isolate in our analysis are generated from two scenarios of supersymmetry based on stop quark and electroweakino production, treated in Chapters 5 and 6 respectively.

4.1.1 Event similarity in collider space

High-energy LHC collisions produce particles which are related by their kinematic attributes and decay topologies reflecting their partonic origins. BSM signals may be rarer or more complex than SM data, generating potentially fewer connections in a network between BSM events as a measure of their similarity in the space. Events are connected if they possess similar values of the underlying variables used as a basis to construct the graph. For example, the missing energy values reconstructed for LHC events should reveal clusters of SM events at lower values than the supersymmetry events, which are expected to be few and scattered at higher values than the bulk of the distribution. By constructing a network from these events instead of treating their values of missing energy in isolation, we obtain new information derived from the *similarity* connections: the information we know about a particular event grows with its connectedness, as every connected neighbour effectively passes information about its own kinematics on to the event of interest. Network properties thus encode information not only about an event's kinematics as a discriminator between event types, but also about the properties of all the events nearby which are connected to that event.

We consider that an event is likely to share significant kinematic properties with a distinct subset of the dataset. This "similar" subset may be clustered in a localised region of kinematic space, or be located approximately the same distance from a central event. BSM signals may possess characteristic structural patterns when viewed holistically in a graph network, which are conveniently calculated using the network metrics. Distributions arise from the sampling of continuous manifolds underlying SM and supersymmetry processes, which lend differences to their signatures measured at the LHC. Using similarities and differences in networks of events, Chapters 5 and 6 show that signal events of a particular process type may be identifiable by a combination of network metric calculations. These calculations are sensitive to structurally distinct processes and may be strong discriminators of BSM scenarios.

Information about similarity is obtained from factors including the density of connections between nodes, the types of nodes that share connections, and the geometry and topology of the sampling of events. Analysing geometries requires knowledge of the metric dimensions of shapes, whereas topological study refers to the intrinsic shape properties of metric-free structures [93]. A structure may have shapes which are invariant under continuous deformation of the space, making these shapes homotopic, and belonging to the study of topology. The distributions of signal and background events in collider space can reveal both topological and geometric differences, each possessing some features which are invariant under spatial deformation and some which are metric-dependent. We make network topologies the broad focus of our study, where topological disparities between signal and background events are independent of metric dimensions and so their visibility in analysis is less dependent on the choice of parameters in network construction. Topological properties are therefore more relevant to a first proof-of-principle. Furthermore, this focus aligns with our long-term goal of developing a standard analysis technique with applications across a broad range of physics scenarios, which should not require fine-tuning of network parameters to achieve sensitivity.

4.1.2 Applications of graph networks

Networks have been created from social, trade, biological, communication and climate data [92, 94, 95] to represent social media, political influence, neural pathways and the transfer of information across the internet, amongst other real-world applications. Community structure is a common property of social networks, where groups of nodes possess dense internal connections but are weakly connected between groups. Modelling how networks are generated with different topologies advanced in 1998 with the *small world* model created by

Watts and Strogatz, generated by re-wiring lattice structures which start with strictly non-random patterns of edges and progressively become more random [96]. The Watts-Strogatz method creates networks with strongly-connected local clusters of nodes, which are not typically present in other random networks, lending it practical applications in modelling neural networks, power grids and international collaborations because they also share this feature.

Random networks in general have properties typical of average graphs, and in particular the Erdős–Rényi random graph model has well-studied properties and strong practical applications, dating back to 1959 [97]. Erdős–Rényi random graphs $G_r(n,p)$ are created by specifying only two parameters: the total number of nodes n and the probability p that a random pair of nodes is connected. A typical use of these random networks is to compare them with real networks to rigorously evaluate the likelihood that non-trivial topological structures have emerged in the real data. This comparison in network analysis is treated similar to testing a null hypothesis for evaluating LHC signal significance using likelihood functions described in Section 3.3.3. In our LHC networks, we quantify the difference between SM and SM+BSM network metrics and therefore compare signal distributions with SM-only networks instead of random networks, although we do not exclude this potentially useful addition in future work.

Network metrics identify a range of graph characteristics which are not discernable by eye. Some distinguishable features include clustering, which refers to the presence of a set of closely-connected nodes, and centrality, which indicates the "popularity" of a node and is calculated from algorithms based on its number of connections with others. The most common centrality measure of popularity is *degree*, which is defined for a single node as the total sum of its edges. Another useful centrality metric is *betweenness*, which detects those nodes forming the most important pathways for travelling along connected paths through the network. These and other characteristic properties are defined in Section 4.3. Features such as clusters in high-dimensional data space can be difficult to interpret, prompting the need for effective network data analysis by defining rigorous indicators of these characteristics and calculating their significance.

Network metrics identify topological structures effectively in a diverse range of studies, including several on galaxy distributions [98–100]. Analyses of the spatial distributions of galaxies using network analysis revealed that distinct topologies can be distinguished using tools such as clustering and centrality measures of importance. In particular, filamentary structures bridging large clusters are identified from betweenness centrality in the topology of distinct galaxy classes, in agreement with known filament properties determined using other means such as n-point correlation functions and visual interpretations [101]. The authors conclude that network representations of galaxy distributions improve their ability to quantify topological features in the universe.

Characteristics which are present in some real networks but not others are an ideal method of discriminating between datasets. Constructing networks from real data may produce a typical property of structure and dynamics known as "scale-free" behaviour, discovered from a range of data types including social, biological, technological, transportation and information networks [102–104]. Scale-free networks possess the property that the degree distribution follows a power law, where the fraction of nodes with degree k is equal to $k^{-\alpha}$ for some $\alpha > 1$, and real networks may weakly or strongly possess this property. Highly central "hub" nodes are prevalent in real social and biological networks, and absent in others such as random networks and lattices. When hubs are absent, the degree centrality distribution is narrower as it does not possess a tail of highly-connected nodes, whereas scale-free networks with power-law degree distributions must possess hub nodes [105]. Other network metrics are also sensitive to hub behaviour and can discriminate between types of data with and without hub nodes, as these nodes heavily influence large-scale structure.

The types of analysis that can be performed with graph networks make network techniques advantageous in a particle collider search. The primary advantages are outlined below.

- Networks require few assumptions about the nature of the data. By designing a search that applies network analysis directly to LHC data with no additional constraints, we produce results which are independent of any assumptions besides the signal region definitions. One reason is because networks are computationally effective in treating discrete data, requiring no estimators of continuous functions pertaining to theory models.
- Networks encode information contained in the data at a scale chosen in the analysis design, which can be adjusted to describe an appropriate level of detail and complexity. A remarkable feature of a complex network is its heterogeneity, which is explored easily at different scales.
- The observational data used to create a network can be replaced by simulations, allowing for direct comparison. We can therefore apply network analysis in LHC searches where our strategy typically requires us to compare simulated backgrounds with measured collider data.

- The tools contained in network analysis provide immediate classification of individual events as well as more holistic structures such as clusters, comprising a wide range of possible discriminators. The complexity of large-scale structures is typically difficult to access via alternative analysis strategies.

Despite these benefits, our strategy for applying graph network analysis in an event-by-event LHC search for new physics is novel and has not been attempted until now. The existing studies into network analysis applications in collider physics are few and take a different approach.

Existing collider physics applications

The similarity between two collider events was evaluated at MIT in 2019 by developing a metric for the collider space based on the *Earth Mover's Distance* [106, 107]. This is only one example of a distance metric in collider space; many more standard distance metrics can be calculated from LHC data in an n-dimensional kinematic space. We investigate several useful metrics in our study, and apply the concept to construct a network of events, which has never before been attempted for the purpose of LHC analysis. Recently, graph networks have been applied as part of deep learning studies in particle physics [108–111], though these applications are distinct from our analysis approach. Some notable studies include modelling kinematics within an event for classification [112, 113], reconstructing tracks in silicon detectors [114], pile-up subtraction at the LHC [115], investigating multiparticle correlators [116], and particle reconstruction in calorimeters [117]. However, network applications to particle physics searches beyond deep learning capabilities remain largely unexplored.

4.1.3 Analysis strategy

We explore differences between the network properties of SM events and supersymmetry events due to their distinct origins by calculating powerful discriminating variables defined by graph theory. The network metrics are not highly correlated, and conveniently exist in both local and global form, where local metrics are calculated event-by-event and global metrics produce a single value characteristic of the entire network. Averaging over the local metrics for every node typically produces a global metric; for example, *transitivity* is a global metric calculated by averaging the local clustering [100]. We focus on local metrics only, leaving global metrics to future study.

Local network metrics can apply to LHC searches in two ways: the first is a model-dependent approach tuned to one BSM scenario, and the second is model-independent anomaly detection. In the first case, we can construct SRs from the local event-by-event variables to improve sensitivity to a rare physics signature. This allows us to quantify the increase in sensitivity derived from the network variables compared with a standard search. An advantage of local metrics is that they provide a new basis of variables on which we can place selection cuts in place of standard kinematics to increase signal yield as a straightforward test of their analysis power. In the second case, we can create a model-agnostic search by first calculating the network variables for the SM-only network before searching for deviations from these values in data, covering a large signal phase space. We focus our preliminary analysis on the model-dependent limit-setting case to generate results which are comparable with the standard methods described in Chapter 3. A possible approach to a model-independent anomaly detection network analysis is presented in Chapter 8.

Developing network analysis strategies in the novel context of ATLAS searches is a new technique which requires careful tests to prove that network metrics are meaningful when calculated from graphs in collider space. Considerations must be made for the particle collider context: for example, we require variables which possess indistinguishable distributions when calculated from the background-only network compared with the background part of the signal-background network. Adding signal should not change the background distribution significantly, otherwise the shapes of the background-only network metrics cannot meaningfully be compared with metrics calculated from real data. One possible method of achieving a consistent background shape is to search for signal which populates mostly different regions of collider space and is therefore not heavily connected to the background.

In addition, particle collider networks must contain a method of representing a larger number of events by a smaller set, because a simulated set of events cannot be made to contain the precise number of expected events for every given process type. Since the effective number of LHC events in a simulated sample is different to the size of the sample, our networks represent the effective number of events by adding event weights, a process which must also be tested to ensure these weights represent the intended information correctly. Conveniently, weights may be added to a network in two distinct ways, where one type of weight alters the relative importance of the nodes and the other alters the relative importance of the links. Tests of the weighting method and the background-only versus background-signal distributions, along with other tests which are significant to particle collider networks, are described in Section 4.2.3 and performed in Chapters 5-7.

4.2 Network construction

We build network graphs with LHC events from data or simulation where a pair of events is connected if they are nearby in the space of our chosen n-dimensional graphs. We define the concept of nearby events using a free parameter called *linking length*, which we set to the maximum distance between two connected events, thereby choosing a scale on which we can observe complexity in the patterns of connection. Given an n-dimensional dataset comprising a set of N entries with n attributes, we choose a method of constructing a network that reveals non-trivial behaviour in the network metrics. The main parameters involved in construction include the choice of n variables which are used to create the graph space, the method of calculating distance in that space, and the value of the distance nominated as the optimal choice of linking length for observing network complexity.

4.2.1 Definitions

A *network* is a mathematical data structure that is comprised of nodes connected by either directed or undirected edges. A directed edge passes from some node A to another node B where no edge may exist from B to A, whereas an undirected edge must connect two nodes in both directions equally. We consider finite, undirected graph networks where two events cannot be connected in one direction only. The graph is denoted by $G = (\mathscr{N}, \mathscr{E})$, where $\mathscr{N}$ is the set of nodes, and $\mathscr{E} \subseteq \{\{i, j\} : i \neq j \in \mathscr{N}\}$ is the edge list. Figure 4.1 depicts a toy graph network visualisation comprising nine nodes and thirteen undirected edges.

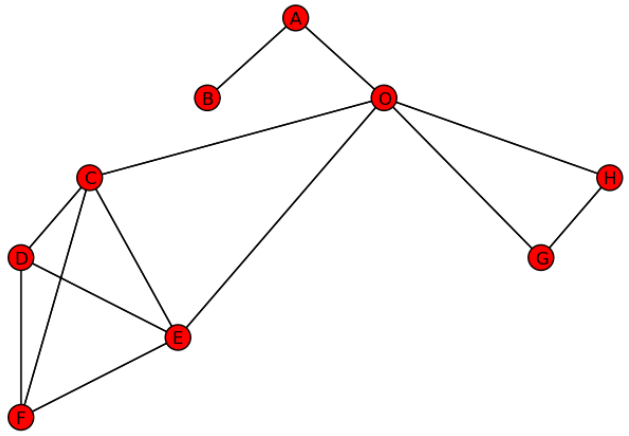


Fig. 4.1 A pictorial representation of a graph network in two dimensions, where node O is a hub with a high degree centrality of 5, B is isolated and has low centrality, and C, D, E and F form a highly connected cluster.

A convenient form of representing large graphs is the *adjacency matrix*, which has the list of nodes as the rows and columns. Connected nodes have a 1 in the relevant adjacency matrix entry, indicating the presence of an edge, while disconnected nodes have a 0. We may write this as $\mathbf{A} = (a_{ij})_{i,j\in\mathscr{N}}$, where $a_{ij} \in \{0,1\}$ and $a_{ij} = 1$ if $\{i, j\} \in \mathscr{E}$, and we note that the adjacency matrix will be symmetric in our case because our LHC event networks are undirected. The *neighbours* of a node $v \in \mathscr{N}$ are those that are linked to the node by an edge, defined via:

$$\mathscr{N}_v = \{i \in \mathscr{N} : a_{iv} = 1\} \tag{4.1}$$

These are also referred to as the members of v's *punctured neighbourhood*. We can define an *extended adjacency matrix* $\mathbf{A}^+ = (a^+_{ij})_{i,j\in\mathscr{N}} = \mathbf{A} + \mathbf{I}$ with:

$$a^+_{ij} = a_{ij} + \delta_{ij} \tag{4.2}$$

where $\mathbf{I}$ is the identity matrix, and δ_{ij} is the Kronecker delta symbol. This then allows us to define the *unpunctured neighbourhood* of v as the set of nodes which includes both the neighbours of v, and v itself:

$$\mathcal{N}^+ = \left\{ i \in \mathcal{N} : a^+_{iv} = 1 \right\} = \mathcal{N}_v \cup \{v\} \tag{4.3}$$

In our analysis, we will define the adjacency matrix for LHC events by defining a distance between events in the space of kinematic variables that are measured for each event. Assuming some distance d_{ij} between the nodes in this space of variables, we then define the adjacency matrix:

$$a_{ij} = \begin{cases} 1, & \text{if } d_{ij} \leq l, \\ 0, & \text{otherwise,} \end{cases} \tag{4.4}$$

where l is the linking length. This prescription clearly leaves many choices open for how to proceed, including the choice of kinematic variables, the choice of distance metric, and the choice of linking length for a given analysis. Both the choice of distance metric and the choice of kinematic variables will change the topological structure mapped by the network, and sensible choices might lead to greater differences between the behaviour of SM processes and new physics processes within the network. The variety of options is an advantage of network analysis, as it allows the user to select prescriptions best suited to the data.

The potential sensitivity of a network analysis search for BSM physics in collider data depends upon the topological differences between the SM and the SM+BSM networks, which in turn builds upon differences encoded in their distributions in the chosen n-dimensional space of kinematic attributes. To test the effect of spread in the kinematic space on the optimum choice of distance metric and linking length, we generated toy data samples simulating LHC signal and background event distributions in several dimensions. An example of two toy simulations comprising 1000 events each is visualised in two dimensions in Figure 4.2. Extensions of this simulated data are then plotted in graph form to compare possible distance calculations in Section 4.2.2.

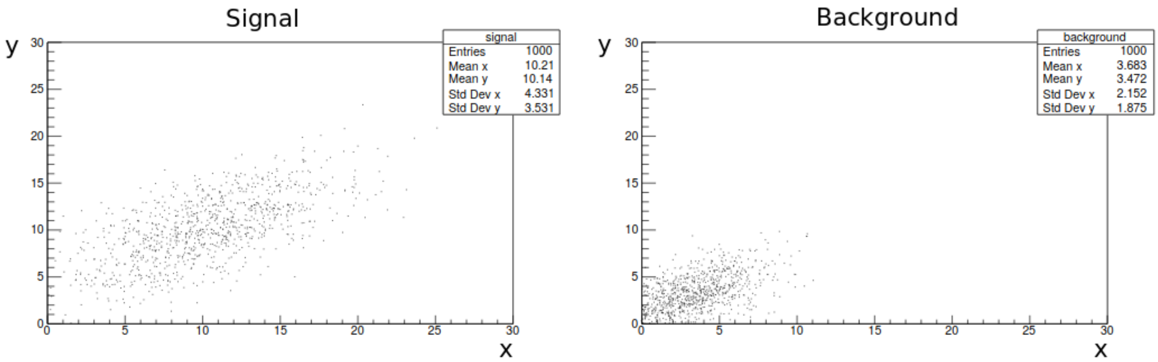


Fig. 4.2 Example distributions are plotted in two dimensions x and y, simulated to represent the difference between BSM signal and SM background events in kinematic space. Real signal and background events are likely to populate a different, although overlapping, region of collider space.

4.2.2 Network parameters

The network construction leaves open the choice of kinematic variables, where the spatial distribution of LHC data may differ significantly depending on the chosen basis. It is a challenging task to optimise this basis to maximise the discriminating power of the network for a particular dataset, and one that does not necessarily translate well into good separation for other datasets. In a proof-of-principle study, measuring the effects on signal-background separation of the choice of basis variables for several datasets is beyond our scope. We instead gain greater insight into the sensitivity of the strategy by varying other factors that have a more predictable and quantifiable influence on the network. Our focus lies in varying spatial distance measures, linking length and choice of local network metrics while we study a fixed set of basis variables typical in supersymmetry analyses.

Distance metrics

The choice of distance metric sets the length between events in the space by tracing a path along a manifold. Testing several different distance metrics, we search for topological features of the resulting networks which distinguish signal from background. For two vectors u and v in the space of n kinematic variables for an analysis, some useful distance metrics include:

- **The Euclidean distance:** $d_{\text{euc}} = \sqrt{\sum_{i=1}^{n}(u_i - v_i)^2}$.

- **The Chebyshev distance:** $d_{\text{cheb}} = \max|u_i - v_i|$, i.e. the maximum of the difference between similar kinematic variables for the two chosen points.
- **The Bray-Curtis distance:** $d_{\text{bray}} = \sum_{i=1}^{n} \frac{|u_i - v_i|}{\sum_{j=1}^{n}|u_j| + \sum_{j=1}^{n}|v_j|}$.
- **The cityblock distance:** Also known as the Manhattan distance, the cityblock distance is given by $d_{\text{city}} = \sum_{i=1}^{n}|u_i - v_i|$.
- **The cosine distance:** $d_{\cos} = 1 - \frac{u \cdot v}{\sqrt{u \cdot u}\sqrt{v \cdot v}}$.
- **The Canberra distance:** $d_{\text{can}} = \sum_{i=1}^{n} \frac{|u_i - v_i|}{|u_i| + |v_i|}$.
- **The Mahalanobis distance:** $d_{\text{mah}} = \sqrt{(u-v)V^{-1}(u-v)^T}$, where V^{-1} is the inverse of the sample covariance matrix (calculated over the entire dataset of events).
- **The correlation distance:** $d_{\text{corr}} = 1 - \frac{(u-\bar{u}) \cdot (v-\bar{v})}{\sqrt{(u-\bar{u})^2}\sqrt{(v-\bar{v})^2}}$, where $\bar{u}$ is the mean of the elements of the vector u.

Other distances exist [118], but the above list was found to be sufficient for our purposes and relatively easy to calculate. Each distance was calculated from toy datasets, both with and without "event" weights, to evaluate its characteristics and any relevant constraints. All were discovered to be useful except the Canberra distance, which was ineffective and is not referred to in the later analysis. When events are weighted, a vital consideration follows that the weights must be propagated into the Mahalanobis covariance matrix. An example of a toy distribution for the cityblock distance calculated from an extended version of the data in Figure 4.2 is contained in Figure 4.3.

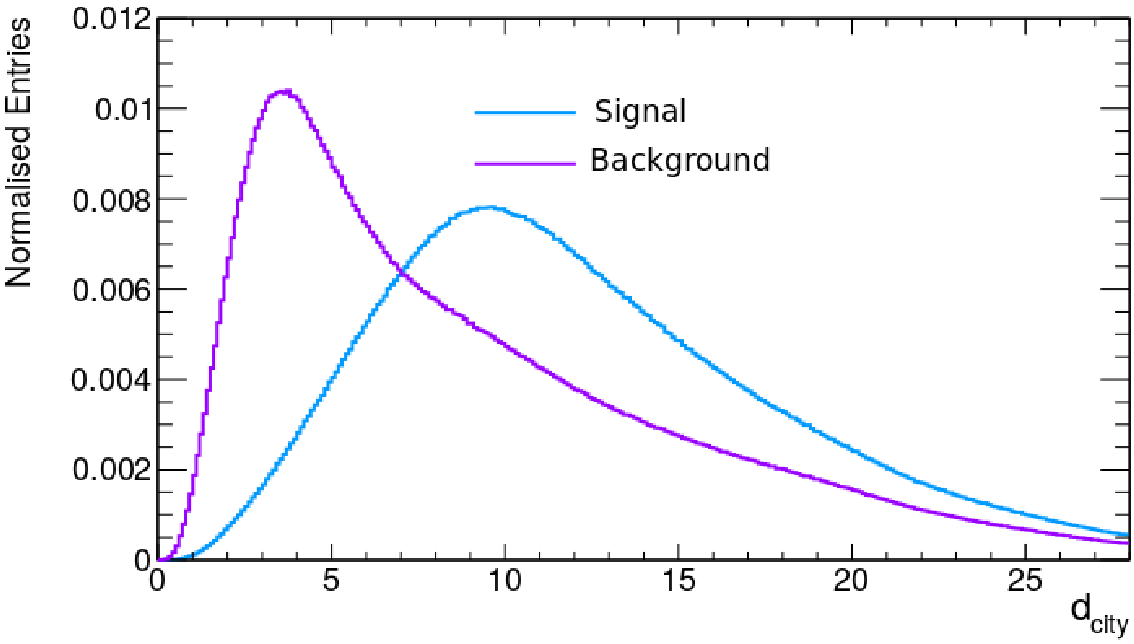


Fig. 4.3 The distribution of cityblock distances calculated between every node pair in an example network created from 10000 “signal” and 10000 “background” events in a 5-dimensional space.

Although significant differences clearly exist between these distance distributions without any need to construct and analyse a network, they are only comparable here because they contain the same number of “signal” and “background” events. In real data, signal contributions are rare and therefore insignificantly small by comparison to the number of background events present in the same histogram bin, regardless of how each is distributed in collider space. Therefore, network approaches are required to build upon these differences and create statistically significant measures.

Linking length choice

The linking length parameter sets the scale of connections depending on typical distances in collider space, and is therefore distance-metric-dependent. An appropriate linking length reveals non-trivial complexity, and may take one of several possible values. A small length will result in a disconnected network showing little structure, and one that is too large will not reveal central nodes or local peculiarities. In networks where important topological structures exist only on a narrow range of spatial scales, the linking length is a straightforward choice.

However, others contain spatially diverse topological structures where non-trivial complexity appears on a range of scales. For LHC event networks, SM backgrounds are less likely to produce spatially diverse structures than networks containing both SM background and BSM signal. Since these are intended as tools to isolate signal, they are given linking lengths which set the spatial scale to a typical distance at which signal events are non-trivially connected, but background events are even more connected. This strategy ensures that the networks detect the difference in levels of connectivity between background and signal to help distinguish them topologically. The linking length is therefore set to a characteristic scale of separation of background events, equivalent to selecting a value on the upper edge of the background distribution peak in Figure 4.3.

Node weights

Networks manage weights on either edges or nodes, where the two strategies can represent different effects. Edge weights adjust the strength of connections between nodes when some are more likely to "interact" or be related for reasons not captured by their spatial locations. For example, neural networks often apply edge weights to represent the number of synapses and gap junctions between neurons [96], and transportation networks weight edges by the amount of traffic flow between locations [119]. Conversely, node weights adjust the effective amount of data represented by a single node. In the LHC case we clearly require node weights without edge weights, as introduced in Section 4.1.3. Node weights affect how the network metrics are calculated, following a technique described in Section 4.3.

4.2.3 Node-weighted networks

Node weights can treat certain events as representative of a much larger sample when some types have higher predicted rates of production at the LHC than is reflected in their sampling. The weights effectively increase the number of these nodes to signify that more are present than could be generated in an MC sample. We find the node-weighting strategy useful when changing the absolute numbers of nodes in the networks is not viable for computational reasons. The popular theory where edge weights are used instead is not equivalent to using node weights, and although some methods exist of translating node weights into edge weights [120], in general measurements made on these networks are not invariant under this translation. For a sample of events in kinematic space represented by the yellow points in

Figure 4.4, we may re-sample with less data and report the blue points instead, so that a network of the blue data with high node weights accurately represents the yellow set.

Fig. 4.4 Blue points are a smaller sampling from the distribution which created the yellow points, and receive higher weighting to represent a larger effective number of kinematically similar events.

Simulating rare kinematics

MC simulations naturally produce fewer events with rare kinematics, which can result in poor sampling of events in the tails of the kinematic distributions. We increase the statistics in these poorly-sampled regions by binning the kinematic distribution into slices and increasing sampling in the rare tail slices, then assigning these events a low weight according to their predicted cross-section. High weights are assigned instead to the events sampled within the bulk distribution slices where more events are expected. Weights ensure that the kinematic distributions of the sliced samples retain the shapes of the inclusive distributions and are consistent with physics arguments. Consequently, events in the tails become a more reliable representation of the space.

For comparison, we generated inclusive samples in the same regions of phase space to compare with binned ("sliced") events, verifying that the shapes of their kinematic distributions are consistent. Figure 4.5 compares the distributions of the average transverse momentum of outgoing particles, $\hat{p}_{\mathrm{T}}$, for inclusive and sliced electroweakino production event samples simulated with `Pythia` and `Delphes`. The sliced histogram contains a combination of events from three distinct bins whose limits were user-specified at generator-level. The `Pythia` event generator assigns a cross-section to every bin, so that events in a given slice possess a known proportion of the total stop production process cross-section.

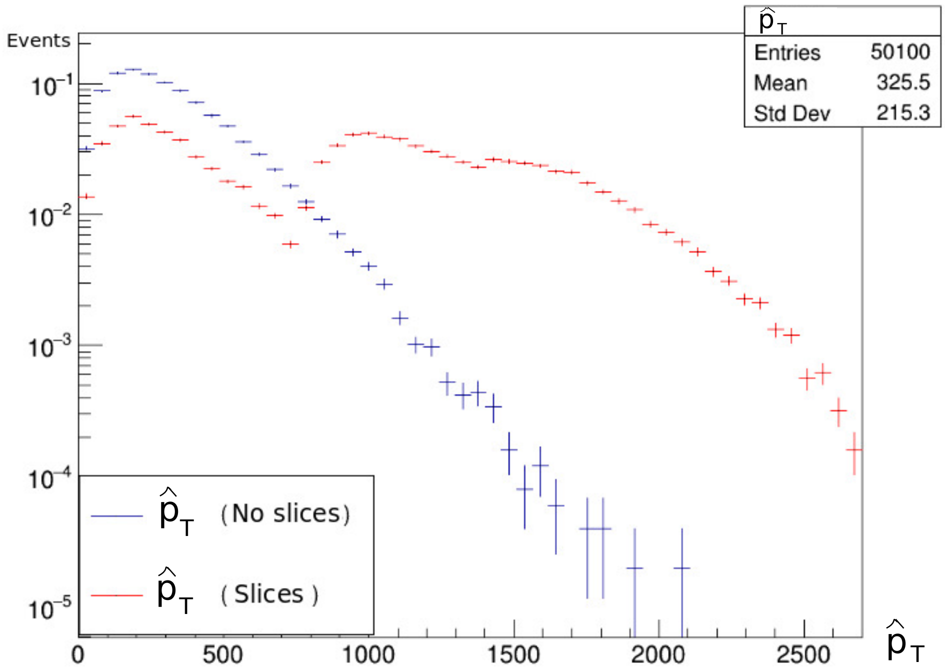


Fig. 4.5 A comparison of a typical electroweakino production process generated with Pythia for $m_{\tilde{\chi}_1^\pm} = m_{\tilde{\chi}_2^0} = 400$ GeV and $m_{\tilde{\chi}_1^0} = 0$ GeV (details in Section 5.1), with events biased to high momentum using slices, shown in the $\hat{p}_\mathrm{T}$ distribution.

The excess of events at high $\hat{p}_\mathrm{T}$ is the effect of biasing the event sample towards highly energetic events. To return the distribution to the inclusive sample shape, we assign cross-section weights to every event according to the prescription from Section 3.3.2, producing the red histogram from Figure 4.6.

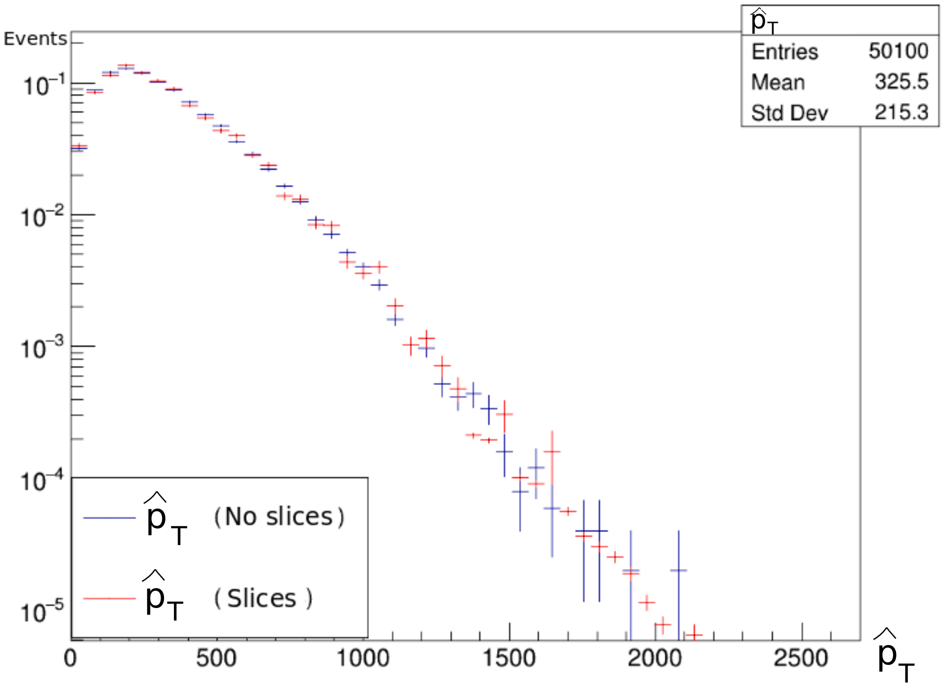


Fig. 4.6 Electroweakino production with slices in $\hat{p}_T$ re-weighted from Figure 4.5 to match the un-biased process cross-section shown in blue with reduced statistical fluctuations in the tail region.

Monte Carlo re-weighting

The weighting prescription must account for non-trivial event weights derived from multiple sources, including integrated luminosities for the simulated LHC processes and the cross-sections of slices. The weight on each slice therefore becomes a product of three factors: the LHC luminosity weight, the MC cross-section weight and a normalisation factor. The first is the LHC weight scaling the process to the correct number of events expected at a given luminosity. This is equal to the product of the process cross-section and the integrated luminosity, divided by the initial number of stop events before preselection N_{init}. Preselection criteria applied at generator- and detector-level decrease the proportion of events suitable for the analysis, and the algorithm must then generate a larger initial number N_{init} before achieving the required number that fulfil the criteria.

The second factor depends on the MC cross-section output by `Pythia` per slice, updated after each run with reduced uncertainty. The resulting event weight for a given slice equals this cross-section divided by the total number of events in the slice before preselection. The third factor is normalisation intended to scale all slices by a value proportional to the inverse of the MC cross-section factor, cancelling its magnitude so that the full weighting procedure produces a sum of all event weights equal to the number of events expected from the LHC. The normalisation factor for a given process is equal to the total number of process events after preselection, typically $N =$10,000 in our analysis, divided by the sum of the MC cross-section weight factors from all slices. The outcome is to set the sum of the weights to N events multiplied by the LHC weight factor rescaling to the correct final count.

The product of the three factors produces event weights which vary by several orders of magnitude, and we must calculate the network metrics using robust algorithms under such heavy re-weighting of nodes. We verify that the metrics are reliable under this re-weighting by methods which will be revisited in Section 4.3.1. One component is a straightforward comparison between the network metric distributions of the sliced metrics and those calculated from inclusive samples, reminiscent of the comparison of the kinematic distributions in Figures 4.5 and 4.6. The inclusive samples receive only the first weight factor equal to the LHC luminosity weighting, which is large compared with the combination of all three factors on the sliced events. Robust network metrics must handle these large weight disparities so that their distributions are unchanged in shape when calculated from inclusive compared with sliced events of the same process type. Understanding the comparison between network variables calculated from sliced and inclusive samples requires us to consider the role of node weights in network calculations.

4.3 Network metric definitions

We employ several local network metrics that draw their calculations from different key properties of a network, investigating their power of discrimination between signal and background in LHC data when combined in a single analysis. Our choice of local metrics involves attributes defined event-by-event, unlike global metrics which calculate properties of the whole network. We therefore produce binned histograms plotting the distribution of our metrics where the number of entries per histogram is equal to the number of events in the network, characterising the importance of events relative to their neighbours instead of characterising the shape of the network in a set of single quantities. Focusing only on local metrics, we can combine the local calculations for all events in the network and the different properties they measure to form a holistic picture of network geometry and topology.

For the reasons described in the previous section, any network metrics we calculate must incorporate node weights, and must therefore be invariant under the transformation between changing the number of nodes of some type by some factor and multiplying the weights on the original set of nodes by a matching factor. Figure 4.7 represents the weight on a node by its size, as a pictorial depiction of relative importance, where the more highly weighted nodes should receive greater contribution to the distribution of any of the network metrics. Nodes with weight 3 and 4 also happen to be more central than the boundary nodes with weight 1, which means that to calculate the network metrics, we must count 4 nodes with number of edges equal to 2 for the largest node and 3 for the second-largest. By weighting nodes, we assume that each of the effective nodes which are represented at the same network position possesses the same number of links to external nodes. We require a method of computing network metrics which treats every node as a sub-network of interlinked nodes with identical connections, so that we create a network of interacting networks.

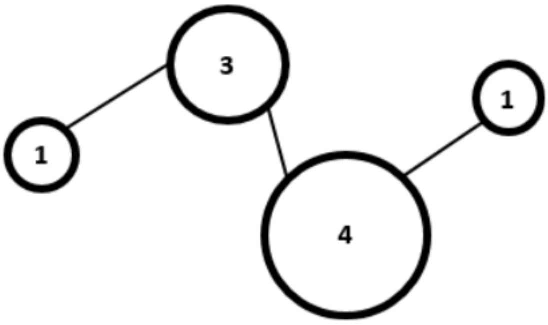


Fig. 4.7 A pictorial representation of relative node weights, where weight is represented by node size. The central nodes with high weights have a large impact on the network topology, which must be considered in calculations of network metrics.

The convenient method of calculating the network metrics with node weights represented correctly is to define typical graph theory metrics using the node-splitting invariance (n.s.i) method [121]. This and similar methods have proven useful in generating precise estimations of network characteristics in climate research to evaluate patterns in rainfall and ocean currents, and in economic models to provide accurate information on trade balance [122–124]. The n.s.i approach assumes that every node v represents a portion of the domain of interest which is represented by its weight w_v. If the domain of interest is equivalent to the underlying n-dimensional graph space given in the LHC case by a set of n kinematic attributes, then this continuous space is divided into $\mathscr{N}$ portions, where $\mathscr{N}$ is the number of events in the sample. Each node $v \in \mathscr{N}$ then receives a weight w_v such that the sum of the weights, $\sum_{v\in\mathscr{N}} w_v$, is equal to the effective number of events.

Under the n.s.i assumptions, events do not have a well-defined location in collider space, and their possible locations are constrained by the boundary of the area w_{nu}. Figure 4.8 represents three separate sets of nodes by the orange, blue and green regions of the domain, labelled V_i, V_j, and V_k. Each node v represents a portion of the domain belonging to the region assigned to its node set, and edges can connect regions which are either adjacent (orange and blue) or distinct (for example, blue and green). If the largest distance apart between points within the same area w_v is less than the linking length, then nodes lying within that area would be connected regardless of n.s.i, lending strength to the claim that n.s.i is an appropriate method of representing a large sample of effective nodes by a smaller sample of network nodes with large weights. In this case, any inhomogeneities inside the region represented by an individual network node are irrelevant.

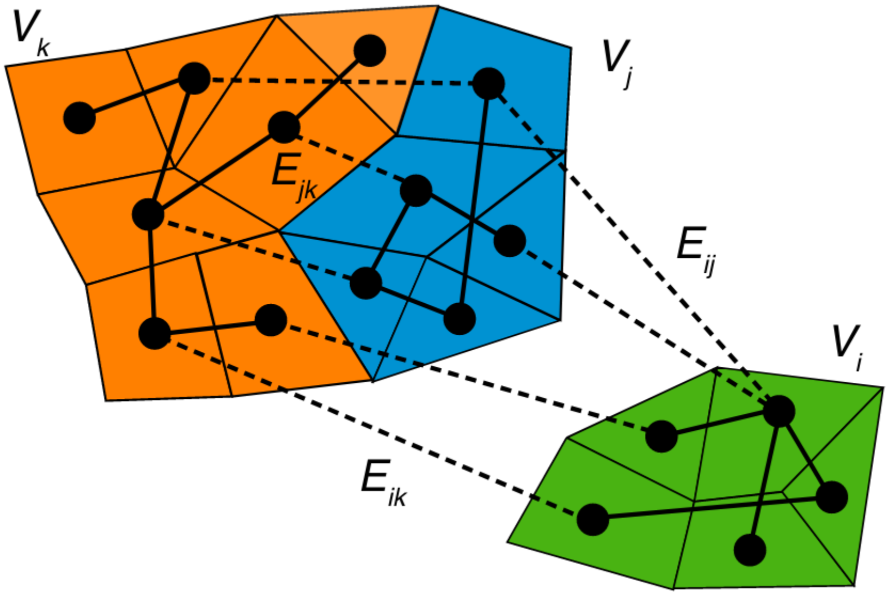


Fig. 4.8 A network representation of a data distribution where nodes are located in distinct regions of the underlying graph space [121].

The representation of a larger set of nodes by a smaller set is called a *refinement*, and proceeds via the n.s.i technique of node splitting [121]. The concept of splitting replaces a node v with two connected nodes $v^{'}$ and $v^{''}$ assigned the weights $w_{v^{'}} + w_{v^{''}} = w_v$. In Figure 4.9, the node splitting operation transforms graph G into graph $G^{'}$ by assigning new nodes with smaller weights which sum to the original and are connected by an edge to each other and identically to both external nodes to which v was connected. The operation can be repeated iteratively many times, first on the graph $G^{'}$ to produce a graph $G^{''}$ with three split shaded nodes, and so on to create many more nodes. The inverse operation is *twin merging*, which decreases the number of linked nodes to transform the graph iteratively back into G.

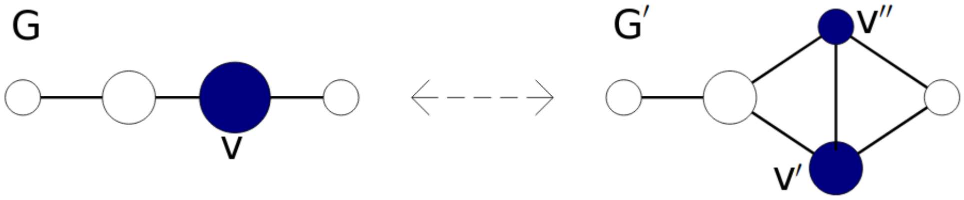


Fig. 4.9 The dual operations of node splitting and merging transform graph G into $G^{'}$ and vice versa, respectively. The n.s.i technique applies both operations to calculate network characteristics from node-weighted networks [121].

The network metrics are conveniently calculated under the n.s.i paradigm using a custom version of the `pyunicorn` package [122]. `Pyunicorn` is an open source analysis toolbox designed for data analysis and modelling in complex network theory. We use these metrics to search for non-trivial shape properties and investigate the complexity of structures on different scales.

Centrality metrics

- **The n.s.i degree:** For a given node v, this is the weighted version of the degree centrality (DC), given by:

  $$k_v^* = \frac{\sum_{i \in \mathcal{N}_v^+} w_i}{(W+1)}, \tag{4.5}$$

  where $W = \sum_{i \in \mathcal{N}} w_i$ is the sum of the weights of all nodes in the network. In our analysis, W is equivalent to the total number of events expected at the LHC for our assumed integrated luminosity.

  Relative degree in a toy network diagram is visible using the scale in Figure 4.10, where central nodes are defined as those with more edge links.

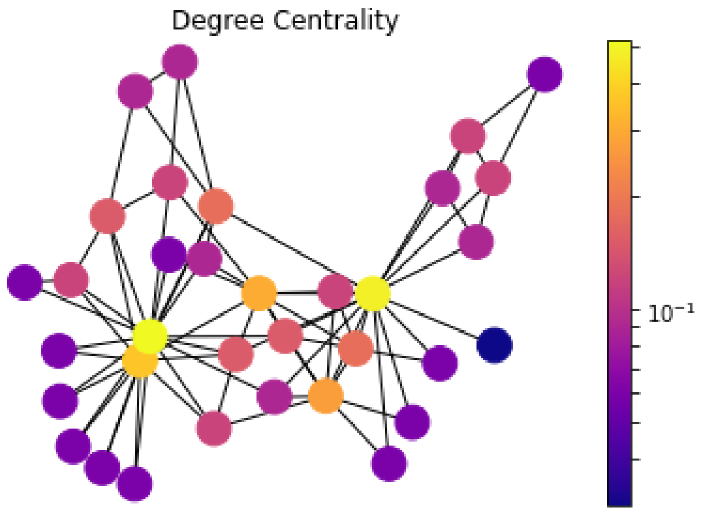


Fig. 4.10 The degree centrality calculated for a toy node distribution using the `NetworkX` Python library [125].

- **The n.s.i average and maximum neighbours degree:** The average neighbours degree of a node v represents the average size of the network region that an event linked to v is linked to. The n.s.i measure of this quantity is given by:

$$k^*_{nn,v} = \frac{1}{(W+1)} \frac{\sum_{i\in\mathscr{N}_v^+} w_i k_i^*}{k_v^*}. \tag{4.6}$$

  An n.s.i *maximum* neighbors degree can be defined as:

$$k^*_{nnmax,v} = \max_{i\in\mathscr{N}_v^+} k_i^*. \tag{4.7}$$

- **The n.s.i betweenness centrality:** The *shortest path betweenness centrality* of a node v is a measure of the proportion of shortest paths between all randomly selected pairs of nodes in the network that pass through node v. Betweenness can be explained using a real-life traffic network: if there is only one straight road connecting two cities, then all traffic will travel by this road to arrive at the other city. The betweenness of nodes lying on this road is therefore higher than other nodes, which does not strictly coincide with the degree centrality metric, as shown in comparing Figures 4.10 and 4.11.

  If we randomly choose nodes labelled by a and b, we have that:

$$BC_v = \langle n_{ab}(v)/n_{ab} \rangle_{ab} \in [0,1], \tag{4.8}$$

where n_{ab} is the total number of shortest paths from a to b, $n_{ab}(v)$ is the number of those paths that pass through v, and we have defined the average of a function of node pairs by $\langle h(i,j)\rangle_{ij} = \frac{1}{N^2}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} h(i,j)$. A formal expression for this quantity is written by first noting that n_{ab} can be written as a sum over the *tuples* $(t_0,...,t_{d_{ab}})$, with $t_0 = a$ and $t_{d_{ab}} = b$ (d_{ab} is the number of links between the nodes a and b on the shortest path), where each tuple in the sum gives a contribution of 1 if every node t_l in the tuple is linked to its successor t_{l+1}, or 0 if at least one node is not linked to its successor. Both of these conditions are met if we simply take the product of elements of the adjacency matrix for each pair of nodes in the tuple, allowing us to write

$$n_{ab} = \sum_{(t_o,...,t_{d_{ab}})\in\mathcal{N}^{d_{ab}+1}, t_0=a, t_{d_{ab}}=b} \prod_{l=1}^{d_{ab}} a_{t_{l-1}t_l}. \tag{4.9}$$

$n_{ab}(v)$ is given by a similar formula, except that, for some m in $1...d_{ab}-1$, t_m must equal v:

$$n_{ab}(v) = \sum_{m=1}^{d_{ab}-1} \sum_{(t_o,...,t_{d_{ab}})\in\mathcal{N}^{d_{ab}+1}, t_0=a, t_m=v, t_{d_{ab}}=b} \prod_{l=1}^{d_{ab}} a_{t_{l-1}t_l}. \tag{4.10}$$

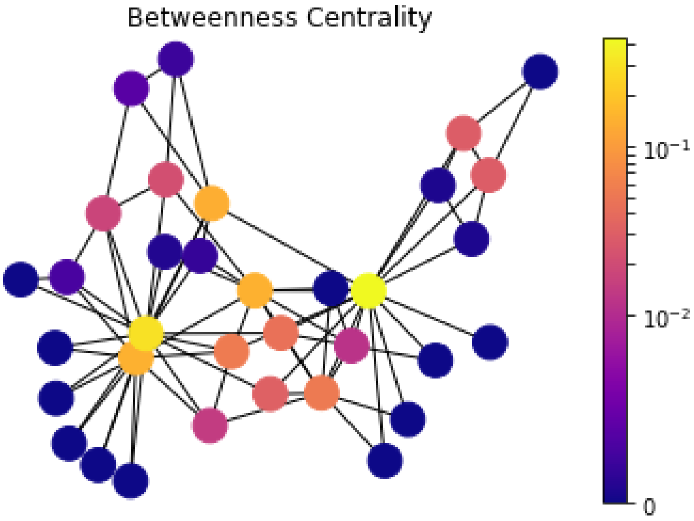


Fig. 4.11 The betweenness centrality metric is high for nodes which are common to many paths between randomly chosen nodes, calculated with the `NetworkX` Python library.

It is possible to make an n.s.i version of this quantity based on a *weighted average* instead of an average (which would be consistent with the above formula in the limit that all node weights are 1), but the `pyunicorn` package instead uses a weighted sum, giving the n.s.i betweenness centrality as:

$$BC^*_v = \langle n^*_{ab}(v)/n^*_{ab}\rangle^{wsum}_{ab} \in [0, W^2/w_v], \tag{4.11}$$

where we have defined the *weighted sum* of a function of pairs of nodes $\langle h(i,j)\rangle^{wsum}_{ij} = \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} w_i h(i,j) w_j$, and $n^*_{ab}(v)$ and n^*_{ab} are defined below. The n.s.i betweenness centrality values obtained for our examples below do not come close to saturating the maximum value of W^2/w_v.

Formulae for $n^*_{ab}(v)$ and n^*_{ab} can be derived as follows. If a node s is hypothetically split into two nodes $s' + s''$, any shortest path through s becomes a pair of shortest paths (one of which passes through s', and the other of which passes through s''). In addition, a shortest path from s'' to some $b \neq s'$ will never meet s'. Thus, the betweenness centrality can be made invariant under node splitting by making each path's contribution proportional to the product of the weights of the inner nodes, but with the condition that we skip the weight w_v in this product when calculating $n_{ab}(v)$. Formally, we can write a modified n^*_{ab} as:

$$n^*_{ab} = \sum_{m=1}^{d_{ab}-1} a_{t_0t_1}\prod_{l=2}^{d_{ab}}(w_{t_{l-1}}a_{t_{l-1}t_l}), \tag{4.12}$$

and a modified $n^*_{ab}(v)$ as:

$$n^*_{ab}(v) = \frac{1}{w_v}\sum_{m=1}^{d_{ab}-1}\sum_{(t_o,\dots,t_{d_{ab}})\in\mathcal{N}^{d_{ab}+1},t_0=a,t_m=v,t_{d_{ab}}=b}\left(a_{t_0t_1}\prod_{l=2}^{d_{ab}}(w_{t_{l-1}}a_{t_{l-1}t_l})\right). \tag{4.13}$$

A geometric interpretation of the n.s.i betweenness can be obtained by considering the set of nodes of our node-weighted network $\{\mathcal{N}\}$ as a sample from a population of points $\{\mathcal{N}_0\}$. Each node v in the network then represents some small cell $\mathcal{R}_v$ of points in the geometric vicinity of v in $\{\mathcal{N}_0\}$. The n.s.i betweenness can be interpreted as an estimate of the probability density that a randomly-chosen shortest path between two randomly-chosen points in the population network passes through a specific randomly-chosen point in $\mathscr{R}_v$. This is ultimately a measure of the importance of the node v in the network.

- **The n.s.i closeness centrality:** Another measure of node importance is the *closeness centrality*, which is defined for a node v by $CC_v = 1/\langle d_{vi}\rangle_i$, where d_{vi} is the number of links on a shortest path from v to i, or, if there is no path, ∞, and we have defined the same average over nodes that was used previously. A larger value of CC_v for a node indicates a smaller average number of links from v to another node in the network. Although closeness also measures centrality, it detects different properties to the degree and betweenness, as seen in Figure 4.12. The n.s.i version of this metric is given by:

$$CC_v^* = \frac{W}{\sum_{i\in\mathcal{N}} w_i d_{vi}^*} \in [0,1], \tag{4.14}$$

where d_{vi}^* is an *n.s.i distance function* given by:

$$d_{vv}^* = 1 \text{ and } d_{vi}^* = d_{vi} \text{ for } i \neq v. \tag{4.15}$$

The n.s.i distance function can be justified as follows. A weighted version of CC_v should give us the inverse average distance of v from other weight units or points rather than from other nodes. But for this to become n.s.i, each node must have a unit (instead of zero) distance to itself since, after an imagined split of a node $s \to s's''$, the two parts s' and s'' of s have unit not zero distance.

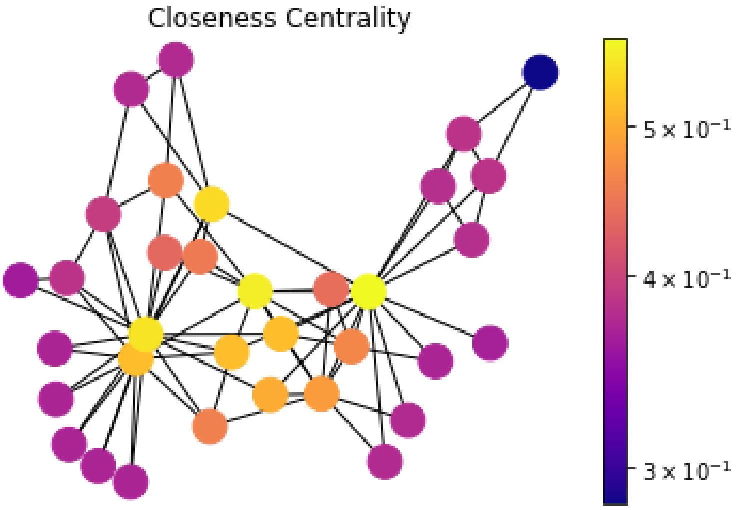


Fig. 4.12 The closeness centrality metric is high for a node which is more directly connected to other nodes, so that its shortest paths pass through fewer nodes, calculated with the `NetworkX` Python library.

- **The n.s.i exponential closeness centrality:** A limitation of the closeness centrality is that it receives very low values for nodes which are very close to most of the other nodes, but very far away from at least one of them. To prevent outlying nodes from skewing the closeness calculation for "typical" nodes, we can use the *exponential closeness centrality*, defined as $CC_{EC,v} = \langle 2^{-d_{vi}} \rangle_i$. The n.s.i measure is given by:

$$CC^*_{EC,v} = \left\langle 2^{-d^*_{vi}} \right\rangle_i^w \in [0,1]. \tag{4.16}$$

  where we have used the notation $\langle g(v) \rangle_v^w = \frac{1}{W} \sum_{v \in \mathcal{N}} w_v g(v)$.

- **The n.s.i harmonic closeness centrality:** The harmonic closeness centrality reverses the sum and reciprocal operations in the definition of closeness centrality, such that $1/d_{vi}$ contributes zero to the sum if there is no path from v to i. The n.s.i harmonic close centrality is given by:

$$CC^*_{HC,v} = \langle 1/d^*_{vi} \rangle_i^w \in [0,1]. \tag{4.17}$$

Correlation metrics

Correlation metrics measure properties of grouping in contrast to the importance measured by centralities. Local correlation is estimated from the clustering coefficient, which measures the arrangement of a network into local groups. It quantifies the abundance of triangular patterns of connection, measuring the short-range order within the local vicinity of the node. A triangle is formed when a pair of connected nodes are both connected to a third node, and the number of these shapes relative to the total number of node triples is a useful descriptor of local structure. In Figure 4.13, the two edges r_1 and r_2 may form a triangular closed triple with r_{12}, only if this edge exists. If the edge r_{12} is disconnected, then the three nodes do not contribute positively to the clustering coefficient metric.

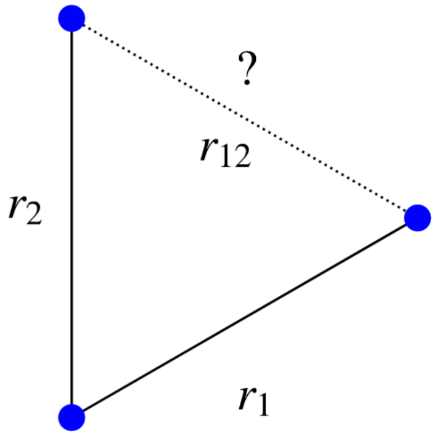


Fig. 4.13 A schematic diagram for describing clustering coefficient in a graph network, where connecting the edge r_{12} results in the addition of a closed triple, or triangle [100].

Correlation metrics are designed to detect the differences between the two networks shown in Figure 4.14. Highly correlated networks have a large clustering coefficient, and may take the form of community groups, a regular lattice structure or a single large cluster. In a lattice network, a high clustering coefficient may indicate a small world network if the linking length is small; correlation metrics are used to measure small-worldedness, which produces several interesting properties [96] mentioned previously in Section 4.1.2. We calculate two clustering coefficients, *local* and *local Soffer* clustering.

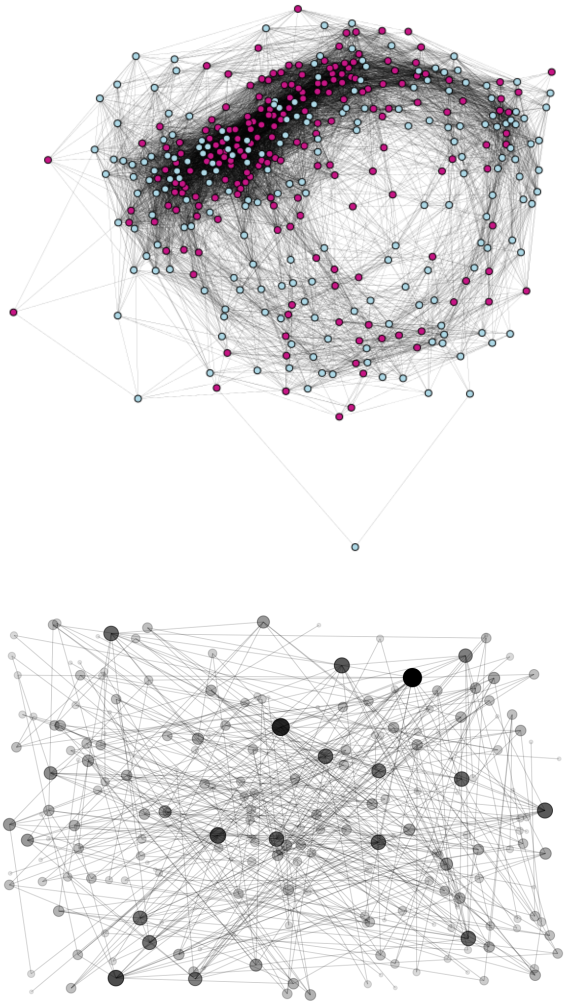


Fig. 4.14 Above: two types of simulated LHC event samples mixed in one network as an example of strong clustering in real data. Below: a subsection of a random network showing minimal clustering. Created with the `NetworkX` library.

- **The n.s.i local clustering coefficient:** The local clustering coefficient of a node v is the probability that two nodes drawn at random from those linked to v are linked with each other. It is given by:

$$C_v = \frac{\sum_{i\in\mathcal{N}_v}\sum_{j\in\mathcal{N}_v} a_{ij}}{k_v(k_v-1)} = \frac{N^2\left\langle a_{vi}a_{ij}a_{jv}\right\rangle_{ij}}{k_v(k_v-1)}. \tag{4.18}$$

  The n.s.i version is given by:

$$C^*_v = \frac{W^2 \left\langle a^+_{vi} a^+_{ij} a^+_{jv} \right\rangle^w_{ij}}{k^{*2}_v} \in \left[\frac{w_v (2k^*_v - w_v)}{k^{*2}_v}, 1 \right] \subseteq [0,1]. \tag{4.19}$$

- **The n.s.i local Soffer clustering coefficient:** An alternative form of the clustering coefficient proposed by Soffer and Vázquez [126] includes a correction that reduces the impact of degree correlations:

  $$C_{s,v} = \frac{N^2 \left\langle a_{vi} a_{ij} a_{jv} \right\rangle_{ij}}{\sum_{i \in \mathcal{N}_v} (\min(k_i, k_v) - 1)} \in [C_v, 1]. \tag{4.20}$$

  The n.s.i version of this is given by:

  $$C^*_{s,v} = \frac{W^2 \left\langle a^+_{vi} a^+_{ij} a^+_{jv} \right\rangle^w_{ij}}{\sum_{i \in \mathcal{N}^+_v} w_i \min(k^*_i, k^*_v) \in [C^*_v, 1]}. \tag{4.21}$$

Different clustering metrics detect different network properties, just as each centrality metric contributes new information. In addition, local clustering and centrality metrics measure vastly different network characteristics, and are typically not highly correlated. Local Soffer clustering in particular is an algorithm that was developed to minimise correlations with the degree centrality metric [126]. In real networks the clustering coefficient generally decreases with node degree, which is a signature of the network hierarchical structure. Degree-correlation biases in the clustering coefficient definition are removed by introducing the local Soffer clustering, designed to accurately represent clustering topologies in real networks which does not necessarily decrease with degree.

4.3.1 Impact of large weights on node-splitting invariant metrics

The network metrics calculated by n.s.i methods must assume that a weighted node can be equivalently represented by a group of unweighted nodes under two conditions:

1. full *internal* connectivity, and
2. identical *external* connectivity,

where internal connections occur between "effective nodes" that are represented by the same weighted node in a subsample, and external connections are edges between different subsampled nodes. To satisfy condition 1, a group of X internal nodes must all lie within a distance of the linking length l in order to be represented in a weighted network by an event of weight X. Satisfying condition 2 requires that all X nodes are also situated within distance l from the same set of other nodes outside the group. Such strict assumptions are unlikely to be satisfied exactly in any real network, but exact accordance with both conditions is unnecessary for n.s.i measures to become reliable under node weighting. Several tests act as valuable indicators of the reliability of node-weighted network metrics calculated under the n.s.i paradigm, which are introduced here and explored further in Appendix A to support the robustness of our weighted LHC network metric calculations.

Our aim is to understand the impact of disparate weights on n.s.i metric calculations in a controlled environment to better interpret the results presented in Chapters 5-7. We test the power of node-weighted networks to accurately represent large event samples by comparing two samples: one given large event weights and the other given weights which are either the same or smaller by one or two orders of magnitude. We call these simulated datasets *Toy 1* and *Toy 2* respectively. The datasets should be sampled from the same distribution so that they represent the same model and region of phase space, and should also contain the same number of effective events. They are sampled from the same simple five-dimensional Gaussian distribution, where Toy 1 contains fewer events than Toy 2 such that $N_1 < N_2$. The weight on every Toy 1 event must therefore be larger, $w_1 > w_2$, to retain the same effective event number $N_{eff} = N_1 \times w_1 = N_2 \times w_2$. We consider three different approaches for generating these samples:

1. first generate Toy 2 and select its elements at random to create a subsample Toy 1,
2. generate Toy 1 and Toy 2 samples independently, or
3. first generate a small sample Toy 1 and design a larger second sample to have points nearby to those in the small set.

The first and third methods are more contrived than method two but allow for more controlled tests of several factors. For example, the third method may be used to place points in a larger Toy 2 sample precisely so that they will satisfy the connectivity conditions above, so that these two conditions may be gradually broken to measure their effects on the network metrics. Applying method one can instead allow us to consider the contrived situation where subsampled points are positioned exactly upon points in the larger sample. In this section, we consider the more generalised test using independent datasets as in method two.

The networks containing these samples from which we calculate the metrics can be created in two ways:

1. by combining both samples in the same network, or
2. by creating separate networks for the samples.

Adding samples with different weights to the same network is the more appropriate test for our LHC case, where we simulate backgrounds with significantly different weights to the signal and create SM+BSM networks in which signal and background events are interconnected. Therefore, we combine Toy 1 and Toy 2 in the same network to permit interconnections between the two data types with different weights. In addition, we vary the weight difference by iterating this test three times: one with equal weights $w_1 = w_2$, a second with $w_1 = 10 \cdot w_2$, and a third with $w_1 = 100 \cdot w_2$. In each case, the total number of effective events is $N_{eff} = 20{,}000$ and the relative numbers of nodes and their relative weights are contained in Table 4.1.

Table 4.1 Test combinations of samples Toy 1 and Toy 2 comprising numbers of events N_1 and N_2 respectively. Tests are designed to evaluate the impact of weight ratio w_1/w_2 on the network metrics.

	Test 1	Test 2	Test 3
N_1	10000	1000	200
N_2	10000	10000	20000
w_1/w_2	1	10	100

Figure 4.15 displays three examples of network metrics comparing the three tests, where Toy 1 and Toy 2 distributions are plotted independently. These examples show clearly identifiable shape properties in the network metric distributions. The distributions calculated for both Toy samples have consistent shape across all three tests for each of the three metrics. Some fluctuations are present in the tails of the distributions of the Toy 1 sample with only $N_1 = 200$, which can be attributed to statistical error due to low sampling. We note that the linking length l is different in each network because the three plots show examples using different distance calculations.

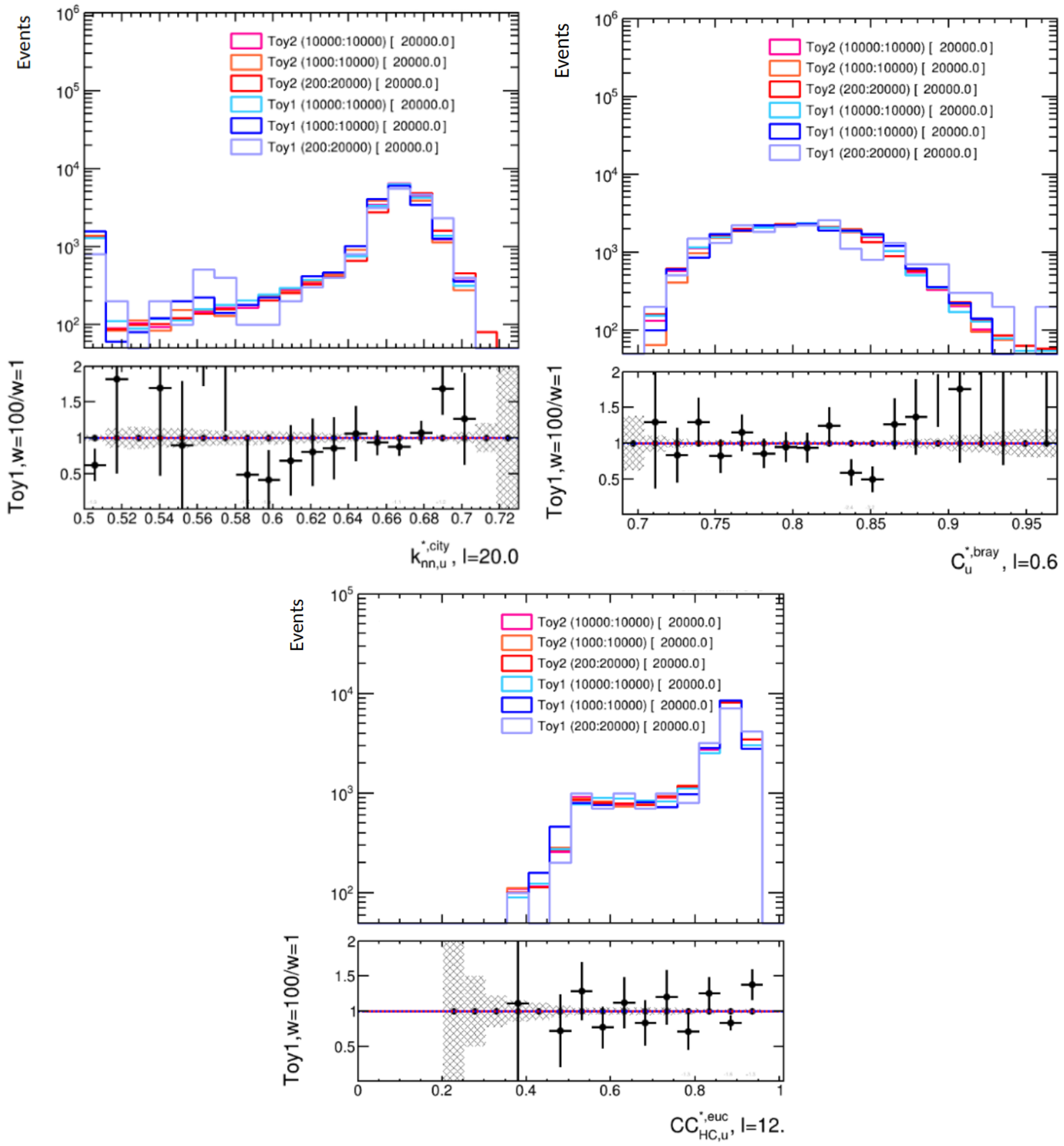


Fig. 4.15 Examples of metric histograms for the average neighbors degree with cityblock distance, the local clustering coefficient with braycurtis distance and the harmonic closeness centrality with Euclidean distance, calculated from three test networks in Table 4.1. The legend reads: $Sample(N_1 : N_2)[N_{eff}]$. Significance plots show the ratio between Toy 1 events per bin at weight 100 and at weight 1.

The consistent shapes of these metric disributions indicates some robustness of n.s.i network metric calculations under conditions where the maximum node weight ratio is 100 : 1. Further tests are addressed in Appendix A.

4.4 Significance calculations

We aim to increase exclusion sensitivities using network variables in combination with standard approaches to LHC analyses, and require valid means of comparing network techniques with other methods to determine the significance of a potential discovery.

The binomial $Z_{\rm bi}$ value we use to evaluate the significance of our analysis calculates the expectation for an absense of signal in a ratio of Poisson means, and is a binomial Z value based on one-sided Gaussian standard deviations. The calculation relies on a double number counting experiment, where one of the two expectations comes from a measurement that might contain signal and the second from an auxiliary measurement for a signal-free background [127–129]. The values were calculated using the BinomialExpZ function belonging to the number counting utilities in the RooStats package [130]. A full statistical analysis is beyond the scope of our proof-of-principle study, as we do not have access to a full background analysis, complete simulation of the detector systems, or the systematic uncertainties available to the experimental collaborations. The exclusion and discovery reach should be explored in a full evaluation in future ATLAS studies.

The exclusion sensitivities evaluated from the binomial significance are useful when compared with standard analyses. The binomial significance for the best signal yields is calculated in cases where selection cuts are placed on both network metrics and kinematic variables and compared with cases where selections are placed only on the kinematic variables. Chapter 5 presents the analysis and results for the study of supersymmetric electroweakino production, and Chapter 6 presents the analysis for stop pair production. In addition, Chapter 5 further evaluates network sensitivity by combining binomial significance results for standard selection cuts with significance derived from a standard BDT method. The result is a more comprehensive study of the sensitivity achieved with network analysis techniques.

Chapter 5

Case study 1: Electroweakino analysis

5.1 Introduction

A collider search for new physics is a novel context for graph network techniques, so the applications in this context must be evaluated using simulated datasets before approaching real LHC data. Estimates of the exclusion sensitivities with simulated events will indicate the future network analysis potential. We consider the analysis potential in two case studies searching for supersymmetry, using model-dependent examples to evaluate whether network techniques may extend to broader BSM searches. The electroweakino example is the focus of this chapter.

If the masses of gluinos and squarks are out of reach at the LHC, the direct production of neutralinos and chargions may dominate sparticle production. The electroweakinos are the linear combinations of the superpartners of unbroken electroweak gauge bosons and supersymmetric Higgs particles, appearing as the mass eigenstates, the neutralinos ($\tilde{\chi}^0_{i=1,2,3,4}$) and charginos ($\tilde{\chi}^\pm_{i=1,2}$). The composition of these eigenstates by the mixing of the *bino*, *wino*, and *Higgsino* states implies the resulting phenomenology, so LHC analyses must assume the relative masses and mixings to design targeted searches. We assume that electroweakinos are produced via wino-like $\tilde{\chi}^\pm_1 - \tilde{\chi}^0_2$ production, and decay into SM gauge bosons and the lightest neutralinos in a model where these neutralinos are pure bino states. The model contains masses $m_{\tilde{\chi}^\pm_1} = m_{\tilde{\chi}^0_2} = 400$ GeV, and $m_{\tilde{\chi}^0_1} = 0$ GeV. The leptonic decays of both gauge bosons produces a three-lepton final state, which has been targeted in recent searches, and is shown in Figure 5.1. The model was not excluded in the three-lepton search channel in 139 fb^{-1} analyses conducted by the ATLAS collaboration until very recently [131]. Our network analysis assumes a comparable 150 fb^{-1} of integrated luminosity.

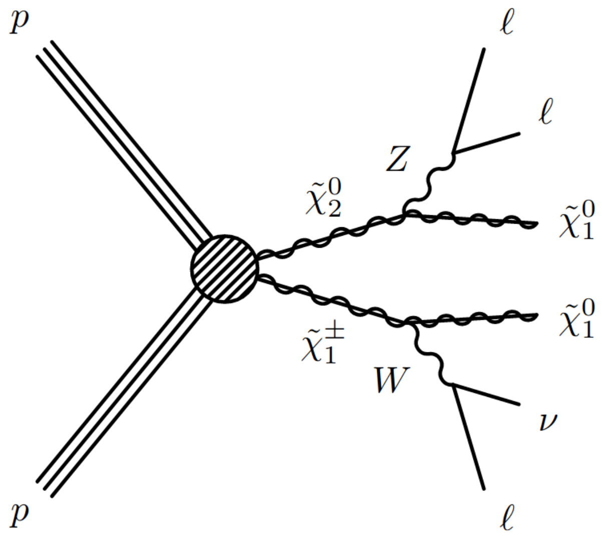


Fig. 5.1 The Feynman diagram of the simplified supersymmetric model considered in the prototype electroweakino analysis.

The background for the three-lepton electroweakino signal is by far dominated by the diboson WZ background, and both signal and background are simulated with `Pythia 8` [132]. We use the `CTEQ6L1` PDF set [133], with an LHC detector simulation performed using `Delphes3` [72, 134, 135] supplied with the default ATLAS detector card. Jets are reconstructed using the anti-k$_\mathrm{T}$ algorithm with a radius parameter $R = 0.4$ [136] using the FastJet package [137]. The electroweakino signal is normalised to the relevant integrated luminosity using the next-to-leading-order plus next-to-leading-log cross-section provided in Refs. [138, 139], and the WZ sample uses the next-to-next-to-leading order cross-section presented in Ref. [140].

5.2 Electroweakino analysis design

Preselection criteria are used to narrow the search to a typical three-lepton signal region, where we require exactly three light leptons (electrons or muons). The preselection requires transverse momentum $p_\mathrm{T} > 25$ GeV and pseudorapidity $|\eta| < 2.5$. The default light lepton reconstruction is provided by `Delphes3`, with no additional pseudorapidity or isolation requirements. The default settings restrict electrons, photons and muons to $|\eta| < 2.5$, and restrict jets and missing transverse energy to $|\eta| < 4.9$. There must be no b-tagged jets in the event, and at most 1 non-b-tagged jet with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.4$. Finally, the event must contain the dilepton invariant mass of an opposite-sign same-flavour pair satisfying $|m_{ll} - m_Z| < 10$ GeV.

5.2.1 Kinematic basis for network construction

The basis of variables underlying the network is selected from those which typically discriminate well between diboson events and electroweakino events, showing the greatest difference between the WZ background and our benchmark signal point. The variables are defined in Section 3.3.1 and listed in brief as follows:

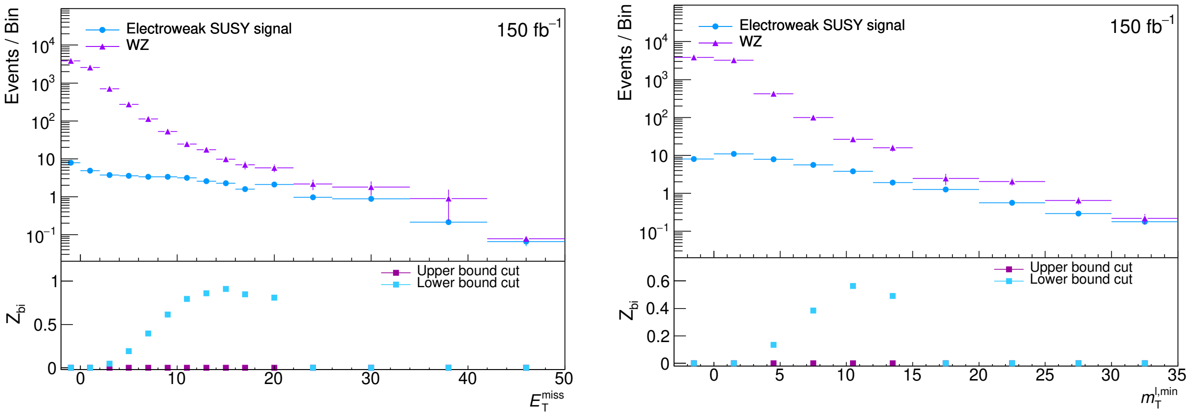
- $E_{\mathrm{T}}^{\mathrm{miss}}$: the missing transverse energy,
- $m_{\mathrm{T}}^{l,\mathrm{min}}$: the transverse mass minimised by selecting one lepton,
- $p_{\mathrm{T}}(Z)$: the reconstructed transverse momentum of the Z boson,
- $\Delta\Phi(l_Z^+, l_Z^-)$: the azimuthal angle between the two leptons associated with the Z boson
- $\Delta\Phi(Z, l_W)$: the azimuthal angle between the reconstructed Z boson and the lepton coming from the W boson.

We avoid large statistical uncertainty in the rare tail regions of the kinematic distributions by sampling highly in these regions using the slicing process described in Section 4.2.3. We apply this procedure to the diboson background event samples only, unlike the simulations in Chapter 6, where all samples are generated in slices including both stop quark signal and top quark background. Once the slices are sampled, the preselection criteria select a set of events that amounts to just over 10,000 signal and 10,000 background events before weighting to construct the networks. The distributions of the kinematic variables for the electroweakino signal and WZ background are shown in Figure 5.2, where the WZ events were generated in slices to reduce tail fluctuations and uncertainties.

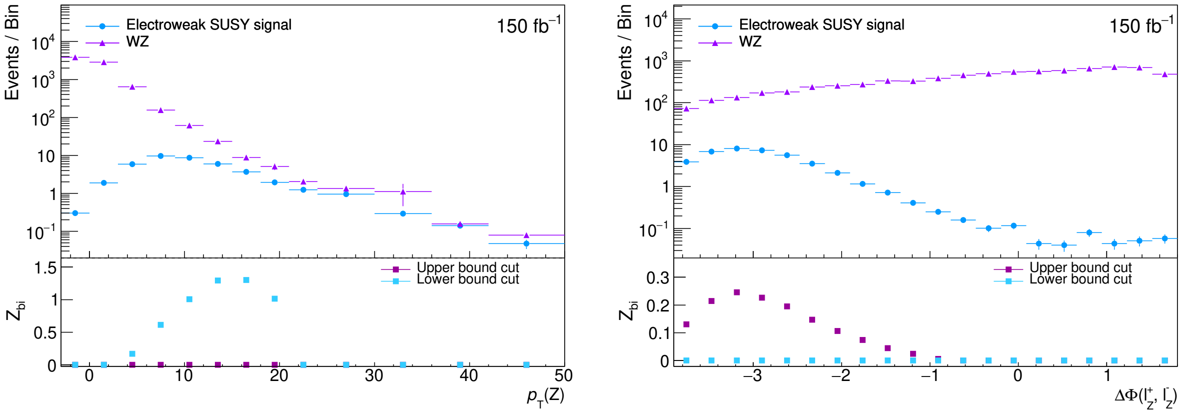
A possible problem with constructing a network from a basis of variables which have values on different scales is that those with higher average values will contribute more to the analysis results calculated later from the network. Even if significant exclusion sensitivity is obtained from variables which have typically smaller scales, the relative difference in their values is dwarfed by that of higher-scale variables. Re-scaling each variable by its median value reduces this concern. Additionally, a median scaling is more effective than a mean scaling because it avoids high sensitivity to uncertain tail effects. However, any technique must carefully consider the dimensions and bounds of the analysis variables, ensuring that a sample with values symmetric around zero are not eliminated. For example, a typical angular variable in collider analyses may have a chosen range of $(-\pi, \pi)$, and may contribute analysis sensitivity equal to a variable with bounds $(0, 2\pi)$. Our re-scaling method was designed with this consideration in mind.

The re-scaling method contains the following steps. First, we determine the scale of each variable by calculating its median value for the background simulated events. We then re-scale all entries in both signal and background by subtracting the median of the background sample and normalising by the median absolute deviation (MAD). The MAD is calculated by first subtracting the median of the background dataset from each point to produce a new dataset, then calculating the median of the new dataset. Consequently, the basis variables convey the same information into our networks, but care has been taken to weight them evenly to avoid mistakenly eliminating some valuable discrimination power from the later analysis. The variables plotted in Figure 5.2 are normalised by the median scaling technique.

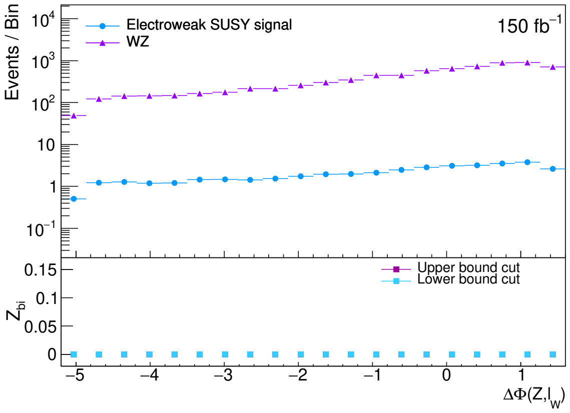
The lower panel of each figure shows the binomial significance Z_{bi} for either an upper cut or lower cut on the variable at the value given on the horizontal axis, assuming a total systematic uncertainty of 15% (chosen to be consistent with numbers quoted by the ATLAS Run-2 searches in the 3-lepton channel [131, 141]). This significance calculation also includes the statistical uncertainty of the background which is added to the systematic uncertainty in quadrature. When the signal or background weighted event yield drops below 3 events, the Z_{bi} is set to 0. Differences in shape between the signal and SM distributions are clearly apparent, but no selection on a single kinematic variable is able to achieve an expected exclusion sensitivity at 95% confidence level (corresponding to $Z_{\text{bi}} = 1.64$) for our chosen benchmark point.



(a) E_T^{miss} (b) $m_T^{l,\text{min}}$



(c) $p_T(Z)$ (d) $\Delta\Phi(l_Z^+, l_Z^-)$



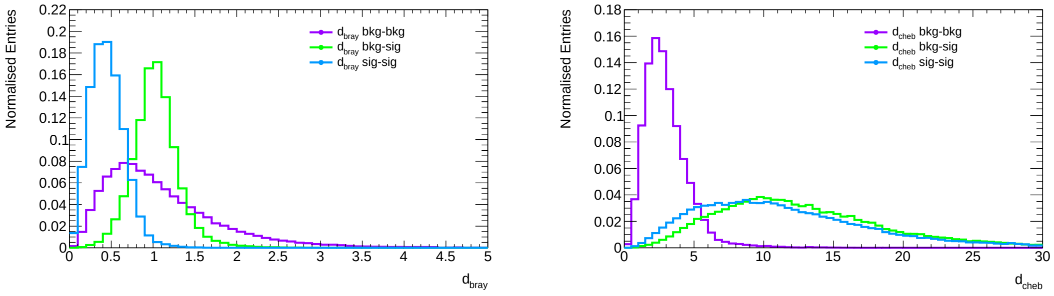
(e) $\Delta\Phi(Z, l_W)$

Fig. 5.2 Event rates as a function of the kinematic variables used in our prototype electroweakino search, with the median scaling applied. Events in the overflow bin are not shown in the distribution but are included in the Z_{bi} calculation.

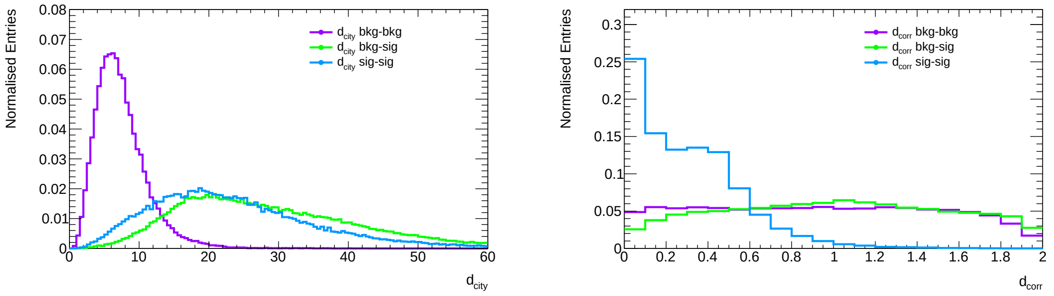
Parallelising the network metric calculations will in future allow us to analyse networks containing $\mathcal{O}(100,000)$ events before weighting, which is a useful step we plan to pursue after our present proof-of-principle calculations. The metrics are computationally expensive to evaluate, not only because the adjacency matrix must always be repeatedly accessed to enumerate the edges between each node and every other node in the network, but also because some metrics comprise intricate methods of calculation. For example, the betweenness centrality metric requires a computationally intensive calculation of every shortest path in the network. After parallelising, which is made available through the `pyunicorn` package, we are confident that the network variables can be calculated using a looser preselection within the ATLAS and CMS collaborations. Looser preselection requirements are beneficial for several reasons; for example, tighter kinematic requirements in the preselection increases the similarity between signal and background and therefore weakens the analysis sensitivity. Additionally, looser criteria permit the analysis to cover a broader parameter space and search for less specific signals.

5.2.2 Distance calculations

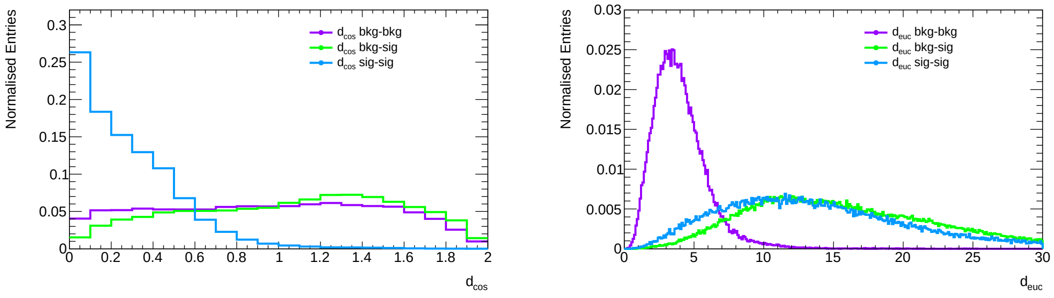
We create networks from the re-scaled data by calculating the distance metrics defined in Section 4.2.2, letting events become nodes and defining edges between nodes according to the prescription for choosing a linking length discussed in the previous chapter. Figure 5.3 shows distributions of these distances for signal and background events normalised to unit area. Shape properties of the distributions differ depending on the type of events chosen; for example, distances between two signal events may operate on a scale that is different to background-background and signal-background distances. Therefore, we distinguish between event types in the distance distributions, obtaining further information regarding the optimal setup of our networks.



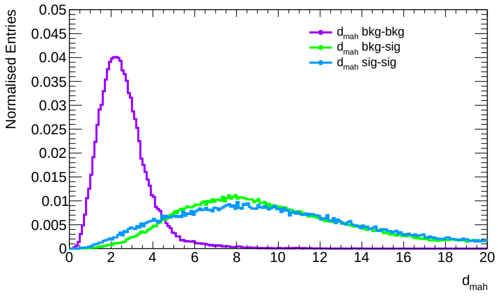
(a) Bray-Curtis (b) Chebyshev



(c) cityblock (d) correlation



(e) cosine (f) Euclidean



(g) Mahalanobis

Fig. 5.3 Distributions of the distance metrics used in our prototype electroweakino search.

The signal-signal and background-background distributions in these distances are somewhat different, indicating different typical scales at which events are clustered or otherwise distributed in collider space. In most cases, the plots reveal that background events are more likely to be found clustered close to other background events, while signal events are less likely to be close to any other events, including other signal. This is true for Chebyshev, cityblock, Euclidean and Mahalanobis distances. In the other three cases, however, the signal-signal and background-background roles are reversed; the Bray-Curtis, correlation and cosine metrics all have more concentrated signal-signal distributions at low distances, while background events are typically "further away" from other events. The correlation and cosine metrics in particular show this distinct shape property; these metrics measure angular separation, indicating that signal events may have more closely aligned directionality in the chosen space. In this case, the characteristic angular distances of the signal model may contribute to distinguishing them from background in an angular distance network. We conclude that a combination of several networks created from different distance metrics which contribute non-trivial additional information may be useful.

The difference between these signal and background shapes in the distance distributions suggests that the condition for adding an edge could be flipped so events are connected only when their distance in collider space is greater than the linking length, instead of smaller. Some network metric distributions are later shown for networks with flipped conditions, and we compare these with the metrics calculated from standard networks. Although the main results of this analysis proceed from the standard definitions for consistency, the flipped conditions are an example of an additional type of network parameter which may contribute greater power in distinguishing signal from background for some BSM models.

The linking length relates the characteristic scale of separation of background events based on Figure 5.3 and informs which entries are 1 or 0 in the adjacency matrix. For each distance metric d_{ij} between pairs of events i and j, we define the adjacency matrix for the graph network by using the definition in Equation 4.4. For distances where the background is closer in collider space to other background events than signal, our choice of linking length creates networks which contain background more densely clustered than signal. In others, for example correlation and cosine distance networks, the signal is more densely connected. Our final choice of linking length for every distance metric is summarised in Table 5.1.

The networks then contain complex structures on the chosen scale, with properties we can calculate from the event-by-event values provided by the network metrics defined in Section 4.3. Our analysis with these metrics proceeds in two stages, designed to reflect how it could be performed in practice at the LHC:

Table 5.1 Linking length values used for each distance metric for our prototype electroweakino analysis.

Distance metric	**Linking length**
d_{bray}	0.7
d_{cheb}	4.8
d_{city}	12
d_{corr}	0.6
d_{cos}	0.6
d_{euc}	6.4
d_{mah}	4.8

- In this section (5.2), we design signal regions that would be sensitive to our chosen electroweakino benchmark point. Since we know which events are background and which are signal in our simulated networks, we can design optimal signal regions from the network metrics calculated directly from signal-plus-background networks. The same process of using MC samples to design signal regions would be performed in the ATLAS and CMS collaborations. After choosing optimal regions, we calculate the hypothetical significance in each region by counting the number of signal and background events, assuming that the background contribution to the signal-plus-background network metrics is similar to their metric distributions calculated from background-only networks. If this is true, then the significance calculation is valid, as described in Section 4.1.3. Alternatively, meaningful significance results can be calculated by optimising the selections on network metrics calculated from a signal-plus-background network compared with a background-only network, using a separate background sample for the latter. However, we do not pursue this second option in this section, because since we have determined that all variables in our optimal analyses below satisfy the earlier assumption on the background shape.

- In Section 5.3, the second option above is pursued as a realistic example of an LHC exclusion test, where an independent set of background events becomes a background-only network. These events represent the *null hypothesis*, or mock LHC data that would be measured in absence of a signal. We then compare the yields from mock LHC data network metrics with yields from signal-plus-background network analysis in our search region to determine the exclusion significance of our benchmark model. Therefore, the exclusion significance now incorporates the effect of statistical fluctuations in real LHC data. This test could be repeated on a variety of signal models, to generate exclusion limits in, for example, simplified model parameter planes.

Numerous possible event selections can contribute to a network analysis, including selections on: the network metrics only, the original kinematic variables only, or a combination of both. The large possible number of combinations of distance metric and network metric we have chosen for the electroweakino study provides $7 \times 8 = 56$ new variables, excluding the option of flipped linking criteria. Adding this latter option doubles the total number of possibilities. All network selections may be combined with others on either different network variables or the original kinematic variables to improve exclusion sensitivity.

5.2.3 Signal region selection

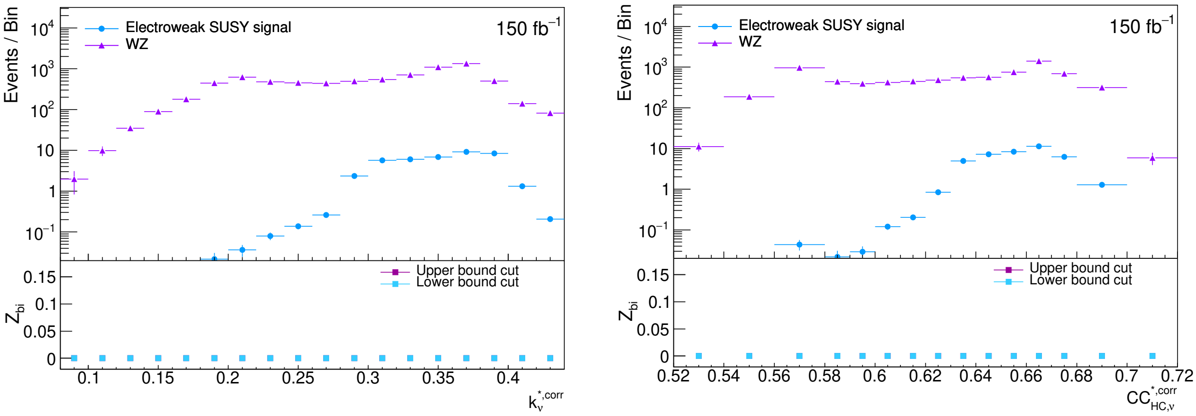
Contrasting with the distributions of the kinematic variables in Figure 5.2, the distributions for the network metrics in the signal-plus-background network are revealed in Figures 5.5-5.4. We have selected several useful network metrics for the Euclidean, Chebyshev, cityblock, Bray-Curtis and cosine distance metrics. The supersymmetric events consistently show lower values of k_v^*, as well as the clustering coefficient and closeness variables. This suggests that the supersymmetric events form fewer connections with other events than the SM events, and that the events they do connect with are also sparsely connected. This arises from a combination of factors including the different typical lepton four-momenta expected in the supersymmetry case (given that the final state has a higher multiplicity), different signal distributions for variables which have well-defined kinematic endpoints for the background, and the smaller number of supersymmetry events limiting the number of possible signal connections. These network metrics may also be useful for other BSM cases, because the rare and complex nature of the signal events for our benchmark model is a common feature among BSM model processes. More complex supersymmetry models favoured by recent global fits significantly depart from simplified model assumptions, and possess especially rare and complex signals.

The local network metric distributions provide greater discrimination between signal and background than can be obtained with the original kinematic variables alone. Optimal kinematic selections for potential analyses come from placing a single upper or lower cut on each variable in turn, and determining which variables lend greatest significance. The cut value may take one of a range of possible values we considered. Iteratively, cuts are placed on every variable and retained if they improve significance until the number of signal and background events passing the combined selections decreases to less than three each. Several examples of promising selections and their Z_{bi} values and event yields are shown in Table 5.2

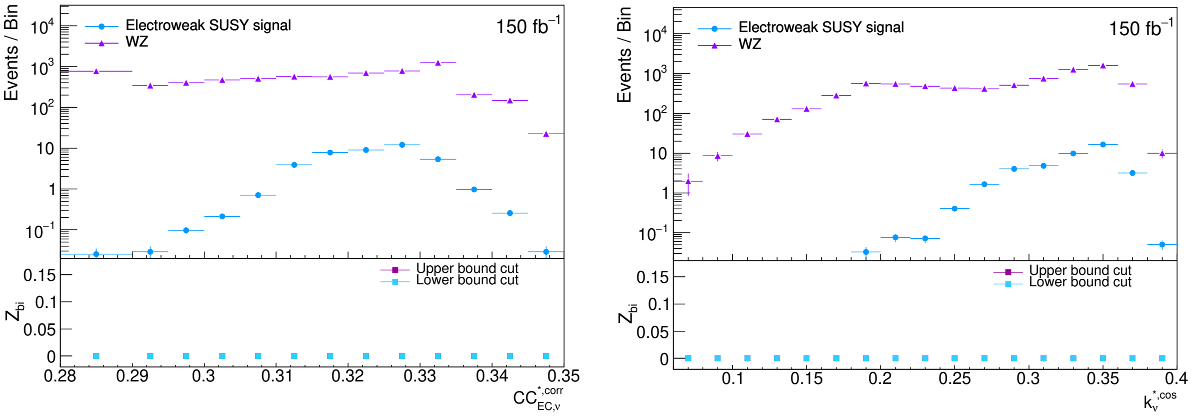
for $Z_{\text{bi}} > 1.64$. Binomial significance is calculated from both the statistical uncertainty on the event yields and an assumed systematic uncertainty of 15%.

Table 5.2 Examples of binomial significances for search regions that provide exclusion sensitivity without reducing yields below 3 events. The errors only include the statistical error.

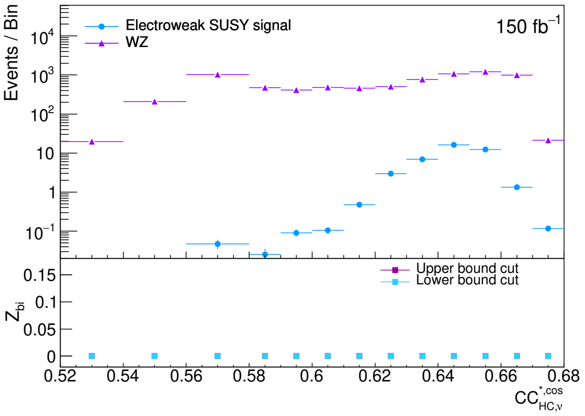
Requirement	N_{signal}	$N_{\text{background}}$	Z_{bi}
$k_v^{*,\text{euc}} < 0.003, CC_{EC,v}^{*,\text{corr}} > 0.324$	7.78 ± 0.17	6.53 ± 1.29	1.96
$k_v^{*,\text{euc}} < 0.003, CC_{HC,v}^{*,\text{corr}} > 0.656$	8.45 ± 0.18	7.52 ± 1.45	1.96
$k_v^{*,\text{city}} < 0.004, k_v^{*,\text{cos}} > 0.343$	5.86 ± 0.15	3.20 ± 1.05	1.89
$k_v^{*,\text{city}} < 0.004, CC_{HC,v}^{*,\text{cos}} > 0.646$	7.43 ± 0.16	5.67 ± 1.42	1.89
$k_v^{*,\text{city}} < 0.004, k_v^{*,\text{corr}} > 0.359$	7.34 ± 0.16	5.43 ± 1.10	2.04
$k_v^{*,\text{city}} < 0.004, CC_{EC,v}^{*,\text{corr}} > 0.323$	8.13 ± 0.17	7.15 ± 1.44	1.93
$k_v^{*,\text{city}} < 0.004, CC_{HC,v}^{*,\text{corr}} > 0.656$	8.17 ± 0.17	7.27 ± 1.45	1.92



(a) $k_v^{*,\mathrm{corr}}$ (b) $CC_{HC,v}^{*,\mathrm{corr}}$

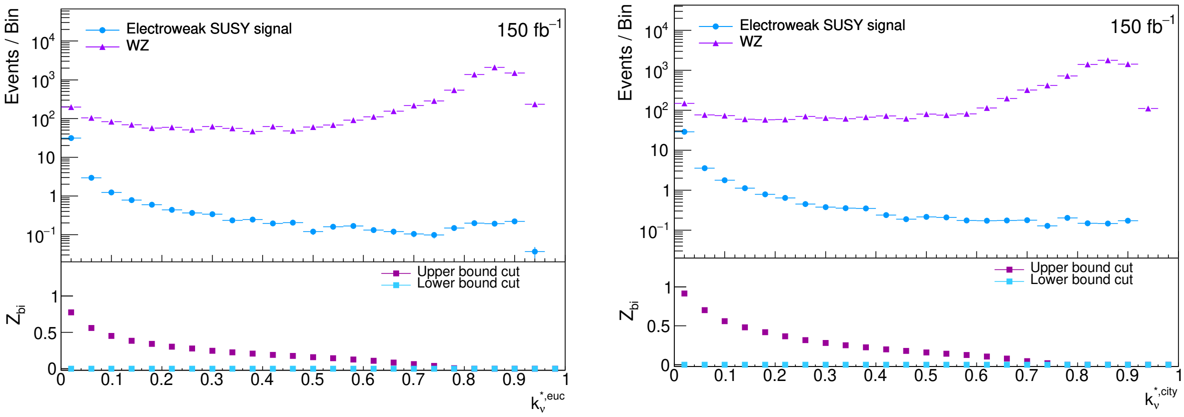


(c) $CC_{EC,v}^{*,\mathrm{corr}}$ (d) $k_v^{*,\mathrm{cos}}$



(e) $CC_{HC,v}^{*,\mathrm{cos}}$

Fig. 5.4 Event rates as a function of useful network metrics for our prototype electroweakino analysis calculated using d_{corr} and d_{cos}. Events in the overflow bin are not shown in the distribution but are included in the Z_{bi} calculation.

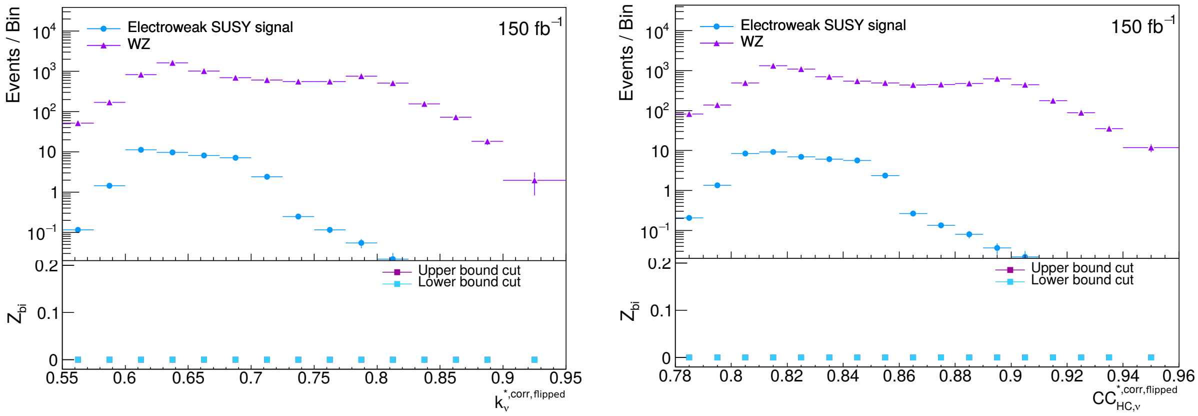


(a) $k_v^{*,\mathrm{euc}}$ (b) $k_v^{*,\mathrm{city}}$

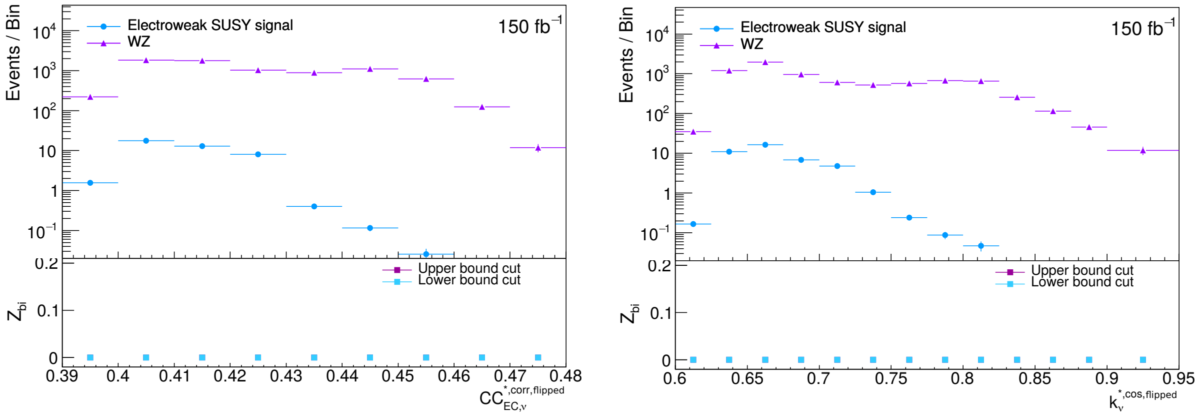
Fig. 5.5 Event rates as a function of useful network metrics for our prototype electroweakino analysis calculated using d_{city}. Events in the overflow bin are not shown in the distribution but are included in the Z_{bi} calculation.

The network metrics possess different shape properties for the background and signal for this benchmark model, when calculated from a combined signal-plus-background network. For a more complete view of the distribution shapes, we checked the shapes of the network distributions with those calculated using the flipped criterion, connecting nodes when they are a greater distance apart than the linking lengths in Table 5.1. The flipped metric distributions are displayed in Figure 5.6 for the same set of network metrics chosen in Figure 5.4 with correlation and cosine distances. Both angular distances showed signal grouped closer than the background in collider space, which was spread “further apart” by these measures, so the flipped correlation and flipped cosine networks are designed to increase connectivity in the background events compared with the signal. Flipping the connectivity for networks with these distances produces similar signal-background separation to the standard Euclidean and cityblock network definitions with the degree centralities shown in Figure 5.5.

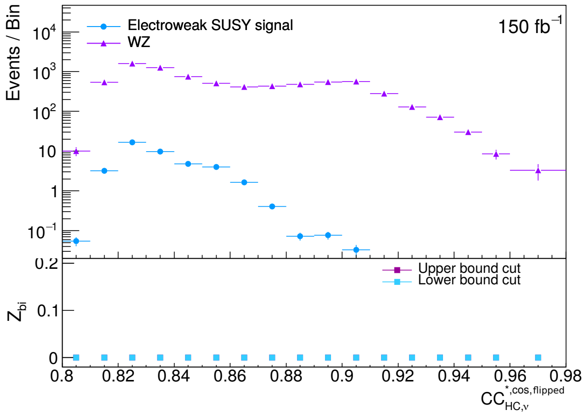
The flipped metrics are not used in the prototype analysis to calculate signal yields, but are an interesting method of verifying the shape properties of the network metrics, in particular for the correlation and cosine distances with reversed signal and background connectivity. Flipping the network connectivity definition in these cases reveals similar distributions to the metrics from the standard constructions, but exhibits evidence of signal events with lower average centralities.



(a) $k_v^{*,\mathrm{corr}}$ (b) $CC_{HC,v}^{*,\mathrm{corr}}$



(c) $CC_{EC,v}^{*,\mathrm{corr}}$ (d) $k_v^{*,\mathrm{cos}}$



(e) $CC_{HC,v}^{*,\mathrm{cos}}$

Fig. 5.6 Event rates as a function of flipped network metrics calculated using d_{corr} and d_{cos}. Events in the overflow bin are not shown in the distribution but are included in the Z_{bi} calculation.

5.3 Results of realistic electroweakino exclusion test

The next stage of our electroweakino analysis is designed to reveal any differences between the network metrics for a signal-plus-background network and a background-only network. We do this by comparing the yields in the signal regions derived in the previous section with those found when there is no longer any signal in the networks. To represent collider data without signal, we use the mock LHC background-only dataset. The setup produces a realistic example of an LHC exclusion test, allowing us to calculate the exclusion sensitivity we would obtain by using real LHC data for this benchmark model. The sensitivity calculation is again performed using binomial significance, calculated with an error on the background yield that includes both the statistical uncertainty and an additional 15% systematic uncertainty.

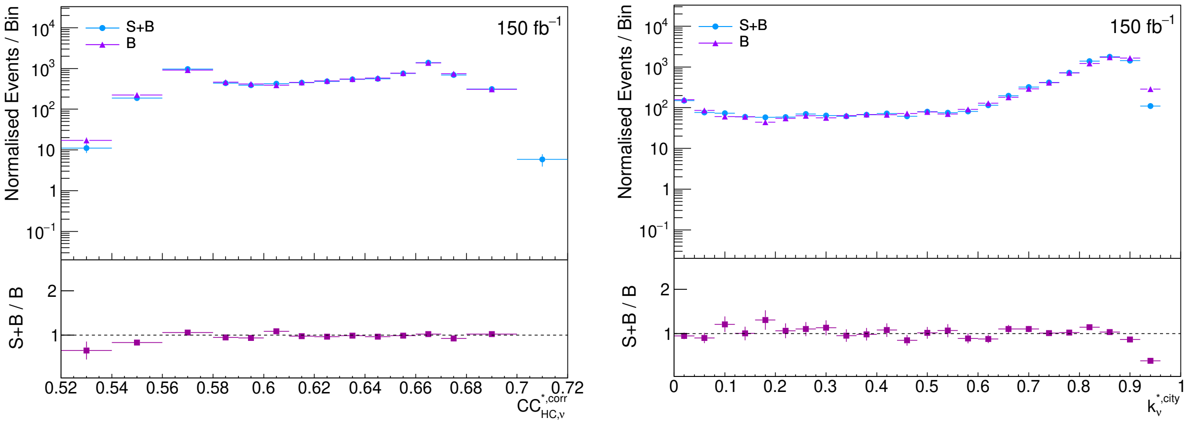
We compare yields in search regions defined with network metrics only with regions constructed using only conventional kinematic variables, representing standard cut-and-count analysis regions. The kinematic-only selections are loosely inspired by the regions in the ATLAS 36.1 fb^{-1} search [142], but generally have tighter requirements on $E_{\mathrm{T}}^{\mathrm{miss}}$ and $M_T^{l,\mathrm{min}}$. Additionally, our preselection requirements include $n_{jets} < 2$, unlike the ATLAS search region which was inclusive in light jet multiplicity, but had “binned” regions considering the 0 and > 0 light jet categories separately. Tightening these requirements is designed to increase signal yield in kinematic regions for a more competitive comparison with network methods.

Table 5.3 compares the expected sensitivity of these regions with that obtained in the standard kinematic search regions. The Z_{bi} values differ from those of the previous section, due to statistical fluctuations in the mock LHC dataset. Network methods provide exclusion sensitivity for this benchmark supersymmetry model, outperforming the analysis based exclusively on the original kinematic variables. The two types of analysis may be performed in combination; results revealed that a further kinematic selection $p_{\mathrm{T}}(Z) > 160$ GeV provides additional signal-background discrimination that enhances the role of the network metrics. However, if this cut is instead applied in the preselection, the same search regions defined by the network metrics produces lower yield, suggesting that the preselection affects the network topology non-trivially. In this case, the operations of selecting data and constructing a network do not commutatively produce the same metric distributions. The measurable impact of preselection criteria on network topology and subsequently on signal yields in network regions is an avenue for further investigation intended to extend our ability to accurately predict when network metrics will provide useful power of discrimination, re-visited in Chapter 8.

When comparing yields with background-only datasets, we must compare background distributions for network metrics calculated from the background-only networks with those calculated from signal-plus-background metrics. Verifying a consistent background shape indicates that the addition of rare signal events does not substantially change the connection properties of background events. If this is true, then the procedure we designed to define search regions using only signal-plus-background networks is realistic. Figure 5.7 superimposes background distributions taken from the two networks, and reveals that the general shape properties are the same. We consider this sufficient evidence to validate our method of defining search regions for a mock LHC study.

Table 5.3 Yields in our electroweakino search regions for our mock LHC data set ($N_{\text{b-only}}$) and our mock MC set ($N_{\text{s+b}}$). Also shown is the sensitivity of search regions using only the original kinematic variables. The errors quoted are statistical only, while the Z_{bi} calculation uses a relative background uncertainty.

Requirement(s)	$N_{\text{b-only}}$	$N_{\text{s+b}}$	Z_{bi}
$k_v^{*,\text{euc}} < 0.003, CC_{EC,v}^{*,\text{corr}} > 0.324$	7.27 ± 1.17	14.31 ± 1.3	1.76
$k_v^{*,\text{euc}} < 0.003, CC_{HC,v}^{*,\text{corr}} > 0.656$	8.43 ± 1.36	15.97 ± 1.46	1.73
$k_v^{*,\text{city}} < 0.004, k_v^{*,\text{cos}} > 0.343$	3.48 ± 0.85	9.06 ± 1.06	1.90
$k_v^{*,\text{city}} < 0.004, CC_{HC,v}^{*,\text{cos}} > 0.646$	6.21 ± 1.29	13.1 ± 1.42	1.77
$k_v^{*,\text{city}} < 0.004, k_v^{*,\text{corr}} > 0.359$	6.08 ± 1.13	12.77 ± 1.11	1.79
$k_v^{*,\text{city}} < 0.004, CC_{EC,v}^{*,\text{corr}} > 0.323$	7.77 ± 1.46	15.29 ± 1.45	1.74
$k_v^{*,\text{city}} < 0.004, CC_{HC,v}^{*,\text{corr}} > 0.656$	7.17 ± 1.31	15.45 ± 1.46	2.01
$k_{nn,v}^{*,\text{city}} < 0.009, CC_{HC,v}^{*,\text{corr}} > 0.659$	4.4 ± 0.87	9.5 ± 0.85	1.63
$p_{\text{T}}(Z) > 160\,\text{GeV}$, $E_{\text{T}}^{\text{miss}} > 100\,\text{GeV}$, $m_{\text{T}}^{l,\text{min}} > 150\,\text{GeV}$	93.93 ± 7.66	105.28 ± 7.55	0.47
$p_{\text{T}}(Z) > 160\,\text{GeV}$, $E_{\text{T}}^{\text{miss}} > 200\,\text{GeV}$, $m_{\text{T}}^{l,\text{min}} > 150\,\text{GeV}$	28.55 ± 3.83	34.99 ± 3.62	0.64
$p_{\text{T}}(Z) > 160\,\text{GeV}$, $E_{\text{T}}^{\text{miss}} > 300\,\text{GeV}$, $m_{\text{T}}^{l,\text{min}} > 150\,\text{GeV}$	7.05 ± 1.59	13.41 ± 1.78	1.48



(a) $CC^{*,\mathrm{corr}}_{HC,v}$ (b) $k^{*,\mathrm{city}}_{v}$

Fig. 5.7 Event rates as functions of $C^{*,\mathrm{euc}}_{s,v}$ and $k^{*,\mathrm{city}}_{v}$ for WZ background events calculated from either the signal-plus-background or background-only network. These two networks use different sets of simulated WZ events. Events in the overflow bin are not shown in the distribution.

5.4 Discussion

The network variables reveal exclusion potential for an electroweakino example using a benchmark point that has recently been excluded by a three lepton ATLAS search performed using 139 fb^{-1} of data. Comparing network techniques with prototype cut-and-count analyses, the network analysis performs well, due to the differences in shape between signal and background in the network variables. The analysis includes the degree centrality metric calculated for cityblock, Euclidean, cosine and correlation distance networks, and closeness centrality metrics calculated for correlation and cosine distance networks. The closeness metrics include both exponential and harmonic closenesses. Several other network metrics may provide additional power, but were excluded from the proof-of-principle analysis. Many network metrics are not highly correlated, designed to calculate different characteristics of network topology, and may therefore combine to increase exclusion potential.

A more comprehensive comparison of network variables with standard search techniques comes from comparing their sensitivity with that of a Boosted Decision Tree (BDT). In this section, results are compared first with a BDT trained on the original kinematic variables with the same preselection as the networks, and then with a BDT trained on the original variables (with preselection) plus the addition of one promising network metric.

Training the BDTs was performed using the ROOT package TMVA [143] with its default settings. We trained both BDTs on half the electroweakino signal and WZ background samples, randomly selected, so that they could be tested on the other half, retaining good sample size. The BDT output distribution provides binomial significance Z_{bi}, calculated using the same method as before, which is evaluated for upper and lower cuts on the BDT output variable shown alongside the full sample BDT distribution in Figure 5.8. The maximum Z_{bi} is 3.63, which is slightly higher than the network variable cuts proposed in Table 5.2.

Whereas a BDT performs admirably when required to exclude the same signal that it was trained on, its performance is expected to decrease when generalising to other signal points which were not part of its training set. An advantage of the graph network approach is that its performance generalises across diverse models and search regions because it does not rely on model-dependent training.

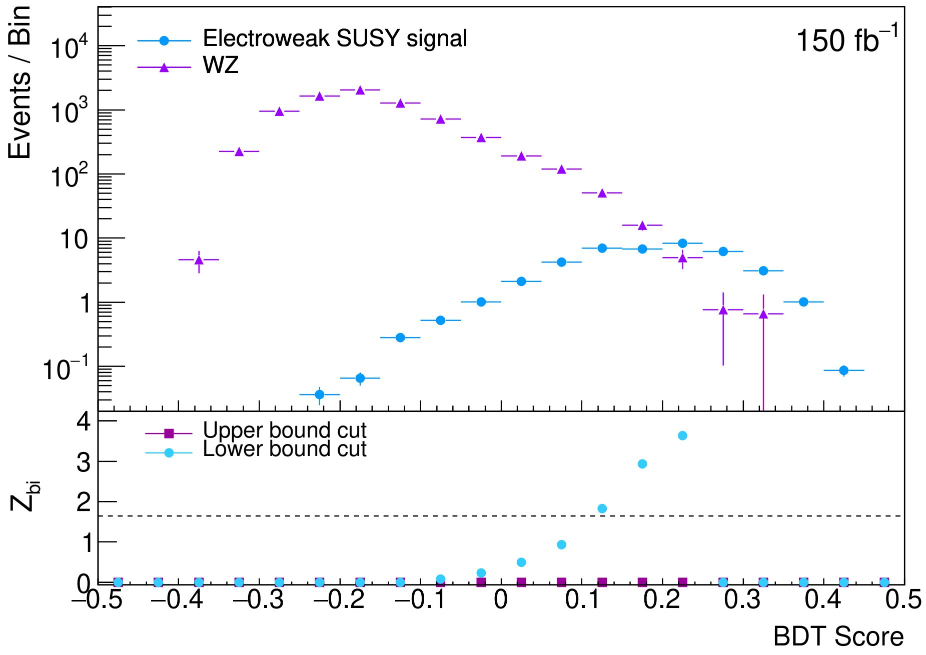


Fig. 5.8 Distributions of the BDT score for electroweakino signal and WZ background events, in the case where the BDT is only trained on kinematic variables. In the lower panel the Z_{bi} is shown for cumulative upper or lower bound cuts on the score. Events in the overflow bin are not shown in the distribution but are included in the Z_{bi} calculation.

The second BDT we consider was created from the same setup as the first, but with the addition of one network variable: $k_V^{*,\text{euc}}$. The distribution of the second BDT output variable is plotted in Figure 5.9, where a clear improvement in Z_{bi} is observed, reaching up to 3.98.

Adding only one network metric that revealed promising behaviour in the initial analysis noticeably increases BDT performance.

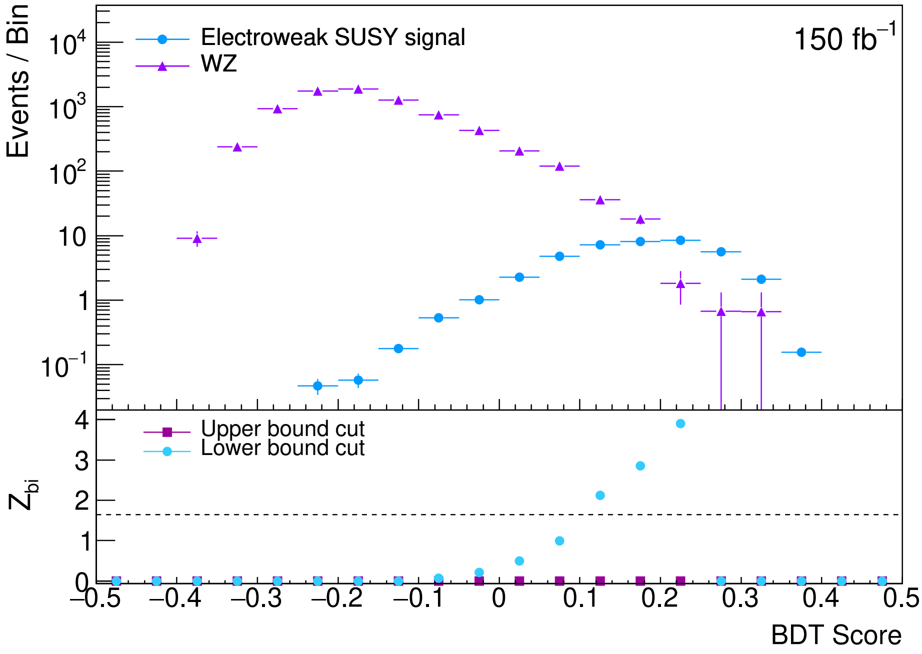


Fig. 5.9 Distributions of the BDT score for electroweakino signal and WZ background events after $k_v^{*,\mathrm{euc}}$ is added to the BDT. In the lower panel the Z_{bi} is shown for cumulative upper or lower bound cuts on the score. Events in the overflow bin are not shown in the distribution but are included in the Z_{bi} calculation.

The promising performance of several network metrics in the electroweakino case study motivates further research. We progress towards checking network applications to real ATLAS data measured at the LHC, but precede this investigation with a second test on simulated events from a different region of phase space. Chapter 6 presents a supersymmetric top quark case study, for comparison with network construction and metric calculations for a different model and search region. Following these results, Chapter 7 discusses our application to real ATLAS data, checking the shapes of the metric distributions against simulated background expectations.

Chapter 6

Case study 2: Stop quark analysis

6.1 Introduction

Our second demonstration of network analysis applied to an LHC search is a search for supersymmetric top quarks. One purpose of this second example is to determine whether the apparent gains in sensitivity seen in Chapter 5 are common to other final states and signal topologies. We choose the stop production example as a comparison with the electroweakino study to evaluate the differences which may arise when more challenging signal models are employed. Comparing with a different signal model also reveals whether the same network metrics that produce useful results in the electroweakino case are widely useful, or if new combinations of distance and network metrics are more valuable for distinguishing events distributed in a different collider space.

The squarks are produced in strong interactions with significantly larger production cross-sections than non-coloured sparticles of equal mass (such as neutralinos and charginos), so they are a primary focus of hadron collider search efforts. However, the benchmark squark production model we choose is also motivated by its challenging properties; the kinematic properties of the top quarks produced are highly similar to those expected in the SM. Consequently, this chapter contributes a more thorough test of LHC network analysis. The model assumes that $\tilde{t}_1\tilde{t}_1$ production dominates at the LHC, and the decay of the stop is presumed to occur with 100% branching ratio to either a top quark and a lightest neutralino, $\tilde{\chi}_1^0$, or a b quark and a lightest chargino. We consider a model with $m_{\tilde{t}_1} = 500$ GeV, and $m_{\tilde{\chi}_1^0} = 280$ GeV. The decay topology we analyse corresponds to a one lepton final state, identified in Figure 6.1. Since this work began, the ATLAS Collaboration released a zero lepton stop pair production search utilising 139 fb^{-1} of data which excludes our chosen benchmark model [144]. However, the results remain valid to show that, for a given set of kinematic variables in the 1 lepton case, network measures can improve standard analysis performance. It remains to be seen how useful this would be in the 0 lepton case, and whether our 1 lepton network-based analysis would outperform the 0 lepton analysis.

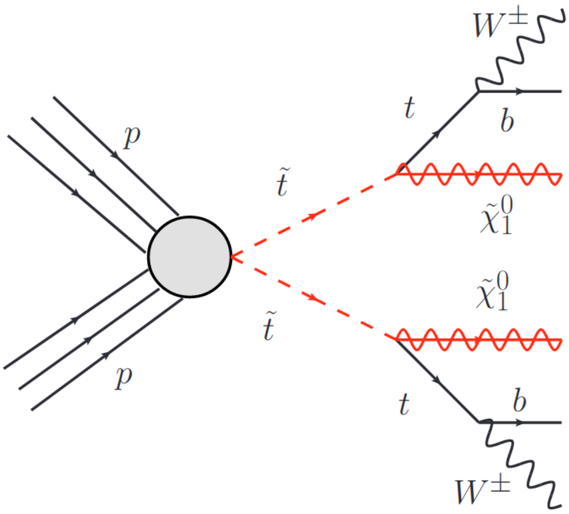


Fig. 6.1 The Feynman diagram of the simplified supersymmetric model considered in the prototype stop pair analysis.

The dominant SM background in this final state is top quark production, including both top pair and single top production. In a real analysis, there would be a small contribution from events containing a W boson produced in association with jets, but we neglect this in our preliminary study. The signal and background are again simulated using `Pythia 8` and `Delphes3` with the ATLAS detector card. The cross-section used for normalisation of the stop sample is the next-to-next-to-leading-order plus approximate next-to-next-to-leading log cross-section derived from Refs. [145–148]. The top sample uses the next-to-next-to-leading-order plus next-to-next-to-leading log $t\bar{t}$ cross-section derived from Refs. [149–155]. The weight on the top background is given approximately by this cross-section, because the subdominant contribution from single top events is scaled by the same value as the $t\bar{t}$ contribution.

6.2 Stop analysis design

Before building the network we require there to be exactly one electron or muon, with transverse momentum $p_\mathrm{T} > 25$ GeV, at least 2 b-jets (with $p_\mathrm{T} > 30$ GeV), and a missing transverse energy, $E_\mathrm{T}^\mathrm{miss}$, greater than 100 GeV. Our prototype analysis selects the most constraining final state: the 1 lepton state in which one of the top quarks decays hadronically and the other decays leptonically. Increasing the constraints on the model to search for a more SM-like signal creates a challenging environment for identifying signal events using any analysis strategy.

6.2.1 Kinematic basis for network construction

We examined a number of kinematic variables typical of stop searches, and select the following six because they show comparatively good discrimination between signal and background processes for the benchmark model. These are described briefly here and defined more fully in Section 3.3.1, and are used to construct networks of stop and top events in six-dimensional collider space.

- $p_\mathrm{T}^{j_1}$: the leading jet transverse momentum,
- $E_\mathrm{T}^\mathrm{miss}$: the missing transverse energy,
- $m_\mathrm{T}^{b,\mathrm{min}}$: the minimum transverse mass formed by the two b-jets and missing transverse energy,
    - $m_T^b = \sqrt{2 p_\mathrm{T}^b E_\mathrm{T}^\mathrm{miss} \left[1 - \cos(\Delta\phi)\right]}$, where p_T^b is the tranverse momentum of each b-jet, and $\Delta\phi$ is the difference in ϕ between the each b-jet and the missing transverse momentum,
- m_{bl}^min: the minimum invariant mass formed by the lepton and each of the two b-jets,
    - For the top pair production process, this has a maximum value which is lower than the range of possible signal values,
- H_T: the scalar sum of the moduli of the transverse momenta for the lepton and the 2 b-jets;
- am_T2: the asymmetric m_{T2} defined in Refs. [61, 62, 156–158].

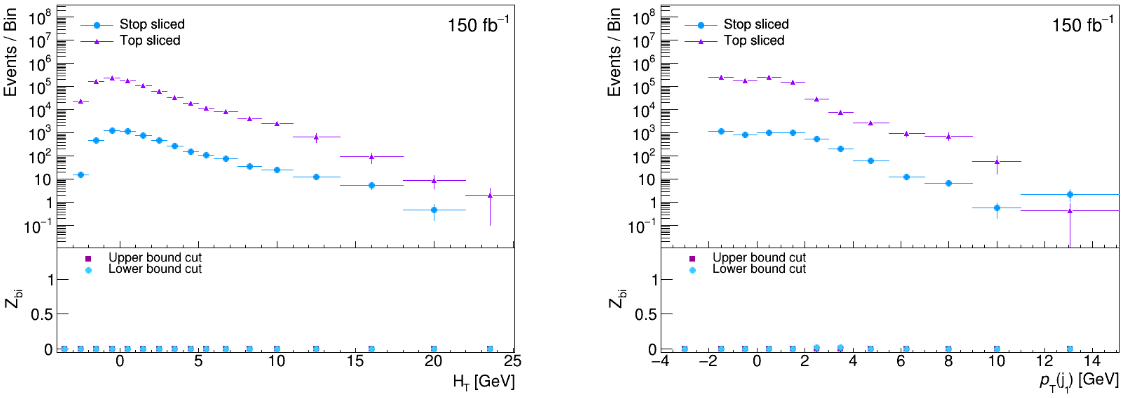
Both the top background and the stop signal samples were generated using the slicing procedure to increase tail statistics in the kinematic distributions. All samples in the stop case study are generated in slices of H_T, the scalar sum of the transverse momenta, unlike the slices in the electroweakino background samples which were placed in the transverse momentum variable $\hat{p}_\mathrm{T}$. The scalar sum H_T is better suited to the stop example; Appendix B discusses the slicing method applied to simulated events used in this chapter, including its motivation and application.

We show histograms of event rates as functions of these variables after the pre-selection in Figure 6.2, with the samples scaled using the "background median" procedure motivated in Section 5.2.1. Although these distributions revealed the greatest separation between signal and background compared with other options for the kinematic basis variables, there remains strong similarity in the shapes of the stop and top distributions. The background is substantially higher than the signal across the entire range in every variable, and the signal is clearly more difficult to isolate using the original kinematic variables than the electroweakino signal, where selections on the kinematic variables alone shown in Chapter 5 carried some exclusion power.

6.2.2 Distance calculations

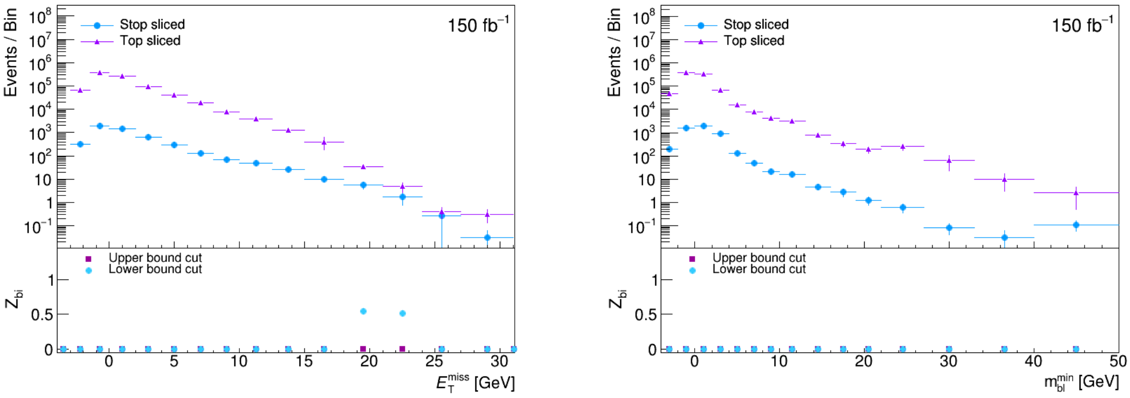
Histograms of the distance metrics are provided for our stop example in Figure 6.3, and were calculated between every pair of events in the six-dimensional collider space. Those distance metrics which did not transpire to provide useful discrimination with any local network metric are suppressed. To construct the networks, we select linking lengths using the method described in Chapter 4 and employed in Chapter 5, producing the distances in Table 6.1. The correlation and cosine metrics have signal more concentrated at small distances from other events than the background, which was also true in the electroweakino example. Although these angular distances are useful in the stop analysis, we retain the standard condition for connecting two events with an edge only if they are closer in distance than the linking length and not if they are further away, avoiding flipped networks for simplicity.

Figure 6.3 contains distance distributions calculated from inclusive samples with no slicing, although we were also required to calculate distances between pairs of events in the sliced samples to create the networks. The inclusive samples are used here instead because they represent the distribution of real events in collider space, and guide our choice of linking length. They represent the distances between events in sliced samples without requiring the



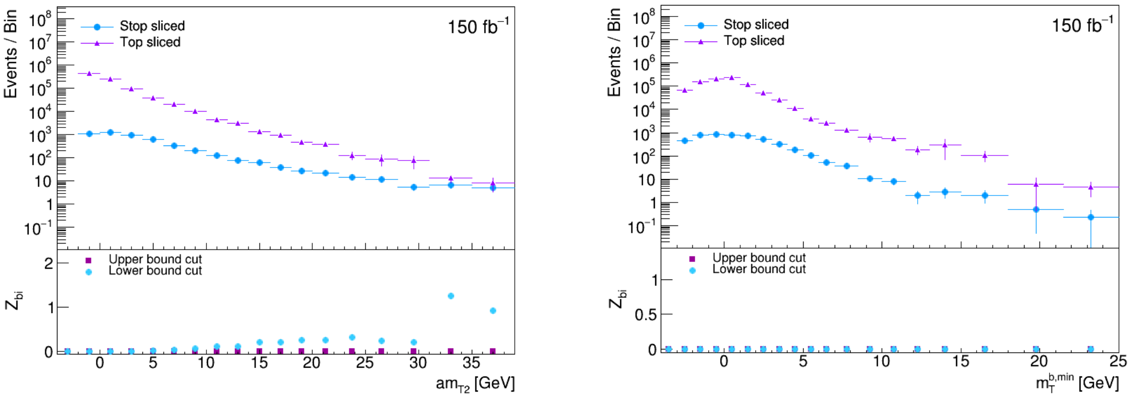
(a) H_{T}

(b) $p_{\mathrm{T}}(j_1)$



(c) $E_{\mathrm{T}}^{\mathrm{miss}}$

(d) m_{bl}^{min}

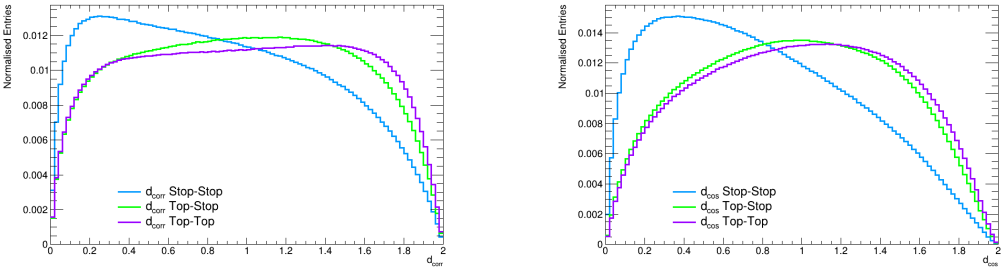


(e) am_{T2}

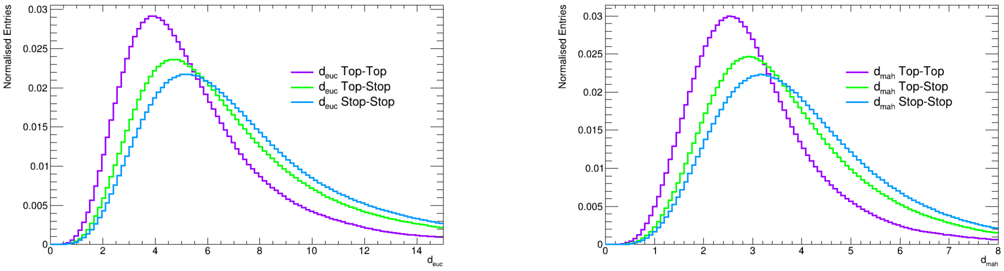
(f) $m_{\mathrm{T}}^{b,\mathrm{min}}$

Fig. 6.2 Event rates as functions of the kinematic variables for the stop simplified model example that show the most difference between the signal and the background. Each variable has been scaled by the “background median” procedure described in the text. Events in the overflow bin are not shown in the distribution but are included in the Z_{bi} calculation.

distance distributions to be weighted, and pairwise distances are numerous enough that low numbers of events do not cause difficulty here as they do for the inclusive samples in the tails of the original kinematic variables.



(a) correlation (b) cosine



(c) Euclidean (d) Mahalanobis

Fig. 6.3 Distributions of the distance metrics we use in our prototype stop search, calculated using inclusive samples.

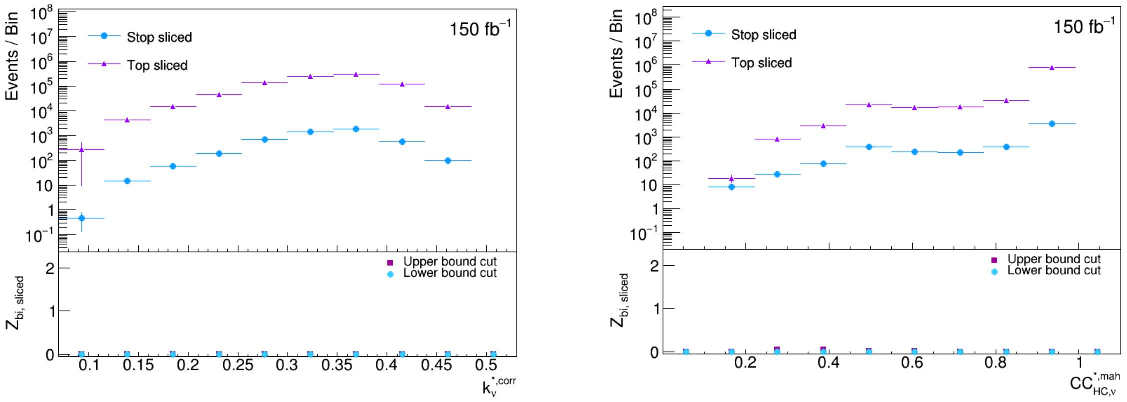
Table 6.1 Linking length values used for each distance metric for our prototype stop analysis.

Distance metric	**Linking length**
d_{corr}	0.7
d_{cos}	0.8
d_{euc}	5.5
d_{mah}	4.0

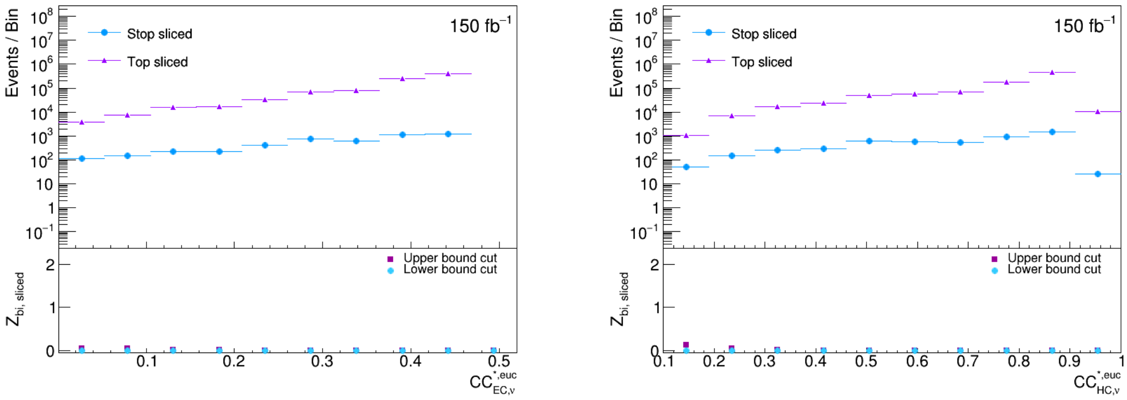
6.3 Network metric distributions

In Figure 6.4, we show the network metrics that were used in the previous example, and that are considered robust under the theoretical reasoning described in Appendix A.1. Although the distributions show some differences in shape, there is not enough separation between the signal and background distributions to render them useful in stop searches. We found no combination of selections on these variables plus the original kinematic variables that gives sensitivity for exclusion at the LHC. It remains possible that different choices of the original kinematic variables used to build the network might change this picture, but it is clear that the use of local network metrics does not automatically give sensitivity to BSM physics signals.

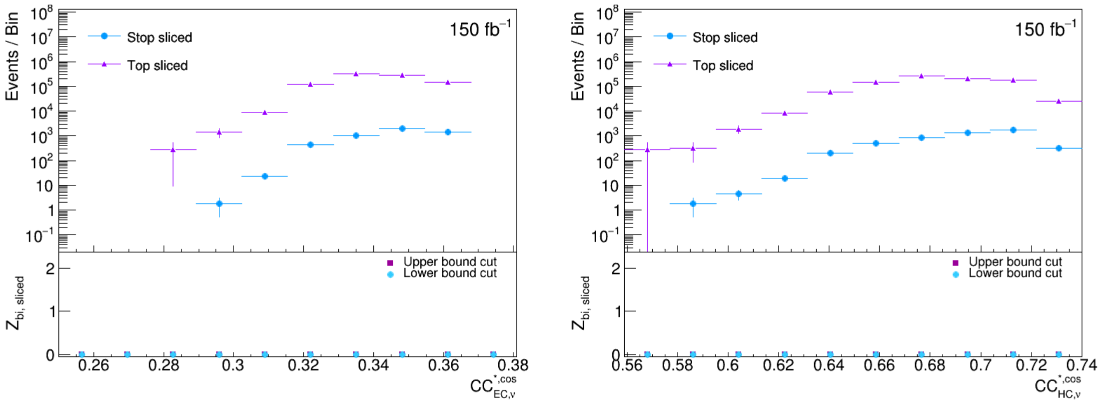
Some network metric distributions which were not as powerful in the electroweakino search in Chapter 5 now show much greater signal-background separation for the stop signal model. Figure 6.5 contains several of these metrics: betweenness centrality distributions for correlation, cosine and Mahalanobis distances, local and local soffer clustering coefficients for Euclidean distance, and average neighbours degree for cosine. The betweenness measures in particular fall off much faster for top events than stop, especially if the network is built using correlation and cosine distance metrics. All three betweenness plots show background distributions decreasing faster than signal, but in particular the tails in the cosine and correlation background betweennesses sink appreciably lower than the signal, producing higher values of Z_{bi} than the Mahalanobis betweenness. More modest discrimination comes from the local and local soffer clustering coefficients and average neighbours degree. If these metrics are trustworthy under the conditions of LHC network construction, then these distributions can define powerful signal regions with high yield. To be considered reliable, the metrics must be robust under node weighting with n.s.i conditions satisfied as described in Section 4.3.1. The n.s.i external connectivity assumption is a focus of our current tests outlined in Appendix A.1, and we recommend caution on the use of these metrics until further investigations are performed.



(a) $k_{v}^{*,\text{corr}}$ (b) $CC_{HC,v}^{*,\text{mah}}$

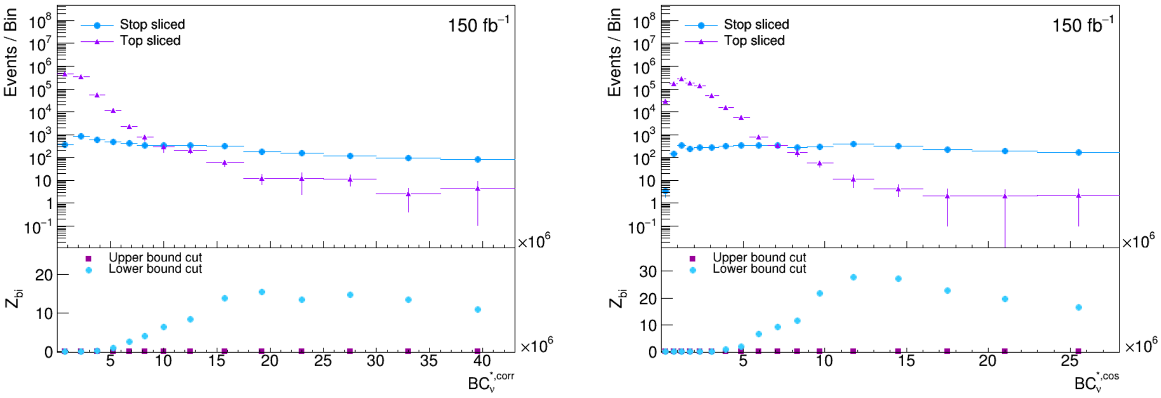


(c) $CC_{EC,v}^{*,\text{euc}}$ (d) $CC_{HC,v}^{*,\text{euc}}$

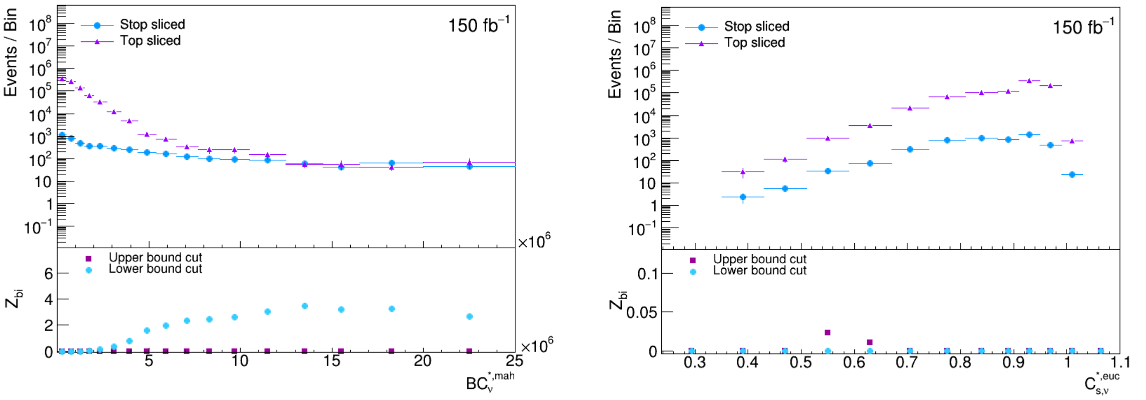


(e) $CC_{EC,v}^{*,\text{cos}}$ (f) $CC_{HC,v}^{*,\text{cos}}$

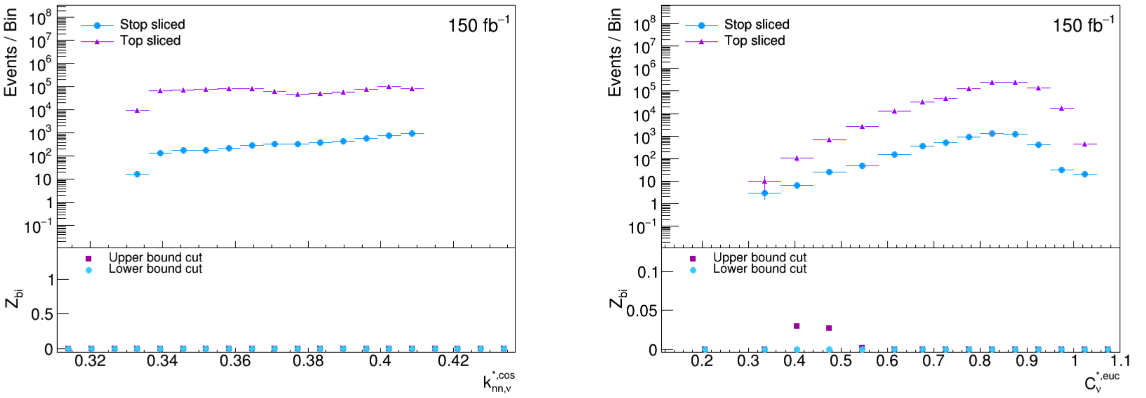
Fig. 6.4 Event rates as functions of the network variables for the stop simplified model example. Events in the overflow bin are not shown in the distribution but are included in the Z_{bi} calculation.



(a) $BC_v^{*,\text{corr}}$ (b) $BC_v^{*,\text{cos}}$



(c) $BC_v^{*,\text{mah}}$ (d) $C_{s,v}^{*,\text{euc}}$



(e) $k_{nn,v}^{*,\text{cos}}$ (f) $C_v^{*,\text{euc}}$

Fig. 6.5 Event rates as functions of the network variables for the stop simplified model example that show the most difference between the signal and the background either when alone or when combined in cuts with other variables. Events in the overflow bin are not shown in the distribution but are included in the Z_{bi} calculation.

We consider the possiblity that the network metrics in Figure 6.5 are affected by bias introduced in the n.s.i calculations when the networks have highly disparate weights. The weight differences are largely a product of the slicing procedure described in Section 4.2.3, which contributes up to two orders of magnitude difference between event weights. This prompts a straightforward method of testing the weights in this context using a comparison between sliced and inclusive event distributions. In Appendix B, we check the agreement between inclusive samples and sliced samples of stop and top events for network metrics presented in the stop case study to consider the effects of slicing on the metric distributions. If the slicing procedure is causing skew in the network metrics, then comparing them with the metrics calculated from inclusive samples should reveal this skew, because the inclusive samples have consistent weights on every event, differing only in stop and top weights according to their respective LHC production cross-sections. Strong similarity between these distributions indicates that the metrics can somewhat be trusted in the context of this case study.

For the metrics which are known to be robust under large node weighting, results may improve if the stop networks contain more events, which would allow the linking length to be optimised for signal sensitivity. Our current limitation on the number of events results from the computational complexity of local metric calculations for large networks. As more signal and background events are added, the number of edges in the network becomes very large, increasing the computational load of our parallelised calculations in the `pyunicorn` package. We aim to study how the prototype analyses change when choosing different linking lengths for each distance metric, where shorter linking lengths may significantly reduce computational complexity because they would produce sparsely-connected networks. In this case, however, background and signal events will often have a low degree and be difficult to discriminate. We address the possibilities for future increases in computing power in Chapter 8. Given our present computational limitations, the next stage of our investigations is to evaluate the shape properties of the metric distributions calculated from networks of real ATLAS data, using appropriate numbers of events under the constraints of such complex calculations.

Chapter 7

ATLAS electroweakino network analysis

7.1 Introduction

The outcomes of the graph network analysis application to supersymmetry searches in Chapters 5 and 6 must now be corroborated with evidence showing that network techniques are valid in a real ATLAS search. The next stage in our event-by-event LHC-style study is to verify that there is significant agreement between the shapes of the network variable distributions calculated from real data and those calculated from simulated models in the same region. We select another electroweakino search region, with dominant diboson WZ background, which was recently revealed by ATLAS to contain no excess of signal events [159]. Ultimately, we aim to create a framework within data-driven collider phenomenology that does not depend on a specialised set of observables, but can instead be built around whichever tools are available for the given physics scenario and search strategy. Therefore, our approach to an ATLAS search for supersymmetry in this chapter is to create networks from the same original variables used to exclude our specific target model. We pursue a narrow search region as a necessary example prior to broader applications in future.

Several key aspects of the previous analyses are verified in this chapter. Whereas LHC data always acquire unitary weights per event, MC backgrounds have non-unitary weights which may span several orders of magnitude. This chapter is therefore not only a check that graph theory variables are calculable on LHC events, but also that the n.s.i method of calculating graph variables is reliable in the presence of vastly different node weights. As well as comprising weighted events, the simulated backgrounds also comprise a set of multiple dominant backgrounds instead of a single background as per the previous analyses.

Although only the WZ background again represents more than 95% of the events in the region selected for this chapter, we aim for a more thorough understanding of network techniques applied to real data and therefore consider the full background set.

The existing search has been unblinded and excludes our chosen model in the signal region, so we compare the network distributions in the same region with those calculated from the MC backgrounds, expecting strong agreement. The region was designed to search for a supersymmetric electroweak signal that differs from the signal simulated in the preliminary electroweakino study in Chapter 5, and the original kinematic variables underlying the networks are also different. The construction provides a useful method of verifying network metric shape properties, as the unblinded analysis predicts that data and background network distributions should agree within the sensitivity of the existing analysis.

7.2 Experimental data and target model

The network analysis in this chapter exploits 139 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data collected with the ATLAS experiment in Run 2 during 2015 and 2016. The data is selected from an ATLAS electroweakino analysis targeting chargino-neutralino production. Whereas the previous analysis searched for several final states with two or three charged leptons, we choose only a single decay topology with a three-lepton final state in the presence of additional initial state radiation (ISR) jets. The decay topology is depicted in Figure 7.1, where the production of $\tilde{\chi}_1^{\pm}\tilde{\chi}_2^0$ is associated with an ISR jet, labelled j. To define a signal region for this final state that is consistent with the region in the previous analysis, and to incorporate powerful original variables into network techniques, we adopt the recursive jigsaw reconstruction (RJR) variables. The RJR method was introduced in Section 3.3.1 and is again discussed in Section 7.2.1. The ATLAS analysis comfortably excluded the search region using these variables.

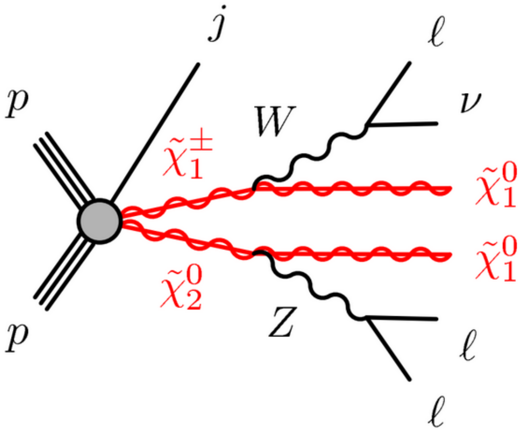


Fig. 7.1 A Feynman diagram showing the target $3l$ signal topology involving the decay of $\tilde{\chi}_1^{\pm}\tilde{\chi}_2^0$ via leptonically decaying W and Z bosons [159].

The experimental signature for this decay topology identifies events with a three-lepton final state along with a set of additional kinematic properties. The signal region defined by these properties is derived from the RJ variables in the following section. We target supersymmetry scenarios with compressed mass spectra, where the sparticles have smaller mass splittings and lower intrinsic $E_{\mathrm{T}}^{\mathrm{miss}}$. Compressed scenarios implicate multiple supersymmetric states nearly degenerate in mass which produce soft decay products, and are therefore an experimental challenge for ATLAS and can only be probed using highly sensitive search strategies. Methods such as RJR may achieve even greater sensitivity to small BSM signals when combined with graph network techniques in these regions.

The strategy targets compressed supersymmetry scenarios using a standard approach that selects events with a partially resolved sparticle system recoiling from high-p_{T} ISR jets, motivating an ISR topology. The high transverse momenta in the target events are derived from the momentum kick passed to the weakly interacting particles from ISR jets. Therefore, the source of the large momenta of LSPs comes not from decay but instead from recoiling against ISR, a feature that causes LSPs to receive higher momentum than their neutrino backgrounds because they possess greater mass. The radiation of quarks and gluons that produces ISR states also contributes extra jets to the measured event. Therefore, our decay topology of interest is distinguished by both its high jet multiplicity and its large transverse momenta, which are typical phenomenological features of supersymmetry.

7.2.1 Basis of variables for network construction

The basis of RJ variables underlying the networks is outlined below. RJR boosts from the laboratory frame define a set of observables in the centre-of-mass (CM) frame which target properties of the visible and invisible systems and are maximally uncorrelated. The compressed spectra motivate the compressed ISR decay tree in Figure 7.2. An RJR approach defines a decay tree that is split into the lab state, decay states, visible states and invisible states, and an ISR system splits the decay tree into the CM frame, the ISR frame and the system frame S.

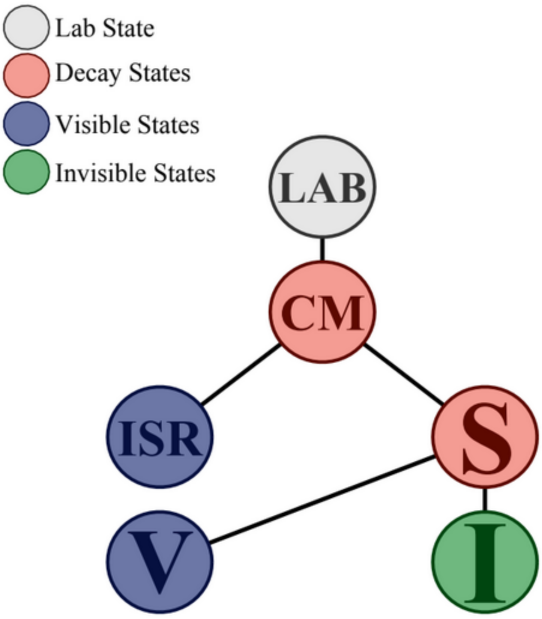


Fig. 7.2 A recursive jigsaw decay tree for a compressed ISR system. The signal sparticle system S decays to a set of visible momenta V, and the invisible momentum I recoils from the ISR system [159].

The variables that characterise the compressed ISR system are built upon the transverse momenta of the ISR and invisible systems, where the invisible system “I” contains the LSPs, $\tilde{\chi}$. An ISR system produces events with at least one high-p_{T} jet. Selected RJR variables are defined as follows:

- $p_{\mathrm{T,\,ISR}}$ is the magnitude of the vector-sum of the transverse momenta of all ISR associated jets evaluated in the CM frame, which gives the momentum of the ISR system.

- $p_{\mathrm{T,\,I}}$ is the magnitude of the vector-sum of the transverse momenta of the I system in the CM frame, which corresponds to the missing transverse momentum without the ISR recoil.
- R_{ISR} is chosen for its sensitivity to the ratio $m_{\tilde{\chi}}/m_{\tilde{p}}$, where $\tilde{p}$ denotes the parent sparticles. It is a measure of the amount of momentum in the invisible system that can be attributed to the ISR kick, and is calculated from the transverse momenta of the ISR and I systems according to Equation 7.1.

  $$R_{\mathrm{ISR}} = \frac{|\vec{p}_{\mathrm{T,\,I}} \cdot \hat{p}_{\mathrm{T,\,ISR}}|}{|\vec{p}_{\mathrm{T,\,ISR}}|} \tag{7.1}$$

  R_{ISR} is a projection of the invisible system onto the ISR boost-axis in units of the ISR system momentum.
- $\Delta\phi_{\mathrm{ISR,\,I}}$ is the azimuthal angular difference between the ISR-associated jets and the invisible system in the CM frame.
- m_{T} is the transverse mass of the S (V+I) system, including the mass of all jets associated with the visible system.

We use these variables not only to define the signal region, but also as the basis to construct a network. The RJ variables increase sensitivity to the chosen signal compared with other choices of kinematic variables, and applying network analysis with this set of highly sensitive underlying variables is intended to lend the greatest power to the network approach.

7.3 Background estimation

The three lepton search region is dominated by the WZ diboson irreducible background whose shape is taken from MC simulation and fitted to data in dedicated control regions using normalisation factors. For the compressed decay tree selection, the normalisation factor is 1.13 ± 0.13 as calculated by the ATLAS Collaboration [159]. Additional contributions to the modelled background come from sources including VVV (triboson), $t\bar{t}V$ and other processes with a Higgs boson.

Backgrounds are used to construct separate networks from the data, and the network metrics are calculated from each separately. The metrics from the data-only networks are then compared with those calculated from the background-only networks to determine if they are a reliable indicator of the event types present in data. We aim to identify distinct shape properties within the background-only networks which depend on the composition of events and should be shared identically in the data networks only if signal is absent.

7.3.1 Negative node weights

Some background events are simulated in ways that require negative node weights, which are designed to cancel their contributions with positively weighted nodes. For example, Section 3.3.2 describes integrating a differential cross-section at NLO with a combination of divergences, requiring correction terms with the same infrared pole structure as the real matrix element. The correction factors subtract the over-estimated quantities, producing negatively weighted events. Any such events that pass preselection propagate their negative weights into the network, assigning nodes with “importance” less than zero. These weights must then become part of the n.s.i network metric calculations, where they are used to evaluate the characteristics both of the negatively-weighted nodes and of the other nodes whose local properties are affected.

The network calculations may become biased by negative weights and produce nonsensical results. Amongst these metrics is the betweenness centrality, which no longer reliably represents a well-defined quantity in a naive MC background network with arbitrary negative weights. The n.s.i betweenness metric involves a weighted product which can become negative if the number of negative contributions to the product is odd, and yet betweenness is by definition non-negative: it is defined as the proportion of possible shortest paths passing through a given node, calculated in the unweighted case from the ratio between two non-negative quantities. Although other metrics such as degree centrality do not become nonsensical when nodes are assigned negative weights, we aim to achieve a consistent method of constructing our networks and therefore choose to remove the negative weights from all network calculations.

Several options are available for calculating the metrics which manage this difficulty. A procedure may be adopted from certain scenarios involving *edge* weighted networks, where we may first select the smallest or most negative weight and add its absolute value to all node weights, transforming the network into one with weights greater than or equal to zero. However, in an LHC network the physical significance of the weights is lost when all events are scaled in relative importance by a common value; negatively weighted events are no longer subtracted, and their contributions do not cancel the infrared divergences. Furthermore, this naive approach cannot always preserve the network characteristics even in the edge weight case, where shortest paths in the new network are insufficient to represent those in the old one. Other options were explored for achieving our desired outcome.

We eliminate the negative weights before building our networks by selecting one negative event and locating a set of one or more positively-weighted events with maximally similar values of the original variables used to construct the underlying graph space. We then sum their weights with the negative weight until its contribution to the simulated dataset is cancelled. The same procedure is applied to every negative weight, so that the final set of events produces a network with only positive node weights. The method assumes that all negative events are located close-by in collider space with one or more positive counterparts. After cancellation, the network should more accurately contain the physics information required to represent a real network of LHC events.

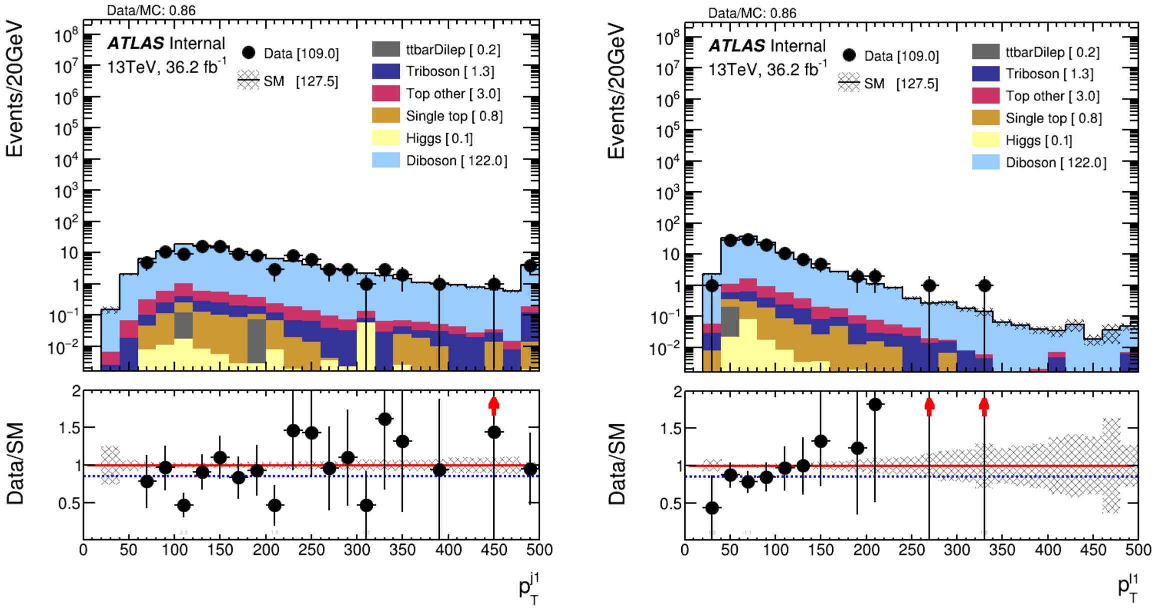
7.4 Search region

Preselection criteria are applied to the variables in Section 7.2.1 to define the search region. These criteria are taken from the ATLAS search for neutralinos and charginos belonging to the aforementioned supersymmetry model, with a compressed mass spectrum targeting a three lepton final state. The "loose" identification criteria for this region require one *b*-tagged jet plus the set of additional constraints which are presented in Table 7.1. After applying these criteria, the number of events remaining in data is 109, which we consider sufficient for a first test of network analysis given that distances are calculated between every pair of events. The loose cut is the least constrained of the regions defined in the existing ATLAS study.

Table 7.1 Selection criteria for the compressed $3l$ ISR signal region with loose constraints [159].

Variable	$p_{\mathrm{T,\,ISR}}$ [GeV]	$p_{\mathrm{T,\,I}}$ [GeV]	R_{ISR}	$\Delta\phi_{\mathrm{ISR,\,I}}$	m_{T} [GeV]
Selection	>100	>80	$>0.55, < 1.0$	>2.0	>100

Kinematic variable plots are considered alongside our network analysis for two main purposes. First, we verify that the region selection is sufficient for our aims by ensuring that the original variables are populated by enough data events to achieve smooth data distributions, and enough background events for the region to be considered well-modelled. In all subsequent figures, the number of background events passing preselection is approximately 25,000, which slightly exceeds the maximum number added to a single network in the preceding chapters. Second, we compare the similarity between data and backgrounds in the original variable distributions with their agreement in the network metric plots in Section 7.6. The similarity here then provides a baseline against which the network metric distributions must compete, assuming that no new physics is manifest in this region that can be detected by our sensitive network variables but not by the original methods. The distributions in Figure 7.3 are a standard choice of two typical variables representing the kinematics in the region. Although the data distributions show notable statistical fluctuations and minor deviations from the expected backgrounds, the MC modelling and numbers of events are sufficient. These deviations guide our expectations for the agreement in later network metric distributions.



(a) p_T^{jet1} (b) p_T^{lep1}

Fig. 7.3 Event rates as functions of the transverse momentum for the (a) jet and (b) lepton with largest transverse momentum.

7.5 Distance calculations

The networks of ATLAS data and backgrounds in this chapter are constructed using four distance calculations selected from the set considered in Chapters 5 and 6 as a representative sample. We select a range of angular and non-angular distances: the correlation, cosine, Euclidean and Mahalanobis. In addition, our test process involves checking the Bray-Curtis, Chebyshev and cityblock distances, which are found to have distributions similar in shape to the Euclidean distance and are therefore disregarded here to simplify the presentation. All distances between events in background-only networks are entirely distinct from those between events in data-only networks, as the two types never overlap. Linking lengths were chosen by applying the same method used in Chapters 5 and 6, and are shown in Table 7.2.

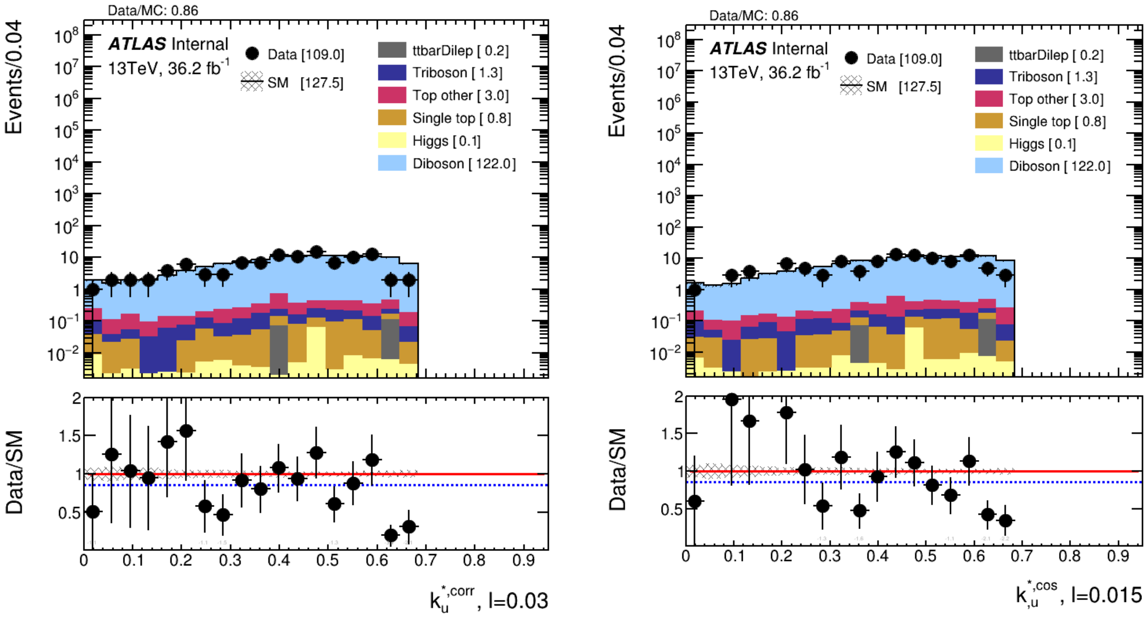
Table 7.2 Linking length values used for each distance metric for our ATLAS data test of network analysis.

Distance metric	**Linking length**
d_{corr}	0.03
d_{cos}	0.015
d_{euc}	70
d_{mah}	2.5

7.6 Results of network analysis

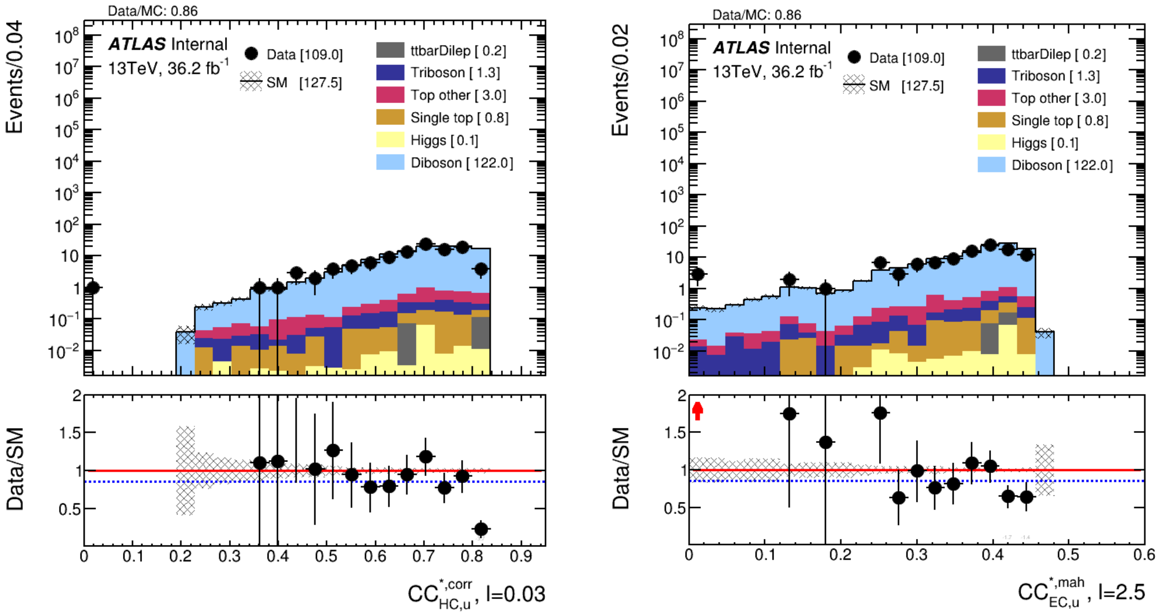
The local network metric distributions for data and backgrounds are compared in Figures 7.4 and 7.5. They show strong agreement between the ATLAS data and simulated MC backgrounds, as expected in a search region containing no other evidence of signal. We interpret this agreement as evidence that our analysis provides effective measures of real event properties. Not only do the network metrics show highly similar distributions between data and backgrounds for robust n.s.i calculations, but also for the n.s.i calculations which remain uncertain according to Appendix A. A sample of several of the metrics which we treat with caution is presented in Figure 7.5. The local and local Soffer clustering coefficients for the Euclidean distance in particular show distinct shape properties, suggesting they may measure different non-trivial network characteristics. The betweenness centrality is a computationally intensive calculation and is not yet measured at an appropriate linking length for networks of such large size.

We aim to verify that these shape properties are a feature of the physics scenario and depend on the distribution of events by comparing with false datasets. To make the distinction, we replace the background events with a set of the same number of event-like points selected at random between the maximum and minimum values of the true MC background set. Figure 7.6 presents a subset of the metrics shown in Figures 7.4 and 7.5, comparing the true LHC data distributions with false backgrounds. The false events are defined in five



(a) $k_v^{*,\text{corr}}$

(b) $k_v^{*,\text{cos}}$

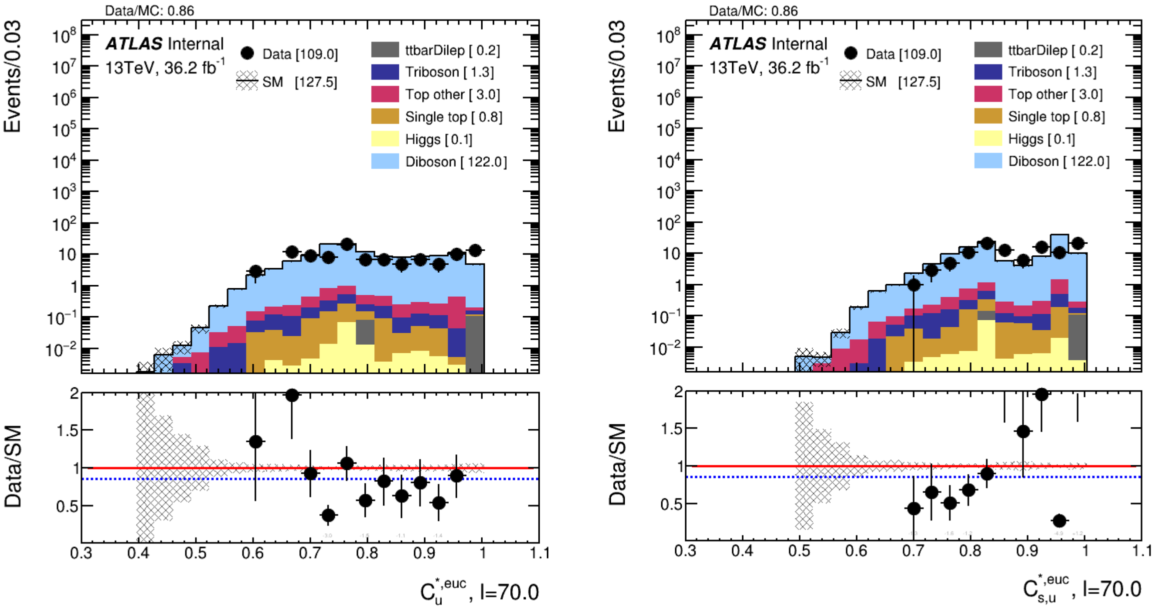


(c) $CC_{HC,v}^{*,\text{corr}}$

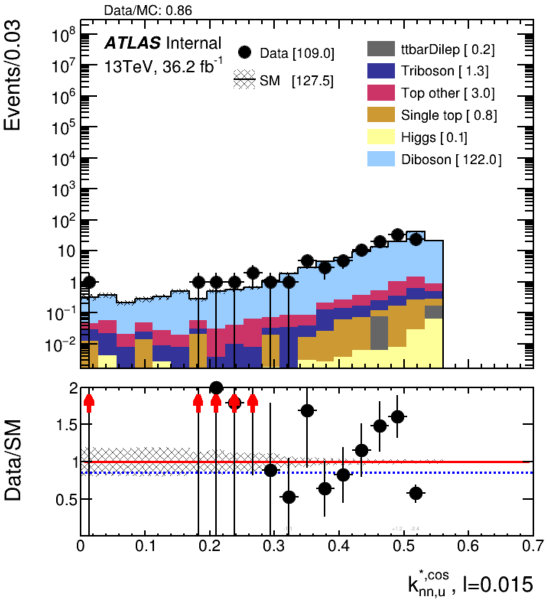
(d) $CC_{EC,v}^{*,\text{mah}}$

Fig. 7.4 Event rates as functions of the network variables for degree and closeness centrality metrics calculated for data-only networks with 109 events, compared with distributions of $\sim 25,000$ dominant SM background events normalised to fit. No events fall in the overflow bin.

variables representing the original set, and replace the underlying variables used in network construction. The value of every event in each variable is selected randomly from a flat



(a) $C_v^{*,\mathrm{euc}}$ (b) $C_{s,v}^{*,\mathrm{euc}}$



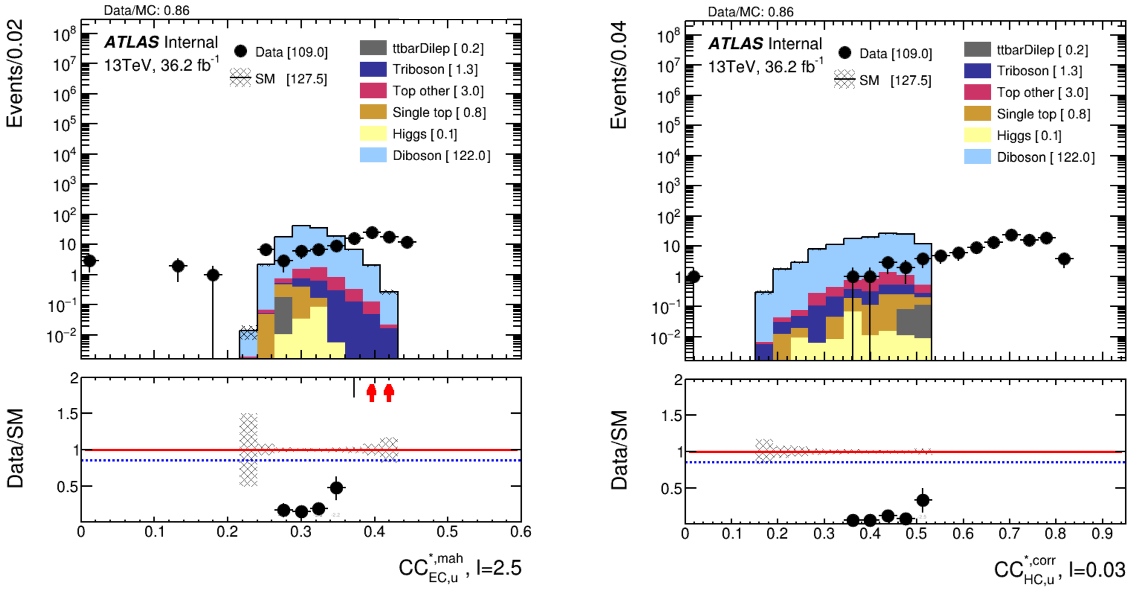
(c) $k_{nn,v}^{*,\cos}$

Fig. 7.5 Event rates as functions of the network variables for clustering, soffer clustering and average neighbors degree metrics calculated for data-only networks with 109 events, compared with distributions of $\sim 25,000$ dominant SM background events normalised to fit. No events fall in the overflow bin.

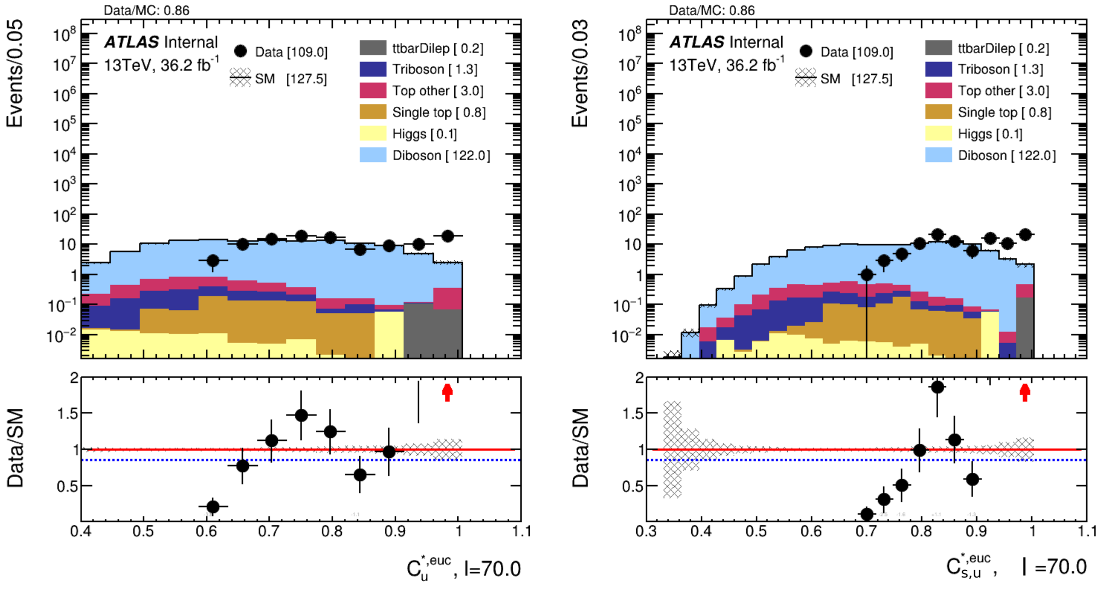
distribution within the same ranges as the original variables, while total event weight is held constant.

The resulting network distributions are sufficiently different to convince us that the shape properties of the metric distributions cannot easily be reproduced with a random network. Furthermore, several iterations of generating false events and plotting the same network metric distributions shows different behaviour every time. We conclude that the data and background distributions do not only agree because the network metrics must somehow inherit the same shape properties in this space of original variables. Instead, their particular event distributions within the space are influencing the metric calculations. In this case, non-trivial information is contained both in the strong agreement between data and backgrounds and in the metric shape properties.

The network analysis techniques we apply in this chapter offer a new method of distinguishing data from backgrounds when signal is present, and of increasing the confidence level of future exclusions for BSM scenarios in LHC analyses in the absence of signal. Our current progress tests network techniques for a BSM analysis but could also be considered a preliminary test of network applications to SM precision measurements, where the parameters of the SM are extracting from sensitive variables. Further investigation will help us to evaluate the significance of network results and to determine the confidence levels they can contribute to future analyses.



(a) $CC^{*,\mathrm{mah}}_{EC,v}$ (b) $CC^{*,\mathrm{corr}}_{HC,v}$



(c) $C^{*,\mathrm{euc}}_{v}$ (d) $C^{*,\mathrm{euc}}_{s,v}$

Fig. 7.6 Event rates as functions of the network variables for closeness centrality and clustering metrics calculated for data-only networks with 109 events, compared with distributions of background-like events selected randomly from flat distributions and normalised to the total background weight sum in Figures 7.4 and 7.5. No events fall in the overflow bin.

Chapter 8

Conclusions and outlook

Graph network analysis provides powerful tools for calculating properties of the structures of complex systems, and has now for the first time been applied to LHC searches for new physics. Our proof-of-principle studies suggest that network techniques may lend BSM analyses with greater sensitivity to supersymmetry signals. Furthermore, these methods naturally extend to searches for a wide range of BSM physics scenarios, and further to applications in model-independent searches.

Using networks to discriminate signal from background processes at the LHC depends on a large set of convenient variables defined by mathematical graph theory. The range and usefulness of network variables is enhanced by constructing networks from several different metrics for calculating the distances between events in the underlying graph space. The usefulness is enhanced further by selecting appropriate basis variables to define this space, not only from the basic set of kinematic collider variables, but also from more advanced measures derived using techniques such as RJR. Network techniques therefore build on the sensitivity obtained from whichever variables define the space, which may include options that already lend strong discovery potential to existing searches. After we construct a series of networks with different distance definitions, we calculate the network metrics to incorporate additional sensitivity derived from network structures. Whereas typical strategies in LHC analyses design variables targeting properties of isolated events, our novel approach is sensitive to topological structures arising from the distribution of all surrounding events. Although our network metrics measure properties which are defined locally for every event, we obtain more information about each event than we would obtain by treating it in isolation by considering its network connections.

The local network metrics are assigned as new event attributes, and the network analysis sensitivity is evaluated in Chapters 5 and 6 using standard methods of selecting events passing cuts on these attributes to produce signal and background yields. From calculations of binomial significance, our results in Chapter 5 indicate that graph network methods offer comfortable exclusion potential for the electroweakino production scenario. Furthermore, adding a local network metric to a BDT which was originally trained only on the conventional kinematic variables increases its performance. These results indicate that the technique provides a promising alternative to current search methods, and motivate a deeper analysis into further options for distance metrics, linking lengths and preselections.

The stop production scenario in Chapter 6 contains a signal process with kinematics closely resembling those of the top background, and the network metrics which are robust under the n.s.i assumptions do not prove to be sensitive to exclusion of this scenario. However, several metrics which remain uncertain under n.s.i assumptions and are undergoing further investigation revealed a significant difference in behaviour between the signal and background events, suggesting that some network measures may also provide significant exclusion sensitivity in the stop analysis. The betweenness centrality distributions calculated with the correlation, cosine and Mahalanobis distances show unique shapes for the signal compared with the background, a property that we may exploit in future work. The signal distributions are visibly flatter with longer high-betweenness tails, whereas background events are grouped at low betweenness. Exploiting these distinct properties will require further tests of the betweenness and other remaining metrics to evaluate their behaviour under the n.s.i external connectivity assumption. We hope to demonstrate with certainty in which scenarios both internal and external connectivity assumptions are satisfied, and also consider the possibility of developing alternative forms of the network metrics which are well-defined under violations of these two assumptions.

The results of network analysis with ATLAS data in Chapter 7 provide further evidence suggesting that these techniques have promising potential in real LHC searches. Our progress confirms that n.s.i metrics produce reasonable distributions when calculated from LHC data networks compared with the expected backgrounds. Further, we find agreement between data-only and background-only metrics in a region that is already excluded by ATLAS, therefore agreeing with expectation. We use these preliminary tests as a guide for future research into the discovery potential of ATLAS data networks across a range of signal models, and consider which methods may best isolate a generic signal in anomaly detection. A point of interest in this venture is the impact of parameters such as preselection criteria on the exclusion potential for a signal model. To determine the constraints of such parameters, faster and less computationally-intensive calculations will enable us to search through larger and more varied ATLAS datasets. In the future, we consider the option of performing a full ATLAS analysis using network variables to achieve a complete comparison with existing results.

Another parameter motivating faster calculations with larger networks is the linking length, which sets the scale of interest, and is currently assigned an appropriate value from the distance distributions. Larger linking length produces greater connectivity and therefore more computationally-intensive calculations. The network recipe depends on the linking length to produce non-trivial connection topologies, but this parametric representation of the particle collider event distribution is necessary only if there exists a physical reason for the scale. Should there be no physical constraint on the scale, this parameter may be an unnecessary construct that can be replaced by shape-finding algorithms which are designed to remove scale dependence using the Hessian matrix. These algorithms perform smoothing operations with multiple options for the scale to identify structures which are consistent across several scales. Shape-finders and other algorithm-based network techniques may help to evaluate the parameter dependence of collider event networks in future work.

Once the parameter-dependence of network characteristics and topologies is better understood, the network analysis presented here can be expanded in several broad ways. Global network metrics differ from local metrics in their assignment of a single value to all nodes, evaluated at a chosen scale. For example, transitivity is a global metric calculated by averaging the local clustering coefficient over all nodes, and can be calculated for a single set of events in a constant underlying collider space across a wide range of linking lengths. Plotting the global metric against linking length then pinpoints the scale of any scale-dependent new physics visible in this metric. The method may prove useful in future anomaly detection applications, where we design a broad analysis with no optimisation on a particular scenario that is manifest at a known scale. It is possible that an SM-only network has distinct topology from an SM-plus-signal network, where the signal is given by one of many possible BSM scenarios, and that these distinctions are visible at one or more network scales. A looser preselection and larger linking length would help to broaden an anomaly search, but both options increase the number of events per network. This future analysis is further motivation to develop computational improvements, possibly including the parallelisation of the `pyunicorn` package. Since the BSM parameter space is so broad, enhancing the performance of network techniques with many events and developing broader network metric-based searches will save time and efforts and lead to a faster discovery.

Extracting information about the topologies of collider event networks may contribute significant discovery potential for BSM physics, and benefits from straightforward combination with other analysis strategies. The metrics in this work may in future combine with topological data analysis (TDA), an approach for extracting information from high-dimensional datasets using a framework that targets distance metric-independent shape properties. Furthermore, machine learning models may be trained to optimise networks for discrimination between BSM and SM event topologies, once they are better known. Expanding our study of event topologies to incorporate these advances, as well as data from future developments in collider experiments, provides interesting avenues for high-energy particle physics analyses.

Appendix A

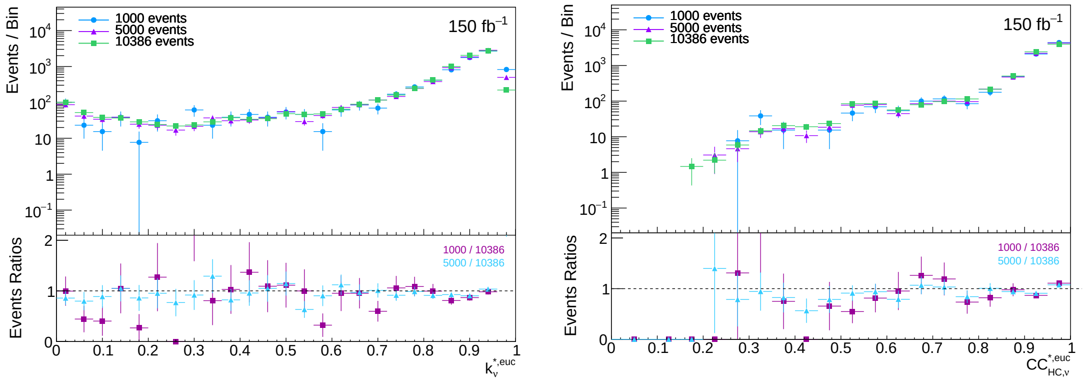
Network techniques

A.1 Justifying the robustness of n.s.i network metrics

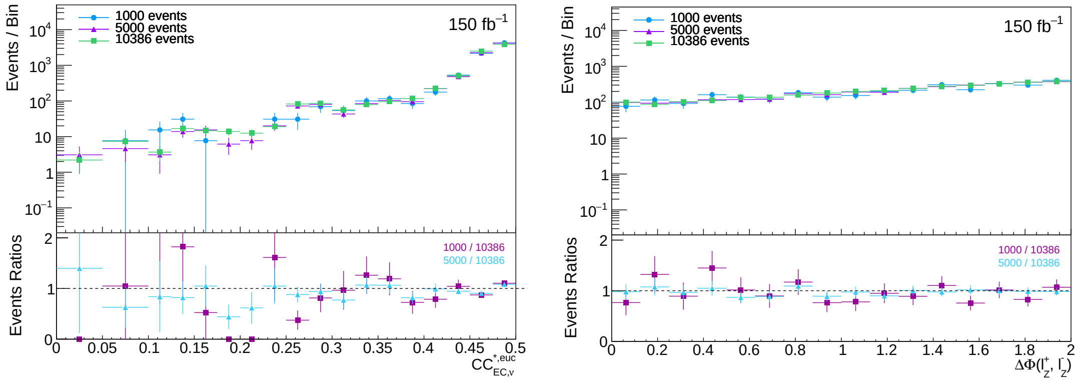
We have tested the reliability of the n.s.i network metrics in our study when events are reweighted. Building on the considerations explained in Section 4.3.1, in A.1.1 we consider how the network variables behave when weights are redistributed amongst different numbers of nodes in an LHC context. Later, A.1.2 examines the n.s.i simplifying assumptions from a purely theoretical perspective, evaluating their applicability to the present example.

A.1.1 Empirical Tests

In networks containing intermingled MC signal and background events together, we consider the impact on the network metric distributions caused by changing only the number of background events. This test is designed to indicate how the network metric calculations may be affected by a redistribution of node weights. Although the test in Section 4.3.1 performed a similar function, we obtain further useful information here by delving specifically into the LHC case, where we create networks that we are likely to encounter in a real LHC analysis instead of using more general toy datasets. We first consider several local network metrics from the electroweakino study in Chapter 5 for networks constructed using the 11,197 signal events combined with the full set of 10,486 background events. Both sets of events are inclusive (not sliced), so for simplicity all events have one of only two possible weights: the signal or background LHC luminosity weighting factor. To consider the effect of increased “node merging”, which is described in Section 4.3 as the inverse operation to node splitting, we randomly select subsamples of background events to create with only 1,000 or 5,000 nodes, weighting these nodes higher so that they represent the full background set. Consequently, a single background node in these smaller samples represents a larger number of true events, which are considered “merged”. Overlaying the distributions in Figure A.1 corroborates the results of the test in Section 4.3.1, revealing that larger event samples produce less sparse tails and smoother distributions but maintain the shape in the bulk.



(a) $k_v^{*,\mathrm{euc}}$ (b) $CC_{HC,v}^{*,\mathrm{euc}}$

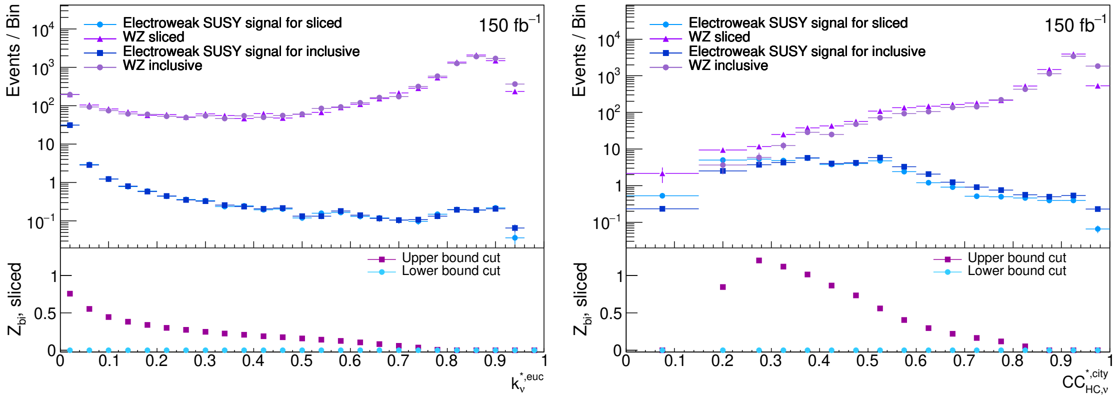


(c) $CC_{EC,v}^{*,\mathrm{euc}}$ (d) $\Delta\Phi(l_Z^+, l_Z^-)$

Fig. A.1 Event rates for the electroweakino case study as a function of the network metrics calculated using d_{euc}, calculated using different background sample sizes. Events in the overflow bin are not shown in the distribution.

We now consider the possible impacts on the metric distributions caused by the increase in the range of weights introduced by the slicing procedure, which assigns low weights to tail events. The shapes of the n.s.i network variable distributions appear to be robust under these changes, revealing little difference in shape when weights are redistributed by the slicing method. Many metrics show similar agreement between the distributions calculated from sliced and inclusive samples, suggesting that for this range of LHC events and variables, the network metrics behave reliably. Considering Figure A.2, which shows $k_v^{*,\mathrm{euc}}$ and $CC_{HC,v}^{*,\mathrm{city}}$ for either an inclusive WZ sample or a WZ sample sliced in p_T, we conclude that the small changes in the network distributions in the signal are caused by the small changes in the background, but are not a sign of unstable behaviour. Further discussion is considered on the impact of slicing in Appendix B, where we consider the maximum difference between node weights in the Chapter 6 stop events.

Although the results presented so far are promising, without comparing with a network of fully unweighted MC events, we cannot produce conclusive empirical evidence that metrics from node-weighted networks can represent metrics belonging to true unweighted LHC data networks.



(a) $k_v^{*,\mathrm{euc}}$

(b) $C_v^{*,\mathrm{euc}}$

Fig. A.2 Event rates as functions of $k_v^{*,\mathrm{euc}}$ and $CC_{HC,v}^{*,\mathrm{city}}$ for the electroweakino simplified model example, comparing the network variables calculated from inclusive WZ samples with those calculated from WZ samples sliced in p_T. Events in the overflow bin are not shown in the distribution but are included in the Z_bi calculation. In both cases the SUSY signal sample is the same, and the legend indicates the associated WZ sample its network was built with.

A.1.2 Theoretical Considerations

We further explore the possibility of introducing artificial distortions in weighted networks by appealing to theoretical arguments. Our approach to defining complex network metrics on a node-weighted network was first developed by Heitzig et al. [121], who derive network metrics for node-weighted networks using an axiomatic approach. The authors assign a node weight proportional to the size of the underlying domain that the node represents. Additionally, the n.s.i calculations in `pyunicorn` require that the local network metrics are invariant under the node splitting operation described in Section 4.3. Chapter 4 describes node-splitting as replacing any node s (of weight w_s) with two nodes s' and s'' (weights $w_{s'}$ and $w_{s''}$), where $w_s = w_{s'} + w_{s''}$. The new nodes s' and s'' are connected both to each other and to the same set of nodes as the original node s. If G is the network which contained node s, and G' is the refined network which contains s' and s'', then all n.s.i network metrics θ^* must satisfy

$$\theta^*(v \in G) = \theta^*(v \in G') \text{ for } v \neq s, s', s'' \tag{A.1}$$

and

$$\theta^*(s) = \theta^*(s') = \theta^*(s'') \tag{A.2}$$

We are concerned with how well n.s.i calculations from a smaller, lower-resolution node-weighted network G_{lr} may approximate a larger, higher-resolution, unweighted network G_{hr}. True LHC data is unweighted, so our network calculations for weighted background events must be compatible with the unweighted data networks. To evaluate the success of n.s.i calculations, we consider their assumptions: first, that a weighted node may represent a group of nodes with identical internal connectivity, and second, that the group of nodes possesses identical external connectivity. Figure A.3 shows a diagram where the conditions are both satisfied, pictured in red (internal condition) and purple (external condition) respectively. Some tests of these conditions are discussed in Section 4.3.1. To test them further in this section, we first acknowledge that we can only expect one or both to be satisfied in a lower-resolution network if the traits can appear in the corresponding higher-resolution network: the larger network should contain tight groups of almost identical nodes. If so, then the inverse operation of node merging may be performed on these groups to produce the approximation. Dense clusters possess highly specific topology, which may be missing from the low-resolution networks when n.s.i calculations are applied incorrectly.

We illustrate a small example of two data samples, one larger and unweighted and one smaller with high weights, to discuss the effects of constructing networks from these samples on the n.s.i conditions. In Figure A.4, the 10 blue circles are a smaller sample from the same distribution as the 100 yellow dots. A node-weighted network constructed from the blue circles may relate ideally to a network built (with the same linking length) from the yellow dots, so that every blue node corresponds to a group of 10 fully internally connected and identically externally connected yellow nodes. In such a small sample, this clearly becomes less likely, although it remains possible. If the linking length is long enough, then the assumption of full internal connectivity will still hold almost perfectly. The networks in Chapters 5 and 6 are typically well-connected: for example, electroweakino networks in Chapter 5 have link densities ranging from 0.213 to 0.492, which are high values increasing

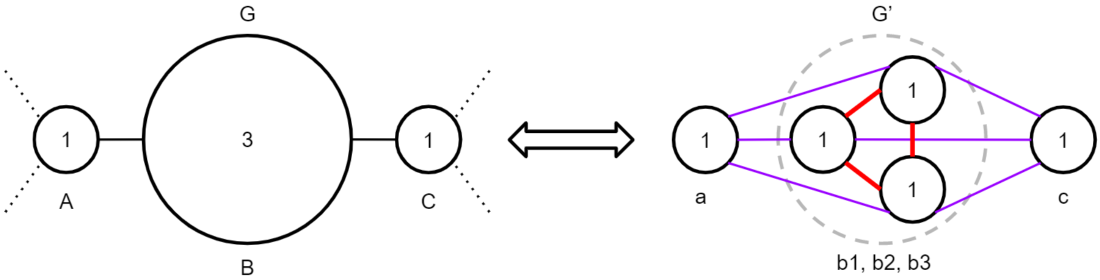


Fig. A.3 The principle of node splitting invariance means that a node-weighted graph G is equivalent to a refined graph G' where all nodes have been split into a number of unweighted nodes proportional to the weight. Here node B is split into b1, b2 and b3. The nodes b1, b2 and b3 are assumed to have i) full internal connectivity (red links) and ii) identical external connectivity (purple links).

the likelihood that a group of proximate nodes from G_{hr} most accurately represented by the same weighted node from G_{lr} are all inter-linked.

Large link density and high linking length are not in general strong enough to account for the second n.s.i condition requiring identical external connectivity. For example, the two blue circles A and B in Figure A.4 fall within the linking length, but only some of their nearby yellow dots also do. If LHC data is tightly clustered, equivalent to tight clusters of the yellow dots centred on the blue circles, then the identical external connectivity assumption will almost perfectly hold. However, without strong evidence that this is true in the LHC datasets we consider, we choose only to rely on n.s.i network metrics that are robust regardless of external connectivity. In Section A.1.2, we justify why the n.s.i degree and n.s.i closeness, harmonic closeness and exponential closeness centralities are reliable.

Justifying the choice of network metrics used

For a complex network based on geometric proximity, we interpret the weight of a node as representing approximately the density of nodes in a nearby region. With a long enough linking length, which we argue is true in this thesis, all potential nodes represented by a weighted node are connected, satisfying the first n.s.i assumption. Close proximity causes ndoes to share many external neighbors, but their external connections are not necessarily identical. Assuming identical external connectivity may, for some network metrics (e.g. betweenness centrality), cause a node-weight-dependent bias in the metric distributions. Identifying the cases where this is true requires further work. However, we argue that the n.s.i degree centrality and the three variants of the n.s.i closeness centrality are certainly safe

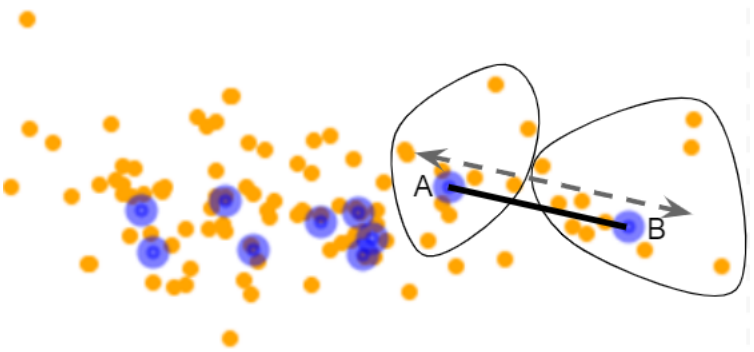


Fig. A.4 Visualizing in 2-D two data samples representing the same distribution. The graph G_{hr} (higher-resolution) is then constructed from the yellow nodes and the weighted blue nodes form the lower-resolution version G_{lr}. Circles around nodes A and B contain approximately the 10 closest yellow nodes, which would be assigned uniquely to them under n.s.i operations. The linking length L (grey dotted line) connects A and B and any other node pairs within distance $d < L$.

to use in this analysis. This conclusion agrees with preliminary empirical tests (here and in previous chapters), but further work is required to investigate the conditions under which network metrics become unreliable. These conditions may include network configuration, network size, distribution of node weights and link density. The investigation is left as an important line of future work, as it may support future use of metrics discounted here, such as betweenness centrality.

With a node $v \in G_{lr}$ representing the nodes $v_1, ..., v_{w_v} \in G_{hr}$, we assume the set of neighbors of a weighted node v in G_{lr} is to a good approximation representative of the set of neighbors of $v_1, ..., v_w$ in G_{hr}. Consequently, the n.s.i degree of v will be approximately the average of the degree of all $v_1, ..., v_{w_v}$. The approximation becomes less accurate if $v_1, ..., v_{w_v}$ have a sparse, non-symmetric or otherwise non-trivial structure, but this is primarily a problem for smaller, weighted MC samples approximating finer-grained data sets, and is not compounded by the n.s.i approximation.

The n.s.i closeness centrality of a node u measures the average length d_{ui} of shortest paths from u to any other node in the network i. The three versions of closeness centrality differ in how they perform this average; closeness centrality sums the length of the shortest paths from u to each other node in the network, and takes the inverse of this sum; harmonic closeness centrality sums the inverses of the distances of shortest paths from u to every node in the network; exponential closeness centrality sums the inverses of $2^{d_{ui}}$. In all cases node weights along each path are multiplied together to approximate the number of shortest paths being traversed (see [121] for precise formulae). Since all three of these closeness centralities take into account only the length of these shortest paths (not their uniqueness, as discussed below), we argue that they are not significantly distorted by the assumption of identical external connectivity. If we reconfigure an underlying higher-resolution network G_{hr} to conform with n.s.i assumptions, by removing some external links from some nodes in a group, and replacing them with external links identical to other external links in the group, this will not cause a significant change in the closeness centrality of any of the nodes in the cluster. So long as full internal connectivity holds, any of the new external links assigned to a node in order to satisfy external connectivity would have had a path length of 2, and now have this path length reduced to 1. Due to the fact that the network is constructed via geometric proximity, we expect that any of the broken links will likely still be connected by a path length of 2. Therefore, the assumption of identical external connectivity will only cause some minor changes in the average path lengths starting from a node, and these changes will to some extent cancel out.

In contrast, an incorrect assumption of identical external connectivity may cause higher-weighted nodes to have a lower betweenness centrality. This is because unlike closeness centrality, betweenness depends also on the uniqueness of shortest paths passing through a node. In calculating betweenness centrality, the contribution of a shortest path from a to b which passes through v is always scaled by the inverse of the total number of shortest paths that pass from a to b. While nearby nodes in a higher-resolution network may have a diverse and distinct set of shortest paths passing through each of them, n.s.i forces all nodes in the same group to lie along an identical set of shortest paths (see the identical external purple links for $b1, b2, b3$ in Figure A.3). Thus n.s.i betweenness centrality may be systematically lower for larger-weight nodes, if the assumption of identical external connectivity does not actually hold.

Finally, we note that improvements in the method of construction of our node-weighted network could lead to increased accuracy of the n.s.i assumption of identical external connectivity. For example, we could generate a node-weighted network by simulating a large sample of MC events and clustering events together into a larger weighted event only where a tight cluster of events naturally exists. This would ensure that a weighted node accurately represents the density of a very local area of the kinematic space. Efficient algorithms such as k-means clustering could be used to generate a set of representative weighted nodes from a larger sample. Methods considered in [107] to find the smallest representative set of events could also be applied. Building a more robust node-weighted network in this way may significantly increase the power of this analysis by permitting the use of important topological network measures omitted here, in particular the betweenness centrality and clustering coefficient.

Appendix B

Event weights in sliced samples

B.1 Stop example: slicing procedure

The slicing procedure is the primary cause of disparate event weights in Chapters 5 and 6, and is therefore responsible for a large part of the discussion in Appendix A. The stop example in Chapter 6 in particular comprises event samples where the difference in weight between slices is large, where the selections in H_T alter the cross-sections more than the selections in $\hat{p}_\mathrm{T}$ for the electroweakino production example. The maximum difference between weights in the electroweakino case study is therefore somewhat lower and less problematic in network calculations. In this appendix, the slices selected for the stop example are justified and their application is described.

When abnormally high bins of H_T are selected at parton-level in MC simulations with `Pythia 8`, events are generated from a kinematic distribution biased towards rare energetic particles, so that the probability of high momentum final state particles is greater than we observe in nature. We must then add small weights to the rare events, of which there now exists too great a number, and reproduce the expected numbers of events in measurements at the LHC. The two weight factors associated with slicing are the slice cross-section factor and the normalisation factor, described in Section 4.2.3. The product of these two factors is labelled *weight factor* in Table B.1, which contains weights associated with slicing applied to the signal and background samples in the stop example. These slices introduce the maximum weight difference between any two slices from the same sample. If the n.s.i calculations are biased by large differences in node weights, then the difference between these slices is likely to be responsible for the greatest contribution to this bias.

Table B.1 The two slices with maximally different weights from the stop and top simulated event samples used in the stop case study in Chapter 6, where N_{events} is the number of events in the slice after preselection.

	Slice (H_T, GeV)	$N_{\textbf{events}}$	**Weight factor**
Stop	300-400	1600	1.40×10^{-1}
	>2500	20	2.68×10^{-3}
Top	0-150	200	1.31×10^{-1}
	>2000	60	1.30×10^{-3}

The combination of all slices comprising the full set of stop or top events in selected kinematic space produces a smoother distribution in not only the sliced variable H_T, but also the other kinematic variables on which the network analysis is dependent. Figure B.1 shows the distribution of events by slice in the seven lowest-H_T slices for the H_T variable (upper plot), where the dark blue entries represent the combination of re-weighted events from the event slices and the black distribution is an inclusive sample. The plots are shown as a guide to convey the effect of slicing on the contents of the kinematic distributions for variables underlying the stop networks. They build upon the information presented in Figures 4.5 and 4.6, which revealed that smoothness increases significantly in the $\hat{p}_T$ distribution when electroweakino events are generated in $\hat{p}_T$ slices.

The sliced events produce realistic distributions after re-weighting not only in the variable that was binned at generator-level, but also in the other kinematic variables. Since events containing many highly energetic particles are likely to belong to the H_T tail plus the tail regions of several other variables, including jet momenta and transverse mass, the rare kinematic regions are also populated in other variables. For example, the same set of events sliced in H_T is present in the lower plot from Figure B.1, showing the transverse momentum of the first jet p_T^{j1}. The sample of re-weighted events in dark blue is composed again of the various coloured entries and compares with the inclusive entries in black.

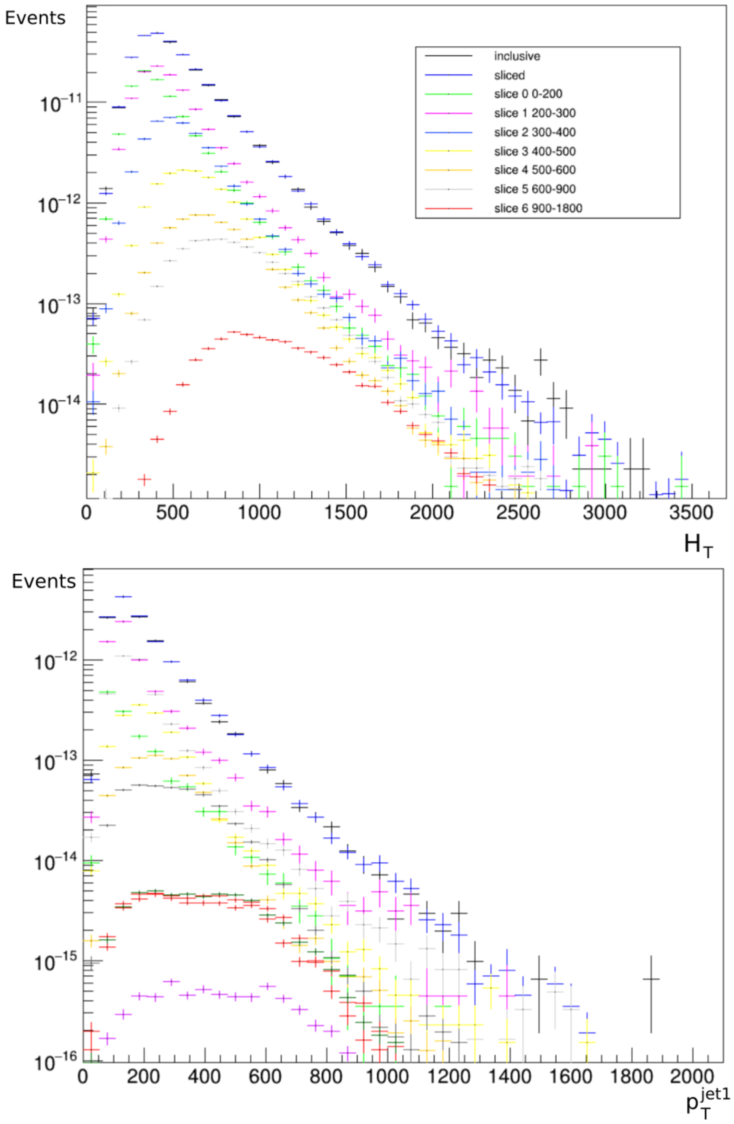
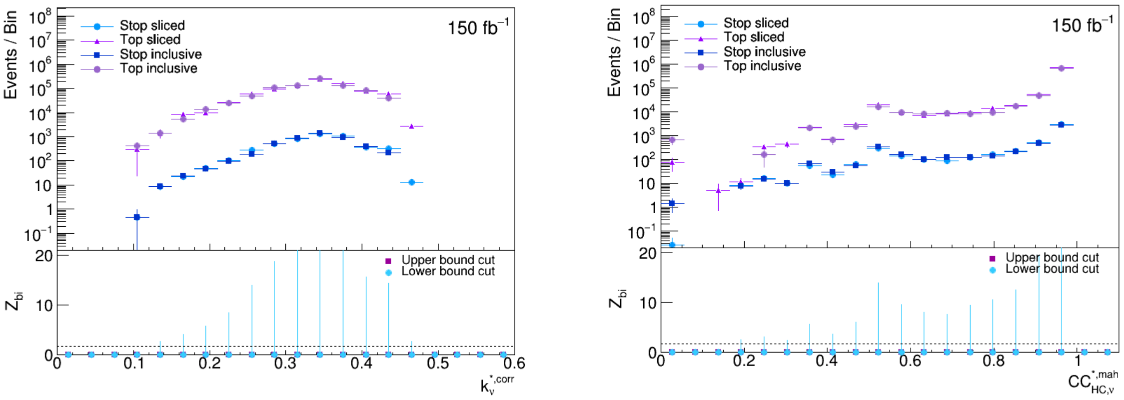


Fig. B.1 Distributions of events across variables H_T and $p_{\mathrm{T}}^{\mathrm{j1}}$ for sample slices in H_T.

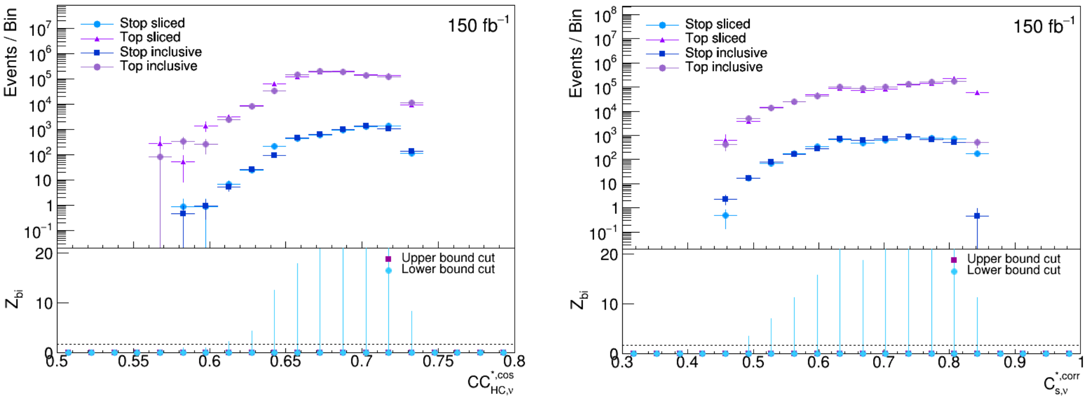
These slices are deemed sufficient for reliably modelling the kinematic tails so that the rare regions are accurately represented in networks of limited size. The computational intensity of network calculations limits the event count to 10,000 signal and 10,000 background events, which contains a significant portion of re-weighted kinematic tail events as a result of the slicing procedure. After constructing the networks, n.s.i local network metrics are calculated for both inclusive and sliced samples generated from the search region defined by preselection criteria in Chapter 6. Their shape properties should agree to within the statistical deviations expected from low event counts in the tails, if weight disparities introduced by slicing has little effect on n.s.i calculations. The inclusive and sliced distributions are compared in Figures B.2 and B.3.

The regions showing least agreement in Figure B.3 between sliced and inclusive distributions are those with fewer events in the inclusive samples, particularly in the tail regions where the inclusive distributions are unpopulated. For example, the tails of the betweenness centrality distributions for the inclusive background samples are negligibly populated, which is again true in the Euclidean local and local Soffer clustering distributions, where the sliced networks better represent these regions by filling out the tails. We see evidence of agreement between inclusive and sliced distributions not only in the network metrics which are considered reliable following the discussion in Section 4.3.1 and Appendix A, but also in the metrics which should be treated with caution awaiting further investigation.



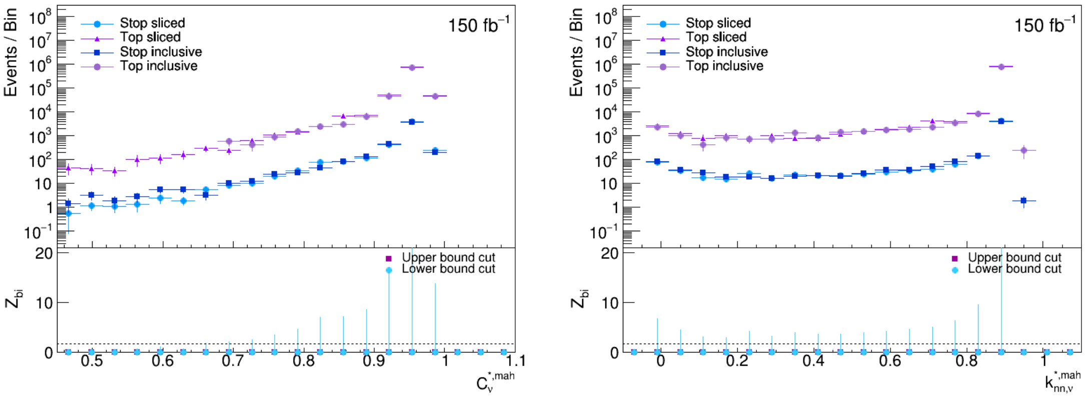
(a) $k_{v}^{*,\mathrm{corr}}$

(b) $CC_{HC,v}^{*,\mathrm{mah}}$



(c) $CC_{HC,v}^{*,\mathrm{cos}}$

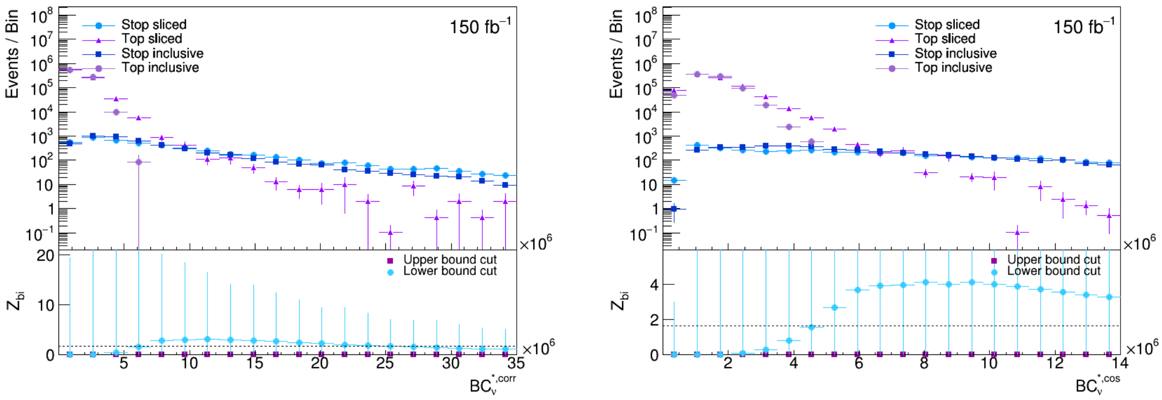
(d) $C_{s,v}^{*,\mathrm{corr}}$



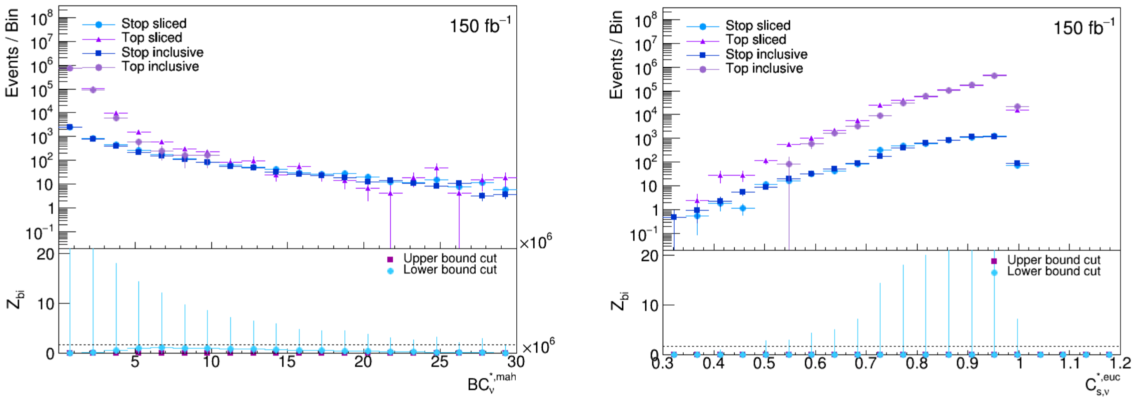
(e) $C_{v}^{*,\mathrm{mah}}$

(f) $k_{nn,v}^{*,\mathrm{mah}}$

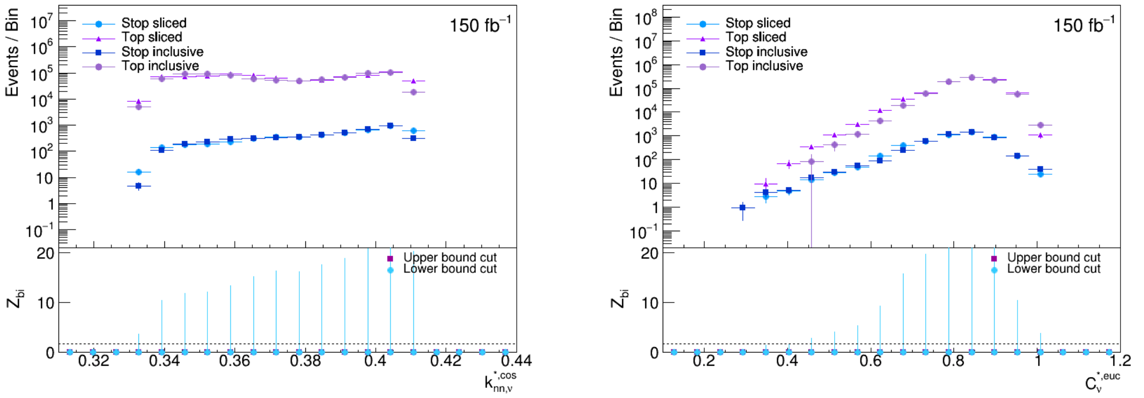
Fig. B.2 Event rates as functions of the network variables for the stop simplified model example. Events in the overflow bin are not shown in the distribution but are included in the Z_{bi} calculation.



(a) $BC_v^{*,\mathrm{corr}}$ (b) $BC_v^{*,\mathrm{cos}}$



(c) $BC_v^{*,\mathrm{mah}}$ (d) $C_{s,v}^{*,\mathrm{euc}}$



(e) $k_{nn,v}^{*,\mathrm{cos}}$ (f) $C_v^{*,\mathrm{euc}}$

Fig. B.3 Event rates as functions of the network variables for the stop simplified model example that show the most difference between the signal and the background either when alone or when combined in cuts with other variables. Events in the overflow bin are not shown in the distribution but are included in the Z_{bi} calculation.

References

[1] A. Mullin, S. Nicholls, H. Pacey, M. Parker, M. White, and S. Williams. Does SUSY have friends? A new approach for LHC event analysis. *JHEP*, 02:160, 2021. doi: 10.1007/JHEP02(2021)160.

[2] E. Noether. Invariante variationsprobleme. *Nachrichten von der Gesellschaft der Wis-senschaften zu Gottingen, Mathematisch-Physikalische Klasse*, pages 235–257, 1918.

[3] Sandbox Studio. The Standard Model of particle physics, 07 2015. Fermilab/SLAC Publication.

[4] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716(1207.7214):1–29, 09 2012.

[5] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716(1207.7235):30–61, 09 2012.

[6] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19(21):1264–1266, 1967. doi: 10.1103/PhysRevLett.19.1264.

[7] S.L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22(4):579–588, 1961. doi: 10.1016/0029-5582(61)90469-2.

[8] C.N. Yang and R.L. Mills. Conservation of Isotopic Spin and Isotopic Gauge Invariance. *Phys. Rev.*, 96(1):191, 1954. doi: 10.1103/PhysRev.96.191.

[9] J. Goldstone, A. Salam, and S. Weinberg. Broken Symmetries. *Phys. Rev.*, 127(9): 965–970, 1962. doi: 10.1103/PhysRev.127.965.

[10] C. Quigg. Spontaneous Symmetry Breaking as a Basis of Particle Mass. *Rept. Prog. Phys.*, 70(7):1019–1054, 2007. doi: 10.1088/0034-4885/70/7/R01.

[11] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, (10):531–533, 1963.

[12] M. Kobayashi and T. Maskawa. CP violation in the renormalisable theory of weak interaction. *Prog. Theor. Phys.*, (49):652–657, 1973.

[13] C. Colangelo, S. Durr, A. Juttner, L. Lellouch, H. Leutwyler, V. Lubicz, S. Necco, C.T. Sachrajda, S. Simula, A. Vladikas, U. Wenger, and H. Wittig. Review of lattice results concerning low energy particle physics. *Eur. Phys. J. C*, (71:1675), 2010.

[14] Z. Maki, M. Nakagawa, and S. Sakata. Remarks on the unified model of elementary particles. *Prog. Theor. Phys.*, (28):870–880, 1962.

[15] LHCb Collaboration. Observation of CP violation in charm decays. *Phys. Rev. Lett.*, (211803), March 2019.

[16] O. Buchmueller, C. Doglioni, and L.T. Wang. Search for dark matter at colliders. *Nature Phys.*, 13(3):217–223, 2017. doi: 10.1038/nphys4054.

[17] B. Lillard and T.M.P. Tait. A High Quality Composite Axion. *JHEP*, 11:199, 2018. doi: 10.1007/JHEP11(2018)199.

[18] N. Haba, H. Ishida, R. Takahashi, and Y. Yamaguchi. Hierarchy problem, gauge coupling unification at the Planck scale, and vacuum stability. *Nucl. Phys. B*, 900(C): 244–258, 2015. doi: 10.1016/0029-5582(61)90469-2.

[19] P. Grange, P.F. Mathio, B. Mutet, and E. Werner. Aspects of fine-tuning of the Higgs mass within finite field theories. *Phys. Rev. D*, 88(125015), Dec 2013.

[20] S.P. Martin. A Supersymmetry Primer, 1997. doi: 10.1142/9789812839657.0001.

[21] J.S. Ellis, Hagelin, D.V. Nanopoulos, K.A. Oliver, and M. Srednicki. Supersymmetric relics from the Big Bang. *Nucl. Phys. B*, (238):453–76, 1984.

[22] N. Sakai. Naturalness in Supersymmetric GUTs. *Zeit. Phys. C*, (11):153, 1981.

[23] S. Cassel, D.M. Ghilencea, S. Kraml, A. Lessa, and G.G. Ross. Fine-tuning implications for complementary dark matter and LHC SUSY searches. *J. High Energy Phys.*, 5(120):1–17, 2011. doi: 10.1007/JHEP05(2011)120.

[24] S. Dimopoulos, S. Raby, and F. Wilczek. Supersymmetry and the Scale of Unification. *Phys. Rev. D*, (24):1681–83, 1981.

[25] B. de Wit. Supergravity, 12 2002. Lecture notes: Les Houches Summer School.

[26] I.J.R. Aitchison. Supersymmetry and the MSSM: An elementary introduction, 2004. doi: 0505105v1.

[27] A. Djouadi et al. The Minimal supersymmetric Standard Model: Group summary report. In *GDR (Groupement De Recherche) - Supersymetrie*, 12 1998.

[28] A. Brignole, Luis E. Ibanez, and C. Munoz. Towards a theory of soft terms for the supersymmetric Standard Model. *Nucl. Phys. B*, 422:125–171, 1994. doi: 10.1016/0550-3213(94)00068-9. [Erratum: Nucl.Phys.B 436, 747–748 (1995)].

[29] E. Hardy. Is natural SUSY natural? *J. High Energy Phys.*, 2013(10):1–30, 2013. doi: 10.1007/JHEP10(2013)133.

[30] M. Papucci, J.T. Ruderman, and A. Weiler. Natural SUSY endures. October 2011.

[31] D. Croon, T.E. Gonzalo, L. Graf, N. Košnik, and G. White. GUT Physics in the era of the LHC. *Front. in Phys.*, 7:76, 2019. doi: 10.3389/fphy.2019.00076.

[32] Y. Muramatsu and Y. Shigekami. SUSY threshold corrections to quark and lepton mixing inspired bu SO(10) GUT models. *J. High Energy Phys.*, 2(103):1–24, 02 2020. doi: 10.1007/JHEP02(2020)103.

[33] S. Amoroso, ATLAS, and CMS Collaborations. Precision measurements at the LHC. Proceedings of Science, 14th International Symposium on Radiative Corrections (RADCOR2019), February 2020.

[34] C. Gütschow and Z. Marshall. Setting limits on supersymmetry using simplified models, 2012.

[35] G. Aad et al. Comparison between simulated and observed LHC beam backgrounds in the ATLAS experiment at $E_{beam} = 4$ TeV. *JINST*, (13), 2018. doi: 10.1088/1748-0221/13/12/P12006.

[36] B. Stelzer. Run-2 Physics Results from ATLAS at the LHC. The ATLAS Collaboration, June 2019.

[37] G. Aad et al. Luminosity public results run 2, 2020. URL http:https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2#Luminosity_summary_plots_for_201. 2020-11-20.

[38] M.D. Schwartz. TASI Lectures on Collider Physics. Harvard University, Sep 2017.

[39] B. Salvachua. Overview of proton-proton physics during Run II. CERN, 2019. Evian Les Bains, France, 7-14.

[40] S. Fartoukh and N. Karastathis. First report to the LMC on Run III machine configurations (p-p runs). CR3 working group, 2019.

[41] C. Borschensky, M. Krämer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn, and X. Portell. Squark and gluino production cross sections in p p collisions at $\sqrt{s} = 13$, 14, 33 and 100 TeV. *The European Physical Journal C*, 74(12), Dec 2014. ISSN 1434-6052. doi: 10.1140/epjc/s10052-014-3174-y. URL http://dx.doi.org/10.1140/epjc/s10052-014-3174-y.

[42] A. Abada, M. Abbrescia, et al. HE-LHC: The High-Energy Large Hadron Collider. *Eur. Phys. J. Spec. Topics*, 4(228):1109–1382, 2019. doi: 10.1140/epjst/e2019-900088-6.

[43] H. Baer, V. Barger, J.S. Gainer, P. Huang, M. Savoy, H. Serve, and X. Tata. What hadron collider is required to discover or falsify natural supersymmetry? *Phys. Lett. B*, 774, 2017. doi: https://doi.org/10.1016/j.physletb.2017.09.086.

[44] P. Athron et al. A global fit of the MSSM with GAMBIT. *Eur. Phys. J.*, C77(12):879, 2017. doi: 10.1140/epjc/s10052-017-5196-8.

[45] P. Athron et al. GAMBIT: The Global and Modular Beyond-the-Standard-Model Inference Tool. *Eur. Phys. J.*, C77(11):784, 2017. doi: 10.1140/epjc/s10052-017-5513-2, 10.1140/epjc/s10052-017-5321-8. [Addendum: Eur. Phys. J.C78,no.2,98(2018)].

[46] P. Athron et al. Global fits of GUT-scale SUSY models with GAMBIT. *Eur. Phys. J.*, C77(12):824, 2017. doi: 10.1140/epjc/s10052-017-5167-0.

[47] E. Bagnaschi et al. Likelihood Analysis of the pMSSM11 in Light of LHC 13-TeV Data. *Eur. Phys. J.*, C78(3):256, 2018. doi: 10.1140/epjc/s10052-018-5697-0.

[48] J. C. Costa et al. Likelihood Analysis of the Sub-GUT MSSM in Light of LHC 13-TeV Data. *Eur. Phys. J.*, C78(2):158, 2018. doi: 10.1140/epjc/s10052-018-5633-3.

[49] Standard Model Summary Plots Spring 2020. Technical Report ATL-PHYS-PUB-2020-010, CERN, Geneva, May 2020. URL http://cds.cern.ch/record/2718937.

[50] G. Aad et al. Summary plots from the ATLAS Standard Model physics group, 2020.

[51] D. Froidevaux. *Integration of detectors into a large experiment: Examples from ATLAS and CMS*, pages 693–783. Springer, Cham., September 2020.

[52] The ATLAS collaboration. Operation and performance of the ATLAS semiconductor tracker. *Journal of Instrumentation*, 9(08):P08009–P08009, Aug 2014. ISSN 1748-0221. doi: 10.1088/1748-0221/9/08/p08009. URL http://dx.doi.org/10.1088/1748-0221/9/08/P08009.

[53] The ATLAS TRT collaboration. The ATLAS TRT end-cap detectors. *Journal of Instrumentation*, 3(10):P10003–P10003, oct 2008. doi: 10.1088/1748-0221/3/10/p10003. URL https://doi.org/10.1088%2F1748-0221%2F3%2F10%2Fp10003.

[54] The ATLAS TDAQ Collaboration. The ATLAS Data Acquisition and High Level Trigger system. *JINST*, 11:06008, June 2016. doi: https://doi.org/10.1088/1748-0221/11/06/P06008.

[55] The ATLAS Collaboration. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics, 2008.

[56] D. Saunders. Data reconstruction in modern particle physics. CERN School of Computing, 2016.

[57] D. Kar. *Experimental particle physics: Analysis objects*, pages 3–1 – 3–32. IOP Publishing, August 2019.

[58] K. Kaneta, S. Kang, and H. S. Lee. Diphoton channel at the lhc experiments to find a hint for a new heavy gauge boson. *Int. J. Mod. Phys. A*, 31(27):1650159, 2015.

[59] G. Kasieczka et al. The LHC Olympics 2020: A Community Challenge for Anomaly Detection in High Energy Physics. 1 2021.

[60] Aachen-Annecy-Birmingham-CERN-Helsinki-London(QMC)-Paris(CdF)-Riverside-Rome-Rutherford-Saclay(CEN)-Vienna Collaboration. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV. *Phys. Lett. B*, (122):103–116, January 1983.

[61] C.G. Lester and D.J. Summers. Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders. *Phys. Lett.*, B463:99–103, 1999. doi: 10.1016/S0370-2693(99)00945-4.

[62] C.G. Lester and B. Nachman. Bisection-based asymmetric M_{T2} computation: a higher precision calculator than existing symmetric methods. *JHEP*, 03:100, 2015. doi: 10.1007/JHEP03(2015)100.

[63] A. Barr, C.G. Lester, and P. Stephens. m_{T2}: the truth behind the glamour. *J. Phys. G*, (29):2343–2363, April 2003.

[64] P. Jackson and C. Rogan. Recursive jigsaw reconstruction: HEP event analysis in the presence of kinematic and combinatoric ambiguities. *Physical Review D*, 96(11), Dec 2017. ISSN 2470-0029. doi: 10.1103/physrevd.96.112007. URL http://dx.doi.org/10.1103/PhysRevD.96.112007.

[65] P. Jackson, C. Rogan, and M. Santoni. Sparticles in motion: Analyzing compressed SUSY scenarios with a new method of event reconstruction. *Phys. Rev. D*, 95(3): 035031, 2017. doi: 10.1103/PhysRevD.95.035031.

[66] C. Rogan. Restframes software library for HEP event analysis, 2016. URL http://www.RestFrames.com.

[67] J. Alwall, P. Schuster, and N. Toro. Simplified Models for a First Characterization of New Physics at the LHC. *Phys. Rev.*, D79:075020, 2009. doi: 10.1103/PhysRevD.79.075020.

[68] S. Agostinelli et al. Geant4—a simulation toolkit. *Nucl. Inst. Meth. Phys. Res. Sec. A: Accelerators, Spectrometers, Detectors and Associated Equipment*.

[69] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, D.K. Abhayasinghe, S.H. Abidi, O.S. Abouzeid, N.L. Abraham, H. Abramowicz, et al. Observation of $H \to bb^{-}$ decays and VH production with the ATLAS detector. *Physics Letters B*, 786: 59–86, Nov 2018. ISSN 0370-2693. doi: 10.1016/j.physletb.2018.09.013. URL http://dx.doi.org/10.1016/j.physletb.2018.09.013.

[70] ATLAS Collaboration. ATLAS Event Displays: Higgs boson decaying to two b-quarks. General Photo, Aug 2018. URL https://cds.cern.ch/record/2636049.

[71] T. Sjöstrand, S. Mrenna, and P. Skands. A Brief Introduction to PYTHIA 8.1. 2007. doi: 10.1016/j.cpc.2008.01.036.

[72] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi. DELPHES 3, A modular framework for fast simulation of a generic collider experiment. *JHEP*, 02:057, 2014. doi: 10.1007/JHEP02(2014)057.

[73] P.Z. Skands. QCD for collider physics, 2010. European School of High Energy Physics (ESHEP 2010).

[74] Georges Aad et al. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J. C*, 73(3):2304, 2013. doi: 10.1140/epjc/s10052-013-2304-2.

[75] G. Aad et al. The ATLAS Simulation Infrastructure. *Eur. Phys. J. C*, 70:823–874, 2010. doi: 10.1140/epjc/s10052-010-1429-9.

[76] P.Z. Skands. Tuning Monte Carlo Generators: The Perugia Tunes. *Phys. Rev. D*, 82: 074018, 2010. doi: 10.1103/PhysRevD.82.074018.

[77] J. Neyman and E.S. Pearson. IX. on the problem of the most efficient tests of statistical hypotheses. *Phil. Trans. R. Soc. Lond. A*, (231):694–706, February 1933.

[78] P. Athron et al. Combined collider constraints on neutralinos and charginos. *Eur. Phys. J.*, C79(5):395, 2019. doi: 10.1140/epjc/s10052-019-6837-x.

[79] G. Aad et al. Searches for squarks and gluinos with the ATLAS detector. *Phys. Scr.*, (95):094003, December 2020.

[80] A.J. Larkoski, I. Moult, and B. Nachman. Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning. *Phys. Rept.*, 841:1–63, 2020. doi: 10.1016/j.physrep.2019.11.001.

[81] G. Kasieczka, S. Marzani, G. Soyez, and G. Stagnitto. Towards Machine Learning Analytics for Jet Substructure. *JHEP*, 09:195, 2020. doi: 10.1007/JHEP09(2020)195.

[82] K. Datta, A. Larkoski, and B. Nachman. Automating the Construction of Jet Observables with Machine Learning. *Phys. Rev. D*, 100(9):095016, 2019. doi: 10.1103/PhysRevD.100.095016.

[83] P. Thaprasop, K. Zhou, J. Steinheimer, and C. Herold. Unsupervised Outlier Detection in Heavy-Ion Collisions. *Phys. Scripta*, 96(6):064003, 2021. doi: 10.1088/1402-4896/abf214.

[84] Deep Learning for Pion Identification and Energy Calibration with the ATLAS Detector. 7 2020.

[85] J. Shlomi, P. Battaglia, and J.R. Vlimant. Graph neural networks in particle physics. *Machine Learning: Science and Technology*, October 2020. URL https://arxiv.org/abs/2007.13681.

[86] A.A. Pol, V. Berger, G. Cerminara, C. Germain, and M. Pierini. Anomaly detection with conditional variational autoencoders. *IEEE ICMLA*, October 2020.

[87] J. Raine. Deep generative models for fast shower simulation with the atlas experiment. 24th International Conference on Computing in High Energy and Nuclear Physics, October 2019.

[88] O. Cerri, T.Q. Nguyen, M. Pierini, M. Spiropulu, and J.R. Vlimant. Variational autoencoders for new physics mining at the large hadron colliders. *J. High Energy Phys.*, (36), June 2019.

[89] J. Elmsheuser, C. Anastopoulos, J. Boyd, J. Catmore, and H. Gray. Evolution of the ATLAS analysis model for Run-3 and prospects for HL-LHC. 24th International Conference on Computing in High Energy and Nuclear Physics, 2020.

[90] M.E.J. Newman. The structure and function of complex networks. *SIAM rev*, 45 (2):167–256, 2003. doi: 10.1137/S003614450342480. URL https://doi.org/10.1137/S003614450342480.

[91] G. Nikulin, C. Jones, F. Giorgi, G. Asrar, M. Buchner, R. Cerezo-Mota, O.B. Christensen, M. Déqué, J. Fernandez, A. Hänsler, E. van Meijgaard, P. Samuelsson, M.B. Sylla, and L. Sushama. Precipitation climatology in an ensemble of CORDEX-Africa regional climate simulations. *Journal of Climate*, 25:6057–6078, 2012.

[92] J.H. Feldhoff, S. Lange, J. Volkholz, J.F. Donges, J. Kurths, and F.W. Gerstengarbe. Complex networks for climate model evaluation with application to statistical versus dynamical modeling of South American climate. *Climate Dynamics*, 85(44):1567–1581, May 2015. doi: 10.1007/s00382-014-2182-9. URL https://doi.org/10.1007/s00382-014-2182-9.

[93] R. de Regt, S. Apunevych, C. von Ferber, Y. Holovatch, and B. Novosyadlyj. Network analysis of the COSMOS galaxy field. *Mon. Not. Roy. Astron. Soc.*, 477(3):4738–4748, Mar 2018. doi: 10.1093.mnras.sty801.

[94] M. Wiedermann, J. F. Donges, J. Heitzig, and J. Kurths. Node-weighted interacting network measures improve the representation of real-world complex systems. *Europhysics Letters 102*, 102(1):28007, Jan 2013. doi: 10.1209/0295-5075/102/28007.

[95] C. Zhou, L. Zemanová, G. Zamora, C.C. Hilgetag, and J. Kurths. Hierarchical organisation unveiled by functional connectivity in complex brain networks. *Phys. Rev. Lett.*, 97(23):238103, Dec 2006. doi: 10.1103/PhysRevLett.97.238103.

[96] D.J. Watts and S.H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393(2):440–442, Jun 1998. doi: 10.1038/30918. URL https://doi.org/10.1038/30918.

[97] P. Erdős and A. Rényi. On random graphs I. *Publicationes Mathematicae*, 5(6): 290–297, 1959. URL https://www.renyi.hu/~p_erdos/1959-11.pdf.

[98] S. Hong and A. Dey. Network analysis of cosmic structures: network centrality and topological environment. *Mon. Not. Roy. Astron. Soc.*, 450, 2015.

[99] S. Hong, B. Coutinho, A. Dey, A.L. Barabási, M. Vogelsberger, L. Hernquist, and K. Gebhardt. Discriminating Topology in Galaxy Distributions using Network Analysis. *Mon. Not. Roy. Astron. Soc.*, 459(3):2690–2700, 2016. doi: 10.1093/mnras/stw803.

[100] S. Hong, D. Jeong, H.S. Hwang, J. Kim, S.E. Hong, C. Park, A. Dey, M. Milosavljevic, K. Gebhardt, and K.S. Lee. Constraining cosmology with big data statistics of cosmological graphs. *Mon. Not. Roy. Astron. Soc.*, 493, 2020.

[101] M. Tsizh, B. Novosyadlyj, and Y. Holovatch. Large-scale structures in the ΛCDM Universe: network analysis and machine learning. *Mon. Not. Roy. Astron. Soc.*, 495, 2020.

[102] A.D. Broido and A. Clauset. Scale-free networks are rare. *Nature Communications*, 10(1):1017, Mar 2019. doi: https://doi.org/10.1038/s41467-019-08746-5.

[103] R. Khanin and E. Wit. How scale-free are biological networks. *J. Comp. Biol.*, 13(1): 810–818, Mar 2006. doi: 10.1089/cmb.2006.13.810.

[104] G. Lima-Mendez and J. van Helden. The powerful law of the power law and other myths in network biology. *Mol. Biosyst.*, 5(12):1482–93, Dec 2009. doi: 10.1039/b908681a.

[105] L.A. Adamic, B.A. Huberman, A.L Barabási, R. Albert, H. Jeong, and G. Bianconi. Power-law distribution of the World Wide Web. *Science*, 287(5461):2115, Mar 2000. doi: 10.1126/science.287.5461.2115a.

[106] P.T. Komiske, E.M. Metodiev, and J. Thaler. Metric Space of Collider Events. *Phys. Rev. Lett.*, 123(4):041801, 2019. doi: 10.1103/PhysRevLett.123.041801.

[107] P.T. Komiske, R. Mastandrea, E.M. Metodiev, P. Naik, and J. Thaler. Exploring the Space of Jets with CMS Open Data. 2019.

[108] E.A. Moreno, T.Q. Nguyen, J.R. Vlimant, O. Cerri, H.B. Newman, A. Periwal, M. Spiropulu, J.M. Duarte, and M. Pierini. Interaction Networks for the Identification of Boosted $H \to b\bar{b}$ Decays. 2019.

[109] E.A. Moreno, O. Cerri, J.M. Duarte, H.B. Newman, T.Q. Nguyen, A. Periwal, M. Pierini, A. Serikova, M. Spiropulu, and J.R. Vlimant. JEDI-net: a jet identification algorithm based on interaction networks. 2019.

[110] H. Qu and L. Gouskos. ParticleNet: Jet Tagging via Particle Clouds. arXiv:1902.08570, 2019.

[111] I. Henrion, J. Brehmer, J. Bruna, K. Cho, K. Cranmer, G. Louppe, and G. Rochette. Neural message passing for jet physics. 2017.

[112] M. Abdughani, J. Ren, L. Wu, and J.M. Yang. Probing stop pair production at the LHC with graph neural networks. *JHEP*, 08:055, 2019. doi: 10.1007/JHEP08(2019)055.

[113] N. Choma, F. Monti, L. Gerhardt, T. Palczewski, Z. Ronaghi, Prabhat, W. Bhimji, M.M. Bronstein, S.R. Klein, and J. Bruna. Graph Neural Networks for IceCube Signal Classification. 2018.

[114] S. Farrell et al. Novel deep learning methods for track reconstruction. In *4th International Workshop Connecting The Dots 2018 (CTD2018) Seattle, Washington, USA, March 20-22, 2018*, 2018. URL http://lss.fnal.gov/archive/2018/conf/fermilab-conf-18-598-cd.pdf.

[115] J. Arjona Martínez, O. Cerri, M. Pierini, M. Spiropulu, and J.R. Vlimant. Pileup mitigation at the Large Hadron Collider with graph neural networks. *Eur. Phys. J. Plus*, 134(7):333, 2019. doi: 10.1140/epjp/i2019-12710-3.

[116] P.T. Komiske, E.M. Metodiev, and J. Thaler. Cutting Multiparticle Correlators Down to Size. 2019.

[117] S.R. Qasim, J. Kieseler, Y. Iiyama, and M. Pierini. Learning representations of irregular particle-detector geometry with distance-weighted graph networks. *Eur. Phys. J.*, C79(7):608, 2019. doi: 10.1140/epjc/s10052-019-7113-9.

[118] E. Deza and M.M. Deza, editors. *Dictionary of Distances*. Elsevier, Amsterdam, 2006. ISBN 978-0-444-52087-6. doi: https://doi.org/10.1016/B978-044452087-6/50000-7. URL http://www.sciencedirect.com/science/article/pii/B9780444520876500007.

[119] T. Opsahl, V. Colizza, P. Panzarasa, and J.J. Ramasco. Prominence and control: The weighted rich-club effect. *Phys. Rev. Lett.*, 101(16):168702, Oct 2008. doi: 10.1103/PhysRevLett.101.168702.

[120] C. Chekuri, A. Ene, and A. Vakilian. Node-weighted Network Design in Planar and Minor-closed Families of Graphs. Oct 2019.

[121] J. Heitzig, J. F. Donges, Y. Zou, N. Marwan, and J. Kurths. Node-weighted measures for complex networks with spatially embedded, sampled, or differently sized nodes. *European Physical Journal B*, 85(1):38, Jan 2012. doi: 10.1140/epjb/e2011-20678-7.

[122] J.F. Donges et al. Unified functional network and nonlinear time series analysis for complex systems science: The pyunicorn package. *Chaos*, 25:113101, 2015. doi: 10.1063/1.4934554.

[123] R. Albert and A. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, 2002. doi: 10.1103/RevModPhys.74.47.

[124] M. Gerlach and E.G. Altmann. Testing statistical laws in complex systems. *Phys. Rev. Lett.*, 122(16):168301, 2019. doi: 10.1103/PhysRevLett.122.168301.

[125] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using networkx. In Gaël Varoquaux, Travis Vaught, and Jarrod Millman, editors, *Proceedings of the 7th Python in Science Conference*, pages 11 – 15, Pasadena, CA USA, 2008.

[126] S.N. Soffer and A. Vázquez. Network clustering coefficient without degree-correlation biases. *Phys. Rev. E*, 71:057101, May 2005. doi: 10.1103/PhysRevE.71.057101. URL https://link.aps.org/doi/10.1103/PhysRevE.71.057101.

[127] K. Cranmer. Statistical challenges for searches for new physics at the LHC. *Statistical Problems in Particle Physics, Astrophysics and Cosmology*, May 2006. doi: 10.1142/9781860948985_0026. URL http://dx.doi.org/10.1142/9781860948985_0026.

[128] R.D. Cousins, J.T. Linnemann, and J. Tucker. Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a poisson process. *ucl. Inst. Meth. Phys. Res. Sec. A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 595 (2):480–501, Oct 2008. ISSN 0168-9002. doi: 10.1016/j.nima.2008.07.086. URL http://dx.doi.org/10.1016/j.nima.2008.07.086.

[129] J.T. Linnemann. Measures of Significance in HEP and Astrophysics, 2003.

[130] L. Moneta, K. Belasco, K. Cranmer, S. Kreiss, A. lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf. The roostats project, 2010.

[131] G. Aad et al. Search for chargino-neutralino production with mass splittings near the electroweak scale in three-lepton final states in $\sqrt{s}$=13 TeV pp collisions with the ATLAS detector. *Phys. Rev. D*, 101(7):072001, 2020. doi: 10.1103/PhysRevD.101.072001.

[132] T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C.O. Rasmussen, and P.Z. Skands. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015. doi: 10.1016/j.cpc.2015.01.024.

[133] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, Pavel M. Nadolsky, and W.K. Tung. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002. doi: 10.1088/1126-6708/2002/07/012.

[134] M. Selvaggi. DELPHES 3: A modular framework for fast-simulation of generic collider experiments. *J. Phys. Conf. Ser.*, 523:012033, 2014. doi: 10.1088/1742-6596/523/1/012033.

[135] A. Mertens. New features in Delphes 3. *J. Phys. Conf. Ser.*, 608(1):012045, 2015. doi: 10.1088/1742-6596/608/1/012045.

[136] M. Cacciari, G.P. Salam, and G. Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008. doi: 10.1088/1126-6708/2008/04/063.

[137] M. Cacciari, G.P. Salam, and G. Soyez. FastJet User Manual. *Eur. Phys. J. C*, 72: 1896, 2012. doi: 10.1140/epjc/s10052-012-1896-2.

[138] B. Fuks, M. Klasen, D.R. Lamprea, and M. Rothering. Gaugino production in proton-proton collisions at a center-of-mass energy of 8 TeV. *JHEP*, 10:081, 2012. doi: 10.1007/JHEP10(2012)081.

[139] B. Fuks, M. Klasen, D.R. Lamprea, and M. Rothering. Precision predictions for electroweak superpartner production at hadron colliders with RESUMMINO. *Eur. Phys. J. C*, 73:2480, 2013. doi: 10.1140/epjc/s10052-013-2480-0.

[140] Massimiliano Grazzini, Stefan Kallweit, Dirk Rathlev, and Marius Wiesemann. $W^{\pm}Z$ production at hadron colliders in NNLO QCD. *Phys. Lett.*, B761:179–183, 2016. doi: 10.1016/j.physletb.2016.08.017.

[141] M. Aaboud et al. Search for electroweak production of supersymmetric particles in final states with two or three leptons at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector. *Eur. Phys. J. C*, 78(12):995, 2018. doi: 10.1140/epjc/s10052-018-6423-7.

[142] M. Aaboud et al. Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using 36 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector. *JHEP*, 06:108, 2018. doi: 10.1007/JHEP06(2018)108.

[143] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, M. Backes, T. Carli, O. Cohen, A. Christov, D. Dannheim, K. Danielowski, S. Henrot-Versille, M. Jachowski, K. Kraszewski, A. Krasznahorkay Jr., M. Kruk, Y. Mahalalel, R. Ospanov, X. Prudent, A. Robert, D. Schouten, F. Tegenfeldt, A. Voigt, K. Voss, M. Wolter, and A. Zemla. Tmva - toolkit for multivariate data analysis, 2007.

[144] G. Aad et al. Search for a scalar partner of the top quark in the all-hadronic $t\bar{t}$ plus missing transverse momentum final state at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 80(8):737, 2020. doi: 10.1140/epjc/s10052-020-8102-8.

[145] W. Beenakker, C. Borschensky, M. Krämer, A. Kulesza, and E. Laenen. NNLL-fast: predictions for coloured supersymmetric particle production at the LHC with threshold and Coulomb resummation. *JHEP*, 12:133, 2016. doi: 10.1007/JHEP12(2016)133.

[146] W. Beenakker, M. Kramer, T. Plehn, M. Spira, and P. M. Zerwas. Stop production at hadron colliders. *Nucl. Phys.*, B515:3–14, 1998. doi: 10.1016/S0550-3213(98)00014-5.

[147] W. Beenakker, S. Brensing, M. Kramer, A. Kulesza, E. Laenen, and I. Niessen. Supersymmetric top and bottom squark production at hadron colliders. *JHEP*, 08:098, 2010. doi: 10.1007/JHEP08(2010)098.

[148] W. Beenakker, C. Borschensky, R. Heger, M. Krämer, A. Kulesza, and E. Laenen. NNLL resummation for stop pair-production at the LHC. *JHEP*, 05:153, 2016. doi: 10.1007/JHEP05(2016)153.

[149] M. Beneke, P. Falgari, S. Klein, and C. Schwinn. Hadronic top-quark pair production with NNLL threshold resummation. *Nucl. Phys.*, B855:695–741, 2012. doi: 10.1016/j.nuclphysb.2011.10.021.

[150] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason. Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation. *Phys. Lett.*, B710:612–622, 2012. doi: 10.1016/j.physletb.2012.03.013.

[151] P. Bärnreuther, M. Czakon, and A. Mitov. Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$. *Phys. Rev. Lett.*, 109:132001, 2012. doi: 10.1103/PhysRevLett.109.132001.

[152] M. Czakon and A. Mitov. NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels. *JHEP*, 12:054, 2012. doi: 10.1007/JHEP12(2012)054.

[153] M. Czakon and A. Mitov. NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction. *JHEP*, 01:080, 2013. doi: 10.1007/JHEP01(2013)080.

[154] M. Czakon, P. Fiedler, and A. Mitov. Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha\frac{4}{S})$. *Phys. Rev. Lett.*, 110:252004, 2013. doi: 10.1103/PhysRevLett.110.252004.

[155] M. Czakon and A. Mitov. Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders. *Comput. Phys. Commun.*, 185:2930, 2014. doi: 10.1016/j.cpc.2014.06.021.

[156] A.J. Barr, B. Gripaios, and C.G. Lester. Transverse masses and kinematic constraints: from the boundary to the crease. *JHEP*, 11:096, 2009. doi: 10.1088/1126-6708/2009/11/096.

[157] P. Konar, K. Kong, K.T. Matchev, and M. Park. Dark Matter Particle Spectroscopy at the LHC: Generalizing M(T2) to Asymmetric Event Topologies. *JHEP*, 04:086, 2010. doi: 10.1007/JHEP04(2010)086.

[158] Y. Bai, H.C. Cheng, J. Gallicchio, and J. Gu. Stop the Top Background of the Stop Search. *JHEP*, 07:110, 2012. doi: 10.1007/JHEP07(2012)110.

[159] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, D. K. Abhayasinghe, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, H. Abramowicz, and et al. Search for chargino-neutralino production using recursive jigsaw reconstruction in final states with two or three charged leptons in proton-proton collisions at $s = \sqrt{13}$ TeV with the ATLAS detector. *Physical Review D*, 98(9), Nov 2018. ISSN 2470-0029. doi: 10.1103/physrevd.98.092012. URL http://dx.doi.org/10.1103/PhysRevD.98.092012.