# Does SUSY have friends? A new approach for LHC event analysis

**Anna Mullin**, Holly Pacey, Andy Parker, Martin White, Sarah Williams. ArXiv:1912.10625

UNIVERSITY OF CAMBRIDGE

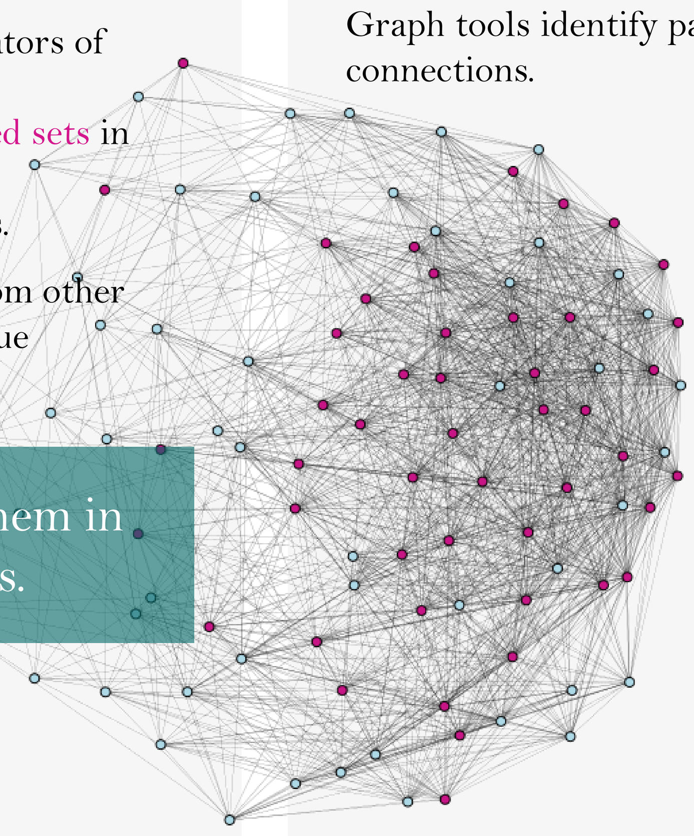
THE UNIVERSITY of ADELAIDE

## Network analysis for BSM searches

Graph network analysis can:

- Treat discrete multi-dimensional data without estimators of continuous functions from theory.
- Encode similarities and differences between unordered sets in pairwise connections.
- Reveal rich relations or interactions between datasets.

Beyond-Standard Model (BSM) events look different from other events in chosen kinematic variables, lending them unique topologies in networks.



## Comparing events instead of treating them in isolation reveals complex structures.

Features of complex graphs:

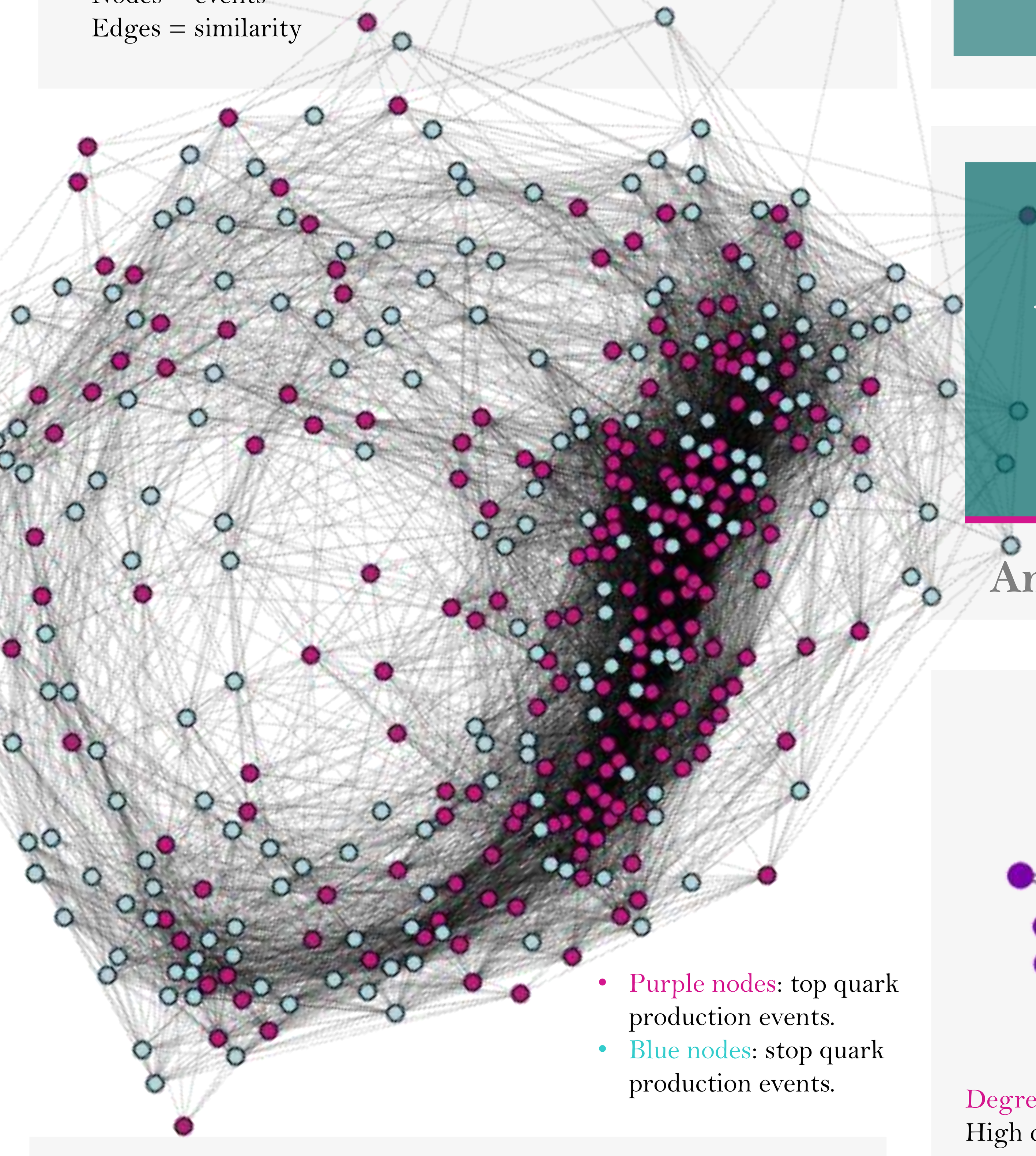
- Clustered vs random
- Central vs fringe
- Filament structures
- Connected paths

We quantify these properties by calculating network metrics, conveniently defined by graph theory.

The network metrics can be local (evaluate per event) or global (a network average). We focus first on local metrics to perform an event-by-event analysis, giving a value for each event like traditional analysis variables. We use simulated LHC collision events.

Network metrics target a range of features, typically selecting unique topologies from BSM events, which are rarer and more complex.

Nodes = events
Edges = similarity



- Purple nodes: top quark production events.
- Blue nodes: stop quark production events.

## Graph tools can gain sensitivity to anomalous event topologies.

Graph tools identify patterns in SM-only vs SM+BSM network connections.

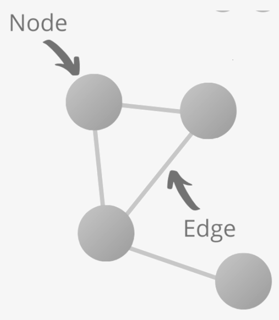
How do we optimize BSM analyses?

- Additional observables,
- Model-independence,
- Targeting rare, complex decays.

Calculating graph network variables contributes additional discrimination between signal and background, without relying on optimization for a chosen model.

Graph network definitions:

- Node: an N-dimensional datapoint.
- Edge: an indicator that two nodes in the N-dimensional space are connected.




Therefore, we require definitions for:

1. Locations of nodes (LHC events) in N-dim space,
2. Distances between all possible pairs of events in the dataset,
3. A binary measure of "similarity" between two nodes/events.

**Definitions:**

Distance metric: the chosen path through the space, for example:

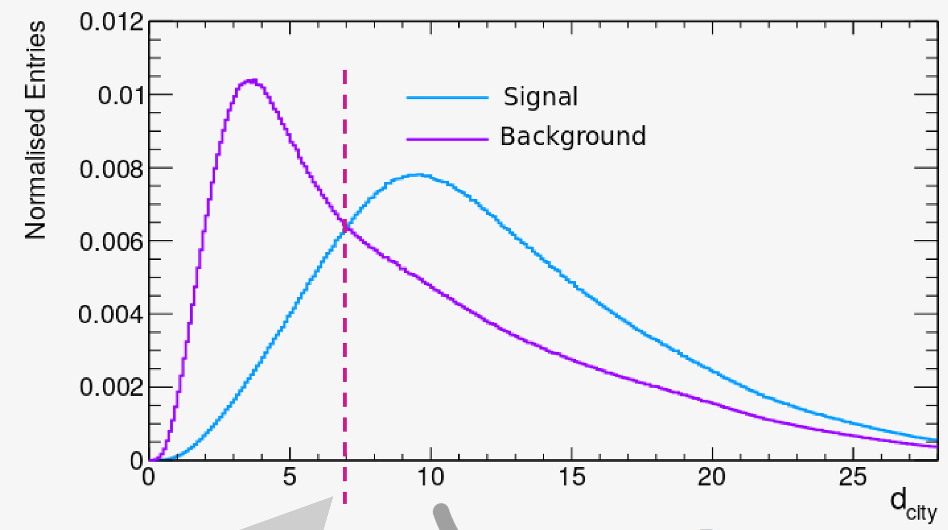
- Euclidean distance: $d_{\text{euc}} = \sqrt{\sum_{i=1}^{n}(u_i - v_i)^2}$.
- Cosine distance: $d_{\cos} = 1 - \frac{u\cdot v}{\sqrt{u\cdot u}\sqrt{v\cdot v}}$.

Linking length: the maximum distance between two connected nodes.

## Edges are information sharing.

For example, to test a toy dataset comprised of 10,000 'BSM' and 10,000 'SM' events in a 5-dimensional simulated kinematic space, we calculate 'cityblock distance':

$$d_{\text{city}} = \sum_{i=1}^{n} |u_i - v_i|$$




Choose to link pairwise events with d > 7.

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 1 \\ 1 & 0 & 1 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}$$

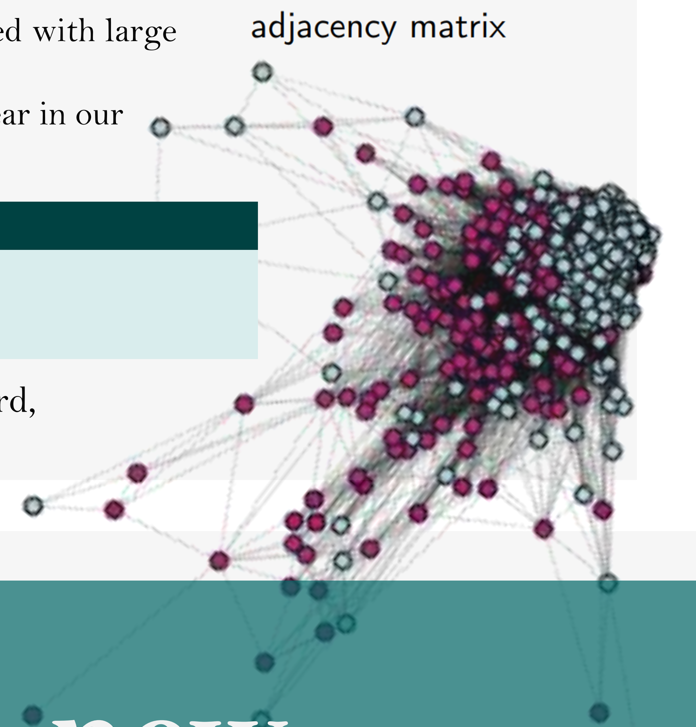
Square, symmetric adjacency matrix

Aim: to compare signal-plus-background networks with background-only networks.
Problem: networks cannot scale to contain enough events to represent the required differences in cross-sections.
Background processes are simulated with large differences in weights.
Solution: weight factors must appear in our networks.

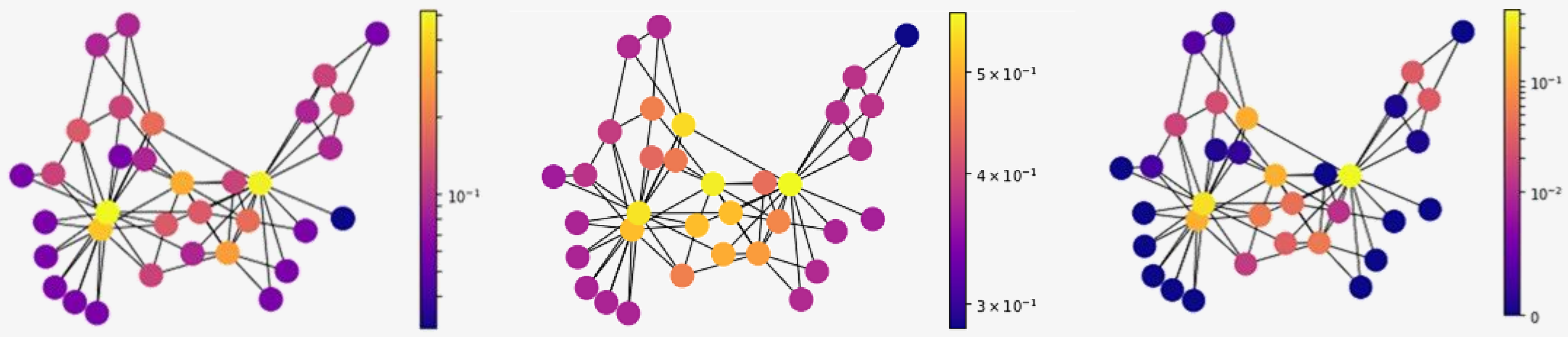
**Networks provide two options:**

- Add weights to the edges
- Add weights to the nodes

Node weights are non-standard, but the only viable option.



### Network metric sample distributions and their definitions



**Degree centrality**
High degree = large number of links.

Node-weighted degree for node $v$: $k^*_v = \frac{\sum_{i\in\mathcal{N}_v^+} w_i}{W}$

Where $W$ is the sum of all node weights over all events $i$.

**Closeness centrality**
High closeness = smaller average number of links to all other nodes.

Node-weighted closeness: $CC^*_v = \frac{W}{\sum_{i\in\mathcal{N}} w_i d^*_{vi}} \in [0,1]$

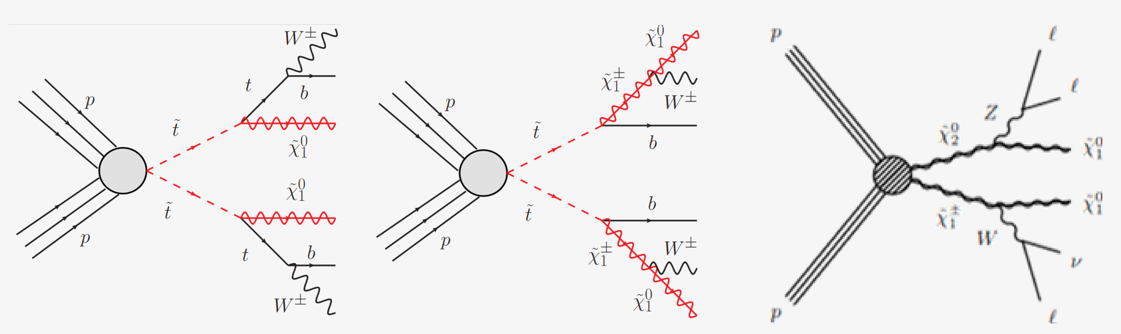
Where $d$ is the number of links on a shortest path.

**Betweenness centrality**
High betweenness = high probability that a shortest path passes through this node.

Node-weighted betweenness: $BC^*_v = \langle n^*_{ab}(v)/n^*_{ab}\rangle^{wsum}_{ab} \in [0, W^2/w_v]$

Where $n$ is the number of shortest paths between randomly chosen nodes $a$ and $b$.

## Our analysis: a supersymmetry search example



Stop quark search + Electroweakino search

6D kinematic graph space:

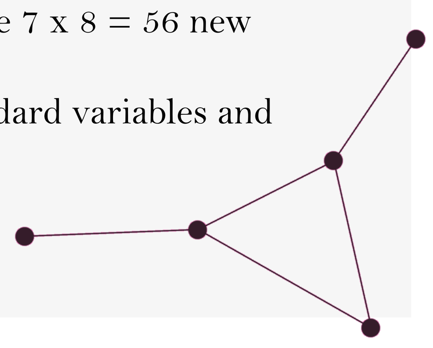
- Leading jet transverse momentum,
- Missing transverse energy,
- Minimum transverse mass (of the two b-jets),
- Minimum invariant mass of the lepton and two b-jets,
- Scalar sum of the transverse momenta,
- Asymmetric mT2.

5D kinematic graph space:

- Missing transverse energy,
- Transverse momentum of the Z boson,
- Azimuthal angle between the two leptons associated with the Z boson,
- Minimum transverse mass,
- Azimuthal angle between the Z boson and lepton from the W boson.

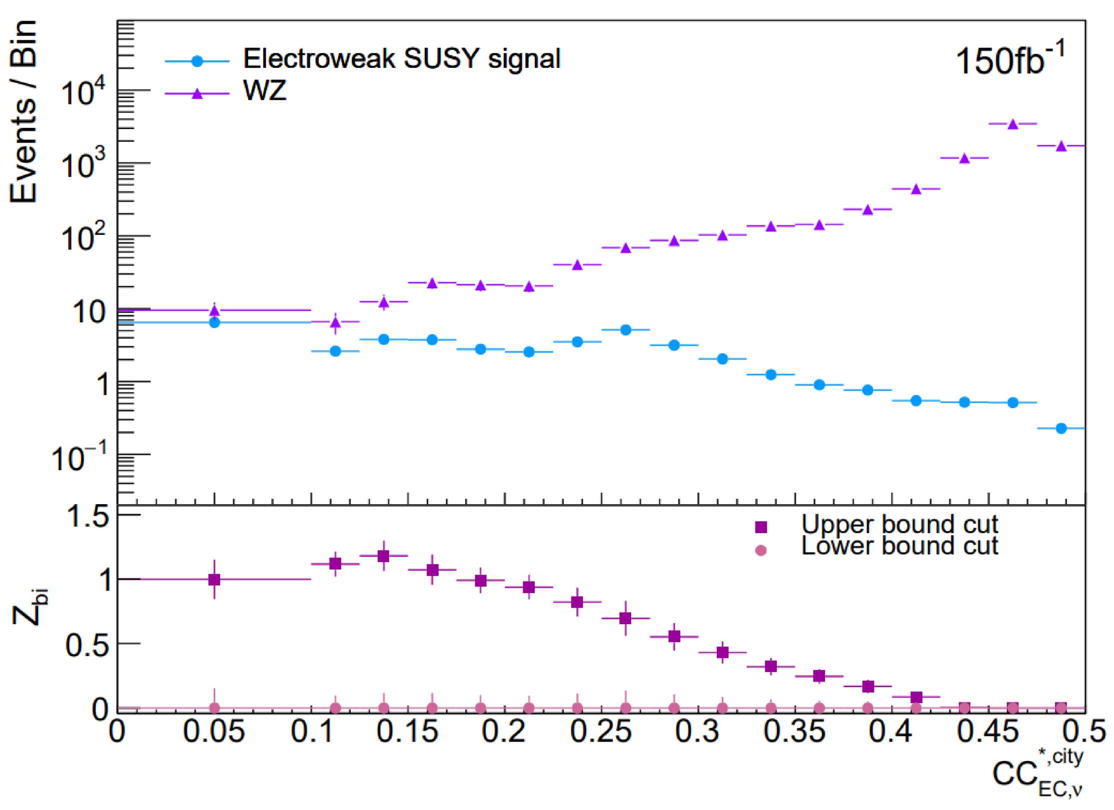
Benefits:

- A wide range of new network variables are available, increasing the likelihood that some are useful discriminators.
  - We chose seven distance metrics to define 7 x 8 = 56 new variables.
- Cuts can be placed on a combination of standard variables and network metrics to increase signal yield.



## Results from network calculations

We show discrimination between signal and background network metric distributions in several new metrics. For example, below is the closeness calculated from the cityblock distance in our electroweak study.

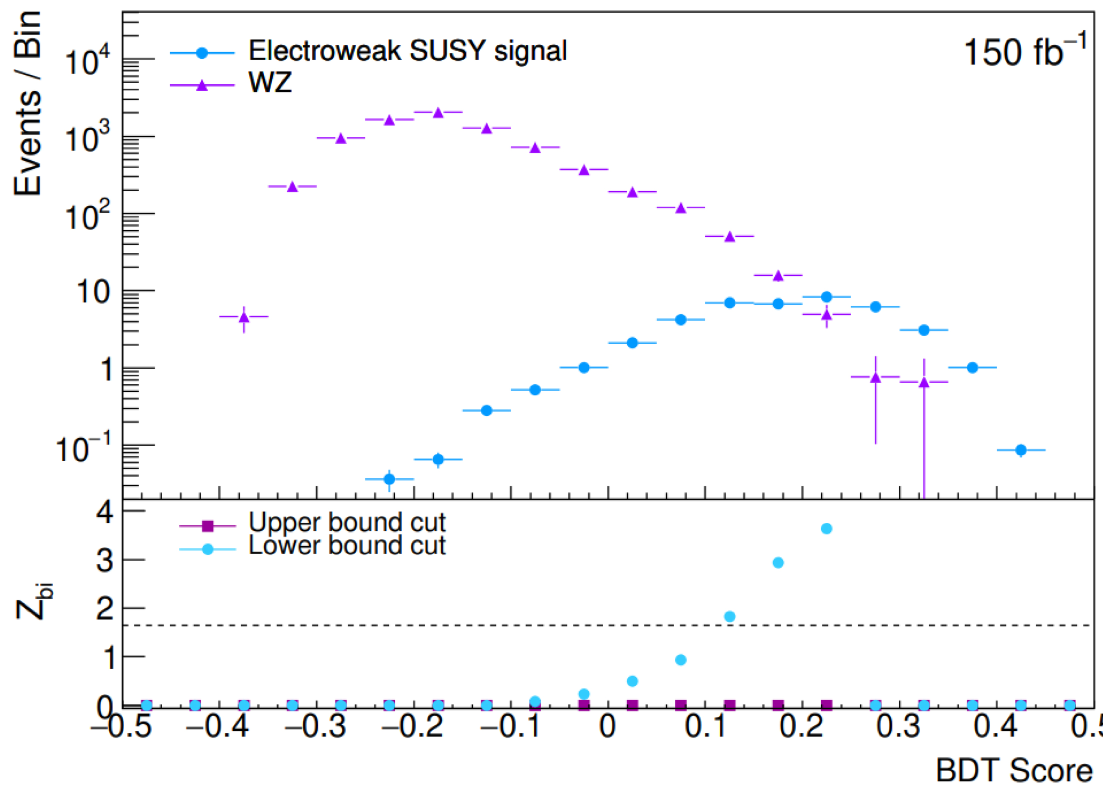



An example cut and count yield using the above distribution:

| Requirement | $N_{\text{signal}}$ | $N_{\text{background}}$ | $Z_{\text{bi}}$ |
|---|---|---|---|
| $k^{*,\text{euc}}_v < 0.003, CC^{*,\text{corr}}_{EC,v} > 0.324$ | $7.78\pm0.17$ | $6.53\pm1.29$ | 1.96 |

In addition: we can combine network metrics with machine learning, e.g. in a boosted decision tree trained on standard variables + network metrics.

Binomial significance, as from a number counting experiment




Above shows an example boosted decision tree, which improved performance when trained on degree centrality in addition to standard kinematic variables. Performance increased from a Zbi of 3.63 to 3.98.

## Further work

Ideas are welcome to guide our choice of BSM model as we move to ATLAS datasets.

We are also interested in optimizing graph techniques for large datasets. Graph calculations are computationally limited, and rely on node-splitting and node-merging algorithms to accurately represent large LHC datasets with only a subset of events. We are testing in what contexts these are reliable, for every network metric.

References: Hong, S, Coutinho, B, Dey, A, Barabast, AL, Vogelsberger, M, Hernquist, L & Gebhardt, K 2016, 'Discriminating topology in galaxy distributions using network analysis', *Monthly Notices of the Royal Astronomical Society*, vol. 459, no. 3, pp. 2690-2700.
ATLAS Collaboration 2008, 'The ATLAS Experiment at the CERN Large Hadron Collider', *Journal of Instrumentation*, vol. 3.