Application: Supersymmetry and Searches at Colliders

Supersymmetry (SUSY) was (is) a popular model for physics <u>beyond the Standard Model</u> (BSM).

This proposal introduces partners for every particle in the SM through symmetry between bosons (integer spin) and fermions (½-integer spin).

It may seem a fairly innocuous extension but the underlying principle is a <u>unique</u> extension of the

Poincaré group for spacetime.

Poincare algebra in $SO(1,3)$

P^μ - translations

$M^{\mu\nu}$ - Lorentz boosts + rotations

satisfy:

$$[P^\mu, P^\nu] = 0$$

$$-i\,[M_{\mu\nu}, P_\rho] = g_{\mu\nu} P_\nu - g_{\nu\rho} P_\mu$$

$$-i\,[M_{\mu\nu}, M_{\rho\sigma}] = g_{\mu\rho} M_{\nu\sigma} - g_{\nu\rho} M_{\mu\sigma} + g_{\nu\sigma} M_{\mu\rho} - g_{\mu\sigma} M_{\nu\rho}$$

Supersymmetry is represented by an operators $Q, \bar{Q}$ that relate particles differing by ½ integer spin. Q and $\bar{Q}$ are Weyl spinors and the central objects in the anti-commutation relations involving the (extended) Pauli matrices $\sigma^{\mu}_{\alpha\dot\alpha}$

$$\{ Q_\alpha , \bar{Q}_{\dot\alpha} \} = 2\sigma^{\mu}_{\alpha\dot\alpha} P_\mu$$

In addition

$$[M^{\mu\nu}, Q_\alpha] = \sigma^{\mu\nu}{}_{\alpha}{}^{\beta} Q_\beta$$

$$[M^{\mu\nu}, \bar{Q}^{\dot\alpha}] = (\bar\sigma^{\mu\nu})^{\dot\alpha}{}_{\dot\beta}\, Q^{\dot\beta}$$

where $\sigma^{\mu\nu}{}_{\alpha}{}^{\beta} = \frac{i}{4}\left(\sigma^\mu \bar\sigma^\nu - \sigma^\nu \bar\sigma^\mu\right)_\alpha{}^\beta$

and

$$[Q_\alpha, P^\mu] = 0$$

$$\{Q_\alpha, Q_\beta\} = 0$$

We can then set up models in which

$$[Q_\alpha, \phi(x)] = \psi_\alpha(x)$$

scalar field $\qquad$ fermion field

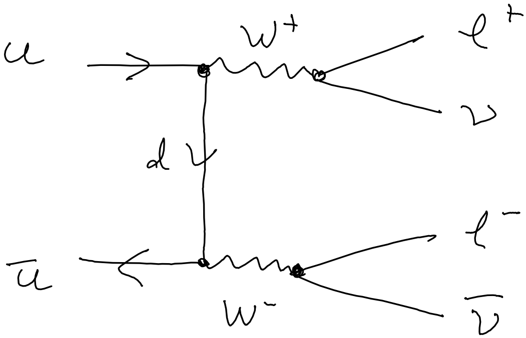
SUSY in reality

* In unbroken SUSY, all super-partners have the same mass. The symmetry must break to match the observation that there is no SUSY below currently accessible energies.

  - SUSY breaking can happen in a variety of scenarios all with different parameters, $\underbrace{\text{particle spectra}}_{\text{mass hierachies}}$.

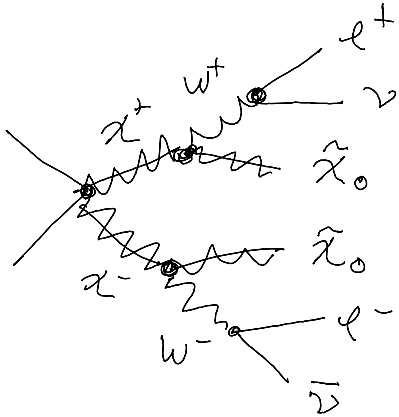
  - supersymmetric particles can decay to standard model partners or into light states decoupled from the standard model (c.f. Neutrinos).

For this reason we may have susy (or indeed any BSM / New Physics process) affecting the production of the same final state possible in the SM, e.g.



final state

$2\ell + \underbrace{\text{MET}}_{\text{missing transverse energy}}$



final state

2ℓ + MET

($\chi^{\pm}, \chi^{0}$ susy particles)

Try to use <u>simulated</u> data to generate training data for classifier that takes final state information and predicts if process looks like SUSY or not.

<u>NB</u>: SUSY has almost entirely been ruled out in scenarios where SUSY breaking scale ~ EW scale ...

[10-20 years ago MSSM (minimal supersym. standard model) considered (by many) to be a serious candidate for a theory BSM]

In this example we can use it as an application in which we have:

1) many input "features" describing detector level observables

2) Simulated events for two theoretical models ("SM" or "SUSY").

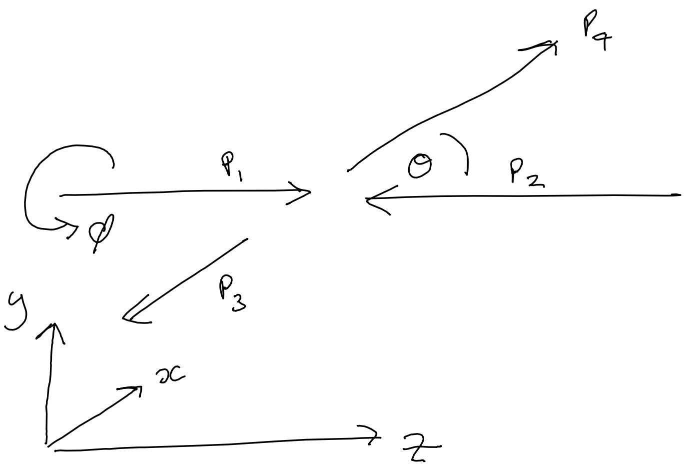
we will use to investigate

a) processing a large dataset (5M events)

b) loss function regularisation methods

c) difference between linear and non-linear classifiers.

Kinematic variables



e.g. $2 \to 2$

$$p = (E, p_x, p_y, p_z)$$

$$= \left(m_T \cosh(y),\ p_T \cos\phi,\ p_T \sin\phi,\ m_T \sinh(y)\right)$$

where $\quad p_T$ = tranverse momentum

E_T = transverse energy

y = rapidity

ϕ = azimuthal angle

if particle is massless (e.g. photon)

$$p \overset{m=0}{=} p_T \left(\cosh(y), \cos(\phi), \sin(\phi), \sinh(y) \right)$$

The rapidity

$$y = \frac{1}{2} \log\left(\frac{E+p_z}{E-p_z}\right)$$

$$\approx -\log\left(\tan\left(\theta/2\right)\right) =: \eta$$

η - pseudo rapidity.

Analyzing classifier peformance :

Reciever Operating Characteristics (ROC)

and Area under the Curve (AUC)

We consider a binary classifier as before where the sigmoid function

$$\sigma(\underset{\text{input}}{z(x)}) = \frac{1}{1+e^{-z(x)}} = P(x \mid y=1)$$

is taken for the classifier output representing the probability the outcome is 1

Take th predicted value to be 1 if the probability is > threshold, T (50% would be obvious choice) $\hat{y}(x, T) = \Theta\left(P(x|y=1) - T\right)$

Let's plot against unseen test data (x_i, y_i). We can compute

$$\hat{y}(x_i, T) - y_i$$

which for a given value of T will be either

$\hat{y}$	y	$\hat{y} - y$	
1	1	0	True Positive
1	0	1	False Positive
0	1	-1	False Negative
0	0	0	True Negative

Aside Confusion Matrix

$\hat{y} \backslash y$	1	0
1	"TP"	"FP"
0	"FN"	"TN"

For a given test set the probability distribution we may define

TP, the number of true positives

TN, the number of true negatives

SP, the number of positive values in the "signal" data set (i.e. y_i test data)

SN, # of -ve values in y_i

N, the number of test samples

FP, the number of false positives

FN, the number of false negatives

which satisfy:

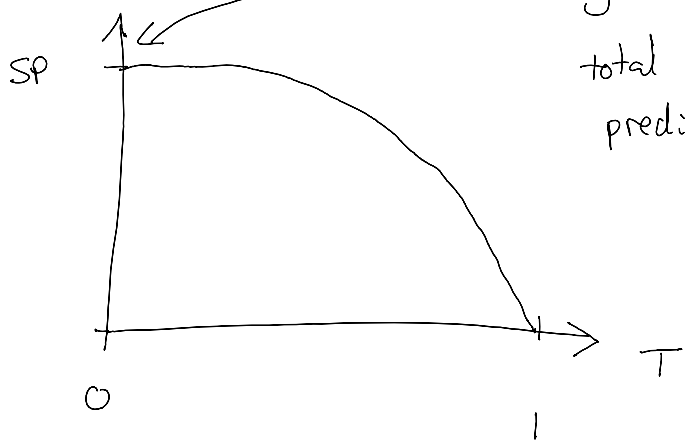
$$SP + SN = N$$

$$TP + FN = SP$$

$$TN + FP = SN$$

We may then plot TP as a function of the threshold T, (e.g.)

if $T = 0$ all $\hat{y} = 1$ hence $TP = SP$
total number of tue predictions $= N$



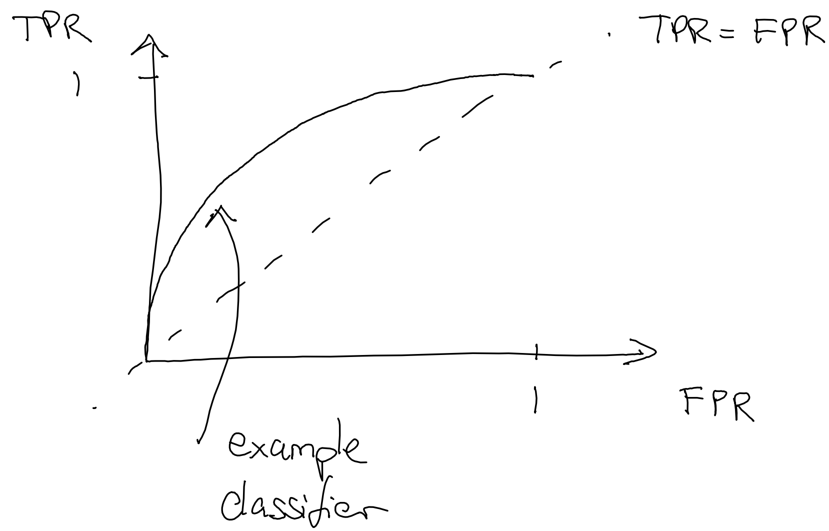
The <u>True Positive Rate</u> is

$$TPR = \frac{TP}{TP + FN} = \frac{TP}{SP}$$

and <u>False Positive Rate</u>

$$FPR = \frac{FP}{FP + TN} = \frac{FP}{SN}$$

The ROC curve is FPR vs TPR and will take the form



A perfect classifier will always have TPR = 1 while a worst case classifier will have TPR = FPR (50:50 chance...)

Problems with this analysis

1) Input data is simply labelled "SUSY" — clearly this is insufficient to describe the particular SUSY model (breaking scenario ... setting of free parameters ...) or the details of the simulation (Leading order, parton shower? ...)

NB: MSSM can have ~100 free params compared to ~18 for SM.

2) Large amount of data not really needed

3) No attempt to understand why certain values lead to better classification