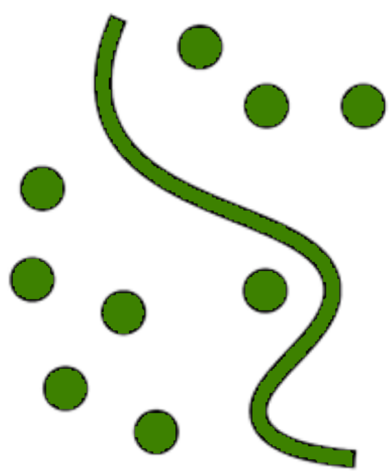
# Generic event classification using rapidity-mass matrices and machine learning

https://arxiv.org/abs/1805.11650


S. Chekanov

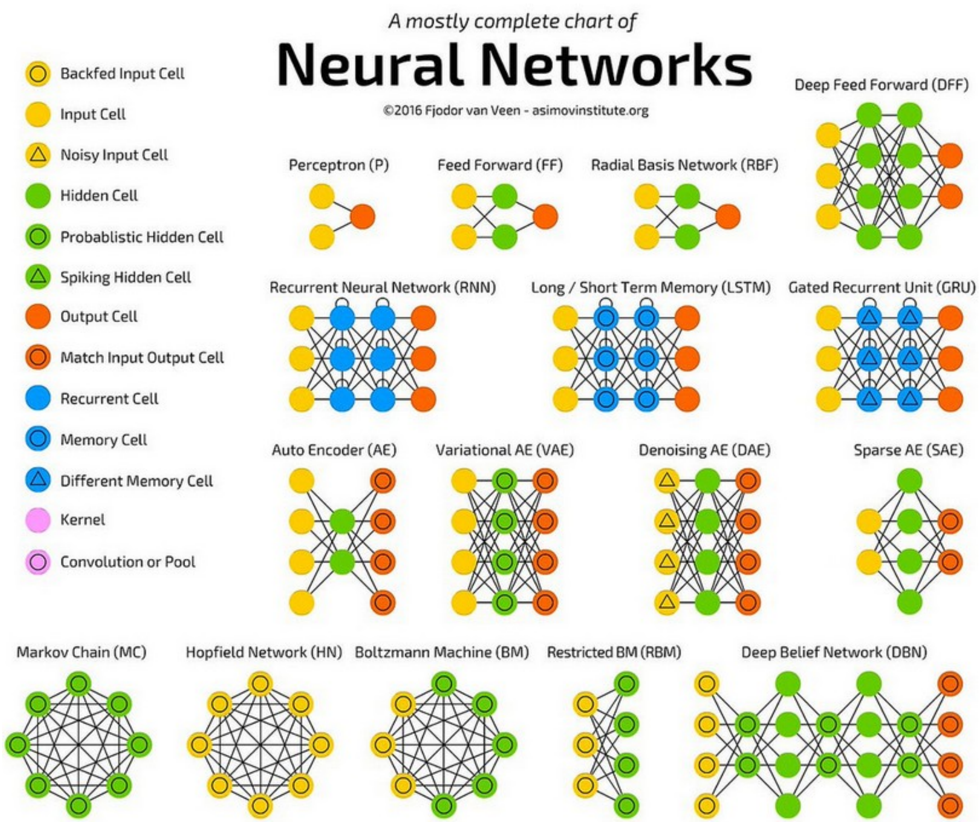

**ANL Theory Seminar**

(September 25, 2018)



# Using Artificial Neural Networks (ANN)

- Many new  neural networks invented in recent years
- Pushed by leading industries for object identification in images/video




When applies to HEP:

Success of ANN in HEP is loosely related to the choice ANN
(i.e. one ANN vs another ANN)

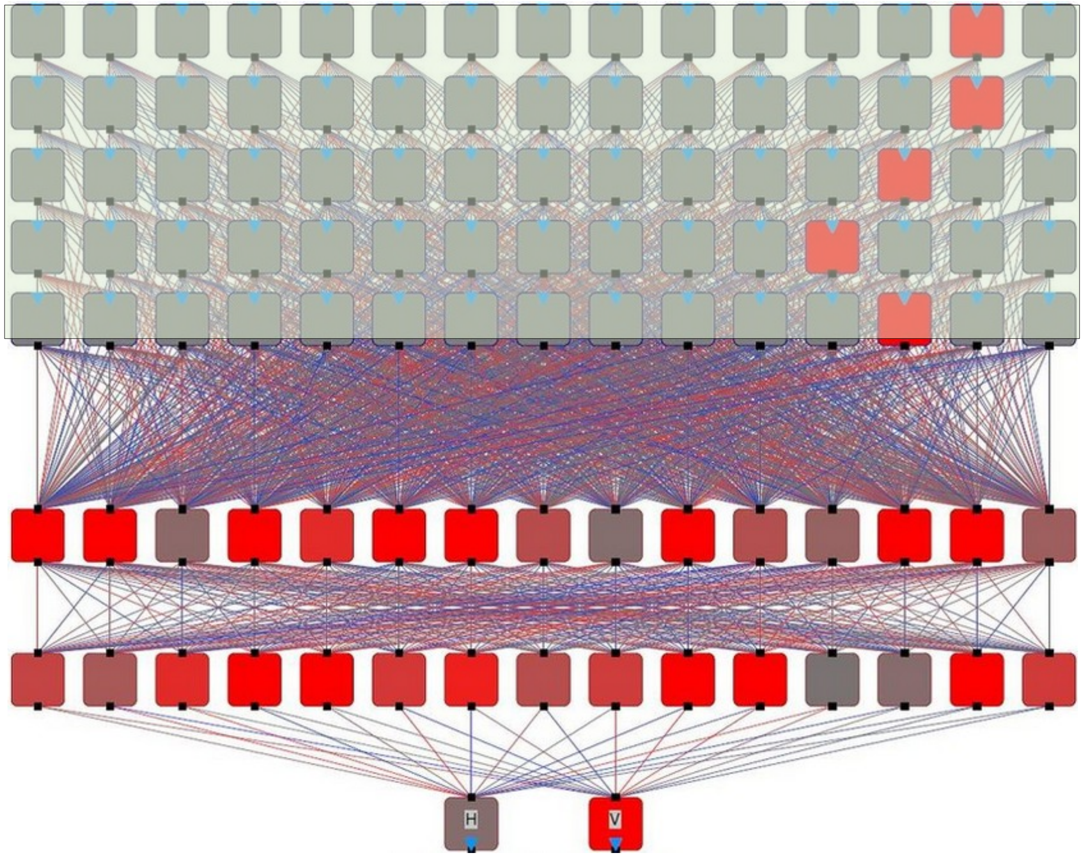
How to prepare a meaningful input?

90% of time spent on hand-crafting input variables

Credit the Asimov Institute & J.Stirrup

# Most common usage of ANN: Image identification

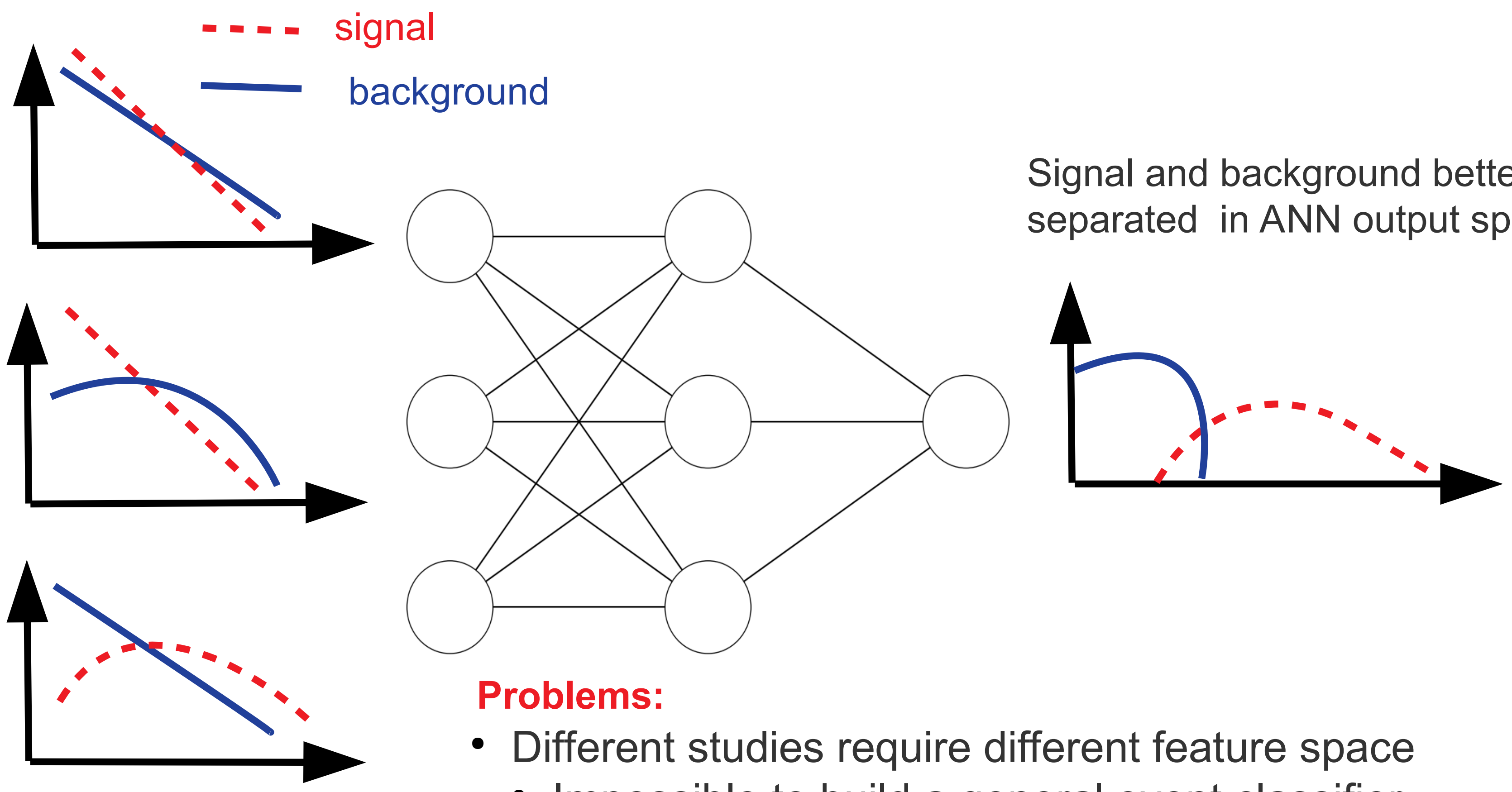
Image with pixels




Event classification using imaging of collision events. S.Chekanov (ANL) et al

# ANN in HEP

- Extensively used in HEP in the last ~25 years




"feature space"

Signal and background better separated in ANN output space

**Problems:**

- Different studies require different feature space
  - Impossible to build a general event classifier
- No unique "feature space" even for a single analysis
  - Pick and Use (PaU) method

# Feature space for event classifications

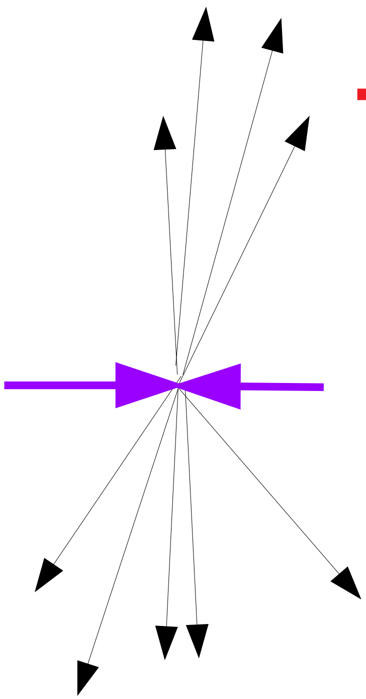
- **Event classification depends on prepared inputs**
  - Identify variables with background and signal “features”
  - Data and dimensionality reduction
  - Data re-scale (the range between 0 and 1 is a popular choice),
  - Data normalization (to avoid cases when some of input values overweight others)
  - etc.
- **ANN are suppose to simplify analysis  but:**
  - Preparing analysis for NN  is time consuming
  - Need to hand-pick variables, study them etc.. No uniqueness of input variables.

- **Idea: create a general image-like transformation of lists with 4-momenta to data structures that reflect most significant features of hadronic-final state**
  - General representation of collision event. Single and double- particle densities
  - Natural language for machine learning → leverage algorithms from leading industries
  - Easy to visualize for humans
  - Leverage algorithms for image identification from  leading industries

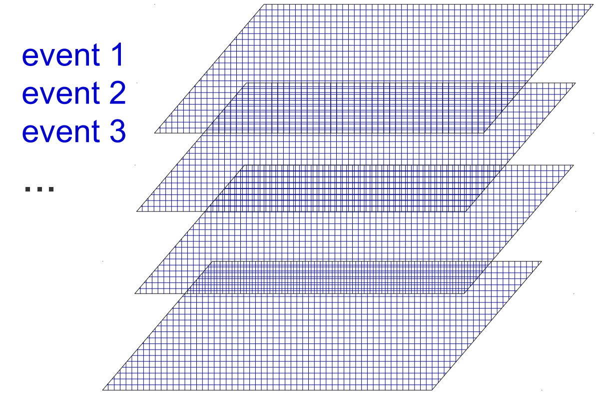
FREE BONUS!

# Mapping collision events to 2D arrays

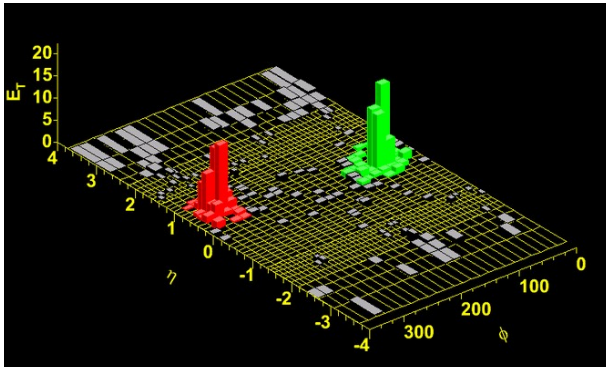
List with 4-momenta of N particles



**Matrices:**
- Fixed size
- Dimensionless
- Lorentz invariant
- Fixed range of values
- Single particle densities
- Two-particle correlations
- Independent cells
- Cells connected by proximity due to a well-defined hierarchy
- Easy to visualize




**NOT GOOD for our goal**



# Rapidity-mass matrix (RMM)

jets ⟵⟶ muons ⟵⟶ .. electrons, photons

$$
\begin{pmatrix}
\mathbf{e_T^{miss}} & m_T(j_1) & m_T(j_2) & \ldots m_T(j_N) & m_T(\mu_1) & m_T(\mu_2) & \ldots m_T(\mu_N) \\
h_L(j_1) & \mathbf{e_T(j_1)} & m(j_1,j_2) & \ldots m(j_1,j_N) & m(j_1,\mu_1) & m(j_1,\mu_2) & \ldots m(j_1,\mu_N) \\
h_L(j_2) & h(j_1,j_2) & \delta\mathbf{e_T(j_2)} & \ldots m(j_2,j_N) & m(j_2,\mu_1) & m(j_2,\mu_2) & \ldots m(j_2,\mu_N) \\
\ldots & \ldots & \ldots & \ldots, & \ldots & \ldots & \ldots \\
h_L(j_N) & h(j_1,j_N) & \ldots & \ldots \delta\mathbf{e_T(j_N)} & m(j_N,\mu_1) & m(j_N,\mu_2) & \ldots m(j_N,\mu_N) \\
h_L(\mu_1) & h(\mu_1,j_1) & h(\mu_1,j_2) & \ldots h(\mu_1,j_N) & \mathbf{e_T(\mu_1)} & m(\mu_1,\mu_2) & m(\mu_1,\mu_N) \\
h_L(\mu_2) & h(\mu_2,j_1) & h(\mu_1,j_2) & \ldots h(\mu_2,j_N) & h(\mu_1,\mu_2) & \delta\mathbf{e_T(\mu_2)} & m(\mu_2,\mu_N) \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \\
h_L(\mu_N) & h(\mu_N,j_1) & h(\mu_N,j_2) & \ldots h(\mu_N,j_N) & h(\mu_N,\mu_1) & h(\mu_N,\mu_2) & \delta\mathbf{e_T(\mu_N)}
\end{pmatrix}
$$

$e_T^{miss}$ – missing ET of events

$m_T(i)$ - transverse mass of object "i"

$e_T(i)$ - transverse energy (ordered)

$\delta e_T(i)$ – transverse energy imbalances

$m(i,j)$ – two-particle invariant masses

(scaled by $1/\sqrt{s}$)

$h_L(i)$ - cosh(y)-1 (y is rapidity) – Lorentz factor

$h(i,j)$ - $\cosh(0.5(y_i - y_j))$ -1 – rapidity difference

(scaled by a constant)

What does this matrix represent?

Event classification using imaging of collision events. S.Chekanov (ANL) et al

# Events with missing particles (large missing ET)

Missing transverse energy

$$\begin{pmatrix}
\mathbf{e_T^{miss}} & m_T(j_1) & m_T(j_2) & \dots m_T(j_N) & m_T(\mu_1) & m_T(\mu_2) & \dots m_T(\mu_N) \\
h_L(j_1) & \mathbf{e_T(j_1)} & m(j_1,j_2) & \dots m(j_1,j_N) & m(j_1,\mu_1) & m(j_1,\mu_2) & \dots m(j_1,\mu_N) \\
h_L(j_2) & h(j_1,j_2) & \delta\mathbf{e_T(j_2)} & \dots m(j_2,j_N) & m(j_2,\mu_1) & m(j_2,\mu_2) & \dots m(j_2,\mu_N) \\
\dots & \dots & \dots & \dots, & \dots & \dots & \dots \\
h_L(j_N) & h(j_1,j_N) & \dots & \dots\delta\mathbf{e_T(j_N)} & m(j_N,\mu_1) & m(j_N,\mu_2) & \dots m(j_N,\mu_N) \\
h_L(\mu_1) & h(\mu_1,j_1) & h(\mu_1,j_2) & \dots h(\mu_1,j_N) & \mathbf{e_T}(\mu_1) & m(\mu_1,\mu_2) & m(\mu_1,\mu_N) \\
h_L(\mu_2) & h(\mu_2,j_1) & h(\mu_1,j_2) & \dots h(\mu_2,j_N) & h(\mu_1,\mu_2) & \delta\mathbf{e_T}(\mu_2) & m(\mu_2,\mu_N) \\
\dots & \dots & \dots & \dots & \dots & \dots & \\
h_L(\mu_N) & h(\mu_N,j_1) & h(\mu_N,j_2) & \dots h(\mu_N,j_N) & h(\mu_N,\mu_1) & h(\mu_N,\mu_2) & \delta\mathbf{e_T}(\mu_N)
\end{pmatrix}$$

Sensitive to particles that include decays to invisible particles

Missing transverse mass for each particle type


$$\begin{pmatrix}
\mathbf{e_T^{miss}} & m_T(j_1) & m_T(j_2) & \dots m_T(j_N) & m_T(\mu_1) & m_T(\mu_2) & \dots m_T(\mu_N) \\
h_L(j_1) & \mathbf{e_T(j_1)} & m(j_1,j_2) & \dots m(j_1,j_N) & m(j_1,\mu_1) & m(j_1,\mu_2) & \dots m(j_1,\mu_N) \\
h_L(j_2) & h(j_1,j_2) & \delta\mathbf{e_T(j_2)} & \dots m(j_2,j_N) & m(j_2,\mu_1) & m(j_2,\mu_2) & \dots m(j_2,\mu_N) \\
\dots & \dots & \dots & \dots, & \dots & \dots & \dots \\
h_L(j_N) & h(j_1,j_N) & \dots & \dots\delta\mathbf{e_T(j_N)} & m(j_N,\mu_1) & m(j_N,\mu_2) & \dots m(j_N,\mu_N) \\
h_L(\mu_1) & h(\mu_1,j_1) & h(\mu_1,j_2) & \dots h(\mu_1,j_N) & \mathbf{e_T}(\mu_1) & m(\mu_1,\mu_2) & m(\mu_1,\mu_N) \\
h_L(\mu_2) & h(\mu_2,j_1) & h(\mu_1,j_2) & \dots h(\mu_2,j_N) & h(\mu_1,\mu_2) & \delta\mathbf{e_T}(\mu_2) & m(\mu_2,\mu_N) \\
\dots & \dots & \dots & \dots & \dots & \dots & \\
h_L(\mu_N) & h(\mu_N,j_1) & h(\mu_N,j_2) & \dots h(\mu_N,j_N) & h(\mu_N,\mu_1) & h(\mu_N,\mu_2) & \delta\mathbf{e_T}(\mu_N)
\end{pmatrix}$$


SM particles with neutrino decays, exotic particles (SUSY, Dark Matter candidates etc)

Example: reconstruction of transverse masses (W → μ ν)

# Transverse momenta of all objects

$$
\begin{pmatrix}
\mathbf{e_T^{miss}} & m_T(j_1) & m_T(j_2) & \ldots m_T(j_N) & m_T(\mu_1) & m_T(\mu_2) & \ldots m_T(\mu_N) \\
h_L(j_1) & \mathbf{e_T(j_1)} & m(j_1,j_2) & \ldots m(j_1,j_N) & m(j_1,\mu_1) & m(j_1,\mu_2) & \ldots m(j_1,\mu_N) \\
h_L(j_2) & h(j_1,j_2) & \delta\mathbf{e_T(j_2)} & \ldots m(j_2,j_N) & m(j_2,\mu_1) & m(j_2,\mu_2) & \ldots m(j_2,\mu_N) \\
\ldots & \ldots & \ldots & \ldots, & \ldots & \ldots & \ldots \\
h_L(j_N) & h(j_1,j_N) & \ldots & \ldots \delta\mathbf{e_T(j_N)} & m(j_N,\mu_1) & m(j_N,\mu_2) & \ldots m(j_N,\mu_N) \\
h_L(\mu_1) & h(\mu_1,j_1) & h(\mu_1,j_2) & \ldots h(\mu_1,j_N) & \mathbf{e_T}(\mu_1) & m(\mu_1,\mu_2) & m(\mu_1,\mu_N) \\
h_L(\mu_2) & h(\mu_2,j_1) & h(\mu_1,j_2) & \ldots h(\mu_2,j_N) & h(\mu_1,\mu_2) & \delta\mathbf{e_T}(\mu_2) & m(\mu_2,\mu_N) \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \\
h_L(\mu_N) & h(\mu_N,j_1) & h(\mu_N,j_2) & \ldots h(\mu_N,j_N) & h(\mu_N,\mu_1) & h(\mu_N,\mu_2) & \delta\mathbf{e_T}(\mu_N)
\end{pmatrix}
$$

**Diagonal elements:**

**Transverse momentum of leading in Et objects**

**Transverse momentum imbalances**

$$\delta e_T(i_n) = \frac{E_T(i_{n-1}) - E_T(i_n)}{E_T(i_{n-1}) + E_T(i_n)}$$

**Transverse energy imbalances:**

- Sensitive to interactions of partons in the medium of heavy ion collisions
- Can be used to separate dijet QCD events from more complex events

**Can be use reconstruct:**

- Transverse energies of all objects
- $H_T$ of events
- Energies $E(i) = e_T(i)\sqrt{s}(h_L(i)/C + 1)$

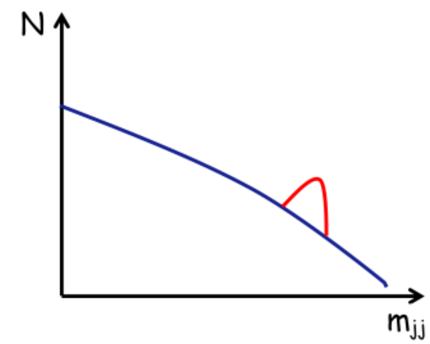
# Invariant masses of particle (jet) pairs

$$
\begin{pmatrix}
\mathbf{e_T^{miss}} & m_T(j_1) & m_T(j_2) & \ldots m_T(j_N) & m_T(\mu_1) & m_T(\mu_2) & \ldots m_T(\mu_N) \\
h_L(j_1) & \mathbf{e_T(j_1)} & m(j_1,j_2) & \ldots m(j_1,j_N) & m(j_1,\mu_1) & m(j_1,\mu_2) & \ldots m(j_1,\mu_N) \\
h_L(j_2) & h(j_1,j_2) & \delta\mathbf{e_T(j_2)} & \ldots m(j_2,j_N) & m(j_2,\mu_1) & m(j_2,\mu_2) & \ldots m(j_2,\mu_N) \\
\ldots & \ldots & \ldots & \ldots, & \ldots & \ldots & \ldots \\
h_L(j_N) & h(j_1,j_N) & \ldots & \ldots \delta\mathbf{e_T(j_N)} & m(j_N,\mu_1) & m(j_N,\mu_2) & \ldots m(j_N,\mu_N) \\
h_L(\mu_1) & h(\mu_1,j_1) & h(\mu_1,j_2) & \ldots h(\mu_1,j_N) & \mathbf{e_T}(\mu_1) & m(\mu_1,\mu_2) & m(\mu_1,\mu_N) \\
h_L(\mu_2) & h(\mu_2,j_1) & h(\mu_1,j_2) & \ldots h(\mu_2,j_N) & h(\mu_1,\mu_2) & \delta\mathbf{e_T}(\mu_2) & m(\mu_2,\mu_N) \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \\
h_L(\mu_N) & h(\mu_N,j_1) & h(\mu_N,j_2) & \ldots h(\mu_N,j_N) & h(\mu_N,\mu_1) & h(\mu_N,\mu_2) & \delta\mathbf{e_T}(\mu_N)
\end{pmatrix}
$$

**m(i,j)** - Invariant masses of all objects (jets and identified particles)

Peaks in invariant masses will be shown as enhanced top-right cells

“Gold- standard” for search for new physics



# Longitudinal flow in events

$$\begin{pmatrix} \mathbf{e_T^{miss}} & m_T(j_1) & m_T(j_2) & \ldots m_T(j_N) & m_T(\mu_1) & m_T(\mu_2) & \ldots m_T(\mu_N) \\ h_L(j_1) & \mathbf{e_T(j_1)} & m(j_1,j_2) & \ldots m(j_1,j_N) & m(j_1,\mu_1) & m(j_1,\mu_2) & \ldots m(j_1,\mu_N) \\ h_L(j_2) & h(j_1,j_2) & \delta\mathbf{e_T(j_2)} & \ldots m(j_2,j_N) & m(j_2,\mu_1) & m(j_2,\mu_2) & \ldots m(j_2,\mu_N) \\ \ldots & \ldots & \ldots & \ldots, & \ldots & \ldots & \ldots \\ h_L(j_N) & h(j_1,j_N) & \ldots & \ldots\delta\mathbf{e_T(j_N)} & m(j_N,\mu_1) & m(j_N,\mu_2) & \ldots m(j_N,\mu_N) \\ h_L(\mu_1) & h(\mu_1,j_1) & h(\mu_1,j_2) & \ldots h(\mu_1,j_N) & \mathbf{e_T}(\mu_1) & m(\mu_1,\mu_2) & m(\mu_1,\mu_N) \\ h_L(\mu_2) & h(\mu_2,j_1) & h(\mu_1,j_2) & \ldots h(\mu_2,j_N) & h(\mu_1,\mu_2) & \delta\mathbf{e_T}(\mu_2) & m(\mu_2,\mu_N) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \\ h_L(\mu_N) & h(\mu_N,j_1) & h(\mu_N,j_2) & \ldots h(\mu_N,j_N) & h(\mu_N,\mu_1) & h(\mu_N,\mu_2) & \delta\mathbf{e_T}(\mu_N) \end{pmatrix}$$

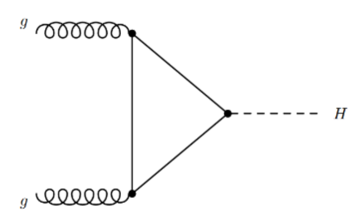
**$h_L$ = cosh(y) -1:**

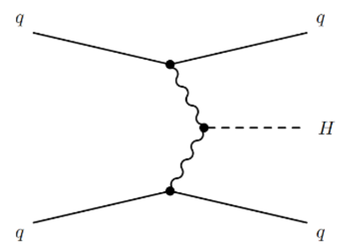
- Large values for forward physics
- Small value (~0) for central production

**cosh(y) = γ** is Lorentz factor

<u>Example:</u>

- VBF Higgs production has large values in the first column (centrally produced)
- ..



# Rapidity differences

$$
\begin{pmatrix}
\mathbf{e}_{\mathbf{T}}^{\mathbf{miss}} & m_T(j_1) & m_T(j_2) & \ldots m_T(j_N) & m_T(\mu_1) & m_T(\mu_2) & \ldots m_T(\mu_N) \\
h_L(j_1) & \mathbf{e}_{\mathbf{T}}(\mathbf{j_1}) & m(j_1,j_2) & \ldots m(j_1,j_N) & m(j_1,\mu_1) & m(j_1,\mu_2) & \ldots m(j_1,\mu_N) \\
h_L(j_2) & h(j_1,j_2) & \delta \mathbf{e}_{\mathbf{T}}(\mathbf{j_2}) & \ldots m(j_2,j_N) & m(j_2,\mu_1) & m(j_2,\mu_2) & \ldots m(j_2,\mu_N) \\
\ldots & \ldots & \ldots & \ldots, & \ldots & \ldots & \ldots \\
h_L(j_N) & h(j_1,j_N) & \ldots & \ldots \delta \mathbf{e}_{\mathbf{T}}(\mathbf{j_N}) & m(j_N,\mu_1) & m(j_N,\mu_2) & \ldots m(j_N,\mu_N) \\
h_L(\mu_1) & h(\mu_1,j_1) & h(\mu_1,j_2) & \ldots h(\mu_1,j_N) & \mathbf{e}_{\mathbf{T}}(\mu_1) & m(\mu_1,\mu_2) & m(\mu_1,\mu_N) \\
h_L(\mu_2) & h(\mu_2,j_1) & h(\mu_1,j_2) & \ldots h(\mu_2,j_N) & h(\mu_1,\mu_2) & \delta \mathbf{e}_{\mathbf{T}}(\mu_2) & m(\mu_2,\mu_N) \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \\
h_L(\mu_N) & h(\mu_N,j_1) & h(\mu_N,j_2) & \ldots h(\mu_N,j_N) & h(\mu_N,\mu_1) & h(\mu_N,\mu_2) & \delta \mathbf{e}_{\mathbf{T}}(\mu_N)
\end{pmatrix}
$$

h(i,j) ~ cosh ( $y_j$ - $y_i$) -1

h(i,j) ~ 0 for collimated 2-particles

Rapidity difference used for:

- Dijet searches (CMS, ATLAS)
- Probing parton dynamics (CMS)

# Useful features

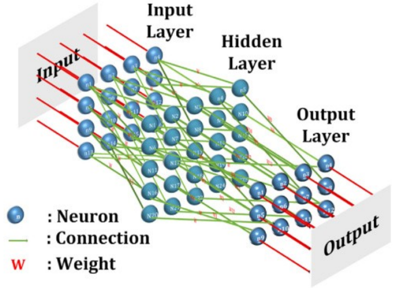
- Dimensionless, Lorentz invariant (except for Lorenz factors for all objects)

- Small linear correlations between RMM cells
  - No redundant information
  - According to Monte Carlo simulations:
    - Pearson correlation coefficient >0.5 for 0.3% of cell pairs
    - Out of 0.4%, 50% correlation is seen between mT and m(i,j)

- Well-defined hierarchy by construction
  - Cells connected by proximity
  - Should look as "images", not as random noise.. → good for visualization

- Natural language for machine learning:
  - Each cell maps to a fixed node/neutron.
  - Normalization and standardization

RMM




*<u>However:</u>*

*- RMM is a sparse matrix for single events. Keep non-zero values and their indexes!*

*- Averaging aver many events makes more visually appealing images*

# Monte Carlo simulations

## Several processes from Pythia8 (LO+PS)

- **Dijet QCD:**
  - All 2→2 processes (10)
- **Top production:**
  - g g -> t tbar
  - q qbar -> t tbar
- **Charged Higgs production**
  - b g -> H+- t
- **Double boson production**
  - f fbar -> gamma*/Z0 gamma*/Z0
  - f fbar' -> Z0 W+-
  - f fbar -> W+ W-
- **SM Higgs production**

http://atlaswww.hep.anl.gov/hepsim/


Show all | p →← p | 8 TeV | 13 TeV | 14 TeV | 27 TeV | 33 TeV | 100 TeV | $e^+ \rightarrow\leftarrow e^-$ | 250 GeV | 380 GeV | 500 GeV | 1 TeV | 3 TeV | $\mu^+ \rightarrow\leftarrow \mu^-$ | 1 TeV

Get involved | Full Search | Experiments | Manual | Mirrors | Tools | About | Login

HEP.ANL.GOV

HepSim

Repository with Monte Carlo simulations for particle physics

- March 15 2018: Charged Higgs event samples
- Sep,22 2017: Z+Higgs → nunu+XX event samples
- Sep,15 2017: Higgs → mu+mu- event samples

Show 25 entries

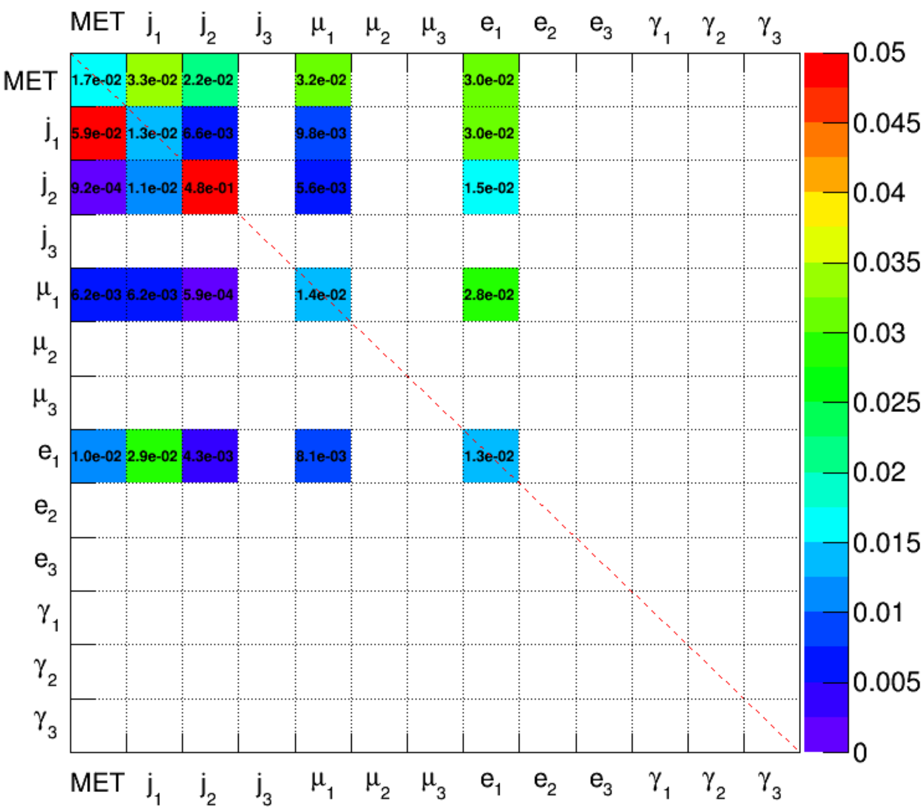
Previous 1 2 3 4 5 ... 13 Next Search:

| Id | →← | E [TeV] | Dataset name | Generator | Process | Topic | Files | Created |
|---|---|---|---|---|---|---|---|---|
| 328 | pp | 13 | tev13pp_pythia8_rmm | PYTHIA8 | Various SM/BSM process for ML | SM | Info | 2018/09/16 |
| 327 | pp | 13 | tev13pp_qcd_pythia8_proio | PYTHIA8 | QCD dijets for ProIO tests | SM | Info | 2018/08/27 |
| 326 | pp | 13 | tev13pp_qcd_pythia8_proio_tests | PYTHIA8 | QCD dijets for tests of ProIO | SM | Info | 2018/08/20 |
| 325 | e-p | 0.035 | gev35ep_pythia8_dis1q2ct14lo | PYTHIA8 | DIS events at Q2>1 GeV2 | SM | Info | 2018/07/25 |
| 323 | pp | 13 | tev13pp_mg5_chaHT_tbeta_hw | MADGRAPH/PY8 | H- top with H- to HW and tan(beta)=1-7 | Exotics | Info | 2018/06/13 |
| 322 | pp | 13 | tev13pp_mg5_chaHT_tbeta_tb | MADGRAPH/PY8 | H- top with H- to tb and tan(beta)=1-7 | Exotics | Info | 2018/06/13 |
| 321 | pp | 13 | tev13pp_mg5_chaHW_tbeta_tb | MADGRAPH/PY8 | H+ W- with H+ decay to t-bbar tan(beta)=1-7 | Exotics | Info | 2018/06/06 |
| 320 | pp | 13 | tev13pp_mg5_chaHW_tbeta_hw | MADGRAPH/PY8 | H+ W- with H+ decay to HW for tan(beta)=1-7 | Exotics | Info | 2018/06/06 |
| 318 | pp | 13 | tev13pp_pythia8_gamgam | PYTHIA8 | Higgs to gamma gamma | SM | Info | 2018/04/20 |

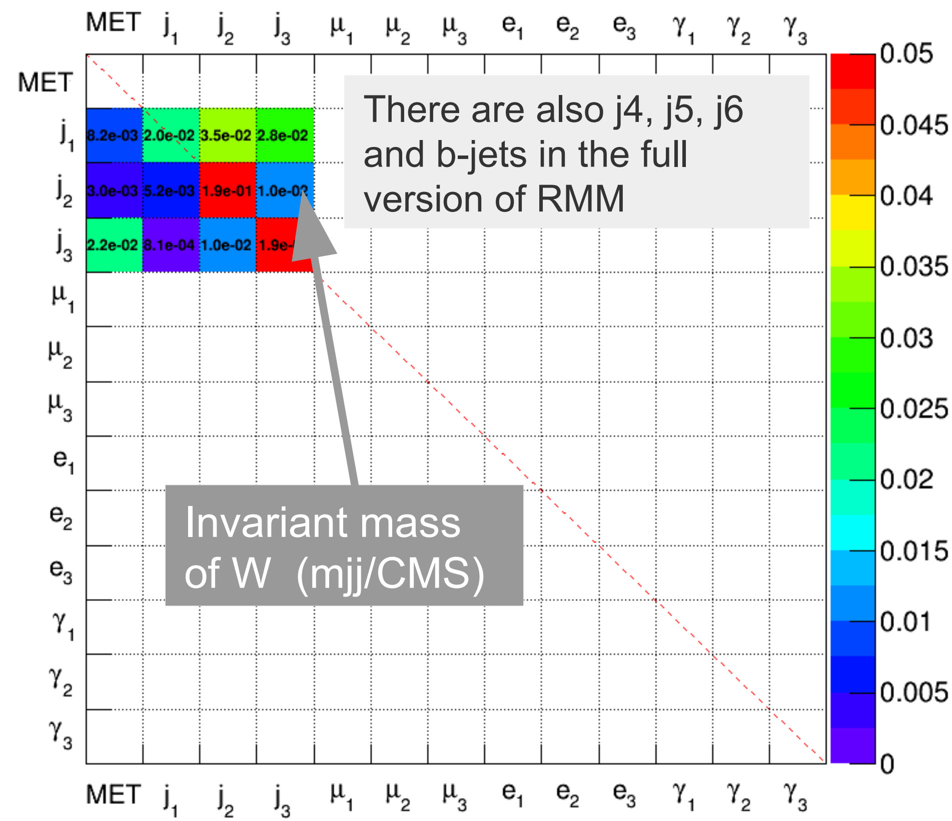
**All LO processes and all top/W/H decays enabled**

# Example: Two events with $t\bar{t}$

$t\bar{t} \rightarrow Wb\,W\bar{b} \rightarrow e\ nu\ b\ \mu\ nu\ \bar{b}$



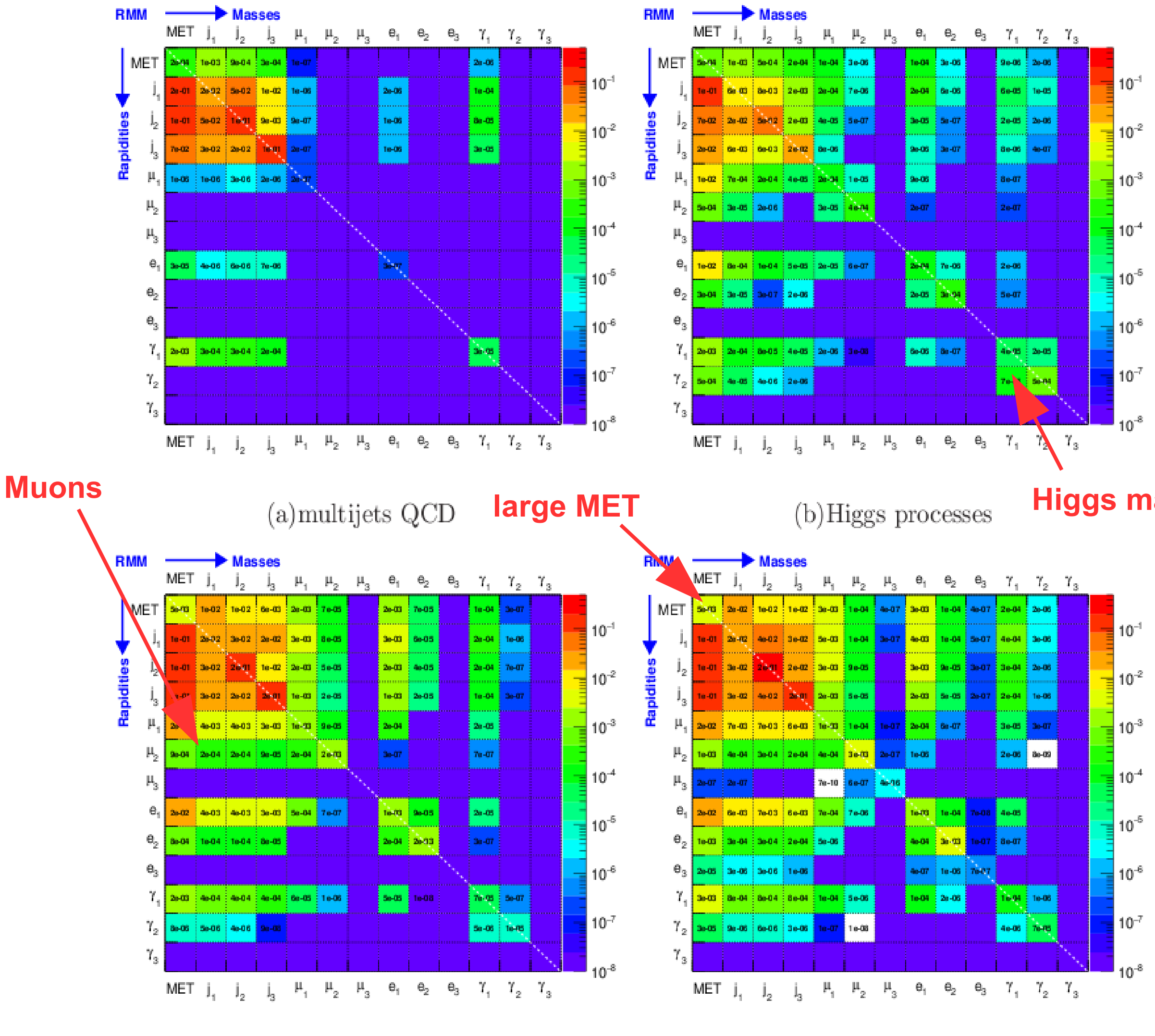

$t\bar{t} \rightarrow Wb\,W\bar{b} \rightarrow 6$ jets




Cell with MET is “fired”.
Also μ and e leptons

No MET and leptons
But many jets

# Average values of cells for 4 processes




(a) multijets QCD

(b) Higgs processes

(c) Top production

(d) $H^{+}t$ production

**Considered:**

- jets, muons, electrons, photons
- up to 3 objects

**Pythia8 simulation**:
All QCD processes included with all decays of top, W, H
50,000 events per RMM

**H+ t is similar to $t\bar{t}$**

This similarity was made intentionally by allowing H+ → W H, where H → $b\bar{b}$

Event classification using imaging of collision events. S.Chekanov (ANL) et al

# Practical example: Separation of H+ from $t\bar{t}$

disabled links during the NN training since they are extracted “features”

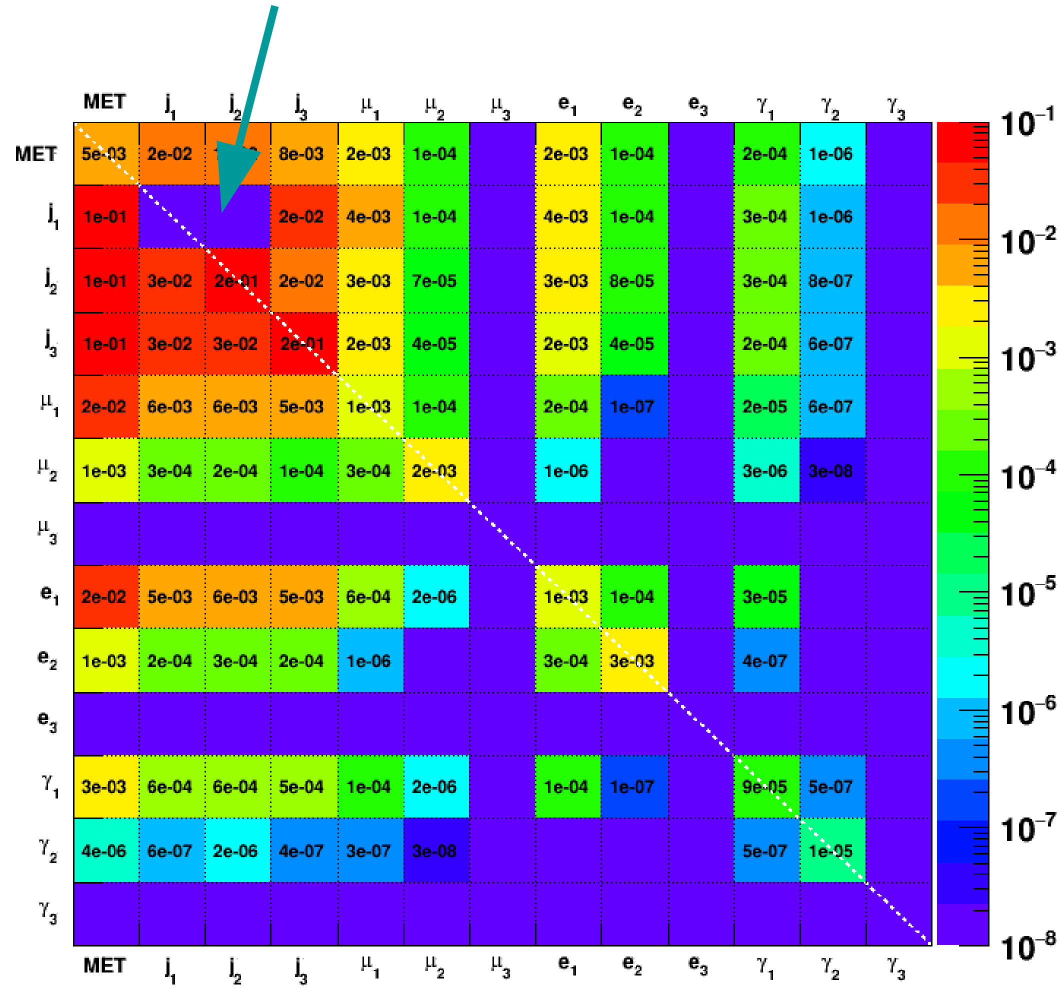
Reconstruct invariant mass of 2 jets for $t\bar{t}$ (background) and H+ (signal)

H+ (600 GeV) → H W
where H→ bbar

Should see a bump at ~600 GeV for H+ events

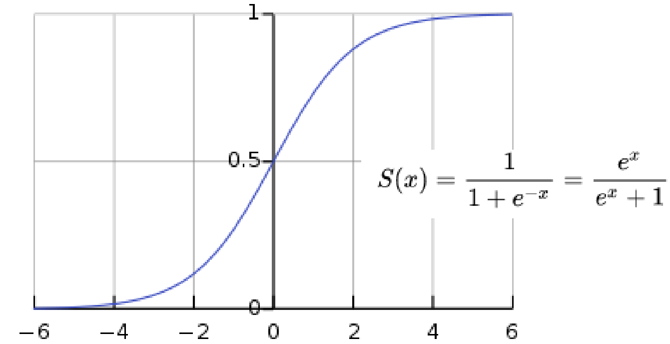
Invariant mass can be calculated from RMM itself by summing up cells at (3,2) for all RMM (and scale by 13000)

To avoid biases, disable cells (2,2) and (3,2) during the NN training



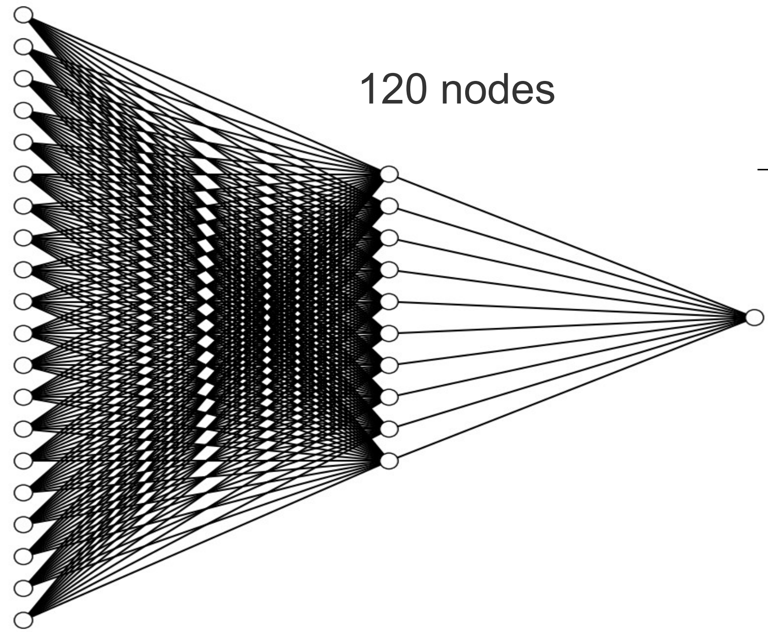
# Using RMM for ANN

A simple backpropogarion NN with “sigmoid” activation




169 nodes

120 nodes

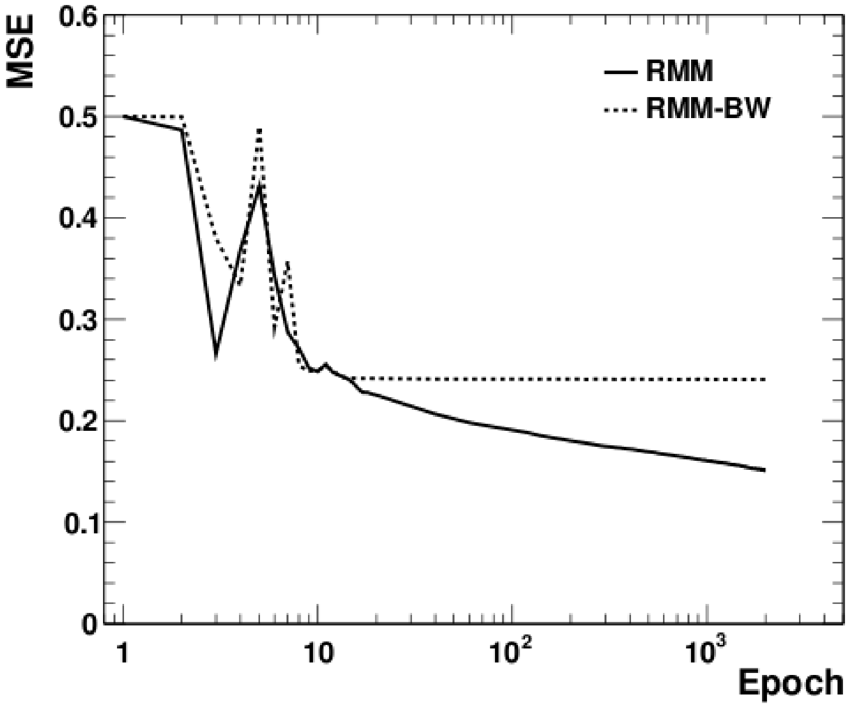


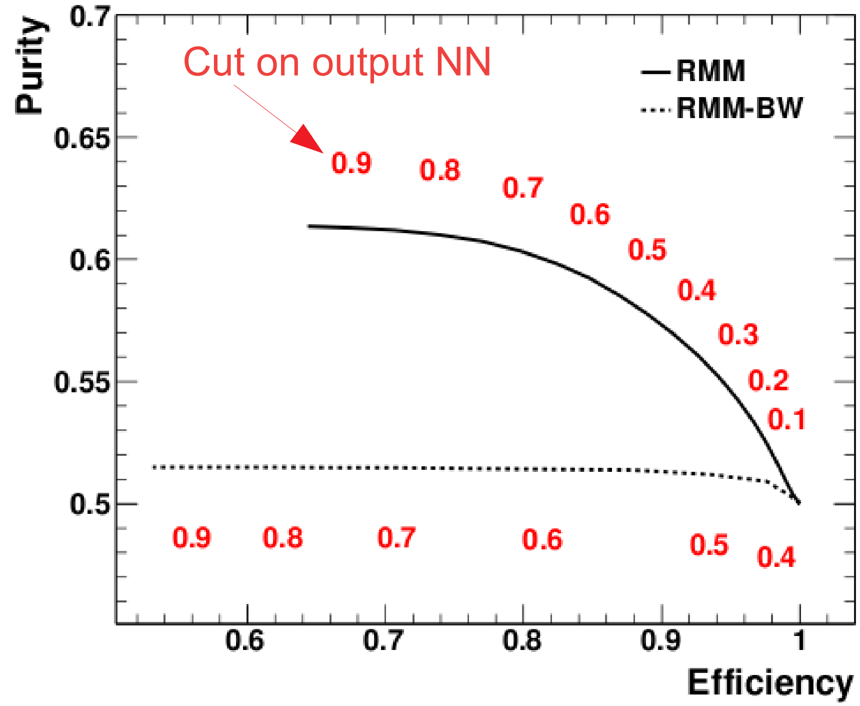
10k Pythia8 events used to create 10k RMM (13x13)

output: 0 ($t\bar{t}$) or 1 (H+)

- Use 10k events with ttbar, and 10k with H+
- Create cross validation for ANN
- Stop training when MSE < than for cross validated ANN

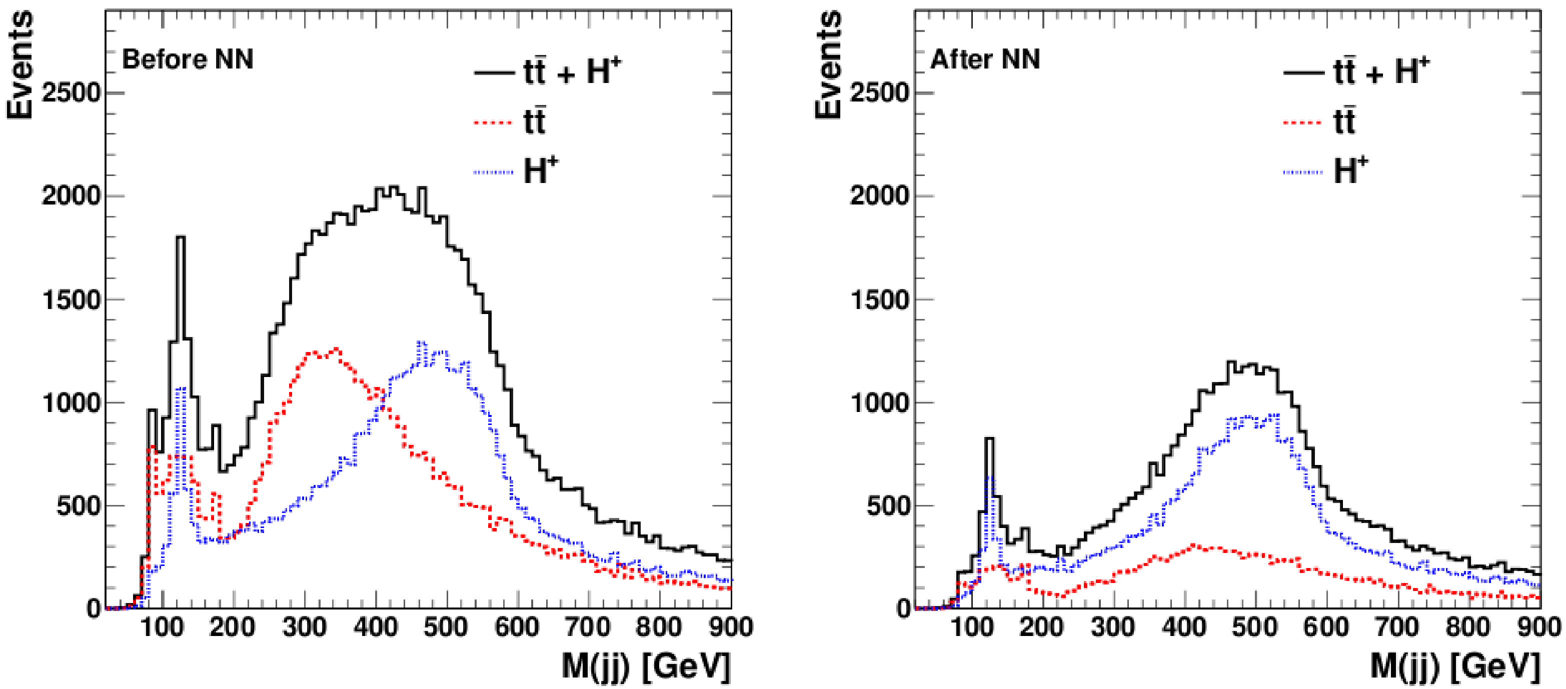
# Using RMM for neural networks




**Well trained:**
Mean Squared Error (MSE) decreased even in the case when only "activated" cells are used, i.e. without the actual values (dashed)
→ analogy with "black-and-white" images (RMM-BW)

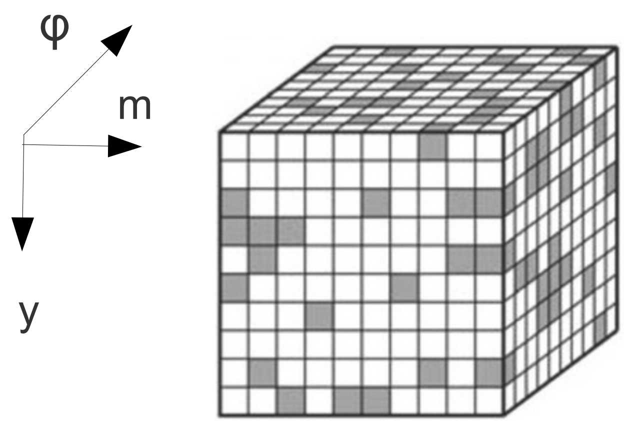
# Practical example: Separation of H+ from ttbar




- The NN based on RMM helps to reduce S/B by a factor 3.
- Signal efficiency is reduced by 30%
- Small shift for ttbar  (may require better tuning of disabled links)

# Extending RMM

- RMM includes information on single and two-particle densities
  - but no phi due to rotational symmetry)
- Can be extended to 3D matrices to include φ,  3-particle densities etc.




**Plus:**

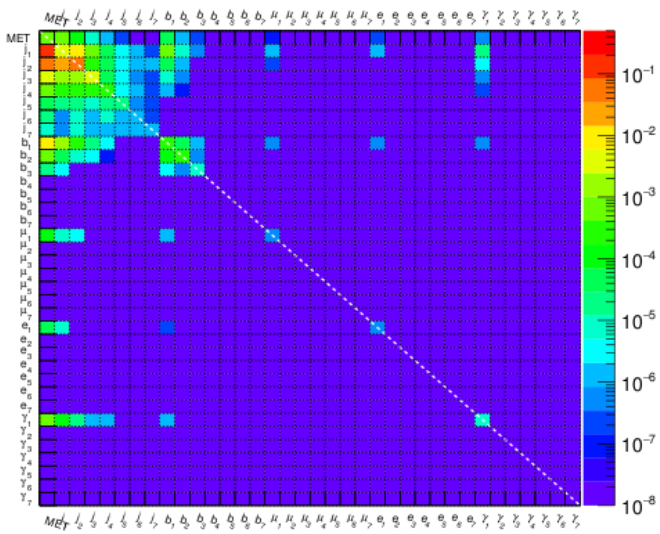
- Add tau, leptons with + and – charges (separately), b-jets
- Increase multiplicity of each object to ~10-20 (empty cells are not stored)
- Add more complex (and well reconstructed) types: J/Phi, W, Z, Higgs

# Generic event identification

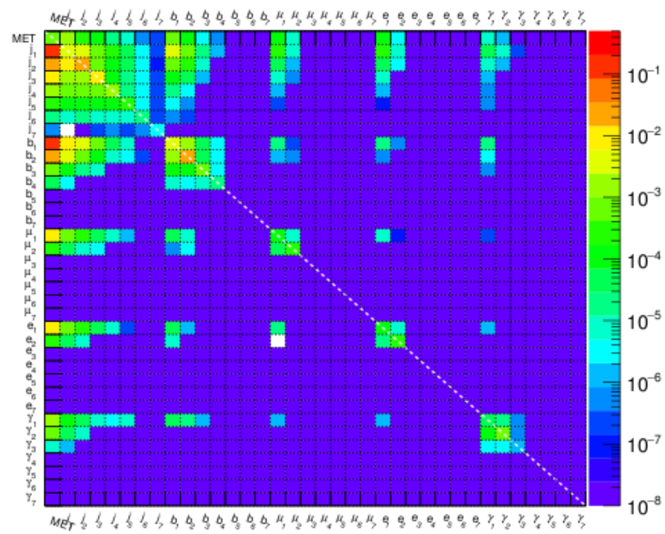
- Premise of the RMM - generality. Includes single & 2-particle densities
- No need to hand-pick input variables for every event topology/decay
- Good choice for general event classifiers?

**Example:**

- 5 processes with (1) SM QCD (2) Higgs (3) H+ (4) ttbar (5) Double bosons
- Create RMM using Np=7 and 6 objects using b-jets



Multi-jet QCD



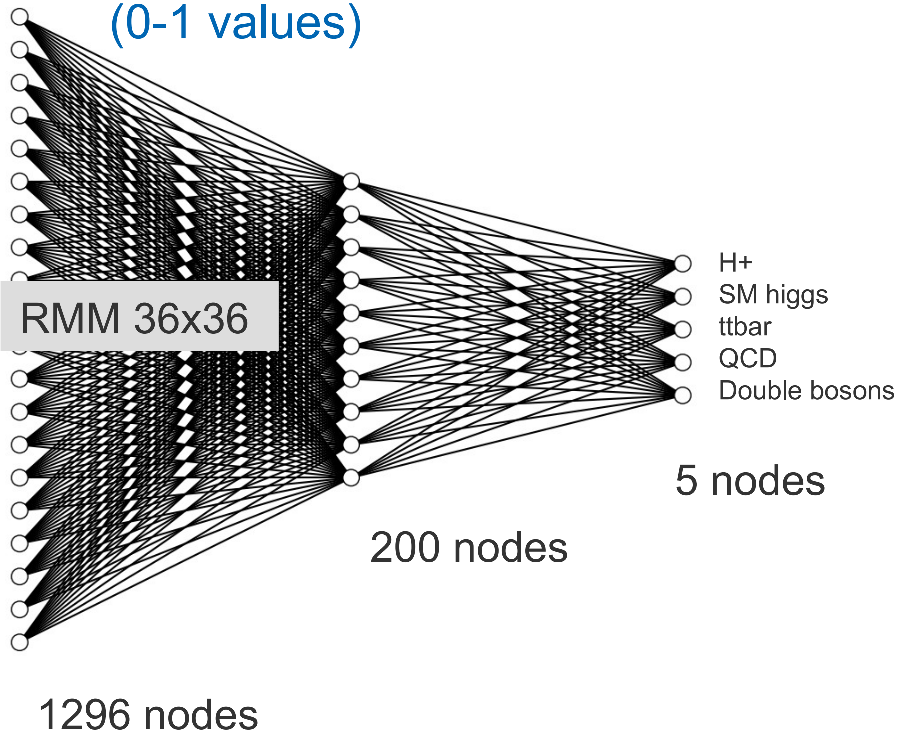
Higgs productions (all decays)

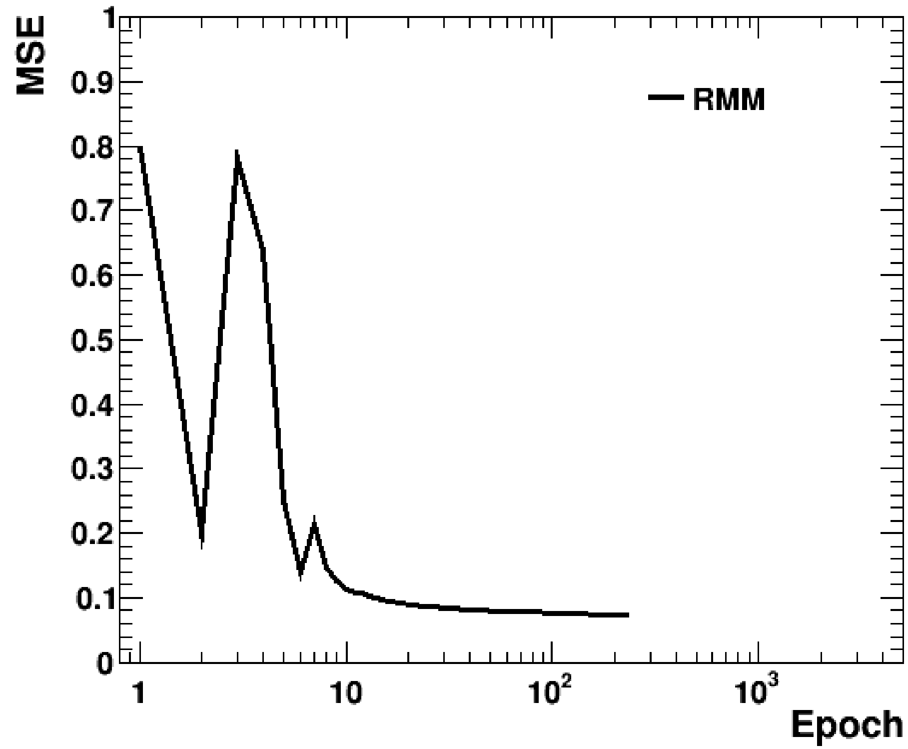
Shows average values of cells for 50k events

Event classification using imaging of collision events. S.Chekanov (ANL) et al

# ANN training

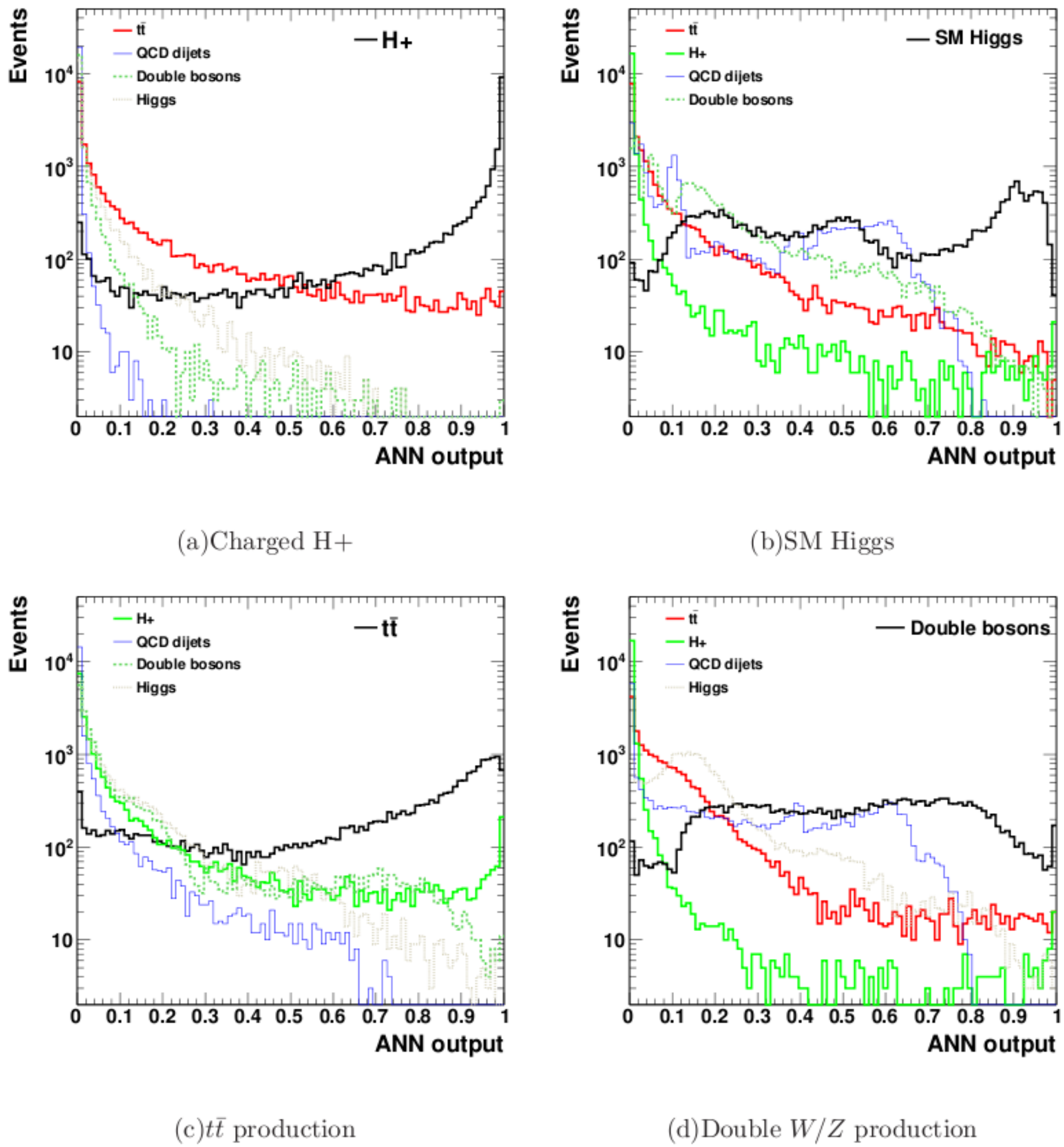
Backpropogation NN with Signoid function, 5 outputs for each process (0-1 values)




Well trained:
Mean Squared Error (MSE) decreased from 0.8 to 0.07

# Result of NN training




(a)Charged H+ (b)SM Higgs

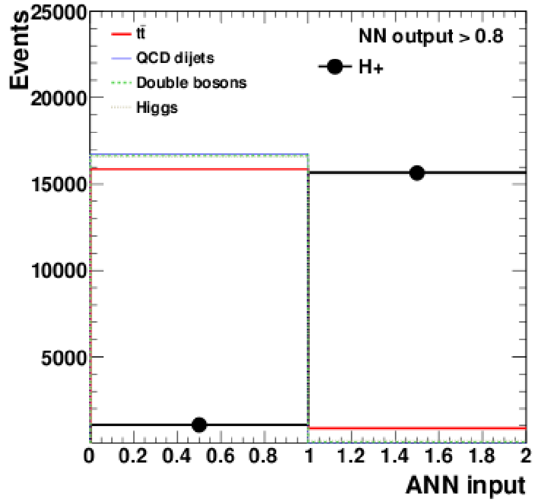
(c)$t\bar{t}$ production (d)Double $W/Z$ production

Good event separation of “signal” events (black line) from other processes

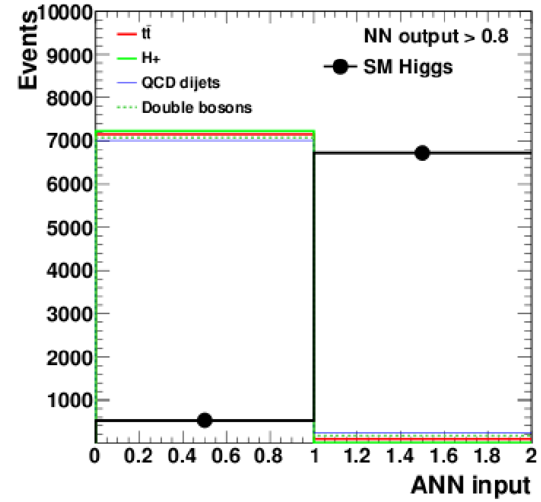
Event classification using imaging of collision events. S.Chekanov (ANL) et al

# Purity of event identification

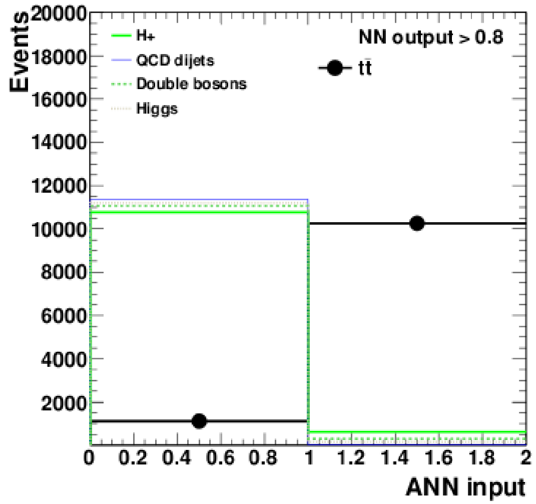
Purity of event classification is 80%-90% assuming 0.5 cut on output node



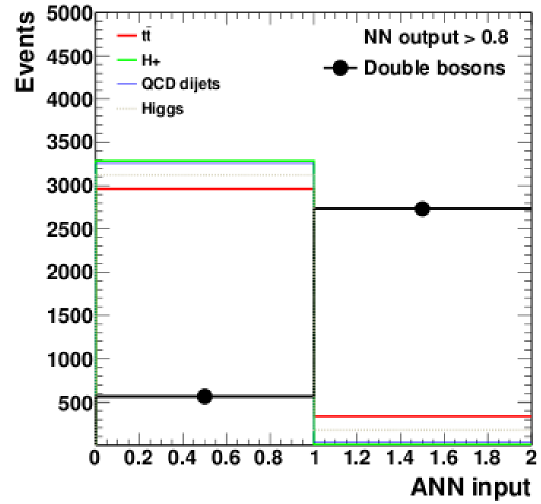

(a)Charged H+




(b)SM Higgs



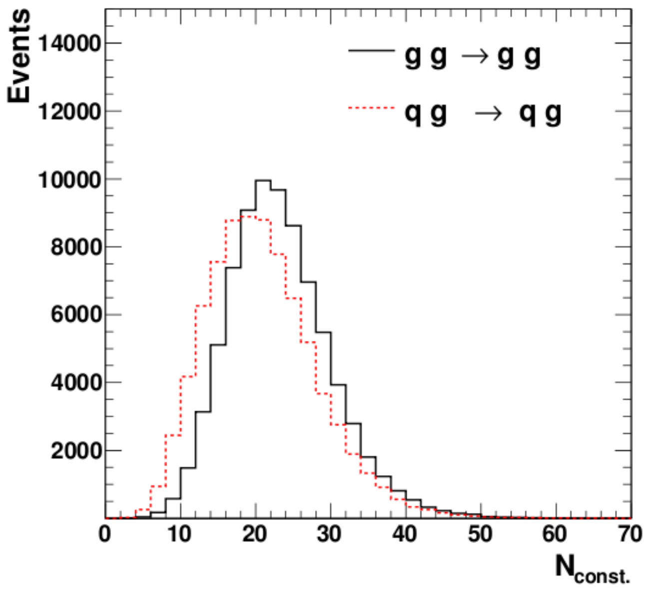

(c)$t\bar{t}$ production




(d)Double $W/Z$ production

# QCD dijet separation: Challenging case!

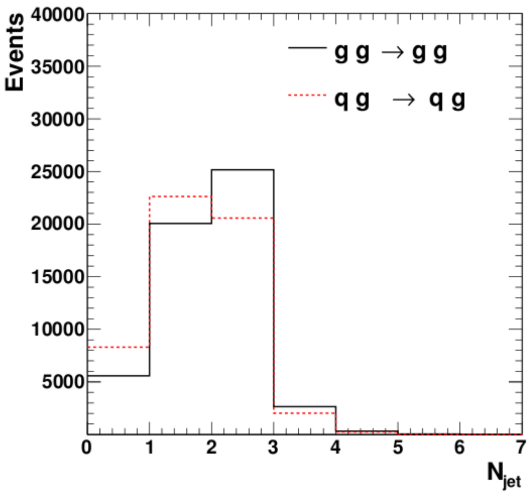
Separate **gg** from **qg** final states (dijets) → Distributions are nearly identical.
Presence of **g** instead of **q** leads to broader jets and changes in jet kinematics / shape

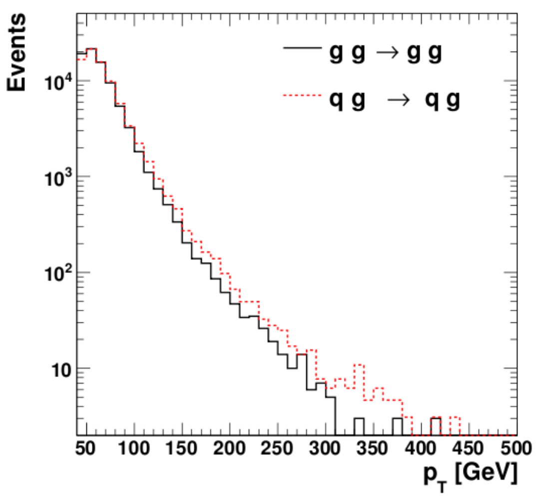



Well-known difference: Number of jet constituents is larger for gluon jets than for quark jets due to difference in color factors ($C_A$ =3 vs $C_F$ = 3/4)

But there are many other distributions that can be used for ANN. How to choose them?

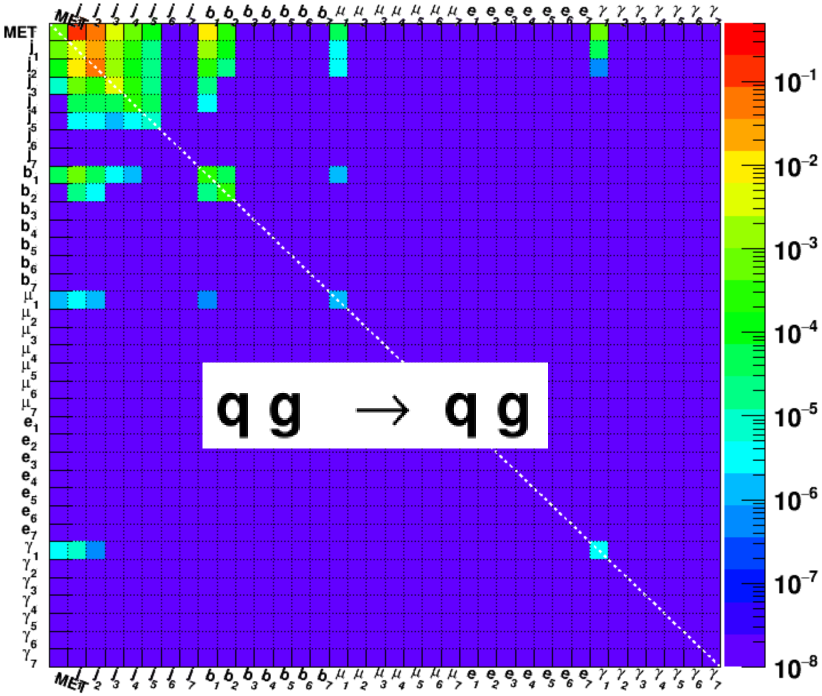
Use hand-crafted variables using Pick-and-Use approach?

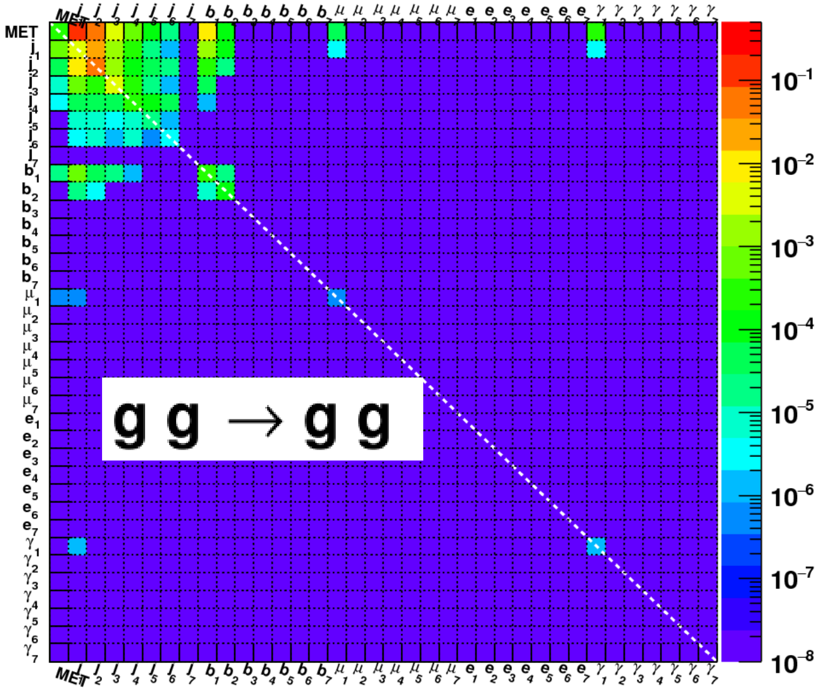



Event classification using imaging of collision events. S.Chekanov (ANL) et al

# Creating RMM for gg and qg events




**gg** process compared to **qg** has:

- softer pT
- more jets
- reduced photon rate

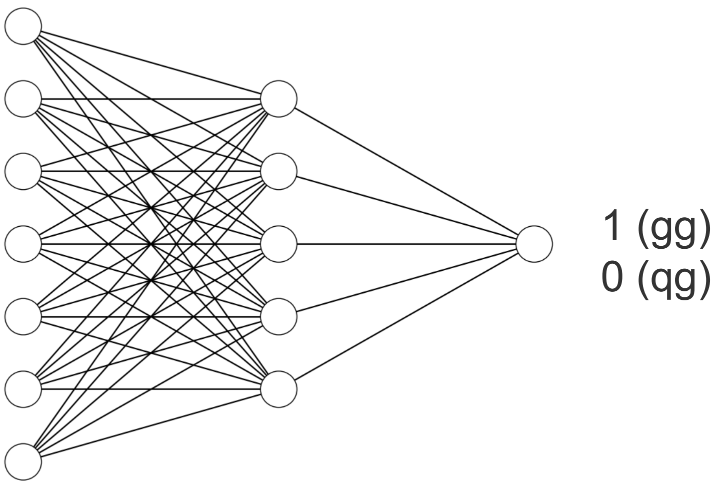
..

Event classification using imaging of collision events. S.Chekanov (ANL) et al

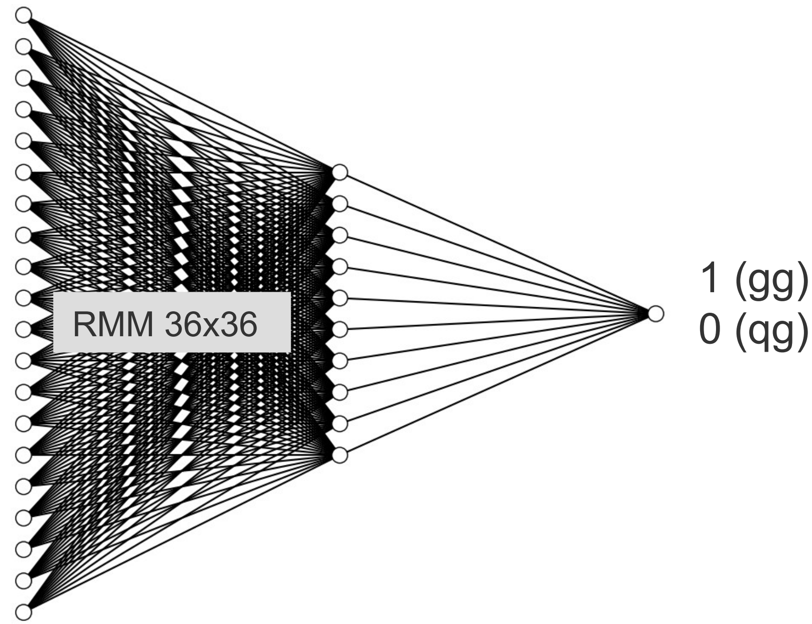
# QCD dijet separation: Challenging case!

Two approaches using machine learning:

**Traditional PaU**
- hand-crafted input variables (7 nodes)
- hidden layer (5 nodes)
- output with 1 (gg) or 0 (qg)



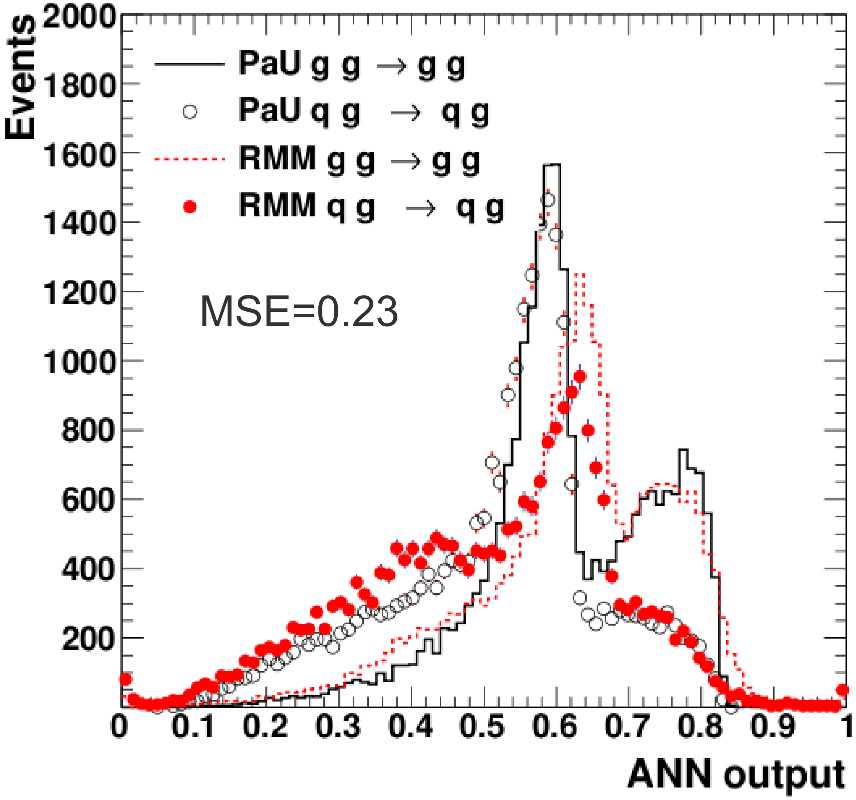

**RMM**
- RMM matrix as input (36x36+2)
- hidden layer (200 nodes)
- output with 1 (gg) or 0 (qg)



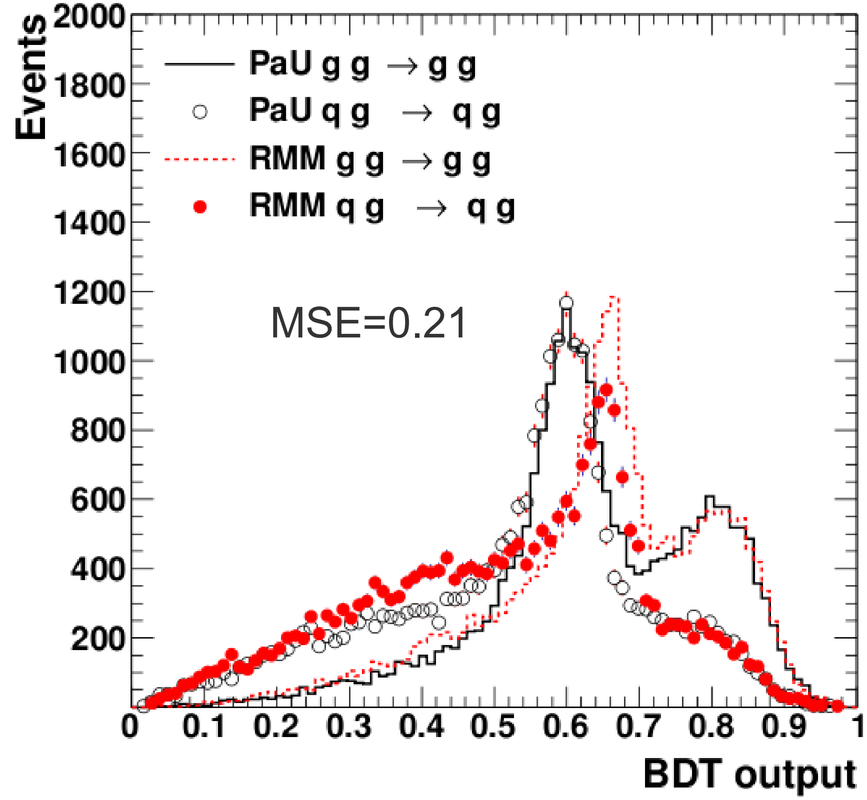

Alternatively: Use Boosted Decision tree (BDT) using PaU and RMM
100 trees, depth 7, stochastic gradient (arXiv:1609.06119)

# Machine learning for gg and qg separation

**AAN**




**BDT**




- Reasonable separation of **gg** from **qg**
- RMM over-performs “hand-crafted pick-and-use” (PaU) method
- RMM has separation purity 67% vs 63% for PaU assuming cut at 0.5
- BDT confirms this conclusion

Event classification using imaging of collision events. S.Chekanov (ANL) et al

# Conclusions

- RMM improves event classification due comprehensive (nearly independent) single and two-particle densities for all particle types

- Same RMM transformation can be plugged into very different problems to produce good results with minimal tweaking
  - Unless you do something exotic and single and double densities of reconstructed objects are not sufficient

- Unlike hand-crafted inputs for machine learning, RMM can identify events with rather unexpected features. For example, **qg** events have an enhanced production of isolated photons. This contributes to RMM, but often escape attention
  - No need to worry about different decay channels (and their kinematics)

- If you are interested in a package that transforms events to RMM – contact me