

Ricerca di nuova fisica prodotta in associazione ad un quark top ed energia trasversa mancante identificati tramite algoritmi di machine learning

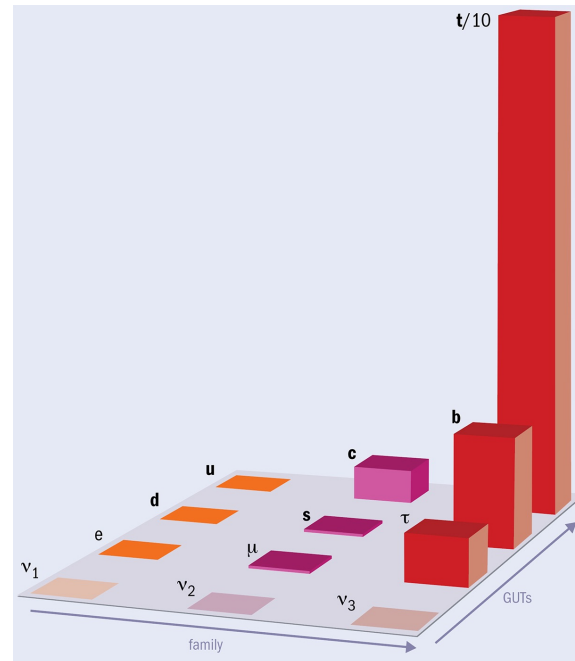
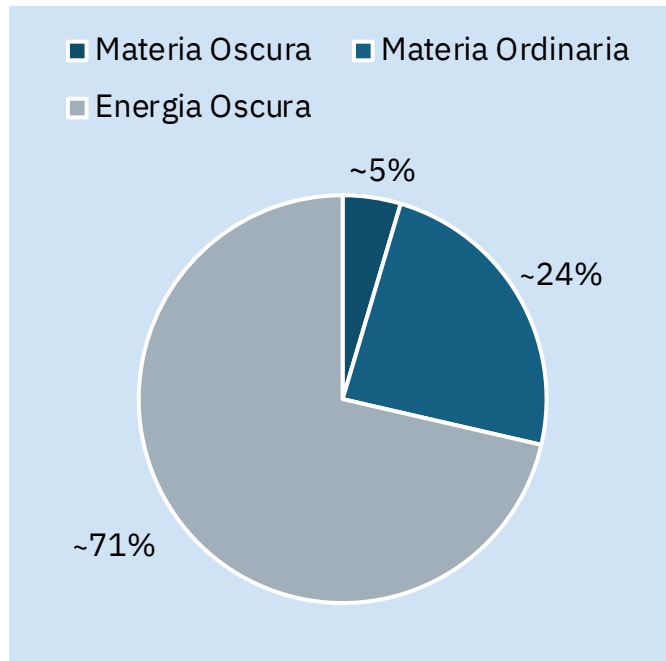
Benedetta Argiento, per la collaborazione CMS

IFAE 2025 – Cagliari, 09 Aprile 2025

Introduzione

Domande aperte del Modello Standard:

- *formali*: gerarchia e interazione gravitazionale
- *sperimentali*: materia ed energia oscura



⇒ *Ricerca di nuova fisica:*

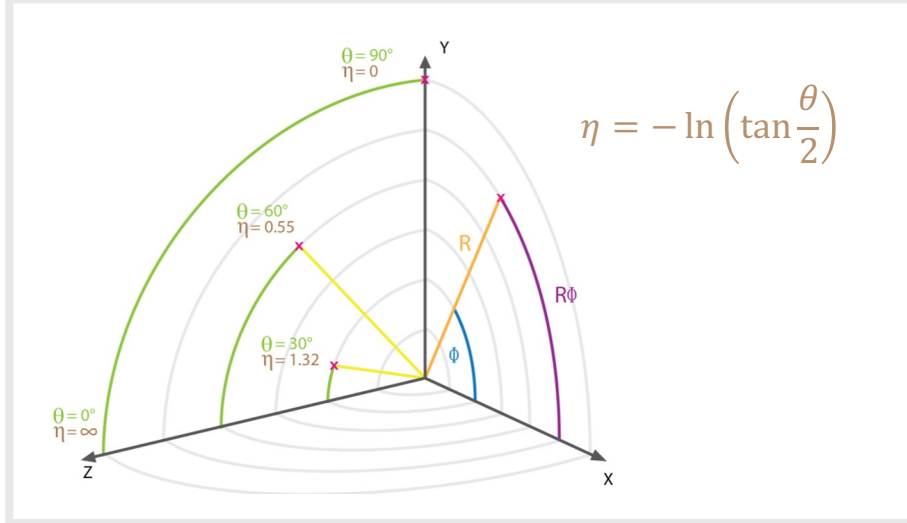
- Modelli che prevedono nuovi bosoni o fermioni pesanti (VLQ)
- Ricerca di Materia Oscura (Dark Matter – DM) con Simplified Models

⇒ *Segnatura interessanti:*

- Quark di terza generazione → rapporto privilegiato con il bosone di Higgs
- Particelle non rivelabili → energia trasversa mancante

Ricostruzione Top quark

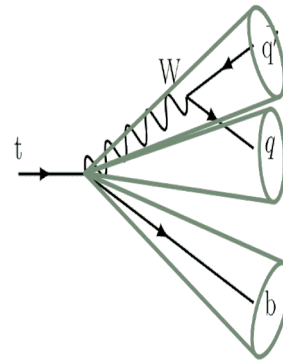
All'interno di CMS i prodotti dell'adronizzazione vengono identificati con dei coni nel piano (η, ϕ)



Algoritmo **anti-KT**:

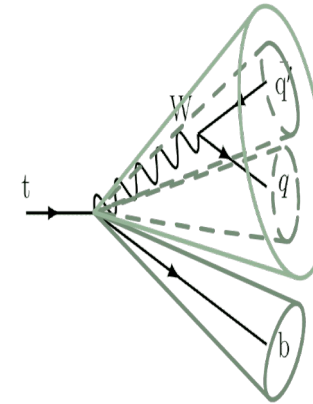
- $R=0.4 \Rightarrow$ Jet stretto (AK4)
- $R=0.8 \Rightarrow$ Jet largo (AK8)

RESOLVED



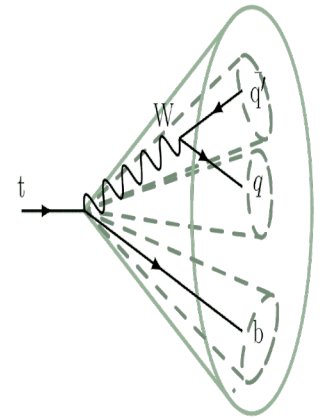
2 jet Ak4 e b-jet

SEMI-RESOLVED



W-jet e b-jet

MERGED



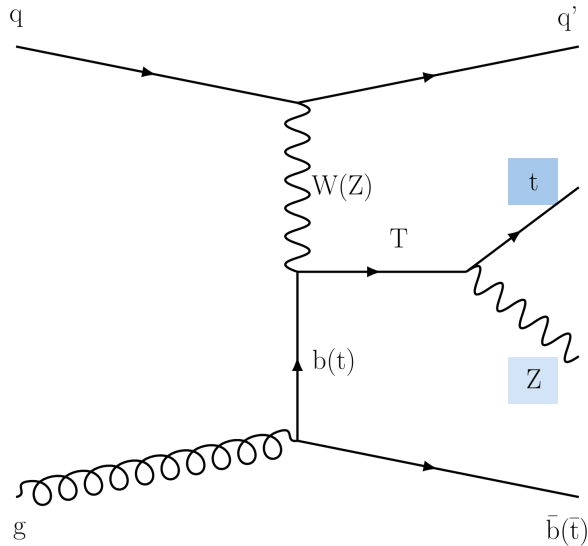
1 jet AK8

p_T

Vector-Like Quark/ $T \rightarrow tZ(\nu\nu)$

CMS-B2G-19-004

Stato finale:



- un forward jet (q')
- quark b taggato

Decadimento adronico del quark top:

- *fully-resolved*
- *partially-merged*
- *fully-merged*

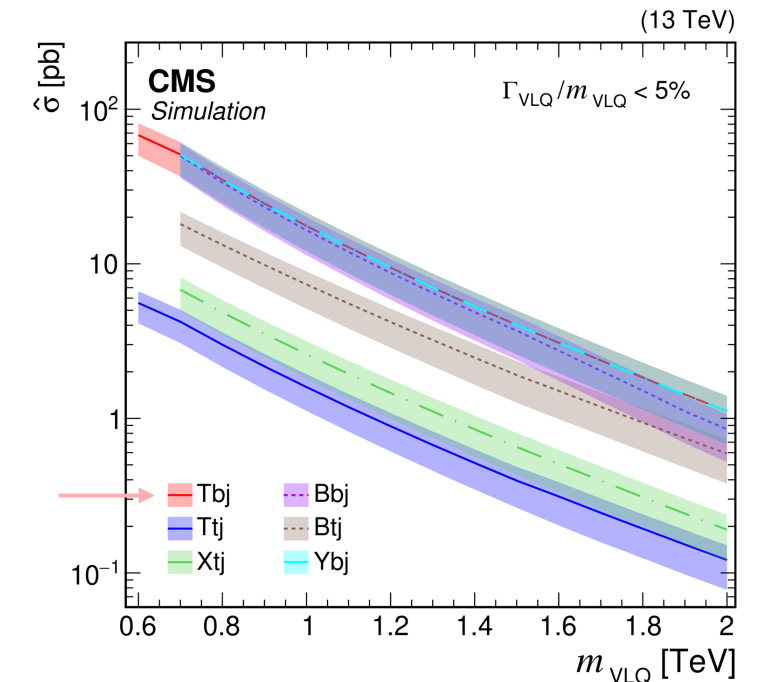
- $Z \rightarrow \nu\nu$: energia mancante nello stato finale

Background:

1. $Z \rightarrow \nu\nu + jets$
2. $W \rightarrow l\nu + jets$
3. Coppe $t\bar{t}$

Diverse ipotesi di massa per T
($0.6 \rightarrow 1.8 \text{ TeV}$) e per la larghezza
di decadimento (1 al 30% di m_T)

CMS-EX0-23-006



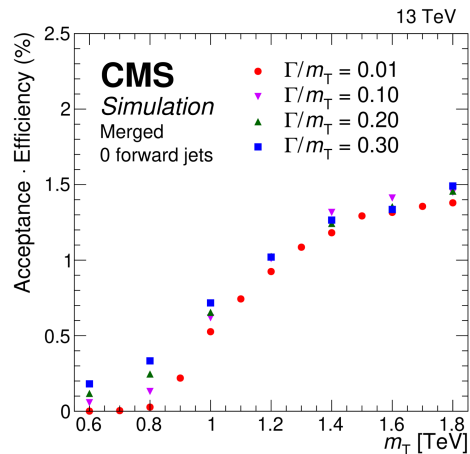
Strategia di analisi $/T \rightarrow tZ(\nu\nu)$

CMS-B2G-19-004

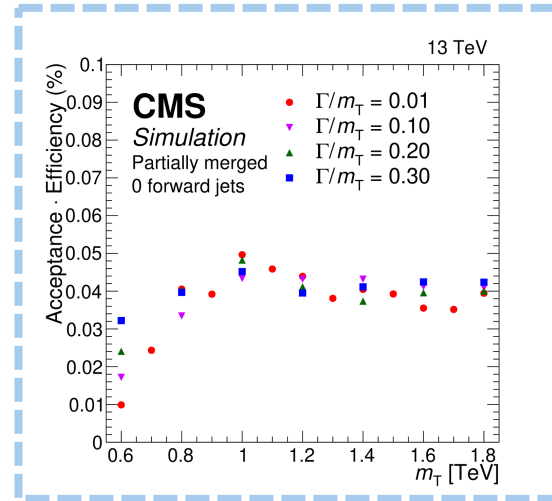
- $p_T^{miss} > 200 \text{ GeV} \rightarrow$ neutrini
- $\min \Delta\Phi_{\{MET, jet\}} > 0.6 \rightarrow$ rimozione eventi QCD
- Veto elettroni e muoni

Gerarchia ottimizzata per raggiungere il miglior limite di esclusione previsto sull'intero intervallo di massa:

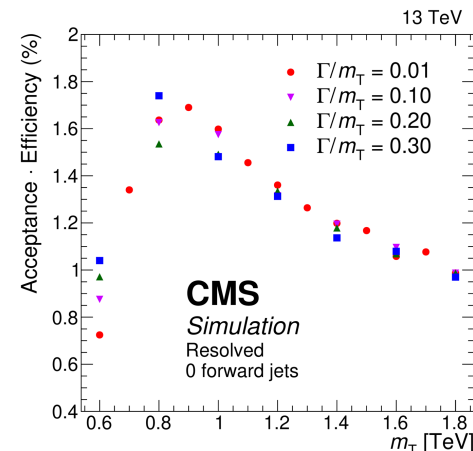
1.



2.



3.



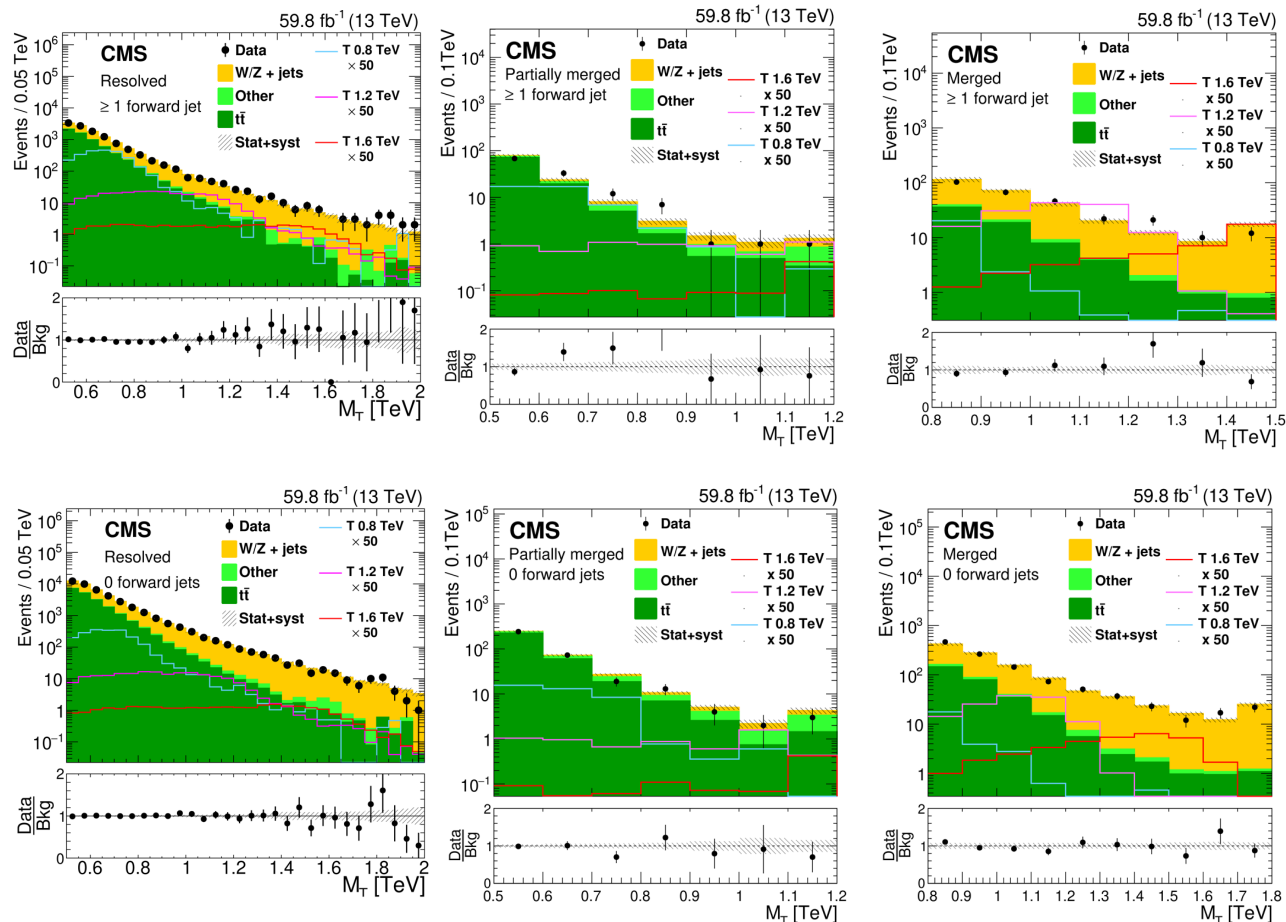
Nel caso *partially merged* c'è un ordine di grandezza in meno nell'efficienza di selezione rispetto al caso *resolved* e *merged*

⇒ necessità di migliorare la ricostruzione del top

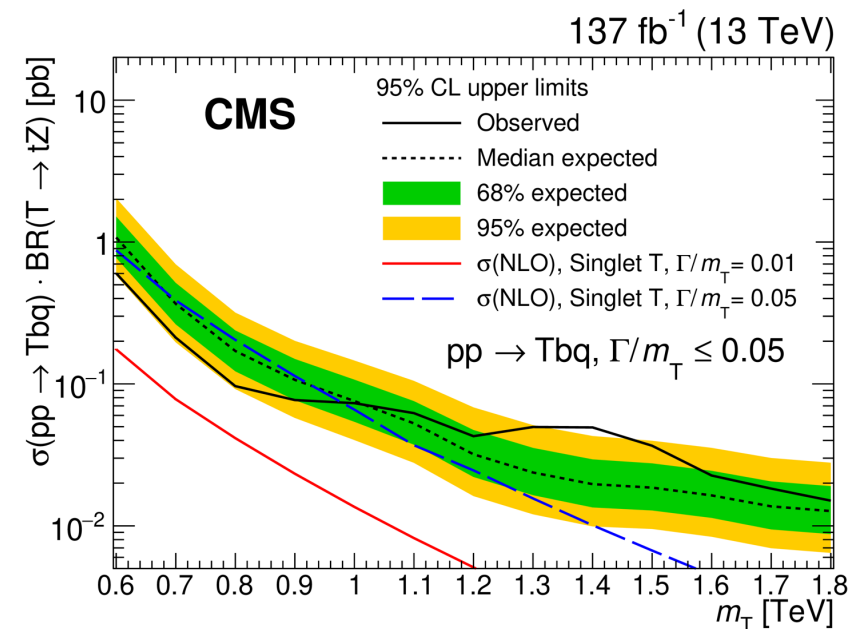
Risultati $T \rightarrow tZ(\nu\nu)$

CMS-B2G-19-004

#forward jet



p_T

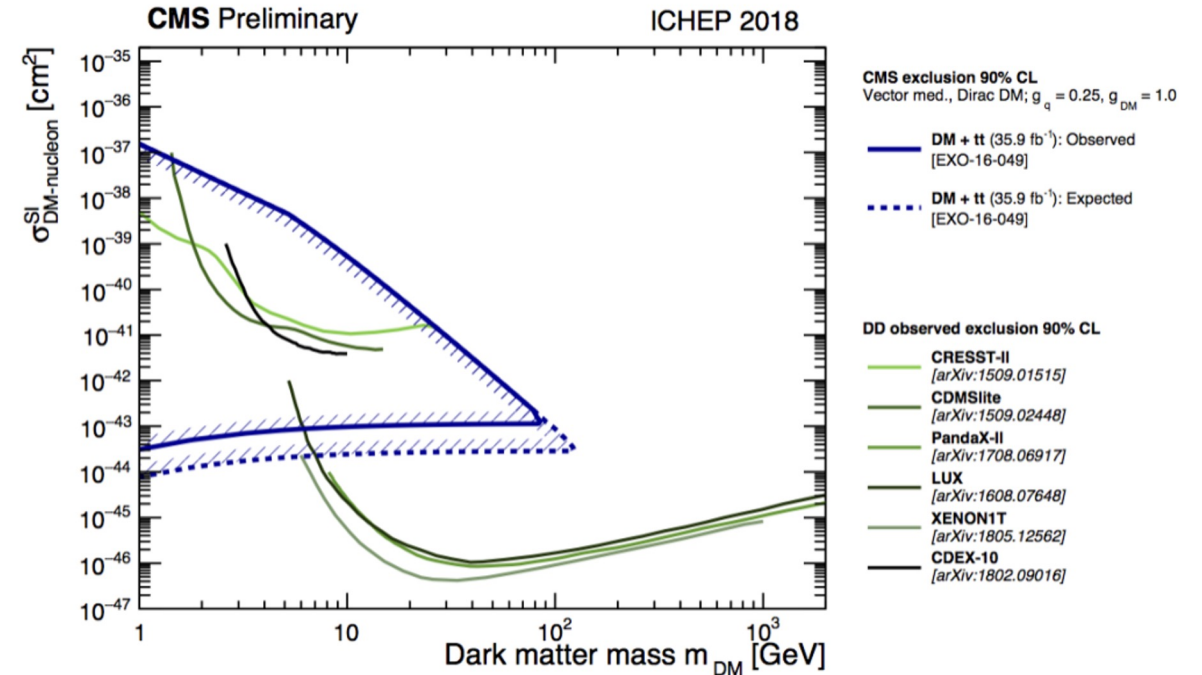
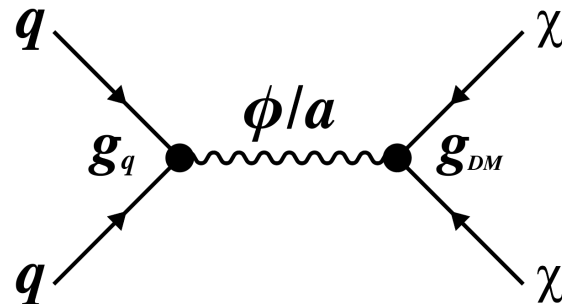
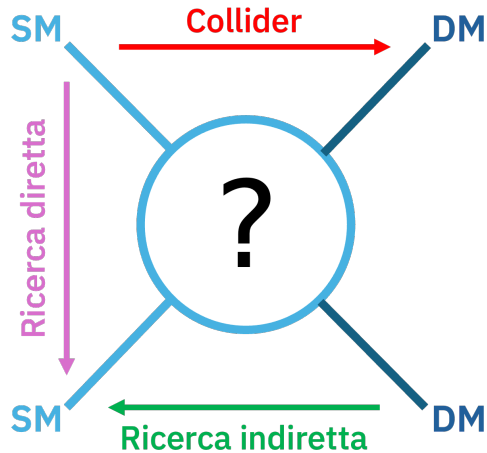


Fit simultaneo sulla massa nelle 6 regioni
 $\Rightarrow$ nessuna evidenza di segnale
 $\Rightarrow$ upper limit: 2,5 σ per $m_T = 1,4 \text{ TeV}$ a larghezza ridotta

Ricerca di DM ai collider/Simplified models

Ipotesi:

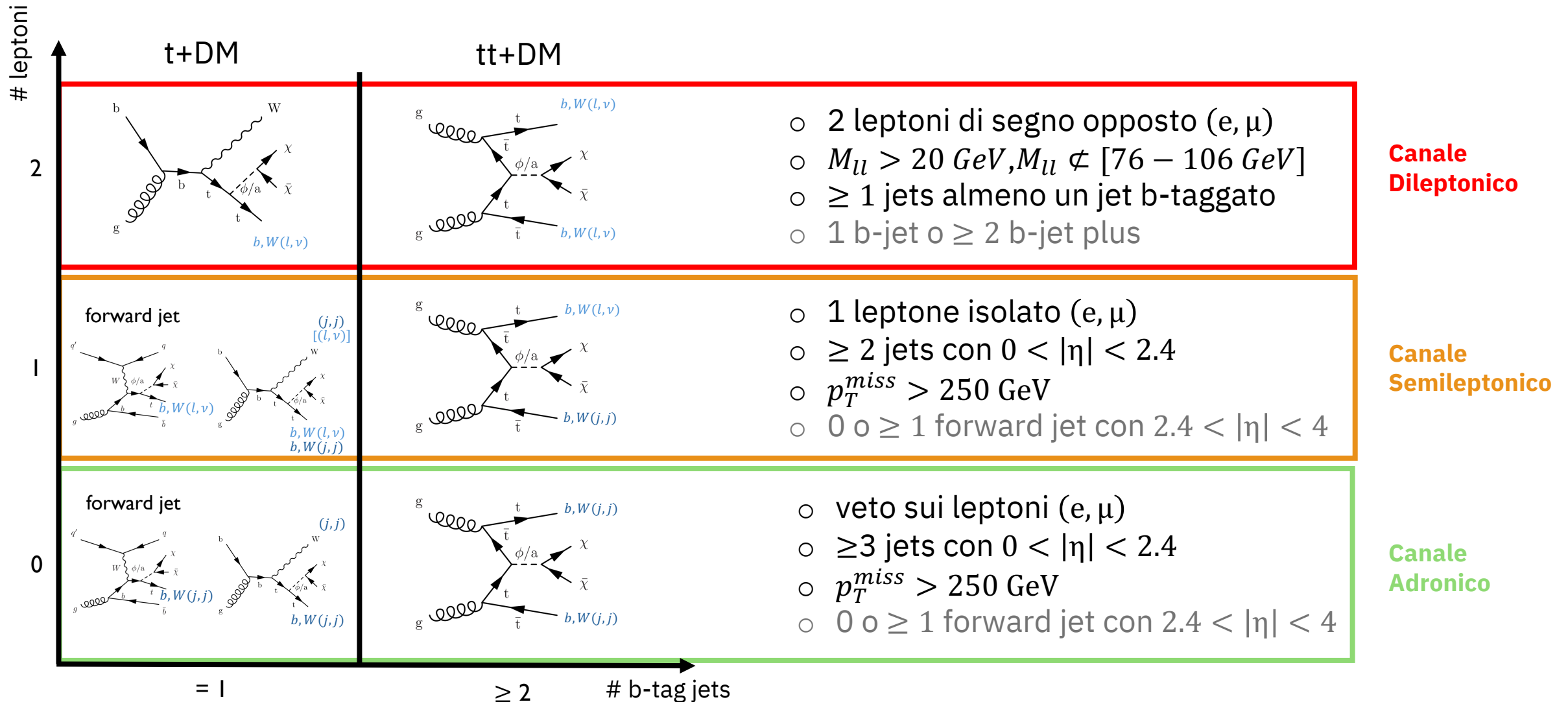
- ϕ/a (scalare/pseudoscalare)
- $m_{\phi/a} = [50, 500] \text{ GeV}$
- $g_{\text{DM}} = 1$
- $m_\chi = 1 \text{ GeV}$



Limiti alle sezioni d'urto χ -nucleone in funzione della massa della particella di DM: *ricerca diretta a CMS* e *ricerca indiretta*

Ricerca a CMS $/tDM$ & $t\bar{t}DM$

CMS-EX0-22-014

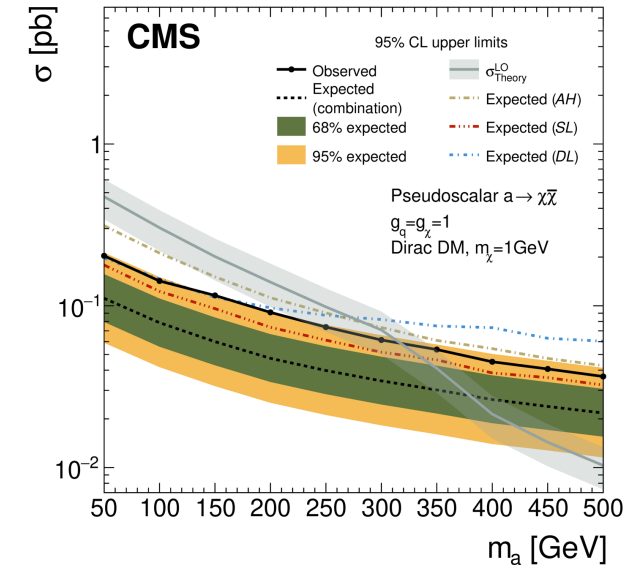
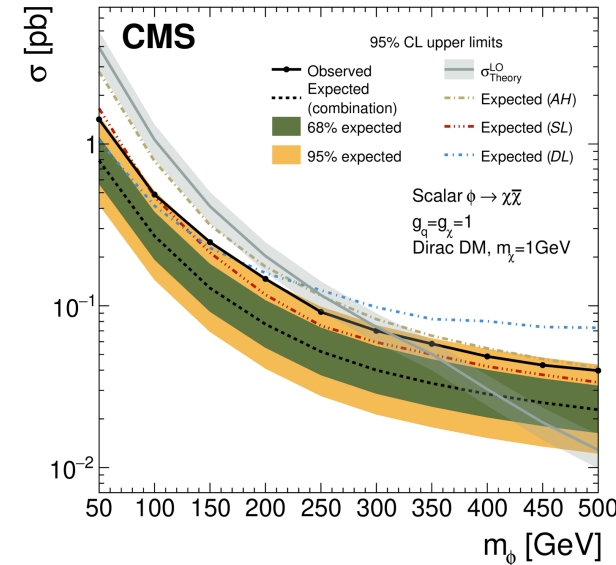
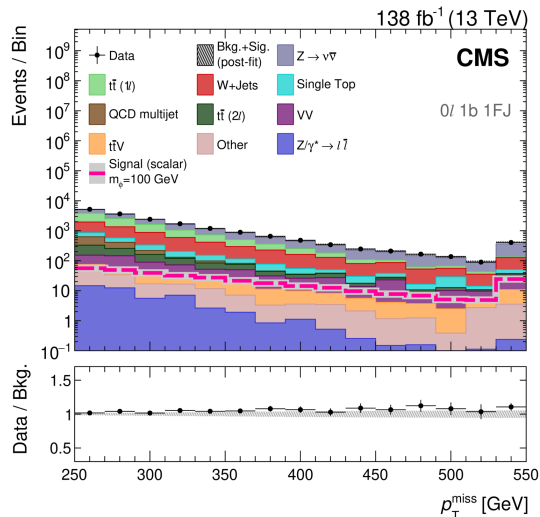
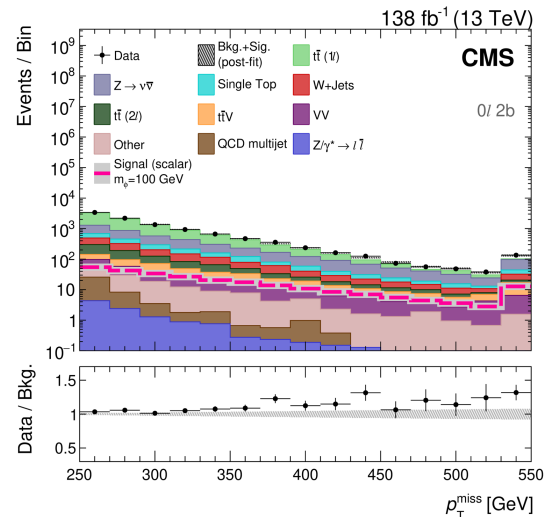
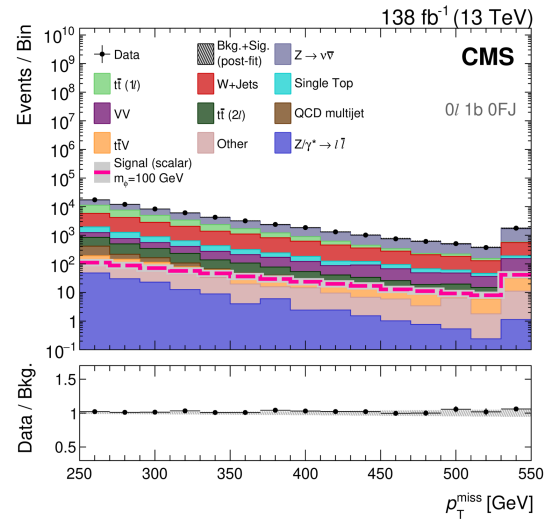


Risultati/ tDM & $t\bar{t}DM$

CMS-EX0-22-014

Fit simultaneo:

- p_T^{miss} per le regioni AH e SL;
- Neural Network per regione DL



Eccesso diffuso compatibile con il segnale è estato osservato nei dati
 $\Rightarrow$ Upper limit: *esclusione* fino a 310 GeV per mediatore pseudoscalare e fino a 320 GeV per mediatore scalare

TROTA/*Top Reconstruction: an Object Tagger Algorithm*

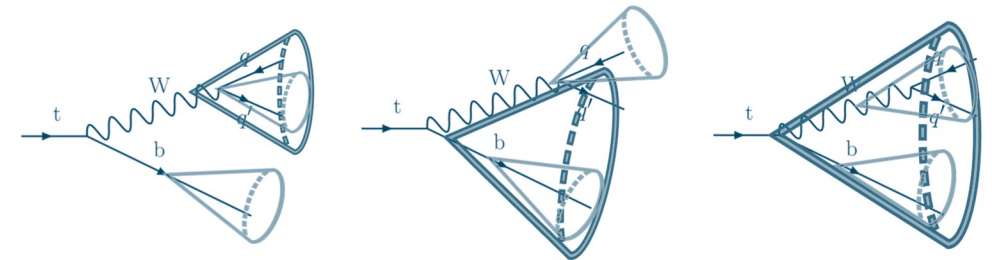
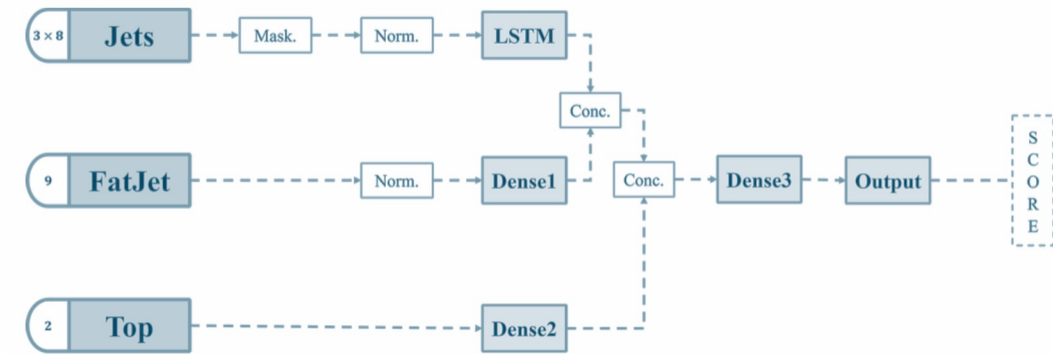
Ogni candidato top è composto da un numero variabile di oggetti:

1. **Top Resolved**: 3 AK4 jet $\rightarrow$ basso p_T
2. **Top Mixed**: combinazione di AK4 e AK8
3. **Top Merged**: 1 jet AK8 $\rightarrow$ Top tagger di ParticleNet

La **DNN** per la categorizzazione del top mixed è elaborata due diversi input tenendo conto delle diverse configurazioni possibili:

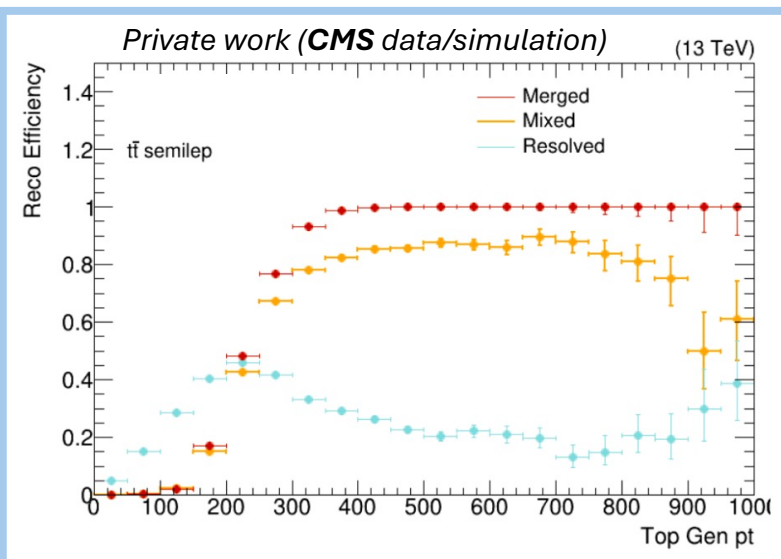
- Variabili cinematiche e di alto livello dei **jet**
 $\Rightarrow$ **LSTM**;
- Variabili cinematiche e di alto livello del **fatjet**
 $\Rightarrow$ **DNN**;
- i due output vengono combinati in una *DNN comune*, che fornisce un unico risultato.

$\Rightarrow$ **score** sulla categorizzazione del top.

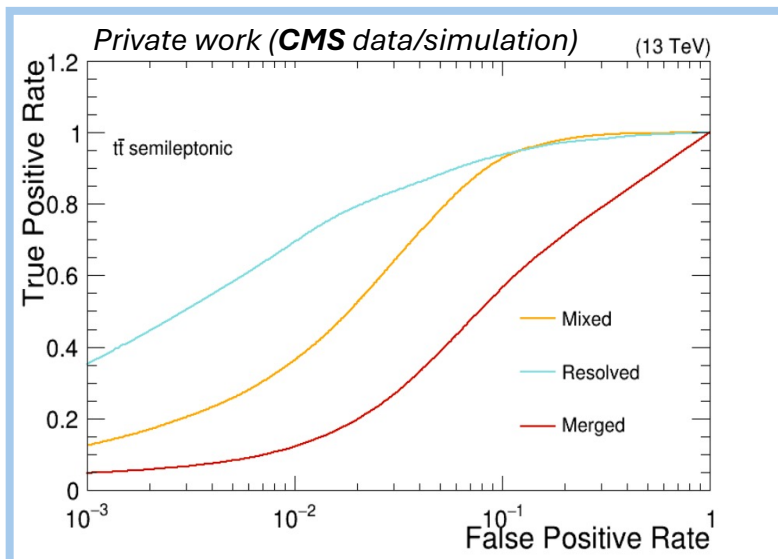


TROTA/prestazioni su $t\bar{t}$ semilep

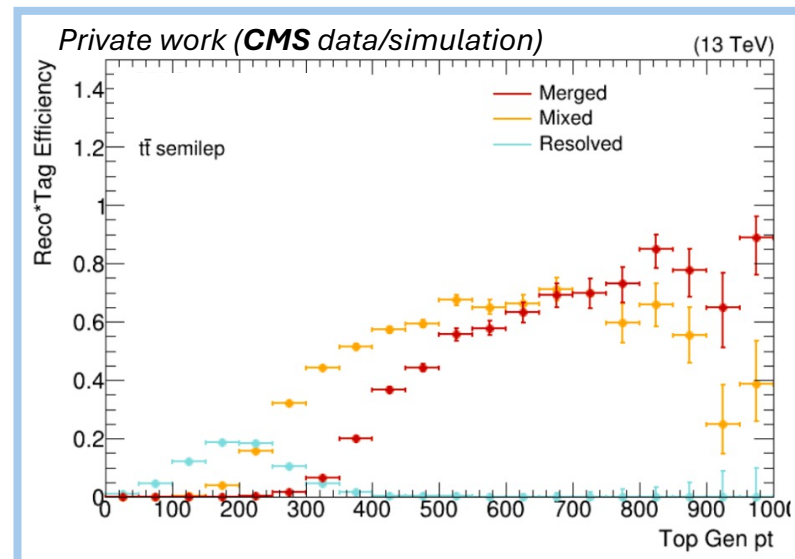
work in progress



Ricostruzione del quark top con best score



Identificazione: si evidenzia una migliore identificazione delle configurazione resolved e mixed rispetto alla configurazione merged ricostruita solamente con ParticleNet



Efficienza complessiva con tagging efficiency con un mistag rate del 1% $\Rightarrow$ con TROTA si evidenzia una migliore ricostruzione della configurazione resolved e mixed anche a basso p_T

tDM + TROTA/strategia di analisi

work in progress

Preselezione:

- $p_T^{miss} > 200$ GeV
- No leptoni con $p_T > 25$ GeV e $|\eta| < 2,5$
- forward jet $p_T > 50$ GeV e $|\eta| > 3$

Selezione con la ricostruzione del quark top di TROTA:

❑ Configurazione **Resolved**:

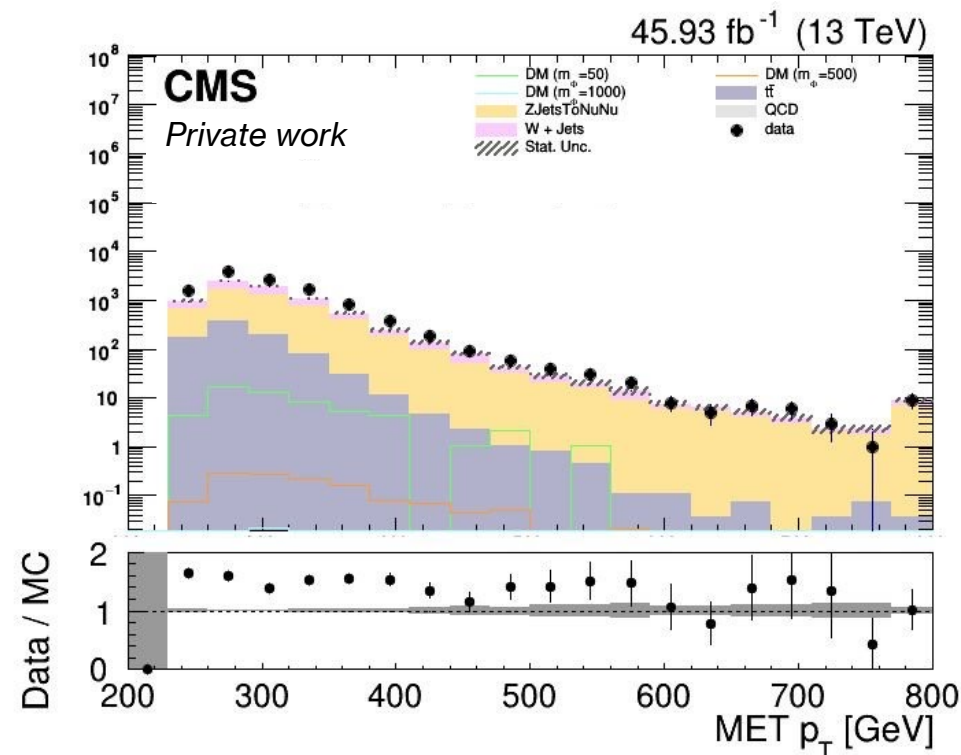
- almeno un top resolved con $p_T > 160$ GeV
- nessun top mixed e merged

❑ Configurazione **Semi-Resolved**:

- almeno un top mixed con $160 < p_T < 350$ GeV
- nessun top merged e al più un top resolved

❑ Configurazione **Merged**:

- almeno un top merged con $p_T > 350$ GeV
- al più un top mixed e un top resolved



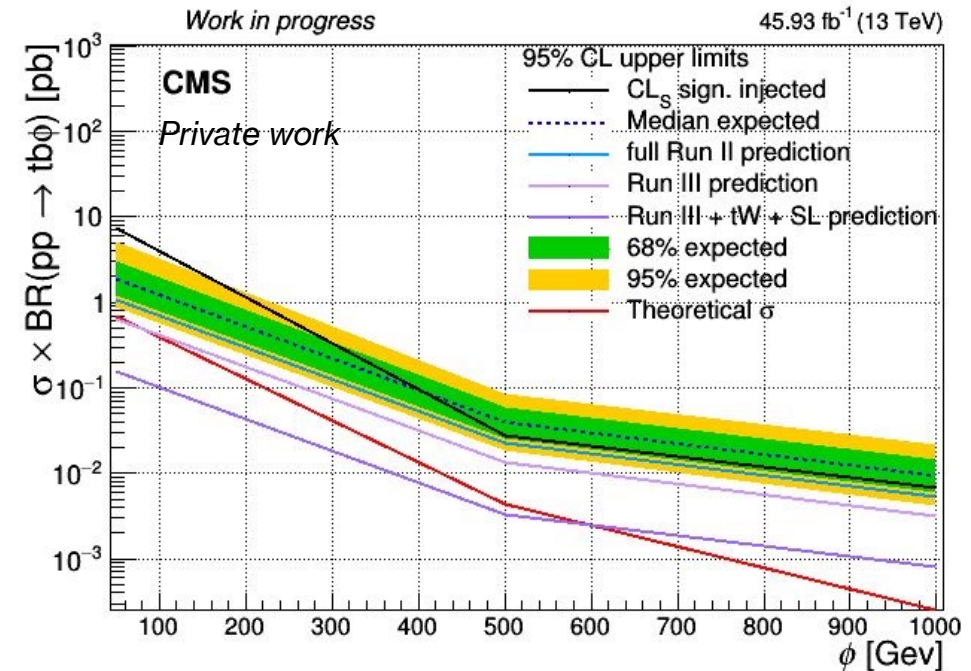
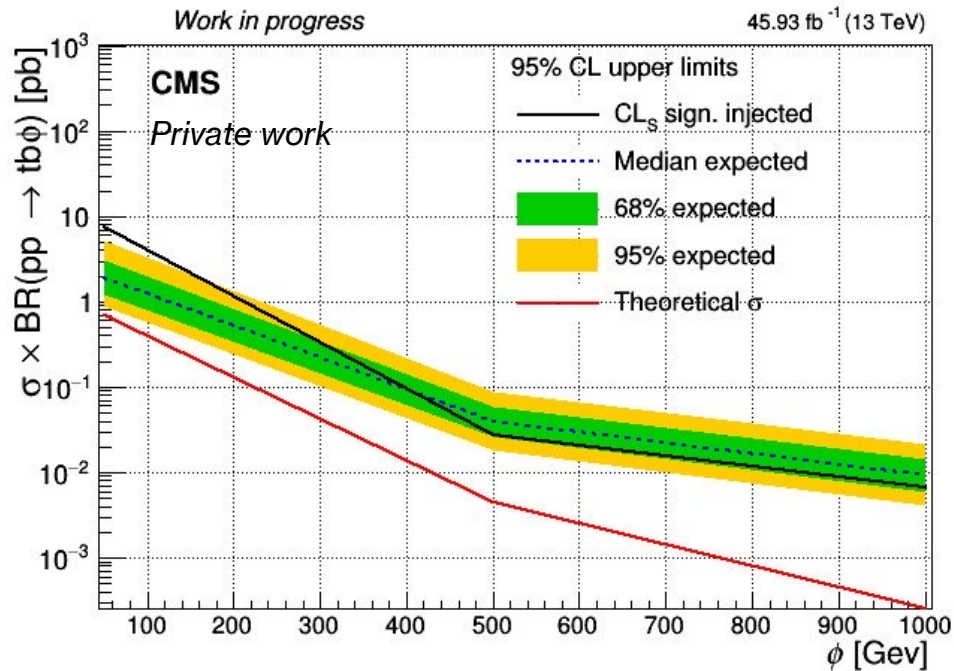
Fit simultaneo sulle MET per le sei regioni ricostruite (Resolved, Semi-Resolved e Merged con e senza Forward Jet)

tDM + TROTA/risultati

work in progress

È stato utilizzato un *sottocampione dei dati del Run II*, non sufficiente per un'esclusione diretta del modello in considerazione.

Facendo proiezioni sull'intero dataset
⇒ *risultati comparabili* all'analisi attuale
Estrapolando con l'aggiunta del canale tW
⇒ sensibilità anche all'ipotesi di **500 GeV**



Conclusioni

- Il data taking del Run-II ha permesso di sondare *numerosi canali ad energie mai viste prima*, cercando di rispondere alle domande aperte del MS.

- Molto spazio delle fasi è ancora da esplorare e LHC continua a cercare nuova fisica in canali sempre più rari o energia maggiore
→ miglioramento delle tecniche sperimentali;

- Per i decadimenti con stati finali adronici di particelle del MS c'è molto spazio per tecniche nuove di ricostruzione e identificazione
→ algoritmi ML;

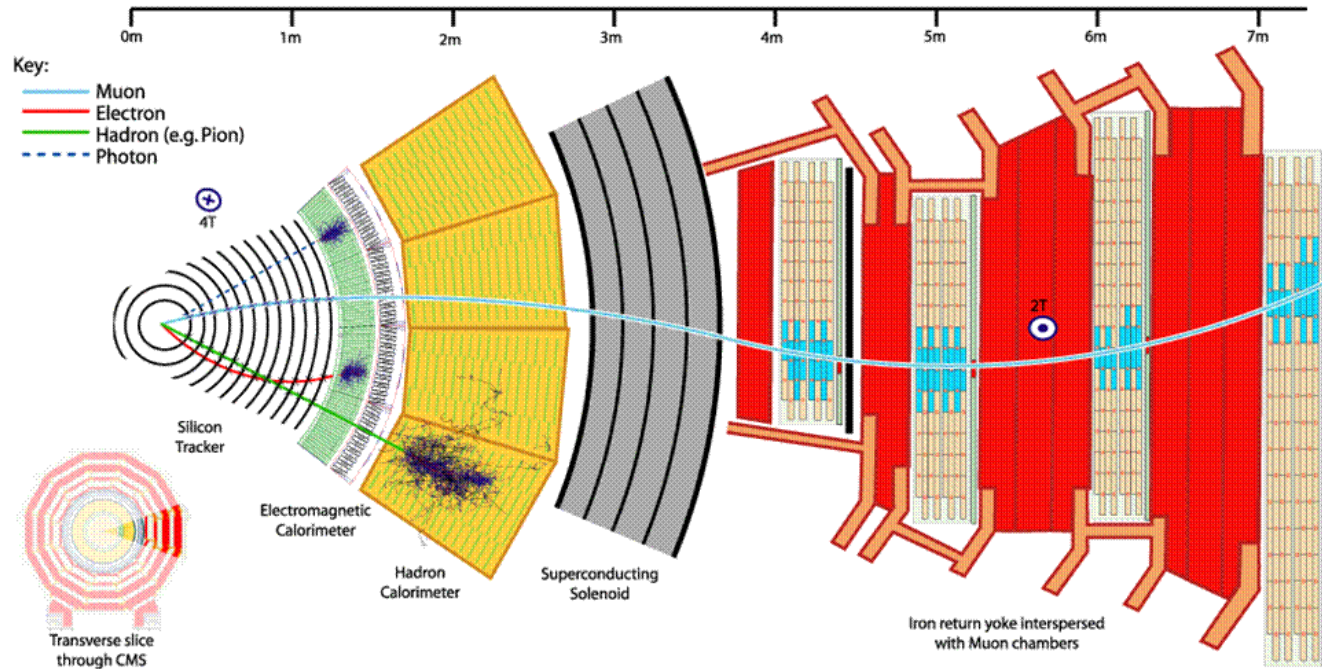
- C'è ancora margine di miglioramento, specie in vista delle analisi del Run-III attualmente in corso e per HL-LHC, le cui condizioni saranno ancora più difficili in termini di background.

*Grazie per
l'attenzione*

Benedetta Argiento – argiento@na.infn.it

BACKUP

CMS/Compact Muon Solenoid



CMS scheme sliced orthogonally to beam axis

- Silicon Tracker
- Electromagnetic calorimeter (ECAL)
- Hadron calorimeter (HCAL)
- Superconducting solenoid
- Muon system

TROTA/Top resolved

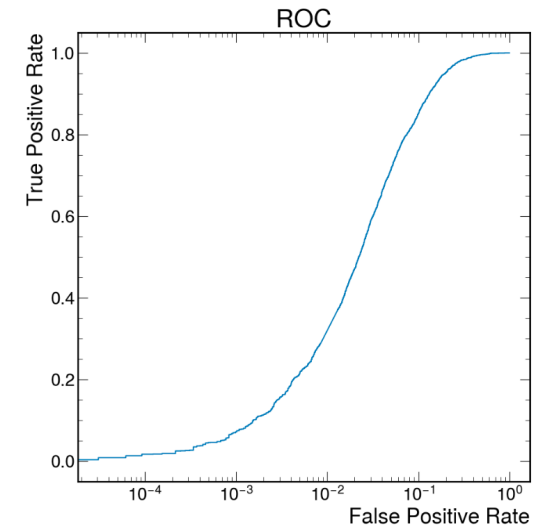
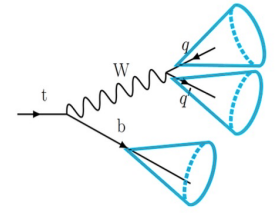
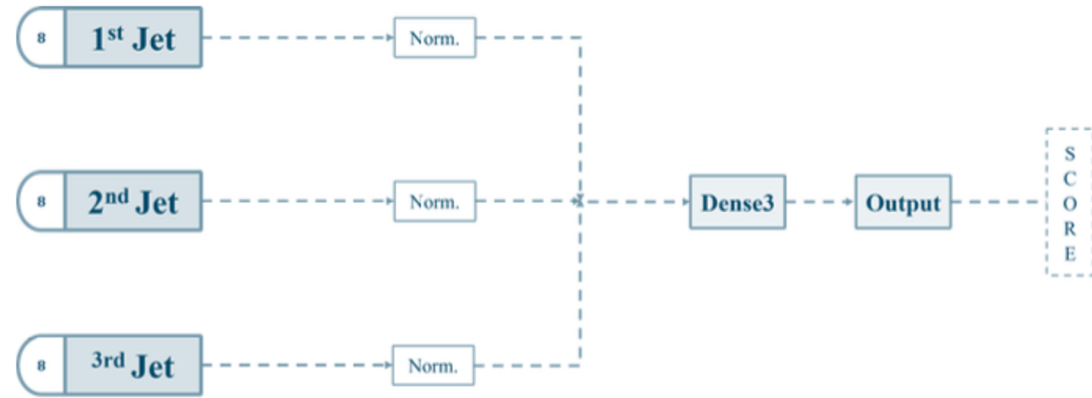
Buone prestazioni per top a basso p_T .

Input:

- area del jet;
- b-tagging score di ParticleNet;
- massa del jet;
- jet p_T ;
- $\Delta\eta(jet, \sum_{i=0}^2 p^{jet,i})$
- $\Delta\phi(jet, \sum_{i=0}^2 p^{jet,i})$

Architettura:

- batch normalization layer;
- dense layer con 25 unità e ReLU come funzione di attivazione;
- dense layer con 30 unità e ReLU come funzione di attivazione;
- dense layer con 1 unità e una sigmoide come funzione di attivazione utilizzato come output layer.

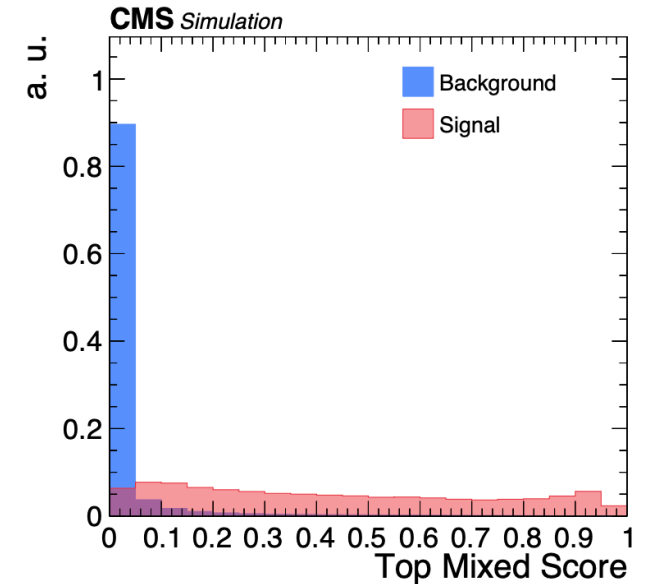
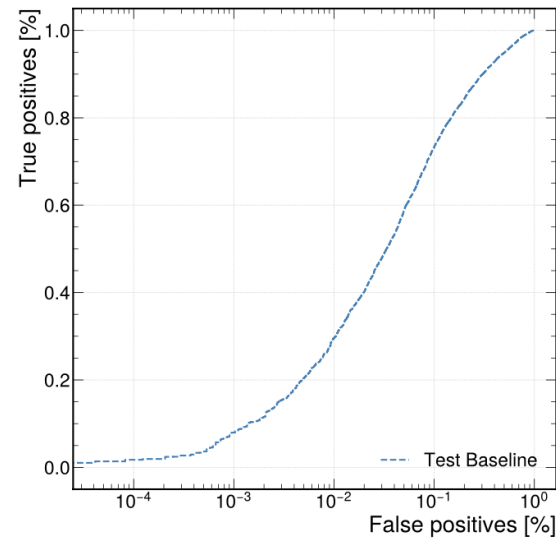


TROTA/Top mixed

Range di p_T tra la regione resolved e quella mixed.

Input:

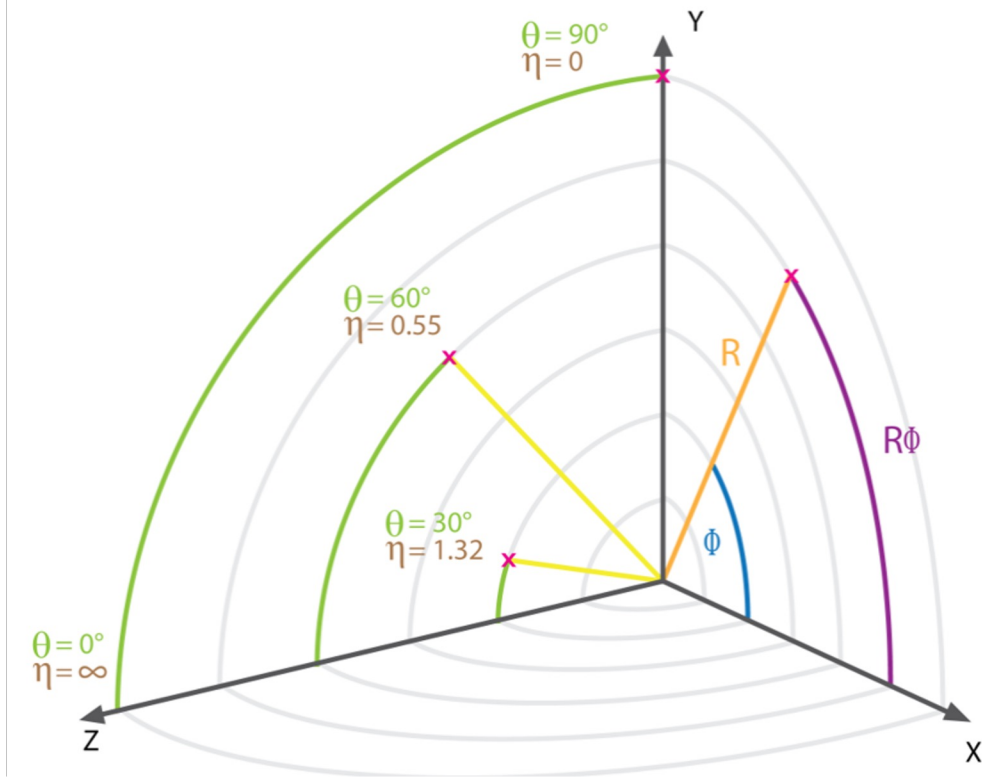
- area del jfatet;
- massa del fatjet;
- fatjet p_T , η , ϕ ;
- fatjet DeepJet b-tagging score;
- fatjet ParticleNet QCD tagging score;
- fatjet ParticleNet t tagging score;
- fatjet ParticleNet W tagging score.



Architettura:

- Jet AK4: Elaborati con strato LSTM (10 unità, attivazione "tanh") per catturare dipendenze sequenziali.
- Jet AK8: Elaborati con strato denso (9 unità, attivazione ReLU) per caratteristiche aggregate.
- Input Combinato: Massa AK4 + combinazione AK4/AK8 elaborati con strato denso (1 unità, attivazione ReLU).
- Elaborazione Finale: Tensore concatenato processato da strato denso (5 unità, attivazione ReLU).

CMS/coordinate system



Coordinate system of CMS inside LHC, where the z axis is along the beam and the x axis is towards the LHC centre

The (r, θ, ϕ) coordinate identifies a Lorentz invariant metric boosted along the Z axis:

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

$$\Rightarrow \textbf{Angular distance: } \Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$$

Other used quantities are:

$$\Rightarrow \textbf{Transverse momentum: } p_T = \sqrt{p_x^2 + p_y^2}$$

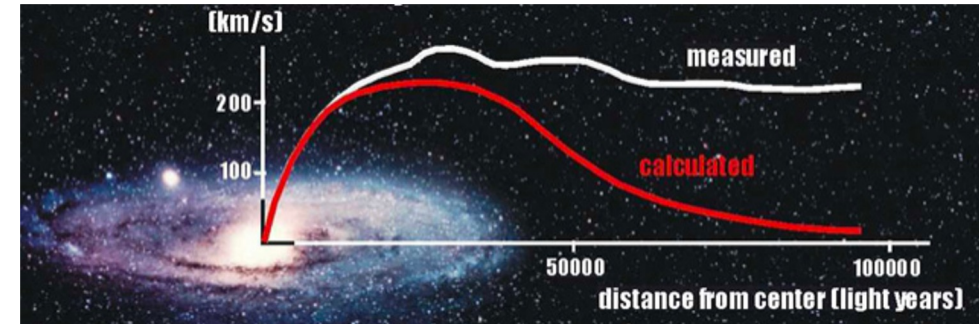
$$\Rightarrow \textbf{Transverse energy: } E_T = E \sin \theta$$

Dark matter/evidence

Curva di rotazione, discrepanza rispetto alla previsione:

- ⇒ previsto una diminuzione
- ⇒ le misure mostrano un andamento asintotico

Lensing gravitazionale: distribuzione di materia in grado di curvare la traiettoria della luce
⇒ Non solo massa visibile



Velocità di rotazione delle galassie in funzione della distanza dal centro



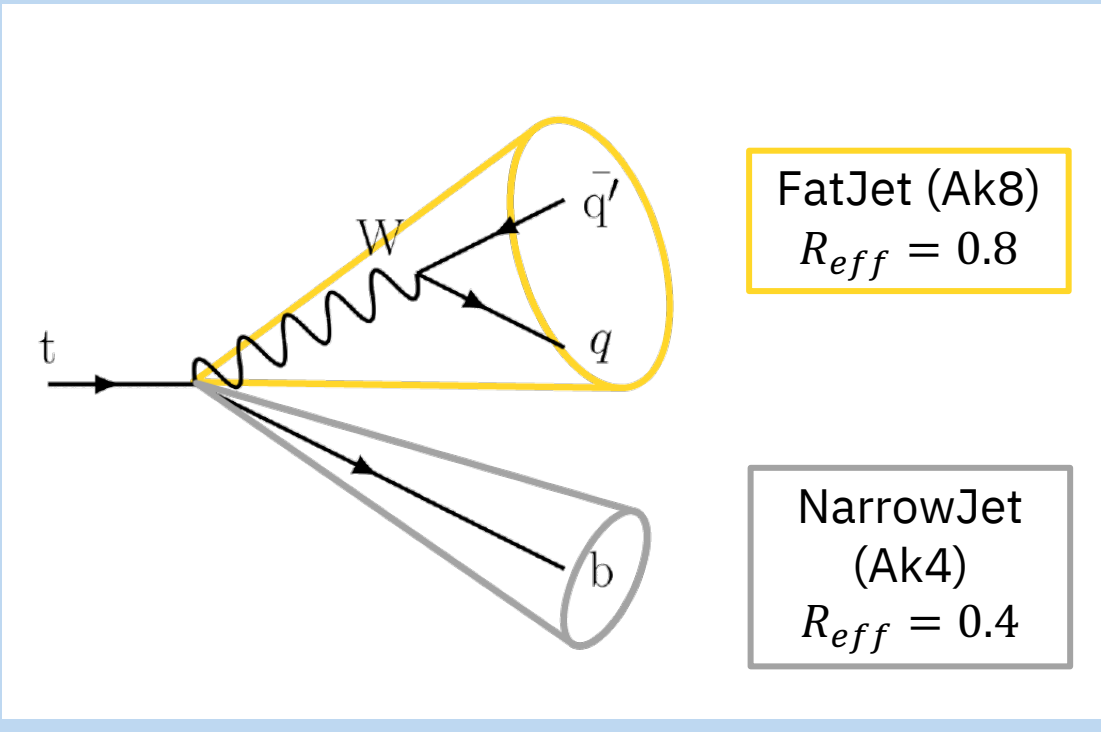
Esempio di lensing gravitazionale (Einstein Ring)

Jet clustering

Eur. Phys. J. C 76, 600

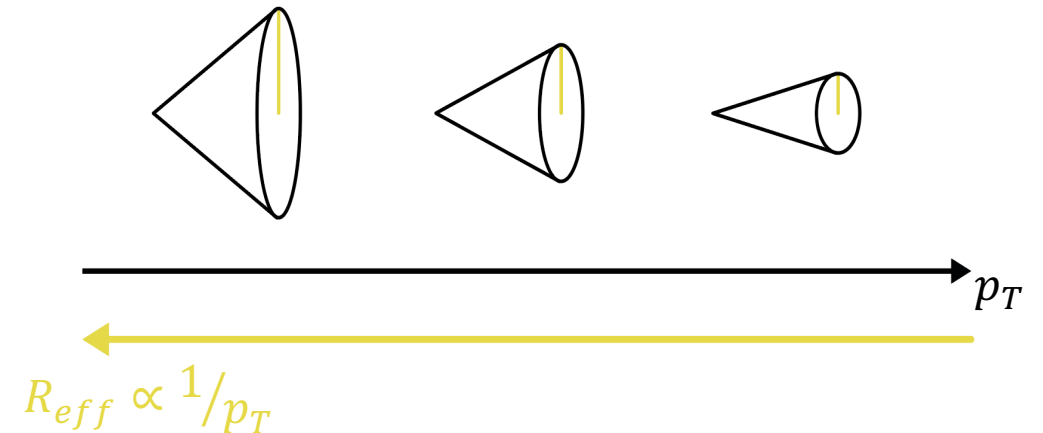
Algoritmi di ricombinazione sequenziale: in input inizia con un elenco di candidati ricostruiti all'interno del rivelatore

A raggio fisso:



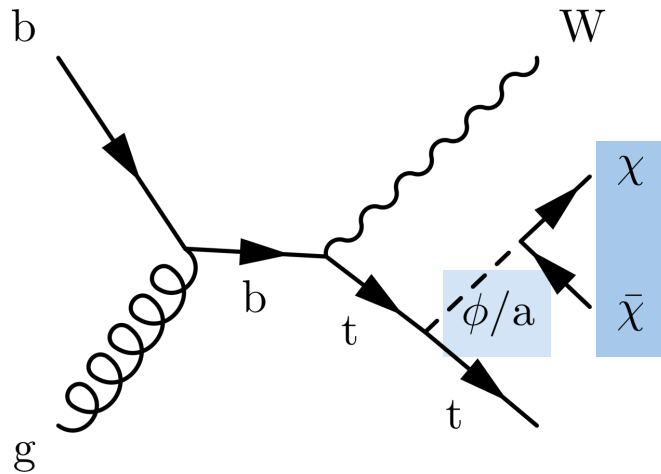
HOTVR (Heavy Object Tagging with Variable Radius)

$$R_{eff} = \begin{cases} \rho/p_T; \\ R_{min} \text{ per } \rho/p_T < R_{min}; \\ R_{max} \text{ per } \rho/p_T > R_{max}. \end{cases}$$



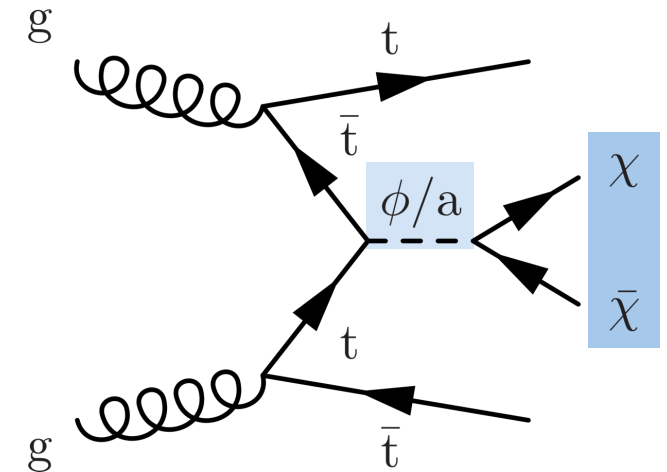
Simplified models $/tDM$ & $t\bar{t}DM$

Stato finale:



eccesso in MET come segnale

mediatore dell'interazione un bosone scalare o pseudoscalare con diverse ipotesi di massa



Canale **Adronico (AH)** / **Semileptonico (SL)** :

- 0 o 1 leptone nello stato finale
- categorizzazione basata sui jet e b-jet
- analisi della MET dopo la selezione sulla base di variabili discriminanti

Canale **Dileptonico (DL)** :

- 2 leptoni nello stato finale
- MET elevata anche per i neutrini
⇒ più complicata la discriminazione
- Multi Variate Analysis

Jet clustering

Eur. Phys. J. C 76, 600

Algoritmi di ricombinazione sequenziale: in input inizia con un elenco di candidati ricostruiti all'interno del rivelatore

$$d_{ij} = \min[p_{T,i}^{2n}, p_{T,j}^{2n}] \Delta R_{ij}^2$$

$$d_{iB} = p_{T,i}^{2n} R_{eff}^2$$

con

d_{ij} : distanza tra due pseudo jet

ΔR_{ij}^2 : distanza angolare nel piano (η, ϕ)

d_{iB} : distanza di un pseudo jet dal fascio

A raggio fisso ($R_{eff} = (0.4, 0.8)$):

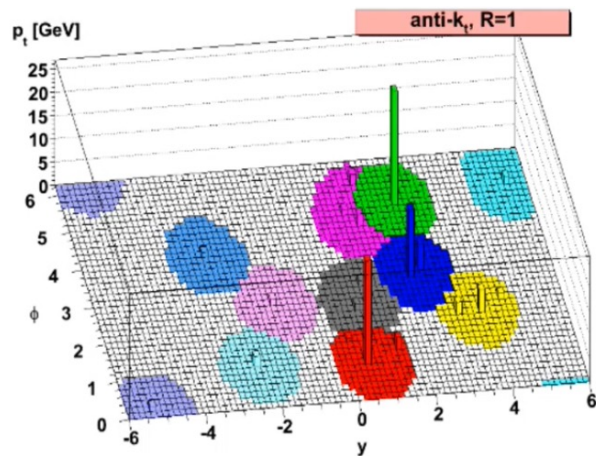
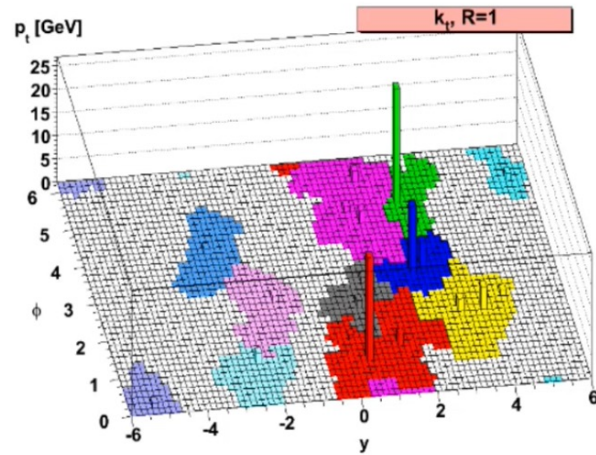
- Anti-kT: $n = -1$;
- Cambridge/Aachen (CA): $n = 0$;
- kT: $n = +1$.

HOTVR (*Heavy Object Tagging with Variable Radius*), $n = 2$:

$$R_{eff} = \begin{cases} \rho/p_T ; \\ R_{min} & \text{per } \rho/p_T < R_{min}; \\ R_{max} & \text{per } \rho/p_T > R_{max}. \end{cases}$$

dove ρ rappresenta la pendenza di R_{eff} .

Ricostruzione dei jet/anti-KT



⇒ **Collinear safety**: l'algoritmo non è influenzato dalla scissione o fusione artificiale di partoni collinari

⇒ **Infrared safety**: l'algoritmo rimane insensibile all'emissione di gluoni morbidi, che hanno un impatto minimo sul momento del jet e sulla collimazione

Usando la metrica: $d_{ij} = \min\left(\frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2}\right) \frac{(\Delta R_{ij})^2}{R^2}$; $d_{iB} = \frac{1}{p_{Ti}^2}$

- Se $d_{ij} < d_{iB}$: le particelle si fondono in un'unica pseudo-particella.