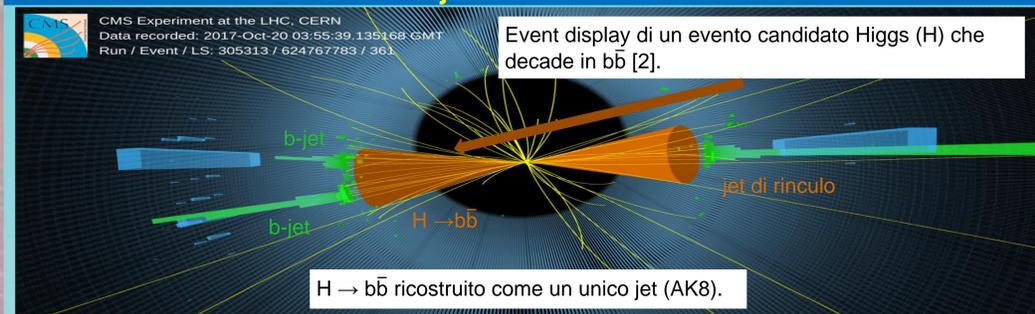


## Abstract

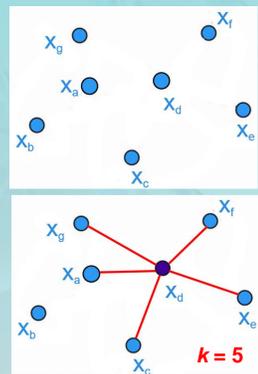
Un aspetto fondamentale delle ricerche svolte all'esperimento CMS riguarda l'identificazione dei jet prodotti nelle collisioni protone-protone ad alta energia. I bosoni W/Z, il bosone di Higgs ed il quark top possono essere prodotti con elevato boost di Lorentz ed, in tali circostanze, i loro prodotti di decadimento possono essere ricostruiti come fat jet, ossia anti-k<sub>t</sub> jet [1] di raggio 0.8 (AK8). L'identificazione della particella, che dà inizio al fat-jet, ricopre pertanto un ruolo cruciale nel distinguere i top quark ed i bosoni boostati dal fondo dominante di QCD. Diversi algoritmi di identificazioni di AK8, basati su sofisticate tecniche di machine learning, sono stati sviluppati dalla collaborazione CMS. In questo talk verrà fornita una loro panoramica in termini di prestazione ed utilizzo all'interno della collaborazione.

## Fisica dei jet boostati a CMS

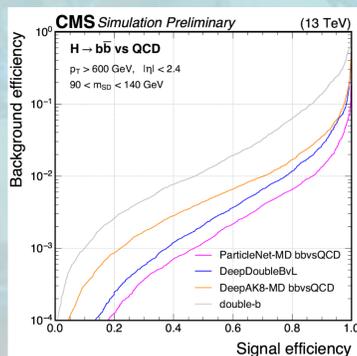
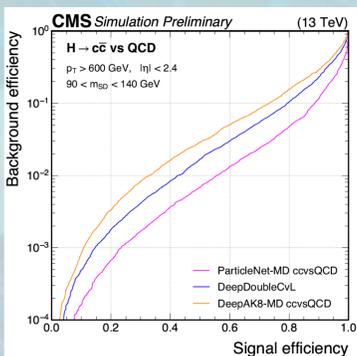


Algoritmi di identificazione per il Run2:

- Addestramento fatto su simulazioni MonteCarlo (MC) di risonanze X (H, Z, W) che decadono in b $\bar{b}$ , c $\bar{c}$  o q $\bar{q}$  (X  $\rightarrow$  b $\bar{b}$ /c $\bar{c}$ /q $\bar{q}$ ) contro il fondo di QCD.
- Addestramento effettuato in maniera tale da essere indipendente dalla massa della particella risonante.
- **DeepDoubleX** [3]: *Deep Neural Network* (NN).
- Input: component del jet + vertici secondari + informazioni globali del jet
- **ParticleNet-MD** [4]: *Dynamic Graph Convolutional NN*
- Stessi input di DeepDoubleX **più le coordinate spaziali**.
- Rielabora le proprietà di ogni costituente X<sub>i</sub> analizzando i k corrispondenti più vicini (nel piano  $\eta$ - $\phi$ ).
- Output: score per discriminare X  $\rightarrow$  b $\bar{b}$ /c $\bar{c}$  dalla QCD



$$PN-MD_{BB(CC)vsQCD} = \frac{p(X \rightarrow b\bar{b} (c\bar{c}))}{p(X \rightarrow b\bar{b} (c\bar{c})) + p(QCD)}$$

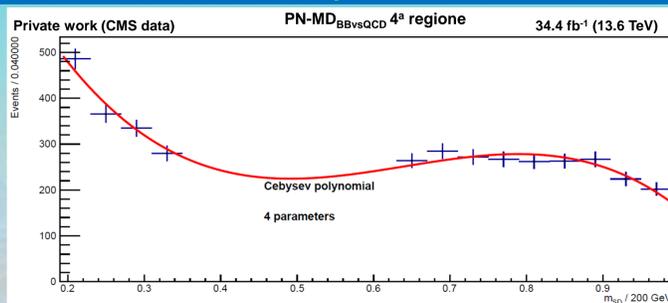


ParticleNet-MD è risultato migliore di DeepDoubleX per il Run 2 [5].

- DeepDoubleX non più supportato per il Run 3.
- ParticleNet-MD con nuovi sample MC basati sulle condizioni di Run3.
- ParticleNet-MD necessita di una nuova validazione per i dati di Run3.
- Validazione eseguita indipendentemente per ogni anno, a partire dal 2022.

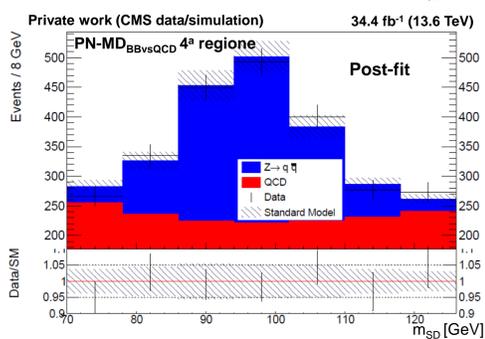
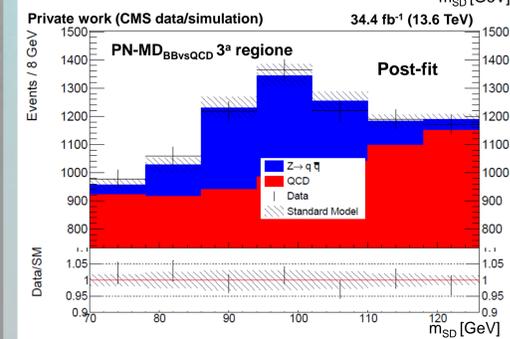
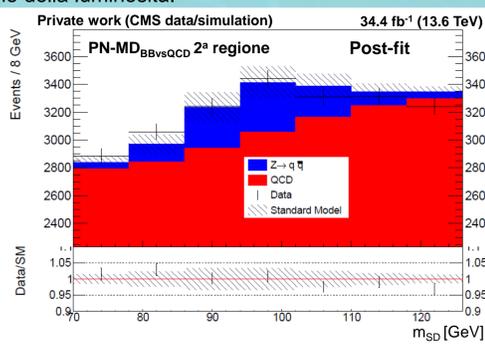
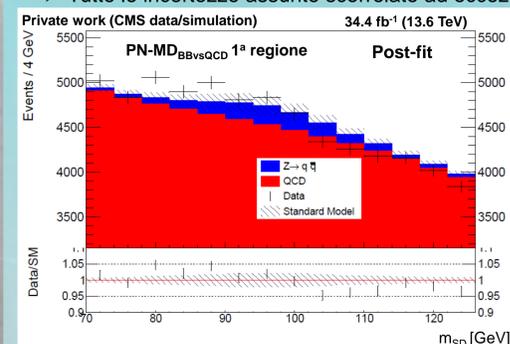
## Stima data-driven della QCD per il 2022

- QCD stimata da un fit nelle regioni laterali della m<sub>SD</sub> sui dati
- Fit: polinomi di Chebyshev. Ordine stabilito mediante il test di Fisher (CL al 5%).
- Due incertezze sistematiche che tengono conto delle incertezze sui parametri del fit e della riproducibilità del fit stesso.



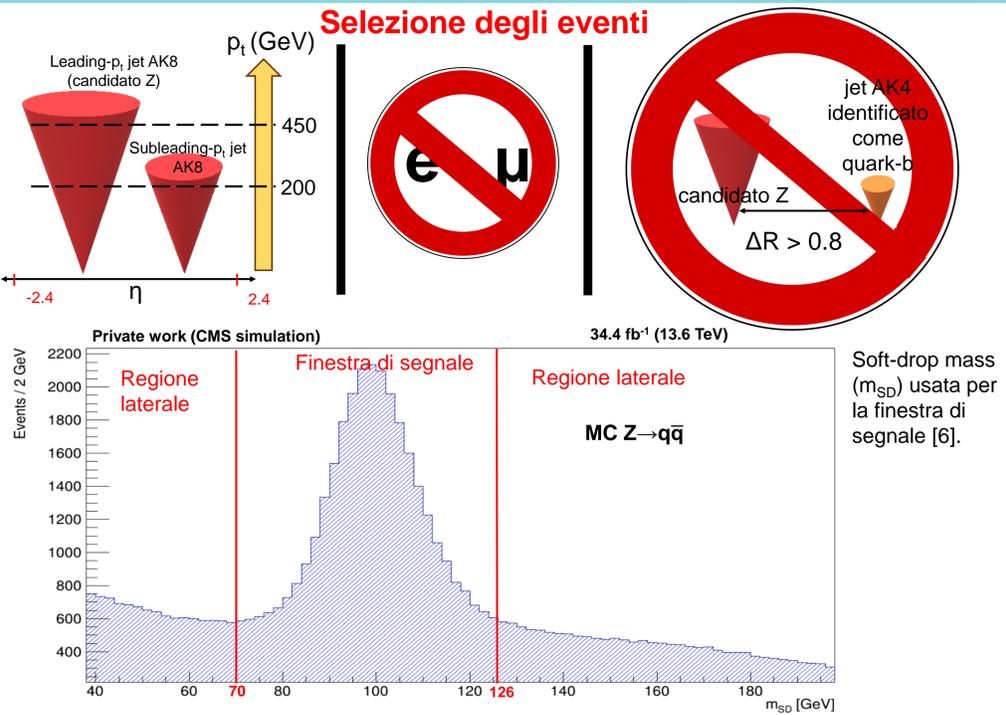
## Likelihood fit

- Effettuato un Likelihood fit simultaneo delle distribuzioni di m<sub>SD</sub> relative alle 4 regioni ad alto score di PN-MD<sub>BBvsQCD</sub> per migliorare l'accordo dati 2022 - atteso SM:
- Associata una **signal strength** (r) indipendente per il campione MC di Z  $\rightarrow$  q $\bar{q}$  in ogni regione dello score.
- Incertezze considerate nel fit:
  - statistiche e correzioni di scala dei jet per Z  $\rightarrow$  q $\bar{q}$ ;
  - fit e riproducibilità del fit per QCD;
  - luminosità.
- Tutte le incertezze assunte scorrelate ad eccezione della luminosità.

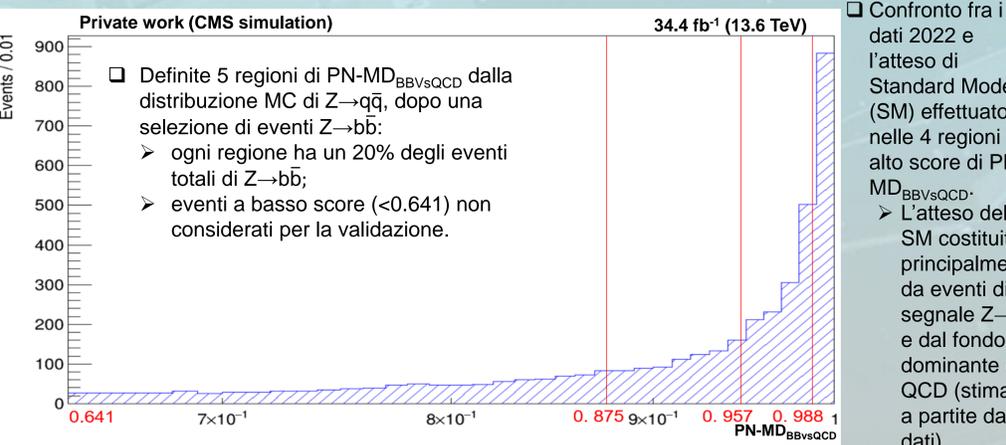


## Validazione dello score X $\rightarrow$ b $\bar{b}$ di ParticleNet-MD sui dati del 2022

Validazione con eventi Z  $\rightarrow$  b $\bar{b}$ -like ad alto boost di Lorentz.

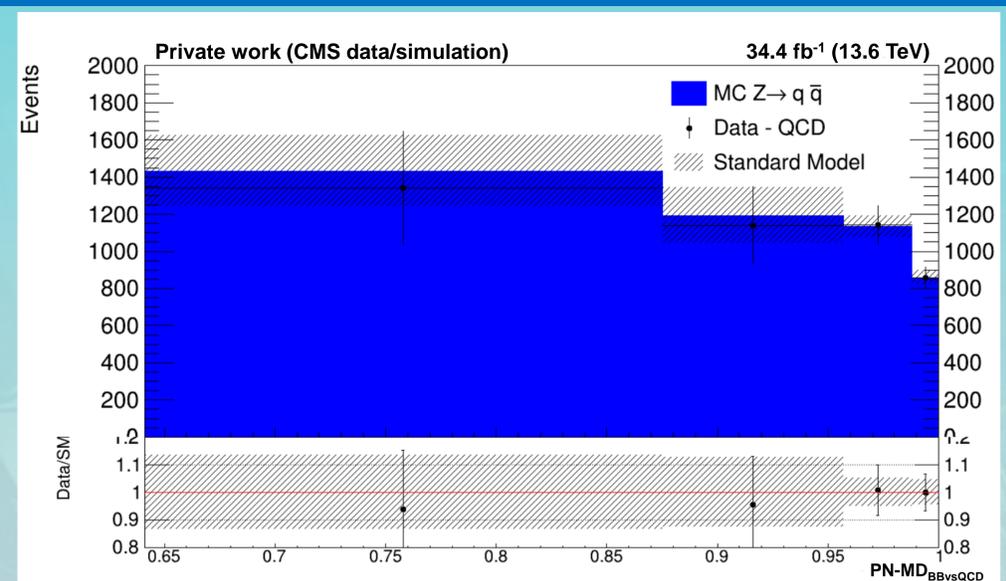


Soft-drop mass (m<sub>SD</sub>) usata per la finestra di segnale [6].



Confronto fra i dati 2022 e l'atteso di Standard Model (SM) effettuato nelle 4 regioni ad alto score di PN-MD<sub>BBvsQCD</sub>. L'atteso dello SM costituito principalmente da eventi di segnale Z  $\rightarrow$  q $\bar{q}$  e dal fondo dominante di QCD (stimato a partite dai dati)

## Risultati e conclusioni



- Buon accordo fra dati e l'atteso di SM.
- Signal strengths compatibili con 1 all'interno dell'incertezza.
- Stima data-driven mostra una buona modellizzazione del fondo di QCD.
- Osservato un pulling dei nuisance parameters associati alla QCD al più del 4%

|  | r    | errore |
|--|------|--------|
| 0.641 < PN-MD <sub>BBvsQCD</sub> ≤ 0.875 | 1.3  | 0.5    |
| 0.875 < PN-MD <sub>BBvsQCD</sub> ≤ 0.957 | 1.3  | 0.4    |
| 0.957 < PN-MD <sub>BBvsQCD</sub> ≤ 0.988 | 1.31 | 0.15   |
| 0.988 < PN-MD <sub>BBvsQCD</sub> ≤ 1     | 1.04 | 0.09   |

Lo score PN-MD<sub>BBvsQCD</sub> mostra una buona discriminazione fra QCD e Z  $\rightarrow$  q $\bar{q}$ , mettendo in evidenza il picco della Z ad alto PN-MD<sub>BBvsQCD</sub>

## Referenze e contatti

donato.troiano@cern.ch  
 donato.troiano@ba.infn.it

[1] M Cacciari *et al.*, The anti-k<sub>t</sub> jet clustering algorithm, JHEP 04 (2008) 063  
 [2] <https://cds.cern.ch/record/2714889>  
 [3] E. Bols *et al.*, Jet Flavour Classification Using DeepJet, JINST 15 (2020) 12, P12012  
 [4] Huilin Qu *et al.*, Jet tagging via particle clouds, PHYSICAL REVIEW D 101, 056019 (2020)  
 [5] CMS-PAS-BTV-22-001  
 [6] A. J. Larkoski *et al.*, Soft Drop, JHEP 05 (2014) 146