

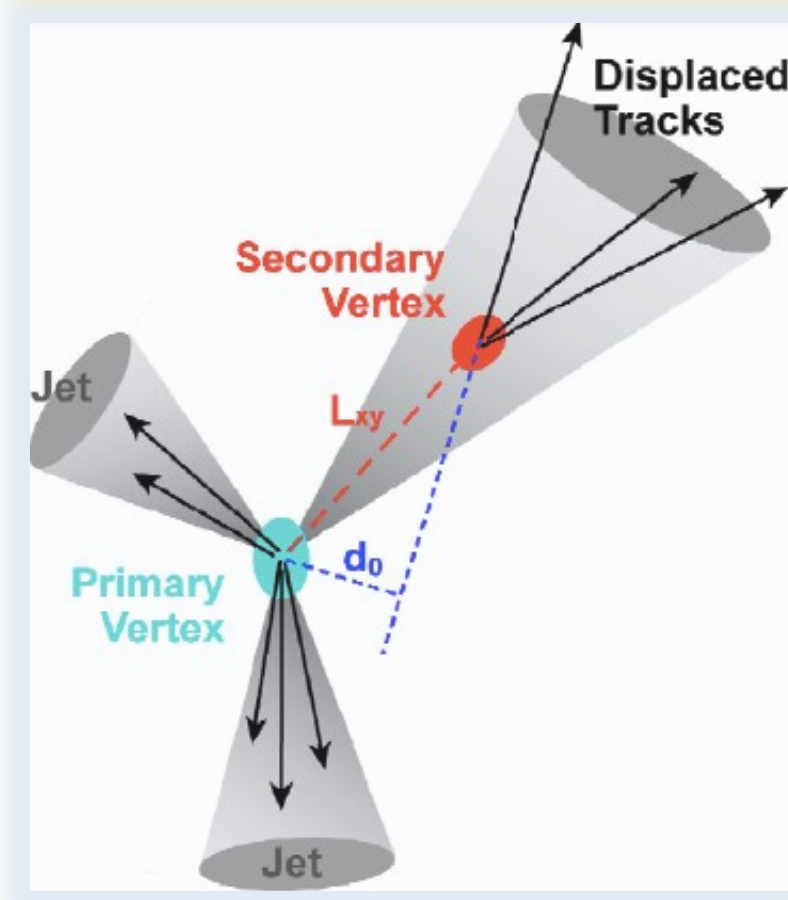
Calibrazione dei jet leggeri in ATLAS mediante algoritmi di b -tagging basati su reti neurali ricorsive e grafiche

Leonardo Tofflin^{1,2}

Incontri di Fisica delle Alte Energie 2024, 3-5 Aprile 2024 (Firenze, Italia)

Identificazione dei b -jet

- Algoritmi di b -tagging basati sulle peculiarità degli **adroni B**
 - vita media relativamente lunga ~ 1.5 ps
 - alta massa ~ 5 GeV
 - molteplicità delle tracce dei prodotti di decadimento
 - costituzione dei b -jet



- Calibrazione richiesta per correggere differenze tra dati e Monte Carlo (MC) su cui gli algoritmi sono istruiti:
 - modelli di frammentazione del quark b
 - simulazione dello sciame partonico
 - risposta del rivelatore

- Efficienza di (mis-)identificazione per jet di **sapore f** : $\epsilon_f = \frac{N_{pass}^f}{N_{all}^f}$

$$SF^f = \frac{\epsilon_{data}^f}{\epsilon_{MC}^f}$$

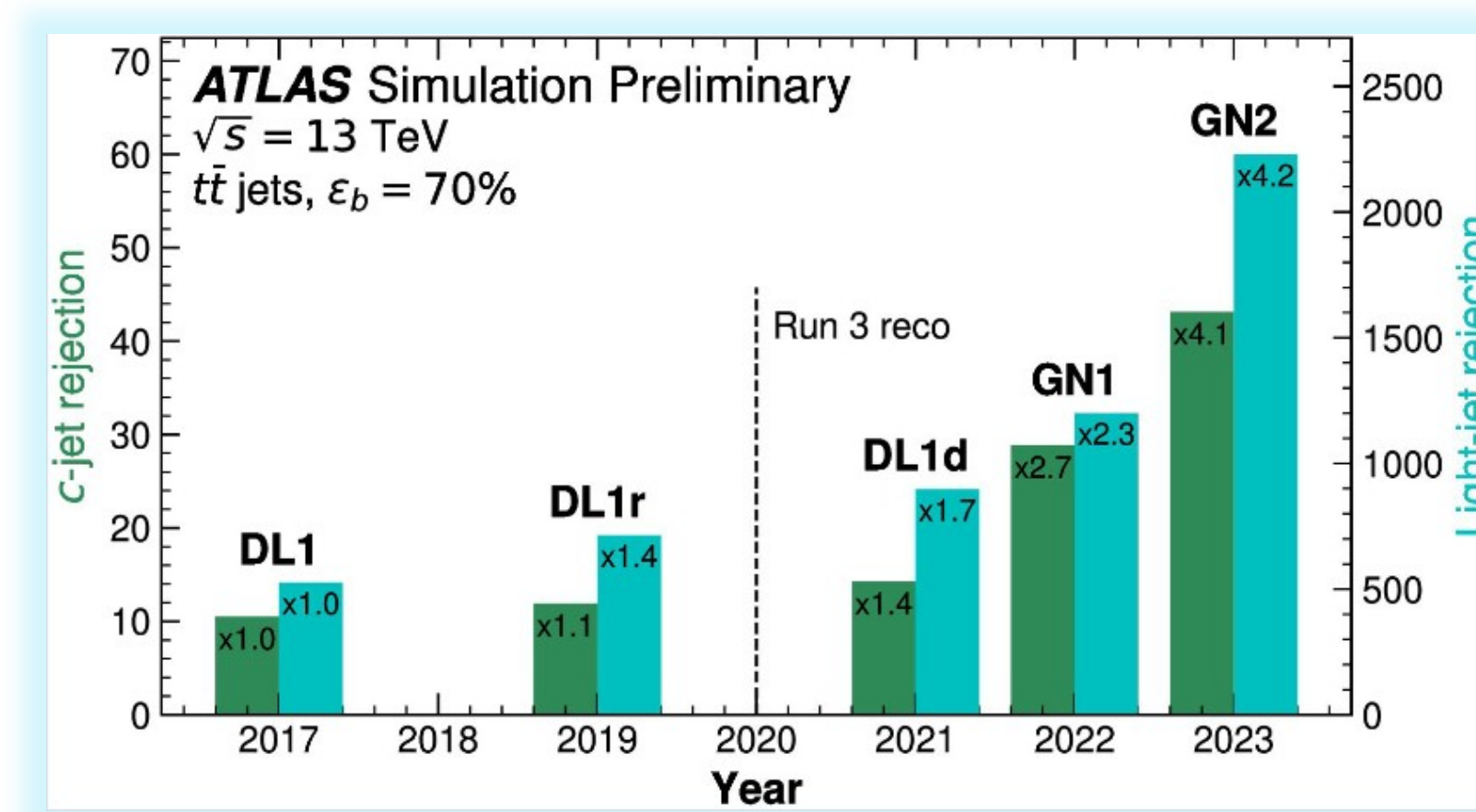
- Fattore di calibrazione (o fattore di scala, "scale factor"):

f-jet dopo la selezione di b -tagging

f-jet in totale nel campione di dati o MC

Algoritmi di b -tagging in ATLAS

- Algoritmi di b -tagging a base di **reti neurali profonde** (DL1, DL1r, DL1d) e **grafiche** (GN1, GN2)
 - nel prossimo futuro, GN2 come nuovo algoritmo standard raccomandato
 - costante miglioramento nella reiezione dei jet leggeri
 - rateo di mis-identificazione dell'ordine di 0.04% per $\epsilon_b=70\%$ con GN2



DL1d

- output dei trigger di basso livello e cinematiche dei jet presi come input per costruire un discriminante singolo [3]:

$$D_b = \log \left[\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_u} \right]$$

- rete neurale** (NN) **ricorsiva** "deep-forward" istruita usando **Keras** [4]
- backend **Theano** [5] come interfaccia utente
- ottimizzatore **Adam** [6]
- output: probabilità per un jet di essere un **b , c o leggero**

GN2

- fortemente basato sull'architettura dei trasformatori [7]
- rete neurale grafica** (GNN)
- nessun input da algoritmi di basso livello
- aggiunta del contributo dei leptoni tau nel discriminante singolo:

$$D_b = \log \left(\frac{p_b}{f_c p_c + f_\tau p_\tau + (1 - f_c - f_\tau) p_u} \right)$$

- apprendimento simultaneo** di sapore dei jet, identificazione dei vertici e dell'origine delle tracce su un campione di dati maggiore rispetto a DL1d
 - aggiunta di un layer di "dropout" per stabilizzare l'apprendimento
- migliore prestazione e adattabilità
- possibilità di aggiungere informazioni su leptoni e particelle neutre nei jet per il futuro tagger "**GN3**"
- output: probabilità per un jet di essere un **b , c o leggero**

Tecniche per la calibrazione dei jet leggeri

- Calibrazione dei jet leggeri sviluppata dal gruppo "**Flavour Tagging**" (FTAG) nella **Collaborazione ATLAS** al CERN [1]
 - estrazione dell'efficienza di errata identificazione dei jet leggeri dai dati di **Run-3**
 - estrazione degli **scale factor** per i **jet leggeri** in **eventi Z+jets**
 - studi su possibili miglioramenti delle performance del tagger mediante inclusione dei muoni soffici

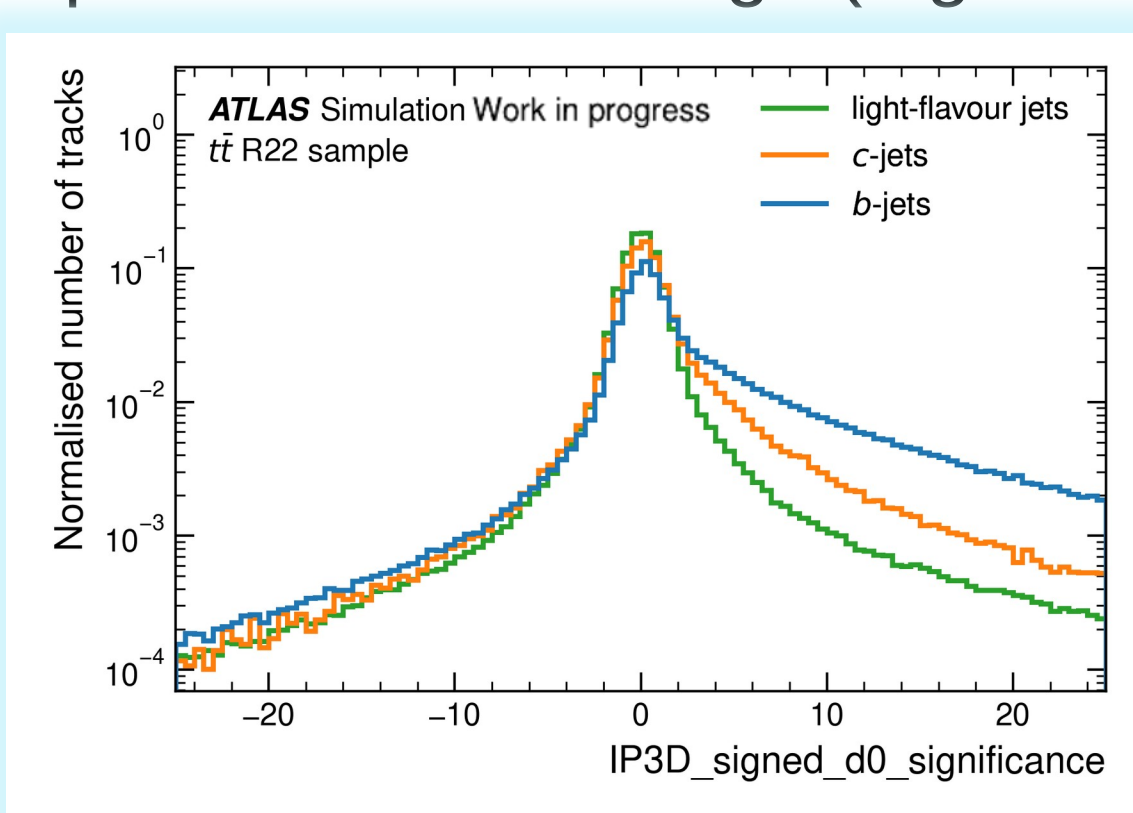
- Problematica sperimentale**: numero di jet leggeri superanti la selezione del b -tagger troppo **basso** per poter stimare ϵ_{light} dai dati

- Strategia: **tagger "negativo"** per **ridurre contaminazione dai b -jet** e arricchire il campione in jet leggeri

- algoritmo "flip"**: inversione del segno del parametro d'impatto (IP) e della lunghezza di decadimento delle tracce assegnate ai jet

- assunzione: mis-identificazione dei jet leggeri causata principalmente da **effetti di risoluzione sul parametro d'impatto** (errata modellizzazione delle risoluzioni di d_0 and z_0)

- distribuzione degli IP simmetrica intorno a 0 per jet leggeri, asimmetrica per b -jet per la presenza di particelle a vita lunga (e.g. S o Λ)

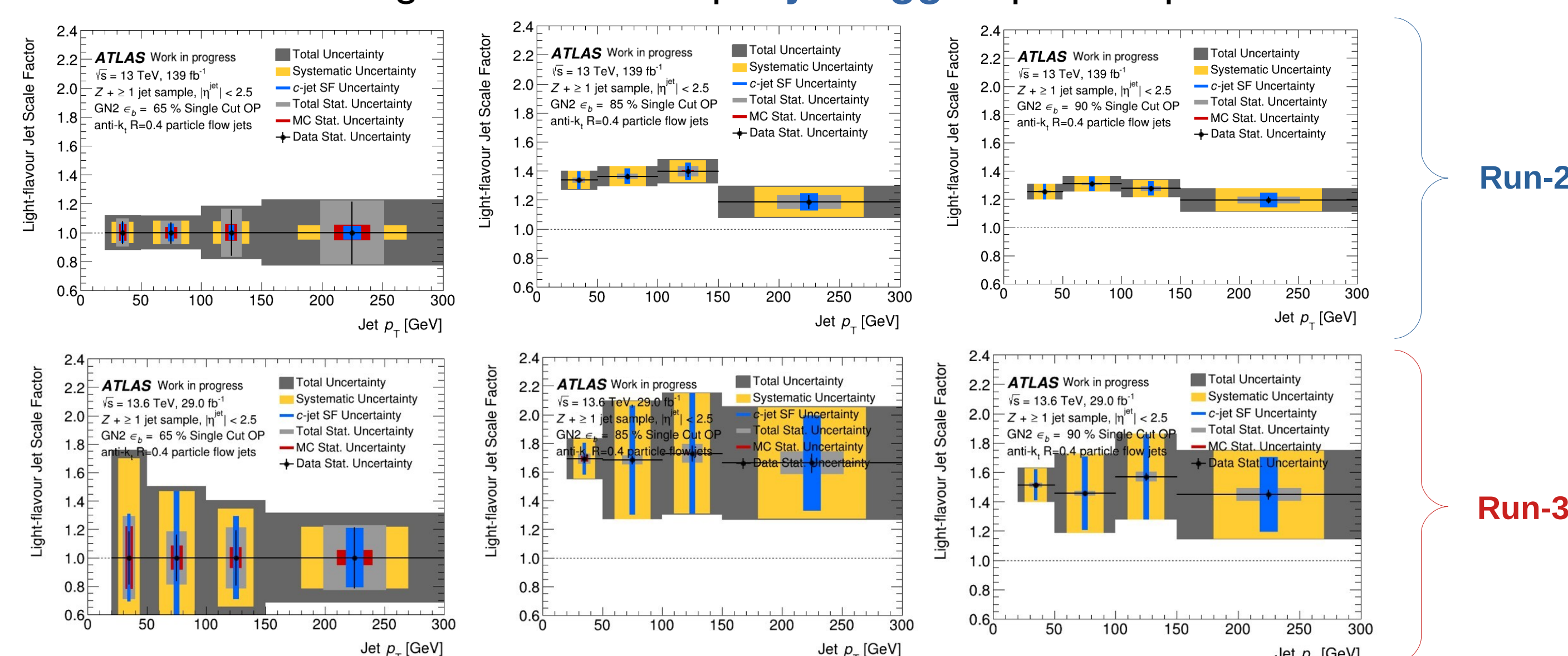


- con inversione del segno di IP: stesso rateo di mis-identificazione dei jet leggeri ed **efficienza per b -jet fortemente ridotta**

- fit** della distribuzione della massa del vertice secondario (m_{SV}) e del **discriminante** del tagger in intervalli pseudo-continui di p_T dei jet

Risultati preliminari per GN2

- Test preliminari** sulla calibrazione dei dati di **Run-2** e **Run-3** con il tagger GN2
 - insieme ridotto di incertezze sistematiche
 - risultati definitivi sugli **scale factor** per **jet leggeri** previsti prossimamente

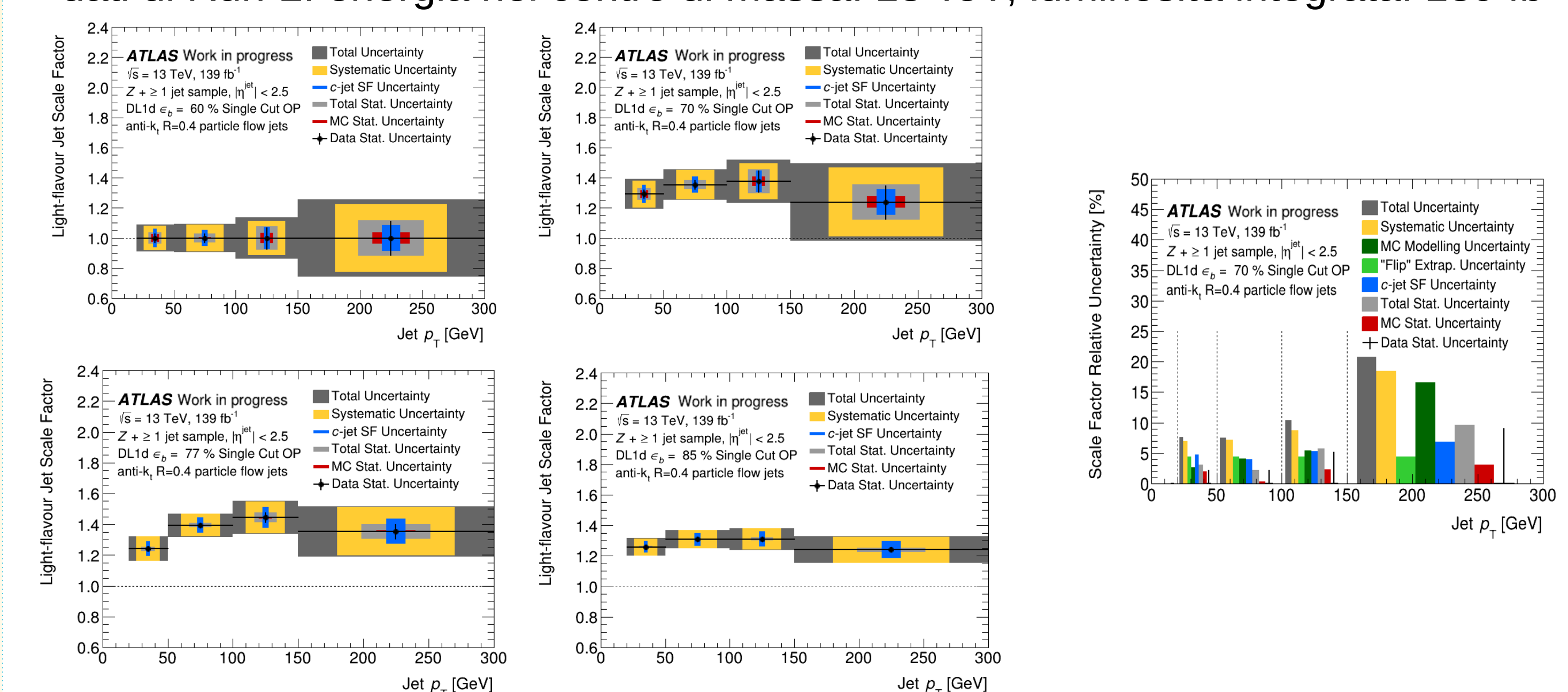


Referenze:

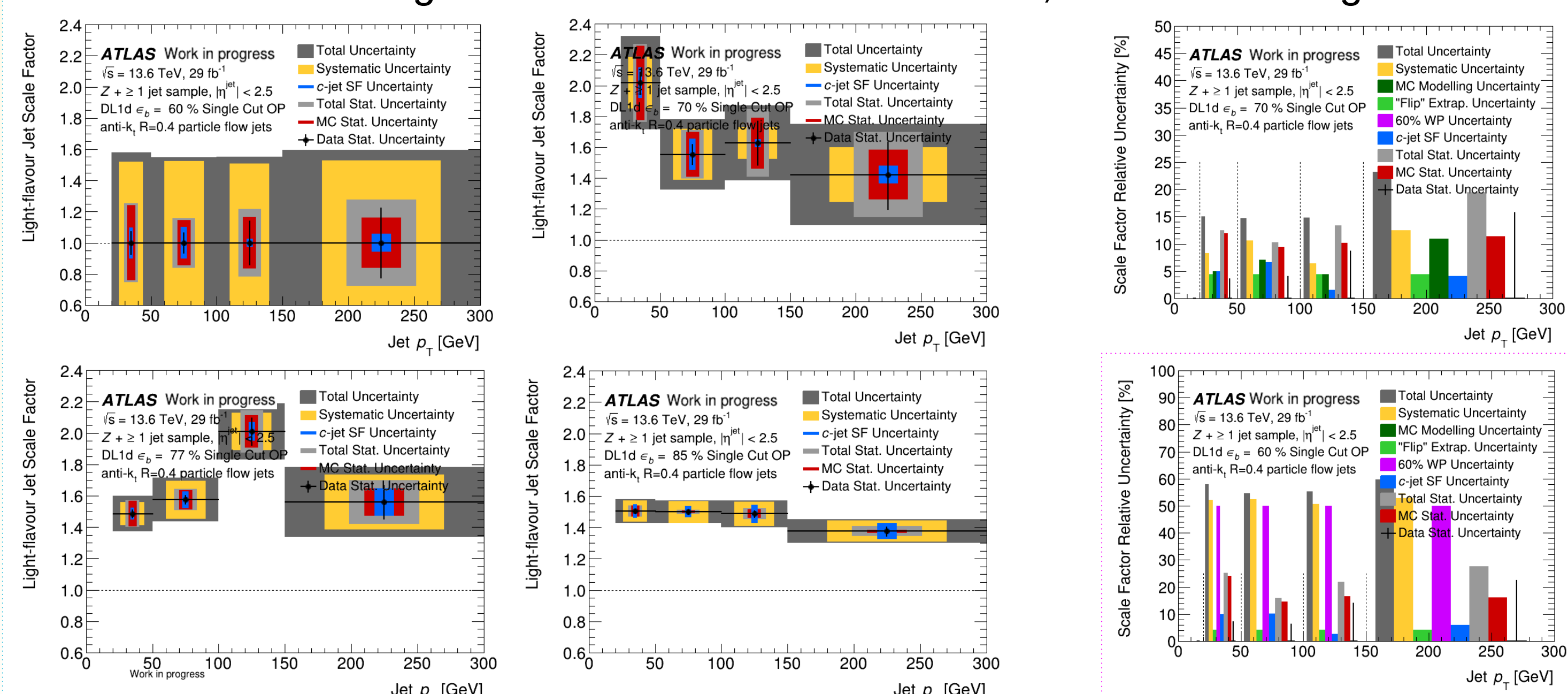
- [1] ATLAS Collaboration: *JINST* **3** (2008) S08003.
- [2] Collezione di grafici del gruppo ATLAS Flavour Tagging: [FTAG-2023-01](#).
- [3] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. ATL-PHYS-PUB-2017-013 (2017). <https://cds.cern.ch/record/2273281>.
- [4] F. Chollet et al., *Keras*. <https://keras.io> (2015).
- [5] Theano Development Team, *Theano: a Python framework for fast computation of mathematical expressions* (2016), arXiv: 1605.02688 [cs.SC].
- [6] D. P. Kingma, J. Ba, *Adam: A method for stochastic optimization* (2014). arXiv: 1412.6980 [cs.SC].
- [7] A. Vaswani et al., *Attention Is All You Need*, arXiv:1706.03762.

Risultati per DL1d

- Scale factor per jet leggeri** definiti per il tagger DL1d
- dati di Run-2: energia nel centro di massa: 13 TeV; luminosità integrata: 139 fb⁻¹



- dati di Run-3: energia nel centro di massa: 13.6 TeV; luminosità integrata: 29 fb⁻¹



- non convergenza del fit per $\epsilon_b=60\%$ a causa di **bassa statistica**
- scale factor per $\epsilon_b=60\%$ fissati a 1.0