Simplex Demixing: Disentangling Multiple Light-Flavor Jets at Colliders

ML4Jets 2026

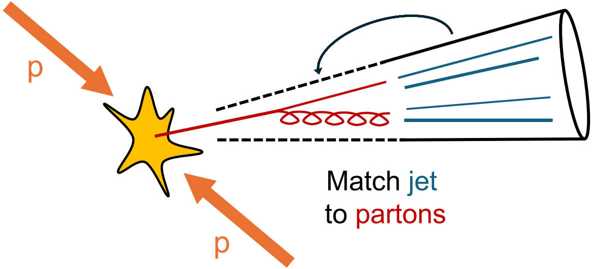
Gregorio de la Fuente

09/14/2026

Joint work with Jesse Thaler

Link to our paper: 2607.24921

Naively, jet flavor is assigned using parton labels from a Monte Carlo generator



This is unphysical. Can we come up with a hadron-level definition of jet flavor instead?

[Gras, Höche, Kar, Larkoski, Lönnblad, Plätzer, Siódmok, Skands, Soyez, Thaler, 1704.03878]

Toward a hadron-level definition of jet flavor

2015: Jet flavor as an enriched region of phase space.

– Well-defined, but impractical.

[Gras et al., 1704.03878]

2018: An operational definition of quark and gluon jets.

– Practical, but limited to two flavors.

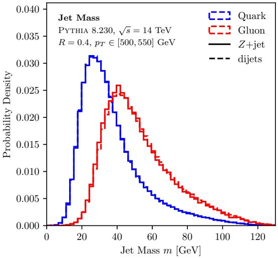
[Komiske, Metodiev, Thaler, 1809.01140]

Today: **A practical definition of multiple jet flavors.**

[de la Fuente, Thaler, 2607.24921]

Quark jet: a phase space region (as defined by an unambiguous hadronic fiducial cross section measurement) that yields an enriched sample of quarks (as interpreted by some suitable, though fundamentally ambiguous, criterion).

Adapted from [Gras et al., 1704.03878]

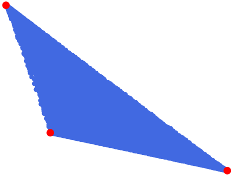


New perspective: jet flavors are the vertices of a simplex

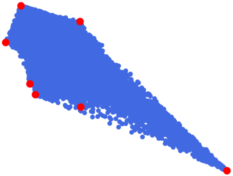
The simplex is the bounding shape of the output of a ML classifier trained on jet samples.

Each jet gets mapped to a point within the simplex, and the flavors are the vertices of the simplex. We extract...

u, d, and g flavors from 3 artificial samples.



7 flavors from 14 quasi-realistic dijet samples.



Our method works for an arbitrary number of samples M and flavors $T \leq M$, with enough information. [de la Fuente, Thaler, 2607.24921]

2018: The operational definition of **two** flavors

Classification without labels (CWoLa). Train a classifier on $p_1(x), p_2(x)$ using the cross-entropy loss:

$$L[c_1, c_2] = -\int d^D x \, [p_1(x) \ln c_1(x) + p_2(x) \ln c_2(x)] \,,$$

subject to $c_1(x), c_2(x) \geq 0$, $c_1(x) + c_2(x) = 1$ (probability 1-simplex).

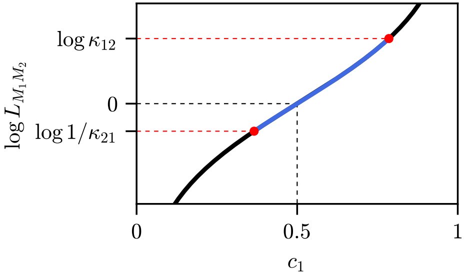
At $\delta L = 0$, the output $c_1(x)$ is monotonic in the likelihood ratio $L_{12}(x) \equiv p_1(x)/p_2(x)$.

The jets lie between

$$\kappa_{12} \equiv \min_x L_{12},$$

$$\kappa_{21} \equiv \min_x \frac{1}{L_{12}}.$$

[Metodiev, Nachman, Thaler, 1708.02949]



2018: The operational definition of **two** flavors

Jet topics. Assume a mixture model of separable distributions:

$$p_1(x) = f_1 p_q(x) + (1 - f_1) p_g(x),$$
$$p_2(x) = f_2 p_q(x) + (1 - f_2) p_g(x).$$

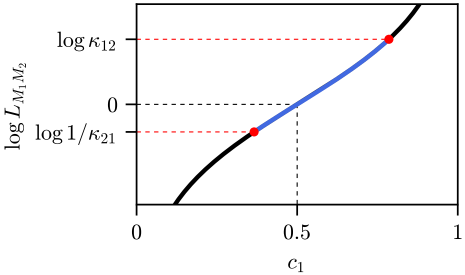
[Metodiev, Thaler, 1802.00008] [Dillon, Faroughy, Kamenik, 1904.04200] [Alvarez, Lamagna, Szewc, 1911.09699]

If $f_1 \neq f_2$, the endpoints give:

$$f_1 = \frac{1 - \kappa_{12}}{1 - \kappa_{12}\kappa_{21}},$$
$$f_2 = \frac{\kappa_{21}(1 - \kappa_{12})}{1 - \kappa_{12}\kappa_{21}}.$$

[Komiske, Metodiev, Thaler, 1809.01140]

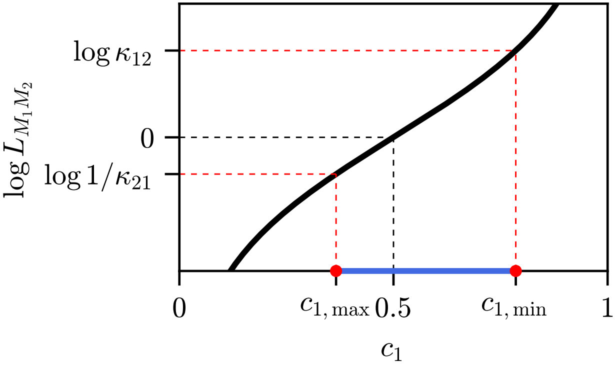
[Katz-Samuels, Blanchard, Scott, 1710.01167]



The classifier distribution is a **geometric object**

Every jet gets mapped between $c_{1,min}$ and $c_{1,max}$: $c_{1,min}$ is quark-like, while $c_{1,max}$ is gluon-like (or vice versa).

A line segment = 1-simplex



Can get f_1 and f_2 from $c_1(x)$ directly!

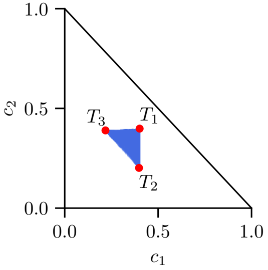
2026: For M samples, the distribution is a $(T-1)$-simplex

As an example, if we train a classifier on $M = 3$ jet samples $p_1(x)$, $p_2(x)$, $p_3(x)$:

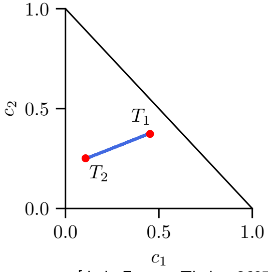
$$L[c_1, c_2, c_3] = -\int d^D x \, [p_1(x) \ln c_1(x) + p_2(x) \ln c_2(x) + p_3(x) \ln c_3(x)]$$

subject to $c_1(x), c_2(x), c_3(x) \geq 0$, $c_1(x) + c_2(x) + c_3(x) = 1$ (probability 2-simplex).

T $=$ **3** flavors, **2-simplex**:



T $=$ **2** flavors, **1-simplex**:

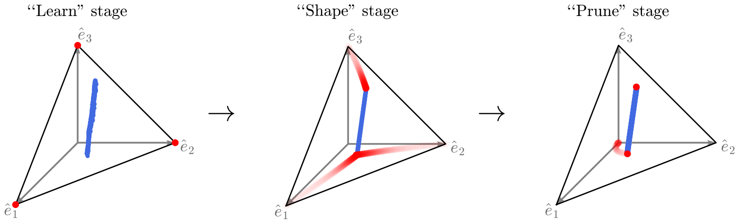


[de la Fuente, Thaler, 2607.24921]

[Developed independently by Bonnet-Guerrini, Ioannou-Nikolaides, Petersen, Piuri, 2607.24943] Talk today!

We extract the vertices using **Simplex Demixing**

Our approach uses regularization terms to attract vertex weights (in red) towards the distribution (in blue).



From the final vertex positions, we solve for the mixing fractions of operational flavors $p_t(x)$ in the mixtures $p_m(x)$,

$$p_m(x) = F_{mt} p_t(x).$$

[de la Fuente, Thaler, 2607.24921]

Case study 1: u, d, and g flavors from $M = 3$ samples

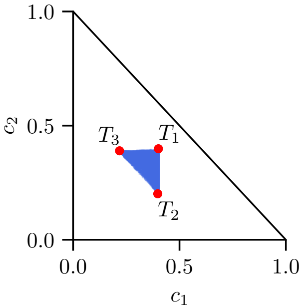
We use artificial mixed samples of PYTHIA-generated u-quark, d-quark, and gluon jets: $p_1(x)$, $p_2(x)$, and $p_3(x)$

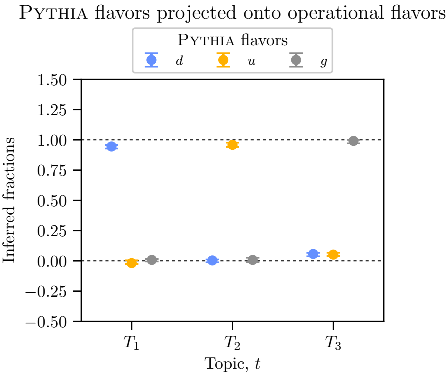
[Bierlich et al., 2203.11601],

and a Particle Flow Network (PFN-ID) architecture.

[Komiske, Metodiev, Thaler, 1810.05165].

The classifier learns a 2-simplex with $T = 3$ flavors:

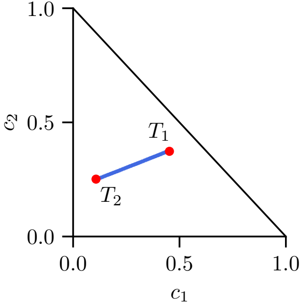


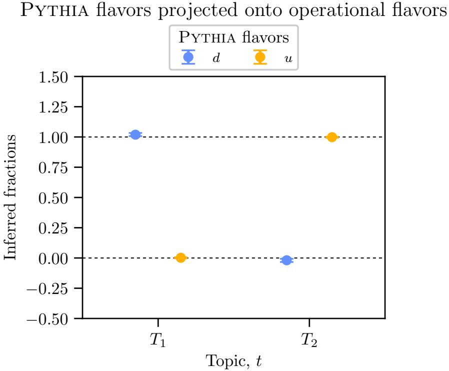


Edge case: what if the gluon flavor is missing?

Use three artificial mixed samples of PYTHIA-generated u-quark and d-quark jets (no gluon jets).

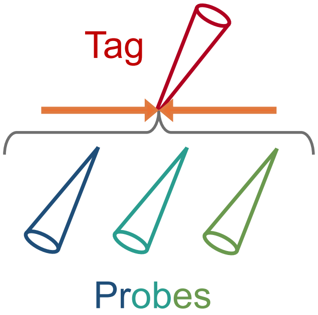
The classifier learns a 1-simplex with $T = 2$ latent flavors.





Case study 2: multiple flavors from $M = 14$ samples

Tag-and-probe: each PYTHIA dijet event is split into a "tag" jet and a "probe" jet by randomizing over ϕ.

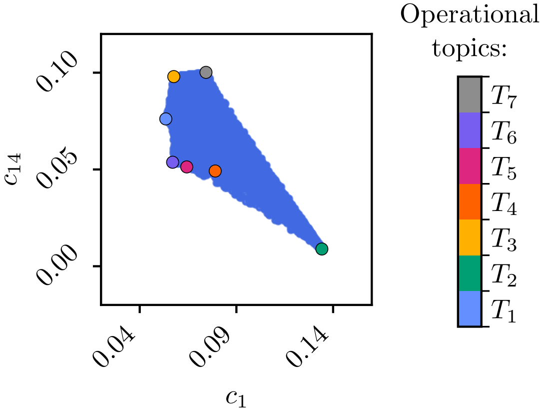


Each sample is defined by collecting all probe jets with the same tag-jet flavor and in the same probe-jet η bin.

(7 light flavors) $\times$ (2 η bins) = 14 samples

We classify the mixtures assuming $T = 7$ vertices

The classifier distribution is bounded by 7 vertices within the probability 13-simplex, given by $c_1 + \ldots + c_{14} = 1$. One of the 2D projections is shown below:

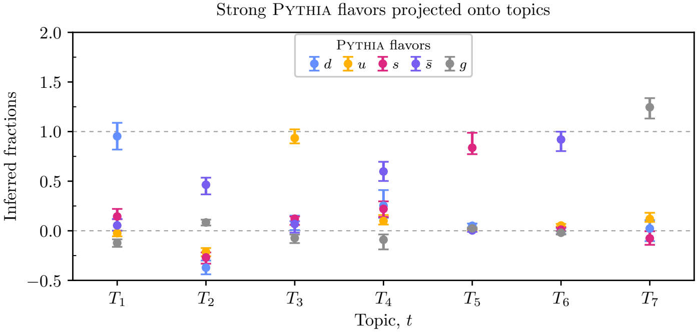


The d, u, s, $\bar{s}$, and g flavors map to five vertices

Like before, we write the unphysical PYTHIA flavors $p_c(x)$ as a combination of operational flavors (topics) $p_t(x)$,

$$p_c(x) = \widehat{p}(t|c)\, p_t(x).$$

We bootstrap the training data to get error bars on the $\widehat{p}(t|c)$.

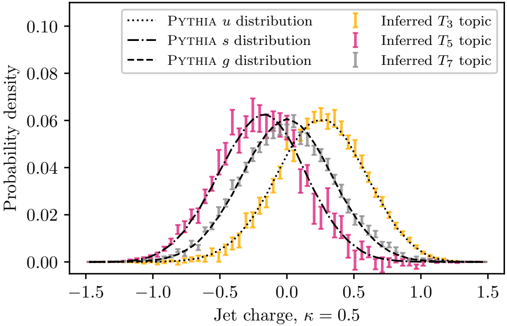


The topic substructure distributions match PYTHIA

We use F_{mt} to solve for the topic distributions $p_t(\mathcal{O})$ from the binned mixture distributions $p_m(\mathcal{O})$ over any observable $\mathcal{O}$.

$$p_t(\mathcal{O}) = (F^+)_{tm}\, p_m(\mathcal{O}).$$

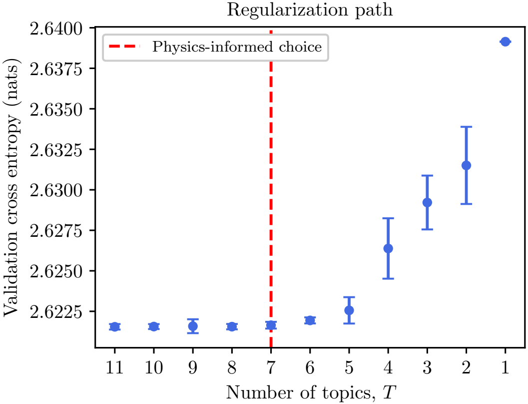
The bootstrap yields the error bars on the bins.



Physics conclusion: we can extract multiple jet flavors and their properties from data.

$T = 7$ is a single point along the regularization path

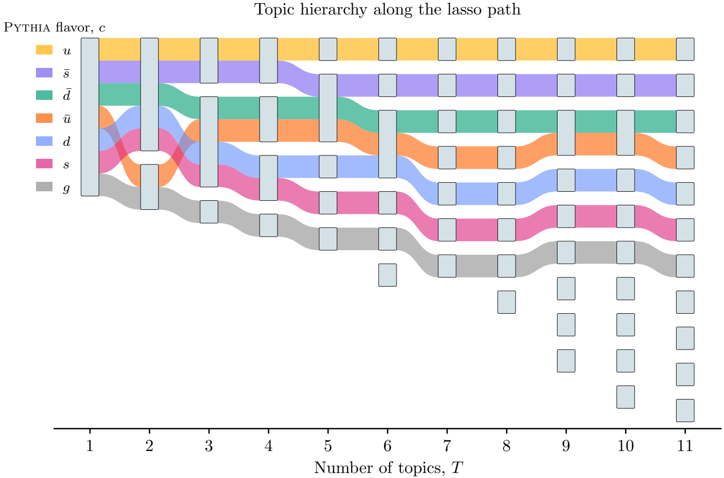
We record the validation loss performance for different values of T.



1. How consistent are our results for other T values on the plateau?
2. How does the model performance degrade at lower T?

The classifier remains interpretable for other values of T

For each T, we write the unphysical Pythia flavors $p_c(x)$ as a combination of operational flavors $p_t(x)$, $p_c(x) = \widehat{p}(t|c)p_t(x)$.

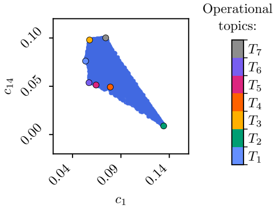


d, u, s, $\bar{s}$, and g are strongly identifiable across different T.

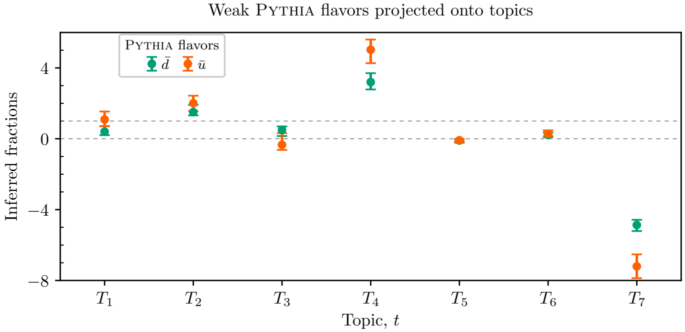
Summary. Email me at grego137@mit.edu!

1. We have provided a hadron-level, data-driven definition of multiple jet flavors.
2. Jet flavors are vertices of a classifier distribution in our framework.
3. In a toy example, we extracted down, up, and gluon flavors that agree with the PYTHIA parton-level flavors.
4. In a quasi-realistic example, we identified the down, up, strange, anti-strange, and gluon flavors in a data-driven way.
5. **Next steps:** make the procedure fully deployable on real data.

Read our paper here!



The $\bar{d}$ and $\bar{u}$ flavors are weakly identifiable



The $\bar{d}$ and $\bar{u}$ flavors are weakly identifiable because they have nearly the same abundance in each mixture.