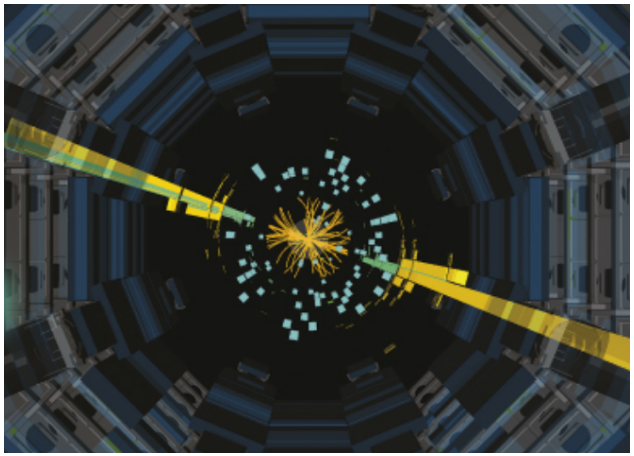
Jet Topology

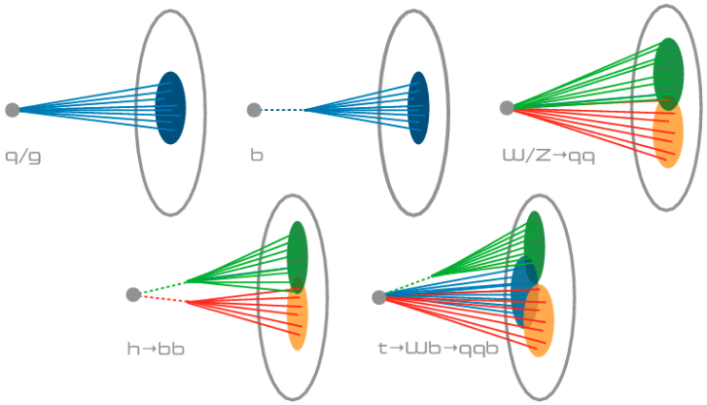
Lingfeng Li (HKUST)

Based on arXiv.2006.12446
In collaboration with Tao Liu and Si-Jun Xu

Jan. 15, 2021
IAS Program on High Energy Physics

Introduction: What are Jets?





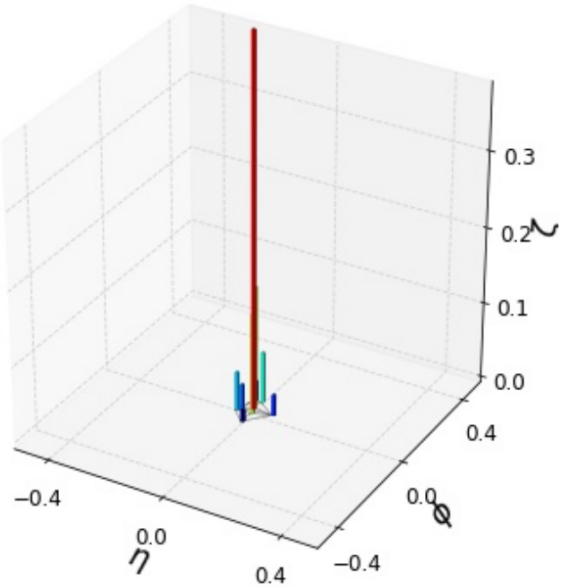
Perturbative interactions + non-perturbative confinement processes.

Strong motivation exists to look into its structure.

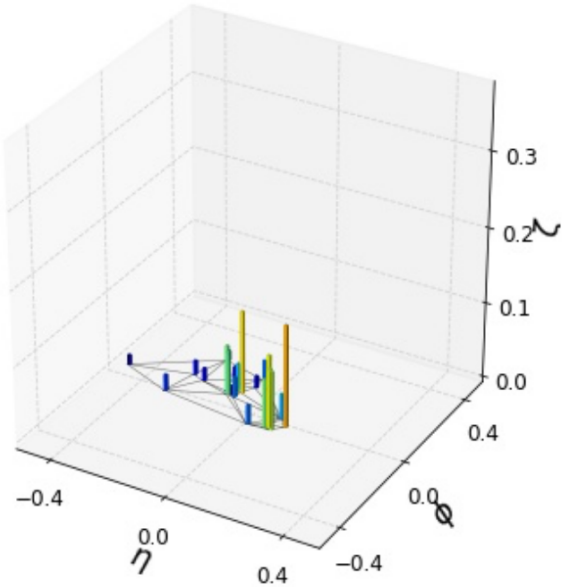
Jet profile inherits from the kinematics of the shower-produced partons. It is in essence a manifestation of the nature (flavor, QCD charge, momentum, etc.) of its ancestral particle $\Rightarrow$ assist jet tagging & deepen our understanding on jet dynamics.

[Moreno et al.(2020)Moreno, Nguyen, Vlimant, Cerri, Newman, Periwal, Spiropulu, Duarte, and Pierini]

A First Look At Different Jets

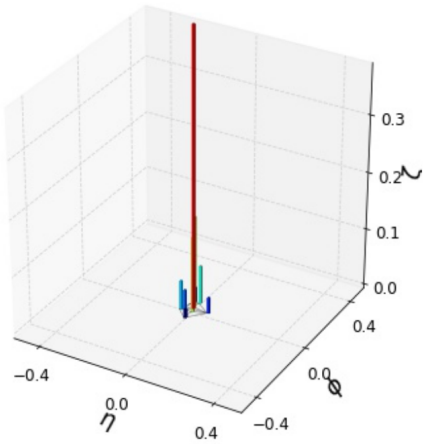


Parashorea chinensis

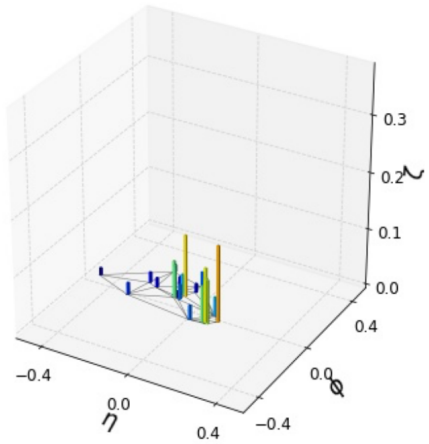


Ficus microcarpa

Describe A Jet with Its Topology



quark jet



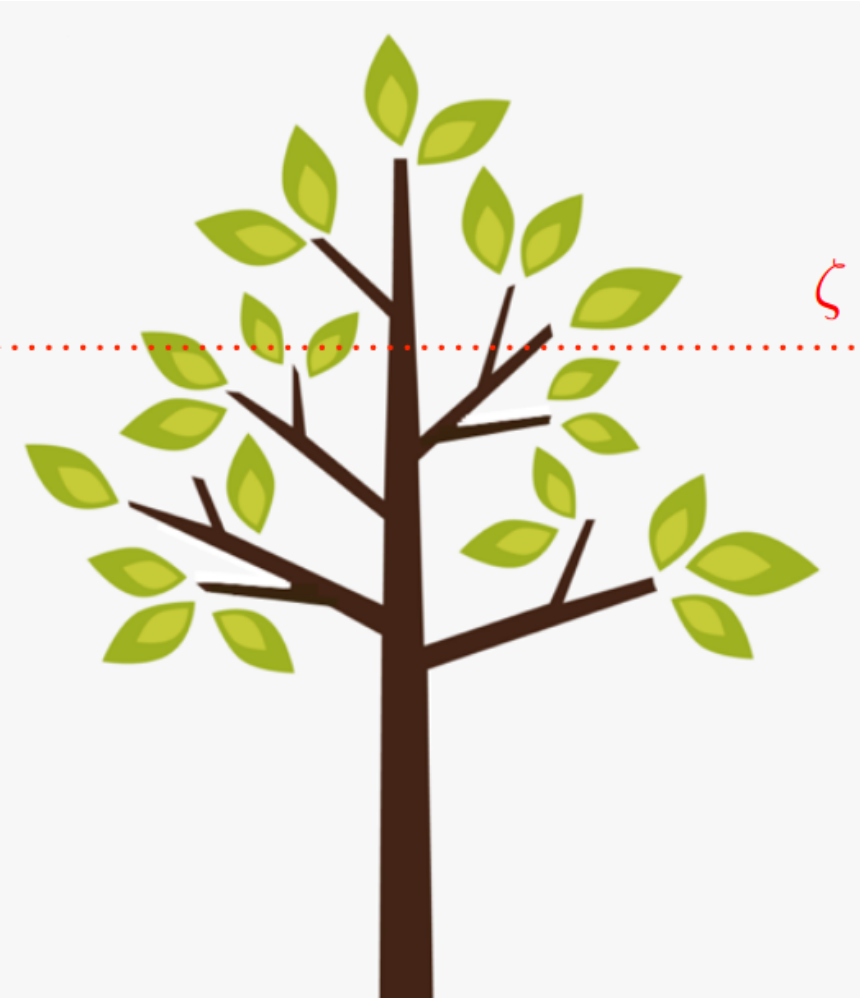
gluon jet

- "geometric" features: shape, size, position, etc.
- "topological" features: multiplicity, connectivity

What make these trees look different? - Structure of their branches

The Topology of a Tree

How to describe a real tree in terms of topology?



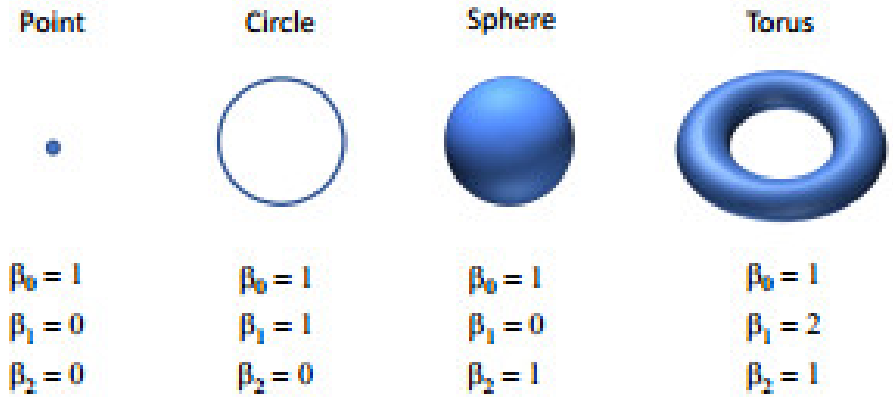
ζ: The threshold where the topological feature is determined (here is the physical height)

- Top ($\zeta \to \inf$): nothing
- Current value: 3 disconnected branches
- Bottom ($\zeta \to 0$): the trunk

Betti Numbers

Betti number [Betti(1870)] is the rank of the i-th Homology group in algebraic topology. The latter characterizes the topology of a space based on the relationship between the cycles and its boundaries.

- $\beta_0 \approx \#$ of branches
- $\beta_1 \approx \#$ of 2D "holes"
- $\beta_2 \approx \#$ of 3D "cavities"
- .....

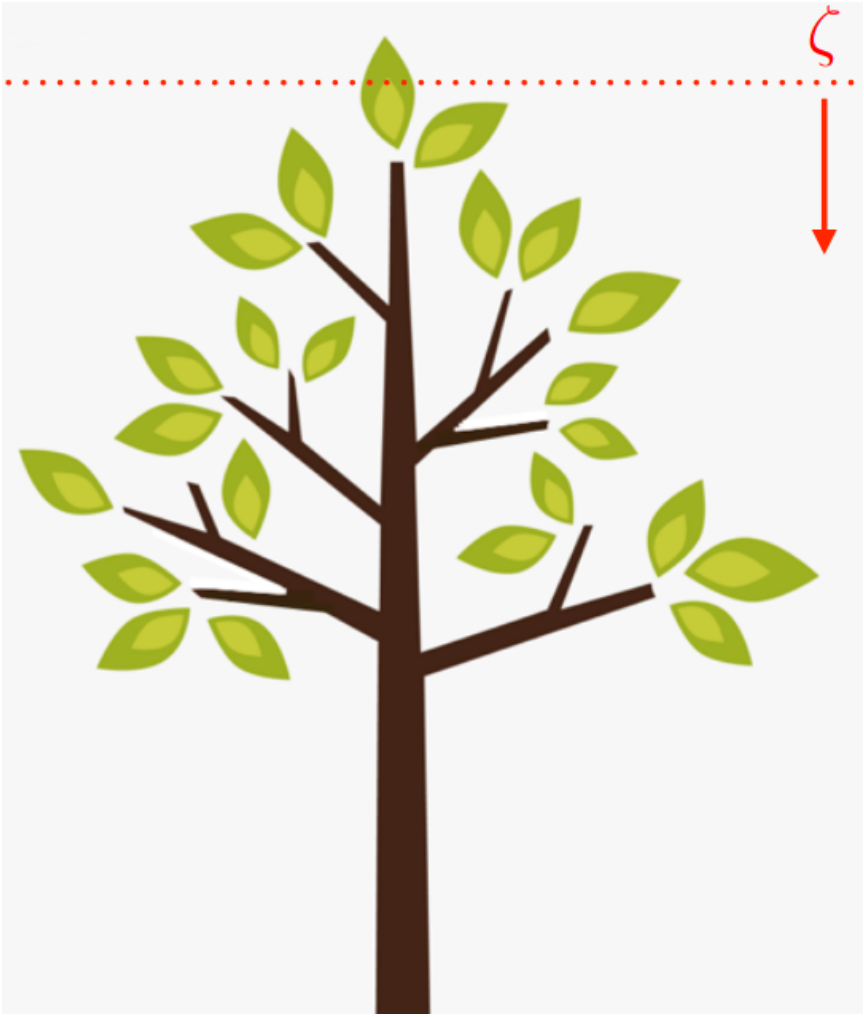


plot from [Siopsis(2019)]

Euler-Poincaré formula

$$\chi = \beta_0 - \beta_1 + \beta_2 - \beta_3 \ldots$$

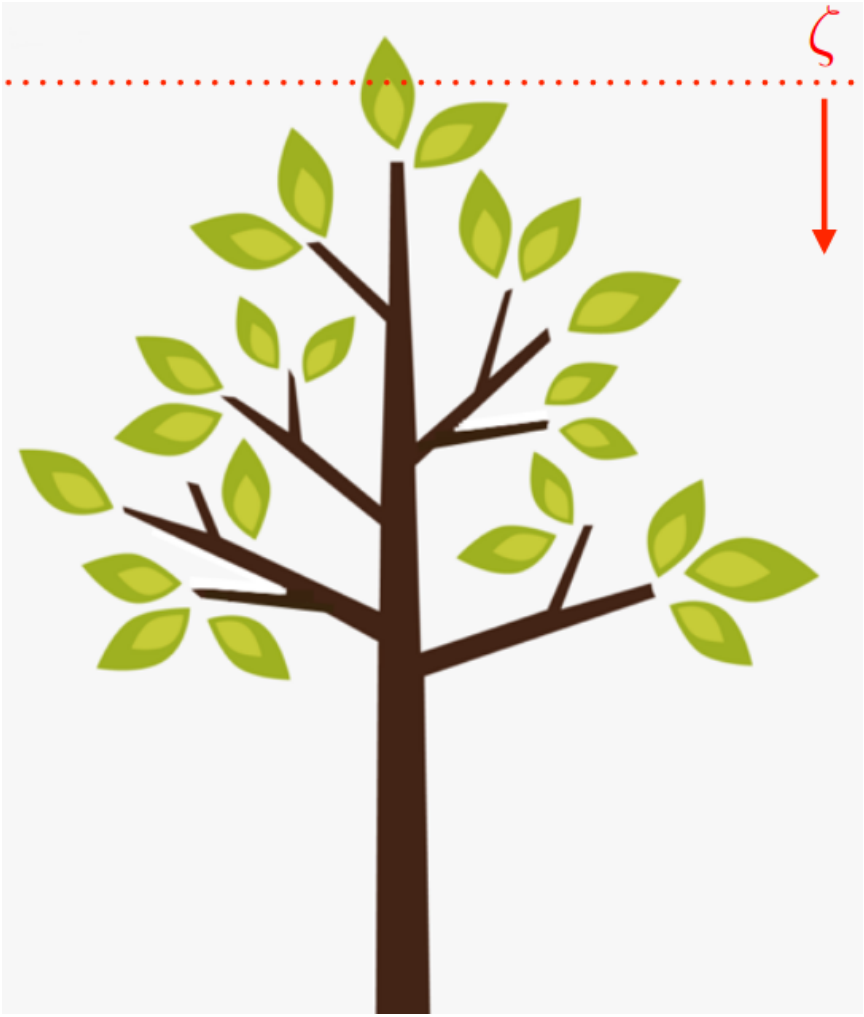
The Topology of a Tree (II)



As the threshold ζ decreases from above the top, one can keep the record of how each feature (e.g. a branch) evolves:

- birth (the height of its tip)
- growth (the increase of its weight above the threshold)
- death (the height of the place where it merges with a bigger branch)

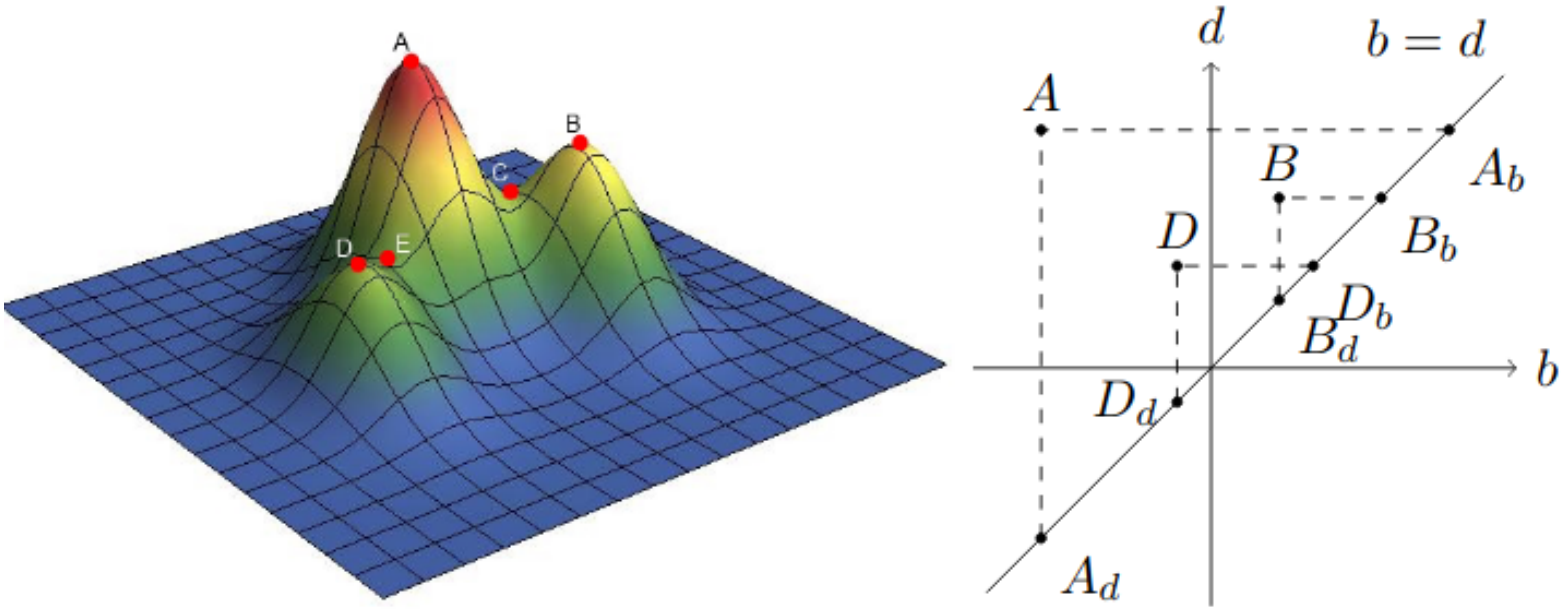
The Topology of a Tree (III)



Some other structure exists, e.g., a cavity. In terms of a jet it correspond to the feature from β_1.

Persistent Homology

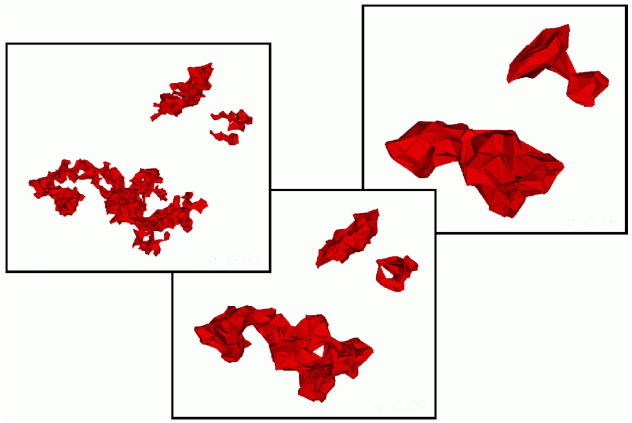
More generally, given a scanning scale ζ and a function:



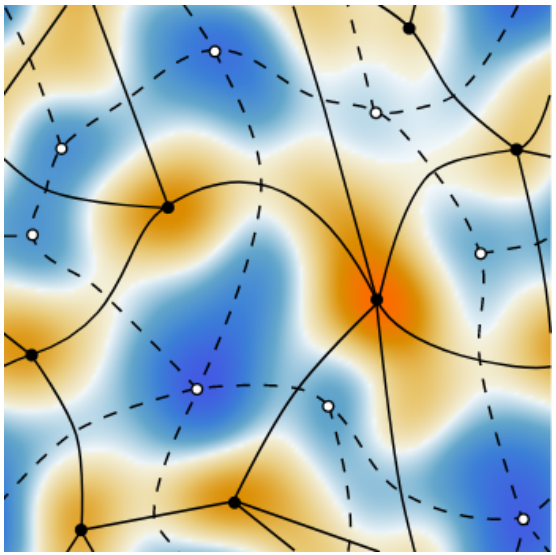
[Feldbrugge et al.(2019)Feldbrugge, van Engelen, van de Weygaert, Pranav, and Vegter]

The evolution (birth/death) of features now encoded in a plane.

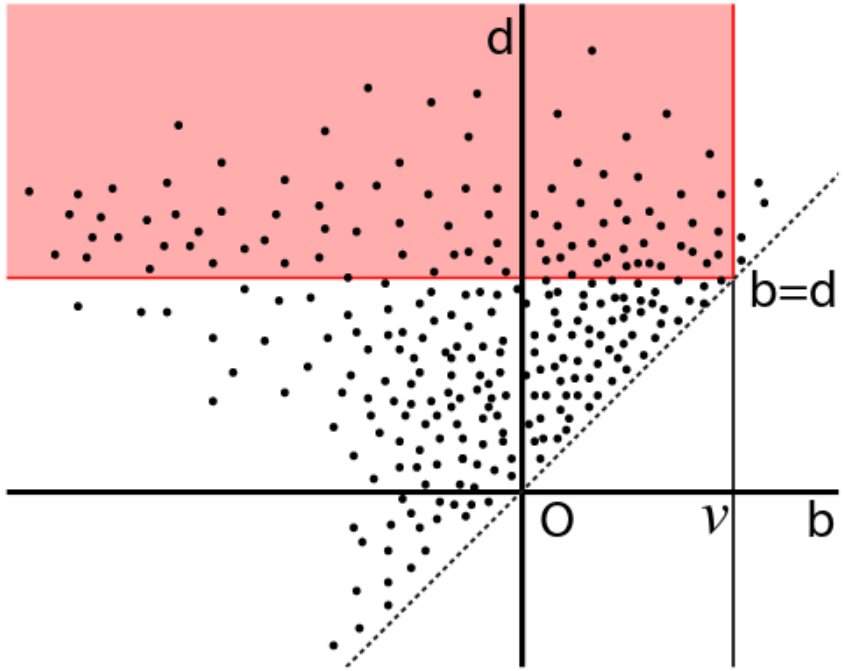
Persistent Homology (II)



In cosomolgy[van de Weygaert et al.(2011), Cole and Shiu(2018)]

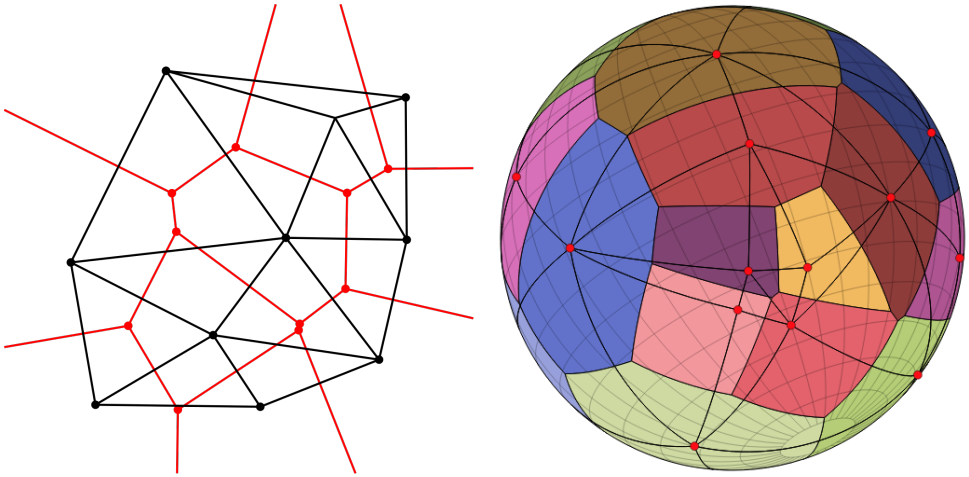


Related to the topology of a graph by the Morse theorem [Morse(1925)]



The sum of features in the red area is the corresponding Betti number $\beta_i(\zeta)$.

First Step: Delaunay Triangulation(DT)

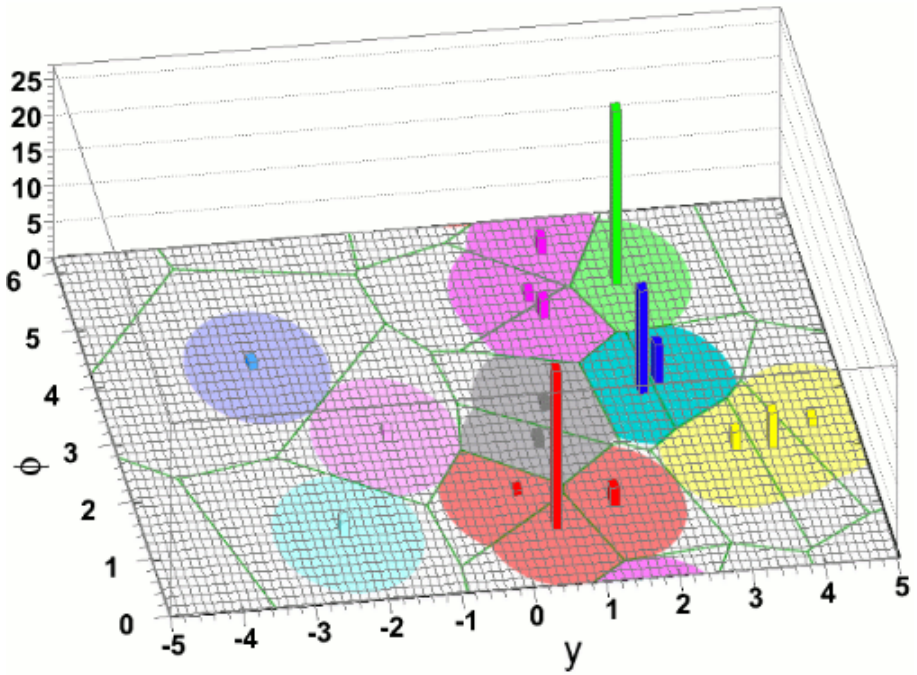


[Wikipedia & jasondavies.com]

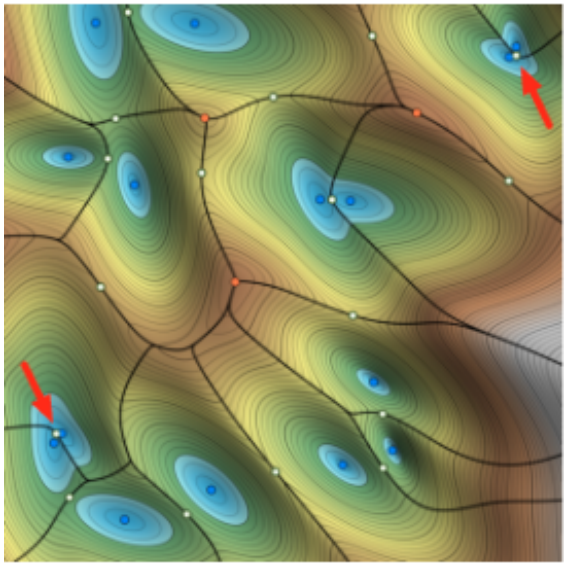
Delaunay Triangulation (DT) is the dual of a Voronoi Diagram (VD)

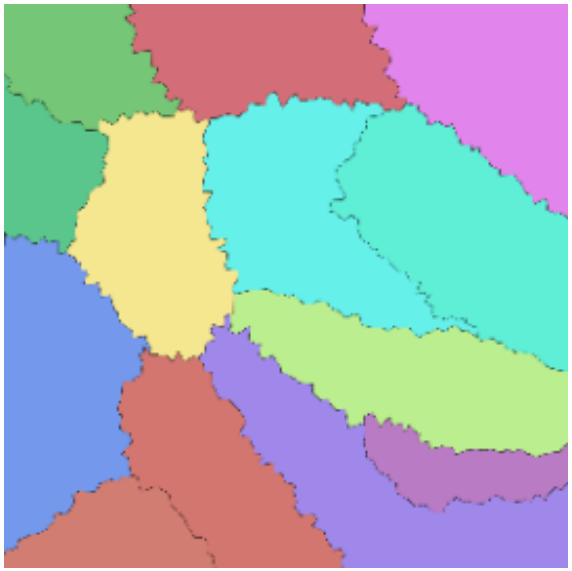
The cell of Voronoi diagram (VD) in jet physics has a natural explanation as the passive catchment area for each jet constituent, for the kt jet clustering

[Cacciari et al.(2008)Cacciari, Salam, and Soyez]



Delaunay Triangulation (II)



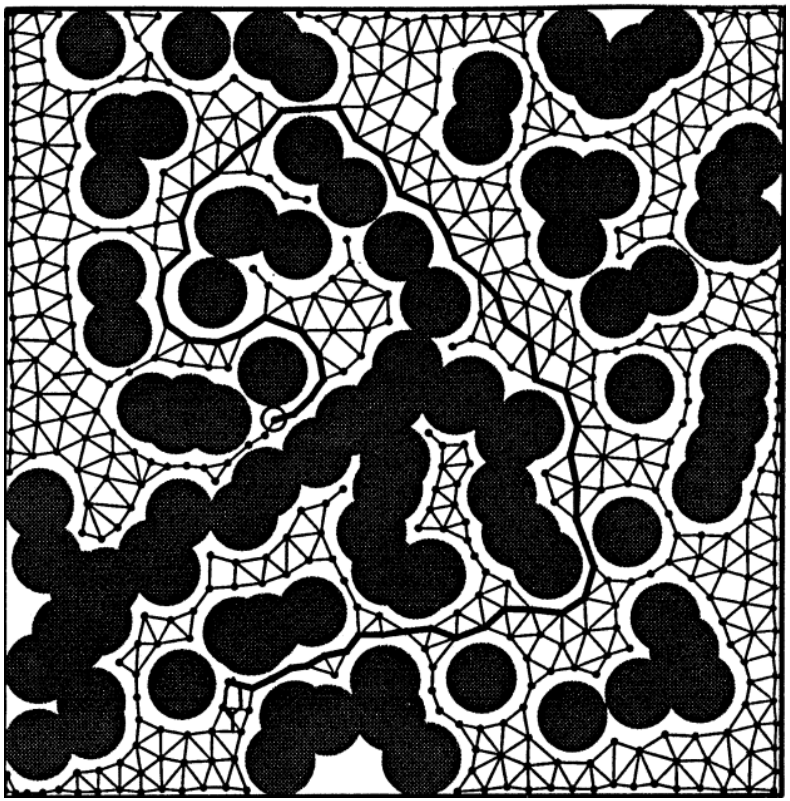


Widely used in Computational Topology

[Martinetz and Schulten(1994), Correa and Lindstrom(2011),

Maljovec et al.(2013)Maljovec, Saha, Lindstrom, Bremer, Wang, Correa, and Pascucci]...

DT authentically recovers the topology of a certain object in the limit of even sampling.



Relation With Morphological Studies

Scalar MF of a 2D shape

- W_0: Area ($\int_K d\vec{r}$)
- W_1: Circumference ($\int_{\partial K} d\vec{r}$)
- W_2: Euler Characteristic ($\int_{\partial K} \kappa d\vec{r}$)

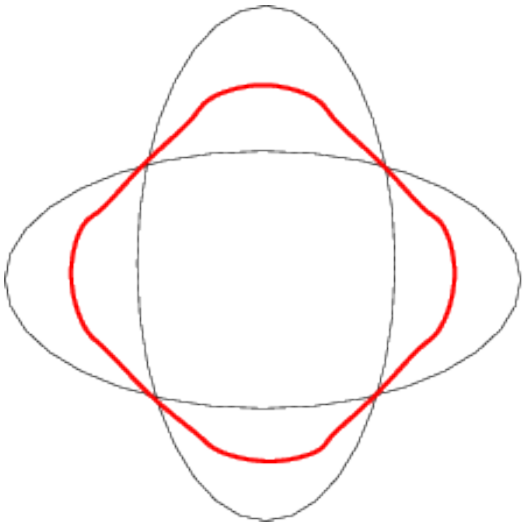
Also vector & tensor MFs

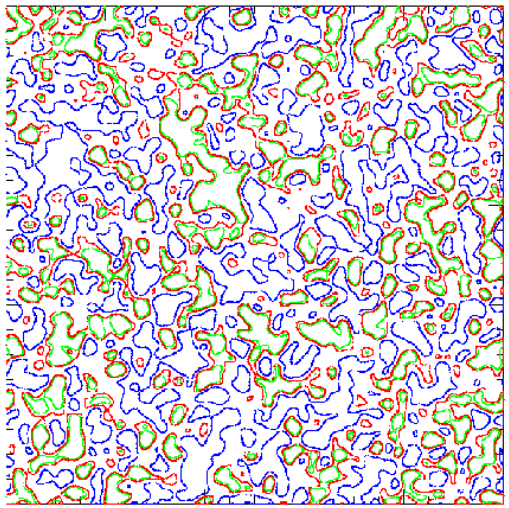
In the context of Minkowski Functionals (MF):[Minkowski(1903), Schmalzing and Gorski(1998)]

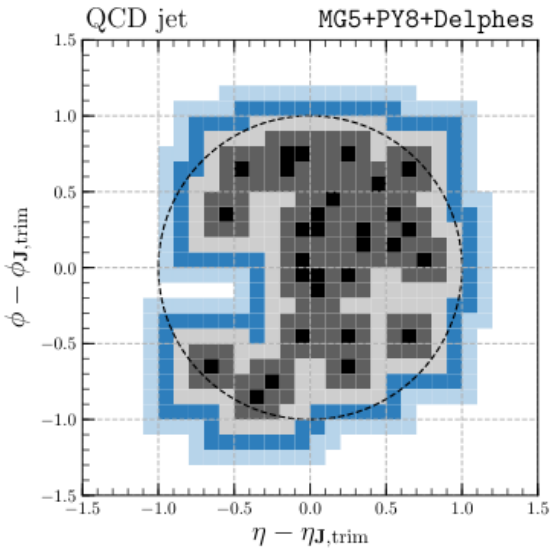
Application in cosmology:

[Mecke et al.(1994)Mecke, Buchert, and Wagner,

Goyal et al.(2020)Goyal, Chingangbam, and Appleby]...







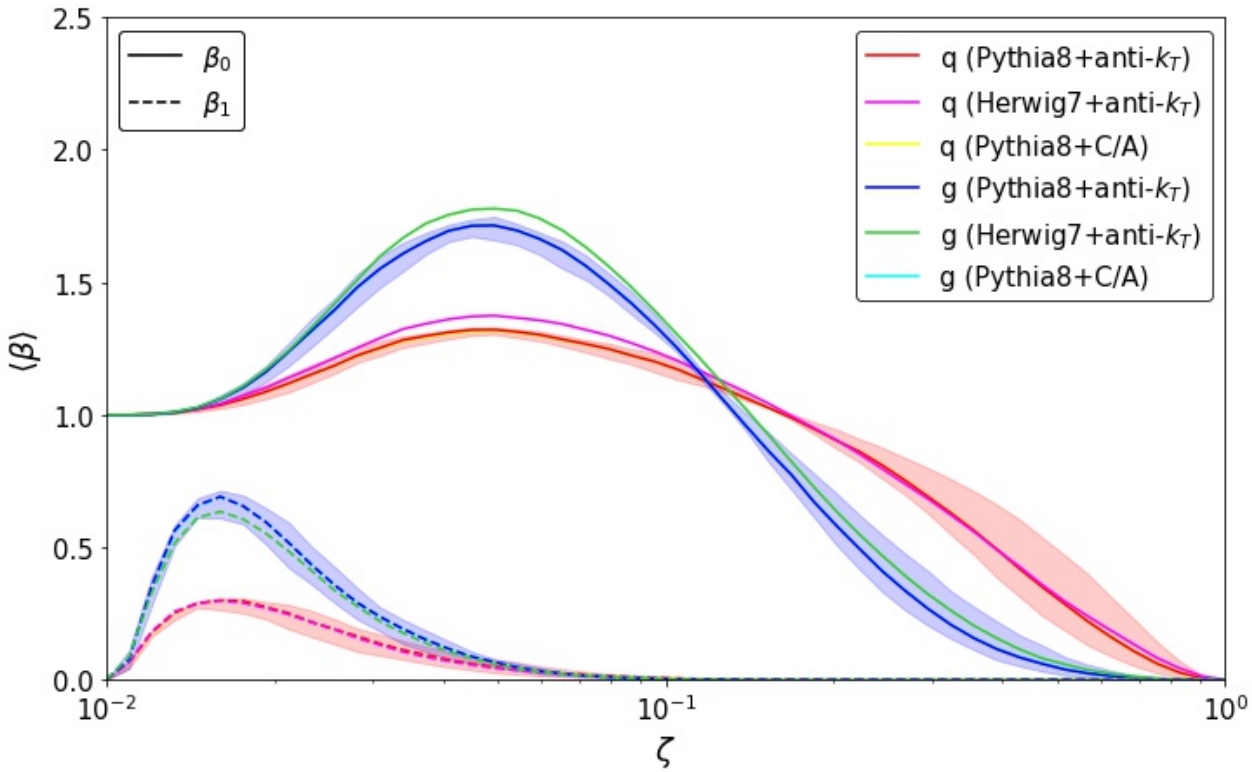
See also Mihoko's talk yesterday.

[Lim and Nojiri(2020)]

Numerical Study: Set up

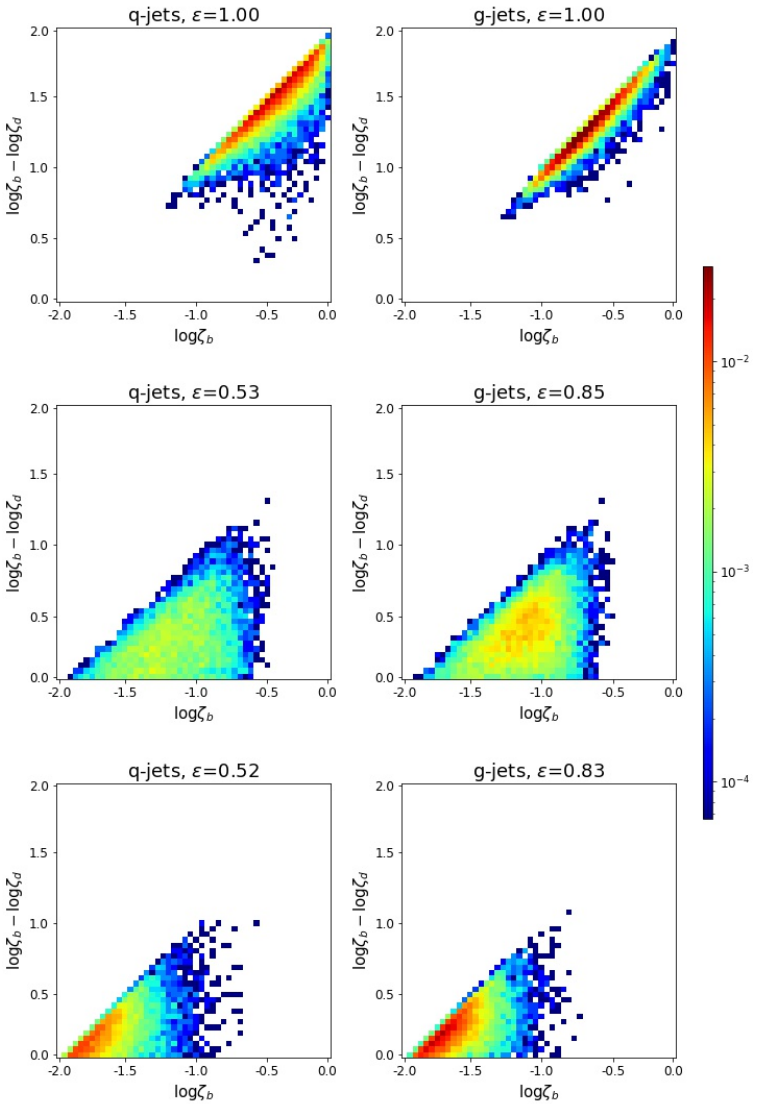
- The sample jets are generated from the $pp \to Z + q/g$ events with $Z \to \nu\bar{\nu}$.
- Showered with Pythia 8 and clustered with anti-kT algorithm ($\Delta R = 0.6$), unless otherwise specified.
- The q and g jets are then equally selected from five 50GeV-wide, non-overlapped bins, with the jet p_T ranging from 100-350 GeV.
- The scale threshold is defined to be p_T (normalized by $p_{T,\text{jet}}$).
- To ensure these topological invariants to be collinear- and infrared-safe technically, we graph the jets by merging their collinear constituents with $\Delta R_{ij} \leqslant 0.01$ and then vetoing the soft ones with $\zeta_i \leqslant 0.01$.
- No dedicated detector simulation (part of the detector effects have been resolved into the said cutoffs of ΔR_{ij} and ζ_i)

Betti Number Evolution



- As ζ decreases, the q-jet $\langle\beta_0\rangle$ rises up earlier $\Rightarrow$ slow evolution (vs. g later/ higher/ sharper)
- ζ further decreases $\Rightarrow$ soft jet constituents connect parts, decreasing of $\langle\beta_0\rangle$ and increasing of $\langle\beta_1\rangle$ (vs. g, more holes)
- Converges to $\beta_0 = 1$, $\beta_1 = 0$ when $\zeta \to 10^{-2}$
- Robust against the parton-shower models and jet clustering algorithms

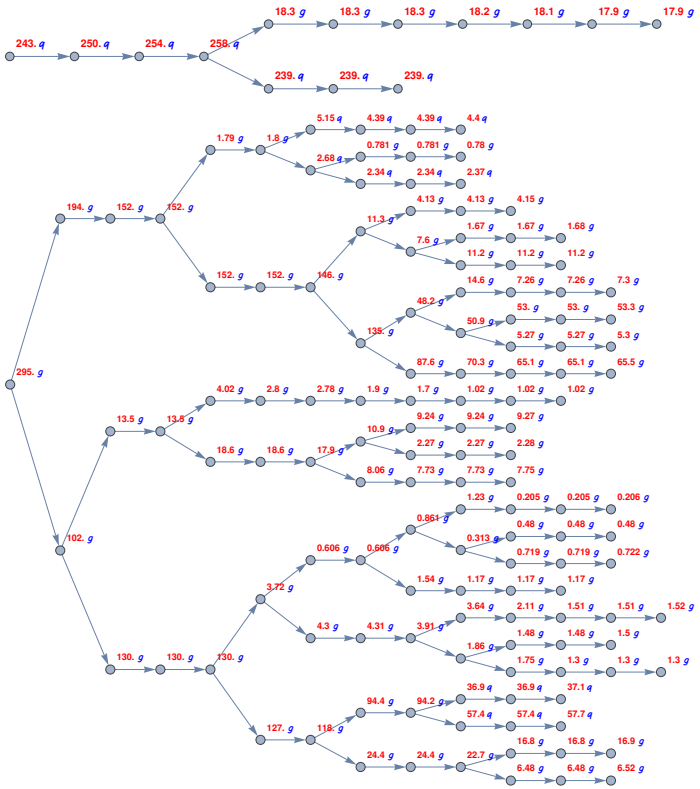
Persistent Homology of Different Jets



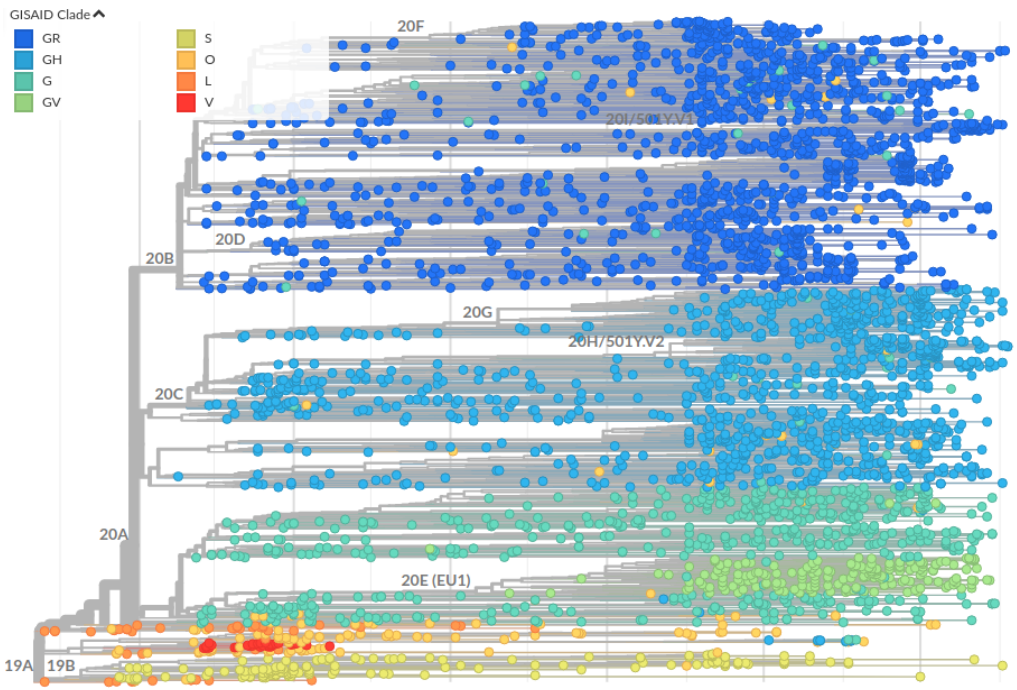
q Jets vs. g Jets

- First β_0 features:
  q: Larger ζ_b and longer lifetime (ζ_b/ζ_d) vs. g: Smaller ζ_b and shorter lifetime
- Second β_0 and first β_1 features:
  q: $\lesssim 55\%$ a chance to develop vs. g: $> 80\%$
  q: Larger ζ_b/ζ_d vs. g: Smaller ζ_b/ζ_d

Phylogenetic Tree Structure

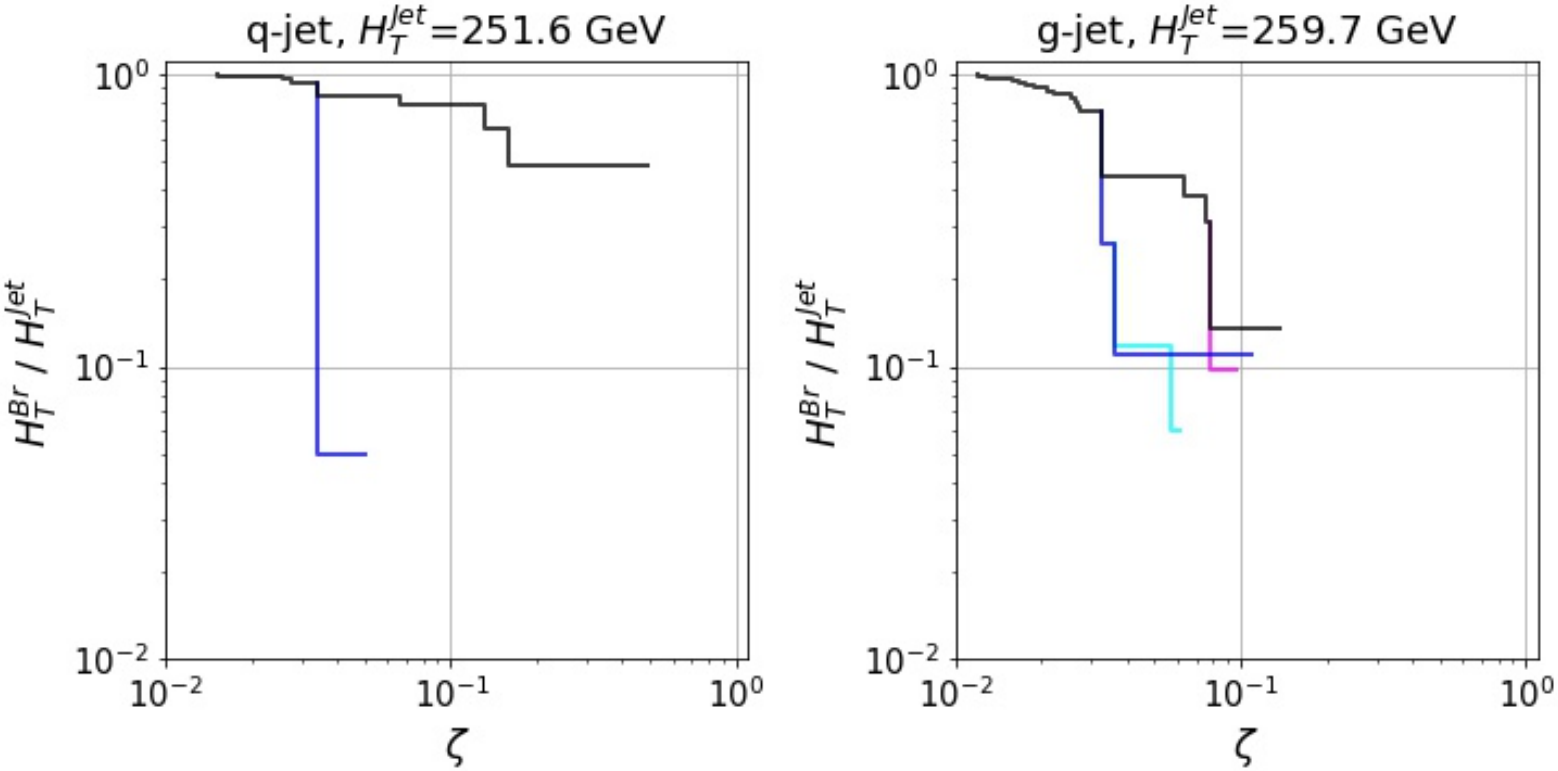


PY8 shower history of q and g



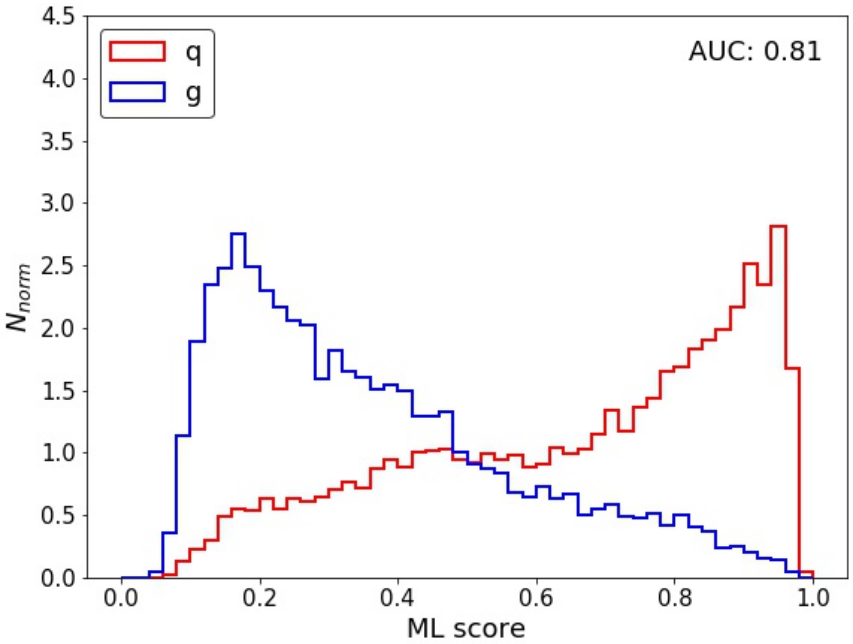
Phylogenetic Tree of COVID-19 [GISAID(2008)]

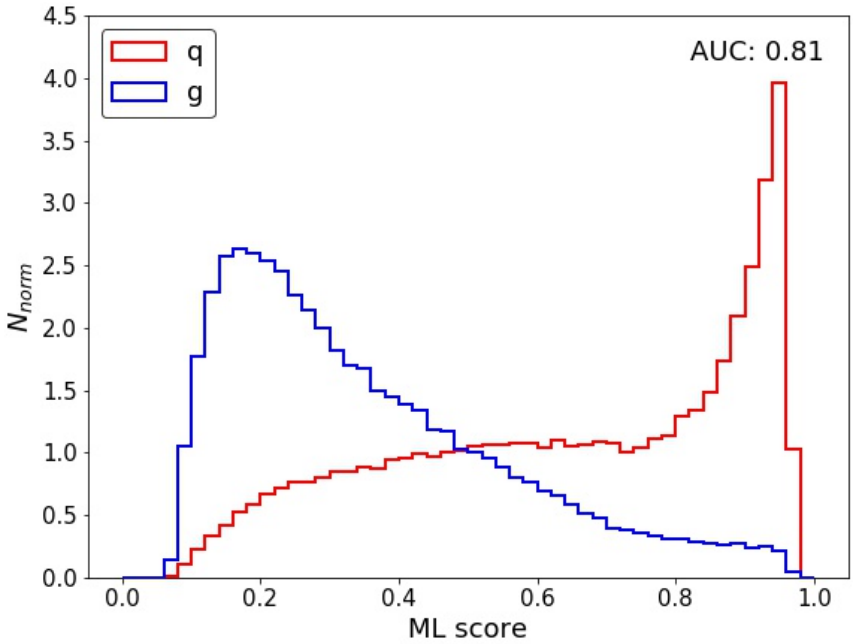
Phylogenetic Tree Structure (II)



x-axis: threshold z; y-axis: branch weight (scalar sum of p_T)

Classification Based on Tree Structure





Self-generated samples
$p_T \in [100, 350]$ GeV

Public samples
$p_T \in [500, 550]$ GeV

[Komiske et al.(2019)Komiske, Metodiev, and Thaler]

Inputs: the first six β_0 features (birth, death, weight at death, ID of merger) + the first five β_1 feature

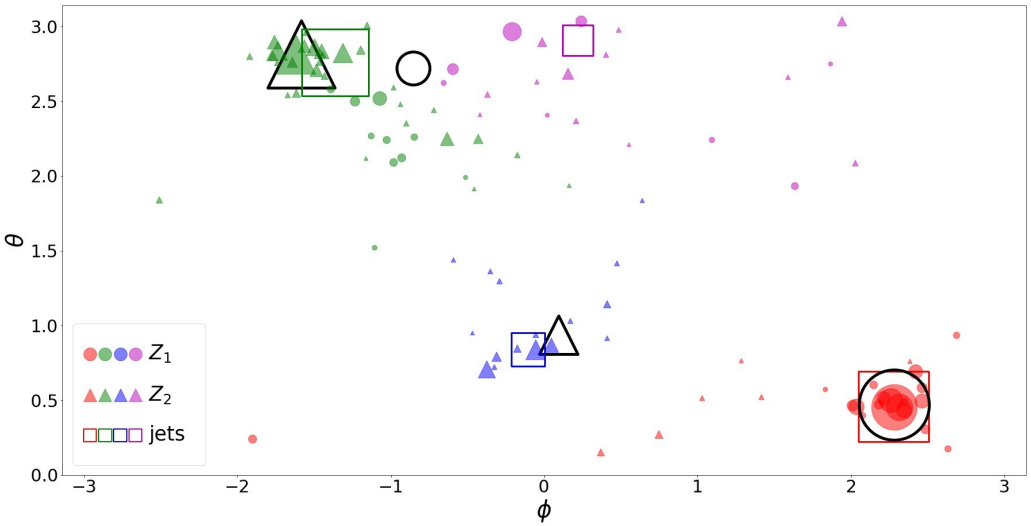
Outlook: Further Jet Studies

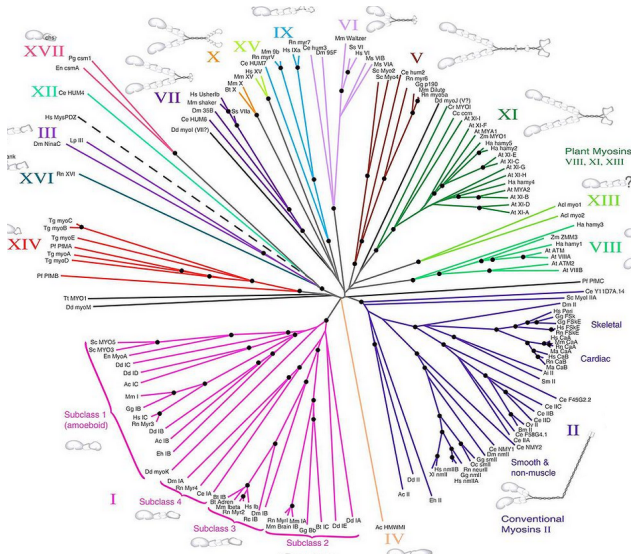
- Further optimization is possible:
  - Better reconstruction of the jet profile than DT?
  - Synergizing the missed "geometric/morphological" messages in this study such as spatial distribution of jet branches, into the construction of $q - g$ jet classifiers?
  - Proper simulation of detector effects: persistence may serve as a grooming tool to suppress soft contaminations
- New taggers for heavy jets (top, Higgs, W/Z, etc.)
- A new handle to probe jet dynamics

Outlook: Event-Level Studies

Extend the application of this tool to study the collider-event topology [Li et al.(2020)Li, Li, Liu, and Xu]:

- Event branches (or "Topological" jets) and their evolution
- These topological features are insusceptible to Lorentz boost along beam direction $\Rightarrow$ well-applied in both ee and pp colliders.





plot from Wikipedia

Summary

- We introduce persistent homology to characterize topological structure of jets
- Topological invariants measure multiplicity and connectivity of jet branches at a given scale threshold, while their persistence records the evolution of each topological feature as this threshold varies
- Reconstruction of the topology based jet phylogenetic tree
- We demonstrate these points in a benchmark scenario of light-quark versus gluon jets and then develop classifiers to distinguish between them based on these topological features. We show that encouraging sensitivities can be achieved.
- We hope - this topological method will open a new angle to look into jet (and event) physics.

Enrico Betti.
Sopra gli spazi di un numero qualunque di dimensioni.
1870.
doi: 10.1007/BF02420029.

Matteo Cacciari, Gavin P. Salam, and Gregory Soyez.
The Catchment Area of Jets.
JHEP, 04:005, 2008.
doi: 10.1088/1126-6708/2008/04/005.

Alex Cole and Gary Shiu.
Persistent Homology and Non-Gaussianity.
JCAP, 03:025, 2018.
doi: 10.1088/1475-7516/2018/03/025.

Carlos Correa and Peter Lindstrom.
Towards robust topology of sparsely sampled data.
IEEE Transactions on Visualization and Computer Graphics,
17(12):1852–1861, 2011.

Job Feldbrugge, Matti van Engelen, Rien van de Weygaert, Pratyush Pranav, and Gert Vegter.
Stochastic Homology of Gaussian vs. non-Gaussian Random Fields: Graphs towards Betti Numbers and Persistence Diagrams.
JCAP, 1909(09):052, 2019.
doi: 10.1088/1475-7516/2019/09/052.

GISAID.
Global initiative on sharing all influenza data (www.gisaid.org), 2008.
URL https://www.gisaid.org.

Priya Goyal, Pravabati Chingangbam, and Stephen Appleby.
Morphology of CMB fieldseffect of weak gravitational lensing.
JCAP, 2002:020, 2020.
doi: 10.1088/1475-7516/2020/02/020.

Patrick T. Komiske, Eric M. Metodiev, and Jesse Thaler.
Energy Flow Networks: Deep Sets for Particle Jets.
JHEP, 01:121, 2019.
doi: 10.1007/JHEP01(2019)121.

Lingfeng Li, Ying-Ying Li, Tao Liu, and Si-Jun Xu.
Learning Physics at Future e^-e^+ Colliders with Machine.
4 2020.

Sung Hak Lim and Mihoko M. Nojiri.
Morphology for Jet Classification.
10 2020.

Dan Maljovec, Avishek Saha, Peter Lindstrom, Peer-Timo Bremer, Bei Wang, Carlos Correa, and Valerio Pascucci.
A comparative study of morse complex approximation using different neighborhood graphs.
In *Workshop on Topological Methods in Data Analysis and Visualization (accepted)*, 2013.

Thomas Martinetz and Klaus Schulten.
Topology representing networks.
Neural Networks, 7(3):507–522, 1994.

K.R. Mecke, T. Buchert, and H. Wagner.
Robust morphological measures for large scale structure in the universe.
Astron. Astrophys., 288:697–704, 1994.

Hermann Minkowski.
Volumen und oberfläche.
Mathematische Annalen, 57:447–495, 1903.
URL http://eudml.org/doc/158108.

Eric A. Moreno, Thong Q. Nguyen, Jean-Roch Vlimant, Olmo Cerri, Harvey B. Newman, Avikar Periwal, Maria Spiropulu, Javier M. Duarte, and Maurizio Pierini.
Interaction networks for the identification of boosted $H \to b\bar{b}$ decays.

Phys. Rev. D, 102(1):012010, 2020.
doi: 10.1103/PhysRevD.102.012010.

Marston Morse.
Relations between the critical points of a real function of independent variables.
Transactions of the American Mathematical Society, 27(3): 345–396, 1925.

Jens Schmalzing and Krzysztof M. Gorski.
Minkowski functionals used in the morphological analysis of cosmic microwave background anisotropy maps.
Mon. Not. Roy. Astron. Soc., 297:355, 1998.
doi: 10.1046/j.1365-8711.1998.01467.x.

George Siopsis.
Quantum topological data analysis with continuous variables.

Foundations of Data Science, 1:419, 2019.

doi: 10.3934/fods.2019017.

Rien van de Weygaert et al.
Alpha, Betti and the Megaparsec Universe: on the Topology of the Cosmic Web, volume 14, pages 60–101.
2011.