# Snowmass2021 - Letter of Interest

# *Topological Aspects of Jets and Events at Colliders*

**Thematic Areas:**

■ (TF07) Collider phenomenology
■ (CompF3) Machine Learning
■ (EF01) EW Physics: Higgs Boson properties and couplings
■ (EF04) EW Precision Physics and constraining new physics

**Contact Information:**

Lingfeng Li (Jockey Club Institute for Advanced Study, The Hong Kong University of Science and Technology, Hong Kong S.A.R.) [iaslfli@ust.hk]
Tao Liu (Department of Physics, The Hong Kong University of Science and Technology, Hong Kong S.A.R.) [taoliu@ust.hk]
Si-Jun Xu (Department of Physics, The Hong Kong University of Science and Technology, Hong Kong S.A.R.) [sxuaw@connect.ust.hk]

**Abstract:** In the previous work[1] we introduced persistent Betti numbers to characterize topological structure of jets. These topological invariants measure multiplicity and connectivity of jet branches at a given scale threshold, while their persistence records evolution of each topological feature as this threshold varies. This study provides a topological tool to develop jet taggers and opens a new angle to look into jet physics. As an application, we plan to use this tool to construct topological light-quark and gluon-jet classifiers and topological taggers for boosted heavy jets. Moreover, we plan to extend the study on jet topology to jet morphology by properly including geometric elements into this topological observable scheme. At last, we would apply this methodology to event-level data analysis at colliders.

# 1 Introduction

The spray of constituents in a jet are a manifestation of the nature (flavor, QCD charge, four momentum, etc.) of its ancestral particle. Their energy profile inherits from the kinematics of the shower-produced partons. Measuring jet structure thus may decipher the relevant information on its ancestral particle. These features can on the one hand assist tagging jet flavor, and on the other hand, deepen our understanding on jet dynamics. We introduce persistent Betti numbers[2;3] to characterize topological structure of jets. These topological invariants measure multiplicity and connectivity of jet branches at a given scale threshold, while their persistence records evolution of each topological feature as this threshold varies.

The topological features of jet structure can be captured using a connected planar graph, where the jet constituents are viewed as sparse samplings of jet profile and define the vertices of this graph. Here we take Delaunay triangulation (DT)[4] to make the jet DT graph $G_{\rm ref}$. The topological features of jet branches above and below $\zeta$ are then encoded as Betti numbers of the superlevel $G(\zeta)$ and sublevel sets $\bar{G}(\zeta)$, defined as

$$G(\zeta) = G_{\rm ref}\{\text{jet constituents } i \mid \zeta_i \geq \zeta\}\ ,\ \bar{G}(\zeta) = G_{\rm ref}\{\text{jet constituents } i \mid \zeta_i < \zeta\}\ . \tag{1}$$

Betti number $\beta_i$[5] is the rank of the $i$-th Homology group in algebraic topology. In our case, $\beta_0$, $\beta_1$ and $\beta_2$ count the numbers of the connected components, holes and voids of the topological space, respectively. For the convenience of discussions, we define $\zeta = \frac{p_T}{p_T^{\rm Jet}}$ as a normalized scale. The averaged Betti numbers of jet constituents above $\zeta$, for $q$ and $g$ jets simulated and clustered with different methods, are shown in the right panel of Fig. 1 (See[1] for more details).

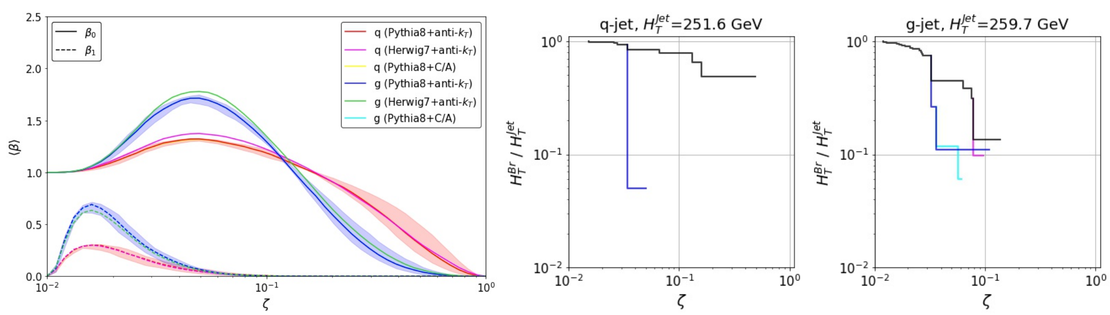


Figure 1: Left: Averaged Betti numbers of the superlevel sets above $\zeta$, for the light-quark and gluon jets ($100\text{GeV} < p_T < 350\text{GeV}$). Right: Branch phylogenetic trees for the two typical light-quark (w/ two branches) and gluon (w/ four branches) jets.

Morse theory[6] plays another important role when studying topological structures of jets. It states in our context that the topology of $G(\zeta)$ changes only if $\zeta$ passes some vertex of $G_{\rm ref}$ or jet constituent. This allows the birth, growth and death of each topological feature to be persistently recorded by evaluating the impact of every $G_{\rm ref}$ vertex passed by $\zeta$, as $\zeta$ varies, and hence makes sense of its evolution. Particularly, with the persistent knowledge of $\beta_0$, we will be able to build the branch phylogenetic tree for each jet. We show the ones in the right panel Fig. 1 for the two typical light-quark and gluon jets.

# 2 Proposal

In this letter of interest we propose several important directions for next-step explorations of jet topological strucutres.

- Build topological $q$- and $g$-jet classifiers: To develop such jet classifiers, we need to properly synergize the observables in this topological scheme (including the information carried by the branch phylogenetic trees), and maybe some complementary others known to us before. This could be achieved by using the deep neuron network (DNN) techniques. Additionally, one advantage of persistent topology is that the persistence length can be applied to strengthening the robustness of the relevant topological observables against topological noise. This point has been extensively appreciated by people working on topological computation[3;7]. In our context, we can apply, e.g., the relative lifetime of a jet branch, as a trimming tool to remove the jet branches with a short lifetime. This will allow us to focus on the more robust topological features in each jet, and may benefit to the classifier construction by suppressing topological contaminations from environment (pileups, underlying events, detector noise, etc.).

- Topological taggers for boosted heavy jets: Previously we focused on light QCD jets. But, it is a natural thinking to extend the use of topological taggers to boosted heavy jets such as $W^{\pm}$, $Z$, Higgs and top jets, given their important role in searching for new physics at Large Hadron Collider and even future hadron collider.

- Develop more complete jet morphology by properly including geometric elements in this topological observable scheme: Recall, in integral geometry the shape of a $D$-dimensional geometric object is characterized with $D + 1$ quantifiers, named Minkowski functionals[8]. If this object is smooth and closed, one of these quantifiers is reduced to Euler-characteristic, via Gauss-Bonnet theorem, an alternating sum of Betti numbers, while the other ones are purely geometric. This indicates that topological variants are not complete in characterizing jet structure and more complete jet morphology could be developed by incorporating geometric Minkowski functionals. Given the role of Minkowski functionals in analyzing morphology of the Cosmic Microwave Background (CMB) map[9–11], doing so will also enrich the dictionary between the collider observable scheme and the CMB one which was built recently[12].

- Apply the tool of persistent homology for the event-level data analysis at colliders. In the event-level analysis, the constituents in each event will be all projected to the detector sphere or cylinder. The persistent Betti numbers are expected to be applied to characterize topological structure of the whole event. One analogue with a reversed process is the generalization of the event shape N-jettiness[13] to the jet shape N-subjettiness[14]. Notably, this topological observable scheme is generically insusceptible to the boost of collision events along beam direction, because of its topological nature, and hence can be well-applied to both hadron and lepton colliders. This application can be also extended to the resonance search, where the color of new particles may result in distinguishable topological structures in their decay products. The persistent homology thus may serve as a new type of color-flow observables[15].

## References


[1] L. Li, T. Liu, and S.-J. Xu, "Jet Topology," 6 2020.

[2] H. Edelsbrunner, D. Letscher, and A. Zomorodian, "Topological persistence and simplification," in *Proceedings 41st annual symposium on foundations of computer science*, pp. 454–463, IEEE, 2000.

[3] A. Zomorodian and G. Carlsson, "Computing persistent homology," *Discrete & Computational Geometry*, vol. 33, no. 2, pp. 249–274, 2005.

[4] B. Delaunay, "Sur la sphere vide," *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk*, vol. 7, pp. 793–800, 1934.

[5] E. Betti, "Sopra gli spazi di un numero qualunque di dimensioni," 1870.

[6] M. Morse, "Relations between the critical points of a real function of n independent variables," *Transactions of the American Mathematical Society*, vol. 27, no. 3, pp. 345–396, 1925.

[7] H. Carr, J. Snoeyink, and U. Axen, "Computing contour trees in all dimensions," *Computational Geometry*, vol. 24, no. 2, pp. 75 – 94, 2003. Special Issue on the Fourth CGC Workshop on Computational Geometry.

[8] H. Minkowski, "Volumen und oberfläche," *Mathematische Annalen*, vol. 57, pp. 447–495, 1903.

[9] K. Mecke, T. Buchert, and H. Wagner, "Robust morphological measures for large scale structure in the universe," *Astron. Astrophys.*, vol. 288, pp. 697–704, 1994.

[10] S. Winitzki and A. Kosowsky, "Minkowski functional description of microwave background Gaussianity," *New Astron.*, vol. 3, p. 75, 1998.

[11] J. Schmalzing and K. M. Gorski, "Minkowski functionals used in the morphological analysis of cosmic microwave background anisotropy maps," *Mon. Not. Roy. Astron. Soc.*, vol. 297, p. 355, 1998.

[12] L. Li, Y.-Y. Li, T. Liu, and S.-J. Xu, "Learning Physics at Future $e^-e^+$ Colliders with Machine," 4 2020.

[13] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, "N-Jettiness: An Inclusive Event Shape to Veto Jets," *Phys. Rev. Lett.*, vol. 105, p. 092002, 2010.

[14] J. Thaler and K. Van Tilburg, "Identifying Boosted Objects with N-subjettiness," *JHEP*, vol. 03, p. 015, 2011.

[15] J. Gallicchio and M. D. Schwartz, "Seeing in Color: Jet Superstructure," *Phys. Rev. Lett.*, vol. 105, p. 022001, 2010.