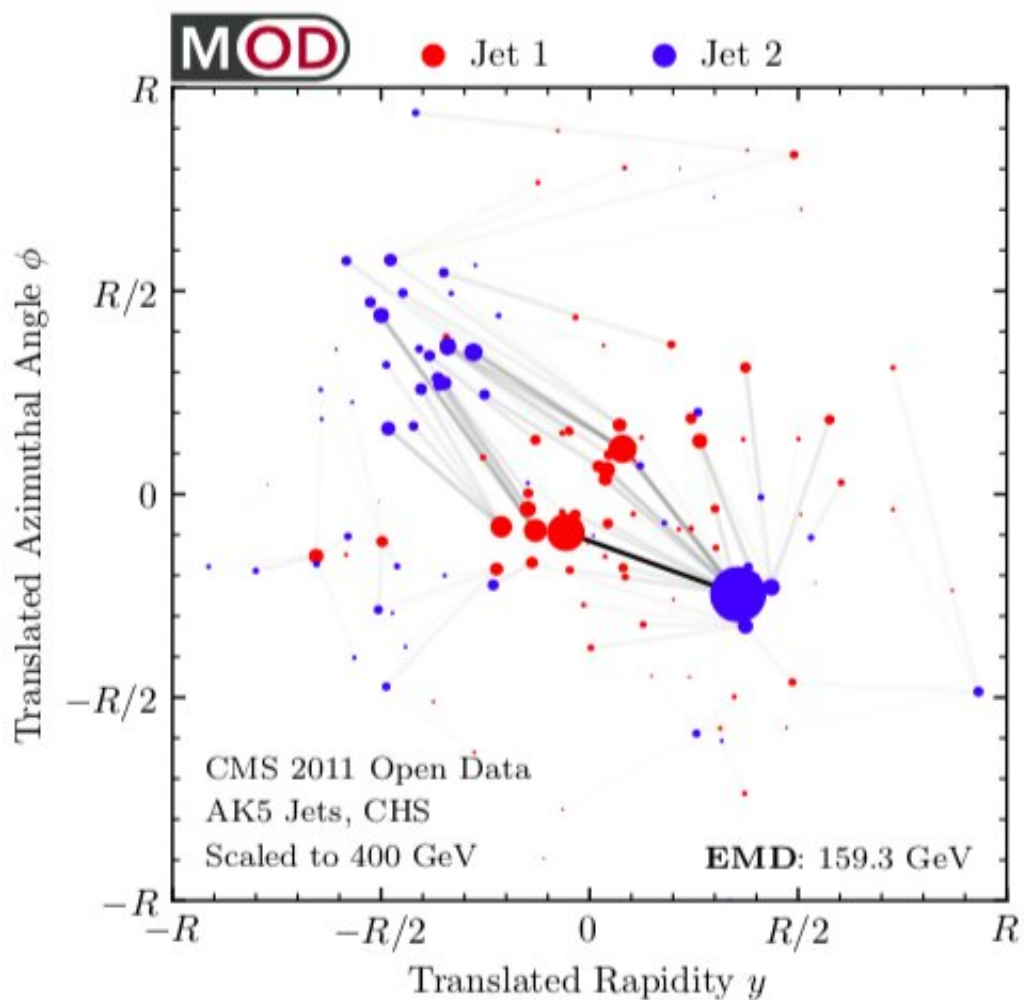


A new metric to capture the similarity between collider events

August 9 2019, by Ingrid Fadelli



A static illustration of the optimal transport between two jets from the CMS Open Data. Credit: Komiske, Metodiev & Thaler.

Researchers at the Massachusetts Institute of Technology (MIT) have recently developed a metric that can be used to capture the space of collider events based on the earth mover's distance (EMD), a measure used to evaluate dissimilarity between two multi-dimensional probability distributions. The metric they proposed, outlined in a paper published in *Physical Review Letters*, could enable the development of new powerful tools to analyze and visualize collider data, which do not rely on a choice of observables.

"Our research is motivated by a remarkably simple question: When are two [particle collisions](#) similar?" Eric Metodiev, one of the researchers who carried out the study, told Phys.org. "At the Large Hadron Collider (LHC), protons are smashed together at extremely high energies and each collision produces a complex mosaic of particles. Two collider events can look similar, even if they consist of different numbers and types of particles. This is analogous to how two mosaics can look similar, even if they are made up of different numbers and colors of tiles."

In their study, Metodiev and his colleagues set out to capture the similarity between collider events in a way that is conceptually useful for particle physics. To do this, they employed a strategy that merges ideas related to optimal transport theory, which is often used to develop cutting-edge image recognition tools, with insights from [quantum field theory](#), a construct that describes fundamental particle interactions.

"Our new result is a quantitative method to determine the distance (via a 'metric') between two collision events," Metodiev said. "Once you know the distance between every pair of collider events, you can then triangulate the entire space of LHC data. We hope this way of processing information from the LHC will yield new insights into the fundamental interactions of nature."

Essentially, the metric developed by the researchers represents the 'work'

required to rearrange one collider event into another. It is based on the EMD, a method that is typically used to develop computer vision tools that compare the similarities between two objects or images.

EDM works by trying to rearrange one event into another by moving "dirt," or in this case particle energies, around. Typically, the more work is required to successfully perform this rearrangement, the more two events, objects or images are dissimilar.

"The reason why this notion of similarity is so useful in particle physics is that it aligns with the way that we perform theoretical calculations," Patrick Komiske, another researcher involved in the study, told Phys.org. "In quantum field theory, you cannot predict exactly what will happen in any particular collision event, but you can predict the probability to produce certain patterns of particle debris. To define what you mean by a pattern, though, you need a notion of similarity, which turns out to be exactly what our metric provides."

In their paper, Metodiev, Komiske and their colleague Jesse Thaler specifically applied their metric to jets; sprays of particles that commonly arise from high-energy quarks and gluons. While the properties of individual jets have been studied extensively for over the past four decades, their metric allowed the researchers to study the relationship between pairs of jets, thus unveiling new and complementary information about the jet formation process.

"Having a universal notion of similarity between events is very useful for a variety of collider tasks," Metodiev said. "One common task at the LHC is to classify different types of collisions, in the same way as you would classify an image as containing a cat, dog, or unicorn. Using our metric to classify jets as arising from a quark, gluon, or something more exotic, we achieve a performance that approaches that of modern machine learning techniques."

In a series of evaluations, the researchers demonstrated the effectiveness of their method in capturing the similarity of collider events. Their technique achieved remarkable results, with accuracy levels comparable to those attained by state-of-the-art machine learning models.

In addition to potentially helping researchers to classify collider events, the metric developed by Metodiev and his colleagues could be used to visualize collider data in an entirely new way. Traditionally, in [particle physics](#), researchers focus on a single attribute of a collection of collider events (i.e. the 'forest') or at the detailed properties of one individual collider event (i.e. the 'trees'). Since the new metric allows users to group similar collider events together, it enables the observation of the 'forest' and individual 'trees' simultaneously, by identifying events that best capture the main features of the dataset as a whole.

"In addition, from a more mathematical perspective, once you have a notion of distance, you can study the geometry of the space of events, which provides a new way to think about existing concepts in collider physics dating back to the 1970s," Metodiev added. "For example, to avoid infinities in quantum field theory calculations, one simply has to ensure that the event geometry is sufficiently smooth, without any singular points. In the future, we plan to develop new collider observables and techniques based on this geometric perspective."

The metric developed by Metodiev, Komiske and Thaler could have numerous interesting applications. It could even be used to search for irregularities in LHC datasets using a strategy known as anomaly detection, which could ultimately help to unearth new physics evidence.

In the short term, the researchers plan to use their metric to rediscover known aspects of the standard model in the new geometric language they proposed. Ultimately, however, their technique could unveil evidence of the existence of new particles or forces, as well as previously unknown

aspects of the standard model itself.

"With our notion of similarity, we can identify not only the most common event configurations, but also the most exotic ones, and it possible that these exotic events could provide hints for physics beyond the standard model," Thaler told Phys.org. "We are currently working on benchmarking this idea with public collider data. Since 2014, the CMS experiment at the LHC has been releasing subsets of their data for unrestricted use, including all of the information necessary to calculate our metric. This gives us an opportunity to explore the space of events on real [collider](#) data."

More information: Patrick T. Komiske et al. Metric Space of Collider Events, *Physical Review Letters* (2019). [DOI: 10.1103/PhysRevLett.123.041801](#)
journals.aps.org/prl/abstract/..._ysRevLett.123.041801

© 2019 Science X Network

Citation: A new metric to capture the similarity between collider events (2019, August 9)
retrieved 19 April 2024 from
<https://phys.org/news/2019-08-metric-capture-similarity-collider-events.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--