Seeking new physics, scientists borrow from social networks

July 26 2019, by Jennifer Chu

MIT physicists find a way to relate hundreds of thousands of particle collisions, similar to a social network. Credit: Chelsea Turner, MIT

When two protons collide, they release pyrotechnic jets of particles, the details of which can tell scientists something about the nature of physics

and the fundamental forces that govern the universe.

Enormous particle accelerators such as the Large Hadron Collider can generate billions of such collisions per minute by smashing together beams of protons at close to the speed of light. Scientists then search through measurements of these collisions in hopes of unearthing weird, unpredictable behavior beyond the established playbook of physics known as the Standard Model.

Now MIT physicists have found a way to automate the search for strange and potentially new physics, with a technique that determines the degree of similarity between pairs of collision events. In this way, they can estimate the relationships among hundreds of thousands of collisions in a proton beam smashup, and create a geometric map of events according to their degree of similarity.

The researchers say their new technique is the first to relate multitudes of particle collisions to each other, similar to a social network.

"Maps of social networks are based on the degree of connectivity between people, and for example, how many neighbors you need before you get from one friend to another," says Jesse Thaler, associate professor of physics at MIT. "It's the same idea here."

Thaler says this social networking of particle collisions can give researchers a sense of the more connected, and therefore more typical, events that occur when protons collide. They can also quickly spot the dissimilar events, on the outskirts of a collision network, which they can further investigate for potentially new physics. He and his collaborators, graduate students Patrick Komiske and Eric Metodiev, carried out the research at the MIT Center for Theoretical Physics and the MIT Laboratory for Nuclear Science. They detail their new technique this week in the journal *Physical Review Letters*.

Seeing the data agnostically

Thaler's group focuses, in part, on developing techniques to analyze open data from the LHC and other particle collider facilities in hopes of digging up interesting physics that others might have initially missed.

"Having access to this public data has been wonderful," Thaler says. "But it's daunting to sift through this mountain of data to figure out what's going on."

Physicists normally look through collider data for specific patterns or energies of collisions that they believe to be of interest based on theoretical predictions. Such was the case for the discovery of the Higgs boson, the elusive elementary particle that was predicted by the Standard Model. The particle's properties were theoretically outlined in detail but had not been observed until 2012, when physicists, knowing approximately what to look for, found signatures of the Higgs boson hidden amid trillions of proton collisions.

But what if particles exhibit behavior beyond what the Standard Model predicts, that physicists have no theory to anticipate?
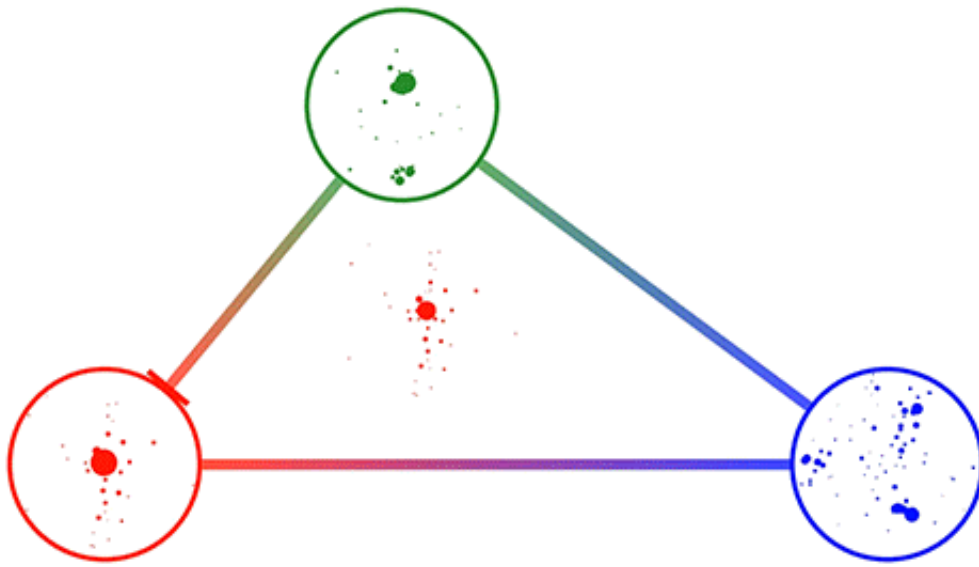
Three particle collision events, in the form of jets, obtained from the CMS Open Data, form a triangle to represent an abstract "space of events." The animation depicts how one jet can be optimally rearranged into another. Credit: Massachusetts Institute of Technology

Thaler, Komiske, and Metodiev have landed on a novel way to sift through collider data without knowing ahead of time what to look for. Rather than consider a single collision event at a time, they looked for ways to compare multiple events with each other, with the idea that perhaps by determining which events are more typical and which are less so, they might pick out outliers with potentially interesting, unexpected behavior.

"What we're trying to do is to be agnostic about what we think is new physics or not," says Metodiev. "We want to let the data speak for itself."

Moving dirt

Particle collider data are jam-packed with billions of proton collisions, each of which comprises individual sprays of particles. The team realized these sprays are essentially [point clouds](#)—collections of dots, similar to the point clouds that represent scenes and objects in computer vision. Researchers in that field have developed an arsenal of techniques to compare point clouds, for example to enable robots to accurately identify objects and obstacles in their environment.

Metodiev and Komiske utilized similar techniques to compare point clouds between pairs of collisions in particle collider data. In particular, they adapted an existing algorithm that is designed to calculate the optimal amount of energy, or "work" that is needed to transform one point cloud into another. The crux of the algorithm is based on an abstract idea known as the "earth's mover's distance."

"You can imagine deposits of energy as being dirt, and you're the earth mover who has to move that dirt from one place to another," Thaler explains. "The amount of sweat that you expend getting from one configuration to another is the notion of distance that we're calculating."

In other words, the more energy it takes to rearrange one point cloud to resemble another, the farther apart they are in terms of their similarity. Applying this idea to particle collider data, the team was able to calculate the optimal energy it would take to transform a given point cloud into another, one pair at a time. For each pair, they assigned a number, based on the "distance," or degree of similarity they calculated between the two. They then considered each point cloud as a single point and arranged these points in a social network of sorts.

The team has been able to construct a social network of 100,000 pairs of collision events, from open data provided by the LHC, using their technique. The researchers hope that by looking at collision datasets as networks, scientists may be able to quickly flag potentially interesting

events at the edges of a given network.

"We'd like to have an Instagram page for all the craziest events, or point clouds, recorded by the LHC on a given day," says Komiske. "This technique is an ideal way to determine that image. Because you just find the thing that's farthest away from everything else."

Typical collider datasets that are made publicly available normally include several million events, which have been preselected from an original chaos of billions of collisions that occurred at any given moment in a particle accelerator. Thaler says the team is working on ways to scale up their technique to construct larger networks, to potentially visualize the "shape," or general relationships within an entire dataset of particle collisions.

In the near future, he envisions testing the technique on historical data that physicists now know contain milestone discoveries, such as the first detection in 1995 of the top quark, the most massive of all known elementary particles.

"The top quark is an object that gives rise to these funny, three-pronged sprays of radiation, which are very dissimilar from typical sprays of one or two prongs," Thaler says. "If we could rediscover the top quark in this archival data, with this technique that doesn't need to know what new physics it is looking for, it would be very exciting and could give us confidence in applying this to current datasets, to find more exotic objects."

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology