

# LHC students face data drought

Computer simulations are the only option when the world's largest particle accelerator isn't working.

Last November, Sara Bolognesi stood before a committee at the University of Turin in Italy and defended her PhD thesis in experimental high-energy physics. The 180-page document is a treatise on finding the Higgs boson, part of the mechanism believed to endow all other matter with mass. The pages are crammed with dozens of figures and tables, but something is missing: real data.

That's because the Large Hadron Collider (LHC), the world's largest particle accelerator at CERN, outside Geneva in Switzerland, is broken. The 4.6-billion Swiss franc (US\$4.3-billion) collider is designed to accelerate protons to near the speed of light and smash them together in four giant detectors spread around its 27-kilometre circumference. Physicists once hoped that the LHC would start its collisions in late 2006, but last September, after a series of delays and soon after the machine was switched on, an electrical short caused extensive damage along a sector of the machine. Repairs have taken longer than expected, and, as of last week, the LHC was not scheduled to restart before mid-November.

The long delays have ended the dreams of a generation of graduate students hoping to use fresh data for their theses. With no machine to deliver results, "people are doing experimental PhDs and effectively doing very little experimenting", says Will Reece, a graduate student at Imperial College London working on a detector known as LHCb. "It's a strange situation."

Strange but not unprecedented, says Rolf-Dieter Heuer, CERN's director-general. During the mid-1980s, physicists were focused on building the Large Electron-Positron collider, the predecessor to the LHC. Over that period, Heuer says, graduate students sometimes wrote theses based on data from detector tests. Today, many of the same physicists work on the LHC project.

But although the electron collider took a few years to build, construction of the LHC took more than a decade, and most testing for the current detectors ended years ago. Aside from a trickle of data created by stray cosmic rays hitting the detectors, there will be no data to be analysed until the collider restarts. "It's a mess," says Burton Richter, a Nobel-prizewinning physicist at the SLAC National Accelerator Laboratory in Menlo Park, California.

European graduate students such as



Sara Bolognesi hopes to move from simulated experiments to real data collection at CERN.

Bolognesi face strict time constraints for completing their PhDs. Most universities require a thesis to be submitted within three to four years, and that means that students cannot wait for their data. Instead, their analyses are being done with data from 'Monte Carlo' simulations — computer programs that replicate what might come out of real collisions.

Not everybody thinks that the simulated data are a problem. "I don't feel that bad about not having data in my thesis," says Carsten Hof, a graduate student at Aachen University in Germany, who is finishing his PhD on software that will automatically analyse real collisions. "All the bugs we found and fixed now will also be fixed for real data." Hof adds that the data drought may actually be an advantage. "You look at everything with no bias," he says.

Heuer says that the situation reflects the growing size and sophistication of high-energy physics experiments. Whereas early experiments could be done in days by a handful of people, machines such as the LHC take thousands of researchers years to complete. The current generation of students may not be familiar with real data, he says, but they have extensive experience in building the huge detectors needed to capture them. Future PhD students will work on software without touching the innards of the detectors, he points out. As long as students get a taste of what's involved with each stage of the project, he says, "I don't think that people are losing anything."

Others are more worried. Although Monte

Carlo simulations can reproduce the uncertainties seen in real data, they will never contain a big surprise. That means simulated data can never be as good as the real thing, says Gustaaf Brooijmans, a physicist at Columbia University in New York. "It's like a badly written murder mystery," he says. "In the first chapter you're given enough information that you know who did it, and then you read the rest of the book, and, lo and behold, you get the right answer."

For this reason, Columbia and other US institutions require students to use real data in their PhD theses. That solves the data dilemma, but creates a new problem: US students working on the LHC must move to finish their theses. For students such as Ketino Kaadze of Kansas State University in Manhattan, this meant travelling from Geneva to Batavia, Illinois, the home of the world's other major particle collider, the Tevatron.

Kaadze says that she was initially nervous about the move from one experiment to the other, but she has found it valuable. Although it will take her longer to complete her PhD, she is glad to have made the switch. "I think it's very important to have this experience," she says.

Now at CERN for a postdoctoral fellowship, Bolognesi worries that she will be at a disadvantage compared with students like Kaadze. "Two years from now, I will have to search for work," she says. "I hope they will not discriminate against me." By the time she is looking for a job, the LHC should have completed its first run, and Bolognesi will hopefully have completed a first of her own — an analysis of real collisions. ■

Geoff Brumfiel