

Institut Pascal

Learning to Discover

19th-29th April 2022



Generator workshop introduction

See also Denis Ullmo Institut Pascal director
[presentation](#) in the agenda Tuesday 20th morning

Program



- 19-20 : Representation Workshop
- 21-22 : Dealing with uncertainties workshop
- 25-26 : Generators workshop
- 27-29 : final conference

- Each workshop : ~55 on-site participants + ~75 remote
- Final conference : 110 on-site + 80 remote

Monday



09:00 → 12:30

Generative Models workshop: Monday morning



09:00

Generator Workshop introduction

15m



09:15

Overview of generator model for detector simulation

1h 25m



Detailed simulations of physics processes are a cornerstone of most physics measurements and searches at experiments such as those at the Large Hadron Collider (LHC) as well as in other scientific domains. However, the large volumes of data collected at the LHC precludes a requirement for even more simulated data in order to test various hypotheses to a high precision. As the LHC moves to the era of high luminosity, the large scale production of simulated particle physics collisions becomes even more important.

One of the bottlenecks in the generation of these data is the time required to simulate the response of the detector to particles created in collisions. This is in particular the case for calorimeter systems, which capture the energy of showers of particles.

The current state of the art simulations are performed using the Geant 4 toolkit, which provides a detailed simulation of the interaction of each individual particle with the detector material. This is a very time and CPU intensive process which requires a detailed description of the detector. For example, simulations of particle physics collision events in the ATLAS detector can take several minutes per event.

Fast simulation methods have been used by experiments for many years to reduce the required CPU time, with several non machine learning approaches actively used by the experiments at the LHC. Although these fast simulation methods often perform well, they are still far from the levels of the detailed simulation and efforts to improve them are ongoing. With the advent of deep generative modelling interest in bringing modern techniques to solve this challenge has grown, with the aim to improve upon the current fast simulation approaches. Deep generative modelling has shown great success in other domains, and the hope is to provide a simulation almost as accurate and varied as Geant 4, but without the required CPU time at inference. In this talk an outline of the problem and its challenges will be presented, followed by an overview of several recent approaches and an outlook on future developments.

Speaker: Johnny Raine (Université de Genève)

10:40

Generator model for 4-momenta events

20m



Event generation and Hadronization make up a significant fraction of our high energy physics simulation chain. Therefore there exists a strong interest in using generative models to supplement these simulation tasks. In both applications the data is commonly expressed as a list of 4-momenta. This presents an additional challenge to any generative model, as 4-momenta have inherent complex correlations that need to be correctly reproduced.

Several generative models for simulating 4-momenta are presented and discussed, with special focus on how the problem of complex correlations is addressed.

Speaker: Sascha Diefenbacher (Universität Hamburg)

11:00

Coffee Break

30m

11:30

Break out sessions

1h



12:30 → 14:00

Welcome buffet

1h 30m



14:00 → 17:30

Generative Models workshop: Monday afternoon



14:00

GNNs for generating molecules and PDE solutions

1h 30m



Graph Neural Networks (GNNs) have proven to be a versatile tool to predict properties of molecules, generative molecules and even predict solutions of a partial differential equation (PDE). Many physical application domains also exhibit symmetries which can be incorporated into the GNNs through equivariant convolutions or data augmentation. In this talk I will explain how this tool can be leveraged to generate molecules from their equilibrium distribution, possibly conditioned on some properties, and even predict solutions of a PDE.

Speaker: Max Welling (U. Amsterdam / MSR)

15:30

Coffee break

30m

16:00

Breakout sessions

1h 30m

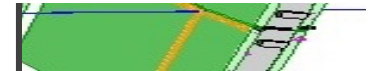


17:30 → 18:30

Happy hour

1h

Cathedral (IPA)



Tuesday

09:00 → 12:30 **Generative Models workshop: Tuesday morning**



09:00

The calorimeter challenge and ODD detector

🕒 1h 30m



There has been many recent developments in the application of machine-learning techniques to fast simulation of cascades in the calorimeters. This is usually the most time consuming part of the event simulation in high energy physics experiments. Most current efforts are focused and fine tuned to specific detectors, which makes it difficult to compare. We present a first fast calorimeter simulation challenge, with a three-diffculty level dataset. The purpose of this challenge is to spur the development and benchmarking of fast and high-fidelity calorimeter shower generation using deep learning methods. It will be possible to directly compare new deep learning approaches on common benchmarks. It is expected that participants will make use of cutting-edge techniques in generative modeling with deep learning, e.g. GANs, VAEs and normalizing flows.

As a follow up to this challenge, we would like to implement a benchmark calorimeter, within the Open Data Detector (ODD). Tracking system is already implemented in the ODD, and is an evolution of the successful detector from the Tracking Machine Learning Challenges. The calorimeter system will allow not only to benchmark the fast simulation of particle cascades, but creates unique opportunities for development and comparison of different reconstruction techniques, and particle flow methods.

Speaker: Anna Zaborowska (CERN)

10:30

Coffee break

🕒 30m

11:00

Breakout sessions

🕒 1h 30m



12:30 → 14:00 **Lunch at the cafeteria**

🕒 1h 30m



14:00 → 17:30 **Generative Models workshop: Tuesday afternoon**



14:00

Breakout sessions

🕒 1h 30m



15:30

Coffee break

🕒 30m

16:00

Conclusion from breakout sessions

🕒 1h 30m



Social Events



☐ Mon 25th

- Welcome buffet in la Cathedrale
- 17h40 Happy hour (=beer) in la Cathedrale
- Dinner on your own

☐ Tues 26th

- Lunch at the cafeteria
- Dinner on your own

☐ Thursday 28th

- 6:30PM visit and dinner at Musée d'Orsay in Paris
- You should already have confirmed.

Covid



- ☐ No official mask mandate any more in France except in medical premises
- ☐ However, we would all like to attend on-site till the last day
- ☐ Mask wearing recommended at all time
- ☐ Use outside areas (terrace and rooftop) when possible
- ☐ CO₂ monitored in the amphitheaters

Spirit



- ❑ Long talks, do not hesitate to ask questions during the talk:
 - On zoom raise your hand
- ❑ Long coffee breaks
- ❑ Break-out sessions
- ❑ Un-organised spontaneous discussions

3rd of a series



- 15-26 Jul 2019: Real time analysis workshop
- 14-25 Oct 2019: Advanced Pattern recognition
- Learning to discover : ~~Jul 2020, Feb 2021, Feb 2022~~, **April 2022!**

Premices



- ❑ The whole Institut Pascal (half-building) is yours
- ❑ Two amphitheatres the big 120 one and the small 70 one (we use the big for plenary for social distancing)
- ❑ 40 offices (« Learning to Discover ») pick yours (possibly shared).
 - You can write your name on the sheet outside
- ❑ Small meeting rooms, free coffee machines...
- ❑ If you need to come early morning (before 8:45) or late afternoon (after 5:30) or in the week-end, please ask Sabrina an access card

slack



- ❑ Primary channel of communication
- ❑ 191 participants as of this morning. Make sure you're connected (ask your neighbour)

The screenshot displays the Slack interface. On the left is a dark purple sidebar with the 'LEARNING' section expanded, showing a list of channels. The '# sci_representation' channel is selected and highlighted in blue. The main area shows the channel's history. At the top, it says '# sci_representation' with a dropdown arrow and a header bar showing 10 members. Below this, there are pinned items: '1 élément(s) épinglé(s)' and a button to 'Ajouter un marque-page'. The chat history is divided by dates: 'Jeudi 7 avril' and 'Vendredi 8 avril'. On Thursday, David Rousseau joined and defined the channel description. On Friday, Savannah Thais joined with 7 others. A pinned message from David Rousseau states: 'Live notes will be on this [google doc](#) (please do not advertise). Anyone here can edit or comment.' Below this, a message from Slackbot says: 'D'accord ! Je vous le proposerai de nouveau la prochaine fois.' At the bottom, a message from Francois Lanusse says: 'a rejoint #sci_representation.' The bottom of the screen shows the message input area with a toolbar containing icons for bold, italic, strikethrough, link, unlink, list, list, code, and a send button. The text 'Envoyer un message #sci_representation' is visible above the input field.

LEARNING

- Fils de discussion
- Tous les messages dire...
- Mentions et réactions
- Éléments enregistrés
- Slack Connect
- Plus
- Favoris
- sabrina.soccard
- Canaux
 - # general
 - # organisation
 - # outings
 - # random
 - # sci_conference
 - # sci_generators
 - # sci_representation**
 - # sci_uncertainties
 - # transports_and_regio...
 - + Ajouter des canaux
- Messages directs

sci_representation 10

1 élément(s) épinglé(s) + Ajouter un marque-page

Jeudi 7 avril

David Rousseau 23 h 35
a rejoint #sci_representation.

David Rousseau 23 h 35
a défini la description du canal : Discussion pertaining to representation workshop

Vendredi 8 avril

Savannah Thais 0 h 20
a rejoint #sci_representation, avec 7 autres personnes.

Hier

★ Épinglé par vous

David Rousseau 23 h 18
Live notes will be on this [google doc](#) (please do not advertise). Anyone here can edit or comment.

👁 Visible par vous seulement

Slackbot 23 h 18
D'accord ! Je vous le proposerai de nouveau la prochaine fois.

Francois Lanusse 23 h 35
a rejoint #sci_representation.

B I S | | | | </>

Envoyer un message #sci_representation

11 A, Generator Learning, David Rousseau, introduction

Zoom



- ❑ Zoom in meeting mode (no waiting room, everyone sees every one else connection)
 - Link, meeting ID and code distributed on slack
 - Please do not advertise beyond workshop and conference
- ❑ Anyone of you can start zoom in amphi and meeting rooms : « join » room number + code
 - Do not hesitate to exercise it yourself
- ❑ 6 break out room created. Anyone can navigate between these rooms and the main.
- ❑ Zoom chat (volatile) only to be used for connection issues
- ❑ Use slack for scientific exchanges : #sci_uncertainties for this workshop (other specific channel can be created on request)

Live Document



- ❑ Google doc link posted in slack
#sci_generators (please do not advertise
beyond workshop attendees)
- ❑ Live notes
- ❑ Break-out sessions proposed themes
- ❑ Anyone can edit / comment. Feel free to
add info. Propose break-out sessions.

Break-out sessions



- ☐ Semi auto organised
- ☐ Theme added in google doc, you can add your own and vote
- ☐ Org committee will assign physical and zoom room
- ☐ You take it from there
- ☐ (sketchy) report Tuesday end of afternoon

(not) Facebook



- ❑ List of participants on indico lists who is here or remote for which workshop and conference
- ❑ Google sheet with one page per participant shared on slack
 - Please duplicate first (dummy) page and briefly introduce yourself, recent work, interest

Workshop summary at the final conference



- Each workshop will be summarised in 30' at the final conference
- Johnny Raine kindly volunteered for this one
 - He might ask for your inputs

Outcome



- ❑ Work started or continued from collaboration enriched in these workshops may be published in special edition of Computing and Software for the Big Science.
- ❑ If interested, talk to one of the organiser

Organising committee



- ❑ Sabrina Soccard, Program Manager at Institut Pascal
- ❑ Peter Battaglia, senior scientist at Google Deepmind
- ❑ Anja Butter, ITP Heidelberg : generator models and uncertainties in ML for Particle Physics
- ❑ Cécile Germain : Emeritus professor (computer science) Université Paris-Saclay, LRI-CNRS, and INRIA TAU team, organiser of the 2014 HiggsML and 2018-2019 TrackML challenge
- ❑ Tobias Golling, associate professor at Université de Genève, generative models and anomaly detection for particle physics
- ❑ Vladimir Vava Gligorov, CNRS/IN2P3 LPNHE, leader of the LHCb Real Time analysis project, organiser of the Real Time Analysis Institut Pascal Workshop
- ❑ Eilam Gross, Weizmann Institute, organiser of the Hammers and Nails 2017 and 2019 workshop
- ❑ Michael Kagan, SLAC, co-coordinator of CERN Interexperiment Machine Learning group
- ❑ Danilo Rezende, Researcher on Probabilistic Methods for Decision Making, Senior Staff Researcher and Team Lead at Google Deepmind
- ❑ David Rousseau, CNRS/IN2P3 IJCLab : former co-coordinator of ATLAS Machine Learning group, co-coordinator of the LHC Interexperiment Machine Learning Group, organiser of the 2014 Higgs Machine Learning challenge, and of the 2018-2019 TrackML challenge
- ❑ Andreas Salzburger, CERN, coordinator of the ATLAS software upgrade group, organiser of the 2018-2019 TrackML challenge and organiser of the Advanced Pattern Recognition Institut Pascal Workshop
- ❑ Savannah Thais, Princeton, representation learning for Particle Physics, AI and ethics
- ❑ Jean-Roch Vlimant, Caltech, co-coordinator of CMS Machine Learning group, organiser of the 2018-2019 TrackML challenge
- ❑ Slava Voloshynovskiy, head of Stochastic Information Processing group, Université de Genève

Sponsors



□ Thanks to our sponsors (see indico Menu)

- Institut Pascal
- Paris Saclay Center for Data Science
- TrackML
- DataIA
- CNRS/IN2P3
- Recept
- Dark Matter Lab