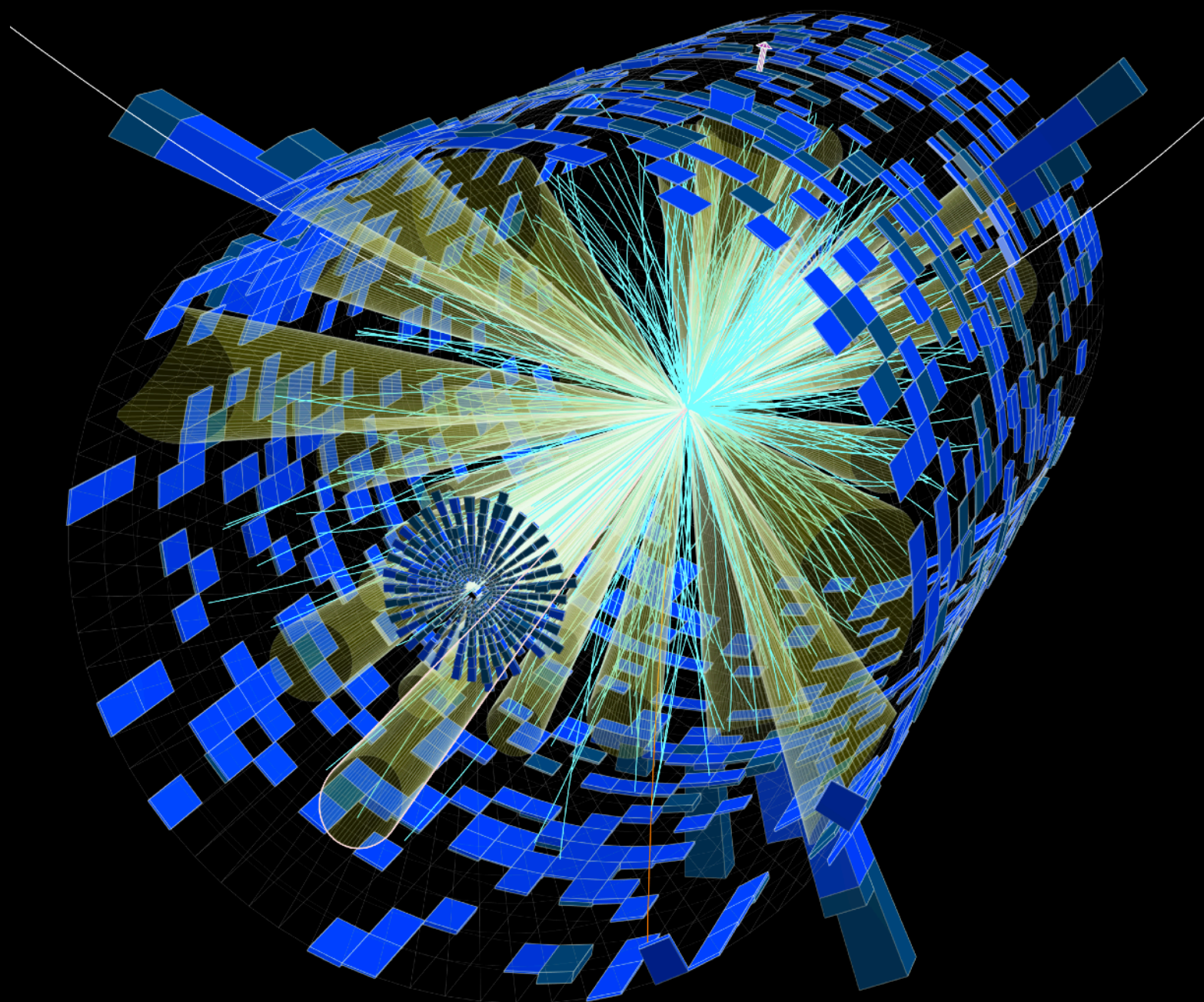




MACHINE LEARNING

PARTICLE / ASTRO / COSMO



@KyleCranmer

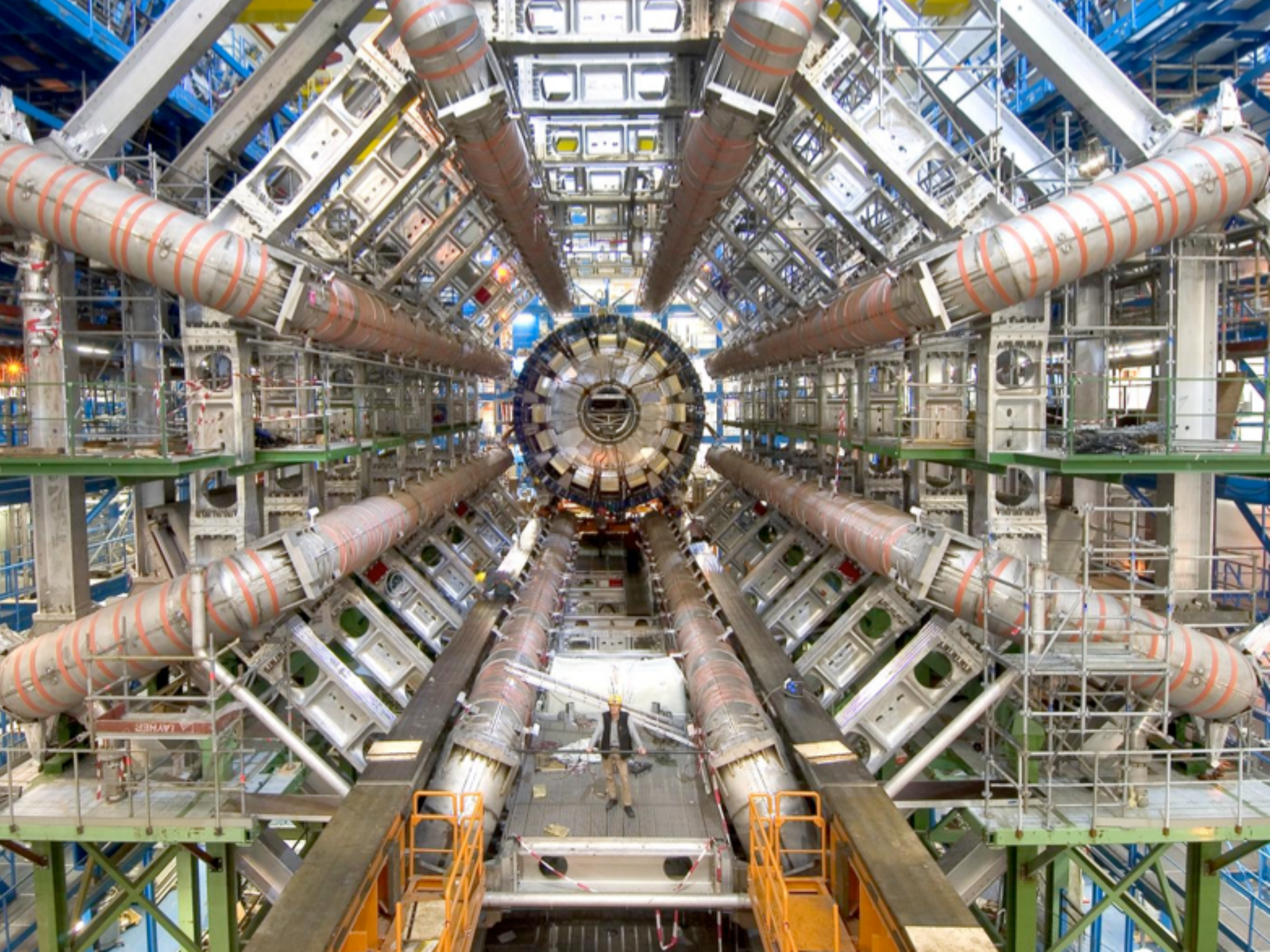
New York University
Department of Physics
Center for Data Science
CILVR Lab

GOALS

The goal of these workshops is to provide a setting to promote open and informal discussions and the exchange of information needed to **help assess the promise and challenges of an emerging field.**

Targeted Attendees

- experts on machine learning techniques and applications in interdisciplinary studies, including physics, biology, computer science, mathematics, etc.
- Physical Review editorial staff from across the journals



I CARE ABOUT PUBLISHING, DATA, CODE, & CYBERINFRASTRUCTURE



Welcome to **INSPIRE**, the High Energy Physics information system. [Find out more](#)

HEP :: [HEPNAMES](#) :: [INSTITUTIONS](#) :: [CONFERENCES](#) :: [JOBS](#) :: [EXPERIMENTS](#)

HEP Search

High-Energy Physics Literature Database

Use "find" for SPIRES-style search ([other tips](#))

[Brief format](#) [Search](#) [Easy Search](#) [Advanced Search](#)

[find j "Phys.Rev.Lett..105" :: more](#)

HOW TO SEARCH

SPIRES syntax is (mostly) supported (requires "find")

- [find a richter, b and t quark and date > 1984](#)
- [find j phys.rev.,D50,1140 or j jhep.0903,112](#)
- [find eprint arxiv:1007.5048](#) (Note the plots available on the detailed record)
- [find fulltext "quark-gluon plasma"](#) (Note new "fulltext" operator)
- [find a ellis and refersto a witten](#) (Note "refersto")
- [find a kane and citedby title SUSY and topcite 200+](#) (Note "citedby")

New techniques:

- [1985 richter quark multiplicity](#)
- [arXiv:1007.5048](#)
- [citedby:author:ellis -refersto:author:witten](#)
- [author:randall | author:sundrum cited:450->1350](#)

Additional Help:

- [More search tips and full help](#)

INSPIRE UPDATES

See our blog at blog.inspirehep.net for updates on new features and other news. You can also follow us at [@inspirehep](https://twitter.com/inspirehep) on twitter. To send us feedback use feedback@inspirehep.net. The data in INSPIRE is updated daily. To request corrections to data in INSPIRE, write us at help@inspirehep.net. INSPIRE superseded SPIRES in 2012.



Repository for publication-related High-Energy Physics data

This new site replaces the old site at <http://hepdata.cedar.ac.uk>.

Search on 8716 publications and 76099 data tables.

Search for a paper, author, experiment, reaction [Search](#) [Advanced](#)

e.g. reaction $P\bar{P} \rightarrow LQ\bar{L}Q^*X$, title has "photon collisions", collaboration is LHCf or D0.

Data from the LHC

Experiment	View Data
 ATLAS	View Data
 ALICE	View Data
 CMS	View Data
 LHCb	View Data



Reproducible research data analysis platform

Flexible	Scalable	Reusable	Free
Run many computational workflow engines.	Support for remote compute clouds.	Containerise once, reuse elsewhere. Cloud-native.	Free Software. GPL licence. Made with ❤ at CERN.
		 	



Video Guides [GitHub Help](#) [GitHub.com](#)

Making Your Code Citable

© 10 minute read

Digital Object Identifiers (DOI) are the backbone of the academic reference and metrics system. If you're a researcher writing software, this guide will show you how to make the work you share on GitHub citable by archiving one of your GitHub repositories and assigning a DOI with the data archiving tool [Zenodo](#).

ProTip: This tutorial is aimed at researchers who want to cite GitHub repositories in academic literature. Provided you've already set up a GitHub repository, this tutorial can be completed without installing any special software. If you haven't yet created a project on GitHub, start first by [uploading your work](#) to a repository.

Intro

- [Choosing Your Repo](#)
- [Login to Zenodo](#)
- [Check Repo Settings](#)
- [Create a New Release](#)
- [Minting a DOI](#)
- [Finishing up](#)

☐ Software (21242)

The Promise of an Emerging Field

JULY 2012



ImageNet Classification with Deep Convolutional Neural Networks

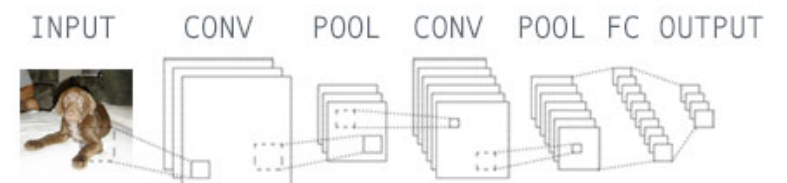
Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

Abstract

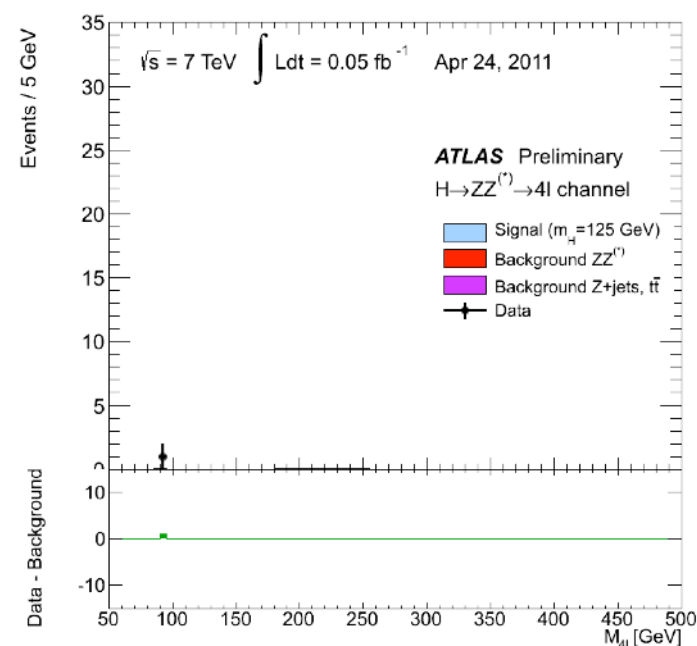
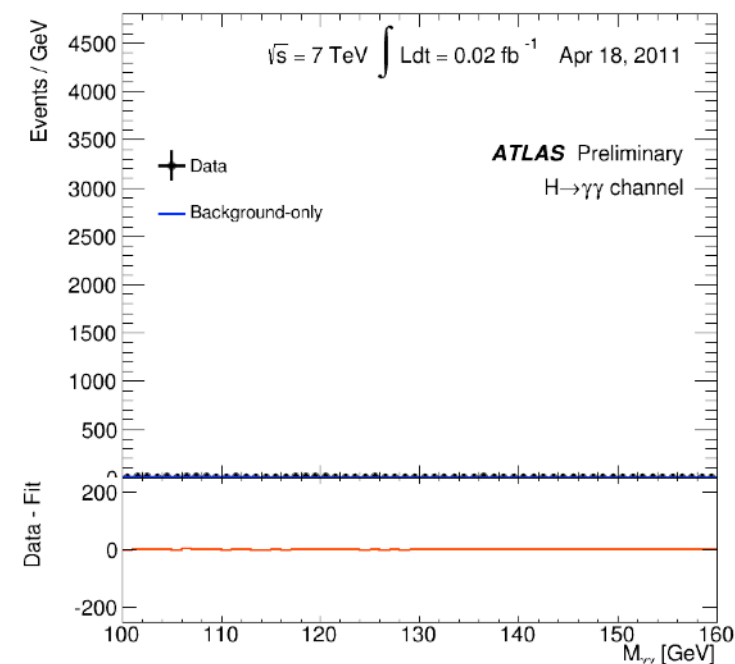
We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called “dropout” that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.



Dog:	94%
Cat:	31%
Bird:	2%
Boat:	0%



Dog:	37%
Cat:	91%
Bird:	21%
Boat:	1%



HEP ML WORKSHOPS

- [2nd IML Machine Learning Workshop \(2018\)](#) (April 9-12, 2018)
- [Machine Learning for Phenomenology \(2018\)](#) (April 3-6, 2018)
- [4th International Connecting The Dots Workshop \(2018\)](#) (March 20-22, 2018)
- [Machine Learning for Jet Physics](#) (December 11-13, 2017)
- [Deep Learning for Physical Sciences at NIPS](#) (December 8, 2017)
- [18th International Workshop on Advanced Computing and Analysis Techniques in Physics Research \(ACAT 2017\)](#) (August 21-25, 2017)
- [Hammers & Nails - Machine Learning & HEP](#) (July 19-28, 2017)
- [CMS Machine Learning Workshop \(2017\)](#) (July 5-6, 2017) restricted CMS
- [ATLAS Machine Learning Workshop \(2017\)](#) (June 6-9, 2017) restricted ATLAS
- [Workshop on Machine Learning and b-tagging](#) (May 23-26, 2017) restricted ATLAS
- [DS@HEP 2017](#) (May 8-12, 2017)
- [2nd S2I2 HEP/CS Workshop \(Parallel Session\)](#) (May 1-3, 2017)
- [CERN openlab workshop on Machine Learning and Data Analytics](#) (April 27, 2017)
- [First IML Workshop on Machine Learning](#) (March 20-22, 2017)
- [DS@HEP at the Simons Foundation](#) (July 5-7, 2016)
- [ALICE Mini-Workshop 2016: Statistical Methods and Machine Learning Tutorial](#) (May 18, 2016) restricted ALICE
- [ATLAS Machine Learning Workshop \(2016\)](#) (March 29-31, 2016) restricted ATLAS
- [Heavy Flavour Data Mining workshop](#) (February 18-20, 2016)
- [Data Science @ LHC 2015](#) (November 9-13, 2015)

HEPML PAPERS

- M. Stoye, J. Brehmer, L. Gilles, J. Paez, and K. Cranmer, "[Likelihood-free inference with an improved cross-entropy estimator](#)," arXiv:1808.00973 [stat.ML]. (August 2, 2018)
- D. Bourgeois, C. Fitzpatrick, and S. Stahl, "[Using holistic event information in the Trigger](#)," arXiv:1808.00711 [physics.ins-det]. (August 2, 2018)
- M. Andrews, M. Paulini, S. Gleyzer, and B. Poczos, "[End-to-End Physics Event Classification with the CMS Open Data: Applying Image-based Deep Learning on Detector Data to Directly Classify Collision Events at the LHC](#)," arXiv:1807.11916 [hep-ex]. (July 31, 2018)
- J. Lin, M. Freytsis, I. Moutl, and B. Nachman, "[Boosting \$H \rightarrow b\bar{b}\$ with Machine Learning](#)," arXiv:1807.10768 [hep-ph]. (July 27, 2018)
- K. Albertsson *et al.*, "[Machine Learning in High Energy Physics Community White Paper](#)," arXiv:1807.02876 [physics.comp-ph]. (July 8, 2018)
- D. Guest, K. Cranmer, and D. Whiteson, "[Deep Learning and its Application to LHC Physics](#)," arXiv:1806.11484 [hep-ex]. (June 29, 2018)
- J. Brehmer, K. Cranmer, G. Louppe, and J. Pavez, "[Constraining Effective Field Theories with Machine Learning](#)," arXiv:1805.00013 [hep-ph]. (April 30, 2018)
- J. Brehmer, K. Cranmer, G. Louppe, and J. Pavez, "[A Guide to Constraining Effective Field Theories with Machine Learning](#)," arXiv:1805.00020 [hep-ph]. (April 30, 2018)
- CMS Collaboration, A. M. Sirunyan *et al.*, "[Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV](#)," arXiv:1712.07158 [physics.ins-det]. (December 19, 2017)
- V. Estrade, C. Germain, I. Guyon and D. Rousseau, "[Adversarial learning to eliminate systematic errors: a case study in High Energy Physics](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- D. Weitekamp III, T. Nguyen, D. Anderson, R. Castello, M. Pierini, M. Spiropulu and J. Vlimant, "[Deep topology classifiers for a more efficient trigger selection at the LHC](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- L. Hertel, L. Li, P. Baldi and J. Bian, "[Convolutional Neural Networks for Electron Neutrino and Electron Shower Energy Reconstruction in the NOvA Detectors](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- M. Stoye, J. Kieseler, M. Verzetti, H. Qu, L. Gouskos, A. Stakia and the CMS Collaboration, "[DeepJet: Generic physics object based jet multiclass classification for LHC experiments](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- B. Hooberman, A. Farbin, G. Khatkhat, V. Pacela, M. Pierini, J. Vlimant, M. Spiropulu, W. Wei, M. Zhang and S. Vallecorsa, "[Calorimetry with Deep Learning: Particle Classification, Energy Regression, and Simulation for High-Energy Physics](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- M. Paganini, L. de Oliveira and B. Nachman, "[Survey of Machine Learning Techniques for High Energy Electromagnetic Shower Classification](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- L. de Oliveira, M. Paganini, and B. Nachman, "[Tips and Tricks for Training GANs with Physics Constraints](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- S. Farrell, P. Calafiura, M. Mudigonda, Prabhat, D. Anderson, J. Bendavid, M. Spiropoulou, J. Vlimant, S. Zheng, G. Cerati, L. Gray, J. Kowalkowski, P. Spentzouris, A. Tsaris and D. Zurawski, "[Particle Track Reconstruction with Deep Learning](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- I. Henrion, K. Cranmer, J. Bruna, K. Cho, J. Brehmer, G. Louppe and G. Rochette, "[Neural Message Passing for Jet Physics](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- T. Cheng, "[Recursive Neural Networks in Quark/Gluon Tagging](#)," arXiv:1711.02633 [hep-ph]. (November 7, 2017)
- S. Chang, T. Cohen, and B. Ostidek, "[What is the Machine Learning?](#)," arXiv:1709.10106 [hep-ph]. (September 28, 2017)
- M. Frate, K. Cranmer, S. Kalia, A. Vandenberg-Rodes, and D. Whiteson, "[Modeling Smooth Backgrounds and Generic Localized Signals with Gaussian Processes](#)," arXiv:1709.05681 [physics.data-an]. (September 17, 2017)
- E. M. Metodiev, B. Nachman, and J. Thaler, "[Classification without labels: Learning from mixed samples in high energy physics](#)," arXiv:1708.02949 [hep-ph]. (August 9, 2017)
- J. Bendavid, "[Efficient Monte Carlo Integration Using Boosted Decision Trees and Generative Deep Neural Networks](#)," arXiv:1707.00028 [hep-ph]. (June 30, 2017)
- T. Cohen, M. Freytsis, and B. Ostidek, "[\(Machine\) Learning to Do More with Less](#)," arXiv:1706.09451 [hep-ph]. (June 28, 2017)
- M. Paganini, L. de Oliveira, and B. Nachman, "[CaloGAN: Simulating 3D High Energy Particle Showers in Multi-Layer Electromagnetic Calorimeters with Generative Adversarial Networks](#)," arXiv:1705.02355 [hep-ex]. (May 5, 2017)
- C. Shimmmin, P. Sadowski, P. Baldi, E. Weik, D. Whiteson, E. Goul, and A. Sgaard, "[Decorrelated Jet Substructure Tagging using Adversarial Neural Networks](#)," arXiv:1703.03507 [hep-ex]. (March 9, 2017)
- G. Louppe, K. Cho, C. Becot, and K. Cranmer, "[QCD-Aware Recursive Neural Networks for Jet Physics](#)," arXiv:1702.00748 [hep-ph]. (February 2, 2017)
 - Lecture: "[QCD-Aware Neural Networks for Jet Physics](#)," by [Kyle Cranmer](#)
- L. M. Dery, B. Nachman, F. Rubbo, and A. Schwartzman, "[Weakly Supervised Classification in High Energy Physics](#)," JHEP 05 (2017) 145, arXiv:1702.00414 [hep-ph]. (February 1, 2017)
- L. de Oliveira, M. Paganini, and B. Nachman, "[Learning Particle Physics by Example: Location-Aware Generative Adversarial Networks for Physics Synthesis](#)," arXiv:1701.05927 [stat.ML]. (January 20, 2017)
- L.-G. Pang, K. Zhou, N. Su, H. Petersen, H. Stocker, X.-N. Wang, "[An equation-of-state-meter of QCD transition from deep learning](#)," arXiv:1612.04262 [hep-ph]. (December 13, 2016)
- P. T. Komiske, E. M. Metodiev, and M. D. Schwartz, "[Deep learning in color: towards automated quark/gluon jet discrimination](#)," JHEP 01 (2017) 110, arXiv:1612.01551 [hep-ph]. (December 5, 2016)
- MicroBooNE Collaboration, R. Acciarri *et al.*, "[Convolutional Neural Networks Applied to Neutrino Events in a Liquid Argon Time Projection Chamber](#)," JINST 12 (2017) no. 03, P03011, arXiv:1611.05531 [physics.ins-det]. (November 16, 2016)
- M. Kagan, L. d. Oliveira, L. Mackey, B. Nachman, and A. Schwartzman, "[Boosted Jet Tagging with Jet-Images and Deep Neural Networks](#)," EPJ Web Conf. 127 (2016) 00009. (November 15, 2016)
- G. Bertone, M. P. Deisenroth, J. S. Kim, S. Liem, R. Ruiz de Austri, and M. Welling, "[Accelerating the BSM interpretation of LHC data with machine learning](#)," arXiv:1611.02704 [hep-ph]. (November 8, 2016)
- G. Louppe, M. Kagan, and K. Cranmer, "[Learning to Pivot with Adversarial Networks](#)," arXiv:1611.01046 [stat.ME]. (November 3, 2016)
 - Lecture: "[\(Machine\) Learning to Pivot with Adversarial Networks](#)," by [Gilles Louppe](#)
- J. Barnard, E. N. Dawe, M. J. Dolan, and N. Rajcic, "[Parton Shower Uncertainties in Jet Substructure Analyses with Deep Neural Networks](#)," Phys. Rev. D95 (2017) no. 1, 014018, arXiv:1609.00607 [hep-ph] (September 2, 2016)
- A. Rogozhnikov, "[Reweighting with Boosted Decision Trees](#)," J. Phys. Conf. Ser. 762 (2016) no. 1, 012036, arXiv:1608.05806 [physics.data-an]. (August 20, 2016)
- S. Caron, J.S. Kim, K. Rolbiecki, R. Ruiz de Austri, B. Stienen "[The BSM-AI project: SUSY-AI -- Generalizing LHC limits on supersymmetry with machine learning](#)," EPJ C (2017) 77:257, arXiv:1605.02797 [hep-ph]. (May 9, 2016)
- A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M. D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, and P. Vahle, "[A Convolutional Neural Network Neutrino Event Classifier](#)," JINST 11 (2016) no. 09, P09001, arXiv:1604.01444 [hep-ex]. (April 5, 2016)
- P. Baldi, K. Cranmer, T. Faucett, P. Sadowski, and D. Whiteson, "[Parameterized neural networks for high-energy physics](#)," Eur. Phys. J. C76 (2016) no. 5, 235, arXiv:1601.07913 [hep-ex]. (January 28, 2016)
- L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, "[Jet-images deep learning edition](#)," JHEP 07 (2016) 069, arXiv:1511.05190 [hep-ph]. (November 16, 2015)
- A. Rogozhnikov, A. Bukva, V. Gligorov, A. Ustyuzhanin, M. Williams, "[New approaches for boosting to uniformity](#)," arXiv:1410.4140 [hep-ex]. (October 15, 2014)
- P. Baldi, P. Sadowski, and D. Whiteson, "[Searching for Exotic Particles in High-Energy Physics with Deep Learning](#)," Nature Commun. 5 (2014) 4308, arXiv:1402.4735 [hep-ph]. (February 19, 2014)
- J. Stevens, M. Williams, "[uBoost: A boosting method for producing uniform selection efficiencies from multivariate classifiers](#)," arXiv:1305.7248 [nucl-ex]. (May 30, 2013)
- V. Gligorov, M. Williams, "[Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree](#)," arXiv:1210.6861 [physics]. (October 25, 2012)
- B. H. Denby, "[Neural Networks and Cellular Automata in Experimental High-energy Physics](#)," Comput. Phys. Commun. 49 (1988) 429–448. (September 20, 1987)

HEPML PAPERS

- M. Stoye, J. Brehmer, L. Gilles, J. Paez, and K. Cranmer, "[Likelihood-free inference with an improved cross-entropy estimator](#)," arXiv:1808.00973 [stat.ML]. (August 2, 2018)
- D. Bourgeois, C. Fitzpatrick, and S. Stahl, "[Using holistic event information in the Trigger](#)," arXiv:1808.00711 [physics.ins-det]. (August 2, 2018)
- M. Andrews, M. Paulini, S. Gleyzer, and B. Poczos, "[End-to-End Physics Event Classification with the CMS Open Data: Applying Image-based Deep Learning on Detector Data to Directly Classify Collision Events at the LHC](#)," arXiv:1807.11916 [hep-ex]. (July 31, 2018)
- J. Lin, M. Freytsis, I. Moutl, and B. Nachman, "[Boosting \$H \rightarrow b\bar{b}\$ with Machine Learning](#)," arXiv:1807.10768 [hep-ph]. (July 27, 2018)
- K. Albertsson *et al.*, "[Machine Learning in High Energy Physics Community White Paper](#)," arXiv:1807.02876 [physics.comp-ph]. (July 8, 2018)
- D. Guest, K. Cranmer, and D. Whiteson, "[Deep Learning and its Application to LHC Physics](#)," arXiv:1806.11484 [hep-ex]. (June 29, 2018)
- J. Brehmer, K. Cranmer, G. Louppe, and J. Pavez, "[Constraining Effective Field Theories with Machine Learning](#)," arXiv:1805.00013 [hep-ph]. (April 30, 2018)
- J. Brehmer, K. Cranmer, G. Louppe, and J. Pavez, "[A Guide to Constraining Effective Field Theories with Machine Learning](#)," arXiv:1805.00020 [hep-ph]. (April 30, 2018)
- CMS Collaboration, A. M. Sirunyan *et al.*, "[Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV](#)," arXiv:1712.07158 [physics.ins-det]. (December 19, 2017)
- V. Estrade, C. Germain, I. Guyon and D. Rous, (December 8, 2017)
- D. Weitekamp III, T. Nguyen, D. Anderson, R. Workshop at NIPS (2017). (December 8, 2017)
- L. Hertel, L. Li, P. Baldi and J. Bian, [Convoluti](#) NIPS (2017). (December 8, 2017)
- M. Stoye, J. Kieseler, M. Verzetti, H. Qu, L. G. Sciences Workshop at NIPS (2017). (December 8, 2017)
- B. Hooberman, A. Farbin, G. Khattak, V. Pace [Physics](#), Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- M. Paganini, L. de Oliveira and B. Nachman, [S](#) (December 8, 2017)
- L. de Oliveira, M. Paganini, and B. Nachman, (December 8, 2017)
- S. Farrell, P. Calafiura, M. Mudigonda, Prabha [Learning](#), Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- I. Henrion, K. Cranmer, J. Bruna, K. Cho, J. B. (December 8, 2017)
- T. Cheng, "[Recursive Neural Networks in Quantum Physics](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- S. Chang, T. Cohen, and B. Ostidiek, "[What is Deep Learning for Physical Sciences?](#)" Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- M. Frate, K. Cranmer, S. Kalia, A. Vandenberg, (December 8, 2017)
- E. M. Metodiev, B. Nachman, and J. Thaler, "[QCD-Aware Neural Networks](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- J. Bendavid, "[Efficient Monte Carlo Integration](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- T. Cohen, M. Freytsis, and B. Ostidiek, "[\(Machine Learning\) for High-Energy Physics](#)," Proceedings of the Deep Learning for Physical Sciences Workshop at NIPS (2017). (December 8, 2017)
- M. Paganini, L. de Oliveira, and B. Nachman, (December 8, 2017)
- C. Shimmin, P. Sadowski, P. Baldi, E. Weik, D. (December 8, 2017)
- G. Louppe, K. Cho, C. Becot, and K. Cranmer, (December 8, 2017)
 - Lecture: [QCD-Aware Neural Networks](#)
- L. M. Dery, B. Nachman, F. Rubbo, and A. Sch, (December 8, 2017)
- L. de Oliveira, M. Paganini, and B. Nachman, (December 8, 2017)
- L.-G. Pang, K. Zhou, N. Su, H. Petersen, H. S, (December 8, 2017)
- P. T. Komiske, E. M. Metodiev, and M. D. Sch, (December 8, 2017)
- MicroBooNE Collaboration, R. Acciarri *et al.*, (December 8, 2017)
- M. Kagan, L. d. Oliveira, L. Mackey, B. Nachm, (December 8, 2017)
- G. Bertone, M. P. Deisenroth, J. S. Kim, S. Lie, (December 8, 2017)
- G. Louppe, M. Kagan, and K. Cranmer, "[Learn](#)" (December 8, 2017)
 - Lecture: [\(Machine\) Learning to Pivot](#)
- J. Barnard, E. N. Dawe, M. J. Dolan, and N. R, (December 8, 2017)
- A. Rogozhnikov, "[Reweighting with Boosted D](#)" (December 8, 2017)
- S. Caron, J.S. Kim, K. Rolbiecki, R. Ruiz de Austri, B. Stienen, "[The BSM-AI project: SUSY-AI -- Generalizing LHC limits on supersymmetry with machine learning](#)", EPJ C (2017) 77:257, arXiv:1605.02797 [hep-ph]. (May 9, 2016)
- A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M. D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, and P. Vahle, "[A Convolutional Neural Network Neutrino Event Classifier](#)," JINST 11 (2016) no. 09, P09001, arXiv:1604.01444 [hep-ex]. (April 5, 2016)
- P. Baldi, K. Cranmer, T. Faucett, P. Sadowski, and D. Whiteson, "[Parameterized neural networks for high-energy physics](#)," Eur. Phys. J. C76 (2016) no. 5, 235, arXiv:1601.07913 [hep-ex]. (January 28, 2016)
- L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, "[Jet-images deep learning edition](#)," JHEP 07 (2016) 069, arXiv:1511.05190 [hep-ph]. (November 16, 2015)
- A. Rogozhnikov, A. Bukva, V. Gligorov, A. Ustyuzhanin, M. Williams, "[New approaches for boosting to uniformity](#)," arXiv:1410.4140 [hep-ex]. (October 15, 2014)
- P. Baldi, P. Sadowski, and D. Whiteson, "[Searching for Exotic Particles in High-Energy Physics with Deep Learning](#)," Nature Commun. 5 (2014) 4308, arXiv:1402.4735 [hep-ph]. (February 19, 2014)
- J. Stevens, M. Williams, "[uBoost: A boosting method for producing uniform selection efficiencies from multivariate classifiers](#)," arXiv:1305.7248 [nucl-ex]. (May 30, 2013)
- V. Gligorov, M. Williams, "[Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree](#)," arXiv:1210.6861 [physics]. (October 25, 2012)
- B. H. Denby, "[Neural Networks and Cellular Automata in Experimental High-energy Physics](#)," Comput. Phys. Commun. 49 (1988) 429–448. (September 20, 1987)

Most of these papers are proof-of-concept papers by small teams with representative, but simplified setting.

Follow up by large collaborations in progress, but tends to be slow.

Disruptive to traditional norms of publishing, data sharing, credit, and collaboration membership.

Deep Learning and Its Application to LHC Physics

Dan Guest,¹ Kyle Cranmer,² and Daniel Whiteson¹

¹Department of Physics and Astronomy, University of California, Irvine, California 92697, USA

²Physics Department, New York University, New York, NY 10003, USA

Annu. Rev. Nucl. Part. Sci. 2018. 68:1–22

<https://doi.org/10.1146/annurev-nucl-101917-021019>

Copyright © 2018 by Annual Reviews.
All rights reserved

Keywords

deep learning, LHC, machine learning, particle physics

Abstract

Machine learning has played an important role in the analysis of high-energy physics data for decades. The emergence of deep learning in 2012 allowed for machine learning tools which could adeptly handle higher-dimensional and more complex problems than previously feasible. This review is aimed at the reader who is familiar with high energy physics but not machine learning. The connections between machine learning and high energy physics data analysis are explored, followed by an introduction to the core concepts of neural networks, examples of the key results demonstrating the power of deep learning for analysis of LHC data, and discussion of future prospects and concerns.

A Roadmap for HEP Software and Computing R&D for the 2020s

HEP Software Foundation¹

ABSTRACT: Particle physics has an ambitious and broad experimental programme for the coming decades. This programme requires large investments in detector hardware, either to build new facilities and experiments, or to upgrade existing ones. Similarly, it requires commensurate investment in the R&D of software to acquire, manage, process, and analyse the shear amounts of data to be recorded. In planning for the HL-LHC in particular, it is critical that all of the collaborating stakeholders agree on the software goals and priorities, and that the efforts complement each other. In this spirit, this white paper describes the R&D activities required to prepare for this software upgrade.

REVIEW

<https://doi.org/10.1038/s41586-018-0361-2>

Machine learning at the energy and intensity frontiers of particle physics

Alexander Radovic^{1*}, Mike Williams^{2*}, David Rousseau³, Michael Kagan⁴, Daniele Bonacorsi^{5,6}, Alexander Himmel⁷, Adam Aurisano⁸, Kazuhiro Terao⁴ & Taritree Wongjirad⁹

Our knowledge of the fundamental particles of nature and their interactions is summarized by the standard model of particle physics. Advancing our understanding in this field has required experiments that operate at ever higher energies and intensities, which produce extremely large and information-rich data samples. The use of machine-learning techniques is revolutionizing how we interpret these data samples, greatly increasing the discovery potential of present and future experiments. Here we summarize the challenges and opportunities that come with the use of machine learning at the frontiers of particle physics.

Machine Learning in High Energy Physics Community White Paper

July 10, 2018

Abstract: Machine learning is an important applied research area in particle physics, beginning with applications to high-level physics analysis in the 1990s and 2000s, followed by an explosion of applications in particle and event identification and reconstruction in the 2010s. In this document we discuss promising future research and development areas in machine learning in particle physics with a roadmap for their implementation, software and hardware resource requirements, collaborative initiatives with the data science community, academia and industry, and training the particle physics community in data science. The main objective of the document is to connect and motivate these areas of research and development with the physics drivers of the High-Luminosity Large Hadron Collider and future neutrino experiments and identify the resource needs for their implementation. Additionally we identify areas where collaboration with external communities will be of great benefit.

Editors: Sergei Gleyzer²⁶, Paul Seyfert¹¹, Steven Schramm²⁸

Contributors: Kim Albertsson¹, Piero Altoe², Dustin Anderson³, Michael Andrews⁴, Juan Pedro Araque Espinosa⁵, Adam Aurisano⁶, Laurent Basara⁷, Adrian Bevan⁸, Wahid Bhimji⁹, Daniele Bonacorsi¹⁰, Paolo Calafiura⁹, Mario Campanelli⁸, Louis Capps², Federico Carminati¹¹, Stefano Carrazza¹¹, Taylor Childers¹², Elias Coniavitis¹³, Kyle Cranmer¹⁴, Claire David¹⁵, Douglas Davis¹⁶, Javier Duarte¹⁷, Martin Erdmann¹⁸, Jonas Eschle¹⁹, Amir Farbin²⁰, Matthew Feickert²¹, Nuno Filipe Castro⁵, Conor Fitzpatrick²², Michele Floris¹¹, Alessandra Forti²³, Jordi Garra-Tico²⁴, Jochen Gemmler²⁵, Maria Girone¹¹, Paul Glayshe¹⁵, Sergei Gleyzer²⁶, Vladimir Gligorov²⁷, Tobias Golling²⁸, Jonas Graw², Lindsey Gray¹⁷, Dick Greenwood²⁹, Thomas Hacker³⁰, John Harvey¹¹, Benedikt Hegner¹¹, Lukas Heinrich¹⁴, Ben Hooberman³¹, Johannes Junggeburth³², Michael Kagan³³, Meghan Kane³⁴, Konstantin Kanishchev⁷, Przemysław Karpiński¹¹, Zahari Kassabov³⁵, Gaurav Kaul³⁶, Dorian Keira³, Thomas Keck²⁵, Alexei Klimentov³⁷, Jim Kowalkowski¹⁷, Luke Kreczko³⁸, Alexander Kurepin³⁹, Rob Kutschke¹⁷, Valentin Kuznetsov⁴⁰, Nicolas Köhler³², Igor Lakomov¹¹, Kevin Lannon⁴¹, Mario Lassnig¹¹, Antonio Limosani⁴², Gilles Louppe¹⁴, Aashrita Mangu⁴³, Pere Mato¹¹, Helge Meinhard¹¹, Dario Menasce⁴⁴, Lorenzo Moneta¹¹, Seth Moortgat⁴⁵, Meenakshi Narain⁴⁶, Mark Neubauer³¹, Harvey Newman³, Hans Pabst³⁶, Michela Paganini⁴⁷, Manfred Paulini⁴, Gabriel Perdue¹⁷, Uzziel Perez⁴⁸, Attilio Picazio⁴⁹, Jim Pivarski⁵⁰, Harrison Prosper⁵¹, Fernanda Psihas⁵², Alexander Radovic⁵³, Ryan Reece⁵⁴, Aurelius Rinkevicius⁴⁰, Eduardo Rodrigues⁶, Jamal Rorie⁵⁵, David Rousseau⁵⁶, Aaron Sauers¹⁷, Steven Schramm²⁸, Ariel Schwartzman³³, Horst Severini⁵⁷, Paul Seyfert¹¹, Filip Siroky⁵⁸, Konstantin Skazytkin³⁹, Mike Sokoloff⁶, Graeme Stewart⁵⁹, Bob Stienen⁶⁰, Ian Stockdale⁶¹, Giles Strong⁵, Savannah Thais⁴⁷, Karen Tomko⁶², Eli Upfal⁴⁶, Emanuele Usai⁴⁶, Andrey Ustyuzhanin⁶³, Martin Vala⁶⁴, Sofia Vallecorsa⁶⁵, Mauro Verzetti⁶⁶, Xavier Vilasís-Cardona⁶⁷, Jean-Roch Vlimant³, Ilija Vukotic⁶⁸, Sean-Jiun Wang²⁶, Gordon Watts⁶⁹, Michael Williams⁷⁰, Wenjing Wu⁷¹, Stefan Wunsch²⁵, Omar Zapata⁷²

Why should physicists care?



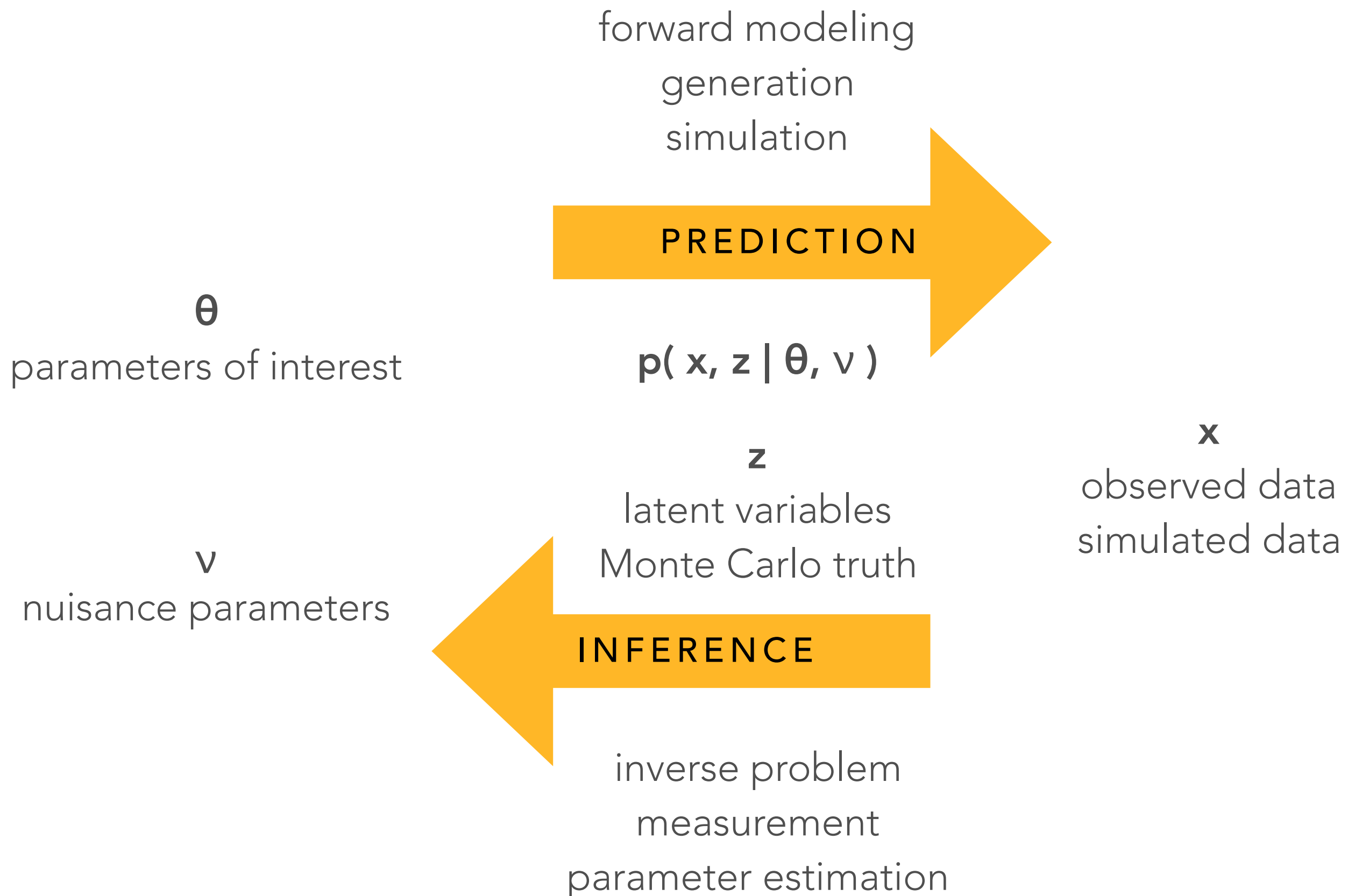
Deep Learning for Physical Sciences

Workshop at the 31st Conference on Neural Information Processing Systems (NIPS)

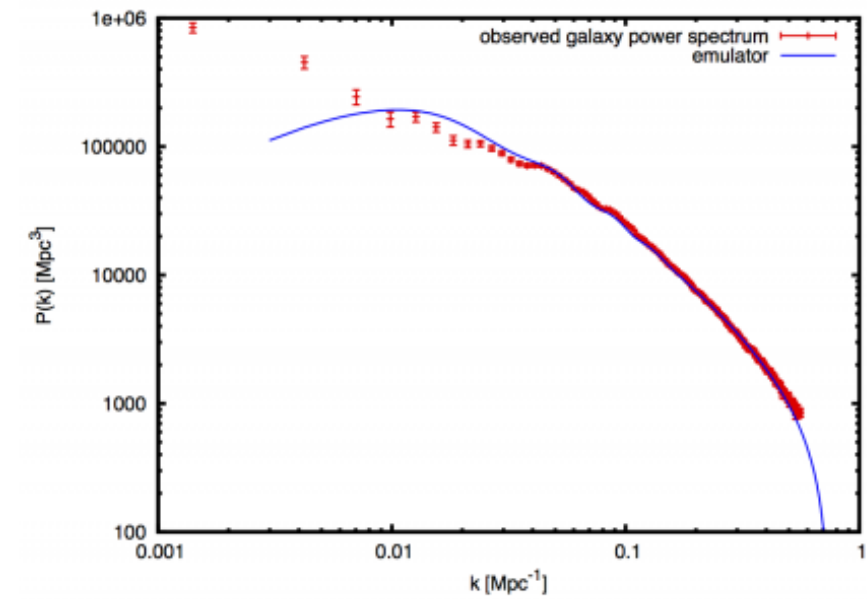
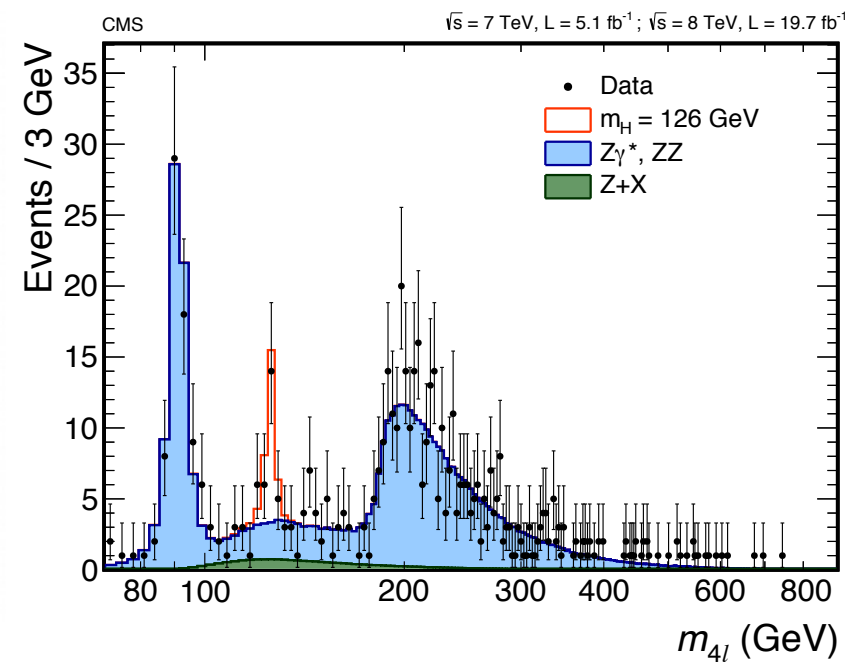
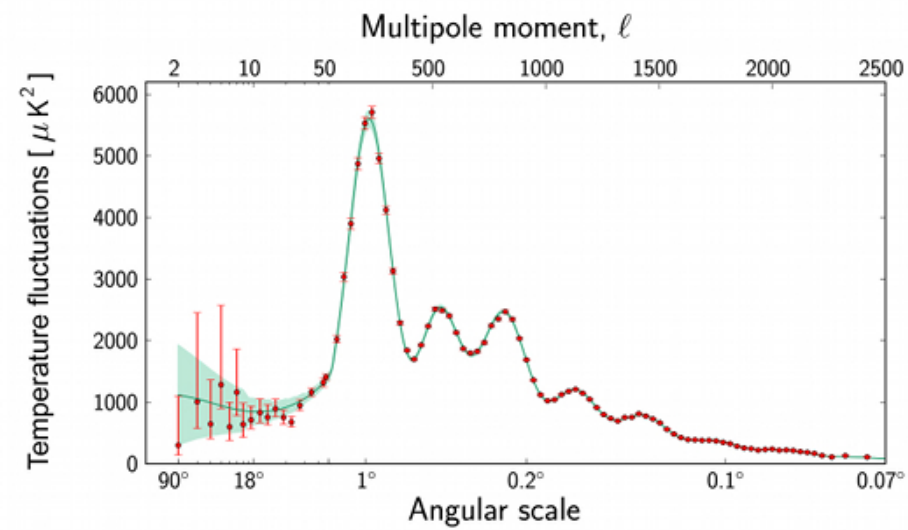
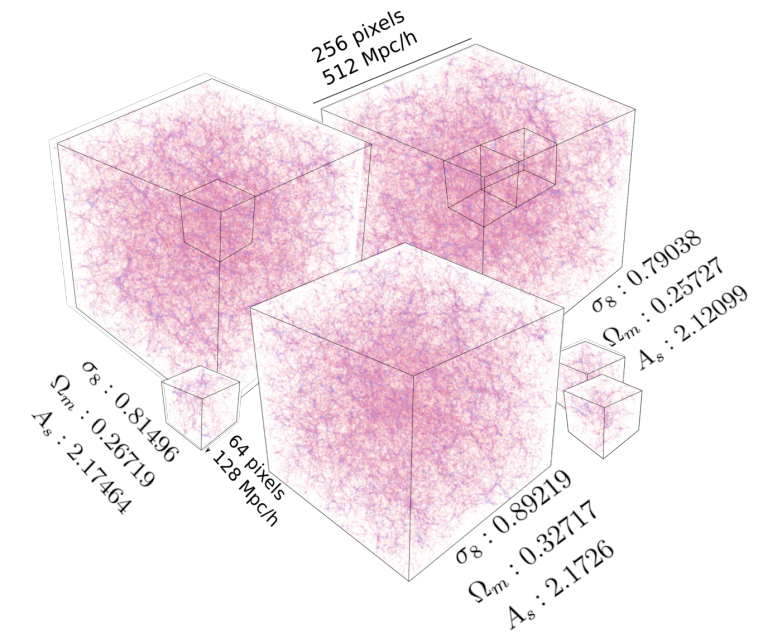
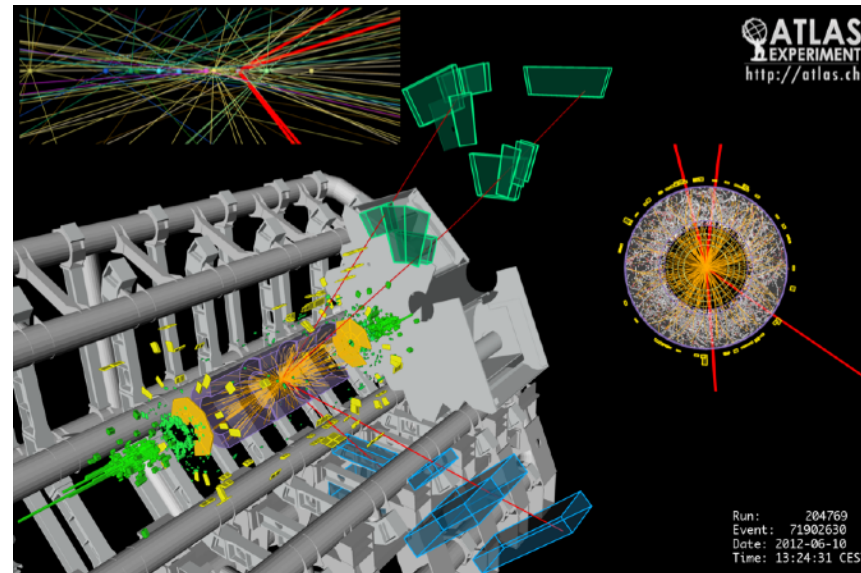
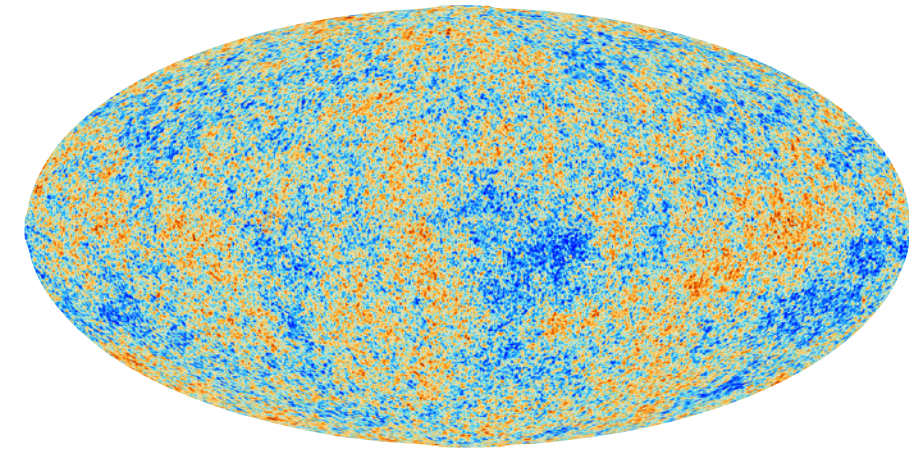
December 8, 2017

[About](#)[Schedule](#)[Call for papers](#)[Organizers](#)[Location](#)

THE PLAYERS



PREDICTION: THE FORWARD MODEL



PANEL DISCUSSION

1. There is a lot of low hanging fruit, we can use M.L. to

- improve what we normally do
- speed up accelerate what we normally do
- near-term impact: execution, application

2. More profound changes to how we approach physics

- new capabilities to be exploited
- attack previously intractable problems
- long-term transformation: exploratory, innovations

Simulators can produce labeled training data for supervised learning

(note: some simulation are very computationally expensive)

PREDICTION IN PARTICLE PHYSICS

1) We begin with Quantum Field Theory

$$\begin{aligned}
 \mathcal{L}_{SM} = & \underbrace{\frac{1}{4} \mathbf{W}_{\mu\nu} \cdot \mathbf{W}^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}}_{\text{kinetic energies and self-interactions of the gauge bosons}} \\
 & + \underbrace{\bar{L} \gamma^\mu (i \partial_\mu - \frac{1}{2} g \boldsymbol{\tau} \cdot \mathbf{W}_\mu - \frac{1}{2} g' Y B_\mu) L + \bar{R} \gamma^\mu (i \partial_\mu - \frac{1}{2} g' Y B_\mu) R}_{\text{kinetic energies and electroweak interactions of fermions}} \\
 & + \underbrace{\frac{1}{2} \left| (i \partial_\mu - \frac{1}{2} g \boldsymbol{\tau} \cdot \mathbf{W}_\mu - \frac{1}{2} g' Y B_\mu) \phi \right|^2 - V(\phi)}_{W^\pm, Z, \gamma, \text{ and Higgs masses and couplings}} \\
 & + \underbrace{g'' (\bar{q} \gamma^\mu T_a q) G_\mu^a}_{\text{interactions between quarks and gluons}} + \underbrace{(G_1 \bar{L} \phi R + G_2 \bar{L} \phi_c R + h.c.)}_{\text{fermion masses and couplings to Higgs}}
 \end{aligned}$$

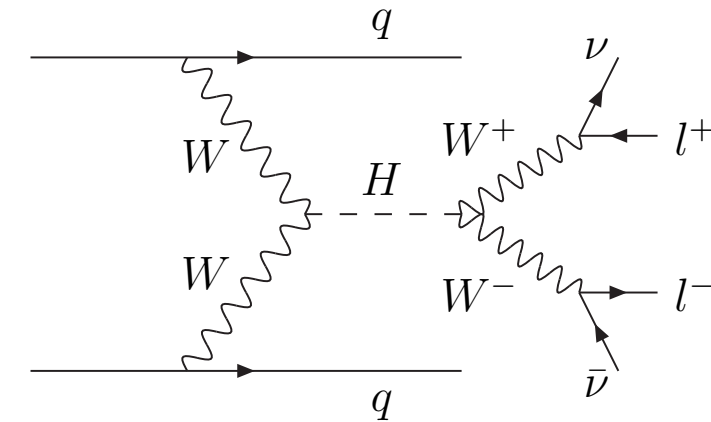
PREDICTION IN PARTICLE PHYSICS

$$\begin{aligned}
 \mathcal{L}_{SM} = & \underbrace{\frac{1}{4}\mathbf{W}_{\mu\nu} \cdot \mathbf{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu}}_{\text{kinetic energies and self-interactions of the gauge bosons}} \\
 & + \underbrace{\bar{L}\gamma^\mu(i\partial_\mu - \frac{1}{2}g\boldsymbol{\tau} \cdot \mathbf{W}_\mu - \frac{1}{2}g'Y B_\mu)L + \bar{R}\gamma^\mu(i\partial_\mu - \frac{1}{2}g'Y B_\mu)R}_{\text{kinetic energies and electroweak interactions of fermions}} \\
 & + \underbrace{\frac{1}{2}\left|(i\partial_\mu - \frac{1}{2}g\boldsymbol{\tau} \cdot \mathbf{W}_\mu - \frac{1}{2}g'Y B_\mu)\phi\right|^2 - V(\phi)}_{W^\pm, Z, \gamma, \text{ and Higgs masses and couplings}} \\
 & + \underbrace{g''(\bar{q}\gamma^\mu T_a q)G_\mu^a}_{\text{interactions between quarks and gluons}} + \underbrace{(G_1\bar{L}\phi R + G_2\bar{L}\phi_c R + h.c.)}_{\text{fermion masses and couplings to Higgs}}
 \end{aligned}$$

1) We begin with Quantum Field Theory

2) Theory gives detailed prediction for high-energy collisions

hierarchical: $2 \rightarrow \mathcal{O}(10) \rightarrow \mathcal{O}(100)$ particles



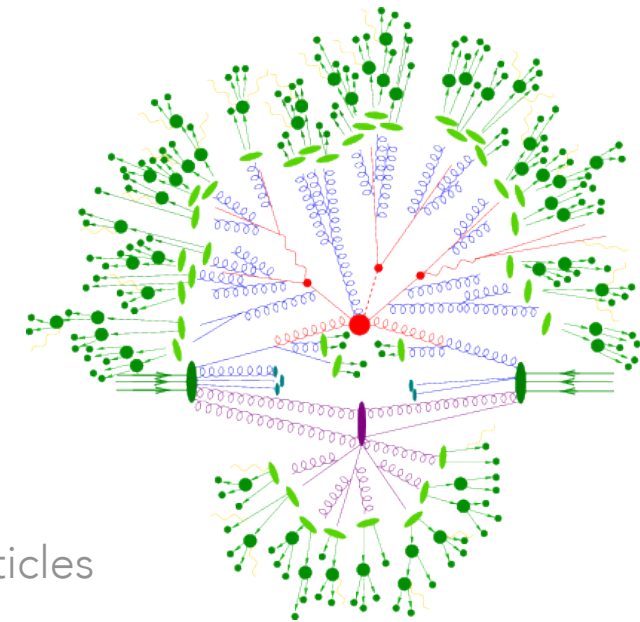
PREDICTION IN PARTICLE PHYSICS

$$\begin{aligned}
 \mathcal{L}_{SM} = & \underbrace{\frac{1}{4} \mathbf{W}_{\mu\nu} \cdot \mathbf{W}^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}}_{\text{kinetic energies and self-interactions of the gauge bosons}} \\
 & + \underbrace{\bar{L} \gamma^\mu (i \partial_\mu - \frac{1}{2} g \boldsymbol{\tau} \cdot \mathbf{W}_\mu - \frac{1}{2} g' Y B_\mu) L + \bar{R} \gamma^\mu (i \partial_\mu - \frac{1}{2} g' Y B_\mu) R}_{\text{kinetic energies and electroweak interactions of fermions}} \\
 & + \underbrace{\frac{1}{2} |(i \partial_\mu - \frac{1}{2} g \boldsymbol{\tau} \cdot \mathbf{W}_\mu - \frac{1}{2} g' Y B_\mu) \phi|^2 - V(\phi)}_{W^\pm, Z, \gamma, \text{ and Higgs masses and couplings}} \\
 & + \underbrace{g'' (\bar{q} \gamma^\mu T_a q) G_\mu^a}_{\text{interactions between quarks and gluons}} + \underbrace{(G_1 \bar{L} \phi R + G_2 \bar{L} \phi_c R + h.c.)}_{\text{fermion masses and couplings to Higgs}}
 \end{aligned}$$

1) We begin with Quantum Field Theory

2) Theory gives detailed prediction for high-energy collisions

hierarchical: $2 \rightarrow \text{O}(10) \rightarrow \text{O}(100)$ particles



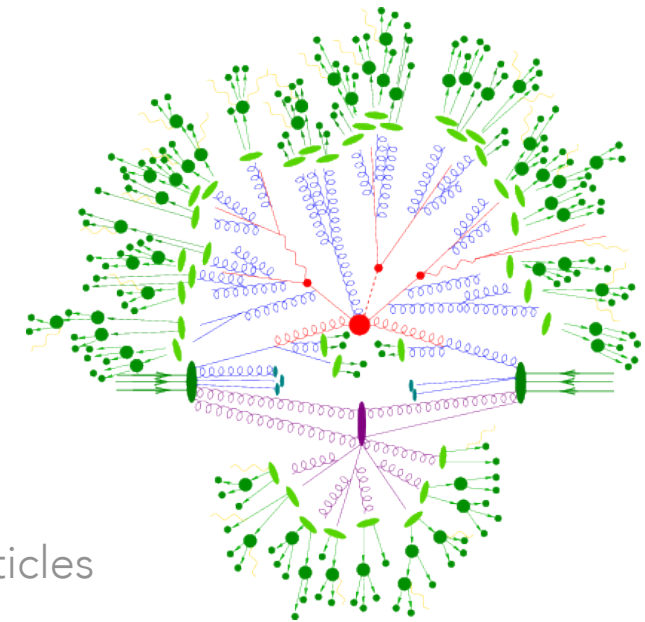
PREDICTION IN PARTICLE PHYSICS

$$\begin{aligned} \mathcal{L}_{SM} = & \underbrace{\frac{1}{4} \mathbf{W}_{\mu\nu} \cdot \mathbf{W}^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}}_{\text{kinetic energies and self-interactions of the gauge bosons}} \\ & + \underbrace{\bar{L} \gamma^\mu (i \partial_\mu - \frac{1}{2} g \boldsymbol{\tau} \cdot \mathbf{W}_\mu - \frac{1}{2} g' Y B_\mu) L + \bar{R} \gamma^\mu (i \partial_\mu - \frac{1}{2} g' Y B_\mu) R}_{\text{kinetic energies and electroweak interactions of fermions}} \\ & + \underbrace{\frac{1}{2} |(i \partial_\mu - \frac{1}{2} g \boldsymbol{\tau} \cdot \mathbf{W}_\mu - \frac{1}{2} g' Y B_\mu) \phi|^2 - V(\phi)}_{W^\pm, Z, \gamma, \text{ and Higgs masses and couplings}} \\ & + \underbrace{g'' (\bar{q} \gamma^\mu T_a q) G_\mu^a}_{\text{interactions between quarks and gluons}} + \underbrace{(G_1 \bar{L} \phi R + G_2 \bar{L} \phi_c R + h.c.)}_{\text{fermion masses and couplings to Higgs}} \end{aligned}$$

1) We begin with Quantum Field Theory

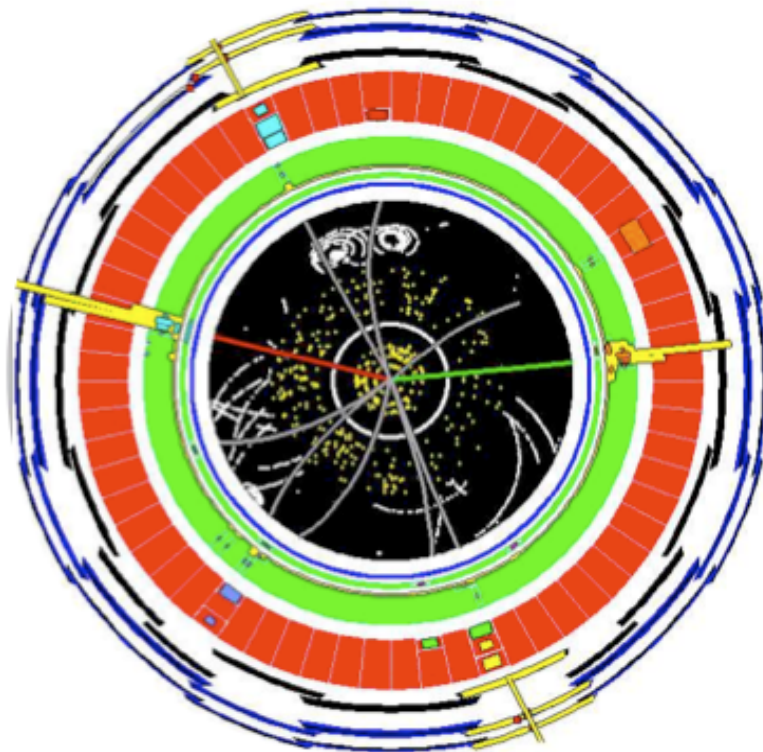
2) Theory gives detailed prediction for high-energy collisions

hierarchical: $2 \rightarrow \mathcal{O}(10) \rightarrow \mathcal{O}(100)$ particles

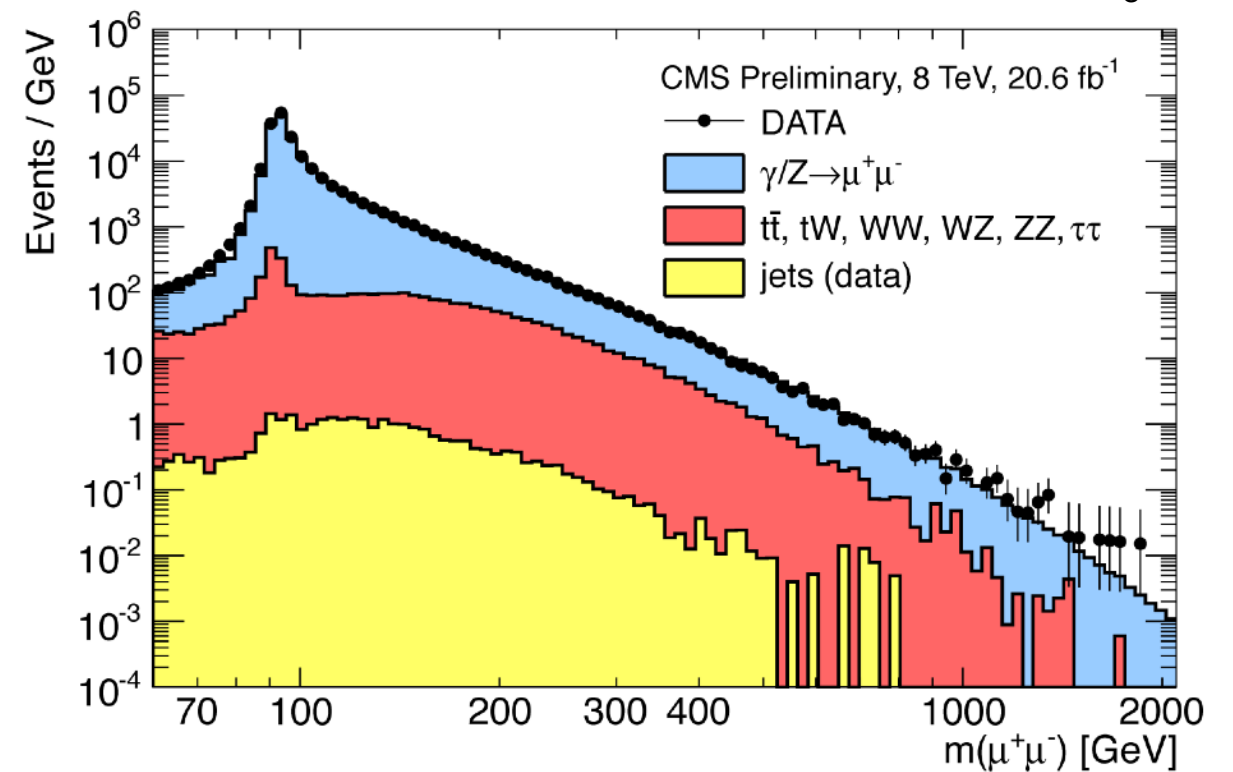
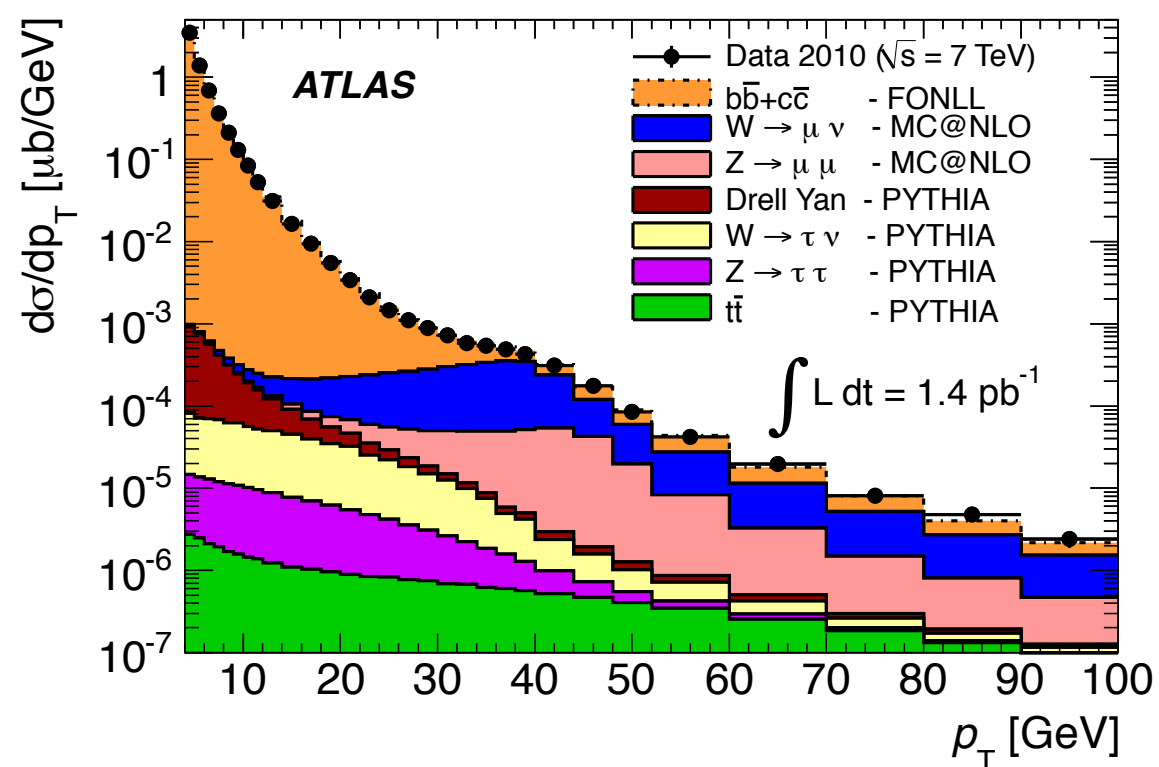
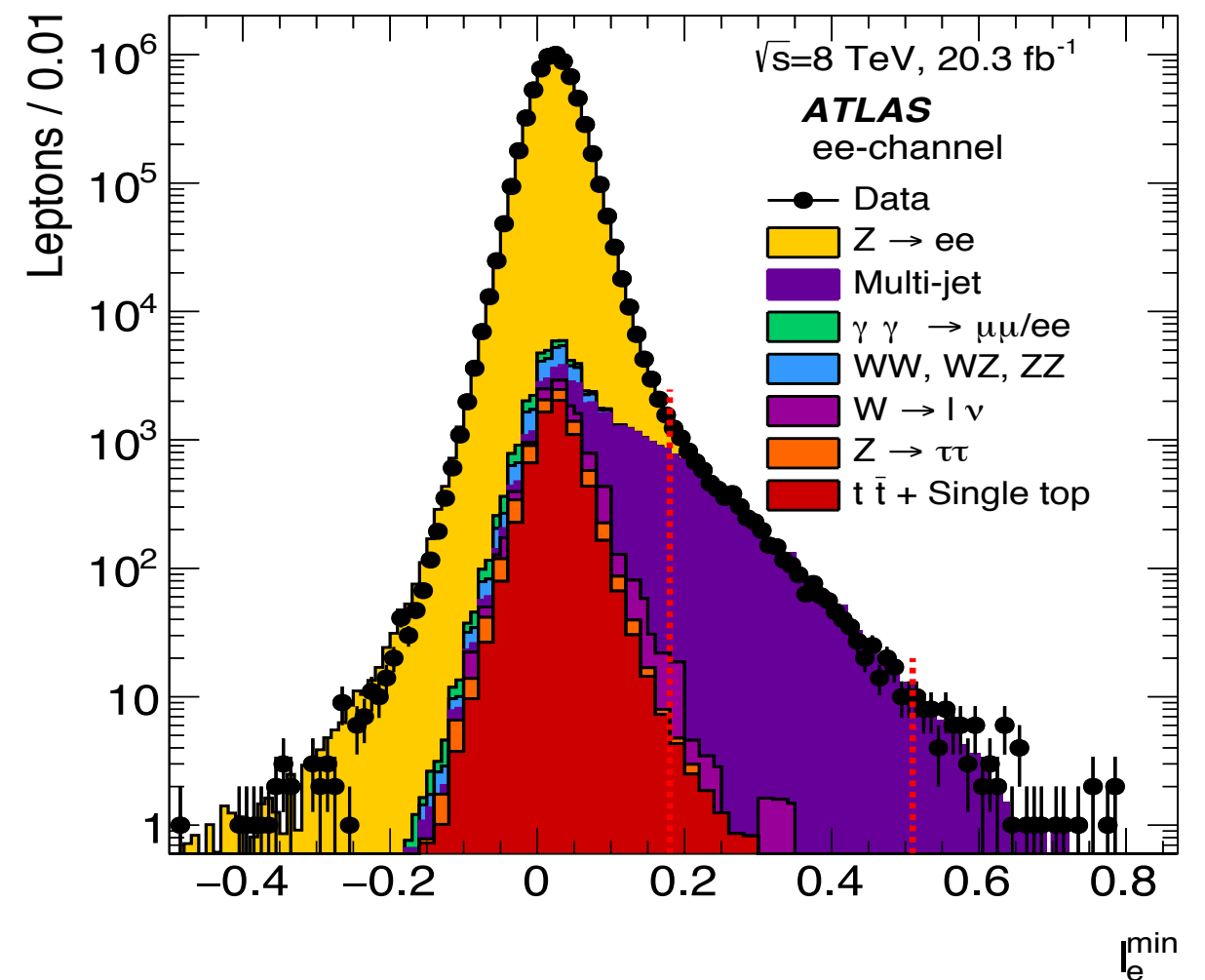
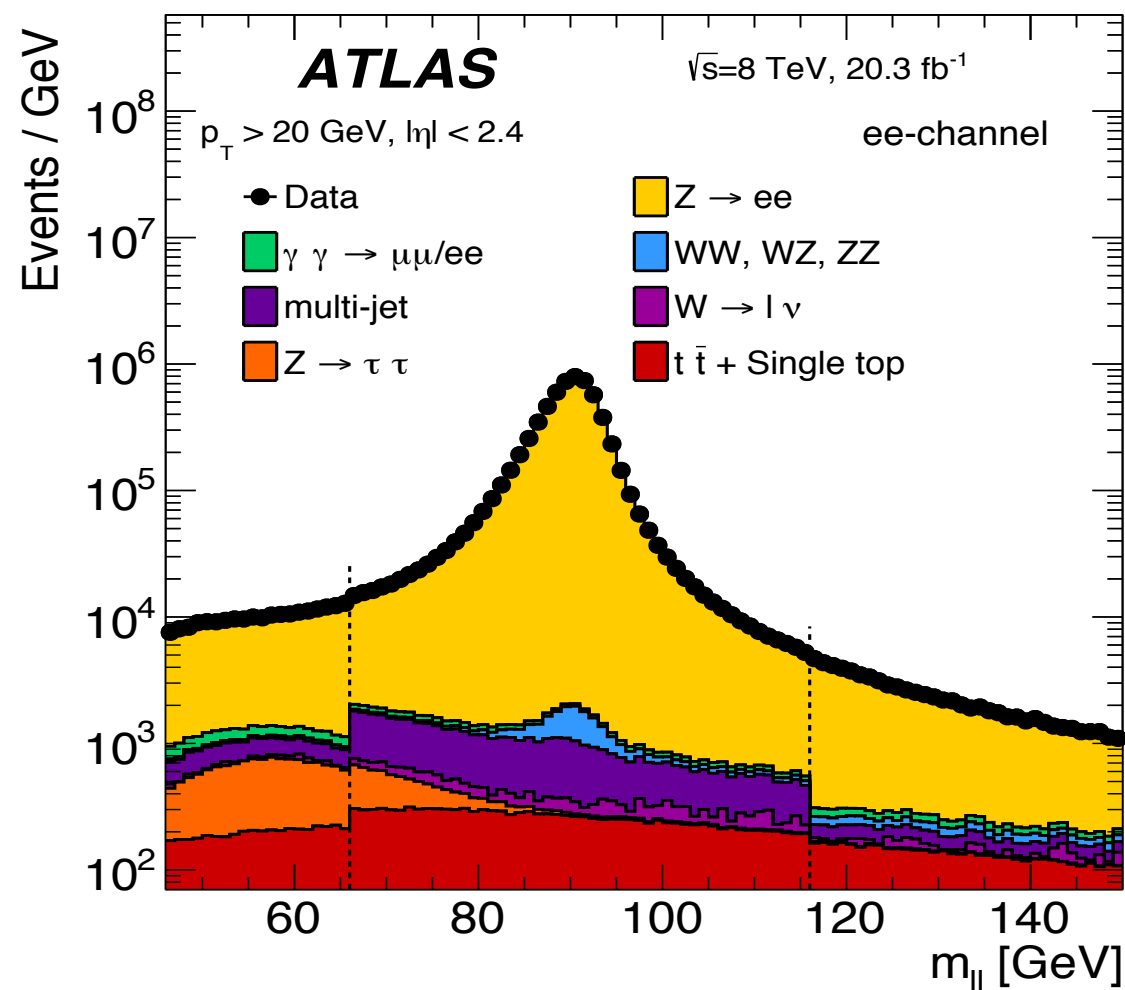


3) The interaction of outgoing particles with the detector is simulated.

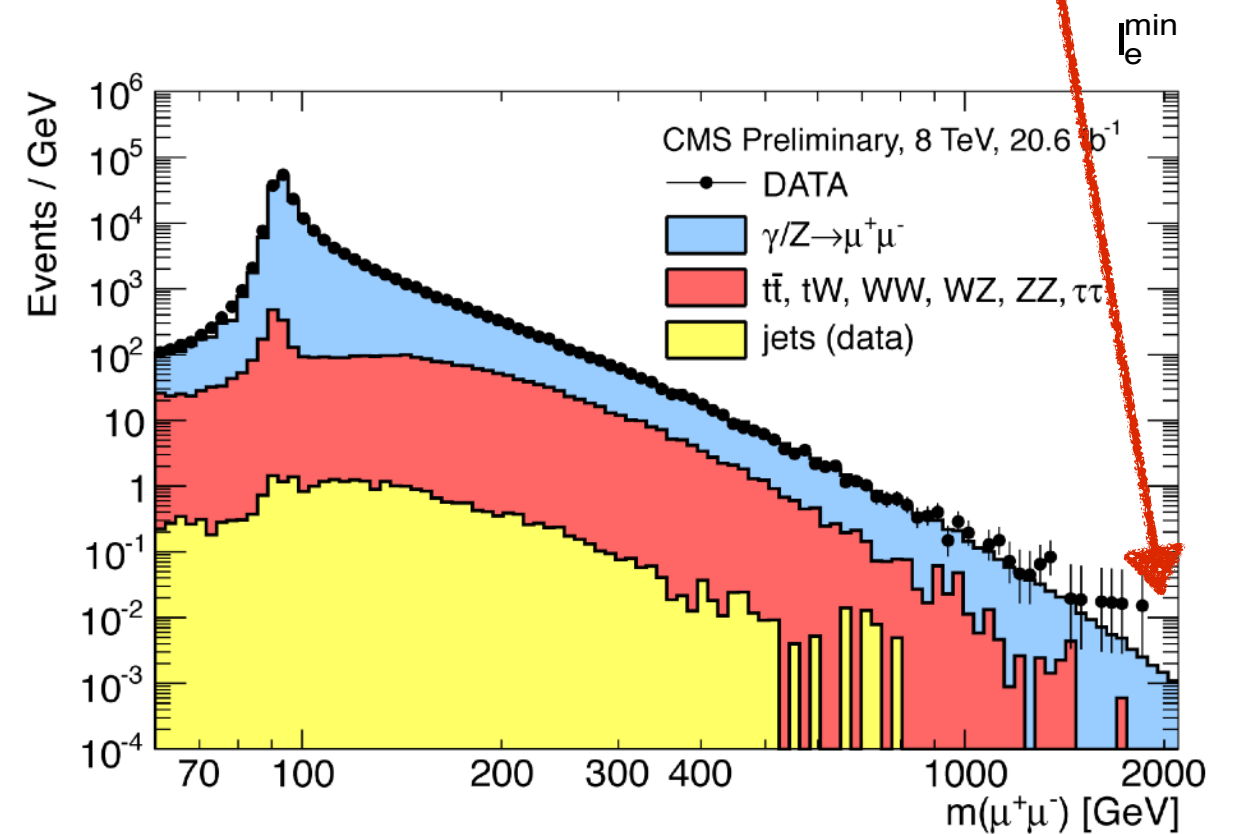
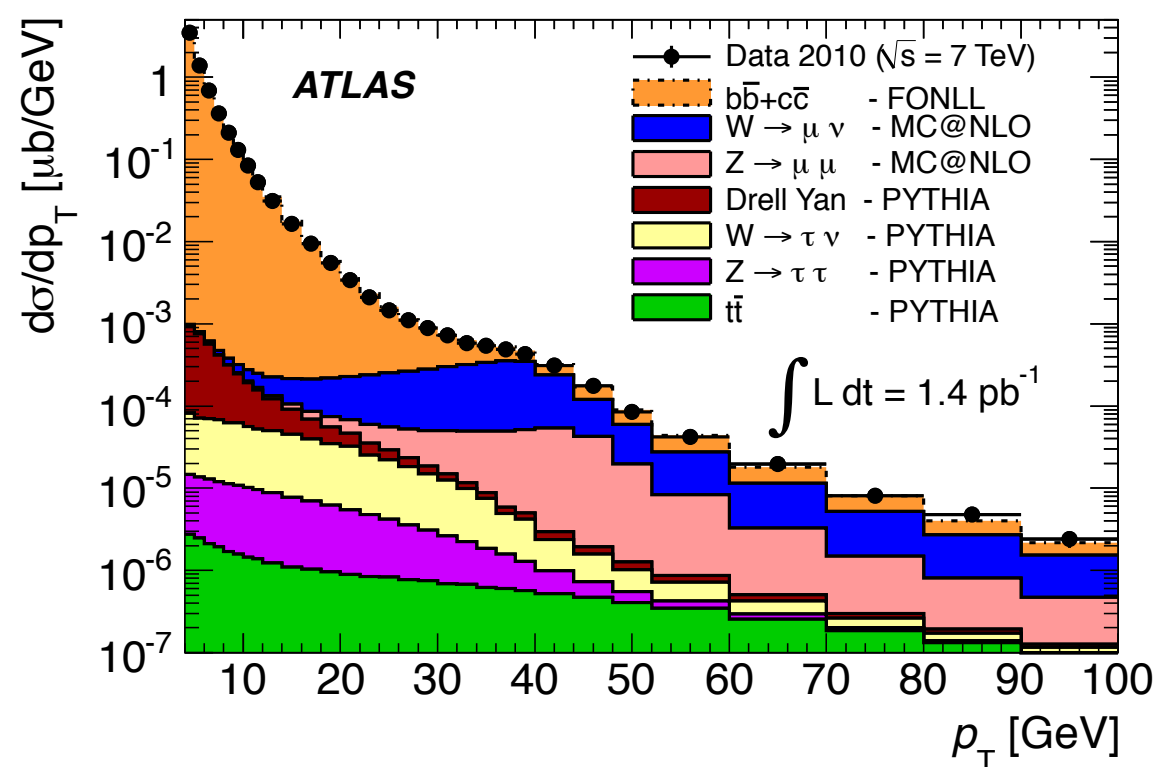
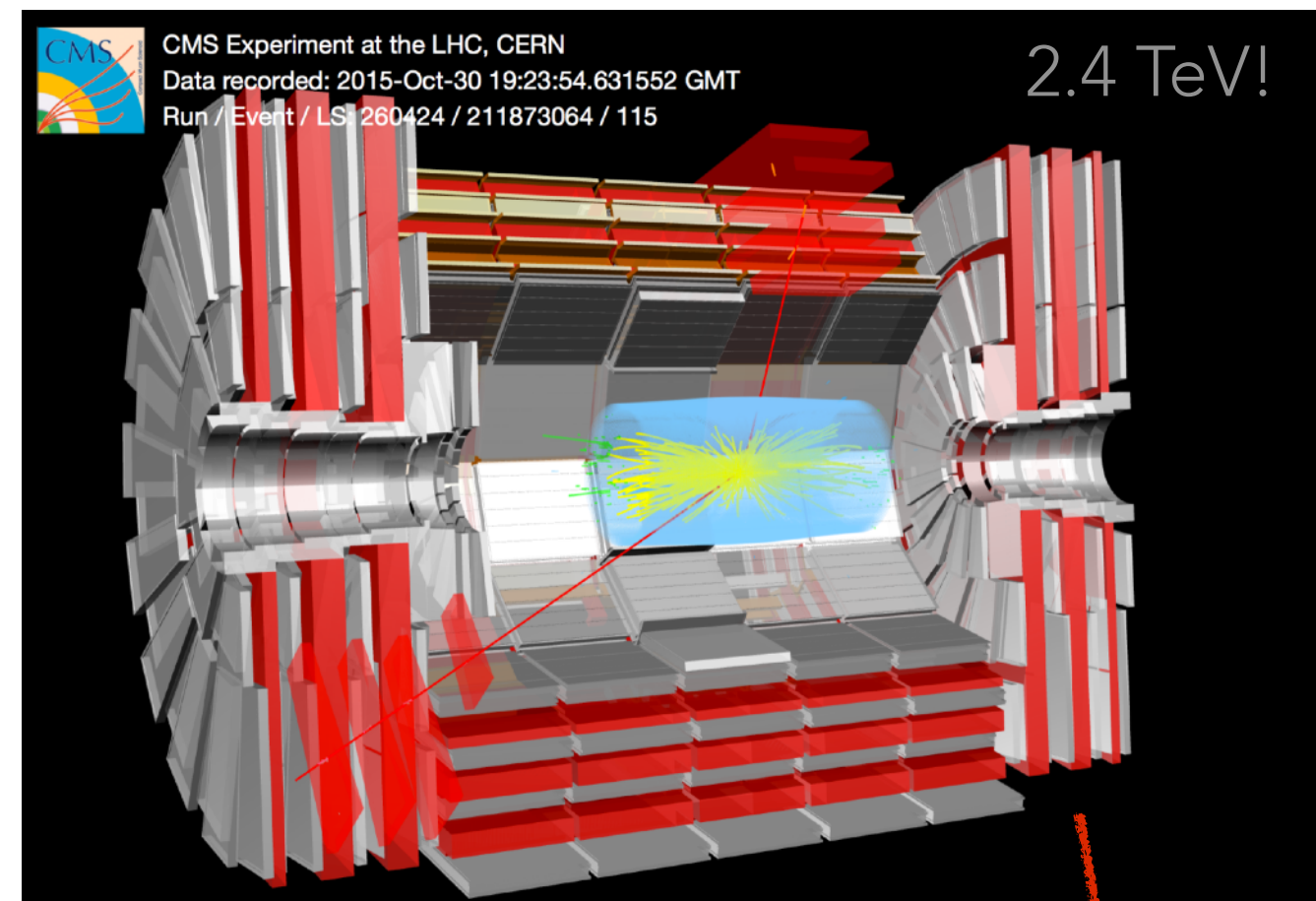
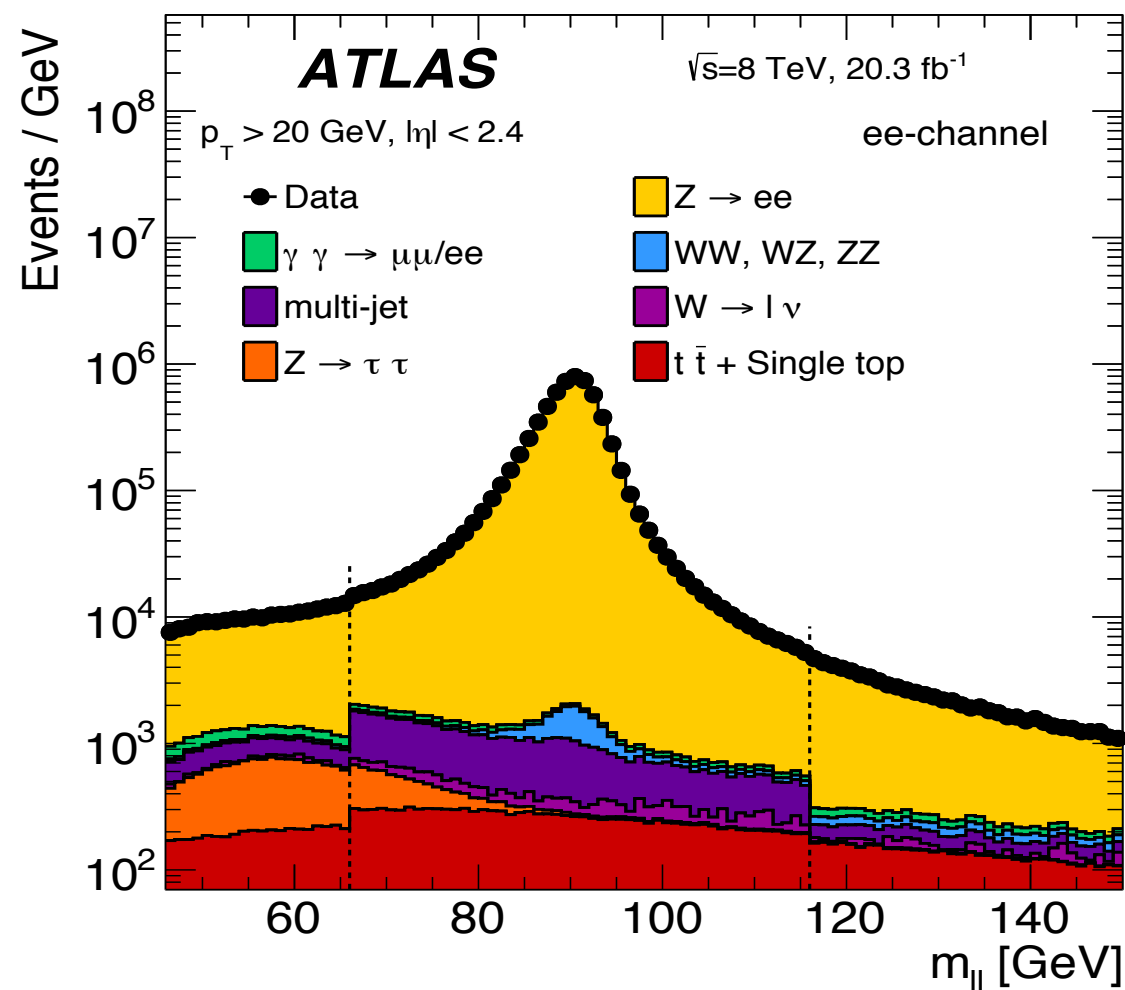
>100 million sensors



HIGH FIDELITY SIMULATION



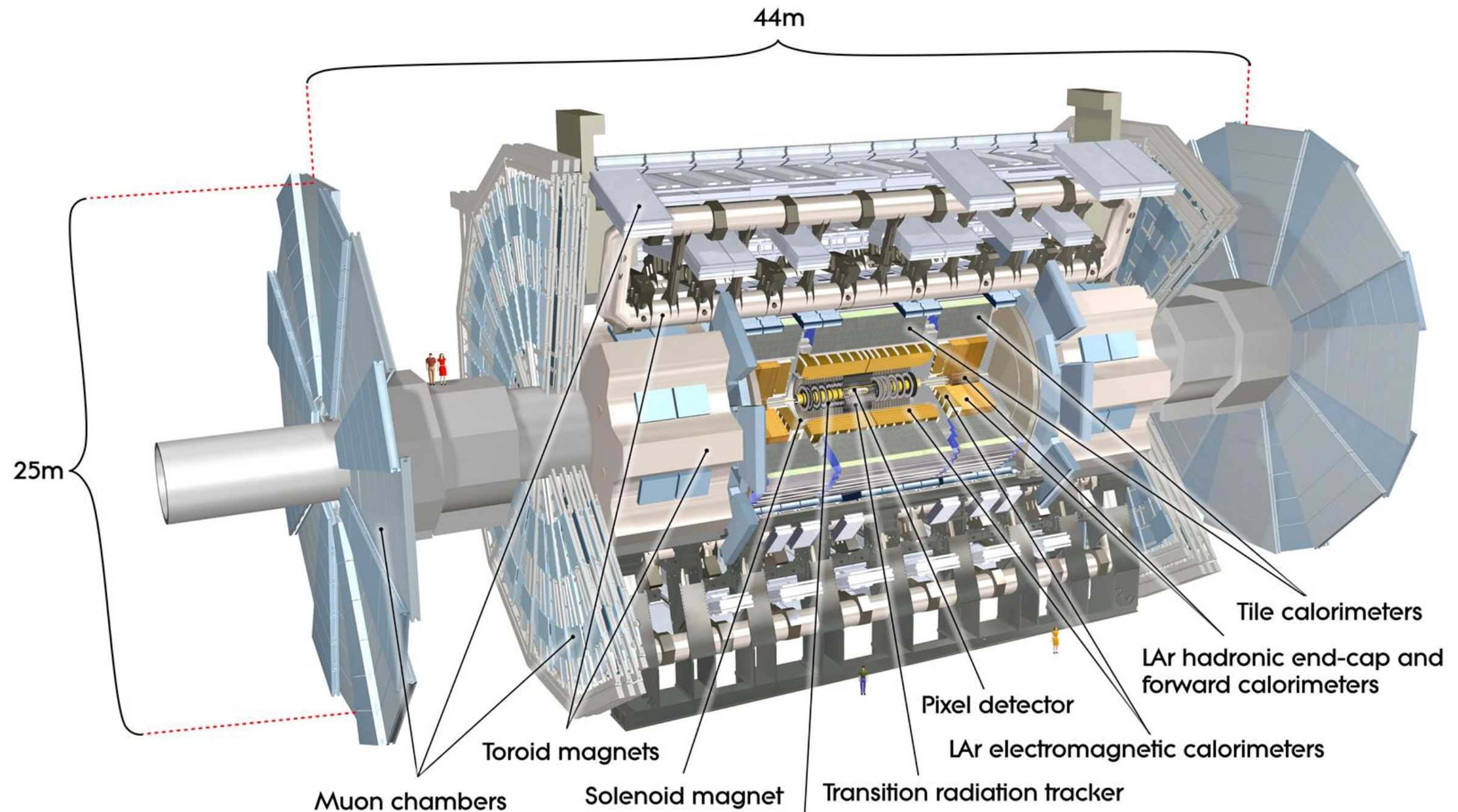
HIGH FIDELITY SIMULATION



HIGH FIDELITY SIMULATION

Detector is 44m long

- Detector resolves details at $< \text{mm}$ scale; Simulation accurate!



HIGH FIDELITY SIMULATION

Detector is 44m long

- Detector resolves details at $< \text{mm}$ scale; Simulation accurate!

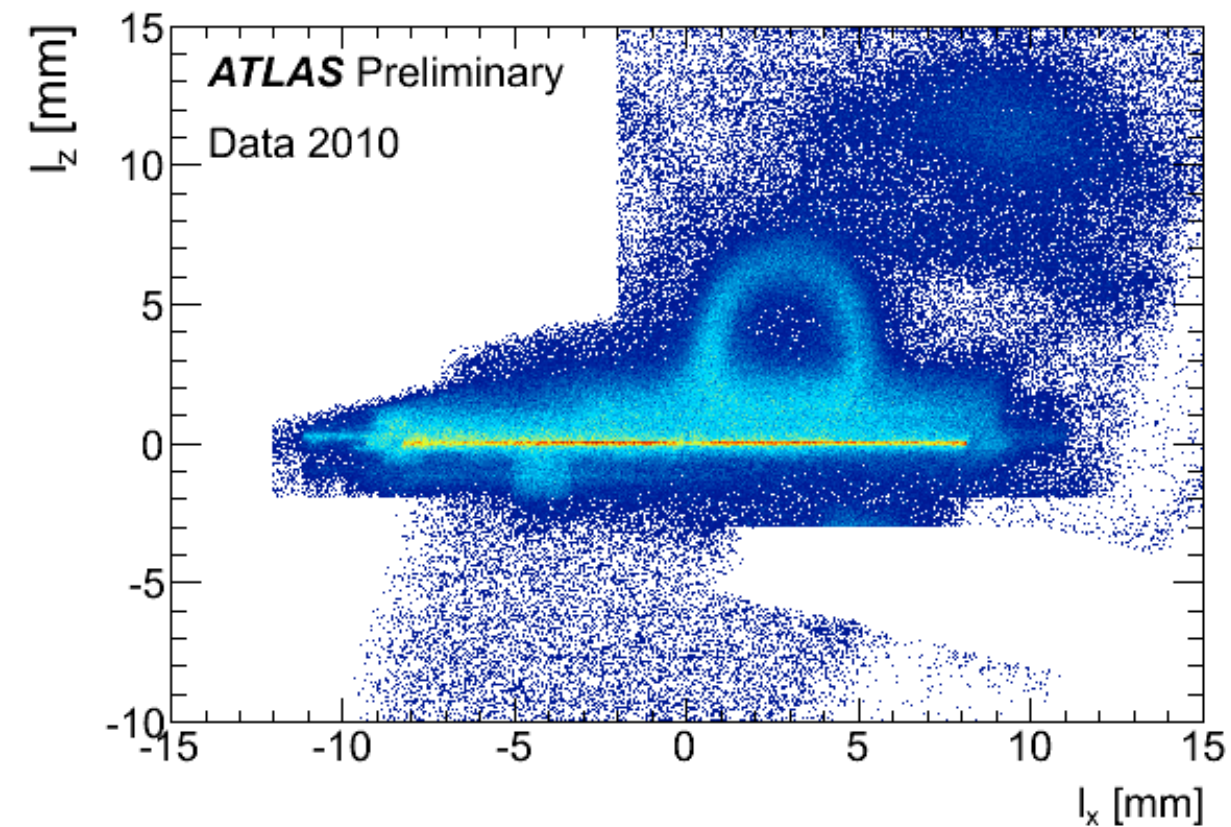
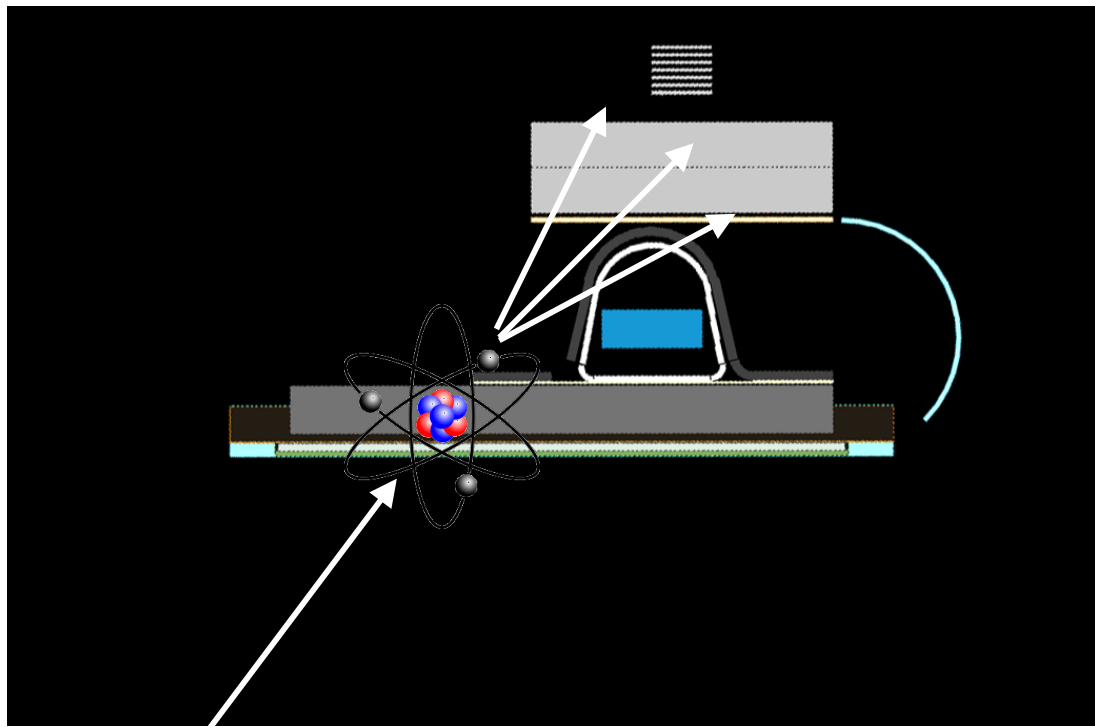


Figure:
ATLAS pixel model as described in simulation (left), tomography from vertices built from tracks for hadronic interactions (right)

Immediate Impact

NEUTRINOLESS DOUBLE BETA DECAY

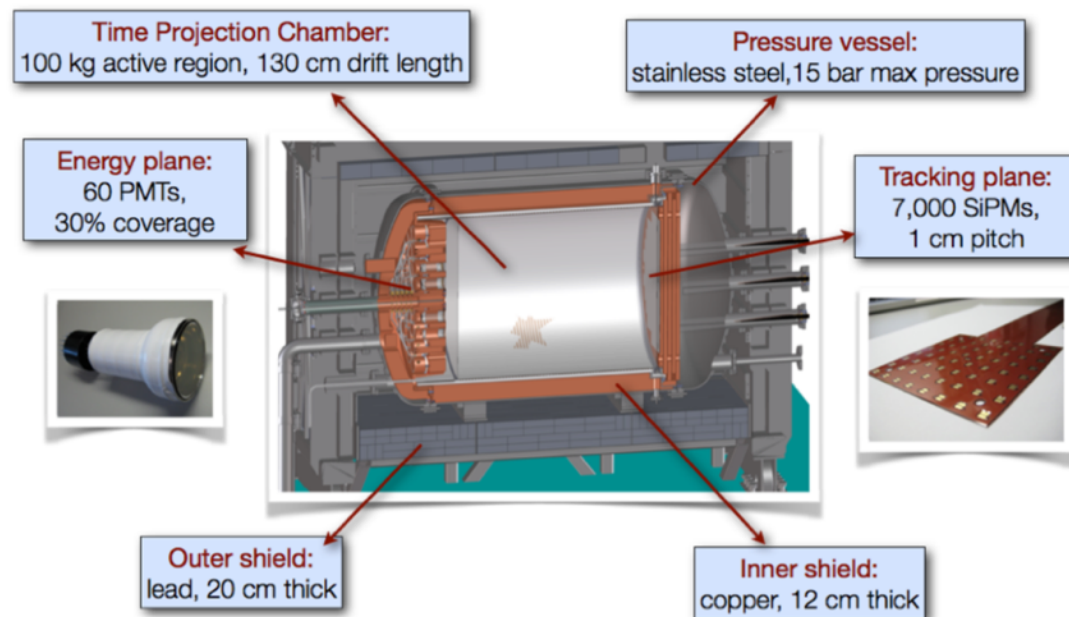
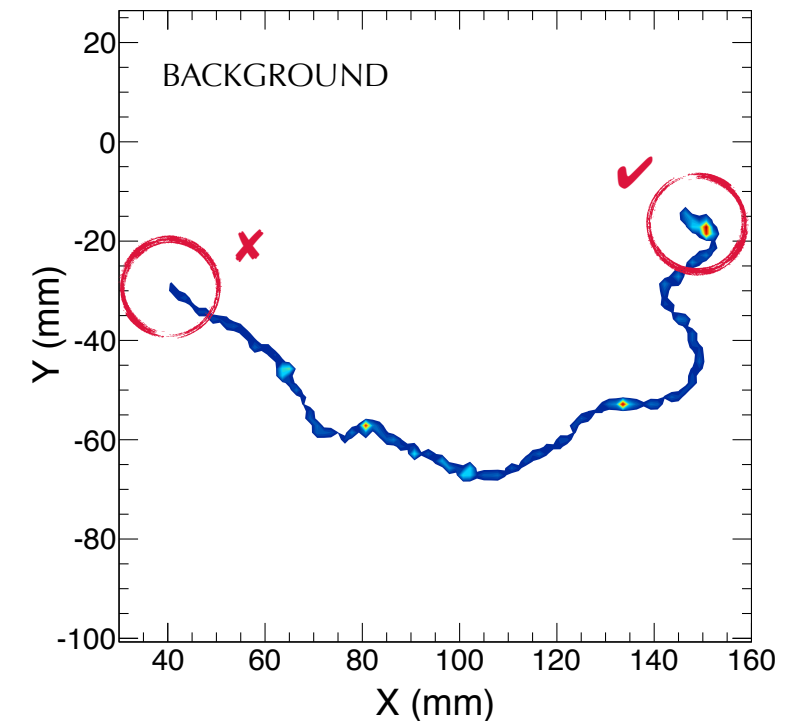
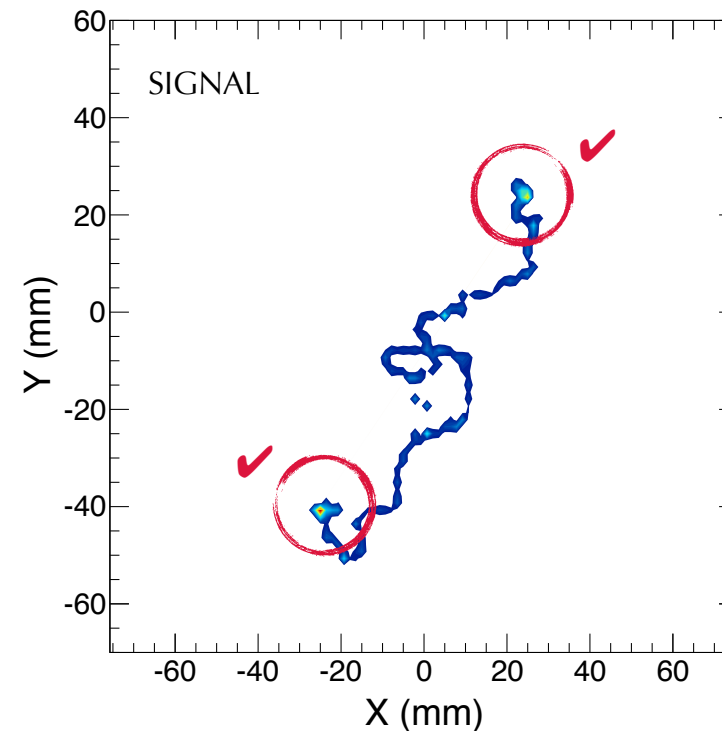
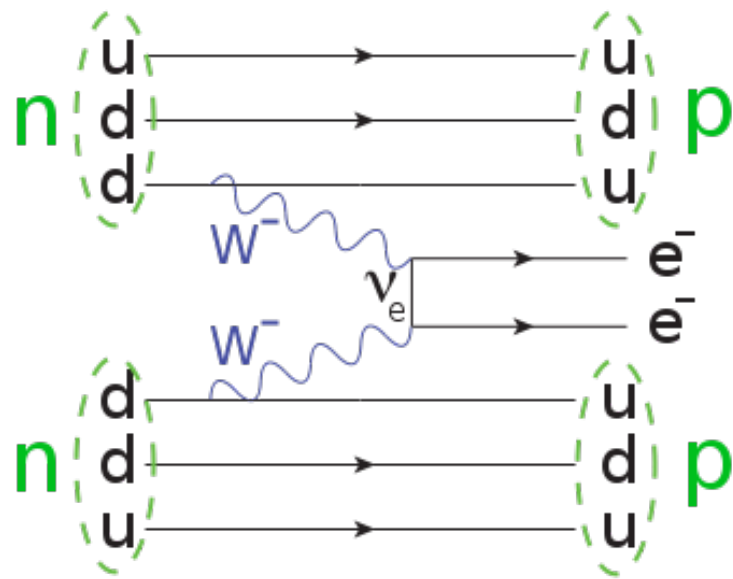
Background rejection in NEXT using deep neural networks

next

[arXiv:1609.06202](https://arxiv.org/abs/1609.06202)

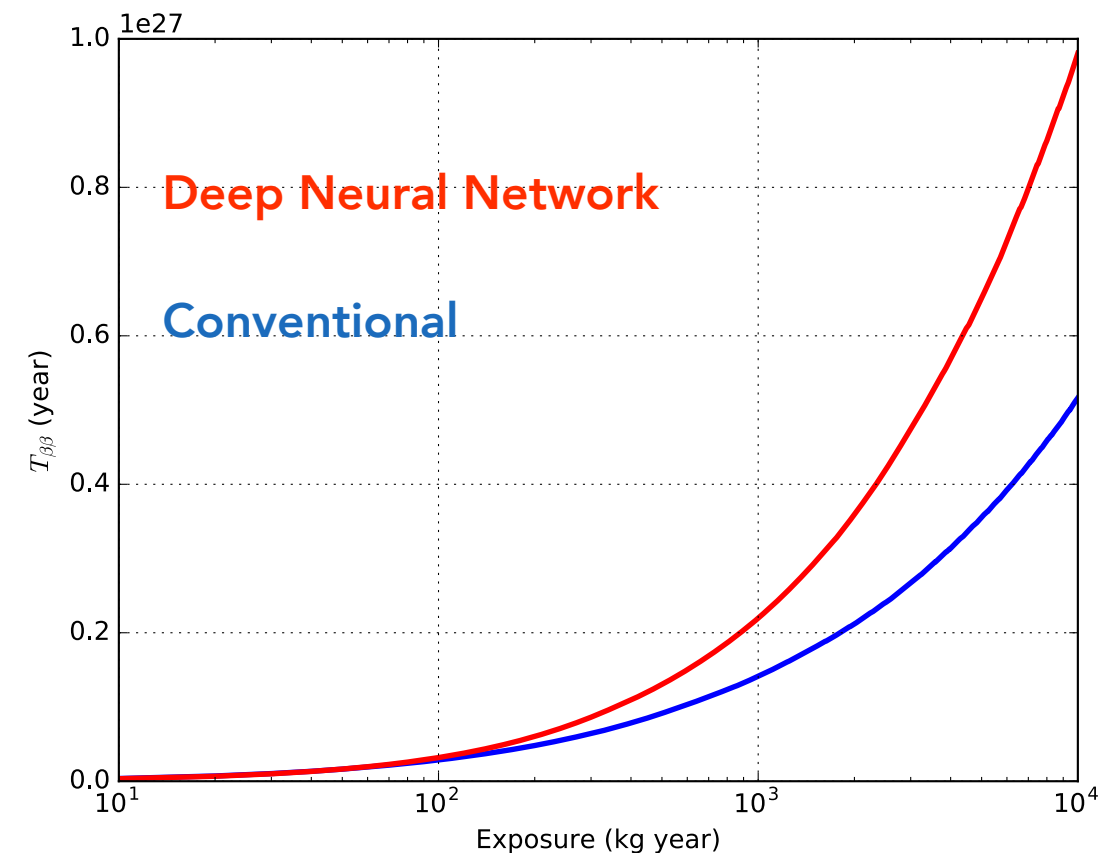
NEXT collaboration

NEXT is a high pressure Xenon (HPXe) Time Projection Chamber (TPC).

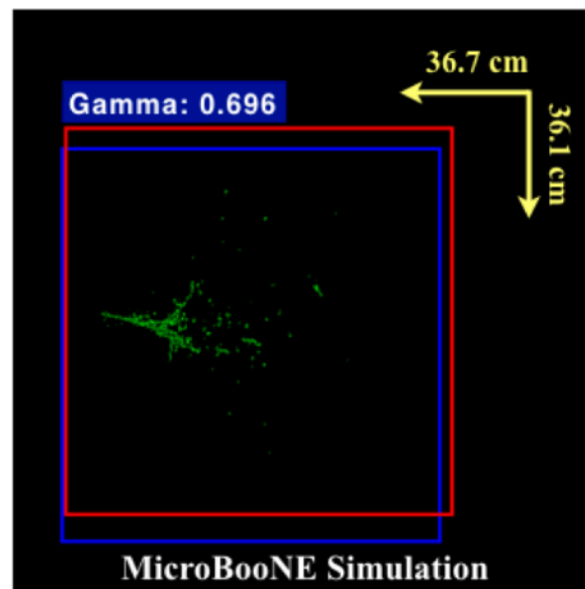
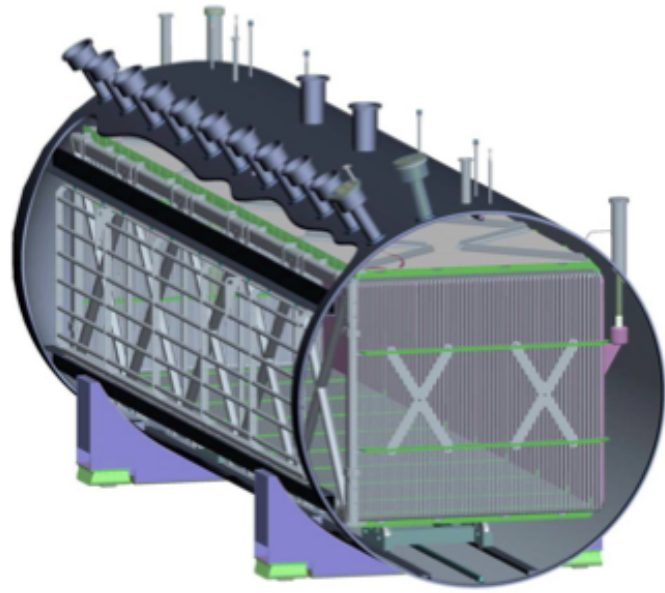


NEXT-100

Neutrinoless double beta decay; from 2017.



NEUTRINOS



Applications of Convolutional Neural Networks and Deep Learning to the Intensity Frontier

NOvA

- CVN-based event classification is used to categorize events as cosmic-ray, neutral-current, ν_e -charged-current, ν_μ -CC, and ν_τ -CC.
- CVN-based algorithms contributed one of the most important improvements in our recent electron neutrino appearance results; equivalent to a 30% increase in statistics.
- Now exploring impacts of these methods elsewhere in reconstruction: clustering, prong-by-prong particle identification, and energy estimation.

MicroBooNE

- Potential for applications to highly-granular liquid argon TPC detectors is huge and is being vigorously explored.

A convolutional neural network neutrino event classifier

A. Aurisano^a, A. Radovic^b, D. Rocco^c, A. Himmel^d, M.D. Messier^a, E. Niner^d, G. Pawloski^c, F. Psihas^a, A. Sousa^a and P. Vahle^b

Published 1 September 2016 • © 2016 IOP Publishing Ltd and Sissa Medialab srl

[Journal of Instrumentation](#), Volume 11, September 2016

Editors' Suggestion

Constraints on Oscillation Parameters from ν_e Appearance and ν_μ Disappearance in NOvA

P. Adamson *et al.* (NOvA Collaboration)
Phys. Rev. Lett. **118**, 231801 – Published 5 June 2017

arXiv.org > hep-ex > arXiv:1706.04592

High Energy Physics – Experiment

Search for active–sterile neutrino mixing using neutral–current interactions in NOvA

NOvA Collaboration: P. Adamson, L. Aliaga, D. Ambrose, N. Anfimov, A. Antoshkin, E. Arrieta-Díaz, K. Augsten, A.

Convolutional neural networks applied to neutrino events in a liquid argon time projection chamber

R. Acciarri^a, C. Adams^a, R. An^b, J. Asaadi^c, M. Auger^a, L. Bagby^a, B. Baller^a, G. Barr^a, M. Bass^a, F. Bay^a + Show full author list

Published 14 March 2017 • © 2017 IOP Publishing Ltd and Sissa Medialab srl

[Journal of Instrumentation](#), Volume 12, March 2017

102 Total downloads



Turn on MathJax

Get permission to re-use this article

MicroBooNE Simulation

Proton: 0.922

19.2 cm

19.2 cm

Muon: 0.903

MicroBooNE Simulation

18.7 cm

18.6 cm

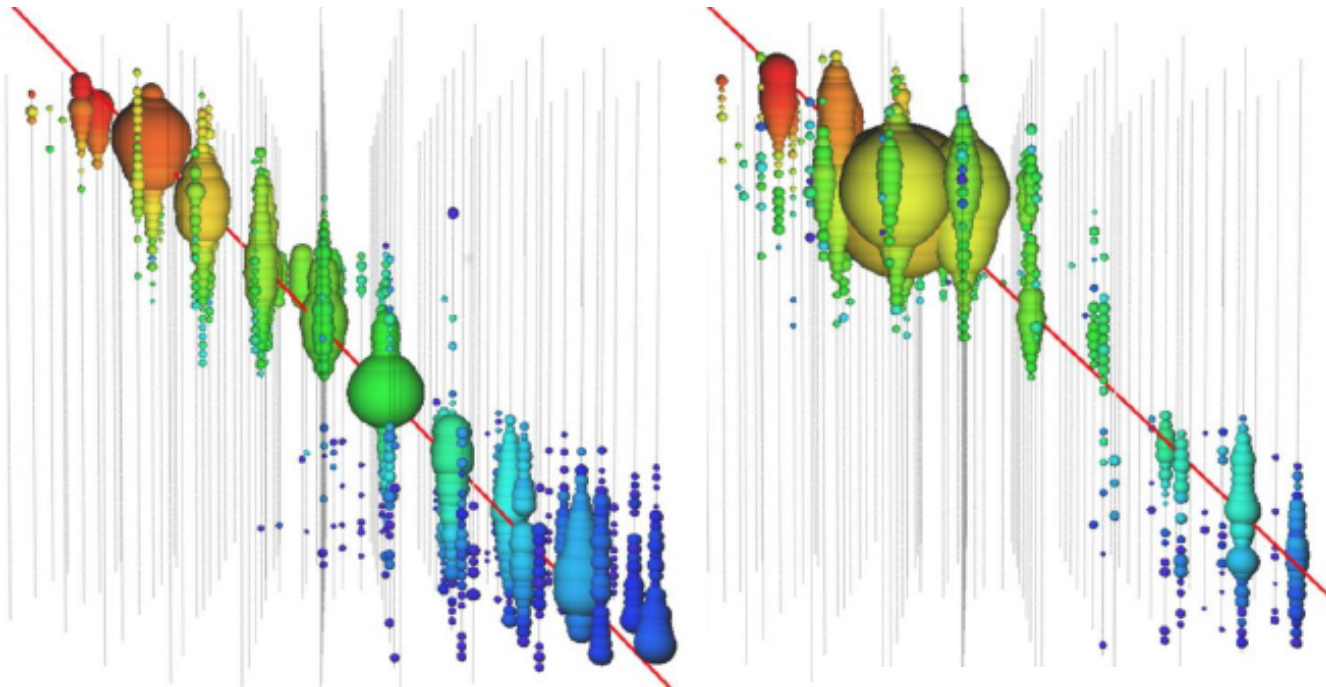
Pion: 0.983

19.9 cm

19.5 cm

MicroBooNE Simulation

HIGH ENERGY NEUTRINOS!



Graph Neural Networks for IceCube Signal Classification

Nicholas Choma^{*}, Federico Monti[†], Lisa Gerhardt[‡], Tomasz Palczewski[§], Zahra Ronaghi[‡], Prabhat[‡], Wahid Bhimji[‡], Michael M. Bronstein[¶], Spencer R. Klein^{§||}, Joan Bruna^{*} for the IceCube collaboration^{**}

^{*}*Courant Institute of Mathematical Sciences
New York University
New York NY 10012 USA*

[†]*Institute of Computational Science
Università della Svizzera italiana
Lugano 6900, Switzerland*

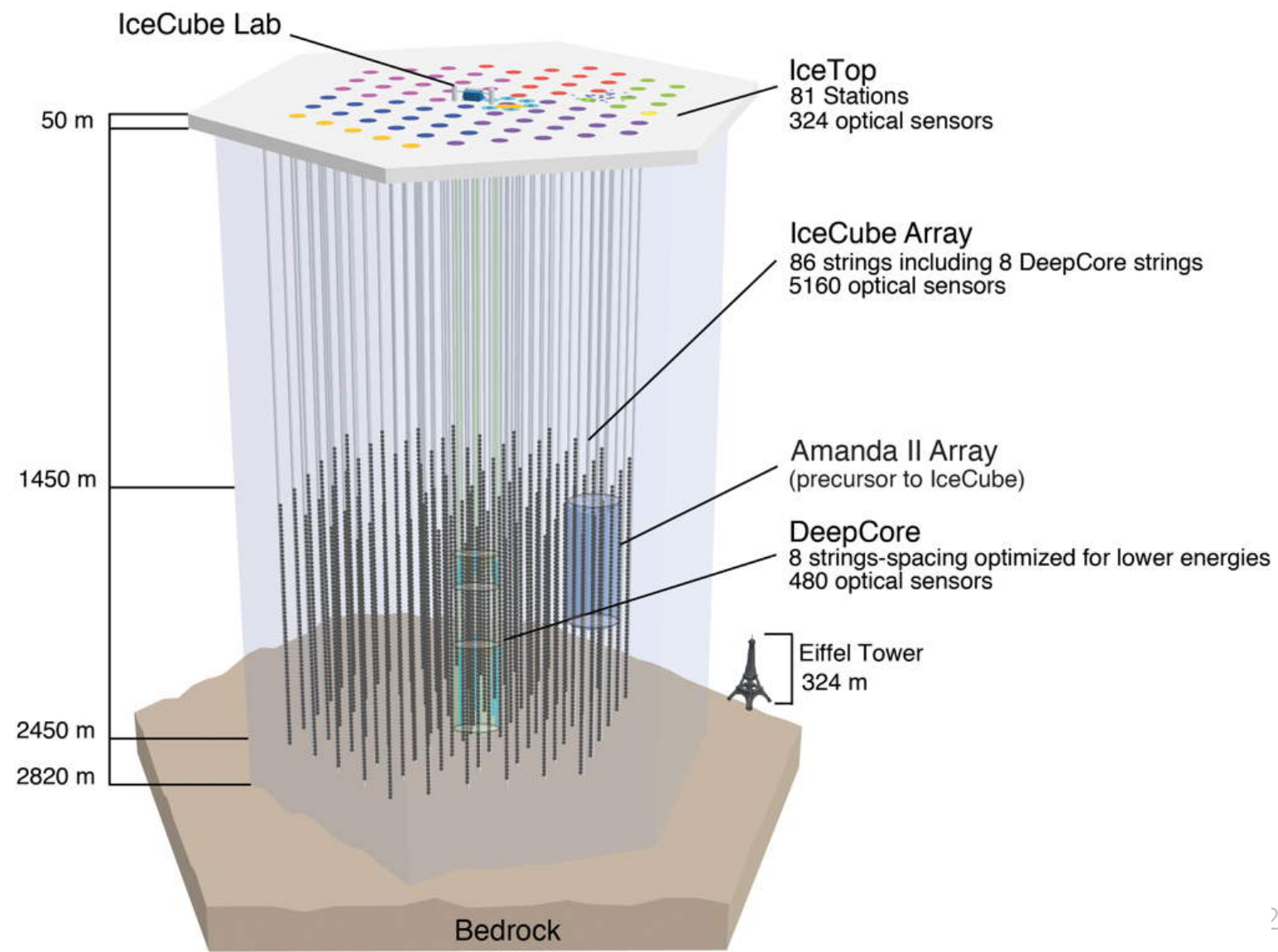
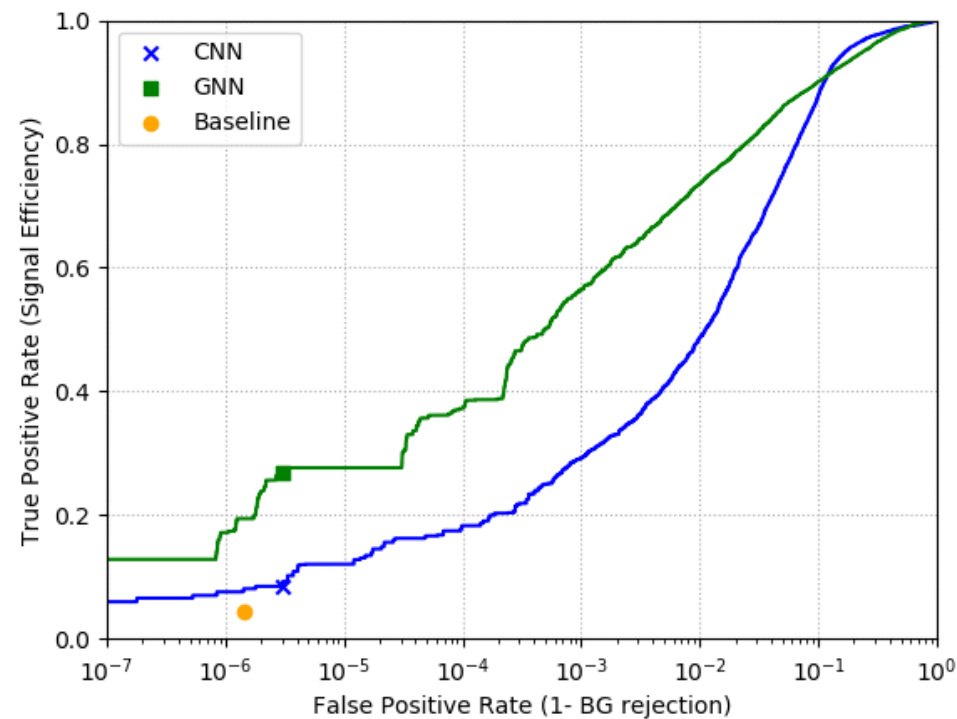
[§]*Department of Physics
University of California, Berkeley
Berkeley CA 94720 USA*

[‡]*NERSC
Lawrence Berkeley National Laboratory
Berkeley CA 94720 USA*

[¶]*Department of Computing
Imperial College
London SW7 2AZ, UK*

^{||}*Nuclear Science Division
Lawrence Berkeley National Laboratory
Berkeley CA 94720 USA*

nc2201@cims.nyu.edu, federico.monti@usi.ch, lgerhardt@lbl.gov, tpalczewski@berkeley.edu, zronaghi@lbl.gov, prabhat@lbl.gov, wbhimji@lbl.gov, m.bronstein@imperial.ac.uk, srklein@lbl.gov, bruna@cims.nyu.edu



JETS

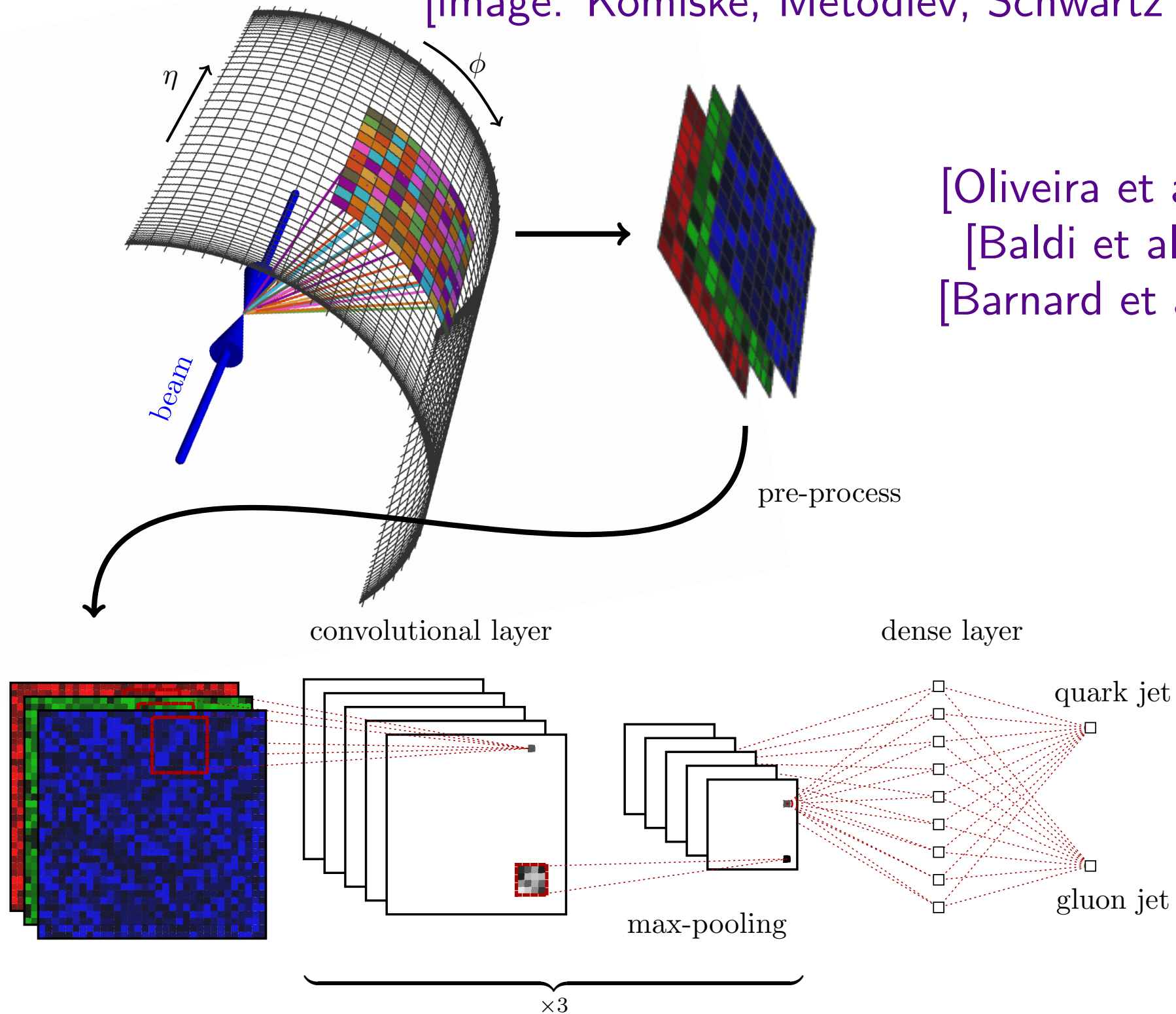
Run: 329716
Event: 857582452
2017-07-14 10:48:51 CEST



ATLAS
EXPERIMENT

JET IMAGES

[image: Komiske, Metodiev, Schwartz arxiv:1612.01551]



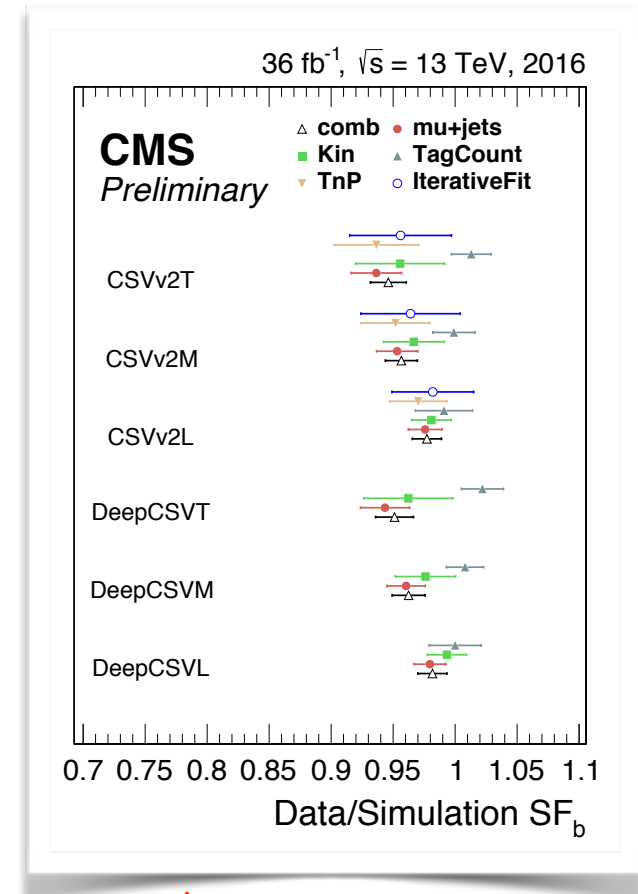
[Oliveira et al arXiv:1511.05190]

[Baldi et al arXiv:1603.09349]

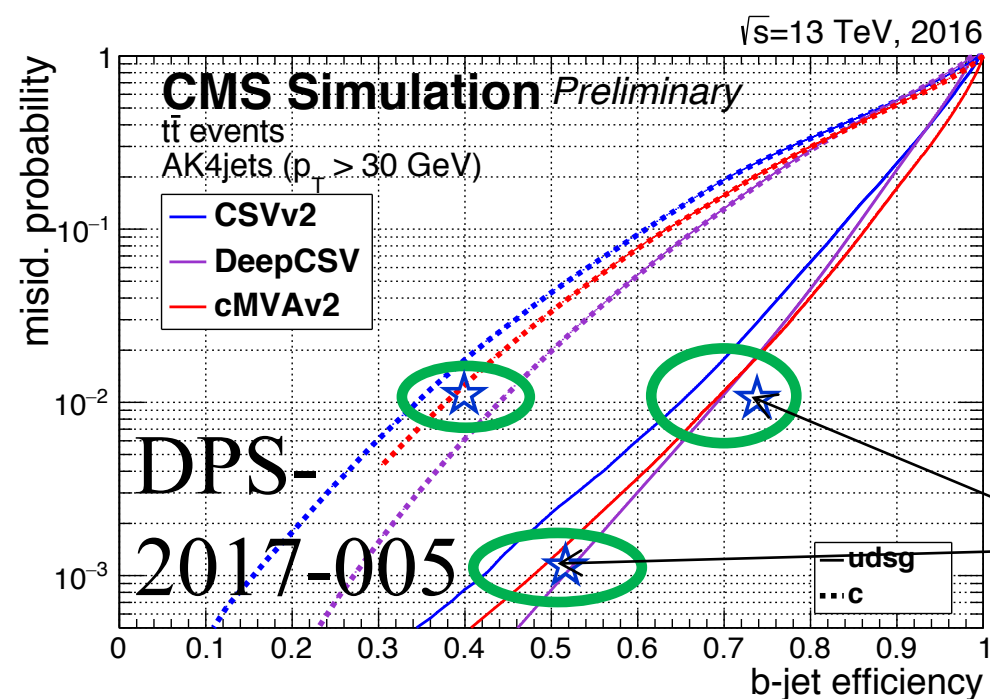
[Barnard et al arXiv:1609.00607]

New deep learned flavour tagger

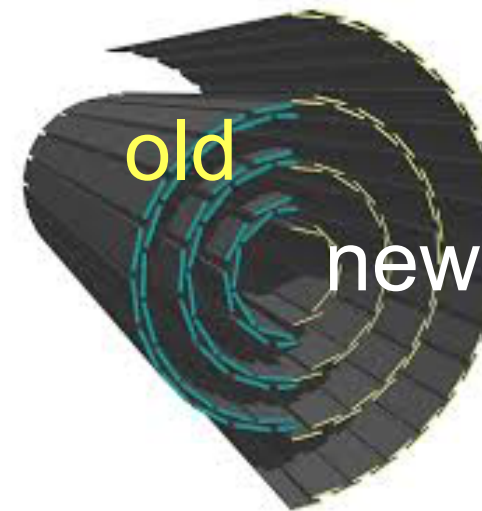
- Diverse samples
 - QCD and $t\bar{t}$
- Optimal training: Large jet sample to avoid over-fitting
 - 50M jets!
- Use complete standard CSV b-tag “Tag info” (from $\sim 30 \rightarrow 60$)
- Dense Deep Neural Network (Dense)



↑ systematics
under control



☆ Old tagger with new pixel detector
in simulation



Similar impact as the new inner pixel! ←

Higgs boson observed decaying to b quarks – at last!

By [ATLAS Collaboration](#), 9th July 2018

The Brout-Englert-Higgs mechanism solves the apparent theoretical impossibility of weak vector bosons (W and Z) to have mass. The discovery of the Higgs boson in 2012 via its decays into photon, Z and W pairs was a triumph of the Standard Model built upon this mechanism. The Higgs field can also be used in an elegant way to provide mass to charged fermions (quarks and leptons) through interactions involving “Yukawa couplings”, with strength proportional to the particle mass. The [observation of the Higgs boson decaying into pairs of \$\tau\$ leptons](#) provided the first direct evidence of this type of interaction.

Six years after its discovery, ATLAS has observed about 30% of the Higgs boson decays predicted in the Standard Model. However, the favoured decay of the Higgs boson into a pair of b quarks ($H \rightarrow b\bar{b}$), which is expected to account for almost 60% of all possible decays, had remained elusive up to now. Observing this decay mode and measuring its rate is a mandatory step to confirm (or not...) the mass generation for fermions via Yukawa interactions, as predicted in the Standard Model.

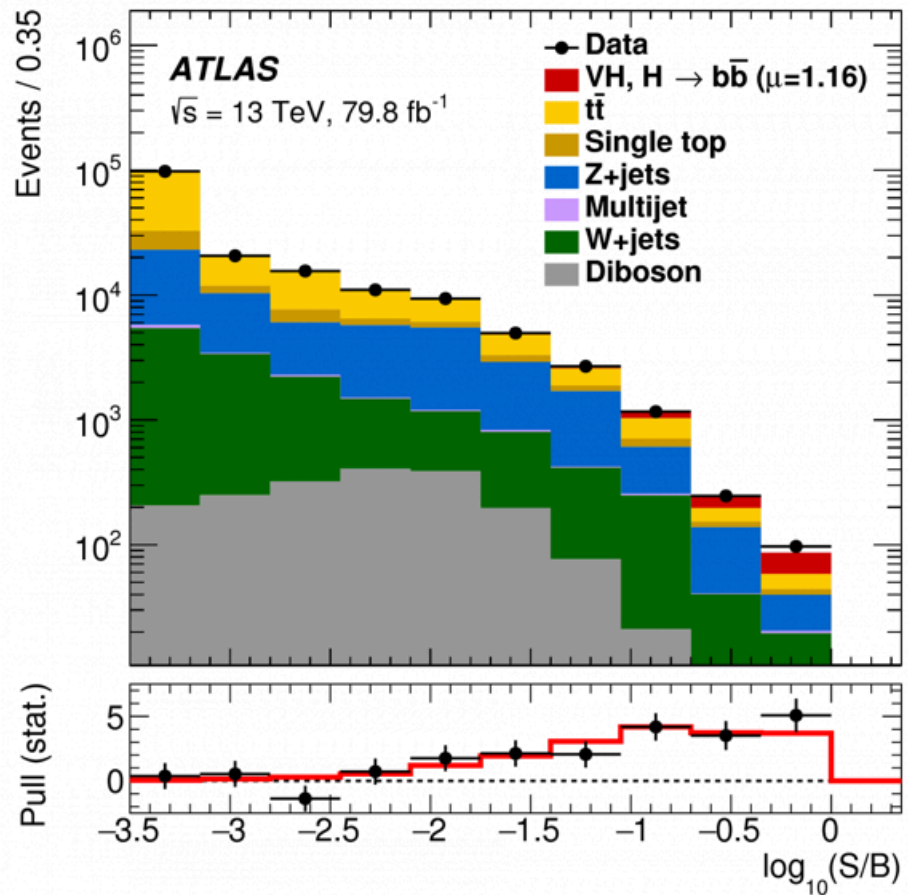
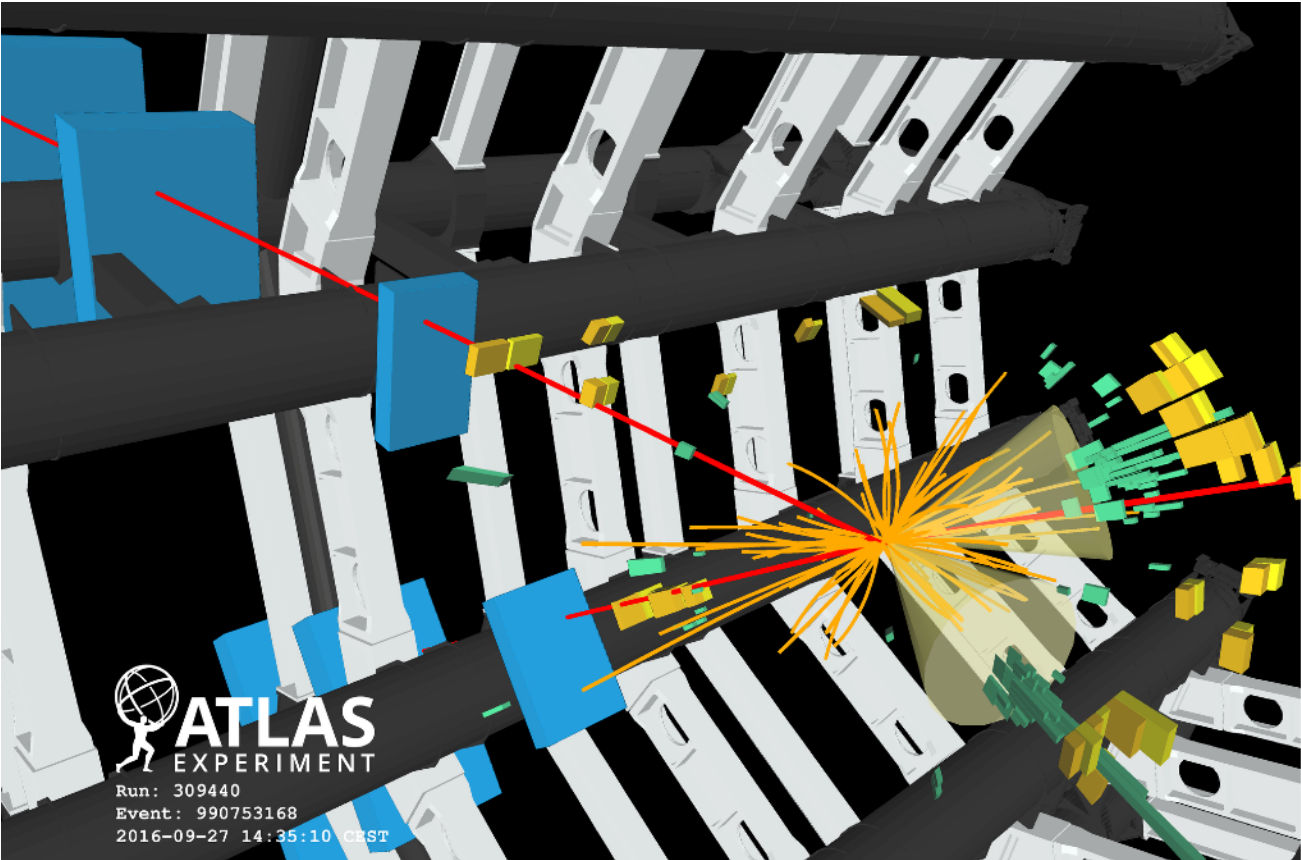


Figure 3: Distribution showing the combination of all BDT outputs reordered in terms of $\log(S/B)$, where S and B are the signal and background yields, respectively. The signal is shown in red, and the different backgrounds in various colours. The data are shown as points with error bars. The lower panel shows the “pull”, i.e. the ratio of data minus background to the statistical uncertainty of the background. (Image: ATLAS Collaboration/CERN)

In Progress

 Featured Prediction Competition

TrackML Particle Tracking Challenge

High Energy Physics particle tracking in CERN detectors

\$25,000

Prize Money



CERN · 656 teams · 2 months ago

[Overview](#)[Data](#)[Kernels](#)[Discussion](#)[Leaderboard](#)[Rules](#)

Overview

Description

Evaluation

Timeline

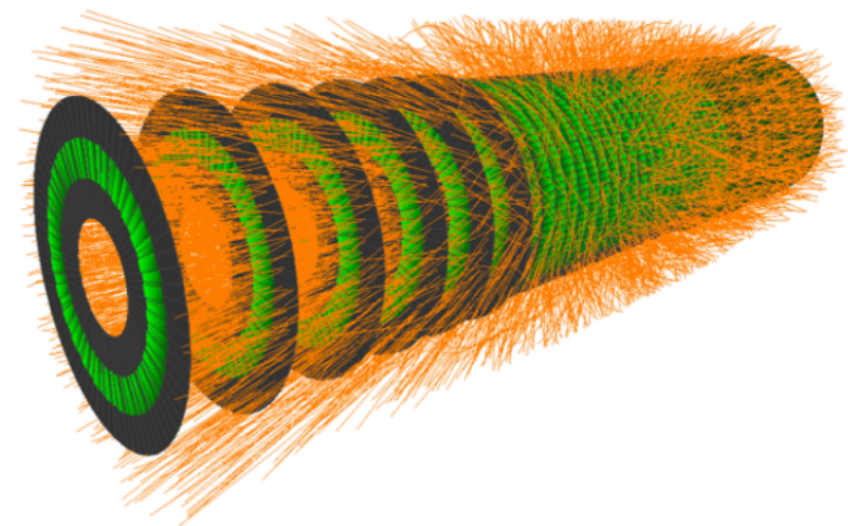
Prizes

About The Sponsors

To explore what our universe is made of, scientists at CERN are colliding protons, essentially recreating mini big bangs, and meticulously observing these collisions with intricate silicon detectors.

While orchestrating the collisions and observations is already a massive scientific accomplishment, analyzing the enormous amounts of data produced from the experiments is becoming an overwhelming challenge.

Event rates have already reached hundreds of millions of collisions per second, meaning physicists must sift through tens of petabytes of data per year. And, as the resolution of detectors improve, ever better software is needed for real-time pre-processing and filtering of the most promising events, producing even more data.

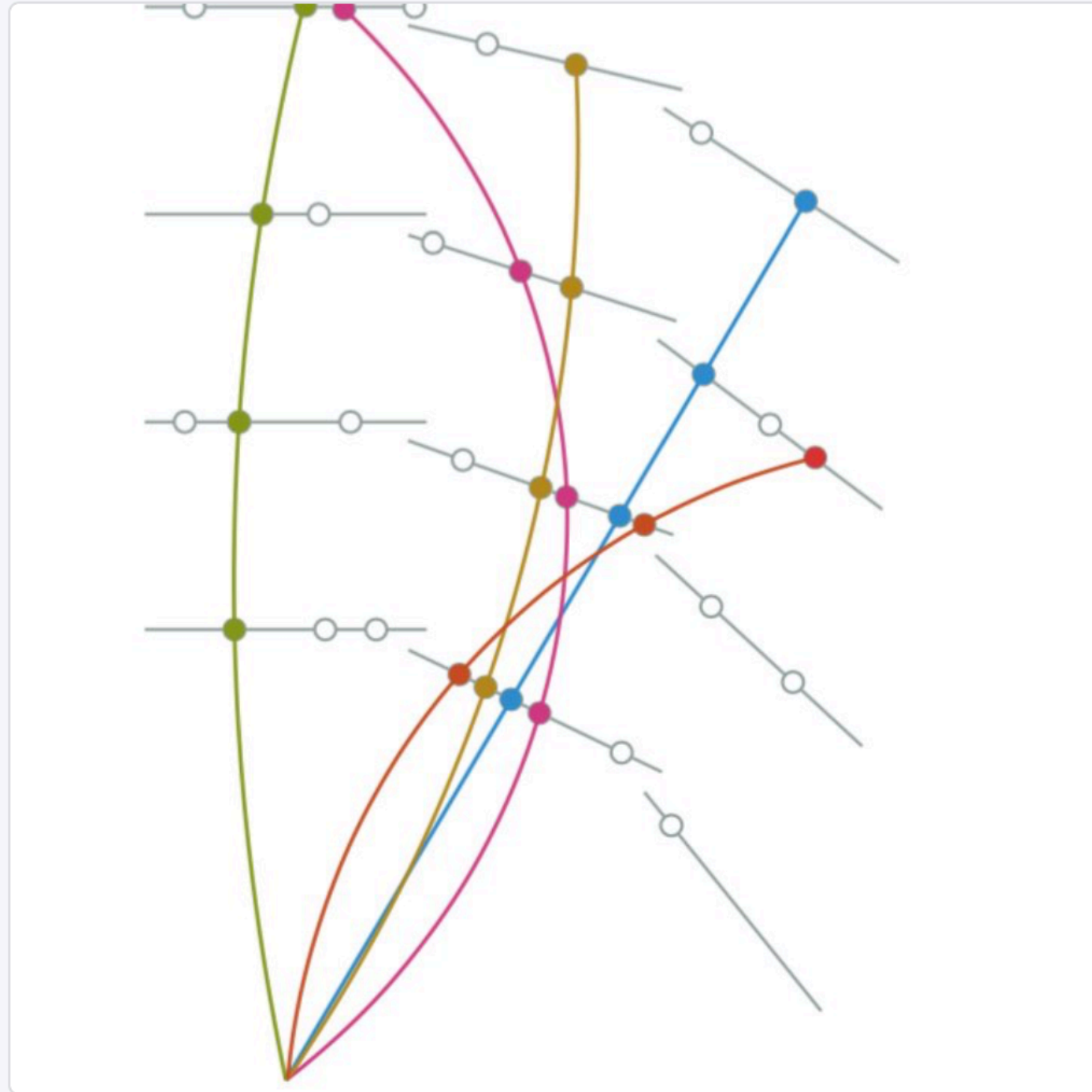


THE PARTICLE TRACKING CHALLENGE



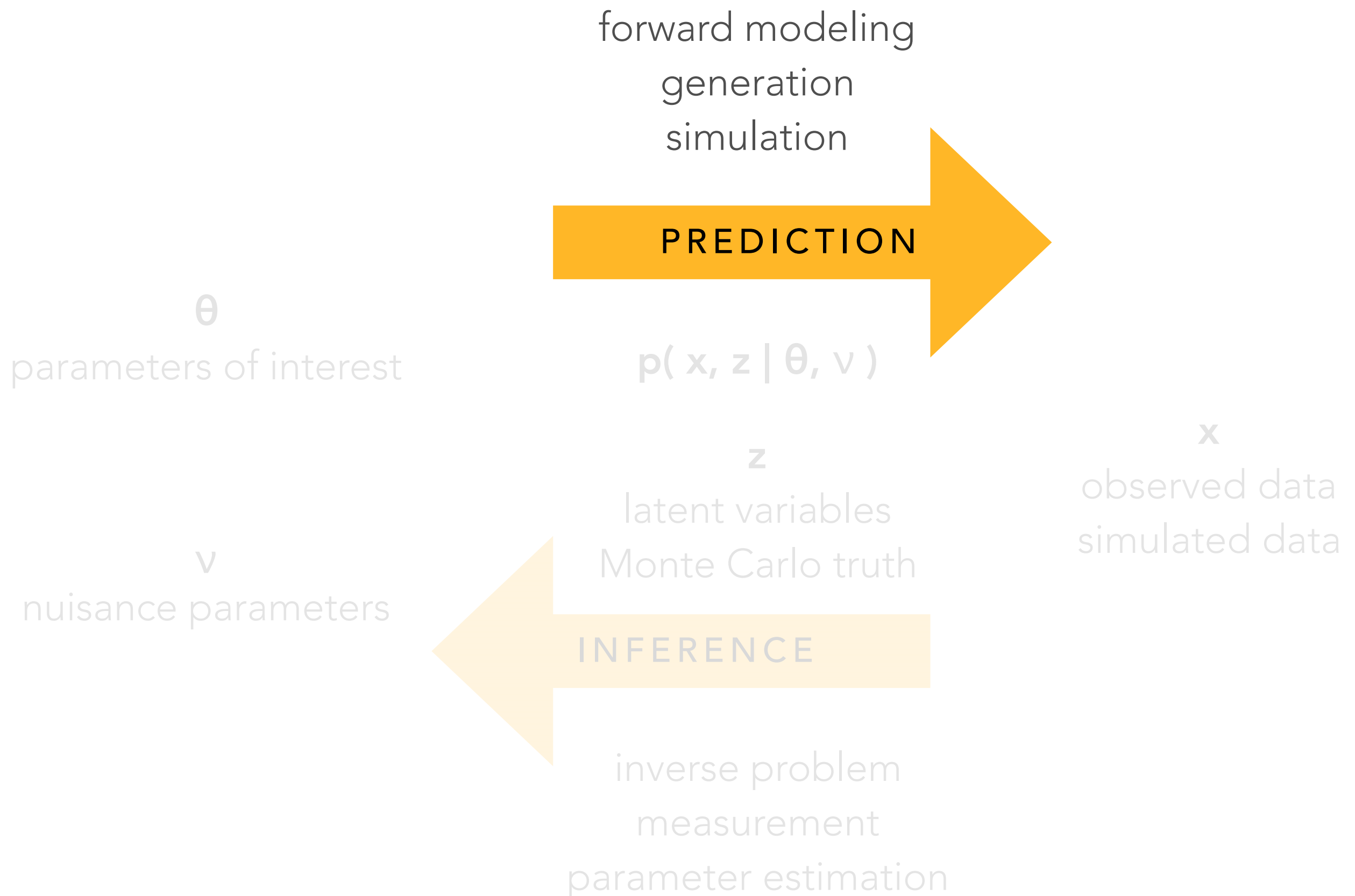
trackml @trackmlhc · Sep 12

So how it works: you're given 3D points coordinates, which you have to group into tracks which are arc of helices : 100.000 points into 10.000 tracks. Very simplified 2D view [#trackml](#)





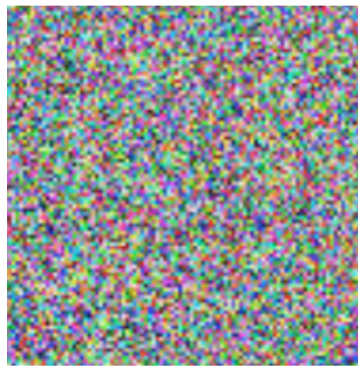
THE PLAYERS



GENERATIVE ADVERSARIAL NETWORKS

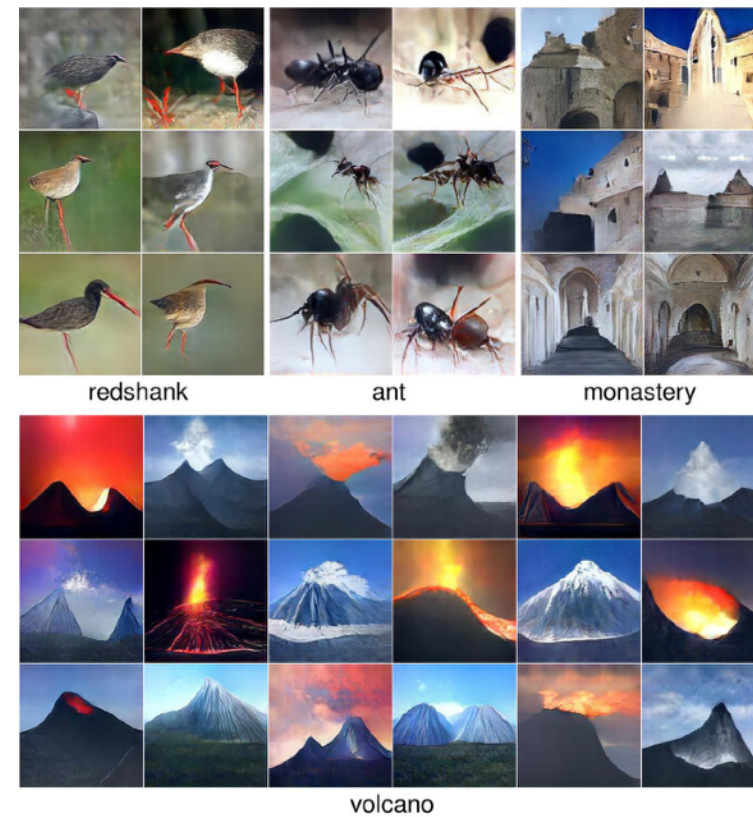
Z

Noise $\sim N(0,1)$



Generative
Model

X



A STEVEN SPIELBERG FILM

catch me

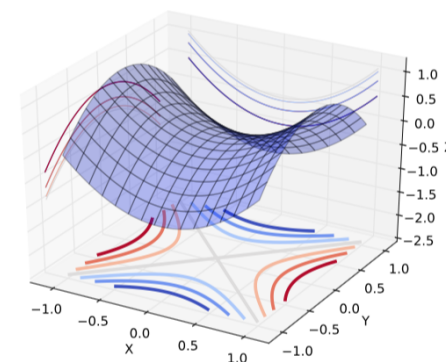
if you can

Leo is G

Tom is D

- We want to
 - For fixed G , find D which **maximizes** $V(D, G)$,
 - For fixed D , find G which **minimizes** $V(D, G)$;
- In other words, we are looking for the *saddle point*

$$(D^*, G^*) = \max_D \min_G V(D, G).$$

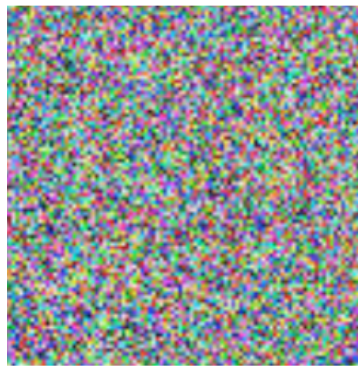


GENERATIVE ADVERSARIAL NETWORKS

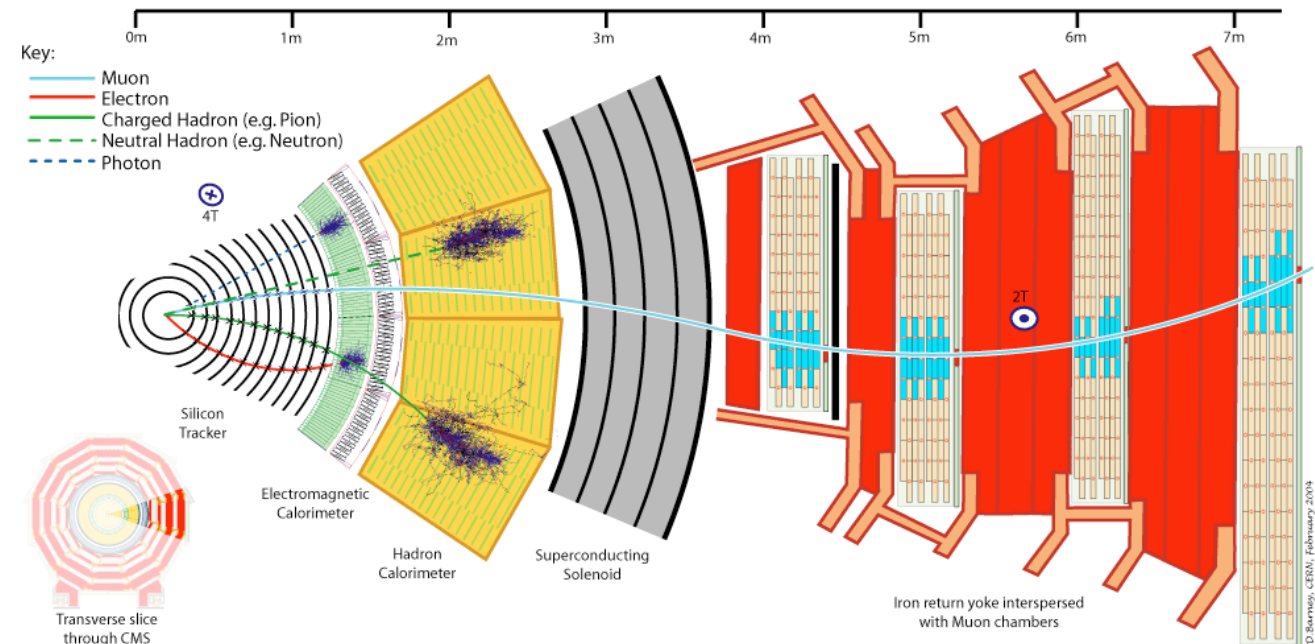
Z

X

Noise $\sim N(0,1)$



Generative Model



Leo is G



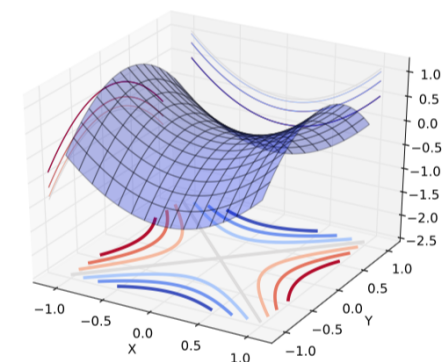
Tom is D

catch me

if you can

- We want to
 - For fixed G , find D which **maximizes** $V(D, G)$,
 - For fixed D , find G which **minimizes** $V(D, G)$;
- In other words, we are looking for the *saddle point*

$$(D^*, G^*) = \max_D \min_G V(D, G).$$



LEARNING THE SIMULATION (FOR SPEED)

CaloGAN: Simulating 3D High Energy Particle Showers in Multi-Layer Electromagnetic Calorimeters with Generative Adversarial Networks

Creating Virtual Universes Using Generative Adversarial Networks

Mustafa Mustafa^{*1}, Deborah Bard¹, Wahid Bhimji¹, Rami Al-Rfou², and Zarija Lukić¹

¹Lawrence Berkeley National Laboratory, Berkeley, CA 94720

²Google Research, Mountain View, CA 94043

Michela Paganini^{a,b}, Luke de Oliveira^a, and Benjamin Nachman^a

^aLawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA, 94720, USA

^bDepartment of Physics, Yale University, New Haven, CT 06520, USA

E-mail: michela.paganini@yale.edu, lukedeoliveira@lbl.gov, bnachman@cern.ch

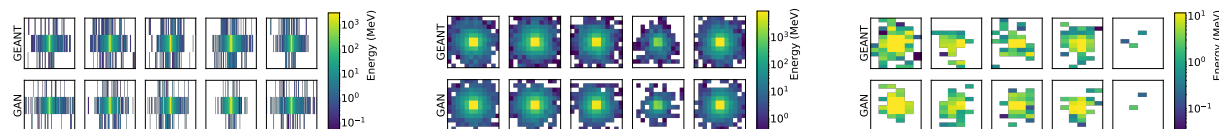


Figure 9: Five randomly selected e^+ showers per calorimeter layer from the training set (top) and the five nearest neighbors (by euclidean distance) from a set of CALOGAN candidates.

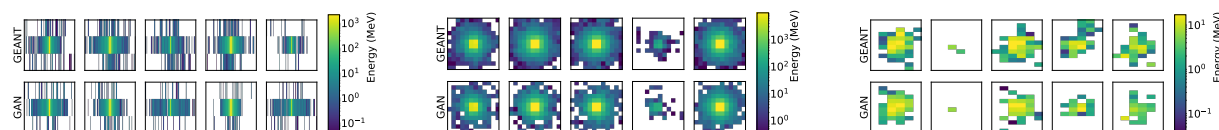


Figure 10: Five randomly selected γ showers per calorimeter layer from the training set (top) and the five nearest neighbors (by euclidean distance) from a set of CALOGAN candidates.

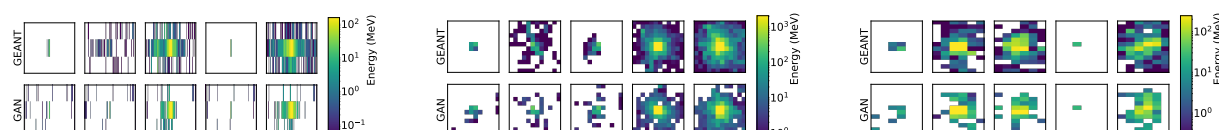
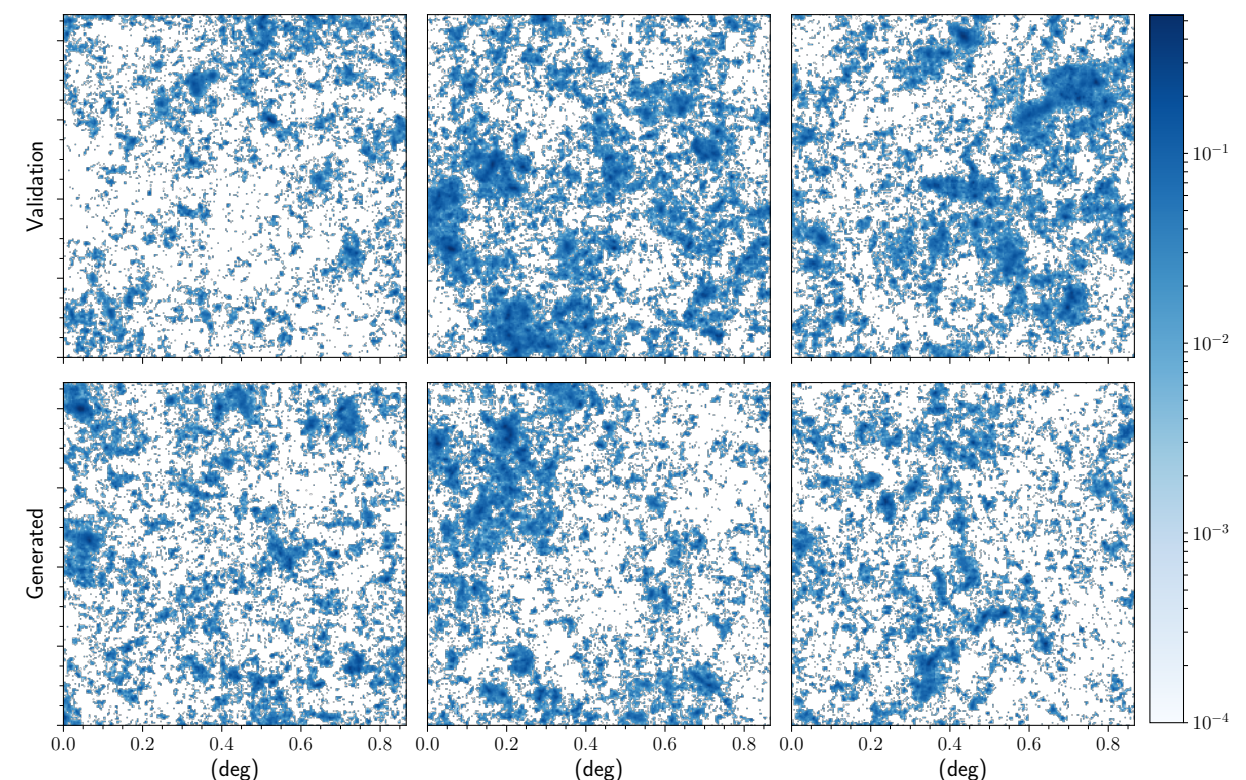


Figure 11: Five randomly selected π^+ showers per calorimeter layer from the training set (top) and the five nearest neighbors (by euclidean distance) from a set of CALOGAN candidates.



ATLAS PUB Note

ATL-SOFT-PUB-2018-001

10th July 2018



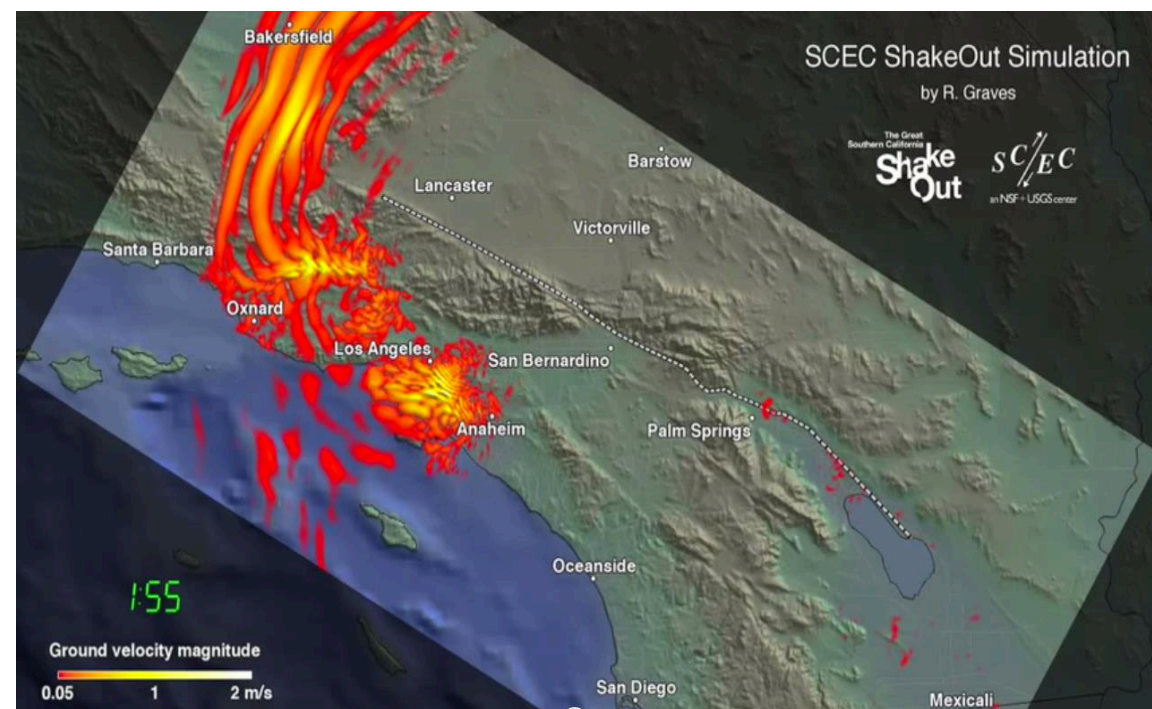
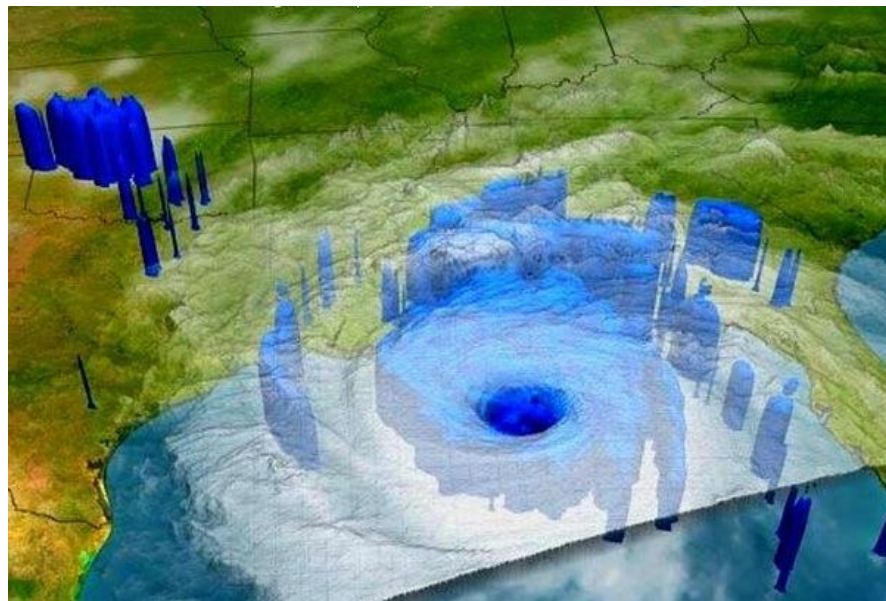
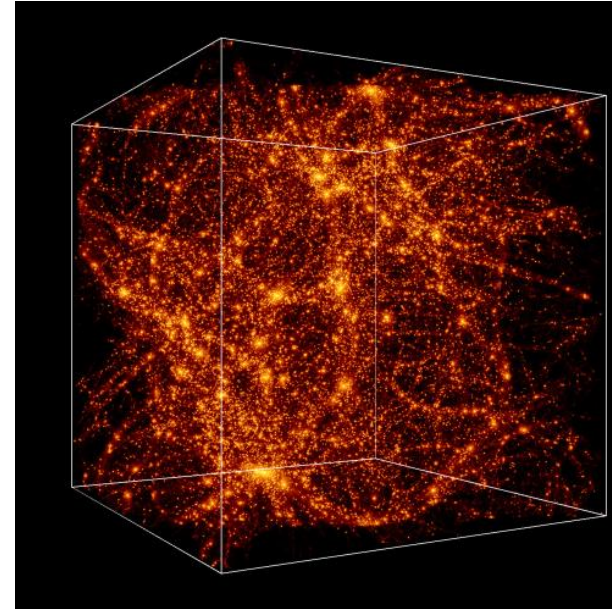
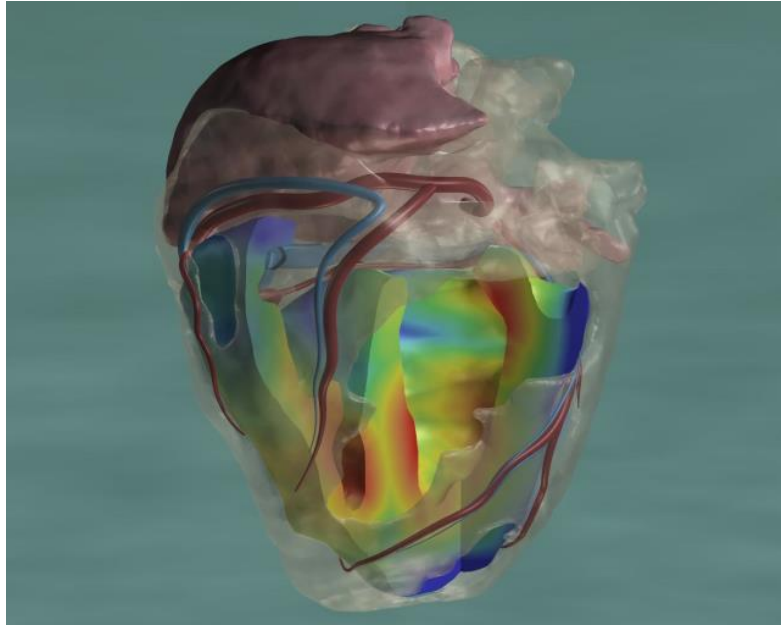
Deep generative models for fast shower simulation
in ATLAS

The bigger picture



Max
Welling

Generative Models: Simulators



WHY WE SHOULD CARE

Many areas of science have simulations based on some well-motivated mechanistic model.

However, the aggregate effect of many interactions between these low-level components leads to an intractable inverse problem.

The developments in machine learning and AI go way beyond improved classifiers and have the potential to effectively bridge the microscopic - macroscopic divide & aid in the inverse problem.

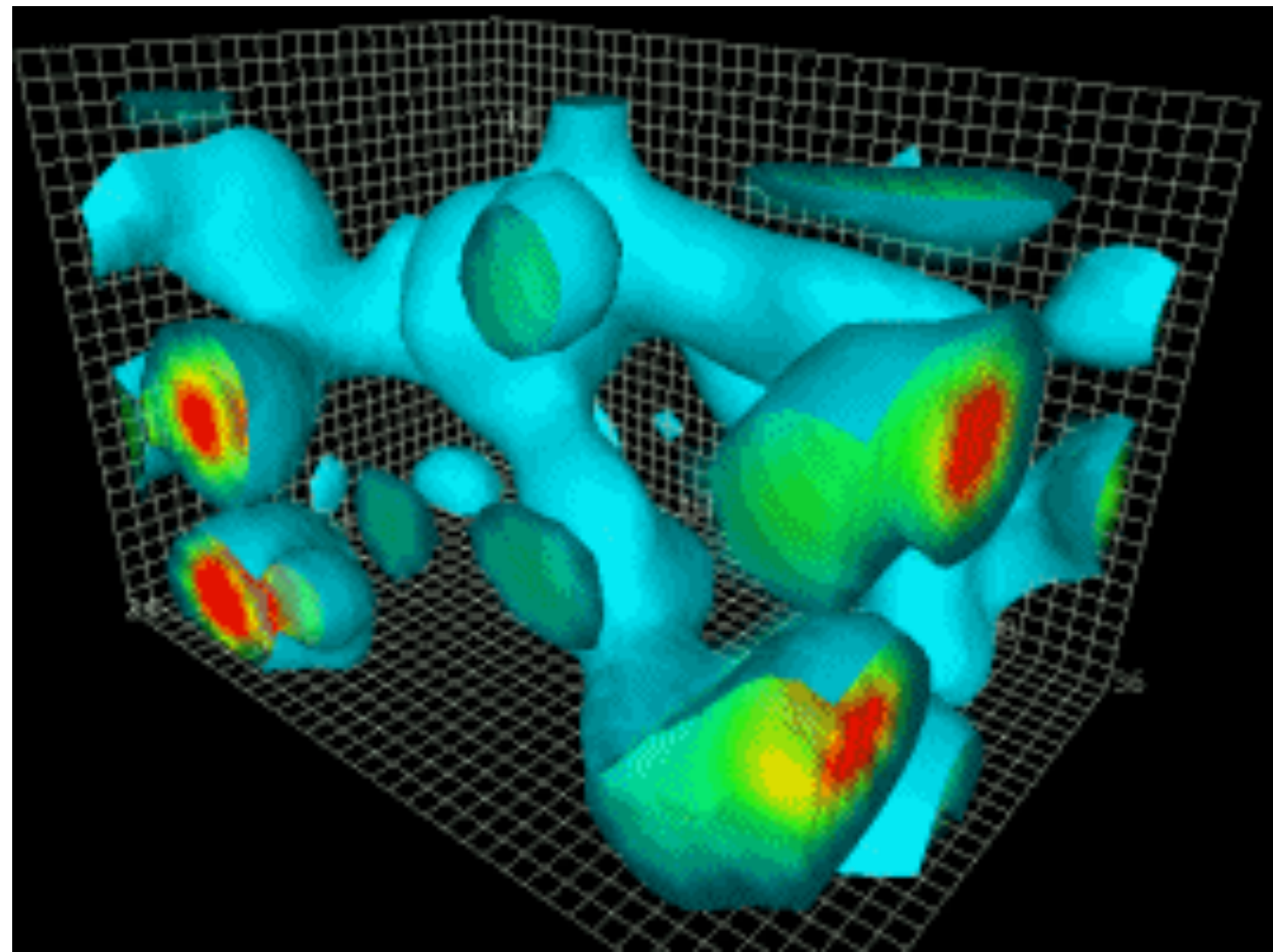
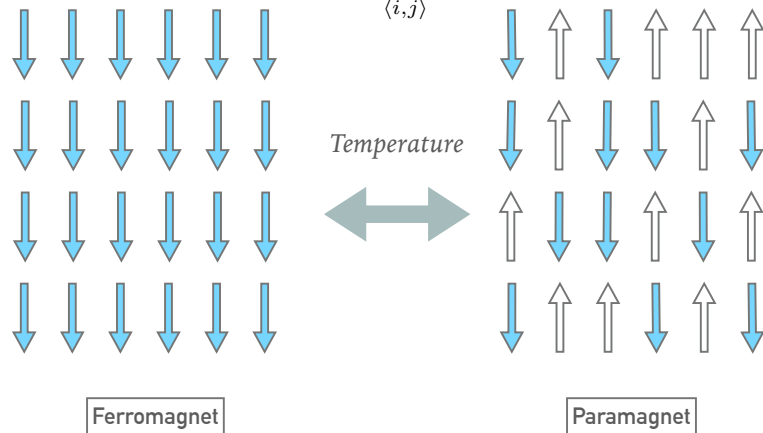
- they can provide effective statistical models for macroscopic phenomena that are tied back to the low-level microscopic (reductionist) model
- generative models and likelihood-free inference are two particularly exciting areas

LATTICE FIELD THEORY

PHASES, PHASE TRANSITIONS, AND THE ORDER PARAMETER

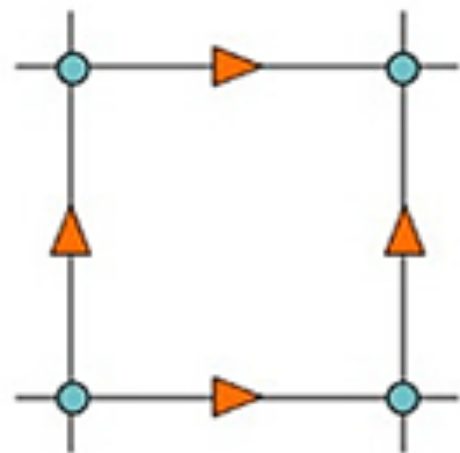
Ising ferromagnet in two dimensions

$$E = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$$

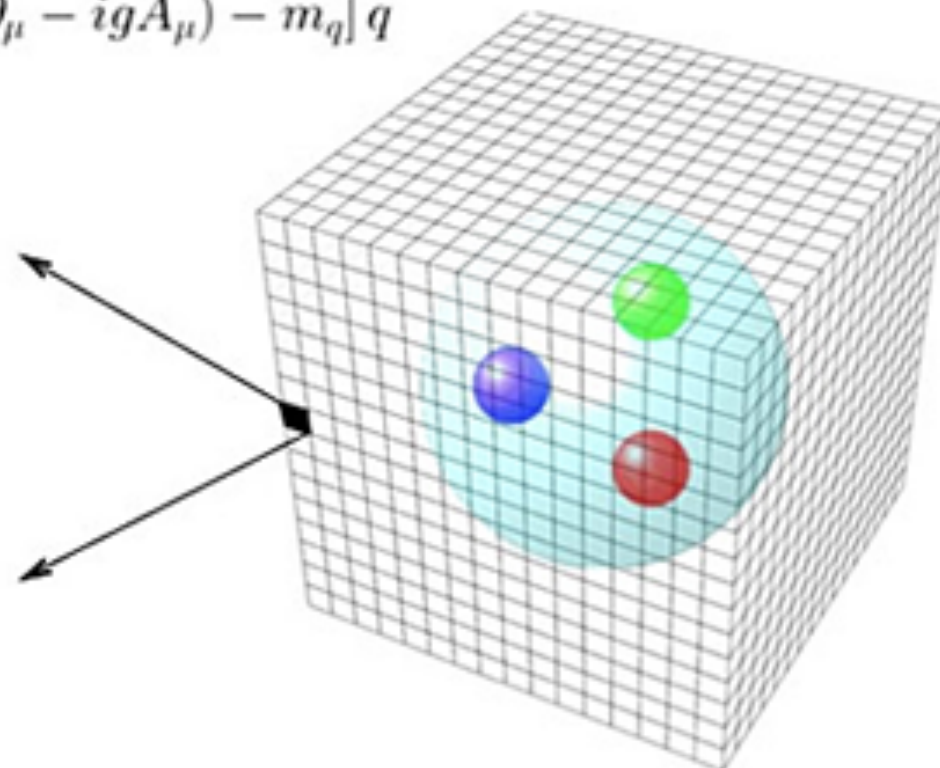


QCD Lagrangian

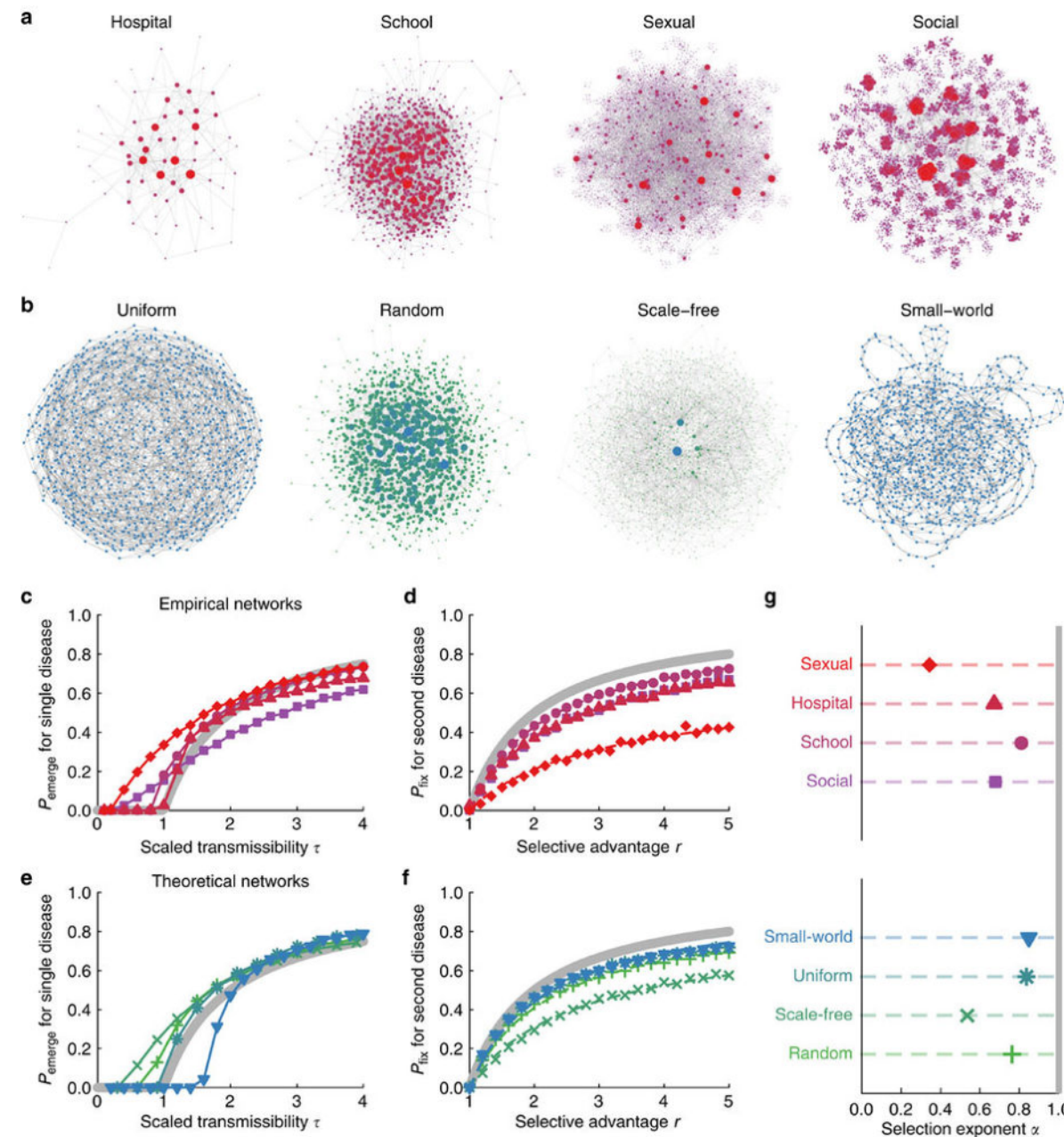
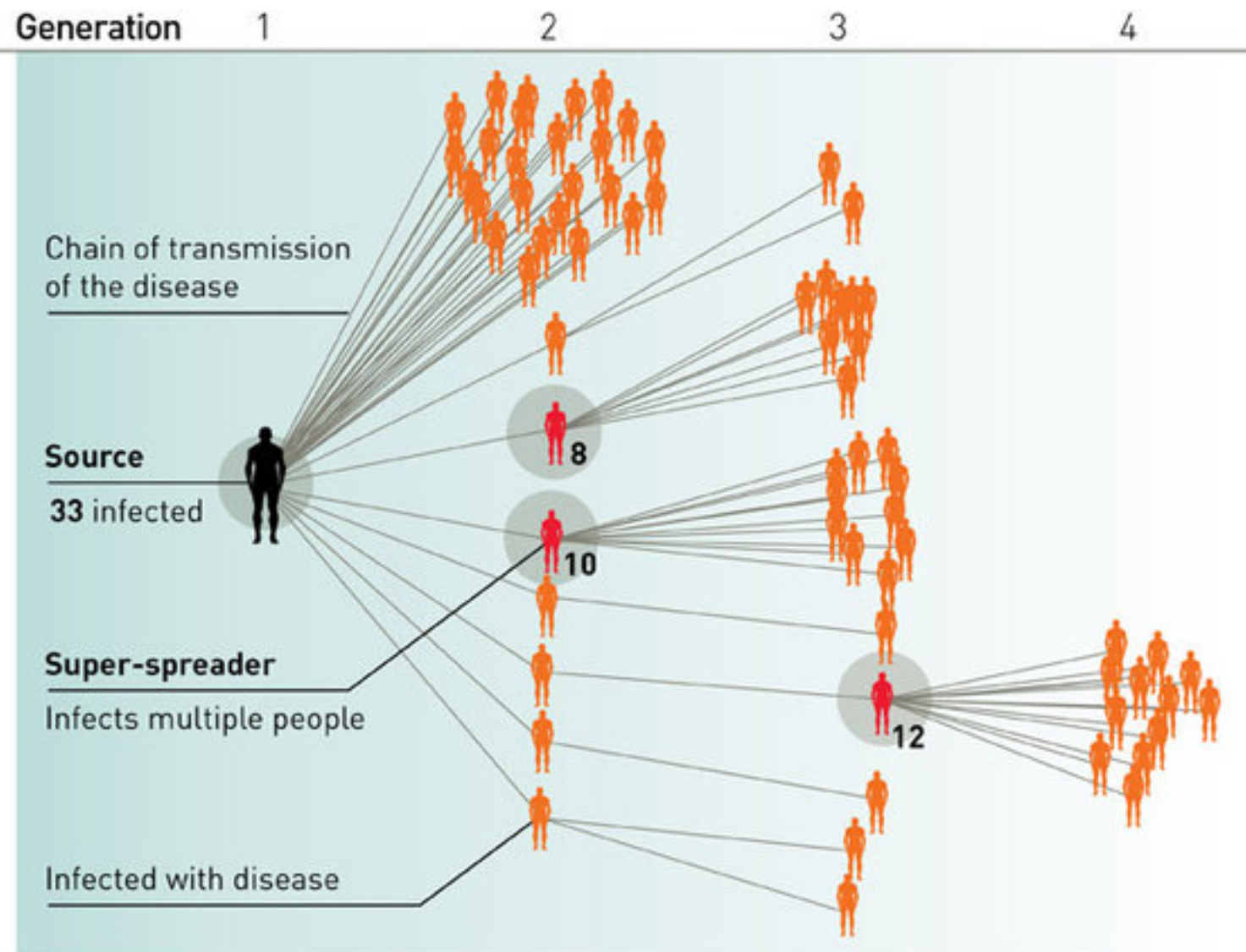
$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \sum_{q=u,d,s,c,b,t} \bar{q} [i\gamma^\mu (\partial_\mu - igA_\mu) - m_q] q$$



● quark ▲ gluon



EPIDEMIOLOGY & POPULATION GENETICS



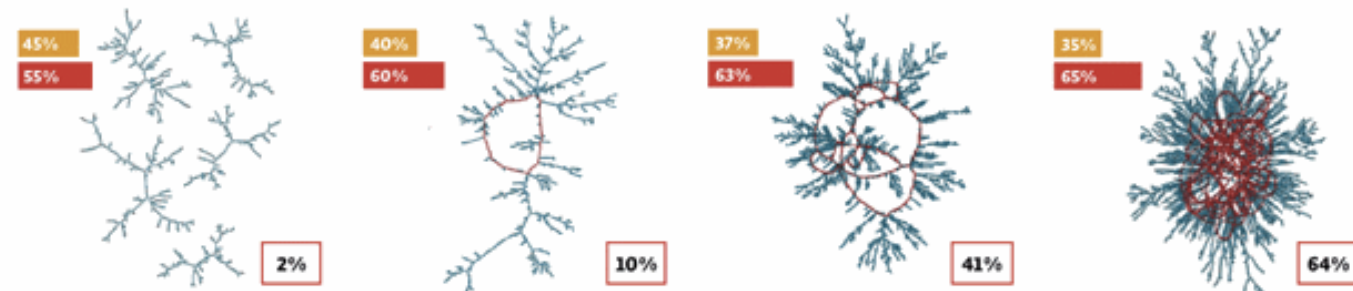
Small Change, Big Effects

KEY
1 partner
2 or 3 partners

Percent of people that are connected in the network through their sexual partnerships

Modest variations in the concurrency rate—the proportion of people in overlapping sexual partnerships—can have a dramatic effect on a population's vulnerability to HIV.

When the concurrency rate is 55%, only 2% of this population is connected to the broader sexual network required for HIV transmission (top). But when concurrency reaches 65%, an astonishing 64% of the population is vulnerable, even though the number of sexual partners remains constant.



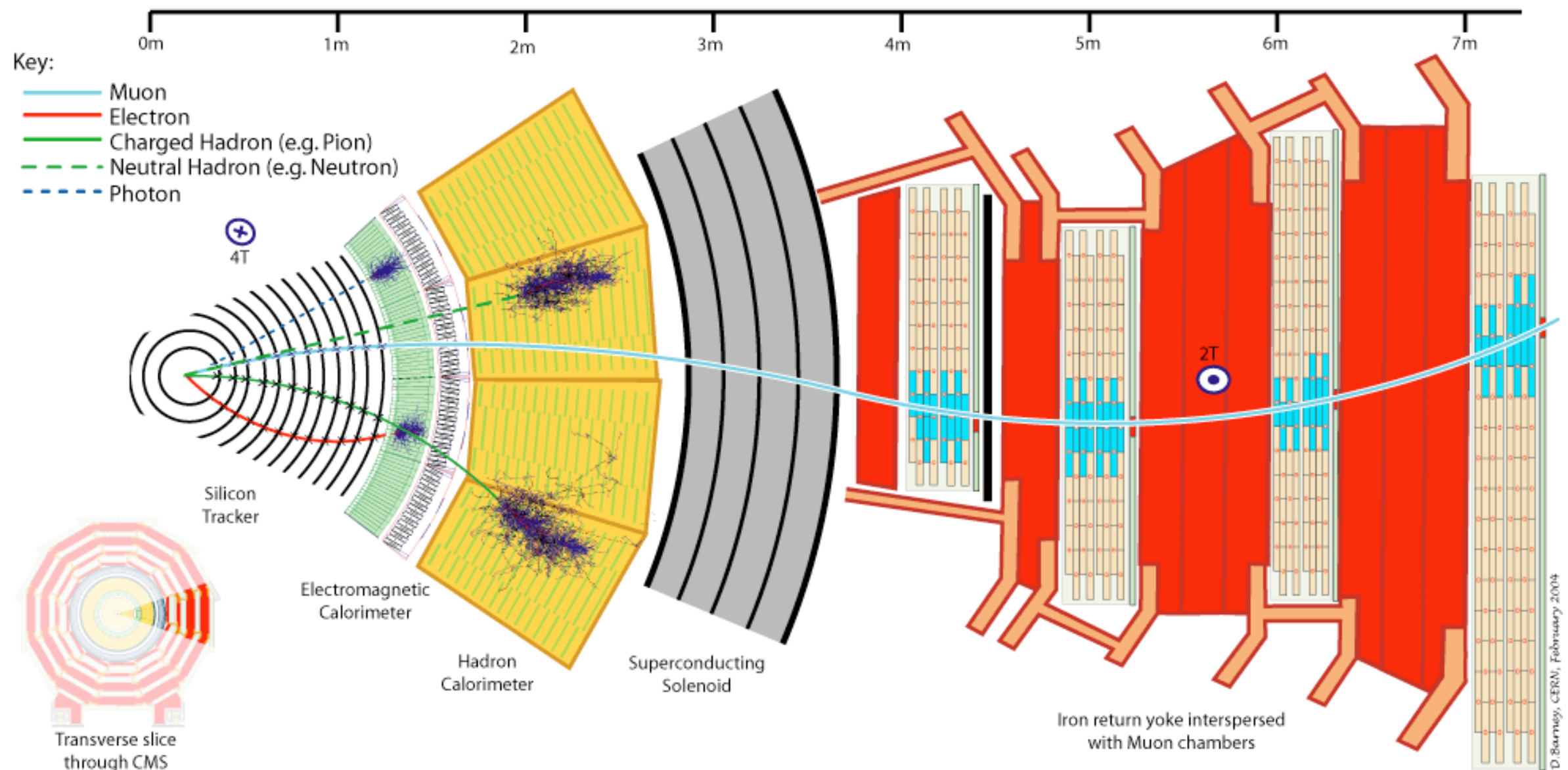
Source: Morris, et al. The Relationship Between Concurrent Partnerships and HIV Transmission, 2008. See www.aidsstar-one.com/.

DETECTOR SIMULATION

Conceptually: $\text{Prob}(\text{detector response} \mid \text{particles})$

Implementation: Monte Carlo integration over micro-physics

Consequence: evaluation of the likelihood is intractable



DETECTOR SIMULATION

Conceptually: $\text{Prob}(\text{detector response} \mid \text{particles})$

Implementation: Monte Carlo integration over micro-physics

Consequence: evaluation of the likelihood is intractable

This motivates a new class of algorithms for what is called **likelihood-free inference**, which only require ability to generate samples from the simulation in the “forward mode”

A COMMON THEME

ABC

resources on approximate
Bayesian computational
methods

 Search

Home

Home

This website keeps track of developments in approximate Bayesian computation (ABC) (a.k.a. likelihood-free), a class of computational statistical methods for Bayesian inference under intractable likelihoods. The site is meant to be a resource both for biologists and statisticians who want to learn more about ABC and related methods. Recent publications are under Publications 2012. A comprehensive list of publications can be found under Literature. If you are unfamiliar with ABC methods see the Introduction. Navigate using the menu to learn more.

[ABC in Montreal](#)

[ABC in Montreal \(2014\)](#)

ABC in Montreal

Approximate Bayesian computation (ABC) or likelihood-free (LF) methods have developed mostly beyond the radar of the machine learning community, but are important tools for a large and diverse segment of the scientific community. This is particularly true for systems and population biology, computational neuroscience, computer vision, healthcare sciences, but also many others.

Interaction between the ABC and machine learning community has recently started and contributed to important advances. In general, however, there is still significant room for more intense interaction and collaboration. Our workshop aims at being a place for this to happen.

ICML 2017 Workshop on Implicit Models

Workshop Aims

Probabilistic models are an important tool in machine learning. They form the basis for models that generate realistic data, uncover hidden structure, and make predictions. Traditionally, probabilistic models in machine learning have focused on prescribed models. Prescribed models specify a joint density over observed and hidden variables that can be easily evaluated. The requirement of a tractable density simplifies their learning but limits their flexibility --- several real world phenomena are better described by simulators that do not admit a tractable density. Probabilistic models defined only via the simulations they produce are called implicit models.

Arguably starting with generative adversarial networks, research on implicit models in machine learning has exploded in recent years. This workshop's aim is to foster a discussion around the recent developments and future directions of implicit models.

Implicit models have many applications. They are used in ecology where models simulate animal populations over time; they are used in phylogeny, where simulations produce hypothetical ancestry trees; they are used in physics to generate particle simulations for high energy processes. Recently, implicit models have been used to improve the state-of-the-art in image and content generation. Part of the workshop's focus is to discuss the commonalities among applications of implicit models.

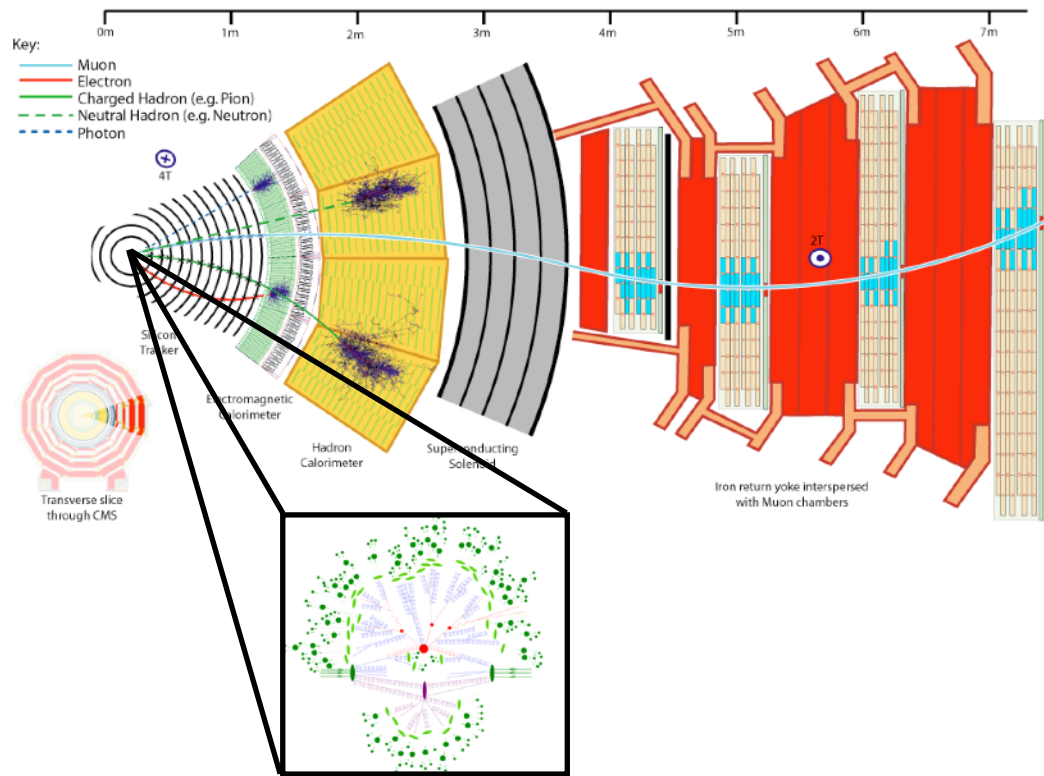
Of particular interest at this workshop is to unite fields that work on implicit models. For example:

- **Generative adversarial networks** (a NIPS 2016 workshop) are implicit models with an adversarial training scheme.
- Recent advances in **variational inference** (a NIPS 2015 and 2016 workshop) have leveraged implicit models for more accurate approximations.
- **Approximate Bayesian computation** (a NIPS 2015 workshop) focuses on posterior inference for models with implicit likelihoods.
- Learning implicit models is deeply connected to **two sample testing, density ratio and density difference** estimation.

We hope to bring together these different views on implicit models, identifying their core challenges and combining their innovations.

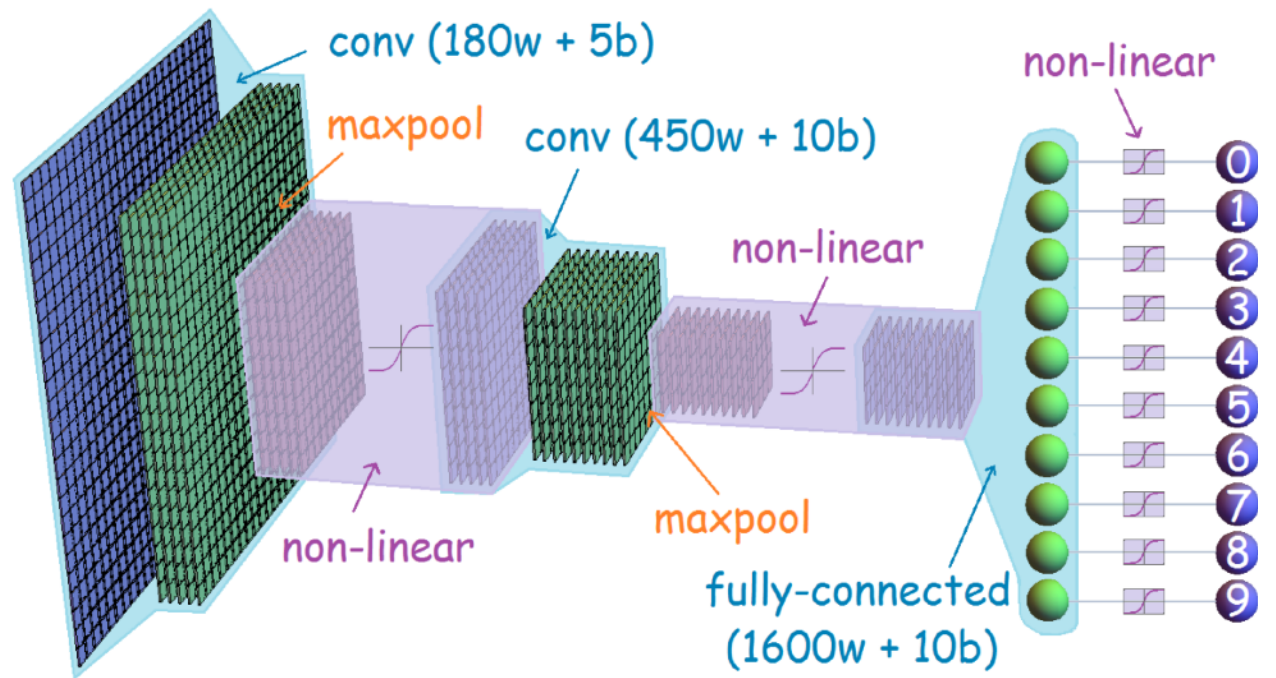
APPROACHES TO LIKELIHOOD-FREE INFERENCE

Use simulator
(much more efficiently)



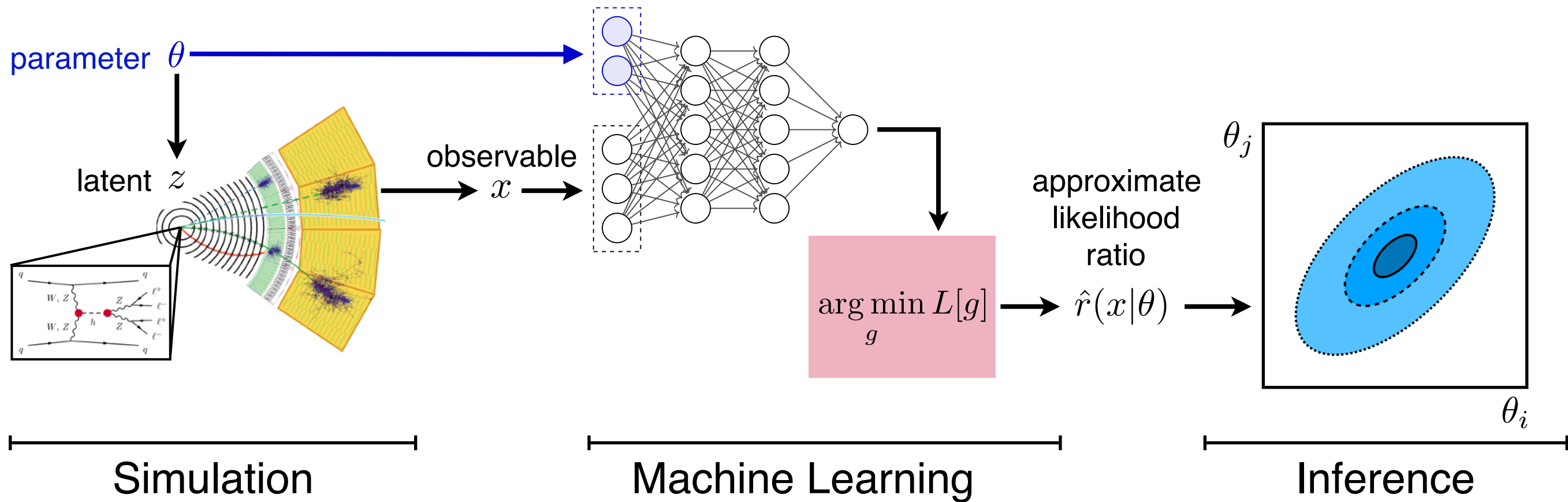
- Approximate Bayesian Computation (ABC)
- Probabilistic Programming
- Adversarial Variational Optimization (AVO)

Learn simulator
(with deep learning)

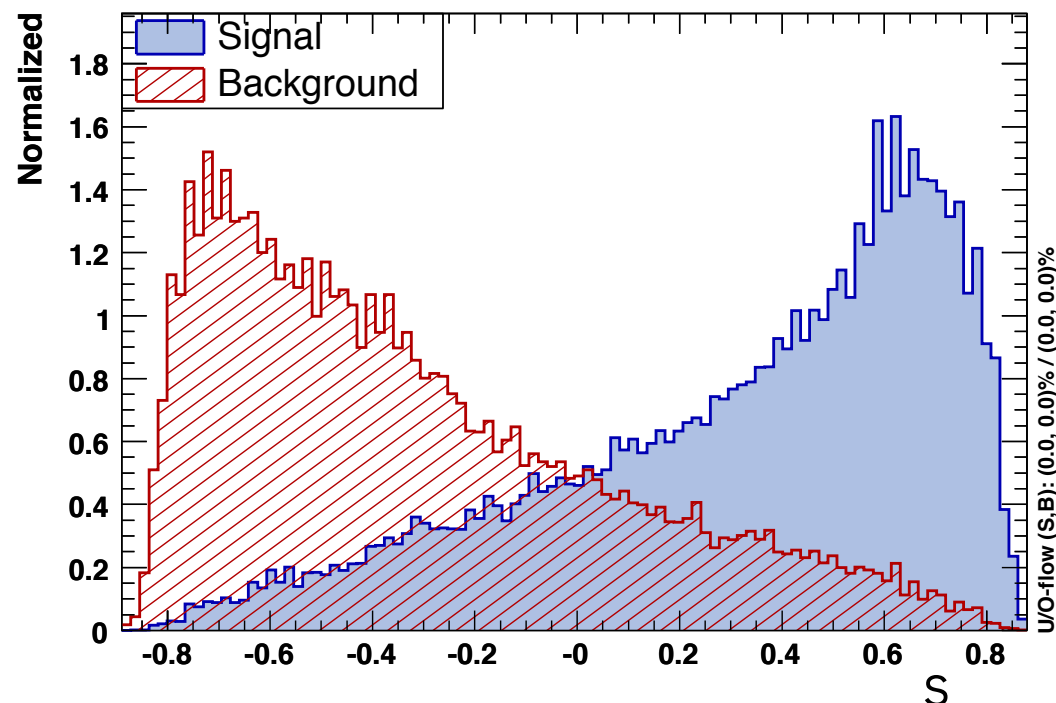
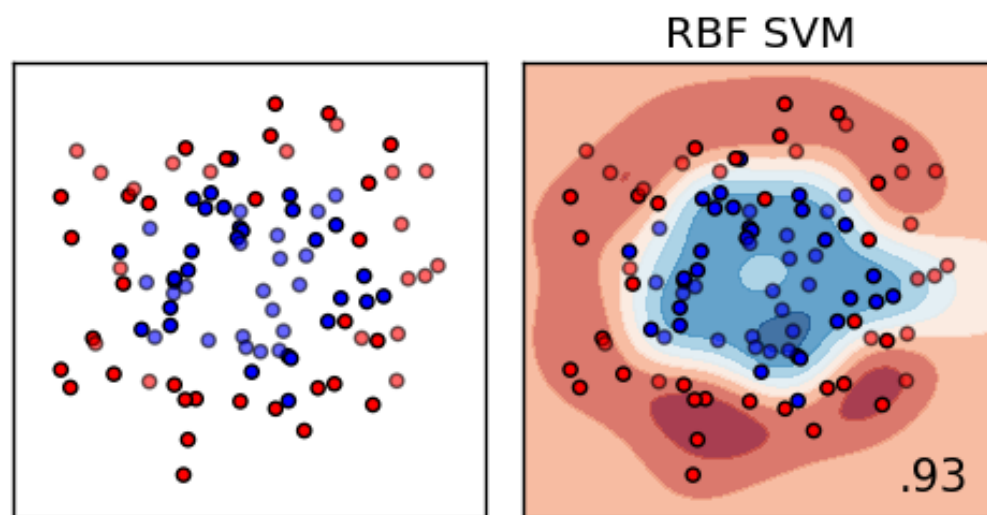
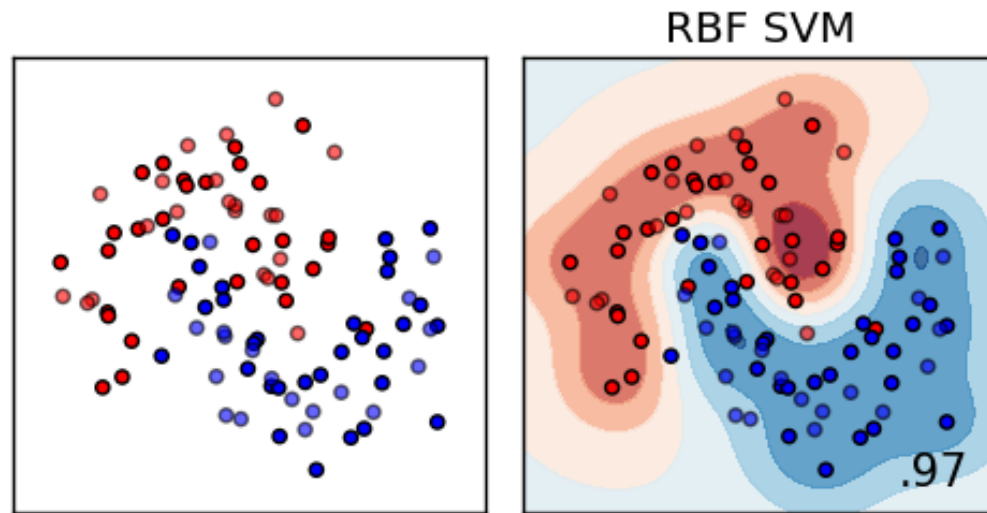


- Generative Adversarial Networks (GANs), Variational Auto-Encoders (VAE)
- Likelihood ratio from classifiers (CARL)
- Autoregressive models, Normalizing Flows

LEARNING FROM SIMULATED DATA



LIKELIHOOD RATIO TRICK



- **binary classifier**: find function $s(x)$ that minimizes **loss**:

$$L[s] = \int p(x|H_0) (0 - s(x))^2 dx + \int p(x|H_1) (1 - s(x))^2 dx$$

$$\approx \frac{1}{N} \sum_{i=1}^N (y_i - s(x_i))^2$$

- i.e. approximate the optimal classifier

$$s(x) = \frac{p(x|H_1)}{p(x|H_0) + p(x|H_1)}$$

- which is 1-to-1 with the likelihood ratio

$$\frac{p(x|H_1)}{p(x|H_0)}$$

EXTENDING THE LIKELIHOOD RATIO TRICK

A binary classifier approximates

$$s(x) = \frac{p(x|H_1)}{p(x|H_0) + p(x|H_1)}$$

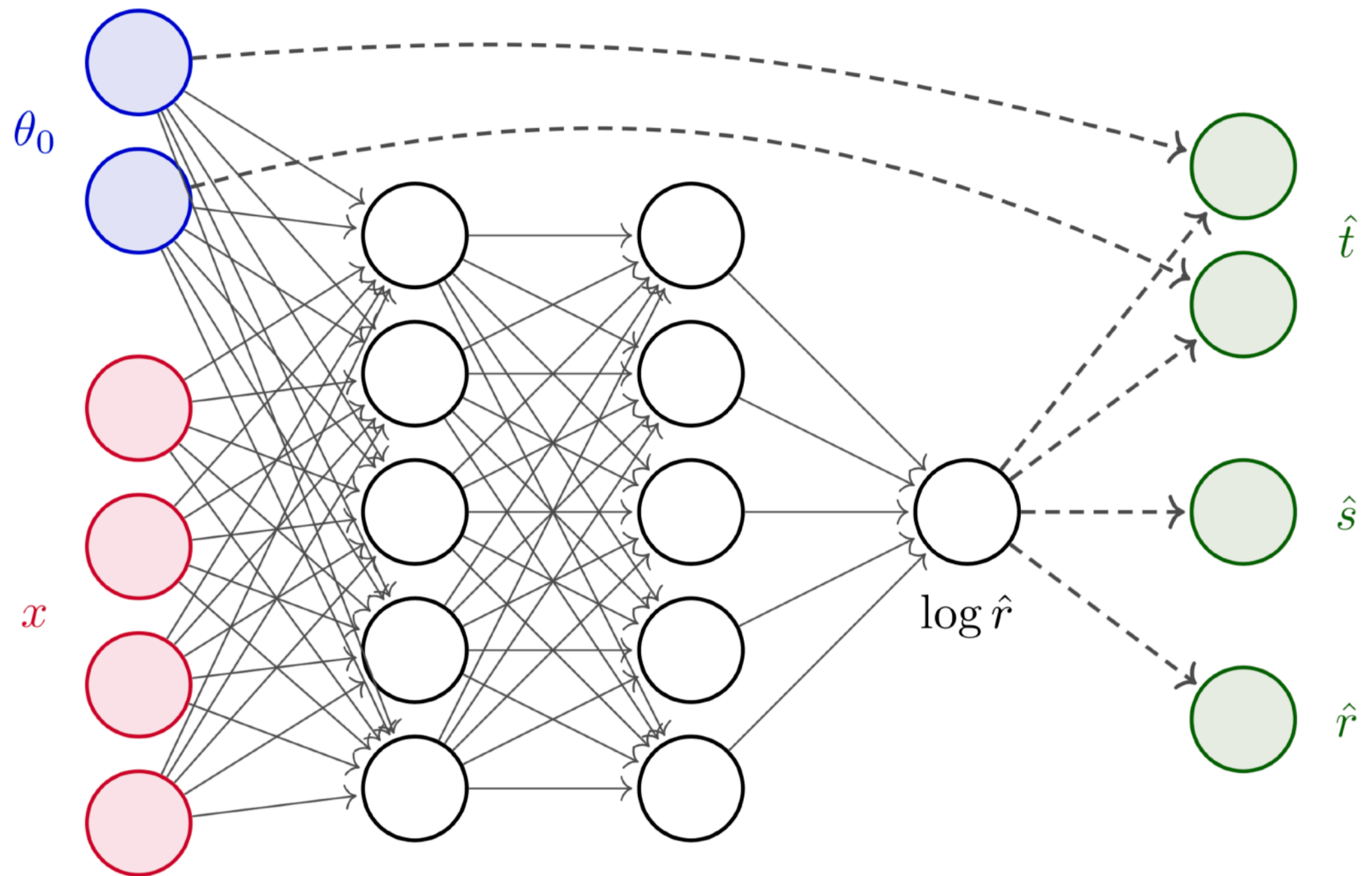
Which is one-to-one with the likelihood ratio

$$\frac{p(x|H_1)}{p(x|H_0)} = 1 - \frac{1}{s(x)}$$

Can do the same thing for any two points θ_0 & θ_1 in parameter space Θ . I call this a **parametrized classifier**

$$s(x; \theta_0, \theta_1) = \frac{p(x|\theta_1)}{p(x|\theta_0) + p(x|\theta_1)}$$

PARAMETERIZED CLASSIFIERS



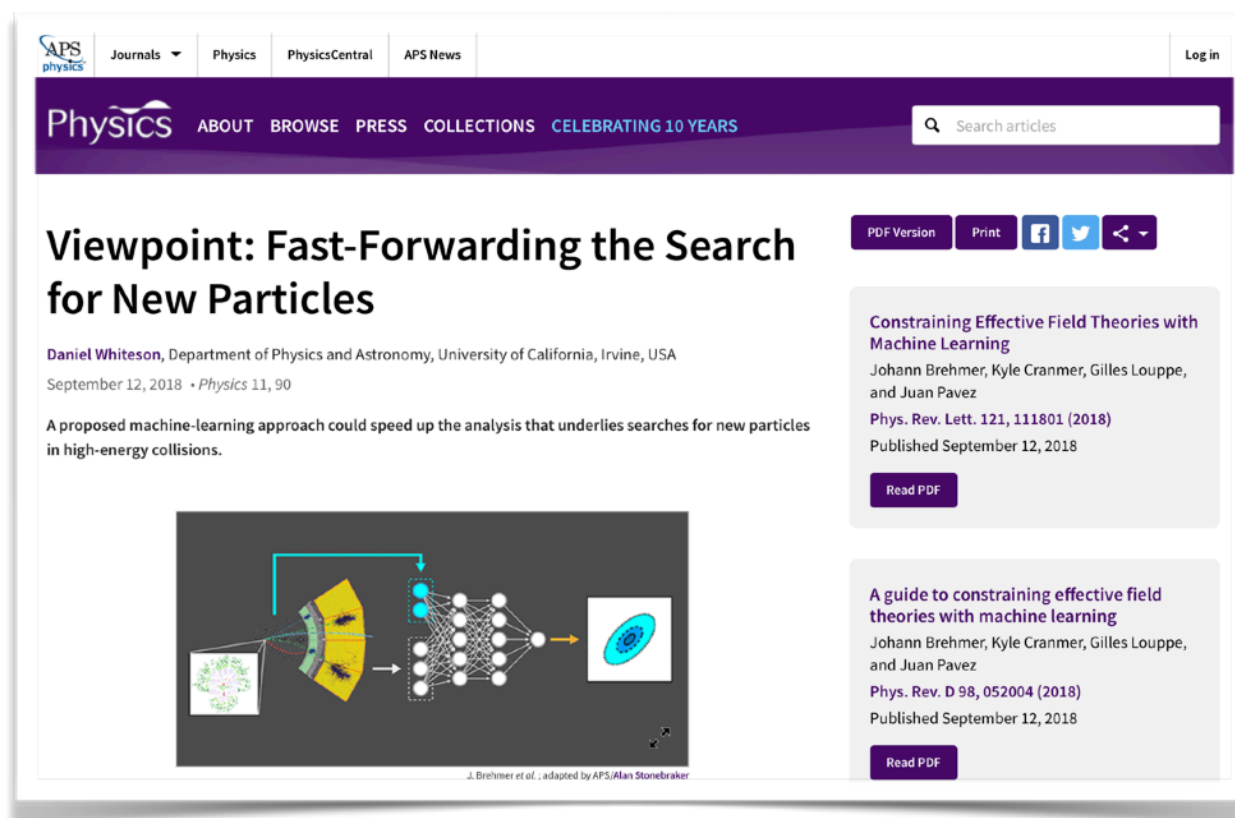
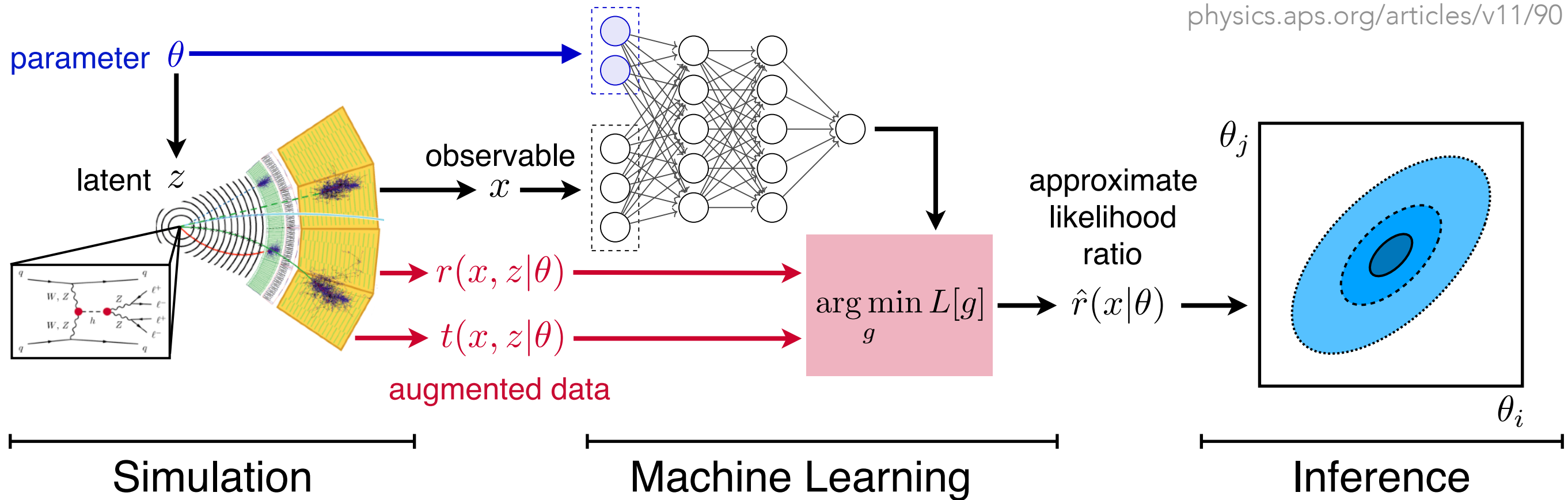
LEARNING WITH AUGMENTED DATA

arXiv:1805.12244

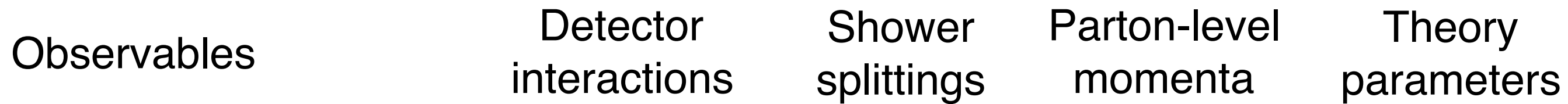
PRL, arXiv:1805.00013

PRD, arXiv:1805.00020

physics.aps.org/articles/v11/90



MINING GOLD FROM THE SIMULATION



$$x \longleftarrow z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$$

$$p(x|\theta) = \int dz_d \int dz_s \int dz_p p(x|z_d)$$

$$p(z_d|z_s)$$

$$p(z_s|z_p)$$

$$p(z_p|\theta)$$

This parton-level likelihood can be evaluated!

⇒ We can calculate the “joint” likelihood ratio conditional on a specific evolution:

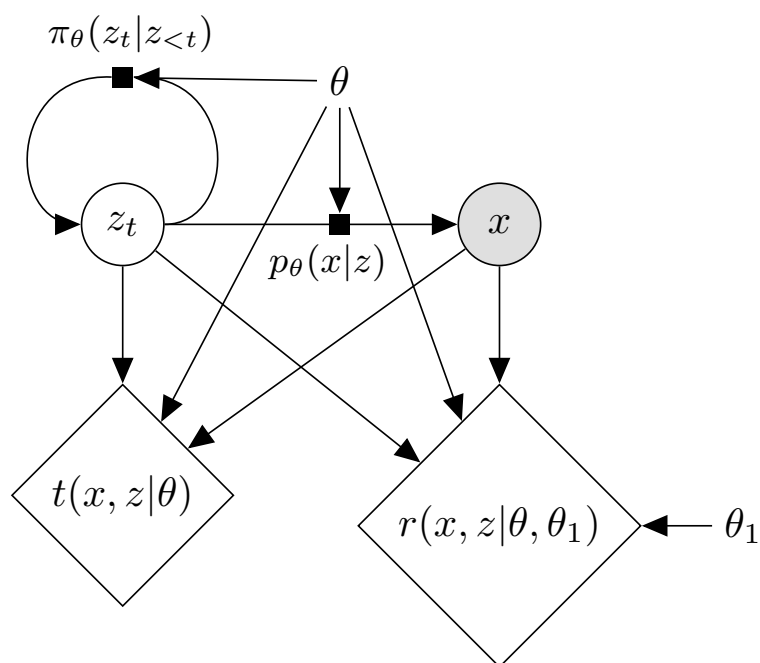
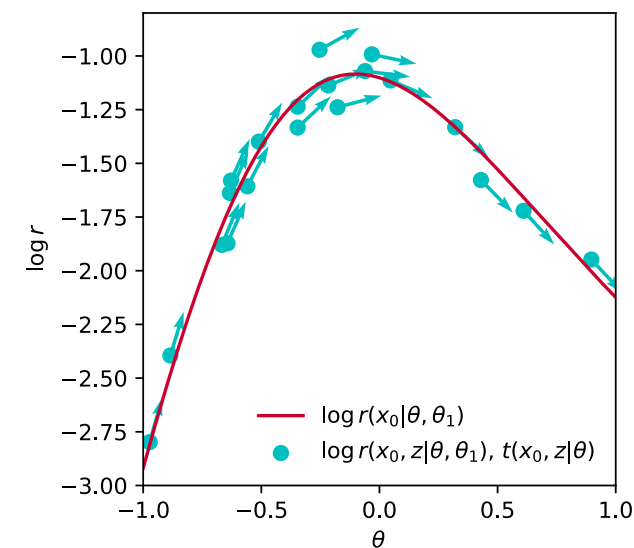
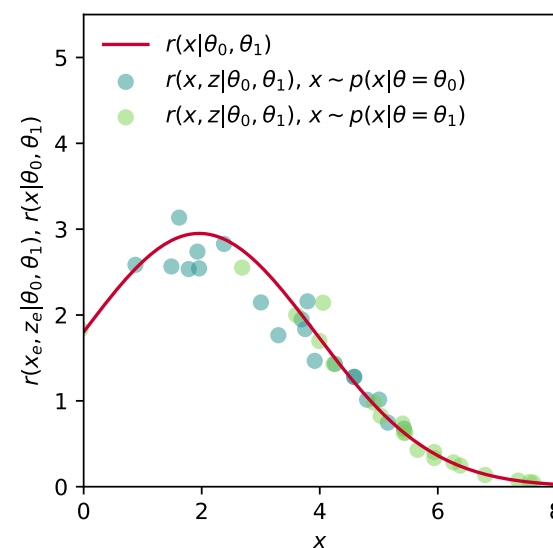
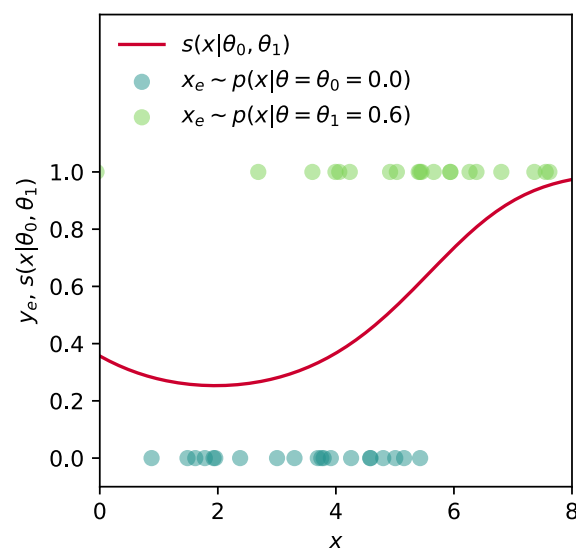
$$r(x, z|\theta_0, \theta_1) \equiv \frac{p(x, z_d, z_s, z_p|\theta_0)}{p(x, z_d, z_s, z_p|\theta_1)} = \frac{p(x|z_d)}{p(x|z_d)} \frac{p(z_d|z_s)}{p(z_d|z_s)} \frac{p(z_s|z_p)}{p(z_s|z_p)} \frac{p(z_p|\theta_0)}{p(z_p|\theta_1)}$$

(“How much more likely is this specific evolution assuming θ_0 compared to θ_1 ?”)

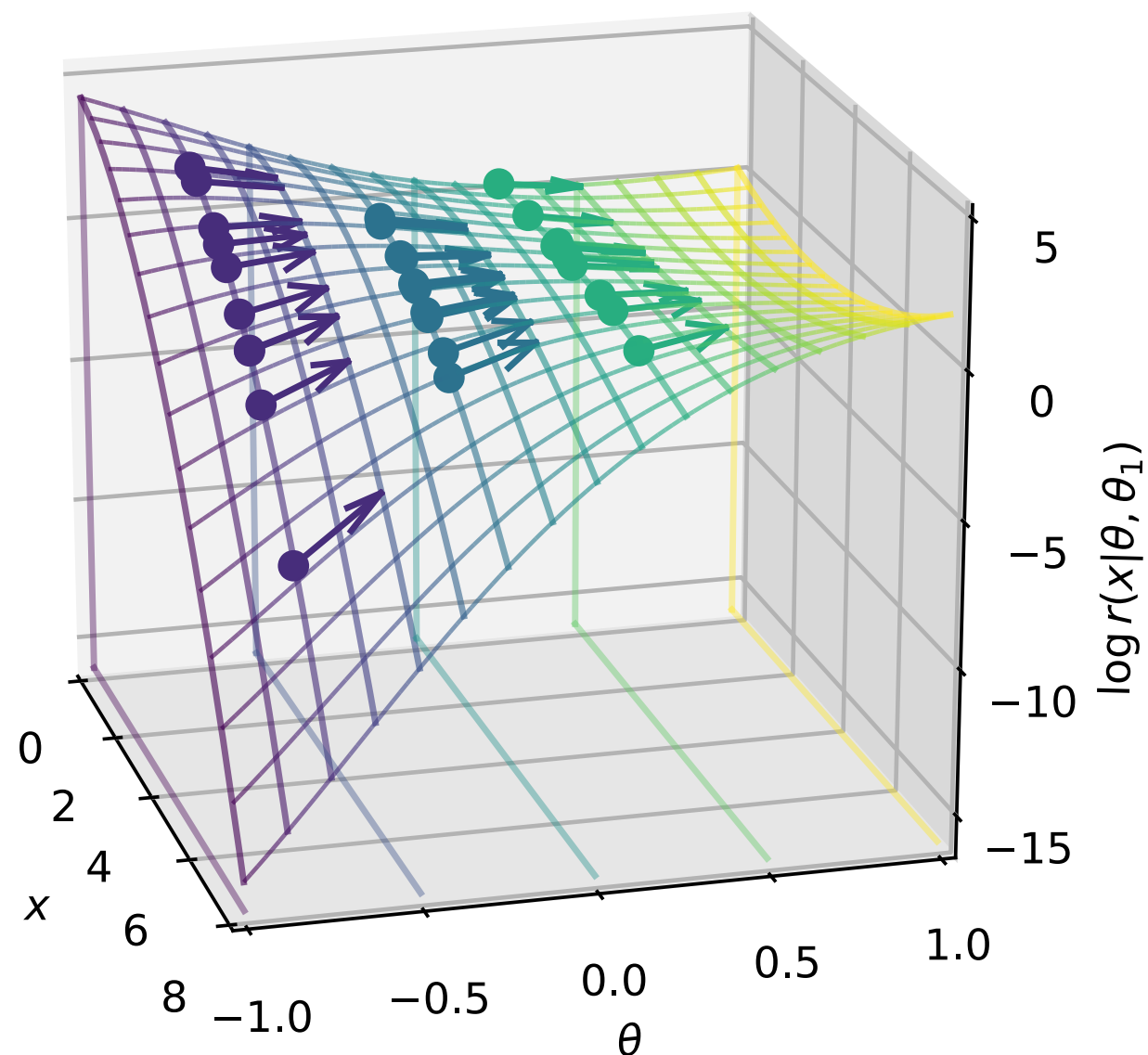
⇒ Similarly, we can calculate the joint score

$$t(x, z|\theta_0) \equiv \nabla_{\theta} \log p(x, z_d, z_s, z|\theta)$$

PUTTING IT ALL TOGETHER



can think of simulator
as policy π_θ in language of
reinforcement learning



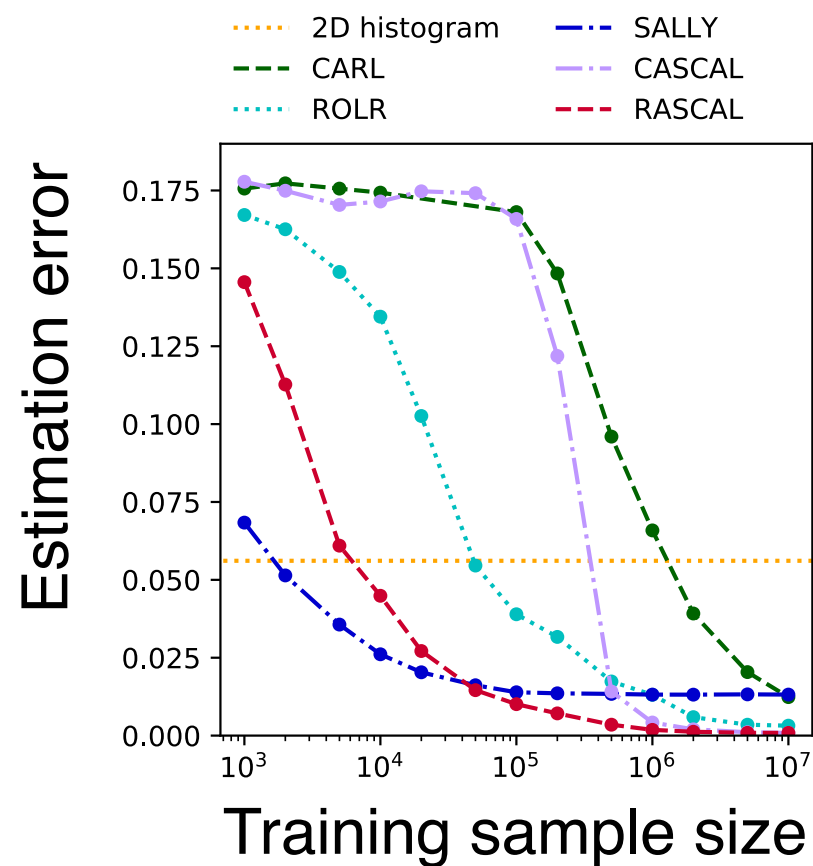
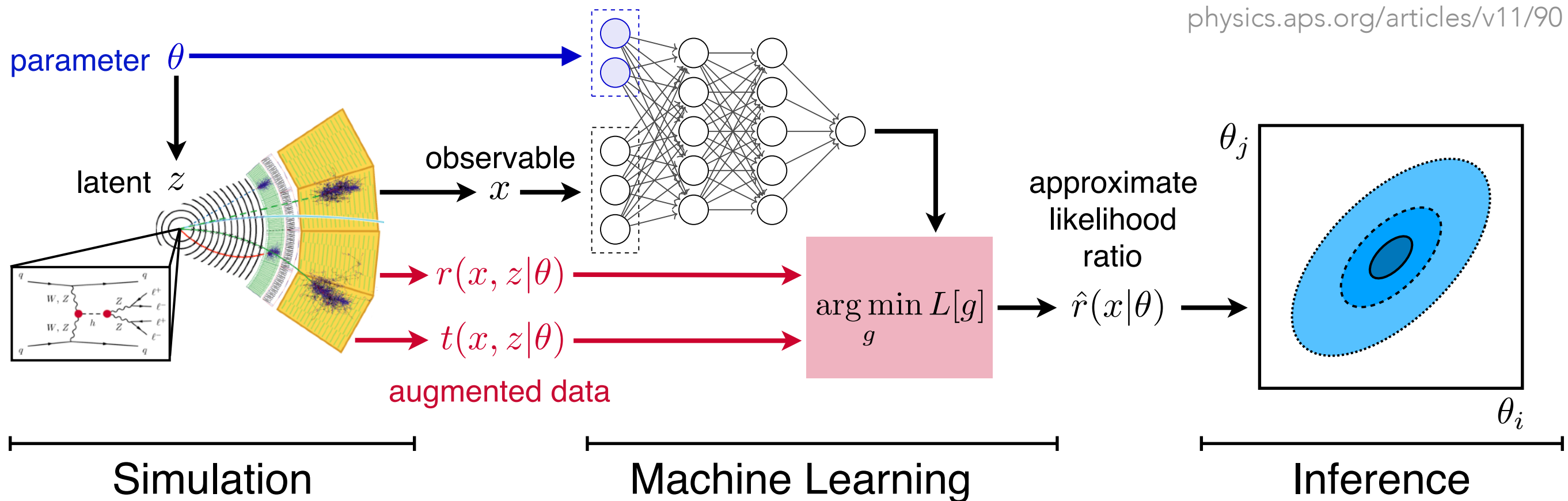
LEARNING WITH AUGMENTED DATA

arXiv:1805.12244

PRL, arXiv:1805.00013

PRD, arXiv:1805.00020

physics.aps.org/articles/v11/90



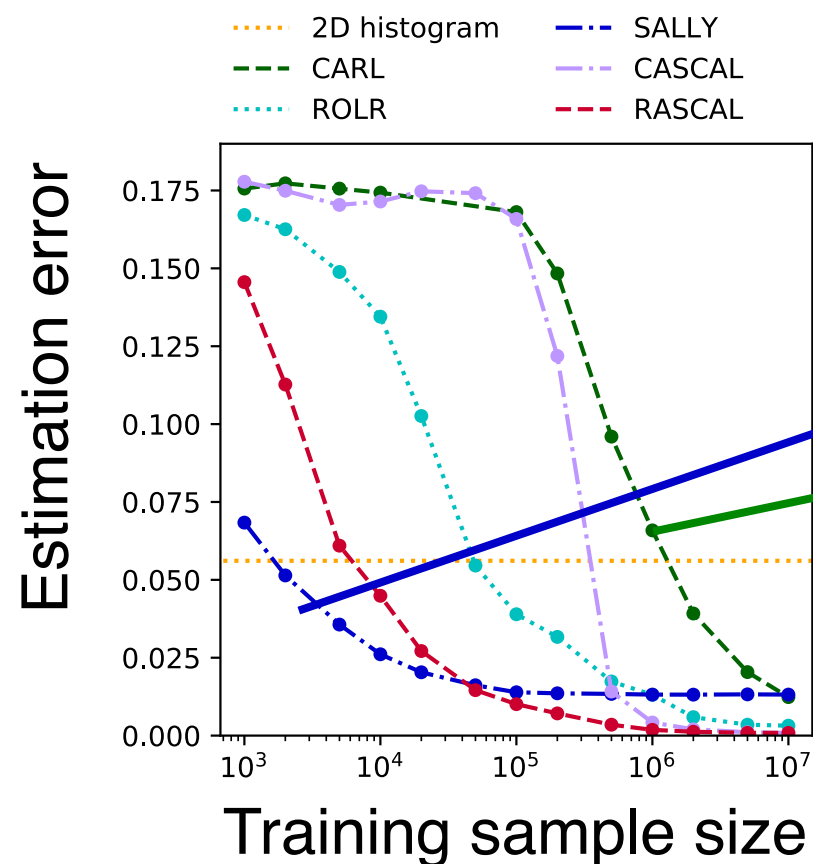
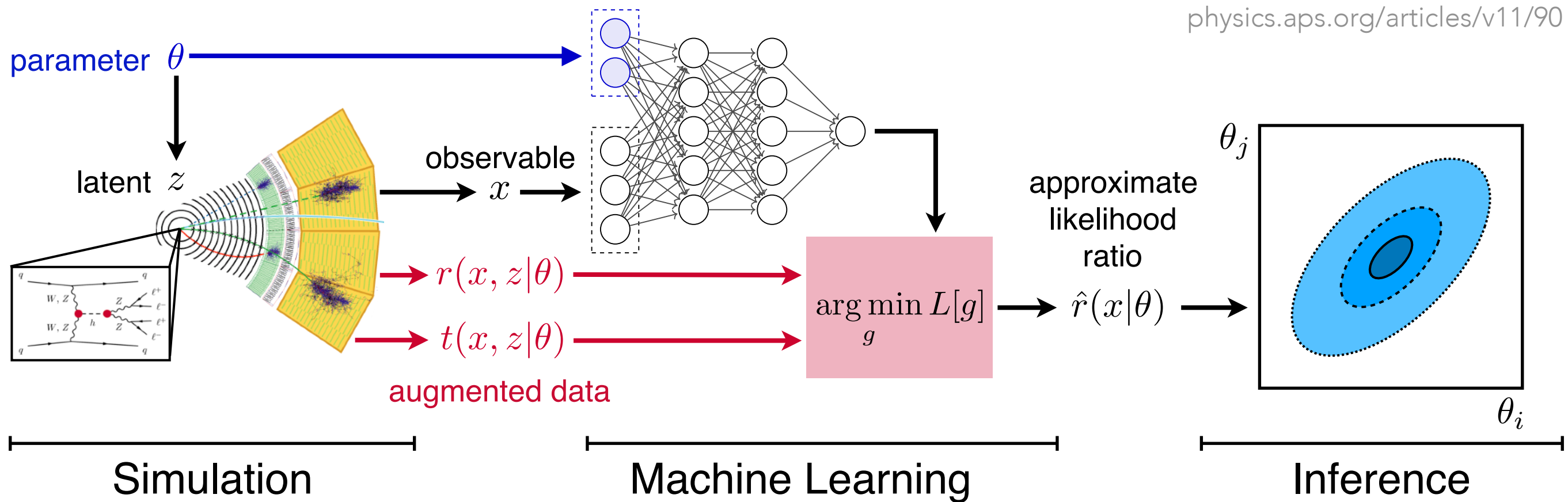
LEARNING WITH AUGMENTED DATA

arXiv:1805.12244

PRL, arXiv:1805.00013

PRD, arXiv:1805.00020

physics.aps.org/articles/v11/90

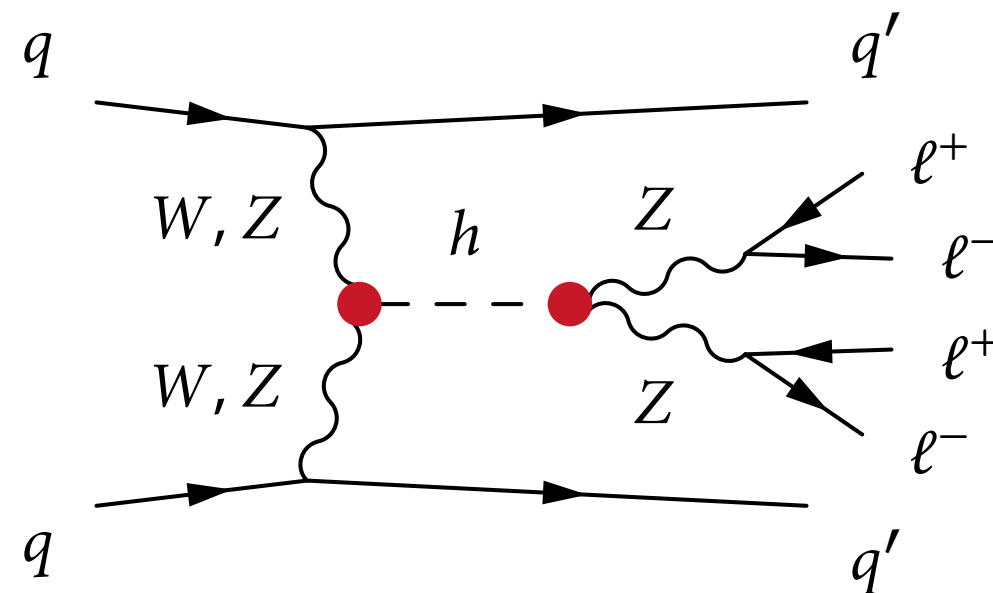
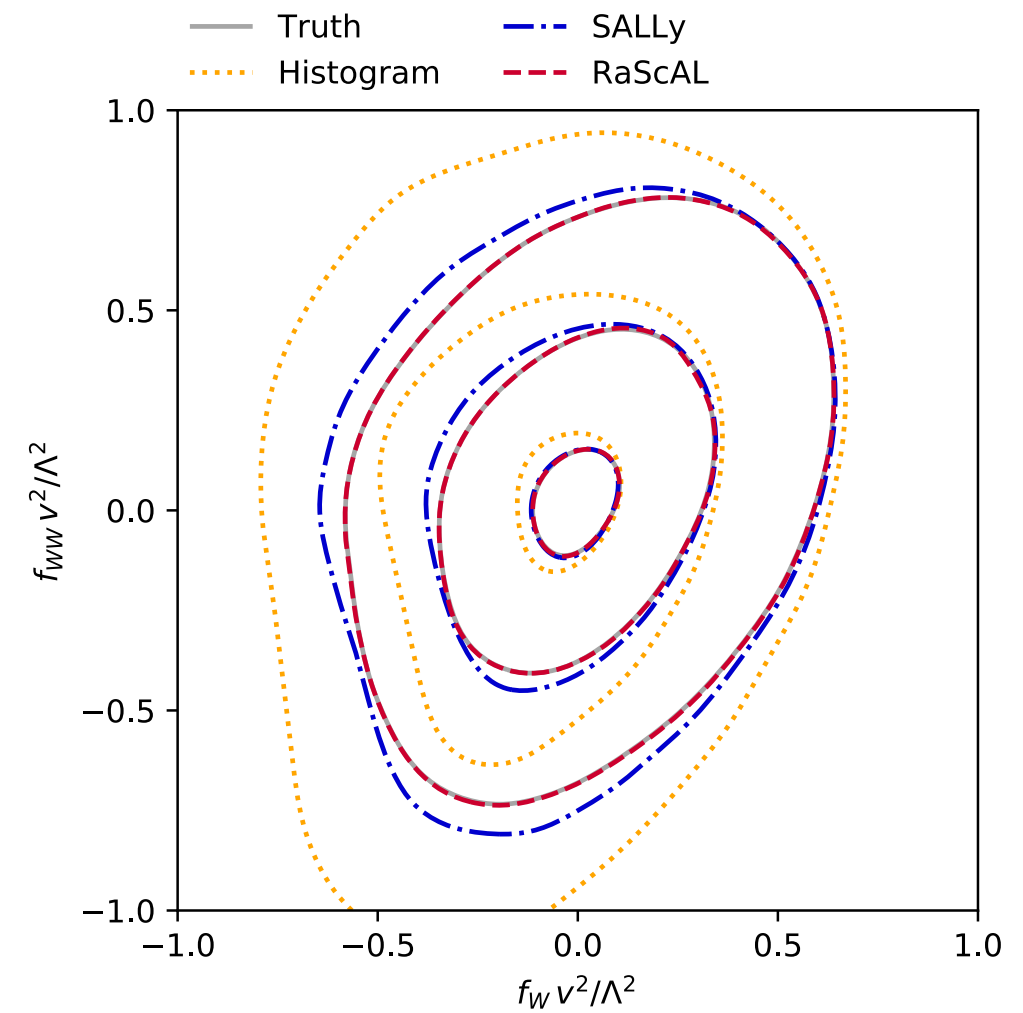
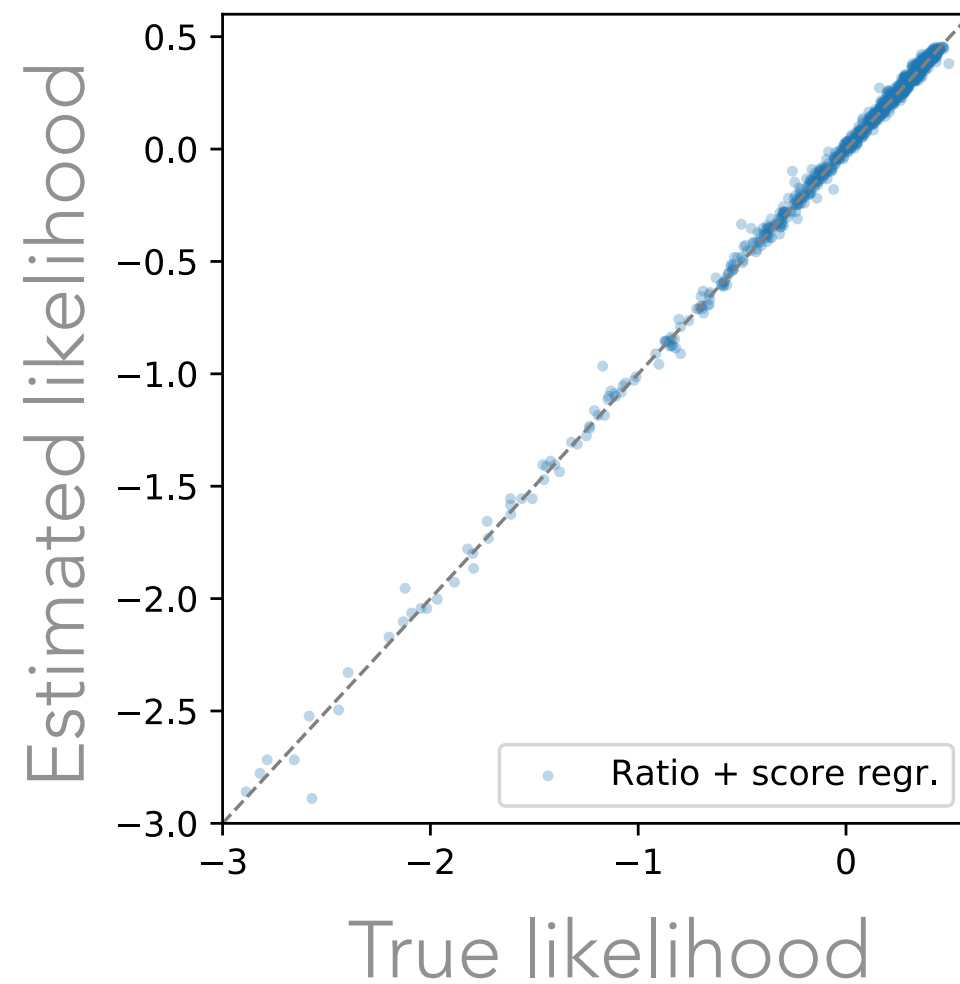


New techniques
require less data than
without augmented data

Traditional Histogram

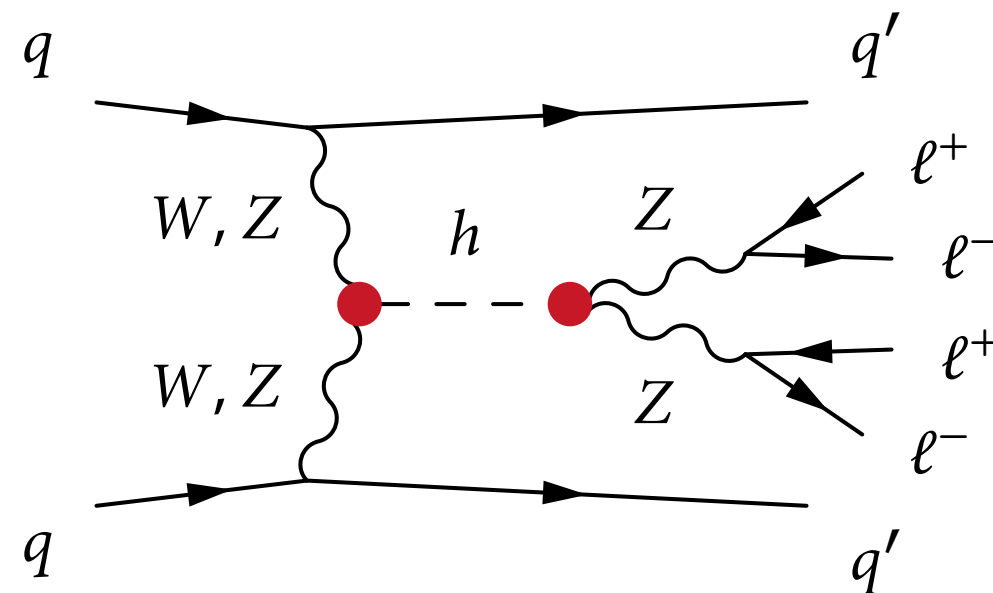
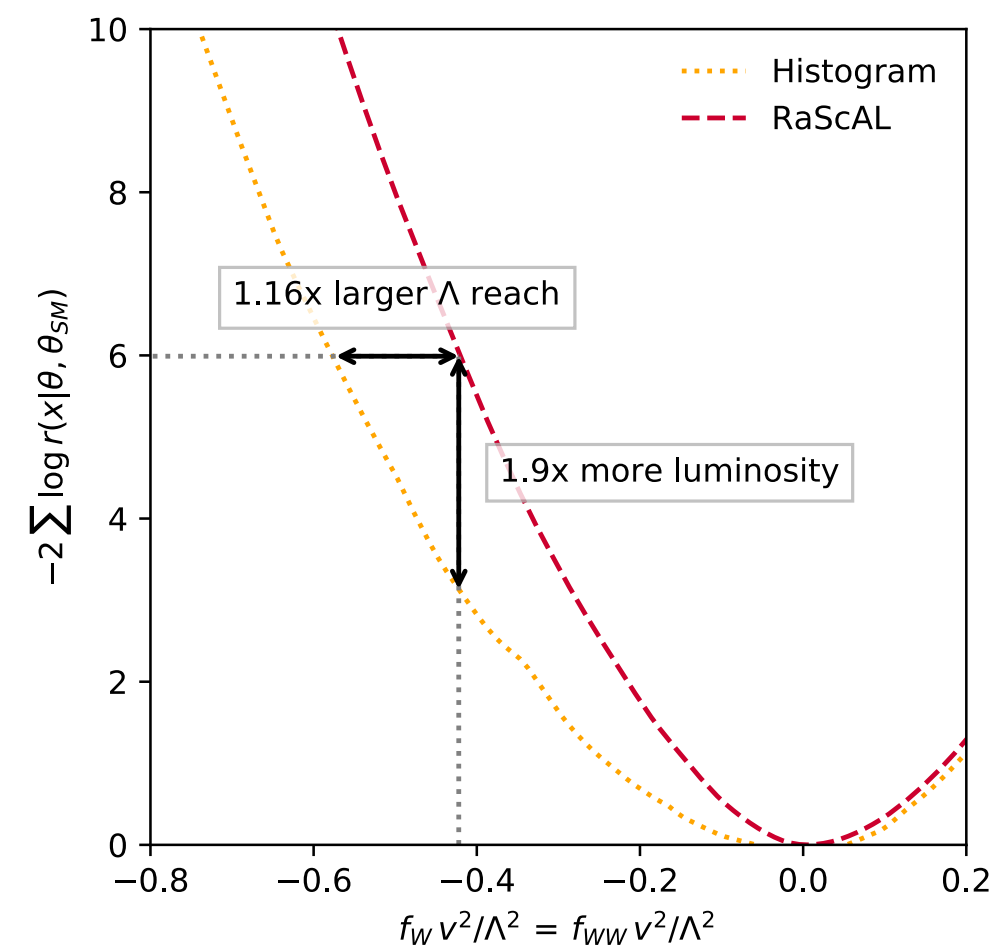
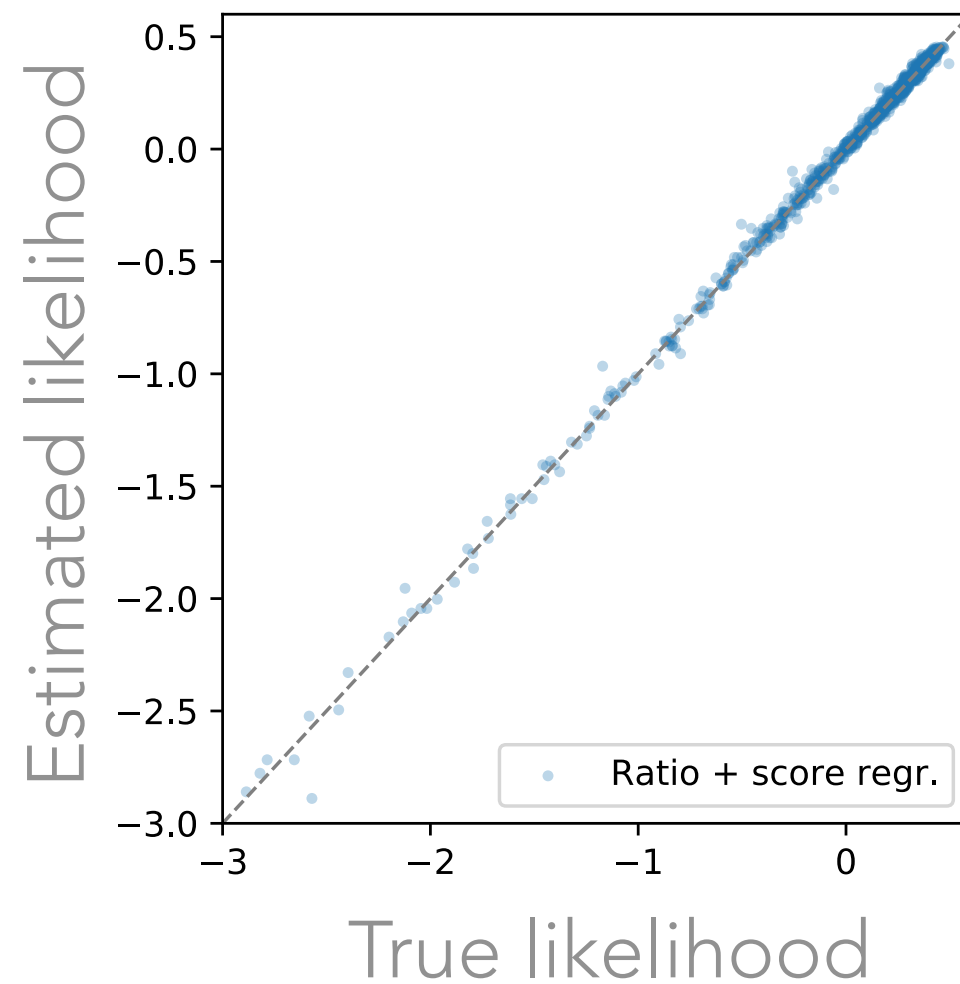
MACHINE LEARNING THE HIGGS EFFECTIVE FIELD THEORY

(based on a 42-Dim observation $\mathbf{x}$)



MACHINE LEARNING THE HIGGS EFFECTIVE FIELD THEORY

(based on a 42-Dim observation x)



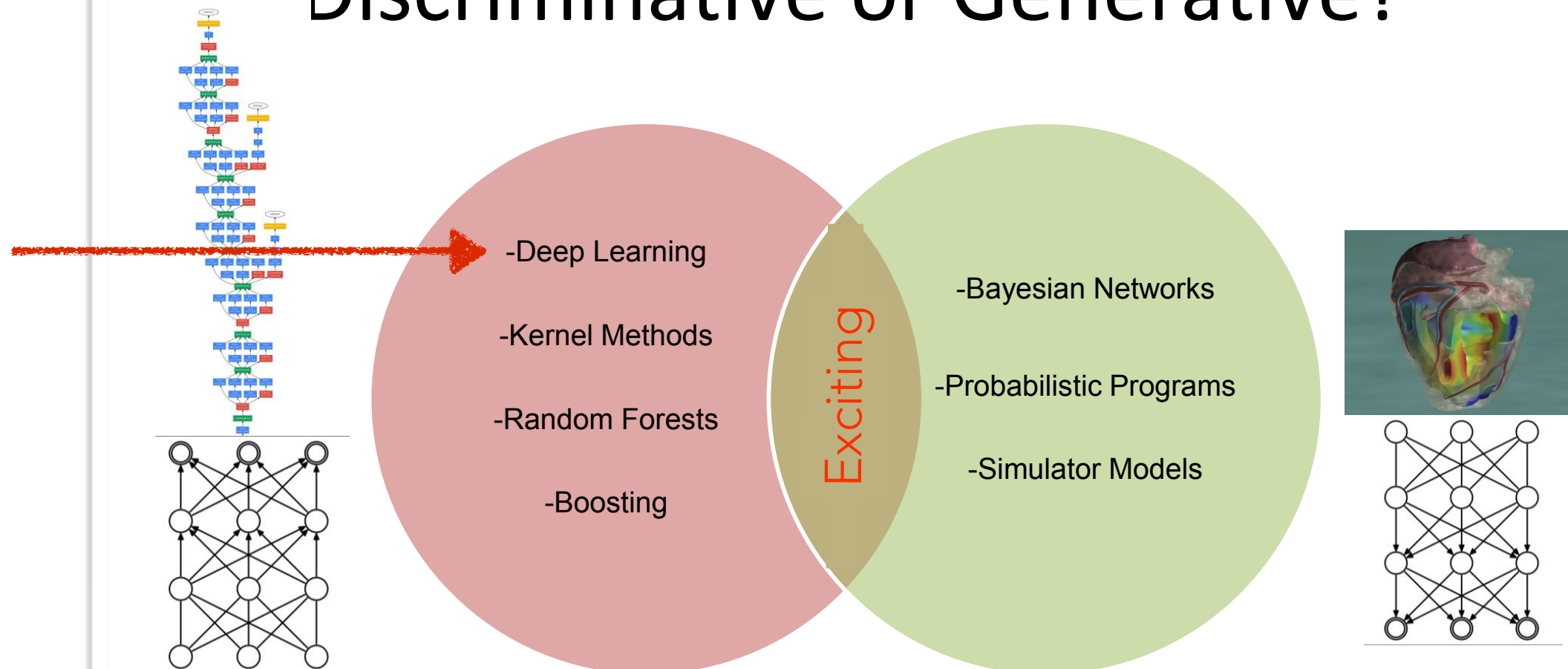
Physics-Aware Machine Learning

We can leverage both the power of deep learning and inject our expert / domain knowledge



Max Welling

Discriminative or Generative?



- Advantages discriminative models:
 - Flexible map from input to target (low bias)
 - Efficient training algorithms available
 - Solve the problem you are evaluating on.
 - Very successful and accurate!

- Advantages generative models:
 - Inject expert knowledge
 - Model causal relations
 - Interpretable
 - Data efficient
 - More robust to domain shift
 - Facilitate un/semi-supervised learning

Reductionist



mechanistic models
clear causal structure

Emergent



descriptive models
unclear causal structure



nuclear & particle
physics

astrophysics

health

climate

documents

language

cosmology

connectome

perception

lattice

protein folding

psychology

simulations

quantum chemistry

systems biology

ecology

Reductionist

Maybe AI should start with problems where causal structure is clear and mechanistic models are available?

Emergent



mechanistic models
clear causal structure

descriptive models
unclear causal structure

nuclear & particle
physics

astrophysics

ecology

health

climate

documents

language

cosmology

connectome

perception

lattice

protein folding

psychology

simulations

quantum chemistry

systems biology

Oct
17

Learning Physics with Deep Neural Networks

Mathematics and Physical Sciences

Speaker



Stéphane Mallat, Ph.D.

École Normale Supérieure de Paris (ENS)

Can we learn physical properties from data? Machine learning offers a solution. It has many similarities with physics, requiring the approximation of functionals which depend on large numbers of variables, such as millions of pixels in images, letters in text, or particles in a physical system. Machine learning algorithms have considerably improved in the last 10 years through the processing of massive amounts of data. In particular, deep neural networks have spectacular applications, such as image classification and medical, industrial and physical data analysis.

In this lecture, Stéphane Mallat will show how machine learning can be applied to statistical physics, turbulent fluids and quantum chemistry. Beyond applications, he will highlight common mathematical approaches in physics and machine learning to overcome the issue of dimensionality. Two central pillars of such approaches are finding symmetries and separating phenomena at different scales. He will show that these pillars also govern the architecture and properties of deep convolutional neural networks.

HAVE A PLAN TO TELL YOU WHERE YOU ARE GOING

GOALS

DEFINE THE MILESTONES THAT WILL ASSIST WITH REACHING YOUR GOALS

OBJECTIVES

DECIDE THE PLAN OF ACTION TO ACHIEVE YOUR OBJECTIVES

STRATEGIES

IDENTIFY THE TOOLS YOU WILL USE TO IMPLEMENT YOUR STRATEGIES

TACTICS

GOALS & STRATEGIES

Goals:

- Use machine learning to do better science

Strategies:

- Import domain knowledge into models (inductive bias)
- Export knowledge from learned models
- Leverage machine learning for intractable inverse problems
- Incorporate traditional scientific concerns into the learning paradigm
 - include impact of domain shift / systematics uncertainties into objective
 - maintain an actionable, scientifically-useful notion of “interpretability”
 - use real-world data for training when possible
 - be data efficient
- Modify codebase to facilitate use of these techniques

GOALS & STRATEGIES

Goals:

- Use machine learning to do better science

Strategies:

- Import domain knowledge into models (inductive bias)
- Export knowledge from learned models
- Leverage machine learning for intractable inverse problems
- Incorporate traditional scientific concerns into the learning paradigm
 - include impact of domain shift / systematics uncertainties into objective
 - maintain an actionable, scientifically-useful notion of “interpretability”
 - use real-world data for training when possible
 - be data efficient
- Modify codebase to facilitate use of these techniques

POSED AS QUESTIONS,
THESE STRATEGIES DEFINE
INTERESTING RESEARCH
DIRECTIONS.

STRATEGIES & TACTICS

Strategy: Import domain knowledge into models

- **Tactic:** exploit symmetries in the data
- **Tactic:** exploit geometric structure of the data
- **Tactic:** exploit causal structure of the generative process
- **Tactic:** exploit hierarchical / compositional structure
- **Tactic:** exploit Markov property of the generative process
- **Tactic:** exploit tangent space of statistical manifold

Strategy: Export knowledge from learned models

- **Tactic:** learnable components that can be interpreted

Strategy: maintain an actionable, scientifically-useful notion of “interpretability”

- **Tactic:** compose model from reusable components that perform a specific task and can be individually characterized & validated

STRATEGIES & TACTICS

Strategy: Leverage machine learning for intractable inverse problems

- **Tactic:** Use the likelihood ratio trick to convert a discriminative classifier into a density ratio
- **Tactic:** Use autoregressive models & normalizing flows for conditional density estimation
- **Tactic:** Use universal probabilistic programming
- **Tactic:** Approximate gradients of non-differentiable, black-box models (AVO, RELAX, ...)

Strategy: include impact of systematics uncertainties into objective

- **Tactic:** Design loss functions more relevant to scientific goals
- **Tactic:** Adversarial training for continuous domain adaptation & fairness ("learning to pivot")

Strategy: use real-world data for training when possible

- **Tactic:** Weakly supervised learning

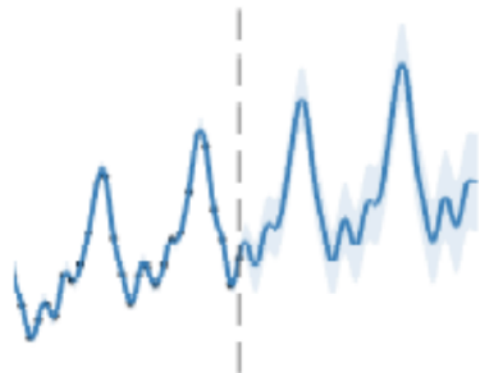
Strategy: be data efficient

- **Tactic:** exploiting domain knowledge can dramatically reduce number of parameters

PHYSICS-AWARE MACHINE LEARNING

Physics goes into the construction of a “Kernel” that defines M.L. model

- Vocabulary of kernels + grammar for composition = powerful modeling



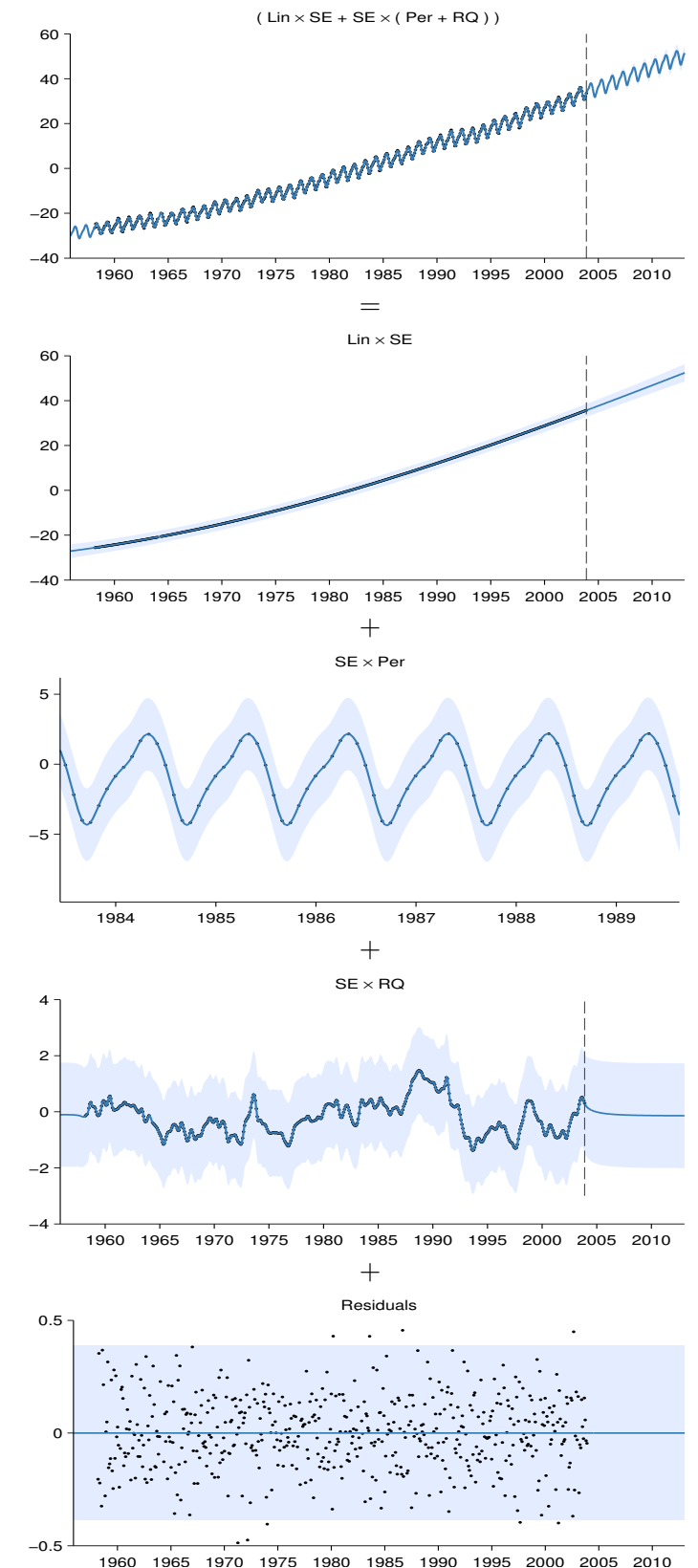
Structure Discovery in Nonparametric Regression through Compositional Kernel Search

David Duvenaud, James Robert Lloyd, Roger Grosse, Joshua B. Tenenbaum, Zoubin Ghahramani
International Conference on Machine Learning, 2013
[pdf](#) | [code](#) | [poster](#) | [bibtex](#)

Exploiting compositionality to explore a large space of model structures

Roger Grosse, Ruslan Salakhutdinov, William T. Freeman, Joshua B. Tenenbaum
Conference on Uncertainty in Artificial Intelligence, 2012
[pdf](#) | [code](#) | [bibtex](#)

Mauna Loa atmospheric CO₂



Publishing Challenges

WHO IS THE AUDIENCE FOR THE PAPERS?

Papers that innovate new methods motivated by physics problems have an audience problem:

- hard to address both ML & physics community simultaneously (language, motivation, background,...)
- communities have very different styles and norms for what constitutes a good paper
- Same method may also apply to other domains

Multiple contributions aimed at different audiences

- targeted communication, but overlapping content

MORE ATOMIC PUBLISHABLE UNITS

Abstracted Problem statement

Key Idea

Instances of problem / motivation in domain context

Experiments

- toy experiments
 - pedagogical known solution
- real-world experiments
- replication / verification of reproducibility

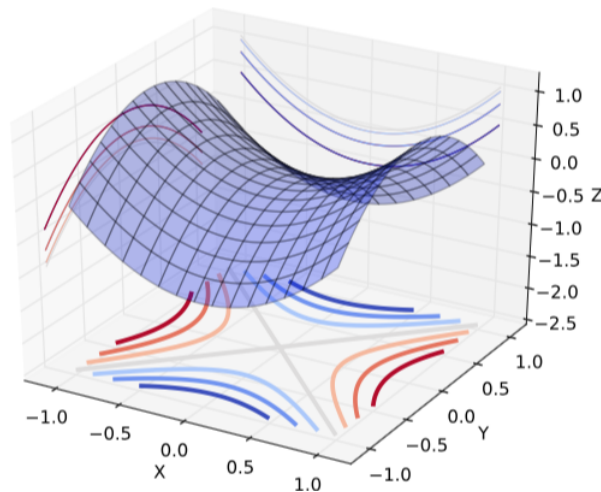
Examples of Physics+ML: A Symbiotic Relationship

INCORPORATING SYSTEMATIC UNCERTAINTY

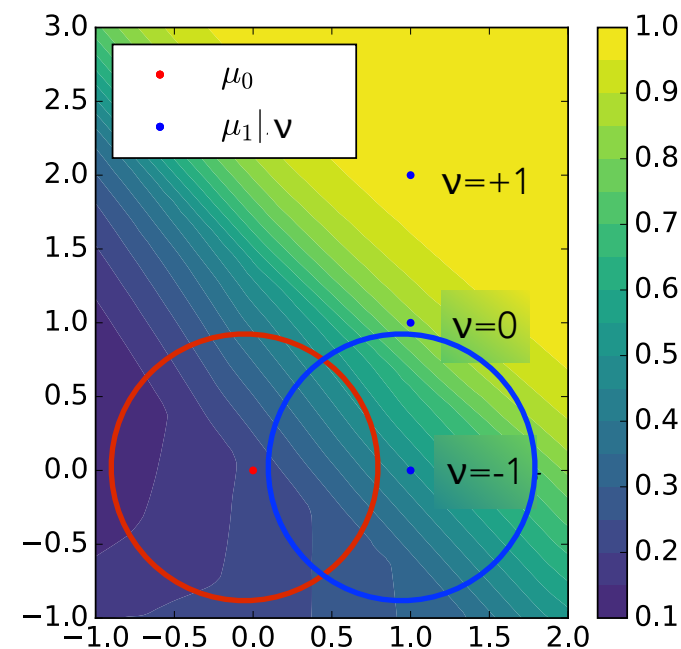
Typically classifier $\mathbf{f}(\mathbf{x})$ trained to minimize loss $\mathbf{L}_f$.

- want classifier output to be insensitive to systematics (nuisance parameter $\mathbf{v}$)
- introduce an **adversary** $\mathbf{r}$ that tries to predict $\mathbf{v}$ based on $\mathbf{f}$.
- setup as a minimax game:

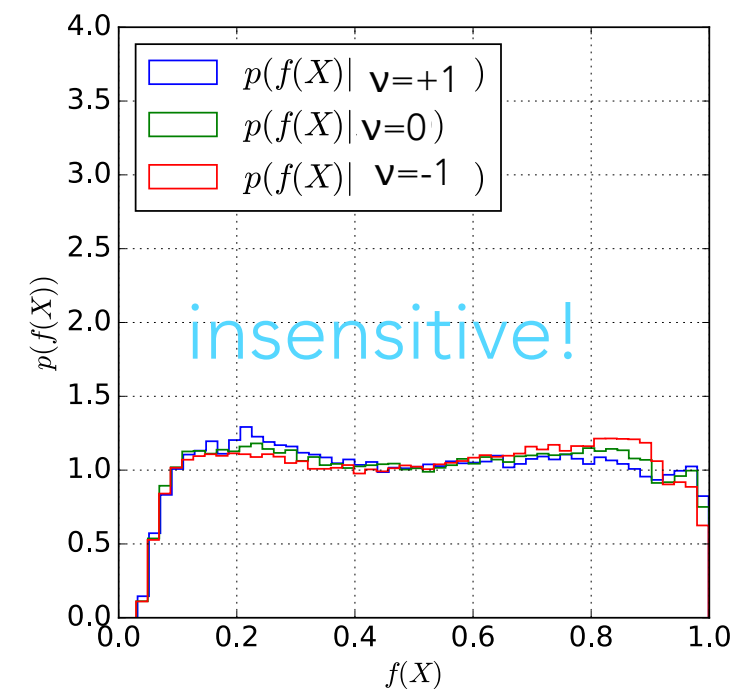
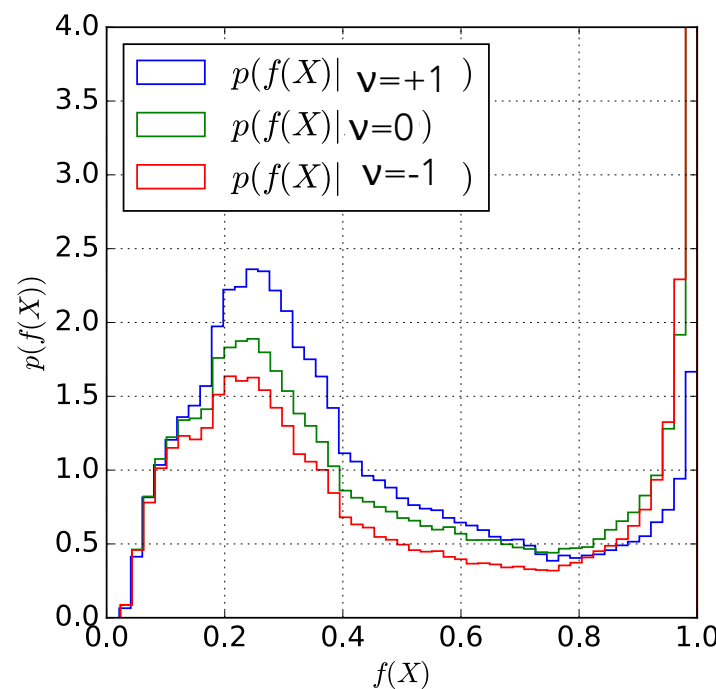
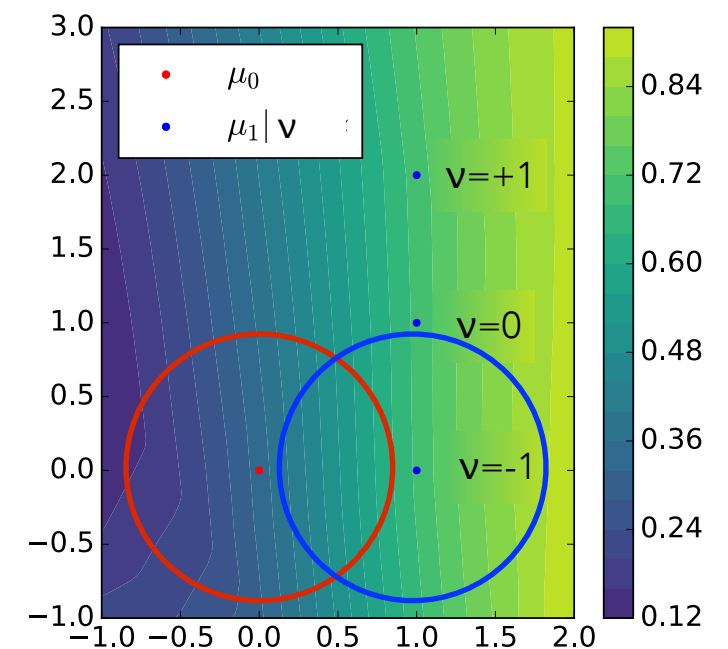
$$(f^*, r^*) = \max_r \min_f V(f, r)$$



normal training



adversarial training



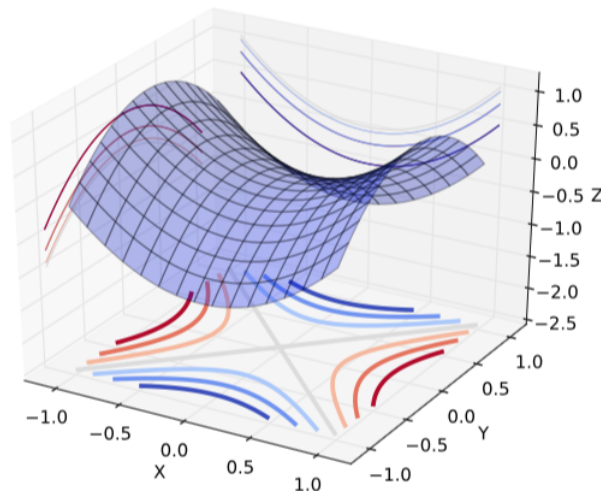
insensitive!

INCORPORATING SYSTEMATIC UNCERTAINTY

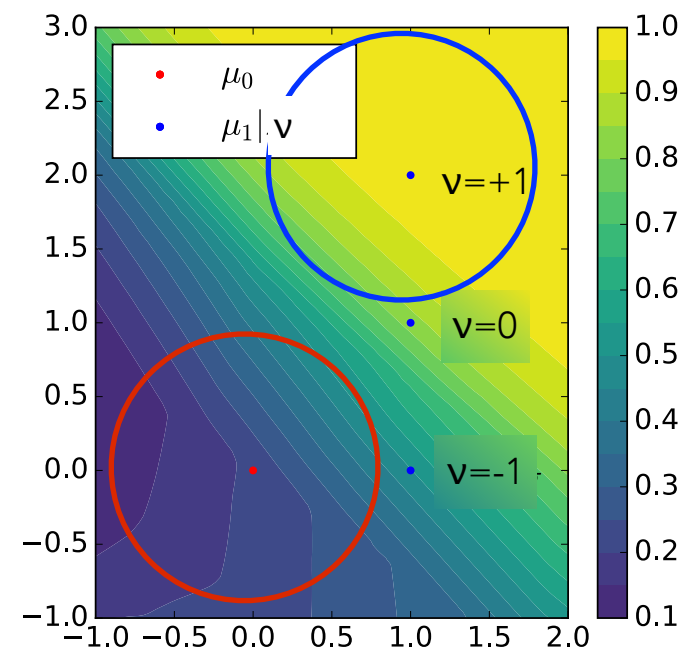
Typically classifier $\mathbf{f}(\mathbf{x})$ trained to minimize loss $\mathbf{L}_f$.

- want classifier output to be insensitive to systematics (nuisance parameter $\mathbf{v}$)
- introduce an **adversary** $\mathbf{r}$ that tries to predict $\mathbf{v}$ based on $\mathbf{f}$.
- setup as a minimax game:

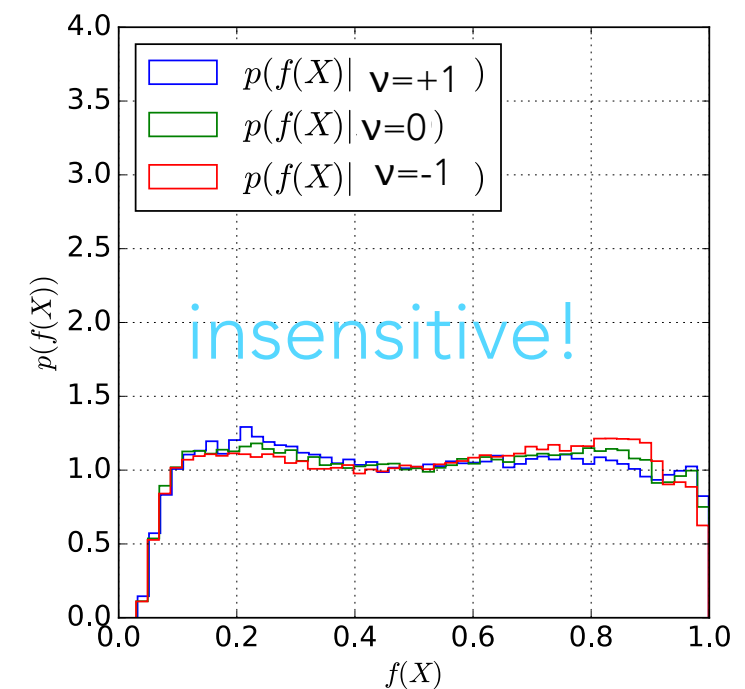
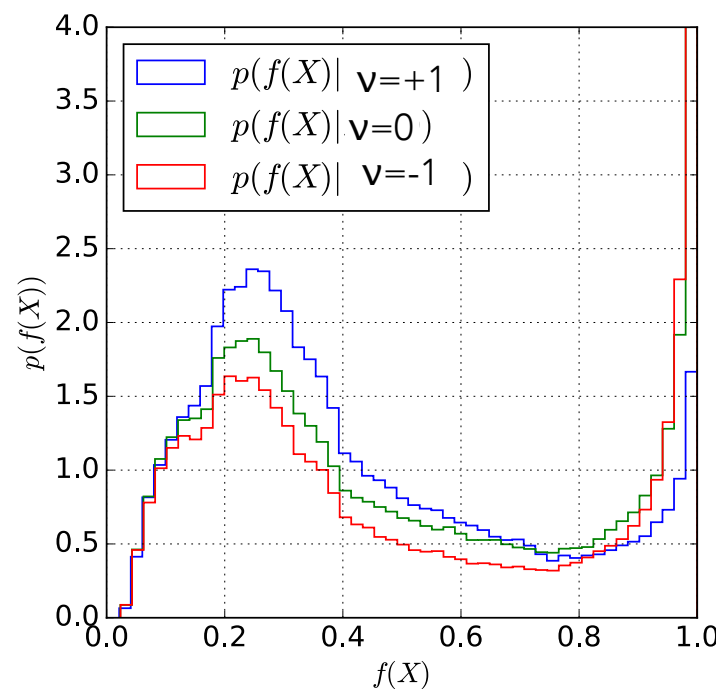
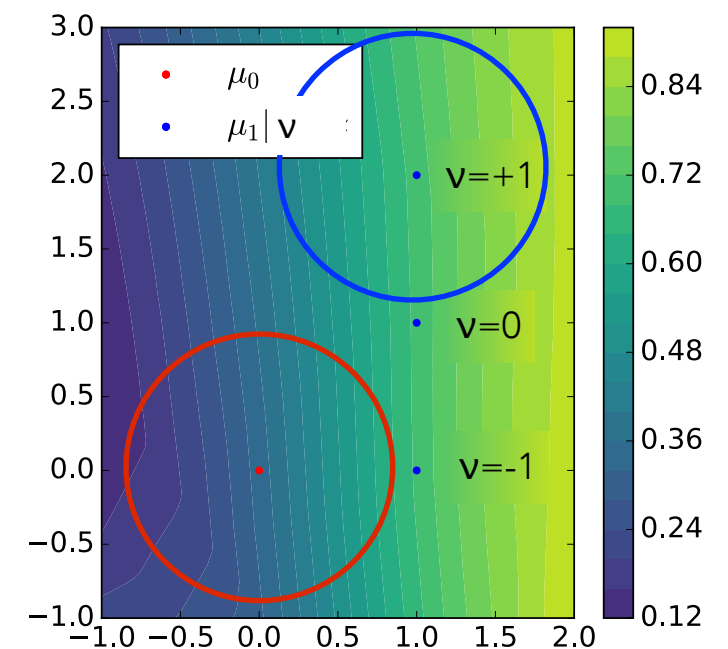
$$(f^*, r^*) = \max_r \min_f V(f, r)$$



normal training



adversarial training

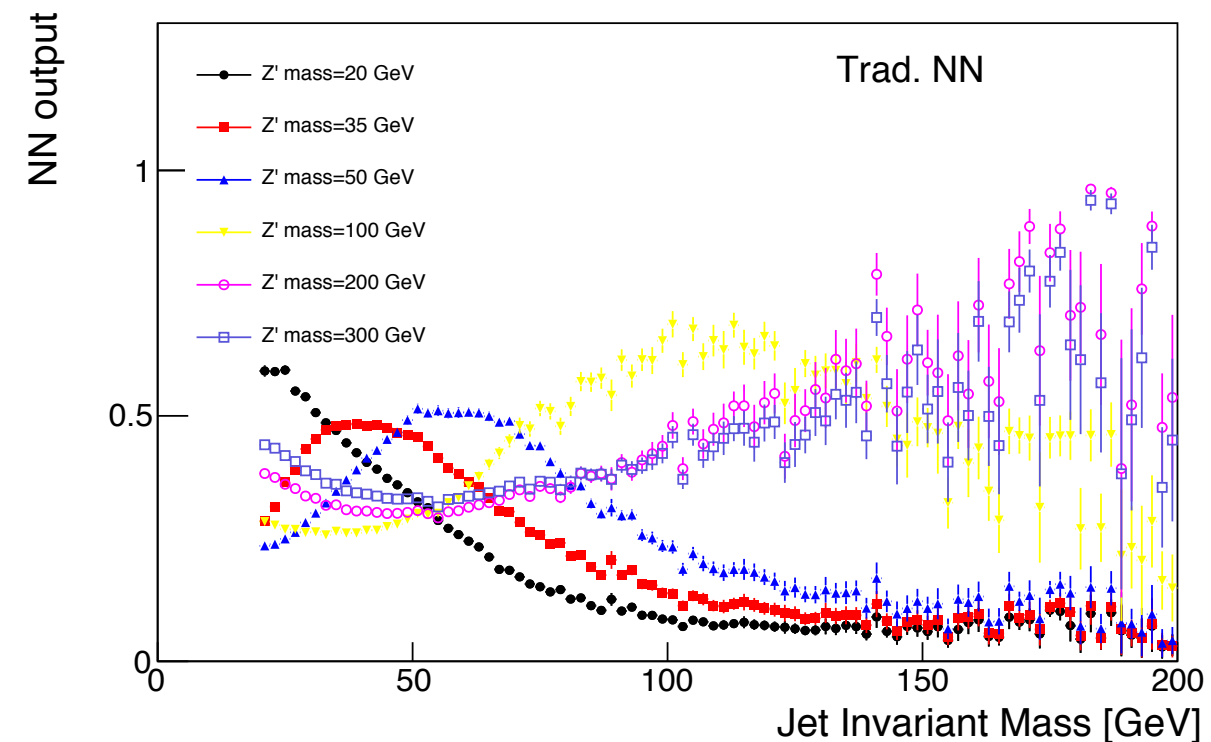
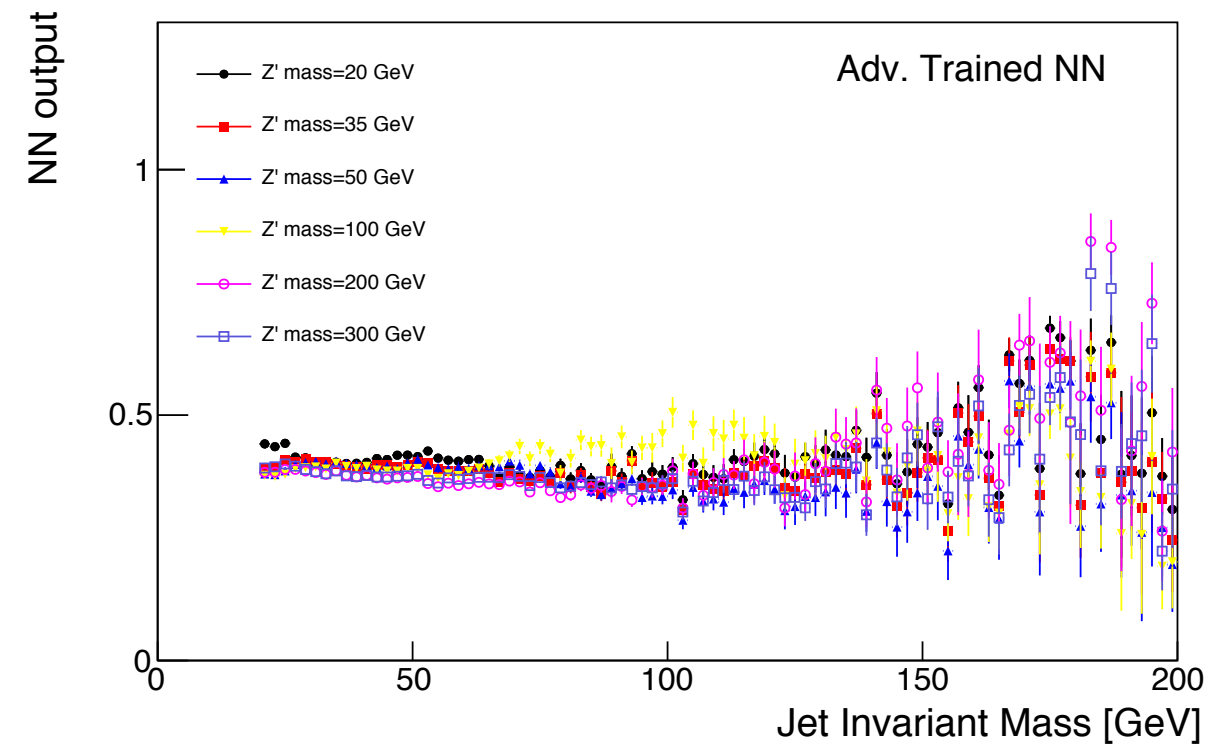


FAIR CLASSIFIERS

Adversarial approach of “Learning to Pivot” can also be used to train a classifier that is independent from some other continuous variable.

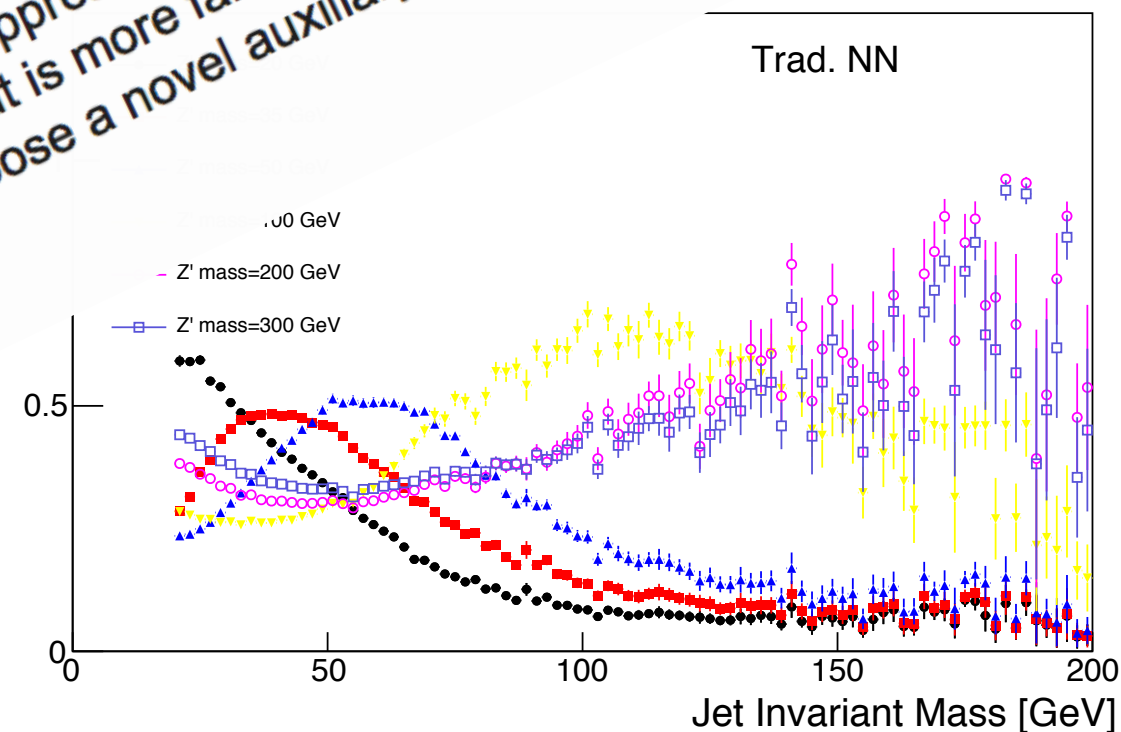
- fairness to continuous attribute
- motivation for doing this is related to robustness to uncertainties and interpretability

K.C, J. Pavez, and G. Louppe, arXiv:1506.02169
P. Baldi, K.C, T. Faucett, P. Sadowski, D. Whiteson arXiv:1601.07913
G. Louppe, M. Kagan, K.C, arXiv:1611.01046
Shimmin, et. al. arXiv:1703.03507



Adversarial approach of
 “Learning to Pivot” can also be
 used to train a classifier
 independent from
 continuous

- Fader networks: Manipulating images by sliding attributes**
[G Lample, N Zeghidour, N Usunier...](#) - [Advances in Neural ...](#), 2017 - [papers.nips.cc](#)
 This paper introduces a new encoder-decoder architecture that is trained to reconstruct images by disentangling the salient information of the image and the values of attributes directly in the latent space. As a result, after training, our model can generate different ...
 ☆ Related articles All 5 versions
- Fairness GAN**
[P Sattigeri, SC Hoffman, V Chenthamarakshan...](#) - [arXiv preprint arXiv ...](#), 2018 - [arxiv.org](#)
 In this paper, we introduce the Fairness GAN, an approach for generating a dataset that is plausibly similar to a given multimedia dataset, but is more fair with respect to protected attributes in allocative decision making. We propose a novel auxiliary classifier GAN that ...
 ☆ Related articles All 3 versions



Adversarial Variational Optimization of Non-Differentiable Simulators

Gilles Louppe¹ and Kyle Cranmer¹¹New York University

Complex computer simulators are increasingly used across fields of science as generative models tying parameters of an underlying theory to experimental observations. Inference in this setup is often difficult, as simulators rarely admit a tractable density or likelihood function. We introduce Adversarial Variational Optimization (AVO), a likelihood-free inference algorithm for fitting a non-differentiable generative model incorporating ideas from empirical Bayes and variational inference. We adapt the training procedure of generative adversarial networks by replacing the differentiable generative network with a domain-specific simulator. We solve the resulting non-differentiable min-max problem by minimizing variational upper bounds of the two adversarial objectives. Effectively, the procedure results in learning a proposal distribution over simulator parameters, such that the corresponding marginal distribution of the generated data matches the observations. We present results of the method with simulators producing both discrete and continuous data.

Similar to W-GAN setup, but instead of using a neural network as the generator, use the actual simulation (eg. Pythia, GEANT)

Continue to use a neural network discriminator / critic.

Difficulty: the simulator isn't differentiable, but there's a **trick!**

Allows us to efficiently fit / **tune simulation** with stochastic gradient techniques!

Leo is G Tom is D

NEW! AVO

Adversarial Variational Optimization of Non-Differentiable

Gilles Louppe¹ and

Complex con

tyi
oft
Adv
diffe
We
gener
max
the pr
corresp
results

Synthesizing Programs for Images using Reinforced Adversarial Learning

Yaroslav Ganin¹ Tejas Kulkarni² Igor Babuschkin² S. M. Ali Eslami² Oriol Vinyals²

Abstract

Advances in deep generative networks have led to impressive results in recent years. Nevertheless, such models can often waste their capacity on the minutiae of datasets, presumably due to weak inductive biases in their decoders. This is where graphics engines may come in handy since they abstract away low-level details and represent images as high-level programs. Current methods that combine deep learning and renderers are limited by hand-crafted likelihood or distance functions, a need for large amounts of supervision, or difficulties in scaling their inference algorithms to richer datasets. To mitigate these issues, we present SPIRAL, an adversarially trained agent that generates a program which is executed by a graphics engine to interpret and sample images. The goal of this agent is to fool a discriminator network that distinguishes between real and rendered images.

Similar

o, but
I network
actual
ANT)

network

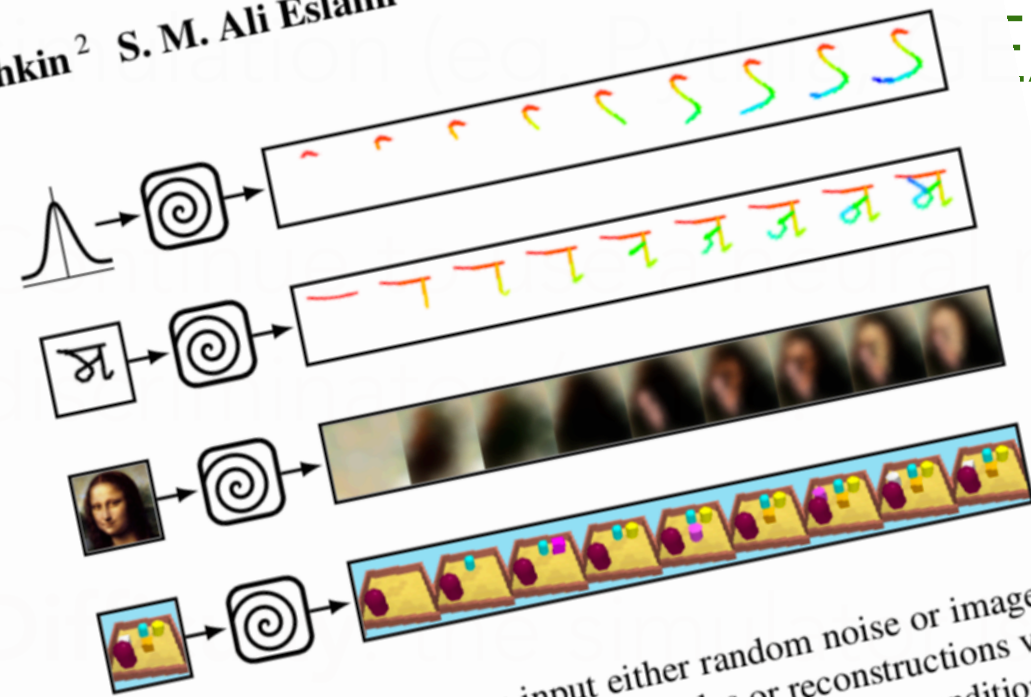


Figure 1. **SPIRAL** takes as input either random noise or images and iteratively produces plausible samples or reconstructions via graphics program synthesis. The first row depicts an unconditional run given random noise. The second, third and fourth rows depict conditional execution given an image with a handwritten character, the Mona Lisa, and objects arranged in a 3D scene.

ick!

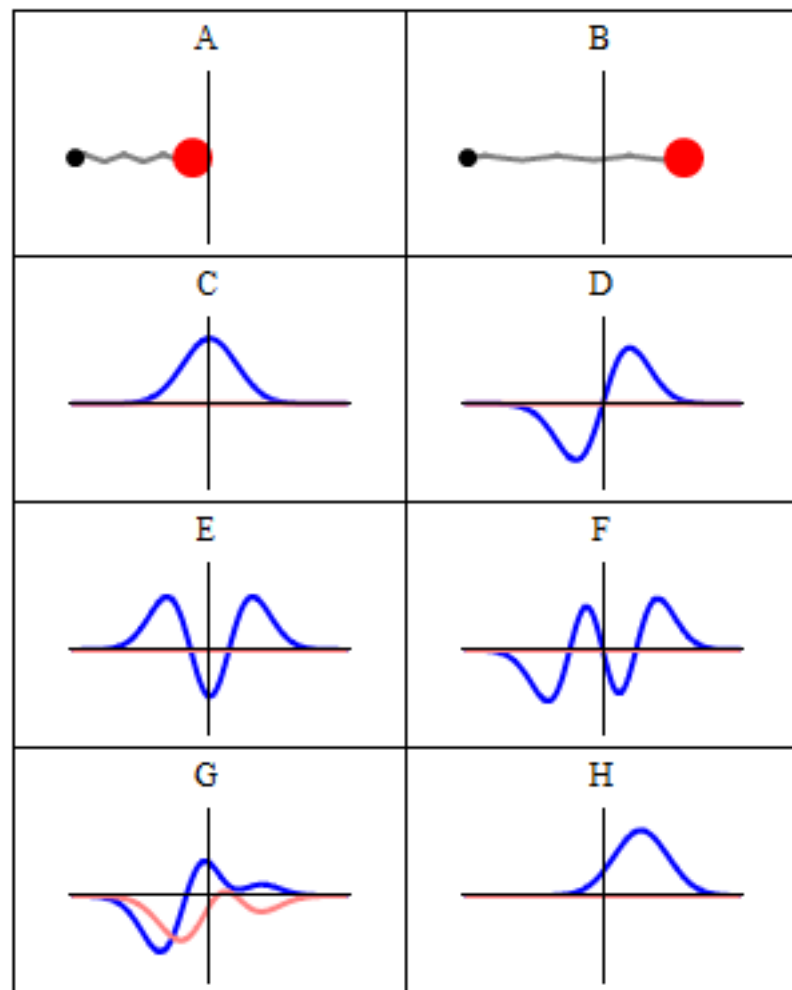
ufficiently fit /
same simulation with stochastic
gradient techniques!

Leo is G Tom is D

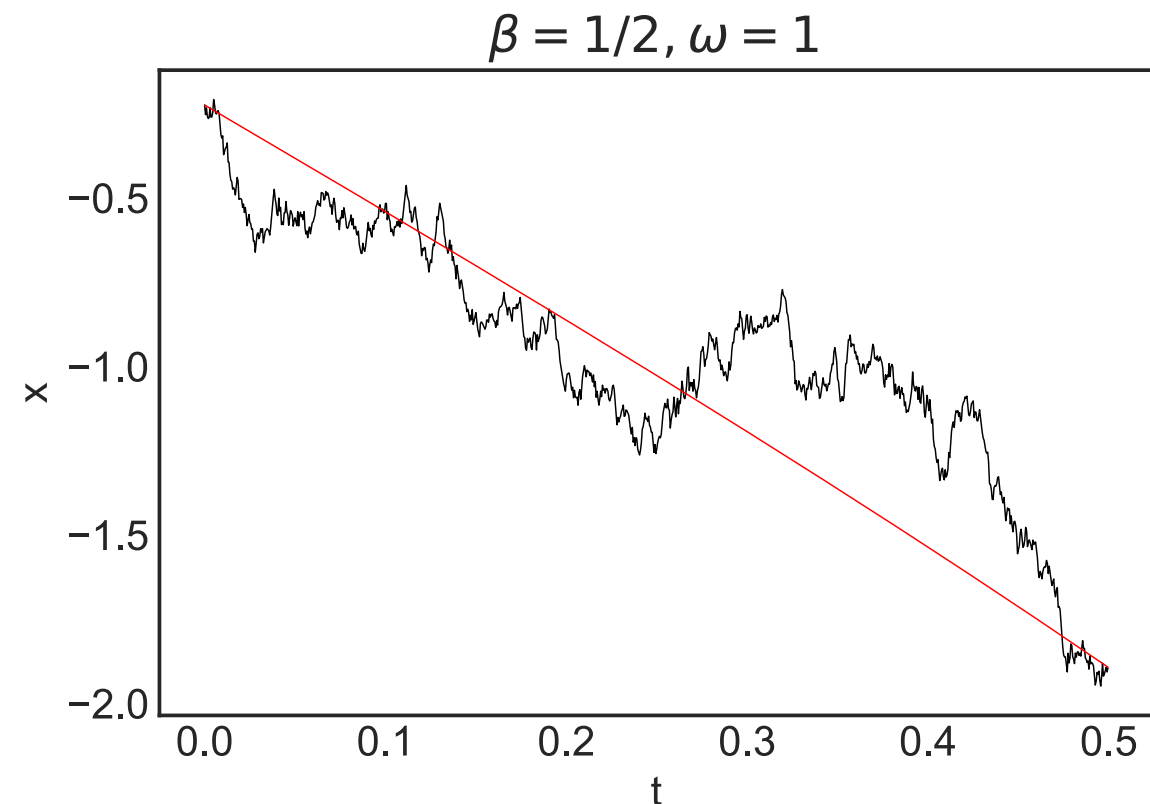
SIMPLE HARMONIC OSCILLATOR

Start simple:

- 0+1 Quantum Field Theory = 1-D Quantum Mechanics



Use path integral to sample

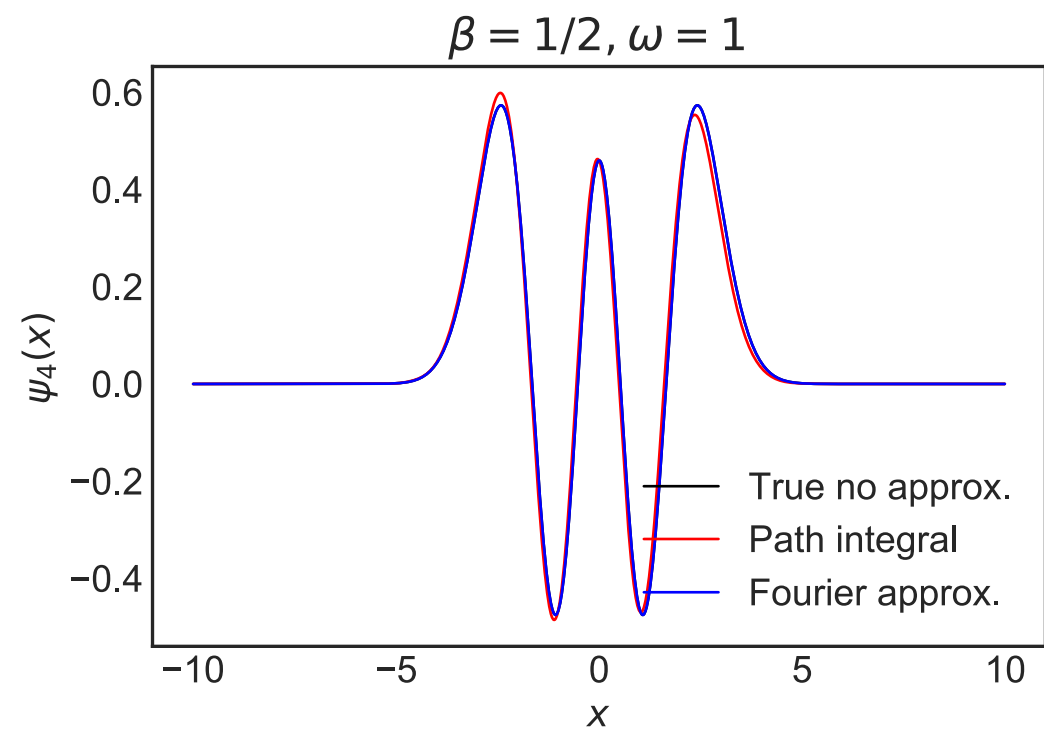
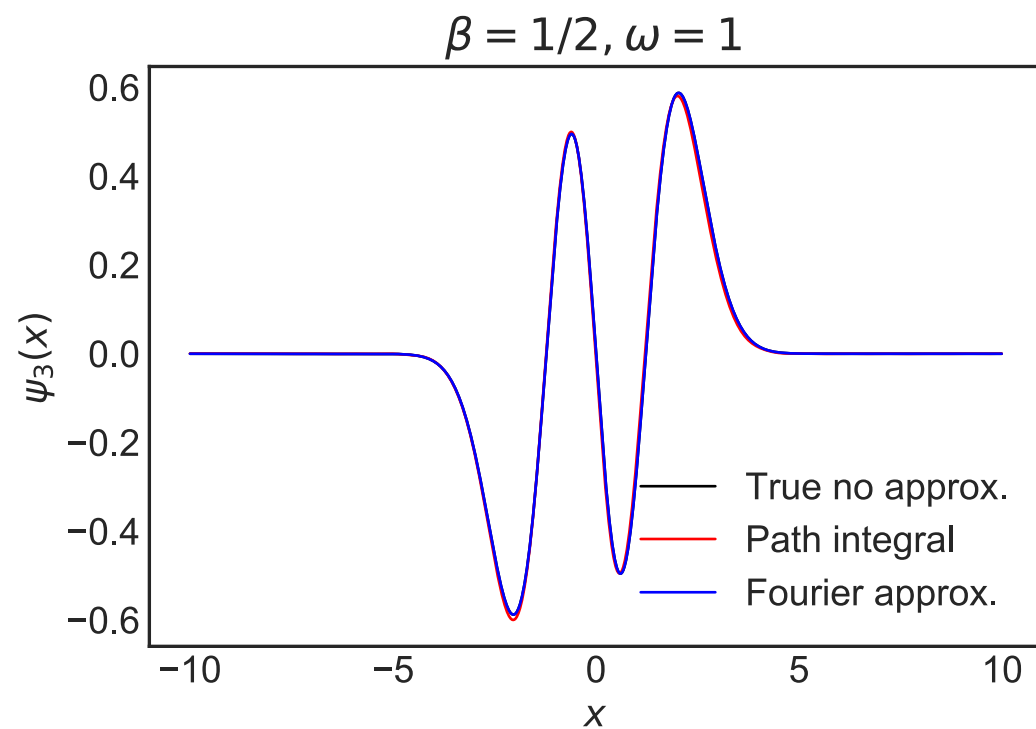
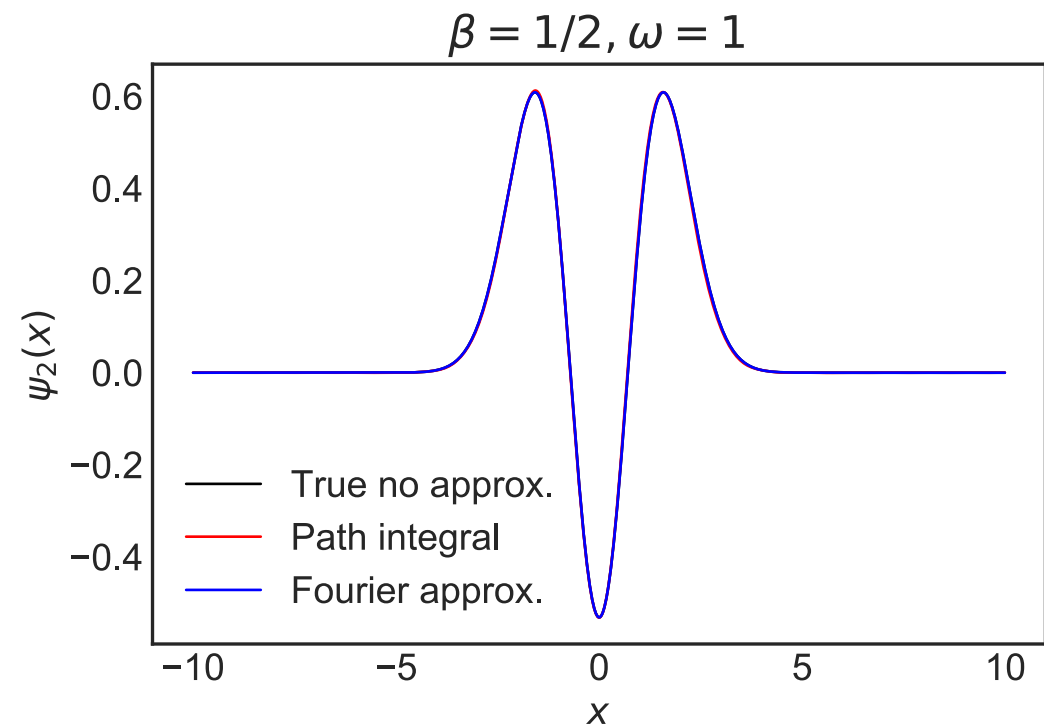
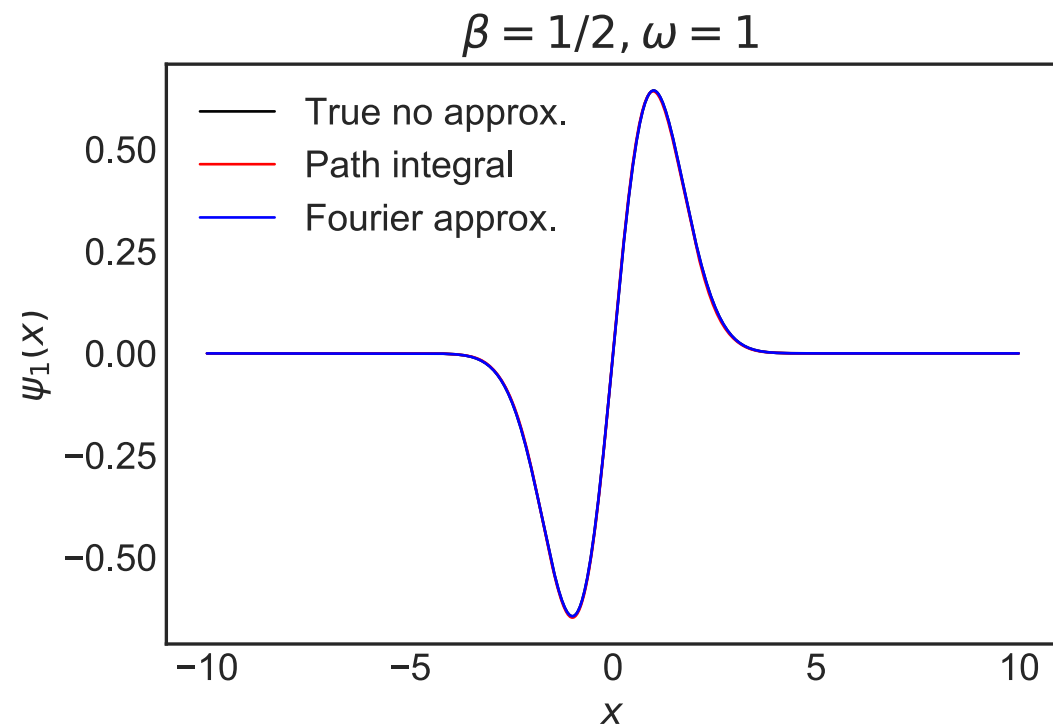


Minimize the following loss function

$$L[\tilde{\rho}] = \sum_{(x_i, x'_i) \sim \rho(x, x')} \left[\sum_i -\log \tilde{\lambda}_i \tilde{\psi}_i(x_i) \tilde{\psi}_i^*(x'_i) \right] + C \sum_{i < j} \sum_k a_k^{(i)} a_k^{(j)}$$

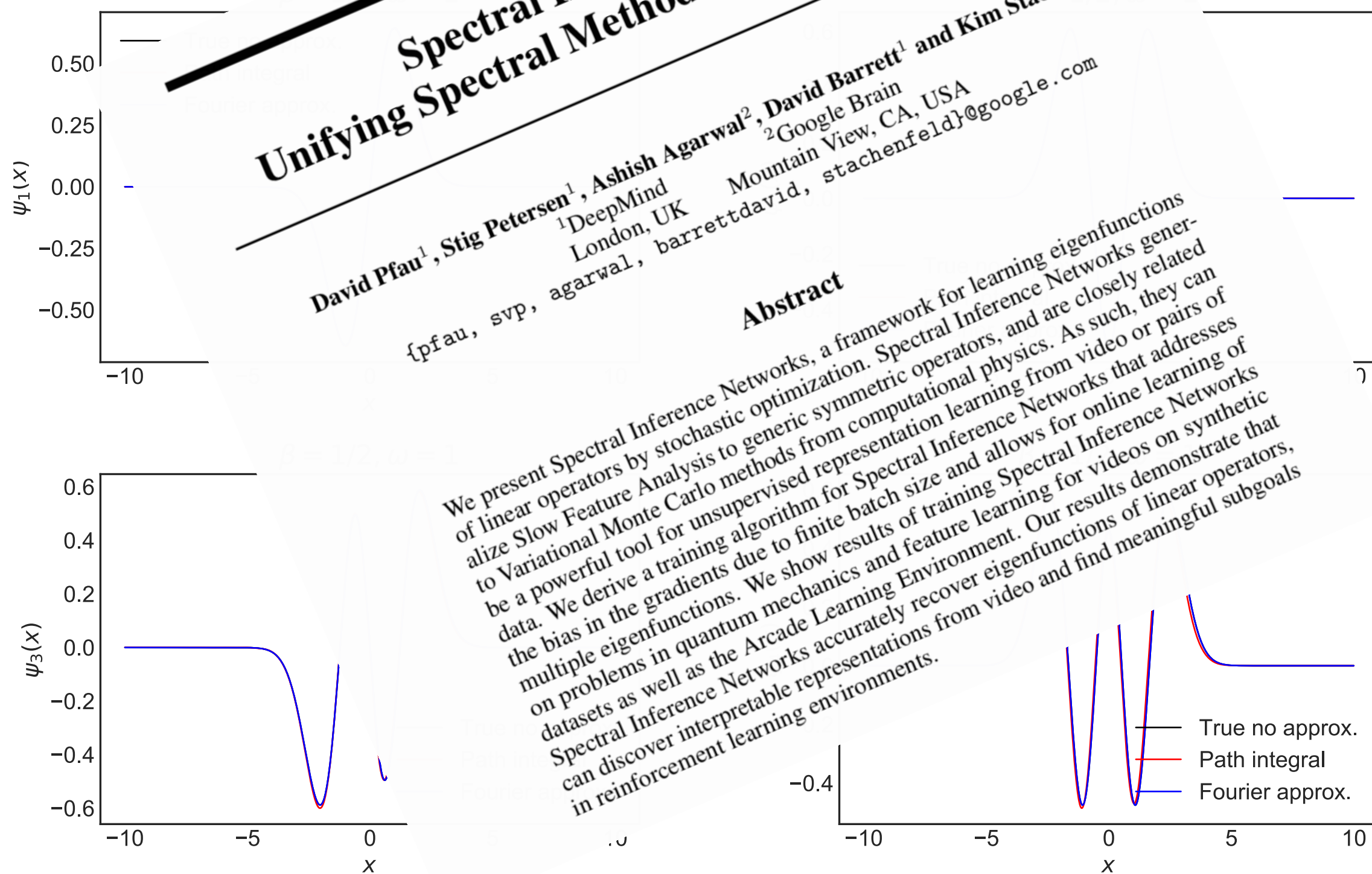
FITTING THE ENERGY EIGENSTATES

Can extract the eigenstates as well



FITTING THE ENERGY EIGENSTATES

Can extract the eigenstates



Backdrop: Stochastic Backpropagation



Giuseppe Carleo

@gppcarleo

Following



Somehow, I overlooked this (very cool) recent work by @KyleCranmer arxiv.org/pdf/1806.01337... ! Incidentally, our quantum calculations are typically "problems where evaluation of the loss on a small minibatch leads to a poor approximation of the real loss", worth trying it!

11:03 AM - 7 Jun 2018

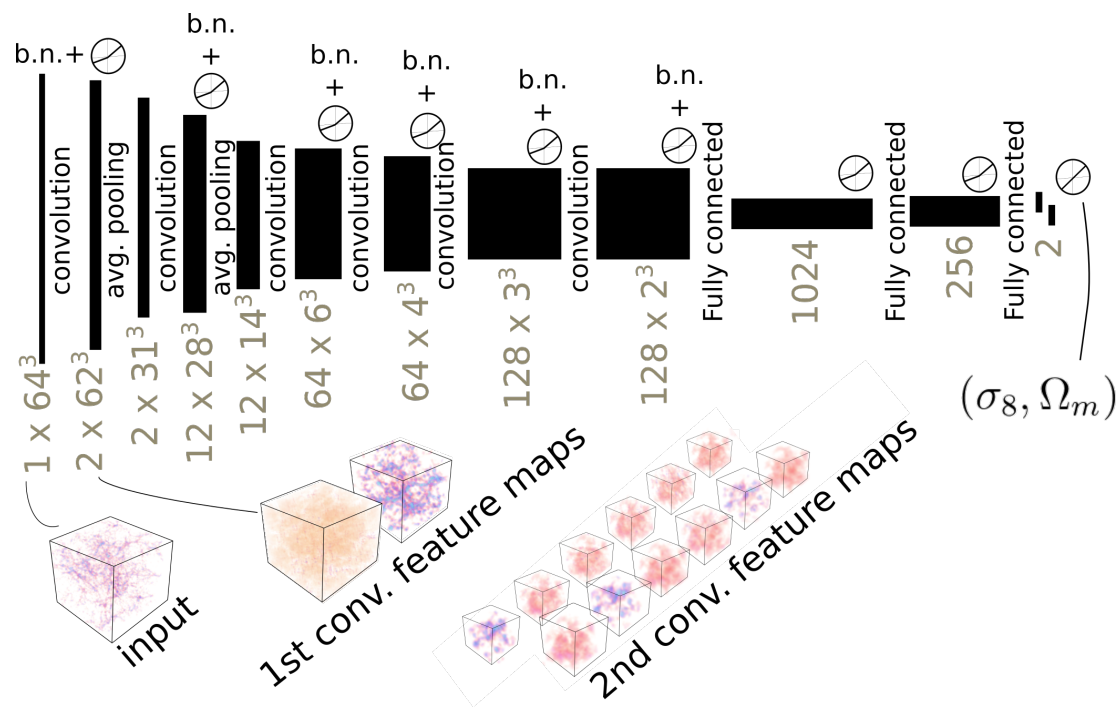
Estimating Cosmological Parameters from the Dark Matter Distribution

Siamak Ravanbakhsh*
 Junier Oliva*
 Sebastien Fromenteau†
 Layne C. Price†
 Shirley Ho†
 Jeff Schneider*
 Barnabás Póczos*

MRAVANBA@CS.CMU.EDU
 JOLIVA@CS.CMU.EDU
 SFROMENT@ANDREW.CMU.EDU
 LAYNEP@ANDREW.CMU.EDU
 SHIRLEYH@ANDREW.CMU.EDU
 JEFF.SCHNEIDER@CS.CMU.EDU
 BAPOCZOS@CS.CMU.EDU

* School of Computer Science, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, USA

† McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Carnegie 5000 Forbes Ave., Pittsburgh, PA 15213, USA



Simulations are expensive, very few training examples

- each one is huge, crop into smaller boxes
- cropping loses larger scale structure

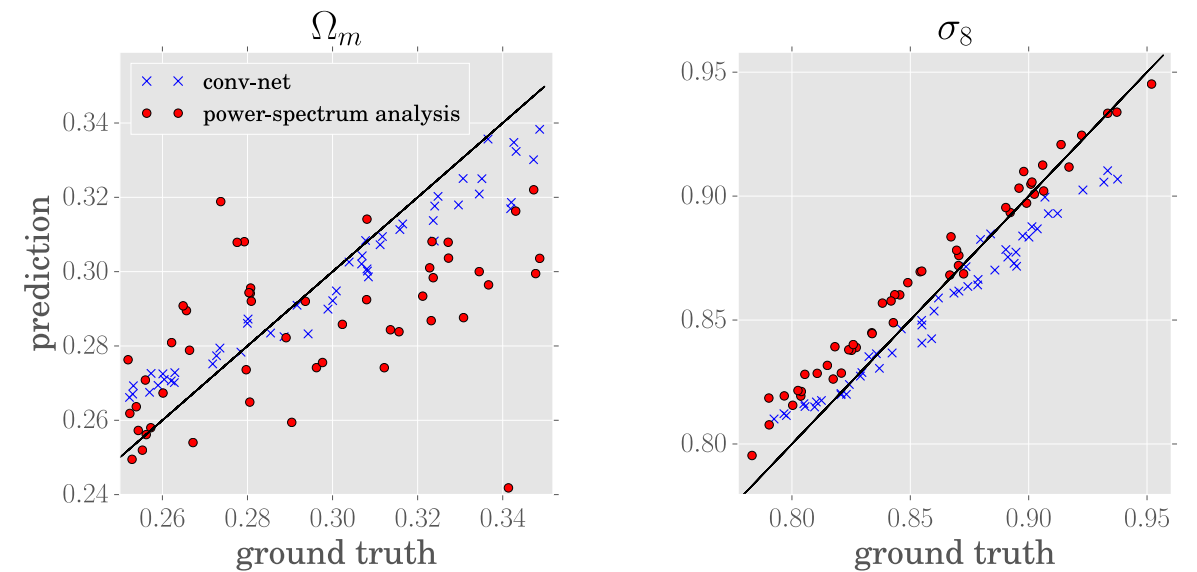
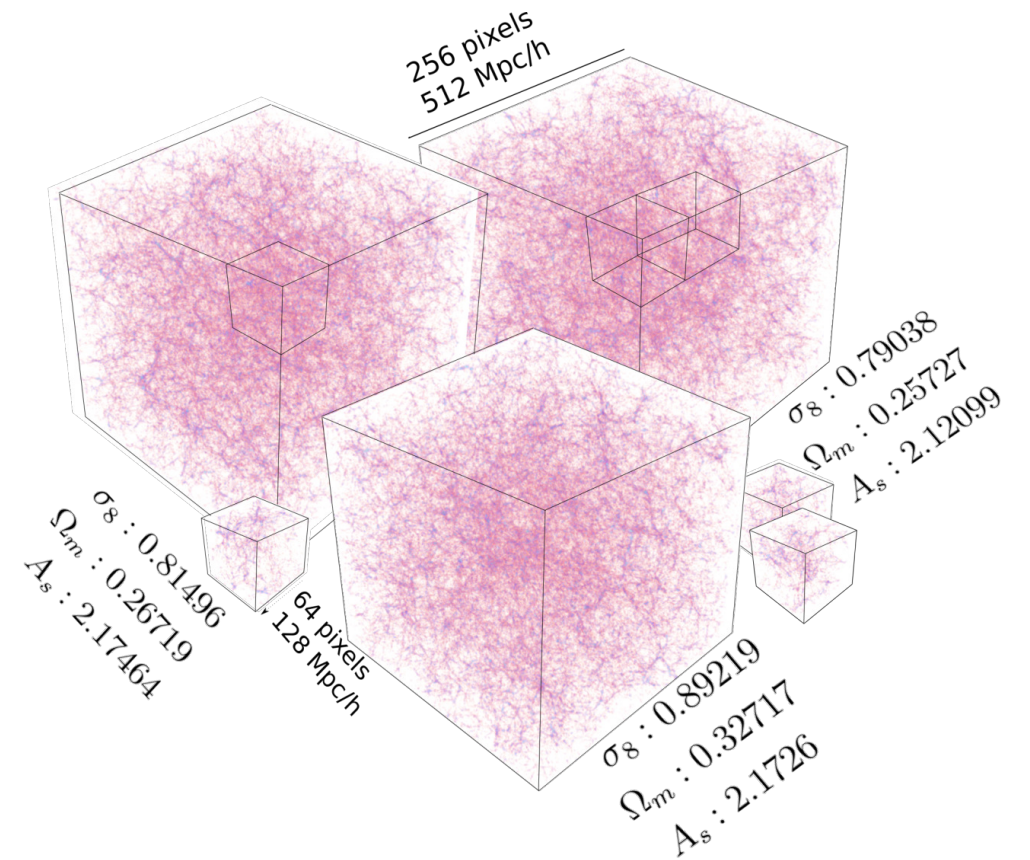


Figure 2. Prediction and ground truth of Ω_m and σ_8 using 3D conv-net and analysis of the power-spectrum on 50 test cube instances.

Machine learning action parameters in lattice quantum chromodynamics

Phiala E. Shanahan,^{1,2} Daniel Trewartha,² and William Detmold³

¹Department of Physics, College of William and Mary, Williamsburg, VA 23187-8795, USA

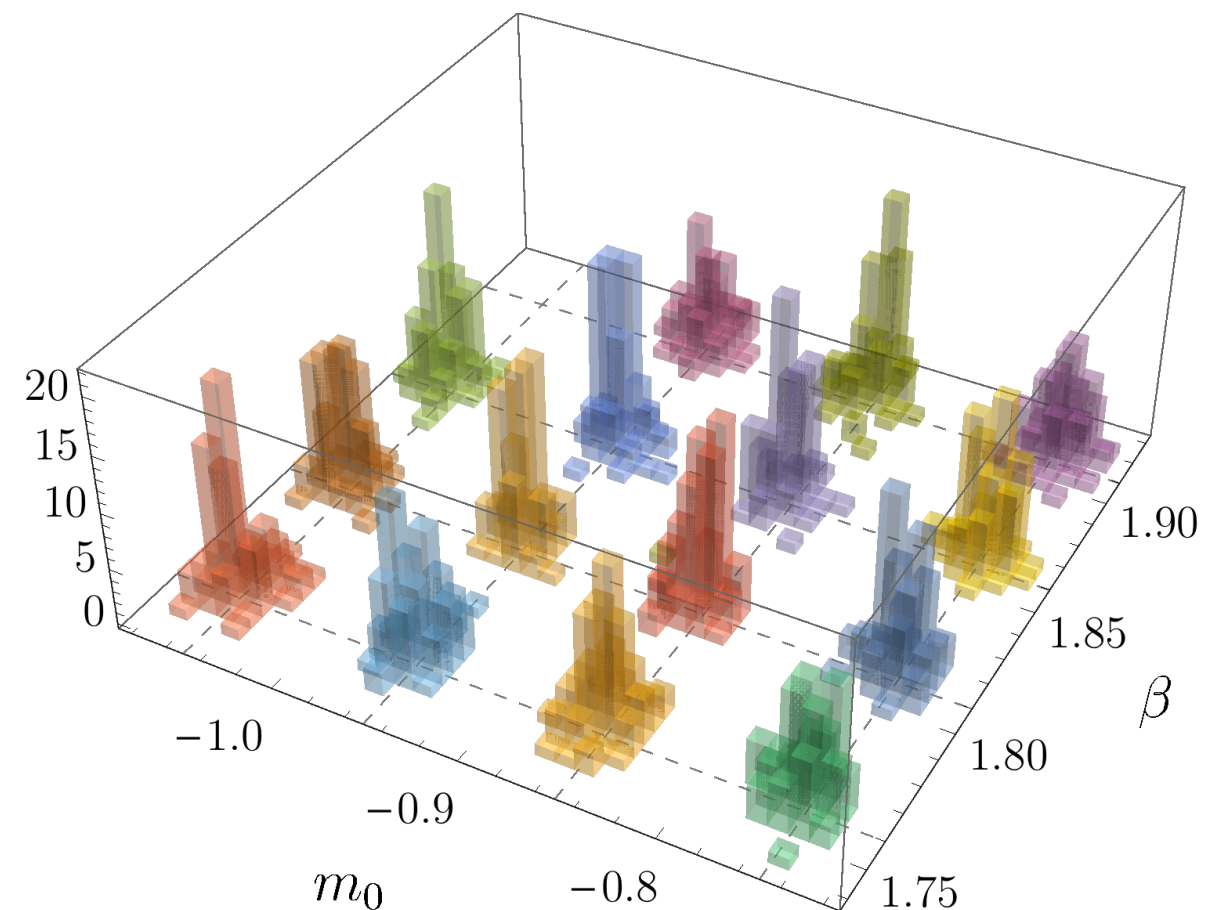
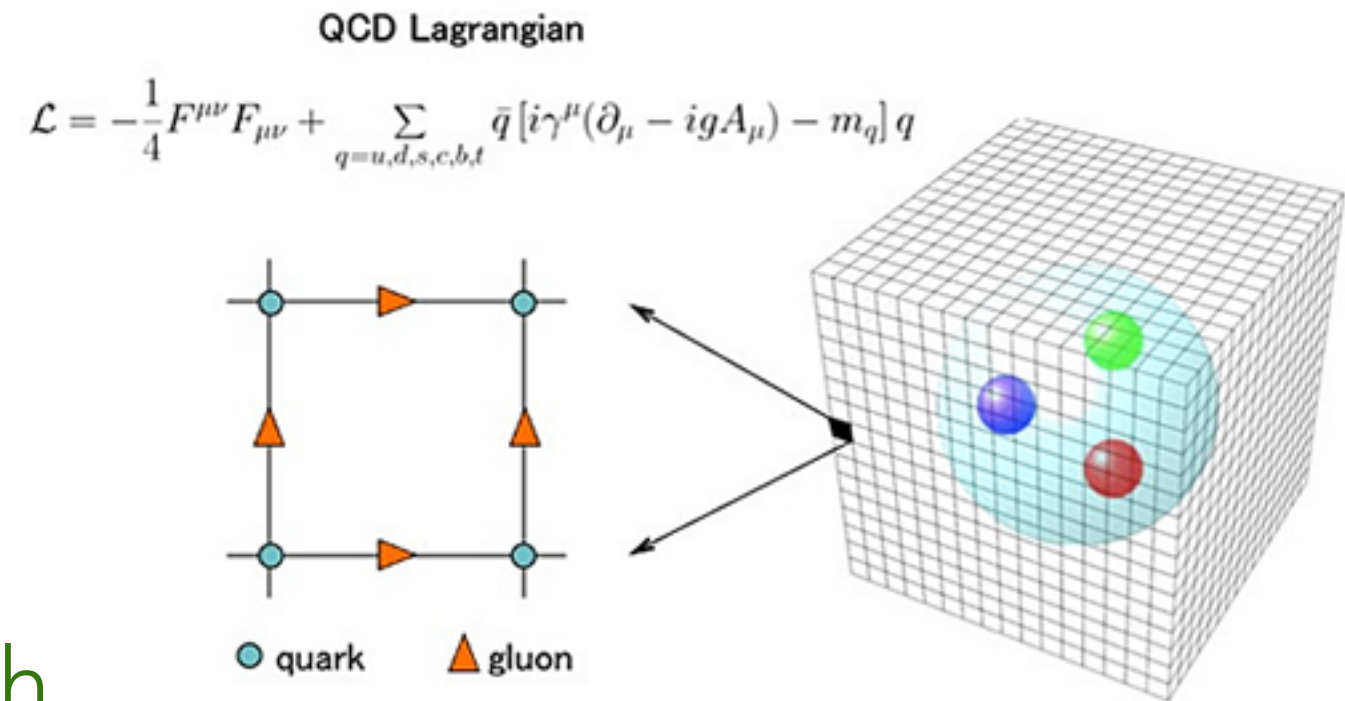
²Jefferson Laboratory, 12000 Jefferson Avenue, Newport News, VA 23606, USA

³Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

(Dated: January 18, 2018)

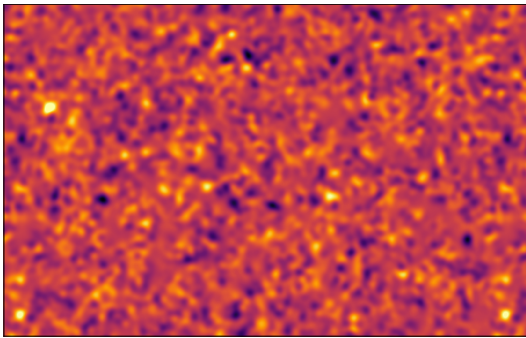
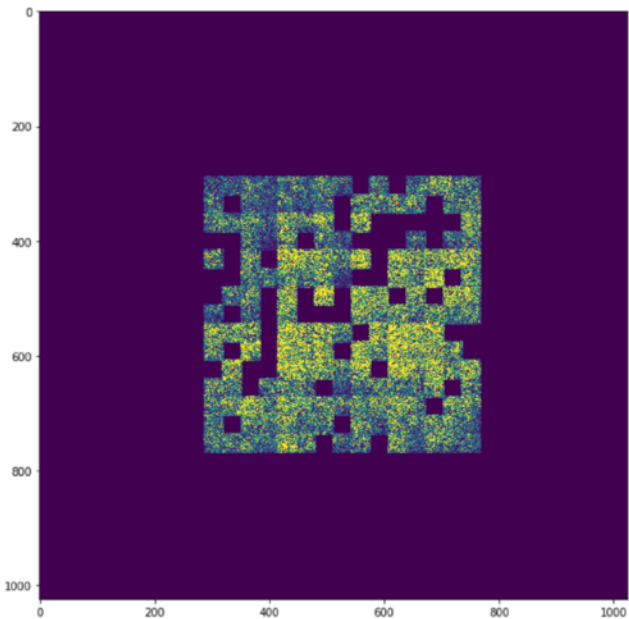
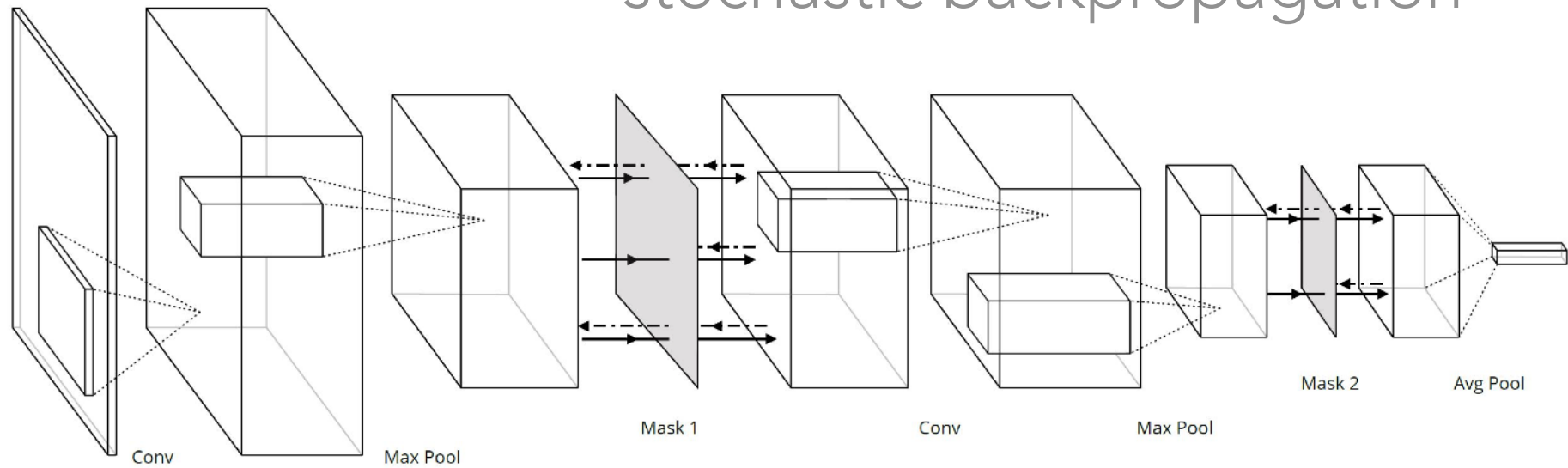
Each of the 10^7 lattice locations has data $x_i \in \mathbb{R}^{32}$ with non-trivial data with continuous local symmetry.

- $O(100)$ of training examples
- each one is huge
- cropping is a problem because data is hierarchical

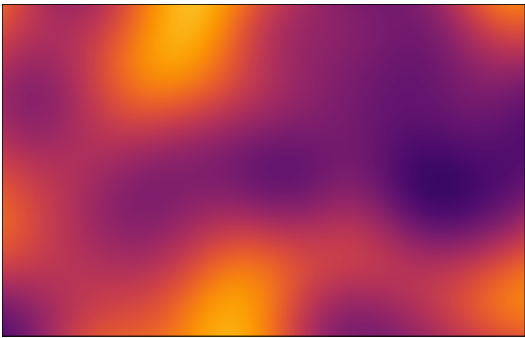


BACKDROP

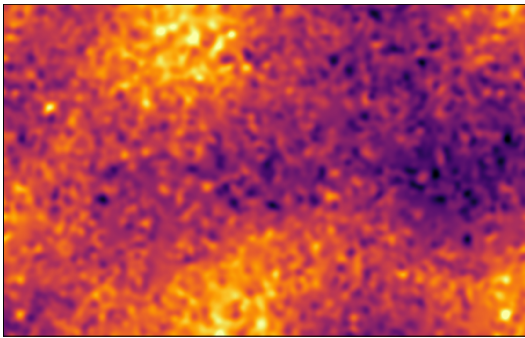
identity on forward pass,
stochastic backpropagation



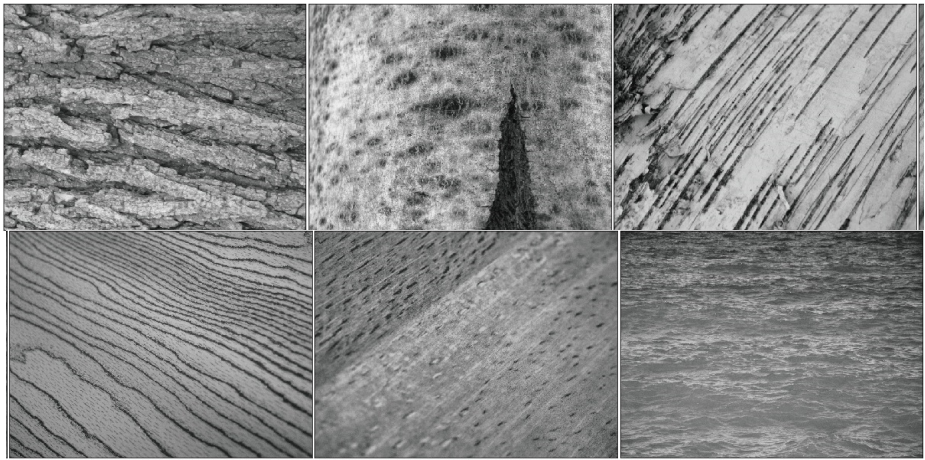
(a) Small scale GP



(b) Large scale GP

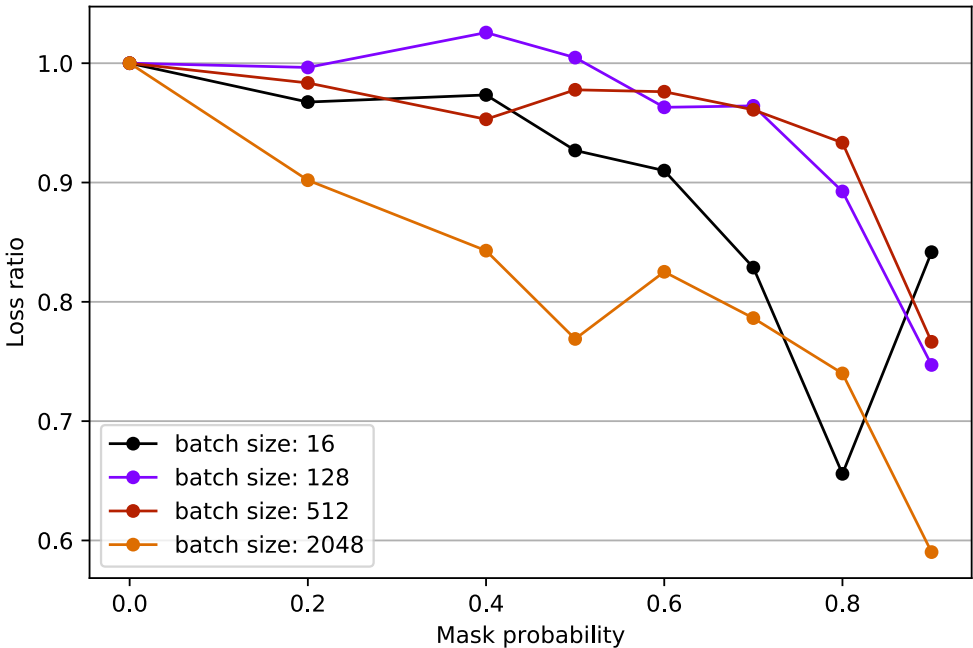


(c) Final image



(p_s, p_l)	Total	Small	Large
(0, 0)	45.9	64.7	74.1
(0, 0.75)	54.1	67.4	82.9
(0, 0.94)	49.1	55.3	88.2
(0.99, 0)	57.1	77.9	73.8
(0.999, 0)	57.4	68.2	84.1
(0.99, 0.75)	53.5	55.0	98.2
(0.99, 0.94)	66.2	78.7	84.6
(0.999, 0.75)	54.7	61.5	88.8
(0.999, 0.94)	55.0	58.5	95.0
512 × 512 crop	45.3	50.9	87.9
256 × 256 crop	34.3	53.9	55.3

← Significant →
improvements in
generalization



JETS

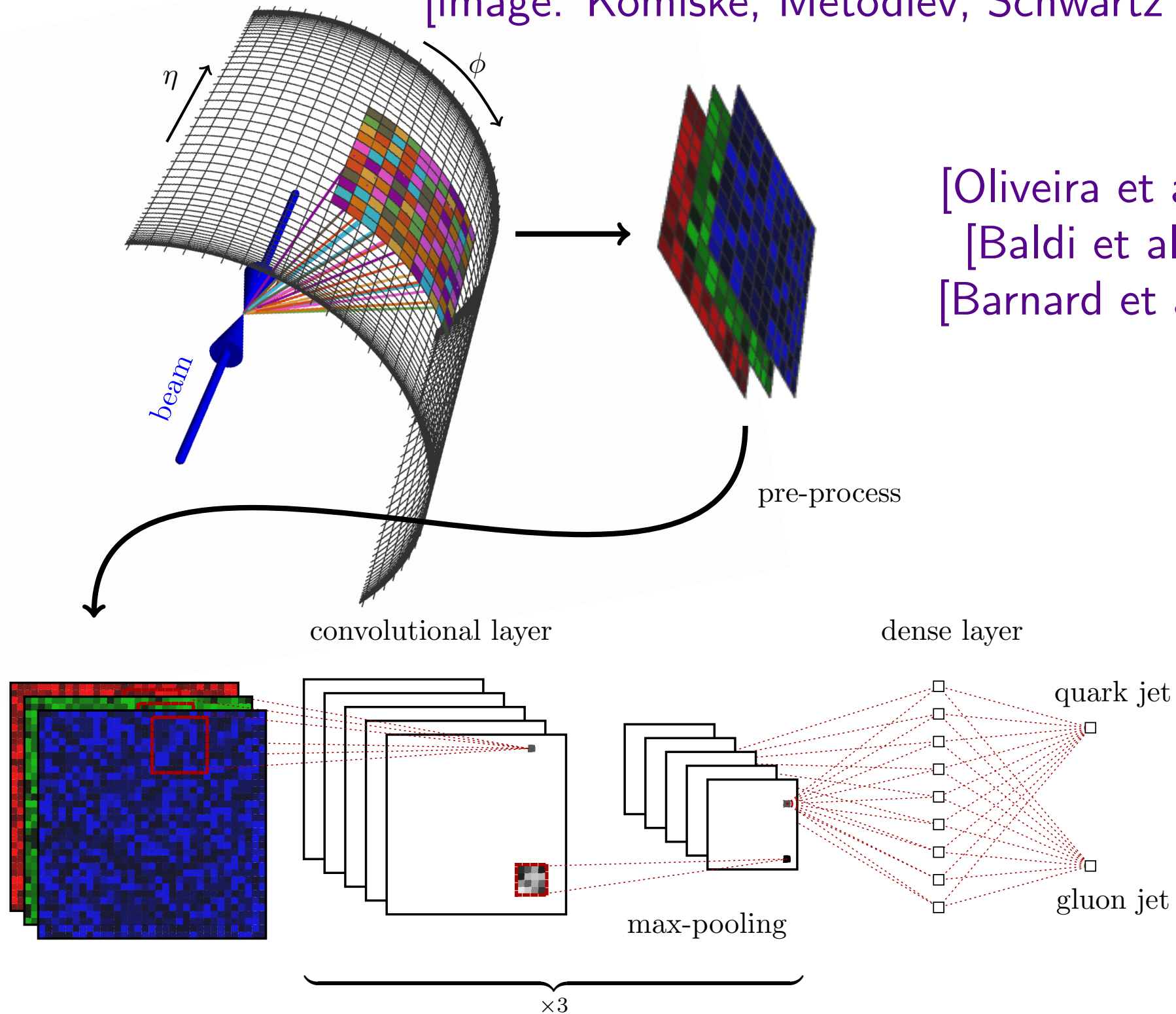
Run: 329716
Event: 857582452
2017-07-14 10:48:51 CEST



ATLAS
EXPERIMENT

JET IMAGES

[image: Komiske, Metodiev, Schwartz arxiv:1612.01551]



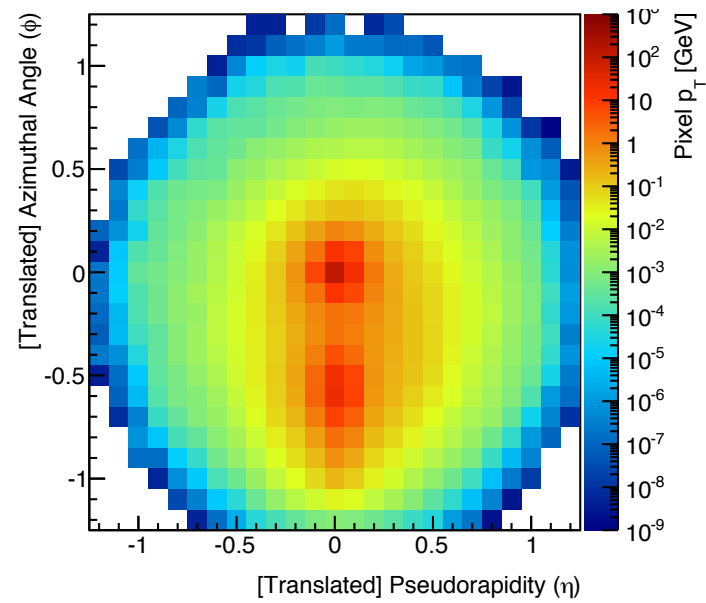
[Oliveira et al arXiv:1511.05190]

[Baldi et al arXiv:1603.09349]

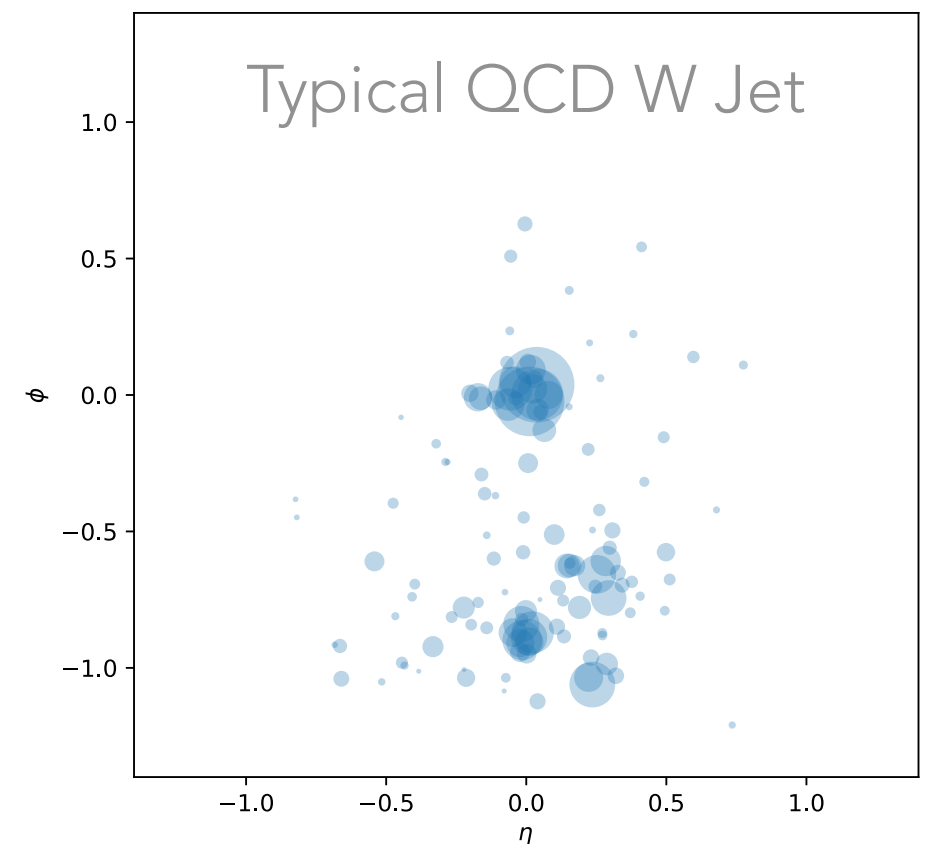
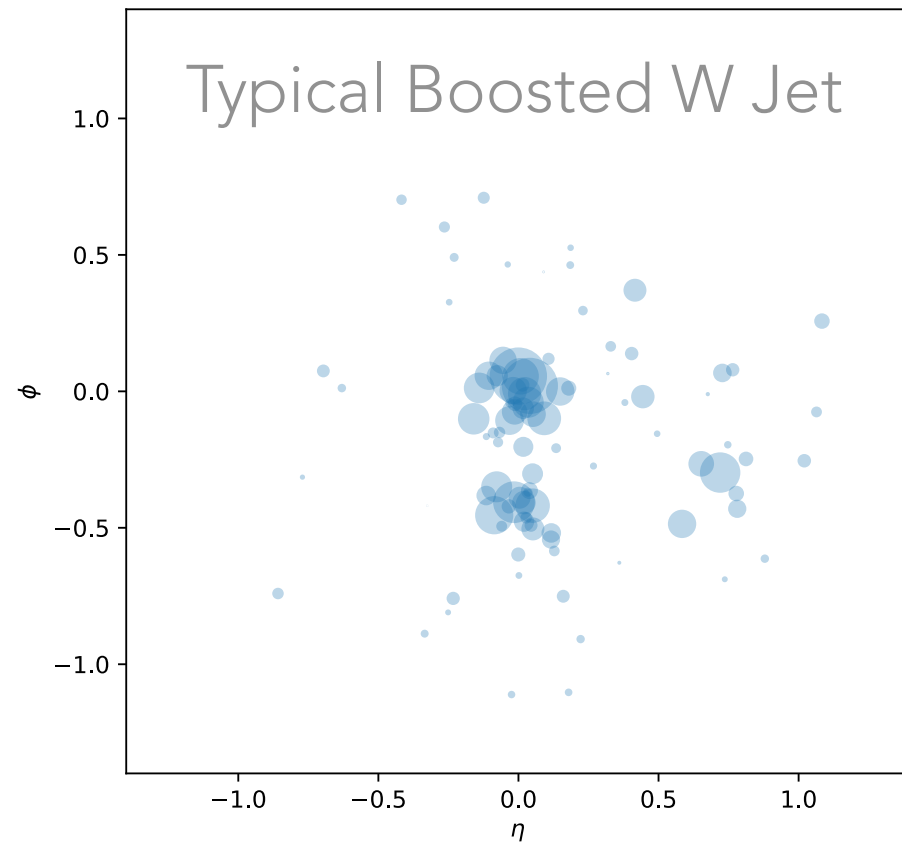
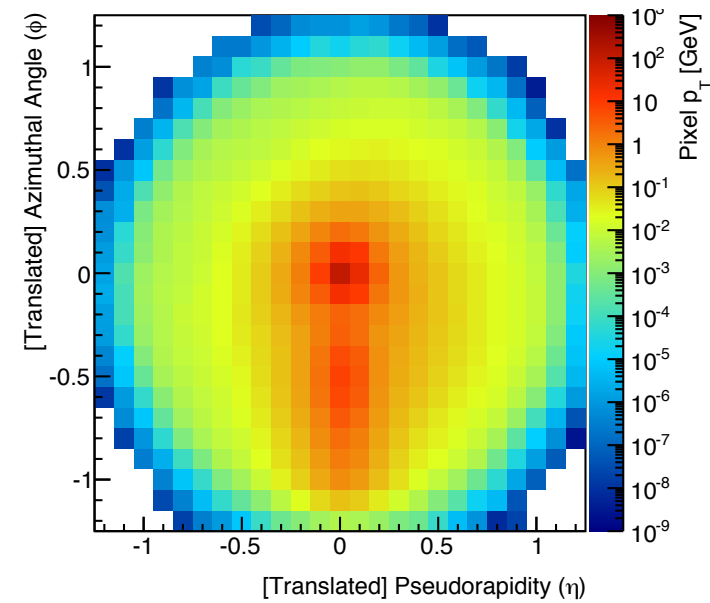
[Barnard et al arXiv:1609.00607]

CLASSIFICATION

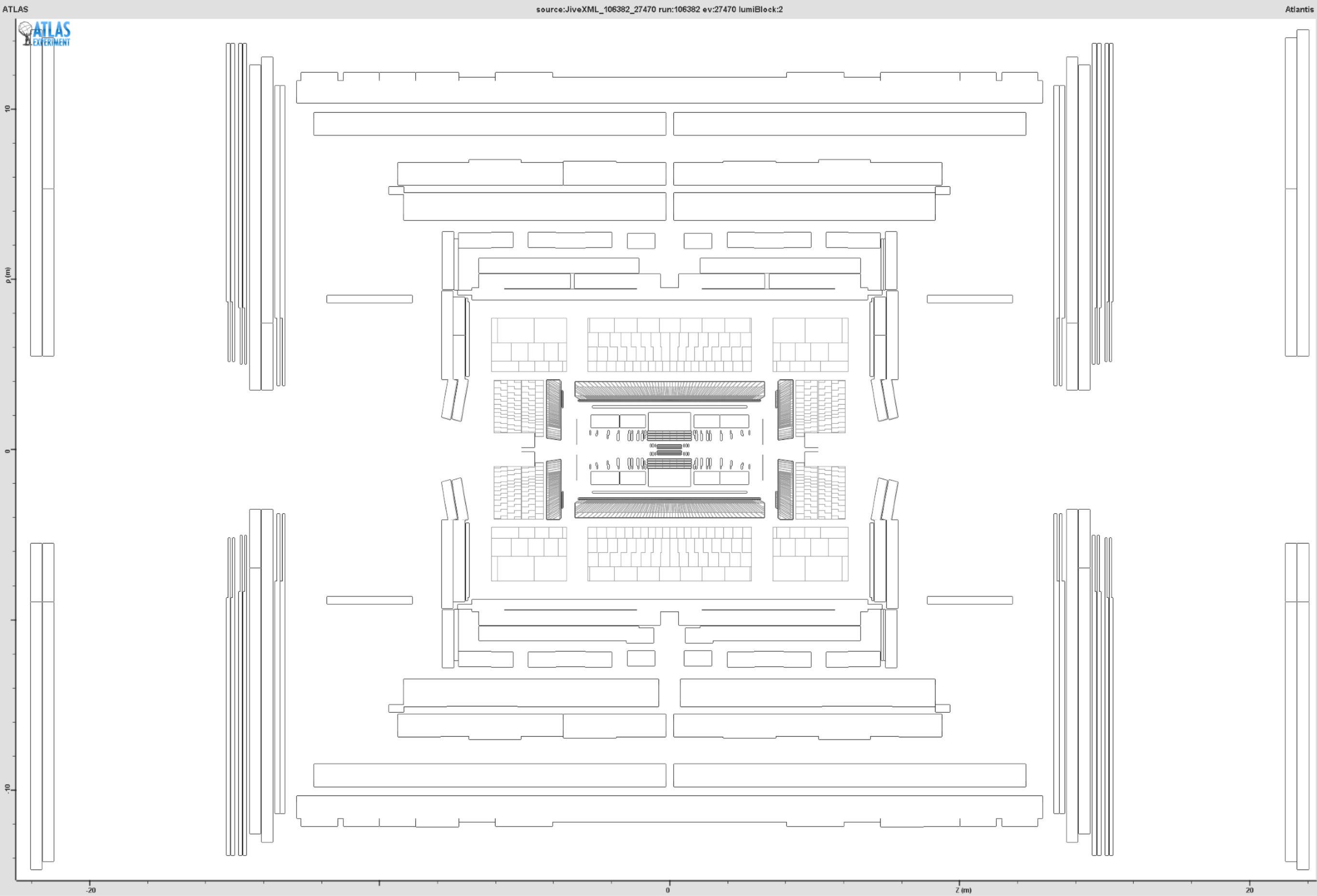
Average Boosted W Jet



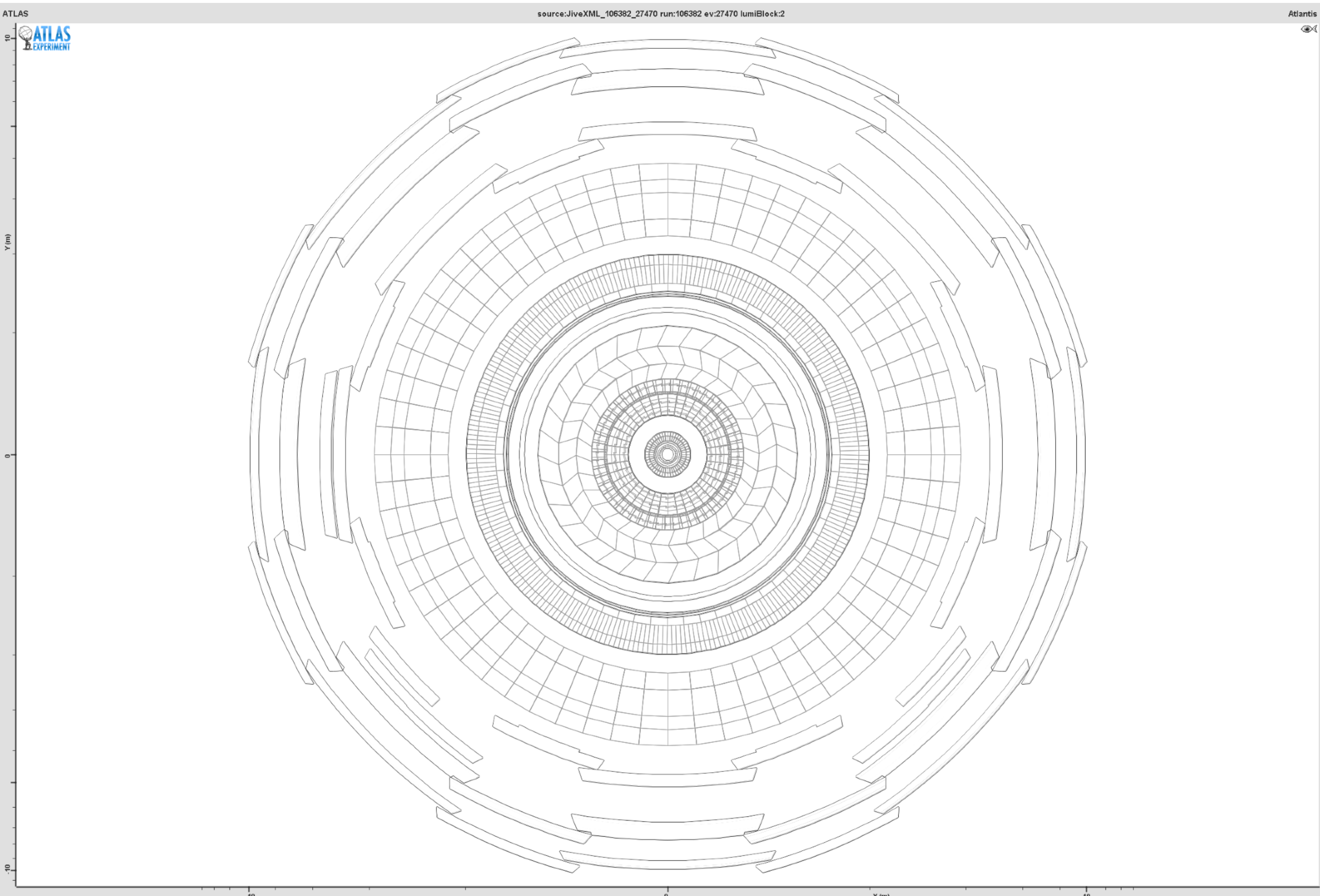
Average QCD Jet



NON-UNIFORM GEOMETRY



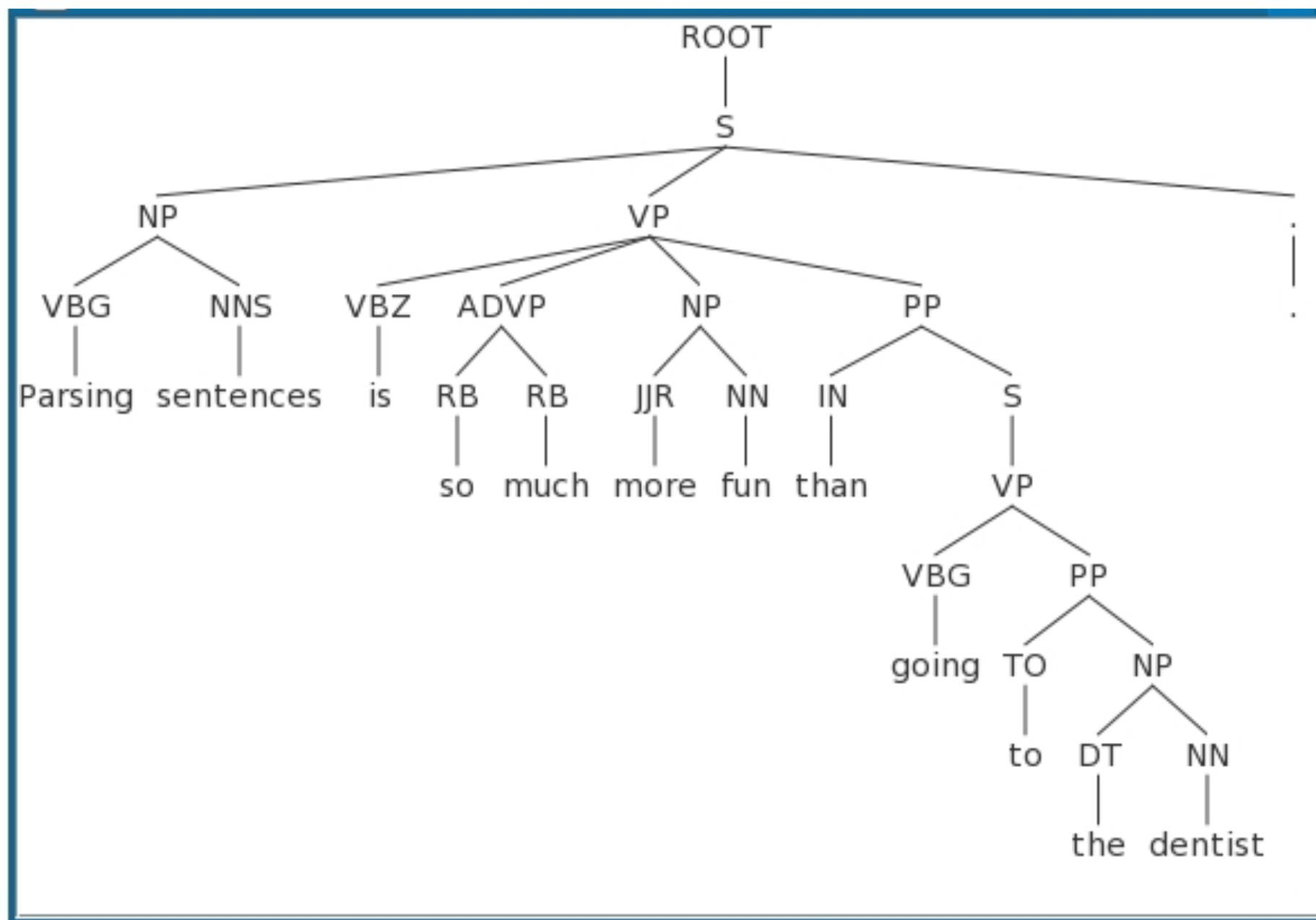
NON-UNIFORM GEOMETRY



FROM IMAGES TO SENTENCES

Recursive Neural Networks showing great performance for Natural Language Processing tasks

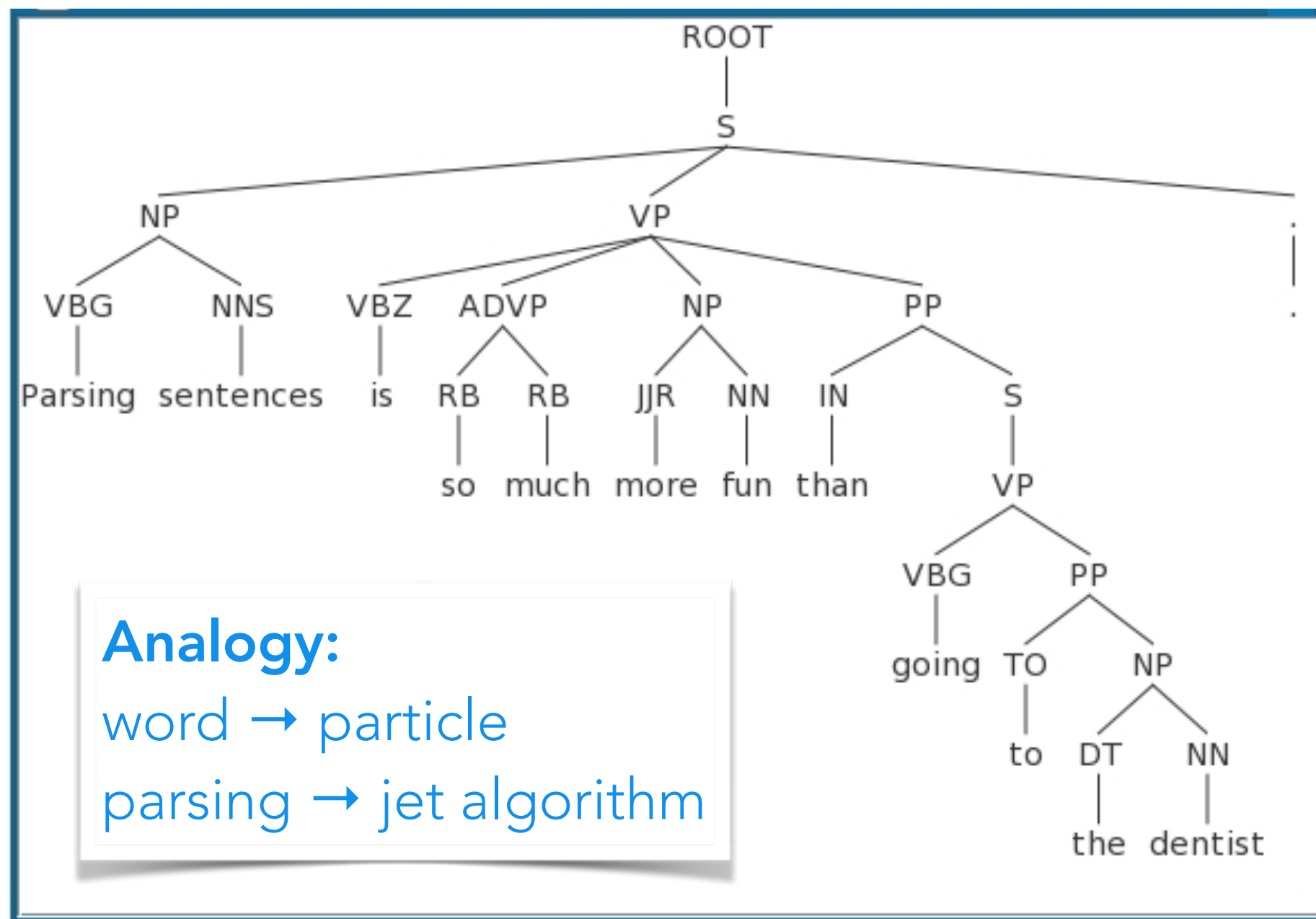
- neural network's topology given by parsing of sentence!



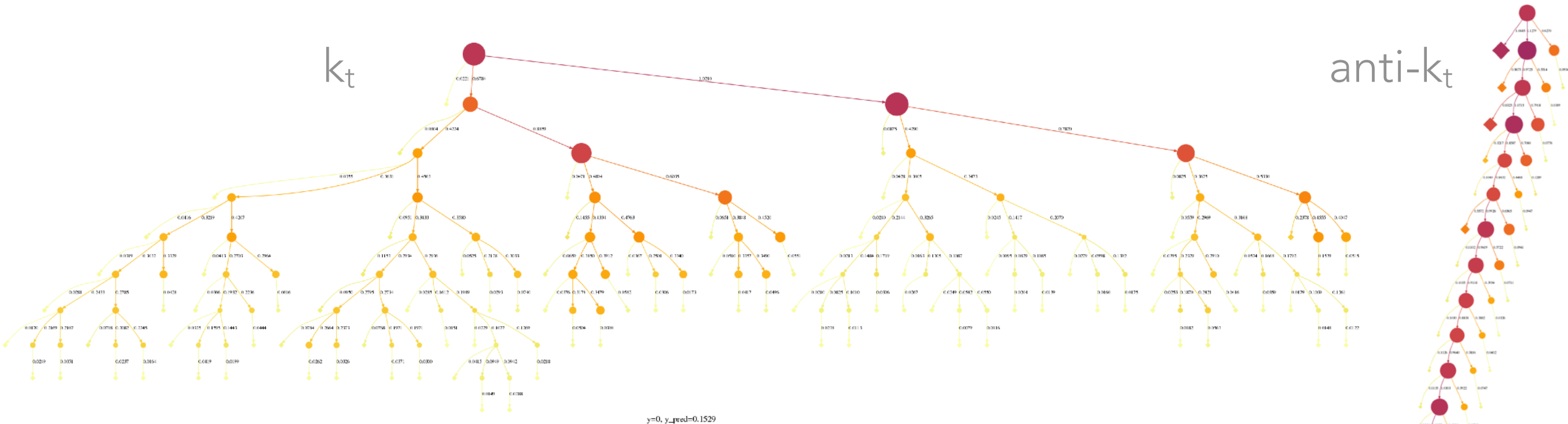
FROM IMAGES TO SENTENCES

Recursive Neural Networks showing great performance for Natural Language Processing tasks

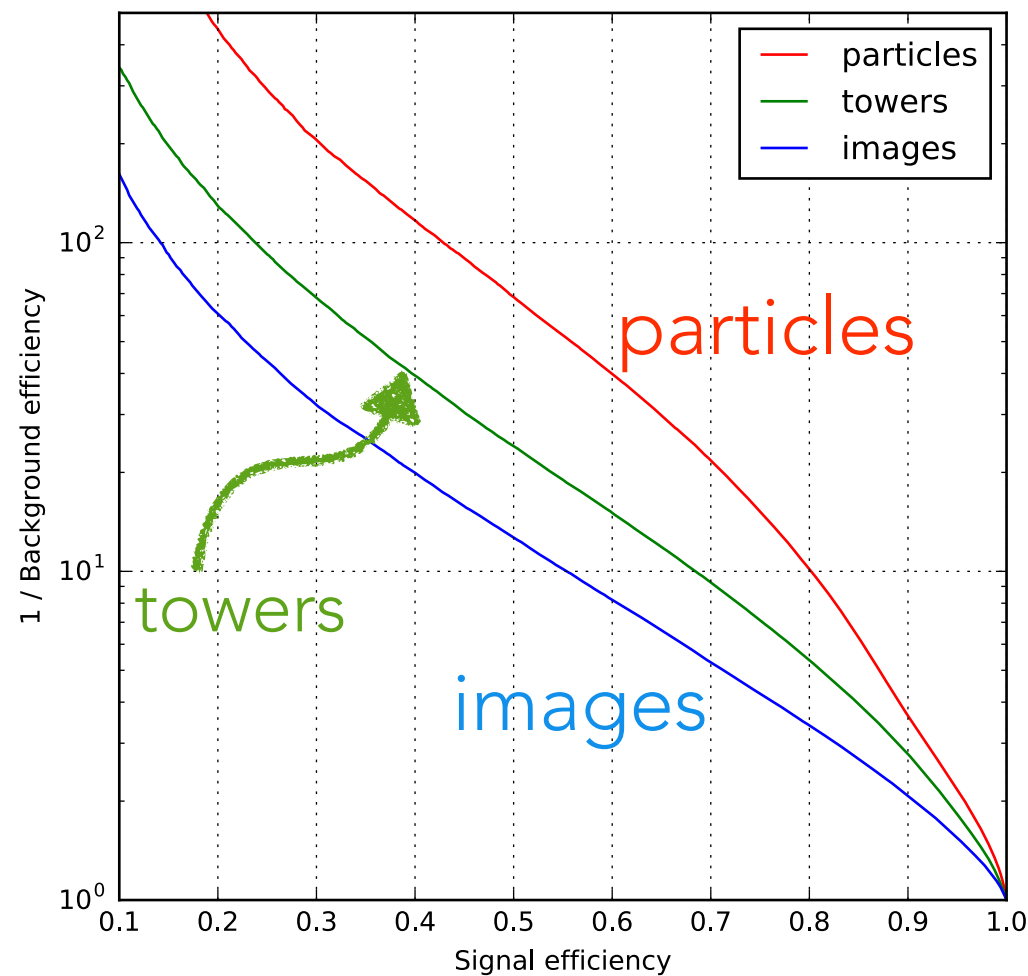
- neural network's topology given by parsing of sentence!



QCD-INSPIRED RECURSIVE NEURAL NETWORKS



$y=0, y_{\text{pred}}=0.1529$



- Generative process is a tree-like, $\sim$ stationary Markov Process
- Physics algorithms exist to estimate the tree
- Tree-RNN needs much less data to train!



DYNAMIC BATCHING

DEEP LEARNING WITH DYNAMIC COMPUTATION GRAPHS

Moshe Looks, Marcello Herreshoff, DeLesley Hutchins & Peter Norvig
Google Inc.
{madscience,marcelloh,delesley,pnorvig}@google.com

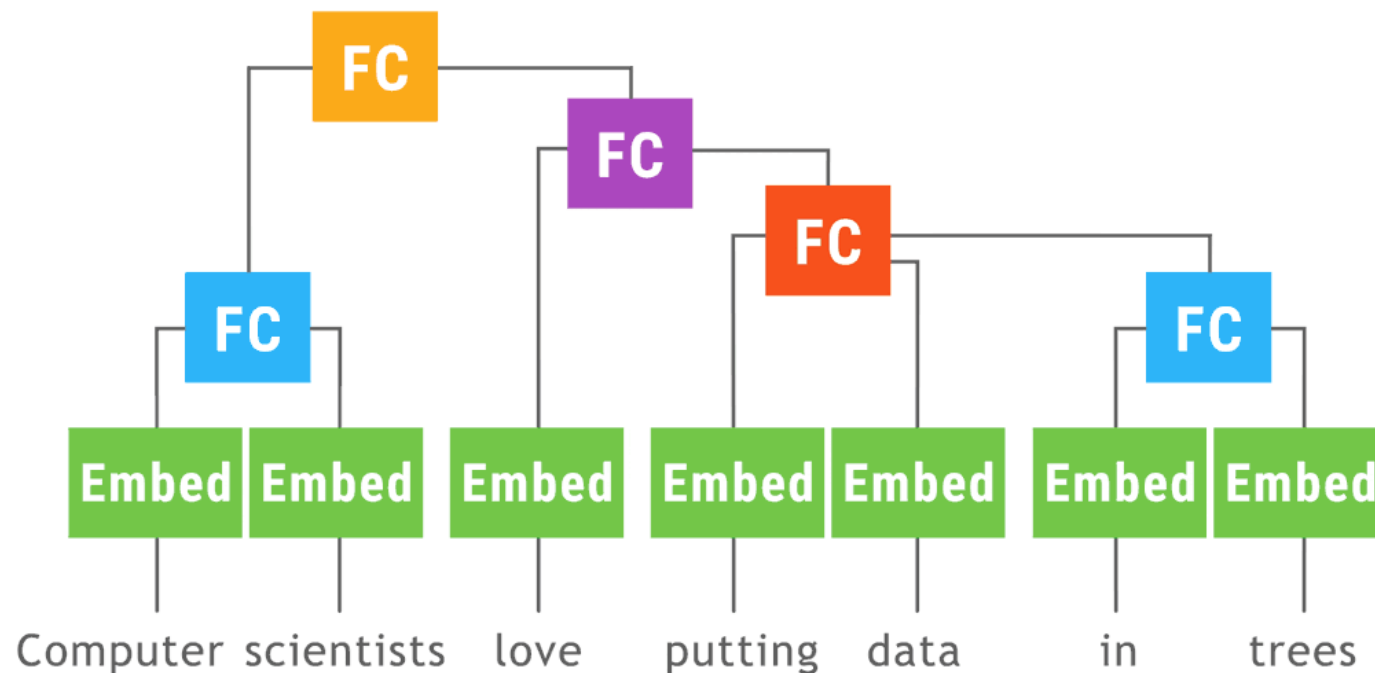
arXiv:1702.02181

TensorFlow Fold

TensorFlow Fold is a library for creating [TensorFlow](#) models that consume structured data, where the structure of the computation graph depends on the structure of the input data. For example, [this model](#) implements [TreeLSTMs](#) for sentiment analysis on parse trees of arbitrary shape/size/depth.

Fold implements [dynamic batching](#). Batches of arbitrarily shaped computation graphs are transformed to produce a static computation graph. This graph has the same structure regardless of what input it receives, and can be executed efficiently by TensorFlow.

- [Download and Setup](#)
- [Quick Start Notebook](#)
- [Documentation](#)



Excerpt from "QCD-Aware Recursive Neural Networks for Jet Physics"
arXiv:1702.00748

Appendix C: Implementation details

While tree-structured networks appear to be a principled choice in natural language processing, they often have been overlooked in favor of sequence-based networks on the account of their technical incompatibility with batch computation [33]. Because tree-structured networks use a different topology for each example, batching is indeed often impossible in standard implementations, which prevents them from being trained efficiently on large datasets. For this reason, the recursive jet embedding we introduced in Sec. III A would undergo the same technical issues if not implemented with caution.

In our implementation, we achieve batch computation by noticing that activations from a same level in a recursive binary tree can be performed all at once, provided all necessary computations from the deeper levels have already been performed. This principle extends to the synchronized computation across multiple trees, which enables the batch computation of our jet embeddings across many events. More specifically, the computation of jet embeddings is preceded by a traversal of the recursion trees for all jets in the batch, and whose purpose is to unroll and regroup computations by their level of recursion. Embeddings are then reconstructed level-wise in batch, in a bottom-up fashion, starting from the deepest level of recursion across all trees.

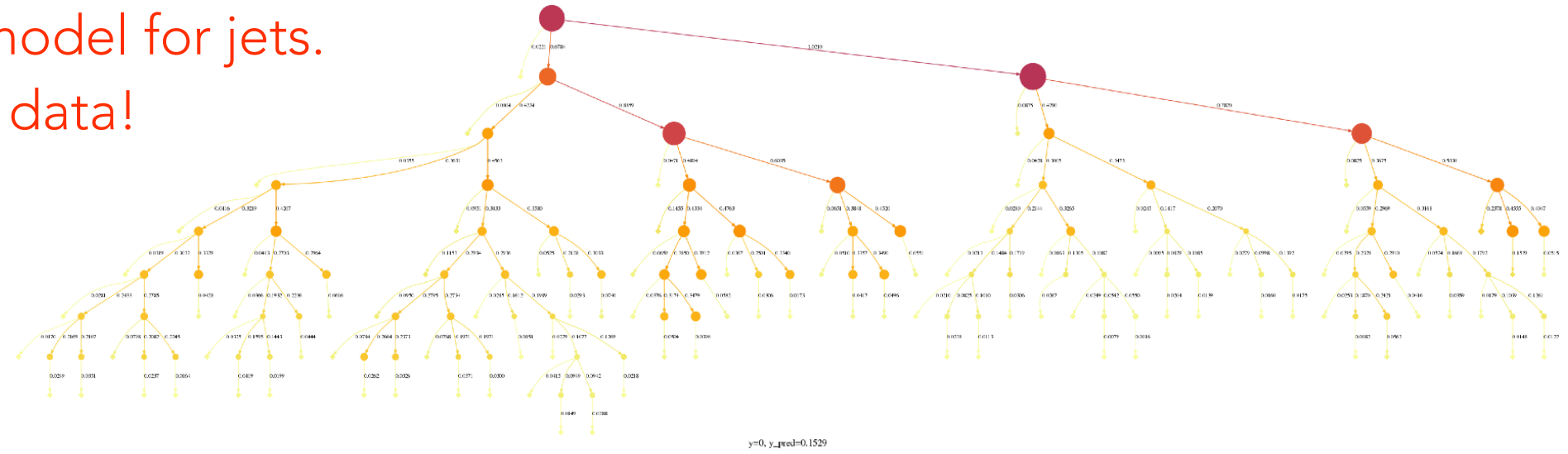
Finally, learning is carried out through gradients obtained by the automatic differentiation of the full model chain on a batch of events (i.e., the recursive computation of jet embeddings, the sequence-based recurrence to form the event embeddings, and the forward pass through the classifier). The implementation is written in native Python code and makes use of Autograd [42] for the easy derivation of the gradients over dynamic structures. Code is available at ² under BSD license for further technical details.

² <https://github.com/glouppe/recnn>

[33] Samuel R. Bowman, Jon Gauthier, Abhinav Rastogi, Raghav Gupta, Christopher D Manning, and Christopher Potts. A fast unified model for parsing and sentence understanding. *arXiv preprint arXiv:1603.06021*, 2016.

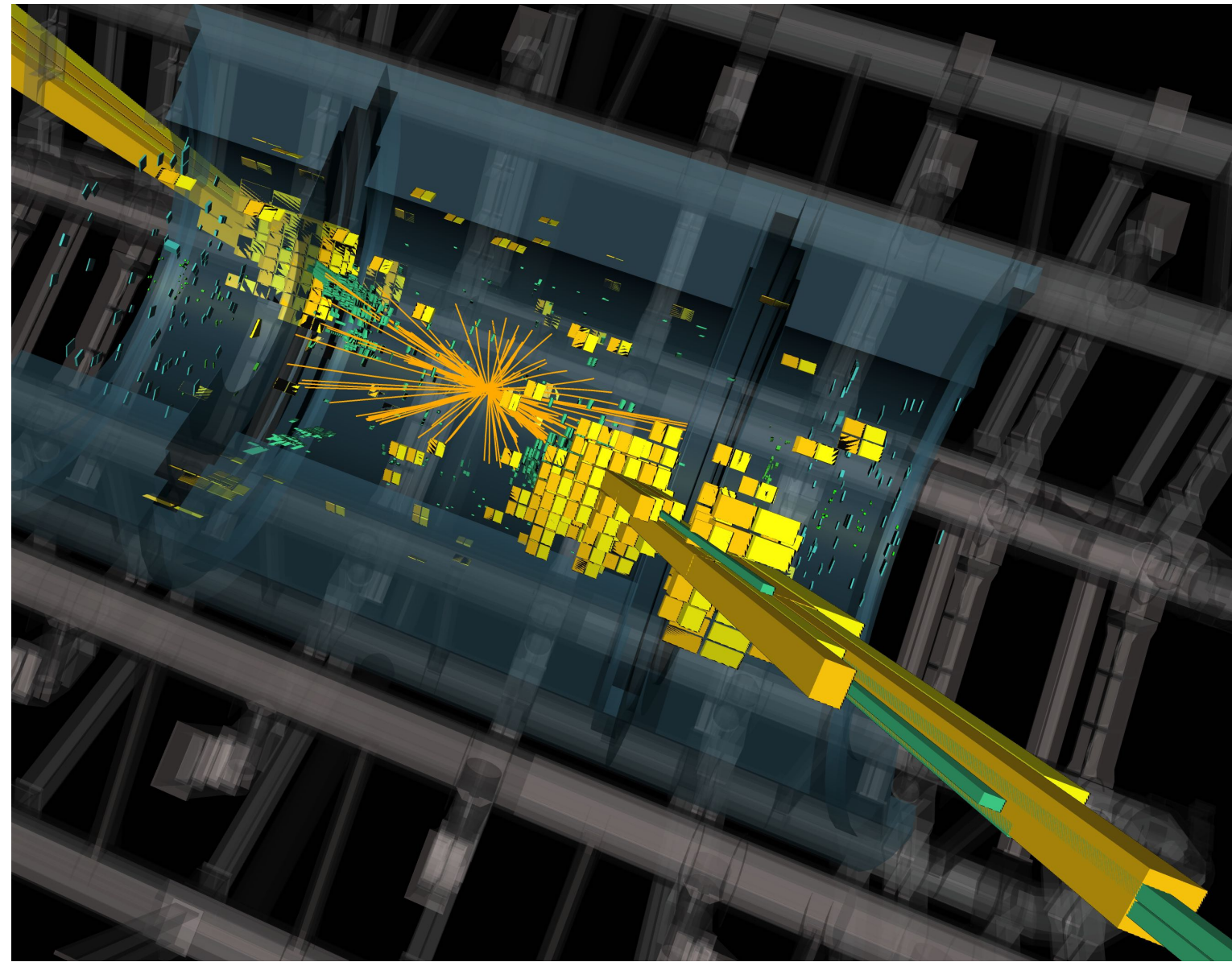
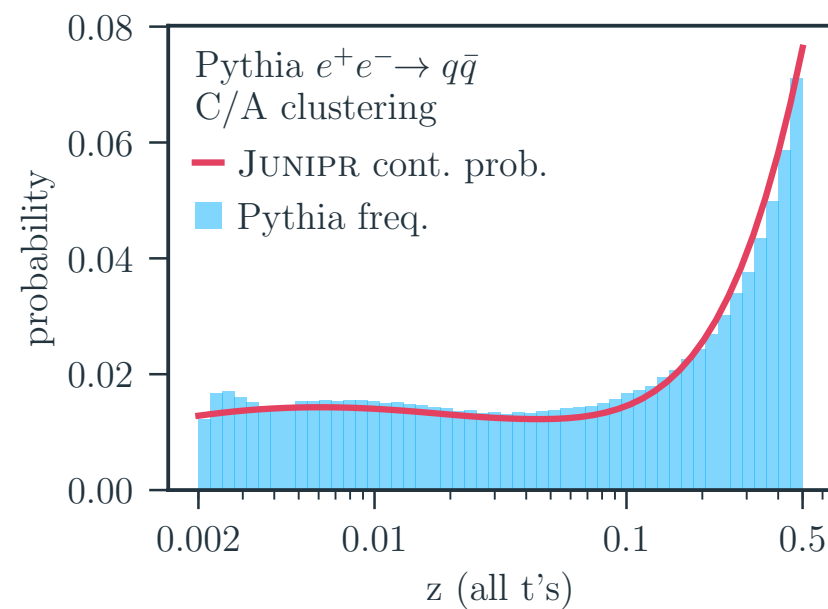
GENERATIVE MODELS (FROM DATA)

JUNIPR is a generative model for jets.
Can train on real data!



$$P_{\text{jet}}(\{p_1, \dots, p_n\}) = \left[\prod_{t=1}^{n-1} P_t(k_1^{(t+1)}, \dots, k_{t+1}^{(t+1)} | k_1^{(t)}, \dots, k_t^{(t)}) \right] \times P_n(\text{end} | k_1^{(n)}, \dots, k_n^{(n)}).$$

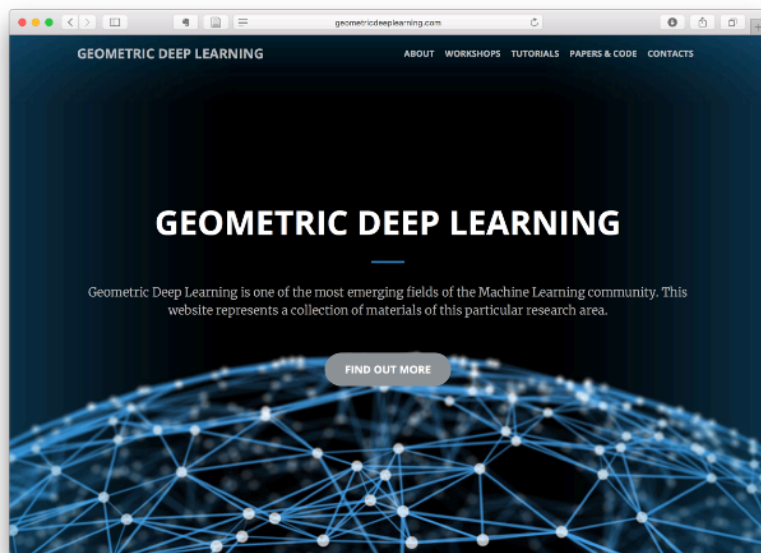
... and it is interpretable



Neural Message Passing for Jet Physics

Isaac Henrion, Johann Brehmer, Joan Bruna,
Kyunghyun Cho, Kyle Cranmer, Gilles Louppe,
Gaspar Rochette

Courant Institute & Center for Data Science



Isaac Henrion

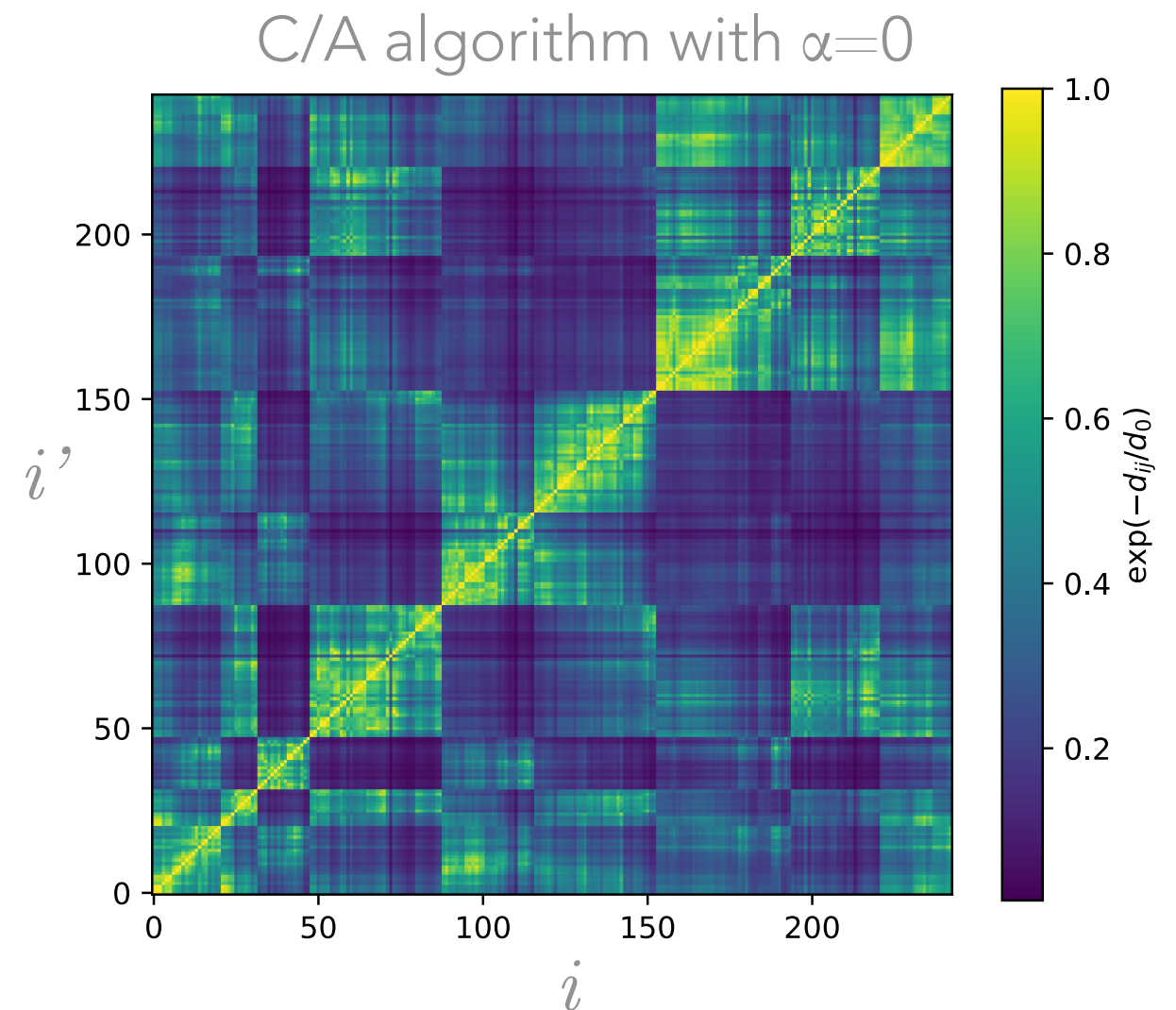
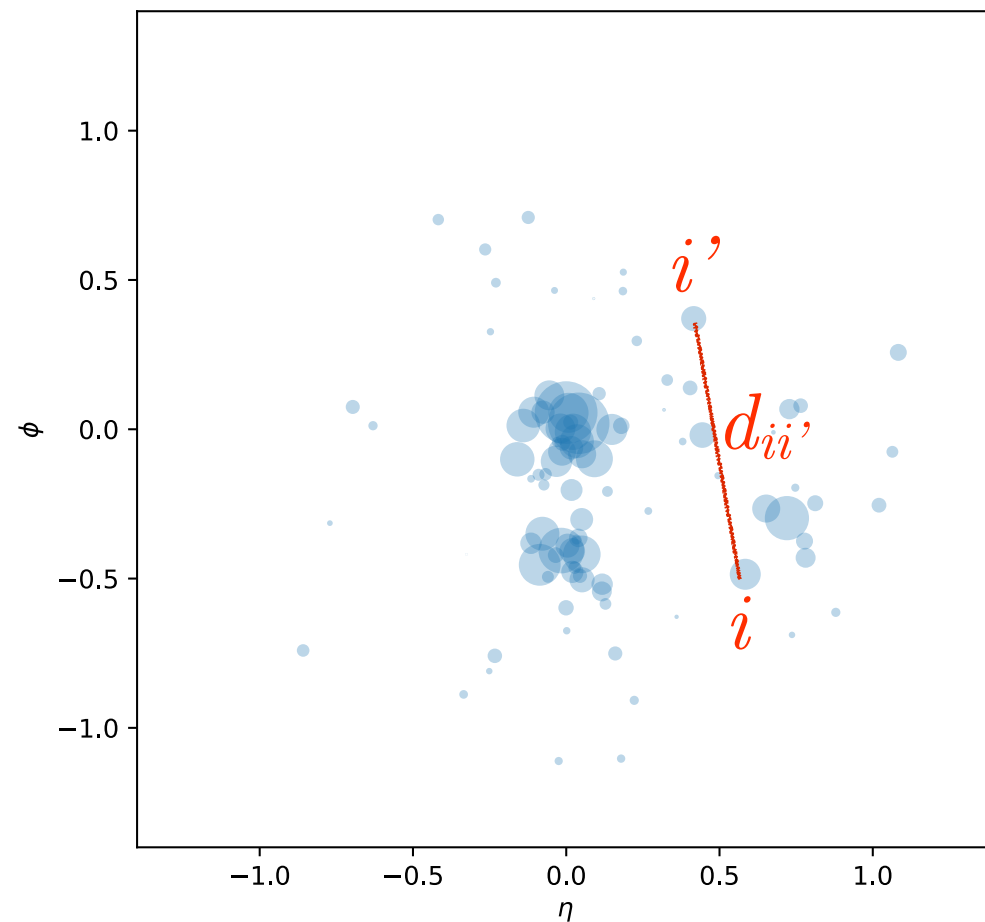
Paper: https://dl4physicalsciences.github.io/files/nips_dlps_2017_29.pdf

Talk: https://dl4physicalsciences.github.io/files/nips_dlps_2017_slides_henrion.pdf

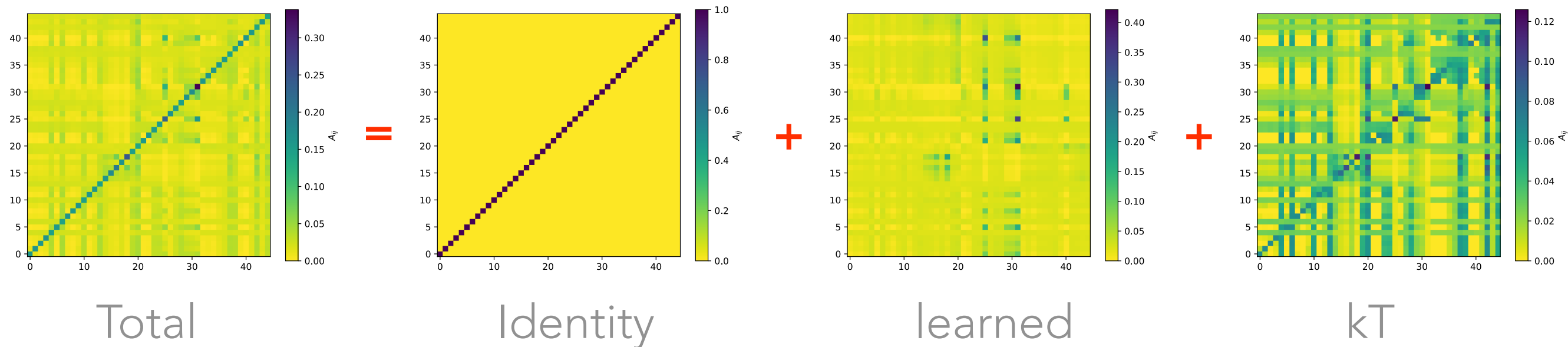
JETS AS A GRAPH

Using message passing neural networks over a fully connected graph on the particles

- Two approaches for adjacency matrix for edges
 - **import** physics knowledge by using metric of jet algorithms $d_{ii'}^\alpha = \min(p_{ti}^{2\alpha}, p_{ti'}^{2\alpha}) \frac{\Delta R_{ii'}^2}{R^2}$
 - learn adjacency matrix and **export** new jet algorithm

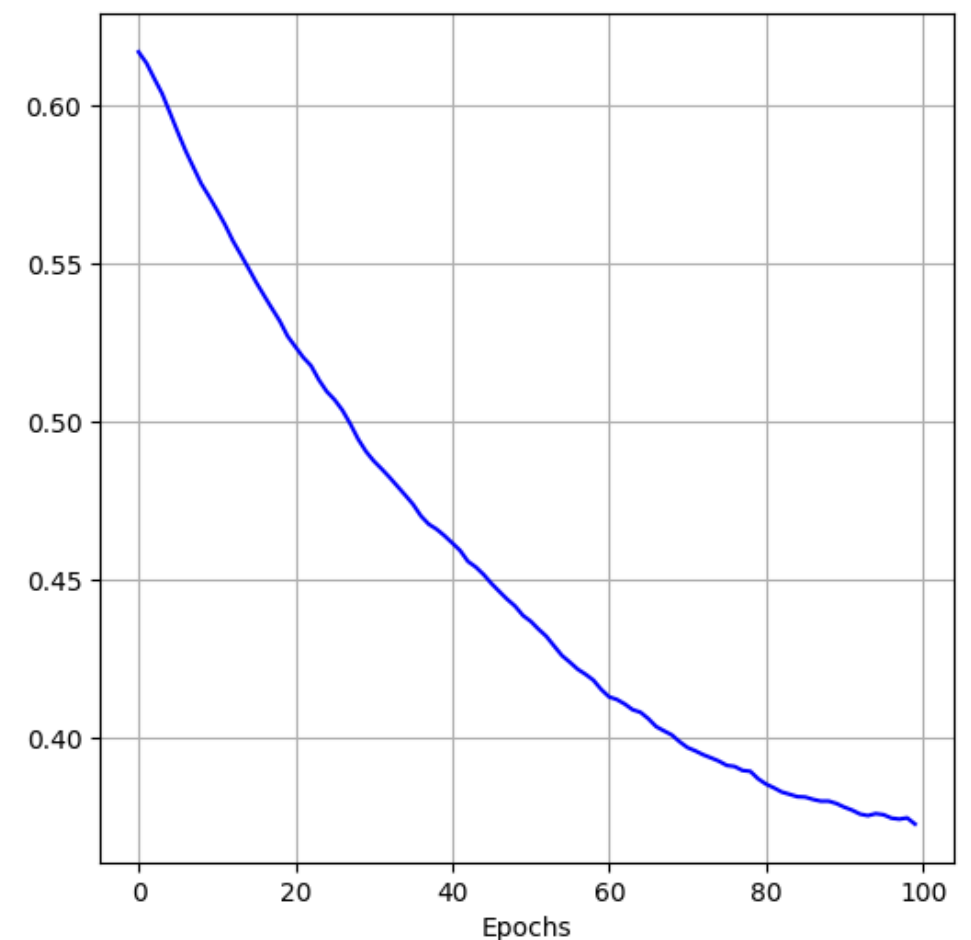


LEARNED ADJACENCY DEFINES A NEW JET ALGORITHM



The adjacency matrix starts off as purely physics-inspired (kT) and then learned adjacency matrix becomes more important

- weight given to physics adjacency matrix slowly decays, but stays relevant →



CONCLUSION

The developments in machine learning have the potential to effectively bridge the microscopic - macroscopic divide & aid in the inverse problem.

- leverage expert knowledge of the generative process
- learn surrogates that extract relevant features for inference task

Several strategies to incorporate domain knowledge into the model

- these define cross-disciplinary research directions

Non-traditional setting for publishing & collaboration