Theoretical Physics for Machine Learning – ICLR 2019

Workshop Proposal for the Seventh International Conference on Learning Representations

Daniel A. Roberts,[a] David J. Schwab,[b] Riccardo Zecchina,[c] and Yann LeCun,[ad]

[a] *Facebook AI Research, Facebook, New York, NY 10003, USA*
[b] *Initiative for the Theoretical Sciences, CUNY Graduate Center, New York, NY 10016, USA*
[c] *Department of Decision Sciences, Bocconi University, Milano MI, Italy*
[d] *Courant Institute of Mathematical Sciences, New York University, New York, NY 10003, USA*

danr@fb.com, dschwab@gc.cuny.edu, riccardo.zecchina@unibocconi.it, yann@fb.com

Abstract

At the broadest level, the program of machine learning is the selection and fitting of mathematical models to describe desired phenomena, often coded in a probability distribution. This can also be considered to be a description of the program of physics, provided we specify that the phenomena are *natural*. It's our view that the tools, intuitions, and frameworks of physics can have a much broader domain of applicability than traditionally considered, and the purpose of this workshop is to investigate such connections to machine learning and artificial intelligence. In particular, we are interested in exploring how perspectives from physics can help our understanding of the success of modern machine learning and guide the development of future methods.

Proposal

Machine learning is a field of computer science, mathematics, statistics, and—we contend—physics, that effectively solves tasks too complicated to program directly from first principles. Machine learning has undergone a renaissance in recent years, with a succession of experimental triumphs [1–5] in the realm of *deep learning* [6], a widely successful framework for building multi-purpose models using neural networks. Much of this success, however, has been empirical and lacks a satisfying theoretical framework for understanding why certain approaches succeed while others fail.

The trail of groundbreaking empirical results over the past half-dozen years represents an exciting opportunity for theoretical physicists. Physics thrives on a tight interplay between theory and experiment, and we see machine learning as providing a trove of experimental results in search of theoretical understanding. Theoretical physics is not only about building the best models of natural phenomena, but also about understanding the relationship between the details, e.g. parameters of the models, and the statistics of the large-scale phenomena, such as unexpected generalization performance.

This workshop will investigate the use of ideas from theoretical physics —in particular, high energy theory, condensed matter theory, and statistical mechanics—in machine learning domains such as image classification [1], game playing [2–5], and language modeling [7]. The experimental "data" that physicists collect is typically different from that ordinarily collected for machine learning tasks, with the goal for physicists to be able to open the "black box" [8] of their models so they may be intuitively understood. An improved theoretical understanding should yield more effective approaches to challenging problems in machine learning. Although this research direction is still new, there have already been fruitful developments [8–20], and we anticipate talks focusing on successful applications of this paradigm.

Potential Interest

Machine learning has already attracted recent attention from physicists, e.g. see [21], a review of machine learning written for physicists and co-written by one of the organizers. As further evidence, the figure below shows the overall exponential increase in arXiv preprints with "machine learning" or "deep learning" in the title or abstract, and in green shows the subset of physics articles tracking that trend. Many of these articles apply machine learning tools to physics; but some are about the topic of this workshop, applying physics tools, principles, and methodology to machine learning.

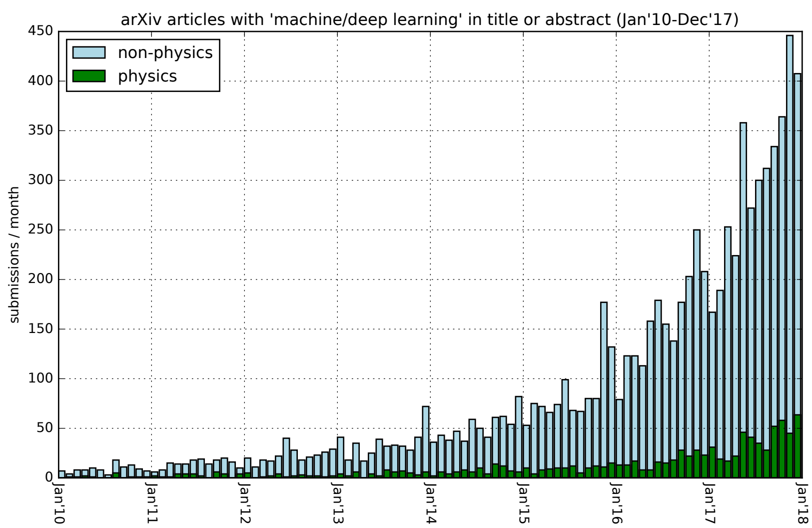

Figure 1: Preprints with "machine learning" or "deep learning" in the abstract or title posted on the arXiv over time. We are grateful to Paul Ginsparg for making this plot.

Motivated by the growing excitement in the physics community, for over a year one of the organizers has been running an online journal club, *hep-ai* (`https://hep-ai.org`), focused on discussions of machine learning and AI generally led by physicists with interests and sympathies in AI. This journal club has grown quickly to over 150 members drawn from graduate students, postdocs, and professors, and has fostered useful interactions between the members of the machine learning and physics communities. As on average 30 members tend to participate in the talks and discussions held every other week, we anticipate that many member of the *hep-ai* group will be enthusiastic about participating in such a workshop.

We also feel that this represents an opportunity for the ICLR community to encourage the broader interested group of physicists to attend a mainstream machine learning conference and interact with researchers working on AI. We hope that by mixing AI researchers and physicists we can jointly make progress on better understanding machine learning algorithms and developing principles behind why they work. This in turn should provide guidance on building new models that exhibit unprecedented performance on a wider variety of tasks.

Potential Topics

Let us discuss a few directions that we anticipate may be covered over the course of the workshop.

- *Prior knowledge: Symmetry and beyond.* In order to improve learning on a particular task, one must make use of the particular properties of that task, essentially embedding prior knowledge into the model. For example, leveraging a problem's symmetries—e.g. by introducing a network architecture with constraints on the parameters that reflect that symmetry—is crucial for obtaining competitive performance in deep learning. For example, recurrent neural networks [22, 23], reflect the time translation symmetry of natural language texts, and analogous properties are used in convolutional networks [24] for image classification. Given a problem with a potentially new type of symmetry, it is unknown in general how best to incorporate that symmetry into the architecture. Our experience in physics indicates that symmetry can be used as an organizing principle that guides us in constructing models. Can symmetries play a similarly powerful role in machine learning?

  In physics, prior knowledge can be embedded in *tensor networks*, a very successful variational ansatz motivated by physical constraints on entanglement. They give an exponential reduction in the number of parameters required to represent a state [25, 26]. This intuition has even played a major role in understanding black hole physics in the context of holography [27–29]. Similar prior knowledge is available for classical probability distributions, and should be equally important for developing and understanding machine learning models. There already is some very promising initial work in this direction, e.g. [30, 31] and [32]. What other priors can physics provide for machine learning?

- *The dynamics of optimization: SGD and landscapes.* Optimization algorithms used in deep learning constitute a class of novel dynamical systems. Stochastic Gradient Descent (SGD) and its variants are the leading algorithms used to optimize deep networks. These dynamics live on a complex and poorly understood loss landscape. The dynamics of SGD itself may be responsible for the generalization properties of deep networks [33], but the mechanism for this is not well understood. In [34] fluctuation-dissipations relations for SGD were derived, leading to a novel way of annealing the learning rate. Can we make further progress in the non-equilibrium setting? What about when the underlying distribution is time-dependent? Furthermore, can our understanding of very general dynamical frameworks (i.e., action principles and their connections to symmetries) give more than a descriptive explanation?

  Focusing on the loss landscape, [35] suggested that typical deep learning models shares some properties with that of glassy systems. More recently, building on statistical physics analytical and algorithmic tools, it was shown that the landscape of neural networks models compared to that of spin glasses, is enriched by the existence of high local entropy regions, or wide valleys leading to wide flat minima, which are rare and yet accessible by learning algorithms [36–38]. These dense regions have good generalization properties. Glass physicists have also begun to apply tools for the characterization of glassy dynamics to understanding learning on complicated neural network loss landscapes, identifying important differences when models are over- vs. under- parameterized [39, 40] .

Analyzing nontrivial phase structures and complicated potentials is an area of expertise for physicists, and there's a clearly a wealth of additional experiments, analysis, and theory necessary to understand these phenomena. Can insights developed in these communities carry over to the landscapes encountered in machine learning?

- *Effective Theory, Renormalization, and Machine Learning.* The study of physics can, in some sense, be described as the search for simplified descriptions of nature that retain predictive power. In the context of thermodynamically large systems, this is exemplified by the renormalization group, which makes rigorous the terms "relevant" and "irrelevant" as applied to the degrees of freedom of the problem. Renormalization allows one to determine what microscopic aspects of a physical system govern macroscale behavior. Is this perspective useful for understanding how high-level knowledge can be extracted by building up from pixel-level information? Since modern deep neural networks utilize large numbers of neurons, can ideas from renormalization be used to provide simplified yet predictive descriptions of their behavior? Connections have already been made between renormalization and deep learning [41], but there's much more that can be done in establishing effective descriptions of the important models of interest.

Organizers

The organizers have had many years of experience working as researchers across the fields of artificial intelligence and theoretical physics.

Dan Roberts

Dan is AI research scientist at Facebook AI Research (FAIR) in NYC. His academic background in theoretical physics includes a Ph.D. from MIT and a postdoc at the Institute for Advanced Study in Princeton, New Jersey.

His research has focused on the interplay between physics and computation. In theoretical physics, he studied the relationship between black holes [42], quantum chaos [43–45], computational complexity [29,46], randomness [47], and how the laws of physics are related to fundamental limits of computation. As an AI researcher at FAIR, he's begun to apply tools, insights, and frameworks from theoretical physics to gain insight into machine learning and artificial intelligence.

Additionally, Dan founded and is the organizer of *hep-ai* (`https://hep-ai.org`), an online journal club focused on discussions of machine learning and AI papers from a theoretical physicist's perspective. The club has over 150 members consisting mostly of physics graduate students, postdocs, and professors, and the club's talks occurring every other week tend to attract on average over 30 participants.

David Schwab

David is an Assistant Professor of Biology and Physics in the Initiative for the Theoretical Sciences at the CUNY Graduate Center. He is also core faculty at the Center for the Physics of Biological Function, an NSF-funded Physics Frontiers Center joint between Princeton University and CUNY. He is a Simons Investigator in the Mathematical Modeling of Living Systems and a recipient of a K25 Fellowship from the NIH. Previously, he was an Assistant Professor of Physics at Northwestern University, and a postdoctoral research scholar at Princeton University. He has a Ph.D. in condensed matter theory and theoretical biophysics from UCLA.

He co-authored a highly influential paper relating a form of unsupervised deep learning to the renormalization group from statistical physics which attracted the attention of a number of physicists [41]. He is also known for his work on using tensor networks from many-body quantum systems for supervised learning problems [30]. David recently co-authored an extensive review on machine learning for physicists [21].

Outside of machine learning, David is known for his work on the fundamental energetic costs of computation in biological systems. He also proposed an explanation for the ubiquitous observation of Zipf's law in high-dimensional data. Finally, David has worked extensively with experimental collaborators to understand the sophisticated computations performed by directionally-selective ganglion cells.

Riccardo Zecchina

Riccardo Zecchina is a Full Professor at Bocconi University and holder of the Vodafone Chair in Machine Learning and Data Science. Previously, he was a full Professor at Politecnico di Torino. He has been research scientist and head of the Statistical Physics Group at the International Centre for Theoretical Physics in Trieste (1997-2007).

His research interests lie at the interface between statistical physics, computer science, machine learning and information theory. He's interested in fundamental aspects as well as in the development of statistical physics algorithms for optimization and learning problems. His current research interests are focused on machine learning and out-of-equilibrium phenomena. His past studies include combinatorial optimization, probabilistic and message-passing algorithms, statistical physics of complex systems (disordered systems), out-of-equilibrium dynamics, analysis of algorithms, and interdisciplinary applications of statistical physics (learning algorithms, inverse problems in systems biology, source coding, game theoretical models).

Yann LeCun

Yann LeCun is VP and Chief AI Scientist at Facebook and Silver Professor at NYU affiliated with the Courant Institute and the Center for Data Science. He was the founding Director of Facebook AI Research and of the NYU Center for Data Science. He received an EE Diploma from ESIEE (Paris) in 1983, a PhD in Computer Science from Universit Pierre et Marie Curie (Paris) in 1987. After a postdoc at the University of Toronto, he joined AT&T Bell Laboratories. He became head of the Image Processing Research Department at AT&T Labs-Research in 1996, and joined NYU in 2003 after a short tenure at the NEC Research Institute. In late 2013, LeCun became Director of AI Research at Facebook, while remaining on the NYU Faculty part-time. He was visiting professor at Collge de France in 2016.

His research interests include machine learning and artificial intelligence, with applications to computer vision, natural language understanding, robotics, and computational neuroscience. He is best known for his work in deep learning and the invention of the convolutional network method which is widely used for image, video and speech recognition. He is a member of the US National Academy of Engineering, the recipient of the 2014 IEEE Neural Network Pioneer Award, the 2015 IEEE Pattern Analysis and Machine Intelligence Distinguished Researcher Award, the

2016 Lovie Award for Lifetime Achievement, the University of Pennsylvania Pender Award, and honorary doctorates from IPN, Mexico and EPFL.

Logistics

We propose to have a full day consisting of six half-hour invited talks, four half-hour sessions of contributed talks (which we may subdivide into additional talks, depending on interest), an hour long poster session, and a panel discussion. A tentative schedule for the day is suggested below.

Tentative Schedule

- 09:00 - 09:30: Invited talk 1
- 09:30 - 10:00: Contributed talk(s) 1
- 10:00 - 10:30: Invited talk 2
- 10:30 - 11:00: Coffee break
- 11:00 - 11:30: Invited talk 3
- 11:30 - 12:00: Contributed talk(s) 2
- 12:00 - 12:30: Invited talk 4
- 12:30 - 13:30: Lunch
- 13:30 - 14:00: Invited talk 5
- 14:00 - 14:30: Contributed talk(s) 3
- 14:30 - 15:00: Invited talk 6
- 15:30 - 16:30: Poster session
- 16:30 - 17:00: Contributed talk(s) 4
- 17:00 - 18:00: Panel discussion

Attendance and Speakers

A rough prediction for the size of the workshop is 60 people. This is an estimate given the size (150 people) and rate of growth of the *hep-ai* journal club discussed above, and given the strong interest in the machine learning community of interacting with physicists. As interest in the physics community in machine learning and in the machine learning community in physics ideas continues to increase, this may be somewhat of an underestimate.

Furthermore, we are considering the following researchers as keynote speakers for the six invited talk slots.

- Yasaman Bahri, Google Brain

- Kyle Cranmer, NYU
- Surya Ganguli, Stanford
- Jared Kaplan, Johns Hopkins University
- Pankaj Mehta, Boston University
- Jeffrey Pennington, Google Brain
- Xiaoliang Qi, Stanford University
- Levent Sagun, ENS-Paris, CEA-Saclay, and EPFL
- Jesse Thaler, MIT
- Naftali Tishby, Hebrew University of Jerusalem
- Sho Yaida, FAIR
- Lenka Zdeborová, CEA/Saclay

Plan for Funding

Our plan is to request funding from various organizations that support that these sorts of workshops. We expect that Facebook is very likely to contribute, and we will approach other organizations once we have a better sense of our expenses. We will use the funds to pay for flights and accommodations for our speakers.

References

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, pp. 1097–1105. 2012.

[2] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, *et al.*, "Human-level control through deep reinforcement learning," *Nature* **518** no. 7540, (2015) 529–533.

[3] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, *et al.*, "Mastering the game of go with deep neural networks and tree search," *Nature* **529** no. 7587, (2016) 484–489.

[4] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, *et al.*, "Mastering the game of go without human knowledge," *Nature* **550** no. 7676, (2017) 354.

[5] N. Brown and T. Sandholm, "Superhuman ai for heads-up no-limit poker: Libratus beats top professionals," *Science* (2017) .

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521** no. 7553, (2015) 436–444.

[7] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," `arXiv:1301.3781 [cs.CL]`.

[8] R. Shwartz-Ziv and N. Tishby, "Opening the Black Box of Deep Neural Networks via Information," `arXiv:1703.00810 [cs.LG]`.

[9] S. S. Schoenholz, J. Gilmer, S. Ganguli, and J. Sohl-Dickstein, "Deep Information Propagation," *ArXiv e-prints* (Nov., 2016) , `arXiv:1611.01232 [stat.ML]`.

[10] J. Lee, Y. Bahri, R. Novak, S. S. Schoenholz, J. Pennington, and J. Sohl-Dickstein, "Deep Neural Networks as Gaussian Processes," *ArXiv e-prints* (Oct., 2017) , `arXiv:1711.00165 [stat.ML]`.

[11] D. Levy, M. D. Hoffman, and J. Sohl-Dickstein, "Generalizing Hamiltonian Monte Carlo with Neural Networks," *ArXiv e-prints* (Nov., 2017) , `arXiv:1711.09268 [stat.ML]`.

[12] J. Pennington, S. S. Schoenholz, and S. Ganguli, "Resurrecting the sigmoid in deep learning through dynamical isometry: theory and practice," *ArXiv e-prints* (Nov., 2017) , `arXiv:1711.04735 [cs.LG]`.

[13] G. Yang and S. S. Schoenholz, "Mean Field Residual Networks: On the Edge of Chaos," *ArXiv e-prints* (Dec., 2017) , `arXiv:1712.08969`.

[14] E. D. Cubuk, B. Zoph, S. S. Schoenholz, and Q. V. Le, "Intriguing Properties of Adversarial Examples," *ArXiv e-prints* (Nov., 2017) , `arXiv:1711.02846 [stat.ML]`.

[15] S. S. Schoenholz, J. Pennington, and J. Sohl-Dickstein, "A Correspondence Between Random Neural Networks and Statistical Field Theory," *ArXiv e-prints* (Oct., 2017) , `arXiv:1710.06570 [stat.ML]`.

[16] S. L. Smith, P.-J. Kindermans, and Q. V. Le, "Don't Decay the Learning Rate, Increase the Batch Size," *ArXiv e-prints* (Nov., 2017) , `arXiv:1711.00489 [cs.LG]`.

[17] S. L. Smith and Q. V. Le, "A Bayesian Perspective on Generalization and Stochastic Gradient Descent," *ArXiv e-prints* (Oct., 2017) , `arXiv:1710.06451 [cs.LG]`.

[18] S. L. Smith, D. H. P. Turban, S. Hamblin, and N. Y. Hammerla, "Offline bilingual word vectors, orthogonal transformations and the inverted softmax," *ArXiv e-prints* (Feb., 2017) , `arXiv:1702.03859 [cs.CL]`.

[19] S. L. Smith, "Monte Carlo Sort for unreliable human comparisons," *ArXiv e-prints* (Dec., 2016) , `arXiv:1612.08555 [cs.AI]`.

[20] M. Arjovsky, A. Shah, and Y. Bengio, "Unitary Evolution Recurrent Neural Networks," *ArXiv e-prints* (Nov., 2015) , `arXiv:1511.06464 [cs.LG]`.

[21] P. Mehta, M. Bukov, C.-H. Wang, A. G. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, "A high-bias, low-variance introduction to machine learning for physicists," *arXiv preprint arXiv:1803.08823* (2018) .

[22] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences* **79** no. 8, (1982) 2554–2558.

[23] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation* **9** no. 8, (1997) 1735–1780.

[24] Y. LeCun, B. E. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. E. Hubbard, and L. D. Jackel, "Handwritten digit recognition with a back-propagation network," in *Advances in neural information processing systems*, pp. 396–404. 1990.

[25] G. Evenbly and G. Vidal, "Tensor network states and geometry," *Journal of Statistical Physics* **145** no. 4, (2011) 891–918.

[26] G. Vidal, "Entanglement Renormalization," *Phys. Rev. Lett.* **99** no. 22, (2007) 220405, `arXiv:cond-mat/0512165 [cond-mat]`.

[27] B. Swingle, "Constructing holographic spacetimes using entanglement renormalization," `arXiv:1209.3304 [hep-th]`.

[28] T. Hartman and J. Maldacena, "Time Evolution of Entanglement Entropy from Black Hole Interiors," *JHEP* **05** (2013) 014, `arXiv:1303.1080 [hep-th]`.

[29] A. R. Brown, D. A. Roberts, L. Susskind, B. Swingle, and Y. Zhao, "Holographic Complexity Equals Bulk Action?," *Phys. Rev. Lett.* **116** no. 19, (2016) 191301, `arXiv:1509.07876 [hep-th]`.

[30] E. M. Stoudenmire and D. J. Schwab, "Supervised Learning with Quantum-Inspired Tensor Networks," `arXiv:1605.05775 [stat.ML]`.

[31] V. Pestun and Y. Vlassopoulos, "Tensor network language model," `arXiv:1710.10248 [cs.CL]`.

[32] Y.-Z. You, Z. Yang, and X.-L. Qi, "Machine Learning Spatial Geometry from Entanglement Features," `arXiv:1709.01223 [cond-mat.dis-nn]`.

[33] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning requires rethinking generalization," `arXiv:1611.03530 [cs.LG]`.

[34] S. Yaida, "Fluctuation-dissipation relations for stochastic gradient descent," *arXiv preprint arXiv:1810.00004* (2018) .

[35] Choromanska, Anna and Henaff, Mikael and Mathieu, Michaël and Ben Arous, Gérard and LeCun, Yann, "The Loss Surfaces of Multilayer Networks," `arXiv:1412.0233 [cs.LG]`.

[36] C. Baldassi, A. Ingrosso, C. Lucibello, L. Saglietti, and R. Zecchina, "Subdominant dense clusters allow for simple learning and high computational performance in neural networks with discrete synapses," *Physical review letters* **115** no. 12, (2015) 128101.

[37] C. Baldassi, C. Borgs, J. T. Chayes, A. Ingrosso, C. Lucibello, L. Saglietti, and R. Zecchina, "Unreasonable effectiveness of learning neural networks: From accessible states and robust ensembles to basic algorithmic schemes," *Proceedings of the National Academy of Sciences* **113** no. 48, (2016) E7655–E7662.

[38] P. Chaudhari, A. Choromanska, S. Soatto, Y. LeCun, C. Baldassi, C. Borgs, J. Chayes, L. Sagun, and R. Zecchina, "Entropy-sgd: Biasing gradient descent into wide valleys," *arXiv preprint arXiv:1611.01838* (2016) .

[39] M. Baity-Jesi, L. Sagun, M. Geiger, S. Spigler, G. B. Arous, C. Cammarota, Y. LeCun, M. Wyart, and G. Biroli, "Comparing dynamics: Deep neural networks versus glassy systems," *arXiv preprint arXiv:1803.06969* (2018) .

[40] M. Geiger, S. Spigler, S. d'Ascoli, L. Sagun, M. Baity-Jesi, G. Biroli, and M. Wyart, "The jamming transition as a paradigm to understand the loss landscape of deep neural networks," *arXiv preprint arXiv:1809.09349* (2018) .

[41] P. Mehta and D. J. Schwab, "An exact mapping between the variational renormalization group and deep learning," *arXiv preprint arXiv:1410.3831* (2014) .

[42] D. A. Roberts, D. Stanford, and L. Susskind, "Localized shocks," *JHEP* **03** (2015) 051, `arXiv:1409.8180 [hep-th]`.

[43] D. A. Roberts and D. Stanford, "Two-dimensional conformal field theory and the butterfly effect," *Phys. Rev. Lett.* **115** no. 13, (2015) 131603, `arXiv:1412.5123 [hep-th]`.

[44] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, "Chaos in quantum channels," *JHEP* **02** (2016) 004, `arXiv:1511.04021 [hep-th]`.

[45] D. A. Roberts and B. Swingle, "Lieb-Robinson Bound and the Butterfly Effect in Quantum Field Theories," *Phys. Rev. Lett.* **117** no. 9, (2016) 091602, `arXiv:1603.09298 [hep-th]`.

[46] A. R. Brown, D. A. Roberts, L. Susskind, B. Swingle, and Y. Zhao, "Complexity, action, and black holes," *Phys. Rev.* **D93** no. 8, (2016) 086006, `arXiv:1512.04993 [hep-th]`.

[47] D. A. Roberts and B. Yoshida, "Chaos and complexity by design," *JHEP* **04** (2017) 121, `arXiv:1610.04903 [quant-ph]`.