Sidak Pal Singh

Nationality: Indian, Date of birth: 21st June, 1995.

Education

2020 - now **PhD Machine Learning**, *Eidgenössische Technische Hochschule (**ETH**) Zürich, Switzerland.*
Advisors: Thomas Hofmann (ETH Zürich) & Bernhard Schölkopf (ETH Zürich, MPI Tübingen).
Received the *Center for Learning Systems (**CLS**)* fellowship for joint PhD at ETH and MPI Tübingen.

2017 - 2020 **MSc Data Science**, *École Polytechnique Fédérale de Lausanne (**EPFL**), Switzerland.*
Awarded **Research Scholarship** to cover tuition fees and living expenses. GPA: *5.87/6*

2013 - 2017 **B.Tech Computer Science**, *Indian Institute of Technology (**IIT**) Roorkee, India.*
Department Medal for best project & ranked amongst **top 5%** with CGPA: *9.27/10*

Publications

ArXiv Lorenzo Noci*, Sotiris Anagnostidis*, Luca Biggio*, Antonio Orvieto*, **Sidak Pal Singh*** & Aurelien Lucchi. *Signal Propagation in Transformers: Theoretical Perspectives and the Role of Rank Collapse*, CoRR abs/2206.03126, 2022.

ICLR **Sidak Pal Singh**, Aurelien Lucchi, Thomas Hofmann & Bernhard Schölkopf. *Phenomenology of Double Descent in Finite-Width Neural Networks*, 10th International Conference on Learning Representations (ICLR), 2022.

NeurIPS **Sidak Pal Singh**, Gregor Bachmann & Thomas Hofmann. *Analytic Insights into Structure and Rank of Neural Network Hessian Maps,* NeurIPS, 2021 - Advances in Neural Information Processing Systems.

NeurIPS **Sidak Pal Singh**, & Dan Alistarh. *Efficient Second-Order Approximation for Neural Network Compression,* NeurIPS, 2020 - Advances in Neural Information Processing Systems.

NeurIPS **Sidak Pal Singh** & Martin Jaggi. *Model Fusion via Optimal Transport,* NeurIPS 2020 - Advances in Neural Information Processing Systems.

AISTATS, ICLR **Sidak Pal Singh**, Andreas Hug, Aymeric Dieuleveut & Martin Jaggi. *Context Mover's Distance & Barycenters: Optimal transport of contexts for building representations,* AISTATS 2020 & ICLR, DeepGenStruct Workshop, 2019.

ArXiv **Sidak Pal Singh**, Angela Fan & Michael Auli. *GLOSS: Generative Latent Optimization of Sentence Representations,* CoRR abs/1907.06385, 2019.

ICWS Rohit Ranchal, **Sidak Pal Singh**, Pelin Angin, Ajay Mohindra, Hui Lei & Bharat Bhargava. *RaaS and Hierarchical Aggregation Revisited*. 24th International Conference on Web Services, 2017.

IJCAI **Sidak Pal Singh,** Sopan Khosla, Sajal Rustagi, Manisha Patel, Dhaval Patel. *SL-FII: Syntactic & Lexical Constraints for Disease Mention Recognition in News Headlines*. International Joint Conference on Artificial Intelligence, BAI workshop, 2016.

Award & Honors

2021 **ELLIS**, Nominated as a PhD student in the European Lab for Learning and Intelligent Systems.

2021 **Princeton**, Accepted as a participant to Princeton Deep Learning Theory Summer School.

2020 **Heidelberg Laureate Forum**, Selected to attend as one of the top 225 young researchers in Mathematics and Computer Science worldwide.

2020 **CMMRS**, Selected to participate in the Cornell, Maryland, Max Planck pre-doctoral school.

2020 **EPFL**, Awarded EDIC Fellowship for PhD program in Computer Science (declined).

2019 **NeurIPS**, Awarded travel grant and selected amongst top 50% reviewers.

2018, 19, 20 **Google Brain, FAIR, Amazon, Sony, RIKEN & Salesforce**, Selected for research internship.

2019 **Computational Aspects of Geometry**, *IMT, Toulouse*, selected for the travel grant.

2018 **Machine Learning for Programming**, *Oxford, UK*, student grant to attend ML4P workshop.

2017 **Microsoft Research (MSR)**, *India*, selected for the fulltime Research Fellow position.

2016 **Honda Y-E-S (Young Engineer & Scientists) Award, Y-E-S Plus Award**, *10,000 $*, received financial support to pursue research in Japan (14 students selected all over India).

2016 **Google Venkat Panchapakesan Scholarship**, *750 $*, invited to visit Google & YouTube Headquarters, USA (6 students selected all over India).

2013 **IIT JEE**, Secured an All India Rank of 298 in the Joint Entrance Exam (MAINS) out of 1,400,000 students.

Research Experience

Oct 2019 - Aug 2020 **Research Assistant**, *Dan Alistarh's group, IST Austria*, Vienna.
- Designed ways to practically utilize the curvature information given by the Hessian in deep learning, in particular, for model compression.

Mar 2019 - Aug 2019 **Research Scholar**, *Martin Jaggi's group, EPFL*, Lausanne, Switzerland.
- Developed a novel method to ensemble different trained models into a single model. This is done via layer-wise averaging of weights by first aligning the inter-model neurons with Optimal Transport.

Sep 2018 - Feb 2019 **PhD Research Intern**, *Michael Auli's group, Facebook AI Research (FAIR)*, Menlo Park, USA.
- Worked on building non-compositional embeddings for application in text representation and generation.

Sep 2017 - Aug 2018 **Research Scholar**, *Martin Jaggi's group, EPFL*, Lausanne, Switzerland.
- Proposed a method to represent each entity as a *probability distribution over their context embeddings*, which allows us to use Optimal Transport to compare entities and represent a composition of entities. This resulted in strong empirical results for tasks such as word entailment and sentence similarity tasks.

May 2016 - Jul 2016 **Machine Learning Research Intern**, *Marco Cuturi's group, Kyoto University*, Japan.
- Designed a loss function for training Generative Adversarial Networks (GANs) based on *entropy regularized Wasserstein* distances. Learned the ground metric via Large Margin Nearest Neighbors (LMNN) on the feature activations and implemented the entire system using the *Chainer* library.

May 2015 - Jul 2015 **Summer Intern**, *Bharat Bhargava's group, Purdue University*, West Lafayette, USA.
- Developed a method to estimate the relevance of reviews using their metadata, with a particular focus on reviews with limited votes. Integrated this with *Rating as a Service* (RaaS) architecture and provided a *RESTful API* for interaction (written in *Node.js* with *MongoDB* for persistence). [paper] [code]

Feb 2016 - Apr 2016 **Research Assistant**, *Dhaval Patel's group, IIT Roorkee*, Roorkee, India.
- Designed a system to recognize disease mentions in news headlines by utilizing lexical and syntactic constraints to identify signaling word roots. Given the unsupervised nature of this task, simulated annotations from a manually prepared list of 95 diseases. [paper] [code]

Dec 2015 - Jan 2016 **Winter Intern**, *Koyel Mukherjee's group, Xerox Research Centre*, Bangalore, India.
- Prototyped a multimodal trip planning system integrating dynamic ridesharing with scheduled transportation services. Used *k-medoids* to find clusters of landmarks in the road network graph. [code]

Other Projects

Apr 2019 - Jun 2019 **Sparse Approximate Inverse Preconditioner (SPAI)**.
- Studied Grote & Huckle's proposed SPAI method and reproduced the given results. Also, verified the performance of preconditioners when used together with a linear system solver like BICGSTAB. [report]

May 2019 - Jun 2019 **Robust regression with Huber loss**.
- Reformulated the objective as a quadratic program and used kernel trick in the dual. [code] [report]

May 2018 **Toy-PyTorch: Neural Networks library from scratch**.
- Implemented a mini deep learning framework from scratch by relying only on `torch.Tensor`. Obtained a better understanding of the entire pipeline of training neural networks with backpropagation. [code]

Nov 2017 - Dec 2017 **Aerial Image Segmentation using U-Net**.
- Developed a fully convolutional approach with U-Net that *won the in-class Kaggle* competition. [report]

Nov 2017 - Dec 2017 **Markov Chain Monte Carlo for Ising Perceptron**.
- Designed a Metropolis-Hastings based Markov Chain to model the perceptron classification problem. Used simulated annealing for optimal choice of parameters & reaching closer to global minima. [report]

Sep 2017 **Data Study Group: Health & Well-being**, *The Alan Turing Institute*, London, UK.
- Inferred patient risk from mammography using transfer learning with a pre-trained Inception_v3. Visualized attention maps to understand the working of internal CNN layers. [slides]

Relevant Courses

Undergraduate Machine Learning, Artificial Intelligence, Advanced Graph Theory, Theory of Computation, Data Structures, Algorithms, Principles of Programming Languages.

Masters Advanced probability & applications, Computational Linear Algebra, Learning Theory, Multivariate Statistics, Deep learning, Optimization for Machine Learning, Advanced Algorithms, Statistics for data science, Markov chains & algorithmic applications, Information theory and signal processing.

PhD Beyond i.i.d. learning: Causality, dynamics, and interactions, Critiques of Scientific Objectivity.

Invited Talks

TOPML "Effective Number of Parameters in Neural Networks via Hessian Rank" at the workshop on the Theory of Overparameterized Machine Learning (TOPML) (April 2022).

EPFL "Analytic Insights into Structure and Rank of Neural Network Hessian Maps" at the MLO lab (July 2021).

DLCT "Model Fusion via Optimal Transport" at the Deep Learning: Classics and Trends reading group (September 2020).

Google "WoodFisher: Efficient Second-Order Approximation for Neural Network Compression" at the Google Sparsity reading group (September 2020).

Service

Academic Reviewer for ICML 2019, 2020, 2022; NeurIPS 2019, 2020, 2021, 2022; ICLR 2020, 2021, 2022. Student volunteer for NeurIPS 2017, 2019.

TA Teaching Assistant for Computational Intelligence Lab course at ETH Zürich, Spring 2021.

Organizer Head co-organizer for the ELLIS reading group on Mathematics of Deep Learning (2022).

Founder Started and organized the Machine Learning Reading Group (MLRG) in IIT Roorkee (2016).

Mentorship Received Star Mentor Award for guiding 10 first-year undergraduates (2015).

Technical skills

- **Languages**: C/C++, Python, Java.
- **Software Packages**: PyTorch, Jax, Chainer, sklearn, Numpy, MATLAB, Docker, Git, Android, LaTeX.
- **Operating Systems**: GNU/Linux, macOS, Windows.