

EDUCATION

- Swiss Federal Institute of Technology in Zurich (ETHz), Zurich, Switzerland** 03/2021 - (ongoing)
- Ph.D. in Deep Learning advised by Thomas Hofmann
 - Inductive bias of machine learning models and the implication of model size and dataset size to generalization
 - Transformers, Representation learning, and Efficient Inference
- Swiss Federal Institute of Technology in Zurich (ETHz), Zurich, Switzerland** 09/2018 - 10/2020
- Data Science Master (2-year degree; 120 ECTS): Grade: 5.86/6
 - Thesis title: “[Direct-Search for a Class of Stochastic Min-Max Problems](#)”
- National Technical University of Athens (NTUA), Athens, Greece** 09/2013 - 08/2018
- BSc & MSc in Electrical and Computer Engineering (5-year joint degree; 300 ECTS): Grade: 9.90/10 (*top 1%*)
 - Computer Systems, Computer Software and Networks, Signals and Automatic Control
 - Thesis title: “[Device profiling for anomaly detection](#)”

PROFESSIONAL EXPERIENCE

- Research Scientist Intern, Meta Zurich** 06/2024 - 11/2024
- Efficient Inference in vision-language models
- Student Researcher, Google DeepMind Zurich** 10/2023 - 12/2023
- Multimodal model evaluation/fine-tuning
- Research Internship, von Neumann Lab, Huawei Research Center Zurich** 11/2020 - 02/2021
- Implementation and analysis of non-blocking operations in GraphBLAS.
- Lab & Assistant, Data Analytics Lab, ETH Zurich** 09/2019 - (ongoing)
- Head teaching assistant for the courses Deep Learning and Computational Intelligence Lab. I also serve as the IT coordinator and the web manager.
- Research Assistant, Systems Lab, ETH Zurich** 11/2018 - 07/2019
- Automatic translation of front-end primitives to a hardware-agnostic intermediate language, for efficient implementation of big data applications.
- Teaching & Research Assistant, Computation and Reasoning Laboratory, NTUA, Greece** 09/2016 – 08/2018
- Lab Assistant, CSLAB: Computing Systems Laboratory, NTUA, Greece** 09/2017 – 08/2018

SELECTED PUBLICATIONS

- [Anagnostidis S, Bachmann G., Kim Y., Kohler J., Georgopoulos M., Sanakoyeu A., Du Y., Pumarola A., et al. FlexiDiT: Your Diffusion Transformer Can Easily Generate High-Quality Samples with Less Compute](#) [to appear]
- [Anagnostidis S., Bachmann G, Schlag I., Hofmann T. \(2024\) Navigating Scaling Laws: Compute Optimality in Adaptive Model Training](#), ICML (Spotlight)
- [Anagnostidis, S., Pavllo, D., Biggio, L., Noci, L., Lucchi, A., & Hoffmann, T. \(2023\) Dynamic Context Pruning for Efficient and Interpretable Autoregressive Transformers](#). NeurIPS (Spotlight)
- Bachmann, G., [Anagnostidis, S.](#), & Hofmann, T. (2023) [Scaling MLPs: A Tale of Inductive Bias](#). NeurIPS
- Köpf, A., Kilcher, Y., von Rütte, D., [Anagnostidis, S.](#), Tam, Z. R., Stevens, K., et al. (2023) [OpenAssistant Conversations--Democratizing Large Language Model Alignment](#). NeurIPS (Oral)

- Noci, L., [Anagnostidis, S.](#), Biggio, L., Orvieto, A., Singh, S. P., & Lucchi, A. (2022) **Signal propagation in transformers: Theoretical perspectives and the role of rank collapse**. NeurIPS
- Delitzas, A., Parelli, M., Hars, N., Vlassis, G., [Anagnostidis, S.](#), Bachmann, G., & Hofmann, T. (2023) **Multi-CLIP: Contrastive Vision-Language Pre-training for Question Answering in 3D Scenes**. BMVC (Oral)
- von Rütte D., [Anagnostidis S.](#), Bachmann G., Hofmann T. (2024) **A Language Model's Guide Through Latent Space**. ICML
- Imfeld M., Graldi J., Giordano M., Hofmann T., [Anagnostidis S.](#), Pal Singh S. (2024) **Transformer Fusion with Optimal Transport**. ICLR

SUPERVISING STUDENTS (REVERSE CHRONOLOGICAL ORDER)

Han Yang	Msc Thesis: MegaPortrait: Revisiting Diffusion Control for High-fidelity Portrait Generation [to appear] Msc Project: NovelPortrait: Towards View-consistent Relightable Portrait Generation [arxiv]
Alexander Theus & Olin Geimer	Msc Project: Intra-Fusion: Towards Meta-Pruning via Contextualising Neuron Importance [ICLR Spotlight]
Moritz Imfeld & Jacopo Graldi & Marco Giordano	Msc Project: Transformer Fusion with Optimal Transport [ICLR]
Ben Arous Elior	Msc Thesis: Image-Editing Specialists: A Multi-Reward Approach for Diffusion Models [to appear] Msc Project: Harnessing Synthetic Datasets: The Role of Shape Bias in Deep Neural Network Generalization [NeurIPS SyntheticData4ML]
Dimitri von Rutte	Msc Project: A Language Model's Guide Through Latent Space [ICML]
Sarnthein-Lotichius Felix	Msc Thesis: On Representation Learning in Self-Distillation With No Labels [ICML , Thesis]
Sabrina Herbst	Summer Student Intern: Pruning Large Language Models
Arvid Berg & Simon Schläpfer	Msc Project: ViT-nessing Mode Connectivity after Permutations
Kevin Blin	Msc Thesis: Object-Centric and Spatial bias in World Models [Thesis]
Soyuer Bartu	Msc Project: Contrastive Representation Learning For Remote Sensing Datasets and Downstream Applications
Thierry Meier	Bachelor Thesis: Spectral Similarity of 1D 1H NMR Spectra

HONORS & ACADEMIC AWARDS

- **Onassis Foundation, Scholarship** 2019-2021
- **Honorary award** given to the three graduates with the highest grades for my NTUA bachelor & master
- **Nikolaos Kritikos Scholarship** for Excellence in Mathematics
- **Thomaidio Award** 2014, 2015, 2016, 2017, 2018
- **National Physics Competition 2012**, 7th place

SKILLS

- Programming Python, C/C++, Java, SML/Haskell, MPI, Assembly (ARM, 8086), OpenMPI
- Tools Linux, Matlab, Tensorflow, Pytorch

LANGUAGES

- **English** (fluent)
- **German** (advanced)
- **Russian** (basic)
- **Greek** (native)

FULL PUBLICATION LIST (REVERSE CHRONOLOGICAL ORDER)

- [Anagnostidis S.](#), Bachmann G., Kim Y., Kohler J., Georgopoulos M., Sanakoyeu A., Du Y., Pumarola A., Thabet A., Schönfeld E. (2024)
FlexiDiT: Your Diffusion Transformer Can Easily Generate High-Quality Samples with Less Compute [to appear]
- Kim Y., [Anagnostidis S.](#), Du Y., Schönfeld E., Kohler J., Georgopoulos M., Pumarola A., Thabet A., Sanakoyeu A. (2024)
Autoregressive Distillation of Diffusion Transformers [to appear]
- Bachmann G., [Anagnostidis S.](#), Pumarola A., Georgopoulos M., Sanakoyeu A., Du Y., Schönfeld E., Thabet A., Kohler J. (2024)
Judge Decoding: Faster Speculative Sampling Requires Going Beyond Model Alignment. [to appear]
- Yang H., [Anagnostidis S.](#), Simsar E., Hofmann T. (2024)
MegaPortrait: Revisiting Diffusion Control for High-fidelity Portrait Generation. [to appear]
- [Anagnostidis S.](#), Bulian J. (2024)
How Susceptible are LLMs to Influence in Prompts? COLM
- von Rütte D., [Anagnostidis S.](#), Bachmann G., Hofmann T. (2024)
A Language Model's Guide Through Latent Space. ICML, ICLR Workshop
- [Anagnostidis S.](#), Bachmann G, Schlag I., Hofmann T. (2024)
Navigating Scaling Laws: Compute Optimality in Adaptive Model Training, ICML (Spotlight)
- Imfeld M., Graldi J., Giordano M., Hofmann T., [Anagnostidis S.](#), Pal Singh S. (2024)
Transformer Fusion with Optimal Transport. ICLR
- Theus A., Geimer O., Wicke F., Hofmann T., [Anagnostidis S.](#), Singh SP. (2024)
Towards Meta-Pruning via Optimal Transport. ICLR (Spotlight)
- [Anagnostidis S.](#), Pavllo, D., Biggio, L., Noci, L., Lucchi, A., & Hoffmann, T. (2023)
Dynamic Context Pruning for Efficient and Interpretable Autoregressive Transformers. NeurIPS (Spotlight)
- Bachmann, G., [Anagnostidis S.](#), & Hofmann, T. (2023)
Scaling MLPs: A Tale of Inductive Bias. NeurIPS
- Köpf A., Kilcher Y., von Rütte D., [Anagnostidis S.](#), Tam ZR., Stevens K., Barhoum A., Duc NM., Stanley O., Nagyfi R., Shahul ES, Suri S., Glushkov D., Dantuluri A., Maguire A., Schuhmann C., Nguyen H., Mattick A. (2023)
OpenAssistant Conversations--Democratizing Large Language Model Alignment. NeurIPS (Oral)
- Benarous E., [Anagnostidis S.](#), Biggio L., Hofmann T. (2023)
Harnessing Synthetic Datasets: The Role of Shape Bias in Deep Neural Network Generalization. NeurIPS SyntheticData4ML
- Delitzas A., Parelli M., Hars N., Vlassis G., [Anagnostidis S.](#), Bachmann G., Hofmann T. (2023)
Multi-CLIP: Contrastive Vision-Language Pre-training for Question Answering tasks in 3D Scenes. BMVC

(Oral), CVPR workshop

- Sarnthein F., Bachmann G., [Anagnostidis S.](#), Hofmann T. (2023)
Random Teachers are Good Teachers. ICML
- [Anagnostidis S.](#), Lucchi A., Hofmann T. (2023)
Mastering Spatial Graph Prediction of Road Networks. ICCV
- Noci, L., [Anagnostidis S.](#), Biggio, L., Orvieto, A., Singh, S. P., & Lucchi, A. (2022)
Signal propagation in transformers: Theoretical perspectives and the role of rank collapse. NeurIPS
- [Anagnostidis S.](#), Thomsen A., Kacprzak T., Tröster T., Biggio L., Refregier A., Hofmann T. (2022)
Cosmology from Galaxy Redshift Surveys with PointNet. NeurIPS Workshop
- [Anagnostidis S.](#), Bachmann G., Noci L., Hofmann T. (2022)
The curious case of benign memorization. ICLR
- Mastoras A., [Anagnostidis S.](#), Yzelman AJN. (2022)
Nonblocking execution in GraphBLAS. ACM TACO, IPDPSW
- [Anagnostidis S.](#), Lucchi A., Diouane Y. (2021)
Direct-search for a class of stochastic min-max problems. AISTATS

ACADEMIC SERVICE

Reviewer CVPR'25, NeurIPS'24 (outstanding reviewer), ICLR'24, ICML'24, NeurIPS'23 (outstanding reviewer), ICLR'23

INTERESTS

- Love various sports including basketball, tennis, cycling, football and [superkondi](#). Going to run an IRONMAN Triathlon in the future.
- Love classical music, I have been playing the violin since I was 5 years old.
- A true believer of open source; leading machine learning team for open-assistant, on open source replication/extension of ChatGPT.