Tomasz Kacprzak

tomasz.kacprzak@phys.ethz.ch
tomaszkacprzak.github.io

Employment

2021 - present	**Senior Data Scientist**, Swiss Data Science Center, Paul Scherrer Institute
2018 - 2021	**Senior Scientist**, Institute for Particle Physics and Astrophysics, ETH Zurich
2014 - 2018	**Post-Doctoral Research Associate**, Institute for Astronomy, ETH Zurich

Education

2010 - 2014	**PhD in Physics and Astronomy**, University College London, UK Advisors: Prof Sarah Bridle, Prof John Shawe-Taylor
2009 - 2010	**MSc Machine Learning**, First Class, University College London, UK
2004 - 2009	**MSc Computer Science**, First Class, Wroclaw University of Technology, PL

Awards

2024 - 2025	NERSC Generative AI for Science Project: *DES+DESC Generative AI for Cosmology* **Principal Investigator**, 7k A100 GPU Perlmutter Node Hours, ≈40k USD Awarded by the National Energy Research Scientific Computing Center, USA
2023 - 2025	NERSC Exascale Science Applications Program, Project: *DESLearning* **Principal Investigator**, 0.5 NERSC FTE algorithmic scaling and technical support
2020 - 2021	Computational Production Project: *Measuring Dark Energy with Deep Learning* **Principal Investigator**, 750k GPU-node hours, ≈500k CHF Awarded by the Swiss National Supercomputing Center
2017 - 2019	Collaborative Grant: *Deep Learning for Observational Cosmology* **Principal Investigator**, hired staff: N. Perraudin (postdoc), J. Fluri (PhD) Awarded by Swiss Data Science Center, ≈750k CHF
2019	Workshop: *Artificial Intelligence Methods in Cosmology*, June 9-12 Main applicant and lead organizer, ≈5k CHF Awarded by ETHZ Congressi Stefano Franscini

Major Achievements in AI and HPC for Physics and Cosmology

- Delivered first deep learning cosmology measurements using large-scale simulations on state-of-the-art GPUs (ETHZ Leonard, CSCS Piz Daint) and parallel scaling on NERSC Perlmutter
- First SDSC expert on the DGX A100 GPU nodes (PSI Gewendolen), utilized for AI workflows (DeepLSS) and massive non-convex inverse problem custom solvers (LaueOT)
- Introduced first generative AI models for cosmic web simulations, including large-scale 3D volumetric training data, using GANs and DDPMs (Piz Daint, Perlmutter)
- Produced CosmoGridV1, the first massive simulations set for simulations-based inference and AI training, 750TB of data public at www.cosmogrid.ai, ran on Piz Daint, large production project
- Created first approximate Bayesian computation methodology and pipelines for galaxy population constraints from image simulations, multiple runs on ETHZ Euler cluster $\mathcal{O}(1m)$ CPU hours
- Introduced first calibration methodology for gravitational shear using large image simulation on HPC clusters, currently adapted as standard in all weak gravitational lensing surveys

Leadership and Service

2023- **Member of the SKA Switzerland Consortium**
- NCCR *D+Cosmos* grant submission: coordinated pillar for Data Science and AI *Digital Frontier* (14 PIs, budget ≈15m CHF)
- Developing multiprobe SBI projects for combining SKA and LSST/Euclid

2022- **Simulations Working Group Coordinator in the Dark Energy Survey**
- Leading a group of 30 DES members, 10 cosmology projects with SBI
- Providing simulated theory prediction for multiprobe SBI projects
- Coordinating submissions for competitive computing resources grants
- Organizing bi-weekly calls for project updates, planning and invited talks

2012- **Member of the Science Committee and Builder of the Dark Energy Survey**
- Coordinating new innovation-oriented projects, designing publication policy
- Permanent *Builder* status awarded after 2 years FTE of DES infrastructure work
- In-person participation in 15 collaboration meetings (USA, Europe, South America)
- Co-wrote IM3SHAPE, a weak lensing measurement code for DES SV and Y1
- Performed simulations-based shear calibration in DES SV (NERSC Cori)
- Chair of the Early Career Scientist Committee, organizing career events (2016-2017)

2025- **Euclid Collaboration: Co-Lead of Work Package *Field Level Inference***
- Organizing projects for field-level inference in multiprobe cosmology
- Design of SBI simulations for Cosmological Simulations Working Group
- Internal reviewer for papers in the area of SBI and higher-order statistics

2022- **Student Supervision and Mentoring**
- Idea originator and lead advisor for 4 post-doc, 8 PhD, and 13 MSc projects
- 9/13 MSc projects resulted in articles published in peer-reviewed journals

2018- **Service in Astrophysics and Computer Science**
- Science Organizing Committee Member *UniverseAI*, Athens 2-6 June 2025
- Reviewer for the *Application Track* at *Supercomputing* conference 2024 and 2025
- Paper reviewer for Nature Astronomy, PNAS, Physical Review, JCAP, etc.
- External examiner for PhD thesis Justine Zeghal (APC, Paris, FR)
- Proposal reviewer for Swiss Data Science Center calls for collaborative projects
- Proposal reviewer for UKRI Future leaders programme

2019 **Lead Organizer of Workshop *Artificial Intelligence Methods in Cosmology***
- Monte Verita, Ascona, June 2019, 46 participants, 6 invited speakers
- Lead grant writer, speaker invitations, program development, logistics

Recent presentations (selected)

2024/10 *X Meeting on Fundamental Cosmology*, Sevilla, ES, **invited review talk**

2023/11 *Debating the Potential of ML in Astro. Surveys*, IAP Paris, FR, **invited review talk**

2023/11 *Mathematics and Informatics Colloquium*, Uni Basel, CH, invited seminar

2023/05 *ML X Astrophysics Symposium*, Flatiron Institute, New York, USA, invited talk

2023/03 *University Observatory Munich Colloquium*, DE, invited talk + 2-day tutorials

2022/07 *Key Challenges in Galaxy and CMB lensing*, Cambridge, UK, invited guest talk

2022/06 *Bayesian Deep Learning in Cosmology*, Paris, FR, **invited keynote talk**

2022/06 *Space Science Data Center Seminar*, Rome, IT, invited seminar talk

2022/04 *Berkeley ML and Science Forum*, Berkeley, USA, invited seminar talk

2021/10 *Cosmology seminar*, University of Geneva, CH, invited seminar talk

2019/09 *Cosmology seminar*, SLAC, Stanford, USA, invited talk

Publications

Authored 24 papers on own original ideas, including 8 as the lead author. Significantly contributed to further 13 papers by students, and 33 international collaboration papers. My H-Index is 41. Highlighted publications:

1. **DeepSphere: Efficient spherical convolutional neural network with HEALPix sampling**
   N. Perraudin, M. Defferrard, T. Kacprzak, R. Sgier
   Astronomy and Computing, 2019, 27, 130
   Highest-cited paper, 220 citations, benchmark dataset for spherical CNNs

2. **Cosmological constraints with deep learning from KiDS-450 weak lensing maps**
   J. Fluri, T. Kacprzak, A. Lucchi, A. Refregier, A. Amara, T. Hofmann, A. Schneider
   Physical Review D, 2019, 100, 6
   First constraints on survey data with deep learning trained on map simulations, 151 citations, popularized by the media in the MIT Technology Review Magazine "Facial recognition algorithms can be deployed to hunt for dark matter", 19/09/2019, and ETH News "Artificial intelligence probes dark matter in the universe", 18/09/2019.

3. **DeepLSS: breaking parameter degeneracies in large scale structure with deep learning**
   T. Kacprzak, J. Fluri
   Physical Review X, 2022, 2, 031029,
   High impact journal (factor 14.5), popular article in APS Physics Magazine, "Machine Learning Pins Down Cosmological Parameters", 19/08/2022.

4. **Laue Indexing with Optimal Transport**
   T. Kacprzak, S. Samothrakitis, C. B. Larsen, J. Kopeček, M. Strobl, E. Polatidis, G. Obozinski
   Submitted to IEEE Pattern Analysis and Machine Intelligence, 2404.06478
   Custom non-convex solver for demixing, using full DGX A100 node, large-scale optimal transport

5. **Fast Cosmic Web Simulations with Generative Adversarial Networks**
   A. C. Rodriguez, T. Kacprzak, A. Lucchi, A. Amara, R. Sgier, +3 authors
   Computational Astrophysics and Cosmology, 2018, 5, 1, 4, 11
   First generative AI model for cosmic web simulations, seminal paper, 157 citations

6. **CosmoGridV1: a simulated wCDM theory prediction for map-level cosmology**
   T. Kacprzak, J. Fluri, A. Schneider, A Refregier, J Stadel
   Journal of Cosmology and Astroparticle Physics, 2023, 02, 050, 29
   Largest theory prediction for cosmological SBI, 750TB of data, CSCS large production project

7. **Fast Point Spread Function Modeling with Deep Learning**
   J. Herbel, T. Kacprzak, A. Amara, A. Refregier, A. Lucchi
   Journal of Cosmology and Astroparticle Physics, 2018, 07, 54
   One of the first deep learning models for telescope images, CSCS small project, 68 citations

8. **Cosmological N-body simulations: a challenge for scalable generative models**
   N. Perraudin, A. Srivastava, A. Lucchi, T. Kacprzak, T. Hofmann, A. Refregier
   Computational Astrophysics and Cosmology, 2019, 6, 5
   Novel patch-based generative adversarial networks for super-resolution of 3D volumetric simulations

9. **Simulation-based inference of deep fields**
   B. Moser, T. Kacprzak, S. Fischbacher, A. Refregier, D. Grimm, L. Tortorelli
   Journal of Cosmology and Astroparticle Physics, 2024, 05, 049, 37
   Massive approximate Bayesian computation, multiple ETHZ Euler runs of $\mathcal{O}(1m)$ CPU-hours

10. **Scalable Approximate Algorithms for Optimal Transport Linear Models**
    T. Kacprzak, F. Kamper, G. Janka, M. Heiss, A. Dillner, S. Takahama
    In press, submitted to Journal of Machine Learning Research
    Latest paper, a new model bridging the gap between linear regression and optimal transport

Supervision

I was the **idea originator and lead advisor for 3 post-doc projects, 8 PhD projects, 13 MSc projects.** 9/13 MSc projects resulted in articles in peer-reviewed journals, marked with ↗.

2024 - present	Jozef Bucko, post-doc, *Generative AI for cosmology*
2022 - present	Arne Thomsen, PhD, *Cosmology with deep learning of combined probes in DES*
2025	Silvan Fischbacher, PhD, *Optimal transport for subhalo abundance matching* ↗
2024	Silvan Fischbacher, PhD, *GalSBI: galaxy population evolution modelling with SBI* ↗
2023	Beatrice Moser, PhD, *Evolution of galaxy samples with ABC modelling in DES* ↗
2023	Virginia Ajani, post-doc, *Peak statistics of combined probes in DES*
2022	Silvan Fischbacher, MSc, *Redshift requirements for shear with intrinsic alignment* ↗
2022	Gaspard Aymerich, MSc, *Interpretability of deep-learning methods in weak lensing* ↗
2022	Ting Tan, MSc, *Assessing theoretical uncertainties for cosmology from weak lensing* ↗
2022	Dominik Zürcher, PhD, *Dark energy survey year 3 results: cosmology with peaks* ↗
2022	Janis Fluri, PhD, *Full wCDM analysis of KiDS-1000 lensing using deep learning* ↗
2018	Nathanael Perraudin, post-doc data scientist, *Deep learning on the sphere* ↗
2020	Timothy Wing Hei Yiu, MSc, *A tomographic mass map emulator of KiDS-1000* ↗
2020	Benjamin Suter, MSc, *Cosmology with machine learning and human-designed statistics*
2019	Dominik Zürcher, PhD, *Forecast for non-Gaussian statistics in large-scale surveys* ↗
2019	Janis Fluri, PhD, *Constraints with deep learning from KiDS-450 weak lensing maps* ↗
2019	Conrad Schwanitz, MSc, *Interpretability measures for deep learning on lensing maps*
2019	Sajanth Subramaniam, MSc, *Systematics-invariant constraints with deep learning*
2018	Jörg Herbel, PhD, *Fast point spread function modeling with deep learning* ↗
2018	Sandro Marcon, MSc, *Emulation of cosmological mass maps with conditional GANs* ↗
2018	Ankit Srivastava, MSc, *N-body simulations: a challenge for scalable GANs* ↗
2018	Janis Fluri, PhD, *Constraints from noisy convergence maps through deep learning* ↗
2018	Alex Stauffer, MSc, *Approximate Bayesian computation in cosmology with ABCpy*
2018	Jonathan Rosenthal, MSc, *Generative temporal models for cosmology*
2017	Janis Fluri, MSc, *Lensing peak statistics in the era of large cosmological surveys* ↗
2017	Andres Rodrigues, MSc, *Fast cosmic web with generative adversarial networks* ↗
2017	Jorit Schmelzle, MSc, *Cosmological model discrimination with deep learning*

Teaching

2019 - 2020	Lecturer for UG course *Statistical Methods and Analysis in Experimental Physics* Topical block: Bayesian methods, machine learning, simulations-based inference Tasks: lectures, preparing assignments, leading the tutorials for approx. 50 students
2019	Lecturer for the UG course *Astrophysics 1*, topical block: *Introduction to Cosmology*
2019 - 2020	Guest lecturer for UG course *Introduction to Data Science* at University of Zurich Topics: deep learning, convolutional neural networks, generative models
2017 - 2018	Course coordinator for undergraduate module *Physics 1 and 2*, approx. 300 students Tasks: creation of exercises, preparation of exams and coordination of marking
2016 - 2017	Teaching assistant for MSc course *Advanced Statistical Methods in Cosmology* Tasks: curriculum development, creating assignments, leading tutorials
2017	Leader of the tutorial sessions, masters-level module *Cosmological Probes*