Dr. rer. nat. Jonas Geiping

Full Name	Jonas Alexander Geiping	Offine Phone	+1 301 405 2671
Address	8125 Paint Branch Drive 20742 College Park, MD United States	Email Permanent Email Website	jgeiping@umd.edu jonas.geiping@gmail.com jonasgeiping.github.io
Date of Birth	25th of March 1992, Berlin	Google Scholar	206vNCEAAAAJ

Biography

My background is in mathematical optimization and its applications to deep learning. My research focuses on fundamental questions of safety and efficiency in modern machine learning. On the safety side, I am interested in testing the security of machine learning models, making models safer to use and deploy, and modifying them to mitigate negative effects on users of ML models and bystanders. On the efficiency side I am interested in making models that can be trained with smaller compute footprints and improved reasoning capabilities, especially in language, which will hopefully help to democratize the use of machine learning.

Experience

Oct. 2023 -
ELLIS Institute Tübingen
Hector Endowed ELLIS Fellow

I will start on the 15th of October as principal investigator at the ELLIS institute.

Sept. 2021 - Oct. 2023
University of Maryland, College Park
Postdoctoral Researcher

Research into privacy and security in machine learning and into deep learning as a science, from an optimization perspective with applications in federated learning, stochastic training of neural networks and topics in security.

May 2021 July 2021
University of Siegen
Postdoctoral Associate

Postdoctoral Research: optimization and generalization in deep learning.

Oct. 2016 March 2021
University of Siegen
Research Associate

Research in the fields of mathematical optimization, machine learning and computer vision with Prof. Michael Möller with work done in non-convex composite optimization, convex relaxations, optimization for computer vision, learning of optimization objectives and bi-level optimization problems in security.

Aug. 2019 Nov. 2019
University of Maryland, College Park
Short-Term Scholar Exchange Visitor

Visiting the group of Prof. Tom Goldstein, joint research in topics of data poisoning for machine learning models and empirical evaluation of deep learning theory.

Oct 2014 - June 2016
University of Münster (WWU) - Cells-in-Motion Cluster of Excellence
Research Assistant

CiM flexible fund project *FF-2014-06* - " Analysis of cell-cell interactions during neuronal migration in the developing cortex by live cell imaging and cell shape quantification"

Education

Dec 2016 - April 2021
University of Siegen
Dr. rer. nat. (PhD), Computer Science

Advisor: *Prof. Michael Möller*
Thesis: *Modern Optimization Techniques in Computer Vision - From Variational Models to Machine Learning Security*

Oct 2014 - Sept. 2016
Westfälische-Wilhelms Universität Münster
M.Sc., Mathematics

Advisor: Prof. Martin Burger
Thesis: *Image Analysis of Neural Tissue Development: Variational Methods for Segmentation and 3D-Reconstruction from large pinhole confocal fluorescence microscopy*

Oct 2011 - Sept. 2014
Westfälische-Wilhelms Universität Münster
B.Sc., Mathematics

Advisor: Prof. Martin Burger
Thesis: *Topology-Preserving Segmentation Methods and Application to Mitotic Cell Tracking*

Selected Publications

Jonas Geiping and Tom Goldstein. Cramming: Training a Language Model on a single GPU in one day. In *Proceedings of the 40th International Conference on Machine Learning*, pages 11117–11143. PMLR, July 2023. URL https://proceedings.mlr.press/v202/geiping23a.html.

Jonas Geiping, Hartmut Bauermeister*, Hannah Dröge*, and Michael Moeller. Inverting Gradients - How easy is it to break privacy in federated learning? In *Advances in Neural Information Processing Systems*, volume 33, December 2020. URL https://proceedings.neurips.cc/paper/2020/hash/c4ede56bbd98819ae6112b20ac6bf145-Abstract.html.

Jonas Geiping*, Liam H. Fowl*, W. Ronny Huang, Wojciech Czaja, Gavin Taylor, Michael Moeller, and Tom Goldstein. Witches' Brew: Industrial Scale Data Poisoning via Gradient Matching. In *International Conference on Learning Representations*, April 2021. URL https://openreview.net/forum?id=01olnfLIbD.

John Kirchenbauer*, **Jonas Geiping***, Yuxin Wen, Jonathan Katz, Ian Miers, and Tom Goldstein. A Watermark for Large Language Models. In *Proceedings of the 40th International Conference on Machine Learning*, pages 17061–17084. PMLR, July 2023. URL https://proceedings.mlr.press/v202/kirchenbauer23a.html.

Recent Invited Talks

Sept 2023	UMass Amherst *NLP Seminar Talk - Efficient Training of Language Models and Watermarking of Generated Text*
July 2023	Intel Labs *Training Language Models on a Budget*
Mar 2023	Morgan Stanley Research, New York *Efficient Training of Language Models and Watermarking of Generated Text*
Feb 2023	ENSTA Paris | Institut Polytechnique de Paris *Cramming: Training a Language Model on a Single GPU in One Day*
Oct 2022	Qualcomm AI Research - DistributedML Seminar *Privacy and Security Analysis in Federated Learning*
May 2022	Federated Learning One World Seminar *New Threat Models for Privacy Attacks in Federated Learning*
Feb 2022	RIKEN AIP - 5th TrustML Young Scientist Seminar *Attacks on Privacy in Federated Learning Scenarios*
June 2021	Google ML privacy testing team *Inverting Gradients - A Privacy Question for Federated Learning?*

Recently in the News

Feb 2023	New York Times *How ChatGPT Could Embed a 'Watermark' in the Text It Generates*
Feb 2023	Computerphile *Ch(e)at GPT? - Computerphile*
Feb 2023	WIRED *How to Detect AI-Generated Text, According to Researchers*
Feb 2023	Nature - News Feature *What ChatGPT and generative AI mean for science*
Feb 2023	heise.de *Wie Wasserzeichen Texte von KI-Chatbots sichtbar machen könnten*
Jan 2023	Marktechpost *New AI Research from the University of Maryland Investigates Cramming Challenge*
Jan 2023	MIT Technology Review *A watermark for chatbots can expose text written by an AI*
Jan 2023	Marktechpost *Researchers at the University of Maryland Propose Cold Diffusion*
Dec 2022	TechCrunch *Image-generating AI can copy and paste from training data, raising IP concerns*

Teaching

I have prepared and discussed exercises and homework in deep learning, convex optimization and variational image processing while at the University of Siegen, and have held lectures a number of times as a substitute. These courses were first developed during this time and I have learned a lot from being part of their conceptualization. I have also taught as a teaching assistant in numerical linear algebra at the University of Münster.

Community Work

I have reviewed and continue to review for the large machine learning conferences and journals and computer vision conferences. I have been named "Top Reviewer" at the Thirty-sixth Conference on Neural Information Processing Systems (NeurIPS 2022) and "Highlighted Reviewer" at the Tenth International Conference on Learning Representations (ICLR 2022).
I am a proponent of openly available publications and code. All of my research is available as preprints on `arxiv.org` and our code implementations have accrued consistent interest from the community, measured in several hundred stars and forks on github. Implementations can be found for code written predominantly by me at `https://github.com/JonasGeiping/` and for my colleagues on their respective pages.

Full List of Publications

Arpit Bansal*, Hong-Min Chu*, Avi Schwarzschild, Soumyadip Sengupta, Micah Goldblum, **Jonas Geiping**, and Tom Goldstein. Universal Guidance for Diffusion Models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 843–852, 2023. URL `https://openaccess.thecvf.com/content/CVPR2023W/GCV/html/Bansal_Universal_Guidance_for_Diffusion_Models_CVPRW_2023_paper.html`.

Eitan Borgnia‡, Valeriia Cherepanova‡, Liam Fowl‡, Amin Ghiasi‡, **Jonas Geiping**‡, Micah Goldblum‡, Tom Goldstein‡, and Arjun Gupta‡. Strong Data Augmentation Sanitizes Poisoning and Backdoor Attacks Without an Accuracy Tradeoff. In *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3855–3859, June 2021a. doi: 10.1109/ICASSP39728.2021.9414862.

Eitan Borgnia*, **Jonas Geiping***, Valeriia Cherepanova*, Liam Fowl*, Arjun Gupta*, Amin Ghiasi, Furong Huang, Micah Goldblum, and Tom Goldstein. DP-InstaHide: Provably Defusing Poisoning and Backdoor Attacks with Differentially Private Data Augmentations. In *ICLR 2021 Workshop on Security and Safety in Machine Learning Systems*, March 2021b. URL `http://arxiv.org/abs/2103.02079`.

Ping-Yeh Chiang‡, **Jonas Geiping**‡, Micah Goldblum‡, Tom Goldstein‡, Renkun Ni‡, Steven Reich‡, and Ali Shafahi‡. Witchcraft: Efficient PGD Attacks with Random Step Size. In *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3747–3751, May 2020. doi: 10.1109/ICASSP40776.2020.9052930.

Ping-yeh Chiang, Renkun Ni, David Yu Miller, Arpit Bansal, **Jonas Geiping**, Micah Goldblum, and Tom Goldstein. Loss Landscapes are All You Need: Neural Network Generalization Can Be Explained Without the Implicit Bias of Gradient Descent. In *The Eleventh International Conference on Learning Representations*, February 2023. URL https://openreview.net/forum?id=QC10RmRbZy9.

Hong-Min Chu, **Jonas Geiping**, Liam H. Fowl, Micah Goldblum, and Tom Goldstein. Panning for Gold in Federated Learning: Targeted Text Extraction under Arbitrarily Large-Scale Aggregation. In *International Conference on Learning Representations*, February 2023. URL https://openreview.net/forum?id=A9WQaxYsfx.

Liam Fowl*, Ping-yeh Chiang*, Micah Goldblum*, **Jonas Geiping**, Arpit Bansal, Wojtek Czaja, and Tom Goldstein. Preventing Unauthorized Use of Proprietary Data: Poisoning for Secure Dataset Release. In *ICLR 2021 Workshop on Security and Safety in Machine Learning Systems*, February 2021a. URL http://arxiv.org/abs/2103.02683.

Liam Fowl*, **Jonas Geiping***, Wojciech Czaja, Micah Goldblum, and Tom Goldstein. Robbing the Fed: Directly Obtaining Private Data in Federated Learning with Modified Models. In *International Conference on Learning Representations*, September 2021b. URL https://openreview.net/forum?id=fwzUgo0FM9v.

Liam Fowl*, Micah Goldblum*, Ping-yeh Chiang*, **Jonas Geiping**, Wojciech Czaja, and Tom Goldstein. Adversarial Examples Make Strong Poisons. In *Advances in Neural Information Processing Systems*, volume 34, pages 30339–30351. Curran Associates, Inc., 2021c. URL https://proceedings.neurips.cc/paper/2021/hash/fe87435d12ef7642af67d9bc82a8b3cd-Abstract.html.

Liam H. Fowl*, **Jonas Geiping***, Steven Reich, Yuxin Wen, Wojciech Czaja, Micah Goldblum, and Tom Goldstein. Decepticons: Corrupted Transformers Breach Privacy in Federated Learning for Language Models. In *International Conference on Learning Representations*, February 2023. URL https://openreview.net/forum?id=r0BrY4BiEX0.

Kanchana Vaishnavi Gandikota, **Jonas Geiping**, Zorah Lähner, Adam Czapliński, and Michael Moeller. A Simple Strategy to Provable Invariance via Orbit Mapping. In *Asian Conference on Computer Vision (ACCV)*, Macau, December 2022. arXiv. doi: 10.48550/arXiv.2209.11916. URL http://arxiv.org/abs/2209.11916.

Jonas Geiping and Tom Goldstein. Cramming: Training a Language Model on a single GPU in one day. In *Proceedings of the 40th International Conference on Machine Learning*, pages 11117–11143. PMLR, July 2023. URL https://proceedings.mlr.press/v202/geiping23a.html.

Jonas Geiping and Michael Moeller. Composite Optimization by Nonconvex Majorization-Minimization. *SIAM Journal on Imaging Sciences*, pages 2494–2528, January 2018. doi: 10.1137/18M1171989. URL https://epubs.siam.org/doi/10.1137/18M1171989.

Jonas Geiping and Michael Moeller. Parametric Majorization for Data-Driven Energy Minimization Methods. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 10262–10273, 2019. URL http://openaccess.thecvf.com/content_ICCV_2019/html/Geiping_Parametric_Majorization_for_Data-Driven_Energy_Minimization_Methods_ICCV_2019_paper.html.

Jonas Geiping*, Hendrik Dirks*, Daniel Cremers, and Michael Moeller. Multiframe Motion Coupling for Video Super Resolution. In Marcello Pelillo and Edwin Hancock, editors, *Energy Minimization Methods in Computer Vision and Pattern Recognition*, Lecture Notes in Computer Science, pages 123–138. Springer International Publishing, 2018. ISBN 978-3-319-78199-0. doi: 10.1007/978-3-319-78199-0_9.

Jonas Geiping*, Hartmut Bauermeister*, Hannah Dröge*, and Michael Moeller. Inverting Gradients - How easy is it to break privacy in federated learning? In *Advances in Neural Information Processing Systems*, volume 33, December 2020a. URL https://proceedings.neurips.cc/paper/2020/hash/c4ede56bbd98819ae6112b20ac6bf145-Abstract.html.

Jonas Geiping, Fjedor Gaede, Hartmut Bauermeister, and Michael Moeller. Fast Convex Relaxations using Graph Discretizations. In *31st British Machine Vision Conference (BMVC 2020, Oral Presentation)*, Virtual, September 2020b. URL https://www.bmvc2020-conference.com/conference/papers/paper_0694.html.

Jonas Geiping, Liam Fowl, Gowthami Somepalli, Micah Goldblum, Michael Moeller, and Tom Goldstein. What Doesn't Kill You Makes You Robust(er): Adversarial Training against Poisons and Backdoors. In *ICLR 2021 Workshop on Security and Safety in Machine Learning Systems*, February 2021a. URL http://arxiv.org/abs/2102.13624.

Jonas Geiping*, Liam H. Fowl*, W. Ronny Huang, Wojciech Czaja, Gavin Taylor, Michael Moeller, and Tom Goldstein. Witches' Brew: Industrial Scale Data Poisoning via Gradient Matching. In *International Conference on Learning Representations*, April 2021b. URL https://openreview.net/forum?id=01olnfLIbD.

Jonas Geiping, Micah Goldblum, Phil Pope, Michael Moeller, and Tom Goldstein. Stochastic Training is Not Necessary for Generalization. In *International Conference on Learning Representations*, September 2021c. URL https://openreview.net/forum?id=ZBESeIUB5k.

Jonas Geiping, Micah Goldblum, Gowthami Somepalli, Ravid Shwartz-Ziv, Tom Goldstein, and Andrew Gordon Wilson. How Much Data Are Augmentations Worth? An Investigation into Scaling Laws, Invariance, and Implicit Regularization. In *International Conference on Learning Representations*, February 2023. URL https://openreview.net/forum?id=3aQs3MCSexD.

Micah Goldblum*, **Jonas Geiping***, Avi Schwarzschild, Michael Moeller, and Tom Goldstein. Truth or backpropaganda? An empirical investigation of deep learning theory. In *Eighth International Conference on Learning Representations (ICLR 2020, Oral Presentation)*, April 2020. URL https://iclr.cc/virtual_2020/poster_HyxyIgHFvr.html.

Andreas Görlitz, **Jonas Geiping**, and Andreas Kolb. Piecewise Rigid Scene Flow with Implicit Motion Segmentation. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1758–1765, November 2019. doi: 10.1109/IROS40897.2019.8968018.

W. Ronny Huang*, **Jonas Geiping***, Liam Fowl, Gavin Taylor, and Tom Goldstein. MetaPoison: Practical General-purpose Clean-label Data Poisoning. In *Advances in Neural Information Processing Systems*, volume 33, Vancouver, Canada, December 2020. URL https://proceedings.neurips.cc//paper_files/paper/2020/hash/8ce6fc704072e351679ac97d4a985574-Abstract.html.

John Kirchenbauer*, **Jonas Geiping***, Yuxin Wen, Jonathan Katz, Ian Miers, and Tom Goldstein. A Watermark for Large Language Models. In *Proceedings of the 40th International Conference on Machine Learning*, pages 17061–17084. PMLR, July 2023. URL https://proceedings.mlr.press/v202/kirchenbauer23a.html.

Jie Li, Yow-Ting Shiue, Yong-Siang Shih, and **Jonas Geiping**. Augmenters at SemEval-2023 Task 1: Enhancing CLIP in Handling Compositionality and Ambiguity for Zero-Shot Visual WSD through Prompt Augmentation and Text-To-Image Diffusion. In *Proceedings of the The 17th International Workshop on Semantic Evaluation (SemEval-2023)*, pages 44–49, Toronto, Canada, July 2023. Association for Computational Linguistics. URL https://aclanthology.org/2023.semeval-1.5.

Jovita Lukasik*, **Jonas Geiping***, Michael Moeller**, and Margret Keuper**. Differentiable Architecture Search: A One-Shot Method? In *AutoML Conference 2023*, August 2023. URL https://openreview.net/forum?id=LV-5kHj-uV5.

Khalid Saifullah*, Yuxin Wen*, **Jonas Geiping**, Micah Goldblum, and Tom Goldstein. Seeing in Words: Learning to Classify through Language Bottlenecks. *ICLR TinyPapers*, May 2023. URL https://openreview.net/forum?id=_QreMdMNIz-.

Pedro Sandoval-Segura, Vasu Singla, Liam Fowl, **Jonas Geiping**, Micah Goldblum, David Jacobs, and Tom Goldstein. Poisons that are learned faster are more effective. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 197–204, June 2022a. doi: 10.1109/CVPRW56347.2022.00033.

Pedro Sandoval-Segura, Vasu Singla, **Jonas Geiping**, Micah Goldblum, Tom Goldstein, and David W. Jacobs. Autoregressive Perturbations for Data Poisoning. In *Advances in Neural Information Processing Systems*, December 2022b. URL https://openreview.net/forum?id=1vusesyN7E.

Gowthami Somepalli, Vasu Singla, Micah Goldblum, **Jonas Geiping**, and Tom Goldstein. Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6048–6058, 2023. URL https://openaccess.thecvf.com/content/CVPR2023/html/Somepalli_Diffusion_Art_or_Digital_Forgery_Investigating_Data_Replication_in_Diffusion_CVPR_2023_paper.html.

Yuxin Wen*, **Jonas Geiping***, Liam Fowl, Micah Goldblum, and Tom Goldstein. Fishing for User Data in Large-Batch Federated Learning via Gradient Magnification. In *Proceedings of the 39th International Conference on Machine Learning*, pages 23668–23684. PMLR, June 2022a. URL https://proceedings.mlr.press/v162/wen22a.html.

Yuxin Wen*, **Jonas Geiping***, Liam Fowl, Hossein Souri, Rama Chellappa, Micah Goldblum, and Tom Goldstein. Thinking Two Moves Ahead: Anticipating Other Users Improves Backdoor Attacks in Federated Learning. In *AdvML Frontiers Workshop at 39th International Conference on Machine Learning*, Baltimore, Maryland, USA, 2022b. arXiv. doi: 10.48550/arXiv.2210.09305. URL https://advml-frontier.github.io/pdf/31/CameraReady/backdoor_fl.pdf.

Yuxin Wen, Arpit Bansal, Hamid Kazemi, Eitan Borgnia, Micah Goldblum, **Jonas Geiping**, and Tom Goldstein. Canary in a Coalmine: Better Membership Inference with Ensembled Adversarial Queries. In *International Conference on Learning Representations*, February 2023a. URL https://openreview.net/forum?id=b7SBTEBFnC.

Yuxin Wen, **Jonas Geiping**, Micah Goldblum, and Tom Goldstein. STYX: Adaptive Poisoning Attacks Against Byzantine-Robust Defenses in Federated Learning. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5, June 2023b. doi: 10.1109/ICASSP49357.2023.10096606.

Recent Preprints

Randall Balestriero, Mark Ibrahim, Vlad Sobal, Ari Morcos, Shashank Shekhar, Tom Goldstein, Florian Bordes, Adrien Bardes, Gregoire Mialon, Yuandong Tian, Avi Schwarzschild, Andrew Gordon Wilson, **Jonas Geiping**, Quentin Garrido, Pierre Fernandez, Amir Bar, Hamed Pirsiavash, Yann LeCun, and Micah Goldblum. A Cookbook of Self-Supervised Learning. *arxiv:2304.12210[cs]*, April 2023. doi: 10.48550/arXiv.2304.12210. URL http://arxiv.org/abs/2304.12210.

Arpit Bansal, Eitan Borgnia, Hong-Min Chu, Jie S. Li, Hamid Kazemi, Furong Huang, Micah Goldblum, **Jonas Geiping**, and Tom Goldstein. Cold Diffusion: Inverting Arbitrary Image Transforms Without Noise. *arxiv:2208.09392[cs]*, August 2022. doi: 10.48550/arXiv.2208.09392. URL http://arxiv.org/abs/2208.09392.

Neel Jain, John Kirchenbauer, **Jonas Geiping**, and Tom Goldstein. How to Do a Vocab Swap? A Study of Embedding Replacement for Pre-trained Transformers. November 2022. URL https://openreview.net/forum?id=MsjB2ohCJO1.

Neel Jain*, Khalid Saifullah*, Yuxin Wen, John Kirchenbauer, Manli Shu, Aniruddha Saha, Micah Goldblum, **Jonas Geiping**, and Tom Goldstein. Bring Your Own Data! Self-Supervised Evaluation for Large Language Models. *arxiv:2306.13651[cs]*, June 2023a. URL http://arxiv.org/abs/2306.13651.

Neel Jain, Avi Schwarzschild, Yuxin Wen, Gowthami Somepalli, John Kirchenbauer, Ping-yeh Chiang, Micah Goldblum, Aniruddha Saha, **Jonas Geiping**, and Tom Goldstein. Baseline Defenses for Adversarial Attacks Against Aligned Language Models. *arxiv:2309.00614[cs]*, September 2023b. doi: 10.48550/arXiv.2309.00614. URL http://arxiv.org/abs/2309.00614.

John Kirchenbauer*, **Jonas Geiping***, Yuxin Wen, Manli Shu, Khalid Saifullah, Kezhi Kong, Kasun Fernando, Aniruddha Saha, Micah Goldblum, and Tom Goldstein. On the Reliability of Watermarks for Large Language Models. *arxiv:2306.04634[cs]*, June 2023. doi: 10.48550/arXiv.2306.04634. URL http://arxiv.org/abs/2306.04634.

Renkun Ni, Ping-yeh Chiang, **Jonas Geiping**, Micah Goldblum, Andrew Gordon Wilson, and Tom Goldstein. K-SAM: Sharpness-Aware Minimization at the Speed of SGD. *arxiv:2210.12864[cs]*, October 2022. doi: 10.48550/arXiv.2210.12864. URL http://arxiv.org/abs/2210.12864.

Pedro Sandoval-Segura, **Jonas Geiping**, and Tom Goldstein. JPEG Compressed Images Can Bypass Protections Against AI Editing. *arxiv:2304.02234[cs]*, April 2023a. doi: 10.48550/arXiv.2304.02234. URL http://arxiv.org/abs/2304.02234.

Pedro Sandoval-Segura, Vasu Singla, **Jonas Geiping**, Micah Goldblum, and Tom Goldstein. What Can We Learn from Unlearnable Datasets? *arxiv:2305.19254[cs]*, May 2023b. doi: 10.48550/arXiv.2305.19254. URL http://arxiv.org/abs/2305.19254.

Manli Shu, Jiongxiao Wang, Chen Zhu, **Jonas Geiping**, Chaowei Xiao, and Tom Goldstein. On the Exploitability of Instruction Tuning. *arxiv:2306.17194[cs]*, June 2023. doi: 10.48550/arXiv.2306.17194. URL http://arxiv.org/abs/2306.17194.

Gowthami Somepalli, Vasu Singla, Micah Goldblum, **Jonas Geiping**, and Tom Goldstein. Understanding and Mitigating Copying in Diffusion Models. *arxiv:2305.20086[cs]*, May 2023. doi: 10.48550/arXiv.2305.20086. URL http://arxiv.org/abs/2305.20086.

Yuxin Wen*, Neel Jain*, John Kirchenbauer, Micah Goldblum, **Jonas Geiping**, and Tom Goldstein. Hard Prompts Made Easy: Gradient-Based Discrete Optimization for Prompt Tuning and Discovery. February 2023a. URL https://arxiv.org/abs/2302.03668v1.

Yuxin Wen, John Kirchenbauer, **Jonas Geiping**, and Tom Goldstein. Tree-Ring Watermarks: Fingerprints for Diffusion Images that are Invisible and Robust. *arxiv:2305.20030[cs]*, May 2023b. doi: 10.48550/arXiv.2305.20030. URL http://arxiv.org/abs/2305.20030.

*Shared authorship denoted by * and †. Alphabetic ordering denoted by ‡.*