

# Zachary Novack

[zacharynovack.github.io](https://zacharynovack.github.io)

znovack@ucsd.edu

---

## RESEARCH INTERESTS

Generative Audio/Music, Controllable Generation, Generative Model Acceleration, Audio-Language Reasoning

---

## EDUCATION

*Ph.D. in Computer Science*

Fall 2022 - Present

University of California – San Diego, San Diego, CA

Advisors: Julian McAuley, Taylor Berg-Kirkpatrick

*M.S. in Computer Science*

Fall 2022 - Spring 2024

University of California – San Diego, San Diego, CA

Advisors: Julian McAuley, Taylor Berg-Kirkpatrick

*B.S. in Statistics & Machine Learning*

August 2018 - May 2022

Carnegie Mellon University, Pittsburgh, PA

Advisors: Zachary Lipton, Simon DeDeo

- Minor in Sonic Arts (music technology)
- 3.93/4.0 GPA

---

## SELECTED RESEARCH

**DITTO: Diffusion Inference-Time T-Optimization for Music Generation.**

Zachary Novack, Julian McAuley, Taylor Berg-Kirkpatrick, Nicholas J. Bryan.

ICML (**Oral, Top 1.5%**) 2024.

**Presto! Distilling Steps and Layers for Accelerating Music Generation.**

Zachary Novack, Ge Zhu, Jonah Casebeer, Julian McAuley, Taylor Berg-Kirkpatrick, Nicholas J. Bryan

ICLR (**Spotlight, Top 5.1%**) 2025.

**Are you really listening? Boosting Perceptual Awareness in Music-QA Benchmarks**

Yongyi Zang, Sean O'Brien, Taylor Berg-Kirkpatrick, Julian McAuley, **Zachary Novack**

ISMIR (**Best Paper Award Winner**) 2025.

**Fast Text-to-Audio Generation with Adversarial Post-Training**

Zachary Novack, Zach Evans, Zack Zukowski, Josiah Taylor, CJ Carr, Julian Parker, Adnan Al-Sinan,

Gian Marco Iodice, Julian McAuley, Taylor Berg-Kirkpatrick, Jordi Pons

WASPAA 2025.

**DITTO-2: Distilled Diffusion Inference-Time T-Optimization for Music Generation.**

Zachary Novack, Julian McAuley, Taylor Berg-Kirkpatrick, Nicholas J. Bryan.

ISMIR 2024.

**CHiLS: Zero-Shot Image Classification with Hierarchical Label Sets.**

Zachary Novack, Julian McAuley, Zachary Lipton, Saurabh Garg.

ICML 2023.

## SELECTED ACCOLADES

---

Best Paper Award, ISMIR 2025	Fall 2025
UCSD CSE Doctoral Award for Excellence in Leadership	Spring 2025
Malcolm R. Stacey Memorial Fellowship	Spring 2024
1 <sup>st</sup> Place: Adobe Intern Project Expo	August 2023
NSF Graduate Research Fellowship - Honorable Mention	Spring 2022
Phi Beta Kappa Member	October 2021 - Present
<a href="#">Andrew Carnegie Society Scholar</a>	September 2021 - Present
Small Undergraduate Research Grant (SURG)	June 2021
Dietrich Senior Honors Research Fellowship	May 2021
1 <sup>st</sup> Place: Statistics & Data Science Research Showcase	May 2021
Summer Undergraduate Research Fellowship (SURF)	June 2020
2 <sup>nd</sup> Place: 15-112 Term Project Showcase	April 2019
Dean's List: High Honors	December 2018 - May 2022
<a href="#">Quantitative Social Science Scholar</a>	August 2018 - May 2022
Paul Mellon Memorial Presidential Scholarship	August 2018 - May 2022

## INVITED TALKS

---

### **Presto! Distilling Steps and Layers for Accelerating Music Generation**

- Boston AI Music Meetup Host: Christian Steinmetz, January 2025
- MIT HAI-Res Group Host: Anna Huang, January 2025
- UCSD GenAI Summit Host: Rose Yu, February 2025
- National University of Singapore (NUS)  
Speech and Music AI Workshop Host: Wang Ye, April 2025
- Sony Research Host: Christian Simon, June 2025

### **DITTO: Diffusion Inference-Time T-Optimization for Music Generation**

- University of Rochester AI Audio Lab Host: Zhiyao Duan, December 2023
- MIT AI Music Reading Group Host: Lancelot Blanchard, February 2024
- Spotify MIQ Reading Group Host: Yu Wang, February 2024
- UC San Diego AI Seminar Host: Rose Yu, March 2024
- ICML Oral July 2024
- BISH Bash – Adobe Host: Oriol Nieto, August 2024

### **Unsupervised Lead Sheet Generation via Semantic Compression**

- AES AI & the Musician Symposium June 2024

### **Disentangling the Mechanisms Behind Implicit Regularization in SGD**

- HOOML Workshop, NeurIPS December 2022

## PAPERS

---

### *Tutorials*

- [Connecting Music Audio and Natural Language](#)  
SeungHeon Doh, Ilaria Manco, **Zachary Novack**, Jong Wook Kim, Ke Chen  
International Society of Music Information Retrieval (ISMIR), 2024

### Preprints

- **Steering Autoregressive Music Generation with Recursive Feature Machines**  
Daniel Zhao, Daniel Beaglehole, Taylor Berg-Kirkpatrick, Julian McAuley, **Zachary Novack**, 2025
- **MuseTok: Symbolic Music Tokenization for Generation and Semantic Understanding**  
Jingyue Huang, **Zachary Novack**, Phillip Long, Yupeng Hou, Ke Chen, Taylor Berg-Kirkpatrick, Julian McAuley, 2025
- **SoundReactor: Frame-level Online Video-to-Audio Generation**  
Koichi Saito, Julian Tanke, Christian Simon, Masato Ishii, Kazuki Shimada, **Zachary Novack**, Zhi Zhong, Akio Hayakawa, Takashi Shibuya, Yuki Mitsufuji, 2025
- **Zephyrus: An Agentic Framework for Weather Science**  
Sumanth Varambally, Marshall Fisher, Jas Thakker, Yiwei Chen, Zhirui Xia, Yasaman Jafari, Ruijia Niu, Manas Jain, Veeramakali Vignesh Manivannan, **Zachary Novack**, Luyu Han, Srikanth Eranky, Salva Rühling Cachay, Taylor Berg-Kirkpatrick, Duncan Watson-Parris, Yi-An Ma, Rose Yu, 2025
- **MusiCRS: Benchmarking Audio-Centric Conversational Recommendation**  
Rohan Surana, Amit Namburi, Gagan Mundada, Abhay Lal, **Zachary Novack**, Julian McAuley, Junda Wu, 2025
- **Bob’s Confetti: Phonetic Memorization Attacks in Music and Video Generation**  
**Zachary Novack\***, Jaechul Roh\*, Yuefeng Peng, Niloofar Mireshghallah, Taylor Berg-Kirkpatrick, Amir Houmansadr, 2025
- **WildFX: A DAW-Powered Pipeline for In-the-Wild Audio FX Graph Modeling**  
Qihui Yang, Taylor Berg-Kirkpatrick, Julian McAuley, **Zachary Novack**, 2025

### Conference Papers

- **Repurposing Marigold for Zero-Shot Metric Depth Estimation via Defocus Blur Cues**  
Chinmay Talegaonkar, Nikhil Gandudi Suresh, **Zachary Novack**, Yash Belhe, Priyanka Nagasamudra, Nicholas Antipa  
**(Spotlight, Top 3.2%)** at Neural Information Processing Systems (NeurIPS), 2025
- **WildScore: Benchmarking MLLMs in-the-Wild Symbolic Music Reasoning**  
Gagan Mundada, Yash Vishe, Amit Namburi, Xin Xu, **Zachary Novack**, Julian McAuley, Junda Wu  
Empirical Methods in Natural Language Processing (EMNLP), 2025
- **Fast Text-to-Audio Generation with Adversarial Post-Training**  
**Zachary Novack**, Zach Evans, Zack Zukowski, Josiah Taylor, CJ Carr, Julian Parker, Adnan Al-Sinan, Gian Marco Iodice, Julian McAuley, Taylor Berg-Kirkpatrick, Jordi Pons  
IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2025
- **Are you really listening? Boosting Perceptual Awareness in Music-QA Benchmarks**  
Yongyi Zang, Sean O’Brien, Taylor Berg-Kirkpatrick, Julian McAuley, **Zachary Novack**  
**(Best Paper Award Winner)** at International Society of Music Information Retrieval (ISMIR), 2025
- **Video-Guided Text-to-Music Generation Using Public Domain Movie Collections**  
Haven Kim, **Zachary Novack**, Weihang Xu, Julian McAuley, Hao-Wen Dong  
International Society of Music Information Retrieval (ISMIR), 2025
- **Aligning Text-to-Music Evaluation with Human Preferences**  
Yichen Huang, **Zachary Novack**, Koichi Saito, Jiatong Shi, Shinji Watanabe, Yuki Mitsufuji, John Thickstun, Chris Donahue  
International Society of Music Information Retrieval (ISMIR), 2025
- **Deriving Representative Structure from Music Corpora.**  
Ilana Shapiro, Ruanqianqian Huang, **Zachary Novack**, Cheng-i Wang, Hao-Wen Dong, Taylor Berg-Kirkpatrick, Shlomo Dubnov, Sorin Lerner

Special Track for AI, Arts, and Creativity at International Joint Conference on Artificial Intelligence (IJCAI), 2025

- **Presto: Distilling Steps and Layers for Accelerating Music Generation.**  
Zachary Novack, Ge Zhu, Jonah Casebeer, Julian McAuley, Taylor Berg-Kirkpatrick, Nicholas J. Bryan  
(**Spotlight, Top 5.1%**) at International Conference on Learning Representations (ICLR), 2025
- **CoLLAP: Contrastive Long-form Language-Audio Pretraining with Musical Temporal Structure Augmentation.**  
Junda Wu, Warren Li, Zachary Novack, Amit Namburi, Carol Chen, Julian McAuley  
(**Oral**) at International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2025  
SoCal NLP Symposium, 2024
- **PDMX: A Large-Scale Public Domain MusicXML Dataset for Symbolic Music Processing.**  
Phillip Long, Zachary Novack, Taylor Berg-Kirkpatrick, Julian McAuley  
International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2025  
NeurIPS Workshop on Creativity & Generative AI, 2024
- **FUTGA-MIR: Enhancing Fine-grained and Temporally-aware Music Understanding with Music Information Retrieval**  
Junda Wu, Zachary Novack, Amit Namburi, Jiaheng Dai, Hao-Wen Dong, Zhouhang Xie, Carol Chen, Julian McAuley  
International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2025
- **DITTO-2: Distilled Diffusion Inference-Time T-Optimization for Music Generation.**  
Zachary Novack, Julian McAuley, Taylor Berg-Kirkpatrick, Nicholas J. Bryan  
International Society of Music Information Retrieval (ISMIR), 2024
- **DITTO: Diffusion Inference-Time T-Optimization for Music Generation.**  
Zachary Novack, Julian McAuley, Taylor Berg-Kirkpatrick, Nicholas J. Bryan  
(**Oral Top 1.5%**) at International Conference on Machine Learning (ICML), 2024
- **CHiLS: Zero-Shot Image Classification with Hierarchical Label Sets**  
Zachary Novack, Julian McAuley, Zachary Lipton, Saurabh Garg  
International Conference on Machine Learning (ICML), 2023  
ICLR Workshop on Multimodal Representation Learning, 2023
- **Disentangling the Mechanisms Behind Implicit Regularization in SGD**  
Zachary Novack, Simran Kaur, Tanya Marwah, Saurabh Garg, Zachary Lipton  
International Conference on Learning Representations (ICLR), 2023  
**Spotlight** and **Best Poster** at NeurIPS Workshop on The Benefits of Higher-Order Optimization in Machine Learning, 2022

#### *Workshops*

- **FUTGA: Towards Fine-grained Music Understanding through Temporally-Enhanced Generative Augmentation**  
Junda Wu, Zachary Novack, Amit Namburi, Jiaheng Dai, Hao-Wen Dong, Zhouhang Xie, Carol Chen, Julian McAuley  
3rd Workshop on NLP for Music and Audio, 2024
- **Unsupervised Lead Sheet Generation via Semantic Compression**  
Zachary Novack, Nikita Srivatsan, Taylor Berg-Kirkpatrick, Julian McAuley  
AES International Symposium on AI and the Musician, 2024

#### *Nonrefereed Papers*

- **Down the Rabbit Hole: Modeling Twitter Dynamics through Bayesian Inference**  
Zachary Novack  
Senior Honors Thesis (Carnegie Mellon University), 2022
- **Personalized Sequential Recommendation for Adaptive Itemization in MOBA Games**  
Zachary Novack  
Web Mining and Recommender Systems (CSE 258) Course Project (UC San Diego), 2022
- **Towards Generalizable Deep Speech Anonymization**  
Aaron Broukhim, Zachary Novack  
Deep Generative Models (CSE 291) Course Project (UC San Diego), 2022
- **Approximating Optimal Transport via GANs for Recourse Disparity Analysis**  
Zachary Novack, Qi Xuan Teo, Ryan Steed  
Probabilistic Graphic Models (10-708) Course Project (Carnegie Mellon University), 2022
- **Tracking Political Sentiment on Cold War China in Congressional Speeches**  
Zachary Novack, Eden Hu, and Mason Lin  
**1st Place** at Statistics and Data Science Research Showcase (Carnegie Mellon University), 2021
- **Lunch at the EigenSalad Bar: Linear Approaches to Dimensionality Reduction for Image Processing**  
Zachary Novack  
Numerical Linear Algebra (21-344) Course Project (Carnegie Mellon University), 2021

#### *Blog Posts*

- **Armchair Statistics: Benford's Law and other Misconceptions in the Age of Data**  
Zachary Novack  
Carnegie Mellon University Triple Helix, 2021

## TEACHING EXPERIENCE

---

#### *Co-Instructor*

University of California - San Diego, San Diego, CA

- CSE 253: Machine Learning for Music Spring 2025  
*Prof. Julian McAuley*

#### *Graduate Teaching Assistant*

University of California - San Diego, San Diego, CA

- CSE 258: Web Mining and Recommender Systems Fall 2023, 2024  
*Prof. Julian McAuley*

#### *Undergraduate Teaching Assistant*

Carnegie Mellon University, Pittsburgh, PA

- 10-600: Machine Learning Primer Summer 2022  
*Prof. Matthew Gormley*
- 10-301/601: Introduction to Machine Learning Fall 2021 - Summer 2022  
*Prof. Matthew Gormley and Henry Chai*
- 85-340: Research Methods for Social Psychology Fall 2021  
*Prof. David Creswell*
- 36-225: Introduction to Probability Theory Summer 2021  
*Prof. Peter Freeman*
- 36-226: Introduction to Statistical Inference Spring 2021  
*Prof. Peter Freeman and Nynke Niezink*

- 88-300: Programming for Social Scientists Summer 2020 - Spring 2021  
Prof. Mark Patterson

## WORK EXPERIENCE

---

**Stability AI – Audio Group** October 2024 – Present  
Research Scientist Intern under CJ Carr and Zach Evans

**Adobe – Audio Group** June 2023 - October 2024  
Research Scientist Intern under Nicholas Bryan

- Investigating methods for [interactive control](#) (ICML 2024) and [efficient generation](#) (ISMIR 2024) for audio-domain generative music models.

**ACMI Lab** (CMU) Spring 2021 - Spring 2023  
Research Assistant under Zachary Lipton

- Developed new method to leverage hierarchical class information for zero-shot prediction in CLIP models (ICML 2023).
- Performed large-scale verification study validate explicit regularization mechanisms for SGD across modern image benchmarks and model types (ICLR 2023).

**Laboratory for Social Minds** (CMU) Summer 2020 - Fall 2022  
Research Assistant under Simon DeDeo

- Designed a temporal Bayesian framework to analyze social media addiction.
- Investigated ideological network evolution on the fringe web forums /pol/ (4chan) and The Red Pill (Reddit).

**Unisys Corporation** Summer 2020 - Spring 2021  
AI/ML Intern

- Designed time-series models (ARIMA, LSTM, Facebook Prophet) for computer resource utilization prediction under distribution shift

## ACADEMIC SERVICE

---

**Organizing Committee:** ISMIR (2025, Jam Session Host), AI4Music Workshop - NeurIPS (2025)

**Program Committee:** NLP4MusA Workshop - ISMIR (2024)

**Reviewer:** ICLR (2023-2024), ICASSP (2023-2024), NeurIPS (2023-2025), ISMIR (2024-2025)

**Ph.D. Admissions Committee:** UCSD (CSE, 2023)

**Ph.D. Visit Day Committee:** UCSD (CSE, 2023-2024 AI/ML Area Lead, 2025 General Chair)

## MENTORSHIP

---

**Independent Researcher: Yongyi Zang**

- Now **Head of Research @ Smule**
- Are you really listening? Boosting Perceptual Awareness in Music-QA Benchmarks, ISMIR 2025

**Bachelors in Computer Science, UCSD Student: Phillip Long**

- PDMX: A Large-Scale Public Domain MusicXML Dataset for Symbolic Music Processing, ICASSP 2025

**PhD in Computer Science, UCSD Student: Ilana Shapiro**

- Deriving Representative Structure from Music Corpora, IJCAI 2025

**PhD in Computer Science, UCSD Student: Haven Kim**

- Video-Guided Text-to-Music Generation Using Public Domain Movie Collections, ISMIR 2025

**PhD in Computer Science, UCSD Student: Jingyue Huang**

- MuseTok: Symbolic Music Tokenization for Generation and Semantic Understanding, under submission

**Visiting Researcher, CMU Student: Yichen Huang**

- Aligning Text-to-Music Evaluation with Human Preferences, ISMIR 2025

**Bachelors in Computer Science, UCSD Student: Amit Namburi**

- Now **Masters in Computer Science, UCSD**
- PDMX: A Large-Scale Public Domain MusicXML Dataset for Symbolic Music Processing, ICASSP 2025
- FUTGA: Towards fine-grained music understanding through temporally-enhanced generative augmentation, ICASSP 2025
- CoLLAP: Contrastive Long-form Language-Audio Pretraining with Musical Temporal Structure Augmentation, ICASSP 2025

**Masters in Computer Engineering, UCSD Student: Qihui Yang**

- WildFX: A DAW-Powered Pipeline for In-the-Wild Audio FX Graph Modeling, under submission

## MUSICAL ACTIVITIES

---

### Teaching Experience

*Front Ensemble Technician* Fall 2023 - Present  
[POW Percussion Ensemble](#), Anaheim, CA

*Audio Team* Summer 2023 - Summer 2024  
[Pacific Crest Drum & Bugle Corps](#), Diamond Bar, CA

- Facilitated design and live interfacing with large-scale audio rig for 150 active performers

*Front Ensemble Coordinator* Fall 2019 - Summer 2020  
[Gateway Senior High School](#), Monroeville, PA

- Led rehearsals and designed pedagogical structure for the front ensemble (non-mobile percussion) in Gateway's marching band and indoor percussion programs, working with a group of 10-15 students from ages 14-18.

*Performer and Composer* Spring 2019 - Spring 2020  
[Exploded Ensemble](#), Carnegie Mellon University, Pittsburgh, PA

- Designed large-scale Max/MSP programs for multimedia interactive performances
- Composed electro-acoustic pieces for mixed instrumentation ensembles

*Percussion Arranger* Fall 2018 - Spring 2019  
[Tomball High School Indoor Percussion](#), Tomball, TX

- Arranged musical production for large percussion ensemble in order to compete in the Winter Guard International (WGI) national circuit

### Projects

[RoboPierre](#) Spring 2020

Adaptive Impressionist Music via Generative Modeling

- Developed interactive web app to randomly generate polyphonic music trained on impressionistic composers
- Implemented using Google Magenta's Polyphony RNN and custom stochastic voice leading algorithm

*ThereMyn*

Spring 2019

Motion-Controlled Monophonic Synthesizer

- Used infrared distance monitor to drive audio signal creation
- Created front-end GUI to translate audio signals into a usable motion-controlled synthesizer

## SKILLS

---

*Programming Languages and Packages*

- Python (Pytorch, Tensorflow, Scikit-Learn, PySpark, CVXPY), R (dplyr, tscout, zoo), C, Matlab, SQL (postgres, MySQL), Stan, Git, Shell, Max/MSP/Jitter

*Other Skills*

- AWS (S3, EC2, EMR), Microsoft Azure, Docker, Agile, Jira, Grafana, Ableton Live

## SELECTED COURSES

---

**UC San Diego**

Deep Generative Models, Search and Optimization, Information Visualization, Recommender Systems, Computing Education, Math for Robotics

**Carnegie Mellon University**

Probabilistic Graphical Models, Convex Optimization, Multimedia Signal Processing, ML w/Large Datasets, Real Analysis, Numerical Linear Algebra, Probability & Statistics, Statistical Computing, Linear Algebra, Philosophy of ML, Algorithms & Data Structures