

Yinghao Ma (马英浩)

[Email](#) | +44 (0)7741575063 | [nicolaus625.github.io](#) | [Linkedin](#) | [Google Scholar](#)

EDUCATION

Queen Mary University of London (QMUL)	09/2022 – 09/2027 (expected)
<i>PhD: AI & Music, School of Electronic Engineering and Computer Science. Supervised by Dr Emmanouil Benetos</i>	
• Research interests:	Large language models (LLMs) for music understanding & generation
• Awards and Honours:	Google PhD Fellowship 2025-2026 (Machine Perception track)
• Leadership:	Co-founder of the Multimodal Art Projection community (Hugging-face);
Carnegie Mellon University (CMU)	09/2020 – 08/2022
<i>MS: Music & Technology, School of Music. Supervised by Prof. Richard M. Stern</i>	
• Overall GPA:	4.03/4.00 (Top 1 of the grade)
• Awards and Honours:	Fellowship for master students that covers 50% of tuition fee U.S. National Music Honor Society member. (Theta Xi, Pi Kappa Lambda reward).
• Music Background:	Recorded Chinese musical version of Beethoven's serenade for 250th anniversary of his birth during COVID to cheer others up. Released on CMU DL course web
Peking University (PKU)	09/2016 – 07/2020
<i>BS: Mathematics & Applied Mathematics (Probability Theory), School of Mathematical Sciences</i>	
• Awards and Honours:	Outstanding graduates of the School of Mathematical Science at PKU Preliminary excellence prize for S.-T.Yau College Student Mathematics Contests
• Music Background:	One of the student conductors in the orchestra of the Chinese Music Institute at PKU The amateur highest level of Chinese flutes, China Conservatory of Music

(CO)-FIRST AUTHOR PUBLICATION

- Ma, Z.*, **Ma, Y.***, Zhu, Y.*, , et al. "MMAR: A Challenging Benchmark for Deep Reasoning in Speech, Audio, Music, and Their Mix." Advances in Neural Information Processing Systems (NeurIPS), 2025.
- **Ma, Y.**, Li, S., Yu, J. et al. "CMI-Bench: A Comprehensive Benchmark for Evaluating Music Instruction Following." International Society for Music Information Retrieval (ISMIR), 2025.
- Li, Y.*, Zhang, G.*, **Ma, Y.***, et al. "OmniBench: Towards The Future of Universal Omni-Language Models." Advances in Neural Information Processing Systems (NeurIPS), 2025.
- **Ma, Y.**, Øland, A., Ragni, A., et al. "Foundation models for music: A survey." [Huggingface Daily Paper top3. Submission to ACM Computing Survey Journal]
- Qu, X.*, Bai, Y.* , **Ma, Y.***, et al. "MuPT: A Generative Symbolic Music Pretrained Transformer." International Conference on Learning Representations (ICLR) 2025.
- Deng, Z.*, **Ma, Y.***, Liu, Y. et al. "MusiLingo: Bridging Music and Text with Pre-trained Language Models for Music Captioning and Query Response." North American Chapter of the Association for Computational Linguistics 2024.
- Li, D.* , **Ma, Y.***, et al. "Mertech: Instrument Playing Technique Detection Using Self-supervised Pretrained Model with Multi-task Finetuning." International Conference on Acoustics, Speech & Signal Processing (ICASSP) 2024.
- Li, Y.*, Yuan, R.*; Zhang, G.* , **Ma, Y.***; et al. "MERT: Acoustic Music Understanding Model with Large-Scale Self-supervised Training." International Conference on Learning Representations (ICLR) 2024.
- Yuan, R.* , **Ma, Y.***, Li, Y.* , et al. "MARBLE: Music Audio Representation Benchmark for universal Evaluation." Advances in Neural Information Processing Systems (NeurIPS), 2023.
- **Ma, Y.**, Yuan, R., Li, Y., et al. "On the Effectiveness of Speech Self-Supervised Learning for Music." International Society for Music Information Retrieval (ISMIR), 2023.

RESEARCH EXPERIENCE

Controlled Genre-Specific Acoustic Music Generation	12/2024 – 04/2025
<i>Collabourate with Dr Chenghua Lin, University of Manchester</i>	
• Predict loss value of music LLM based on base/medium language model performance.	
OmniBench: Towards The Future of Universal Omni-Language Models	
<i>Collabourate with Dr Wenhao Huang, 01.ai</i>	06/2024 – 09/2024
• Constructed music questions in OmniBench to assess reasoning capabilities across tri-modal inputs (image, audio, text), setting a new standard for evaluating omni-language models in complex, integrated scenarios.	
• Analysis large audio-language models' performance in multi-modal reasoning tasks, uncovering critical limitations in instruction-following and reasoning across modalities.	
Symbolic Music Scaling (SMS) Law for a Symbolic Music GPT	01/2024 – 03/2024

Collabourate with Dr Jie Fu, Hong Kong University of Science and Technology

- Developed and integrated the SMS Law into Symbolic Music GPT, optimising model scalability and efficiency, which led to enhanced performance with constrained computational resources.
- Analyses overfitted loss curve to predict the early stop points during training.

Bridging Music & Text with Pre-trained Models for Music Captioning and QA 07/2023 – 11/2023

Supervised by Dr Emmanouil Benetos, Centre for Digital Music, Queen Mary University of London

Instrument Playing Technique (IPT) Detection on World Music 06/2023 – 09/2023

Supervised by Dr Emmanouil Benetos, Centre for Digital Music, Queen Mary University of London

MARBLE: Music Audio Representation Benchmark for universal Evaluation 01/2023 – 06/2023

Supervised by Dr Emmanouil Benetos, Centre for Digital Music, Queen Mary University of London

- Designing the downstream tasks, datasets, evaluation metrics and state-of-the-art results.
- Implemented the mir_eval metrics with torchmetrics and developing utilisation for sequential tasks.
- Established a fair, reproducible and universal music information retrieval benchmark for future work.

Acoustic Music Understanding Model with Large-Scale Self-supervised Training 08/2022 – 05/2023

Supervised by Dr Emmanouil Benetos, Centre for Digital Music, Queen Mary University of London

- Built self-supervised learning systems, acquiring 50k+ downloading of checkpoints on Huggingface.
- Replaced the pseudo-tag from MFCCs to Chroma music features for harmonic information.
- Utilising deep features like Encodex instead of k-means for scaling up models to 1 B parameters.

Learnable Front End for Music, Speech and Audio (Master thesis) 09/2021 – 07/2022

Master Thesis, Supervised by Prof. Richard Stern, Carnegie Mellon University

Chinese Flute Playing Technique Classification Based on FCNNs (undergraduate thesis) 02/2020 – 05/2020

Undergraduate Thesis, Supervised by Prof. Xiaoou Chen, Peking University

Correspondence between Speech Melody and Pitch Contour in Sichuan Folk Songs 07/2019 – 09/2019

Research Assistant, Supervised by Prof. Zhiyao Duan, University of Rochester

Automatic Musical Instrument Recognition and Timbre Recognition 02/2019 – 07/2019

Research Assistant, Supervised by Prof. Xiaoou Chen, Peking University

WORK EXPERIENCE

Tempo, Beat and Downbeat Detection in Chinese Pop Songs (internship) 06/2020 – 08/2020

Algorithm Engineer, Beijing Deepmusic Technology Co. (Beijing, China)

Cover Song Detection & Evaluation of Automatic Speech Recognition (internship) 05/2021 – 08/2021

Algorithm Engineer, Tencent Holdings Limited. (Beijing, China)

Teaching Assistant & Guest Lecturer of Machine Learning for Signal Processing 08/2021 – 12/2021

Teaching Assistant of Digital Signal Processing 02/2022 – 05/2022

Teaching Assistant of Applied Statistics 09/2023 – 12/2023

Teaching Assistant of Deep Learning for Audio & Music 02/2024 – 04/2024

Multimodal Understanding for Acoustic Music (internship) 08/2024 – 10/2024

Research intern, Research and Development, Yamaha Corporation (Hamamatsu, Japan), by Dr. Akira Maezawa

Exploring Flow Matching for Real-time Controlable Music Generation (internship) 07/2025 – 09/2025

Research intern, Microsoft Research. (Redmond, USD), supervised by Dr. Hannes Gamper

SELECTED ACADEMIC ACHIEVEMENT

Advanced Digital Signal Processing

- 1st rank of the class. Proofread lecture notes' errata, listed in acknowledgements.
- Signal sampling, interpolation, STFT, classical estimation and maximum entropy of PSD function, LPC, adaptive filtering, microphone array and beam forming. MFCC, Introduction to Wavelet Analysis.

Introduction to Deep Learning (A), Scalability Machine Learning (A+), Convex Optimisation (A+)

LEADERSHIP & SKILLS

- One of the organisers of ISMIR 2025 satellite event on [Large Language Models for Music and Audio](#).
- One of the student conductors at the Chinese Music Institute at Peking University.
- Organised charitable activities. social research and publicity for the disabled and people with rare diseases.
- Python (>5k LOC); MATLAB, C (>500 LOC)