

Research Statement

My research goal is to develop generative AI systems that make music creation more expressive, controllable, and accessible—systems that expand the endless possibilities of music by amplifying marginalized traditions and reimagining creative futures. While recent advances in music generation have achieved impressive fidelity, they often overlook cultural diversity and experimental aesthetics, leaving much of the world’s musical richness underrepresented. My work aims to bridge this gap by building models that understand the nuance of human musicianship, support a wide range of creative controls, and integrate seamlessly into co-creative workflows. Along this journey, I have greatly benefited from the support of the research community—through opportunities such as the New-to-ISMIR Paper Mentoring Program, the ICML NewinML workshop, and an ISMIR travel grant, I gained new horizons and perspectives that have shaped my research direction. To further grow as a researcher and contribute back to this community, I intend to pursue a PhD.

I first engaged with this challenge through the lens of *Nanyin*, a UNESCO-recognized Chinese musical tradition. As part of my Master’s research, I constructed a pipa-centric dataset and designed a heterogeneous graph neural network capable of reconstructing heterophonic accompaniments. This project not only addressed the scarcity of digital resources for Nanyin but also underscored my broader interest in how generative models can support low-resource and culturally diverse musical practices. It taught me that music AI must balance technical innovation with cultural sensitivity, and it inspired my long-term vision of human-centered generative music systems. I believe that revitalizing endangered traditions through AI is not a matter of static preservation but of dynamic reinvention—much like how jazz evolved through fusion and continual innovation.

Building on this foundation, I have been exploring expressive and unified generative modeling frameworks that connect symbolic, audio, and multimodal representations. For example, I proposed narrative-driven multi-track generation for Nanyin by modeling its heterophonic textures. These works reflect my belief that advances in representation learning and tokenization are essential for capturing musical nuance, whether in the ornamentations of traditional performance or in the irregular structures of electronic genres I personally enjoy, such as glitch and IDM. I plan to continue probing how dynamic audio chunking and flexible tokenization can open new avenues for both cultural and experimental music.

Another strand of my research investigates co-creative generative interfaces. My analysis of

interpolation failures in MusicVAE revealed how conventional architectures encode deep-seated stylistic biases, excluding more experimental aesthetics. Motivated by this, I have begun designing interpolation-driven exploration tools that allow users to traverse the “in-between” spaces of musical style, transforming model limitations into new creative opportunities. I see this as part of a larger agenda of designing interactive, human-centered systems that respond to the needs of both professional musicians and beginners—systems that might accompany a live performance in real time, or guide novices by dynamically shaping musical energy and structure.

Looking forward, I envision generative models that are not only technically capable but also musically aware, combining interpretability, adaptability, and collaboration. This vision extends to several concrete directions: integrating physics-informed synthesis with diffusion and token-based models to capture the subtle gestures of master performers; developing embodied interfaces for AI-driven virtual bands where users can inhabit avatars and perform in real time; and building evaluation protocols that move beyond static metrics to reflect the dynamics of true human–AI collaboration.

By pursuing these directions, my aim is to expand the role of generative AI in music from a passive tool into an active creative partner—one that empowers marginalized traditions, preserves cultural knowledge, and opens new spaces for improvisation, performance, and discovery.