

Adaptive representation learning for the gestural control of deep audio generative models

Sarah Nabi^{1,2}, Philippe Esling¹, Frédéric Bevilacqua² and Geoffroy Peeters³

¹Artificial Creative Intelligence and Data Science (ACIDS), Sound Analysis-Synthesis, UMR 9912 STMS - IRCAM - CNRS - Sorbonne, Paris

²Sound Music Movement Interaction (ISMM), UMR 9912 STMS - IRCAM - CNRS - Sorbonne, Paris

³Information Processing and Communications Laboratory (LTCI), Telecom Paris, Institut Polytechnique de Paris, CNRS - UMR 9217, Paris

sarah.nabi@ircam.fr

Problem statement

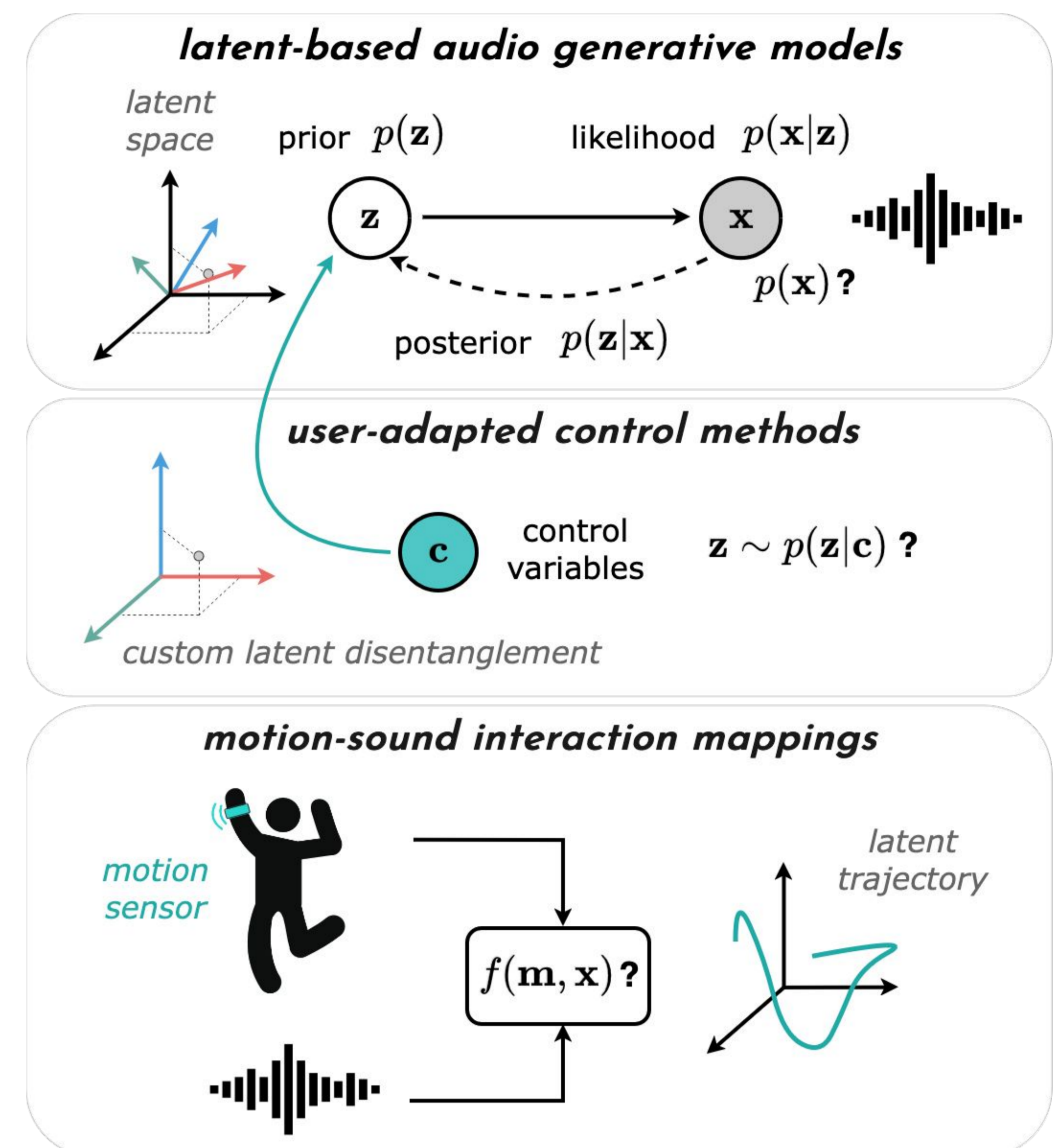
In recent years, **deep generative models** have achieved impressive results in generating high-quality audio samples offering promising tools for musical creation. However, the **lack of control and interactivity** of these approaches preclude their creative use. Existing control methods mainly rely on massive sets of labelled examples and fail to address the **need for intuitive and personalised control**. Musical creation is an iterative and exploratory process where motion and sound are intrinsically linked and can only rely on very limited datasets that match closely a given artistic intent.

Research objectives

The goal of this PhD research is to leverage **deep audio generative models as new musical instruments suitable for creative endeavors, which could be easily customized in terms of both control and gestural interaction**. We aim to provide

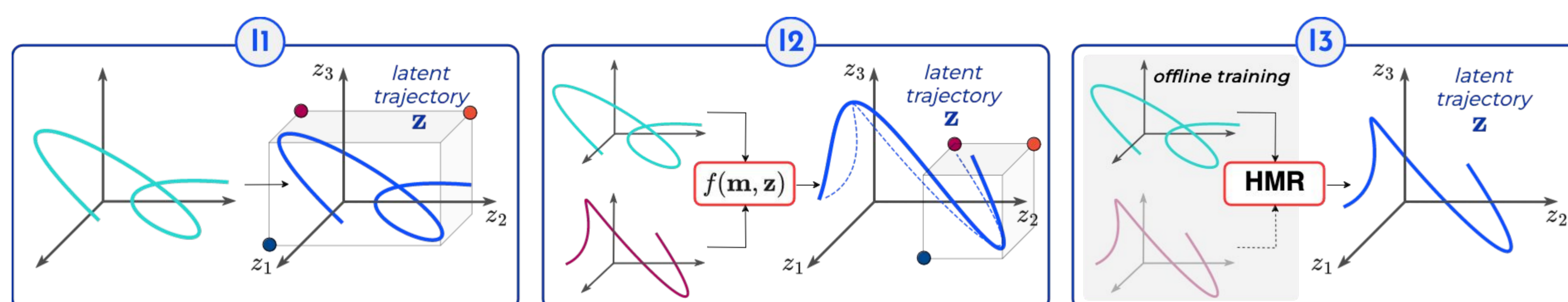
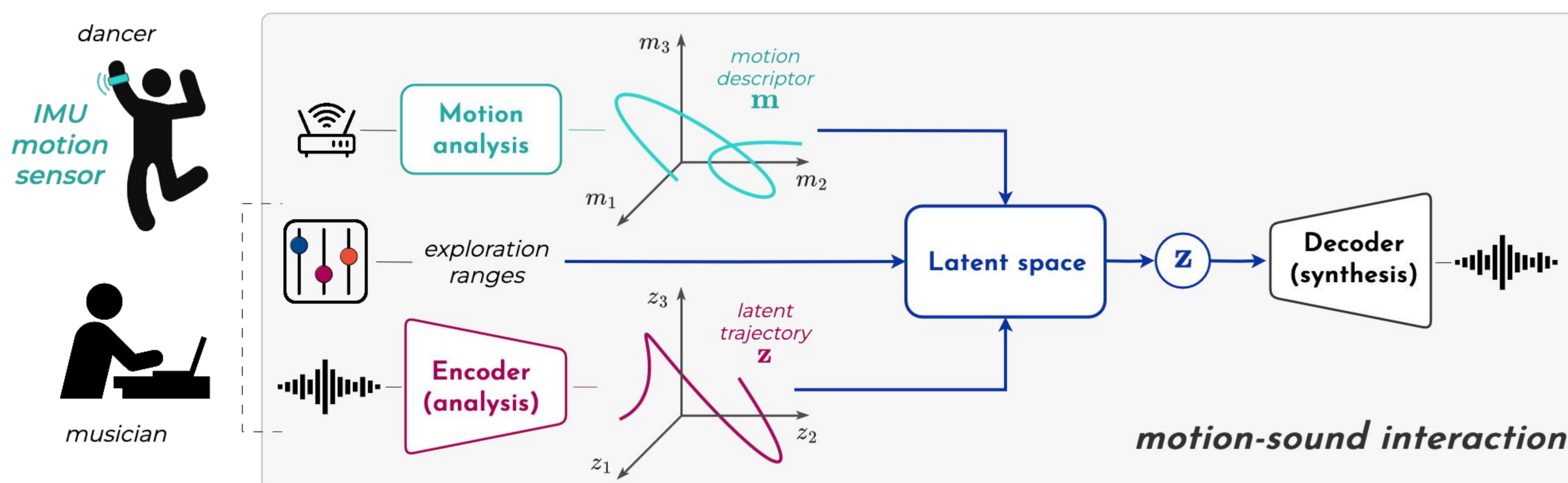
- **user-adapted control methods** able to transform latent representations of audio synthesis models based on **few examples**
- **interactive design spaces with a multimodal motion-to-sound generation approach** for more personalized human-machine interaction

Our approach



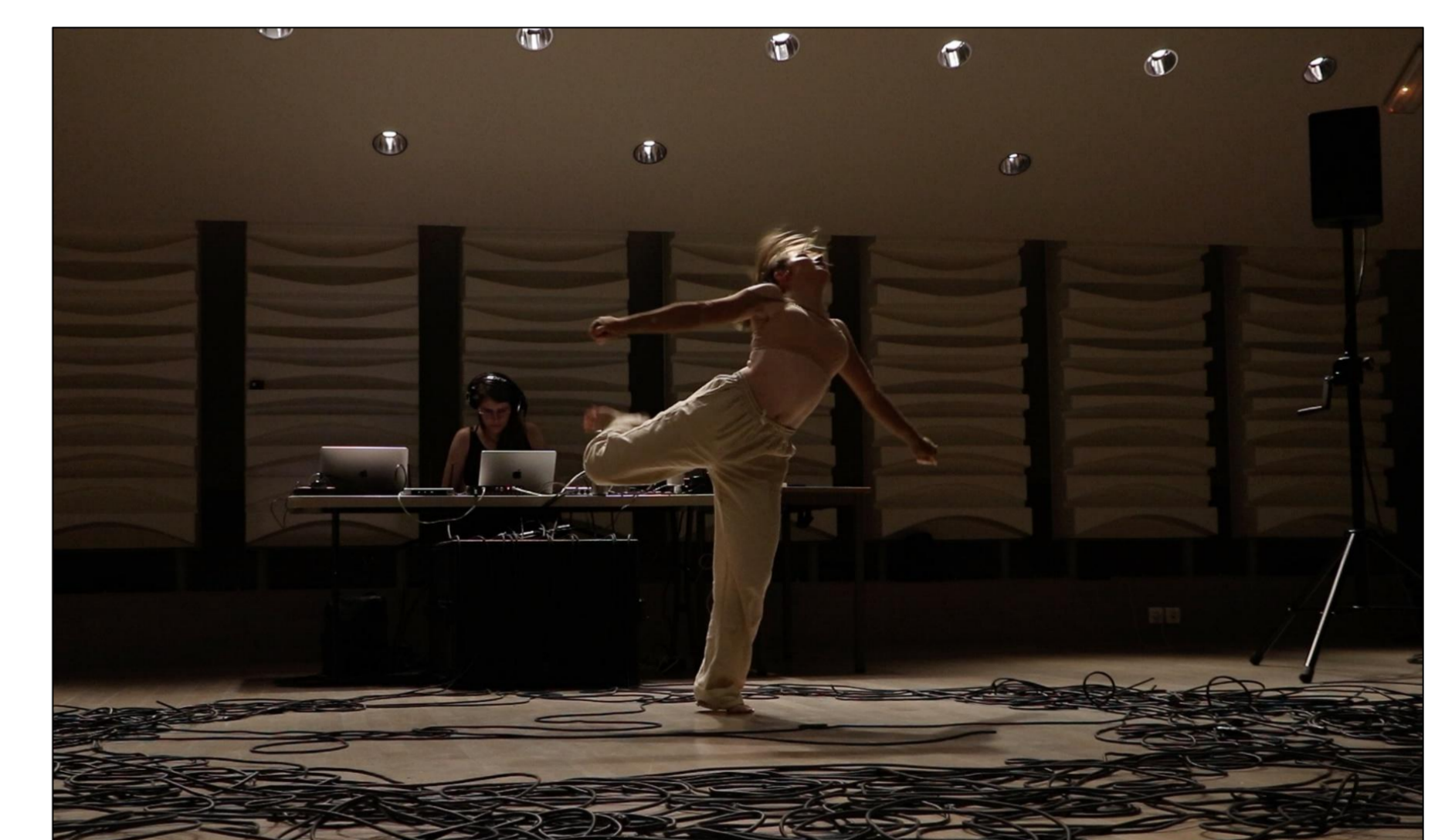
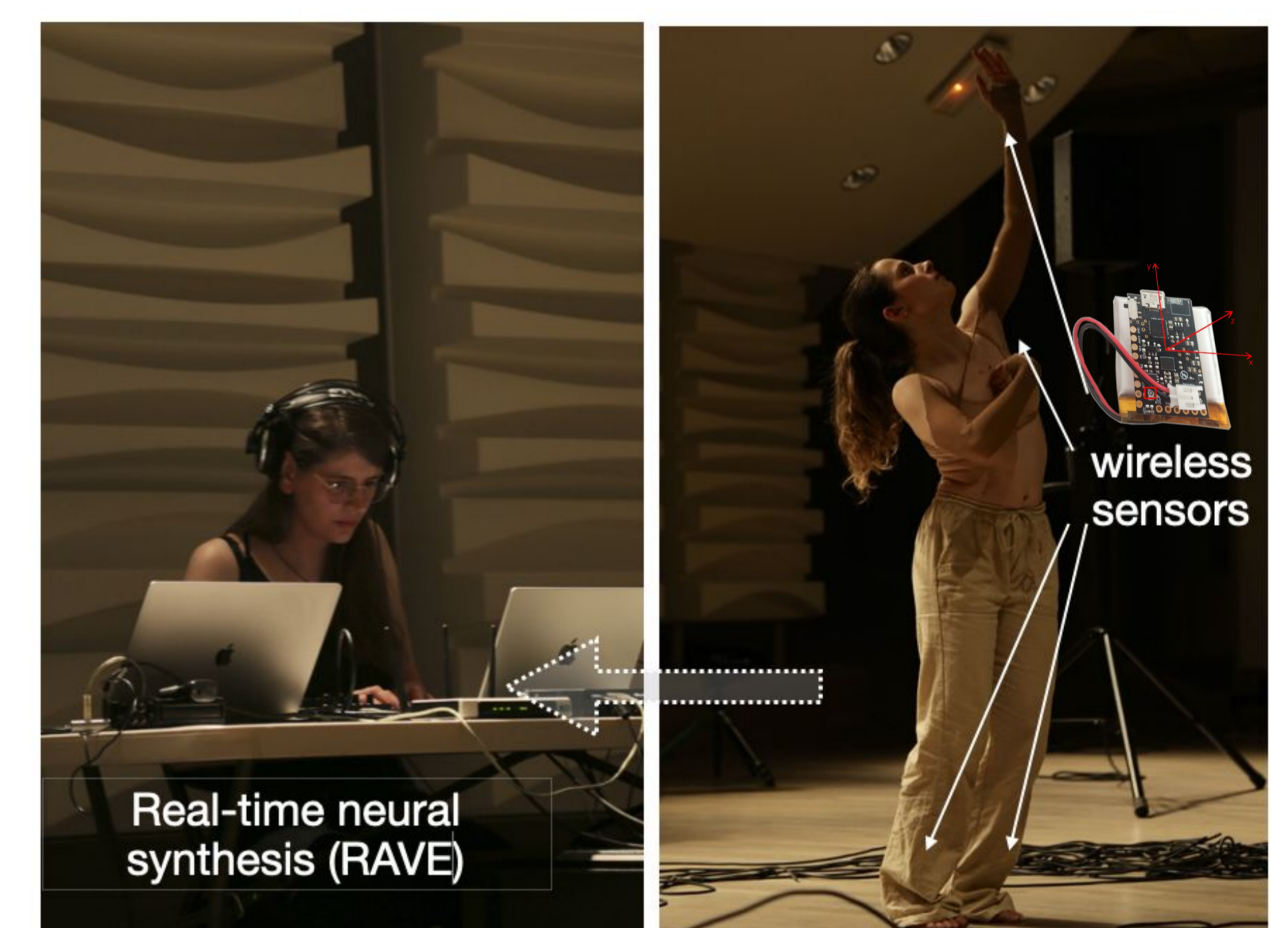
Research axis 1 - Control parameters discovery through embodied exploration of deep audio latent spaces

Proposed method



- We developed a motion-sound interactive system integrating deep audio generative model and proposed 3 embodied interaction methods to “creatively” explore latent spaces in real time.
- We established an art-research collaboration with the dancer/choreographer Marie Bruand and adopted a *performance-led research* design approach to document the creative design process and initiate discussions towards the use of AI in dance-music practices.
- We highlighted current limitations and open research questions to investigate in future works.

➔ **paper submitted to the MOCO'24 conference (under review)**



The interactive dance/music performance *Prélude* was submitted and accepted for the art festival *Nuit Blanche 2023* in Paris.



online examples linked to our paper submission to the MOCO'24 conference

