Prof. R. Wattenhofer

Scalar Quantization for Audio Compression

Current neural audio codecs rely on a modified Autoencoder architecture to compress audio waveforms into a compressed latent representation. To further compress the latent representation a technique called residual vector quantization (RVQ) is applied. RVQ is based on vector quantization (VQ) which is also found in image compression. Although VQ can theoretically achieve optimal compression for any source, it quickly becomes computationally impractical for higher-dimensional data due to the exponential growth of both the codebook size and the search complexity. Therefore, image compressors have started to shift away from VQ towards scalar quantization, a different kind of latent quantization.

In this project, we want to explore the benefits of scalar quantization within the field of audio compression. To this end, we will develop a novel neural audio codec using scalar quantization techniques instead of vector quantization, evaluating a different compression technique compared to the current prevailing paradigm.
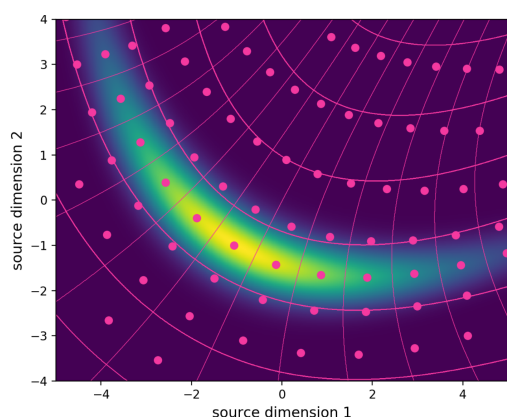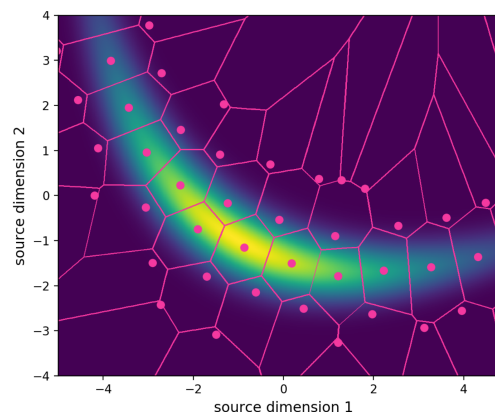

Figure 1: Nonlinear scalar quantization

Figure 2: Vector quantization

Requirements: Strong programming skills in languages such as Python. Preferably with interests in sound processing and data compression techniques.

Weekly meetings will be scheduled to address questions, discuss progress, and brainstorm future ideas.

Contact

In a few short sentences, please describe your interest in this project and any relevant coding experience or background (e.g., projects or coursework).

- Till Aczel: `taczel@ethz.ch`, ETZ G60.1

- Luca Lanzendörfer: `lanzendoerfer@ethz.ch`, ETZ G93