Appendix

January 8, 2026

Giada Manfredi

4. Experimental results

4.1. Additional experimental results

This experiment was conducted on the audio sample generated by MusicGen using the prompt: "*Pop dance track with catchy melodies, tropical percussion, and upbeat rhythms, perfect for the beach.*"

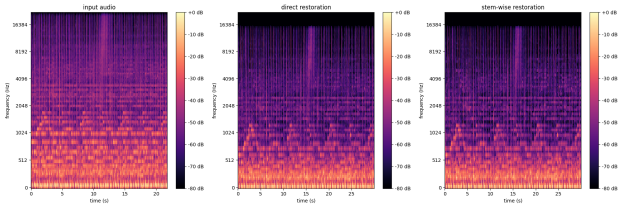

Figure 1.

Qualitative inspection of spectrograms in (Figure 1) shows that low-intensity frequency components are noticeably attenuated after restoration, while high-energy musical components remain largely unaffected. Differences between direct restoration and hybrid restoration are subtle in full-band spectrograms but become evident under quantitative analysis.

To quantify noise reduction, we analyze the 25th percentile of spectral amplitude across frequency bins, which emphasizes low-energy components typically associated with noise and artifacts.

	Input audio	Direct restoration	Stem-wise restoration
25th percentile value	-57.32 dB	-68.77 dB	**-70.45 dB**

These results indicate that applying DSP independently to separated sources enables more effective suppression of low-level noise without degrading musical content.

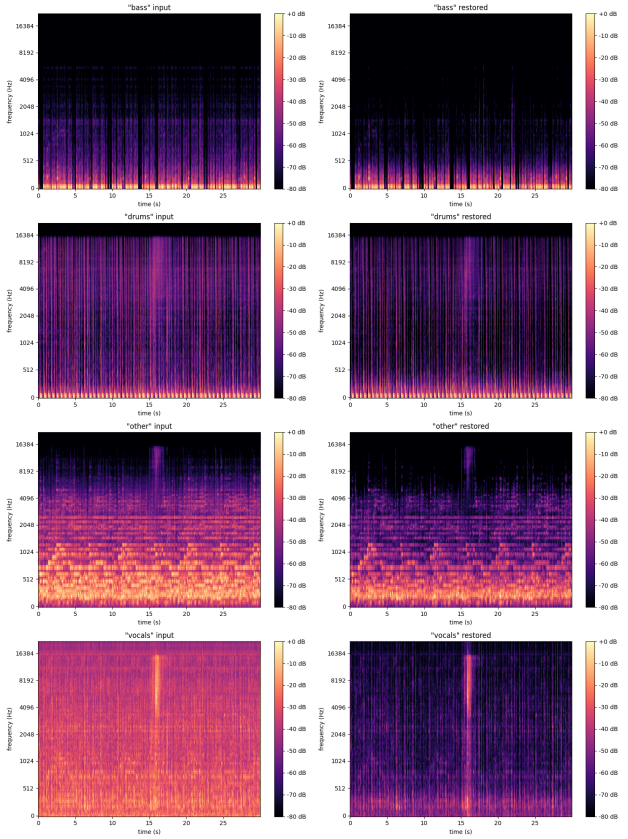

Figure 2. The four stems obtained with Demucs, shown before and after restoration. The most substantial change occurs in the "vocals" stem, as the analyzed audio contained no vocal track leading Demucs to allocate the majority of the noise to this stem.

Email: Giada Manfredi <manfredi.2139294@studenti.uniroma1.it>.

Machine Learning 2025, Sapienza University of Rome, 2nd semester a.y. 2024/2025.

4.2. Regarding the audio analyzed in the report

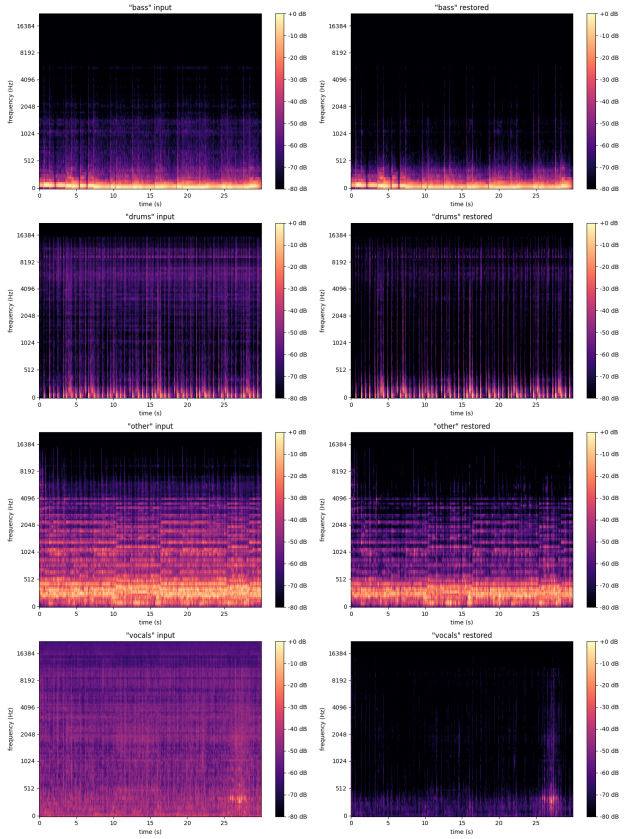

Figure 3. The four stems obtained with Demucs, shown before and after restoration. In this case as well, the most substantial change occurs in the "vocals" stem, as the analyzed audio contained no vocal track leading Demucs to allocate the majority of the noise to this stem.