MDX
Music Demixing Workshop

Source Separation using a Score-based Generative Model

Simon Rouard[*1, 2] **and Gaëtan Hadjeres**[†1]

1 Sony CSL **2** CentraleSupélec and ENS Paris-Saclay

Abstract

Diffusion models (or score-based generative models) recently achieved impressive results on audio and speech generation tasks in the waveform domain, such as vocoding, super-resolution or unconditional generation of instrument sounds. In this work, we investigate the use of a general conditional diffusion models to the Music Source Separation task, where our approach simply consists in generating the desired stem conditioned on the input mixture. Compared to other approaches that generate in the waveform domain, our approach does not rely on the design of a specific loss function, is stochastic by nature and allows to trade off speed vs accuracy. We illustrate our approach on the standard MusDB dataset with promising results.

---

[*]co-first author
[†]co-first author