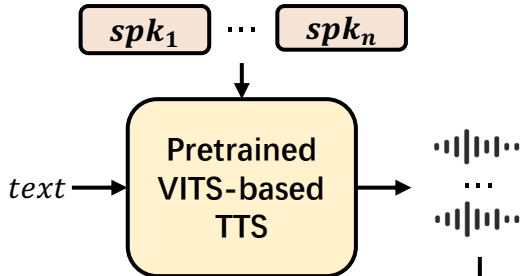
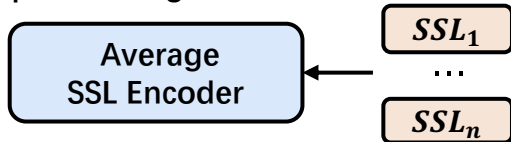


Step1: Get parallel speech



Step2: Training AVENet



Waveform



SSL Encoder

Average
SSL Encoder

Step3: Training
and Inference VC

VITS-based
VC