**CLAP Score Distributions
(Higher is Better)**

Left panel legend: Fine-tuned VAE Captions, MusicCaps Captions, LP-MusicCaps Captions

Right panel legend: Fine-tuned VAE Captions, Zero-shot Captions, Base VAE Captions

Y-axis: Density

X-axis: CLAP Score