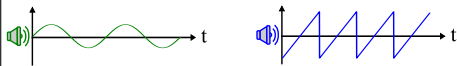
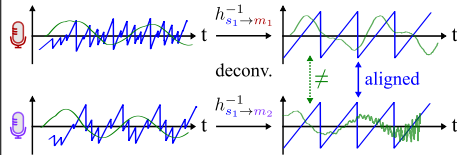


(a) room

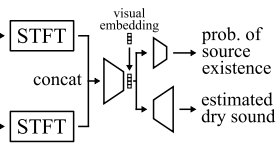


(b) dry sound from each source



(c) received audios

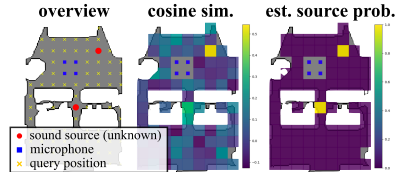
(d) deconvolved



(e) network

(f) average cosine similarity on 1000 scenes

	query has source \uparrow	query has no source \downarrow
received audio from different mics	0.02	0.02
delayed based on distance	0.22	0.06
deconvolved with RIRs	0.41	0.05



(g) example scene