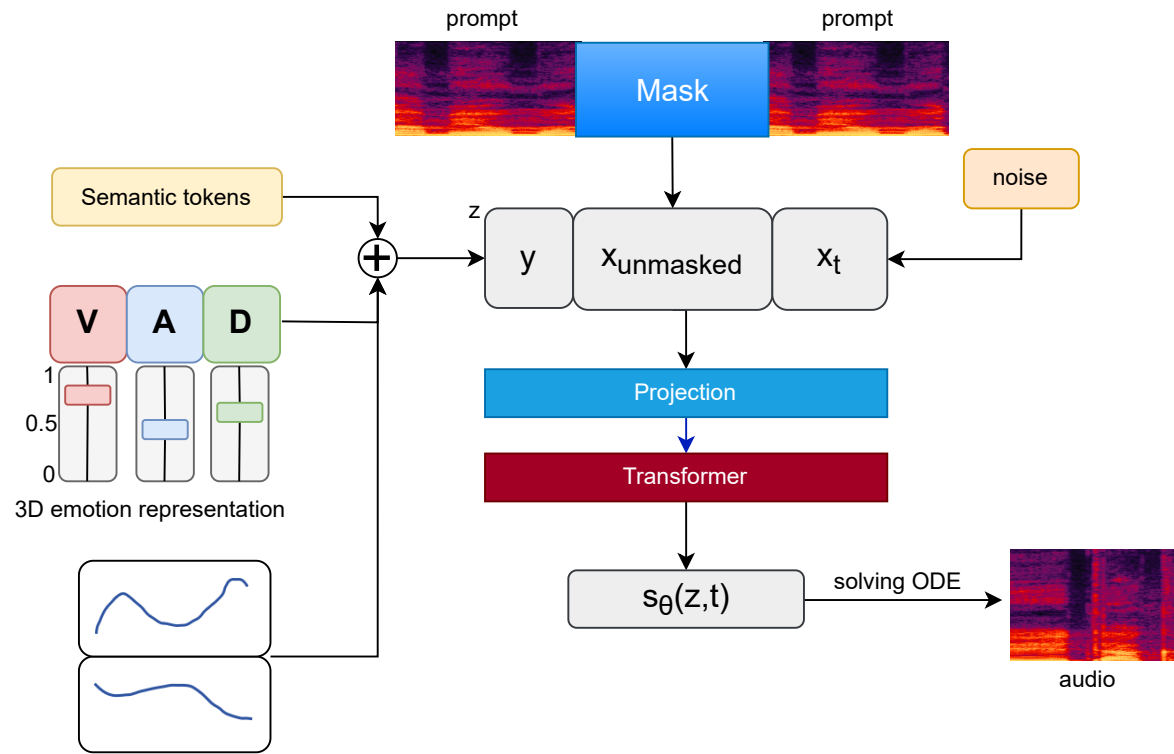


(a) Stage 1 Flexible control of expressiveness via keyframes



(b) Stage 2 Prosody-guided zero-shot synthesis