Noise-Robust Speaker&Emotion Separator ADV. Emotion (b) Orthogonal Constraint Classifier **Emotion** Denoiser \mathcal{H}_s Speaker Speaker Encoder **Emotion** Classifier Teacher Clean Reference **Emotion** \mathcal{H}_e^{cln} Extractor Emotion **Emotion Embedder** Classifier Speaker Extractor Stop (a) Knowledge Distillation Gradient Gradient Reversal **Emotion** Conditional **Emotion Noisy** Reference VQ Diffusion Encoder Classifier Loss --> Function \mathcal{H}_e^{den} \mathcal{H}_e $\mathcal{H}_e \oplus \mathcal{H}_s$ Phoneme Transformer **Content Adaptor** Mel Decoder Flow Post-Net Vocoder Sequence Encoder

Predicted Mel-spectrum

Output Expressive Speech