

Learning Interpretable Representation for Controllable Polyphonic Music Generation

ISMIR 2020 Poster Presentation

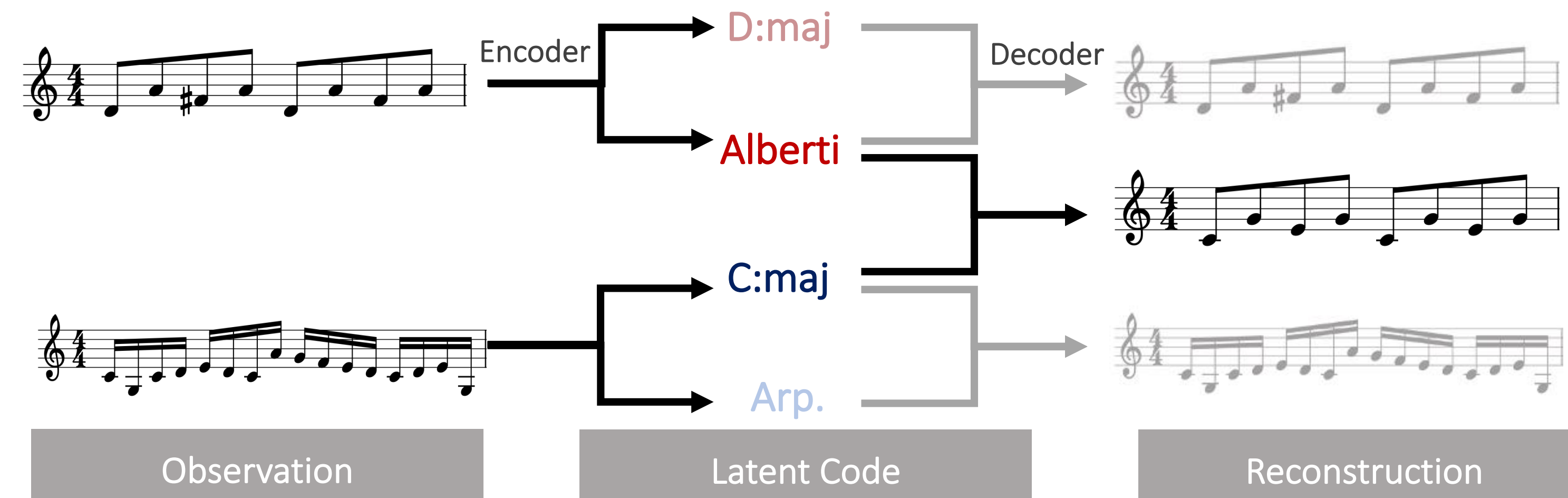
Ziyu Wang, Dingsu Wang, Yixiao Zhang, Gus Xia

Music X Lab, Computer Science Department, NYU Shanghai

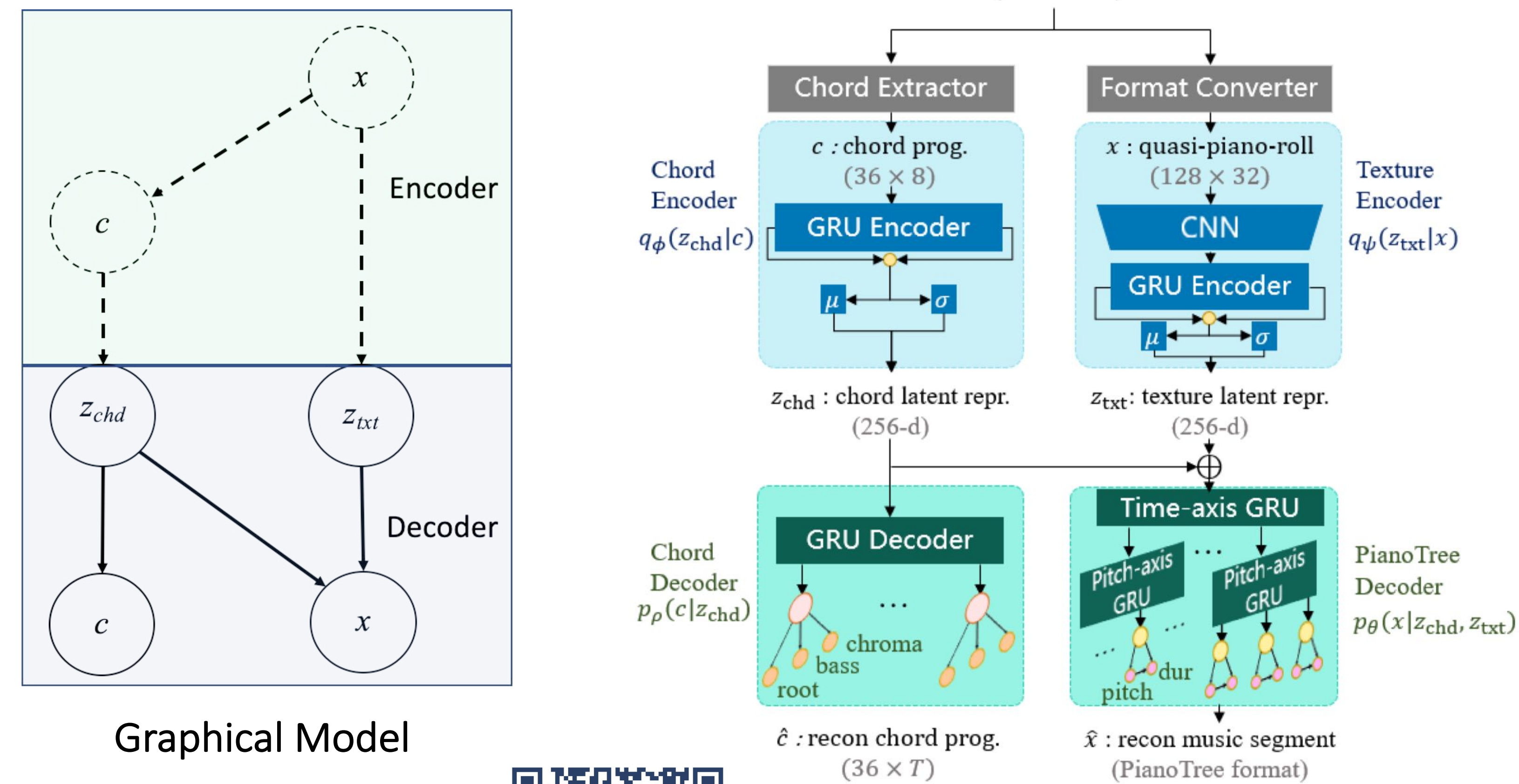
{ziyu.wang, dingsu.wang, yixiao.zhang, gxia}@nyu.edu



Motivation

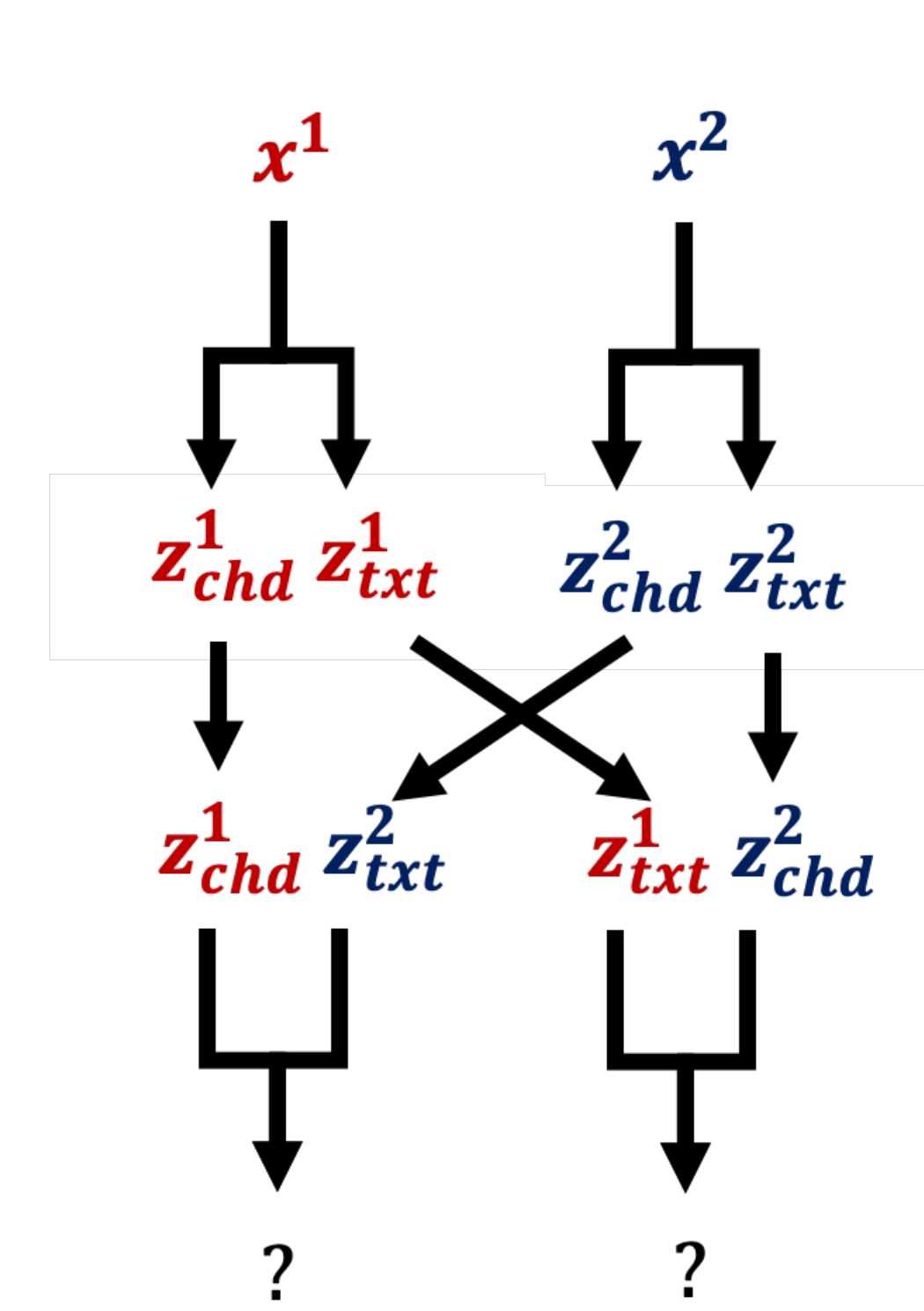


Model Architecture



Controlled Music Generation by Disentanglement

Task 1: Compositional Style Transfer



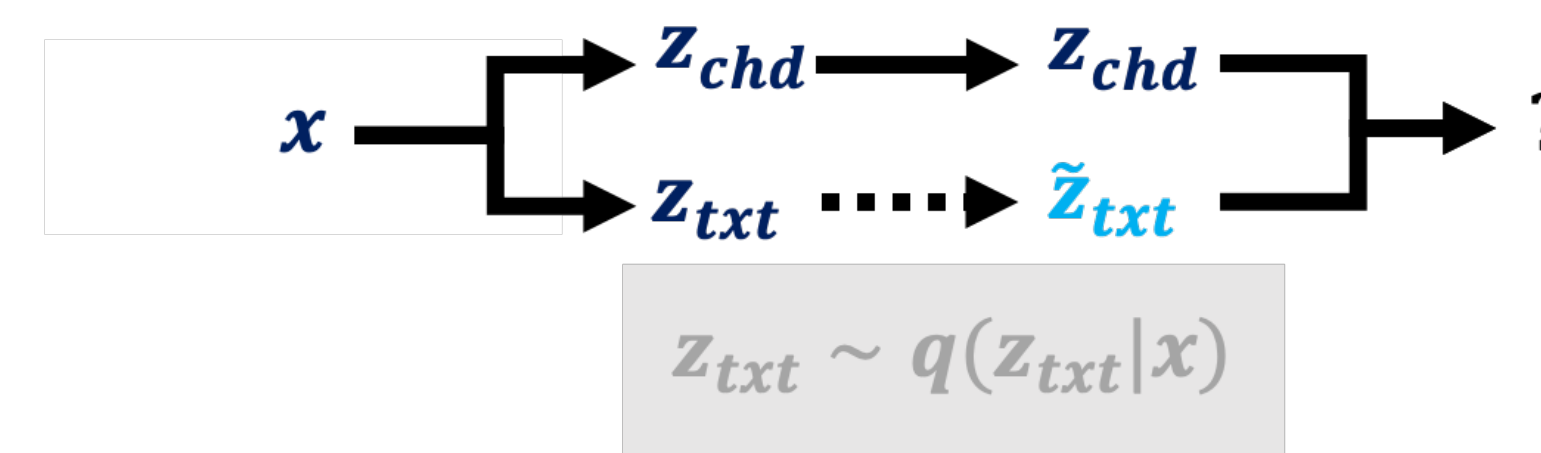
Original 1
 z_{chd}^1, z_{txt}^1
Right-hand: Syncopated alto melody throughout
Left hand: 8th note arpeggio
Two-voice melody
characteristic cut-off
Rising contour

Original 2
 z_{chd}^2, z_{txt}^2
Right-hand: lyrical soprano melody throughout
Left hand: 8th note "inverse Alberti"
Long bass note throughout
Change to 16th note arpeggio

Generation 1
 z_{chd}^2, z_{txt}^1
New melody still preserving syncopated style
Preserving two-voice melody
Preserving arpeggio contour
Preserving the cut-off
Preserving pitch contour

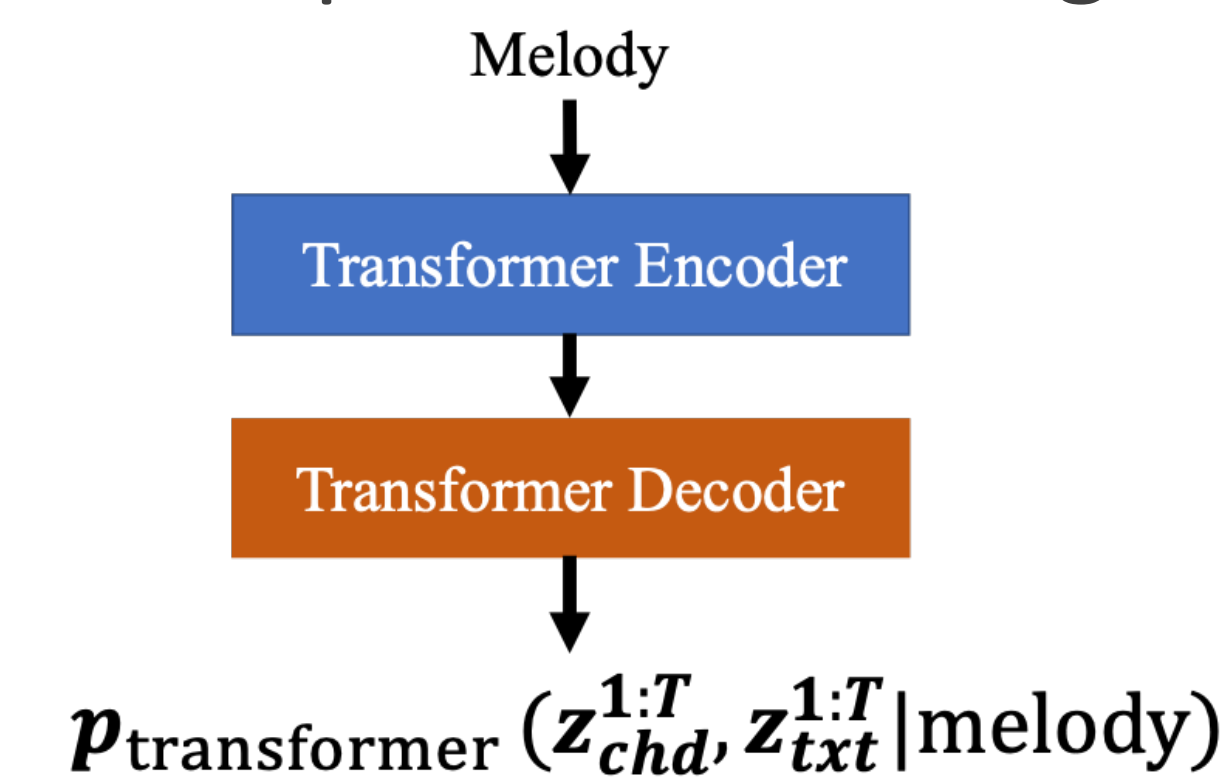
Generation 2
 z_{chd}^1, z_{txt}^2
New melody still preserving lyrical feel
Preserving "inverse Alberti" contour
Change to 16th note arpeggio

Task 2: Texture Variation via Sampling



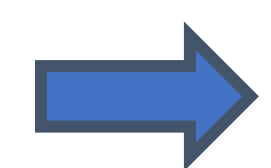
Improvising melody
Preserving the cut-off
Minor rhythmic variations

Task 3: Accompaniment Arrangement



Mel. Condition
Gm Cm Fm Bb Eb Bb Cm Bb Ab Bb Eb
Mimicking rhythmic pattern
Secondary Melody as a transition between two phrases
Acc. Condition
Same bass pattern
Entering new section

Source Code & MIDI Demo



Model Diagram