

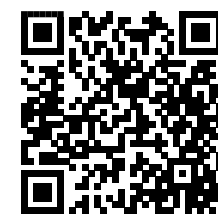
Composer Vector: Style-steering Symbolic Music Generation in a Latent Space

Xunyi Jiang, Julian McAuley, Xin Xu*

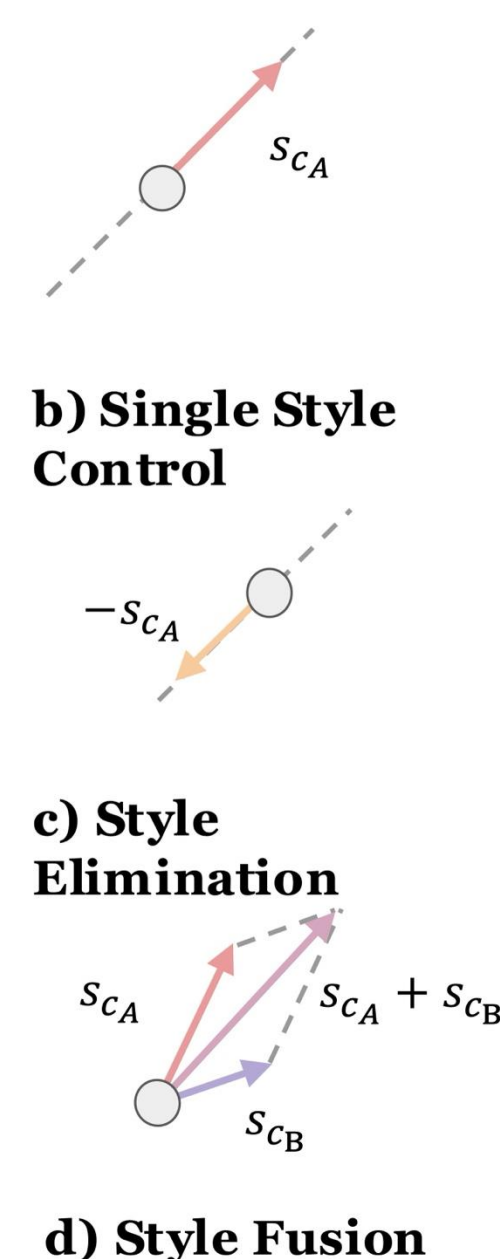
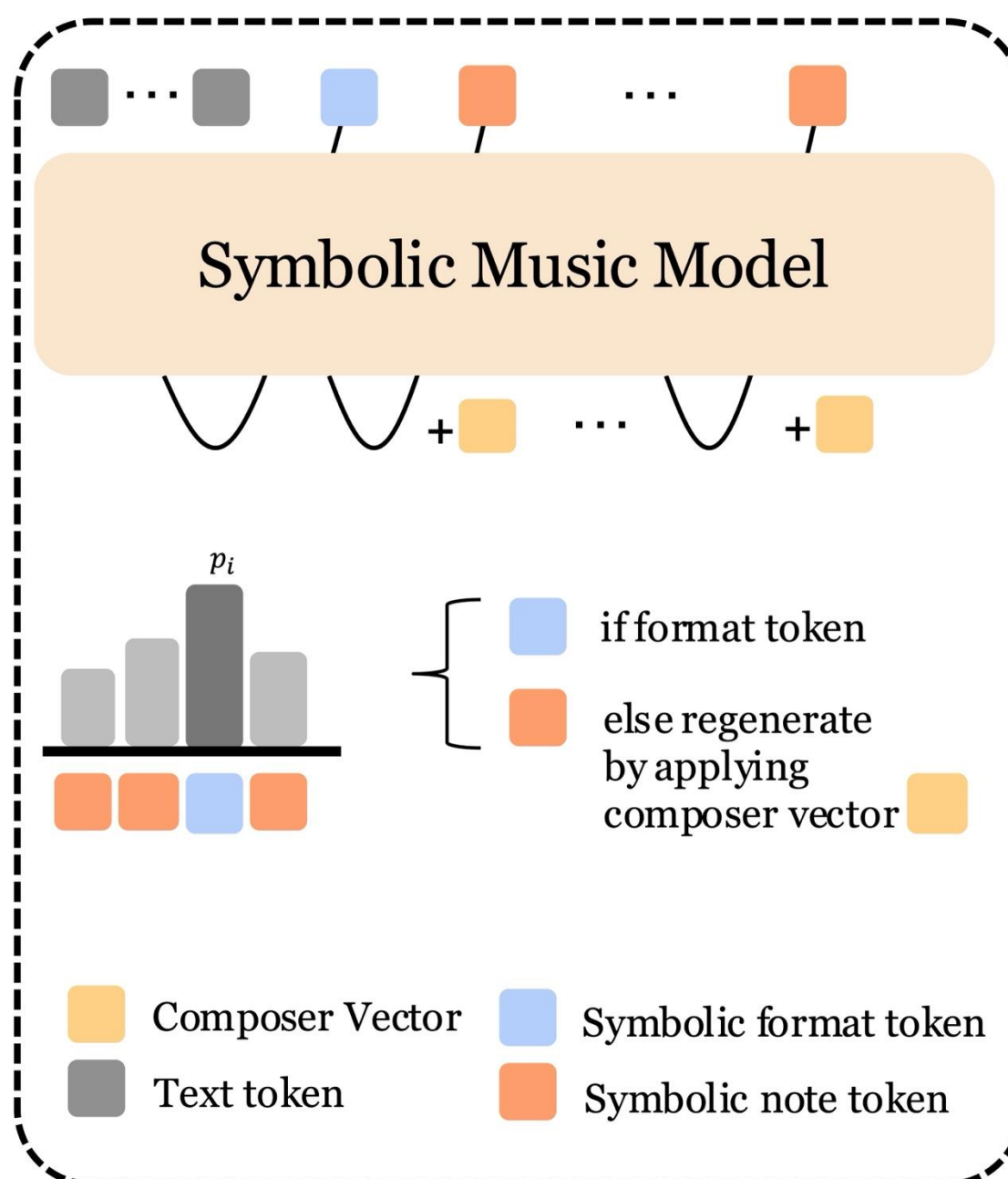
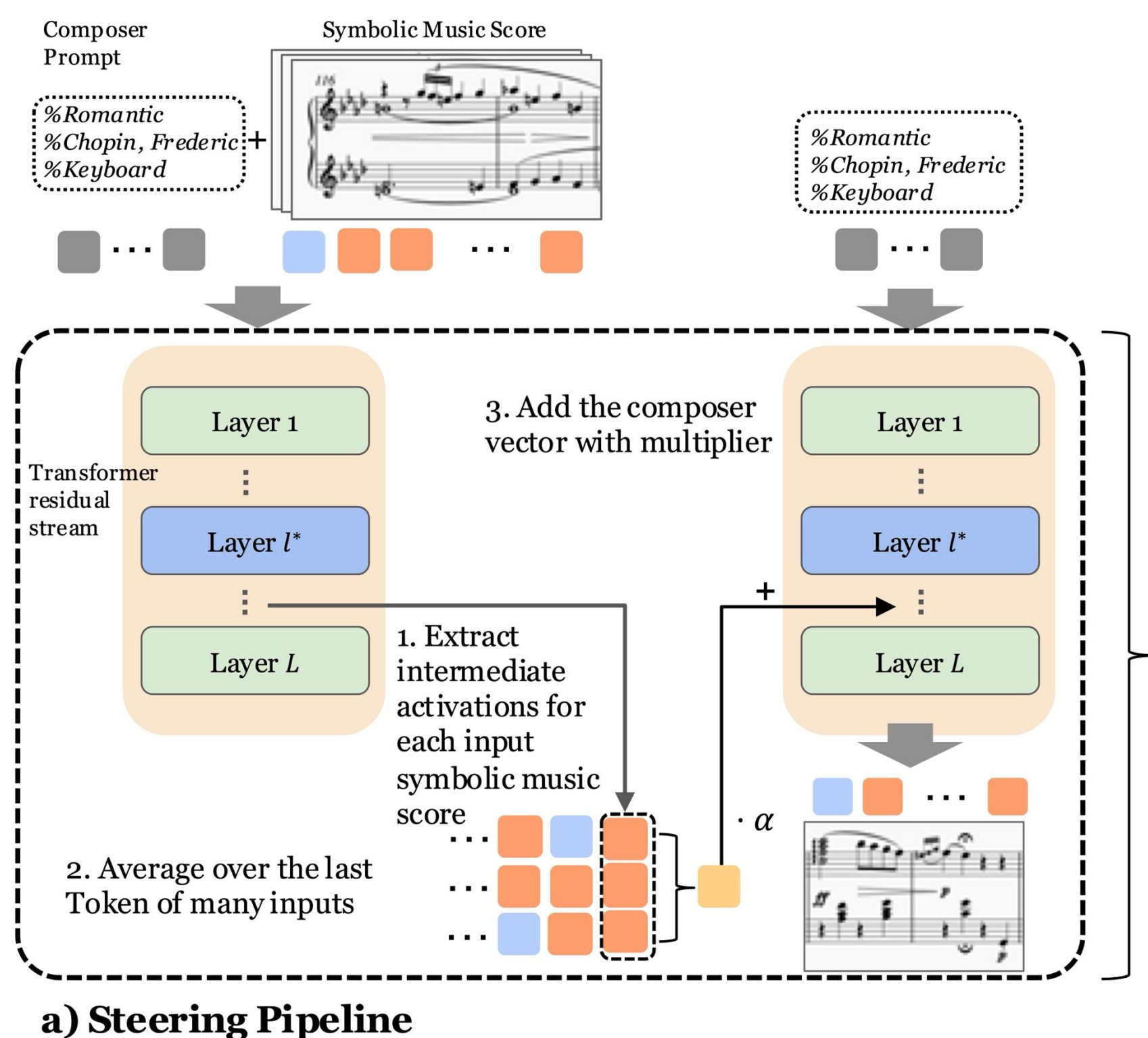
University of California, San Diego



Demo Page



GitHub



Overview

- **Problem:** Symbolic music models sound convincing, but precise *composer-style* control and style blends usually needs extra training and labels.
- **Idea: Composer Vector.** Compute a style direction for a composer from hidden states and **add it at inference** to steer style. **No fine-tuning!**
- **Why it works:** Composer identity is separable in deep layers; adding the vector with a coefficient α smoothly shifts style while preserving format.
- **Results:** Across Notagen and ChatMusician, steering **raises CLAP/CLaMP similarity** and **composer classifier's probability** for the target style.

Vectors can be **linearly fused** for blended styles, and α gives continuous control.

Composer Vector

Composer Vector construction.

Extract hidden state from piece-prompt corpus $\mathcal{D} = \{x_i^c \oplus p_i^c\}$, where x_i^c is a textual prompt of composer and p_i^c is a symbolic music score in ABC format.

Then compute piece-level hidden states at the best layer l^* and average them to get the composer style vector $s_c = \frac{1}{N_c} \sum_{i=1}^{N_c} h_T^{(l^*)}(x_i^c \oplus p_i^c) \in R^d$, where T is the length of sequence. We define $s_c = \sum_i w_i s_{c_i}$ as style fusion.

Inference-time steering.

Inject at layer l^* in the residual stream during decoding:

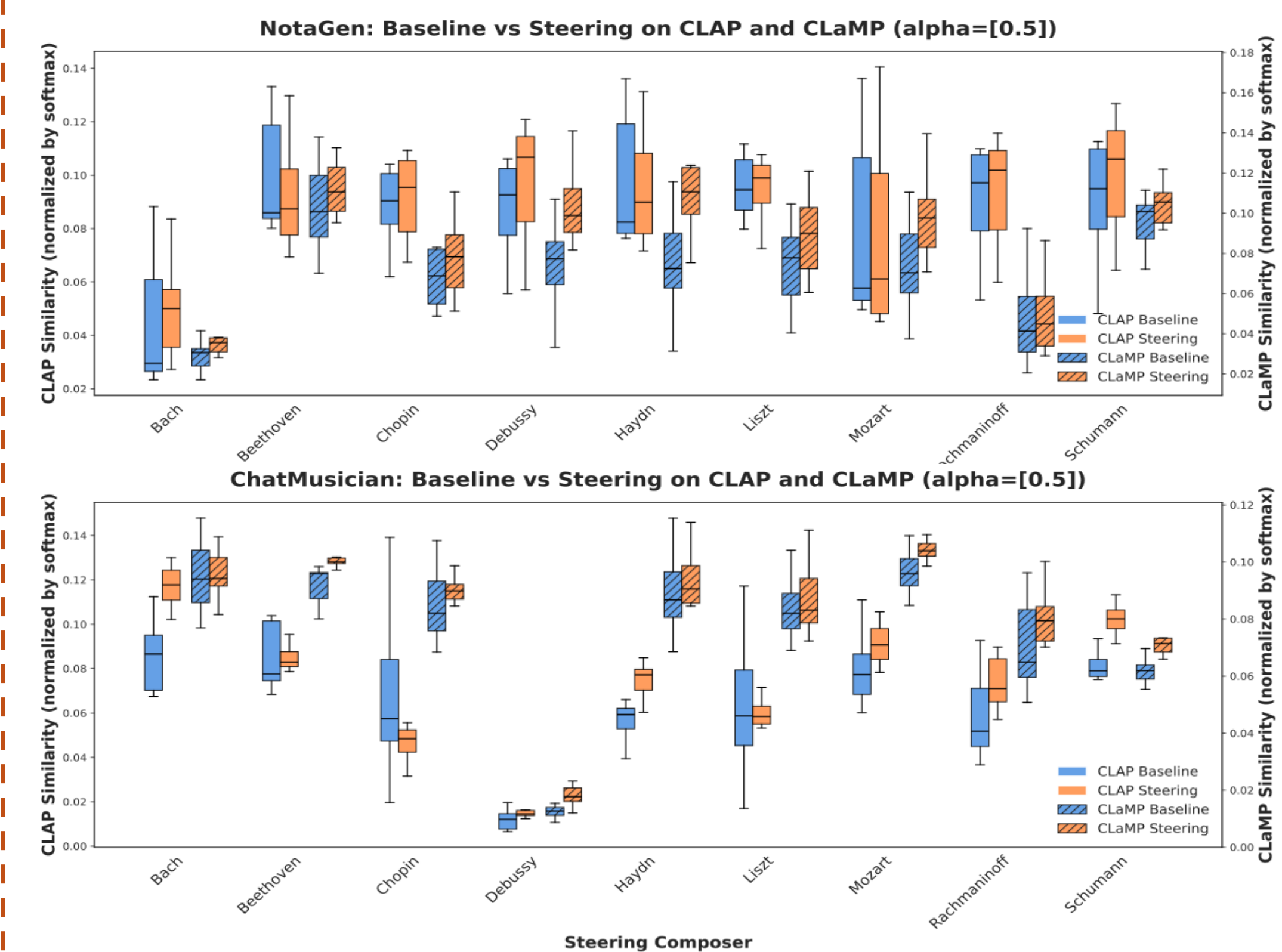
$$\widehat{h_t^{(l^*)}}(x, y_{<t}) = h_t^{(l^*)}(x, y_{<t}) + \alpha s_c,$$

$$\widehat{h_t^{(l^*)}}(x, y_{<t}) \leftarrow \frac{|h_t^{(l^*)}(x, y_{<t})|_2}{|\widehat{h_t^{(l^*)}}(x, y_{<t})|_2} \widehat{h_t^{(l^*)}}(x, y_{<t})$$

Format preservation.

Only steer music content tokens (notes, durations, dynamics); skip format tokens (baselines, control symbols) to keep scores valid.

Controlling Symbolic Music Generation with Composer Vector



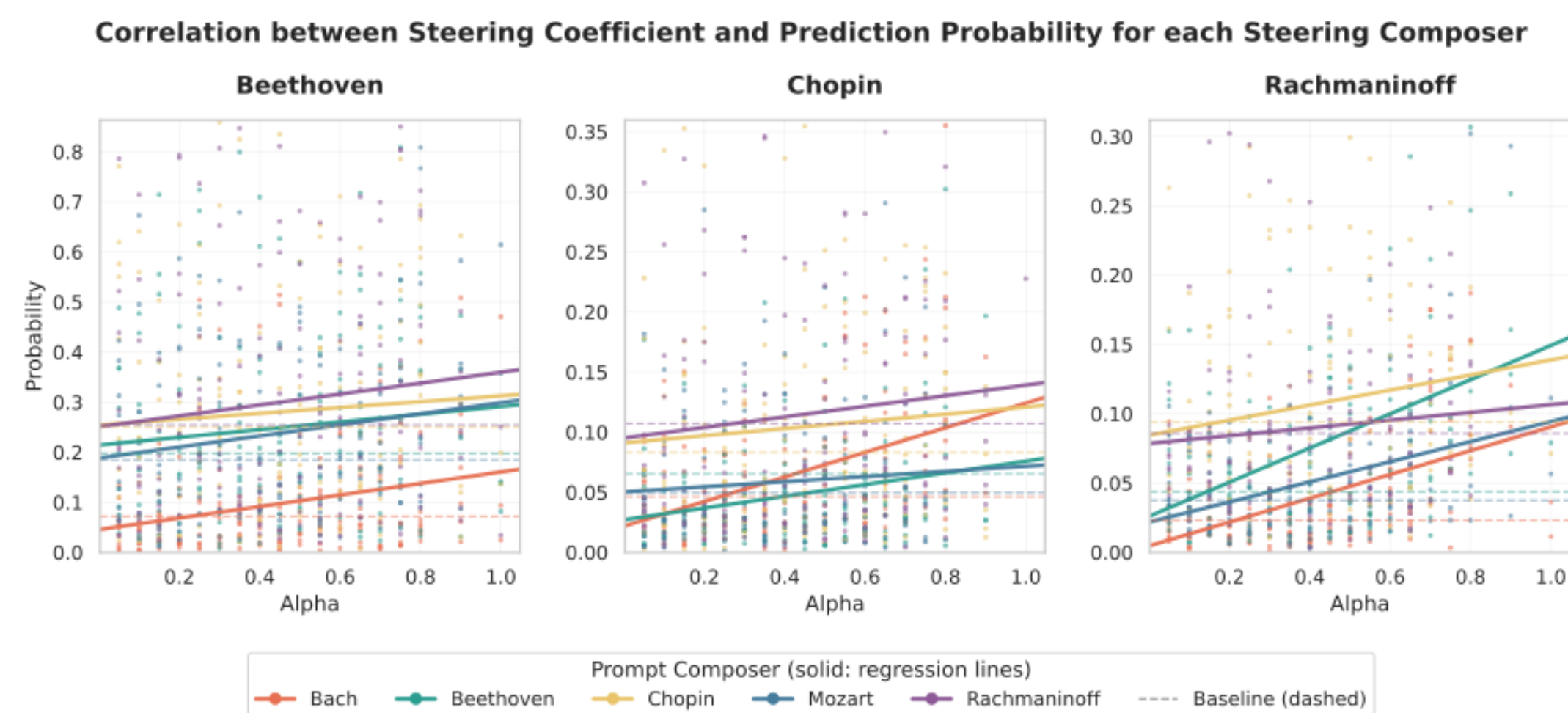
- **CLAP and CLaMP3 similarity increase after steering.**

- **Composer classifier prediction probability increases after steering.**

Method	Bach Mozart	Haydn Beethoven	Haydn Liszt	Beethoven Mozart	Beethoven Bach	Schubert Ravel	Schumann Bach	Chopin Bach	Debussy Beethoven
Notagen	5.9	26.3	18.7	4.9	11.7	6.6	11.2	14.9	33.7
+vector	14.0	34.2	26.2	16.0	17.9	14.4	29.2	22.5	52.2
ChatMusician	3.4	4.5	5.0	6.0	25.9	1.2	29.5	23.0	1.4
+vector	16.4	30.1	15.1	13.2	54.3	14.6	57.0	53.4	20.8

Composer Vector can Provide Continuous Style Control

Target composer prediction probability is positive correlated with steering coefficient, providing continuous control in composer style generation.



Multi-style Fusion through Composer Vector

One composer's probability consistently increases while the other decreases, demonstrating that linear composition of composer vectors enables continuous, interpretable style interpolation.

