

# Supplementary Material accompanying TISMIR submitted paper titled “Identifying Melodic Motifs and Stable Notes from Gestural Information in Indian Vocal Performances”

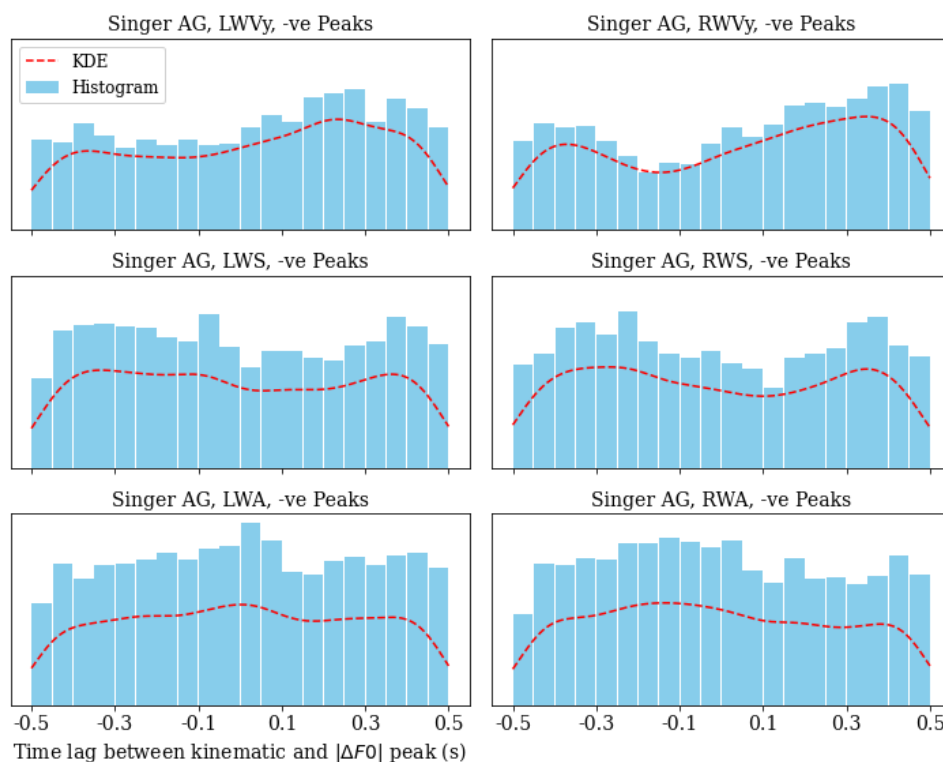
This document contains extra figures and tables organised by the section number they correspond to in the main paper.

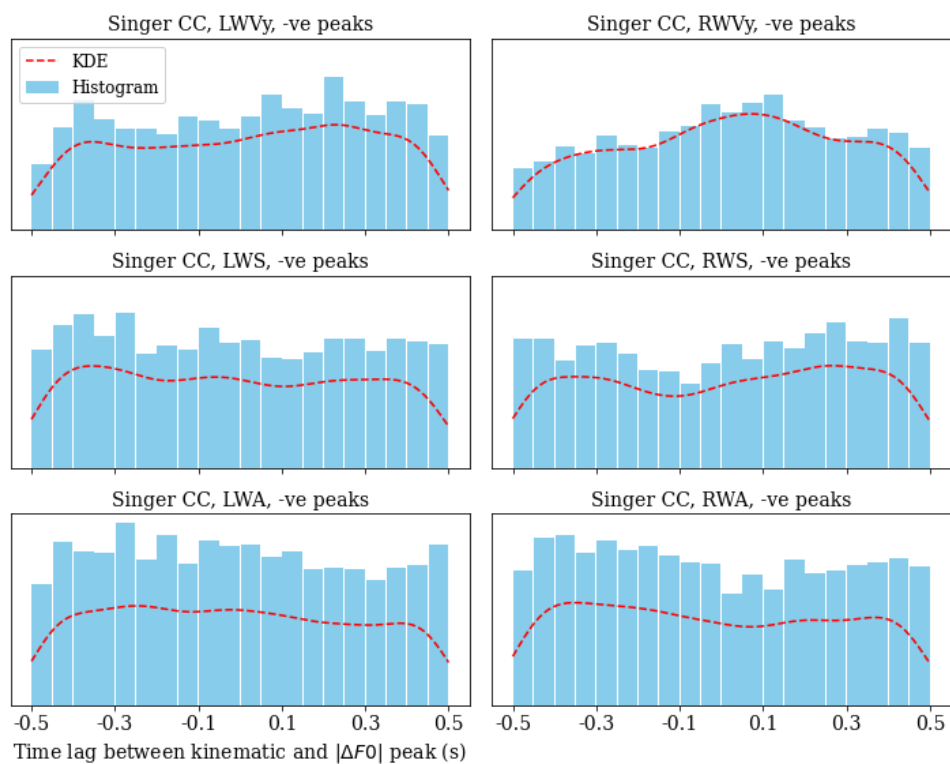
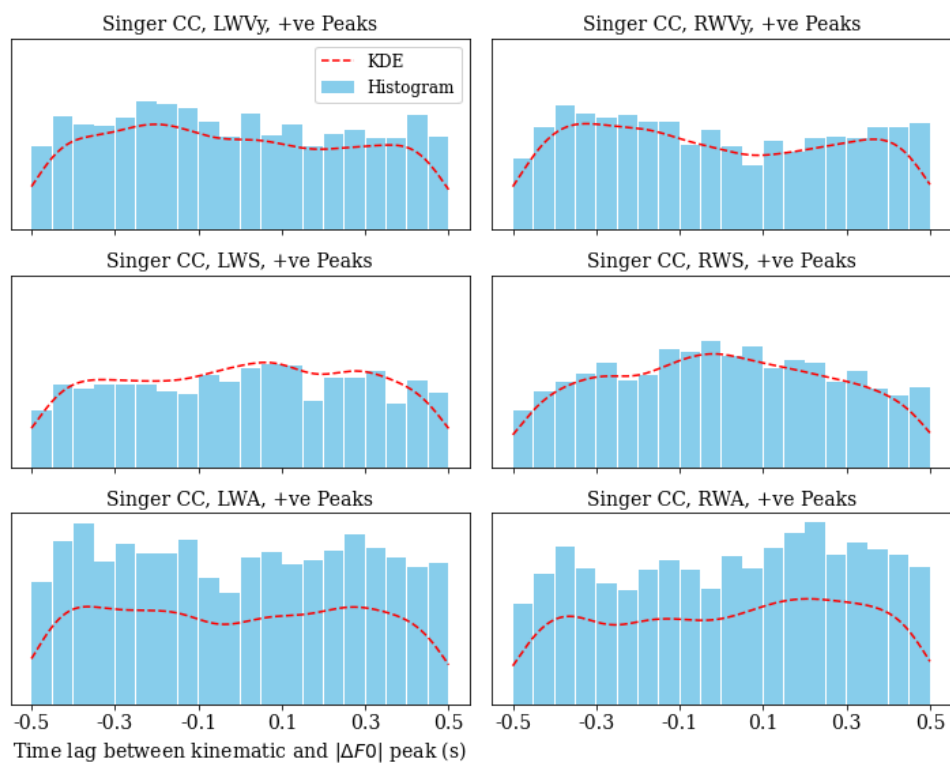
# Section 4S: Relating F0 with Kinematic Changes

## 4S.2 Temporal Coupling Study

The following figures present distributions for the time offset values (in seconds) observed between kinematic peaks and corresponding  $|\Delta F0|$  peaks, within the 1s windows centered at kinematic peaks. The plots are provided for singers AG and CC since they show the highest values for magnitude coupling in the main paper's Table 2.

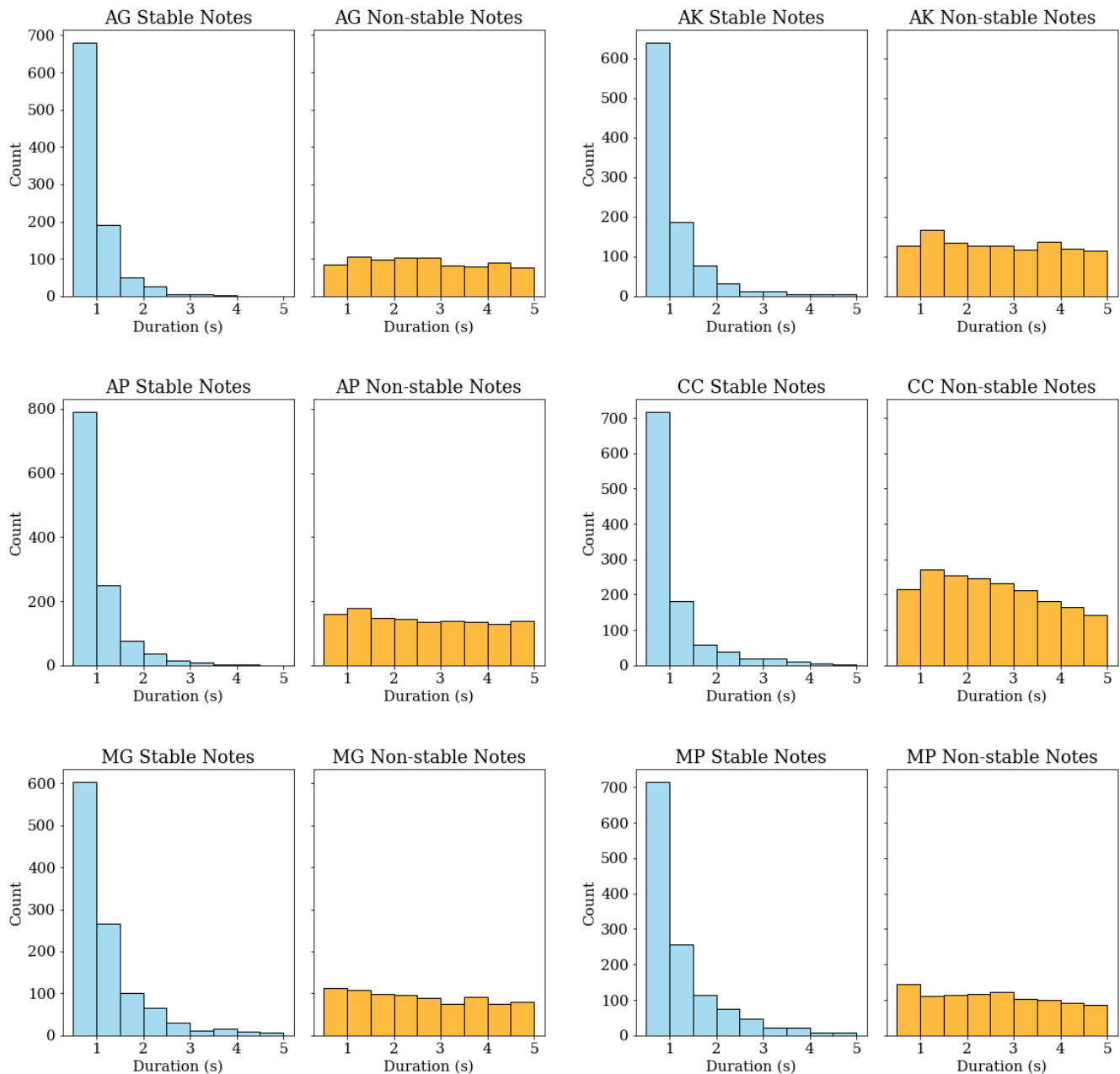
Negative time lags indicate the  $|\Delta F0|$  peak leading the kinematic peak, while positive lags indicate the  $|\Delta F0|$  peak lagging the kinematic peak. Positive and negative kinematic peaks are studied separately for each of height (Vy), and acceleration; peaks and valleys are separated for speed parameter. We observe a nearly uniform distribution in every case indicating the absence of a consistent lag or lead.

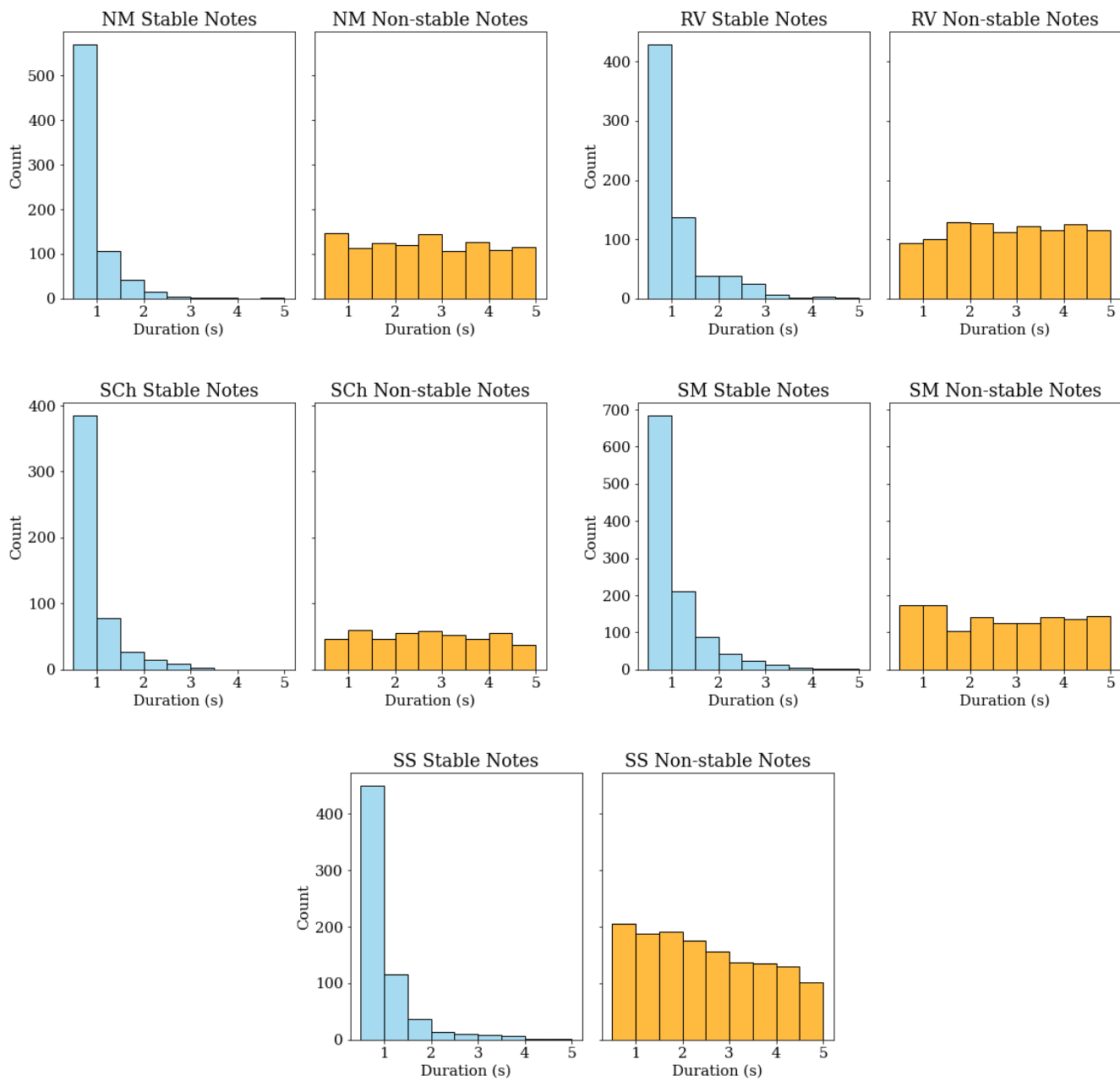




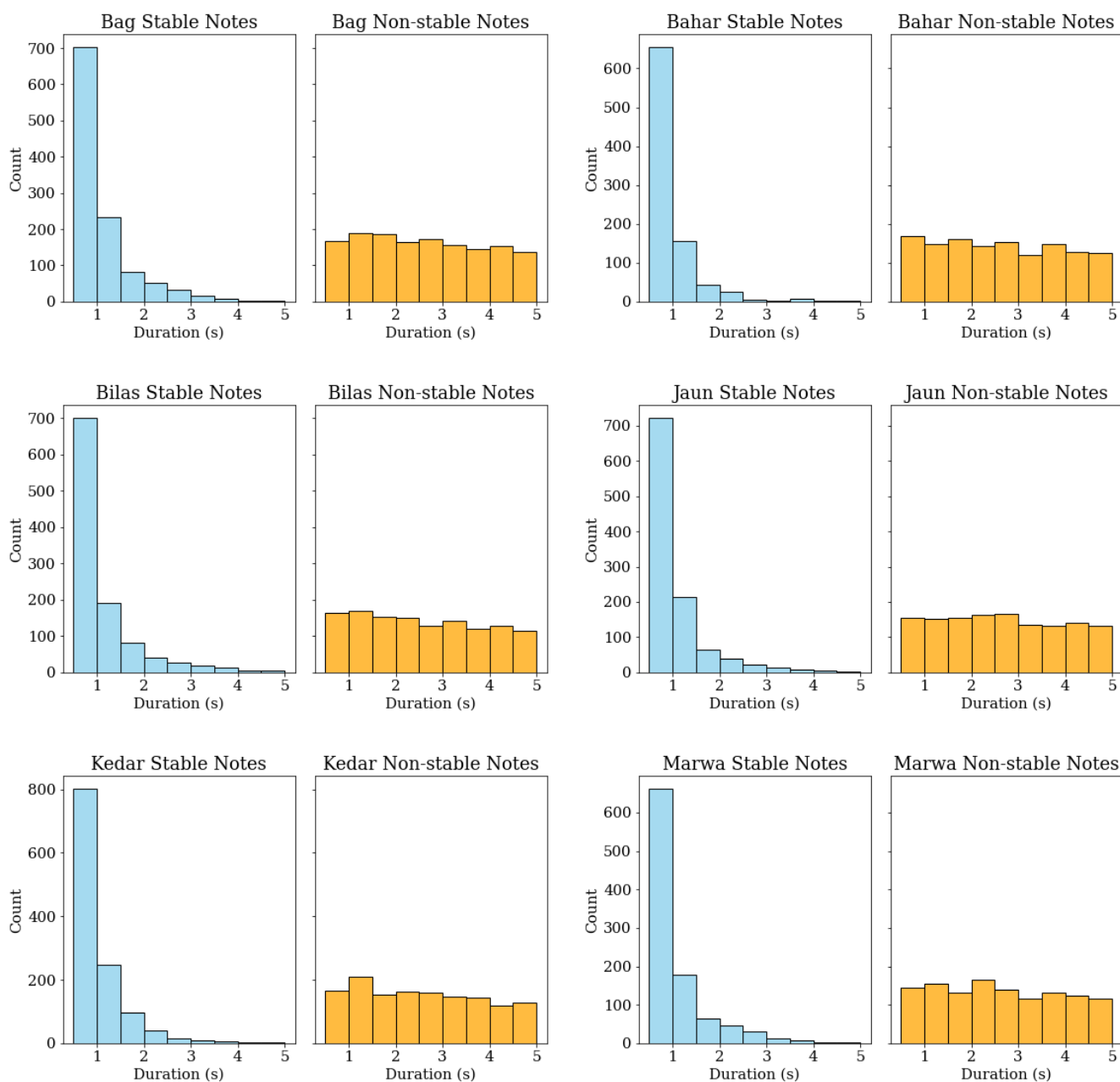
## Section 5S: F0 segmentation methods

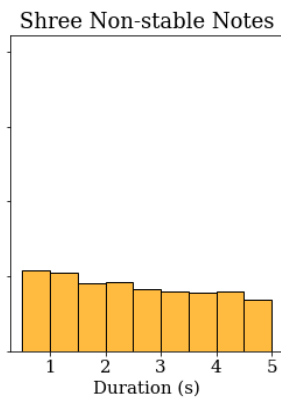
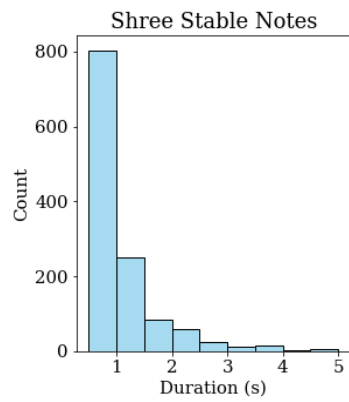
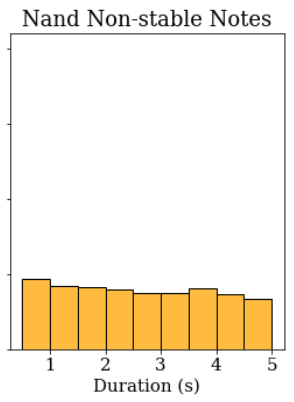
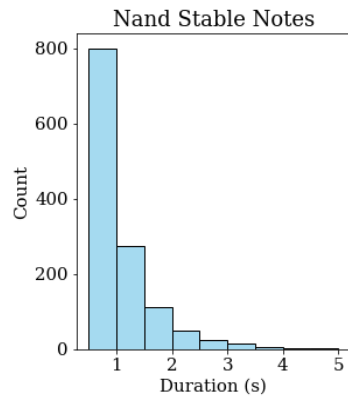
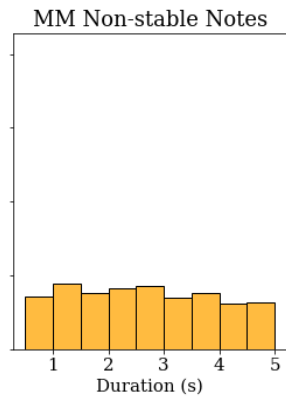
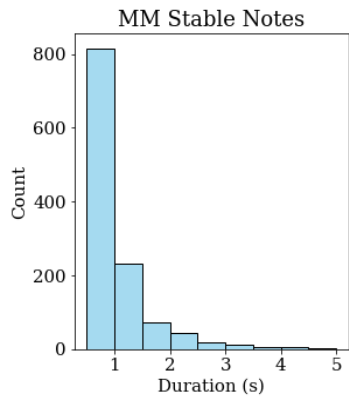
5S.1: Singerwise distribution of stable note and non-stable segment durations across ragas. All singers follow the trend of a distribution that is skewed towards lower durations for stable note segments.





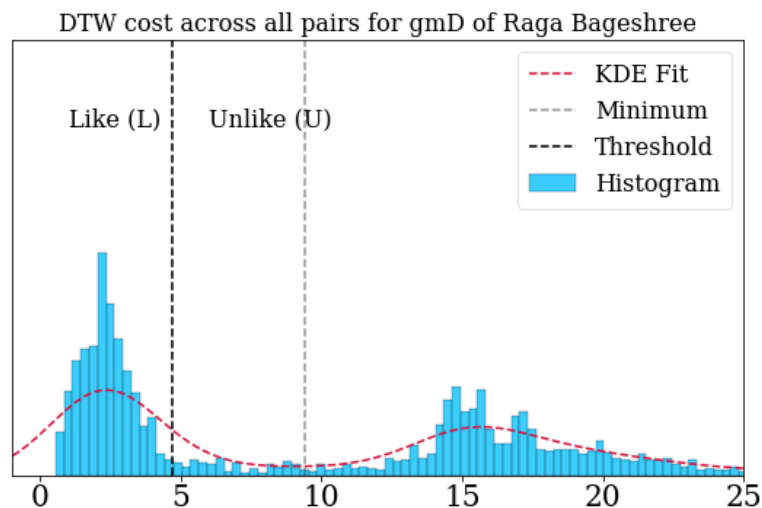
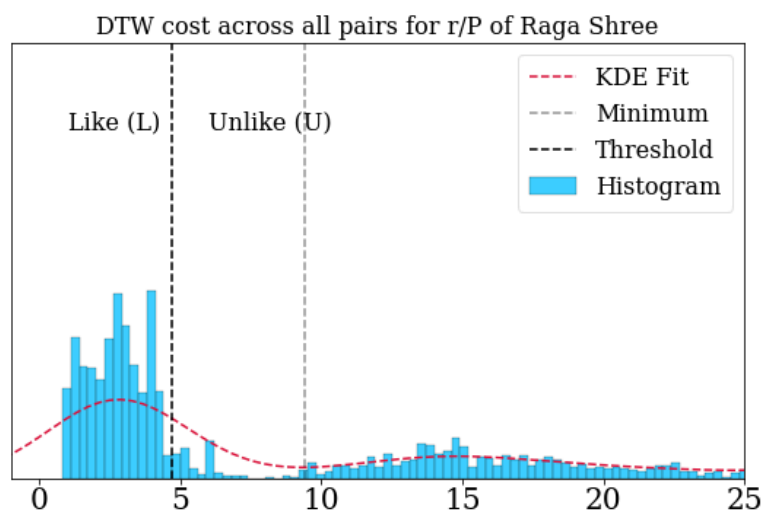
5S.2: Ragawise distribution of stable note and non-stable segment durations across singers. All ragas follow the trend of a distribution that is skewed towards lower durations for stable note segments.





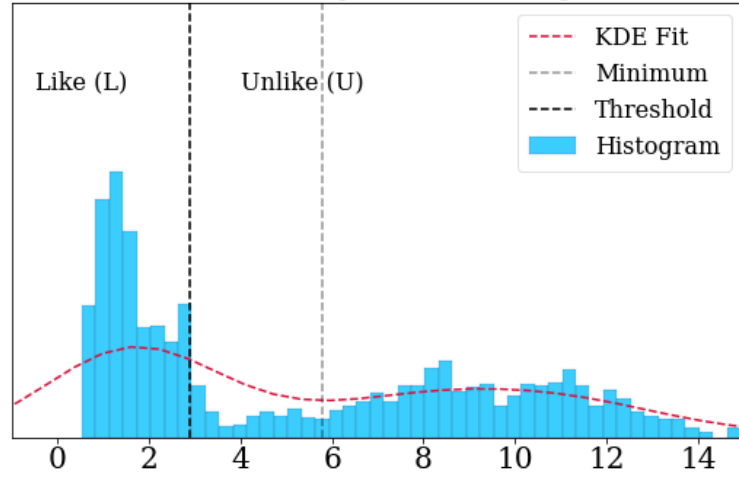
## 5S.3 Histograms and KDE Plots for each raga phrase

The following plots show the histogram for the DTW cost obtained from the audio-based subsequence alignment. (Figure 9 of the paper shows the distribution computed for the raga Nand P\|R phrase.) The KDE fit is superposed on each distribution. The KDE is bimodal and the threshold to label L and U is set at a half of DTW cost at the valley between the peaks in all cases. Only for nDN (Bahar), we do not get a bimodal distribution so the threshold is set manually at one-eighth the location of the minimum. This threshold ensures that the data is roughly balanced between Like and Unlike segments and that the like segments are indeed melodically similar to the pakad templates based on informal checks by an expert listener.

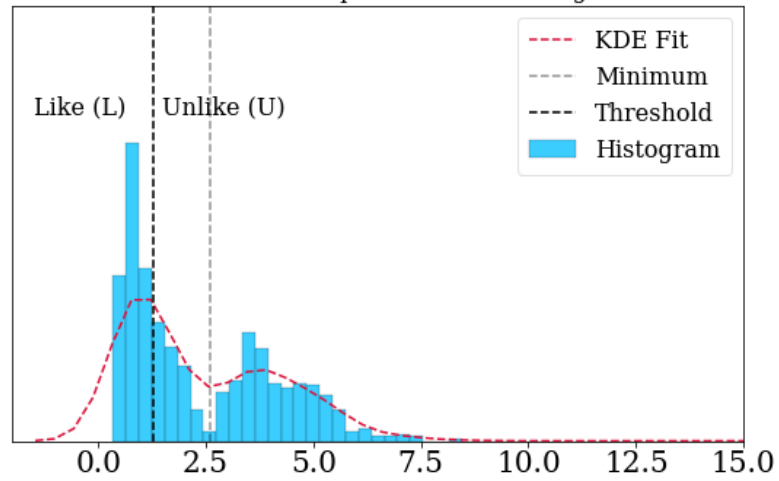




DTW cost across all pairs for P\R of Raga Nand



DTW cost across all pairs for nDN of Raga MM



DTW cost across all pairs for nDN of Raga Bahar

