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 organized by the section number they correspond to in the main paper.

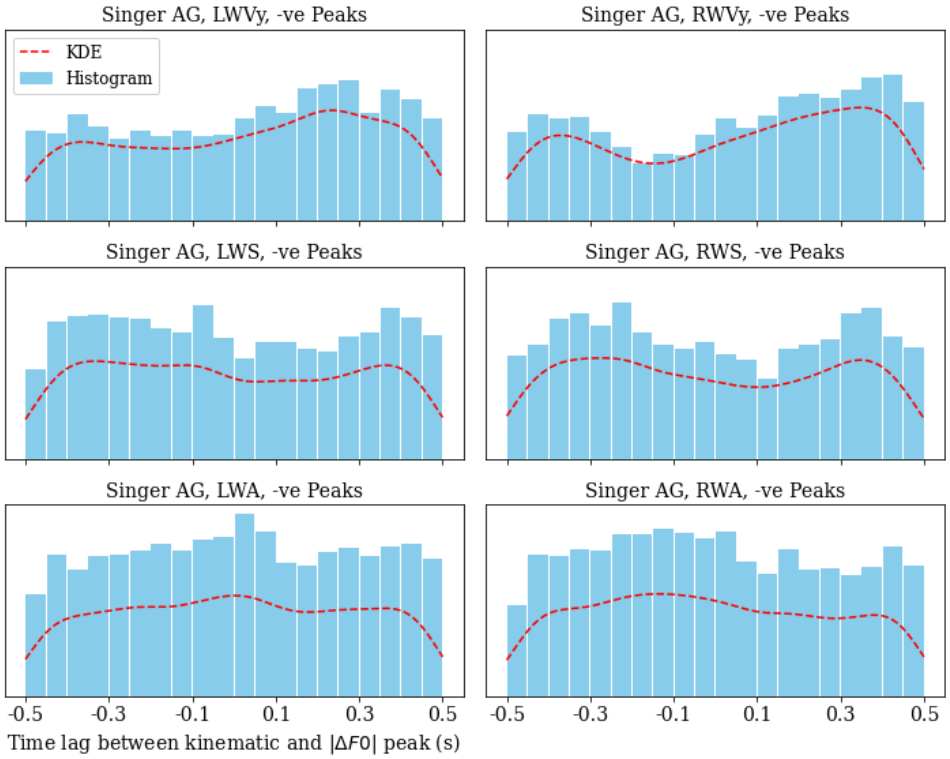
Contents

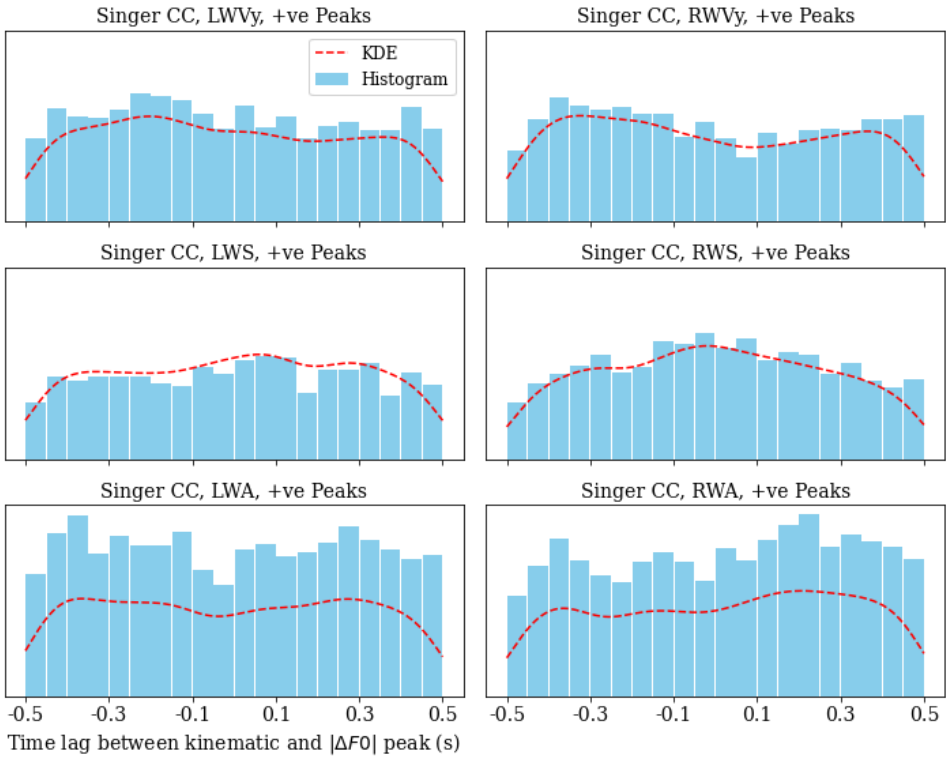
Section 4S: Relating F0 with Kinematic Changes

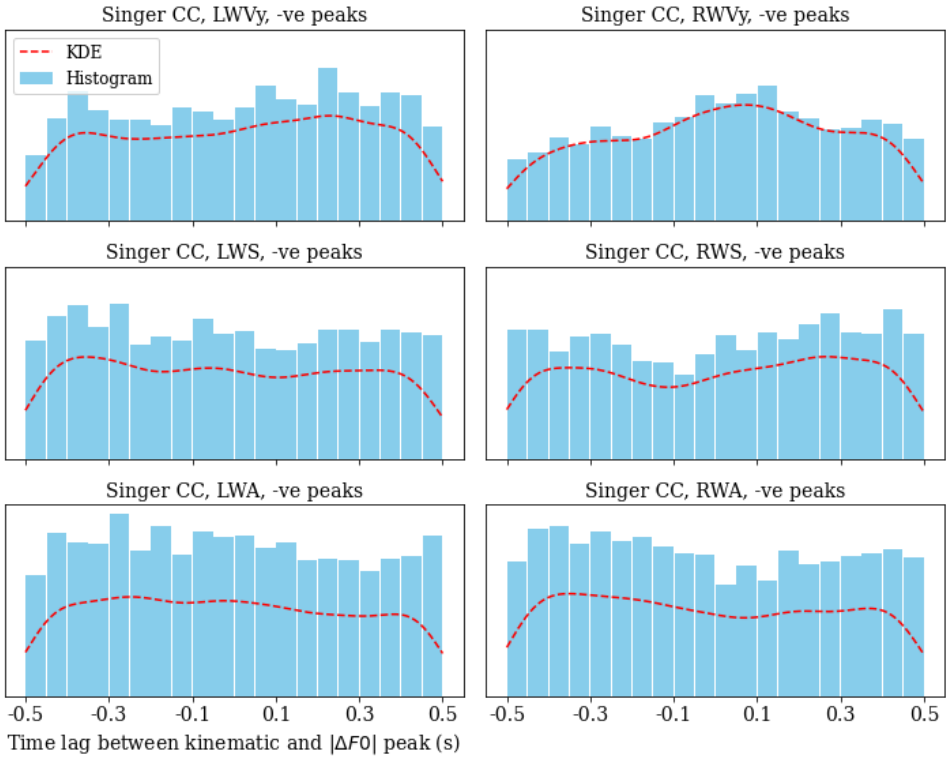
4S.2 Temporal Coupling Study

The following figures present distributions for the time offset values (in seconds) observed between kinematic peaks and corresponding |ΔF0| 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 |ΔF0| peak leading the kinematic peak, while positive lags indicate the |ΔF0| 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.

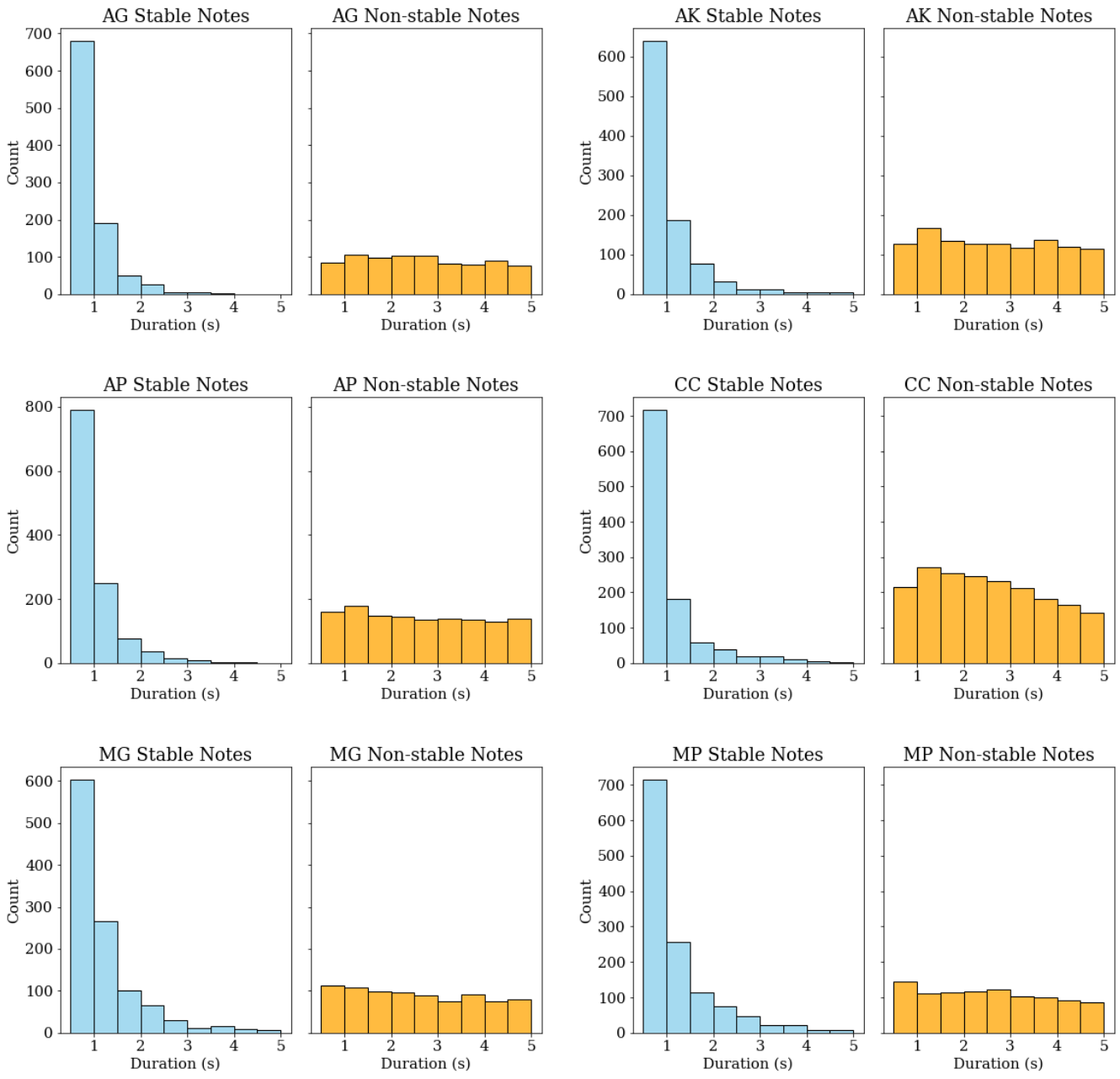


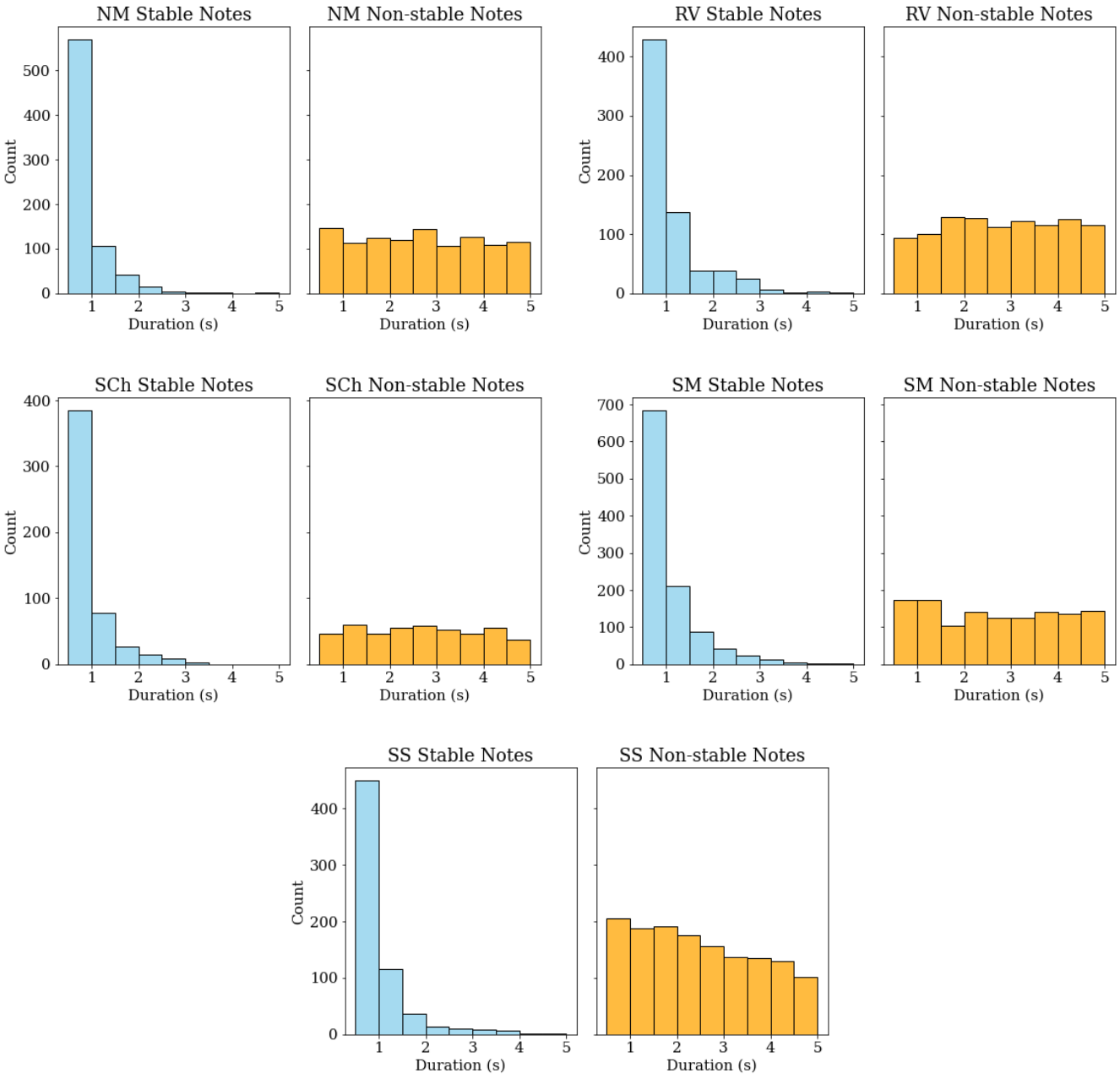

Singer CC, LWVy, +ve Peaks
KDE
Histogram
Singer CC, RWVy, +ve Peaks
Singer CC, LWS, +ve Peaks
Singer CC, RWS, +ve Peaks
Singer CC, LWA, +ve Peaks
Singer CC, RWA, +ve Peaks
-0.5
-0.3
-0.1
0.1
0.3
0.5
Time lag between kinematic and |ΔF0| peak (s)


Singer CC, LWVy, -ve peaks
KDE
Histogram
Singer CC, RWVy, -ve peaks
Singer CC, LWS, -ve peaks
Singer CC, RWS, -ve peaks
Singer CC, LWA, -ve peaks
Singer CC, RWA, -ve peaks
-0.5
-0.3
-0.1
0.1
0.3
0.5
Time lag between kinematic and |ΔF0| peak (s)

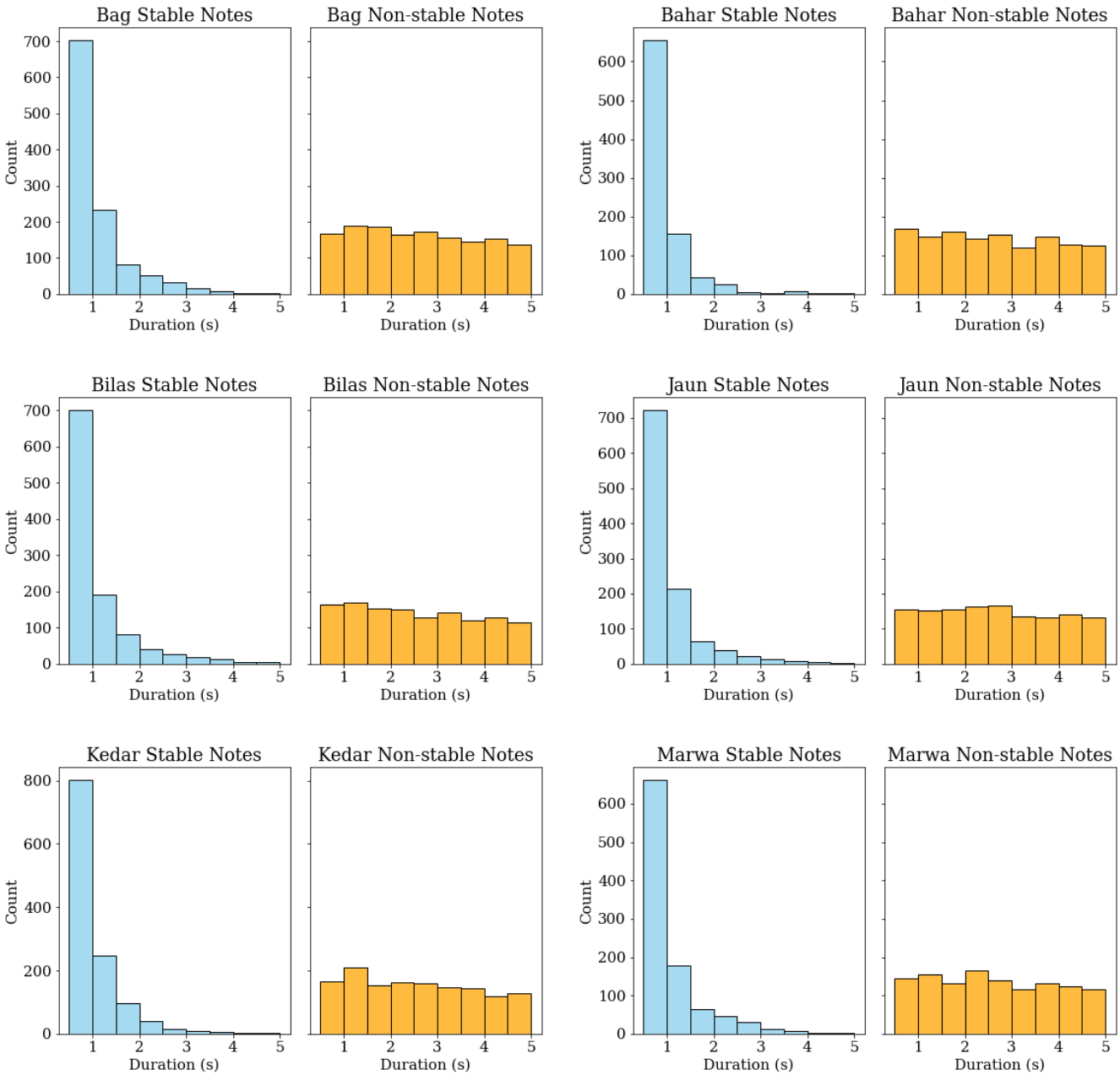
Section 5S: F0 segmentation methods

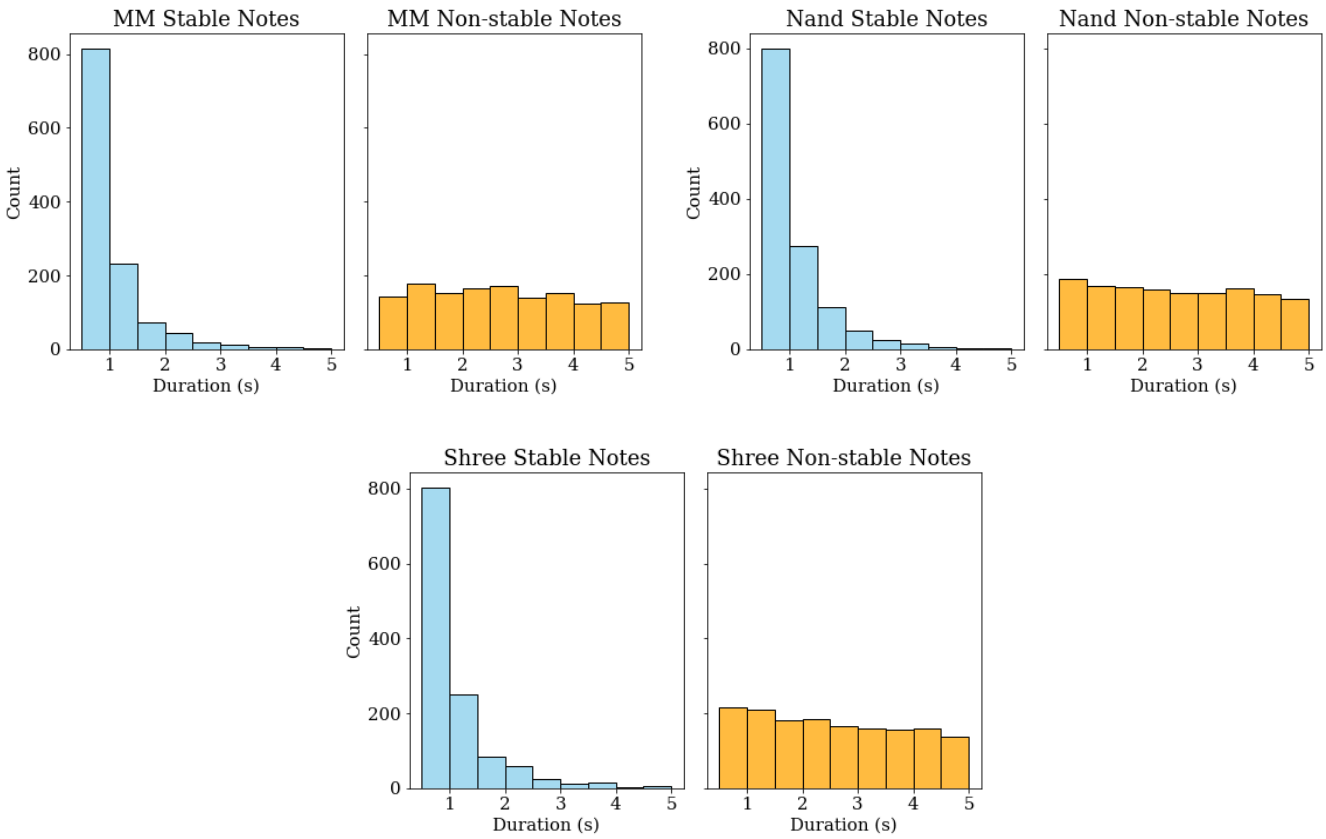
5S.1: Singerwise distribution of stable note and non-stable segment durations across ragas. All singers follow the trend of a distribution that is skewed towards lower durations for stable note segments.




NM Stable Notes
NM Non-stable Notes
RV Stable Notes
RV Non-stable Notes
SCh Stable Notes
SCh Non-stable Notes
SM Stable Notes
SM Non-stable Notes
SS Stable Notes
SS Non-stable Notes
Count
Duration (s)

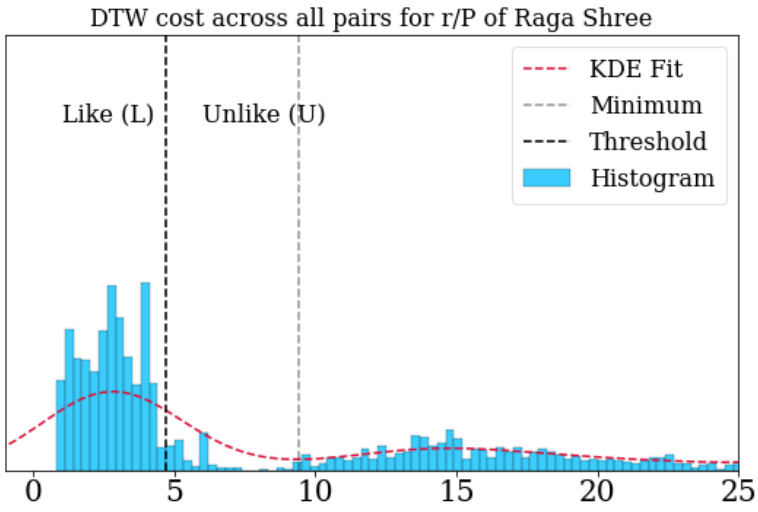
5S.2: Ragawise distribution of stable note and non-stable segment durations across singers. All ragas follow the trend of a distribution that is skewed towards lower durations for stable note segments.

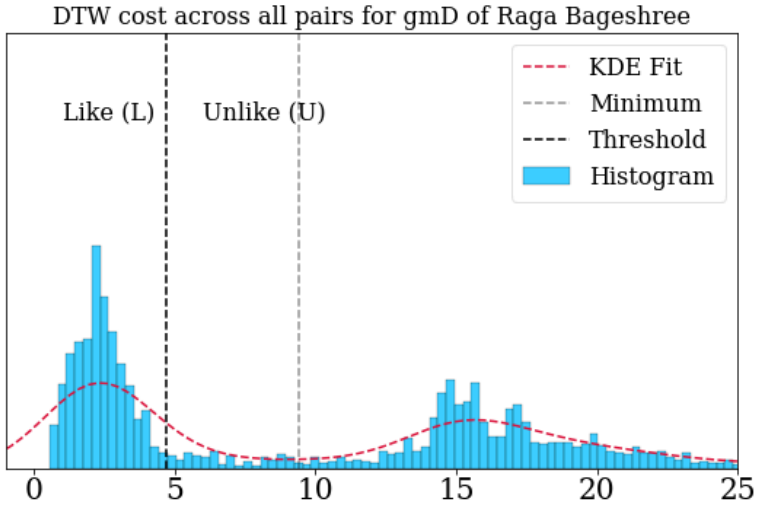


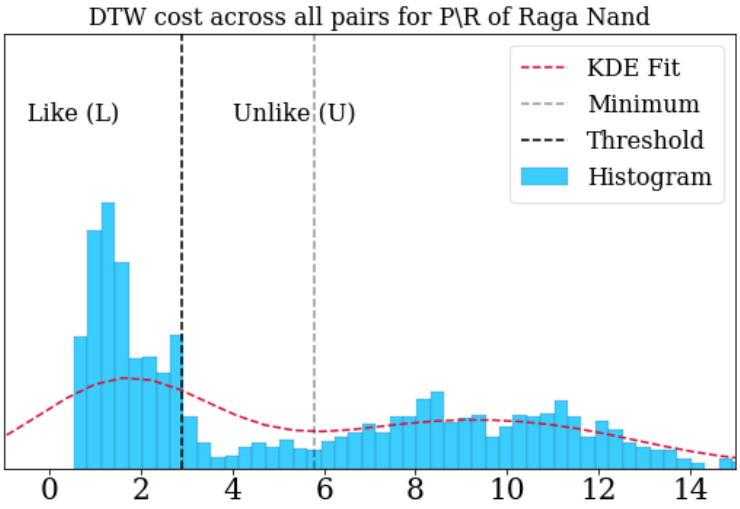

MM Stable Notes
MM Non-stable Notes
Nand Stable Notes
Nand Non-stable Notes
Shree Stable Notes
Shree Non-stable Notes
Count
Duration (s)
0
200
400
600
800
1
2
3
4
5

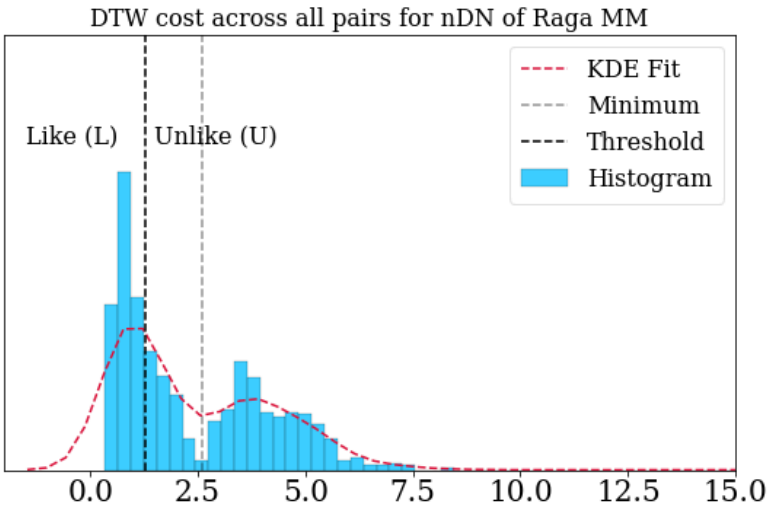
5S.3 Histograms and KDE Plots for each raga phrase

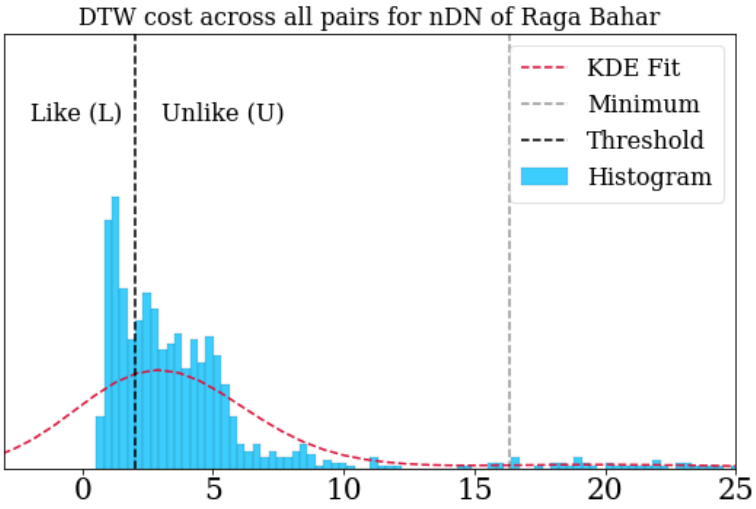
The following plots show the histogram for the DTW cost obtained from the audio-based subsequence alignment. (Figure 9 of the paper shows the distribution computed for the raga Nand P\R phrase.) The KDE fit is superposed on each distribution. The KDE is bimodal and the threshold to label L and U is set 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
KDE Fit
Minimum
Threshold
Histogram
Like (L)
Unlike (U)
0
2
4
6
8
10
12
14


DTW cost across all pairs for nDN of Raga MM
KDE Fit
Minimum
Threshold
Histogram
Like (L)
Unlike (U)
0.0
2.5
5.0
7.5
10.0
12.5
15.0


DTW cost across all pairs for nDN of Raga Bahar
Like (L)
Unlike (U)
KDE Fit
Minimum
Threshold
Histogram
0
5
10
15
20
25

Section 7S: Results for Singer-specific Gesture based phrase detection

In the following tables, we report the F1-scores after stratified 10-fold cross-validation for the raga-phrase detection framework for each phrase. Bold values indicate results that are statistically significant when compared to a chance classifier which predicts 'Like' with probability equal to the percentage of like segments ($p < 0.05$). In some of the cases, the singer did not sing the particular phrase in the pakad, or the alap instances were too few in number (less than 15 in total). In such cases, the singer's data was not considered in the experiment and the tables contain a blank '-' for them. Also, for ragas MM and Bahar, the overall number of segments was low because of which we decided to do a 3-fold cross validation rather than a 10-fold cross validation, since the implementation requires at least 1 example from each class in the testing data in every fold.

Raga Shree, Phrase r/P

	AG	**AK**	**AP**	**CC**	**MG**	**MP**	**NM**	**RV**	**SCh**	**SM**	**SS**
Count	31	38	41	52	28	27	30	47	36	36	49
% Like	71	57.9	43.9	30.8	50	48.1	56.7	61.7	38.9	61.1	42.9
DTW-D (1)	83	71.2	0	11.8	80.0	48.3	74.4	**76.3**	81.5	75.9	0
DTW-I (1)	88.	73.3	0	40.0	59.3	23.5	73.7	**76.3**	63.6	75.9	0
DTW-LR (2)	76	71.2	48	43.5	76.9	43.5	72.2	**76.3**	**92.3**	75	58.8
DTW-Ind (36)	81	**78.0**	**77.8**	**85**	**82.8**	**69.6**	**82.4**	74.6	85.7	74.4	**66.7**
DTW-Ind-W (18)	**87.0**	60.5	**77.8**	78.6	75.9	64	74.3	73.3	89.7	**78.3**	42.9

Raga Bageshree, Phrase gmD

	AG	**AK**	**AP**	**CC**	**MG**	**MP**	**NM**	**RV**	**SCh**	**SM**	**SS**
Count	30	28	34	35	19	23	28	32	24	36	42
% Like	60	60.7	47.1	31.4	68.4	60.8	46.4	56.3	58.3	50	57.1
DTW-D (1)	**87.2**	**77.3**	0	28.6	86.7	75.7	0	61.9	73.7	58.3	**72.7**
DTW-I (1)	75	75.6	0	26.7	77.4	75.7	0	**68.1**	73.7	41.2	**72.7**
DTW-LR (2)	**87.2**	76.2	55.2	60	77.4	70.6	34.8	**68.1**	**81.3**	62.5	70.8
DTW-Ind (36)	**87.2**	64.7	71.0	60.9	**88.9**	**83.9**	**72**	64.7	71.4	**73.7**	64
DTW-Ind-W (18)	78.9	60.6	**74.1**	**66.7**	84.6	78.6	60.9	66.7	72.2	66.7	65.3

Raga Nand, Phrase P\R

R	AG	AK	AP	CC	MG	MP	NM	RV	SCh	SM	SS
P\R Count	23	33	22	27	25	21	35	37	-	32	40
P\R % Like	78.3	42.4	40.9	25.9	48	47.6	42.8	43.2	-	53.1	42.5
P\R DTW-D (1)	**87.8**	0	0	54.5	14.3	**62.5**	0	0	-	69.4	9.5
P\R DTW-I (1)	**87.8**	0	46.2	61.5	37.5	16.7	47.6	64	-	72.3	0
P\R DTW-LR (2)	**87.8**	**66.7**	46.2	61.5	66.7	40	30	56	-	75.0	38.5
P\R DTW-Ind (36)	80	60	**55.6**	**72.7**	**72.7**	52.6	**64.5**	**66.7**	-	**80.0**	54.5
P\R DTW-Ind-W (18)	87.2	62.5	50	54.5	72	60	60	56.2	-	76.5	**60.6**

Raga MM, Phrase nDN (3-fold CV)

	AG	AK	AP	CC	MG	MP	NM	RV	SCh	SM	SS
nDN Count	19	16	-	16	-	-	18	-	-	-	21
nDN % Like	68.4	56.3	-	56.3	-	-	61.1	-	-	-	28.6
nDN DTW-D (1)	81.1	78.6	-	72.2	-	-	75.6	-	-	-	0
nDN DTW-I (1)	81.1	72.2	-	72.2	-	-	18	-	-	-	0
nDN DTW-LR (2)	73.3	**95.2**	-	86.9	-	-	69	-	-	-	48.9
nDN DTW-Ind (36)	89.6	77.8	-	64.1	-	-	76.7	-	-	-	13.3
nDN DTW-Ind (18)	90.7	84.1	-	82.1	-	-	78.6	-	-	-	26.7

Raga Bahar, Phrase nDN (3-fold CV)

	AG	AK	AP	CC	MG	MP	NM	RV	SCh	SM	SS
nDN Count	-	-	-	44	-	-	29	-	-	37	30
nDN %Like	-	-	-	38.6	-	-	62	-	-	56.8	36.7
nDN DTW-D (1)	-	-	-	61.8	-	-	75.1	-	-	71.7	11.1
nDN DTW-I (1)	-	-	-	25	-	-	75.5	-	-	71.7	52.4
nDN DTW-LR (2)	-	-	-	21.6	-	-	76.3	-	-	63.6	22.9
nDN DTW-Ind (36)	-	-	-	57.8	-	-	**81.1**	-	-	61	20.6
nDN DTW-Ind (18)	-	-	-	47	-	-	73.7	-	-	64.8	22.9