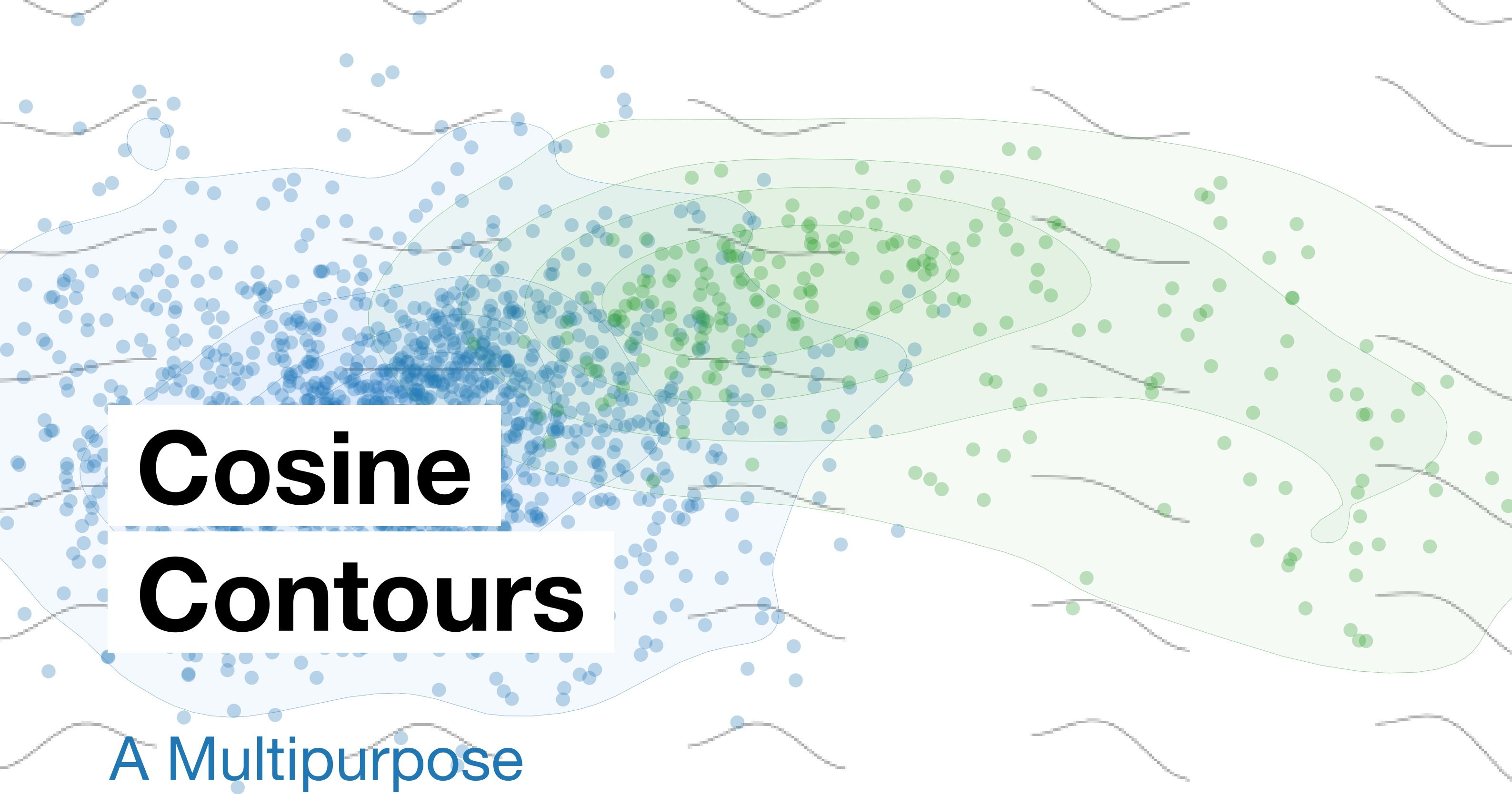
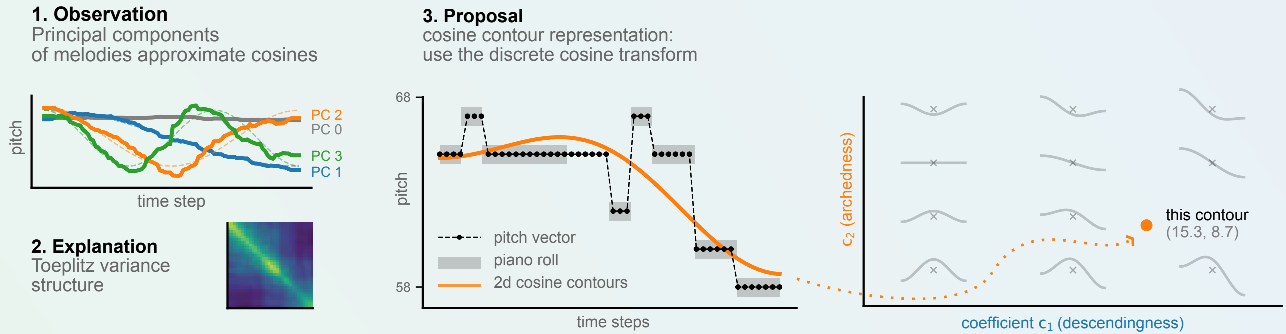


Cosine Contours

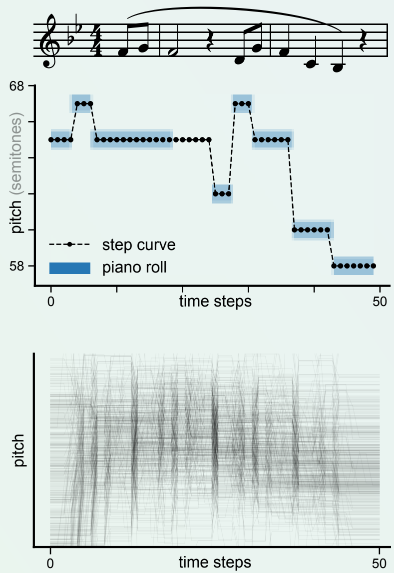
A Multipurpose Representation for Melodies

Cornelissen, Zuidema & Burgoyne
Institute for Logic, Language & Computation,
University of Amsterdam

Graphical Abstract



Data
and representation



Melodic segments represented as fixed-length vectors of pitches.

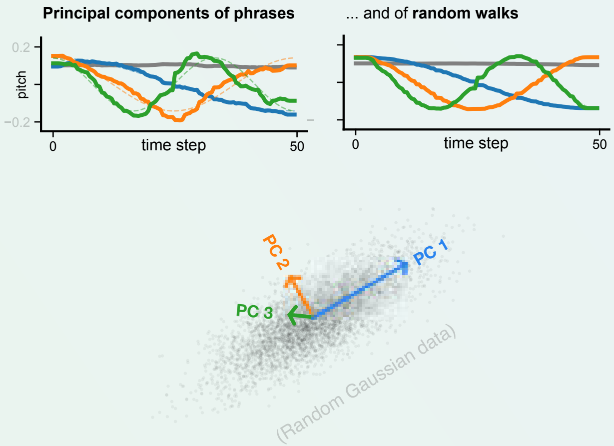
Essen Chinese and German folksongs from Essen (subsets); *phrases*.

Densmore: *song* contours from various Native American cultures from the Densmore collections

Chant *motif* (neume/word/syllable) contours from CantusCorpus; *phrase* and *song* contours from GregoBaseCorpus.

Observation

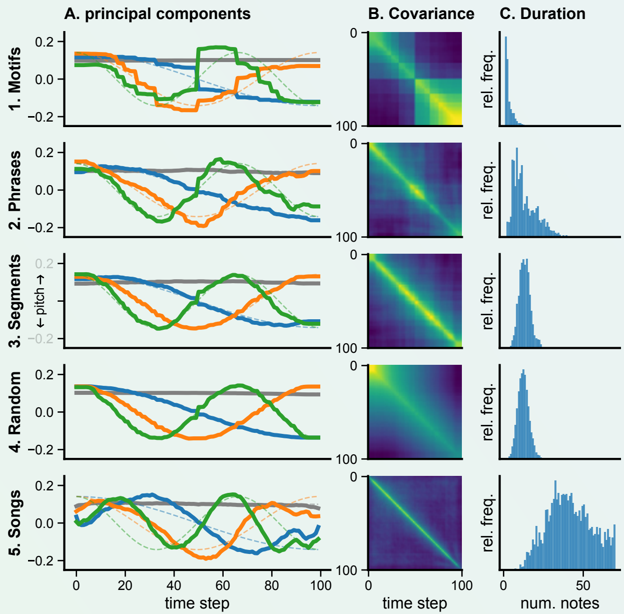
PCs are cosines



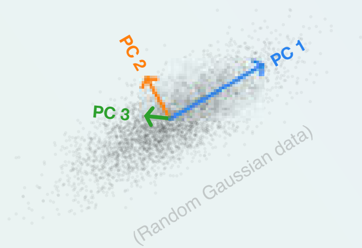
Principal components of melodic phrases approximate cosine functions of increasing frequency

Principal components are vectors in the same space as the data: vector of 50 pitches which we visualize as contours.

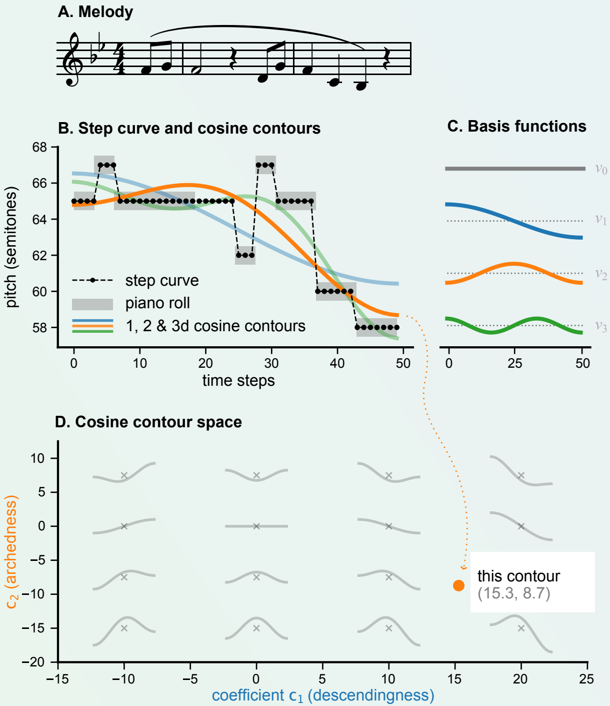
Explanation
Toeplitz covariance



The covariance matrix roughly resembles a *Toeplitz* matrix, which asymptotically have Fourier basis functions as eigenvectors.



Proposal
Cosine contours



Cosine contours represent a melodic contour as a combination of cosine functions using the discrete cosine transform.

Near-optimal representation: most variance captured in few dimensions (like PCA).

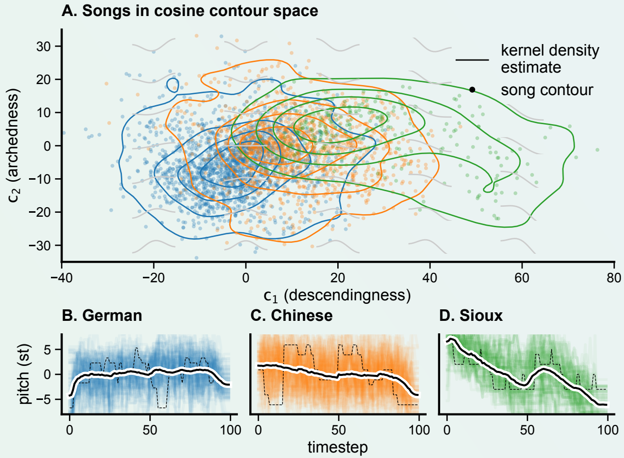
Data-independent: contours from different traditions live in the same low-dimensional space (unlike PCA).

Intuitive and interpretable contour space

Variable level of abstraction: increase the dimension to move from the rough shape to the exact melody.

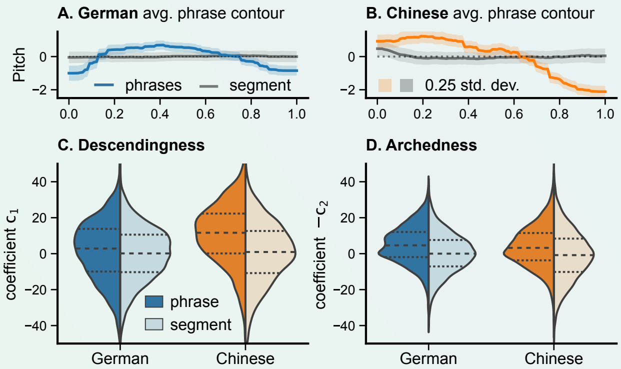
Case study 1

Visualizing different traditions



Songs of three cultures represented in the cosine contour space **(A)** show substantial variability. The average of all contours in a tradition **(B–D)** also illustrates this (thick black lines; dashed lines highlight one contour).

Case study 2
Melodic arch hypothesis

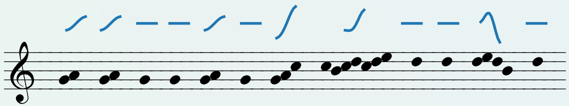


Phrases of German (A) and Chinese (B) songs tend to be more descending and arched compared to random segments from the same melodies, as visible from their average contours. This can be quantified by comparing the first **(C)** and second **(D)** coefficients of their cosine representations.

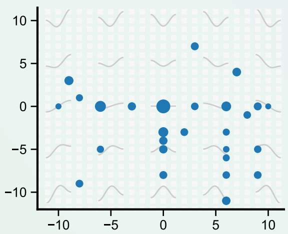
Case study 3

Mode classification

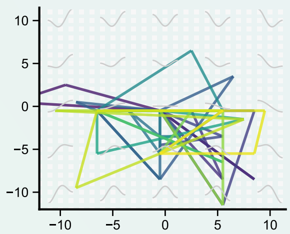
A. Chant and cosine contours



B. tf–idf vector visualized



C. Chant as a walk



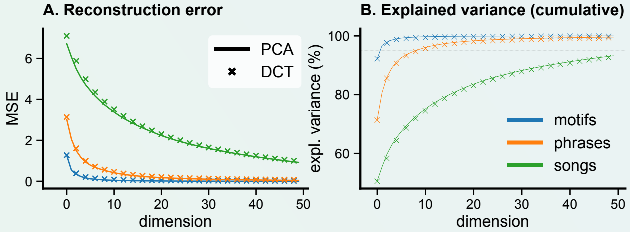
D. Mode classification results (accuracy)

	Responsory		Antiphon	
Neumes	52	**74**	30	**49**
Syllables	76	**79**	35	**53**
Words	**81**	73	**83**	**83**
Contour repr.	Parson	Cosine	Parson	Cosine

Motifs used for mode classification in Gregorian chant. **(A)** A chant is segmented into motifs derived from the notation (neumes) or lyrics (syllables, words). The blue curves show the two-dimensional cosine contours for those motifs. **(B)** We discretize the contour space and represent the chant as a vector of tf–idf weighed motif frequencies ('grid cell frequencies'). Dots illustrate the nonzero entries of this vector for the chant shown above. **(C)** The chant is now a walk through contour space, but our 'bag of motifs' ignores order. **(D)** Using these vectors to classify mode, we outperform a previous study using

Evaluation

Optimality



DCT approximates PCA, the optimal transform, in terms of the reconstruction error **(A)** and the explained variance ratio **(B).** The reconstruction error is the mean squared error between an contour and a lower dimensional reconstruction.