Introduction to Music Processing

Hierarchical Models

Key-Scape Plots, Pitch-Scape Clustering and RBNs

Dr Robert Lieck
robert.lieck@durham.ac.uk

Key-Scape Plots

Prélude No. 1 in C Major

from “Das Wohltemperierte Klavier” Book I
BWV 846

Johann Sebastian Bach
(1685 - 1750)

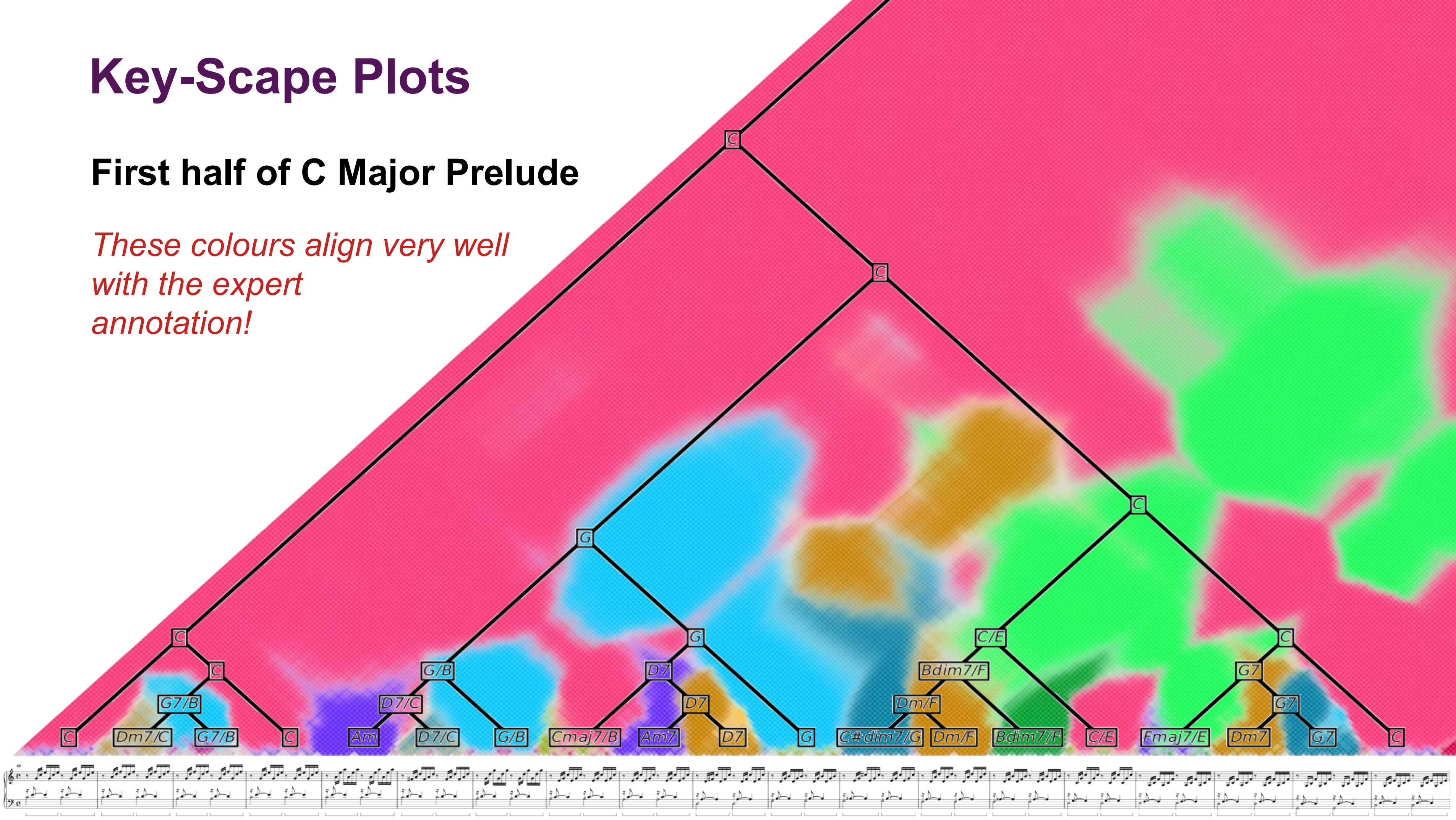
Durham University

https://www.youtube.com/watch?v=frxT2qB1POQ

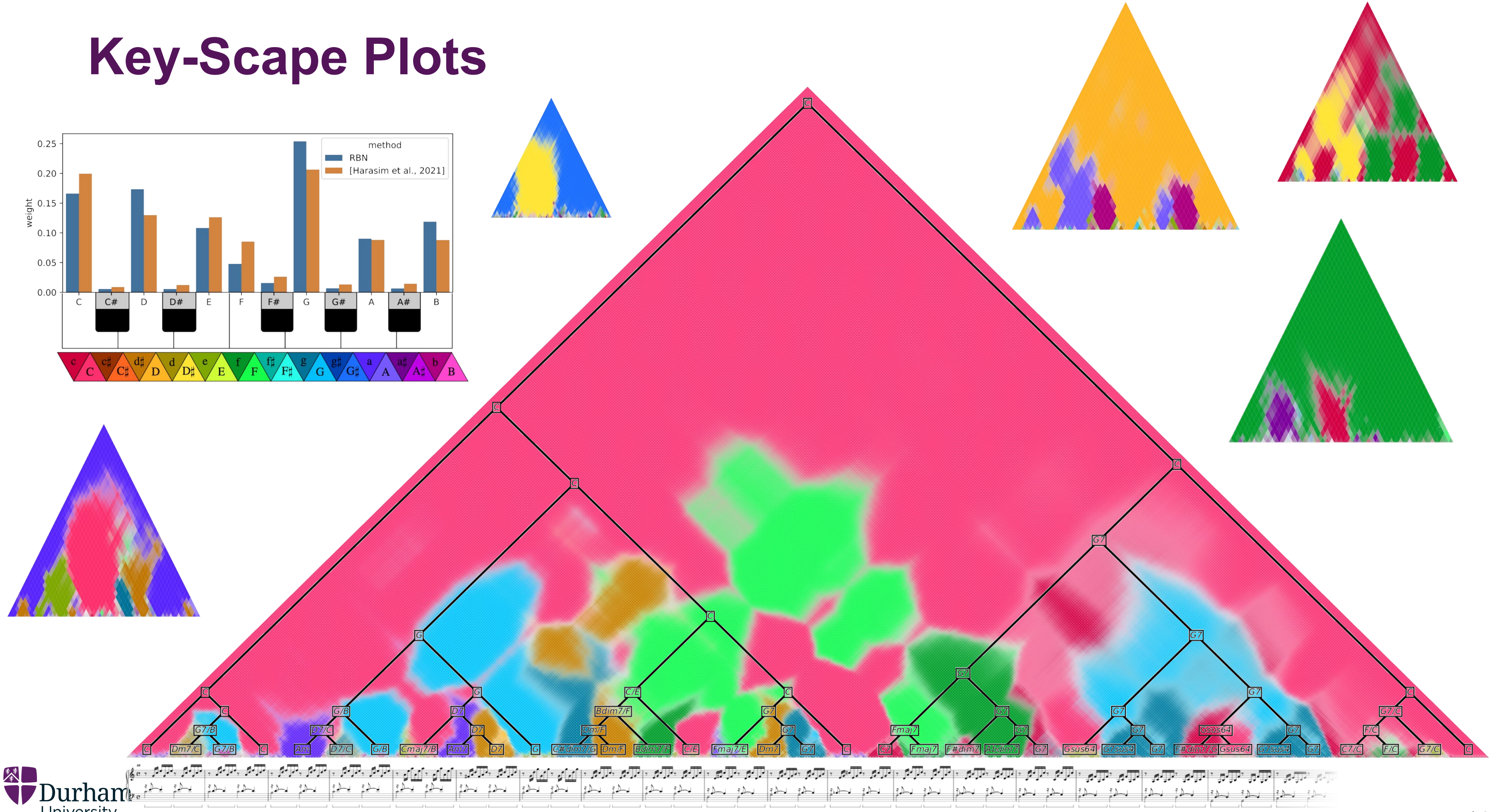
Key-Scape Plots

First half of C Major Prelude

These colours align very well with the expert annotation!



Key-Scape Plots



Durham University

Key-Scape Plots: Visualising Pitch Scapes

Compute Pitch Scape:

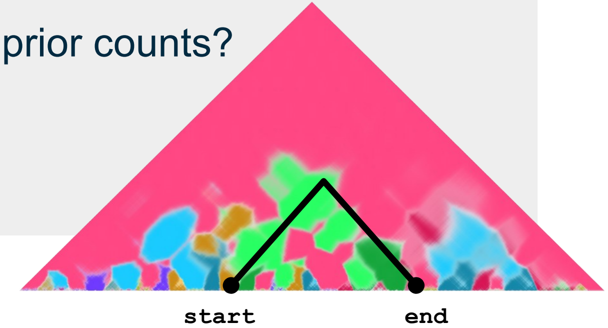
For each point `(start, end)`:

1) get section of music

2) extract the pitch-class distribution / chroma vector

   → how to weight different notes?

   → use prior counts?



Visualise Pitch Scape:

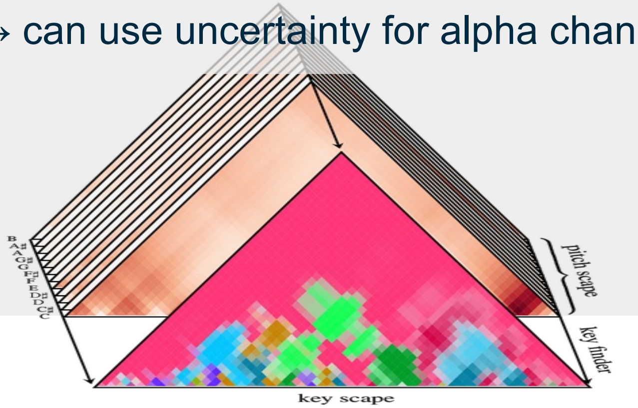
1) find the key via template matching

   → can use matching scores for softmax

2) pick colour according to the key

   → need to decide on colour palette

   → can use uncertainty for alpha channel



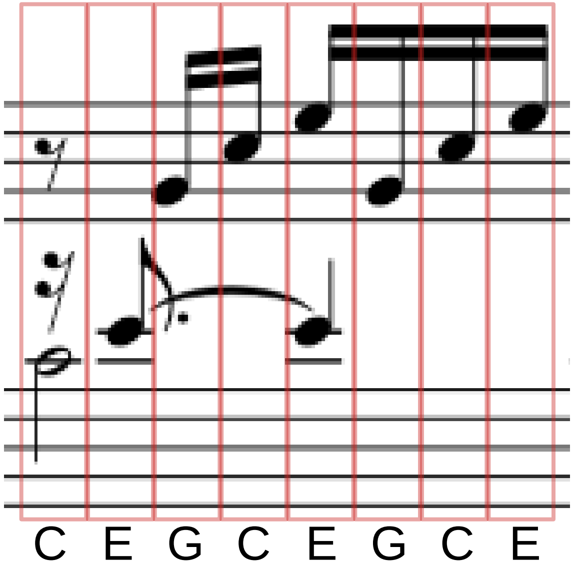
Durham University

Computing Pitch Scapes

How to weight different notes?

Pitch-class counts:

C	C#	D	D#	E	F	F#	G	G#	A	A#	B
1											
1				1							
1				1			1				
2				1							
1				2							
1				1			1				
2				1							
1				2							

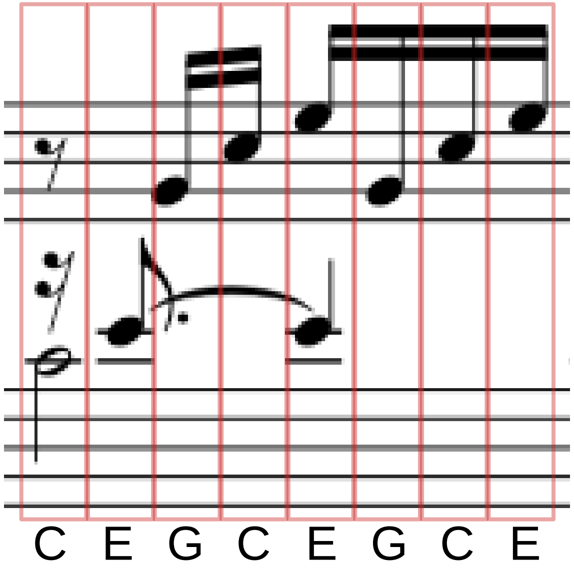


Computing Pitch Scapes

How to weight different notes?

Normalised pitch-class counts:

C	C#	D	D#	E	F	F#	G	G#	A	A#	B
1											
1/2				1/2							
1/3				1/3			1/3				
2/3				1/3							
1/3				2/3							
1/3				1/3			1/3				
2/3				1/3							
1/3				2/3							

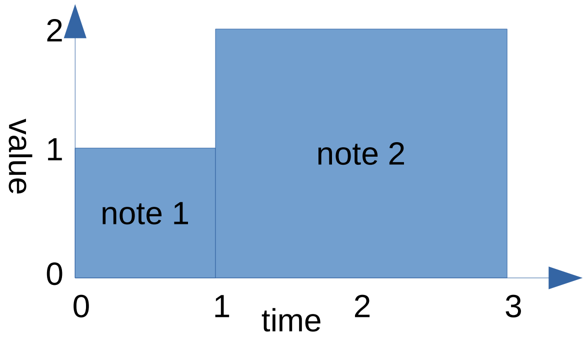


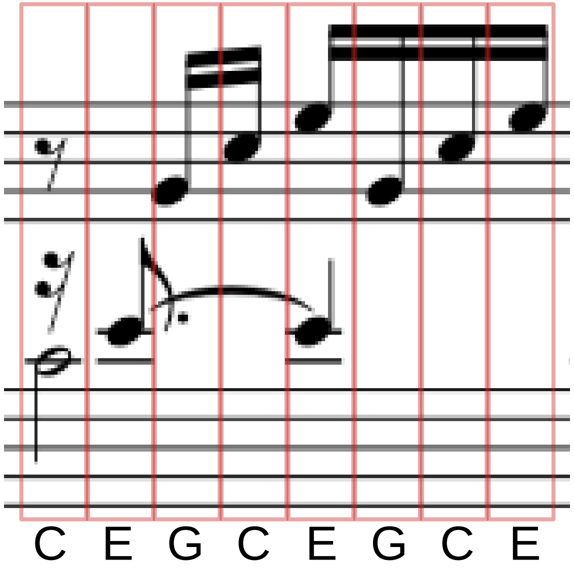
Computing Pitch Scapes

How to weight different notes?

How do we sum over time?

- Summing PC counts or normalised counts?
- Take duration into account?
- Normalise output to get PCD?
- *What is the “sum” from 0 to 3 below?*
  - Use duration as weight
    → area/integral over pitch-class “density”





Computing Pitch Scapes

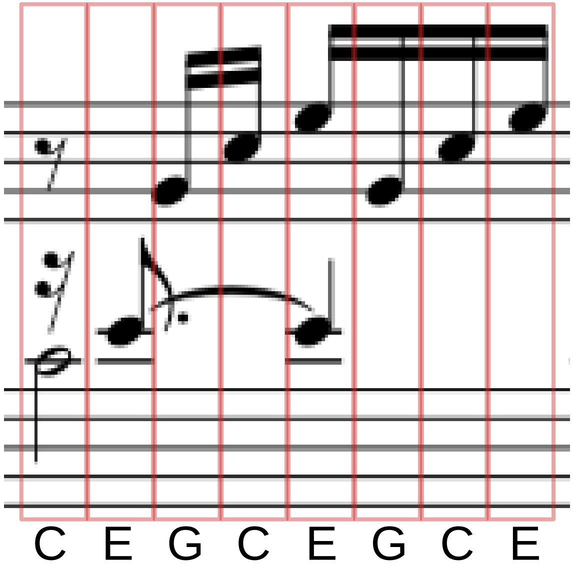
How to weight different notes?

How to use prior counts?

- In *k*-Markov and *n*-gram models we used prior counts to avoid problems due to low counts (the zero count problem and noisy estimates).
- In pitch scapes, we need to add prior counts ***after*** summing over time.

  → assume uniform distribution without data

  → give less weight to short events (e.g. ornamentations)

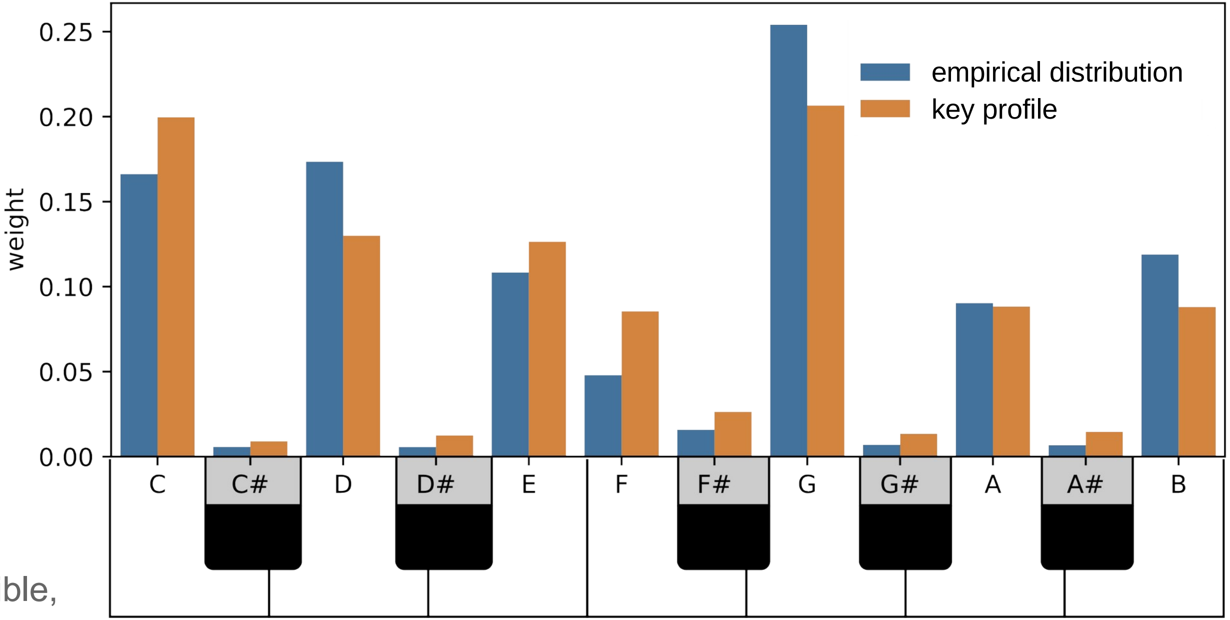


Durham University

Visualising Pitch Scapes

Keyfinding

- compare empirical distribution against key profile
  (different possibilities for how to obtain profiles)
- go through all possible rotations/transpositions and modes (major/minor)
- compute distance/score
  (different distance metrics possible, typically Euclidean)
- pick best match or compute softmax



Durham University

Visualising Pitch Scapes

$$\boxed{\text{score}(\text{distribution} = d \mid \text{mode} = m, \text{transposition} = t) := \; -\,\text{distance}(d,\; \text{profile}(m, t))}$$

Uncertainty Quantification

- Interpret scores as logits (un-normalised log-likelihoods)
- Get distribution over modes and transpositions
  (use temperature tau and possibly prior information)

$$p(m, t \mid d) := \frac{\exp[\text{score}(d \mid m, t)/\tau] \cdot \text{prior}(m, t)}{\sum_{m',t'} \exp[\text{score}(d \mid m', t')/\tau] \cdot \text{prior}(m', t')}$$

- Entropy is bounded uncertainty measure

$$H(m, t|d) = -\sum_{m,t} \log p(m, t \mid d)$$

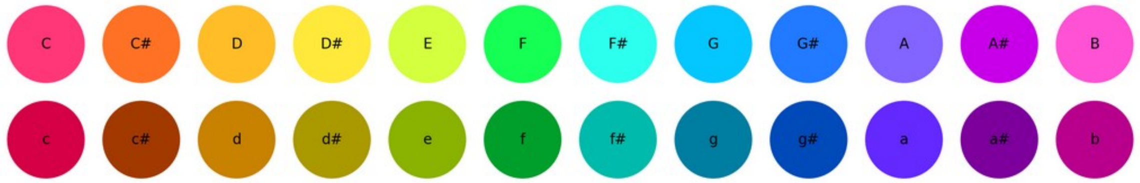
$$H_{\text{min}} = 0$$
$$H_{\text{max}} = \log(|m| \cdot |t|)$$

Durham University

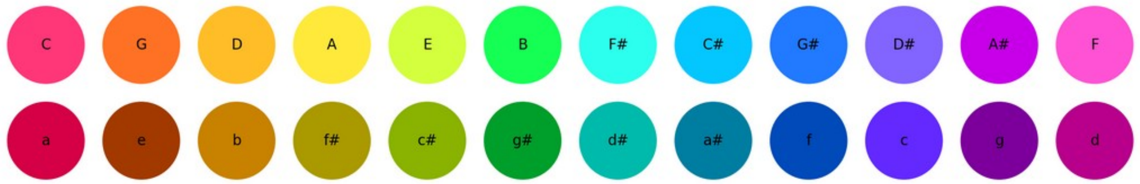
Visualising Pitch Scapes

Colour Palettes

- Chromatic circle
  → keys ordered by frequency
  → parallel keys have similar colour



- Circle of fifths
  → keys ordered by fifths/signature
  → relative key have similar colour



Durham University

Other applications of scape plots…

(here repetitive structure but also see other references provided at the end)

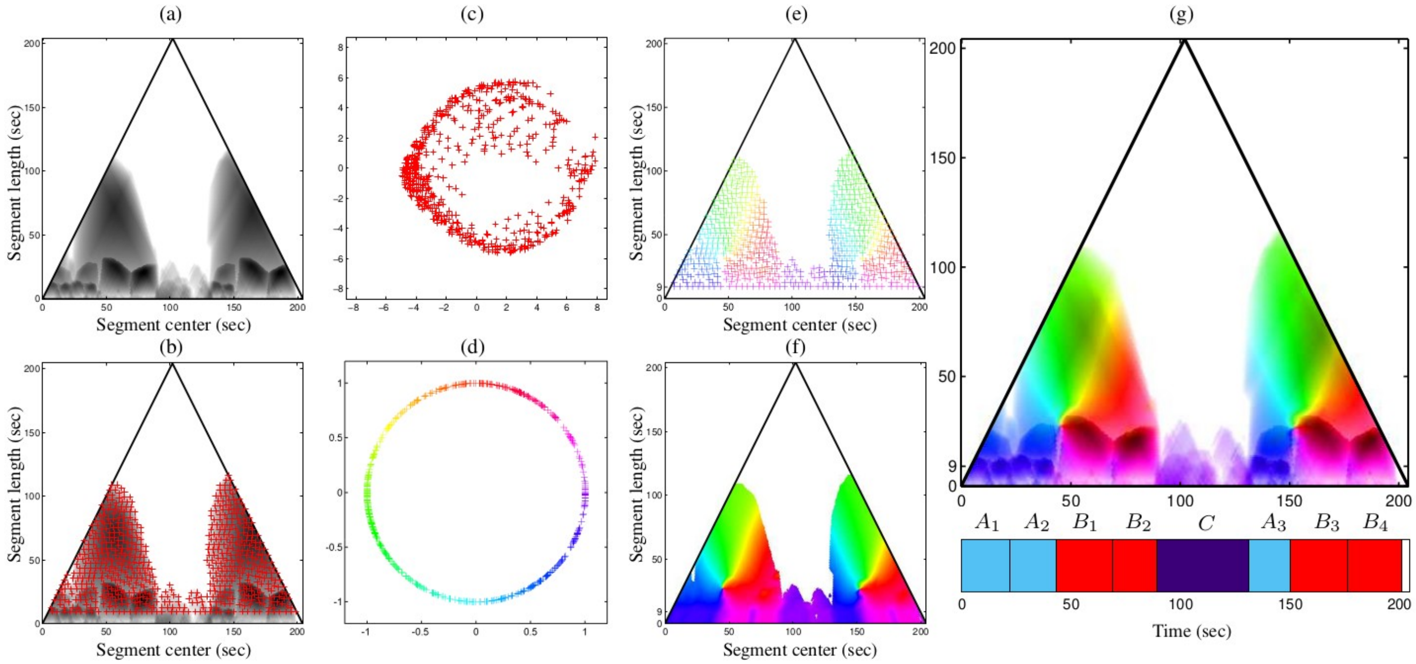
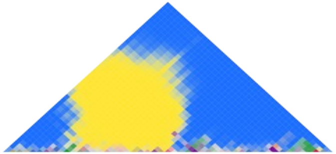
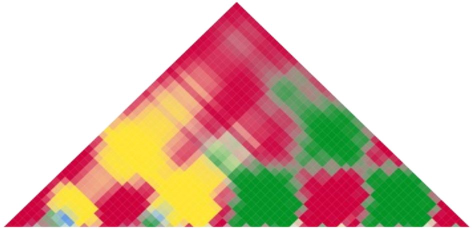


Figure 4: Illustration of the pipeline for computing the structure scape plot for Brahms. **(a)** Fitness scape plot. **(b)** Fitness scape plot with sampled anchor points. **(c)** Anchor points projected onto the first two principal components. **(d)** Anchor points projected to the unit circle colored with the resulting hue value. **(e)** Hue-colored anchor points. **(f)** Hue-colored scape plot using interpolation techniques. **(g)** Structure scape plot combining fitness (lightness) and cross-segment relation (hue) information.

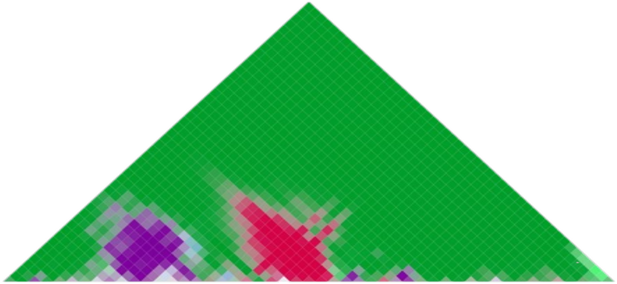
Durham University

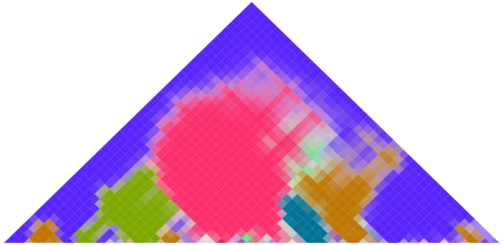
Müller M, Jiang N (2012) A Scape Plot Representation for Visualizing Repetitive Structures of Music Recordings. In: ISMIR. Citeseer, pp 97–102

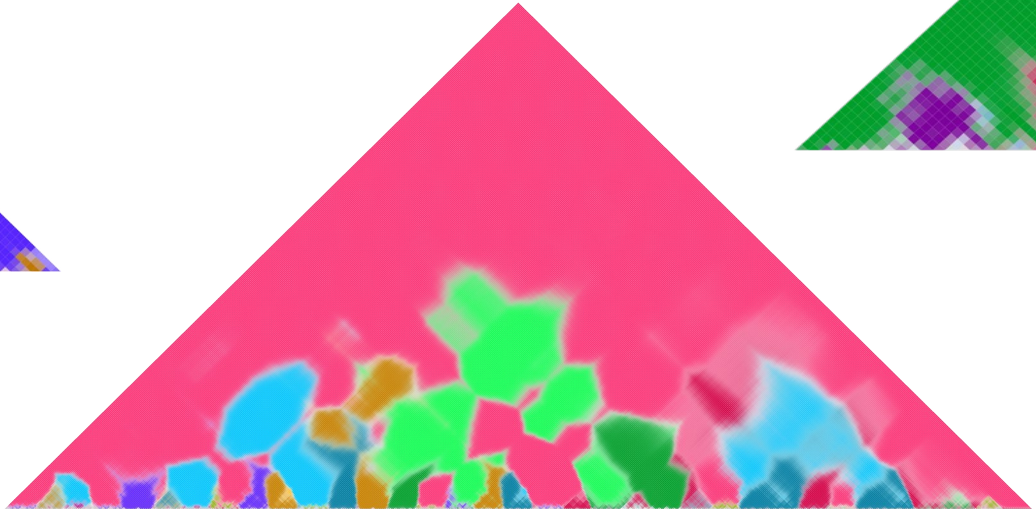




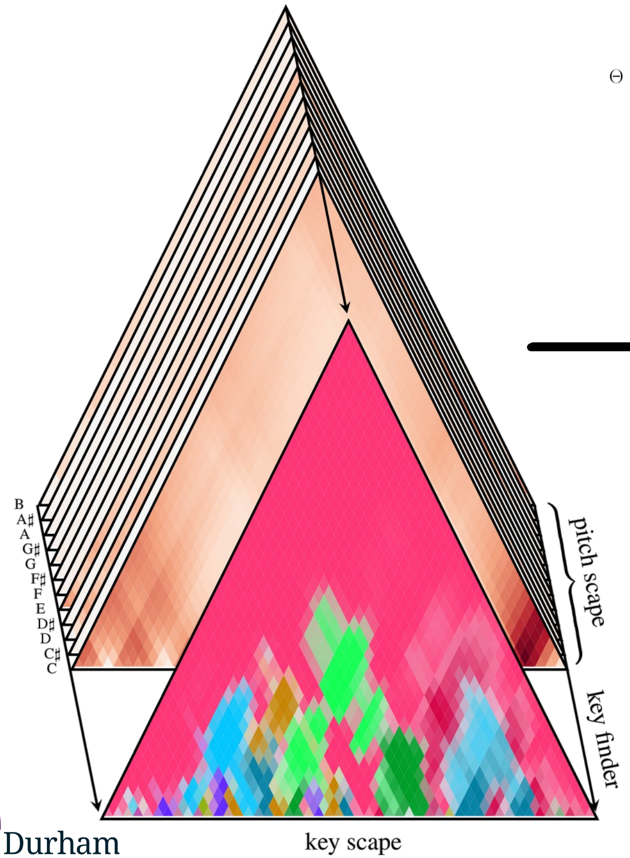
What next…?

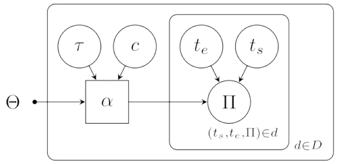






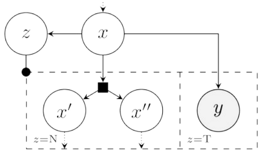
Close and Distant Reading



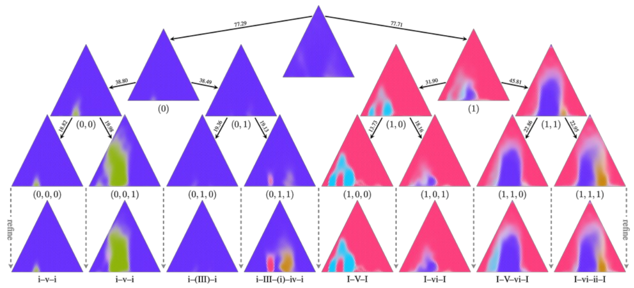


zoom out

zoom in

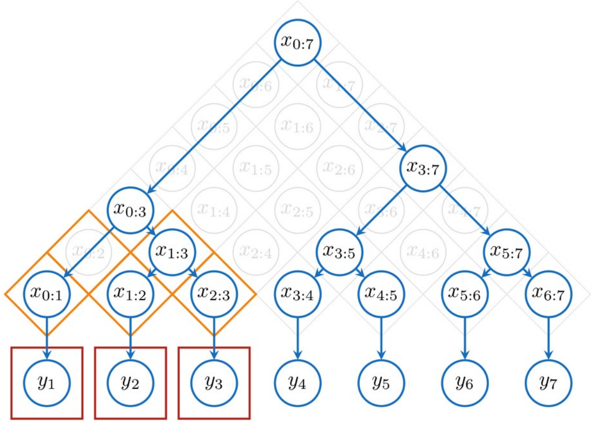


Pitch Scape Clustering



Lieck & Rohrmeier (2020)

Recursive Bayesian Networks



Lieck & Rohrmeier (2021)

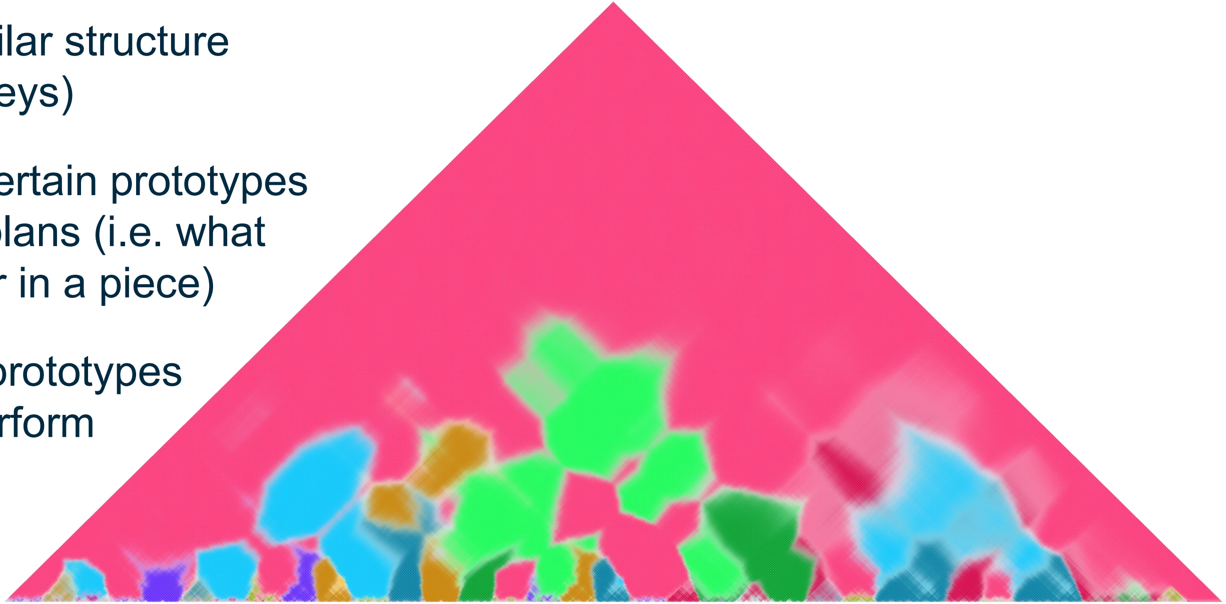
Durham University

Pitch-Scape Clustering

Pitch-Scape Clustering

Basic Idea

- Key-scape plots look very structured
- Different pieces show similar structure (but possibly in different keys)
- Music theory postulates certain prototypes for harmonic modulation plans (i.e. what local keys typically appear in a piece)
- How can we define such prototypes of key-scape plots and perform clustering to find them?



Lieck R, Rohrmeier M (2020) Modelling Hierarchical Key Structure With Pitch Scapes

Pitch-Scape Clustering

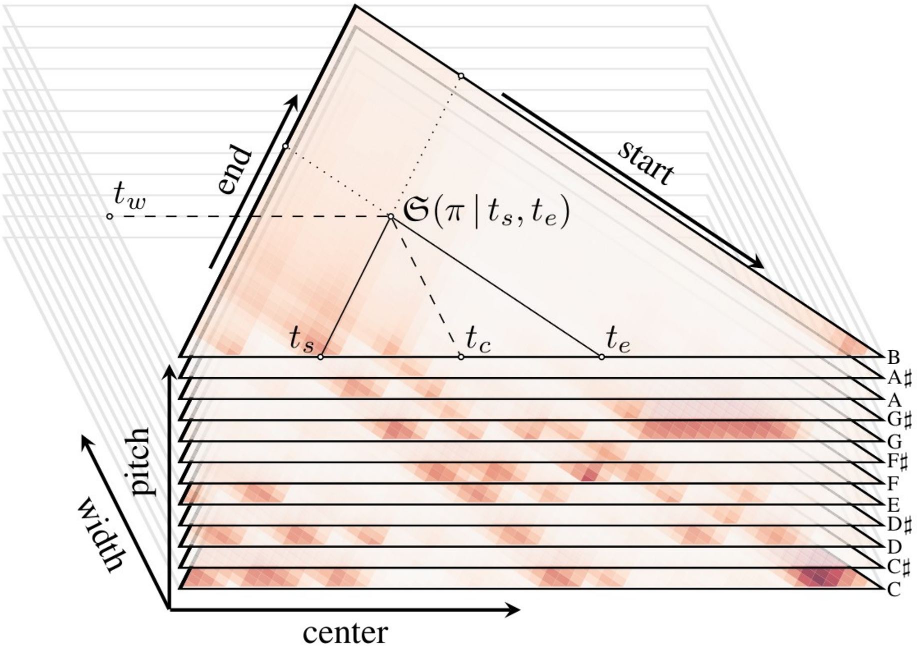
Time-Continuous Pitch-Scape Model

- Define pitch-class distribution for each (continuous) start and end time

$$\sum_{\pi=0}^{11} \mathfrak{S}(\pi \mid t_s, t_e) = 1$$

- Empirical value from actual pitch-class counts

$$\mathfrak{S}(\pi \mid t_s, t_e) := \underbrace{\frac{1}{t_e - t_s + 12c}}_{\textit{normalisation}} \Big[\overbrace{c}^{\textit{prior counts}} + \underbrace{\int_{t_s}^{t_e} \delta(\pi \mid t)\, dt}_{\textit{overall pitch class counts}} \Big]$$



Pitch-Scape Clustering

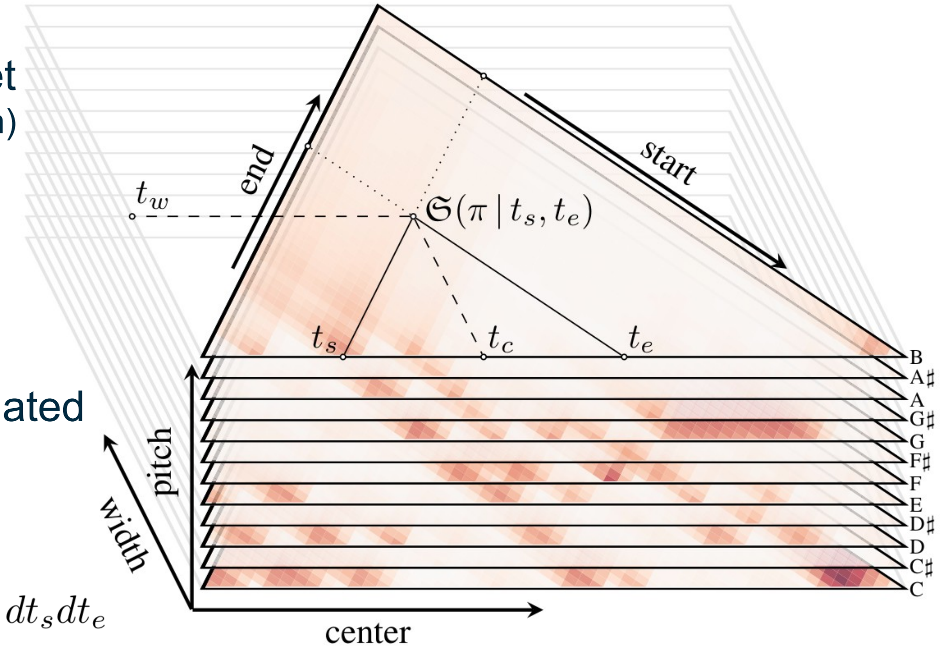
Define Prototypes

- Define function alpha as prior for a Dirichlet
  (conjugate prior of categorical/pitch-class distribution)

$$p(\Pi \mid \alpha, t_s, t_e) = \mathrm{Dir}(\Pi; \alpha(t_s, t_e))$$

- Posterior for given pitch scape is approximated by sum over all points in the scape

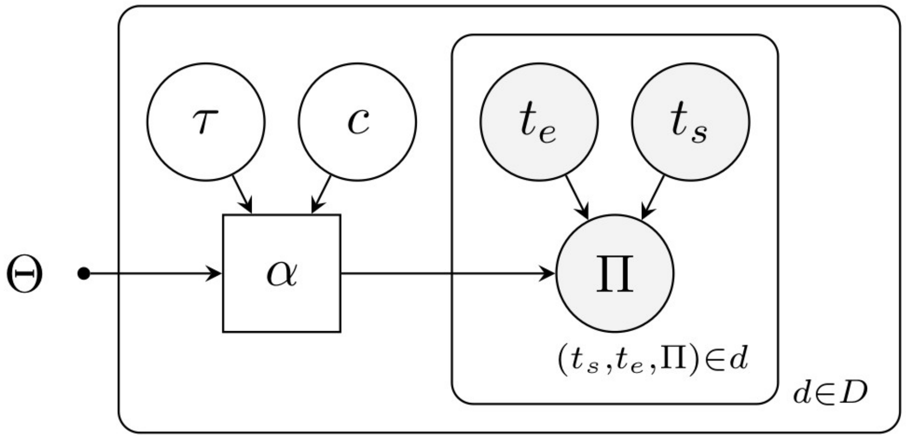
$$\log p(\mathfrak{S} \mid \alpha) = \frac{2}{T^2} \iint_{0 \le t_s < t_e \le T} \log \mathrm{Dir}\big(\mathfrak{S}(t_s, t_e); \alpha(t_s, t_e)\big) \, dt_s dt_e$$

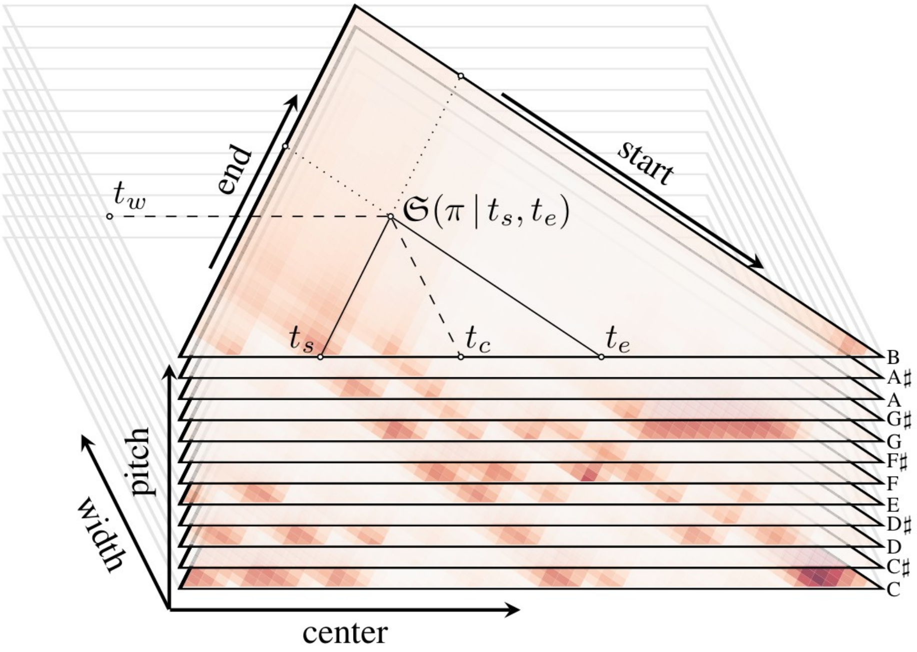


Pitch-Scape Clustering

Training a Mixture Model

- Define a mixture over prototypes and transpositions
- Each prototype has parameters that define alpha
- Perform gradient descent on log-posterior

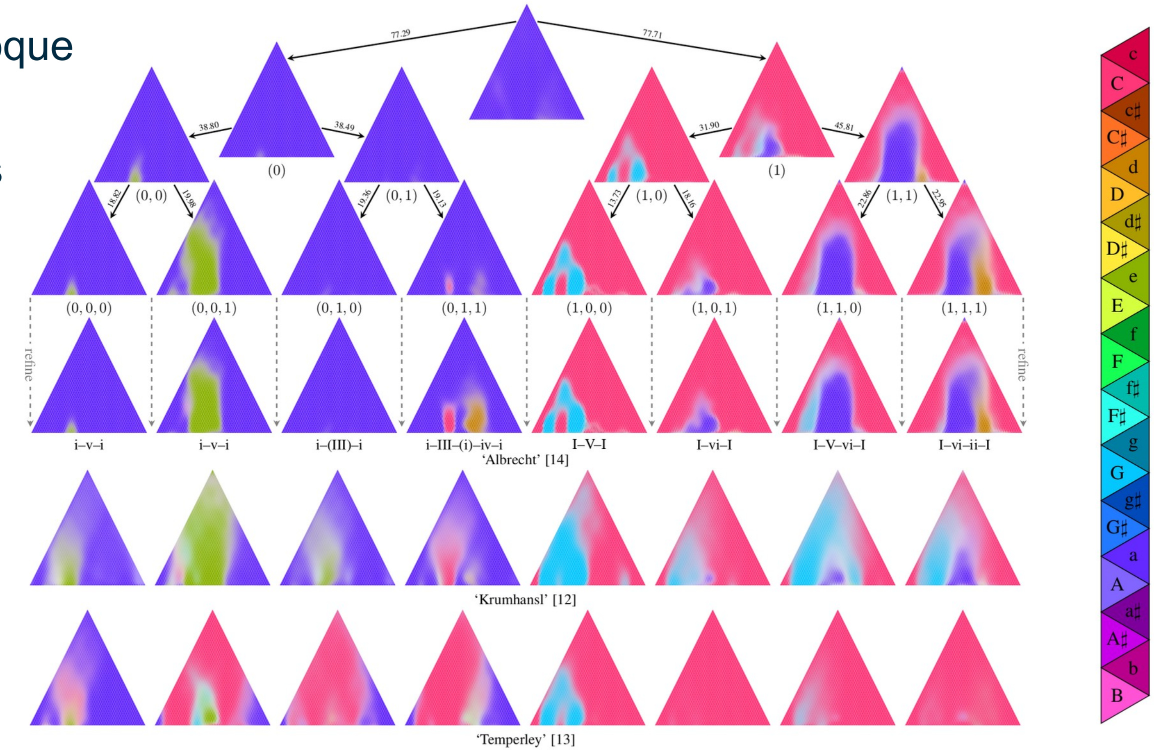




Pitch-Scape Clustering

Prototypical Modulation Plans

- Hierarchical clustering for Baroque corpus
- Split into major and minor keys
- Prototypes major
  - I–V–I
  - I–vi–I
  - I–V–vi–I
  - I–vi–ii–I
- Prototypes minor:
  - i–v–i
  - i–(III)–i
  - i–III–(i)–iv–i

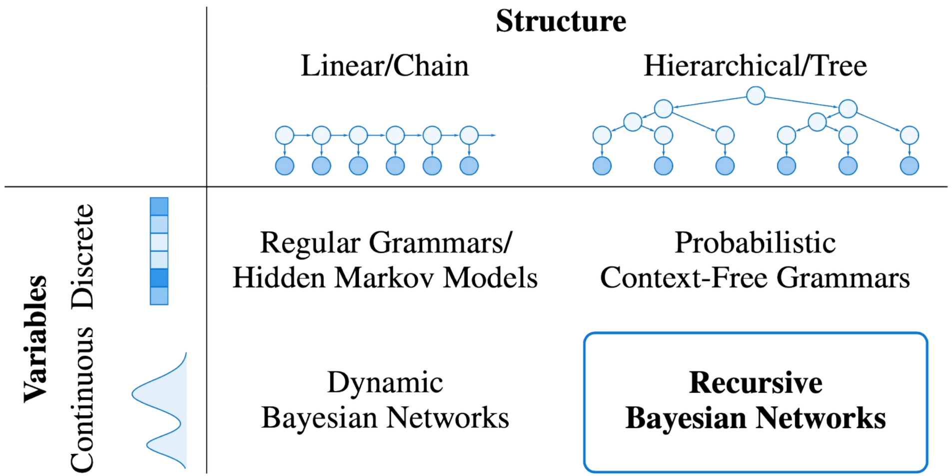


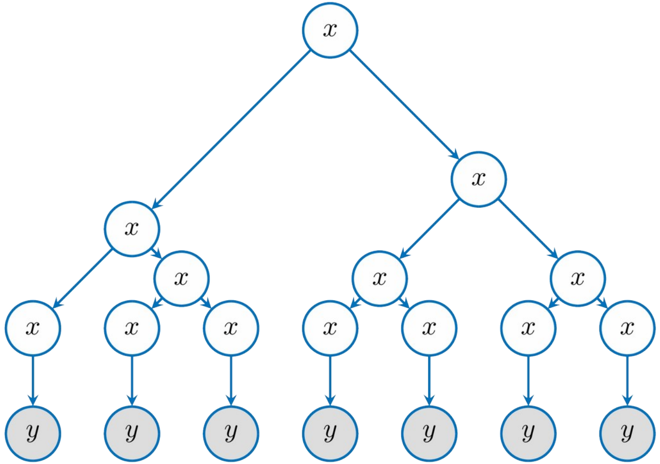
Durham University

Recursive Bayesian Networks

Recursive Bayesian Networks

PCFGs with Continuous Variables





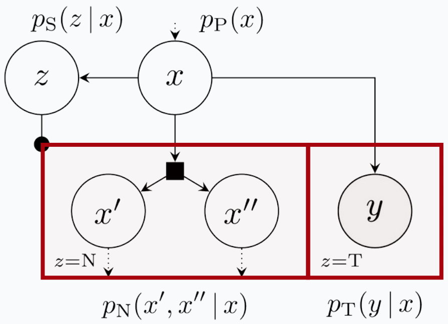
Durham University

- Lieck R, Rohrmeier M (2021) Recursive Bayesian Networks: Generalising and Unifying Probabilistic Context-Free Grammars and Dynamic Bayesian Networks. In: Advances in Neural Information Processing Systems 34 (NeurIPS 2021)

Recursive Bayesian Networks

The generative view

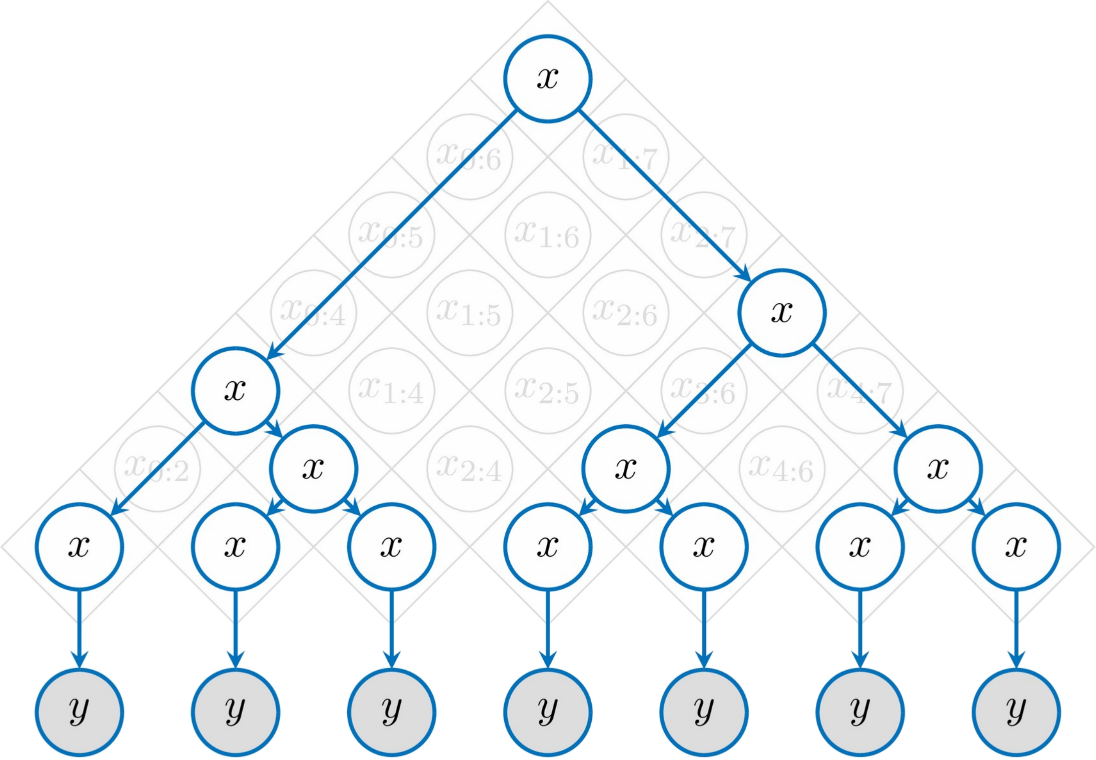
Chomsky Normal Form



Separate:

I. Generating structure

II. Generating variables



Durham University

Recursive Bayesian Networks

Inference in RBNs

- **Two Challenges**
  I. Unknown structure
  II. Unknown variables

Exponential No. Structures

Catalan numbers ($C_{100} \approx 10^{57}$)

(1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, …)

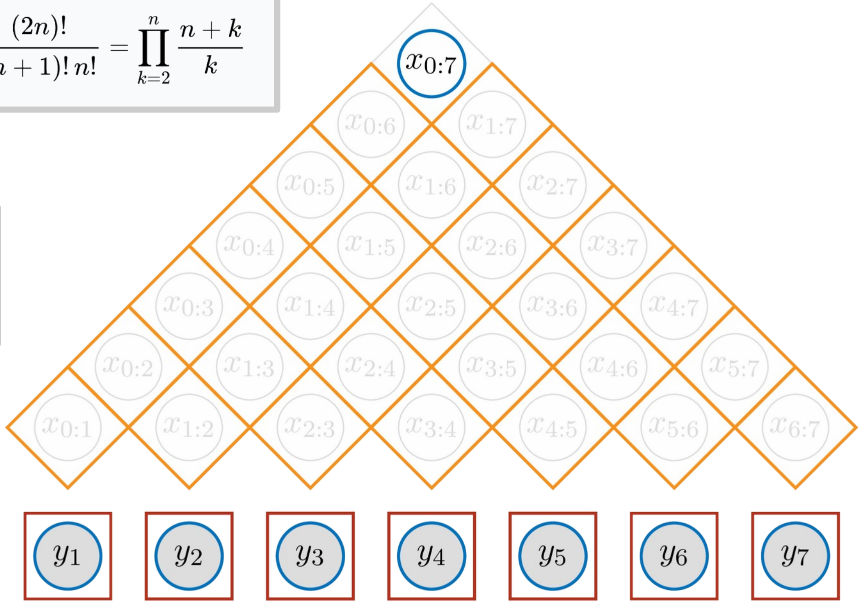
$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{(n+1)!\,n!} = \prod_{k=2}^{n} \frac{n+k}{k}$$

- **Goal is to compute**
  - Marginal data likelihood

$$p(\mathbf{Y}) = \int \beta_{0:n}(x_{0:n})\, p_{\mathrm{P}}(x_{0:n})\, dx_{0:n}$$

  - Inside probabilities

$$\beta_{i:k}(x_{i:k}) := p(\mathbf{Y}_{i:k} \mid x_{i:k})$$



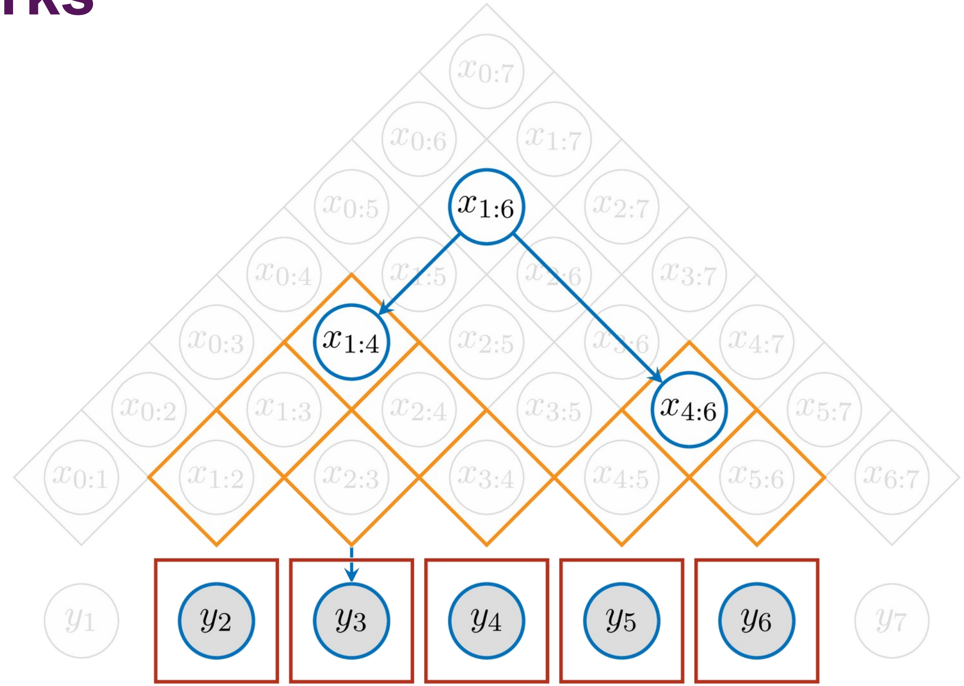
Durham University

Recursive Bayesian Networks

Inference in RBNs

- **Compute recursively**
  - dynamic programming
- **Separate marginalisation**
  - over structure (Chlg. I)
  - over variables (Chlg. II)
- **Complexity:** $\mathcal{O}(n^3\,k)$

  (n: sequence length, k: size of alphabet)



$$\beta_{i:i+1}(x_{i:i+1}) = p_{\mathrm{S}}(z_{i:i+1}{=}\mathrm{T} \mid x_{i:i+1})\, p_{\mathrm{T}}(y_{i+1} \mid x_{i:i+1})$$

$$\beta_{i:k}(x_{i:k}) = p_{\mathrm{S}}(z_{i:k}{=}\mathrm{N} \mid x_{i:k}) \sum_{j=i+1}^{k-1} \iint p_{\mathrm{N}}(x_{i:j}, x_{j:k} \mid x_{i:k})\, \beta_{i:j}(x_{i:j})\, \beta_{j:k}(x_{j:k})\, dx_{i:j}\, dx_{j:k}$$

Durham University

Recursive Bayesian Networks

Dealing with continuous variables

Analytic Solution

- Solve integrals in closed form
- Requires (self) conjugate distributions
- Only possible with linear Gaussians

Analytic Approximation

- Choose parameterised form that affords closed-form solution
- Fit parameters to true distributions
- Variational inference

Particle Approximation

- Represent distributions by finite samples of ”particles”
- Replace marginalisation by averaging over particles
- Importance sampling & Markov chain Monte Carlo

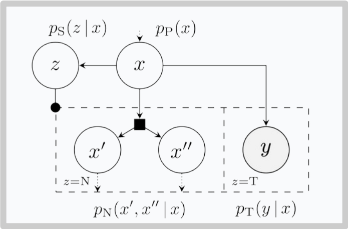
Recursive Bayesian Networks

Gaussian RBNs

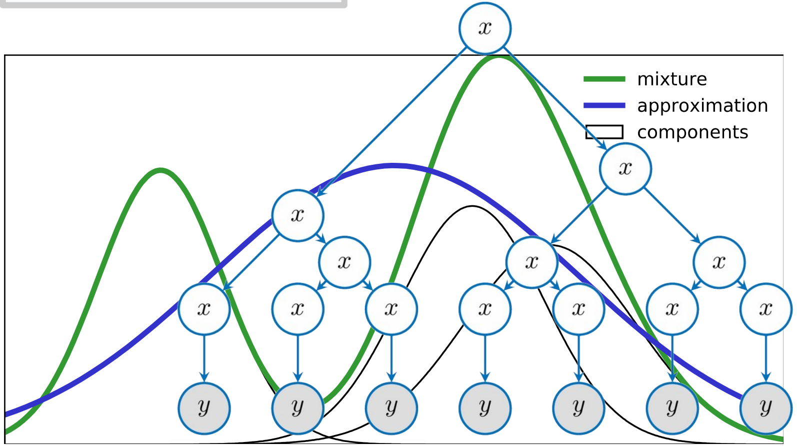
- **Only Gaussian distributions**
  - **exp. many structures**
  - **exp. many mixture components**

- **Solution**
  - Adaptive approximation

$$\beta_{i:k}(x_{i:k}) \approx c_{i:k}\,\mathcal{N}(x_{i:k}, \mu_{i:k}, \Sigma_{i:k})$$



$$p_P(x) := \mathcal{N}(x, \mu_P, \Sigma_P)$$
$$p_N(x', x'' \mid x) := \mathcal{N}(x', x, \Sigma_{NL})\mathcal{N}(x'', x, \Sigma_{NR})$$
$$p_T(y \mid x) := \mathcal{N}(y, x, \Sigma_T)$$
$$p_S(z{=}T \mid x) := p_{\text{term}}\ .$$



$$\beta_{i:k}(x_{i:k}) = p_S(z_{i:k}{=}N \mid x_{i:k}) \sum_{j=i+1}^{k-1} \iint p_N(x_{i:j}, x_{j:k} \mid x_{i:k})\, \beta_{i:j}(x_{i:j})\, \beta_{j:k}(x_{j:k})\, dx_{i:j}\, dx_{j:k}$$

References

1) Sapp CS (2001) Harmonic Visualizations of Tonal Music. In: Proc. International Computer Music Conference (ICMC). Havana, Cuba

2) Sapp CS (2005) Visual hierarchical key analysis. Computers in Entertainment 3:3. https://doi.org/10.1145/1095534.1095544

3) Sapp CS (2007) Comparative Analysis of Multiple Musical Performances. In: ISMIR. pp 497–500

4) Müller M, Jiang N (2012) A Scape Plot Representation for Visualizing Repetitive Structures of Music Recordings. In: ISMIR. Citeseer, pp 97–102

5) Park S, Kwon T, Lee J, et al (2019) A Cross-Scape Plot Representation for Visualizing Symbolic Melodic Similarity. In: Flexer A, Peeters G, Urbano J, Volk A (eds) Proceedings of the 20th International Society for Music Information Retrieval Conference, ISMIR 2019, Delft, The Netherlands, November 4-8, 2019. pp 423–430

6) Lieck R, Rohrmeier M (2020) Modelling Hierarchical Key Structure With Pitch Scapes. In: Proceedings of the 21st International Society for Music Information Retrieval Conference. Montréal, Canada, pp 811–818

7) Lieck R, Rohrmeier M (2021) Recursive Bayesian Networks: Generalising and Unifying Probabilistic Context-Free Grammars and Dynamic Bayesian Networks. In: Advances in Neural Information Processing Systems 34 (NeurIPS 2021)

8) Viaccoz C, Harasim D, Moss FC, Rohrmeier M (2022) Wavescapes: A visual hierarchical analysis of tonality using the discrete Fourier transform. Musicae Scientiae 10298649211034906