Durham University

Introduction to Music Computing

Sequential Models

Chord Recognition and HMMs

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

Figures from

Meinard Müller, Fundamentals of Music Processing, Springer 2015

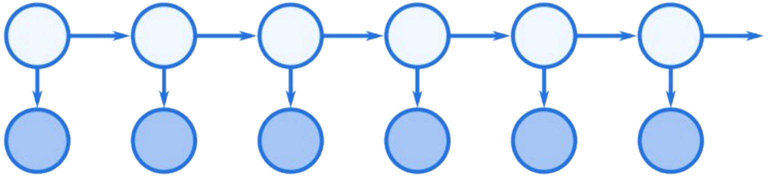
Chapter 5: Chord Recognition

International Audio Laboratories Erlangen

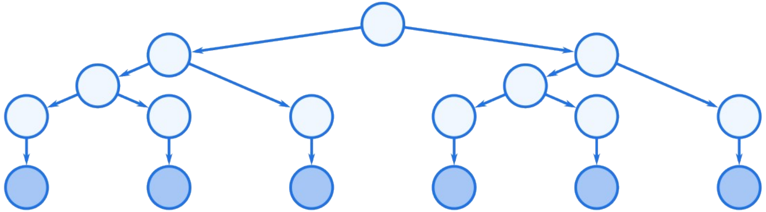
www.music-processing.de

Computational Models of Music

- **Sequential Models** (*n*-gram and (hidden) Markov models)



- **Hierarchical Models** (context-free grammars)



- **Neural Networks** (RNNs, Transformers, WaveNet)

Chord Recognition

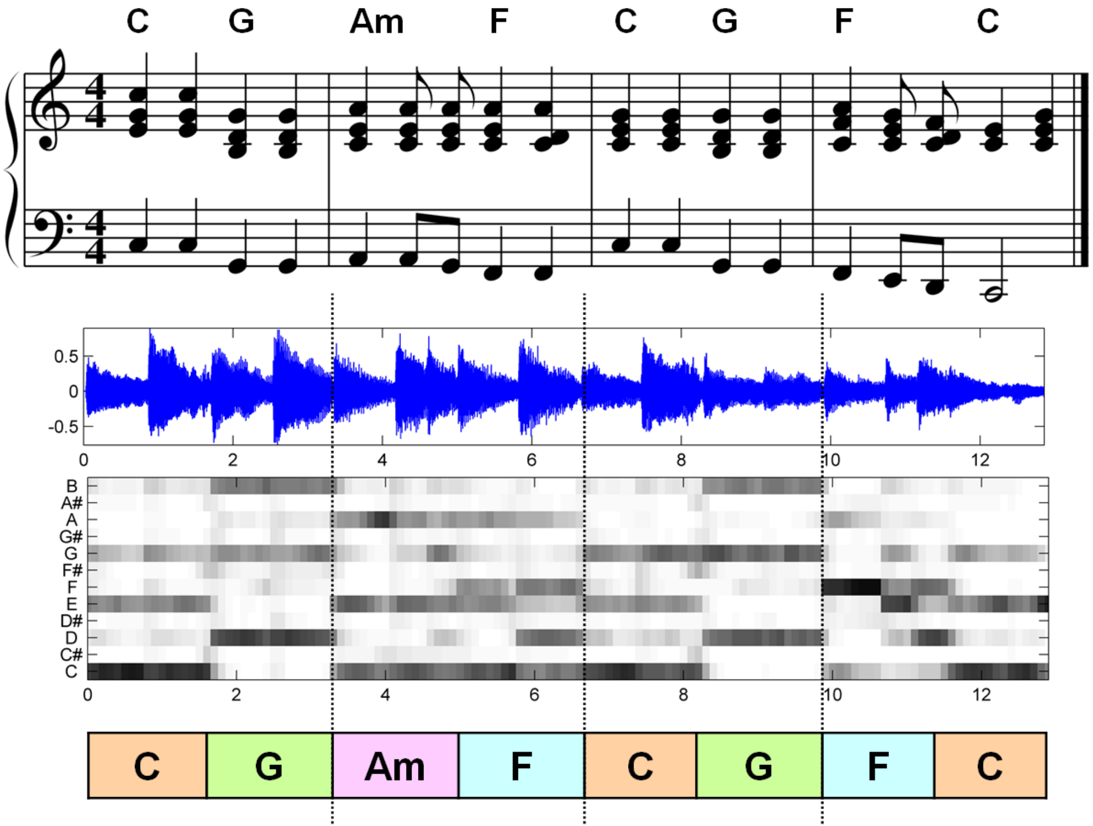


Fig. 5.1

Recap: Basic Theory of Harmony

Tuning Systems

Overtones
(Fifths, Thirds)
(Fifths)

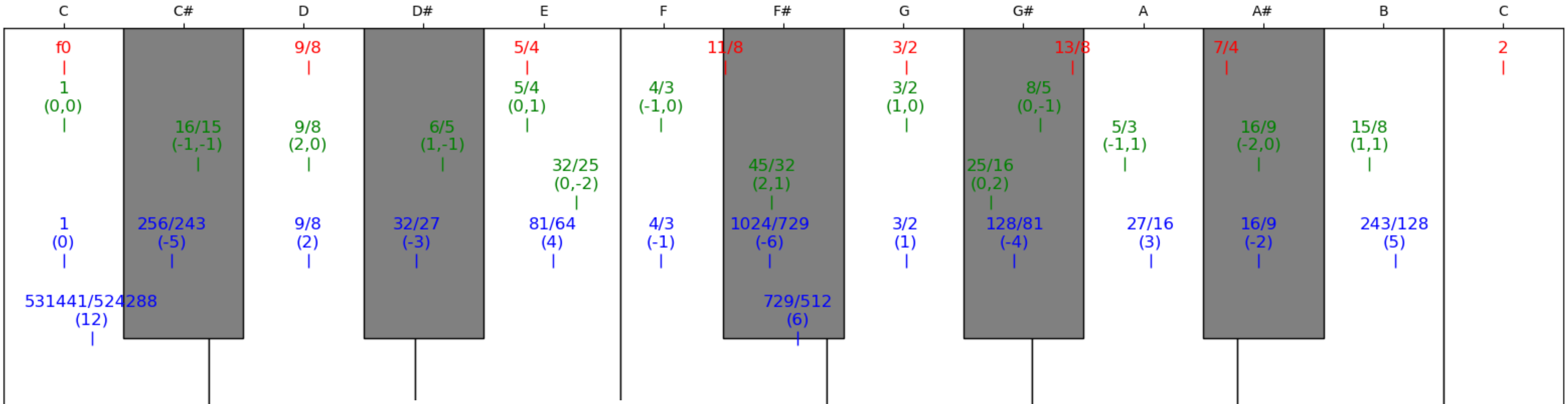


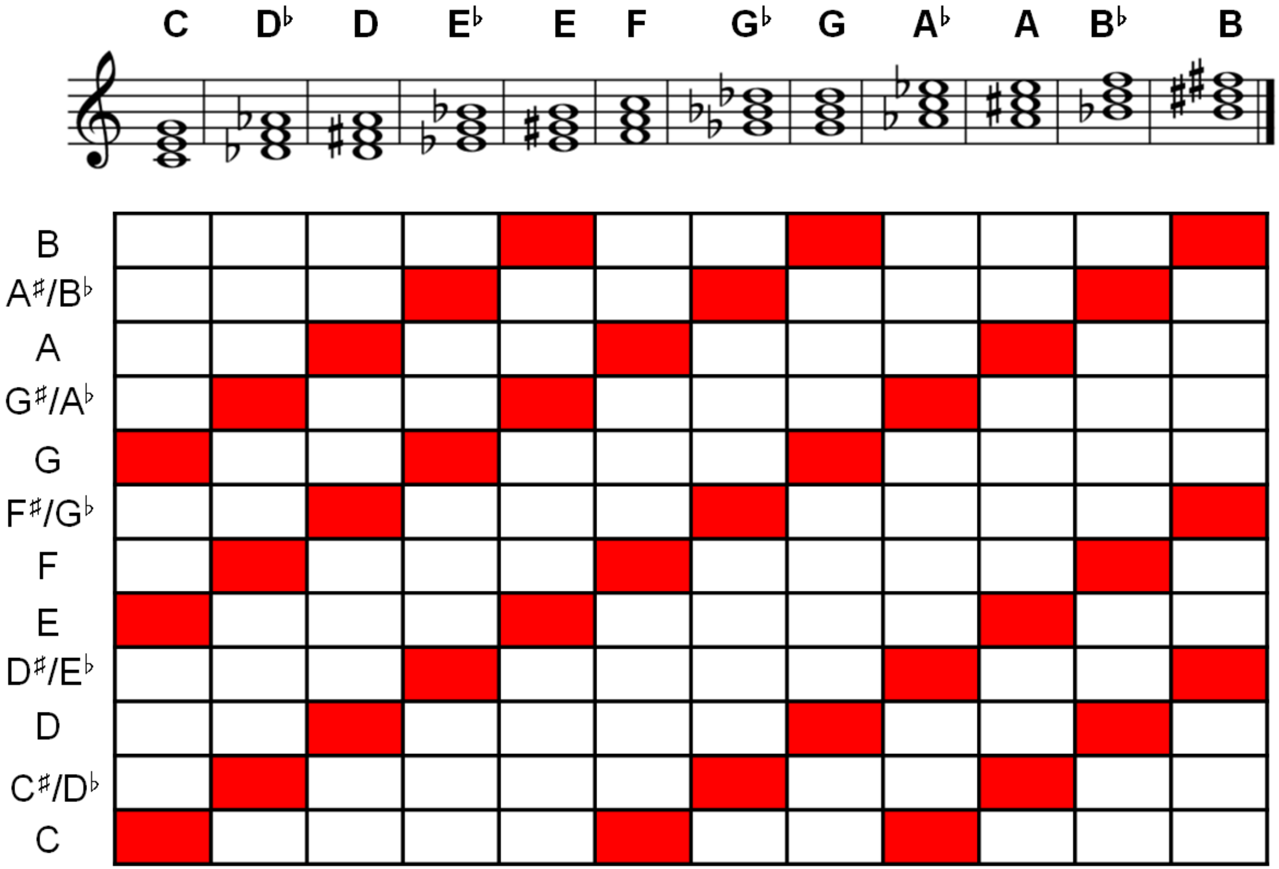
Fig 5.3 [FMP15]

Δ	Interval name	Interval	JI ratio	Pyt. ratio
0	(Perfect) unison	C4 – C4	1:1	1:1
1	Minor second	C4 – D♭4	15:16	$3^5:2^8$
2	Major second	C4 – D4	8:9	$2^3:3^2$
3	Minor third	C4 – E♭4	5:6	$3^3:2^5$
4	Major third	C4 – E4	4:5	$2^6:3^4$
5	(Perfect) fourth	C4 – F4	3:4	$3:2^2$
6	Tritone	C4 – F♯4	32:45	$2^9:3^6$ or $3^6:2^{10}$
7	(Perfect) fifth	C4 – G4	2:3	2:3
8	Minor sixth	C4 – A♭4	5:8	$3^4:2^7$
9	Major sixth	C4 – A4	3:5	$2^4:3^3$
10	Minor seventh	C4 – B♭4	5:9	$3^2:2^4$
11	Major seventh	C4 – B4	8:15	$2^7:3^5$
12	(Perfect) octave	C4 – C5	1:2	1:2

Durham University

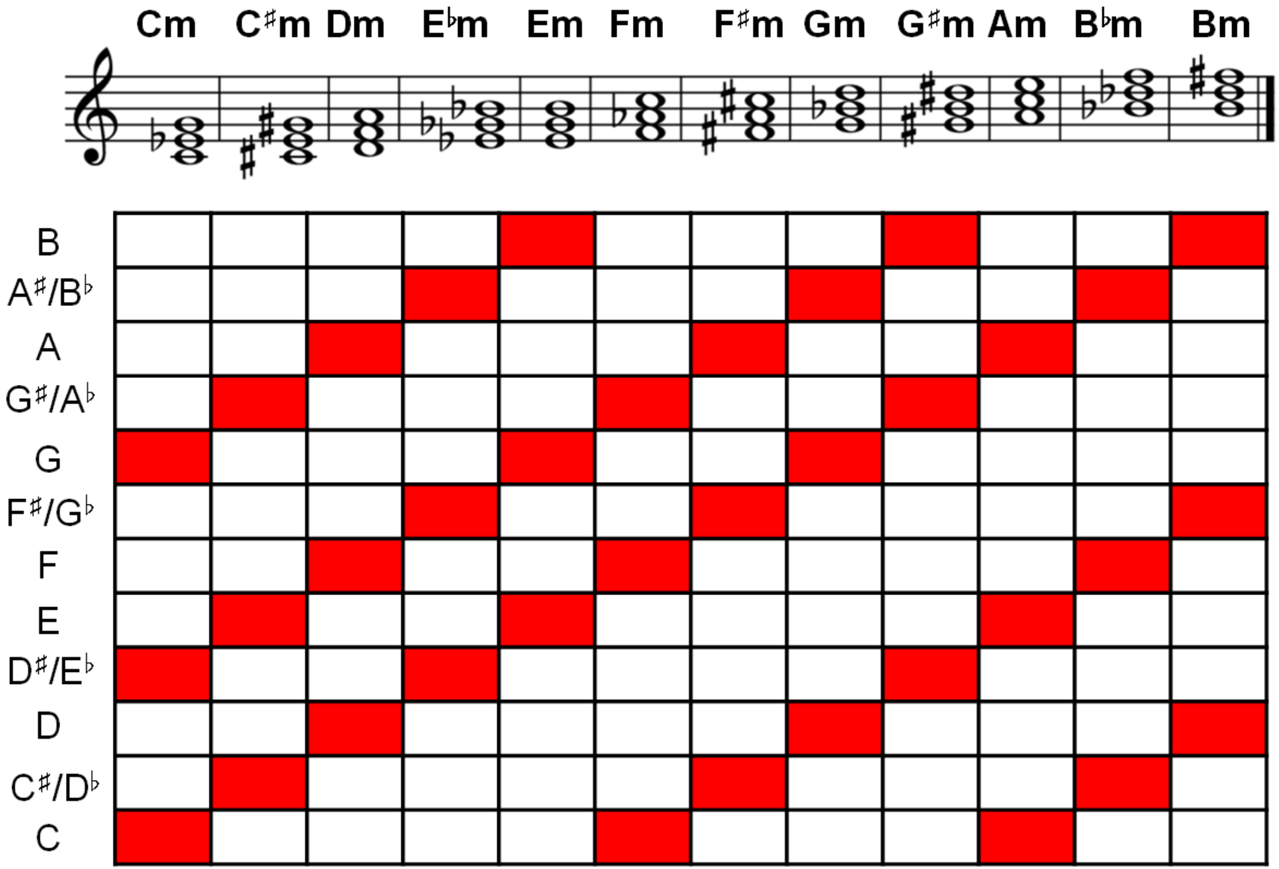
Basic Theory of Harmony

Fig. 5.6



Basic Theory of Harmony

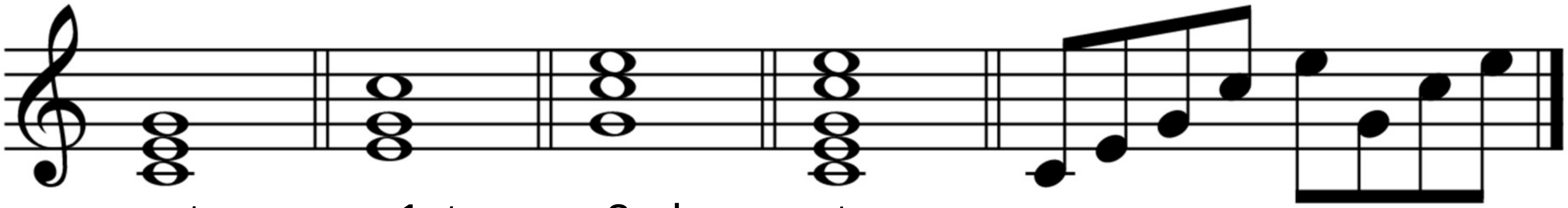
Fig. 5.6



Durham University

Basic Theory of Harmony

Fig. 5.7



root position | 1st inversion | 2nd inversion | octave doubling | arpeggiated

This is all a C major chord!

Durham University

Basic Theory of Harmony

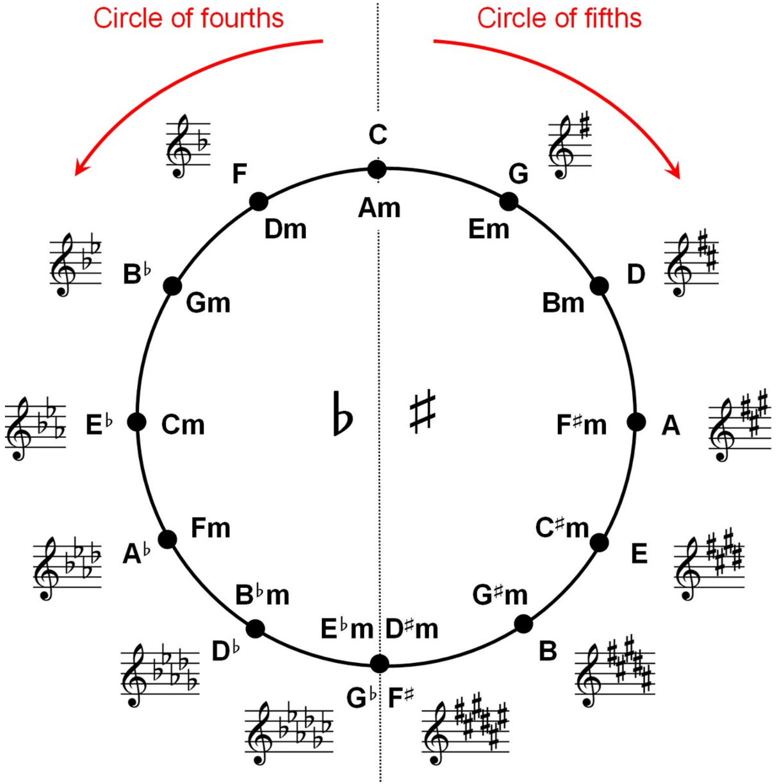


Fig. 5.10

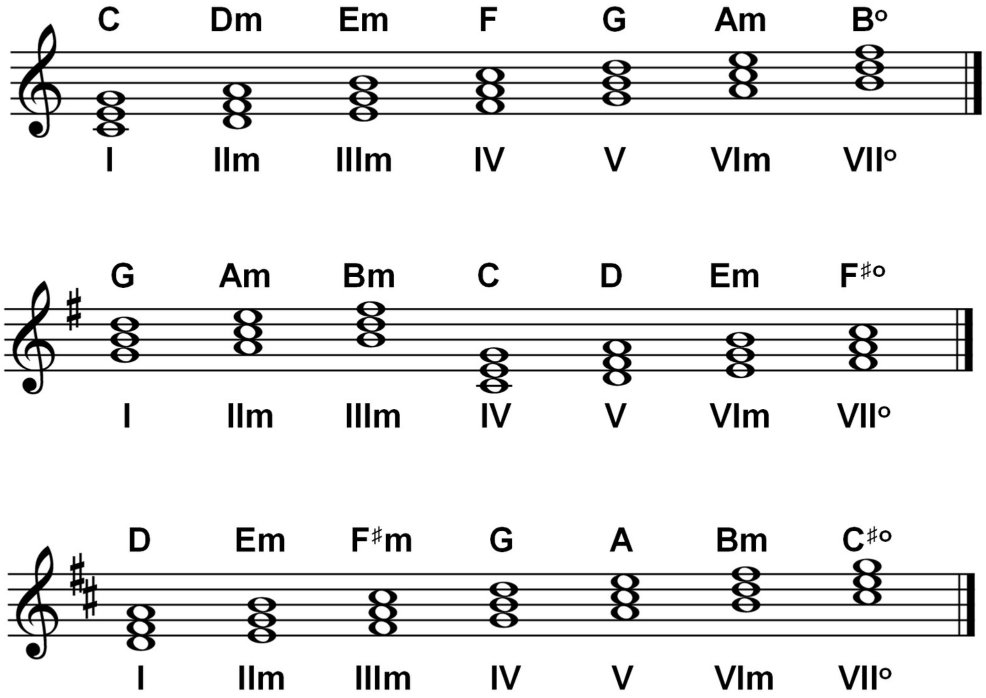
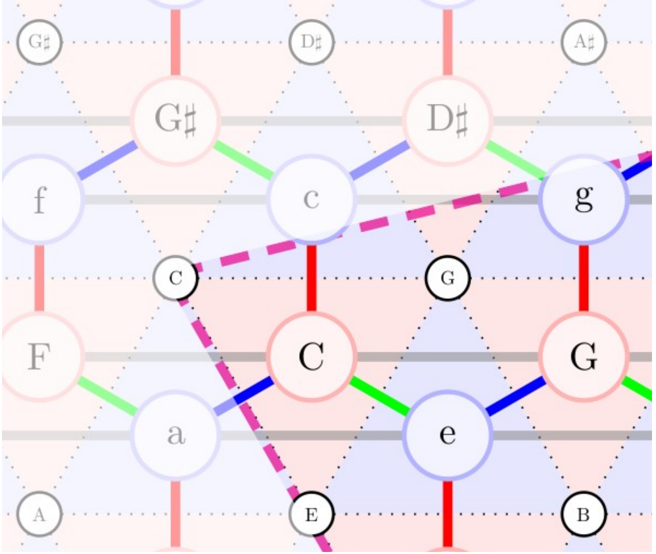


Fig. 5.11

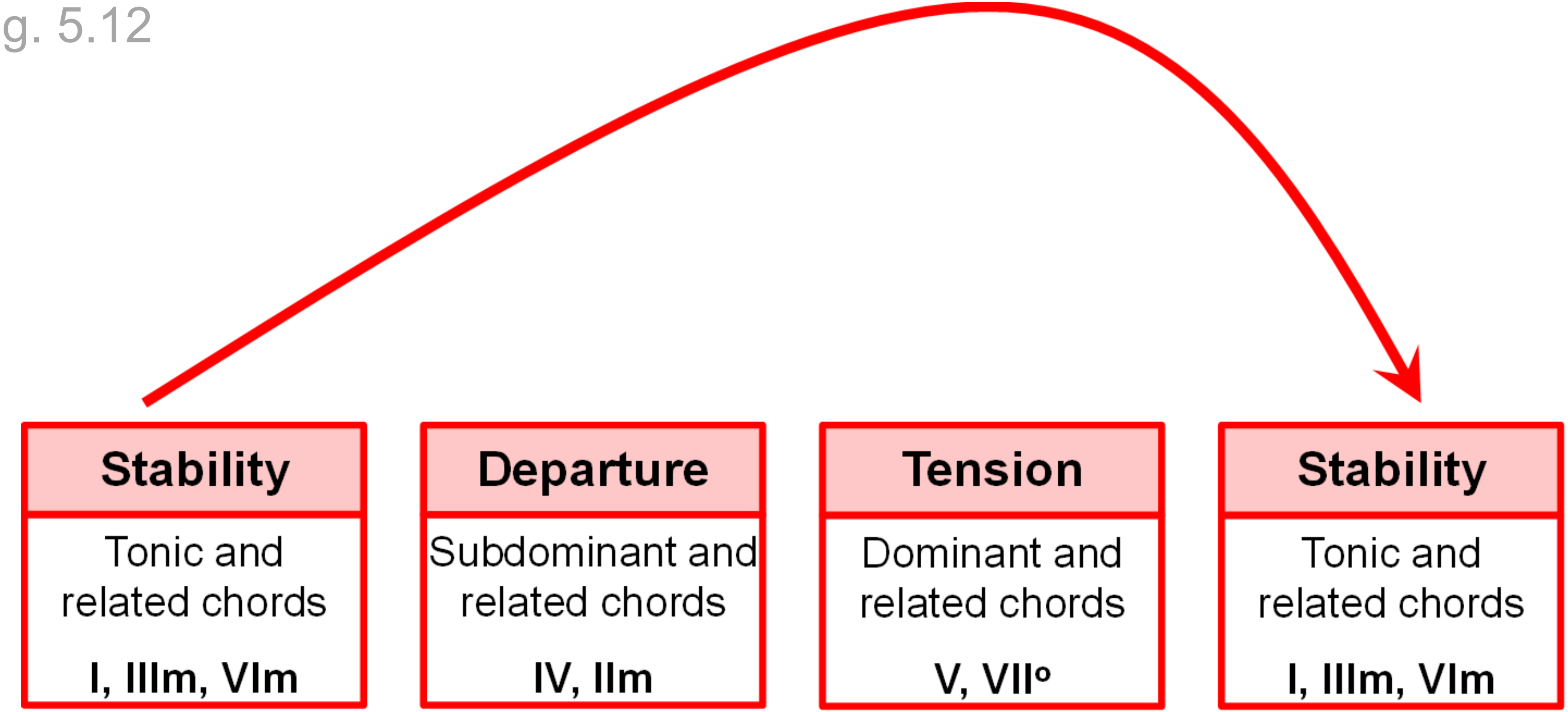
Tonnetz

- Two versions (equivalent/dual)
  - **pitch-based:** pitches as hexagonal faces (triads at corners)
  - **triad-based:** triads as triangular faces (pitches at corners)
- Neo-Riemannian operations with minimal voice-leading
  - <u>R</u>elative major/minor
    C major (C, E, <u>G</u>) ↔ a minor (<u>A</u>, C, E)
  - <u>P</u>arallel major/minor
    C major (C, <u>E</u>, G) ↔ c minor (C, <u>Eb</u>, G)
  - <u>L</u>eading-Tone Exchange
    C major (C, <u>E</u>, G) ↔ e minor (E, G, <u>B</u>)



Basic Theory of Harmony

Fig. 5.12

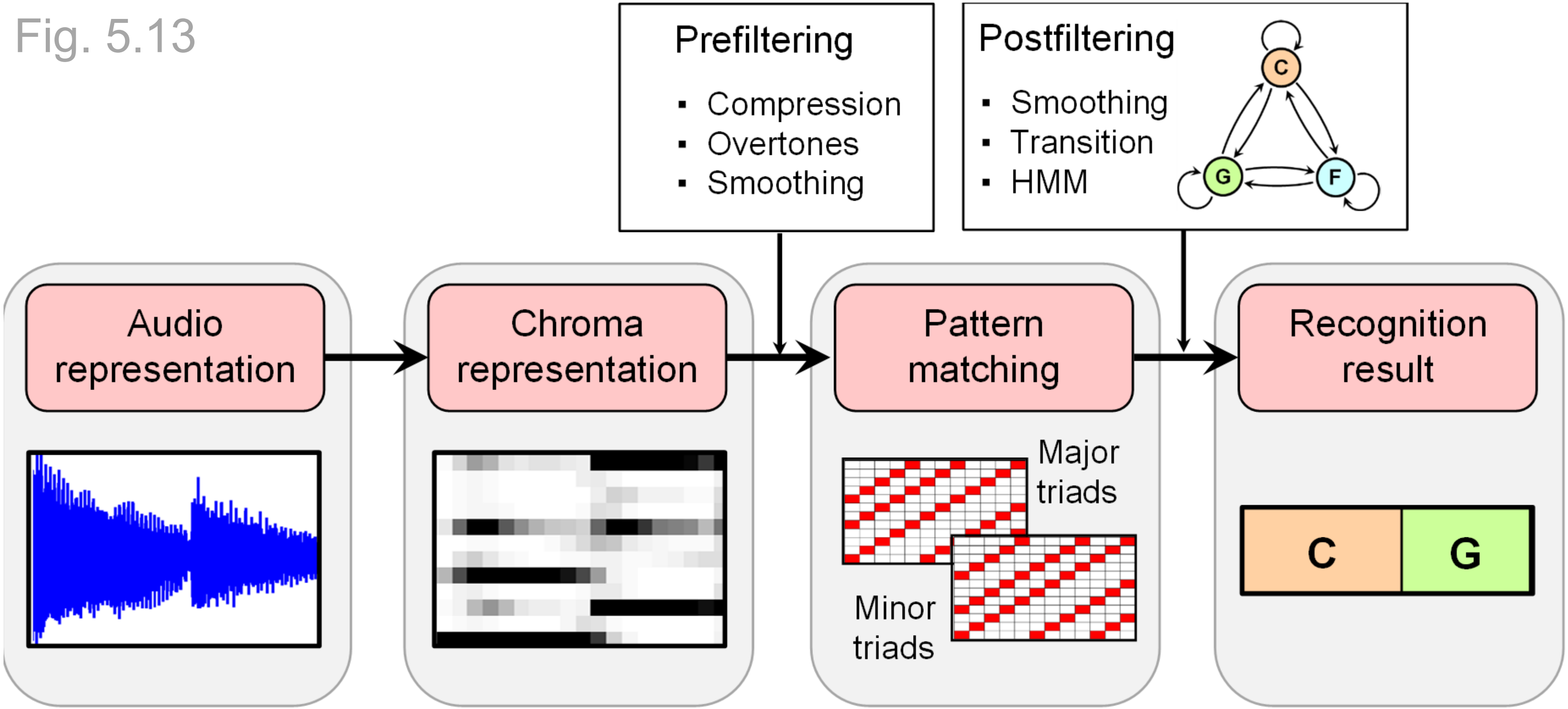


Durham University

Template-Based Chord Recognition

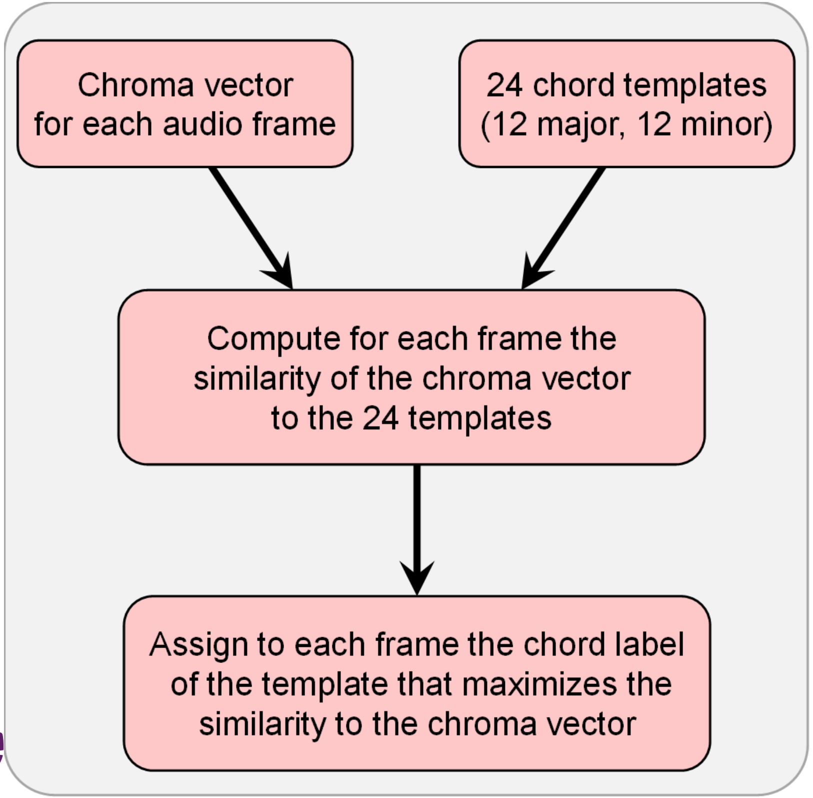
Template-Based Chord Recognition

Fig. 5.13



Template-Based Chord Recognition

Fig. 5.14



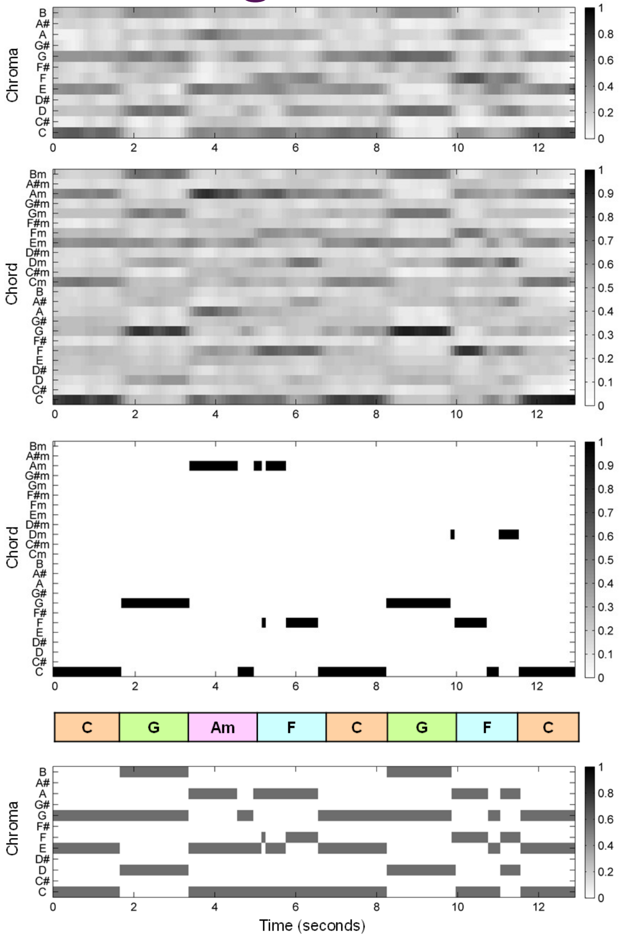
	C	C$^\sharp$	D	…	Cm	C$^\sharp$m	Dm	…
B	0	0	0	…	0	0	0	…
A$^\sharp$	0	0	0	…	0	0	0	…
A	0	0	1	…	0	0	1	…
G$^\sharp$	0	1	0	…	0	1	0	…
G	1	0	0	…	1	0	0	…
F$^\sharp$	0	0	1	…	0	0	0	…
F	0	1	0	…	0	0	1	…
E	1	0	0	…	0	1	0	…
D$^\sharp$	0	0	0	…	1	0	0	…
D	0	0	1	…	0	0	1	…
C$^\sharp$	0	1	0	…	0	1	0	…
C	1	0	0	…	1	0	0	…

Template-Based Chord Recognition

Fig. 5.15

Template-based chord recognition using binary templates for the 24 major and minor chords.

The audio recording consists of the first four measures of the Beatles song “Let It Be”.

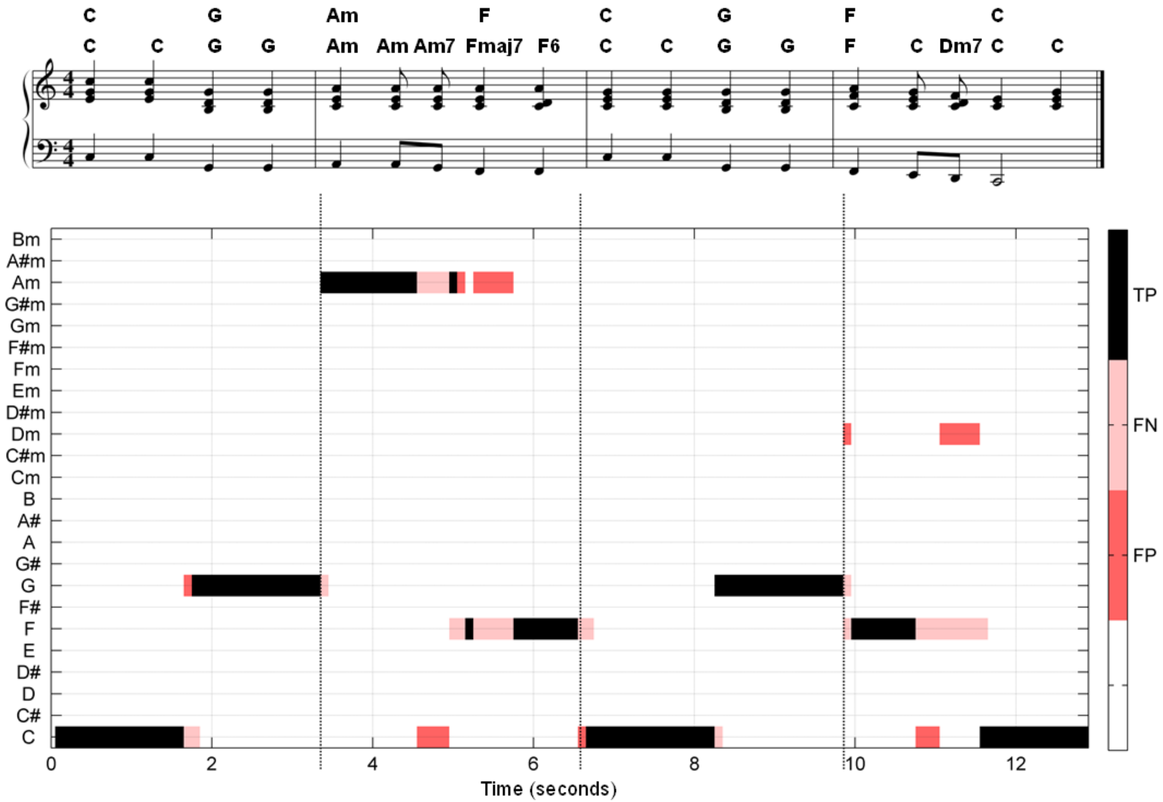


- **Chroma representation**
- **Similarity values between the chroma vectors and the 24 chord templates**
- **Chord recognition result**
- **Manually specified chord annotations**
- **Normalized binary templates of the chord recognition result.**

Durham University

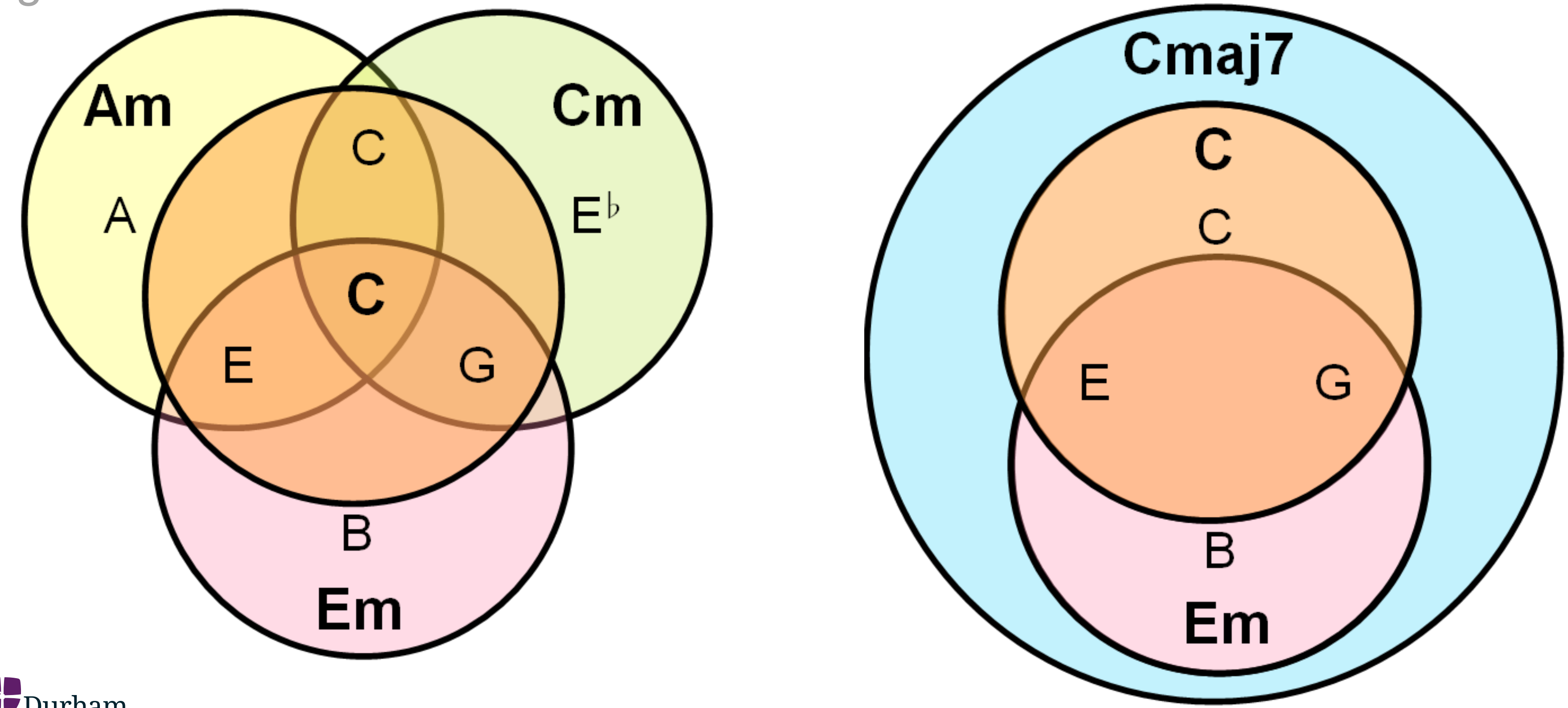
Template-Based Chord Recognition

Fig. 5.16



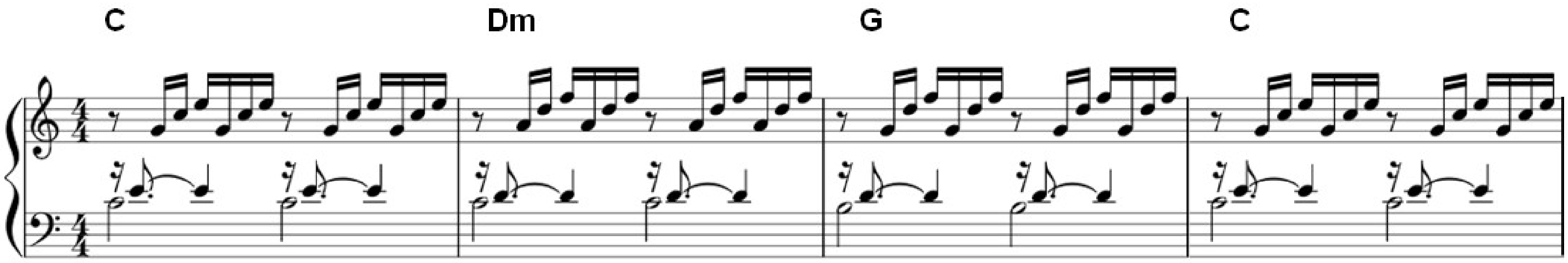
Template-Based Chord Recognition

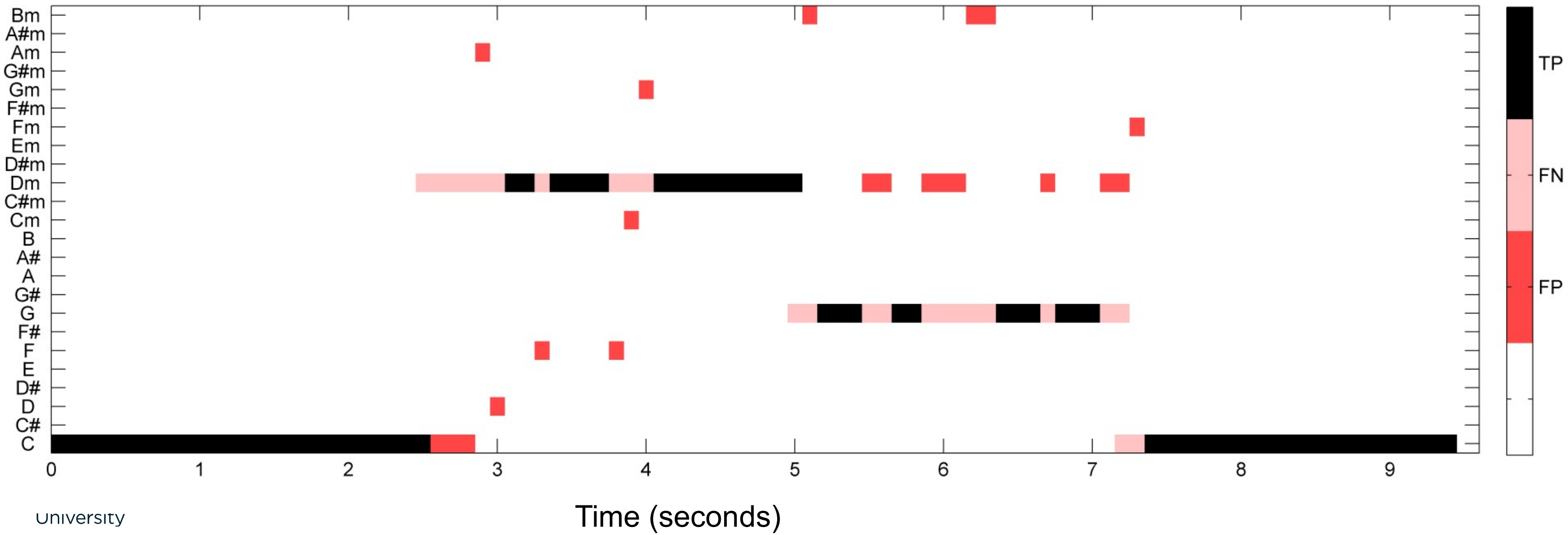
Fig. 5.17



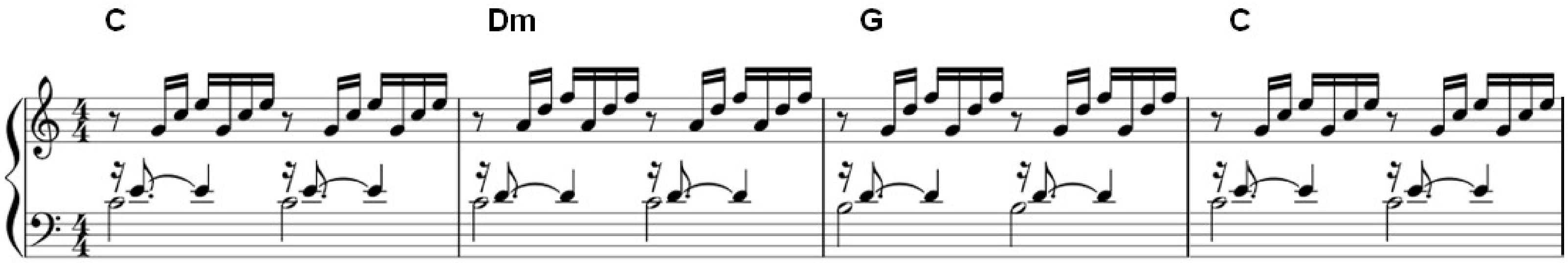
Template-Based Chord Recognition

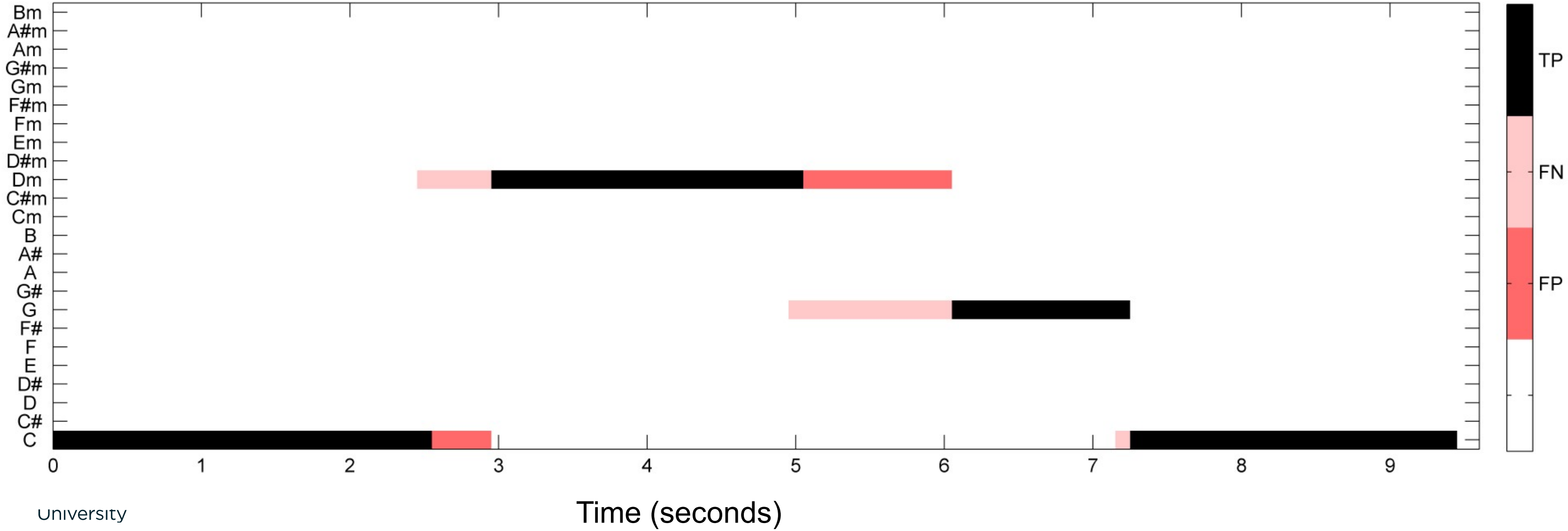
Fig. 5.20





Template-Based Chord Recognition

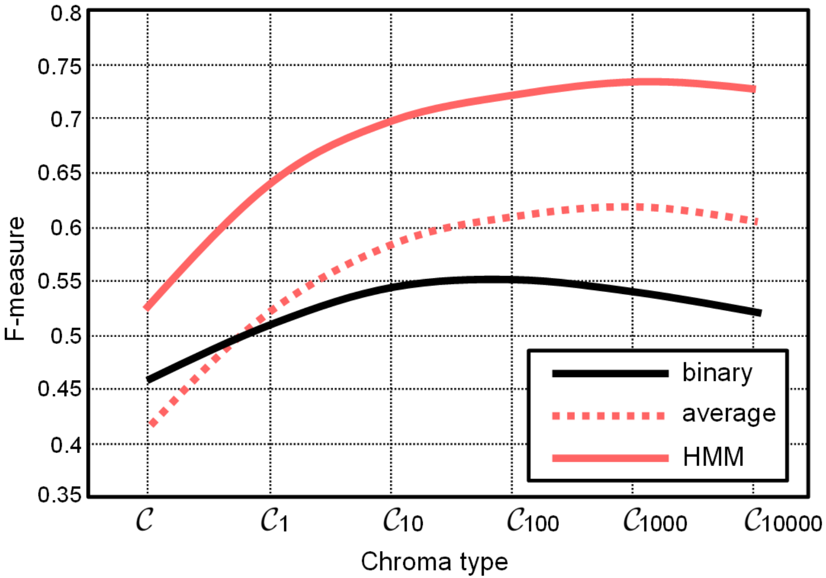




with prefiltering

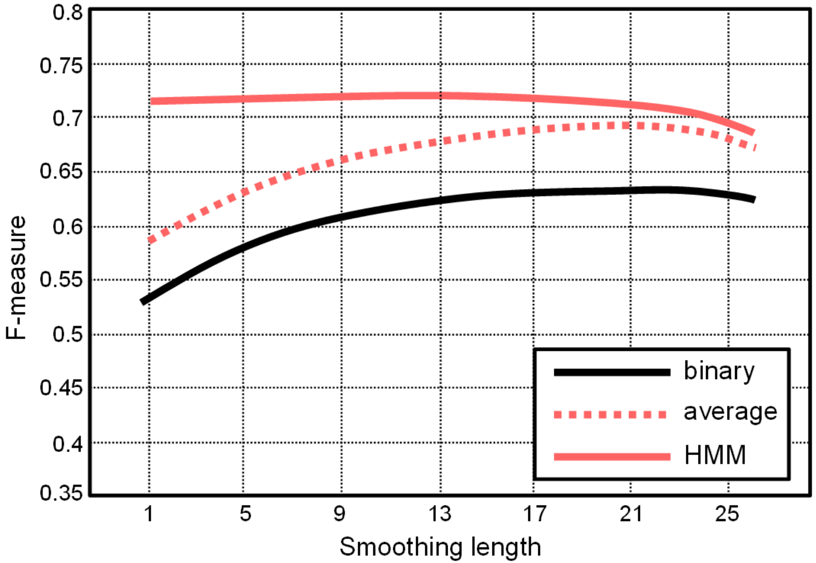
Template-Based Chord Recognition

Fig. 5.22



Logarithmic compression

Fig. 5.23

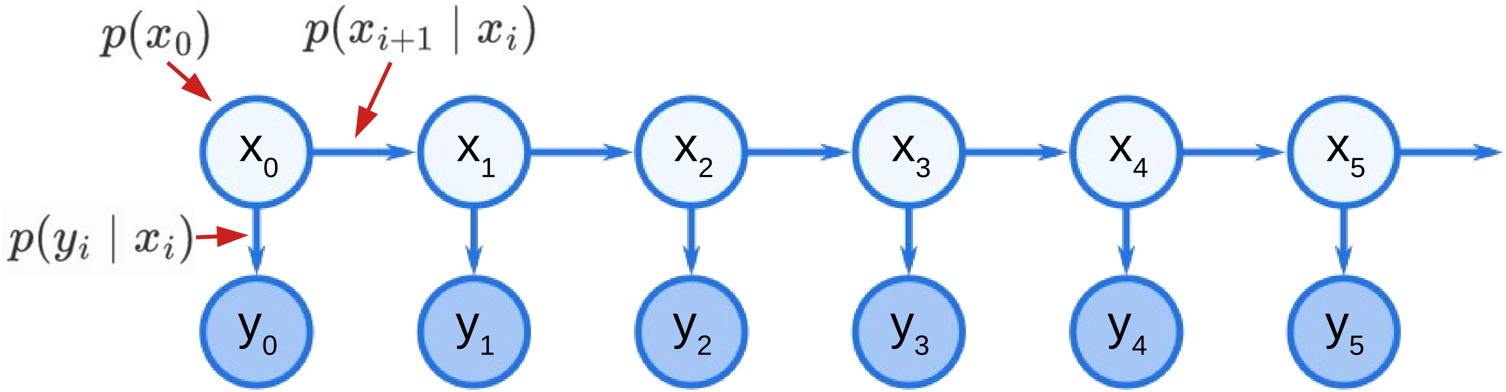


Prefiltering by smoothing

Durham University

HMM-Based Chord Recognition

Hidden Markov Models (HMMs)



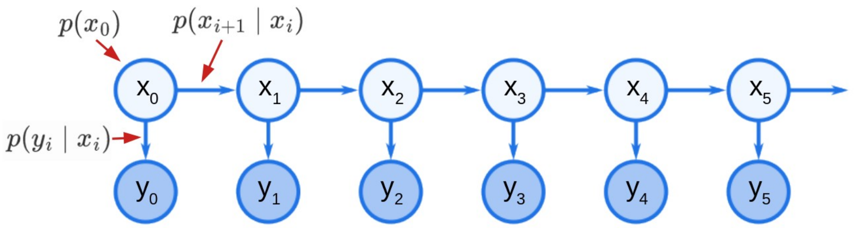
$$p(x_0) := \text{prior/starting distribution}$$
$$p(x_{i+1} \mid x_i) := \text{transition probabilities}$$
$$p(y_i \mid x_i) := \text{observation probabilities}$$

$x \in \mathcal{X}$ discrete e.g. $\{0, 1, \ldots, k\}$

$y \in \mathcal{Y}$ discrete or continuoue

Durham University

HMM-Based Chord Recognition



$x \in \mathcal{X} = \{C, F, G\}$

$y \in \mathcal{Y} = \mathbb{R}^{12}$

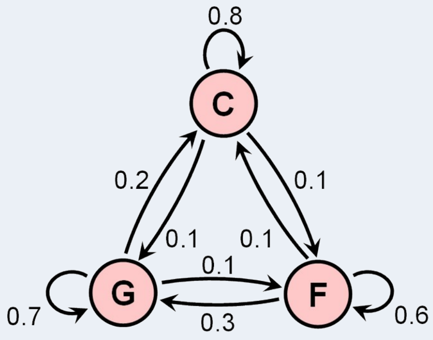


Fig. 5.24

$p(x_{i+1} \mid x_i)$

A	α_1	α_2	α_3
α_1	0.8	0.1	0.1
α_2	0.2	0.7	0.1
α_3	0.1	0.3	0.6

$p(x_0)$

C	α_1	α_2	α_3
	0.6	0.2	0.2

$p(y_i \mid x_i)$

B	β_1	β_2	β_3
α_1	0.7	0	0.3
α_2	0.1	0.9	0
α_3	0	0.2	0.8

Fig. 5.28

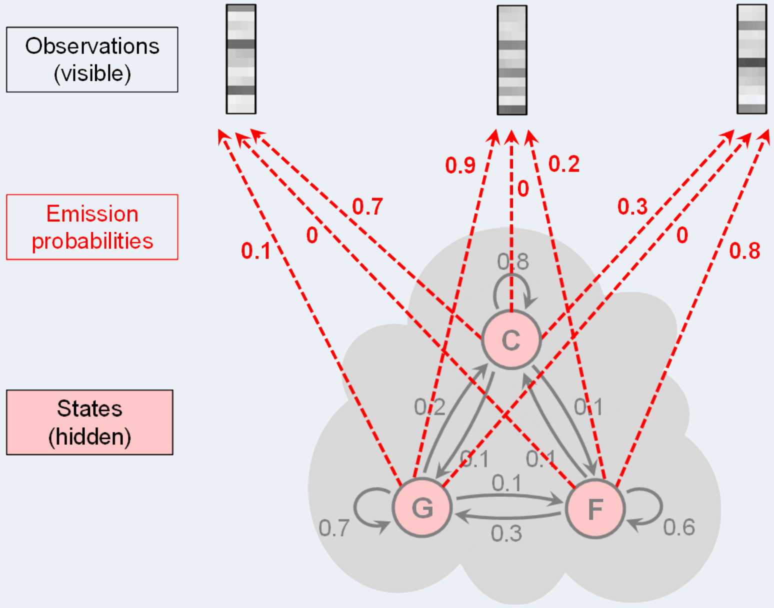
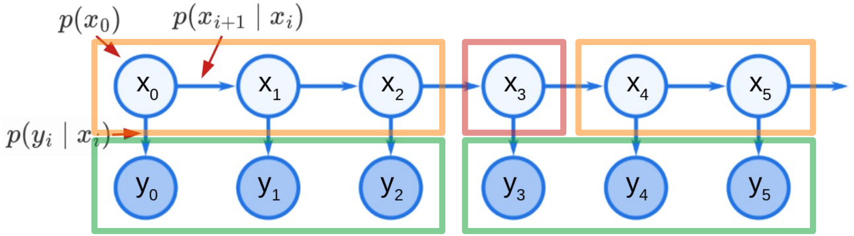


Fig. 5.25

Durham University

HMM-Based Chord Recognition

Forward-Backward Algorithm



forward

$$\alpha(x_i) := p(y_0, \ldots, y_{i-1}, x_i)$$
$$= \sum_{x_{i-1}} p(x_i \mid x_{i-1})\, p(y_{i-1} \mid x_{i-1})\, p(y_0, \ldots, y_{i-2}, x_{i-1})$$
$$= \sum_{x_{i-1}} p(x_i \mid x_{i-1})\, p(y_{i-1} \mid x_{i-1})\, \alpha(x_{i-1})$$
$$\alpha(x_0) = p(x_0)$$

backward

$$\beta(x_i) := p(y_i, y_{i+1}, \ldots, y_n \mid x_i)$$
$$= \sum_{x_{i+1}} p(x_{i+1} \mid x_i)\, p(y_i \mid x_i)\, p(y_{i+1}, \ldots, y_n \mid x_{i+1})$$
$$= \sum_{x_{i+1}} p(x_{i+1} \mid x_i)\, p(y_i \mid x_i)\, \beta(x_{i+1})$$
$$\beta(x_n) = p(y_n \mid x_n)$$

$$\alpha(x_i)\beta(x_i) = p(y_0, \ldots, y_n, x_i)$$

data likelihood

$$\sum_{x_i} \alpha(x_i)\beta(x_i) = p(y_0, \ldots, y_n) = \ell(\text{data})$$

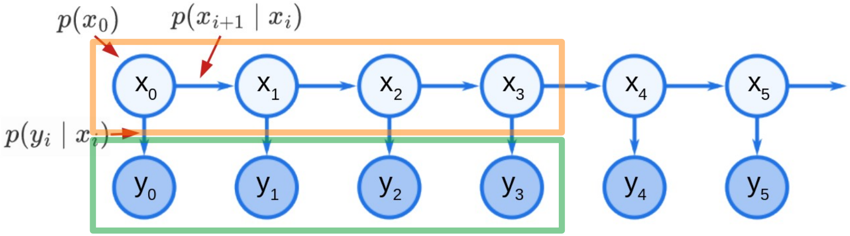
marginals

$$\frac{\alpha(x_i)\beta(x_i)}{\ell(\text{data})} = p(x_i \mid y_0, \ldots, y_n)$$

Durham University

HMM-Based Chord Recognition

Viterbi Algorithm



maximum posterior estimate

$$\max_{x_0,...,x_n} p(x_0, \ldots, x_n \mid y_0, \ldots, y_n)$$

$$= \max_{x_0,...,x_n} \frac{p(y_0, \ldots, y_n, x_0, \ldots, x_n)}{p(y_0, \ldots, y_n)}$$

$$= \frac{1}{\ell(data)} \max_{x_0,...,x_n} p(y_0, \ldots, y_n, x_0, \ldots, x_n)$$

Viterbi

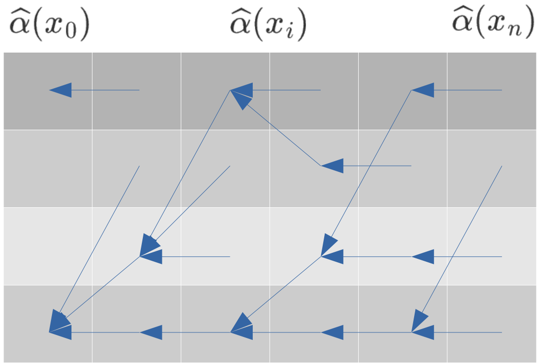
$$\widehat{\alpha}(x_i) := \max_{x_0,...,x_{i-1}} p(y_0, \ldots, y_i, x_0, \ldots, x_i)$$

$$= p(y_i \mid x_i) \max_{x_{i-1}} p(x_i \mid x_{i-1}) \max_{x_0,...,x_{i-2}} p(y_0, \ldots, y_{i-1}, x_0, \ldots, x_{i-1})$$

$$= p(y_i \mid x_i) \max_{x_{i-1}} p(x_i \mid x_{i-1}) \, \widehat{\alpha}(x_{i-1})$$

$$\widehat{\alpha}(x_0) = p(y_0, x_0) = p(y_0 \mid x_0) \, p(x_0)$$

$$\max_{x_n} \widehat{\alpha}(x_n) = \max_{x_0,...,x_n} p(y_0, \ldots, y_n, x_0, \ldots, x_n)$$



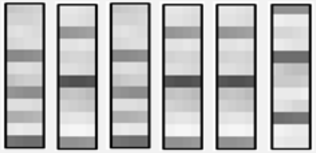
HMM-Based Chord Recognition

Fig. 5.28

Input

Observation sequence

$O = (o_1, o_2, o_3, o_4, o_5, o_6)$



β_1 β_3 β_1 β_3 β_3 β_2

Viterbi algorithm

D	$o_1 = \beta_1$	$o_2 = \beta_3$	$o_3 = \beta_1$	$o_4 = \beta_3$	$o_5 = \beta_3$	$o_6 = \beta_2$
α_1	0.4200	0.1008	0.0564	0.0135	0.0033	0
α_2	0.0200	0	0.0010	0	0	0.0006
α_3	0	0.0336	0	0.0045	0.0022	0.0003

E	$o_1 = \beta_1$	$o_2 = \beta_3$	$o_3 = \beta_1$	$o_4 = \beta_3$	$o_5 = \beta_3$
α_1	$\boxed{1}$	$\boxed{1}$	1	1	1
α_2	1	1	1	1	$\boxed{3}$
α_3	1	3	$\boxed{1}$	$\boxed{3}$	3

$\boxed{i_6 = 2}$

Output

Optimal state sequence

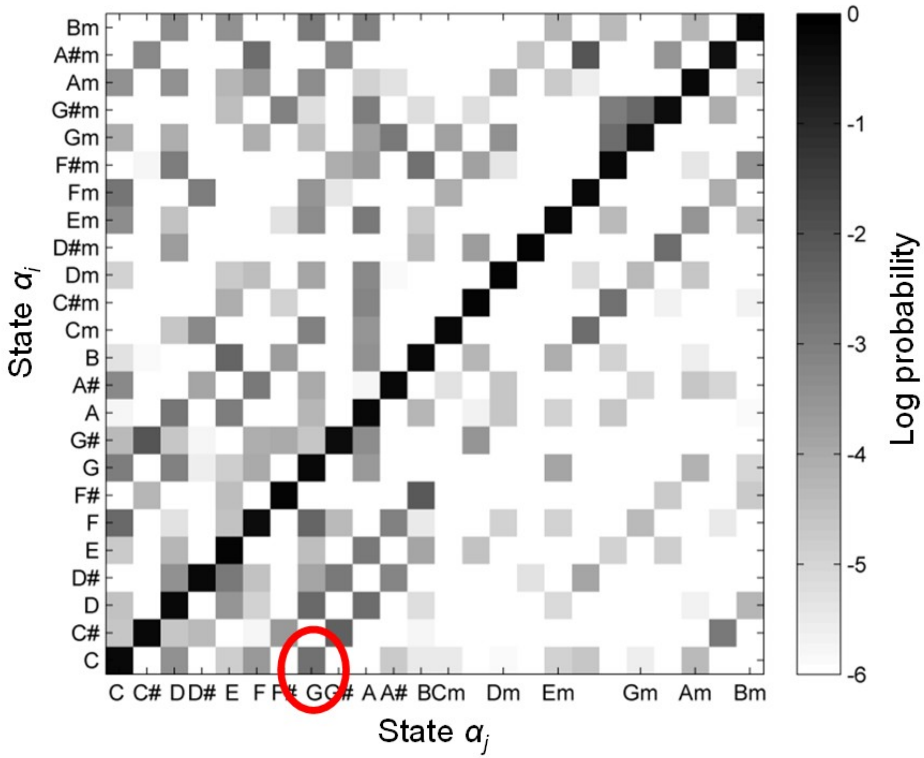
$S^* = (\alpha_1, \alpha_1, \alpha_1, \alpha_3, \alpha_3, \alpha_2)$

(different naming convention!!)

Durham University

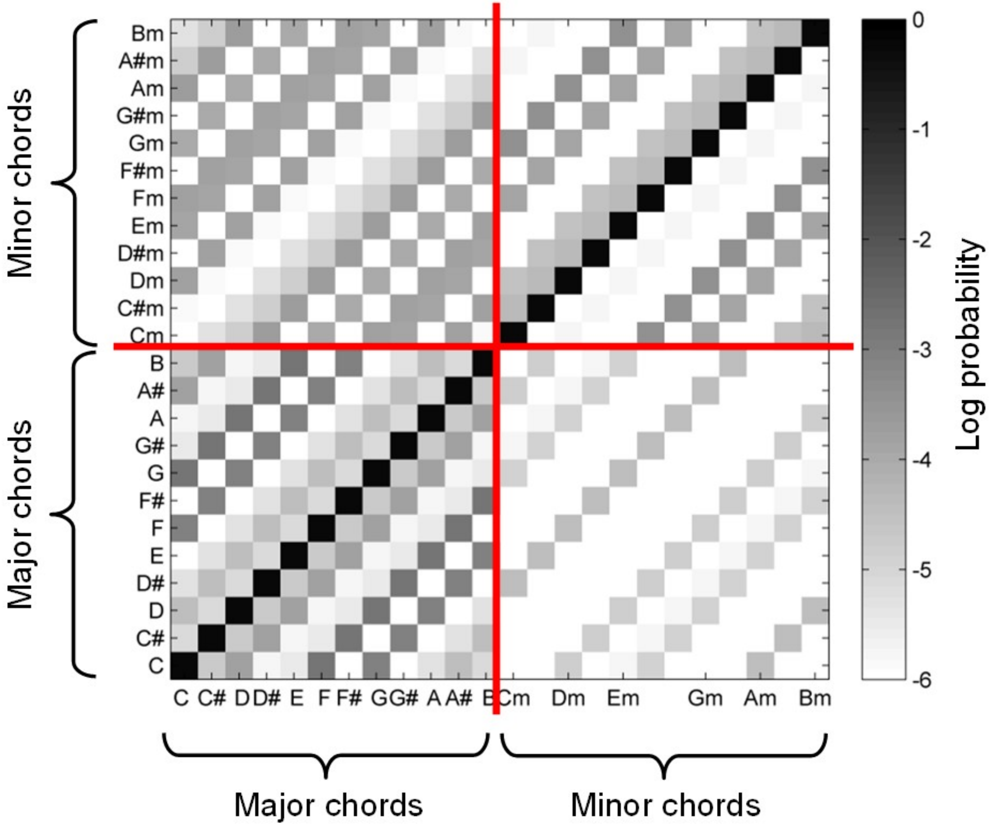
HMM-Based Chord Recognition

Fig. 5.29



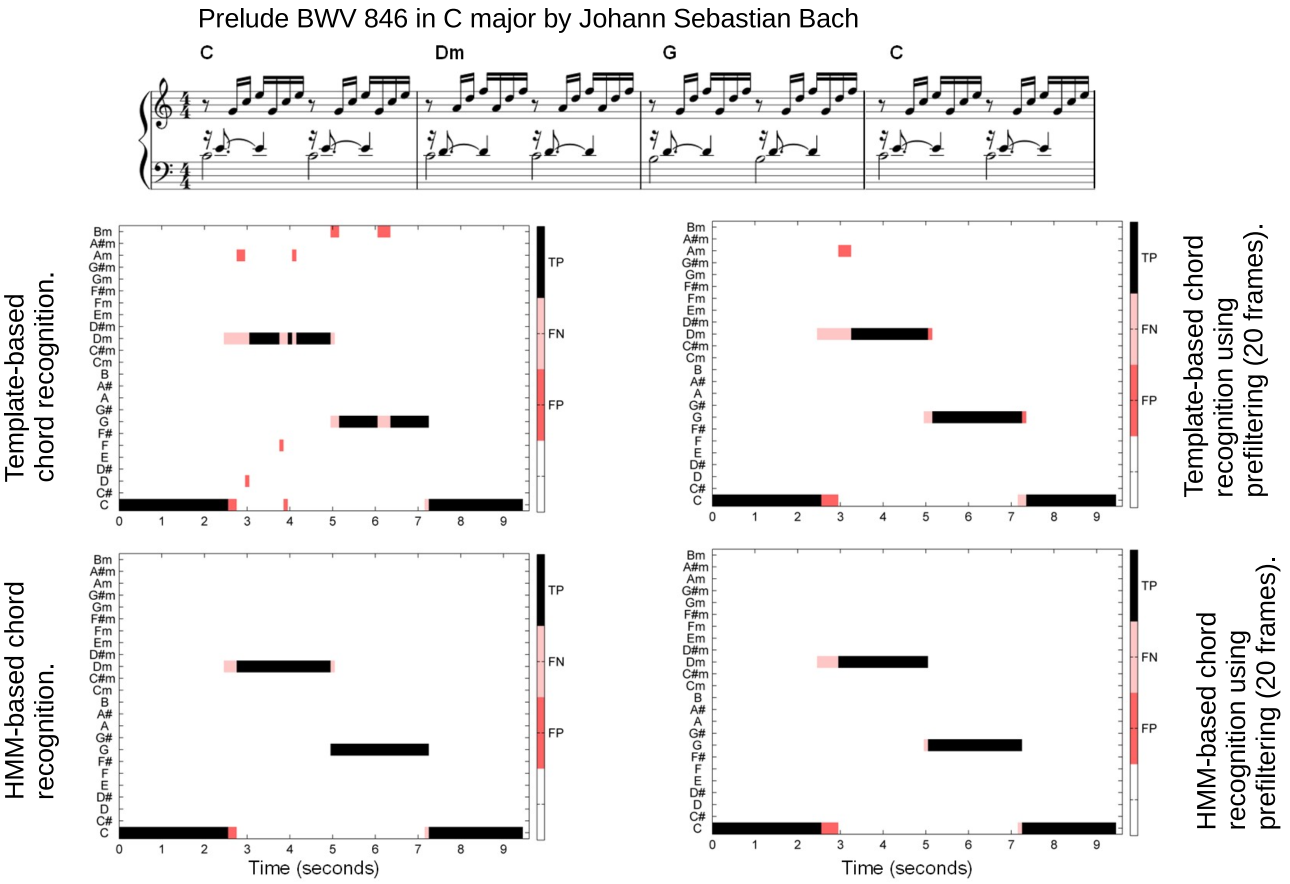
HMM-Based Chord Recognition

Fig. 5.30



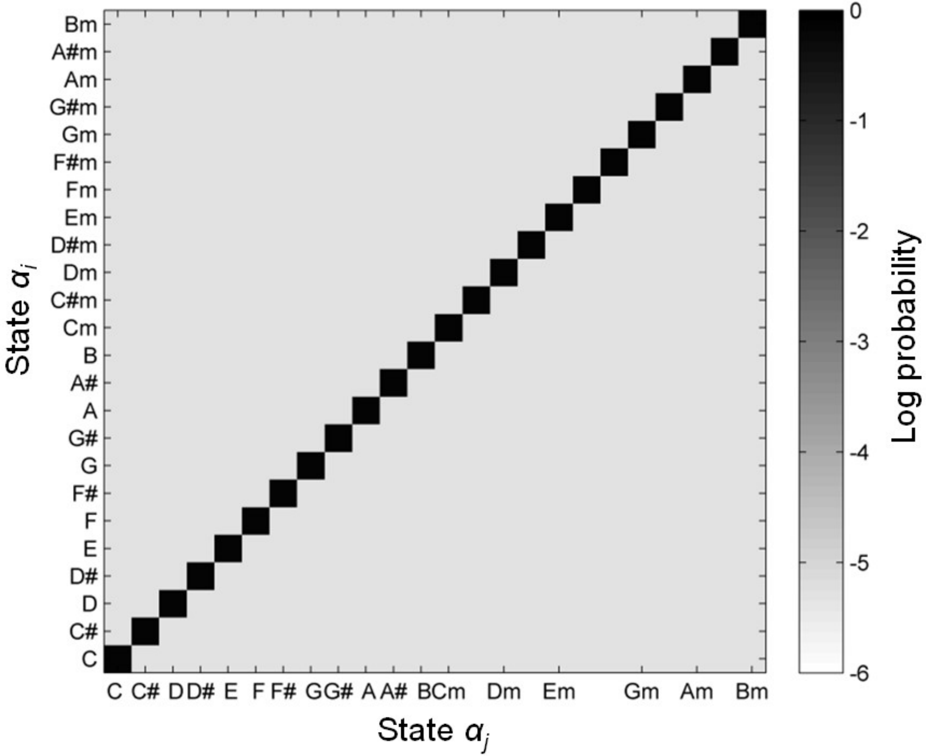
HMM-Based Chord Recognition

Fig. 5.31



HMM-Based Chord Recognition

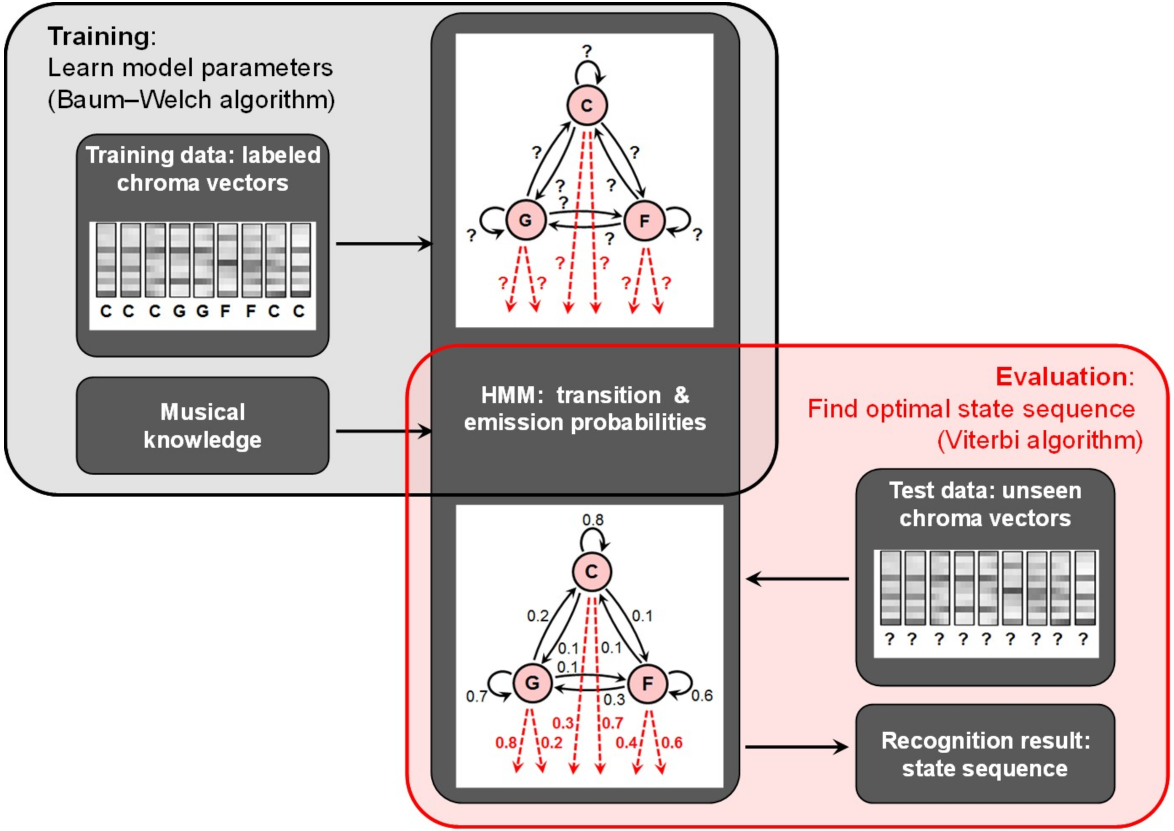
Fig. 5.32



Durham University

Further Notes

Fig. 5.33



References

1) Meinard Müller (2015) Fundamentals of music processing: Audio, analysis, algorithms, applications. Springer
2) Bishop CM (2007) Pattern Recognition and Machine Learning (Information Science and Statistics), 1st ed. 2006. Corr. 2nd printing. Springer