Engraving Oriented Joint Estimation of Pitch Spelling and Local and Global Keys

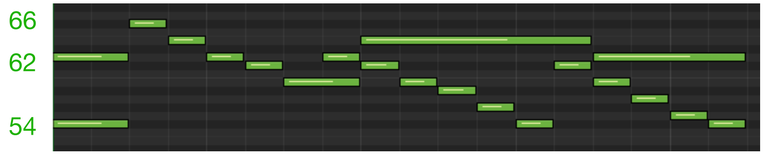
Augustin Bouquillard

Florent Jacquemard

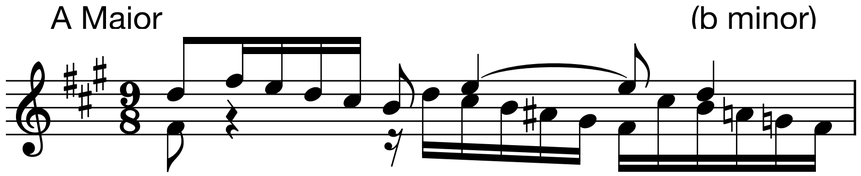
TENOR Zurich 2024

Pitch Spelling

Given MIDI representation of pitches
in 0..128 semitones…

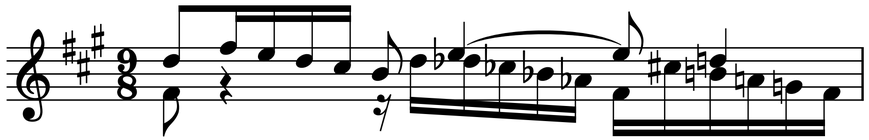


…find appropriate note names
in A..G, + accidentals



Bach, Fugue in A Major BWV864, measure 33, RH

Why the above spelling, and not this one?



pitch class	spelling_1	spelling_2	spelling_3
0	D𝄫	C	B♯
1	D♭	C♯	[B𝄪]
2	E𝄫	D	C𝄪
3	[F𝄫]	E♭	D♯
4	F♭	E	D𝄪
5	G𝄫	F	E♯
6	G♭	F♯	[E𝄪]
7	A𝄫	G	F𝄪
8	A♭	G♯	
9	B𝄫	A	G𝄪
10	[C𝄫]	B♭	A♯
11	C♭	B	A𝄪

It depends on the context of occurrence (of the note):

- global tonality of the piece (Key)
- the local tonality,
- the harmonic context,
- the voice-leading structure
  (ascending or descending melodic movements…),
- …

Related Work

Algorithms

- E. Cambouropoulos. *Pitch spelling: A computational model.* MP (20) 2003.
- D. Temperley. *The cognition of basic musical structures*. MIT press 2004.
- E. Chew and Y.-C. Chen. *Real-time pitch spelling using the spiral array*. CMJ (29) 2005.
- D. Meredith. *The PS13 pitch spelling algorithm*. JNMR (35) 2006.

Combinatoric Optimization (in appropriate data structures)

- A. K. Honingh. *Compactness in the Euler-lattice: A parsimonious pitch spelling model.* MS (13) 2009.
- B. Wetherfield. *The minimum cut pitch spelling algorithm: Simplifications and developments.* TENOR 2020.

Statistical models, trained on datasets

- G. Teodoru and C. Raphael. *Pitch spelling with conditionally independent voices.* ISMIR 2007.
- F. Foscarin, N. Audebert, and R. Fournier-S'Niehotta. *PKSpell: Data-driven pitch spelling and key signature estimation*. ISMIR 2021

Principle of our approach

We consider the accidentals, as they would appear on the engraved score,
following the rules for common Western Music Notation:

The accidentals in the Key Signature are omitted by default.
An accidental holds good for the duration of a bar.
~~*It applies only to the pitch at which it is written: each additional octave requires a further accidental.*~~

E. Gould. *Behind Bars: The definitive guide to music notation.* Faber Music, 2011

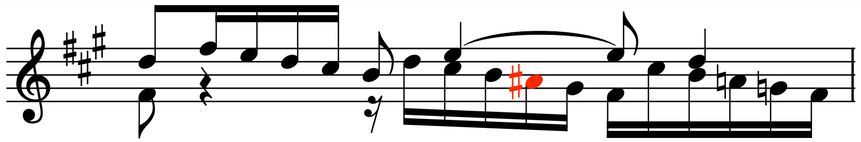
we forget about this
second part
for reasoning modulo 12

Principles (roughly):

1. try to have as few accidentals as possible (for the sake of readability)
2. accidentals are also useful indications of the composer intention
   (local key changes, harmony…)

➡ strong interdependency between Pitch-Spelling, and (local and global) Key Estimation (KE).

Shortest Path Computation



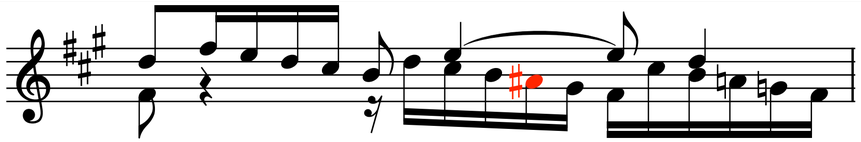
initial state s_0
= key signature

C	#
D	
E	
F	#
G	#
A	
B	

0

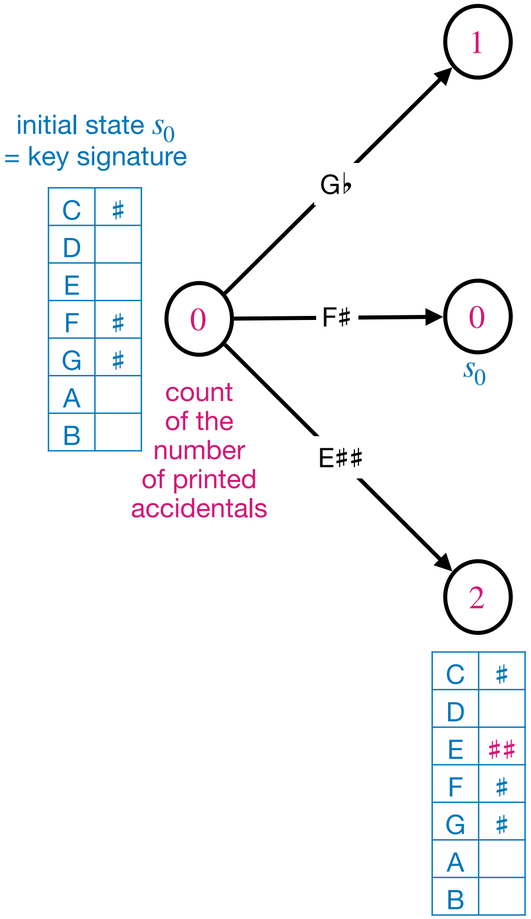
count of the number of printed accidentals

Shortest Path Computation



note$_1$

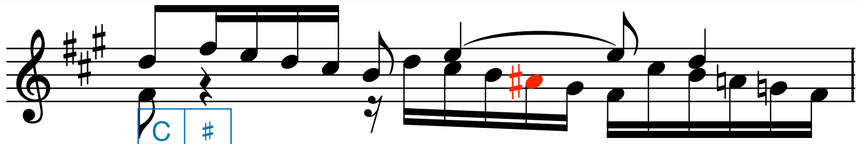
C	#
D	
E	
F	#
G	♭
A	
B	



C	#
D	
E	#
F	#
G	#
A	
B	

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

Shortest Path Computation



note$_1$

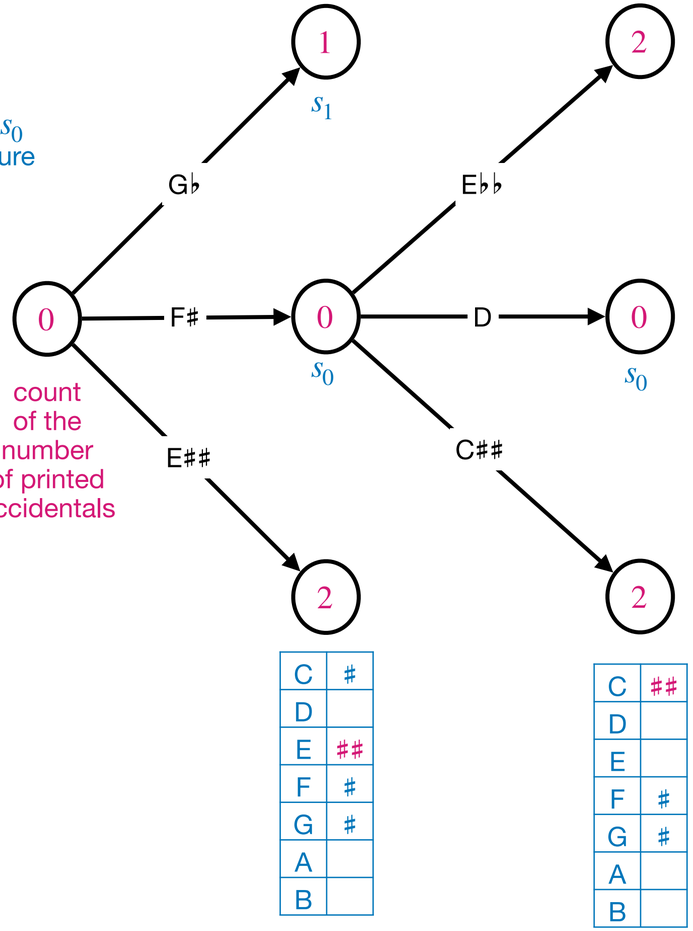
C	#
D	
E	
F	#
G	♭
A	
B	

note$_2$

C	#
D	
E	♭♭
F	#
G	#
A	
B	

initial state s_0
= key signature

C	#
D	
E	
F	#
G	#
A	
B	

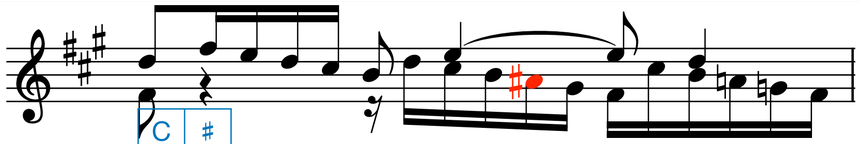


count of the number of printed accidentals

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

C	##
D	
E	
F	#
G	#
A	
B	

Shortest Path Computation



note$_1$ note$_2$ … note$_{12}$

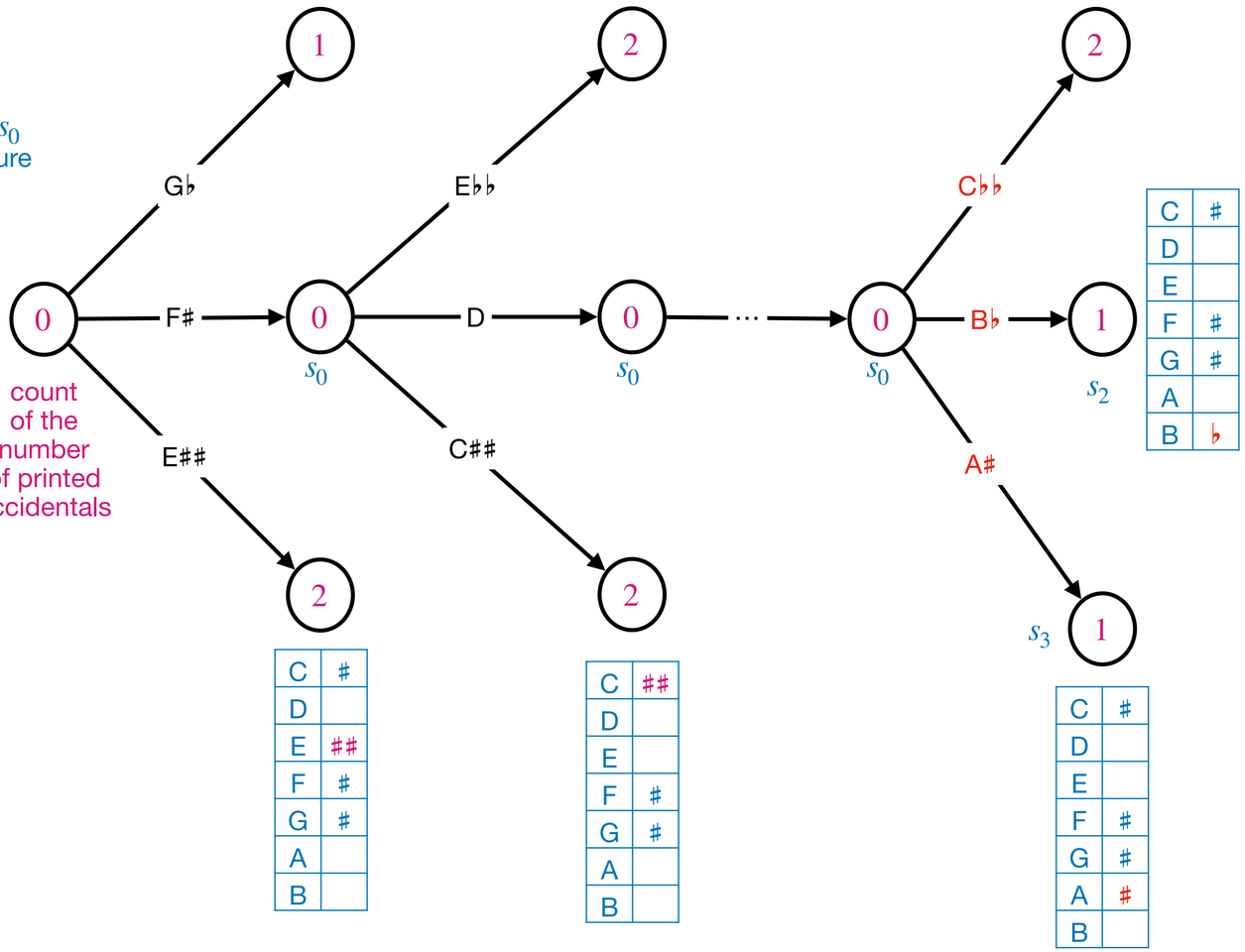
C	#
D	
E	
F	#
G	♭
A	
B	

C	#
D	
E	♭♭
F	#
G	#
A	
B	

initial state s_0
= key signature

C	#
D	
E	
F	#
G	#
A	
B	

count of the number of printed accidentals



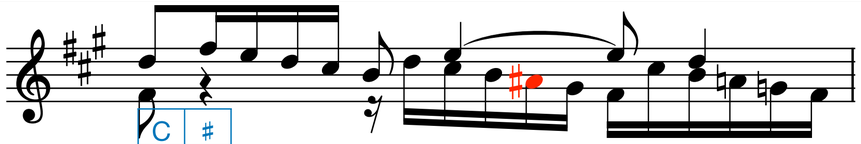
C	#
D	
E	##
F	#
G	#
A	
B	

C	##
D	
E	
F	#
G	#
A	
B	

C	#
D	
E	
F	#
G	#
A	#
B	

C	#
D	
E	
F	#
G	#
A	
B	♭

Shortest Path Computation



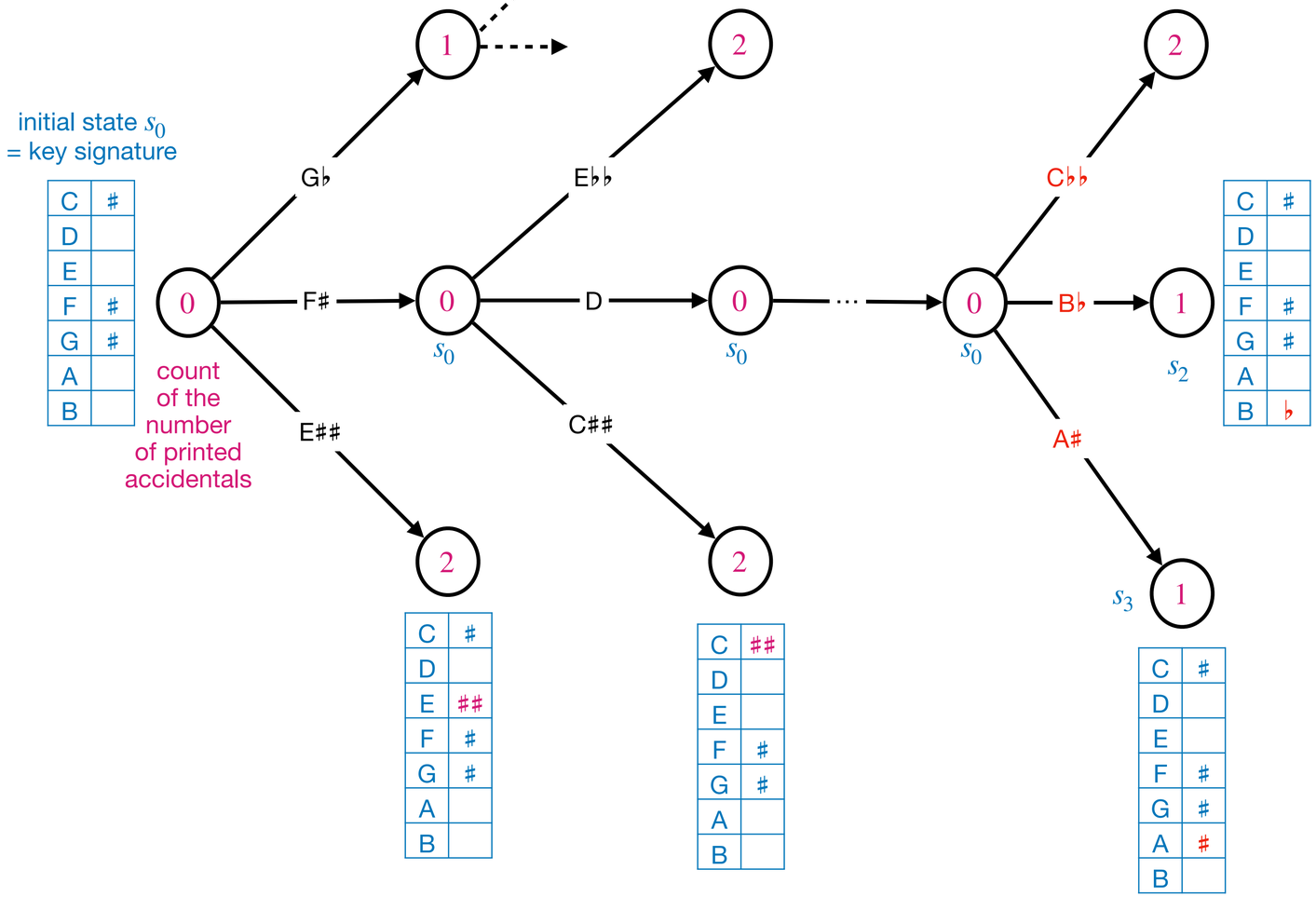
note$_1$ note$_2$ … note$_{12}$

C	#
D	
E	
F	#
G	♭
A	
B	

C	#
D	
E	♭♭
F	#
G	#
A	
B	

initial state s_0
= key signature

C	#
D	
E	
F	#
G	#
A	
B	



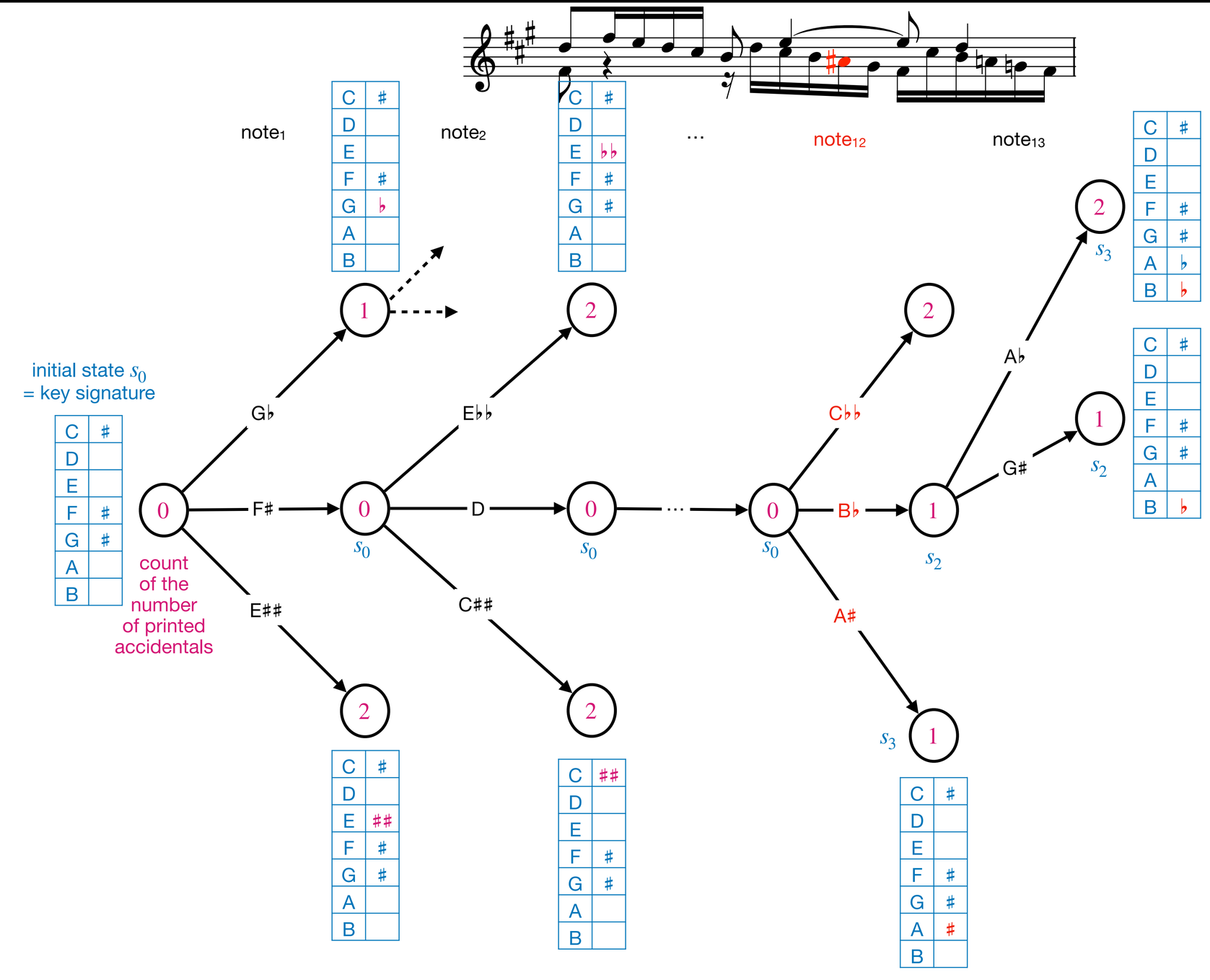
count of the number of printed accidentals

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

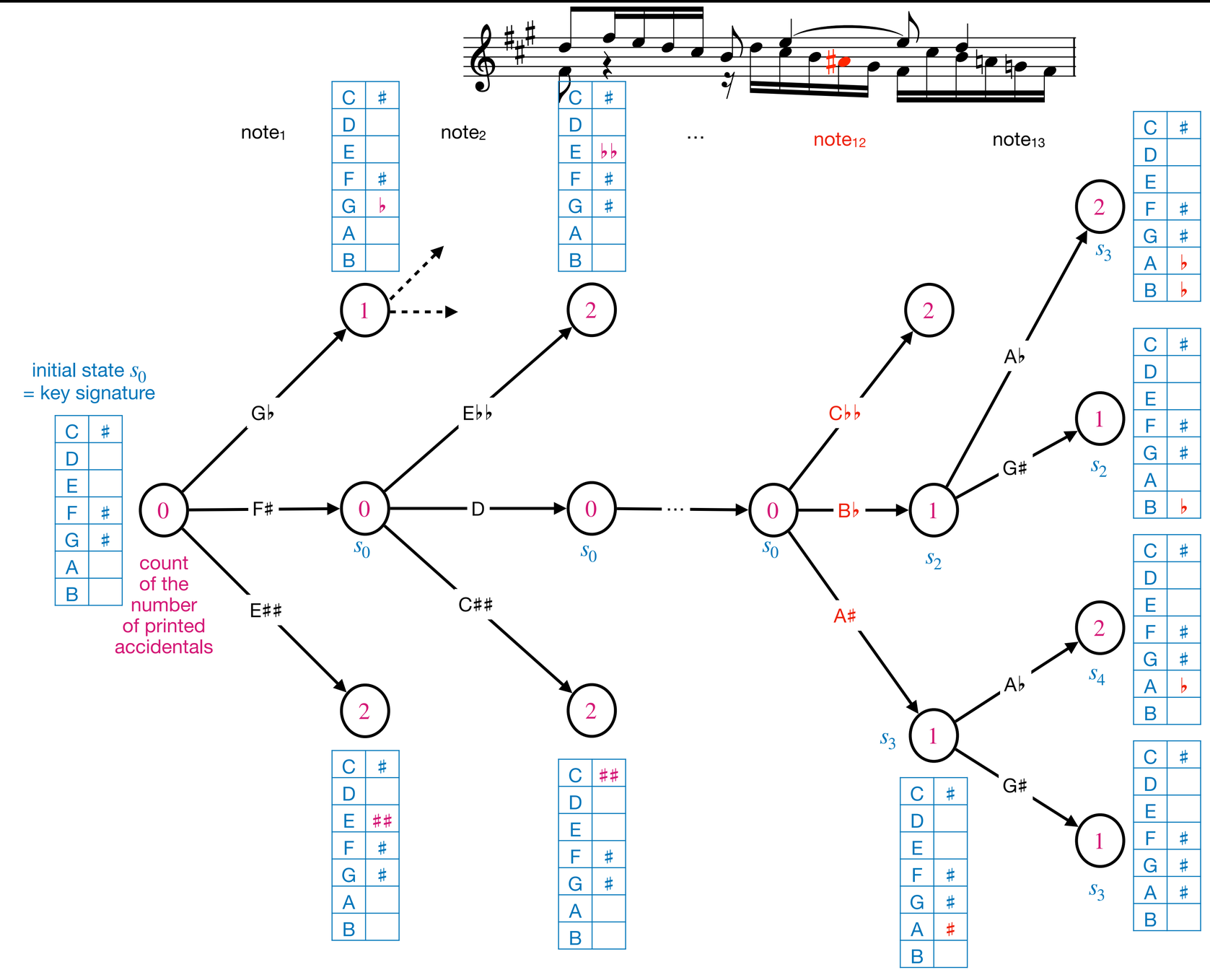
C	##
D	
E	
F	#
G	#
A	
B	

C	#
D	
E	
F	#
G	#
A	#
B	

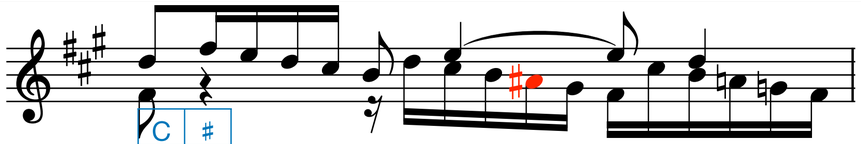
Shortest Path Computation



Shortest Path Computation



Shortest Path Computation



note$_1$ note$_2$ … note$_{12}$ note$_{13}$ … note$_{18}$ note$_{19}$ note$_{20}$ note$_{21}$

initial state s_0
= key signature

C	#
D	
E	
F	#
G	#
A	
B	

count of the number of printed accidentals

0 F♯ 0 D 0 … 0 B♭ 1 G♯ 1 s_2

0 G♭ 1

0 E♭♭ 2

0 C♭♭ 2 A♭ 2 s_3

0 E♯♯ 2

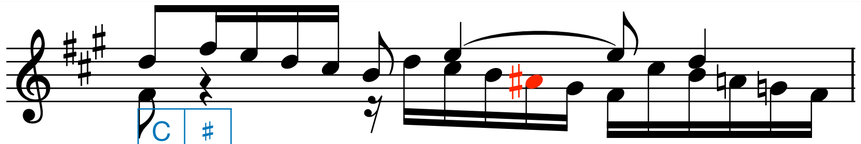
0 C♯♯ 2

0 A♯ 1 A♭ 2 s_4

1 G♯ 1 s_3

1 B 1 A♮ 2 G♮ 3 F♯ 3

Shortest Path Computation



note$_1$ note$_2$ … note$_{12}$ note$_{13}$ … note$_{18}$ note$_{19}$ note$_{20}$ note$_{21}$

initial state s_0 = key signature

C	#
D	
E	
F	#
G	#
A	
B	

count of the number of printed accidentals

G♭ F♯ E♯♯ E♭♭ D C♯♯ C♭♭ B♭ A♯ A♭ G♯ B A G F♯ (tie)

C	#
D	
E	
F	#
G	♭
A	
B	

C	#
D	
E	♭♭
F	#
G	#
A	
B	

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

C	##
D	
E	
F	#
G	#
A	
B	

C	#
D	
E	
F	#
G	#
A	#
B	

C	#
D	
E	
F	#
G	#
A	♭
B	♭

C	#
D	
E	
F	#
G	#
A	
B	♭

C	#
D	
E	
F	#
G	#
A	♭
B	

C	#
D	
E	
F	#
G	#
A	#
B	

C	#
D	
E	
F	#
G	#
A	
B	♮

C	#
D	
E	
F	#
G	#
A	
B	

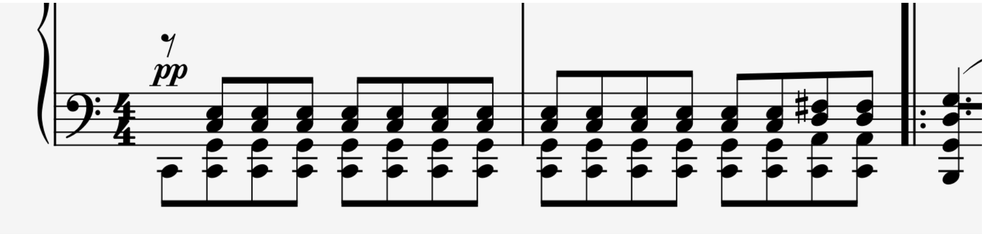
C	#
D	
E	
F	#
G	♮
A	♮
B	

C	#
D	
E	
F	#
G	#
A	
B	

C	#
D	
E	
F	#
G	♮
A	♮
B	

Combinatoric explosion

- the graph is built on-the-fly,
  only the necessary states are constructed.
- in the worst case, 3^n where n = nb of notes per mesure



Beethoven, Sonata 21 ”Waldstein”, measures 1-3, LH

makes the computation explode!

additional rule:

two simultaneous notes in the same pitch class must have the same name.

here, notes are called simultaneous if they have the same onset and are not grace notes.

This rule is ensured with a mapping: 0..11 → { A,…,G } along with the state.

It reduces combinatorial explosion in cases like the above.

Tie Breaking

For tie breaking, we consider local tonalities.

Tie Breaking

We can compute the best paths for every measure and every key signature.

Tie Breaking

~~We can compute the best paths for every measure and every key signature~~.

We can compute the best paths for every measure and every Key = KS + mode (major or harmonic minor)

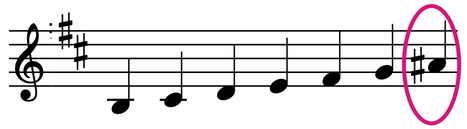
Tie Breaking

We can compute the best paths for every measure and every Key = KS + mode (major or harmonic minor)

for the cost of paths:

- count printed accidental
- discount accidental on lead degree in minor scales (it is added to the initial state)

e.g. B min



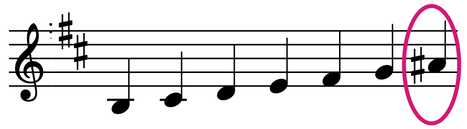
Tie Breaking

We can compute the best paths for every measure and every Key = KS + mode (major or harmonic minor)

for the cost of paths:

- count printed accidental
- discount accidental on lead degree in minor scales (it is added to the initial state)

e.g. B min



	bar_0	bar_1	…
D♭ maj	$paths_{0,0}$	$paths_{1,0}$	
A♭ maj	$paths_{0,1}$	$paths_{1,1}$	
…	…	…	
B♭ min	$paths_{0,12}$	$paths_{1,12}$	
…	…	…	

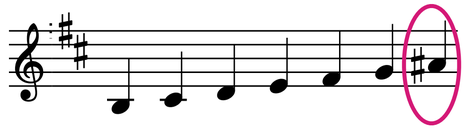
Procedure

step1: We can compute the best paths for every measure and every Key = KS + mode (major or harmonic minor)

for the cost of paths:

- count printed accidental
- discount accidental on lead degree in minor scales (it is added to the initial state)

e.g. B min



	bar_0	bar_1	…
D♭ maj	$paths_{0,0}$	$paths_{1,0}$	
A♭ maj	$paths_{0,1}$	$paths_{1,1}$	
…	…	…	
B♭ min	$paths_{0,12}$	$paths_{1,12}$	
…	…	…	

step 2: compute a grid with 1 estimated local Key for each global Key and each measure

it is the Key with best rank according to 3 rankings:

- rank in the corresponding column in the table of step 1
- distance to estimated Key for previous measure
- distance to candidate global Key

	bar_0	bar_1	…
D♭ maj	$key_{0,0}$	$key_{1,0}$	
A♭ maj	$key_{0,1}$	$key_{1,1}$	
…	…	…	
B♭ min	$key_{0,12}$	$key_{1,12}$	
…	…	…	

Weber distance between tonalities

G. Weber. *Versucht einer geordneten Theory der Tonsetzkunst*. B. Schott's Sohnen, 1818

KS		-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
		C♭	G♭	D♭	A♭	E♭	B♭	F	C	G	D	A	E	B	F♯	C♯	a♭	e♭	b♭	f	c	g	d	a	e	b	f♯	c♯	g♯	d♯	a♯
-7	C♭	0	1	2	2	3	4	4	5	6	6	7	8	8	9	10	1	2	3	3	4	5	5	6	7	7	8	9	9	10	11
-6	G♭	1	0	1	2	2	3	4	4	5	6	6	7	8	8	9	2	1	2	3	3	4	5	5	6	7	7	8	9	9	10
-5	D♭	2	1	0	1	2	2	3	4	4	5	6	6	7	8	8	2	2	1	2	3	3	4	5	5	6	7	7	8	9	9
-4	A♭	2	2	1	0	1	2	2	3	4	4	5	6	6	7	8	1	2	2	1	2	3	3	4	5	5	6	7	7	8	9
-3	E♭	3	2	2	1	0	1	2	2	3	4	4	5	6	6	7	2	1	2	2	1	2	3	3	4	5	5	6	7	7	8
-2	B♭	4	3	2	2	1	0	1	2	2	3	4	4	5	6	6	3	2	1	2	2	1	2	3	3	4	5	5	6	7	7
-1	F	4	4	3	2	2	1	0	1	2	2	3	4	4	5	6	3	3	2	1	2	2	1	2	3	3	4	5	5	6	7
0	C	5	4	4	3	2	2	1	0	1	2	2	3	4	4	5	4	3	3	2	1	2	2	1	2	3	3	4	5	5	6
1	G	6	5	4	4	3	2	2	1	0	1	2	2	3	4	4	5	4	3	3	2	1	2	2	1	2	3	3	4	5	5
2	D	6	6	5	4	4	3	2	2	1	0	1	2	2	3	4	5	5	4	3	3	2	1	2	2	1	2	3	3	4	5
3	A	7	6	6	5	4	4	3	2	2	1	0	1	2	2	3	6	5	5	4	3	3	2	1	2	2	1	2	3	3	4
4	E	8	7	6	6	5	4	4	3	2	2	1	0	1	2	2	7	6	5	5	4	3	3	2	1	2	2	1	2	3	3
5	B	8	8	7	6	6	5	4	4	3	2	2	1	0	1	2	7	7	6	5	5	4	3	3	2	1	2	2	1	2	3
6	F♯	9	8	8	7	6	6	5	4	4	3	2	2	1	0	1	8	7	7	6	5	5	4	3	3	2	1	2	2	1	2
7	C♯	10	9	8	8	7	6	6	5	4	4	3	2	2	1	0	9	8	7	7	6	5	5	4	3	3	2	1	2	2	1
-7	a♭	1	2	2	1	2	3	3	4	5	5	6	7	7	8	9	0	1	2	2	3	4	4	5	6	6	7	8	8	9	10
-6	e♭	2	1	2	2	1	2	3	3	4	5	5	6	7	7	8	1	0	1	2	2	3	4	4	5	6	6	7	8	8	9
-5	b	3	2	1	2	2	1	2	3	3	4	5	5	6	7	7	2	1	0	1	2	2	3	4	4	5	6	6	7	8	8
-4	f	3	3	2	1	2	2	1	2	3	3	4	5	5	6	7	2	2	1	0	1	2	2	3	4	4	5	6	6	7	8
-3	c	4	3	3	2	1	2	2	1	2	3	3	4	5	5	6	3	2	2	1	0	1	2	2	3	4	4	5	6	6	7
-2	g	5	4	3	3	2	1	2	2	1	2	3	3	4	5	5	4	3	2	2	1	0	1	2	2	3	4	4	5	6	6
-1	d	5	5	4	3	3	2	1	2	2	1	2	3	3	4	5	4	4	3	2	2	1	0	1	2	2	3	4	4	5	6
0	a	6	5	5	4	3	3	2	1	2	2	1	2	3	3	4	5	4	4	3	2	2	1	0	1	2	2	3	4	4	5
1	e	7	6	5	5	4	3	3	2	1	2	2	1	2	3	3	6	5	4	4	3	2	2	1	0	1	2	2	3	4	4
2	b	7	7	6	5	5	4	3	3	2	1	2	2	1	2	3	6	6	5	4	4	3	2	2	1	0	1	2	2	3	4
3	f♯	8	7	7	6	5	5	4	3	3	2	1	2	2	1	2	7	6	6	5	4	4	3	2	2	1	0	1	2	2	3
4	c♯	9	8	7	7	6	5	5	4	3	3	2	1	2	2	1	8	7	6	6	5	4	4	3	2	2	1	0	1	2	2
5	g♯	9	9	8	7	7	6	5	5	4	3	3	2	1	2	2	8	8	7	6	6	5	4	4	3	2	2	1	0	1	2
6	d♯	10	9	9	8	7	7	6	5	5	4	3	3	2	1	2	9	8	8	7	6	6	5	4	4	3	2	2	1	0	1
7	a♯	11	10	9	9	8	7	7	6	5	5	4	3	3	2	1	10	9	8	8	7	6	6	5	4	4	3	2	2	1	0

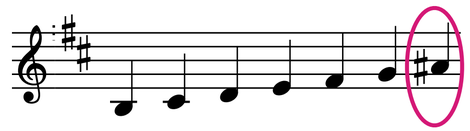
Procedure

step1: We can compute the best paths for every measure and every Key = KS + mode (major or harmonic minor)

for the cost of paths:

- count printed accidental
- discount accidental on lead degree in minor scales (it is added to the initial state)

e.g. B min



	bar_0	bar_1	…
D♭ maj	$paths_{0,0}$	$paths_{1,0}$	
A♭ maj	$paths_{0,1}$	$paths_{1,1}$	
…	…	…	
B♭ min	$paths_{0,12}$	$paths_{1,12}$	
…	…	…	

step 2: compute a grid with 1 estimated local Key for each global Key and each measure

it is the Key with best rank according to 3 rankings:

- rank in the corresponding column in the table of step 1
- distance to estimated Key for previous measure
- distance to candidate global Key

	bar_0	bar_1	…
D♭ maj	$key_{0,0}$	$key_{1,0}$	
A♭ maj	$key_{0,1}$	$key_{1,1}$	
…	…	…	
B♭ min	$key_{0,12}$	$key_{1,12}$	
…	…	…	

step 3: compute a new table with 1 best path for each key and each measure

for the cost of paths:

- count printed accidental, no discount (initial state = Key Signature)
- accidental not in the scale of the estimated local Key

	bar_0	bar_1	…
D♭ maj	$path'_{0,0}$	$path'_{1,0}$	
A♭ maj	$path'_{0,1}$	$path'_{1,1}$	
…	…	…	
B♭ min	$path'_{0,12}$	$path'_{1,12}$	
…	…	…	

step 4: select the row in the second table with the smallest cumulated cost

it contains:

- a spelling (following the best paths)
- a key
- one local key for each measure

Algorithm PSE

Input: a sequence of notes,
each note is given by:

- one MIDI pitch values in 0..128
- whether the note is simultaneous with the next
- the measure number

Output:

- a spelling: one name for each note.
- one global Key
  Key = Key Signature + Mode (here *major* or *harmonic minor).*
- one local Key for each measure
  Key Signature + Mode.

step 5 (post processing): rewrite the passing notes

with rules from:

D. Meredith

The PS13 pitch spelling algorithm

JNMR (35) 2006.

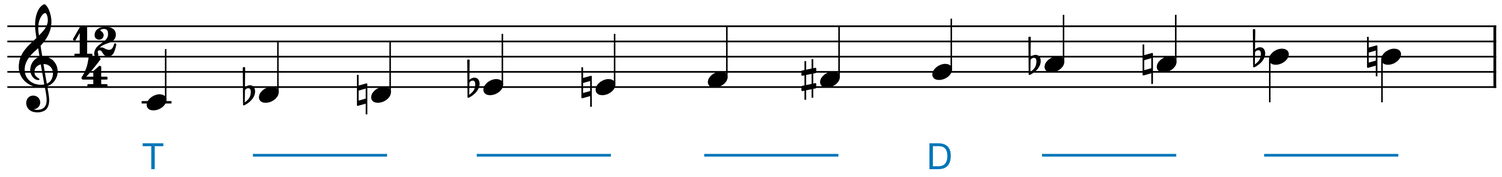
like the following

broderie down	C C♭ C	→	C B C
broderie up	C C♯ C	→	C D♭ C
descending_{11}	C C♭ A	→	C B A
descending_{12}	C C𝄫 A♭	→	C B♭ A♭
descending_{21}	C A♯ A	→	C B♭ A
descending_{22}	C A♯ A♭	→	C B♭ A♭
ascending_{11}	A A♯ C	→	A B♭ C
ascending_{12}	A♭ A♯ C	→	A♭ B♭ C
ascending_{21}	A C♭ C	→	A B C
ascending_{22}	A C♭ C♯	→	A B C♯

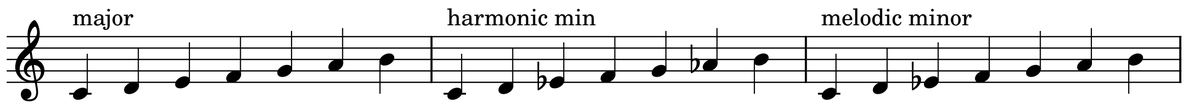
Deterministic Variant

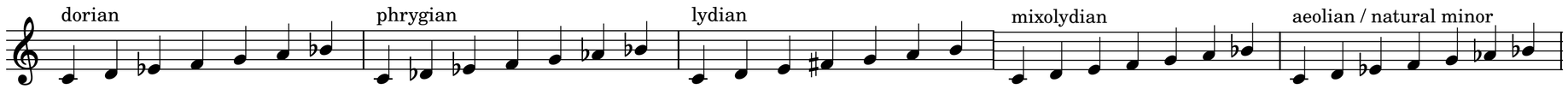
called PS13b

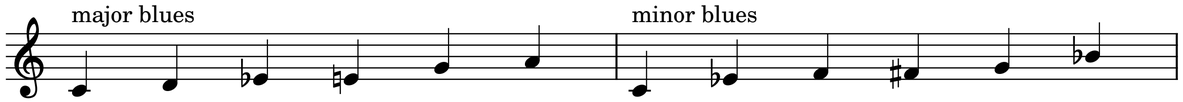
during the best path computation,
for each pitch-class, choose the (unique) name in the chromatic harmonic scale



this scale embeds:







Implementation
in C++20, for efficiency and integration

Evaluation
Python binding with pybind11, for the evaluation
- parse MusicXML score (with Music21 toolkit)
- extract input for the algorithm
- run algorithm PSE or PS13b
- compare note names and Key with original score
- generate feedback score (with Music21 toolkit)

Evaluation Data 216 464 notes from:
- monophonic (complex) dataset: Lamarque-Goudard
  textbook for learning rhythm in music schools.
  - 250 extracts from Bach and Scarlatti to Wolf, Duparc, Debussy, Ibert...

- piano dataset: ASAP (220 pieces or movements, MusicXML scores and MIDI recordings or performances)
  evaluation on 110 pieces:
  - Bach: Well Tempered Clavier - BWV856, 873
  - Mozart: some sonata movements + K 475 Fantaisie
  - Beethoven: some sonata movements
  - 13 Chopin Etudes (from opus 10 and 25)
  - 8 Rachmaninov preludes (from opus 23 and 32)

Evaluation Results

accuracy	number of spelled notes	pitch spelling PSE	pitch spelling PS13b	key signature estimation PSE	key signature estimation PS13b	key signature estimation Krumhansl-Schmuckler
Bach WTC ASAP corpus	55530	99.50%	98.27%	99.09%	98.29%	87.27%
5 movements from Mozart Sonatas present in ASAP	10043	99.11%	97.30%	100%	100%	80%
Fantaisie K. 745 plus 5 movements from Mozart Sonatas	13830	97.65%	95.97%	80%	80%	60%
33 movements from Beethoven Sonatas	87292	97.64%	95.65%	92.32%	95.71%	66.15%
13 Etudes by Chopin	25103	96.71%	96.03 %	96.15%	96.15 %	84.62%
4 Rachmaninov Preludes	7022	98.76%	97.49%	100%	100%	100%
Lamarque-Goudard	27687	98.46%	98.23%	76.90%	74.30%	50.60%

Comparison with other system:
unfair since we use information on measures (unlike most systems)

PKSpell (state-of-art on Meredith benchmark MuseData)
reaches accuracy, on 33 pieces from ASAP:

- 96.50% for Pitch Spelling
- 90.30% for Key Signature Estimation

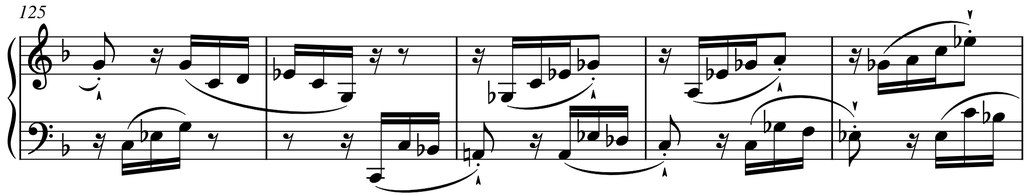
PSE
on 110 pieces from ASAP:

- 98.19% for Pitch Spelling
- 95.58% for Key Signature Estimation

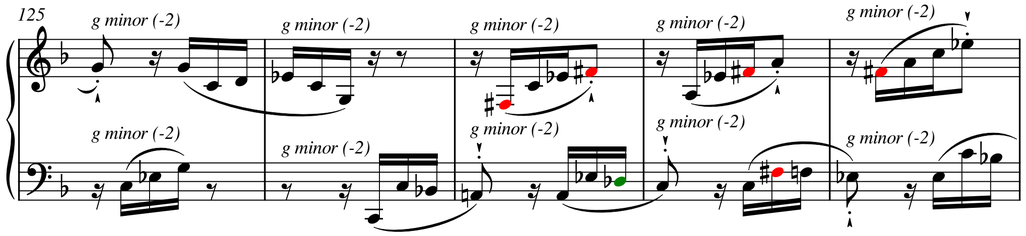
Feedback

https://gitlab.inria.fr/pse/pseval

Beethoven, Sonata 17 ”Tempest”, 3d mvt, measures 126-147
(original)

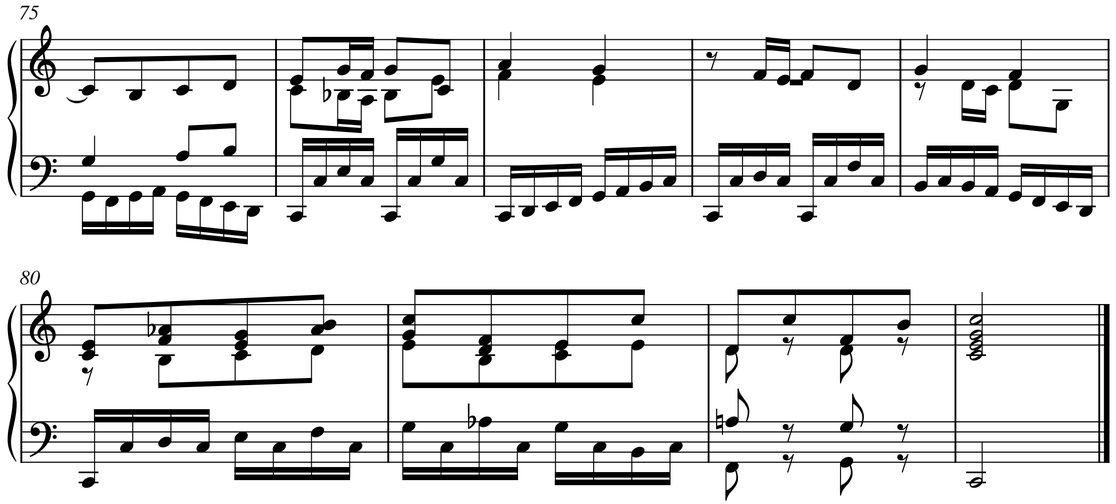


Beethoven, Sonata 17 ”Tempest”, 3d mvt, measures 125-129
(respelled with PSE and annotated)

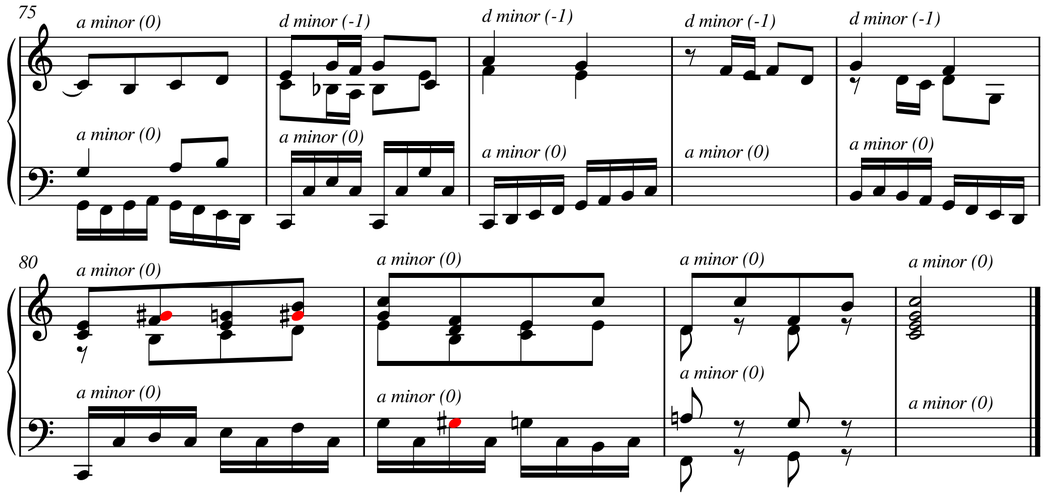


red note: spelling mistake.
green note: spelling mistake fixed by rewriting.

Bach, Fugue in C major, Das Wohltemperierte Clavier, last measures
(original)



Bach, Fugue in C major, Das Wohltemperierte Clavier, last measures
(respelled with PSE and annotated)



Conclusion, Further Work

Summary:

- Pitch Spelling algorithm based on engraving rules for accidentals
- It also estimates global Key and local Keys (one for each measure)
- Exhaustive search and efficient deterministic variant
- Evaluation on challenging datasets

algorithm for quantized music (bar number must be identified)
it makes sense for backend of transcription (after transcription)

Future work :

- for best path computation, consider:
  - note durations
  - metric weight

- jazz Pitch Spelling, with more modes