DATA QUALITY MATTERS

Iterative Corrections on a Corpus of Mendelssohn String Quartets and Implications for MIR Analysis

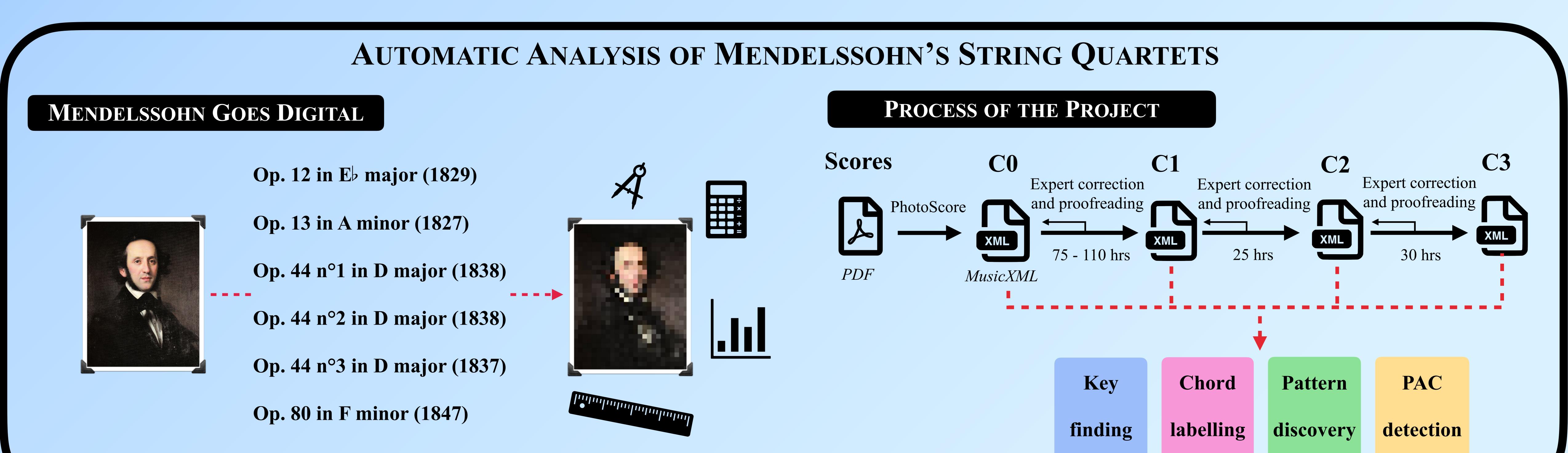
COMPUTATIONAL TONAL STUDY (CTS) GROUP

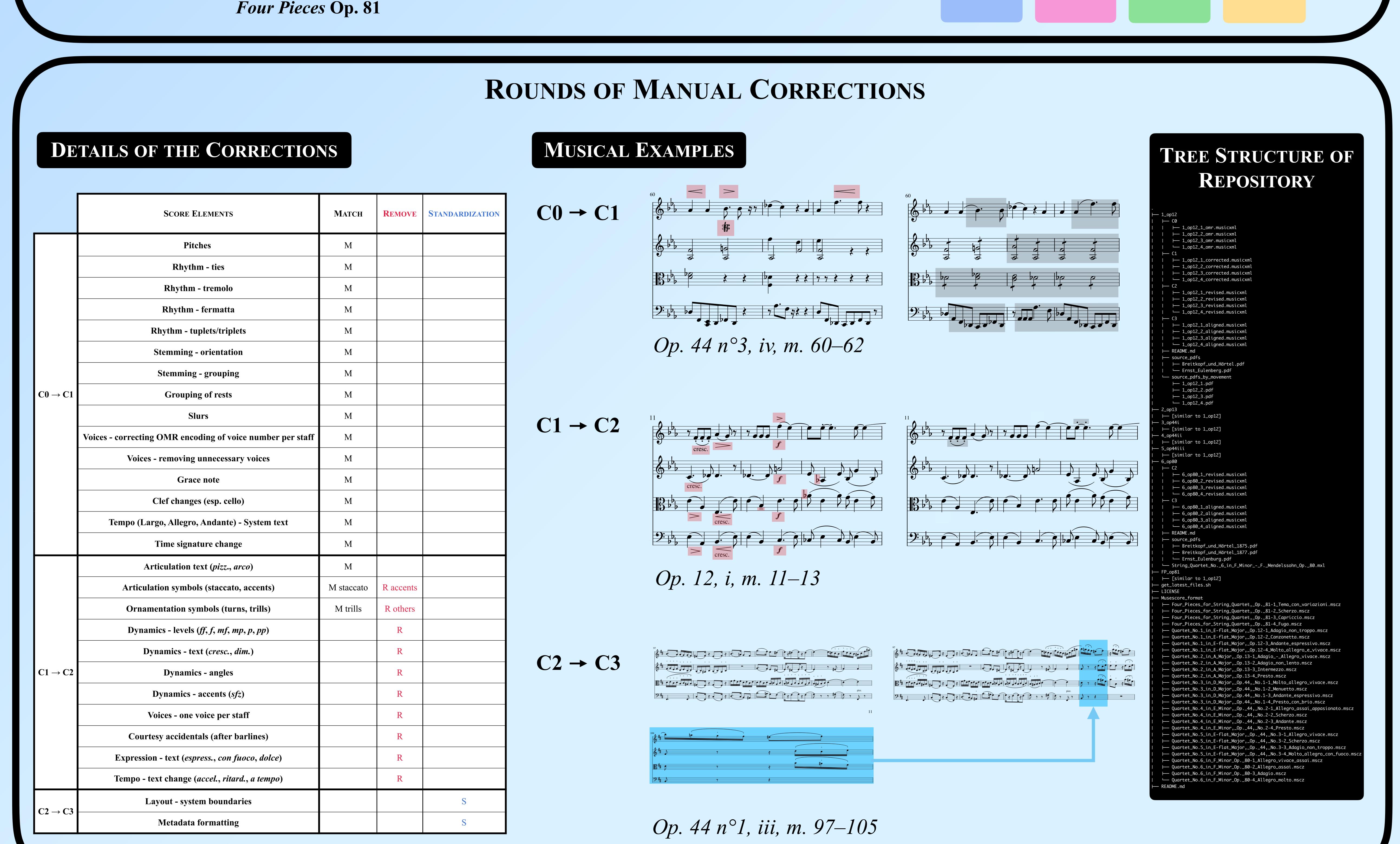
Schulich School of Music, McGill University, Canada

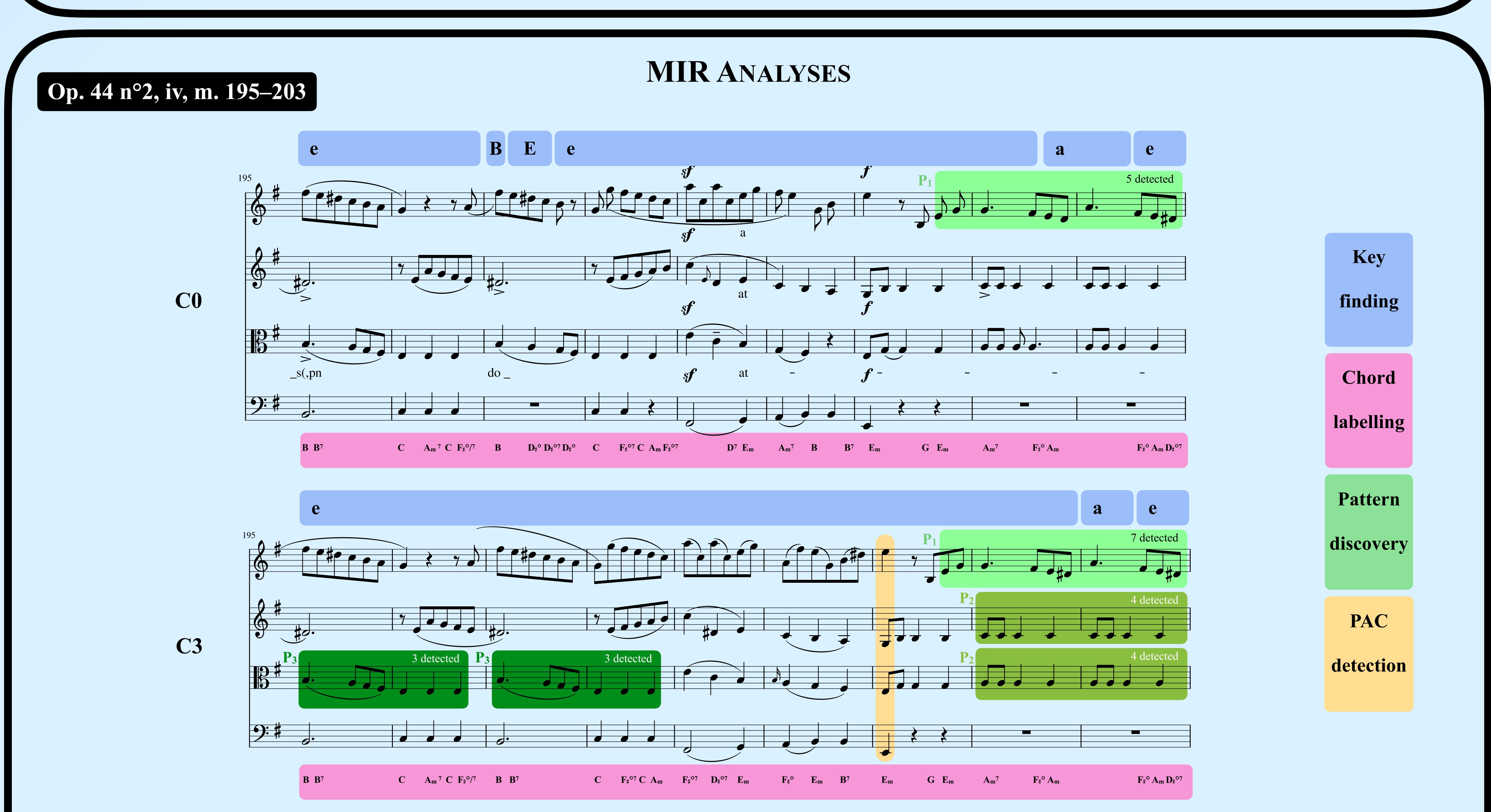
Néstor Nápoles López Jacob deGroot-Maggetti Timothy de Reuse Samuel Howes Yaolong Ju Suzuka Kokubu Sylvain Margot Finn Upham

CRIStAL, UMR 9189, CNRS, Université de Lille

Laurent Feisthauer







GENERAL RESULTS

COMPARISON PAIR	Changes in Local Key Annotations median(%) [min(%), max(%)]	Changes in Chord Labels median(%) [min(%), max(%)]	Changes in Monophonic Pattern median(%) [min(%), max(%)]			Changes in PAC Detection variation(%) number		
			Num. Patterns	Coverage	Cardinality	Changes in PACs detected	New PACs detected	PACs lost
$\mathbf{C0} \longrightarrow \mathbf{C1}$	46.8% [9.9%, 71.1%]	69.1% [17.5%, 96.7%]	85% [16%, 2100%]	110% [21%, 1100%]	8.3% [-22%, 29%]	154.5% 22 to 56	177.3% 39	22.7%
C1 → C2 + C1 proofreading	0.4% [0.0%, 9.8%]	0.7% [0.0%, 41.1%]	0.0% [-5.1%, 7.1%]	0.041% [0.85%, 9.8%]	0.0% [-3.4%, 3.5%]	-3.6% 56 to 54	1.8%	5.4%
C2 → C3 + C2 proofreading	0.0% [0.0%, 3.0%]	0.0% [0.0%, 12.5%]	0.0% [0.0%, 0.0%]	0.0% [0.0%, 0.0%]	0.0% [0.0%, 0.0%]	1.9% 54 to 55	3.7%	1.9%

OMR MAIN ERRORS

Most Common Errors

- Accidentals Rests
 - Missing
 - Misplaced
 - Tremolo ▶ Extra
 - ▶ Staccato Slur

Articulations

▶ Tie

- Hidden
- Fast passages
- Sixteenths
- Eighths
- ▶ Stems are troublesome

Most Impactful Errors

- Key signatures
- Clefs

