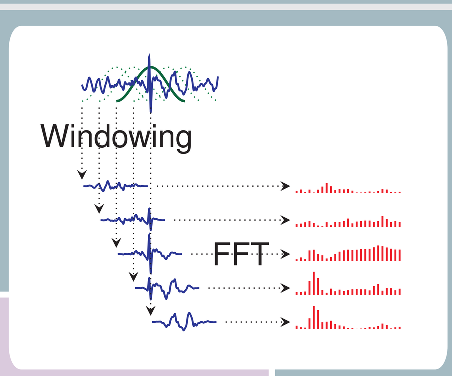
Least-Squares Local Tuning Frequency Estimation for Choir Music

V. Gnann, M. Kitza et al. Institut für Nachrichtentechnik RWTH Aachen University

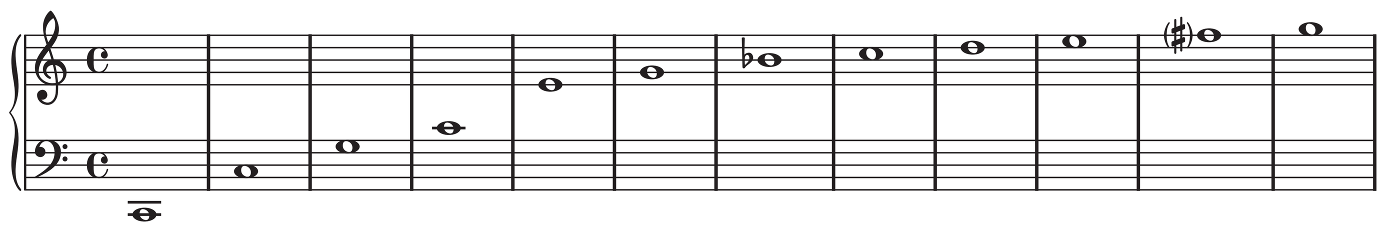
Motivation

- Unaccompanied choirs tend to lower their pitch while singing, often not noticed by their conductors.
- Objective: Tuning pitch display as a mobile app which works
  - independent from the actual tones
  - also in polyphonic situations.



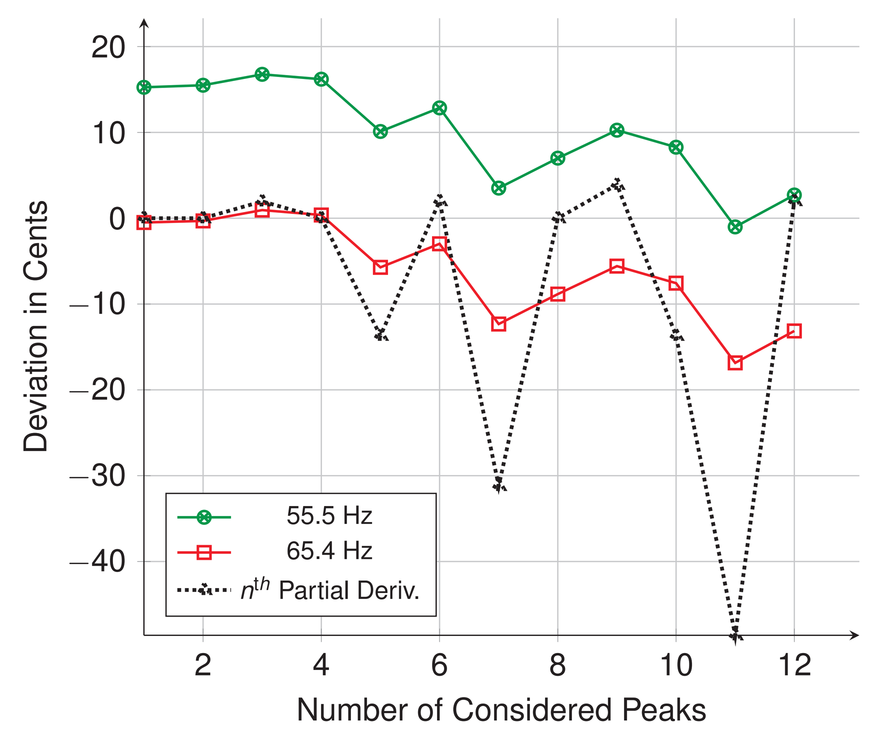
Spectrogram

- STFT with long windows, low sampling rate, and without overlap.
  → high frequency resolution
- Frequency restriction weakens influence of high partials
  - necessary because some high partials have a big deviation to the next equally-tempered tone
  - Restrict analysis to fundamental frequency range: ≈ 80 – 1300 Hz.



Partial	1	2	3	4	5	6	7	8	9	10	11	12
f_{harm}	65.4	130.8	196.2	261.6	327.0	392.4	457.8	523.3	588.7	654.1	719.5	784.8
$f_{temp.}$	65.4	130.8	196.0	261.6	329.6	392.0	466.2	523.3	587.3	659.3	740.0	784.0
Cent	0.0	0.0	1.96	0.0	-13.7	2.0	-31.2	0.0	3.9	-13.7	-48.6	2.0

Sawtooth Results: Number of Considered Peaks

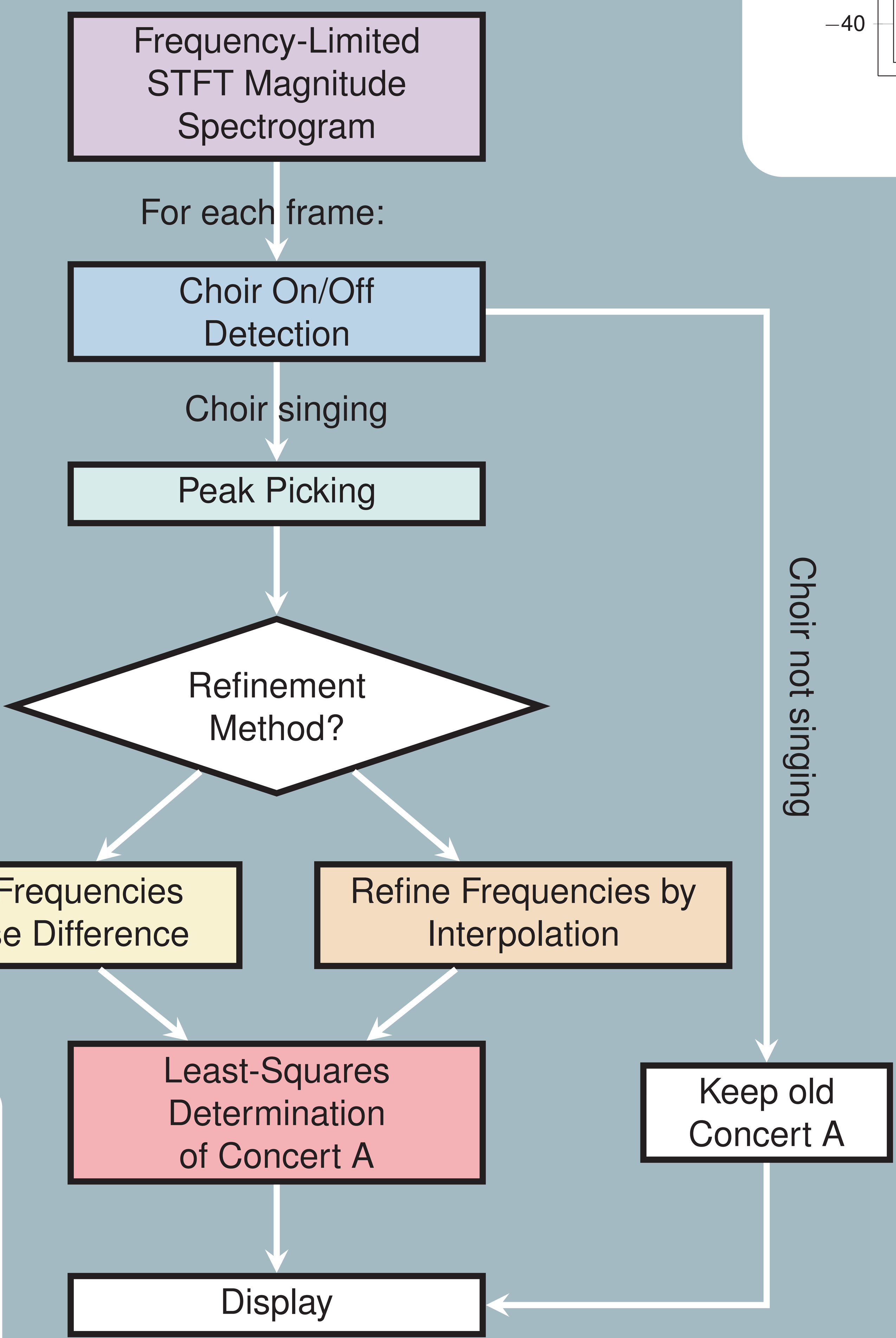


Choir On/Off Detection

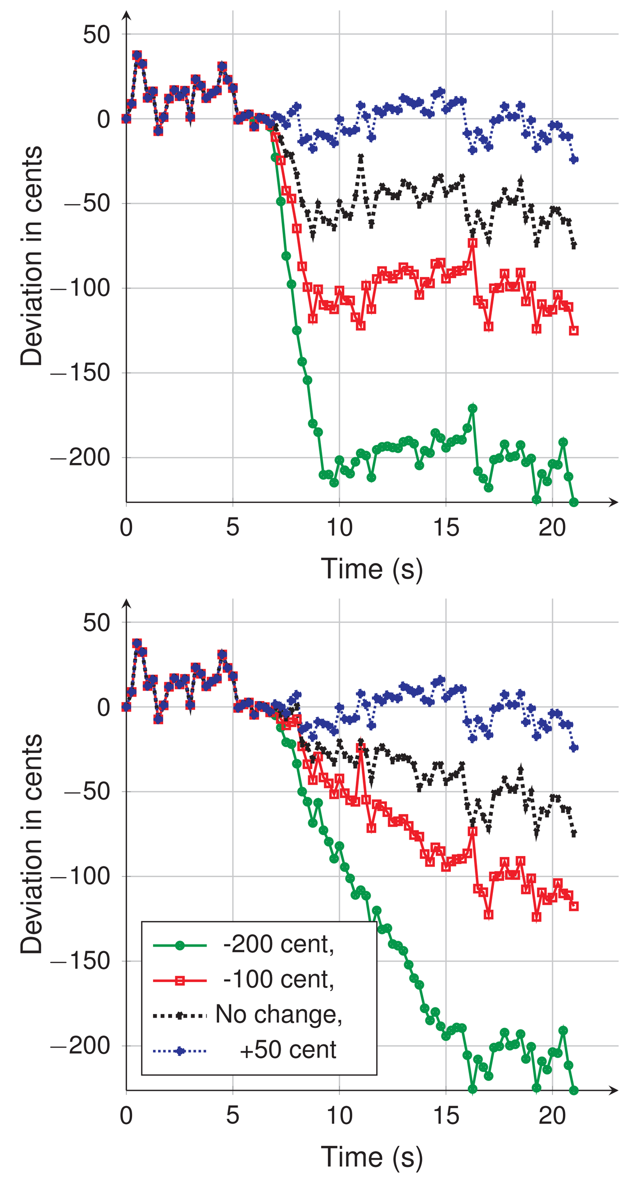
- Prevents noice-induced drifting of the measured tuning.
- Combination of
  - energy measure
  - spectral flatness measure
- Thresholds adjustable (due to different noise conditions)

Peak picking

- Analyzes the spectrogram to find local maxima (peaks)
- max. 6 maxima with a minimum distance between them

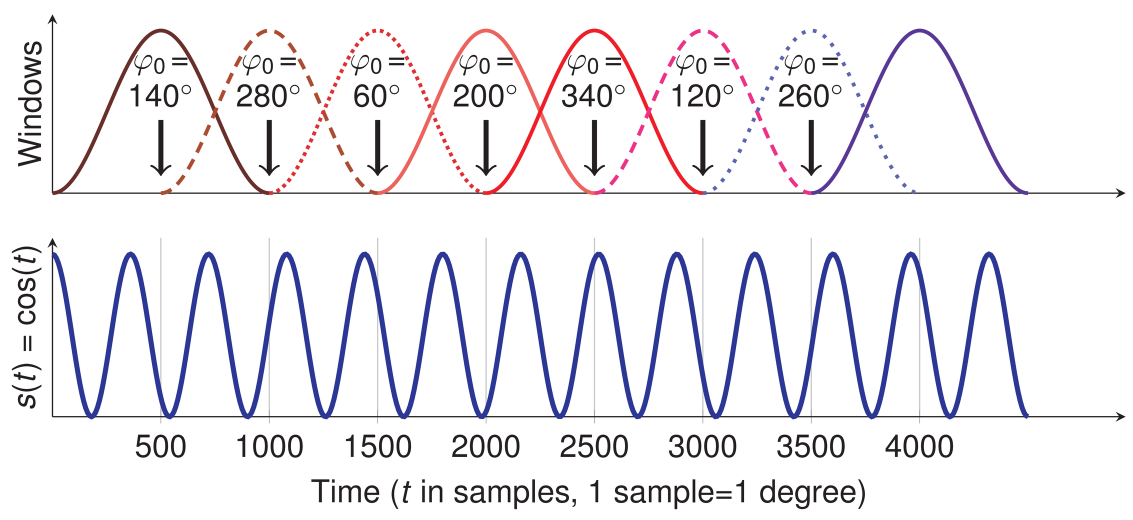


Result: Pitch Shifted Signals



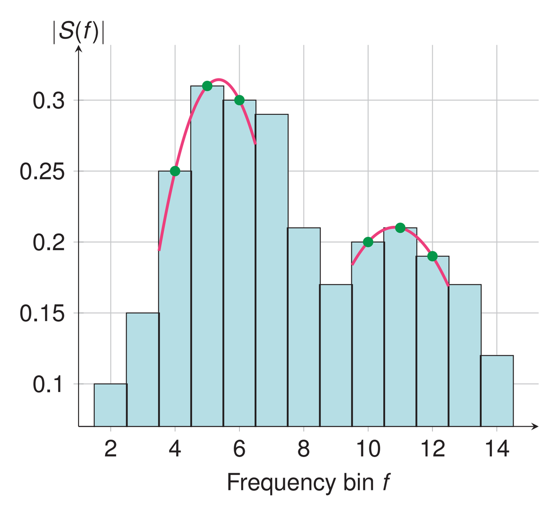
Peak refinement by phase difference

- Measurement of the *instantaneous* frequency of the peak frequency bin.
- Instantaneous frequency can be calculated from the phase difference.
  $\rightarrow f_i(n,k) = f_s \left(\frac{k}{N} + \frac{\mathrm{princarg}(\varphi(n,k) - \varphi(n-1,k))}{2\pi} \right)$



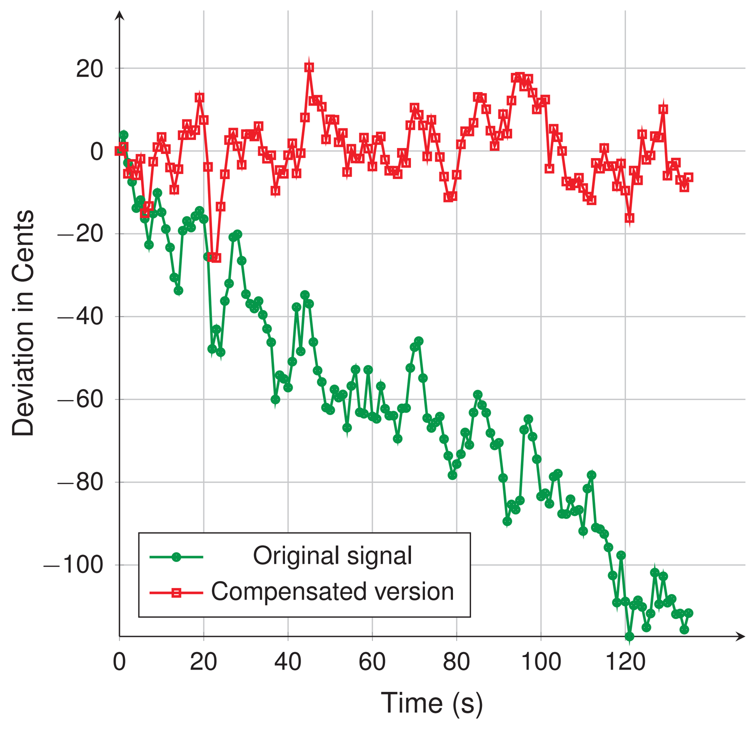
Peak refinement by interpolation

- Paraboloid through peak position and surrounding frequency bins
- Paraboloid peak = refined peak position



Results: Real-World Recording

- Gospel Choir
- Pitch decline: about one semitone
- Original + Compensation with a pitch shifter controlled with our algorithm



Optimal Tuning Pitch

- Least-Squares Determination
- $f_A = \frac{\sum_i f_i \cdot 2^{\frac{s(f_i)}{12}}}{\sum_i 2^{\frac{s(f_i)}{6}}}$
  with f_i = peak frequencies,
  $s(\cdot)$ = corresponding semitone
- $n_{cent} = 1200 \cdot \log_2 \left(\frac{f_A}{f_{ARef}} \right)$