Modeling Meeting Turns

Dan Ellis <dpwe@ee.columbia.edu>
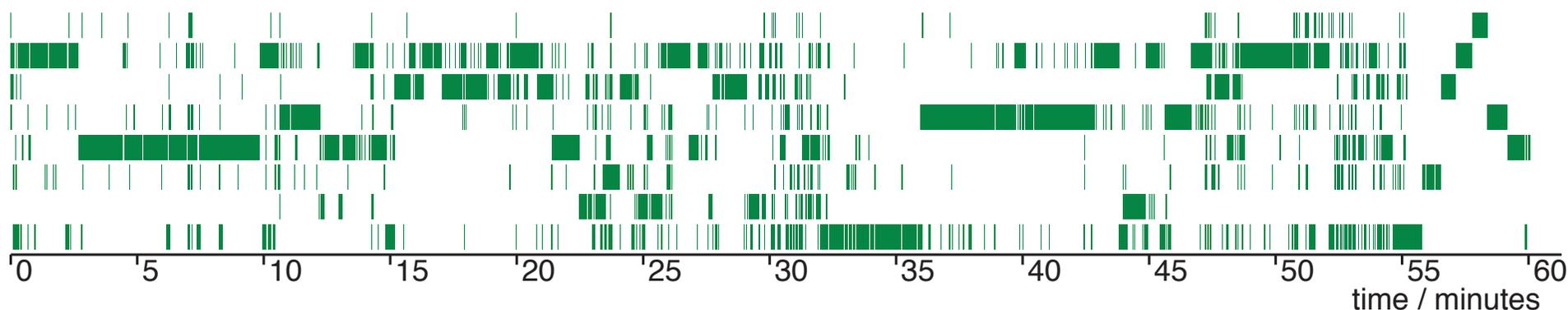LabROSA, Columbia University & ICSI

- Meeting turns visualization
- Turn-pattern segmentation
- 'Talkativity' modeling

Lab
ROSA
Laboratory for the Recognition and
Organization of Speech and Audio

Meeting Turn Visualization

- Speaker turns form patterns on multi-minute timescales:

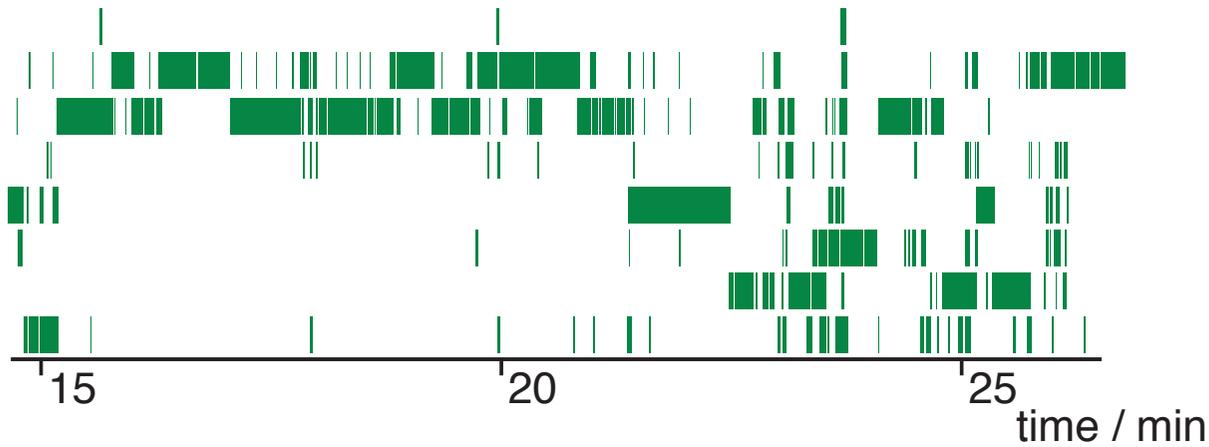**mr04: Hand-marked speaker turns**

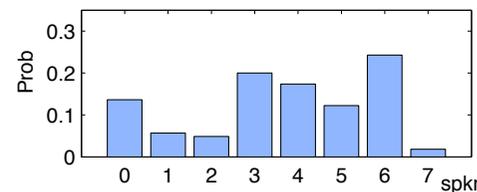- Points of pattern change are 'significant'?
  - topics?
  - modes?

Modeling meeting segments
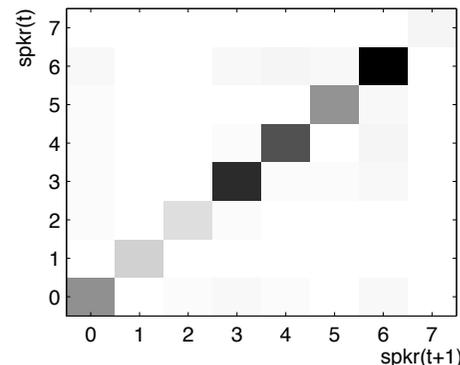
- Model speaker activity patterns like states
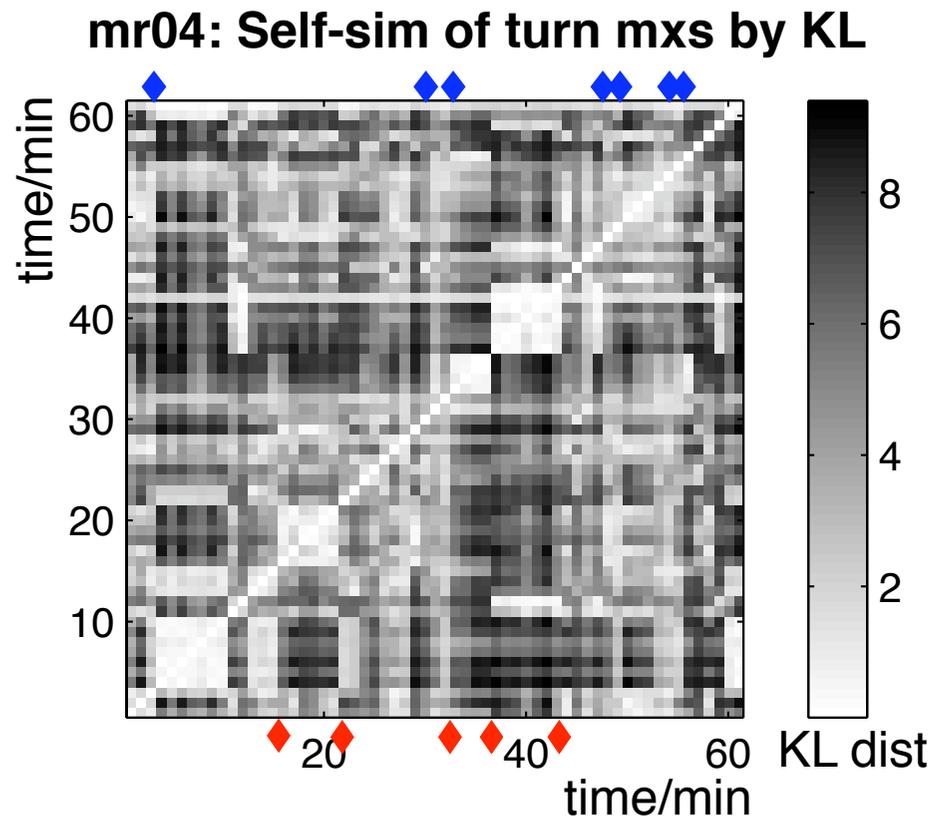

- Prior vector:
  $P(spkr^i)$

- 'Transition' matrix:
  $P(spkr^i_t, spkr^j_{t-1})$

Self-similarity

- Display $Dist(minute_i, minute_j)$
 as KL distance of speaker distributions

mr04: Self-sim of turn mxs by KL

Laboratory for the Recognition and
Organization of Speech and Audio

BIC Segmentation

- ## BIC (Bayesian Information Criterion):
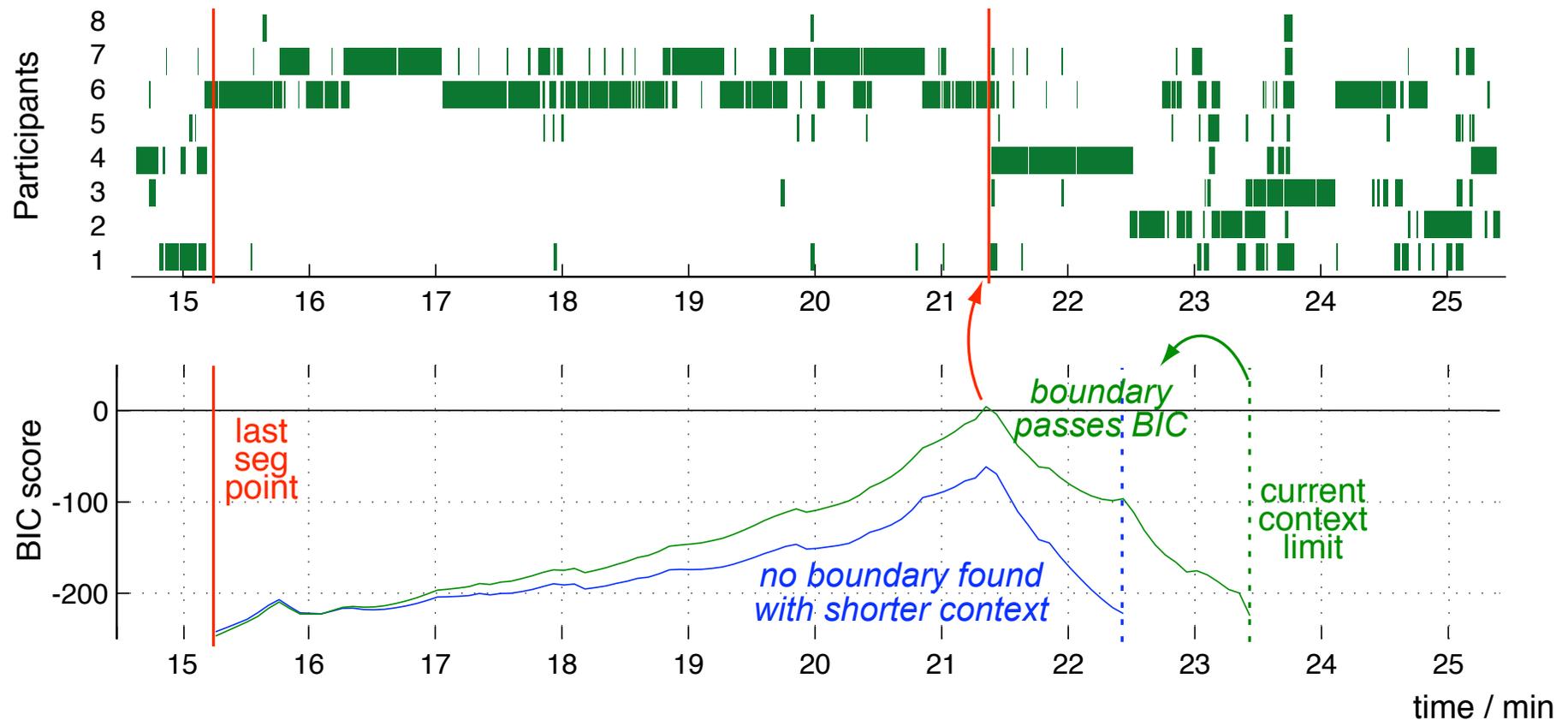 Compare more and less complex models

$$\log \frac{L(X_1; M_1) L(X_2; M_2)}{L(X; M_0)} \gtrless \frac{\lambda}{2} \log(N) \Delta \#(M)$$

- ## For segmentation:
  - Grow context window from current boundary
  - For each window, test every possible segmentation
  - When BIC is positive, mark new segment
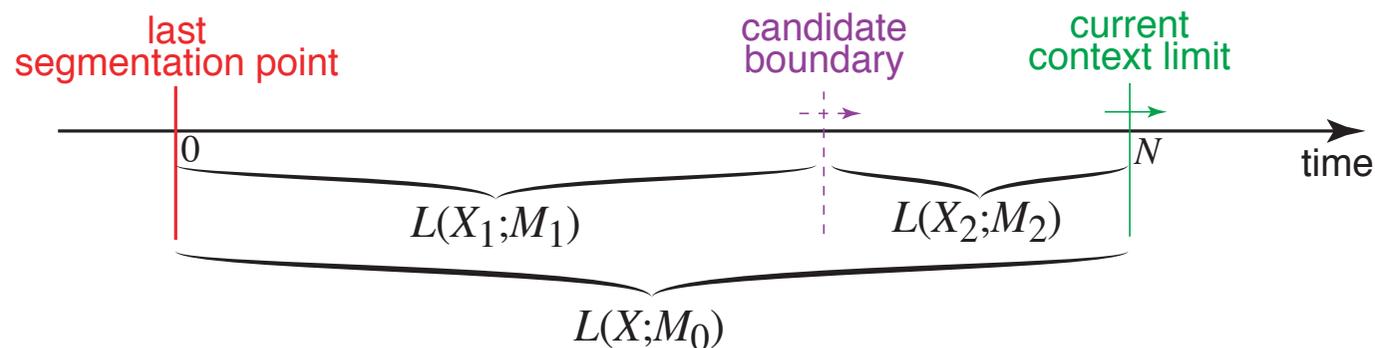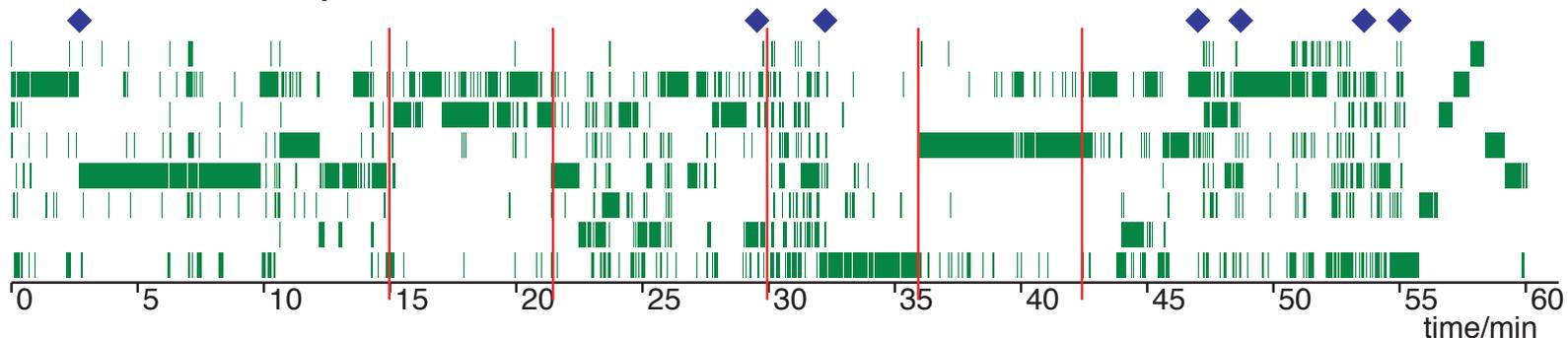
BIC Segmentation

- Example of boundary finding:

BIC Segmentation

- Appears to find shifts in turn patterns:

mr04: Hand-marked speaker turns vs. time + auto/manual boundaries

- Evaluate against topic boundaries
  (6 meetings, 36 boundaries)
  - 15 (42%) agree to within ± 2 minutes
  - 16 'false alarm' insertions

"Talkativity"

- Factors affecting how much one person speaks in a given meeting:
  - relevance/interest of topic to speaker ← *indexable*
  - competition with other speakers ← *confounding*
  - innate tendency to talk - "talkativity" T_s

- Model of expected 'airtime' consumed by each participant s in meeting m:

$$P_{sm} = \frac{T_s}{\sum_{t \in S_m} T_t}$$

  - given $\{T_s\}$, deviations from expected values factor out competition, baseline talkativity

Estimating "Talkativity"
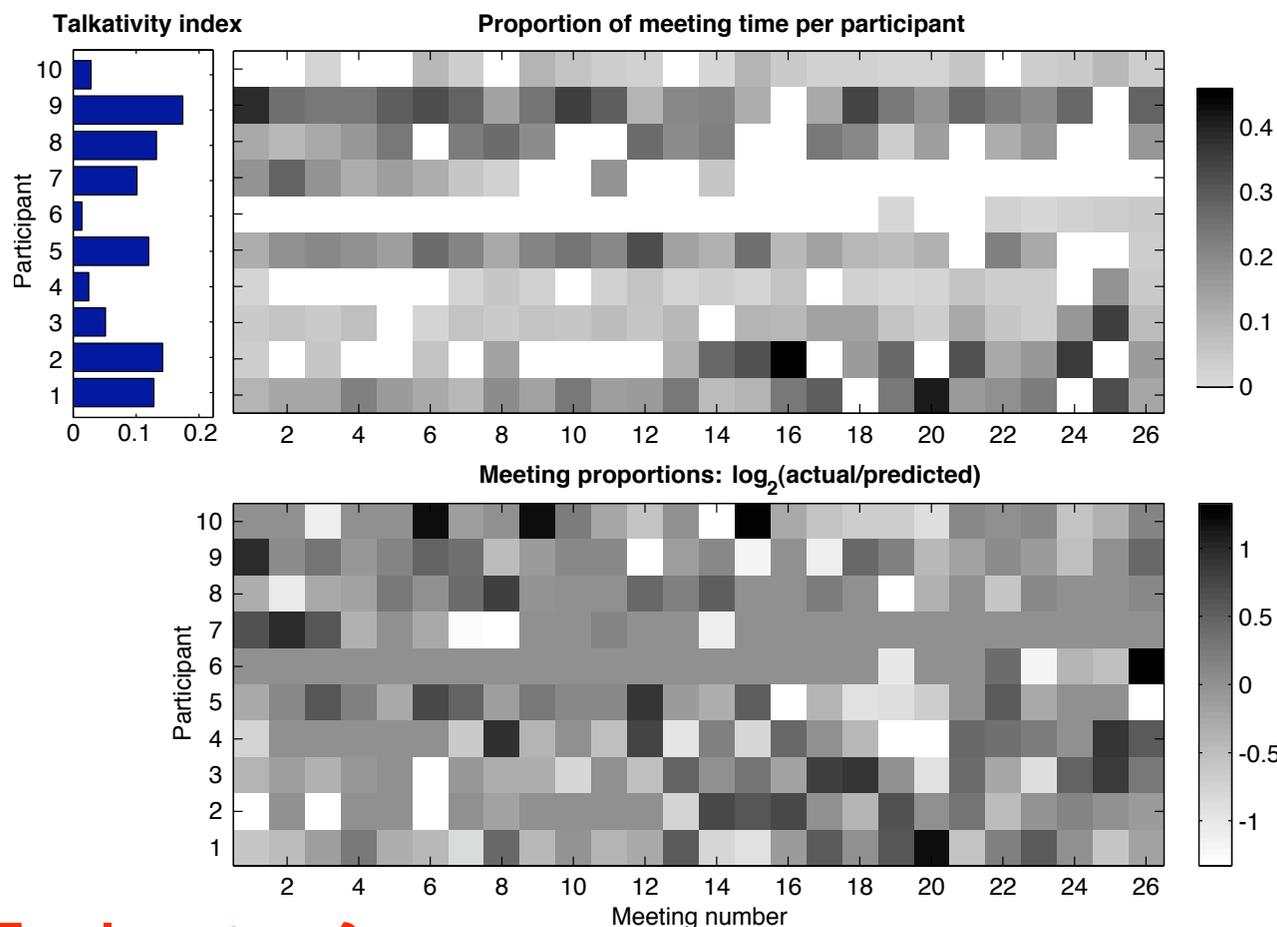
- **Find best-fitting $\{T_s\}$ to fit meeting set**

$$T_s = avg_{m \in M_s} \frac{P_{sm} \sum_{t \in S_m, t \neq s} T_t}{1 - P_{sm}}$$

  - ○ Iteratively recalculate $\{T_s\}$ until (fast) convergence

  - ○ 26 meetings (mr* set), 10 common participants, avg 6.9 participants/meeting

- **Calculate actual:predicted ratios**

Lab
ROSA
Laboratory for the Recognition and
Organization of Speech and Audio

"Talkativity" Results

- Meeting proportions & ratio to prediction

Evaluation?