

Fit:
 obs [14, 6, 216]
 ↑
 features
 OF THE SAME
 WAV FILES
 word

em-step

obs [m]
 xi-sum [states x #frames] ← forward
 A x (α_t (β_t x B[:, t+1])^T) → partial-sum
 xi-sum +=

Backward:
 β(:, -1) = ones...
 β_t = A · (β(:, t+1) x β_{t+1})
 β_t = $\frac{\beta_t}{Z_{\beta t}}$

A ← expected-A ← xi-sum
 expected prior → y₀
 y_t ←
 α_t x β_t
 M, cov

forward:
 A [states x states]
 α₀ = B(:, 0) · prior
 α_t = B(:, t) x (A^T · α_{t-1})
 Z_t = 1 / (1 + α_t)
 log-likelihood += log(Z_t)

state-likelihood:
 B [state x #frames] = f @ 1 5 3
 M[:, s], cov[:, s]
 obs[:, f] @ 1 2 3

α [states x frames]
 log-likelihood - 7290