Parameter Estimation of Hidden Markov Models

Segmentation K-means algorithm

Emilie Nielsen, Noah Koesgaard, Ludvig Marcussen & Michelle Rix

UNIVERSITY OF COPENHAGEN

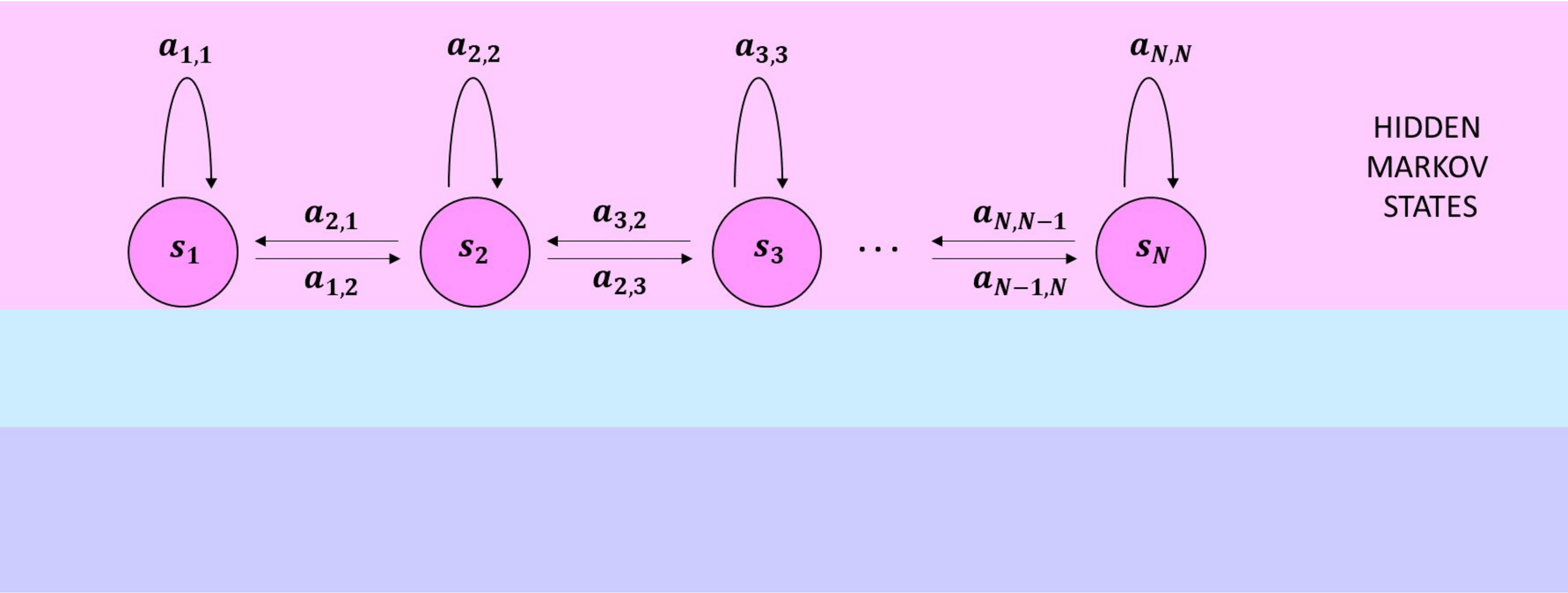
The Article

The Segmental *K*-Means Algorithm for Estimating Parameters of Hidden Markov Models

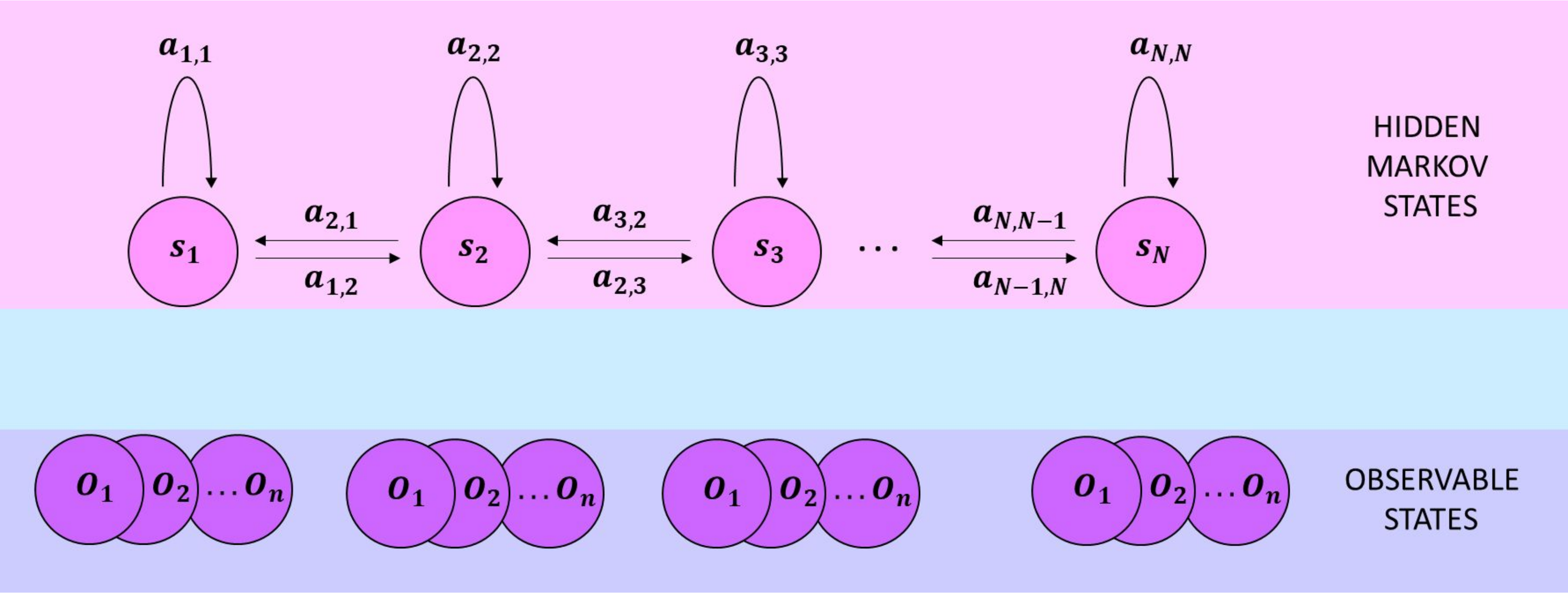
BIING-HWANG JUANG AND L. R. RABINER

Abstract—Statistical analysis techniques using hidden Markov models have found widespread use in many problem areas. This correspondence discusses and documents a parameter estimation algorithm for data sequence modeling involving hidden Markov models. The algorithm which we call the segmental *K*-means method uses the state-optimized joint likelihood for the observation data and the underlying Markovian state sequence as the objective function for estimation. We prove the convergence of the algorithm and compare it with the traditional Baum–Welch reestimation method. We also point out the increased flexibility this algorithm offers in the general speech modeling framework.

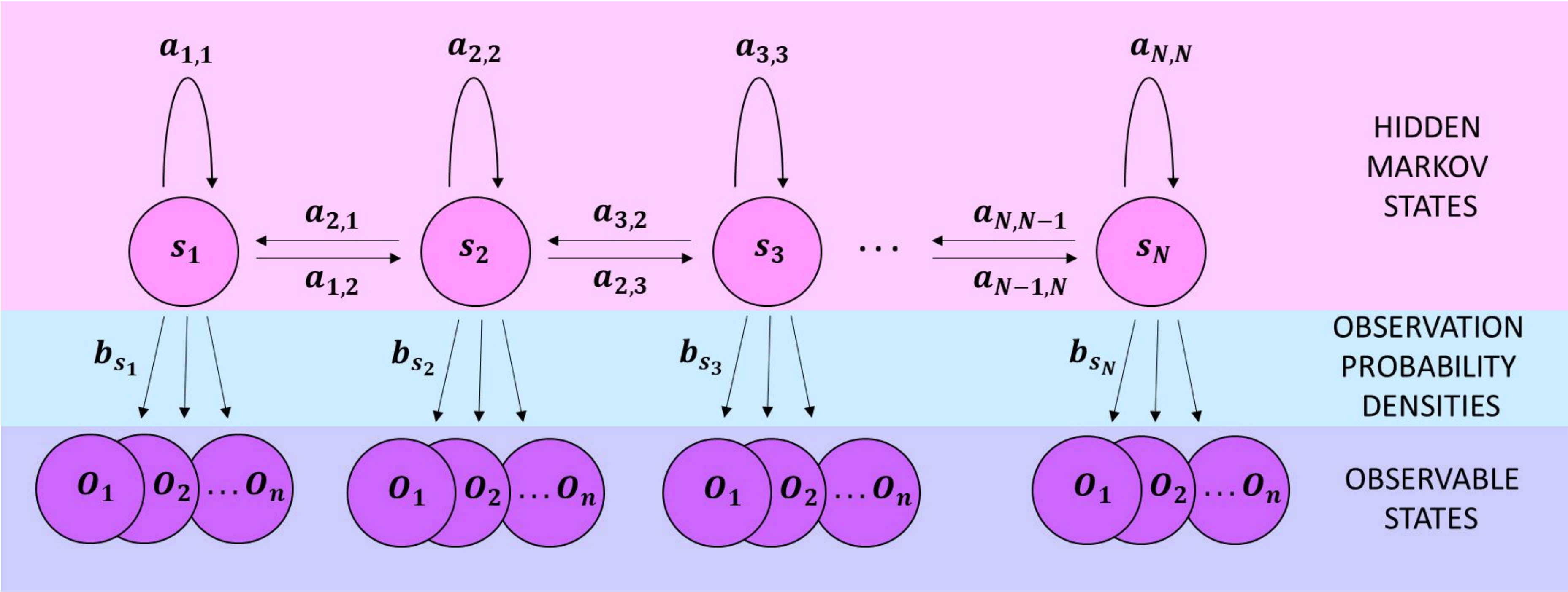
Hidden Markov Model



Hidden Markov Model



Hidden Markov Model



The Problem in a Nutshell

- Structure of the hidden markov model is known.
- How do we estimate the parameters of the model based on the observed sequence?

Estimating Parameters - The Usual Method (Baum-Welch)

$$P(s|A,\pi) = \pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)},s_t)}$$

Estimating Parameters - The Usual Method (Baum-Welch)

$$P(s|A,\pi) = \pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)},s_t)}$$

$$f(x|\lambda) = \sum_{s} \pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)},s_t)} b_{s_t}(x_t)$$

Estimating Parameters - The Usual Method (Baum-Welch)

$$P(s|A,\pi) = \pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)},s_t)}$$

Requires significant computation

$$f(x|\lambda) = \sum_{s} \pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)},s_t)} b_{s_t}(x_t)$$

Estimating Parameters - The Usual Method (Baum-Welch)

$$P(s|A,\pi) = \pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)},s_t)}$$

Requires significant computation

$$f(x|\lambda) = \sum_{s} \pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)},s_t)} b_{s_t}(x_t)$$

May encounter numerical difficulties

Estimating Parameters - “New” Method

$$f(x|\lambda) = \sum_{s} \pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)}, s_t)} b_{s_t}(x_t)$$

$$\max[f(x, s|\lambda)] = \max[\pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)}, s_t)} b_{s_t}(x_t)]$$

Estimating Parameters - “New” Method

$$f(x|\lambda) = \sum_{s} \pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)}, s_t)} b_{s_t}(x_t)$$

Significantly fewer computations

$$\max[f(x, s|\lambda)] = \max[\pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)}, s_t)} b_{s_t}(x_t)]$$

Estimating Parameters - “New” Method

$$f(x|\lambda) = \sum_{s} \pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)}, s_t)} b_{s_t}(x_t)$$

Significantly fewer computations

$$\max[f(x, s|\lambda)] = \max[\pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)}, s_t)} b_{s_t}(x_t)]$$

Encounters problems with similar likelihoods

Estimating Parameters - “New” Method

Algorithm in three steps:

1. Initialization of parameters in λ

Estimating Parameters - “New” Method

Algorithm in three steps:

1. Initialization of parameters in λ

2. Segmentation step:

$$\max[f(x, s|\lambda)] = \max[\pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)}, s_t)} b_{s_t}(x_t)]$$

Estimating Parameters - “New” Method

Algorithm in three steps:

1. Initialization of parameters in λ

2. Segmentation step:

$$\max[f(x, s|\lambda)] = \max[\pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)}, s_t)} b_{s_t}(x_t)]$$

3. Optimization step:

$$\bar{\lambda} = \arg\max_{\lambda}\{\max_{s}[f(x, s|\lambda)]\}$$

Estimating Parameters - “New” Method

Algorithm in three steps:

1. Initialization of parameters in λ

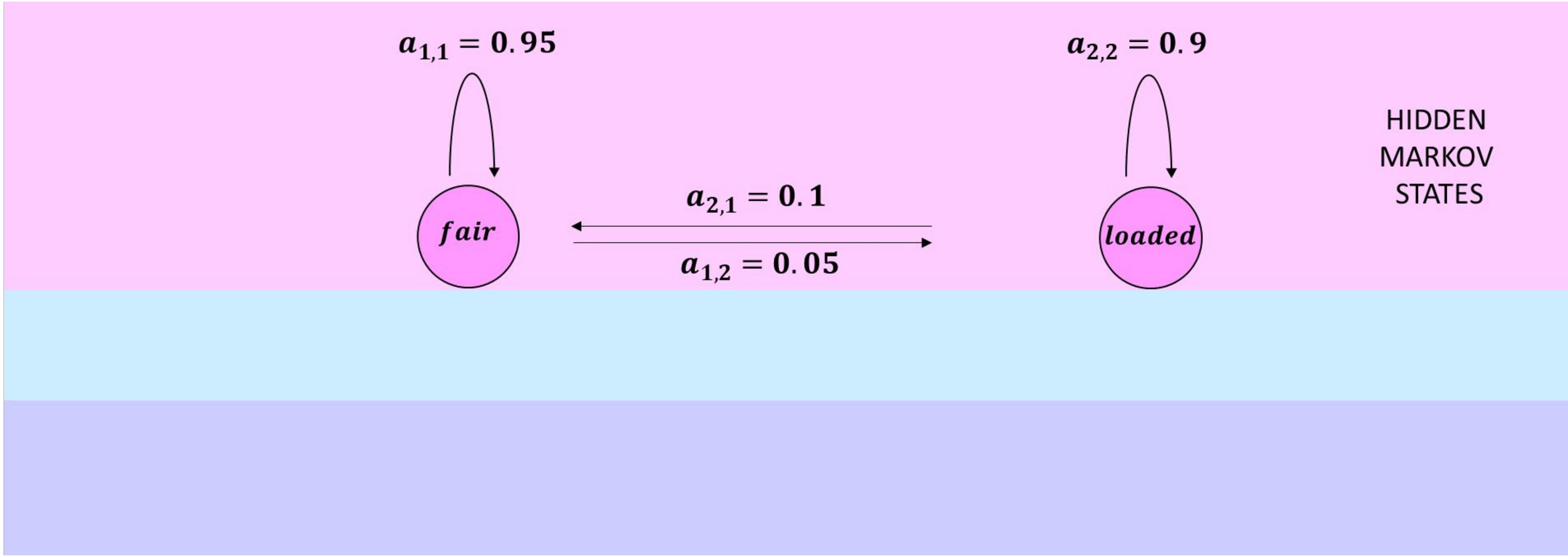
2. Segmentation step:

$$\max[f(x, s|\lambda)] = \max[\pi_{s_0} \prod_{t=1}^{T} a_{(s_{(t-1)}, s_t)} b_{s_t}(x_t)]$$

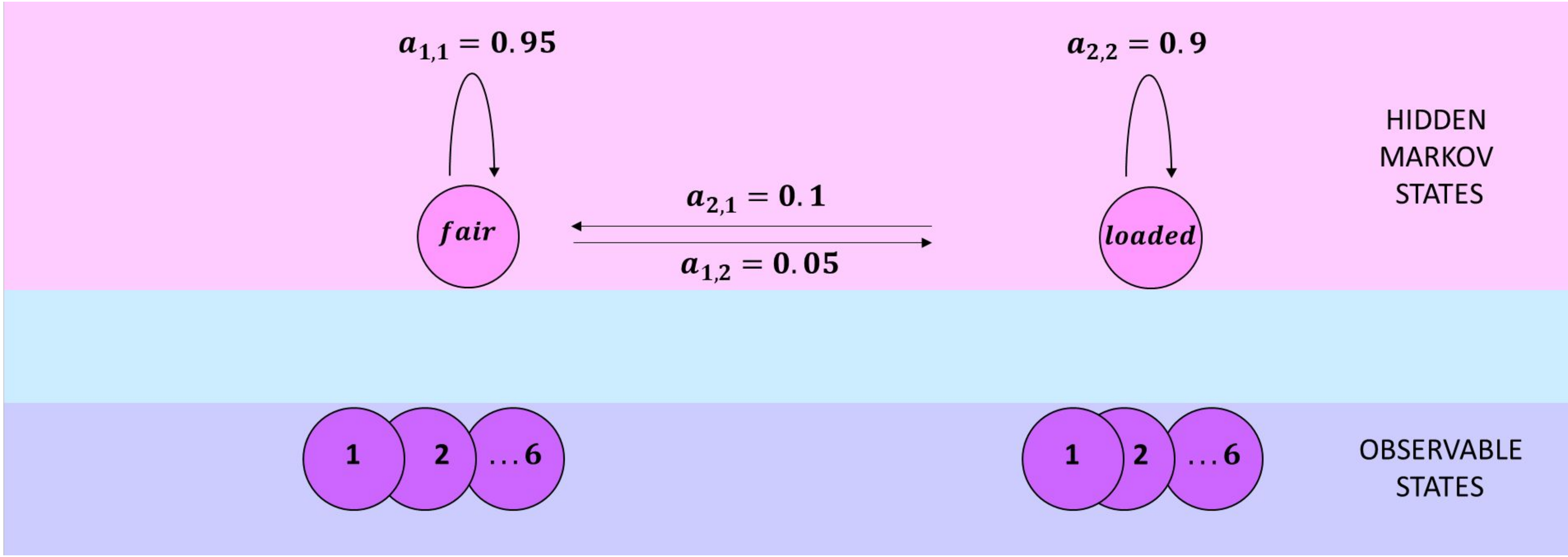
3. Optimization step:

$$\bar{\lambda} = \arg\max_{\lambda}\{\max_{s}[f(x, s|\lambda)]\}$$

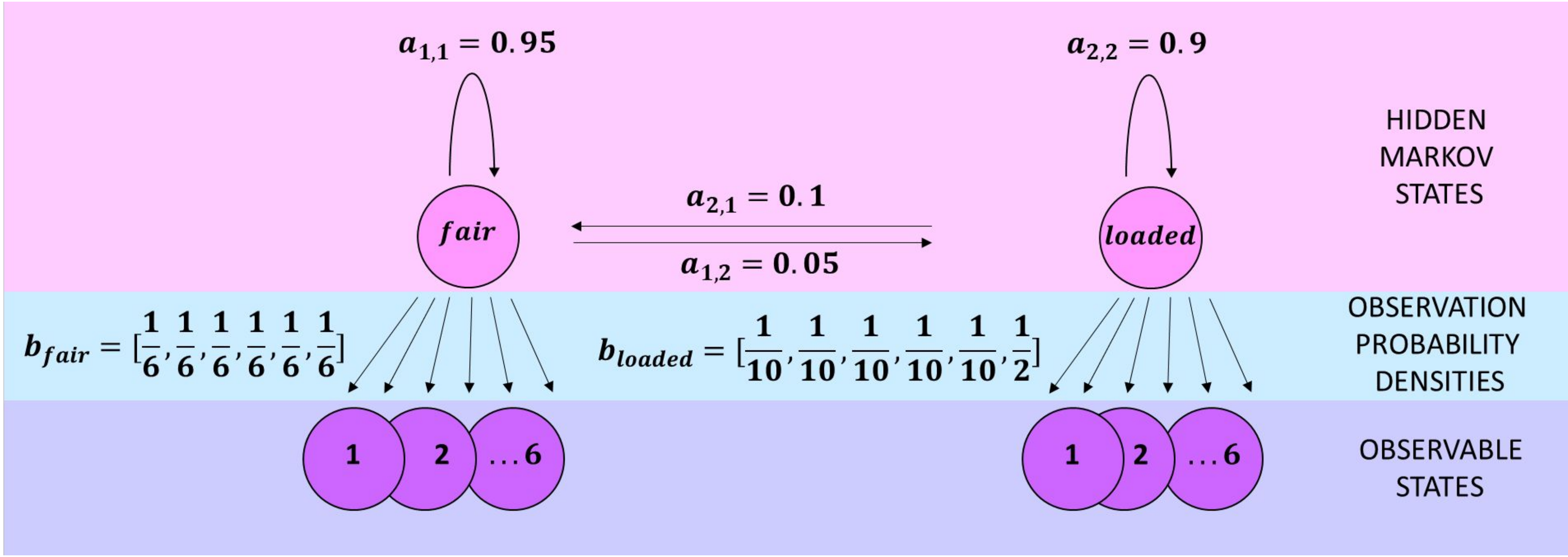
Example calculation: The Occasionally Dishonest Casino



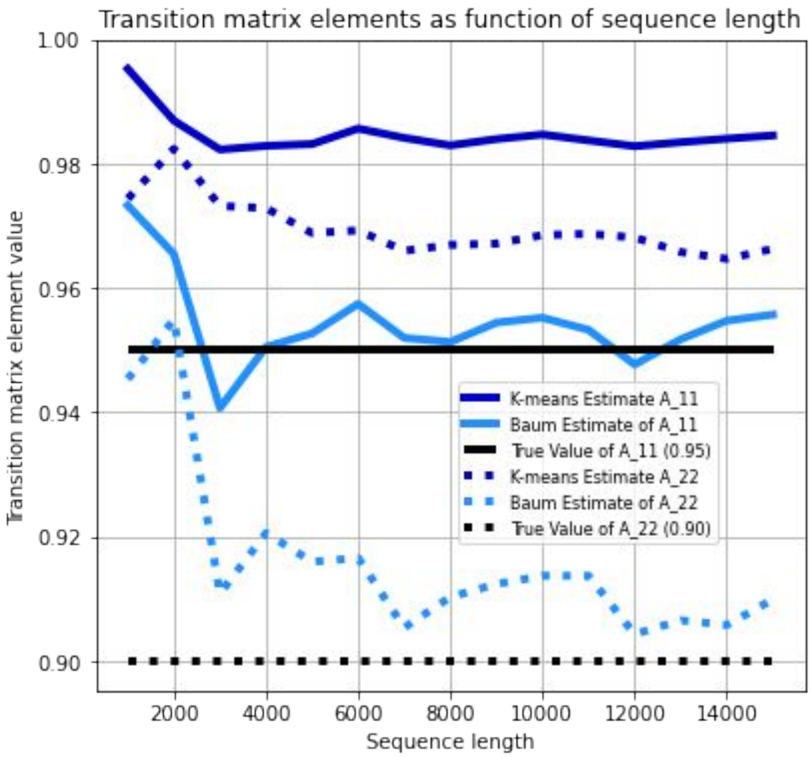
Example calculation: The Occasionally Dishonest Casino



Example calculation: The Occasionally Dishonest Casino



Example calculation: The Occasionally Dishonest Casino



Example calculation: The Occasionally Dishonest Casino