

Assignment 1

Thomai Stathopoulou (*thomai@kth.se*)

Dario Vidas (*vidas@kth.se*)

September 26, 2015

HMM $\lambda = q, A, B$:

$$q = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix} \quad A = \begin{bmatrix} 0.99 & 0.01 \\ 0.03 & 0.97 \end{bmatrix} \quad B = \begin{bmatrix} b_1(x) \\ b_2(x) \end{bmatrix} \quad \begin{matrix} b_1(x) = \mathcal{N}_1(0, 1) \\ b_2(x) = \mathcal{N}_2(3, 2) \end{matrix}$$

Calculation of $P(S_t = j)$

Theoretical calculation of $P(S_t = j)$:

$$\begin{aligned} P(S_1 = 1) &= q_1 = 0.75 \\ P(S_1 = 2) &= q_2 = 0.25 \\ \\ P(S_2 = 1) &= a_{11}P(S_1 = 1) + a_{21}P(S_1 = 2) = 0.75 \\ P(S_2 = 2) &= a_{12}P(S_1 = 1) + a_{22}P(S_1 = 2) = 0.25 \\ \\ P(S_3 = 1) &= a_{11}P(S_2 = 1) + a_{21}P(S_2 = 2) = 0.75 \\ P(S_3 = 2) &= a_{12}P(S_2 = 1) + a_{22}P(S_2 = 2) = 0.25 \\ &\vdots \end{aligned}$$

Corresponding measured relative frequencies for 10,000 samples:

$$\begin{aligned} S_t = 1 & \quad f = 0.7661 \\ S_t = 2 & \quad f = 0.2339 \end{aligned}$$

Expected Value and Variance

Theoretical calculation of expected value and variance:

$$\begin{aligned} E[X_t] &= E_{S_t}[E_{X_t}[X_t|S_t]] \\ &= \sum_{S_t} [P[S_t]E_{X_t}[X_t|S_t]] \\ &= \sum_{S_t} \sum_{X_t} [P[S_t]P[X_t|S_t]] \\ &= b_1(x)P[S_t = 1] + b_1(x)P[S_t = 2] + b_2(x)P[S_t = 1] + b_2(x)P[S_t = 2] \\ &= 0.75\mu_1 + 0.25\mu_1 + 0.75\mu_2 + 0.25\mu_2 \\ &= 0.75 \end{aligned}$$

$$\begin{aligned}
\text{var}[X_t] &= E_{S_t}[\text{var}_{X_t}[X_t|S_t]]\text{var}_{S_t}[E_{X_t}[X_t|S_t]] \\
&= \sum_{S_t} [\text{var}_{X_t}[X_t|S_t]P[S_t]] + E_{S_t}[E_{X_t}^2[X_t|S_t]] - (E_{S_t}[E_{X_t}[X_t|S_t]])^2 \\
&= \sigma_1^2 P[S_t = 1] + \sigma_2^2 P[S_t = 2] + \sum_{S_t} [P[S_t]E_{X_t}^2[X_t|S_t]] - \left(\sum_{S_t} [P[S_t]E_{X_t}[X_t|S_t]]\right)^2 \\
&= 0.75 + 4 \cdot 0.25 + P[S_t = 1]\mu_1^2 + P[S_t = 2]\mu_2^2 - (P[S_t = 1]\mu_1 + P[S_t = 2]\mu_2)^2 \\
&= 1.75 + 0.25 \cdot 3^2 - (0.25 \cdot 3)^2 \\
&= 3.4375
\end{aligned}$$

After producing a sequence of 100,000 observations, the expected value and variation are calculated: 0.7499 and 3.4465 respectively.

HMM Output Sequence

Figure 1 shows the output produced from a sequence of states produced by our HMM λ . It is obvious in the figure that the output values are scattered around 0 and 3, which are the two means of the output distributions of states 1 and 2 respectively. Additionally, the spread around value 3, is larger than the spread around value 0, which is in accordance two the variances of the two distributions ($\sigma_1 = 1, \sigma_2 = 2 = 2\sigma_1$). Finally, since at any t there is a 75% chance of being at state 1 and a 25% chance of being at state 2, we can see that the output follows the same behavior (75% of the output follows $\mathcal{N}_1(\mu_1, \sigma_1)$ and the rest 25% of the output follows $\mathcal{N}_2(\mu_2, \sigma_2)$).

HMM $\lambda = q, A, B$:

$$q = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix} \quad A = \begin{bmatrix} 0.99 & 0.01 \\ 0.03 & 0.97 \end{bmatrix} \quad B = \begin{bmatrix} b_1(x) \\ b_2(x) \end{bmatrix} \quad \begin{array}{l} b_1(x) = \mathcal{N}_1(0, 1) \\ b_2(x) = \mathcal{N}_2(0, 2) \end{array}$$

HMM Output Sequence

In Figure 2 we can see that, now the outputs are scattered around the common mean and the only thing that varies is the spread around the mean. This time the state producing the output is not so obvious, since the two distributions have many common values and the only case when the state is obvious is when there is a large spread caused by the second Gaussian.

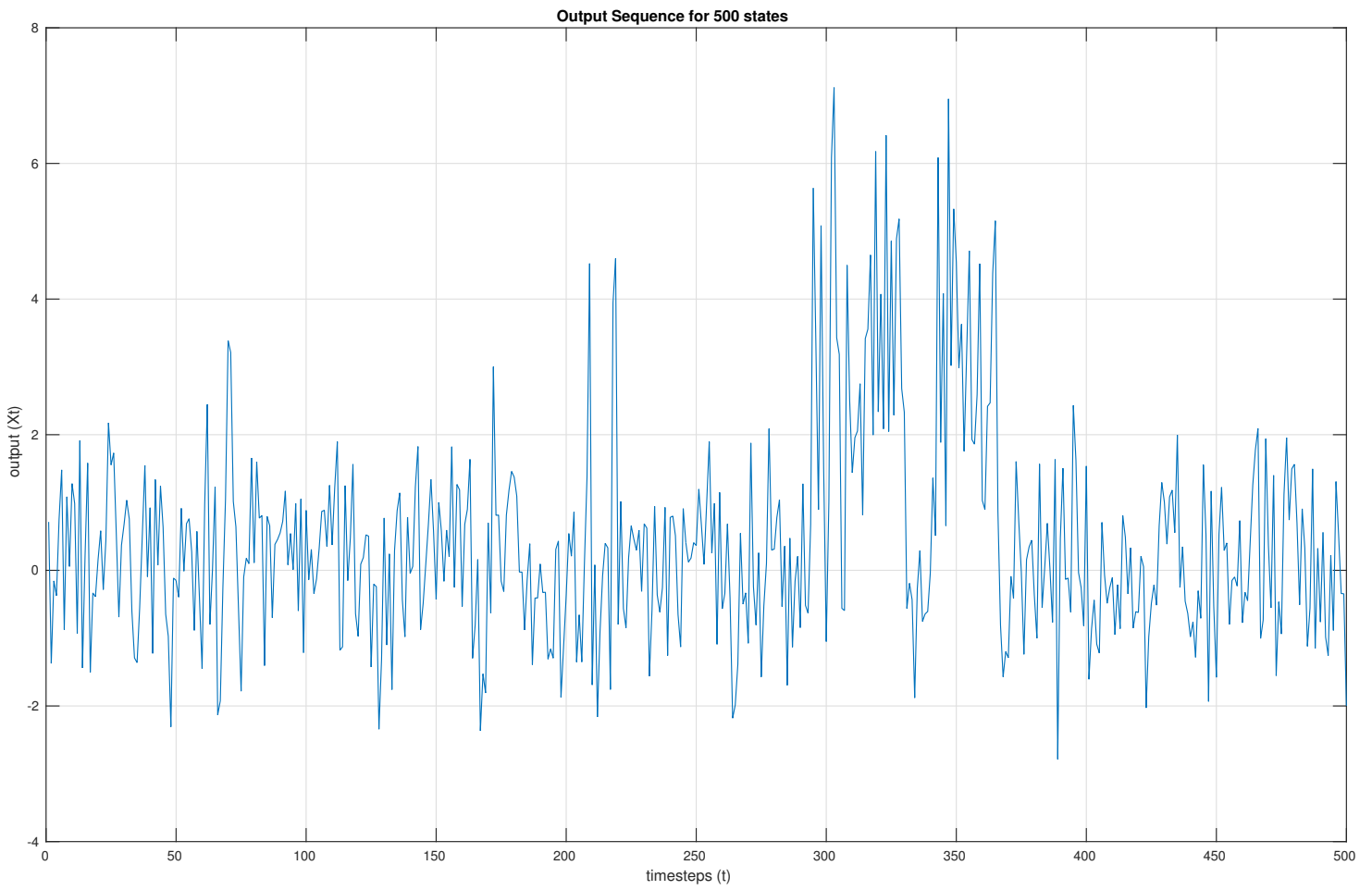


Figure 1: Output Sequence of 500 States

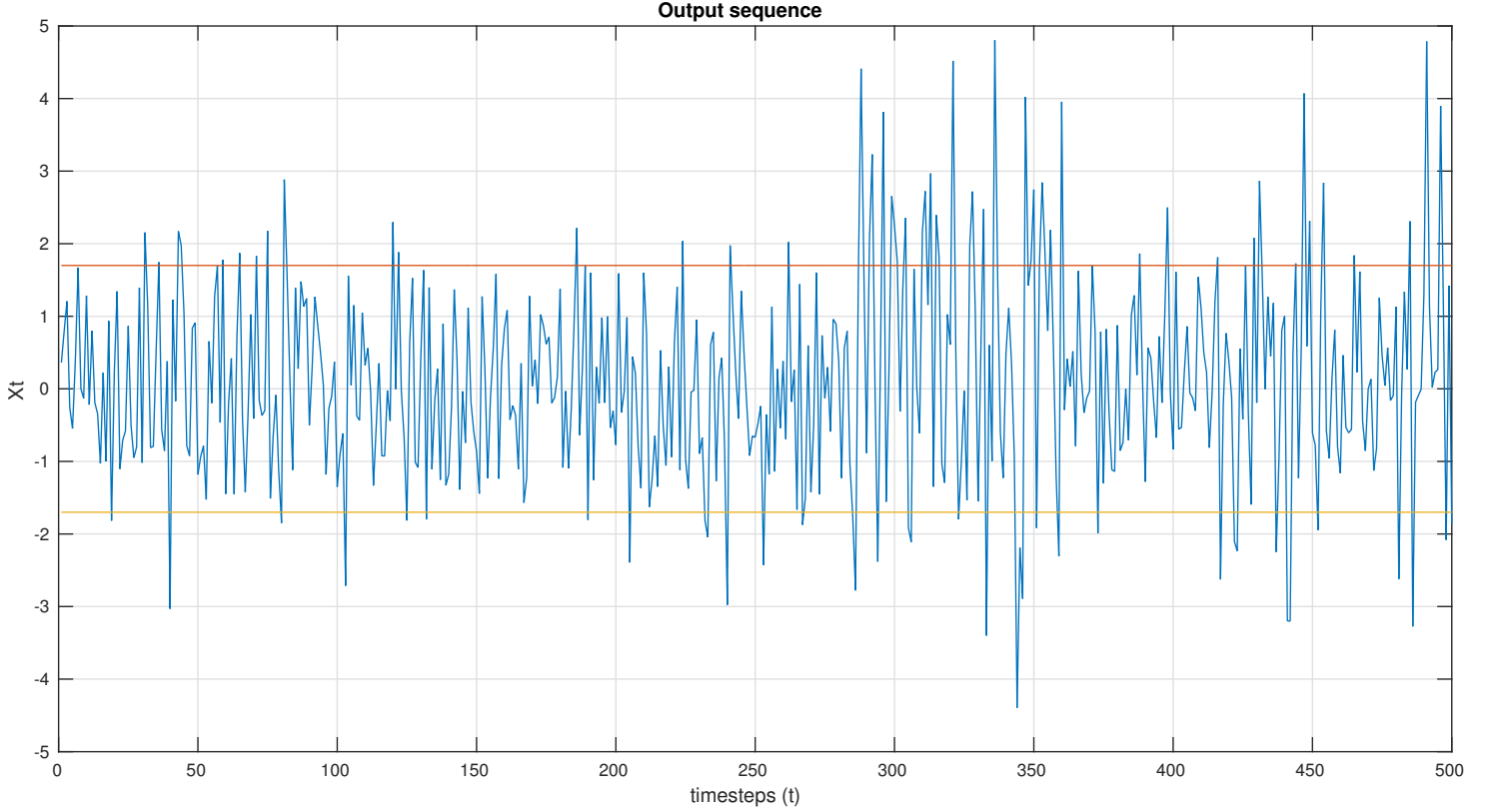


Figure 2: Output Sequence of 500 States

In Figure 3 we see the output distributions based on the states. It is obvious that a threshold can be chosen, so that we can decide the originating state of the output at any step of the output sequence. However, it is obvious that such a classification will have a very large error.

Finite Duration HMM

The finite duration HMM differs from what we've been using so far, in that it contains an extra state, which signifies the end of the sequence. Hence the rand function should check at each timestep, if this state has been reached and if so, stop producing states.

In order to verify that our function works properly for a finite HMM, we need only define a finite HMM, produce a number of state sequences and check if at least one of them has length shorter than what asked by the

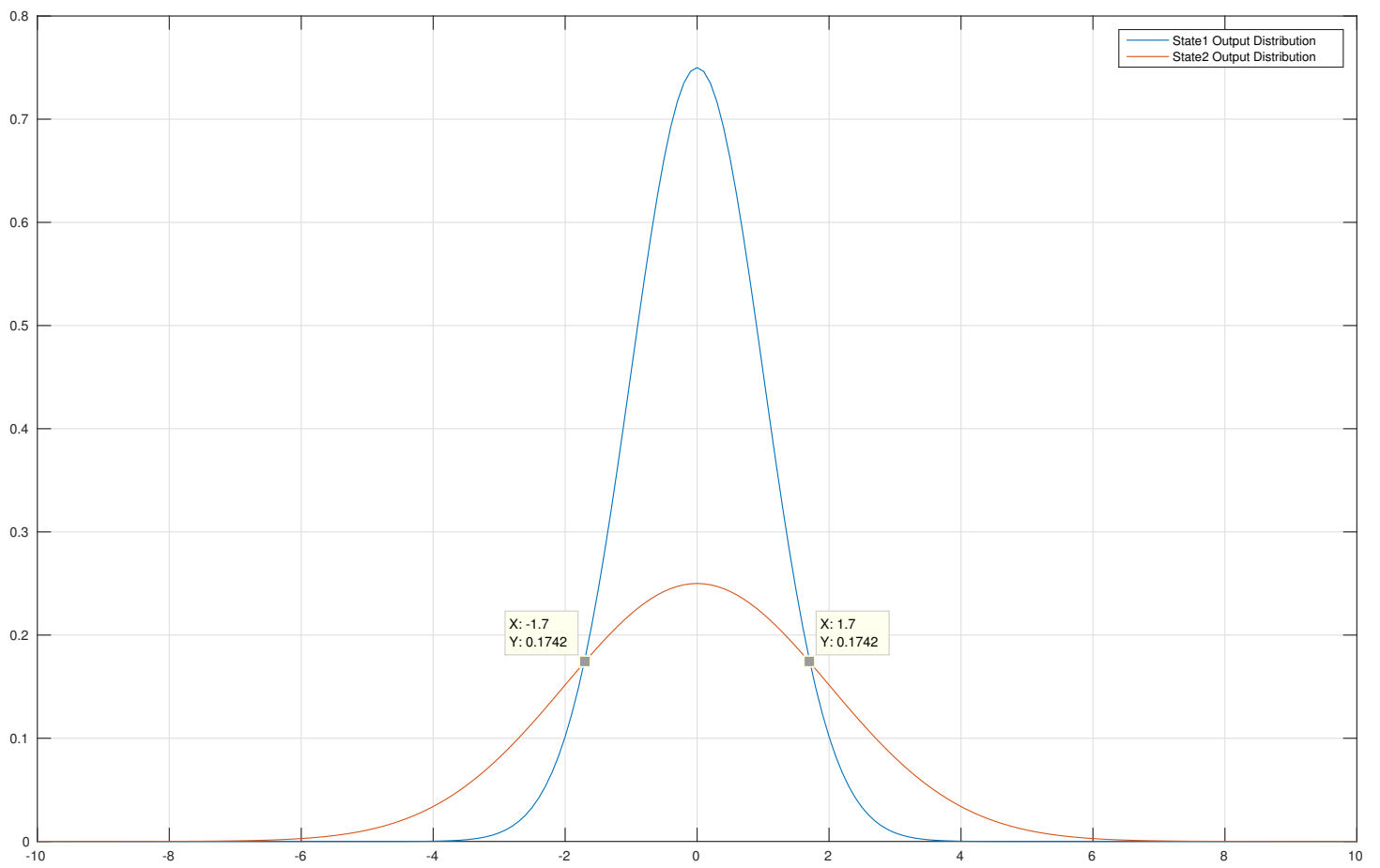


Figure 3: Output Distributions

parameters.

HMM $\lambda = q, A, B$:

$$q = \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} \quad A = \begin{bmatrix} 0.9 & 0.05 & 0.05 \\ 0.05 & 0.9 & 0.05 \end{bmatrix} \quad B = \begin{bmatrix} b_1(x) \\ b_2(x) \end{bmatrix} \quad \begin{matrix} b_1(x) = \mathcal{N}_1(0, 1) \\ b_2(x) = \mathcal{N}_2(3, 2) \end{matrix}$$

For this HMM, after producing 1000 sequences of (required) length 100, we see that, 99.7% of the sequences stopped before reaching the 100th state. This indicates that the finite functionality of rand works properly.

It has been noticed that, the mean length of the 1000 sequences is $19.848 \sim 20$. Examining the transition matrix of our HMM, we can see that at any t, we have a 5% (or 1/20) chance of reaching the end state. That means that, statistically, 1 out of twenty states will end the sequence.

Multidimensional Observations

Defining HMM with multidimensional outputs:

HMM $\lambda = q, A, B$:

$$q = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix} \quad A = \begin{bmatrix} 0.99 & 0.01 \\ 0.03 & 0.97 \end{bmatrix} \quad B = \begin{bmatrix} b_1(x) \\ b_2(x) \end{bmatrix} \quad \begin{matrix} b_1(x) = \mathcal{N}_1\left(\begin{bmatrix} 5 \\ 20 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\right) \\ b_2(x) = \mathcal{N}_2\left(\begin{bmatrix} 0 \\ 10 \end{bmatrix}, \begin{bmatrix} 1 & 6 \\ 6 & 3 \end{bmatrix}\right) \end{matrix}$$

Similar to task 4, we produce an output sequence of length 500 and plot the output values. Each dimension of the vector is a different line. In Figure 4 we can see how the outputs are changing while the states change and while following the Gaussian distribution for each state and each dimension.

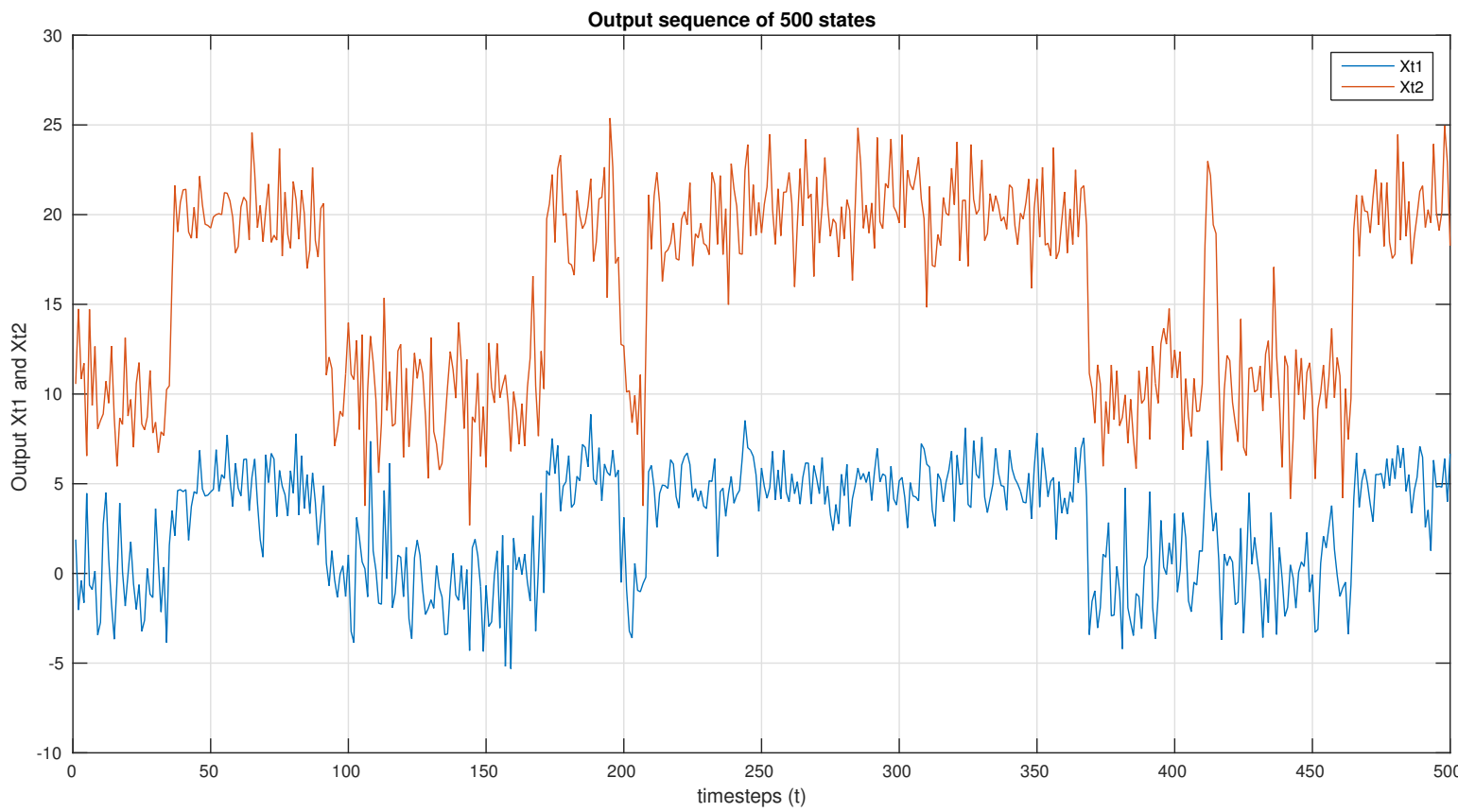


Figure 4: Both dimensions of the Output Sequence of 500 States